آمار و احتمالات مهندسی

دکتر نادر نعمت‌الهی
دانشگاه علامه طباطبایی

فهرست مطالب

فصل اول

آمار توصیفی

۱.۱ آمار چیست

در اصطلاح عامیانه آمار به معنای ثبت و نمایش اطلاعات عددی در مورد یک موضوع، مثلاً ثبت و نمایش تعداد بیکاران، تعداد تصادفات رانندگی، میزان محصولات کشاورزی، میزان صدور نفت، جمعیت شهر تهران و غیره می‌باشد. ولی علم آمار امروزه دارای مفهومی بسیار وسیعتر از این کاربرد عامیانه است. مفاهیم عامیانهٔ آمار زیر مجموعه‌ای از آمار مصطلح بین آماردانان است. از نقطه نظر علمی آمار به مجموعهٔ روشهایی برای جمع آوری، تنظیم و خلاصه کردن اطلاعات عددی و غیر عددی و انجام استنباط و نتیجه‌گیری بوسیلهٔ تجزیه و تحلیل آنها، اطلاق می‌شود. در این اثر با چنین مفاهیمی مواجه هستیم. بحث خود را با آمار توصیفی آغاز می‌کنیم و با ارائه مطالبی از احتمالات به مبحث استنباط آماری می‌پردازیم.

در بررسی صفاتی از یک جمعیت به علت بزرگ بودن جمعیت، پر هزینه بودن بررسی و سایر معضلات، اغلب اوقات در بررسی تمامی اعضای جمعیت با مشکل مواجه می‌شویم. برای رفع این مشکل از نمونه استفاده می‌کنیم. یعنی از جمعیت اعضای مناسبی را انتخاب می‌کنیم و به مطالعهٔ صفات مورد نظر در این اعضا می‌پردازیم. اگر نمونه‌گیری با دقت آماری کافی صورت پذیرد آنگاه نتایج حاصل برای نمونه به سادگی برای جمعیت قابل تعمیم است.

مثال ۱.۱.۱ در بررسی آماری کیفیت کالاهای تولیدی یک کارخانه با مشکل پر هزینه بودن، وقت گیر بودن، از بین رفتن کالای مورد بررسی و غیره مواجه هستیم. از این رو به جای بررسی

تک تک محصولات یک کارخانه فقط تعدادکمی از آنها را به تصادف انتخاب و مورد بررسی قرار می‌دهیم و با تعیین شاخصهایی در این نمونه به بررسی پارامترهای مورد نظر در تولید کل محصولات تولیدی می‌پردازیم. مثلاً اگر از تولیدات لامپ یک کارخانهٔ لامپ سازی در یک روز ۱۰۰ لامپ را به طور تصادفی انتخاب کرده و مشاهده کنیم که ۵ عدد از آنها معیوب هستند، در این صورت ۵٪ لامپهای مشاهده شده معیوب هستند. هم اکنون این سؤال مطرح می‌شود که آیا از روی این مشاهدات می‌توان نتیجه گرفت که ۵٪ از کل لامپهای تولیدی این کارخانه معیوب می‌باشند؟ توسط روشهای آماری می‌توان به این سؤال و کلاً این نوع نتیجه گیریها پاسخ داد.

با توجه به مطالب گفته شده، در یک بررسی آماری با دو بخش آمار توصیفی و آمار استنباطی مواجه هستیم. در آمار توصیفی با تهیه، تنظیم و خلاصه کردن اطلاعات عددی و غیر عددی سروکار داریم و حال آنکه تجزیه و تحلیل و نتیجه گیری آماری به آمار استنباطی برمی‌گردد. برای اینکه بتوان از روی مشاهدات قسمتی از جمعیت یک نتیجه گیری و قضاوت منطقی و علمی در مورد کل جمعیت داشته باشیم از نظریهٔ احتمالات استفاده می‌کنیم. بحث این فصل به آمار توصیفی اختصاص دارد.

۲.۱ آمار توصیفی

روشهایی که بوسیلهٔ آنها می‌توان اطلاعات جمع آوری شده را تنظیم، طبقه‌بندی و خلاصه نمود و آنها را بوسیلهٔ نمودارهایی نمایش داد به آمار توصیفی موسوم است. برای معرفی این روشها نیاز به برخی اصطلاحات داریم که در ذیل به معرفی آنها می‌پردازیم.

جمعیت مجموعهٔ تمام افراد یا اشیایی که مطالعات آماری در مورد یک یا چند صفت آنها در یک مکان و زمان معین انجام می‌گیرد به **جمعیت** موسوم است. هر یک از این افراد یا اشیا را یک عضو جمعیت می‌نامند و تعداد اعضای جمعیت را اندازهٔ جمعیت می‌نامند.

برای مثال اگر بخواهیم معدل دانشجویان یک دانشکده در یک نیمسال را مورد بررسی قرار دهیم آنگاه جمعیت مورد نظر کلیهٔ دانشجویان آن دانشکده می‌باشند و صفت مورد مطالعه معدل نیمسال تحصیلی آنها است. همین طور اگر بخواهیم میزان کالری موجود در غذاهای کنسرو شده در یک کارخانه کنسرو سازی در یک روز معین را مورد بررسی قرار دهیم آنگاه جمعیت مورد نظر

تمامی غذاهای کنسرو شده کارخانه در آن روز و صفت مورد مطالعه میزان کالری موجود در آنها می‌باشد.

چنانچه متذکر شدیم، مثلاً در بررسی میزان کالری موجود در غذاهای کنسرو شده، به علت هزینه زیاد، وقت گیر بودن، نداشتن امکانات کافی یا از بین رفتن برخی از محصولات تولیدی، مطالعه و بررسی کلیه محصولات تولیدی بسیار مشکل و برخی اوقات غیر ممکن است. بنابراین به جای مطالعهٔ کل محصولات تولیدی کارخانه قسمتی از آن را به عنوان نمونه انتخاب می‌کنیم و به مطالعهٔ آن می‌پردازیم.

نمونه زیر مجموعه‌ای از جمعیت که طبق یک قاعده و ضابطهٔ خاصی برای مطالعهٔ صفتی از جمعیت انتخاب می‌شود را یک **نمونه** گویند. تعداد اعضای نمونه به اندازهٔ نمونه موسوم است.

در بررسیهای آماری سعی می‌کنند در انتخاب نمونه دقت کافی انجام گیرد تا با بررسی چنین نمونه مناسبی نتایج حاصله از آن را بتوان با دقت زیاد برای جمعیت تعمیم داد. در هر صورت بایستی نمونه انتخاب شده یک الگوی مناسب از جمعیت باشد. برای مثال اگر بخواهیم در مورد میزان درآمد افراد ساکن در یک شهر مطالعه‌ای را انجام دهیم، بایستی نمونه ما بگونه‌ای انتخاب شود که شامل افراد با درآمد کم، متوسط و زیاد به نسبت موجود در جمعیت باشد.

داده‌ها در یک بررسی آماری، بایستی صفت مورد مطالعه را به صورت اعداد و ارقام نمایش دهیم. اگر صفت مورد مطالعه کمی، مانند وزن، حجم، درجهٔ حرارت و غیره باشد آنگاه این عمل به سادگی با اندازه‌گیری امکان‌پذیر است. اما اگر صفت مورد مطالعه کیفی، مانند گروه خون، شغل، رنگ چشم و غیره باشد آنگاه بایستی با یک قاعده معین این مسائل کیفی را با اعداد و ارقام نشان داد. در هر صورت این اعداد و ارقام را **داده‌ها** گویند که به دو صورت **گسسته** و **پیوسته** می‌باشند. داده‌های گسسته داده‌هایی هستند که بین دو مقدار متصور آنها هیچ عدد دیگری وجود نداشته باشد، مانند تعداد فرزندان یک خانواده که شامل مقادیر ۰، ۱، ۲ و ... است و همچنین صفت شغل افراد که به آن مثلاً اعداد ۱، ۲، ۳ و ... را نسبت می‌دهیم و بین این مقادیر عدد دیگری در رابطه با صفت مورد نظر وجود ندارد. داده‌های پیوسته داده‌هایی هستند که بین هر دو مقدار متصور آنهاهمواره عدد دیگری وجود دارد، مانند وزن افرادکه بین دو نفر با وزنهای نزدیک به هم همواره می‌توان فردی را با وزنی بین وزن دو فرد یاد شده در جمعیت یافت. از جمله داده‌های گسسته می‌توان داده‌های

مربوط به صفات گروه خون، رنگ، نژاد، شغل، تعداد کالاهای تولیدی و غیره را برشمرد و از جمله داده‌های پیوسته می‌توان داده‌های مربوط به صفات وزن، طول قد، فشار گاز، قطر لوله تولیدی یک کارخانه و غیره را برشمرد.

در آمار بعد از جمع آوری داده‌ها به بررسی آماری بر روی آنها می‌پردازیم. در مرحلهٔ نخست با توجه به اهداف بررسی، داده‌ها را تنظیم، طبقه بندی و خلاصه می‌کنیم به طوری که بتوانیم اطلاعات مفیدی برای نیل به اهداف و نتایج مورد نظر به دست آوریم. انجام این کار در سه مرحله به شرح زیر صورت می‌پذیرد:

الف – تنظیم و طبقه‌بندی داده‌ها در یک جدول،

ب – ترسیم نمودارهای گوناگون از روی مقادیر ارائه شده در جدول،

ج – خلاصه کردن داده‌ها به یک یا چند عدد موسوم به شاخص یا آماره.

سه موضوع فوق از موضوعات اساسی بحث آمار توصیفی است که در ذیل به معرفی و بررسی آنها می‌پردازیم.

۳.۱ جدولهای آماری

نخستین گام در خلاصه کردن داده‌ها، طبقه‌بندی و تنظیم آنها در یک جدول موسوم به جدول آماری است. یک جدول آماری بایستی به نحوی تنظیم شود که بتوان از آن به راحتی اطلاعات نهفته در داده‌ها را استخراج کرد. متداولترین جدول آماری **جدول فراوانی** است که در آن داده‌ها، تعداد موجود از هر داده و درصد موجود از هر داده مشخص می‌شود. بنابراین یک جدول فراوانی شامل موارد زیر است

الف – فراوانی و فراوانی نسبی فرض کنید n داده از k نوع (k طبقه) داشته باشیم و تعداد این داده‌ها در این k طبقه به ترتیب f_1، f_2، ... و f_k باشند. به f_1، f_2، ... و f_k فراوانی‌های طبقات و به $r_1=\frac{f_1}{n}$، $r_2=\frac{f_2}{n}$، ... و $r_k=\frac{f_k}{n}$ فراوانی‌های نسبی طبقات گوئیم. واضح است که

$$1\le f_i\le n \quad , \quad 0\le r_i\le 1 \quad , \quad \forall i=1,2,\ldots,k$$

$$f_1+f_2+\ldots+f_k=\sum_{i=1}^{k} f_i=n \quad , \quad r_1+r_2+\ldots+r_k=\sum_{i=1}^{k} r_i=1$$

ب – فراوانی تجمعی و فراوانی تجمعی نسبی اگر در هر طبقه فراوانی آن طبقه و طبقات قبل از آن را با یکدیگر جمع کنیم فراوانی تجمعی آن طبقه به دست می‌آید و اگر در هر طبقه فراوانی نسبی آن طبقه و طبقات قبل از آن را با یکدیگر جمع کنیم فراوانی تجمعی نسبی آن طبقه حاصل می‌شود، یعنی

$$\boxed{\begin{array}{l} \text{فراوانی تجمعی طبقهٔ jام} = g_j=f_1+f_2+\ldots+f_j=\sum_{i=1}^{j} f_i \\ \text{فراوانی تجمعی نسبی طبقه jام} = s_j=r_1+r_2+\ldots+r_j=\sum_{i=1}^{j} r_i \end{array}}$$

واضح است که $g_k=n$ و $s_k=1$.

مثال ۱.۳.۱ اگر از ۱۰۰ کارمند یک اداره، ۵۰ نفر دارای حقوق کم، ۳۰ نفر دارای حقوق متوسط و ۲۰ نفر دارای حقوق زیاد باشند آنگاه فراوانی‌های سه طبقه حقوق کم، متوسط و زیاد در بین این کارمندان به ترتیب $f_1=50$، $f_2=30$ و $f_3=20$ با فراوانی‌های نسبی $r_1=0/5$، $r_2=0/3$ و $r_3=0/2$ و فراوانی‌های تجمعی $g_1=50$، $g_2=80$ و $g_3=100$ و فراوانی‌های تجمعی نسبی $s_1=0/5$، $s_2=0/8$ و $s_3=1$ می‌باشد.

یک جدول فراوانی جدولی است که ستونهای آن شامل نوع داده‌ها (طبقات)، فراوانی، فراوانی نسبی، فراوانی تجمعی و فراوانی تجمعی نسبی داده‌ها باشد. در قسمتهای بعد این جدول را برای داده‌های گسسته و پیوسته به ترتیب تشکیل می‌دهیم.

۱.۳.۱ جدول فراوانی برای داده‌های گسسته

با ارائه مثالهای زیر نحوهٔ تشکیل جدول فراوانی را برای داده‌های گسسته بیان می‌کنیم.

مثال ۲.۳.۱ صنعتگری چهار نوع قطعه A، B، C و D تولید می‌کند. اگر او در یک روز ۲۰ قطعه از این قطعات را به شرح زیر تولید کرده باشد

B, C, C, A, D, C, C, B, D, C, A, C, D, C, B, C, C, B, D, D

یک جدول فراوانی برای این قطعات تشکیل دهید.

حل ابتدا چهار نوع قطعه A، B، C و D را به ترتیب با اعداد ۱، ۲، ۳ و ۴ متناظر می‌کنیم. این اعداد تنها برای نامگذاری این قطعات است و نمی‌توان روی آنها چهار عمل اصلی حساب را انجام داد.

سپس بوسیلهٔ شمارش، فراوانی هر قطعه را محاسبه کرده و با محاسبهٔ فراوانی نسبی، فراوانی تجمعی و فراوانی تجمعی نسبی، جدول فراوانی به صورت جدول ۱.۱ را به دست می‌آوریم.

نوع قطعه	x_i	f_i	r_i	g_i	s_i
A	۱	۲	۰/۱	۲	۰/۱
B	۲	۴	۰/۲	۶	۰/۳
C	۳	۹	۰/۴۵	۱۵	۰/۷۵
D	۴	۵	۰/۲۵	۲۰	۱
جمع		۲۰	۱/۰		

جدول ۱.۱ جدول فراوانی ۲۰ قطعهٔ تولیدی صنعتگر در یک روز

از این جدول اطلاعات زیادی را می‌توان استخراج کرد. برای مثال عدد ۰/۲ در ستون فراوانی نسبی به معنای آن است که ۲۰٪ از قطعات تولیدی آن روز از نوع B می‌باشند و عدد ۰/۳ در ستون فراوانی تجمعی نسبی به معنای آن است که ۳۰٪ از قطعات تولیدی آن روز از نوع A یا B می‌باشند.

مثال ۳.۳.۱ تعداد قرص‌های سرما خوردگی که در ۵۰ خانواده در عرض یک ماه زمستان مصرف می‌شود عبارت‌اند از

۷, ۵, ۳, ۳, ۴, ۵, ۳, ۲, ۸, ۳, ۳, ۲, ۴, ۴, ۳, ۶, ۸, ۶, ۷, ۴, ۵, ۴, ۶, ۴, ۵

۲, ۳, ۴, ۲, ۷, ۳, ۵, ۴, ۶, ۲, ۲, ۳, ۴, ۵, ۴, ۸, ۴, ۳, ۲, ۲, ۶, ۴, ۵, ۷, ۸

یک جدول فراوانی برای این داده‌ها تشکیل دهید. چند درصد خانواده‌ها بیش از ۴ قرص در ماه مصرف می‌کنند؟

حل داده‌ها از طریق شمارش تعداد قرصهای سرماخوردگی مصرف شده بوسیلهٔ اعداد ۲، ۳، ... و ۸ به دست آمده‌اند و روی آنها می‌توان چهار عمل اصلی حساب را انجام داد. همانند مثال قبل جدول فراوانی برای این داده‌ها را به صورت جدول ۲.۱ تشکیل می‌دهیم.

با توجه به عدد ۰/۶۰ در ستون فراوانی تجمعی نسبی، ۶۰٪ از خانواده‌ها حداکثر ۴ قرص در ماه مصرف می‌کنند و بنابراین ۴۰٪ از خانواده‌ها بیش از ۴ قرص در ماه مصرف می‌کنند.

x_i	f_i	r_i	g_i	s_i
۲	۸	۰/۱۶	۸	۰/۱۶
۳	۱۰	۰/۲۰	۱۸	۰/۳۶
۴	۱۲	۰/۲۴	۳۰	۰/۶۰
۵	۷	۰/۱۴	۳۷	۰/۷۴
۶	۵	۰/۱۰	۴۲	۰/۸۴
۷	۴	۰/۰۸	۴۶	۰/۹۲
۸	۴	۰/۰۸	۵۰	۱/۰۰
جمع	۵۰	۱/۰۰		

جدول ۲.۱ جدول فراوانی قرصهای سرماخوردگی مصرف شده ۵۰ خانواده در یک ماه

۲.۳.۱ جدول فراوانی برای داده‌های پیوسته

با ارائه مثالهای زیر نحوهٔ تشکیل جدول فراوانی برای داده‌های پیوسته را بیان می‌کنیم.

مثال ۴.۳.۱ وزنهای ۴۰ قالب کره که به نزدیکترین عدد صحیح گرد شده‌اند به قرار زیر است

۵۲	۳۵	۲۴	۴۷	۳۶	۵۱	۳۴	۳۸	۴۶	۳۳
۴۷	۳۶	۳۸	۵۰	۴۷	۳۴	۴۱	۴۰	۴۲	۴۰
۲۶	۲۹	۳۰	۳۲	۳۰	۳۵	۳۷	۳۷	۴۱	۲۱
۳۱	۳۰	۲۶	۳۵	۴۵	۲۳	۴۳	۳۱	۳۴	۴۳

یک جدول فراوانی برای این داده‌ها تشکیل دهید.

حل داده‌ها از نوع پیوسته می‌باشند. در این حالت داده‌ها را به تعدادی رده (فاصله) با طول مساوی تقسیم کرده و در هر رده فراوانی داده‌ها را می‌شماریم. روش به دست آوردن تعداد رده‌ها و طول هر رده به ترتیب در زیر آورده شده است.

۱- برای به دست آوردن تعداد رده‌ها یک قاعدهٔ عمومی وجود ندارد و معمولاً تعداد رده‌ها را بین ۵ تا ۲۵ رده اختیار می‌کنند. یک قاعدهٔ مفید استفاده از دستور استورگس (Sturges) است که در آن تعداد رده k از رابطهٔ زیر محاسبه می‌شود

$$\boxed{k=۱+۳/۳۲۲\log_{۱۰}^{n}}$$

که در آن n تعداد کل داده‌هاست. چون حاصل این عدد اعشاری است آن را به عدد صحیح بزرگتر از آن گرد می‌کنند. بنابراین در این مثال داریم که

$$k=۱+۳/۳۲۲\log_{۱۰}^{۴۰}=۶/۳۲۲\simeq ۷$$

۲- با توجه به اینکه وزنها به نزدیکترین عدد صحیح گرد شده‌اند بنابراین عدد ۳۵ در داده‌ها در واقع عددی در فاصلهٔ (۳۵/۵-۳۴/۵) می‌باشد. عدد ۰/۵ را میزان تغییر پذیری مقادیر داده‌ها نامیده و به صورت زیر آن را محاسبه می‌کنیم

$$\text{میزان تغییر پذیری داده‌ها} = S = \frac{\text{واحد گرد شدهٔ داده‌ها}}{۲} = \frac{۱}{۲} = ۰/۵$$

۳- کوچکترین، بزرگترین و دامنهٔ واقعی داده‌ها را به صورت زیر محاسبه می‌کنیم

$$\min = \text{کوچکترین داده} - S = ۲۱-۰/۵ = ۲۰/۵$$

$$\max = \text{بزرگترین داده} + S = ۵۲+۰/۵ = ۵۲/۵$$

$$\text{دامنه} = R = \max-\min = ۵۲/۵-۲۰/۵ = ۳۲$$

۴- طول هر رده را از تقسیم دامنهٔ R بر تعداد رده k به دست می‌آوریم و عدد حاصل را که ممکن است دارای چند رقم اعشار باشد مطابق واحد گرد شدهٔ داده‌ها به عدد بالاتر گرد می‌کنیم. بنابراین

$$\text{طول رده} = w = \frac{R}{k} = \frac{۳۲}{۷} = ۴/۵۷۱۴ \approx ۵$$

چون داده‌ها به عدد صحیح گرد شده‌اند و اولین عدد صحیح بزرگتر از ۴/۵۷۱۴ عدد ۵ می‌باشد پس طول رده به ۵ گرد شده است.

حال بایستی ۷ رده هر یک به طول ۵ را تشکیل دهیم. برای این منظور در جدول فراوانی یک ستون به نام رده‌ها ایجاد می‌کنیم و در آن اولین رده به طول ۵ را به صورت ۲۵/۵-۲۰/۵ (با شروع از min با طول w) در نظر گرفته و ردهٔ بعدی را به صورت ۳۰/۵-۲۵/۵ (شروع از انتهای ردهٔ قبل با طول w) در نظر می‌گیریم و این عمل را تا تشکیل ۷ رده ادامه می‌دهیم. برای شمارش فراوانی هر طبقه در جدول فراوانی از ستونی به نام خط و نشان استفاده می‌کنیم که در این ستون با خط زدن هر داده و مشخص کردن مکان آن در طبقهٔ مربوطه بوسیلهٔ یک نشان، فراوانی داده‌ها در

طبقات مشخص می‌گردند. توجه کنید که اگر بوسیلهٔ انجام مراحل فوق آخرین رده دارای فراوانی صفر باشد، آن ردهٔ آخر را حذف می‌کنیم. در جدول فراوانی داده‌های پیوسته نقاط وسط رده‌ها را با x_i نمایش می‌دهیم و آن را نمایندهٔ رده می‌نامیم. با انجام عملیات گفته شده در بالا جدول فراوانی مربوط به وزن قالبهای کره به صورت جدول ۳.۱ به دست می‌آید

رده‌ها	خط و نشان	x_i	f_i	r_i	g_i	s_i
۲۰/۵-۲۵/۵	𝍷𝍷𝍷	۲۳	۳	۰/۰۷۵	۳	۰/۰۷۵
۲۵/۵-۳۰/۵	𝍸 𝍷	۲۸	۶	۰/۱۵	۹	۰/۲۲۵
۳۰/۵-۳۵/۵	𝍸 𝍸	۳۳	۱۰	۰/۲۵	۱۹	۰/۴۷۵
۳۵/۵-۴۰/۵	𝍸 𝍷𝍷𝍷	۳۸	۸	۰/۲۰	۲۷	۰/۶۷۵
۴۰/۵-۴۵/۵	𝍸 𝍷	۴۳	۶	۰/۱۵	۳۳	۰/۸۲۵
۴۵/۵-۵۰/۵	𝍸	۴۸	۵	۰/۱۲۵	۳۸	۰/۹۵
۵۰/۵-۵۵/۵	𝍷𝍷	۵۳	۲	۰/۰۵	۴۰	۱/۰۰
جمع			۴۰	۱/۰۰		

جدول ۳.۱ جدول فراوانی وزنهای ۴۰ قالب کره

با استفاده از این جدول مشخص می‌شود که ۲۵٪ از قالبهای کره دارای وزنی بین ۳۵/۵-۳۰/۵ می‌باشند و ۴۷/۵٪ از قالبهای کره دارای وزنی کمتر از ۳۵/۵ می‌باشند.(چرا؟)

مثال ۵.۳.۱ داده‌های زیر یک نمونه ۵۰ تایی از اندازه نیروی پارگی نخهای کتان می‌باشد

۲۶/۴	۲۳/۰	۲۳/۹	۳۰/۲	۲۹/۵	۳۲/۷	۲۵/۰	۲۷/۱	۲۸/۳	۲۱/۲
۲۳/۵	۲۹/۲	۳۶/۸	۲۹/۰	۱۷/۳	۲۹/۳	۲۱/۹	۲۹/۴	۳۳/۷	۲۷/۳
۱۸/۶	۳۴/۸	۲۸/۷	۲۶/۸	۲۹/۶	۳۳/۵	۳۷/۵	۲۱/۸	۲۹/۵	۲۰/۶
۲۸/۳	۳۷/۰	۳۳/۲	۳۱/۳	۲۲/۷	۲۲/۲	۲۹/۶	۲۷/۵	۳۴/۱	۲۵/۴
۲۴/۰	۳۸/۴	۲۳/۶	۳۴/۵	۲۵/۴	۲۸/۱	۲۴/۸	۲۸/۹	۲۴/۶	۳۶/۹

یک جدول فراوانی برای این داده‌ها تشکیل دهید.

حل با توجه به روش گفته شده در مثال ۴.۳.۱، محاسبات زیر را برای این مثال انجام می‌دهیم

$$k=۱+۳/۳۲۲\log_{۱۰}^{۵۰}=۶/۶۴۴\approx ۷$$

$$S=\frac{۰/۱}{۲}=۰/۰۵$$

$$min=۱۷/۳-۰/۰۵=۱۷/۲۵$$

$$max=۳۸/۴+۰/۰۵=۳۸/۴۵$$

$$R=۳۸/۴۵-۱۷/۲۵=۲۱/۲$$

$$w=\frac{۲۱/۲}{۷}=۳/۰۲۸۶\approx ۳/۱$$

بنابراین بایستی ۷ رده هر یک به طول ۳/۱ را تشکیل دهیم. جدول فراوانی حاصل در جدول ۴.۱ آورده شده است.

رده‌ها	خط و نشان	x_i	f_i	r_i	g_i	s_i
۱۷/۲۵-۲۰/۳۵	||	۱۸/۸	۲	۰/۰۴	۲	۰/۰۴
۲۰/۳۵-۲۳/۴۵	卌 ||	۲۱/۹	۷	۰/۱۴	۹	۰/۱۸
۲۳/۴۵-۲۶/۵۵	卌 卌	۲۵/۰	۱۰	۰/۲۰	۱۹	۰/۳۸
۲۶/۵۵-۲۹/۶۵	卌 卌 卌 ||	۲۸/۱	۱۷	۰/۳۴	۳۶	۰/۷۲
۲۹/۶۵-۳۲/۷۵	|||	۳۱/۲	۳	۰/۰۶	۳۹	۰/۷۸
۳۲/۷۵-۳۵/۸۵	卌 |	۳۴/۳	۶	۰/۱۲	۴۵	۰/۹۰
۳۵/۸۵-۳۸/۹۵	卌	۳۷/۴	۵	۰/۱۰	۵۰	۱/۰۰
جمع			۵۰	۱/۰۰		

جدول ۴.۱ جدول فراوانی اندازهٔ نیروی پارگی ۵۰ نخ کتان

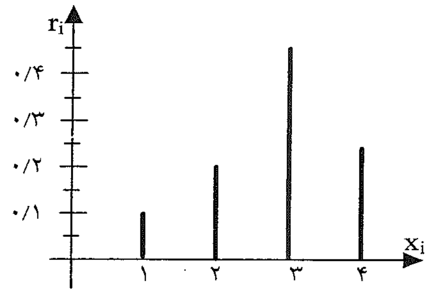
۴.۱ نمودارهای آماری

دومین گام در خلاصه کردن داده‌ها تبدیل جدول فراوانی به نمودارهایی است که بوسیلهٔ این نمودارها بتوان اطلاعات نهفته در داده‌ها و جدول فراوانی را تا حدودی به طور عینی و بدون توضیح و تشریح اضافی دید. نمودارهای مختلف و گوناگونی در آمار وجود دارند که در این قسمت بعضی از آنها را برای داده‌های گسسته و پیوسته می‌آوریم.

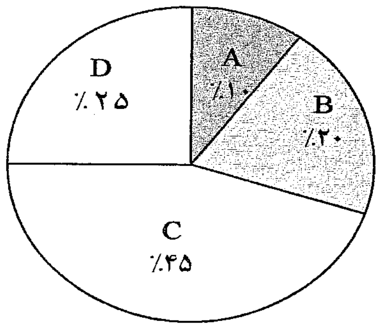
۱.۴.۱ نمودارهای آماری برای داده‌های گسسته

برای داده‌های گسسته دو نوع نمودار میله‌ای و دایره‌ای را در زیر معرفی می‌کنیم.

الف- نمودار میله‌ای در این نمودار دو محور عمود بر هم در نظر می‌گیریم و بر روی محور افقی مقادیر x_iها و بر روی محور عمودی مقادیر فراوانی نسبی r_iها را نمایش می‌دهیم. سپس در هر مقدار x_i میله‌ای به ارتفاع فراوانی نسبی r_i مربوط به آن طبقه رسم می‌کنیم. نمودار حاصله را نمودار میله‌ای داده‌ها گویند. برای مثال در شکل ۱.۱ نمودار میله‌ای داده‌های مثال ۲.۳.۱ رسم شده است. در این نمودار علاوه بر مشاهدهٔ اطلاعات جدول فراوانی، می‌توان مقایسه‌ای بین طبقات مختلف انجام داد.



شکل ۱.۱ نمودار میله‌ای مربوط به ۲۰ قطعه تولیدی یک صنعتگر در یک روز



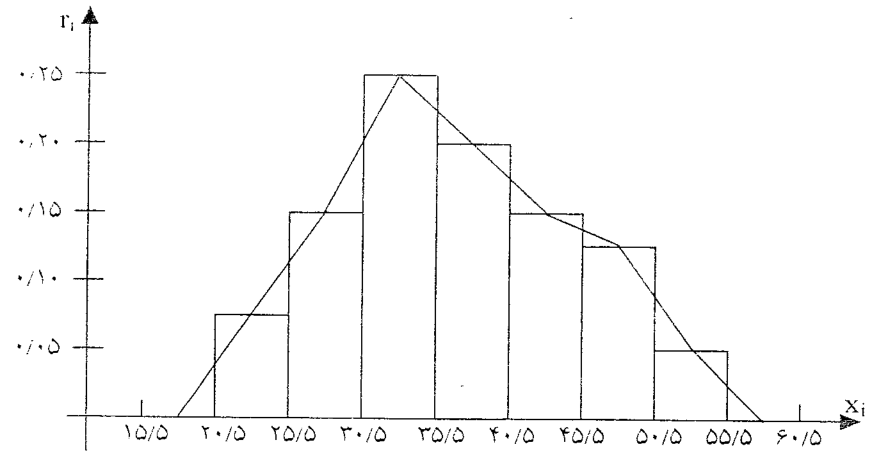
شکل ۲.۱ نمودار دایره‌ای مربوط به ۲۰ قطعهٔ تولیدی یک صنعتگر در یک روز

ب- نمودار دایره‌ای در این نمودار دایره‌ای را رسم کرده و این دایره را به تعداد طبقات جدول فراوانی به قطعهایی تقسیم می‌کنیم، به طوری که اندازهٔ هر قطاع متناسب با فراوانی نسبی طبقهٔ مربوطه باشد. برای مثال، در مثال ۲.۳.۱ طبقهٔ دوم (قطعهٔ نوع B) دارای فراوانی نسبی ۰/۲ می‌باشد و برای این طبقه یک قطاع به اندازهٔ $۷۲^\circ=۰/۲\times۳۶۰^\circ$ را در دایره در نظر می‌گیریم و این عمل را برای طبقات دیگر نیز تکرار می‌کنیم. شکل ۲.۱ نمودار دایره‌ای داده‌های مثال ۲.۳.۱ را نشان می‌دهد.

۲.۴.۱ نمودارهای آماری برای داده‌های پیوسته

برای داده‌های پیوسته سه نمودار را در زیر معرفی می‌کنیم

الف- هیستوگرام (نمودار ستونی) هیستوگرام نموداری متشکل از تعدادی مستطیل است که تعداد این مستطیل‌ها برابر تعداد رده‌های جدول فراوانی می‌باشد. قاعده هر مستطیل روی محور افقی قرار دارد و طول آن برابر طول واقعی رده است، که هر چه باشد آن را یک واحد در نظر می‌گیریم و مرکز آن نمایندهٔ رده است. ارتفاع هر مستطیل برابر فراوانی نسبی مربوط به آن رده است. برای مثال هیستوگرام مربوط به مثال ۴.۳.۱ در شکل ۳.۱ رسم شده است. توجه کنید که چون عرض هر مستطیل برابر یک واحد در نظر گرفته شده و ارتفاع هر مستطیل برابر فراوانی نسبی ردهٔ مربوطه می‌باشد پس مجموع مساحت تمام مستطیل‌های هیستوگرام برابر یک واحد مربع است.

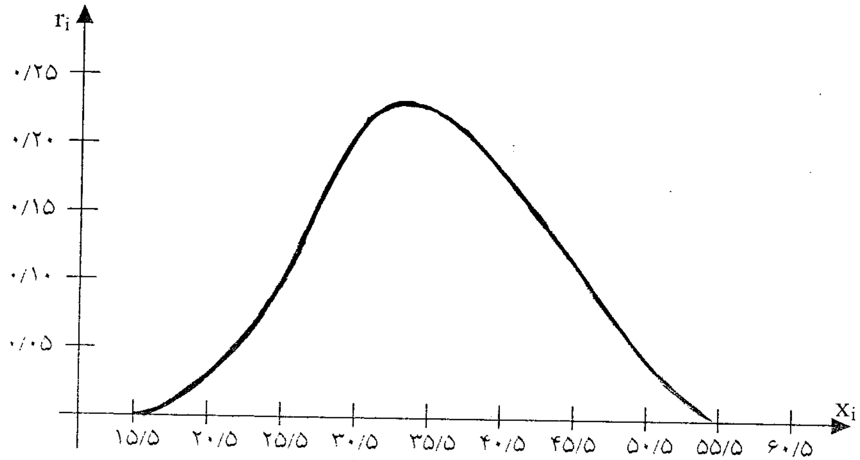


شکل ۳.۱ نمودار هیستوگرام و چندبر فراوانی مربوط به وزنهای ۴۰ قالب کره

ب- چندبر فراوانی اگر وسط قاعده‌های بالای مستطیل‌های هیستوگرام را بوسیلهٔ خطوط مستقیم به طور متوالی به یکدیگر وصل کرده و ابتدای آن را به وسط ردهٔ ماقبل و انتهای آن را به وسط ردهٔ مابعد مستطیل‌های هیستوگرام وصل کنیم یک چند ضلعی بوجود می‌آید که آن را چندبر فراوانی گویند و مساحت زیر این چندبر فراوانی نیز یک واحد مربع است. در شکل ۳.۱ چندبر فراوانی مربوط به مثال ۴.۳.۱ نمایش داده شده است.

ج- منحنی فراوانی اگر تعداد داده‌ها زیاد و طول رده‌ها کوچک باشند در این صورت تعداد رده‌ها زیاد شده و در نتیجه تعداد اضلاع چندبر فراوانی نیز زیاد می‌شود و به یک منحنی نزدیک می‌شود.

مساحت زیر این منحنی یک واحد مربع است و به آن منحنی فراوانی گویند. در شکل ۴.۱ منحنی فراوانی تقریبی مربوط به داده‌های مثال ۴.۳.۱ رسم شده است.



شکل ۴.۱ نمودار منحنی فراوانی تقریبی مربوط به وزنهای ۴۰ قالب کره

۵.۱ خلاصه کردن داده‌ها در چند عدد

با تشکیل جدول فراوانی و رسم نمودارها می‌توان اطلاعات نهفته در داده‌ها را تا حدودی مشخص کرد. با این حال برای اینکه بتوانیم نتایج کلی دربارهٔ صفت مورد مطالعه به دست آورده و این نتایج را به سادگی گزارش کنیم، بهتر است که داده‌ها را در یک یا چند عدد خلاصه کنیم. چنین اعدادی را شاخص یا معیار گویند و به دو نوع، شاخص‌های تمرکز و شاخص‌های پراکندگی، موسوم‌اند. در این بخش این نوع شاخص‌ها را معرفی می‌کنیم.

۱.۵.۱ شاخص‌های تمرکز

شاخص‌های تمرکز مقادیری هستند که معمولاً در حوالی مرکز منحنی فراوانی قرار گرفته‌اند. مهمترین شاخص‌های تمرکز میانگین، میانه و نما می‌باشند که در این قسمت به شرح آنها می‌پردازیم.

الف- میانگین فرض کنید داده‌های x_1، x_2، ... و x_k به ترتیب دارای فراوانیهای f_1، f_2، ... و

f_k باشند و تعداد کل این داده‌ها برابر n باشد. مجموع داده‌ها تقسیم بر تعداد داده‌ها را میانگین (حسابی) این داده‌ها گویند، یعنی

$$\boxed{\bar{x}=\frac{1}{n}\sum_{i=1}^{k} f_i x_i}$$

در صورتی که داده‌ها پیوسته باشند، x_iها را نمایندهٔ رده‌ها در نظر می‌گیریم. توجه کنید که میانگین به عنوان یک شاخص تمرکز برای داده‌های عددی مناسب می‌باشد.

مثال ۱.۵.۱ برای داده‌های مثال ۳.۳.۱، میانگین داده‌ها را به دست آورید.

حل با استفاده از جدول ۲.۱، برای محاسبهٔ میانگین یک ستون f_ix_i که از ضرب x_i (نماینده رده) در فراوانی طبقه حاصل می‌شود، به جدول اضافه می‌کنیم تا محاسبهٔ میانگین به راحتی انجام گیرد. در جدول ۵.۱ این محاسبات انجام گرفته و میانگین این داده‌ها به صورت زیر به دست می‌آید

x_i	f_i	f_ix_i
۲	۸	۱۶
۳	۱۰	۳۰
۴	۱۲	۴۸
۵	۷	۳۵
۶	۵	۳۰
۷	۴	۲۸
۸	۴	۳۲
جمع	۵۰	۲۱۹

جدول ۵.۱ جدول محاسبه میانگین

$$\sum_{i=1}^{k} f_ix_i=۲۱۹ \Rightarrow \bar{x}=\frac{۲۱۹}{۵۰}=۴/۳۸$$

یعنی هر خانواده به طور متوسط ۴/۳۸ قرص در ماه مصرف می‌کند.

مثال ۲.۵.۱ برای داده‌های مثال ۴.۳.۱، میانگین داده‌ها را به دست آورید.

حل با استفاده از جدول ۳.۱ و در نظر گرفتن x_i به عنوان نمایندهٔ رده در فرمول میانگین و تشکیل یک جدول همانند جدول ۵.۱ به سادگی دیده می‌شود که

$$\sum_{i=1}^{k} f_ix_i=۱۴۷۵ \Rightarrow \bar{x}=\frac{۱۴۷۵}{۴۰}=۳۶/۸۷۵$$

یعنی هر قالب کره به طور متوسط ۳۶/۸۷۵ وزن دارد.

ب- میانه اگر داده‌ها را به طور غیر نزولی مرتب کنیم آنگاه عدد m را میانهٔ این داده‌ها گویند در صورتی که تقریباً نصف داده‌ها در سمت چپ و نصف داده‌ها در سمت راست این عدد قرار گیرند، روش محاسبهٔ میانه برای داده‌های گسسته و پیوسته متفاوت می‌باشد و در زیر هر یک را به طور جداگانه بررسی می‌کنیم.

محاسبهٔ میانه برای داده‌های گسسته فرض کنید $x_{(۱)}$، $x_{(۲)}$، ... و $x_{(n)}$ شکل مرتب شده داده‌های $x_۱$، $x_۲$، ... و x_n به طور غیر نزولی باشند. اگر تعداد داده‌ها n فرد باشد آنگاه $m=x_{(\frac{n+۱}{۲})}$، یعنی داده‌ای که در وسط قرار دارد میانه محسوب می‌شود و اگر n زوج باشد، آنگاه $m=\frac{x_{(\frac{n}{۲})}+x_{(\frac{n}{۲}+۱)}}{۲}$، یعنی نصف مجموع دو داده‌ای که در وسط قرار دارند میانه محسوب می‌شود.

مثال ۳.۵.۱ میانه را برای داده‌های ۱۲, ۱۸, ۲۵, ۱۲, ۱۷, ۱۹, ۲۱, ۲۱, ۱۷, ۱۹, ۱۵ به دست آورید.

حل ابتدا داده‌ها را به طور غیرنزولی مرتب می‌کنیم یعنی ۱۲, ۱۵, ۱۷, ۱۷, ۱۸, ۱۹, ۱۹, ۲۱, ۲۱, ۲۵, ۱۲. چون تعداد داده‌ها n=۱۱ فرد می‌باشد بنابراین $m=x_{(۶)}=۱۸$.

مثال ۴.۵.۱ میانه را برای داده‌های ۲, ۱۲, ۸, ۹, ۱۱, ۲, ۳, ۳, ۷, ۵ به دست آورید.

حل داده‌های مرتب شده عبارت‌اند از ۱۲ ,۱۱ ,۹ ,۸ ,۷ ,۵ ,۳ ,۳ ,۲ ,۲. چون تعداد داده‌ها n=۱۰ زوج می‌باشد بنابراین $m=\frac{x_{(۵)}+x_{(۶)}}{۲}=\frac{۵+۷}{۲}=۶$.

مثال ۵.۵.۱ میانه را برای داده‌های مثال ۳.۳.۱ به دست آورید.

حل چون تعداد داده‌ها n=۵۰ زوج می‌باشد بنابراین با توجه به ستون فراوانی جدول ۲.۱ داریم که

$$m=\frac{x_{(۲۵)}+x_{(۲۶)}}{۲}=\frac{۴+۴}{۲}=۴$$

یعنی تقریباً نیمی از خانواده‌ها حداکثر ۴ قرص و نیمی از خانواده‌ها حداقل ۴ قرص در ماه مصرف می‌کنند.

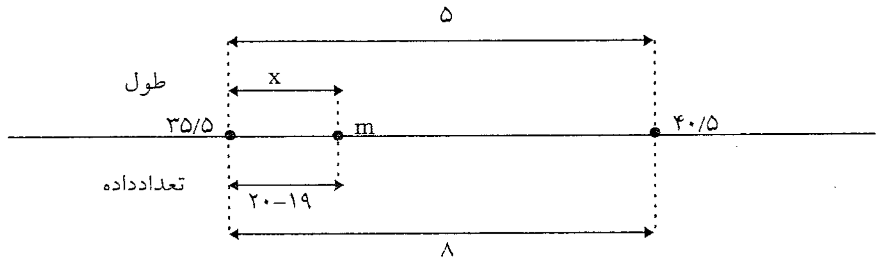
محاسبهٔ میانه برای داده‌های پیوسته روش به دست آوردن میانه برای داده‌های پیوسته را

ابتدا با یک مثال شرح می‌دهیم و سپس یک رابطهٔ کلی برای محاسبهٔ میانه ارائه می‌دهیم.

مثال ۶.۵.۱ میانه را برای داده‌های مثال ۴.۳.۱ به دست آورید.

حل برای پیدا کردن میانه برای این داده‌ها با استفاده از جدول ۳.۱ مراحل زیر را انجام می‌دهیم.

۱- ابتدا در جدول فراوانی نخستین رده‌ای که فراوانی تجمعی نسبی آن بزرگتر یا مساوی ۰/۵۰ باشد را در نظر می‌گیریم. این رده که میانه درون آن قرار دارد را ردهٔ میانه می‌نامیم. در این مثال ردهٔ میانه ۴۰/۵-۳۵/۵ می‌باشد.



شکل ۵.۱ محاسبه میانه برای داده‌های پیوسته

۲- فرض می‌کنیم که در ردهٔ میانه داده‌ها در فواصل مساوی توزیع شده‌اند در این صورت با توجه به شکل ۵.۱ می‌توان تناسب زیر را تشکیل داد.

طول	تعداد رده
۵	۸
$x=\frac{(۲۰-۱۹)۵}{۸}=۰/۶۲۵$	۲۰−۱۹

بنابراین با توجه به شکل ۵.۱ میانه به صورت زیر محاسبه می‌گردد.

$$m = \text{کران پائین ردهٔ میانه} + x = ۳۵/۵+\frac{(۲۰-۱۹)۵}{۸} = ۳۵/۵+۰/۶۲۵ = ۳۶/۱۲۵$$

با توجه به عملیات انجام شده در مثال ۶.۵.۱ می‌توان فرمول زیر را برای محاسبهٔ میانه ارائه داد.

$$\boxed{m=L_{۰/۵}+\frac{(۰/۵n-g_{۰/۵})w}{f_{۰/۵}}}$$

در این فرمول $L_{۰/۵}$ کران پائین ردهٔ میانه، n تعداد داده‌ها، $g_{۰/۵}$ فراوانی تجمعی ردهٔ قبل از ردهٔ میانه، $f_{۰/۵}$ فراوانی ردهٔ میانه و w طول رده می‌باشد.

مثال ۷.۵.۱ در مثال ۵.۳.۱ میانهٔ داده‌ها را محاسبه کنید.

حل با توجه به جدول ۴.۱ ردهٔ میانه ۲۹/۶۵-۲۶/۵۵ می‌باشد و عملیات مربوط به میانه عبارت‌اند از

$$L_{۰/۵}=۲۶/۵۵ \;,\; n=۵۰ \;,\; g_{۰/۵}=۱۹ \;,\; f_{۰/۵}=۱۷ \;,\; w=۳/۱$$

بنابراین

$$m=۲۶/۵۵+\frac{(۰/۵\times ۵۰-۱۹)۳/۱}{۱۷}=۲۶/۵۵+۱/۰۹۴=۲۷/۶۴۴$$

توجه کنید که میانه تحت تأثیر داده‌های بسیار بزرگ و یا بسیار کوچک قرار نمی‌گیرد، در حالیکه میانگین شدیداً تحت تأثیر چنین داده‌هایی قرار می‌گیرد.

ج- نما داده‌ای که فراوانی آن از سایر داده‌ها بیشتر باشد را نما یا مد نامیده و با نماد M نمایش می‌دهیم. همانند میانه روش محاسبهٔ نما برای داده‌های گسسته و پیوسته متفاوت می‌باشد که در زیر هر یک را جداگانه شرح می‌دهیم.

محاسبه نما برای داده‌های گسسته در این حالت ابتدا فراوانی داده‌ها را پیدا می‌کنیم و داده‌ای که فراوانی آن بیشتر باشد را به عنوان نما اختیار می‌کنیم. اگر دو دادهٔ غیر مجاور دارای فراوانی یکسان و بیش از سایر فراوانی‌ها باشند، هر دو را به عنوان نما اختیار کرده و داده‌ها را دو نمایی گوئیم و اگر این دو داده مجاور یکدیگر باشند نصف مجموع آنهارا به عنوان نما اختیار می‌کنیم. اگر تمام داده‌ها دارای فراوانی یکسان باشند گوئیم که داده‌ها بدون نما هستند.

مثال ۸.۵.۱ برای داده‌های ۲,۸,۳,۷,۳,۹,۲,۳,۷,۹,۳,۷ چون فراوانی داده ۳ از سایر داده‌ها بیشتر است پس نما برابر است با $M=۳$.

مثال ۹.۵.۱ برای داده‌های ۵,۶,۷,۳,۷,۹,۶,۵,۷,۵ چون فراوانی دو دادهٔ غیر مجاور ۵و۷ از سایر داده‌ها بیشتر است پس داده‌ها دو نمایی هستند و $M_1=۵$ و $M_2=۷$.

مثال ۱۰.۵.۱ برای داده‌های ۳,۸,۲,۵,۳,۶,۵,۸,۳,۲,۵ چون فراوانی دو دادهٔ مجاور ۳و۵ از سایر داده‌ها بیشتر است پس نما برابر است با $M=\frac{۳+۵}{۲}=۴$.

مثال ۱۱.۵.۱ برای داده‌های ۳,۶,۷,۴,۷,۶,۳,۴ چون فراوانی همهٔ داده‌ها یکسان است پس داده‌ها بدون نما هستند.

توجه کنید که نما می‌تواند به عنوان معیار تمرکز داده‌های گسسته که از یک متغیر کیفی حاصل شده‌اند مورد استفاده قرار گیرد.

مثال ۱۲.۵.۱ برای داده‌های مثال ۲.۳.۱ با توجه به جدول ۱.۱، قطعهٔ نوع C فراوانی بیشتری نسبت به سایر قطعات دارد و بنابراین نما برابر $M=۳$ و یا قطعه نوع C می‌باشد.

محاسبهٔ نما برای داده‌های پیوسته در این حالت داده‌ها را در یک جدول فراوانی مرتب می‌کنیم و رده‌ای که فروانی آن از سایر رده‌ها بیشتر است را به عنوان ردهٔ نمایی اختیار می‌کنیم. حال می‌توان نمایندهٔ این رده یعنی x_i را به عنوان نما اختیار کرد و یا اگر بخواهیم نما را به طور دقیق‌تر در این رده محاسبه کنیم از فرمول زیر استفاده می‌کنیم

$$\boxed{M=L_M+(\frac{D_۱}{D_۱+D_۲})w}$$

در این فرمول L_M کران پائین ردهٔ نمایی، $D_۱$ اختلاف فراوانی‌های نسبی ردهٔ نمایی و ردهٔ قبل از آن، $D_۲$ اختلاف فراوانی‌های نسبی ردهٔ نمایی و ردهٔ بعد از آن و w طول رده می‌باشد.

مثال ۱۳.۵.۱ در مثال ۴.۳.۱ نمای داده‌ها را محاسبه کنید.

حل با توجه به جدول ۳.۱ ردهٔ نمایی ۳۰/۵-۳۵/۵ می‌باشد و بنابراین محاسبات برای یافتن نما عبارت‌اند از

$$L_M=۳۰/۵\ ,\ D_۱=۰/۲۵-۰/۱۵=۰/۱۰\ ,\ D_۲=۰/۲۵-۰/۲۰=۰/۰۵\ ,\ w=۵$$

بنابراین

$$M=۳۰/۵+(\frac{۰/۱۰}{۰/۱۰+۰/۰۵})۵=۳۰/۵+۳/۳۳۳=۳۳/۸۳۳$$

مثال ۱۴.۵.۱ در مثال ۵.۳.۱ نمای داده‌ها را محاسبه کنید.

حل با توجه به جدول ۴.۱ ردهٔ نمایی ۲۶/۵۵-۲۹/۶۵ می‌باشد و بنابراین محاسبات برای یافتن نما عبارت‌اند از

$$L_M=۲۶/۵۵\ ,\ D_۱=۰/۳۴-۰/۲۰=۰/۱۴\ ,\ D_۲=۰/۳۴-۰/۰۶=۰/۲۸\ ,\ w=۳/۱$$

بنابراین

$$M=۲۶/۵۵+(\frac{۰/۱۴}{۰/۱۴+۰/۲۸})۳/۱=۲۶/۵۵+۱/۰۳۳=۲۷/۵۸۳$$

۲.۵.۱ شاخص‌های پراکندگی

بوسیلهٔ شاخص های تمرکز می‌توان میزان تمرکز داده‌ها را در یک عدد خلاصه کرد. اما یک مسئله مهم در ارتباط با داده‌های آماری، میزان تغییرات و پراکندگی آنها است، بدین معنی که اندازه‌گیری تا چه اندازه از فردی به فرد دیگر تغییر می‌کند. در این قسمت شاخص های پراکندگی را به عنوان معیاری برای سنجش میزان تغییرات داده‌ها معرفی می‌کنیم. ابتدا به مثال زیر توجه کنید.

مثال ۱۵.۵.۱ نمرات ۱۰ دانش‌آموز در دو امتحان تستی درسهای ریاضی و زبان به صورت زیر است:

درس ریاضی : ۰,۴,۴,۱۲,۱۲,۱۲,۱۴,۱۶,۱۶,۲۰

درس زبان : ۸,۸,۹,۹,۱۲,۱۲,۱۲,۱۳,۱۳,۱۴

همان طور که دیده می‌شود در هر دو کلاس میانگین نمره $\bar{x}=۱۱$، میانه $m=۱۲$ و نما $M=۱۲$ می‌باشد، اما نحوهٔ تغییر پذیری نمرات در دو درس نسبت به میانگین ۱۱ یا میانه و نمای ۱۲ متفاوت می‌باشد. در درس ریاضی پراکندگی نمرات زیاد ولی در درس زبان پراکندگی نمرات کمتر است.

با توجه به مثال بالا برای اینکه بتوانیم این دو درس را با یکدیگر مقایسه کنیم بایستی از شاخص های دیگری بنام شاخص های پراکندگی استفاده کنیم که در زیر مهمترین این شاخص ها را معرفی می‌کنیم.

الف– دامنهٔ داده‌ها اگر $x_{(۱)}$ کوچکترین داده و $x_{(n)}$ بزرگترین داده باشد آنگاه

$$\boxed{R=x_{(n)}-x_{(۱)}}$$

را دامنهٔ داده‌ها گویند. برای مثال در مثال ۱۵.۵.۱ دامنهٔ نمرات دروس ریاضی و زبان به ترتیب $R_۱=۲۰-۰=۲۰$ و $R_۲=۱۴-۸=۶$ می‌باشد که نشان می‌دهد نمرات درس زبان دارای پراکندگی کمتری است.

با وجود اینکه این شاخص وسعت پراکندگی داده‌ها را معرفی می‌کند اما یک شاخص مناسب و کافی برای سنجش تغییر پذیری داده‌ها نیست زیرا تنها به کوچکترین و بزرگترین داده وابسته است.

ب– میانگین انحرافات اختلاف مثبت داده x_i از میانگین یعنی $|x_i-\bar{x}|$ را انحراف از میانگین داده x_i گویند و میانگین تمام انحرافات یعنی $D=\frac{۱}{n}\sum_{i=۱}^{k}f_i|x_i-\bar{x}|$ را میانگین

انحرافات گویند. میانگین انحرافات یک شاخص مناسب برای سنجش تغییر پذیری داده‌ها است، زیرا به تمامی داده‌ها و انحرافات آنها وابسته است. اما روش محاسبهٔ آن مشکل می‌باشد و به دلیل وجود قدر مطلق در فرمول آن، نمی‌توان آن را ساده کرد تا محاسبات آن ساده شوند و به جای آن از شاخص مناسب زیر استفاده می‌کنیم.

ج- واریانس و انحراف استاندارد میانگین مجذور انحرافات را واریانس می‌نامند و با نماد S_b^2 نمایش می‌دهند، یعنی

$$\boxed{S_b^2=\frac{1}{n}\sum_{i=1}^{k} f_i(x_i-\bar{x})^2}$$

در مبحث استنباط آماری واریانس را از مجموع مجذور انحرافات داده‌ها تقسیم بر n−۱ به دست می‌آورند و آنرا با نماد S^2 نمایش می‌دهند، یعنی

$$\boxed{S^2=\frac{1}{n-1}\sum_{i=1}^{k} f_i(x_i-\bar{x})^2}$$

دلیل آماری این کار در مبحث استنباط آماری مشخص خواهد شد. در این کتاب هر جا صحبت از واریانس می‌کنیم منظور S^2 می‌باشد. جذر واریانس یعنی $S=\sqrt{S^2}$ را انحراف استاندارد گویند و انحراف استاندارد شاخص مناسبی برای سنجش پراکندگی می‌باشد. در قضیهٔ زیر روش ساده‌ای برای محاسبهٔ واریانس می‌آوریم.

قضیهٔ ۱.۱ واریانس را می‌توان از فرمول زیر محاسبه کرد

$$\boxed{S^2=\frac{1}{n-1}\left[\sum_{i=1}^{k} f_i x_i^2-\frac{1}{n}\left(\sum_{i=1}^{k} f_i x_i\right)^2\right]}$$

اثبات این قضیه در تمرین ۱۳ از متعلمین خواسته شده است.

مثال ۱۶.۵.۱ در مثال ۴.۳.۱ واریانس و انحراف استاندارد داده‌ها را محاسبه کنید.

حل با استفاده از جدول ۳.۱، جدول ۶.۱ را برای محاسبهٔ واریانس و انحراف استاندارد تشکیل می‌دهیم. با استفاده از این جدول و قضیهٔ ۱.۱ داریم

$$S^2=\frac{1}{39}\left[56965-\frac{(1475)^2}{40}\right]=66/0096\quad,$$
$$S=\sqrt{66/0096}=8/12$$

x_i	f_i	f_ix_i	$f_ix_i^2$
۲۳	۳	۶۹	۱۵۸۷
۲۸	۶	۱۶۸	۴۷۰۴
۳۳	۱۰	۳۳۰	۱۰۸۹۰
۳۸	۸	۳۰۴	۱۱۵۵۲
۴۳	۶	۲۵۸	۱۱۰۹۴
۴۸	۵	۲۴۰	۱۱۵۲۰
۵۳	۲	۱۰۶	۵۶۱۸
جمع	۴۰	۱۴۷۵	۵۶۹۶۵

جدول ۶.۱ جدول محاسبهٔ واریانس

روش تبدیل داده‌ها اگر داده‌ها بزرگ باشند محاسبات مربوط به واریانس مشکل می‌شود. در این حالت می‌توان برای محاسبهٔ واریانس از روش تبدیل داده‌ها استفاده کرد. در این روش داده‌های x_1، x_2، ... و x_k را بوسیلهٔ تبدیل زیر به داده‌های y_1، y_2، ... و y_k تبدیل می‌کنیم

$$\boxed{y_i=\frac{x_i-a}{b}\qquad i=1,2,\dots,k}$$

که در آن اعداد a و b مقادیر ثابتی هستند و به ترتیب برای تغییر مبدأ و تغییر واحد اندازه‌گیری داده‌ها به کار برده می‌شوند. معمولاً در داده‌های پیوسته a را برابر نماینده ردهٔ نمایی و b را برابر طول رده یعنی w می‌گیرند. سپس میانگین و واریانس داده‌های جدید y_1، y_2، ... و y_k را محاسبه کرده و به سادگی می‌توان میانگین و واریانس داده‌های اصلی x_1، x_2، ... و x_k را از فرمولهای زیر به دست آورد.

$$\boxed{\bar{x}=a+b\bar{y}\quad,\quad S_x^2=b^2S_y^2\quad,\quad S_x=bS_y}$$

مثال ۱۷.۵.۱ در مثال ۵.۳.۱ میانگین، واریانس و انحراف استاندارد داده‌ها را از روش تبدیل داده‌ها محاسبه کنید.

حل با استفاده از جدول ۴.۱ و قرار دادن a=۲۸/۱ و b=۳/۱ جدول ۷.۱ را برای محاسبهٔ میانگین، واریانس و انحراف استاندارد تشکیل می‌دهیم. با استفاده از این جدول می‌توان محاسبات

زیر را انجام داد

x_i	f_i	$y_i=\frac{x_i-۲۸/۱}{۳/۱}$	f_iy_i	$f_iy_i^۲$
۱۸/۸	۲	-۳	-۶	۱۸
۲۱/۹	۷	-۲	-۱۴	۲۸
۲۵/۰	۱۰	-۱	-۱۰	۱۰
۲۸/۱	۱۷	۰	۰	۰
۳۱/۲	۳	۱	۳	۳
۳۴/۳	۶	۲	۱۲	۲۴
۳۷/۴	۵	۳	۱۵	۴۵
جمع	۵۰		۰	۱۲۸

جدول ۷.۱ جدول محاسبهٔ واریانس به روش تبدیل داده‌ها

$$\bar{y}=\frac{۰}{۵۰}=۰\ ,\ S_y^۲=\frac{۱}{۴۹}\left[۱۲۸-\frac{(۰)^۲}{۵۰}\right]=۲/۶۱۲۲$$

$$\bar{x}=۲۸/۱+۳/۱(۰)=۲۸/۱\ ,\ S_x^۲=(۳/۱)^۲(۲/۶۱۲۲)=۲۵/۱۰۳\ ,\ S_x=۵/۰۱$$

د- ضریب تغییر واریانس و انحراف استاندارد به واحد اندازه‌گیری داده‌ها بستگی دارند. برای مقایسه دو سری داده بایستی از شاخص‌هایی استفاده کنیم که به واحد اندازه‌گیری داده‌ها بستگی نداشته باشند. یکی از این شاخص‌ها ضریب تغییر می‌باشد که به صورت $CV=\frac{S}{\bar{x}}$ تعریف می‌شود و معمولاً بر حسب درصد بیان می‌شود.

مثال ۱۸.۵.۱ کارخانه‌ای دو نوع لامپ تولید می‌کند. لامپ نوع اول دارای میانگین طول عمر ۲۰۰ ساعت با انحراف استاندارد ۱۱ ساعت و لامپ نوع دوم دارای میانگین طول عمر ۲۴۰ ساعت با انحراف استاندارد ۱۲ ساعت است. کدام نوع لامپ بهتر است؟

حل

$$\bar{x}_۱=۲۰۰\ ,\ S_۱=۱۱\ ,\ CV_۱=\frac{۱۱}{۲۰۰}=۰/۰۵۵\ \Rightarrow\ CV_۱=\%۵/۵$$

$$\bar{x}_۲=۲۴۰\ ,\ S_۲=۱۲\ ,\ CV_۲=\frac{۱۲}{۲۴۰}=۰/۰۵\ \Rightarrow\ CV_۲=\%۵$$

بنابراین لامپ نوع دوم بهتر است زیرا دارای میانگین طول عمر بیشتر و ضریب تغییر کمتری است.

۶.۱ تمرینات

۱ چه تفاوتی بین آمار توصیفی و آمار استنباطی وجود دارد؟

۲ در هر یک از موارد زیر جمعیت، نمونه و صفت مربوطه را مشخص کنید.

الف- آموزگاری می‌خواهد میانگین نمرات درس ریاضی کلاس پنجم دبستان را، در دو منطقهٔ آموزشی مقایسه کند.

ب- مدیر یک فروشگاه زنجیره‌ای می‌خواهد میانگین فروش روزانه محصول جدیدی را در یکی از فروشگاهها برآورد کند.

ج- یک شرکت اتومبیل سازی می‌خواهد میزان مصرف بنزین در صد کیلومتر اتومبیلهای تولیدی شرکت در سال گذشته را برآورد کند.

۳ کدام یک از داده‌های زیر گسسته و کدام یک پیوسته هستند؟

الف- میزان بارندگی بر حسب سانتی متر، در یک شهر در طول ماههای سال.

ب- تعداد دانشجویان ورودی به دانشگاه در سالهای مختلف.

ج- طول عمر لامپهای تلویزیونی که توسط یک کارخانه تولید می‌شود.

د- اندازهٔ طول ۱۰۰۰ گلولهٔ تولید شده در یک کارخانهٔ اسلحه سازی.

ه) تعداد سهام فروخته شده در بازار بورس در هر روز.

۴ تعداد اتومبیلهای تولید شده توسط یک شرکت اتومبیل سازی در ۱۸ روز کاری به شرح زیر می‌باشد

۱۱ ۱۲ ۱۴ ۱۳ ۱۲ ۱۵ ۱۶ ۱۱ ۱۲

۱۴ ۱۰ ۱۲ ۱۵ ۱۲ ۱۲ ۱۳ ۱۱ ۱۷

یک جدول فراوانی برای این داده‌ها تشکیل دهید و نمودار میله‌ای داده‌ها را رسم کنید.

۵ کاشی‌های تولیدی یک کارخانه کاشی‌سازی از نظر کیفیت در سه گروه درجهٔ یک، درجهٔ ۲ و درجهٔ ۳ طبقه‌بندی می‌شوند. جدول زیر توزیع کاشی‌های تولیدی این کارخانه را در هفتهٔ گذشته نشان می‌دهد.

کیفیت	درجهٔ ۱	درجهٔ ۲	درجهٔ ۳
تعداد برحسب هزار جعبه	۲۰	۱۱	۹

یک نمودار دایره‌ای برای این داده‌ها رسم کنید.

۶ زمانهای شعله‌ور شدن نوعی از مواد پارچه‌ای که در معرض شعلهٔ آتش قرار گرفته‌اند به نزدیکترین صدم ثانیه گرد شده‌اند و به صورت زیر داده شده‌اند.

۵/۲۱	۴/۹۰	۳/۷۸	۱۱/۲۵	۶/۶۰	۲/۶۵	۶/۷۵	۵/۵۰	۴/۷۹	۲/۵۸
۱/۷۶	۳/۴۹	۳/۷۵	۳/۹۰	۸/۷۹	۷/۸۶	۵/۸۴	۵/۹۲	۶/۲۰	۲/۵۱
۹/۲۰	۶/۷۷	۳/۱۰	۵/۳۳	۵/۹۲	۴/۷۱	۸/۸۰	۴/۵۶	۱/۵۲	۴/۰۴
۱/۲۰	۵/۶۲	۶/۴۳	۸/۶۴	۹/۶۵	۶/۲۵	۷/۴۰	۲/۴۶	۱/۳۸	۶/۴۳
۶/۸۵	۹/۷۰	۱/۷۰	۷/۴۱	۵/۰۹	۹/۴۵	۴/۷۲	۶/۹۰	۳/۸۷	۱/۵۸
۲/۸۰	۵/۱۱	۶/۴۰	۷/۹۶	۴/۱۱	۱۲/۸۰	۳/۶۲	۱/۴۷	۴/۵۴	۴/۳۲
۷/۳۵	۴/۵۰	۳/۲۴	۱۰/۶۰	۶/۳۷	۱/۴۲	۲/۴۶	۲/۱۱	۵/۱۲	۲/۲۰
۱۱/۷۵	۲/۵۰	۱/۷۹	۳/۸۱	۵/۴۰	۱/۹۲	۸/۷۵	۲/۳۲	۵/۱۵	۴/۱۹

الف – یک جدول فراوانی برای این داده‌ها تشکیل دهید.

ب – هیستوگرام و چندبر فراوانی را برای این داده‌ها رسم کنید.

۷ در ۱۵۰۰ اندازه‌گیری، کوچکترین و بزرگترین عددهای به دست آمده به ترتیب ۱۰/۸ و ۱۱/۹ سانتیمتر بوده‌اند. برای تشکیل جدول فراوانی برای این داده‌ها، طول ردهٔ مناسب را پیدا کنید.

۸ در یک مرکز کامپیوتر دانشگاهی، در مدت ۶۰ روز، تعداد توقفهای ناشی از اشتباه ماشین در هر روز ثبت شده‌اند و داده‌های زیر به دست آمده‌اند

۰ ۱ ۱ ۰ ۱ ۰ ۱ ۱ ۳ ۰ ۲ ۰ ۶ ۳ ۴ ۰ ۰ ۵ ۸ ۱

۰ ۴ ۲ ۱ ۳ ۳ ۴ ۶ ۱ ۰ ۰ ۲ ۱ ۲ ۱ ۰ ۰ ۰ ۲ ۲

۲ ۲ ۰ ۴ ۴ ۲ ۰ ۲ ۰ ۱ ۱ ۰ ۰ ۰ ۰ ۰ ۲ ۱ ۳ ۰

الف – جدول فراوانی را برای این داده‌ها تشکیل دهید.

ب – نمودار میله‌ای و نمودار دایره‌ای را برای این داده‌ها رسم کنید.

ج – در چه نسبتی از روزها تعداد توقفها بیشتر از ۳ بار بوده است.

۹ شرکتی دارای ۵۰ اتومبیل است که بیمهٔ بدنه شده‌اند. تعداد مراجعات به شرکت بیمه برای دریافت خسارت این اتومبیلها در سال گذشته به صورت زیر بوده است

تعداد خسارتها	۰	۱	۲	۳	۴	۵	۶
تعداد اتومبیلها	۲۱	۱۳	۵	۴	۲	۳	۲

الف – میانگین، میانه و نمای تعداد خسارتها را محاسبه کنید.

ب – واریانس و انحراف استاندارد خسارتها را محاسبه کنید.

۱۰ فراوانی رده‌های نامساوی ۱۰/۵-۱۳/۵، ۱۳/۵-۱۷/۵، ۱۷/۵-۱۹/۵ و ۱۹/۵-۲۴/۵ به ترتیب ۵، ۷، ۲ و ۶ می‌باشد. با استفاده از رابطه

فراوانس نسبی رده = عرض مستطیل × طول رده

هیستوگرام را به گونه‌ای رسم کنید که مساحت تمام مستطیل‌های هیستوگرام یک واحد مربع شود.

۱۱ اگر میانگین یک سری داده‌های m تایی برابر $\bar{x}$ و میانگین یک سری داده‌های n تایی برابر $\bar{y}$ باشد، ثابت کنید که میانگین آمیخته این دو سری از داده‌ها برابر $\frac{m\bar{x}+n\bar{y}}{m+n}$ است.

۱۲ اگر بر روی داده‌های x_1، x_2، ... و x_k با فراوانیهای f_1، f_2، ... و f_k تبدیل $y_i=\frac{x_i-a}{b}$ $\quad i=1,\ldots,k$ را انجام دهیم، نشان دهید که رابطهٔ بین میانگین و واریانس y_iها و x_iها به صورت زیر می‌باشد

$$\bar{x}=a+b\bar{y} \quad , \quad S_x^2=b^2S_y^2$$

۱۳ قضیهٔ ۱.۱ را اثبات کنید.

۱۴ داده‌های زیر قطر ۵۰ بلبرینگ ساخته شده توسط یک کارخانه بر حسب اینچ می‌باشد

۰/۷۳۷	۰/۷۲۹	۰/۷۲۶	۰/۷۳۵	۰/۷۴۱	۰/۷۳۶	۰/۷۴۰	۰/۷۴۳	۰/۷۳۸	۰/۷۳۱
۰/۷۳۵	۰/۷۳۹	۰/۷۴۲	۰/۷۳۳	۰/۷۲۴	۰/۷۳۵	۰/۷۳۶	۰/۷۳۷	۰/۷۲۸	۰/۷۳۶
۰/۷۳۲	۰/۷۳۴	۰/۷۳۵	۰/۷۳۸	۰/۷۲۸	۰/۷۴۰	۰/۷۴۲	۰/۷۳۶	۰/۷۴۵	۰/۷۳۳
۰/۷۳۵	۰/۷۲۷	۰/۷۳۸	۰/۷۳۴	۰/۷۳۰	۰/۷۳۹	۰/۷۳۲	۰/۷۳۰	۰/۷۳۳	۰/۷۳۹
۰/۷۴۴	۰/۷۳۶	۰/۷۳۶	۰/۷۳۲	۰/۷۳۴	۰/۷۲۷	۰/۷۳۵	۰/۷۳۲	۰/۷۳۵	۰/۷۴۱

الف – یک جدول فراوانی برای این داده‌ها تشکیل دهید و هیستوگرام و چندبر فراوانی داده‌ها را رسم کنید.

ب – میانگین، میانه و نما و انحراف استاندارد داده‌ها را محاسبه کنید.

ج- چند درصد داده‌ها در فاصلهٔ $(\bar{x}-s,\bar{x}+s)$ و چند درصد داده‌ها در فاصلهٔ $(\bar{x}-2s,\bar{x}+2s)$ قرار دارند؟

۱۵ نشان دهید که ضریب تغییر به واحد اندازه‌گیری داده‌ها بستگی ندارد، یعنی اگر داده‌ها را در عدد ثابت b ضرب کنیم، ضریب تغییر ثابت می‌ماند.

۱۶ تعداد گلهای به ثمر رسیده توسط دو تیم A و B در طول یک دوره مسابقات به صورت زیر است. کدام تیم وضع بهتری دارد؟

تعداد گلها در مسابقه	۰	۱	۲	۳	۴
تعداد بازیهای تیم A	۵۴	۱۸	۱۶	۱۰	۸
تعداد بازیهای تیم B	۳۴	۱۸	۱۲	۱۰	۶

۱۷ جدول زیر جدول فراوانی مربوط به وزن تعدادی از دانش‌آموزان یک دبیرستان را نشان می‌دهد

رده‌ها	x_i	f_i	r_i	g_i	s_i
۳۰/۵-۳۵/۵			۰/۰۶		
۳۵/۵-۴۰/۵		۷			
				۱۹	
					۰/۶۴
			۰/۱۶		
				۴۶	
جمع		۵۰	۱/۰۰		

الف- جدول را کامل کنید.

ب- میانگین، میانه، نما و انحراف استاندارد را محاسبه کنید.

۱۸ عدد Q_p که $0<p<1$ را چندک مرتبهٔ p ام داده‌ها گویند هرگاه تقریباً $100p$٪ داده‌ها قبل از آن قرار گیرند. در حالت خاص $Q_1=Q_{0/25}$ و $Q_2=Q_{0/5}$ و $Q_3=Q_{0/75}$ را به ترتیب چارکهای اول و دوم و سوم داده‌ها گویند.

الف- نشان دهید که برای داده‌های گسسته $Q_p=(1-w)x_{(r)}+wx_{(r+1)}$ که در آن $x_{(1)}$ و $x_{(2)}$ ... و $x_{(n)}$ داده‌های مرتب شده به طور غیر نزولی، $r=[(n+1)p]$ و $w=(n+1)p-r$ می‌باشد.

ب- نشان دهید که برای داده‌های پیوسته $Q_p=L_p+\frac{(np-g_p)\,w}{f_p}$ که در آن L_p کران پائین رده‌ای است که فراوانی تجمعی نسبی آن بزرگتر یا مساوی p است که به آن ردهٔ Q_p گوئیم، g_p فراوانی تجمعی ردهٔ قبل از ردهٔ Q_p، f_p فراوانی ردهٔ Q_p و w طول رده می‌باشد.

۱۹ در تمرین ۴، میانه و چارک اول داده‌ها را محاسبه کنید.

۲۰ در تمرین ۶، میانگین، انحراف استاندارد، میانه و چارک سوم داده‌ها را محاسبه کنید.

۲۱ در تمرین ۸، چارک دوم و $Q_{0/9}$ را محاسبه کنید.

فصل دوم

احتمال

۱.۲ مقدمه

چنانچه در فصل اول مشاهده شد در یک بررسی آماری از جمعیت مورد نظر نمونه‌ای انتخاب می‌کنیم و به کمک روشهای آمار توصیفی به بررسی مشخصات اصلی نهفته در این نمونه می‌پردازیم. هدف اصلی در این بررسی آماری، تجزیه و تحلیل اطلاعات حاصل از نمونه به منظور به دست آوردن نتایج و انجام استنباط هایی در مورد جمعیت مورد مطالعه است. چون نمونه شامل قسمتی از اطلاعات موجود در جمعیت است، پس نمی‌توان به درستی نتایج به دست آمده از روی آن به طور قطع اطمینان داشت. بنابراین نیاز به روشهایی داریم که میزان درستی نتایج حاصل را بتوان بوسیله آنها سنجید. به کمک نظریه احتمال می‌توان به چنین روشهایی دست یافت. در این فصل مفاهیم اساسی احتمال را بررسی می‌کنیم.

۲.۲ فضای نمونه و پیشامد

در مسائل علمی اغلب اوقات با آزمایشهایی مواجه می‌شویم که اگر تحت شرایط مشابهی تکرار شوند نتایج مختلفی را به دست می‌دهند، که نشانگر تأثیر یک عامل اتفاقی در نتیجه آزمایش است. برای مثال در بررسی طول عمر لامپهای تولیدی یک کارخانه و یا بررسی میزان تأثیر یک داروی جدید روی بیماران و... با چنین آزمایشهایی مواجه می‌شویم که به آنها آزمایشهای تصادفی گویند.

تعریف ۱.۲ آزمایشی که تحت شرایط یکسان بتوان آن را تکرار کرد و نتیجه آن قبل از انجام آزمایش قابل تعیین نبوده ولی کلیه نتایج ممکن آن قابل تعیین باشد را یک **آزمایش تصادفی** گویند.

مثال ۱.۲.۲ هر کدام از موارد زیر یک آزمایش تصادفی است.

الف- پرتاب یک سکه

ب- پرتاب یک تاس

ج- پرتاب متوالی یک سکه تا مشاهده یک شیر

د- اندازه‌گیری درجه حرارت بدن یک بیمار در یک شبانه روز

هـ- اندازه‌گیری طول عمر یک لامپ

در هر یک از آزمایشهای مثال ۱.۲.۲ نتیجه آزمایش از قبل قابل تعیین نیست ولی کلیه نتایج آزمایش قابل تعیین می‌باشند. مثلاً در پرتاب یک تاس یکی از اعداد ۱، ۲، ۳، ۴، ۵ یا ۶ ظاهر می‌شود ولی قبل از انجام آزمایش نمی‌توان گفت که کدام عدد رخ می‌دهد.

تعریف ۲.۲ مجموعه تمام نتایج یک آزمایش تصادفی را **فضای نمونه** گویند و آن را با نماد S نمایش می‌دهند.

مثال ۲.۲.۲ در مثال ۱.۲.۲ فضای نمونه آزمایشهای تصادفی عبارت اند از

الف- در پرتاب یک سکه داریم $S=\{H,T\}$ که در آن H نمایانگر رخداد "شیر" و T نمایانگر رخداد "خط" است و در پرتاب دو سکه داریم $S=\{HH,HT,TH,TT\}$

ب- در پرتاب یک تاس داریم $S=\{۱,۲,۳,۴,۵,۶\}$

همینطور فضای نمونه حاصل از پرتاب دو تاس عبارت است از

$$S=\{(۱,۱),(۱,۲),\dots,(۶,۶)\}=\{(x,y)\ |\ x,y=۱,۲,\dots,۶\}$$

ج- فضای نمونه حاصل از پرتاب یک سکه تا مشاهده یک شیر عبارت است از

$$S=\{H,TH,TTH,TTTH,\dots\}$$

د- اگر درجه حرارت بدن یک بیمار در یک شبانه روز را اندازه‌گیری کنیم آنگاه فضای نمونه عبارت است از $S=[۳۵,۴۲]$ که یک فاصله بسته است.

هـ- اگر طول عمر لامپ تولیدی یک کارخانه که حداکثر ۱۰۰۰۰ ساعت طول عمر دارد را اندازه‌گیری کنیم آنگاه فضای نمونه فاصله $S=[۰,۱۰۰۰۰]$ است.

آزمایش تصادفی ← فضای نمونه ← پیشامد (تعاریف)

مثال ۳.۲.۲ از خط تولید یک کارخانه ۳ محصول را به طور تصادفی انتخاب می‌کنیم. این محصولات ممکن است خراب یا سالم باشند.

الف- اگر خراب بودن محصول را با D و سالم بودن آن را با N نمایش دهیم، آنگاه فضای نمونه موردنظر عبارت است از

$$S_۱=\{NNN,NND,NDN,DNN,DDN,DND,NDD,DDD\}$$

ب- اگر به تعداد قطعات خراب در بین ۳ قطعه انتخابی توجه کنیم آنگاه فضای نمونه عبارت است از $S_۲=\{۰,۱,۲,۳\}$

مثال ۳.۲.۲ نشان می‌دهد که در یک آزمایش تصادفی، ممکن است بیش از یک فضای نمونه داشته باشیم و جنبه موردنظر از آزمایش تصادفی است که فضای نمونه را تعیین می‌کند. همچنین با توجه به مثالهای بالا می‌توان فضاهای نمونه را به طور کلی به دو گروه زیر تقسیم نمود

۱- فضای نمونه گسسته که شامل دو حالت زیر است

الف- **فضای نمونه متناهی** که تعداد اعضای آن متناهی است، مانند فضای نمونه در مثال ۲.۲.۲(الف) و (ب).

ب- **فضای نمونه نامتناهی شمارش‌پذیر** که یک مجموعه نامتناهی اما شمارش‌پذیر است، مانند مثال ۲.۲.۲(ج).

۲- فضای نمونه پیوسته که اعضای آن به صورت یک فاصله از اعداد حقیقی یا یک سطح در فضای دو بعدی یا... است، مانند مثال ۲.۲.۲ (د) و (هـ).

در ادامه نخست بیشتر در مورد فضای نمونه متناهی بحث می‌کنیم و سپس در مورد فضاهای نمونه دیگر در طول فصل بحث می‌کنیم.

تعریف ۳.۲ در یک فضای نمونه متناهی، هر زیر مجموعه از فضای نمونه را یک **پیشامد** می‌نامند. پیشامدی که تنها دارای یک عضو باشد به پیشامد ساده موسوم است و پیشامدی با تعداد اعضای بیشتر از یک عضو را پیشامد مرکب گوئیم. اگر پیشامدی دارای هیچ عضوی نباشد آن را پیشامد محال یا تهی می‌نامیم و پیشامدی که برابر فضای نمونه S باشد به پیشامد حتمی موسوم است.

مثال ۴.۲.۲ اگر یک جفت تاس را یک بار پرتاب کنیم، آنگاه فضای نمونه آن عبارت است از

$$S=\{(x,y)\ |\ x,y=۱,۲,\dots,۶\}$$

الف- اگر E_1 پیشامد این باشد که مجموع اعداد روی دو تاس کمتر از ۳ باشد آنگاه E_1 یک پیشامد ساده است و $E_1=\{(۱,۱)\}$

ب- اگر E_2 پیشامد این باشد که مجموع اعداد روی دو تاس بیش از ۱۰ باشد آنگاه E_2 یک پیشامد مرکب است و $E_2=\{(۵,۶),(۶,۵),(۶,۶)\}$

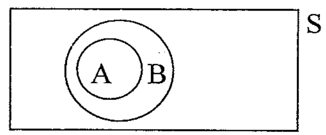
ج- اگر E_3 پیشامد این باشد که مجموع اعداد روی دو تاس حداقل ۱۳ باشد آنگاه E_3 یک پیشامد محال است و $E_3=\emptyset$

د- اگر E_4 پیشامد این باشد که مجموع اعداد روی دو تاس حداکثر ۱۲ باشد آنگاه E_4 یک پیشامد حتمی است و $E_4=S$

وقوع یک پیشامد گوئیم پیشامد A به وقوع پیوسته است هر گاه نتیجه آزمایش تصادفی منجر به مشاهده عضوی از پیشامد A گردد. برای مثال در مثال ۴.۲.۲ اگر در پرتاب دو تاس نتیجه (۶,۵) را مشاهده کنیم آنگاه گوئیم پیشامد E_2 به وقوع پیوسته و پیشامد E_1 رخ نداده است.

۱.۲.۲ اعمال روی پیشامدها

چون پیشامدها زیر مجموعه‌ای از فضای نمونه هستند پس می‌توان همانند مجموعه‌ها اعمال جبری را روی آنها انجام داد. در این حالت فضای نمونه مجموعهٔ مرجع می‌باشد و توسط نمودار ون می‌توان پیشامدها و فضای نمونه را به صورت شکل ۱.۲ نمایش داد. بعضی از اعمال روی پیشامدها عبارت‌اند از

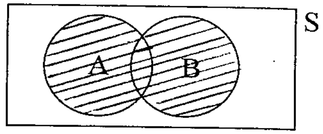


شکل ۱.۲ $A\subset B$

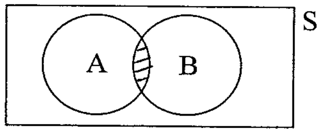
الف- زیر پیشامد پیشامد A را زیر پیشامد، پیشامد B گوئیم هر گاه وقوع A، وقوع B را نتیجه دهد (شکل ۱.۲) و آنرا با نماد $A\subset B$ نمایش می‌دهیم.

ب- دو پیشامد مساوی دو پیشامد A و B را مساوی گوئیم هر گاه وقوع یکی وقوع دیگری را نتیجه دهد، یعنی $A=B \Leftrightarrow (A\subset B\ ,\ B\subset A)$

پ- اجتماع دو پیشامد پیشامد $A\cup B=\{x \mid x\in A$ یا $x\in B\}$ را اجتماع دو پیشامد A و B گوئیم و وقوع $A\cup B$ به معنای وقوع حداقل یکی از دو پیشامد A یا B است (شکل ۲.۲).



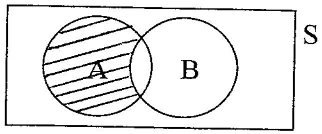
شکل ۲.۲ پیشامد $A\cup B$



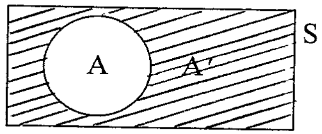
شکل ۳.۲ پیشامد $A\cap B$

ت- اشتراک دو پیشامد پیشامد $A\cap B=\{x \mid x\in A$ و $x\in B\}$ را اشتراک دو پیشامد A و B گوئیم و وقوع $A\cap B$ به معنای وقوع همزمان هر دو پیشامد A و B است (شکل ۳.۲).

ث- تفاضل دو پیشامد پیشامد $A-B=\{x \mid x\in A\ ,\ x\notin B\}$ را تفاضل پیشامد B از A گوئیم و وقوع $A-B$ به معنای وقوع «فقط A و نه B» است (شکل ۴.۲).



شکل ۴.۲ پیشامد $A-B$



شکل ۵.۲ پیشامد A'

ج- متمم یک پیشامد پیشامد $A'=\{x \mid x\in S\ ,\ x\notin A\}$ را متمم پیشامد A گوئیم و وقوع A' به معنای عدم وقوع پیشامد A است (شکل ۵.۲).

اشتراک و اجتماع بیش از دو پیشامد نیز به نحو مشابهی تعریف می‌گردد. اجتماع پیشامدهای A_1، A_2، ... و A_m شامل اعضایی است که حداقل به یکی از A_iها متعلق باشد و اشتراک آنها شامل اعضایی است که در همه پیشامدهای A_1، A_2، ... و A_m باشند. یعنی

$$\bigcap_{i=1}^{m} A_i = A_1\cap A_2\cap\ldots\cap A_m \qquad \bigcup_{i=1}^{m} A_i = A_1\cup A_2\cup\ldots\cup A_m$$

تعریف ۴.۲ دو پیشامد A و B را ناسازگار (جدا) گوئیم هر گاه $A\cap B=\emptyset$، یعنی دو پیشامد را ناسازگار گوئیم هر گاه هر دو نتوانند همزمان اتفاق بیفتند.

مثال ۵.۲.۲ در پرتاب یک تاس اگر A پیشامد مشاهده عدد زوج و B پیشامد مشاهده عدد فرد باشد آنگاه

$$A=\{۲,۴,۶\}\ ,\ B=\{۱,۳,۵\}\ ,\ A\cap B=\emptyset$$

پیشامد ناسازگار

بنابراین A و B ناسازگار هستند.

تعریف ۵.۲ پیشامدهای A_1، A_2، A_3 و... را دو به دو ناسازگار گوئیم هرگاه به ازای هر $i\neq j$

$$A_i \cap A_j = \emptyset$$

کاربرد: موارد عملی (کار)

۳.۲ احتمال

احتمال وقوع یک پیشامد به معنای شانس وقوع آن پیشامد در انجام یک آزمایش تصادفی است. برای محاسبه احتمال تعبیرهای مختلفی از جمله فراوانی نسبی، هم شانسی و شخصی وجود دارد که هر کدام از این تعبیرها می‌تواند برای بکارگیری احتمال در مسایل عملی مفید باشند ولی به هر کدام انتقادهایی وارد است. در زیر ابتدا روش محاسبه احتمال به طریق فراوانی نسبی را می‌آوریم و سپس تعریف ریاضی احتمال که بر اساس اصول موضوع احتمالات قرار دارد را می‌آوریم.

فرض کنید یک آزمایش تصادفی را n مرتبه تحت شرایط یکسان تکرار کنیم و تعداد دفعاتی که در این n آزمایش پیشامد A بوقوع پیوسته را با $f_n(A)$ نمایش دهیم. بنابراین $r_n(A)=\frac{f_n(A)}{n}$ فراوانی نسبی وقوع پیشامد A می‌باشد و انتظار داریم که با زیاد شدن تعداد آزمایشات n، $r_n(A)$ به یک عدد ثابتی نزدیک شود که این عدد ثابت را احتمال وقوع پیشامد A گوئیم و آن را با $P(A)$ نمایش می‌دهیم یعنی

$$P(A)=\lim_{n\to+\infty} r_n(A)$$

اما انجام آزمایش تصادفی به تعداد دفعات زیاد ممکن است عملی نباشد و یا ممکن است $r_n(A)$ به یک عدد ثابت نزدیک نشود. با این وجود در بیشتر موارد عملی از این تعبیر احتمال برای محاسبه احتمال استفاده می‌شود و این تعبیر با اصول موضوع احتمال که در زیر ارائه می‌شوند، نیز سازگاری دارد. برای مثال اگر سکه‌ای را ۱۰۰۰ مرتبه پرتاب کنیم و مشاهده کنیم که ۴۹۵ مرتبه شیر و ۵۰۵ مرتبه خط مشاهده شده است در این صورت فراوانی نسبی مشاهده شیر و خط به ترتیب ۰/۴۹۵ و ۰/۵۰۵ می‌باشد و می‌توان احتمال وقوع هر کدام از این دو پیشامد را $0/5=\frac{1}{2}$ در نظر گرفت.

تعریف ریاضی تابع احتمال تابع احتمال تابعی از فضای نمونه S به داخل مجموعه اعداد حقیقی R به صورت $P:S\to R$ است. به عبارت دقیقتر احتمال تابعی مانند P است که به هر پیشامد A از فضای نمونه S عدد حقیقی $P(A)$ را به گونه‌ای نسبت می‌دهد که در ۳ اصل موضوع زیر صدق

کند

۱- $P(S)=1$

۲- برای هر پیشامد A در S، $P(A)\geq 0$

۳- اگر A_1، A_2، A_3 و... پیشامدهای دو به دو ناسازگار باشند آنگاه

$$P(A_1\cup A_2\cup A_3\cup\dots)=P(A_1)+P(A_2)+P(A_3)+\dots$$

در ادامه تابع احتمال را روی هر یک از فضاهای نمونه گسسته و پیوسته به دست می‌آوریم.

۱.۳.۲ مدل احتمال روی فضای نمونه متناهی

فرض کنید $S=\{e_1,e_2,\dots,e_n\}$ یک فضای نمونه متناهی غیرتهی باشد. یک مدل احتمال روی این فضای نمونه عبارت است از نسبت دادن اوزان (احتمالات) نامنفی p_1، p_2،... و p_n به نقاط فضای نمونه S به طوری که مجموع تمام این اعداد برابر یک شود یعنی

S	e_1	e_2	...	e_n
احتمال	p_1	p_2	...	p_n

$$0\leq p_i\leq 1\ ,\ i=1,2,\dots,n \quad , \quad \sum_{i=1}^{n} p_i=1$$

این اعداد متضمن ارزیابی وقوع پیشامدهای ساده یک آزمایش تصادفی می‌باشند و بایستی به گونه‌ای نسبت داده شوند که پیشامد ساده‌ای که شانس وقوع آن کمتر است عدد نسبت داده شده به صفر نزدیکتر و پیشامد ساده‌ای که شانس وقوع آن بیشتر است عدد نسبت داده شده به یک نزدیکتر باشد. اگر در یک فضای نمونه پیشامدهای ساده شانس یکسان برای اتفاق افتادن داشته باشند در این صورت بایستی اعداد (احتمالات) یکسان به این نقاط نسبت داده شود. برای مثال در پرتاب یک تاس مدل احتمال برابر است با

S	۱	۲	۳	۴	۵	۶
احتمال	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

حال اگر A یک پیشامد در فضای نمونه S باشد، احتمال پیشامد A برابر مجموع تمام احتمالات نسبت داده شده به پیشامدهای ساده تشکیل دهنده A در نظر گرفته می‌شود، یعنی

$$A=\{e_1,e_2,\dots,e_m\}\subset S \Rightarrow P(A)=p_1+p_2+\dots+p_m=\sum_{i=1}^{m} p_i$$

به راحتی می‌توان نشان داد که تعریف فوق در ۳ اصل احتمال صدق می‌کند.

مثال ۱.۳.۲ سکه‌ای را دو بار پرتاب می‌کنیم احتمال آوردن حداقل یک شیر را بیابید.

حل فضای نمونه حاصل از این آزمایش برابر است با $S=\{TT,TH,HT,HH\}$

اگر سکه سالم باشد آنگاه شانس رخداد شیر یا خط با هم برابر است و در نتیجه به هر کدام از نقاط فضای نمونه شانس یکسان w را نسبت می‌دهیم و بنابراین ۴w=۱ و یا $w=\frac{1}{4}$. پس مدل احتمال برای این آزمایش عبارت است از

S	TT	TH	HT	HH
احتمال	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

حال اگر A پیشامد مشاهده حداقل یک شیر باشد آنگاه $A=\{TH,HT,HH\}$ و در نتیجه

$$P(A)=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=\frac{3}{4}$$

مثال ۲.۳.۲ یک تاس به شکلی است که احتمال آوردن عدد زوج در آن دو برابر عدد فرد می‌باشد. احتمال آوردن عدد بیشتر از ۳ در پرتاب این تاس را بیابید.

حل فضای نمونه حاصل از این آزمایش برابر است با $S=\{1,2,3,4,5,6\}$

اگر به اعداد فرد شانس w و به اعداد زوج شانس 2w را نسبت دهیم، چون بایستی جمع احتمالات برابر یک شود پس بایستی $w=\frac{1}{9}$ باشد و بنابراین مدل احتمال* برای این آزمایش تصادفی عبارت است از

S	۱	۲	۳	۴	۵	۶
احتمال	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{9}$

حال اگر B پیشامد مشاهده عدد بیشتر از ۳ باشد آنگاه $B=\{4,5,6\}$ و در نتیجه

$$P(B)=\frac{2}{9}+\frac{1}{9}+\frac{2}{9}=\frac{5}{9}$$

برای حالتی که شانس وقوع برای همه یکسان باشد

مدل احتمال یکنواخت در مثالهای بالا ضرایب وزنی w در حکم احتمال پیشامدهای ساده می‌باشند. با مقایسه دو مثال ۱.۳.۲ و ۲.۳.۲ مشاهده می‌شود که اگر نقاط فضای نمونه $S=\{e_1,e_2,\dots,e_n\}$ دارای شانس مساوی برای انتخاب شدن باشند آنگاه احتمال وقوع هر پیشامد $A=\{e'_1,e'_2,\dots,e'_m\}$ در S عبارت است از

برای حالتی که شانس وقوع برای همه یکسان باشد

$$P(A)=\frac{m}{n}=\frac{n(A)}{n(S)}=\frac{\text{تعداد اعضای پیشامد A}}{\text{تعداد اعضای فضای نمونه S}}=\frac{\text{تعداد حالات مطلوب}}{\text{تعداد کل حالات}}$$

این مدل احتمال را مدل احتمال یکنواخت گویند.

مثال ۳.۳.۲ یک جفت تاس را پرتاب می‌کنیم. احتمال آوردن مجموع هفت را به دست آورید.

حل در این آزمایش فضای نمونه دارای $n(S)=6^2=36$ عضو است که همگی دارای شانس یکسان برای به وقوع پیوستن هستند. حال اگر A پیشامد مشاهده مجموع هفت باشد آنگاه

$$A=\{(1,6),(2,5),\dots,(6,1)\}$$

و $n(A)=6$، بنابراین

$$P(A)=\frac{n(A)}{n(S)}=\frac{6}{36}=\frac{1}{6}$$

مثال ۴.۳.۲ جعبه‌ای شامل ۳ توپ سفید، ۴ توپ سیاه و ۵ توپ قرمز است. یک توپ به تصادف از این جعبه خارج می‌کنیم

الف- احتمال اینکه توپ انتخابی قرمز باشد را بیابید.

ب- احتمال اینکه توپ انتخابی سفید باشد را بیابید.

حل در این آزمایش فضای نمونه دارای $n(S)=12$ عضو می‌باشد که شانس انتخاب هر توپ با یکدیگر مساوی است.

الف- اگر R پیشامد مشاهده توپ قرمز باشد آنگاه $n(R)=5$ و $P(R)=\frac{n(R)}{n(S)}=\frac{5}{12}$

ب- اگر W پیشامد مشاهده توپ سفید باشد آنگاه $n(W)=3$ و $P(W)=\frac{n(W)}{n(S)}=\frac{3}{12}=\frac{1}{4}$

تذکر توجه کنید که اگر نتوان احتمالات مساوی را به نقاط فضای نمونه نسبت داد، بایستی از طریق تجربه و آزمایش ضرایب وزنی w را به نقاط فضای نمونه نسبت داده و مدل احتمال را تعیین کنیم.

۴.۲ چند قانون احتمال

با استفاده از ۳ اصل احتمال می‌توان نتایج زیر را که برای محاسبه احتمالات مفید می‌باشند، به دست آورد.

قضیه ۱.۲ $P(\emptyset)=0$

اثبات با قرار دادن $A_۱=S$ و $A_۲=A_۳=\dots=\emptyset$ در اصل ۳ احتمال قضیه به سادگی اثبات می‌شود.

قضیه ۲.۲ اگر $B_۱$، $B_۲$،...و B_n پیشامدهای دو به دو ناسازگار باشند آنگاه

$$P(B_۱\cup B_۲\cup\dots\cup B_n)=P(B_۱)+P(B_۲)+\dots+P(B_n)$$

اثبات با قرار دادن $A_۱=B_۱$، $A_۲=B_۲$، ...، $A_n=B_n$ و $A_i=\emptyset\quad \forall i>n$ در اصل ۳ احتمال قضیه به سادگی اثبات می‌شود.

نتیجه ۱.۲ اگر A و B دو پیشامد ناسازگار باشد آنگاه

$$\boxed{P(A\cup B)=P(A)+P(B)} \qquad (۱.۲)$$

مثال ۱.۴.۲ یک جفت تاس را پرتاب می‌کنیم احتمال آوردن مجموع هفت یا مجموع بیش از ده را بیابید.

حل در این آزمایش $n(S)=۶^۲=۳۶$ و اگر A پیشامد مشاهده مجموع هفت باشد آنگاه $n(A)=۶$ و $P(A)=\frac{۶}{۳۶}$ و اگر B پیشامد مشاهده مجموع بیش از ده باشد آنگاه $n(B)=۳$ و $P(B)=\frac{۳}{۳۶}$. چون $A\cap B=\emptyset$ پس A و B ناسازگار هستند و بنابراین

$$P(A\cup B)=P(A)+P(B)=\frac{۶}{۳۶}+\frac{۳}{۳۶}=\frac{۱}{۴}$$

قضیه ۳.۲ اگر A یک پیشامد و A' متمم آن باشد ($A\cup A'=S$) آنگاه

$$\boxed{P(A)+P(A')=۱ \quad \text{یا} \quad P(A')=۱-P(A)} \qquad (۲.۲)$$

اثبات چون $A\cup A'=S$ و $A\cap A'=\emptyset$ پس با قرار دادن $B=A'$ در نتیجه ۱.۲ قضیه اثبات می‌شود.

مثال ۲.۴.۲ اگر سکه‌ای را ۶ بار پرتاب کنیم آنگاه احتمال آوردن حداقل یک شیر را بیابید.

حل در این آزمایش $n(S)=۲^۶=۶۴$ و اگر A پیشامد مشاهده حداقل یک شیر باشد آنگاه A' پیشامد مشاهده هیچ شیر است یعنی $A'=\{TTTTTT\}$ و $n(A')=۱$ در نتیجه $P(A')=\frac{۱}{۶۴}$ و بنابراین

$$P(A)=۱-P(A')=۱-\frac{۱}{۶۴}=\frac{۶۳}{۶۴}$$

با استفاده از نتیجه ۱.۲ و اصول احتمال می‌توان نتایج زیر را به سادگی اثبات کرد. اثبات این نتایج را به خواننده واگذار می‌کنیم.

قضیه ۴.۲ اگر A و B دو پیشامد دلخواه باشند آنگاه

$$\boxed{P(B-A)=P(B)-P(A\cap B)} \qquad (۳.۲)$$

نتیجه ۲.۲ اگر A و B دو پیشامد باشند که $A\subset B$ آنگاه

الف- $P(B-A)=P(B)-P(A)$

ب- $P(A)\le P(B)$

نتیجه ۳.۲ برای هر پیشامد A داریم که $۰\le P(A)\le ۱$

قضیه ۵.۲ اگر A و B دو پیشامد دلخواه باشند آنگاه

$$\boxed{P(A\cup B)=P(A)+P(B)-P(A\cap B)} \qquad (۴.۲)$$

مثال ۳.۴.۲ در یک زندان معین معلوم شده است که $\frac{۲}{۳}$ از زندانیها دارای سن کمتر از ۲۵ سال و $\frac{۳}{۵}$ از زندانیها مرد و $\frac{۵}{۸}$ از زندانیها زن یا دارای سن حداقل ۲۵ سال می‌باشند. احتمال اینکه یک زندانی که به طور تصادفی انتخاب شده است زنی با حداقل سن ۲۵ سال باشد را بیابید.

حل اگر M پیشامد این باشد که زندانی انتخاب شده مرد باشد و A پیشامد این باشد که زندانی انتخاب شده دارای سن کمتر از ۲۵ سال باشد آنگاه

$$P(A)=\frac{۲}{۳} \quad , \quad P(A')=۱-\frac{۲}{۳}=\frac{۱}{۳}$$
$$P(M)=\frac{۳}{۵} \quad , \quad P(M')=۱-\frac{۳}{۵}=\frac{۲}{۵}$$
$$P(M'\cup A')=\frac{۵}{۸} \quad , \quad P(M'\cap A')=?$$

بنابراین با استفاده از فرمول (۴.۲) داریم که

$$P(M'\cap A')=P(M')+P(A')-P(M'\cup A')=\frac{۲}{۵}+\frac{۱}{۳}-\frac{۵}{۸}=\frac{۱۳}{۱۲۰}$$

مثال ۴.۴.۲ احتمال آنکه یک هواپیمای جدید، تایید طراحی را به دست آورد برابر با ۰/۱۶، کارآیی استفاده از مواد را کسب کند برابر با ۰/۲۴ و هر دو را کسب کند ۰/۱۱ است.

الف- احتمال آنکه لااقل یکی از دو تأیید را به دست آورد را بیابید.

ب- احتمال آنکه فقط یکی از دو تأیید را به دست آورد را بیابید.

حل اگر A پیشامد تأیید طراحی و B پیشامد کارآیی استفاده از مواد باشند آنگاه

$$P(A)=۰/۱۶ \quad , \quad P(B)=۰/۲۴ \quad , \quad P(A\cap B)=۰/۱۱$$

الف- $P(A\cup B)=P(A)+P(B)-P(A\cap B)=۰/۱۶+۰/۲۴-۰/۱۱=۰/۲۹$

ب- پیشامد مورد نظر $(A-B)\cup(B-A)=(A\cup B)-(A\cap B)$ می‌باشد. بنابراین با توجه به نتیجه ۲.۲(الف) داریم که

$$P((A\cup B)-(A\cap B)) = ۰/۲۹ - ۰/۱۱ = ۰/۱۸$$

۵.۲ قواعد شمارش

چنانچه در بخش قبل مشاهده شد برای محاسبه احتمال وقوع یک پیشامد نیاز به محاسبه تعداد اعضای آن و تعداد اعضای فضای نمونه داریم. اغلب اوقات شمارش اعضای یک مجموعه کار دشواری است و برای انجام این کار نیاز به شناسایی برخی از اصول و قوانین داریم که در ذیل به معرفی آنها می‌پردازیم.

اصل ضرب فرض کنید که یک کار را بتوان با دو عمل پیاپی A و B انجام داد. اگر عمل A به m طریق و به دنبال آن عمل B بتواند به n طریق انجام پذیرد آنگاه این کار بـه mn طـریق انـجام می‌پذیرد.

مثال ۱.۵.۲ با ارقام ۰، ۱، ۲ و ۳ چند عدد دو رقمی می‌توان نوشت در صورتی که

الف- تکرار ارقام مجاز باشد. ب- تکرار ارقام مجاز نباشد.

حل نوشتن یک عدد دو رقمی شامل دو عمل، انتخاب رقم دهگان (A) و انتخاب رقم یکان (B) می‌باشد. بنابراین

الف- رقم دهگان می‌تواند یکی از ارقام ۱، ۲ یا ۳ به ۳ طریق و رقم یکان یکی از ارقام ۰، ۱، ۲ یا ۳ به ۴ طریق باشد، بنابراین

$$\begin{array}{ccccc} A & & B & & \\ \boxed{۳} & \times & \boxed{۴} & = & ۱۲ \end{array}$$

ب- رقم دهگان می‌تواند یکی از ارقام ۱، ۲ یا ۳ به ۳ طریق و رقم یکان می‌تواند رقم ۰ یا یکی از دو رقم باقی مانده از ارقام ۱، ۲ یا ۳ باشد، بنابراین

$$\begin{array}{ccccc} A & & B & & \\ \boxed{۳} & \times & \boxed{۳} & = & ۹ \end{array}$$

اصل جمع فرض کنید یک کار را بتوان با دو عمل A یا B انجام داد. اگر عمل A به m طریق و عمل B به n طریق انجام پذیرند و این دو عمل نتوانند همزمان اتفاق بیفتند آنگاه این کار به m+n طریق انجام می‌پذیرد.

مثال ۲.۵.۲ به چند طریق می‌توان از بین ۴ دانشجوی کـامپیوتر و ۵ دانشـجوی ریـاضی، دو دانشجو را انتخاب کرد به طوری که نفر اول به عنوان سرگروه و نفر دوم به عنوان دستیار باشد و هر دو نفر از یک رشته باشند.

حل طبق اصل ضرب دو دانشجو از رشته کامپیوتر به ۱۲=۳×۴ طریق یا از رشته ریـاضی بـه ۲۰=۴×۵ طریق انتخاب می‌شوند. بنابراین طبق اصل جمع این دو نفر را می‌توان به ۳۲=۲۰+۱۲ طریق انتخاب کرد.

اصول جمع و ضرب را می‌توان برای بیش از دو عمل، مثلاً k عمل A_1، A_2، ...، و A_k نیز گسترش داد.

مثال ۳.۵.۲ یک تاس سالم را ۴ بار به طور مستقل پرتاب می‌کنیم. احتمال اینکه در هیچکدام از پرتابها عدد مضرب ۳ مشاهده نشود را بیابید؟

حل در این آزمایش $n(S)$ برابـر تـعداد حـالات مـمکن پـرتاب ۴ بـار یک تـاس است، پس $n(S)=۶\times۶\times۶\times۶=۶^۴$ و اگر A پیشامد این باشد که عدد مضرب ۳ مشاهده نشود آنگاه در هر پرتاب ۴ حالت ۱، ۲، ۴ و ۵ مورد نظر ما می‌باشد، بنابراین $n(A)=۴\times۴\times۴\times۴=۴^۴$ و در نتیجه

$$P(A)=\frac{n(A)}{n(S)}=\frac{۴^۴}{۶^۴}=\frac{۱۶}{۸۱}$$

مثال ۴.۵.۲ تحت هر یک از شرایط زیر تعداد اعداد ۳ رقمی بدون تکرار ارقام که با ارقام ۰، ۱، ۲، ۳، ۴ و ۵ می‌توان ساخت را به دست آورید.

الف- اعداد فرد باشند.

ب- اعداد بزرگتر از عدد ۳۳۰ باشند.

پ- اعداد زوج باشند.

حل الف- در این حالت ابتدا رقم یکان به ۳ طریق (ارقام ۱، ۳ و ۵) انتخاب می‌شود. سپس رقم صدگان که به غیر از صفر و رقم یکان انتخاب شده می‌باشد به ۴ طریق انتخاب شده و در آخر رقم دهگان نیز به ۴ طریق انتخاب می‌شود و بنابراین تعداد طریق انتخاب ۴۸=۳×۴×۴ است.

ب- این اعداد به دو صورت می‌باشند. اعدادی که رقم صدگان آنها ۴ یا ۵ است که تعداد طـریق انتخاب آنها ۴۰=۴×۵×۲ است، یا اعدادی که رقم صدگان آنها ۳ است که تعداد طریق انتخاب آنها

۸=۴×۲×۱ است. پس طبق اصل جمع تعداد طریق‌های انتخاب برابر ۴۸=۸+۴۰ است.

پ- چون تعداد اعداد زوج متمم تعداد اعداد فرد است و تعداد کل اعداد بدون تکرار ارقام ۱۰۰=۴×۵×۵ است پس تعداد اعداد زوج ۵۲=۴۸-۱۰۰ است. آیا می‌توانید تعداد اعداد زوج را به طور مستقیم محاسبه کنید؟

جایگشتها در بعضی از مسایل می‌خواهیم تعداد طریق قرار گرفتن ۶ نفر در یک صف و یا تعداد طریق انتخاب ۲ نفر از بین ۶ نفر را به دست آوریم. برای این منظور می‌توان از اصول شمارش و مفهوم جایگشت استفاده کنیم.

