中国科学技术大学
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
地址:中国安徽省合肥市 电话:0551-63602184 传真:0551-63631760 网址:http://www.ustc.edu.cn

Chapter 1

1. $P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\dots$

2. $P(B_i|A)=\dfrac{P(B_i)P(A|B_i)}{\sum_j P(B_j)P(A|B_j)}$

$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}$.

$\Gamma(x)=\int_0^{\infty}e^{-t}t^{x-1}dt\ (x>0)$

$B(x,y)=\int_0^1 t^{x-1}(1-t)^{y-1}dt$.

$B(x,y)=\dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

$\Gamma(\frac{n}{2})=(n-2)!!\,2^{-\frac{n-1}{2}}\sqrt{\pi}\quad(2\nmid n)$.

Chapter 2

1. Poisson 分布: $P(X=i)=\dfrac{e^{-\lambda}\lambda^i}{i!}$

正态分布: $f(x)=(\sqrt{2\pi}\sigma)^{-1}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

指数分布: $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0.\end{cases}$

均匀分布: $f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0, & 其他\end{cases}$

二维正态分布: $f(x_1,x_2)=(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1}\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-a)^2}{\sigma_1^2}-\frac{2\rho(x_1-a)(x_2-b)}{\sigma_1\sigma_2}+\frac{(x_2-b)^2}{\sigma_2^2}\right)\right]$.

2、边缘分布: $P(X_1=a_{1k})=\sum_{j_2,\dots,j_n}P(k,j_2,\dots,j_n)$.

$f_1(x_1)=\int_{-\infty}^{+\infty}f(x_1,x_2)dx_2$

$f_2(x_2)=\int_{-\infty}^{+\infty}f(x_1,x_2)dx_1$.

3、$f(x_1,x_2)=f_2(x_2)f_1(x_1|x_2)$.

$f(x_1,\dots,x_n)=g(x_1,\dots,x_n)h(x_{k+1},\dots,x_n|x_1,\dots,x_k)$

4、若连续型随机向量$(X_1,\dots,X_n)$的概率密度函数$f(x_1,\dots,x_n)$可表为n个函数$g_1,\dots,g_n$之积，其中g_i只依赖于x_i，即
$f(x_1,\dots,x_n)=g_1(x_1)\dots g_n(x_n)$，则$X_1,\dots,X_n$相互独立，且$X_i$ ~~的分布~~ 的边缘密度函数$f_i(x_i)$与$g_i(x_i)$只差常数因子$(C_1C_2\dots C_n=1)$.

5、若$X_1,\dots,X_n$相互独立，则$Y_1=g_1(X_1,\dots,X_m)$和$Y_2=g_2(X_{m+1},\dots,X_n)$独立.

6、设$Y=X_1+X_2$，则$l(y)=\int_{-\infty}^{+\infty}f(y-x,x)dx$.

地址:中国安徽省合肥市 电话:0551-63602184 传真:0551-63631760 网址:http://www.ustc.edu.cn

统计三大分布.

$D(\chi^2)=nD(X_i^2)=n\{E(X_i^4)-[E(X_i^2)]^2\}=2n.$

$E(\chi^2)=n.$

χ^2分布: $K_n(x)=\begin{cases}\frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}}e^{-\frac{x}{2}}x^{\frac{n-2}{2}}, & x>0\\ 0, & x\le 0.\end{cases}$

若$X_1,\dots,X_n$相互独立,都服从正态分布$N(0,1)$,则$Y=X_1^2+\dots+X_n^2$服从自由度为n的卡方分布χ_n^2.

1° 设X_1, X_2独立, $X_1\sim\chi_m^2$, $X_2\sim\chi_n^2$, 则$X_1+X_2\sim\chi_{m+n}^2$

2° 设$X_1,\dots,X_n$独立,且服从指数分布,则

$$X=2\lambda(X_1+\dots+X_n)\sim\chi_{2n}^2.$$

3° 设$X_1,\dots,X_n$独立同分布,有公共的正态分布$N(\mu,\sigma^2)$. 记$\bar{X}=\frac{X_1+\dots+X_n}{n}$,

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

t分布: $t_n(y)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\frac{y^2}{n}\right)^{-\frac{n+1}{2}}$.

t分布以标准正态分布为极限分布.

