Lecture 15: Order Statistics

Statistics 104

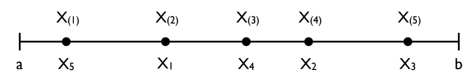
Colin Rundel

March 14, 2012

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

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



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

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

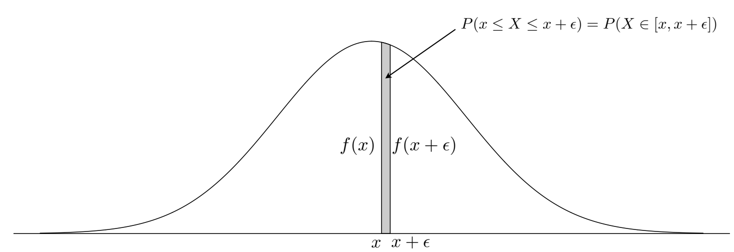
Similarly, $X_{(n)}$ is the largest X and

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

Notation Detour

For a continuous random variable we can see that

$$\begin{aligned} f(x)\epsilon &\approx P(x \le X \le x+\epsilon) = P(X \in [x, x+\epsilon]) \\ \lim_{\epsilon \to 0} f(x)\epsilon &= \lim_{\epsilon \to 0} P(X \in [x, x+\epsilon]) \\ f(x) &= \lim_{\epsilon \to 0} P(X \in [x, x+\epsilon])/\epsilon \end{aligned}$$



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

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

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

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

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

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

$$f_{(1)}(x) = \frac{d}{dx}(1 - F(x))^n = n(1 - F(x))^{n-1}\frac{dF(x)}{dx} = nf(x)(1 - F(x))^{n-1}$$

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

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

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

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.

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

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.

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

Beta Function - Expectation

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

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

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

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

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

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

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...

Limit Distributions of Maxima and Minima

Previous we have shown that

$$\begin{aligned} F_{(1)}(x) &= P(X_{(1)} < x) = 1 - (1 - F(x))^n \\ F_{(n)}(x) &= P(X_{(n)} < x) = F(x)^n \end{aligned}$$

When n tends to infinity we get

$$\lim_{n\to\infty} F_{(1)}(x) = \lim_{n\to\infty} 1 - (1 - F(x))^n = \begin{cases} 0 & \text{if } F(x) = 0 \\ 1 & \text{if } F(x) > 0 \end{cases}$$

$$\lim_{n\to\infty} F_{(n)}(x) = \lim_{n\to\infty} F(x)^n = \begin{cases} 1 & \text{if } F(x) = 1 \\ 0 & \text{if } F(x) < 1 \end{cases}$$

Limit Distributions of Maxima and Minima, cont.

These results show that the limit distributions are degenerate as they only take values of 0 or 1. To avoid the degeneracy we would like to use a simple transform that such that the limit distributions are not degenerate.

Let's consider simple linear transformations

$$\begin{aligned} \lim_{n\to\infty} F_{(n)}(a_n + b_n x) &= \lim_{n\to\infty} F(a_n + b_n x)^n = F'(x) \\ \lim_{n\to\infty} F_{(1)}(c_n + d_n x) &= \lim_{n\to\infty} 1 - (1 - F(c_n + d_n x))^n = F''(x) \end{aligned}$$

$$\begin{aligned} F_{(n)}(a_n + b_n x) &= P(X_{(n)} < a_n + b_n x) = P\left(\frac{X_{(n)} - a_n}{b_n} < x\right) \\ F_{(1)}(c_n + d_n x) &= P(X_{(1)} < c_n + d_n x) = P\left(\frac{X_{(1)} - c_n}{d_n} < x\right) \end{aligned}$$

Maximum of Exponentials, cont.

