一、条件概率与事件独立性

$P(AB) = P(B)P(A|B)$

$P(A) = \sum_i P(B_i)P(A|B_i)$

$P(B_i|A) = \frac{P(AB_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_j P(B_j)P(A|B_j)}$

独立：$P(A|B) = P(A)$，$P(B|A) = P(B)$

$P(AB) = P(A)P(B)$ ⇒ 主要用作证明独立性

A, B 独立 ⇒ $A, \bar{B}$ 独立

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

二、常用离散型分布

1. 二项分布与伯努利分布 $b(n,p)$ ⇒ n次试验，每次成功概率为p

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

$E[X] = np$，$Var(X) = np(1-p)$

2、几何分布 $X \sim Ge(p)$ 每次成功概率为p，求第一次成功的次数

$$P(X=k) = (1-p)^{k-1} \cdot p$$

无记忆性： $P(X=8 \mid X>5) = P(X=3)$

$$P(X>s+t \mid X>s) = P(X>t)$$

$E[X] = \frac{1}{p}$， $Var(X) = \frac{1-p}{p^2}$

3、泊松分布 $X \sim P(\lambda)$ ⇒ 描述稀有事件发生概率

$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$， $E[X] = \lambda$， $Var(X) = \lambda$

二项分布的泊松近似

p小，n大时， $b(n,p) \approx P(np)$， $b(5000, 0.001) \approx P(5)$

泊松定理： n重伯努利试验，$B(n, p_n)$，

若 $np_n \to \lambda$，则

$$\lim_{n \to \infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

例：某疾病发病率 0.001，5000人，则发病不超过5人概率：

$$P(X \le 5) \approx P(Y \le 5) = \sum_{k=0}^{5} \frac{5^k}{k!} e^{-5} = 0.616$$

4、负二项分布：独立事件，每次成功概率为p，$Nb(r,p)$

出现r次成功所需的总实验次数为X

$$P(X=k)=C_{k-1}^{r-1}p^r(1-p)^{k-r}$$ (最后一次必然成功，前k-1次可选择)

$r=1 \Rightarrow$ 退化为几何分布

$E[X]=\frac{r}{p}$ $\quad var(X)=\frac{r(1-p)}{p^2}$

(记忆：几何分布的期望和方差都乘以r)

三、期望，方差，相关系数

1. $E[(X-c)^2]=E[(X-E[X]+E[X]-c)^2] \geq var(X)$

2. 切比雪夫不等式：

$$\forall \varepsilon>0,\quad P(|X-E[X]|\geq\varepsilon)\leq\frac{Var(X)}{\varepsilon^2}$$

3. X, Y 独立 $\Rightarrow E[XY]=E[X]E[Y]$，$Var(X+Y)=Var(X)+Var(Y)$

一般的 X, Y $\Rightarrow E[aX+bY]=aE[X]+bE[Y]$

4. 协方差 $cov(X,Y)=E[XY]-E[X]E[Y]$

性质：$cov(X,c)=0$

$cov(a_1X+a_2Y,Z)=a_1cov(X,Z)+a_2cov(Y,Z)$

$cov(X,Y)^2\le Var(X)Var(Y)$

相关系数：$\rho=\frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}\in[-1,1]$

$E[XY]=E[X]E[Y]$

独立 $\Rightarrow$ 不相关 $\not\Rightarrow$ 独立

5. $g(x)$ 单调可微，则 $\begin{cases} X\text{ 与 }Y\text{ 独立} \Rightarrow g(X)\text{ 与 }g(Y)\text{ 独立} \\ X\text{ 与 }Y\text{ 不相关} \not\Rightarrow g(X)\text{ 与 }g(Y)\text{ 不相关} \end{cases}$

