Ch 9.4 正态总体抽样分布定理

回归前一次课

统计量：次序统计量、及其分布函数、密度函数

Γ - 函 数 $\Gamma(\alpha)=\int_0^{+\infty} x^{\alpha-1}e^{-x}dx$， $\Gamma(1)=1$ 和 $\Gamma(1/2)=\sqrt{\pi}$，以及 $\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$

$\Gamma(\alpha,\lambda)$分布：

$$f(x)=\begin{cases}\dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

若随机变量$X\sim\Gamma(\alpha,\lambda)$, 则有$E(X)=\alpha/\lambda$和$\mathrm{Var}\,(X)=\alpha/\lambda^2$

独立随机变量$X\sim\Gamma(\alpha_1,\lambda)$和$Y\sim\Gamma(\alpha_2,\lambda)$,则$X+Y\sim\Gamma(\alpha_1+\alpha_2,\lambda)$

若随机变量$X\sim N(0,1)$, 则有$X^2\sim\Gamma(1/2,1/2)$

统计三大分布

$\chi^2(n)$分布: $Y = X_1^2 + X_2^2 + \cdots + X_n^2$

随机变量$X \sim \chi^2(n)$, 则$E(X) = n$和$\mathrm{Var}(X) = 2n$;

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

$t(n)$分布: $T = X/\sqrt{Y/n}$；当$n \to \infty$时$t(n)$近似于$N(0,1)$

$F(m,n)$分布：$F = \frac{X/m}{Y/n}$

若随机变量$F \sim F(m,n)$, 则$1/F = F(n,m)$.

分布可加性

- 如果 $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$ 和 $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$，且 X 与 Y 独立，那么 $X \pm Y \sim \mathrm{N}(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$;
- 如果 $X \sim B(n_1, p)$ 和 $Y \sim B(n_2, p)$，且 X 与 Y 独立，那么 $X + Y \sim B(n_1 + n_2, p)$;
- 如果$X \sim P(\lambda_1)$和$Y \sim P(\lambda_2)$,且X与Y独立, 那么$X + Y \sim P(\lambda_1 + \lambda_2)$;
- 如果$X \sim \Gamma(\alpha_1, \lambda)$ 和$Y \sim \Gamma(\alpha_2, \lambda)$，且$X$与$Y$独立，那么$X + Y \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$;
- 如果$X \sim \chi^2(m)$和$Y \sim \chi^2(n)$,且X与Y独立, 那么$X + Y \sim \chi^2(m + n)$.

正态分布的抽样分布定理一

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

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

正态分布的抽样分布定理二

定理： 设$X_1, X_2, \cdots, X_n$是来自总体$N(\mu, \sigma^2)$的样本,其样本均值和修正样本方差分别为

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

则有$\bar{X}$和S^2相互独立, 且

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

正态分布的抽样分布定理三

定理： 设$X_1, X_2, \cdots, X_m$是来自总体$N(\mu, \sigma^2)$的样本,其样本均值和修正样本方差分别为

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

则

$$\frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1)$$

正态分布的抽样分布定理四

定 理： 设$X_1, X_2, \cdots, X_m$和$Y_1, Y_2, \cdots, Y_n$分别来自总体$N(\mu_X, \sigma_X^2)$和$N(\mu_Y, \sigma_Y^2)$的两个独立样本, 其修正样本方差分别为S_X^2和S_Y^2, 则

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

正态分布的抽样分布定理五

定理： 设$X_1, X_2, \cdots, X_m$和$Y_1, Y_2, \cdots, Y_n$分别来自总体$N(\mu_X, \sigma^2)$和$N(\mu_Y, \sigma^2)$的两个独立样本，令其样本均值分别$\bar{X}$和$\bar{Y}$，修正样本方差分别为S_X^2和S_Y^2, 则

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

课堂练习

设$X_1, X_2, \cdots, X_{10}$是总体$N(\mu, 1/4)$ 的样本

i) 若$\mu = 0$, 求$P(\sum_{i=1}^{10} X_i^2 \geq 4)$;

ii) 若μ未知, 求$P\left(\sum_{i=1}^{10} (X_i - \bar{X})^2 \geq 1\right)$.

