2016 年秋季学期《概率与数理统计》期中考试试题答案

1. 某芯片制造厂生产的一批芯片的寿命满足正态分布，其中 $\mu=1.4\times10^6,\sigma=3\times10^5$，单位为小时 。
   (a) 某芯片采购商要求 90%以上的芯片寿命至少有 1×10^6 小时，请问该批芯片是否满足要求？ （5 分）
   (b) 随机抽查这批芯片中的 100 片，请问至少有 20 片的寿命小于 1.8×10^6 小时的概率为多少？（10 分）

解：

(a)

$$P(\mathrm{T}\ge10^6)=1-P(\mathrm{T}\le10^6)=1-\Phi(\frac{10^6-1.4\times10^6}{3\times10^5})$$
$$=1-\Phi(-\frac{4}{3})=\Phi(\frac{4}{3})\approx90.88\%$$

(b) $P(\mathrm{T}<1.8*10^6)=\Phi(4/3)=0.9088$

故，100 片芯片中寿命小于 1.8×10^6 的满足二项分布 $(n,p)=100,0.9088$ ，故

$$P(\mathrm{N}\ge20)=\sum_{n=20}^{100}\binom{100}{n}(0.9088)^n(1-0.9088)^{100-n}$$

或者用正态分布去估计二项分布。$\mu=np=90.88,\sigma^2=np(1-p)=8.28$

所以：

$$P(\mathrm{N}\ge20)=P(\mathrm{N}\ge19.5)=\mathrm{p}\{\frac{N-90.88}{2.87}\ge\frac{19.5-90.88}{2.87}\}$$
$$\approx\mathrm{P}\{\mathrm{Z}\ge-24.9\}=1-\Phi(-24.9)\approx1$$

2. 设随机变量 X 的概率密度函数是

$$f(x)=\begin{cases}\dfrac{2x}{\pi^2},0<x<\pi\\0,\quad 其他\end{cases}$$

求 $Y=\sin X$ 的概率密度。（10 分）

解：

X在$(0,\pi)$取值时，$Y=sin\ X$在$(0,1)$取值，故若$y<0$或$y>1$，则$f_Y(y)=0$.若$0\leqslant y\leqslant 1$，Y的分布函数为

$$\begin{aligned}F_Y(y) &= P(Y\leqslant y)=P(0\leqslant Y\leqslant y)=P(0\leqslant sin\ X\leqslant y)\\ &= P((0\leqslant X\leqslant arcsin\ y)\bigcup(\pi-arcsin\ y\leqslant X\leqslant\pi))\\ &= \int_0^{arcsin\ y}\frac{2x}{\pi^2}dx+\int_{\pi-arcsin\ y}^{\pi}\frac{2x}{\pi^2}dx=\frac{1}{\pi^2}(arcsin\ y)^2+1-\frac{1}{\pi^2}(\pi-arcsin\ y)^2\\ &= \frac{2}{\pi}arcsin\ y.\end{aligned}$$

故当$0<y<1$时，$f_Y(y)=\frac{d}{dy}F_Y(y)=\frac{2}{\pi\sqrt{1-y^2}}$. 因此，$Y$的密度函数为

$$f_Y(y)=\begin{cases}\frac{2}{\pi\sqrt{1-y^2}}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

3.　X 的概率密度函数如下：

$$f(x)=\begin{cases}ce^{-2x}, 0<x<\infty\\ 0, \quad x<0\end{cases}$$

(a) 求 c。　（4 分）
(b) 求 $P\{X>2\}$ 的值。　（5 分）
(c) 对任意 a>0，求证：$P\{X\geq a\}\leq\dfrac{E(X)}{a}$　（6 分）

解：
(a) c=2

(b) $P(X>2)=1-F(2)=1-\int_0^2 2e^{-2x}dx=e^{-4}$

(c)
$$E[X]=\int_0^\infty xf(x)dx=\int_0^a xf(x)dx+\int_a^\infty xf(x)dx$$
$$\geq\int_a^\infty xf(x)dx\geq\int_a^\infty af(x)dx=a\int_a^\infty f(x)dx=aP\{X\geq a\}$$

4.　设

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}(-\infty<x<+\infty)$$

$$g(x)=\begin{cases}\cos(x) & |x|<\pi\\ 0 & |x|\geq\pi\end{cases}$$

$$f(x,y)=\varphi(x)\varphi(y)+\frac{1}{2\pi}g(x)g(y),\quad (-\infty<x,y<+\infty)$$