$P(g(X)\le x_1,g(Y)\le x_2)=P(X\le g^{-1}(x_1),Y\le g^{-1}(x_2))$

$=P(X\le g^{-1}(x_1))P(Y\le g^{-1}(x_2))=P(g(X)\le x_1)P(g(Y)\le x_2)$

两个反例：后面正态分布一节给出

6. $\rho=\pm1\Leftrightarrow X$ 与 Y 为线性关系

$Y=aX+b$ $\quad cov(X,Y)=a\,cov(X,X)=a\,var(X)$

$Var(Y)=a^2var(X)$, $\rho=\frac{a\,var(X)}{|a|\,var(X)}=\pm1$

四、条件期望（离散型）

$$E[X \mid Y=y] = \sum_i x_i \, P(X=x_i \mid Y=y)$$

例：几何分布的无记忆性 $X \sim Ge(p)$

$$E[X \mid X>k] = \sum_{m=k+1}^{+\infty} m \, P(X=m \mid X>k)$$

$$= \sum_{m=k+1}^{+\infty} m \cdot P(X=m-k) = k + E[X]$$

例：Possion 分布 $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$，相互独立

$$P(X=k \mid X+Y=n) = \frac{P(X=k, Y=n-k)}{P(X+Y=n)} = \frac{\frac{\lambda_1^k}{k!} e^{-\lambda_1} \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2}}{\frac{(\lambda_1+\lambda_2)^n}{n!} e^{-(\lambda_1+\lambda_2)}}$$

$$= C_n^k \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(1-\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-k} \sim b\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$$

$\therefore E[X \mid X+Y=n] = \frac{n\lambda_1}{\lambda_1+\lambda_2}$

可加性：X, Y 独立 $\Rightarrow X+Y \sim P(\lambda_1+\lambda_2)$

$E[X|Y] \Rightarrow$ 代表随机变量，为关于Y的函数

若我们求出 $E[X|Y=y] = f(y)$

则 $E[X|Y] = f(Y)$

重期望公式：$E[X] = E[E[X|Y]] = E[X|Y>a]P(Y>a) + E[X|Y\le a]P(Y\le a)$

Wald 方程：$X_1, X_2 \cdots X_n$ i.i.d，N 取整数

N 与 $\{X_k\}$ 独立，则 $E[\sum_{k=1}^{N} X_k] = E[X_1]E[N]$

PF：$E[\sum_{k=1}^{N} X_k] = E[E[\sum_{k=1}^{N} X_k | N]]$

$= E[N E[X_1]] = E[N]E[X_1]$

五. 连续型随机变量

1. 分布函数 $F(x) = P(X \le x)$：单调, 有界, 右连续

2. 密度函数 $p(x) = F'(x)$

3. 期望 $E[X] = \int_{-\infty}^{+\infty} x\, p(x)\, dx$

4. 方差 $var(X) = \int_{-\infty}^{+\infty} x^2 p(x)\, dx - \left(\int_{-\infty}^{+\infty} x\, p(x)\, dx\right)^2$

5. 常见分布

(1) 均匀分布 $X \sim U(a,b)$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases} \qquad p(x) = \begin{cases} \frac{1}{b-a} & a \le x < b \\ 0 & \text{其它} \end{cases}$$

$$E[X] = \int_a^b x\, p(x)\, dx = \frac{a+b}{2}$$

$$var(X) = \int_a^b x^2 p(x)\, dx - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

(2) 指数分布　$X \sim Exp(\lambda)$

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad p(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$E[X] = \frac{1}{\lambda}$, $Var(X) = \frac{1}{\lambda^2}$, $E[X^2] = \frac{2}{\lambda^2}$

无记忆性：$P(Y > s+t \mid Y > s) = P(Y > t)$

例：计算 $\int_0^{+\infty} 5x^2 e^{-3x} dx$

原式 $= \frac{5}{3} \int_0^{+\infty} x^2 \, 3e^{-3x} dx = \frac{5}{3} E[X^2]$　$X \sim Exp(3)$

$= \frac{5}{3} \cdot \frac{2}{9} = \frac{10}{27}$

(3) 正态分布　$X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$

$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$, $p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$E[X] = \mu$, $var(X) = \sigma^2$

6. 随机变量函数的分布与期望

$Y=g(X)$, $\quad F_Y(y)=P(g(X)\le y)$, $E[Y]=\int g(x)p(x)dx$

例: $X\sim N(0,1)$, 求 $Y=X^2$, $Z=|X|$ 的分布

$y\ge 0$, $F_Y(y)=P(X^2\le y)=P(-\sqrt{y}\le X\le\sqrt{y})$

$$=2(\Phi(\sqrt{y})-\Phi(0))=2\Phi(\sqrt{y})-1$$

$$p_Y(y)=\begin{cases}\varphi(\sqrt{y})\,y^{-\frac{1}{2}} & y\ge 0\\ 0 & y<0\end{cases}$$

