Appendix B

The Algebra of Expectations

B.1 EXPECTATIONS OF RANDOM VARIABLES

A very prominent place in theoretical statistics is occupied by the concept of the mathematical expectation of a random variable X. If the distribution of X is discrete, the the expectation (or expected value) of X is defined to be

$$E(X) = \sum_x xp(x)$$

where the sum is taken over all of the different values that the variable X can assume. and

$$\sum_x p(x) = 1.00.$$

For a continuous random variable X ranging over all the real numbers, the expectation is defined by

$$E(X) = \int_{-\infty}^{\infty} xf(x)\, d(x)$$

where

$$\int_{-\infty}^{\infty} f(x)\, d(x) = 1.00.$$

In essence, the expectation defined in either of these ways is a kind of weighted sum of values, and thus the rules of summation have very close parallels in the rules for the algebraic treatment of expectations. These rules apply either to discrete or to continuous random variables if particular boundary conditions exist; for our purposes these rules can be used without our going further into these special qualifications, however.

RULE 1. If a is some constant number, then

$$E(a) = a.$$

That is, if the same constant value a were associated with each and every elementary event in some sample space, the expectation or mean of the values would most certainly be a.

RULE 2. If a is some constant real number and X is a random variable with expectation $E(X)$, then

$$E(aX) = aE(X).$$

Suppose that a new random variable is formed by multiplying each value of X by the constant number a. Then the expectation of the new random variable is just a times the expectation of X. This is very simple to show for a discrete random variable X: By definition,

$$E(aX) = \sum_x axp(ax).$$

However, the probability of any value aX must be exactly equal to the probability of the corresponding X value, and so

$$E(aX) = \sum_x axp(x) = a\sum_x xp(x) = aE(X).$$

RULE 3. If a is a constant real number and X is a random variable, then

$$E(X + a) = E(X) + a.$$

This can be shown very simply for a discrete variable. Here,

$$\begin{aligned} E(X + a) &= \sum_x (x + a)p(x + a) \\ &= \sum_x xp(x + a) + a\sum_x p(x + a). \end{aligned}$$

However, $p(X + a) = p(X)$ for each value of X, so that

$$E(X + a) = E(X) + a\sum_x p(x) = E(X) + a.$$

The expectations of functions of random variables, such as

$$E[(X + 2)^2]$$
$$E(\sqrt{X + b})$$
$$E(b^X),$$

to give only a few examples, are subject to the same algebraic rules as summations. That is, the operation indicated within the punctuation is to be carried out *before* the expectation is taken. It is most important that this be kept in mind during any algebraic argument involving expectations. In general,

$$E[(X + 2)^2] \neq [E(X) + E(2)]^2$$
$$E(\sqrt{X}) \neq (\sqrt{E(X)})$$
$$E(b^X) \neq b^{E(X)}$$

and so forth.

The next few rules concern two (or more) random variables, symbolized by X and Y.

RULE 4. If X is a random variable with expectation $E(X)$, and Y is a random variable with expectation $E(Y)$, then

$$E(X + Y) = E(X) + E(Y).$$

Verbally, this rule says that the expectation of a sum of two random variables is the sum of their expectations. Once again, the proof is simple for two discrete variables X and Y. Consider the new random variable $(X + Y)$. The probability of a value of $(X + Y)$ involving a *particular* X value and a *particular* Y value is the joint probability $p(x,y)$. Thus,

$$E(X + Y) = \sum_x \sum_y (x + y)p(x,y).$$

Notice that here the expectation involves the sum over all possible *joint* events (x,y). This could be written as

$$E(X + Y) = \sum_x \sum_y (x + y)p(x,y)$$
$$= \sum_x \sum_y xp(x,y) + \sum_x \sum_y yp(x,y).$$

However, for any fixed x.

$$\sum_y p(x,y) = p(x)$$

and for any fixed y,

$$\sum_x p(x,y) = p(y).$$

Thus,

$$p(x + y) = \sum_x xp(x) + \sum_y yp(y) = E(X) + E(Y).$$

In particular, one of the random variables may be in a functional relation to the other. For example, let $Y = 3X^2$. Then

$$\begin{aligned} E(X + Y) &= E(X + 3X^2) \\ &= E(X) + E(3X^2) \\ &= E(X) + 3E(X^2). \end{aligned}$$

This principle lets one *distribute* the expectation over an expression which itself has the form of a sum. We will make a great deal of use of this principle.

