14.31/14.310 Lecture 7

Probability---functions of RVs

There are various methods one can use to figure out the distribution of a function of random variables. Which methods one can use on a particular problem depend on whether the original random variable is discrete or continuous, whether there is just one random variable or a random vector, and whether the function is invertible or not. We will not learn all of the methods here. Instead we'll learn one important method and also see a lot of examples that can be applied somewhat generally.

Probability---functions of RVs

X is a random variable with $f_X(x)$ known. We want the distribution of $Y = h(X)$. Then,

$$F_Y(y) = \int_{\{x:\, h(x) \le y\}} f_x(x)\, dx$$

If Y is also continuous, then

$$f_Y(y) = dF_Y(y)/dy$$

Probability---functions of RVs

X is a random variable with $f_X(x)$ known. We want the distribution of $Y = h(X)$. Then,

$$F_Y(y) = \int_{\{x:\, h(x) \leq y\}} f_X(x)\, dx$$

First, find the CDF by integrating over the appropriate region

If Y is also continuous, then

$$f_Y(y) = dF_Y(y)/dy$$

Probability---functions of RVs

X is a random variable with $f_X(x)$ known. We want the distribution of $Y = h(X)$. Then,

$$F_Y(y) = \int_{\{x:\, h(x) \le y\}} f_X(x)\,dx$$

First, find the CDF by integrating over the appropriate region

If Y is also continuous, then

$$f_Y(y) = dF_Y(y)/dy$$

Then take the derivative to find the PDF

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

Recall we need to "integrate over the appropriate region."

Easier said than done, perhaps, but we will argue in steps what is the appropriate region.

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

First note that the support of X is [-1,1], which implies that the induced support of Y is [0,1].

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

First note that the support of X is $[-1,1]$, which implies that the induced support of Y is $[0,1]$.

Remember this---we will use it again in a few slides.

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

$F_Y(y) = P(Y <= y)$ by definition

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

$F_Y(y) = P(Y <= y)$ by definition (first step)

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

$F_Y(y) = P(Y <= y)$ by definition

$= P(X^2 <= y)$ plugging in function

Probability---example

$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$

$Y = X^2$. What is f_Y?

$F_Y(y) = P(Y <= y)$ by definition

$= P(X^2 <= y)$ plugging in function

$= P(-\sqrt{y} \le X \le \sqrt{y})$ solving for X

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

$F_Y(y) = P(Y <= y)$ by definition

$= P(X^2 <= y)$ plugging in function

$= P(-\sqrt{y} \leq X \leq \sqrt{y})$ solving for X

$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\,dx$ integrating over appropriate area

Probability---example

$$f_X(x) = \begin{cases} 1/2 & \text{for } -1 <= x <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$Y = X^2$. What is f_Y?

$F_Y(y) = P(Y <= y)$ by definition

$= P(X^2 <= y)$ plugging in function

$= P(-\sqrt{y} \le X \le \sqrt{y})$ solving for X

$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\,dx$ integrating over appropriate area

$= \sqrt{y}$ for $0 <= y <= 1$

Probability---example

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ \sqrt{y} & \text{for } 0 <= y <= 1 \\ 1 & \text{for } y > 1 \end{cases}$$

Since Y is continuous, we can just take the derivative of F_Y to get f_Y.

$$f_Y(y) = \begin{cases} 1/(2\sqrt{y}) & \text{for } 0 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Probability---example

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ \sqrt{y} & \text{for } 0 <= y <= 1 \\ 1 & \text{for } y > 1 \end{cases}$$

This is where we use that fact we noted earlier

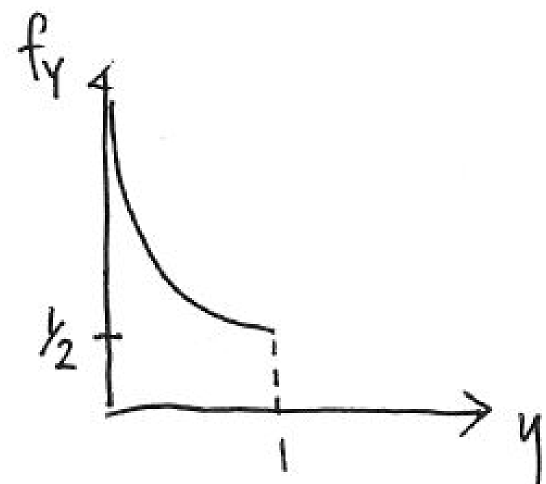
Since Y is continuous, we can just take the derivative of F_Y to get f_Y.

$$f_Y(y) = \begin{cases} 1/(2\sqrt{y}) & \text{for } 0 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Probability---example

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ \sqrt{y} & \text{for } 0 <= y <= 1 \\ 1 & \text{for } y > 1 \end{cases}$$

Since Y is continuous, we can just take the derivative of F_Y to get f_Y.



$$f_Y(y) = \begin{cases} 1/(2\sqrt{y}) & \text{for } 0 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Probability---important examples we'll see

1. Linear transformation of a single random variable
2. Probability integral transformation
3. Convolution
4. Order statistics

Probability---linear transformation

There may be lots of reasons why we care about the distribution of a linear transformation of a random variable. Perhaps the random variable is measured in the wrong or inconvenient units. (What's the distribution of the length of Steph Curry's shots in meters, instead of feet?) Perhaps some formula dictates a linear relationship between two variables, and we know how one is distributed. (The number of heating degree days in the month of February can be approximated as 28x(65-average high temp).) Perhaps some theory predicts a linear relationship between variables.

