清华大学本科生考试试题专用纸

考试课程　　　　　　　　　　　　　　　年　月　日

卷 **B** 《概率统计考试题》(04.06) 班级_____ 学号______ 姓名________

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

一（33 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 $F_i(x)$和 $f_i(x)$分别是 X_i 的 *df* 和 *pdf*, $i=1,2$. 则 $F_1(x)F_2(x)$是分布函数 （ √ ）
而 $f_1(x)f_2(x)$ 不是密度函数．（ √ ）

2．设两个相互独立的事件 A 和 B 都不发生的概率为 1/9, 事件 A 发生 B 不发生的概率与事件 A 不发生 B 发生的概率相等, 则 $P(A)=$ $\underline{\ \frac{2}{3}\ }$.

3．设$(X_1,X_2)\sim N(0,0,\sigma^2,\sigma^2,0.5)$，令$Y_1=X_1-\frac{1}{2}X_2$， $Y_2=\frac{1}{2}X_1-X_2$.
则 Y_1和 Y_2都有正态分布且分布参数相同 (√), 但不独立 (√).

4．设 *rv* $X\sim t(n)$ $(n>1)$, $Y=\dfrac{1}{X^2}$, 则 $Y\sim$ ($F(n,1)$).

5．设 *rv* X 和 Y 都服从~ N(0, 1), 则 $X+Y$ 服从正态分布 ($N(0,2)$)；
而 X^2和Y^2都服从χ^2分布 ($\chi^2(1)$)

6．设总体 X 的方差存在, $X_1, X_2, \ldots X_{n-1}$ 是其简单样本， $Y_k=\frac{1}{k}\sum_{i=1}^{k}X_i$， $k=n, n+1$，
则用 Y_n 及 Y_{n+1} 估计 EX 时，Y_{n+1}比 Y_n 有效 (√).

7．设 $X_1, X_2, \cdots, X_n$ 为 *iid*, $\sim N(0,\sigma^2)$．令 $\bar{X}=\dfrac{1}{n}\sum_{j=1}^{n}X_j$， $Y_j=X_j-\bar{X}$, $j=1,2$.
则 Y_1与 Y_2的相关系数 $\underline{\ -\frac{1}{n-1}\ }$.

8．如果总体 $X\sim P(\lambda)$, 则当 $n\to\infty$ 时, $Y_n=\dfrac{1}{n}\sum_{i=1}^{n}X_i^2$ 几乎处处收敛于 $\underline{\ \frac{1}{\lambda}+\frac{1}{\lambda^2}\ }$.

二（8分）设 A、B 是两个事件，且 $0<P(A)<1$，$P(B\mid A)=P(B\mid\bar{A})$，问 A、B 是否独立？请说明理由.

$\frac{P(AB)}{P(A)} = \frac{P(B\bar{A})}{P(\bar{A})} = \frac{\quad}{1-P(A)}$

三（10分）. 设 $X_1, X_2, \ldots, X_n \ldots$ 为 *iid*.

$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) = P(B|\bar{A})$

(1) 如果 $X_1 \sim$ 0-1 分布，参数为 p. 试对固定正整数 $k\le n$，求如下概率：

$P(A) = \frac{P(AB)}{P(B|A)}$

$P(\sum_{i=1}^{n} X_i = k)$、$P(\sum_{i=1}^{n} X_i = k,\ X_n = 0)$ 及 $P(\min\{n:\ X_n \neq 0, n=1,2,\ldots\} = k)$.

(2) 如果 $X_1 \sim N(0,\sigma^2)$，试求 $\left(\frac{X_1+X_2}{X_1-X_2}\right)^2$ 的分布.

$\therefore P(A)P(B) = P(AB)$

$f(x) = \begin{cases} \lambda e^{-\lambda x} & x>0 \\ 0 & 其他 \end{cases}$

四（10分）设某网络服务器首次失效时间(寿命)$X \sim \mathrm{Ex}(\lambda)$，现随机购得 4 台中.

$Y \sim B(4, P)$

(1) 令 Y : 4 台中寿命小于此类服务器期望寿命的服务器台数，求 Y 的最可能值.

$P = \int_0^{\lambda} \lambda e^{-\lambda x} = 1-e^{-\lambda^2}$

(2) 事件 A : 4 台中最早失效时间大于此类服务器期望寿命. 求 A 的概率.

$\min(x_1\ x_2\ x_3\ x_4) > \lambda \quad P = \left(\int_{\lambda}^{+\infty} f(x)dx\right)^4 = e^{-4\lambda^2}$

$\therefore E(Y) = np = 4(1-e^{-\lambda^2})$

五（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，则误差 $\varepsilon_j = x_j - y_j$ 在区间 $(-0.5\times10^{-5},\ 0.5\times10^{-5})$ 内随机取值. 视 ε_j 为从区间内的均匀分布随机变量，令累积误差 $\eta_n = \sum_{j=1}^{n}\varepsilon_j$，试利用中心极限定理，当 n=10,000 时有 99.7%以上的把握给出 $|\eta_n|$ 的近似估计的上界.　注 $2\Phi(3)-1=0.9974$.

$M = 5.087\times10^{-2}$

六（9分）设(X,Y)的 *pdf* 为 $f(x,y) = c\cdot\exp\{-n(x+y)\}\, I(0<x<y<+\infty)$，其中 n 为已知正整数，c 为待定常数.

$f(x,y) = \begin{cases} ce^{-n(x+y)} & 0<x<y<+\infty \\ & 其他 \end{cases}$

(1). 求 c；　(2). 求条件密度 $f_{Y|X}(y|x)$；　(3). X 与 Y 是否独立，为什么？

$c\int_0^{+\infty} e^{-nx}dx \int_x^{+\infty} e^{-ny}dy = c\int_0^{+\infty}\frac{1}{n}e^{-2nx}dx = \frac{c}{2n^2} = 1 \quad \therefore c = 2n^2$

七（20分）设某糖厂用自动包装机集箱外运糖果. 某日开工后在生产线上抽测 9 箱，得数据 99.3, 98.7, 100.5, 101.2, 98.3, 99.7, 99.5, 102.1, 100.5（kg）. 认为包装的每箱糖重为正态分布 $N(\mu, \sigma^2)$，参数 μ 和 σ^2 都未知. 取 $\alpha = 0.05$.

(1) 试写出 μ 的最大似然估计值和标准差 σ 的置信度为 0.95 的双侧置信区间.

(2) 如果规定包装机每箱装糖重量为 100kg 且方差未知. 问由抽测数据能否认为生产线上每箱装糖重量不低于规定重量？

附表　$z_{0.05} = 1.64$,　$z_{0.025} = 1.96$

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