《Probability and Statistics》

2022-2023（2）期末考试卷参考答案及评分标准

Some possible useful critical values：

$z_{0.025} = 1.96$；$z_{0.05} = 1.645$；

$t_{0.025,8} = 2.306$；$t_{0.025,9} = 2.262$；$t_{0.05,8} = 1.86$；$t_{0.05,9} = 1.833$；

$\chi^2_{0.025,8} = 17.534$; $\chi^2_{0.025,9} = 19.022$; $\chi^2_{0.975,8} = 2.18$; $\chi^2_{0.975,9} = 2.70$;

I. **(20 points)** In a game, A, B and C take turns flipping a fair coin. The first one to get a head wins. Assume that A flips first, then B, then C, then A, and so on. Let X be the number of flips in total when the game ends.

a) Find the probability that B wins in his/her fifth flip.

b) Find the probability that C wins.

c) Calculate $\mathbb{P}(X \leq 2023)$.

d) Find the expected value of X.

Solution.

a) The probability that B win in his/her fifth flip is

$$\left(\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\right)^4\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2^{14}}.$$ ……………………………5 分

b) Let C_n be the event that C wins in his/her nth flip. Then

$$\mathbb{P}(C_n) = \left(\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\right)^n = \frac{1}{8^n}$$

and

$$\mathbb{P}(\text{C wins}) = \sum_{n=1}^{+\infty}\mathbb{P}(C_n) = \sum_{n=1}^{+\infty}\frac{1}{8^n} = \frac{1}{7}.$$ ……………………10 分

c) Since $\mathbb{P}(X = n) = \frac{1}{2^n}$, we have

$$\mathbb{P}(X \leq 2023) = \sum_{i=1}^{2023}\mathbb{P}(X = i) = \sum_{i=1}^{2023}\frac{1}{2^i} = 1 - 2^{-2023}.$$ ……………15 分

d) From c), we know that X is a geometric random variable with parameter $\frac{1}{2}$. Thus

……………………………20 分

$$\mathbb{E}(X) = \frac{1}{1/2} = 2.$$

II. **(10 points)** You ask your roommate to water a sickly plant while you are on vacation. Without water, it will die with probability 0.8; with water, it will die with probability 0.15. You are 90 percent certain that your roommate will remember to water the plant.

a. What is the probability that the plant will be alive when you return?

b. If the plant is dead upon your return, what is the probability that your roommate forgot to water it?

Solution.

a) $\mathbb{P}(\text{alive}) = \mathbb{P}(\text{alive} \mid \text{remember})\mathbb{P}(\text{remember}) + \mathbb{P}(\text{alive} \mid \text{forget})\mathbb{P}(\text{forget}) = (1 - 0.15) \times 0.9 + (1 - 0.8) \times 0.1 = 0.785.$ ……………………………5 分

b) $\mathbb{P}(\text{forget} \mid \text{die}) = \frac{\mathbb{P}(\text{forget,die})}{\mathbb{P}(\text{die})} = \frac{\mathbb{P}(\text{die} \mid \text{forget})\mathbb{P}(\text{forget})}{1-\mathbb{P}(\text{alive})} = \frac{0.8\times0.1}{1-0.785} = \frac{16}{43} \approx 0.372.$ ………10 分

III. (10 points) Balls numbered 1 through 2023 are in an urn. Suppose that n, $n \le 2023$, of them are randomly selected without replacement. Let Y denote the largest number selected.

a. Find the probability mass function of Y.

b. Calculate $\mathbb{E}(Y)$.

