- independence

① Prove: If A is independent of B then A is independent of B'.

$P(A\cap B)=P(A)P(B)$ $\qquad$ $P(A)=P(A\cap B)+P(A\cap B')$

$\qquad\qquad\qquad\qquad\qquad\qquad =P(A)P(B)+P(A\cap B')$

$\Rightarrow \quad P(A)(1-P(B))=P(A)P(B')=P(A\cap B')$

② $E_1,\cdots E_n$ are independent if for every subset $E_1',\cdots E_r'$ $(r\le n)$

$P(E_1'\cap\cdots E_r')=P(E_1')\cdots P(E_r')$

$\Rightarrow P(E_1\mid f(E_2,\cdots E_n))=\dfrac{P(E_1)P(f(E_2,\cdots,E_n))}{P(f(E_2,\cdots E_n))}=P(E_1)$

$f(E_2,\cdots,E_n)$ is an event obtained by performing basic operations on sets $E_2,\cdots,E_n$.

pairwise independent 两两独立

$\Uparrow$

mutually independent 相互独立

$\forall\, i_1,i_2\cdots i_k\le 1,2,\cdots,n$ $\quad P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k})$

conditionally independent 条件独立 在已知某个特定信息后,两事件无关联

$P(A|C)\cdot P(B|C)=P(A\cap B|C)$

- the law of total probability

$P(A)=\sum_{i=1}^{n}P(A\cap B_i)=\sum_{i=1}^{n}P(A|B_i)P(B_i)$

- Bayer's rule

$$P(B_j|A)=\frac{P(A\cap B_j)}{P(A)}=\frac{P(A|B_j)\cdot P(B_j)}{\sum_{i=1}^{n}P(A|B_i)P(B_i)}$$

[e.g.] C: you have cancer $\qquad$ T: you test positive

$P(T|C)=0.95$ $\qquad P(T'|C')=0.95$ $\qquad P(C)=0.004$

compute $P(T)$. $P(C|T)$

$P(T)=P(T|C)\cdot P(C)+P(T|C')\cdot P(C')$

$\quad =P(T|C)\cdot P(C)+(1-P(T'|C'))\cdot(1-P(C))=0.0536$

$$P(C|T)=\frac{P(C\cap T)}{P(T)}=\frac{P(C)\cdot P(T|C)}{P(T)}=0.07089$$

· the r-th moment [矩]

1. 矩估计

		样本 (A_r)	总体
① $E(X^r)$	r阶原点矩	$\frac{1}{n}\sum_{i=1}^{n} x_i$	$E(X)$
		$\frac{1}{n}\sum_{i=1}^{n} x_i^r$	$E(X^r)$
② $E[(X-b)^r]$	r阶中心矩	$\frac{1}{n}\sum_{i=1}^{n}(x_i-b)^r$	$E[(X-E(X))^k]$

r=2 $E[(X-E(X))^2] = D[X]$

③ $E[(X)_r] \triangleq E[X(X-1)\cdots(X-r+1)]$ r-th factorial Moment

2. moment generating function [矩母函数]

X be a discrete random variable. if there exists a $h>0$ s.t.

$$E[e^{tX}] = \sum_{i\in N} e^{tx_i} f(x_i) \quad \text{exists} \quad \forall t \in (-h,h)$$

$\Rightarrow M(t) = E[e^{tX}]$, $t\in(-h,h)$ is the mgf of X

(连续型: $E[e^{tX}] = \int_{-\infty}^{\infty} f(x)e^{tx}\,dx$)

$$M'(t) = \frac{dE(e^{tX})}{dt} = E[e^{tX}\cdot X] \qquad M'(0) = E(X)$$

$$\vdots \qquad\qquad \vdots$$

$$M^{(r)}(t) = E[X^r e^{tX}] \qquad M^{(r)}(0) = E[X^r]$$

$Var[X] = M''(0) - [M'(0)]^2$

[e.g.1] $X \sim U(a,b)$ (a,b未知), $X_1,\cdots X_n$ 为来自总体的简单随机样本. 试用矩估计法估计a,b.

① $\overline{X} = EX = \frac{a+b}{2}$

② $\begin{cases} A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = E(X^2) = \int_a^b x^2 \frac{1}{b-a}\,dx = \frac{a^2+ab+b^2}{3} \\ B_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 = DX = \frac{(b-a)^2}{12} \text{ (better)} \end{cases}$

$\Rightarrow \begin{cases} \hat{a} = \overline{X} - \sqrt{3B_2} \\ \hat{b} = \overline{X} + \sqrt{3B_2} \end{cases}$ *由于B_2可由$X_1,\cdots X_n$推出, 故简写为B_2

3. X is a nonnegative random variable and $c>0$, then $P(X\geq c)\leq \frac{E[X]}{c}$

Proof:

$$E[X]=\sum_{k\geq 0} k f(k) \geq \sum_{k\geq c} c f(k) \geq c\sum_{k\geq c} f(k) = c\,P(X\geq c)$$

$$\Rightarrow \frac{E[X]}{c} \geq P(X\geq c)$$

- Distribution

1. Standardized random variable

$$Z=\frac{X-\mu}{\sigma} \quad (E[X]=\mu,\ Var[X]=\sigma^2)$$

$$\begin{cases} E[Z]=\frac{1}{\sigma}[E(X)-\mu]=\frac{1}{\sigma}(\mu-\mu)=0 \\ var(Z)=E[(Z-E(Z))^2]=E[Z^2]=E[\frac{1}{\sigma^2}(X-\mu)^2] \end{cases}$$

$$=\frac{1}{\sigma^2}E[(X-\mu)^2]=\frac{1}{\sigma^2}\cdot\sigma^2=1$$

2. Bernoulli Distribution $f(x)=p^x(1-p)^{1-x}$ $x\in\{0,1\}$

$E[X]=1\cdot p+0\cdot(1-p)=p$ $Var[X]=(1-p)^2\cdot p+(0-p)^2(1-p)=p(1-p)$

$M(t)=E[e^{tX}]=\boxed{e^t\cdot p+(1-p)}$ $t\in\mathbb{R}$

3. Binomial Distribution $f(X=k)=\binom{n}{x}p^x(1-p)^{n-x}$

$E[X]=\sum_{k=0}^{n} kP(X=k)=np$ $Var[X]=np(1-p)$

$$M(t)=\sum_{k=0}^{n} e^{tk}\cdot\binom{n}{k}p^k(1-p)^{n-k}=\boxed{[pe^t+(1-p)]^n}$$

