四 大 分布

$N(\mu, \sigma^2)$

μ 1次 μ ([illegible])

σ^2 2次 μ

$\bar{x}$ 来估计 μ

$\sim N(\mu, \frac{\sigma^2}{n})$

$\sum_{i=1}^{n}(x_i-\mu)^2$

正态分布的平方和

$$
\begin{tikzcd}
Z \arrow[r, "2"'] & \chi^2_1 \text{（单正态）} \\
t_n \arrow[u, "n\to\infty"] \arrow[r, "2"'] & F_{1,n} \text{（二正态）.}
\end{tikzcd}
$$

对称、

不对称、

八大抽样分布

(一) 单正态分布

$x_1, \cdots, x_n \leftarrow X \sim N(\mu, \sigma^2)$

μ 估计. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

- σ^2 已知 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ ①
- σ^2 未知 $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ ②

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

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

σ^2 估计. $X_i \sim N(\mu, \sigma^2)$.

- μ 已知 $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2_n$ ③
- μ 未知 $\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \equiv \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ ④

只要与 S^2 有关的，自由度下降1.

公式②的推导要求 $\bar{X}$ 与 S^2 独立！

（但此处 $\bar{X}$ 与 S^2 独立仅对正态总体而言！）

$x_1, x_2, \cdots, x_i, \cdots, x_n$ 独立.

$x_1-\bar{x}, x_2-\bar{x}, \cdots, x_i-\bar{x}, \cdots, x_n-\bar{x}$.

$\quad \| \qquad\quad \| \qquad\qquad\quad \| \qquad\qquad\quad \|$

$\;\; Y_1 \qquad\; Y_2 \qquad\qquad\; Y_i \qquad\qquad\; Y_n$

$Y_1+Y_2+\cdots+Y_n = x_1+\cdots+x_n-n\bar{x} \equiv 0$

（有一个关系式，自由度下降1）

注：有两个概念的判定必须知道总体的分布
- 判定独立.
- 最大似然估计.

八大抽样分布.

(二) 二正态总体.

$X_1, \cdots, X_m$ ↙ $X\sim N(\mu_1,\sigma_1^2)$　　$Y_1, \cdots, Y_n$ ↙ $Y\sim N(\mu_2,\sigma_2^2)$

$\mu_1-\mu_2$ 估计.

$\bar{X}-\bar{Y}\sim N(\mu_1-\mu_2,\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n})$

- σ_1^2, σ_2^2 已知

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\sim N(0,1) \quad ⑤$$

- σ_1^2, σ_2^2 未知但相等为 σ^2

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim t_{m+n-2} \quad ⑥$$

其中 $S_w^2=\frac{m-1}{m+n-2}S_1^2+\frac{n-1}{m+n-2}S_2^2$

$S_w=\sqrt{S_w^2}$.

$\frac{\sigma_1^2}{\sigma_2^2}$ 估计.

- μ_1, μ_2 已知.

$$\frac{\frac{\sum_{i=1}^{m}(X_i-\mu_1)^2}{m\,\sigma_1^2}}{\frac{\sum_{i=1}^{n}(Y_i-\mu_2)^2}{n\,\sigma_2^2}}=\frac{\frac{\sum_{i=1}^{m}(X_i-\mu_1)^2}{m}\Big/\frac{\sum_{i=1}^{n}(Y_i-\mu_2)^2}{n}}{\sigma_1^2/\sigma_2^2}\sim F_{m,n} \quad ⑦$$

- μ_1, μ_2 未知

$$\frac{\frac{\sum_{i=1}^{m}(X_i-\bar{X})^2}{(m-1)\,\sigma_1^2}}{\frac{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}{(n-1)\,\sigma_2^2}}=\frac{\frac{\sum_{i=1}^{m}(X_i-\bar{X})^2}{m-1}\Big/\frac{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}{n-1}}{\sigma_1^2/\sigma_2^2}=\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\sim F_{m-1,n-1} \quad ⑧$$

第九周作业详解.

6.9 设S^2是从正态总体$N(\mu,\sigma^2)$中抽取的容量为16的样本方差，其中μ,σ^2未知. 求$P\{\frac{S^2}{\sigma^2}\le 2.04\}$，$D(S^2)$.

