PROBABILITY AND STATISTICS I
HOMEWORK VII

Jianyu Dong   2019511017

April, 13 2021

1

$\mathbf{Var}(X+Y)$ is smaller that $\mathbf{Var}(X-Y)$ for the following reason.
By the theorem, we have that

$$\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) + 2Cov(X,Y),$$
$$\mathbf{Var}(X-Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) - 2Cov(X,Y).$$

Since X and Y are negatively correlated, we could get that $\rho(X,Y) < 0$, then we have that $Cov(X,Y) = \rho(X,Y)\sigma_X\sigma_Y < 0$, so we could get that

$$\mathbf{Var}(X+Y) < \mathbf{Var}(X) + \mathbf{Var}(Y) < \mathbf{Var}(X-Y)$$

2

Assume that there exist two random variables X and Y satisfying the properties given by the problem. By the theorem, we could get that

$$\mathbf{Var}(X) = E(X^2) - E(X)^2 = 1,\ \mathbf{Var}(Y) = E(Y^2) - E(Y)^2 = 25$$

Then we could determine that

$$\sigma_X = \sqrt{\mathbf{Var}(X)} = 1,\ \sigma_Y = \sqrt{\mathbf{Var}(Y)} = 5$$

By the theorem, we could get that

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - 6 = -6$$

Thus, we get that

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = -\frac{6}{5} < -1$$

Which is in contradiction with the theorem that $\rho(X,Y) \geqslant -1$, so the assumption is false. So two random variables X and Y cannot possibly have the following properties: $E(X) = 3,\ E(Y) = 2,\ E(X^2) = 10,\ E(Y^2) = 29,\ and\ E(XY) = 0.$

3

Since the random variables X+Y and X-Y have finite variance $\sigma^2_{X+Y}, \sigma^2_{X-Y} < \infty$, we could get that

$$\begin{aligned} Cov(X+Y, X-Y) &= \mathbf{E}((X+Y)(X-Y)) - \mathbf{E}(X+Y)\mathbf{E}(X-Y) \\ &= \mathbf{E}(X^2) - \mathbf{E}(Y^2) - \mathbf{E}(X)^2 + \mathbf{E}(Y)^2 = \mathbf{Var}(X) - \mathbf{Var}(Y) = 0 \end{aligned}$$

By the definition, we have that

$$\rho(X+Y, X-Y) = \frac{Cov(X+Y, X-Y)}{\sigma_{X+Y}\sigma_{X-Y}} = 0$$

Thus, X+Y and X-Y are uncorrelated.

4

By the definition, we could get that

$$\mathbf{E}(X \mid Y) = \int_{-\infty}^{\infty} x f_X(x \mid y)\, dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} x f(x,y)\, dx$$

If $\mathbf{E}(X \mid Y)$ is a constant for all values of Y, we could get that $\mathbf{E}(X \mid Y) = \mathbf{E}(X)$ is independent with the random variable Y. Then we could determine that

$$\mathbf{E}(X) \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} x f(x,y)\, dx\, dy$$

Which shows that

$$\mathbf{E}(X)\mathbf{E}(Y) = \mathbf{E}(XY)$$

Since $0 < \mathbf{Var}(X) < \infty, 0 < \mathbf{Var}(Y) < \infty$, we could have that

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y} = 0$$

So that if $\mathbf{E}(X \mid Y)$ is a constant for all values of Y, then X and Y are uncorrelated.

5

By the theorem, we have that

$$\mathbf{E}(\mathbf{E}(Y \mid X)) = \mathbf{E}(Y)$$

So we could get that

$$\mathbf{E}(X_n) = \int_0^1 dx_1 \int_{x_1}^1 \frac{1}{1-x_1}\, dx_2 \int_{x_2}^1 \frac{1}{1-x_2}\, dx_3 \cdots \int_{x_{n-1}}^1 \frac{x_n}{1-x_{n-1}}\, dx_n = \frac{2^n - 1}{2^n}.$$

6

Since the joint distribution of X and Y is the uniform distribution on the region $x^2 + y^2 < 1$, using the fact that $\iint f(x,y)\, dx\, dy = 1$, we could determine the joint p.d.f. is

$$f(x,y) = \begin{cases} \frac{1}{\pi}, & for\ x^2 + y^2 < 1, \\ 0, & elsewhere. \end{cases}$$

For $-1 < y < 1$, we could determine that

$$\mathbf{E}(X \mid Y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{x}{\pi}\, dx = 0.$$

For $|y| \geqslant 1$, it is easily to get that $\mathbf{E}(X \mid Y) = 0$.
Then, we could calculate that

$$\mathbf{E}(X \mid Y) = 0.$$

7

By the definition, we could determine that

$$f_Y(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{x+y}{x+\frac{1}{2}}, & for\ 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, \\ 0, & elsewhere. \end{cases}$$