$E[X]=M'(0)=np$ $E[X^2]=n(n-1)p^2+np$ $Var[X]=np(1-p)$

4. Hypergeometric Distribution $f(x)=\frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{x}}$ $N=N_1+N_2$

$E[X]=n\frac{N_1}{N}$ $Var[X]=n\frac{N_1}{N}\cdot\frac{N_2}{N}\cdot\frac{N-n}{N-1}$

when $N>>n$, $p=\frac{N_1}{N}$, then $E[X]=np$ $Var[X]=np(1-p)$

5. Geometric Distribution $f(x)=(1-p)^{x-1}\cdot p$

$E[X]=\sum_{x=1}^{\infty}(1-p)^{x-1}\cdot x\cdot p=\frac{1}{p}$ $Var[X]=\frac{1-p}{p^2}$

E[X]	1	E(X)+1
P(x)	p	1-p

$$M(t)=\boxed{\frac{pe^t}{1-(1-p)e^t}} \quad (t<-\ln(1-p))$$

$E[X]=p+(1-p)(E[X]+1)\Rightarrow E[X]=\frac{1}{p}$

* Negative Binomial Distribution

$r=1$ 几何分布

X: the trial number at which the r-th success is observed.

$$P(X=x) = \binom{x-1}{r-1} \cdot p^{r-1} (1-p)^{(x-1)-(r-1)} \cdot p = \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad (x \geq r)$$

$$E[X_r] = p \cdot (1+E[X_{r-1}]) + (1-p)(1+E[X_r])$$

$$E[X_r] = \frac{1}{p} + E[X_{r-1}] \Rightarrow E[X_r] = \frac{r}{p}$$

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

$$M(t) = E[e^{tX}] = \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

$$= p^r e^{tr} \sum_{x=r}^{\infty} \binom{x-1}{r-1} [e^t(1-p)]^{x-r}$$

$$= \boxed{\frac{(pe^t)^r}{[1-e^t(1-p)]^r}}$$

[if $|e^t(1-p)|<1$, i.e. $t<-\ln(1-p)$]

$$*\ (1-w)^{-r} = [1+(-w)]^r = \sum_{k=0}^{\infty} \binom{-r}{k} \cdot (-w)^k$$

$$= \sum_{k=0}^{\infty} (-1)^k \cdot \binom{r+k-1}{k} \cdot (-1)^k \cdot w^k$$

$$= \sum_{k=0}^{\infty} \binom{r+k-1}{k} w^k$$

$$= \sum_{x=r}^{\infty} \binom{x-1}{x-r} w^{x-r} = \sum_{x=r}^{\infty} \binom{x-1}{r-1} w^{x-r}$$

6. Possion Distribution

① $[0,t_1]$ $[t_1,t_2]$ 事件独立 ② $\lim_{h\to 0} \frac{P(1 \text{ event in } h)}{h} = \lambda$ ③ $\lim_{h\to 0} \frac{P(\geq 2 \text{ events in } h)}{h} = 0$

$$P(X=x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

$$p = \frac{\lambda}{n}. \quad P(X=x) = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} \qquad n\to\infty$$

$$\lim_{n\to\infty} P(X=x) = \frac{n!}{x!(n-x)!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

$$\lim_{n\to\infty} = \frac{n^x}{x!} \cdot \frac{\lambda^x}{n^x} \cdot e^{-\lambda} \cdot 1 = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

$$M(t)=\sum_{x=0}^{\infty} e^{tx} e^{-\lambda}\cdot\frac{\lambda^x}{x!}$$

$$=e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^t\lambda)^x}{x!}=e^{-\lambda}\cdot e^{e^t\lambda}=\boxed{e^{\lambda(e^t-1)}}$$

$E[X]=M'(0)=\lambda \qquad Var[X]=E[X^2]-(E[X])^2=\lambda$

7. Uniform Distribution $E[X]=\frac{1}{2}$. $Var[X]=\frac{1}{12}$ $\quad M[t]=\begin{cases}\frac{e^t-1}{t} & t\neq 0\\ 1 & t=0\end{cases}$

X is uniform $[0,1] \Rightarrow Y=a+(b-a)X$ is uniform $[a,b]$

pdf $=\frac{d}{dy}F_Y(y)=\frac{d}{dy}\left(\frac{y-a}{b-a}\right)=\frac{1}{b-a} \quad (a<y<b)$ $\qquad M_Y(t)=e^{at}M_X((b-a)t)$

$E[Y]=a+\frac{1}{2}(b-a)=\frac{a+b}{2}$

$Var[Y]=(b-a)^2\,Var[X]=\frac{(b-a)^2}{12}$

$$M(t)=\int_a^b e^{ty}\frac{1}{b-a}dy=\frac{1}{b-a}\cdot\frac{1}{t}(e^{tb}-e^{ta})=\begin{cases}\boxed{\frac{e^{tb}-e^{ta}}{t(b-a)}} & t\neq 0\\ 1 & t=0\end{cases}$$

8. (100p)th Percentile

$$p=\int_{-\infty}^{\pi_p} f(x)\,dx=F(\pi_p)$$

[e.g.] $Y\sim U(a,b)$

$$p=\int_{-\infty}^{\pi_p} f(x)\,dx=\int_{-\infty}^{\pi_p}\frac{1}{b-a}dx=\frac{\pi_p-a}{b-a} \Rightarrow \pi_p=p(b-a)+a$$

