# 浙江大学 2012–2013 学年秋冬学期

## 《概率论与数理统计》期末考试试卷

课程号：061B9090/061Q0029，开课学院：数学系，任课教师：________________

考试试卷：A 卷 √、B 卷（请在选定项上打 √）

考试形式：闭 √、开卷（请在选定项上打 √），允许带计算器入场

考试日期：2013 年 1 月 20 日，考试时间 120 分钟

**诚信考试，沉着应考，杜绝违纪。**

**请注意：本试卷共六大题，四页，两大张。**

**请勿将试卷拆开或撕页！如发生此情况责任自负！**

**考生姓名：______________ 学号：______________ 所属院系：______________**

| 题序 | 一 | 二 | 三 | 四 | 五 | 六 | 总分 |
|---|---|---|---|---|---|---|---|
| 得分 | | | | | | | |
| 评卷人 | | | | | | | |

注：$\Phi(1)=0.84,\ \Phi(1.645)=0.95,\ \Phi(1.96)=0.975, \Phi(2.31)=0.99$，

$t_{0.05}(8)=1.86,\ t_{0.025}(8)=2.31,\ t_{0.05}(16)=1.75,\ t_{0.025}(16)=2.12,$

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

$c^2_{0.05}(4)=9.49,\ c^2_{0.05}(3)=7.82, F_{0.025}(9,7)=4.82, F_{0.025}(7,9)=4.2.$

**一．填空题（每小格 3 分，共 39 分。每个分布要求写出参数）：**

1. 设事件 $A,B,C$ 相互独立，已知 $P(A)=0.5, P(A\cup B)=0.6, P(\bar{C}\mid A)=0.4$, 则 P(B)= ____0.2____， $P(A\cup B\cup C)=$ ____0.84____.

2. 在区间 $(0,q)$ (参数 $q>0$)内独立重复观测 5 次，记为 $X_1, \mathbf{L}, X_5, X_i \sim U(0,q)$，(1) 设 $q=2$，则最大观测值小于 1.8 且最小观测值大于 0.4 的概率为____0.16807____; (2) 设 $q>0$ 未知，5 次观测值为 1.18,0.48,1.59,0.13,1.76，则 $q$ 的矩估计值是____2.056____.

3.某超市从开门到第 1 位顾客进入所需时间 $X$(分钟)的概率密度函数 $f(x)=\begin{cases}\frac{1}{5}e^{-x/5}, x>0\\ 0, \quad 其他.\end{cases}$

则超市开门后的 10 分钟内至少有 1 人进入的概率为____$1-e^{-2}$____; 从开门到第 1 位顾客进入平均花____5____分钟.

4．设某地区男性成年人的身高 $X$（厘米）与体重 $Y$（公斤）服从二元正态分布，$X\sim N(169.5,10.5^2),Y\sim N(57.3,16.2^2),\ r_{XY}=0.6$，从该地区独立随机选 $n$ 名男子，测得身高体重为 $(X_1,Y_1)\mathbf{L}(X_n,Y_n)$，记 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\bar{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$。则 $\bar{X}$ 服从 $\underline{N(169.5,\frac{10.5^2}{n})}$ 分布，$Cov(\bar{X},\bar{Y})=$ $\underline{102.06/n}$，当 $n\to\infty$ 时，$\frac{1}{n}\sum_{i=1}^{n}\frac{(X_i-169.5)(Y_i-57.3)}{10.5\times16.2}\xrightarrow{P}$ $\underline{\quad 0.6\quad}$.

5. 设总体 $X\sim N(m,s^2),\ m,s^2$ 均未知，$X_1,\mathbf{L},X_9$ 为来自 $X$ 的简单随机样本，$\bar{X}$ 和 $S$ 分别是样本均值和样本标准差，(1) 若根据样本观测值，$\bar{x}=7.076,s=1.2$，则 $s^2$ 的置信度为 0.95 的双侧置信区间为 $\underline{\quad(0.657\ ,\ 5.284)\quad}$，检验假设 $H_0:m=8,H_1:m\neq8$ 的 P_值为 $\underline{\quad 0.05\quad}$；(2) 设 $X_{10}$ 是从总体中独立抽取的另一次观测，则 $\frac{3(X_{10}-\bar{X})}{\sqrt{10}S}$ 服从 $\underline{\quad t(8)\quad}$ 分布.

