Learning Objectives

- Distribution of $X_{(1)}$
- Distribution of $X_{(n)}$
- Distribution of $(X_{(1)}, \ldots, X_{(n)})$
- Distribution of $X_{(r)}$
- Joint Distribution of $X_{(r)}$ and $X_{(s)}$
- Distribution of sample range $R = X_{(n)} - X_{(1)}$

Order Statistics (O.S.)

- $X_1, X_2, \ldots, X_n$: random sample of size n
- F_X: cdf (continuous) of the population
- $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$: corresponding order statistics such that

$$X_{(1)} < X_{(2)} < \ldots < X_{(n)}$$

- $X_{(r)}$: r^{th} order statistic, $1 \leq r \leq n$.

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

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F$. Then, the distribution of $X_{(1)}$ is

$$f_{X_{(1)}}(y_1) = n[1 - F(y_1)]^{n-1} f(y_1).$$

Proof:

$$P(X_{(1)} > y_1) = P(X_1 > y_1, \ldots, X_n > y_n) = \prod_{i=1}^{n} P(X_i > y_i)$$
$$= [1 - F(y_1)]^n$$

So, $F_{X_{(1)}}(y_1) = 1 - [1 - F(y_1)]^n$.
Since F is absolutely continuous, we have the pdf of $X_{(1)}$ as

$$f_{X_{(1)}}(y_1) = n[1 - F(y_1)]^{n-1} f(y_1).$$

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

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F$. Then, the distribution of $X_{(n)}$ is

$$f_{X_{(n)}}(y_n) = n[F(y_n)]^{n-1} f(y_n).$$

Proof:

$$\begin{aligned} P(X_{(n)} \leq y_n) = P(X_1 \leq y_1, \ldots, X_n \leq y_n) &= \prod_{i=1}^{n} P(X_i \leq y_i) \\ &= [F(y_n)]^n \end{aligned}$$

So, $F_{X_{(n)}}(y_n) = [F(y_n)]^n$.
Since F is absolutely continuous, we have the pdf of $X_{(n)}$ as

$$f_{X_{(n)}}(y_n) = n[F(y_n)]^{n-1} f(y_n).$$

Examples

- Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} Exp(\lambda)$. Then, $X_{(1)} \sim Exp(n\lambda)$.

- Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} U[0, \theta]$. Then, the distribution of $X_{(n)}$ is

$$f_{X_{(n)}}(y_n) = \frac{n y_n^n}{\theta^n}, \ 0 \le y_n \le 1.$$

Joint distribution of $X_{(1)}, \ldots, X_{(n)}$

Let $X_1, \ldots, X_n$ is a random sample from continuous distribution F with pdf f. Define

$$Y_i = X_{(i)},\ i = 1, \ldots, n.$$

Then, the joint distribution of $(Y_1, \ldots, Y_n)$ is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = n!\prod_{i=1}^{n} f(y_i),\ -\infty < y_1 < \ldots < y_n < \infty.$$

Proof: We know that the joint pdf of $\boldsymbol{X} = (X_1, \ldots, X_n)$ is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i)$$

Joint distribution of $X_{(1)}, \ldots, X_{(n)}$ contd.

There are $n\,!$ inverse images of the above transformation, such as,

$$
\begin{aligned}
A_1 &= \{\boldsymbol{x} : x_1 < x_2 < x_3 < \ldots < x_n\} \\
A_2 &= \{\boldsymbol{x} : x_2 < x_1 < x_3 < \ldots < x_n\} \\
\vdots & \qquad\qquad \vdots \\
A_n &= \{\boldsymbol{x} : x_n < x_{n-1} < \ldots < x_1\}
\end{aligned}
$$

In the first region A_1, the inverse image is

$$
\begin{aligned}
x_1 &= y_1 \\
x_2 &= y_2 \\
\vdots &\quad \vdots \\
x_n &= y_n
\end{aligned}
$$

and the corresponding Jacobian $J = 1$.

Joint distribution of $X_{(1)}, \ldots, X_{(n)}$ contd.

In the second region A_2, the inverse image is

$$\begin{aligned} x_1 &= y_2 \\ x_2 &= y_1 \\ \vdots &\quad \vdots \\ x_n &= y_n \end{aligned}$$

and the corresponding Jacobian $J = -1$.
Similarly, in the last region A_n, the inverse image is

$$\begin{aligned} x_1 &= y_n \\ x_2 &= y_{n-1} \\ \vdots &\quad \vdots \\ x_n &= y_1 \end{aligned}$$

and the corresponding Jacobian $J = (-1)^n$.

Joint distribution of $X_{(1)}, \ldots, X_{(n)}$ contd.

Thus, in all the cases, $|J| = 1$.

Further, the density in every region is $\prod_{i=1}^{n} f(y_i)$.

Since there are $n\,!$ region, the pdf of $(Y_1, \ldots, Y_n)$ is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = n\,! \prod_{i=1}^{n} f(y_i), \; -\infty < y_1 < \ldots < y_n < \infty.$$

Distribution of $X_{(r)}$, $1 \leq r \leq n$

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F$. Then, the distribution of $Y_r = X_{(r)}$ is

$$f_{Y_r}(y_r) = \frac{n\,!}{(r-1)\,!(n-r)\,!}[F(y_r)]^{r-1}[1-F(y_r)]^{n-r}f(y_r),$$

where $-\infty < y_r < \infty$.