$z\ge 0$, $F_Z(z)=P(|X|\le z)=P(-z\le X\le z)$

$$=2\Phi(z)-1$$

$$p_Z(z)=\begin{cases}2\varphi(z) & z\ge 0\\ 0 & z<0\end{cases}$$

六. 多维连续随机变量

1. 联合分布函数

$F(x_1, x_2, \cdots x_n) = P(X_1 \le x_1, \cdots X_n \le x_n)$

$P(a < X < b, c < Y < d) = F(b,d) - F(a,d) - F(b,c) + F(a,c)$

2. 联合密度函数

$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)\,du\,dv$, $p(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y)$

边缘密度函数 $p_X(x) = \int_{-\infty}^{+\infty} p(x,y)\,dy$

例: $p(x,y) = \begin{cases} 6e^{-2x-3y} & x>0, y>0 \\ 0 & 其它 \end{cases}$, 求 $P(X>Y)$

$P(X>Y) = \int_0^{+\infty} dx \int_0^{x} 6e^{-2x-3y}\,dy$

$= \int_0^{+\infty} 2e^{-2x}(1-e^{-3x})\,dx = \frac{3}{5}$

3. 独立性 $X_1, X_2, \cdots X_n$ 独立

$\Leftrightarrow p(x_1, \cdots x_n) = p(x_1)\,p(x_2) \cdots p(x_n) \Leftrightarrow P(X_1 \le x_1, \cdots X_n \le x_n) = \prod_{i=1}^{n} P(X_i \le x_i)$

例：$p(x,y)=\begin{cases} 8xy & 0\le x\le y\le 1 \\ 0 & \text{其它} \end{cases}$ 判断X，Y是否独立

$$p_X(x)=\int_x^1 8xy\,dy=4x(1-x^2)\mathbb{1}_{\{0\le x\le 1\}}$$

$$p_Y(y)=\int_0^y 8xy\,dx=4y^3\mathbb{1}_{\{0\le y\le 1\}}$$

当 $0<y<x<1$ 时，$p(x,y)=0\ne p_X(x)p_Y(y)$，故不独立

4. 期望

$X=(X_1,\cdots X_n)$，$Z=g(X_1,\cdots X_n)$

则 $E[Z]=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x_1,\cdots x_n)\,p(x_1\cdots x_n)\,dx_1\cdots dx_n$

例：$p(x,y)=\begin{cases} \frac{1}{4} & 0<x,y<2 \\ 0 & \text{其它} \end{cases}$ $Z=\min\{X,Y\}$，求 $E[Z^2]$

$$E[Z^2]=\int_0^2 dx\left(\int_0^x y^2\cdot\frac{1}{4}dy+\int_x^2 x^2\cdot\frac{1}{4}dy\right)$$

$$=\frac{2}{3}$$

七、二维正态分布 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

1、$p(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)}$

$\Downarrow$ 会读数

$X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$

但反之不成立（X, Y 正态 $\not\Rightarrow$ (X, Y) 联合正态）

2、$\rho = 0 \Leftrightarrow X, Y$ 不相关 $\Leftrightarrow X, Y$ 独立

但一般情况下 不相关 $\not\Rightarrow$ 不独立

例：$X \sim N(0, 1)$，$Y = X^2$，显然 X, Y 不独立

$\operatorname{cov}(X, Y) = E[X^3] - E[X]E[X^2] = 0 \Rightarrow X, Y$ 不相关

$\rho = \pm 1 \Rightarrow X, Y$ 线性相关，$Y = aX + b$

$\rho = 0 \Rightarrow X, Y$ 不存在线性相关，但可能非线性相关

3、$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \Rightarrow (aX + bY, cX + dY)$ 也为二元正态

$aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2)$

4、$X_1, \cdots X_n$ 独立，$X_i \sim N(\mu_i, \sigma_i^2)$（独立可加性）

则 $X_1 + \cdots + X_n \sim N(\sum \mu_i, \sum \sigma_i^2)$

(\) 条件分布与条件期望

1. $F(x \mid Y=y) = P(X \le x \mid Y=y) = \int_{-\infty}^{x} \frac{p(u,y)}{p_Y(y)} du$

$p_{X|Y}(x \mid y) = \frac{p(x,y)}{p_Y(y)}$

例：正态分布（少背）

$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$,

则 $Y \mid X=x \sim N(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2))$

2. 连续随机变量全概率公式

$p(x,y) = p_X(x)\, p(y \mid x)$

$p_Y(y) = \int_{-\infty}^{+\infty} p(x,y)\, dx = \int_{-\infty}^{+\infty} p_X(x)\, p(y \mid x)\, dx$

贝叶斯：$p(x \mid y) = \frac{p(x,y)}{p_Y(y)} = \frac{p_X(x)\, p(y \mid x)}{\int_{-\infty}^{+\infty} p_X(x)\, p(y \mid x)\, dx}$

