Lancaster University College
at Beijing Jiaotong University

2019/20 Examinations

Course code: WB73L004Q

Course name: Probability Theory and Mathematical Statistics (B)

Final examination (January)

INSTRUCTIONS TO STUDENTS

1) Duration of the exam: 120 minutes
2) This paper contains 3 pages. There are 8 questions.
3) You must answer all questions.
4) This is a closed book exam. No books or notes may be brought into the exam room.
5) A scientific calculator is allowed in the examination. Other electronic devices are not allowed in the exam room.
6) Some values that might be useful:

$$\Phi(1) = 0.8413, \qquad \Phi(1.5921) = 0.9443,$$

$$t_{0.025}(15) = 2.1314.$$

1. (20pt)

(1) Two events A and B are independent, and $P(A) = 0.6$, $P(B) = 0.5$, then $P(A|A \cup \overline{B}) =$ 0.75 .

(2) If n is large and p is small, then the binomial distribution $B(n, p)$ can be approximated by *Poisson* distribution with parameter $\lambda = np$.

(3) If random variables $X \sim B\left(18, \frac{1}{3}\right)$, $Y \sim P(3)$, and X and Y are independent, then $\text{Var}(X - Y) =$ 7.

(4) Suppose that random variable X has $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$. By Chebyshev's inequality, $P(|X - \mu| \geq 3\sigma) \leq$ 1/9.

(5) Given observed data of a random sample: $(14, 20, 2, 16, 3)$, its sample mean is *11*, sample variance is *65*.

(6) Random variables $X_1, X_2, X_3, X_4, X_5, X_6$ are independent and all have $N(0,1)$. If $\frac{c(X_1 - X_2)}{\sqrt{(X_3 - X_4)^2 + (X_5 - X_6)^2}} \sim t(m)$, then $c =$ ____$\sqrt{2}$____ and $m =$ ___2___.

(7) One class of 16 students had an English test. The sample mean and standard deviation of scores are 80 and 8, respectively. A 95% confidence interval of the mean score is (rounded to the nearest thousandth) *(75.737, 84.263)*.

2. **(10pt)** Random point (X, Y) is uniformly distributed in the unit disk $x^2 + y^2 < 1$. And we define two discrete random variables U and V as follows:

$$U = \begin{cases} 1, & X^2 + Y^2 < \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

$$V = \begin{cases} 1, & X > 0 \text{ and } Y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

(1) Determine the joint PF of U and V.

(2) Are U and V independent?

Solution: (1) It is clear that $P(U = 1) = P\left(X^2 + Y^2 < \frac{1}{2}\right) = \frac{1}{2}$. And,

$$P(U = 0) = P\left(X^2 + Y^2 \geq \frac{1}{2}\right) = \frac{1}{2}.$$

$$P(U = 1, V = 1) = P\left(X^2 + Y^2 < \frac{1}{2}, (X > 0 \text{ and } Y > 0)\right) = \frac{1}{8}.$$

$$P(U = 1, V = 0) = P\left(X^2 + Y^2 < \frac{1}{2}\right) - P\left(X^2 + Y^2 < \frac{1}{2}, (X > 0 \text{ and } Y > 0)\right) = \frac{3}{8}.$$

$$P(U = 0, V = 1) = P\left(X^2 + Y^2 \geq \frac{1}{2}, (X > 0 \text{ and } Y > 0)\right) = \frac{1}{8}.$$

$$P(U=0,V=0)=P\left(X^2+Y^2\ge\tfrac{1}{2}\right)-P\left(X^2+Y^2\ge\tfrac{1}{2},(X>0\text{ and }Y>0)\right)=\tfrac{3}{8}.$$

probability	V=0	V=1
U=0	$\frac{3}{8}$	$\frac{1}{8}$
U=1	$\frac{3}{8}$	$\frac{1}{8}$

(2) It is easy to see that U and V are independent.

3. **(12pt)** The joint PDF of X and Y is given by

$$f(x,y)=\begin{cases}\dfrac{1}{2x^2y}, & x\ge 1,\dfrac{1}{x}\le y\le x,\\ 0\ , & \text{otherwise.}\end{cases}$$

(1) Find the PDF of Y.

(2) Evaluate $E(\frac{Y}{X})$.

