考试科目： 概率论与数理统计 开课单位： 数学系

考试时长： 2023年2月18日14:00-16:00 命题教师： 概率统计教学组

题号	1	2	3	4	5	6	7	8
分值	20 分	20 分	10 分	10 分	10 分	10 分	10 分	10 分

本试卷共三大部分，满分（100）分

（考试结束后请将试卷、答题卡一起交给监考老师）

(附注: 下分位数)

$u_{0.95} = 1.645, u_{0.975} = 1.96, u_{0.99} = 2.33, u_{0.995} = 2.58, t_{0.95}(18) = 1.73, t_{0.975}(18) = 2.10$

$t_{0.95}(20) = 1.72, t_{0.99}(8) = 2.90, t_{0.99}(9) = 2.82, t_{0.995}(8) = 3.36, t_{0.995}(9) = 3.25, t_{0.975}(20) = 2.09$

$t_{0.95}(35) = 1.690, t_{0.975}(35) = 2.03, t_{0.95}(36) = 1.689, t_{0.975}(36) = 2.028$

$F_{0.95}(9, 9) = 3.18, F_{0.95}(10, 9) = 3.14, F_{0.95}(10, 10) = 2.98, F_{0.975}(9, 9) = 4.03, F_{0.975}(10, 9) = 3.96$

$F_{0.975}(10, 10) = 3.72, \chi^2_{0.025}(8) = 2.18, \chi^2_{0.05}(8) = 2.73, \chi^2_{0.95}(8) = 15.51, \chi^2_{0.975}(8) = 17.53$

$\chi^2_{0.025}(9) = 2.70, \chi^2_{0.05}(9) = 3.33, \chi^2_{0.95}(9) = 16.92, \chi^2_{0.975}(9) = 19.02, \chi^2_{0.005}(25) = 10.52$

$\chi^2_{0.01}(25) = 15.52, \chi^2_{0.995}(25) = 46.93, \chi^2_{0.99}(25) = 44.31, \chi^2_{0.005}(26) = 11.16, \chi^2_{0.01}(26) = 12.20$

$\chi^2_{0.995}(26) = 48.29, \chi^2_{0.99}(26) = 45.64, \chi^2_{0.05}(25) = 14.61, \chi^2_{0.1}(25) = 16.47, \chi^2_{0.95}(25) = 37.65$

$\chi^2_{0.9}(25) = 34.38, \chi^2_{0.05}(26) = 15.38, \chi^2_{0.1}(26) = 17.29, \chi^2_{0.95}(26) = 38.89, \chi^2_{0.9}(26) = 35.56$

第一部分 选择题（每题4 分，总共20 分）

Part One – Single Choice (4 marks each question, 20 marks in total)

1. 假设A和B是两个随机事件，且$P(A \mid B) = 0.6, P(B \mid A) = 0.5$, $P(A \cup B) = 0.8$。则$P(A)$的值为（　　）

   Suppose A and B are two events with $P(A \mid B) = 0.6, P(B \mid A) = 0.5$ and $P(A \cup B) = 0.8$. Find the value of $P(A)$.

   A. 0.6　　B. 0.5　　C. 0.4　　D. 0.3

2. 假设$Z \sim N(0,1)$. 则$E(1+Z+Z^2+Z^3)$ 的值为（　　）

   Suppose $Z \sim N(0,1)$. Find $E(1+Z+Z^2+Z^3)$.

   A. 0　　B. 1　　C. 2　　D. 3

3. 设随机变量$X_1, X_2, ..., X_n$互相独立，则根据辛钦大数定律，当$X_1, X_2, ..., X_n$满足以下哪个条件时，$\frac{1}{n}\sum_{i=1}^{n} X_i$ 会依概率收敛于其共同的数学期望.

   A. 有相同的数学期望　　B. 服从同一离散分布
   C. 服从同一泊松分布　　D. 服从同一连续分布

   Suppose the random variables $X_1, X_2, ..., X_n$ are independent. According to the law of Khinchin's large number, $\frac{1}{n}\sum_{i=1}^{n} X_i$ converges to its common mathematical expectation in probability if $X_1, X_2, ..., X_n$ satisfy the following condition (　　).