3、独立随机变量 X, Y, $Z=X+Y$ 分布

卷积公式 $P(Z \le z) = P(X+Y \le z)$

$$= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} P_X(x) P_Y(y)\, dy = \int_{-\infty}^{+\infty} P_X(x) F_Y(z-x)\, dx$$

$$P_Z(z) = \int_{-\infty}^{+\infty} P_X(x) P_Y(z-x)\, dx$$

4、换元公式

设 $(y_1, \cdots y_n) = g(x_1, \cdots x_n)$，$g$ 有反函数

令 $J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$

则 $P_Y(y_1, \cdots y_n) = P_X(x_1, \cdots x_n)\, |J|$ → 注意行列式加绝对值

例：Box-Muller 方法　X, Y 独立 $\sim U(0,1)$

$\begin{cases} u = (-2\ln X)^{\frac{1}{2}} \cos 2\pi Y \\ v = (-2\ln X)^{\frac{1}{2}} \sin 2\pi Y \end{cases}$，则 U, V 独立 $\sim N(0,1)$

(可利用均匀分布生成正态分布)

附: $\begin{cases} x = \exp\left(-\frac{u^2+v^2}{2}\right) \\ y = \frac{1}{2\pi}\arctan\frac{v}{u} \end{cases} \Rightarrow \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = -\frac{1}{2\pi} e^{-\frac{u^2+v^2}{2}}$

$X, Y \sim U[0,1]$

$\therefore p_{UV}(u,v) = p_{XY}(x,y)\frac{1}{2\pi} e^{-\frac{u^2+v^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}$

$\therefore U, V$ 独立且 $\sim N(0,1)$

例: 独立随机变量 乘除法公式 $(Z=UV,\ W=\frac{U}{V}) \longrightarrow$ X关于Y的条件密度函数

5、条件期望 $E[X|Y] = \int x \cdot p_X(x|Y)\, dx$

例: $X_1 \sim Exp(\lambda_1)$, $X_2 \sim Exp(\lambda_2)$, X_1, X_2 独立,

求 $E[X_1 \mid X_1 < X_2]$

$$P(X_1 < X_2) = \int_0^{+\infty}\int_{x_1}^{+\infty} \lambda_1 e^{-\lambda_1 x_1} \lambda_2 e^{-\lambda_2 x_2}\, dx_2\, dx_1 = \frac{\lambda_1}{\lambda_1+\lambda_2}$$

$$P(X_1 \le x,\ X_1 < X_2) = \int_0^{x}\int_{x_1}^{+\infty} \lambda_1 e^{-\lambda_1 x_1} \lambda_2 e^{-\lambda_2 x_2}\, dx_2\, dx_1 = \frac{\lambda_1}{\lambda_1+\lambda_2}\left(1-e^{-(\lambda_1+\lambda_2)x}\right)$$

$$\therefore P(X_1 \le x \mid X_1 < X_2) = \frac{P(X_1 \le x,\ X_1 < X_2)}{P(X_1 < X_2)} = 1 - e^{-(\lambda_1+\lambda_2)x}$$

$$\therefore p_{X_1}(x \mid X_1 < X_2) = (\lambda_1+\lambda_2) e^{-(\lambda_1+\lambda_2)x}$$

$$\therefore E[X_1 \mid X_1 < X_2] = \int_0^{+\infty} x \cdot p_{X_1}(x \mid X_1 < X_2)\, dx = \frac{1}{\lambda_1+\lambda_2}$$

九、大数定律与中心极限定理

1、依概率收敛

$\{X_n\}$，X 随机变量，

$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0 \Rightarrow X_n \xrightarrow{P} X$

2、伯努利大数定律 $B(n,p)$

S_n：n次伯努利试验中发生的次数

则 $\frac{S_n}{n} \xrightarrow{P} p_0$

3、切比雪夫大数定律

$X_1, \cdots X_n \cdots$ 两两不相关，$Var(X_i) \le C$，$\forall i$，$S_n = \sum_{i=1}^{n} X_i$

则 $P\left(\left|\frac{S_n - E[S_n]}{n}\right| > \varepsilon\right) \xrightarrow{n\to\infty} 0 \quad \forall \varepsilon > 0$

Pf：$P\left(\left|\frac{S_n - E[S_n]}{n}\right| > \varepsilon\right) \le \frac{E[(S_n - E[S_n])^2]}{\varepsilon^2 n^2} \underset{\text{不相关}}{=} \frac{\sum Var(X_i)}{\varepsilon^2 n^2}$

