Lancaster University College
at Beijing Jiaotong University

2019/20 Examinations

Course code: WB73L004Q

Course name: Probability Theory and Mathematical Statistics (B)

Resit examination (March)

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:

$$\Phi(0.5634) = 0.7134, \qquad \Phi(1) = 0.8413,$$

$$\Phi(1.3363) = 0.9093, \qquad t_{0.05}(4) = 2.1318.$$

1. (20pt)

(1) The joint PDF of X and Y is $f(x,y) = \begin{cases} ae^{-2x-4y}, & x > 0, y > 0, \\ 0\ , & \text{otherwise.} \end{cases}$ then $a =$ 8.

(2) A family is chosen at random from all three-child families. If it is known that the family has a boy among the three children, then the conditional probability that the chosen family has one boy and two girls is 3/7.

(3) If random variables $X \sim B(4, 0.3)$, $Y \sim B(11, 0.3)$, and X, Y are independent, then $X + Y \sim$ $B(15, 0.3)$.

(4) If random variables X, Y have a joint continuous distribution, then $P(X = Y) =$ 0.

(5) If random variable $X \sim U(1,3)$, then $\mathrm{Var}(X) =$ 1/3.

(6) Suppose that X_1, X_2 have a bivariate normal distribution. Then X_1, X_2 are independent if and only if $\mathrm{Cov}(X_1, X_2)$ $= 0$. (or ρ_{X_1,X_2})

(7) Let $X_1, X_2, \cdots, X_n$ be a random sample from X with $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$. Then $E[(X_2 - X_1)^2]$ $= 2\sigma^2$, and if $c\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2$ is an unbiased estimator of σ^2 then $c = \frac{1}{2(n-1)}$.

(8) Let X_1, X_2, X_3, X_4 be a random sample from distribution $N(\mu, \sigma^2)$. $\bar{X}$ is the sample mean. Suppose that $c\sum_{i=1}^{4}(X_i - \bar{X})^2 \sim \chi^2(m)$. Then $c = \frac{1}{\sigma^2}$ and $m =$ 3.

2. **(10pt)** In a world, 40% of products are highly successful, 35% are moderately successful, and 25% are poor products. In addition, 95% of highly successful products receive good reviews, 60% of moderately successful products receive good reviews, and 10% of poor products receive good reviews.

(1) What is the probability that a product attains a good review?

(2) If a new product attains a good review, what is the probability that it will be a highly successful product?

(3) If a product does not attain a good review, what is the probability that it will be a highly successful product?

Solution: Let H denote the event of being highly successful, M the event of being moderately successful, and P the event of being poor. Let G denote the event of having a good review.

(1) $P(G) = P(G|H)P(H) + P(G|M)P(M) + P(G|P)P(P)$

$= 0.95 \times 0.4 + 0.6 \times 0.35 + 0.1 \times 0.25 = 0.615.$

(2) $P(H|G) = \frac{P(HG)}{P(G)} = \frac{P(G|H)P(H)}{P(G)} = \frac{0.95\times 0.4}{0.615}$=0.6179.

(3) $P(H|\bar{G})=\frac{P(H\bar{G})}{P(\bar{G})}=\frac{P(H)-P(HG)}{1-P(G)}=\frac{0.4-0.38}{1-0.615}=0.0325.$

3. **(10pt)** Two random variables X and Y are independent, $X \sim N(2,4)$, $Y \sim N(3,5)$. Determine $P(3X+2Y\leq 22)$.

Solution: It is clear that $3X+2Y \sim N(3\times 2+2\times 3, 9\times 4+4\times 5)=N(12,56)$

$$P(3X+2Y\leq 22)=P\left(\frac{3X+2Y-12}{\sqrt{56}}\leq\frac{22-12}{\sqrt{56}}\right)=\Phi(1.3363)$$

$$=0.9093.$$

4. **(12pt)** The joint PDF of X and Y is given by

$$f(x,y)=\begin{cases}12\ \ y^2, & 0\leq x\leq 1, 0\leq\ \ y\leq x,\\ 0\ , & \text{otherwise.}\end{cases}$$

(1) Find $\mathrm{Cov}(X,Y)$.

(2) Find $P(X+Y\leq 1)$.

Solution: (1)

$$E(X)=\iint\ \ xf(x,y)dxdy=\int_0^1\int_0^x x\cdot 12y^2dy\,dx=\int_0^1 4x^4dx=\frac{4}{5}$$

$$E(Y)=\iint yf(x,y)dxdy=\int_0^1\int_0^x y\cdot 12y^2dy\,dx=\int_0^1 3x^4dx=\frac{3}{5}$$

$$E(XY)=\iint xyf(x,y)dxdy=\int_0^1\int_0^x xy\cdot 12y^2dy\,dx=\int_0^1 3x^5dx=\frac{1}{2}$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{2}-\frac{4}{5}\frac{3}{5}=\frac{1}{50}.$$

(2)

$$P(X+Y\leq 1)=\int_0^{\frac{1}{2}}\int_y^{1-y}12\ \ y^2dx\,dy=\int_0^{\frac{1}{2}}12\ \ y^2(1-2y)dy$$

$$=[4y^3-6y^4]|_0^{1/2}=\frac{1}{8}.$$

5. **(13pt)** Let X denote the total number of successes in 15 Bernoulli trials, with probability of success $p\ =\ 0.3$ on each trial.

