清华大学本科生考试试题专用纸

卷 **A**　　考试课程　概率论与数理统计　　　　　2007 年 6 月 25 日

记号：~服从；　:= 记为；　*iid* 独立同分布；　*df* 分布函数；　*pdf* 分布密度函数；

rv 随机变量；$r\vec{v}$ 随机向量；　B(n, *p*)二项分布；　P(λ) Poisson 分布；

Ge(*p*) 几何分布；　Ex(λ) 指数分布；　U(a,b)均匀分布；　N(μ,σ^2)正态分布。

一(36 分)填空与判断正误（正确时填√，错误时填×；填入的分布必须带参数）

1. 设事件 A，B 满足 0<P(B)<1，$P(A|B)=P(A|\overline{B})$，则必有事件 A 与 B 相互独立（　　）；此时如令

$$X=\begin{cases}1 & 如果A发生\\ -1 & 反之\end{cases},\quad Y=\begin{cases}1 & 如果B发生\\ -1 & 反之\end{cases}$$

则一定有$P(X=-1|Y=-1)=P(X=-1)$　（　　）

2. 已知$P(\overline{A})=0.3$，$P(B)=0.4$，$P(\overline{B}A)=0.5$，则$P\big(B\big|(A\cup\overline{B})\big)$=________，此时 A 和 B 独立（　　）

3. 有一批同型号产品，已知其中由一廠生产的占 30%，二廠生产的占 50%，三廠生产的占 20%，又知这三个廠产品的次品率分别为 2%、1%和 1%。问从这批产品中任取一件是次品的概率等于___________

4. 设(X_1,X_2)~$N(0,0,\sigma^2,\sigma^2,0.5)$，令$Y_1=X_1-X_2$，$Y_2=X_1+X_2$，则$P(X_1<X_2)$=___________；

Y_1和Y_2不独立（　　），又如果$\sigma^2=1$，此时$f_{X_2|X_1}(1|0)$=___________

5. 设 rv X 和 Y *iid*，如 X~$U_{[0,2]}$，则 P{max{X,Y}≤1}=___________

6. 设总体 X 的 *pdf* 为$f_X(x)=\frac{1}{2}e^{-|x|},-\infty<x<\infty$，$X_1,X_2,\ldots,X_n$ 为总体的简单随机样本，样本均值为$\overline{X}$，则 $D\overline{X}$=____________

7. 设某产品的寿命 X 的 *pdf* 为$f(x)=\begin{cases}0 & x\le 1\\ \lambda e^{-\lambda(x-1)} & x>1\end{cases}$，这里$\lambda$是未知参数。又设

$X_1,X_2,\ldots,X_n$ 为该总体 X 的简单样本，则λ的矩估计量$\hat{\lambda}_M$=__________，而极大似然估计量$\hat{\lambda}_L$=_________。

二(**8** 分) 设 rv $X_k=2\sin(\omega_0 k+\Theta)$，k=1,2，其中$\omega_0$是常数，rv$\Theta\sim U_{[-\pi,\pi]}$，求

(1) EX_k　　(2) ω_0取何值时 X_1 和 X_2 不相关？

三(**8** 分) 设 rv X 和 Y 独立，且 X~U[0,2]，Y~U[0,1]，求 U=X+Y 的分布.求 DZ，其中 rv Z 如下定义：$Z=\begin{cases}1 & X>Y\\ -1 & X\le Y\end{cases}$.

四(**10** 分)设 $X_1,X_2,\ldots,X_n$，n>2，独立同分布，~N(μ,σ^2)

(1) 求 X_1-X_2 与 X_1 的相关系数 r

(2) 写出 $X_1-\overline{X}$ 的 *pdf*，并求 D($|X_1-\overline{X}|$)，其中 $\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$

五(**8** 分) 设总体 X 的 pdf 为 $f(x,\theta)=\begin{cases}\dfrac{1}{2\theta} & 0<x<\theta\\ \dfrac{1}{2(1-\theta)} & \theta\le x<1\\ 0 & \text{else}\end{cases}$。$X_1,X_2,\ldots,X_n$ 是来自总体 X 的简单样本。求参数θ的矩估计量$\hat{\theta}$；它是否为θ的无偏估计？说明理由。