解答下列问题：

(d) $f(x,y)$ 是概率密度函数。(设二维随机变量 (X,Y) 的概率密度函数为 $f(x,y)$)（8 分）

(e) (X,Y) 关于 X ,Y 的边缘分布为正态分布。（5 分）

(f) X 与 Y 的相关系数 $\rho_{XY}=0$ 。（5 分）

解：

(a)

当$|x|<\pi,|y|<\pi$时,

$$\begin{aligned}f(x,y)&=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}+\frac{1}{2\pi}e^{-\pi^2}cos\ xcos\ y\\&=\frac{1}{2\pi}(e^{-\frac{x^2+y^2}{2}}+e^{-\pi^2}cos\ xcos\ y)\geqslant 0\end{aligned}$$

而当$|x|\geqslant\pi,|y|\geqslant\pi$或其他情况时，显然有$f(x,y)\geqslant 0$.且有

$$\begin{aligned}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}+\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}e^{-\pi^2}cos\ xcos\ y\\&=1+0=1.\end{aligned}$$

所以$f(x,y)$是概率密度函数.

(b)

因为

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}(-\infty<x<\infty),$$

同理

$$f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}(-\infty<x<\infty),$$

所以X~$N(0,1),Y$~$N(0,1)$.

(c)

因为

$$
\begin{aligned}
Cov(X,Y) &= E(XY) - E(X)E(Y) = E(XY) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy[\varphi(x)\varphi(y) + \frac{1}{2\pi}e^{-\pi^2}g(x)g(y)]dxdy \\
&= 0,
\end{aligned}
$$

所以$\rho_{XY} = 0.$

5. (a) 若 X 是非负整数随机变量，请证明：

$$E[X] = \sum_{n=1}^{\infty} P\{X \geq n\} = \sum_{n=0}^{\infty} P\{X > n\} \quad （7 分）$$

提示：定义随机变量序列$I_n, n \geq 1,$

$$I_n = \begin{cases} 1, & if\ \ n \leq X \\ 0, & if\ \ n > X \end{cases}$$

再用I_n表示 X 。

(b) 若 X 和Y都是非负整数随机变量，请证明：

$$E[XY] = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} P\{X \geq n, Y \geq m\} \quad （8 分）$$

解：

(a) 设$I_n = \begin{cases} 1, & if\ \ n \leq X \\ 0, & if\ \ n > X \end{cases}$，故$\mathrm{E}[I_k] = \mathrm{P}\{\mathrm{n} \leq \mathrm{X}\}$，所以$\mathrm{X} = \sum_{n=1}^{\infty} I_k$，

故，$\mathrm{E[X]} = \sum_{n=1}^{\infty} P\{X \geq n\} = \sum_{n=0}^{\infty} P\{X > n\}$

(b) 设 $I_n = \begin{cases} 1, & if\ \ n \leq X \\ 0, & if\ \ n > X \end{cases}$， $J_m = \begin{cases} 1, & if\ \ m \leq Y \\ 0, & if\ \ m > Y \end{cases}$， 故 $\mathrm{XY} = \sum_{n=1}^{\infty} I_n \sum_{m=1}^{\infty} J_m = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} I_n J_m$

所以，$\mathrm{E[XY]} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} E[I_n J_m] = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} P\{X \geq n, Y \geq m\}$

6. 证明题：

(a) 设 $X_1, X_2, ..., X_n$ 是 n 个相互独立的随机变量，X_i（$i = 1, 2, \ldots, n$）均服从于 0—1 伯努利分布，即 $P(X_i = 1) = p$，$P(X_i = 0) = 1 - p$。已知随机变量 $X = X_1 + X_2 + \ldots + X_n$，证明：$X$ 的方差$Var(X) = np(1-p)$。（7 分）

(b) 设随机变量 X 服从于二项分布 $X \sim B(n,p)$，试证明当 $n \to \infty$，$p \to 0$，X 近似服从于泊松分布。（10 分）

$$C_n^k p^k (1-p)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda} \text{，其中 } \lambda = np$$

解：

(a)