6. 在研究我国人均消费水平问题上，考虑人均国民收入 $x$（千元）对人均消费金额 $Y$（千元）的影响。设 $Y\sim N(a+bx,s^2),\ a,b,s^2$ 均未知，$(x_1,y_1)\mathbf{L}(x_{19},y_{19})$ 是 1980-1998 年的数据，已知 $\bar{x}=2.32,\ \ \bar{y}=1.09,\ \ \sum_{i=1}^{19}(x_i-\bar{x})^2=73.980,\ \ \sum_{i=1}^{19}(y_i-\bar{y})^2=15.343$，$\sum_{i=1}^{19}(x_i-\bar{x})(y_i-\bar{y})=33.291$，采用最小二乘估计，则回归方程 $\hat{y}=$ $\underline{\quad 0.046+0.45x\quad}$.

二．（11 分）有 A,B 两盒，A 盒中有 1 个红球 1 个白球，B 盒中有 4 件正品 2 件次品。先从 A 盒中采用<u>放回</u>抽样取 2 球，$X$ 表示从 A 盒中取到的红球数，若 $X$=1 时，则从 B 盒中采用<u>不放回</u>抽样取 3 件产品；若 $X\neq1$ 时，从 B 盒中采用<u>不放回</u>抽样取 2 件产品。$Y$ 表示从 B 盒中取到的次品数。(1)已知 $X$=1，求 $Y$ 的条件分布律；(2) 求 $Y$ 的分布律.

（1）$P(Y=0\,|\,X=1)=C_2^0C_4^3/C_6^3=0.2$；$P(Y=1\,|\,X=1)=C_2^1C_4^2/C_6^3=0.6$；

$P(Y=2\,|\,X=1)=C_2^2C_4^1/C_6^3=0.2$ ……………………..5 分

（2）$P(Y=0)=P(X=1)P(Y=0\,|\,X=1)+P(X\neq1)P(Y=0\,|\,X\neq1)=\frac{1}{2}\times\frac{1}{5}+\frac{1}{2}\times\frac{1}{5}=\frac{3}{10}$

$P(Y=1)=P(X=1)P(Y=1\,|\,X=1)+P(X\neq1)P(Y=1\,|\,X\neq1)=\frac{1}{2}\times\frac{3}{5}+\frac{1}{2}\times\frac{8}{15}=\frac{17}{30}$

$P(Y=2)=P(X=1)P(Y=2\,|\,X=1)+P(X\neq1)P(Y=2\,|\,X\neq1)=\frac{1}{2}\times\frac{1}{5}+\frac{1}{2}\times\frac{1}{15}=\frac{4}{30}$

……………………..11 分

三.（12 分）设总体 X 服从参数为 $l$ 的泊松分布，$X_1,\mathbf{L},X_{200}$ 为来自 X 的简单随机样本，

$\overline{X}$ 是样本均值;（1）若 $l=2$，求 $P(X_1\geq 2)$ 的值，以及 $P(\overline{X}>2.1)$ 的近似值。（2）若 $l>0$ 未知，判断统计量 $T=\dfrac{1}{200}\sum\limits_{i=1}^{200}X_i(X_i-1)$ 是否为 $l^2$ 的无偏估计量，说明理由.

（1）$P(X_1\geq 2)=1-P(X_1<2)=1-P(X_1=0)-P(X_1=1)=1-3e^{-2}=0.594$ ……(4 分)

由中心极限定理 $\overline{X}\overset{\text{近似}}{\sim}N(2,0.1^2)$，$P(\overline{X}>2.1)\approx 1-\Phi(\dfrac{2.1-2}{0.1})=0.16$……(8 分)

（2）$E(T)=E\{\dfrac{1}{200}\sum\limits_{i=1}^{200}X_i(X_i-1)\}=E(X(X-1))=E(X^2)-E(X)=D(X)+E^2(X)-E(X)$

$$=l+l^2-l=l^2$$

所以 T 是 $l^2$ 的无偏估计量。 …………….(12 分)

