北京大学数学科学学院期中考试试题

20 18 -20 19 学年第 二 学期

可能用到的数据：

a. 设标准正态分布的累积分布函数为$\Phi(x)$, $\Phi(1.645) = 0.95$, $\Phi(1.96) = 0.975$.

1. (10分) 设随机事件A,B相互独立，A,C互不相容，并且 $P(A) = P(B) = P(C) = 0.3$, 试求下列事件的概率$A - B$, $A \cup B$, $A - C$, $A \cup C$.

   Solution: 记全集为Ω, 空集为ϕ. 由题知道$A \cap C = \phi$, $P(AB) = P(A)P(B)$.

$$P(A - B) = P(A) - P(AB) = P(A) - P(A)P(B) = 0.21$$

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(AB0 \\ &= P(A) + P(B) - P(A)P(B) \\ &= 0.3 + 0.3 - 0.3 \times 0.3 = 0.51 \end{aligned}$$

$$P(A - C) = P(A) - P(AC) = P(A) - P(\phi) = 0.3 - 0 = 0.3$$

$$\begin{aligned} P(A \cup C) &= P(A) + P(C) - P(AC) \\ &= P(A) + P(B) - P(\phi) \\ &= 0.3 + 0.3 - 0 = 0.6 \end{aligned}$$

□

2. (10分) 设$P(A) = a$, $P(B) = b$, 其中$a, b > 0$. 求证

$$\frac{a+b-1}{b} \leq P(A|B) \leq \frac{a}{b}$$

   Proof: 由$AB \subset A$知, $P(AB) \leq P(A) = a$.

   因为 $P(AB) = P(A) + P(B) - P(A \cup B)$, 且$0 \leq P(A \cup B) \leq 1$ 有

$$P(AB) \geq a + b - 1$$

   又 $P(A|B) = \frac{P(AB)}{P(B)}$, 因此有：

$$\frac{a+b-1}{b} \leq P(A|B) \leq \frac{a}{b}$$

□

3. (10分) 甲乙两艘轮船驶向一个不能同时停泊两艘轮船的码头，它们在一昼夜内到达的时间是等可能的。如果甲船停泊时间是一小时，乙船停泊时间是两小时，求它们中任何一艘都不需要等候码头空出的概率。

   Solution:

   设X为甲到达的时间，Y为乙到达的时间，则$X, Y \sim U(0, 24)$. $D = \{(x, y)|0 \leq x, y \leq 1\}$. 现在来看需要等待的情况： 如果甲先到, 即$X < Y$, 乙船需要等待进港即乙船在甲船停泊的1小时内

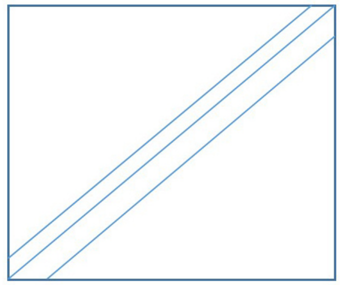

Figure 1: 事件示意图

到达了，即$Y < X+1$. 如果乙先到, 即$X > Y$, 甲船需要等待即甲船在乙船停泊的2小时内到达了，即$X < Y+2$. 对应于下图中间的两个条带，

于是按照几何概型，得到不需要等待的概率

$$P(A) = \frac{m(A)}{m(D)} = \frac{\frac{1}{2}\times 22^2 + \frac{1}{2}\times 23^2}{24^2} = \frac{1013}{1152} \approx 0.8793$$

□

4. (10分) 设随机变量$X \sim N(\mu, \sigma^2)$, 求 $E(|X|)$.

Solution:

$$E(|X|) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} |x| e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

