

Theorem 5.4.4 Let $X_{(1)}, \dots, X_{(n)}$ denote the order statistics of a random sample, X_1, \dots, X_n , from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the pdf of $X_{(j)}$ is

$$(5.4.4) \quad f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j}.$$

Proof: We first find the cdf of $X_{(j)}$ and then differentiate it to obtain the pdf. As in Theorem 5.4.3, let Y be a random variable that counts the number of X_1, \dots, X_n less than or equal to x . Then, defining a “success” as the event $\{X_j \leq x\}$, we see that $Y \sim \text{binomial}(n, F_X(x))$. (Note that we can write $P_i = F_X(x_i)$ in Theorem 5.4.3. Also, although X_1, \dots, X_n are continuous random variables, the counting variable Y is discrete.) Thus,

$$F_{X_{(j)}}(x) = P(Y \geq j) = \sum_{k=j}^n \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k},$$

and the pdf of $X_{(j)}$ is

$$\begin{aligned} f_{X_{(j)}}(x) &= \frac{d}{dx} F_{X_{(j)}}(x) \\ &= \sum_{k=j}^n \binom{n}{k} \left(k [F_X(x)]^{k-1} [1 - F_X(x)]^{n-k} f_X(x) \right. \\ &\quad \left. - (n-k) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \right) \quad (\text{chain rule}) \\ &= \binom{n}{j} j f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j} \\ &\quad + \sum_{k=j+1}^n \binom{n}{k} k [F_X(x)]^{k-1} [1 - F_X(x)]^{n-k} f_X(x) \\ &\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \quad \begin{pmatrix} k = n \text{ term} \\ \text{is 0} \end{pmatrix} \\ &= \frac{n!}{(j-1)!(n-j)!} f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j} \\ (5.4.5) \quad &+ \sum_{k=j}^{n-1} \binom{n}{k+1} (k+1) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \quad \begin{pmatrix} \text{change} \\ \text{dummy} \\ \text{variable} \end{pmatrix} \\ &\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x). \end{aligned}$$

Noting that

$$(5.4.6) \quad \binom{n}{k+1} (k+1) = \frac{n!}{k!(n-k-1)!} = \binom{n}{k} (n-k),$$

we see that the last two sums in (5.4.5) cancel. Thus, the pdf $f_{X_{(j)}}(x)$ is given by the expression in (5.4.4). \square

Example 5.4.5 (Uniform order statistic pdf) Let X_1, \dots, X_n be iid uniform(0, 1), so $f_X(x) = 1$ for $x \in (0, 1)$ and $F_X(x) = x$ for $x \in (0, 1)$. Using (5.4.4), we see that the pdf of the j th order statistic is

$$\begin{aligned} f_{X_{(j)}}(x) &= \frac{n!}{(j-1)!(n-j)!} x^{j-1}(1-x)^{n-j} \quad \text{for } x \in (0, 1) \\ &= \frac{\Gamma(n+1)}{\Gamma(j)\Gamma(n-j+1)} x^{j-1}(1-x)^{(n-j+1)-1}. \end{aligned}$$

Thus, the j th order statistic from a uniform(0, 1) sample has a beta($j, n - j + 1$) distribution. From this we can deduce that

$$EX_{(j)} = \frac{j}{n+1} \quad \text{and} \quad \text{Var } X_{(j)} = \frac{j(n-j+1)}{(n+1)^2(n+2)}. \quad \|$$

The joint distribution of two or more order statistics can be used to derive the distribution of some of the statistics mentioned at the beginning of this section. The joint pdf of any two order statistics is given in the following theorem, whose proof is left to Exercise 5.26.

Theorem 5.4.6 Let $X_{(1)}, \dots, X_{(n)}$ denote the order statistics of a random sample, X_1, \dots, X_n , from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$. Then the joint pdf of $X_{(i)}$ and $X_{(j)}$, $1 \leq i < j \leq n$, is

$$(5.4.7) \quad f_{X_{(i)}, X_{(j)}}(u, v) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} \times [F_X(v) - F_X(u)]^{j-1-i} [1 - F_X(v)]^{n-j}$$

for $-\infty < u < v < \infty$.

The joint pdf of three or more order statistics could be derived using similar but even more involved arguments. Perhaps the other most useful pdf is $f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n)$, the joint pdf of all the order statistics, which is given by

$$f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = \begin{cases} n! f_X(x_1) \cdots f_X(x_n) & -\infty < x_1 < \cdots < x_n < \infty \\ 0 & \text{otherwise.} \end{cases}$$

The $n!$ naturally comes into this formula because, for any set of values x_1, \dots, x_n , there are $n!$ equally likely assignments for these values to X_1, \dots, X_n that all yield the same values for the order statistics. This joint pdf and the techniques from Chapter 4 can be used to derive marginal and conditional distributions and distributions of other functions of the order statistics. (See Exercises 5.27 and 5.28.)

We now use the joint pdf (5.4.7) to derive the distribution of some of the functions mentioned at the beginning of this section.