<u>清 华 大 学 本 科 生 考 试 试 题 专 用 纸</u>

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；

$\boldsymbol{rv}$ 随机变量；　$\boldsymbol{r\vec{v}}$ 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（33 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$F_i(x)$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)=________________。

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$\boldsymbol{rv}$ $X\sim t(n)$　$(n>1)$，$Y=\frac{1}{X^2}$，则$Y\sim$（　　　　　　　）。

5. 设设$\boldsymbol{rv}$ X和Y都服从$\sim N(0,1)$，则$X+Y$服从正态分布（　　　　）；而X^2和Y^2都服从χ^2分布（　　　　）

6. 设总体 X 的方差存在,$X_1,X_2,\dots,X_{n-1}$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$X_1,X_2,\dots,X_{n-1}$为 *iid*，$\sim N(0,\sigma^2)$，令$\bar{X}=\frac{1}{n}\sum_{j=1}^{n}X_j$ ，$Y_j=X_j-\bar{X}$，*j*=1,2。则Y_1与Y_2的相关系数_______________

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

二（8 分）设 A、B 是两个时间，且 0<P(A)<1，P(B|A)=P(B|Ā)，问 A、B 是否独立？请说明理由。

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

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

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

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

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

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

(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。

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，得到的 x_j 近似数为 y_j，则误差 $\varepsilon_j = x_j - y_j$ 在区间$(-0.5 \times 10^{-5}, 0.5 \times 10^{-5})$内随机取值，视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 $2\Phi(3)-1=0.9974$.**

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 为已知整数，c 为待定常数.

七（20 分）设某糖厂用自动包装机集箱外运糖果，某日开工后再生产线上抽测 9 箱，得数据 99.3，100.5，101.2，98.3，99.7，99.5，102.1，100.5（kg）.认为包装的每箱重为正态分布 N(μ,σ^2)，参数 μ 和 σ^2 都未知，取 α=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) 如果规定包装机没箱装糖重量为 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

$\chi^2_\alpha(n)$	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. √，√

2. $\frac{2}{3}$

3. √，√

4. F(n,1)

5. √，√

6 √

7. $\frac{1}{n+1}$

8. $\lambda^2+\lambda$

二．解：P(B)=P(B|A)P(A)+P(B|Ā)P(Ā)=P(B|A)[P(A)+P(Ā)]= P(B|A)，所以 A 和 B 独立。

三．(1) $C_n^k p^k(1-p)^{n-k}$，$C_{n-1}^k p^k(1-p)^{n-k}$，$p(1-p)^{k-1}$

(2) $\left(\frac{X_1+X_2}{X_1-X_2}\right)^2 \sim F(1,1)$

五．解: $\mathrm{E}\varepsilon_j=0, \mathrm{D}\varepsilon_j=\frac{1}{12}$，$\mathrm{P}(|\eta_n|<a)=0.997$，

$$\mathrm{P}\left(-a<\Sigma\varepsilon_j<a\right)=\mathrm{P}\left(\frac{-a}{\sqrt{1000\times\frac{1}{12}}}<\frac{\Sigma\varepsilon_j}{\sqrt{n}\sigma}<\frac{a}{\sqrt{1000\times\frac{1}{12}}}\right)=2\Phi\left(\frac{a}{5\sqrt{\frac{10}{3}}}\right)-1$$

$$=0.997$$

$$\therefore a=15\sqrt{\frac{10}{3}}$$

六．(1) $c=2n^2$

(2) $f_{Y|X}(y|1)=\mathrm{e}^{-n(y-1)}$

(3) $2n^2\mathrm{e}^{-n(x+y)}\neq\frac{1}{n^2}\mathrm{e}^{-n(x+y)}$，所以 X、Y 不独立