# A basic property of order statistics - distributions of arbitrary order

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

June 11, 2024

## 1 Introduction

Let $X_1, X_2, \ldots, X_n$ denote a random sample of size $n$ from a cumulative distribution function $F(.)$. Then $Y_1 \leq Y_2 \leq \cdots \leq Y_n$ , where the $\mathrm{Y}_i$ are arranged in order of increasing magnitudes and are defined to be the order statistics corresponding to the random sample $X_1, X_2, \ldots, X_n$. Thus the $Y_i$ are functions of the random sample $X_1, X_2, \ldots, X_n$ and are in order. While the random sample itself has independent properties, the order statisitics are not independent since if $Y_i \geq y$ then it must be that $Y_{i+1} \geq y$.

Obtaining expressions for the distributions of the minimum and maximum of independent random variables with the same cumulative distribution is straightforward while getting expressions for the cumulative and density distributions of an arbitrary order statistic is more tricky. We start with the easy cases and then move on to the general case.

## 2 Distribution of the maximum and minimum

Let $X_1, X_2, \ldots, X_n$ be $n$ random variables and let $Y_1 = \min\{X_1, X_2, \ldots, X_n\}$ and $Y_n = \max\{X_1, X_2, \ldots, X_n\}$ . The distribution of $Y_n$ is found as follows. $F_{Y_n}(y) = \mathbb{P}[Y_n \leq y] = \mathbb{P}[X_1 \leq y; \ldots; X_n \leq y]$ since the largest of the $X_i$ is less than $y$ if and only if all of the $X_i$ are less than or equal to $y$. Now because the $X_i$ ae assumed independent we have that:

$$F_{Y_n}(y) = \mathbb{P}[X_1 \leq y; \ldots; X_n \leq y] = \prod_{i=1}^{n} \mathbb{P}[X_i \leq y] = \prod_{i=1}^{n} F_{X_i}(y) \tag{1}$$

If we now assume that the $X_i$ have the same distribution, $F_X(.)$, we get the nice result:

$$\boxed{\prod_{i=1}^{n} F_{X_i}(y) = [F_X(y)]^n} \tag{2}$$

We can now easily find the density function of $Y_n$ assuming the $X_i$ are independent and identically distributed with common density function $f_X(.)$ and cumulative distribution function $F_X(.)$:

$$\boxed{f_{Y_n}(y) = \frac{d}{dy} F_{Y_n}(y) = n[F_X(y)]^{n-1} f_X(y)} \tag{3}$$

Moving on to the distributions for the minimum $Y_1$ we see that:

$$F_{Y_1}(y) = \mathbb{P}[Y_1 \leq y] = 1 - \mathbb{P}[Y_1 > y] = 1 - \mathbb{P}[X_1 > y; \ldots; X_n > y] \tag{4}$$

which must be the case since $Y_1$ is greater than $y$ if and only if every $X_i > y$. Once again, assuming independence of the $X_i$ we have that:

$$\boxed{F_{Y_1}(y) = 1 - \mathbb{P}[X_1 > y; \ldots; X_n > y] = 1 - \prod_{i=1}^{n} \mathbb{P}[X_i > y] = 1 - \prod_{i=1}^{n} [1 - F_{X_i}(y)]} \tag{5}$$

Again if we assume that the $X_i$ are identically distributed with common cumulative function $F_X(.)$ we have that:

$$\boxed{F_{Y_1}(y) = 1 - \prod_{i=1}^{n} [1 - F_{X_i}(y)] = 1 - [1 - F_X(y)]^n} \tag{6}$$

If the common probability density function is $f_X(.)$ we have that:

$$\boxed{f_{Y_1}(y) = \frac{d}{dy}F_{Y_1}(y) = n[1 - F_X(y)]^{n-1}\, f_X(y)} \tag{7}$$