设X_1, X_2独立, $X_1\sim\chi_n^2$, $X_2\sim N(0,1)$, 则$Y=\frac{X_2}{\sqrt{X_1/n}}\sim t_n$.

$E(t)=0$

$D(t)=\frac{n}{n-2}$ (n>

设$X_1,\dots,X_n$独立同分布,有公共正态分布$N(\mu,\sigma^2)$,则

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

F分布: ~~f_{mn}~~ $f_{mn}(y)=m^{\frac{m}{2}}n^{\frac{n}{2}}\frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}y^{\frac{m}{2}-1}(my+n)^{-\frac{m+n}{2}}$ $(y>0)$.

设X_1, X_2独立, $X_1\sim\chi_n^2$, $X_2\sim\chi_m^2$, 则$Y=\frac{X_2/m}{X_1/n}\sim F_{mn}$.

$y\le 0$时,有$f_{mn}(y)=0$.

设$X_1,\dots,X_n, Y_1,\dots,Y_m$独立, X_i各有分布$N(\mu_1,\sigma_1^2)$, Y_i各有分布$N(\mu_2,\sigma_2^2)$,则 $\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}\sim F_{m-1,n-1}$.

若$\sigma_1^2=\sigma_2^2$,则 $\frac{[(\bar{X}-\bar{Y})-(\mu_1-\mu_2)]}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim t_{n+m-2}$,

其中$S_w=\sqrt{\frac{\sum_{i=1}^n(X_i-\bar{X})^2+\sum_{j=1}^m(Y_j-\bar{Y})^2}{n+m-2}}$

1° 若 $F\sim F(n_1,n_2)$，则 $\frac{1}{F}\sim F(n_2,n_1)$.

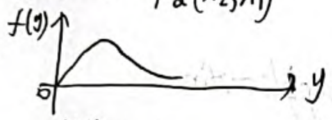
2° 若 $t\sim t(n)$，则 $t^2\sim F(1,n)$.

上 α 分位点的性质:

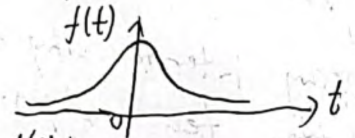
$$t_{1-\alpha}(n)=-t_\alpha(n).$$

$$F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$$

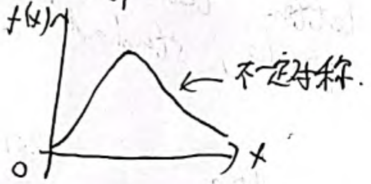
χ^2 分布:



t 分布:



F 分布.



Chapter 3 1. 若 $\int_{-\infty}^{+\infty}|x|f(x)\,dx<\infty$，则 $E(X)=\int_{-\infty}^{+\infty}xf(x)\,dx$.

2. $E(X_1+\dots+X_n)=E(X_1)+\dots+E(X_n)$.

若 $X_1,\dots,X_n$ 独立，则 $E(X_1\dots X_n)=E(X_1)\dots E(X_n)$.

$E(cX)=cE(X)$.

3. 条件数学期望: $E(X|Y=y)=\int_{-\infty}^{+\infty}x\,p_{X|Y}(x|y)\,dx$.

$\varphi(y)$，关于 y 的函数.

4. X,Y 为 r.v.，$EX,EY,Eg(Y)$ 存在，则

1° $a\le X\le b\Rightarrow a\le E(X|Y=y)\le b$

2° $E(C_1X_1+C_2X_2|Y=y)=C_1E(X_1|Y=y)+C_2E(X_2|Y=y)$.

3° $E[E(X|Y)]=EX$.