   A. They have the same expectation
   B. They follow the same discrete distribution
   C. Then follow the same Poisson distribution
   D. They follow the same discrete distribution

4. 设随机变量X, Y均服从标准正态分布，则以下哪个是正确的？

   A. $X+Y \sim N(0,2)$　　B. $X^2+Y^2 \sim \chi^2(2)$
   C. $\frac{X^2}{Y^2} \sim F(1,1)$　　D. X^2和Y^2均服从$\chi^2(1)$分布

   Let the random variables X and Y follow the standard normal distribution, then which of the following statement is true?

   A. $X+Y \sim N(0,2)$　　B. $X^2+Y^2 \sim \chi^2(2)$
   C. $\frac{X^2}{Y^2} \sim F(1,1)$　　D. Both X^2 and Y^2 follow to the distribution $\chi^2(1)$

5. 已知$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, 为检验总体$X$的方差是否大于$Y$的方差，则应作假设（　　）.

   Suppose $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, which of the follow assumptions should be made in order to test whether the variance of X is greater than the variance of Y?

   A. $H_0 : \sigma_1^2 > \sigma_2^2;\ H_1 : \sigma_1^2 \leq \sigma_2^2$　　B. $H_0 : \sigma_1^2 \geq \sigma_2^2;\ H_1 : \sigma_1^2 < \sigma_2^2$
   C. $H_0 : \sigma_1^2 < \sigma_2^2;\ H_1 : \sigma_1^2 \geq \sigma_2^2$　　D. $H_0 : \sigma_1^2 \leq \sigma_2^2;\ H_1 : \sigma_1^2 > \sigma_2^2$

第二部分　填空题（每空2 分，总共20 分）

Part Two – Blank Filling (2 marks each blank, 20 marks in total)

1. 掷三粒骰子，则出现的三个点数中最小的点数为3的概率是__________.

   Throw three dice, what is the probability that the smallest of the three points is 3?

2. 已知随机变量$X \sim N(2, \sigma^2)$, 且有$P(2 < X < 4) = 0.3$, 则$P(X < 0) =$ __________.

   Suppose $X \sim N(2, \sigma^2)$ and $P(2 < X < 4) = 0.3$, then $P(X < 0) =$ __________.

3. 已知随机变量X的密度函数为$f(x) = Ae^{-x^2+2x-1}, -\infty < x < \infty$,其中$A$为常数,则$E[X^2]$的具体的值= __________.

   Suppose X's density function is $f(x) = Ae^{-x^2+2x-1}, -\infty < x < \infty$, where A is a constant, then the specific value of $E[X^2] =$ __________.

4. 设随机变量X的频率函数为$P\{X = -2\} = \dfrac{1}{2}$, $P\{X = 1\} = a$, $P\{X = 3\} = b$, 且$E[X] = 0$, 则$D[X] =$ __________.

   Suppose X's frequency function is $P\{X = -2\} = \dfrac{1}{2}$, $P\{X = 1\} = a$, $P\{X = 3\} = b$, and $E[X] = 0$, then $D[X] =$ __________.

5. 设随机变量(X, Y)在单位圆$D : x^2 + y^2 \le 1$内服从均匀分布，则X和Y的相关系数$\rho_{XY} =$ __________.

   Suppose $(X, Y) \sim U(D)$ with $D : x^2 + y^2 \le 1$, then the correlation coefficient $\rho_{XY} =$ __________.

6. 设(X, Y)服从分布$N(3, 1, 1^2, 1^2, 0)$, 令$U = X + Y, V = X - Y$, 则(U, V)服从分布__________.

   Let (X, Y) have $N(3, 1, 1^2, 1^2, 0)$ distribution. If $U = X + Y, V = X - Y$, then (U, V) is distributed as __________.

7. 设$X_1, X_2, \cdots, X_n, \cdots$是相互独立的随机变量序列，且服从参数为$\lambda = 2$的泊松分布，则$\dfrac{1}{n}\sum_{i=1}^{n} X_i^2$依概率收敛于__________.