تعریف ۶.۲ ترتیبی را که می‌توان اشیاء یک مجموعه را در کنار یکدیگر قرار داد یک جایگشت گویند.

مثال ۵.۵.۲ جایگشتها و تعداد جایگشتهای مختلف سه حرف a، b و c را به دست آورید.

حل جایگشتهای مختلف این ۳ حرف عبارت‌اند از abc-acb-bac-bca-cab-cba. برای به دست آوردن تعداد جایگشتها، این ۳ حرف را می‌خواهیم در ۳ مکان قرار دهیم. مکان اول به ۳ طریق و مکان دوم به دو طریق و مکان سوم به یک طریق می‌توانند اشغال شوند بنابراین تعداد کل جایگشتها برابر ۳!=۱×۲×۳ می‌باشد.

در حالت کلی داریم که

> اگر n عنصر متمایز را بخواهیم در یک صف کنار یکدیگر قرار دهیم تعداد جایگشتهای مختلف این عناصر برابر است با $n(n-1)(n-2)\dots(2)(1)=n!$

مثال ۶.۵.۲ چهار پزشک و پنج مهندس می‌خواهند در یک صف کنار یکدیگر قرار گیرند.

الف- احتمال اینکه مهندس‌ها در یک طرف صف و پزشک‌ها در طرف دیگر صف قرار گیرند را بیابید.

ب- احتمال اینکه پزشک‌ها و مهندس‌ها یک در میان در صف قرار گیرند را بیابید.

ج- احتمال اینکه دو مهندس بخصوص همواره کنار یکدیگر قرار گیرند را بیابید.

د- احتمال اینکه دو مهندس بخصوص هیچگاه کنار یکدیگر قرار نگیرند را بیابید.

حل در این حالت $n(S)=۹!$ یعنی تعداد کل حالات قرار گرفتن این ۹ نفر در صف می‌باشد.

الف- اگر A پیشامد قرار گرفتن پزشک‌ها در یک طرف صف (به !۴ طریق) و مهندس‌ها در طرف

دیگر صف (به !۵ طریق) باشد، چون شروع صف می‌تواند با پزشک‌ها یا مهندس‌ها باشد پس $n(A)=۲!۴!۵!$ و در نتیجه $P(A)=\frac{n(A)}{n(S)}=\frac{۲!۴!۵!}{۹!}$.

ب- اگر B پیشامد قرار گرفتن پزشک‌ها و مهندس‌ها یک در میان در صف باشد آنگاه مهندس‌ها به !۵ طریق در صف قرار می‌گیرند و پزشک‌ها به !۴ طریق در بین مهندس‌ها قرار می‌گیرند پس $P(B)=۵!۴!$ و در نتیجه $P(B)=\frac{n(B)}{n(S)}=\frac{۴!۵!}{۹!}$.

ج- اگر C پیشامد قرار گرفتن دو مهندس بخصوص در کنار یکدیگر باشد، این دو مهندس به !۲ طریق کنار یکدیگر قرار می‌گیرند و با در نظر گرفتن این دو مهندس به عنوان یک عنصر، در کل ۸ عنصر داریم که می‌توانند در صف قرار گیرند پس $n(C)=۸!۲!$ و در نتیجه

$$P(C)=\frac{n(C)}{n(S)}=\frac{۸!۲!}{۹!}=\frac{۲}{۹}.$$

د- پیشامد مورد نظر متمم C یعنی C' است، در نتیجه $P(C')=۱-P(C)=۱-\frac{۲}{۹}=\frac{۷}{۹}$

مثال ۷.۵.۲ به چند طریق می‌توان با حروف a,b,c,d,e کلمات دو حرفی ساخت در صورتی که

الف- تکرار حروف مجاز نباشد. ب- تکرار حروف مجاز باشد.

حل الف- حالتهای مختلف ساختن کلمات دو حرفی بدون تکرار در شکل ۶.۲ مشاهده می‌گردد. با استفاده از اصل ضرب این تعداد به صورت زیر محاسبه می‌شود

ab	ac	ad	ae
ba	bc	bd	be
ca	cb	cd	ce
da	db	dc	de
ea	eb	ec	ed

$$۵\times۴=\frac{(۵\times۴)\times(۳\times۲\times۱)}{(۳\times۲\times۱)}=\frac{۵!}{۳!}=\frac{۵!}{(۵-۲)!}$$

شکل ۶.۲ تبدیلات ۲ عنصر از ۵ عنصر

این مقدار را تبدیل ۲ از ۵ گوئیم و با نماد $P_۲^۵$ نمایش می‌دهیم.

ب- با استفاده از اصل ضرب این تعداد برابر $۵^۲=۵\times۵$ است.

در حالت کلی داریم که

اگر از بین n عنصر متمایز بخواهیم r عنصر را انتخاب کرده و در یک صف قرار دهیم در این صورت

الف- اگر تکرار عناصر مجاز نباشد آنگاه تعداد راه‌های، ممکن برابر است با

$$P_r^n = \frac{n!}{(n-r)!}$$

ب- اگر تکرار عناصر مجاز باشد آنگاه تعداد راه‌های ممکن برابر است با n^r

به P_r^n تبدیل r از n گویند و بایستی همواره $0 \le r \le n$ باشد. چون ۱=!۰ پس $P_n^n = n!$ است.

مثال ۸.۵.۲ کلمه COMPUTER را در نظر بگیرید.

الف- تعداد کلمات ۸ حرفی بدون تکرار حروف که از حروف این کلمه می‌توان ساخت برابر است با $P_8^8 = 8!$.

ب- تعداد کلمات ۵ حرفی بدون تکرار حروف که از حروف این کلمه می‌توان ساخت برابر است با $P_5^8 = \frac{8!}{(8-5)!} = \frac{8!}{3!}$

پ- تعداد کلمات ۵ حرفی با مجاز بودن تکرار حروف که از حروف این کلمه می‌توان ساخت برابر است با 8^5.

مثال ۹.۵.۲ تعداد جایگشتها مختلف حروف کلمه BALL را به دست آورید.

حل در اینجا با عوض کردن جای دو حرف L در کلمه تغییری ایجاد نمی‌شود و جایگشتهای مختلف این حروف عبارت‌اند از

BALL - BLAL - BLLA - LBLA - LBAL - LLBA

ABLL - ALBL - ALLB - LALB - LABL - LLAB

که تعداد آنها ۱۲ است. اگر این ۴ حرف متمایز می‌بودند آنگاه تعداد جایگشتها ۲۴=!۴ می‌شد اما در هر یک از جایگشتهای بالا اگر جای دو حرف L را که به !۲ انجام می‌پذیرد، عوض کنیم تغییری در کلمات بوجود نمی‌آید و بنابراین تعداد جایگشتهای حاصل برابر $12 = \frac{4!}{2!}$ است.

مثال ۱۰.۵.۲ تعداد جایگشتهای مختلف حروف کلمه PEPPER را به دست آورید.

حل در اینجا ۳ حرف P یکسان و ۲ حرف E یکسان و ۱ حرف R داریم. بنابراین تعداد جایگشتهای مختلف حروف این کلمه برابر $60 = \frac{6\times5\times4}{2} = \frac{6!}{3!2!1!}$ است.

در حالت کلی داریم که

اگر n عنصر وجود داشته باشند که n_1 تای آنها از نوع اول و n_2 تای آنها از نوع دوم و ... و n_r تای آنها از نوع rام باشند که $n_1+n_2+\dots+n_r=n$ آنگاه تعداد جایگشتهای این عناصر برابر است با

$$\binom{n}{n_1,n_2,\dots,n_r} = \frac{n!}{n_1!n_2!\dots n_r!}$$

مثال ۱۱.۵.۲ می‌خواهیم ۳ کتاب ریاضی عمومی ۱، ۵ کتاب معادلات دیفرانسیل و ۴ کتاب آمار مهندسی را در کنار یکدیگر در یک قفسه قرار دهیم. احتمال اینکه هر ۳ کتاب ریاضی عمومی ۱ پهلوی هم قرار گیرند را بیابید.

حل در اینجا ۳ کتاب ریاضی عمومی ۱ مانند ۳ حرف M، ۵ کتاب معادلات مانند ۵ حرف E و ۴ کتاب آمار مانند ۴ حرف S می‌باشند پس تعداد طریق قرار گرفتن آنها در یک قفسه $n(S) = \frac{12!}{3!5!4!}$ است. حال اگر A پیشامد قرار گرفتن ۳ کتاب ریاضی عمومی ۱ پهلوی هم باشد آنگاه $n(A) = \frac{10!}{1!5!4!}$ زیرا ۳ کتاب ریاضی عمومی ۱ در حکم یک گروه متصل MMM می‌باشند. بنابراین

$1 \to 1+5+4=10$

$$P(A) = \frac{n(A)}{n(S)} = \frac{10!}{1!5!4!}\times\frac{3!5!4!}{12!} = \frac{3!10!}{12!} = \frac{1}{22}$$

ترکیب اگر در قرار دادن اعضای متمایز یک مجموعه در کنار یکدیگر (و یا انتخاب اعضا از یک مجموعه) ترتیب قرار گرفتن اعضا در کنار یکدیگر (ترتیب انتخاب اعضا) مهم نباشد، در این صورت جایگشت حاصله را ترکیب گویند.

مثال ۱۲.۵.۲ حروف a,b,c,d,e را در نظر بگیرید

الف- از این حروف چند کلمه دو حرفی بدون تکرار حروف می‌توان ساخت به طوری که ترتیب قرار گرفتن حروف مهم نباشد.

ب- از این حروف چند کلمه هفت حرفی می‌توان ساخت به طوری که ترتیب قرار گرفتن حروف مهم نباشد.

حل الف- این مثال همانند مثال ۷.۵.۲ است با این تفاوت که ترتیب قرار گرفتن حروف مهم نیست، یعنی از دو حالت ab و ba فقط بایستی یکی را انتخاب کرد و... بنابراین تعداد کل حالات از تقسیم P_2^5 بر !۲ (تعداد جایگشتهای دو حرف انتخابی) به دست می‌آید یعنی

$$\text{تعداد کلمات} = \frac{P_2^5}{2!} = \frac{5!}{2!(5-2)!} = 10$$

این تعداد را با $\binom{5}{2}$ یا C_2^5 نمایش داده و آن را ترکیب ۲ از ۵ گویند.

ب- در شکل ۷.۲-الف بعضی از کلمات ۷ حرفی مورد نظر مشخص شده‌اند. این حالات را می‌توان به صورت دیگری همانند شکل ۷.۲-ب نمایش داد که در آن تعداد xهای سمت چپ خط اول نمایانگر تعداد حرف a، تعداد xهای بین دو خط اول و دوم از سمت چپ نمایانگر تعداد حرف b و...

<table>
<tr><td>a a b b c d e</td><td>x x | x x | x | x | x</td></tr>
<tr><td>a b b c d e e</td><td>x | x x | x | x | x x</td></tr>
<tr><td>b c c d e e e</td><td>| x | x x | x | x x x</td></tr>
<tr><td>a b d d d e e</td><td>x | x | | x x x | x x</td></tr>
<tr><td>a a b c c e e</td><td>x x | x | x x | | x x</td></tr>
<tr><td>c c c c e e e</td><td>| | x x x x | | x x x</td></tr>
<tr><td>الف</td><td>ب</td></tr>
</table>

شکل ۷.۲ ترکیبات با تکرار

بنابراین تعداد کلمات مختلف از قرار دادن ۷ حرف x و چهار خط | به دست می‌آید که این تعداد برابر با $\frac{11!}{7!4!}$ و یا $\frac{11!}{7!4!} = \binom{5+7-1}{7}$ است.

در حالت کلی داریم که

اگر از بین n عنصر متمایز بخواهیم r عنصر را انتخاب کنیم به طوری که ترتیب انتخاب مهم نباشد در این صورت

الف- اگر تکرار عناصر مجاز نباشد آنگاه تعداد راه‌های ممکن برابر است با

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

ب- اگر تکرار عناصر مجاز باشد آنگاه تعداد راه‌های ممکن برابر است با

$$\binom{n+r-1}{r}$$

به $\binom{n}{r}$ ترکیب r از n گویند و همواره بایستی $0 \le r \le n$ باشد.

مثال ۱۳.۵.۲ از بین ۴ پزشک و ۲ پرستار می‌خواهیم یک کمیته ۴ نفری تشکیل دهیم.

الف- احتمال اینکه اعضای کمیته شامل ۲ پزشک و ۲ پرستار باشد را بیابید.

ب- احتمال اینکه اعضای کمیته شامل حداقل ۲ پرستار باشد را بیابید.

حل چون در انتخاب افراد ترتیب مهم نیست، بنابراین تعداد انتخاب ۴ نفر از این ۷ نفر برابر $n(S) = \binom{7}{4}$ است.

الف- انتخاب ۲ پزشک به $\binom{4}{2}$ و انتخاب ۲ پرستار به $\binom{3}{2}$ طریق انجام می‌شود بنابراین اگر A پیشامد انتخاب ۲ پزشک و ۲ پرستار باشد آنگاه $n(A) = \binom{3}{2}\binom{4}{2}$ و در نتیجه

$$P(A) = \frac{\binom{3}{2}\binom{4}{2}}{\binom{7}{4}}$$

ب- انتخاب حداقل ۲ پرستار به معنای انتخاب ۲ یا ۳ پرستار است، بنابراین اگر B پیشامد انتخاب حداقل ۲ پرستار باشد آنگاه طبق اصل جمع $n(B) = \binom{3}{2}\binom{4}{2} + \binom{3}{3}\binom{4}{1}$ است و در نتیجه

$$P(B) = \frac{\binom{3}{2}\binom{4}{2} + \binom{3}{3}\binom{4}{1}}{\binom{7}{4}}$$

مثال ۱۴.۵.۲ از جعبه‌ای که شامل ۵ مهره آبی، ۳ مهره سفید و ۴ مهره قرمز است، ۶ مهره به تصادف یک به یک و بدون جایگذاری انتخاب می‌کنیم.

الف- احتمال اینکه ۲ مهره آبی و ۱ مهره سفید انتخاب شوند را بیابید.

ب- احتمال اینکه از هر رنگ به تعداد مساوی مهره انتخاب شود را بیابید.

حل در اینجا ترتیب انتخاب مهره‌ها مهم نیست، پس $n(S) = \binom{12}{6}$

الف- اگر A پیشامد انتخاب ۲ مهره آبی و ۱ مهره سفید باشد آنگاه

$$P(A) = \frac{n(A)}{n(S)} = \frac{\binom{5}{2}\binom{3}{1}\binom{4}{3}}{\binom{12}{6}}$$

ب- اگر B پیشامد انتخاب از هر رنگ به تعداد مساوی باشد آنگاه

$$P(B) = \frac{n(B)}{n(S)} = \frac{\binom{5}{2}\binom{3}{2}\binom{4}{2}}{\binom{12}{6}}$$

مثال ۱۵.۵.۲ تعداد جوابهای صحیح و غیرمنفی معادله $x_1 + x_2 + \dots + x_n = r$ که در آن $x_i \ge 0$ ، $i = 1, 2, \dots, n$ را به دست آورید.

حل مساله مانند این است که بخواهیم r مهره را در n جعبه قرار دهیم بطوریکه تکرار مهره‌ها در جعبه‌ها مجاز و ترتیب قرار گرفتن مهره‌ها مهم نباشد. بنابراین با توجه به مثال ۱۲.۵.۲(ب) تعداد حالات ممکن برابر $\binom{n+r-1}{r}$ می‌باشد. توجه کنید که با مقایسه با مثال ۱۲.۵.۲(ب) مهره‌ها همان xها و دیواره‌های وسط جعبه‌ها همان خطوط | می‌باشند.

مثال ۱۶.۵.۲ مدیر یک شرکت خصوصی می‌خواهد ۵ سکه بهار آزادی را به عنوان پاداش بین ۳ کارمند A و B و C تقسیم کند احتمال اینکه به کارمند A حداقل ۲ سکه پاداش دهد را بیابید.

حل تعداد کل حالات پاداش دادن از حل معادله $x_1+x_2+x_3=5 \quad x_i \geq 0$ به دست می‌آید پس $n(S)=\binom{3+5-1}{5}$ اگر A پیشامد این باشد که کارمند A حداقل ۲ سکه دریافت کند آنگاه تعداد راههای ممکن از حل معادله $x_1+x_2+x_3=5,\ x_1\geq 2,\ x_2,x_3\geq 0$ و یا حل معادله $y_1+x_2+x_3=3,\ y_1=x_1-2\geq 0,\ x_2,x_3\geq 0$ به دست می‌آید بنابراین $n(A)=\binom{3+3-1}{3}$ و در نتیجه

$$P(A)=\frac{n(A)}{n(S)}=\frac{\binom{5}{3}}{\binom{7}{5}}=\frac{10}{21}$$

۶.۲ مدل احتمال روی فضای نمونه نامتناهی

در بخشهای قبل در مورد محاسبه احتمال در فضای نمونه متناهی بحث کردیم. در این قسمت در مورد محاسبه احتمال در فضاهای نمونه نامتناهی شمارش‌پذیر و پیوسته بحث می‌کنیم.

۱.۶.۲ فضای نمونه نامتناهی شمارش‌پذیر

در این حالت فضای نمونه به صورت یک مجموعه نامتناهی شمارش‌پذیر مانند $S=\{e_1,e_2,e_3,\ldots\}$ است که به هر یک از نقاط e_i احتمالات $0\leq p_i\leq 1$ را به گونه‌ای نسبت می‌دهیم که $\sum_{i=1}^{+\infty} p_i=1$. در این فضا هر پیشامد زیرمجموعه‌ای از فضای نمونه است و احتمال هر پیشامد را همانند حالت فضای نمونه متناهی محاسبه می‌کنیم.

مثال ۱.۶.۲ سکه‌ای را آنقدر پرتاب می‌کنیم تا یک شیر مشاهده کنیم و سپس توقف می‌کنیم.

الف- احتمال اینکه تعداد فردی پرتاب لازم باشد را بیابید.

ب- احتمال اینکه حداقل ۷ پرتاب لازم باشد را بیابید.

حل فضای نمونه و احتمالات نسبت داده شده به نقاط فضای نمونه به صورت زیر می‌باشد

S	H	TH	TTH	$TTTH$	....
احتمال	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	....

توجه کنید که مجموع کل احتمالات برابر ۱ می‌شود زیرا با توجه به اینکه این مجموع یک سری هندسی را تشکیل می‌دهد داریم که

$$\sum_{i=1}^{+\infty} p_i=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\ldots=\frac{\frac{1}{2}}{1-\frac{1}{2}}=1$$

الف- اگر e_i نمایانگر تعداد i پرتاب تا رسیدن به شیر باشد و A پیشامد تعداد فردی پرتاب باشد آنگاه $A=\{e_1,e_3,e_5,\ldots\}$ و در نتیجه

$$P(A)=\frac{1}{2}+\frac{1}{8}+\frac{1}{32}+\ldots=\frac{\frac{1}{2}}{1-\frac{1}{4}}=\frac{2}{3}$$

ب- اگر B پیشامد حداقل ۷ پرتاب باشد آنگاه $B=\{e_7,e_8,e_9,\ldots\}$ و در نتیجه

$$P(B)=\left(\frac{1}{2}\right)^7+\left(\frac{1}{2}\right)^8+\left(\frac{1}{2}\right)^9+\ldots=\frac{\left(\frac{1}{2}\right)^7}{1-\frac{1}{2}}=\left(\frac{1}{2}\right)^6=\frac{1}{64}$$

تذکر توجه کنید که در این حالت نمی‌توان یک مدل احتمال یکنواخت روی فضای نمونه نامتناهی شمارش‌پذیر پیاده کرد. زیرا اگر $p_i=p \quad \forall i=1,2,3,\ldots$ در این صورت

$$\sum_{i=1}^{+\infty}p_i=\sum_{i=1}^{+\infty}p=+\infty$$

۲.۶.۲ فضای نمونه پیوسته

در یک حالت خاص فضای نمونه پیوسته را می‌توان به صورت یک فاصله کراندار $S=[a,b]$ از اعداد حقیقی (یا یک سطح محدود شده در فضای دو بعدی یا...) در نظر گرفت. در این حالت هر پیشامد می‌تواند به صورت یک زیر فاصله یا اجتماعی از زیرفاصله‌ها (یا یک زیرسطح در فضای دو بعدی یا...) باشد و اگر کل احتمال را به عنوان یک واحد جرم که به طور پیوسته و یکنواخت روی فاصله (یا سطح یا...) توزیع شده است، در نظر بگیریم آنگاه احتمال هر پیشامد A در این فضا را می‌توان به صورت زیر محاسبه می‌کرد

$$P(A)=\frac{\text{طول زیر فاصله } A}{\text{طول فاصله } S} \quad (\text{یا}=\frac{\text{مساحت ناحیه } A}{\text{مساحت ناحیه } S} \quad \text{یا} \ldots)$$

تذکر فضای نمونه پیوسته و همچنین پیشامدها و محاسبه احتمال در این فضا دارای مفاهیمی وسیعتر از موارد گفته شده در بالا می‌باشند. این مفاهیم در کتابهای پیشرفته احتمال مورد بررسی قرار می‌گیرند.

مثال ۲.۶.۲ عددی را به تصادف از فاصله اعداد حقیقی $[۱,۴]$ انتخاب می‌کنیم.

الف- احتمال اینکه عدد انتخابی در فاصله $[۲,۳/۵]$ باشد را بیابید.

ب- احتمال اینکه عدد انتخابی دقیقاً ۲ باشد را بیابید.

حل الف- در اینجا $S=[۱,۴]$ و $A=[۲,۳/۵]$ بنابراین

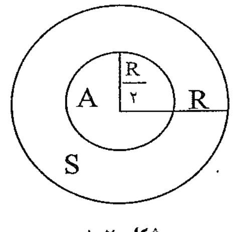
$$P(A)=\frac{\text{طول فاصله } A}{\text{طول فاصله } S}=\frac{۳/۵-۲}{۴-۱}=\frac{۱/۵}{۳}=\frac{۱}{۲}$$

ب- چون در فاصله $[۱,۴]$ بی نهایت عدد وجود دارد و انتخاب یک عدد بخصوص در بین این اعداد غیرممکن است پس $P(\{۲\})=۰$.

تذکر با توجه به مثال بالا، احتمال هر پیشامد تک عضوی در هر فضای نمونه پیوسته صفر می‌باشد.

مثال ۳.۶.۲ از داخل دایره‌ای به شعاع R نقطه‌ای را به تصادف انتخاب می‌کنیم. احتمال اینکه فاصله نقطه انتخابی تا مرکز دایره کمتر از فاصله آن تا محیط دایره باشد را بیابید.

حل با توجه به شکل ۸.۲ فضای نمونه S شامل کلیه نقاط درون دایره است و نقاطی که درون دایره به شعاع $\frac{R}{۲}$ باشند فاصله‌اشان تا مرکز کمتر از فاصله‌اشان تا محیط دایره است و بنابراین پیشامد A مورد نظر ما را تشکیل می‌دهند. در نتیجه



شکل ۸.۲

$$P(A)=\frac{\text{مساحت ناحیه } A}{\text{مساحت ناحیه } S}=\frac{\text{مساحت دایره به شعاع } \frac{R}{۲}}{\text{مساحت دایره به شعاع } R}=\frac{\pi\left(\frac{R}{۲}\right)^{۲}}{\pi R^{۲}}=\frac{۱}{۴}$$

۷.۲ احتمال شرطی

در بعضی از مسایل نیاز به محاسبه احتمال رخداد پیشامد B را داریم مشروط بر اینکه پیشامد A اتفاق افتاده باشد. برای درک این موضوع به مثال زیر توجه کنید.

مثال ۱.۷.۲ یک تاس را که شانس رخداد عدد زوج در آن دو برابر عدد فرد می‌باشد را یک بار پرتاب می‌کنیم.

الف- احتمال رخداد یک عدد زوج را بیابید.

ب- اگر بدانیم عدد حاصل از پرتاب تاس از ۳ بزرگتر بوده است، احتمال رخداد یک عدد زوج را بیابید.

حل الف- مدل احتمال در این آزمایش تصادفی عبارت است از

S	۱	۲	۳	۴	۵	۶
احتمالات	$\frac{۱}{۹}$	$\frac{۲}{۹}$	$\frac{۱}{۹}$	$\frac{۲}{۹}$	$\frac{۱}{۹}$	$\frac{۲}{۹}$

اگر B پیشامد رخداد عدد زوج باشد در این صورت $B=\{۲,۴,۶\}$ و بنابراین $P(B)=\frac{۶}{۹}=\frac{۲}{۳}$.

ب- بر اساس اطلاع داده شده فضای نمونه جدید عبارت است از $A=\{۴,۵,۶\}$ که چون شانس مشاهده عدد زوج دو برابر عدد فرد است، پس مدل احتمال روی این فضای نمونه جدید عبارت است از

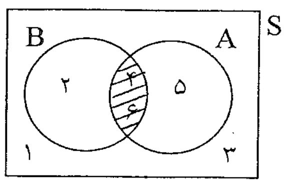
A	۴	۵	۶
احتمالات	$\frac{۲}{۵}$	$\frac{۱}{۵}$	$\frac{۲}{۵}$

بنابراین احتمال وقوع پیشامد B به شرط آنکه پیشامد A به وقوع پیوسته باشد برابر $\frac{۴}{۵}$ است که آن را با نماد $P(B\mid A)=\frac{۴}{۵}$ نمایش می‌دهند.

احتمال شرطی $P(B\mid A)$ را می‌توان از همان فضای نمونه اولیه به صورت زیر محاسبه کرد. با توجه به شکل ۹.۲ داریم که $P(B\mid A)\propto P(A\cap B)$ و در نتیجه $P(B\mid A)=kP(A\cap B)$. اما همواره داریم که $P(A\mid A)=۱$

و در نتیجه $۱=P(A\mid A)=kP(A\cap A)=kP(A)$

و اگر $P(A)\neq ۰$ باشد آنگاه $k=\frac{۱}{P(A)}$ و بنابراین



شکل ۹.۲ احتمال شرطی

$$P(B\mid A)=\frac{P(A\cap B)}{P(A)} \quad , \quad P(A)\neq ۰$$

برای مثال در مثال ۱.۷.۲ داریم که

$$A=\{4,5,6\} \Rightarrow P(A)=\frac{5}{9}$$

$$A\cap B=\{4,6\} \Rightarrow P(A\cap B)=\frac{4}{9}$$

در نتیجه

$$\frac{P(A\cap B)}{P(A)}=\frac{\frac{4}{9}}{\frac{5}{9}}=\frac{4}{5}=P(B\mid A)$$

که با مقدار به دست آمده از مثال فوق مطابقت دارد.

تعریف ۷.۲ احتمال شرطی پیشامد B به شرط وقوع پیشامد A که آن را با نماد $P(B\mid A)$ نمایش می‌دهند، به صورت زیر تعریف می‌شود.

$$\boxed{P(B\mid A)=\frac{P(A\cap B)}{P(A)} \quad , \quad P(A)\neq 0}$$ (۵.۲)

مثال ۲.۷.۲ جعبه‌ای شامل ۳ مهره سفید و ۴ مهره سیاه و ۲ مهره قرمز است. از این جعبه یک مهره به تصادف خارج می‌کنیم اگر این مهره سفید نباشد احتمال اینکه این مهره سیاه باشد را بیابید.

حل اگر B پیشامد سیاه بودن مهره و W' پیشامد سفید نبودن مهره باشد آنگاه احتمال مطلوب عبارت است از

$$P(B\mid W')=\frac{P(B\cap W')}{P(W')}=\frac{P(B)}{P(W')}=\frac{\frac{4}{9}}{\frac{6}{9}}=\frac{2}{3}$$

مثال ۳.۷.۲ جدول زیر تعداد قطعات سالم و معیوب تولیدی توسط دو کارخانه ۱ و ۲ را نشان می‌دهد. اگر یک قطعه به طور تصادفی انتخاب شود و این قطعه سالم باشد، احتمال اینکه از کارخانه ۱ انتخاب شده باشد را بیابید.

	کارخانه ۱	کارخانه ۲	جمع
معیوب	۱۵	۵	۲۰
سالم	۴۵	۳۵	۸۰
جمع	۶۰	۴۰	۱۰۰

حل اگر A پیشامد انتخاب قطعه سالم و B پیشامد انتخاب قطعه از کارخانه ۱ باشد در این صورت

$$P(B\mid A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{45}{100}}{\frac{80}{100}}=\frac{45}{80}$$

قانون ضرب احتمال از رابطه (۵.۲) نتیجه زیر به دست می‌آید.

نتیجه ۴.۲ اگر A و B دو پیشامد باشند که بتوانند همزمان اتفاق بیفتند آنگاه

$$\boxed{P(A\cap B)=P(A)\,P(B\mid A) \quad , \quad P(A)\neq 0}$$ (۶.۲)

این فرمول را قانون ضرب احتمال گویند.

مثال ۴.۷.۲ در مثال ۲.۷.۲ اگر دو مهره به تصادف یک به یک و بدون جایگذاری از جعبه خارج کنیم، مطلوب است

الف- احتمال اینکه مهره انتخابی اول سفید و دومی سیاه باشد را بیابید.

ب- احتمال اینکه مهره انتخابی اول سیاه و دومی قرمز باشد را بیابید.

حل اگر قرار دهیم

$W_i =$ پیشامد اینکه مهره انتخابی iام سفید باشد

$B_i =$ پیشامد اینکه مهره انتخابی iام سیاه باشد i=۱,۲

$R_i =$ پیشامد اینکه مهره انتخابی iام قرمز باشد

در این صورت $P(W_1)$ به معنای احتمال انتخاب اولین مهره سفید و $P(B_2\mid W_1)$ به معنای احتمال انتخاب دومین مهره سیاه به شرط آنکه بدانیم اولین مهره انتخابی سفید بوده است. بنابراین

الف- $P(W_1\cap B_2)=P(W_1)P(B_2\mid W_1)=\frac{3}{9}\times\frac{4}{8}=\frac{1}{6}$

ب- $P(B_1\cap R_2)=P(B_1)P(R_2\mid B_1)=\frac{4}{9}\times\frac{2}{8}=\frac{1}{9}$

نتیجه ۴.۲ را می‌توان به حالت کلی‌تر زیر تعمیم داد.

نتیجه ۵.۲ اگر A_1 ، A_2 ، ... و A_k پیشامدهایی باشند که بتوانند همزمان اتفاق بیفتند آنگاه

$$\boxed{\begin{aligned}P(A_1\cap A_2\cap A_3\cap\dots\cap A_k)&=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1\cap A_2)\\&\dots P(A_k\mid A_1\cap A_2\cap\dots\cap A_{k-1})\end{aligned}}$$ (۷.۲)

مثال ۵.۷.۲ در مثال ۲.۷.۲ فرض کنید ۳ مهره یک به یک و بدون جایگذاری از جعبه خارج کنیم. مطلوب است

الف- احتمال اینکه مهره انتخابی اول قرمز، دومی سفید و سومی قرمز باشد را بیابید.

ب- احتمال اینکه دو مهره قرمز و یک سفید انتخاب شوند را بیابید.

حل الف- با توجه به پیشامدهای معرفی شده در مثال ۴.۷.۲ داریم که

$$P(R_1\cap W_2\cap R_3)=P(R_1)P(W_2\mid R_1)P(R_3\mid R_1\cap W_2)=\frac{۲}{۹}\times\frac{۳}{۸}\times\frac{۱}{۷}=\frac{۱}{۸۴}$$

ب- در این حالت ترتیب انتخاب مهره‌ها مهم نیست. اگر A پیشامد انتخاب ۲ مهره قرمز و یک مهره سفید از جعبه باشد در این صورت

$$P(A)=\frac{n(A)}{n(S)}=\frac{\binom{۲}{۲}\binom{۳}{۱}}{\binom{۹}{۳}}=\frac{۳}{۸۴}=\frac{۱}{۲۸}$$

پیشامدهای مستقل در مثال ۴.۷.۲ فرض کنید پس از انتخاب مهره اول آن را به جعبه بازگردانده و سپس مهره دوم را انتخاب کنیم. در این صورت انتخاب مهره اول تأثیری در انتخاب مهره دوم ندارد و داریم که

$$P(B_2\mid W_1)=P(B_2)=\frac{۴}{۹}\Rightarrow\frac{P(W_1\cap B_2)}{P(W_1)}=P(B_2)$$

و یا

$$P(W_1\cap B_2)=P(W_1)P(B_2)$$

در این حالت پیشامدهای فوق را از یکدیگر مستقل گویند.

تعریف ۸.۲ دو پیشامد A و B را از یکدیگر مستقل گویند اگر و فقط اگر

$$\boxed{P(A\cap B)=P(A)P(B)}\qquad(۸.۲)$$

مثال ۶.۷.۲ یک ایستگاه آتش نشانی دارای دو ماشین آتش نشان است که به طور مستقل کار می‌کنند و احتمال اینکه یک ماشین آتش‌نشان در موقع نیاز موجود باشد ۰/۹۹ است. احتمال اینکه موقع نیاز حداقل یکی از این دو ماشین موجود باشند را بیابید.

حل اگر $A_i,\ i=۱,۲$ پیشامد این باشد که ماشین iام موقع نیاز موجود باشد آنگاه $P(A_1)=P(A_2)=۰/۹۹$ و A_1 و A_2 از یکدیگر مستقل هستند، بنابراین

$$P(A_1\cup A_2)=P(A_1)+P(A_2)-P(A_1\cap A_2)$$
$$=P(A_1)+P(A_2)-P(A_1)(A_2)=۰/۹۹+۰/۹۹-(۰/۹۹)^۲=۰/۹۹۹۹$$

استقلال سه پیشامد سه پیشامد A، B و C را مستقل گوئیم اگر و فقط اگر روابط زیر برقرار باشند

۱- $P(A\cap B)=P(A)P(B)$

۲- $P(A\cap C)=P(A)P(C)$

۳- $P(B\cap C)=P(B)P(C)$

۴- $P(A\cap B\cap C)=P(A)P(B)P(C)$

سه رابطه اول می‌گویند که A و B و C بایستی دو به دو از یکدیگر مستقل باشند.

رابطه استقلال و ناسازگاری دو پیشامد همان طور که در قسمتهای قبل مشاهده کردیم دو پیشامد در صورتی ناسازگار هستند که نتوانند همزمان اتفاق بیفتند و دو پیشامد در صورتی مستقل هستند که بتوانند همزمان اتفاق بیفتند اما تأثیری روی یکدیگر نداشته باشند. حال اگر دو پیشامد بخواهند هم ناسازگار و هم مستقل باشند آنگاه بایستی $P(A\cap B)=۰$ و $P(A\cap B)=P(A)P(B)$ و یا بایستی $P(A)=۰$ یا $P(B)=۰$ باشد.

مثال ۷.۷.۲ یک جفت تاس را دو مرتبه پرتاب می‌کنیم. احتمال آنکه مجموع اعداد روی دو تاس ۵ و ۱۰ شود را بیابید.

حل اگر قرار دهیم

$A_i=$ پیشامد اینکه در پرتاب iام مجموع ۵ شود

$B_i=$ پیشامد اینکه در پرتاب iام مجموع ۱۰ شود i=۱,۲

در این صورت احتمال مطلوب عبارت است از

$$P((A_1\cap B_2)\cup(B_1\cap A_2))=P(A_1\cap B_2)+P(B_1\cap A_2)$$ (به دلیل ناسازگاری)

$$=P(A_1)P(B_2)+P(B_1)P(A_2)$$ (به دلیل استقلال)

$$=\frac{۴}{۳۶}\times\frac{۳}{۳۶}+\frac{۳}{۳۶}\times\frac{۴}{۳۶}=\frac{۱}{۵۴}$$

۸.۲ فرمول احتمال بیز و فرمول تفکیک احتمال

یکی از فرمولهای مهم احتمال فرمول احتمال بیز می‌باشد که ابتدا آن را با ذکر یک مثال تشریح می‌کنیم.

مثال ۱.۸.۲ فرض کنید ۴۰٪ افراد یک شهر را مردان و ۶۰٪ آنان را زنان تشکیل دهند. همچنین فرض کنید ۵۰٪ از مردان و ۳۰٪ از زنان سیگاری باشند. اگر شخصی از بین افراد سیگاری به تصادف انتخاب شود احتمال اینکه این شخص مرد باشد را بیابید.

حل اگر پیشامدهای زیر را تعریف کنیم

$M=$ پیشامد اینکه شخص انتخابی مرد باشد

$W =$ پیشامد اینکه شخص انتخابی زن باشد

$A =$ پیشامد اینکه شخص انتخابی سیگاری باشد

در این صورت از مفروضات مسأله داریم که

$$P(W)=۰/۶۰ \quad , \quad P(M)=۰/۴۰$$
$$P(A\mid W)=۰/۳۰ \quad , \quad P(A\mid M)=۰/۵۰$$

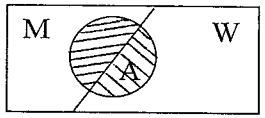
و می‌خواهیم $P(M\mid A)$ را محاسبه کنیم. برای این منظور داریم که

$$P(M\mid A)=\frac{P(M\cap A)}{P(A)}=\frac{P(M)P(A\mid M)}{P(A)} \qquad (۹.۲)$$

برای محاسبه $P(A)$ با توجه به شکل ۱۰.۲ داریم که $A=(A\cap M)\cup(A\cap W)$

بنابراین

$$P(A)=P(A\cap M)+P(A\cap W)$$
$$=P(M)P(A\mid M)+P(W)P(A\mid W)$$



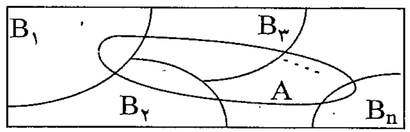
شکل ۱۰.۲

که با توجه به مفروضات مساله $P(A)$ قابل محاسبه است. این فرمول را فرمول تفکیک احتمال گوئیم زیرا با تفکیک پیشامد A به دو پیشامد مجزا، $P(A)$ را قابل محاسبه کردیم. با قرار دادن مقدار $P(A)$ در فرمول (۹.۲) فرمول زیر که به فرمول احتمال بیز معروف است به دست می‌آید.

$$P(M\mid A)=\frac{P(M)P(A\mid M)}{P(M)P(A\mid M)+P(W)P(A\mid W)}$$
$$=\frac{(۰/۴۰)(۰/۵۰)}{(۰/۴۰)(۰/۵۰)+(۰/۶۰)(۰/۳۰)}\approx ۰/۵۳$$

برای به دست آوردن فرمول احتمال بیز در حالت کلی، فرض کنید $B_۱$، $B_۲$، ... و B_n پیشامدهایی باشند که دو به دو ناسازگار بوده و اجتماع تمام آنها برابر فضای نمونه S باشد. یعنی

$$B_i\cap B_j=\emptyset \qquad \forall i\neq j$$
$$S=\bigcup_{i=۱}^{n} B_i$$



شکل ۱۱.۲ تفکیک پیشامدها

اصطلاحاً پیشامدهای $B_۱$، $B_۲$، ... و B_n را یک افراز برای فضای نمونه S گویند و یا گویند فضای نمونه S به پیشامدهای $B_۱$، $B_۲$، ... و B_n افراز یا تفکیک شده است. حال اگر A پیشامدی با احتمال مثبت باشد داریم که

$$A=A\cap S=A\cap\left(\bigcup_{i=۱}^{n}B_i\right)=(A\cap B_۱)\cup(A\cap B_۲)\cup\ldots\cup(A\cap B_n)$$

که در آن پیشامدهای $(A\cap B_۱)$، $(A\cap B_۲)$، ... و $(A\cap B_n)$ دو به دو ناسازگار هستند. بنابراین

$$\boxed{\begin{aligned}P(A)&=P(A\cap B_۱)+P(A\cap B_۲)+\ldots+P(A\cap B_n)\\&=P(B_۱)P(A\mid B_۱)+P(B_۲)P(A\mid B_۲)+\ldots+P(B_n)P(A\mid B_n)\\&=\sum_{i=۱}^{n}P(B_i)P(A\mid B_i)\end{aligned}} \qquad (۱۰.۲)$$

فرمول فوق را فرمول تفکیک احتمال (و یا فرمول احتمال کل) گویند. همچنین

$$\boxed{P(B_j\mid A)=\frac{P(B_j\cap A)}{P(A)}=\frac{P(B_j)P(A\mid B_j)}{\sum_{i=۱}^{n}P(B_i)P(A\mid B_i)} \qquad j=۱,۲,\ldots,n} \qquad (۱۱.۲)$$

فرمول فوق را فرمول احتمال بیز گویند.

مثال ۲.۸.۲ دو جعبه وجود دارند که جعبه اول شامل ۲ مهره سفید و ۳ مهره قرمز است و جعبه دوم شامل ۳ مهره سفید و ۴ مهره قرمز است. از جعبه اول یک مهره چشم بسته انتخاب کرده و به داخل جعبه دوم می‌اندازیم و سپس از جعبه دوم به تصادف یک مهره خارج می‌کنیم. مطلوب است

الف- احتمال اینکه مهره خارج شده از جعبه دوم سفید باشد را بیابید.

ب- اگر مهره خارج شده از جعبه دوم قرمز باشد، احتمال اینکه مهره خارج شده از جعبه اول سفید بوده باشد را بیابید.

حل اگر پیشامدهای زیر را تعریف کنیم

$W_i =$ پیشامد اینکه از جعبه iام مهره سفید خارج شود

$R_i =$ پیشامد اینکه از جعبه iام مهره قرمز خارج شود $\qquad i=۱,۲$

در این صورت

الف- $P(W_۲)$ مورد سؤال است که از فرمول تفکیک احتمال داریم که

$$P(W_۲)=P(W_۱\cap W_۲)+P(R_۱\cap W_۲)$$

$$=P(W_1)\,P(W_2 \mid W_1)+P(R_1)P(W_2 \mid R_1)$$

$$=\frac{2}{5}\times\frac{4}{8}+\frac{3}{5}\times\frac{3}{8}=\frac{17}{40}$$

ب- $P(W_1 \mid R_2)$ مورد سئوال است که از فرمول احتمال بیز داریم که

$$P(W_1 \mid R_2)=\frac{P(W_1)P(R_2 \mid W_1)}{P(W_1)P(R_2 \mid W_1)+P(R_1)P(R_2 \mid R_1)}$$

$$=\frac{\frac{2}{5}\times\frac{4}{8}}{\frac{2}{5}\times\frac{4}{8}+\frac{3}{5}\times\frac{5}{8}}=\frac{8}{23}$$

مثال ۳.۸.۲ فرض کنید که سه صندوق وجود دارد که هر کدام دارای دو کشو است. در هر یک از کشوهای صندوق اول یک سکه طلا وجود دارد و در یکی از کشوهای صندوق دوم یک سکه طلا و در کشو دیگر آن یک سکه نقره وجود دارد و در هر یک از کشوهای صندوق سوم یک سکه نقره وجود دارد. یکی از صندوقهای را به تصادف انتخاب می‌کنیم و یکی از کشوهای آن را باز می‌کنیم، اگر سکه داخل این کشو طلا باشد احتمال اینکه کشوی دیگر این صندوق نیز شامل سکه طلا باشد را بیابید.

حل اگر قرار دهیم

$B_i =$ پیشامد اینکه صندوق iام انتخاب شود $\quad i=1,2,3$

$A =$ پیشامد اینکه سکه طلا مشاهده گردد

در این صورت

$$P(B_1)=P(B_2)=P(B_3)=\frac{1}{3},\ P(A \mid B_1)=1,\ P(A \mid B_2)=\frac{1}{2},\ P(A \mid B_3)=0$$

و $P(B_1 \mid A)$ مورد سئوال است. بنابراین از فرمول احتمال بیز داریم که

$$P(B_1 \mid A)=\frac{P(B_1)P(A \mid B_1)}{\sum_{i=1}^{3} P(B_i)P(A \mid B_i)}=\frac{\frac{1}{3}\times 1}{\frac{1}{3}\times 1+\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times 0}=\frac{2}{3}$$

مثال ۴.۸.۲ یک موسسه مشاوره‌ای ماشینهای مورد نیازش را از سه آژانس با احتمالهای ۲۰٪ از D، ۲۰٪ از E و ۶۰٪ از F کرایه می‌کند. اگر ۱۰٪ ماشینهای آژانس D، ۱۲٪ ماشینهای آژانس E و ۴٪ از ماشینهای آژانس F لاستیک خراب داشته باشند، مطلوب است

الف- احتمال آنکه موسسه یک ماشین با لاستیک خراب کرایه کرده باشد را بیابید.

ب- اگر ماشین کرایه شده بوسیله موسسه دارای لاستیک خراب باشد احتمال آنکه از آژانس F کرایه کرده باشد را بیابید.

حل اگر قرار دهیم

$B_i =$ پیشامد اینکه موسسه از آژانس نوع iام ماشین کرایه کند $\quad i=D,E,F$

$A =$ پیشامد اینکه ماشین کرایه شده توسط موسسه دارای لاستیک خراب باشد

در این صورت

الف- (فرمول تفکیک احتمال)

$$P(A)=P(B_D\cap A)+P(B_E\cap A)+P(B_F\cap A)$$

$$=P(B_D)P(A \mid B_D)+P(B_E)P(A \mid B_E)+P(B_F)P(A \mid B_F)$$

$$=(0/20)(0/10)+(0/20)(0/12)+(0/60)(0/04)=0/068$$

ب- (فرمول بیز)

$$P(B_F \mid A)=\frac{P(B_F)P(A \mid B_F)}{P(A)}=\frac{(0/60)(0/04)}{0/068}=\frac{6}{17}$$

۹.۲ مسائل حل شده

مثال ۱.۹.۲ سکه‌ای را یک بار پرتاب می‌کنیم اگر شیر آمد آن سکه را یک بار دیگر پرتاب می‌کنیم و اگر خط آمد یک تاس را پرتاب می‌کنیم. فضای نمونه حاصل از این آزمایش را معین کنید و پیشامد A مربوط به مشاهده عدد کمتر از ۴ در پرتاب تاس را مشخص کنید.

حل

$$S=\{(H,T),(H,H),(T,1),(T,2),\ldots,(T,6)\}$$

$$A=\{(T,1),(T,2),(T,3)\}$$

مثال ۲.۹.۲ فرض کنید A و B و C سه پیشامد باشند. پیشامدهای زیر را برحسب این سه پیشامد یا متممهای آنها بنویسید

الف- فقط A اتفاق بیفتد.

ب- حداکثر دو تا از این سه پیشامد اتفاق بیفتد.

ج- حداقل دو تا از این سه پیشامد اتفاق بیفتند.

د- دقیقاً دو تا از این سه پیشامد اتفاق بیفتند.

حل الف- $A-(B\cup C)=A\cap(B'\cap C')$

ب- $(A\cap B\cap C)'=A'\cup B'\cup C'$

ج- $(A\cap B)\cup(A\cap C)\cup(B\cap C)$

د- $\left[(A\cap B)\cup(A\cap C)\cup(B\cap C)\right]-(A\cap B\cap C)$

مثال ۳.۹.۲ فضای نمونه هر یک از آزمایشهای تصادفی زیر را مشخص کنید و تعیین کنید که کدام یک متناهی، نامتناهی شمارش‌پذیر یا پیوسته هستند.

الف- پاسخ دادن به یک آزمون تستی ۴ جوابه که شامل ۲۰ سئوال است.

ب- تعداد تلفات حوادث رانندگی یک شهر معین در سال آینده.

ج- مدت زمانی که یک تلویزیون رنگی بدون آنکه عیبی پیدا کند و بدون تعمیر به کار ادامه دهد.

حل الف- فضای نمونه متناهی

$$S=\{(x_۱,x_۲,...,x_{۲۰}) \mid x_i=A,B,C,D\ ;\ i=۱,۲,...,۲۰\}$$

ب- فضای نمونه نامتناهی شمارش‌پذیر $S=\{۰,۱,۲,۳,....\}$

ج- فضای نمونه پیوسته $S=[۰,+\infty)$

مثال ۴.۹.۲ چهار نفر برای ریاست یک شورا کاندید شده‌اند. شانس کاندید A دو برابر کاندید B، شانس کاندید B دو برابر کاندید C و شانس کاندید C و D یکی است. مطلوب است

الف- احتمال اینکه کاندید A یا B انتخاب شوند را بیابید.

ب- احتمال اینکه کاندید A انتخاب نشود را بیابید.

حل در این مساله مدل احتمال برابر است با

S	A	B	C	D
احتمالات	$\frac{۴}{۸}$	$\frac{۲}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$

الف- $P(A\cup B)=P(A)+P(B)=\frac{۴}{۸}+\frac{۲}{۸}=\frac{۳}{۴}$

ب- $P(A')=۱-P(A)=۱-\frac{۴}{۸}=\frac{۴}{۸}=\frac{۱}{۲}$

مثال ۵.۹.۲ در آزمایشی سکه نااریبی را دو بار می‌اندازیم. اگر هر دو بار شیر بیاید آزمایش را

پایان می‌دهیم و در غیر اینصورت سکه را یک بار دیگر انداخته و آزمایش را پایان می‌دهیم. احتمال اینکه یک خط مشاهده شود را بیابید.

حل مدل احتمال برابر است با

S	HH	HTT	HTH	THT	THH	TTT	TTH
احتمالات	$\frac{۱}{۴}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$	$\frac{۱}{۸}$

اگر A پیشامد مشاهده یک خط باشد در این صورت $A=\{HTH,THH\}$ و بنابراین $P(A)=\frac{۱}{۴}$.

مثال ۶.۹.۲ شخصی که در جایگاه بنزین توقف می‌کند، بازبینی لاستیکهای ماشین را با احتمال ۰/۱۲ و بازبینی روغن ماشین را با احتمال ۰/۲۹ و بازبینی هر دو را با احتمال ۰/۰۷ تقاضا می‌کند.

الف- احتمال اینکه این شخص درخواست بازبینی لاستیک یا روغن ماشینش داشته باشد را بیابید.

ب- احتمال اینکه این شخص هیچکدام از بازبینی‌ها را درخواست نکند را بیابید.

حل اگر A پیشامد بازبینی لاستیک ماشین و B پیشامد بازبینی روغن ماشین باشد در این صورت

الف- $P(A\cup B)=P(A)+P(B)-P(A\cap B)=۰/۱۲+۰/۲۹-۰/۰۷=۰/۳۴$

ب- $P(A'\cap B')=P((A\cup B)')=۱-P(A\cup B)=۱-۰/۳۴=۰/۶۶$

مثال ۷.۹.۲ اگر A و B و C سه پیشامد دلخواه باشند ثابت کنید که

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$$

حل $P(A\cup B\cup C)=P(A\cup(B\cup C))=P(A)+P(B\cup C)-P(A\cap(B\cup C))$

$$=P(A)+\left[P(B)+P(C)-P(B\cap C)\right]-P((A\cap B)\cup(A\cap C))$$

$$=P(A)+P(B)+P(C)-P(B\cap C)-\left[P(A\cap B)+P(A\cap C)-P(A\cap B\cap C)\right]$$

$$=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$$

مثال ۸.۹.۲ سه نفر به طور مستقل در کشف پیام‌های رمز کار می‌کنند. احتمال کشف رمز برحسب تجربه‌ای که از کار آنها داریم به ترتیب $\frac{۱}{۳}$، $\frac{۱}{۴}$ و $\frac{۱}{۵}$ است. احتمال آنکه پیام رمز توسط آنها کشف شود را بیابید.

حل اگر A، B و C به ترتیب پیشامد این باشد که نفر اول، دوم و سوم پیام رمز را کشف کنند در

این صورت A، B و C از یکدیگر مستقل هستند و بنابراین

$$P(A)=\frac{۱}{۵} \quad , \quad P(B)=\frac{۱}{۴} \quad , \quad P(C)=\frac{۱}{۳}$$

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$$
$$=P(A)+P(B)+P(C)-P(A)P(B)-P(A)P(C)-P(B)P(C)+P(A)P(B)P(C)$$
$$=\frac{۱}{۵}+\frac{۱}{۴}+\frac{۱}{۳}-\frac{۱}{۲۰}-\frac{۱}{۱۵}-\frac{۱}{۱۲}+\frac{۱}{۶۰}=\frac{۳۶}{۶۰}=\frac{۳}{۵}$$

مثال ۹.۹.۲ یک مدار شامل سه فیوز است و به گونه‌ای طراحی شده است که اگر حداقل ۲ فیوز سالم باشد آنگاه مدار متصل خواهد بود و در غیر این صورت مدار قطع می‌گردد. اگر این سه فیوز به طور مستقل عمل کنند و احتمال سالم ماندن آنها به ترتیب $\frac{۱}{۲}$، $\frac{۲}{۳}$ و $\frac{۳}{۴}$ باشد، احتمال اینکه مدار متصل باشد را بیابید.

حل اگر A، B و C به ترتیب پیشامد این باشد که فیوز اول، دوم و سوم سالم باشند در این صورت با استفاده از مثالهای ۲.۹.۲ (پ) و ۷.۹.۲ داریم که

$$P((A\cap B)\cup(A\cap C)\cup(B\cap C))$$
$$=P(A\cap B)+P(A\cap C)+P(B\cap C)-۲P(A\cap B\cap C)$$
$$=P(A)P(B)+P(A)P(C)+P(B)P(C)-۲P(A)P(B)P(C)$$
$$=\frac{۱}{۳}+\frac{۳}{۸}+\frac{۱}{۲}-۲(\frac{۱}{۴})=\frac{۱۷}{۲۴}$$

مثال ۱۰.۹.۲ چهار زوج برای یک تئاتر ۸ بلیط در یک ردیف خریداری کرده‌اند.

الف- احتمال اینکه هر زوج پهلوی یکدیگر بنشینند را بیابید.

ب- احتمال اینکه تمام مردها پهلوی هم و در سمت راست زنها باشند را بیابید.

حل الف- اگر A پیشامد این باشد که هر زوج پهلوی یکدیگر بنشینند در این صورت

$$n(S)=۸! \quad , \quad n(A)=۴!(۲!)^۴ \Rightarrow P(A)=\frac{n(A)}{n(S)}=\frac{۴!(۲!)^۴}{۸!}=\frac{۱}{۱۰۵}$$

ب- اگر B پیشامد این باشد که تمام مردها پهلوی هم و در سمت راست زنها باشند در این صورت

$$n(B)=۴!\,۴! \Rightarrow P(B)=\frac{n(B)}{n(S)}=\frac{۴!\,۴!}{۸!}=\frac{۱}{۷۰}$$

مثال ۱۱.۹.۲ سه نفر شیرازی، دو نفر تهرانی و چهار نفر اصفهانی می‌خواهند در یک صف

پهلوی هم قرار گیرند. به فرض آنکه هم شهریها از هم متمایز نباشند، احتمال اینکه دو تهرانی در دو سر صف قرار بگیرند را بیابید.

حل چون هم شهریها از یکدیگر متمایز نیستند پس $n(S)=\frac{۹!}{۳!۲!۴!}$ و اگر A پیشامد این باشد که دو تهرانی در دو سر صف بایستند آنگاه $n(A)=\frac{۷!}{۳!۴!}$ و بنابراین $P(A)=\frac{n(A)}{n(S)}=\frac{۲!۷!}{۹!}=\frac{۱}{۳۶}$.

مثال ۱۲.۹.۲ ده صندلی در یک ردیف قرار گرفته‌اند و ۲ نفر می‌خواهند روی این صندلیها بنشینند.

الف- احتمال اینکه این دو نفر پهلوی هم قرار گیرند را بیابید.

ب- احتمال اینکه بین این دو نفر یک صندلی قرار بگیرد را بیابید.

حل الف- اگر A پیشامد قرار گرفتن دو نفر پهلوی هم باشد، چون به ۹ طریق می‌توان دو صندلی پهلوی هم را انتخاب کرد پس

$$n(S)=P_۲^{۱۰}=\frac{۱۰!}{۸!}=۹۰ \quad , \quad n(A)=۹\times ۲!=۱۸ \Rightarrow P(A)=\frac{n(A)}{n(S)}=\frac{۱۸}{۹۰}=\frac{۱}{۵}$$

ب- اگر B پیشامد این باشد که بین این دو نفر یک صندلی قرار بگیرد، چون به ۸ طریق می‌توان این دو صندلی را انتخاب کرد پس

$$n(B)=۸\times ۲!=۱۶ \Rightarrow P(B)=\frac{n(B)}{n(S)}=\frac{۱۶}{۹۰}=\frac{۸}{۴۵}$$

مثال ۱۳.۹.۲ یک کشتی دارای یک چوب پرچم است که سه موضع مختلف دارد که در هر یک از آنها می‌توان یک پرچم برافراشت. فرض کنید که کشتی برای علامت دادن، چهار پرچم (از چهار نوع مختلف) داشته باشد و یک پرچم وقتی در موضعهای مختلف افراشته شود علامتهای متفاوت نشان دهد. چند علامت مختلف می‌توان ساخت؟

حل اگر قرار دهیم $i=۰,۱,۲,۳$ علامتهایی که با i پرچم می‌توان ساخت $= A_i$

در این صورت

$$n(A_۰)=۱ \quad , \quad n(A_۱)=\binom{۳}{۱}P_۱^۴=۳\times۴=۱۲$$
$$n(A_۲)=\binom{۳}{۲}P_۲^۴=۳\times۴\times۳=۳۶ \quad , \quad n(A_۳)=\binom{۳}{۳}P_۳^۴=۴\times۳\times۲=۲۴$$

بنابراین تعداد علامتهای مختلف برابر است با

$$\sum_{i=۰}^{۳} n(A_i)=۱+۱۲+۳۶+۲۴=۷۳$$

مثال ۱۴.۹.۲ احتمال اینکه روزهای تولد یک خانواده ده نفری در سال، تمام روزهای هفته را

شامل شوند را بیابید.

حل تعداد کل راه‌هایی که ده نفر می‌توانند در روزهای مختلف هفته متولد شوند از حل معادله زیر به دست می‌آید.

$$x_1+x_2+\dots+x_7=10 \quad , \quad x_i\ge 0$$

و بنابراین $n(S)=\binom{7+10-1}{10}=\binom{16}{10}$. اگر A پیشامد این باشد که روزهای تولد این ۱۰ نفر تمام روزهای هفته را شامل شود در اینصورت $n(A)$ از حل معادله زیر به دست می‌آید

$$x_1+x_2+\dots+x_7=10 \quad , \quad x_i\ge 1$$

و یا با قرار دادن $y_i=x_i-1\ge 0$ داریم که

$$y_1+y_2+\dots+y_7=3 \quad , \quad y_i\ge 0$$

بنابراین $n(A)=\binom{7+3-1}{3}=\binom{9}{3}$ و در نتیجه

$$P(A)=\frac{n(A)}{n(S)}=\frac{\binom{9}{3}}{\binom{16}{10}}=\frac{3}{286}$$

مثال ۱۵.۹.۲ به چند طریق می‌توان n مهره نامتمایز را در k جعبه قرار داد بطوریکه هر جعبه شامل حداقل ۲ مهره باشد $(n\ge 2k)$.

حل مساله مانند حل معادله زیر است

$$x_1+x_2+\dots+x_k=n \quad , \quad x_i\ge 2$$

و یا با قرار دادن $y_i=x_i-2\ge 0$ داریم که $y_1+y_2+\dots+y_K=n-2k$

بنابراین تعداد راه‌های ممکن برابر است با $\binom{k+n-2k-1}{n-2k}=\binom{n-k-1}{n-2k}$

مثال ۱۶.۹.۲ می‌خواهیم ترتیب حروف کلمه TALLAHASSEE را عوض کنیم

الف- احتمال اینکه هر سه حرف A پهلوی هم قرار گیرند را بیابید.

ب- احتمال اینکه هیچکدام از سه حرف A پهلوی هم قرار نگیرند را بیابید.

حل الف- اگر A پیشامد این باشد که هر سه حرف A پهلوی هم قرار گیرند در این صورت

$$n(S)=\frac{11!}{1!3!2!1!2!2!} \quad , \quad n(A)=\frac{9!}{1!1!2!1!2!2!} \Rightarrow P(A)=\frac{n(A)}{n(S)}=\frac{9!3!}{11!}=\frac{3}{55}$$

ب- توجه کنید که اگر B پیشامد این باشد که هیچکدام از سه حرف A پهلوی هم قرار نگیرند آنگاه B متمم پیشامد A نیست. برای محاسبه n(B) ابتدا حروف غیر از A یعنی TLLHSSEE

را به $\frac{8!}{1!2!1!2!2!}$ مرتب می‌کنیم و سپس سه حرف A را می‌توان به $\binom{9}{3}$ طریق در بین و انتهای این حروف قرار داد به طوری که هیچ حرف A پهلوی هم نباشد بنابراین

$$n(B)=\binom{9}{3}\times\frac{8!}{1!2!1!2!2!}$$

و در نتیجه

$$P(B)=\frac{n(B)}{n(S)}=\binom{9}{3}\frac{8!3!}{11!}=\frac{56}{110}=\frac{28}{55}$$

مثال ۱۷.۹.۲ یک تاس سالم را آنقدر پرتاب می‌کنیم تا برای اولین بار عدد شش بیاید. مطلوب است

الف- احتمال اینکه تعداد پرتابهای لازم مضرب ۳ باشد را بیابید.

ب- احتمال اینکه حداقل ۳ پرتاب لازم باشد را بیابید.

حل اگر e_k نمایانگر این باشد که در k امین پرتاب برای اولین بار عدد شش بیاید در این صورت مدل احتمال برای این مساله عبارت است از

S	e_1	e_2	e_3	e_4	....
احتمالات	$\frac{1}{6}$	$(\frac{5}{6})(\frac{1}{6})$	$(\frac{5}{6})^2(\frac{1}{6})$	$(\frac{5}{6})^3(\frac{1}{6})$	....

الف- اگر A پیشامد این باشد که تعداد پرتابهای لازم مضرب ۳ است در این صورت

$$A=\{e_3,e_6,e_9,\dots\}\Rightarrow P(A)=(\tfrac{5}{6})^2(\tfrac{1}{6})+(\tfrac{5}{6})^5(\tfrac{1}{6})+\dots=\frac{(\frac{5}{6})^2(\frac{1}{6})}{1-(\frac{5}{6})^3}=\frac{25}{91}$$

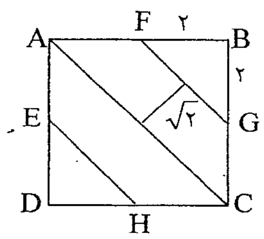
ب- اگر B پیشامد این باشد که حداقل ۳ پرتاب لازم است در این صورت

$$B=\{e_3,e_4,e_5,\dots\}\Rightarrow P(B)=(\tfrac{5}{6})^2(\tfrac{1}{6})+(\tfrac{5}{6})^3(\tfrac{1}{6})+\dots=\frac{(\frac{5}{6})^2(\frac{1}{6})}{1-\frac{5}{6}}=\frac{25}{36}$$

مثال ۱۸.۹.۲ نقطه‌ای را به تصادف از داخل مربع ABCD به ضلع ۴ واحد انتخاب می‌کنیم. احتمال اینکه فاصله نقطه انتخابی از قطر AC کمتر از $\sqrt{2}$ باشد را بیابید.

حل با توجه به شکل ۱۲.۲ اگر A پیشامد این باشد که فاصله نقطه انتخابی از قطر AC کمتر از $\sqrt{2}$ باشد، در این صورت

$$P(A)=\frac{S_{AEHCGF}}{S_{ABCD}}=۱-\frac{۲S_{BFG}}{S_{ABCD}}$$
$$=۱-\frac{۲(\frac{۱}{۲}\times۲\times۲)}{۴\times۴}=۱-\frac{۱}{۴}=\frac{۳}{۴}$$



شکل ۱۲.۲

مثال ۱۹.۹.۲ خانواده‌ای ۳ فرزند دارد. اگر فرزند اول و آخر از یک جنس باشد، احتمال همجنس بودن تمام فرزندان را بیابید.

حل اگر قرار دهیم

A = پیشامد همجنس بودن تمام فرزندان خانواده = $\{BBB\,,GGG\}$

B = پیشامد اینکه فرزند اول و آخر از یک جنس باشند = $\{BBB\,,BGB\,,GBG,GGG\}$

در این صورت

$$P(A)=\frac{۲}{۸}\quad,\quad P(B)=\frac{۴}{۸}\quad,\quad P(A\cap B)=\frac{۲}{۸}$$

و $P(A\mid B)$ مورد سئوال است. بنابراین

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}=\frac{۲}{۴}=\frac{۱}{۲}$$

مثال ۲۰.۹.۲ از جعبه‌ای محتوی ۹ کارت با شماره‌های ۱ تا ۹، ۲ کارت را به تصادف بیرون می‌آوریم، اگر بدانیم که مجموع دو عدد زوج است، احتمال اینکه هر دو عدد فرد باشند را بیابید.

حل اگر قرار دهیم

A = پیشامد اینکه مجموع اعداد روی دو کارت زوج باشد

B = پیشامد اینکه اعداد روی دو کارت هر دو فرد باشند

در این صورت $B\subset A$ و بنابراین

$$P(B\mid A)=\frac{P(A\cap B)}{P(A)}=\frac{P(B)}{P(A)}=\frac{\frac{۲۰}{۷۲}}{\frac{۳۲}{۷۲}}=\frac{۲۰}{۳۲}=\frac{۵}{۸}$$

مثال ۲۱.۹.۲ احتمال اینکه مرد متأهلی نمایش مخصوصی از تلویزیون را تماشا کند ۰/۴ و احتمال اینکه همسر او همان نمایش را تماشا کند ۰/۵ و احتمال اینکه این مرد برنامه‌ای را تماشا کند که همسر او در حال تماشای آن است ۰/۷ می‌باشد. مطلوب است

الف - احتمال اینکه هر دو نفر برنامه را تماشا کنند را بیابید.

ب - احتمال اینکه این خانم برنامه‌ای را تماشا کند که شوهرش در حال تماشای آن است را بیابید.

ج - احتمال اینکه حداقل یک نفر از این زوج برنامه را تماشا کنند را بیابید.

حل قرار می‌دهیم

M = پیشامد اینکه مرد برنامه تلویزیونی را تماشا کند

W = پیشامد اینکه همسر او برنامه تلویزیونی را تماشا کند

در این صورت

$$P(M)=۰/۴\quad,\quad P(W)=۰/۵\quad,\quad P(M\mid W)=۰/۷$$

الف -
$$P(M\cap W)=P(W)P(M\mid W)=۰/۵\times۰/۷=۰/۳۵$$

ب -
$$P(W\mid M)=\frac{P(M\cap W)}{P(M)}=\frac{۰/۳۵}{۰/۴}=\frac{۷}{۸}$$

ج -
$$P(M\cup W)=P(M)+P(W)-P(M\cap W)=۰/۴+۰/۵-۰/۳۵=۰/۵۵$$

مثال ۲۲.۹.۲ در یک طرح الکترونیکی، به تجربه دیده شده است که اگر یک کارگر جدید به برنامه‌های کارآموزی شرکت توجه کند، با احتمال ۰/۸۶ از تولید سهم خواهد برد و احتمال متناظر برای کارگر جدیدی که به برنامه‌های کارآموزی شرکت توجه نداشته باشد برابر با ۰/۳۵ است. اگر ۰/۸۰ همه کارگران جدید به برنامه‌های کارآموزی شرکت توجه داشته باشند، احتمال آنکه یک کارگر جدید از تولید سهم ببرد را بیابید.

حل اگر قرار دهیم

A = پیشامد اینکه کارگر جدید به برنامه‌های کارآموزی شرکت توجه کند

B = پیشامد اینکه کارگر جدید از تولید سهم ببرد

در این صورت

$$P(B\mid A)=۰/۸۶\quad,\quad P(B\mid A')=۰/۳۵\quad,\quad P(A)=۰/۸۰$$

و $P(B)$ مورد سئوال است از فرمول تفکیک احتمال داریم که

$$P(B)=P(A\cap B)+P(A'\cap B)=P(A)P(B\mid A)+P(A')P(B\mid A')$$
$$=(۰/۸۰)(۰/۸۶)+(۰/۲۰)(۰/۳۵)=۰/۷۵۸$$

مثال ۲۳.۹.۲ اگر A، B و C سه پیشامد مستقل باشند، ثابت کنید که

الف- A' و B' از یکدیگر مستقل هستند.

ب- A' و C از یکدیگر مستقل هستند.

ج- $A\cup B$ و C از یکدیگر مستقل هستند.

حل الف- با استفاده از تمرین ۵.۱۰.۲ داریم که

$$P(A'\cap B')=P((A\cup B)')=1-P(A\cup B)=1-\left[P(A)+P(B)-P(A\cap B)\right]$$
$$=1-\left[P(A)+P(B)-P(A)P(B)\right]=\left[1-P(B)\right]-P(A)\left[1-P(B)\right]$$
$$=\left[1-P(B)\right]\left[1-P(A)\right]=P(B')P(A')$$

بنابراین A' و B' از یکدیگر مستقل هستند

ب-
$$P(A'\cap C)=P(C-A)=P(C)-P(A\cap C)=P(C)-P(A)P(C)$$
$$=P(C)\left[1-P(A)\right]=P(C)P(A')$$

بنابراین A' و C از یکدیگر مستقل هستند.

ج-

$$P((A\cup B)\cap C)=P((A\cap C)\cup(B\cap C))=P(A\cap C)+P(B\cap C)-P(A\cap B\cap C)$$
$$=P(A)P(C)+P(B)P(C)-P(A)P(B)P(C)$$
$$=P(C)\left[P(A)+P(B)-P(A)P(B)\right]$$
$$=P(C)\left[P(A)+P(B)-P(A\cap B)\right]=P(C)P(A\cup B)$$

بنابراین $A\cup B$ و C از یکدیگر مستقل هستند.

مثال ۲۴.۹.۲ سه تیرانداز هر کدام یک تیر به یک هدف شلیک می‌کنند. احتمال آنکه تیرانداز اول به هدف بزند ۰/۴ است و همین احتمال برای تیراندازهای دوم و سوم به ترتیب ۰/۵ و ۰/۷ است. احتمال اینکه دو تیر به هدف بخورد و یک تیر به خطا برود را بیابید.

حل اگر قرار دهیم $i=۱,۲,۳$ پیشامد اینکه تیر انداز iام به هدف بزند = A_i

در این صورت پیشامدهای A_1، A_2 و A_3 از یکدیگر مستقل هستند و به سادگی دیده می‌شود که

A_1'، A_2، A_3 و A_1، A_2'، A_3 و همچنین A_1، A_2، A_3' از یکدیگر مستقل هستند و

$P(A_1)=۰/۴\ ,\ P(A_2)=۰/۵\ ,\ P(A_3)=۰/۷$.

بنابراین احتمال مورد نظر برابر است با

$$P(A_1\cap A_2\cap A_3')+P(A_1\cap A_2'\cap A_3)+P(A_1'\cap A_2\cap A_3)=$$
$$P(A_1)P(A_2)P(A_3')+P(A_1)P(A_2')P(A_3)+P(A_1')P(A_2)P(A_3)=$$
$$(۰/۴)(۰/۵)(۰/۳)+(۰/۴)(۰/۵)(۰/۷)+(۰/۶)(۰/۵)(۰/۷)=۰/۴۱$$

مثال ۲۵.۹.۲ احتمال اینکه مردی تا ۲۰ سال دیگر زنده باشد ۰/۶ و همین احتمال برای همسر او ۰/۹ است. مطلوب است محاسبه احتمال اینکه

الف- هیچکدام از آنها تا ۲۰ سال دیگر زنده نباشند.

ب- فقط همسر تا ۲۰ سال دیگر زنده باشد.

حل اگر قرار دهیم پیشامد اینکه مرد تا ۲۰ سال دیگر زنده باشد = M

پیشامد اینکه زن تا ۲۰ سال دیگر زنده باشد = W

در این صورت M و W از یکدیگر مستقل هستند و $P(M)=۰/۶$ و $P(W)=۰/۹$. بنابراین

الف- $P(M'\cap W')=P(M')P(W')=(۰/۴)(۰/۱)=۰/۰۴$

ب- $P(W\cap M')=P(W)P(M')=(۰/۹)(۰/۴)=۰/۳۶$

مثال ۲۶.۹.۲ سرایداری یک دسته کلید ۸تایی برای بازکردن در ۸اتاق را دارد که هر کلید تنها در یک اتاق را باز می‌کند. اگر در ۴۰ درصد از این اتاقها قفل نباشند و او ۳کلید را به طور تصادفی همراه آورده باشد، احتمال اینکه او بتواند وارد اتاقی شود را بیابید.

حل اگر قرار دهیم پیشامد اینکه در اتاق باز باشد = A

پیشامد اینکه سرایدار بتواند وارد اتاق شود = B

در این صورت $P(B)$ مورد سؤال است که یک مساله تفکیک احتمال است. بنابراین

$$P(B)=P(A\cap B)+P(A'\cap B)=P(A)P(B\mid A)+P(A')P(B\mid A')$$
$$=(۰/۴۰)(۱)+(۰/۶۰)(\frac{۳}{۸})=۰/۶۲۵$$

مثال ۲۷.۹.۲ یک آزمایش تشخیص سرطان با احتمال ۹۹ درصد برای بیماران سرطانی پاسخ مثبت می‌دهد و با احتمال ۵ درصد برای بیماران غیرسرطانی پاسخ مثبت می‌دهد. از بین بیماران

یک بیمارستان که ۷ درصد آنها سرطانی هستند بیماری را به تصادف انتخاب کرده و آزمایش فوق برای وی پاسخ مثبت داده است. احتمال اینکه این بیمار سرطانی باشد را بیابید.

حل قرار می‌دهیم

$C=$ پیشامد اینکه بیمار مبتلا به سرطان باشد

$A=$ پیشامد اینکه آزمایش برای بیمار پاسخ مثبت دهد

در این صورت

$$P(A\mid C)=۰/۹۹ \quad , \quad P(A\mid C')=۰/۰۵ \quad , \quad P(C)=۰/۰۷$$

و بنابراین $P(C\mid A)$ مورد سئوال است که یک مساله بیز است، در نتیجه

$$P(C\mid A)=\frac{P(C)P(A\mid C)}{P(C)P(A\mid C)+P(C')P(A\mid C')}=\frac{(۰/۰۷)(۰/۹۹)}{(۰/۰۷)(۰/۹۹)+(۰/۹۳)(۰/۰۵)}$$
$$=۰/۵۹۸$$

مثال ۲۸.۹.۲ جعبه‌ای شامل ۲ توپ سفید و ۳ توپ سیاه می‌باشد و همچنین ۳ کارت وجود دارد که روی آنها شماره‌های ۱، ۲ و ۳ یادداشت گردیده است. یک کارت به تصادف از بین این ۳ کارت انتخاب و به تعداد عدد روی کارت مشاهده شده از جعبه به تصادف توپ انتخاب می‌کنیم.

الف- احتمال اینکه تمامی توپهای انتخاب شده از جعبه سفید باشند را بیابید.

ب- اگر تمامی توپهای انتخاب شده از جعبه سفید باشند احتمال اینکه روی کارت عدد ۲ مشاهده شده باشد را بیابید.

حل قرار می‌دهیم

$B_i=$ پیشامد اینکه کارت شماره i انتخاب شود $\quad i=۱,۲,۳$

$W=$ پیشامد اینکه تمامی توپهای انتخابی سفید باشند

الف- از فرمول تفکیک احتمال داریم که

$$P(W)=P(B_۱\cap W)+P(B_۲\cap W)+P(B_۳\cap W)=\sum_{i=۱}^{۳}P(B_i)P(W\mid B_i)$$
$$=\frac{۱}{۳}\times\frac{\binom{۲}{۱}}{\binom{۵}{۱}}+\frac{۱}{۳}\times\frac{\binom{۲}{۲}}{\binom{۵}{۲}}+\frac{۱}{۳}\times ۰=\frac{۵}{۳۰}=\frac{۱}{۶}$$

ب- از فرمول بیز داریم که

$$P(B_۲\mid W)=\frac{P(B_۲)P(W\mid B_۲)}{P(W)}=\frac{\frac{۱}{۳}\times\frac{\binom{۲}{۲}}{\binom{۵}{۲}}}{\frac{۱}{۶}}=\frac{۱}{۵}$$

مثال ۲۹.۹.۲ دو تهیه کننده A و B برای یک شرکت تولیدی یک قطعه معین را تهیه می‌کنند. در گذشته ۵ درصد از قطعات تهیه شده بوسیله A و ۹ درصد از قطعات تهیه شده بوسیله B معیوب بوده‌اند. A چهار برابر B از قطعه مذکور را تهیه می‌کند. فرض کنید یک قطعه تهیه شده باشد و ملاحظه شود که سالم است. احتمال اینکه این قطعه بوسیله A تهیه شده باشد را بیابید.

حل قرار می‌دهیم

$A=$ پیشامد اینکه قطعه بوسیله A تهیه شود

$B=$ پیشامد اینکه قطعه بوسیله B تهیه شود

$C=$ پیشامد اینکه قطعه سالم باشد

در این صورت

$$P(A)=\frac{۴}{۵} \quad , \quad P(B)=\frac{۱}{۵} \quad , \quad P(C\mid A)=۰/۹۵ \quad , \quad P(C\mid B)=۰/۹۱$$

بنابراین

$$P(A\mid C)=\frac{P(A)P(C\mid A)}{P(A)P(C\mid A)+P(B)P(C\mid B)}=\frac{\frac{۴}{۵}\times ۰/۹۵}{\frac{۴}{۵}\times ۰/۹۵+\frac{۱}{۵}\times ۰/۹۱}=۰/۸۰۷$$

مثال ۳۰.۹.۲ فروشگاهی ۱۰۰ دستگاه کامپیوتر از چهار عرضه کننده A، B، C و D خریداری کرده است. تعداد دستگاههای سالم و معیوب هر عرضه کننده در جدول زیر آمده است.

	A	B	C	D
سالم	۲۱	۱۶	۱۹	۲۹
معیوب	۲	۳	۴	۶

الف- اگر یک دستگاه به تصادف از میان کامپیوترها انتخاب شود احتمال اینکه این دستگاه معیوب باشد را بیابید.

ب- احتمال اینکه دستگاه معیوبی که به تصادف انتخاب شده است متعلق به عرضه کننده A باشد را بیابید.

ج- احتمال اینکه دستگاه انتخاب شده یا سالم باشد و یا متعلق به عرضه کننده A نباشد را

بیابید.

حل قرار می‌دهیم

$B_i =$ پیشامد اینکه دستگاه کامپیوتر از عرضه کننده i خریداری شود $\quad i=A,B,C,D$

$E =$ پیشامد اینکه دستگاه کامپیوتر خریداری شده معیوب باشد

در این صورت

الف-

$$P(E)=\sum_{i=A}^{D} P(B_i)P(E\mid B_i)=\frac{۲۳}{۱۰۰}\times\frac{۲}{۲۳}+\frac{۱۹}{۱۰۰}\times\frac{۳}{۱۹}+\frac{۲۳}{۱۰۰}\times\frac{۴}{۲۳}+\frac{۳۵}{۱۰۰}\times\frac{۶}{۳۵}$$
$$=۰/۱۵$$

ب-

$$P(B_A\mid E)=\frac{P(B_A)P(E\mid B_A)}{P(E)}=\frac{\frac{۲۳}{۱۰۰}\times\frac{۲}{۲۳}}{۰/۱۵}=\frac{۲}{۱۵}$$

ج-

$$P(E'\cup B'_A)=P((E\cap B_A)')=۱-P(E\cap B_A)=۱-P(B_A)P(E\mid B_A)$$
$$=۱-\frac{۲۳}{۱۰۰}\times\frac{۲}{۲۳}=۱-۰/۰۲=۰/۹۸$$

مثال ۳۱.۹.۲ ظرفی محتوی ۳ مهره آبی و ۷ مهره سفید است. مهره‌ای بـه تـصادف انـتخاب می‌کنیم. اگر رنگ این مهره سفید باشد آن را دوباره در ظرف قرار می‌دهیم و دو مهره سفید دیگر به ظرف اضافه می‌کنیم. اگر مهره استخراجی آبی باشد آن را در ظرف قرار نداده و مهره دیگری نیز در ظرف قرا نمی‌دهیم. سپس برای بار دوم دو مهره از ظرف بیرون می‌آوریم.

الف- احتمال اینکه دو مهره انتخاب شده در بار دوم سفید باشند را بیابید.

ب- اگر دو مهره انتخاب شده در بار دوم آبی باشند، احتمال اینکه مهره انتخاب شده بار اول سفید باشد را بیابید.

حل قرار می‌دهیم

$B_1 =$ پیشامد اینکه مهره انتخابی در بار اول آبی باشد

$W_1 =$ پیشامد اینکه مهره انتخابی در بار اول سفید باشد

$B_2 =$ پیشامد اینکه دو مهره انتخابی در بار دوم آبی باشند

$W_2 =$ پیشامد اینکه دو مهره انتخابی در بار دوم سفید باشند

در این صورت

الف-

$$P(W_2)=P(B_1\cap W_2)+P(W_1\cap W_2)=P(B_1)P(W_2\mid B_1)+P(W_1)P(W_2\mid W_1)$$
$$=\frac{۳}{۱۰}\times\frac{\binom{۷}{۲}}{\binom{۹}{۲}}+\frac{۷}{۱۰}\times\frac{\binom{۹}{۲}}{\binom{۱۲}{۲}}=\frac{۷}{۴۰}+\frac{۲۱}{۵۵}=\frac{۱۲۲۵}{۲۲۰۰}=\frac{۴۹}{۸۸}$$

ب-

$$P(W_1\mid B_2)=\frac{P(W_1)P(B_2\mid W_1)}{P(W_1)P(B_2\mid W_1)+P(B_1)P(B_2\mid B_1)}$$
$$=\frac{\frac{۷}{۱۰}\times\frac{\binom{۳}{۲}}{\binom{۱۲}{۲}}}{\frac{۷}{۱۰}\times\frac{\binom{۳}{۲}}{\binom{۱۲}{۲}}+\frac{۳}{۱۰}\times\frac{\binom{۲}{۲}}{\binom{۹}{۲}}}=\frac{۴۲}{۵۳}$$

مثال ۳۲.۹.۲ جعبه A شامل ۴ توپ قرمز، ۲ توپ سفید و ۶ توپ سیاه است و جعبه B شامل ۳ توپ قرمز و ۵ توپ سفید است. یک تاس متعادل پرتاب می‌شود، اگر ۱ یا ۶ ظاهر شود یک توپ از ظرف A و در غیر اینصورت یک توپ از ظرف B خارج می‌کنیم.

الف- احتمال اینکه یک توپ سفید استخراج شود را بیابید.

ب- اگر یک توپ قرمز به تصادف خارج شود احتمال اینکه در پرتاب تاس ۱ یا ۶ ظاهر شده باشد را بیابید.

حل قرار می‌دهیم

$A =$ پیشامد اینکه در پرتاب تاس عدد ۱ یا ۶ ظاهر شود

$W =$ پیشامد اینکه یک توپ سفید از جعبه خارج شود

$R =$ پیشامد اینکه یک توپ قرمز از جعبه خارج شود

در این صورت

الف-

$$P(W)=P(A\cap W)+P(A'\cap W)=P(A)P(W\mid A)+P(A')P(W\mid A')$$
$$=\frac{۲}{۶}\times\frac{۲}{۱۲}+\frac{۴}{۶}\times\frac{۵}{۸}=\frac{۱۷}{۳۶}$$

ب-
$$P(A\mid R)=\frac{P(A)P(R\mid A)}{P(A)P(R\mid A)+P(A')P(R\mid A')}=\frac{\frac{2}{6}\times\frac{4}{12}}{\frac{2}{6}\times\frac{4}{12}+\frac{4}{6}\times\frac{3}{8}}=\frac{4}{13}$$

مثال ۳۳.۹.۲ شخصی به تصادف یکی از اعداد صحیح ۱، ۲ و ۳ را انتخاب می‌کند و سپس به تعداد عدد انتخاب شده یک تاس را پرتاب می‌کند.

الف- احتمال اینکه مجموع ۵ بیاورد را بیابید.

ب- اگر این شخص مجموع ۴ آورده باشد، احتمال اینکه عدد انتخاب شده ۲ باشد را بیابید.

حل قرار می‌دهیم

$B_i=$ پیشامد اینکه شخص عدد i را انتخاب کرده باشد $\quad i=1,2,3$

$A=$ پیشامد اینکه شخص مجموع ۵ بیاورد

$C=$ پیشامد اینکه شخص مجموع ۴ بیاورد

در این صورت

الف- داریم که

$$A\mid B_1=\{5\}$$
$$A\mid B_2=\{(1,4),(2,3),(3,2),(4,1)\}$$
$$A\mid B_3=\{(1,1,3),(1,3,1),(3,1,1),(1,2,2),(2,1,2),(2,2,1)\}$$

بنابراین

$$P(A)=\sum_{i=1}^{3}P(B_i)P(A\mid B_i)=\frac{1}{3}\times\frac{1}{6}+\frac{1}{3}\times\frac{4}{36}+\frac{1}{3}\times\frac{6}{216}=\frac{11}{108}$$

ب- داریم که

$$C\mid B_1=\{4\}$$
$$C\mid B_2=\{(1,3),(2,2),(3,1)\}$$
$$C\mid B_3=\{(1,1,2),(1,2,1),(2,1,1)\}$$

بنابراین از فرمول بیز داریم که

$$P(B_2\mid C)=\frac{P(B_2)P(C\mid B_2)}{\sum_{i=1}^{3}P(B_i)P(C\mid B_i)}=\frac{\frac{1}{3}\times\frac{3}{36}}{\frac{1}{3}\times\frac{1}{6}+\frac{1}{3}\times\frac{3}{36}+\frac{1}{3}\times\frac{3}{216}}=\frac{6}{19}$$

مثال ۳۴.۹.۲ یک ظرف شامل ۳ مهره قرمز و ۷ مهره سفید است. یک مهره از ظرف خارج می‌کنیم و هر رنگی که باشد مهره دیگری از رنگ مخالف آن در ظرف قرار می‌دهیم. سپس مهره دیگری از ظرف خارج می‌کنیم. اگر این مهره با مهره قبلی که به دست آمده بود هم رنگ باشد،

احتمال اینکه هر دو سفید باشند را بیابید.

حل پیشامدهای زیر را در نظر می‌گیریم

$R_i=$ پیشامد اینکه در بار iام مهره قرمز خارج شود $\quad i=1,2$

$W_i=$ پیشامد اینکه در بار iام مهره سفید خارج شود

$A=$ پیشامد اینکه هر دو مهره انتخاب شده هم رنگ باشند

در این صورت $W_1\cap W_2\subset A$ و $P(W_1\cap W_2\mid A)$ مورد سئوال است. بنابراین

$$P(W_1\cap W_2\mid A)=\frac{P(W_1\cap W_2\cap A)}{P(A)}=\frac{P(W_1\cap W_2)}{P(W_1\cap W_2)+P(R_1\cap R_2)}$$
$$=\frac{P(W_1)P(W_2\mid W_1)}{P(W_1)P(W_2\mid W_1)+P(R_1)P(R_2\mid R_1)}=\frac{\frac{7}{10}\times\frac{6}{10}}{\frac{7}{10}\times\frac{6}{10}+\frac{3}{10}\times\frac{2}{10}}=\frac{7}{8}$$

مثال ۳۵.۹.۲ جعبه A شامل ۳ مهره سفید و ۵ مهره قرمز و جعبه B شامل ۶ مهره سفید و ۲ مهره قرمز می‌باشد. یک جعبه را به تصادف انتخاب کرده و یک مهره از آن خارج می‌کنیم و بدون توجه به رنگ در جعبه سوم C قرار می‌دهیم که خود شامل ۳ مهره سفید و ۳ مهره قرمز است. سپس از جعبه C یک مهره خارج می‌کنیم. احتمال قرمز بودن این مهره را بیابید.

حل پیشامدهای زیر را در نظر می‌گیریم

$W_i=$ پیشامد اینکه از جعبه iام مهره سفید خارج شود $\quad i=A,B,C$

$R_i=$ پیشامد اینکه از جعبه iام مهره قرمز خارج شود

$A=$ پیشامد اینکه جعبه A در بار اول انتخاب شود

$B=$ پیشامد اینکه جعبه B در بار اول انتخاب شود

در این صورت $P(R_C)$ مورد سئوال است. بنابراین

$$P(R_C)=P(A)P(R_C\mid A)+P(B)P(R_C\mid B)=\frac{1}{2}P(R_C\mid A)+\frac{1}{2}P(R_C\mid B)$$

اما از طرفی داریم که

$$P(R_C\mid A)=P(R_A\cap R_C)+P(W_A\cap R_C)$$
$$=P(R_A)P(R_C\mid R_A)+P(W_A)P(R_C\mid W_A)=\frac{5}{8}\times\frac{4}{7}+\frac{3}{8}\times\frac{3}{7}=\frac{29}{56}$$
$$P(R_C\mid B)=P(R_B\cap R_C)+P(W_B\cap R_C)$$

$$= P(R_B)P(R_C \mid R_B) + P(W_B)P(R_C \mid W_B) = \frac{2}{8}\times\frac{4}{7}+\frac{6}{8}\times\frac{3}{7}=\frac{26}{56}$$

بنابراین

$$P(R_c)=\frac{1}{2}\times\frac{29}{56}+\frac{1}{2}\times\frac{26}{56}=\frac{55}{112}$$

مثال ۳۶.۹.۲ سه جعبه را در نظر بگیرید به طوری که جعبه اول شامل ۴ مهره سفید و ۲ مهره قرمز و جعبه دوم شامل ۴ مهره سفید و ۸ مهره قرمز و جعبه سوم شامل ۳ مهره سفید و ۱ مهره قرمز باشد. اگر یک مهره به تصادف از هر جعبه انتخاب شود، احتمال اینکه مهره انتخاب شده از جعبه اول سفید باشد به شرط آنکه دقیقاً ۲ مهره سفید انتخاب شده باشد را بیابید.

حل قرار می‌دهیم

$W_i =$ پیشامد اینکه از جعبه iام یک مهره سفید انتخاب شود $\qquad i=1,2,3$

$R_i =$ پیشامد اینکه از جعبه iام یک مهره قرمز انتخاب شود

$A =$ پیشامد اینکه دقیقاً ۲ مهره سفید انتخاب شود

در این صورت $P(W_1 \mid A)$ مورد سئوال است. چون انتخاب مهره از جعبه‌ها بطور مستقل انجام می‌گیرد، بنابراین

$$P(W_1 \mid A)=\frac{P(W_1\cap A)}{P(A)}=\frac{P(W_1\cap W_2\cap R_3)+P(W_1\cap R_2\cap W_3)}{P(W_1\cap W_2\cap R_3)+P(W_1\cap R_2\cap W_3)+P(R_1\cap W_2\cap W_3)}$$

$$=\frac{P(W_1)P(W_2)P(R_3)+P(W_1)P(R_2)P(W_3)}{P(W_1)P(W_2)P(R_3)+P(W_1)P(R_2)P(W_3)+P(R_1)P(W_2)P(W_3)}$$

$$=\frac{(\frac{4}{6}\times\frac{4}{12}\times\frac{1}{4})+(\frac{4}{6}\times\frac{8}{12}\times\frac{3}{4})}{(\frac{4}{6}\times\frac{4}{12}\times\frac{1}{4})+(\frac{4}{6}\times\frac{8}{12}\times\frac{3}{4})+(\frac{2}{6}\times\frac{4}{12}\times\frac{3}{4})}=\frac{112}{136}=\frac{14}{17}$$

مثال ۳۷.۹.۲ جعبه A شامل ۳ مهره سفید و ۴ مهره قرمز و جعبه B شامل ۳ مهره قرمز و ۲ مهره سفید می‌باشند. یکی از جعبه‌ها را به تصادف انتخاب کرده مهره‌ای به تصادف از آن خارج می‌کنیم و در جعبه بعدی قرار می‌دهیم. سپس از این جعبه ۲ مهره بدون جایگذاری خارج می‌کنیم. اگر این ۲ مهره سفید باشند احتمال اینکه مهره‌ای که منتقل شده است قرمز باشد را بیابید.

حل قرار می‌دهیم

$W_i =$ پیشامد اینکه در بار اول از جعبه iام مهره سفید خارج شود

$R_i =$ پیشامد اینکه در بار اول از جعبه iام مهره قرمز خارج شود $\qquad i=A,B$

$WW_i =$ پیشامد اینکه در بار دوم از جعبه iام دو مهره سفید خارج شود

$A =$ پیشامد اینکه در بار اول جعبه A انتخاب شده باشد

$B =$ پیشامد اینکه در بار اول جعبه B انتخاب شده باشد

$D =$ پیشامد مورد نظر مسئله

در این صورت

$$P(D)=P(A\cap D)+P(B\cap D)=P(A)P(D\mid A)+P(B)P(D\mid B)$$

اما از طرفی $P(A)=P(B)=\frac{1}{2}$ و

$$P(D\mid A)=P(R_A\mid WW_B)=\frac{P(R_A)P(WW_B\mid R_A)}{P(R_A)P(WW_B\mid R_A)+P(W_A)P(WW_B\mid W_A)}$$

$$=\frac{\frac{4}{7}\times\frac{\binom{2}{2}}{\binom{6}{2}}}{\frac{4}{7}\times\frac{\binom{2}{2}}{\binom{6}{2}}+\frac{3}{7}\times\frac{\binom{3}{2}}{\binom{6}{2}}}=\frac{4}{13}$$

$$P(D\mid B)=P(R_B\mid WW_A)=\frac{P(R_B)P(WW_A\mid R_B)}{P(R_B)P(WW_A\mid R_B)+P(W_B)P(WW_A\mid W_B)}$$

$$=\frac{\frac{3}{5}\times\frac{\binom{3}{2}}{\binom{8}{2}}}{\frac{3}{5}\times\frac{\binom{3}{2}}{\binom{8}{2}}+\frac{2}{5}\times\frac{\binom{4}{2}}{\binom{8}{2}}}=\frac{9}{21}$$

بنابراین

$$P(D)=\frac{1}{2}\times\frac{4}{13}+\frac{1}{2}\times\frac{9}{21}=\frac{201}{546}=\frac{67}{182}$$

۱۰.۲ تمرینات

۱ در هر یک از حالات زیر فضای نمونه آزمایش تصادفی را معین کنید.

الف- پرتاب یک سکه ۳ مرتبه

ب- پرتاب یک تاس ۳ مرتبه

ج- جواب دادن به یک آزمون تستی ۴ جوابه شامل ۲۰ سئوال

د- تعداد ذرات رادیواکتیو که از یک شمارنده در یک ساعت عبور می‌کنند

ه- یک تاس را پرتاب می‌کنیم اگر عدد زوج آمد تاس را دو مرتبه پرتاب می‌کنیم و در غیر اینصورت یک سکه را پرتاب می‌کنیم

و- زمان ملاقات دو نفر در بین ساعات یک تا دو بعد از ظهر

ز- انتخاب یک نقطه درون مربعی به ضلع ۴ سانتیمتر که اضلاع آن روی محورهای مختصات و در ناحیه اول باشد

ح- پرتاب متوالی یک تاس تا مشاهده یک عدد مضرب ۳.

۲ در آزمایشهای تمرین ۱ نوع فضای نمونه را از لحاظ متناهی یا نامتناهی شمارش‌پذیر و یا پیوسته معین کنید.

۳ در هر یک از قسمتهای تمرین ۱ پیشامدهای زیر را معین کنید.

الف- مشاهده حداکثر یک شیر

ب- مشاهده مجموع ۵ در پرتاب ۳ مرتبه تاس

ج- تنها ۵ سئوال اول صحیح پاسخ داده شوند

د- حداکثر ۲۰ ذره رادیواکتیو از شمارنده عبور کنند

ه- مشاهده شیر در پرتاب سکه

و- زمان ملاقات در بین ساعات ۱/۳۰ تا ۱/۴۵ باشد

ز- نقطه انتخابی درون مربعی به ضلع ۲ سانتیمتر و یک رأس روی مبدأ مختصات باشد

ح- حداقل ۵ پرتاب برای مشاهده عدد مضرب ۳ لازم باشد.

۴ سکه‌ای را چهار مرتبه پرتاب می‌کنیم. فرض کنید E پیشامد مشاهده شیر در بار اول و آخر و F پیشامد شیر در بار دوم و سوم باشد. پیشامدهای زیر را مشخص کنید

$$E \quad , \quad F \quad , \quad E\cup F \quad , \quad E\cap F \quad , \quad F-E$$

۵ ثابت کنید برای هر دو پیشامد E و F داریم که

$$(E\cap F)'=E'\cup F' \quad , \quad (E\cup F)'=E'\cap F'$$

این قوانین به قوانین دمرگان معروف هستند. این قوانین را به بیش از دو پیشامد تعمیم دهید.

۶ چهار کارت به شماره‌های ۱، ۲، ۳ و ۴ وجود دارند که شانس انتخاب شماره ۴ دو برابر انتخاب شماره ۲ و شانس انتخاب شماره‌های فرد ۳ برابر انتخاب شماره ۴ می‌باشد. مدل احتمال برای انتخاب یک کارت از بین این کارتها را معین کنید. احتمال اینکه عدد انتخاب شده مربع کامل باشد را بیابید.

۷ فرض کنید که یک تاس طوری ساخته شده باشد که در آن احتمال آمدن هر عدد متناسب با خود آن عدد باشد. مدل احتمال برای پرتاب یک بار این تاس را معین کنید و با استفاده از آن احتمال آمدن یک عدد زوج و احتمال آمدن یک عدد بزرگتر از ۴ را حساب کنید.

۸ فرض کنید که یک تاس طوری ساخته شده باشد که در آن احتمال آمدن هر عدد متناسب با عکس خود آن عدد باشد. مدل احتمال برای پرتاب یک بار این تاس را معین کنید و با استفاده از آن احتمال آمدن یک عدد فرد و احتمال آمدن یک عدد کمتر از ۴ را حساب کنید.

۹ با فرض $P(B)=0/7$ ، $P(A)=0/6$ و $S=A\cup B$ مطلوب است محاسبه

$$P(A\cap B) \quad , \quad P(B-A) \quad , \quad P(B'-A')$$

۱۰ فرض کنید A و B دو پیشامد ناسازگار باشند و $P(A)=0/3$ و $P(B)=0/4$ مطلوب است محاسبه

$$P(A\cup B) \quad , \quad P(A-B) \quad , \quad P(A'-B')$$

۱۱ احتمال اینکه به بیماری که به دکتر مراجعه کرده، قرص داده شود ۰/۳۷ و احتمال اینکه به وی آمپول داده شود ۰/۲۸ است و احتمال اینکه به وی هم قرص و هم آمپول داده شود ۰/۱۵ است. احتمال اینکه به وی قرص یا آمپول و یا هر دو داده شود را بیابید.

۱۲ فرض کنید A و B دو پیشامد باشند به طوری که $P(A)=\frac{3}{8}$، $P(B)=\frac{1}{2}$ و $P(A\cap B)=\frac{1}{4}$ مطلوب است محاسبه

$$P(A\cup B)\ ,\ P(A')\ ,\ P(B')\ ,\ P(A'\cap B')\ ,\ P(A'\cup B')\ ,\ P(A\cap B')\ ,\ P(A'\cap B)$$

۱۳ اگر پلاک اتومبیلها از ۵ شماره غیرصفر و یک حرف الفبای فارسی تشکیل شود، تعداد اتومبیلهایی را که می‌توان شماره کرد را تعیین کنید.

۱۴ به چند طریق می‌توان ۶ کتاب ریاضی مختلف و ۵ کتاب فیزیک مختلف را در یک قفسه قرار داد به طوری که کتابهای یک رشته در کنار هم قرار گیرند و کتابهای ریاضی سمت چپ کتابهای فیزیک

باشند؟

۱۵ به چند طریق می‌توان ۹ کتاب را در یک قفسه در کنار هم قرارداد به طوری که سه تای آنها همیشه پهلوی هم قرار گیرند؟

۱۶ چند عدد سه رقمی از ارقام ۱، ۲، ۴، ۶، ۸ و ۹ می‌توان نوشت به طوری که

الف – محدودیتی نداشته باشیم

ب – اعداد زوج باشند.

ج – اعداد مضرب ۵ باشند.

د – اعداد کوچکتر از ۴۰۰ باشند.

۱۷ دوازده زن و ده مرد در یک محل زندگی می‌کنند که از بین آنها سه زن و دو مرد با هم خواهر و برادر هستند. به چند طریق این عده می‌توانند با هم ازدواج کنند؟

۱۸ تعداد جایگشتهای مختلف حروف کلمه STATISTICS را به دست آورید.

۱۹ به چند طریق می‌توان ۱۲ نفر را به کمیته‌های ۵، ۳ و ۴ نفره تقسیم نمود؟

۲۰ از یک گروه شامل ۹ مرد و ۳ زن می‌خواهیم یک کمیته ۴ نفره تشکیل دهیم

الف – احتمال اینکه کمیته شامل فقط یک زن باشد را بیابید.

ب – احتمال اینکه کمیته شامل حداقل یک زن باشد را بیابید.

ج – احتمال اینکه کمیته شامل هم مرد و هم زن باشد را بیابید.

۲۱ ۱۲ بسته محتوی جوایز در دست داریم که قرار است به دانش آموزی یکی از این ۱۲ بسته به عنوان جایزه داده شود. دانش آموز نمی‌داند که در ۳ بسته کیف بغلی، در ۵ بسته دفترچه و در ۴ بسته خودکار وجود دارد.

الف – اگر دانش‌آموز یک بسته را به تصادف انتخاب کند، احتمال اینکه جایزه او کیف بغلی باشد را بیابید.

ب – اگر دانش‌آموز دو بسته را به تصادف انتخاب کند، احتمال اینکه جایزه او کیف بغلی باشد را بیابید.

۲۲ ده مهره غیر متمایز را به چند طریق می‌توانید در چهار جعبه متمایز قرار دهید در صورتی که بخواهیم در هر جعبه لااقل یک مهره قرار گیرد؟

۲۳ یک خانواده ۳ نفری پنجشنبه یا جمعه متولد شده‌اند. احتمال اینکه هم پنجشنبه و هم جمعه جشن تولد داشته باشند را بیابید.

۲۴ با فرض $r \leq n$ ثابت کنید که

$$\binom{n}{r} = \binom{n}{n-r} \quad , \quad \binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

۲۵ بیست اتومبیل در مسابقه‌ای شرکت می‌کنند. هشت مرسدس بنز، هفت پژو و بقیه شورلت می‌باشند. هر گاه فقط کارخانه سازنده اتومبیل در نظر گرفته شود، مطلوب است

الف – احتمال اینکه یکی از اتومبیلهای مرسدس بنز مقام اول را به دست آورد را بیابید.

ب – احتمال اینکه مقامهای اول و دوم توسط اتومبیلهای مرسدس بنز به دست آیند را بیابید.

۲۶ یک آسانسور از طبقه هم کف با ۸ نفر حرکت می‌کند، به طوری که در طبقه ششم هیچکس در آسانسور باقی نمی‌ماند. به چند طریق این افراد می‌توانسته‌اند از آسانسور پیاده شوند؟ اگر افراد شامل ۵ مرد و ۳ زن بوده باشند، جواب چیست؟

۲۷ فروشگاهی ۵۰ دستگاه ضبط صوت را از تولیدکننده‌ای دریافت می‌کند. چهار دستگاه به طور تصادفی انتخاب شده و آزمایش می‌شوند. اگر هر چهار دستگاه سالم باشند فروشگاه محموله را قبول می‌کند و در غیر اینصورت ۵۰ دستگاه را عودت می‌دهد. فرض کنید محموله حاوی ۵ دستگاه نقص‌دار است.

الف – احتمال اینکه محموله توسط فروشگاه پذیرفته شود را بیابید.

ب – احتمال اینکه دو دستگاه آزمایش شده اول دارای نقص و دو دستگاه بعدی بدون نقص باشند را بیابید.

۲۸ جعبه‌ای شامل ۳ مهره قرمز و ۷ مهره سیاه است. افراد A و B مهره‌ها را به ترتیب یکی یکی و بدون جایگذاری خارج می‌کنند تا اینکه یک مهره قرمز انتخاب شود. احتمال اینکه مهره قرمز توسط A انتخاب شود را بیابید.

۲۹ محموله‌ای شامل ۱۰ جفت کفش است. اگر ۸ لنگه کفش از این محموله به تصادف انتخاب شود، مطلوب است

الف – احتمال اینکه هیچکدام از این ۸ لنگه کفش جفت نباشند را بیابید.

ب – احتمال اینکه دقیقاً دو تا از این ۸ لنگه کفش جفت باشند را بیابید.

۳۰ اگر A_1، A_2، A_3، ... و A_n پیشامدهای دلخواهی باشند ثابت کنید که

$$P(\bigcup_{i=1}^{n} A_i)=\sum_{i=1}^{n} P(A_i)-\sum_{i=1}^{n-1}\sum_{j>i}^{n}P(A_i\cap A_j)+\sum_{i=1}^{n-2}\sum_{j>i}^{n-1}\sum_{k>j}^{n}P(A_i\cap A_j\cap A_k)$$
$$+\dots+(-1)^n P(A_1\cap A_2\cap\dots\cap A_n)$$

۳۱ صد مهره وجود دارد که از شماره ۱ تا ۱۰۰ شماره‌گذاری شده‌اند و صد جعبه وجود دارد که آنها نیز از شماره ۱ تا ۱۰۰ شماره‌گذاری شده‌اند. هر مهره را به تصادف تنها در یکی از این جعبه‌ها قرار می‌دهیم به طوری که هر جعبه تنها شامل یک مهره شود. احتمال اینکه حداقل شماره یک مهره با شماره جعبه آن یکی باشد را بیابید.

۳۲ شما به طرف هدفی آنقدر تیراندازی می‌کنید تا به هدف برخورد کند. فرض کنید که در هر تیراندازی احتمال اینکه شما به هدف بزنید ۰/۹ است و تیراندازیها از یکدیگر مستقل هستند. احتمال این را به دست آورید که

الف– بیش از دو تیراندازی لازم باشد.

ب– تعداد تیراندازیهای لازم مضربی از ۳ باشد.

۳۳ حسن امتحان کتبی گواهی نامه رانندگی را آنقدر می‌دهد تا پذیرفته شود. فرض کنید که احتمال قبول شدنش در هر بار که امتحان می‌دهد ۰/۱ باشد و امتحان‌ها از یکدیگر مستقل باشند. احتمال این را به دست آورید که

الف– حداقل ۴ بار امتحان لازم باشد.

ب– حداکثر ۱۰ بار امتحان لازم باشد.

۳۴ نقطه‌ای به تصادف بین ۰ و ۱ روی محور xها در صفحه (x,y) انتخاب می‌کنیم سپس دایره‌ای به مرکز مبدأ مختصات رسم می‌کنیم که بر آن نقطه بگذرد. احتمال اینکه مساحت دایره رسم شده کمتر از $\frac{\pi}{2}$ باشد را بیابید.

۳۵ یک خط کش ۳۰ سانتی متری به تصادف در یک نقطه در امتداد طولش به دو قطعه شکسته شده است. احتمال اینکه قطعه طویلتر لااقل دو برابر قطعه کوتاهتر باشد را بیابید.

۳۶ فرض کنید که از سه پیشامد A، B و C پیشامدهای A و B مستقل هستند و پیشامدهای B و C ناسازگارند و $P(A)=0/5$، $P(B)=0/3$ و $P(C)=0/1$. پیشامدهای زیر را به صورت

مجموعه‌ها بنویسید و احتمال هر یک از آنها را محاسبه کنید.

الف– B و C رخ دهند.

ب– حداقل یکی از دو پیشامد A و B رخ دهد

ج– B رخ ندهد.

د– هر سه پیشامد رخ دهند.

