ME 102: Lecture 7

S. Gopalakrishnan

August 29, 2016

Recap of last lecture

- Cumulative Distribution Function

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

- Probability mass function

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

- Probability density function

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

The function $f(x)$ is the PDF.

- Joint probability density function

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

The function $f(x, y)$ is called the joint PDF.

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)

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

Independent Random Variables

The random variables X and Y are said to be independent if for any two sets of real numbers A and B

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\} \tag{1}$$

In other words, X and Y are independent if, for all A and B, the events $E_A = \{X \in A\}$ and $F_B = \{Y \in B\}$ are independent.

It can be shown by using the three axioms of probability that Equation 1 will follow if and only if for all a, b

$$P\{X \leq a, Y \leq b\} = P\{X \leq a\}, P\{Y \leq b\}$$

Hence, in terms of the joint distribution function F of X and Y , we have that X and Y are independent if

$$F(a, b) = F_X(a)F_Y(b) \quad \text{for all } a, b$$

When X and Y are discrete random variables, the condition of independence Equation 1 is equivalent to

$$p(x, y) = p_X(x)p_Y(y) \quad \text{for all } x, y \tag{2}$$

where p_X and p_Y are the probability mass functions of X and Y. The equivalence follows because, if Equation 1 is satisfied, then we obtain Equation 2 by letting A and B be, respectively, the one-point sets $A = \{x\}, B = \{y\}$.

Furthermore, if Equation 2 is valid, then for any sets A, B

$$\begin{aligned}
P\{X \in A, Y \in B\} &= \sum_{y \in B} \sum_{x \in A} p(x, y) \\
&= \sum_{y \in B} \sum_{x \in A} p_X(x) p_Y(y) \\
&= \sum_{y \in B} p_Y(y) \sum_{x \in A} p_X(x) \\
&= P\{Y \in B\} P\{X \in A\}
\end{aligned}$$

and thus Equation 1 is established.
In the jointly continuous case, the condition of independence is equivalent to

$$f(x, y) = f_X(x) f_Y(y) \qquad \text{for all } x, y$$

Loosely speaking, X and Y are independent if knowing the value of one does not change the distribution of the other. Random variables that are not independent are said to be dependent.

Example

Suppose that X and Y are independent random variables having the common density function.

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

Find the density function of the random variable X/Y

We start by determining the distribution function of X/Y . For $a > 0$

$$\begin{aligned}
F_{X/Y}(a) &= P\{X/Y \leq a\} \\
&= \iint_{x/y \leq a} f(x,y)dxdy \\
&= \iint_{x/y \leq a} e^{-x}e^{-y}dxdy \\
&= \int_0^\infty \int_0^{ay} e^{-x}e^{-y}dxdy \\
&= \int_0^\infty (1-e^{-ay})e^{-y}dy \\
&= \left[-e^{-y} + \frac{e^{-(a+1)y}}{a+1} \right]_0^\infty \\
&= 1 - \frac{1}{a+1}
\end{aligned}$$

Differentiation yields that the density function of X/Y is given by

$$f_{X/Y}(a) = \frac{1/}{(a+1)^2}, \qquad 0 < a < \infty$$

We can also define joint probability distributions for n random variables in exactly the same manner as we did for $n = 2$. For instance, the joint cumulative probability distribution function $F(a_1, a_2, \ldots, a_n)$ of the n random variables $X_1, X_2, \ldots, X_n$ is defined by

$$F(a_1, a_2, \ldots, a_n) = P\{X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n\}$$

If these random variables are discrete, we define their joint probability mass function $p(x_1, x_2, \ldots, x_n)$ by

$$p(x_1, x_2, \ldots, x_n) = P\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\}$$

Further, the n random variables are said to be jointly continuous if there exists a function $f(x_1, x_2, \ldots, x_n)$, called the joint probability density function, such that for any set C in n-space

$$P\{(X_1, X_2, \ldots, X_n) \in C\} = \int \int_{(x_1, x_2, \ldots, x_n) \in C} \cdots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$$

In particular, for any n sets of real numbers $A_1, A_2, \ldots, A_n$

$$P\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\} = \int_{A_n} \int_{A_{n-1}} \cdots \int_{A_1} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$$

The concept of independence may, of course, also be defined for more than two random variables. In general, the n random variables $X_1, X_2, \ldots, X_n$ are said to be independent if, for all sets of real numbers $A_1, A_2, \ldots, A_n$,

$$P\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\} = \prod_{i=1}^{n} P\{X_i \in A_i\}$$

As before, it can be shown that this condition is equivalent to

$$P\{X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n\} = \prod_{i=1}^{n} P\{X_i \leq a_i\} \quad \text{for all } a_1, a_2, \ldots, a_n$$

Finally, we say that an infinite collection of random variables is independent if every finite subcollection of them is independent.

