ME 102: Lecture 6

S. Gopalakrishnan

August 22, 2016

Random Variables

- When a random experiment is performed, we are often not interested in all of the details of the experimental result but only in the value of some numerical quantity determined by the result.
- These quantities of interest that are determined by the result of the experiment are known as random variables.
- Since the value of a random variable is determined by the outcome of the experiment, we may assign probabilities of its possible values.

Example of fair dice

Let X denote the random variable that is defined as the sum of two fair dice

$$
\begin{aligned}
P\{X=2\} &= P\{(1,1)\} &&= \frac{1}{36}\\
P\{X=3\} &= P\{(1,2),(2,1)\} &&= \frac{2}{36}\\
P\{X=4\} &= P\{(1,3),(2,2),(3,1)\} &&= \frac{3}{36}\\
P\{X=5\} &= P\{(1,4),(2,3),(3,2),(4,1)\} &&= \frac{4}{36}\\
P\{X=6\} &= P\{(1,5),(2,4),(3,3),(4,2),(5,1)\} &&= \frac{5}{36}\\
P\{X=7\} &= P\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} &&= \frac{6}{36}\\
P\{X=8\} &= P\{(2,6),(3,5),(4,4),(5,3),(6,2)\} &&= \frac{5}{36}\\
P\{X=9\} &= P\{(3,6),(4,5),(5,4),(6,3)\} &&= \frac{4}{36}\\
P\{X=10\} &= P\{(4,6),(5,5),(6,4)\} &&= \frac{3}{36}\\
P\{X=11\} &= P\{(5,6),(6,5)\} &&= \frac{2}{36}\\
P\{X=12\} &= P\{(6,6)\} &&= \frac{1}{36}
\end{aligned}
$$

In other words, the random variable X can take on any integral value between 2 and 12 and the probability that it takes on each value is given in previous slide. Since X must take on some value, we must have

$$1 = P(S) = P\left(\bigcup_{i=2}^{12}\{X = i\}\right) = \sum_{i=2}^{12} P\{X = i\}$$

Suppose that an individual purchases two electronic components each of which may be either defective or acceptable. In addition, suppose that the four possible results – $(d, d), (d, a), (a, d), (a, a)$ – have respective probabilities .09, .21, .21, .49 [where (d, d) means that both components are defective, (d, a) that the first component is defective and the second acceptable, and so on]. If we let X denote the number of acceptable components obtained in the purchase, then X is a random variable taking on one of the values 0, 1, 2 with respective probabilities

$$P\{X = 0\} = 0.09$$
$$P\{X = 1\} = 0.42$$
$$P\{X = 2\} = 0.49$$

If we were mainly concerned with whether there was at least one acceptable component, we could define the random variable I by

$$I = \begin{cases} 1 \text{ if } X = 1 \text{ or } 2 \\ 0 \text{ if } X = 0 \end{cases}$$

If A denotes the event that at least one acceptable component is obtained, then the random variable I is called the indicator random variable for the event A, since I will equal 1 or 0 depending upon whether A occurs. The probabilities attached to the possible values of I are

$$P\{I = 1\} = 0.91$$
$$P\{I = 0\} = 0.09$$

Random variables whose set of possible values can be written either as a finite sequence $x_1, \ldots, x_n$, or as an infinite sequence $x_1, \ldots$ are said to be discrete. For instance, a random variable whose set of possible values is the set of nonnegative integers is a discrete random variable. However, there also exist random variables that take on a continuum of possible values. These are known as continuous random variables. One example is the random variable denoting the lifetime of a car, when the car's lifetime is assumed to take on any value in some interval (a, b).

Distribution Function

The cumulative distribution function, or more simply the distribution function, F of the random variable X is defined for any real number x by

$$F(x) = P\{X \leq x\}$$

That is, $F(x)$ is the probability that the random variable X takes on a value that is less than or equal to x.v

Notation: We will use the notation $X \sim F$ to signify that F is the distribution function of X .

All probability questions about X can be answered in terms of its distribution function F .

