复习

一. 概率

概率 $P(A)=1-P(\bar{A})$ $P(A\cup B)=P(A)+P(B)-P(AB)$

条件概率 $P(A|B)=\dfrac{P(AB)}{P(B)}$

设 $A_1, A_2, \cdots, A_n$ 为全空间的一个划分 则

$$P(B)=\sum_{i=1}^{n}P(B|A_i)P(A_i) \quad \text{(全概率公式)}$$

$$P(A_i|B)=\frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{n}P(B|A_k)P(A_k)} \quad \text{(贝叶斯公式)}$$

事件独立 $P(AB)=P(A)P(B)$ 或 $P(A|B)=P(A)$.

1. 一考生回答一道有 $m\,(m\geqslant 2)$ 个选择的选择题. 设他知道正确答案的概率为 p；不知道时就猜. 已知该生回答正确, 则他知道正确答案的条件概率为多少？

解: 设 $A=\{$回答正确$\}$ $B=\{$知道正确答案$\}$ 则

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

$$=\frac{1\cdot p}{1\cdot p+\frac{1}{m}\cdot(1-p)}$$

$$=\frac{mp}{mp+1-p}.$$

□

2. 设 $A_1, \cdots, A_n$ 相互独立，证明

$$P\left(\bigcup_{k=1}^{n} A_k\right) = 1 - \prod_{k=1}^{n} P(\overline{A_k}).$$

证. $P\left(\bigcup_{k=1}^{n} A_k\right) = 1 - P\left(\overline{\bigcup_{k=1}^{n} A_k}\right) = 1 - P\left(\bigcap_{k=1}^{n} \overline{A_k}\right)$

$$\overset{\text{独立性}}{=} 1 - \prod_{k=1}^{n} P(\overline{A_k}).$$

$\square$

3. 求 n 阶行列式的展开式中任取一项，这项至少含一个主对角线元素的概率

解：$a_{1 i_1} a_{2 i_2} \cdots a_{n i_n}$. $i_k \neq i_l, \ \forall\ k \neq l$

求 $P(j = i_j \text{ for some } j)$.

令 $A_j = \{ j = i_j \}$

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j) - \sum_{i \neq j} P(A_i A_j) + \sum_{i \neq j \neq k} P(A_i A_j A_k) + \cdots$$

$$P(A_j) = \frac{(n-1)!}{n!}, \quad P(A_i A_j) = \frac{(n-2)!}{n!}, \cdots$$

$$\Rightarrow P\left(\bigcup_{j=1}^{n} A_j\right) = n \cdot \frac{(n-1)!}{n!} - C_n^2 \frac{(n-2)!}{n!} + C_n^3 \frac{(n-3)!}{n!} + \cdots$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} \cdots$$

$$= \sum_{j=1}^{n} \frac{(-1)^{j+1}}{j!}$$

$\square$

4. 袋中有M个球，含N个白球，不放回取n个球，

$A_i=\{$第i次取到白球$\}$，

$B_k=\{$取了k个白球$\}$.

求 $P(A_i\mid B_k)$.

√ √ * $\quad n=3$

* √ √ $\quad i=2$

$k=2$

√ √ * $\quad n=3$

√ * √ $\quad i=1$

$k=2$

解：

$$P(A_i\mid B_k)=\frac{P(A_iB_k)}{P(B_k)}$$

$$P(B_k)=\frac{C_N^k\,C_{M-N}^{n-k}}{C_M^n}$$

$$P(A_iB_k)=P(A_1B_k)=P(B_k\mid A_1)P(A_1)$$

$$=\frac{C_{N-1}^{k-1}\,C_{M-N}^{n-k}}{C_{M-1}^{n-1}}\cdot\frac{N}{M}$$

$$\Rightarrow\quad P(A_i\mid B_k)=\frac{C_{N-1}^{k-1}}{C_{M-1}^{n-1}}\cdot\frac{N}{M}\cdot\frac{C_M^n}{C_N^k}=\frac{k}{n}\qquad\square$$

二、随机变量

1、分布函数. $F_X(x) = P(X \le x)$

联合分布与边缘分布. $\begin{cases} p_{i\cdot} = \sum_j p_{ij} \\ f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy \end{cases}$

随机变量函数的分布.(按定义求).