4° X,Y 独立 $\Rightarrow E(Y|X)=E(Y)$

5° $E(g(X)Y|X)=g(X)E(Y|X)$

6° $E(c|X)=c$

7° $E(g(X)|X)=g(X)$.

8° $E(Y-E(Y|X))^2\le E(Y-g(X))^2$

5、$Var(X)=E(X^2)-(EX)^2$.

$Var(X+c)=Var(X)$

$Var(cX)=c^2Var(X)$

若$X_1\sim X_n$独立，则$Var(X_1+\dots+X_n)=Var(X_1)+\dots+Var(X_n)$.

6、k阶原点矩：$\alpha_k=E(X^k)$.

k阶中心矩：$\mu_k=E[(X-EX)^k]$.

7、偏度系数：$\frac{\mu_3}{\mu_2^{\frac{3}{2}}}$

峰度系数：$\beta=\frac{\mu_4}{\mu_2^2}-3$. 正态分布的峰度系数为0.

8、$|E(XY)|\le\sqrt{EX^2EY^2}$ 等号成立 $\Leftrightarrow\exists a,b,a^2+b^2\neq0$, s.t. $P(aX+bY=0)=1$.

9、协方差：$Cov(X,Y)=E[(X-m_1)(Y-m_2)]$

$Cov(X,Y)=0$时，称X，Y不相关. $Cov(X,Y)=E(XY)-(EX)(EY)$.

10、相关系数：$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}$.

$|\rho_{XY}|=1$时，X，Y线性相关.

11、协方差矩阵：$\Sigma=\begin{pmatrix}\sigma_{11}&\sigma_{12}\\\sigma_{21}&\sigma_{22}\end{pmatrix}$.

1° 半正定.

2° Σ退化$\Leftrightarrow\exists a_1,a_2,a_1^2+a_2^2\neq0$, s.t. $P(\sum_{i=1}^{2}a_i(X_i-\mu)=0)=1$.

12、设$Y_1,Y_2\sim N(0,1)$且独立，$ad-bc\neq0$,

$\begin{cases}X_1=aY_1+bY_2+\mu_1\\X_2=cY_1+dY_2+\mu_2\end{cases}$，则$Cov(\cancel{X,Y}_{X_1,X_2})=ac+bd$,

$(X_1,X_2)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$，其中$\sigma_1^2=a^2+b^2$; $\sigma_2^2=c^2+d^2$; $\rho=\frac{ac+bd}{\sigma_1\sigma_2}$.

X_1,X_2独立$\Leftrightarrow X_1,X_2$不相关.

1101C-08 201412·2500

13、大数定理：设 $X_1, X_2, \ldots$ 独立同分布，$\mu = EX_1$，则
$\forall \varepsilon > 0$，有 $\lim_{n\to\infty} P(|\overline{X}_n - \mu| \geq \varepsilon) = 0$.

14、设 $X_1, X_2, \ldots, X_n, \ldots$ 独立同分布，X_i 的分布是
$P(X_i = 1) = p$，$P(X_i = 0) = 1-p$.
则 $\forall x \in R$，有 $\lim_{n\to\infty} P\left(\frac{1}{\sqrt{np(1-p)}}(X_1 + \ldots + X_n - np) \leq x\right) = \Phi(x)$.
$P(t_1 \leq X_1 + \ldots + X_n \leq t_2) = \Phi(y_2) - \Phi(y_1)$，
其中 $y_i = \frac{t_i - np}{\sqrt{np(1-p)}}$

15、贝叶斯估计：$h(\theta | X_1, \ldots, X_n) = h(\theta) f(X_1, \theta) \ldots f(X_n, \theta) / p(X_1, \ldots, X_n)$
$\tilde{\theta} = \int_0^1 \theta h(\theta | X_1, \ldots, X_n) d\theta = E[h(\theta | X_1, \ldots, X_n)]$.

16、若 $\forall (\theta_1, \ldots, \theta_k)$，$E_{\theta_1, \ldots, \theta_k}[\hat{g}(X_1, \ldots, X_n)] = g(\theta_1, \ldots, \theta_k)$，
则 $\hat{g}$ 是 $g(\theta_1, \ldots, \theta_k)$ 的一个无偏估计量.

