$Var[Y|X=x]$
$= E[Y^2|X=x] - E^2[Y|X=x]$

3项分布 $p_Z = 1 - p_X - p_Y$

$h(y|x) = \frac{(n-x)!}{y!(n-x-y)!}\left(\frac{p_Y}{1-p_X}\right)^y\left(\frac{p_Z}{1-p_X}\right)^{n-x-y}$

$b[n-x, \frac{p_Y}{1-p_X}]$ $u = (n-x)\frac{p_Y}{1-p_X}$

$g(x|y) = \frac{(n-y)!}{x!(n-x-y)!}\left(\frac{p_X}{1-p_Y}\right)^x\left(\frac{p_Z}{1-p_Y}\right)^{n-x-y}$

$b[n-y, \frac{p_X}{1-p_Y}]$ $\rho^2 = \frac{p_X p_Y}{(1-p_X)(1-p_Y)}$

$\rho = -\sqrt{\frac{p_X p_Y}{(1-p_X)(1-p_Y)}}$ XX R.V CTS. $\int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} f(x,y)\,dx = 1$

M-PDF $f_X(x) = \int f(x,y)\,dy$

$f_Y(y) = \int f(x,y)\,dx$ $E[g(X,Y)] = \int dy \int g(x,y) f(x,y)\,dx$

$E[X] = \int dy \int x f(x,y)\,dx$
$= \int x f_X(x)\,dx$ $Var[X] = \int dy \int (x - E[X])^2 f(x,y)\,dx$

$\int (x-E[X])^2 f_X(x)\,dx$

$E[XY] = \int dy \int xy f(x,y)\,dx$ Conditional

$P(Y\in A|X=x) = \int_{y\in A} h(y|x)\,dy$ $P(X\in A|Y=y) = \int_{x\in A} g(x|y)\,dx$

$E[Y|X=x] = \int y\, h(y|x)\,dy$

$Var[Y|X=x] = E[Y^2|X=x] - E^2[Y|X=x]$

$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp[-\frac{1}{2}q(x,y)]$

$q(x,y) = \frac{1}{1-\rho^2}\left[\left(\frac{x-u_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-u_x}{\sigma_x}\right)\left(\frac{y-u_y}{\sigma_y}\right) + \left(\frac{y-u_y}{\sigma_y}\right)^2\right]$

$X \sim N(u_x, \sigma_x^2)$ $Y \sim N(u_y, \sigma_y^2)$

$X|Y \sim N(u_x + \frac{\sigma_x}{\sigma_y}\rho(y-u_y), (1-\rho^2)\sigma_x^2)$

$Y|X \sim N(u_y + \frac{\sigma_y}{\sigma_x}\rho(x-u_x), (1-\rho^2)\sigma_y^2)$

Inde <=> 无关 for ↑ 双正态

3-d normal 截面 ① ⊥切 $f(x_0,y) = f_X(x_0)\,h(y|x_0)$

$0 < z_0 < \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$, $\exp[-\frac{1}{2}q(x,y)] = z_0 \cdot 2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}$

$q(x,y)\cdot(1-\rho^2) = -2(1-\rho^2)\ln(z_0 2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})$

Discrete Case

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

$P(X=x) = f(x)$ $g(y) = f[v(y)]$

注意 X, Y 各自空间 S_X, S_Y!

CTS Case

$F(x) = \int f(t)\,dt$ $F'(x) = \frac{dF(x)}{dx} = f(x)$

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

$f(x)$, $Y = u(X)$, $X = v(Y)$.

RNG $Y \sim U(0,1)$ 生成随机数

CDF $F(a)=0, F(b)=1$ CDF ↑

$X = F^{-1}(Y)$.

X is CTS R.V. $S_X = (a,b)$

CDF $F(x)$↑ 则 $Y = F(X) \sim U(0,1)$

Mutually Inde. $f(x_1 \dots x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n)$. 必要条件

$S = S_{X_1} \times \dots \times S_{X_n}$

若 $X_1 \dots X_n$ are inde

任取 n 个都是 inde.

i.i.d = 独立 + 分布一致

$Y = u_1(X_1) \dots u_n(X_n)$

$EY = E(u_1X_1) \dots E(u_nX_n)$.

$X_1, X_2 \dots X_n$
有 $u_1, u_2 \dots u_n$
$\sigma_1^2, \dots \sigma_n^2$ $\bar{X} = \sum_{i=1}^n \frac{1}{n}X_i$

$Y = \sum_{i=1}^n a_i X_i$ $E\bar{X} = u$

$EY = \sum_{i=1}^n a_i u_i$ $Var\,\bar{X} = \frac{\sigma^2}{n}$.