X与Y独立时, $\begin{cases} P(X+Y=k) = \sum_i P(X=i)\,P(Y=k-i) \\ f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(x)\, f_Y(z-x)\,dx. \end{cases}$

2、数字特征

期望 $EX = \begin{cases} \sum_i x_i p_i \\ \int_{-\infty}^{+\infty} x f(x)\,dx \end{cases}$ $E\,g(X) = \begin{cases} \sum_i g(x_i)\, p_i \\ \int_{-\infty}^{+\infty} g(x)\, f(x)\,dx \end{cases}$

方差 $DX = E(X-EX)^2 = EX^2 - (EX)^2$

协方差 $Cov(X,Y) = EXY - EX\,EY.$

相关系数 $\rho_{XY} = \dfrac{Cov(X,Y)}{\sqrt{DX\,DY}}$

性质
- $E(aX+b) = aEX + b$, $D(aX+b) = a^2 DX$
- $D(X \pm Y) = DX + DY \pm 2\,Cov(X,Y)$
- 若X与Y不相关,则 $E(XY) = EX\,EY$, $D(X+Y) = DX+DY.$
- 独立 $\overset{\longrightarrow}{\underset{\not\longleftarrow}{}}$ 不相关

3、常见随机变量

$q:=1-p$ 名称	期望	方差
$B(n,p)$ $p_k=C_n^k p^k q^{n-k}$ $0\le k\le n$.	np	npq
$P(\lambda)$ $p_k=\frac{\lambda^k}{k!}e^{-\lambda}$, $k\geq 0$	λ	λ
几何分布 $p_k=pq^{k-1}$, $k\geq 1$	$\frac{1}{p}$	$\frac{q}{p^2}$
$U(a,b)$ $f(x)=\frac{1}{b-a}$, $a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$E(\lambda)$ $f(x)=\lambda e^{-\lambda x}$, $x>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$N(\mu,\sigma^2)$ $f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

1. 设每袋水泥重量 $X\sim N(50,2.5^2)$ (kg)，卡车载重为2吨. 为了以0.95的概率不超载，一车最多能装多少袋水泥？

解：设装 n 袋水泥，则总重量

$$Y=\sum_{i=1}^{n}X_i \sim N(50n,\ 2.5^2 n)$$

$$P(Y\le 2000)\geq 0.95$$

$$左=P\left(\frac{Y-50n}{2.5\sqrt{n}}\le\frac{2000-50n}{2.5\sqrt{n}}\right)=\Phi\left(\frac{2000-50n}{2.5\sqrt{n}}\right)$$

查表 $\Phi(1.64)\approx 0.9495$ $\Phi(1.65)\approx 0.9505$

$$\therefore \quad \frac{2000-50n}{2.5\sqrt{n}}\geq 1.645 \Rightarrow n\le 39.483$$

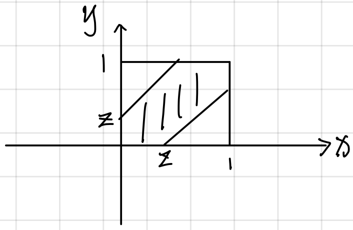
故一车最多装39袋水泥. □

2. 在线段 $(0,1)$ 内任取两点，求两点间距离的分布函数、概率密度、期望、方差.

解：令 $X, Y \sim U(0,1)$ 且独立，$Z = |X-Y|$.

$\forall\ 0<z<1$,

$$P(Z \le z) = \iint_{|x-y|\le z} f(x,y)\,dx\,dy$$



$$= \iint_{\substack{|x-y|\le z \\ 0<x<1 \\ 0<y<1}} dx\,dy$$

$$= 1-(1-z)^2 = z(2-z)$$

$$\therefore\ F_Z(z) = \begin{cases} 0, & z<0 \\ z(2-z), & 0\le z<1 \\ 1, & z\ge 1 \end{cases}$$

$$f_Z(z) = \begin{cases} 2(1-z), & 0\le z\le 1 \\ 0, & \text{其它} \end{cases}$$

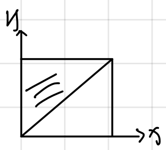
两种方法求期望和方差：

方法一：$EZ = \int z f_Z(z)\,dz = \int_0^1 2z(1-z)\,dz = \frac{1}{3}$

$$EZ^2 = \int z^2 f_Z(z)\,dz = \int_0^1 2z^2(1-z)\,dz = \frac{1}{6}$$

$$DZ = EZ^2 - (EZ)^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$$

方法二：$EZ = E|X-Y| = \iint |x-y| f(x,y)\,dx\,dy$



$$= \int_0^1\int_0^1 |x-y|\,dx\,dy = 2\iint_{y>x} (y-x)\,dx\,dy$$

$$= 2\int_0^1\int_0^y (y-x)\,dx\,dy = 2\int_0^1 \frac{y^2}{2}\,dy = \frac{1}{3}$$

$$EZ^2 = E|X-Y|^2 = \int_0^1\int_0^1 (x-y)^2\,dx\,dy = \frac{1}{6}$$

$$DZ = EZ^2 - (EZ)^2 = \frac{1}{18}.$$

□

3. 将n个球放入M个盒子中，设每个球放入各个盒子是等可能的. 求有球的盒子数X的期望.

