

概率论与数理统计

ruanyl@buaa.edu.cn

统计量，样本分布

$$\mathbf{X} = (X_1, \dots, X_n)^T, \bar{\mathbf{X}} = (\bar{X}_n, \dots, \bar{X}_n)^T, \mathbf{1} = (1, \dots, 1)^T, \mathbf{a} = (a, \dots, a)^T$$

$$(\mathbf{X} - \bar{\mathbf{X}})^T \cdot \mathbf{1} = 0,$$

$$\|\mathbf{X} - \mathbf{a}\|^2 = \|(\mathbf{X} - \bar{\mathbf{X}}) - (\mathbf{a} - \bar{\mathbf{X}})\|^2 = \|\mathbf{X} - \bar{\mathbf{X}}\|^2 + \|\mathbf{a} - \bar{\mathbf{X}}\|^2.$$

$$\sum_{i=1}^n (X_i - a)^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2 + n (\bar{X}_n - a)^2, \quad \forall a.$$

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

大数定律

$$\frac{1}{n} \sum_{i=1}^n X_i^k \approx E(X^k), \quad k = 1, 2, \dots$$

样本均值,

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

令

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

样本方差

$$\frac{S_n^2}{n} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad \frac{S_n^2}{n-1} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

容易验证若 X_1, \dots, X_n 是来自期望为 μ , 方差为 σ^2 的独立同分布样本, 那么

$$E(\bar{X}_n) = \mu, \text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}, E\left(\frac{S_n^2}{n-1}\right) = \sigma^2.$$

若总体为正态分布, 利用统计学基本定理

$$\text{Var}\left(\frac{S_n^2}{n-1}\right) = \text{Var}\left(\frac{\sigma^2}{n-1} \frac{S_n^2}{\sigma^2}\right) = \frac{2\sigma^4}{n-1}.$$

回顾 Gamma 分布

$$X \sim \text{Gamma}(\alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

$$EX = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\beta x} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} = \frac{\alpha}{\beta}.$$

$$\text{Var}(X) = E(X^2) - (EX)^2 = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}.$$

$$\text{Gamma}(1, \beta) = \text{Exp}(\beta), \quad \text{Gamma}(n, \beta) = \underbrace{\text{Exp}(\beta) + \cdots + \text{Exp}(\beta)}_{n \text{ independent copies}},$$

$$\underbrace{\text{Gamma}(\alpha_1, \beta) + \cdots + \text{Gamma}(\alpha_n, \beta)}_{n \text{ independent copies}} = \text{Gamma}(\alpha_1 + \cdots + \alpha_n, \beta)$$

$$X_1, \dots, X_n \sim \text{iid } \mathcal{N}(0, 1), \quad X_1^2 + \cdots + X_n^2 \sim \text{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right) \triangleq \chi_n^2.$$

$$E(\chi_n^2) = n, \quad \text{Var}(\chi_n^2) = 2n.$$

Theorem 3. 若 $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ 为独立同分布随机样本, 那么

(1) 相互独立

$$\bar{X}_n, S_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

(2)

$$\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

(3)

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

$Y \sim \chi_n^2, Z \sim \mathcal{N}(0, 1)$ 相互独立,

$$\frac{Z}{(Y/n)^{1/2}}$$

的分布称为自由度为 n 的 t -分布, 记为 t_n .

若 $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ 为独立同分布随机样本, 那么

$$\frac{(n(n-1))^{1/2} (\bar{X}_n - \mu)}{\left(\sum_{i=1}^n (X_i - \bar{X}_n)^2\right)^{1/2}} = \frac{(\bar{X}_n - \mu) / (\sigma/\sqrt{n})}{\left(\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2 / (n-1)\right)^{1/2}} \sim t_{n-1}$$

$Y \sim \chi_m^2, W \sim \chi_n^2$ 相互独立,

$$\frac{Y/m}{W/n}$$

的分布称为自由度为 (m, n) 的 F -分布, 记为 $F_{m,n}$.

置信区间

待估参数	其他参数	枢轴量 V	V 的分布
μ	σ^2 已知	$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$	$\mathcal{N}(0, 1)$
μ	σ^2 未知	$\frac{\bar{X}_n - \mu}{\hat{\sigma}/\sqrt{n}}$	t_{n-1}
σ^2	μ 已知	$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2$	χ_n^2
σ^2	μ 未知	$\sum_{i=1}^n \left(\frac{X_i - \bar{X}_n}{\sigma} \right)^2$	χ_{n-1}^2

这里

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