Γ $f_n(x) = \dfrac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}$ $\longrightarrow$ n个 events

$f_\alpha(x) = \dfrac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{C}$ α是连续时间 α>0 不用整数

$\int_0^\infty f_\alpha(x)dx = 1 \Rightarrow \int_0^\infty \frac{\lambda^\alpha}{C} x^{\alpha-1} e^{-\lambda x} dx = \frac{\lambda^\alpha}{c}\int_0^\infty x^{\alpha-1} e^{-\lambda x} dx$ 令 $\lambda x = y$

则 $x = \frac{y}{\lambda}$ $dx = \frac{dy}{\lambda}$ $\Rightarrow = \frac{\cancel{\lambda^\alpha}}{c}\int_0^\infty (\frac{y}{\cancel{\lambda}})^{\alpha-1} e^{-y}\frac{dy}{\cancel{\lambda}}$

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

$\therefore C = \underbrace{\int_0^\infty y^{\alpha-1} e^{-y} dy}_{\Gamma(\alpha)}$

$\begin{cases} \alpha \text{ is integer}, & \Gamma(n) = (n-1)! \\ \alpha \text{ is not integer} & \Gamma(\alpha) = \underbrace{\int_0^\infty y^{\alpha-1} e^{-y} dy}_{\text{finite and sum to 1.}} \end{cases}$

Check: $\int_0^1 y^{\alpha-1} e^{-y} dy + \int_1^\infty y^{\alpha-1} e^{-y} dy$

$\wedge \qquad \wedge$

(1) $\alpha \in [0,1]$ $\int_0^1 y^{\alpha-1} dy$ $\int_1^\infty e^{-y} dy$

$\int_0^1 y^{\alpha-1} dy = \frac{y^\alpha}{\alpha}\Big|_0^1 = \frac{1}{\alpha}$

$\int_1^\infty e^{-y} dy = -e^{-y}\Big|_1^\infty = e^{-1}$

$\alpha \in (1,\infty)$ $\int_0^1 y^{\alpha-1} e^{-y} dy$ $\int_0^1 y^{\alpha-1} e^{-y} dy$

$\wedge$ $y^{\alpha-1} \le y^{[\alpha]}$ → 整数部分

$\qquad\qquad\qquad\quad n$

$\int_0^1 e^{-y} dy$ $\wedge$

$\quad ||$ $\int_0^\infty y^n e^{-y} dy$

$-e^{-y}\Big|_0^1$

$= 1-e^{-1}$

$\int_0^\infty y^n e^{-y}\,dy = n!$

$\Gamma(n)=(n-1)!$

$$f_\alpha(y) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

$\begin{cases} \alpha\text{: shape parameter } (\alpha=1\text{是指数分布}) \\ \lambda - \text{rate} \end{cases}$

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

α是整数时是 α 个事件的等待时间

α不是整数时无意义.

$\Theta = \frac{1}{\lambda}$ ~ scale parameter (Θ越大，分布越分散spread)

$$f_\alpha(y) = \frac{\quad}{\Gamma(\alpha)}$$

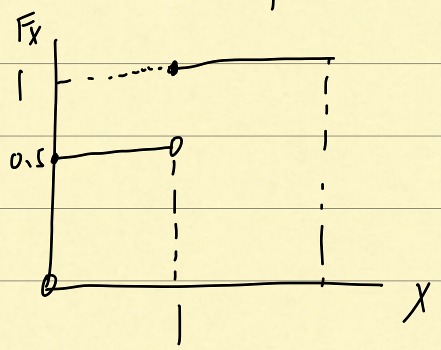
α越大等待时间越长

$r\uparrow \rightarrow$ slower convergence to 1

$\chi^2(r)$

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

已知 CDF 求 quantile $\quad F(\pi_p) = p \quad \pi_p$是quantile.



$$F(x) = P_r(X \le x) = \begin{cases} 0 & x<0 \\ \frac{1}{2} & 0 \le x \le 1 \\ 1 & x \ge 1 \end{cases}$$

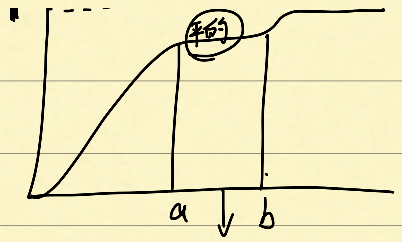
$F(\pi_{0.75}) = 0.75$

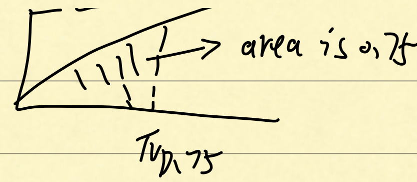
media (上下相等)

$\pi_{0.5}$

$F(x)$

1





flat: $P_r(a<x<b) = F(b) - F(a) = 0$

$F(\pi_{\tilde{p}}) = \tilde{p}$ $\qquad r=2 \quad \theta=2.$

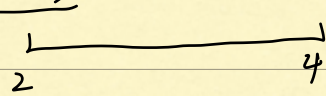
$\chi^2(2) = \quad Gamma(1, \theta=2)$

$CDF = R.$

$F(x) = R$

$X \sim U(a,b)$

uniform distribution.

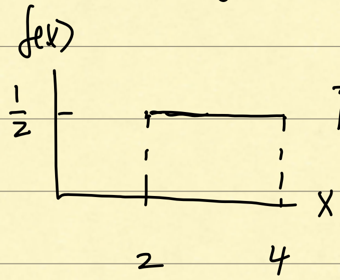


random number. equally likely for each value.

