Chapter 5
Order Statistics

Dr. Mohamed Rhaima

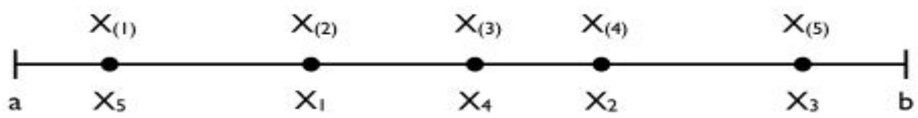
Department of Statistics and Operations Research

November, 25, 2019

Order Statistics

Let X_1, X_2, X_3, X_4, X_5 be iid random variables with a distribution F with a range of (a, b). We can relabel these X's such that their labels correspond to arranging them in increasing order so that

$$X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq X_{(4)} \leq X_{(5)}$$



In the case where the distribution F is continuous we can make the stronger statement that

$$X_{(1)} < X_{(2)} < X_{(3)} < X_{(4)} < X_{(5)}$$

Since $P(X_i = X_j) = 0$ for all $i \neq j$ for continuous random variables.

Order Statistics, cont.

For $X_1, X_2, \ldots, X_n$ iid random variables X_k is the kth smallest X, usually called the kth order statistic.

$X_{(1)}$ is therefore the smallest X and

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

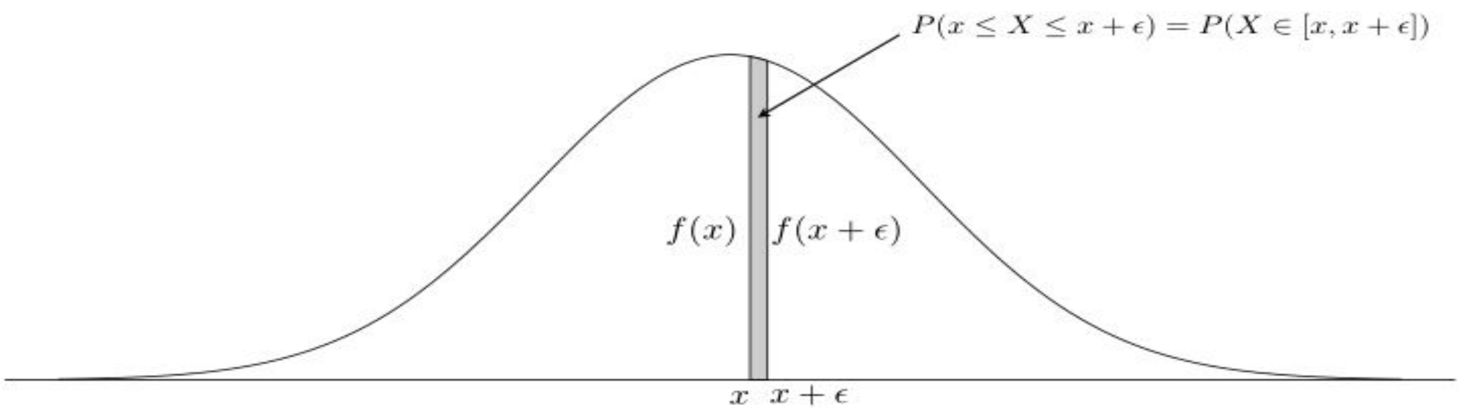
Similarly, $X_{(n)}$ is the largest X and

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

Notation Detour

For a continuous random variable we can see that

$$f(x)\epsilon \approx P(x \leq X \leq x+\epsilon) = P(X \in [x, x+\epsilon])$$
$$\lim_{\epsilon \to 0} f(x)\epsilon = \lim_{\epsilon \to 0} P(X \in [x, x+\epsilon])$$
$$f(x) = \lim_{\epsilon \to 0} P(X \in [x, x+\epsilon])/\epsilon$$



Density of the maximum

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the maximum is

$$
\begin{aligned}
P(X_{(n)} \in [x, x+\epsilon]) &= P(\text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and all others} < x) \\
&= \sum_{i=1}^{n} P(X_i \in [x, x+\epsilon] \text{ and all others} < x) \\
&= nP(X_1 \in [x, x+\epsilon] \text{ and all others} < x) \\
&= nP(X_1 \in [x, x+\epsilon])P(\text{all others} < x) \\
&= nP(X_1 \in [x, x+\epsilon])P(X_2 < x) \cdots P(X_n < x) \\
&= nf(x)\epsilon F(x)^{n-1}
\end{aligned}
$$