设$X_1, X_2, \cdots, X_{25}$是总体$N(12, \sigma^2)$的样本

i) 若$\sigma = 2$, 求$P\left(\sum_{i=1}^{25} X_i/25 \geq 12.5\right)$;

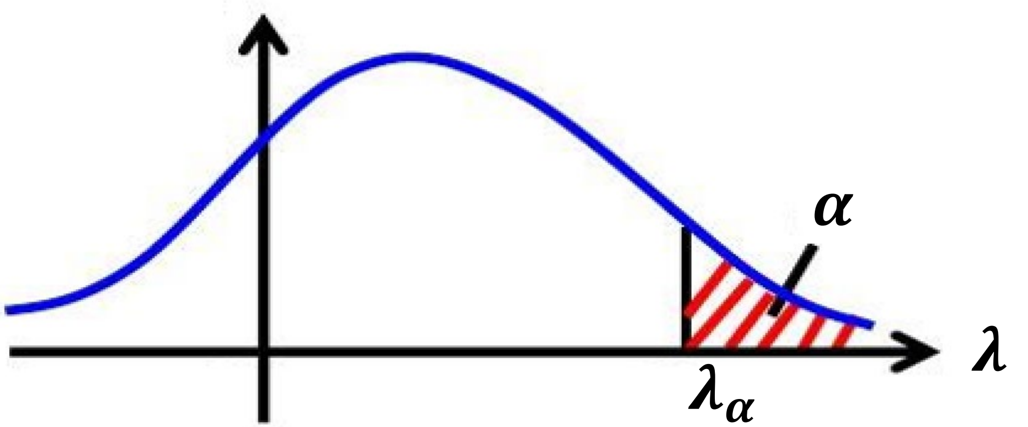
ii) 若 σ 未知但知道修正样本方差为 $S^2 = 5.57$，求 $P\left(\sum_{i=1}^{25} X_i/25 \geq 12.5\right)$.

分位数(点)

对给定$\alpha \in (0,1)$ 和随机变量X, 称满足

$$P(X > \lambda_\alpha) = \alpha$$

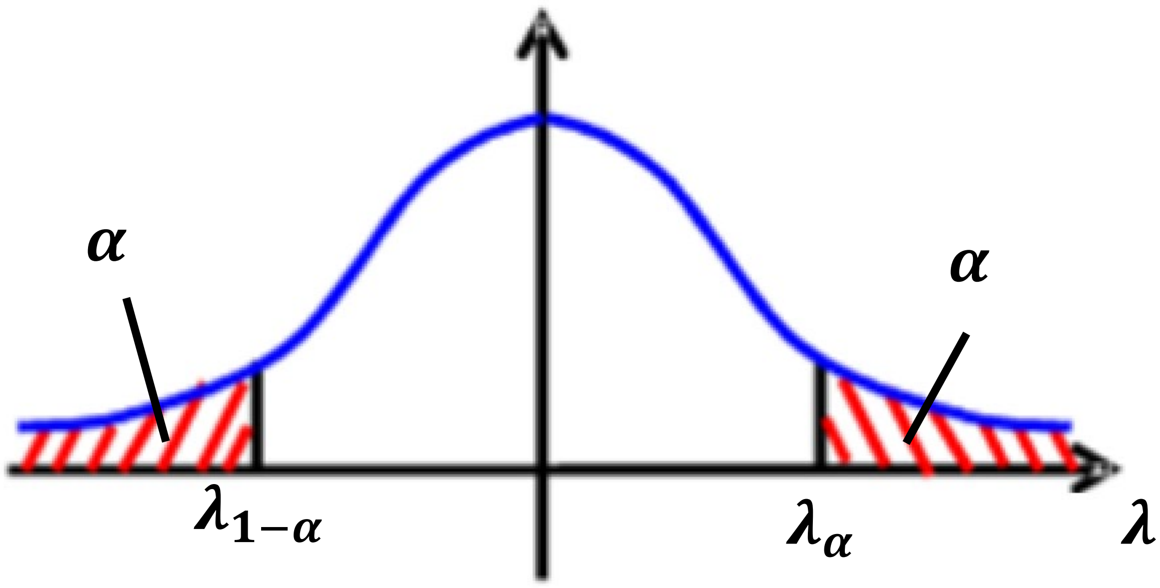
的实数λ_α为上侧α分位数(点)



