Functions of a random variable

April 24, 2020

A few times in this class we've considered functions of random variables (for example X^2), but typically we have just computed the expectation of this (for example $\mathbb{E}[X^2]$). This was useful because the moments of random variable give us information about the original random variable. But sometimes we might be interested in $g(X)$, itself, for some function g. In these cases we would like to actually compute the pdf/pmf of $g(X)$. In Chapter 6, we'll discuss several ways to do this.

In these notes we will compute the pdf of $g(Y)$ for $g:\mathbb{R}\to\mathbb{R}$ and Y a random variable.

In later sections we will also consider the pdf of $g(X,Y)$ for $g:\mathbb{R}^2\to\mathbb{R}$ and X,Y random variables.

The book organizes these in a different order than I will.

1 Functions of Discrete Random variables

We'll focus very little on functions of Discrete Random variables, because it doesn't end up being so interesting. But we'll start with an example because it's a little bit easier to understand than the continuous case.

Example 1.1. *Let X be a discrete random variable with pmf:*

$$p_X(-1)=.2 \quad p_X(0)=.1 \quad p_X(1)=.4 \quad p_X(2)=.3.$$

Compute the pmf of X^2.

Solution: The random variable X can take the values $-1,0,1,2$, so the random variable X^2 can take the values $0^2,1^2,2^2$ (note we used that $1^2=(-1)^2$). Then we can compute the probability it takes on each of these values:

$$\mathbb{P}(X^2=0)=\mathbb{P}(X=0)=.1$$

$$\mathbb{P}(X^2=1)=\mathbb{P}(X=-1)+\mathbb{P}(X=1)=.2+.4=.6$$

$$\mathbb{P}(X^2=4)=\mathbb{P}(X=2)=.3$$

So the pmf is

$$p_{X^2}(0)=.2 \quad p_{X^2}(1)=.6 \quad p_{X^2}(4)=.3$$

This example shows the general idea of computing the pmf of $g(X)$ for some function X. First you determine the values $g(X)$ can take, by applying the function g to the values X can take. Then to determine the pmf of $g(X)$ at some value y, add up the probabilities that X takes on each value that is mapped by g to y. Mathematically, we're adding up the probabilities X takes a value in the inverse image of y, which we label $g^{-1}(y)$.

2 Functions of Continuous Random variables

We'll most focus on the continuous case, which you can think of similar to the discrete case, but we cannot merely compute the probability a continuous random variable takes a fixed value, because this is always 0. We instead will work with the CDF.

2.1 Linear Functions

We have actually briefly seen a special case of computing the pdf of $g(X)$, when $g(x)=ax+b$ is a linear function. We used this idea to compute the expectation and variance of normal and exponential random variables. Let's look at this example again.

EXAMPLE 2.1. *Let X be a random variable with pdf $f_X(x)$. Let a,b be real numbers. What is the pdf of $aX+b$?*

Solution: We'll actually compute the CDF of $aX+b$, then differentiate it to get the pdf. The derivation is slightly different if a is positive or negative. Let's assume $a>0$ first.

$$\begin{aligned} F_{aX+b}(t) &= \mathbb{P}(aX+b \le t) \\ &= \mathbb{P}\left(X \le \frac{t-b}{a}\right) \\ &= F_X\left(\frac{t-b}{a}\right) \end{aligned}$$

The first line is the definition of the CDF. In the second line we solve for X. In the third line, we recognize the previous line was actually the CDF of X evaluated at $\frac{t-b}{a}$.

Now we differentiate:

$$f_{aX+b}(t) = \frac{d}{dt}F_{aX+b}(t) = \frac{d}{dt}F_X\left(\frac{t-b}{a}\right) = f_X\left(\frac{t-b}{a}\right)\frac{1}{a}$$

where the last equality uses the chain rule.

Now we consider the case $a<0$:

$$\begin{aligned} F_{aX+b}(t) &= \mathbb{P}(aX+b \le t) \\ &= \mathbb{P}\left(X \ge \frac{t-b}{a}\right) \\ &= 1 - F_X\left(\frac{t-b}{a}\right) \end{aligned}$$

The difference is when we solve for X, we are dividing by a negative number so we flip the inequality. Then we can differentiate as before:

$$f_{aX+b}(t) = \frac{d}{dt} F_{aX+b}(t) = \frac{d}{dt}\left(1 - F_X\left(\frac{t-b}{a}\right)\right) = -f_X\left(\frac{t-b}{a}\right)\frac{1}{a}$$

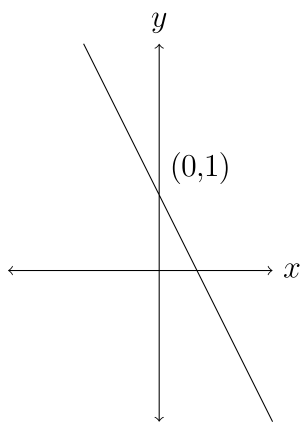
We can combine these two cases as:

$$f_{aX+b}(t) = \frac{1}{|a|} f_X\left(\frac{t-b}{a}\right)$$

EXAMPLE 2.2. *Let X be an exponential random variable with mean 10. What is the pdf of $-2X+1$?*

Solution: The first thing we should do is determine where the support of the pdf of $-2X+1$ is positive. We do this by applying the function $-2x+1$ to the set where the pdf of X is positive. In this case the pdf of X is positive for $x>0$, so the pdf of $-2X+1$ will be positive for when evaluated at values less than 1.