۳۷ از دانشجویان سال اول دانشکده‌ای، ۲۵٪ از درس ریاضی، ۱۵٪ از درس فیزیک و ۱۰٪ از هر دو درس ریاضی و فیزیک مردود شده‌اند. دانشجویی را به تصادف انتخاب می‌کنیم. مطلوب است محاسبه احتمال اینکه

الف– اگر از درس فیزیک مردود شده است، از درس ریاضی نیز مردود شده باشد.

ب– اگر از درس ریاضی مردود شده است، از درس فیزیک نیز مردود شده باشد.

۳۸ احتمال به صدا در آمدن هر یک از سه آژیر خطر مستقلی که در یک فروشگاه نصب شده‌اند به هنگام بروز آتش سوزی برابر ۰/۹۵ است.

الف– احتمال اینکه هر سه آژیر خطر به هنگام آتش سوزی به صدا در آیند را بیابید.

ب– احتمال اینکه حداقل یکی از ۳ آژیر خطر به صدا درآید را بیابید.

۳۹ از ظرفی که دارای ۶ مهره سیاه و ۴ مهره سفید است، ۳ مهره را یکی یکی و بدون جایگذاری انتخاب می‌کنیم. احتمال اینکه هر سه مهره از یک رنگ باشند را بیابید. احتمال اینکه از هر دو رنگ مشاهده شود را بیابید.

۴۰ در ظرفی ۳ مهره سفید و ۵ مهره سیاه داریم. از این ظرف دو مهره یک به یک و بدون جایگذاری بیرون می‌آوریم. نشان دهید احتمال اینکه مهره اول سفید باشد برابر است با احتمال اینکه مهره دوم سفید باشد.

۴۱ جعبه‌ای شامل ۶ مهره سفید و ۹ مهره سیاه است. اگر ۴ مهره یک به یک و بدون جایگذاری از این جعبه خارج کنیم، احتمال اینکه ۲ مهره انتخابی اول سفید و ۲ مهره انتخابی آخر سیاه باشند را بیابید.

۴۲ یک ظرف دارای ۵ مهره سفید و ۷ مهره سیاه است. در هر بار یک مهره به تصادف از ظرف انتخاب می‌کنیم و آن را با ۲ مهره همرنگ خودش به ظرف باز می‌گردانیم. احتمال اینکه دو مهره انتخابی اول سیاه و دو مهره انتخابی بعدی سفید باشند را بیابید.

۴۳ در تمرین ۳۹ اگر مهره‌ها را یکی یکی و با جایگذاری انتخابی کنیم. احتمالات مورد نظر را به

دست آورید.

۴۴ یک تولید کننده هواکش‌های الکتریکی، موتورهای مورد نیاز خود را از دو شرکت تهیه می‌کند. ۷۵٪ از موتورها از شرکت A و ۲۵٪ بقیه از شرکت B خریداری می‌شود. فرض کنید که ۵٪ از موتورهای شرکت A و ۳٪ از موتورهای شرکت B خراب باشند. این تولید کننده یک موتور خراب در بین موتورها پیدا کرده است، چقدر احتمال دارد که این موتور از شرکت B خریداری شده باشد؟

۴۵ دو جعبه را در نظر بگیرید که جعبه اول شامل ۲ توپ سفید و ۳ توپ سیاه و جعبه دوم شامل ۳ توپ سفیدو ۲ توپ سیاه می‌باشند. یکی از جعبه‌ها را به تصادف انتخاب کرده و از آن ۲ توپ به تصادف خارج می‌کنیم

الف – احتمال اینکه ۲ توپ خارج شده سیاه باشند را بیابید.

ب – اگر ۲ توپ خارج شده سفید باشند، احتمال اینکه جعبه دوم انتخاب شده باشد را بیابید.

۴۶ شرکتی دارای دو کارخانه است که دوربین عکاسی تولید می‌کنند. ۲۰٪ تولیدات کارخانه A و ۵٪ تولیدات کارخانه B معیوب هستند. تولیدات A در هفته دو برابر تولیدات B است.

الف – احتمال اینکه دوربینی که به طور تصادفی از میان دوربین‌های تولید شده انتخاب می‌شود، سالم باشد را بیابید.

ب – اگر دوربین انتخاب شده خراب باشد، احتمال اینکه از تولیدات کارخانه A باشد را بیابید.

۴۷ در کارخانه‌ای کارگران در سه شیفت صبح، عصر و شب کار می‌کنند. آمار نشان می‌دهد که به ترتیب ۴۰٪، ۴۰٪ و ۲۰٪ تولیدات توسط شیفت‌های صبح، عصر و شب تولید می‌شود و به ترتیب ۵٪، ۱۰٪ و ۲۰٪ از تولیدات شیفت‌های صبح، عصر و شب معیوب هستند. اگر کالای معیوبی را به طور تصادفی انتخاب کنیم، احتمال اینکه توسط شیفت صبح تولید شده باشد را بیابید.

۴۸ فرض کنید ۵ مرد از ۱۰۰ مرد و ۲۵ زن از ۱۰۰ زن یک شهر که تعداد زنها و مردهای آن شهر یکسان هستند، کوررنگ باشند. از این شهر یک نفر را به تصادف انتخاب می‌کنیم. چنانچه شخص انتخاب شده کوررنگ باشد، احتمال اینکه مرد باشد را بیابید.

۴۹ در یک دانشگاه وزن ۴٪ مردان و ۱٪ زنان از ۶۰ کیلوگرم بیشتر است و می‌دانیم ۶۰٪ از دانشجویان زن می‌باشند. دانشجویی را به تصادف انتخاب کرده و وزن او از ۶۰ کیلوگرم بیشتر

است. مطلوب است محاسبه احتمال اینکه این دانشجو زن باشد.

۵۰ کیسه‌ای حاوی یک سکه ۵ ریالی، یک سکه ۱۰ ریالی و یک سکه ۲۰ ریالی است. با توجه به اینکه ۵ ریالی سکه‌ای تقلبی با دو روی شیر است، از این کیسه سکه‌ای به تصادف انتخاب کرده و آن را چهار بار متوالی پرتاب می‌کنیم. اگر نتیجه چهار شیر باشد، احتمال اینکه سکه پرتاب شده همان سکه تقلبی باشد را بیابید.

۵۱ ظرف A شامل ۳ مهره قرمز و ۵ مهره سفید و ظرف B شامل ۲ مهره قرمز و ۱ مهره سفید و ظرف C شامل ۲ مهره قرمز و ۳ مهره سفید می‌باشند. یک ظرف به تصادف انتخاب می‌کنیم و دو مهره از آن به تصادف انتخاب می‌کنیم. اگر دو مهره انتخابی یکی قرمز و دیگری سفید باشند، احتمال اینکه از ظرف A انتخاب شده باشند را بیابید.

۵۲ دوجعبه وجود دارند که جعبه اول شامل ۳ مهره سفید و ۲ مهره سیاه است و جعبه دوم شامل ۲ مهره سفید و ۳ مهره سیاه است. از جعبه اول ۲ مهره به تصادف انتخاب کرده و در داخل جعبه دوم می‌اندازیم و سپس از جعبه دوم ۳ مهره به تصادف و بدون جایگذاری خارج می‌کنیم.

الف – احتمال اینکه ۳ مهره خارج شده از جعبه دوم، ۲ سفید و ۱ سیاه باشند را بیابید.

ب – اگر هر ۳ مهره خارج شده از جعبه دوم سیاه باشند، احتمال اینکه ۲ مهره خارج شده از جعبه اول سیاه باشند را بیابید.

۵۳ فرض کنید یک موسسه بیمه افراد جامعه را به سه گروه افراد دارای ریسک بالا، متوسط و پائین تقسیم نموده است که به ترتیب ۲۰، ۵۰ و ۳۰ درصد جامعه را تشکیل می‌دهند. اطلاعات این موسسه نشان می‌دهد که احتمال تصادف کردن این گروهها در طول سال به ترتیب ۳۰، ۱۵ و ۵ درصد می‌باشد.

الف – اگر از هر یک از این گروهها یک نفر به تصادف انتخاب شوند احتمال اینکه هیچکدام در طول سال تصادفی نداشته باشند را بیابید

ب – اگر از افراد بیمه شده یک نفر انتخاب شود احتمال اینکه در طول سال تصادفی داشته باشد را بیابید.

ج – اگر از افراد بیمه شده یک نفر در طول سال تصادفی نداشته باشد، احتمال اینکه این شخص از گروه با ریسک متوسط باشد را بیابید.

۵۴ سه سکه وجود دارد که سکه اول یک سکه سالم است، سکه دوم دو شیری (هر دو طرف آن شیر) است و سکه سوم شانس مشاهده شیر در آن سه برابر خط است. یکی از سکه‌ها را به تصادف انتخاب می‌کنیم.

الف- احتمال اینکه در پرتاب یک مرتبه این سکه شیر مشاهده شود را بیابید.

ب- اگر در پرتاب دو مرتبه این سکه دو شیر مشاهده شود، احتمال اینکه سکه سالم انتخاب شده باشد را بیابید.

۵۵ فروشگاهی دارای سه شعبه A، B و C است که به ترتیب ۵۰، ۷۵ و ۱۰۰ کارمند دارند و به ترتیب ۵۰، ۶۰ و ۷۰ درصد کارمندان آن شعب زن هستند. فرض کنید استعفا در میان تمام کارمندان شانس یکسان داشته باشند.

الف- اگر از هر شعبه یک نفر استعفا دهد، احتمال آنکه هر سه نفر زن باشند را بیابید.

ب- اگر از این فروشگاه یک نفر استعفا دهد، احتمال آنکه زن باشد را بیابید.

ج- اگر یک زن از این فروشگاه استعفا دهد، احتمال آنکه از شعبه C باشد را بیابید.

۵۶ یک دستگاه حقیقت سنج به مظنون وصل می‌گردد و می‌دانیم اگر شخص گناهکار باشد با احتمال ۹۰٪ دستگاه در او اثر می‌کند و اگر بی‌گناه باشد ۹۹٪ در او اثر دارد. اگر یک مظنون از یک گروه مظنون که فقط ۰/۰۵ آنها تاکنون جنایی مرتکب شده‌اند انتخاب گردد و دستگاه نشان دهد که او گناهکار است، احتمال اینکه او بی گناه باشد را بیابید.

۵۷ جعبه A محتوی ۳ توپ قرمز و ۲ توپ سفید و جعبه B محتوی ۲ توپ قرمز و ۵ توپ سفید است. جعبه‌ای را به تصادف انتخاب و توپی از آن خارج می‌کنیم و پس از مشاهده رنگ آن، توپ را در جعبه دیگری می‌اندازیم و سپس توپی از این جعبه به تصادف خارج می‌کنیم. احتمال اینکه هر دو توپ خارج شده هم رنگ باشند را بیابید.

۵۸ در ظرف A، ۵ مهره قرمز و ۵ مهره سیاه و در ظرف B، ۴ مهره قرمز و ۸ مهره سیاه و در ظرف C، ۳ مهره قرمز و ۶ مهره سیاه وجود دارند. یک مهره از A خارج می‌کنیم و آن را در B قرار می‌دهیم و سپس یک مهره از B خارج می‌کنیم و در C قرار می‌دهیم. حال اگر یک مهره از C خارج کنیم احتمال قرمز بودن آن را بیابید.

فصل سوم

متغیرهای تصادفی

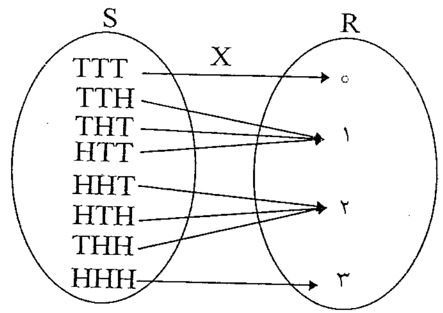
۱.۳ مفهوم متغیر تصادفی

در اغلب آزمایشهای تصادفی ما به جای نتایج حاصل از آزمایش به توابعی از نتایج علاقه‌مند هستیم. چنین توابعی که روی فضای نمونه تعریف می‌شوند به متغیر تصادفی موسوم‌اند و برای پی بردن به مفهوم آنها ابتدا به مثال زیر توجه کنید.

مثال ۱.۱.۳ آزمایش پرتاب یک سکه ۳ مرتبه را در نظر بگیرید. فضای نمونه حاصل از این آزمایش عبارت است از

$$S=\{TTT, TTH, THT, HTT, HHT, HTH, THH, HHH\}$$

فرض کنید در این آزمایش به تعداد شیرهای مشاهده شده به شما امتیاز داده شود. بنابراین برای شما فقط شمارش تعداد شیرها مورد نظر است که این تعداد می‌تواند یکی از مقادیر ۰، ۱، ۲ و ۳ باشد. به عبارت دیگر ما به جزئیات مربوط به فضای نمونه علاقه‌مند نیستیم بلکه فقط به یک توصیف عددی از نتیجه علاقه‌مندیم. برای این منظور به هر یک از نقاط فضای نمونه عددی حقیقی را نسبت می‌دهیم و این عمل را بوسیله یک تابع حقیقی که آن را متغیر تصادفی گویند انجام می‌دهیم. مثلاً اگر متغیر تصادفی (تابع) X برابر تعداد شیرهای مشاهده شده در ۳ مرتبه پرتاب سکه باشد، در این صورت



$X : S \to R$

$X(THT)=۱$

$X(HTH)=۲$

$\vdots \qquad \vdots$

تعریف ۱.۳ یک متغیر تصادفی تابعی از فضای نمونه به مجموعه اعداد حقیقی است به طوری که به هر نقطه از فضای نمونه یک عدد حقیقی را نسبت می‌دهد. برای نمایش متغیر تصادفی از یکی از حروف بزرگ مانند X، Y و... استفاده می‌شود و برای نمایش یکی از مقادیری که متغیر تصادفی اختیار می‌کند از حروف کوچک معادل آن یعنی x، y و... استفاده می‌شود. مجموعه مقادیر و یا برد متغیر تصادفی X را با S_X نمایش می‌دهند و آن را تکیه گاه X گویند.

با استفاده از متغیرهای تصادفی می‌توان کلیه مباحث احتمال که در فصل قبل بیان شد را به نحو ساده‌تری بیان کرد و این مباحث را نیز تعمیم داد. برای مثال در مثال ۱.۱.۳ در صورتی که گفته شود متغیر تصادفی X دارای مقدار ۲ است، که آن را با مجموعه $\{\omega \in S \mid X(\omega)=۲\}$ و یا به طور ساده‌تر با نماد (X=۲) نمایش می‌دهند، به این معنی است که یکی از اعضای پیشامد $A=\{HHT, HTH, THH\}$ را مشاهده کرده‌ایم و بنابراین (X=۲) یک پیشامد است و می‌توان احتمال آن را به صورت زیر محاسبه کرد، $P(X=۲)=P(A)=\frac{۳}{۸}$. همچنین $(X\le ۱)$ معادل پیشامد $B=\{TTT, TTH, THT, HTT\}$ است و بنابراین $P(X\le ۱)=P(B)=\frac{۴}{۸}$ و یا $(X\in(۱/۵,۲/۵])=(X=۲)\equiv A$ بنابراین $P(X\in(۱/۵,۲/۵])=P(X=۲)=\frac{۳}{۸}$.

توجه توجه کنید که برای هر زیر مجموعه C از اعداد حقیقی منظور از $(X\in C)$ پیشامد این است که مقادیر متغیر تصادفی X در مجموعه C قرار گیرند.

در زیر چند مثال از متغیرهای تصادفی می‌آوریم.

مثال ۲.۱.۳ سه مهره به شماره‌های ۱، ۲ و ۳ را در ۳ جعبه به شماره‌های ۱، ۲ و ۳ به طور تصادفی می‌ریزیم به گونه‌ای که هر جعبه تنها شامل یک مهره باشد. اگر متغیر تصادفی X برابر تعداد جورها

(تعداد مهره‌هایی که در جعبه با شماره متناظر خودشان قرار گرفته‌اند) باشد، مطلوب است

الف- احتمال اینکه دقیقاً یک جور داشته باشیم را بیابید.

ب- احتمال اینکه حداقل دو جور داشته باشیم را بیابید.

حل فهرست نقاط فضای نمونه که عبارت از قرار گرفتن سه مهره ۱، ۲ و ۳ در جعبه‌ها می‌باشد و مقادیری که متغیر تصادفی X به نقاط نسبت می‌دهد عبارت است از

S	۱۲۳	۱۳۲	۲۱۳	۲۳۱	۳۱۲	۳۲۱
x	۳	۱	۱	۰	۰	۱

بنابراین $S_X=\{۰,۱,۳\}$ و در نتیجه

الف- $P(X=۱)$ مورد سؤال است پس $P(X=۱)=P(\{۱۳۲,۲۱۳,۳۲۱\})=\frac{۳}{۶}=\frac{۱}{۲}$

ب- $P(X\ge ۲)$ مورد سؤال است پس $P(X\ge ۲)=P(X=۳)=P(\{۱۲۳\})=\frac{۱}{۶}$

مثال ۳.۱.۳ سکه‌ای که شانس مشاهده شیر در آن دو برابر خط است را آنقدر پرتاب می‌کنیم تا یک خط مشاهده کنیم و سپس توقف می‌کنیم. اگر متغیر تصادفی M برابر تعداد پرتابهای لازم تا رسیدن به یک خط باشد، احتمال اینکه حداقل ۴ پرتاب لازم باشد را بیابید.

حل فهرست نقاط فضای نمونه، مقادیری که متغیر تصادفی M به این نقاط نسبت می‌دهد و احتمالات مربوطه عبارت است از

S	T	HT	HHT	$HHHT$	....
M	۱	۲	۳	۴	....
احتمالات	$\frac{۱}{۳}$	$(\frac{۲}{۳})(\frac{۱}{۳})$	$(\frac{۲}{۳})^۲(\frac{۱}{۳})$	$(\frac{۲}{۳})^۳(\frac{۱}{۳})$	....

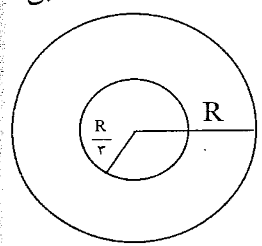
در اینجا $S_M=\{۱,۲,۳,۴,...\}$ و $P(M\ge ۴)$ مورد سؤال است بنابراین

$$P(M\ge ۴)=P(M=۴)+P(M=۵)+...$$

$$=(\frac{۲}{۳})^۳(\frac{۱}{۳})+(\frac{۲}{۳})^۴(\frac{۱}{۳})+...=\frac{(\frac{۲}{۳})^۳(\frac{۱}{۳})}{۱-\frac{۲}{۳}}=(\frac{۲}{۳})^۳=\frac{۸}{۲۷}$$

مثال ۴.۱.۳ از داخل دایره‌ای به شعاع R نقطه‌ای به تصادف انتخاب می‌کنیم و متغیر تصادفی Y

را برابر فاصله نقطه انتخابی تا مرکز دایره در نظر می‌گیریم. احتمال اینکه فاصله نقطه انتخابی تا مرکز دایره کمتر از $\frac{1}{3}$ شعاع دایره باشد را بیابید.



حل در اینجا $S_Y=[0,R]$ و $P(Y<\frac{R}{3})$ مورد سئوال است. بنابراین

$$P(Y<\frac{R}{3})=\frac{\text{مساحت دایره به شعاع } \frac{R}{3}}{\text{مساحت دایره به شعاع } R}=\frac{\pi\left(\frac{R}{3}\right)^2}{\pi R^2}=\frac{1}{9}$$

متغیرهای تصادفی گسسته و پیوسته متغیر تصادفی که مجموعه مقادیر آن متناهی یا نامتناهی شمارش‌پذیر باشد را متغیر تصادفی گسسته و متغیر تصادفی که مجموعه مقادیر آن یک فاصله عددی یا اجتماع چند فاصله عددی باشد را متغیر تصادفی پیوسته می‌نامند. برای مثال، متغیرهای تصادفی مثالهای ۲.۱.۳ و ۳.۱.۳ از نوع گسسته و متغیر تصادفی مثال ۴.۱.۳ از نوع پیوسته می‌باشد.

در بخشهای بعد توزیع احتمال هر یک از متغیرهای تصادفی گسسته و پیوسته را می‌آوریم.

۲.۳ توزیع احتمالات گسسته

در مثال ۲.۱.۳ چون $(X=x)$ یک پیشامد است پس می‌توان احتمال این پیشامد را محاسبه کرد. مثلاً احتمال پیشامد $(X=0)$ برابر احتمال پیشامد $A=\{231, 312\}$ است که برابر $\frac{2}{6}$ است یعنی $P(X=2)=P(A)=\frac{2}{6}$. با به دست آوردن احتمالات دیگر مربوط به نقاط ۱ و ۳ از مجموعه مقادیر متغیر تصادفی X یعنی $S_X=\{0,1,3\}$، جدول زیر که به جدول توزیع احتمال متغیر تصادفی گسسته X موسوم است را به دست می‌آوریم

x	0	1	3
$P(X=x)$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{1}{6}$

جدول فوق یک تابع از S_X به اعداد حقیقی در فاصله $[0,1]$ برقرار می‌کند که آن را با تابع $f_X(x)$ نمایش داده و به آن **تابع احتمال متغیر تصادفی X** گویند، بنابراین

$$f_X(x)=P(X=x)$$

این تابع یک تابع غیر منفی است و مجموع آن روی کلیه مقادیری که می‌تواند اختیار کند برابر ۱

است.

تعریف ۲.۳ تابع $f_X(x)=P(X=x)$ را تابع احتمال متغیر تصادفی X گویند هر گاه

۱ـ برای هر $x\in R$ داشته باشیم که $f_X(x)\geq 0$

۲ـ $\sum_{x\in R} f_X(x)=1$

برای محاسبه احتمال پیشامد $(X\in C)$ که C زیر مجموعه‌ای از اعداد حقیقی است، به صورت زیر عمل می‌کنیم.

$$\boxed{P(X\in C)=\sum_{x\in C} f_X(x)} \qquad (1.3)$$

مثال ۱.۲.۳ دو تاس را با هم پرتاب می‌کنیم و متغیر تصادفی X را برابر مجموع اعداد روی دو تاس مشاهده شده در نظر می‌گیریم.

الف- تابع احتمال X را به دست آورید.

ب- احتمال اینکه مجموع اعداد روی دو تاس حداکثر ۴ شود را بیابید.

ج- احتمال اینکه مجموع اعداد روی دو تاس بین ۶ و ۹ شود را بیابید.

حل الف- تکیه گاه X برابر $S_X=\{2,3,4,\ldots,12\}$ است و برای مثال

$P(X=4)=P(\{(1,3),(2,2),(3,1)\})=\frac{3}{36}$

با محاسبه احتمالات مربوط به نقاط دیگر S_X، تابع احتمال X به صورت زیر به دست می‌آید.

x	2	3	4	5	6	7	8	9	10	11	12
$f_X(x)=P(X=x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

ب- $P(X\leq 4)$ مورد سئوال است بنابراین از فرمول (۱.۳) داریم که

$$P(X\leq 4)=f_X(2)+f_X(3)+f_X(4)=\frac{1}{36}+\frac{2}{36}+\frac{3}{36}=\frac{1}{6}$$

ج- با استفاده از فرمول (۱.۳) داریم که

$$P(6<X<9)=f_X(7)+f_X(8)=\frac{6}{36}+\frac{5}{36}=\frac{11}{36}$$

مثال ۲.۲.۳ تابع زیر را در نظر بگیرید.

$$f_Y(y)=k\left(\frac{1}{6}\right)^y \qquad y=0,1,2,\ldots$$

الف– مقدار k را چنان تعیین کنید که $f_Y(y)$ تابع احتمال متغیر تصادفی گسسته Y باشد.

ب– احتمالات زیر را محاسبه کنید $\quad P(Y\le\frac{5}{2}) \quad , \quad P(Y\ge\frac{11}{3})$

حل الف– با توجه به تعریف ۲.۳ بایستی $k\ge 0$ و همچنین

$$1=\sum_{y=0}^{+\infty}f_Y(y)=\sum_{y=0}^{+\infty}k\left(\frac{1}{6}\right)^y=k\left[1+\frac{1}{6}+\left(\frac{1}{6}\right)^2+\ldots\right]=k\left[\frac{1}{1-\frac{1}{6}}\right]=k\left(\frac{6}{5}\right)$$

بنابراین بایستی $k=\frac{5}{6}$ باشد.

ب– با استفاده از فرمول (۱.۳) داریم که

$$P(Y\le\frac{5}{2})=f_Y(0)+f_Y(1)+f_Y(2)=\frac{5}{6}\left[1+\frac{1}{6}+\frac{1}{36}\right]=\frac{5}{6}\times\frac{43}{36}=\frac{215}{216}$$

$$P(Y\ge\frac{11}{3})=\sum_{y=4}^{+\infty}f_Y(y)=\sum_{y=4}^{+\infty}\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^y=\frac{\frac{5}{6}\left(\frac{1}{6}\right)^4}{1-\frac{1}{6}}=\left(\frac{1}{6}\right)^4=\frac{1}{1296}$$

مثال ۳.۲.۳ فرمولی برای تابع احتمال تعداد شیر وقتی سکه‌ای را ۴ مرتبه پرتاب می‌کنیم به دست آورید.

حل در اینجا فضای نمونه دارای $n(S)=2^4$ عضو هم شانس است. اگر متغیر تصادفی X را برابر تعداد شیرهای مشاهده شده در ۴ مرتبه پرتاب سکه در نظر بگیریم در این صورت $S_X=\{0,1,2,3,4\}$ و $P(X=x)$ به این معنی است که در ۴ مرتبه پرتاب سکه می‌خواهیم x شیر مشاهده کنیم که تعداد حالات مساعد آن برابر $\binom{4}{x}$ است. بنابراین

$$f_X(x)=P(X=x)=\frac{\binom{4}{x}}{2^4} \qquad x=0,1,2,3,4$$

تابع توزیع (تجمعی)

تعریف ۳.۳ اگر X یک متغیر تصادفی گسسته با تابع احتمال $f_X(x)$ باشد آنگاه تابع توزیع (تجمعی) X که با نماد $F_X(x)$ نمایش داده می‌شود، برای هر $x\in R$ به صورت زیر تعریف می‌شود

$$\boxed{F_X(x)=P(X\le x)=\sum_{t\le x}f_X(t)} \qquad (۲.۳)$$

مثال ۴.۲.۳ در مثال ۲.۱.۳ تابع توزیع متغیر تصادفی X را به دست آورید.

حل در قسمت قبل دیدیم که تابع احتمال متغیر تصادفی مثال ۲.۱.۳ عبارت است از

x	۰	۱	۳
$f_X(x)$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{1}{6}$

برای محاسبه تابع توزیع ابتدا آن را در چند نقطه دلخواه محاسبه می‌کنیم.

$$F_X(0.5)=P(X\le 0.5)=\sum_{t\le 0.5}f_X(t)=f_X(0)=\frac{2}{6}$$

$$F_X(2.3)=P(X\le 2.3)=\sum_{t\le 2.3}f_X(t)=f_X(0)+f_X(1)=\frac{2}{6}+\frac{3}{6}=\frac{5}{6}$$

به راحتی دیده می‌شود که برای هر $x<0$ ، $F_X(x)=0$ و برای هر $0\le x<1$ ، $F_X(x)=\frac{2}{6}$ و....

بنابراین با انجام محاسبات تابع توزیع X به صورت زیر به دست می‌آید.

$$F_X(x)=P(X\le x)=\begin{cases}0 & x<0\\ \frac{2}{6} & 0\le x<1\\ \frac{5}{6} & 1\le x<3\\ 1 & 3\le x\end{cases}$$

مثال ۵.۲.۳ در مثال ۳.۲.۳ تابع توزیع متغیر تصادفی X را به دست آورده و تابع احتمال و تابع توزیع را رسم کنید.

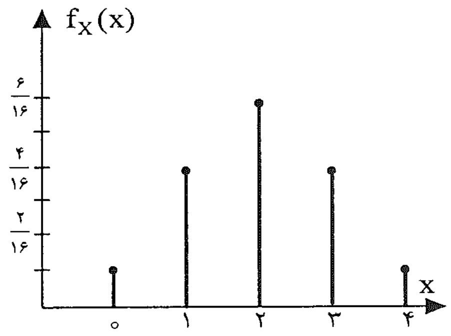
حل فرم جدولی تابع احتمال مثال ۳.۲.۳ به صورت زیر می‌باشد.

x	۰	۱	۲	۳	۴
$f_X(x)=P(X=x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

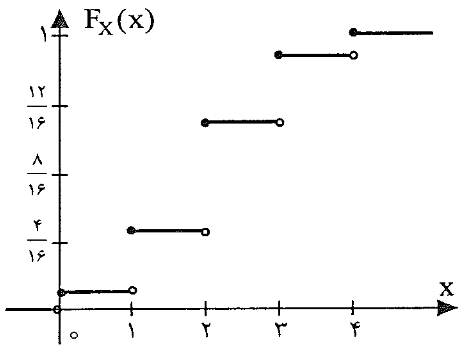
بنابراین با انجام عملیات مشابه مثال قبل تابع توزیع متغیر تصادفی X به صورت زیر به دست می‌آید.

$$F_X(x)=P(X\leq x)=\begin{cases} \circ & x<\circ \\ \frac{۱}{۱۶} & \circ\leq x<۱ \\ \frac{۵}{۱۶} & ۱\leq x<۲ \\ \frac{۱۱}{۱۶} & ۲\leq x<۳ \\ \frac{۱۵}{۱۶} & ۳\leq x<۴ \\ ۱ & ۴\leq x \end{cases}$$

معمولاً نمودار تابع احتمال را به صورت یک نمودار میله‌ای رسم می‌کنند و نمودار تابع توزیع یک متغیر تصادفی گسسته یک تابع پله‌ای است که این دو نمودار در زیر رسم شده‌اند.



نمودار میله‌ای تابع احتمال مثال ۵.۲.۳



نمودار پله‌ای تابع توزیع مثال ۵.۲.۳

خواص تابع توزیع با توجه به تعریف تابع توزیع و همچنین با توجه به نمودار بالا خواص زیر را در مورد تابع توزیع می‌توان بیان کرد.

الف- برای هر $x\in R$ ، $\circ\leq F_X(x)\leq ۱$.

ب- تابع توزیع یک تابع غیرنزولی است یعنی

$$x_۱<x_۲ \Rightarrow F_X(x_۱)\leq F_X(x_۲)$$

ج-
$$F_X(-\infty)=\lim_{x\to-\infty}F_X(x)=\circ$$
$$F_X(+\infty)=\lim_{x\to+\infty}F_X(x)=۱$$

د- تابع توزیع همواره از سمت راست پیوسته است یعنی

$$\lim_{x\to a^+}F_X(x)=F_X(a)$$

توجه کنید که از روی تابع توزیع یک متغیر تصادفی گسسته می‌توان تابع احتمال آن را توسط فرمول زیر به دست آورد.

$$\boxed{f_X(x)=F_X(x)-F_X(x^-)} \qquad (۳.۳)$$

که در آن $F_X(x^-)$ حد سمت چپ تابع توزیع در نقطه x است یعنی

$$F_X(x^-)=\lim_{t\to x^-}F_X(t)$$

مثلاً در مثال ۵.۲.۳ داریم که

$$f_X(۲)=F_X(۲)-F_X(۲^-)=\frac{۱۱}{۱۶}-\frac{۵}{۱۶}=\frac{۶}{۱۶}$$

و همچنین

$$f_X(۱/۵)=F_X(۱/۵)-F_X(۱/۵^-)=\frac{۵}{۱۶}-\frac{۵}{۱۶}=\circ$$

همچنین توجه کنید که هر نوع احتمالی را می‌توان توسط تابع توزیع از فرمولهای زیر محاسبه کرد:

۱- $P(a<X\leq b)=F_X(b)-F_X(a)$

۲- $P(a<X<b)=F_X(b^-)-F_X(a)$

۳- $P(a\leq X<b)=F_X(b^-)-F_X(a^-)$

۴- $P(a\leq X\leq b)=F_X(b)-F_X(a^-)$

۵- $P(X>a)=۱-F_X(a)$

۶- $P(X=b)=F_X(b)-F_X(b^-)$

مثال ۶.۲.۳ یک کلاس آمار ۸ شاگرد دارد که ۵ نفر آنها ۱۹ ساله و ۳ نفر آنها ۲۱ ساله هستند. از این کلاس ۲ شاگرد به تصادف و بدون جایگذاری انتخاب می‌کنیم و متغیر تصادفی X را برابر میانگین سن ۲ شاگرد انتخابی در نظر می‌گیریم. تابع احتمال و تابع توزیع متغیر تصادفی X را به دست آورده و $P(۱۹<X<۲۱)$ را محاسبه کنید.

حل در اینجا $S_X=\{۱۹,۲۰,۲۱\}$ و

$$f_X(۱۹)=P(X=۱۹)=\frac{\binom{۵}{۲}}{\binom{۸}{۲}}=\frac{۱۰}{۲۸}$$

$$f_X(۲۰)=P(X=۲۰)=\frac{\binom{۵}{۱}\binom{۳}{۱}}{\binom{۸}{۲}}=\frac{۱۵}{۲۸} \quad , \quad f_X(۲۱)=\frac{\binom{۳}{۲}}{\binom{۸}{۲}}=\frac{۳}{۲۸}$$

بنابراین تابع احتمال X برابر است با

x	۱۹	۲۰	۲۱
$f_X(x)$	$\frac{۱۰}{۲۸}$	$\frac{۱۵}{۲۸}$	$\frac{۳}{۲۸}$

و در نتیجه تابع توزیع X به صورت زیر به دست می‌آید

$$F_X(x)=P(X\leq x)=\begin{cases} ۰ & x<۱۹ \\ \frac{۱۰}{۲۸} & ۱۹\leq x<۲۰ \\ \frac{۲۵}{۲۸} & ۲۰\leq x<۲۱ \\ ۱ & ۲۱\leq x \end{cases}$$

برای محاسبه $P(۱۹<X<۲۱)$ به دو صورت زیر می‌توان عمل کرد.

$$P(۱۹<X<۲۱)=f_X(۲۰)=\frac{۱۵}{۲۸}$$
$$P(۱۹<X<۲۱)=F_X(۲۱^-)-F_X(۱۹)=\frac{۲۵}{۲۸}-\frac{۱۰}{۲۸}=\frac{۱۵}{۲۸}$$

۳.۳ توزیع احتمالات پیوسته

متغیر تصادفی که مجموعه مقادیر آن یک فاصله عددی یا اجتماع چند فاصله عددی باشد را متغیر تصادفی پیوسته گویند. در مورد متغیر تصادفی پیوسته بایستی توجه کرد که احتمال اینکه یک متغیر تصادفی پیوسته بخواهد فقط یک مقدار بخصوص از مجموعه مقادیرش را بگیرد برابر صفر است. به مثال زیر توجه کنید.

مثال ۱.۳.۳ نقطه‌ای را به تصادف از فاصله اعداد حقیقی $[۰,۲]$ انتخاب می‌کنیم و متغیر تصادفی X را برابر نقطه انتخاب شده در فاصله $[۰,۲]$ در نظر می‌گیریم. در این صورت X یک متغیر تصادفی پیوسته است و برای هر $r\in[۰,۲]$ ، $P(X=r)=۰$. زیرا بین نقاط ۰ و ۲ بی‌نهایت نقطه وجود دارد و احتمال انتخاب یک نقطه بخصوص بسیار ناچیز است.

در حالت کلی اگر b هر عدد حقیقی و X هر متغیر تصادفی پیوسته باشد آنگاه $P(X=b)=۰$ و در نتیجه $P(a<X\leq b)=P(a<X<b)+P(X=b)=P(a<X<b)$ بنابراین توزیع احتمال یک متغیر تصادفی پیوسته را نمی‌توان به صورت یک جدول نمایش داد. در این حالت توزیع احتمال متغیر تصادفی پیوسته X را به صورت یک تابع $f_X(x)$ نمایش داده و آن را

تابع چگالی احتمال X می‌نامند.

برای به دست آوردن تابع چگالی احتمال یک متغیر تصادفی پیوسته معمولاً ابتدا تابع توزیع آن را به دست می‌آورند، زیرا تابع توزیع احتمال را در فواصل محاسبه می‌کند و محاسبه این احتمالات در حالت پیوسته امکان‌پذیر است. تابع توزیع یک متغیر تصادفی پیوسته X به صورت $F_X(x)=P(X\leq x)$ تعریف می‌شود. با توجه به آنکه در حالت گسسته تابع توزیع به صورت $F_X(x)=P(X\leq x)=\sum_{t\leq x} f_X(t)$ محاسبه می‌شد، با قرار دادن تابع چگالی احتمال به جای تابع احتمال و تبدیل مجموع به انتگرال می‌توان تابع توزیع در حالت پیوسته را به طور مشابه و به صورت زیر محاسبه کرد:

$$\boxed{F_X(x)=P(X\leq x)=\int_{-\infty}^{x} f_X(t)dt} \qquad (۴.۳)$$

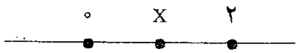
تابع $f_X(x)$ که در رابطه (۴.۳) صدق می‌کند را **تابع چگالی احتمال** و تابع $F_X(x)$ در این رابطه را **تابع توزیع متغیر تصادفی پیوسته** X گویند. همچنین با توجه به قضیه اساسی حساب از رابطه (۴.۳) نتیجه می‌شود که

$$\boxed{F'_X(x)=\frac{d}{dx}F_X(x)=f_X(x)} \qquad (۵.۳)$$

با توجه به روابط (۴.۳) و (۵.۳) با داشتن تابع توزیع X و مشتق گرفتن از آن به راحتی می‌توان تابع چگالی احتمال X را به دست آورد و برعکس با داشتن تابع چگالی احتمال X و انتگرال گرفتن از آن می‌توان تابع توزیع X را به دست آورد.

مثال ۲.۳.۳ در مثال ۱.۳.۳ تابع توزیع متغیر تصادفی پیوسته X را به دست آورده و از روی آن تابع چگالی احتمال X را محاسبه کنید و سپس این دو تابع را رسم کنید.

حل با توجه به نمودار روبرو و مفهوم تابع توزیع



اگر $x<۰$ آنگاه $F_X(x)=P(X\leq x)=۰$ و اگر $۰\leq x<۲$ آنگاه

$$F_X(x)=P(X\leq x)=\frac{[۰,x] \text{ طول فاصله}}{[۰,۲] \text{ طول فاصله}}=\frac{x}{۲}$$

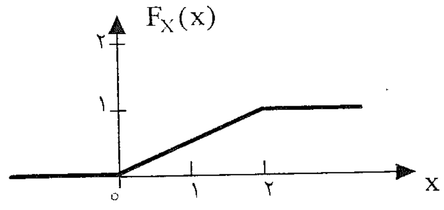
و اگر $x\geq ۲$ آنگاه $F_X(x)=P(X\leq x)=۱$. بنابراین

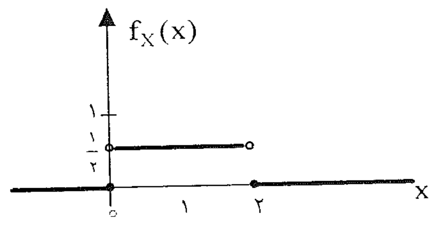
$$F_X(x)=P(X\le x)=\begin{cases}0 & x<0\\ \dfrac{x}{2} & 0\le x<2\\ 1 & 2\le x\end{cases}$$

و در نتیجه

$$f_X(x)=F_X'(x)=\begin{cases}\dfrac{1}{2} & 0<x<2\\ 0 & \text{سایر نقاط}\end{cases}$$

نمودار این دو تابع در زیر رسم شده است. همان طور که دیده می‌شود نمودار تابع توزیع یک متغیر تصادفی پیوسته یک نمودار پیوسته است.





خواص تابع چگالی احتمال و تابع توزیع با توجه به روابط (۴.۳) و (۵.۳) بین تابع چگالی احتمال و تابع توزیع می‌توان خواص زیر را برای تابع چگالی احتمال $f_X(x)$ بیان کرد.

الف- برای هر $x\in R$، $f_X(x)\ge 0$

ب- $\int_{-\infty}^{+\infty} f_X(x)\,dx=1$

همچنین تمامی خواص گفته شده برای تابع توزیع یک متغیر تصادفی گسسته در بخش ۲.۳، برای تابع توزیع یک متغیر تصادفی پیوسته برقرار است و علاوه بر آن در حالت پیوسته این تابع همواره پیوسته است.

برای محاسبه احتمالات مربوط به یک متغیر تصادفی پیوسته می‌توان از رابطه زیر استفاده کرد

$$\boxed{P(a<X\le b)=\int_a^b f_X(x)\,dx=F_X(b)-F_X(a)} \qquad (۶.۳)$$

که در آن a می‌تواند $-\infty$ و b می‌تواند $+\infty$ باشد.

مثال ۳.۳.۳ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}\dfrac{c}{x^2} & 1<x<10\\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار c را تعیین کنید.

ب- تابع توزیع X را به دست آورید.

ج- احتمالات زیر را محاسبه کنید

$$P(X>2)\quad,\quad P(1<X\le 5)\quad,\quad P([X]=3)$$

حل الف- با توجه به خواص تابع چگالی احتمال بایستی $c\ge 0$ و همچنین

$$1=\int_{-\infty}^{+\infty} f_X(x)dx=\int_{-\infty}^{1} f_X(x)dx+\int_1^{10} f_X(x)dx+\int_{10}^{+\infty} f_X(x)dx$$

$$=0+\int_1^{10}\frac{c}{x^2}dx+0=\frac{-c}{x}\Big|_1^{10}=-\frac{c}{10}+c=\frac{9}{10}c$$

پس بایستی $c=\frac{10}{9}$ باشد

ب- اگر $x<1$ آنگاه $F_X(x)=\int_{-\infty}^{x} f_X(t)dt=0$ و اگر $1\le x<10$ آنگاه

$$F_X(x)=\int_{-\infty}^{1} f_X(t)dt+\int_1^{x} f_X(t)dt=0+\int_1^{x}\frac{10}{9t^2}dt=\frac{-10}{9t}\Big|_1^{x}=\frac{-10}{9x}+\frac{10}{9}$$

و اگر $x\ge 10$ آنگاه

$$F_X(x)=\int_{-\infty}^{1} f_X(t)dt+\int_1^{10} f_X(t)dt+\int_{10}^{x} f_X(t)dt=0+1+0=1$$

بنابراین

$$F_X(x)=\begin{cases}0 & x<1\\ \dfrac{10}{9}\left(1-\dfrac{1}{x}\right) & 1\le x<10\\ 1 & 10\le x\end{cases}$$

ج- با توجه به رابطه (۶.۳) داریم که

$$P(X>2)=\int_2^{+\infty} f_X(x)dx=\int_2^{10}\frac{10}{9x^2}dx=\frac{-10}{9x}\Big|_2^{10}=\frac{-1}{9}+\frac{5}{9}=\frac{4}{9}$$

$$P(1<X\le 5)=F_X(5)-F_X(1)=\frac{10}{9}\left(1-\frac{1}{5}\right)-0=\frac{8}{9}$$

$$P([X]=۳)=P(۳\le X<۴)=F_X(۴)-F_X(۳)=\frac{۱۰}{۹}(۱-\frac{۱}{۴})-\frac{۱۰}{۹}(۱-\frac{۱}{۳})=\frac{۵}{۵۴}$$

مثال ۴.۳.۳ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases} cx^{۲}e^{-۲x} & x>۰\\ ۰ & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید.

ب – تابع توزیع X را به دست آورید و $P(X>۵)$ را محاسبه کنید.

حل الف – با توجه به خواص تابع چگالی بایستی $c\ge ۰$ و همچنین با انتگرال‌گیری به روش جزء به جزء داریم که

$$۱=\int_{-\infty}^{+\infty} f_X(x)dx = c\int_{۰}^{+\infty} x^{۲}e^{-۲x}dx = c\left[\frac{-۱}{۲}x^{۲}-\frac{۱}{۲}x-\frac{۱}{۴}\right]e^{-۲x}\Big|_{۰}^{+\infty}$$

$$=۰-(-\frac{c}{۴})=\frac{c}{۴}$$

بنابراین بایستی c=۴ باشد (توجه کنید که برای هر عدد طبیعی n و هر عدد حقیقی $\alpha>۰$ داریم که $\lim_{x\to+\infty} x^{n}e^{-\alpha x}=۰$).

ب – اگر $x<۰$ آنگاه $F_X(x)=۰$ و اگر $x\ge ۰$ آنگاه

$$F_X(x)=\int_{-\infty}^{x} f_X(t)dt=\int_{۰}^{x} ۴t^{۲}e^{-۲t}dt = ۴\left[-\frac{۱}{۲}t^{۲}-\frac{۱}{۲}t-\frac{۱}{۴}\right]e^{-۲t}\Big|_{۰}^{x}$$

$$=۱-(۲x^{۲}+۲x+۱)e^{-۲x}$$

بنابراین

$$F_X(x)=\begin{cases}۰ & x<۰\\ ۱-(۲x^{۲}+۲x+۱)e^{-۲x} & x\ge ۰\end{cases}$$

و در نتیجه

$$P(X>۵)=۱-P(X\le ۵)=۱-F_X(۵)=۶۱e^{-۱۰}=۰/۰۰۳$$

۴.۳ توزیع احتمالات توأم دو متغیره

در بخش‌های قبل متغیرهای تصادفی مورد مطالعه یک بعدی بودند. در بعضی از مسائل ممکن است که نتایجی از چندین متغیر تصادفی را به طور همزمان لازم داشته باشیم. به مثال زیر توجه کنید.

مثال ۱.۴.۳ فرض کنید سکه‌ای را ۳ مرتبه پرتاب کنیم و قرار دهیم:

$X=$ تعداد شیرهای مشاهده شده در ۳ مرتبه پرتاب سکه

$Y=$ تعداد شیرهای مشاهده شده در پرتاب سوم سکه

حال پیشامدهای (X=۱) و (Y=۱) را در نظر بگیرید. اگر وقوع این دو پیشامد در یکدیگر تأثیری نداشته باشد آنگاه دانستن تابع احتمال $f_X(x)=P(X=x)$ و $f_Y(y)=P(Y=y)$ به تنهایی تمام اطلاعات را در مورد متغیرهای تصادفی X و Y به ما می‌دهد. اما همانطور که دیده می‌شود وقوع هر یک از این دو پیشامد در دیگری تأثیر می‌گذارد و بنابراین نیاز به دانستن اطلاعاتی در مورد وقوع همزمان این دو پیشامد داریم.

اگر X و Y دو متغیر تصادفی باشند توزیع احتمال برای وقوع همزمان آنها به صورت تابع دو متغیره $f_{X,Y}(x,y)$ نشان داده می‌شود و معمولاً آن را توزیع احتمال توأم X و Y گویند. اگر X و Y متغیرهای تصادفی گسسته باشند این تابع به فرم زیر تعریف می‌شود.

$$\boxed{f_{X,Y}(x,y)=P(X=x,Y=y)} \qquad (۷.۳)$$

یعنی $f_{X,Y}(x,y)$ احتمال این است که نتایج x و y به طور همزمان اتفاق بیفتند. مثلاً در مثال ۱.۴.۳، $f_{X,Y}(۱,۱)=P(X=۱,Y=۱)$ به معنای احتمال این است که در ۳ مرتبه پرتاب سکه دقیقاً در پرتاب سوم یک شیر مشاهده کنیم که این احتمال برابر $\frac{۱}{۸}$ است. با توجه به رابطه (۷.۳) تعریف زیر را داریم.

تعریف ۴.۳ تابع $f_{X,Y}(x,y)$ را یک تابع احتمال توأم متغیرهای تصادفی گسسته X و Y گویند هرگاه

الف – برای هر x و y داشته باشیم $f_{X,Y}(x,y)\ge ۰$.

ب – $\sum_{x}\sum_{y} f_{X,Y}(x,y)=۱$.

برای محاسبه احتمال قرار گرفتن (X,Y) در یک ناحیه A در صفحه xy به صورت زیر عمل می‌کنیم.

$$\boxed{P((X,Y)\in A)=\sum\sum_{(x,y)\in A} f_{X,Y}(x,y)} \qquad (۸.۳)$$

مثال ۲.۴.۳ در مثال ۱.۴.۳ تابع احتمال توأم X و Y را به دست آورید و $P((X,Y)\in A)$ را محاسبه کنید که در آن $A=\{(x,y) \mid x\le y\}$.

حل توجه کنید که $S_X=\{۰,۱,۲,۳\}$ و $S_Y=\{۰,۱\}$ بنابراین

$$f_{X,Y}(۰,۰)=P(X=۰,Y=۰)=P(\{TTT\})=\frac{۱}{۸}$$
$$f_{X,Y}(۰,۱)=P(X=۰,Y=۱)=۰$$
$$f_{X,Y}(۱,۰)=P(X=۱,Y=۰)=P(\{HTT,THT\})=\frac{۲}{۸}$$

با محاسبه احتمالات مربوط به نقاط دیگر، جدول زیر را که به جدول توزیع احتمالات توأم X و Y موسوم است به دست می‌آوریم.

y \ x	۰	۱	۲	۳
۰	$\frac{۱}{۸}$	$\frac{۲}{۸}$	$\frac{۱}{۸}$	۰
۱	۰	$\frac{۱}{۸}$	$\frac{۲}{۸}$	$\frac{۱}{۸}$

توجه کنید که
$$\sum_{x=۰}^{۳}\sum_{y=۰}^{۱} f_{X,Y}(x,y)=۱$$

چون $A=\{(x,y) \mid x\le y\}=\{(۰,۰),(۰,۱),(۱,۱)\}$ بنابراین با توجه به رابطه (۸.۳) داریم که
$$P(X\le Y)=f_{X,Y}(۰,۰)+f_{X,Y}(۰,۱)+f_{X,Y}(۱,۱)=\frac{۱}{۸}+۰+\frac{۱}{۸}=\frac{۱}{۴}$$

مثال ۳.۴.۳ از داخل جعبه‌ای که شامل ۳ توپ آبی، ۲ توپ قرمز و ۴ توپ سبز است، دو توپ به تصادف یک به یک و بدون جایگذاری انتخاب می‌کنیم و قرار می‌دهیم

X = تعداد توپهای آبی مشاهده شده در ۲ توپ انتخابی

Y = تعداد توپهای قرمز مشاهده شده در ۲ توپ انتخابی

الف– تابع احتمال توأم (X,Y) را به دست آورید.

ب– $P(X+Y\le ۱)$ را محاسبه کنید.

حل الف– توجه کنید که $S_X=S_Y=\{۰,۱,۲\}$ و برای مثال $f_{X,Y}(۰,۰)$ به معنای احتمال این است که در ۲ توپ انتخابی ما هیچ توپ آبی، هیچ توپ قرمز و ۲ توپ سبز داشته باشیم که احتمال آن

برابر است با
$$f_{X,Y}(۰,۰)=P(X=۰,Y=۰)=\frac{\binom{۳}{۰}\binom{۲}{۰}\binom{۴}{۲}}{\binom{۹}{۲}}=\frac{۶}{۳۶}$$

به همین ترتیب
$$f_{X,Y}(۰,۱)=P(X=۰,Y=۱)=\frac{\binom{۳}{۰}\binom{۲}{۱}\binom{۴}{۱}}{\binom{۹}{۲}}=\frac{۸}{۳۶}$$
$$f_{X,Y}(۱,۰)=P(X=۱,Y=۰)=\frac{\binom{۳}{۱}\binom{۲}{۰}\binom{۴}{۱}}{\binom{۹}{۲}}=\frac{۱۲}{۳۶}$$

با محاسبه احتمالات مربوط به نقاط دیگر جدول توزیع احتمالات توأم X و Y به صورت زیر به دست می‌آید.

y \ x	۰	۱	۲
۰	$\frac{۶}{۳۶}$	$\frac{۱۲}{۳۶}$	$\frac{۳}{۳۶}$
۱	$\frac{۸}{۳۶}$	$\frac{۶}{۳۶}$	۰
۲	$\frac{۱}{۳۶}$	۰	۰

جدول ۱.۳ جدول توزیع احتمالات توأم

توجه کنید که تابع احتمال توأم X و Y را به صورت زیر نیز می‌توان نوشت

$$f_{X,Y}(x,y)=\frac{\binom{۳}{x}\binom{۲}{y}\binom{۴}{۲-x-y}}{\binom{۹}{۲}} \qquad x,y=۰,۱,۲ \;,\; ۰\le x+y\le ۲$$

ب– با توجه به رابطه (۸.۳) داریم که
$$P(X+Y\le ۱)=f_{X,Y}(۰,۰)+f_{X,Y}(۰,۱)+f_{X,Y}(۱,۰)=\frac{۶}{۳۶}+\frac{۸}{۳۶}+\frac{۱۲}{۳۶}=\frac{۲۶}{۳۶}$$

برای متغیرهای تصادفی پیوسته تعریف زیر را داریم.

تعریف ۵.۳ تابع $f_{X,Y}(x,y)$ را یک **تابع چگالی احتمال توأم** متغیرهای تصادفی پیوسته X و Y گویند هرگاه

الف – برای هر x و y داشته باشیم $f_{X,Y}(x,y)\geq 0$.

ب – $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx\,dy = 1$.

برای محاسبه احتمال قرار گرفتن (X,Y) در یک ناحیه A در صفحه xy به صورت زیر عمل می‌کنیم.

$$\boxed{P((X,Y)\in A) = \iint_A f_{X,Y}(x,y)\,dx\,dy} \qquad (9.3)$$

مثال ۴.۴.۳ تابع زیر را در نظر بگیرید.

$$f_{X,Y}(x,y) = \begin{cases} cx(1+2y) & 0<x<3\ ,\ 0<y<1 \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – مقدار c را به گونه‌ای تعیین کنید که این تابع یک تابع چگالی احتمال توأم متغیرهای تصادفی X و Y باشد.

ب – $P(1<X<2,\ 0<Y<\frac{1}{2})$ و $P(X+2Y\geq 3)$ را محاسبه کنید.

حل الف – با توجه به خواص تابع چگالی احتمال توأم، بایستی $c\geq 0$ و همچنین

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx\,dy = \int_0^1\int_0^3 cx(1+2y)\,dx\,dy$$

$$= c\int_0^1 (1+2y)\left[\int_0^3 x\,dx\right]dy = c\int_0^1 (1+2y)\left[\frac{1}{2}x^2\right]_0^3 dy$$

$$= \frac{9}{2}c\int_0^1 (1+2y)\,dy = \frac{9}{2}c\left[y+y^2\right]_0^1 = 9c$$

بنابراین بایستی $c=\frac{1}{9}$ باشد.

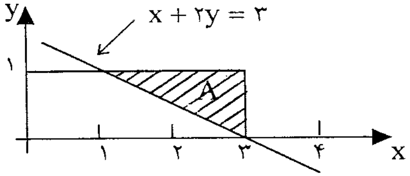
ب – با توجه به رابطه (۹.۳) داریم که

$$P(1<X<2,\ 0<Y<\tfrac{1}{2}) = \int_0^{\frac{1}{2}}\int_1^2 \frac{1}{9}x(1+2y)\,dx\,dy$$

$$= \frac{1}{9}\int_0^{\frac{1}{2}} (1+2y)\left[\frac{1}{2}x^2\right]_1^2 dy = \frac{1}{9}\times\frac{3}{2}\int_0^{\frac{1}{2}}(1+2y)\,dy$$

$$= \frac{1}{6}\left[y+y^2\right]_0^{\frac{1}{2}} = \frac{1}{6}\times\frac{3}{4} = \frac{1}{8}$$

برای محاسبه $P(X+2Y\geq 3)$ ابتدا بایستی ناحیه $A=\{(x,y)\mid x+2y\geq 3\}$ را مشخص کنیم. با توجه به نمودار زیر داریم که



$$A=\{(x,y)\mid 1<x<3,\ \frac{3-x}{2}<y<1\}$$
$$=\{(x,y)\mid 3-2y<x<3,\ 0<y<1\}$$

بنابراین

$$P(X+2Y\geq 3) = \int_0^1\int_{3-2y}^3 \frac{1}{9}x(1+2y)\,dx\,dy = \frac{1}{9}\int_0^1 (1+2y)\left[\frac{1}{2}x^2\right]_{3-2y}^3 dy$$

$$= \frac{2}{9}\int_0^1 (1+2y)(3y-y^2)\,dy$$

$$= \frac{2}{9}\left[\frac{3}{2}y^2+\frac{5}{3}y^3-\frac{1}{2}y^4\right]_0^1 = \frac{2}{9}\times\frac{8}{3} = \frac{16}{27}$$

توزیع احتمالات حاشیه‌ای (کناری)

با داشتن تابع احتمال (چگالی احتمال) توأم متغیرهای تصادفی X و Y می‌توان تابع احتمال (چگالی احتمال) X به تنهایی و Y به تنهایی را محاسبه کرد که به آنها توابع احتمال (چگالی احتمال) حاشیه‌ای گویند. اگر X و Y متغیرهای تصادفی گسسته باشند، تابع احتمال حاشیه‌ای X و تابع احتمال حاشیه‌ای Y به صورت زیر به دست می‌آیند.

$$\boxed{f_X(x) = \sum_y f_{X,Y}(x,y) \quad ,\quad f_Y(y) = \sum_x f_{X,Y}(x,y)} \qquad (10.3)$$

و اگر X و Y متغیرهای تصادفی پیوسته باشند، تابع چگالی احتمال حاشیه‌ای X و تابع چگالی احتمال حاشیه‌ای Y به صورت زیر به دست می‌آیند.

$$\boxed{f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy \quad ,\ f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx} \qquad (11.3)$$

به سادگی می‌توان نشان داد که توابع به دست آمده در (۱۰.۳) و (۱۱.۳) تمامی خواص تابع احتمال و تابع چگالی احتمال متغیرهای تصادفی X و Y را دارند.

مثال ۵.۴.۳ در مثال ۳.۴.۳ تابع احتمال حاشیه‌ای X و تابع احتمال حاشیه‌ای Y را به دست آورید.

حل با توجه به اینکه $S_X = S_Y = \{۰, ۱, ۲\}$ بنابراین با توجه به جدول ۱.۳ و رابطه (۱۰.۳) داریم که

$$f_X(۰) = \sum_{y=۰}^{۲} f_{X,Y}(۰,y) = f_{X,Y}(۰,۰)+f_{X,Y}(۰,۱)+f_{X,Y}(۰,۲) = \frac{۶}{۳۶}+\frac{۸}{۳۶}+\frac{۱}{۳۶} = \frac{۱۵}{۳۶}$$

$$f_X(۱) = \sum_{y=۰}^{۲} f_{X,Y}(۱,y) = \frac{۱۲}{۳۶}+\frac{۶}{۳۶}+۰ = \frac{۱۸}{۳۶}$$

$$f_X(۲) = \sum_{y=۰}^{۲} f_{X,Y}(۲,y) = \frac{۳}{۳۶}+۰+۰ = \frac{۳}{۳۶}$$

بنابراین تابع احتمال حاشیه‌ای X عبارت است از

x	۰	۱	۲
$f_X(x)$	$\frac{۱۵}{۳۶}$	$\frac{۱۸}{۳۶}$	$\frac{۳}{۳۶}$

به همین ترتیب تابع احتمال حاشیه‌ای Y به صورت زیر به دست می‌آید.

y	۰	۱	۲
$f_Y(y)$	$\frac{۲۱}{۳۶}$	$\frac{۱۴}{۳۶}$	$\frac{۱}{۳۶}$

می‌توان عملیات فوق را در جدول توزیع احتمالات توأم به صورت زیر خلاصه کرد

y \ x	۰	۱	۲	$f_Y(y)$
۰	$\frac{۶}{۳۶}$	$\frac{۱۲}{۳۶}$	$\frac{۳}{۳۶}$	$\frac{۲۱}{۳۶}$
۱	$\frac{۸}{۳۶}$	$\frac{۶}{۳۶}$	۰	$\frac{۱۴}{۳۶}$
۲	$\frac{۱}{۳۶}$	۰	۰	$\frac{۱}{۳۶}$
$f_X(x)$	$\frac{۱۵}{۳۶}$	$\frac{۱۸}{۳۶}$	$\frac{۳}{۳۶}$	

توزیع احتمالات شرطی

در فصل دوم احتمال شرطی را به صورت زیر تعریف کردیم

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad , \quad P(B) \neq ۰$$

اگر X و Y متغیرهای تصادفی گسسته باشند و قرار دهیم $A \equiv (X=x)$ و $B \equiv (Y=y)$ در این صورت

$$P(X=x \mid Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad , \quad f_Y(y) \neq ۰$$

به سادگی می‌توان نشان داد که تابع $\frac{f_{X,Y}(x,y)}{f_Y(y)}$ به عنوان تابعی از x برای y ثابت، تمامی شرایط یک تابع احتمال را دارد که به آن تابع احتمال شرطی X به شرط Y=y گویند و آن را با نماد زیر نمایش می‌دهند.

$$\boxed{f_{X \mid Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad , \quad f_Y(y) \neq ۰} \qquad (۱۲.۳)$$

به همین ترتیب تابع احتمال شرطی Y به شرط X=x به صورت زیر تعریف می‌شود.

$$\boxed{f_{Y \mid X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad , \quad f_X(x) \neq ۰} \qquad (۱۳.۳)$$

اگر X و Y متغیرهای تصادفی پیوسته باشند، به طور مشابه تابع چگالی احتمال شرطی X به شرط Y=y توسط رابطه (۱۲.۳) و تابع چگالی احتمال شرطی Y به شرط X=x توسط رابطه (۱۳.۳) تعریف می‌شوند. همچنین برای محاسبه احتمالات شرطی از رابطه زیر استفاده می‌کنیم.

$$\boxed{P(a < X < b \mid Y=c) = \begin{cases} \sum\limits_{a<x<b} f_{X \mid Y}(x \mid c) & \text{اگر X و Y گسسته باشند} \\ \int_a^b f_{X \mid Y}(x \mid c)\,dx & \text{اگر X و Y پیوسته باشند} \end{cases}} \qquad (۱۴.۳)$$

مثال ۶.۴.۳ در مثال ۳.۴.۳ تابع احتمال شرطی X به شرط Y=۰ را به دست آورده و $P(X \leq ۱ \mid Y=۰)$ را محاسبه کنید.

حل با توجه به رابطه (۱۲.۳) داریم که

$$f_{X \mid Y}(x \mid ۰) = \frac{f_{X,Y}(x,۰)}{f_Y(۰)} = \frac{f_{X,Y}(x,۰)}{\frac{۲۱}{۳۶}} \qquad x = ۰, ۱, ۲$$

بنابراین با توجه به جدول ۱.۳ داریم که

x	0	1	2
$f_{X\mid Y}(x\mid 0)$	$\frac{6}{21}$	$\frac{12}{21}$	$\frac{3}{21}$

و از رابطه (۱۴.۳) نتیجه می‌شود که

$$P(X\le 1 \mid Y=0)=\sum_{x=0}^{1} f_{X\mid Y}(x\mid 0)=\frac{6}{21}+\frac{12}{21}=\frac{6}{7}$$

مثال ۷.۴.۳ در مثال ۴.۴.۳ توابع چگالی احتمال $f_X(x)$، $f_Y(y)$ و $f_{X\mid Y}(x\mid y)$ را بـه دست آورده و $P(1<X<2\mid Y=\frac{1}{2})$ را محاسبه کنید.

حل با توجه به رابطه (۱۱.۳) داریم که

$$f_X(x)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy=\int_0^1 \frac{1}{9}x(1+2y)\,dy=\frac{1}{9}x\left[y+y^2\right]_0^1=\frac{2}{9}x$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx=\int_0^3 \frac{1}{9}x(1+2y)\,dx=\frac{1}{9}(1+2y)\left[\frac{1}{2}x^2\right]_0^3=\frac{1}{2}(1+2y)$$

بنابراین

$$f_X(x)=\begin{cases}\frac{2}{9}x & 0<x<3\\ 0 & \text{سایر نقاط}\end{cases} \quad , \quad f_Y(y)=\begin{cases}\frac{1}{2}(1+2y) & 0<y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

همچنین از رابطه (۱۲.۳) داریم که

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{\frac{1}{9}x(1+2y)}{\frac{1}{2}(1+2y)}=\frac{2}{9}x$$

بنابراین

$$f_{X\mid Y}(x\mid y)=\begin{cases}\frac{2}{9}x & 0<x<3\ ,\ 0<y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

و با توجه به رابطه (۱۴.۳) داریم که

$$P(1<X<2\mid Y=\tfrac{1}{2})=\int_1^2 f_{X\mid Y}(x\mid \tfrac{1}{2})\,dx=\int_1^2 \frac{2}{9}x\,dx=\frac{1}{9}x^2\Big|_1^2=\frac{1}{3}$$

متغیرهای تصادفی مستقل در مثال ۷.۴.۳ دیده می‌شود که $f_{X\mid Y}(x\mid y)$ به y بستگی نداشت و در حقیقت $f_{X\mid Y}(x\mid y)=f_X(x)$. در این حالت متغیر تصادفی Y تأثیری روی متغیر تصادفی X ندارد و گویند این دو متغیر تصادفی از یکدیگر مستقل هستند. در این حالت داریم که

$$f_{X\mid Y}(x\mid y)=f_X(x)\ \Rightarrow\ \frac{f_{X,Y}(x,y)}{f_Y(y)}=f_X(x)\ \Rightarrow\ f_{X,Y}(x,y)=f_X(x)f_Y(y)$$

تعریف ۶.۳ فرض کنید X و Y متغیرهای تصادفی با تابع احتمال (چگالی احتمال) توأم $f_{X,Y}(x,y)$ و توابع احتمال (چگالی احتمال) حاشیه‌ای $f_X(x)$ و $f_Y(y)$ باشند. متغیرهای تـصادفی X و Y را از لحاظ آماری مستقل گویند اگر و فقط اگر

$$\boxed{f_{X,Y}(x,y)=f_X(x)f_Y(y) \qquad \text{برای هر } x,y} \qquad (۱۵.۳)$$

برای مثال متغیرهای تصادفی مثال ۴.۴.۳ از یکدیگر مستقل هستند زیرا برای آنها رابـطه (۱۵.۳) برقرار است (مثال ۷.۴.۳ را ملاحظه کنید) ولی متغیرهای تصادفی مثال ۳.۴.۳ از یکدیگر مستقل نیستند زیرا در این مثال داریم که

$$f_X(0)=\frac{15}{36} \quad , \quad f_Y(0)=\frac{21}{36} \quad , \quad f_{X,Y}(0,0)=\frac{6}{36}$$

و بنابراین $f_{X,Y}(0,0)\neq f_X(0)f_Y(0)$.

مثال ۸.۴.۳ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند.

$$f_{X,Y}(x,y)=\begin{cases}cxy & 0<y<2x^2\ ,\ 0<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار c را تعیین کنید.

ب- $P(X+Y<1)$ را محاسبه کنید.

ج- $P(0<Y<\frac{1}{3}\mid X=\frac{1}{2})$ و $P(\frac{1}{2}<X<1\mid Y=\frac{1}{3})$ را محاسبه کنید.

د- آیا X و Y مستقل هستند؟ چرا؟

حل الف- در این مثال کران متغیرهای X و Y به یکدیگر وابسته است و ناحیه‌ای را که می‌توان روی آن انتگرال گرفت عبارت است از

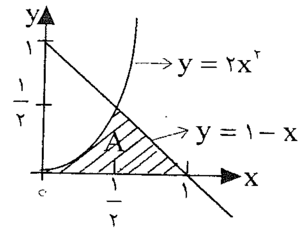
$$B=\{(x,y)\mid 0<x<1,\ 0<y<2x^2\}=\left\{(x,y)\ \Big|\ \sqrt{\frac{y}{2}}<x<1,\ 0<y<2\right\}$$

بنابراین بایستی $c\ge 0$ و همچنین

$$1=\int_0^1\int_0^{2x^2} cxy\,dy\,dx=c\int_0^1 x\left[\frac{1}{2}y^2\right]_0^{2x^2}dx=2c\int_0^1 x^5\,dx$$

$$=\frac{c}{3}x^6\Big|_0^1=\frac{c}{3} \quad\Rightarrow\quad c=3$$

ب- برای محاسبه احتمال بایستی ناحیه $A=\{(x,y) \mid x+y<1\}$ را معین کنیم. با توجه به نمودار زیر داریم که



$$A=\{(x,y)\mid \sqrt{\frac{y}{2}}<x<1-y\ ,\ 0<y<\frac{1}{2}\}$$

بنابراین

$$P(X+Y<1)=\int_0^{\frac12}\int_{\sqrt{\frac{y}{2}}}^{1-y} 3xy\,dx\,dy = 3\int_0^{\frac12} y\left[\frac{1}{2}x^2\right]_{\sqrt{\frac{y}{2}}}^{1-y}dy$$

$$=\frac{3}{2}\int_0^{\frac12} y\left(1-\frac{5}{2}y+y^2\right)dy$$

$$=\frac{3}{2}\left[\frac{1}{2}y^2-\frac{5}{6}y^3+\frac{1}{4}y^4\right]_0^{\frac12}=\frac{7}{128}$$

ج- ابتدا توابع چگالی احتمال حاشیه‌ای و شرطی را به دست می‌آوریم (به حدود انتگرالها و توابع توجه کنید).

$$f_X(x)=\int_0^{2x^2} 3xy\,dy=\frac{3}{2}x\left[y^2\right]_0^{2x^2}=6x^5$$

$$f_Y(y)=\int_{\sqrt{\frac{y}{2}}}^{1} 3xy\,dx=\frac{3}{2}y\left[x^2\right]_{\sqrt{\frac{y}{2}}}^{1}=\frac{3}{2}y\left(1-\frac{y}{2}\right)$$

بنابراین

$$f_X(x)=\begin{cases}6x^5 & 0<x<1\\ 0 & \text{سایر نقاط}\end{cases}\ ,\quad f_Y(y)=\begin{cases}\frac{3}{2}y\left(1-\frac{y}{2}\right) & 0<y<2\\ 0 & \text{سایر نقاط}\end{cases}$$

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{3xy}{\frac{3}{2}y\left(1-\frac{y}{2}\right)}=\frac{4x}{2-y}$$

$$f_{Y\mid X}(y\mid x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{3xy}{6x^5}=\frac{y}{2x^4}$$

بنابراین

$$f_{X\mid Y}(x\mid y)=\begin{cases}\dfrac{4x}{2-y} & 0<y<2\ ,\ \sqrt{\dfrac{y}{2}}<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

$$f_{Y\mid X}(y\mid x)=\begin{cases}\dfrac{y}{2x^4} & 0<x<1\ ,\ 0<y<2x^2\\ 0 & \text{سایر نقاط}\end{cases}$$

با توجه به این توابع، احتمالات مورد نظر به صورت زیر به دست می‌آیند.

$$P(0<Y<\tfrac{1}{3}\mid X=\tfrac{1}{2})=\int_0^{\frac13} f_{Y\mid X}(y\mid \tfrac{1}{2})dy=\int_0^{\frac13} 8y\,dy=4y^2\Big|_0^{\frac13}=\frac{4}{9}$$

$$P(\tfrac{1}{2}<X<1\mid Y=\tfrac{1}{3})=\int_{\frac12}^{1} f_{X\mid Y}(x\mid \tfrac{1}{3})dx=\int_{\frac12}^{1}\frac{4x}{\frac{5}{3}}dx=\frac{6}{5}x^2\Big|_{\frac12}^{1}=\frac{9}{10}$$

د- چون $f_{X,Y}(x,y)\neq f_X(x)f_Y(y)$ پس X و Y از یکدیگر مستقل نیستند. در ضمن چون حدود متغیرهای تصادفی X و Y به یکدیگر وابسته است پس X و Y از یکدیگر مستقل نیستند.

۵.۳ توزیع احتمالات چند متغیره

تمام بحث بخش قبل در مورد توزیع احتمالات توأم دو متغیر تصادفی را می‌توان به n متغیر تصادفی تعمیم داد. فرض کنید $f(x_1,x_2,\dots,x_n)$ تابع احتمال توأم متغیرهای تصادفی گسسته X_1، X_2، ... و X_n باشد. تابع احتمال حاشیه‌ای X_1 و تابع احتمال توأم حاشیه‌ای X_1 و X_2 به صورت زیر به دست می‌آیند.

$$f_{X_1}(x_1)=\sum_{x_2}\sum_{x_3}\cdots\sum_{x_n} f(x_1,x_2,\dots,x_n)$$

$$f_{X_1,X_2}(x_1,x_2)=\sum_{x_3}\sum_{x_4}\cdots\sum_{x_n} f(x_1,x_2,\dots,x_n)$$

همچنین اگر $f(x_1,x_2,\dots,x_n)$ تابع چگالی احتمال توأم متغیرهای X_1، X_2،... و X_n باشد، تابع چگالی احتمال حاشیه‌ای X_2 و تابع چگالی احتمال توأم حاشیه‌ای X_2 و X_3 به صورت زیر به دست می‌آیند.

$$f_{X_2}(x_2)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1,\dots,x_n)\,dx_1\,dx_3\dots dx_n$$

$$f_{X_2,X_3}(x_2,x_3)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,\dots,x_n)\,dx_1\,dx_4\dots dx_n$$

همچنین تابع احتمال (چگالی احتمال) توأم شرطی X_1، X_2 و X_3 به شرط $X_4=x_4$، ... و $X_n=x_n$ به صورت زیر به دست می‌آید.

$$f_{X_1,X_2,X_3|X_4,\dots,X_n}(x_1,x_2,x_3|x_4,\dots,x_n)=\frac{f(x_1,x_2,\dots,x_n)}{f_{X_4,\dots,X_n}(x_4,\dots,x_n)}$$

تعریف ۶.۳ را می‌توان برای استقلال n متغیر تصادفی به صورت زیر تعمیم داد.

تعریف ۷.۳ فرض کنید X_1، X_2، ... و X_n متغیرهای تصادفی با تابع احتمال (چگالی احتمال) توأم $f(x_1,x_2,\dots,x_n)$ و توابع احتمال (چگالی احتمال) حاشیه‌ای $f_{X_1}(x_1)$، $f_{X_2}(x_2)$، ... و $f_{X_n}(x_n)$ باشند. متغیرهای تصادفی X_1، X_2، ... و X_n را دو به دو از لحاظ آماری مستقل گویند اگر و فقط اگر برای هر x_1، x_2، ... و x_n

$$f(x_1,x_2,\dots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\dots f_{X_n}(x_n)$$

مثال ۱.۵.۳ فرض کنید متغیرهای تصادفی X_1، X_2 و X_3 دارای تابع چگالی احتمال توأم زیر باشند

$$f(x_1,x_2,x_3)=\begin{cases}c & 0<x_1<x_2<x_3<1\\ 0 & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید.

ب – تابع چگالی احتمال حاشیه‌ای X_2 را به دست آورید.

ج – تابع چگالی احتمال توأم X_1 و X_3 را به دست آورید و $P(2X_1<X_3)$ را محاسبه کنید.

د – تابع چگالی احتمال شرطی X_1 به شرط $(X_2,X_3)=(x_2,x_3)$ را به دست آورید.

حل الف –
$$1=c\int_0^1\int_0^{x_3}\int_0^{x_2}dx_1\,dx_2\,dx_3=c\int_0^1\int_0^{x_3}x_2\,dx_2\,dx_3$$
$$=c\int_0^1\frac{1}{2}x_3^2\,dx_3=\frac{c}{6}x_3^3\Big|_0^1=\frac{c}{6}$$

بنابراین c=۶.

ب –
$$f_{X_2}(x_2)=\int_{x_2}^1\int_0^{x_2}6\,dx_1\,dx_3=\int_{x_2}^1 6x_2\,dx_3=6x_2(1-x_2)\qquad 0<x_2<1$$

ج –
$$f_{X_1,X_3}(x_1,x_3)=\int_{x_1}^{x_3}6\,dx_2=6(x_3-x_1)\qquad 0<x_1<x_3<1$$

بنابراین
$$P(2X_1<X_3)=\int_0^1\int_0^{\frac{x_3}{2}}6(x_3-x_1)\,dx_1\,dx_3=6\int_0^1\left[x_1x_3-\frac{1}{2}x_1^2\right]_0^{\frac{x_3}{2}}dx_3$$
$$=6\left(\frac{3}{8}\right)\int_0^1 x_3^2\,dx=\frac{3}{4}x_3^3\Big|_0^1=\frac{3}{4}$$

د –
$$f_{X_2,X_3}(x_2,x_3)=\int_0^{x_2}6\,dx_1=6x_2\qquad 0<x_2<x_3<1$$

$$f_{X_1|X_2,X_3}(x_1|x_2,x_3)=\frac{f(x_1,x_2,x_3)}{f_{X_2,X_3}(x_2,x_3)}=\frac{6}{6x_2}=\frac{1}{x_2}\qquad 0<x_1<x_2<x_3<1$$

۶.۳ مسائل حل شده

مثال ۱.۶.۳ تابع احتمال متغیر تصادفی X به صورت زیر داده شده است. مقدار k را پیدا کنید.

$$f_X(x)=\frac{k}{2^x}\qquad x=0,1,2,3,4$$

حل بایستی $k\geq 0$ و همچنین
$$1=\sum_{x=0}^{4}f_X(x)=\sum_{x=0}^{4}\frac{k}{2^x}=k\left(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}\right)=k\left(\frac{31}{16}\right)$$

بنابراین $k=\frac{16}{31}$.

مثال ۲.۶.۳ فروشگاهی ۶ دستگاه تلویزیون دارد که ۲ دستگاه آن معیوب است. هتلی ۳ دستگاه آن را به طور تصادفی خریداری می‌نماید. اگر X تعداد تلویزیونهای معیوب باشد که توسط هتل خریداری شده است. تابع احتمال X را به دست آورید.

حل در اینجا $S_X=\{0,1,2\}$ و بنابراین

$$f_X(0)=P(X=0)=\frac{\binom{2}{0}\binom{4}{3}}{\binom{6}{3}}=\frac{4}{20}=\frac{1}{5}$$

$$f_X(1)=P(X=1)=\frac{\binom{2}{1}\binom{4}{2}}{\binom{6}{3}}=\frac{12}{20}=\frac{3}{5}$$

$$f_X(۲)=P(X=۲)=\frac{\binom{۲}{۲}\binom{۴}{۱}}{\binom{۶}{۳}}=\frac{۴}{۲۰}=\frac{۱}{۵}$$

بنابراین تابع احتمال X برابر است با

x	۰	۱	۲
$f_X(x)$	$\frac{۱}{۵}$	$\frac{۳}{۵}$	$\frac{۱}{۵}$

مثال ۳.۶.۳ سکه‌ای به گونه‌ای طراحی شده است که احتمال شیر آمدن آن $\frac{۳}{۴}$ و احتمال خط آمدن آن $\frac{۱}{۴}$ است. سکه را ۳ مرتبه پرتاب می‌کنیم و متغیر تصادفی X را برابر تعداد شیرهای مشاهده شده در ۳ پرتاب سکه در نظر می‌گیریم. تابع احتمال X را به دست آورید و سپس تابع توزیع X و $P(X>۱)$ را بیابید.

حل در اینجا $S_X=\{۰,۱,۲,۳\}$ و بنابراین

$$f_X(۰)=P(X=۰)=P(\{TTT\})=\frac{۱}{۴}\times\frac{۱}{۴}\times\frac{۱}{۴}=\frac{۱}{۶۴}$$

$$f_X(۱)=P(X=۱)=P(\{TTH,THT,HTT\})=۳P(\{TTH\})$$
$$=۳\times\frac{۱}{۴}\times\frac{۱}{۴}\times\frac{۳}{۴}=\frac{۹}{۶۴}$$

با به دست آوردن احتمالات دیگر، تابع احتمال X به صورت زیر به دست می‌آید.

x	۰	۱	۲	۳
$f_X(x)$	$\frac{۱}{۶۴}$	$\frac{۹}{۶۴}$	$\frac{۲۷}{۶۴}$	$\frac{۲۷}{۶۴}$

با استفاده از فرمول (۲.۳) تابع توزیع X به صورت زیر به دست می‌آید.

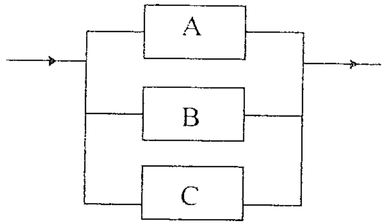
$$F_X(x)=P(X\le x)=\begin{cases} ۰ & x<۰ \\ \frac{۱}{۶۴} & ۰\le x<۱ \\ \frac{۱۰}{۶۴} & ۱\le x<۲ \\ \frac{۳۷}{۶۴} & ۲\le x<۳ \\ ۱ & ۳\le x \end{cases}$$

بنابراین

$$P(X>۱)=۱-P(X\le ۱)=۱-F_X(۱)=۱-\frac{۱۰}{۶۴}=\frac{۵۴}{۶۴}$$

مثال ۴.۶.۳ فرض کنید که در سیستم زیر اجزاء A و B و C مستقل از یکدیگر کار کنند.

احتمال اینکه A کارکند ۰/۹، B کار کند ۰/۸ و C کار کند ۰/۷ است. اگر متغیر تصادفی X نشان دهنده تعداد اجزائی باشد که در یک زمان معین کار می‌کنند، تابع احتمال X را به دست آورید.



حل اگر A، B و C به ترتیب پیشامد این باشند که اجزاء A، B و C در یک زمان معین کار کنند آنگاه A و B و C از یکدیگر مستقل هستند. در اینجا $S_X=\{۰,۱,۲,۳\}$ بنابراین

$$f_X(۰)=P(X=۰)=P(A'\cap B'\cap C')=P(A')P(B')P(C')$$
$$=(۰/۱)(۰/۲)(۰/۳)=۰/۰۰۶$$

$$f_X(۱)=P(X=۱)=P(A\cap B'\cap C')+P(A'\cap B'\cap C)+P(A'\cap B\cap C')$$
$$=(۰/۹)(۰/۲)(۰/۳)+(۰/۱)(۰/۲)(۰/۷)+(۰/۱)(۰/۸)(۰/۳)=۰/۰۹۲$$

به همین ترتیب تابع احتمال X به صورت زیر نتیجه می‌شود.

x	۰	۱	۲	۳
$f_X(x)$	۰/۰۰۶	۰/۰۹۲	۰/۳۹۸	۰/۵۰۴

مثال ۵.۶.۳ تحقیق کنید که تابع زیر یک تابع توزیع است و تابع احتمال برای W را تعیین کنید و با استفاده از آن $P(۳<W\le ۵)$ را محاسبه کنید.

$$F_W(w)=\begin{cases} ۰ & w<۳ \\ \frac{۱}{۳} & ۳\le w<۴ \\ \frac{۱}{۲} & ۴\le w<۵ \\ \frac{۲}{۳} & ۵\le w<۶ \\ ۱ & ۶\le w \end{cases}$$

حل این تابع یک تابع غیرنزولی و از سمت راست پیوسته است و همچنین $۰\le F_W(w)\le ۱$، $F_W(-\infty)=۰$ و $F_W(+\infty)=۱$ است. بنابراین $F_W(w)$ یک تابع توزیع است. با استفاده از فرمول (۳.۳) تابع احتمال W به صورت زیر به دست می‌آید.

w	۳	۴	۵	۶
$f_W(w)$	$\frac{۱}{۳}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۳}$

و بنابراین

$$P(۳< W\leq ۵)=f_W(۴)+f_W(۵)=\frac{۱}{۶}+\frac{۱}{۶}=\frac{۱}{۳}$$

مثال ۶.۶.۳ در یک ظرف ۴ مهره به شماره‌های ۱، ۲، ۳ و ۴ وجود دارد. از این ظرف دو مهره به تصادف و بدون جایگذاری انتخاب می‌کنیم و متغیر تصادفی X را برابر بزرگترین شماره روی دو مهره انتخابی در نظر می‌گیریم تابع احتمال X را به دست آورید.

حل در این مساله فضای نمونه به صورت

$$S=\{\{۱,۲\},\{۱,۳\},\{۱,۴\},\{۲,۳\},\{۲,۴\},\{۳,۴\}\}$$

می‌باشد و $S_X=\{۲,۳,۴\}$ بنابراین

$$f_X(۲)=P(X=۲)=P(\{\{۱,۲\}\})=\frac{۱}{۶}$$

$$f_X(۳)=P(X=۳)=P(\{\{۱,۳\},\{۲,۳\}\})=\frac{۲}{۶}$$

به همین ترتیب داریم که

x	۲	۳	۴
$f_X(x)$	$\frac{۱}{۶}$	$\frac{۲}{۶}$	$\frac{۳}{۶}$

مثال ۷.۶.۳ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر است

$$f_X(x)=\begin{cases} c(۱+x) & ۲<x<۵ \\ ۰ & \text{سایر نقاط} \end{cases}$$

الف- مقدار c را تعیین کنید.

ب- تابع توزیع X را به دست آورید.

حل الف- بایستی $c\geq ۰$ و همچنین

$$۱=\int_{-\infty}^{+\infty} f_X(x)\,dx=\int_۲^۵ c(۱+x)dx=c\left[x+\frac{۱}{۲}x^۲\right]_۲^۵$$

$$=c\left[\left(۵+\frac{۲۵}{۲}\right)-(۲+۲)\right]=c\left(\frac{۲۷}{۲}\right)$$

بنابراین $c=\frac{۲}{۲۷}$.

ب- اگر $x<۲$ آنگاه $F_X(x)=۰$ و اگر $x\geq ۵$ آنگاه $F_X(x)=۱$ و اگر $۲\leq x<۵$ آنگاه

$$F_X(x)=\int_{-\infty}^{x} f_X(t)dt=\int_۲^x \frac{۲}{۲۷}(۱+t)\,dt=\frac{۲}{۲۷}\left[t+\frac{۱}{۲}t^۲\right]_۲^x=\frac{۲}{۲۷}\left[x+\frac{۱}{۲}x^۲-۴\right]$$

بنابراین

$$F_X(x)=\begin{cases} ۰ & x<۲ \\ \frac{۲}{۲۷}\left[\frac{۱}{۲}x^۲+x-۴\right] & ۲\leq x<۵ \\ ۱ & ۵\leq x \end{cases}$$

مثال ۸.۶.۳ اگر متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد، تابع توزیع X را به دست آورید و $P(\frac{۱}{۲}<X<\frac{۳}{۲})$ را محاسبه کنید.

$$f_X(x)=\begin{cases} x & ۰<x<۱ \\ ۲-x & ۱\leq x<۲ \\ ۰ & \text{سایر نقاط} \end{cases}$$

حل اگر $x<۰$ آنگاه $F_X(x)=۰$ و اگر $۰\leq x<۱$ آنگاه

$$F_X(x)=\int_{-\infty}^{x} f_X(t)dt=\int_۰^x t\,dt=\frac{۱}{۲}t^۲\Big|_۰^x=\frac{۱}{۲}x^۲$$

و اگر $۱\leq x<۲$ آنگاه

$$F_X(x)=\int_{-\infty}^{x} f_X(t)dt=\int_۰^۱ t\,dt+\int_۱^x (۲-t)dt=\frac{۱}{۲}+\left[۲t-\frac{۱}{۲}t^۲\right]_۱^x=۲x-\frac{۱}{۲}x^۲-۱$$

بنابراین

$$F_X(x)=\begin{cases} ۰ & x<۰ \\ \frac{۱}{۲}x^۲ & ۰\leq x<۱ \\ ۲x-\frac{۱}{۲}x^۲-۱ & ۱\leq x<۲ \\ ۱ & ۲\leq x \end{cases}$$

در نتیجه

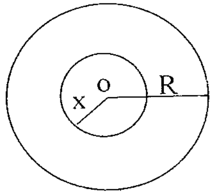
$$P\left(\frac{۱}{۲}<X<\frac{۳}{۲}\right)=F_X\left(\frac{۳}{۲}\right)-F_X\left(\frac{۱}{۲}\right)=\left[۲\left(\frac{۳}{۲}\right)-\frac{۱}{۲}\left(\frac{۳}{۲}\right)^۲-۱\right]-\left[\frac{۱}{۲}\left(\frac{۱}{۲}\right)^۲\right]=\frac{۳}{۴}$$

مثال ۹.۶.۳ نقطه‌ای به تصادف از داخل دایره‌ای به شعاع R انتخاب می‌کنیم و متغیر تصادفی X را برابر فاصله نقطه انتخابی تا مرکز دایره در نظر می‌گیریم.

الف- تابع توزیع X را به دست آورید.

ب- تابع احتمال X را به دست آورید و $P(\frac{R}{۳}<X<\frac{R}{۲})$ را محاسبه کنید.

حل الف- اگر $x<0$ آنگاه $F_X(x)=P(X\le x)=0$ و اگر $0\le x<R$ آنگاه
با توجه به شکل مقابل داریم که



$$F_X(x)=P(X\le x)=\frac{\text{مساحت دایره به شعاع } x}{\text{مساحت دایره به شعاع } R}=\frac{\pi x^2}{\pi R^2}=\left(\frac{x}{R}\right)^2$$

و اگر $x\ge R$ آنگاه $F_X(x)=P(X\le x)=1$ و بنابراین

$$F_X(x)=\begin{cases}0 & x<0\\ \left(\frac{x}{R}\right)^2 & 0\le x<R\\ 1 & R\le x\end{cases}$$

ب- با توجه به رابطه (۵.۳) داریم که

$$f_X(x)=F'_X(x)=\begin{cases}\frac{2x}{R^2} & 0<x<R\\ 0 & \text{سایر نقاط}\end{cases}$$

و در نتیجه

$$P\left(\frac{R}{3}<X<\frac{R}{2}\right)=\int_{\frac{R}{3}}^{\frac{R}{2}}\frac{2x}{R^2}dx=\left[\left(\frac{x}{R}\right)^2\right]_{\frac{R}{3}}^{\frac{R}{2}}=\frac{1}{4}-\frac{1}{9}=\frac{5}{36}$$

مثال ۱۰.۶.۳ میانه متغیر تصادفی X عددی است مانند m که در روابط $P(X\le m)\ge\frac{1}{2}$ و $P(X\ge m)\ge\frac{1}{2}$ صدق کند. اگر X یک متغیر تصادفی پیوسته باشد آنگاه میانه در رابطه $F_X(m)=\int_{-\infty}^{m}f_X(x)dx=\frac{1}{2}$ صدق می‌کند. در هریک از موارد زیر میانه متغیر تصادفی X که دارای تابع احتمال و یا چگالی احتمال زیر است را پیدا کنید.

الف- $f_X(x)=e^{-x}e^{-e^{-x}} \qquad -\infty<x<+\infty$

ب- $f_X(x)=\binom{4}{x}\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{4-x} \qquad x=0,1,2,3,4$

حل الف- متغیر تصادفی X از نوع پیوسته است. بنابراین بایستی $\int_{-\infty}^{m}e^{-x}e^{-e^{-x}}dx=\frac{1}{2}$ باشد. با تغییر متغیر $y=e^{-x}$ داریم که

$$\int_{e^{-m}}^{+\infty}e^{-y}dy=\frac{1}{2}\Rightarrow -e^{-y}\Big|_{e^{-m}}^{+\infty}=\frac{1}{2}\Rightarrow e^{-e^{-m}}=\frac{1}{2}\Rightarrow m=-Ln(Ln2)$$

ب- متغیر تصادفی X از نوع گسسته با تابع احتمال زیر است

x	۰	۱	۲	۳	۴
$f_X(x)$	$\frac{1}{81}$	$\frac{8}{81}$	$\frac{24}{81}$	$\frac{32}{81}$	$\frac{16}{81}$

بنابراین

$$P(X\le 3)=1-P(X=4)=1-\frac{16}{81}=\frac{65}{81}\ge\frac{1}{2}$$
$$P(X\ge 3)=\frac{32}{81}+\frac{16}{81}=\frac{48}{81}\ge\frac{1}{2}$$

و در نتیجه m=۳ است.

مثال ۱۱.۶.۳ اگر متغیر تصادفی X دارای تابع توزیع زیر باشد، تابع چگالی احتمال و $P(X>3)$ را محاسبه کنید

$$F_X(x)=\begin{cases}0 & x<0\\ 1-(1+x)e^{-x} & x\ge 0\end{cases}$$

حل با توجه به رابطه (۵.۳) داریم که

$$f_X(x)=F'_X(x)=-e^{-x}+(1+x)e^{-x}=xe^{-x}$$

در نتیجه

$$f_X(x)=\begin{cases}xe^{-x} & x>0\\ 0 & \text{سایر نقاط}\end{cases}$$

و همچنین

$$P(X>3)=\int_{3}^{+\infty}xe^{-x}dx=(-x-1)e^{-x}\Big|_{3}^{+\infty}=4e^{-3}$$

مثال ۱۲.۶.۳ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}ke^{-2x}(1-e^{-2x}) & x>0\\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار k را تعیین کنید.

ب- تابع توزیع X را به دست آورید و $P(X>1)$ را محاسبه کنید.

حل الف- بایستی $k\ge 0$ و همچنین

$$1=\int_{0}^{+\infty}ke^{-2x}(1-e^{-2x})dx=k\int_{0}^{+\infty}(e^{-2x}-e^{-4x})dx$$
$$=k\left[-\frac{1}{2}e^{-2x}+\frac{1}{4}e^{-4x}\right]_{0}^{+\infty}=k\left(\frac{1}{2}-\frac{1}{4}\right)=k\left(\frac{1}{4}\right)$$

بنابراین $k=4$.

ب- اگر $x<0$ آنگاه $F_X(x)=0$ و اگر $x\ge 0$ آنگاه

$$F_X(x)=\int_{-\infty}^{x} f_X(t)\,dt=\int_{0}^{x} 4e^{-2t}(1-e^{-2t})\,dt$$

$$=4\left[-\frac{1}{2}e^{-2t}+\frac{1}{4}e^{-4t}\right]_{0}^{x}$$

$$=4\left[-\frac{1}{2}e^{-2x}+\frac{1}{4}e^{-4x}+\frac{1}{2}-\frac{1}{4}\right]=1-e^{-2x}(2-e^{-2x})$$

بنابراین

$$F_X(x)=\begin{cases}0 & x<0\\ 1-e^{-2x}(2-e^{-2x}) & x\ge 0\end{cases}$$

و در نتیجه

$$P(X>1)=1-F_X(1)=e^{-2}(2-e^{-2})=0/252$$

مثال ۱۳.۶.۳ اگر متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}k(x^2+y^2) & x=1,2\ ,\ \ y=0,1,2\\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار k را پیدا کنید.

ب- $P(Y<2\mid X=1)$ را محاسبه کنید.

ج- آیا X و Y مستقل هستند؟ چرا؟

حل الف- $1=\sum_{x=1}^{2}\sum_{y=0}^{2}k(x^2+y^2)=k(1+2+5+4+5+8)=k(25)$

بنابراین $k=\frac{1}{25}$ و در نتیجه جدول توزیع احتمالات توأم X و Y به صورت زیر به دست می‌آید.

y \ x	1	2	$f_Y(y)$
0	$\frac{1}{25}$	$\frac{4}{25}$	$\frac{5}{25}$
1	$\frac{2}{25}$	$\frac{5}{25}$	$\frac{7}{25}$
2	$\frac{5}{25}$	$\frac{8}{25}$	$\frac{13}{25}$
$f_X(x)$	$\frac{8}{25}$	$\frac{17}{25}$	

ب- با توجه به رابطه (۱۳.۳) تابع احتمال شرطی $f_{Y\mid X}(y\mid 1)$ به صورت زیر به دست می‌آید

y	0	1	2
$f_{Y\mid X}(y\mid 1)$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{5}{8}$

بنابراین

$$P(Y<2\mid X=1)=\sum_{y=0}^{1}f_{Y\mid X}(y\mid 1)=\frac{1}{8}+\frac{2}{8}=\frac{3}{8}$$

ج- چون $f_X(1)=\frac{8}{25}$، $f_Y(0)=\frac{5}{25}$ و $f_{X,Y}(1,0)=\frac{1}{25}$ در نتیجه

$$f_{X,Y}(1,0)\neq f_X(1)f_Y(0)$$

بنابراین X و Y از یکدیگر مستقل نیستند.

مثال ۱۴.۶.۳ تابع چگالی احتمال توأم متغیرهای تصادفی X و Y به صورت زیر است:

$$f_{X,Y}(x,y)=\begin{cases}cy\,e^{-y(x+1)} & x>0\ ,\ \ y>0\\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار c را تعیین کنید.

ب- $P(XY>1)$ و $P(X<2\mid Y=3)$ را محاسبه کنید.

حل الف-

$$1=\int_{0}^{+\infty}\int_{0}^{+\infty}cye^{-y(x+1)}\,dx\,dy=c\int_{0}^{+\infty}ye^{-y}\left[\int_{0}^{+\infty}e^{-yx}\,dx\right]dy$$

$$=c\int_{0}^{+\infty}ye^{-y}\left(\frac{1}{y}\right)dy=c\int_{0}^{+\infty}e^{-y}\,dy=c\left[-e^{-y}\right]_{0}^{+\infty}=c(1)=c$$

بنابراین c=۱.

ب-

$$P(XY>1)=P\left(X>\frac{1}{Y}\right)=\int_{0}^{+\infty}\int_{\frac{1}{y}}^{+\infty}ye^{-y(x+1)}\,dx\,dy$$

$$=\int_{0}^{+\infty}ye^{-y}\left[\int_{\frac{1}{y}}^{+\infty}e^{-yx}\,dx\right]dy=\int_{0}^{+\infty}ye^{-y}\left(\frac{1}{y}e^{-1}\right)dy$$

$$=e^{-1}\int_{0}^{+\infty}e^{-y}\,dy=e^{-1}$$

برای محاسبه $P(X<2\mid Y=3)$ بایستی ابتدا تابع $f_{X\mid Y}(x\mid y)$ را به دست آوریم.

$$f_Y(y)=\int_{0}^{+\infty}ye^{-y(x+1)}\,dx=ye^{-y}\left[\int_{0}^{+\infty}e^{-yx}\,dx\right]=ye^{-y}\left(\frac{1}{y}\right)=e^{-y}\quad y>0$$

$$f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{ye^{-y(x+1)}}{e^{-y}}=ye^{-yx}$$

بنابراین

$$P(X<2|Y=3)=\int_0^2 f_{X|Y}(x|3)\,dx=\int_0^2 3e^{-3x}\,dx=-e^{-3x}\,|_0^2=1-e^{-6}$$

مثال ۱۵.۶.۳ بر یک طرف یک سکه سالم عدد ۱ و بر طرف دیگر آن عدد ۲ نقاشی شده است. این سکه ۳ بار پرتاب می‌شود. اگر متغیر تصادفی X برابر مجموع دو عدد حاصل از دو پرتاب اول و متغیر تصادفی Y برابر مجموع دو عدد حاصل از دو پرتاب آخر باشد، تابع احتمال توأم X و Y را به دست آورید و $P(X+Y>6)$ را محاسبه کنید. آیا X=Y است؟ آیا $f_X(x)=f_Y(y)$ است؟

حل در اینجا $S_X=S_Y=\{2,3,4\}$ و همچنین

$$f_{X,Y}(2,2)=P(X=2,Y=2)=P(\{1,1,1\})=\frac{1}{8}$$
$$f_{X,Y}(2,3)=P(X=2,Y=3)=P(\{1,1,2\})=\frac{1}{8}$$

با انجام محاسبات مربوط به نقاط دیگر، جدول زیر را به دست می‌آوریم.

$y \backslash x$	۲	۳	۴	$f_Y(y)$
۲	$\frac{1}{8}$	$\frac{1}{8}$	۰	$\frac{2}{8}$
۳	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{4}{8}$
۴	۰	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$
$f_X(x)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	

بنابراین

$$P(X+Y>6)=f_{X,Y}(3,4)+f_{X,Y}(4,3)+f_{X,Y}(4,4)=\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{3}{8}$$

با توجه به جدول توزیع احتمالات توأم X و Y دیده می‌شود که $f_X(x)=f_Y(y)$ اما X برابر Y نیست. در این حالت گویند که متغیرهای تصادفی X و Y هم توزیع هستند.

مثال ۱۶.۶.۳ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند.

$$f_{X,Y}(x,y)=\begin{cases}\frac{c}{y} & 0<x<y\ ,\ 0<y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

الف - مقدار c را تعیین کنید.

ب - $P(X+Y>\frac{1}{2})$ را محاسبه کنید.

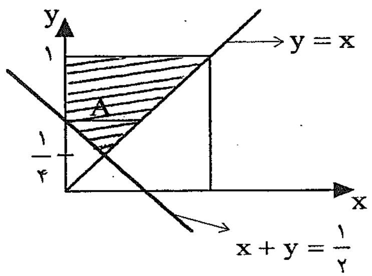
ج - آیا X و Y مستقل هستند؟ چرا؟

حل الف -

$$1=\int_0^1\int_0^y\frac{c}{y}\,dx\,dy=\int_0^1\frac{c}{y}\,[x]_0^y\,dy=\int_0^1 c\,dy=cy\,|_0^1=c$$

بنابراین c=۱.

ب - با توجه به نمودار زیر داریم که



$$P(X+Y>\frac{1}{2})=\iint_A f_{X,Y}(x,y)\,dx\,dy=\int_{\frac14}^{\frac12}\int_{\frac12-y}^{y}\frac{1}{y}\,dx\,dy+\int_{\frac12}^{1}\int_0^y\frac{1}{y}\,dx\,dy$$
$$=\int_{\frac14}^{\frac12}(2-\frac{1}{2y})\,dy+\int_{\frac12}^{1}dy=\left[2y-\frac{1}{2}Lny\right]_{\frac14}^{\frac12}+[y]_{\frac12}^{1}$$
$$=(1+\frac{1}{2}Ln2)-(\frac{1}{2}+\frac{1}{2}Ln4)+\frac{1}{2}=1-\frac{1}{2}Ln2=0/653$$

ج - با توجه به اینکه

$$f_X(x)=\int_x^1\frac{1}{y}\,dy=Lny\,|_x^1=-Lnx \qquad 0<x<1$$
$$f_Y(y)=\int_0^y\frac{1}{y}\,dx=\frac{x}{y}\,|_0^y=1 \qquad 0<y<1$$

بنابراین $f_{X,Y}(x,y)\neq f_X(x)f_Y(y)$ پس X و Y از یکدیگر مستقل نیستند.

مثال ۱۷.۶.۳ فرض کنید برای متغیرهای تصادفی X و Y داشته باشیم که

$$f_X(x)=\begin{cases}c_1x & x=1,2\\ 0 & \text{سایر نقاط}\end{cases}$$

$$f_{Y|X}(y|x)=\begin{cases} c_2 \binom{x}{y} (\frac{1}{2})^x & x=1,2\ ,\ y=0,1,\dots,x \\ 0 & \text{سایر نقاط}\end{cases}$$

الف – مقدار c_1 و c_2 را تعیین کنید.

ب – $P(Y<2|X=2)$ و $P(X>1|Y=1)$ را محاسبه کنید.

حل الف – $f_X(x)$ یک تابع احتمال است بنابراین

$$1=\sum_{x=1}^{2} f_X(x)=c_1+2c_1=3c_1 \Rightarrow c_1=\frac{1}{3}$$

و برای هر مقدار ثابت x تابع $f_{Y|X}(y|x)$ یک تابع احتمال است بنابراین برای x=1 داریم که

$$1=\sum_{y=0}^{1} f_{Y|X}(y|1)=\sum_{y=0}^{1} c_2\binom{1}{y}(\frac{1}{2})=c_2(\frac{1}{2})+c_2(\frac{1}{2})=c_2 \Rightarrow c_2=1$$

ب – تابع احتمال توأم X و Y به صورت زیر به دست می‌آید.

$$f_{X,Y}(x,y)=f_X(x)f_{Y|X}(y|x)=\frac{1}{3}x\binom{x}{y}(\frac{1}{2})^x \qquad x=1,2\ ,\ y=0,1,\dots,x$$

که فرم جدولی آن و همچنین توابع احتمال شرطی $f_{Y|X}(y|2)$ و $f_{X|Y}(x|1)$ به صورت زیر به دست می‌آیند.

$y \backslash x$	۱	۲	$f_Y(y)$
۰	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$
۱	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$
۲	۰	$\frac{1}{6}$	$\frac{1}{6}$
$f_X(x)$	$\frac{2}{6}$	$\frac{4}{6}$	

y	۰	۱	۲
$f_{Y\|X}(y\|2)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

x	۱	۲
$f_{X\|Y}(x\|1)$	$\frac{1}{3}$	$\frac{2}{3}$

بنابراین

$$P(Y<2|X=2)=\sum_{y=0}^{1} f_{Y|X}(y|2)=\frac{1}{4}+\frac{2}{4}=\frac{3}{4}$$

$$P(X>1|Y=1)=f_{X|Y}(2|1)=\frac{2}{3}$$

مثال ۱۸.۶.۳ فرض کنید که متغیرهای تصادفی Xو Y دارای تابع چگالی احتمال توأم زیر باشند که در آن $\lambda>0$ مقداری ثابت است.

$$f_{X,Y}(x,y)=\begin{cases} cx\, e^{-\lambda(x+y)} & x>0\ ,\ y>0\\ 0 & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید.

ب – $P(X>2)$ را محاسبه کنید.

ج – آیا X و Y مستقل هستند؟ چرا؟

حل الف – با استفاده از انتگرال‌گیری به روش جزء به جزء داریم که

$$1=\int_0^{+\infty}\int_0^{+\infty} cxe^{-\lambda(x+y)}dxdy = c\left[\int_0^{+\infty} xe^{-\lambda x}dx\right]\left[\int_0^{+\infty}e^{-\lambda y}dy\right]$$

$$=c\left[(-\frac{x}{\lambda}-\frac{1}{\lambda^2})e^{-\lambda x}\right]_0^{+\infty}\left[-\frac{1}{\lambda} e^{-\lambda y}\right]_0^{+\infty}=c(\frac{1}{\lambda^2})(\frac{1}{\lambda})=\frac{c}{\lambda^3}$$

بنابراین $c=\lambda^3$

ب – برای محاسبه احتمال، ابتدا $f_X(x)$ را به دست می‌آوریم.

$$f_X(x)=\int_0^{+\infty}\lambda^3 xe^{-\lambda(x+y)}dy=\lambda^2xe^{-\lambda x}\left[\int_0^{+\infty}\lambda e^{-\lambda y}dy\right]=\lambda^2 xe^{-\lambda x}\quad x>0$$

بنابراین

$$P(X>2)=\int_2^{+\infty}\lambda^2 xe^{-\lambda x}dx=\left[(-\lambda x-1)e^{-\lambda x}\right]_2^{+\infty}=(2\lambda+1)e^{-2\lambda}$$

ج – با توجه به اینکه $f_X(x)=\lambda^2xe^{-\lambda x}\ \ x>0$ و همچنین

$$f_Y(y)=\int_0^{+\infty}\lambda^3xe^{-\lambda(x+y)}dx=\lambda e^{-\lambda y}\left[(-\lambda x-1)e^{-\lambda x}\right]_0^{+\infty}=\lambda e^{-\lambda y}\quad y>0$$

پس $f_{X,Y}(x,y)=f_X(x)f_Y(y)$ و در نتیجه X و Y از یکدیگر مستقل هستند.

مثال ۱۹.۶.۳ جعبه‌ای شامل ۴ ترانزیستور است و می‌دانیم ۲ عدد از آنها معیوب هستند. ترانزیستورها را یک به یک آزمایش نموده تا هر دو معیوب مشخص شوند و سپس توقف می‌کنیم. اگر X برابر تعداد آزمایشهای لازم تا مشخص شدن اولین معیوب و Y برابر تعداد آزمایشهای اضافی تا مشخص شدن دومین معیوب باشند، تابع احتمال توأم X و Y را به دست آورید و $P(X\le 2|Y=1)$ را محاسبه کنید.