Example

Suppose that the successive daily changes of the price of a given stock are assumed to be independent and identically distributed random variables with probability mass function given by

$$P\{\text{daily change is } i\} = \begin{cases} -3 & \text{with the probability } 0.05 \\ -2 & \text{with the probability } 0.10 \\ -1 & \text{with the probability } 0.20 \\ 0 & \text{with the probability } 0.30 \\ 1 & \text{with the probability } 0.20 \\ 2 & \text{with the probability } 0.10 \\ 3 & \text{with the probability } 0.05 \end{cases}$$

Then the probability that the stock's price will increase successively by 1, 2, and 0 points in the next three days is

$$P\{X_1 = 1, X_2 = 2, X_3 = 0\} = (.20)(.10)(.30) = .006$$

where we have let X_i denote the change on the ith day.

Conditional Distributions

The relationship between two random variables can often be clarified by consideration of the conditional distribution of one given the value of the other.
The conditional probability of E given F is defined, provided that $P(F) > 0$, by

$$P(E|F) = \frac{P(EF)}{P(F)}$$

Hence, if X and Y are discrete random variables, it is natural to define the conditional probability mass function of X given that $Y = y$, by

$$\begin{aligned} p_{X|Y}(x|y) &= P\{X = x|Y = y\} \\ &= \frac{P\{X = x, Y = y\}}{P\{Y = y\}} \\ &= \frac{p(x, y)}{p_Y(y)} \end{aligned}$$

for all values of y such that $p_Y(y) > 0$.

Example

Suppose that $p(x, y)$, the joint probability mass function of X and Y , is given by

$$p(0,0) = .4, \quad p(0,1) = .2, \quad p(1,0) = .1, \quad p(1,1) = .3.$$

Calculate the conditional probability mass function of X given that $Y = 1$.

$$P\{Y = 1\} = \sum_x p(x,1) = p(0,1) + p(1,1) = 0.5$$

Hence

$$P\{X = 0|Y = 1\} = \frac{p(0,1)}{P\{Y = 1\}} = \frac{2}{5}$$

$$P\{X = 1|Y = 1\} = \frac{p(1,1)}{P\{Y = 1\}} = \frac{3}{5}$$

Expectation

If X is a discrete random variable taking on the possible values $x_1, x_2, \ldots,$ then the expectation or expected?? value of X, denoted by $E[X]$, is defined by

$$E[X] = \sum_i x_i P\{X = x_i\}$$

Frequency Interpretation

Consider a random variable X that must take on one of the values $x_1, x_2, \ldots, x_n$ with respective probabilities $p(x_1), p(x_2), \ldots, p(x_n)$; and think of X as representing our winnings in a single game of chance. If we continually play this game, then the proportion of time that we win x_i will be $p(x_i)$. Since this is true for all i, $i = 1, 2, \ldots, n$, it follows that our average winnings per game will be

$$\sum_{i=1}^{n} (x_i) p(x_i) = E[X]$$

Suppose that we play N games where N is very large. Then in approximately $Np(x_i)$ of these games, we shall win x_i , and thus our total winnings in the N games will be

$$\sum_{i=1}^{n} x_i N_p(x_i)$$

implying that our average winnings per game are

$$\sum_{i=1}^{n} \frac{x_i N p(x_i)}{N} = \sum_{i=1}^{n} x_i p(x_i) = E[X]$$

Some examples

If X is the outcome of the roll of a fair die. The expectation is

$$E[x] = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6})\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

If I is an indicator random variable for the event A, that is, if

$$I = \begin{cases} 1 & \text{if A occurs} \\ 0 & \text{ if A does not occur} \end{cases}$$

then

$$E[I] = 1P(A) + 0P(A^c) = P(A)$$

Hence, the expectation of the indicator random variable for the event A is just the probability that A occurs.

Continuous Random Variable – Expectation

Suppose that X is a continuous random variable with probability density function f. Since, for small dx

$$f(x)dx \approx P\{x < X < x + dx\}$$

it follows that a weighted average of all possible values of X, with the weight given to x equal to the probability that X is near x, is just the integral over all x of $xf(x)dx$. Hence, it is natural to define the expected value of X by

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

Example

Suppose that you are expecting a message at some time past 5 P.M. From experience you know that X, the number of hours after 5 P.M. until the message arrives, is a random variable with the following probability density function:

$$f(x) = \begin{cases} \frac{1}{1.5} & \text{if } 0 < x < 1.5 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$E[X] = \int_{0}^{1.5} \frac{x}{1.5} dx = 0.75$$

Hence, on average, you would have to wait three-fourths of an hour.