Probability---linear transformation

Let X have PDF $f_X(x)$. Let $Y = aX + b$, $a \neq 0$. How is Y distributed?

$F_Y(y) = P(Y <= y) = P(aX+b <= y)$

Probability---linear transformation

Let X have PDF $f_X(x)$. Let $Y = aX + b$, $a \neq 0$. How is Y distributed?

$$F_Y(y) = P(Y <= y) = P(aX+b <= y)$$

$$= \begin{cases} P(X <= (y-b)/a) & \text{if } a > 0 \\ P(X >= (y-b)/a) & \text{if } a < 0 \end{cases}$$

Probability---linear transformation

Let X have PDF $f_X(x)$. Let $Y = aX + b$, $a \neq 0$. How is Y distributed?

$$F_Y(y) = P(Y <= y) = P(aX+b <= y)$$

$$= \begin{cases} P(X <= (y-b)/a) & \text{if } a > 0 \\ P(X >= (y-b)/a) & \text{if } a < 0 \end{cases}$$

$$= \begin{cases} \int_{-\infty}^{y-b/a} f_x(x)\,dx & a > 0 \\ \int_{y-b/a}^{\infty} f_x(x)\,dx = 1 - \int_{-\infty}^{y-b/a} f_x(x)\,dx & a < 0 \end{cases}$$

Probability---linear transformation

Let X have PDF $f_X(x)$. Let $Y = aX + b$, $a \neq 0$. How is Y distributed?

So take the derivative to get the PDF:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} f_X\left(\frac{y-b}{a}\right)\frac{1}{a} & a > 0 \\ -f_X\left(\frac{y-b}{a}\right)\frac{1}{a} & a < 0 \end{cases}$$

Probability---linear transformation

Let X have PDF $f_X(x)$. Let $Y = aX + b$, $a \neq 0$. How is Y distributed?

So take the derivative to get the PDF:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} f_X\left(\frac{y-b}{a}\right)\frac{1}{a} & a > 0 \\ -f_X\left(\frac{y-b}{a}\right)\frac{1}{a} & a < 0 \end{cases}$$

In other words,

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

Probability---probability integral transformat^n

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let Y = $F_X(X)$. How is Y distributed?

Probability---probability integral transformat^n

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let Y = $F_X(X)$. How is Y distributed?

Strange that we would use a CDF, which describes the distribution of a random variable, to transform a random variable. But why not? It's a function.

Probability---probability integral transformat^n

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let $Y = F_X(X)$. How is Y distributed?

First note that, whatever the support of X, Y lives on $[0,1]$. Why?

Probability---probability integral transformat^n

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let Y = $F_X(X)$. How is Y distributed?

First note that, whatever the support of X, Y lives on [0,1]. Why? CDFs always have values between 0 and 1.

Probability---probability integral transformat^n

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let $Y = F_X(X)$. How is Y distributed?

First note that, whatever the support of X, Y lives on $[0,1]$. Why? CDFs always have values between 0 and 1.

Also note that F_X is invertible. (We noted earlier that F_X is non-decreasing. In fact, it will be invertible if X is continuous over a connected set.)

Probability---probability integral transformat[n]

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let $Y = F_X(X)$. How is Y distributed?

So $F_Y(y) = P(Y <= y) = P(F_X(X) <= y)$

$= P(X <= F_X^{-1}(y))$

$= F_X(F_X^{-1}(y))$

$= y \qquad 0 <= y <= 1$

Probability--probability integral transformat^n

Let X, continuous, have PDF $f_X(x)$ and CDF $F_X(x)$. Let $Y = F_X(X)$. How is Y distributed?