17. 区间估计

1° 已知 σ，估计 μ.
$$\left[\overline{X}_n - \frac{z_{\alpha/2}\sigma}{\sqrt{n}}, \overline{X}_n + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right].$$

2° 未知 σ，估计 μ.
$$\left[\overline{X}_n - \frac{t_{\alpha/2}(n-1)S}{\sqrt{n}}, \overline{X}_n + \frac{t_{\alpha/2}(n-1)S}{\sqrt{n}}\right].$$

3° 估计 σ^2.
$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right].$$

单侧置信区间.

1° 已知 σ. $\overline{X}_n \pm \frac{z_\alpha \sigma}{\sqrt{n}}$.

2° 未知 σ. $\overline{X}_n \pm \frac{t_\alpha(n-1)S}{\sqrt{n}}$

3° 估计 σ^2: 上限: $\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$ 下限: $\frac{(n-1)S^2}{\chi^2_{\alpha}(n-1)}$.

估计 $\mu_1-\mu_2$: ~~设 $\sigma_1^2/\sigma_2^2=b^2$~~.

1° 已知 σ_1^2 和 σ_2^2: $\left[(\overline{X}_n-\overline{Y}_m)-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}, (\overline{X}_n-\overline{Y}_m)+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\right]$.

2° $\sigma_1^2=\sigma_2^2$: $\left[(\overline{X}_n-\overline{Y}_m)-t_{\alpha/2}S_w\sqrt{\frac{1}{n}+\frac{1}{m}}, (\overline{X}_n-\overline{Y}_m)+t_{\alpha/2}S_w\sqrt{\frac{1}{n}+\frac{1}{m}}\right]$.

3° $\sigma_1^2=b\sigma_2^2$: $\left[(\overline{X}_n-\overline{Y}_m)-t_{\alpha/2}S_b\sqrt{\frac{b^2}{n}+\frac{1}{m}}, \dots\right]$

其中 $S_b^2=\frac{(n-1)S_1^2/b^2+(m-1)S_2^2}{n+m-2}$.

估计 σ_1^2/σ_2^2. $\left[\frac{S_1^2/S_2^2}{F_{\alpha/2}}, \frac{S_1^2/S_2^2}{F_{1-\alpha/2}}\right]$.

Chapter ~~A~~6

1. 正态分布的显著性检验.

1° 已知 σ 时, μ 的检验.

$H_0:\mu=\mu_0$ vs $H_1:\mu\neq\mu_0$.

$W_\alpha=\{|Z|\geq Z_{\alpha/2}\}$. 其中 $Z=\frac{\overline{X}_n-\mu_0}{\sigma/\sqrt{n}}$.

单边检验: $H_0:\mu\leq\mu_0$. $W_\alpha=\{Z\geq Z_\alpha\}$.

2° 未知 σ 时, μ 的检验.

令 $T=\frac{\overline{X}_n-\mu_0}{S/\sqrt{n}}$, 则 $W_\alpha=\{|T|\geq t_{\alpha/2}(n-1)\}$.

单边: $W_\alpha=\{T\geq t_\alpha(n-1)\}$

2. 均值比较的显著性检验.

1° 已知 σ_1^2,σ_2^2, μ_1,μ_2 的检验.

$H_0:\mu_1=\mu_2$ vs $H_1:\mu_1\neq\mu_2$.

$Z=\frac{\overline{X}_n-\overline{Y}_m}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}$. $W_\alpha=\{|Z|\geq Z_{\alpha/2}\}$.

单边: $H_0:\mu_1\leq\mu_2$ vs $H_1:\mu_1>\mu_2$. $W_\alpha=\{Z\geq Z_\alpha\}$.

2° 已知 $\sigma_1^2=\sigma_2^2$ 时，$\mu_1-\mu_2$ 的检验.

$H_0:\mu_1=\mu_2$ vs $H_1:\mu_1\neq\mu_2$.