## 3 Distributions for arbitrary order statistics

We now move on to the more complex case of arbirtrary order statistics. We let $Y_1 \leq Y_2 \leq \cdots \leq Y_n$ be the order statistics from a cumulative distribution function $F(.)$. The marginal cumulative distribution function of $Y_\alpha$ for $\alpha = 1, 2, \ldots, n$ is then given by:

$$\boxed{F_{Y_\alpha}(y) = \sum_{j=\alpha}^{n} \binom{n}{j} [F(y)]^j \,[1 - F(y)]^{n-j}} \tag{8}$$

Before proving this let's see if (8) makes sense for $Y_1$ and $Y_n$.

$$F_{Y_n}(y) = \sum_{j=n}^{n} \binom{n}{j} [F(y)]^j \,[1 - F(y)]^{n-j} = [F(y)]^n \tag{9}$$

which agrees with (2). Also,

$$\begin{aligned} F_{Y_1}(y) =& \sum_{j=1}^{n} \binom{n}{j} [F(y)]^j \,[1 - F(y)]^{n-j} \\ =& \sum_{j=0}^{n} \binom{n}{j} [F(y)]^j \,[1 - F(y)]^{n-j} - [1 - F(y)]^n \\ =& \Big[F(y) + 1 - F(y)\Big]^n - [1 - F(y)]^n \\ =& 1 - [1 - F(y)]^n \end{aligned} \tag{10}$$

which agrees with (6).

To prove (8) we fix $y$ and let $Z_i = \mathbb{I}_{(-\infty,y)}(X_i)$ where $\mathbb{I}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$

is the standard indicator function. With this set up we have that:

$$\sum_{i=1}^{n} Z_i = \text{ the number of } X_i \leq y \tag{11}$$

Clearly the sum $\sum_{i=1}^{n} Z_i$ has a binomial distribution with parameters $n$ and $F(y)$ ie $B(n, F(y))$. The crux of the proof revolves around the observation that $F_{Y_\alpha}(y) = \mathbb{P}[Y_\alpha \leq y] = \mathbb{P}[\sum_{i=1}^{n} Z_i \geq \alpha]$. In essence this is an assertion that the two events $\{Y_\alpha \leq y\}$ and $\{\sum_i Z_i \geq \alpha\}$ are equivalent.

If the $\alpha$th order statistic is less than or equal to $y$, then it must be the case that the number of $X_i$ less than or equal to $y$ is greater than or equal to $\alpha$ and conversely. To see this recall that the $\alpha$th order statistic represents an ordering of the $X_i$ from smallest to largest eg $X_2 \leq X_4 \leq X_1 \leq X_3$ so that we would identify $X_2$ with $Y_1$, $X_4$ with $Y_2$, $X_1$ with $Y_3$ and $X_3$ with $Y_4$. Thus if $\alpha = 3$ we can see if $Y_3 \leq y$ then there are 3 $X_i$ less than or equal to $y$ ie the number of such $X_i$ is trivially at least 3. Conversely, if the number of $X_i \leq y$ is at least 3, this means that $Y_3 \leq y$.

With this background we now see that:

$$F_{Y_\alpha}(y) = \mathbb{P}[Y_\alpha \leq y] = \mathbb{P}[\sum_{i=1}^{n} Z_i \geq \alpha] = \sum_{j=\alpha}^{n} \binom{n}{j} [F(y)]^j \, [1 - F(y)]^{n-j} \tag{12}$$

since $\sum_{i=1}^{n} Z_i$ is $B(n, F(y))$.

An equivalent but simpler way of deriving (12) is to note that the event $\{Y_\alpha \leq y\}$ implies that at least $\alpha$ among the $n$ variables are $\leq y$. Thus there are at least $\alpha$ “successes “ in $n$ independent trials and so we get the same binomial sum. The reason we have at least $\alpha$ “successes” is that there may be multiple random variables $X_i \leq y$.