   Let $X_1, X_2, \cdots, X_n, \cdots$ be a sequence of independent random variables and have a Poisson distribution with parameter λ. Then $\dfrac{1}{n}\sum_{i=1}^{n} X_i^2$ converge in probability to __________.

8. 设$X_1, X_2, \cdots, X_n$为二项分布总体$X \sim b(6, p)$的一个样本, 若$\bar{X} - kS^2$为参数$6p^2$的一个无偏估计，则$k =$ __________.

   Let $X_1, X_2, \cdots, X_n$ be a random sample from a population $X \sim b(m, p)$. If $\bar{X} - kS^2$ is an unbiased estimator of the parameter mp^2, then $k =$ __________.

9. 设$X_1, X_2, \cdots, X_n$ 为正态总体$N(\mu, \sigma^2)$ 的一个样本，则$\frac{n(\bar{X}-\mu)^2}{\sigma^2} + \frac{(n-1)S^2}{\sigma^2}$ 服从分布_________.

   Let $X_1, X_2, \cdots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ population, then $\frac{n(\bar{X}-\mu)^2}{\sigma^2} + \frac{(n-1)S^2}{\sigma^2}$ is distributed as _________.

10. 设$X_1, X_2, \cdots, X_n$ 是来自均匀分布总体$X \sim U[\theta, 3\theta+4]$ 的样本，则未知参数θ 的矩估计量$\hat{\theta} =$ _________.

    Let $X_1, X_2, \cdots, X_n$ be a random sample from a uniform distribute population $X \sim U[\theta, 3\theta+4]$, the method of moments estimate of unknown parameter θ is $\hat{\theta} =$ _________.

第三部分 解答题（每题10分，总共60 分）

Part Three–Question Answering (10 marks each question, 60 marks in total)

1. 设(X, Y)为连续型随机变量，其联合密度函数为

$$f(x) = \begin{cases} \dfrac{x}{5} + Cy, & 0 < x < 1, 1 < y < 5, \\ 0, & \text{其他}. \end{cases}$$

(1) 求常数C的值;

(2) X和Y是否独立？说明理由;

(3) 求$P\{X + Y > 3\}$.

Suppose the joint density function of continuous random variable (X, Y) is

$$f(x) = \begin{cases} \dfrac{x}{5} + Cy, & 0 < x < 1, 1 < y < 5, \\ 0, & \text{otherwise}. \end{cases}$$

(1) Find the constant C;

(2) Are X and Y independent? Explain your reason;

(3) Find $P\{X + Y > 3\}$.

2. 设$X_i (i = 1, 2, 3)$是互相独立的泊松随机变量，均值分别为λ_i，$i = 1, 2, 3$. 令$X = X_1 + X_2$, $Y = X_2 + X_3$, 称(X, Y)为二元泊松随机变量.

(1) 求$E[X]$和$D[Y]$;

(2) 求Cov(X, Y);

(3) 求(X,Y)的联合频率函数$P\{X=i,Y=j\}$.

Suppose $X_i \sim P(\lambda_i)$ with $i=1,2,3$, and X_1,X_2,X_3 are independent. Let $X=X_1+X_2$, $Y=X_2+X_3$.

(1) Find $E[X]$ and $D[Y]$;

(2) Find Cov(X, Y);

(3) Find the joint frequency function $P\{X=i,Y=j\}$ of (X,Y).

3. 设二维随机变量(X,Y)在区域$D:=\{(x,y)|0<x<1,x^2<y<\sqrt{x}\}$上服从均匀分布，令$U=\begin{cases}1, & X\le Y,\\ 0, & X>Y.\end{cases}$

(1) 写出(X,Y)的联合密度函数;

(2) X和Y是否相互独立? 给出理由;

(3) 求$Z=U+X$的分布函数$F(z)$.

Suppose $(X,Y)\sim U(D)$ with $D:=\{(x,y)|0<x<1,x^2<y<\sqrt{x}\}$. Let $U=\begin{cases}1, & X\le Y,\\ 0, & X>Y.\end{cases}$

(1) Find the joint density function $f(x,y)$ of (X,Y);

(2) Are X and Y independent? Explain your reason;

(3) Find the distribution function $F_Z(z)$ of $Z=U+X$.

