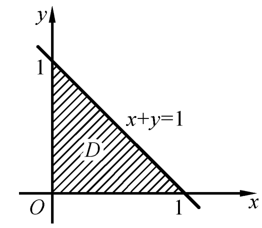
厦门大学《概率统计 I》期末试卷

________学院________系________年级________专业

主考教师：　　试卷类型：（A 卷）

一、**(14 分)** 设二维随机变量（X，Y）在以（0，0)，（0，1)，（1，0）为顶点的三角形区域上服从均匀分布，求 Cov（X，Y），ρ_{XY}.

【解】如图，$S_D=\frac{1}{2}$，故（X，Y）的概率密度为



题 18 图

$$f(x,y)=\begin{cases}2, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

$$E(X)=\iint_D xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{1-x}x\cdot 2\mathrm{d}y=\frac{1}{3}$$

$$E(X^2)=\iint_D x^2 f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{1-x}2x^2\mathrm{d}y=\frac{1}{6}$$

从而 $D(X)=E(X^2)-[E(X)]^2=\frac{1}{6}-\left(\frac{1}{3}\right)^2=\frac{1}{18}$.

同理 $E(Y)=\frac{1}{3},D(Y)=\frac{1}{18}$.

而 $$E(XY)=\iint_D xyf(x,y)\mathrm{d}x\mathrm{d}y=\iint_D 2xy\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{1-x}2xy\mathrm{d}y=\frac{1}{12}.$$

所以

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)\cdot E(Y)=\frac{1}{12}-\frac{1}{3}\times\frac{1}{3}=-\frac{1}{36}.$$

从而 $$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}=\frac{-\frac{1}{36}}{\sqrt{\frac{1}{18}}\times\sqrt{\frac{1}{18}}}=-\frac{1}{2}$$

二、（**10 分**）甲乙两个剧院在竞争 1000 名观众。假定每个观众完全随意地选择一个剧院，且观众选择剧院是彼此独立的，用中心极限定理计算，每个剧院应设置多少个座位，才能保证

因缺少座位而使观众离去的概率小于 1%？ $\Phi(2.33)=0.99$

解： 设甲剧院应设 M 个座位。令

$$X_i=\begin{cases}1, & \text{第i个观众选择甲剧院，}\\ 0, & \text{第i个观众选择乙剧院，}\end{cases}\quad i=1,2,\cdots,1000$$

$X_i(i=1,2,\cdots,1000)$相互独立同分布，$P\{X_i=1\}=P\{X_i=0\}=\frac{1}{2}$

$E（X_i）=\frac{1}{2}$，$D（X_i）=\frac{1}{4}$.$P\{X\le M\}=P\{\sum_{i=1}^{1000}X_i\le M\}\approx\Phi（\frac{M-500}{5\sqrt{10}}）\ge 99\%$，

$\frac{M-500}{5\sqrt{10}}\ge 2.33$，$M\ge 537$.

三、(14 分) （1）设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta(0<\theta<\frac{1}{2})$是未知参数，利用总体的如下样本值 3，1，3，0，3，1，2，3 求 θ 的矩估计值

（2）设总体 $X\sim N(1,\sigma^2)$，σ^2（$\sigma^2>0$）未知，x_1,x_2,Λ,x_n 为一相应的样本值。求 σ^2 的最大似然估计。

【解】

$(1)E(X)=3-4\theta$,令$E(X)=\bar{x}$得$\hat{\theta}=\frac{3-\bar{x}}{4}$

又 $$\bar{x}=\sum_{i=1}^{8}\frac{x_i}{8}=2$$

所以 θ 的矩估计值 $\hat{\theta}=\frac{3-\bar{x}}{4}=\frac{1}{4}$.

（2）似然函数为 $L(\mu)=\prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-1)^2}{2\sigma^2}}\right]=\frac{1}{\left(\sqrt{2\pi}\sigma\right)^n}e^{-\frac{\sum_{i=1}^{n}(x_i-1)^2}{2\sigma^2}}$，相应的对数似然函数为

$$\ln L(\mu)=-\frac{\sum_{i=1}^{n}(x_i-1)^2}{2\sigma^2}-\ln\left(\sqrt{2\pi}\sigma\right)^n。$$

令对数似然函数对 σ^2 的一阶导数为零，得到 σ^2 的最大似然估计值为

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-1)^2\ 。$$

四、(12 分) 10，以 X 表示耶路撒冷新生儿的体重（以克计），设 $X\sim N(\mu,\sigma^2)$，μ,σ^2 均未知。现测得一容量为 30 的样本，得样本均值为 3189，样本**标准差**为 488。试检验假设（$\alpha=0.1$）：

