Cheat Sheet

1. 条件概率

事件的差:

$$A - B = A - AB = A\bar{B}$$

乘法公式:

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

全概率公式:

$$P(B) = \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

贝叶斯公式:

$$P(A_i|B) = \frac{P(A_iB)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}$$

$$\text{后验概率} = \frac{\text{先验概率} \times \text{似然度}}{\text{证据概率}}$$

独立性:

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

2. 离散型随机变量

琴生不等式:

利用琴生不等式有, 对于连续凸函数 g 有

$$g(E(X)) \leqslant E(g(X))$$

例如有

$$(E(X))^2 \leqslant E(X^2) \text{ 和 } e^{E(X)} \leqslant E(e^X)$$

方差:

$\operatorname{Var}(X) = D(X) = E[(X-E(X))^2] = E(X^2) - [E(X)]^2$

Bhatia-Davis 不等式:

对于 $X \in [a,b]$, 有

$\operatorname{Var}(X) \leqslant (b-E(X))(E(X)-a) \leqslant (b-a)^2/4$

Bernoulli 分布:

分布列为 $P(X=1)=p, P(X=0)=1-p$, 记作 $X \sim \operatorname{Ber}(p)$

有 $E(X)=p, \operatorname{Var}(X)=p(1-p)$

二项分布:

重复进行了 n 次 Bernoulli 试验, 记作事件 A, 随机变量 X 表示事件 A 发生的次数.

分布列为 $P(X=k)=\dbinom{n}{k}p^k(1-p)^{n-k}$, 记作 $X \sim B(n,p)$

有 $E(X)=np, \operatorname{Var}(X)=np(1-p)$

用级数求和, 两边求导的方法进行计算证明. 用到二项展开式 $(1+x)^n = \displaystyle\sum_{i=1}^{n}\binom{n}{k}x^k$

几何分布:

重复进行了 n 次 Bernoulli 试验, 记作事件 A, 随机变量 X 表示事件 A 首次发生的试验次数.

分布列为 $P(X=k)=(1-p)^{k-1}p$, 记作 $X \sim G(p)$

有 $E(X)=\dfrac{1}{p}$ 和 $\operatorname{Var}(X)=\dfrac{1-p}{p^2}$

等比数列求和 $S_n = a_1\dfrac{1-q^n}{1-q}$

依旧使用级数 $(1-x)^{-1} = \displaystyle\sum_{k=1}^{\infty}x^k$ 和两边求导来证明.

几何分布拥有无记忆性: $P(X>m+n|X>m)=P(X>n)$

Pascal / 负二项分布:

重复进行了 n 次 Bernoulli 试验, 记作事件 A, 随机变量 X 表示事件 A 第 r 次发生的试验次数.

分布列为 $P(X=k)=\binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}\cdot p=\binom{k-1}{r-1}p^r(1-p)^{k-r}$, 则称 X 为服从参数 p 和 r 的负二项分布.

重点在于 $(1-q)^{-r}=\sum_{t=0}^{\infty}\binom{t+r-1}{r-1}q^t$, 即 $p^{-r}=\sum_{k=r}^{\infty}\binom{k-1}{r-1}(1-p)^{k-r}$

有 $E(X)=\frac{r}{p}$ 和 $\mathrm{Var}(X)=\frac{r(1-p)}{p^2}$

证明的时候只要想办法往概率求和等于一上转化即可.

还需要一个性质: $\frac{k}{r}\cdot\binom{k-1}{r-1}=\binom{k}{r}$

泊松分布:

分布列为 $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, 记作 $X\sim P(\lambda)$

我们知道泰勒展开式 $e^{\lambda}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}$

有 $E(X)=\lambda$ 和 $\mathrm{Var}(X)=\lambda$

泊松定理:

对任意给定的常数 $\lambda>0$, n 为任意正整数, 设 $np=\lambda$, 则对任意给定非负整数 k 有

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

对于随机变量 $X\sim B(n,p)$, 当 n 较大而 p 较小时, 令 $\lambda=np$, 则有

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda}$$

即可以用泊松分布来近似计算二项分布.