Solution.

a) If $n \le k \le 2023$, then

$$\mathbb{P}(Y=k) = \frac{\binom{k-1}{n-1}}{\binom{2023}{n}}$$

otherwise $\mathbb{P}(Y=k)=0$. ……………………………5 分

b) From a), we see that

$$1 = \sum_{k=n}^{2023} \mathbb{P}(Y=k) = \sum_{k=n}^{2023} \frac{\binom{k-1}{n-1}}{\binom{2023}{n}} \Rightarrow \sum_{k=n}^{2023} \binom{k-1}{n-1} = \binom{2023}{n}.$$

Therefore,

$$\mathbb{E}(Y) = \sum_{k=n}^{2023} k \cdot \mathbb{P}(Y=k) = \sum_{k=n}^{2023} \frac{k \cdot \binom{k-1}{n-1}}{\binom{2023}{n}} = \frac{n}{\binom{2023}{n}} \sum_{k=n}^{2023} \binom{k}{n}$$

$$\text{set } j=k+1 \quad = \frac{n}{\binom{2023}{n}} \sum_{j=n+1}^{2024} \binom{j-1}{(n+1)-1} = \frac{n}{\binom{2023}{n}} \binom{2024}{n+1}$$

$$= \frac{2024 \cdot n}{n+1}.$$

……………………………10 分

IV. (10 points) Let W have the exponential distribution with mean 1. Let $0 < w_1 < w_2 < \infty$. Set $g(w)=0$ for $0 \le w < w_1, g(w)=1$ for $w_1 \le w < w_2$, and $g(w)=2$ for $w \ge w_2$. If W can be used to construct a random variable $Y=g(W)$ such that $P(Y=0,1,2)=\frac{1}{3}$. Find w_1 and w_2

Solution.

The probability function of $Y=g(W)$ is given by

$$P(Y=0) = P(0 \le W < w_1) = 1-e^{-w_1},$$
$$P(Y=1) = P(w_1 \le W < w_2) = e^{-w_1} - e^{-w_2},$$
$$P(Y=2) = P(W \ge w_2) = e^{-w_2}.$$

……………………………5 分

Thus Y is uniformly distributed on $\{0,1,2\}$ provided that $1-e^{-w_1}=1/3$ and $e^{-w_2}=1/3$; that is, $w_1=\log(3/2)$ and $w_2=\log 3$. …… ……10 分

V. (10 points) Let U and V be independent random variables, each uniformly distributed on $[0,1]$. Determine the mean and variance of the random variable $Y=3U^2-2V$

Solution.

Now $E(U^2)=\int_0^1 u^2 du = 1/3$ and $E(U^4)=\int_0^1 u^4 du = 1/5$, so $\mathrm{var}(U^2) = E(U^4) - [E(U^2)]^2 = 1/5 - 1/9 = 4/45$. Also, $E(V)=1/2$ and $E(V^2)=1/3$, so $\mathrm{var}(V) = E(V^2) - [E(V)]^2 = 1/3 - 1/4 = 1/12$. ……………………………5 分

Consequently, $E(Y) = E(3U^2-2V) = 3E(U^2) - 2E(V) = 1-1 = 0$ and $\mathrm{var}(Y) = \mathrm{var}(3U^2-2V) = 9\mathrm{var}(U^2) + 4\mathrm{var}(V) = 9(4/45) + 4(1/12) = 4/5 + 1/3 = 17/15$.

……………………………10 分

VI. (10 points) The number of parking tickets issued in a certain city on any given weekday has a Poisson distribution with parameter $\mu = 50$. What is the approximate probability that between 30 and 70 tickets are given out on a particular day (Use the CLT)? ($\Phi(2.90) = 0.9981, \Phi(2.83) = 0.9977$)

Solution.

With $Y =$ # of tickets, Y has approximately a normal distribution with $\mu = 50$ and $\sigma = \sqrt{\mu} = 7.071$. ………………………………5 分

So, using a continuity correction from $[30,70]$ to $[29.5,70.5]$,

$$P(30 \le Y \le 70) \approx P\left(\frac{29.5-50}{7.071} \le Z \le \frac{70.5-50}{7.071}\right) = P(-2.90 \le Z \le 2.90)$$
$$= 2\Phi(2.90) - 1 = 0.9962.$$ ………………………………10 分

VII. (10 points) Consider a random sample $X_1, X_2, \dots, X_n$ from the pdf

$$f(x;k) = \begin{cases} kx^{k-1} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

where $k > 0$.

a. Use the method of moments to obtain an estimator of k.

b. Obtain the maximum likelihood estimator of k.