By the definition, we have that

$$\mathbf{E}(Y \mid X) = \int_{-\infty}^{\infty} y f_Y(y \mid x)\, dy = \begin{cases} \frac{3x+2}{6x+3}, & for\ 0 \leqslant x \leqslant 1, \\ 0, & elsewhere. \end{cases}$$

By the definition, we could calculate that

$$\mathbf{Var}(Y \mid X) = \mathbf{E}((Y - \mathbf{E}(Y \mid X))^2 \mid X) = \int_{-\infty}^{\infty} (y - \mathbf{E}(Y \mid X))^2 f_Y(y \mid x)\, dy$$
$$= \frac{6x^2 + 6x + 1}{18(2x+1)^2},\ for\ 0 \leqslant x \leqslant 1,$$

zero elsewhere.

8

Since U is a uniform distribution on (0,a), we could determine the probability distribution function is

$$f_U(u) = \frac{1}{a},\ for\ 0 < u < a,$$

zero elsewhere. Then we could determine that

$$\mathbf{E}(U) = \int_{-\infty}^{\infty} u f_U(u)\, du = \frac{a}{2}, \qquad \mathbf{E}(U^2) = \int_{-\infty}^{\infty} u^2 f_U(u)\, du = \frac{a^2}{3},$$
$$\mathbf{E}(U^3) = \int_{-\infty}^{\infty} u^3 f_U(u)\, du = \frac{a^3}{4}, \qquad \mathbf{E}(U^4) = \int_{-\infty}^{\infty} u^4 f_U(u)\, du = \frac{a^4}{5}.$$

So by the theorem, we could get that

$$\mathbf{Var}(U) = \mathbf{E}(U^2) - \mathbf{E}(U)^2 = \frac{a^2}{12},\ \mathbf{Var}(U^2) = \mathbf{E}(U^4) - \mathbf{E}(U^2)^2 = \frac{4a^4}{45}.$$

By the definition, we could get that

$$\rho(U, U^2) = \frac{Cov(U, U^2)}{\sigma_U \sigma_{U^2}} = \frac{\mathbf{E}(U^3) - \mathbf{E}(U)\mathbf{E}(U^2)}{\sqrt{\mathbf{Var}(U)\mathbf{Var}(U^2)}} = \frac{\sqrt{15}}{4}.$$

9

Assume Y follows the binomial distribution with the parameters n=9 and p=$\frac{1}{3}$, then we could determine the m.g.f. of Y is

$$M_Y(t) = E(e^{tY}) = \left(\frac{1}{3}e^t + \frac{2}{3}\right)^9,$$

the same as the m.g.f. of the random variable X. By the theorem, we could get that X follows the binomial distribution with the parameters n=9 and p=$\frac{1}{3}$. So the mean and variance of X are

$$\mu = np = 3,\ \sigma^2 = np(1-p) = 2.$$

The p.m.f. of X is

$$f_X(x) = \begin{cases} \binom{9}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{9-x}, & for\ x \in \{0,1,2,3,4,5,6,7,8,9\} \\ 0, & elsewhere. \end{cases}$$

So we could get that

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = P(3 - 2\sqrt{2} < X < 3 + 2\sqrt{2}) = \sum_{x=1}^{5} \binom{9}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{9-x}.$$

10

10.1

If the sampling is with replacement, we could get that the support of X is {0,1,2,3}. We could calculate that

$$P(X=0)=\binom{3}{0}\left(\frac{48}{52}\right)^3=\frac{1728}{2197}, \qquad P(X=1)=\binom{3}{1}\left(\frac{4}{52}\right)\left(\frac{48}{52}\right)^2=\frac{432}{2197},$$
$$P(X=2)=\binom{3}{2}\left(\frac{4}{52}\right)^2\left(\frac{48}{52}\right)=\frac{36}{2197}, \quad P(X=3)=\binom{3}{3}\left(\frac{4}{52}\right)^3=\frac{1}{2197}.$$

So if the sampling is with replacement, the p.m.f. of X is

$$p_X(x)=P(X=x)=\begin{cases}\frac{1728}{2197}, & for\ x=0,\\ \frac{432}{2197}, & for\ x=1,\\ \frac{36}{2197}, & for\ x=2,\\ \frac{1}{2197}, & for\ x=3.\end{cases}$$

10.2

If the sampling is without replacement, we could get that the support of X is {0,1,2,3}. We could calculate that

$$P(X=0)=\frac{\binom{48}{3}}{\binom{52}{2}}=\frac{4324}{5525}, \qquad P(X=1)=\frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{4}}=\frac{1128}{5525},$$
$$P(X=2)=\frac{\binom{4}{2}\binom{48}{1}}{\binom{52}{3}}=\frac{72}{5525}, \quad P(X=3)=\frac{\binom{4}{3}}{\binom{52}{3}}=\frac{1}{5525}.$$