4. 已知总体X 的密度函数为$f(x)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1,\\ 0, & 其他,\end{cases}$ 其中$\theta>0$ 为未知参数. 设$X_1,X_2,\cdots,X_n$ 为总体的一个样本, $x_1,x_2,\cdots,x_n$ 为一相应的样本值.

(1) 求未知参数θ 的矩估计量和矩估计值;

(2) 求未知参数θ 的最大似然估计量和最大似然估计值.

Suppose the probability density function of a population X is $f(x) = \begin{cases} \sqrt{\theta}x^{\sqrt{\theta}-1}, & 0 \leqslant x \leqslant 1, \\ 0, & \text{otherwise,} \end{cases}$ where $\theta > 0$ is unknown parameter. Suppose $X_1, X_2, \cdots, X_n$ are samples, and $x_1, x_2, \cdots, x_n$ are the observed values.

(1) Find the Moment Estimation and the Moment Estimation value of the unknown parameter θ;

(2) Find the Maximum Likelihood Estimation and the Maximum Likelihood Estimation value of the unknown parameter θ.

5. 假设人体身高服从正态分布, 今抽测甲、乙两地区18 岁$\sim$ 25 岁女青年身高的数据如下: 甲地区抽取10 名, 样本均值1.64 m, 样本标准差0.2 m; 乙地区抽取10名, 样本均值1.62 m, 样本标准差0.4 m.

(1) 求两正态总体方差比的置信水平为95% 的置信区间;

(2) 假设甲地区和乙地区女青年身高的方差相等且未知,求两正态总体均值差的置信水平为95% 的置信区间.

Assume that the human body height follows the normal distribution, the height data of young women aged 18 $\sim$ 25 in area A and area B are as follows: 10 women were selected from area A, the sample mean is 1.64 m, and the sample standard deviation is 0.2 m; Ten samples were selected from region B, the sample mean is 1.62 m, and the sample standard deviation is 0.4 m.

(1) Find a 95% confidence interval for the variance ratio of two normal populations?(The result is accurate to two decimal places).

(2) Find a 95% confidence interval the difference between the two normal population means, under the assumption the height variances of area A and area B are equal(The result is accurate to two decimal places).

6. 设某次考试的考生成绩服从正态分布, 从中随机的抽取36 位考生的成绩, 算得平均成绩66.5 分, 标准差为15 分. 问在显著性水平0.05 下, 是否可以认为这次考试全体考生的平均成绩为70 分? 给出检验过程.

Assume that the scores of students in a certain exam follows the normal distribution, and the scores of 36 students are randomly selected, where the average score is 66.5, the standard deviation is 15. With the significance level of $\alpha = 0.05$, can we assume that the average score of all candidates in this test is 70? Give the test procedure.

答案

一、选择题

A C C D D

二、填空题

1. $\dfrac{37}{216} \approx 0.17$

2. 0.2

3. $\dfrac{3}{2}$

4. $\dfrac{9}{2}$

5. 0

6. $N(4, 2, (\sqrt{2})^2, (\sqrt{2})^2, 0)$

7. 6

8. 6

9. $\chi^2(n)$

10. $\dfrac{\overline{X}}{2} - 1$

1. 设(X,Y)为连续型随机变量，其联合密度函数为

$$f(x)=\begin{cases}\dfrac{x}{5}+Cy, & 0<x<1, 1<y<5,\\ 0, & \text{其他}.\end{cases}$$

(1) 求常数C的值;

(2) X和Y是否独立?

(3) 求$P\{X+Y>3\}$.

解：(1)

$$\begin{aligned}1&=\int_0^1\int_1^5\Big(\frac{x}{5}+Cy\Big)dydx\\&=12C+\frac{2}{5}\end{aligned}$$

因此

$$C=\frac{1}{20}$$

(2) 因为联合密度函数不可分解，因此X和Y不独立；

或者求出$f_X(x)$, $f_Y(y)$, 然后说明$f(x,y)\neq f_X(x)f_Y(y)$.