So $F_Y(y) = P(Y <= y) = P(F_X(X) <= y)$

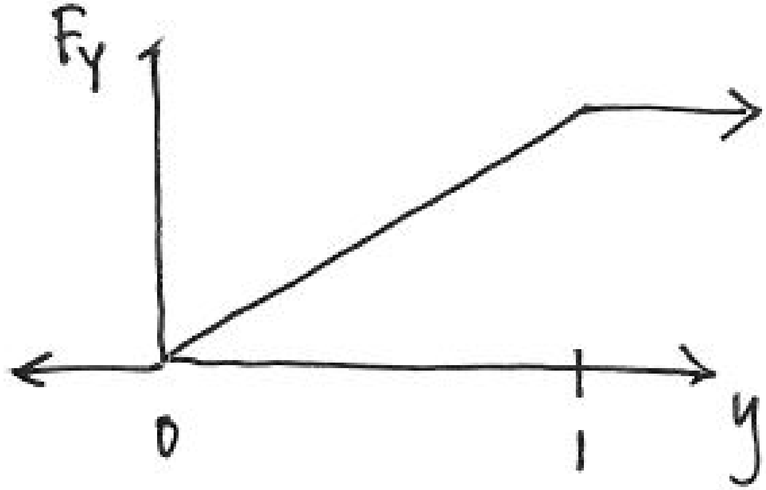
$= P(X <= F_X^{-1}(y))$

$= F_X(F_X^{-1}(y))$

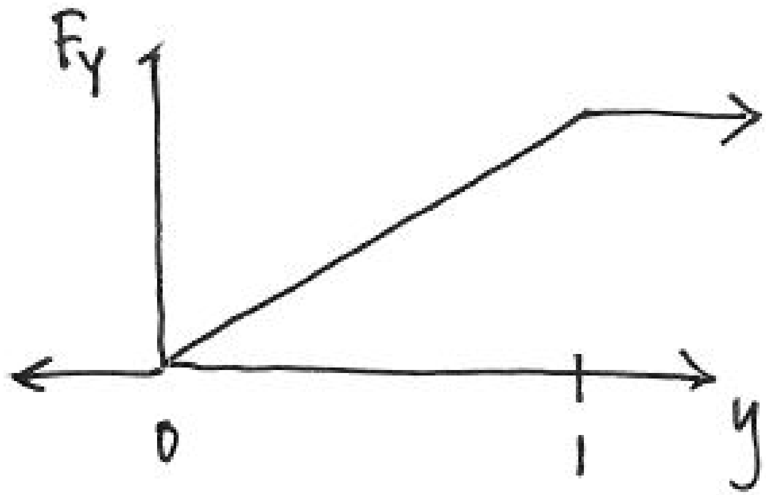
$= y \qquad 0 <= y <= 1$

What random variable has a CDF that looks like that?

Probability---probability integral transformatn



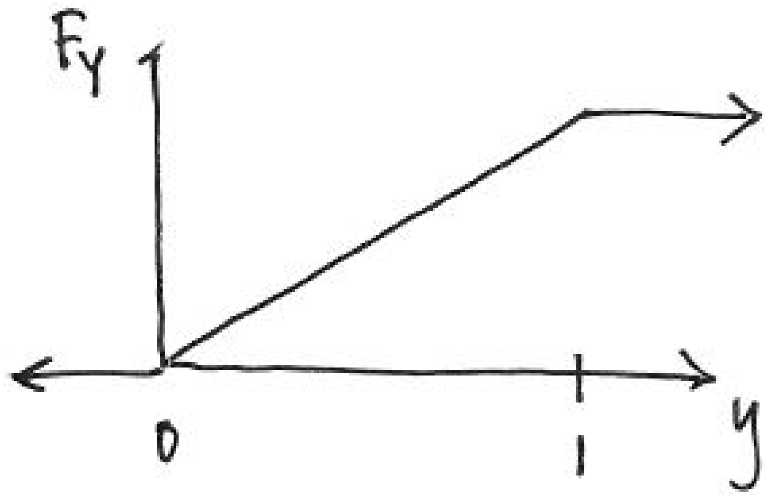
Probability---probability integral transformat^n



A U[0,1] random variable!

So a continuous random variable transformed by its own CDF will always have a U[0,1] distribution.

Probability--probability integral transformatn



A U[0,1] random variable!

So a continuous random variable transformed by its own CDF
will always have a U[0,1] distribution.

Pretty cool.

Probability---probability integral transformat^n

How about the other way? Can we transform a U[0,1] random variable by the inverse of a CDF and get a random variable with that CDF?

Probability--probability integral transformat^n

How about the other way? Can we transform a U[0,1] random variable by the inverse of a CDF and get a random variable with that CDF?

Yes! (assuming the random variable is continuous and meets certain regularity conditions)

Probability--probability integral transformat^n

How about the other way? Can we transform a U[0,1] random variable by the inverse of a CDF and get a random variable with that CDF?

Yes! (assuming the random variable is continuous and meets certain regularity conditions)

Again, pretty cool.

Probability---probability integral transformat^n

Interesting, perhaps, but how could this be useful?

Probability---probability integral transformat^n

Interesting, perhaps, but how could this be useful?

One example: performing computer simulations

Probability---probability integral transformat^n

Suppose we were writing a computer program to simulate, say, the spread of some virus over time in a school population. To perform the simulation, we would need random draws from a uniform distribution to model the proportion of the school population that was infected initially, random draws from an exponential distribution to model the physical proximity of children during a PE class, and random draws from a beta distribution to model humidity inside the school on different days.