Solution: (1) For $0<y<1$, we have

$$f_Y(y)=\int_{\frac{1}{y}}^{\infty}\frac{1}{2x^2y}dx=\frac{1}{2y}\int_{\frac{1}{y}}^{\infty}\frac{1}{x^2}dx=\frac{1}{2y}\left[-\frac{1}{x}\right]\Big|_{\frac{1}{y}}^{\infty}=\frac{1}{2}.$$

For $y\ge 1$, we have

$$f_Y(y)=\int_{y}^{\infty}\frac{1}{2x^2y}dx=\frac{1}{2y}\int_{y}^{\infty}\frac{1}{x^2}dx=\frac{1}{2y}\left[-\frac{1}{x}\right]\Big|_{y}^{\infty}=\frac{1}{2y^2}.$$

Thus,

$$f_Y(y)=\begin{cases}\dfrac{1}{2}, & 0<y<1,\\ \dfrac{1}{2y^2}, & y\ge 1,\\ 0, & \text{otherwise.}\end{cases}$$

(2)

$$E\left(\frac{Y}{X}\right)=\int_1^{\infty}\int_{\frac{1}{x}}^{x}\frac{y}{x}\frac{1}{2x^2y}dy\,dx=\int_1^{\infty}\int_{\frac{1}{x}}^{x}\frac{1}{2x^3}dy\,dx$$

$$=\int_1^{\infty}\frac{1}{2x^3}\left(x-\frac{1}{x}\right)dx=\frac{1}{2}\int_1^{\infty}(x^{-2}-x^{-4})dx=\frac{1}{2}\left[-x^{-1}+\frac{1}{3}x^{-3}\right]\Big|_1^{\infty}$$

$$=\frac{1}{2}\left[x^{-1}-\frac{1}{3}x^{-3}\right]\Big|_{\infty}^{1}=\frac{1}{3}.$$

4. **(10pt)** Suppose that the PDF of a random variable X is

$$f_X(x) = \begin{cases} 3x^2, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

The conditional PDF of Y given $X = x$ is

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{x}, & 0 < y < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(1) Find the joint PDF of X and Y.

(2) Find $P(X + Y < 1)$.

Solution: (1) Let $f(x, y)$ be the joint PDF of X and Y. Then

$$f(x, y) = f_X(x) f_{Y|X}(y|x) = \begin{cases} 3x, & 0 < y < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(2) $$P(X + Y < 1) = \int_0^{\frac{1}{2}} \int_y^{1-y} 3x dx\, dy = \int_0^{\frac{1}{2}} \left[\frac{3}{2}x^2\right] |_y^{1-y} dy$$

$$= \frac{3}{2} \int_0^{\frac{1}{2}} (1 - 2y) dy = \frac{3}{2} [y - y^2]|_0^{\frac{1}{2}} = \frac{3}{8}.$$

5. **(10pt)** Suppose that random variables X, Y, and Z are independent and all have the standard normal distribution. Evaluate $P(3X + 2Y < 6Z - 7)$.

Solution: Let $W = 3X + 2Y - 6Z$. Then

$$P(3X + 2Y < 6Z - 7) = P(W < -7).$$

We have $W \sim N(0, 9 + 4 + 36) = N(0, 49)$. Thus $\frac{W}{7} \sim N(0,1)$. The desired result is

$$P(W < -7) = P\left(\frac{W}{7} < -1\right) = \Phi(-1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587.$$

6. **(12pt)** Let random variables X and Y be independent, and $\mathrm{Var}(X) = \mathrm{Var}(Y) = \sigma^2$. Let $U = 2X + Y$ and $V = 2X - Y$. Determine $\rho_{U,V}$.

Solution: We have

$$\mathrm{Var}(U) = 4\mathrm{Var}(X) + \mathrm{Var}(Y) = 5\sigma^2.$$

$$\mathrm{Var}(V) = 4\mathrm{Var}(X) + \mathrm{Var}(Y) = 5\sigma^2.$$

$$\mathrm{Cov}(U,V) = \mathrm{Cov}(2X+Y, 2X-Y)$$
$$= \mathrm{Cov}(2X,2X) + \mathrm{Cov}(Y,2X) - \mathrm{Cov}(2X,Y) - \mathrm{Cov}(Y,Y)$$
$$= 4\mathrm{Var}(X) - \mathrm{Var}(Y) = 3\sigma^2.$$

Hence, $\rho_{U,V} = \frac{\mathrm{Cov}(U,V)}{\sqrt{\mathrm{Var}(U)\mathrm{Var}(V)}} = \frac{3\sigma^2}{\sqrt{5\sigma^2 \cdot 5\sigma^2}} = \frac{3}{5}.$

7. **(12pt)** An election is held in a small town between two candidates A and B. An initial counting of the votes shows 1422 votes for A and 1405 votes for B. However, further counting reveals that 101 votes are illegal and have to be thrown out. Suppose that the illegal votes are independent from each other and each illegal vote is equally probable for A or for B. Use the central limit theorem with correction for continuity to approximately evaluate the probability that the removal of the illegal votes changes the result of the election. Note that there are only 3 results: A wins, B wins, or they tie.