$P_r(2.7\sim2.9) = P_r(3.5\sim3.7)$

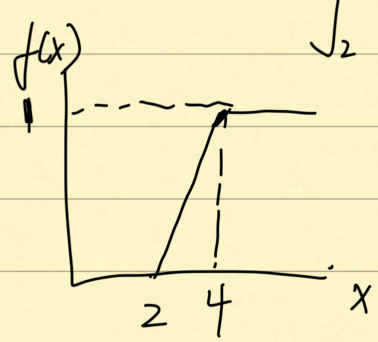
$P_r(x<X<x+\Delta) = f(x)\Delta$ (Pr与Δ的长度成正比)

$f(x) = \quad \int_2^4 c\, dx = 2c = 1 \quad c = \frac{1}{2}$ $\qquad f(x) = \begin{cases} c=1 & 0<x<1 \\ 0 & o/w \end{cases}$



$F(x) = P_r(X \le x) \begin{cases} 0, & x<2 \\ \int_2^x \frac{1}{2} dx & 2 \le x < 4 \\ 1 & x \ge 4 \end{cases}$

$\int_2^x \frac{1}{2} dy = \frac{y}{2}\Big|_2^x = \frac{x}{2} - \frac{2}{2} = \frac{x}{2} - 1$



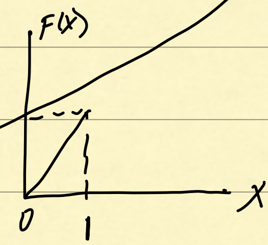
$X \sim U(a+b)$

Skewness of uniform distribution : 0

mean: $\frac{a+b}{2} = E(X)$

$$F(X) = \int_0^1 f(x) = \int_0^1 c \quad = cx\Big|_0^1 = c - 0 = c = 1$$

$$2\int_0^t 1 + 2. \qquad F(x) = \int_0^x 1 dx = x.$$



$f(x) \sim U(0,1)$ $\qquad 2x+2 \sim U(2,4)$

$W = 2U+2$ $\qquad P(U < X) = X$

CDF of W: $Pr(W<X) = Pr(2U+2<X)$

$$= Pr\left(U < \frac{X-2}{2}\right)$$

$$= \int_0^{\frac{X-2}{2}} 1\, dX = \frac{X-2}{2}$$

$$0 < \frac{X-2}{2} < 1$$

when : $2 < X < 4$

$$= \frac{X}{2} - 1$$

$W \sim U(a,b)$: $f_W(x) = \frac{1}{b-a}$ $\qquad$ Var: $\frac{(b-a)^2}{12}$ $\quad E(W) = \frac{a+b}{2}$

$U \sim U(0,1)$: $W = (b-a)\cdot U + a$

Normal distribution

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad E(X)=\mu \quad \sigma=\sigma$$

μ是对称轴 → $\mu\uparrow$ → right
→ $\mu\downarrow$ → left.

σ → $\uparrow$ → spread to μ. →
→ $\downarrow$. → close. to μ →

$\sigma\downarrow$ → convergence to | faster.

PDF of Normal can't be expressed by formula

Normal distribution

$$M(t)=\int_{-\infty}^{+\infty}e^{tx}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

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

$$-\frac{(x-\mu)^2}{2\sigma^2}+tx=-\frac{x^2-2x\mu+\mu^2-2t\sigma^2x}{2\sigma^2}$$

$$=-\frac{\left[x^2-2x(\overbrace{\mu+\sigma^2t}^{\mu_0})+\mu_0^2\right]-\mu_0^2+\mu^2}{2\sigma^2}$$

$$=-\frac{(x-\mu_0)^2}{2\sigma^2}+\frac{\mu_0^2-\mu^2}{2\sigma^2}$$

$$\frac{\mu_0^2-\mu^2}{2\sigma^2}=\frac{(\mu+\sigma^2t)^2-\mu^2}{2\sigma^2}=\frac{\mu^2+2\mu\sigma^2t+(\sigma^2t)^2-\mu^2}{2\sigma^2}$$

$$=\frac{2\mu\sigma^2t+\sigma^4t^2}{2\sigma^2}$$

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

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\left\{-\frac{(x-\mu_0)^2}{2\sigma^2}\right\}\exp\left\{\frac{\sigma^2t^2}{2}+\mu t\right\}dx$$

$$M(t)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty} \exp\left[-\frac{\quad}{2\sigma^2}\right]\cdot\exp\left[\frac{\quad}{2}\right]$$

$$M(t)=\exp\left\{\frac{\sigma^2t^2}{2}+\mu t\right\}\times\underbrace{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu_0)^2}{2\sigma^2}\right]}_{\int_{-\infty}^{+\infty}f(x)=1.}.$$

(trick).

M(t)推导原理:凑一个完整的CDF出来 =1.

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

$$M'(t)=\exp\left\{\mu t+\frac{\sigma^2t^2}{2}\right\}(\mu+t\sigma^2)\qquad E(X)=\mu$$

$$M''(t)=\exp\left\{\mu t+\frac{\sigma^2t^2}{2}\right\}(\mu+t\sigma^2)^2+\exp\left\{\mu t+\frac{\sigma^2t^2}{2}\right\}\times\sigma^2\qquad Var(X)=\sigma^2$$

skewness $=0$

$$\varphi(z)=f_Z(x)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{z^2}{2}\right\}\qquad \Phi(z)=\int_{-\infty}^{x}f(x)dx$$

$$Z\sim N(0,1)\quad X=Z\sigma+\mu\sim(\mu,\sigma^2)$$

$$F_X(x)=Pr(X<x)=Pr(\sigma Z+\mu<x)=Pr\left(Z<\frac{x-\mu}{\sigma}\right)=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

PDF: $$f_X(x)=\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]'_x=\varphi\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma}=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(\frac{x-\mu}{\sigma})^2}{2\sigma^2}\right\}\times\frac{1}{\sigma}$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