حل در اینجا $S_X=S_Y=\{1,2,3\}$. اگر مشاهده ترانزیستور سالم را با S و خراب را با D نمایش دهیم در این صورت

$f_{X,Y}(۱,۱)=P(DD)=\frac{۲}{۴}\times\frac{۱}{۳}=\frac{۱}{۶}$ $\qquad f_{X,Y}(۱,۲)=P(DSD)=\frac{۲}{۴}\times\frac{۲}{۳}\times\frac{۱}{۲}=\frac{۱}{۶}$

$f_{X,Y}(۱,۳)=P(DSSD)=\frac{۲}{۴}\times\frac{۲}{۳}\times\frac{۱}{۲}\times ۱=\frac{۱}{۶}$

با محاسبه احتمالات مربوط به نقاط دیگر، جدول توزیع احتمالات توأم و تابع احتمال $f_{X\mid Y}(x\mid ۱)$ به صورت زیر به دست می‌آیند.

$y \backslash x$	۱	۲	۳	$f_Y(y)$
۱	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۱}{۶}$	$\frac{۳}{۶}$
۲	$\frac{۱}{۶}$	$\frac{۱}{۶}$	۰	$\frac{۲}{۶}$
۳	$\frac{۱}{۶}$	۰	۰	$\frac{۱}{۶}$
$f_X(x)$	$\frac{۳}{۶}$	$\frac{۲}{۶}$	$\frac{۱}{۶}$	

x	۱	۲	۳
$f_{X\mid Y}(x\mid ۱)$	$\frac{۱}{۳}$	$\frac{۱}{۳}$	$\frac{۱}{۳}$

بنابراین

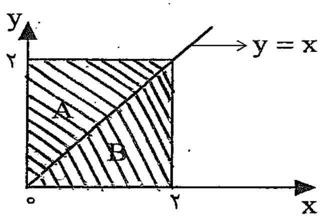
$$P(X\le ۲\mid Y=۱)=\sum_{x=۱}^{۲} f_{X\mid Y}(x\mid ۱)=\frac{۱}{۳}+\frac{۱}{۳}=\frac{۲}{۳}$$

مثال ۲۰.۶.۳ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند.

$$f_{X,Y}(x,y)=\begin{cases} e^{-y}(۱-e^{-x}) & ۰<x<y<+\infty \\ e^{-x}(۱-e^{-y}) & ۰<y<x<+\infty \\ ۰ & \text{سایر نقاط} \end{cases}$$

مطلوب است محاسبه احتمال $P(X\le ۲, Y\le ۲)$ و $P(X\le ۲\mid Y=۳)$.

حل با توجه به نمودار زیر داریم که



$$P(X\le ۲, Y\le ۲)=\iint_A e^{-y}(۱-e^{-x})\,dx\,dy+\iint_B e^{-x}(۱-e^{-y})\,dx\,dy$$

$$=\int_۰^۲\int_۰^y e^{-y}(۱-e^{-x})\,dx\,dy+\int_۰^۲\int_۰^x e^{-x}(۱-e^{-y})\,dy\,dx$$

$$=۲\int_۰^۲\int_۰^y e^{-y}(۱-e^{-x})\,dx\,dy=۲\int_۰^۲ e^{-y}\left[x+e^{-x}\right]_۰^y dy$$

$$=۲\int_۰^۲ e^{-y}(y+e^{-y}-۱)\,dy=۲\left[-ye^{-y}-\frac{۱}{۲}e^{-۲y}\right]_۰^۲$$

$$=۲\left[-۲e^{-۲}-\frac{۱}{۲}e^{-۴}+\frac{۱}{۲}\right]=۱-۴e^{-۲}-e^{-۴}=۰/۴۴$$

برای محاسبه $P(X\le ۲\mid Y=۳)$ ابتدا تابع $f_{X\mid Y}(x\mid y)$ را به دست می‌آوریم.

$$f_Y(y)=\int_۰^y e^{-y}(۱-e^{-x})\,dx+\int_y^{+\infty} e^{-x}(۱-e^{-y})\,dx$$

$$=e^{-y}\left[x+e^{-x}\right]_۰^y+(۱-e^{-y})\left[-e^{-x}\right]_y^{+\infty}$$

$$=e^{-y}(y+e^{-y}-۱)+(۱-e^{-y})e^{-y}=ye^{-y} \qquad ۰<y<+\infty$$

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\begin{cases} \dfrac{۱-e^{-x}}{y} & ۰<x<y<+\infty \\ \dfrac{e^{-x}(۱-e^{-y})}{ye^{-y}} & ۰<y<x<+\infty \\ ۰ & \text{سایر نقاط} \end{cases}$$

بنابراین

$$P(X\le ۲\mid Y=۳)=\int_۰^۲ f_{X\mid Y}(x\mid ۳)\,dx=\int_۰^۲\frac{۱-e^{-x}}{۳}dx=\frac{۱}{۳}\left[x+e^{-x}\right]_۰^۲$$

$$=\frac{۱+e^{-۲}}{۳}=۰/۳۷۸$$

مثال ۲۱.۶.۳ جعبه A شامل ۲ مهره سفید و ۳ مهره سیاه و جعبه B شامل ۱ مهره سفید و ۲ مهره سیاه است. بدون جایگذاری ۲ مهره از جعبه A به جعبه B منتقل می‌کنیم و سپس از جعبه B بدون جایگذاری ۲ مهره خارج می‌کنیم. اگر X برابر تعداد مهره‌های سفید منتقل شده به جعبه B و Y برابر تعداد مهره‌های سفید خارج شده از جعبه B باشند، تابع احتمال توأم X و Y را به دست آورده و $P(X+Y>۲)$ را محاسبه کنید.

حل در اینجا $S_X=S_Y=\{۰,۱,۲\}$ و همچنین

$$f_{X,Y}(۰,۰)=P(X=۰,Y=۰)=P(X=۰)P(Y=۰\mid X=۰)$$

$$=\frac{\binom{۳}{۲}}{\binom{۵}{۲}}\times\frac{\binom{۴}{۲}}{\binom{۵}{۲}}=۰/۱۸$$

$$f_{X,Y}(۰,۱)=P(X=۰,Y=۱)=P(X=۰)P(Y=۱\mid X=۰)$$

$$=\frac{\binom{۳}{۲}}{\binom{۵}{۲}}\times\frac{\binom{۱}{۱}\binom{۴}{۱}}{\binom{۵}{۲}}=۰/۱۲$$

$$f_{X,Y}(۱,۰)=P(X=۱,Y=۰)=P(X=۱)P(Y=۰\mid X=۱)$$

$$=\frac{\binom{۲}{۱}\binom{۳}{۱}}{\binom{۵}{۲}}\times\frac{\binom{۳}{۲}}{\binom{۵}{۲}}=۰/۱۸$$

با انجام محاسبات مشابه جدول توزیع احتمالات توأم X و Y به صورت زیر به دست می‌آید.

y \ x	۰	۱	۲
۰	۰/۱۸	۰/۱۸	۰/۰۱
۱	۰/۱۲	۰/۳۶	۰/۰۶
۲	۰	۰/۰۶	۰/۰۳

بنابراین

$$P(X+Y>۲)=f_{X,Y}(۲,۱)+f_{X,Y}(۱,۲)+f_{X,Y}(۲,۲)$$

$$=۰/۰۶+۰/۰۶+۰/۰۳=۰/۱۵$$

مثال ۲۲.۶.۳ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند.

$$f_{X,Y}(x,y)=\begin{cases}cx(x-y) & -x<y<x\ ,\ ۰<x<۲\\ ۰ & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید.

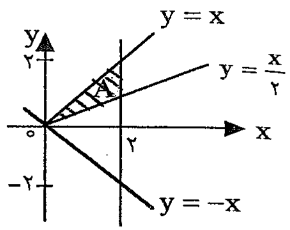
ب – $f_Y(y)$ و $P(۲Y>X)$ و $P(Y<\frac{۱}{۲}\mid X=۱)$ را محاسبه کنید.

حل الف –

$$۱=c\int_۰^۲\int_{-x}^{x}x(x-y)\,dy\,dx=c\int_۰^۲\left[x^۲y-\frac{x}{۲}y^۲\right]_{-x}^{x}dx$$

$$=c\int_۰^۲ ۲x^۳dx=\frac{c}{۲}x^۴\Big|_۰^۲=۸c\quad\Rightarrow\quad c=\frac{۱}{۸}$$

ب – با توجه به نمودار زیر داریم که



$$f_Y(y)=\begin{cases}\int_y^۲\frac{۱}{۸}x(x-y)\,dx & ۰<y<۲\\ \int_{-y}^۲\frac{۱}{۸}x(x-y)\,dx & -۲<y<۰\end{cases}$$

$$=\begin{cases}\frac{۱}{۸}\left[\frac{۱}{۳}x^۳-\frac{y}{۲}x^۲\right]_y^۲ & ۰<y<۲\\ \frac{۱}{۸}\left[\frac{۱}{۳}x^۳-\frac{y}{۲}x^۲\right]_{-y}^۲ & -۲<y<۰\end{cases}$$

بنابراین

$$f_Y(y)=\begin{cases}\frac{۱}{۸}\left[\frac{۵}{۶}y^۳-۲y+\frac{۸}{۳}\right] & -۲<y<۰\\ \frac{۱}{۸}\left[\frac{۱}{۶}y^۳-۲y+\frac{۸}{۳}\right] & ۰<y<۲\\ ۰ & \text{سایر نقاط}\end{cases}$$

با توجه به نمودار بالا داریم که

$$P(۲Y>X)=\int_۰^۲\int_{\frac{x}{۲}}^{x}\frac{۱}{۸}x(x-y)\,dy\,dx=\int_۰^۲\frac{۱}{۸}\left[x^۲y-\frac{x}{۲}y^۲\right]_{\frac{x}{۲}}^{x}dx$$

$$=\frac{۱}{۶۴}\int_۰^۲x^۳dx=\frac{۱}{۶۴}\left[\frac{۱}{۴}x^۴\right]_۰^۲=\frac{۱}{۱۶}$$

برای محاسبه $P(Y<\frac{۱}{۲}\mid X=۱)$ بایستی ابتدا $f_{Y\mid X}(y\mid x)$ را محاسبه کنیم.

$$f_X(x)=\int_{-x}^{x}\frac{۱}{۸}x(x-y)\,dy=\frac{۱}{۸}\left[x^۲y-\frac{x}{۲}y^۲\right]_{-x}^{x}=\frac{۱}{۴}x^۳\qquad ۰<x<۲$$

$$f_{Y\mid X}(y\mid x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{\frac{۱}{۸}x(x-y)}{\frac{۱}{۴}x^۳}=\frac{۱}{۲}\,\frac{x-y}{x^۲}\qquad ۰<x<۲\ ,\ -x<y<x$$

بنابراین

$$P(Y<\tfrac{۱}{۲}\mid X=۱)=\int_{-۱}^{\frac{۱}{۲}}f_{Y\mid X}(y\mid ۱)\,dy=\int_{-۱}^{\frac{۱}{۲}}\frac{۱}{۲}(۱-y)\,dy$$

$$= \frac{1}{2}\left[y - \frac{1}{2}y^{2}\right]_{-1}^{\frac{1}{2}} = \frac{15}{16}$$

مثال ۲۳.۶.۳ دو جعبه وجود دارند که جعبه اول دارای ۳ مهره سفید و ۲ مهره قرمز است و جعبه دوم دارای ۲ مهره سفید و ۳ مهره قرمز است. از جعبه اول یک مهره به تصادف انتخاب و در داخل جعبه دوم قرار می‌دهیم و سپس از جعبه دوم ۲ مهره یک به یک و بدون جایگذاری انتخاب می‌کنیم. اگر X برابر تعداد مهره‌های سفید انتخاب شده از جعبه دوم و Y برابر تعداد مهره‌های قرمز انتخاب شده از جعبه دوم باشند، تابع احتمال توأم X و Y را به دست آورید.

حل در اینجا $S_X = S_Y = \{0, 1, 2\}$ و همواره داریم که $X+Y=2$، چون از جعبه دوم یا مهره سفید و یا مهره قرمز خارج می‌شود. اگر W_1 و R_1 به ترتیب پیشامد انتخاب مهره سفید و مهره قرمز از جعبه اول و WW_2، RR_2 و WR_2 به ترتیب پیشامد انتخاب دو مهره سفید، دو مهره قرمز و انتخاب یک مهره سفید و یک مهره قرمز از جعبه دوم باشند در این صورت

$$f_{X,Y}(0,2) = P(RR_2) = P(W_1 \cap RR_2) + P(R_1 \cap RR_2)$$
$$= \frac{3}{5} \times \frac{\binom{3}{2}}{\binom{6}{2}} + \frac{2}{5} \times \frac{\binom{4}{2}}{\binom{6}{2}} = \frac{21}{75}$$
$$f_{X,Y}(1,1) = P(WR_2) = P(W_1 \cap WR_2) + P(R_1 \cap WR_2)$$
$$= \frac{3}{5} \times \frac{\binom{3}{1}\binom{3}{1}}{\binom{6}{2}} + \frac{2}{5} \times \frac{\binom{2}{1}\binom{4}{1}}{\binom{6}{2}} = \frac{43}{75}$$

و به همین ترتیب $f_{X,Y}(2,0) = P(WW_2) = \frac{11}{75}$ و در نتیجه تابع احتمال توأم X و Y به صورت زیر به دست می‌آید.

y \ x	0	1	2
0	0	0	$\frac{11}{75}$
1	0	$\frac{43}{75}$	0
2	$\frac{21}{75}$	0	0

مثال ۲۴.۶.۳ فرض کنید متغیرهای تصادفی X، Y و Z دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y,Z}(x,y,z) = \begin{cases} k(x+y)e^{-z} & 0<x<1 \ , \ 0<y<2 \ , \ z>0 \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – مقدار k را تعیین کنید و $P(X<Y, Z>1)$ را محاسبه کنید.

ب – نشان دهید که هر سه متغیر X، Y و Z از یکدیگر مستقل نیستند اما Z از X و Z از Y مستقل می‌باشند و X از Y مستقل نمی‌باشد.

حل الف –
$$1 = \int_0^1\int_0^2\int_0^{+\infty} k(x+y)e^{-z}\,dz\,dy\,dx = k\int_0^1\int_0^2 (x+y)\,[-e^{-z}]_0^{+\infty}\,dy\,dx$$
$$= k\int_0^1\int_0^2 (x+y)\,dy\,dx = k\int_0^1 [xy+\tfrac{1}{2}y^2]_0^2\,dx = k\int_0^1 [2x+2]\,dx$$
$$= k[x^2+2x]_0^1 = k(3)$$

بنابراین $k=\frac{1}{3}$.

$$P(X<Y, Z>1) = \int_0^1\int_x^2\int_1^{+\infty} \frac{1}{3}(x+y)e^{-z}\,dz\,dy\,dx$$
$$= \int_0^1\int_x^2 \frac{1}{3}(x+y)\,[-e^{-z}]_1^{+\infty}\,dy\,dx$$
$$= \frac{1}{3}e^{-1}\int_0^1\int_x^2 (x+y)\,dy\,dx = \frac{1}{3}e^{-1}\int_0^1 \left[xy+\frac{1}{2}y^2\right]_x^2 dx$$
$$= \frac{1}{3}e^{-1}\int_0^1 (2x+2-\frac{3}{2}x^2)\,dx = \frac{1}{3}e^{-1}(\frac{5}{2}) = \frac{5}{6}e^{-1} = 0/307$$

ب –
$$f_{X,Z}(x,z) = \int_0^2 \frac{1}{3}(x+y)e^{-z}\,dy = \frac{1}{3}e^{-z}\left[xy+\frac{1}{2}y^2\right]_0^2 = \frac{2x+2}{3}e^{-z}$$
$$0<x<1 \ , \ z>0$$

$$f_X(x) = \int_0^{+\infty} \frac{2x+2}{3}e^{-z}\,dz = \frac{2x+2}{3} \qquad 0<x<1$$
$$f_Z(z) = \int_0^1 \frac{2x+2}{3}e^{-z}\,dx = e^{-z}\left[\frac{x^2+2x}{3}\right]_0^1 = e^{-z} \qquad z>0$$

به همین ترتیب به سادگی به دست می‌آوریم که

$$f_{Y,Z}(y,z) = \frac{2y+1}{6}e^{-z} \qquad 0<y<2 \ , \ z>0$$

$$f_Y(y)=\frac{2y+1}{6} \qquad 0<y<2$$

$$f_{X,Y}(x,y)=\frac{x+y}{3} \qquad 0<x<1\ ,\ 0<y<2$$

بنابراین $f_{X,Y,Z}(x,y,z)\neq f_X(x)f_Y(y)f_Z(z)$ پس X و Y و Z از یکدیگر مستقل نیستند اما $f_{X,Z}(x,z)=f_X(x)f_Z(z)$ و $f_{Y,Z}(y,z)=f_Y(y)f_Z(z)$ پس X و Z و همچنین Y و Z از یکدیگر مستقل هستند و $f_{X,Y}(x,y)\neq f_X(x)f_Y(y)$ یعنی X و Y از یکدیگر مستقل نیستند.

۷.۳ تمرینات

۱ یک تاس سالم را آنقدر پرتاب می‌کنیم تا برای اولین بار عدد ۶ بیاید

الف - اگر A پیشامد مشاهدهٔ عدد ۶ و B پیشامد مشاهدهٔ عدد غیر از ۶ باشد، فضای نمونه و احتمال هر پیشامد ساده را پیدا کنید.

ب - فرض کنید X شمارهٔ پرتابهای لازم برای مشاهدهٔ اولین ۶ باشد. S_X را پیدا کنید. آیا X گسسته است؟

ج - تابع احتمال و تابع توزیع X را به دست آورید و احتمال اینکه X مضرب ۳ باشد را بیابید.

۲ سکه سالمی را سه مرتبه پرتاب می‌کنیم. اگر متغیر تصادفی X را برابر تفاضل تعداد خطهای به دست آمده از تعداد شیرها در نظر بگیریم، تابع احتمال X را به دست آورید و $P(X>-1)$ را محاسبه کنید.

۳ محموله‌ای شامل ۵ پروژکتور است که دوتای آن آسیب دیده است. به طور تصادفی ۳ پروژکتور این محموله خریداری می‌شود. اگر تعداد پروژکتورهای آسیب دیده که خریداری شده‌اند را با متغیر تصادفی X نشان دهیم، تابع احتمال و تابع توزیع X را به دست آورید.

۴ کدامیک از توابع زیر می‌توانند تابع احتمال یک متغیر تصادفی گسسته X باشند؟ چرا؟

الف - $f_X(x)=\frac{x}{15} \qquad x=0,1,2,3,4,5$

ب - $f_X(x)=\frac{5-x^2}{6} \qquad x=0,1,2,3$

ج - $f_X(x)=\binom{3}{x}\left(\frac{1}{2}\right)^3 \qquad x=0,1,2,3$

د - $f_X(x)=\frac{x+1}{25} \qquad x=1,2,3,4,5$

۵ سه تاس سالم را با هم پرتاب می‌کنیم. اگر متغیر تصادفی X برابر مجموع اعداد روی ۳ تاس مشاهده شده باشد، تابع احتمال X را به دست آورید و احتمال اینکه مجموع اعداد روی ۳ تاس حداقل ۱۵ باشد را بیابید.

۶ دو کیسه هر کدام محتوی ۴ مهره است که از یک تا چهار شماره گذاری شده‌اند. یک مهره از کیسهٔ اول و یک مهره از کیسه دوم بیرون می‌آوریم.

الف - اگر متغیر تصادفی X مجموع دو رقم نوشته شده بر روی این دو مهره باشد، تابع احتمال و تابع توزیع X را به دست آورید.

ب - اگر متغیر تصادفی Y اختلاف دو رقم نوشته شده بر روی این دو مهره باشد، تابع احتمال و تابع توزیع Y را به دست آورید.

۷ دو جعبه وجود دارند که جعبهٔ اول شامل ۳ مهرهٔ قرمز و ۴ مهرهٔ سیاه است و جعبهٔ دوم شامل ۲ مهرهٔ قرمز و ۳ مهرهٔ سیاه است. یک مهره از جعبهٔ اول به تصادف انتخاب کرده و آن را در جعبهٔ دوم قرار می‌دهیم و سپس از جعبهٔ دوم ۲ مهره به تصادف خارج می‌کنیم. اگر متغیر تصادفی X برابر تعداد مهره‌های قرمز انتخاب شده از جعبهٔ دوم باشد، تابع احتمال X و تابع توزیع X را به دست آورید و $P(X\geq 1)$ را محاسبه کنید.

۸ از داخل مربع ABCD به ضلع $\sqrt{2}$ واحد نقطه‌ای را به تصادف انتخاب می‌کنیم و متغیر تصادفی X را برابر فاصلهٔ نقطهٔ انتخابی تا قطر AC در نظر می‌گیریم.

الف - S_X را پیدا کنید. آیا X پیوسته است؟ چرا؟

ب - تابع توزیع X را به دست آورید و از روی آن تابع چگالی احتمال X را بیابید.

ج - $P(0\leq X\leq \frac{1}{2})$ را محاسبه کنید.

۹ نشان دهید که تابع زیر تابع توزیع یک متغیر تصادفی پیوستهٔ X است و $P(X\leq\frac{\pi}{4})$ و میانهٔ

توزیع را بیابید.

$$F_X(x)=\begin{cases}0 & x<0\\ 1-\cos x & 0\le x<\frac{\pi}{2}\\ 1 & \frac{\pi}{2}\le x\end{cases}$$

۱۰ تابع چگالی احتمال متغیر تصادفی X به صورت زیر داده شده است

$$f_X(x)=\begin{cases}kx^2 & -k<x<k\\ 0 & \text{سایر نقاط}\end{cases}$$

الف – مقدار k را تعیین کنید.

ب – تابع توزیع X را به دست آورید و $P(|X|\le 1)$, $P([X]=-1)$ را محاسبه کنید.

۱۱ مقدار k را چنان تعیین کنید که تابع زیر یک تابع چگالی احتمال باشد.

$$f_X(x)=\frac{k}{1+x^2}\qquad -\infty<x<+\infty$$

۱۲ اگر تابع توزیع یک متغیر تصادفی X به صورت زیر باشد:

$$F_X(x)=\begin{cases}0 & x<2\\ c-\frac{4}{x^2} & x\ge 2\end{cases}$$

الف – مقدار c را تعیین کنید و تابع چگالی احتمال X را به دست آورید.

ب – $P(X<3)$ و $P(4<X\le 5)$ را محاسبه کنید.

۱۳ کدامیک از توابع زیر یک تابع توزیع می‌باشند:

$$F_Y(y)=\begin{cases}0 & y<-1\\ 1-y^2 & -1\le y<\frac{1}{2}\\ \frac{1}{2}+y^2 & \frac{1}{2}\le y<1\\ 1 & 1\le y\end{cases}$$

$$F_X(x)=\begin{cases}0 & x<0\\ x^2 & 0\le x<\frac{1}{2}\\ 1-3(1-x)^2 & \frac{1}{2}\le x<1\\ 1 & 1\le x\end{cases}$$

۱۴ مقدار c را چنان تعیین کنید که توابع زیر یک تابع چگالی احتمال باشند و در هر حالت تابع توزیع مربوطه را به دست آورید.

$$f_X(x)=\begin{cases}c(x+1) & -1<x\le 3\\ 4c & 3<x<4\\ 0 & \text{سایر نقاط}\end{cases}$$

$$f_Y(y)=\begin{cases}cy & 0<y\le 1\\ c & 1\le y\le 2\\ cy+3c & 2\le y<3\\ 0 & \text{سایر نقاط}\end{cases}$$

۱۵ مقدار نانی که (بر حسب صد کیلوگرم) یک نانوا در طول روز می‌فروشد یک متغیر تصادفی X با تابع چگالی احتمال زیر است

$$f_X(x)=\begin{cases}cx & 0<x<3\\ c(6-x) & 3\le x<6\\ 0 & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید و تابع توزیع X را به دست آورید.

ب – اگر در یک روز بدانیم که بیشتر از ۳۰۰ کیلو نان فروخته شده، احتمال اینکه در آن روز بین ۱۵۰ و ۴۵۰ کیلو نان فروخته شود را بیابید.

۱۶ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد که در آن α و β مقادیر ثابتی هستند $\beta>0$ و $-\infty<\alpha<+\infty$ است.

$$f_X(x)=\begin{cases}c\left[1-\left|\frac{x-\alpha}{\beta}\right|\right] & \alpha-\beta<x<\alpha+\beta\\ 0 & \text{سایر نقاط}\end{cases}$$

الف – مقدار c را تعیین کنید.

ب – تابع توزیع X را به دست آورید.

۱۷ تابع احتمال توأم متغیرهای تصادفی X و Y به صورت زیر داده شده است:

$$f_{X,Y}(x,y)=\begin{cases} c(x+y) & x=0,1,2,3 \ , \ y=0,1,2 \\ 0 & \text{سایر نقاط} \end{cases}$$

مقدار c را تعیین کنید و احتمالات $P(XY\le 2)$, $P(X\le 1 \mid Y=1)$ را محاسبه کنید.

۱۸ تابع چگالی احتمال توأم متغیرهای تصادفی X و Y به صورت زیر می‌باشد:

$$f_{X,Y}(x,y)=\begin{cases} cx(1+3y^2) & 0<x<2 \ , \ 0<y<1 \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – مقدار c را تعیین کنید و احتمالات زیر را محاسبه کنید.

$$P\left(0<X<1, \frac{1}{4}<Y<\frac{1}{2}\right) \ , \ P(X+Y\ge 2)$$

ب – آیا X و Y مستقل هستند؟ چرا؟

۱۹ از یک ظرف میوه که شامل ۳ پرتقال، ۲ سیب و ۳ موز است چهار میوه را به تصادف انتخاب می‌کنیم. اگر X تعداد پرتقالهای انتخابی و Y تعداد سیبهای انتخابی باشند،

الف – تابع احتمال توأم X و Y را به دست آورید.

ب – $P(2X+Y<3)$, $P(Y\le 1 \mid X=2)$ را محاسبه کنید.

۲۰ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} \dfrac{6-x-y}{8} & 0<x<2 \ , \ 2<y<4 \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – نشان دهید که X و Y از یکدیگر مستقل نیستند.

ب – $P(2<Y<3 \mid X=1)$ را محاسبه کنید.

۲۱ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} cx & 0<x<1 \ , \ 0<y<1-x \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – مقدار c را تعیین کنید.

ب – توابع چگالی احتمال $f_X(x)$ و $f_Y(y)$ و توابع چگالی احتمال شرطی $f_{X|Y}(x \mid y)$ و $f_{Y|X}(y \mid x)$ را به دست آورید.

۲۲ فرض کنید برای متغیرهای تصادفی X و Y داشته باشیم که

$$f_X(x)=\frac{e^{-\mu}\mu^x}{x!} \qquad x=0,1,2,\dots$$

$$f_{Y|X}(y \mid x)=\binom{x}{y} p^y(1-p)^{x-y} \qquad y=0,1,\dots,x$$

که در آن $\mu>0$ و $0<p<1$ مقادیر ثابتی هستند. نشان دهید که تابع احتمال کناری Y برابر است با

$$f_Y(y)=\frac{e^{-\mu p}(\mu p)^y}{y!} \qquad y=0,1,2,\dots$$

۲۳ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} k & |x^2-y^2|<1 \\ 0 & \text{سایر نقاط} \end{cases}$$

مقدار k را تعیین کرده و توابع چگالی احتمال حاشیه‌ای X و Y را به دست آورید.

۲۴ دو تاس سالم با هم پرتاب می‌شوند. تابع احتمال توأم X و Y را در هر یک از حالات زیر به دست آورید:

الف – X برابر بزرگترین عدد مشاهده شده روی دو تاس و Y برابر مجموع اعداد مشاهده شده روی دو تاس باشند.

ب – X برابر عدد مشاهده شده روی تاس اول و Y برابر بزرگترین عدد مشاهده شده روی دو تاس باشند.

ج – X برابر کوچکترین و Y برابر بزرگترین عدد مشاهده شده روی دو تاس باشند.

۲۵ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} c\left(x^2+\dfrac{xy}{2}\right) & 0<x<1 \ , \ 0<y<2 \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – مقدار c را تعیین کنید و $f_X(x)$ را به دست آورید.

ب – $P(X>Y)$, $P\left(Y>\frac{1}{2} \mid X<\frac{1}{2}\right)$ را محاسبه کنید.

۲۶ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} c(y^2-x^2)e^{-y} & -y\le x\le y \quad , \quad 0<y<\infty \\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار c را تعیین کرده و توابع $f_{X|Y}(x|y)$, $f_{Y|X}(y|x)$ را به دست آورید.

ب- $P(Y>2X)$ را محاسبه کنید.

۲۷ جعبه‌ای شامل ۳ مهره است که بر روی آنها شماره‌های، ۱، ۲ و ۳ نوشته شده است. ۲ مهره یک به یک و بدون جایگذاری از این جعبه خارج می‌کنیم. اگر متغیر تصادفی X را برابر شمارهٔ اولین مهرهٔ انتخابی از جعبه و متغیر تصادفی Y را برابر شمارهٔ بزرگتر در بین دو مهرهٔ انتخابی در نظر بگیریم، تابع احتمال توأم X و Y را به دست آورید و $P(X\le 2 \mid Y=3)$ را محاسبه کنید. آیا X و Y مستقل هستند؟ چرا؟

۲۸ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} \frac{1}{2}xy & 0<x<y<a \\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار a را تعیین کنید و $P(X+Y\le 1)$, $P(0<X<\frac{1}{2} \mid Y=1)$ را محاسبه کنید.

۲۹ فرض کنید برای دو متغیر تصادفی پیوستهٔ X و Y داشته باشیم که

$$f_X(x)=\begin{cases} a & 0<x<1 \\ 0 & \text{سایر نقاط}\end{cases}$$

$$f_{Y|X}(y\mid x)=\begin{cases} b & 0<x<1 \quad , \quad x<y<x+1 \\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار a و b را تعیین کنید و $P(X+Y<1)$ را محاسبه کنید.

ب- $f_Y(y)$ را به دست آورید. آیا X و Y مستقل هستند؟ چرا؟

۳۰ ظرف A شامل یک مهرهٔ سفید و یک مهرهٔ سیاه و ظرف B شامل دو مهرهٔ سفید و دو مهرهٔ سیاه است. به تصادف و با جایگذاری دو مهره از A خارج می‌کنیم، اگر هم رنگ باشند یک مهره سفید و اگر هم رنگ نباشند یک مهرهٔ سیاه به ظرف B اضافه می‌کنیم و سپس به تصادف از ظرف B یک مهره خارج می‌کنیم. فرض کنید متغیرهای تصادفی X و Y به ترتیب تعداد مهره‌های سفید خارج شده

از A و B باشند.

الف- تابع احتمال توأم X و Y را به دست آورید.

ب- $P(XY=2)$ و $P(X\le 1 \mid Y=1)$ و $P(Y=0 \mid X\ge 2)$ را محاسبه کنید.

۳۱ فرض کنید که متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} \frac{1}{2} & \text{برای نقاط درون مربعی به رئوس } \begin{cases}(-a,0),(a,0)\\(0,-a),(0,a)\end{cases} \\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار a را تعیین کنید و توابع چگالی احتمال حاشیه‌ای X و Y را به دست آورید.

ب- $P(X+Y<0)$ را محاسبه کنید.

۳۲ عدد X را به تصادف از بین اعداد {۱ , ۲ , ۳ , ۴ , ۵} انتخاب می‌کنیم و سپس عددی را به تصادف از بین اعداد {۱ , ۲ , ... , X} انتخاب می‌کنیم و آن را Y می‌نامیم.

الف- تابع احتمال توأم X و Y و تابع احتمال شرطی X به شرط $Y=2$ را به دست آورید.

ب- آیا X و Y مستقل هستند؟ چرا؟

۳۳ اگر X و Y و Z دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y,Z}(x,y,z)=\begin{cases} kxy^2z & 0<x<1 \ , \ 0<y<1 \ , \ 0<z<2 \\ 0 & \text{سایر نقاط}\end{cases}$$

الف- مقدار k را تعیین کنید و $P(X<\frac{1}{4}, Y>\frac{1}{2}, 1<Z<2)$ را محاسبه کنید.

ب- آیا X و Y و Z از یکدیگر مستقل هستند؟ چرا؟

۳۴ ظرفی شامل ۳ مهرهٔ سفید و ۱ مهرهٔ سیاه است. از این ظرف ابتدا یک مهره بیرون می آوریم. اگر این مهره سفید باشد، ۲ مهرهٔ دیگر و اگر سیاه باشد یک مهرهٔ دیگر به تصادف و بدون جایگذاری از ظرف خارج می‌کنیم و متغیرهای تصادفی X و Y را به ترتیب تعداد مهره‌های سفید و سیاه باقی مانده در ظرف در نظر می‌گیریم.

الف- تابع احتمال توأم X و Y را به دست آورید و $P(X<2 \mid Y=0)$ را محاسبه کنید.

ب- آیا X و Y مستقل هستند؟ چرا؟

۳۵ نشان دهید که تابع

$$f_{X,Y}(x,y)=\frac{1}{2\pi(1+x^2+y^2)^{\frac{3}{2}}} \qquad -\infty<x<+\infty\ ,\ -\infty<y<+\infty$$

یک تابع چگالی احتمال توأم است و توابع چگالی احتمال حاشیه‌ای X و Y را به دست آورید.

فصل چهارم

امید ریاضی

۱.۴ مفهوم و تعریف امید ریاضی

در فصل سوم متغیرهای تصادفی و توزیع احتمال آنها را مورد بررسی قرار دادیم. در این فصل امید ریاضی یک متغیر تصادفی و توابعی از متغیرهای تصادفی را بررسی می‌کنیم. برای پی‌بردن به مفهوم امید ریاضی به مثال زیر توجه کنید.

مثال ۱.۱.۴ فرض کنید جریمهٔ توقف اتومبیل در محل پارک ممنوع یک هزار تومان باشد و در یک شهر ۴۰ درصد افراد هیچگاه، ۳۰ درصد افراد یک‌بار، ۲۰ درصد افراد دوبار و ۱۰ درصد افراد سه‌بار در ماه به جهت پارک در محل پارک ممنوع جریمه شوند. به طور متوسط انتظار دارید که هر نفر در این شهر چه مبلغی را در ماه برای جریمه توقف اتومبیل در محل پارک ممنوع پرداخت نماید.

حل فرض کنید ۵۰۰ نفر از افراد این شهر را در نظر بگیریم. بر اساس درصدهای داده شده انتظار داریم که ۲۰۰ نفر آنها هیچ مبلغی، ۱۵۰ نفر آنها یک هزار تومان، ۱۰۰ نفر آنها دو هزار تومان و ۵۰ نفر آنها سه هزار تومان جریمه شوند. بنابراین انتظار داریم که ۵۰۰ نفر مبلغ زیر را (بر حسب هزار تومان) جریمه پرداخت کنند

$$(500\times 0/40\times 0)+(500\times 0/30\times 1)+(500\times 0/20\times 2)+(500\times 0/10\times 3)=500$$

بنابراین بطور متوسط انتظار داریم که هر نفر $1=\frac{500}{500}$ هزار تومان در ماه جریمه پرداخت نماید.

با توجه به عملیات مثال قبل دیده می‌شود که متوسط مبلغ جریمه یک هزار تومان ربطی به تعداد ۵۰۰ نفر ندارد، زیرا اگر تعداد را یک میلیون نفر نیز می‌گرفتیم در جواب تغییری حاصل

نمی‌گردید. در حقیقت این مقدار یک عدد انتظاری (حدی) می‌باشد. در این مثال ما یک متغیر تصادفی X داریم که برابر مبلغ جریمه شخص در یک ماه بر حسب هزار تومان است و تابع احتمال آن به صورت زیر می‌باشد

x	۰	۱	۲	۳
$f_X(x)$	۰/۴۰	۰/۳۰	۰/۲۰	۰/۱۰

عدد زیر را که در حقیقت میانگین وزنی مبلغ جریمه می‌باشد امید ریاضی X یا میانگین X و یا مقدار مورد انتظار X می‌نامند و آنرا با نمادهای $E(X)$ یا μ یا μ_X نمایش می‌دهند.

$$\mu=E(X)=(0\times 0/40)+(1\times 0/30)+(2\times 0/20)+(3\times 0/10)=1$$
$$=\sum_{x=0}^{3} x f_X(x)$$

اگر متغیر تصادفی X پیوسته باشد امید ریاضی آن با تبدیل مجموع به انتگرال در فرمول بالا محاسبه می‌گردد.

تعریف ۱.۴ فرض کنید X یک متغیر تصادفی با تابع احتمال یا چگالی احتمال $f_X(x)$ باشد. امید ریاضی X یا میانگین X به صورت زیر تعریف می‌شود

$$E(X)=\sum_{x} x f_X(x) \qquad \text{اگر X گسسته باشد}$$
$$E(X)=\int_{-\infty}^{+\infty} x f_X(x)\,dx \qquad \text{اگر X پیوسته باشد} \qquad (۱.۴)$$

در صورتی که مجموع یا انتگرال فوق همگرا نباشد گوئیم امید ریاضی X وجود ندارد.

مثال ۲.۱.۴ فرض کنید بخواهیم ۳ نفر را از بین ۵ مهندس و ۳ تکنسین انتخاب کنیم. امید ریاضی تعداد مهندسین انتخابی در بین این ۳ نفر را به دست آورید.

حل اگر متغیر تصادفی X برابر تعداد مهندسین انتخابی در بین ۳ نفر انتخاب شده باشد آنگاه $S_X=\{0,1,2,3\}$ که تابع احتمال آن به صورت زیر به دست می‌آید

$$f_X(x)=P(X=x)=\frac{\binom{5}{x}\binom{3}{3-x}}{\binom{8}{3}} \qquad x=0,1,2,3$$

و یا

x	۰	۱	۲	۳
$f_X(x)$	$\frac{1}{56}$	$\frac{15}{56}$	$\frac{30}{56}$	$\frac{10}{56}$

بنابراین

$$\mu=E(X)=\sum_{x=0}^{3} x f_X(x)=(0\times\frac{1}{56})+(1\times\frac{15}{56})+(2\times\frac{30}{56})+(3\times\frac{10}{56})=\frac{105}{56}\approx 1/9$$

یعنی اگر بخواهیم ۳ نفر را از این گروه انتخاب کنیم به طور متوسط انتظار داریم که ۱/۹ آنها مهندس باشند. (توجه کنید که امید ریاضی X ممکن است مقداری باشد که با مجموعهٔ مقادیر X متفاوت است).

مثال ۳.۱.۴ فرض کنید متغیر تصادفی X نشانگر طول عمر نوعی لاستیک بر حسب سال باشد که دارای تابع چگالی احتمال زیر است

$$f_X(x)=\begin{cases}\frac{1}{2}e^{-\frac{1}{2}x} & x>0\\ 0 & \text{سایر نقاط}\end{cases}$$

متوسط طول عمر این نوع لاستیک را پیدا کنید.

حل متوسط طول عمر این نوع لاستیک E(X) می‌باشد بنابراین

$$\mu=E(X)=\int_{-\infty}^{+\infty} x f_X(x)\,dx=\int_{0}^{+\infty} x(\frac{1}{2}e^{-\frac{1}{2}x})dx=(-x-2)e^{-\frac{1}{2}x}\Big|_{0}^{+\infty}=0-(-2)=2$$

پس انتظار داریم که این نوع لاستیک به طور متوسط ۲ سال کار کند.

۲.۴ امید ریاضی تابعی از یک یا چند متغیر تصادفی

در بعضی از مسائل نیاز به محاسبهٔ امید ریاضی تابعی از یک متغیر تصادفی X مانند g(X) داریم. به عنوان مثال g(X) می‌تواند X^2 یا $2X+3$ یا... باشد. برای محاسبهٔ امید ریاضی g(X) از قضیهٔ زیر استفاده می‌کنیم.

قضیهٔ ۱.۴ فرض کنید X یک متغیر تصادفی با تابع احتمال یا تابع چگالی احتمال $f_X(x)$ باشد. امید ریاضی تابع g(X) به صورت زیر به دست می‌آید

$$
\begin{array}{ll}
E[g(X)] = \sum_{x} g(X) f_X(x) & \text{اگر } X \text{ گسسته باشد} \\
E[g(X)] = \int_{-\infty}^{+\infty} g(X) f_X(x)\,dx & \text{اگر } X \text{ پیوسته باشد}
\end{array} \qquad (۲.۴)
$$

مثال ۱.۲.۴ فرض کنید متغیر تصادفی X دارای تابع احتمال زیر باشد

x	۱-	۰	۱	۲
$f_X(x)$	۰/۱	۰/۱	۰/۵	۰/۳

امید ریاضی تابع $g(X)=(X-۱)^{۲}$ را به دست آورید.

حل X یک متغیر تصادفی گسسته است. بنابراین از رابطهٔ (۲.۴) داریم که

$$
\begin{aligned}
E[(X-۱)^{۲}] &= \sum_{x=-۱}^{۲} (x-۱)^{۲} f_X(x) \\
&= (-۱-۱)^{۲}(۰/۱)+(۰-۱)^{۲}(۰/۱)+(۱-۱)^{۲}(۰/۵)+(۲-۱)^{۲}(۰/۳)=۰/۸
\end{aligned}
$$

مثال ۲.۲.۴ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$
f_X(x)=\begin{cases} \frac{۳}{۱۶}\sqrt{x} & ۰<x<۴ \\ ۰ & \text{سایر نقاط} \end{cases}
$$

امید ریاضی $Y=۵X-۴$ را به دست آورید.

حل X یک متغیر تصادفی پیوسته است، بنابراین

$$
E(۵X-۴)=\int_{۰}^{۴}(۵x-۴)\frac{۳}{۱۶}\sqrt{x}\,dx=\frac{۳}{۱۶}\left[۲\sqrt{x^{۵}}-\frac{۸}{۳}\sqrt{x^{۳}}\right]_{۰}^{۴}=\frac{۳}{۱۶}\left[۶۴-\frac{۶۴}{۳}\right]=۸
$$

با توجه به قضیهٔ ۱.۴ می‌توان مفهوم امید ریاضی را به تابعی از دو متغیر تصادفی به صورت زیر تعمیم داد.

تعریف ۲.۴ فرض کنید X و Y دو متغیر تصادفی با تابع احتمال یا تابع چگالی احتمال توأم $f_{X,Y}(x,y)$ باشند. امید ریاضی تابع $g(X,Y)$ به صورت زیر تعریف می‌شود.

$$
\begin{array}{ll}
E[g(X,Y)] = \sum_{x}\sum_{y} g(X,Y) f_{X,Y}(x,y) & \text{اگر } X \text{ و } Y \text{ گسسته باشند} \\
E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(X,Y) f_{X,Y}(x,y)\,dx\,dy & \text{اگر } X \text{ و } Y \text{ پیوسته باشند}
\end{array} \qquad (۳.۴)
$$

تعریف فوق را می‌توان به سادگی برای امید ریاضی تابعی از چند متغیر تصادفی تعمیم داد.

مثال ۳.۲.۴ جعبه‌ای شامل ۲ مهرهٔ سفید و ۳ مهرهٔ سیاه است. ابتدا یک مهره از این جعبه انتخاب می‌کنیم و متغیر تصادفی X را برابر تعداد مهره‌های سفید در این یک مهره انتخاب شده در نظر می‌گیریم. سپس از مابقی مهره‌های جعبه دو مهرهٔ دیگر بدون جایگذاری انتخاب می‌کنیم و متغیر تصادفی Y را برابر تعداد مهره‌های سفید مشاهده شده در این دو مهره انتخابی در نظر می‌گیریم. تابع احتمال توأم X و Y را به دست آورید و $E(X^{۲}Y)$ را محاسبه کنید.

حل در اینجا $S_X=\{۰,۱\}$ و $S_Y=\{۰,۱,۲\}$ و همچنین داریم که

$$
\begin{aligned}
f_{X,Y}(۰,۰) &= P(X=۰,Y=۰)=P(X=۰)P(Y=۰\mid X=۰) \\
&= \frac{۳}{۵}\times\frac{\binom{۲}{۲}}{\binom{۴}{۲}}=\frac{۳}{۳۰}=۰/۱
\end{aligned}
$$

$$
\begin{aligned}
f_{X,Y}(۱,۰) &= P(X=۱,Y=۰)=P(X=۱)P(Y=۰\mid X=۱) \\
&= \frac{۲}{۵}\times\frac{\binom{۳}{۲}}{\binom{۴}{۲}}=\frac{۶}{۳۰}=۰/۲
\end{aligned}
$$

با انجام محاسبات مشابه، جدول توزیع احتمالات توأم X و Y به صورت زیر به دست می‌آید

y \ x	۰	۱	$f_Y(y)$
۰	۰/۱	۰/۲	۰/۳
۱	۰/۴	۰/۲	۰/۶
۲	۰/۱	۰	۰/۱
$f_X(x)$	۰/۶	۰/۴	

بنابراین

$$
\begin{aligned}
E(X^{۲}Y) &= \sum_{x=۰}^{۱}\sum_{y=۰}^{۲} x^{۲} y f_{X,Y}(x,y) \\
&= (۰)(۰/۱)+(۰)(۰/۴)+(۰)(۰/۱)+(۰)(۰/۲)+(۱)(۰/۲)+(۲)(۰)=۰/۲
\end{aligned}
$$

مثال ۴.۲.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$
f_{X,Y}(x,y)=\begin{cases} \frac{۱۶y}{x^{۳}} & x>۲\ ,\ ۰<y<۱ \\ ۰ & \text{سایر نقاط} \end{cases}
$$

امید ریاضی $g(X,Y)=\frac{X+1}{Y}$ را محاسبه کنید.

حل $E\left(\frac{X+1}{Y}\right)=\int_{2}^{+\infty}\int_{0}^{1}\left(\frac{x+1}{y}\right)\left(\frac{16y}{x^{3}}\right)dy\,dx=16\int_{2}^{+\infty}\frac{x+1}{x^{3}}\left[\int_{0}^{1}dy\right]dx$

$$=16\int_{2}^{+\infty}\left(\frac{1}{x^{2}}+\frac{1}{x^{3}}\right)dx=16\left[\frac{-1}{x}-\frac{1}{2x^{2}}\right]_{2}^{+\infty}=16\left[(0)+\left(\frac{1}{2}+\frac{1}{8}\right)\right]=10$$

توجه کنید اگر در تعریف ۲.۴ قرار دهیم g(X,Y)=X و یا g(X,Y)=Y آنگاه امید ریاضی X یا Y را می‌توان توسط تابع احتمال یا تابع چگالی احتمال توأم X و Y به صورت زیر محاسبه کرد

اگر X و Y گسسته باشند

$$E(X)=\sum_{x}\sum_{y}xf_{X,Y}(x,y)\ ,\ E(Y)=\sum_{x}\sum_{y}yf_{X,Y}(x,y) \qquad (۴.۴)$$

اگر X و Y پیوسته باشند

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf_{X,Y}(x,y)dxdy$$

$$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf_{X,Y}(x,y)dxdy$$

مثال ۵.۲.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}\frac{1}{y}e^{-y} & 0<x<y<+\infty\\ 0 & \text{سایر نقاط}\end{cases}$$

آیا می‌توان E(X) را توسط تابع چگالی احتمال حاشیه‌ای X محاسبه کرد؟ E(X) را با استفاده از رابطهٔ (۴.۴) محاسبه کنید.

حل با توجه به اینکه $f_X(x)=\int_{-\infty}^{+\infty}f_{X,Y}(x,y)dy=\int_{x}^{+\infty}\frac{1}{y}e^{-y}dy$

و انتگرال فوق قابل محاسبه نیست پس نمی‌توان $f_X(x)$ را به دست آورده و از روی آن E(X) را محاسبه کرد. اما با توجه به رابطهٔ (۴.۴) داریم که

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf_{X,Y}(x,y)dxdy=\int_{0}^{+\infty}\int_{0}^{y}\frac{x}{y}e^{-y}dxdy$$

$$=\int_{0}^{+\infty}\frac{1}{y}e^{-y}\left[\frac{1}{2}x^{2}\right]_{0}^{y}dy=\frac{1}{2}\int_{0}^{+\infty}ye^{-y}dy=\frac{1}{2}\left[(-y-1)e^{-y}\right]_{0}^{+\infty}$$

$$=\frac{1}{2}\left[(0)-(-1)\right]=\frac{1}{2}$$

۳.۴ قوانین امید ریاضی

در این بخش قضیه‌هایی را برای ساده کردن محاسبهٔ امید ریاضی تابعی از یک یا چند متغیر تصادفی می‌آوریم. اثبات این قضایا بسیار ساده می‌باشد و بعضی از آنها را در حالت پیوسته ثابت می‌کنیم و مابقی را به خواننده واگذار می‌کنیم.

قضیهٔ ۲.۴ اگر g(X) و h(X) توابعی از متغیر تصادفی X باشند که امید ریاضی آنها موجود است و a و b اعداد ثابتی باشند آنگاه

$$E[ag(X)+bh(X)]=aE[g(X)]+bE[h(X)]$$

اثبات $E[ag(X)+bh(X)]=\int_{-\infty}^{+\infty}[ag(x)+bh(x)]f_X(x)dx$

$$=a\int_{-\infty}^{+\infty}g(x)f_X(x)dx+b\int_{-\infty}^{+\infty}h(x)f_X(x)dx$$

$$=aE[g(X)]+bE[h(X)]$$

نتیجهٔ ۱.۴ اگر a و b اعداد ثابتی باشند آنگاه

$$E(aX+b)=aE(X)+b$$

مثال ۱.۳.۴ مثال ۱.۲.۴ را با استفاده از قضیهٔ ۲.۴ حل کنید.

حل در مثال ۱.۲.۴ تابع احتمال X عبارت بود از

x	۱-	۰	۱	۲
$f_X(x)$	۰/۱	۰/۱	۰/۵	۰/۳

بنابراین

$$E(X)=(-1)(0.1)+(0)(0.1)+(1)(0.5)+(2)(0.3)=1$$

$$E(X^{2})=(-1)^{2}(0.1)+(0)^{2}(0.1)+(1)^{2}(0.5)+(2)^{2}(0.3)=1.8$$

و در نتیجه از قضیه ۲.۴ داریم که

$$E[(X-1)^{2}]=E[X^{2}-2X+1]=E(X^{2})-2E(X)+1=1.8-2(1)+1=0.8$$

مثال ۲.۳.۴ مثال ۲.۲.۴ را با استفاده از نتیجهٔ ۱.۴ حل کنید.

حل در مثال ۲.۲.۴ تابع چگالی احتمال X عبارت بود از

$$f_X(x)=\begin{cases}\frac{۳}{۱۶}\sqrt{x} & ۰<x<۴\\ ۰ & \text{سایر نقاط}\end{cases}$$

بنابراین

$$E(X)=\int_{۰}^{۴}x(\frac{۳}{۱۶}\sqrt{x})\,dx=\frac{۳}{۱۶}\times\frac{۲}{۵}\sqrt{x^{۵}}\Big|_{۰}^{۴}=\frac{۱۲}{۵}$$

و در نتیجه از نتیجه ۱.۴ داریم که

$$E(۵X-۴)=۵E(X)-۴=۵(\frac{۱۲}{۵})-۴=۸$$

قضیهٔ ۳.۴ اگر $g(X,Y)$ و $h(X,Y)$ توابعی از متغیرهای تصادفی X و Y باشند که امید ریاضی آنها موجود است و a و b اعداد ثابتی باشند آنگاه

$$E[ag(X,Y)+bh(X,Y)]=aE[g(X,Y)]+bE[h(X,Y)]$$

نتیجهٔ ۲.۴ اگر X و Y دو متغیر تصادفی باشند آنگاه

$$E(X\pm Y)=E(X)\pm E(Y)$$

همانطور که در نتیجهٔ ۲.۴ ملاحظه شد، امید مجموع یا تفاضل دو متغیر تصادفی برابر مجموع یا تفاضل امیدهای آنها می‌باشد. اما در حالت کلی امید حاصلضرب دو متغیر تصادفی برابر حاصلضرب امیدهای آنها نیست و تنها در حالتی که دو متغیر تصادفی از یکدیگر مستقل باشند این رابطه برقرار است. به قضیهٔ زیر توجه کنید.

قضیهٔ ۴.۴ اگر X و Y دو متغیر تصادفی مستقل باشند آنگاه

$$E(XY)=E(X)E(Y)$$

اثبات چون X و Y متغیرهای تصادفی مستقل هستند بنابراین برای هر X و Y داریم که

$$f_{X,Y}(x,y)=f_X(x)f_Y(y)$$

و در نتیجه

$$\begin{aligned}E(XY)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_{X,Y}(x,y)\,dxdy\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_X(x)f_Y(y)\,dxdy\\&=\left[\int_{-\infty}^{+\infty}xf_X(x)\,dx\right]\left[\int_{-\infty}^{+\infty}yf_Y(y)\,dy\right]=E(X)E(Y)\end{aligned}$$

مثال ۳.۳.۴ در مثال ۴.۲.۴ نشان دهید که رابطهٔ $E(XY)=E(X)E(Y)$ برقرار است.

حل در مثال ۴.۲.۴ تابع چگالی احتمال توأم X و Y عبارت بود از

$$f_{X,Y}(x,y)=\begin{cases}\frac{۱۶y}{x^{۳}} & x>۲\ ,\ ۰<y<۱\\ ۰ & \text{سایر نقاط}\end{cases}$$

بنابراین

$$f_X(x)=\int_{۰}^{۱}\frac{۱۶y}{x^{۳}}dy=\frac{۸y^{۲}}{x^{۳}}\Big|_{۰}^{۱}=\frac{۸}{x^{۳}}\qquad x>۲$$

$$f_Y(y)=\int_{۲}^{+\infty}\frac{۱۶y}{x^{۳}}dx=\frac{-۸y}{x^{۲}}\Big|_{۲}^{+\infty}=۲y\qquad ۰<y<۱$$

در نتیجه برای هر x و y داریم که $f_{X,Y}(x,y)=f_X(x)f_Y(y)$ یعنی X و Y از یکدیگر مستقل هستند و طبق قضیهٔ ۴.۴ داریم که $E(XY)=E(X)E(Y)$. در ضمن با انجام محاسبات ساده دیده می‌شود که $E(XY)=\frac{۸}{۳}$ و $E(X)=۴$ و $E(Y)=\frac{۲}{۳}$ که صحت رابطهٔ مذکور را نشان می‌دهد.

توجه کنید که عکس قضیهٔ ۴.۴ در حالت کلی برقرار نیست. یعنی برای دو متغیر تصادفی X و Y می‌توانیم رابطهٔ $E(XY)=E(X)E(Y)$ را داشته باشیم اما این دو متغیر از یکدیگر مستقل نباشند. به مثال زیر توجه کنید.

مثال ۴.۳.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر باشند. نشان دهید که X و Y از یکدیگر مستقل نیستند اما رابطهٔ $E(XY)=E(X)E(Y)$ برقرار است.

y \ x	۱-	۰	۱	$f_Y(y)$
۱-	۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$	$\frac{۲}{۴}$
۱	۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
$f_X(x)$	$\frac{۱}{۴}$	$\frac{۲}{۴}$	$\frac{۱}{۴}$	

حل با توجه به اینکه $f_X(-۱)f_Y(-۱)=\frac{۱}{۴}\times\frac{۱}{۴}\neq ۰=f_{X,Y}(-۱,-۱)$ پس X و Y مستقل نیستند. اما $E(XY)=(-۱)(-۱)(۰)+\dots+(۱)(۱)(۰)=۰$ و

$E(X)=E(Y)=-\frac{1}{4}+0+\frac{1}{4}=0$ پس $E(XY)=E(X)E(Y)$.

قوانین امید ریاضی را به سادگی می‌توان به چند متغیر تصادفی تعمیم داد. مثلاً اگر X_1، X_2، ... و X_n، n متغیر تصادفی باشند و a_1، a_2، ... و a_n اعداد ثابتی باشند آنگاه

$$E(\sum_{i=1}^{n} a_i X_i)=\sum_{i=1}^{n} a_i E(X_i)$$

و اگر این n متغیر تصادفی از یکدیگر مستقل باشند آنگاه

$$E(\prod_{i=1}^{n} X_i)=\prod_{i=1}^{n} E(X_i)$$

۴.۴ امیدهای ریاضی خاص

در این بخش امیدهای ریاضی توابعی از متغیرهای تصادفی که مفهومی خاص را دارند بررسی می‌کنیم.

گشتاورهای یک متغیر تصادفی در قضیهٔ ۱.۴ اگر قرار دهیم $g(X)=X^r$ که در آن r یک عدد صحیح نامنفی است، آنگاه امید ریاضی این تابع را rامین **گشتاور حول مبدأ متغیر تصادفی** X گویند و آن را با نماد μ'_r نمایش می‌دهند، یعنی

$\boxed{\mu'_r=E(X^r)}$ گشتاور rام X حول مبدأ (۵.۴)

توجه کنید که $\mu'_0=1$ و $\mu'_1=E(X)=\mu$ که همان امید ریاضی X و یا میانگین X است. اگر در قضیهٔ ۱.۴ قرار دهیم $g(X)=(X-\mu)^r$ آنگاه امید ریاضی این تابع را گشتاور مرتبهٔ rام X حول میانگین و یا گشتاور مرکزی X گویند و آن را با نماد μ_r نمایش می‌دهند یعنی

$\boxed{\mu_r=E[(X-\mu)^r]}$ گشتاور مرکزی مرتبهٔ rام X (۶.۴)

توجه کنید که $\mu_0=1$ و $\mu_1=0$ می‌باشد.

واریانس گشتاور مرکزی مرتبهٔ دوم X را واریانس X گویند و با نمادهای σ^2 یا σ_X^2 یا $Var(X)$ نمایش می‌دهند، یعنی

$\boxed{\sigma^2=\sigma_X^2=Var(X)=E[(X-\mu)^2]}$ واریانس X (۷.۴)

واریانس یک متغیر تصادفی، میزان پراکندگی توزیع متغیر تصادفی نسبت به میانگین آن را

می‌سنجد. هر چه واریانس بزرگتر باشد، میزان پراکندگی توزیع متغیر تصادفی نسبت به میانگین بیشتر می‌باشد و هر چه واریانس کوچکتر باشد این میزان کمتر است. جذر واریانس یعنی σ را انحراف معیار گویند. با استفاده از قوانین امید ریاضی می‌توان فرم ساده‌تری برای محاسبهٔ واریانس به دست آورد که آن را در قضیهٔ زیر می‌آوریم.

قضیهٔ ۵.۴ واریانس یک متغیر تصادفی X با میانگین μ به صورت زیر به دست می‌آید

$$\boxed{Var(X)=E[(X-\mu)^2]=E(X^2)-[E(X)]^2} \quad (8.4)$$

مثال ۱.۴.۴ در مثال ۲.۱.۴ واریانس متغیر تصادفی X را به دست آورید.

حل در مثال ۲.۱.۴ دیدیم که $\mu=E(X)=\frac{105}{56}$ و

x	0	1	2	3
$f_X(x)$	$\frac{1}{56}$	$\frac{15}{56}$	$\frac{30}{56}$	$\frac{10}{56}$

بنابراین

$$E(X^2)=\sum_{x=0}^{3} x^2 f_X(x)=(0^2)(\frac{1}{56})+(1^2)(\frac{15}{56})+(2^2)(\frac{30}{56})+(3^2)(\frac{10}{56})=\frac{225}{56}$$

و در نتیجه

$$Var(X)=E(X^2)-[E(X)]^2=\frac{225}{56}-(\frac{105}{56})^2=\frac{1575}{3136}=0/502$$

مثال ۲.۴.۴ میانگین و واریانس متغیر تصادفی X که دارای تابع چگالی احتمال زیر است را پیدا کنید.

$$f_X(x)=\begin{cases}4xe^{-2x} & x>0\\ 0 & \text{سایر نقاط}\end{cases}$$

حل به وسیلهٔ انتگرال‌گیری به روش جزء به جزء داریم که

$$\mu=E(X)=\int_0^{+\infty} x(4xe^{-2x})dx=\left[(-2x^2-2x-1)e^{-2x}\right]_0^{+\infty}=(0)-(-1)=1$$

$$E(X^2)=\int_0^{+\infty} x^2(4xe^{-2x})dx=\left[(-2x^3-3x^2-3x-\frac{3}{2})e^{-2x}\right]_0^{+\infty}=(0)-(-\frac{3}{2})=\frac{3}{2}$$

بنابراین

$$\sigma^2=Var(X)=\frac{3}{2}-(1)^2=\frac{1}{2}$$

کواریانس در تعریف ۲.۴ اگر قرار دهیم $g(X,Y)=(X-\mu_X)(Y-\mu_Y)$ آنگاه امید ریاضی این تابع را کواریانس X و Y گویند و آن را با نمادهای σ_{XY} و یا $COV(X,Y)$ نمایش می‌دهند، یعنی

$$\boxed{\sigma_{XY}=COV(X,Y)=E\left[(X-\mu_X)(Y-\mu_Y)\right] \quad \text{کواریانس}} \quad (۹.۴)$$

کواریانس X و Y رابطهٔ دو متغیر تصادفی X و Y را نشان می‌دهد. اگر X و Y هم جهت باشند (یعنی هر دو با هم افزایش و یا هر دو با هم کاهش یابند) آنگاه کواریانس X و Y مثبت است و اگر X و Y در خلاف جهت هم باشند آنگاه کواریانس X و Y منفی است. با استفاده از قوانین امید ریاضی می‌توان فرم ساده‌تری برای محاسبه کواریانس به دست آورد که آن را در قضیهٔ زیر می‌آوریم.

قضیهٔ ۶.۴ کواریانس دو متغیر تصادفی X و Y به صورت زیر به دست می‌آید

$$\boxed{COV(X,Y)=E\left[(X-\mu_X)(Y-\mu_Y)\right]=E(XY)-E(X)E(Y)} \quad (۱۰.۴)$$

نتیجهٔ ۳.۴ با استفاده از قضیهٔ ۴.۴ و قضیهٔ ۶.۴، اگر X و Y دو متغیر تصادفی مستقل باشند آنگاه $COV(X,Y)=۰$. ولی عکس این مطلب برقرار نیست، یعنی اگر $COV(X,Y)=۰$ آنگاه دلیلی ندارد که X و Y متغیرهای تصادفی مستقل باشند (مثال ۴.۳.۴ را ملاحظه کنید).

مثال ۳.۴.۴ در مثال ۳.۲.۴ کواریانس متغیرهای تصادفی X و Y را به دست آورید.

حل در مثال ۳.۲.۴ جدول توزیع احتمالات توأم X و Y به صورت زیر به دست آمد

y \ x	۰	۱	$f_Y(y)$
۰	۰/۱	۰/۲	۰/۳
۱	۰/۴	۰/۲	۰/۶
۲	۰/۱	۰	۰/۱
$f_X(x)$	۰/۶	۰/۴	

بنابراین

$$E(X)=(۰)(۰/۶)+(۱)(۰/۴)=۰/۴$$
$$E(Y)=(۰)(۰/۳)+(۱)(۰/۶)+(۲)(۰/۱)=۰/۸$$
$$E(XY)=۰+۰+۰+۰+۰/۲+۰=۰/۲$$

در نتیجه

$$COV(X,Y)=E(XY)-E(X)E(Y)=(۰/۲)-(۰/۸)(۰/۴)=-۰/۱۲$$

چون کواریانس منفی است پس X و Y در خلاف جهت یکدیگر حرکت می‌کنند. یعنی اگر در انتخاب مهرهٔ اول تعداد سفید به یک مهره افزایش یابد آنگاه در انتخاب ۲ مهره بعدی تعداد مهره‌های سفید انتخابی کاهش می‌یابد.

خواص واریانس و کواریانس با استفاده از قوانین امید ریاضی به سادگی می‌توان خواص زیر را

برای واریانس و کواریانس نتیجه گرفت که اثبات آنها را به خواننده واگذار می‌کنیم. فرض کنید a، b و c اعداد ثابت و X و Y متغیرهای تصادفی باشند، در این صورت

الف- $Var(c)=۰ \quad , \quad Var(aX+b)=a^{۲}Var(X)$

ب- $COV(X,X)=Var(X)$

ج- $COV(X,Y)=COV(Y,X)$

د- $COV(X,c)=۰$

ه- $COV(aX+b,cY+d)=ac\,COV(X,Y)$

و- $Var(aX+bY+c)=a^{۲}Var(X)+b^{۲}Var(Y)+۲abCOV(X,Y)$

ز- اگر X و Y مستقل باشند آنگاه

$$Var(aX+bY+c)=a^{۲}Var(X)+b^{۲}Var(Y)$$

ح- اگر $X_۱$ و $X_۲$ دو متغیر تصادفی باشند آنگاه

$$COV(aX_۱+bX_۲,Y)=aCOV(X_۱,Y)+bCOV(X_۲,Y)$$

خاصیت (ه) می‌گوید که اگر مبدأ اندازه‌گیری X و Y را تغییر دهیم، کواریانس آنها تغییر نمی‌کند ولی اگر واحد اندازه‌گیری X و Y را تغییر دهیم، کواریانس آنها تغییر می‌کند.

مثال ۴.۴.۴ در مثال ۳.۴.۴، $Var(۲X-۳Y+۴)$ را محاسبه کنید.

حل با استفاده از جدول توزیع احتمالات توأم در مثال ۳.۴.۴ داریم که $COV(X,Y)=-۰/۱۲$

و همچنین $E(X)=E(X^{۲})=۰/۴ \Rightarrow Var(X)=۰/۴-(۰/۴)^{۲}=۰/۲۴$

$$E(Y)=۰/۸,\ E(Y^{۲})=۱ \Rightarrow Var(Y)=۱-(۰/۸)^{۲}=۰/۳۶$$

بنابراین

$$Var(۲X-۳Y+۴)=۴Var(X)+۹Var(Y)-۱۲COV(X,Y)$$
$$=۴(۰/۲۴)+۹(۰/۳۶)-۱۲(-۰/۱۲)=۵/۶۴$$

ضریب همبستگی در خاصیت (ه) کواریانس مشاهده کردیم که کواریانس بستگی به واحد اندازه‌گیری X و Y دارد. برای اینکه معیاری برای سنجش میزان رابطه دو متغیر تصادفی X و Y پیدا کنیم که به واحد اندازه‌گیری X و Y بستگی نداشته باشد، کواریانس بین متغیرهای $\frac{X}{\sigma_X}$ و $\frac{Y}{\sigma_Y}$ را محاسبه می‌کنیم که σ_Y,σ_X به ترتیب انحراف معیارهای X و Y هستند، یعنی

$$COV(\frac{X}{\sigma_X},\frac{Y}{\sigma_Y})=\frac{COV(X,Y)}{\sigma_X\sigma_Y}=\frac{COV(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

این معیار را ضریب همبستگی متغیرهای تصادفی X و Y می‌نامند و آن را با نمادهای ρ یا $\rho(X,Y)$ نمایش می‌دهند. بنابراین

ضریب همبستگی X و Y (۱۱.۴)

$$\boxed{\rho=\rho(X,Y)=\frac{COV(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}}$$

ضریب همبستگی دو متغیر X و Y میزان رابطهٔ خطی دو متغیر تصادفی X و Y را می‌سنجد. با استفاده از قوانین امید ریاضی و خواص واریانس و کواریانس می‌توان خواص زیر را برای ضریب همبستگی اثبات کرد که اثبات آنها را به عنوان تمرین به خواننده واگذار می‌کنیم. فرض کنید a و b اعداد ثابت و X و Y دو متغیر تصادفی باشند، در این صورت

الف- $\rho(aX+b,cY+d)=\rho(X,Y)$

ب- همواره داریم که $-۱\le\rho\le ۱$

ج- اگر $Y=aX+b$ و $a>۰$ آنگاه $\rho>۰$

د- اگر $Y=aX+b$ و $a<۰$ آنگاه $\rho<۰$

هـ- اگر X و Y مستقل باشند آنگاه $\rho=۰$

خاصیت (الف) می‌گوید که ضریب همبستگی X و Y به مبدأ و واحد اندازه‌گیری X و Y بستگی ندارد. توجه کنید که اگر $\rho=۰$ باشد آنگاه دلیلی ندارد که X و Y از یکدیگر مستقل باشند (نتیجهٔ ۳.۴ را ملاحظه کنید). در این حالت یعنی حالتی که $\rho=۰$ باشد، متغیرهای تصادفی X و Y را ناهمبسته گویند.

مثال ۵.۴.۴ در مثال ۳.۴.۴ ضریب همبستگی X و Y را به دست آورید.

حل در مثال ۴.۴.۴ مشاهده کردیم که

$$Var(X)=۰/۲۴\ ,\quad Var(Y)=۰/۳۶\ ,\quad COV(X,Y)=-۰/۱۲$$

بنابراین

$$\rho=\frac{COV(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{-۰/۱۲}{\sqrt{(۰/۲۴)(۰/۳۶)}}=\frac{-۱}{\sqrt{۶}}=-۰/۴۰۸$$

یعنی X و Y دارای رابطه در خلاف جهت یکدیگر هستند ولی این رابطه خیلی شدید نیست.

مثال ۶.۴.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند. ضریب همبستگی X و Y را محاسبه کنید.

$$f_{X,Y}(x,y)=\begin{cases} e^{-y} & ۰<x<y<+\infty \\ ۰ & \text{سایر نقاط} \end{cases}$$

حل برای محاسبهٔ ضریب همبستگی X و Y ابتدا توابع چگالی احتمال حاشیه‌ای X و Y را به دست می‌آوریم.

$$f_X(x)=\int_x^{+\infty}e^{-y}dy=\left[-e^{-y}\right]_x^{+\infty}=e^{-x}\qquad x>۰$$

$$f_Y(y)=\int_۰^{y}e^{-y}dx=e^{-y}[x]_۰^{y}=ye^{-y}\qquad y>۰$$

به وسیلهٔ انتگرال‌گیری جزء به جزء می‌توان نشان داد که برای هر عدد صحیح نامنفی n داریم که $\int_۰^{+\infty}x^ne^{-x}dx=n!$ بنابراین

$$E(X)=\int_۰^{+\infty}xe^{-x}dx=۱!=۱\ ,\quad E(X^۲)=\int_۰^{+\infty}x^۲e^{-x}dx=۲!=۲$$

و در نتیجه $Var(X)=۲-(۱)^۲=۱$. همچنین

$$E(Y)=\int_۰^{+\infty}y^۲e^{-y}dy=۲!=۲\ ,\quad E(Y^۲)=\int_۰^{+\infty}y^۳e^{-y}dy=۳!=۶$$

و در نتیجه $Var(Y)=۶-(۲)^۲=۲$. همچنین

$$E(XY)=\int_۰^{+\infty}\int_۰^{y}xye^{-y}dxdy=\int_۰^{+\infty}ye^{-y}\left[\frac{۱}{۲}x^۲\right]_۰^{y}dy$$
$$=\frac{۱}{۲}\int_۰^{+\infty}y^۳e^{-y}dy=\frac{۳!}{۲}=۳$$

و در نتیجه $COV(X,Y)=۳-(۱)(۲)=۱$. بنابراین

$$\rho=\frac{COV(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{۱}{\sqrt{(۱)(۲)}}=\frac{۱}{\sqrt{۲}}=۰/۷۰۷$$

۵.۴ امید ریاضی و واریانس شرطی

همانند تعریف امید ریاضی و تعریف واریانس، می‌توان امید ریاضی شرطی و واریانس شرطی را تعریف کرد. فرض کنید X و Y دو متغیر تصادفی باشند. امید ریاضی شرطی X^r به شرط Y=y به صورت زیر تعریف می‌شود

$$E(X^r \mid Y=y)=\begin{cases}\sum_{x} x^r f_{X\mid Y}(x\mid y) & \text{اگر X و Y گسسته باشند}\\ \int_{-\infty}^{+\infty} x^r f_{X\mid Y}(x\mid y)\,dx & \text{اگر X و Y پیوسته باشند}\end{cases} \qquad (۱۲.۴)$$

به همین ترتیب امید ریاضی Y^r به شرط X=x به صورت زیر تعریف می‌شود

$$E(Y^r \mid X=x)=\begin{cases}\sum_{y} y^r f_{Y\mid X}(y\mid x) & \text{اگر X و Y گسسته باشند}\\ \int_{-\infty}^{+\infty} y^r f_{Y\mid X}(y\mid x)\,dy & \text{اگر X و Y پیوسته باشند}\end{cases} \qquad (۱۳.۴)$$

واریانس شرطی X به شرط Y=y به صورت زیر تعریف می‌شود

$$\begin{aligned}Var(X\mid Y=y)&=E\left\{\left[X-E(X\mid Y=y)\right]^2 \mid Y=y\right\}\\ &=E(X^2\mid Y=y)-\left[E(X\mid Y=y)\right]^2\end{aligned} \qquad (۱۴.۴)$$

به همین ترتیب واریانس شرطی Y به شرط X=x به صورت زیر تعریف می‌شود

$$\begin{aligned}Var(Y\mid X=x)&=E\left\{\left[Y-E(Y\mid X=x)\right]^2 \mid X=x\right\}\\ &=E(Y^2\mid X=x)-\left[E(Y\mid X=x)\right]^2\end{aligned} \qquad (۱۵.۴)$$

مثال ۱.۵.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر باشند. $E(X^2 \mid Y=3)$ را محاسبه کنید.

y \ x	۲	۳	۴	$f_Y(y)$
۲	۱/۸	۱/۸	۰	۲/۸
۳	۱/۸	۲/۸	۱/۸	۴/۸
۴	۰	۱/۸	۱/۸	۲/۸
$f_X(x)$	۲/۸	۴/۸	۲/۸	

حل تابع احتمال شرطی X به شرط Y=۳ به صورت زیر به دست می‌آید

x	۲	۳	۴
$f_{X\mid Y}(x\mid 3)$	۱/۴	۲/۴	۱/۴

بنابراین

$$E(X^2\mid Y=3)=\sum_{x=2}^{4} x^2 f_{X\mid Y}(x\mid 3)=(2^2)(\tfrac{1}{4})+(3^2)(\tfrac{2}{4})+(4^2)(\tfrac{1}{4})=\frac{38}{4}=9.5$$

مثال ۲.۵.۴ در مثال ۶.۴.۴، $Var(X\mid Y=y)$ را محاسبه کنید.

حل در مثال ۶.۴.۴ داشتیم که

$$f_{X,Y}(x,y)=\begin{cases}e^{-y} & 0<x<y<+\infty\\ 0 & \text{سایر نقاط}\end{cases}\quad,\quad f_Y(y)=\begin{cases}ye^{-y} & y>0\\ 0 & \text{سایر نقاط}\end{cases}$$

بنابراین

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{e^{-y}}{ye^{-y}}=\frac{1}{y} \qquad 0<x<y<+\infty$$

در نتیجه

$$E(X\mid Y=y)=\int_0^y x(\tfrac{1}{y})\,dx=\frac{1}{y}\left[\tfrac{1}{2}x^2\right]_0^y=\frac{y}{2}$$

$$E(X^2\mid Y=y)=\int_0^y x^2(\tfrac{1}{y})\,dx=\frac{1}{y}\left[\tfrac{1}{3}x^3\right]_0^y=\frac{y^2}{3}$$

و بنابراین

$$Var(X\mid Y=y)=\frac{y^2}{3}-(\frac{y}{2})^2=\frac{1}{12}y^2 \qquad y>0$$

۶.۴ مسائل حل شده

مثال ۱.۶.۴ از جعبه‌ای محتوی ۸ لامپ که ۲ تای آنها سوخته است ۳ لامپ را به تصادف انتخاب می‌کنیم. اگر X نشان دهندهٔ تعداد لامپهای سوخته باشد، امید ریاضی X را به دست آورید.

حل تابع احتمال X به صورت زیر به دست می‌آید

$$f_X(x)=P(X=x)=\frac{\binom{2}{x}\binom{6}{3-x}}{\binom{8}{3}}\quad,\quad x=0,1,2$$

بنابراین

x	۰	۱	۲
$f_X(x)$	۱۰/۲۸	۱۵/۲۸	۳/۲۸

و در نتیجه

$$E(X)=\sum_{x=0}^{2} x f_X(x)=(0\times\frac{10}{28})+(1\times\frac{15}{28})+(2\times\frac{3}{28})=\frac{21}{28}=\frac{3}{4}$$

مثال ۲.۶.۴ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases} a+bx^{2} & 0<x<1 \\ 0 & \text{سایر نقاط} \end{cases}$$

اگر $E(X)=\frac{3}{5}$ باشد، مقادیر a و b و تابع توزیع X را به دست آورید.

حل برای اینکه مقادیر a و b را به دست آوریم، دو شرط داریم. شرط اول این است که

$$1=\int_{-\infty}^{+\infty} f_X(x)\,dx=\int_{0}^{1}(a+bx^{2})\,dx=\left[ax+\frac{b}{3}x^{3}\right]_{0}^{1}=a+\frac{b}{3}$$

و شرط دوم این است که

$$\frac{3}{5}=E(X)=\int_{0}^{1}x\,(a+bx^{2})\,dx=\left[\frac{a}{2}x^{2}+\frac{b}{4}x^{4}\right]_{0}^{1}=\frac{a}{2}+\frac{b}{4}$$

با حل دستگاه دو معادله و دو مجهول بالا مقادیر $a=\frac{3}{5}$ و $b=\frac{6}{5}$ به دست می‌آید. همچنین برای مقادیر $0\le x<1$ داریم که

$$F_X(x)=\int_{0}^{x}\left(\frac{3}{5}+\frac{6}{5}t^{2}\right)dt=\frac{3}{5}t+\frac{2}{5}t^{3}\Big|_{0}^{x}=\frac{3}{5}x+\frac{2}{5}x^{3}$$

بنابراین تابع توزیع X برابر است با

$$F_X(x)=\begin{cases} 0 & x<0 \\ \frac{3}{5}x+\frac{2}{5}x^{3} & 0\le x<1 \\ 1 & 1\le x \end{cases}$$

مثال ۳.۶.۴ نشان دهید که تابع زیر یک تابع احتمال متغیر تصادفی گسسته X است. آیا E(X) وجود دارد؟

$$f_X(x)=\begin{cases} \frac{1}{x(x+1)} & x=1,2,3,\ldots \\ 0 & \text{سایر نقاط} \end{cases}$$

حل برای اینکه این تابع یک تابع احتمال باشد بایستی مجموع آن یک شود یعنی

$$\sum_{x=1}^{+\infty}\frac{1}{x(x+1)}=\sum_{x=1}^{+\infty}\left[\frac{1}{x}-\frac{1}{x+1}\right]=\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\ldots=1$$

زیرا این سری یک سری تلسکوپی با مجموع یک است. اما E(X) وجود ندارد زیرا

$$E(X)=\sum_{x=1}^{+\infty}\frac{x}{x(x+1)}=\sum_{x=1}^{+\infty}\frac{1}{x+1}=+\infty$$

که در مقایسه با سری هارمونیک $\sum_{x=1}^{+\infty}\frac{1}{2x}$، یک سری واگرا است.

مثال ۴.۶.۴ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=c\left(\frac{1}{x}\right)^{\theta+1} \qquad x>1,\ \theta>1$$

مقدار c را تعیین کنید و امید ریاضی X را به دست آورید.

حل

$$1=\int_{-\infty}^{+\infty} f_X(x)\,dx=c\int_{1}^{+\infty}\left(\frac{1}{x}\right)^{\theta+1}dx=\frac{-c}{\theta x^{\theta}}\Big|_{1}^{+\infty}=\frac{c}{\theta}$$

بنابراین $c=\theta$ و در نتیجه

$$E(X)=\int_{1}^{+\infty}\theta x\left(\frac{1}{x}\right)^{\theta+1}dx=\frac{-\theta}{(\theta-1)x^{\theta-1}}\Big|_{1}^{+\infty}=\frac{\theta}{\theta-1}$$

مثال ۵.۶.۴ جعبه‌ای دارای ۲ مهرهٔ سفید و ۳ مهرهٔ سیاه است. از داخل جعبه یک مهره به تصادف خارج می‌کنیم و آن را با دو مهرهٔ دیگر از همان رنگ به جعبه بازمی‌گردانیم. سپس دو مهرهٔ دیگر بدون جایگذاری و به تصادف از جعبه خارج می‌کنیم و متغیر تصادفی X را برابر تعداد مهره‌های سفید مشاهده شده در این دو مهره انتخابی در نظر می‌گیریم. تابع احتمال X و واریانس X را به دست آورید.

حل در این مثال $S_X=\{0,1,2\}$ و اگر W و B را به ترتیب پیشامد انتخاب مهرهٔ سفید و سیاه از جعبه دربار اول در نظر بگیریم در این صورت

$$P(X=0)=P(W)P(X=0\mid W)+P(B)P(X=0\mid B)$$

$$=\frac{2}{5}\times\frac{\binom{3}{2}}{\binom{7}{2}}+\frac{3}{5}\times\frac{\binom{5}{2}}{\binom{7}{2}}=\frac{12}{35}$$

$$P(X=1)=P(W)P(X=1\mid W)+P(B)P(X=1\mid B)$$

$$=\frac{2}{5}\times\frac{\binom{4}{1}\binom{3}{1}}{\binom{7}{2}}+\frac{3}{5}\times\frac{\binom{2}{1}\binom{5}{1}}{\binom{7}{2}}=\frac{18}{35}$$

به همین ترتیب با محاسبه $P(X=2)$ داریم که

x	۰	۱	۲
$f_X(x)$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{5}{35}$

و بنابراین

$$E(X)=\left(0\times\frac{12}{35}\right)+\left(1\times\frac{18}{35}\right)+\left(2\times\frac{5}{35}\right)=\frac{28}{35}$$

$$E(X^2) = (0\times\frac{12}{35}) + (1\times\frac{18}{35}) + (4\times\frac{5}{35}) = \frac{38}{35}$$

$$Var(X) = \frac{38}{35} - (\frac{28}{35})^2 = \frac{546}{1225} = \frac{78}{175}$$

مثال ۶.۶.۴ یک متغیر تصادفی پیوسته X دارای تابع چگالی احتمال زیر است.

$$f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{سایر نقاط} \end{cases}$$

امید ریاضی تابع $g(X) = e^{-2X-3}$ را محاسبه کنید.

حل $E(e^{-2X-3}) = \int_0^{+\infty} e^{-2x-3}(e^{-x})\,dx = \int_0^{+\infty} e^{-3x-3}dx = -\frac{1}{3}e^{-3x-3}\Big|_0^{+\infty} = \frac{1}{3}e^{-3}$

مثال ۷.۶.۴ ظرفی دارای ۳ گوی با شماره‌های ۱ و ۲ و ۳ می‌باشد. ابتدا یک گوی از ظرف خارج می‌شود، سپس یک سکهٔ سالم به تعداد دفعات شمارهٔ گوی خارج شده پرتاب می‌شود.

الف – امید ریاضی تعداد شیرهای مشاهده شده را بدست آورید.

ب – اگر بدانیم حداقل یک شیر مشاهده شده، احتمال اینکه حداکثر ۲ شیر مشاهده شود چقدر است.

حل الف – اگر قرار دهیم X = تعداد شیرهای مشاهده شده در پرتاب سکه

$B_i =$ پیشامد اینکه گوی شمارهٔ i از ظرف خارج شود ، $i=1,2,3$

در این صورت $S_X = \{0,1,2,3\}$ و بنابراین

$$P(X=0) = \sum_{i=1}^{3} P(X=0 \mid B_i)P(B_i) = \frac{1}{3}(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}) = \frac{7}{24}$$

$$P(X=1) = \sum_{i=1}^{3} P(X=1 \mid B_i)P(B_i) = \frac{1}{3}(\frac{1}{2}+\frac{2}{4}+\frac{3}{8}) = \frac{11}{24}$$

به همین ترتیب تابع احتمال X به صورت زیر به دست می‌آید

x	۰	۱	۲	۳
$f_X(x)$	$\frac{7}{24}$	$\frac{11}{24}$	$\frac{5}{24}$	$\frac{1}{24}$

و در نتیجه

$$E(X) = \frac{11}{24} + \frac{10}{24} + \frac{3}{24} = 1$$

ب –

$$P(X \le 2 \mid X \ge 1) = \frac{P(1 \le X \le 2)}{P(X \ge 1)} = \frac{f_X(1)+f_X(2)}{1-f_X(0)} = \frac{\frac{16}{24}}{\frac{17}{24}} = \frac{16}{17}$$

مثال ۸.۶.۴ تابع چگالی احتمال متغیر تصادفی X به صورت زیر داده شده است

$$f_X(x) = \begin{cases} k_1 x & 0 < x < 1 \\ k_2(2-x) & 1 \le x < 2 \\ 0 & \text{سایر نقاط} \end{cases}$$

الف – اگر E(X)=۱ باشد مقدار k_1 و k_2 را تعیین کنید.

ب – اگر $X > \frac{1}{2}$ باشد احتمال آنکه X از $\frac{3}{2}$ کمتر باشد را بیابید.

حل الف – برای تعیین مقادیر k_1 و k_2 دو شرط زیر را داریم

$$1 = \int_{-\infty}^{+\infty} f_X(x)\,dx = \int_0^1 k_1 x\,dx + \int_1^2 k_2(2-x)\,dx = \frac{k_1}{2} + \frac{k_2}{2}$$

$$1 = E(X) = \int_{-\infty}^{+\infty} x f_X(x)\,dx = \int_0^1 k_1 x^2 dx + \int_1^2 k_2 x(2-x)\,dx = \frac{k_1}{3} + \frac{2k_2}{3}$$

با حل دو معادله و دو مجهول بالا مقادیر $k_1 = k_2 = 1$ را به دست می‌آوریم.

ب –

$$P(X < \frac{3}{2} \mid X > \frac{1}{2}) = \frac{P(\frac{1}{2} < X < \frac{3}{2})}{P(X > \frac{1}{2})} = \frac{\int_{\frac{1}{2}}^{1} x\,dx + \int_1^{\frac{3}{2}}(2-x)\,dx}{1 - \int_0^{\frac{1}{2}} x\,dx} = \frac{\frac{3}{8}+\frac{3}{8}}{1-\frac{1}{8}} = \frac{6}{7}$$

مثال ۹.۶.۴ امید ریاضی متغیر تصادفی X که دارای تابع احتمال زیر است را پیدا کنید.

$$f_X(x) = (\frac{1}{2})^x \qquad x = 1,2,3,\ldots$$

حل قرار می‌دهیم S=E(X) بنابراین

$$S = E(X) = \sum_{x=1}^{+\infty} x(\frac{1}{2})^x = \frac{1}{2} + 2(\frac{1}{2})^2 + 3(\frac{1}{2})^3 + \ldots$$

و در نتیجه

$$\frac{1}{2}S = (\frac{1}{2})^2 + 2(\frac{1}{2})^3 + 3(\frac{1}{2})^4 + \ldots$$

با تفاضل این دو رابطه داریم که

$$S - \frac{1}{2}S = \frac{1}{2} + (\frac{1}{2})^2 + (\frac{1}{2})^3 + \ldots = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

بنابراین $\frac{1}{2}S = 1$ و یا $S = 2$.

مثال ۱۰.۶.۴ مقدار k را بگونه‌ای تعیین کنید که تابع زیر یک تابع چگالی احتمال متغیر تصادفی

X باشد و واریانس X را بیابید.

$$f_X(x)=\frac{k}{\beta}\left\{1-\left(\frac{x-\alpha}{\beta}\right)^2\right\} \qquad \alpha-\beta<x<\alpha+\beta \qquad -\infty<\alpha<+\infty,\ \beta>0$$

حل

$$1=\int_{\alpha-\beta}^{\alpha+\beta}\frac{k}{\beta}\left\{1-\left(\frac{x-\alpha}{\beta}\right)^2\right\}dx=k\left\{\left[\frac{x}{\beta}-\frac{1}{3}\left(\frac{x-\alpha}{\beta}\right)^3\right]_{\alpha-\beta}^{\alpha+\beta}\right\}=k\left\{2-\frac{2}{3}\right\}=\frac{4k}{3}$$

بنابراین $k=\frac{3}{4}$ و در نتیجه با تغییر متغیر $y=\frac{x-\alpha}{\beta}$ داریم که

$$E(X)=\frac{3}{4}\int_{\alpha-\beta}^{\alpha+\beta}\frac{x}{\beta}\left\{1-\left(\frac{x-\alpha}{\beta}\right)^2\right\}dx=\frac{3}{4}\int_{-1}^{1}(\alpha+\beta y)(1-y^2)dy$$

$$=\frac{3}{4}\left[\alpha y+\frac{\beta}{2}y^2-\frac{\alpha}{3}y^3-\frac{\beta}{4}y^4\right]_{-1}^{1}=\frac{3}{4}\times\left(\frac{4}{3}\alpha\right)=\alpha$$

$$E(X^2)=\frac{3}{4}\int_{\alpha-\beta}^{\alpha+\beta}\frac{x^2}{\beta}\left\{1-\left(\frac{x-\alpha}{\beta}\right)^2\right\}dx=\frac{3}{4}\int_{-1}^{1}(\alpha+\beta y)^2(1-y^2)dy$$

$$=\frac{3}{4}\left[\alpha^2y+\alpha\beta y^2+\frac{\beta^2-\alpha^2}{3}y^3-\frac{\alpha\beta}{2}y^4-\frac{\beta^2}{5}y^5\right]_{-1}^{1}$$

$$=\frac{3}{4}\left(\frac{4\alpha^2}{3}+\frac{4\beta^2}{15}\right)=\alpha^2+\frac{1}{5}\beta^2$$

بنابراین

$$Var(X)=\left(\alpha^2+\frac{1}{5}\beta^2\right)-\alpha^2=\frac{1}{5}\beta^2$$

مثال ۱۱.۶.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}x+y & 0<x<1\ ,\ 0<y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

ضریب همبستگی X و Y را محاسبه کنید.

حل ابتدا توابع چگالی احتمال حاشیه‌ای X و Y را به دست می‌آوریم

$$f_X(x)=\int_0^1(x+y)dy=\left[xy+\frac{1}{2}y^2\right]_0^1=x+\frac{1}{2} \qquad 0<x<1$$

$$f_Y(y)=\int_0^1(x+y)dx=\left[\frac{1}{2}x^2+xy\right]_0^1=\frac{1}{2}+y \qquad 0<y<1$$

بنابراین X و Y دارای توزیع یکسان هستند.

$$E(X)=\int_0^1 x\left(x+\frac{1}{2}\right)dx=\left[\frac{1}{3}x^3+\frac{1}{4}x^2\right]_0^1=\frac{1}{3}+\frac{1}{4}=\frac{7}{12}$$

$$E(X^2)=\int_0^1 x^2\left(x+\frac{1}{2}\right)dx=\left[\frac{1}{4}x^4+\frac{1}{6}x^3\right]_0^1=\frac{1}{4}+\frac{1}{6}=\frac{5}{12}$$

بنابراین

$$Var(X)=\frac{5}{12}-\left(\frac{7}{12}\right)^2=\frac{11}{144}$$

و در نتیجه $E(Y)=\frac{7}{12}$ و $Var(Y)=\frac{11}{144}$. همچنین

$$E(XY)=\int_0^1\int_0^1 xy(x+y)dxdy=\int_0^1 y\left[\frac{1}{3}x^3+\frac{1}{2}x^2y\right]_0^1dy$$

$$=\int_0^1 y\left(\frac{1}{3}+\frac{1}{2}y\right)dy=\left[\frac{1}{6}y^2+\frac{1}{6}y^3\right]_0^1=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$$

بنابراین

$$COV(X,Y)=\frac{1}{3}-\left(\frac{7}{12}\times\frac{7}{12}\right)=\frac{-1}{144}$$

و در نتیجه

$$\rho=\frac{\frac{-1}{144}}{\sqrt{\frac{11}{144}\times\frac{11}{144}}}=\frac{-1}{11}$$

مثال ۱۲.۶.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر باشند.

(x,y)	$(1,0)$	$(1,1)$	$(1,2)$	$(2,0)$	$(2,1)$	$(2,2)$
$f_{X,Y}(x,y)$	$\frac{1}{6}$	b	$\frac{1}{6}$	a	$\frac{2}{6}$	0

اگر $E(XY)=\frac{7}{6}$ باشد مقادیر a و b را تعیین کنید و $E(Y^2\mid X=1)$ را محاسبه کنید.

حل برای تعیین a و b دو شرط زیر را داریم

$$1=\sum_{x=1}^{2}\sum_{y=0}^{2}f_{X,Y}(x,y)=\frac{4}{6}+a+b$$

$$\frac{7}{6}=E(XY)=\sum_{x=1}^{2}\sum_{y=0}^{2}xyf_{X,Y}(x,y)=0+b+\frac{2}{6}+0+\frac{4}{6}+0=b+1$$

با حل دو معادلهٔ بالا $a=\frac{1}{6}$ و $b=\frac{1}{6}$ به دست می‌آیند. با توجه به رابطهٔ

$$f_{Y\mid X}(y\mid 1)=\frac{f_{X,Y}(1,y)}{f_X(1)} \qquad y=0,1,2$$

از جدول فوق به دست می‌آوریم که

$$f_X(1)=\sum_{y=0}^{2}f_{X,Y}(1,y)=\frac{3}{6}$$

و در نتیجه

y	0	1	2
$f_{Y\mid X}(y\mid 1)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

و بنابراین

$$E(Y^2\mid X=1)=(0\times\frac{1}{3})+(1\times\frac{1}{3})+(4\times\frac{1}{3})=\frac{5}{3}$$

مثال ۱۳.۶.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}8xy & 0<x<y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

$Var(X\mid Y=y)$ و $COV(X,Y)$ را محاسبه کنید.

حل ابتدا توابع چگالی احتمال $f_X(x)$ و $f_Y(y)$ را به دست می‌آوریم

$$f_X(x)=\int_x^1 8xy\,dy=4xy^2\Big|_x^1=4x(1-x^2)\qquad 0<x<1$$

$$f_Y(y)=\int_0^y 8xy\,dx=4x^2y\Big|_0^y=4y^3\qquad 0<y<1$$

بنابراین

$$E(X)=\int_0^1 4x^2(1-x^2)\,dx=\left[\frac{4}{3}x^3-\frac{4}{5}x^5\right]_0^1=\frac{4}{3}-\frac{4}{5}=\frac{8}{15}$$

$$E(Y)=\int_0^1 4y^4\,dy=\frac{4}{5}y^5\Big|_0^1=\frac{4}{5}$$

$$E(XY)=\int_0^1\int_0^y 8x^2y^2\,dx\,dy=\int_0^1\frac{8}{3}y^2\left[x^3\right]_0^y dy=\int_0^1\frac{8}{3}y^5\,dy=\frac{4}{9}y^6\Big|_0^1=\frac{4}{9}$$

در نتیجه

$$COV(X,Y)=\frac{4}{9}-(\frac{8}{15}\times\frac{4}{5})=\frac{4}{225}$$

همچنین

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{8xy}{4y^3}=\frac{2x}{y^2}\qquad 0<x<y<1$$

بنابراین

$$E(X\mid Y=y)=\int_0^y\frac{2x^2}{y^2}\,dx=\frac{2x^3}{3y^2}\Big|_0^y=\frac{2}{3}y$$

$$E(X^2\mid Y=y)=\int_0^y\frac{2x^3}{y^2}\,dx=\frac{x^4}{2y^2}\Big|_0^y=\frac{1}{2}y^2$$

در نتیجه

$$Var(X\mid Y=y)=\frac{1}{2}y^2-(\frac{2}{3}y)^2=\frac{1}{18}y^2$$

مثال ۱۴.۶.۴ فرض کنید متغیرهای تصادفی X و Y مستقل بوده و هر یک دارای واریانس یکسان ۴ باشند. مطلوب است محاسبهٔ

$$Var(2X-3Y+1)\ ,\ COV(X+2Y+3,X-2Y+3)\ ,\ COV(X+Y,X-Y)$$

حل با استفاده از خواص واریانس و کواریانس داریم که

$$COV(X+Y,X-Y)=COV(X,X)-COV(X,Y)+COV(Y,X)-COV(Y,Y)$$
$$=Var(X)-Var(Y)=4-4=0$$

$$COV(X+2Y+3,X-2Y+3)=COV(X+2Y,X-2Y)$$
$$=Var(X)-2COV(X,Y)+2COV(Y,X)-4Var(Y)$$
$$=Var(X)-4Var(Y)=4-16=-12$$

$$Var(2X-3Y+1)=4Var(X)+9Var(Y)-12COV(X,Y)$$
$$=4(4)+9(4)-12(0)=52$$

مثال ۱۵.۶.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}\frac{k}{y}y^2e^{-y}x^{y-1} & 0<x<1\ ,\ 0<y\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار k را تعیین کرده و $P(X<\frac{1}{2}\mid Y=2)$ و $Var(X\mid Y=y)$ را محاسبه کنید.

حل ابتدا مقدار k را به دست می‌آوریم

$$1=\int_0^{+\infty}\int_0^1\frac{k}{2}y^2e^{-y}x^{y-1}\,dx\,dy=\frac{k}{2}\int_0^{+\infty}y^2e^{-y}\left[\frac{x^y}{y}\right]_0^1 dy=\frac{k}{2}\int_0^{+\infty}ye^{-y}\,dy$$

$$=\frac{k}{2}\left[(-y-1)e^{-y}\right]_0^{+\infty}=\frac{k}{2}(0-(-1))=\frac{k}{2}$$

بنابراین k=۲. برای محاسبهٔ احتمال و واریانس شرطی بایستی ابتدا $f_{X\mid Y}(x\mid y)$ را به دست آورد.

$$f_Y(y)=\int_0^1 y^2e^{-y}x^{y-1}\,dx=y^2e^{-y}\left[\frac{x^y}{y}\right]_0^1=ye^{-y}\qquad y>0$$

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{y^2e^{-y}x^{y-1}}{ye^{-y}}=yx^{y-1}\qquad 0<x<1\ ,\ 0<y$$

بنابراین
$$P(X<\tfrac{1}{2}\mid Y=2)=\int_0^{\frac12} f_{X\mid Y}(x\mid 2)dx=\int_0^{\frac12}2x\,dx=\left[x^2\right]_0^{\frac12}=\frac{1}{4}$$

و
$$E(X\mid Y=y)=\int_0^1 yx^y dx=\frac{y}{y+1}\left[x^{y+1}\right]_0^1=\frac{y}{y+1}$$

$$E(X^2\mid Y=y)=\int_0^1 yx^{y+1}dx=\frac{y}{y+2}\left[x^{y+2}\right]_0^1=\frac{y}{y+2}$$

و در نتیجه
$$Var(X\mid Y=y)=\frac{y}{y+2}-\left(\frac{y}{y+1}\right)^2=\frac{y}{(y+2)(y+1)^2}$$

مثال ۱۶.۶.۴ ظرفی حاوی ۷ مهره است که دو تا از آنها با ۱، سه تا از آنها با ۲ و دو تا از آنها با ۳ شماره‌گذاری شده‌اند. دو مهره بدون جایگذاری از این ظرف خارج می‌کنیم و متغیر تصادفی X را برابر شمارهٔ کوچکتر و متغیر تصادفی Y را برابر شمارهٔ بزرگتر در دو مهرهٔ انتخابی در نظر می‌گیریم. کواریانس X و Y را به دست آورید.

حل در این مثال $S_X=\{1,2,3\}$ و $S_Y=\{1,2,3\}$ و همچنین

$$f_{X,Y}(1,1)=P(X=1,Y=1)=P(\{(1,1)\})=\frac{2}{7}\times\frac{1}{6}=\frac{1}{21}$$

$$f_{X,Y}(1,2)=P(X=1,Y=2)=P(\{(1,2),(2,1)\})=2\left(\frac{2}{7}\times\frac{3}{6}\right)=\frac{6}{21}$$

به همین ترتیب تابع احتمال توأم X و Y به صورت زیر به دست می‌آید

y \ x	۱	۲	۳	$f_Y(y)$
۱	$\frac{1}{21}$	۰	۰	$\frac{1}{21}$
۲	$\frac{6}{21}$	$\frac{3}{21}$	۰	$\frac{9}{21}$
۳	$\frac{4}{21}$	$\frac{6}{21}$	$\frac{1}{21}$	$\frac{11}{21}$
$f_X(x)$	$\frac{11}{21}$	$\frac{9}{21}$	$\frac{1}{21}$	

بنابراین
$$E(X)=\frac{11}{21}+\frac{18}{21}+\frac{3}{21}=\frac{32}{21}$$

$$E(Y)=\frac{1}{21}+\frac{18}{21}+\frac{33}{21}=\frac{52}{21}$$

$$E(XY)=\frac{1}{21}+\frac{12}{21}+\frac{12}{21}+\frac{12}{21}+\frac{36}{21}+\frac{9}{21}=\frac{82}{21}$$

در نتیجه
$$COV(X,Y)=\frac{82}{21}-\left(\frac{32}{21}\times\frac{52}{21}\right)=\frac{58}{441}=0/13$$

مثال ۱۷.۶.۴ فرض کنید X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}cxy & x>0,\ y>0,\ 0<x+y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار c را تعیین کرده و $E(X^3\mid Y=y)$ را محاسبه کنید.

حل ابتدا مقدار c را تعیین می‌کنیم.

$$1=\int_0^1\int_0^{1-y}cxy\,dxdy=c\int_0^1 y\left[\frac{1}{2}x^2\right]_0^{1-y}dy=\frac{c}{2}\int_0^1 y(1-y)^2dy$$

$$=\frac{c}{2}\left[\frac{1}{2}y^2-\frac{2}{3}y^3+\frac{1}{4}y^4\right]_0^1=\frac{c}{2}\left[\frac{1}{2}-\frac{2}{3}+\frac{1}{4}\right]=\frac{c}{24}$$

بنابراین $c=24$. همچنین

$$f_Y(y)=\int_0^{1-y}24xy\,dx=12x^2y\Big|_0^{1-y}=12y(1-y)^2\qquad 0<y<1$$

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{24xy}{12y(1-y)^2}=\frac{2x}{(1-y)^2}\qquad 0<x<1-y,\ 0<y<1$$

در نتیجه

$$E(X^3\mid Y=y)=\int_0^{1-y}x^3\left[\frac{2x}{(1-y)^2}\right]dx=\left[\frac{2x^5}{5(1-y)^2}\right]_0^{1-y}=\frac{2(1-y)^3}{5}$$

مثال ۱۸.۶.۴ جعبه‌ای شامل ۴ مهره به شماره‌های ۱، ۲، ۳ و ۴ می‌باشد. از این جعبه دو مهره به تصادف و بدون جایگذاری انتخاب می‌کنیم و متغیر تصادفی X را برابر شمارهٔ اولین مهرهٔ انتخابی و متغیر تصادفی Y را برابر تفاضل بین بزرگترین و کوچکترین شماره‌های انتخابی روی مهره‌ها در نظر می‌گیریم. ضریب همبستگی X و Y و $Var(2X-3Y+4)$ را محاسبه کنید.

حل فضای نمونه حاصل از این آزمایش شامل ۱۲ نقطهٔ هم شانس است و $S_X=\{1,2,3,4\}$ و $S_Y=\{1,2,3\}$ و همچنین

$$f_{X,Y}(1,1)=P(X=1,Y=1)=P(\{(1,2)\})=\frac{1}{12}$$

$$f_{X,Y}(2,1)=P(X=2,Y=1)=P(\{(2,1),(2,3)\})=\frac{2}{12}$$

$$f_{X,Y}(3,1)=P(X=3,Y=1)=P(\{(3,2),(3,4)\})=\frac{2}{12}$$

با انجام محاسبات مربوط به نقاط دیگر، جدول توزیع احتمالات توأم X و Y به صورت زیر به دست می‌آید

y \ x	۱	۲	۳	۴	$f_Y(y)$
۱	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{3}{6}$
۲	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{6}$
۳	$\frac{1}{12}$	۰	۰	$\frac{1}{12}$	$\frac{1}{6}$
$f_X(x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

بنابراین

$$E(X)=\frac{1}{4}(1+2+3+4)=\frac{5}{2}\quad,\quad E(Y)=\frac{3}{6}+\frac{4}{6}+\frac{3}{6}=\frac{5}{3}$$

$$E(X^2)=\frac{1}{4}(1+4+9+16)=\frac{15}{2}\quad,\quad E(Y^2)=\frac{3}{6}+\frac{8}{6}+\frac{9}{6}=\frac{10}{3}$$

$$Var(X)=\frac{15}{2}-\left(\frac{5}{2}\right)^2=\frac{5}{4}\quad,\quad Var(Y)=\frac{10}{3}-\left(\frac{5}{3}\right)^2=\frac{5}{9}$$

همچنین

$$E(XY)=\frac{1}{12}[1+2+3+4+4+6+6+4+8+12]=\frac{25}{6}$$

$$COV(X,Y)=\frac{25}{6}-\left(\frac{5}{2}\times\frac{5}{3}\right)=0$$

بنابراین $\rho(X,Y)=0$ اما توجه کنید که X و Y از یکدیگر مستقل نیستند زیرا

$$f_X(1)f_Y(1)=\frac{1}{4}\times\frac{3}{6}\neq\frac{1}{12}=f_{X,Y}(1,1)$$

همچنین

$$Var(2X-3Y+4)=4Var(X)+9Var(Y)-12COV(X,Y)$$
$$=4\left(\frac{5}{4}\right)+9\left(\frac{5}{9}\right)-12(0)=10$$

مثال ۱۹.۶.۴ تابع چگالی احتمال متغیرهای تصادفی X و Y به صورت زیر داده شده است

$$f_{X,Y}(x,y)=\begin{cases}\frac{1}{y}e^{-(y+\frac{x}{y})} & x>0\ ,\ y>0\\ 0 & \text{سایر نقاط}\end{cases}$$

مقادیر $E(X^2\mid Y=y)$ و $COV(X,Y)$ را به دست آورید.

حل توجه کنید که تابع چگالی احتمال حاشیه‌ای $f_X(x)$ را نمی‌توان برای محاسبهٔ $E(X)$ به دست آورد. بنابراین با استفاده از رابطهٔ (۴.۴) داریم که

$$E(X)=\int_0^{+\infty}\int_0^{+\infty}x\left[\frac{1}{y}e^{-(y+\frac{x}{y})}\right]dxdy=\int_0^{+\infty}e^{-y}\left[(-x-y)e^{-\frac{x}{y}}\right]_0^{+\infty}dy$$

$$=\int_0^{+\infty}ye^{-y}dy=(-y-1)e^{-y}\Big|_0^{+\infty}=1$$

اما

$$f_Y(y)=\int_0^{+\infty}\frac{1}{y}e^{-(y+\frac{x}{y})}dx=e^{-y}\left[-e^{-\frac{x}{y}}\right]_0^{+\infty}=e^{-y}\qquad y>0$$

بنابراین

$$E(Y)=\int_0^{+\infty}ye^{-y}dy=1$$

و همچنین

$$E(XY)=\int_0^{+\infty}\int_0^{+\infty}xy\left(\frac{1}{y}e^{-(y+\frac{x}{y})}\right)dxdy=\int_0^{+\infty}ye^{-y}\left[\int_0^{+\infty}\frac{x}{y}e^{-\frac{x}{y}}dx\right]dy$$

$$=\int_0^{+\infty}ye^{-y}(y)dy=\left[(-y^2-2y-2)e^{-y}\right]_0^{+\infty}=2$$

در نتیجه

$$COV(X,Y)=2-(1\times 1)=1$$

همچنین

$$f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{\frac{1}{y}e^{-(y+\frac{x}{y})}}{e^{-y}}=\frac{1}{y}e^{-\frac{x}{y}}\qquad x>0\ ,\ y>0$$

بنابراین

$$E(X^2\mid Y=y)=\int_0^{+\infty}\frac{x^2}{y}e^{-\frac{x}{y}}dx=\left[(-x^2-2xy-2y^2)e^{-\frac{x}{y}}\right]_0^{+\infty}=2y^2$$

مثال ۲۰.۶.۴ نشان دهید که اگر X و Y مستقل باشند آنگاه $E(Y\mid X=x)$ ثابت است (به x بستگی ندارد) و آن مقدار ثابت را حساب کنید. اگر $E(Y\mid X=x)$ ثابت باشد آیا X و Y لزوماً مستقل هستند؟ اگر $E(Y\mid X=x)$ و $E(X\mid Y=y)$ هر دو ثابت باشند آیا X و Y لزوماً مستقل

هستند؟

حل اگر X و Y مستقل باشند آنگاه $f_{Y|X}(y|x)=f_Y(y)$ و در نتیجه

$$E(Y|X=x)=\int_{-\infty}^{+\infty} y f_{Y|X}(y|x)\,dy=\int_{-\infty}^{+\infty} y f_Y(y)\,dy=E(Y)$$

پس $E(Y|X=x)$ بستگی به x ندارد (مقداری ثابت است) و مقدار آن برابر $E(Y)$ است. اما اگر متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر باشند

y \ x	-۱	۰	۱	$f_Y(y)$
-۱	۰	$\frac{1}{4}$	۰	$\frac{1}{4}$
۰	$\frac{1}{4}$	۰	$\frac{1}{4}$	$\frac{2}{4}$
۱	۰	$\frac{1}{4}$	۰	$\frac{1}{4}$
$f_X(x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

در این صورت طبق مثال ۴.۳.۴، X و Y از یکدیگر مستقل نیستند و همچنین به راحتی دیده می‌شود که

$$E(X|Y=y)=0 \quad y=-1,0,1 \quad , \quad E(Y|X=x)=0 \quad x=-1,0,1$$

یعنی $E(Y|X=x)$ و $E(X|Y=y)$ مقادیر ثابتی هستند ولی X و Y از یکدیگر مستقل نیستند.