3. 连续型随机变量

非负随机变量期望:

$E[X]=\int_0^{\infty} P(X>t)\mathrm{d}t$

证明:

首先观察得到 $X=\int_0^X 1\mathrm{d}t=\int_0^{\infty} 1(X>t)\mathrm{d}t$

$$
\begin{aligned}
E[X] &= E[\int_0^{\infty} 1(X>t)\mathrm{d}t] \\
&= \int_{-\infty}^{+\infty}\int_0^{\infty} 1(x>t)f(x)\mathrm{d}t \\
&= \int_0^{+\infty}[\int_{-\infty}^{+\infty} 1(X>t)f(x)\mathrm{d}x]\mathrm{d}t \\
&= \int_0^{+\infty}[\int_{-\infty}^{t} 1(X>t)f(x)\mathrm{d}x+\int_t^{+\infty} 1(X>t)f(x)\mathrm{d}x]\mathrm{d}t \\
&= \int_0^{+\infty}[\int_t^{+\infty} f(x)\mathrm{d}x]\mathrm{d}t \\
&= \int_0^{+\infty} P(X>t)\mathrm{d}t
\end{aligned}
$$

均匀分布:

概率密度 $f(x)=\dfrac{1}{b-a}, x\in[a,b]$, 则记作 $X\sim U(a,b)$

而分布函数则为 $F(x)=\dfrac{x-a}{b-a}, a<x<b$

期望和方差分别为 $E(X)=\dfrac{a+b}{2}, \mathrm{Var}(X)=\dfrac{(b-a)^2}{12}$

指数分布:

$f(x)=\lambda e^{-\lambda x}, x\geqslant 0$, 而 $F(x)=1-e^{-\lambda x}$, 记作 $X\sim e(\lambda)$

期望和方差为 $E(X)=\dfrac{1}{\lambda}, \mathrm{Var}(X)=\dfrac{1}{\lambda^2}$

证明用分部积分法.

指数分布具有无记忆性, 且是唯一具有无记忆性的连续型随机变量:

$P(X>s+t|X>t)=P(X>s)$

正态分布:

概率密度为 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, 记作 $X\sim N(\mu,\sigma^2)$

标准正态分布为 $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, 记作 $X\sim N(0,1)$

期望和方差为 $E(X)=\mu, \mathrm{Var}(X)=\sigma^2$

正态分布的估计:

对于标准正态分布 $X\sim N(0,1)$ 和任意 $\epsilon>0$, 有

$$P(X\geqslant\epsilon)\leqslant\frac{1}{2}e^{-\frac{\epsilon^2}{2}}$$

还有 $P(|X|\geqslant\epsilon)\leqslant\min\{1,\sqrt{\frac{2}{\pi}}\frac{1}{\epsilon}e^{-\frac{\epsilon^2}{2}}\}$

4. 多维随机变量及其分布

随机变量独立性:

$$F(x,y)=F_X(x)F_Y(y)$$

对于离散型随机变量来说, 即等价于 $p_{i,j}=p_{i\cdot}p_{\cdot j}$

对于连续型随机变量来说, 即等价于 $f(x,y)=f_X(x)f_Y(y)$

二维正态分布:

令 $\mu=\begin{pmatrix}\mu_x\\ \mu_y\end{pmatrix}$ 和 $\Sigma=\begin{pmatrix}\sigma_x^2 & \rho\sigma_x\sigma_y\\ \rho\sigma_x\sigma_y & \sigma_y^2\end{pmatrix}$

则 X 和 Y 的联合概率密度函数为

$$f(x,y)=(2\pi)^{-2/2}|\Sigma|^{-\frac{1}{2}}\exp(-\frac{1}{2}(\xi-\mu)^T\Sigma^{-1}(\xi-\mu))$$

其中 $\Sigma^{-1}=\frac{1}{(1-\rho^2)\sigma_x^2\sigma_y^2}\begin{pmatrix}\sigma_y^2 & -\rho\sigma_x\sigma_y\\ -\rho\sigma_x\sigma_y & \sigma_x^2\end{pmatrix}$

极大极小分布:

假设 X,Y 相互独立.

