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；　*$\vec{rv}$* 随机向量；　$B(n,p)$ 二项分布；　$P(\lambda)$ Poisson 分布；

$Ge(p)$ 几何分布；　$Ex(\lambda)$ 指数分布；$U(a,b)$ 均匀分布；$N(\mu,\sigma^2)$ 正态分布。

一（36 分）．填空与判正误(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

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

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 $P(B|(A \cup \overline{B})) =$ 0.25．此时 A 与 B 独立．(×)

3. 有一批同型号产品，已知其中由一厂生产的占 30%，二厂生产的占 50%，三厂生产的占 20%，又知这三个厂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 1.3%

4. 设 $(X_1, X_2) \sim N(0, 0, \sigma^2, \sigma^2, 0.5)$，令 $Y_1 = X_1 - X_2$，$Y_2 = X_1 + X_2$．

则 $P(X_1 < X_2) =$ 0.5；Y_1 和 Y_2 不独立 (×)．又如果 $\sigma^2 = 1$ 此时 $f_{X_2|X_1}(1|0) =$ $\frac{\sqrt{2}}{\sqrt{3\pi}} e^{-\frac{1}{3}}$ 4．

5. 设 *rv* X 与 Y *iid*．如 $X \sim U_{[0,2]}$，则 $P\{\max\{X, Y\} \le 1\} =$ $\frac{1}{4}$．

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 $D\overline{X} =$ $\frac{2}{n}$．

7. 设某产品的寿命 X 的 *pdf* 为

$$f(x) = \begin{cases} 0, & \text{当 } x \le 1, \\ \lambda e^{-\lambda(x-1)}, & \text{当 } x > 1. \end{cases}$$

这里 λ 是未知参数，又设 $X_1, X_2, \cdots, X_n$ 为该总体 X 的简单样本，则 λ 的矩估计量 $\hat{\lambda}_M =$ $1/(\overline{X}-1)$ 而极大似然估计量 $\hat{\lambda}_L =$ $1/(\overline{X}-1)$．

二（8分）. 设 $rv\ X_k = 2\sin(\omega_0 k + \Theta)$，$k = 1, 2$，其中 ω_0 是常数，$rv\ \Theta \sim U_{[-\pi\ \pi]}$．求

(1) EX_k； 0 (2) ω_0 取何值时 X_1 和 X_2 不相关？ $\frac{\pi}{2}+k\pi$

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

三（8分）设 $rv\ X$ 和 Y 独立，且 $X \sim U[0, 2]$，$Y \sim U[0, 1]$．求 $U = X + Y$ 的分布．求 DZ $\frac{3}{4}$ 其中．又 $rv\ Z$ 如下定义：$Z = \begin{cases} 1 & X > Y \\ -1 & X \le Y \end{cases}$.

四（10分）设 $X_1, X_2, \cdots, X_n$，$n>2$，独立同分布，~ $N(\mu, \sigma^2)$，.

(1) 求 $X_1 - X_2$ 与 X_1 的相关系数 r $\frac{\sqrt{2}}{2}$

(2) 写出 $X_1 - \bar{X}$ 的 pdf，并求 $D(|X_1 - \bar{X}|)$，其中 $\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ $E|Y| = \sqrt{\frac{2(n-1)}{n\pi}}\,\sigma$，$E|Y|^2 = \frac{n-1}{n}\sigma^2$

$N(0, \frac{n-1}{n}\sigma^2)$ $D|Y| = \frac{\pi-2}{\pi}\cdot\frac{n-1}{n}\sigma^2$

五（8分）. 设总体 X 的 pdf 为

$$f(x,\theta) = \begin{cases} \frac{1}{2\theta}, & 0 < x < \theta \\ \frac{1}{2(1-\theta)} & \theta \le x < 1, \\ 0 & 其它 \end{cases}$$

$\hat{\theta}_M = \frac{1}{2}(4\bar{X}-1)$ $E\hat{\theta}_M = \theta$

$X_1, X_2 \cdots X_n$ 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 的无偏估计量，并说明理由。

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 $(-0.5\times10^{-5},\ 0.5\times10^{-5})$ 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，当 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. $x_1, x_2, \cdots, x_9$ 是此正态总体的大小为 9 的简单样本的观测值，测得 $\bar{x} = 3.64$，$s^2 = 0.64$.

（1）求 $X_1, X_2, \cdots, X_9$ 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