解：令 $X_i = \begin{cases} 1 & \text{第}i\text{个盒子中有球} \\ 0 & \text{其它} \end{cases}$ $\quad 1 \le i \le M$

则 $X = \sum_{i=1}^{M} X_i$

$P(X_i = 0) = \left(1 - \frac{1}{M}\right)^n$

$\therefore \ EX_i = P(X_i = 1) = 1 - \left(1 - \frac{1}{M}\right)^n$

$EX = \sum_{i=1}^{M} EX_i = M\left[1 - \left(1 - \frac{1}{M}\right)^n\right]$. $\square$

4. $X_1, \cdots, X_n$ i.i.d. 证

$$\rho_{X_i - \overline{X},\, X_j - \overline{X}} = -\frac{1}{n-1}, \quad i \ne j.$$

证：$\mathrm{Cov}(X_i - \overline{X}, X_j - \overline{X})$

$= \mathrm{Cov}(X_i, X_j) + \mathrm{Cov}(\overline{X}, \overline{X}) - \mathrm{Cov}(X_i, \overline{X}) - \mathrm{Cov}(X_j, \overline{X})$

$\triangleq \mathrm{I} + \mathrm{II} - \mathrm{III} - \mathrm{IV}$.

$\mathrm{I} = 0$. $\quad \mathrm{II} = D\overline{X} = \frac{1}{n}\sigma^2$

$\mathrm{Cov}(\overline{X}, \overline{X}) = \mathrm{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} X_i, \overline{X}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Cov}(X_i, \overline{X}) = \mathrm{Cov}(X_i, \overline{X})$, $\ i = 1, \cdots, n$

$\therefore \ \mathrm{III} = \mathrm{IV} = \frac{1}{n}\sigma^2$ 即

$\mathrm{Cov}(X_i - \overline{X}, X_j - \overline{X}) = -\frac{1}{n}\sigma^2$

$D(X_i - \overline{X}) = DX_i + D\overline{X} - 2\,\mathrm{Cov}(X_i, \overline{X})$

$= \sigma^2 + \frac{1}{n}\sigma^2 - \frac{2}{n}\sigma^2 = \frac{n-1}{n}\sigma^2$, $\quad 1 \le i \le n$.

$\therefore \ \rho_{X_i - \overline{X},\, X_j - \overline{X}} = \dfrac{\mathrm{Cov}(X_i - \overline{X}, X_j - \overline{X})}{\sqrt{D(X_i - \overline{X})\, D(X_j - \overline{X})}} = -\frac{1}{n-1}$. $\square$

5. X，Y 同分布，$f(x)=\begin{cases}\frac{3}{8}x^2, & 0<x<2 \\ 0, & 其它\end{cases}$

$A=\{X>a\}$ 与 $B=\{Y>a\}$ 独立，且 $P(A\cup B)=\frac{3}{4}$，求 a.

解：$\frac{3}{4}=P(A\cup B)=1-P(\bar{A}\cap\bar{B})$

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

$\because\ P(A)=P(B)$

$\therefore\ P(\bar{A})=\frac{1}{2}\qquad P(X\le a)=\frac{1}{2}$

又

$$P(X\le a)=\int_0^a \frac{3}{8}x^2\,dx=\frac{a^3}{8}$$

$\therefore\ \frac{3}{8}a^3=\frac{1}{2}\qquad a=4^{\frac{1}{3}}$. □

6. $|X|\le 1$. $P(X=-1)=\frac{1}{8}$，$P(X=1)=\frac{1}{4}$，在 $\{-1<X<1\}$ 发生的条件下，$X\sim U(-1,1)$. 求 X 分布函数.