مثال ۲۱.۶.۴ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}\frac{2}{x}e^{-2x} & 0<y<x<+\infty\\ 0 & \text{سایر نقاط}\end{cases}$$

مقادیر $Var(Y|X=x)$ و $\rho(X,Y)$ را به دست آورید.

حل با توجه به رابطهٔ (۴.۴) و استفاده از انتگرال‌گیری به روش جزء به جزء داریم که

$$E(Y)=\int_0^{+\infty}\int_0^x y\left(\frac{2}{x}e^{-2x}\right)dy\,dx=\int_0^{+\infty}xe^{-2x}dx=\left[\left(-\frac{1}{2}x-\frac{1}{4}\right)e^{-2x}\right]_0^{+\infty}=\frac{1}{4}$$

$$E(Y^2)=\int_0^{+\infty}\int_0^x y^2\left(\frac{2}{x}e^{-2x}\right)dy\,dx=\frac{2}{3}\int_0^{+\infty}x^2e^{-2x}dx$$

$$=\frac{2}{3}\left[\left(-\frac{1}{2}x^2-\frac{1}{2}x-\frac{1}{4}\right)e^{-2x}\right]_0^{+\infty}=\frac{1}{6}$$

بنابراین $Var(Y)=\frac{1}{6}-\left(\frac{1}{4}\right)^2=\frac{5}{48}$. همچنین

$$f_X(x)=\int_0^x\frac{2}{x}e^{-2x}dy=2e^{-2x} \qquad x>0$$

بنابراین

$$E(X)=\int_0^{+\infty}2xe^{-2x}dx=\frac{1}{2} \quad \text{و} \quad E(X^2)=\int_0^{+\infty}2x^2e^{-2x}dx=\frac{1}{2}$$

و در نتیجه $Var(X)=\frac{1}{2}-\left(\frac{1}{2}\right)^2=\frac{1}{4}$. از طرفی

$$E(XY)=\int_0^{+\infty}\int_0^x xy\left[\frac{2}{x}e^{-2x}\right]dy\,dx=\int_0^{+\infty}x^2e^{-2x}dx=\frac{1}{4}$$

بنابراین $COV(X,Y)=\frac{1}{4}-\left(\frac{1}{2}\times\frac{1}{4}\right)=\frac{1}{8}$ و در نتیجه

$$\rho(X,Y)=\frac{\frac{1}{8}}{\sqrt{\frac{1}{4}\times\frac{5}{48}}}=\sqrt{\frac{3}{5}}$$

همچنین

$$f_{Y|X}(y|x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{\frac{2}{x}e^{-2x}}{2e^{-2x}}=\frac{1}{x} \qquad 0<y<x<+\infty$$

بنابراین

$$E(Y|X=x)=\int_0^x y\left(\frac{1}{x}\right)dy=\frac{1}{2}x$$

$$E(Y^2|X=x)=\int_0^x y^2\left(\frac{1}{x}\right)dy=\frac{1}{3}x^2$$

و در نتیجه

$$Var(Y|X=x)=\frac{1}{3}x^2-\left(\frac{1}{2}x\right)^2=\frac{1}{12}x^2$$

۷.۴ تمرینات

۱ فرض کنید متغیر تصادفی X دارای تابع احتمال زیر باشد

$$f_X(x)=\begin{cases}\frac{1}{4} & x=2,4,8,16\\ 0 & \text{سایر نقاط}\end{cases}$$

مطلوب است محاسبهٔ $E(X)$، $E\left(\frac{1}{X}\right)$، $E(2^X)$ و $Var(X)$.

۲ اگر تابع چگالی احتمال متغیر تصادفی X به صورت زیر باشد

$$f_X(x)=\begin{cases}k(2-x) & 0<x<2\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار k را تعیین کنید و $E(X)$ و $Var(X)$ را محاسبه کنید.

۳ فرض کنید متغیر تصادفی X دارای تابع توزیعی (تجمعی) زیر باشد

$$F_X(x)=\begin{cases}0 & x<1\\ k(x^2-1) & 1\le x<4\\ 1 & 4\le x\end{cases}$$

مقدار k را تعیین کنید و واریانس X را بدست آورید.

۴ شخصی می‌خواهد اتومبیل خود را به مبلغ ۱۰۰۰۰ تومان بیمه کند، شرکت بیمه تخمین می‌زند که کل مبلغ را با احتمال ۰/۰۰۲ و نصف آن را با احتمال ۰/۰۱ و ۲۵٪ آن را با احتمال ۰/۱ باید بپردازد. شرکت بیمه چه حق بیمه‌ای باید در نظر بگیرد تا سودی معادل ۱۰۰ تومان داشته باشد.

۵ اگر متغیر تصادفی X دارای میانگین $\frac{1}{3}$ و تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}ax+b & 0<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

مقادیر a و b را تعیین کنید و واریانس X را محاسبه کنید.

۶ اگر X دارای تابع چگالی احتمال زیر باشد، در صورت وجود E(X) را محاسبه کنید.

الف-
$$f_X(x)=\begin{cases}\frac{1}{4}xe^{-\frac{x}{2}} & x>0\\ 0 & \text{سایر نقاط}\end{cases}$$

ب-
$$f_X(x)=\begin{cases}c(1-x^2) & -1<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

ج-
$$f_X(x)=\begin{cases}\frac{5}{x^2} & x>5\\ 0 & x\le 5\end{cases}$$

۷ اگر X یک متغیر تصادفی با تابع احتمال زیر باشد

x	-3	6	9
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

ابتدا $E(X)$ و $E(X^2)$ را به دست آورید و با استفاده از قوانین امید ریاضی $E\left[(2X+1)^2\right]$ را محاسبه کنید.

۸ اگر برای یک متغیر تصادفی X داشته باشیم که

$$E\left[(X-2)^2\right]=6 \quad , \quad E\left[(X-1)^2\right]=10$$

در این صورت مقادیر μ و σ^2 را محاسبه کنید.

۹ از جعبه‌ای که شامل ۴ توپ سیاه و ۲ توپ سبز می‌باشد ۳ توپ به طور متوالی انتخاب می‌گردد و هر توپ قبل از انتخاب توپ دیگر در جعبه قرار می‌گیرد. امید ریاضی تعداد توپهای سبز انتخاب شده را به دست آورید.

۱۰ اگر θ یک مقدار ثابت معلوم باشد و

$$f_X(x)=\begin{cases}\frac{1}{2}\theta & 0<x<1\\ \frac{1}{2} & 1\le x\le 2\\ \frac{1}{2}(1-\theta) & 2<x<3\\ 0 & \text{سایر نقاط}\end{cases}$$

تابع توزیع X و میانگین و واریانس X را به دست آورید.

۱۱ اگر X دارای تابع احتمالی به صورت زیر باشد

x	0	1	2
$P(X=x)$	p	$1-2p$	p

به ازای چه مقداری از p واریانس X ماکزیمم می‌شود؟

۱۲ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}x^2 & 0<x\le 1\\ c & 1<x<a\\ 0 & \text{سایر نقاط}\end{cases}$$

اگر $E(X)=\frac{5}{4}$ باشد مقادیر a و c را تعیین کنید و $Var(2X+3)$ را محاسبه کنید.

۱۳ با فرض اینکه متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}k(1-|x|) & |x|<1\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار k و $P(-\frac{1}{2}<X<\frac{1}{2})$ و $E(X)$ را به دست آورید.

۱۴ یک کلاس آمار ۱۰ شاگرد دارد که ۵ نفر آنها ۱۹ ساله، ۳ نفر آنها ۲۰ ساله و ۲ نفر آنها ۲۱ ساله هستند. از این کلاس ۲ شاگرد به تصادف و بدون جایگذاری انتخاب می‌کنیم. فرض کنید متغیر تصادفی X میانگین سن دو شاگرد انتخابی باشد. تابع احتمال متغیر تصادفی X را به دست آورید و واریانس X را محاسبه کنید.

۱۵ میانگین و واریانس متغیر تصادفی X که دارای تابع چگالی احتمال زیر می‌باشد را به دست آورید.

الف - $x>0$ ، $f_X(x)=\frac{\lambda^3}{2}x^2e^{-\lambda x}$ ، $\lambda>0$ مقداری ثابت

ب - $-\infty<x<+\infty$ $f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$

ج - $-\infty<x<+\infty$ $f_X(x)=\frac{1}{2\beta}e^{-\frac{|x|}{\beta}}$ ، $\beta>0$ مقداری ثابت

۱۶ میانگین متغیر تصادفی X که دارای تابع احتمال زیر می‌باشد را به دست آورید.

الف - $x=0,1,\dots,n$ $f_X(x)=\binom{n}{x}p^x(1-p)^{n-x}$

n عدد صحیح نامنفی و $0<p<1$ مقداری ثابت

ب - $x=0,1,2,\dots$ $f_X(x)=\frac{e^{-\lambda}\lambda^x}{x!}$ ، $\lambda>0$ مقداری ثابت

ج - $x=0,1,2,\dots$ $f_X(x)=p(1-p)^x$ ، $0<p<1$ مقداری ثابت

۱۷ سکه سالمی را سه مرتبه پرتاب می‌کنیم، اگر X برابر تعداد شیرهای مشاهده شده در ۳ مرتبه پرتاب سکه و Y برابر تعداد شیرهای مشاهده شده در ۲ پرتاب اول باشند، تابع احتمال توأم X و Y را به دست آورده و ضریب همبستگی X و Y را محاسبه کنید.

۱۸ اگر X و Y دو متغیر تصادفی با تابع چگالی احتمال توأم زیر باشند.

$$f_{X,Y}(x,y)=\begin{cases}\frac{3}{5}x(x+y) & 0<x<1\ ,\ 0<y<2\\ 0 & \text{سایر نقاط}\end{cases}$$

مطلوب است محاسبهٔ $E(X)$ ، $E(Y)$ ، $E(X+Y)$ و $E(XY)$.

۱۹ فرض کنید متغیرهای تصادفی X و Y دارای تابع احتمال توأم زیر باشند. نشان دهید که $\rho(X,Y)=0$ اما X و Y از یکدیگر مستقل نیستند.

y \ x	-۱	۰	۱
-۱	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$
۰	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{2}{16}$
۱	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$

۲۰ اگر X و Y دو متغیر تصادفی با تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}e^{-(x+y)} & x\ge 0\ ,\ y\ge 0\\ 0 & \text{سایر نقاط}\end{cases}$$

ضریب همبستگی X و Y و $Var(2X-1)$ و $Var(4X+3Y-2)$ را محاسبه کنید.

۲۱ توزیع احتمال توأم دو متغیر تصادفی X و Y به صورت زیر می‌باشد

y \ x	۱	۳	۹
۲	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{12}$
۴	$\frac{1}{4}$	$\frac{1}{4}$	۰
۶	$\frac{1}{8}$	$\frac{1}{24}$	$\frac{1}{12}$

الف - امید ریاضی X و امید ریاضی Y را به دست آورید.

ب - $COV(X,Y)$ را محاسبه کنید، آیا X و Y از یکدیگر مستقل هستند؟

۲۲ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}1 & 0<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

اگر $Y=X^2$ باشد $\rho(X,Y)$ و $\rho(X+1,2Y-1)$ را محاسبه کنید.

۲۳ اگر متغیرهای تصادفی X و Y از یکدیگر مستقل بوده و به ترتیب دارای میانگینهای ۲ و ۳ و واریانسهای ۴ و ۵ باشند، امید ریاضی XY و $(X+Y)(X-Y)$ را حساب کنید.

۲۴ فرض کنید X و Y متغیرهای تصادفی مستقل با توابع چگالی احتمال $f_X(x)=\frac{8}{x^3}\quad x>2$ و $f_Y(y)=2y\quad 0<y<1$ باشند. امید ریاضی تابع Z=XY را به دست آورید.

۲۵ الف- اگر X و Y دارای میانگینهای ۱ و ۲ و واریانسهای ۳ و ۴ و کواریانس ۱- باشند، اولاً میانگین و واریانس Z=۲X-Y+۱ را محاسبه کنید و ثانیاً مقدار m را به گونه‌ای تعیین کنید که U=mX+Y و V=X-mY ناهمبسته باشند.

ب- اگر بین X و Y رابطهٔ خطی X+۲Y-۱=۰ برقرار باشد و $Var(X)=4$، مقادیر $COV(X,Y)$ و $\rho(X,Y)$ را محاسبه کنید.

۲۶ فرض کنید که

$$f_{X,Y}(x,y)=\begin{cases}6(1-x-y) & 0<x<1\ ,\ 0<y<1-x\\ 0 & \text{سایر نقاط}\end{cases}$$

کواریانس X و Y و $Var(Y\mid X=x)$ را به دست آورید.

۲۷ فرض کنید که

$$f_X(x)=\begin{cases}1 & 0<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

$$f_{Y\mid X}(y\mid x)=\binom{n}{y}x^y(1-x)^{n-y}\qquad y=0,1,\ldots,n$$

تابع احتمال Y را به دست آورده و E(Y) را محاسبه کنید.

۲۸ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}k & x^2<y<1\ ,\ 0<x<1\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار k را تعیین کنید و $E(X\mid Y=y)$ و $\rho(X,Y)$ را محاسبه کنید.

۲۹ اگر X_1، X_2، ... و X_n و همچنین Y_1، Y_2، ... و Y_m متغیرهای تصادفی باشند، نشان دهید که

الف- $$COV\Big(\sum_{i=1}^{n}a_iX_i,\sum_{j=1}^{m}b_jY_j\Big)=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ib_j\,COV(X_i,Y_j)$$

ب- $$Var\Big(\sum_{i=1}^{n}a_iX_i\Big)=\sum_{i=1}^{n}a_i^2\,Var(X_i)+2\sum_{i=1}^{n}\sum_{j=i+1}^{n}a_i\,a_j\,COV(X_i,X_j)$$

۳۰ تابع چگالی احتمال توأم X و Y به صورت زیر می‌باشد

$$f_{X,Y}(x,y)=\begin{cases}k(x+y) & 0<x+y<1\ ,\ 0<x<1\ ,\ 0<y<1\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار k را تعیین کرده و $COV(X,Y)$ و $Var(Y\mid X=x)$ را محاسبه کنید.

۳۱ فرض کنید که

$$f_{X,Y}(x,y)=\begin{cases}\frac{3}{2} & 0<x<1\quad -(x-1)^2<y<(x-1)^2\\ 0 & \text{سایر نقاط}\end{cases}$$

نشان دهید که X و Y از یکدیگر مستقل نیستند و $E(X\mid Y=y)$ و $E(Y\mid X=x)$ را محاسبه کنید.

۳۲ جعبه‌ای شامل ۳ توپ سفید، ۲ توپ قرمز و ۲ توپ سیاه است. از این جعبه به تصادف ۲ توپ یک به یک و بدون جایگذاری انتخاب کرده و قرار می‌دهیم

X = تعداد توپهای سفید مشاهده شده

Y = تعداد توپهای قرمز مشاهده شده

الف- تابع احتمال توأم X و Y را به دست آورید.

ب- ضریب همبستگی X و Y و $E(Y\mid X=x)$ را محاسبه کنید.

۳۳ فرض کنید X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases}c(y-x)^a & 0<x<y<1\ ,\ a>0\\ 0 & \text{سایر نقاط}\end{cases}$$

مقدار c را تعیین کرده و $P\big(0<X<\frac{1}{4}\mid Y=\frac{1}{2}\big)$ و $Var(X\mid Y=y)$ را محاسبه کنید.

۳۴ جعبه‌ای دارای ۴ مهرهٔ قرمز و ۳ مهرهٔ سفید است و سکه‌ای وجود دارد که شانس مشاهدهٔ شیر در آن دو برابر خط است. ابتدا سکه را یک مرتبه پرتاب می‌کنیم اگر شیر مشاهده شد ۲ مهرهٔ قرمز و اگر خط مشاهده شد ۲ مهرهٔ سفید به جعبه اضافه می‌کنیم و سپس ۳ مهره از جعبه خارج می‌کنیم و قرار می‌دهیم

$X=$ تعداد مهره‌های سفید مشاهده شده در این ۳ مهره

$Y=$ تعداد مهره‌های قرمز مشاهده شده در این ۳ مهره

تابع احتمال توأم X و Y را به دست آورده و $\rho(X,Y)$ را محاسبه کنید.

۳۵ تابع چگالی احتمال توأم X و Y به صورت زیر است

$$f_{X,Y}(x,y)=\begin{cases} cxye^{-(x^2+y^2)} & x>0\ ,\ y>0 \\ 0 & \text{سایر نقاط} \end{cases}$$

مقدار c را تعیین کنید و $P(X<Y<2X)$ و $COV(X,Y)$ را محاسبه کنید.

۳۶ اگر تابع چگالی احتمال توأم X و Y به صورت زیر باشد

$$f_{X,Y}(x,y)=\begin{cases} kxy & x>0\ ,\ y>0\ ,\ x^2+y^2\le 1 \\ 0 & \text{سایر نقاط} \end{cases}$$

مقدار k را تعیین کنید و $P(X\le Y)$ و $Var(X\mid Y=y)$ را محاسبه کنید.

۳۷ اگر

$$f_{X\mid Y}(x\mid y)=\begin{cases} \dfrac{k+a}{y} & 0<x<y \\ 0 & \text{سایر نقاط} \end{cases}\quad ,\quad f_Y(y)=\begin{cases} 2y & 0<y<k-a \\ 0 & \text{سایر نقاط} \end{cases}$$

مقادیر a و k را تعیین کنید و $E(Y\mid X=x)$ و ضریب همبستگی X و Y را محاسبه کنید.

۳۸ ظرف A شامل ۱ گلولهٔ سفید و ۱ گلولهٔ سیاه، ظرف B شامل ۲ گلولهٔ سفید و ۲ گلولهٔ سیاه است. به تصادف و با جایگذاری ۲ گلوله از A خارج می‌کنیم، اگر هم رنگ باشند ۱ گلوله سفید و اگر هم رنگ نباشند یک گلولهٔ سیاه به ظرف B اضافه می‌کنیم و سپس به تصادف از B یک گلوله خارج می‌کنیم. فرض کنید X و Y به ترتیب تعداد گلوله‌های سفید خارج شده از A و B باشند. تابع احتمال توأم X و Y را به دست آورده و $P(X+Y\le 1)$، $Var(Y\mid X=0)$ و $Var(X+Y-2)$ را محاسبه کنید.

۳۹ فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی احتمال توأم زیر باشند

$$f_{X,Y}(x,y)=\begin{cases} 1 & 0<x<1\ ,\ x<y<x+1 \\ 0 & \text{سایر نقاط} \end{cases}$$

$E(X\mid Y=y)$ و ضریب همبستگی X و Y را به دست آورید.

فصل پنجم

برخی توزیعهای احتمال

۱.۵ مقدمه

در فصل سوم در مورد به دست آوردن توزیع احتمال یک متغیر تصادفی بحث کردیم. در این فصل می‌خواهیم توزیع احتمال چند متغیر تصادفی بخصوص را به دست آوریم.

بعضی از متغیرهای تصادفی دارای یک الگوی خاص هستند و می‌توان برای آنها یک نوع توزیع احتمال را در نظر گرفت. برای مثال در انجام آزمایشات پرتاب یک تاس ۷ مرتبه و پرتاب یک نیزه به طرف هدف، فرض کنید که

$X=$ تعداد مشاهده عدد ۴ در ۷ مرتبه پرتاب تاس

$Y=$ تعداد برخورد به هدف نیزه در ۵ مرتبه پرتاب نیزه

همانطور که دیده می‌شود این دو متغیر تصادفی دارای یک شکل بخصوص هستند، در حقیقت هر دو تعداد موفقیتها در n آزمایش مستقل را بیان می‌کنند. حال اگر توزیع احتمال برای تعداد موفقیتها در n آزمایش مستقل را به دست آوریم آنگاه توزیع احتمال برای تعداد زیادی متغیر تصادفی مشابه با X و Y گفته شده در بالا را به دست آورده‌ایم.

در این فصل ابتدا چند توزیع احتمال گسسته خاص و سپس چند توزیع احتمال پیوسته خاص را مورد بررسی قرار می‌دهیم.

۲.۵ توزیع برنولی[1]

آزمایشی را در نظر بگیرید که دارای دو نتیجه موفقیت با احتمال p و شکست با احتمال $q=1-p$ باشد. چنین آزمایشی را آزمایش برنولی گویند. برای مثال پرتاب یک سکه یک آزمایش برنولی است که در آن موفقیت مشاهده شیر با $p=\frac{1}{2}$ و مشاهده خط شکست با $q=\frac{1}{2}$ می‌باشد. همچنین در پرتاب یک تاس که در آن موفقیت مشاهده عدد ۴ با $p=\frac{1}{6}$ و مشاهده عدد غیر از ۴ شکست با $q=\frac{5}{6}$ است. اگر متغیر تصادفی X را به صورت زیر تعریف کنیم

$$X=\begin{cases}1 & \text{اگر نتیجه آزمایش موفقیت باشد}\\ 0 & \text{اگر نتیجه آزمایش شکست باشد}\end{cases}$$

در این صورت متغیر تصادفی X را متغیر تصادفی برنولی گوئیم و آن را با نماد $X\sim B(1,p)$ نمایش داده و گوئیم X دارای توزیع برنولی با پارامتر p است (عدد ۱ نمایانگر انجام یک بار آزمایش است). تابع احتمال متغیر تصادفی برنولی به صورت زیر به دست می‌آید

x	0	1
$f_X(x)=P(X=x)$	$1-p$	p

که آن را می‌توان در فرمول زیر خلاصه کرد

$$\boxed{\text{تابع احتمال توزیع برنولی}\qquad f_X(x)=p^x(1-p)^{1-x},\quad x=0,1} \qquad (1.5)$$

در قضیه زیر امید ریاضی و واریانس این توزیع را به دست می‌آوریم.

قضیه ۱.۵ اگر $X\sim B(1,p)$ آنگاه $E(X)=p$ و $Var(X)=pq$.

اثبات

$$E(X)=(0)(1-p)+(1)(p)=p$$

$$E(X^2)=(0^2)(1-p)+(1^2)(p)=p$$

بنابراین

$$Var(X)=p-p^2=p(1-p)=pq$$

مثال ۱.۲.۵ یک تاس را یک مرتبه پرتاب می‌کنیم و قرار می‌دهیم:

1- Bernoulli Distribution

$$X=\begin{cases}1 & \text{اگر عدد ۴ مشاهده شود}\\ 0 & \text{در غیر اینصورت}\end{cases}$$

در این صورت $X\sim B(1,\frac{1}{6})$ و همچنین

$$f_X(x)=\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{1-x},\quad x=0,1$$

$$E(X)=\frac{1}{6},\quad Var(X)=\frac{1}{6}\left(\frac{5}{6}\right)=\frac{5}{36}$$

۳.۵ توزیع دو جمله‌ای[1]

آزمایش تصادفی که از انجام چند آزمایش مستقل برنولی بوجود می‌آید را آزمایش دو جمله‌ای گویند. برای مثال انتخاب ۵ مهره با جایگذاری از ظرفی که دارای ۸ مهره سفید و ۶ مهره قرمز است و موفقیت برای ما مشاهده مهره قرمز است یک آزمایش دو جمله‌ای با احتمال موفقیت $\frac{6}{14}$ می‌باشد.

به طور کلی یک آزمایش دو جمله‌ای دارای خواص زیر است.

۱ـ آزمایش دو جمله‌ای از انجام n آزمایش مستقل برنولی بوجود آمده است.

۲ـ در آزمایشات برنولی احتمال موفقیت p (و شکست $q=1-p$) می‌باشد که در تمام آزمایشات برنولی مقداری ثابت است.

مثال ۱.۳.۵ سکه‌ای که شانس مشاهده شیر در آن دو برابر خط است را ۴ مرتبه پرتاب می‌کنیم. اگر موفقیت مشاهده شیر باشد، در این صورت یک آزمایش دو جمله‌ای با پارامترهای $n=4$ و $p=\frac{2}{3}$ داریم.

مثال ۲.۳.۵ بسکتبالیستی ۶۰٪ از توپهایش گل می‌شود. اگر او ۵ پرتاب مستقل انجام دهد و موفقیت برای او گل شدن توپ باشد، در این صورت یک آزمایش دو جمله‌ای با پارامترهای $n=5$ و $p=0/6$ داریم.

تعریف ۱.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

$$\boxed{X = \text{تعداد موفقیتها در n آزمایش مستقل برنولی}}$$

آنگاه X را یک متغیر تصادفی دو جمله‌ای گویند. اگر احتمال موفقیت در هر آزمایش برنولی p

1- Binomial Distribution

باشد، آنگاه گوئیم که X دارای توزیع دو جمله‌ای با پارامترهای n و p است و آن را با نماد $X\sim B(n,p)$ نمایش می‌دهیم.

برای به دست آوردن تابع احتمال توزیع دو جمله‌ای ابتدا مثال زیر را می‌آوریم.

مثال ۳.۳.۵ سکه‌ای که شانس مشاهده شیر در آن دو برابر خط است را ۴ مرتبه پرتاب می‌کنیم اگر

X= تعداد شیرهای مشاهده شده در ۴ مرتبه پرتاب سکه

الف - $P(X=۳)$ را محاسبه کنید.

ب - تابع احتمال X را به دست آورید.

حل با توجه به تعریف ۱.۵ داریم که $X\sim B(۴,\frac{۲}{۳})$ بنابراین

الف - $P(X=۳)$ بدین معنی است که در ۴ مرتبه پرتاب سکه می‌خواهیم ۳ شیر مشاهده کنیم که یک حالت خاص آن HHHT با احتمال $(\frac{۱}{۳})(\frac{۲}{۳})^۳$ است و تعداد حالات مشاهده ۳ شیر در ۴ مرتبه پرتاب سکه برابر $\binom{۴}{۳}$ است. بنابراین

$$P(X=۳)=\binom{۴}{۳}(\frac{۲}{۳})^۳(\frac{۱}{۳})^{۴-۳}$$

ب - با انجام عملیات مشابه قسمت الف تابع احتمال X به صورت زیر به دست می‌آید.

$$f_X(x)=P(X=x)=\binom{۴}{x}(\frac{۲}{۳})^x(\frac{۱}{۳})^{۴-x}\quad , \quad x=۰,۱,۲,۳,۴$$

در حالت کلی اگر $X\sim B(n,p)$ آنگاه $P(X=x)$ بدین معنی است که در n آزمایش مستقل برنولی ما x موفقیت و n-x شکست داشته باشیم که در یک حالت خاص احتمال آن برابر $p^x q^{n-x}$ است و تعداد حالاتی که می‌توان در n آزمایش x موفقیت مشاهده کرد برابر $\binom{n}{x}$ است و بنابراین تابع احتمال X برابر است با

$$\boxed{f_X(x)=\binom{n}{x}p^x q^{n-x}\ ,\ x=۰,۱,۲,...,n \qquad \text{تابع احتمال توزیع دوجمله‌ای}} \qquad (۲.۵)$$

توجه کنید که توزیع برنولی یک حالت خاص توزیع دو جمله‌ای با پارامتر n=۱ است. اگر در n آزمایش برنولی مستقل، متغیر تصادفی X_i برابر نتیجه i-امین آزمایش باشد آنگاه $X=X_۱+X_۲+...+X_n$ تعداد موفقیتها در این n آزمایش برنولی مستقل می‌باشد و $X\sim B(n,p)$ است. در زیر امید ریاضی و واریانس توزیع دو جمله‌ای را به دست می‌آوریم.

قضیه ۲.۵ اگر $X\sim B(n,p)$ آنگاه $E(X)=np$ و $Var(X)=npq$.

اثبات با توجه به مطلب بالا داریم که $X=X_۱+X_۲+...+X_n$ که $X_۱$، $X_۲$، ... و X_n از یکدیگر مستقل و هر یک دارای توزیع برنولی با پارامتر p هستند. بنابراین طبق خواص امید ریاضی و واریانس داریم که

$$E(X)=E(\sum_{i=۱}^{n}X_i)=\sum_{i=۱}^{n}E(X_i)=\sum_{i=۱}^{n}p=np$$

$$Var(X)=Var(\sum_{i=۱}^{n}X_i)=\sum_{i=۱}^{n}Var(X_i)=\sum_{i=۱}^{n}pq=npq$$

مثال ۴.۳.۵ یک تاس را ۵ بار پرتاب می‌کنیم.

الف - احتمال اینکه عدد ۴ دقیقاً ۳ بار مشاهده شود را بیابید؟

ب - احتمال اینکه عدد ۴ حداکثر ۲ بار مشاهده شود را بیابید؟

حل اگر متغیر تصادفی X برابر تعداد مشاهده شیر در ۵ بار پرتاب سکه باشد آنگاه $X\sim B(۵,\frac{۱}{۶})$ است که تابع احتمال آن به صورت زیر می‌باشد

$$f_X(x)=\binom{۵}{x}(\frac{۱}{۶})^x(\frac{۵}{۶})^{۵-x}\quad , \quad x=۰,۱,...,۵$$

بنابراین

الف- $P(X=۳)=f_X(۳)=\binom{۵}{۳}(\frac{۱}{۶})^۳(\frac{۵}{۶})^{۵-۳}=۰/۰۳۲$

ب- $P(X\le ۲)=\sum_{x=۰}^{۲}f_X(x)=\binom{۵}{۰}(\frac{۱}{۶})^۰(\frac{۵}{۶})^۵+\binom{۵}{۱}(\frac{۱}{۶})^۱(\frac{۵}{۶})^۴+\binom{۵}{۲}(\frac{۱}{۶})^۲(\frac{۵}{۶})^۳$

$$=۰/۹۶۴۵$$

برای محاسبه احتمالات در توزیع دو جمله‌ای جدول (I) در ضمیمه ارایه گردیده است که در این جدول $P(X\le r)$ برای مقادیر $n=۵,۱۰,۱۵,۲۰$ و $p=۰/۱,۰/۲,۰/۲۵,...,۰/۹$ محاسبه گردیده است. برای مثال اگر $X\sim B(۱۰,۰/۳)$ باشد آنگاه از جدول (I) داریم که $P(X\le ۴)=۰/۸۴۹۷$.

مثال ۵.۳.۵ یک آزمون انگلیسی شامل ۲۰ سئوال پنج گزینه‌ای است که در هر سئوال تنها یک گزینه درست می‌باشد. شخصی که اصلاً انگلیسی نمی‌داند در این آزمون شرکت می‌کند و سئوالات را به تصادف پاسخ می‌دهد.

الف – انتظار دارید که چند سئوال را درست پاسخ دهد؟

ب – احتمال اینکه حداقل ۱۰ سئوال را درست پاسخ دهد را بیابید.

ج – احتمال اینکه بین ۲ تا ۷ سئوال را درست پاسخ دهد را بیابید.

د – احتمال اینکه دقیقاً ۹ سئوال را درست پاسخ دهد را بیابید.

حل اگر متغیر تصادفی X برابر تعداد سئوالات درس پاسخ داده شده در بین ۲۰ سئوال باشد آنگاه $X\sim B(۲۰,۰/۲)$ و بنابراین

الف– $\mu=E(X)=np=۲۰(۰/۲)=۴$

پس انتظار داریم ۴ سئوال را درست پاسخ دهد.

ب– $P(X\geq ۱۰)=۱-P(X<۱۰)=۱-P(X\leq ۹)=۱-۰/۹۹۷۴=۰/۰۰۲۶$

ج– $P(۲<X<۷)=P(۲<X\leq ۶)=P(X\leq ۶)-P(X\leq ۲)=۰/۹۱۳۳-۰/۲۰۶۱$

$=۰/۷۰۷۲$

د– $P(X=۹)=P(X\leq ۹)-P(X\leq ۸)=۰/۹۹۷۴-۰/۹۹۰۰=۰/۰۰۷۴$

۴.۵ توزیع فوق هندسی[1]

در بخش قبل اگر آزمایشات برنولی تکرار شده، از یکدیگر مستقل نباشند آنگاه آزمایش به دست آمده دیگر آزمایش دو جمله‌ای نمی‌باشد. برای مثال در انتخاب ۵ مهره بدون جایگذاری از ظرفی که دارای ۸ مهره سفید و ۶ مهره قرمز است و موفقیت برای ما مشاهده مهره قرمز است، احتمال موفقیت در هر بار آزمایش تغییر می‌کند و دیگر آزمایش دو جمله‌ای نمی‌باشد. برای اینکه در حالت مستقل نبودن آزمایشات برنولی یا انتخاب بدون جایگذاری، نحوه انجام آزمایش و توزیع احتمال مربوطه را به دست آوریم به مثال زیر توجه کنید.

مثال ۱.۴.۵ از بین ۳ شیمی‌دان و ۴ فیزیک‌دان می‌خواهیم یک کمیته ۵ نفری را انتخاب کنیم. تابع احتمال برای تعداد شیمی‌دانهای انتخابی در کمیته را به دست آورید.

حل اگر متغیر تصادفی X برابر تعداد شیمی‌دانهای انتخابی در کمیته ۵ نفری (تعداد موفقیتها در n

1- Hyper Geometric Distribution

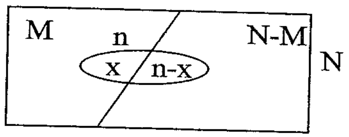
آزمایش غیرمستقل) باشد آنگاه بواسطه انتخاب بدون جایگذاری افراد، X مقادیر $S_X=\{۱,۲,۳\}$ را می‌گیرد و

$$P(X=۱)=\frac{\binom{۳}{۱}\binom{۴}{۴}}{\binom{۷}{۵}}\quad,\quad P(X=۲)=\frac{\binom{۳}{۲}\binom{۴}{۳}}{\binom{۷}{۵}}$$

$$P(X=۳)=\frac{\binom{۳}{۳}\binom{۴}{۲}}{\binom{۷}{۵}}$$

و بنابراین

$$f_X(x)=P(X=x)=\frac{\binom{۳}{x}\binom{۴}{۵-x}}{\binom{۷}{۵}}\quad,\quad x=۱,۲,۳$$



در حالت کلی جمعیت مورد نظر ما به دو قسمت، جمعیت موفقیت و جمعیت شکست تقسیم می‌شود که یکی Mتایی و دیگر N-Mتایی است که N اندازه کل جمعیت است. می‌خواهیم n عضو از این جمعیت Nتایی را بدون جایگذاری انتخاب کنیم. چنین آزمایشی را آزمایش فوق هندسی گویند.

تعریف ۲.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

X = تعداد موفقیتها در یک آزمایش فوق هندسی
= تعداد موفقیتها در nآزمایش غیر مستقل

آنگاه X را یک متغیر تصادفی فوق هندسی گوییم و آن را با نماد $X\sim HG(N,M,n)$ نشان می‌دهیم که در آن N برابر تعداد اعضای جمعیت، M برابر تعداد اعضای جمعیت موفقیت و n برابر تعداد اعضای نمونه انتخابی است.

برای به دست آوردن تابع احتمال توزیع فوق هندسی توجه کنید که P(X=x) بدین معنی است که در انتخاب n عضو از جمعیت Nتایی (به $\binom{N}{n}$ طریق) می‌خواهیم x عضو از جمعیت موفقیت Mتایی (به $\binom{M}{x}$ طریق) و مابقی از جمعیت شکست (به $\binom{N-M}{n-x}$ طریق) باشند.

بنابراین

$$P(X=x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

حال چون در ترکیب $\binom{M}{x}$ بایستی $۰ \le x \le M$ و در ترکیب $\binom{N-M}{n-x}$ بایستی $۰ \le n-x \le N-M$ باشد پس نتیجه می‌شود که $max\,(۰,n-N+M)\le x\le min\,(n,M)$

بنابراین

تابع احتمال توزیع فوق هندسی

$$f_X(x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}\ , \quad max\,(۰,n-N+M)\le x\le min\,(n,M) \qquad (۳.۵)$$

برای مثال در مثال (۱.۴.۵) داریم که $X\sim HG(۷,۳,۵)$

$$f_X(x)=\frac{\binom{۳}{x}\binom{۷-۳}{۵-x}}{\binom{۷}{۵}}\ , \quad max\,(۰,۵-۷+۳)\le x\le min\,(۵,۳)\ \Rightarrow\ ۱\le x\le ۳$$

مثال ۲.۴.۵ در انتخاب ۵ قطعه از بین ۴۰ قطعه که ۳ تای آنها خراب است احتمال این را پیدا کنید که حداکثر یک قطعه انتخاب شده خراب باشد.

حل اگر متغیر تصادفی X برابر تعداد قطعات خراب انتخاب شده در بین ۵ قطعه انتخابی باشد آنگاه $X\sim HG(۴۰,۳,۵)$ و $S_X=\{۰,۱,۲,۳\}$. بنابراین

$$P(X\le ۱)=f_X(۰)+f_X(۱)=\frac{\binom{۳}{۰}\binom{۳۷}{۵}}{\binom{۴۰}{۵}}+\frac{\binom{۳}{۱}\binom{۳۷}{۴}}{\binom{۴۰}{۵}}=۰/۹۶۳۵$$

در قضیه زیر امید ریاضی و واریانس توزیع فوق هندسی را بدون اثبات می‌آوریم.

قضیه ۳.۵ اگر $X\sim HG(N,M,n)$ آنگاه

$$E(X)=\frac{nM}{N}\ , \quad Var\,(X)=\frac{nM}{N}(۱-\frac{M}{N})(\frac{N-n}{N-۱})$$

۱.۴.۵ تقریب توزیع فوق هندسی بوسیله توزیع دوجمله‌ای

اگر در آزمایش فوق هندسی n نسبت به N عدد کوچکی باشد آنگاه انتخاب اعضاء در مراحل مختلف دارای احتمالات تقریباً یکسان هستند یعنی انتخاب اعضاء را می‌توان تقریباً با جایگذاری در نظر گرفت و در نتیجه ما یک آزمایش دو جمله‌ای خواهیم داشت. در این حالت می‌توان توزیع فوق هندسی را با توزیع دوجمله‌ای با پارامترهای n و $p=\frac{M}{N}$ تقریب زد. در حقیقت

اگر زمانی که M و N به سمت بی نهایت میل کنند، مقدار $\frac{M}{N}=p$ ثابت باشد آنگاه می‌توان نشان داد که

$$\lim_{N\to\infty}\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}=\binom{n}{x}p^x(۱-p)^{n-x}$$

مثال ۳.۴.۵ در یک منطقه ۴۰۰۰ از ۱۰۰۰۰ رای دهنده با مالیات فروش مخالف می‌باشند. اگر ۱۵ نفر از افرادی که می‌توانند رای دهند به طور تصادفی انتخاب شوند و از آنها در مورد عقیده‌شان سئوال شود، احتمال اینکه حداکثر ۷ نفر از آنها موافق مالیات فروش باشند را بیابید.

حل اگر متغیر تصادفی X برابر تعداد رای دهندگان در بین ۱۵ نفر باشد که موافق مالیات فروش هستند، آنگاه $X\sim HG(۱۰۰۰۰,۶۰۰۰,۱۵)$ و چون n=۱۵ نسبت به N=۱۰۰۰۰ عدد کوچکی است پس تقریباً $X\sim B\,(۱۵,۰/۶)$ که در آن $p=\frac{M}{N}=\frac{۶۰۰۰}{۱۰۰۰۰}=۰/۶$. بنابراین با استفاده از جدول (I) داریم که

$$P(X\le ۷)\approx\sum_{x=۰}^{۷}\binom{۱۵}{x}(۰/۶)^x(۰/۴)^{۱۵-x}=۰/۲۱۳۱$$

۵.۵ توزیع پواسون[1]

آزمایشی که تعداد موفقیتها در یک فاصله زمانی یا ناحیه مشخص را به دست دهد یک آزمایش پواسون نامیده می‌شود. این فاصله زمانی می‌تواند هر فاصله زمانی مانند ثانیه، دقیقه و... باشد و ناحیه مشخص نیز می‌تواند یک فاصله خطی یا مساحت یا حجم یا... باشد.

تعریف ۳.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

X = تعداد موفقیتها در یک فاصله زمانی یا ناحیه مشخص

آنگاه X را یک متغیر تصادفی پواسون گویند.

برای مثال اگر X تعداد تلفن هایی باشد که در یک ساعت به رییس یک شرکت زده می‌شود و یا X تعداد غلطهای تایپی در هر صفحه یک کتاب باشد و یا X تعداد تاکسی‌هایی باشد که در یک ساعت معین از یک چهار راه عبور می‌کنند، آنگاه X یک متغیر تصادفی پواسون است.

یک آزمایش پواسون بایستی دارای خواص زیر باشد

1- Poisson Distribution

۱ـ تعداد موفقیتهایی که در یک فاصله زمانی یا ناحیه مشخص اتفاق می‌افتد از تعداد موفقیتهایی که در یک فاصله زمانی دیگر یا ناحیه دیگر اتفاق می‌افتد مستقل باشد.

۲ـ احتمال اینکه یک موفقیت در یک فاصله زمانی کوتاه یا ناحیه کوچک روی دهد متناسب با طول فاصله زمانی یا ناحیه مشخص باشد و بستگی به تعداد موفقیتها در خارج از فاصله زمانی یا ناحیه نداشته باشد.

۳ـ احتمال روی دادن بیش از یک موفقیت در یک فاصله زمانی کوتاه یا ناحیه مشخص کوچک، قابل صرفنظر باشد.

توزیع احتمال یک متغیر تصادفی پواسون به پارامتر μ یعنی میانگین تعداد موفقیتها در فاصله زمانی یا ناحیه مورد نظر بستگی دارد. اگر X دارای توزیع پواسون با میانگین μ باشد آن را با نماد $X\sim P(\mu)$ نمایش می‌دهیم. در تعریف و قضیه زیر تابع احتمال و امید ریاضی و واریانس یک متغیر تصادفی پواسون را بدون اثبات می‌آوریم.

تعریف ۴.۵ اگر متغیر تصادفی X برابر تعداد موفقیتها در یک فاصله زمانی یا ناحیه مشخص باشد آنگاه $X\sim P(\mu)$ و

تابع احتمال توزیع پواسون

$$f_X(x)=\frac{e^{-\mu}\mu^x}{x!} \quad , \quad x=0,1,2,... \qquad (4.5)$$

قضیه ۴.۵ اگر $X\sim P(\mu)$ آنگاه $E(X)=\mu$ و $Var(X)=\mu$.

مثال ۱.۵.۵ یک دستگاه چاپگر کامپیوتری به طور متوسط در هر ماه ۲ بار سرویس می‌شود.

الف – احتمال اینکه در یک ماه کمتر از ۲ بار سرویس شود را بیابید.

ب – احتمال اینکه در سه ماه این چاپگر حداقل ۲ بار سرویس شود را بیابید.

حل الف – اگر متغیر تصادفی X برابر تعداد دفعاتی باشد که چاپگر در یک ماه سرویس می‌شود آنگاه $X\sim P(2)$ و بنابراین

$$P(X<2)=f_X(0)+f_X(1)=\frac{e^{-2}2^0}{0!}+\frac{e^{-2}2^1}{1!}=0/406$$

ب – اگر متغیر تصادفی Y برابر تعداد دفعاتی باشد که چاپگر در سه ماه سرویس می‌شود آنگاه طبق خاصیت دوم آزمایش پواسون $Y\sim P(6)$ و بنابراین

$$P(Y\geq 2)=1-P(Y<2)=1-\{f_Y(0)+f_Y(1)\}=1-\left\{\frac{e^{-6}6^0}{0!}+\frac{e^{-6}6^1}{1!}\right\}=0/9826$$

برای محاسبه احتمالات در توزیع پواسون جدول (II) در ضمیمه ارایه گردیده است که در این جدول $P(X\leq r)$ برای مقادیر $\mu = 0/1, 0/2, ..., 1, 1/5, 2, ..., 18$ محاسبه شده است. برای مثال اگر $X\sim P(5)$ آنگاه از این جدول به دست می‌آوریم که $P(X\leq 3)=0/2650$

مثال ۲.۵.۵ به طور متوسط در هر ده دقیقه ۶ مشتری به پای صندوق پرداخت یک فروشگاه می‌رسند.

الف – احتمال اینکه در ده دقیقه حداکثر ۴ مشتری به پای صندوق برسند را بیابید.

ب – احتمال اینکه در پنج دقیقه حداقل ۲ مشتری به پای صندوق برسند را بیابید.

حل الف – اگر X = تعداد مشتریانی که در ده دقیقه به پای صندوق می‌رسند

آنگاه $X\sim P(6)$ و $P(X\leq 4)=0/2851$

ب – اگر Y = تعداد مشتریانی که در پنج دقیقه به پای صندوق می‌رسند

آنگاه $Y\sim P(3)$ و $P(Y\geq 2)=1-P(Y<2)=1-P(Y\leq 1)=1-0/1991=0/8009$

۱.۵.۵ تقریب توزیع دو جمله‌ای بوسیله توزیع پواسون

اگر در آزمایش دو جمله‌ای n عدد بزرگی باشد و p به صفر نزدیک باشد آنگاه می‌توان توزیع دو جمله‌ای را با توزیع پواسون با پارامتر $\mu=np$ تقریب زد. در حقیقت زمانی که $n\to\infty$ و $p\to 0$، مقدار $\mu=np$ مقداری ثابت باشد آنگاه می‌توان نشان داد که

$$\lim_{n\to\infty}\binom{n}{x}p^x(1-p)^{n-x}=\frac{e^{-\mu}\mu^x}{x!}$$

همچنین اگر p به یک نزدیک باشد می‌توان در آزمایش دو جمله‌ای مفاهیم موفقیت و شکست را با یکدیگر عوض کنیم و در نتیجه تقریب بالا را در این مورد هم بکار ببریم.

مثال ۳.۵.۵ اطلاعات یک شرکت بیمه نشان می‌دهد که ۰/۰۰۱ جمعیت در هر سال از نوع معینی تصادف می‌میرند. احتمال اینکه این شرکت مجبور باشد برای بیشتر از ۱۵ نفر از ۱۰۰۰۰ بیمه گزار شرکتش در مقابل خطرات چنین تصادفهایی در سال مفروض غرامت بپردازد را بیابید.

حل اگر متغیر تصادفی X برابر تعداد بیمه گزاران در بین ۱۰۰۰۰ نفر باشد که در سال از تصادفی معین می‌میرند آنگاه $X\sim B(10000, 0/001)$ است که چون n عددی بزرگ و p به صفر نزدیک

است پس تقریباً $X\sim P(۱۰)$ که در آن $\mu = np = ۱۰۰۰۰\,(۰/۰۰۱) = ۱۰$. بنابراین

$$P(X > ۱۵) = ۱ - P(X\leq ۱۵) \approx ۱ - ۰/۹۵۱۳ = ۰/۰۴۸۷$$

۶.۵ توزیع دو جمله‌ای منفی[1]

آزمایشی را در نظر بگیرید که دارای خواص آزمایش دو جمله‌ای باشد با این تفاوت که آزمایشات مستقل برنولی را آنقدر تکرار می‌کنیم تا به یک تعداد معینی از موفقیتها دست یابیم. چنین آزمایشی را آزمایش دوجمله‌ای منفی گویند. برای اینکه در این حالت نحوه انجام آزمایش و توزیع احتمال مربوطه را به دست آوریم به مثال زیر توجه کنید.

مثال ۱.۶.۵ یک تیرانداز ۷۰٪ از تیرهای خود را به هدف می‌زند. احتمال اینکه در ششمین پرتاب چهارمین تیر او به هدف بخورد را بیابید.

حل اگر متغیر تصادفی X برابر تعداد تیرهای پرتاب شده تا رسیدن به چهارمین برخورد به هدف تیرها باشد آنگاه احتمال مورد نظر برابر $P(X=۶)$ است. احتمال $P(X=۶)$ بدین معنی است که در پرتاب ششم تیر به هدف برخورد کند (یعنی موفقیت یا S داشته باشیم) و در ۵ پرتاب اولیه ۳ تیر به هدف برخورد کرده و ۲ تیر به هدف برخورد نکند (یعنی شکست یا F داشته باشیم). یک حالت خاص برای وقوع چنین پیشامدی SFSSFS می‌باشد که احتمال آن برابر است با

$$(۰/۷)(۰/۳)(۰/۷)(۰/۷)(۰/۳)(۰/۷) = (۰/۷)^{۴}(۰/۳)^{۲}$$

و تعداد حالاتی که می‌توان در ۵ پرتاب اولیه ۳ برخورد به هدف داشته باشیم برابر $\binom{۵}{۳}$ است و

بنابراین $$P(X=۶) = \binom{۵}{۳}(۰/۷)^{۴}(۰/۳)^{۲} = \binom{۶-۱}{۴-۱}(۰/۷)^{۴}(۰/۳)^{۶-۴}$$

تعریف ۵.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

$$\boxed{X = \text{تعداد آزمایشات مستقل برنولی تا رسیدن به r موفقیت}}$$

آنگاه X را یک متغیر تصادفی دو جمله‌ای منفی گویند. اگر احتمال موفقیت در هر آزمایش برنولی p باشد آنگاه گوئیم X دارای توزیع دو جمله‌ای منفی با پارامترهای r و p است و آن را نماد

1- Negative Binomial Distribution

$X\sim NB\,(r,p)$ نمایش می‌دهیم.

برای به دست آوردن تابع احتمال توزیع دو جمله‌ای منفی توجه کنید که $P(X=x)$ بدین معنی است که در xامین آزمایش ما یک موفقیت داشته باشیم و در $(x-۱)$ آزمایش اولیه ما $(r-۱)$ موفقیت و $(x-r)$ شکست داشته باشیم که احتمال آن برابر است با $\binom{x-۱}{r-۱}p^{r-۱}q^{x-r}p$. بنابراین

تابع احتمال توزیع دو جمله‌ای منفی

$$f_X(x) = \binom{x-۱}{r-۱}p^{r}q^{x-r}\quad,\quad x = r, r+۱, r+۲, \ldots \qquad (۵.۵)$$

قضیه ۵.۵ اگر $X\sim NB\,(r,p)$ آنگاه $E(X)=\frac{r}{p}$ و $Var\,(X)=\frac{rq}{p^{۲}}$.

مثال ۲.۶.۵ فرض کنید که ۴۰٪ از ماهیهای یک دریاچه از نوع بخصوصی باشند. اگر هر بار یک ماهی گرفته و نوع آن را مشخص کرده و دوباره به دریاچه برگردانیم.

الف- انتظار دارید که در چندمین صید ماهی، چهارمین ماهی از نوع فوق مشاهده شود؟

ب- احتمال اینکه در دهمین بار، چهارمین ماهی از نوع فوق مشاهده شود را بیابید.

حل اگر متغیر تصادفی X برابر تعداد صید ماهی تا رسیدن به چهارمین ماهی نوع بخصوص باشد آنگاه $X\sim NB\,(۴, ۰/۴)$ و در نتیجه

الف- $$E(X) = \frac{r}{p} = \frac{۴}{۰/۴} = ۱۰$$

پس انتظار داریم در دهمین صید، چهارمین ماهی از نوع فوق صید شود.

ب- $$P(X=۱۰) = f_X(۱۰) = \binom{۱۰-۱}{۴-۱}(۰/۴)^{۴}(۰/۶)^{۱۰-۴} = ۰/۱۰۰۳$$

۷.۵ توزیع هندسی[1]

اگر در توزیع دو جمله‌ای منفی $r=۱$ باشد آنگاه یک توزیع احتمال برای تعداد آزمایشات تا رسیدن به یک موفقیت را به دست می‌آوریم. این نوع آزمایش را آزمایش هندسی گویند و توزیع مربوطه را توزیع هندسی گویند. برای مثال پرتاب یک سکه تا رسیدن به یک شیر یک آزمایش هندسی است.

تعریف ۶.۵ اگر متغیر تصادفی X به صورت زیر تعریف شود

1- Geometric Distribution

X = تعداد آزمایشات مستقل برنولی تا رسیدن به یک موفقیت

آنگاه X را یک متغیر تصادفی هندسی گویند. اگر احتمال موفقیت در هر آزمایش برنولی p باشد آنگاه گوئیم X دارای توزیع هندسی با پارامتر p است و آن را با نماد $X\sim G(p)$ نمایش می‌دهیم و تابع احتمال آن عبارت است از

تابع احتمال توزیع هندسی $f_X(x)=pq^{x-1} \quad , \quad x=1,2,3,...$ (۶.۵)

مثال ۱.۷.۵ فرض کنید احتمال قبولی یک نفر در امتحان رانندگی ۰/۷ باشد. احتمال اینکه این شخص در امتحان رانندگی (الف) در مرتبه سوم، (ب) حداکثر در سومین بار قبول شود را بیابید.

حل اگر متغیر تصادفی X برابر تعداد امتحانهای رانندگی شخص تا قبول شدن باشد آنگاه $X\sim G(0/7)$. بنابراین

الف- $P(X=3)=f_X(3)=(0/7)(0/3)^{3-1}=0/063$

ب- $P(X\le 3)=1-P(X>3)=1-\sum_{x=4}^{+\infty}(0/7)(0/3)^{x-1}=1-\frac{(0/7)(0/3)^3}{1-0/3}$
$=1-(0/3)^3=0/973$

قضیه ۶.۵ اگر $X\sim G(p)$ آنگاه

$$E(X)=\frac{1}{p} \quad , \quad Var(X)=\frac{q}{p^2} \quad , \quad P(X\le r)=1-q^r$$

۸.۵ توزیع یکنواخت گسسته[1]

ساده‌ترین توزیع احتمال گسسته توزیعی است که در آن متغیر تصادفی گسسته X تمام مقادیرش را با احتمالات یکسان اختیار کند. چنین توزیع احتمالی را توزیع یکنواخت گسسته گویند.

تعریف ۷.۵ اگر متغیر تصادفی X مقادیر x_1، x_2، ... و x_K را با احتمالات یکسان $\frac{1}{K}$ اختیار کند، آنگاه گوئیم X دارای توزیع یکنواخت گسسته با پارامتر K است و آن را با نماد $X\sim DU(K)$ نمایش می‌دهیم و تابع احتمال آن عبارت است از

1- Discrete Uniform Distribution

تابع احتمال توزیع یکنواخت گسسته
$f_X(x)=P(X=x)=\frac{1}{K} \quad , \quad x=x_1,x_2,...,x_K$ (۷.۵)

قضیه ۷.۵ اگر $X\sim DU(K)$ آنگاه

$$\mu=E(X)=\frac{1}{K}\sum_{i=1}^{K}x_i$$
$$Var(X)=\frac{1}{K}\sum_{i=1}^{K}(x_i-\mu)^2$$

اثبات چون X مقادیر x_1، x_2، ... و x_K را با احتمالات یکسان $\frac{1}{K}$ اختیار می‌کند بنابراین از تعریف امید ریاضی و واریانس نتیجه به راحتی به دست می‌آید.

مثال ۱.۸.۵ یک صفحه هدف زنی دایره‌ای شکل به ۱۵ قطاع مساوی تقسیم شده و با شماره‌های ۱ تا ۱۵ متمایز گردیده است. اگر X برابر عددی باشد که تیر در قطاع مربوط به آن اصابت می‌کند، توزیع احتمال X را به دست آورده و میانگین و واریانس X را محاسبه کنید. احتمال اینکه تیر در قطاع با شماره کمتر از ۱۰ برخورد کند را بیابید.

حل $X\sim DU(15) \quad , \quad f_X(x)=\frac{1}{15} \quad x=1,2,...,15$

$$\mu=E(X)=\frac{1}{15}\sum_{x=1}^{15}x=\frac{15(15+1)}{2\times 15}=8 \quad , \quad \sigma^2=\frac{1}{15}\sum_{x=1}^{15}(x-8)^2=\frac{56}{3}$$

$$P(X<10)=\sum_{x=1}^{9}f_X(x)=\frac{9}{15}$$

۹.۵ توزیع یکنواخت پیوسته[1]

در بخشهای قبل توزیع احتمال چند متغیر تصادفی گسسته خاص را مورد بررسی قرار دادیم. از این بخش به بعد چند توزیع احتمال پیوسته خاص را مورد بررسی قرار خواهیم داد.

ساده‌ترین توزیع احتمال پیوسته توزیع یکنواخت (پیوسته) است که در زیر آن را تعریف می‌کنیم.

تعریف ۸.۵ گوئیم متغیر تصادفی X دارای توزیع یکنواخت در فاصله (a,b) است و آن را با نماد

1- Continuous Uniform Distribution

$X\sim U(a,b)$ نمایش می‌دهیم. هرگاه دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\begin{cases}\dfrac{۱}{b-a} & a<x<b\\ ۰ & \text{سایر نقاط}\end{cases}\qquad \text{تابع چگالی احتمال توزیع یکنواخت}\qquad (۸.۵)$$

تابع توزیع متغیر تصادفی یکنواخت بصورت زیر به دست می‌آید.

$$F_X(x)=\begin{cases}۰ & x<a\\ \dfrac{x-a}{b-a} & a\le x<b\\ ۱ & b\le x\end{cases}$$

قضیه ۸.۵ اگر $X\sim U(a,b)$ آنگاه $E(X)=\frac{a+b}{۲}$ و $Var(X)=\frac{(b-a)^۲}{۱۲}$.

اثبات

$$E(X)=\int_a^b \frac{x}{b-a}dx=\frac{x^۲}{۲(b-a)}\Big|_a^b=\frac{a+b}{۲}$$

$$E(X^۲)=\int_a^b \frac{x^۲}{b-a}dx=\frac{x^۳}{۳(b-a)}\Big|_a^b=\frac{a^۲+ab+b^۲}{۳}$$

$$Var(X)=\frac{a^۲+ab+b^۲}{۳}-\left(\frac{a+b}{۲}\right)^۲=\frac{(a-b)^۲}{۱۲}$$

مثال ۱.۹.۵ فرض کنید B عددی تصادفی از فاصله $[-۳,۳]$ باشد. احتمال اینکه معادله درجه دوم $x^۲+Bx+۱=۰$ حداقل یک ریشه حقیقی داشته باشد را بیابید؟

حل $B\sim U(-۳,۳)$ و تابع احتمال آن عبارت است از $f_B(b)=\frac{۱}{۶},\ -۳<b<۳$

برای اینکه معادله فوق حداقل یک ریشه حقیقی داشته باشد بایستی $\Delta=B^۲-۴\ge۰$. بنابراین

$$P(B^۲-۴\ge۰)=P(B^۲\ge۴)=P(|B|\ge۲)=۱-P(-۲<B<۲)=۱-\int_{-۲}^{۲}\frac{۱}{۶}dx=۱-\frac{۴}{۶}=\frac{۱}{۳}$$

۱۰.۵ توزیع نمایی[1]

توزیع نمایی یکی از توزیعهای مهم آماری است که در زیر آن را معرفی می‌کنیم.

1- Exponential Distribution

تعریف ۹.۵ گوئیم متغیر تصادفی X دارای توزیع نمایی با پارامتر θ است و آن را با نماد $X\sim E(\theta)$ نمایش می‌دهیم، هرگاه تابع چگالی احتمال آن به صورت زیر باشد

$$\text{تابع چگالی احتمال توزیع نمایی}\qquad f_X(x)=\begin{cases}\dfrac{۱}{\theta}e^{-x/\theta} & x>۰\ ,\ \theta>۰\\ ۰ & \text{سایر نقاط}\end{cases}\qquad (۹.۵)$$

معمولاً اگر متغیر تصادفی X نمایانگر طول عمر یک قطعه باشد آنگاه می‌توان X را متغیر تصادفی نمایی در نظر گرفت، توزیع نمایی در آمار کاربرد فراوان دارد. از جمله کاربردهای آن در نظریه اعتماد، نظریه صف و زمان انتظار می‌باشد. در قضیه زیر امید و واریانس این توزیع را به دست می‌آوریم.

قضیه ۹.۵ اگر $X\sim E(\theta)$ آنگاه $E(X)=\theta$ و $Var(X)=\theta^۲$.

اثبات

$$E(X)=\int_۰^{+\infty}x\left(\frac{۱}{\theta}e^{-x/\theta}\right)dx=\left[(-x-\theta)e^{-x/\theta}\right]_۰^{+\infty}=\theta$$

$$E(X^۲)=\int_۰^{+\infty}x^۲\left(\frac{۱}{\theta}e^{-x/\theta}\right)dx=\left[(-x^۲-۲x\theta-۲\theta^۲)e^{-x/\theta}\right]_۰^{+\infty}=۲\theta^۲$$

$$Var(X)=۲\theta^۲-\theta^۲=\theta^۲$$

مثال ۱.۱۰.۵ فرض کنید متغیر تصادفی X نشانگر طول عمر نوعی لاستیک برحسب ساعت باشد که دارای توزیع نمایی با میانگین ۵۰۰ ساعت است. اگر بدانیم که طول عمر یک لاستیک از این نوع از ۷۰۰ ساعت کمتر بوده است، احتمال اینکه طول عمر آن از ۳۰۰ ساعت بیشتر باشد را بیابید.

حل می‌دانیم که $X\sim E(۵۰۰)$ بنابراین

$$P(X>۳۰۰\mid X<۷۰۰)=\frac{P(X>۳۰۰,X<۷۰۰)}{P(X<۷۰۰)}=\frac{P(۳۰۰<X<۷۰۰)}{P(X<۷۰۰)}$$

$$=\frac{\int_{۳۰۰}^{۷۰۰}\frac{۱}{۵۰۰}e^{-\frac{x}{۵۰۰}}dx}{\int_{۰}^{۷۰۰}\frac{۱}{۵۰۰}e^{-\frac{x}{۵۰۰}}dx}=\frac{۰/۳۰۲۲}{۰/۷۵۳۴}=۰/۴۰۱۱$$

مثال ۲.۱۰.۵ طول عمر هر دستگاه کامپیوتر دارای توزیع نمایی با میانگین ۱۷۰۰ ساعت می‌باشد. اگر آزمایشگاهی ۲۰ دستگاه کامپیوتر داشته باشد، احتمال اینکه حداقل دو دستگاه از آنها

قبل از ۱۷۰۰ ساعت خراب شوند را بیابید.

حل قرار می‌دهیم طول عمر یک دستگاه کامپیوتر برحسب ساعت = X

تعداد دستگاههای کامپیوتر در بین ۲۰ دستگاه که دارای طول عمر کمتر از ۱۷۰۰ ساعت هستند = Y

در این صورت $X\sim E(۱۷۰۰)$ و $Y\sim B(۲۰,p)$ که در آن

$$p=P(X<۱۷۰۰)=\int_{۰}^{۱۷۰۰}\frac{۱}{۱۷۰۰}e^{-\frac{x}{۱۷۰۰}}dx=۱-e^{-۱}=۰/۶۳۲۱$$

بنابراین احتمال مورد نظر برابر است با

$$P(Y\geq ۲)=۱-P(Y<۲)=۱-\{f_Y(۰)+f_Y(۱)\}$$

$$=۱-\left\{\binom{۲۰}{۰}(۰/۶۳۲۱)^{۰}(۰/۳۶۷۹)^{۲۰}+\binom{۲۰}{۱}(۰/۶۳۲۱)^{۱}(۰/۳۶۷۹)^{۱۹}\right\}=۰/۹۹۹۹$$

۱.۱۰.۵ رابطه توزیع نمایی و توزیع پواسون

در بخش ۵ دیدیم که یک آزمایش پواسون یک مدل مناسب برای توزیع تعداد اتفاقات (موفقیتها) در یک زمان معین می‌باشد. در این بخش می‌خواهیم نشان دهیم که در یک آزمایش پواسون توزیع زمان طی شده تا وقوع اولین اتفاق و یا توزیع زمان طی شده بین دو اتفاق متوالی از یک توزیع نمایی پیروی می‌کند.

در یک آزمایش پواسون با پارامتر μ، که μ میانگین تعداد اتفاقات در یک واحد زمانی است، قرار می‌دهیم:

تعداد اتفاقات درفاصله زمانی [۰,t] = X

زمان تا رسیدن به اولین اتفاق = Y

در این صورت طبق خاصیت دوم آزمایش پواسون $X\sim P(\mu t)$ و همچنین

$$P(Y>t)=P(X=۰)=e^{-\mu t}$$

بنابراین

$$F_Y(t)=P(Y\leq t)=۱-e^{-\mu t}$$

و در نتیجه

$$f_Y(t)=F'_Y(t)=\mu e^{-\mu t}\quad t>۰$$

پس $Y\sim E(\frac{۱}{\mu})$. یعنی در یک آزمایش پواسون با میانگین μ زمان تا رسیدن به اولین اتفاق دارای توزیع نمایی با پارامتر $\frac{۱}{\mu}$ است. این مطلب را می‌توان برای زمان بین دو اتفاق متوالی به طور مشابه اثبات کرد.

مثال ۳.۱۰.۵ به طور متوسط تعداد ۵ تلفن در یک ساعت به تلفنخانه یک شرکت زده می‌شود.

الف- احتمال اینکه در یک ساعت حداقل ۲ تلفن زده شود را بیابید.

ب- احتمال اینکه تلفن بعدی لااقل بعد از ۱۵ دقیقه زده شود را بیابید.

ج- احتمال اینکه تلفن بعدی قبل از ۱۰ دقیقه زده شود را بیابید.

حل الف- اگر متغیر تصادفی X برابر تعداد تلفنهایی باشد که در یک ساعت به شرکت زده می‌شود، آنگاه $X\sim P(۵)$ و بنابراین

$$P(X\geq ۲)=۱-P(X\leq ۱)=۱-۰/۰۴۰۴=۰/۹۵۹۶$$

ب- اگر متغیر تصادفی Y برابر زمان بین دو تلفن متوالی برحسب ساعت باشد آنگاه $Y\sim E(\frac{۱}{۵})$ و بنابراین احتمال مورد نظر برابر است با

$$P(Y>\frac{۱}{۴})=\int_{\frac{۱}{۴}}^{+\infty}۵e^{-۵y}dy=e^{-\frac{۵}{۴}}=۰/۲۸۶۵$$

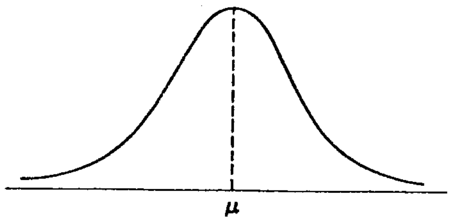
ج-

$$P(Y<\frac{۱}{۶})=\int_{۰}^{\frac{۱}{۶}}۵e^{-۵y}dy=۱-e^{-\frac{۵}{۶}}=۰/۵۶۵۴$$

۱۱.۵ توزیع نرمال[1]

مهمترین توزیع پیوسته در آمار و احتمال توزیع نرمال است. نمودار این توزیع در شکل (۱.۵) رسم شده است و کاملاً نسبت به یک حد متوسط μ متقارن است و به آن منحنی نرمال گوئیم.

اغلب متغیرهای تصادفی پیوسته در طبیعت و صنعت دارای نمودار توزیع به‌فرم فوق می‌باشند. نمودار منحنی نرمال بستگی به دو پارامتر μ و σ^2 دارد و در زیر خواهیم دید که μ میانگین توزیع و σ^2 واریانس توزیع می‌باشد.



شکل ۱.۵ منحنی نرمال

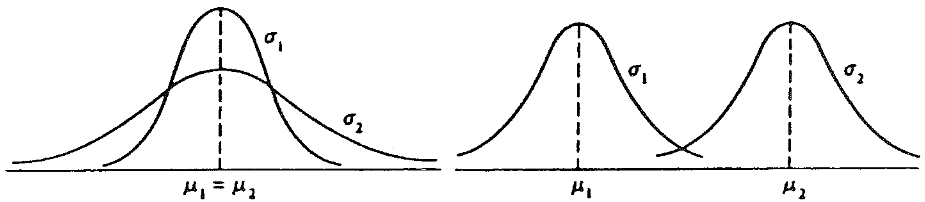
تعریف ۱۰.۵ گوئیم متغیر تصادفی X دارای توزیع نرمال با میانگین μ و واریانس σ^2 است و آن را با نماد $X\sim N(\mu,\sigma^2)$ نمایش می‌دهیم، هرگاه تابع چگالی احتمال آن به صورت زیر باشد

1- Normal Distribution

تابع چگالی احتمال توزیع نرمال

$$f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\quad -\infty<x<+\infty,\ \mu\in R\ ,\ \sigma>0 \qquad (10.5)$$

منحنی نرمال کاملاً بوسیله مقادیر μ و σ^2 مشخص می‌شود. با افزایش σ^2 پراکندگی توزیع افزایش می‌یابد و با افزایش μ منحنی به سمت راست انتقال پیدا می‌کند. در شکل ۲.۵ منحنی نرمال در حالتهای مختلف μ و σ^2 رسم شده است.



شکل ۲.۵ – ب منحنی‌های نرمال با $\mu_1=\mu_2\ ,\ \sigma_1<\sigma_2$

شکل ۲.۵ – الف منحنی‌های نرمال با $\mu_1<\mu_2\ ,\ \sigma_1=\sigma_2$

خواص منحنی نرمال با استفاده از فرم تابع چگالی نرمال و همچنین نمودار آن خواص زیر را در مورد منحنی نرمال می‌توان بیان و اثبات کرد (اثبات به عنوان تمرین)

۱ـ منحنی تنها دارای یک نقطه ماکزیمم در نقطه $x=\mu$ است.

۲ـ منحنی دارای دو نقطه عطف در نقاط $x=\mu\pm\sigma$ است.

۳ـ منحنی نسبت به خط $x=\mu$ متقارن است، یعنی $f_X(\mu-a)=f_X(\mu+a)$

۴ـ در دو طرف حد متوسط μ منحنی به مجانب خود یعنی محور xها نزدیک می‌شود، یعنی

$\lim_{x\to\pm\infty} f_X(x)=0$

۵ـ سطح محصور بین منحنی و محور طولها برابر یک واحد است، یعنی $\int_{-\infty}^{+\infty} f_X(x)dx=1$

۱.۱۱.۵ توزیع نرمال استاندارد

توزیع نرمالی که میانگین آن صفر و واریانس آن یک باشد را توزیع نرمال استاندارد می‌نامند و آن را با نماد $Z\sim N(0,1)$ نمایش می‌دهند. تابع چگالی احتمال توزیع نرمال استاندارد را با $\phi(z)$ و تابع توزیع آن را با $\Phi(z)$ نمایش می‌دهند بنابراین

$$\phi(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\quad -\infty<z<+\infty\quad ,\quad \Phi(z)=P(Z\le z)=\int_{-\infty}^{z}\phi(t)dt$$

در زیر خواص این توزیع و رابطه آن با توزیعهای نرمال غیراستاندارد را بررسی می‌کنیم.

قضیه ۱۰.۵ اگر $Z\sim N(0,1)$ آنگاه $E(Z)=0$ و $Var(Z)=1$.

اثبات

$$E(Z)=\int_{-\infty}^{+\infty}\frac{z}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz=-\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\Big|_{-\infty}^{+\infty}=0-0=0$$

با استفاده از روش جزء به جزء داریم که

$$E(Z^2)=\int_{-\infty}^{+\infty}\frac{z^2}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz=\frac{1}{\sqrt{2\pi}}\left[-ze^{-\frac{1}{2}z^2}\right]_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz$$

$$=0+1=1$$

بنابراین $Var(Z)=1-(0)^2=1$.

قضیه ۱۱.۵ اگر $X\sim N(\mu,\sigma^2)$ آنگاه $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$.

اثبات

$$F_Z(z)=P(Z\le z)=P(\frac{X-\mu}{\sigma}\le z)=P(X\le\mu+\sigma z)=F_X(\mu+\sigma z)$$

بنابراین

$$f_Z(z)=\frac{d}{dz}F_Z(z)=\frac{d}{dz}F_X(\mu+\sigma z)=\sigma F_X(\mu+\sigma z)$$

یعنی

$$f_Z(z)=\frac{\sigma}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(\mu+\sigma z-\mu)^2}=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}=\phi(z)$$

و در نتیجه $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$.

قضیه ۱۲.۵ اگر $X\sim N(\mu,\sigma^2)$ آنگاه $E(X)=\mu$ و $Var(X)=\sigma^2$.

اثبات طبق قضیه ۱۱.۵ داریم که $X=\mu+\sigma Z$ که $Z\sim N(0,1)$، بنابراین طبق خواص امید ریاضی و واریانس داریم که

$$E(X)=E(\mu+\sigma Z)=\mu+\sigma E(Z)=\mu$$

$$Var(X)=Var(\mu+\sigma Z)=\sigma^2 Var(Z)=\sigma^2$$

۲.۱۱.۵ سطح زیر منحنی نرمال

همانگونه که در فصل سوم مشاهده کردیم در توزیعهای پیوسته احتمال قرار گرفتن متغیر تصادفی پیوسته X در یک فاصله (a,b) برابر سطح زیر منحنی تابع چگالی احتمال $f_X(x)$ از a تا

b است که به صورت زیر محاسبه می‌شود

$$P(a<X<b)=\int_a^b f_X(x)\,dx$$

در توزیع نرمال به واسطه پیچیده بودن فرم تابع چگالی احتمال، نمی‌توان به روشهای معمول این انتگرال را محاسبه کرد. به این منظور در توزیع نرمال استاندارد این نوع انتگرالها بوسیله روشهای محاسبات عددی محاسبه و در جدولی ارایه می‌گردد و سپس با استفاده از قضیه ۱۱.۵ این مساحتها در هر توزیع نرمالی محاسبه می‌گردد.

برای محاسبه احتمالات در توزیع نرمال استاندارد جدول (III) در ضمیمه ارایه گردیده است که در این جدول $\Phi(z)=P(Z\le z)$ برای مقادیر $-۳/۵\le z\le ۳/۵$ محاسبه شده است. برای مثال از جدول (III) احتمالات زیر را به دست می‌آوریم.

$$P(Z\le ۲/۳۵)=\Phi(۲/۳۵)=۰/۹۹۰۶$$

$$P(Z>-۱/۴)=۱-P(Z\le -۱/۴)=۱-\Phi(-۱/۴)=۱-۰/۰۸۰۸=۰/۹۱۹۲$$

$$P(-۰/۵۵<Z\le ۱/۷)=P(Z\le ۱/۷)-P(Z\le ۰/۵۵)$$

$$=\Phi(۱/۷)-\Phi(-۰/۵۵)=۰/۹۵۵۴-۰/۲۹۱۲=۰/۶۶۴۲$$

حال اگر $X\sim N(\mu,\sigma^2)$ باشد آنگاه طبق قضیه ۱۱.۵، $Z=\frac{X-\mu}{\sigma}\sim N(۰,۱)$ و بنابراین از این رابطه می‌توان برای محاسبه احتمالات در هر توزیع نرمالی استفاده کرد. به مثالهای زیر توجه کنید.

مثال ۱.۱۱.۵ اگر $X\sim N(۱۰,۱۶)$ باشد $P(۸<X\le ۱۵)$ را محاسبه کنید.

حل در اینجا $\mu=۱۰$ و $\sigma^2=۱۶$ است. باکم کردن مقدار μ از سه طرف نامساوی درون پرانتز و تقسیم سه طرف بر عدد مثبت σ داریم که

$$P(۸<X\le ۱۵)=P(\frac{۸-۱۰}{۴}<\frac{X-\mu}{\sigma}\le\frac{۱۵-۱۰}{۴})=P(-۰/۵<Z\le ۱/۲۵)$$

$$=P(Z\le ۱/۲۵)-P(Z\le -۰/۵)=\Phi(۱/۲۵)-\Phi(-۰/۵)$$

$$=۰/۸۹۴۴-۰/۳۰۸۵=۰/۵۸۵۹$$

مثال ۲.۱۱.۵ قطر داخلی پیستونهایی که توسط یک کارخانه ساخته می‌شود، دارای توزیع نرمال با میانگین ۳۰ میلیمتر و انحراف معیار ۰/۰۲ میلیمتر است. احتمال اینکه قطر داخلی یک پیستون بیش از ۲۹/۹۸ میلیمتر باشد را بیابید.

حل اگر متغیر تصادفی X اندازه قطر داخلی پیستون برحسب میلیمتر باشد آنگاه $X\sim N(۳۰,۰/۰۲^2)$ و بنابراین

$$P(X>۲۹/۹۸)=P(\frac{X-\mu}{\sigma}>\frac{۲۹/۹۸-۳۰}{۰/۰۲})=P(Z>-۱)$$

$$=۱-P(Z\le -۱)=۱-۰/۱۵۸۷=۰/۸۴۱۳$$

مثال ۳.۱۱.۵ یک کارگاه تولید لوله، لوله هایی را تولید می‌کند که طول این لوله‌ها دارای توزیع نرمال با میانگین ۸ و انحراف معیار ۰/۵ متر می‌باشد. مطلوب است

الف- چند درصد لوله‌ها دارای طولی بین ۷/۵ تا ۹ متر هستند؟

ب- اگر در یک روز این کارگاه ۲۰۰ لوله تولید کند، چه تعداد از این لوله‌ها طولی بیش از ۹/۲ متر دارند؟

ج- ۹۷/۵٪ از لوله‌های تولید شده در این کارگاه طولشان از چه مقداری کمتر است؟

د- اگر در یک روز بدانیم که طول لوله‌های تولید شده در این کارگاه حداقل ۷/۸ متر بوده است، چند درصد لوله‌های تولید شده در این روز طولی کمتر از ۸/۴ متر دارند.

هـ- اگر ۵ لوله را یک به یک و با جایگذاری انتخاب کرده و اندازه‌گیری کنیم احتمال اینکه حداکثر یک عدد از آنها دارای طولی بیش از ۹ متر باشد را بیابید.

حل اگر متغیر تصادفی X را برابر طول لوله تولید شده برحسب متر در نظر بگیریم آنگاه $X\sim N(۸,۰/۵^2)$ و بنابراین

الف- برای محاسبه درصد بایستی احتمال مورد نظر را در ۱۰۰٪ ضرب کنیم، بنابراین

$$P(۷/۵<X<۹)=(\frac{۷/۵-۸}{۰/۵}<\frac{X-\mu}{\sigma}<\frac{۹-۸}{۰/۵})=P(-۱<Z<۲)$$

$$=\Phi(۲)-\Phi(-۱)=۰/۹۷۷۲-۰/۱۵۸۷=۰/۸۱۸۵$$

بنابراین درصد مورد نظر برابر است با $۰/۸۱۸۵(۱۰۰\%)=۸۱/۸۵\%$

ب- برای محاسبه تعداد بایستی احتمال موردنظر را در تعداد ۲۰۰ لوله ضرب کنیم، بنابراین

$$P(X>۹/۲)=P(\frac{X-\mu}{\sigma}>\frac{۹/۲-۸}{۰/۵})=P(Z>۲/۴)=۱-P(Z\le ۲/۴)$$

$$=۱-۰/۹۹۱۸=۰/۰۰۸۲$$

بنابراین تعداد موردنظر $= ۰/۰۰۸۲ \times ۲۰۰ = ۱/۶۴ \approx ۲$

پس تقریباً ۲ لوله دارای طولی بیش از ۹/۲ متر هستند.

ج- فرض کنید مقدار موردنظر a باشد بنابراین بایستی $P(X<a) = ۰/۹۷۵$ و در نتیجه

$$۰/۹۷۵ = P(X<a) = P(\frac{X-\mu}{\sigma} < \frac{a-۸}{۰/۵}) = P(Z < \frac{a-۸}{۰/۵}) = \Phi(\frac{a-۸}{۰/۵})$$

بنابراین با استفاده از جدول (III) داریم که

$$\frac{a-۸}{۰/۵} = ۱/۹۶ \Rightarrow a = ۸ + ۰/۵(۱/۹۶) = ۸/۹۸$$

د-

$$P(X<۸/۴ \mid X\ge ۷/۸) = \frac{P(۷/۸\le X<۸/۴)}{P(X\ge ۷/۸)} = \frac{P(-۰/۴\le Z<۰/۸)}{P(Z\ge -۰/۴)}$$

$$= \frac{\Phi(۰/۸)-\Phi(-۰/۴)}{۱-\Phi(-۰/۴)} = \frac{۰/۷۸۸۱-۰/۳۴۴۶}{۱-۰/۳۴۴۶} = ۰/۶۷۶۷$$

بنابراین درصد موردنظر برابر است با $۰/۶۷۶۷\ (٪۱۰۰) = ٪۶۷/۶۷$

ه- اگر متغیر تصادفی Y برابر تعداد لوله‌هایی در بین ۵ لوله باشد که دارای طولی بیش از ۹ متر باشند آنگاه $Y\sim B(۵,p)$ که در آن

$$p = P(X>۹) = P(\frac{X-\mu}{\sigma} > \frac{۹-۸}{۰/۵}) = P(Z>۲) = ۱-P(Z\le ۲) = ۱-۰/۹۷۷۲ = ۰/۰۲۲۸$$

بنابراین احتمال مورد نظر برابر است با

$$P(Y\le ۱) = f_Y(۰)+f_Y(۱) = \binom{۵}{۰}(۰/۰۲۲۸)^{۰}(۰/۹۷۷۲)^{۵} + \binom{۵}{۱}(۰/۰۲۲۸)^{۱}(۰/۹۷۷۲)^{۴} = ۰/۹۹۵$$

۳.۱۱.۵ تقریب توزیع دوجمله‌ای بوسیله توزیع نرمال

احتمالات مربوط به توزیع دوجمله‌ای را به راحتی می‌توان از تابع احتمال توزیع دوجمله‌ای و یا به وسیله جدول (I) محاسبه کرد. اما اگر n عدد بزرگی باشد، دیگر نمی‌توان از این جدول استفاده کرد. برای nهای بزرگ می‌توان توزیع دوجمله‌ای را بوسیله توزیع نرمال توسط قضیه زیر تقریب زد.

قضیه ۱۳.۵ اگر X یک متغیر تصادفی دوجمله‌ای با میانگین $\mu = np$ و واریانس $\sigma^2 = npq$ باشد آنگاه توزیع حدی $Z = \frac{X-np}{\sqrt{npq}}$ موقعی که $n\to +\infty$ یک توزیع نرمال استاندارد است.

معمولاً اگر $n\ge ۳۰$ باشد و یا مقدار p به $\frac{۱}{۲}$ نزدیک باشد، تقریب توزیع دوجمله‌ای بوسیله توزیع نرمال یک تقریب مناسب می‌باشد. همچنین چون توزیع دوجمله‌ای یک توزیع گسسته و توزیع نرمال یک توزیع پیوسته است، در موقع تقریب زدن بایستی در محاسبه احتمالات از عمل تصحیح پیوستگی به صورت زیر استفاده کنیم.

$$P(X=K) = P(K-\frac{۱}{۲}<X<K+\frac{۱}{۲}) \quad , \quad P(X\le K) = P(X<K+\frac{۱}{۲})$$

مثال ۴.۱۱.۵ اگر $X\sim B(۲۰\,,\,۰/۶)$ باشد، احتمال $P(۱۰\le X\le ۱۴)$ را به صورت دقیق و تقریبی محاسبه کنید.

حل با استفاده از جدول (I) مقدار دقیق برابر است با

$$P(۱۰\le X\le ۱۴) = P(X\le ۱۴) - P(X\le ۹) = ۰/۸۷۴۴ - ۰/۱۲۷۵ = ۰/۷۴۶۹$$

همچنین چون $\mu = np = ۲۰(۰/۶) = ۱۲$ و $\sigma^2 = npq = ۲۰(۰/۶)(۰/۴) = ۴/۸$ پس مقدار تقریبی با استفاده از عمل تصحیح پیوستگی به صورت زیر به دست می‌آید

$$P(۱۰\le X\le ۱۴) = P(۹/۵<X<۱۴/۵) = P(\frac{۹/۵-۱۲}{\sqrt{۴/۸}} < \frac{X-np}{\sqrt{npq}} < \frac{۱۴/۵-۱۲}{\sqrt{۴/۸}})$$

$$\approx P(-۱/۱۴<Z<۱/۱۴) = \Phi(۱/۱۴)-\Phi(-۱/۱۴)$$

$$= ۰/۸۷۲۹-۰/۱۲۷۱ = ۰/۷۴۵۸$$

که اختلاف این دو مقدار ۰/۰۰۱۱ می‌باشد.

مثال ۵.۱۱.۵ احتمال اینکه یک قطعه الکترونیکی درکمتر از ۱۰۰۰ ساعت استفاده مداوم از کار بیفتد، برابر ۰/۲۵ است. احتمال اینکه در بین ۲۰۰ عدد از این قطعات کمتر از ۴۵ قطعه در کمتر ۱۰۰۰ ساعت استفاده مداوم از کار بیفتند را بیابید.

حل اگر X تعداد قطعات الکترونیکی در بین ۲۰۰ قطعه باشد که در کمتر از ۱۰۰۰ ساعت استفاده مداوم از کار می‌افتند، آنگاه $X\sim B(۲۰۰\,,\,۰/۲۵)$ و بنابراین

$$P(X<۴۵) = P(X<۴۴/۵) = P(\frac{X-np}{\sqrt{npq}} < \frac{۴۴/۵-۵۰}{\sqrt{۳۷/۵}}) \approx P(Z<-۰/۹) = ۰/۱۸۴۱$$

۴.۱۱.۵ توزیع مجموع متغیرهای تصادفی نرمال

در بعضی مسایل نیاز به محاسبه توزیع مجموعی از چند متغیر تصادفی نرمال مستقل داریم که این توزیع در قضیه زیر آورده شده است.

قضیه ۱۴.۵ اگر X_1، X_2،... و X_n متغیرهای تصادفی مستقل باشند و

$$X_i \sim N(\mu_i , \sigma_i^2) \quad , \quad i = 1, 2, \dots, n$$

و قرار دهیم $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n = \sum_{i=1}^{n} a_iX_i$ در این صورت

$$\boxed{Y = \sum_{i=1}^{n} a_iX_i \sim N\left(\sum_{i=1}^{n} a_i\mu_i \, , \, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)} \tag{11.5}$$

مثال ۶.۱۱.۵ نمره یک درس در امتحان میان ترم دارای میانگین ۵ و انحراف معیار ۳ و در امتحان پایان ترم دارای میانگین ۶ و انحراف معیار ۴ می‌باشد. فرض کنید نمرات دو امتحان از یکدیگر مستقل بوده و هر یک دارای توزیع نرمال باشند. شخصی در این درس قبول می‌شود که ۲ برابر نمره میان ترم او بعلاوه ۲ برابر نمره پایان ترم او از ۱۸ کمتر نباشد. اگر در این درس ۵۰ نفر ثبت نام کرده باشند، چند نفر آنها رد می‌شوند؟

حل اگر متغیرهای تصادفی X_1 و X_2 به ترتیب نمرات امتحان شخص در میان ترم و پایان ترم باشد و Y نمره کل او باشد آنگاه

$$X_1 \sim N(5, 3^2) \quad , \quad X_2 \sim N(6, 4^2)$$

$$Y = 2X_1 + 2X_2 \sim N(2(5) + 2(6), 2^2(3^2) + 2^2(4^2))$$

بنابراین $Y \sim N(22, 100)$ و در نتیجه

$$P(Y < 18) = P\left(\frac{Y-\mu}{\sigma} < \frac{18-22}{10}\right) = P(Z < -0/4) = 0/3446$$

بنابراین تعدادی که در این درس رد می‌شوند برابر است با $0/3446 \times 50 = 17/23 \approx 17$

۱۲.۵ مسائل حل شده

مثال ۱.۱۲.۵ یک تولیدکننده قطعات کوچک، اجناس خود را در بسته‌های ۲۰ تایی برای مصرف

کنندگانش می‌فرستد. فرض کنید هر قطعه یا معیوب است و یا سالم و احتمال معیوب بودن هر قطعه ۰/۰۵ می‌باشد.

الف – به طور متوسط در هر بسته چند قطعه معیوب وجود دارد؟

ب – احتمال اینکه بسته دلخواهی شامل هیچ قطعه معیوبی نباشد را بیابید.

حل اگر X برابر تعداد قطعات معیوب در بین ۲۰ قطعه یک بسته باشد آنگاه $X \sim B(20, 0/05)$ و بنابراین

الف – $\mu = E(X) = np = 20(0/05) = 1$

ب – $P(X=0) = f_X(0) = \binom{20}{0}(0/05)^0(0/95)^{20} = 0/358$

مثال ۲.۱۲.۵ احتمال آنکه شدت نسبی احساس صوت یک تقویت کننده بیشتر از ۲ دسیبل (dB) باشد برابر ۰/۰۵ است. احتمال اینکه در بین ۱۰ عدد از این تقویت کننده‌ها که به طور مستقل انتخاب شوند

الف – یک تقویت کننده دارای شدت نسبی احساس صوت بیش از ۲ دسیبل باشد؟

ب – حداکثر ۲ تقویت کننده دارای شدت نسبی احساس صوت بیش از ۲ دسیبل باشند؟

ج – حداقل ۲ تقویت کننده دارای شدت نسبی احساس صوت بیش از ۲ دسیبل باشند؟

حل اگر X برابر تعداد تقویت کننده‌هایی در بین ۱۰ تقویت کننده باشد که دارای شدت نسبی احساس صوت بیش از ۲ دسیبل باشد آنگاه $X \sim B(10, 0/05)$ و بنابراین

الف – $P(X=1) = f_X(1) = \binom{10}{1}(0/05)^1(0/95)^9 = 0/315$

ب – $P(X \le 2) = \sum_{x=0}^{2} f_X(x) = \sum_{x=0}^{2} \binom{10}{x}(0/05)^x(0/95)^{10-x} = 0/9885$

ج – $P(X \ge 2) = 1 - P(X < 2) = 1 - \sum_{x=0}^{1} f_X(x)$

$$= 1 - \sum_{x=0}^{1} \binom{10}{x}(0/05)^x(0/95)^{10-x} = 0/0861$$

مثال ۳.۱۲.۵ فرض کنید ۸۰٪ لامپهای ساخته شده توسط یک کارخانه معیوب باشد. اگر ۱۵ عدد از لامپها را به تصادف انتخاب کنیم، مطلوب است محاسبه احتمال اینکه

الف – حداکثر ۸ لامپ سوخته باشد.

ب – ۱۰ لامپ یا بیشتر سوخته باشد.

ج – تعداد لامپهای سوخته از ۸ کمتر و از ۱۲ بیشتر نباشد.

حل اگر X برابر تعداد لامپهای سوخته در بین ۱۵ لامپ باشد آنگاه $X\sim B(۱۵,۰/۸)$ و بنابراین از جدول (I) داریم که

الف – $P(X\le ۸)=۰/۰۱۸۱$

ب – $P(X\ge ۱۰)=۱-P(X<۱۰)=۱-P(X\le ۹)=۱-۰/۰۶۱۱=۰/۹۳۸۹$

ج – $P(۸\le X\le ۱۲)=P(X\le ۱۲)-P(X\le ۷)=۰/۶۰۲۰-۰/۰۰۴۲=۰/۵۹۷۸$

مثال ۴.۱۲.۵ میانگین و واریانس یک توزیع دوجمله‌ای به ترتیب ۳ و ۲ است. احتمال اینکه X بزرگتر یا مساوی ۲ باشد را بیابید.

حل داریم که $E(X)=np=۳ \ , \quad Var(X)=np(۱-p)=۲$

بنابراین $۳(۱-p)=۲$ و یا $p=\frac{۱}{۳}$ و n=۹، در نتیجه $X\sim B(۹,\frac{۱}{۳})$ پس

$$P(X\ge ۲)=۱-P(X<۲)=۱-\sum_{x=۰}^{۱}\binom{۹}{x}\left(\frac{۱}{۳}\right)^x\left(\frac{۲}{۳}\right)^{۹-x}=\frac{۱۶۸۶۷}{۱۹۶۸۳}=۰/۸۵۷$$

مثال ۵.۱۲.۵ فرض کنید موتورهای هواپیما به طور مستقل از یکدیگر با احتمال $\frac{۱}{۵}$ خراب می‌شوند و اگر حداقل نصف از موتورهای هواپیما کار کند، هواپیما سالم به زمین خواهد نشست. کدامیک از هواپیماهای چهار موتوره یا دو موتوره احتمال بیشتری برای یک پرواز موفق دارند.

حل اگر X و Y به ترتیب تعداد موتورهای سالم در طول پرواز برای هواپیماهای ۴ و ۲ موتوره باشند آنگاه $X\sim B(۴,\frac{۴}{۵})$ و $Y\sim B(۲,\frac{۴}{۵})$ و در نتیجه

$$P(X\ge ۲)=۱-P(X<۲)=۱-\sum_{x=۰}^{۱}\binom{۴}{x}\left(\frac{۴}{۵}\right)^x\left(\frac{۱}{۵}\right)^{۴-x}=\frac{۶۰۸}{۶۲۵}=۰/۹۷۲۸$$

$$P(Y\ge ۱)=۱-P(Y<۱)=۱-\binom{۲}{۰}\left(\frac{۴}{۵}\right)^۰\left(\frac{۱}{۵}\right)^۲=\frac{۲۴}{۲۵}=۰/۹۶$$

بنابراین یک هواپیمای چهارموتوره دارای شانس بیشتری برای یک پرواز موفق است.

مثال ۶.۱۲.۵ یک مهندس کنترل کیفیت از هر ۲۴ باطری ماشین که آماده بارگیری است، ۳

باطری را بررسی می‌کند. اگر در بین ۲۴ باطری، ۶ باطری دارای نقص جزئی باشند، احتمال آنکه در نمونه‌ای که بازرسی می‌کند

الف – هیچیک از باطریها نقص جزئی نداشته باشند را بیابید.

ب – فقط یکی از باطریها دارای نقص جزئی باشد را بیابید.

ج – لااقل ۲ باطری دارای نقص باشند را بیابید.

حل اگر X تعداد باطریهایی در بین ۲۴ باطری باشد که دارای نقص جزئی می‌باشند آنگاه $X\sim HG(۲۴,۶,۳)$ بنابراین

الف – $$P(X=۰)=f_X(۰)=\frac{\binom{۶}{۰}\binom{۱۸}{۳}}{\binom{۲۴}{۳}}=\frac{۸۱۶}{۲۰۲۴}=۰/۴۰۳۲$$

ب – $$P(X=۱)=f_X(۱)=\frac{\binom{۶}{۱}\binom{۱۸}{۲}}{\binom{۲۴}{۳}}=\frac{۹۱۸}{۲۰۲۴}=۰/۴۵۳۶$$

ج – $P(X\ge ۲)=۱-P(X<۲)=۱-\{f_X(۰)+f_X(۱)\}=۱-۰/۸۵۶۸=۰/۱۴۳۲$

مثال ۷.۱۲.۵ کارخانه‌ای کالای تولید شده را در جعبه‌های ۲۵ تایی به بازار عرضه می‌کند. قبل از عرضه به بازار از جعبه یک نمونه ۳تایی انتخاب و اگر کالای معیوبی دیده نشد به بازار عرضه می‌گردد ولی اگر کالای معیوب در نمونه دیده شد جعبه به کارخانه بازگردانده می‌شود. احتمال اینکه جعبه‌ای که شامل ۴ کالای معیوب است به بازار عرضه شود را بیابید.

حل اگر X تعداد کالاهای معیوب در نمونه ۳ تایی انتخابی از جعبه شامل ۲۵ کالا باشد آنگاه $X\sim HG(۲۵,۴,۳)$ و بنابراین احتمال موردنظر برابر است با

$$P(X=۰)=f_X(۰)=\frac{\binom{۴}{۰}\binom{۲۱}{۳}}{\binom{۲۵}{۳}}=\frac{۱۳۳۰}{۲۳۰۰}=۰/۵۷۸۳$$

مثال ۸.۱۲.۵ یک تولید کننده لوازم برقی ادعا می‌کند که فقط ۱۰٪ تولیداتش معیوب هستند.

الف – اگر ۲۰ قطعه با جایگذاری از خط تولید برای بازرسی به تصادف انتخاب کنیم، احتمال اینکه حداقل ۴ قطعه معیوب مشاهده شود را بیابید.

ب – اگر در جعبه‌ای ۴۰ قطعه وجود داشته باشد و ۱۰ قطعه بدون جایگذاری از جعبه انتخاب شوند، احتمال اینکه حداکثر یک قطعه معیوب در بین ۱۰ قطعه مشاهده شود را بیابید.

حل الف – اگر X تعداد قطعات معیوب در بین ۲۰ قطعه انتخابی باشد که به طور مستقل انتخاب می‌شوند آنگاه $X\sim B(۲۰,۰/۱)$ و از جدول (I) داریم که

$$P(X\ge ۴)=۱-P(X\le ۳)=۱-۰/۸۶۷=۰/۱۳۳$$

ب – اگر X تعداد قطعات معیوب در بین ۱۰ قطعه انتخابی از جعبه ۴۰ قطعه‌ای باشد آنگاه $X\sim HG(۴۰,۴,۱۰)$ و بنابراین

$$P(X\le ۱)=\sum_{x=۰}^{۱} f_X(x)=\sum_{x=۰}^{۱}\frac{\binom{۴}{x}\binom{۳۶}{۱۰-x}}{\binom{۴۰}{۱۰}}=۰/۷۴۴۱$$

مثال ۹.۱۲.۵ ۸۰٪ لامپهای تولید شده توسط یک کارخانه سالم می‌باشند. از بین ۵۰۰۰ لامپ یک نمونه ۱۰تایی به تصادف انتخاب می‌کنیم، احتمال اینکه دقیقاً ۳ لامپ سالم باشد را بیابید.

حل اگر X تعداد لامپهای سالم در بین ۱۰ لامپ انتخابی باشد آنگاه $X\sim HG(۵۰۰۰,۴۰۰۰,۱۰)$ و چون n=۱۰ نسبت به N=۵۰۰۰ عدد کوچکی است پس تقریباً $X\sim B(۱۰,۰/۸۰)$ و بنابراین

$$P(X=۳)\approx\binom{۱۰}{۳}(۰/۸)^{۳}(۰/۲)^{۷}=۰/۰۰۰۷۸۶$$

مثال ۱۰.۱۲.۵ یک ماشین نویس به طور متوسط ۲ غلط در هر صفحه تایپ می‌کند. احتمال اینکه او در صفحه بعدی

الف – چهار یا بیشتر غلط تایپ کند را بیابید.

ب – هیچ غلطی نداشته باشد را بیابید.

حل اگر X تعداد غلطهای تایپی در یک صفحه باشد آنگاه $X\sim P(۲)$ و بنابراین از جدول (II) داریم که

الف – $P(X\ge ۴)=۱-P(X<۴)=۱-P(X\le ۳)=۱-۰/۸۵۷۱=۰/۱۴۲۹$

ب – $P(X=۰)=f_X(۰)=\frac{e^{-۲}۲^{۰}}{۰!}=e^{-۲}=۰/۱۳۵۳$

مثال ۱۱.۱۲.۵ در یک شرکت تولیدی معین حوادث به نسبت یک حادثه در هر ۲ ماه اتفاق می‌افتند. اگر فرض کنیم که حوادث به طور مستقل رخ دهند، میانگین تعداد حوادث در سال را بیابید. احتمال اینکه در ماه معینی هیچ حادثه‌ای رخ ندهد را بیابید.

حل اگر X و Y به ترتیب تعداد حوادث در یک سال و یک ماه باشند آنگاه $X\sim P(۶)$ و $Y\sim P(\frac{۱}{۲})$ بنابراین $\mu=E(X)=۶$ یعنی میانگین حوادث در سال ۶ می‌باشد و همچنین

$$P(Y=۰)=f_Y(۰)=\frac{e^{-\frac{۱}{۲}}(\frac{۱}{۲})^{۰}}{۰!}=e^{-\frac{۱}{۲}}=۰/۶۰۶۵$$

مثال ۱۲.۱۲.۵ اگر هر دو دقیقه یک نفر وارد کتابخانه‌ای شود، مطلوب است محاسبه احتمال اینکه

الف – حداقل یک نفر بین ساعت دوازده تا دوازده و پنج دقیقه وارد کتابخانه شوند.

ب – یک تا چهار نفر بین ساعت دوازده تا دوازده و پنج دقیقه وارد کتابخانه شوند.

حل اگر X تعداد افرادی باشد که بین ساعت دوازده تا دوازده و پنج دقیقه وارد کتابخانه می‌شوند آنگاه $X\sim P(۲/۵)$ و بنابراین با استفاده از جدول (II) داریم که

الف – $P(X\ge ۱)=۱-P(X<۱)=۱-f_X(۰)=۱-e^{-۲/۵}=۰/۹۱۷۹$

ب – $P(۱\le X\le ۴)=P(X\le ۴)-P(X\le ۰)=۰/۸۹۱۲-۰/۰۸۲۱=۰/۸۰۹۱$

مثال ۱۳.۱۲.۵ اگر $X\sim P(\mu)$ با استفاده از $E(X(X-۱))$، امید ریاضی و واریانس X را محاسبه کنید.

حل تابع احتمال X برابر است با $f_X(x)=\frac{e^{-\mu}\mu^{x}}{x!}\quad ,\quad x=۰,۱,۲,\ldots$

همچنین بسط مک لورن e^{μ} برابر است با $e^{\mu}=\sum_{y=۰}^{+\infty}\frac{\mu^{y}}{y!}$، بنابراین با تغییر متغیر y=x-۱ داریم که

$$E(X)=\sum_{x=۰}^{+\infty}\frac{xe^{-\mu}\mu^{x}}{x!}=\mu e^{-\mu}\sum_{x=۱}^{+\infty}\frac{\mu^{x-۱}}{(x-۱)!}=\mu e^{-\mu}\sum_{y=۰}^{+\infty}\frac{\mu^{y}}{y!}=\mu e^{-\mu}e^{\mu}=\mu$$

همچنین با تغییر متغیر y=x-۲ داریم که

$$E(X(X-۱))=\sum_{x=۰}^{+\infty}\frac{x(x-۱)e^{-\mu}\mu^{x}}{x!}=\mu^{۲}e^{-\mu}\sum_{x=۲}^{+\infty}\frac{\mu^{x-۲}}{(x-۲)!}$$

$$=\mu^{۲}e^{-\mu}\sum_{y=۰}^{+\infty}\frac{\mu^{y}}{y!}=\mu^{۲}e^{-\mu}e^{\mu}=\mu^{۲}$$

بنابراین

$$Var(X)=E(X^{۲})-[E(X)]^{۲}=E(X(X-۱))+E(X)-[E(X)]^{۲}=\mu^{۲}+\mu-\mu^{۲}=\mu$$

مثال ۱۴.۱۲.۵ درظرفی ۹۹ مهره سفید و یک مهره سیاه داریم. از این ظرف ۲۰۰ مرتبه با جایگذاری مهره بیرون می‌آوریم. احتمال اینکه ۷ مرتبه مهره سیاه دیده شود را بیابید.

حل اگر X تعداد مهره‌های سیاه مشاهده شده در بین ۲۰۰ مهره انتخابی باشد آنگاه $X\sim B(۲۰۰,۰/۰۱)$ و چون n=۲۰۰ عددی بزرگ و p=۰/۰۱ به صفر نزدیک می‌باشد بنابراین تقریباً $X\sim P(۲)$ که در آن $\mu = np = ۲۰۰(۰/۰۱) = ۲$. در نتیجه

$$P(X=۷)\approx \frac{e^{-۲}۲^{۷}}{۷!} = ۰/۰۰۳۴$$

مثال ۱۵.۱۲.۵ معلم فراموشکاری به خاطر نمی‌آورد که کدام یک از ۱۲ کلیدی که در دست دارد مربوط به دفتر کار اوست. اگر کلیدها را به تصادف و با جایگذاری امتحان کند.

الف – به طور متوسط باید چند کلید را برای باز شدن در دفتر کارش امتحان کند؟

ب – احتمال اینکه دفتر کارش تنها بعد از سه امتحان باز شود را بیابید.

حل اگر X تعداد کلیدهای امتحان شده تا رسیدن به کلید دفتر کار معلم باشد آنگاه $X\sim G(\frac{۱}{۱۲})$ و بنابراین

الف – $E(X)=\frac{۱}{p}=۱۲$، بنابراین به طور متوسط ۱۲ کلید امتحان می‌شود.

ب – $P(X=۳)=f_X(۳)=(\frac{۱}{۱۲})(\frac{۱۱}{۱۲})^{۳-۱}=\frac{۱۲۱}{۱۷۲۸}=۰/۰۷۰۲$

مثال ۱۶.۱۲.۵ در دهه گذشته، نسبت پرتابهای موفق در یک پایگاه آزمون پرتاب ۰/۸۵ بوده است. فرض کنید که آزمایشی طرح ریزی شده که مستلزم ۳ پرتاب موفق است. احتمال اینکه (الف) دقیقاً ۵ بار، (ب) کمتر از ۵ پرتاب لازم باشد را بیابید.

حل اگر X تعداد پرتابهای لازم تا رسیدن به ۳ پرتاب موفق باشد آنگاه $X\sim NB(۳,۰/۸۵)$ و بنابراین

الف – $P(X=۵)=f_X(۵)=\binom{۵-۱}{۳-۱}(۰/۸۵)^{۳}(۰/۱۵)^{۲}=۰/۰۸۲۹$

ب – $P(X<۵)=f_X(۳)+f_X(۴)=\binom{۳-۱}{۳-۱}(۰/۸۵)^{۳}(۰/۱۵)^{۰}$

$$+\binom{۴-۱}{۳-۱}(۰/۸۵)^{۳}(۰/۱۵)^{۱}=۰/۸۹۰۵$$

مثال ۱۷.۱۲.۵ کارگاهی روزانه ۱۰۰ قطعه الکترونیکی تولید می‌کند که در یک روز خاص ۱۵

قطعه آن معیوب است. بازرسی جهت کنترل کیفیت ۱۰ قطعه از قطعات تولیدی آن روز خاص را مورد بررسی قرار داده و چنانچه حداکثر ۲ قطعه معیوب مشاهده نماید، کیفیت را مطلوب گزارش می‌نماید.

الف – اگر نمونه بدون جایگذاری انتخاب شود احتمال آنکه کیفیت تولید مطلوب گزارش شود را بیابید.

ب – اگر نمونه با جایگذاری انتخاب شود احتمال آنکه هشتمین قطعه انتخابی اولین قطعه معیوب مشاهده شده باشد را بیابید.

حل الف – اگر X تعداد قطعات معیوب در بین ۱۰ قطعه انتخابی بدون جایگذاری باشد آنگاه $X\sim HG(۱۰۰,۱۵,۱۰)$ و بنابراین احتمال موردنظر برابر است با

$$P(X\le ۲)=\sum_{x=۰}^{۲}\frac{\binom{۱۵}{x}\binom{۸۵}{۱۰-x}}{\binom{۱۰۰}{۱۰}}=۰/۸۲۹۵$$

ب – اگر X تعداد قطعات مشاهده شده تا رسیدن به اولین قطعه معیوب باشد آنگاه $X\sim G(۰/۱۵)$ و بنابراین

$$P(X=۸)=(۰/۱۵)(۰/۸۵)^{۸-۱}=۰/۰۴۸۱$$

مثال ۱۸.۱۲.۵ فرض کنید ۲۰٪ انسانها چشمانی میشی دارند. اگر افرادی که به هواپیما سوار می‌شوند به تصادف انتخاب شده باشند، احتمال اینکه هشتمین مسافری که به هواپیما سوار می‌شود سومین مسافری باشد که چشمان او میشی است را بیابید.

