Economics 520, Fall 2009

Lecture Note 3: Functions of Random Variables (CB 1.4–1.6, 2.1)

Often it is convenient to work with numbers rather than events. To do this we use random variables:

Definition 1 *A <u>random variable</u> is a function from the sample space to the real numbers.*

RV=random variable

Example 1: Toss a coin three times. The sample space is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

We can define a random variable X by taking the number of heads in the three tosses. Thus the random variable can take on values 0,1,2,3. So, for example, $X(\{HHT\}) = 2$.

□

Definition 2 *The <u>cumulative distribution function</u> (CDF) of a random variable X, denoted by $F_X(x)$, is*

$$F_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}).$$

be careful about the distinction between X and x

This is often written more informally as $P(X \leq x)$.

sometimes we also write *Pr* for *P*

Example 1 continued: Here the CDF is

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & 3 \leq x. \end{cases}$$

□

Some useful properties of CDFs:

Proposition 1 *The function $F(x)$ is a CDF if and only if:*

1. $\lim_{x \to -\infty} F(x) = 0$ *and* $\lim_{x \to \infty} F(x) = 1$.
2. $F(x)$ *is nondecreasing.*
3. $F(x)$ *is right-continuous; that is, for every* x_0, $\lim_{x \downarrow x_0} F(x) = F(x_0)$.

$x \downarrow x_0$ means x approaching x_0 from above

Definition 3 *A random variable X is <u>discrete</u> if $F_X(x)$ is a step function of x. A random variable X is <u>continuous</u> if $F_X(x)$ is a continuous function of x.*

Discrete random variables take on a countable (finite or infinite) number of values. Our example above is clearly a discrete random variable. Continuous random variables take on values in some continuous range.

Definition 4 *The probability mass function (PMF) of a discrete random variable X is given by*

$$f_X(x) = P(X = x).$$

So, for any set A, the probability

$$P(\{\omega \in \Omega : X(\omega) \in A\}) = P(X \in A) = \sum_{x \in A} f_X(x).$$

In particular, $P(X \le b) = \sum_{x \le b} f_X(x)$. Also, note that with $A = (-\infty, \infty)$, we have $\sum_{x \in A} f_X(x) = P(\omega \in \Omega) = 1$, so the probability function always adds up to unity.

Example 1 continued: The probability mass function here is

$$f_X(x) = \begin{cases} 1/8 & x = 0 \\ 3/8 & x = 1 \\ 3/8 & x = 2 \\ 1/8 & x = 3 \\ 0 & \text{otherwise.} \end{cases}$$

□

For continuous random variables, which take on an uncountable number of values, there is an analogous idea:

Definition 5 *A function $f_X(x)$ is a probability density function (PDF) of a continuous random variable if for all events A, the probability*

$$P(X \in A) = P(\{\omega \in \Omega : X(\omega) \in A\}) = \int_A f_X(x) dx.$$

Note that the probability density function is not unique. Because we only care about integrals over the probability density function, we can change its value at a countable number of points without changing any of the associated probabilities, and thus without changing the distribution of the random variable.

Also note that the probability density function integrates to one, in contrast to the probability function which sums to one.[1]

[1] Since PDF and PMF work similarly, many people just use "probability density function" to refer to both continuous and discrete cases. In measure-theoretic probability, there is no need to make a distinction between the two, so this saves a bit of excess terminology.

By the definition of PDF, we have that

$$P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

So, if f is continuous,

$$\frac{d}{dx}F_X(x) = f_X(x).$$

by the First Fundamental Theorem of Calculus

Example 2: Consider the experiment of picking a point randomly on the interval from zero to two and defining the random variable X as the distance to zero. It may be reasonable to assign the probability to the point picked being in any subinterval as being proportional to the length of that interval. In that case the CDF is:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x/2 & 0 \leq x < 2 \\ 1 & 2 \leq x. \end{cases}$$

and the PDF is:

$$f_X(x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \leq x < 2 \\ 0 & 2 \leq x. \end{cases}$$

□

Not all random variables taking on an uncountable number of values are continuous. There are mixed random variables, which are partly continuous and partly discrete. For example, a variable such as hours worked per year, or expenditures on cars, might have some positive mass at 0, and be continuously distributed for values greater than zero. This creates some conceptual difficulties in defining what the probability (mass or density) function is, but one can always work with the CDF.