Example

Suppose X has the following probability mass function

$$p(0) = .2, \quad p(1) = .5, \quad p(2) = .3$$

Calculate $E[X^2]$. Letting $Y = X^2$, we have that Y is a random variable that can take on one of the values $0^2, 1^2, 2^2$ with respective probabilities

$$p_Y(0) = P\{Y = 0^2\} = .2$$

$$p_Y(1) = P\{Y = 1^2\} = .5$$

$$p_Y(4) = P\{Y = 2^2\} = .3$$

Hence,

$$E[X2] = E[Y] = 0(.2) + 1(.5) + 4(.3) = 1.7$$

Example

The time, in hours, it takes to locate and repair an electrical breakdown in a certain factory is a random variable, call it X ? whose density function is given by

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

If the cost involved in a breakdown of duration x is x^3, what is the expected cost of such a breakdown?
Letting $Y = X^3$ denote the cost, we first calculate its distribution function as follows.
For $0 \leq a \leq 1$,

$$\begin{aligned} F_Y(a) &= P\{Y \leq a\} \\ &= P\{X^3 \leq a\} \\ &= P\{X \leq a^{\frac{1}{3}}\} \\ &= \int_0^{a^{\frac{1}{3}}} dx \\ &= a^{\frac{1}{3}} \end{aligned}$$

By differentiating $F_Y(a)$, we obtain the density of Y.

$$f_Y(a) = \frac{1}{3} a^{-\frac{2}{3}}, \quad 0 \leq a \leq 1$$

Hence,

$$
\begin{aligned}
E[X^3] = E[Y] &= \int_{-\infty}^{\infty} a f_Y(a) da \\
&= \int_0^1 a \frac{1}{3} a^{-\frac{2}{3}} da \\
&= \frac{1}{3} \int_0^1 a^{\frac{1}{3}} da \\
&= \frac{1}{3} \frac{3}{4} a^{\frac{4}{3}} \Big|_0^1 \\
&= \frac{1}{4}
\end{aligned}
$$

EXPECTATION OF A FUNCTION OF A RANDOM VARIABLE

(a) If X is a discrete random variable with probability mass function $p(x)$, then for any real–valued function g,

$$E[g(X)] = \sum_{x} g(x)p(x)$$

(b) If X is a continuous random variable with probability density function $f(x)$, then for any real–valued function g,

$$E[g(x)] = \int_{\infty}^{\infty} g(x)f(x)dx$$

A Corollary

If a and b are constants, then

$$E[aX + b] = aE[X] + b$$

Proof In the discrete case,

$$\begin{aligned} E[aX + b] &= \sum_{x} (ax + b)p(x) \\ &= a\sum_{x} xp(x) + b\sum_{x} p(x) \\ &= aE[X] + b \end{aligned}$$

In the continuous case,

$$\begin{aligned} E[aX+b] &= \int_{-\infty}^{\infty} (ax+b)f(x)dx \\ &= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{\infty}^{\infty} f(x)dx \\ &= aE[X]+b \end{aligned}$$

If we take a $=0$ we see that,

$$E[b]=b$$

if we take $b=0$, then we obtain

$$E[aX]=aE[X]$$

The expected value of a random variable X, $E[X]$, is also referred to as the mean or the first moment of X . The quantity $E[X^n], n \geq 1$, is called the nth moment of X . We note that

$$E[X^n] = \begin{cases} \sum_x x^n p(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x^n f(x)dx & \text{if } X \text{ is continuous} \end{cases}$$

Expected Value of Sums of Random Variables

If X and Y are random variables and g is a function of two variables, then

$$
\begin{aligned}
E[g(X,Y)] &= \sum_y \sum_x g(x,y)p(x,y) && \text{in the discrete case} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy && \text{in the continuous case}
\end{aligned}
$$

For example, if $g(X,Y) = X + Y$, then, in the continuous case,

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy \\
&= E[X] + E[Y]
\end{aligned}
$$

A similar result can be shown in the discrete case and indeed, for any random variables X and Y

$$E[X+Y] = E[X] + E[Y]$$

By repeatedly applying previous equation we can show that the expected value of the sum of any number of random variables equals the sum of their individual expectations.

$$\begin{aligned} E[X + Y + Z] &= E[(X + Y) + Z] \\ &= E[X + Y] + E[Z] \\ &= E[X] + E[Y] + E[Z] \end{aligned}$$

And in general, for any n

$$E[X_1 + X_2 \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

Example

A construction firm has recently sent in bids for 3 jobs worth (in profits) 10, 20, and 40 (thousand) dollars. If its probabilities of winning the jobs are respectively .2, .8, and .3, what is the firm's expected total profit?