Probability---probability integral transformat^n

Suppose we were writing a computer program to simulate, say, the spread of some virus over time in a school population. To perform the simulation, we would need random draws from a uniform distribution to model the proportion of the school population that was infected initially, random draws from an exponential distribution to model the physical proximity of children during a PE class, and random draws from a beta distribution to model humidity inside the school on different days.

But the computer language you were using only generated random draws from U[0,1].

Probability---probability integral transformat^n

Note that random number generators, tables of random digits, and many other sources of random (and pseudo-random) numbers are giving you *uniform* random numbers.

Probability---probability integral transformat^n

Note that random number generators, tables of random digits, and many other sources of random (and pseudo-random) numbers are giving you *uniform* random numbers.

So, if you knew (or could look up) the CDFs of exponential and beta random variables, you could compute the inverses of those CDFs and then use those functions to transform the random draws from the U[0,1] into random draws from exponential and beta distributions.

Probability---convolution

A convolution in the context of probability refers to the sum of independent random variables. We have already seen one example where we cared about the sum of independent random variables (although we didn't know they were independent at the time).

Probability---convolution

A convolution in the context of probability refers to the sum of independent random variables. We have already seen one example where we cared about the sum of independent random variables (although we didn't know they were independent at the time)---the headache example. We were interested in the sum there because I could take the two pills sequentially, so the distribution of the sum of their effective lives was of interest.

Probability---convolution

A convolution in the context of probability refers to the sum of independent random variables. We have already seen one example where we cared about the sum of independent random variables (although we didn't know they were independent at the time)---the headache example. We were interested in the sum there because I could take the two pills sequentially, so the distribution of the sum of their effective lives was of interest.

Such questions can arise in many contexts: the total value of two investments, the total number of successes in two independent sets of trials, etc.

Probability---convolution

Convolutions generalize naturally in two ways:

- sum of N, not 2, independent random variables
- linear function of independent random variables

Probability---convolution

Convolutions generalize naturally in two ways:

- sum of N, not 2, independent random variables
- linear function of independent random variables

We'll do the simple version, sum of two independent random variables.

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

We will proceed similarly to the headache example.

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

We will proceed similarly to the headache example.

One difference: in the headache example, we were given the joint PDF and here we're not.

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

We will proceed similarly to the headache example.

One difference: in the headache example, we were given the joint PDF and here we're not. But we can easily get the joint PDF because we know the random variables are independent: $f_{XY}(x,y) = f_X(x)f_Y(y)$

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

Recall that, in the headache example, we just set up the double integral to get the P(X+Y <= z), i.e., the CDF of Z, and then took the derivative of that to get the PDF.

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

Recall that, in the headache example, we just set up the double integral to get the P(X+Y <= z), i.e., the CDF of Z, and then took the derivative of that to get the PDF.

That method works, as well as some others.

Probability---convolution

Let X be continuous with PDF f_X, Y continuous with PDF f_Y. X and Y are independent. Let Z be their sum. What is the PDF of Z?

Recall that, in the headache example, we just set up the double integral to get the P(X+Y <= z), i.e., the CDF of Z, and then took the derivative of that to get the PDF.

That method works, as well as some others.

So, we get $F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x) f_Y(y)\,dx\,dy$

Probability---convolution

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x) f_Y(y)\, dx\, dy$$

So $$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\, dy \qquad -\infty < z < \infty$$

Probability---order statistics

I told you that the uniform was my favorite distribution. Well, order statistics are my favorite function of random variables. If that's not enough motivation for you, keep in mind that order statistics can be very useful in economic modeling (we'll see an example in auctions) and they also are the basis for some important estimators.

Probability---order statistics

Let $X_1, \ldots, X_n$ be continuous, independent, identically distributed, with PDF f_X. (We often abbreviate "independent, identically distributed" as "i.i.d." A group of i.i.d. random variables is also called a random sample.) Let $Y_n = \max\{X_1, \ldots, X_n\}$. This is called the n^{th} order statistic.

Probability---order statistics

Let $X_1, \ldots, X_n$ be continuous, independent, identically distributed, with PDF f_X. (We often abbreviate "independent, identically distributed" as "i.i.d." A group of i.i.d. random variables is also called a random sample.) Let $Y_n = \max\{X_1, \ldots, X_n\}$. This is called the n^{th} order statistic. (We can also define the 1^{st} order statistic as the smallest value, the 2^{nd} order statistic as the second smallest value, and so forth.)

Probability---order statistics

Let $X_1, \ldots, X_n$ be continuous, independent, identically distributed, with PDF f_X. (We often abbreviate "independent, identically distributed" as "i.i.d." A group of i.i.d. random variables is also called a random sample.) Let $Y_n = \max\{X_1, \ldots, X_n\}$. This is called the n^{th} order statistic.

How is the n^{th} order statistic distributed?

