武汉大学 2019-2020 第一学期

概率统计 B 期终试题（A 卷）答案

一、（12 分）已知 $P(A)=0.5, P(B)=0.6, P(A\mid B)=0.5,$ 求 $P(\overline{A\bigcup B})$ 和 $P(A\overline{B}|(AUB))$ 。

解：由乘法公式, 得

$$P(AB)=P(A\mid B)P(B)=0.6\times 0.5=0.3.$$

则有

$$P(A\bigcup B)=P(A)+P(B)-P(AB)=0.5+0.6-0.3=0.8.$$

由此得 $P(\overline{A\bigcup B})=0.2.$

由于 $P(A\overline{B})=P(A)-P(AB)=0.5-0.3=0.2$ 且 $A\overline{B}\subset A\subset AUB$, 得

$$P(A\overline{B}\mid A\bigcup B)=\frac{P(A\overline{B})}{P(A\bigcup B)}=\frac{0.2}{0.8}=0.25.$$

二、（12 分）一批外表完全一样的元件，来自甲乙丙三厂，各占比例为 $5:3:2$，已知他们各自的次品率分别为 $0.02, 0.01, 0.03$；从这批元件中任取一件；求 ⑴它是次品的概率？ ⑵若它是次品，它来自甲乙丙三厂的概率各是多少？

解：设事件 A,B,C 依次表示元件来自于甲乙丙三厂, 事件 D 表示产品是次品, 则由题目条件知 $P(A)=0.5,P(B)=0.3,P(A)=0.2$ 及 $P(D|A)=0.02,P(D|B)=0.01,P(D|C)=0.03$.

（1）由全概率公式, 得这批元件中任取一件是品的概率

$$P(D)=P(D|A)P(A)+P(D|B)P(B)+P(D|C)P(C)=0.019.$$

（2）若它是次品，来自甲乙丙三厂的概率分别为

$$P(A|D)=\frac{P(D|A)P(A)}{P(D)}=\frac{10}{19},P(B|D)=\frac{P(D|B)P(B)}{P(D)}=\frac{3}{19},P(A|D)=\frac{P(D|C)P(C)}{P(D)}=\frac{6}{19}.$$

三、（12 分）若公司经理每天上班的时间在 9 到 10 点的任意时刻，而秘书在 8:30 到 9:30 的任意时刻；以 A 记事件“两人到达时间相差不超过 20 分钟”。 （1）求 A 发生的概率。⑵ 平常的一周（5 天）中，求 A 恰好发生三次的概率。

解：设随机变量 X 表示公司经理上班的时间, Y 表示秘书上班的时间, 记

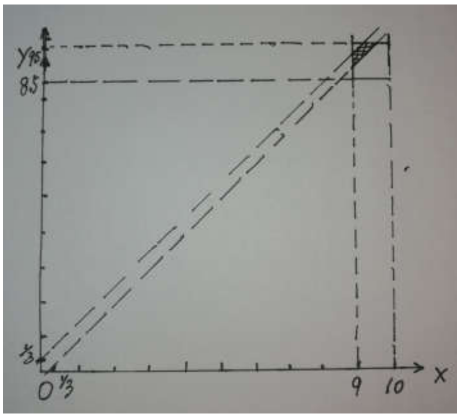
$$D\!: 9\le X\le 10,\ 8.5\le Y\le 9.5.$$

那么事件 A“两人到达时间相差不超过 20 分钟”表示为

$$A:|X-Y|\le\frac{1}{3}, \text{且 } 8\le X\le 9, 8.5\le Y\le 9.5.$$

应用几何概率，如下图所示，A 的概率即为 A 占 D 中面积的比例，得

$$P(A)=\frac{1}{3}.$$



（2）A 恰好发生三次的概率为

$$C_5^3\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2=\frac{40}{243}.$$

%%% 应用几何概率方法时，应当把图画出来.

