4.7: Transformations of Random Variables

Taylor

University of Virginia

Motivation

We transform random variables all the time. So far we talked about how to find exact and approximate expressions for the means and variances of transformed rvs in the linear and non-linear case, respectively. We've talked about how mgfs can help us finding the exact distribution (not just moments) of linear transformations of special rvs. And in the last section we kind of alluded to what we're going to do now.

Motivation

Here was the situation from last time. Let $X \sim f(x)$, and let $Y = cX$, with $c > 0$. What is $f(y)$?

We did this, we basically found the cdf easily, and then differentiated to find the density

$$F_Y(y) = P(Y \le y) = P(cX \le y) = P(X \le y/c) = F_X(y/c)$$

so $f_Y(y) = f_X(y/c)\frac{1}{c}$ if $Y = cX$

Motivation

Let's generalize this to monotonic function (or 1-1, so it'll have an inverse, which is what we need). Let $Y = g(X)$. I'm going to call this the original transformation. So the inverse transformation is $X = g^{-1}(Y)$

When $g(\cdot)$ is increasing...

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

and when $g(\cdot)$ is decreasing...

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

Motivation

When $g(\cdot)$ is increasing...

$$\frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(g^{-1}(y)) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)$$

and when $g(\cdot)$ is decreasing...

$$\frac{d}{dy}F_Y(y) = -\frac{d}{dy}F_X(g^{-1}(y)) = f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right|$$

(because the negative sign cancels out with the negative slope of $\frac{d}{dy}g^{-1}(y)$)

Note

Note: you only need this for cts distributions. If your rv is discrete it's way simpler...

$$p_Y(y) = P(Y = y) = P(g(X) = y) = P(X = g^{-1}(y)) = p_X(g^{-1}(y))$$

you just plug in $g^{-1}(y)$ everywhere you see an x in the pmf of X

Note

Finally, here's a statement of the transformation theorem.

Theorem

Let Y and X be two cts rvs. If $Y = g(X)$, where $g(\cdot)$ is a monotonic, and differentiable function, then

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|$$

Example

We touched on lognormal distributions a bit before. Now let's find its density.

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Let $Y = \exp[X]$. Find $f_Y(y)$

Example

We touched on lognormal distributions a bit before. Now let's find its density.

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Let $Y = \exp[X]$. Find $f_Y(y)$

$$f_Y(y) = f_X(\log y)\left|\frac{1}{y}\right| = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\log y - \mu)^2}{2\sigma^2}\right]\frac{1}{y}$$

we can drop the absolute value because $y = e^x > 0$ for any x

Example

Earlier we used the Geometric random variable $X \in \{1, 2, \ldots\}$. It's pmf was $p(x) = p(1-p)^{x-1}$. In this case X represents the *total number of births* observed before a success. Or it's the *trial number on which we stopped*.

Define $Y = X - 1$. What's the pmf for Y?

Example

Earlier we used the Geometric random variable $X \in \{1, 2, \ldots\}$. It's pmf was $p(x) = p(1-p)^{x-1}$. In this case X represents the *total number of births* observed before a success. Or it's the *trial number on which we stopped*.

Define $Y = X - 1$. What's the pmf for Y?

$$p_Y(y) = p_X(x(y)) = p(1-p)^{(y+1)-1} = p(1-p)^y$$

Sometimes people also call Y a Geometric random variable. But in this case it represents *the number of failures* before the success. It's really annoying, and unncessarily confusing. So when people say 'geometric rv,' you should ask for clarification.