For example, suppose we wanted to compute $P\{a < X \leq b\}$. This can be accomplished by first noting that the event $\{X \leq b\}$ can be expressed as the union of the two mutually exclusive events $\{X \leq a\}$ and $\{a < X \leq b\}$. Therefore, applying Axiom 3, we obtain that

$$P\{X \leq b\} = P\{X \leq a\} + P\{a < X \leq b\}$$

$$P\{a < X \leq b\} = F(b) - F(a)$$

Example

Suppose the random variable X has distribution function

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-x^2} & x > 0 \end{cases}$$

What is the probability that X exceeds 1?

Example

Suppose the random variable X has distribution function

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-x^2} & x > 0 \end{cases}$$

What is the probability that X exceeds 1?
The desired probability is computed as follows:

$$\begin{aligned} P\{X > 1\} &= 1 - P\{X \leq 1\} \\ &= 1 - F(1) \\ &= e^{-1} \\ &= .368 \end{aligned}$$

Types of Random Variables

As was previously mentioned, a random variable whose set of possible values is a sequence is said to be discrete. For a discrete random variable X , we define the probability mass function $p(a)$ of X by

$$p(a) = P\{X = a\}$$

The probability mass function $p(a)$ is positive for at most a countable number of values of a. That is, if X must assume one of the values $x_1, x_2, \ldots,$ then

$$\begin{aligned} p(x_i) &> 0, && i = 1, 2, \ldots \\ p(x) &= 0, && \text{all other values of x} \end{aligned}$$

Since X must take on one of the values x_i , we have

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

Consider a random variable X that is equal to 1, 2, or 3. If we know that

$$p(1) = \frac{1}{2} \text{ and } p(2) = \frac{1}{3}$$

then it follows (since $p(1) + p(2) + p(3) = 1$) that $p(3) = \frac{1}{6}$

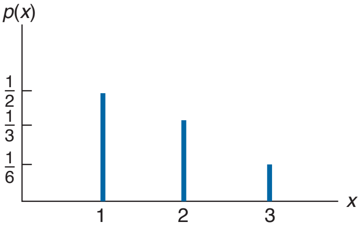

Figure: Graph of $p(x)$

The cumulative distribution function F can be expressed in terms of $p(x)$ by

$$F(a) = \sum_{\text{all } x \leq a} p(x)$$

If X is a discrete random variable whose set of possible values are $x_1, x_2, x_3, \ldots,$ where $x_1 < x_2 < x_3 < \ldots,$ then its distribution function F is a step function. That is, the value of F is constant in the intervals $[x_{i-1}, x_i)$ and then takes a step (or jump) of size $p(x_i)$ at x_i .

For instance, suppose X has a probability mass function given as

$$p(1) = \frac{1}{2} \qquad p(2) = \frac{1}{3} \qquad p(3) = \frac{1}{6}$$

then the cumulative distribution function F of X is given by

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{2} & 1 \leq a < 2 \\ \frac{5}{6} & 2 \leq a < 3 \\ 1 & 3 \leq a \end{cases}$$

For instance, suppose X has a probability mass function given as

$$p(1) = \frac{1}{2} \qquad p(2) = \frac{1}{3} \qquad p(3) = \frac{1}{6}$$

then the cumulative distribution function F of X is given by

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{2} & 1 \leq a < 2 \\ \frac{5}{6} & 2 \leq a < 3 \\ 1 & 3 \leq a \end{cases}$$

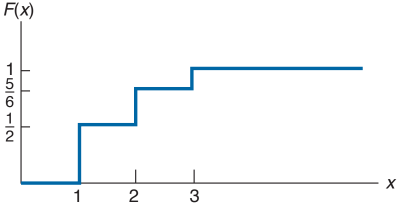

Figure: Graph of $F(x)$

Continuous Random Variables

We often must consider random variables whose set of possible values is an interval. Let X be such a random variable. We say that X is a continuous random variable if there exists a nonnegative function $f(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set B of real numbers

$$P\{X \in B\} = \int_B f(x)dx$$

The function $f(x)$ is called the probability density function of the random variable X.

