清华大学本科生考试试题专用纸

考试课程：概率论与数理统计（A 卷）　　　　考试日期：　2023 年 1 月 3 日下午

班级 ____________　学号 ________________　姓名 _______________　任课教师姓名 _______________

试卷说明：

1、本试卷有三页（含附录页），共 9 道大题.

2、所有题目都要答在答题纸上且标明题号，首页答题纸上注明 A/B 卷.

3、考试结束时请将答卷拍照合并成一个 pdf 文件上传至网络学堂作业栏（限时 20 分钟）.

---

1. （24 分）填空题

（1）　设随机事件 A 和 B 独立，$P(A)=\dfrac{1}{3}$，$P(A+B)=\dfrac{1}{2}$，则 $P(B)=(\quad)$.

（2）　随机变量 X 的概率密度函数 $f(x)=\begin{cases} a\,x, & 0\le x\le 2 \\ 0, & \text{其它} \end{cases}$，则 $P(1<X<2)=(\quad)$.

（3）　随机变量 X 的分布列为 $P(X=k)=2a\cdot 0.8^{k-1}$（$k=1,2,\cdots$），则常数 $a=(\quad)$.

（4）　设 $X\sim N(1,9)$，$Y\sim N(-1,4)$，若 $P(X>3)=P(Y\le \dfrac{a}{3})$，则 $a=(\quad)$

（5）　已知随机变量 $X\sim U(0,2)$，则 $E(\min\{X,1\})=(\quad)$.

（6）　随机向量 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，其中 $\mu_1=-3$，$\mu_2=0$，$\sigma_1^2=9$，$\sigma_2^2=4$，$\rho=0.5$，则 $\mathrm{Var}(X+\dfrac{Y}{2})=(\quad)$.

（7）　设 $X_1,\cdots,X_n$ 是来自二项分布总体 X 的随机样本，$X\sim B(100,\dfrac{1}{5})$，若样本容量 $n=8$，则 $E\left(\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2\right)=(\quad)$.

（8）　设 $X_1,\cdots,X_n$ 为来自正态总体 $N(\mu,4)$ 的随机样本，若使 $P\left(\left|\bar{X}-\mu\right|<\dfrac{1}{2}\right)\ge 0.96$ 成立，则样本容量 n 至少要达到 $(\quad)$.

2. （8 分）某厂有甲、乙、丙三车间生产同一种产品，产量分别占总产量的 60%，30% 和 10%，各车间的次品率分别是 3%，5% 和 7%．试求：

（1）　在该厂产品中任取一件，恰为次品的概率.

（2）　若发现一件产品为次品，该次品来自甲车间的概率.

3. （10 分）随机向量 (X,Y) 的联合概率密度函数为 $f(x,y)=\begin{cases} \dfrac{x}{4}, & 0\le x,y\le 2 \\ 0, & \text{其它} \end{cases}$

（1）　求 X,Y 的边际密度函数 $f_X(x)$ 和 $f_Y(y)$.

（2） 验证随机变量 X,Y 相互独立.

（3） 计算相关系数 $\mathrm{Corr}(X+Y,X-Y)$.

4. （8 分）设 $X_1,\cdots,X_n$ 是来自区间 $[\theta,4]$ 上均匀分布总体的随机样本.

（1） 求参数 θ 的矩估计量 $\hat{\theta}$ 和极大似然估计量 θ^*.

（2） 当 $n=5$ 时，一组样本观测值为 $1,4,3,1,1$，求样本均值与样本方差.

5. （4 分）某宴会提供一坛 5 升装的白酒，假定所有饮酒者每次倒酒的量独立同分布，其期望为 100 毫升，标准差为 7 毫升，那么倒了 49 次后该坛酒仍有剩余的概率约为多少？（需给出计算过程）

6. （6 分）随机变量 $X_1,\cdots,X_n$ 独立同分布且 $X_1+\cdots+X_n\sim N(0,n)$，证明：$X_i$（$i=1,\cdots,n$）皆服从标准正态分布.

7. （22 分）设随机样本 $X_1,\cdots,X_n$ 来自具有概率密度函数为 $f(x)=\begin{cases}\theta x^{\theta-1}, & 0<x<1\\ 0, & \text{其它}\end{cases}$ 的分布，其中 $\theta>0$ 是未知参数.

（1） 求 θ 的极大似然估计量 θ^*.