Solution: The total number of legal votes is 1422+1405-101=2726. Hence if the final votes for A become 1363 or fewer, then the result of the election is changed, which is equivalent to there being 1422-1363=59 or more illegal votes for A. Let X be the number of illegal votes for A. Then $X \sim B(101, 0.5)$.

$$E(X) = 101 \cdot 0.5 = 50.5$$
$$\mathrm{Var}(X) = 101 \cdot 0.5 \cdot 0.5 = 25.25.$$

By the central limit theorem, X approximately has $N(50.5,\ 25.25)$. Using correction for continuity, the desired probability is

$$P(X \geq 59) = P(X > 58.5) = P(\frac{X - 50.5}{\sqrt{25.25}} > \frac{58.5 - 50.5}{\sqrt{25.25}})$$
$$\approx 1 - \Phi\left(\frac{58.5-50.5}{\sqrt{25.25}}\right) = 1 - \Phi(1.5921) = 1 - 0.9443 = 0.0557.$$

8. **(14pt)** Suppose that $X_1, \cdots, X_n$ form a random sample from a continuous X which has the following PDF with parameter $\theta > 0$:

$$f(x|\theta) = \begin{cases} \frac{2}{\sqrt{\pi\theta}} \exp\left(-\frac{x^2}{\theta}\right), & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

It is known that $\int_0^\infty t^{-\frac{1}{2}} e^{-t} dt = \sqrt{\pi}$ and $\int_0^\infty t^{\frac{1}{2}} e^{-t} dt = \frac{\sqrt{\pi}}{2}$.

(1) Find $\mu_2 = E(X^2)$, and derive a moment estimator of θ.

(2) Derive the maximum likelihood estimator of θ.

Solution: (1)

$$\mu_2 = E(X^2) = \int_0^\infty x^2 \frac{2}{\sqrt{\pi\theta}} \exp(-\frac{x^2}{\theta}) dx$$

Let $t = \frac{x^2}{\theta}$, then $x = \sqrt{\theta t}$ and $dx = \frac{\sqrt{\theta}}{2\sqrt{t}} dt$. Then

$$\mu_2 = E(X^2) = \int_0^\infty \theta t \frac{2}{\sqrt{\pi\theta}} \exp(-t) \frac{\sqrt{\theta}}{2\sqrt{t}} dt$$

$$= \frac{\theta}{\sqrt{\pi}} \int_0^\infty \sqrt{t} \exp(-t) dt = \frac{\theta}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} = \frac{\theta}{2}.$$

Thus $\theta = 2\mu_2$, and a moment estimator of θ is

$$\hat{\theta}_{MOM} = 2A_2 = \frac{2}{n} \sum_{i=1}^n X_i^2 .$$

(2) Given observed data $x_1, \ldots, x_n$, the likelihood function

$$L(\theta) = \prod_{i=1}^n \frac{2}{\sqrt{\pi\theta}} \exp(-\frac{x_i^2}{\theta}) = 2^n \pi^{-\frac{n}{2}} \theta^{-\frac{n}{2}} \exp(-\frac{\sum_{i=1}^n x_i^2}{\theta})$$

$$\ln L(\theta) = \ln(2^n \pi^{-\frac{n}{2}}) - \frac{n}{2} \ln\theta - \frac{\sum_{i=1}^n x_i^2}{\theta}$$

$$\frac{d \ln L(\theta)}{d\theta} = -\frac{n}{2}\frac{1}{\theta} + \frac{\sum_{i=1}^n x_i^2}{\theta^2} = \frac{1}{\theta}(\frac{\sum_{i=1}^n x_i^2}{\theta} - \frac{n}{2})$$

Obviously, $\frac{d \ln L(\theta)}{d\theta} = 0$ if and only if $\theta = \frac{2}{n}\sum_{i=1}^{n} x_i^2$.

Therefore, the maximum likelihood estimator of θ is

$$\hat{\theta}_{ML} = \frac{2}{n}\sum_{i=1}^{n} X_i^2 .$$

In this problem, the two estimators are equal.