# 4 Deriving the probability density function of the $\alpha$th order statistic

The aim is to show that if $X_1, X_2, \ldots, X_n$ is a random sample from the probability density function $f(.)$ with cumulative distribution function $F(.)$ and $Y_1 \leq Y_2 \leq \cdots \leq Y_n$ are the corresponding order statistics then:

$$\boxed{f_{Y_\alpha}(y) = \frac{n!}{(\alpha - 1)!(n - \alpha)!} [F(y)]^{\alpha - 1} \, [1 - F(y)]^{n-\alpha} \, f(y)} \tag{13}$$

It ought to be clear that in order to get (13) we need to differentiate (12) but because of the summation structure of (12) some telescoping must occur to just get one term out of the sum. Differentiating we have:

$$
\begin{aligned}
f_{Y_\alpha}(y) =& \frac{d}{dy} F_{Y_\alpha}(y) \\
=& \sum_{j=\alpha}^{n} \left\{ \binom{n}{j} j [F(y)]^{j-1} \left[1 - F(y)\right]^{n-j} f(y) + \binom{n}{j} (n-j) [F(y)]^{j} [1 - F(y)]^{n-j-1} \times -f(y) \right\} \\
=& \sum_{j=\alpha}^{n} \frac{n!}{(n-j)!(j-1)!} [F(y)]^{j-1} \left[1 - F(y)\right]^{n-j} f(y) + \\
& \sum_{j=\alpha}^{n-1} \frac{-n!}{(n-j-1)!j!} [F(y)]^{j} [1 - F(y)]^{n-j-1} f(y) \\
=& \sum_{j=\alpha}^{n} A_j + \sum_{j=\alpha}^{n-1} B_j
\end{aligned}
\tag{14}
$$

Note that in the expansion of (8) before differentiation the final term is $[F(y)]^n$ so that upon differentiation we have a $(n-1)$ exponent. This is why the second sum only runs to $n-1$. We have that $A_\alpha = \frac{n!}{(n-\alpha)!(\alpha-1)!} [F(y)]^{\alpha-1} [1 - F(y)]^{n-\alpha} f(y)$ which is the only term that remains in the sum. The telescoping can be seen by noting that :

$$
\begin{aligned}
A_{\alpha+k} =& \frac{n!}{(n-\alpha-k)!\ (\alpha+k-1)!} [F(y)]^{\alpha+k-1} [1 - F(y)]^{n-\alpha-k} f(y) \\
B_{\alpha+k-1} =& \frac{-n!}{(n-\alpha-k)!\ (\alpha+k-1)!} [F(y)]^{\alpha+k-1} [1 - F(y)]^{n-\alpha-k} f(y)
\end{aligned}
\tag{15}
$$

Thus $A_{\alpha+k} + B_{\alpha+k-1} = 0$ and the sum telescopes to $A_\alpha = \frac{n!}{(n-\alpha)!(\alpha-1)!} [F(y)]^{\alpha-1} [1 - F(y)]^{n-\alpha} f(y)$ as advertised.

An alternative way of getting (13) is to use the limit definition of the derivative as follows:

$$
\begin{aligned}
f_{Y_\alpha}(y) &= \lim_{\Delta y \to 0} \frac{F_{Y_\alpha}(y+\Delta y) - F_{Y_\alpha}(y)}{\Delta y} \\
&= \lim_{\Delta y \to 0} \frac{\mathbb{P}[y < Y_\alpha \le y + \Delta y]}{\Delta y} \\
&= \lim_{\Delta y \to 0} \frac{\mathbb{P}[\ (\alpha-1) \text{ of the } X_i \le y;\text{ one } X_i \text{ in } (y, y+\Delta y)\ ;\ (n-\alpha) \text{ of the } X_i > y + \Delta y\ ]}{\Delta y} \\
&= \lim_{\Delta y \to 0} \left\{ \frac{n!}{(\alpha-1)!\,1!\,(n-\alpha)!} \frac{[F(y)]^{\alpha-1}[F(y+\Delta y) - F(y)]\,[1-F(y+\Delta y)]^{n-\alpha}}{\Delta y} \right\} \\
&= \frac{n!}{(\alpha-1)!(n-\alpha)!}[F(y)]^{\alpha-1}[1-F(y)]^{n-\alpha}\, f(y)
\end{aligned} \tag{16}
$$