(3)

$$\begin{aligned}P\{X+Y>3\}&=\int_0^1\int_{3-x}^5\Big(\frac{x}{5}+\frac{y}{20}\Big)dydx\\&=\frac{11}{15}\end{aligned}$$

2. 设$X_i(i=1,2,3)$是互相独立的泊松随机变量，均值分别为λ_i, $i=1,2,3$. 令$X=X_1+X_2$, $Y=X_2+X_3$, 称(X,Y)为二元破泊松随机变量.

(1) 求$E[X]$和$D[Y]$;

(2) 求Cov(X, Y);

(3) 求(X,Y)的联合频率函数$P\{X=i,Y=j\}$.

解：(1)

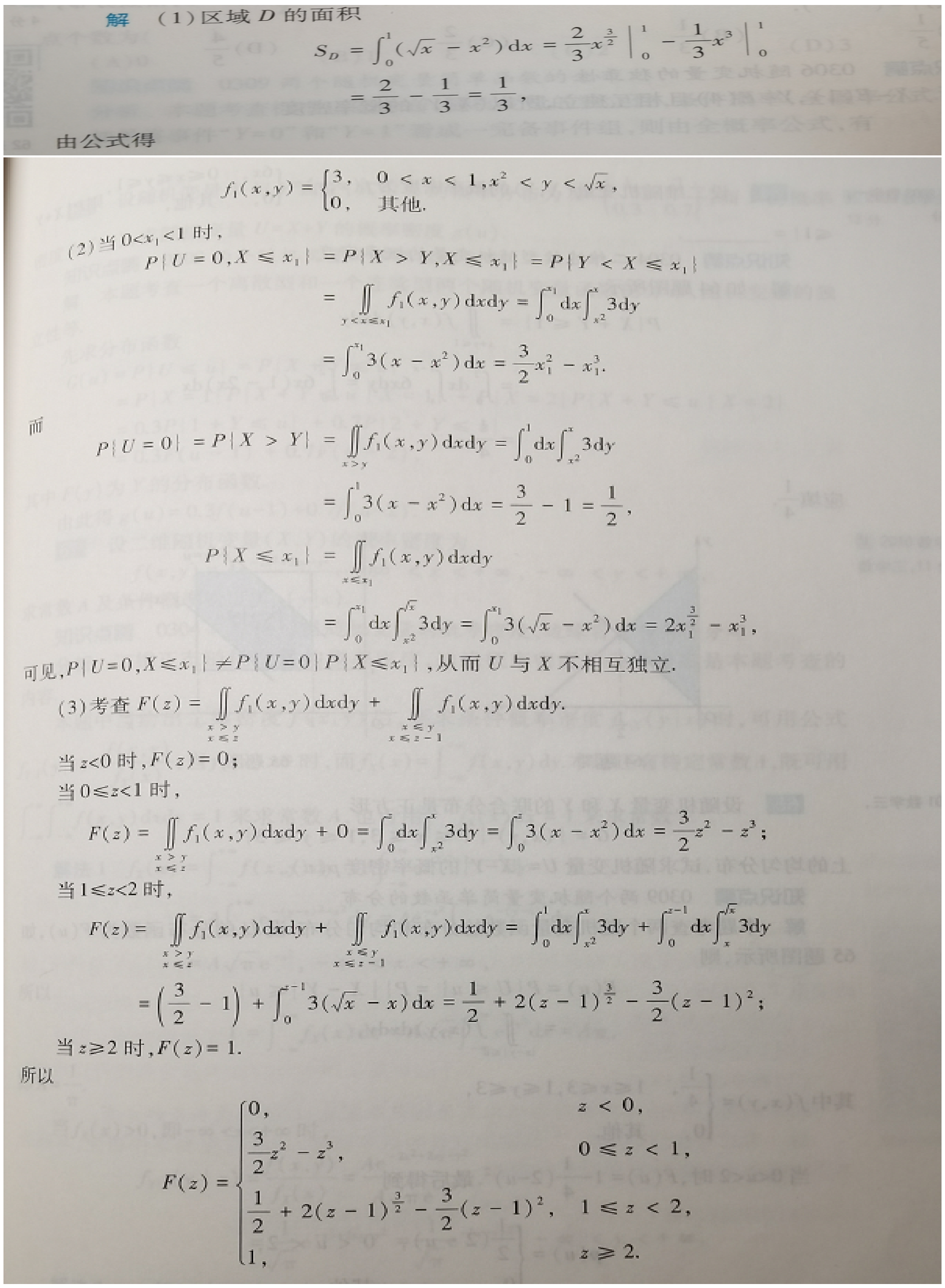
$$E[X]=\lambda_1+\lambda_2$$
$$D[Y]=\lambda_2+\lambda_3$$