Probability---order statistics

How is the n^{th} order statistic distributed?

$$F_n(y) = P(Y_n <= y) = P(X_1 <= y, X_2 <= y, \ . \ . \ . \ . \ , X_n <= y)$$

by definition of Y_n

Probability---order statistics

How is the n^{th} order statistic distributed?

$$F_n(y) = P(Y_n <= y) = P(X_1 <= y, X_2 <= y, \ldots, X_n <= y)$$

by definition of Y_n

$$= P(X_1 <= y)P(X_2 <= y) \ldots P(X_n <= y)$$

due to independence

Probability---order statistics

How is the n^{th} order statistic distributed?

$F_n(y) = P(Y_n <= y) = P(X_1 <= y, X_2 <= y, \ldots, X_n <= y)$

by definition of Y_n

$= P(X_1 <= y)P(X_2 <= y) \ldots P(X_n <= y)$

due to independence

$= F_X(y)^n$

due to identical distribution

Probability---order statistics

How is the n^{th} order statistic distributed?

$F_n(y) = P(Y_n <= y) = P(X_1 <= y, X_2 <= y, \ldots, X_n <= y)$

by definition of Y_n

$= P(X_1 <= y)P(X_2 <= y) \ldots P(X_n <= y)$

due to independence

$= F_X(y)^n$

due to identical distribution

So, $f_n(y) = dF_n(y)/dy = n(F_X(y))^{n-1} f_X(y)$

Probability---order statistics

How is the 1st order statistic distributed?

A similar calculation will lead to this:

$$f_1(y) = n(1-F_X(y))^{n-1} f_X(y)$$

Probability---order statistics

So we have the following:

$$f_n(y) = n(F_X(y))^{n-1} f_X(y)$$

$$f_1(y) = n(1-F_X(y))^{n-1} f_X(y)$$

What do these distributions look like if we have a random sample from, say, a U[0,1] distribution?

Probability---order statistics

So we have the following:

$$f_n(y) = n(F_X(y))^{n-1} f_X(y)$$

$$f_1(y) = n(1-F_X(y))^{n-1} f_X(y)$$

What do these distributions look like if we have a random sample from, say, a U[0,1] distribution? Depends on n.

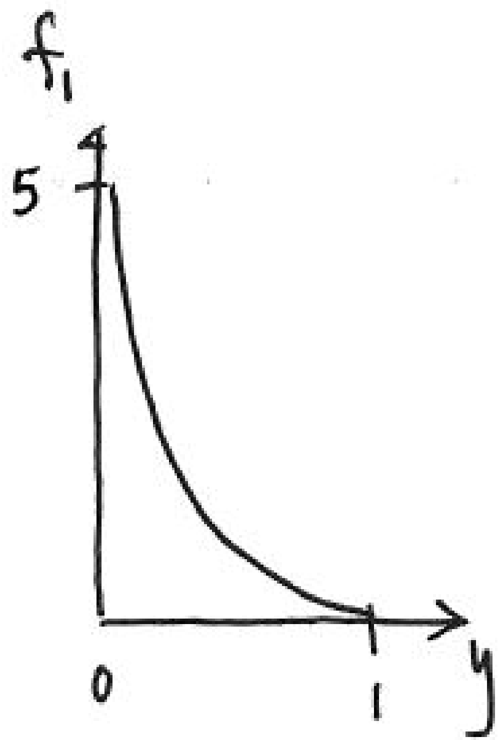
For n = 5: $f_n(y) = 5y^4$ $\quad$ $0 <= y <= 1$

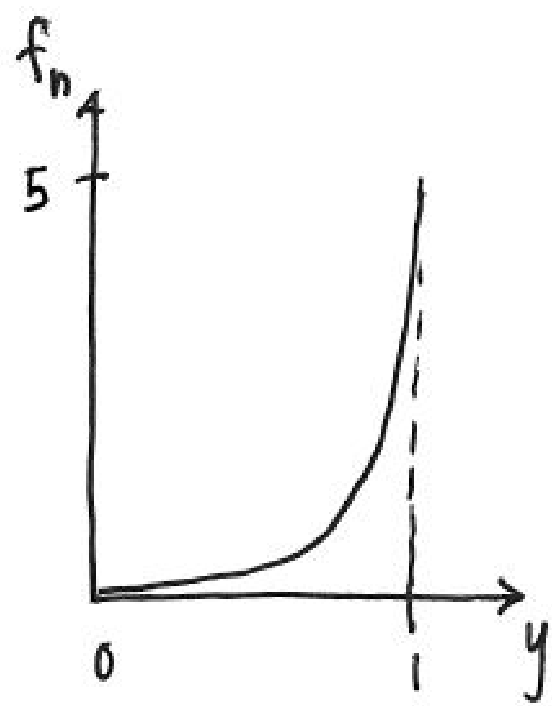
$f_1(y) = 5(1-y)^4$ $\quad$ $0 <= y <= 1$

