中 国 科 学 技 术 大 学

2019—2020学年第一学期考试试卷

考试科目 概率论与数理统计　　　得分 ________

所在系 ______ 　　姓名 ________ 　　学号 ________

考试时间: 2020年1月9日下午2:30—4:30;　使用简单计算器

一. 填空判断选择题 (每题3分,答题请写在试卷上):

1 设$Z(t) = X + Yt$, 其中X与Y相互独立，且都服从$N(0, \sigma^2)$, 则$Z(t)$与$Z(s)(s < t)$的相关系数= ________.

2 三人独立地去破译一份密码，已知各人能译出的概率分别为1/5, 1/3, 1/4, 则破译出密码的概率是________.

3 设随机变量 X 和 Y 都服从正态分布, 且他们不相关, 则 ________

(A) X 与 Y一定独立　　(B) (X, Y)服从二元正态

(C) X和Y未必独立　　(D) $X + Y$服从一维正态分布

4 设$X_1, \ldots, X_{16}$服从正态分布$N(\mu, 0.16)$, 检验问题

$$H_0 : \mu = 0.5 \leftrightarrow H_1 : \mu > 0.5,$$

显著水平为0.05. 检验的拒绝域是________, $\mu = 0.65$时犯第二类错误的概率为 ________.

5 设$X \sim N(\mu, \sigma^2)$, 样本$X_1, \cdots, X_n$来自总体X的一组样本, $\overline{X}$是样本均值, 则$c =$ ________时, $c(X_1 + X_2 + X_3 - 3\overline{X})^2$为$\sigma^2$的无偏估计.

6 总体参数θ的置信水平为95%的置信区间是指________.

(A) θ落在该区间的概率为95%　　(B) θ落在该区间的概率为5%

(C) 在用同样方法构造的多个区间中包含θ 的区间比例为95%

(D) 在用同样方法构造的多个区间中包含θ 的区间比例为5%

7 用一组有30个观测值的样本去估计模型$y = a + bx + \epsilon$. 在0.05的显著性水平下对b的显著性作t检验, 则b显著不为0 的条件是其统计量T ________.

(A) $T > t_{29}(0.05)$　　(B) $|T| > t_{29}(0.025)$

(C) $T > t_{28}(0.05)$　　(D) $|T| > t_{28}(0.025)$

8 设$X_1, \ldots, X_n$为来自概率密度函数为$f(x) = \theta x^{\theta-1} I(0 < x < 1)$的总体的样本, θ未知, 则$U = e^{-1/\theta}$的极大似然估计为________.

9 设总体$X \sim N(\mu, \sigma^2)$, σ^2未知, $x_1, \cdots, x_n$为来自X的样本值, 先对μ进行假设检验。若在显著性水平$\alpha = 0.01$下拒绝了$H_0 : \mu = \mu_0$, 则当显著性水平改为$\alpha = 0.05$时, 下列结论正确的是

(A) 必拒绝H_0　　(B)必接受H_0

(C)第一类错误的概率变小　　(D) 可能接受也可能拒绝H_0

装 订 线

二. (12分) 某次考试共有100道选择题，每道题都是4选1.

(1)设某考生一点也不懂, 完全靠蒙，求答对20题以上40题以下的概率是多少？

(2)设另一考生平常模拟考试的正确率时85%, 求他答对90题以上的概率有多大？

三. (15分) 设$X_1, \ldots, X_n$为来自均匀总体$U(-\theta, \theta)$的简单随机样本,

(1) 求θ的矩估计$\hat{\theta}_1$和极大似然估计$\hat{\theta}_2$;

(2) $\hat{\theta}_1$ 和 $\hat{\theta}_2$是否是无偏的？若否，请修正。

(3) 请问修正后的估计那个更有效？

四. (15分) 博尔特科维茨研究了从1875年到1894年20年间14个军团（共计280个军团)士兵被马踢死的人数记录如下：

军团死亡人数	0	1	2	3	4	≥ 5
观察数	144	91	32	11	2	0

博尔特科维茨推断这个每年每军团死亡数是服从Poisson分布.

(1) 求出Poisson分布的参数λ的极大似然估计.

(2) 试检验在水平0.05下博尔特科维茨的推断是否正确？

五. (13分) 做一下试验比较人对红光或绿光的反应时间(单位s). 随机抽取了8人做实验, 结果如下:

红光	0.30	0.23	0.41	0.53	0.24	0.36	0.38	0.51
绿光	0.43	0.32	0.58	0.46	0.27	0.41	0.38	0.61

试问是否可以认为人对红光的反应比绿光快？ ($\alpha = 0.05$)

六. (15分) 考虑日本汇率与汽车出口数量的数据: (X:汇率(日元/美元), Y: 数量(万辆))

年度	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
X	168	145	128	138	145	135	127	111	102	94
Y	661	631	610	588	583	575	567	502	446	379

已知$\sum X_i = 1293, \sum Y_i = 5542, \sum X_i^2 = 171617, \sum Y_i^2 = 3139490, \sum X_i Y_i = 732776.$ 满足经典线性回归模型的假设 ($\alpha = 0.05$).

(1) 求X与Y的样本相关系数.

(2) 求出一元线性回归方程, 即求a和b的最小二乘估计. 试给出参数的经济意义.

(3) 检验回归模型的水平α的显著性检验.

(4) 若1996年日元对美元是104.8, 预测当年的出口汽车数量和其$1-\alpha$的置信区间.