$\le \frac{C}{n\varepsilon^2} \xrightarrow{n\to\infty} 0$

4. 中心极限定理 (Linderberg-Levi)

$\{X_n\}$ i.i.d. $E[X_i]=\mu$, $var(X_i)=\sigma^2$

则 $\frac{\sum X_i - E[\sum X_i]}{\sqrt{var(\sum X_i)}} \sim N(0,1)$ (依分布收敛)

也即 $\frac{\sum X_i - n\mu}{\sqrt{n}\,\sigma} \sim N(0,1)$

注：$\lim\limits_{n\to\infty} P\left(\frac{\sum X_i - n\mu}{\sqrt{n}\,\sigma} \le y\right) = \underset{N(0,1)}{\Phi(y)}$ ~~为依概率收敛~~

De-Moivre-Laplace：考虑 $X_n \sim B(n,p)$

$\frac{X_n - np}{\sqrt{np(1-p)}} \sim N(0,1)$

注：考试时解答题必须写清定理使用条件

(说明模型为二项分布/独立同分布)

十、泊松过程

1、计数过程

$N(t)$：$[0,t]$ 内某事件发生个数

2、独立增量性

$N(t_1)$，$N(t_2)-N(t_1)$，$N(t_3)-N(t_2)$ …… 独立

3、平稳增量性

$N(t_2)-N(t_1)$ 与 $N(t_2-t_1)$ 同分布

4、泊松过程

满足2和3的计数过程，且 $P(N(t)=k)=\frac{(\lambda t)^k}{k!}e^{-\lambda t}$

即 $N(t)\sim P(\lambda t)$

$E[N(t)]=\lambda t$，$Var(N(t))=\lambda t$

5、与二项分布的联系　　$0<s<t,\ 0\le k\le n$

$P(N(s)=k \mid N(t)=n) = C_n^k \left(\frac{s}{t}\right)^k \left(1-\frac{s}{t}\right)^{n-k}$

即 $N(t)=n$ 条件下，$N(s) \sim B(n, \frac{s}{t})$

6、与指数分布的联系

设第 n 个事件的到达时间 S_n

设第 $k-1$ 个事件发生到第 k 个事件发生的时间间隔为 T_k

则 $P(T_1 \le t) = 1 - P(T_1 > t) = 1 - P(N(t)=0) = 1 - e^{-\lambda t}$

$\therefore T_k \sim T_1 \sim Exp(\lambda)$ i.i.d

$S_n = \sum_{k=1}^{n} T_k$　　$S_n \le t \Leftrightarrow N(t) \ge n$

由指数分布可加性知 $S_n \sim Gamma(n, \lambda)$

7、与均匀分布的联系

给定 $N(t)=n$ 条件下，n 个时间的到达时刻 $S_1, \cdots S_n$

与 n 个独立的 $(0, t)$ 上的均匀分布的顺序统计量

具有相同分布

8. 泊松过程的分裂(随机选择不变性)。

顾客数 $\sim$ Possion (λ)

每个顾客买电视机概率为 P，λ, p, d，

求售出 k 台电视机的概率

$$P(X=k)=\sum_{n=k}^{+\infty} P(X=k \mid N=n)\, P(N=n)$$

$$=\sum_{n=k}^{+\infty} C_n^k p^k (1-p)^{n-k} \cdot \frac{\lambda^n e^{-\lambda}}{n!}=\frac{(\lambda p)^k}{k!} e^{-\lambda p}$$

$\therefore X \sim P(\lambda p)$

$X \sim$ Possion (λ)，X 中个体中逐个独立的以概率为 p 保留

筛选得到 Y 个个体，则 $Y \sim$ Possion (λp)