Probability---order statistics

$f_1(y) = 5(1-y)^4 \qquad 0 <= y <= 1$

$f_n(y) = 5y^4 \qquad 0 <= y <= 1$



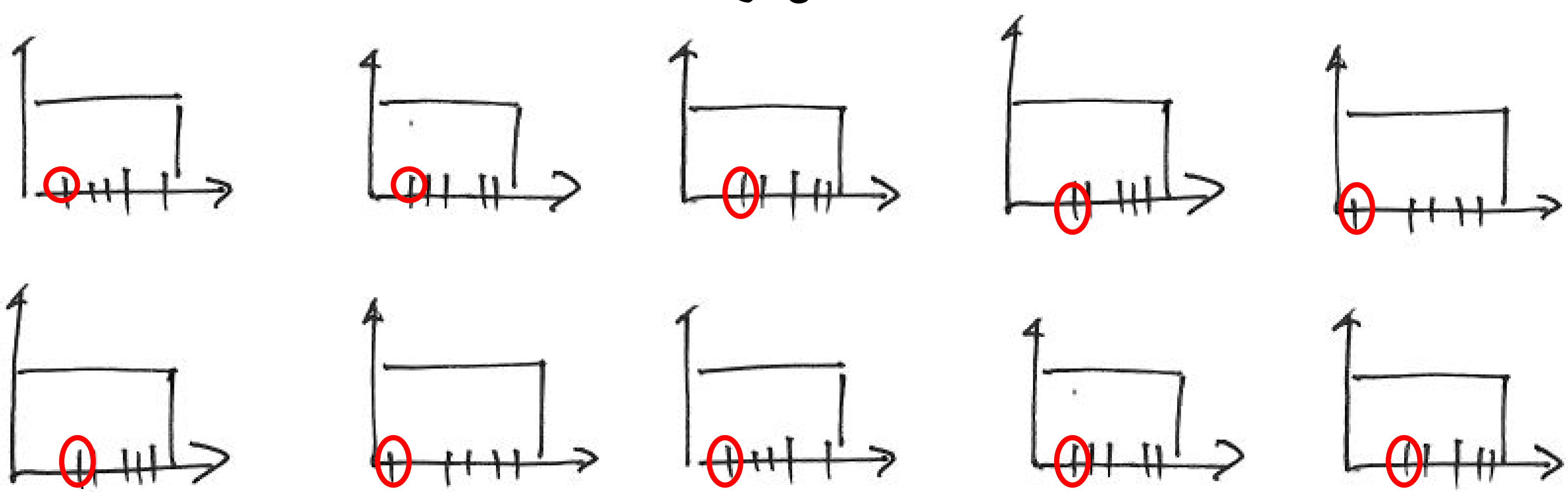


Probability---order statistics

Think of it like this:

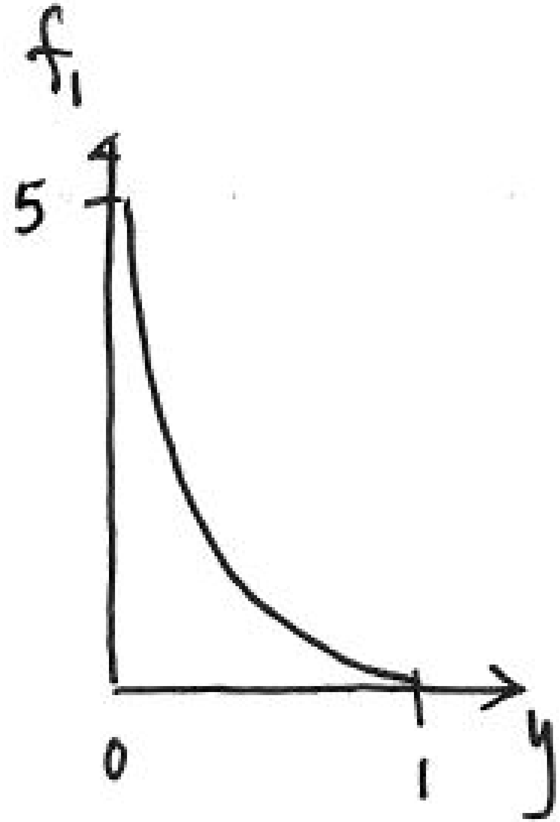
You have a random sample of size 5 from a U[0,1] distribution. How is the smallest realization from that random sample distributed?

What is the PDF of these guys?



Probability---order statistics

You'll get something with the same support, [0,1], but with probability concentrated near 0.