$Var\,Y = \sum_{i=1}^n a_i^2 \sigma_i^2$

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

$X_1, X_2 \dots X_n$ 有 MGF M_{X_i}

$Y = \sum_{i=1}^n a_i X_i$

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

$Y = \sum_{i=1}^n X_i$, $M_Y(t) = \prod_{i=1}^n M(t)$
$= (M(t))^n$

$\bar{X} = \sum_{i=1}^n \frac{1}{n}X_i$, $M_{\bar{X}}(t)$
$= \prod_{i=1}^n M(\frac{1}{n}t) = [M(\frac{t}{n})]^n$

$X_1 \dots X_n$ 有 $r_1 \dots r_n$
$X_i \sim \chi^2(r_i)$ $Y = X_1 + \dots X_n$
$\sim \chi^2(r_1 + \dots r_n)$.

$Z_1 \dots Z_n$ 有 $N(0,1)$

$W = Z_1^2 + \dots Z_n^2 \sim \chi^2(n)$.

$X_1 \dots X_n$ 分别有 $N(u_i, \sigma_i^2)$

$W = \sum_{i=1}^n \left(\frac{X_i - u_i}{\sigma_i}\right)^2 \sim \chi^2(n)$. $Y = \sum_{i=1}^n a_i X_i$

$X_1 \dots X_n$ 分别有 $u_1 \dots u_n$, $\sigma_1^2 \dots \sigma_n^2$.

$Y \sim N(\sum_{i=1}^n a_i u_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$

$X_1 \dots X_n$ 均匀 (u, σ^2)

$\bar{X} \sim N(u, \frac{\sigma^2}{n}) \iff \frac{\bar{X}-u}{\sigma/\sqrt{n}} \sim N(0,1)$

$E[S^2] = \sigma^2$. $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$

$W = \sum_{i=1}^n \left(\frac{X_i - u}{\sigma}\right)^2$, $Z = \frac{\bar{X}-u}{\sigma/\sqrt{n}}$, $W = \frac{n-1}{\sigma^2}S^2 + Z^2$

$W \sim \chi^2(n)$, $Z^2 \sim \chi^2(1)$.

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

$T = \frac{Z}{\sqrt{U/r}}$ $Z \sim N(0,1)$ $U \sim \chi^2(r)$ 学生氏!

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

标准 $f(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{e^{\frac{x^2}{2}}}$ $T \sim t(r)$.

$r=1$ $f(x) = \frac{1}{\pi(1+x^2)}$ $X_1 \dots X_n$ 有 $N(u, \sigma^2)$

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

$T = \frac{Z}{\sqrt{U/(n-1)}} = \frac{\bar{X}-u}{S/\sqrt{n}} \sim t(n-1)$

$\frac{\bar{X}-u}{\sigma/\sqrt{n}} \sim N(0,1) \iff \bar{X} \sim N(u, \frac{\sigma^2}{n}) \iff \sum_{i=1}^n X_i \sim N(nu, n\sigma^2)$

依分布 cvg $\lim_{n\to\infty} F_n(z) = F(z)$ $Z_n \xrightarrow{d} Z$

$|P(Z_n \le z) - P(Z \le z)| < \varepsilon$

$\frac{\bar{X}-u}{\sigma/\sqrt{n}} \sim N(0,1)$ $\bar{X} \sim N(u, \frac{\sigma^2}{n})$ $\sum_{i=1}^n X_i \sim N(nu, n\sigma^2)$

$P(Y=k) = P(k - \frac{1}{2} < Y < k + \frac{1}{2})$ $Y \sim b(n,p)$.

$\frac{Y/n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0,1)$ $P(Y \le k) = \Phi\left(\frac{k + \frac{1}{2} - np}{\sqrt{npq}}\right)$

$P(Y < k) = \Phi\left(\frac{k - \frac{1}{2} - np}{\sqrt{npq}}\right)$ $X_1 \dots X_n \sim$ poisson (λ). $Y = \sum_{i=1}^n X_i$

$Y \sim$ Poisson $(n\lambda)$ 注意 interval 变化

由离散 → CTS

$P(|X-u| \ge k\sigma) \le \frac{1}{k^2}$, $P(|X-u| \ge \varepsilon) \le \frac{\sigma^2}{\varepsilon^2}$

依概率 cvg $\lim_{n\to\infty} P(|Z_n - Z| \ge \varepsilon) = 0$. $Z_n \xrightarrow{P} Z$

$\lim_{n\to\infty} P(|\bar{X} - u| \ge \varepsilon) = 0$.