令$\frac{x-\mu}{\sigma^2} = t$, 则$x = \mu + \sigma t$, $dx = \sigma dt$. 于是

$$\begin{aligned}
E(|X|) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} |\mu + \sigma t| e^{-\frac{1}{2}t^2} dt \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\frac{u}{\sigma}}^{+\infty} (\mu + \sigma t) e^{-\frac{1}{2}t^2} dt - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{u}{\sigma}} (\mu + \sigma t) e^{-\frac{1}{2}t^2} dt \\
&= \frac{\mu}{\sqrt{2\pi}}\left[\int_{-\frac{u}{\sigma}}^{+\infty} e^{-\frac{1}{2}t^2} dt - \int_{-\infty}^{-\frac{u}{\sigma}} e^{-\frac{1}{2}t^2} dt\right] \\
&\quad + \frac{\sigma}{\sqrt{2\pi}}\left[\int_{-\frac{u}{\sigma}}^{+\infty} t e^{-\frac{1}{2}t^2} dt - \int_{-\infty}^{-\frac{u}{\sigma}} t e^{-\frac{1}{2}t^2} dt\right] \\
&= \mu\left[\Phi(\frac{\mu}{\sigma}) - \Phi(-\frac{\mu}{\sigma})\right] + \frac{2\sigma}{\sqrt{2\pi}}\int_{-\frac{u}{\sigma}}^{+\infty} t e^{-\frac{1}{2}t^2} dt \\
&= \mu\left[2\Phi(\frac{\mu}{\sigma}) - 1\right] + \sqrt{\frac{2}{\pi}}\sigma e^{-\frac{u^2}{2\sigma^2}}
\end{aligned}$$

□

5. (10分) 设某昆虫生产k个卵的概率服从参数为λ的泊松分布 $p_k = \frac{\lambda^k}{k!}e^{-\lambda} \quad (\lambda > 0)$, 又设一个虫卵能孵化成成虫的概率为$p$. 若卵能否孵化为成虫是相互独立的，问：此昆虫的下一代有m条成虫的概率是多少？

Solution: 设下一代卵的个数为随机变量X，孵化为成虫的个数为随机变量Y, 由题意：$X \sim P(\lambda)$, $Y \sim B(n,p)$. 由全概公式

$$
\begin{aligned}
P(Y=m) &= \sum_{k=m}^{+\infty} P(Y=m, X=k) = \sum_{k=m}^{+\infty} P(Y=m|X=k)P(X=k) \\
&= \sum_{k=m}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} p^m q^{k-m} \\
&= \frac{(\lambda p)^m}{m!} e^{-\lambda} \sum_{k=m}^{+\infty} \frac{\lambda^{k-m}}{(k-m)!} q^{k-m} \\
&= \frac{(\lambda p)^m}{m!} e^{-\lambda} e^{\lambda q} \\
&= \frac{(\lambda p)^m}{m!} e^{-\lambda p)}
\end{aligned}
$$

即Y服从参数为λp的泊松分布。

□

6. (10分) 设某地区的肝癌发病率为0.0001. 现有某种肝癌检测仪器的准确率为0.99 (即患有肝癌的人仪器以99%的准确度判断其有病，而没患肝癌的人检测结果中99%显示其没有病)。现某人采用此仪器检测显示有病，问其真正患肝癌的概率是多少？

Solution: 记A为“此人患癌症”，B为“此人被仪器检测出癌症”，由已知得到：

$$P(A) = 0.0001, P(B|A) = 0.99, P(B|\overline{A}) = 1 - P(\overline{B}|\overline{A}) = 0.01$$

由贝叶斯公式

$$
\begin{aligned}
P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})} \\
&= \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.01 \times 0.9999} = \frac{1}{102} \approx 0.098
\end{aligned}
$$

□

7. (10分) 设(X, Y)的联合密度为

$$f(x,y) = \begin{cases} Ce^{-(x+2y)}, & x \geq 0, y \geq 0 \\ 0, & \text{其他} \end{cases}$$

(1). 求常数C;

(2). 求$P(0 < X < 1, 0 < Y < 0.5)$.

Solution: 注意到密度函数可分解，于是X, Y独立。

(1). 由概率的规范性

$$
\begin{aligned}
\int_D f(x,y)dxdy &= C\int_0^{+\infty} e^{-x}dx \int_0^{+\infty} e^{-2y}dy \\
&= C\left[(-e^{-x})|_0^{+\infty}\right]\left[(-\frac{1}{2}e^{-2y})|_0^{+\infty}\right] = \frac{C}{2} = 1
\end{aligned}
$$

于是$C = 2$.