已知正态分布 $y=g(x)$,求 y的PDF,⇒ 先求CDF,将 y变为$g(x)$ ⇒ 再求导.
反解

π_p: $Pr(X\le\pi_p)=p$ $\qquad$ $\pi_{2.5\%}\approx-1.96$

$\quad\, Pr(\sigma Z+\mu\le\pi_p)=p$ $\qquad$ $\pi_{5\%}\approx-1.645$

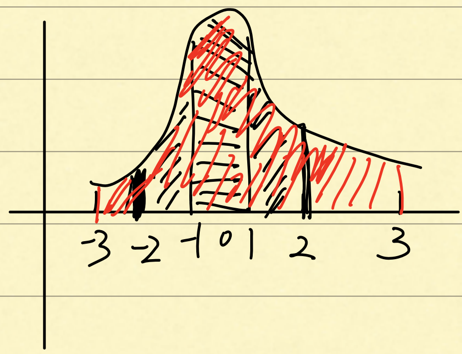
$\quad\, Pr\left(Z\le\frac{\pi_p-\mu}{\sigma}\right)=p$ $\qquad$ in R: $\pi_{5\%}=$ qnorm (0.05)

$\pi_p=\sigma\cdot\hat{\pi}_p+\mu$ (得到标准正态后×σ+μ)

$F_Z\left(\frac{\pi_p-\mu}{\sigma}\right)=p$

$\frac{\pi_p-\mu}{\sigma}=\tilde{\pi}_p$

$\pi_p=\sigma\tilde{\pi}_p+\mu$



$P\{\sigma<Z<\sigma\}=0.68$

$P\{2\sigma<Z<2\sigma\}=0.95$

$P\{3\sigma<Z<3\sigma\}\approx 0.997$

$Z\sim N(0,1)$

MGF of Z^2 = MGF of χ_1^2 $\quad Z^2\sim\chi_1^2$

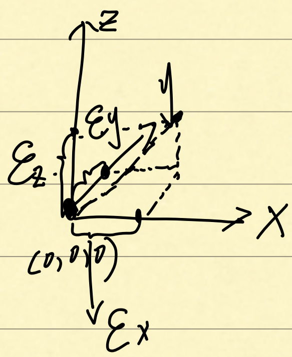
$M_{Z^2}(t)=(1-2t)^{-\frac{1}{2}}\sim\chi^2(1)$

$\chi^2(n)=Z_1^2+Z_2^2+\dots+Z_n^2 \quad Z_i\sim \text{iid } N(0,1)$

independent identical distribution

$M_{\chi^2(n)}(t)=M_{Z_1^2}(t)\times\dots\times M_{Z_n^2}(t)=(1-2t)^{-\frac{n}{2}}=\text{MGF of }\chi^2(n)$

chi-square : Sum of n independent square standard normal variables.



$(\varepsilon_x,\varepsilon_y,\varepsilon_z)$

$\varepsilon_x,\varepsilon_y,\varepsilon_z\sim\text{iid } N(0,0.1^2)$

$\sqrt{\varepsilon_x^2+\varepsilon_y^2+\varepsilon_z^2}\quad 0$

$\Pr\left(\sqrt{\varepsilon_x^2+\varepsilon_y^2+\varepsilon_z^2}>0.3\right)$

0.1 0.1

$$Pr\left\{\left(\frac{\varepsilon_x}{0.1}\right)^2+\left(\frac{\varepsilon_y}{0.1}\right)^2+\left(\frac{\varepsilon_z}{0.1}\right)^2 > \left(\frac{0.3}{0.1}\right)^2\right\}$$

$$Pr\{ Z_x^2 + Z_y^2 + Z_z^2 > 9 \}$$

$Z_x \downarrow$ standard normal

$\chi^2(3)$

$Pr\left(\underbrace{\chi^2(3) > 9}_{右边}\right) = 0.029$ in R: pchisq(9, df=3, lower.tail=false) $\downarrow$ 找upper tail

有0.029的机率使测量误差超过0.3

$\chi^2(r) = \text{Gamma}\left(\frac{r}{2}, 2\right)$ (Chi-square is special case of Gamma)

in R: 1 - pgamma(9, shape=$\frac{3}{2}$, scale=2)

sample mean ~ population mean

$E(\bar{x}) = \mu$, $SD(\bar{x}) = \frac{\sigma}{\sqrt{n}}$ $\quad n \uparrow \to \bar{X} \to \mu \to SD \to 0$

$$P\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \le z\right) \Rightarrow P(Z \le z) \quad Z \sim N(0,1)$$

$$w = \frac{\bar{X} - E(\bar{X})}{SD(\bar{X})} \quad E(w) = 0 \quad Var(w) = 1.$$

$\downarrow$ standardized $\bar{X}$

still a random variable

$$\bar{\varepsilon} = \frac{\varepsilon_1 + \varepsilon_2 + \cdots \varepsilon_n}{n}, \quad G \sim \text{Gamma}(\text{shape}=n, \text{rate}=n)$$

$$\frac{\bar{\varepsilon}-E(\bar{\varepsilon})}{\sqrt{var(\bar{\varepsilon})}} = \frac{\bar{\varepsilon}-\frac{1}{n}}{\sqrt{\frac{\frac{1}{n^2}}{n}}} = \frac{\bar{\varepsilon}-\frac{1}{n}}{\frac{1}{n\sqrt{n}}}$$

Central limit Therom: sample mean → standardized (rescale it) → follow $N(0,1)$

new rV: $\dfrac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ → $\dfrac{\text{standard deviation of single variable}}{\sqrt{\#\ rVs.}}$

MGF → take $n \to 0$ → 计算 → 看是否与 standardized 相同

大数定律 $E(\bar{X})=\mu$, $sd(\bar{X})=0=\frac{\sigma}{\sqrt{n}}$ when $n \to 0$.

how likely it is still same Pr of success 0.37.