$\lim_{n\to\infty} M_n(t) = M(t)$ $Z_n \xrightarrow{d} Z$

$1 + \dots n^2 = \frac{n(n+1)(2n+1)}{6}$ $\sum \frac{1}{x!} = e$

$1 + \dots n^3 = \left(\frac{n(n+1)}{2}\right)^2$ 指数 $E[X^k] = k!\,\lambda^{-k}$

Gamma $E[X^k] = \alpha(\alpha+1)\dots(\alpha+k-1)\theta^k$

No Memory 指/几何 dx 横切 dy 竖切

$E[XY] = \int_a^b dx \int_c^d f(x,y)\cdot xy \cdot dy$ 注意 interval

$\int u\,dv = uv - \int v\,du$ $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$

Approximation discrete 没等号左加

右减 正常右减左加

Random Variable.

PMF, $f(x)>0$, $S \to (0,1]$

$\sum_{x\in S} f(x)=1$, $P(X\in A)=\sum_{x\in A} f(x)$.

cdf $F(x)=P(X\le x)=\sum_{x=x,\,x\in S} f(x)$

$P(a<X\le b)=F(b)-F(a)$.

mean $E[g(x)]=\sum_{x\in S} g(x)f(x)$.

$M^r(0)=E[X^r]$ $\quad Var=E[(X-u)^2] = E[X^2]-E^2[X]$

注 X可以用 x^2、x^3 替代 $\quad E[X(X-1)]+E[X]-E^2[X]$

MGF $E[e^{tX}]=\sum_{x\in S} e^{tx}f(x)$. $-h<t<h$.

$M(0)=1$, $M'(0)=u$, $M''(0)=u^2+\sigma^2$

MGF一样 → distribution一样.

算$E[X^r]$可用. Bernoulli

PMF $p^x q^{1-x}$, $u=p$, $\sigma^2=pq$

MGF $q+pe^t$, $X\sim b(1,p)$

Binomial

PMF $C\binom{n}{x} p^x q^{n-x}$ $\quad u=np$. $\sigma^2=npq$.

MGF $(q+pe^t)^n$ $\quad X\sim b(n,p)$

CDF $\sum_{y=0}^{[x]} C\binom{n}{y} p^y q^{n-y}$

$F(x)=P(X\le x)$

Geometric

PMF $q^{x-1}p$ $\quad u=\frac{1}{p}$

MGF $\frac{pe^t}{1-qe^t}$ $\quad \sigma^2=\frac{q}{p^2}$

CDF $P(X>k)=(1-p)^k$

$P(X\le k)=1-(1-p)^k$

$t<-\ln(1-p)$.

Hypergeometric

PMF $\frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$ $\quad u=n\frac{N_1}{N}$

$\sigma^2=n\frac{N_1}{N}\cdot\frac{N_2}{N}\cdot\frac{N-n}{N-1}$

Negative Binomial

PMF $\binom{x-1}{r-1}p^r q^{x-r}$ $\quad u=\frac{r}{p}$

$\sigma^2=\frac{rq}{p^2}$ $\quad$ MGF $\frac{(pe^t)^r}{(1-qe^t)^r}$

uniform PMF $\frac{1}{m}$

$u=\frac{m+1}{2}$ $\quad \sigma^2=\frac{m-1}{2}$

从1、2、…m抽一个.

Poisson $u=\sigma^2=\lambda$

PMF $\frac{\lambda^x e^{-\lambda}}{x!}$ $\quad$ 注意λ取用是否是取 $\lambda\cdot T$

MGF $e^{\lambda(e^t-1)}$. $X\sim Poisson(\lambda)$

CTS R.V.

PDF $S\to(0,\infty)$

$f(x)>0$, $x\in S$

$P(X=a)=0$ $\quad \int_S f(x)dx=1$

无需PDF的CTS or bdd.

$P(a\le X\le b)=\int_a^b f(x)dx$.

CDF $F(x)\triangleq P(X\le x)=\int_{-\infty}^{x} f(t)dt$

$F'(x)=f(x)$. $u=\int_{-\infty}^{\infty} x f(x)dx$.