极大分布:

$$F_Z(z) = P(Z \leqslant z) = P(\max(X,Y) \leqslant z) = P(X \leqslant z, Y \leqslant z) = F_X(z)F_Y(z)$$

极小分布:

$$F_Z(z) = 1 - (1 - F_X(z))(1 - F_Y(z))$$

和的分布:

对于 $Z = X + Y$,

通用的有 $F_Z(z) = \iint_{x+y\leqslant z} f(x,y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{z} f(x,u-x)\mathrm{d}u = \int_{-\infty}^{z}\int_{-\infty}^{+\infty} f(x,u-x)\mathrm{d}x\mathrm{d}u$

两边求导可得 $f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x$

假设 X,Y 相互独立, 有著名的卷积公式

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)\mathrm{d}y$$

乘除的分布:

$$f_{XY}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f(x, \frac{z}{x})\mathrm{d}x$$

$$f_{X/Y}(z) = \int_{-\infty}^{+\infty} |x| f(x, xz)\mathrm{d}x$$

联合分布函数:

设有 $U = u(X,Y), V = v(X,Y)$

$$f_{UV}(u,v) = f_{XY}(x(u,v), y(u,v)) \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

多维随机变量柯西不等式:

$$E[XY] \leqslant \sqrt{E[X^2]E[Y^2]}$$

协方差:

定义协方差为 $\mathrm{Cov}(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$

则有 $\operatorname{Var}(X+Y)=\operatorname{Var}(X)+\operatorname{Var}(Y)+2\operatorname{Cov}(X,Y)$

对任意 X_1,X_2,Y 还有性质 $\operatorname{Cov}(X_1+X_2,Y)=\operatorname{Cov}(X_1,Y)+\operatorname{Cov}(X_2,Y)$

并且 $\operatorname{Cov}(aX,bY)=ab\operatorname{Cov}(X,Y)$

若 X 和 Y 独立, 则 $\operatorname{Cov}(X,Y)=0$, 反之则不然.

协方差不等式:

$[\operatorname{Cov}(X,Y)]^2\leqslant\operatorname{Var}(X)\operatorname{Var}(Y)$, 当且仅当 $Y=aX+b$ 等号成立.

相关系数:

$$\rho_{XY}=\frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}$$

条件概率:

先有 $f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}$

再有 $F_{X|Y}(x|y)=P(X\leqslant x|Y=y)=\displaystyle\int_{-\infty}^{x}f_{X|Y}(u|y)\mathrm{d}u$

乘法公式:

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y)$$

5. 集中不等式

Markov 不等式:

对任意随机变量 $X\geqslant 0,\epsilon>0$, 有

$$P(X\geqslant\epsilon)\leqslant\frac{E(X)}{\epsilon}$$

证明:

$$E[X]=E[X|X\geqslant\epsilon]P(X\geqslant\epsilon)+E[X|X\leqslant\epsilon]P(X\leqslant\epsilon)\geqslant\epsilon P(X\geqslant\epsilon)$$

推论: 对于单调增的非负函数 g 有 $P(X\geqslant\epsilon)\leqslant\dfrac{E[g(X)]}{g(\epsilon)}$

Chebyshev 不等式:

$$P(|X-\mu|>\epsilon)\leqslant\frac{\mathrm{Var}(X)}{\epsilon^2}$$

Cantelli 不等式:

$$P(X-\mu\geqslant\epsilon)\leqslant\frac{\sigma^2}{\sigma^2+\epsilon^2}\text{ 和 }P(X-\mu\leqslant-\epsilon)\leqslant\frac{\sigma^2}{\sigma^2+\epsilon^2}$$

证明过程中令 $Y=X-\mu$, 并添加了一个 t 变量, 用以之后求最值.

Chebyshev 不等式推论:

独立同分布随机变量 $E[X_i]=\mu,\mathrm{Var}(X_i)\leqslant\sigma^2$

$$P(|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|\geqslant\epsilon)\leqslant\frac{\sigma^2}{n\epsilon^2}$$

Young 不等式:

对于 $\frac{1}{p}+\frac{1}{q}=1$, 有

$$ab\leqslant\frac{1}{p}a^p+\frac{1}{q}b^q$$

Holder 不等式:

$$E[|XY|]\leqslant(E[|X|^p])^{\frac{1}{p}}(E[|Y|^q])^{\frac{1}{q}}$$

Chernoff 不等式:

矩生成函数为 $M_X(t)=E[e^{tX}]$

Chernoff 方法为:

$$P(X\geqslant E[X]+\epsilon)=P(e^{tX}\geqslant e^{tE[X]+t\epsilon})\leqslant e^{-t\epsilon-tE[X]}E[e^{tX}]$$

Chernoff 引理:

对于 $X\in[0,1]$ 的期望 $\mu=E[X]$, 对任意 $t>0$ 有

$$E[e^{tX}]\leqslant\exp(t\mu+\frac{t^2}{8})$$

使用凸函数性质 $e^{tX}=e^{tX+(1-X)0}\leqslant Xe^t+(1-X)e^0$

因此对 $X\in[a,b]$ 有 $E[e^{tX}]\leqslant\exp(\mu t+\dfrac{t^2(b-a)^2}{8})$

亚高斯随机变量:

将有界随机变量和高斯随机变量统一起来.

