

MIT OpenCourseWare  
<http://ocw.mit.edu>

14.30 Introduction to Statistical Methods in Economics  
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

# 14.30 Introduction to Statistical Methods in Economics

## Lecture Notes 9

Konrad Menzel

March 10, 2009

### 1 Functions of Random Variables

In this part of the lecture we are going to look at functions of random variables,  $Y = u(X)$ . Note that  $Y$  is again a random variable: since  $X$  is a mapping from the sample space  $S$  into the real numbers,

$$X : S \rightarrow \mathbb{R}$$

and  $u : \mathbb{R} \rightarrow \mathbb{R}$ , the composition of  $u$  and  $X$  is also a mapping from  $S$  into the real numbers:

$$Y = u \circ X : S \rightarrow \mathbb{R}$$

**Example 1** *If  $X$  is the life of the first spark plug in a lawnmower, and  $Y$  the life of the second, we may be interested in the sum of the two,  $Z = X + Y$ .*

**Example 2** *Before coming to MIT, I applied for several German fellowships, so I would receive a monthly stipend of  $X$  Euros, depending on which fellowship I was going to get, and the exchange rate in, say, September 2005 was going to be  $Y$  Dollars per Euro. Each quantity was uncertain at the time I was applying, but since I was going to spend the money in the US, the main quantity of interest was the dollar amount  $Z = X * Y$  I was going to receive (at least in terms of the Dollar exchange rate, I could not complain).*

We now want to know how to obtain the density and c.d.f. for the transformed random variable  $u(X)$ , so we can treat each problem involving a *function* of a random variable in the same way as a question involving only the random variable itself with a known p.d.f.

We'll consider three cases:

1. underlying variable  $X$  discrete
2. underlying variable  $X$  continuous
3.  $X$  continuous *and*  $u(X)$  strictly increasing

The last case is of course a special case of the second, but we'll see that it's much easier to work with.

## 1.1 Discrete Case - "2-Step" Method

If  $X$  is a discrete random variable with p.d.f.  $f_X(x)$ , and  $Y = u(X)$ , where  $u(\cdot)$  is a known deterministic function. Then

$$f_Y(y) = P(Y = y) = P(u(X) = y) = \sum_{x:u(x)=y} f_X(x)$$

### Example 3

$$f_X(x) = \begin{cases} \frac{1}{5} & \text{if } x \in \{-2, -1, 0, 1, 2\} \\ 0 & \text{otherwise} \end{cases}$$

Then if  $Y = g(X) = |X|$ ,

$$f_Y(y) = \begin{cases} f_X(0) = \frac{1}{5} & \text{if } y = 0 \\ f_X(-1) + f_X(1) = \frac{2}{5} & \text{if } y = 1 \\ f_X(-2) + f_X(2) = \frac{2}{5} & \text{if } y = 2 \\ 0 & \text{otherwise} \end{cases}$$

Note that if  $X$  is discrete, then  $Y$  is also discrete.

## 1.2 Continuous Case - "2-Step" Method

If  $X$  is a continuous random variable with p.d.f.  $f_X(x)$ , and  $Y = u(X)$ , then the c.d.f. of  $Y$  is given by

$$F_Y(y) = P(Y \leq y) = P(u(X) \leq y) = \int_{x:u(x) \leq y} f_X(x) dx$$

If  $Y$  is also continuous, then

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$

Note that even if  $X$  is continuous,  $Y$  need not be continuous.

**Example 4**  $Y = \lceil X \rceil$ , the largest integer smaller than  $X$ , is discrete regardless whether  $X$  is continuous or discrete.

### Example 5

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Let's look at

$$Y = X^2$$

From  $X \in [-1, 1]$ , it follows that  $Y = [0, 1]$ . How do we get the density of  $Y$ ? For  $y \in [0, 1]$ , the c.d.f. is

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2} dx = \sqrt{y}$$

In sum

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ \sqrt{y} & \text{if } y \in [0, 1) \\ 1 & \text{if } y \geq 1 \end{cases}$$

Since  $Y$  is continuous, we can derive the density

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & \text{if } y \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

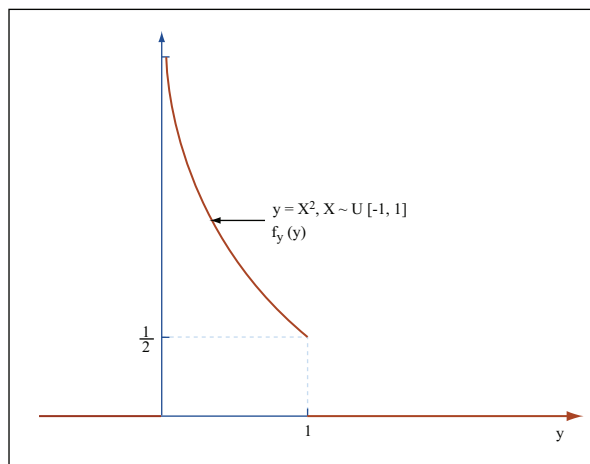


Image by MIT OpenCourseWare.