Proof: We know that the joint distribution of $(Y_1, \ldots, Y_n)$ is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = n\,!\prod_{i=1}^{n} f(y_i), \; -\infty < y_1 < \ldots < y_n < \infty.$$

In order to find the marginal pdf of Y_r, $f_{\boldsymbol{Y}}(\boldsymbol{y})$ has to be integrated w.r.t. $(y_1, \ldots, y_{r-1}, y_{r+1}, \ldots, y_n)$, i.e.,

Distribution of $X_{(r)}$ contd.

$$f_{Y_r}(y_r) = \int\limits_{y_r}^{\infty} \cdots \int\limits_{y_{n-2}}^{\infty} \int\limits_{y_{n-1}}^{\infty} \int\limits_{-\infty}^{y_r} \cdots \int\limits_{-\infty}^{y_2} n\,! \prod_{i=1}^{n} f(y_i) dy_1 \ldots dy_{r-1} dy_n dy_{n-1} \ldots dy_{r+1}$$

Now, integrating

$$\begin{aligned}
f(y_1)\ w.r.t.\ y_1 &\longrightarrow F(y_2) \\
F(y_2)f(y_2)\ w.r.t.\ y_2 &\longrightarrow \frac{1}{2}[F(y_3)]^2 \\
\frac{1}{2}F^2(y_3)f(y_3)\ w.r.t.\ y_3 &\longrightarrow \frac{1}{3\,!}[F(y_4)]^3 \\
&\vdots \\
???\ w.r.t.\ y_{r-1} &\longrightarrow \frac{1}{(r-1)\,!}[F(y_r)]^{r-1}.
\end{aligned}$$

Distribution of $X_{(r)}$ contd.

Further, integrating

$$
\begin{aligned}
f(y_n)\ w.r.t.\ y_n &\longrightarrow [1 - F(y_{n-1})] \\
[1 - F(y_{n-1})]f(y_{n-1})\ w.r.t.\ y_{n-1} &\longrightarrow \frac{1}{2}[1 - F(y_{n-2})]^2 \\
\frac{1}{2}[1 - F(y_{n-1})]^2 f(y_{n-2})\ w.r.t.\ y_{n-2} &\longrightarrow \frac{1}{3\,!}[1 - F(y_{n-3})]^3 \\
&\vdots \\
???\ w.r.t.\ y_{r+1} &\longrightarrow \frac{1}{(n-r)\,!}[1 - F(y_r)]^{n-r}.
\end{aligned}
$$

Thus combining the above terms, we get:

$$
f_{Y_r}(y_r) = \frac{n\,!}{(r-1)\,!(n-r)\,!}[F(y_r)]^{r-1}[1 - F(y_r)]^{n-r} f(y_r),
$$

$-\infty < y_r < \infty$.

Remark

- Note that

$$f_{Y_r}(y_r) = \frac{n\,!}{(r-1)\,!(n-r)\,!}[F(y_r)]^{r-1}[1-F(y_r)]^{n-r}f(y_r)$$
$$= \frac{1}{Be(r, n-r+1)}[F(y_r)]^{r-1}[1-F(y_r)]^{n-r}f(y_r), -\infty < y_r < \infty.$$

- **Special case:** $X_1, \ldots, X_n \overset{i.i.d.}{\sim} U[0,1]$.

$$f_{Y_r}(y_r) = \frac{1}{Be(r, n-r+1)} y_r^{r-1}(1-y_r)^{n-r},\ 0 < y_r < 1.$$

Thus, $Y_r \sim Be(r, n-r+1)$.

Joint distribution of $X_{(r)}$ and $X_{(s)}$

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F$. Then, the distribution of

$$(Y_r, Y_s) = (X_{(r)}, X_{(s)}), \quad r < s,$$

is

$$f_{Y_r, Y_s}(y_r, y_s) = \frac{n\,!}{(r-1)\,!(n-s)\,!(s-r-1)\,!} [F(y_r)]^{r-1}[1-F(y_s)]^{n-r}[F(y_s)-F(y_r)]^{s-r-1} f(y_r) f(y_s),$$

where $-\infty < y_r < y_s < \infty$.

Distribution of sample range $R = X_{(n)} - X_{(1)}$

Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F$. Then, the desired pdf of $R = X_{(n)} - X_{(1)}$ is

$$f_R(r) = n(n-1) \int_r^\infty [F(s) - F(s-r)]^{n-2} f(s-r) f(s) ds.$$

Proof: With $r = 1$ and $s = n$, the joint pdf of $Y_1 = X_{(1)}$ and $Y_n = X_{(n)}$ is

$$f_{Y_1,Y_n}(y_1, y_n)$$
$$= n(n-1)[F(y_n) - F(y_1)]^{n-2} f(y_1) f(y_n), \; -\infty < y_1 < y_n < \infty$$

Distribution of sample range R contd.

Now, define the transformation $R = Y_n - Y_1$ and $S = Y_n$. The inverse transformation is $Y_1 = S - R$ and $Y_n = S$, and the corresponding Jacobian is $|J| = 1$.

Thus, the joint pdf of R and S is

$$f_{R,S}(r,s) = n(n-1)[F(s) - F(s-r)]^{n-2} f(s-r) f(s), \; 0 < r < s$$

The desired pdf of R is

$$f_R(r) = n(n-1) \int_r^{\infty} [F(s) - F(s-r)]^{n-2} f(s-r) f(s) ds$$

Distribution of sample range R contd.

- **Special case:** $X_1, \ldots, X_n \overset{i.i.d.}{\sim} U[0,1]$.

The pdf of R is

$$f_R(r) = n(n-1)\int_r^1 r^{n-2} ds = n(n-1)r^{n-2}(1-r),\ 0 < r < 1$$

Thus, $R \sim Be(n-1, 2)$.