四、若随机变量 (X,Y) 的联合概率密度为 $f(x)=\begin{cases}\frac{2}{\pi}e^{-\frac{1}{2}(x^2+y^2)} & x>0,y>0\\ 0 & \text{其它}\end{cases}$，(1)求随机变量 X 和 Y 的边沿缘概率密度 $f_x(x);f_y(y)$；并判别他们是否独立？（2）求 $Z=\sqrt{X^2+Y^2}$ 的概率密度。

解：（1）当 $x>0$，得

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy=\frac{2}{\pi}\int_0^{\infty}e^{-\frac{1}{2}(x^2+y^2)}dy=\frac{1}{\pi}e^{-\frac{x^2}{2}}\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy=\sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}},$$

故

$$f_X(x)=\begin{cases}\sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}x^2} & x>0\\ 0 & \mathrm{x}\le 0\end{cases},$$

同理可得

$$f_Y(y)=\begin{cases}\sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}y^2} & y>0\\ 0 & y\le 0\end{cases}.$$

由于 $f(x,y)=f_X(x)f_Y(y)$，故随机变量 X 与 Y 相互独立.

（2）首先给出随机变量 Z 的分布函数

$$F_Z(z)=P(Z\le z)=\iint_{x+y\le z}f(x,y)dxdy,\ -\infty<z<\infty.$$

当 $z>0$ 时，运用极坐标变换，得

$$F_Z(z)=\frac{2}{\pi}\iint_{x^2+y^2\le z^2}e^{-\frac{1}{2}(x^2+y^2)}dxdy$$
$$=\frac{2}{\pi}\int_0^{\frac{\pi}{2}}\int_0^z e^{-\frac{r^2}{2}}rdrd\theta=1-e^{-\frac{1}{2}z^2}.$$

故

$$F_Z(z)=\begin{cases}1-e^{-\frac{1}{2}z^2} & z>0\\ 0 & z\le 0\end{cases}.$$

那么 $Z=X+Y$ 的密度函数为

$$f_Z(z)=\begin{cases}ze^{-\frac{1}{2}z^2} & z>0\\ 0 & z\le 0\end{cases}.$$

或者在 (2) 中使用积分转换法:

(2) 由题意，使用积分转换法，设 h 是任一连续有界的辅助函数，$g(x,y)=\sqrt{x^2+y^2}$，得

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(g(x,y))f(x,y)dxdy=\frac{2}{\pi}\int_0^{\infty}\int_0^{\infty}h\left(\sqrt{x^2+y^2}\right)e^{-\frac{1}{2}(x^2+y^2)}dxdy$$

运用极坐标变换，令 $z=\sqrt{x^2+y^2}$，得

$$\frac{2}{\pi}\int_0^{\infty}\int_0^{\infty}h\left(\sqrt{x^2+y^2}\right)e^{-\frac{1}{2}(x^2+y^2)}dxdy=\frac{2}{\pi}\int_0^{\frac{\pi}{2}}d\theta\int_0^{\infty}h(z)e^{-\frac{1}{2}z^2}zdz=\int_0^{\infty}h(z)ze^{-\frac{1}{2}z^2}dz,$$

因此，$Z=\sqrt{X^2+Y^2}$ 的密度函数为

$$f_Z(z)=\begin{cases}ze^{-\frac{1}{2}z^2} & z>0\\ 0 & z\le 0\end{cases}.$$

%%% 如果你到现在是还是不能理解积分转换法的步骤，那么就不要用这种作法，而选用前面的标准作法.

五、（12 分）某机器一天正常工作的概率为 0.8，已知：正常时一台机器每天获利 8 万元，故障时每台机器亏损 2 万元。现有 100 台此机器；（1）求每天的平均利润。若希望平均利润达到 700 万元，提出一个解决办法。（2）现有情况下，为保证一天的利润不低于 3000 万元的概率大于 0.977，问要增加工多少台机器？（已知 $\Phi(2.0)=0.977$）

解（1)对 100 台机器编号，设 X_i 表示第 i 台机器获得的利润，则 $EX_i=0.8\times 8-0.2\times 2=6.$

那么 100 台机器的平均利润为

$$E\left(\sum_{i=1}^{100} X_i\right)=600.$$

若希望平均利润达到 700 万元，则可提高机器正常工作的概率到 0.9, 或增加机器台数到 177 台或以上。

（2）仍用 X_i 表示第 i 台机器获得的利润，则

$$EX_i=0.8\times 8-0.2\times 2=6,\ DX_i=16.$$

设增加的机器台数为 n，由题意，需 $P(\sum_{i=1}^{100+n} X_i\geq 3000)\geq 0.977$。那么由中心极限定理，以及 $E\left(\sum_{i=1}^{100+n} X_i\right)=600+6n,\ D\left(\sum_{i=1}^{100+n} X_i\right)=1600+16n$，得