In words, the equation states that the probability that X will be in B may be obtained by integrating the probability density function over the set B.

Since X must assume some value, $f(x)$ must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

All probability statements about X can be answered in terms of $f(x)$. For instance, letting $B = [a, b]$, we obtain that

$$P\{a \leq X \leq b\} = \int_{a}^{b} f(x)dx$$

if we let $a = b$ in the above, then

$$P\{a \leq X \leq a\} = \int_{a}^{a} f(x)dx = 0$$

In words, this equation states that the probability that a continuous random variable will assume any particular value is zero.

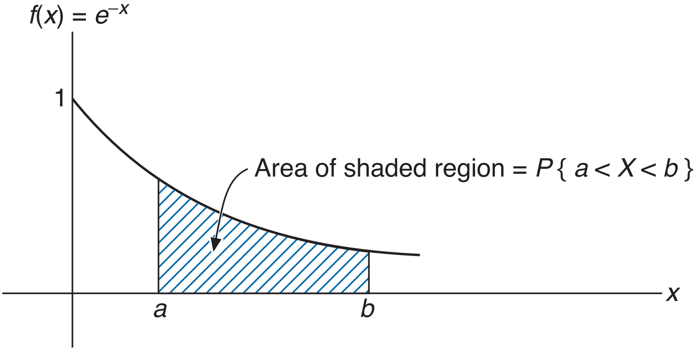

Figure: The probability density function $f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

The relationship between the cumulative distribution $F(\cdot)$ and the probability density $f(\cdot)$ is expressed by

$$F(a) = P\{X \in (-\infty, a]\} = \int_{-\infty}^{a} f(x)dx$$

Differentiating both sides yields

$$\frac{d}{da}F(a) = f(a)$$

That is, the density is the derivative of the cumulative distribution function.

A somewhat more intuitive interpretation of the density function may be obtained as follows:

$$P\left\{a-\frac{\epsilon}{2} \leq X \leq a+\frac{\epsilon}{2}\right\}=\int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x) d x=\epsilon f(a)$$

where ϵ is small. In other words, the probability that X will be contained in an interval of length ϵ around the point a is approximately $\epsilon f(a)$. From this, we see that $f(a)$ is a measure of how likely it is that the random variable will be near a.

Example

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of C? Find $P\{X > 1\}$.

Example

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of C? Find $P\{X > 1\}$.
Since f is a probability density function, we must have that $\int_{-\infty}^{\infty} f(x)dx = 1$, implying

$$C \int_0^2 (4x - 2x^2)dx = 1$$

$$C \left[2x^2 - \frac{2x^3}{3} \right]_{x=0}^{x=2} = 1$$

Hence $C = \frac{3}{8}$

Example

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of C? Find $P\{X > 1\}$.
Since f is a probability density function, we must have that $\int_{-\infty}^{\infty} f(x)dx = 1$, implying

$$C \int_0^2 (4x - 2x^2)dx = 1$$

$$C \left[2x^2 - \frac{2x^3}{3} \right]_{x=0}^{x=2} = 1$$

Hence $C = \frac{3}{8}$

$$P\{X > 1\} = \int_1^{\infty} f(x)dx = \frac{3}{8} \int_1^2 (4x - 2x^2)dx = \frac{1}{2}$$

Jointly Distributed Random Variables

To specify the relationship between two random variables, we define the joint cumulative probability distribution function of X and Y by

$$F(x, y) = P\{X \leq x, Y \leq y\}$$

A knowledge of the joint probability distribution function enables one, at least in theory, to compute the probability of any statement concerning the values of X and Y . For instance, the distribution function of X, call it F_X, can be obtained from the joint distribution function F of X and Y as follows:

$$\begin{aligned} F_X(x) &= P\{X \leq x\} \\ &= P\{X \leq x, Y < \infty\} \\ &= F(x, \infty) \end{aligned}$$