附录 分布及分位数: $\Phi(3.46) = 0.9997$, $\Phi(-2/\sqrt{3}) = 0.124, \Phi(1.4) = 0.92$, $u_{0.025} = 1.960$, $u_{0.05} = 1.645$, $t_{0.025}(13) = 2.16$, $t_{14}(0.025) = 2.145$, $t_{13}(0.05) = 1.771$, $t_7(0.05) = 1.89$, $t_7(0.025) = 2.36$, $t_8(0.025) = 2.306$, $t_9(0.025) = 2.262$

答案： 1. $\frac{1+ts}{\sqrt{(1+t^2)(1+s^2)}}$, 2. 3/5, 3. C, 4. $\bar{X} > 0.6645, \Phi(0.145)$ 5. $\frac{n}{3(n-3)}$ 6. C 7. D 8. $\exp\{\frac{1}{n}\sum \ln X_j\}$ 9. A

二、 (1)令X_i=I(第i题答对),

$$P(20 \le \sum_{i=1}^{100} X_i \le 40) = P(\frac{20-25}{\sqrt{100*0.25*0.75}} \le \frac{\sum_{i=1}^{100} X_i - 25}{\sqrt{100*0.25*0.75}} \le \frac{40-25}{\sqrt{100*0.25*0.75}})$$
$$= \Phi(2\sqrt{3}) - \Phi(-2/\sqrt{3}) = 0.8785.$$

(2)令Y_i=I(第i题答对),

$$P(\sum_{i=1}^{100} Y_i \ge 90) = P(\frac{\sum_{i=1}^{100} Y_i - 85}{\sqrt{100*0.85*0.15}} \ge \frac{90-85}{\sqrt{100*0.85*0.15}})$$
$$= 1 - \Phi(1.4) = 0.08.$$

三

(1)令$Y_i = |X_i|$,则$Y_i \sim U(0,\theta), EY_i = \theta/2$.

所以矩估计$\hat{\theta}_1 = 2\bar{Y}$.

似然函数$L(y,\theta) = (\frac{1}{\theta})^n \prod_{i=1}^n I(0 < Y_i < \theta) = (\frac{1}{\theta})^n I(Y_{(n)} < \theta)$.

所以极大似然估计$\hat{\theta}_2 = \max Y_i = Y_{(n)}$.

(2)$E\hat{\theta}_1 = 2EY_1 = \theta$, 所以$\hat{\theta}_1$是无偏的.

$\hat{\theta}_2 \sim f(y) = 1/\theta(y/\theta)^{n-1}$, $E\hat{\theta}_2 = \int_0^\theta n(y/\theta)^n dy = \frac{n}{n+1}\theta$.

$\hat{\theta}_2$不是无偏的,修正$\hat{\theta}_2^* = \frac{n+1}{n}\hat{\theta}_2$.

(3)$var(\hat{\theta}_1) = \frac{4}{n}var(Y_1) = \frac{\theta^2}{3n}$,

$var(\hat{\theta}_2^*) = \frac{(n+1)^2}{n^2}(E\hat{\theta}_2^2 - (E\hat{\theta}_2)^2) = \frac{\theta^2}{n(n+2)}$,

$var(\hat{\theta}_1) \ge var(\hat{\theta}_2)$, 当且仅当$n=1$时取等, 所以修正后极大似然估计更有效.

四.

(1)令n_i为死亡人数为i的观察数,似然函数$L = \prod_i (e^{-\lambda}\frac{\lambda^i}{i!})^{n_i} = C_1 e^{-\lambda\sum_i n_i}\lambda^{\sum_i in_i}$,

对数似然函数$l = C_2 - \lambda\sum_i n_i + \sum_i in_i log\lambda$,

$\hat{\lambda} = \frac{\sum_i in_i}{n} = 0.7$.

(2)重新分组,

H_0: 服从参数为0.7的Poisson分布.

军团死亡人数	0	1	2	3	≥ 4
观察数ν_i	144	91	32	11	2
理论频数np_i	139.04	97.33	34.07	7.95	1.61

$\chi^2=\sum_{i=1}^{5}\frac{(np_i-\nu_i)^2}{np_i}=1.98$, $\chi^2\le\chi_3^2(0.05)=7.81$,无法拒绝$H_0$.

五.

记红光反应时间为X_i, 绿光反应时间为Y_i, $Z_i=X_i-Y_i$,

Z	-0.13	-0.09	-0.17	0.07	-0.03	-0.05	0	-0.10

H_0:$\mu\ge 0$, $H_1:\mu<0$

$\bar{Z}=-0.0625$, $S^2=0.00585$, $T=\frac{\sqrt{n}(\bar{Z}-\mu_0)}{S}=\frac{\sqrt{8}(-0.0625)}{\sqrt{0.00585}}=-2.31<-t_7(0.05)=-1.85$.

拒绝H_0, 红光反应速度快.

六

(1)

$$r=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-(\sum x)^2}\sqrt{n\sum y^2-(\sum y)^2}}=0.932.$$

(2)

$$\hat{b}=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-(\sum x)^2}=3.654,$$

$$\hat{a}=\bar{Y}-\hat{b}\bar{X}=81.738.$$

日本汇率每增长1, 汽车出口数量增加3.654万辆.

(3) 提出假设

$H_0:b=0, H_1:b\ne 0$

$$F=\frac{SSR/1}{SSE/(n-2)}=\frac{59179.84}{8933.76/8}=52.99>F_{1,8}(0.05)=5.317.$$

拒绝H_0, 模型显著.

(4) $\hat{y}_0=81.738+3.654*104.8=464.677$,

$Se=\sqrt{\frac{SSE}{n-2}}=33.417$,

预测区间为

$$\hat{y}_0\pm t_{0.025}(n-2)Se\sqrt{1+\frac{1}{n}+\frac{(x_0-\bar{X})^2}{\sum(X_i-\bar{X})^2}}$$

[379.03,550.33].