解：$\forall\ -1<x<1$，

$$\begin{aligned}F(x)&=P(X\le x)=P(X=-1)+P(-1<X\le x)\\&=\frac{1}{8}+P(-1<X\le x,-1<X<1)\\&=\frac{1}{8}+P(-1<X\le x\mid -1<X<1)P(-1<X<1)\\&=\frac{1}{8}+\frac{x+1}{2}\times\left(1-\frac{1}{8}-\frac{1}{4}\right)\\&=\frac{1}{8}+\frac{5}{16}(x+1)=\frac{5x+7}{16}\end{aligned}$$

$$\therefore\ F(x)=\begin{cases}0, & x<-1\\ \frac{5x+7}{16}, & -1\le x<1\\ 1, & x\ge 1\end{cases}$$

□

7. 设 X 概率密度 $f(x)=\begin{cases} \frac{1}{a}x^2, & 0<x<3 \\ 0, & 其它 \end{cases}$

令 $Y=\begin{cases} 2, & X\le 1 \\ X, & 1<X<2 \\ 1, & X\ge 2. \end{cases}$

(1) 求 Y 分布函数. (2) 求 $P(X\le Y)$.

解: (1) $\int_{-\infty}^{+\infty} f(x)\,dx = \int_0^3 \frac{1}{a}x^2\,dx = 1 \Rightarrow a=9.$

$1\le Y\le 2.$

$\forall\ 1<y<2,$

$$P(Y\le y) = P(Y\le y,\ X\le 1) + P(Y\le y,\ 1<X<2) + P(Y\le y,\ X\ge 2)$$

$$= 0 + P(1<X\le y) + P(X\ge 2)$$

$$= \int_1^y \frac{1}{9}x^2\,dx + \int_2^3 \frac{1}{9}x^2\,dx$$

$$= \frac{1}{27}(y^3-1+27-8) = \frac{y^3+18}{27}$$

$\therefore\ F_Y(y)=\begin{cases} 0, & y<1 \\ \frac{y^3+18}{27}, & 1\le y<2 \\ 1, & y\ge 2 \end{cases}$

(2) $$P(X\le Y) = P(X\le Y,\ X\le 1) + P(X\le Y,\ 1<X<2) + P(X\le Y,\ X\ge 2)$$

$$= P(X\le 1) + P(1<X<2) + 0$$

$$= P(X<2)$$

$$= \int_0^2 \frac{1}{9}x^2\,dx = \frac{8}{27}.$$

□

8. $X \sim E(\lambda_1)$, $Y \sim E(\lambda_2)$, 且独立, 求 $Z = X + Y$ 概率密度.

解: $z>0$, $f(z) = \int f_X(x) f_Y(z-x)\, dx$

$$= \int_0^z \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 (z-x)} dx$$

$$= \lambda_1 \lambda_2 e^{-\lambda_2 z} \int_0^z e^{-(\lambda_1-\lambda_2)x} dx$$

1) $\lambda_1 = \lambda_2 = \lambda$,

$$f(z) = \lambda^2 z e^{-\lambda z}, \quad z>0$$

2) $\lambda_1 \neq \lambda_2$,

$$f(z) = \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2} e^{-\lambda_2 z}\left[1 - e^{-(\lambda_1-\lambda_2)z}\right]$$

$$= \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}\left[e^{-\lambda_2 z} - e^{-\lambda_1 z}\right] \qquad \square$$

9. $X \sim N(1, 3^2)$, $Y \sim N(0, 4^2)$, $\rho_{XY} = -\frac{1}{2}$.

$Z = \frac{X}{3} + \frac{Y}{2}$. 求 1) EZ, DZ 2) ρ_{XZ}

解: 1) $EZ = \frac{1}{3}EX + \frac{1}{2}EY = \frac{1}{3}\times 1 + \frac{1}{2}\times 0 = \frac{1}{3}$

$$DZ = \frac{1}{9}DX + \frac{1}{4}DY + 2\,Cov\left(\frac{X}{3}, \frac{Y}{2}\right)$$

$$= \frac{1}{9}DX + \frac{1}{4}DY + \frac{1}{3}\rho_{XY}\sqrt{DX\,DY}$$

$$= \frac{1}{9}\times 9 + \frac{1}{4}\times 16 - \frac{1}{3}\times\frac{1}{2}\times 3\times 4$$

$$= 3$$

2) $Cov(X, Z) = Cov(X, \frac{1}{3}X + \frac{1}{2}Y)$

$= \frac{1}{3}DX + \frac{1}{2}Cov(X, Y)$

$= \frac{1}{3} \times 9 - \frac{1}{2} \times \frac{1}{2} \times 3 \times 4$

$= 0$

$\therefore \rho_{XZ} = 0$ $\square$

三、大数定律和中心极限定理

- 若 1) $X_1, X_2, \cdots$ 独立，$EX_i = \mu$，$DX_i = \sigma^2$

  或 2) $X_1, X_2, \cdots$ i.i.d.，$EX_i = \mu$，

  则 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$.