（1）$H_0:\mu\ge 3315$，$H_1:\mu<3315$。$t_{0.1}(29)=1.3114$

（2）$H'_0:\sigma\le 525$, $H'_1:\sigma>525$。$\chi^2_{0.1}(29)=42.557$

解:（1）这是一个方差未知的正态总体的均值检验，属于左边检验问题，检验统计量为

$$t=\frac{\bar{x}-3315}{s/\sqrt{n}}。$$

代入本题具体数据，得到$t=\dfrac{3189-3315}{488/\sqrt{30}}=-1.4142$。

检验的临界值为$-t_{0.1}(29)=-1.3114$。

因为$t=-1.4142<-1.3114$，所以样本值落入拒绝域中，故拒绝原假设H_0，即认为$\mu<3315$。

（2）题中所要求检验的假设实际上等价于要求检验假设

$$H'_0:\sigma^2\le 525^2,\quad H'_1:\sigma^2>525^2$$

这是一个正态总体的方差检验问题，属于右边检验。
检验统计量为

$$\chi^2=\frac{(n-1)s^2}{525^2}。$$

代入本题中的具体数据得到$\chi^2=\dfrac{(30-1)\times 488^2}{525^2}=25.0564$。

检验的临界值为$\chi^2_{0.1}(29)=42.557$。

因为$\chi^2=25.0564<42.557$，所以样本值没有落入拒绝域，因此接受原假设，即认为标准差不大于525

五、（10 分）．为比较两个学校同一年级学生数学课程的成绩，随机地抽取学校 A 的 9 个学生，得分数的平均值为$\bar{x}_A=81.31$，方差为$s_A^2=60.76$；随机地抽取学校 B 的 15 个学生，得分数的平均值为$\bar{x}_B=78.61$，方差为$s_B^2=48.24$。设样本均来自正态总体且方差相等，参数均未知，两样本独立。求均值差$\mu_A-\mu_B$的置信水平为 0.95 的置信区间。

$t_{0.025}(22)=2.0739$

解：根据两个正态总体均值差的区间估计的标准结论，均值差$\mu_A-\mu_B$的置信水平为 0.95 的置信区间为

$$\left((\bar{x}_A-\bar{x}_B)\pm s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{0.025}(n_1+n_2-2)\right)=\left(2.7\pm s_w\sqrt{\frac{1}{9}+\frac{1}{15}}t_{0.025}(22)\right)$$

$$=\left(2.7\pm s_w\sqrt{\frac{1}{9}+\frac{1}{15}}t_{0.025}(22)\right)=\left(2.7\pm 7.266\times\sqrt{\frac{1}{9}+\frac{1}{15}}\times 2.0739\right)$$

$$=(2.7\pm 6.35)=(-3.65,\quad 9.05)$$

六、(14 分) 随机抽取了 10 个家庭，调查了他们的家庭月收入 x（单位：百元）和月支出 y(单位：百元)，记录于下表：

x	20	15	20	25	16	20	18	19	22	16
y	18	14	17	20	14	19	17	18	20	13

经计算可得 $\sum_{i=1}^{10}x_i=191,\sum_{i=1}^{10}y_i=170,\sum_{i=1}^{10}x_i^2=3731,\sum_{i=1}^{10}x_iy_i=3310,\sum_{i=1}^{10}y_i^2=2948,$

求：(1) 求 y 与 x 的一元线性回归方程.

(2) 对所得的回归方程作显著性检验.（α=0.05） $F_{0.05}(1,8)=5.32,t_{0.025}(8)=2.306$

【解】(1) 经计算可得

$$\sum_{i=1}^{10}x_i=191,\sum_{i=1}^{10}y_i=170,\sum_{i=1}^{10}x_i^2=3731,\sum_{i=1}^{10}x_iy_i=3310,\sum_{i=1}^{10}y_i^2=2948,$$

$$S_{xx}=82.9,S_{xy}=63,S_{yy}=58.$$

$$故\hat{b}=\frac{S_{xy}}{S_{xx}}=0.7600,\hat{a}=\frac{170}{10}-0.76\times\frac{191}{10}=2.4849,$$

从而回归方程：$\hat{y}=2.4849+0.76x.$

(2) $Q_A=\frac{S_{xy}^2}{S_{xx}}=47.8770,Q_E=Q_T-Q_A=58-47.877=10.1230,$

$$F=\frac{Q_A}{Q_E/n-2}=37.8360>F_{0.05}(1,8)=5.32.$$

故拒绝 H_0,即两变量的线性相关关系是显著的.