### 1.3 Change of Variables Formula for One-to-One Transformations

It is in general not very convenient to derive the density of  $Y$  from the density  $f_X(x)$  of  $X$  through the c.d.f.s, in particular because this involves one integration and one differentiation. So one might wonder whether there is a more direct connection between the p.d.f.s.

Before going to the more general case, suppose  $u(x) = ax$  for some constant  $a > 0$ . Then the c.d.f. of  $Y = u(X) = aX$  is given by

$$F_Y(y) = \int_{ax \leq y} f_X(x) dx = \int_{x \leq \frac{y}{a}} f_X(x) dx = F_X\left(\frac{y}{a}\right)$$

Using the chain rule, we can derive the p.d.f. of  $Y$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y}{a}\right) = \frac{1}{a} \frac{d}{dx} F_X(x) = \frac{1}{a} f_X(x)$$

What is a good intuition for this? - if  $a > 1$ , we could think of the transformation as stretching the axis on which the random variable falls. This moves any pair of points on the axis apart by a factor of  $a$ , but leaves constant the probability that the variable falls between the points. Therefore, the distribution of  $Y$  is "thinned out" by a factor of  $\frac{1}{a}$  compared to the distribution of  $X$ . One could visualize this by thinking about a lump of dough containing a number of raisins - the more we spread out the dough, the sparser the distribution of the raisins in the dough will be with respect to the surface of the table.

For differentiable monotone transformations  $u(\cdot)$  of  $X$  we have the following result

**Proposition 1** *Let  $X$  be a continuous random variable with known density  $f_X(x)$  such that  $P(a \leq X \leq b) = 1$ , and  $Y = u(X)$ . If  $u(\cdot)$  is strictly increasing and differentiable on an interval  $[a, b]$ , and has an inverse  $s(y) = u^{-1}(y)$ , then the density of  $Y$  is given by*

$$f_Y(y) = \begin{cases} f_X(s(y)) \left| \frac{d}{dy} s(y) \right| & \text{if } u(a) \leq y \leq u(b) \\ 0 & \text{otherwise} \end{cases}$$

Notice that an analogous result is true if  $u(x)$  is strictly *decreasing* on  $[a, b]$ .

**Example 6** *Let  $X$  be uniform on  $[0, 1]$ , so it has p.d.f.*

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the p.d.f. of  $Y = X^2$ ? We can see that on the support of  $X$ ,  $u(x) = x^2$  is strictly increasing and differentiable, so that we can use the inverse of  $u(\cdot)$ ,  $s(y) = \sqrt{y}$  to obtain the p.d.f. of  $Y$ ,

$$f_Y(y) = f_X(s(y)) \left| \frac{d}{dy} s(y) \right| = f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} = \begin{cases} \frac{1}{2\sqrt{y}} & \text{if } 0 \leq \sqrt{y} \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

This is similar to one example we did above, except that in the previous case, the support of  $X$  was  $[-1, 1]$ , so that  $u(x) = x^2$  was not monotone on the support of  $X$ .

It is very important to note that this formula works only for differentiable one-to-one - i.e. monotone - transformations. In other cases, we have to stick to the more cumbersome "2-step" methods for the discrete and continuous case, respectively.

## 1.4 Probability Integral / Quantile Transformation

For continuous random variables, there is an interesting - and also very useful - result: the "c.d.f. of a c.d.f." is that of a uniform variable in the following sense:

**Proposition 2** Let  $X$  be a continuous random variable with c.d.f.  $F_X(x)$ . Then the c.d.f. evaluated at a random draw of  $X$ ,  $F_X(X)$  is uniformly distributed, i.e.

$$F_X(X) \sim U[0, 1]$$

You should notice that a function of a random variable is itself a random variable (we'll discuss this in more detail later on).

PROOF: Since the c.d.f. takes values only between zero and 1, we can already see that the c.d.f.  $G(\cdot)$  of  $F(X)$  satisfies

$$\begin{aligned} G(F(X)) &= P(F(X) \leq x) = 0 & \text{if } x < 0 \\ G(F(X)) &= P(F(X) \leq x) = 1 & \text{if } x > 1 \end{aligned}$$

Without loss of generality (i.e. only avoiding a few uninteresting extra definitions or case distinctions), suppose  $F(\cdot)$  is strictly monotonic - keeping in mind that all c.d.f.s are nondecreasing. This means that there is an inverse function  $F^{-1}(\cdot)$ , i.e. a function such that  $F^{-1}(F(x)) = x$ .