若 X 满足 $E[e^{(X-E[X])t}] \leqslant \exp(\frac{bt^2}{2})$

则称为亚高斯随机变量.

高斯分布是参数为 σ^2 的亚高斯分布.

对于满足 $E[X_i]=0$ 的亚高斯随机变量有

$E[\max\limits_{i\in[n]} X_i] \leqslant \sqrt{2b\ln n}$

根据 Jensen 不等式有 $\exp(tE[\max\limits_{i\in[n]} X_i]) \leqslant E[\exp(t\max\limits_{i\in[n]} X_i)]$

6. 大数定理以及中心极限定理

依概率收敛:

$\lim\limits_{n\to\infty} P(|X_n-a|<\epsilon)=1$ 或 $\lim\limits_{n\to\infty} P(|X_n-a|\geqslant\epsilon)=0$

则称 $\{X_n\}$ 依概率收敛于 a, 记作 $X_n \xrightarrow{P} a$

大数定律:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} E[X_i]$$

即看 $P(|\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-E[X_i])|\geqslant\epsilon)\to 0$

Markov 大数定律:

$$\frac{1}{n^2}\mathrm{Var}(\sum_{i=1}^{n} X_i)\to 0$$

Chebyshev 大数定律:

$\mathrm{Var}(X_n)\leqslant c$ 即 $\frac{c}{n\epsilon^2}\to 0$

辛钦大数定律:

每个随机变量的期望 $E[X_i]=\mu$ 存在.

Bernoulli 大数定律:

对于 $X_n \sim B(n,p)$ 可以看作是一系列的 Bernoulli 随机变量, 然后就有 $\dfrac{X_n}{n} \xrightarrow{P} p$

判断随机变量序列满足大数定律:

独立同分布, 则用辛钦大数定律; 否则用 Markov 大数定律.

依分布收敛:

$$\lim_{n\to\infty} F_{Y_n}(y) = F_Y(y)$$

中心极限定理:

对于独立同分布的随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 的期望 $E[X_i]=\mu$ 和方差 $\mathrm{Var}(X_i)=\sigma^2$

则 $Y_n = \dfrac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$

变形公式为:

$$\sum_{i=1}^n X_i \xrightarrow{d} N(n\mu, n\sigma^2) \text{ 和 } \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{d} N(\mu, \frac{\sigma^2}{n})$$

7. 统计基本概念

样本均值:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$$

且有 $E[\bar{X}]=\mu, \mathrm{Var}(\bar{X}) = \dfrac{\sigma^2}{n}, \bar{X} \xrightarrow{d} N(\mu, \dfrac{\sigma^2}{n})$

样本方差:

$$S_0^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{X}^2$$

则有 $E[S_0^2] = \dfrac{n-1}{n}\sigma^2$

修正后的样本方差:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \text{ 即 } S^2 = \frac{n}{n-1}S_0^2$$

则有 $E[S^2] = \sigma^2$

样本 k 阶原点矩:

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

样本 k 阶中心矩:

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

第 k 次序统计量:

$$F_k(x) = \sum_{r=k}^{n}\binom{n}{r}[F(x)]^r[1-F(x)]^{n-r}$$

$$f_k(x) = \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}[1-F(x)]^{n-k}f(x)$$

Beta 函数:

$$\mathrm{Beta}(\alpha_1, \alpha_2) = \int_0^1 x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}\mathrm{d}x$$

Gamma 函数:

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}\mathrm{d}x$$

$$\Gamma(1) = 1, \Gamma(\tfrac{1}{2}) = \sqrt{\pi}, \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1), \Gamma(n+1) = n!$$

$$\mathrm{Beta}(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$$

$$\mathrm{Beta}(\alpha_1, \alpha_2) = \frac{\alpha_1 - 1}{\alpha_1 + \alpha_2 - 1}\mathrm{Beta}(\alpha_1 - 1, \alpha_2)$$