注：每分钟 k 次 $\rightarrow$ 强度为 $\lambda=k$ 的 Possion 过程

X_t 为 t 分钟到达的人数，服从 $P(\lambda t)$ Possion 分布

十一、马尔可夫链

1. 概念

$P(X_{n+1}=j \mid X_0=i_0, \cdots X_n=i) = P(X_{n+1}=j \mid X_n=i)$

时齐性：$P_{ij} = P(X_{n+1}=j \mid X_n=i)$ 与 n 无关

2. C—K 方程　　$P_{ij}^{(n)} = P(X_{m+n}=j \mid X_m=i)$

$P^{(m+n)} = P^{(m)}P^{(n)} = P^{m+n}$　　n 步转移概率

$P_{ij}^{(m+n)} = \sum_k P_{ik}^{(m)} P_{kj}^{(n)}$

3. 状态的分类

常返：有限步转移后最终能回到 i（从 i 出发）

暂态：有限步后不一定能回到 i

首达概率：$f_{ij}^{(n)} = P(X_n=j, X_k \neq j, k=1 \cdots n-1 \mid X_0=i)$

$f_{ij} = \sum_n f_{ij}^{(n)}$

i 常返 $\Leftrightarrow f_{ii}=1$

4. 可达与互通

$i \leftrightarrow j$：$\exists n_1, n_2$，$P_{ij}^{(n_1)} > 0$，$P_{ji}^{(n_2)} > 0$

i常返且$i \to j \Rightarrow i \leftrightarrow j$且$j$常返

可分解为若干连通分支

5. 周期：i出发，只可能在d的整数倍次回到i

则称d为i的周期

6. 稳态收敛定理

不可约(单个常返类)非周期马氏链存在唯一平稳分布

且收敛到这个分布　　$\pi = \pi P$

7. i的平均返回时间 $\mu_i = \sum_{n=1}^{+\infty} n f_{ii}^{(n)}$

$\mu_i < +\infty$ 正常返

$\mu_i = +\infty$ 零常返

十二、蒙特卡洛方法

布丰投针

$$\iint_D f(x)\,dx \approx S_D \cdot \frac{f(x_1)+\cdots+f(x_n)}{n}$$

例：若能生成密度函数为 $2x\,I_{\{0<x<1\}}$ 的随机数

求 $\int_0^1 e^{x^2}dx$

$$\int_0^1 e^{x^2}dx = \int_0^1 \frac{e^{x^2}}{2x}\,2x\,dx = E\left[\frac{e^{x^2}}{2x}\right]$$

$$\approx \frac{\sum_i \frac{e^{x_i^2}}{2x_i}}{n}$$

例：通过 $U[0,1]$ 得到 $p(x)=2x\,I_{\{0<x<1\}}$ 的伪随机数

$F(x)=\int p(x)=x^2 \Rightarrow X=F^{-1}(U)$

$\therefore \{x_i\} \sim U[0,1] \Rightarrow \{\sqrt{x_i}\}$ 为密度为 $f(x)$ 的伪随机数.

十、数理统计相关概念

样本：$\{X_i\}$ 为样本观察值，服从独立同分布假设

样本均值：$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

样本方差：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ ⇒ 无偏

样本矩：$a_k = E[X^k]$　$\hat{a}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$（原点矩）

经验分布函数

$S(x) = X_1, \cdots X_n$ 中不大于 x 的样本个数

$F_n(x) = \frac{S(x)}{n} \xrightarrow{n\to\infty} F(x) = P(X \le x)$

十一、次序统计量

$X_1 \cdots X_n \sim X$，$X_{(i)}$ → 由小到大排的第 i 个

$X_{(k)}$ 密度函数 $P_k(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1}(1-F(x))^{n-k} p(x)$

$(X_{(i)}, X_{(j)})$ 联合密度函数

$p_{i,j}(y, z) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F(y)^{i-1}(F(z)-F(y))^{j-i-1}(1-F(z))^{n-j} p(y)p(z)$

十二、点估计

1、矩估计

令 $E[X] \approx \frac{1}{n}\sum X_i = \overline{X}$

$E[X^2] \approx \frac{1}{n}\sum X_i^2$

$\Rightarrow var(X) \approx \frac{1}{n}\sum X_i^2 - \overline{X}^2 \approx S^2$

例：$X_1, \ldots X_n \overset{iid}{\sim} U(a,b)$，求a和b的矩估计

$E[X] = \frac{a+b}{2} \approx \overline{X}$, $var(X) = \frac{(a-b)^2}{12} \approx S^2$

$\Rightarrow \begin{cases} \hat{a} \approx \overline{X} - \sqrt{3}S \\ \hat{b} \approx \overline{X} + \sqrt{3}S \end{cases}$