(b) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(X_1+X_2,X_2+X_3)=\mathrm{Cov}(X_1,X_2+X_3)+\mathrm{Cov}(X_2,X_2+X_3)$
$=\mathrm{Cov}(X_2,X_2)=\mathrm{Var}(X_2)=\lambda_2$

(c) 利用条件期望的性质

$$\begin{aligned}
P\{X=i,Y=j\} &= \sum_k P\{X=i,Y=j \mid X_2=k\}P\{X_2=k\} \\
&= \sum_k P\{X_1=i-k,X_3=j-k \mid X_2=k\}\mathrm{e}^{-\lambda_2}\lambda_2^k/k! \\
&= \sum_k P\{X_1=i-k,X_3=j-k\}\mathrm{e}^{-\lambda_2}\lambda_2^k/k! \\
&= \sum_k P\{X_1=i-k\}P\{X_3=j-k\}\mathrm{e}^{-\lambda_2}\lambda_2^k/k! \\
&= \sum_{k=0}^{\min(i,j)} \mathrm{e}^{-\lambda_1}\frac{\lambda_1^{i-k}}{(i-k)!}\mathrm{e}^{-\lambda_3}\frac{\lambda_3^{j-k}}{(j-k)!}\mathrm{e}^{-\lambda_2}\frac{\lambda_2^k}{k!}
\end{aligned}$$

3. 设二维随机变量(X,Y)在区域$D := \{(x,y)|0 < x < 1, x^2 < y < \sqrt{x}\}$上服从均匀分布，令$U = \begin{cases} 1, & X \le Y, \\ 0, & X > Y. \end{cases}$ (1) 写出(X,Y)的联合密度函数; (2) X和Y是否相互独立? 给出理由; (3) 求$Z = U + X$的分布函数$F(z)$.



解 (1)区域 D 的面积

$$S_D = \int_0^1 (\sqrt{x} - x^2)\,\mathrm{d}x = \frac{2}{3}x^{\frac{3}{2}}\Big|_0^1 - \frac{1}{3}x^3\Big|_0^1 = \frac{2}{3} - \frac{1}{3} = \frac{1}{3},$$

由公式得

$$f_1(x,y) = \begin{cases} 3, & 0 < x < 1, x^2 < y < \sqrt{x}, \\ 0, & \text{其他}. \end{cases}$$

(2)当 $0<x_1<1$ 时,

$$P\{U=0, X \le x_1\} = P\{X > Y, X \le x_1\} = P\{Y < X \le x_1\} = \iint\limits_{y<x\le x_1} f_1(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^{x_1}\mathrm{d}x\int_{x^2}^{x} 3\mathrm{d}y = \int_0^{x_1} 3(x - x^2)\,\mathrm{d}x = \frac{3}{2}x_1^2 - x_1^3.$$

而

$$P\{U=0\} = P\{X > Y\} = \iint\limits_{x>y} f_1(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^1 \mathrm{d}x\int_{x^2}^{x} 3\mathrm{d}y = \int_0^1 3(x - x^2)\,\mathrm{d}x = \frac{3}{2} - 1 = \frac{1}{2},$$

$$P\{X \le x_1\} = \iint\limits_{x\le x_1} f_1(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^{x_1}\mathrm{d}x\int_{x^2}^{\sqrt{x}} 3\mathrm{d}y = \int_0^{x_1} 3(\sqrt{x} - x^2)\,\mathrm{d}x = 2x_1^{\frac{3}{2}} - x_1^3,$$

可见, $P\{U=0, X\le x_1\} \ne P\{U=0\}P\{X \le x_1\}$, 从而 U 与 X 不相互独立.

