统计量

一. 统计量是什么

统计量

定义 2.1. 由样本算出的量是统计量 (Statistic), 或, 统计量 是样本的函数.

统计量是样本的transformation.

对这一定义我们作如下几点说明:

(1) 统计量只与样本有关, 不能与未知参数有关.
例如 $X \sim N\left(a, \sigma^2\right), X_1, \cdots, X_n$ 是从总体 X 中抽取的i.i.d.样本,
则 $\sum_{i=1}^n X_i$ 和 $\sum_{i=1}^n X_i^2$ 都是统计量,
当 a 和 σ^2 皆为未知参数时, $\sum_{i=1}^n\left(X_i-a\right)$ 和 $\sum_{i=1}^n X_i^2 / \sigma^2$ **都不是统计量**.

（2）由于样本具有两重性, 即样本既可以看成具体的数, 又可以看成随机变量;
统计量 是样本的函数, 因此统计量 也具有两重性.
正因为统计量 可视为随机变量(或随机向量), 因此才有概率分布可言, 这是我们利用统计量进行统计推断的依据.

二、若干常用的统计量

1. 样本均值

设 $X_1, \cdots, X_n$ 是从某总体 X 中抽取的样本,则称

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

为样本均值(Sample mean). 它分别反映了总体数学期望的信息.
就是单纯的n分之一.

2. 样本方差

设 $X_1, \cdots, X_n$ 是从某总体 X 中抽取的样本,则称

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

为样本方差 (Sample variance).
它分别反映了总体方差的信息, 而 S_n 反映了总体标准差的信息.
是n-1分之一,不要自己添加一个维度.

3. 样本矩

设 $X_1, \cdots, X_n$ 为从总体 F 中抽取的样本, 则称

$$a_{n,k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \quad k = 1, 2, \cdots$$

为样本 k 阶原点矩. 特别 $k = 1$ 时, $a_{n,1} = \bar{X}$, 即样本均值.

称

$$m_{n,k} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^k, \quad k = 2, 3, \cdots$$

为样本 k 阶中心矩(中心=均值). 特别 $k = 2$ 时, $m_{n,2} = S_n^2$, 即样本方差. 样本的原点矩和中心矩统称为样本矩 (Sample moments).

4. 二维随机向量的样本矩

设 $(X_1, Y_1), \cdots, (X_n, Y_n)$ 为从二维总体 $F(x, y)$ 中抽取的样本, 则

\begin{aligned} & \bar{X}=\frac{1}{n} \sum_{i=1}^{n} X_{i}, \quad S_{X}^{2}=\frac{1}{n-1

m{1 / 2}= \begin{cases}X{\left(\frac{n+1}{2}\right)} & \text { 当 } \mathrm{n} \text { 为奇数 } \tag{2.1}\ \frac{1}{2}\left[X{(n / 2)}+X{(n / 2+1)}\right] & \text { 当 } \mathrm{n} \text { 为偶数 }\end{cases}

样本中位数 (Sample median)反映总体中位数的信息. 当总体分布**关于某点对称时**, 对称中心既

\begin{equation*}

\hat{V}=S_{n} / \bar{X} \tag{2.2}

\end{equation*}

为样本变异系数 (Sample coefficient of variation). 它反映了总体变异系数(Population coeffic

\begin{equation*}

\hat{\beta}{1}=\frac{m{n, 3}}{m{n, 2}^{3 / 2}}=\sqrt{n} \sum{i=1}^{n}\left(X{i}-\bar{X}\right)^{3} /\left(\sum{i=1}^{n}\left(X_{i}-\bar{X}\right)^{2}\right)^{3 / 2} \tag{2.3}

\end{equation*}

为样本偏度系数 (Sample skewness). 它反映了总体偏度系数的信息, 总体偏度系数(Population ske

\begin{equation*}

\hat{\beta}{2}=\frac{m{n, 4}}{m{n, 2}^{2}}-3=n \sum{i=1}^{n}\left(X{i}-\bar{X}\right)^{4} /\left(\sum{i=1}^{n}\left(X_{i}-\bar{X}\right)^{2}\right)^{2}-3 \tag{2.4}

\end{equation*}

为样本峰度系数($Samplekurtosis$). 它反映了总体峰度系数β_k的信息. 总体峰度系数($Populationkurto$