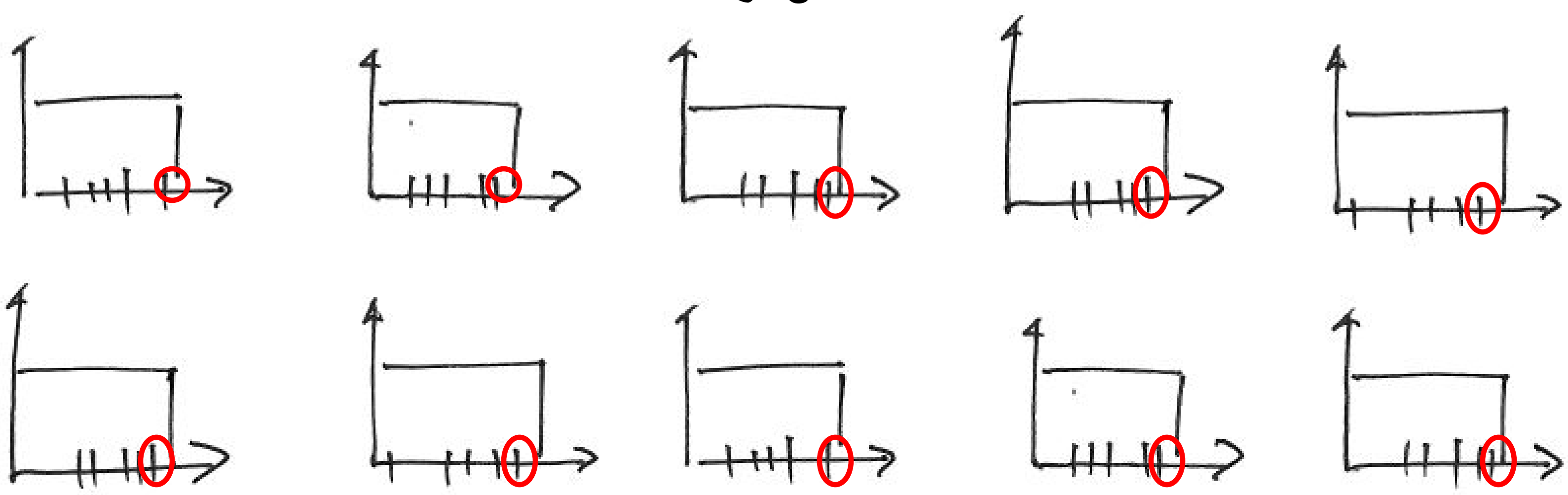


Probability---order statistics

Think of it like this:

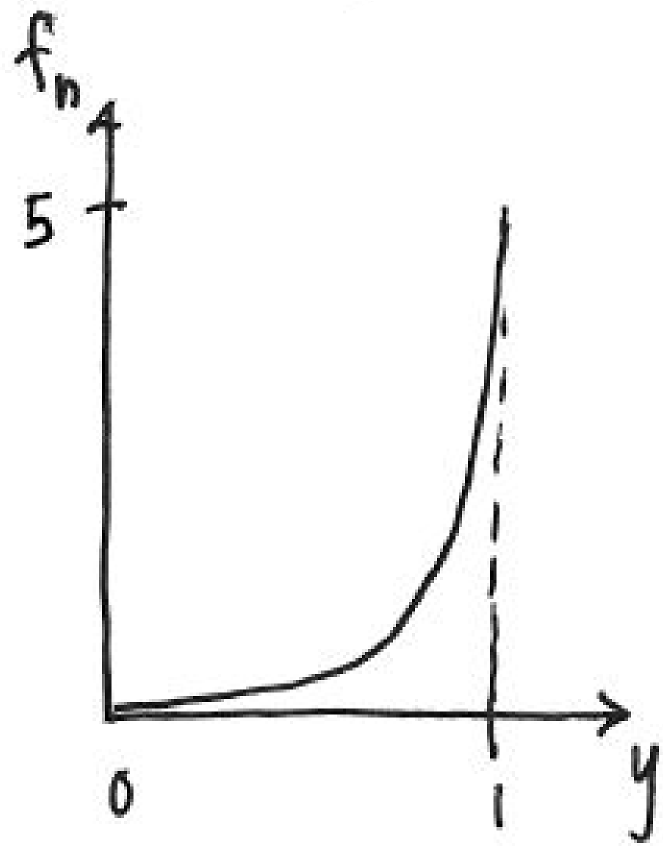
You have a random sample of size 5 from a U[0,1] distribution. How is the largest realization from that random sample distributed?

What is the PDF of these guys?



Probability---order statistics

You'll get something with the same support, [0,1], but with probability concentrated near 1.