$\sigma^2=\int_{-\infty}^{\infty}(x-u)^2 f(x)dx$

MGF $\int_S e^{tx}f(x)dx$

$p=\int_{-\infty}^{\pi_p} f(x)dx=F(\pi_p)$. $\quad P(Z<z_\alpha)=1-P(Z>z_\alpha)=1-\alpha$

z_α是$(100-\alpha)$百分位数

uniform

PDF $\frac{1}{b-a}$ $\quad$ MGF $\begin{cases}\frac{e^{tb}-e^{ta}}{t(b-a)}, & t\ne0\\ 1, & t=0\end{cases}$

$u=\frac{a+b}{2}$, $\sigma^2=\frac{(b-a)^2}{12}$

CDF $\frac{x-a}{b-a}$

Exponential Distribution 首次

$\theta=\frac{1}{\lambda}$ $\quad \lambda$: avg occurences/unit interval

PDF $=\frac{1}{\theta}e^{-\frac{x}{\theta}}$ $\quad$ MGF $=\frac{1}{1-\theta t}$

$u=\theta$, $\sigma^2=\theta^2$. CDF $=1-e^{-\frac{x}{\theta}}$

Gamma Distribution 第α次

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

MGF $\frac{1}{(1-\theta t)^\alpha}$ $\quad \Gamma(\frac{1}{2})=\sqrt{\pi}$

$\Gamma(t)=\int_0^\infty y^{t-1}e^{-y}dy$ $\quad \Gamma(n)=(n-1)!$

$\Gamma(t)=(t-1)\Gamma(t-1)$ $\quad u=\alpha\theta$, $\sigma^2=\alpha\theta^2$

Chi-square $\theta=2$, $\alpha=\frac{r}{2}$

PDF $\frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}}x^{\frac{r}{2}-1}e^{-\frac{x}{2}}$ $\quad$ CDF $\int_0^x f(t)dt$

$X\sim\chi^2(r)$

MGF $(1-2t)^{-\frac{r}{2}}$

Normal Distribution

PDF $\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-u)^2}{2\sigma^2}}$ $\quad \int_{-\infty}^{\infty}f(x)dx=1$

CDF $\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz$

MGF $e^{ut+\frac{1}{2}\sigma^2t^2}$

$u=u$, $\sigma^2=\sigma^2$ $\quad X\sim N(u,\sigma^2)$

Standard $Y\sim N(0,1)$ $\quad$ PDF $\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$ $\quad Y\sim N(u,\sigma^2)$ $\quad X=\frac{Y-u}{\sigma}\sim N(0,1)$

$X\sim N(u,\sigma^2)\Rightarrow\left(\frac{X-u}{\sigma}\right)^2\sim\chi^2(1)$

XX R.V. $\quad$ Joint PMF

$S\subseteq S_X\times S_Y$ $\quad f(x,y)>0$, $(x,y)\in S$ $\quad f(x,y)$.

$\sum_{(x,y)\in S}f(x,y)=1$ $\quad P(X=x,Y=y)$

Marginal PMF

$f_X(x)=P(X=x)=\sum_{y\in S_Y(x)}f(x,y)$

对y求和!

Trinomial

$f(x,y)=\frac{n!}{x!y!(n-x-y)!}p_X^x p_Y^y(1-p_X-p_Y)^{n-x-y}$

$f_X(x)=\binom{n}{x}p_X^x(1-p_X)^{n-x}$

$f_Y(y)=C\binom{n}{y}p_Y^y(1-p_Y)^{n-y}$

$X\sim b(n,p_X)$, $Y\sim b(n,p_Y)$

Inde $f(x,y)=f_X(x)f_Y(y)$.

If X and Y inde $\Rightarrow S=S_X\times S_Y$ → Rectangular

$E[g(X,Y)]=\sum_{(x,y)\in S}g(x,y)f(x,y)$

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

$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E[XY]-E[X]E[Y]$ $\quad$ inde⇒ρ=0

>0 正相关 <0 负相关 =0 无关

inde ⇒ 无关 $\quad$ 无关 ⇏ inde (有特例)

$\rho=\frac{Cov(X,Y)}{\sigma_X\cdot\sigma_Y}\in[-1,1]$, if $\rho=\pm1$

线性 $Y-E[Y]=\lambda[X-E[X]]$

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

$E[XY]=u_Xu_Y+\rho\sigma_X\sigma_Y$ $\quad$ Conditional 分布

$g(x|y)=\frac{f(x,y)}{f_Y(y)}$ $\quad h(y|x)=\frac{f(x,y)}{f_X(x)}$

if X, Y inde $g(x|y)=f_X(x)$

$h(y|x)=f_Y(y)$

$E[g(Y)|X=x]=\sum_{y\in S_Y(x)}g(y)h(y|x)$

if $g(Y)=Y$ $\quad y\,h(y|x)$