9. Exponential Distribution

指数分布 事件的时间间隔概率 [计时]

泊松分布 某段时间内,事件具体发生的概率 [计数]

$f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}} \quad (0\leq x<\infty,\ \theta>0)$

$P(X>k)=\int_k^{\infty}\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=e^{-\frac{k}{\theta}}$

CDF:

$$F(x)=\begin{cases}0 & -\infty<x<0\\ 1-e^{-\frac{x}{\theta}} & 0\leq x<\infty\end{cases}$$

$E[X]=\theta,\ Var[X]=\theta^2$

$$M(t)=\int_0^{\infty}e^{tx}\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=\boxed{\frac{1}{1-t\theta}} \quad \left(t<\frac{1}{\theta}\right)$$

$$P(X>t+s\mid X>s)=\frac{P(X>t+s)}{P(X>s)}=e^{-\frac{t}{\theta}}=P(X>t) \quad \text{无记忆性}$$

[几何分布 $P(X>i+j\mid X>i)=\frac{(1-p)^{i+j}}{(1-p)^i}=(1-p)^j=P(X>j)$]

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

[e.g.] Customer arrive in a certain shop according to a possion process at a mean rate of 20 per hour. P(the shopkeeper will have to wait more than 5 minutes for the arrival of the first customer?

$\lambda = \frac{1}{3}$ person per minute 发生率 $\theta = \frac{1}{\lambda} = 3$

$\Rightarrow f(x) = \frac{1}{3} e^{-\frac{x}{3}} \ (x \geq 0)$ $\quad F(x) = 1 - e^{-\frac{x}{3}}$ $\quad \Rightarrow P(X>5) = 1 - F(5) \approx 0.189$

10. Erlang Distribution [第α次发生所需时间]

CDF

$$F_W(t) = 1 - \underbrace{P(W>t)}_{P(Y<\alpha)} = 1 - \sum_{k=0}^{\alpha-1} e^{-\lambda t} \frac{\lambda^k t^k}{k!}$$

PDF

$$f_W(t) = F'_W(t) = \frac{\lambda^\alpha}{(\alpha-1)!} t^{\alpha-1} e^{-\lambda t} \quad (t \geq 0)$$

$$= \frac{1}{(\alpha-1)!\,\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} \quad (0 \leq x < \infty). \quad \theta = \frac{1}{\lambda}$$

平均时间间隔 $= \frac{1}{\lambda}$

$E[W] = \alpha\theta \qquad Var[W] = \alpha\theta^2$

11. Gamma Distribution $\alpha \in Z \Rightarrow \alpha \in R$

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}}$$

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty \frac{1}{\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} = 1 \Rightarrow \Gamma(\alpha) = \int_0^\infty \frac{1}{\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} dx$$

$$= \int_0^\infty y^{\alpha-1} e^{-y} dy$$

① $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

$$\Gamma(\alpha) = [-y^{\alpha-1} e^{-y}]_0^\infty + \int_0^\infty (\alpha-1) y^{\alpha-2} e^{-y} dy = (\alpha-1) \int_0^\infty y^{\alpha-2} e^{-y} dy = (\alpha-1)\Gamma(\alpha-1)$$

② $\alpha = n$. $\Gamma(\alpha) = (n-1)!$

$$\Gamma(n) = (n-1)\,\Gamma(n-1) \cdots = (n-1)(n-2) \cdots 1 \cdot \Gamma(1) = (n-1)!$$

$$\Gamma(1) = \int_0^\infty e^{-y} dy = -e^{-y}\Big|_0^\infty = 1$$

$$M(t) = E[e^{tX}] = \int_0^\infty e^{tx} \cdot \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} dx = \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^\infty x^{\alpha-1} \exp[-(\tfrac{1}{\theta}-t)x] dx$$

let $\beta = \frac{1}{\theta} - t$, $M(t) = \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^\infty x^{\alpha-1} e^{-\beta x} dx$ $(t < \frac{1}{\theta})$

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy,\ M(t) = \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot \int_0^\infty (\tfrac{y}{\beta})^{\alpha-1} e^{-y} (\tfrac{dy}{\beta}) = \frac{1}{\beta^\alpha} \cdot \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot \Gamma(\alpha)$$

let $y = \beta x$ [x, y $(0, \infty)$ 不变]

$$= \frac{1}{(\theta\beta)^\alpha} = \frac{1}{(1-\theta t)^\alpha}$$

$E[X] = \alpha\theta$. $Var[X] = \alpha(\alpha+1)\theta^2 - [\alpha\theta]^2 = \alpha\theta^2$ $\quad E[Y^k] = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \theta^k = \frac{(\alpha+k-1)!}{(\alpha-1)!} \theta^k$

12. Chi-square Distribution 卡方分布 $\quad X \sim \chi^2(r)$

$f(x) = \frac{1}{\Gamma(\frac{r}{2}) 2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}$ $\quad 0 < x < \infty$ $\quad$ [在Gama分布中取 $\alpha = \frac{r}{2}$, $\theta = 2$]

$M(t) = (1-2t)^{-\frac{r}{2}}$ $\quad \mu = \alpha\theta = r$ $\quad \sigma^2 = \alpha\theta^2 = 2r$

$X_1 \sim X_n$ 相互独立. $X_i \sim N(0,1)$ $\quad X = \sum_{i=1}^n X_i^2 \sim \chi^2(n)$ $\quad n$: degrees of freedom

$E[X^2] = n$, $D(\chi^2) = 2n$

13. Normal Distribution $\quad X \sim N(\mu, \sigma^2)$

$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$ $\quad x \in \mathbb{R}$

$$M(t) = \int_{-\infty}^\infty \frac{e^{tx}}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) dx$$

$$= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\tfrac{1}{2\sigma^2}(x^2 - 2(\mu+\sigma^2 t)x + \mu^2)] dx$$