This rule may also be extended to any finite number of random variables:

RULE 5. Given some finite number of random variables, the expectation of the sum of those variables is the sum of their individual expectations. Thus,

$$E(X + Y + Z) = E(X) + E(Y) + E(Z)$$

and so on.

In particular, some of these random variables may also be in functional relations to others. Let $Y = 6X^4$, and let $Z = \sqrt{2X}$. Then

$$E(X + Y + Z) = E(X) + 6E(X^4) + E(\sqrt{2X}).$$

B.2 THE VARIANCE OF A RANDOM VARIABLE.

More useful rules involve the variance of a random variable. The variance is defined by

$$\text{var}(X) = \sigma_X^2 = E[X - E(X)]^2$$

or

$$\sigma_X^2 = E(X^2) - [E(X)]^2.$$

The standard deviation of the random variable X is then $\sigma_X = \sqrt{\text{var}(X)}$.

The following rules give the effect of a transformation of X on the variance.

RULE 6. If a is some constant real number, and if X is a random variable with expectation $E(X)$ and variance σ_X^2, then the random variable $(X + a)$ has variance σ_X^2.

This can be shown as follows:

$$\text{var}(X + a) = E[(X + a)^2] - [E(X + a)]^2.$$

By Rule 3 above, and expanding the squares, we have

$$\begin{aligned} E[(X + a)^2] - [E(X + a)]^2 &= E[(X^2 + 2Xa + a^2)] - [E(X) + a]^2 \\ &= E(X^2 + 2aX + a^2) - [E(X)]^2 - 2aE(X) - a^2. \end{aligned}$$

Then by Rules 5 and 1, we have

$$\begin{aligned} \text{var}(X + a) &= E(X^2) + 2aE(X) + a^2 - [E(X)]^2 - 2aE(X) - a^2 \\ &= E(X^2) - [E(X)]^2 \\ &= \sigma_X^2. \end{aligned}$$

RULE 7. If a is some constant real number, and if X is a random variable with variance σ_X^2, the variance of the random variable aX is

$$\text{var}(aX) = a^2\sigma^2{}_X.$$

In order to show this, we take

$$\begin{aligned} \text{var}(aX) &= E[(aX)^2] - [E(aX)]^2 \\ &= a^2E(X^2) - a^2[E(X)]^2 \\ &= a^2(E(X^2) - [E(X)]^2) \\ &= a^2\sigma_X^2. \end{aligned}$$

In short, adding a constant value to each value of a random variable leaves the variance unchanged, but multiplying each value by a constant multiplies the variance by the square of the constant.

RULE 8. If X and Y are independent random variables, with variances σ_X^2 and σ_Y^2 respectively, then the variance of the sum $X + Y$ is

$$\sigma_{(X+Y)}^2 = \sigma_X^2 + \sigma_Y^2.$$

Similarly, the variance of $X - Y$ is

$$\sigma_{(X-Y)}^2 = \sigma_X^2 + \sigma_Y^2.$$

This principle also extends to any number of independent random variables (see Eq. C.2.6 in Appendix C).

B.3 INDEPENDENCE AND COVARIANCE

The next rule is a most important one that applies only to *independent* random variables (see Section C.1 in Appendix C).

RULE 9. Given random variable X with expectation $E(X)$ and the random variable Y with expectation $E(Y)$, then if X and Y are independent,

$$E(XY) = E(X)E(Y).$$

This rule states that if random variables are *statistically independent*, the expectation of the product of these variables is the product of their separate expectations. An important corollary to this principle is:

If $E(XY) \neq E(X)E(Y)$, the variables X and Y are not independent.