Similarly, the cumulative distribution function of Y is given by

$$F_Y(y) = F(\infty, y)$$

In the case where X and Y are both discrete random variables whose possible values are, respectively, $x_1, x_2, \ldots,$ and $y_1, y_2, \ldots,$ we define the joint probability mass function of X and Y, $p(x_i, y_j)$, by

$$p(x_i, y_j) = P\{X = x_i, Y = y_j\}$$

The individual probability mass functions of X and Y are easily obtained from the joint probability mass function by the following reasoning. Since Y must take on some value y_j , it follows that the event $\{X = x_i\}$ can be written as the union, over all j, of the mutually exclusive events $\{X = x_i, Y = y_j\}$. That is,

$$\{X = x_i\} = \bigcup_j P\{X = x_i, Y = y_j\}$$

and so, using Axiom 3 of the probability function, we see that

$$\begin{aligned} P\{X = x_i\} &= P\left(\bigcup_j \{X = x_i, Y = y_j\}\right) \\ &= \sum_j P\{X - x_i, Y = y_j\} \\ &= \sum_j p(x_i, y_j) \end{aligned}$$

Similarly, we can obtain $P\{Y = y_j\}$ by summing $p(xi_, y_j)$ over all possible values of x_i, that is,

$$\begin{aligned} P\{Y = y_j\} &= P\left(\bigcup_i \{X = x_i, Y = y_j\}\right) \\ &= \sum_i P\{X - x_i, Y = y_j\} \\ &= \sum_i p(x_i, y_j) \end{aligned}$$

Hence, specifying the joint probability mass function always determines the individual mass functions. However, it should be noted that the reverse is not true. Namely, knowledge of $P\{X = x_i\}$ and $P\{Y = y_j\}$ does not determine the value of $P\{X = x_i, Y = y_j\}$.

Example

Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries. If we let X and Y denote, respectively, the number of new and used but still working batteries that are chosen, then the joint probability mass function of X and Y , $p(i,j) = P\{X = i, Y = j\}$, is given by

$$p(0,0) = \binom{5}{3} \Big/ \binom{12}{3} = \frac{10}{220}$$

$$p(0,1) = \binom{4}{1}\binom{5}{2} \Big/ \binom{12}{3} = \frac{40}{220}$$

$$p(0,2) = \binom{4}{2}\binom{5}{1} \Big/ \binom{12}{3} = \frac{30}{220}$$

$$p(0,3) = \binom{4}{3} \Big/ \binom{12}{3} = \frac{4}{220}$$

$$p(1,0) = \binom{3}{1}\binom{5}{2} \Big/ \binom{12}{3} = \frac{30}{220}$$

$$p(1,1) = \binom{3}{1}\binom{4}{1}\binom{5}{1} \Big/ \binom{12}{3} = \frac{60}{220}$$

$$p(1,2) = \binom{3}{1}\binom{4}{2} \Big/ \binom{12}{3} = \frac{18}{220}$$

$$p(2,0) = \left(\begin{array}{c} 3 \\ 2 \end{array}\right)\left(\begin{array}{c} 5 \\ 1 \end{array}\right)\Big/\left(\begin{array}{c} 12 \\ 3 \end{array}\right) = \frac{15}{220}$$

$$p(2,1) = \left(\begin{array}{c} 3 \\ 2 \end{array}\right)\left(\begin{array}{c} 4 \\ 1 \end{array}\right)\Big/\left(\begin{array}{c} 12 \\ 3 \end{array}\right) = \frac{12}{220}$$

$$p(3,0) = \left(\begin{array}{c} 3 \\ 3 \end{array}\right)\Big/\left(\begin{array}{c} 12 \\ 3 \end{array}\right) = \frac{1}{220}$$

$$p(2,0) = \binom{3}{2}\binom{5}{1} \Big/ \binom{12}{3} = \frac{15}{220}$$

$$p(2,1) = \binom{3}{2}\binom{4}{1} \Big/ \binom{12}{3} = \frac{12}{220}$$

$$p(3,0) = \binom{3}{3} \Big/ \binom{12}{3} = \frac{1}{220}$$

These probabilities can most easily be expressed in tabular form as

TABLE 4.1 $P\{X = i, Y = j\}$

i \ j	0	1	2	3	Row Sum $= P\{X = i\}$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
Column Sums $= P\{Y = j\}$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

