

پروژه درس آمار و روش تحقیق در نساچی

۱. پروژه شامل دو بخش است. بخش اول مربوط به تحلیل توصیفی داده ها است. بخش دوم در مورد فاصله اطمینان و آزمون های آماری می باشد. در این پروژه توصیف آماری داده ها و خروجی های بدست آمده اهمیت دارد.
۲. این پروژه بایستی با نرم افزار R انجام شود و در صورتی که با هر نرم افزار دیگر مانند متلب یا اکسل و یا به صورت دستی محاسبه شود نمره ای نخواهد داشت.
۳. نتایج تحلیل های آماری باید در قالب یک گزارش pdf شامل جداول، نمودار ها و خروجی های برنامه به همراه توصیف هرکدام از آن ها نوشته شود. همچنین هر یک از نمودارها باید به فرمت png ذخیره سازی و ارسال شود.
۴. کدهای R در فایل pdf قرار داده نشود و به صورت یک فایل R file باشد. فایل داده ی مورد استفاده نیز در همان قالب استفاده شده ارسال گردد. تمامی این فایل ها به صورت zip فشرده سازی و در سامانه قرار داده شود.
۵. داده ی تهیه شده برای هر گروه باید مختص همان گروه و مرتبط با رشته تحصیلی شما باشد و از یک داده برای دو گروه استفاده نشود. در ابتدای پروژه در مورد داده ها و متغیر های آن توضیح مختصری داده شود.
۶. داده ای متشکل از حداقل ۵۰ سطر و سه متغیر (دو متغیر کمی پیوسته و یک متغیر کیفی یا گسسته) تهیه نموده و آن را به فرمت .csv. و یا .txt. ذخیره سازی و در محیط R فراخوانی کنید (توجه داشته باشید اگر داده ها دارای اسم سر ستون هستند آنها را هنگام فراخوانی در نظر بگیرید).

نمونه داده:

sepal.Length	sepal.width	Petal.Length	Petal.width	species	type
5.1	3.5	1.4	0.2	setosa	2
4.9	3.0	1.4	0.2	setosa	2
4.7	3.2	1.3	0.2	setosa	2
4.6	3.1	1.5	0.2	setosa	1
5.0	3.6	1.4	0.2	setosa	1
5.4	3.9	1.7	0.4	setosa	3
6.7	3.3	5.7	2.5	virginica	3
6.7	3.0	5.2	2.3	virginica	1
6.3	2.5	5.0	1.9	virginica	3
6.5	3.0	5.2	2.0	virginica	2
6.2	3.4	5.4	2.3	virginica	1
5.9	3.0	5.1	1.8	virginica	1

بخش اول (آمار توصیفی):

۱. برای متغیرهای کیفی:

- جدول فراوانی (به همراه ستون فراوانی، درصد فراوانی، فراوانی تجمعی)
- نمودار میله ای (به همراه تخصیص اسم دلخواه به هر میله) بر حسب درصد فراوانی
- نمودار دایره ای بر حسب فراوانی

۲. برای متغیرهای پیوسته:

- شاخص های تمرکز (میانگین، میانه، نما)
- شاخص های پراکندگی (مینیمم، ماکسیمم، دامنه، انحراف معیار، واریانس)

۳. برای متغیرهای پیوسته نمودار مستطیلی با ۱۰ رده و نمودار جعبه ای (box plot) رسم کنید.

۴. بین دو متغیر پیوسته مقدار کوواریانس و ضریب همبستگی محاسبه کنید.

بخش دوم:

۱. یک فاصله اطمینان ۹۵ درصد برای میانگین یک متغیر کمی محاسبه کنید (به دو روش).

۲. یک فاصله اطمینان ۹۵ درصد برای واریانس نیز محاسبه کنید.

۳. برای یک متغیر پیوسته آزمون زیر را برای میانگین جامعه در سطح معناداری ۰,۰۵ انجام دهید. مقدار μ_0 را برابر با مجموع میانگین آن متغیر و ۱۰ در نظر بگیرید و نتیجه ی آن را بیان کنید.

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$$

۴. برای دو متغیر کمی پیوسته:

الف) آزمون همبستگی بین آنها را در سطح اطمینان ۹۵ درصد انجام دهید.

ب) یک مدل رگرسیون خطی ساده بر روی داده ها برازش دهید و برآورد پارامترهای مدل و تفسیر مناسب آن را بیان کنید.

ج) با توجه به خروجی بدست آمده آزمون های زیر را مورد صفر بودن ضرایب رگرسیونی را انجام دهید:

$$\begin{cases} H_0: \alpha = 0 \\ H_1: \alpha \neq 0 \end{cases} \quad \begin{cases} H_0: \beta = 0 \\ H_1: \beta \neq 0 \end{cases}$$

همچنین فاصله اطمینان ۹۵ درصد برای پارامترهای α و β بدست آورده و فرمول آن را نیز بیان کنید.