$X \sim N(\mu+\sigma^2 t, \sigma^2)$

$$= \exp\left(\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}\right) \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[\frac{-[x-(\mu+\sigma^2 t)]^2}{2\sigma^2}\right] dx$$

$$= \exp(\mu t + \frac{\sigma^2 t^2}{2})$$

$E[X] = M'(0) = \mu$. $\quad Var[X] = (\mu^2 + \sigma^2) - \mu^2 = \sigma^2$

* Standard normal distribution $\quad X \sim N(0,1)$

$Z_\alpha \Rightarrow P(Z \geq Z_\alpha) = \alpha$ $\quad Z_\alpha = \phi^{-1}(1-\alpha)$ $\quad 100(1-\alpha)$-th percentile

$X \sim N(u, 6^2)$. then $Z = \frac{X-u}{6}$ is $N(0,1)$

Proof:

$$P(Z \le z) = P\left(\frac{X-u}{6} \le z\right) = P(X \le z6+u) = \int_{-\infty}^{z6+u} \frac{1}{6\sqrt{2\pi}} \exp\left[-\frac{(x-u)^2}{26^2}\right] dx$$

let $w = \frac{x-u}{6}$, then $P(Z \le z) = \int_{-\infty}^{z} \frac{1}{6\sqrt{2\pi}} e^{-\frac{w^2}{2}} \cdot 6\, dw$

$$= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw \Rightarrow u' = 0,\ 6'^2 = 1$$

$$* P(a \le X \le b) = P\left(\frac{a-u}{6} \le \frac{X-u}{6} \le \frac{b-u}{6}\right) = \phi\left(\frac{b-u}{6}\right) - \phi\left(\frac{a-u}{a}\right)$$

* Normal and χ^2 Distribution

$X \sim N(u, 6^2)$ $\quad V = \frac{(X-u)^2}{6^2} = Z^2$ is $\chi^2(1)$ $\quad V \sim \chi^2(1)$

Proof:

$G(v)$: the cdf $\quad G(v) = P(Z^2 \le v) = \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \int_0^v \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} dy$

$(z = \sqrt{y})$

$$g(v) = G'(v) = \frac{1}{\sqrt{2\pi v}} e^{-\frac{v}{2}}$$

$$= \frac{1}{\sqrt{\pi} \cdot \sqrt{2}} v^{\frac{1}{2}-1} \cdot e^{-\frac{v}{2}}$$

$X \sim \chi^2(1)$: $f(x) = \frac{1}{\Gamma(\frac{1}{2})\sqrt{2}} x^{\frac{1}{2}-1} e^{-\frac{x}{2}}$

$$\int_0^\infty g(v)\, dv = 1 \Rightarrow \frac{1}{\sqrt{2\pi}} \int_0^\infty (2x)^{\frac{1}{2}-1} e^{-x} \cdot 2dx$$

$\int_0^\infty x^{\alpha-1} e^{-x} dx$

$$= \frac{1}{\sqrt{\pi}} \int_0^\infty x^{\frac{1}{2}-1} \cdot e^{-x} dx = \frac{1}{\sqrt{\pi}} \Gamma(\alpha) = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1$$

$\Rightarrow \Gamma(\frac{1}{2}) = \sqrt{\pi}$

$\Rightarrow$ cdf $g(v)$ $\chi^2(r)$ $r=1$ 的情況

• Bivarite Distribution of Discrete Type [二维]

1. Marginal Probability mass function 边际概率质量函数

$f_X(x) = \sum_y f(x,y) = P(X=x)$

independent $f(x,y) = f_X(x) \cdot f_Y(y)$ ☆点、和集合均可证明独立

① X, Y independent $\Rightarrow P(X\in A, Y\in B) = P(X\in A)P(X\in B)$

$$P(X\in A, Y\in B) = \sum_{y\in B} P(X\in A, Y=y)$$
$$= \sum_{y\in B}\sum_{x\in A} P(X=x, Y=y) = \sum_{y\in B}\sum_{x\in A} P(X=x)\cdot P(Y=y)$$
$$= \sum_{y\in B} P(X\in A)P(Y=y) = P(X\in A)P(Y\in B)$$

② 集合独立 ⇒ 点独立

取 $A=\{x\}$, $B=\{y\}$ 利用单点集即可证明

2. Joint cumulative distribution function

$F(x,y) = P(X\le x, Y\le y)$ [X, Y independent $\Leftrightarrow F(x,y) = F_X(x)\cdot F_Y(y)$]

Lemma: $F(x,y) = F_X(x)\cdot F_Y(y) \Rightarrow \{X=x_i\}, \{Y\le y\}$ independent $\forall x_i \in S_X, y$

$P(X=x_i, Y\le y) = P(X\le x_i, Y\le y) - P(X\le x_{i-1}, Y\le y)$ (假设 x, y 已升序排列)

$= [P(X\le x_i) - P(X\le x_{i-1})]P(Y\le y) = P(X=x_i)P(Y\le y)$ $(i=1, x_0=x_1-1)$

① $F(x,y) = F_X(x)F_Y(y) \Rightarrow X, Y$ independent

$P(X=x_i, Y=y_j) = P(X=x_i, Y\le y_j) - P(X=x_i, Y\le y_{j-1}) = P(X=x_i)P(Y=y_j)$

i.e. $f(x_i, y_j) = f_X(x_i) f_Y(y_j)$ $(j=1, y_0=y_1-1)$

$$F(x,y) = \int_{s=-\infty}^{x}\int_{t=-\infty}^{y} f(s,t)\,dt\,ds = \int_{s=-\infty}^{x}\int_{t=-\infty}^{y} f_X(s)\cdot f_Y(t)\,dt\,ds$$
$$= \int_{s=-\infty}^{x} f_X(s)\int_{t=-\infty}^{y} f_Y(t)\,dt\,ds$$
$$= F_X(x)\,F_Y(y)$$