对称分布的分位数

随机变量X的概率密度函数关于y轴对称, 则有

$$\lambda_{1-\alpha} = -\lambda_{\alpha}$$

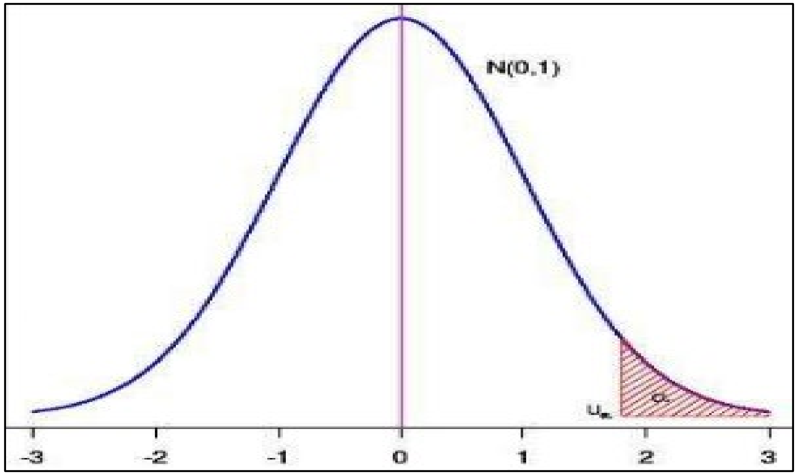


正态分布的分位数

对正态分布$X \sim N(0,1)$, 给定$\alpha \in (0,1)$, 满足

$$P(X > \mu_\alpha) = \int_{\mu_\alpha}^{\infty} f(x)dx = \alpha$$

的点 **μ_α称为正态分布上侧α分位点**



性质：

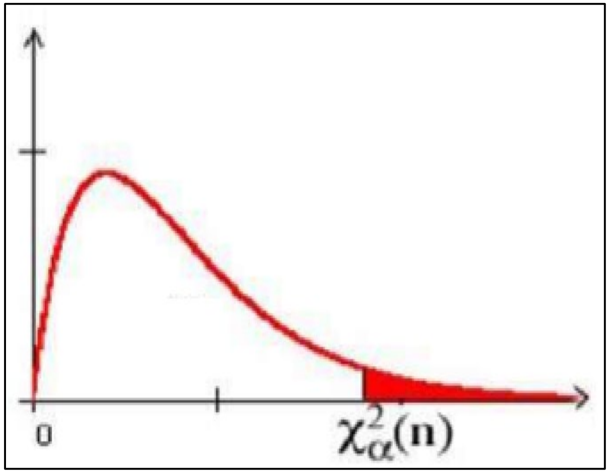
- $\mu_{1-\alpha} = -\mu_\alpha$
- $\Phi(\mu_\alpha) = 1 - \alpha$

χ^2分布的分位数

对χ^2分布$X \sim \chi^2(n)$, 给定$\alpha \in (0,1)$, 满足

$$P\left(X \geq \chi_\alpha^2(n)\right) = \alpha$$

的点$\boldsymbol{\chi_\alpha^2(n)}$**称为**$\boldsymbol{\chi^2(n)}$**分布上侧**$\boldsymbol{\alpha}$**分位点**