(1) Determine $P(X\ =\ 4)$.

(2) Determine approximately the value of $P(X\ =\ 4)$ by using the central limit theorem with the correction for continuity.

Solution: (1) It is clear that $X\sim B(15,0.3)$. So

$$P(X=4)=C_{15}^{4}0.3^{4}0.7^{11}=\frac{15\cdot 14\cdot 13\cdot 12}{24}0.3^{4}0.7^{11}$$

$$=\frac{15\cdot 7\cdot 13\cdot 3^{4}\cdot 7^{11}}{10^{15}}=\frac{15\cdot 13\cdot 3^{4}\cdot 7^{12}}{10^{15}}=0.2186$$

(2) $E(X)=15\cdot 0.3=4.5$, $\mathrm{Var}(X)=15\cdot 0.3\cdot 0.7=3.15$. By the central limit theorem, X approximately has $N(4.5,\ 3.15)$. Using correction for continuity,

$$P(X=4)=P(3.5<X<4.5)=P(\frac{3.5-4.5}{\sqrt{3.15}}<\frac{X-4.5}{\sqrt{3.15}}<0)$$

$$\approx\Phi(0)-\Phi\left(\frac{-1}{\sqrt{3.15}}\right)=\Phi\left(\frac{1}{\sqrt{3.15}}\right)-0.5=\Phi(0.5634)-0.5$$

$$=0.7134-0.5=0.2134.$$

6. **(12pt)** Suppose that X_1,X_2,X_3,X_4,X_5 are a random sample from distribution $N(\mu,\sigma^2)$, where μ and σ^2 are both unknown. We have the following observed data of the sample: $(x_1,x_2,x_3,x_4,x_5)=(29,20,11,16,14)$. Find a 90% confidence interval for μ.

Solution: We have $n=5$, $\bar{x}=\frac{29+20+11+16+14}{5}=18$. $\alpha=0.1$

$$s^2=\frac{1}{4}[(29-18)^2+(20-18)^2+(11-18)^2+(16-18)^2+(14-18)^2]$$

=48.5.

$$\mathrm{s}=\sqrt{s^2}=\sqrt{48.5}=6.9642$$

$$t_{\frac{\alpha}{2}}(n-1)=t_{0.05}(4)=2.1318$$

$$\frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)=\frac{6.9642}{\sqrt{5}}\cdot 2.1318=6.6395.$$

Therefore a 90% confidence interval for μ is

$$\left(\bar{x}-\frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\bar{x}+\frac{s}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)=(18-6.6395,18+6.6395)=$$

$$(11.3605,24.6395).$$

7. **(10pt)** Let X_1, X_2 be a random sample from X with $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$. We want $Y = a_1X_1 + a_2X_2$ to be an estimator of μ, where a_1, a_2 are real constants.

(1) For Y to be unbiased, what requirement must a_1, a_2 fulfil?

(2) Find $\text{Var}(Y)$.

(3) Determine a_1, a_2 if we want Y to be unbiased and the most efficient.

Solution: (1) $E(Y) = a_1E(X_1) + a_2E(X_2) = (a_1 + a_2)\mu$. For $E(Y) = \mu$, it must be satisfied that $a_1 + a_2 = 1$.

(2) $\text{Var}(Y) = a_1^2\text{Var}(X_1) + a_2^2\text{Var}(X_2) = (a_1^2 + a_2^2)\sigma^2$.

(3) Now suppose that Y is unbiased and the most efficient. Then $a_1 + a_2 = 1$, and

$\text{Var}(Y) = (a_1^2 + a_2^2)\sigma^2 = (a_1^2 + (1 - a_1)^2)\sigma^2 = (2a_1^2 - 2a_1 + 1)\sigma^2$.

It is easy to verify that when $a_1 = \frac{1}{2}$, $\text{Var}(Y)$ attains its minimum, i.e. the most efficient. Therefore in that situation, $a_1 = a_2 = \frac{1}{2}$.

8. **(13pt)** Let $X_1, X_2, \cdots, X_n$ be a random sample from X whose PDF is

$$f(x) = \begin{cases} \theta x^{-(\theta+1)}, & x > 1, \\ 0 \quad, & \text{otherwise}, \end{cases}$$

where $\theta > 1$ is an unknown parameter.

(1) Find the moment estimator of θ.

(2) Find the maximum likelihood estimator of θ.

Solution: (1) $\mu_1 = E(X) = \int_1^\infty x\theta x^{-\theta-1}dx = \int_1^\infty \theta x^{-\theta}dx = \frac{\theta}{-\theta+1}x^{-\theta+1}|_1^\infty = \frac{\theta}{\theta-1}$.

Therefore, $\theta = \frac{\mu_1}{\mu_1 - 1}$. And the moment estimator of θ is

$$\hat{\theta}_{MOM} = \frac{A_1}{A_1 - 1}, \text{ where } A_1 = \frac{1}{n}\sum_{i=1}^n X_i.$$

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

$$L(\theta) = \prod_{i=1}^n \theta x_i^{-(\theta+1)} = \theta^n (\prod_{i=1}^n x_i)^{-(\theta+1)}$$

$$lnL(\theta) = nln\theta - (\theta + 1)\sum_{i=1}^n \ln(x_i)$$

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

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

Therefore, the maximum likelihood estimator of θ is

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