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）

分数	阅卷人

一、（16 分）（1）设随机变量 X 和 Y 的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1+xy}{3} & 0<x<2,0<y<1\\ 0 & \text{其　他}\end{cases}$$

求 X 和 Y 的相关系数；

（2）对随机变量 X 和 Y，已知 D(X)=2，D(Y)=3，Cov(X,Y)=1, Z=3X-2Y+1 ，U=X+4Y-3，求 Z 和 U 的相关系数。

解（1）$f_X(x)=\dfrac{1}{3}+\dfrac{x}{6},0<x<2$；$f_Y(y)=\dfrac{2}{3}+\dfrac{2y}{3},0<y<1$

E(X)=10/9，E(Y)=5/9，E(XY)=17/27,　　（3 分）

Cov(X,Y)=1/81.　　（1 分）

$$E(X^2)=\frac{14}{9},D(X)=\frac{26}{81};E(Y^2)=\frac{7}{18},D(Y)=\frac{13}{162}.\rho_{XY}=\frac{1}{13}$$

（5 分）

(2) Cov(Z,U)=3DX+10Cov(X,Y)-8DY=-8.　　（2 分）

D(Z)=9DX+4DY-12Cov(X,Y)=18　　（2 分）

D(U)=DX+16DY+8Cov(X,Y)=58　　（2 分）

$$\rho_{Z,U}=-\frac{4}{3\sqrt{29}}$$

（1 分）

分数	阅卷人

二、（14分）某药厂生产的某种药品医治一种疑难的血液病,医院检验员任意抽查100个服用此药品的病人，

（1） 若此药品对这种疾病治愈率是0.8，问多于75人治愈的概率是多少？

（2）若有80人治愈,求治愈率p的置信水平为0.95的置信区间？$\Phi(1.25)=0.8944$，$\Phi(1.96)=0.975$

解 (1) X服从$B(100,0.8)$，

$$P\{\sum_{i=1}^{100}X_i>75\}=1-P\{X\le 75\}\approx 1-\Phi\left(\frac{75-100\times 0.8}{\sqrt{100\times 0.8\times 0.2}}\right)$$

（4分）

$$=1-\Phi(-1.25)=\Phi(1.25)=0.8944.$$

（3分）

(3) X服从b(100,p)

$$\left|\frac{\mathrm{X}-100\mathrm{p}}{\sqrt{100\mathrm{p}（1-p)}}\right|<1.96$$

$$（\mathrm{X}-100\mathrm{p}）^2<1.96^2\times 100p(1-p)$$

$$64-163.841\mathrm{p}+103.841\mathrm{p}^2<0$$

（4分）

$$\mathrm{p}\in（0.7889-\sqrt{0.006},0.7889+\sqrt{0.006}）=（0.711,0.8667）$$

（3分）

分数	阅卷人

三、（12 分）设总体 X 的概率密度为

$$f(x)=\frac{1}{\sigma}\exp\left\{-\frac{x+1}{\sigma}\right\},\qquad x>-1$$

其中σ是未知参数，从总体中抽取 n=5 的简单随机样本，样本值为 1，1，2，3，3，求σ的矩估计（6 分）和最大似然估计（6 分）。

计算一阶矩

$$EX=\int_{-1}^{\infty}\frac{x}{\sigma}\exp\left\{-\frac{x+1}{\sigma}\right\}dx=\int_{0}^{\infty}\frac{1}{\sigma}(\sigma t-1)e^{-t}\,dt=\sigma-1$$

所以

$$\hat{\sigma}-1=\bar{X}=\frac{1}{5}(1+1+2+3+3)=2$$

σ的矩估计量为

$$\hat{\sigma}=3.$$

数据的似然函数为

$$L=\prod_{i=1}^{5}\frac{1}{\sigma}\exp\left\{-\frac{X_i+1}{\sigma}\right\}=\left(\frac{1}{\sigma}\right)^5\exp\left\{-\frac{1}{\sigma}\sum_{i=1}^{5}(X_i+1)\right\}=\frac{1}{\sigma^5}\exp\left\{-\frac{15}{\sigma}\right\}$$