Solution: Since such a random variable represents the number of successes in n independent trials when each trial has a common probability p of being a success, we may write

$$X = X_1 + \cdots + X_n$$

where the X_i are independent Bernoulli random variables such that

$$X_i = \begin{cases} 1, & \text{if the } i\text{th trial is a success} \\ 0, & \text{otherwise} \end{cases}$$

Hence, from Equation (2.16) we obtain

$$\mathrm{Var}(X) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)$$

But

$$\begin{aligned} \mathrm{Var}(X_i) &= E[X_i^2] - (E[X_i])^2 \\ &= E[X_i] - (E[X_i])^2 \qquad \text{since } X_i^2 = X_i \\ &= p - p^2 \end{aligned}$$

(b)

$$\begin{aligned} P\{X = i\} &= \frac{n!}{(n-i)!\,i!} p^i (1-p)^{n-i} \\ &= \frac{n!}{(n-i)!\,i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i} \end{aligned}$$

Now, for n large and p small

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \qquad \frac{n(n-1)\cdots(n-i+1)}{n^i} \approx 1, \qquad \left(1 - \frac{\lambda}{n}\right)^i \approx 1$$

Hence, for n large and p small,

$$P\{X = i\} \approx e^{-\lambda} \frac{\lambda^i}{i!}$$

7. 设 $X_1, X_2, \ldots, X_n$ 是 n 个相互独立的连续型随机变量，其累计概率分布函数为 $F(x)$，概率密度函数为 $F'(x) = f(x)$。现将 $X_1, X_2, \ldots, X_n$ 按由小到大的顺序排列，$X_{(i)}$ 表示排列在第 i 的随机变量，则 $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ 称为 X 的次序统计量。试求 $X_{(i)}$ 的概率密度函数 $f_{X_{(i)}}(x)$。（10 分）

解：

Example 2.38 Let $X_1, \ldots, X_n$ be independent and identically distributed continuous random variables with probability distribution F and density function $F' = f$. If we let $X_{(i)}$ denote the ith smallest of these random variables, then $X_{(1)}, \ldots, X_{(n)}$ are called the *order statistics*. To obtain the distribution of $X_{(i)}$, note that $X_{(i)}$ will be less than or equal to x if and only if at least i of the n random variables $X_1, \ldots, X_n$ are less than or equal to x. Hence,

$$P\{X_{(i)} \le x\} = \sum_{k=i}^{n} \binom{n}{k} (F(x))^k (1 - F(x))^{n-k}$$

Differentiation yields that the density function of $X_{(i)}$ is as follows:

$$\begin{aligned} f_{X_{(i)}}(x) &= f(x) \sum_{k=i}^{n} \binom{n}{k} k (F(x))^{k-1} (1 - F(x))^{n-k} \\ &\quad - f(x) \sum_{k=i}^{n} \binom{n}{k} (n-k) (F(x))^k (1 - F(x))^{n-k-1} \end{aligned}$$

$$\begin{aligned} &= f(x) \sum_{k=i}^{n} \frac{n!}{(n-k)!(k-1)!} (F(x))^{k-1} (1 - F(x))^{n-k} \\ &\quad - f(x) \sum_{k=i}^{n-1} \frac{n!}{(n-k-1)!k!} (F(x))^k (1 - F(x))^{n-k-1} \\ &= f(x) \sum_{k=i}^{n} \frac{n!}{(n-k)!(k-1)!} (F(x))^{k-1} (1 - F(x))^{n-k} \\ &\quad - f(x) \sum_{j=i+1}^{n} \frac{n!}{(n-j)!(j-1)!} (F(x))^{j-1} (1 - F(x))^{n-j} \\ &= \frac{n!}{(n-i)!(i-1)!} f(x) (F(x))^{i-1} (1 - F(x))^{n-i} \end{aligned}$$

The preceding density is quite intuitive, since in order for $X_{(i)}$ to equal x, $i - 1$ of the n values $X_1, \ldots, X_n$ must be less than x; $n - i$ of them must be greater than x; and one must be equal to x. Now, the probability density that every member of a specified set of $i - 1$ of the X_j is less than x, every member of another specified set of $n - i$ is greater than x, and the remaining value is equal to x is $(F(x))^{i-1}(1 - F(x))^{n-i} f(x)$. Therefore, since there are $n!/[(i-1)!(n-i)!]$ different partitions of the n random variables into the three groups, we obtain the preceding density function. ■