清华大学试题专用纸

考试课程：概率论与数理统计　考试时间：2021年1月10日19：00—21：00

姓名____________学号 **20**__________班级_______**.**

一 填空题（每空3分，共30分；答案均写在试卷上，注意标清题号）

1.线上内容中提到过的贝特朗奇论(Bertrand’s paradox)中，不同理解下得到的概率值分别是__________。

2. 将一枚均匀的硬币独立地抛掷16次，正面出现次数记为X，希望 对概率$P(6\le X\le 10)$做尽可能准确的估计，用中心极限定理的结果为____________________，用切比雪夫不等式的结果为________________________。

3. (X,Y)的联合密度函数$p(x,y)=e^{-x-y}I_{x>0,y>0}$，$Z=\begin{cases}Y, & 若X\ge Y\\ 2X, & 若X<Y\end{cases}$，则$E(Z^2)=$__________________。

4. 二维随机变量$(X,Y)\sim N(0,0,4,4,0.5)$，则$E(X^2\mid X+Y=4)=$__________________________。

5. 设$X_1,X_2,\ldots,X_8$为相互独立的$N(0,1)$随机变量，则$P\left(\dfrac{X_1-X_2}{|X_1+X_2+\cdots+X_8|}\ge \underline{\qquad\qquad\qquad}\right)=0.05$。

6. 设$X_1,\ldots,X_{100}$是来自总体$N(\mu,4)$的样本。对期望进行假设检验，$H_0:\mu=0$，$H_1:\mu>0$，若取拒绝域为$\{(x_1,\ldots,x_{100}):\bar{x}>0.2\}$，则此检验犯第一类错误的概率为____________________________，当$\mu=0.6$时，此检验犯第二类错误的概率为_______________________，$\bar{x}=0.2$的p值为__________________________。

7. 对均匀总体$U(0,\theta)$做假设检验，原假设与备择假设分别为$H_0:\theta=5$，$H_1:\theta<5$，以$x_{(n)}=\max\{x_1,\cdots,x_n\}$为检验统计量，显著性水平$\alpha=0.064$，若样本容量$n=3$，则拒绝域为______________________________。

二.（10分）学生在做有4个选项的选择题时，如果他知道正确答案就一定可以答对，如果不知道问题的正确答案，就等可能地随机猜1个选项。假设学生知道正确答案的概率是0.6，不知道正确答案的概率是0.4。

（1）求答对的概率，（2）现从卷面上看题是答对了，求学生知道正确答案的概率。

三.（10分）设随机变量$X\sim N(0,2)$，$Y=|X|$，求随机变量Y的分布函数、密度函数、期望和方差。

四（10分）X,Y均服从参数为1的指数分布，且相互独立。令$U=\max(X,Y)$，$V=\min(X,Y)$

（1）分别求U、V的密度函数，（2）计算U、V的相关系数。

五.（10分）泊松分布总体$X\sim P(\lambda)$，$X_1,\cdots,X_n$为来自该总体的样本，

（1）用最大似然估计法给出总体期望的估计量，（2）试给出参数λ^2的无偏估计量。

六.（10分）设总体$X\sim Ge(p)$，$X_1,X_2,\cdots,X_n$为来自该总体的简单随机样本。定义$Y=\begin{cases}1, & X_1=1\\ 0, & X_1>1\end{cases}$

（1）计算$E\left(Y\middle|X_1+X_2+\cdots+X_n\right)$

（2）判断Y与$E\left(Y\middle|X_1+X_2+\cdots+X_n\right)$是否为参数$p$的无偏估计量？若不是，可否进行无偏校正。

七.（15分）$X_1,X_2,\cdots,X_n$是来自总体$X\sim N\left(\mu_1,\sigma^2\right)$的样本，$Y_1,Y_2,\cdots,Y_m$是来自总体$Y\sim N\left(\mu_2,c\cdot\sigma^2\right)$的样本，$n=4,m=6,c=4.5,\bar{x}=52,\bar{y}=47,s_X^2=30,s_Y^2=99$。

（1）给出参数$\mu_1-\mu_2$的矩估计量，并计算该估计量的期望和方差；

（2）若已知$\sigma^2=25$，给出参数$\mu_1-\mu_2$的95%置信水平的双侧置信区间；

（3）若σ^2未知，给出参数$\mu_1-\mu_2$的90%置信水平的单侧置信区间的置信下界。

八.（5 分）给定n个不同的数$x_1,\cdots,x_n$，设$y_1,\cdots,y_n$是$x_1,\cdots,x_n$按照升序排列后的结果，即$y_1<y_2<\cdots<y_n$。用随机化快速排序算法对$x_1,\cdots,x_n$进行排序，每一次都是从所有可能的元素中独立且均匀地选取基准元素，对$1\le i<j\le n$，定义随机变量X_{ij}，如果在算法过程中y_i与y_j进行了比较$X_{ij}=1$，否则$X_{ij}=0$。求$E\left(X_{ij}\right)$。

备注 1. 本考卷的样本均为简单随机样本，样本均值$\bar{X}=\frac{X_1+\cdots+X_n}{n}$，样本方差为$S^2=\frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2$

备注 2. 正态总体的样本均值和样本方差相互独立，且$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，其中$n$为样本容量

备注 3. 泊松分布随机变量$X\sim P(\lambda)$的分布列 $P(X=k)=e^{-\lambda}\lambda^k/k!$

备注 4. 指数分布随机变量$X\sim Exp(\lambda)$，$F_X(x)=\left(1-e^{-\lambda x}\right)I_{x>0}$，$E(X)=1/\lambda, Var(X)=1/\lambda^2$

备注 5. 第三题解答中标准正态随机变量的分布函数和密度函数分别可用$\Phi(x)$和$\varphi(x)$表示

备注 6. 分布函数和分位数（题目解答要严格按照下面给出数值进行计算）

$\Phi(1.28)=0.9,\quad \Phi(1.44)=0.925,\quad \Phi(1.65)=0.95,\quad \Phi(1.96)=0.975,$
$\Phi(1)=0.84,\quad \Phi(1.25)=0.89,\quad \Phi(1.5)=0.93,\quad \Phi(1.75)=0.96,\quad \Phi(2)=0.98,\quad \Phi(3)=0.999$

$P(t(1)>3.10)=0.1, P(t(2)>1.89)=0.1, P(t(8)>1.40)=0.1, P(t(9)>1.38)=0.1, P(t(10)>1.37)=0.1$

$P(t(1)>6.32)=0.05, P(t(2)>2.92)=0.05, P(t(8)>1.86)=0.05, P(t(9)>1.83)=0.05, P(t(10)>1.81)=0.05$

$P(F(1,1)>40)=0.1, P(F(1,2)>8.5)=0.1, P(F(1,5)>4.1)=0.1, P(F(2,1)>50)=0.1, P(F(5,1)>57)=0.1$

$P(F(1,1)>161)=0.05, P(F(1,2)>19)=0.05, P(F(1,5)>6.6)=0.05, P(F(2,1)>200)=0.05, P(F(5,1)>230)=0.05$