② X, Y independent $\Rightarrow F(x,y) = F_X(x)F_Y(y)$

$F(x,y) = P(X\le x, Y\le y) = P(X\le x)P(Y\le y) = F_X(x)\cdot F_Y(y)$

3. $E[X] = \sum_{x \in S_X} x f_X(x) = \sum_{(x,y) \in S} x f(x,y)$

$E[u(X_1, X_2)] = \sum\sum_{(x_1,y_1) \in S} u(x_1, x_2) \cdot f(x_1, x_2)$

4. covariance

$\operatorname{cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY}$

$= E[XY - \mu_X Y - \mu_Y Y + \mu_X \mu_Y]$

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

- $\operatorname{cov}(X,Y) \begin{cases} > 0 & \text{positively correlated} \\ = 0 & \text{uncorrelated} \\ < 0 & \text{negatively correlated} \end{cases}$

- X, Y independent $\Rightarrow \operatorname{cov}(X,Y) = 0$ (反之不一定成立)

$E[XY] = \sum_{(x,y) \in S} xy\, f(x,y) = \sum_{(x,y) \in S} xy\, f_X(x) f_Y(y)$

$= \sum_{x \in S_X} x f_X(x) \cdot \sum_{y \in S_Y} y f_Y(y) = E[X]E[Y]$

- $\operatorname{cov}(X,X) = \operatorname{Var}[X]$

5. The Correlation Coefficient

$\rho = \frac{\operatorname{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

① $\rho \in [-1, 1] \Longleftrightarrow \operatorname{Cov}(X,Y)^2 \le (\sigma_X \sigma_Y)^2$

let $V = X - \mu_X$, $W = Y - \mu_Y$

$E[(V+tW)^2] = E[V^2] + 2tE[VW] + t^2 E[W^2]$

$= \sigma_X^2 + 2t\,\mathrm{COV}(X,Y) + t^2\sigma_Y^2 \geq 0$ (1)

$\forall t \in \mathbb{R}$, (1) is true

$\Rightarrow t = -\dfrac{\mathrm{COV}(X,Y)}{\sigma_Y^2}$ $\quad f(t) = \dfrac{\mathrm{COV}^2(X,Y)}{\sigma_Y^2} - \dfrac{2\,\mathrm{COV}^2(X,Y)}{\sigma_Y^2} + \sigma_X^2 \geq 0$

$\Rightarrow \mathrm{COV}(X,Y)^2 \leq (\sigma_X \sigma_Y)^2$

• $\rho = \pm 1$ $\quad Y - E[Y] = K(X - E[X])$ $\quad \exists K \in \mathbb{R}$

$\rho = \pm 1$. $\mathrm{Cov}(X,Y) = \pm\sigma_X\sigma_Y$ $\quad$ find K $\Rightarrow$ find t s.t $E[(V+tW)^2] = 0$

Since $E[(V+tW)^2] \geq 0$, check $f(t)_{min} = f\left(-\dfrac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}\right) = 0$

$\Rightarrow V + \left[-\dfrac{\mathrm{COV}(X,Y)}{\sigma_Y^2}\right]W = 0 \Rightarrow K = \dfrac{\mathrm{COV}(X,Y)}{\sigma^2 Y}$

6. Conditional Probability

$A = \{X = x\}.\ B = \{Y = y\}$

$$g(x|y) = P(X = x \mid Y = y) = \frac{P(A \cap B)}{P(B)} = \frac{f(x,y)}{f_Y(y)}$$

• if $f(x,y) = f_X(x) f_Y(y) \Rightarrow g(x|y) = f_X(x)$

$$\sum_{y \in S_Y} h(y|x) = \sum_{y \in S_Y} \frac{f(x,y)}{f_X(x)} = \frac{\sum_{y \in S_Y} f(x,y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1$$

• $$E[u(X)|Y=y] = \sum_{x \in S_X} u(x)\, g(x|y)$$

$$= \sum_{x \in S_X} x \cdot g(x|y)$$

• $$Var[X|Y=y] = \sum_{x \in S_X} (x - E[X|Y=y])^2 g(x|y)$$

[e.g.] X be the number of success in n Bernoulli trials.

Y be the number of success in the first $m<n$ trials.

Z be the number of success in the last $n-m$ trials.

[Z is independent of Y]

① $g(x|y) = \dfrac{f(x,y)}{f_Y(y)} = \dfrac{f(x-y,y)}{f_Y(y)} \leftarrow x = Y+Z$

$Z \in [0, n-m]$
$\downarrow$
$n-m+y$

$= \dfrac{f_Y(y)\, f_Z(x-y)}{f_Y(y)} = f_Z(x-y)$

② $E[X|Y=y] = \sum_{x=y} x\, f_Z(x-y)$ $\quad Z \sim B(n-m, p)$

$= \sum_{z=0}^{n-m} (y+z) f_Z(z)$

$= y \sum_{z=0}^{n-m} f_Z(z) + \sum_{z=0}^{n-m} z f_Z(z)$

$= y \cdot 1 + E[Z] = y + (n-m)p$

③ $Var[X|Y=y] = \sum_{x=y}^{n-m+y} [x - y - (n-m)p]^2 f_Z(x-y)$

$= \sum_{z=0}^{n-m} [z-(n-m)p]^2 f_Z(z) = Var[Z]$
$= (n-m)p(1-p)$

7. Joint Probability Density Function

$P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy \quad A \subset \mathbb{R}^2$

$E[u(X,Y)] = \iint_{(x,y)\in S} u(x,y) f(x,y)$ $\quad Var[X] = E[(X-E[X])^2]$
$\quad Var[Y] = E[(Y-E[Y])^2]$

$Cov(X,Y) = E[XY] - E[X]E[Y]$

$E[XY] = \iint_{(x,y)\in S} xy\, f(x,y)\,dx\,dy$, $\rho(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}}$

[e.g.] $f(x,y) = \frac{1}{\theta^2}\exp(-\frac{y}{\theta})$ $\quad 0<x<y<\infty$

① $E[X]$, $Var[X]$, $E[X|Y=y]$, $Var[X|Y=y]$

② $E[Y]$, $Var[Y]$, $E[Y|X=x]$, $Var[Y|X=x]$

$f_X(x) = \int_{y=x}^{\infty} f(x,y)\,dy = \theta e^{-\frac{x}{\theta}}$ $\quad X \sim E(\theta) \Rightarrow E[X]=\theta$, $Var[X]=\theta^2$