Jointly continuous functions

We say that X and Y are jointly continuous if there exists a function $f(x, y)$ defined for all real x and y, having the property that for every set C of pairs of real numbers (that is, C is a set in the two-dimensional plane)

$$P\{(X, Y) \in C\} = \iint_{(x,y)\in C} f(x, y)dxdy$$

The function $f(x, y)$ is called the joint probability density function of X and Y . If A and B are any sets of real numbers, then by defining $C = \{(x, y) : x \in A, y \in B\}$, we see that

$$P\{X \in A, Y \in B\} = \int_B \int_A f(x, y)dxdy$$

Since

$$\begin{aligned} F(a, b) &= P\{X \in (-\infty, a], Y \in (-\infty, b]\} \\ &= \int_{-\infty}^{b} \int_{-\infty}^{a} f(x, y)dxdy \end{aligned}$$

upon differentiation,

$$f(a, b) = \frac{\partial^2}{\partial a \partial b} F(a, b)$$

wherever the partial derivatives are defined. Another interpretation of the joint density function is obtained as follows:

$$\begin{aligned} P\{a < X < a + da, b < Y < b + db\} &= \int_b^{b+db} \int_a^{a+da} f(x, y) dx dy \\ &= f(a, b) da db \end{aligned}$$

when da and db are small and $f(x, y)$ is continuous at a, b. Hence $f(a, b)$ is a measure of how likely it is that the random vector (X, Y) will be near (a, b).

If X and Y are jointly continuous, they are individually continuous, and their probability density functions can be obtained as follows:

$$\begin{aligned} P\{X \in A\} &= P\{X \in X, Y \in (-\infty, \infty)\} \\ &= \int_A \int_{-\infty}^{\infty} f(x, y) dxdy \\ &= \int_A f_X(x) dx \end{aligned}$$

where

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

is thus the probability density function of X . Similarly, the probability density function of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

compute the following

(a) $P\{X > 1, Y < 1\}$

(b) $P\{X < Y\}$

(c) $P\{X < a\}$.

Example

The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

compute the following

(a) $P\{X > 1, Y < 1\}$

(b) $P\{X < Y\}$

(c) $P\{X < a\}$.

Solution:

(a)

$$\begin{aligned} P\{X > 1, Y < 1\} &= \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dxdy \\ &= \int_0^1 2e^{-2y}(-e^{-x}|_1^\infty)dy \\ &= e^{-1}\int_0^1 2e^{-2y}dy \\ &= e^{-1}(1 - e^{-2}) \end{aligned}$$

Example

(b)

$$
\begin{aligned}
P\{X < Y\} &= \iint_{(x,y):x<y} 2e^{-x}e^{-2y}\,dxdy \\
&= \int_0^\infty \int_0^y 2e^{-x}e^{-2y}\,dxdy \\
&= \int_0^\infty 2e^{-2y}(1-e^{-y})dy \\
&= \int_0^\infty 2e^{-2y}dy - \int_0^\infty 2e^{-3y}dy \\
&= 1 - \frac{2}{3} = \frac{1}{3}
\end{aligned}
$$

Example

(b)

$$
\begin{aligned}
P\{X < Y\} &= \iint_{(x,y):x<y} 2e^{-x}e^{-2y}\,dxdy \\
&= \int_0^{\infty}\int_0^{y} 2e^{-x}e^{-2y}\,dxdy \\
&= \int_0^{\infty} 2e^{-2y}(1-e^{-y})\,dy \\
&= \int_0^{\infty} 2e^{-2y}\,dy - \int_0^{\infty} 2e^{-3y}\,dy \\
&= 1 - \frac{2}{3} = \frac{1}{3}
\end{aligned}
$$

(c)

$$
\begin{aligned}
P\{X < a\} &= \int_0^{a}\int_0^{\infty} 2e^{-x}e^{-2y}\,dxdy \\
&= \int_0^{a} e^{-x}\,dx \\
&= 1 - e^{-a}
\end{aligned}
$$