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ and $a_n = \log(n)/\lambda$, $b_n = 1/\lambda$ then the cdf of $X_{(n)}$ is given by

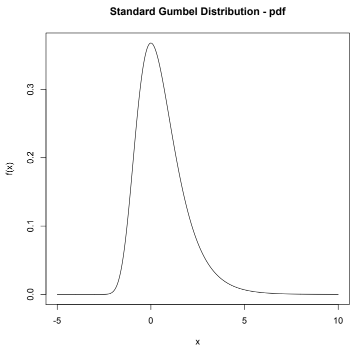
$$\begin{aligned} F_{(n)}(a_n + b_n x) &= F((\log(n) + x)/\lambda)^n \\ &= \left(1 - e^{-\lambda(\log(n)+x)/\lambda}\right)^n \\ &= \left(1 - e^{-\log(n)}e^{-x}\right)^n \\ &= \left(1 - e^{-x}/n\right)^n \end{aligned}$$

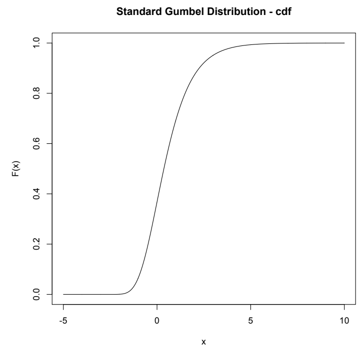
$$\lim_{n\to\infty} F_{(n)}(a_n + b_n x) = \exp(-e^{-x})$$

This is known as the standard Gumbel distribution.

Gumbel Distribution

Let $X \sim \text{Gumbel}(0,1)$ then





$$F(x) = \exp(-e^{-x})$$
$$f(x) = e^{-x}\exp(-e^{-x})$$

$$E(X) = \pi/\sqrt{6}$$
$$\text{Median}(X) = -\log(\log(2))$$

Maximum of Exponentials, cont.

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ and $a_n = \log(n)/\lambda$, $b_n = 1/\lambda$ then if n is large we can use the Standard Gumbel to calculate properties of $X_{(n)}$.

$$\begin{aligned} \text{Median}(X_{(n)}) &= m_{(n)} \\ P(X_{(n)} < m_{(n)}) &= 1/2 \end{aligned}$$

$$\begin{aligned} P\left(\frac{X_{(n)} - a_n}{b_n} < m_G\right) &= 1/2 \\ P(X_{(n)} < a_n + b_n m_G) &= 1/2 \end{aligned}$$

$$\begin{aligned} m_{(n)} &= a_n + b_n m_G \\ &= \frac{1}{\lambda}\log n - \frac{1}{\lambda}\log\log 2 \end{aligned}$$

Maximum of Exponentials, Example

In 2009 Usain Bolt broke the world record in the 100 meters with a time of 9.58 seconds in Berlin, Germany. If we imagine that the running speed in m/s of competitive sprinters is given by an Exponential distribution with $\lambda = 1$. How many sprinters would need to run to have a 50/50 chance of beating Usian Bolt's record (having a faster running speed)?

Let $X \sim \text{Exp}(1)$ then we need to find n such that $\text{Median}(X_{(n)}) \geq 100/9.58$

Maximum of Uniforms

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Unif}(0,1)$ and $a_n = 1$, $b_n = 1/n$ then the cdf of $X_{(n)}$ is given by

$$\begin{aligned} F_{(n)}(x) &= F(x)^n \\ &= x^n \end{aligned}$$

$$\begin{aligned} F_{(n)}(a_n + b_n x) &= F(a_n + b_n x)^n \\ &= (1 + x/n)^n \\ \lim_{n\to\infty} F_{(n)}(a_n + b_n x) &= e^x \end{aligned}$$

This is example of the Reverse Weibul distribution.

Maximum of Paretos

This is a distribution we have not seen yet, but is useful for describing many physical processes. It's key feature is that it has long tails meaning it goes to 0 slower than a distribution like the normal.

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} \text{Pareto}(\alpha, k)$ and $a_n = 0$, $b_n = kn^{1/\alpha}$ then the cdf of X is

$$F_X(x) = \begin{cases} 1 - \left(\frac{k}{x}\right)^\alpha & \text{if } x \geq k, \\ 0 & \text{otherwise} \end{cases}$$

The cdf of $X_{(n)}$ is then given by

$$F_{(n)}(x) = F(x)^n = \left(1 - \left(\frac{k}{x}\right)^\alpha\right)^n$$

$$F_{(n)}(a_n + b_n x) = F(a_n + b_n x)^n = \left(1 - \left(\frac{k}{kxn^{1/\alpha}}\right)^\alpha\right)^n = \left(1 - \frac{x^{-\alpha}}{n}\right)^n$$

$$\lim_{n\to\infty} F_{(n)}(a_n + b_n x) = e^{-x^\alpha}$$

This is example of the Fréchet distribution.