One can obtain even more complicated expressions using this last technique. For instance, the joint density of $Y_\alpha$ and $Y_\beta$ for $1 \le \alpha < \beta \le n$ multiplied by $\Delta x\, \Delta y$ approximates the relevant probability. That is,

$$
\begin{aligned}
f_{Y_\alpha, Y_\beta}(x,y)\Delta x \Delta y \approx &\mathbb{P}[x < Y_\alpha \le x + \Delta x; y < Y_\beta \le y + \Delta y] \\
\approx &\mathbb{P}[\ (\alpha-1) \text{ of the } X_1 \le x;\text{ one } X_i \\
&\quad \text{in } (x, x+\Delta x]\ ;\ (\beta-\alpha-1 \text{ of the } X_i \text{ in } (x+\Delta x, y]; \\
&\quad \text{one of } X_i \text{ in } (y, y+\Delta y](n-\beta) \text{ of the } X_i > y + \Delta y] \\
\approx &\frac{n!}{(\alpha-1)!\,1!\,(\beta-\alpha-1)!1!(n-\beta)!} \\
&\times [F(x)]^{\alpha-1}[F(y) - F(x+\Delta x)]^{\beta-\alpha-1}[1-F(y+\Delta y)]^{n-\beta} f(x)\Delta x f(y)\Delta y
\end{aligned} \tag{17}
$$

It is important to note that this differential development does not contradict the basic proposition that $Y_\alpha \le x$ implies that there are at least $\alpha$ $X_i \le x$, for we have arbitrarily small $\Delta x$ and there are $(\alpha - 1)$ of the $X_i \le x$ and one $X_i$ in $(x, x + \Delta x]$ so that approaching the limit from the right we will have $\alpha$ instances of $X_i \le x$.

Hence

$$
f_{Y_\alpha, Y_\beta}(x,y) = \frac{n!}{(\alpha-1)!(\beta-\alpha-1)!(n-\beta)!}[F(x)]^{\alpha-1}[F(y)-F(x)]^{\beta-\alpha-1}\,[1-F(y)]^{n-\beta} f(x) f(y) \tag{18}
$$

for $x < y$ and $f_{Y_\alpha,Y_\beta}(x,y) = 0$ for $x \geq y$

Using the indicator function the RHS of (118) can be written as:

$$\frac{n!}{(\alpha-1)!(\beta-\alpha-1)!(n-\beta)!}[F(x)]^{\alpha-1}[F(y)-F(x)]^{\beta-\alpha-1}\,[1-F(y)]^{n-\beta}f(x)f(y)\mathbb{I}_{(x,\infty)}(y) \tag{19}$$

Note the multinomial structure of the scaling factor in (17). A multinomial coefficient has the structure $\frac{n!}{\prod_{i=1}^{k} n_i}$ where $n = \sum_{i=1}^{k} n_i$.

# 5 Obtaining a density by integrating out a variable

If there is any justice in the universe one should be able to go from $f_{Y_\alpha,Y_\beta}(x,y)$ to $f_{Y_\beta}(y)$ by integrating over $x$. Thus if we integrate (18) over $x$ we should get:

$$f_{Y_\beta}(y) = \frac{n!}{(\beta-1)!(n-\beta)!}]F(y)]^{\beta-1}\,[1-F(y)]^{n-\beta}\,f(y) \tag{20}$$