六(**10** 分) 在计算机上作大型科学计算，需对十进制的x_j的小数点后第 6 位作四舍五入，得到x_j的近似数y_j，则认为随机误差$\varepsilon_j=x_j-y_j$在区间(−0. 5×10^{-5}, 0. 5×10^{-5})内均匀分布，累积误差为$\eta_n=\sum_{j=1}^{n}\varepsilon_j$。试求

(1)用切贝雪夫不等式估计，当 n=10000 时给出$|\eta_n|$不超过 0. 0005 的概率的上界；

(2)用中心极限定理，n=10000 时，以 99. 7%以上的把握给出$|\eta_n|$的近似估计(估计上界)。

七(**20** 分) 设冶炼厂的某项污染指标 X~ N(μ,σ^2)，σ^2 未知，$x_1,x_2,\ldots,x_9$ 是此正态总体的大小为 9 的简单样本的观测值，测得$\overline{x}=3.64$，s^2=0.64.

(1) 求 $X_1,X_2,\ldots,X_9$ 中恰有 2 个小于该总体 X 的期望的概率 p

(2) 问 μ 是否明显低于 μ_0=4.00（取显著性水平 α=0.05）？

附表　$z_{0.05}$=1.64,　$z_{0.025}$=1.96

$\chi^2_\alpha(n)$	n=8　n=9	$\chi^2_\alpha(n)$	n=8　n=9	$t_\alpha(n)$	n=8　n=9
α=0.95	2.733　3.325	α=0.05	15.507　16.919	α=0.05	1.8595　1.8331
α=0.975	2.180　2.700	α=0.025	17.535　19.023	α=0.025	2.3060　2.2622

参考答案

一．1. $\underline{\surd}$，$\underline{\surd}$

2. $\underline{0.25}$，$\underline{\times}$

3. $\underline{1.3\%}$

4. $\underline{1/2}$，$\underline{\times}$，$\sqrt{\frac{2}{3\pi}}e^{\frac{-2}{3}}$

5. $\underline{1/4}$

6. $\underline{2/n}$

7. $\frac{1}{\bar{X}-1}$，$\frac{1}{\bar{X}-1}$

二．$EX_k=0$ ， $\omega_0=2k\pi+\frac{\pi}{2}$或$2k\pi+2\pi-\frac{\pi}{2}$

三．(1) $F_U(u)=\begin{cases} 0 & u\le 0 \\ \frac{1}{4}u^2 & 0<u\le 1 \\ \frac{1}{2}u-\frac{1}{4} & 1<u\le 2 \\ 1-\frac{1}{2}(3-u)^2 & 2<u\le 3 \\ 1 & u\ge 3 \end{cases}$

(2) $EZ=P(Z=1)-P(Z=-1)=\frac{3}{4}-\frac{1}{3}=\frac{1}{2}$，$EZ^2=1$

故$DZ=EZ^2-(EZ)^2=1-\frac{1}{4}=\frac{3}{4}$

四．(1) $r=\frac{\sqrt{2}}{2}$

(2) $X_1-\bar{X}\sim N\left(0,\frac{n-1}{n}\sigma^2\right)$，故$D(|X_1-\bar{X}|)=\frac{n-1}{n}\sigma^2$

五．$\hat{\theta}=\frac{4\bar{X}-1}{2}$，是无偏估计

六．(1) $P(|\eta_n|<0.0005)\ge\frac{2}{3}$

(2) 由中心极限定理：$\zeta_n\stackrel{\text{def}}{=}\frac{\sum_{j=1}^{n}(\varepsilon_j-0)}{\sqrt{n}D\varepsilon_j}\sim N(0,1)$，所以$\eta_n\sim N(0,\sigma^2)$，其中$\sigma=\frac{1}{12}\times 10^{-8}$

设上界为 c，

则$P(|\eta_n|<c)=P(-c<\eta_n<c)=99.7\%$

$\therefore c=3\sigma=2.5\times 10^{-7}$

七．(1) $C_9^2\left(\frac{1}{2}\right)^9=\frac{36}{2^9}=\frac{9}{2^7}$

(2) 是。