(3)考查 $F(z) = \iint\limits_{\substack{x>y \\ x\le z}} f_1(x,y)\,\mathrm{d}x\mathrm{d}y + \iint\limits_{\substack{x\le y \\ x \le z-1}} f_1(x,y)\,\mathrm{d}x\mathrm{d}y$.

当 $z<0$ 时, $F(z)=0$;

当 $0\le z<1$ 时,

$$F(z) = \iint\limits_{\substack{x>y \\ x\le z}} f_1(x,y)\,\mathrm{d}x\mathrm{d}y + 0 = \int_0^z \mathrm{d}x\int_{x^2}^{x} 3\mathrm{d}y = \int_0^z 3(x - x^2)\,\mathrm{d}x = \frac{3}{2}z^2 - z^3;$$

当 $1\le z<2$ 时,

$$F(z) = \iint\limits_{\substack{x>y \\ x\le z}} f_1(x,y)\,\mathrm{d}x\mathrm{d}y + \iint\limits_{\substack{x\le y \\ x \le z-1}} f_1(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^1 \mathrm{d}x\int_{x^2}^{x} 3\mathrm{d}y + \int_0^{z-1}\mathrm{d}x\int_x^{\sqrt{x}} 3\mathrm{d}y$$

$$= \left(\frac{3}{2} - 1\right) + \int_0^{z-1} 3(\sqrt{x} - x)\,\mathrm{d}x = \frac{1}{2} + 2(z-1)^{\frac{3}{2}} - \frac{3}{2}(z-1)^2;$$

当 $z\ge 2$ 时, $F(z)=1$.

所以

$$F(z) = \begin{cases} 0, & z < 0, \\ \frac{3}{2}z^2 - z^3, & 0 \le z < 1, \\ \frac{1}{2} + 2(z-1)^{\frac{3}{2}} - \frac{3}{2}(z-1)^2, & 1 \le z < 2, \\ 1, & z \ge 2. \end{cases}$$

4. 已知总体X 的密度函数为$f(x)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\leqslant x\leqslant 1,\\ 0, & \text{其他},\end{cases}$ 其中$\theta>0$ 为未知参数. 设$X_1,X_2,\cdots,X_n$ 为总体的一个样本, $x_1,x_2,\cdots,x_n$ 为一相应的样本值.

(1) 求未知参数θ 的矩估计量和矩估计值;

(2) 求未知参数θ 的最大似然估计量和最大似然估计值.

(1) $\mu_1=\int_0^1 x\sqrt{\theta}x^{\sqrt{\theta}-1}\,\mathrm{d}x=\int_0^1\sqrt{\theta}x^{\sqrt{\theta}}\mathrm{d}x=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$.
由此得

$$\theta=\left(\frac{\mu_1}{1-\mu_1}\right)^2.$$

在上式中以$\bar{X}$ 代替μ_1, 得θ 的矩估计量和矩估计值分别为

$$\hat{\theta}=\left(\frac{\bar{X}}{1-\bar{X}}\right)^2,\quad \hat{\theta}=\left(\frac{\bar{x}}{1-\bar{x}}\right)^2.$$

(2) 设$x_1,x_2,\cdots,x_n$ 是一个样本值. 似然函数为

$$L=\prod_{i=1}^n\left(\sqrt{\theta}x_i^{\sqrt{\theta}-1}\right)=\theta^{n/2}\left(\prod_{i=1}^n x_i\right)^{\sqrt{\theta}-1},$$

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

令

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L=\frac{n}{2\theta}+\frac{1}{2\sqrt{\theta}}\sum_{i=1}^n\ln x_i=0,$$

得θ 的最大似然估计值为

$$\hat{\theta}=\frac{n^2}{\left(\sum\limits_{i=1}^n\ln x_i\right)^2},$$

θ 的最大似然估计量为

$$\hat{\theta}=\frac{n^2}{\left(\sum\limits_{i=1}^n\ln X_i\right)^2}.$$

5. 假设人体身高服从正态分布, 今抽测甲、乙两地区18 岁～ 25 岁女青年身高的数据如下: 甲地区抽取10 名, 样本均值1.64 m, 样本标准差0.2 m; 乙地区抽取10名, 样本均值1.62 m, 样本标准差0.4 m.