$$\frac{3000-600-6n}{4\sqrt{100+n}}\leq -2\ 。$$

解得 $n\geq 431.$

法二：

（1）设 100 台机器一天正常工作的台数为 X，则 $X\sim B(100,0.8)$，且 $EX=80$. 设一天的利润为 Y，则

$$Y=8X-2(100-X)=10X-200.$$

那么平均利润为

$$EY=E(10X-200)=10EX-200=600.$$

若希望平均利润达到 700 万元，则可增加机器台数或提高机器正常工作的的概率。

（2）设机器工作台数为 n，$X\sim B(n,0.8)$，$EX=0.8n,\ DX=0.16n$，得

$$EY=E(10X-2n)=6n,\quad DY=D(10X-2n)=100\times 0.16n=16n.$$

由题意，需 $P(Y\geq 3000)\geq 0.977$. 即

$$P(10X-2n\geq 3000)=P(X\geq \frac{3000+2n}{10})\geq 0.977.$$

那么由中心极限定理,及 $X\sim B(n,0.8)$，得

$$\left(\frac{3000+2n}{10}-0.8n\right)\Big/0.4n=\frac{3000-6n}{4\sqrt{n}}\leq -2$$

解得 $n \geq 531$. 故至少要加 431 台机器

或者在 (2) 中直接写:

（2）设机器工作台数为 n, $X \sim B(n,0.8)$, $EX = 0.8n, DX = 0.16n$, 得

$$EY = E(10X-2n) = 6n,\ \ DY = D(10X-2n) = 100 \times 0.16n = 16n.$$

由题意，需 $P(Y \geq 3000) \geq 0.977$. 即 $P(\frac{Y-6n}{4\sqrt{n}} \geq \frac{3000-6n}{4\sqrt{n}}) \geq 0.977$.那么由中心极限定理，得 $\frac{3000-6n}{4\sqrt{n}} \leq -2$，解得 $n \geq 531$. 故至少要加 431 台机器

%%% 注意，最后一步的增加台数不用计算器并不好算.

六、(12 分) 若 $X_1, X_2, \cdots, X_8$ 是正态总体 $N(\mu,\sigma^2)$ 的样本，（1）求常数 a,b,c,d （这里 $abc \neq 0$），使 $Y = a(X_1-X_2)^2 + b(2X_3-X_4-X_5)^2 + c(3X_6-2X_7-X_8)^2 \sim \chi^2(d)$；

（2）若 $Z = \sum_{i=1}^{8}(X_i-\mu)^2$，求 Z 的期望与方差

解:（1）由于

$$Y = (\sqrt{a}(X_1-X_2))^2 + (\sqrt{b}(2X_3-X_4-X_5))^2 + (\sqrt{c}(3X_6-2X_7-X_8))^2 \sim \chi^2(d),$$

且

$$D\left(\sqrt{a}(X_1-X_2)\right) = 2a\sigma^2, D\left(\sqrt{b}(2X_3-X_4-X_5)\right) = 6b\sigma^2, D\left(\sqrt{c}(3X_6-2X_7-X_8)\right) = 14c\sigma^2,$$

而且 $abc \neq 0$，故得

$$a = \frac{1}{2\sigma^2}, b = \frac{1}{6\sigma^2}, c = \frac{1}{14\sigma^2}, d = 3.$$

（2）因 $X_i \sim N(\mu,\sigma^2)$, 故 $\frac{X_i-\mu}{\sigma} \sim N(0,1), i=1,2,\cdots,8,$. 从而 $\frac{Z}{\sigma^2} \sim \chi^2(8)$, 并且

$$EZ = 8\sigma^2, DZ = 16\sigma^4.$$

%%% 卡方分布的期望与方差可以记下来直接运用.