$$f_{(n)}(x) = nf(x)F(x)^{n-1}$$

Density of the minimum

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the minimum is

$$
\begin{aligned}
P(X_{(1)} \in [x, x+\epsilon]) &= P(\text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and all others} > x) \\
&= \sum_{i=1}^{n} P(X_i \in [x, x+\epsilon] \text{ and all others} > x) \\
&= nP(X_1 \in [x, x+\epsilon] \text{ and all others} > x) \\
&= nP(X_1 \in [x, x+\epsilon])P(\text{all others} > x) \\
&= nP(X_1 \in [x, x+\epsilon])P(X_2 > x) \cdots P(X_n > x) \\
&= nf(x)\epsilon(1 - F(x))^{n-1}
\end{aligned}
$$

$$f_{(1)}(x) = nf(x)(1 - F(x))^{n-1}$$

Density of the kth Order Statistic

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the kth order statistic is

$$
\begin{aligned}
P(X_{(k)} \in [x, x+\epsilon]) &= P(\text{one of the } X\text{'s} \in [x, x+\epsilon] \text{ and exactly } k-1 \text{ of the others} < x) \\
&= \sum_{i=1}^{n} P(X_i \in [x, x+\epsilon] \text{ and exactly } k-1 \text{ of the others} < x) \\
&= nP(X_1 \in [x, x+\epsilon] \text{ and exactly } k-1 \text{ of the others} < x) \\
&= nP(X_1 \in [x, x+\epsilon])P(\text{exactly } k-1 \text{ of the others} < x) \\
&= nP(X_1 \in [x, x+\epsilon]) \left(\binom{n-1}{k-1} P(X < x)^{k-1} P(X > x)^{n-k} \right)
\end{aligned}
$$

$$
f_{(k)}(x) = nf(x)\binom{n-1}{k-1} F(x)^{k-1}(1 - F(x))^{n-k}
$$

Cumulative Distribution of the min and max

For $X_1, X_2, \ldots, X_n$ iid continuous random variables with pdf f and cdf F the density of the kth order statistic is

$$\begin{aligned} F_{(1)}(x) &= P(X_{(1)} < x) = 1 - P(X_{(1)} > x) \\ &= 1 - P(X_1 > x, \ldots, X_n > x) = 1 - P(X_1 > x) \cdots P(X_n > x) \\ &= 1 - (1 - F(x))^n \end{aligned}$$

$$\begin{aligned} F_{(n)}(x) &= P(X_{(n)} < x) = 1 - P(X_{(n)} > x) \\ &= P(X_1 < x, \ldots, X_n < x) = P(X_1 < x) \cdots P(X_n < x) \\ &= F(x)^n \end{aligned}$$

$$\begin{aligned} f_{(1)}(x) &= \frac{d}{dx}(1 - F(x))^n = n(1 - F(x))^{n-1}\frac{dF(x)}{dx} = nf(x)(1 - F(x))^{n-1} \\ f_{(n)}(x) &= \frac{d}{dx}F(x)^n = nF(x)^{n-1}\frac{dF(x)}{dx} = nf(x)F(x)^{n-1} \end{aligned}$$

Theorem

Let $X_1, \ldots, X_n$ be independent random variables with a common continuous distribution. Let $X_{(1)}, \ldots, X_{(n)}$ be their order statistics. For each x, let $N_n(x)$ be the number of i such that $X_i \leq x$. Then $N_n(x)$ is a binomial random variable with parameters n and $F(x)$. Furthermore,

$$\mathbb{P}(X_{(j)} \leq x) = \mathbb{P}(N_n(x) \geq j).$$

This result can be stated even more explicitly in terms of the binomial probabilities. In this form it says that if $\mathbb{P}(X_i \leq x) = F(x)$, then

$$\mathbb{P}(X_{(j)} \leq x) = \sum_{k=j}^{n} \binom{n}{k} F(x)^k \left(1 - F(x)\right)^{n-k}$$

Order Statistic of Standard Uniforms

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Unif}(0,1)$ then the density of $X_{(n)}$ is given by

$$
\begin{aligned}
f_{(k)}(x) &= nf(x)\binom{n-1}{k-1}F(x)^{k-1}(1-F(x))^{n-k} \\
&= \begin{cases} n\binom{n-1}{k-1}x^{k-1}(1-x)^{n-k} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

This is an example of the Beta distribution where $r = k$ and $s = n - k + 1$.

$$X_{(k)} \sim \text{Beta}(k, n-k+1)$$

Beta Distribution

The Beta distribution is a continuous distribution defined on the range $(0, 1)$ where the density is given by

$$f(x) = \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}$$

where $B(r, s)$ is called the Beta function and it is a normalizing constant which ensures the density integrates to 1.

$$\begin{aligned}
1 &= \int_0^1 f(x)dx \\
1 &= \int_0^1 \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}dx \\
1 &= \frac{1}{B(r,s)} \int_0^1 x^{r-1}(1-x)^{s-1}dx \\
B(r,s) &= \int_0^1 x^{r-1}(1-x)^{s-1}dx
\end{aligned}$$