$f_Y(y) = \int_{x=0}^{y} f(x,y)\,dx = \frac{y}{\theta^2} e^{-\frac{y}{\theta}}$ $\quad$ Erlang $(\alpha=2) \Rightarrow E[Y] = 2\theta$, $Var[Y] = 2\theta^2$

$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{1}{y}$ $\quad X|Y=y \sim U(0,y) \Rightarrow E[X|Y=y] = \frac{y}{2}$ $\quad Var[X|Y=y] = \frac{y^2}{12}$

$f_{Y|X}(y|x) = \frac{1}{\theta} e^{-\frac{y-x}{\theta}}$ $\quad E[Y|X=x] = E[X+Z] = \theta + x$

$Y = X+Z$ $\quad Var[Y|X=x] = Var[X+Z] = Var[Z] = \theta^2$

constant

8. tower Property $\quad E[E[X|Y]] = E[X]$

proof: $E[E[X|Y]] = \int_{S_Y} E[X|Y=y] f_Y(y)\,dy = \int_{S_Y}\int_{\mathbb{R}} x\, g(x|y) f_Y(y)\,dx\,dy$

$= \int_{\mathbb{R}} x \int_{S_Y} f(x,y)\,dy\,dx$

$= \int_{\mathbb{R}} x\, f_X(x)\,dx = E[X]$

$Var[Y] = E[Var[Y|X]] + Var[E[Y|X]]$

$E[Y^2] = E[E[Y^2|X]] = E[Var[Y|X] + (E[Y|X])^2] = E[Var[Y|X]] + E[E[Y|X]^2]$

$Var[Y] = E[Y^2] - (E[Y])^2 = E[Var[Y|X]] + E[E[Y|X]^2] - (E[Y])^2$

$= E[Var[Y|X]] + E[E[Y|X]^2] - (E[E[Y|X]])^2$

9. Bivariate Normal Distribution $\quad = E[Var[Y|X]] + Var[E[Y|X]]$

$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp[-\frac{q(x,y)}{2}]$, $q(x,y) = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]$ $\quad \sigma_X>0, \sigma_Y>0, |\rho|<1$

*边际分布均为正态、$X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$

☆ $E[XY] = E[E[XY|X]]$

$= E[X \cdot E[Y|X]]$

$E[X|Y=y] = \mu_X + \frac{\sigma_X}{\sigma_Y}\rho(y-\mu_Y)$ $\quad Var[X|Y=y] = (1-\rho^2)\sigma_X^2$

$\longrightarrow E[Y|X=x] = \mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x-\mu_X)$

☆ independence $\Leftrightarrow$ uncorrelation

proof:

① $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp(-\frac{u^2}{2}) \int_{-\infty}^{\infty} \exp[-\frac{(v-\rho u)^2}{2(1-\rho^2)}]\,dy$ $\quad u = \frac{x-\mu_X}{\sigma_X}, v = \frac{y-\mu_Y}{\sigma_Y} \Rightarrow q(x,y) = \frac{1}{1-\rho^2}(u^2 - 2\rho uv + v^2)$

$= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp(-\frac{u^2}{2}) \underbrace{\int_{-\infty}^{\infty} \exp[-\frac{t^2}{2(1-\rho^2)}]}_{\sqrt{2\pi(1-\rho^2)}}\,dt \quad (t = v - \rho u) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp[-\frac{(x-\mu_X)^2}{2\sigma_X^2}]$

② $g(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}}{\frac{1}{\sqrt{2\pi}\sigma_Y}} \cdot \frac{\exp[-\frac{q(x,y)}{2}]}{\exp[-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}]} = \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}} \exp[-\frac{(x-B)^2}{2(1-\rho^2)\sigma_X^2}]$ $\quad B = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)$

$\Rightarrow E[X|Y=y] = \mu_X + \frac{\sigma_X}{\sigma_Y}\rho(y-\mu_Y)$ $\quad Var[X|Y=y] = (1-\rho^2)\sigma_X^2$

③ 1° independent $\Rightarrow \rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = 0$

2° $\rho = 0$ $\quad f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp[-\frac{1}{2}(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2})]$

$= \frac{1}{\sqrt{2\pi}\sigma_X} \exp[-\frac{(x-\mu_X)^2}{2\sigma_X^2}] \cdot [\frac{1}{\sqrt{2\pi}\sigma_Y} \exp[-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}]] = f_X(x) \cdot f_Y(y)$

10. function of one random variable $\quad$ [one-to-one mapping]

$Y = u(X) \Leftrightarrow X = v(Y)$

<1> discrete case

$g(y) = P(Y=y) = P(u(X)=y) = P(X=v(y))$

$P(X=x) = f(x) \Rightarrow g(y) = f(v(y))$

<2> continuous case → 反函数唯一

[monotonic function]

$$g(y) = f(v(y))\left|\frac{dv(y)}{dy}\right|$$

① increasing

$$G(y) = P(Y \le y) = P(u(x) \le y) = P(X \le v(y)) = \int_{c_1}^{v(y)} f(x)\,dx$$

$$g(y) = G'(y) = f(v(y))\,v'(y) = f(v(y))\left|\frac{dv(y)}{dy}\right|$$

② decrease $G(y) = P(X \ge v(y)) = 1 - \int_{c_1}^{v(y)} f(x)\,dx$

$$g(y) = -f(v(y))\frac{dv(y)}{dy} = f(v(y))\left|\frac{dv(y)}{dy}\right|$$

· random number generator

<1> $Y \sim U(0,1)$ $F(x)$ strictly increasing $a<x<b$ $(F(a)=0, F(b)=1)$ $\Rightarrow$ $X = F^{-1}(Y)$ is a continuous-type random variable with cdf $F(x)$

proof: $P(X \le x) = P(F^{-1}(Y) \le x)$ 均递增

$$= P(F(F^{-1}(Y)) \le F(x)) = P(Y \le F(x))$$

$Y \sim U(0,1) \Rightarrow P(X \le x) = P(Y \le F(x)) = F(x)$

<2> cdf $F(x)$ $a<x<b$ strictly increasing

$Y = F(x) \sim U(0,1)$

proof: $F(a)=0$, $F(b)=1$ $\Rightarrow P(Y \le y) = P(F(x) \le y)$ $y \in (0,1)$

$$= P(X \le F^{-1}(y))$$

$P(X \le x) = F(x) \Rightarrow P(X \le F^{-1}(y)) = F(F^{-1}(y)) = y$ $\sim U(0,1)$