- 若 $X_1, X_2, \cdots$ i.i.d. $EX_i = \mu$，$DX_i = \sigma^2$，则 $\forall x \in R$，

$$P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \le x \right) \longrightarrow \Phi(x).$$

例、设 $X_1, \cdots, X_{100} \overset{i.i.d.}{\sim} B(1, p)$，则

$\frac{1}{100}\sum_{i=1}^{100} X_i \approx p$. ✓ $\qquad \sum_{i=1}^{100} X_i \sim B(100, p)$. ✓

$P\left(a < \sum_{i=1}^{100} X_i < b \right) \approx \Phi(b) - \Phi(a)$ ×.

$P\left(a < \sum_{i=1}^{100} X_i < b \right) \approx \Phi\left(\frac{b - 100p}{\sqrt{100p(1-p)}} \right) - \Phi\left(\frac{a - 100p}{\sqrt{100p(1-p)}} \right)$. ✓

四、数理统计

1. 三个概念：总体，样本，统计量（$\bar{X}$，S^2）.

2. 三个分布：$\chi^2(n)$，$t(n)$，$F(n_1, n_2)$.

3. 三个结论：$X_1, \cdots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ 则

   ① $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

   ② $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

   ③ $\bar{X}$ 与 S^2 独立.

4. 二个方法：矩估计；极大似然估计

5. 三个标准：无偏性；有效性；一致性

6. 置信区间 $N(\mu, \sigma^2)$ 估计 μ $\begin{cases} \sigma^2 \text{ 已知} \\ \sigma^2 \text{ 未知} \end{cases}$

   估计 σ^2：μ 未知

例. 设 $X \sim t(n)$，则 $\frac{1}{X^2} \sim$ ______. $F(n,1)$

$$X = \frac{Y}{\sqrt{\frac{Y_1^2 + Y_2^2 + \cdots + Y_n^2}{n}}} \qquad Y, Y_1, \cdots, Y_n \overset{i.i.d}{\sim} N(0,1)$$

$$\frac{1}{X^2} = \frac{(Y_1^2 + \cdots + Y_n^2)/n}{Y^2} = \frac{\chi^2(n)/n}{\chi^2(1)}$$

例. $(X_1, \cdots, X_n)$ 为取自总体为 $B(1,p)$ 的样本，则 $P(n\bar{X} = k) =$ ______.

$C_n^k p^k (1-p)^{n-k}$

$n\bar{X} = X_1 + \cdots + X_n \sim B(n, p)$.

例. $(X_1,\cdots,X_n)$ 为取自总体为 $N(0,1)$ 的样本，则 $a=$ $\underline{n}$ ，$b=$ $\underline{n-1}$ 时，

$a\bar{X}^2+bS^2\sim\chi^2(n)$.

$$a\bar{X}^2+bS^2=a\bar{X}^2+\frac{b}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$
$$=a\bar{X}^2+\frac{b}{n-1}\sum_{i=1}^{n}X_i^2-\frac{bn}{n-1}\bar{X}^2$$

例. $x_1,\cdots,x_n$ 为来自总体 $N(\mu,\sigma^2)$ 的样本观测值，$\bar{x}=9.5$，参数 μ 的置信度为 95% 的置信区间上限为 10.8，则 μ 的置信度为 95% 的置信区间下限为 ______. 8.2.

σ^2 已知：$\frac{\bar{X}-\mu}{\sigma}\sqrt{n}\sim N(0,1)$ $\Rightarrow\left(\bar{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\ \bar{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right)$

σ^2 未知：$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，$\frac{\bar{X}-\mu}{S}\cdot\sqrt{n}\sim t(n-1)$

$\Rightarrow\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\ \bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$.

$9.5\times2-10.8=8.2$

例. 设正态总体均值 μ 的置信区间长度为 $l=2\frac{S}{\sqrt{n}}t_{\alpha}(n-1)$，则其置信水平为 ____. $1-2\alpha$.

例. 总体 $X\sim N(\mu,\sigma^2)$. σ^2 未知时，μ 的置信区间长度为 l_1；在置信水平不变的条件下，用 σ^2 的无偏估计 S^2 作为 σ^2 的的已知值，所得 μ 的置信区间长度为 l_2，则 l_1 ____ l_2.

$l_1=\frac{2S}{\sqrt{n}}t_{\alpha/2}(n-1)$ $\qquad$ $l_2=\frac{2S}{\sqrt{n}}u_{\alpha/2}$

$P(t(n)>t_0)\geqslant P(N(0,1)\geqslant t_0)$

$\because t_{\alpha/2}(n-1)\geqslant u_{\alpha/2}$. $\quad\therefore l_1\geqslant l_2$.