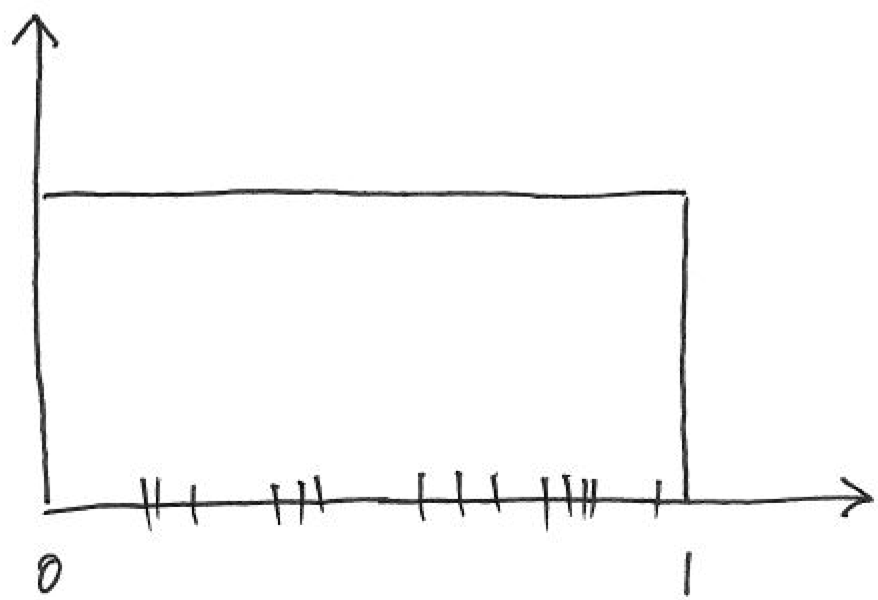


Probability---order statistics

What if n is larger then 5?

Probability---order statistics

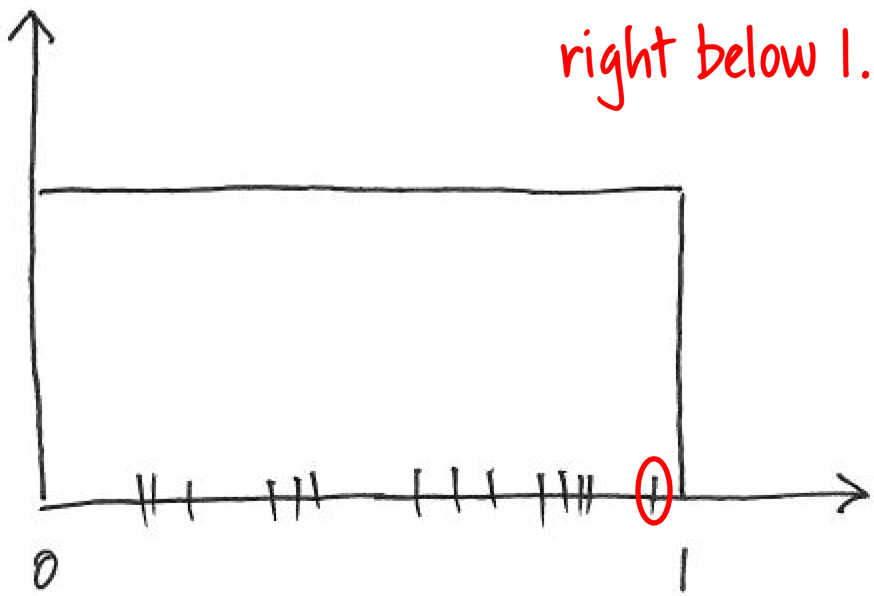
What if n is larger then 5?



Probability---order statistics

What if n is larger then 5?

This guy is more likely to be near 1---its distribution will be more concentrated right below 1.

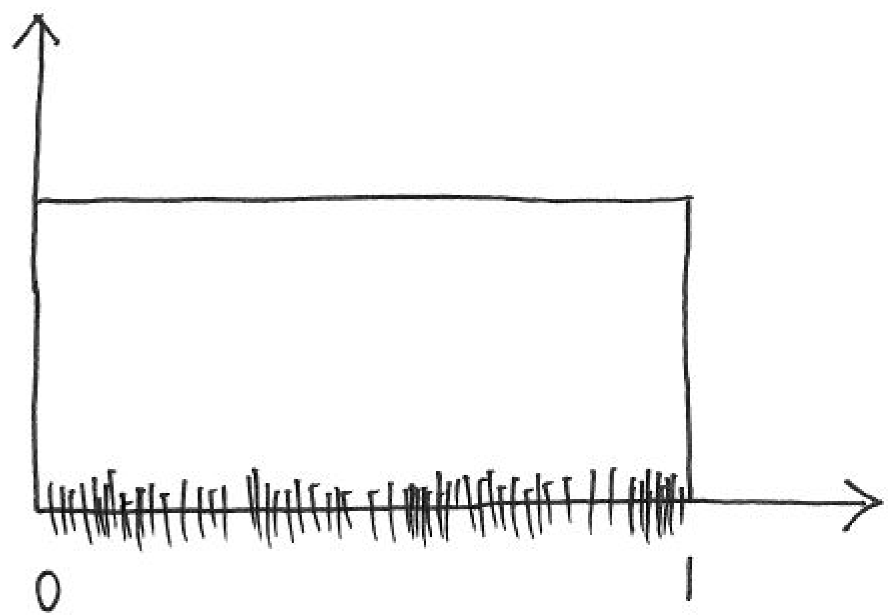


Probability---order statistics

What if n is really large?

Probability---order statistics

What if n is really large?



Probability---order statistics

What if n is really large?

This guy is even more likely to be near 1---its distribution will be even more concentrated right below 1.