Lecture 9

Expectation

The *expectation* EX—some times also written as $E(X)$—of a random variable X is defined formally as

$$EX = \sum_x xP\{X = x\}.$$

If X has infinitely-many possible values, then the preceding sum must also be defined. This happens, for example, if

$$\sum_x |x|f(x) < \infty.$$

Also, EX is always defined if $\mathrm{P}\{X \geq 0\} = 1$ [but EX might be ∞] , or if $\mathrm{P}\{X \leq 0\} = 1$ [but EX might be $-\infty$] .

Other equivalent terms for "expectation" are "mean," "expected value," and "mean value." People use these terms interchangeably.

Example 1. If X takes the values ± 1 with respective probabilities $1/2$ each, then $EX = 0$. For example, suppose you play a fair game; if you win then you win \$1; else, you lose \$1. If X denotes your total winnings after one play [negative win means loss] then $EX = 0$.

Example 2. There are N tickets with face values $x_1, \ldots, x_N$. We select one at random and let X denote the face value of the randomly-selected ticket. Then,

$$EX = \left(\frac{1}{N} \times x_1\right) + \cdots + \left(\frac{1}{N} \times x_N\right) = \frac{x_1 + \cdots + x_N}{N}.$$

That is, in this example, the expected draw is the ordinary average of the box.

Example 3 (Constant random variables)**.** Every number c is a random variable [$P\{c = x\} = 0$ if $x \neq 0$, and $P\{c = c\} = 1$]. Therefore, $Ec = c$.

Example 4 (Indicators)**.** Let A denote an event, and form a random variable I_A by setting $I_A := 1$ if A occurs, and $I_A := 0$ otherwise. Note $P\{I_A = 1\} = P(A)$ and $P\{I_A = 0\} = P(A^c)$. Therefore,

$$E(I_A) = P(A).$$

Theorem 1 (Addition rule for expectations)**.** *Suppose X and Y are random variables, defined on the same sample space, such that EX and EY are well defined and finite. Then, for all constants α and β,*

$$E(\alpha X + \beta Y) = \alpha EX + \beta EY.$$

Proof. If we could prove that

$$E(cZ) = cE(Z) \tag{10}$$

for every constant c and random variable Z. Then, $\alpha EX = E(\alpha X)$ and $\beta EY = E(\beta Y)$. Therefore, we would have to show that $E(X' + Y') = E(X') + E(Y')$ where $X' := \alpha X$ and $Y' := \beta Y$. That is, (10) reduces the problem to $\alpha = \beta = 1$.

In order to prove (10) we compute:

$$\begin{aligned} E(cZ) &= \sum_x xP\{cZ = x\} = \sum_x xP\left\{Z = \frac{x}{c}\right\} \\ &= c\sum_x \left(\frac{x}{c}\right) P\left\{Z = \frac{x}{c}\right\}. \end{aligned}$$

Change variables [$u := x/c$] to find that

$$E(cZ) = c\sum_u uP\{Z = u\} = cEZ,$$

as asserted.

It remains to prove the theorem with $\alpha = \beta = 1$. First of all note that

$$P\{X + Y = a\} = \sum_b P\{X = b\,, Y = a - b\} := \sum_b f(b\,, a - b).$$

[f denotes the joint mass function of $(X\,, Y)$.] Therefore,

$$E(X + Y) = \sum_a a\mathrm{P}\{X + Y = a\} = \sum_a a \sum_b f(b\,, a - b).$$

We can write a as $(a-b)+b$ and substitute:

$$E(X+Y) = \underbrace{\sum_a (a-b) \sum_b f(b\,, a-b)}_{T_1} + \underbrace{\sum_a \sum_b b f(b\,, a-b)}_{T_2}.$$

In order to compute T_1, we change the label of the variables $[c := a-b]$:

$$T_1 = \sum_c c \underbrace{\sum_b f(b\,, c)}_{\text{marginal of } Y} = \sum_c c \mathrm{P}\{Y=c\} = EY.$$