$X_i \sim \text{Binom}(n=1708, P=0.37)$

$$\hat{p} = \frac{X_i}{n} = \bar{X}$$

$$E(\hat{p}) = \frac{n \cdot P}{n} = P = 0.37$$

$$Var(\hat{p}) = \frac{nP(1-P)}{n^2} = \frac{0.37 \times 0.63}{1708}$$

$$\frac{\sqrt{1708}(\bar{X}-0.37)}{\quad}$$

$$\frac{\hat{p}-E(\hat{p})}{sd(\hat{p})} = \frac{\hat{p}-0.37}{\sqrt{\frac{0.37\times0.63}{1708}}} = \frac{\sqrt{1708}(\hat{p}-0.37)}{\sqrt{0.37\times0.63}}$$

$$\frac{\sqrt{1708}(\hat{p}-0.37)}{\sqrt{0.37\times0.63}} \sim N(0,1)$$

$$P\left[\frac{\sqrt{1708}(\hat{p}-0.37)}{\quad}, \frac{\sqrt{1708}(0.36-0.37)}{\quad}\right]$$

$Pr(\hat{p} \le 0.36) = Pr\left[\frac{\cdots}{\sqrt{0.37 \times 0.63}} \le \frac{\cdots}{\sqrt{0.37 \times 0.63}} \right]$

$\Phi\left(\frac{-0.01 \times \sqrt{1708}}{\sqrt{0.37 \times 0.63}} \right)$

Calculate in R : $\text{pnorm}\left(\frac{-0.01 \times \sqrt{1708}}{\sqrt{0.37 \times 0.63}} \right)$

discrete: $P(X=x)$ $\qquad \sum_{x \in range(x)} f(x) = 1$ $\quad X$ is Discrete

continous: $F'(x)$ $\qquad \int_{-\infty}^{+\infty} f(x)\,dx = 1$ $\quad X$ is continous

极大似然估计

参数估计：估计概率模型的参数 对于参数/参区间.

极大似然估计

原理

X	1	2
P	θ	$1-\theta$

$\frac{3}{5}$ $\quad \frac{2}{5}$

① ①
② ②

n个球

$\Rightarrow$ 不确定 θ，要抽样 $\boxed{①\ ①\ ②\ ①\ ②}$

$\theta \quad \theta(1-\theta)\theta(1-\theta)$

极大似然原理: $L(\theta) = \theta^3(1-\theta)^2$ 这组样本概率值

偏执地通过样本估计总体，样本①:②=3:2

不一定正确，但这就是极大估计。

求 θ 最大值.

$\ln L(\theta) = 3\ln\theta + 2\ln(1-\theta)$

$\frac{d\ln L(\theta)}{d\theta} = \frac{3}{\theta} - \frac{2}{1-\theta} = 0$

$\hat{\theta} = \frac{3}{5}$

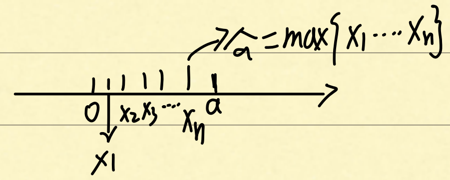
例2 $X \sim U(0,a)$ a未知

$f(x) = \begin{cases} \frac{1}{a} & (0,a) \\ 0 & 其他 \end{cases}$ $\qquad x_1, x_2, \cdots x_n$

$f(x_1)$ $f(x_2)$ $f(x_n)$ $L(a)=\frac{1}{a^n}$

$L(a)=\frac{1}{a}\cdot\frac{1}{a}\cdots\frac{1}{a}$

$\frac{1}{a^n}$最大 a^n最小 a最小



$L(Y,\theta)$ 结果 参数

① 人→不知男女→去了网吧 (从结果出发) Y
Θ 男 90% (更倾向于把人的性别估计为男) $\hat{\Theta}$=男
10%

② 人→不知男女→去了美甲店 (从结果出发) Y →更愿意相信她是女人 $\hat{\Theta}$=女

连续:

$f(x;\theta_1,\cdots,\theta_k)$ $X_1,\cdots X_n$为 samples 则 $f(x_i;\theta_1,\cdots,\theta_k)$ 为 X_i的PDF

$L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)=f(x_1;\theta_1,\cdots,\theta_k)\cdot f(x_2;\theta_1,\cdots,\theta_k)\cdot f(x_3;\theta_1,\cdots,\theta_k)\cdot f(x_4;\theta_1,\cdots,\theta_k)\cdots f(x_n;\theta_1,\cdots,\theta_k)$

离散 累乘

$L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k)=\prod P_r(x_i)$

L(something) 是一个概率函数.

if $\hat{\Theta}$=men, 固定Θ ; Y=去网吧, X=美甲店 $\Rightarrow L(Y,\hat{\Theta})>L(X,\hat{\Theta})$ ∴ 他是去网吧的人 (可能)

in the same idea.

if X固定美甲店, $\hat{\Theta}_1$=men, $\hat{\Theta}_2$=women $\Rightarrow L(X,\hat{\Theta}_2)>L(X,\hat{\Theta}_1)$ ∴更愿相信$\hat{\Theta}_2$ 她是女人

∴ 他更可能是女的

极大似然估计 △可能性更大△ 当把 $x_1,\cdots,x_n$固定,而把L看做$\theta_1,\cdots,\theta_k$的函数时,称L为似然函数。对于不同$(\theta_1,\cdots,\theta_k)$的取值,仅仅在了观察结果$(x_1,\cdots,x_n)$

已知的条件下，$(\theta_1,\cdots,\theta_k)$ 各种值的似然程度

已知结果推参数 $\Rightarrow$ 疫情发生了 $\xrightarrow{\theta_1}$ 美国导致的 △

结果 $\xrightarrow{\theta_2}$ 武汉毒所

$\xrightarrow{\theta_3}$ 海鲜.

$$L(x_1,\cdots,x_n,\theta_1^*,\cdots,\theta_k^*)=\max_{\theta_1,\cdots\theta_k} L(x_1,\cdots,x_n;\theta_1,\cdots,\theta_k) \text{ 的 } \underbrace{(\theta_1^*,\cdots,\theta_k^*)}$$