The high level scheme of doing this is to integrate by parts multiple times and one finds that, structurally the “uv “ term is always zero and what is left ultimately collapses to what is required. Along the way a constant arises and it cancels in a nice way to give the final leading constant. The real substance of the process is to work out the number of iterations. We let:

$$C = \frac{n!}{(\alpha-1)!(\beta-\alpha-1)!(n-\beta)!} \tag{21}$$

Thus our required integration of (18) becomes:

$$\begin{aligned} f_{Y_\beta}(y) =& C[1-F(y)]^{n-\beta}f(y)\int_{-\infty}^{x}[F(x)]^{\alpha-1}[F(y)-F(x)]^{\beta-\alpha-1}f(x)\,dx \\ =& C[1-F(y)]^{n-\beta}f(y)\,J \end{aligned} \tag{22}$$

where $J = \int_{-\infty}^{x}[F(x)]^{\alpha-1}[F(y)-F(x)]^{\beta-\alpha-1}f(x)\,dx$.

We now integrate by parts with $u = [F(x)]^{\alpha-1}$ so that $du = (\alpha-1)[F(x)]^{\alpha-2}f(x)\,dx$. Also $dv = [F(y)-F(x)]^{\beta-\alpha-1}f(x)\,dx$ so that $v = -\frac{[F(y)-F(x)]^{\beta-\alpha}}{\beta-\alpha}$

As noted above:

$$uv\Big|_{x=-\infty}^{x=y} = -[F(x)]^{\alpha-1}\frac{[F(y)-F(x)]^{\beta-\alpha}}{\beta-\alpha}\Bigg|_{x=-\infty}^{x=y} = 0 \tag{23}$$

since $\lim_{x\to-\infty} F(x) = 0$ is a property of cumulative distribution functions and $F(.)$ is bounded. It is clear that at each iteration we will choose our $u$ to be the previous one with an index 1 lower and $v$ will increase by 1. This structure ensures that on each iteration our "uv" term is zero so we only need worry about the $\int v\,du$ term.

Thus at the first iteration we have:

$$J = \frac{\alpha-1}{\beta-1}\int_{-\infty}^{x}[F(x)]^{\alpha-2}[F(y)-F(x)]^{\beta-\alpha}f(x)\,dx \tag{24}$$

After the second iteration we have:

$$J = \frac{\alpha-1}{\beta-\alpha}\ \frac{\alpha-2}{\beta-\alpha+1}\int_{-\infty}^{x}[F(x)]^{\alpha-3}[F(y)-F(x)]^{\beta-\alpha+1}f(x)\,dx \tag{25}$$

if the number of iterations is $\gamma$ where $\gamma$ runs from 1 to $\alpha-1$ because when $\gamma=\alpha-1$ the exponent of $F(x)$ will be zero and that term is then just 1. When $\gamma = \alpha-1$ the exponent of $[F(y)-F(x)]$ becomes $\beta-\alpha-1+\gamma = \beta-\alpha-1+\alpha-1 = \beta-2$ (see (22) )

Thus after $\gamma$ iterations we have:

$$
\begin{aligned}
J =& \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{\alpha-(\alpha-1)}{\beta-\alpha-1+\alpha-1} \int_{-\infty}^{x} [F(x)]^{0}[F(y)-F(x)]^{\beta-2} f(x)\, dx \\
=& \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{1}{\beta-2} \int_{-\infty}^{x} [F(y)-F(x)]^{\beta-2} f(x)\, dx \\
=& \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{1}{\beta-2} \int_{F(y)}^{0} u^{\beta-2}\, (-du) \text{ with } u = F(y)-F(x) \\
=& \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{1}{\beta-2} \int_{0}^{F(y)} u^{\beta-2}\, du \\
=& \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{1}{\beta-2} \left[ \frac{u^{\beta-1}}{\beta-1} \right]_{0}^{F(y)} \\
=& \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{1}{\beta-2} \times \frac{1}{\beta-1} \times [F(y)]^{\beta-1}
\end{aligned}
\tag{26}
$$