As regards T_2, we reorder the sums first, and then relabel $[c := a-b]$:

$$T_2 = \sum_b b \sum_a f(b\,, a-b) = \sum_b b \underbrace{\sum_c f(b\,, c)}_{\text{marginal of } X} = \sum_b b P\{X=x\} = EX.$$

The theorem follows. □

The addition rule of expectation yields the following by induction:

Corollary 1. *If $X_1, \ldots, X_n$ are all random variables with well-defined finite expectations, then*

$$E(X_1 + \cdots + X_n) = EX_1 + \cdots + EX_n.$$

Example 5 (Method of the indicators)**.** The "method of the indicators" is an application of the previous example and the previous corollary in the following way: If $A_1, \ldots, A_N$ are events, then

$$E\left(I_{A_1} + \cdots + I_{A_N}\right) = P(A_1) + \cdots + P(A_N).$$

You should note that $I_{A_1} + \cdots + I_{A_N}$ denotes the number of events among $A_1, \ldots, A_N$ which occur.

We will make several uses of the method of the indicators in this course. The following is our first example.

Example 6 (The mean of the binomials)**.** Let X have the binomial distribution with parameters n and p. What is EX? Of course, we can use the definition of expectations and write

$$EX = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}.$$

But this is not a simple-to-understand expression. Here is a more effective approach: We can realize $X = I_{A_1} + \cdots + I_{A_n}$ where $A_1, \ldots, A_n$ are independent events with $P(A_1) = \cdots = P(A_n) = p$. In particular,

$$EX = P(A_1) + \cdots + P(A_n) = np.$$

Thus, for example, in 10 tosses of a fair coin we expect $np = 10 \times \frac{1}{2} = 5$ heads.

Example 7. For all constants a and b, and all random variables X,

$$E(aX + b) = aE(X) + b,$$

provided that EX is well defined. For instance, suppose the temperature of a certain object, selected on a random day, is a random variable with mean 34° Fahrenheit. Then, the expected value of the temperature of the object in Celsius is

$$\frac{5}{9} \times (EX - 32) = \frac{5}{9} \times (34 - 32) = \frac{10}{9}.$$

Sort out the details.

Finally, two examples to test the boundary of the theory so far.

Example 8 (A random variable with infinite mean)**.** Let X be a random variable with mass function,

$$P\{X = x\} = \begin{cases} \dfrac{1}{Cx^2} & \text{if } x = 1, 2, \ldots, \\ 0 & \text{otherwise,} \end{cases}$$

where $C = \sum_{j=1}^{\infty}(1/j^2)$. Then,

$$EX = \sum_{j=1}^{\infty} j \cdot \frac{1}{Cj^2} = \infty.$$

But $\mathrm{P}\{X < \infty\} = \sum_{j=1}^{\infty} 1/(Cj^2) = 1$.

Example 9 (A random variable with an undefined mean)**.** Let X be a random with mass function,

$$P\{X = x\} = \begin{cases} \dfrac{1}{Dx^2} & \text{if } x = \pm 1, \pm 2, \ldots, \\ 0 & \text{otherwise,} \end{cases}$$

where $D = \sum_{j=\pm 1, \pm 2, \cdots}(1/j^2)$. Then, EX is undefined. If it were defined, then it would be

$$\lim_{n,m\to\infty} \left(\sum_{j=-m}^{-1} \frac{j}{Dj^2} + \sum_{j=1}^{n} \frac{j}{Dj^2} \right) = \frac{1}{D} \lim_{n,m\to\infty} \left(\sum_{j=-m}^{-1} \frac{1}{j} + \sum_{j=1}^{n} \frac{1}{j} \right).$$

But the limit does not exist. The rough reason is that if N is large, then $\sum_{j=1}^{N}(1/j)$ is very nearly $\ln N$ plus a constant (Euler's constant). "Therefore," if n, m are large, then

$$\left(\sum_{j=-m}^{-1} \frac{1}{j} + \sum_{j=1}^{n} \frac{1}{j}\right) \approx -\ln m + \ln n = \ln\left(\frac{n}{m}\right).$$

If $n = m \to \infty$, then this is zero; if $m \gg n \to \infty$, then this goes to $-\infty$; if $n \gg m \to \infty$, then it goes to $+\infty$.