求出让结果发生的最大参数。

从而使得估计值看起来真的像是真的参数值，这个估计值 $(\theta_1^*,\cdots\theta_k^*)$ 叫做 $(\theta_1,\cdots,\theta_k)$ 的极大似然估计。一般取自然对数，把乘积变成和 $\ln L=\sum_{i=1}^{n}\ln f(x_i;\theta_1,\cdots,\theta_k)$ 求 $\ln L$ 最大时的 θ_i 取值（对各参数求偏导）

设：$(x_1,\cdots,x_n)$ 是从总体中 $N(\mu,\sigma^2)$ 中抽取的样本，则似然函数为

$$L=\prod_{i=1}^{n}\left[(\sqrt{2\pi\sigma^2})^{-1}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\right] \quad \left(\prod_{i=1}^{n} f(x_i)\right)$$

$$\ln L=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\begin{cases}\dfrac{\partial \ln L}{\partial \mu}=\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0\\[2ex]\dfrac{\partial \ln L}{\partial \sigma^2}=-\dfrac{n}{\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2=0\end{cases} \Rightarrow \begin{cases}\mu^*=\sum_{i=1}^{n}x_i/n=\bar{x}\\[1ex](\sigma^*)^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/n=B_k \text{ （即样本中心矩）}\end{cases}$$

σ^2 是整体 看成 $\sigma^2=t$

PDF 必须连续 可导

$Pr(X=x,Y=y)$ Marginal PMF $Pr(X=x)$

PMF $Pr(Y=y)$

$$\left.\begin{array}{l}f(0,0)\\ f(0,1)\\ f(1,0)\end{array}\right\} =\frac{1}{3},\ =\frac{1}{3},\ =\frac{1}{3}$$

marginal $\begin{cases}Pr\{X=0\}=f(0,0)+f(0,1)=\frac{2}{3}\\ Pr\{X=1\}=f(1,0)=\frac{1}{3}\end{cases}$

$Pr(X=1)=Pr(Y=1)=\frac{1}{3}$

$Pr(X=Y=1)\neq Pr(X=1)Pr(Y=1)=\frac{1}{9}$

marginal of X — marginal of Y

$$f_X(x)=\sum_{\{y:(x,y)\in S\}} f(x,y) \qquad f_Y(y)=\sum_{\{x:(x,y)\in S\}} f(x,y)$$

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

$Pr(X>Y) = \sum_{\{(x,y):(x,y)\in S,\ x>y\}} = f(1,0) = \frac{1}{3}$

\# accidents ~ poison(λ)

X ~ # serious accidents Y ~ # non serious accidents

$X, Y \in \{0, 1, 2, \cdots\}$

$P\{X=x, Y=y\} = P\{X+Y=x+y, Y=y\}$ = conditional prob

$P(X=x \mid X+Y=x+y) =$

non-serious serious ~ Binomial (n, P)

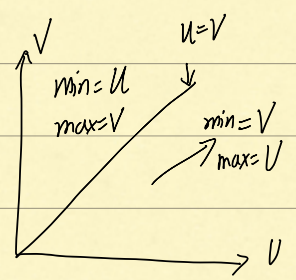
☆例题

$X+Y$ $\quad X \sim pois(\lambda P)$ $\quad Y \sim pois(\lambda(1-P))$

$X+Y \sim Pois(\lambda = \lambda P + \lambda(1-P))$

$U, V \sim$ iid $U(0,1)$ $\quad f_{U,V}(u,v) = f_U(u) \cdot f_V(v) = 1$

joint PDF? $\Rightarrow Pr(X<u, Y<v) = Pr\{\min(U,V)<u, \max(U,V)<v\}$



$u \geq v: Pr(\min(U,V)<u, \max(U,V)<v)$

$= Pr(\min(U,V)<v, \max(U,V)<v)$

$V<U$

$U<V$

not serious: $1-p$

accident occur in a poisson process of rate λ, each accident is serious with prob $p>0$ independent of process of accident occurance and of other accidents. Find the joint and marginal
(给的 serious、non-serious 发生得[illegible])
pmf's of the number of serious and non-serious accidents in a time interval of 1. Are these two random variables independent?

X # serious accidents $\quad Y$ # non-serious accident $\quad x, y \in \{0,1,2,\cdots\}$ any number

joint PMF $P(X=x, Y=y) = P(\underbrace{X+Y=x+y}_{}, X=x)$

总事件 ~ poison(λ)

poisson(λ)

$= P(X+Y=x+y)\, P(X=x \mid X+Y=x+y)$ $\qquad X \sim \text{binomial}(n, p)$

$= \dfrac{e^{-\lambda}\lambda^{x+y}}{(x+y)!} P(X=x \mid X+Y=x+y)$ $\qquad \underline{n=x+y}$

$$P(X=x \mid X+Y=x+y) = \binom{x+y}{x} p^x (1-p)^y$$

$$= \frac{e^{-\lambda}\lambda^{x+y}}{(x+y)!} \times \binom{x+y}{x} p^x (1-p)^y$$

$$= \frac{e^{-\lambda}\lambda^{x+y}}{(x+y)!} \times \frac{(x+y)!}{y!\, x!} p^x (1-p)^y$$

$$= \frac{e^{-p\lambda}(p\lambda)^x}{x!} \times \frac{e^{-(1-p)\lambda}((1-p)\lambda)^y}{y!}$$

X, Y independent: PMF can be factorised into two functions

first function of X, second function of Y

$\lambda_1 = p\lambda \qquad \lambda_2 = (1-p)\lambda$

$$P(X=x) = \sum_{y=0}^{\infty} P(X=x, Y=y) = \frac{e^{-p\lambda}(p\lambda)^x}{x!} \underbrace{\sum_{y=0}^{\infty} \frac{e^{-(1-p)\lambda}((1-p)\lambda)^y}{y!}}_{1}$$