حل اگر X تعداد مسافرینی که به هواپیما سوار می‌شوند تا رسیدن به سومین مسافر چشم میشی باشد آنگاه $X\sim NB(۳,۰/۲)$ و بنابراین

$$P(X=۸)=\binom{۸-۱}{۳-۱}(۰/۲)^{۳}(۰/۸)^{۵}=۰/۰۵۵$$

مثال ۱۹.۱۲.۵ سه نفر با هم در یک قهوه خانه سکه پرتاب می‌کنند، آن یکی که در اقلیت باشد پول چای را می‌دهد. اگر سه سکه یک جور بیاید پرتاب سکه دوباره تکرار می‌گردد. احتمال اینکه کمتر از ۴ بار پرتاب سکه لازم باشد را بیابید.

حل اگر X تعداد پرتابهای سکه‌ها تا رسیدن به اولین غیر یک جور باشد آنگاه $X\sim G(p)$ که در آن $p=۱-P(\{HHH,TTT\})=۱-\frac{۲}{۸}=\frac{۳}{۴}$ بنابراین

$$P(X<4)=P(X\le 3)=1-(\frac{1}{4})^{3}=\frac{63}{64}=0/9844$$

مثال ۲۰.۱۲.۵ مدت زمانی که طول می‌کشد تا شخصی فاصله خانه تا ایستگاه قطار را پیاده طی کند دارای توزیع یکنواخت بین ۱۵ تا ۲۰ دقیقه است. اگر شخص ساعت هفت و سی دقیقه از منزل خارج شود احتمال اینکه او سوار قطاری شود که ساعت هفت و چهل و هشت دقیقه به ایستگاه می‌رسد را بیابید.

حل اگر X زمان طی کردن فاصله خانه تا ایستگاه توسط شخص باشد آنگاه $X\sim U(15,20)$ و بنابراین احتمال موردنظر برابر است با

$$P(X<18)=\int_{15}^{18}\frac{1}{20-15}dx=\frac{3}{5}=0/6$$

مثال ۲۱.۱۲.۵ اگر $X\sim U(-1,1)$ مطلوب است محاسبه $P(|X|>\frac{1}{2})$ و $P(\sin\frac{\pi X}{2}>\frac{1}{2})$.

حل

$$f_X(x)=\frac{1}{2}\quad,\quad -1<x<1$$

$$P(|X|>\frac{1}{2})=1-P(|X|\le\frac{1}{2})=1-\int_{-\frac{1}{2}}^{\frac{1}{2}}\frac{1}{2}dx=1-\frac{1}{2}=\frac{1}{2}$$

$$P(\sin\frac{\pi X}{2}>\frac{1}{2})=P(\frac{\pi}{6}<\frac{\pi X}{2}<\frac{5\pi}{6})=P(\frac{1}{3}<X<\frac{5}{3})=P(X>\frac{1}{3})=\int_{\frac{1}{3}}^{1}\frac{1}{2}dx=\frac{1}{3}$$

مثال ۲۲.۱۲.۵ شعاع کره‌ای یک عدد تصادفی بین ۲ و ۴ است. میانگین حجم آن را بیابید. احتمال اینکه حجم آن حداکثر 36π شود را بیابید.

حل اگر X شعاع کره باشد آنگاه $X\sim U(2,4)$ و چون حجم کره برابر $V=\frac{4}{3}\pi X^{3}$ است پس

$$E(V)=E(\frac{4}{3}\pi X^{3})=\frac{4}{3}\pi\int_{2}^{4}\frac{x^{3}}{2}dx=\frac{\pi}{6}x^{4}|_{2}^{4}=\frac{\pi}{6}(240)=40\pi$$

$$P(V<36\pi)=P(\frac{4}{3}\pi X^{3}<36\pi)=P(X^{3}<27)=P(X<3)=\int_{2}^{3}\frac{dx}{2}=\frac{1}{2}$$

مثال ۲۳.۱۲.۵ مدت زمان برحسب دقیقه که قطار تهران مشهد تأخیر دارد متغیر تصادفی نمایی X با میانگین ۱۰ دقیقه است. مطلوب است تعیین $P(X>E(X))$.

حل چون $X\sim E(10)$ پس $E(X)=10$ و در نتیجه

$$P(X>E(X))=P(X>10)=\int_{10}^{+\infty}\frac{1}{10}e^{-\frac{x}{10}}dx=-e^{-\frac{x}{10}}|_{10}^{+\infty}=e^{-1}=0/3679$$

مثال ۲۴.۱۲.۵ فرض کنید X زمان تعمیر یک تلویزیون برحسب ساعت باشد که دارای توزیع نمایی با میانگین ۲ ساعت است.

الف - اگر در یک روز یک تعمیرکار تلویزیونها را یک به یک و مستقلاً تعمیر کند، احتمال اینکه سومین تلویزیون تعمیری اولین تلویزیونی باشد که قبل از ۲ ساعت تعمیر می‌شود را بیابید.

ب - اگر بدانیم که در ۴ ساعت حداقل ۲ تلویزیون تعمیر شده است، احتمال اینکه در همان ۴ ساعت حداکثر ۳ تلویزیون تعمیر شده باشد را بیابید.

حل الف - اگر Y تعداد تلویزیونهای تعمیری تا رسیدن به اولین تلویزیونی باشد که قبل از ۲ ساعت تعمیر می‌شود آنگاه $Y\sim G(p)$ که در آن

$$p=P(X<2)=\int_{0}^{2}\frac{1}{2}e^{-\frac{x}{2}}dx=-e^{-\frac{x}{2}}|_{0}^{2}=1-e^{-1}=0/6321$$

بنابراین

$$P(Y=3)=f_Y(3)=(0/6321)(0/3679)^{3-1}=0/0856$$

ب - اگر M تعداد تلویزیونهایی باشد که در مدت ۴ ساعت تعمیر می‌شوند آنگاه به واسطه رابطه توزیع نمایی و توزیع پواسون، M دارای توزیع پواسون با پارامتر $\mu=4(\frac{1}{2})=2$ است. بنابراین از جدول (II) داریم که

$$P(M\le 3\mid M\ge 2)=\frac{P(2\le M\le 3)}{P(M\ge 2)}=\frac{P(M\le 3)-P(M\le 1)}{1-P(M\le 1)}=\frac{0/8571-0/4060}{1-0/4060}=0/7594$$

مثال ۲۵.۱۲.۵ اگر $X\sim G(p)$ و $Y\sim E(\theta)$ آنگاه θ را برحسب p چنان تعیین کنید که $P(X>1)=P(Y>1)$.

حل

$$P(X>1)=1-P(X\le 1)=1-\{1-q^{1}\}=q=1-p$$

$$P(Y>1)=\int_{1}^{+\infty}\frac{1}{\theta}e^{-\frac{x}{\theta}}=-e^{-\frac{x}{\theta}}|_{1}^{+\infty}=e^{-\frac{1}{\theta}}$$

بنابراین

$$e^{-\frac{۱}{\theta}} = ۱-p \Rightarrow -\frac{۱}{\theta} = Ln(۱-p) \Rightarrow \theta = \frac{-۱}{Ln(۱-p)}$$

مثال ۲۶.۱۲.۵ فرض کنید متغیر تصادفی X مدت زمان تا توقف یک دستگاه برحسب ۵۰ روز باشد که دارای توزیع نمایی با پارامتر یک است.

الف – احتمال آنکه دستگاه حداقل ۱۰۰ روز بدون توقف کار کند را بیابید.

ب – احتمال آنکه دستگاه در ۱۵۰ روز حداکثر ۲ بار توقف کند را بیابید.

حل الف – چون $f_X(x) = e^{-x}, x > ۰$ پس

$$P(X>۲) = \int_{۲}^{+\infty} e^{-x}dx = -e^{-x} |_{۲}^{+\infty} = e^{-۲} = ۰/۱۳۵۳$$

ب – اگر Y تعداد دفعاتی باشد که دستگاه در ۱۵۰ روز توقف می‌کند، آنگاه بواسطه رابطه توزیع نمایی و پواسون، Y دارای توزیع پواسون با پارامتر $\mu = ۳(۱) = ۳$ است پس از جدول (II) داریم که $P(Y \le ۲) = ۰/۴۲۳۲$.

مثال ۲۷.۱۲.۵ فرض کنید به طور متوسط ۸ تاکسی در یک ساعت به یک ایستگاه وارد می‌شوند.

الف – احتمال آنکه در نیم ساعت اول حداقل یک تاکسی وارد ایستگاه شود را بیابید.

ب – احتمال آنکه زمان بین ورود دو تاکسی بیشتر از دو برابر میانگین زمانهای ورود باشد را بیابید.

حل الف – اگر X تعداد تاکسی‌هایی باشد که در نیم ساعت وارد ایستگاه می‌شوند آنگاه $X \sim P(۴)$ و بنابراین

$$P(X \ge ۱) = ۱ - P(X < ۱) = ۱ - f_X(۰) = ۱ - e^{-۴} = ۰/۹۸۱۷$$

ب – اگر Y زمان بین ورود دو تاکسی برحسب ساعت باشد آنگاه $Y \sim E(\frac{۱}{۸})$ و $E(Y) = \frac{۱}{۸}$ و در نتیجه

$$P(Y > ۲E(Y)) = P(Y > \frac{۱}{۴}) = \int_{\frac{۱}{۴}}^{+\infty} ۸e^{-۸y}\, dy = -e^{-۸y} |_{\frac{۱}{۴}}^{+\infty} = e^{-۲} = ۰/۱۳۵۳$$

مثال ۲۸.۱۲.۵ طول عمر موتورهای ساخته شده توسط کارخانه‌ای برحسب سال، دارای توزیع نرمال با میانگین ۱۰ سال و انحراف معیار ۲ سال است. فرض کنید که کارخانه سازنده، موتورها را به مدت ۶ سال ضمانت کرده است.

الف – این کارخانه چند درصد موتورهایی که فروخته است را باید تعویض نماید؟

ب – این کارخانه مدت ضمانت را چقدر باید انتخاب کند، تا حداکثر یک درصد موتورهای فروخته شده تعویض گردد.

حل اگر X طول عمر موتور ساخته شده توسط این کارخانه باشد آنگاه $X \sim N(۱۰, ۲^۲)$ بنابراین

الف – $P(X<۶) = P(\frac{X-\mu}{\sigma} < \frac{۶-۱۰}{۲}) = P(Z < -۲) = ۰/۰۲۲۸$

بنابراین درصد موردنظر برابر است با $۰/۰۲۲۸ \times ٪۱۰۰ = ٪۲/۲۸$

ب – اگر a مدت ضمانت موردنظر باشد آنگاه

$$۰/۰۱ \ge P(X<a) = P(\frac{X-\mu}{\sigma} < \frac{a-۱۰}{۲}) = P(Z < \frac{a-۱۰}{۲}) = \Phi(\frac{a-۱۰}{۲})$$

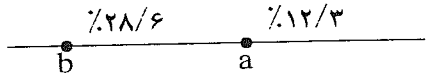
بنابراین با استفاده از جدول (III) داریم که $\frac{a-۱۰}{۲} \le -۲/۳۳ \Rightarrow a \le ۵/۳۴$

مثال ۲۹.۱۲.۵ فرض کنید متغیر تصادفی X نمرات دانش‌آموزان یک کلاس در امتحان آمار باشد که دارای توزیع نرمال با میانگین ۷۰ و انحراف معیار ۱۰ است.

الف – اگر ٪۱۲/۳ از دانش‌آموزان این کلاس نمره A و ٪۲۸/۶ از آنها نمره B بیاورند، پایین‌ترین نمره A و پایین‌ترین نمره B را در این کلاس تعیین کنید.

ب – $P(\frac{۸}{X} > ۰/۱)$ را محاسبه کنید.

حل الف – اگر a پایین‌ترین نمره A و b پایین‌ترین نمره B باشد آنگاه طبق شکل زیر داریم که



$$۰/۱۲۳ = P(X>a) = P(\frac{X-\mu}{\sigma} > \frac{a-۷۰}{۱۰}) = ۱ - \Phi(\frac{a-۷۰}{۱۰})$$

بنابراین $\Phi(\frac{a-۷۰}{۱۰}) = ۰/۸۷۷$ و یا $\frac{a-۷۰}{۱۰} = ۱/۱۶$ و یا $a = ۸۱/۶$ همچنین

$$P(X>b) = ۰/۲۸۶ + ۰/۱۲۳ = ۰/۴۰۹$$

در نتیجه

$$۰/۴۰۹ = P(X>b) = P(\frac{X-\mu}{\sigma} > \frac{b-۷۰}{۱۰}) = ۱ - \Phi(\frac{b-۷۰}{۱۰})$$

بنابراین $\Phi(\frac{b-۷۰}{۱۰}) = ۰/۵۹۱$ و یا $\frac{b-۷۰}{۱۰} = ۰/۲۳$ و یا $b = ۷۲/۳$.

ب – $P(\frac{۸}{X} > ۰/۱) = P(X < ۸۰) = P(\frac{X-\mu}{\sigma} < \frac{۸۰-۷۰}{۱۰}) = \Phi(۱) = ۰/۸۴۱۳$

مثال ۳۰.۱۲.۵ فرض کنید وزن آهن موجود در یک کیلوگرم سنگ آهن یک معدن متغیر تصادفی X است که دارای توزیع نرمال با میانگین صفر و انحراف معیار مجهول σ گرم می‌باشد.

الف- اگر $P(X\ge ۵/۶)=۰/۲۸۷۷$ مقدار $\sigma^۲$ را بیابید.

ب- احتمال اینکه در یک کیلوگرم سنگ آهن بین ۵ تا ۱۱ گرم آهن وجود داشته باشد را به دست آورید.

ج- هنگامی یک معدن مورد بهره‌برداری قرار می‌گیرد که حداقل ۴ گرم آهن در یک کیلوگرم سنگ آهن آن موجود باشد. احتمال اینکه از بین ۶ معدن که به طور مستقل انتخاب می‌شوند، حداقل ۲ معدن مورد بهره برداری قرار نگیرد را بیابید.

حل الف- چون $۰/۲۸۷۷=P(X\ge ۵/۶)$ پس

$$۰/۷۱۲۳=P(X<۵/۶)=P(\frac{X-\mu}{\sigma}<\frac{۵/۶-۰}{\sigma})=\Phi(\frac{۵/۶}{\sigma})$$

بنابراین $۰/۵۶=\frac{۵/۶}{\sigma}$ و یا $\sigma=۱۰$ و در نتیجه $\sigma^۲=۱۰۰$

ب- $$P(۵<X<۱۱)=P(\frac{۵-۰}{۱۰}<\frac{X-\mu}{\sigma}<\frac{۱۱-۰}{۱۰})=P(۰/۵<Z<۱/۱)$$

$$=\Phi(۱/۱)-\Phi(۰/۵)=۰/۸۶۴۳-۰/۶۹۱۵=۰/۱۷۲۸$$

ج- اگر Y تعداد معدنهایی در بین ۶ معدن باشد که مورد بهره برداری قرار نمی‌گیرند آنگاه $Y\sim B(۶,p)$ که در آن

$$p=P(X<۴)=P(\frac{X-\mu}{\sigma}<\frac{۴-۰}{۱۰})=P(Z<۰/۴)=۰/۶۵۵۴$$

بنابراین
$$P(Y\ge۲)=۱-P(Y<۲)=۱-\{f_Y(۰)+f_Y(۱)\}$$

$$=۱-\{(۰/۳۴۴۶)^۶+\binom{۶}{۱}(۰/۶۵۵۴)(۰/۳۴۴۶)^۵\}=۰/۹۷۹۲$$

مثال ۳۱.۱۲.۵ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_X(x)=\frac{۱}{\sqrt{۳۲\pi}}e^{-\frac{۱}{۳۲}(x+۷)^۲}\quad,\quad -\infty<x<+\infty$$

مطلوب است محاسبه $P(X^۲\le۴۹)$ و $P(|X-۳|<۱۲)$.

حل با توجه به فرم تابع چگالی داریم که $X\sim N(-۷,۱۶)$ بنابراین

$$P(X^۲\le۴۹)=P(|X|\le۷)=P(-۷\le X\le۷)=P(\frac{-۷+۷}{۴}\le\frac{X-\mu}{\sigma}\le\frac{۷+۷}{۴})$$

$$=P(۰\le Z\le ۳/۵)=\Phi(۳/۵)-\Phi(۰)=۱-۰/۵=۰/۵$$

$$P(|X-۳|<۱۲)=P(-۱۲<X-۳<۱۲)=P(-۹<X<۱۵)$$

$$=P(\frac{-۹+۷}{۴}<\frac{X-\mu}{\sigma}<\frac{۱۵+۷}{۴})=P(-۰/۵<Z<۵/۵)$$

$$=\Phi(۵/۵)-\Phi(-۰/۵)=۱-۰/۳۰۸۵=۰/۶۹۱۵$$

مثال ۳۲.۱۲.۵ الف- فرض کنید که تولیدات چای در بسته‌های ۱۰۰ گرمی بسته بندی شوند. اگر وزن بسته‌ها دارای توزیع نرمال باشد و مشاهده کنیم که ۵ درصد بسته‌ها وزنی کمتر از ۸۳/۶ و ۲ درصد آنها وزنی بیش از ۱۲۰/۵ گرم دارند، میانگین و واریانس توزیع را به دست آورید.

ب- تصمیم گرفته می‌شود که بسته‌ها را به صورت بسته‌های ۱۰۰۰ گرمی درهم ادغام کنیم (۱۰ بسته را به تصادفی با هم ادغام می‌کنیم). چند درصد بسته‌ها دارای وزنی بیش از ۱۱۰۰ گرم خواهند داشت؟

حل الف- اگر X وزن بسته‌های چای باشد آنگاه $X\sim N(\mu,\sigma^۲)$ و بنابراین

$$۰/۰۵=P(X<۸۳/۶)=P(\frac{X-\mu}{\sigma}<\frac{۸۳/۶-\mu}{\sigma})=\Phi(\frac{۸۳/۶-\mu}{\sigma})$$

$$۰/۰۲=P(X>۱۲۰/۵)=P(\frac{X-\mu}{\sigma}>\frac{۱۲۰/۵-\mu}{\sigma})=۱-\Phi(\frac{۱۲۰/۵-\mu}{\sigma})$$

بنابراین از جدول (III) داریم که

$$\frac{۸۳/۶-\mu}{\sigma}=-۱/۶۴$$
$$\frac{۱۲۰/۵-\mu}{\sigma}=۲/۰۵$$

که با حل دستگاه معادلات فوق $\mu=۱۰۰$، $\sigma=۱۰$ به دست می‌آیند.

ب- اگر Y وزن ۱۰ بسته ادغام شده باشد آنگاه $Y=۱۰X\sim N(۱۰۰۰,۱۰۰۰۰)$ و بنابراین

$$P(Y>۱۱۰۰)=P(\frac{Y-\mu}{\sigma}>\frac{۱۱۰۰-۱۰۰۰}{۱۰۰})=P(Z>۱)=۱-\Phi(۱)$$
$$=۱-۰/۸۴۱۳=۰/۱۵۸۷$$

در نتیجه ۱۵/۸۷٪ وزنی بیش از ۱۱۰۰ گرم دارند.

مثال ۳۳.۱۲.۵ احتمال اینکه بیماری بر اثر عمل جراحی ساده قلب خوب شود ۰/۹ است. احتمال

اینکه بین و خود ۸۴ تا ۹۵ بیمار از ۱۰۰ بیمار بعدی که به این عمل مبادرت می‌کنند، بهبود یابند را بیابید.

حل اگر X تعداد بیمارانی در بین ۱۰۰ بیمار باشد که از عمل جراحی بهبود می‌یابند آنگاه $X\sim B(۱۰۰,۰/۹)$ و از تقریب نرمال داریم که

$$P(۸۴\le X\le ۹۵)=P(۸۳/۵<X<۹۵/۵)$$
$$=P(\frac{۸۳/۵-۹۰}{۳}<\frac{X-np}{\sqrt{npq}}<\frac{۹۵/۵-۹۰}{۳})\approx P(-۲/۱۷<Z<۱/۸۳)$$
$$=\Phi(۱/۸۳)-\Phi(-۲/۱۷)=۰/۹۶۶۴-۰/۰۱۵۰=۰/۹۵۱۴$$

مثال ۳۴.۱۲.۵ یک نمونه تصادفی ۷۲تایی از تابع چگالی احتمال زیر را در نظر بگیرید. احتمال اینکه حداقل ۳۰ عضو نمونه از ۳ بزرگتر باشند تقریباً چقدر است؟

$$f_X(x)=\begin{cases}\frac{۱}{x^۲} & x>۱\\ ۰ & \text{سایر نقاط}\end{cases}$$

حل اگر Y تعدادی از نمونه ۷۲تایی باشد که از ۳ بزرگتر هستند آنگاه $Y\sim B(۷۲,p)$ که در آن

$$p=P(X>۳)=\int_۳^{+\infty}\frac{۱}{x^۲}dx=\frac{-۱}{x}\Big|_۳^{+\infty}=\frac{۱}{۳}$$

بنابراین با استفاده از تقریب نرمال داریم که

$$P(Y\ge ۳۰)=P(Y>۲۹/۵)=P(\frac{Y-np}{\sqrt{npq}}>\frac{۲۹/۵-۲۴}{۴})\approx P(Z>۱/۳۸)$$
$$=۱-\Phi(۱/۳۸)=۱-۰/۹۱۶۲=۰/۰۸۳۸$$

مثال ۳۵.۱۲.۵ یک دستگاه برش با برش طولی و عرضی صفحات فلزی، صفحات مستطیل شکلی را تولید می‌کند. برشهای طولی و عرضی به ترتیب از توزیع نرمال مستقل با میانگینهای مساوی ۲ و واریانسهای ۴ و ۵ پیروی می‌کنند.

الف- احتمال اینکه محیط یک صفحه بریده شده از ۹/۲ کمتر باشد را بیابید.

ب- اگر ۱۰ صفحه مستقلاً بریده شوند احتمال اینکه حداقل ۲ عدد از آنها دارای محیطی کمتر از ۸ باشند را بیابید.

حل الف- اگر X، Y و W به ترتیب طول، عرض و محیط صفحه بریده شده باشند آنگاه X و Y از

یکدیگر مستقل هستند و $X\sim N(۲,۴)$، $Y\sim N(۲,۵)$ و $W=۲X+۲Y\sim N(۸,۳۶)$

بنابراین $P(W<۹/۲)=P(\frac{W-\mu}{\sigma}<\frac{۹/۲-۸}{۶})=\Phi(۰/۲)=۰/۵۷۹۳$

ب- اگر M تعداد صفحات بریده شده در بین ۱۰ صفحه باشد که دارای محیطی کمتر از ۸ هستند آنگاه $M\sim B(۱۰,p)$ که در آن $p=P(W<۸)=P(\frac{W-\mu}{\sigma}<\frac{۸-۸}{۶})=\Phi(۰)=۰/۵$

بنابراین از جدول (I) داریم که $P(Y\ge ۲)=۱-P(Y\le ۱)=۱-۰/۰۱۰۷=۰/۹۸۹۳$

مثال ۳۶.۱۲.۵ نمرات درسهای ۴ واحدی ریاضی، ۳ واحدی آمار و ۳ واحدی زبان به ترتیب دارای توزیع نرمال با میانگینهای ۱۲، ۱۴ و ۱۵ و انحراف معیارهای ۲، $\sqrt{۲}$ و $\sqrt{۲}$ می‌باشد.

الف- اگر ۵۰ نفر این ۳ درس را انتخاب کرده باشند، معدل چند نفرشان در این ۳ درس از ۱۴/۵ بیشتر است؟

ب- اگر ۴۰ نفر از دانشجویانی که این ۳ درس را انتخاب کرده‌اند به طور مستقل انتخاب کنیم، احتمال اینکه حداقل ۱۵ نفر از آنها دارای معدلی کمتر از ۱۳/۵ در این ۳ درس باشند را بیابید.

حل اگر $X_۱$، $X_۲$ و $X_۳$ به ترتیب نمرات دروس ریاضی، آمار و زبان و Y معدل یک نفر در این ۳ درس باشد آنگاه $X_۱\sim N(۱۲,۴)$، $X_۲\sim N(۱۴,۲)$ و $X_۳\sim N(۱۵,۲)$ و

$$Y=\frac{۴X_۱+۳X_۲+۳X_۳}{۱۰}\sim N(۱۳/۵,۱)$$

الف- $P(Y>۱۴/۵)=P(\frac{Y-\mu}{\sigma}>\frac{۱۴/۵-۱۳/۵}{۱})=P(Z>۱)=۱-\Phi(۱)$

$$=۱-۰/۸۴۱۳=۰/۱۵۸۷$$

بنابراین تعداد $۸\approx۷/۹۳۵=(۰/۱۵۸۷)\,۵۰$ نفر از ۵۰ نفر معدلی بیش از ۱۴/۵ دارند.

ب- اگر M تعداد دانشجویان در بین ۴۰ نفر باشند که دارای معدل کمتر از ۱۳/۵ هستند آنگاه $M\sim B(۴۰,p)$ که در آن

$$p=P(Y<۱۳/۵)=P(\frac{Y-\mu}{\sigma}<\frac{۱۳/۵-۱۳/۵}{۱})=\Phi(۰)=\frac{۱}{۲}$$

بنابراین با استفاده از تقریب نرمال داریم که

$$P(M\ge ۱۵)=P(M>۱۴/۵)=P(\frac{M-np}{\sqrt{npq}}>\frac{۱۴/۵-۲۰}{\sqrt{۱۰}})\approx P(Z>-۱/۷۴)$$

$$= ۱-\Phi(-۱/۷۴) = ۱-۰/۰۴۰۹ = ۰/۹۵۹۱$$

۱۳.۵ تمرینات

۱ دانشگاهی n دانشجو دارد که ۷۰٪ آنها به سئوالات یک امتحان هوش در کمتر از ۲۵ دقیقه جواب می‌دهند. احتمال اینکه حداکثر ۲ نفر از ۱۲ دانشجویی که به تصادف انتخاب کرده‌ایم طی مدت ۲۵ دقیقه امتحان را به پایان نرسانند را بیابید.

۲ ویزیتوری برای فروش کالای خاصی به درب منازل مراجعه می‌کند، وی ۱۰ درصد شانس دارد که بتواند کالای خود را به خانم خانه‌دار آن منزل بفروشد. اگر وی در یک روز به ۲۰ منزل مراجعه کرده باشد

الف – احتمال اینکه کمتر از سه فروش داشته باشد را بیابید.

ب – احتمال اینکه یک فروش داشته باشد را بیابید.

ج – احتمال اینکه بیشتر از ۵ فروش داشته باشد را بیابید

۳ سکه سالمی را یکبار پرتاب می‌کنیم، اگر متغیر تصادفی X نشان دهنده تعداد شیر در این آزمایش باشد، توزیع احتمال، میانگین و واریانس X را به دست آورید.

۴ نوعی داروی جدید موجب واکنش ناخوشایند در ۵٪ از مصرف کنندگان می‌شود. اگر این دارو به نمونه‌ای تصادفی مرکب از ۱۳ مصرف کننده داده شود، احتمال هر یک از پیشامدهای زیر را حساب کنید.

الف – هیچ واکنش ناخوشایند.

ب – یک واکنش ناخوشایند.

۵ یک کارشناس امور ترافیک گزارش می‌دهد که ۷۵ درصد از وسایط نقلیه سنگین که از جاده کمربندی استفاده می‌کنند، رانندگان غیربومی دارند، از بین ۵ وسیله نقلیه سنگین که از جاده کمربندی استفاده خواهند نمود، احتمال اینکه حداقل ۳ وسیله آن راننده‌ای غیر بومی داشته باشند را بیابید.

۶ در یک توزیع دوجمله‌ای میانگین برابر ۳ و واریانس برابر ۲ می‌باشد. احتمال اینکه هیچ پیروزی نداشته باشیم را بیابید.

۷ با استفاده از تابع احتمال توزیع دوجمله‌ای و تعریف امید ریاضی ابتدا $E(X)$ و $E(X(X-1))$ را محاسبه و سپس $E(X^2)$ و $\mathrm{Var}(X)$ را برای متغیر تصادفی دوجمله‌ای X پیدا کنید.

۸ در بین ۱۲ باطری خورشیدی در ویترین یک فروشگاه، ۹ باطری صفحه تخت کانونی و بقیه متمرکز هستند. اگر شخصی به طور تصادفی ۴ عدد از این باطریها را برای بررسی انتخاب کند، احتمال اینکه ۳ عدد از آنها صفحه تخت باشند را بیابید.

۹ از ۲۸ نفر اعضای یک گروه آموزشی، ۱۸ نفر اتومبیلهای ساخت خارج و ۱۰ نفر اتومبیلهای ساخت داخل دارند. اگر ۵ نفر اعضاء این گروه را به تصادف انتخاب کنیم، احتمال اینکه حداقل ۳ نفر آنها اتومبیلهای ساخت خارج داشته باشند را بیابید.

۱۰ جعبه‌ای حاوی ۱۰ مقاومت مشابه است که هفت عدد آن سالم و ۳ عدد خراب است. دو مقاومت به تصادف انتخاب شده است. توزیع احتمال را برای تعداد مقاومتهای خراب در نمونه دوتایی محاسبه کنید.

۱۱ از بین ۱۰ موشک چهار عدد انتخاب شده و شلیک می‌گردد. اگر از ۱۰ موشک ۳ تا خراب بوده و شلیک نشوند

الف – انتظار دارید چند موشک در بین چهار موشک خراب باشند؟

ب – احتمال اینکه هر چهار تا شلیک شوند را بیابید.

ج – احتمال اینکه حداکثر ۲ تا شلیک شوند را بیابید.

۱۲ در بین ۱۶ متقاضی کار، ۱۰ نفر آنها دارای مدرک دانشگاهی هستند. اگر ۳ نفر را به طور تصادفی انتخاب کرده و با آنها مصاحبه کنیم، مطلوب است محاسبه احتمال اینکه

الف – هیچ کدام دارای مدرک دانشگاهی نباشند.

ب – حداقل ۲ نفر آنها دارای مدرک دانشگاهی باشند.

۱۳ نصف افراد یک روستا که دارای ۱۲۰۰ نفر جمعیت می‌باشد با انتخاب شدن شخص معینی به نمایندگی روستا موافق می‌باشند. یک نمونه ۱۰تایی از افراد روستا را به تصادف انتخاب می‌کنیم، احتمال اینکه حداقل ۳ نفر موافق این شخص باشند را بیابید.

۱۴ فرض کنید که یک نفر به طور متوسط ۵ بار در سال مبتلا به سرماخوردگی می‌شود. مطلوب است محاسبه احتمال اینکه او در سال آینده

الف- چهار بار مبتلا به سرماخوردگی شود.

ب- حداقل چهار بار مبتلا به سرماخوردگی شود.

ج- حداکثر چهار بار مبتلا به سرماخوردگی شود.

۱۵ تعداد اتومبیلهایی که برای زدن بنزین سوپر به پمپ بنزین معینی مراجعه می‌کنند کاملاً تصادفی بوده و به طور متوسط ۲۴ اتومبیل در ساعت است. اگر برق به مدت ۱۰ دقیقه قطع گردد، احتمال اینکه این پمپ بنزین حداقل ۲ مشتری را در این فاصله از دست بدهد را بیابید.

۱۶ مشتریان یک کتابفروشی به طور متوسط ۶ نفر در ساعت مراجعه می‌کنند. اگر مغازه در ساعت ۹:۳۰ دقیقه صبح باز شود، احتمال اینکه تا ساعت ۱۰ صبح دقیقاً یک نفر و تا ظهر ۱۰ نفر مراجعه کنند را بیابید.

۱۷ تعداد اشعه‌های گامایی که در یک ثانیه از یک مواد رادیواکتیو معین منتشر می‌شوند، متغیر تصادفی پواسون با میانگین ۵/۸ است. اگر برای بیش از ۱۲ اشعه در ثانیه ثبت کننده از کار بیفتد، احتمال اینکه این وسیله در مدت ۲ ثانیه از کار بیفتد را بیابید.

۱۸ فرض کنید به طور متوسط یک تصادف مرگبار اتومبیل در یک روز در شهر معینی رخ می‌دهد.

الف- احتمال اینکه بیش از ۱۰ تصادف از این نوع در هفته رخ دهد را بیابید.

ب- احتمال اینکه بیشتر از ۲ روز بین دو تصادف فاصله بیفتد را بیابید.

۱۹ فرض کنید به طور متوسط از هر ۱۰۰۰ نفر یک نفر ورقه مالیاتی خود را اشتباه پر کند. اگر ۱۰۰۰۰ فرم به طور تصادفی انتخاب شده و امتحان شوند احتمال اینکه ۷ یا ۸ تا از فرمها اشتباه داشته باشند را بیابید.

۲۰ ظرفی دارای ۵ گوی قرمز و ۳ گوی سفید است. از داخل این ظرف یکی یکی و با عمل جایگزینی، آنقدر گوی خارج می‌کنیم تا اولین گوی قرمز خارج شود. اگر X نشان دهنده تعداد دفعات انتخاب گوی باشد

الف- تابع احتمال X، E(X) و Var(X) را به دست آورید.

ب- $P(X\ge 1)$ را محاسبه کنید.

۲۱ سکه سالمی را متوالیاً می‌اندازیم. احتمال اینکه پنجمین خط قبل از دهمین شیر مشاهده شود را بیابید.

۲۲ یادداشتهای یک باشگاه شنا نشان می‌دهد که ۰/۲ استخرهای تازه تأسیس آنها در ظرف سال احتیاج به تعمیر دارد. احتمال اینکه ششمین استخری که در یک سال می‌سازند، اولین استخری باشد که در ظرف یک سال احتیاج به تعمیر داشته باشد را بیابید.

۲۳ سکه‌ای را آنقدر پرتاب می‌کنیم تا ۲ شیر مشاهده شود، اگر بدانیم که حداقل ۳ پرتاب انجام گرفته است احتمال اینکه حداکثر ۶ پرتاب لازم باشد را بیابید.

۲۴ احتمال اینکه تیراندازی به هدف بزند برابر ۰/۸ است. مطلوب است احتمال اینکه

الف- کمتر از ۵ بار لازم باشد تا اولین تیر به هدف بخورد.

ب- حداقل ۵ بار لازم باشد تا سومین تیر به هدف بخورد.

۲۵ ظرفی محتوی ۹ گوی به شماره‌های ۱، ۲،... و ۹ است. از داخل ظرف به طور تصادفی یک گوی را انتخاب می‌کنیم. اگر Xعبارت از شماره گوی انتخابی باشد، تابع احتمال X، E(X) و Var(X) را به دست آورید.

۲۶ زمان لازم برای آنکه معلمی هر روز از منزلش به مدرسه بیاید متغیری تصادفی بین ۲۰ تا ۲۷ دقیقه است. اگر این معلم ساعت ۹ صبح کلاس داشته باشد و ساعت ۸:۳۷ دقیقه صبح منزلش را ترک کند احتمال اینکه به موقع به کلاسش برسد را بیابید.

۲۷ اگر Xدارای توزیع یکنواخت در فاصله (۱,۳-) باشد و Y دارای توزیع نمایی با پارامتر θ باشد، مقدار θ را چنان تعیین کنید که $\sigma_X^2=\sigma_Y^2$.

۲۸ روی پاره خطی به طول l نقطه‌ای را به تصادف انتخاب می‌کنیم. احتمال اینکه پاره خط بزرگتر حداقل ۲ برابر پاره خط کوچکتر شود را بیابید.

۲۹ اگر X دارای توزیع یکنواخت در فاصله (۰,۱) باشد، $E(-LnX)$ را محاسبه کنید.

۳۰ مدت زمانی که یک ساعت بدون وقفه کار می‌کند، متغیر تصادفی نمایی با پارامتر ۵۰ روز می‌باشد.

الف- احتمال اینکه ساعتی کمتر از ۲۰ روز بدون وقفه کار کند را بیابید.

ب- احتمال اینکه ساعتی حداقل ۶۰ روز بدون وقفه کار کند را بیابید.

۳۱ فرض کنید X طول عمر یک لامپ روشنایی باشد که دارای توزیع نمایی با پارامتر θاست. آیا X می‌توانند در رابطه $P(X\le 2)=2P(2<X\le 3)$ صدق کند؟ در صورت امکان برای چه

مقادیری از θ این برابری برقرار می‌شود.

۳۲ متوسط عمر تلویزیونی دارای توزیع نمایی با میانگین ۸ سال است. اگر بدانیم یک تلویزیون ۸ سال کار کرده احتمال اینکه باز کار کند را بیابید.

۳۳ مدت زمانی که طول می‌کشد تا یک نفر در یک آبمیوه فروشی پذیرایی شود یک متغیر تصادفی نمایی با میانگین ۴ دقیقه است. احتمال اینکه یک شخص حداقل در ۴ روز از ۶ روز آینده برای کمتر از ۳ دقیقه پذیرایی شود را بیابید.

۳۴ طول عمر یک قطعه بخصوص برحسب ساعت دارای تابع چگالی زیر می‌باشد.

$$f(t)=\begin{cases}2e^{-2t} & t>0\\ 0 & \text{سایر نقاط}\end{cases}$$

الف- متوسط طول عمر این قطعه را به دست آورید.

ب- اگر پنج قطعه از این قطعات را در پنج دستگاه مختلف که به طور مستقل کار می‌کنند، نصب کنیم احتمال اینکه دقیقاً ۲ قطعه از این ۵ قطعه قبل از ۸ ساعت از بین بروند را بیابید.

۳۵ در یک اتوبان در هر شبانه روز به طور متوسط ۳ تصادف اتومبیل روی می‌دهد.

الف- اگر بدانیم در یک شبانه روز حداقل ۲ تصادف روی داده است، احتمال اینکه در همان شبانه روز حداکثر ۳ تصادف روی داده باشد را بیابید.

ب- احتمال اینکه زمان بین دو تصادف متوالی حداقل ۸ ساعت باشد را بیابید.

۳۶ به یک پمپ بنزین از ساعت ۷ الی ۸ صبح هر روز به طور متوسط ۶ اتومبیل وارد می‌شود. در ساعت ۷:۳۰ یک روز خاص مشاهده گردیده که یک اتومبیل وارد پمپ بنزین شده است.

الف- احتمال اینکه حداکثر تا ساعت ۷:۴۰ همان روز اتومبیل بعدی وارد شود را بیابید.

ب- احتمال اینکه تا ساعت ۸ همان روز هیچ اتومبیل دیگری وارد نشود را بیابید.

۳۷ اگر متغیر تصادفی X دارای توزیع نرمال با میانگین ۳ و انحراف معیار ۳ باشد، احتمالهای زیر را محاسبه کنید.

الف- $P(X\le -1)$ ب- $P(0\le X\le 4)$

ج- $P(X\ge 1/5)$ د- $P(-2\le X\le 1/8)$

۳۸ یک ماشین سازنده نوشابه طوری ساخته شده است که در هر لیوان به طور متوسط ۲۰۷ میلی

لیتر نوشابه می‌ریزد. اگر مقدار نوشابه ریخته شده توسط دستگاه دارای توزیع نرمال با انحراف معیار ۱۵ میلی لیتر باشد.

الف- چند درصد لیوانها شامل بیش از ۲۳۱ میلی لیتر نوشابه هستند.

ب- احتمال اینکه یک لیوان شامل ۱۹۸ تا ۲۱۶ میلی لیتر نوشابه باشد را بیابید.

ج- چند عدد از لیوانها لب ریز خواهند شد اگر لیوانهای ۲۳۷ میلی لیتری را برای ۱۰۰۰ نوشابه بعدی بکار ببریم.

د- ۲۵٪ از لیوانها پایین‌تر از چه مقداری از نوشابه را خواهند داشت؟

۳۹ وقتی میله‌های پلاستیکی از قالب در می‌آیند، به طور خودکار به طولهای ظاهری ۶ اینچ بریده می‌شوند. طولهای واقعی به طور نرمال با میانگین ۶ اینچ و انحراف معیار ۰/۰۶ اینچ توزیع شده‌اند.

الف- چه نسبتی از میله‌ها دارای طولی بین ۵/۹ تا ۶/۱ اینچ می‌باشند؟

ب- اگر ۵۰ لوله بریده شود چه تعداد از آنها دارای طولی بیش از ۵/۸ اینچ هستند؟

۴۰ یک ماشین قطعاتی تولید می‌کند که طول آنها دارای توزیع نرمال با میانگین ۵ و انحراف معیار ۰/۱ سانتیمتر است. تنها قطعاتی مورد قبول قرار می‌گیرند که در فاصله ۵/۰۵±۰/۱۲ سانتیمتر قرار داشته باشند. احتمال اینکه یک قطعه تولید شده بوسیله این ماشین معیوب باشد را بیابید.

۴۱ فرض کنید بهره هوشی دانشجویی که به تصادف از یک دانشگاه انتخاب شده است دارای توزیع نرمال با میانگین ۱۱۰ و انحراف معیار ۲۰ باشد. فاصله‌ای به مرکز میانگین را چنان تعیین کنید که ۵۰٪ بهره هوشی دانشجویان این دانشگاه را شامل شود.

۴۲ اگر X و Y مستقل و هر دو دارای توزیع $N(1,4)$ باشند مطلوب است محاسبه

$$P(X<2Y+1)\quad,\quad P(XY>0)\quad,\quad P(|X-Y|>2)$$

۴۳ برخی از مسافرین هواپیما که قبلاً جا رزرو می‌کنند برای پرواز حاضر نمی‌شوند. شرکت هواپیمایی ممکن است این مسئله را مدنظر داشته باشد و بیشتر از ظرفیت هواپیما برای مسافرین جا رزرو کند. به طور کلی ۴ درصد از مسافرینی که جا رزرو می‌کنند برای پرواز حاضر نمی‌شوند. اگر برای پرواز خاصی ۳۵۰ نفر جا رزرو کرده باشند و ظرفیت هواپیما ۳۴۰ مسافر بیشتر نباشد، احتمال اینکه جا برای کلیه مسافرینی که برای پرواز حاضر می‌شوند وجود داشته باشد را بیابید.

۴۴ یک شرکت یک نوع چای را در جعبه هایی بسته بندی می‌کند که وزن خالص چای در این جعبه‌ها دارای توزیع نرمال با میانگین ۱۰۰ و انحراف معیار ۸ گرم است و وزن جعبه‌ها دارای توزیع نرمال با میانگین ۲۰ و انحراف معیار ۶ گرم می‌باشد.

الف – احتمال اینکه وزن خالص چای در یک جعبه از ۹۲ گرم بیشتر باشد را بیابید.

ب – چند درصد از جعبه‌های چای بسته بندی شده در این شرکت دارای وزن کل (وزن جعبه و چای) کمتر از ۱۳۵ گرم هستند؟

ج – اگر ۴ جعبه را با جایگذاری از این جعبه‌های چای بسته بندی شده انتخاب کنیم احتمال اینکه حداقل یکی از این جعبه‌ها دارای وزن کل بین ۱۱۰ و ۱۳۰ گرم باشد را بیابید.

۴۵ فرض کنید در یک کارخانه برای تولید و عرضه یک محصول ۳ مرحله ساخت، مونتاژ و بسته‌بندی انجام گیرد که زمان انجام این سه مرحله از یکدیگر مستقل بوده و هر یک دارای توزیع نرمال با میانگینهای ۲۰، ۱۲ و ۸ دقیقه و انحراف معیارهای ۶، ۳ و ۲ دقیقه به ترتیب برای ساخت، مونتاژ و بسته بندی باشند.

الف – اگر ۵ محصول به تصادف و با جایگذاری از محصولات این کارخانه انتخاب شوند احتمال اینکه حداقل ۲ عدد از این محصولات دارای زمان تولید و عرضه (ساخت، مونتاژ و بسته بندی) بین ۴۳ و ۵۷ دقیقه باشند را بیابید.

ب – اگر ۵۰ محصول به تصادف و با جایگذاری از محصولات این کارخانه انتخاب شوند احتمال اینکه حداکثر ۲۰ عدد از این محصولات دارای زمان تولید و عرضه بیش از ۵۰ دقیقه باشند را بیابید.

۴۶ در یک کلاس ۷۰ نفری نمرات امتحان درس آمار دارای توزیع نرمال با میانگین مجهول μ و انحراف معیار ۱۰ می‌باشد

الف – اگر ۳۳٪ از افراد این کلاس نمره‌اشان کمتر از ۴۵/۶ باشد، μ را پیدا کنید.

ب – چند نفر در این کلاس نمره اشان بیش از ۷۵ می‌شود.

ج – اگر علاوه بر این امتحان، درس آمار دارای پروژه‌ای باشد که نمره آن دارای توزیع نرمال با میانگین ۱۵ و انحراف معیار ۲ است و کسی در درس آمار قبول می‌شود که جمع نمرات امتحان و پروژه‌اش بیشتر از ۷۰ باشد، چند درصد در این کلاس قبول می‌شوند؟

۴۷ مقاومتی از هشت مولفه تشکیل شده است که به صورت سری به هم لحیم شده‌اند و مقاومت کل برابر مجموع مقاومتهای مؤلفه‌هاست. ۳ مؤلفه دارای میانگین ۲۰۰ و انحراف معیار ۲ اهم است، ۴ مؤلفه دارای میانگین ۱۵۰ و انحراف معیار ۳ اهم است و یک مؤلفه دارای میانگین ۲۵۰ و انحراف معیار ۱ اهم است.

الف – مقاومت ۵ درصد از این مقاومتها از چه مقداری کوچکتر خواهد بود؟

ب – اگر یک نمونه ۴تایی از این مقاومتهای انتخاب کنیم احتمال اینکه حداقل یکی از آنها دارای مقاومتی بیش از ۱۴۴۳ اهم باشد را بیابید.

۴۸ فرض کنید یک امتحان عمومی استخدام طوری طراحی شده است که ۷۰٪ افرادی که ضریب هوشی ۹۰ دارند در این امتحان قبول می‌شوند. اگر ۱۵ نفر از افراد با ضریب هوشی ۹۰ در این امتحان شرکت کنند، احتمال اینکه در بین آنها (الف) حداقل ۱۵ نفر (ب) حداکثر ۶ نفر در امتحان موفق شوند را بیابید.

۴۹ شخص حواس پرتی ۵ کلید شبیه به هم در جیب خود دارد که فقط یکی از آنها کلید درب منزل اوست. وی هنگام رسیدن به منزل یک کلید را به طور تصادفی انتخاب کرده و امتحان می‌کند و در صورت باز نشدن درب، کلید را مجدداً به داخل همان جیب گذاشته وکلید دیگر را امتحان می‌کند. انتظار می‌رود که برای ورود به منزل چند کلید را امتحان کند؟

۵۰ اگر Y دارای توزیع یکنواخت در فاصله $(0,5)$ باشد احتمال اینکه $4x^2+4xY+Y+2=0$ دارای ریشه‌های حقیقی باشد را بیابید.

۵۱ احمد می‌خواهد بسته‌ای شامل ۱۰۰ دیسکت کامپیوتری خود را بفروشد. او نمی‌داند که ۱۰ عدد از این دیسکتها معیوب اند. رضا در صورتی حاضر به خرید آنها می‌شود که نمونه‌ای تصادفی شامل ۱۰ دیسکت، بیش از یک دیسکت معیوب نداشته باشد. احتمال اینکه رضا دیسکتها را بخرد را بیابید.

۵۲ تعداد تلفن‌هایی که به تلفنخانه شرکتی در ساعات اداری می‌شود دارای میانگین ۶ تلفن در هر ۱۰ دقیقه است. اگر تلفنچی برای صرف چای ۵ دقیقه محل کار خود را ترک کند. احتمال اینکه در این مدت (الف) ۲ تلفن (ب) بیش از ۳ تلفن به شرکت شده باشد را بیابید.

۵۳ نمونه‌ای متشکل از ۵۰ قلم کالا در هر ساعت از تولیدات یک ماشین استخراج می‌گردد. اگر در

نمونه دو یا بیشتر معیوب موجود باشد ماشین را برای تنظیم متوقف می‌کنیم. در صورتی که ماشین به طور متوسط ۲ درصد معیوب تولید کند، متوسط فاصله زمان بین دو تنظیم را تعیین کنید.

۵۴ قطر پیچ‌های تولیدی یک کارخانه دارای توزیع نرمال با میانگین ۴ و انحراف معیار ۰/۲ سانتیمتر است. چه نسبتی از پیچ‌ها دارای قطری (الف) بیشتر از ۴/۲ (ب) کمتر از ۳/۸ (ج) بین ۳/۸ تا ۴/۲ سانتیمتر هستند؟

۵۵ دایره‌ای به ۲۵ قطاع مساوی تقسیم و با شماره‌های یک تا ۲۵ متمایز گشته است. اگر توپی را روی محیط این دایره بگردانیم به طوری که این توپ در پایان گردش بتواند در داخل یکی از قطاعها قرار بگیرد و X برابر عددی باشد که هر بار به دست می‌آید، توزیع احتمال X را به دست آورید.

۵۶ در یک شهر ۶٪ راننده‌ها حداقل یک بلیط پارکینگ به دست می‌آورند. احتمال اینکه در بین ۸۰ راننده‌ای که به طور تصادفی در این شهر انتخاب شده‌اند (الف) حداقل ۴ راننده (ب) از ۳ تا ۶ راننده حداقل یک بلیط پارکینگ به دست آورند را بیابید.

۵۷ یک تست چند جوابی دارای ۱۵ سئوال است که هر سئوال دارای چهار جواب احتمالی بوده که فقط یکی از آنها جواب درست است. ورقه شخصی که به طور شانسی علامت زده است، چقدر احتمال دارد که به ۵ تا ۱۰ سئوال پاسخ مثبت داده باشد؟

۵۸ نمره یک درس ۳ واحدی دارای توزیع نرمال با میانگین ۱۲ و انحراف معیار ۲ و نمره یک درس ۲ واحدی دارای توزیع نرمال با میانگین ۱۷ و انحراف معیار ۴ می‌باشد. اگر نمره این دو درس از یکدیگر مستقل باشند.

الف – احتمال اینکه معدل یک نفر در این دو درس از ۱۶ کمتر باشد را بیابید.

ب – اگر از دانشجویانی که در این دو درس ثبت نام کرده‌اند یک به یک و با جایگذاری انتخاب کنیم، احتمال اینکه هفتمین دانشجوی انتخابی سومین دانشجویی باشد که معدل او در این دو درس از ۱۵ بیشتر باشد را بیابید.

۵۹ بازرس کالاهای تولیدی یک کارخانه، هر روز از بین ۱۲ کالایی که برگشت داده می‌شوند ۵ تا را به تصادف انتخاب کرده و آنها را به دقت بازرسی می‌کند. اگر یک روز ۴ کالای برگشتی از بین ۱۲ کالا واقعاً معیوب باشند، احتمال اینکه ۲ تا از کالاهای معیوب در نمونه انتخابی بازرس باشند را

بیابید.

۶۰ فرض کنید ۴۰٪ از ماهیهای یک دریاچه از نوع بخصوصی باشند. اگر هر بار یک ماهی گرفته نوع آن را مشخص و دوباره به دریاچه برگردانیم، احتمال اینکه در دهمین بار پنجمین ماهی از نوع فوق مشاهده شود را بیابید.

۶۱ طول عمر یک نوع کلید الکتریکی دارای توزیع نمایی با میانگین ۲ سال می‌باشد. اگر ۱۰۰ عدد از این کلیدها در دستگاههای مختلف نصب شوند، احتمال اینکه حداکثر ۳۰ عدد از آنها در طول سال اول از بین بروند را بیابید.

۶۲ یک مهندس ایمنی احساس می‌کند که ۳۰٪ همه حوادث صنعتی در کارخانه‌اش به علت کوتاهی کارکنان در فراگیری آموزشها است. احتمال اینکه در بین ۸۴ حادثه صنعتی که در این کارخانه رخ می‌دهند، بیشتر یا مساوی ۲۰ و کمتر یا مساوی با ۳۰ حادثه به دلیل کوتاهی کارکنان از فراگیری آموزشها باشد را بیابید.

۶۳ زمان رسیدن اتوبوسی به ایستگاه یک متغیر تصادفی یکنواخت روی فاصله (a,b) با میانگین ۲ بعد از ظهر و واریانس ۱۲ دقیقه است. مطلوب است تعیین مقادیر a و b.

۶۴ فرض کنید مشتریان یک فروشگاه به طور متوسط ۳۰ نفر در یک ساعت به فروشگاه وارد می‌شوند

الف – احتمال اینکه در عرض ۲ دقیقه هیچ کس یا حداقل ۲ نفر به فروشگاه وارد شوند را بیابید.

ب – احتمال اینکه زمان بین ورود ۲ مشتری حداقل ۲ دقیقه باشد را بیابید.

۶۵ یک الحاقیه در مورد مخالفت اتصال یک ناحیه اطراف شهر که دارای ۱۲۰۰ خانوار است به خود شهر داده شده است. اگر ساکنین نصف آن خانوارها مخالف وابسته شدن به این شهر باشند، احتمال اینکه از یک نمونه تصادفی ۱۰ تایی حداقل ۳ تا موافق الحاقیه باشند را بیابید.

۶۶ سه سکه را که شانس شیر آمدن برای آنها $\frac{1}{2}$، $\frac{1}{3}$ و $\frac{1}{4}$ می‌باشد با هم پرتاب می‌کنیم. احتمال اینکه هر سه شیر بیایند را بیابید. امید ریاضی و واریانس تعداد شیرها را به دست آورید.

۶۷ فرض کنید متغیر تصادفی X زمان مونتاژ کردن یک کالا باشد که دارای توزیع نرمال با میانگین ۲۰ دقیقه و انحراف معیار مجهول σ باشد.

الف- اگر ۹۷/۵ درصد از کالاها کمتر از ۲۹/۸ دقیقه مونتاژ شوند مقدار σ را تعیین کنید.

ب- $P(|X-۱۰|<۵)$ را محاسبه کنید.

ج- اگر ۳ عدد از این کالا به طور مستقل و همزمان مونتاژ شوند، احتمال اینکه حداقل یکی از آنها کمتر از ۲۰ دقیقه مونتاژ شود را بیابید.

۶۸ در یک هتل متقاضیان به طور متوسط ۵ نفر در ساعت مراجعه می‌کنند. فرض کنید تا ۱۰ دقیقه قبل هنوز متقاضی نیامده باشد.

الف- احتمال اینکه متقاضی بعدی کمتر از ۲ دقیقه دیگر بیاید را بیابید.

ب- احتمال اینکه فاصله زمانی آمدن دهمین و یازدهمین متقاضی از دو دقیقه تجاوز نکند را بیابید.

۶۹ فرض کنید طول عمر لامپهای روشنایی شرکتی معین دارای توزیع نرمال با میانگین ۱۰۰۰ و انحراف معیار ۱۰۰ ساعت باشد. همچنین فرض کنید طول عمر لامپهای ساخت شرکتی دیگر نیز دارای توزیع نرمال با میانگین ۹۰۰ و انحراف معیار ۱۵۰ ساعت باشد. اگر از هر یک از این شرکتها یک لامپ خریداری شود، احتمال اینکه حداقل یکی از آنها ۹۸۰ ساعت یا بیشتر عمر کند را بیابید.

فصل ششم

توزیعهای نمونه‌ای

۱.۶ نمونه تصادفی و توزیع نمونه‌ای

همانگونه که در فصل اول اشاره شد آمار به عنوان یک موضوع علمی، شامل مفاهیم و روشهایی جهت جمع آوری، سازماندهی و خلاصه کردن داده‌های به دست آمده از قسمتی از یک جمعیت و انجام استنباط و نتیجه‌گیری از روی آنها در مورد کل جمعیت می‌باشد.

در هر مطالعه آماری با مجموعه افراد یا اشیاء سر و کار داریم که هدف از مطالعه کسب اطلاعات در مورد آنها می‌باشد. این مجموعه را اصطلاحاً **جمعیت** و افراد یا اشیاء تشکیل دهنده آن را اعضاء یا عناصر جمعیت گویند. در بسیاری موارد به علت پرهزینه بودن، وقت‌گیر بودن و... اطلاعات فقط از قسمتی از جمعیت جمع‌آوری می‌گردد و بر اساس اطلاعات حاصله از این قسمت استنباط‌هایی در مورد تمام جمعیت انجام می‌گیرد. انتخاب قسمتی از جمعیت را **نمونه‌گیری** و قسمت انتخاب شده را یک **نمونه** گویند.

در این فصل برخی از مفاهیم نمونه‌گیری که در فصلهای بعد مورد استفاده قرار می‌گیرند را می‌آوریم.

نمونه تصادفی

در مبحث نمونه‌گیری معمولاً جمعیت را با یک متغیر تصادفی X نمایش می‌دهند که این متغیر تصادفی X دارای یک توزیع احتمال بخصوص $f_X(x)$ است. این توزیع احتمال را توزیع احتمال جمعیت گویند. برای مثال اگر جمعیت لامپهای تولیدی یک کارخانه باشد و بخواهیم

مطالعه‌ای روی طول عمر لامپها داشته باشیم آنگاه X را طول عمر لامپ تولیدی کارخانه در نظر می‌گیریم و فرض می‌کنیم $X\sim E(\theta)$. همچنین اگر جمعیت نوزادان باشند و بخواهیم روی قد نوزادان مطالعه‌ای انجام دهیم، در این صورت Xرا طول قد نوزاد در نظر می‌گیریم و فرض می‌کنیم $X\sim N(\mu,\sigma^2)$. حال فرض کنید بخواهیم از این جمعیت یک نمونه به اندازه n انتخاب کنیم. اگر مقدار اولین عضو انتخابی را با X_1 و مقدار دومین عضو انتخابی را با X_2 و... و مقدار nامین عضو انتخابی را X_n نمایش دهیم در این صورت نمونه ما به صورت X_1، X_2،... و X_n می‌باشد که هر کدام دارای همان توزیع احتمال جمعیت یعنی $f_X(x)$ می‌باشند و از یکدیگر مستقل هستند. این نمونه را **نمونه تصادفی** گویند. برای مثال اگر در اندازه‌گیری قد نوزادان نمونه‌ای از ۳ نوزاد انتخاب کرده و اندازه قد نوزاد iام را با X_i $(i=1,2,3)$ نشان دهیم آنگاه X_1، X_2 و X_3 نمونه تصادفی ما را تشکیل می‌دهند که هر کدام دارای همان توزیع جمعیت یعنی $N(\mu,\sigma^2)$ هستند. اگر پس از اندازه‌گیری روی ۳ نوزاد مقادیر $x_1=45$ و $x_2=50$ و $x_3=47$ سانتیمتر را به دست آوریم آنگاه x_1، x_2 و x_3 را مقادیر مشاهده شده این نمونه تصادفی گویند.

برای درک مفهوم نمونه تصادفی به مثال زیر توجه کنید.

مثال ۱.۱.۶ چهار مهره به شماره‌های ۱، ۲، ۳ و ۴ وجود دارد. می‌خواهیم یک نمونه تصادفی به اندازه n=۲ از این مهره‌ها انتخاب کنیم. اگر متغیر تصادفی X نمایانگر شماره روی یک مهره انتخابی باشد آنگاه X دارای تابع احتمال یکنواخت زیر است

x	۱	۲	۳	۴
$f_X(x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

اگر X_1 و X_2 به ترتیب نمونه‌های انتخاب شده در مرتبه اول و دوم باشند آنگاه تعداد نمونه‌های تصادفی ممکن به فرم (X_1,X_2) برابر $16=4^2$ است که هر کدام دارای شانس یکسان $\frac{1}{16}$ برای انتخاب شدن می‌باشند. در جدول ۱.۶ نمونه‌های مختلف همراه با احتمالات آنها آورده شده است.

شماره‌نمونه	(x_1,x_2)	$\bar{x}$	احتمال	شماره‌نمونه	(x_1,x_2)	$\bar{x}$	احتمال
۱	(۱,۱)	۱	$\frac{1}{16}$	۹	(۳,۱)	۲	$\frac{1}{16}$
۲	(۱,۲)	۱/۵	$\frac{1}{16}$	۱۰	(۳,۲)	۲/۵	$\frac{1}{16}$
۳	(۱,۳)	۲	$\frac{1}{16}$	۱۱	(۳,۳)	۳	$\frac{1}{16}$
۴	(۱,۴)	۲/۵	$\frac{1}{16}$	۱۲	(۳,۴)	۳/۵	$\frac{1}{16}$
۵	(۲,۱)	۱/۵	$\frac{1}{16}$	۱۳	(۴,۱)	۲/۵	$\frac{1}{16}$
۶	(۲,۲)	۲	$\frac{1}{16}$	۱۴	(۴,۲)	۳	$\frac{1}{16}$
۷	(۲,۳)	۲/۵	$\frac{1}{16}$	۱۵	(۴,۳)	۳/۵	$\frac{1}{16}$
۸	(۲,۴)	۳	$\frac{1}{16}$	۱۶	(۴,۴)	۴	$\frac{1}{16}$

جدول ۱.۶ جدول نمونه‌های ۲ تایی ممکنه از جمعیت ۴ تایی

مشاهده می‌شود که عناصر اول و دوم نمونه تصادفی یعنی X_1 و X_2 قبل از اینکه مشاهده شوند به عنوان متغیرهای تصادفی در نظر گرفته می‌شوند و مقادیر ۱، ۲، ۳ و ۴ را با توابع احتمال زیر اختیار می‌کنند

i=۱,۲

x_i	۱	۲	۳	۴
$f_{X_i}(x_i)=P(X_i=x_i)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

که دقیقاً همان توزیع احتمال جمعیت یعنی متغیر تصادفی X می‌باشد.

تعریف ۱.۶ فرض کنید متغیر تصادفی X نمایانگر یک جمعیت با توزیع احتمال $f_X(x)$ باشد. یک نمونه تصادفی به اندازه n از این جمعیت عبارت است از جمع آوری n متغیر تصادفی مستقل X_1، X_2،... و X_n که هر کدام دارای توزیع احتمال $f_X(x)$ هستند. مقادیر مشاهده شده این نمونه تصادفی را با x_1، x_2،... و x_n نمایش می‌دهند.

پارامتر و آماره

هر ویژگی یک جمعیت را یک پارامتر جمعیت و یا به اختصار **پارامتر** گویند و ویژگی

متناظر را در نمونه **آماره** می‌نامند. به عنوان مثال میانگین جمعیت μ یک پارامتر جمعیت است در حالی که اگر از جمعیت نمونه‌گیری کرده و میانگین نمونه $\overline{X}=\frac{۱}{n}\sum_{i=۱}^{n}X_i$ را محاسبه کنیم، این میانگین نمونه یک آماره است. به عبارت دقیقتر یک آماره تابعی از نمونه تصادفی است که به پارامتر مجهول بستگی ندارد. توجه کنید که مقدار آماره از یک نمونه به نمونه دیگر تغییر می‌کند ولی مقدار پارامتر جمعیت همواره ثابت است. برای مثال در مثال ۱.۱.۶ میانگین جمعیت مقداری ثابت و برابر است با $\mu=E(X)=\sum_{x=۱}^{۴}x f_X(x)=\frac{۱}{۴}(۱+۲+۳+۴)=۲/۵$
در حالی که میانگین نمونه $\overline{X}$ بسته به اینکه چه نمونه‌ای استخراج شده باشد متغیر خواهد بود.

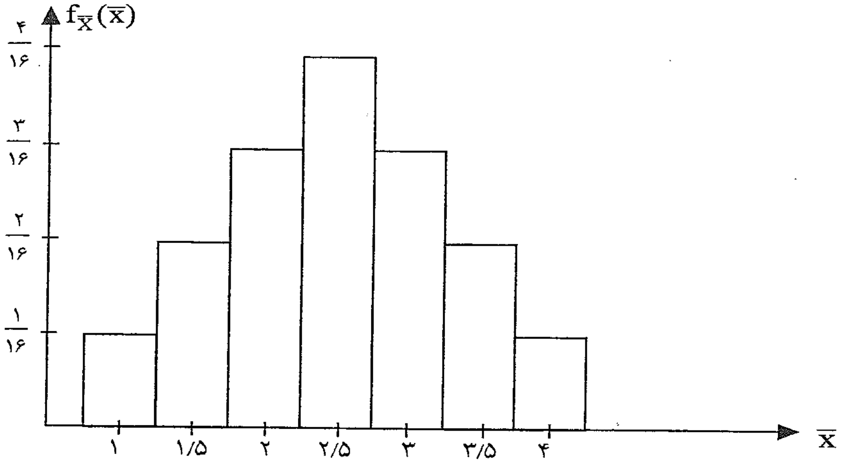
در آمار با انجام نمونه‌گیری و به دست آوردن نمونه تصادفی $X_۱$، $X_۲$،... و X_n آماره‌هایی مانند $\overline{X}$ و $S^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\overline{X})^۲$ را به عنوان میانگین و واریانس نمونه محاسبه کرده و از روی این آماره‌ها استنباط‌هایی در مورد پارامترهای جمعیت مانند میانگین جمعیت μ و واریانس جمعیت $\sigma^۲$ انجام می‌دهیم.

توزیع نمونه‌ای

آماره که تابعی از نمونه تصادفی است خود نیز یک متغیر تصادفی است. هر چند در هر وضعیت مفروض فقط با یک نمونه و یا یک مجموعه از مشاهدات سر و کار داریم و آماره مورد نظر فقط یک مقدار متناظر را دارد، ولی با نمونه های مختلف مقدار آماره مطابق با توزیعی که از روی توزیع نمونه تصادفی معین می‌شود، تغییر می‌کند. نکته مهم این است که رفتار آماره را می‌توان بوسیله یک توزیع احتمال توصیف کرد. بنابراین هر آماره یک متغیر تصادفی است و توزیع احتمال آن را **توزیع نمونه‌ای** گویند.

مثال ۲.۱.۶ در مثال ۱.۱.۶ بر اساس اطلاعات به دست آمده از انتخاب نمونه دو تایی می‌خواهیم میانگین جمعیت μ را حدس بزنیم. برای این منظور شاید بهترین حدس میانگین نمونه دوتایی یعنی $\overline{X}$ باشد. در جدول ۱.۶ میانگین نمونه‌های ممکن دوتایی و احتمالات مربوطه آورده شده است. با توجه به جدول ۱.۶ تابع احتمال و هیستوگرام زیر را برای توزیع احتمال میانگین نمونه $\overline{X}$ داریم.

$\overline{x}$	۱	۱/۵	۲	۲/۵	۳	۳/۵	۴
$P(\overline{X}=\overline{x})$	$\frac{۱}{۱۶}$	$\frac{۲}{۱۶}$	$\frac{۳}{۱۶}$	$\frac{۴}{۱۶}$	$\frac{۳}{۱۶}$	$\frac{۲}{۱۶}$	$\frac{۱}{۱۶}$



توجه کنید که جدول به دست آمده تمام ویژگی‌های یک توزیع احتمال را دارد. تنها اختلاف این توزیع احتمال با توزیعهای احتمال که تاکنون مورد بررسی قرار داده‌ایم در این است که نتایج به دست آمده در اثر انتخاب یک نمونه تصادفی بوده است و به همین دلیل این توزیع احتمال را توزیع نمونه‌ای گویند. در هیستوگرام فوق مشاهده می‌شود که اگر چه توزیع جمعیت مورد نظر یک توزیع یکنواخت است ولی توزیع نمونه میانگین نمونه $\overline{X}$ تغییر شکل داده و به توزیع نرمال گرایش پیدا کرده است.

در بخش‌های بعد توزیع نمونه‌ای چند آماره مهم را که در فصل بعد کاربرد دارند را مورد بررسی قرار می‌دهیم.

۲.۶ توزیع نمونه‌ای میانگین نمونه $\overline{X}$

فرض کنید که از جمعیت X با میانگین μ و واریانس $\sigma^۲$ نمونه‌ای تصادفی $X_۱$، $X_۲$،... و X_n به اندازه n انتخاب کرده باشیم و بخواهیم توزیع نمونه‌ای میانگین نمونه $\overline{X}=\frac{۱}{n}\sum_{i=۱}^{n}X_i$ را به دست آوریم. برای این منظور دو حالت زیر را در نظر می‌گیریم.

الف– اگر جمعیت دارای توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ باشد چون $\overline{X}$ به صورت یک ترکیب خطی از متغیرهای تصادفی نرمال مستقل می‌باشد، بنابراین طبق قضیه ۱۴.۵ داریم که

$$\overline{X}\sim N(\frac{1}{n}\sum_{i=1}^{n}\mu,\frac{1}{n^{2}}\sum_{i=1}^{n}\sigma^{2})$$

و یا

$$\overline{X}\sim N(\mu,\frac{\sigma^{2}}{n})$$

یعنی میانگین نمونه $\overline{X}$ نیز دارای توزیع نرمال با همان میانگین جمعیت μ است و واریانس آن از واریانس هر یک از مشاهدات کمتر است و با افزایش اندازه نمونه مقدار آن کاهش می‌یابد. بنابراین

$$\boxed{Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)} \qquad (۱.۶)$$

ب- اگر جمعیت دارای توزیع نرمال نباشد توزیع $\overline{X}$ به توزیع جمعیت مورد نمونه‌گیری بستگی دارد و با افزایش حجم نمونه n، توزیع نمونه‌گیری $\overline{X}$ تغییر شکل داده و به توزیع نرمال نزدیک می‌شود. این موضوع چنان اهمیت دارد که آن را به صورت قضیه‌ای به نام قضیه حد مرکزی در زیر بیان می‌کنیم.

قضیه ۱.۶ (قضیه حد مرکزی) فرض کنید $X_1, X_2, \ldots$ دنباله‌ای از متغیرهای تصادفی هم توزیع با میانگین μ و واریانس σ^{2} متناهی باشند به قسمی که هر تعداد متناهی از این متغیرهای تصادفی از یکدیگر مستقل باشند. در این صورت توزیع حدی $Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ موقعی که $n\to\infty$ توزیع نرمال استاندارد است.

به عبارت دیگر قضیه حد مرکزی بیانگر این است که وقتی اندازه نمونه n افزایش یابد توزیع میانگین نمونه $\overline{X}$ یک نمونه تصادفی که از هر جمعیتی گرفته شده باشد، توزیع نرمال با میانگین μ و واریانس $\frac{\sigma^{2}}{n}$ است.

تقریب نرمال برای توزیع نمونه‌ای $\overline{X}$ معمولاً موقعی که $n\geq 30$ باشد یک تقریب مناسب است و در صورتی که نمونه تصادفی از جمعیت نرمال انتخاب شده باشد آنگاه طبق قسمت (الف) برای هر مقدار n، $\overline{X}$ دارای توزیع نرمال است.

مثال ۱.۲.۶ یک شرکت تولیدی لاستیک اتومبیل، لاستیک‌هایی تولید می‌کند که طول عمر این لاستیک‌ها دارای توزیع نرمال با میانگین ۲۴ ماه و انحراف معیار ۲ ماه است. احتمال اینکه در یک نمونه تصادفی ۲۵تایی از لاستیک‌ها میانگین طول عمر کمتر از ۲۵ ماه باشد را بیابید.

حل چون جمعیت نرمال است پس $\overline{X}\sim N(24,\frac{2^{2}}{25})$ و در نتیجه

$$P(\overline{X}<25)=P(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<\frac{25-24}{\frac{2}{5}})=P(Z<2/5)=0/9938$$

مثال ۲.۲.۶ فرض کنید مقدار سالهای تحصیل در بین افراد بالغ درکشوری معین دارای میانگین ۱۱/۱ سال و انحراف معیار ۳ سال باشد. احتمال اینکه در یک نمونه تصادفی ۱۰۰ نفری از افراد بالغ متوسط تعداد سالهای تحصیل بین ۱۱ تا ۱۲ سال باشد را بیابید.

حل چون جمعیت نرمال نیست ولی $n>30$ بنابراین تقریباً $\overline{X}\sim N(11/1,\frac{3^{2}}{100})$ و در نتیجه

$$P(11<\overline{X}<12)=P(\frac{11-11/1}{\frac{3}{10}}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<\frac{12-11/1}{\frac{3}{10}})$$
$$\approx P(-0/33<Z<3)=P(Z<3)-P(Z\leq -0/33)$$
$$=0/9987-0/3707=0/628$$

۳.۶ توزیع نمونه‌ای واریانس نمونه S^{2}

یک معیار پراکندگی معقول واریانس نمونه $S^{2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^{2}$ می‌باشد که توزیع نمونه‌ای آن به توزیعی به نام توزیع مربع-کای[1] مربوط می‌شود که در زیر این توزیع را معرفی می‌کنیم.

توزیع مربع-کای

اگر $Z\sim N(0,1)$ آنگاه توزیع $Y=Z^{2}$ را یک توزیع مربع-کای با یک درجه آزادی گویند. همچنین اگر $Z_1, Z_2, \ldots$ و Z_n یک نمونه تصادفی nتایی از جمعیت نرمال استاندارد باشند (یعنی Z_iها از یکدیگر مستقل باشند و $Z_i\sim N(0,1)$) آنگاه توزیع $Y=Z_1^{2}+Z_2^{2}+\ldots+Z_n^{2}$ را یک توزیع مربع-کای با n درجه آزادی گویند و آن را با نماد $Y\sim\chi^{2}_{(n)}$ نمایش می‌دهند.

تابع چگالی احتمال یک متغیر تصادفی مربع-کای با nدرجه آزادی عبارت است از

$$\boxed{f_Y(y)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}y^{\frac{n}{2}-1}e^{-\frac{y}{2}} \qquad y>0} \qquad (۲.۶)$$

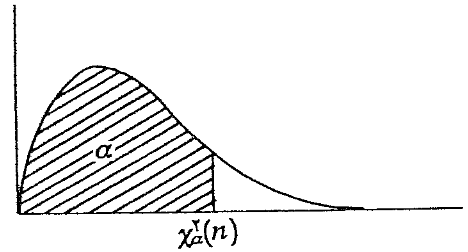
1- Chi-Square distribution

که در آن $\Gamma(.)$ تابع گاما می‌باشد که به صورت زیر تعریف می‌شود.

$$\Gamma(\beta)=\int_{۰}^{\infty} x^{\beta-۱}e^{-x}\,dx \qquad \beta>۰$$

و همواره $\Gamma(\beta)=(\beta-۱)\,\Gamma(\beta-۱)$ و $\Gamma(\frac{۱}{۲})=\sqrt{\pi}$ و اگر β عدد صحیحی باشد آنگاه $\Gamma(\beta)=(\beta-۱)!$.

نمودار تابع چگالی احتمال توزیع مربع-کای در شکل ۱.۶ رسم شده است.



شکل ۱.۶ نمودار تابع چگالی احتمال توزیع $\chi^{۲}_{(n)}$

توزیع مربع-کای یکی از توزیعهای مهم در توزیعهای نمونه‌ای است و کاربرد اصلی آن در مبحث استنباط آماری می‌باشد. برای توزیع مربع-کای جدول (IV) در ضمیمه ارایه گردیده است که در این جدول مقادیر $\chi^{۲}_{\alpha}(n)$ به ازای مقادیر مختلف α و n محاسبه شده است که در آن $\chi^{۲}_{\alpha}(n)$ نقطه‌ای روی محور افقی تابع چگالی احتمال توزیع $\chi^{۲}_{(n)}$ است که مساحت زیر منحنی قبل از این نقطه برابر α است. یعنی

$$Y\sim\chi^{۲}_{(n)} \Rightarrow P(Y\le\chi^{۲}_{\alpha}(n))=\alpha$$

برای مثال از جدول (IV) داریم که $\chi^{۲}_{۰/۹۵}(۱۵)=۲۵/۰$ و $\chi^{۲}_{۰/۸۰}(۲۵)=۳۰/۷$.

مثال ۱.۳.۶ اگر $X\sim\chi^{۲}_{(۱۸)}$ مطلوب است

الف- $P(X\le ۷/۰۱)$ را محاسبه کنید.

ب- اگر $P(X>x)=۰/۰۵$ مقدار x را به دست آورید.

حل الف- چون $\chi^{۲}_{۰/۰۱}(۱۸)=۷/۰۱$ بنابراین $P(X\le ۷/۰۱)=۰/۰۱$

ب- $P(X>x)=۰/۰۵ \Rightarrow P(X\le x)=۰/۹۵ \Rightarrow x=\chi^{۲}_{۰/۹۵}(۱۸)=۲۸/۹$

در قضیه زیر توزیع نمونه‌ای $S^{۲}$ را به دست می‌آوریم.

قضیه ۲.۶ اگر $X_۱, X_۲, \ldots, X_n$ یک نمونه تصادفی از جمعیت نرمال با میانگین μ و واریانس $\sigma^{۲}$ باشند آنگاه

الف-
$$\frac{(n-۱)S^{۲}}{\sigma^{۲}}=\frac{\sum_{i=۱}^{n}(X_i-\bar{X})}{\sigma^{۲}}\sim\chi^{۲}_{(n-۱)}$$

ب- $\bar{X}$ و $S^{۲}$ از یکدیگر مستقل هستند.

مثال ۲.۳.۶ یک جمعیت نرمال دارای واریانس ۶ است. اگر نمونه تصادفی ۲۵ تایی از این جمعیت انتخاب شود احتمال اینکه واریانس نمونه بین ۳/۴۵ و ۱۰/۷۵ باشد را بیابید.

حل

$$P(۳/۴۵<S^{۲}<۱۰/۷۵)=P\left(\frac{۲۴\times ۳/۴۵}{۶}<\frac{(n-۱)S^{۲}}{\sigma^{۲}}<\frac{۲۴\times ۱۰/۷۵}{۶}\right)$$
$$=P(۱۳/۸<\chi^{۲}_{(۲۴)}<۴۳)$$
$$=P(\chi^{۲}_{(۲۴)}<۴۳)-P(\chi^{۲}_{(۲۴)}\le ۱۳/۸)=۰/۹۹-۰/۰۵=۰/۹۴$$

۴.۶ توزیع نمونه‌ای $\frac{\bar{X}-\mu}{S/\sqrt{n}}$

در بخش ۲.۶ مشاهده کردیم که اگر $X_۱, X_۲, \ldots, X_n$ یک نمونه تصادفی از جمعیت نرمال با میانگین μ و واریانس $\sigma^{۲}$ باشند آنگاه $Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ دارای توزیع نرمال استاندارد است. در بعضی موارد مقدار $\sigma^{۲}$ نامعلوم است و به جای آن می‌توان از واریانس نمونه $S^{۲}$ به عنوان یک برآورد یا تخمین استفاده کرد. حال اگر در Z به جای σ مقدار S را قرار دهیم، توزیع نمونه‌ای $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}$ به چه صورت خواهد بود. در این بخش نشان می‌دهیم که توزیع نمونه‌ای T به توزیعی به نام توزیع t مربوط می‌شود که آن را در زیر معرفی می‌کنیم.

توزیع t[1]

فرض کنید $Z\sim N(۰,۱)$ و $Y\sim\chi^{۲}_{(n)}$ و X و Y از یکدیگر مستقل باشند. در این صورت توزیع متغیر تصادفی $T=\frac{Z}{\sqrt{\frac{Y}{n}}}$ را یک توزیع t با n درجه آزادی گویند و آن را با نماد

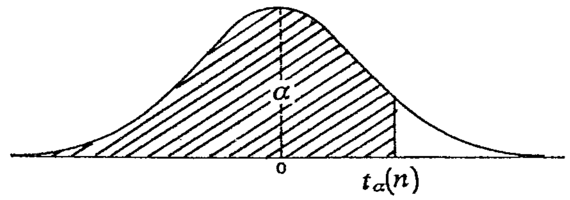
1- t-distribution

$T\sim t_{(n)}$ نمایش می‌دهند.

تابع چگالی احتمال یک متغیر تصادفی t با n درجه آزادی عبارت است از

$$f_T(t)=\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}}(1+\frac{t^2}{n})^{-\frac{n+1}{2}} \qquad -\infty<t<+\infty \qquad (۳.۶)$$

نمودار تابع چگالی احتمال توزیع t در شکل ۲.۶ رسم شده است.



شکل ۲.۶ نمودار تابع چگالی احتمال توزیع $t_{(n)}$

برای توزیع t جدول (V) در ضمیمه ارایه گردیده است که در این جدول مقادیر $t_\alpha(n)$ به ازای مقادیر مختلف α و n محاسبه شده است که در آن $t_\alpha(n)$ نقطه‌ای روی محور افقی تابع چگالی احتمال توزیع $t_{(n)}$ است که مساحت زیر منحنی قبل از این نقطه برابر α است. یعنی

$$T\sim t_{(n)} \Rightarrow P(T<t_\alpha(n))=\alpha$$

برای مثال از جدول (V) داریم که $t_{0/70}(10)=0/542$ و $t_{0/95}(15)=1/75$.

مثال ۱.۴.۶ اگر $T\sim t_{(16)}$ مطلوب است

الف – $P(T>1/34)$ را محاسبه کنید.

ب – اگر $P(T<t)=0/80$ باشد مقدار t را محاسبه کنید.

حل الف – $P(T>1/34)=1-P(T<1/34)=1-0/90=0/10$

ب – $P(T<t)=0/80 \Rightarrow t=t_{0/80}(16)=0/865$

در قضیه زیر توزیع نمونه‌ای $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}$ را به دست می‌آوریم.

قضیه ۳.۶ اگر $\bar{X}$ و S^2 به ترتیب میانگین و واریانس یک نمونه تصادفی به اندازه n از یک

جمعیت نرمال با میانگین μ و واریانس σ^2 باشند آنگاه

$$\boxed{T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t_{(n-1)}} \qquad (۴.۶)$$

اثبات با توجه به فرمول (۱.۶) و قضیه ۲.۶ داریم که

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1) \qquad , \qquad Y=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{(n-1)}$$

و Z و Y از یکدیگر مستقل هستند. پس $T=\frac{Z}{\sqrt{\frac{Y}{n-1}}}\sim t_{(n-1)}$. از طرفی

$$T=\frac{Z}{\sqrt{\frac{Y}{n-1}}}=\frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}}=\frac{\bar{X}-\mu}{S/\sqrt{n}}$$

و بنابراین $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t_{(n-1)}$.

توجه کنید که اگر $n\to\infty$ آنگاه می‌توان نشان داد که توزیع حدی $T=\frac{\bar{X}-\mu}{S/\sqrt{n}}$ یک توزیع نرمال استاندارد است. بنابراین برای $n\geq 30$ توزیع T تقریباً نرمال استاندارد است و به همین علت در جدول (V) مقادیر درجه آزادی بزرگتر از ۳۰ با ∞ نشان داده شده است و مقادیر این ردیف از جدول، با جدول توزیع نرمال یکی است.

مثال ۲.۴.۶ نمرات یک کلاس از دانشجویان دارای توزیع نرمال با میانگین ۱۵ است. اگر از این کلاس یک نمونه ۲۰تایی انتخاب کنیم و مشاهده کنیم که انحراف استاندارد نمرات آنها ۴/۲۸ است، احتمال اینکه میانگین نمرات این افراد از ۱۷ بیشتر باشد را بیابید.

حل در اینجا $\mu=15$، $S=4/28$ و $n=20$ است بنابراین

$$P(\bar{X}>17)=P(\frac{\bar{X}-\mu}{S/\sqrt{n}}>\frac{17-15}{\frac{4/28}{\sqrt{20}}})=P(t_{(19)}>2/09)$$

$$=۱-P(t_{(۱۹)}\le ۲/۰۹)=۱-۰/۹۷۵=۰/۰۲۵$$

۵.۶ توزیع نمونه‌ای اختلاف میانگین‌ها

فرض کنید که دو جمعیت داشته باشیم که جمعیت اول دارای میانگین $\mu_۱$ و واریانس $\sigma_۱^۲$ و جمعیت دوم دارای میانگین $\mu_۲$ و واریانس $\sigma_۲^۲$ باشد. یک نمونه تصادفی $n_۱$ تایی از جمعیت اول انتخاب کرده و میانگین نمونه آن را با $\overline{X}_۱$ و واریانس نمونه آن را با $S_۱^۲$ نمایش می دهیم و یک نمونه تصادفی $n_۲$ تایی از جمعیت دوم انتخاب کرده و میانگین نمونه آن را با $\overline{X}_۲$ و واریانس نمونه آن را با $S_۲^۲$ نمایش می دهیم. همچنین فرض کنید که نمونه گیری از دو جمعیت مستقل از یکدیگر باشند. در این بخش می خواهیم توزیع نمونه‌ای اختلاف میانگین ها یعنی $\overline{X}_۱-\overline{X}_۲$ را به دست آوریم. برای این منظور ۲ حالت زیر را در نظر می گیریم.

حالت اول: واریانس دو جمعیت $\sigma_۱^۲$ و $\sigma_۲^۲$ معلوم باشند

در این حالت بستگی به نرمال بودن جمعیتها دو حالت زیر پیش می آید

الف- اگر دو جمعیت نرمال باشند آنگاه طبق قضیه ۱۴.۵ داریم که

$$\overline{X}_۱\sim N(\mu_۱,\frac{\sigma_۱^۲}{n_۱}) \qquad , \qquad \overline{X}_۲\sim N(\mu_۲,\frac{\sigma_۲^۲}{n_۲})$$

و $\overline{X}_۱$ و $\overline{X}_۲$ از یکدیگر مستقل هستند. بنابراین $\overline{X}_۱-\overline{X}_۲$، که یک ترکیب خطی از متغیرهای تصادفی نرمال است، نیز دارای توزیع نرمال است یعنی

$$\overline{X}_۱-\overline{X}_۲\sim N(\mu_۱-\mu_۲,\frac{\sigma_۱^۲}{n_۱}+\frac{\sigma_۲^۲}{n_۲})$$

و یا

$$\boxed{Z=\frac{(\overline{X}_۱-\overline{X}_۲)-(\mu_۱-\mu_۲)}{\sqrt{\frac{\sigma_۱^۲}{n_۱}+\frac{\sigma_۲^۲}{n_۲}}}\sim N(۰,۱)} \qquad (۵.۶)$$

ب- اگر دو جمعیت نرمال نباشند آنگاه طبق قضیه حد مرکزی برای اندازه نمونه $n_۱\ge ۳۰$ و $n_۲\ge ۳۰$ هر کدام از $\overline{X}_۱$ و $\overline{X}_۲$ دارای توزیع تقریبی نرمال بوده و چون از یکدیگر مستقل هستند بنابراین $\overline{X}_۱-\overline{X}_۲$ نیز تقریباً دارای توزیع نرمال است و نتیجه قسمت (الف) نیز در این حالت

برقرار می باشد.

مثال ۱.۵.۶ دو کارخانه تولید کابل A و B وجود دارند. کابل هایی که توسط کارخانه A تولید می شود، به طور متوسط تحمل ۴۰۰۰ پوند نیروی کششی و انحراف معیار ۳۰۰ پوند را دارند و کابل هایی که توسط کارخانه B تولید می شود، به طور متوسط تحمل ۴۵۰۰ پوند نیرو با انحراف معیار ۲۰۰ پوند را دارند. اگر ۱۰۰ کابل نوع A و ۵۰ کابل نوع B مورد آزمایش قرار گیرند، احتمال اینکه متوسط تحمل نیروی کششی B حداقل ۶۰۰ پوند بیش از نیروی کششی A باشد را بیابید.

حل در اینجا چون $n_A=۱۰۰$ و $n_B=۵۰$ پس تقریباً

$$\overline{X}_A-\overline{X}_B\sim N(۴۰۰۰-۴۵۰۰,\frac{۳۰۰^۲}{۱۰۰}+\frac{۲۰۰^۲}{۵۰})$$

و یا تقریباً $\overline{X}_A-\overline{X}_B\sim N(-۵۰۰\,,\,۱۷۰۰)$ بنابراین

$$P(\overline{X}_B\ge\overline{X}_A+۶۰۰)=P(\overline{X}_A-\overline{X}_B\le -۶۰۰)$$

$$=P(\frac{(\overline{X}_A-\overline{X}_B)-(\mu_A-\mu_B)}{\sqrt{\frac{\sigma_A^۲}{n_A}+\frac{\sigma_B^۲}{n_B}}}\le\frac{-۶۰۰+۵۰۰}{\sqrt{۱۷۰۰}})$$

$$\approx P(Z\le -۲/۴۳)=۰/۰۰۷۵$$

حالت دوم: واریانس دو جمعیت نامعلوم اما مساوی باشند

در این حالت $\sigma_۱^۲=\sigma_۲^۲=\sigma^۲$ می باشد و $\sigma^۲$ که واریانس مشترک دو جمعیت می باشد نامعلوم است. در جمعیت اول می توان از $S_۱^۲$ و در جمعیت دوم می توان از $S_۲^۲$ به عنوان یک برآورد یا تخمین برای $\sigma^۲$ استفاده کرد. اما بهتر است که از یک میانگین وزنی این دو کمیت برای برآورد $\sigma^۲$ استفاده کنیم. بنابراین از واریانس مشترک دو نمونه یعنی

$$\boxed{S_P^۲=\frac{(n_۱-۱)S_۱^۲+(n_۲-۱)S_۲^۲}{(n_۱+n_۲-۲)}} \qquad (۶.۶)$$

به عنوان یک برآورد یا تخمین $\sigma^۲$ استفاده می‌کنیم. حال اگر دو جمعیت نرمال باشند آنگاه با استفاده از قضیه ۲.۶ داریم که

$$\frac{(n_1+n_2-2)S_P^2}{\sigma^2}=\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2_{(n_1+n_2-2)}$$

و بنابراین با استفاده از تعریف توزیع t به راحتی می‌توان نشان داد که

$$\boxed{T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{(n_1+n_2-2)}} \qquad (۷.۶)$$

۶.۶ توزیع نمونه‌ای نسبت واریانسهای نمونه

دو جمعیت مفروض در بخش ۵.۶ را در نظر بگیرید. در این بخش می‌خواهیم توزیع نمونه‌ای نسبت واریانسهای نمونه یعنی $\frac{S_1^2}{S_2^2}$ را به دست آوریم. توزیع نمونه‌ای این نسبت به توزیعی بنام توزیع F مربوط می‌شود که آن را در زیر معرفی می‌کنیم.

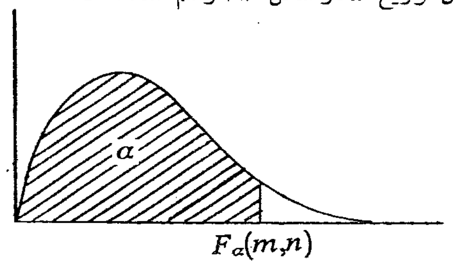
توزیع F[1]

اگر $U\sim\chi^2_{(m)}$ و $V\sim\chi^2_{(n)}$ و متغیرهای تصادفی U و V از یکدیگر مستقل باشند آنگاه توزیع متغیر تصادفی $F=\frac{U/m}{V/n}$ را یک توزیع F با m و n درجه آزادی گویند و آن را با نماد $F\sim F_{(m,n)}$ نمایش می‌دهند.

تابع چگالی احتمال یک توزیع F با درجات آزادی m و n عبارت است از

$$\boxed{f_F(x)=\frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\left(\frac{m}{n}\right)^{\frac{m}{2}}\frac{x^{\frac{m}{2}-1}}{(1+\frac{m}{n}x)^{\frac{m+n}{2}}} \qquad x>0} \qquad (۸.۶)$$

نمودار تابع چگالی احتمال توزیع F در شکل ۳.۶ رسم شده است.



شکل ۳.۶ نمودار تابع چگالی احتمال توزیع $F_{(m,n)}$

1- F-distribution

برای توزیع F جدول (VI) در ضمیمه ارایه گردیده است که در این جدول $F_\alpha(m,n)$ به ازای مقادیر مختلف m و n و $\alpha=0/90, 0/95, 0/99$ محاسبه شده است. که در آن $F_\alpha(m,n)$ نقطه‌ای روی محور افقی تابع چگالی احتمال توزیع $F_{(m,n)}$ است که مساحت زیر منحنی قبل از این نقطه برابر α است. برای مثال از جدول (VI) داریم که $F_{0/90}(6,14)=2/24$ و $F_{0/95}(20,10)=2/77$.

مثال ۱.۶.۶ اگر $F\sim F_{(10,12)}$ مطلوب است

الف - $P(F<2/75)$ را محاسبه کنید.

ب - اگر $P(F>x)=0/01$ باشد مقدار x را محاسبه کنید.

حل الف - چون $F_{0/95}(10,12)=2/75$ پس $P(F<2/75)=0/95$

ب - $P(F>x)=0/01 \Rightarrow P(F\le x)=0/99 \Rightarrow x=F_{0/99}(10,12)=4/30$

در قضیه زیر توزیع نمونه‌ای $\frac{S_1^2}{S_2^2}$ را به دست می‌آوریم.

قضیه ۴.۶ اگر S_1^2 و S_2^2 به ترتیب واریانسهای نمونه‌های تصادفی مستقل به اندازه‌های n_1 و n_2 از جمعیتهای نرمال با واریانسهای σ_1^2 و σ_2^2 باشند آنگاه

$$\boxed{F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F_{(n_1-1,n_2-1)}} \qquad (۹.۶)$$

اثبات از قضیه ۲.۶ نتیجه می‌شود که

$$U=\frac{(n_1-1)S_1^2}{\sigma_1^2}\sim\chi^2_{(n_1-1)} \quad , \quad V=\frac{(n_2-1)S_2^2}{\sigma_2^2}\sim\chi^2_{(n_2-1)}$$

و U و V از یکدیگر مستقل هستند. بنابراین از تعریف توزیع F داریم که

$$F=\frac{U/(n_1-1)}{V/(n_2-1)}\sim F_{(n_1-1,n_2-1)}$$

از طرفی

$$F=\frac{(n_1-1)S_1^2/[\sigma_1^2(n_1-1)]}{(n_2-1)S_2^2/[\sigma_2^2(n_2-1)]}=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

بنابراین نتیجه قضیه حاصل می‌شود.

مثال ۲.۶.۶ اگر S_1^2 و S_2^2 نمایانگر واریانسهای نمونه‌های تصادفی مستقل به اندازه‌های ۲۵ و

۳۱ تایی به ترتیب از جمعیت‌های نرمال با واریانسهای ۱۰ و ۱۵ باشند، احتمال اینکه S_1^2 حداقل ۱/۲۶ برابر S_2^2 باشد چقدر است؟

حل

$$P\left(\frac{S_1^2}{S_2^2} \ge 1/26\right) = P\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \ge 1/26\left(\frac{15}{10}\right)\right) = P\left(F_{(24,30)} \ge 1/89\right)$$

$$= 1 - P\left(F_{(24,30)} < 1/89\right) = 1 - 0/95 = 0/05$$

۷.۶ تمرینات

۱ جعبه‌ای حاوی ۵ مهره است که از ۱ تا ۵ شماره‌گذاری شده‌اند. از این جعبه نمونه‌ای مرکب از ۲ مهره به تصادف و با جایگذاری استخراج می‌شود.

الف - با تشکیل جدول توزیع احتمال توأم X_1 و X_2، توزیع $\overline{X} = \frac{X_1 + X_2}{2}$ را محاسبه کنید.

ب - میانگین و واریانس $\overline{X}$ را از روی توزیع آن محاسبه کنید.

ج - میانگین و واریانس جمعیت را پیداکنید.

د - نمودار توزیع جمعیت (برای یک مشاهده X) و نمودار توزیع نمونه‌ای $\overline{X}$ را رسم کنید.

۲ جعبه‌ای پر از مهره‌های هم شکل است که یک سوم آنها با شماره ۲ و یک سوم با شماره ۴ و یک سوم با شماره ۶ مشخص شده‌اند.

الف - یک مهره استخراج می‌گردد اگر شماره آن X باشد، میانگین μ و انحراف معیار σ جمعیت را بیابید.

ب - نمونه‌ای مرکب از دو مهره با جایگذاری استخراج می‌شود. اگر $\overline{X}$ میانگین نمونه باشد، جدول توزیع احتمال $\overline{X}$ را به دست آورید و از روی آن میانگین و واریانس $\overline{X}$ را حساب کنید.

ج - قسمت (ب) را برای نمونه‌ای متشکل از ۳ مهره حل کنید.

۳ اندازه‌گیری قطر گام نوعی پیچ دارای توزیع نرمال با میانگین ۰/۴۰۰۸ اینچ و انحراف معیار ۰/۰۰۰۳ اینچ است. حدود مشخصات طراحی به صورت ۰/۰۰۱±۰/۴ اینچ ارایه شده است. نمونه‌های ۴ تایی از قطرهای گام گرفته می‌شود. احتمال اینکه میانگین نمونه در داخل حدود مشخصات قرار گیرد را بیابید. احتمال تجاوز میانگین از ۰/۴۰۰۵ را بیابید.

۴ عرض یک شکاف که بر یک قطعه از آلیاژ آلومینیوم که با ریخته‌گری تولید می‌شود، توزیع نرمال با میانگین ۰/۹ اینچ و انحراف معیار ۰/۰۲ اینچ دارد. حدود مشخصات طراحی عبارت از ۰/۰۰۵±۰/۹ اینچ است. هر ساعت نمونه‌هایی ۵تایی از آلیاژ ریخته‌گری گرفته شده و میانگین آن محاسبه می‌شود. چند درصد از این میانگین‌های نمونه در خارج از حدود مشخصات طراحی قرار می‌گیرند؟ حدود را طوری تعیین کنید که درصد میانگین‌های نمونه که خارج از حدود قرار می‌گیرند معادل ۰/۲۷ درصد باشد.

۵ مقدار پولی که مردم شهر معینی درکیف خود دارند بطور متوسط ۹۰۰۰ تومان با انحراف معیار ۲۵۰۰ تومان است. احتمال آنکه گروهی مرکب از ۲۲۵ نفر به طور متوسط بیش از ۹۱۰۰ تومان همراه داشته باشند را بیابید.

۶ فرض کنید طول دانه‌های زنجیر دوچرخه‌ای به طور متوسط ۰/۵۰ اینچ با انحراف معیار ۰/۰۴ اینچ باشد. استانداردهای سازنده دوچرخه اقتضا می‌کند که زنجیر بین ۴۹ و ۵۰ اینچ طول داشته باشد. اگر زنجیرها از ۱۰۰ دانه درست شده باشند، چه نسبتی از آنها مطابق استاندارد می‌باشند؟

۷ اگر اندازه نمونه‌ای ۳۶ و خطای استاندارد آن ۲ باشد، اندازه نمونه چقدر باید باشد تا اینکه خطای استاندارد آن کاهش پیدا نموده و به ۱/۲ برسد؟

۸ در فاصله (۰,۱) تعداد ۷۵ عدد را به تصادف انتخاب می‌کنیم. نشان دهید احتمال اینکه معدل این عددها در فاصله (۰/۴۵,۰/۵۵) قرار گیرد تقریباً برابر ۰/۸۶۶ است.

۹ تعداد مکالمات تلفنی که از طریق مرکزی انجام می‌شود به طور متوسط ۴ مکالمه در دقیقه است. احتمال تقریبی اینکه طی یک ساعت آینده حداقل ۲۵۰ مکالمه از طریق این مرکز انجام شود را بیابید.

۱۰ وزن یک گلوله فلزی دارای میانگین ۵/۰۲ اونس و انحراف معیار ۰/۳۰ اونس است. احتمال اینکه مجموع وزن یک نمونه تصادفی ۱۰۰تایی انتخاب شده از این نوع گلوله‌های فلزی (الف) بین ۴۹۶ تا ۵۰۰ اونس (ب) بیش از ۵۱۰ اونس باشد را بیابید.

۱۱ پژوهشگری می‌خواهد میانگین جمعیتی را با استفاده از نمونه‌ای برآورد کند. وی مایل است

نمونه به اندازه کافی بزرگ باشد تا احتمال اینکه تفاوت میانگین نمونه از میانگین جمعیت بیشتر از ۲۵ درصد انحراف استاندارد نباشد، ۰/۹۵ باشد. نمونه را بایستی چقدر بزرگ بگیرد؟

۱۲ اندازه قد مردان در یک جمعیت دارای توزیع نرمال با میانگین ۱۶۷ و انحراف معیار ۳ سانتیمتر می‌باشد.

الف – اگر ۱۰۰ نفر به طور مستقل از این جمعیت انتخاب شوند احتمال اینکه حداقل ۵۵ نفر آنها طول قدشان کمتر از ۱۶۷ سانتیمتر باشد را بیابید.

ب – اگر ۱۰۰ نفر به طور مستقل از این جمعیت انتخاب شوند احتمال اینکه میانگین قد این ۱۰۰ نفر بیش از ۱۶۷/۶ سانتیمتر باشد را بیابید.

۱۳ قد ۱۰۰۰ دانشجو دارای توزیع نرمال با میانگین ۶۸/۵ اینچ و انحراف معیار ۲/۷ اینچ است. اگر ۲۰۰ نمونه تصادفی ۲۵تایی از این جمعیت اختیار گردد. مطلوب است

الف – میانگین و واریانس توزیع نمونه‌ای میانگین را به دست آورید.

ب – تعداد میانگین نمونه راکه در فاصله ۶۷/۹ تا ۶۹/۲ قرار می‌گیرند را به دست آورید.

۱۴ مطلوب است محاسبه مقادیر زیر

$$\chi^2_{0/05}(28)\ ,\ \chi^2_{0/20}(17)\ ,\ \chi^2_{0/70}(40)\ ,\ \chi^2_{0/90}(24)\ ,\ \chi^2_{0/95}(15)$$

۱۵ اگر متغیر تصادفی X دارای توزیع مربع-کای با ۵ درجه آزادی باشد، مطلوب است محاسبه اعداد a و b در هر یک از حالات زیر

الف – $P(X>b)=0/05$ ب – $P(X<a)=0/10$

ج – $P(X>b)=0/01$ د – $P(a<X<b)=0/95$

۱۶ یک نمونه تصادفی ۱۰تایی از یک جمعیت نرمال با واریانس ۴۲/۵ گرفته شده است، احتمال اینکه انحراف استاندارد نمونه بین ۳/۱۴ و ۸/۹۴ باشد را پیداکنید.

۱۷ طول عمر لامپهای تصویر تلویزیون ساخت کارخانه‌ای دارای توزیع نرمال با میانگین ۲۰۰۰ ساعت و انحراف معیار ۶۰ ساعت است. اگر ۱۰ لامپ تصویر تلویزیون ساخت این کارخانه به طور تصادفی انتخاب شود، احتمال اینکه انحراف استاندارد این ۱۰ لامپ (الف) بیش از ۵۰ ساعت نباشد، (ب) بین ۵۰ و ۷۰ ساعت باشد را بیابید.

۱۸ احتمال اینکه یک نمونه تصادفی ۲۰تایی از یک جمعیت نرمال با واریانس $\frac{19}{3}$ دارای واریانس

نمونه بین ۶/۱ و ۱۴/۶ باشد را بیابید.

۱۹ فرض کنید X_1، X_2، ... و X_{10} نمونه‌ای تصادفی از جمعیت نرمال استاندارد باشند، $P(2/56 < \sum_{i=1}^{10} X_i^2 < 18/3)$ را حساب کنید.

۲۰ مطلوب است محاسبه مقادیر زیر

$$t_{0/90}(15)\ ,\ t_{0/95}(20)\ ,\ t_{0/975}(30)\ ,\ t_{0/80}(50)\ ,\ t_{0/90}(70)$$

۲۱ اگر متغیر تصادفی X دارای توزیع t با ۹ درجه آزادی باشد، مطلوب است محاسبه عدد a در هر یک از حالتهای زیر

الف – $P(X>a)=0/05$ ب – $P(X<a)=0/01$

ج – $P(-a<X<a)=0/95$ د – $P(-a<X<a)=0/99$

۲۲ یک نمونه تصادفی ۲۵ تایی از یک جمعیت نرمال با حد متوسط ۸۰ و انحراف معیار ۵ انتخاب شده است. یک نمونه تصادفی ۳۶ تایی از یک جمعیت نرمال دیگری با حد متوسط ۷۵ و انحراف معیار ۳ انتخاب شده است. احتمال اینکه میانگین نمونه محاسبه شده از نمونه ۲۵ تایی حداقل ۳/۴ از میانگین نمونه محاسبه شده از نمونه ۳۶ تایی بیشتر و از ۵/۹ کوچکتر باشد را بیابید.

۲۳ از دو جمعیت نرمال با میانگین‌های مساوی و واریانسهای ۵۰ و ۴۰، دو نمونه تصادفی ۱۰۰تایی به طور مستقل انتخاب کرده‌ایم. احتمال اینکه تفاضل میانگین‌ها بیش از ۲ باشد را بیابید.

۲۴ لامپهای الکتریکی ساخت کارخانه A دارای طول عمر متوسط ۱۴۰۰ ساعت و انحراف معیار ۲۰۰ ساعت است. لامپهای الکتریکی ساخت کارخانه B دارای طول عمر متوسط ۱۲۰۰ ساعت و انحراف معیار ۱۰۰ ساعت است. اگر از لامپهای ساخت هر کارخانه، یک نمونه تصادفی ۱۰۰تایی مستقل از یکدیگر انتخاب کنیم و مورد آزمایش قرار دهیم، مطلوب است محاسبه احتمال اینکه متوسط عمر لامپهای انتخابی از کارخانه A حداقل

الف – ۱۶۰ ساعت بیش از متوسط عمر لامپهای انتخابی از کارخانه B باشد.

ب – ۲۵۰ ساعت بیش از متوسط عمر لامپهای انتخابی از کارخانه B باشد.

۲۵ فرض کنید $\overline{X}_1$ و $\overline{X}_2$ میانگین‌های دو نمونه به اندازه n از جمعیتی با واریانس σ^2 باشند. مقدار n را چنان تعیین کنید تا احتمال اینکه این دو میانگین نمونه بیشتر از σ اختلاف داشته باشند. تقریباً ۰/۰۱ باشد.

۲۶ فرض کنید در دو جمعیت میانگین مصرف روزانه پروتئین به ترتیب ۱۲۵ و ۱۰۰ گرم باشد. اگر مقادیر مصرف روزانه پروتئین در دو جمعیت دارای توزیع نرمال با انحراف معیار ۱۵ گرم باشد، احتمال اینکه نمونه‌های تصادفی و مستقل ۲۵ نفری از هر جمعیت، دارای تفاوت بین میانگین‌های نمونه کمتر از ۱۲ گرم باشد را بیابید.

۲۷ میانگین نمره هوش دانشجویان سال اول، در یک کالج بخصوص ۵۴۰ با انحراف معیار ۵۰ است. احتمال اینکه دو گروه از دانشجویان انتخابی به طور تصادفی به ترتیب شامل ۳۲ و ۵۰ دانشجو در حد متوسط نمره‌هایشان (الف) بیش از ۲۰ نمره، (ب) مقداری بین ۵ و ۱۰ نمره، فرق داشته باشند را بیابید.

۲۸ میانگین و انحراف معیار نمرات درس آمار دانشجویان یک دانشگاه به ترتیب ۷۲ و ۸ می‌باشد. دو نمونه تصادفی مستقل ۳۶ و ۴۰ تایی انتخاب می‌کنیم، احتمال اینکه تفاضل میانگین نمرات این دو نمونه بین ۲ و ۵ باشد را بیابید.

۲۹ یک نوع گلوله فلزی ساخت کارخانه‌ای به طور متوسط دارای وزن ۰/۵۰ اونس با انحراف معیار ۰/۰۲ اونس است. مطلوب است محاسبه احتمال اینکه دو بسته ۱۰۰۰تایی از این نوع گلوله‌ها بیش از ۲ اونس با یکدیگر تفاوت وزنی داشته باشند.