(2). 同理

$$
\begin{aligned}
P(0 < X < 1, 0 < Y < 0.5) &= 2\left[(-e^{-x})|_0^1\right]\left[(-\frac{1}{2}e^{-2y})|_0^{0.5}\right] \\
&= (1 - e^{-1})^2
\end{aligned}
$$

□

8. (10分) 设a是区间$(0,1)$上的一个定点，随机变量$X \sim U(0,1)$服从均匀分布。以Y表示点X到点a的距离。问a取何值时，X与Y不相关。

Solution: $X \sim U(0,1)$, 则$E(X)=\frac{1}{2}$. 令$Y=|X-a|$, Y作为随机变量X的函数的期望,

$$E(Y)=E(|X-a|)=\int_0^a (a-x)dx+\int_a^1 (x-a)dx=a^2-a+\frac{1}{2}$$

同理，可计算$XY=X|X-a|$作为随机变量X的函数的期望

$$E(XY)=E(X|X-a|)=\int_0^a x(a-x)dx+\int_a^1 x(x-a)dx=\frac{a^3}{3}-\frac{a}{2}+\frac{1}{2}$$

由XY不相关得$E(XY)=E(X)E(Y)$方程

$$\frac{a^3}{3}-\frac{a}{2}+\frac{1}{2}=\frac{1}{2}(a^2-a+\frac{1}{2})$$

整理得方程

$$4a^3-6a^2+1=(2a-1)(2a^2-2a-1)=0$$

解得 $a=\frac{1}{2}$, $a=\frac{1+\sqrt{3}}{2}$,$a=\frac{1-\sqrt{3}}{2}$, 后两个解不在$(0,1)$区间舍去. 所以 $a=\frac{1}{2}$. □

9. (10分) 设(X,Y)的联合密度为

$$p(x,y)=\begin{cases}\frac{1}{8}(x+y), & 0\le x\le 2, 0\le y\le 2\\ 0, & \text{其它}\end{cases}$$

求X与Y的协方差$Cov(X,Y)$及相关系数ρ_{XY}.

Solution: 首先求边缘密度

$$p(x)=\int_0^2 p(x,y)dy=\frac{x+1}{4}$$

由对称性$p(y)=p(x)=\frac{y+1}{4}$

$$E(X)=\int_0^2 xp(x)dx=\int_0^2\frac{x(x+1)}{4}=\frac{7}{6}Var(Y)\quad=\int_0^2(x-\frac{7}{6})^2\frac{x+1}{4}dx=\frac{11}{36}$$

由对称性有$E(Y)=E(X)=\frac{7}{6}, Var(Y)=Var(X)=\frac{11}{36}$

$$\begin{aligned}E(XY)&=\int\int_D xyp(x,y)dxdy\\&=\frac{1}{8}\int_0^2 dx\int_0^2 xy(x+y)dy\\&=\frac{1}{8}\int_0^2(2x^2+\frac{8}{3}x)dx=\frac{4}{3}\end{aligned}$$

于是

$$Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{4}{3}-\left(\frac{7}{6}\right)^2=-\frac{1}{36}$$

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{-\frac{1}{36}}{\frac{11}{36}}=-\frac{1}{11}$$

□

10. (10分) 设某生产线上组装每件产品的时间服从指数分布，平均需要10分钟，且各件产品的组装时间相互独立。如果要保证有95%的可能性，求16小时最多可以组装多少件产品。

Solution: 记X为生产一件产品的时间(以分钟为单位)，由题意$X \sim Exp(\lambda)$, $E(X) = \frac{1}{\lambda} = 10$, 于是$\lambda = \frac{1}{10}$. $Var(X) = \frac{1}{\lambda^2} = 100$.

生产n件产品的时间为部分和$S_n = X_1 + X_2 + \cdots + X_n$, X_i独立同分布；由题意，就是要求最大n, 使得

$$P(S_n \leq 960) \geq 0.95$$

由中心极限定理

$$\frac{S_n - nE(X)}{\sqrt{nVar(X)}} \xrightarrow{d} N(0,1)$$

于是

$$\begin{aligned} P(S_n \leq 960) &= P(\frac{S_n - 10n}{10\sqrt{n}} \leq \frac{960 - 10n}{10\sqrt{n}}) \\ &= \Phi(\frac{96 - n}{\sqrt{n}}) \geq 0.95 \end{aligned}$$

由于$\Phi(1.645) = 0.95$, 于是可以得到方程

$$\frac{96 - n}{\sqrt{n}} = 1.645$$

求解得到$n \approx 81.18$, 向下取整得到最大$n = 81$. □