The basis for Rule 9 can also be shown fairly simply for discrete variables. Since X and Y are independent, $p(x,y) = p(x)p(y)$. Then,

$$E(XY) = \sum_x \sum_y (xy)p(x)p(y) = \sum_x \sum_y xp(x)yp(y).$$

However, for any fixed x, $yp(y)$ is perfectly free to be any value, so that

$$E(XY) = \sum_x xp(x) \sum_y yp(y) = E(X)E(Y).$$

Definition: Given the random variable X with expectation $E(X)$ and the random variable Y with expectation $E(Y)$, then the covariance of X and Y is

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y),$$

the expected value of the product of X and Y minus the product of the expected values.

The covariance is thus a reflection of the departure from independence of X and Y. When X and Y are independent,

$$\text{cov}(X,Y) = 0,$$

by the rule given above. When random variables are *independent*, their covariance is *zero*. However, it is not necessarily true that zero covariance implies that the variables are independent. On the other hand, $\text{cov}(X,Y) \neq 0$ always implies that the variables X and Y are not independent.

Definition: Given two random variables X and Y, then the covariance X and divided by the standard deviation of each variable,

$$\rho_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X\sigma_Y},$$

is known as the coefficient of correlation between X and Y.

The correlation coefficient ρ_{XY} may be any value between -1 and 1. However, nc that if X and Y are independent, $\rho_{XY} = 0$. The converse is not true, however, sin $\rho_{XY} = 0$ does not necessarily imply the independence of X and Y.

By extension of Rule 9 to any finite number of random variables, we have:

RULE 10. Given any finite number of random variables, if all the variables ai independent of each other, the expectation of their product is the product of the separate expectations: thus,

$$E(XYZ) = E(X)E(Y)E(Z),$$

and so on.

EXERCISES—THE ALGEBRA OF EXPECTATIONS

Consider the following probability distributions of discrete random variables. Find the expectatio $E(X)$, for each.

1.

x	$p(x)$
1	.75
0	.25
	1.00

2.

x	$p(x)$
5	4/18
0	12/18
−5	2/18
	18/18

3.

x	$p(x)$
5	.1
4	.4
3	.1
2	.3
1	.1
	1.0

4.

x	$p(x)$
399	.20
154	.20
125	.20
100	.20
−200	.20
	1.00

5.

x	$p(x)$
36	.12
30	.18
24	.20
18	.32
12	.09
6	.05
0	.04
	1.00

Consider a discrete random variable taking on only the values $x_1, x_2, \cdots, x_{10}$. Symbolize the following as expectations:

6. $x_1p(x_1) + x_2p(x_2) + \cdots + x_{10}p(x_{10})$

7. $x_1^2p(x_1) + x_2^2p(x_2) + \cdots + x_{10}^2p(x_{10})$

8. $[x_1^2p(x_1) + x_2^2p(x_2) + \cdots + x_{10}^2p(x_{10})] - [x_1p(x_1) + x_2p(x_2) + \cdots + x_{10}p(x_{10})]^2$

9. $[x_1 - E(X)]^2p(x_1) + [x_2 - E(X)]^2p(x_2) + \cdots + [x_{10} - E(X)]^2p(x_{10})$

10. $(x_1^2 - 4x_1 + 5)p(x_1) + (x_2^2 - 4x_2 + 5)p(x_2) + \cdots + (x_{10}^2 - 4x_{10} + 5)p(x_{10})$

Consider two random variables, X and Y. Simplify the following:

11. $E[(X + 35)/10]$

12. $E[X - 14Y + E(Y) + 7] - E(X + Y - 5)$

13. $E[X^2 - 2XE(X) + E^2(X)]$

14. $E[X^2 + Y^2 - 2(X + Y)^2]$

15. $E(17)$

16. $E[(X - E(X))(Y - E(Y))]$

See if you can prove the following for discrete random variables: (**Hint:** Turn each expectation into the equivalent weighted sum.)

17. $E(aX) = aE(X)$

18. $E(aX + b) = aE(X) + b$

19. $E[X - E(X)] = 0$

For the variables defined in Exercises 1, 2, and 3, find

20. var(X), or σ_X^2

21. $E(5X - 12)$

22. $E[(X^2 + 2)/10]$

23. Suppose that the random variables given in Exercises 3 and 4 were independent. Let us call the variable in Exercise 3, X, and that in exercise 4, Y. Then, find the value of $E(XY)$.