Order statistics

Min and Max

It is useful to understand the following simple result. Let $X_1, \dots, X_n$ be n mutually independent random variables on a probability space $(\Omega, \mathcal{F}, P)$, all sharing the same probability distribution $Q(I) = P(X_i \in I)$. Let $F(s) = Q((-\infty, s]) = P(\{\omega : X_i(\omega) \le s\})$ be the corresponding cumulative distribution function. Define the random variables

$$X_*(\omega) = \min\{X_1(\omega), \dots, X_n(\omega)\},\ X^*(\omega) = \max\{X_1(\omega), \dots, X_n(\omega)\}.$$

If the random variable $X_i, 1 \le i \le n$ are continuous random variables, the cumulative distribution function F_{X^*} of X^* is $F_{X^*}(s) = P(X^* \le s) = F(s)^n$ and the cumulative distribution function F_{X_*} of X_* is $F_{X_*}(s) = P(X_* \le s) = 1 - (1 - F(s))^n$.

Proof: Clearly,

$$P(X^* \le s) = P(X_1 \le s \,\&\, X_2 \le s \,\&\, \dots \,\&\, X_n \le s).$$

Because the X_i are mutually independent and have the same distribution, this is $P(X_1 \le s)^n = F(s)^n$. To compute $P(X_* \le s)$, we write $P(X_* \le s) = 1 - P(X_* > s)$ and observe that

$$P(X_* > s) = P(X_1 > s \,\&\, X_2 > s \,\&\, \dots \,\&\, X_n > s) = P(X_1 > s)^n.$$

The desired follows because $P(X_1 > s) = 1 - P(X_1 \le s) = 1 - F(s)$.

Let $X_1, \dots, X_n$ be n mutually independent continuous random variables on a probability space $(\Omega, \mathcal{F}, P)$, all sharing the same density function f and cumulative distribution $F(s) = \int_{-\infty}^{s} f(x)dx$. The random variables $X_* = \min\{X_1, \dots, X_n\}$ and $X^* = \max\{X_1, \dots, X_n\}$ have respective densities

$$f_{X^*}(x) = nf(x)F(x)^{n-1} \text{ and } f_{X_*}(x) = nf(x)(1 - F(x))^{n-1}$$

at any point x where f is continuous.

What are "order statistics" (continuous random variables):

Let $V = (X_1, \dots, X_n)$ be a vector of independent continuous random variables with identical density functions $f_{X_i} = f$. The density function of V is $f_V(x_1, \dots, x_n) = f(x_1) \dots f(x_n)$. Now, we order the coordinate of V to obtain a new vector

$$U = (X_{(1)}, \ldots, , X_{(n)})$$

where $X_{(i)}$ is the i-th smallest value among $X_1, \ldots, X_n$. So, $X_{(1)} = \min\{X_1, \ldots, X_n\}$, $X_{(2)}$ is the smallest among the remaining variables: if $X_{(1)} = X_j$ for some j then $X_{(2)} = \min\{X_k : k \neq j\}$. Obviously, $X_{(n)} = \max\{X_1, \ldots, X_n\}$ and also $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n-1)} \leq X_{(n)}$.

Density of order statistics for independent identically distributed RVs Let $\Delta = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 < x_2 < \cdots < x_n\}$. The density function for U is

$$f_U(x_1, \ldots, x_n) = n!\ \mathbf{1}_\Delta(x_1, \ldots, x_n)\ f(x_1) \cdots f(x_n).$$

Proof Given $x_1 < \cdots < x_n$, pick ϵ so small that it is smaller than $\min\{x_{i+1} - x_i : i = 1, 2, \ldots, n-1\}$ and compute $P(X_{(1)} \in (x_1, x_1 + \epsilon), \ldots, X_{(n)} \in (x_n, x_n + \epsilon))$. Let $\theta : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a permutation. We can split the event

$$\{X_{(1)} \in (x_1, x_1 + \epsilon), \ldots, X_{(n)} \in (x_n, x_n + \epsilon)\}$$

into the union over all permutations θ of the disjoint events

$$\{X_{(1)} \in (x_1, x_1 + \epsilon), \ldots, X_{(n)} \in (x_n, x_n + \epsilon) \text{ and } X_{(i)} = X_{\theta(i)}, 1 \leq i \leq n\} = \{X_{\theta(1)} \in (x_1$$

Because the variables X_i are independent with identical density function f, given a fixed permutation θ,

$$P(\{X_{\theta(1)} \in (x_1, x_1 + \epsilon), \ldots, X_{\theta(n)} \in (x_n, x_n + \epsilon)\}) \approx f(x_1) f(x_2) \ldots f(x_n) \epsilon^n$$

(observe how we have reorder the factors \$f(x_i)\$ to obtain the right-hand side) and,

$$P(X_{(1)} \in (x_1, x_1 + \epsilon), \ldots, X_{(n)} \in (x_n, x_n + \epsilon)) \approx n! f(x_1) f(x_2) \ldots f(x_n) \epsilon^n.$$

This yields the desired result.

Cumulative distribution and density for each order statistics Fix $1 \leq j \leq n$. Let F be the common cumulative distribution function shared by the X_is. The density function for $Y = X_{(j)}$ is $f_Y(x) = \binom{n}{j-1, n-j, 1} F(x)^{j-1} (1 - F(x)^{n-j} f(x)$.

Sketch of the proof: For $x \in \mathbb{R}$ and small $\epsilon > 0$, we approximate $P(X_{(j)} \in (x, x + \epsilon))$. Pick a $j-1$ subset $J = \{i_1, \ldots, i_{j-1}\}$ of $\{1, \ldots, n\}$ for the indices of those $j-1$ variables that are less than $X_{(j)}$ and pick a remaining i_j for $X_{(j)} = X_{i_j}$. There are $\binom{n}{j-1, n-j, 1}$ possible choices. For each of these choices

$$P(X_\ell < x \text{ for } \ell \in J \text{ and } X_{i_j} \in (x, x + \epsilon) \text{ and } x < X_k \text{ for } k \notin \{i_j\} \cup J) \approx F(x)^{j-1} (1 - F($$

and the desired conclusion follows. This last approximation requires a little bit of care to be done rigorously.

Min and max of independent uniforms Recall the discussion of order statistics and, in particular the formulas for the densities of the max and min of a n independent equidistributed random variables $X_1, \ldots, X_n$. Here we assume that the X_i's are uniform on the interval $(0,1)$. The cumulative distribution function of a uniform is the function

$$F(s) = \begin{cases} 0 & \text{for } x \in (-\infty, 0), \\ s & \text{for } x \in (0,1), \\ 1 & \text{for } x \in [1, +\infty). \end{cases}$$

It follows that the density of $Y = X_{(n)} = \max\{X_1, \ldots, X_n\}$ is $f_Y(y) = nx^{n-1}\mathbf{1}_{0,1)}(y)$. Similarly, $f_Z(z) = n(1-z)^{n-1}\mathbf{1}_{(0,1)}(z)$. These two distribution belong to a family called the Beta distributions.

Now, if we look at the the density of $X_{(j)}$, $j \in \{1, \ldots, n\}$, we find that it is

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} x^{j-1}(1-x)^{n-j}\mathbf{1}_{(0,1)}(x)$$

which is also a beta distribution.