(1) 求两正态总体方差比的置信水平为95% 的置信区间;

(2) 假设甲地区和乙地区女青年身高的方差相等且未知,求两正态总体均值差的置信水平为95% 的置信区间.

解 设$x_1, \cdots, x_{10}$ 为甲地区抽取的女青年身高, $y_1, \cdots, y_{10}$ 为乙地区抽取的女青年身高, 由题设条件, $\bar{x} = 1.64, s_x = 0.2$, $\bar{y} = 1.62, s_y = 0.4$

(1) $\dfrac{\sigma^2_{甲}}{\sigma^2_{乙}}$ 的$1-\alpha$ 的置信区间为

$$\left[\frac{s_x^2}{s_y^2} \cdot \frac{1}{F_{1-\alpha/2}(m-1, n-1)}, \frac{s_x^2}{s_y^2} \cdot \frac{1}{F_{\alpha/2}(m-1, n-1)}\right].$$

此处$\alpha = 0.05, m = n = 10$, 查表得$F_{0.975}(9, 9) = 4.03$, $F_{0.025} = \dfrac{1}{F_{0.975}(9, 9)} = \dfrac{1}{4.03}$, 由此, $\dfrac{\sigma^2_{甲}}{\sigma^2_{乙}}$ 的置信水平为95% 的置信区间为

$$\left[\frac{0.2^2}{0.4^2} \cdot \frac{1}{4.03}, \frac{0.2^2}{0.4^2} \cdot 4.03\right] = [0.0620, 1.0075] \approx [0.06, 1.01].$$

(2)两个正态总体的方差相等, 则

$$s_\omega^2 = \frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2} = \frac{9 \times 0.2^2 + 9 \times 0.4^2}{10+10-2} = \frac{1.8}{18} = 0.1$$

$\mu_x - \mu_y$ 的置信水平为$1-\alpha$的置信区间为

$$\left[(\bar{X} - \bar{Y}) - t_{\frac{\alpha}{2}}(m+n-2)S_\omega\sqrt{\frac{1}{m} + \frac{1}{n}}, (\bar{X}_x - \bar{X}_y) + t_{\frac{\alpha}{2}}(m+n-2)S_\omega\sqrt{\frac{1}{m} + \frac{1}{n}}\right].$$

此处$\alpha = 0.05, m = n = 10$, 查表得$t_{0.975}(18) = 2.10$, 由此, 代入数据得$\mu_x - \mu_y$ 的置信水平为95% 的置信区间为

$$\left[(1.64 - 1.62) - 2.10 \cdot \sqrt{0.02}, (1.64 - 1.62) + 2.10 \cdot \sqrt{0.02}\right] \approx [-0.27, 0.31].$$

6. 设某次考试的考生成绩服从正态分布, 从中随机的抽取36 位考生的成绩, 算得平均成绩66.5 分, 标准差为15 分. 问在显著性水平0.05 下, 是否可以认为这次考试全体考生的平均成绩为70 分? 给出检验过程.

解 设该次考试的考生成绩为X, 则$X \sim N(\mu, \sigma^2)$, 把从X 中抽取的容量为n 的样本?值记为$\bar{X}$, 样本标准差记为S, 本题是在显著性水平$\alpha = 0.05$ 下检验假设

$$H_0 : \mu = 70; H_1 : \mu \neq 70,$$

拒绝域为

$$|t| = \frac{|\bar{x} - 70|}{S}\sqrt{n} \geqslant t_{1-\frac{\alpha}{2}}(n-1).$$

由$n = 36, \bar{x} = 66.5, S = 15, t_{0.975}(36-1) = 2.0301$, 算得

$$|t| = \frac{|66.5 - 70|}{15}\sqrt{36} = 1.4 < 2.0301,$$

可以接受假设$H_0 : \mu = 70$, 即在显著性水平0.05 下, 可以认为这次考试全体考生的平均成绩为70 分.