2、最大似然估计

$P(A|\theta)$　　A发生，则 θ 应为使 $P(A|\theta)$ 最大的值

(1) 离散型　$A = \{X_1 = x_1, \ldots X_n = x_n\}$

$P(A|\theta) = P_\theta(X_1 = x_1) \cdots P_\theta(X_n = x_n)$

(2) 连续型 $P(A|\theta) \approx P(x_1|\theta)\, P(x_2|\theta) \cdots P(x_n|\theta)$

步骤：① $L(\theta)=\prod_{i=1}^{n} P(x_i|\theta)$

② $\frac{d\ln L(\theta)}{d\theta}=0$ 解方程

例：Possion(λ)．

$$L(\lambda)=\prod_i P(X=x_i|\lambda)=\prod_i \frac{\lambda^{x_i}}{x_i!}e^{-\lambda}$$

$$=e^{-n\lambda}\frac{\lambda^{x_1+\cdots+x_n}}{\prod_i x_i!}$$

$$\ln L(\lambda)=\sum_i x_i\ln\lambda-n\lambda-\sum_i \ln(x_i!)$$

$$\frac{d\ln L(\lambda)}{d\lambda}=\frac{1}{\lambda}\sum_i x_i-n=0\Rightarrow\hat{\lambda}=\bar{x}$$

注：MLE 具有不变性 $\theta\to\hat{\theta}$

单调函数 $g(\theta)$ 的 MLE 是 $g(\hat{\theta})$

例：$U[0,\theta]$ $L(\theta)=\begin{cases}\frac{1}{\theta^n} & 0<\max\{x_i\}<\theta \\ 0 & \text{其它}\end{cases}$

$\therefore$ $\hat{\theta}=\max\{x_i\}$ 使 $L(\theta)$ 最大

十三、估计量的评价标准

1、相合性

$$\hat{\theta}_n \xrightarrow[p]{n\to+\infty} \theta$$

例：$X_i \sim b(m,p)$，$\hat{p} = \frac{\bar{X}}{m}$

则 $\bar{X} \xrightarrow{P} E[X] = mp$ （由大数定律）

$\hat{p} = \frac{\bar{X}}{m} \xrightarrow{P} p$ ⇒ 相合

2、无偏性：$E[\hat{\theta}] = \theta$

3、有效性：同为无偏估计，方差越小越有效

4、均方误差：

$$MSE(\hat{\theta}) = E[(\hat{\theta}-\theta)^2]$$

$$= \underbrace{E[(\hat{\theta}-E(\hat{\theta}))^2]}_{方差} + \underbrace{(E[\hat{\theta}]-\theta)^2}_{偏差}$$

无偏估计：$MSE[\hat{\theta}] = Var(\hat{\theta})$

十四、区间估计

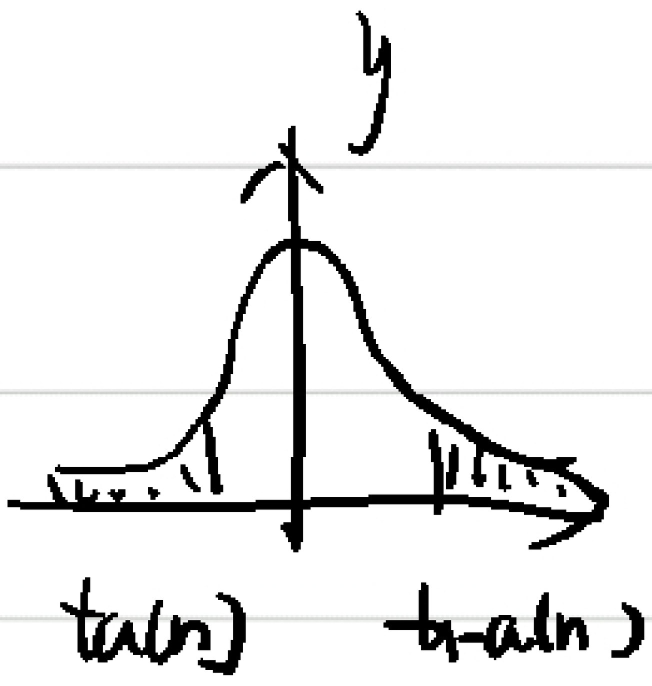
1. 三大抽样分布　$X_i \sim N(0,1)$

$$\chi^2(n) = \sum_{i=1}^{n} X_i^2$$

$$F(n_1, n_2) = \frac{\frac{\chi^2(n_1)}{n_1}}{\frac{\chi^2(n_2)}{n_2}}$$

$$t(n) = \frac{N(0,1)}{\sqrt{\frac{\chi^2(n)}{n}}}$$

2. 分位数　$P(X \le \chi^2_{1-a}(n)) = 1-a$



$$F_a(n_2, n_1) = \frac{1}{F_{1-a}(n_1, n_2)} \quad \left(\frac{1}{F(n_1, n_2)} = F(n_2, n_1)\right)$$