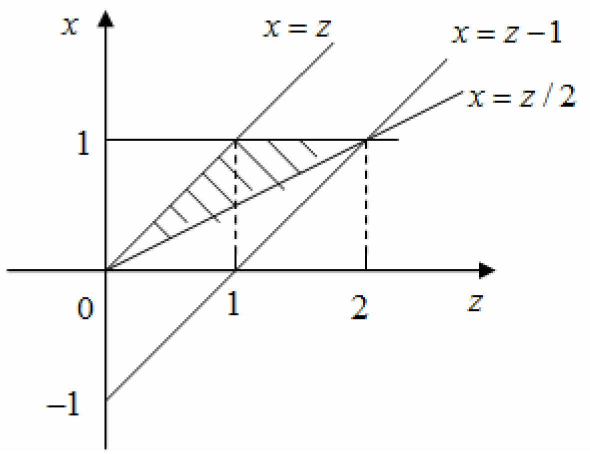
四．(12 分) 设随机变量(X,Y)的密度函数 $f(x,y)=\begin{cases}6(x-y),0<y<x<1\\ 0,其它\end{cases}$，求

（1）$P(Y>0.5)$; （2）X 的边际密度函数 $f_X(x)$ (3) 设 Z=X+Y,求 Z 的密度函数 $f_Z(z)$

（1）$P(Y>0.5)=\iint\limits_{y>0.5}f(x,y)dxdy=\int_{0.5}^{1}dx\int_{0.5}^{x}6(x-y)dy=\int_{0.5}^{1}(3x^2-3x+3/4)=1/8$……(4 分)

（2）$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\begin{cases}\int_0^x 6(x-y)dy=3x^2,0<x<1\\ 0 \qquad ,其它\end{cases}$ ………..….(8 分)

(3) $f_Z(z)=\int_{-\infty}^{\infty}f(x,z-x)dx=\begin{cases}\int_{z/2}^{z}6(2x-z)dx=3z^2/2 \qquad ,0<z<1\\ \int_{z/2}^{1}6(2x-z)dx=3(2-z)^2/2,1<z<2\\ 0 \qquad ,其它\end{cases}$ ………..….(12 分)



五.（12 分）设两个独立正态总体 $X \sim N(m_1, s_1^2), Y \sim N(m_2, s_2^2)$ ,现分别从总体 $X$ 和 $Y$ 中取得容量为 10 和 8 的样本，测得样本均值 $\bar{x} = 148.32, \bar{y} = 141.11$，样本标准差 $s_1 = 6.4$，$s_2 = 5.4$ .(1)以显著水平 0.05 检验假设 $H_0 : s_1^2 = s_2^2, H_1 : s_1^2 \neq s_2^2$； (2) 设 $s_1^2 = s_2^2 = s^2$ 未知，求 $m_1 - m_2$ 的置信度为 95%的双侧置信区间.

(1) 取检验统计量 $F = S_1^2 / S_2^2$,

$H_0$ 的拒绝域为 $F \geq F_{a/2}(n_1 - 1, n_2 - 1)$ 或 $F \leq F_{1-a/2}(n_1 - 1, n_2 - 1)$

$F = S_1^2 / S_2^2 = 6.4^2 / 5.4^2 = 1.04$；

$F_{0.025}(9,7) = 4.82 \quad F_{0.975}(9,7) = 1/4.20 = 0.238$

可见 $F_{1-a/2}(n_1 - 1, n_2 - 1) < F < F_{a/2}(n_1 - 1, n_2 - 1)$，

因此接受原假设，即认为方差相同。 …………………..（6 分）

（2）取枢轴量 $G = \dfrac{\bar{X} - \bar{Y} - (m_1 - m_2)}{S_w \sqrt{1/n_1 + 1/n_2}} \sim t(n_1 + n_2 - 2)$；

设 $P(|G| \leq t_{a/2}(n_1 + n_2 - 2)) = 1 - a$

即得 $m_1 - m_2$ 的 $1 - a$ 的双侧置信区间为 $(\bar{X} - \bar{Y} \pm t_{a/2}(n_1 + n_2 - 2) S_w \sqrt{1/n_1 + 1/n_2})$

$$S_w = \sqrt{\{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2\} / (n_1 + n_2 - 2)} = \sqrt{35.7975} = 5.983$$

$$t_{a/2}(n_1 + n_2 - 2) S_w \sqrt{1/n_1 + 1/n_2} = 2.12 * 5.983 * \sqrt{1/10 + 1/8} = 6.017$$

$$(\bar{X} - \bar{Y} \pm t_{a/2}(n_1 + n_2 - 2) S_w \sqrt{1/n_1 + 1/n_2}) = (7.21 \pm 6.017) = (1.193, 13.227)$$

……………..……(12 分)

