Order statistics

Monday, May 27, 2024 13:00

理解 order statistics

思路: order stats defn 中有 $X_{(j)} < x$ 看上去和 indicator 有一些联系

$\Rightarrow$ 转化为 Bern | Bin

Continuous case:

random sampling $X_1, X_2 \dots X_k \dots X_n$ n R.V.s.

↳ For X_k, a fixed value x.

$$I\{X_k < x\} = \begin{cases} 1 & P(X_k < x) = F_X(x) \\ 0 & 1 - F_X(x) \end{cases}$$

↳ For all R.V.s. count # of trials where $X_k < x$ (i.e., # of successes)

∴ define Y (R.V.) as

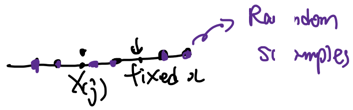
∴ $P(Y = y) = \binom{n}{y} [F_X(x)]^y (1 - F_X(x))^{n-y}$

求 order stat. 的 CDF 可转化为 Bin.

e.g. $F_{X_{(j)}}(x) \overset{\text{defn}}{=} P(X_{(j)} \le x)$ [the j^{th} largest X is smaller / equals to x]

$= P(Y \geq j)$ $\overset{\text{equal.}}{\equiv}$ [at least j X_i (R.U.s) are smaller than x]

$$= \sum_{k=j}^{n} \binom{n}{k} [F_X(x)]^k [(1-F_X(x))]^{n-k}$$



order stat PDF is then:

$$f_{X_{(j)}}(x) = \frac{d}{dx} \sum_{k=j}^{n} \binom{n}{k} \underbrace{F_X(x)^k}_{A} \underbrace{(1-F_X(x))^{n-k}}_{B}$$

$(A \cdot B)' = A'B + B'A$

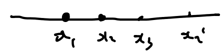
$$= \sum_{k=j}^{n} \binom{n}{k} \Big[k (F_X(x))^{k-1} f_X(x) \cdot [(1-F_X(x))]^{n-k}$$

$$- F_X(x)^k (n-k)(1-F_X(x))^{n-k-1} \cdot f(x) \Big]$$

化简得：（不想推导了!!） see casella P229

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x) [F_X(x)]^{j-1} [1-F_X(x)]^{n-j}$$

Discrete case:

X R.V. 有如下取值 

$P(X=x_i)=p_i$

现在有一个Random sample $X_1, X_2 .. X_n$.

For X_k. $x_k < x_i$ (Fixed x)

$$I\{X_k < x_i\} = \begin{cases} 1 & P(X_k \le x_i) = p_1+p_2+\cdots p_i \quad \text{记作 } P_i \\ 0 & 1-(p_1+p_2+\cdots p_i) \quad 1-P_i \end{cases}$$

$$\therefore P(X_{(j)} \le x_i) = P(Y \ge j) = \sum_{k=j}^{n} \binom{n}{k} P_i^{\,k} (1-P_i)^{n-k}$$

$$\therefore P(X_{(j)} = x_i) = P(X_{(j)} \le x_i) - P(X_{(j)} \le x_{i-1})$$

$$= \sum_{k=j}^{n} \binom{n}{k} \left[P_i^{\,k} (1-P_i)^{n-k} - P_i^{\,k} (1-P_{i-1})^{n-k} \right]$$