Beta Function

The connection between the Beta distribution and the kth order statistic of n standard Uniform random variables allows us to simplify the Beta function.

$$
\begin{aligned}
B(r,s) &= \int_0^1 x^{r-1}(1-x)^{s-1}dx \\
B(k, n-k+1) &= \frac{1}{n\binom{n-1}{k-1}} \\
&= \frac{(k-1)!(n-1-k+1)!}{n(n-1)!} \\
&= \frac{(r-1)!(n-k)!}{n!} \\
&= \frac{(r-1)!(s-1)!}{(r+s-1)!} = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}
\end{aligned}
$$

Beta Function - Expectation

Let $X \sim \text{Beta}(r, s)$ then

$$
\begin{aligned}
E(X) &= \int_0^1 x \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1} dx \\
&= \frac{1}{B(r,s)} \int_0^1 1x^{(r+1)-1}(1-x)^{s-1} dx \\
&= \frac{B(r+1,s)}{B(r,s)} \\
&= \frac{r!(s-1)!}{(r+s)!} \frac{(r+s-1)!}{(r-1)!(s-1)!} \\
&= \frac{r!}{(r-1)!} \frac{(r+s-1)!}{(r+s)!} \\
&= \frac{r}{r+s}
\end{aligned}
$$

Beta Function - Variance

Let $X \sim \text{Beta}(r, s)$ then

$$\begin{aligned} E(X^2) &= \int_0^1 x^2 \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1} dx \\ &= \frac{B(r+2,s)}{B(r,s)} = \frac{(r+1)!(s-1)!}{(r+s+1)!} \frac{(r+s-1)!}{(r-1)!(s-1)!} \\ &= \frac{(r+1)r}{(r+s+1)(r+s)} \end{aligned}$$

$$\begin{aligned} Var(X) &= E(X^2) - E(X)^2 \\ &= \frac{(r+1)r}{(r+s+1)(r+s)} - \frac{r^2}{(r+s)^2} \\ &= \frac{(r+1)r(r+s) - r^2(r+s+1)}{(r+s+1)(r+s)^2} \\ &= \frac{rs}{(r+s+1)(r+s)^2} \end{aligned}$$

Beta Distribution - Summary

If $X \sim \text{Beta}(r, s)$ then

$$f(x) = \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1}$$
$$F(x) = \int_0^x \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1} dx = \frac{B_x(r,s)}{B(r,s)}$$

$$B(r,s) = \int_0^1 x^{r-1}(1-x)^{s-1} dx = \frac{(r-1)!(s-1)!}{(r+s-1)!} = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}$$
$$B_x(r,s) = \int_0^x x^{r-1}(1-x)^{s-1} dx$$

$$E(X) = \frac{r}{r+s}$$
$$Var(X) = \frac{rs}{(r+s)^2(r+s+1)}$$

Minimum of Exponentials

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$, we previously derived a more general result where the X's were not identically distributed and showed that $\min(X_1, \ldots, X_n) \sim \text{Exp}(\lambda_1 + \cdots + \lambda_n) = \text{Exp}(n\lambda)$ in this more restricted case.

Lets confirm that result using our new more general methods

$$\begin{aligned} f_{(1)}(x) &= nf(x)(1 - F(x))^{n-1} \\ &= n\left(\lambda e^{-\lambda x}\right)\left(1 - [1 - e^{-\lambda x}]\right)^{n-1} \\ &= n\lambda e^{-\lambda x}\left(e^{-\lambda x}]\right)^{n-1} \\ &= n\lambda\left(e^{-\lambda x}]\right)^{n} \\ &= n\lambda e^{-n\lambda x} \end{aligned}$$

Which is the density for $\text{Exp}(n\lambda)$.

Maximum of Exponentials

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ then the density of $X_{(n)}$ is given by

$$\begin{aligned} f_{(n)}(x) &= nf(x)F(x)^{n-1} \\ &= n\left(\lambda e^{-\lambda x}\right)\left(1 - e^{-\lambda x}\right)^{n-1} \end{aligned}$$

Which we can't do much with, instead we can try the cdf of the maximum.

Maximum of Exponentials, cont.

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ then the cdf of $X_{(n)}$ is given by

$$\begin{aligned} F_{(n)}(x) &= F(x)^n \\ &= \left(1 - e^{-\lambda x}\right)^n \\ &= \left(1 - \frac{ne^{-\lambda x}}{n}\right)^n \\ F_{(n)}(x) &\approx \exp(-ne^{-\lambda x}) \end{aligned}$$

$$\lim_{n\to\infty} F_{(n)}(x) = \lim_{n\to\infty} \exp(-ne^{-\lambda x}) = 0$$

This result is not unique to the exponential distribution...