$t_{1-a}(n) = -t_a(n)$ → t分布(的密度函数)关于y轴对称

3, 重要结论

$x_1, \cdots x_n \sim N(\mu, \sigma^2)$，则

① $\bar{X}$ 与 S^2 独立

② $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$

③ $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(0,1)$

④ $\frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t(n-1)$

$x_i \sim N(\mu_1, \sigma_1^2)$，$y_i \sim N(\mu_2, \sigma_2^2)$

则 $\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n_1-1, n_2-1)$

4. 区间估计

$P(\hat{\theta}_L \le \theta \le \hat{\theta}_U) = 1-\alpha$, $[\hat{\theta}_L, \hat{\theta}_U]$

(1) $X_i \sim N(\mu, \sigma^2)$, σ 已知, 求 μ 的 $1-\alpha$ 置信区间

$$u = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$$

使得置信区间最短

$$1-\alpha = P(-u_{1-\frac{\alpha}{2}} \le u \le u_{1-\frac{\alpha}{2}})$$
$$= P\left(-u_{1-\frac{\alpha}{2}} \le \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \le u_{1-\frac{\alpha}{2}}\right)$$
$$= P\left(\bar{X} - \frac{\sigma u_{1-\frac{\alpha}{2}}}{\sqrt{n}} \le \mu \le \bar{X} + \frac{\sigma u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$$

$\therefore$ μ 的置信区间为 $\left[\bar{X} - \frac{\sigma u_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X} + \frac{\sigma u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right]$

(2) σ 未知, $t = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t(n-1)$

(3) 估计 σ^2: $\eta = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ $\Rightarrow$ 注意 χ^2 不对称

用 $[\chi^2_{\frac{\alpha}{2}}, \chi^2_{1-\frac{\alpha}{2}}]$

(4) 双正态 = 背 利用 $F^2(n_1, n_2)$

十五、假设检验

1. 检验法则与拒绝域

H_0 H_1

拒绝域 W → 由样本 $X_1, \cdots X_n$ 决定

观测值 $(x_1, \cdots x_n) \in W$ = 拒绝 H_0

2. 两类错误

$$\begin{cases} \alpha = P(\text{拒绝 } H_0 \mid H_0 \text{真}) \\ \beta = P(\text{接受 } H_0 \mid H_0 \text{假}) \end{cases}$$

3. 显著性水平 α = $P(\text{拒绝 } H_0 \mid H_0 \text{真}) \le \alpha$ 条件下

尽可能让 β 小

4. 单正态假设检验 $X_i \sim N(\mu, \sigma^2)$

(1) 若 $H_0: \mu = \mu_0$, $H_1: \mu \ne \mu_0$ (双边)

则令 $u = \dfrac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma} \sim N(0,1)$

双边：$p(|u| \geqslant u_{1-\frac{\alpha}{2}}) = \alpha$

(2) 若 H_0：$\mu \leq \mu_0$，H_1：$\mu > \mu_0$

或 H_0：$\mu = \mu_0$，H_1：$\mu > \mu_0$ (单边)

$\Downarrow$

$p(u \geqslant u_{1-\alpha}) = \alpha$

∴ 当 $u = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma} \geqslant u_{1-\alpha}$ 时拒绝

记忆：$\bar{x}\uparrow \Rightarrow u\uparrow \rightarrow \mu$ 比 μ_0 大 → 拒绝

若 H_0：$\mu \geqslant \mu_0$；H_1：$\mu < \mu_0$

$\Rightarrow p(u \leq u_\alpha) = \alpha \Rightarrow$ $u \leq u_\alpha$ 时拒绝

注：应用时注意检验方式不要写错

例：检验集合量是否显著偏低

要用 H_0：$\mu = a$，H_1：$\mu < a$

不能用 H_0：$\mu \leq a$，H_1：$\mu > a$

(3) σ 未知 : t 检验 $t = \frac{\sqrt{n}(\bar{x}-\mu_0)}{S} \sim t(n-1)$

$P(|t| \geq t_{1-\frac{\alpha}{2}}(n-1)) = \alpha$

(4) 检验 σ^2 : $H_0: \sigma^2 = \sigma_0^2$, $H_1: \sigma^2 \neq \sigma_0^2$

$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

5. 双正态 : 用 F 分布

6. P值 : 能用观测值作出拒绝原假设的最小的 α (使得观测值落在拒绝域中)