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{15}. \quad 即 \quad \frac{15S^2}{\sigma^2}\sim\chi^2_{15}.$$

$$P\{\frac{S^2}{\sigma^2}\le 2.04\}=P\{\frac{15S^2}{\sigma^2}\le 30.6\}\text{查表} 1-0.01=0.99$$

$$D(\frac{15S^2}{\sigma^2})=2\times 15=30$$

$$\frac{15^2}{\sigma^4}D(S^2)=30\Longrightarrow D(S^2)=\frac{30\sigma^4}{15^2}=\frac{2}{15}\sigma^4$$

利用χ^2分布及χ^2分布的期望是自由度，方差是自由度的2倍等知识求解.

6.11. 设(X_1,X_2,X_3,X_4)是来自正态总体$N(0,2^2)$的样本，$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$

求常数a,b使得统计量$X\sim\chi^2_2$

$$X_1-2X_2\sim N(0,20) \qquad 3X_3-4X_4\sim N(0,100)$$

$$\left(\frac{X_1-2X_2}{\sqrt{20}}\right)^2+\left(\frac{3X_3-4X_4}{\sqrt{100}}\right)^2\sim\chi^2_2.$$

$$\therefore a=\frac{1}{20} \qquad b=\frac{1}{100}$$

若不给定自由度，答案不唯一！

6.14 设总体 $X \sim N(\mu, \sigma^2)$，$(X_1, X_2, \cdots, X_n, X_{n+1})$ 是 X 之 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$，$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$，试确定常数 c，使随机变量 $T = c\,\frac{X_{n+1} - \bar{X}_n}{S_n}$ 服从 t 分布并求其自由度.

分析：　一次，有分子分母基本就是 t 分布啦！.

$$X_{n+1} - \bar{X}_n \sim N\left(0,\ \sigma^2 + \frac{\sigma^2}{n} = \frac{n+1}{n}\sigma^2\right)$$

$$\pm\frac{X_{n+1} - \bar{X}_n}{\sqrt{\frac{n+1}{n}\sigma^2}} \sim N(0,1) \qquad \frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$$

独立.

$$\frac{\pm\dfrac{X_{n+1} - \bar{X}_n}{\cancel{\sigma}\sqrt{\frac{n+1}{n}}}}{\sqrt{\dfrac{\cancel{(n-1)}\,S_n^2}{\cancel{(n-1)}\,\cancel{\sigma^2}}}} \sim t_{n-1}$$

化简得 $\pm\sqrt{\frac{n}{n+1}}\,\frac{X_{n+1} - \bar{X}_n}{S_n} \sim t_{n-1}$

$c = \pm\sqrt{\frac{n}{n+1}}$　　自由度为 $n-1$.

6.17 设 $X \sim N(\mu_1, \sigma^2)$ $Y \sim N(\mu_2, \sigma^2)$

$X_1, \cdots, X_m \swarrow X$ $Y_1, \cdots, Y_n \swarrow Y$

令 $S^2 = \frac{m-1}{m+n-2} S_1^2 + \frac{n-1}{m+n-2} S_2^2$. $\alpha . \beta$ 为两个给定的常数，试求统计量

$$Z = \frac{\alpha(\bar{X}-\mu_1) + \beta(\bar{Y}-\mu_2)}{\sqrt{\frac{\alpha^2}{m} + \frac{\beta^2}{n}}\, S}$$ 的分布.

分析： 一次 有分子分母 应是t分布.

$$\bar{X}-\mu_1 \sim N(0, \frac{\sigma^2}{m}) \qquad \bar{Y}-\mu_2 \sim N(0, \frac{\sigma^2}{n})$$

$$\alpha(\bar{X}-\mu_1) + \beta(\bar{Y}-\mu_2) \sim N(0, (\frac{\alpha^2}{m} + \frac{\beta^2}{n})\sigma^2).$$

$$\frac{\alpha(\bar{X}-\mu_1) + \beta(\bar{Y}-\mu_2)}{\sigma\sqrt{\frac{\alpha^2}{m} + \frac{\beta^2}{n}}} \sim N(0,1)$$ ----分子的构造

$$\frac{(m-1)S_1^2}{\sigma^2} \sim \chi^2_{m-1} \qquad \frac{(n-1)S_2^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\frac{(m-1)S_1^2 + (n-1)S_2^2}{\sigma^2} \sim \chi^2_{m+n-2}.$$ ----分母的构造

$$Z = \frac{\dfrac{\alpha(\bar{X}-\mu_1) + \beta(\bar{Y}-\mu_2)}{\cancel{\sigma}\sqrt{\frac{\alpha^2}{m} + \frac{\beta^2}{n}}}}{\sqrt{\dfrac{(m-1)S_1^2 + (n-1)S_2^2}{(m+n-2)\cancel{\sigma^2}}}} = \frac{\alpha(\bar{X}-\mu_1) + \beta(\bar{Y}-\mu_2)}{\sqrt{\frac{\alpha^2}{m} + \frac{\beta^2}{n}}\, S} \sim t_{m+n-2}$$