Solution.

a) $E(X) = \int_{-\infty}^{+\infty} xf(x;k)\,dx = \int_0^1 kx^k\,dx = \frac{k}{k+1}$.

Equating $E(X)$ to $\bar{X}$ gives $\frac{k}{k+1} = \bar{X}$. Solving it we get the moment estimator

$$\hat{k} = \frac{\bar{X}}{1-\bar{X}}.$$ ………………………………5 分

b) The likelihood function is

$L(k) = k^n x_1^{k-1} x_2^{k-1} \cdots x_n^{k-1}, \quad 0 \le x_i \le 1 (i = 1, \dots, n)$.

$\ln L(k) = n\ln k + (k-1)\sum \ln x_i$.

Let $\frac{\mathrm{d}L(k)}{\mathrm{d}k} = \frac{n}{k} + \sum \ln x_i = 0$ to obtain $\hat{k} = -\frac{n}{\sum \ln x_i}$.

And thus the mle is

$$\hat{k} = -\frac{n}{\sum \ln X_i}.$$ ………………………………10 分

VIII. (10 points) Suppose the weight distribution of packages sent in a certain manner is normal. A random sample of $n = 9$ packages yielded a sample mean weight of 10 and a sample standard deviation of 4.5.

a. Assume the true population standard deviation is unknown. Please compute a 95% CI for the true average weight.

b. Assume the true average weight is unknown. Please compute a 95% CI for the true population variance.

c. Assume the true population standard deviation is known, which happens to be 4.5. What sample size is necessary to ensure that the resulting 95% CI for the true average weight has a width of (at most) 2?

Solution.

a) A 95% CI for the true average weight is $\bar{x} \pm t_{0.025,8}\frac{4.5}{\sqrt{9}} = 10 \pm 2.306 \times \frac{4.5}{3} =$ 10 ± 3.459. ………………………………4 分

b) A 95% CI for the true population variance, σ^2, is given by

$\left(\frac{(n-1)s^2}{\chi^2_{0.025,8}}, \frac{(n-1)s^2}{\chi^2_{0.975,8}}\right) = \left(\frac{8\times4.5^2}{17.534}, \frac{8\times4.5^2}{2.18}\right) = (9.2392,\ 74.3119)$. ………………………………7 分

c) A 95% CI has a width $2z_{0.025}\frac{4.5}{\sqrt{n}} = 9 \times 1.96 \times \frac{1}{\sqrt{n}} = \frac{17.64}{\sqrt{n}}$.

Solve $\frac{17.64}{\sqrt{n}} \le 2$ to obtain $n \ge \left(\frac{17.64}{2}\right)^2 = 77.7924$.

Since n must be an integer, n is at least 78. ………………………………10 分

IX. (10 points) Consider the following sample of diameters of $n = 9$ randomly selected ball bearings (滚珠)：

$$0.65, 0.55, 0.62, 0.66, 0.54, 0.58, 0.6, 0.5, 0.7.$$

Suppose the ball bearing diameter is normally distributed with $\sigma = 0.24$

a. Does the data indicate conclusively that the true average diameter bigger than 0.5 at a significance level $\alpha = 0.05$?

b. If the true average diameter is $\mu = 0.6316$ and a level $\alpha = 0.05$ test based on $n = 9$ is used, what is the probability of making a type II error?

Solution.

a) Null hypothesis: $H_0: \mu = 0.5$. Alternative hypothesis: H_α: $\mu > 0.5$.
The testing statistic value $z = \frac{\bar{x}-0.5}{0.24/3}$. Rejection region: $z \geq z_{0.05} = 1.645$.
Substituting $\bar{x} = 0.6$, $z = 1.25$. The computed value does not fall in the rejection region. So we cannot reject H_0, which means that the sample data does not suggest that the true average diameter is bigger than 0.5 at a significance level $\alpha = 0.05$. ……………………………5 分

b) $\beta(0.6316) = \Phi\left(z_{0.05} + \frac{0.5-0.6316}{0.24/3}\right) = \Phi(0) = 0.5.$ ……………………………10 分