# Order Statistics — Quarto Notebook

## Table of contents



## Setup

```r
library(tidyverse)
```

## Order statistics

When ordering a collection of independent continuous random variables with common cdf $F$ (and pdf $f$), let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} F$ with pdf $f$. Assume $f(x) \geq 0$ for $x \in (a, b)$ and 0 otherwise. We define the order statistics as

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

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

## CDFs and PDFs of the minimum and maximum

For the minimum $X_{(1)}$:

$F_{X_{(1)}}(x_1) = \mathbb{P}(X_{(1)} \leq x_1) = 1 - \mathbb{P}(X_{(1)} > x_1)$

$= 1 - \mathbb{P}(\text{all } X_i > x_1)$

$= 1 - [\mathbb{P}(X_1 > x_1)\, \mathbb{P}(X_2 > x_1) \cdots \mathbb{P}(X_n > x_1)]$

$= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > x_1)$

$= 1 - \prod_{i=1}^{n} (1 - F(x_1))$

$= 1 - [1 - F(x_1)]^n \, \mathbb{I}_{(a,b)}(x_1)$

Differentiating,

$f_{X_{(1)}}(x_1) = nf(x_1)\, [1 - F(x_1)]^{n-1} \, \mathbb{I}_{(a,b)}(x_1)$

For the maximum $X_{(n)}$:

$F_{X_{(n)}}(x_n) = \mathbb{P}(X_{(n)} \leq x_n) = \mathbb{P}(\text{all } X_i \leq x_n)$

$= \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_n)$

$= [F(x_n)]^n \, \mathbb{I}_{(a,b)}(x_n)$

$f_{X_{(n)}}(x_n) = n\, [F(x_n)]^{n-1} f(x_n)\, \mathbb{I}_{(a,b)}(x_n)$

## Example: Uniform$(0, 1)$

Suppose we have a random sample of $n$ Uniform values, $X_1, X_2, ..., X_n \sim U(0, 1)$.

$f(x) = \mathbb{I}_{(0,1)}(x)$

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

Then

$f_{X_{(1)}}(x_1) = n(1 - x_1)^{n-1} \, \mathbb{I}_{(0,1)}(x_1)$

$f_{X_{(n)}}(x_n) = nx_n^{n-1} \, \mathbb{I}_{(0,1)}(x_n)$

### Example: Exponential$(\lambda)$

Suppose we have a random sample of $n$ Exponential random variables, $X_1, X_2, ..., X_n \sim \text{Exp}(\lambda)$.

$f(x) = \lambda e^{-\lambda x}\, \mathbb{I}_{(0,\infty)}(x)$

$F(x) = (1 - e^{-\lambda x})\, \mathbb{I}_{(0,\infty)}(x)$

So

$f_{X_{(1)}}(x_1) = n\lambda e^{-n\lambda x_1}\, \mathbb{I}_{(0,\infty)}(x_1)$

$\Rightarrow\;\; X_{(1)} \sim \text{Exp}(n\lambda)$

$f_{X_{(n)}}(x_n) = n\,[1 - e^{-\lambda x_n}]^{n-1}\, e^{-\lambda x_n}\, \mathbb{I}_{(0,\infty)}(x_n)$

An immediate application (with $t > 0$):

$\mathbb{P}(X_{(1)} > t) = e^{-n\lambda t}\, \mathbb{I}_{(0,\infty)}(t)$

### General $k$th order statistic

$$f_{X_{(k)}}(x_k) = \frac{n!}{(k-1)!(n-k)!}\,[F(x_k)]^{k-1}\, f(x_k)\,[1 - F(x_k)]^{n-k}\, \mathbb{I}_{(a,b)}(x_k)$$

#### Joint density of the $j$th and $k$th order statistics $(j < k)$

$$f_{X_{(j)},X_{(k)}}(x_j, x_k) = \frac{n!}{(j-1)!(k-j-1)!(n-k)!}\,[F(x_j)]^{j-1} f(x_j)$$
$$\times\,[F(x_k) - F(x_j)]^{k-j-1} f(x_k)\,[1 - F(x_k)]^{n-k}\, \mathbb{I}_{(a,b)}(x_j)\, \mathbb{I}_{(x_j,b)}(x_k)$$

#### Example: Uniform$(0,1)$ again

$$f_{X_{(k)}}(x_k) = \frac{n!}{(k-1)!(n-k)!}\, x_k^{k-1}(1 - x_k)^{n-k}\, \mathbb{I}_{(0,1)}(x_k)$$

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

#### Joint density of the minimum and maximum for Uniform$(0,1)$

$$f_{X_{(1)},X_{(n)}}(x_1, x_n) = n(n-1)\,(x_n - x_1)^{n-2}\, \mathbb{I}_{(0,1)}(x_1)\, \mathbb{I}_{(x_1,1)}(x_n)$$

## R helpers (optional)

The following chunk uses tidyverse style to simulate samples and empirically verify the formulas.

```r
set.seed(1)
n <- 5
m <- 100000

df <- tibble(id = 1:m) %>%
  mutate(samples = map(id, ~ runif(n))) %>%
  mutate(mins = map_dbl(samples, min),
         maxs = map_dbl(samples, max))

df %>% summarise(mean_min = mean(mins), mean_max = mean(maxs))
```

```
# A tibble: 1 x 2
  mean_min mean_max
     <dbl>    <dbl>
1    0.166    0.833
```

```r
set.seed(1)
n <- 5
lambda <- 2
m <- 100000

df <- tibble(id = 1:m) %>%
  mutate(samples = map(id, ~ rexp(n, rate=lambda))) %>%
  mutate(mins = map_dbl(samples, min))

df %>% summarise(mean_min = mean(mins), theoretical = 1/(n*lambda))
```

```
# A tibble: 1 x 2
  mean_min theoretical
     <dbl>       <dbl>
1   0.1000         0.1
```