Solution

Letting $X_i, i = 1, 2, 3$ denote the firm?s profit from job i , then

$$\text{total profit} = X_1 + X_2 + X_3$$

$$E[\text{total profit}] = E[X_1] + E[X_2] + E[X_3]$$

$$E[X_1] = 10(.2) + 0(.8) = 2$$

$$E[X_2] = 20(.8) + 0(.2) = 16$$

$$E[X_3] = 40(.3) + 0(.7) = 12$$

and thus the firm's expected total profit is 30 thousand dollars.

Example

A secretary has typed N letters along with their respective envelopes. The envelopes get mixed up when they fall on the floor. If the letters are placed in the mixed-up envelopes in a completely random manner (that is, each letter is equally likely to end up in any of the envelopes), what is the expected number of letters that are placed in the correct envelopes?

Solution

Letting X denote the number of letters that are placed in the correct envelope, we can most easily compute $E[X]$ by noting that

$$X = X_1 + X_2 + \cdots + X_N$$

where

$$X_i = \begin{cases} 1 & \text{if the } i\text{th letter is placed in its proper envelope} \\ 0 & \text{otherwise} \end{cases}$$

Now, since the ith letter is equally likely to be put in any of the N envelopes, it follows that

$$P\{X_i = 1\} = P\{i\text{th letter is in its proper envelope}\} = \frac{1}{N}$$

$$E[X_i] = 1P\{X_i = 1\} + 0P\{X_i = 0\} = \frac{1}{N}$$

Therefore

$$E[X] = E[X_1] + \cdots + E[X_N] = \left(\frac{1}{N}\right) N = 1$$

Variance

We expect X to take on values around its mean $E[X]$, it would appear that a reasonable way of measuring the possible variation of X would be to look at how far apart X would be from its mean on the average. One possible way to measure this would be to consider the quantity $E[|X - \mu|]$, where $\mu = E[X]$, and $[X - \mu]$ represents the absolute value of $X - \mu$.

If X is a random variable with mean μ, then the **variance** of X , denoted by $Var(X)$, is defined by

$$Var(X) = E[(X - \mu)^2]$$

An alternative formula for $Var(X)$ can be derived as follows:

$$\begin{aligned} Var(X) &= E[(X - \mu)^2] \\ &= E[X^2 - 2\mu X - \mu^2] \\ &= E[X^2] - E[2\mu X] + E[\mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - \mu^2 \end{aligned}$$

That is,

$$Var(X) = E[X^2] - (E[X])^2$$

Example

Compute $Var(X)$ when X represents the outcome when we roll a fair die.
Solution
Since $P\{X = i\} = 1, i = 1, 2, 3, 4, 5, 6,$ we obtain

$$
\begin{aligned}
E[X^2] &= \sum_{i=1}^{6} i^2 P\{X = i\} \\
&= 1^2(\frac{1}{6}) + 2^2(\frac{1}{6}) + 3^2(\frac{1}{6}) + 4^2(\frac{1}{6}) + 5^2(\frac{1}{6}) + 6^2(\frac{1}{6}) \\
&= \frac{91}{6}
\end{aligned}
$$

$$E[X] = \frac{7}{2}$$

$$
\begin{aligned}
Var(X) &= E[X^2] - (E[X])^2 \\
&= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}
\end{aligned}
$$

An Identity

For any constants a and b

$$Var(aX + b) = a^2 Var(X)$$

Let $\mu = E[X]$ and recall that $E[aX + b] = a\mu + b$. Thus, by the definition of variance, we have

$$\begin{aligned}
Var(aX + b) &= E[(aX + b - E[aX + b])^2] \\
&= E[(aX + b - a\mu - b)^2] \\
&= E[(aX - a\mu)^2] \\
&= E[a^2(X - \mu)^2] \\
&= a^2 E[(X - \mu)^2] \\
&= a^2 Var(X)
\end{aligned}$$

The quantity $\sqrt{Var(X)}$ is called the standard deviation of X.