7.3. 设总体 X 的 $f(x;\theta)=\begin{cases}\dfrac{2\theta^2}{(\theta^2-1)x^3} & ; 1<x<\theta \\ 0 & ;\text{其他}\end{cases}$

$X_1,\cdots,X_n \swarrow X$ 试求参数 θ 的矩估计.

$$EX=\int_{-\infty}^{+\infty}x\cdot f(x;\theta)dx=\int_1^\theta \frac{2\theta^2}{(\theta^2-1)x^2}dx$$

$$=\frac{2\theta^2}{\theta^2-1}\left(1-\frac{1}{\theta}\right)=\frac{2\theta}{\theta+1}$$

解得 $\theta=\dfrac{EX}{2-EX}$ $\quad\because EX \xrightarrow{P} \bar{X}$

$\therefore \hat{\theta}=\dfrac{\bar{X}}{2-\bar{X}}$ → 此为精确写法.

$\bar{X}=EX=\dfrac{2\theta}{\theta+1}$

$\therefore \hat{\theta}=\dfrac{\bar{X}}{2-\bar{X}}$ → 此为简便写法.

只是为了书写方便.

一定不要误认为 $\bar{X}$ 真的等于 EX （若判断题. $\bar{X}=EX$ ✗ ；$E(\bar{X})=EX$ ✓）

7.12 总体 X 的 $f(x;\theta)=\begin{cases}\dfrac{2}{\theta^2}(\theta-x) & ; 0<x<\theta \\ 0 & ;\text{其他}\end{cases}$

$X_1,\cdots,X_n \swarrow X$ 求 θ 的矩估计量.

$$\bar{X}=EX=\int_{-\infty}^{+\infty}xf(x;\theta)dx=\int_0^\theta \frac{2x}{\theta^2}(\theta-x)\,dx$$

$$=\frac{1}{3}\theta$$

$\therefore \hat{\theta}=3\bar{X}$

一般矩估计都用简便写法计算。

7.6 设总体 $X \sim N(\tan u+5, \sigma_0^2)$ 其中 u 未知，且满足 $-\frac{\pi}{2} < u < \frac{\pi}{2}$，$\sigma_0^2$ 已知

$X_1, \cdots, X_n$ 是 X 试求 u 的最大似然估计量.

$$f(x;u) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(x-\tan u-5)^2}{2\sigma_0^2}}$$

$$L(u) = \prod_{i=1}^{n} f(x_i;u) = \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\tan u-5)^2}{2\sigma_0^2}}$$

一定要把 x 改成 x_i.

这里 x_i 已知，具体一组抽样数值.

u 是未知的.

$$l(u) = \ln L(u) = \ln\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n - \frac{\sum_{i=1}^{n}(x_i-\tan u-5)^2}{2\sigma_0^2}$$

$$\frac{dl(u)}{du} = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\tan u-5)\sec^2 u = 0$$

$$\sum_{i=1}^{n} x_i - n\tan u - 5n = 0.$$

$$\tan u = \frac{\sum_{i=1}^{n} x_i - 5n}{n}$$

解得 $\hat{u} = \arctan\left(\frac{\sum_{i=1}^{n} x_i - 5n}{n}\right)$

即 $\hat{u} = \arctan(\bar{X}-5)$

事实上. $X \sim N(u, \sigma_0^2)$. σ_0^2 已知时 u 的最大似然估计 $\hat{u} = \bar{X}$.

最大似然估计具有不变性.

$$\widehat{\tan u}+5 = \bar{X}.$$

$$\widehat{\tan u} = \bar{X}-5$$

$$\hat{u} = \arctan(\bar{X}-5)$$

7.9 设总体X的分布律为

X	-1	0	1	2
$f(x;\theta)$	θ^3	$3\theta^2(1-\theta)$	$3\theta(1-\theta)^2$	$(1-\theta)^3$