Beta 分布:

$f(x) = \frac{x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}}{\mathrm{Beta}(\alpha_1, \alpha_2)}, x \in (0,1)$

记作 $X \sim B(\alpha_1, \alpha_2)$

有 $E[X] = \frac{\alpha_1}{\alpha_1 + \alpha_2}, \mathrm{Var}(X) = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)}$

均匀分布第 k 次序统计量:

若 $X_1, X_2, \cdots, X_n$ 服从 $U(0,1)$, 则

$X_{(k)} \sim B(k, n-k+1)$

Gamma 分布:

$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\lambda x}, x > 0$

记作 $X \sim \Gamma(\alpha, \lambda)$

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

和指数分布比较类似.

Gamma 分布具有可加性: $X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$

特别的, $X \sim \Gamma(\frac{1}{2}, \frac{1}{2})$ 有 $f(x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{1}{2}x}, x > 0$

且有若 $X \sim N(0,1)$, 则 $X^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$

卡方分布:

若 $X_1, X_2, \cdots, X_n$ 是来自总体 $X \sim N(0,1)$ 的一个样本, 称 $Y = X_1^2 + X_2^2 + \cdots + X_n^2$ 为服从自由度为 n 的 χ^2 分布, 记作 $Y \sim \chi^2(n)$

所以也就有 $Y \sim \Gamma(\frac{n}{2}, \frac{1}{2})$

$E[Y] = n, \mathrm{Var}(Y) = 2n$

可加性: $X + Y \sim \chi^2(m+n)$

若随机变量 $X \sim N(0,1)$, 则 $E[X^k] = \begin{cases} (k-1)!!, & k \text{ is even} \\ 0, & k \text{ is odd} \end{cases}$

t 分布:

设 $X \sim N(0,1)$ 和 $Y \sim \chi^2(n)$ 相互独立, 则称

$T = \frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布, 记作 $T \sim t(n)$

F 分布:

设随机变量 $X \sim \chi^2(m)$ 和 $Y \sim \chi^2(n)$ 相互独立, 则称

$F = \frac{X/m}{Y/n}$ 为服从自由度 m, n 的 F 分布, 记作 $F \sim F(m,n)$

五大抽样定理其一:

设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 的样本, 则有

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

在知方差 σ 时能用于估计 μ; 知期望 μ 时能用于估计 σ^2.

五大抽样定理其二:

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

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

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

能用于估计方差 σ^2

五大抽样定理其三:

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

可用于估计期望 μ

五大抽样定理其四:

设 $X_1, X_2, \cdots, X_m$ 和 $Y_1, Y_2, \cdots, Y_n$ 分别来自总体 $N(\mu_X, \sigma^2), N(\mu_Y, \sigma^2)$, 其中两者方差一致, 则

$$\frac{\bar{X}-\bar{Y}-(\mu_X-\mu_Y)}{\sqrt{\frac{(m-1)S_X^2+(n-1)S_y^2}{m+n-2}}\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t(m+n-2)$$

五大抽样定理其五:

设 $X_1, X_2, \cdots, X_m$ 和 $Y_1, Y_2, \cdots, Y_n$ 分别来自总体 $N(\mu_X, \sigma_X^2), N(\mu_Y, \sigma_Y^2)$, 其中两者方差不一定一致, 则

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F(m-1, n-1)$$

分位点:

给定 $\alpha \in (0,1)$ 和随机变量 X, 称 $P(X > \lambda_\alpha) = \alpha$ 的实数 λ_α 为上侧 α 分位点.

8. 参数估计

矩估计:

使用样本 k 阶矩和样本 k 阶中心矩相等进行估计.

最大似然估计

似然函数 $L(\theta) = L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} P(x_i; \theta)$

取其对数 $\ln L(\theta) = \sum_{i=1}^{n} \ln P(x_i; \theta)$ 然后进行求导等于零

并令其等于零, 即可解出 $\hat{\theta}$

我们有 $\hat{\mu} = \mu(\hat{\theta})$ 是 μ 的最大似然估计.

无偏性:

$$E[\hat{\theta}(X_1, X_2, \cdots, X_n)] = \theta$$

则 $\hat{\theta}$ 是 θ 的无偏估计.

有效性:

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

一般方差越小, 无偏估计越好.