解法二:

$$\hat{\sigma}=\sqrt{\frac{s_{yy}-\hat{b}s_{xy}}{n-2}}=\sqrt{\frac{58-0.76\times 63}{8}}=1.124$$

$$|t|=\frac{|\hat{b}|}{\hat{\sigma}}\sqrt{s_{xx}}=\frac{|0.76|}{1.124}\sqrt{82.9}=6.15>t_{0.025}(8)=2.306.$$

故拒绝 H_0,即两变量的线性相关关系是显著的.

七、(10 分) 灯泡厂用 4 种不同的材料制成灯丝，检验灯线材料这一因素对灯泡寿命的影响. 若灯泡寿命服从正态分布，不同材料的灯丝制成的灯泡寿命的方差相同，试根据表中试验结果记录，在显著性水平 0.05 下检验灯泡寿命是否因灯丝材料不同而有显著差异？

		试验批号							
		1	2	3	4	5	6	7	8
灯丝	A_1	1600	1610	1650	1680	1700	1720	1800	

材料	A_2	1580	1640	1640	1700	1750			
水平	A_3	1460	1550	1600	1620	1640	1660	1740	1820
	A_4	1510	1520	1530	1570	1600	1680		

计算得到 $\sum_{i=1}^{4}\sum_{j=1}^{4}x_{ij}^2=69895900$, $T_{.1}=11760, T_{.2}=8310, T_{.3}=13090, T_{.4}=9410. F_{0.05}(3,22)=3.05.$

【解】

$r=4, n=\sum_{i=1}^{r}n_i=26;$

$S_T=\sum_{i=1}^{4}\sum_{j=1}^{4}x_{ij}^2-\frac{T_{..}^2}{n}$=69895900-69700188.46=195711.54,

$S_A=\sum_{i=1}^{4}\frac{1}{n_i}T_{i.}^2-\frac{T_{..}^2}{n}$=69744549.2-69700188.46=44360.7,

$S_E=S_T-S_A$=151350.8,

$F=\frac{S_A/(r-1)}{S_E/(n-r)}=\frac{44360.7/3}{151350.8/22}=2.15$,

$F_{0.05}(3,22)=3.05>F.$

故灯丝材料对灯泡寿命无显著影响.

方差来源	平方和 S	自由度	均方和 $\bar{S}$	F 值
因素影响	44360.7	3	14786.9	2.15
误差	151350.8	22	6879.59	
总和	195711.54	25		

.

八. **(10 分)**统计了日本西部地震在一天中发生的时间段，共观察了 527 次地震，这些地震在一天中的四个时间段的分布如下表

时间段	0 点—6 点	6 点—12 点	12 点—18 点	18 点—24 点
次　数	123	135	141	128

试取 $\alpha=0.05$ 检验假设：地震在各个时间段内发生时等可能的。$\chi^2_{0.05}(3)=7.815$

解：根据题意，要检验以下假设：

H_0：地震的发生时间在（0,24）内是均匀分布的

检验统计量为 $\chi^2=\sum_{i=1}^{4}\frac{f_i^2}{np_i}-n$，其中 $p_i=6/24=0.25$。

代入本题中的数据得到 $\chi^2=\frac{123^2+135^2+141^2+128^2}{527\times0.25}-527=1.417$，检验的临界值为

$\chi^2{}_{0.05}(4-1)=7.815$。因为 $\chi^2=1.417<7.815$，所以样本值没有落入拒绝域，因此接受原假设，即认为地震在各个时间段内发生时等可能的。

九、(6 分) 设总体 X~N（μ,σ^2），X_1，X_2，…，X_{2n}（$n\geq2$）是总体 X 的一个样本，$\overline{X}=\frac{1}{2n}\sum_{i=1}^{2n}X_i$，

令 $Y=\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$，求 E（Y）.

【解】令 $Z_i=X_i+X_{n+i}$, $i=1,2,\ldots,n$.则

$Z_i\sim N(2\mu,2\sigma^2)(1\leq i\leq n)$,且 $Z_1,Z_2,\ldots,Z_n$ 相互独立.

令
$$Z=\sum_{i=1}^{n}\frac{Z_i}{n},\ S^2=\sum_{i=1}^{n}(Z_i-\overline{Z})^2/n-1,$$

则
$$\overline{X}=\sum_{i=1}^{2n}\frac{X_i}{2n}=\frac{1}{2n}\sum_{i=1}^{n}Z_i=\frac{1}{2}\overline{Z},$$

故
$$\overline{Z}=2\overline{X}$$

那么
$$Y=\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2=\sum_{i=1}^{n}(Z_i-\overline{Z})^2=(n-1)S^2,$$

所以
$$E(Y)=(n-1)ES^2=2(n-1)\sigma^2.$$