So if the sampling is without replacement, the p.m.f. of X is

$$p_X(x)=P(X=x)=\begin{cases}\frac{4324}{5525}, & for\ x=0,\\ \frac{1128}{5525}, & for\ x=1,\\ \frac{72}{5525}, & for\ x=2,\\ \frac{1}{5525}, & for\ x=3.\end{cases}$$

11

Since X follows the hypergeometric distribution with parameters (N, M, n), by the definition, we could get the p.m.f. of X is

$$p_X(x)=P(X=x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \ \ for\ \ \max\{0,n-N+M\}\leqslant x\leqslant\min\{n,M\},$$

zero elsewhere.

By the definition, we could get that

$$\begin{aligned}\mathbf{E}(X(X-1))&=\sum_{x=0}^{n}x(x-1)p_X(x)=\sum_{x=2}^{n}x(x-1)\binom{N-M}{n-x}\frac{M!}{(M-x)!x!}\frac{(N-n)!n!}{N!}\\&=\frac{M(M-1)n(n-1)}{N(N-1)}\sum_{x=2}^{n}\binom{N-M}{n-x}\binom{M-2}{x-2}\binom{N-2}{n-2}^{-1}\\&=\frac{M(M-1)n(n-1)}{N(N-1)}\end{aligned}$$

Using the fact that the sum of the probability of a hypergeometric distribution with parameters (N-2,M-2,n-2) is 1.

For a fixed k, by the definition, we could get that

$$\begin{aligned}
\mathbf{E}\left(\binom{X}{k}\right) &= \sum_{x=0}^{n}\binom{x}{k}\binom{N-M}{n-x}\frac{M!}{(M-x)!x!}\frac{(N-n)!n!}{N!} \\
&= \frac{1}{k!}\frac{M!}{(M-k)!}\frac{n!}{(n-k)!}\frac{(N-k)!}{N!}\sum_{x=k}^{n}\binom{N-M}{n-x}\binom{M-k}{x-k}\binom{N-k}{n-k}^{-1} \\
&= \frac{\binom{M}{k}\binom{n}{k}}{\binom{N}{k}}
\end{aligned}$$

Using the fact that the sum of the probability of a hypergeometric distribution with parameters (N-k,M-k,n-k) is 1.

12

12.1

If Y is continuous with conditional p.d.f. by $f_Y(y \mid x)$, then by the definition, we could get that

$$\mathbf{E}(g(X)Y \mid X) = \int_{-\infty}^{\infty} g(x)y \cdot f_Y(g(x)y \mid x)\,dy = g(x)\int_{-\infty}^{\infty} y \cdot f_Y(y \mid x)\,dy = g(X)\mathbf{E}(Y \mid X).$$

Similarly for discrete Y with conditional p.m.f. $p_Y(y \mid x)$ we have

$$E$$

12.2

If Y is continuous with conditional p.d.f. by $f_Y(y \mid x)$, then by the definition, we could get that

$$\mathbf{E}(X\mathbf{E}(Y \mid X)) = \int_{-\infty}^{\infty} x\mathbf{E}(Y \mid x)f_X(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_Y(y \mid x)f_X(x)\,dx\,dy.$$

Since $f_Y(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$, we could get that

$$\mathbf{E}(X\mathbf{E}(Y \mid X)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\,dx\,dy = E(XY)$$

12.3

By the definition, we have that

$$Cov(X, \mathbf{E}(Y \mid X)) = \mathbf{E}(X\mathbf{E}(Y \mid X)) - \mathbf{E}(X)\mathbf{E}(\mathbf{E}(Y \mid X)).$$

By the definition, we could get that

$$\mathbf{E}(X\mathbf{E}(Y \mid X)) = \int_{-\infty}^{\infty} x\int_{-\infty}^{\infty} yf_Y(y \mid x)\,dy f_X(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{X,Y}(x,y)\,dx\,dy = \mathbf{E}(XY)$$

By the theorem, we have that

$$\mathbf{E}(\mathbf{E}(Y \mid X)) = \mathbf{E}(Y)$$

Thus, we get that

$$Cov(X, \mathbf{E}(Y \mid X)) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = Cov(X, Y)$$

13

Since $X_1, \ldots$ be i.i.d. random variables and N is a random variable taking values in Z^+, we could get that

$$\mathbf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathbf{E}(NX_1)$$

Further N and $\{X_i\}$ are independent, by the theorem, we could get that

$$\mathbf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathbf{E}(NX_1) = \mathbf{E}(X_1)\mathbf{E}(N)$$