（2） 判断 $\dfrac{1}{\theta^*}$ 是否是 $\dfrac{1}{\theta}$ 的无偏估计并给出相应理由.

（3） 当 $n=4$ 时，一组样本观测值为 $\dfrac{1}{2},\dfrac{1}{3},\dfrac{1}{5},\dfrac{1}{5}$，利用 Fisher 信息量和这组观测值给出 θ^* 的标准误差的估计.

（4） 利用 θ^* 和（3）中样本观测值给出 θ 的满足 95% 置信要求的一个具体置信区间.

（5） 考虑假设 H_0：$\theta=1$ v.s. H_1：$\theta\neq1$，利用（3）中样本观测值, 在检验水平 $\alpha=0.05$ 下进行似然比检验.

8. （10 分）假设有两个正态总体 A 和 B. 从总体 A 随机抽取 10 个数据，得到 6 个 100 和 4 个 99，从总体 B 随机抽取 2 个数据，得到的是-100 和-200. 请对两个总体均值分别进行单尾 t 分布假设检验：$H_0:\mu=100$ vs $H_1:\mu<100$，取检验水平 $\alpha=0.05$.

（1） 你的结论如何？

（2） 你对题目从提出到结论有何看法？

9. （8 分）设随机变量 X 的分布列为 $P(X=k|\theta)=\theta(1-\theta)^k$（$k=0,1,2,\cdots$），参数 θ 的先验分布是 $(0,1)$ 上的均匀分布. 对 X 作三次独立观测，得到观测值为 $x_1=2$，$x_2=3$，$x_3=5$.

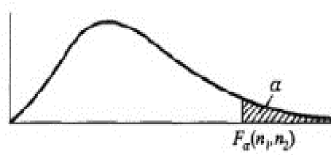
（1） 求 θ 的后验分布概率密度函数.

（2） 求 θ 的后验均值估计.

附录 1：标准正态累积分布函数 $\Phi(x)$

x	0.85	0.92	1.00	1.64	1.80	1.96	2.05	2.33	2.58
$\Phi(x)$	0.802	0.821	0.841	0.950	0.964	0.975	0.980	0.990	0.995

附录 2：记号约定：$0<\alpha<1$, 自由度为 n 的 t 分布为 $t(n)$，自由度为 n 的卡方分布为 $\chi^2(n)$，自由度为 (m,n) 的 F 分布为 $F(m,n)$，上 α 分位点分别为 $t_\alpha(n)$， $\chi^2_\alpha(n)$， $F_\alpha(m,n)$.



$t_{0.05}(1)=6.314$ $t_{0.025}(1)=12.706$ $t_{0.05}(2)=2.970$ $t_{0.025}(2)=4.303$

$t_{0.05}(9)=1.833$ $t_{0.025}(9)=2.262$ $t_{0.05}(10)=1.812$ $t_{0.025}(10)=2.228$

$\chi^2_{0.975}(1)=0.001$ $\chi^2_{0.95}(1)=0.004$ $\chi^2_{0.05}(1)=3.84$ $\chi^2_{0.025}(1)=5.02$

$\chi^2_{0.975}(2)=0.05$ $\chi^2_{0.95}(2)=0.10$ $\chi^2_{0.05}(2)=5.99$ $\chi^2_{0.025}(2)=7.38$

$\chi^2_{0.975}(5)=0.83$ $\chi^2_{0.95}(5)=1.15$ $\chi^2_{0.05}(5)=11.07$ $\chi^2_{0.025}(5)=12.83$

$F_{0.025}(4,3)=15.10$ $F_{0.975}(4,3)=0.10$ $F_{0.05}(4,3)=9.12$ $F_{0.95}(4,3)=0.15$

$F_{0.025}(3,4)=9.98$ $F_{0.975}(3,4)=0.07$ $F_{0.05}(3,4)=6.59$ $F_{0.95}(3,4)=0.11$

附录 3： $e^{-1}\approx 0.3679$， $e^{-2}\approx 0.135$， $e^{-3}\approx 0.050$， $e^{-4}\approx 0.018$， $e^{-5}\approx 0.007$

附录 4： $\log 2\approx 0.7$， $\log 3\approx 1.1$， $\log 5\approx 1.6$

附录 5：分布 $\beta(a,b)$ 的概率密度函数是 $f(x)=\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},\quad 0<x<1.$