令 $T=\dfrac{\overline{X}_n-\overline{Y}_m}{S_w\sqrt{\frac{1}{n}+\frac{1}{m}}}$

则 $W_\alpha=\{|T|\geq t_{\frac{\alpha}{2}}(m+n-2)\}$.

单边：$H_0:\mu_1\leq\mu_2$ vs $H_1:\mu_1>\mu_2$.

$W_\alpha=\{T\geq t_\alpha(m+n-2)\}$.

3° 成对数据.

$H_0:\mu=0$ vs $H_1:\mu\neq 0$.

令 $Z_i=X_i-Y_i$.

$T=\dfrac{\overline{Z}_n}{S_z/\sqrt{n}}\sim t(n-1)$.

则 $W_\alpha=\{|T|\geq t_{\frac{\alpha}{2}}(n-1)\}$

4° 未知 σ_1^2,σ_2^2 时，μ_1,μ_2 的大样本检验.

$Z_0=\dfrac{\overline{X}_n-\overline{Y}_m}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}\approx\dfrac{\overline{X}_n-\overline{Y}_m}{\sqrt{S_1^2/n+S_2^2/m}}\sim N(0,1)$.

5°. $H_0:\sigma_1^2\leq\sigma_0^2$ vs $H_1:\sigma_1^2>\sigma_0^2$.

令 $U_0=\dfrac{(n-1)S^2}{\sigma_1^2}\sim\chi^2(n-1)$.

则 $W=\{U_0\leq\chi^2_{1-\alpha/2}(n-1)\}\cup\{U_0\geq\chi^2_{\alpha/2}(n-1)\}$. ($\chi^2$检验)

6° 令 $F_0=\dfrac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}$，则 $W_\alpha=\{F\leq F_{1-\alpha}(58,40)\}$. $H_0:\sigma_1^2=\sigma_2^2$.

$F=\dfrac{S_1^2}{S_2^2}$ (F检验).

3. 拟合优度检验.

$H_0:X\sim F(x)$ vs $H_1:F(x)$ 不是 X 的分布函数.

1101C-08 201412·2500

取 $t_0<\min\{X_1,\dots,X_n\}$, $t_m>\max\{X_1,\dots,X_n\}$.

取 $t_0<t_1<\dots<t_m$, 令 $I_j=(t_{j-1},t_j]$, $j=1,2,\dots,m$.

记 $\hat{p}_j=\frac{1}{n}\sum_{k=1}^{n} I[X_k\in I_j]$ 作为 $p_j=P(X\in I_j)=F(t_j)-F(t_{j-1})$ 的估计.

令 $U=\sum_{j=1}^{m}\frac{n}{p_j}(\hat{p}_j-p_j)^2$. $U\sim\chi^2(m-1)$ (近似服从)

$W=\{U>\chi^2_\alpha(m-1)\}$.

若 $F(x)$ 中有 r 个未知参数，则先计算出这 r 个参数的最大似然估计. $W=\{U>\chi^2_\alpha(m-r-1)\}$.

要求 $np_j\geq 5$, $1\leq j\leq m$.

4. 列联表.

p_{11}	p_{12}	p_1
p_{21}	p_{22}	p_2
q_1	q_2	

边缘概率.

$$V_n=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{n(\hat{p}_{ij}-\hat{p}_i\hat{q}_j)^2}{\hat{p}_i\hat{q}_j}=\frac{n(n_{11}n_{22}-n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}\sim\chi^2(1)$$

$$W_\alpha=\{V_n\geq\chi^2_\alpha(1)\}.$$

$k\times l$ 列联表

H_0: X,Y 独立 vs H_1: X,Y 不独立.

$$V_n=\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n(\hat{p}_{ij}-\hat{p}_i\hat{q}_j)^2}{\hat{p}_i\hat{q}_j}=n\left(\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}}-1\right)\sim\chi^2((k-1)(l-1))$$

$$W_\alpha=\{V_n\geq\chi^2_\alpha\}.$$