Suppose we have a random variable X with distribution function $F_X(x)$ and PMF or PDF $f_X(x)$, where $f_X(x) > 0$ for $x \in \mathcal{X}$. Sometimes we are interested in the distribution of a function of this random variable, say $Y = g(X)$. The distribution of this new random variable is defined by the equation

$$P(Y \in A) = P(g(X) \in A) = P(\{\omega \in \Omega : g(X(\omega)) \in A\}).$$

We are interested in calculating the PDF and PMF/PDF for this transformation. The following result gives the general case:

Result 1 *Define the mapping $g^{-1}(A)$ by:*

$$g^{-1}(A) = \{x : g(x) \in A\}.$$

Then

$$\begin{aligned} F_Y(y) &= P(\{\omega \in \Omega : Y(\omega) \in (-\infty, y]\}) \\ &= P(\{\omega \in \Omega : g(X(\omega)) \in (-\infty, y]\}) \\ &= P(\{\omega \in \Omega | X(\omega) \in g^{-1}((-\infty, y])\}). \end{aligned}$$

This result is not always that easy to apply. More useful results rely on transformations satisfying particular conditions:

Result 2 *Suppose* $g(x)$ *is a strictly monotone function with inverse* $g^{-1}(\cdot)$*. Let* $\mathcal{Y} = \{y : g(x) = y$ for some $x \in \mathcal{X}\}$*. Then*
(i) If X *is a discrete random variable with PMF* $f_X(x)$ *on* $\mathcal{X}$*, then* Y *is a discrete random variable with PMF*

$$f_Y(y) = f_X(g^{-1}(y)),$$

on $\mathcal{Y}$*,*
(ii) If X *is a continuous random variable with PDF* $f_X(x)$ *on* $\mathcal{X}$*, then* Y *is a continuous random variable with PDF*

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{\partial g^{-1}}{\partial y}(y) \right|,$$

on $\mathcal{Y}$*.*

The argument for the continuous case goes as follows, assuming $g(\cdot)$ is increasing,

$$F_Y(y) = P(Y \le y) = P(g(X) \le y)$$

$$= P(x \le g^{-1}(y)) = F_X(g^{-1}(y)).$$

Now use the chain rule to take the derivative with respect to y to get the desired result.

Example 3: Suppose X has a Poisson distribution with parameter λ. That is, the PMF of X is

This is an example of a *parametric family* of distributions

$$f_X(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise,} \end{cases}$$

for some positive λ. This is a distribution that is often used for counts of events, for example counts of emissions of radioactive particles, counts of job offers, number of patents, etc. Different choices for λ lead to different probability distributions, but they all have a similar form.

Suppose we are interested in the distribution of $Y = g(X) = 2 \cdot X$. The inverse of the transformation is $X = g^{-1}(Y) = Y/2$. The PMF of Y is

$$f_X(x) = \begin{cases} \frac{e^{-\lambda}\lambda^{(y/2)}}{(y/2)!} & y = 0, 2, 4, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

□

note how $\mathcal{Y}$ reflects the transformation

Example 4: Suppose X has an exponential distribution with PDF

$$f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The CDF for this distribution is $F_X(x) = 1 - e^{-x}$. This distribution, or rather its extension with $f_X(x) = \lambda \exp(-\lambda x)$, for positive λ, is widely used for modelling durations such as survival times after heart transplants, or durations of unemployment spells.

Suppose we are interested in the distribution of $Y = g(X) = 1 - e^{-X}$. This is a monotone transformation with inverse $X = g^{-1}(Y) = -\ln(1 - Y)$. The derivative of the inverse of the transformation is

$$\frac{\partial g^{-1}}{\partial y}(y) = \frac{1}{1-y}.$$

The range of Y is $\mathcal{Y} = (0, 1)$. The PDF of Y is

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \cdot \left| \frac{\partial g^{-1}}{\partial y}(y) \right| = \exp(\ln(1-y)) \cdot \frac{1}{1-y} = 1 & 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

This is known as a uniform distribution (the PDF is constant over its range). □

Example 5: Finally consider a non–monotone transformation. Let X be a random variable with PDF

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

a uniform distribution on the interval $[-1, 1]$. Suppose we are interested in the distribution of $Y = g(X) = X^2$. Because of the non–monotonicity we have to do this on a more *ad hoc* basis. In this example we work directly through the cumulative distribution functions. Alternatively we can split things up into intervals where the transformation is monotone. First we calculate the CDF for X:

$$F_X(x) = \begin{cases} 0 & x \leq -1 \\ (x+1)/2 & -1 < x \leq 1 \\ 1 & 1 < x. \end{cases}$$

Then the CDF for $Y = X^2$ is

$$P(Y \leq y) = P(X^2 < y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Hence

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ (\sqrt{y}+1)/2 - (-\sqrt{y}+1)/2 = \sqrt{y} & 0 < y \leq 1 \\ 1 & 1 < y. \end{cases}$$

The PDF is then

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

□