1 数理统计的基本概念

1.1 总体，样本，统计量

总体: 研究对象全体，可以用一个随机变量 X 及其分布 F 表示

个体: 总体中的每一个成员

抽样: 按照一定法则，从总体中抽取部分个体观测的过程，常用有不放回抽样和放回抽样，在样本容量相对总体过小时这两个方法可以视为等同

样本: 抽样选中的部分个体，用 n 个随机变量来表示

样本容量: 样本中个体的数量

样本观测值: 由样本中的 n 个随机变量观察得到的 n 个数据

简单随机抽样满足:1) 代表性: 总体中每个个体都有相同的被抽到的概率，则样本中 n 个随机变量和总体 X 有相同的分布

2）独立性:n 个随机变量相互独立

简单随机样本: 由简单随机抽样得到的样本

统计量: 设 $X_1, X_2, \cdots, X_n$ 是总体 X 的一个样本，若 $T = T(X_1, X_2, \cdots, X_n)$ 是样本的函数, 且不含任意未知参数，则称 T 是一个统计量

常见统计量: 样本均值:

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

样本方差:

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

样本标准差:

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

样本 k 阶原点矩:

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

极大次序统计量:

$$X_{(n)} = max\{X_1, X_2, \cdots, X_n\}$$

极小次序统计量:

$$X_{(1)} = min\{X_1, X_2, \cdots, X_n\}$$

1.2 常用统计量的分布

1.2.1 标准正态

1.2.2 χ^2 分布

1.χ^2 分布: 设 $X_1, X_2, \cdots, X_n$ 相互独立且服从标准正态分布，则

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

服从自由度为 n 的 χ^2 分布，记为 $\chi^2(n)$，其密度函数为:

$$f(x)=\begin{cases}\frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n}{2})}e^{-\frac{x}{2}}x^{\frac{n}{2}-1} & x>0,\\ 0 & x\le 0\end{cases}$$

2. 性质 1）可加性: 若 X∼ $\chi^2(n)$,Y∼ $\chi^2(m)$, 则 $X+Y\sim\chi^2(n+m)$

2)$E[\chi^2(n)]=n, D[\chi^2(n)]=2n$

1.2.3 t 分布

1. 设 X∼N（0，1），Y∼ $\chi^2(n)$, 且 X 和 Y 相互独立，则

$$t=\frac{X}{\sqrt{Y/n}}$$

服从自由度为 n 的 t 分布，记为 t（n），其密度函数为

$$f(X)=\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}(1+\frac{x^2}{n}^{-\frac{n+1}{2}})$$

当 n 充分大时，t 分布可以看作标准正态

1.2.4 F 分布

1.$X\sim\chi^2(n),Y\sim\chi^2(m)$, 且 X，Y 相互独立，则

$$F=\frac{X/n}{Y/m}$$

服从自由度为 n，m 的 F 分布，记为 F(n,m)

2. 性质: 1)$X\sim F(n,m)$, 则 $\frac{1}{X}\sim F(m,n)$

2) 若 $t\sim t(n)$, 则 $t^2\sim F(1,n)$