To determine this region is it useful to draw the graph of $-2x+1$, we then look at what y values are obtained when $x>0$ (the set where the pdf of X is positive)



On the picture, we want to look at what values of the y-axis are hit, when we evaluate $-2x+1$ at $x>0$. The set $x>0$ is chosen because that is where the pdf of X is positive. From the picture, we see each y values less than 1 will be hit. So the pdf of $-2X+1$ is positive when evaluated at points less than 1.

We can then follow the above derivation line by line with the explicit formula

$$f_X(x) = \frac{1}{10} e^{-x/10}$$

for $x>0$ and 0 otherwise, or you can just apply the formula to get:

$$f_{-2X+1}(t) = \frac{1}{|-2|} f_X\left(\frac{t-1}{-2}\right) = \frac{1}{20} e^{(t-1)/20}$$

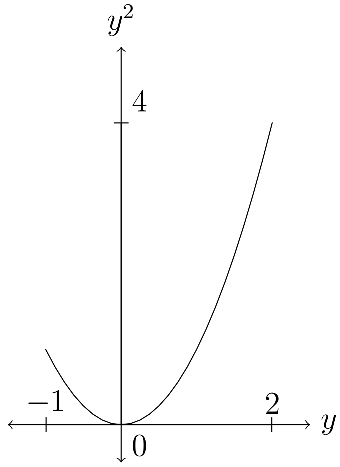
for $t<1$ and 0 otherwise.

2.2 General Functions

The computation of the pdf of $g(X)$ in the general case doesn't follow quite as nicely, so it's best to start with an example. The general steps will look similar to the linear function case, but with some new complications.

EXAMPLE 2.3. *Let Y be a uniformly distributed random variable on the interval $[-1,2]$. Compute the pdf of Y^2.*

Solution: We start be determining the where the pdf of Y^2 will be positive. To do this is is useful to graph the function y^2 on the set $[-1,2]$.

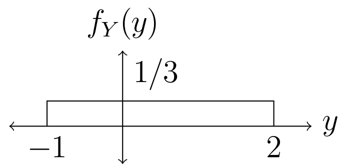


So we see the pdf of Y^2 will be positive between 0 and 4. Furthermore, we see two points get mapped to values between 0 and 1 (namely the values between -1 and 1) and one point point maps to the values between 1 and 4. When we compute the CDF we will notice this.

Before we compute the CDF of Y^2, note that the pdf of Y is

$$f_Y(y) = \frac{1}{3}$$

for $-1 \le y \le 2$, and 0 otherwise. The picture is:



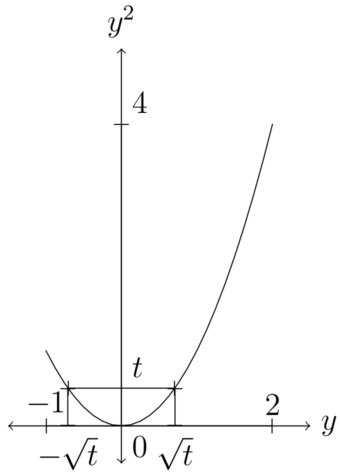
The CDF is:

$$F_{Y^2}(t) = \mathbb{P}(Y^2 \le t)$$
$$= \mathbb{P}(-\sqrt{t} \le Y \le \sqrt{t}) \tag{2.1}$$

if $0 \le t \le 1$ then

$$\mathbb{P}(-t \le Y \le t) = \int_{-\sqrt{t}}^{\sqrt{t}} \frac{1}{3} dy = \frac{2\sqrt{t}}{3}$$

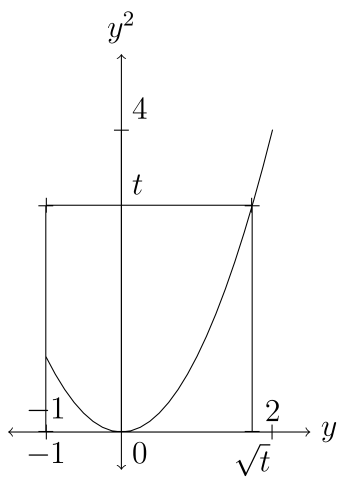
The picture to determine the bounds is:



if $1 \le t \le 4$ then

$$\mathbb{P}(-t \le Y \le t) = \int_{-1}^{\sqrt{t}} \frac{1}{3} dy = \frac{\sqrt{t}}{3} + \frac{1}{3}$$

The picture to determine the bounds is:



Now we put the combine the last two computations to get

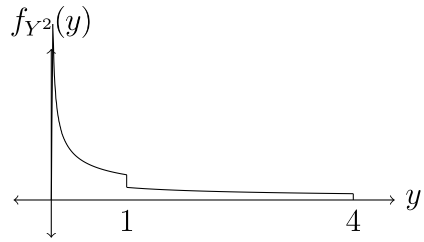
$$f_{Y^2}(t) = \frac{d}{dt} F_{Y^2}(t)$$

So differentiating the two expressions above we get:

$$f_{Y^2}(t) = \begin{cases} \frac{1}{3\sqrt{t}}, & \text{if } 0 \le t \le 1, \\ \frac{1}{6\sqrt{t}}, & \text{if } 1 \le t \le 4, \\ 0, & \text{otherwise.} \end{cases}$$

You can check that this is in fact a pdf (it is positive and integrates to 1). In general it is a good idea to check your work and integrate your final answer, making sure it integrates to 1. If not, maybe you forget to integrate over a certain set.

The graph of the pdf of Y^2 is:



In general idea to compute the pdf of $g(Y)$ when the pdf of Y is known you start with the CDF of $g(Y)$ and try to write in terms of probabilities involving Y. This was done in (2.1). In general is this the hard part of the computation, where you need to be careful to not leave any terms out. Once you have probabilities involving just Y, use the pdf of Y to integrate and compute these probabilities. Then differentiate the CDF to get the pdf.

Mathematically this looks something like

$$F_{g(Y)}(t) = \mathbb{P}(g(Y) \leq t) = \int_{y:g(y)\leq t} f_Y(y)dy$$

$$f_{g(Y)}(t) = \frac{d}{dt} F_{g(Y)}(t)$$

Often the most work goes into determining the set $\{y : g(y) \leq t\}$, (this is notation for y such that $g(y)$ is less than or equal to t).

2.3 Monotonic Functions

We conclude with a special case of functions g, I will always assume the function g is differentiable, which will be a fairly mild assumption. If the derivative g' is positive then g is increasing and if the derivative g' is negative then g is decreasing. In either case, there exist an inverse function (denoted g^{-1}) such that

$$g(g^{-1}(x)) = g^{-1}(g(x)) = x.$$

The linear functions above are monotonic (increasing if $a > 0$ and decreasing is $a < 0$). The function y^2 is not monotonic around 0, and therefore there does not exist an inverse function ($-y$ and y get mapped to the same point).

In general, for monotonic functions the pdf can actually be computed similar to in the linear function case:

If $g' > 0$ then

$$F_g(Y)(t) = \mathbb{P}(g(Y) \leq t) = \mathbb{P}(Y \leq g^{-1}(t)) = F_Y(g^{-1}(t))$$

and if $g' < 0$ then

$$F_g(Y)(t) = \mathbb{P}(g(Y) \leq t) = \mathbb{P}(Y \geq g^{-1}(t)) = 1 - F_Y(g^{-1}(t))$$

the difference between positive and negative derivatives is same as in the linear function case.

Then differentiating we have

$$f_{g(Y)}(t) = f_Y(g^{-1}(t)) \left| \frac{d}{dt} g^{-1}(t) \right|$$

where the second term comes from the chain rule. The absolute value takes care accounts for the case with this derivative is negative.

Note that we don't require the derivative of g to always have the same sign, we only need it to have the same sign for values where the pdf of Y is positive.

Let's do an example

EXAMPLE 2.4. *A circle is drawn with a random radius. The distribution of the radius is an exponential random variable with mean 1.*

What is the pdf the of the area this circle.

Solution: Recall the area of a circle is with radius r is πr^2. So we define a random variable $A = \pi R^2$ where R is an exponential random variable with mean 1. The pdf of R is

$$f_R(r) = e^{-r}$$

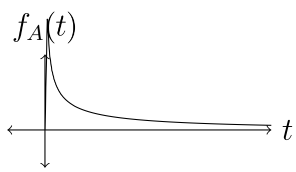
for $r > 0$ and 0 otherwise.

The exponential random variable only takes positive values, for positive r the derivative of πr^2 always positive so we can apply the simplified formula above.

Our function is $g(r) = \pi r^2$ so $g^{-1}(t) = \sqrt{t/\pi}$.

$$f_A(t) = f_R\left(\sqrt{\frac{t}{\pi}}\right)\left|\frac{d}{dt}\left(\sqrt{\frac{t}{\pi}}\right)\right| = e^{-\sqrt{t/\pi}}\frac{1}{2\sqrt{\pi t}}$$

for $t > 0$ and 0 otherwise.

The graph of this new pdf is: 

The formula

$$f_{g(Y)}(t) = f_Y(g^{-1}(t))\left|\frac{d}{dt}g^{-1}(t)\right|$$

is not so easy to remember. But if you know the technique in the general case, and remember that increasing/decreasing functions have an inverse, then you can derive this formula readily from the general argument.