注意：文件只列出了部分要点，没有提到的内容不代表不会考察.

6 样本及抽样分布

样本与统计量

主要内容：常用统计量，三大统计分布：χ^2分布、t分布、F分布，正态分布常用统计量的分布

精确定义

定义：称随机变量$X_1, X_2, \cdots, X_n$构成一个（简单）随机样本，如果这些随机变量

1. 相互独立；
2. 服从相同的分布.

它们共同服从的分布称为总体分布；样本个数n称为样本容量.

常用统计量

定义：对样本$X_1, X_2, \cdots, X_n$，称

$$\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

为样本均值.

常用统计量

定义：对样本$X_1, X_2, \cdots, X_n$, 称

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

为样本方差；称

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

为样本标准差.

常用统计量

样本方差的性质：

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2 - n\overline{X}^2\right).$$

χ^2分布

χ^2分布的性质：

1. 若X服从标准正态分布, 则X^2 服从1个自由度的χ^2分布, 即
$$X^2 \sim \chi_1^2.$$
2. 可加性：设$Y_1 \sim \chi_m^2$, $Y_2 \sim \chi_n^2$, 且两者相互独立, 则
$$Y_1 + Y_2 \sim \chi_{m+n}^2.$$

χ^2分布

定理：设$X_1, X_2, \cdots, X_n$相互独立, 都服从标准正态分布, 则

$$\chi^2 := \sum_{i=1}^{n} X_i^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

服从n个自由度的χ^2分布, 即

$$\chi^2 \sim \chi_n^2.$$

t分布

定理：设两个随机变量X, Y相互独立, 并且

$$X \sim N(0,1), \qquad Y \sim \chi_n^2.$$

则

$$T := \frac{X}{\sqrt{Y/n}}$$

服从具有n个自由度的t分布.

F分布

定理：设两个随机变量Y_1, Y_2相互独立, 并且

$$Y_i \sim \chi_{n_i}^2, \qquad i = 1, 2.$$

则

$$F := \frac{Y_1/n_1}{Y_2/n_2} \sim F_{n_1, n_2}.$$

F分布

F分布的性质：

1. 若$F \sim F_{m,n}$, 则

$$\frac{1}{F} \sim F_{n,m}.$$

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

3. 若$X \sim t_n$, 则

$$X^2 \sim F_{1,n}.$$

单个正态总体的统计量的分布

定理 1. 设$X_1, X_2, \cdots, X_n$是取自正态总体 $N(\mu, \sigma^2)$ 的样本. 则$\overline{X}$与S^2相互独立, 且有

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

$$\frac{\sum_{i=1}^{n}\left(X_i-\mu\right)^2}{\sigma^2} \sim \chi_n^2, \qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2.$$

两个正态总体的统计量的分布

定理 2. 设$X_1, X_2, \cdots, X_m$与$Y_1, Y_2, \cdots, Y_n$分别是取自两个相互独立的正态总体

$$N(\mu_1, \sigma^2), \qquad N(\mu_2, \sigma^2)$$

的样本. 则

$$U := \frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim N(0,1),$$

其中$\overline{X}$, $\overline{Y}$ 分别是两个样本各自的均值.

两个正态总体的统计量的分布

定理 3. 设$X_1, X_2, \cdots, X_m$与$Y_1, Y_2, \cdots, Y_n$分别是取自两个相互独立的正态总体

$$N(\mu_1, \sigma^2), \qquad N(\mu_2, \sigma^2)$$

的样本. 则

$$T := \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2}} \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t_{m+n-2},$$

其中$\overline{X}$, $\overline{Y}$, S_1^2, S_2^2分别是两个样本各自的均值及方差.

两个正态总体的统计量的分布

定理 4. 设$X_1, \cdots, X_m$与$Y_1, \cdots, Y_n$分别是取自两个相互独立的正态总体

$$N(\mu_1, \sigma_1^2), \qquad N(\mu_2, \sigma_2^2)$$

的样本. 则

$$F := \frac{\sum_{i=1}^{m} \left(X_i - \mu_1\right)^2 \Big/ \left(m\sigma_1^2\right)}{\sum_{j=1}^{n} \left(Y_j - \mu_2\right)^2 \Big/ \left(n\sigma_2^2\right)} \sim F_{m,n}$$

两个正态总体的统计量的分布

定理 5. 设$X_1, \cdots, X_m$与$Y_1, \cdots, Y_n$分别是取自两个相互独立的正态总体

$$N(\mu_1, \sigma_1^2), \qquad N(\mu_2, \sigma_2^2)$$

的样本. 则

$$F := \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{m-1,n-1}.$$