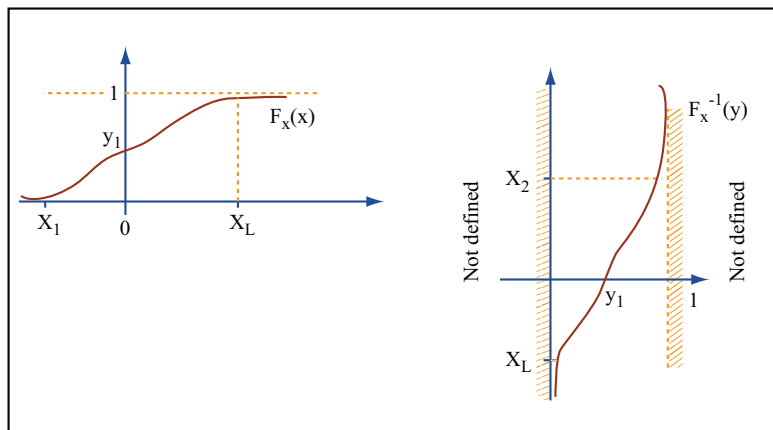


Image by MIT OpenCourseWare.

The inverse function will also be strictly monotonic, so that for  $0 \leq x \leq 1$ , the c.d.f. of the random variable  $F_X(X)$  is given by

$$P\left(F_X(X) \leq x\right) = P\left(F_X^{-1}(F_X(X)) \leq F_X^{-1}(x)\right) = P\left(X \leq F_X^{-1}(x)\right) = F_X(F_X^{-1}(x)) = x$$

(the first equality uses monotonicity of  $F^{-1}(\cdot)$ , and the third the definition of a c.d.f.). Summarizing, the c.d.f.  $G(\cdot)$  of the random variable  $F(X)$  is given by

$$G(F(x)) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

We can easily check that this is also the c.d.f. of a uniform random variable on the interval  $[0, 1]$ , so that  $F(X)$  has the same probability distribution as  $U[0, 1]$   $\square$

What is this result useful for? As an example, there are very efficient ways of generating *uniform* random numbers with a computer. If you want to get a sample of  $n$  draws from a random variable with c.d.f.  $F_X(\cdot)$ , you can

- draw  $U_1, \dots, U_n \sim U[0, 1]$
- transform each uniform draw according to

$$X_i = F_X^{-1}(U_i)$$

By the argument we made before,  $X_1, \dots, X_n$  behave like a random variable with c.d.f.  $F_X(\cdot)$ . This method is known as *integral (or quantile) transformation*.

**Example 7** Say, we have a computer program which allows us to draw a random variable  $U$  from a uniform distribution, but we actually want to obtain a random draw of  $X$  with p.d.f.

$$f_X(x) = \begin{cases} \frac{1}{2}e^{-\frac{1}{2}x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

We can obtain the c.d.f. of  $X$  by integration:

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-\frac{1}{2}x} & \text{if } x \geq 0 \end{cases}$$

so that the inverse of the c.d.f. is given by

$$F_X^{-1}(u) = -2 \log(1 - u) \quad \text{for } u \in [0, 1]$$

If we try this using some statistics software or Excel, a histogram of the draws will look like this:

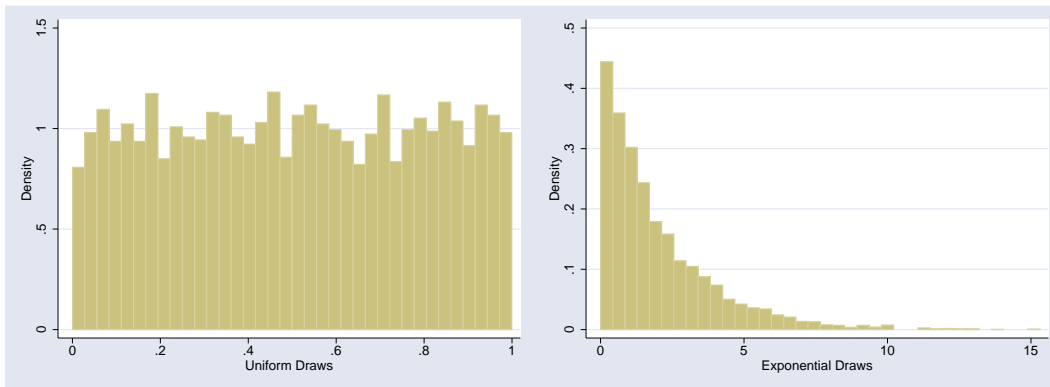


Figure 1: Histogram of 5,000 draws  $U_i$  from a Uniform (left), and the transformation  $X_i = -2 \log(1 - U_i)$  (right)

*If you want to try a few examples on your own in Excel, you can create uniform random draws using the `RAND()` function. You can then plot histograms by clicking yourself through the menus ("Tools">"Data Analysis">"Analysis Tools">"Histogram").*