六.（14 分）对总体进行 100 次独立重复观察，得到观察值 $x_i, i=1,\mathbf{L},100$,其中最小值为 1.01，最大值为 520.1，平均值为 16.7，具体数据分布如下：

| 观察值 $x_i$ 的范围 | $x \le 1.6$ | $1.6 < x \le 2$ | $2 < x \le 4$ | $4 < x \le 10$ | $x > 10$ |
|---|---|---|---|---|---|
| 频数 $n_i$ | 33 | 17 | 23 | 12 | 15 |

(1)若总体 $X$ 的概率密度函数为 $f(x,q)=\begin{cases} q/x^2, x \ge q \\ 0,其它 \end{cases}$，求 $q$ 的极大似然估计值；

（2）在显著水平 0.05 下用 $c^2$ 拟合检验法检验 $H_0$：总体 $X$ 的概率密度 $f(x)=\begin{cases} x^{-2}, x \ge 1 \\ 0,其它 \end{cases}$.

(1)设 $L(q)=\prod_{i=1}^{100} f(x_i)=\prod_{i=1}^{100} q/x_i = q^{100}/(x_1x_2\mathbf{L}\, x_{100})$

$\ln L(q)=100\ln q-\ln(x_1x_2\mathbf{L}\, x_{100})$

$\ln L(q)/dq=100/q>0$， $L(q)$关于$q$增函数，

$\hat{q}_L=\min(x_1,\mathbf{L},x_{100})=1.01$

(2) $q=1$,

$P(X\le 1.6)=\int_1^{1.6} x^{-2}dx=0.375$,同理得$X$落在其他四个区间的概率为:0.125,0.25,0.15,0.1.

| 观察值 $x_i$ 的范围 | $x \le 1.6$ | $1.6 < x \le 2$ | $2 < x \le 4$ | $4 < x \le 10$ | $x > 10$ |
|---|---|---|---|---|---|
| 频数 $n_i$ | 33 | 17 | 23 | 12 | 15 |
| 概率 $p_i$ | 0.375 | 0.125 | 0.25 | 0.15 | 0.1 |
| 理论频数 $np_i$ | 37.5 | 12.5 | 25 | 15 | 10 |

检验统计量$c^2=\sum_{i=1}^{5}\frac{n_i^2}{np_i}-n$， $H_0$的拒绝域$c^2=\sum_{i=1}^{k}\frac{n_i^2}{np_i}-n \ge c_a^2(k-r-1)$

这里 $n$=100, $k$=5, $r$=0

$$c^2=\sum_{i=1}^{k}\frac{n_i^2}{np_i}-n=\frac{33^2}{37.5}+\frac{17^2}{12.5}+\frac{23^2}{25}+\frac{12^2}{15}+\frac{15^2}{10}-100=5.42<c_a^2(k-r-1)=c_{0.05}^2(4)=9.49$$

接受原假设，即认为总体 $X$ 的概率密度是 $f(x)=\begin{cases} x^{-2}, x \ge 1 \\ 0,其它 \end{cases}$。

一填空

5（2）

$$\mathbf{Q}\ X_{10}-\overline{X}=-\frac{1}{9}X_1-\frac{1}{9}X_1-\mathbf{L}-\frac{1}{9}X_9+X_{10}$$

$\therefore X_{10}-\overline{X}$服从正态分布

$$E(X_{10}-\overline{X})=E(X_{10})-E(\overline{X})=m-m=0$$

$$D(X_{10}-\overline{X})=D(X_{10})+D(\overline{X})-2Cov(X_{10},\overline{X})$$

$$\overset{\text{独立}}{=}s^2+\frac{s^2}{9}+0=\frac{10}{9}s^2$$

$$\therefore X_{10}-\overline{X}\sim N(0,\frac{10}{9}s^2)$$

$$\text{记}U=\frac{(X_{10}-\overline{X})-0}{\sqrt{\frac{10}{9}s^2}}=\frac{3(X_{10}-\overline{X})}{s\sqrt{10}}\sim N(0,1)$$

$$\text{记}V=\frac{(9-1)S^2}{s^2}\sim c^2(8)$$

由抽样分布定理2知：$U,V$独立

$$\frac{U}{\sqrt{V/8}}=\frac{\frac{3(X_{10}-\overline{X})}{s\sqrt{10}}}{\sqrt{\frac{(9-1)S^2}{s^2}/8}}=\frac{3(X_{10}-\overline{X})}{S\sqrt{10}}\sim t(8)$$