取对数，求导，

$$\frac{\partial\ln L}{\partial\sigma}=\frac{\partial\left(-5\ln\sigma-\frac{15}{\sigma}\right)}{\partial\sigma}=\frac{-5}{\sigma}+\frac{15}{\sigma^2}=0$$

所以σ的最大似然估计量为

$$\hat{\sigma}_{MLE}=3.$$

分数	阅卷人

四、（12 分）设某种砖头的抗压强度 X 服从 N（μ，σ²），今随机抽取 9 块砖头，测得数据如下：

64 69 49 92 55 97 41 84 88

（1） 求 μ 的置信概率为 0.95 的置信区间；

（2） 求 σ^2 的置信概率为 0.95 的置信区间。

$(t_{0.025}(8)=2.306, \chi^2_{0.025}(8)=17.534, \chi^2_{0.975}(8)=2.18)$

解 $\overline{X}=71, S^2=408.5, S=20.21$ （4 分）

μ 的置信概率为 0.95 的置信区间$(71\pm 2.306\dfrac{20.21}{3})=(55.464,86.535)$ （4 分）

σ^2 的置信概率为 0.95 的置信区间$(\dfrac{8\times 408.5}{17.534},\dfrac{8\times 408.5}{2.18})=(186.38,1499.082)$ （4 分）

分数	阅卷人

五、（12 分）溪流混浊是由于水中有悬浮固体，分别以 X,Y 表示晴天和雨天水的混浊度（以 NTU 单位计）的总体。设 $X\sim N(\mu_X,\sigma_X{}^2)$，$Y\sim N(\mu_Y,\sigma_Y{}^2)$，$\mu_X,\mu_Y,\sigma_X{}^2,\sigma_Y{}^2$ 均未知。今取到总体 X 的容量 $n_1=9$ 的样本，算得样本均值为 $\overline{x}=93$，样本标准差为 $s_X=12.9$；取到总体 Y 的容量为 $n_2=11$ 的样本，算得样本均值为 $\overline{y}=132$，样本标准差为 $s_Y=7.1$，两样本独立。

（1）试检验假设($\alpha=0.05$)：$H_0:\sigma_X^2=\sigma_Y^2$，$H_1:\sigma_X^2\neq\sigma_Y^2$；

（2）如能接受 H_0，接着检验假设($\alpha=0.025$) $H'_0:\mu_X\ge\mu_Y$，$H'_1:\mu_X<\mu_Y$。

$t_{0.025}(18)=2.1009$，$F_{0.025}(8,10)=3.85$，$F_{0.025}(10,8)=4.3$

解：（1）这是一个两个正态总体的方差之比的检验问题，属于双边检验。检验统计量为 $F=\dfrac{s_X^2}{s_Y^2}$ （2 分）

代入本题中的具体数据得到 $F=\dfrac{12.9^2}{7.1^2}=3.301$。（2 分）

检验的临界值为 $F_{0.025}(8,10)=3.85$，$F_{0.975}(8,10)=\dfrac{1}{4.3}=0.2326$。因为 $0.2326<F=3.301<3.85$，所以样本值没有落入拒绝域，因此接受原假设，即认为两总体方差相等。

（2 分）

（2）因为两总体方差相等，所以这是一个方差相等的两个正态总体的均值之差的检验问题，属于左边检验。检验统计量为

$$t=\frac{(\overline{x}-\overline{y})-0}{s_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$$ （2 分）

代入本题中的具体数据得到 $t=\dfrac{(93-132)-0}{10.1\times\sqrt{\dfrac{1}{9}+\dfrac{1}{11}}}=-8.5929$。（2 分）