<3> X random variable $\Rightarrow$ find $Y = u(X)$

[e.g.] $f(x) = \frac{1}{\pi(1+x^2)}$, $x \in (-\infty, \infty)$, $Y = X^2$

find the pdf of Y

$$G(y) = P(Y \le y) = P(X^2 \le y) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)\,dx = 2\int_0^{\sqrt{y}} \frac{1}{\pi(1+x^2)}\,dx$$

$$g(y) = G'(y) = \frac{1}{\pi(1+y)\sqrt{y}}$$

· Several random variables

1. n independent random variables [mutually independent]

$f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(n)$

2. Independently and identically distributed random variables

$g(x_1, \cdots x_n) = f(x_1) \cdots f(x_n)$ [直接拆解成n个相同的分布]

3. $X_1, \cdots, X_n$ independent random variables

· $E[u_1(X_1) \cdots U_n(X_n)] = \sum_{x_1 \in S_{x_1}} u_1(x_1) f_1(x_1) \cdots \sum_{x_n \in S_{x_n}} u_n(x_n) f_n(x_n)$

$= E[u_1(X_1)] \cdots E[U_n(X_n)]$

· $Y = \sum_{i=1}^{n} a_i X_i$

$E[Y] = \sum_{i=1}^{n} a_i E[X_i] = \sum_{i=1}^{n} a_i u_i$

$Var[Y] = E[(\sum_{i=1}^{n} a_i X_i - \sum_{i=1}^{n} a_i u_i)^2] = E[(\sum_{i=1}^{n} a_i (X_i - u_i))^2] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j E[(X_i - u_i)(X_j - u_j)]$

when $i \neq j$, 由独立知 $\Rightarrow = \sum_{i=1}^{n} a_i^2 \sigma_i^2$ ($i=j$, $E[(X_i-u_i)^2] = \sigma_i^2$)

$E[(X_i - u_i)(X_j - u_j)] = E[X_i - U_i] \cdot E[X_j - U_j] = 0$

4. Moment Generating function Technique

(1) $X_1, \cdots, X_n$ are independent random variables, $M_{X_i}(t)$ $t \in [-h_i, h_i]$ ($h_i > 0$), $Y = \sum_{i=1}^{n} a_i X_i$

$M_Y(t) = E[e^{tY}] = E[e^{t(a_1 X_1 + \cdots a_n X_n)}] = E[e^{a_1 t X_1}] \cdots E[e^{a_n t X_n}]$ (by independence)

$= \prod_{i=1}^{n} M_{X_i}(a_i t)$

$\Rightarrow M_{X_i}(t) = M(t)$ ① $M_Y(t) = [M(t)]^n$ $(a=1)$

② $\overline{X} = \sum_{i=1}^{n} (\frac{1}{n}) X_i$ $M_{\overline{X}}(t) = [M(\frac{t}{n})]^n$ $(a = \frac{1}{n})$

(2) $X_1, \cdots, X_n$ be independent chi-square random variables with $r_1, \cdots, r_n$ degrees of freedom $Y = X_1 + X_2 + \cdots + X_n \sim \chi^2(r_1 + \cdots + r_n)$

$X_i \sim \chi^2(r_i)$ $M_{X_i}(t) = \frac{1}{(1-2t)^{\frac{r_i}{2}}}$, $t < \frac{1}{2}$ $M_Y(t) = \frac{1}{(1-2t)^{\frac{1}{2}\sum r_i}}$ $(t < \frac{1}{2})$ i.e. $\sim \chi^2(r_1 + \cdots + r_n)$

$\Rightarrow X_i \sim N(0,1)$, $X_i^2 \sim \chi^2(1)$ $Y = \sum X_i \sim \chi^2(n)$

$X_i \sim N(\mu_i, \sigma^2) \Rightarrow Y = \sum \frac{(X_i - u_i)^2}{\sigma_i^2} \sim \chi^2(n)$

5.

(1) $X_i \sim N(\mu_i, \sigma_i^2)$ mutually independent $Y = \sum_{i=1}^{n} c_i X_i \sim N(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2)$

$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(c_i t) = \prod_{i=1}^{n} \exp(u_i c_i t + \frac{\sigma_i^2 c_i^2 t^2}{2}) = \exp[(\sum_{i=1}^{n} c_i \mu_i) t + (\sum_{i=1}^{n} c_i^2 \sigma_i^2) \frac{t^2}{2}]$

$X_i \sim N(u_i, \sigma_i^2)$

$* X_i \sim N(u, \sigma^2) \Rightarrow \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$ $[c_i = \frac{1}{n}]$

(2) $X_i \sim N(\mu, \sigma^2)$, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

① $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ ② $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$

无偏方差

proof ②: $W = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \sum_{i=1}^{n}\left[\frac{(X_i-\bar{X})+(\bar{X}-\mu)}{\sigma}\right]^2 = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 + \frac{n(\bar{X}-\mu)^2}{\sigma^2} + 2\sum_{i=1}^{n}\frac{(\bar{X}-\mu)(X_i-\bar{X})}{\sigma^2}$

$= \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 + \frac{n(\bar{X}-\mu)^2}{\sigma^2} + \frac{2(\bar{X}-\mu)}{\sigma^2}\sum_{i=1}^{n}(X_i-\bar{X}) = \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 + \frac{n(\bar{X}-\mu)^2}{\sigma^2} \sim \chi^2(n)$

