Order Statistics

May 4, 2020

1 Order Statistics

Given random variables $X_1, X_2, ..., X_n$, it is of interest in some applications to know what the distribution of the largest of the random variable is, or the second largest and so on down to the minimum. These types of quantities are known as order statistics. The book goes into quite a bit of detail with these. We will only focus on two examples, the maximum and the minimum, but we'll briefly introduce all of them, in the following definition:

DEFINITION 1.1. *Let $X_1, X_2, ..., X_n$ be random variables, the associated* ***order statistics*** *are denoted $X_{(1)}, X_{(2)}, ..., X_{(n)}$, and are defined to be ordering of the origin random variables such that*

$$X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}.$$

In particular $X_{(n)} = max\{X_1, X_2, ..., X_n\}$ and $X_{(1)} = min\{X_1, X_2, ..., X_n\}$. So the idea is you sample the n random variables to give you a list of n numbers, then you forget the order you sampled them and rearrange them so they are in increasing order. Then we want to know: what is the probability the maximum (or any other term in the sequence) takes on a certain value? or what is expectation of the maximum? or any other type of question we have asked in this class. These questions can be answered once we know the pdf of $X_{(n)}$ (or whichever r.v. we were interested in).

2 Maximum and Minimum order statistic

To compute the pdf of $X_{(n)} = max\{X_1, X_2, ..., X_n\}$, we can think of $X_{(n)}$ as a function of the random variables and use the ideas we learned about computing the pdf of a function of random variables. In principle, this can be done for any joint distribution of random variables $X_1, X_2, ..., X_n$, but the situation is much nicer when they are iid (independent and identically distributed) and when they are continuous random variables so we will just restrict ourselves to this case.

THEOREM 2.1. *Let $X_1, X_2, ..., X_n$ be independent continuous random variables with the same pdf, $f_X(x)$. Then the pdf of $X_{(n)} = \max\{X_1, X_2, ..., X_n\}$ is:*

$$f_{X_{(n)}}(x) = n(F_X(x))^{n-1} f_X(x)$$

Proof. As usual when computing the pdf of we'll start with the pdf.

$$F_{X_{(n)}}(x) = \mathbb{P}(X_{(n)} \leq x) = \mathbb{P}(max\{X_1, X_2, ..., X_n\} \leq x)$$

then the crucial observation to make is that if the maximum of all the random variables is less than x then all of the random variables must be less than x. Additionally, if all the random variables are less than x then the maximum is less than x so we have:

$$\mathbb{P}(max\{X_1,X_2,...,X_n\}\leq x)=\mathbb{P}(\{X_1\leq x\}\cap\{X_2\leq x\}\cap...\cap\{X_n\leq x\})$$

Then we use that because the X_i's are independent the events above are all independent, and the probability of the intersection of independent events is the product of the individual events.

$$\mathbb{P}(\{X_1\leq x\}\cap\{X_2\leq x\}\cap...\cap\{X_n\leq x\})=\mathbb{P}(\{X_1\leq x\})\mathbb{P}(\{X_2\leq x\}...\mathbb{P}(\{X_n\leq x\})$$

Then because the $X_i's$ have the same distribution, $\mathbb{P}(\{X_i\leq x\})$ is the same for each i, so

$$\mathbb{P}(\{X_1\leq x\})\mathbb{P}(\{X_2\leq x\}...\mathbb{P}(\{X_n\leq x\})=\mathbb{P}(\{X_1\leq x\})^n=(F_X(x))^n$$

Putting this all together:

$$F_{X_{(n)}}(x)=(F_X(x))^n$$

So then we differentiate to get the pdf:

$$f_{X_{(n)}}(x)=\frac{d}{dx}F_{X_{(n)}}(x)=\frac{d}{dx}(F_X(x))^n=n(F_X(x))^{n-1}f_X(x)$$

where the last term comes from the chain rule.

□

Let's apply this in an example.