$X_1,\cdots,X_n \swarrow X$. 试求参数θ的矩估计值和最大似然估计值。

其中样本观测值为 $(-1,1,0,2,2,-1,0,-1)$

$$\bar{X}=EX=(-1)\cdot\theta^3+0\cdot 3\theta^2(1-\theta)+1\cdot 3\theta(1-\theta)^2+2(1-\theta)^3$$

$$=\cancel{-\theta^3}+3\theta+\cancel{3\theta^3}\cancel{-6\theta^2}+2\cancel{-2\theta^3}-6\theta\cancel{+6\theta^2}$$

$$=-3\theta+2$$

$$\bar{x}=\frac{-1+1+0+2+2-1+0-1}{8}=\frac{1}{4}$$

解得 $\hat{\theta}=\frac{2-\bar{X}}{3}$ $\qquad \hat{\theta}=\frac{2-\frac{1}{4}}{3}=\frac{7}{12}$

最大似然估计用于离散型本质是具体一次抽样的事件的概率。

不妨设 -1 出现的个数为 k，0 出现的个数为 l，1 出现的个数为 m，则 2 出现的个数为 $n-k-l-m$.

$$L(\theta)=(\theta^3)^k\left[3\theta^2(1-\theta)\right]^l\left[3\theta(1-\theta)^2\right]^m\left[(1-\theta)^3\right]^{n-k-l-m}$$

$$=3^{l+m}\theta^{3k+m+2l}(1-\theta)^{3n-3k-2l-m}$$

$$l(\theta)=\ln L(\theta)=(l+m)\ln 3+(3k+m+2l)\ln\theta+(3n-3k-2l-m)\ln(1-\theta)$$

$$\frac{dl(\theta)}{d\theta}=\frac{3k+m+2l}{\theta}+\frac{3k+2l+m-3n}{1-\theta}=0$$

$$3n\theta=3k+m+2l$$

$$\hat{\theta}=\frac{3k+m+2l}{3n}$$

此题具体 $k=3,\ m=1,\ l=2$
$n=8$

$$\hat{\theta}=\frac{3\times 3+1+2\times 2}{3\times 8}=\frac{7}{12}$$

事实上：用 k,l,m,n 表示 $\bar{x}=\frac{(-1)k+0\cdot l+1\cdot m+2(n-k-l-m)}{n}$

$$=\frac{2n-3k-m-2l}{n}=2-\frac{3k+m+2l}{n}$$

矩估计 $\hat{\theta}=\frac{2-\bar{x}}{3}=\frac{2-\left(2-\frac{3k+m+2l}{n}\right)}{3}=\frac{3k+m+2l}{3n}=\hat{\theta}$ 最大似然估计

6

7.10 设总体 $X \sim U(\theta, 2\theta)$ $x_1, \cdots, x_n$ 是 X. 求 θ 的矩估计及最大似然估计.

最大似然:
$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & ; \theta \le x \le 2\theta \\ 0. & ; \text{其他}. \end{cases}$$

$$L(\theta) = \begin{cases} \frac{1}{\theta^n} & ; \theta \le x_i \le 2\theta \\ 0 & ; \text{~~其他~~} \end{cases}$$

x_i 已知. θ是未知.

由 $\theta \le x_i \Rightarrow \theta \le x_{(1)}$

由 $2\theta \ge x_i \Rightarrow 2\theta \ge x_{(n)}$

$\Downarrow$

$\theta \ge \frac{x_{(n)}}{2}$

$$\therefore L(\theta) = \frac{1}{\theta^n} \quad ; \quad \frac{x_{(n)}}{2} \le \theta \le x_{(1)}$$

$\frac{x_{(n)}}{2}$ θ $x_{(1)}$

$$\hat{\theta} = \frac{x_{(n)}}{2} \qquad \hat{\theta} = \frac{X_{(n)}}{2}$$

矩估计:

$$\bar{X} = EX = \frac{3\theta}{2} \Rightarrow \hat{\theta} = \frac{2\bar{X}}{3}$$

7.13. 一个盒子中装有白球和黑球. 有放回地取出一个容量为 n 的样本, 其中有 k 个白球. 求盒子里黑球数与白球数之比 R 的最大似然估计量.

黑球个数 : 白球个数 $= R : 1$.

$$L(R) = P(\text{有 } k \text{ 个白球}) = \left(\frac{1}{R+1}\right)^k \left(\frac{R}{R+1}\right)^{n-k}$$

$$l(R) = \ln L(R) = k \ln \frac{1}{R+1} + (n-k)[\ln R - \ln(R+1)]$$
$$= -n \ln(R+1) + (n-k) \ln R$$

$$\frac{dl(R)}{dR} = \frac{-n}{R+1} + \frac{n-k}{R} = 0 \qquad \hat{R} = \frac{n-k}{k}$$

黑球个数

白球个数

此解符合我们的直觉!