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

$$P(Y=y) = \sum_{x=0}^{\infty} P(X=x, Y=y) = e^{-(1-p)\lambda}$$

$X+Y \sim \text{Pois}(\lambda = \lambda p + \lambda(1-p))$

$X \sim \text{Pois}(\lambda p)$ 、、

$y \sim Pois(\lambda(1-p))$

$X \sim \lambda_1,\ Y \sim \lambda_2 \quad X+Y \sim Pois(\lambda_1+\lambda_2)$ MGF can derive it

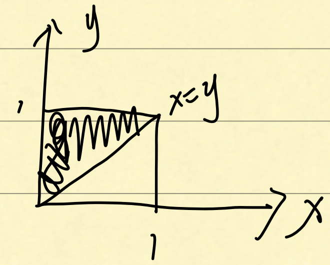
$$Pr(X+Y=k) = \sum_{m=0}^{n} P(X=m) \cdot Pr(Y=k-m)$$

$U, V \sim U(0,1)$ (iid) $\quad X \sim \min\{u,v\} \quad Y \sim \max\{u,v\} \Rightarrow X<Y$

joint pdf of X, Y? marginal pdf of X, Y? X, Y independent?

joint PMF of U, V: $f_{U,V}(u,v) = \underbrace{f_U(u)}_{1} \cdot \underbrace{f_V(v)}_{1} = 1 \quad x, y \in [0,1]$

$Pr(X \le u,\ Y \le v) \Rightarrow Pr(\min(U,V) \le u,\ \max(U,V) \le v)$



$X<Y$

$Pr(\min(U,V)<u,\ \max(U,V)<v)$

$v>u$

$u \ge v$: $Pr(\min(U,V)<u,\ \max(U,V)<v)$

$= Pr(\min(U,V)<v,\ \max(U,V)<v)$

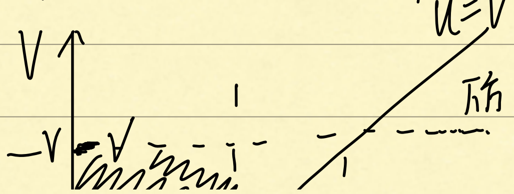
(maximum small than v; minimumum must small than v)

$Pr(U<u,\ V<v)$

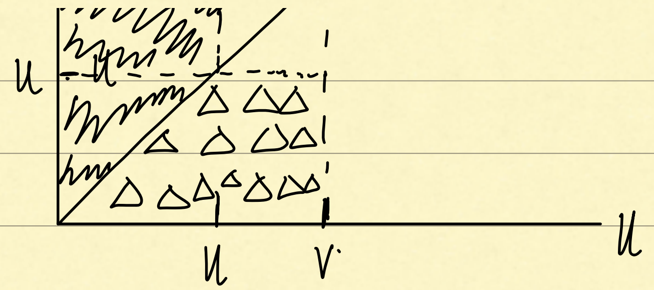
$\Uparrow$

上方 $V>U \Rightarrow \min\{U,V\}=U$
$\max\{U,V\}=V$

$v>u$



下方 $U>V \Rightarrow \min\{U,V\}=V$
$\max\{U,V\}=U$ $\Rightarrow Pr(V<u,\ U<v)$ △



$$F_{X,Y}(u,v) = v^2 - (v-u)^2 = 2uv - u^2 \quad (u<v)$$

$$u \geq v: F_{X,Y}(u,v) = F_{X,Y}(v,v) = v^2 \quad (u \geq v)$$

$$\begin{cases} f_{X,Y}(u,v) = 2, & u<v \\ f_{X,Y}(u,v) = 0, & u \geq v \end{cases}$$

$F(u,v) = 2uv - v^2$ $\quad$ (u) $F(u,1) = 2u - u^2 \quad 0<v<1$

(v) $F(1,v) = F(v,v) = v^2 \quad 0<v<1$

↓

first componet can be larger than 2nd one.

We can only replace 1st componene with v (the smallest v)

$$f_u(u) = 2 - 2u$$

$$f_v(v) = 2v$$

$$f_u(u) \cdot f_v(v) = u(1-u) \cdot v$$

$E(ax+by) = aE(x) + bE(y)$ 不用check independence

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (ax+by) f(x,y)\,dx\,dy = a\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x,y)\,dx + b\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f(x,y)\,dy$$

$$\begin{cases} X = Z \sim N(0,1) \\ E(Z^3) = 0 \quad \text{skewness: (standardised RV)}^{(3)} \end{cases}$$

↳ skewness of standardized normal

$$Var(x+y) = Var(x) + Var(y) + 2cov(X,Y)$$

$$cov(X,Y) = E(XY) - E(X)E(Y) = E((X-\mu_X)(Y-\mu_Y)) = E(XY) - \mu_X\mu_Y$$

Correlation of 2 rVs is the covariance of standardised rVs

$$\rho = corr(X,Y) = E(Z_X Z_Y) - 0 \times 0$$

$$Z_X = \frac{X-\mu_X}{\sigma_X} \quad Z_Y = \frac{Y-\mu_Y}{\sigma_Y}$$

$$-1 \le \rho \le 1$$

$$Y = aX + b \Rightarrow \rho(X,Y) = 1\ (a>0) \quad \rho = -1\ (a<0)$$

corr. cov 只能说明无线性关系, ⇒ 但可以有 非线性关系 (non-linear relationship)

$$\rho = corr = \frac{cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

trinomial — 3 possible outcomes

B Below budget   A: Above budget   E: meets budget

$E(n_A + n_E) = 0.75 \times 3$

$Var(n_A + n_E) = npq = 0.5625$

$corr(n_A, n_E)$

$P_r(n_A = x, n_E = y)$

$$\underbrace{|AA\cdots A|}_{x}\underbrace{EE\cdots E|}_{y}\underbrace{B, B, B \cdots |}_{z}$$

n months

A	E	B
p_A	p_E	p_B

$P_r(n_A = x, n_E = y, n_B = z)$, $z \to n-x-y$

$p_A^x \, p_E^y \, p_B^{n-x-y \leftarrow z}$

$\underbrace{\binom{n}{x}\binom{n-x}{y}}_{\shortparallel} \; x! \, y! \; x(n-x-y)!$

N (# all different arrangements) (unordered)

ordered: $N x! y! (n-x-y)! = n!$

$$N = \left[\binom{n}{x}\binom{n-x}{y}\right] = \binom{n}{x, y, n-x-y}$$ polynomial.