七、(12 分) 已知 X 的概率密度为 $f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(x-\mu)} & x > \mu \\ 0 & x \leq \mu \end{cases}$, $\lambda > 0$，$X_1, X_2, \cdots, X_n$ 是样本，试求参数 μ, λ 的最大似然估计，并判别是否无偏。

解：取样本值为$x_1,x_2,\cdots,x_n$，则有似然函数

$$L=L(\lambda,\mu)=\frac{1}{\lambda^n}e^{-\frac{1}{\lambda}\sum_{i=1}^{n}(x_i-\mu)}.$$

首先求得

$$L'_\mu=\frac{n}{\lambda^{n+1}}e^{-\frac{1}{\lambda}\sum_{i=1}^{n}(x_i-\mu)}>0,$$

而由题目条件可知 $\mu\le\min\{x_1,x_2,\cdots,x_n\}$，故由极大似然估计法，取

$$\widehat{\mu}=\min\{x_1,x_2,\cdots,x_n\}.$$

另一方面，令

$$L'_\lambda=\left(-\frac{n}{\lambda^{n+1}}+\frac{\sum_{i=1}^{n}(x_i-\mu)}{\lambda^{n+2}}\right)e^{-\frac{1}{\lambda}\sum_{i=1}^{n}(x_i-\mu)}=\left(\frac{\sum_{i=1}^{n}(x_i-\mu)}{\lambda}-n\right)\frac{1}{\lambda^{n+1}}e^{-\frac{1}{\lambda}\sum_{i=1}^{n}(x_i-\mu)}=0,$$

得$\lambda=\frac{1}{n}\sum_{i=1}^{n}x_i-\mu$，故由最大似然估计法，取

$$\widehat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}x_i-\min\{x_1,x_2,\cdots,x_n\}.$$

相应的估计量分别为

$$\widehat{\mu}=\min\{X_1,X_2,\cdots,X_n\},\quad \widehat{\lambda}=\overline{X}-\min\{X_1,X_2,\cdots,X_n\}.$$

接下来先求$\widehat{\mu}$的期望,，设F_m,f_m分别表示其分布函数和密度函数，则

$$F_m(x)=1-(1-F(x))^n,$$

.

并由题目条件，令$x>\mu$，得

$$f_m(x)=n(1-F(x))^{n-1}f(x)=\begin{cases}\frac{n}{\lambda}e^{-\frac{n}{\lambda}(x-\mu)} & x>\mu\\ 0 & x\le\mu\end{cases}.$$

那么

$$E\left(\widehat{\mu}\right)=\int_\mu^\infty xf_m(x)dx=\frac{n}{\lambda}\int_\mu^\infty xe^{-\frac{n}{\lambda}(x-\mu)}xf_m(x)dx=\frac{\lambda}{n}+\mu.$$

再求$\widehat{\lambda}$的期望，首先

$$E\left(X\right)=\int_{\mu}^{\infty}xf(x)dx=\frac{1}{\lambda}\int_{\mu}^{\infty}xe^{-\frac{1}{\lambda}(x-\mu)}xf_m(x)dx=\lambda+\mu.$$

从而有

$$E\left(\hat{\lambda}\right)=E\left(\bar{X}-\min\{X_1,X_2,\cdots,X_n\}\right)=\frac{n-1}{n}\lambda.$$

因此, λ 和 μ 的最大似然估计都有偏.

%%% 极大值和极小值的密度函数以及期望都需要做具体的计算, 并不容易.

八、(12 分)某作物的产量近似服从正态分布 $N(\mu,\sigma^2)$，现发现新的种子，取 25 块样田做实验，发现平均亩产为 1864 公斤，样本标准差为 50 公斤；问：此新种子的亩产量是不是显著大于 1800 公斤？（$\alpha=0.05$） 已知：$u_{0.05}=1.65,u_{0.025}=1.96$，

$t_{0.05}(25)=1.708,t_{0.05}(24)=1.712,t_{0.025}(25)=2.060,t_{0.025}(24)=2.064$

解：根据题意做假设检验, 设

$$H_0:\mu=1800,H_1:\mu>1800$$

并取检验统计量为 $t=\dfrac{\overline{X}-1800}{s}\sqrt{n}\sim t(n-1)$. 由题目条件, 取 $n=25,s=50,\alpha=0.05,$

及 $t_\alpha(24)=1.712$, 给出拒接域：$t\geq1.712$.

代入样本均值 $\bar{x}=1864$, 计算得 $t=6.4\geq1.712$. 所以拒接 H_0, 认为此新种子的亩产量显著大于 1800 公斤.