Going back to (22) and using (21) the integration becomes:

$$
\begin{aligned}
f_{Y_\beta}(y) =& C[1-F(y)]^{n-\beta} f(y) \int_{-\infty}^{x} [F(x)]^{\alpha-1}[F(y)-F(x)]^{\beta-\alpha-1} f(x)\, dx \\
=& \frac{n!}{(\alpha-1)!(\beta-\alpha-1)!(n-\beta)!} \times \frac{\alpha-1}{\beta-\alpha} \times \frac{\alpha-2}{\beta-\alpha+1} \times \cdots \times \frac{1}{\beta-2} \times \frac{1}{\beta-1} \\
& \times [F(y)]^{\beta-1}[1-F(y)]^{n-\beta} f(y) \\
=& \frac{n!}{(\alpha-1)!(\beta-\alpha-1)!(n-\beta)!} \times \frac{(\alpha-1)!}{(\beta-\alpha)(\beta-\alpha+1)\ldots(\beta-2)(\beta-1)} \\
& \times [F(y)]^{\beta-1}[1-F(y)]^{n-\beta} f(y) \\
=& \frac{n!}{\cancel{(\alpha-1)!}(\beta-\alpha-1)!(n-\beta)!} \times \frac{\cancel{(\alpha-1)!}}{(\beta-\alpha)(\beta-\alpha+1)\ldots(\beta-2)(\beta-1)} \\
& \times [F(y)]^{\beta-1}[1-F(y)]^{n-\beta} f(y) \\
=& \frac{n!}{(\beta-1)!(n-\beta)!} \times [F(y)]^{\beta-1}[1-F(y)]^{n-\beta} f(y)
\end{aligned}
\tag{27}
$$

which is (20). Note that $(\beta-1)(\beta-2)\ldots(\beta-\alpha+1)(\beta-\alpha) \times (\beta-\alpha-1)! = (\beta-1)!$

## 5.1 Example

Just to emphasise that (27) is not hocus pocus, consider the case where $n = 10, \alpha = 3, \beta = 7$. Let's work through the integration to get $f_{Y_\beta}(y)$ which by (13) ( with $\beta$ being substituted for $\alpha$ ) should be:

$$f_{Y_7}(y) = \frac{10!}{6!3!}[F(y)]^6[1 - F(y)]^3 f(y) \tag{28}$$

From (18) and bearing in mind that the "uv " term in integration by parts will be zero at each iteration:

$$\begin{aligned}
f_{Y_7}(y) &= \int_{-\infty}^{y} f_{Y_3,Y_7}(x,y)\,dx \\
&= \frac{10!}{2!3!3!}[1-F(y)]^3 f(y)\int_{-\infty}^{y}[F(x)]^2[F(y)-F(x)]^3 f(x)\,dx \\
&= \frac{10!}{2!3!3!}[1-F(y)]^3 f(y)\Bigg\{\underbrace{-[F(x)]^2\frac{[F(y)-F(x)]^4}{4}\Bigg|_{x=-\infty}^{y}}_{=0} + \\
&\quad \frac{2}{4}\int_{-\infty}^{y}[F(x)][F(y)-F(x)]^4 f(x)\,dx \\
&= \frac{10!}{2!3!3!}[1-F(y)]^3 f(y)\,\frac{2}{4}\times\frac{1}{5}\int_{-\infty}^{y}[F(y)-F(x)]^5 f(x)\,dx \\
&= \frac{10!}{2!3!3!}[1-F(y)]^3 f(y)\,\frac{2}{4}\times\frac{1}{5}\left[\frac{u^6}{6}\right]_0^{F(y)} \quad \text{where } u = F(y)-F(x) \\
&= \frac{10!}{2!3!3!}[1-F(y)]^3 f(y)\,\frac{2}{4}\times\frac{1}{5}\times\frac{1}{6}\,[F(y)]^6 \\
&= \frac{10!}{\cancel{2!}3!3!}[1-F(y)]^3 f(y)\,\frac{\cancel{2!}}{4.5.6}\,[F(y)]^6 \\
&= \frac{10!}{6!3!}[F(y)]^6[1-F(y)]^3\,f(y)
\end{aligned} \tag{29}$$

as advertised in (28).

# 6 History

Created 11 June 2024