۳۰ مقادیر زیر را محاسبه کنید

$$F_{۰/۹۵}(۷,۱۵)\ , F_{۰/۹۵}(۱۵,۷) , F_{۰/۹۹}(۲۴,۱۹)\ , F_{۰/۹۹}(۱۹,۲۴)\ , F_{۰/۹۰}(۲۸,۱۲)$$

۳۱ از دو جمعیت نرمال با واریانسهای ۲۰ و ۳۰ به ترتیب نمونه‌های تصادفی ۸ و ۱۰ تایی انتخاب کرده‌ایم. احتمال اینکه واریانس نمونه اول بیش از ۲ برابر واریانس نمونه دوم باشد را بیابید.

۳۲ اگر نمونه‌های تصادفی مستقل با اندازه‌های $n_۱=n_۲=۸$ از جمعیتهای نرمال با واریانسهای یکسان به دست آمده باشند، احتمال اینکه یکی از این دو واریانس نمونه حداقل ۷ برابر بزرگتر از دیگری باشد را بیابید.

۳۳ اگر $S_۱^۲$ و $S_۲^۲$ به ترتیب واریانسهای دو نمونه تصادفی مستقل از دو جمعیت نرمال باشند و بدانیم که واریانس جمعیت دوم سه برابر واریانس جمعیت اول است و نمونه‌هایی به اندازه $n_۱=۸$ و $n_۲=۱۲$ انتخاب شده باشند، مطلوب است محاسبه $P(S_۱<\sqrt{۱/۶۳}\,S_۲)$

فصل هفتم

نظریه برآوردیابی

۱.۷ استنباط آماری

همانگونه که در فصل ششم اشاره شد هدف از یک بررسی آماری، جمع آوری و تنظیم اطلاعات از قسمتی از جمعیت و تجزیه و تحلیل و نتیجه‌گیری از روی آنها در مورد کل جمعیت می‌باشد. این تجزیه و تحلیل و نتیجه‌گیری را استنباط آماری می‌گویند.

تعریف ۱.۷ استنباط آماری روشی است که بوسیله آن براساس نتایج حاصله از نمونه انتخابی از جمعیت در مورد کل جمعیت یا پارامترهای ناشناخته جمعیت نتیجه گیریهایی انجام می‌گیرد.

استنباط آماری دارای دو شاخه مهم زیر است

۱ـ برآورد پارامترهای مجهول جمعیت،

۲ـ آزمون فرضهای آماری در مورد پارامترهای مجهول جمعیت.

برآورد پارامتر مجهول جمعیت خود نیز به دو روش انجام می‌گیرد الف – برآورد نقطه‌ای، ب – برآورد فاصله‌ای. برای روشن شدن موضوع به مثال زیر توجه کنید.

مثال ۱.۱.۷ الف – یک بازرس می‌خواهد متوسط میزان پروتئین موجود در یک غذای کنسرو شده را تعیین کند. برای این منظور او یک نمونه n تایی از این غذای کنسرو شده را جمع آوری می‌کند و با اندازه‌گیری میزان پروتئین موجود در آنها و محاسبه متوسط این مقادیر می‌خواهد متوسط میزان پروتئین در کل کنسروها را تعیین کند. این عمل را برآوردیابی گویند.

ب- دو شرکت A و B یک نوع غذای کنسرو شده را تولید می‌کنند. شرکت A مدعی است که میزان پروتئین موجود در کنسروهای این شرکت از شرکت B بیشتر است. این ادعای شرکت A را یک فرض آماری گویند. با جمع آوری دو نمونه از این دو شرکت و مقایسه آنها ادعای شرکت A را رد یا قبول می‌کنیم که این عمل را آزمون فرض آماری گویند.

در این فصل به مبحث برآورد پارامترهای مجهول جمعیت خواهیم پرداخت و در فصل بعد مبحث آزمون فرضهای آماری را در نظر می‌گیریم.

۲.۷ برآورد پارامتر مجهول جمعیت

فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی n تایی با مقادیر مشاهده شده x_1، x_2، ... و x_n از جمعیت X با توزیع احتمال $f_\theta(x)$ باشند که این توزیع احتمال به پارامتر مجهول θ بستگی دارد. هدف از برآوردیابی یافتن کمیتی از روی مقادیر مشاهده شده x_1، x_2، ... و x_n بعنوان تقریبی از پارامتر مجهول θ می‌باشد. این عمل به دو روش انجام می‌گیرد.

الف- برآورد نقطه‌ای در این روش از روی مقدار مشاهده شده x_1، x_2، ... و x_n یک نمونه تصادفی، تنها مقدار مشاهده شده یک آماره را بعنوان تقریبی از پارامتر مجهول جمعیت ارایه می‌دهیم. اگر بر اساس نمونه تصادفی X_1، X_2، ... و X_n آماره مورد نظر ما برای تخمین پارامتر مجهول $T=T(X_1,\dots,X_n)$ باشد آنگاه متغیر تصادفی $T=T(X_1,\dots,X_n)$ را یک **برآوردگر**[1] پارامتر θ و عدد $t=T(x_1,\dots,x_n)$ را یک **برآورد**[2] پارامتر θ گویند. برای مثال برای به دست آوردن متوسط قد افراد یک اداره اگر یک نمونه تصادفی چهار تایی X_1، X_2، X_3 و X_4 با مقادیر مشاهده شده $x_1=172$، $x_2=168$، $x_3=165$ و $x_4=175$ را جمع آوری کنیم آنگاه آماره $\overline{X}=\frac{1}{4}\sum_{i=1}^{4} X_i$ یک برآوردگر میانگین قد افراد اداره یعنی μ است و مقدار مشاهده شده آن یعنی $\overline{x}=\frac{1}{4}\sum_{i=1}^{4} x_i=170$ یک برآورد نقطه‌ای μ است.

چون یک برآوردگر تابعی از نمونه تصادفی است بنابراین برای یک پارامتر مجهول می‌توان برآوردگرهای زیادی را معرفی کرد. حال این سئوال مطرح می‌شود که در بین

1- Estimator
2- Estimate

برآوردگرهای یک پارامتر کدامیک بهترین است. در پاسخ به این سئوال به تعاریف زیر توجه کنید.

تعریف ۲.۷ برآوردگر $T=T(X_1,\dots,X_n)$ را یک برآوردگر ناریب برای θ گویند هر گاه

$$\boxed{E(T)=\theta} \qquad (1.7)$$

به عبارت دیگر اگر متوسط مقدار برآوردگر T به ازای نمونه‌های مختلف برابر پارامتر θ باشد، آنگاه T را برای θ ناریب گویند.

مثال ۱.۲.۷ فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی از جمعیتی با میانگین μ و واریانس σ^2 باشند.

الف- تحت چه شرطی برآوردگر $T=\sum_{i=1}^{n} a_iX_i$ که در آن a_i ها مقادیر ثابتی هستند، یک براوردگر ناریب برای μ است.

ب- نشان دهید که واریانس نمونه $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ یک برآوردگر ناریب برای واریانس جمعیت σ^2 است.

حل الف- می‌دانیم که $E(X_i)=\mu\,, i=1,2,\dots,n$ و بایستی $E(T)=\mu$ باشد. بنابراین طبق قوانین امید ریاضی داریم که

$$\mu=E(T)=E\Big(\sum_{i=1}^{n} a_iX_i\Big)=\sum_{i=1}^{n} a_iE(X_i)=\sum_{i=1}^{n} a_i\mu=\mu\Big(\sum_{i=1}^{n} a_i\Big)$$

بنابراین با شرط $\sum_{i=1}^{n} a_i=1$، برآوردگر T برای μ ناریب است. اگر قرار دهیم $a_i=\frac{1}{n}, i=1,\dots,n$ آنگاه $\overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$ برای μ ناریب است.

ب- به تمرین ۲ مراجعه کنید.

تعریف ۳.۷ اگر T_1 و T_2 دو برآوردگر ناریب θ باشند و $Var(T_1)<Var(T_2)$ آنگاه برآوردگر T_1 را کاراتر از برآوردگر T_2 گویند. بنابراین در بین برآوردگرهای ناریب θ آن برآوردگری که دارای کمترین واریانس باشد را کاراترین برآوردگر گویند.

مثال ۲.۲.۷ در مثال ۱.۲.۷ (الف) نشان دهید که در بین تمام برآوردگرهای $T=\sum_{i=1}^{n} a_iX_i$ با شرط $\sum_{i=1}^{n} a_i=1$، برآوردگر $\overline{X}$ دارای کمترین واریانس است.

حل طبق خواص واریانس برای متغیرهای تصادفی مستقل داریم که

$$Var(T)=Var(\sum_{i=1}^{n} a_iX_i)=\sum_{i=1}^{n} a_i^2 Var(X_i)=\sigma^2\sum_{i=1}^{n} a_i^2$$

$$=\sigma^2\sum_{i=1}^{n}(a_i-\frac{1}{n}+\frac{1}{n})^2=\sigma^2\left[\sum_{i=1}^{n}(a_i-\frac{1}{n})^2+\frac{2}{n}\sum_{i=1}^{n}(a_i-\frac{1}{n})+\frac{1}{n}\right]$$

$$=\sigma^2\left[\sum_{i=1}^{n}(a_i-\frac{1}{n})^2+\frac{1}{n}\right]$$

مقدار درون کروشه موقعی کمترین مقدار خود را به دست می‌آورد که $\sum_{i=1}^{n}(a_i-\frac{1}{n})^2=0$ و یا $a_i=\frac{1}{n}, i=1,...,n$، یعنی $\overline{X}$ در بین برآوردگرهای فوق دارای کمترین واریانس است.

ب- برآورد فاصله‌ای در برآورد نقطه‌ای موقعی برآوردگر $T=T(X_1,...,X_n)$ برای پارامتر θ یک برآوردگر خوب است که مقدار مشاهده شده آن یعنی برآورد $t=T(x_1,...,x_n)$ به پارامتر θ نزدیک باشد. اما چون با تغییر نمونه مقدار برآورد تغییر می‌یابد پس برآورد نقطه‌ای دارای خطای زیاد می‌باشد. بنابراین به جای برآورد نقطه‌ای می‌توان از برآورد فاصله که دارای خطای کمتری است استفاده کرد.

در روش برآورد فاصله‌ای، یک فاصله (L,U) از اعداد حقیقی را به عنوان تقریبی از پارامتر مجهول θ ارایه می‌دهیم که این فاصله با یک احتمال زیاد پارامتر θ را در برداشته باشد. این فاصله را فاصله اطمینان گویند و اگر احتمال قرارگرفتن پارامتر θ در این فاصله $1-\alpha$ باشد، آن را یک فاصله اطمینان $100(1-\alpha)\%$ گویند. L را حد پائین فاصله و U را حد بالای فاصله و $1-\alpha$ را ضریب اطمینان فاصله گویند. چگونگی به دست آوردن فاصله اطمینان در مثال زیر تشریح شده است.

مثال ۳.۲.۷ فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی از جمعیت نرمال با میانگین μ و واریانس σ^2 باشند. که در آن σ^2 مقداری معلوم (مثلاً $\sigma^2=4$) ولی μ پارامتر مجهول است. یک برآورد نقطه‌ای و یک فاصله اطمینان $100(1-\alpha)\%$ برای μ پیدا کنید.

حل پارامتر μ میانگین کل جمعیت است بنابراین طبق مثالهای ۱.۲.۷ و ۲.۲.۷ معقول به نظر می‌رسد که آن را با میانگین نمونه $\bar{x}$ برآورد کنیم بنابراین برآوردگر نقطه‌ای μ عبارت است از

$$\hat{\mu}=\overline{X}$$

برای به دست آوردن یک فاصله اطمینان $100(1-\alpha)\%$ برای μ بایستی فاصله (L,U) را به گونه‌ای تعیین کنیم که

$$\boxed{P(L<\mu<U)=1-\alpha} \qquad (۲.۷)$$

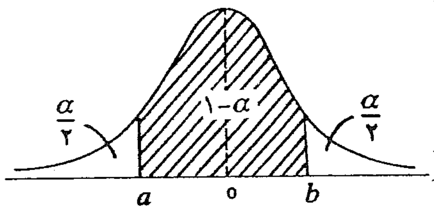
برای این منظور ابتدا تابعی را تعیین می‌کنیم که بستگی به نمونه تصادفی و پارامترهای معلوم و مجهول جمعیت داشته باشد به طوری که توزیع این تابع به پارامترهای مجهول بستگی نداشته باشد. این تابع را **تابع محور** می‌نامند. با استفاده از مطالب بخش ۲.۶ تابع محور مناسب برای این مثال عبارت است از

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

حال برای به دست آوردن فاصله اطمینان برای μ اعداد a<b را به گونه‌ای انتخاب می‌کنیم که

$$P(a<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<b)=1-\alpha$$

با توجه به شکل ۱.۷ نقطه b نقطه‌ای روی محور افقی نمودار تابع چگالی نرمال استاندارد است که مساحت زیر منحنی قبل از این نقطه $1-\frac{\alpha}{2}$ باشد. این نقطه را با $z_{1-\frac{\alpha}{2}}$ نمایش می‌دهیم که از جدول (III) قابل محاسبه است. نقطه a قرینه نقطه b می‌باشد. با قرار دادن این مقادیر a و b در رابطه فوق داریم که



شکل ۱.۷

$$P(-z_{1-\frac{\alpha}{2}}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_{1-\frac{\alpha}{2}})=1-\alpha$$

با حل نامساوی درون پرانتز نسبت به پارامتر μ داریم که

$$P(\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})=1-\alpha$$

با مقایسه این رابطه با رابطه (۲.۷) نتیجه می‌شود که یک فاصله اطمینان $100(1-\alpha)\%$ برای میانگین μ عبارت است از

$$\mu\in\left(\overline{X}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{X}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$

در بخشهای بعد با استفاده از مطالب و مثالهای این بخش پارامترهای مختلف جمعیت را برآورد خواهیم کرد.

۳.۷ برآورد میانگین جمعیت

فرض کنید که از یک جمعیت X با میانگین μ و واریانس σ^2 یک نمونه تصادفی به اندازه n انتخاب کرده باشیم و بخواهیم میانگین جمعیت یعنی μ را برآورد کنیم. طبق مطالب بخش قبل بهترین برآوردگر نقطه‌ای μ عبارت است از $\hat{\mu}=\overline{X}$. برای به دست آوردن یک فاصله اطمینان برای μ دو حالت زیر را در نظر می‌گیریم.

الف- واریانس جمعیت معلوم است در این حالت اگر جمعیت نرمال باشد آنگاه یک تابع محور مناسب عبارت است از

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1) \qquad (۳.۷)$$

و بنابراین با انجام عملیات مشابه مثال ۳.۲.۷ داریم که

> یک فاصله اطمینان $100(1-\alpha)$٪ برای میانگین جمعیت نرمال μ موقعی که واریانس σ^2 معلوم است عبارت است از
>
> $$\mu\in\left(\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$$
>
> که در آن $\bar{x}$ میانگین نمونه تصادفی و $z_{1-\frac{\alpha}{2}}$ مقدار متغیر نرمال استاندارد است که سطح زیر منحنی در سمت چپ آن $1-\frac{\alpha}{2}$ باشد.

توجه کنید که اگر جمعیت نرمال نباشد ولی $n\geq 30$ باشد آنگاه رابطه (۳.۷) هنوز به طور تقریبی برقرار است و در نتیجه فاصله اطمینان فوق یک فاصله اطمینان $100(1-\alpha)$٪ برای μ است.

ب- واریانس جمعیت نامعلوم است در این حالت اگر جمعیت نرمال باشد آنگاه طبق مطالب بخش ۴.۶ یک تابع محور مناسب برای ساختن فاصله اطمینان برای μ عبارت است از

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t_{(n-1)} \qquad (۴.۷)$$

حال با استفاده از این تابع محور و انجام عملیات مشابه با مثال ۳.۲.۷ و جایگزینی s به جای σ و $t_{1-\frac{\alpha}{2}}(n-1)$ به جای $z_{1-\frac{\alpha}{2}}$ داریم که

> یک فاصله اطمینان $100(1-\alpha)$٪ برای میانگین جمعیت نرمال μ موقعی که واریانس σ^2 نامعلوم است عبارت است از
>
> $$\mu\in\left(\bar{x}-t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}\ ,\ \bar{x}+t_{1-\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}\right)$$
>
> که در آن $\bar{x}$ و s به ترتیب میانگین و انحراف استاندارد نمونه تصادفی n تایی و $t_{1-\frac{\alpha}{2}}(n-1)$ مقدار متغیر $t_{(n-1)}$ است که سطح زیرمنحنی در سمت چپ آن $1-\frac{\alpha}{2}$ باشد.

توجه کنید که برای محاسبه s^2 می‌توان از فرمول زیر استفاده کرد.

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right] \qquad (۵.۷)$$

همچنین توجه کنید که اگر $n>30$ باشد آنگاه $t_{1-\frac{\alpha}{2}}(n-1)=z_{1-\frac{\alpha}{2}}$ و بنابراین در حالتی که σ نامعلوم و $n>30$ باشد می‌توان از فاصله اطمینان حالت (الف) با قرار دادن s به جای σ استفاده کرد.

مثال ۱.۳.۷ یک تولیدکننده لامپهای روشنایی، لامپهایی را طول می‌کند که انحراف معیار طول عمر آنها ۴۰ ساعت است. اگر یک نمونه تصادفی ۳۶ تایی دارای حد متوسط عمر ۸۷۰ ساعت باشند، یک برآورد نقطه‌ای و یک فاصله اطمینان ۹۵ درصدی برای حد متوسط عمر تمام لامپهای تولیدی این کارخانه را به دست آورید.

حل در این مثال $\sigma=40$، $n=36$، $\bar{x}=870$ و $1-\alpha=0/95$ می‌باشد بنابراین برآورد نقطه‌ای عبارت است از $\hat{\mu}=\bar{x}=870$. برای به دست آوردن فاصله اطمینان داریم که

$$1-\alpha=0/95 \Rightarrow 1-\frac{\alpha}{2}=0/975$$

و با استفاده از جدول (III) $z_{1-\frac{\alpha}{2}}=z_{0/975}=1/96$، بنابراین

$$\mu \in (\bar{x} - z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}} , \bar{x} + z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}})$$

$$=(۸۷۰ - ۱/۹۶\frac{۴۰}{\sqrt{۳۶}} , ۸۷۰ + ۱/۹۶\frac{۴۰}{\sqrt{۳۶}})=(۸۵۶/۹۳ , ۸۸۳/۰۷)$$

یعنی ۹۵ درصد اطمینان داریم که حد متوسط عمر تمام لامپهای تولیدی این کارخانه در فاصله فوق قرار دارد.

مثال ۲.۳.۷ اگر طول قد کارمندان یک اداره دارای توزیع نرمال باشد، یک فاصله اطمینان ۹۵ درصدی برای میانگین طول قد کارمندان این اداره پیدا کنید در حالیکه یک نمونه ۵ تایی از بین کارمندان انتخاب شده باشد و مقادیر ۱۶۰، ۱۷۰، ۱۶۵، ۱۷۵ و ۱۸۰ به دست آمده باشد.

حل در این مثال σ نامعلوم است و $n=۵$، $\sum_{i=۱}^{۵} x_i = ۸۵۰$، $\sum_{i=۱}^{۵} x_i^۲ = ۱۴۴۷۵۰$ و $۱-\alpha=۰/۹۵$ می‌باشد، بنابراین

$$\bar{x} = \frac{۱}{۵}\sum_{i=۱}^{۵} x_i = \frac{۸۵۰}{۵} = ۱۷۰$$

$$s^۲ = \frac{۱}{۵-۱}\left[\sum_{i=۱}^{۵} x_i^۲ - \frac{۱}{۵}(\sum_{i=۱}^{۵} x_i)^۲\right] = \frac{۱}{۴}\left[۱۴۴۷۵۰ - \frac{(۸۵۰)^۲}{۵}\right] = ۶۲/۵$$

و با استفاده از جدول (V) داریم که $t_{۱-\frac{\alpha}{۲}}(n-۱)=t_{۰/۹۷۵}(۴)=۲/۷۸$ در نتیجه

$$\mu \in (\bar{x} - t_{۱-\frac{\alpha}{۲}}\frac{s}{\sqrt{n}} , \bar{x} + t_{۱-\frac{\alpha}{۲}}\frac{s}{\sqrt{n}})$$

$$=(۱۷۰ - ۲/۷۸\sqrt{\frac{۶۲/۵}{۵}} , ۱۷۰ + ۲/۷۸\sqrt{\frac{۶۲/۵}{۵}})=(۱۶۰/۱۷ , ۱۷۹/۸۳)$$

یعنی ۹۵ درصد اطمینان داریم که میانگین طول قد کارمندان این اداره در فاصله فوق قرار دارد.

توجه کنید به کمک بعضی از ماشینهای محاسبه می‌توان به راحتی مقادیر $\bar{x}$ و s را به طور مستقیم محاسبه نمود.

خطای برآورد میانگین چون اغلب مقدار برآورد نقطه‌ای $\bar{x}$ دقیقاً مساوی μ نیست بنابراین برآورد نقطه‌ای دارای خطا است. با استفاده از حدود فاصله اطمینان می‌توان میزان این خطا یعنی $|\bar{x}-\mu|$ را در دو حالت زیر تعیین کرد.

الف - اگر واریانس $\sigma^۲$ معلوم باشد آنگاه

$$-z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}} \Rightarrow |\bar{x} - \mu| < z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}}$$

ب - اگر واریانس $\sigma^۲$ نامعلوم باشد آنگاه

$$-t_{۱-\frac{\alpha}{۲}}(n-۱)\frac{s}{\sqrt{n}} < \bar{x} - \mu < t_{۱-\frac{\alpha}{۲}}(n-۱)\frac{s}{\sqrt{n}} \Rightarrow |\bar{x} - \mu| < t_{۱-\frac{\alpha}{۲}}(n-۱)\frac{s}{\sqrt{n}}$$

بنابراین اگر $\bar{x}$ را به عنوان برآورد نقطه‌ای μ به کار ببریم آنگاه $۱۰۰(۱-\alpha)$٪ مطمئن هستیم که

الف - خطای برآورد کمتر از $z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}}$ است اگر σ معلوم باشد.

ب - خطای برآورد کمتر از $t_{۱-\frac{\alpha}{۲}}(n-۱)\frac{s}{\sqrt{n}}$ است اگر σ معلوم نباشد.

برای مثال در مثال ۱.۳.۷ چون $z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}} = ۱/۹۶\frac{۴۰}{\sqrt{۳۶}} = ۱۳/۰۷$

بنابراین ۹۵ درصد مطمئن هستیم که خطای برآورد میانگین از ۱۳/۰۷ کمتر است.

تعیین اندازه نمونه در یک بررسی آماری یکی از مهمترین مراحل، تعیین اندازه نمونه n قبل از عمل نمونه‌گیری می‌باشد. اگر یک حداکثر مقدار خطای e برای برآورد میانگین μ برای نمونه‌گیر قابل تحمل باشد آنگاه بوسیله خطای برآورد می‌توان اندازه نمونه n را در دو حالت زیر تعیین کرد.

الف - اگر واریانس $\sigma^۲$ معلوم باشد آنگاه

$$z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}} \le e \Rightarrow n \ge (z_{۱-\frac{\alpha}{۲}}\sigma/e)^۲$$

ب - اگر واریانس $\sigma^۲$ نامعلوم باشد آنگاه

$$t_{۱-\frac{\alpha}{۲}}(n-۱)\frac{s}{\sqrt{n}} \le e \Rightarrow n \ge (t_{۱-\frac{\alpha}{۲}}(n-۱)s/e)^۲$$

که در این حالت ابتدا به وسیله یک نمونه مقدماتی مقداری برای s به دست می‌آوریم و سپس اندازه نمونه واقعی n را از فرمول فوق محاسبه می‌کنیم. در هر دو حالت n را برابر کوچکترین عدد صحیح که در نامساویهای فوق صدق کند انتخاب می‌کنیم. بنابراین

اگر $\bar{x}$ را به عنوان برآورد نقطه‌ای μ به کار ببریم آنگاه $(1-\alpha)100\%$ مطمئن هستیم که خطای برآورد از مقدار مشخص e کمتر است موقعی که اندازه نمونه از رابطه زیر محاسبه گردد

الف - اگر واریانس σ^2 معلوم باشد $\quad n \ge (z_{1-\frac{\alpha}{2}}\sigma/e)^2$

ب - اگر واریانس σ^2 نامعلوم باشد $\quad n \ge (t_{1-\frac{\alpha}{2}}(n-1)s/e)^2$

مثال ۳.۳.۷ در مثال ۱.۳.۷ اگر بخواهیم با اطمینان ۹۵ درصد خطای برآورد حد متوسط طول عمر لامپها از ۱۰ ساعت کمتر باشد، اندازه نمونه را تعیین کنید.

حل
$$n\ge(z_{1-\frac{\alpha}{2}}\sigma/e)^2=\left(\frac{1/96\times 40}{10}\right)^2=61/47$$

بنابراین اندازه نمونه بایستی n=۶۲ باشد.

مثال ۴.۳.۷ در مثال ۲.۳.۷ اگر بخواهیم با اطمینان ۹۵ درصد خطای برآورد میانگین طول قد کارمندان اداره از ۵ سانتیمتر کمتر باشد، اندازه نمونه را تعیین کنید.

حل اگر داده‌های مثال ۲.۳.۷ را به عنوان یک نمونه مقدماتی در نظر بگیریم آنگاه $s^2=62/5$ و در نتیجه

$$n\ge(t_{1-\frac{\alpha}{2}}(n-1)s/e)^2=\left(\frac{2/78}{5}\right)^2\times 62/5=19/321$$

بنابراین اندازه نمونه بایستی n=۲۰ باشد.

۴.۷ برآورد واریانس جمعیت

فرض کنید که از یک جمعیت نرمال با میانگین μ و واریانس σ^2 یک نمونه تصادفی به اندازه n انتخاب کرده باشیم و بخواهیم واریانس جمعیت یعنی σ^2 را برآورد کنیم. بهترین برآوردگر نقطه‌ای σ^2 عبارت است از واریانس نمونه $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ یعنی $\hat{\sigma}^2=S^2$

برای به دست آوردن یک فاصله اطمینان برای σ^2، طبق مطالب بخش ۳.۶ تابع محور مناسب عبارت است از

$$Y=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{(n-1)} \qquad (6.7)$$

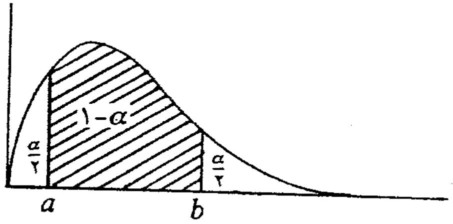
حال با استفاده از این تابع محور، اعداد a و b را چنان تعیین می‌کنیم که

$$P\left(a<\frac{(n-1)S^2}{\sigma^2}<b\right)=1-\alpha$$

با توجه به شکل ۲.۷ مقادیر a و b عبارتند از

$$a=\chi^2_{\frac{\alpha}{2}}(n-1)\ ,\ b=\chi^2_{1-\frac{\alpha}{2}}(n-1)$$

با قرار دادن این مقادیر در رابطه فوق و انجام عملیات مشابه مثال ۳.۲.۷ داریم که



شکل ۲.۷

یک فاصله اطمینان $(1-\alpha)100\%$ برای واریانس جمعیت نرمال σ^2 عبارت است از

$$\sigma^2\in\left(\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\ ,\ \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)$$

که در آن s^2 واریانس یک نمونه تصادفی nتایی است.

مثال ۱.۴.۷ طول یک لوله ساختمانی دارای توزیع نرمال با میانگین μ و واریانس σ^2 است. یک نمونه تصادفی ۲۵ تایی از لوله‌ها جمع آوری شده است و مقادیر $\sum_{i=1}^{25}x_i=252/7$ و $\sum_{i=1}^{25}x_i^2=2579/7$ حاصل شده است. یک برآورد نقطه‌ای و یک فاصله اطمینان ۹۰ درصدی برای واریانس واقعی لوله‌ها به دست آورید.

حل با توجه به مقادیر داده شده و رابطه (۵.۷) داریم که

$$s^2=\frac{1}{25-1}\left[2579/7-\frac{(252/7)^2}{25}\right]=1/06$$

بنابراین برآورد نقطه‌ای σ^2 عبارت است از $\quad \hat{\sigma}^2=s^2=1/06$

همچنین با استفاده از جدول (IV) داریم که

$$1-\alpha=0/90 \Rightarrow \begin{cases}\frac{\alpha}{2}=0/05\\ 1-\frac{\alpha}{2}=0/95\end{cases} \Rightarrow \begin{cases}\chi^2_{\frac{\alpha}{2}}(n-1)=\chi^2_{0/05}(24)=13/8\\ \chi^2_{1-\frac{\alpha}{2}}(n-1)=\chi^2_{0/95}(24)=36/4\end{cases}$$

بنابراین

$$\sigma^2\in\left(\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)},\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)=\left(\frac{24\times 1/06}{36/4},\frac{24\times 1/06}{13/8}\right)=(0/699,\ 1/843)$$

یعنی ۹۰ درصد اطمینان داریم که واریانس واقعی لوله‌ها در فاصله فوق قرار دارد.

۵.۷ برآورد تفاضل میانگین دو جمعیت

فرض کنید دو جمعیت داریم که جمعیت اول دارای میانگین μ_1 و واریانس σ_1^2 و جمعیت دوم دارای میانگین μ_2 و واریانس σ_2^2 باشد. یک نمونه تصادفی n_1 تایی از جمعیت اول انتخاب کرده و میانگین نمونه آن را $\bar{X}_1$ و واریانس نمونه آن را S_1^2 می‌نامیم و یک نمونه تصادفی n_2 تایی از جمعیت دوم انتخاب کرده و میانگین نمونه آن را $\bar{X}_2$ و واریانس نمونه آن را S_2^2 می‌نامیم. همچنین فرض کنید که این دو نمونه تصادفی از یکدیگر مستقل باشند. می‌خواهیم اختلاف میانگین دو جمعیت یعنی $\mu_1-\mu_2$ را برآورد کنیم. بهترین برآوردگر نقطه‌ای $\mu_1-\mu_2$ عبارت از اختلاف میانگینهای دو نمونه می‌باشد یعنی $\mu_1\hat{-}\mu_2=\bar{X}_1-\bar{X}_2$

برای به دست آوردن فاصله اطمینان برای $\mu_1-\mu_2$ دو حالت زیر را در نظر می‌گیریم.

الف- واریانس دو جمعیت معلوم باشند در این حالت σ_1^2 و σ_2^2 مقادیر معلومی هستند. اگر دو جمعیت نرمال باشند، با توجه به مطالب بخش ۵.۶ و رابطه (۵.۶) تابع محور مناسب عبارت است از

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1) \tag{7.7}$$

حال با استفاده از این تابع محور و انجام عملیات مشابه مثال ۳.۲.۷ به دست می‌آوریم که

> یک فاصله اطمینان $100(1-\alpha)\%$ برای تفاضل میانگین دو جمعیت نرمال که واریانسهای آنها معلوم است عبارت است از
>
> $$\mu_1-\mu_2\in\left(\bar{x}_1-\bar{x}_2-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ \bar{x}_1-\bar{x}_2+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$
>
> که در آن $\bar{x}_1$ و $\bar{x}_2$ به ترتیب میانگینهای نمونه‌های تصادفی n_1تایی و n_2تایی از دو جمعیت می‌باشند.

توجه کنید که اگر دو جمعیت نرمال نباشند اما $n_1\geq 30$ و $n_2\geq 30$ باشد آنگاه طبق مطالب بخش ۵.۶، رابطه (۷.۷) هنوز به طور تقریبی برقرار است و بنابراین فاصله اطمینان فوق نیز در این حالت یک فاصله اطمینان $100(1-\alpha)\%$ برای $\mu_1-\mu_2$ است.

ب- واریانس دو جمعیت نامعلوم اما مساوی باشند در این حالت $\sigma_1^2=\sigma_2^2=\sigma^2$ می‌باشد و بایستی ابتدا σ^2 یعنی واریانس مشترک دو جمعیت را با واریانس مشترک دو نمونه به صورت زیر برآورد کنیم.

$$\hat{\sigma}^2=S_P^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} \tag{8.7}$$

اگر دو جمعیت نرمال باشند آنگاه طبق مطالب بخش ۵.۶ و رابطه (۷.۶) تابع محور مناسب عبارت است از

$$T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{(n_1+n_2-2)}$$

حال با استفاده از این تابع محور و انجام عملیات مشابه مثال ۳.۲.۷ به دست می‌آوریم که

> یک فاصله اطمینان $100(1-\alpha)\%$ برای تفاضل میانگین دو جمعیت نرمال که واریانسهای آنها نامعلوم اما مساوی هستند عبارت است از
>
> $$\mu_1-\mu_2\in$$
> $$\left(\bar{x}_1-\bar{x}_2-t_{1-\frac{\alpha}{2}}(n_1+n_2-2)s_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}_1-\bar{x}_2+t_{1-\frac{\alpha}{2}}(n_1+n_2-2)s_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$
>
> که در آن s_P^2 واریانس مشترک دو نمونه تصادفی n_1 و n_2 تایی از دو جمعیت می‌باشد.

توجه کنید که اگر $n_1\geq 30$ و $n_2\geq 30$ و جمعیتها غیر نرمال با واریانسهای نامعلوم باشند آنگاه می‌توان از فاصله اطمینان قسمت (الف) با قرار دادن s_1 و s_2 به جای σ_1 و σ_2 استفاده کرد.

مثال ۱.۵.۷ دو شرکت A و B لامپهای روشنایی تولید می‌کنند که به ترتیب دارای انحراف معیارهای ۲۷ و ۳۱ ساعت است. یک نمونه تصادفی ۴۰ تایی از لامپهای تولیدی هر شرکت انتخاب می‌کنیم و به ترتیب متوسط طول عمر لامپها برای این دو نمونه را ۶۴۹ و ۶۳۵ به دست می‌آوریم. با

ساختن یک فاصله اطمینان ۹۶ درصدی برای اختلاف متوسط طول عمر لامپهای دو شرکت، چه نتیجه‌ای به دست می‌آورید.

حل در این مثال $\sigma_۱=۲۷$، $\sigma_۲=۳۱$، $n_۱=n_۲=۴۰$، $\bar{x}_۱=۶۴۹$، $\bar{x}_۲=۶۳۵$ و $z_{۱-\frac{\alpha}{۲}}=z_{۰/۹۸}=۲/۰۵۴$. بنابراین طبق حالت (الف) داریم که

$$\mu_۱-\mu_۲\in$$
$$\left(۶۴۹-۶۳۵-۲/۰۵۴\sqrt{\frac{(۲۷)^۲}{۴۰}+\frac{(۳۱)^۲}{۴۰}},\ ۶۴۹-۶۳۵+۲/۰۵۴\sqrt{\frac{(۲۷)^۲}{۴۰}+\frac{(۳۱)^۲}{۴۰}}\right)$$
$$=(۱۴-۱۳/۳۵۱,\ ۱۴+۱۳/۳۵۱)=(۰/۶۴۹,\ ۲۷/۳۵۱)$$

چون تمامی فاصله مثبت است پس ۹۶ درصد اطمینان داریم که متوسط طول عمر لامپهای شرکت A از شرکت B بیشتر است.

مثال ۲.۵.۷ دو آزمایشگاه ۱ و ۲ به طور مستقل برای اندازه‌گیری چربی موجود در شیرهای پاستوریزه اقدام می‌نمایند. هر یک تعدادی نمونه انتخاب کرده و نتایج در جدول زیر ثبت شده است. با فرض نرمال بودن دو جمعیت و مساوی بودن واریانسها، یک فاصله اطمینان ۹۵ درصدی برای اختلاف میانگین چربی موجود در شیرهای دو کارخانه به دست آورید. چه نتیجه‌ای می‌گیرید؟

آزمایشگاه ۱	۳	۵	۷	۳	۸	۶	۸	۹	۴	۷
آزمایشگاه ۲	۹	۸	۸	۴	۷	۶	۸	۶		

حل از جدول فوق مقادیر زیر را به دست می‌آوریم

$$n_۱=۱۰,\ \sum x_{۱i}=۶۰,\ \sum x_{۱i}^۲=۴۰۲,\ \bar{x}_۱=۶,\ s_۱^۲=۴/۶۷$$
$$n_۲=۸,\ \sum x_{۲i}=۵۶,\ \sum x_{۲i}^۲=۴۱۰,\ \bar{x}_۲=۷,\ s_۲^۲=۲/۵۷$$

بنابراین

$$s_P^۲=\frac{(n_۱-۱)s_۱^۲+(n_۲-۱)s_۲^۲}{n_۱+n_۲-۲}=\frac{۹(۴/۶۷)+۷(۲/۵۷)}{۱۶}=۳/۷۵$$

در نتیجه $s_P=۱/۹۳۷$ و همچنین $t_{۱-\frac{\alpha}{۲}}(n_۱+n_۲-۲)=t_{۰/۹۷۵}(۱۶)=۲/۱۲$

بنابراین با استفاده از حالت (ب) داریم که

$$\mu_۱-\mu_۲\in\left(۶-۷-۲/۱۲(۱/۹۳۷)\sqrt{\frac{۱}{۱۰}+\frac{۱}{۸}},\ ۶-۷+۲/۱۲(۱/۹۳۷)\sqrt{\frac{۱}{۱۰}+\frac{۱}{۸}}\right)$$

$$=(-۱-۱/۹۴۸,\ -۱+۱/۹۴۸)=(-۲/۹۴۸,\ ۰/۹۴۸)$$

چون فاصله شامل مقادیر منفی و مثبت است پس ۹۵ درصد اطمینان داریم که میانگین چربی موجود در شیرهای دو کارخانه با یکدیگر مساوی است.

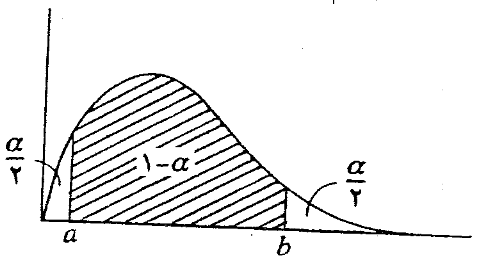
۶.۷ برآورد نسبت واریانس دو جمعیت

دو جمعیت بخش قبل و نمونه‌های تصادفی به دست آمده از آنها را در نظر بگیرید. در این بخش می‌خواهیم نسبت واریانس دو جمعیت یعنی $\frac{\sigma_۲^۲}{\sigma_۱^۲}$ را برآورد کنیم. بهترین برآوردگر نقطه‌ای $\frac{\sigma_۲^۲}{\sigma_۱^۲}$ عبارت از نسبت واریانسهای دو نمونه می‌باشد یعنی

$$\widehat{\frac{\sigma_۲^۲}{\sigma_۱^۲}}=\frac{S_۲^۲}{S_۱^۲}$$

برای به دست آوردن یک فاصله اطمینان برای $\frac{\sigma_۲^۲}{\sigma_۱^۲}$، اگر دو جمعیت نرمال باشند آنگاه طبق مطالب بخش ۶.۶ تابع محور مناسب عبارت است از

$$F=\frac{S_۱^۲}{S_۲^۲}\frac{\sigma_۲^۲}{\sigma_۱^۲}\sim F(n_۱-۱,n_۲-۱) \qquad (۹.۷)$$



شکل ۳.۷

حال با استفاده از این تابع محور و توجه به شکل ۳.۷، با انجام عملیات مشابه مثال ۳.۲.۷ یک فاصله اطمینان $٪۱۰۰(۱-\alpha)$ برای $\frac{\sigma_۲^۲}{\sigma_۱^۲}$ به صورت زیر به دست می‌آید.

$$\frac{\sigma_۲^۲}{\sigma_۱^۲}\in\left(\frac{s_۲^۲}{s_۱^۲}F_{\frac{\alpha}{۲}}(n_۱-۱,n_۲-۱),\ \frac{s_۲^۲}{s_۱^۲}F_{۱-\frac{\alpha}{۲}}(n_۱-۱,n_۲-۱)\right)$$

با توجه به اینکه $F_{\frac{\alpha}{۲}}(n_۱-۱,n_۲-۱)=\frac{۱}{F_{۱-\frac{\alpha}{۲}}(n_۲-۱,n_۱-۱)}$ بنابراین به دست می‌آوریم که

یک فاصله اطمینان $(1-\alpha)100\%$ برای نسبت واریانس دو جمعیت نرمال عبارت است از

$$\frac{\sigma_2^2}{\sigma_1^2}\in\left(\frac{s_2^2}{s_1^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)},\ \frac{s_2^2}{s_1^2}F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right)$$

که در آن s_1^2 و s_2^2 به ترتیب واریانسهای نمونه‌های n_1 و n_2 تایی از دو جمعیت می‌باشند.

مثال ۷.۶.۱ از یک درس یک کلاس صبح و یک کلاس بعد از ظهر تشکیل شده است. در آخر ترم از کلاس صبح ۸ نفر و از کلاس بعدازظهر ۹ نفر به تصادف انتخاب کرده و از آنها امتحان بعمل آورده‌ایم و نتایج در جدول زیر یادداشت گردیده‌اند.

نمرات کلاس صبح	۹	۷	۸	۱۰	۱۲	۱۵	۷	۱۲	
نمرات کلاس بعدازظهر	۳	۷	۹	۲	۱۸	۱۶	۵	۱۱	۱۰

فرض کنید نمرات دو کلاس از یکدیگر مستقل بوده و از توزیع نرمال پیروی کنند. یک فاصله اطمینان ۹۰ درصدی برای نسبت واریانسها و نسبت انحراف معیارهای نمرات دو کلاس به دست آورید. چه نتیجه‌ای می‌گیرید؟

حل از جدول فوق مقادیر زیر به دست می‌آیند

$$n_1=8\ ,\ \sum x_{1i}=80\ ,\ \sum x_{1i}^2=856\ ,\ s_1^2=8$$
$$n_2=9\ ,\ \sum x_{2i}=81\ ,\ \sum x_{2i}^2=969\ ,\ s_2^2=30$$

همچنین از جدول (VI) داریم که

$$1-\alpha=0/90\Rightarrow\begin{cases}F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)=F_{0/95}(7,8)=3/50\\ F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)=F_{0/95}(8,7)=3/73\end{cases}$$

بنابراین فاصله اطمینان ۹۰ درصدی برای $\frac{\sigma_2^2}{\sigma_1^2}$ عبارت است از

$$\frac{\sigma_2^2}{\sigma_1^2}\in\left(\frac{30}{8}\times\frac{1}{3/73},\ \frac{30}{8}\times 3/50\right)=(1/005,\ 13/125)$$

حال اگر از مقادیر این فاصله جذر بگیریم فاصله اطمینان ۹۰ درصدی برای نسبت انحراف معیارها

یعنی $\frac{\sigma_2}{\sigma_1}$ به صورت زیر به دست می‌آید

$$\frac{\sigma_2}{\sigma_1}\in(1/003,\ 3/623)$$

چون این فاصله تماماً از یک بزرگتر است پس ۹۰ درصد اطمینان داریم که $\sigma_2^2>\sigma_1^2$ می‌باشد.

۷.۷ تمرینات

۱ فرض کنید X_1 و X_2 دو برآوردگر نااریب و مستقل پارامتر θ با واریانسهای ۲ و ۳ باشند. اگر $T=a_1X_1+a_2X_2$ آنگاه ضرایب a_1 و a_2 را به گونه‌ای پیدا کنید که T یک برآوردگر نااریب با کمترین واریانس برای θ باشد. با مقایسه واریانس برآوردگرهای X_1 و X_2 و T نتیجه بگیرید که کدامیک از این ۳ برآوردگر نااریب، بهتر می‌باشد.

۲ فرض کنید X_1، X_2، ... و X_n یک نمونه تصادفی n تایی از جمعیت X با میانگین μ و واریانس σ^2 باشند. اگر $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ و $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ به ترتیب میانگین و واریانس این نمونه باشند، مطلوب است

الف - نشان دهید که $E(\overline{X})=\mu$ و $Var(\overline{X})=\frac{\sigma^2}{n}$

ب - نشان دهید که

$$S^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2\right]$$

ج - با استفاده از قسمت (ب) نشان دهید که $E(S^2)=\sigma^2$.

۳ نمونه‌ای به اندازه ۲۵ لامپ روشنایی از یک دسته بزرگ از لامپهای ۴۰واتی گرفته شده است و میانگین عمر لامپهای نمونه ۱۴۱۰ ساعت است. با فرض اینکه عمر لامپها دارای توزیع نرمال با انحراف معیار ۲۰۰ ساعت باشد، یک فاصله اطمینان ۹۵٪ برای میانگین عمر لامپهای این دسته به دست آورید.

۴ الف - از یک جمعیت نرمال با واریانس ۴ یک نمونه تصادفی به اندازه ۲۵ انتخاب کرده‌ایم و میانگین این نمونه ۲۰ شده است. یک فاصله اطمینان ۹۰ درصدی برای میانگین این جمعیت پیدا کنید.

ب - اگر بخواهیم طول فاصله اطمینان را به نصف کاهش دهیم، اندازه نمونه را چه مقدار باید انتخاب کنیم؟

ج- بیشترین خطایی را که ممکن است در تعیین فاصله اطمینان ۹۰ درصدی مرتکب شویم را به دست آورید.

۵ یک نوع خازن الکترونیکی بوسیله یک شرکت ساخته می‌شود و در طی سالها شرکت دریافته است که طول عمر این خازنها دارای توزیع نرمال با انحراف معیار ۲۲۵ ساعت است. میانگین یک نمونه ۳۰ تایی از این خازنها برابر ۱۴۰۷/۵ ساعت است. یک فاصله اطمینان ۹۹ درصدی برای میانگین طول عمر خازنهای این شرکت به دست آورید.

۶ یک نشریه صنعت کامپیوتر می‌خواهد میانگین درآمد سهام عادی را در سال گذشته برای تمام شرکتهای نرم‌افزاری کامپیوتر برآورد کند. نمونه‌ای تصادفی به اندازه ۱۴ شرکت از بین این شرکتها انتخاب شده است و درآمد سهام عادی این شرکتها به شرح زیر می‌باشد.

-۳/۱۵	۱۷/۷۸	-۸/۴۷	۲۴/۸۷	-۶/۵۲	۳۱/۳۸	۸/۳۷
۲۶/۶۸	۲۱/۰۱	۱۱/۲۸	۰/۳۵	۲۴/۸۳	۱۵/۴۵	۷/۲۰

فرض کنید که درآمد سهام عادی این شرکتها از توزیع نرمال پیروی کند. برای میانگین درآمد سهام عادی، یک فاصله اطمینان ۹۰ درصدی به دست آورید.

۷ یک مطالعه آلودگی هوا در یک ایستگاه هواشناسی انجام گرفته است. مقادیر مواد آلی محلول در بنزین که در هوا معلق می‌باشد در هشت نمونه مختلف از هوا به صورت ۲/۲، ۱/۸، ۳/۱، ۲/۰، ۲/۴، ۲/۰، ۲/۱ و ۱/۲ بوده است. فرض کنید جمعیتی که از آن نمونه گرفته شده، نرمال باشد، یک فاصله اطمینان ۹۵ درصدی برای میانگین جمعیت به دست آورید.

۸ قطر ۲۰۰ بولبرینگ ساخت یک کارخانه اندازه‌گیری شده و میانگین ۰/۸۲۴ اینچ و انحراف استاندارد آن ۰/۴۲ اینچ بوده است.

الف- یک فاصله اطمینان ۹۰ درصدی برای میانگین قطر بولبرینگهای ساخت این کارخانه پیدا کنید.

ب- یک فاصله اطمینان ۹۵ درصدی برای واریانس قطر بولبرینگهای ساخت این کارخانه پیدا کنید.

۹ یک نمونه تصادفی ۸تایی از یک نوع سیگار به طور متوسط ۱۸/۶ میلی گرم نیکوتین با انحراف استاندارد ۲/۴ میلی‌گرم دارد. با فرض نرمال بودن میزان نیکوتین در این نوع سیگار، یک فاصله

اطمینان ۹۹ درصدی برای حد متوسط واقعی نیکوتین این نوع سیگار پیدا کنید.

۱۰ یک سازنده ساعت می‌خواهد تغییرپذیری یک نوع ساعت ساخته شده را برآورد کند. برای این منظور یک نمونه از هشت ساعت را برای مدت ۴۸ ساعت به کار می‌اندازد. تعداد ثانیه‌هایی را که هر ساعت پس از ۴۸ ساعت عقب یا جلو افتاده است با یک ساعت دقیق اندازه می‌گیرد و نتایج ۶+، ۴-، ۷-، ۵+، ۹+، ۶-، ۳- و ۲+ را به دست می‌آورد. به فرض اینکه زمانهای ثبت شده برای این نوع ساعت به صورت نرمال توزیع شوند

الف- یک فاصله اطمینان ۹۵ درصدی برای میانگین زمانهای ثبت شده برای این نوع ساعت پس از ۴۸ ساعت به دست آورید.

ب- یک فاصله اطمینان ۹۵ درصدی برای انحراف معیار زمانهای ثبت شده برای این نوع ساعت پس از ۴۸ ساعت به دست آورید.

۱۱ یک نمونه تصادفی ۴۰تایی از شبکه‌های شیمیایی، از بین ۲۰۰ عدد از این شبکه‌ها که انتظار می‌رود انحراف معیار وزنشان $\sigma=$ ۱۲/۲ پوند باشد، استخراج شده و متوسط وزن نمونه ۲۴۰/۸ پوند به دست آمده است. اگر وزن متوسط همه ۲۰۰ شبکه با ۲۴۰/۸ پوند برآورد شود، با اطمینان ۹۹٪ حداکثر مقدار خطای برآورد را بیابید.

۱۲ یک جایگاه سوخت‌گیری برای کامیونها، یادداشتهای انواع داد و ستدها با مشتریانش را نگهداری کرده است. اگر نمونه‌ای از ۱۸ یادداشت میانگین فروشها را ۶۳/۸۴ گالن سوخت دیزل با انحراف استاندارد ۲/۷۵ گالن نشان دهد و مقدار ۶۳/۸۴ به عنوان برآورد متوسط فروش سوخت دیزل نسبت به هر مشتری در جایگاه به کار رود، با اطمینان ۹۹٪ حداکثر مقدار خطای برآورد را پیدا کنید.

۱۳ اگر $\bar{x}$ میانگین نمونه یک نمونه تصادفی به اندازه n از یک جمعیت نرمال با انحراف معیار ۰/۳ باشد، مقدار n را چنان تعیین کنید که فاصله اطمینان ۹۹ درصدی برای میانگین جمعیت به صورت $(\bar{x}-0/1, \bar{x}+0/1)$ باشد.

۱۴ یک متخصص مجرب می‌خواهد معلوم کند که به طور متوسط چه مدت برای کندن ۳ سوراخ در یک فلز معین لازم است. تعداد نمونه مورد نیاز برای رسیدن به اطمینان ۹۵٪ جهت اینکه میانگین نمونه او در فاصله ۱۵ ثانیه از حد متوسط واقعی قرار گیرد را پیدا نمائید. فرض کنید از مطالعات

قبلی $\sigma=۴۰$ را به دست آورده‌ایم.

۱۵ رئیس دانشکده‌ای می‌خواهد از میانگین نمونه تصادفی برای برآورد متوسط زمانی که طول می‌کشد، دانشجویان از کلاسی به کلاس بعدی بروند، استفاده کند و مایل است با اطمینان ۹۹٪ ادعا کند که خطای ماکزیمم ۰/۲۵ دقیقه است. اگر او از تجربه گذشته فرض کند که $\sigma=۱/۴۰$ دقیقه است، اندازه نمونه را چقدر باید انتخاب کند.

۱۶ پژوهشگری پخش حرارت فلزات را مطالعه می‌کند، به این ترتیب که یک سرمیله فلزی را با یک منبع استاندارد به مدت ۱۰ ثانیه حرارت می‌دهد و آنگاه درجه حرارت سر دیگر میله را اندازه‌گیری می‌کند. درجه‌های حرارتی که برای ۵ فلز مورد آزمایش به دست آمده‌اند عبارت از ۲۱۶، ۲۲۱، ۲۲۰، ۲۱۸ و ۲۲۵ می‌باشد. اگر نمونه مشاهده شده از یک جمعیت نرمال باشند، یک فاصله اطمینان ۹۹ درصدی برای σ^2 و یک فاصله اطمینان ۹۹ درصدی برای σ به دست آورید.

۱۷ در تمرین ۹ یک فاصله اطمینان ۹۹ درصدی برای انحراف معیار واقعی نیکوتین این نوع سیگار به دست آورید.

۱۸ انحراف استاندارد طول عمر یک نمونه ۲۰۰ تایی از لامپهای الکتریکی ساخت یک کارخانه ۱۰۰ ساعت است. برای انحراف معیار طول عمر لامپهای ساخت این کارخانه یک فاصله اطمینان ۹۵ درصدی و ۹۹ درصدی به دست آورید.

۱۹ یک نمونه تصادفی ۱۵ تایی از یک جمعیت نرمال انتخاب و نتایج زیر به دست آمده است. یک فاصله اطمینان ۹۹ درصدی برای σ^2 به دست آورید.

۷/۲ ۴/۳ ۱۱/۴ ۱۴/۳ ۵/۶ ۶/۶ ۱۲/۷

۶/۱ ۱۴ ۷/۱ ۱۲/۸ ۱۰ ۱۱/۲ ۱۳/۸ ۱۰/۸

۲۰ صفحه‌های پلاستیکی که توسط یک ماشین تولید می‌شوند به طور متناوب مورد بازبینی قرار می‌گیرند تا تفاوتهای ضخامت آنها بررسی گردد. ناهمگنی در غلظت ماده‌ای که به کار می‌رود، وجود تفاوتهایی در ضخامت صفحه‌ها را غیرقابل اجتناب می‌سازد. در ده صفحه تولید شده در یک نوبت کاری، اندازه‌های ضخامت برحسب میلیمتر به قرار زیر بوده است

۲۲۵، ۲۳۲، ۲۲۸، ۲۲۷، ۲۲۹، ۲۲۵، ۲۳۰، ۲۲۶، ۲۲۸، ۲۲۶

یک فاصله اطمینان ۹۵ درصدی برای انحراف معیار واقعی ضخامت صفحه‌های تولید شده در این

نوبت کاری بسازید.

۲۱ میانگین یک نمونه ۱۰تایی از $N(\mu_1, ۲۲۵)$ برابر ۱۷۰/۲ و میانگین یک نمونه ۱۲تایی از $N(\mu_2, ۲۲۶)$ که مستقل از اولی است برابر ۱۷۶/۷ به دست آمده است. یک فاصله اطمینان ۹۵ درصدی برای $\mu_1-\mu_2$ به دست آورید.

۲۲ یک نمونه تصادفی ۲۵ تایی از یک جمعیت نرمال با انحراف معیار ۵ دارای میانگین ۸۰ است. یک نمونه تصادفی دیگر ۳۶ تایی از جمعیت نرمال دیگری با انحراف معیار ۳ دارای میانگین ۷۵ است. یک فاصله اطمینان ۹۴ درصدی برای تفاضل میانگین این دو جمعیت به دست آورید. چه نتیجه‌ای می‌گیرید؟

۲۳ یک نمونه تصادفی ۱۵۰ تایی از لامپهای نوع A دارای میانگین ۱۴۰۰ ساعت و انحراف معیار ۱۲۰ ساعت است. یک نمونه تصادفی ۲۰۰ تایی از لامپهای نوع B دارای میانگین ۱۲۰۰ ساعت و انحراف معیار ۸۰ ساعت است. یک فاصله اطمینان ۹۵ درصدی و ۹۹درصدی برای اختلاف میانگین دو نوع لامپ به دست آورید. چه نتیجه‌ای می‌گیرید؟

۲۴ دو ماشین A و B جعبه‌های ۸ گرمی از یک ماده را بسته بندی می‌کنند. با استفاده از تجربیات گذشته با این ماشینها می‌پذیریم که انحراف معیار وزن بسته‌های پر شده به وسیله ماشین A و B به ترتیب ۰/۰۴ و ۰/۰۵ گرم است. از جعبه‌های پرشده هر یک از ماشینها صد جعبه به تصادف انتخاب شده و نتایج زیر به دست آمده‌اند.

ماشین A : $n_A=۱۰۰$, $\bar{x}_A=۸/۱۸$

ماشین B : $n_B=۱۰۰$, $\bar{x}_B=۸/۱۵$

برای تفاضل میانگین وزنهای پر شده با دو ماشین A و B یک فاصله اطمینان ۹۹ درصدی به دست آورید.

۲۵ داده‌های زیر نمونه‌ای از نمرات امتحان آمار دو گروه از دانشجویان می‌باشد.

گروه ۱	۲۲/۵	۲۵/۰	۳۰/۰	۲۷/۵	۲۰/۰
گروه ۲	۲۱/۵	۱۷/۵	۱۷/۰	۲۰/۰	

فرض کنید نمرات این دو گروه از یکدیگر مستقل بوده و هر یک دارای توزیع نرمال با واریانسهای مساوی باشند.

الف – یک فاصله اطمینان ۹۰ درصدی برای میانگین نمرات امتحان آمار گروه دوم تشکیل داده و آن را تعبیر کنید.

ب – با تشکیل یک فاصله اطمینان ۹۵ درصدی برای تفاضل میانگین نمرات دو گروه، نتیجه بگیرید که کدام گروه دارای میانگین نمره بهتری است.

۲۶ میزان تقاضا بر حسب کیلوگرم برای ۲ محصول x و y دارای توزیع نرمال با واریانسهای یکسان می‌باشد. میزان تقاضا برای ۶ روز این دو محصول را جمع‌آوری و در جدول زیر ثبت کرده‌ایم

x	۳۰	۲۶	۳۰	۱۹	۲۶	۳۱
y	۴۰	۳۴	۳۰	۲۹	۳۵	۴۲

آیا با اطمینان ۹۵ درصد می‌توان گفت که میانگین تقاضا برای محصول y بیشتر از میانگین تقاضا برای محصول x است؟

۲۷ یک نمونه ۱۵تایی از نوشابه‌های ساخت کارخانه A دارای میانگین ۶۰ و انحراف استاندارد ۳ می‌باشد. نمونه ۲۱تایی از نوشابه‌های ساخت کارخانه B دارای میانگین ۵۸ و انحراف استاندارد ۲ است. فرض کنید دو جمعیت نرمال با واریانسهای مساوی باشند. یک فاصله اطمینان ۹۵ درصدی برای $\mu_A-\mu_B$ تعیین کنید.

۲۸ مفروضات زیر در رابطه با زمان نمایش فیلمهای تولید شده توسط دو شرکت تولید کننده فیلم است

شرکت ۱	۱۰۳	۹۴	۱۱۰	۶۷	۹۸		
شرکت ۲	۹۷	۸۲	۱۲۳	۹۲	۱۷۵	۸۸	۱۱۸

یک فاصله اطمینان ۹۰ درصدی برای اختلاف میانگین زمان نمایش فیلم دو شرکت به دست آورید. فرض کنید که توزیع زمان نمایش فیلم دو شرکت نرمال با واریانسهای یکسان باشد.

۲۹ یک شرکت سازنده لامپ روشنایی ده تا از لامپهای حاوی سیم نوع A و ده تا از لامپهای دارای سیم نوع B را آزمایش می‌کند. نتیجه‌های زیر برای طول عمر این ۲۰ لامپ برحسب ساعت به دست آمده‌اند

سیم A	۱۳۴۰	۱۴۹۷	۱۶۱۴	۱۳۸۰	۱۲۹۳	۱۰۲۸	۱۲۷۰	۱۰۹۴	۱۴۶۶	۱۶۴۳
سیم B	۱۰۹۲	۱۳۸۳	۱۰۶۵	۱۶۲۷	۱۰۶۱	۱۱۴۳	۱۰۱۷	۱۱۳۸	۱۰۲۱	۱۷۱۱

به فرض نرمال بودن جمعیتها و $\sigma_A^2=\sigma_B^2=\sigma^2$، برای تفاضل بین متوسط عمر لامپهای با سیم A و با سیم B یک فاصله اطمینان ۹۵٪ بیابید. آیا این نمونه‌ها فرض برابری میانگینها را تأیید می‌کند.

۳۰ نمرات زیر نمونه‌ای از نمرات درس آمار و احتمال در دو گروه ۱ و ۲ می‌باشد. فرض کنید نمرات دو کلاس دارای توزیع نرمال با واریانسهای مساوی باشند.

گروه ۱	۱۲	۱۰	۱۴	۱۳	۱۱		
گروه ۲	۱۷	۱۵	۱۴	۱۶	۱۷	۱۷	۱۶

الف – فاصله (۱۰/۹۲،۱۳/۰۸) یک فاصله اطمینان چند درصدی برای میانگین نمرات گروه ۱ است؟

ب – آیا با اطمینان ۹۵ درصد می‌توان گفت که میانگین نمره گروه ۲ بیشتر از گروه ۱ است؟ چرا؟

۳۱ برای مقایسه نمرات امتحان آمار دو کلاس از هر کلاس یک نمونه انتخاب کرده‌ایم. فرض کنید از کلاس اول یک نمونه ۲۵ تایی انتخاب کرده و فاصله اطمینان ۹۵ درصدی (۱۲/۹۴،۱۷/۰۶) برای میانگین نمرات این کلاس به دست آمده باشد و از کلاس دوم یک نمونه ۱۶ تایی انتخاب کرده و یک فاصله اطمینان ۹۰ درصدی (۱۰/۲۵،۱۳/۷۵) برای میانگین نمرات این کلاس به دست آمده باشد. با فرض نرمال بودن نمرات دو کلاس و مساوی بودن واریانسها، یک فاصله اطمینان ۹۰ درصدی برای اختلاف میانگین نمرات دو کلاس به دست آورید. چه نتیجه‌ای می‌گیرید؟

۳۲ فرض کنید معدل دانشجویان دو دانشگاه A و B دارای توزیع نرمال باشد. در صورتی که دو نمونه تصادفی انتخابی از این دو دانشگاه به صورت زیر باشد، یک فاصله اطمینان ۹۵ درصدی برای $\frac{\sigma_A^2}{\sigma_B^2}$ به دست آورید.

دانشگاه A	۳	۳/۲	۲/۹	۳/۴	۲/۵	۲/۸	۳	۳/۳	۲/۷	۳/۱
دانشگاه B	۳	۲/۶	۳/۱	۳/۵	۳/۲	۳/۹	۲/۷			

۳۳ در تمرین ۲۸ یک فاصله اطمینان ۹۰ درصدی برای $\frac{\sigma_1}{\sigma_2}$ تشکیل دهید.

۳۴ در تمرین ۳۰ یک فاصله اطمینان ۹۰ درصدی برای $\frac{\sigma_2^2}{\sigma_1^2}$ تشکیل دهید.

۳۵ آزمایشی برای تعیین غلظت دو نوع متفاوت بنزین سربدار و بدون سرب، نتایج زیر را به دست داده است

نوع سربدار : $n_1 = ۲۵ \quad , \quad \bar{x}_1 = ۳۵/۸۴ \quad , \quad s_1^2 = ۱۳۰/۴۵۷۶$

نوع بدون سرب : $n_2 = ۲۵ \quad , \quad \bar{x}_2 = ۳۰/۶۰ \quad , \quad s_2^2 = ۵۳/۰۶۰۴$

به فرض نرمال بودن، یک فاصله اطمینان ۹۰ درصدی برای نسبت واریانسهای دو جمعیت به دست آورید.

۳۶ دو نمونه تصادفی با اندازه‌های ۲۵ و ۱۶ به ترتیب از دانشجویان پسر و دختری که در یک آزمون شرکت کرده‌اند را انتخاب کرده و مشاهده نموده‌ایم که میانگین نمرات دانشجویان پسر ۸۲ و واریانس آن ۶۴ است، در حالیکه میانگین نمرات دانشجویان دختر ۷۸ و واریانس آن ۴۹ می‌باشد. یک فاصله اطمینان ۹۸ درصدی برای نسبت انحراف معیار نمرات پسرها به دخترها به دست آورید. فرض کنید توزیع نمرات نرمال باشد.

فصل هشتم

آزمون فرضهای آماری

۱.۸ مفاهیم اولیه

در فصل قبل یکی از شاخه‌های استنباط آماری یعنی برآورد پارامتر مجهول جمعیت را مورد بررسی قرار دادیم. در این فصل یکی دیگر از شاخه‌های استنباط آماری یعنی آزمون فرضیه‌های آماری را مورد بررسی قرار می‌دهیم. ابتدا تعریف یک فرض آماری را می‌آوریم.

تعریف ۱.۸ یک فرض آماری ادعائی در مورد یک یا چند جمعیت مورد بررسی است که ممکن است درست یا نادرست باشد. به عبارت دیگر فرض آماری یک ادعا یا گزاره‌ای در مورد توزیع یک جمعیت یا توزیع یک متغیر تصادفی است.

برای درک مفهوم فرض آماری به مثال زیر توجه کنید.

مثال ۱.۱.۸ آزمایشی نشان داده است که میزان مؤثربودن نوعی داروی استاندارد روی یک بیماری بخصوص ۶۰ درصد است. یک داروساز ادعا می‌کند که اثر داروی جدیدی که او ساخته است بیشتر از داروی استاندارد می‌باشد. این ادعای داروساز یک فرض آماری است. حال برای اینکه ادعای این داروساز را مورد بررسی قرار دهیم، بایستی این دارو را روی افراد بیمار جمعیت آزمایش کنیم و نتایج را مورد بررسی قرار دهیم. اما آزمایش این دارو روی تمامی افراد بیمار جمعیت معقولانه نمی‌باشد و بجای آن بایستی دارو را روی تعدادی از افراد بیمار جمعیت آزمایش کنیم که این تعداد افراد همان نمونه ما را تشکیل می‌دهند. بنابراین فرض کنید که این دارو را روی ۲۰ بیمار آزمایش کنیم و متغیر تصادفی X را برابر تعداد بیمارانی در بین ۲۰ بیمار در نظر بگیریم

که توسط داروی جدید معالجه شده‌اند. در این صورت $X\sim B(۲۰, p)$ که در آن p مقداری نامعلوم و درصد مؤثر بودن داروی جدید است. مؤثر بودن داروی جدید به این معنی است که $p>۰/۶۰$ است. بنابراین در رابطه با این سئوال که "آیا میزان مؤثر بودن داروی جدید بیشتر از داروی استاندارد است؟" دو حالت (دو فرض) زیر پیش می‌آید.

۱ـ داروی جدید بهتر از داروی استاندارد است $p>۰/۶۰$

۲ـ داروی جدید بهتر از داروی استاندارد نیست $p\leq ۰/۶۰$

حال بایستی بوسیله اطلاعاتی که از نمونه به دست می‌آید در مورد درست بودن یا نبودن این فرضیات نتیجه‌گیری کنیم.

از دو فرضی راکه در یک مسئله آزمون فرض مطرح می‌شود یکی را فرض صفر یا خنثی[1] گفته و آن را با H_0 نمایش داده و دیگری را فرض مقابل[2] گفته و آن را با H_1 نمایش می‌دهند. هرگاه بخواهیم یک ادعا را از طریق تائید آن بوسیله اطلاعات حاصل از نمونه ثابت کنیم، خلاف آن ادعا را در فرض صفر H_0 و خود آن ادعا را در فرض مقابل H_1 قرار می‌دهیم. بنابراین در مثال ۱.۱.۸ با مسئله آزمون فرضهای زیر مواجه می‌شویم

$$\begin{cases} H_0 : p\leq ۰/۶۰ \\ H_1 : p > ۰/۶۰ \end{cases}$$

آزمون فرضهای آماری در یک مسئله آزمون فرضهای آماری اگر اطلاعات به دست آمده از نمونه با فرض H_0 "ناسازگار" باشد در این صورت فرض H_0 را رد می‌کنیم و در مقابل فرض H_1 را می‌پذیریم و اگر اطلاعات به دست آمده از نمونه با فرض H_0 "سازگار" باشد در این صورت گوئیم که اطلاعات به دست آمده از نمونه دلیلی بر رد فرض H_0 ندارد و یا در حقیقت گوئیم فرض H_0 را می‌پذیریم. مثلاً در مثال ۱.۱.۸ اگر درصد افراد بهبود یافته در نمونه از ۶۰ درصد بیشتر باشد فرض H_0 را رد می‌کنیم و ادعای داروساز را می‌پذیریم و در مقابل اگر درصد افراد بهبود یافته در نمونه کمتر از ۶۰ درصد باشد آنگاه دلیلی بر رد فرض H_0 نداریم. در مسئله آزمون فرضها وقتی می‌گوئیم یک فرض آماری رد شده است به این معنی است که از روی نمونه انتخابی فرض آماری با قاطعیت رد شده است، اما وقتی گوئیم یک فرض آماری پذیرفته شده است به این معنی است که نمونه به دست آمده از جمعیت دلیلی بر ردکردن آن فرض آماری را به دست نمی‌دهد. بنابراین در

1- null hypothesis
2- Alternative hypothesis

یک مسئله آزمون فرضها هدف ما ردکردن فرض H_0 و در نتیجه پذیرفتن فرض مقابل H_1 یعنی ادعای موردنظر می‌باشد.

ناحیه بحرانی و آماره آزمون برای انجام یک آزمون آماری نیاز به آماره و ناحیه بحرانی آزمون داریم که با ذکر یک مثال آنها را تشریح می‌کنیم.

مثال ۲.۱.۸ فرض کنید که یک داروی استاندارد ۲۵٪ در درمان یک بیماری مؤثر است و شخصی ادعا می‌کند که داروی ساخته شده توسط او ۵۰٪ در درمان آن بیماری مؤثر می‌باشد. بنابراین برای تحقیق در صحت حرف این شخص با آزمون زیر مواجه می‌شویم

$$\begin{cases} H_0 : p=\frac{۱}{۴} \\ H_1 : p=\frac{۱}{۲} \end{cases}$$

حال این سئوال مطرح می‌شود که چگونه این آزمون را انجام دهیم. همانطور که قبلاً گفته شد، بایستی نمونه‌ای از بیماران را در نظر بگیریم و دارو را روی آنها آزمایش کنیم. فرض کنید که ۲۰ بیمار را انتخاب کنیم و متغیر تصادفی X را برابر تعداد بیماران بهبود یافته توسط داروی جدید در بین ۲۰ بیمار در نظر بگیریم. در این صورت اگر مقادیر مشاهده شده X کوچک باشد، مثلاً کمتر از ۸، آنگاه نمی‌توان H_0 را ردکرد زیرا در این صورت کمتر از ۴۰٪ افراد بیمار بهبود یافته‌اند و نمی‌تواند $p=\frac{۱}{۲}$ باشد. حال فرض کنید قرارداد کنیم که اگر مقادیر مشاهده شده X بزرگتر یا مساوی ۹ باشد آنگاه فرض H_0 را رد خواهیم کرد. یعنی اگر مقادیر مشاهده شده X متعلق به مجموعه $C=\{x \mid x\geq ۹\}$ باشد آنگاه H_0 را ردکنیم و اگر چنین نبود H_0 را رد نکنیم. به این آماره X که بر اساس مقادیر مشاهده شده آن فرض H_0 را رد یا قبول می‌کنیم **آماره آزمون** گویند و به ناحیه C که کلیه مقادیر مربوط به رد فرض H_0 را به دست می‌دهد، **ناحیه بحرانی آزمون** گویند.

تعریف ۲.۸ آماره $T=T(X_1,\ldots,X_n)$ که بر اساس مقادیر مشاهده شده آن یک فرض را رد یا قبول می‌کنیم آماره آزمون گویند و به مجموعه مقادیری از این آماره که به ازای آن فرض H_0 را بایستی ردکرد، ناحیه بحرانی آزمون گویند و با نماد C نمایش می‌دهند. متمم ناحیه بحرانی یعنی C' را ناحیه پذیرش آزمون گویند.

اگر ناحیه بحرانی C یک آزمون مشخص شود در این صورت با جمع آوری نمونه و محاسبه $t=T(x_1,\ldots,x_n)$ می‌توان آزمون آماری را به صورت زیر انجام داد. اگر $T(x_1,\ldots,x_n)\in C$

آنگاه فرض $H_۰$ را رد می‌کنیم و در غیر این صورت آن را می‌پذیریم.

بنابراین در مثال ۲.۱.۸ اگر $x \geq ۹$ فرض $H_۰$ را رد می‌کنیم و در غیر این صورت آن را می‌پذیریم.

خطاهای آزمون

آیا قضاوتی راکه در مثال ۲.۱.۸ انجام دادیم بدون خطا می‌باشد؟ جواب این سئوال منفی است. زیرا ممکن است که واقعاً $H_۰$ درست باشد یعنی داروی جدید نیز ۲۵٪ مؤثر باشد و ما مشاهده کنیم که در این نمونه ۲۰ بیماری ۱۰ بیمار بهبود یافته‌اند و در نتیجه فرض $H_۰$ را ردکنیم. بنابراین ممکن است فرض $H_۰$ درست باشد و ما آن را ردکنیم که این خطا را خطای نوع اول آزمون گویند. در مقابل ممکن است که فرض $H_۰$ درست نباشد یعنی داروی جدید ۵۰٪ مؤثر باشد و ما مشاهده کنیم که ۶ بیمار از بین ۲۰ بیمار بهبود یافته‌اند و در نتیجه فرض $H_۰$ را قبول کنیم. بنابراین ممکن است فرض $H_۰$ نادرست باشد و ما آن را قبول کنیم که این خطا را خطای نوع دوم آزمون گویند.

احتمال خطای نوع اول را با α نمایش می‌دهند و آن را **سطح معنی‌دار یا سطح تشخیص** آزمون گویند و احتمال خطای نوع دوم را با β نمایش می‌دهند. بنابراین

$$\alpha=P(\text{فرض } H_۰ \text{ رد شود} \mid H_۰ \text{ درست است})=P(T(X_۱,\dots,X_n)\in C \mid H_۰ \text{ درست است})$$

$$\beta=P(\text{فرض } H_۰ \text{ قبول شود} \mid H_۱ \text{ درست است})=P(T(X_۱,\dots,X_n)\notin C \mid H_۱ \text{ درست است})$$

توان آزمون احتمال ردکردن فرض $H_۰$ در صورتی که فرض $H_۱$ درست باشد یعنی احتمال رد کردن فرض $H_۰$ به حق را توان آزمون گویند و با β^* نشان می‌دهند. بنابراین

$$\beta^*=P(\text{فرض } H_۰ \text{ ردشود} \mid H_۱ \text{ درست‌است})=۱-P(\text{فرض } H_۰ \text{ قبول‌شود} \mid H_۱ \text{ درست‌است})=۱-\beta$$

مثال ۳.۱.۸ در مثال ۲.۱.۸ احتمال خطای نوع اول و دوم و توان آزمون را محاسبه کنید.

حل در مثال ۲.۱.۸ داشتیم که $C=\{x \mid x\geq ۹\}$ و $X\sim B(۲۰,p)$. بنابراین

$$\alpha=P(X\in C \mid H_۰ \text{ درست})=P(X\geq ۹ \mid p=\tfrac{۱}{۴})$$

$$=۱-P(X\leq ۸ \mid p=\tfrac{۱}{۴})=۱-۰/۹۵۹۱=۰/۰۴۰۹$$

$$\beta=P(X\notin C \mid H_۱ \text{ درست})=P(X<۹ \mid p=\tfrac{۱}{۲})$$

$$=P(X\leq ۸ \mid p=\tfrac{۱}{۲})=۰/۲۵۱۷$$

$$\beta^*=۱-\beta=۱-۰/۲۵۱۷=۰/۷۴۸۳$$

بنابراین احتمال خطای نوع اول کم و احتمال خطای نوع دوم زیاد است.

مثال ۴.۱.۸ در مثال ۲.۱.۸ اگر ناحیه بحرانی به صورت $C=\{x \mid x\geq ۸\}$ باشد، احتمال خطای نوع اول و دوم و توان آزمون را محاسبه کنید.

حل

$$\alpha=P(X\geq ۸ \mid p=\tfrac{۱}{۴})=۱-P(X\leq ۷ \mid p=\tfrac{۱}{۴})=۱-۰/۸۹۸۲=۰/۱۰۱۸$$

$$\beta=P(X<۸ \mid p=\tfrac{۱}{۲})=P(X\leq ۷ \mid p=\tfrac{۱}{۲})=۰/۱۳۱۶$$

$$\beta^*=۱-۰/۱۳۱۶=۰/۸۶۸۴$$

بنابراین احتمال خطای نوع اول افزایش یافت و در مقابل احتمال خطای نوع دوم کاهش یافت.

با مقایسه دو مثال بالا دیده می‌شود که با تغییر دادن ناحیه بحرانی نمی‌توان هم خطای نوع اول و هم خطای نوع دوم را همزمان کاهش داد. موقعی که یکی را کاهش می‌دهیم، دیگری افزایش می‌یابد و برعکس. بنابراین بایستی آن ناحیه بحرانی را انتخاب کنیم که با قراردادن یک حداکثر مقدار برای احتمال خطای نوع اول بتوان احتمال خطای نوع دوم را تا آنجاکه ممکن است کاهش داد و یا به عبارتی تا آنجاکه ممکن است توان آزمون را حداکثر کرد.

مثال ۵.۱.۸ فرض کنید $X\sim N(\mu,۴)$ و آزمون زیر را در نظر بگیرید

$$\begin{cases} H_۰ : \mu=۰ \\ H_۱ : \mu=۱ \end{cases}$$

یک نمونه تصادفی به اندازه n=۲۵ از X در نظر می‌گیریم اگر ناحیه بحرانی به صورت $C=\{(X_۱,\dots,X_{۲۵}) \mid \overline{X}>c\}$ باشد، مقدار c را چنان تعیین کنید که $\alpha=۰/۱$ باشد و احتمال خطای نوع دوم و توان آزمون را محاسبه کنید.

حل می‌دانیم که $\overline{X}\sim N(\mu,\frac{۴}{۲۵})$ بنابراین

$$۰/۱=\alpha=P(\overline{X}>c \mid \mu=۰)=P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}>\frac{c-۰}{\frac{۲}{۵}}\right)=P(Z>۲/۵c)$$

در نتیجه

$$P(Z\leq ۲/۵c)=۰/۹ \Rightarrow ۲/۵c=z_{۰/۹}=۱/۲۸ \Rightarrow c=\frac{۱/۲۸}{۲/۵}=۰/۵۱۲$$

$$\beta=P(\overline{X}\le C\mid \mu=۱)=P(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}>\frac{۰/۵۱۲-۱}{\frac{۲}{۵}})=P(Z<-۱/۲۲)=۰/۱۱۱۲$$

$$\beta^{*}=۱-\beta=۱-۰/۱۱۱۲=۰/۸۸۸۸$$

انواع فرضها فرضهای آماری به طور کلی به دو دسته ساده و مرکب تقسیم می‌شوند. فرضی را ساده گوئیم که تحت آن فرض توزیع جمعیت کاملاً مشخص گردد. مثلاً در مثال ۵.۱.۸ فرض $H_{۰}:\mu=۰$ یک فرض ساده می‌باشد. فرضی را مرکب گوئیم که تحت آن فرض توزیع جمعیت کاملاً مشخص نمی‌گردد. مثلاً در مثال ۱.۱.۸ فرض $H_{۱}:p>۰/۶$ یک فرض مرکب می‌باشد زیرا اگر این فرض درست باشد مقدار p دقیقاً مشخص نمی‌شود و در نتیجه توزیع جمعیت کاملاً مشخص نمی‌گردد.

آزمونهای یک طرفه و دو طرفه فرض کنید θ پارامتر مجهول جمعیت باشد و بخواهیم آزمونهایی در مورد این پارامتر انجام دهیم. آزمون هر فرضیه آماری که در آن فرضیه مقابل یک طرفه باشد، آزمون یک طرفه نامیده می‌شود. برای مثال اگر $\theta_{۰}$ مقدار ثابتی از θ باشد آنگاه هر یک از آزمونهای زیر یک طرفه می‌باشند

$$\begin{cases}H_{۰}:\theta\le\theta_{۰}\\H_{۱}:\theta>\theta_{۰}\end{cases}\qquad\begin{cases}H_{۰}:\theta=\theta_{۰}\\H_{۱}:\theta<\theta_{۰}\end{cases}$$

آزمون هر فرضیه آماری که در آن فرضیه مقابل دو طرفه باشد یعنی

$$\begin{cases}H_{۰}:\theta=\theta_{۰}\\H_{۱}:\theta\neq\theta_{۰}\end{cases}$$

یک آزمون دو طرفه نامیده می‌شود.

مراحل انجام یک آزمون با توجه به مطالب گفته شده در این بخش، برای انجام یک آزمون آماری بایستی مراحل زیر را طی کرد.

۱ـ تعیین فرضهای صفر $H_{۰}$ و مقابل $H_{۱}$.

۲ـ تعیین یک سطح معنی دار α که معمولاً آن را ۰۱/۰ یا ۰۵/۰ یا ۱/۰ می‌گیرند.

۳ـ تعیین آماره آزمون $T=T(X_{۱},\dots,X_{n})$ که معمولاً براساس برآوردگر نقطه‌ای پارامتر مجهول θ می‌باشد.

۴ـ تعیین ناحیه بحرانی آزمون C که بر اساس آماره آزمون، فرض مقابل آزمون و سطح

معنی دار α می‌باشد.

۵ـ محاسبه مقدار مشاهده شده آماره آزمون بر اساس نمونه تصادفی مشاهده شده $x_{۱}$، $x_{۲}$، ... و x_{n}.

۶ـ نتیجه گیری: اگر مقدار محاسبه شده آماره آزمون یعنی $T(x_{۱},\dots,x_{n})$ در ناحیه بحرانی C قرار گرفت آنگاه فرض $H_{۰}$ را رد می‌کنیم و در غیر این صورت فرض $H_{۰}$ را قبول می‌کنیم.

در بخش بعد این مراحل را در انجام آزمون روی پارامترهای مختلف جمعیت بکار می‌بریم.

۲.۸ آزمون فرضهای آماری روی پارامترهای جمعیت

در این بخش با توجه به مطالب بیان شده در بخش قبل آزمون فرضهای آماری روی میانگین و واریانس جمعیت در حالتهای مختلف را انجام می‌دهیم. در ابتدا آزمون فرض روی میانگین یک جمعیت موقعی که واریانس جمعیت معلوم است را بررسی می‌کنیم.

فرض کنید که از یک جمعیت X با میانگین مجهول μ و واریانس معلوم $\sigma^{۲}$ یک نمونه تصادفی $X_{۱}$، $X_{۲}$، ... و X_{n} انتخاب کرده باشیم و بخواهیم آزمونهایی روی میانگین μ انجام دهیم. برای این منظور ۳ حالت زیر را در نظر می‌گیریم.

الف در آزمون فرض $\begin{cases}H_{۰}:\mu=\mu_{۰}\\H_{۱}:\mu>\mu_{۰}\end{cases}$ که در آن $\mu_{۰}$ مقداری معلوم است، اگر $\overline{X}$ برآوردگر μ مقادیر بزرگ را اختیار کند یعنی $\overline{X}>c$ آنگاه فرض $H_{۰}$ را رد می‌کنیم. بنابراین ناحیه بحرانی آزمون به صورت $\overline{X}>c$ است که در آن c به گونه‌ای تعیین می‌گردد که سطح معنی دار آزمون برابر مقدار مشخص شده α باشد. یعنی

$$\alpha=P(\overline{X}>c\mid\mu=\mu_{۰})$$

حال اگر جمعیت نرمال باشد و یا اینکه نرمال نبوده اما $n\ge۳۰$ باشد آنگاه طبق مطالب بخش ۲.۶ داریم که

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(۰,۱)$$

بنابراین برای تعیین c داریم که

$$\alpha=P(\overline{X}>c\mid\mu=\mu_{۰})=P(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}>\frac{c-\mu_{۰}}{\sigma/\sqrt{n}})=P(Z>\frac{c-\mu_{۰}}{\sigma/\sqrt{n}})$$

در نتیجه
$$\frac{c-\mu_0}{\sigma/\sqrt{n}} = z_{1-\alpha} \Rightarrow c=\mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$$

بنابراین ناحیه بحرانی آزمون به صورت $\bar{X} > \mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$ و یا $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}$ می‌باشد.

در نتیجه

> در آزمون $\begin{cases}H_0: \mu=\mu_0\\ H_1: \mu>\mu_0\end{cases}$ فرض H_0 رد شود اگر و فقط اگر
> $$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}$$

توجه کنید که اگر فرض H_0 به صورت $H_0: \mu\le\mu_0$ باشد، آنگاه می‌توان نشان داد که ناحیه بحرانی هنوز به صورت فوق می‌باشد.

ب در آزمون فرض $\begin{cases}H_0: \mu=\mu_0\\ H_1: \mu<\mu_0\end{cases}$ اگر $\bar{X}$ مقادیر کوچک را اختیار کند یعنی $\bar{X}<c$ آنگاه فرض H_0 را رد می‌کنیم. بنابراین ناحیه بحرانی به صورت $\bar{X}<c$ است که در آن c با انجام عملیات مشابه قسمت (الف) برابر $c=\mu_0 - z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$ به دست می‌آید و بنابراین ناحیه بحرانی به صورت $\bar{X}<\mu_0 - z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$ و یا $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < -z_{1-\alpha}$ خواهد بود. بنابراین

> در آزمون $\begin{cases}H_0: \mu=\mu_0\\ H_1: \mu<\mu_0\end{cases}$ فرض H_0 رد شود اگر و فقط اگر
> $$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < -z_{1-\alpha}$$

توجه کنید که اگر فرض H_0 به صورت $H_0: \mu\ge\mu_0$ باشد آنگاه می‌توان نشان داد که ناحیه بحرانی هنوز به صورت فوق می‌باشد.

ج در آزمون فرض $\begin{cases}H_0: \mu=\mu_0\\ H_1: \mu\neq\mu_0\end{cases}$ اگر $\bar{X}$ مقادیر کوچک یا مقادیر بزرگ را اختیار کند یعنی $\bar{X}<c_1$ یا $\bar{X}>c_2$ آنگاه فرض H_0 را رد می‌کنیم. بنابراین ناحیه بحرانی به صورت $\bar{X}<c_1$ یا $\bar{X}>c_2$ است که در آن c_1 و c_2 به صورت زیر تعیین می‌گردند.

$$\alpha=P(\bar{X}<c_1 \text{ یا } \bar{X}>c_2 \mid \mu=\mu_0) \Rightarrow 1-\alpha = P(c_1<\bar{X}<c_2 \mid \mu=\mu_0)$$

بنابراین
$$1-\alpha = P\left(\frac{c_1-\mu_0}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{c_2-\mu_0}{\sigma/\sqrt{n}}\right)$$

و در نتیجه
$$\frac{c_2-\mu_0}{\sigma/\sqrt{n}} = -\frac{c_1-\mu_0}{\sigma/\sqrt{n}} = z_{1-\frac{\alpha}{2}}$$

و یا $c_1=\mu_0 - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ و $c_2=\mu_0 + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$. بنابراین ناحیه بحرانی به صورت $\bar{X} > \mu_0 + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ یا $\bar{X} < \mu_0 - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ و یا معادلاً $\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{1-\frac{\alpha}{2}}$ می‌باشد.

بنابراین

> در آزمون $\begin{cases}H_0: \mu=\mu_0\\ H_1: \mu\neq\mu_0\end{cases}$ فرض H_0 رد شود اگر و فقط اگر
> $$|Z| = \left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{1-\frac{\alpha}{2}}$$

مثال ۱.۲.۸ یک کارخانه تولیدکننده لامپهای روشنایی، لامپهایی تولید می‌کند که طول عمر آنها از توزیع نرمال با حد متوسط ۸۰۰ ساعت و انحراف معیار ۴۰ ساعت پیروی می‌کند. می‌خواهیم آزمون $\mu=800$ در مقابل $\mu\neq 800$ را انجام دهیم. اگر یک نمونه تصادفی ۳۰ تایی از آن لامپها دارای حد متوسط طول عمر ۷۸۸ ساعت باشد، آزمون فوق را در سطح معنی دار ۰/۰۴ انجام دهید.

حل آزمون از نوع (ج) می‌باشد که در آن $\bar{x}=788$، $\sigma=40$، $n=30$، $\mu_0=800$ و $\alpha=0/04$ است بنابراین

$$Z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{788-800}{\frac{40}{\sqrt{30}}} = -1/643 \quad , \quad z_{1-\frac{\alpha}{2}} = z_{0/98} = 2/05$$

چون $|Z| = 1/643 < z_{1-\frac{\alpha}{2}} = 2/05$ بنابراین فرض H_0 رد نمی‌شود، یعنی حد متوسط طول عمر لامپها برابر ۸۰۰ ساعت است.

مثال ۲.۲.۸ تعداد زیادی از بیماران مبتلا به یک بیماری بخصوص را گرد آورده و گزارش کرده‌اند که مدت درمان بیماری به روش استاندارد دارای میانگین ۱۵ روز و انحراف معیار ۳ روز می‌باشد. ادعا شده که یک روش جدید می‌تواند مدت درمان را کوتاه‌تر کند و انحراف معیار درمان

همان ۳ روز می‌باشد. برای روش جدید درمان را بر روی ۷۰ بیمار آزمایش کرده و میانگین مدت درمان ۱۴ روز شده است. آیا در سطح معنی‌دار ۰/۰۲۵ روش جدید بهتر است؟

حل در این مثال با آزمون $\begin{cases} H_0 : \mu = 15 \\ H_1 : \mu < 15 \end{cases}$ مواجه هستیم، پس فرض H_0 را رد می‌کنیم در صورتی که $Z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -z_{1-\alpha}$ از طرفی داریم که

$$\alpha = 0/025 \Rightarrow z_{1-\alpha} = z_{0/975} = 1/96$$

$$Z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{14-15}{\frac{3}{\sqrt{70}}} = -2/789$$

چون $-2/789 = Z < -z_{1-\alpha} = -1/96$ پس فرض H_0 رد می‌شود، یعنی میانگین مدت درمان روش جدید کمتر است.

برای آزمون فرض روی میانگین موقعی که واریانس نامعلوم است، آزمون فرض روی واریانس یک جمعیت، آزمون فرض روی تفاضل میانگین دو جمعیت و یا نسبت واریانس دو جمعیت نیز می‌توان با انجام عملیات مشابه حالتهای (الف) تا (ج) نواحی بحرانی آزمون را تعیین کرد. حاصل این عملیات در جدولهای ۱.۸ و ۲.۸ ارایه گردیده است. در جدول ۱.۸ یک نمونه تصادفی nتایی از یک جمعیت نرمال با میانگین μ و واریانس σ^2 انتخاب شده و آزمونهایی روی μ یا σ^2 انجام گرفته است. در آزمونهای ۱ تا ۱۳ اگر $n \geq 30$ باشد فرض نرمال بودن جمعیت را می‌توان حذف کرد. در جدول ۲.۸ یک نمونه تصادفی n_1 تایی از جمعیت نرمال با میانگین μ_1 و واریانس σ_1^2 و یک نمونه تصادفی n_2 تایی از جمعیت نرمال دیگری با میانگین μ_2 و واریانس σ_2^2 انتخاب شده‌اند و این دو نمونه از یکدیگر مستقل هستند. سپس آزمونهایی روی تفاضل میانگین دو جمعیت و یا نسبت واریانس دو جمعیت انجام گرفته است. در آزمونهای ۱۰ تا ۱۲ اگر $n_1 \geq 30$ و $n_2 \geq 30$ باشد فرض نرمال بودن را می‌توان حذف کرد. در زیر مثالهایی از این آزمونها را می‌آوریم. توجه کنید اگر در مسئله‌ای مقدار α مشخص نشده باشد آن را ۰/۰۵ در نظر می‌گیریم.

شماره آزمون	H_0	H_1	آماره آزمون	ناحیه بحرانی
۱	$\mu=\mu_0$ یا $\mu\leq\mu_0$	$\mu>\mu_0$	σ معلوم، $Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$Z>z_{1-\alpha}$
۲	$\mu=\mu_0$ یا $\mu\geq\mu_0$	$\mu<\mu_0$	σ معلوم، $Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$Z<-z_{1-\alpha}$
۳	$\mu=\mu_0$	$\mu\neq\mu_0$	σ معلوم، $Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$\lvert Z\rvert>z_{1-\frac{\alpha}{2}}$
۴	$\mu=\mu_0$ یا $\mu\leq\mu_0$	$\mu>\mu_0$	σ نامعلوم، $T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$T>t_{1-\alpha}(n-1)$
۵	$\mu=\mu_0$ یا $\mu\geq\mu_0$	$\mu<\mu_0$	σ نامعلوم، $T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$T<-t_{1-\alpha}(n-1)$
۶	$\mu=\mu_0$	$\mu\neq\mu_0$	σ نامعلوم، $T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$\lvert T\rvert>t_{1-\frac{\alpha}{2}}(n-1)$
۷	$\sigma^2=\sigma_0^2$ یا $\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$	$X^2=\frac{(n-1)S^2}{\sigma_0^2}$	$X^2>\chi^2_{1-\alpha}(n-1)$
۸	$\sigma^2=\sigma_0^2$ یا $\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$	$X^2=\frac{(n-1)S^2}{\sigma_0^2}$	$X^2<\chi^2_{\alpha}(n-1)$
۹	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$X^2=\frac{(n-1)S^2}{\sigma_0^2}$	$X^2>\chi^2_{1-\frac{\alpha}{2}}(n-1)$ یا $X^2<\chi^2_{\frac{\alpha}{2}}(n-1)$

جدول ۱.۸ آزمونهای آماری روی میانگین و واریانس یک جمعیت نرمال

شماره آزمون	H_0	H_1	آماره آزمون	ناحیه بحرانی
۱۰	$\mu_1-\mu_2=d_0$ یا $\mu_1-\mu_2\le d_0$	$\mu_1-\mu_2>d_0$	$Z=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ ، σ_1^2 و σ_2^2 معلوم	$Z>z_{1-\alpha}$
۱۱	$\mu_1-\mu_2=d_0$ یا $\mu_1-\mu_2\ge d_0$	$\mu_1-\mu_2<d_0$	$Z=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ ، σ_1^2 و σ_2^2 معلوم	$Z<-z_{1-\alpha}$
۱۲	$\mu_1-\mu_2=d_0$	$\mu_1-\mu_2\neq d_0$	$Z=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ ، σ_1^2 و σ_2^2 معلوم	$\lvert Z\rvert>z_{1-\frac{\alpha}{2}}$
۱۳	$\mu_1-\mu_2=d_0$ یا $\mu_1-\mu_2\le d_0$	$\mu_1-\mu_2>d_0$	$T=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ ، $\sigma_1=\sigma_2$ ولی نامعلوم	$T>t_{1-\alpha}(n_1+n_2-2)$
۱۴	$\mu_1-\mu_2=d_0$ یا $\mu_1-\mu_2\ge d_0$	$\mu_1-\mu_2<d_0$	$T=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ ، $\sigma_1=\sigma_2$ ولی نامعلوم	$T<-t_{1-\alpha}(n_1+n_2-2)$
۱۵	$\mu_1-\mu_2=d_0$	$\mu_1-\mu_2\neq d_0$	$T=\frac{(\bar{X}_1-\bar{X}_2)-d_0}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ ، $\sigma_1=\sigma_2$ ولی نامعلوم	$\lvert T\rvert>t_{1-\frac{\alpha}{2}}(n_1+n_2-2)$
۱۶	$\sigma_1^2=\sigma_2^2$ یا $\sigma_1^2\le\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$F=\frac{S_1^2}{S_2^2}$	$F>F_{1-\alpha}(n_1-1,n_2-1)$
۱۷	$\sigma_1^2=\sigma_2^2$ یا $\sigma_1^2\ge\sigma_2^2$	$\sigma_1^2<\sigma_2^2$	$F=\frac{S_1^2}{S_2^2}$	$F<F_{\alpha}(n_1-1,n_2-1)$
۱۸	$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$F=\frac{S_1^2}{S_2^2}$	$F<F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$ یا $F>F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$

جدول ۲.۸ آزمونهای آماری روی تفاضل میانگینها و نسبت واریانسهای دو جمعیت نرمال

مثال ۳.۲.۸ نمونه تصادفی از پرونده‌های فراوان شرکتی نشان می‌دهد که سفارشات برای قطعه معینی از ماشینها به ترتیب در ۱۰، ۱۲، ۱۹، ۱۴، ۱۵، ۱۸، ۱۱ و ۱۳ روز بایگانی شده است. اگر تعداد روزهای بایگانی از توزیع نرمال پیروی کند، آیا در سطح معنی‌دار ۰/۰۱ می‌توان ادعا کرد که میانگین زمان بایگانی چنین سفارشاتی از ۱۰/۵ روز بیشتر است؟

حل در این مثال با آزمون $\begin{cases}H_0: \mu=10/5\\ H_1: \mu>10/5\end{cases}$ مواجه هستیم که در آن σ^2 نامعلوم است. بنابراین از آزمون شماره ۴ استفاده می‌کنیم. از داده‌ها به دست می‌آوریم که $n=8$، $\sum x_i=112$ و $\sum x_i^2=1640$ بنابراین

$$\bar{x}=\frac{112}{8}=14 \quad , \quad s^2=\frac{1}{7}\left[1640-\frac{(112)^2}{8}\right]=10/286$$

$$\alpha=0/01 \Rightarrow t_{1-\alpha}(n-1)=t_{0/99}(7)=3/0$$

$$T=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}=\frac{14-10/5}{\sqrt{\frac{10/286}{8}}}=3/087$$

چون $3/087=T>t_{1-\alpha}(n-1)=3/0$ پس فرض H_0 رد می‌شود، یعنی میانگین زمان بایگانی بیش از ۱۰/۵ روز است.

مثال ۴.۲.۸ یک تولید کننده قطعات پیش ساخته مدعی است که انحراف معیار مقاومت محصولات او برابر ۱۰ کیلوگرم بر سانتیمتر مربع است. یک نمونه تصادفی ۱۰تایی از این محصولات نتایج $\bar{x}=312$ و $s^2=195$ را به دست داده است. اگر اندازه مقاومت این محصولات دارای توزیع نرمال باشند، آیا نتایج به دست آمده با ادعای تولید کننده سازگار است؟ سطح معنی‌دار را ۰/۰۵ بگیرید.

حل در این مثال با آزمون $\begin{cases}H_0: \sigma^2=100\\ H_1: \sigma^2\neq 100\end{cases}$ مواجه هستیم بنابراین از آزمون شماره ۹ استفاده می‌کنیم. در نتیجه

$$X^2=\frac{(n-1)s^2}{\sigma_0^2}=\frac{9(195)}{100}=17/55$$

$$\alpha=0/05 \Rightarrow \chi^2_{\frac{\alpha}{2}}(n-1)=\chi^2_{0/025}(9)=2/70 \quad , \quad \chi^2_{1-\frac{\alpha}{2}}(n-1)=\chi^2_{0/975}(9)=19/0$$

چون $۲/۷۰ = \chi^2_{\frac{\alpha}{2}}(n-۱) \not> X^2 = ۱۷/۵۵$ و $۱۹/۰ = \chi^2_{۱-\frac{\alpha}{2}}(n-۱) \not< X^2 = ۱۷/۵۵$ پس فرض $H_۰$ رد نمی‌شود یعنی نتایج با ادعای تولیدکننده سازگار است.

مثال ۵.۲.۸ یک نمونه تصادفی به اندازه ۳۶ از یک جمعیت با انحراف معیار ۵/۲ دارای میانگین ۸۱ است. یک نمونه تصادفی دیگر به اندازه ۴۹ از یک جمعیت دیگر با انحراف معیار ۳/۴ دارای میانگین ۷۶ است. آیا در سطح معنی‌دار ۰/۰۶ میانگین این دو جمعیت با هم برابر است؟

حل در این مثال با آزمون $\begin{cases} H_۰ : \mu_۱ = \mu_۲ \\ H_۱ : \mu_۱ \neq \mu_۲ \end{cases}$ مواجه هستیم که واریانسها معلوم می‌باشند. بنابراین از آزمون شماره ۱۲ با $d_۰ = ۰$ استفاده می‌کنیم. از اطلاعات مسئله داریم که

$$n_۱ = ۳۶ \quad , \quad \sigma_۱ = ۵/۲ \quad , \quad \bar{x}_۱ = ۸۱$$

$$n_۲ = ۴۹ \quad , \quad \sigma_۲ = ۳/۴ \quad , \quad \bar{x}_۲ = ۷۶$$

بنابراین

$$Z = \frac{(\bar{x}_۱ - \bar{x}_۲) - d_۰}{\sqrt{\frac{\sigma_۱^۲}{n_۱} + \frac{\sigma_۲^۲}{n_۲}}} = \frac{۸۱ - ۷۶ - ۰}{\sqrt{\frac{(۵/۲)^۲}{۳۶} + \frac{(۳/۴)^۲}{۴۹}}} = ۵/۰۳۳$$

$$\alpha = ۰/۰۶ \Rightarrow z_{۱-\frac{\alpha}{2}} = z_{۰/۹۷} = ۱/۸۸$$

چون $۱/۸۸ = z_{۱-\frac{\alpha}{2}} < |Z| = ۵/۰۳۳$ پس فرض $H_۰$ رد می‌شود، یعنی $\mu_۱ \neq \mu_۲$ و چون $۱/۸۸ < Z = ۵/۰۳۳$ بنابراین نتیجه می‌شود که $\mu_۱ > \mu_۲$ است.

مثال ۶.۲.۸ دو گروه ۴۰ نفری برای انجام یک آزمایش انتخاب شده‌اند، گروه اول را با رژیم غذایی A و گروه دوم را با رژیم غذایی B مورد آزمایش قرار داده‌ایم. میانگین و انحراف استاندارد کاهش وزن در رژیم غذایی A به ترتیب ۱۱ و ۴/۳ کیلوگرم و در رژیم غذایی B به ترتیب ۸ و ۵/۷ کیلوگرم بوده است. در سطح معنی‌دار ۰/۰۵ آیا می‌توان ادعا کرد که میانگین کاهش وزن بوسیله رژیم غذایی A از میانگین کاهش وزن بوسیله رژیم غذایی B به اندازه حداقل یک کیلوگرم بیشتر است؟ فرض کنید کاهش وزن دو نوع رژیم دارای توزیع نرمال با واریانسهای مساوی باشند.

حل در این مثال با آزمون $\begin{cases} H_۰ : \mu_A \leq \mu_B + ۱ \\ H_۱ : \mu_A > \mu_B + ۱ \end{cases}$ مواجه هستیم. بنابراین از آزمون شماره ۱۳ با $d_۰ = ۱$ استفاده می‌کنیم. از اطلاعات مسئله داریم که

$$n_۱ = ۴۰ \quad , \quad \bar{x}_۱ = ۱۱ \quad , \quad s_۱ = ۴/۳$$

$$n_۲ = ۴۰ \quad , \quad \bar{x}_۲ = ۸ \quad , \quad s_۲ = ۵/۷$$

بنابراین

$$s_P^۲ = \frac{۳۹(۴/۳)^۲ + ۳۹(۵/۷)^۲}{۷۸} = ۲۵/۴۹$$

$$T = \frac{(\bar{x}_۱ - \bar{x}_۲) - d_۰}{s_P\sqrt{\frac{۱}{n_۱} + \frac{۱}{n_۲}}} = \frac{۱۱ - ۸ - ۱}{\sqrt{۲۵/۴۹}\sqrt{\frac{۱}{۴۰} + \frac{۱}{۴۰}}} = ۱/۷۷۲$$

$$\alpha = ۰/۰۵ \Rightarrow t_{۱-\alpha}(n_۱ + n_۲ - ۲) = t_{۰/۹۵}(۷۸) = Z_{۰/۹۵} = ۱/۶۴$$

چون $۱/۷۷۲ = T > t_{۱-\alpha}(n_۱ + n_۲ - ۲) = ۱/۶۴$ بنابراین فرض $H_۰$ رد می‌شود و ادعای گفته شده درست می‌باشد.

مثال ۷.۲.۸ یک درس را به دو روش تدریس نموده‌ایم. سپس در روش اول از ۱۶ نفر امتحان به عمل آورده و انحراف استاندارد نمرات ۹ به دست آمد، در حالیکه در روش دوم از ۲۵ نفر امتحان بعمل آورده و انحراف استاندارد نمرات ۱۲ به دست آمده است. آیا در سطح معنی‌دار ۰/۰۱ می‌توان متقاعد شد که پراکندگی نمرات در روش اول کمتر از روش دوم است؟ فرض کنید نمرات دو روش از توزیع نرمال پیروی می‌کنند.

حل در این مثال با آزمون $\begin{cases} H_۰ : \sigma_۱^۲ \geq \sigma_۲^۲ \\ H_۱ : \sigma_۱^۲ < \sigma_۲^۲ \end{cases}$ مواجه هستیم. بنابراین از آزمون شماره ۱۷ استفاده می‌کنیم. از اطلاعات مسئله داریم که

$$n_۱ = ۱۶ \; , \; s_۱ = ۹ \quad , \quad n_۲ = ۲۵ \quad , \quad s_۲ = ۱۲$$

$$F = \frac{s_۱^۲}{s_۲^۲} = \left(\frac{۹}{۱۲}\right)^۲ = ۰/۵۶۲۵$$

$$\alpha = ۰/۰۱ \Rightarrow \quad F_\alpha(n_۱ - ۱, n_۲ - ۱) = \frac{۱}{F_{۱-\alpha}(n_۲ - ۱, n_۱ - ۱)} = \frac{۱}{F_{۰/۹۹}(۲۴, ۱۵)} = ۰/۳۰۴$$

چون $۰/۵۶۲۵ = F \not< F_\alpha(n_۱ - ۱, n_۲ - ۱) = ۰/۳۰۴$ پس فرض $H_۰$ رد نمی‌شود، یعنی پراکندگی نمرات در روش اول کمتر نیست.

مثال ۸.۲.۸ ادعا شده است که وزن قوطی‌های روغنی بخصوص ۱۰ انس است. اگر وزنهای یک نمونه تصادفی ۱۰تایی از این قوطیها به صورت زیر باشد و وزن قوطیها دارای توزیع نرمال باشد، آیا این ادعا را می‌پذیرید؟

۹/۸ , ۱۰/۳ , ۱۰/۴ , ۹/۹ , ۹/۸ , ۱۰/۱ , ۱۰/۳ , ۱۰/۱ , ۹/۷ , ۱۰/۲

حل در این مثال با آزمون $\begin{cases} H_۰ : \mu = ۱۰ \\ H_۱ : \mu \neq ۱۰ \end{cases}$ مواجه هستیم که در آن $\sigma^۲$ نامعلوم است. بنابراین از آزمون شماره ۶ با $\mu_۰ = ۱۰$ و $\alpha = ۰/۰۵$ استفاده می‌کنیم. از اطلاعات مسئله داریم که

$$n = ۱۰ \ , \quad \sum x_i = ۱۰۰/۶ \ , \quad \sum x_i^۲ = ۱۰۱۲/۵۸$$

در نتیجه $\bar{x} = ۱۰/۰۶$ و s=۰/۲۴۶ و همچنین

$$\alpha = ۰/۰۵ \Rightarrow t_{۱-\frac{\alpha}{۲}}(n-۱) = t_{۰/۹۷۵}(۹) = ۲/۲۶$$

$$T = \frac{\bar{x} - \mu_۰}{s/\sqrt{n}} = \frac{۱۰/۰۶ - ۱۰}{\frac{۰/۲۴۶}{\sqrt{۱۰}}} = ۰/۷۷۱$$

چون $۰/۷۷۱ = |T| \ngtr t_{۱-\frac{\alpha}{۲}}(n-۱) = ۲/۲۶$ پس $H_۰$ رد نمی‌شود یعنی ادعا درست می‌باشد.

مثال ۹.۲.۸ در یک آزمون حساب سال پنجم ابتدایی نمره ۸ دانش‌آموز پسر و ۶ دانش‌آموز دختر به صورت زیر به دست آمده است.