t−分布的分位数

对t-分布$X \sim t(n)$, 给定$\alpha \in (0,1)$, 满足

$$P\big(X > t_{\alpha}(n)\big) = \alpha$$

的点**$t_{\alpha}(n)$称为$t(n)$-分布上侧α分位点**

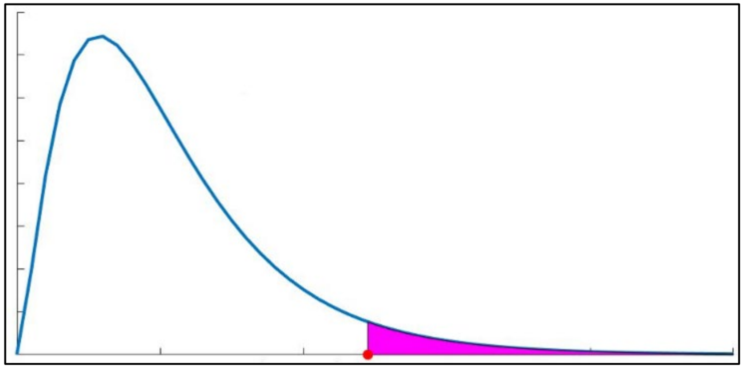
由对称性可知$t_{1-\alpha}(n) = -t_{\alpha}(n)$

F-分布的分位数

对F-分布$X \sim F(m,n)$, 给定$\alpha \in (0,1)$, 满足

$$P[X > F_\alpha(m,n)] = \alpha$$

的点$\boldsymbol{F_\alpha(m,n)}$**称为**$\boldsymbol{F(m,n)}$ **分布上侧**$\boldsymbol{\alpha}$**分位点**



定理：对F-分布的分位点有

$$F_{1-\alpha}(m,n) = \frac{1}{F_\alpha(n,m)}.$$

Ch 10 参数估计

问题

总体X的分布/密度函数为$F(X,\theta)$, θ为未知参数(或未知向量)

现从总体中抽取一样本 $X_1, X_2, \cdots, X_n$

问题：如何依据样本$X_1, X_2, \cdots, X_n$估计参数θ，或θ的函数$g(\theta)$，此类问题称为 **参数估计问题**

研究内容包括: 点估计, 估计量标准, 区间估计

点估计包括：矩估计法、极大似然估计法

矩估计法

总体X的k阶矩: $a_k = E[X^k]$　　样本k阶矩: $A_k = \sum_{i=1}^{n} X_i^k$

用样本矩去估计总体矩求参数θ的方法称为 **矩估计法**

理论基础【辛钦大数定理】 $X_1, X_2, \cdots, X_n$为独立同分布的随机变量, 若$E(X) = \mu$, 则有

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu.$$

推论: 若$E[X^k] = a_k$存在, 则有

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{p} a_k = E[X^k].$$

中心矩估计

总体X的k阶中心矩: $b_k = E\left[\left(X - E(X)\right)^k\right]$

样本k阶中心矩: $B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$

用样本中心矩去估计总体中心矩求参数θ的方法亦称为 **矩估计法**

矩估计方法

总体X的分布函数F包含m个未知参数$\theta_1,\theta_2,\cdots,\theta_m$

- 计算总体X的k阶矩：$a_k=a_k(\theta_1,\theta_2,\cdots,\theta_m)=E[X^k]\ \ k\in[m]$ (a_k一般为$\theta_1,\theta_2,\cdots,\theta_m$的函数)
- 计算样本的k阶矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$
- 令样本矩等于总体矩：

$$A_k=a_k=a_k(\theta_1,\theta_2,\cdots,\theta_m)\ k\in[m]$$

得到m个关于$\theta_1,\theta_2,\cdots,\theta_m$的方程组

- 求解方程组得到估计量$\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m$

例题

设总体X的概率密度函数

$$f(x)=\begin{cases}(\alpha+1)x^{\alpha} & x\in(0,1)\\ 0 & \text{其它}\end{cases}$$

设$X_1,X_2,\cdots,X_n$是来自总体X的样本, 求参数α的矩估计.

例题

设$X_1, X_2, \cdots, X_n$是来自总体X的样本, 以及总体X的密度函数为

$$f(x) = \begin{cases} \dfrac{1}{\theta}\, e^{-(x-\mu)/\theta} & x \geq \mu \\ 0 & \text{其它} \end{cases}$$

其中$\theta > 0$, 求μ和θ的矩估计.

课堂习题

求正态总体$N(\mu,\sigma^2)$的μ,σ^2的矩估计法.

求总体$X \sim U(a,b)$中a,b的矩估计法.