$P_A^x P_E^y (1-P_A-P_E)^{n-x-y} \binom{n}{\ }$

$P(X=x, E=y, B=n-x-y) = P_A^x P_B^y (1-P_A-P_B)^{n-x-y} \binom{n}{x, y, n-x-y}$

trinomial distribution 3 outcomes

$n=3$

$0.25^x \, 0.5^y \, 0.25^{3-x-y} \binom{3}{x, y, 3-x-y}$

$x \backslash y$	0	1	2	3	$P(N_E = x)$
0					
1					
2					
3					

$F(-\infty, y) = 0 = F(x, -\infty) = F(-\infty, -\infty)$

$F(x,y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\, dv$

$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x,y)\, dy = 1$

$(F(x,y))'' = \frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$ 二阶混合偏导.

$\partial x \partial y$

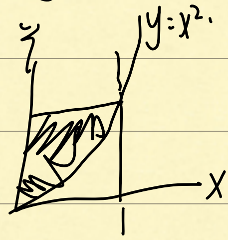
$P(a\le X\le b, c< Y\le d) = F(b,d)-F(a,d)-F(b,c)+F(a,c) = \int_a^b dx\int_c^d f(x,y)dy$.

随机点(X,Y)落在区域D内的概率等于$f(x,y)$在区域D上对x,y的二重积分

$$P((X,Y)\in D)=\iint_D f(x,y)\,dx\,dy$$

marginal PDF: $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy$ $\quad f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx$

$$f(x,y)=\begin{cases}Axy & (0<x<1 \quad x^2<y<1)\\ 0 & 其他\end{cases}$$



① 常数A ② $P(Y>\frac{1}{2}|X>Y)$ ③ $f_X(x)$ 与 $f_Y(y)$

①

$\int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty} Axy\,dy=1$

$\int_0^1 dx\int_{x^2}^1 Axy\,dy=1$

$A\int_0^1 dx\int_{x^2}^1 xy\,dy=1$

$A\int_0^1 x\,dx\int_{x^2}^1 y\,dy=1$

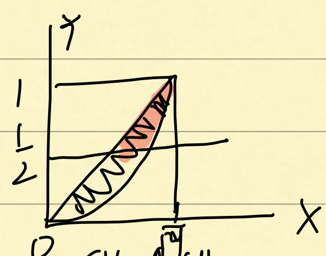
$A\ \frac{1}{2}x^2\Big|_0^1 \int_{x^2}^1 y\,dy=1$

$\int^1 y\,dy = \frac{1}{2}-\frac{1}{2}x^4$

②

$= \dfrac{P(X>Y>\frac{1}{2})}{P(X>Y)}$

$=$



$= \dfrac{\int_{\frac{1}{2}}^1 dy\int_y^{\sqrt{y}} 6xy\,dx}{\int_0^1 dx\int_{x^2}^x 6xy\,dy}$

$= \frac{11}{16} = \dfrac{\frac{11}{14}}{\frac{1}{4}}$

③

$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy$

$=\int_{x^2}^1 f(x,y)\,dy$

$(0<x<1)$.

$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx$

$=\int_0^1 f(x,y)\,dx$

$(0<y<1)$

$\int_{x^2} y\,dy = 2$

$A\int_0^1 \left(\frac{1}{2}x-\frac{1}{2}x^5\right)dx=1$

$A\left(\frac{1}{4}-\frac{1}{12}\right)=1$

$A=6$

→ x范围定

x型 $\int_c^c dx\int_{x_{min}}^{x_{max}} dy$ dx的上下限为常数，先积y

y型 $\int_c^c dy\int_{y_{min}}^{y_{max}} dx$ dy的上下限为常数 先积x算

↓ y的范围不确定

独立性 $F(x,y)=F(x)F(y)$ $f(x,y)=f(x)f(y)$ ⇒ 若x,y独立，$f(x)\cdot g(x)$ 也独立

2维均匀分布

$$f(x,y)=\begin{cases}\frac{1}{S_D} & (x,y)\in D \\ 0 & (x,y)\notin D\end{cases}$$

→ 面积分之一 1维是长度分之一

则称(x,y)服从D上的二维均匀分布 $P((x,y)\in G)=\frac{1}{S_D}\iint_G dxdy=\frac{S_G}{S_D}$

二维正态分布

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

5个参数 $|\rho|<1$

一维：$f(x,y)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$

$(x,y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

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

反之不成立！ X 正态

X 与 Y 独立 $\Leftrightarrow \rho=0$ (仅适用二维正态)

Next week: Thursday 11-12, PT 213; Friday 11:30-1:00, old geology

conditional pmf/pdf of X gire $Y=y$

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

continous: $Pr(X \le x \mid Y=y) = \lim_{\Delta \to 0} Pr(X \le x \mid y \le Y \le y+\Delta)$

discrete: $\frac{Pr(X=x, Y=y)}{Pr(Y=y)}$

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

$$= \lim_{\Delta \to 0} \frac{Pr(X \le x,\ y \le Y \le y+\Delta)}{Pr(y \le Y \le y+\Delta)}$$

$$= \lim_{\Delta \to 0} \frac{Pr(X \le x, Y \le y+\Delta) - Pr(X \le x, Y \le y)}{Pr(Y \le y+\Delta) - Pr(Y \le y)}$$

$$= \lim_{\Delta \to} \frac{(Pr(x, y+\Delta) - F(x,y))/\Delta}{\underbrace{(F_Y(y+\Delta) - F_Y(y))/\Delta}_{f_Y(y)}}$$

$$= \frac{\partial F(x,y)/\partial y}{f_Y(y)}$$

if they are independent:

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

$$\Pr(\text{Vol}=\text{average} \mid \text{height}<\text{average}) = \frac{\Pr(\text{vol}=\text{ave}, \text{hei}<\text{ave})}{\Pr(\text{hei}<\text{ave})}$$

$$= \frac{\frac{2}{31}}{\frac{6}{31}} = \frac{1}{3}$$

$X = Z_1 \qquad Y = \rho Z_1 + \sqrt{1-\rho^2}\, Z_2 \qquad$ (Z_1, Z_2 standard normal independent)

$$X = \sigma_X \cdot Z_1 + \mu_X \sim N(\mu_X, \sigma_X^2)$$

$$Y = \sigma_Y \cdot [\rho Z_1 + \sqrt{1-\rho^2} Z_2] + \mu_Y \sim N(\mu_Y, \sigma_Y^2)$$

$$= \mu_Y + \rho\sigma_Y \cdot Z_1 + \sqrt{1-\rho^2}\,\sigma_Y \cdot Z_2$$

$$(X,Y) \sim BVN(\mu_X, \sigma_X^2, \mu_Y, \sigma_Y^2, \rho)$$

二维正态

Standardised $BVN(0,1,0,1,\rho)$

$Corr(X,Y) = E(Z_X(X) Z_Y(Y)) = \rho$ (ρ不会因为标准化而变化)

$$(Y \mid Z_1 = z) \sim \mu_Y + \rho\sigma_Y z + \sqrt{1-\rho^2}\,\sigma_Y Z_2 \sim N(\mu_Y + \rho\sigma_Y \cdot z, (1-\rho^2)\sigma_Y^2)$$

$$X = x \iff \frac{X-\mu_X}{\sigma} = \frac{x-\mu_X}{\sigma} = z_1$$

$$(Y \mid X = x) \sim N\left(\mu_Y + \rho\sigma_Y \frac{x-\mu_X}{\sigma}, (1-\rho^2)\sigma_Y^2\right)$$

定义: (X, Y), $f(x,y)$ $f_X(x)$, $f_Y(y)$ 若 $f_Y(y) > 0$

X 在 $Y=y$ 条件下 $F(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$ $\quad F(y|x) = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)} dv$

$f(x|y) = \frac{f(x,y)}{f_Y(y)}$ $\left(\frac{\text{联合}}{\text{边缘}}\right)$

$f(x,y) = \frac{1}{\pi^2(1+x^2)(1+y^2)}$ $\quad \begin{cases} f_X(x) = \frac{1}{\pi(1+x^2)} \\ f_Y(y) = \frac{1}{\pi(1+y^2)} \end{cases}$

$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{1}{\pi(1+x^2)}$

$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{1}{\pi(1+y^2)}$

② $f(x,y) = \begin{cases} \frac{1}{\pi r^2} & x^2+y^2 \le r \\ 0 & \text{else} \end{cases}$ $\quad f_X(x) \begin{cases} \frac{2\sqrt{r^2-x^2}}{\pi r^2} & |x| \le r \\ 0 & \text{else} \end{cases}$ $\quad f_Y(y) = \begin{cases} \frac{2\sqrt{r^2-y^2}}{\pi r^2} & |y| < r \\ 0 & \text{else} \end{cases}$

$|y| < r \quad f(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{2\sqrt{r^2-y^2}} & -\sqrt{r^2-y^2} \le x \le \sqrt{r^2-y^2}. \\ 0 \end{cases}$

$F_{X|A}(x) = \begin{cases} 1 & x > b \\ \frac{F_X(x) - F_X(a)}{\underbrace{F_X(b) - F_X(a)}_{\text{constant}}} & a \le x < b \\ 0 & x < a \end{cases} \Rightarrow f_{X|A}(x) = \begin{cases} \overset{\text{differentiate}}{\frac{f_X(x)}{P(A)}} & a \le x < b \\ 0 & \text{otherwise} \end{cases}$

$E[X|A] = \int_{-\infty}^{+\infty} x f_{X|A}(x) dx \quad E[g(x)|A] = \int_{-\infty}^{+\infty} g(x) f_{X|A}(x) dx,$

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

$f(x) = e^{-1} =$

$P(A) = \int_{1}^{\infty} e^{-x} dx = \frac{1}{e}$ $\quad F_X(x) = 1 - e^{-x} = 1 - e^{-1}$

$f_{X|X>1}(x) = \frac{e^{-x}}{\frac{1}{e}} \quad x > 1 = e^{-x+1}$

$F_{X|A}(A) = \frac{F_X(x) - F_X(1)}{} = \frac{1 - e^{-x} - \frac{1}{e}}{} = \frac{1 - e^{-x} - 1 + e^{-1}}{\frac{1}{e}}$

$P(A)$ $\frac{1}{e}$ e

$$= \frac{e^{-1} - e^{-t}}{e^{-1}}$$

$$= 1 - e^{-x} \cdot e^{-1}$$

$$= 1 - e^{-x-1}$$

$X, Y \Rightarrow SBVN\ (\rho = 0.7)$

$P_r(X \geq 2 \mid Y = 2)$

$$Y \mid X = 2 \sim N\left(0 + 0.7 \times 1 \times \frac{2-0}{1},\ (1-\rho^2)\sigma_Y^2\right)$$

$$\sim N(1.4,\ 0.51)$$