سلام

ما در این قسمت یک فرم جدید از توزیع های گسسته را برای حالت دو تایی و سه تایی و در نهایت k تایی بدست آوردیم و احتمال هر یک را محاسبه کردیم در قسمت توزیع مجانبی برای گشتاورر مرکزی مرتبه دوم Y ها نشان دادیم که بخشی از توزیع کی اسکور با یک مقدار ثابت جمع شده و میخواهیم شبیه سازی این قسمت را بدست آوریم که شکل مجانبی گشتاور مرتبه دوم شبیه کی اسکور باید بشود و برای گشتاورهای پایه ی زوج نیز این مطلب تعمیم پیدا کند.

New structure for discrete random variable

Let discrete random variable Y with probability modelP(Y $= a_1) = s_1$ and P(Y $= a_2) = s_2; s_1 + s_2 = 1$. We rewrite this variable as follows

$$Y = a_1X_1 + a_2(1 - X_1)$$

where $X_1 \sim B(1, p_1)$. In the case that Y takes three points a_1, a_2, a_3 with probabilities s_1, s_2, s_3 respectively, a variable representation of Y can be represented in terms of two independent Bernoulli random variables:

$$Y = a_1X_1 + a_2(1 - X_1)X_2 + a_3(1 - X_1)(1 - X_2)$$

where $X_1 \sim B(1, p_1)$ and $X_2 \sim B(1, p_2)$ are independent. It is observed that the expressions formed as a sum of random variables are not independent variables; For example, $(1 - X_1)X_2$ are not independent from $(1 - X_1)(1 - X_2)$. This point affects the calculations to some extent on the difficulty of obtaining the results. The adaptation of this representation for the discrete random variable Y when $a_1 = 1, a_2 = 0, a_3 = 2$ and $s_1 = 1/2, s_2 = s_3 = 1/4$ shows that a variable B(2,1/2) can be summed Don't display Bernoulli random variables. Although it becomes troublesome in the calculations of the characteristics of this variable, but in this article, from the aspect of studying the behavior of the limit distribution of central moments, it gives us a general result.

In general, suppose that Y is a discrete random variable with the following probability model

$$P(Y = a_i) = s_i; i = 1,2, \dots, k$$

Where $\sum_{i=1}^{k} s_i = 1, \quad s_1 \geq s_2 \geq \cdots \geq s_k > 0$. Consider independent Bernoulli random variables as $X_i \sim B(1, p_i), i = 1,2, \dots, (k - 1)$ and $q_i = 1 - p_i, \;\; i = 1,2, \dots k - 1$. The random variable Y can be represented as follows:

(1) $$Y = a_1X_1 + a_2(1 - X_1)X_2 + a_3(1 - X_1)(1 - X_2)X_3 + \cdots + a_k(1 - X_1) \dots (1 - X_{k-1})$$

With the equations as follows, it is easy to show the values of p_i in terms of s_i.

$$P(Y=a_1)=P(X_1=1)=s_1=p_1,$$
$$P(Y=a_2)=P(X_1=0,X_2=1)=s_2=q_1p_2,$$
(2) $$P(Y=a_3)=P(X_1=X_2=0,X_3=1)=s_3=q_1q_2p_3,$$
$$\vdots$$
$$P(Y=a_k)=P(X_1=X_2=\ ...=\ X_{k-1}=0)=s_k=q_1q_2\ldots q_{k-1}$$

From solving the above equations, we have:

$$p_1=s_1,\quad q_1=1-s_1$$

$$p_2=\frac{s_2}{1-s_1},\qquad q_2=1-\frac{s_2}{1-s_1}$$

$$p_3=\frac{s_3}{(1-s_1)\left(1-\frac{s_2}{1-s_1}\right)}=\frac{s_3}{1-s_1-s_2},\quad q_3=1-\frac{s_3}{1-s_1-s_2}$$

$$p_4=\frac{s_4}{(1-s_1)\left(1-\frac{s_2}{1-s_1}\right)\left(1-\frac{s_3}{1-s_1-s_2}\right)}=\frac{s_4}{1-s_1-s_2-s_3},\quad q_3=1-\frac{s_4}{1-s_1-s_2-s_3}$$

$$\vdots$$

$$p_{k-1}=\frac{s_{k-1}}{1-\sum_{i=1}^{k-2}s_i}$$

By (1), let's define W as follows:

$$W=a_2(1-X_1)X_2+a_3(1-X_1)(1-X_2)X_3+\cdots+a_k\prod_{i=1}^{k-1}(1-X_i).$$

Asymptotic distribution

With the following assumptions, we consider the variable Y in expression (1) as two components $Y=a_1X_1+W$. Now $Y_1,Y_2,\ldots,Y_n$ be identically independent random sample from (1), and $M_2=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\bar{Y})^2$ is the second central moment of the sample.

(6) $$\begin{aligned}nM_2&=\sum_{i=1}^{n}(Y_i-\bar{Y})^2\\&=\sum_{i=1}^{n}[a_1(X_{1i}-\overline{X_1})]^2+\sum_{i=1}^{n}(W_i-\overline{W})^2+2a_1\sum_{i=1}^{n}(X_{1i}-\overline{X_1})(W_i-\overline{W})\end{aligned}$$

For the first term of relation (6), we have:

$$\sum_{i=1}^{n} a_1^2 (X_{1i} - \overline{X_1})^2 = a_1^2 n M_2^{X_1}$$

Where $M_2^{X_1}$ is the second central moment from random sample of X_1 i.e $X_{11}, X_{12}, \dots, X_{1n}$. Abbasi (2008) and Georgios Afendras and Nickos Papadatos and Violetta Piperigou (2018) showed the following result under the condition of Bernoulli distribution with $p_1 = \frac{1}{2}$ success probability:

(7) $$a_1^2 n \left(M_2^{X_1} - \frac{1}{4}\right) \xrightarrow{D} a_1^2 \frac{-1}{4} \chi_{(1)}^2 .$$

The second and third terms of (6), by $n \to +\infty$, tend to

(8) $$Var(W) + 2a_1 Cov(X_1, W) = \sigma^2 - a_1^2 p_1 q_1 .$$

We display the limit expressions (7) and (8) as the limit result of (6).

(9) $$nM_2 \xrightarrow{D} a_1^2 \frac{-1}{4} \chi_{(1)}^2 + n a_1^2 \frac{1}{4} + \sigma^2 - a_1^2 p_1 q_1$$