Lecture 13: Order Statistics

Sta230 / Mth 230

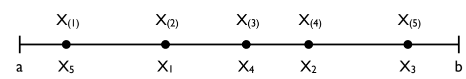
Colin Rundel

March 7, 2014

Order Statistics

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

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



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

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

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.

The first order statistic, $X_{(1)}$ is therefore the smallest X_i and

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

Similarly, the nth order statistic, $X_{(n)}$ is the largest X_i and

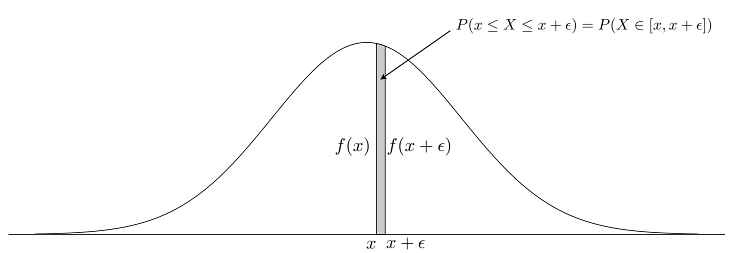
$$X_{(n)} = \max(X_1, \ldots, X_n)$$

Notational Detour

For a continuous random variable

$$f(x)\epsilon \approx P(x \le X \le x + \epsilon) = P(X \in [x, x + \epsilon])$$
$$f(x) \approx P(x \le X \le x + \epsilon) = P(X \in [x, x + \epsilon])/\epsilon$$

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



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

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

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

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

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

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.

Beta Function - Expectation

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

Beta Function - Variance

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

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,

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

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

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

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 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

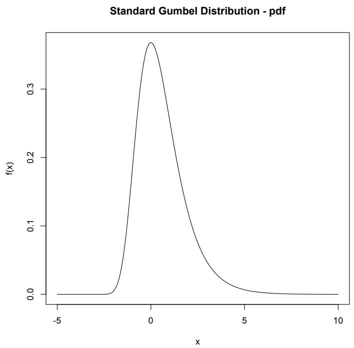
$$
\begin{aligned}
F_{(n)}(a_n + b_n x) &= F((\log(n) + x)/\lambda)^n \\
&= \left(1 - e^{-\lambda(\log(n)+x)/\lambda}\right)^n \\
&= \left(1 - e^{-\log(n)} e^{-x}\right)^n \\
&= \left(1 - e^{-x}/n\right)^n
\end{aligned}
$$

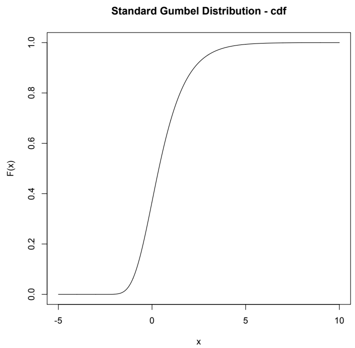
$$\lim_{n\to\infty} F_{(n)}(a_n + b_n x) = \exp(-e^{-x})$$

This is an example of 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 $X_{(n)}$.

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

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

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

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

$$F_{(n)}(x) = F(x)^n = x^n$$

$$F_{(n)}(a_n + b_n x) = F(a_n + b_n x)^n = (a_n + b_n x)^n = (1 + x/n)^n$$

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

This is an 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

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

The cdf of $X_{(n)}$ is 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 an example of the Fréchet distribution.

GEV Distribution

It turns out that these distributions are all special cases of another distribution, the general extreme value distribution.

It turns out that it can be shown that the GEV is the limiting distribution for the, properly normalized, maximum of a sequence of iid random variables.

$$f(x|\mu,\sigma,\xi) = \frac{1}{\sigma}\, t(x)^{\xi+1}\, e^{-t(x)}$$
$$F(x|\mu,\sigma,\xi) = e^{-t(x)}$$

$$t(x) = \begin{cases} \left(1+\left(\frac{x-\mu}{\sigma}\right)\xi\right)^{-1/\xi} & \text{if } \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & \text{if } \xi = 0 \end{cases}$$

GEV Distribution, cont.

If $\xi = 0$,

$$F(x|\mu,\sigma,\xi=0) = e^{-e^{-(x-\mu)/\sigma}} \quad \rightarrow \quad \text{Gumbel dist.}$$

If $\xi > 0$,

$$F(x|\mu,\sigma,\xi>0) = e^{-((x-\mu)/\sigma)^{-1/\xi}} \quad \rightarrow \quad \text{Fréchet dist.}$$

If $\xi < 0$,

$$F(x|\mu,\sigma,\xi>0) = e^{-(-(x-\mu)/\sigma)^{-1/\xi}} \quad \rightarrow \quad \text{Reverse Weibull dist.}$$

GEV Distribution, cont.