检验的临界值为 $t_{0.025}(18)=2.1009$。因为 $t=-8.5929<-2.1009$，所以样本值落入拒绝域，因此拒绝原假设。

（2 分）

分数	阅卷人

六、（12 分）美国《教育统计文摘》1993 年版给出该国 18 岁或以上的人持有学士或更高学位的年龄分布如下

年　龄	18~24	25~34	35~44	45~54	55~64	65 或以上
百分比	5	29	30	16	10	10

在阿拉斯加州随机选择 500 个 18 岁或以上的持有学士或更高学位的一项调查给出如下数据

年　龄	18~24	25~34	35~44	45~54	55~64	65 或以上
人　数	30	150	155	75	35	55

试取 $\alpha = 0.1$ 检验该地区年龄分布是否和全国一样。$\chi^2{}_{0.1}(5) = 9.236$

解：根据题意，要检验以下假设：

H_0：　阿拉斯加州的年龄分布律为

年　龄	18~24	25~34	35~44	45~54	55~64	65 或以上
概　率	0.05	0.29	0.30	0.16	0.10	0.10

检验统计量为 $\chi^2 = \sum_{i=1}^{6} \frac{f_i^2}{np_i} - n$。（6 分）所需计算列表如下：

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
A_1	30	0.05	25	36
A_2	150	0.29	145	155.172
A_3	155	0.30	150	160.167
A_4	75	0.16	80	70.313
A_5	35	0.10	50	24.5
A_6	55	0.10	50	60.5

$\chi^2 = \sum_{i=1}^{6} \frac{f_i^2}{np_i} - n = 506.652 - 500 = 6.652$，（4 分）检验的临界值为 $\chi^2{}_{0.1}(6-1) = 9.236$。因为 $\chi^2 = 6.652 < 9.236$，所以样本值没有落入拒绝域，因此接受原假设，即认为阿拉斯加州的年龄分布与全国的分布一样。（2 分）

分数	阅卷人

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？$F_{0.05}(3,14)=3.34.$

		试验批号				
		1	2	3	4	5
灯丝材料水平	A_1	19	22	20	18	15
	A_2	20	21	33	27	40
	A_3	16	15	18	26	17
	A_4	18	22	19		

【解】

$$r=4, n=\sum_{i=1}^{r} n_i = 18;$$

$S_T=\sum_{i=1}^{4}\sum_{j=1}^{4}x_{ij}^2-\frac{T_{\cdot\cdot}^2}{n}$=8992—386*386/18=714.44,（2 分）

$S_A=\sum_{i=1}^{4}\frac{1}{n_i}T_{i\cdot}^2-\frac{T_{\cdot\cdot}^2}{n}$=318.98,（2 分）

$S_E=S_T-S_A$=395.46（2 分）

$F=\frac{S_A/(r-1)}{S_E/(n-r)}=\frac{318.98/3}{395.46/14}=3.76$,（2 分）

$F_{0.05}(3,14)=3.34<F.$

故灯丝材料对灯泡寿命有显著影响.（2 分）

分数	阅卷人

八、（12 分）某医院用光色比色计检验尿贡时，得尿贡含量与肖光系数读数的结果如下：

尿贡含量 x	2	4	6	8	10
肖光系数 y	64	138	205	285	360

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=0.01）$F_{0.99}(1,3)=34.1,\ t_{0.005}(3)=5.8409$

【解】由数据可以求得，n=5

$\sum_{\alpha} x_{\alpha}=30$，$\overline{x}=6$

$\sum_{\alpha} y_{\alpha}=1052, \overline{y}=210.4$

$\sum_{\alpha} x_{\alpha}^2=220, \sum_{\alpha} x_{\alpha} y_{\alpha}=7790, \sum_{\alpha} y_{\alpha}^2=275990$

$l_{xx}=40, l_{xy}=1478, l_{yy}=54649.2$ （5 分）

则，最小二乘估计为：

$$\hat{\beta}_0=-11.3, \hat{\beta}_1=36.95$$ （2 分）

检验假设 $H_0:\beta_1=0$ （1 分）

$\sigma^2=(l_{yy}-\hat{\beta}_1 l_{xy})/(n-2)=12.36$ （1 分）

$$\mathrm{t}=\left|\frac{\hat{\beta}_1\sqrt{l_{xy}}}{\sqrt{(l_{yy}-\hat{\beta}_1 l_{xy})/(n-2)}}\right|=66.46>t_{0.005}(3)=5.8409,$$

（2 分）

因此，拒绝原假设。（1 分）

即两变量的线性相关关系是显著的.

或者

检验假设 $H_0:\beta_1=0$ 可用统计量

$$F=\frac{\hat{\beta}_1^2 l_{xy}}{(l_{yy}-\hat{\beta}_1 l_{xy})/(n-2)}=4416>F_{1-\alpha}(1,3)=34.1, \alpha=0.01$$

因此，拒绝原假设。

即两变量的线性相关关系是显著的.