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

$M(t) = \frac{1}{(1-2t)^{\frac{n}{2}}} = M_1(t)\cdot\frac{1}{(1-2t)^{\frac{1}{2}}} \Rightarrow M_1(t) = \frac{1}{(1-2t)^{\frac{n-1}{2}}}$, i.e. $\chi^2(n-1)$

* $\bar{X}$代替μ $\bar{X} = \frac{1}{n}\sum_{i=1}^{n}x_i \Rightarrow \sum_{i=1}^{n}(X_i-\bar{X}) = 0$ 由n-1个离差→nth 约束自动满足

6. Student's t Distribution (学生氏分布) 相较于N，T分布容错率更高 r↑ 更接近正态分布

$T = \frac{Z}{\sqrt{\frac{U}{r}}}$ $Z \sim N(0,1)$, $U \sim \chi^2(r)$ Z, U is independent

$f(t) = \frac{\Gamma(\frac{r+1}{2})}{\sqrt{\pi r}\,\Gamma(\frac{r}{2})}\frac{1}{(1+\frac{t^2}{r})^{\frac{r+1}{2}}}$ $-\infty < t < \infty$

proof: $g(z,u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}}u^{\frac{r}{2}-1}e^{-\frac{u}{2}}$ $z \in \mathbb{R}, u \in [0, \infty)$

$F(t) = P(T \le t) = P(Z \le \sqrt{\frac{U}{r}}\,t)$ $f(t) = F'(t)$

$*\ Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ (已知σ) $U = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

总体标准差 样本标准差

$T = \frac{Z}{\sqrt{\frac{U}{n-1}}} = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$ ← 未知σ，用S估计

[n≥30. by CLT → normal distribution]

7. Central Limit Theorem

① $\lim_{n\to\infty} F_n(z) = F(z)$ (cont.) F是cdf，则有 $Z_n \xrightarrow{d} Z$ 依分布收敛

② CLT: $X_1, \dots, X_n \sim N(\mu, \sigma^2)$, $\sigma^2 > 0$

$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} N(0,1)$ $\Rightarrow$ $\begin{cases} \bar{X} = \mu + \frac{\sigma}{\sqrt{n}}\cdot\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} (\mu, \frac{\sigma^2}{n}) \\ S_n = \sum_{i=1}^{n} X_i = n\bar{X} \xrightarrow{d} (n\mu, n\sigma^2) \end{cases}$

$Y = \sum_{i=1}^{n} X_i$ $\frac{Y-\mu_Y}{\sigma_Y}$

③ $M_n(t) = M(t)$ → 分布也近似

Proof: $W_n = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$, $W_n \xrightarrow{d} N(0,1)$

$$M_n(t) = E\left[e^{t\cdot\frac{1}{\sqrt{n}\sigma}\left(\sum_{i=1}^{n} x_i - n\mu\right)}\right] = E\left[\exp\left(\left(\frac{t}{\sqrt{n}}\right)\left(\frac{x_1-u}{\sigma}\right)\right)\cdots\exp\left(\left(\frac{t}{\sqrt{n}}\right)\frac{x_n-u}{\sigma}\right)\right]$$

$$= \left[m\left(\frac{t}{\sqrt{n}}\right)\right]^n \quad -h < \frac{t}{\sqrt{n}} < h,\quad m(t) = E\left[\exp\left(t\left(\frac{x_i-u}{\sigma}\right)\right)\right] \quad (-h<t<h)$$

let $Y = \frac{x_i-u}{\sigma}$ $\quad m(0)=1,\ m'(0)=0.\ \ M''(0)=1$

$$\Rightarrow M(t) = 1 + \frac{m''(\delta)t^2}{2} \quad \delta\in[0,t]$$

$$= 1 + \frac{1}{2}t^2 + \frac{[m''(\delta)-1]t^2}{2}$$

$n\to\infty,\ t_1\to 0 \Rightarrow \lim_{n\to\infty} M_n(t) = \lim_{n\to\infty}\left(1+\frac{t^2}{2n}\right)^n$

$$= e^{\frac{t^2}{2}}$$

i.e. the mgf of $N(0,1)$

8.

Markov's Inequality $P(X\geq c) \leq \frac{E[X]}{c}$ X is a nonnegative random variable and $c>0$

Chebyshev's Inequality $\forall k\geq 1$ $\quad P(|X-\mu|\geq k\sigma) \leq \frac{1}{k^2}$ (if $\epsilon = k\sigma$ $\ P(|X-\mu|\geq\epsilon) \leq \frac{\sigma^2}{\epsilon^2}$)

Proof: let $Y = (X-u)^2$. $c = k^2\sigma^2$, $E[Y] = \sigma^2$

$$\Rightarrow P(|X-u|\geq k\sigma) = P(Y\geq c) \leq \frac{E[Y]}{c} = \frac{1}{k^2}$$

9. $Z_1, \cdots Z_n$ is said to converge in probability to a random variable Z, often denoted by $Z_n \xrightarrow{P} Z$ [依機率收斂] if for any $\epsilon>0$, $\lim_{n\to\infty} P(|Z_n - Z|\geq\epsilon) = 0$

$$\lim_{n\to\infty} P(|\bar{X}-\mu|\geq\epsilon) = 0$$

Proof: $E[\bar{X}] = \mu$. $Var[\bar{X}] = \frac{1}{n}\sigma^2$

$\forall\epsilon>0$

$$P(|\bar{X}-\mu|\geq\epsilon) \leq \frac{1}{\epsilon^2}\cdot\left(\frac{\sigma}{n}\right)^2 = \frac{\sigma^2}{n\epsilon^2}$$

$$0 \leq \lim_{n\to\infty} P(|\bar{X}-\mu|\geq\epsilon) \leq \lim_{n\to\infty}\frac{\sigma^2}{n\epsilon^2} = 0$$