经验分布

经验分布函数

定义 2.2. 设 $X_1,\cdots,X_n$ 为自总体 $F(X)$ 中抽取的i.i.d.样本, 将其按大小排列为 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$, 对任意实数 x, 称下列函数

$$F_n(x)=\begin{cases}0 & x<X_{(1)} & nbsp; \\ \frac{k}{n} & X_{(k)}\le x<X_{(k+1)},k=1,2,\cdots,n-1 \\ 1 & X_{(n)}\le x\end{cases}\tag{2.5}$$

为经验分布函数 (Empirical distibution function).

易见经验分布函数是单调非降右连续函数, 具有分布函数的基本性质.
它在 $x=X_{(k)},k=1,2,3,\cdots,n$ 处有间断, 它是在每个间断点跳跃的幅度为 $1/n$ 的阶梯函数. 若记示性函数

$I_{[A]}(x)=\begin{cases}1 & 当\ x\in A \\ 0 & 其他\end{cases}$ $$则$F_n(x)$可表为(== 要记住这个表述 ==)是朝小的地方找

\begin{equation*}

F{n}(x)=\frac{1}{n} \sum{i=1}^{n} I{\left[X{i} \leq x\right]} \tag{2.6}

\end{equation*}

由定义可知$F_n(x)$是仅依赖于样本$X_1,X_2,\cdots,X_n$的函数, 因此 == 它是统计量 ==. 它可能取值为$

P\left(F{n}(x)=k / n\right)=P\left(\sum{i=1}^{n} Y_{i}=k\right)=\binom{n}{k}[F(x)]^{k}[1-F(x)]^{n-k}

利用二项分布的性质可知$F_n(x)$具有下列 ∗∗大样本性质∗∗ : （1）由中心极限定理, 则当$n\to\infty$时

\frac{\sqrt{n}\left(F_{n}(x)-F(x)\right)}{\sqrt{F(x)(1-F(x))}} \xrightarrow{\mathscr{L}} N(0,1)

（2）由Benoulli大数定律, 则在$n\to\infty$时有

F_{n}(x) \xrightarrow{P} F(x)

(3)由Borel强大数定律, 则在$n\to\infty$时,

P\left(\lim {n \rightarrow \infty} F{n}(x)=F(x)\right)=1

(4)更进一步, 有下列Glivenko − CantelliTheorem(1933) :> [!NOTE]GCTheorem > 定理2.1. 设$F(x)$

> P\left(\lim {n \rightarrow \infty} D{n}=0\right)=1

> $\$\$$