پسرها	۱۰	۱۷	۱۴	۱۲	۱۱	۱۶	۱۸	۱۹
دخترها	۱۶	۱۸	۱۷	۱۳	۱۴	۱۱		

با فرض نرمال بودن نمره‌ها و تساوی واریانسها، آیا به طور متوسط نمرات دانش‌آموزان پسر و دختر یکسان است؟

حل در این مثال با آزمون $\begin{cases} H_۰ : \mu_۱ = \mu_۲ \\ H_۱ : \mu_۱ \neq \mu_۲ \end{cases}$ مواجه هستیم. بنابراین از آزمون شماره ۱۵ با $d_۰ = ۰$ و $\alpha = ۰/۰۵$ استفاده می‌کنیم. از اطلاعات مسئله به دست می‌آوریم که

$$n_۱ = ۸ \ , \quad \bar{x}_۱ = ۱۴/۶۲۵ \ , \quad s_۱^۲ = ۱۱/۴۱$$
$$n_۲ = ۶ \ , \quad \bar{x}_۲ = ۱۴/۸۳ \ , \quad s_۲^۲ = ۶/۹۷$$

بنابراین

$$s_P^۲ = \frac{۷(۱۱/۴۱) + ۵(۶/۹۷)}{۱۲} = ۹/۵۶$$

$$T = \frac{\bar{x}_۱ - \bar{x}_۲ - d_۰}{s_P \sqrt{\frac{۱}{n_۱} + \frac{۱}{n_۲}}} = \frac{۱۴/۶۲۵ - ۱۴/۸۳ - ۰}{\sqrt{۹/۵۶}\sqrt{\frac{۱}{۸} + \frac{۱}{۶}}} = -۰/۱۲۳$$

$$\alpha = ۰/۰۵ \Rightarrow t_{۱-\frac{\alpha}{۲}}(n_۱ + n_۲ - ۲) = t_{۰/۹۷۵}(۱۲) = ۲/۱۸$$

چون $۰/۱۲۳ = |T| \ngtr t_{۱-\frac{\alpha}{۲}}(n_۱ + n_۲ - ۲) = ۲/۱۸$ پس فرض $H_۰$ رد نمی‌شود یعنی به طور متوسط نمرات دانش‌آموزان پسر و دختر یکسان است.

۳.۸ آزمون برازندگی

در آزمونهایی که تاکنون در این فصل انجام داده‌ایم فرض کرده‌ایم که جمعیت دارای یک توزیع احتمال بخصوص است و در مورد پارامترهای مجهول جمعیت مانند μ و $\sigma^۲$ آزمونهایی را انجام داده‌ایم. در این بخش حالتی را در نظر می‌گیریم که در آن توزیع احتمال جمعیت خود نیز مجهول می‌باشد و ما به جمعیت توزیع مشخصی را نسبت داده و آن را مورد آزمون قرار می‌دهیم. یعنی آزمونی را در نظر می‌گیریم که تعیین می‌کند آیا جمعیتی دارای یک توزیع مشخص است یا نه. در بسیاری موارد یک توزیع احتمال برای یک سری مشاهدات در نظر گرفته می‌شود، مثلاً توزیع دو جمله‌ای برای درصدی از جمعیت که دارای خصیصه معینی هستند و یا توزیع یکنواخت برای هم شانس بودن انتخاب افراد جمعیت یا توزیع پواسون برای تعداد غلطهای چاپی در صفحات کتاب و ... می‌خواهیم بدانیم که مشاهدات به دست آمده بر اساس یک نمونه تصادفی از جمعیت، با یک توزیع مفروض مطابقت دارد یا نه؟

توزیعی را که حدس می‌زنیم داده‌ها از آن باشند، توزیع برازنده بر داده‌ها و آزمون لازم را آزمون برازندگی می‌نامند. یکی از این آزمونهای برازندگی، **آزمون مربع-کای** برای برازندگی توزیع می‌باشد که در زیر آن را می‌آوریم.

فرض کنید که نتیجه یک آزمایش تصادفی به یکی از k طبقه دو به دو مجزای $C_۱$، $C_۲$، ... و C_k متعلق باشد به طوری که احتمال متعلق بودن به طبقه C_i برابر مقدار $۰ < p_i < ۱$ باشد که $\sum_{i=۱}^{k} p_i = ۱$. این آزمایش تصادفی را n مرتبه مستقلاً انجام می‌دهیم و قرار می‌دهیم.

O_i = تعداد دفعاتی از این n آزمایش که نتیجه آزمایش به طبقه C_i متعلق باشد $i=1,2,\ldots,k$

در این صورت $i=1,2,\ldots,k$, $O_i \sim B(n,p_i)$ و متغیر تصادفی O_i را فراوانی مشاهده شده[1] می‌نامند. می‌خواهیم این فرض را آزمون کنیم که قانون احتمال آزمایش تصادفی بوسیله p_i° ها مشخص می‌شود یا نه، یعنی

$$\begin{cases} H_0 : p_1 = p_1^\circ, \ldots, p_k = p_k^\circ \\ H_1 : p_i \neq p_i^\circ \quad \text{برای حداقل یک } i=1,2,\ldots,k \end{cases}$$

که در آن p_i° ها مقادیر معینی هستند و $1 > p_i^\circ > 0$ و $\sum_{i=1}^{k} p_i^\circ = 1$. اگر فرض H_0 درست باشد آنگاه $E(O_i)=np_i^\circ$، یعنی اگر H_0 درست باشد و آزمایش را n بار مستقلاً تکرار کنیم، انتظار داریم که در این n آزمایش به طور متوسط فراوانی نتایجی که به طبقه C_i تعلق دارند برابر $e_i = E(O_i) = np_i^\circ$ باشد. عدد e_i را فراوانی مورد انتظار[2] می‌نامند. واضح است که

$$\sum_{i=1}^{k} O_i = \sum_{i=1}^{k} e_i = n$$

برای انجام آزمون H_0 در مقابل H_1 از آماره زیر که به آماره X^2 پیرسون مشهور است استفاده می‌کنیم.

$$\boxed{X^2 = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i}} \qquad (1.8)$$

می‌توان نشان داد که به طور تقریبی $X^2 \sim \chi^2_{(k-1)}$. واضح است که فرض H_0 موقعی پذیرفته می‌شود که اختلاف O_i ها و e_i ها و در نتیجه X^2 کوچک باشد. بنابراین ناحیه بحرانی به صورت $X^2 > c$ خواهد بود. اگر بخواهیم سطح معنی‌دار آزمون برابر α باشد آنگاه ناحیه بحرانی به صورت $X^2 > \chi^2_{1-\alpha}(k-1)$ تبدیل می‌شود، یعنی

> در آزمون برازندگی، فرض H_0 (برازندگی توزیع) رد می‌شود اگر و فقط اگر
> $X^2 > \chi^2_{1-\alpha}(k-1)$

مثال ۱.۳.۸ یک تاس را ۱۲۰ بار پرتاب می‌کنیم و مشاهدات زیر را به دست می‌آوریم. در سطح معنی‌دار $\alpha = 0/05$ آزمون کنید که آیا تاس سالم است

1- observed frequency   2- Expected frequency

شماره تاس	۱	۲	۳	۴	۵	۶	جمع
فراوانی	۱۸	۲۳	۱۶	۲۱	۱۸	۲۴	۱۲۰

حل در این مثال با آزمون زیر مواجه هستیم

$$\begin{cases} H_0 : p_1 = p_2 = \ldots = p_6 = \frac{1}{6} \\ H_1 : p_i \neq \frac{1}{6} \quad i=1,2,\ldots,6 \quad \text{برای حداقل یک} \end{cases}$$

بنابراین $i=1,\ldots,6$, $e_i = np_i^\circ = 120 \times \frac{1}{6} = 20$ و در نتیجه جدول زیر را به دست می‌آوریم

i	۱	۲	۳	۴	۵	۶	جمع
O_i	۱۸	۲۳	۱۶	۲۱	۱۸	۲۴	۱۲۰
e_i	۲۰	۲۰	۲۰	۲۰	۲۰	۲۰	۱۲۰

و از این جدول داریم که

$$X^2 = \sum_{i=1}^{6} \frac{(o_i - e_i)^2}{e_i} = \frac{(18-20)^2}{20} + \ldots + \frac{(24-20)^2}{20} = 2/5$$

$$\chi^2_{1-\alpha}(k-1) = \chi^2_{0/95}(5) = 11/1$$

چون $11/1 = \chi^2_{1-\alpha}(k-1) \not< X^2 = 2/5$ پس فرض H_0 رد نمی‌شود، یعنی تاس سالم است.

تذکر ۱ آزمون برازندگی را می‌توان در مواردی که مقادیر مورد انتظار e_i ها (و یا p_i° ها) به پارامترهای نامعلوم θ_1، θ_2، ... و θ_t وابسته هستند، نیز بکار برد. برای این منظور ابتدا θ_1، θ_2، ... و θ_t را بوسیله مشاهدات O_1، O_2، ... و O_k برآورد می‌کنیم، سپس به وسیلهٔ برآوردهای $\hat\theta_1$، $\hat\theta_2$، ... و $\hat\theta_t$ مقادیر مورد انتظار e_1، e_2، ... و e_k را محاسبه و در فرمول X^2 قرار می‌دهیم. در این حالت فرض H_0 را رد خواهیم کرد اگر و فقط اگر $X^2 > \chi^2_{1-\alpha}(k-1-t)$ که در آن t تعداد پارامترهایی است که توسط مشاهدات برآورد شده‌اند.

تذکر ۲ اگر در آزمون برازندگی مقدار بعضی از مقادیر مورد انتظار کوچکتر از ۵ باشد، بایستی چند طبقهٔ مجاور را با هم ادغام کنیم، تا جمع مقادیر مورد انتظار طبقات جدید بزرگتر یا مساوی ۵ شود.

مثال ۲.۳.۸ تعداد غلطهای چاپی در ۱۰۰ صفحه یک کتاب را شمرده‌ایم و مشاهدات در جدول

زیر آورده شده است. آیا توزیع پواسون در سطح معنی دار ۰/۰۵ بر داده‌ها برازنده است.

تعداد غلطها i	۰	۱	۲	۳	۴	۵	۶
تعداد صفحات o_i	۳۶	۴۰	۱۹	۲	۰	۲	۱

حل اگر X تعداد غلطهای چاپی در یک صفحه کتاب باشد آنگاه آزمون $\begin{cases} H_0 : X\sim P(\mu) \\ H_1 : X\sim P(\mu) \end{cases}$ مورد نظر است. در ابتدا به وسیلهٔ مشاهدات، μ میانگین توزیع پواسون را برآورد می‌کنیم.

$$\hat{\mu}=\bar{x}=\frac{1}{n}\sum_{i=1}^{k} f_i x_i=\frac{1}{100}\left[(0\times 36)+(1\times 40)+\dots+(6\times 1)\right]=1$$

بنابراین تحت فرض H_0 داریم که

$$f_X(x)=\frac{e^{-1}1^x}{x!}=\frac{1}{x!}e^{-1} \qquad x=0,1,2,\dots$$

در نتیجه

$$p_0^\circ=P(X=0)=e^{-1}=0/3679 \;\Rightarrow\; e_0=np_0^\circ=100(0/3679)=36/79$$

$$p_1^\circ=P(X=1)=e^{-1}=0/3679 \;\Rightarrow\; e_1=36/79$$

با محاسبهٔ مقادیر دیگر e_i به طور مشابه، جدول زیر را به دست می‌آوریم.

i	۰	۱	۲	۳	۴	۵	۶	جمع
o_i	۳۶	۴۰	۱۹	۲	۰	۲	۱	۱۰۰
e_i	۳۶/۷۹	۳۶/۷۹	۱۸/۳۹	۶/۱۳	۱/۵۳	۰/۳۱	۰/۰۵	۱۰۰

چون ۳ طبقهٔ آخر دارای مقادیر مورد انتظار کمتر از ۵ هستند پس ۴ طبقهٔ آخر را با هم ادغام می‌کنیم و جدول زیر به دست می‌آید

i	۰	۱	۲	بزرگتر از ۲	جمع
o_i	۳۶	۴۰	۱۹	۵	۱۰۰
e_i	۳۶/۷۹	۳۶/۷۹	۱۸/۳۹	۸/۰۲	۱۰۰

بنابراین

$$X^2=\sum_{i=0}^{3}\frac{(o_i-e_i)^2}{e_i}=\frac{(36-36/79)^2}{36/79}+\dots+\frac{(5-8/02)^2}{8/02}=1/454$$

$$d.f.=k-1-t=4-1-1=2 \;\Rightarrow\; \chi^2_{1-\alpha}(k-1-t)=\chi^2_{0/95}(2)=5/99$$

چون $1/454=X^2 \not> \chi^2_{1-\alpha}(k-1-t)=5/99$ پس H_0 رد نمی‌شود، یعنی توزیع پواسون بر داده‌ها برازنده است.

مثال ۳.۳.۸ طول عمر ۱۰۰۰ لامپ یک کارخانه را اندازه‌گیری کرده‌ایم و اطلاعات جدول زیر به دست آمده است همچنین $\sum x_i=200000$. سرپرست کارخانه ادعا دارد که طول عمر لامپها دارای توزیع نمایی است. آیا ادعای او را در سطح معنی‌دار ۰/۰۱ می‌پذیرید.

تعداد	طول عمر t
۵۴۳	t≤۱۵۰
۲۵۸	۱۵۰<t≤۳۰۰
۱۲۰	۳۰۰<t≤۴۵۰
۴۸	۴۵۰<t≤۶۰۰
۲۰	۶۰۰<t≤۷۵۰
۱۱	۷۵۰<t

حل اگر X طول عمر لامپ تولیدی کارخانه باشد آنگاه آزمون $\begin{cases} H_0 : X\sim E(\theta) \\ H_1 : X\nsim E(\theta) \end{cases}$ مورد نظر است. در ابتدا θ میانگین توزیع نمایی را بوسیله مشاهدات برآورد می‌کنیم.

$$\hat{\theta}=\bar{x}=\frac{1}{n}\sum x_i=\frac{200000}{1000}=200$$

بنابراین تحت فرض H_0 داریم که

$$f_X(x)=\frac{1}{200}e^{-\frac{1}{200}x} \qquad x>0$$

در نتیجه

$$p_1^\circ=P(X\le 150)=\int_0^{150}\frac{1}{200}e^{-\frac{1}{200}x}dx=0/5277 \;\Rightarrow\; e_1=1000p_1^\circ=527/7$$

$$p_2^\circ=P(150<X\le 300)=\int_{150}^{300}\frac{1}{200}e^{-\frac{1}{200}x}dx=0/2492 \;\Rightarrow\; e_2=249/2$$

با محاسبهٔ مقادیر دیگر e_i به طور مشابه، جدول زیر را به دست می‌آوریم.

i	۱	۲	۳	۴	۵	۶	جمع
o_i	۵۴۳	۲۵۸	۱۲۰	۴۸	۲۰	۱۱	۱۰۰۰
e_i	۵۲۷/۷	۲۴۹/۲	۱۱۷/۷	۵۵/۶	۲۶/۳	۲۳/۵	۱۰۰۰

بنابراین

$$X^2=\sum_{i=1}^{۶}\frac{(o_i-e_i)^2}{e_i}=\frac{(۵۴۳-۵۲۷/۷)^2}{۵۲۷/۷}+\ldots=۹/۹۹۶$$

$$d.f.=k-۱-t=۶-۱-۱=۴ \quad\Rightarrow\quad \chi^2_{۱-\alpha}(k-۱-t)=\chi^2_{۰/۹۹}(۴)=۱۳/۳$$

چون $۹/۹۹۶=X^2\not>\chi^2_{۱-\alpha}(k-۱-t)=۱۳/۳$ پس $H_۰$ رد نمی‌شود، یعنی توزیع نمایی برازنده بر داده‌ها است.

۴.۸ تمرینات

۱ در هر کدام از حالتهای زیر، فرض صفر و فرض مقابل را مشخص کنید.

الف - یک تولیدکنندهٔ اتومبیل می‌خواهد ادعای تهیه کننده‌ای را بررسی کند که حداکثر مقاومت سیمهایی که می‌سازد کمتر از ۵۰ اهم باشد.

ب - ادارهٔ تحقیقات ادعا می‌کند رشته‌هایی که برای لامپها درست کرده است، عمر متوسط لامپها را تا بیش از ۳۰۰ ساعت افزایش خواهد داد. یک بازرس علاقه‌مند به بررسی این ادعا می‌باشد.

۲ می‌خواهیم برای یک سکه آزمون $\begin{cases}H_۰: p=\frac{۱}{۲}\\ H_۱: p=\frac{۳}{۴}\end{cases}$ را انجام دهیم. سکه را ده بار پرتاب می‌کنیم و X را تعداد شیرها در این ده پرتاب در نظر می‌گیریم. اگر $X\geq ۸$ فرض $H_۰$ را رد می‌کنیم.

الف - احتمال خطای نوع اول و دوم و توان آزمون را محاسبه کنید.

ب - اگر $X\geq c$ را ناحیهٔ بحرانی بگیریم، c را چنان تعیین کنید تا احتمال خطای نوع دوم بیش از ۰/۰۹ نباشد.

۳ نسبت خانواده‌های ساکن در شهر بخصوصی که از کمپانی A شیر می‌خرند ۰/۶ است. اگر از یک نمونهٔ تصادفی ۱۰ خانواری ۳ یا کمتر از کمپانی A شیر بخرند فرضیهٔ صفر $p=۰/۶$ را به نفع فرضیهٔ مقابل $p<۰/۶$ رد می‌کنیم. احتمال خطای نوع اول را محاسبه نمایید. احتمال خطای نوع دوم را برای مقادیر $p=۰/۳$، $p=۰/۴$ و $p=۰/۵$ محاسبه نمایید.

۴ در صورتی که از یک جمعیت نرمال با واریانس ۴ یک نمونهٔ تصادفی به حجم ۱۶ انتخاب کرده باشیم، در آزمون $\begin{cases}H_۰: \mu=۸\\ H_۱: \mu=۱۰\end{cases}$ در سطح معنی‌دار ۰/۰۵ احتمال خطای نوع دوم را

محاسبه کنید.

۵ فرض کنید متغیر تصادفی X دارای تابع چگالی زیر باشد

$$f_\theta(x)=\begin{cases}\frac{۱}{\theta}e^{-\frac{۱}{\theta}x} & x>۰,\theta>۰\\ ۰ & \text{سایر نقاط}\end{cases}$$

می‌خواهیم آزمون $\begin{cases}H_۰: \theta=۲۰۰\\ H_۱: \theta=۵۰۰\end{cases}$ را انجام دهیم. اگر $X>۳۰۰$ مشاهده شود، فرض $H_۰$ را رد می‌کنیم. احتمال خطای نوع اول و دوم و توان آزمون را محاسبه کنید.

۶ متوسط قد دانشجویان مرد سال اول یک دانشکده بخصوص ۶۸/۵ اینچ با انحراف معیار ۲/۷ اینچ گزارش شده است. اگر یک نمونهٔ تصادفی ۵۰ تایی از دانشجویان سال اول فعلی دارای حد متوسط قد ۶۹/۷ اینچ باشد، آیا در سطح معنی‌دار ۰/۰۲ دلیلی برای تصور تغییر در حد متوسط قد وجود دارد؟

۷ وزارت کار و امور اجتماعی، مزد روزانهٔ کارگران کارخانه را به طور متوسط ۱۳۲۰ تومان با انحراف ۲۵۰ تومان تعیین نموده است. اگر کارخانه‌ای به ۴۰ کارگر خود روزانه به طور متوسط ۱۲۲۰ تومان پرداخت نماید، آیا می‌توان این کارخانه را متهم نمود که کمتر از مزد تعیین شدهٔ وزارت کار و امور اجتماعی پرداخت می‌کند.

۸ لامپهای تلویزیون موجود در بازار دارای میانگین طول عمر ۱۲۰۰ ساعت با انحراف معیار ۳۰۰ ساعت هستند. کارخانه‌ای لامپ جدیدی تولید کرده و مدعی است که میانگین طول عمر لامپهای ساخت کارخانه‌اش بیشتر از ۱۲۰۰ ساعت است. برای بررسی این ادعا یک نمونه ۱۰۰ تایی انتخاب و میانگین طول عمر ۱۲۶۵ ساعت به دست آمده است. آیا با ادعای صاحب کارخانه موافق هستید؟

۹ یک فرایند تولید رنگ موجود است که توزیع تولید روزانه آن نرمال با میانگین ۸۰۰ و انحراف معیار ۳۰ تن است. به منظور ازدیاد تولید اصلاحاتی در این فرایند پیشنهاد شده است و یک نمونهٔ تصادفی ۱۰۰ روزه از تولید فرایند اصلاح شده دارای میانگین ۸۱۲ تن می‌باشد. در سطح معنی‌دار ۰/۰۱ آیا فرایند اصلاح شده میانگین تولید روزانه را افزایش می‌دهد؟

۱۰ نمرات امتحان هوش یک دبیرستان دارای توزیع نرمال با انحراف معیار ۶ است. از یک گروه ۹ نفری امتحان هوش به عمل آورده‌ایم و نتایج زیر به دست آمده است

۶۵ , ۶۴ , ۶۶ , ۶۵ , ۵۹ , ۶۹ , ۶۴ , ۷۲ , ۶۳

مدیر دبیرستان ادعا می‌کند که میانگین نمرات هوش کل دبیرستان از ۶۵ بیشتر است. در سطح‌معنی‌دار ۰/۰۵ آیا ادعای مدیر دبیرستان را قبول می‌کنید؟

۱۱ مطالعهٔ آماری در گذشته نشان داده است که میزان غیبت کارمندان بر اثر بیماری در عرض سال دارای توزیع نرمال با میانگین ۱۵ روز است. پژوهشگری برای ۲۵ کارمند در سال گذشته شمارهٔ روزهای غیبت را به شرح زیر ثبت کرده است

۵	۲۵	۱۰	۰	۳	۵۰	۱۲	۱۴	۴۰	۱۲	۳۲	۵	
۸	۴	۴۷	۲۰	۱۴	۱۸	۱۶	۱۰	۱	۲۲	۵۸	۲۳	۹

در سطح معنی‌دار ۰/۰۵ آزمون کنید که در سال گذشته متوسط روزهای غیبت بیش از ۱۵ روز است.

۱۲ یک کارخانهٔ مواد شیمیایی به گونه‌ای طراحی شده است که روزانه به طور متوسط ۸۰۰ تن محصول داشته باشد. در ۵ روز متوالی، این کارخانه به ترتیب ۷۸۵، ۸۰۵، ۷۹۰، ۷۹۳ و ۸۰۲ تن محصول داشته است. با فرض نرمال بودن میزان محصول، آیا این داده‌ها نشان دهندهٔ کاهش در حد متوسط میزان محصول کارخانه است؟

۱۳ وزن یک نمونهٔ تصادفی ۱۰ تایی از حلب‌های روغن بر حسب کیلوگرم به صورت زیر گزارش شده است

۱۰/۱ , ۱۰/۳ , ۹/۷ , ۱۰/۱ , ۹/۸ , ۱۰/۲ , ۱۰/۳ , ۱۰/۴ , ۹/۸ , ۹/۹

به فرض نرمال بودن وزن حلب‌های روغن، آیا در سطح معنی‌دار ۰/۰۱، می‌توان متقاعد شد که میانگین وزن این حلب‌های روغن ۱۰ کیلو است؟

۱۴ طبق یک الگوی فیزیکی، میانگین افزایش درجهٔ حرارت آبی که به عنوان خنک کننده در یک اتاقک کمپرسور به کار می‌رود، نباید بیشتر از ۵ درجهٔ سانتیگراد باشد. افزایش درجهٔ حرارت در هشت بار استفادهٔ مستقل از کمپرسور، اندازه‌گیری شده و داده‌های زیر به دست آمده‌اند

۵/۱ , ۶/۴ , ۵/۹ , ۶/۵ , ۴/۹ , ۵/۷ , ۶/۳ , ۶/۴

اگر افزایش درجهٔ حرارت دارای توزیع نرمال باشد، آیا این داده‌ها با الگوی فیزیکی متناقض هستند؟

۱۵ به منظور بررسی رشد جسمی کودکان ۱۰ ساله اندازهٔ قامت یک نمونه ۲۰ تایی از این کودکان اندازه‌گیری شد و نتایج زیر به دست آمد

۱۰۰ ۱۰۵ ۱۰۴ ۱۱۳ ۹۷ ۱۰۴ ۱۱۰ ۱۰۳ ۱۰۹ ۱۰۵

۱۰۶ ۱۰۴ ۱۰۵ ۱۱۱ ۹۸ ۱۰۶ ۱۰۳ ۱۱۰ ۱۰۲ ۱۰۸

با فرض نرمال بودن طول قد کودکان ۱۰ ساله مطلوب است

الف – در سطح معنی‌دار ۰/۰۵ آیا می‌توان ادعا کرد که میانگین واقعی قد کودکان ۱۰ ساله غیر از ۱۰۴ است؟

ب – در سطح معنی‌دار ۰/۱ آیا می‌توان ادعا کرد که واریانس واقعی قد کودکان ۱۰ ساله بیشتر از ۵ است؟

۱۶ طول فنرهایی که یک شرکت تولید می‌کند دارای توزیع نرمال با انحراف معیار ۶ سانتیمتر است. سازندهٔ این فنرها ادعا می‌کند که متوسط طول فنرها بیش از ۶۵ سانتیمتر می‌باشد. توزیع کننده‌ای که قصد خرید فنرها را دارد مایل به آزمون ادعای سازنده است. از نظر خریدار خریدن فنرهایی که متوسط طول آنها حداکثر ۶۵ سانتیمتر است مقرون به صرفه نیست. بدین منظور نمونه‌ای مرکب از ۹ فنر را انتخاب می‌کند و طول آنها را اندازه‌گیری می‌کند.

الف – در مورد میانگین طول فنرها، فرضها را فرمول بندی و ناحیهٔ بحرانی آزمون را در سطح معنی‌دار ۰/۰۵ تعیین کنید.

ب – اگر $\mu=$۶۷ احتمال خطای نوع دوم و توان آزمون را محاسبه کنید.

ج – اگر نتایج حاصله از بررسی ۹ فنر به قرار زیر باشد، خریدار چه تصمیمی می‌گیرد

۶۳ ۷۲ ۶۴ ۶۹ ۵۹ ۶۵ ۶۶ ۶۴ ۶۵

۱۷ یک نمونهٔ تصادفی ۸ تایی از یک نوع سیگار دارای حد متوسط نیکوتین ۱۸/۶ میلی گرم با انحراف استاندارد ۲/۴ میلی گرم است. آیا این مطلب با ادعای تولید کننده مبنی بر متجاوز نبودن انحراف معیار نیکوتین از ۶ میلی گرم توافق دارد؟ سطح معنی‌دار را ۰/۰۱ بگیرید، و فرض کنید توزیع نیکوتین در سیگارها از توزیع نرمال پیروی می‌کند؟

۱۸ متوسط درجهٔ حرارت سالانه یک شهر از میانگین متوسط درجهٔ حرارت پانزدهمین روز هر ماه سال اندازه‌گیری شده است. انحراف معیار درجه حرارت سالانه شهری در یک دوره ۱۰۰ ساله ۱۶ درجهٔ فارنهایت بوده است. در مدت ۱۵ سال گذشته انحراف استاندارد درجهٔ حرارت سالانه را محاسبه نموده‌اند که برابر ۱۸ درجهٔ فارنهایت بوده است. آیا این اطلاعات دلیلی بر این است که انحراف معیار درجهٔ حرارت سالیانهٔ این شهر از انحراف معیار درجهٔ حرارت سابق شهر بیشتر است؟

۱۹ آزمایشگری معتقد است که واریانس اندازه‌هایی که در طول آزمایش ثبت می‌کند کوچکتر از ۲ است. در یک آزمایش او اندازه‌های ۴/۱، ۵/۲ و ۱۰/۲ را ثبت کرده است. اگر اندازه‌ها دارای توزیع نرمال باشند آیا می‌توان ادعای آزمایشگر را در سطح معنی‌دار ۰/۰۱ پذیرفت؟

۲۰ قطعهٔ زمینی بوسیلهٔ ۵ دانشجوی نقشه‌بردار، مساحی شده و مساحتهای زیر بر حسب جریب به دست آمده‌اند

۷/۲۷ ۷/۲۴ ۷/۲۱ ۷/۲۸ ۷/۲۳

با فرض نرمال بودن اندازه‌گیری، در سطح معنی‌دار ۰/۰۵ آیا می‌توان ادعا کرد که واریانس اندازه‌گیری ۰/۹ است؟

۲۱ از یک جمعیت نرمال یک نمونه تصادفی ۱۶ تایی انتخاب شده و $\sum x_i = 24$ و $\sum x_i^2 = 40$ به دست آمده است. در سطح معنی‌دار ۰/۰۵ آیا می‌توان ادعا کرد که واریانس این جمعیت بیش از یک است؟

۲۲ نمرات دانشجویان در کلاس آمار متغیر تصادفی X با توزیع $N(\mu_1, 16)$ و نمرات دانشجویان در کلاس ریاضی متغیر تصادفی Y با توزیع $N(\mu_2, 25)$ می‌باشد و X و Y از یکدیگر مستقل هستند. نمونهٔ تصادفی ۶۴ تایی از نمرات آمار دارای میانگین ۸۰ و نمونهٔ تصادفی ۶۵ تایی از نمرات ریاضی دارای میانگین ۸۵ می‌باشد. آیا در سطح معنی‌دار ۰/۰۱ می‌توان گفت که میانگین نمرات دو کلاس یکسان است؟

۲۳ تحلیل‌گری می‌خواهد میانگین طول عمر عاج یک نوع لاستیک اتومبیل را در حالتی که فشار باد لاستیک بیشتر از حد استاندارد است، با هم مقایسه کند. او دو نمونهٔ تصادفی مستقل، مرکب از ۱۵ لاستیک را از خط تولید انتخاب کرده است. لاستیکهای نمونهٔ ۱ را با فشار باد استاندارد و

لاستیکهای نمونه ۲ را با فشار باد بیش از حد استاندارد تنظیم کرده و میانگین طول عمر عاج لاستیکها بر حسب هزار کیلومتر به ترتیب $\bar{x}_1 = 43/0$ و $\bar{x}_2 = 40/7$ به دست آمده است. اگر هر دو جمعیت دارای توزیع نرمال با واریانس مساوی $\sigma_1^2 = \sigma_2^2 = 1/2$ باشند، آیا در سطح معنی‌دار ۰/۰۱ می‌توان وجود اختلاف بین میانگین طول عمر لاستیکها را در شرایط گفته شده پذیرفت؟

۲۴ با استفاده از تجربهٔ گذشته می‌دانیم که دو ماشین A و B که هر دو یک نوع ریسمان تولید می‌کنند، به ترتیب دارای انحراف معیار ۰/۰۴ و ۰/۰۳ هستند. وضعیت دو ماشین تغییر داده می‌شود و علاقه دارند که وضع هر دو تولید یکی شود. برای بررسی آن، ۱۰ تکه ریسمان از ماشین A و ۱۵ تکه ریسمان از ماشین B برداشته و میانگینهای ۲۵/۳۴ و ۲۵/۴۲ را به ترتیب به دست آورده‌اند. آیا میانگین دو نوع ریسمان با یکدیگر مساوی است؟

۲۵ مطالعه‌ای در زمینهٔ مؤثر بودن آزمایشگاه همراه با درس فیزیک صورت پذیرفته است. به دانشجویان اجازه داده شده است که درس فیزیک سه ساعتی را بدون آزمایشگاه یا فیزیک چهار ساعتی را با آزمایشگاه در یک ترم بگیرند. در گروهی که آزمایشگاه داشتند ۱۱ دانشجو شرکت داشته که معدل نمرهٔ آنها در امتحان ۸۵ با انحراف معیار ۴/۷ شده است و در گروه بدون آزمایشگاه ۱۷ دانشجو شرکت داشته که معدل نمرهٔ آنها ۷۹ با انحراف معیار ۶/۱ شده است. فرض کنید که جمعیتها دارای توزیع نرمال با واریانسهای مساوی هستند. آیا در سطح معنی‌دار ۰/۰۱ می‌توان گفت که آزمایشگاه نمرهٔ دانشجویان را به طور متوسط ۸ نمره بالا خواهد برد؟

۲۶ دو نوع محافظ سپر اتومبیل مقایسه می‌شوند، ۶ تا از هر نوع را به ماشین محکمی سوار می‌کنند، سپس هر کدام از ماشینها را با سرعت ۲۰ کیلومتر در ساعت با دیوار سختی برخورد می‌دهند. هزینه‌های تعمیرات به صورت زیر به دست آمده‌اند

نوع ۱	۱۱۹	۱۰۲	۱۶۵	۱۲۳	۱۴۸	۱۰۷
نوع ۲	۱۲۹	۱۳۳	۱۵۱	۱۱۲	۱۱۵	۱۳۴

با فرض نرمال بودن جمعیتها و تساوی واریانسها، آیا در سطح معنی‌دار ۰/۰۱ می‌توان گفت که تفاوت میان میانگینهای دو جمعیت وجود دارد؟

۲۷ میزان تقاضا برای دو محصول A و B دارای توزیع نرمال با واریانسهای یکسان می‌باشد. میزان تقاضا برای ۱۰ روز این دو محصول را جمع آوری کرده و در جدول زیر یادداشت کرده‌ایم

محصول A	۳۰	۲۶	۳۰	۱۹	۲۵	۳۷	۲۷	۳۸	۲۶	۳۱
محصول B	۴۰	۳۴	۲۸	۲۹	۲۶	۳۶	۲۸	۳۷	۳۵	۴۲

آیا در سطح معنی‌دار ۰/۰۵ میزان تقاضا برای دو محصول متفاوت است؟

۲۸ اندازه‌گیری قطر عرضی قلبهای پسران و دختران بالغ نتایج زیر را داده است

گروه	اندازهٔ نمونه	میانگین	انحراف استاندارد
پسران	۱۲	۱۳/۲۱	۱/۰۵
دختران	۹	۱۱	۱/۰۱

با فرض اینکه جمعیتها نرمال با واریانس مساوی باشند، آیا در سطح معنی‌دار ۰/۰۲ داده‌ها گواه بر این هستند که قطر عرضی قلبهای پسران از دختران بیشتر است؟

۲۹ با انجام آزمایش برای تعیین میزان چسبندگی دو نوع روغن خودرو A و B نتایج زیر به دست آمده است

نوع A	۱۰/۲۸	۱۰/۲۷	۱۰/۳۰	۱۰/۳۲	۱۰/۲۷	۱۰/۲۷	۱۰/۲۸	۱۰/۲۹
نوع B	۱۰/۳۱	۱۰/۳۱	۱۰/۲۶	۱۰/۳۰	۱۰/۲۷	۱۰/۳۱	۱۰/۲۹	۱۰/۲۶

با فرض نرمال بودن دو جمعیت مطلوب است

الف– در سطح معنی‌دار ۰/۰۲ مساوی بودن واریانس دو جمعیت را آزمون کنید.

ب– با استفاده از نتیجهٔ الف در سطح معنی‌دار ۰/۰۵ مساوی بودن میانگینهای دو جمعیت را آزمون کنید.

۳۰ طول عمر ۹ باطری اتومبیل ساخت کارخانهٔ A به طور متوسط ۱۳۰۵ ساعت با انحراف استاندارد ۴۲۰ ساعت و طول عمر ۱۶ باطری اتومبیل ساخت کارخانهٔ B به طور متوسط ۱۲۱۰ ساعت با انحراف استاندارد ۳۸۰ ساعت بوده است. تحت فرض نرمال بودن جمعیتها،

الف– در سطح معنی‌دار ۰/۱۰ تساوی واریانسها را بیازمایید.

ب– در سطح معنی‌دار ۰/۱۰، آیا میانگین طول عمر برای باطریهای دو کارخانهٔ یکسان است.

۳۱ یک شرکت تاکسی رانی می‌خواهد از بین دو نوع لاستیک A و B یکی را برای تاکسیهای تندروی خود انتخاب نماید. به این منظور از هر نوع لاستیک نمونه‌ای ۱۲ تایی را انتخاب کرده و

آنها را آنقدر مورد استفاده قرار داده‌اند تا از بین رفته‌اند. نتایج زیر به دست آمده‌اند

نوع A	$\bar{x}_A=$ ۲۳۶۰۰	$s_A=$ ۳۲۰۰
نوع B	$\bar{x}_B=$ ۲۴۸۰۰	$s_B=$ ۳۷۰۰

با فرض نرمال بودن طول عمر لاستیکها، آیا در سطح معنی‌دار ۰/۱ می‌توان گفت که انحراف معیار طول عمر لاستیکهای نوع A از نوع B کمتر است؟

۳۲ در مطالعهٔ جریان ترافیک در دو تقاطع شلوغ بین ساعت ۴ و ۶ بعدازظهر دیده شده که در ۴۰ روز از روزهای معمولی هفته به طور متوسط ۲۴۷/۳ با انحراف استاندارد ۱۵/۲ ماشین از سمت جنوب به تقاطع اول می‌رسند و گردش به چپ دارند، در حالیکه در ۳۰ روز از روزهای معمولی هفته به طور متوسط ۲۵۴/۱ با انحراف استاندارد ۱۸/۷ ماشین از سمت جنوب به تقاطع دوم می‌رسند و گردش به چپ دارند. آیا در سطح معنی‌دار ۰/۰۵ می‌توان ادعا کرد که تغییر پذیری بزرگتری در تقاطع دوم در تعداد ماشینهایی که از سمت جنوب می‌آیند و گردش به چپ دارند، وجود دارد؟

۳۳ دو نمونهٔ تصادفی به اندازه‌های ۱۶ $=n_1$ و ۱۴ $=n_2$ از دو جمعیت نرمال مستقل دارای نتایج زیر است:

$\bar{x}_1=$ ۷۵ , $s_1=$ ۶/۱ , $\bar{x}_2=$ ۶۰ , $s_2=$ ۵/۳

در سطح معنی‌دار ۰/۰۱ آیا می‌توان گفت که واریانس جمعیت اول از دوم کمتر است؟

۳۴ نمره‌های یک درس آمار در ترم بخصوصی به صورت زیر گزارش شده است

نمره	A	B	C	D	F
تعداد	۱۴	۱۸	۱۲	۲۰	۱۶

در سطح معنی‌دار ۰/۰۱ آیا توزیع یکنواخت بر داده‌ها برازنده است؟

۳۵ مشاهدات زیر تعداد قطعات خراب در ۱۵۰ کارتن محتوی قطعات تولید شده توسط یک کارخانه را نشان می‌دهد

تعداد قطعات خراب	۰	۱	۲	۳	۴	۵	۶	۷
تعداد کارتنها	۲۳	۳۹	۴۳	۲۳	۱۰	۷	۴	۱

آیا تعداد قطعات خراب در کارتنهای این کارخانه از توزیع پواسون پیروی می‌کند؟

۳۶ استادی یک امتحان ۵۰ نمره‌ای به کلاس ریاضی عمومی داده است. جدول زیر نمره‌های ۵۰ دانشجو می‌باشد

۳۸	۳۴	۲۲	۴۱	۳۵	۴۵	۳۲	۳۷	۳۰	۲۶
۳۶	۲۸	۳۴	۱۶	۳۱	۳۳	۳۸	۳۱	۴۷	۳۶
۲۹	۲۷	۲۵	۴۳	۳۴	۳۶	۲۹	۳۹	۳۳	۳۱
۴۰	۳۵	۳۳	۳۱	۳۷	۴۴	۳۲	۴۱	۱۹	۳۴
۱۵	۴۹	۴۷	۳۸	۳۲	۲۶	۳۹	۳۰	۴۲	۳۵

الف – یک جدول فراوانی برای داده‌های فوق تشکیل دهید.

ب – آیا در سطح معنی‌دار ۰/۱ می‌توان ادعا کرد که نمرات دارای توزیع یکنواخت است؟

ج – آیا در سطح معنی‌دار ۰/۱ می‌توان ادعا کرد که نمرات دارای توزیع نرمال است؟

۳۷ یک سکه را آنقدر پرتاب می‌کنیم تا اینکه یک شیر بیاید. اگر X برابر تعداد پرتاب این سکه باشد، بعد از تکرار این آزمایش در ۲۵۶ بار، نتایج زیر حاصل می‌شود

x	۱	۲	۳	۴	۵	۶	۷	۸
تعداد پرتاب	۱۳۶	۶۰	۳۴	۱۲	۹	۱	۳	۱

آیا در سطح معنی‌دار ۰/۰۵ می‌توان ادعا کرد که توزیع هندسی با پارامتر $\frac{1}{2}$ بر داده‌ها برازنده است؟

۳۸ داده‌های زیر میزان محصول ذرت را در ۱۰۰ مزرعه نشان می‌دهد. اگر در این مزارع $\sum x_i = ۹۱۴۰۰$ و $s = ۳۳۱/۸$ باشد، آیا میزان محصول ذرت این مزارع از توزیع نرمال پیروی می‌کند؟

تعداد مزارع	محصول (برحسب کیلوگرم)
۳	$۹۹/۵ \leq x < ۲۹۹/۵$
۷	$۲۹۹/۵ \leq x < ۴۹۹/۵$
۱۵	$۴۹۹/۵ \leq x < ۶۹۹/۵$
۲۶	$۶۹۹/۵ \leq x < ۸۹۹/۵$
۲۲	$۸۹۹/۵ \leq x < ۱۰۹۹/۵$
۱۳	$۱۰۹۹/۵ \leq x < ۱۲۹۹/۵$
۹	$۱۲۹۹/۵ \leq x < ۱۴۹۹/۵$
۵	$۱۴۹۹/۵ \leq x < ۱۶۹۹/۵$

فصل نهم

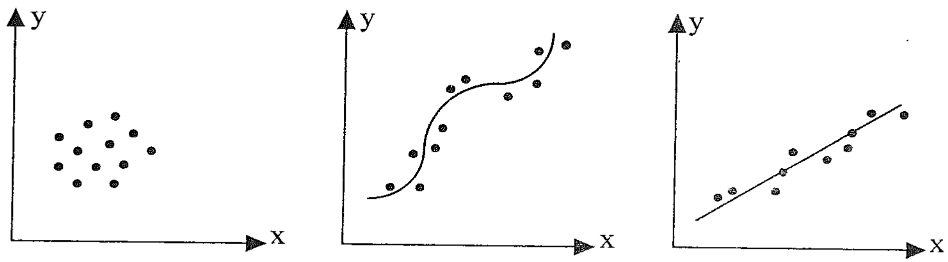
رگرسیون خطی و همبستگی

۱.۹ مقدمه

در اغلب بررسیهای آماری نیاز به پیش بینی مقدار یک متغیر وابسته از روی مقدار یک متغیر مستقل می‌باشد. برای مثال پیش بینی طول قد فرزند از روی طول قد پدرش، یا پیش بینی معدل کل یک دانشجو در یک نیمسال از روی نمرهٔ درس ریاضیات او و یا پیش بینی مقدار مصرف بنزین از روی مسافت طی شدهٔ اتومبیل از این نوع مسائل می‌باشند. چنین مسائلی را مسائل **برگشت یا رگرسیون** گویند. متغیر مستقل را با x و متغیر وابسته را با $Y \mid x$ و یا برای راحتی با Y نمایش می‌دهند. برای مثال در پیش بینی طول قد فرزند از روی طول قد پدرش، طول قد پدر را که یک مقدار ثابت و بخصوص است با x نمایش می‌دهیم و چون برای یک طول قد پدر x فرزندان او می‌توانند طول قدهای متفاوت داشته باشند بنابراین طول قد فرزند او را با متغیر تصادفی $Y \mid x$ یا Y نشان می‌دهیم. به همین ترتیب در پیش بینی مقدار مصرف بنزین، برای مسافت معین طی شدهٔ x، میزان مصرف بنزین را با Y نمایش می‌دهیم.

برای یافتن رابطه بین متغیر مستقل x و متغیر وابستهٔ Y ابتدا یک نمونهٔ تصادفی از جمعیت مورد نظر جمع آوری می‌کنیم. یعنی به ازاء مقادیر x_1، x_2، ... و x_n از متغیر مستقل x مقادیر مربوط به متغیر وابسته Y را اندازه‌گیری می‌کنیم. فرض کنید این مقادیر اندازه‌گیری شده y_1، y_2، ... و y_n باشند. بنابراین نمونهٔ تصادفی ما به صورت زوجهای (x_1,Y_1)، (x_2,Y_2)، ... و (x_n,Y_n) است که مقادیر مشاهدهٔ شده آن عبارت است از (x_1,y_1)، (x_2,y_2)، ... و (x_n,y_n). برای پی بردن به رابطهٔ

بین x و Y ابتدا این مشاهدات که به صورت نقاطی در صفحه هستند را در یک دستگاه مختصات رسم می‌کنیم که به آن نمودار پراکندگی گویند. در شکل ۱.۹ نمودار پراکندگی نقاط برای حالتهای مختلف رسم شده است. توجه کنید که برای یک مقدار x ممکن است چندین مقدار برای y وجود داشته باشد.



الف- x و y رابطه خطی دارند ب- x و y رابطه غیر خطی دارند ج- x و y رابطه‌ای ندارند

شکل ۱.۹ نمودارهای پراکندگی نقاط

حال براین نقاط می‌توان یک خط یا منحنی عبور داد و این خط یا منحنی رابطهٔ بین x و y را مشخص می‌کند. بنابراین در حالت کلی اگر بخواهیم مقدار متغیر Y را از روی مقدار متغیر x پیش بینی کنیم احتیاج به یک رابطه بین x و Y داریم که این رابطه یک معادلهٔ پیش بینی کننده است که به آن معادلهٔ رگرسیون Y روی x گویند.

۲.۹ رگرسیون سادهٔ خطی

هر گاه بین متغیر مستقل x و متغیر وابستهٔ $Y=Y|x$ یک رابطهٔ خطی برقرار باشد گوئیم یک مدل رگرسیون سادهٔ خطی بین x و Y برقرار است. برای تشکیل این رابطهٔ خطی، فرض کنید یک نمونهٔ تصادفی (x_1,Y_1)، (x_2,Y_2)، ... و (x_n,Y_n) با مقادیر مشاهده شدهٔ (x_1,y_1)، (x_2,y_2)، ... و (x_n,y_n) داشته باشیم. توجه کنید که $Y_i=Y|x_i\ ,\ i=1,2,\dots,n$. منظور از رگرسیون خطیِ این است که میانگین $Y|x$ به طور خطی با x در ارتباط باشد یعنی $\mu_{Y|x}=E(Y|x)=\alpha+\beta x$ که به آن خط رگرسیون گویند و در آن α و β پارامترهای نامعلوم هستند که بایستی برآورد شوند. به α و β ضرایب رگرسیون گویند اگر برآورد α و β را با $\hat{\alpha}$ و $\hat{\beta}$ نمایش دهیم در این صورت

مقدار برآورد متغیر وابستهٔ y را با $\hat{y}=\hat{\alpha}+\hat{\beta}x$ نمایش می‌دهند. در زیر روشی برای برآورد ضرایب رگرسیونی α و β از روی نمونه ارائه می‌دهیم و از روی آن بوسیلهٔ $\hat{y}=\hat{\alpha}+\hat{\beta}x$ برای مقدار بخصوص x مقدار متغیر وابستهٔ y را پیش بینی می‌کنیم.

در نمونهٔ تصادفی (x_1,Y_1)، (x_2,Y_2)، ... و (x_n,Y_n) چون همواره مقدار $Y_i=Y|x_i$ برابر $\mu_{Y|x_i}=E(Y|x_i)$ نیست، بنابراین اختلاف آنها برابر یک مقدار تصادفی E_i می‌باشد یعنی

$$\boxed{Y_i=\mu_{Y|x_i}+E_i=\alpha+\beta x_i+E_i\ ,\ i=1,2,\dots,n} \qquad (۱.۹)$$

این مدل را **مدل رگرسیون سادهٔ خطی** گویند و E_i را که یک متغیر تصادفی با میانگین صفر و واریانس σ^2 است را مقدار خطا گویند. اگر مقدار مشاهده شده E_i را با e_i نشان دهیم، در این صورت

$$y_i=\alpha+\beta x_i+e_i\ ,\ i=1,2,\dots,n$$

که e_i را مقدار باقیمانده گویند. حال برای یافتن برآوردهای $\hat{\alpha}$ و $\hat{\beta}$ بگونه‌ای عمل می‌کنیم که مجموع مربعات باقیمانده‌ها یعنی $\sum_{i=1}^{n} e_i^2$ می‌نیمم گردد.

روش حداقل مربعات مجموع مربعات باقیمانده‌ها را معمولاً مجموع مربعات خطاها حول خط رگرسیون گویند و با SSE[1] نمایش می‌دهند، یعنی

$$\boxed{SSE=\sum_{i=1}^{n} e_i^2=\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2} \qquad (۲.۹)$$

مقادیری از $\hat{\alpha}$ و $\hat{\beta}$ که SSE را می‌نیمم کند، برآوردهای حداقل مربعات گویند و روش به دست آوردن آنها را روش **حداقل مربعات**[2] می‌نامند که در زیر به ذکر آن می‌پردازیم.

ابتدا کمیتهای زیر را معرفی می‌کنیم

$$S_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}$$

$$S_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}y_i^2-\frac{\left(\sum_{i=1}^{n}y_i\right)^2}{n}$$

1- Sum of Square Errors 2- Least Squares Method

$$S_{xy}=\sum_{i=۱}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=۱}^{n}x_iy_i-\frac{(\sum_{i=۱}^{n}x_i)(\sum_{i=۱}^{n}y_i)}{n}$$

قضیهٔ ۱.۹ در یک مدل رگرسیون ساده خطی مقادیر $\hat{\alpha}$ و $\hat{\beta}$ که مجموع مربعات خطاها را می‌نیمم می‌کنند عبارت اند از

$$\boxed{\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x} \quad , \quad \hat{\beta}=\frac{S_{xy}}{S_{xx}}} \qquad (۳.۹)$$

اثبات با مشتق‌گیری از SSE نسبت به α و β و مساوی صفر قرار دادن آنها به معادلات زیر می‌رسیم

$$\begin{cases}\sum_{i=۱}^{n}y_i-n\alpha-\beta\sum_{i=۱}^{n}x_i=۰\\ \sum_{i=۱}^{n}x_iy_i-\alpha\sum_{i=۱}^{n}x_i-\beta\sum_{i=۱}^{n}x_i^۲=۰\end{cases}$$

که از حل این دستگاه جوابهای $\hat{\alpha}$ و $\hat{\beta}$ داده شده در (۳.۹) حاصل می‌شوند و می‌توان نشان داد که این مقادیر SSE را می‌نیمم می‌کنند.

بنابراین خط رگرسیون $\mu_{Y|x}=\alpha+\beta x$ بوسیلهٔ خط $\hat{y}=\hat{\alpha}+\hat{\beta}x$ برآورد می‌شود، که $\hat{\alpha}$ و $\hat{\beta}$ از رابطهٔ (۳.۹) به دست می‌آیند.

مثال ۱.۲.۹ برای داده‌های جدول زیر برآورد خط رگرسیون را بیابید. سپس نقاط را در صفحه مشخص نموده و خط برآورد شده را رسم کنید.

x_i	۱	۳	۴	۶	۸	۹	۱۱	۱۴
y_i	۱	۲	۴	۴	۵	۷	۸	۹

حل از جدول فوق مقادیر زیر حاصل می‌شوند

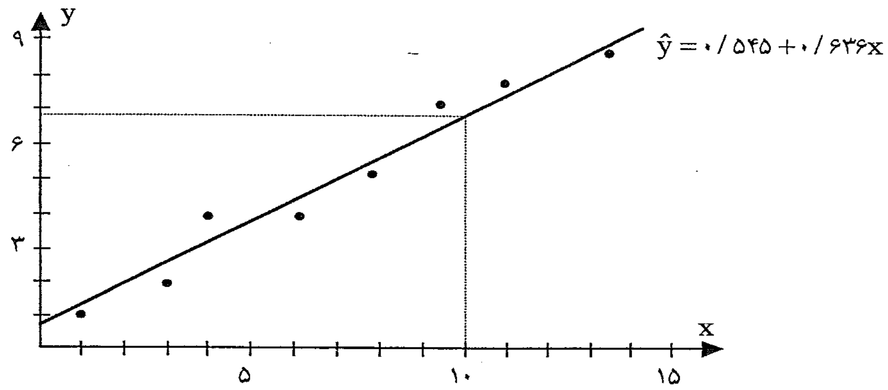
$$\sum x_i=۵۶ \,,\ \sum y_i=۴۰ \,,\ \sum x_i^۲=۵۲۴ \,,\ \sum y_i^۲=۲۵۶ \,,\ \sum x_iy_i=۳۶۴$$

بنابراین $S_{xy}=۳۶۴-\frac{(۵۶)(۴۰)}{۸}=۸۴ \quad , \quad S_{xx}=۵۲۴-\frac{(۵۶)^۲}{۸}=۱۳۲$

$$\hat{\beta}=\frac{S_{xy}}{S_{xx}}=\frac{۸۴}{۱۳۲}=۰/۶۳۶ \quad , \quad \hat{\alpha}=\bar{y}-\hat{\beta}\bar{x}=\frac{۴۰}{۸}-(۰/۶۳۶)\frac{۵۶}{۸}=۰/۵۴۵$$

و در نتیجه

$$\hat{y}=\hat{\alpha}+\hat{\beta}x=۰/۵۴۵+۰/۶۳۶x$$



شکل ۲.۹ نمودار پراکندگی نقاط و خط رگرسیون برآورد شده

مثال ۲.۲.۹ آزمایشی به منظور مطالعه اثر یک داروی معین در پایین آوردن ضربان قلب در افراد بالغ انجام شده است. مقدار داروی تجویز شده بر حسب میلی گرم و تفاوت ضربان قلب پس از استعمال دارو و قبل از آن برای یک نمونه ۱۳ تایی در جدول زیر آورده شده است

مقدار داروی تجویز شده	۰/۵	۰/۷۵	۱/۰	۱/۲۵	۱/۵	۱/۷۵	۲/۰
واکنش در ضربان قلب	۱۰	۸	۱۲	۱۲	۱۴	۱۲	۱۶
مقدار داروی تجویز شده	۲/۲۵	۲/۵	۲/۷۵	۳	۳/۲۵	۳/۵	
واکنش در ضربان قلب	۱۸	۱۷	۲۰	۱۸	۲۰	۲۱	

برآورد خط رگرسیون را به دست آورید. اگر مقدار داروی تجویز شده ۱/۶ باشد، واکنش در ضربان قلب در دقیقه را به چه میزان پیش بینی می‌کنید؟

حل از جدول فوق مقادیر زیر حاصل می‌شوند:

$$\sum x_i=۲۶, \sum x_i^۲=۶۳/۳۷۵, \sum y_i=۱۹۸, \sum y_i^۲=۳۲۲۶, \sum x_iy_i=۴۴۲/۵$$

بنابراین $S_{xy}=۴۴۲/۵-\frac{(۲۶)(۱۹۸)}{۱۳}=۴۶/۵ \ , \ S_{xx}=۶۳/۳۷۵-\frac{(۲۶)^۲}{۱۳}=۱۱/۳۷۵$

$$\hat{\beta}=\frac{۴۶/۵}{۱۱/۳۷۵}=۴/۰۸۸ \quad , \quad \hat{\alpha}=\frac{۱۹۸}{۱۳}-(۴/۰۸۸)\frac{۲۶}{۱۳}=۷/۰۵۵$$

در نتیجه

$$\hat{y}=\hat{\alpha}+\hat{\beta}x=۷/۰۵۵+۴/۰۸۸x$$

حال اگر مقدار داروی تجویز شده $x=۱/۶$ باشد آنگاه پیش بینی می‌کنیم که میزان واکنش ضربان قلب $\hat{y}=۷/۰۵۵+۴/۰۸۸(۱/۶)\approx ۱۳/۶$ در دقیقه باشد.

۳.۹ استنباط آماری روی ضرایب رگرسیونی

در بخش قبل بر اساس نمونهٔ تصادفی $(x_۱,Y_۱)$، $(x_۲,Y_۲)$، ... و (x_n,Y_n) در مدل رگرسیونی

$$Y_i=\alpha+\beta x_i+E_i \qquad i=۱,۲,...,n$$

ضرایب رگرسیونی α و β را به وسیلهٔ برآوردگرهای $\hat{\beta}=\frac{S_{xY}}{S_{xx}}$ و $\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{x}$ برآورد کردیم که در آن $\bar{Y}=\frac{۱}{n}\sum_{i=۱}^{n}Y_i$, $S_{xY}=\sum_{i=۱}^{n}(x_i-\bar{x})(Y_i-\bar{Y})$ دو متغیر تصادفی می‌باشند. حال اگر در مدل رگرسیونی فوق فرض کنیم $i=۱,۲,...,n$, $E_i\sim N(۰,\sigma^۲)$ و $E_۱$، $E_۲$، ... و E_n از یکدیگر مستقل باشند، در این صورت چون Y_i یک ترکیب خطی از متغیرهای تصادفی نرمال E_i می‌باشد، پس $Y_i\sim N(\alpha+\beta x_i,\sigma^۲)$, $i=۱,۲,...,n$. بر اساس توزیع متغیرهای تصادفی Y_i می‌توان توزیع $\hat{\alpha}$، $\hat{\beta}$ و SSE را به دست آورد. توزیع این برآوردگرها را در قضیهٔ زیر بدون اثبات می‌آوریم.

قضیهٔ ۲.۹ در مدل رگرسیونی سادهٔ خطی با فرض $E_i\sim N(۰,\sigma^۲)$, $i=۱,۲,...,n$ داریم که

الف- $\hat{\alpha}\sim N\left(\alpha,\frac{\sigma^۲\sum x_i^۲}{nS_{xx}}\right)$

ب- $\hat{\beta}\sim N\left(\beta,\frac{\sigma^۲}{S_{xx}}\right)$

ج- اگر $S^۲=\frac{SSE}{n-۲}$ آنگاه $\frac{SSE}{\sigma^۲}=\frac{(n-۲)S^۲}{\sigma^۲}\sim\chi^۲_{(n-۲)}$

د- $\hat{\alpha}$ و $S^۲$ و همچنین $\hat{\beta}$ و $S^۲$ از یکدیگر مستقل هستند.

با استفاده از قضیهٔ ۲.۹ می‌توان نتایج زیر را به دست آورد

$$\boxed{\frac{\hat{\beta}-\beta}{\frac{S}{\sqrt{S_{xx}}}}\sim t_{(n-۲)}} \qquad (۴.۹)$$

$$\boxed{\frac{\hat{\alpha}-\alpha}{S\sqrt{\frac{\sum x_i^۲}{nS_{xx}}}}\sim t_{(n-۲)}} \qquad (۵.۹)$$

که در آن

$$\boxed{S^۲=\frac{SSE}{n-۲}=\frac{۱}{n-۲}\sum_{i=۱}^{n}(Y_i-\hat{\alpha}-\hat{\beta}x_i)^۲=\frac{۱}{n-۲}\left[S_{YY}-\hat{\beta}S_{xY}\right]} \qquad (۶.۹)$$

با استفاده از تابع محورهای (۴.۹) و (۵.۹) می‌توان فواصل اطمینان $۱۰۰(۱-\alpha)\%$ برای α و β به دست آورد که این فواصل اطمینان عبارتند از

$$\boxed{\begin{aligned}&\alpha\in\left(\hat{\alpha}-t_{۱-\frac{\alpha}{۲}}(n-۲)S\sqrt{\frac{\sum x_i^۲}{nS_{xx}}}\ ,\ \hat{\alpha}+t_{۱-\frac{\alpha}{۲}}(n-۲)S\sqrt{\frac{\sum x_i^۲}{nS_{xx}}}\right)\\ &\beta\in\left(\hat{\beta}-t_{۱-\frac{\alpha}{۲}}(n-۲)\frac{S}{\sqrt{S_{xx}}}\ ,\ \hat{\beta}+t_{۱-\frac{\alpha}{۲}}(n-۲)\frac{S}{\sqrt{S_{xx}}}\right)\end{aligned}} \qquad (۷.۹)$$

با استفاده از روابط (۴.۹) و (۵.۹) می‌توان آزمونهای آماری را روی α و β انجام داد که این آزمونها در جدول ۱.۹ آورده شده‌اند.

$H_۰$	آمارهٔ آزمون	$H_۱$	ناحیه بحرانی
$\alpha=\alpha_۰$	$T=\frac{\hat{\alpha}-\alpha_۰}{S\sqrt{\frac{\sum x_i^۲}{nS_{xx}}}}$	$\alpha>\alpha_۰$	$T>t_{۱-\alpha}(n-۲)$
		$\alpha<\alpha_۰$	$T<-t_{۱-\alpha}(n-۲)$
		$\alpha\neq\alpha_۰$	$\lvert T\rvert>t_{۱-\frac{\alpha}{۲}}(n-۲)$
$\beta=\beta_۰$	$T=\frac{\hat{\beta}-\beta_۰}{\frac{S}{\sqrt{S_{xx}}}}$	$\beta>\beta_۰$	$T>t_{۱-\alpha}(n-۲)$
		$\beta<\beta_۰$	$T<-t_{۱-\alpha}(n-۲)$
		$\beta\neq\beta_۰$	$\lvert T\rvert>t_{۱-\frac{\alpha}{۲}}(n-۲)$

جدول ۱.۹ آزمونهای آماری روی ضرایب رگرسیونی α و β

مثال ۱.۳.۹ مواد اولیه‌ای که برای ساختن الیاف مصنوعی به کار می‌رود در انبار مرطوبی نگهداری می‌شود. نتایج حاصل از اندازه‌گیریهای رطوبت نسبی در انبار و میزان رطوبت در یک نمونه مواد اولیه (هر دو بر حسب درصد) در ۱۲ روز در جدول زیر ثبت شده است

رطوبت انبار x	۴۲	۳۵	۵۰	۴۳	۴۸	۶۲
رطوبت مواد اولیه y	۱۲	۸	۱۴	۹	۱۱	۱۶
رطوبت انبار x	۳۱	۳۶	۴۴	۳۹	۵۵	۴۸
رطوبت مواد اولیه y	۷	۹	۱۲	۱۰	۱۳	۱۱

الف - برآورد خط رگرسیون را به دست آورید.

ب - یک فاصلهٔ اطمینان ۹۵ درصدی برای β بسازید.

ج - آیا در سطح معنی دار ۰/۰۱ می‌توان ادعا کرد که $\alpha=-۱$ است؟

حل الف - از جدول فوق مقادیر زیر حاصل می‌شوند:

$$\sum x_i=۵۳۳\ ,\ \sum x_i^۲=۲۴۵۲۹\ ,\ \sum y_i=۱۳۲\ ,\ \sum y_i^۲=۱۵۲۶\ ,\ \sum x_iy_i=۶۰۹۳$$

بنابراین

$$S_{xy}=۶۰۹۳-\frac{(۵۳۳)(۱۳۲)}{۱۲}=۲۳۰\quad ,\quad S_{xx}=۲۴۵۲۹-\frac{(۵۳۳)^۲}{۱۲}=۸۵۴/۹۱۷$$

$$S_{yy}=۱۵۲۶-\frac{(۱۳۲)^۲}{۱۲}=۷۴\quad ,\quad \hat{\beta}=\frac{۲۳۰}{۸۵۴/۹۱۷}=۰/۲۶۹$$

$$\hat{\alpha}=\frac{۱۳۲}{۱۲}-(۰/۲۶۹)\frac{۵۳۳}{۱۲}=-۰/۹۴۸\quad ,\quad s^۲=\frac{۱}{۱۰}\left[۷۴-(۰/۲۶۹)(۲۳۰)\right]=۱/۲۱۳$$

در نتیجه

$$\hat{y}=\hat{\alpha}+\hat{\beta}x=-۰/۹۴۸+۰/۲۶۹x$$

ب - با استفاده از فواصل اطمینان (۷.۹) داریم که

$$t_{۱-\frac{\alpha}{۲}}(n-۲)=t_{۰/۹۷۵}(۱۰)=۲/۲۳$$

$$\beta\in\left(۰/۲۶۹-۲/۲۳\sqrt{\frac{۱/۲۱۳}{۸۵۴/۹۱۷}},\ ۰/۲۶۹+۲/۲۳\sqrt{\frac{۱/۲۱۳}{۸۵۴/۹۱۷}}\right)=(۰/۱۸۵,\ ۰/۳۵۳)$$

بنابراین ۹۵ درصد اطمینان داریم که β در فاصلهٔ فوق قرار دارد.

ج - در این مثال با آزمون $\begin{cases}H_۰ : \alpha=-۱\\ H_۱ : \alpha\neq-۱\end{cases}$ مواجه هستیم و با استفاده از جدول ۱.۹ فرض $H_۰$ رد می‌شود اگر و فقط اگر $|T|>t_{۱-\frac{\alpha}{۲}}(n-۲)$ که در آن

$$t_{۱-\frac{\alpha}{۲}}(n-۲)=t_{۰/۹۹۵}(۱۰)=۳/۱۷$$

$$T=\frac{\hat{\alpha}-\alpha_۰}{s\sqrt{\frac{\sum x_i^۲}{nS_{xx}}}}=\frac{-۰/۹۴۸+۱}{\sqrt{\frac{(۱/۲۱۳)(۲۴۵۲۹)}{(۱۲)(۸۵۴/۹۱۷)}}}=۰/۰۳۰۵$$

چون $|T|=۰/۰۳۰۵\not>t_{۱-\frac{\alpha}{۲}}(n-۲)=۳/۱۷$ پس فرض $H_۰$ رد نمی‌شود یعنی $\alpha=-۱$ می‌باشد.

مثال ۲.۳.۹ نمره‌های امتحان میان ترم و پایان ترم یک کلاس ۹ نفره از دانشجویان به صورت زیر است

میان ترم	۶	۵	۷	۷	۴	۶	۴	۵	۳
پایان ترم	۱۰	۸	۱۱	۱۲	۱۱	۹	۱۰	۹	۶

الف - برآورد خط رگرسیونی را برای پیش بینی نمرهٔ پایان ترم از روی نمرهٔ میان ترم به دست آورید.

ب - یک فاصلهٔ اطمینان ۹۰ درصدی برای α پیدا کنید.

ج - آیا در سطح معنی دار ۰/۰۵ می‌توان ادعا کرد که $\beta<۲$ است؟

حل الف - اگر x نمرهٔ میان ترم و y نمرهٔ پایان ترم باشند آنگاه از جدول فوق مقادیر زیر حاصل می‌شوند

$$\sum x_i=۴۷\ ,\ \sum x_i^۲=۲۶۱\ ,\ \sum y_i=۸۶\ ,\ \sum y_i^۲=۸۴۸\ ,\ \sum x_iy_i=۴۶۲$$

بنابراین

$$S_{xy}=۴۶۲-\frac{(۴۷)(۸۶)}{۹}=۱۲/۸۹\quad ,\quad S_{xx}=۲۶۱-\frac{(۴۷)^۲}{۹}=۱۵/۵۶$$

$$S_{yy}=۸۴۸-\frac{(۸۶)^۲}{۹}=۲۶/۲۲\quad ,\quad \hat{\beta}=\frac{۱۲/۸۹}{۱۵/۵۶}=۰/۸۲۸$$

$$\hat{\alpha}=\frac{۸۶}{۹}-(۰/۸۲۸)\frac{۴۷}{۹}=۵/۲۳\quad ,\quad s^۲=\frac{۱}{۷}\left[۲۶/۲۲-(۰/۸۲۸)(۱۲/۸۹)\right]=۲/۲۲$$

در نتیجه

$$\hat{y}=۵/۲۳+۰/۸۲۸x$$

ب- چون $t_{۱-\frac{\alpha}{۲}}(n-۲)=t_{۰/۹۵}(۷)=۱/۹$ بنابراین از (۷.۹) داریم که

$$\alpha\in\left(۵/۲۳-۱/۹\sqrt{\frac{(۲/۲۲)(۲۶۱)}{(۹)(۱۵/۵۶)}},\ ۵/۲۳+۱/۹\sqrt{\frac{(۲/۲۲)(۲۶۱)}{(۹)(۱۵/۵۶)}}\right)=(۱/۳۶۵,۹/۰۹۵)$$

بنابراین ۹۰ درصد اطمینان داریم که α در فاصلهٔ فوق قرار دارد.

ج- در این مثال با آزمون $\begin{cases}H_۰:\beta=۲\\ H_۱:\beta<۲\end{cases}$ مواجه هستیم و با استفاده از جدول ۱.۹ فرض $H_۰$ رد می‌شود اگر و فقط اگر $T<-t_{۱-\alpha}(n-۲)$ که در آن

$$t_{۱-\alpha}(n-۲)=t_{۰/۹۵}(۷)=۱/۹$$

$$T=\frac{\hat{\beta}-\beta_۰}{\frac{S}{\sqrt{S_{xx}}}}=\frac{۰/۸۲۸-۲}{\sqrt{\frac{۲/۲۲}{۱۵/۵۶}}}=-۳/۱۰۳$$

چون $-۳/۱۰۳=T<-t_{۱-\alpha}(n-۲)=-۱/۹$ پس فرض $H_۰$ رد می‌شود یعنی $\beta<۲$ است.

۴.۹ ضریب همبستگی خطی

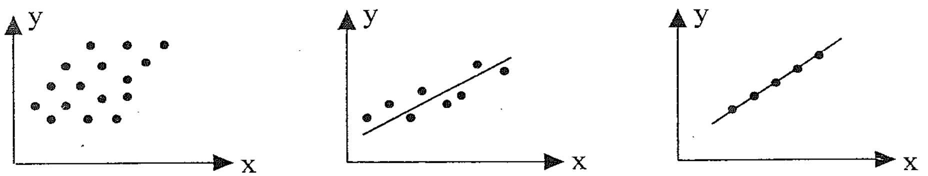
تاکنون فرض کردیم که متغیر مستقل x یک متغیر کنترل شده است و یک متغیر تصادفی نیست. حال فرض کنید که هم متغیر X و هم Y هر دو متغیر تصادفی باشند.

روشهای رگرسیونی موقعی مناسب هستند که متغیر تصادفی Y به متغیر تصادفی X که اغلب به وسیلهٔ پژوهشگر کنترل می‌شود بستگی داشته باشد. برای سنجش میزان وابستگی دو متغیر تصادفی X و Y از معیاری بنام ضریب همبستگی خطی استفاده می‌شود که در بخش ۴.۴ آن را در جمعیت به صورت زیر تعریف کردیم

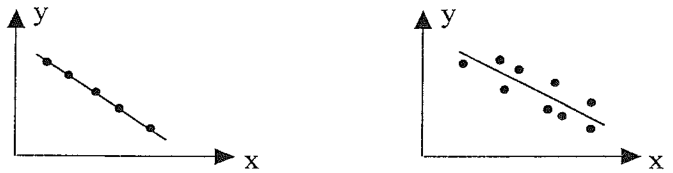
$$\rho=\rho(X,Y)=\frac{COV(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

و در بخش ۴.۴ مشاهده کردیم که ضریب همبستگی خطی به مبداء و واحد اندازه‌گیری داده‌ها بستگی ندارد و همواره $-۱\le\rho\le۱$ می‌باشد. در شکل ۳.۹ حالتهای مختلف از همبستگی خطی X و

Y نشان داده شده است.



الف- $\rho=۱$ همبستگی کامل مثبت　ب- $۰<\rho<۱$ همبستگی مثبت　ج- $\rho=۰$ عدم همبستگی



د- $-۱<\rho<۰$ همبستگی منفی　هـ- $\rho=-۱$ همبستگی کامل منفی

شکل ۳.۹ حالتهای مختلف همبستگی خطی بین متغیرهای X و Y

برای برآورد ضریب همبستگی یک نمونهٔ تصادفی $(X_۱,Y_۱)$، $(X_۲,Y_۲)$، ... و (X_n,Y_n) از (X,Y) را انتخاب می‌کنیم و از روی این نمونهٔ تصادفی کواریانس X و Y، واریانس X و واریانس Y را به صورت زیر برآورد می‌کنیم

$$\hat{\sigma}_X^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\bar{X})^۲=\frac{۱}{n-۱}S_{XX}$$

$$\hat{\sigma}_Y^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(Y_i-\bar{Y})^۲=\frac{۱}{n-۱}S_{YY}$$

$$\hat{\sigma}_{XY}=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\bar{X})(Y_i-\bar{Y})=\frac{۱}{n-۱}S_{XY}$$

حال با قرار دادن این برآوردگرها به جای پارامترهای $\sigma_X^۲,\sigma_Y^۲,\sigma_{XY}$ در فرمول ضریب همبستگی خطی، برآوردگر ضریب همبستگی خطی به صورت زیر به دست می‌آید:

$$\boxed{\hat{\rho}=R=\frac{S_{XY}}{\sqrt{S_{XX}S_{YY}}}} \qquad (۸.۹)$$

اگر مقدار مشاهدهٔ شدهٔ این نمونهٔ تصادفی $(x_۱,y_۱)$، $(x_۲,y_۲)$، ... و (x_n,y_n) باشد آنگاه برآورد

ضریب همبستگی خطی یعنی مقدار مشاهده شده R عبارت از $r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$ خواهد بود که به آن ضریب **همبستگی نمونه** گویند. می‌توان نشان داد که R با تغییر مبدأ و واحد اندازه‌گیری تغییر نمی‌کند و همواره $-۱\le R\le ۱$ است. تعبیر مقادیر R همانند تعبیر مقادیر ρ در شکل ۳.۹ می‌باشد. به مقدار $R^۲$ **ضریب تعیین** گویند و $۱۰۰R^۲$٪ از تغییرات مقادیر Y جهت رابطهٔ خطی آن با متغیر X به حساب می‌آید.

مثال ۱.۴.۹ داده‌های زیر مربوط به مقاومت (بر حسب اهم) و زمان شکست (بر حسب دقیقه) ترانزیستورها با بار اضافی می‌باشد

مقاومت (x)	۴۳	۲۹	۴۴	۳۳	۳۳	۴۷
زمان شکست (y)	۳۲	۲۰	۴۵	۳۵	۲۲	۴۶
مقاومت (x)	۳۴	۳۱	۴۸	۳۴	۴۶	۳۷
زمان شکست (y)	۲۸	۲۶	۳۷	۳۳	۴۷	۳۰

ضریب همبستگی نمونه را به دست آورید و آن را تعبیر کنید.

حل از جدول فوق مقادیر زیر حاصل می‌شوند

$$\sum x_i=۴۵۹\ ,\ \sum x_i^۲=۱۸۰۷۵\ ,\ \sum y_i=۴۰۱\ ,\ \sum y_i^۲=۱۴۳۰۱\ ,\ \sum x_iy_i=۱۵۹۰۷$$

بنابراین

$$S_{xy}=۱۵۹۰۷-\frac{(۴۵۹)(۴۰۱)}{۱۲}=۵۶۸/۷۵\quad ,\quad S_{xx}=۱۸۰۷۵-\frac{(۴۵۹)^۲}{۱۲}=۵۱۸/۲۵$$

$$S_{yy}=۱۴۳۰۱-\frac{(۴۰۱)^۲}{۱۲}=۹۰۰/۹۱۷$$

در نتیجه

$$r=\frac{۵۶۸/۷۵}{\sqrt{(۵۱۸/۲۵)(۹۰۰/۹۱۷)}}=۰/۸۳۲$$

چون مقدار r به یک نزدیک است پس یک رابطهٔ خطی نسبتاً خوبی در جهت مثبت بین X و Y برقرار است. همچنین چون $r^۲=۰/۶۹۳$ پس ۶۹/۳٪ از تغییرات Y جهت رابطهٔ خطی آن با X به حساب می‌آید.

استنباط آماری روی ρ

برای استنباط آماری روی ρ نیاز به داشتن توزیع احتمال R ضریب همبستگی نمونهٔ

تصادفی داریم. فیشر[1] آماردان انگلیسی ثابت کرده است که آماره $W=\frac{۱}{۲}Ln\frac{۱+R}{۱-R}$ برای اندازهٔ نمونهٔ بزرگ تقریباً دارای توزیع نرمال با میانگین تقریبی $\mu_W\approx\frac{۱}{۲}Ln\frac{۱+\rho}{۱-\rho}$ و واریانس تقریبی $\sigma_W^۲\approx\frac{۱}{n-۳}$ می‌باشد. یعنی به طور تقریبی داریم که:

$$\boxed{Z=\frac{W-\mu_W}{\sigma_W}\sim N(۰,۱)}\qquad (۹.۹)$$

با استفاده از تابع محور (۹.۹) می‌توان برای ρ فاصلهٔ اطمینان پیدا کرد و با استفاده از توزیع (۹.۹) می‌توان آزمونهای آماری را روی ρ انجام داد که این آزمونها در جدول ۲.۹ آورده شده‌اند.

$H_۰$	آمارهٔ آزمون	$H_۱$	ناحیهٔ بحرانی آزمون
$\rho=\rho_۰$	$Z=\frac{W-\mu_{W_۰}}{\sigma_W}$	$\rho>\rho_۰$	$Z>z_{۱-\alpha}$
		$\rho<\rho_۰$	$Z<-z_{۱-\alpha}$
		$\rho\neq\rho_۰$	$\lvert Z\rvert>z_{۱-\frac{\alpha}{۲}}$

جدول ۲.۹ آزمونهای آماری روی ρ که $\mu_{W_۰}=\frac{۱}{۲}Ln\frac{۱+\rho_۰}{۱-\rho_۰}$ می‌باشد.

مثال ۲.۴.۹ در مثال ۱.۴.۹ آیا در سطح معنی‌دار ۰/۰۵ می‌توان ادعا کرد که $\rho>۰/۵$ است.

حل در این مثال با آزمون $\begin{cases}H_۰\ :\ \rho=۰/۵\\ H_۱\ :\ \rho>۰/۵\end{cases}$ مواجه هستیم و با استفاده از جدول ۲.۹ فرض $H_۰$ رد می‌شود اگر و فقط اگر $Z>z_{۱-\alpha}$ که در آن

$$r=۰/۸۳۲\quad ,\quad w=\frac{۱}{۲}Ln\frac{۱+r}{۱-r}=\frac{۱}{۲}Ln\frac{۱/۸۳۲}{۰/۱۶۸}=۱/۱۹۵$$

$$\mu_{w_۰}=\frac{۱}{۲}Ln\frac{۱+\rho_۰}{۱-\rho_۰}=\frac{۱}{۲}Ln\frac{۱/۵}{۰/۵}=۰/۵۴۹$$

$$\sigma_w=\sqrt{\frac{۱}{n-۳}}=\sqrt{\frac{۱}{۹}}=۰/۳۳۳\quad ,\quad Z=\frac{w-\mu_{w_۰}}{\sigma_w}=\frac{۱/۱۹۵-۰/۵۴۹}{۰/۳۳۳}=۱/۹۴$$

$$z_{۱-\alpha}=z_{۰/۹۵}=۱/۶۴۵$$

چون $۱/۹۴=Z>z_{۱-\alpha}=۱/۶۴۵$ پس فرض $H_۰$ رد می‌شود یعنی $\rho>۰/۵$ است.

1- Fisher

۵.۹ تمرینات

۱ برای تعیین رابطه بین هزینه حمل یک نوع کالا و فاصلهٔ فروشگاه از محل توزیع کالا، یک نمونهٔ تصادفی شامل ۸ فروشگاه که این کالا را عرضه می‌کنند انتخاب و فاصلهٔ فروشگاه تا محل توزیع کالا و هزینهٔ حمل ۱۰۰ واحد از این کالا در جدول زیر ثبت شده است

x فاصله به کیلومتر	۶	۱۳	۲۷	۱۵	۹	۱۱	۲۱	۱۴
y هزینهٔ حمل بر حسب صد تومان	۲۹	۹۳	۱۵۹	۱۱۵	۶۶	۹۰	۱۳۹	۹۸

برآورد خط رگرسیون هزینهٔ حمل کالا بر حسب فاصله را به دست آورید. اگر فاصلهٔ یک فروشگاه تا محل توزیع ۱۰ کیلومتر باشد، هزینهٔ حمل ۱۰۰ واحد کالا تا این فروشگاه را به چه میزان پیش بینی می‌کنید؟

۲ می‌خواهیم به داده‌های حاصل از قدرت کشش y، ده قطعه پلاستیک که هر یک x دقیقه پخته شده‌اند، یک خط راست برازش کنیم. داده‌ها در جدول زیر ثبت شده‌اند. برآورد خط رگرسیون مورد نظر را به دست آورید.

x	۲۳	۳۵	۴۵	۶۵	۷۵	۹۵	۱۰۵	۱۲۵	۱۵۵	۱۸۵
y	۲	۹/۸	۹/۲	۲۶/۲	۱۷/۱	۲۴/۸	۴۳	۵۵/۳	۳۸/۴	۶۳/۳

۳ از یک نمونهٔ تصادفی از ۱۰ مرد ۳۰ ساله اطلاعات زیر در مورد حقوق سالیانه کنونی (بر حسب صد هزار تومان) و تعداد سالهای تحصیلی رسمی که داشته‌اند به دست آمده است.

x سالهای تحصیل	۱۰	۱۲	۱۳	۱۴	۱۶	۱۶	۱۶	۱۶	۱۸	۲۰
y حقوق سالیانه	۷/۰	۶/۲	۸/۱	۷/۵	۶/۵	۱۰/۵	۸/۰	۱۳/۲	۱۲/۸	۱۶/۵

برآورد خط رگرسیون حقوق سالیانه بر حسب تعداد سالهای تحصیل را به دست آورید. اگر مردی ۱۵ سال تحصیل کرده باشد، حقوق سالیانهٔ او را به چه میزان پیش بینی می‌کنید؟

۴ یک نوع مادهٔ شیمیایی داریم که در x درجه حرارت y گرم آن تجزیه می‌شود. در پنج آزمایش داده‌های زیر به دست آمده‌اند:

x	−۲	−۱	۰	۱	۲
y	۵	۴	۳	۲	۱

الف- برآورد خط رگرسیون مقدار مادهٔ تجزیه شده بر حسب درجه حرارت را به دست

آورید.

ب- نقاط داده شده و خط برآورد شده را در یک دستگاه مختصات رسم کنید.

ج- اگر درجه حرارت ۳ باشد، مقدار ماده‌ای را که تجزیه می‌شود پیش بینی کنید.

۵ در جدول زیر نیروی کشش به کار رفته برای یک نمونه فولاد، بر حسب هزار پوند و طول حاصل از کشش بر حسب یک هزارم اینچ می‌باشد

x نیروی کشش	۱	۲	۳	۴	۵	۶
y طول	۱۴	۳۳	۴۰	۶۳	۷۶	۸۵

الف- با رسم نمودار پراکندگی داده‌ها، تأیید کنید که فرض خطی بودن رگرسیون y روی x معقول است.

ب- برآورد خط رگرسیون را به دست آورده و با استفاده از آن وقتی نیروی کشش ۳/۵ هزار پوند باشد، طول حاصل را پیش بینی کنید.

ج- برای ضرایب رگرسیونی α و β فواصل اطمینان ۹۵ درصدی را تشکیل دهید.

۶ یک مطالعه توسط یک خرده فروش در ارتباط با رابطهٔ بین مخارج تبلیغ و میزان فروش (هر دو بر حسب هزار تومان) به طور هفتگی صورت پذیرفته است و داده‌های زیر به دست آمده است

مخارج تبلیغ	۴۰	۲۰	۲۵	۲۰	۳۰	۵۰	۴۰	۲۰	۵۰	۴۰	۲۵	۵۰
فروش	۳۸۵	۴۰۰	۳۹۵	۳۶۵	۴۷۵	۴۴۰	۴۹۰	۴۲۰	۵۶۰	۵۲۵	۴۸۰	۵۱۰

الف- معادله خط رگرسیون را جهت پیش بینی فروش هفتگی از روی هزینهٔ تبلیغات پیدا کنید.

ب- فاصلهٔ اطمینان ۹۰ درصدی را برای β به دست آورید.

۷ در تمرین ۳، آیا می‌توان در سطح معنی‌دار ۰/۰۱ ادعا کرد که $\beta>0$ است؟

۸ داده‌های زیر مربوط به تعداد کارها بر حسب روز و زمان لازم برای پردازش مرکزی (CPU) می‌باشد.

x تعداد کارها	۱	۲	۳	۴	۵
y زمان (CPU)	۲	۵	۴	۹	۱۰

الف- معادلهٔ خط رگرسیون را جهت پیش بینی زمان (CPU) بر حسب تعداد کارها پیدا

کنید.

ب- یک فاصلهٔ اطمینان ۹۵ درصدی برای ضریب رگرسیونی α به دست آورید.

ج- فرض صفر ۲=β در مقابل ۲<β را در سطح معنی‌دار ۰/۰۵ آزمون کنید.

۹ اثر دما بر رنگ محصولی بررسی شده است. داده‌ها در جدول زیر آورده شده‌اند

دما x	۱۰۰	۱۱۰	۱۲۰	۱۴۰	۱۵۰	۱۷۰	۱۸۰	۲۰۰	۲۳۰	۱۶۰
رنگ y	۳/۵	۷/۴	۷/۱	۱۵/۶	۱۱/۱	۱۴/۹	۲۳/۵	۲۷/۱	۲۲/۱	۳۲/۹

الف- معادلهٔ خط رگرسیون را جهت پیش بینی رنگ بر حسب دما پیدا کنید.

ب- آیا در سطح معنی‌دار ۰/۰۵ می‌توان ادعا کرد که ۰=α است؟

۱۰ داده‌های جدول زیر را در نظر بگیرید

x	۱	۱/۱	۱/۲	۱/۳	۱/۴	۱/۵	۱/۶	۱/۷	۱/۸	۱/۹	۲
y	۸/۱	۷/۸	۸/۵	۹/۸	۹/۵	۸/۹	۸/۶	۱۰/۲	۹/۳	۹/۲	۱۰/۵

الف- معادلهٔ خط رگرسیون را پیدا کنید.

ب- برای ضریب رگرسیونی α و β فواصل اطمینان ۹۰ درصدی را تشکیل دهید.

ج- آیا در سطح معنی‌دار ۰/۰۱ می‌توان ادعا کرد که ۰<α است؟

۱۱ جدول زیر نمره‌های ۷ دانشجو در امتحان ریاضی و زبان می‌باشد.

نمرهٔ ریاضی	۱۲	۱۳	۱۰	۱۳	۱۵	۱۱	۱۲
نمرهٔ زبان	۱۲	۱۴	۱۳	۱۵	۱۱	۱۴	۱۱

ضریب همبستگی نمونه را محاسبه و آن را تعبیر کنید.

۱۲ در تمرین ۱ ضریب همبستگی نمونه را به دست آورید و آن را تعبیر کنید.

۱۳ برای تعیین رابطهٔ خطی بین دو متغیر x و y، نمونهٔ تصادفی به اندازهٔ n=۲۰ از یک جمعیت استخراج و مقادیر x و y را ثبت کرده‌ایم. نتایج به شرح زیر می‌باشند

$$\sum x_i=۳۵ \quad , \quad \sum x_i^۲=۶۸۰ \quad , \quad \sum y_i=۴۸ \quad , \quad \sum y_i^۲=۱۳۴۸ \quad , \quad \sum x_i y_i=۹۶۰$$

ضریب همبستگی نمونه را به دست آورید و آن را تعبیر کنید.

۱۴ در جدول زیر نمرات امتحان پایان ترم درسهای جبر و فیزیک ۱۰ دانشجو که به طور تصادفی انتخاب شده‌اند، درج گردیده است

جبر	۷۵	۸۰	۹۳	۶۵	۸۷	۷۱	۹۸	۶۸	۸۴	۷۷
فیزیک	۸۲	۷۸	۸۶	۷۲	۹۱	۸۰	۹۵	۷۲	۸۹	۷۴

ضریب همبستگی نمونه را به دست آورید و آن را تعبیر کنید.

۱۵ نتایج زیر از اندازه‌گیری حجم عبور و مرور و غلظت مونو اکسید کربن در ۱۵ محل نمونه برداری حاصل شده است

حجم عبور و مرور x	۱۰۰	۱۱۰	۱۲۵	۱۵۰	۱۷۵	۱۹۰
غلظت مونواکسید کربن y	۸/۸	۹/۰	۹/۵	۱۰/۰	۱۰/۵	۱۰/۵

x	۲۰۰	۲۲۵	۲۵۰	۲۷۵	۳۰۰	۳۲۵	۳۵۰	۳۷۵	۴۰۰
y	۱۰/۵	۱۰/۶	۱۱/۰	۱۲/۱	۱۲/۱	۱۲/۵	۱۳/۰	۱۳/۲	۱۴/۵

الف- ضریب همبستگی نمونه را محاسبه و آن را تعبیر کنید.

ب- فرضیهٔ ناهمبسته بودن داده‌ها، یعنی ۰=ρ را در سطح معنی‌دار ۰/۰۵ آزمون کنید.

۱۶ با استفاده از تابع محور (۹.۹) یک فاصلهٔ اطمینان ۱۰۰(۱-α)٪ برای ضریب همبستگی خطی ρ پیدا کنید.

(راهنمایی: توجه کنید که $f(\rho)=Ln\,\frac{۱+\rho}{۱-\rho}$ یک تابع اکیداً صعودی از ρ است.)

۱۷ ضریب همبستگی میان قد و وزن ۱۰۰ نوزاد برابر ۰/۳ است. آیا در سطح معنی‌دار ۰/۱ می‌توان ادعا کرد که ۰/۲۵<ρ است.

۱۸ با استفاده از نتیجهٔ تمرین ۱۶ یک فاصلهٔ اطمینان ۹۰ درصدی برای ضریب همبستگی خطی در تمرین ۱۱ به دست آورید.

۱۹ داده‌های زیر اندازه‌های میزان کربن و شاخص نفوذپذیری ۲۲ ترکیب خاکستر می‌باشند.

میزان کربن (درصد)	۴/۴	۵/۵	۴/۲	۳/۰	۴/۵	۴/۹	۴/۶	۵/۰	۴/۷	۵/۱	۴/۴
شاخص نفوذپذیری	۱۲	۱۴	۱۸	۳۵	۲۳	۲۹	۱۶	۱۲	۱۸	۲۱	۲۷
میزان کربن (درصد)	۴/۱	۴/۹	۴/۷	۵/۰	۴/۶	۳/۶	۴/۹	۵/۱	۴/۸	۵/۲	۵/۲
شاخص نفوذپذیری	۱۳	۱۹	۲۲	۲۰	۱۶	۲۷	۲۱	۱۳	۱۸	۱۷	۱۱

الف- ضریب همبستگی نمونه را به دست آورید و آن را تعبیر کنید.

ب- یک فاصلهٔ اطمینان ۹۵ درصدی برای ρ تشکیل دهید.

۲۰ داده‌های زیر را در نظر بگیرید:

x	۴	۵	۹	۱۴	۱۸	۲۲	۲۴
y	۱۶	۲۲	۱۱	۱۶	۷	۳	۱۷

الف- ضریب همبستگی نمونه را به دست آورید و آن را تعبیر کنید.

ب- فرض $\rho=-۰/۸$ را در سطح معنی‌دار ۰/۰۵ آزمون کنید.

ضمیمه

جداول آماری

جدول I : توزیع دو جمله‌ای

جدول II : توزیع پواسون

جدول III : توزیع نرمال

جدول IV : توزیع مربع-کای

جدول V : توزیع t

جدول VI : توزیع F

جدول I : توزیع دو جمله‌ای

		p									
n	r	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90
5	0	0.5905	0.3277	0.2373	0.1681	0.0778	0.0312	0.0102	0.0024	0.0003	0.0000
	1	0.9185	0.7373	0.6328	0.5282	0.3370	0.1875	0.0870	0.0308	0.0067	0.0005
	2	0.9914	0.9421	0.8965	0.8369	0.6826	0.5000	0.3174	0.1631	0.0579	0.0086
	3	0.9995	0.9933	0.9844	0.9692	0.9130	0.8125	0.6630	0.4718	0.2627	0.0815
	4	1.0000	0.9997	0.9990	0.9976	0.9898	0.9688	0.9222	0.8319	0.6723	0.4095
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	0.3487	0.1074	0.0563	0.0282	0.0060	0.0010	0.0001	0.0000	0.0000	0.0000
	1	0.7361	0.3758	0.2440	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	0.0000
	2	0.9298	0.6778	0.5256	0.3828	0.1673	0.0547	0.0123	0.0016	0.0001	0.0000
	3	0.9872	0.8791	0.7759	0.6496	0.3823	0.1719	0.0548	0.0106	0.0009	0.0000
	4	0.9984	0.9672	0.9219	0.8497	0.6331	0.3770	0.1662	0.0474	0.0064	0.0002
	5	0.9999	0.9936	0.9803	0.9527	0.8338	0.6230	0.3669	0.1503	0.0328	0.0016
	6	1.0000	0.9991	0.9965	0.9894	0.9452	0.8281	0.6177	0.3504	0.1209	0.0128
	7	1.0000	0.9999	0.9996	0.9984	0.9877	0.9453	0.8327	0.6172	0.3222	0.0702
	8	1.0000	1.0000	1.0000	0.9999	0.9983	0.9893	0.9536	0.8507	0.6242	0.2639
	9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9990	0.9940	0.9718	0.8926	0.6513
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	0	0.2059	0.0352	0.0134	0.0047	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.5490	0.1671	0.0802	0.0353	0.0052	0.0005	0.0000	0.0000	0.0000	0.0000
	2	0.8159	0.3980	0.2361	0.1268	0.0271	0.0037	0.0003	0.0000	0.0000	0.0000
	3	0.9444	0.6482	0.4613	0.2969	0.0905	0.0176	0.0019	0.0001	0.0000	0.0000
	4	0.9873	0.8358	0.6865	0.5155	0.2173	0.0592	0.0094	0.0007	0.0000	0.0000
	5	0.9978	0.9389	0.8516	0.7216	0.4032	0.1509	0.0338	0.0037	0.0001	0.0000
	6	0.9997	0.9819	0.9434	0.8689	0.6098	0.3036	0.0951	0.0152	0.0008	0.0000
	7	1.0000	0.9958	0.9827	0.9500	0.7869	0.5000	0.2131	0.0500	0.0042	0.0000
	8	1.0000	0.9992	0.9958	0.9848	0.9050	0.6964	0.3902	0.1311	0.0181	0.0003
	9	1.0000	0.9999	0.9992	0.9963	0.9662	0.8491	0.5968	0.2784	0.0611	0.0023
	10	1.0000	1.0000	0.9999	0.9993	0.9907	0.9408	0.7827	0.4845	0.1642	0.0127
	11	1.0000	1.0000	1.0000	0.9999	0.9981	0.9824	0.9095	0.7031	0.3518	0.0556
	12	1.0000	1.0000	1.0000	1.0000	0.9997	0.9963	0.9729	0.8732	0.6020	0.1841
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9948	0.9647	0.8329	0.4510
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9953	0.9648	0.7941
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	0.1216	0.0115	0.0032	0.0008	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.3917	0.0692	0.0243	0.0076	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.6769	0.2061	0.0913	0.0355	0.0036	0.0002	0.0000	0.0000	0.0000	0.0000
	3	0.8670	0.4114	0.2252	0.1071	0.0160	0.0013	0.0001	0.0000	0.0000	0.0000
	4	0.9568	0.6296	0.4148	0.2375	0.0510	0.0059	0.0003	0.0000	0.0000	0.0000
	5	0.9887	0.8042	0.6172	0.4164	0.1256	0.0207	0.0016	0.0000	0.0000	0.0000
	6	0.9976	0.9133	0.7858	0.6080	0.2500	0.0577	0.0065	0.0003	0.0000	0.0000
	7	0.9996	0.9679	0.8982	0.7723	0.4159	0.1316	0.0210	0.0013	0.0000	0.0000
	8	0.9999	0.9900	0.9591	0.8867	0.5956	0.2517	0.0565	0.0051	0.0001	0.0000
	9	1.0000	0.9974	0.9861	0.9520	0.7553	0.4119	0.1275	0.0171	0.0006	0.0000
	10	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881	0.2447	0.0480	0.0026	0.0000
	11	1.0000	0.9999	0.9991	0.9949	0.9435	0.7483	0.4044	0.1133	0.0100	0.0001
	12	1.0000	1.0000	0.9998	0.9987	0.9790	0.8684	0.5841	0.2277	0.0321	0.0004
	13	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423	0.7500	0.3920	0.0867	0.0024
	14	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793	0.8744	0.5836	0.1958	0.0113
	15	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941	0.9490	0.7625	0.3704	0.0432
	16	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987	0.9840	0.8929	0.5886	0.1330
	17	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9964	0.9645	0.7939	0.3231
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9924	0.9308	0.6083
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9992	0.9885	0.8784
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

جدول II : توزیع پواسون

	μ								
r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6730	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

	μ								
r	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

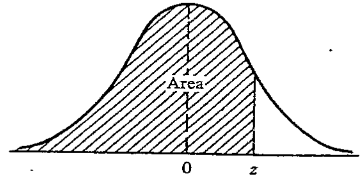
ادامه جدول II : توزیع پواسون

r	μ 5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0	0.0041	0.0025	0.0015	0.0009	0.0006	0.0003	0.0002	0.0001	0.0001
1	0.0266	0.0174	0.0113	0.0073	0.0047	0.0030	0.0019	0.0012	0.0008
2	0.0884	0.0620	0.0430	0.0296	0.0203	0.0138	0.0093	0.0062	0.0042
3	0.2017	0.1512	0.1118	0.0818	0.0591	0.0424	0.0301	0.0212	0.0149
4	0.3575	0.2851	0.2237	0.1730	0.1321	0.0996	0.0744	0.0550	0.0403
5	0.5289	0.4457	0.3690	0.3007	0.2414	0.1912	0.1496	0.1157	0.0885
6	0.6860	0.6063	0.5265	0.4497	0.3782	0.3134	0.2562	0.2068	0.1649
7	0.8095	0.7440	0.6728	0.5987	0.5246	0.4530	0.3856	0.3239	0.2687
8	0.8944	0.8472	0.7916	0.7291	0.6620	0.5925	0.5231	0.4557	0.3918
9	0.9462	0.9161	0.8774	0.8305	0.7764	0.7166	0.6530	0.5874	0.5218
10	0.9747	0.9574	0.9332	0.9015	0.8622	0.8159	0.7634	0.7060	0.6453
11	0.9890	0.9799	0.9661	0.9466	0.9208	0.8881	0.8487	0.8030	0.7520
12	0.9955	0.9912	0.9840	0.9730	0.9573	0.9362	0.9091	0.8758	0.8364
13	0.9983	0.9964	0.9929	0.9872	0.9784	0.9658	0.9486	0.9261	0.8981
14	0.9994	0.9986	0.9970	0.9943	0.9897	0.9827	0.9726	0.9585	0.9400
15	0.9998	0.9995	0.9988	0.9976	0.9954	0.9918	0.9862	0.9780	0.9665
16	0.9999	0.9998	0.9996	0.9990	0.9980	0.9963	0.9934	0.9889	0.9823
17	1.0000	0.9999	0.9998	0.9996	0.9992	0.9984	0.9970	0.9947	0.9911
18		1.0000	0.9999	0.9999	0.9997	0.9994	0.9987	0.9976	0.9957
19			1.0000	1.0000	0.9999	0.9997	0.9995	0.9989	0.9980
20					1.0000	0.9999	0.9998	0.9996	0.9991
21						1.0000	0.9999	0.9998	0.9996
22							1.0000	0.9999	0.9999
23								1.0000	0.9999
24									1.0000

ادامه جدول II : توزیع پواسون

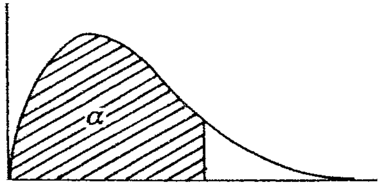
r	μ 10.0	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0
0	0.0000	0.0000	0.0000						
1	0.0005	0.0002	0.0001	0.0000	0.0000				
2	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000		
3	0.0103	0.0049	0.0023	0.0010	0.0005	0.0002	0.0001	0.0000	0.0000
4	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001
5	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003
6	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010
7	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029
8	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071
9	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154
10	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304
11	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549
12	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917
13	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426
14	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081
15	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867
16	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3750
17	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686
18	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622
19	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509
20	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307
21	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615	0.7991
22	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047	0.8551
23	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367	0.8989
24	1.0000	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594	0.9317
25		0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748	0.9554
26		1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848	0.9718
27			0.9999	0.9998	0.9994	0.9983	0.9959	0.9912	0.9827
28			1.0000	0.9999	0.9997	0.9991	0.9978	0.9950	0.9897
29				1.0000	0.9999	0.9996	0.9989	0.9973	0.9941
30					0.9999	0.9998	0.9994	0.9986	0.9967
31					1.0000	0.9999	0.9997	0.9993	0.9982
32						1.0000	0.9999	0.9996	0.9990
33							0.9999	0.9998	0.9995
34							1.0000	0.9999	0.9998
35								1.0000	0.9999
36									0.9999
37									1.0000

جدول III : توزیع نرمال



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

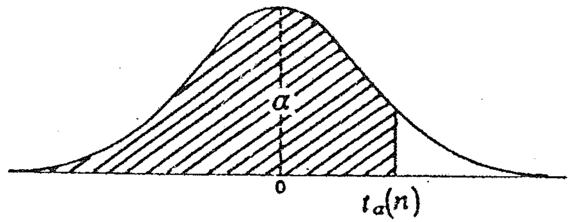
جدول IV : توزیع مربع-کای



$\chi^2_\alpha(n)$

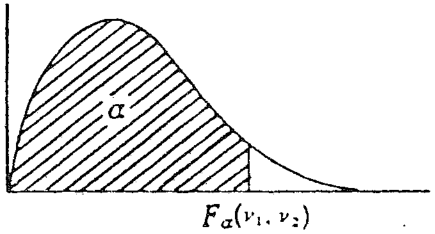
n	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.50}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.115	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

جدول V : توزیع t



n	$t_{.995}$	$t_{.99}$	$t_{.975}$	$t_{.95}$	$t_{.90}$	$t_{.80}$	$t_{.75}$	$t_{.70}$	$t_{.60}$	$t_{.55}$
1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	.727	.325	.158
2	9.92	6.96	4.30	2.92	1.89	1.061	.816	.617	.289	.142
3	5.84	4.54	3.18	2.35	1.64	.978	.765	.584	.277	.137
4	4.60	3.75	2.78	2.13	1.53	.941	.741	.569	.271	.134
5	4.03	3.36	2.57	2.02	1.48	.920	.727	.559	.267	.132
6	3.71	3.14	2.45	1.94	1.44	.906	.718	.553	.265	.131
7	3.50	3.00	2.36	1.90	1.42	.896	.711	.549	.263	.130
8	3.36	2.90	2.31	1.86	1.40	.889	.706	.546	.262	.130
9	3.25	2.82	2.26	1.83	1.38	.883	.703	.543	.261	.129
10	3.17	2.76	2.23	1.81	1.37	.879	.700	.542	.260	.129
11	3.11	2.72	2.20	1.80	1.36	.876	.697	.540	.260	.129
12	3.06	2.68	2.18	1.78	1.36	.873	.695	.539	.259	.128
13	3.01	2.65	2.16	1.77	1.35	.870	.694	.538	.259	.128
14	2.98	2.62	2.14	1.76	1.34	.868	.692	.537	.258	.128
15	2.95	2.60	2.13	1.75	1.34	.866	.691	.536	.258	.128
16	2.92	2.58	2.12	1.75	1.34	.865	.690	.535	.258	.128
17	2.90	2.57	2.11	1.74	1.33	.863	.689	.534	.257	.128
18	2.88	2.55	2.10	1.73	1.33	.862	.688	.534	.257	.127
19	2.86	2.54	2.09	1.73	1.33	.861	.688	.533	.257	.127
20	2.84	2.53	2.09	1.72	1.32	.860	.687	.533	.257	.127
21	2.83	2.52	2.08	1.72	1.32	.859	.686	.532	.257	.127
22	2.82	2.51	2.07	1.72	1.32	.858	.686	.532	.256	.127
23	2.81	2.50	2.07	1.71	1.32	.858	.685	.532	.256	.127
24	2.80	2.49	2.06	1.71	1.32	.857	.685	.531	.256	.127
25	2.79	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
26	2.78	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
27	2.77	2.47	2.05	1.70	1.31	.855	.684	.531	.256	.127
28	2.76	2.47	2.05	1.70	1.31	.855	.683	.530	.256	.127
29	2.76	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
30	2.75	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
40	2.70	2.42	2.02	1.68	1.30	.851	.681	.529	.255	.126
60	2.66	2.39	2.00	1.67	1.30	.848	.379	.527	.254	.126
120	2.62	2.36	1.98	1.66	1.29	.845	.677	.526	.254	.126
∞	2.58	2.33	1.96	1.645	1.28	.842	.574	.524	.253	.126

جدول VI : توزیع $F_{0/90}$



ν_2 \ ν_1 (NUMERATOR DEGREES OF FREEDOM)	1	2	3	4	5	6	7	8	9
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63

Denominator Degrees of Freedom

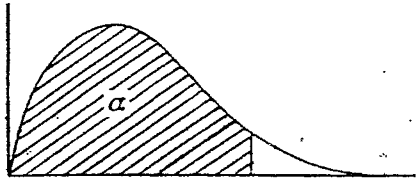
ادامه جدول VI : توزیع $F_{۰/۹۰}$

NUMERATOR DEGREES OF FREEDOM

Denominator Degrees of Freedom

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

ادامه جدول VI : توزیع $F_{۰/۹۵}$



$F_\alpha(\nu_1, \nu_2)$

NUMERATOR DEGREES OF FREEDOM

Denominator Degrees of Freedom

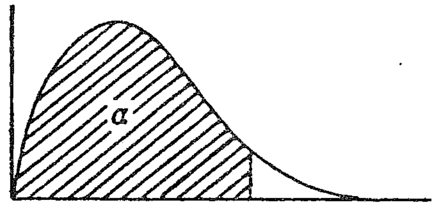
ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.77
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

ادامه جدول VI : توزیع $F_{۰/۹۵}$

ν_2 \ ν_1	NUMERATOR DEGREES OF FREEDOM									
	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Denominator Degrees of Freedom

ادامه جدول VI : توزیع $F_{۰/۹۹}$



$F_\alpha(\nu_1, \nu_2)$

ν_2 \ ν_1	NUMERATOR DEGREES OF FREEDOM								
	1	2	3	4	5	6	7	8	9
1	4,052	4,999.5	5,403	5,625	5,764	5,859	5,928	5,982	6,022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Denominator Degrees of Freedom

ادامه جدول VI : توزیع $F_{0/99}$

ν_2 \ ν_1	NUMERATOR DEGREES OF FREEDOM									
Denominator Degrees of Freedom	10	12	15	20	24	30	40	60	120	∞
1	6,056	6,106	6,157	6.209	6,235	6,261	6.287	6,313	6,339	6,366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

مراجع

1- Hodges, J. L. and Lehmann, E. L., Basic Concepts of Probability and Statistics, 2nd ed., San Francisco, Holden-Day, 1970.

2- Johnson, R. A. and Bhattacharyya, G. K., Statistics: Principles and Methods, 2nd ed., New york, John Wiley and Sons, 1992.

3- Larson, H. J., Introduction to Probability Theory and Statistical Inference, 2nd ed., New york, John Wiley and Sons, 1974.

4- Miller, I., Freund, J. E. and Johnson, R. A., Probability and Statistics for Engineers, 5th ed., Printice Hall, 1994.

5- Mood, A. M., Graybill, F. A. and Boes, D. C., Introduction to the Theory of Statistics, 3rd ed., New york, Mc Graw-Hill Book Company, 1974.

6- Ross, S., A First Course in Probability, 3rd ed., New york, Macmillan Publishing Company, 1989.

7- Walpole, R. E. and Myers, R. H., Probability and Statistics for Engineers and Scientists, 5th ed., New york, Macmillan Publishing Company, 1993.

۸- آمار و احتمال مقدماتی، تألیف دکتر جواد بهبودیان، چاپ سیزدهم (۱۳۷۸).