EXAMPLE 2.2. *Let $X_1,X_2,...,X_n$ be independent random variables uniformly distributed on the interval $[0,1]$. Compute the pdf of $X_{(n)}=\max\{X_1,X_2,...,X_n\}$.*

Solution: The CDF of the X_i's is

$$F_{X_i}(t)=\mathbb{P}(X_i\leq t)=\int_{-\infty}^{t}f_{X_i}(x)dx=\int_0^t 1dx=t$$

for $0\leq t\leq 1$. It's 0 for $t<0$ and 1 for $t>1$.

So the pdf of $X_{(n)}$ is

$$f_{X_{(n)}}(t)=n(t)^{n-1}1$$

for $0\leq t\leq 1$ and 0 otherwise.

We can do a similar derivation for the minimum as for the maximum, but it won't work exactly the same. The key observation for the maximum was to notice that if the maximum is less than some value t then all the random variables must be less than t. It is not true that the minimum being less than t can be related to the entire set of random variable, but we will be able to do something with the opposite inequality. To exploit this we will work with 1 minus the CDF.

THEOREM 2.3. *Let $X_1,X_2,...,X_n$ be independent continuous random variables with the same pdf, $f_X(x)$. Then the pdf of $X_{(1)}=\min\{X_1,X_2,...,X_n\}$ is:*

$$f_{X_{(1)}}(x)=n(1-F_X(x))^{n-1}f_X(x)$$

Proof. As usual when computing the pdf of we'll start with the pdf.

$$1-F_{X_{(1)}}(x)=\mathbb{P}(X_{(n)}>x)=\mathbb{P}(min\{X_1,X_2,...,X_n\}>x)$$

then the crucial observation to make is that if the minimum of all the random variables is greater than x then all of the random variables must be greater than x. Additionally, if all the random variables are greater than x then the minimum is greater than x so we have:

$$\mathbb{P}(max\{X_1,X_2,...,X_n\}>x)=\mathbb{P}(\{X_1>x\}\cap\{X_2>x\}\cap...\cap\{X_n>x\})$$

Then we use that because the X_i's are independent the events above are all independent, and the probability of the intersection of independent events is the product of the individual events.

$$\mathbb{P}(\{X_1>x\}\cap\{X_2>x\}\cap...\cap\{X_n>x\})=\mathbb{P}(\{X_1>x\})\mathbb{P}(\{X_2>x\}...\mathbb{P}(\{X_n>x\})$$

Then because the $X_i's$ have the same distribution, $\mathbb{P}(\{X_i>x\})$ is the same for each i, so

$$\mathbb{P}(\{X_1>x\})\mathbb{P}(\{X_2>x\}...\mathbb{P}(\{X_n>x\})=\mathbb{P}(\{X_1>x\})^n=(1-F_X(x))^n$$

Putting this all together:

$$F_{X_{(1)}}(x)=1-(1-F_X(x))^n$$

So then we differentiate to get the pdf:

$$f_{X_{(1)}}(x)=\frac{d}{dx}F_{X_{(1)}}(x)=-\frac{d}{dx}(1-F_X(x))^n=n(1-F_X(x))^{n-1}f_X(x)$$

where the last term comes from the chain rule.

□

Let's look at an example

EXAMPLE 2.4. *Let $X_1,X_2,...,X_n$ be independent exponential random variables with mean 1. Compute the distribution of $X_{(1)}=\min\{X_1,X_2,...,X_n\}$.*

Solution: The pdf of X_i is

$$f_{X_i}(x)=e^{-x}$$

if $x>0$ and 0 otherwise. So then we have for any $t>0$

$$F_X(t)=\int_0^t e^{-x}dx=1-e^{-t}.$$

From our formula above

$$f_{X_{(1)}}(t)=n(1-(1-e^{-t}))^{n-1}e^{-t}=ne^{-nt}$$

for $t>0$ and 0 otherwise.

Note that $X_{(1)}$ is an exponential random variable with mean $1/n$.