A Tutorial on the Expectation and Variance Operators

Corrections to Dr Ian Rudy (http://www.robinson.cam.ac.uk/iar1/contact.html) please.

The expectation of a random variable is the value the variable takes on average. If we have a sample of values for a random variable *X* then we would estimate the expectation by adding the values and dividing by the number of values. If instead we have a random variable *X* that can take only certain values (say $x_1, x_2, x_3, \ldots, x_n$), and a sample of values of *X* gave corresponding frequencies as $f_1, f_2, f_3, \ldots, f_n$, then we would estimate the expectation from the grouped sample mean:

$$\hat{E}(X) = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} \qquad [1]$$

Hence it should seem reasonable that if our only information about a discrete random variable *X* is its probability mass function *p*(*x*), we would calculate the expectation of *X* by multiplying each possible value of *X* by its probability of occurring, and then summing over all possible values of *X*:

$$E(X) = \sum xp(x) \qquad [2]$$

If instead we have a continuous random variable with a probability density function *p*(*x*), then we would replace the sum by an integral:

$$E(X) = \int xp(x)dx \qquad [3]$$

where the (definite) integral is over all possible values of *X*.

If we wanted the expectation of X^2, it should be apparent that we would use the formula:

$$E(X^2) = \sum x^2 p(x) \text{ or } E(X^2) = \int x^2 p(x)dx \qquad [4]$$

And, extending this idea, if we wanted the expectation of some general function of *X*, say *g*(*X*) we would use the formula:

$$E(g(X)) = \sum g(x)p(x) \text{ or } E(g(X)) = \int g(x)p(x)dx \qquad [5]$$

The variance of a random variable is defined as:

$$\mathrm{var}(X) = E\,[(X - E(X))^2] \qquad [6]$$

It is possible to show (using the formulae below) that this is equivalent to:

$$\mathrm{var}(X) = E(X^2) - [E(X)]^2 \qquad [7]$$

and [7] is often easier to work with than [6].

Mathematicians might call *E*() and var() *operators*. If you have not come across the idea of an operator before, you might think of it as being a function, but with an emphasis that a particular process is being performed on the input. The expectation operator takes a random variable and gives you its average value, the variance operator takes a random variable and gives you its variance.

You should get used to using the expectation and variance operators. They save us from having to write summation and/or integral signs, and allow one to prove results for both discrete and continuous random variables in one go.

The expectation and variance operators obey certain very valuable rules. You should learn these rules and practice using them in questions. The rules can be proven from the definitions above, and are summarised here.

Let X and Y be random variables, and let k be a constant. Then:

$$E(X \pm k) = E(X) \pm k \qquad [8]$$

$$E(kX) = kE(X) \qquad [9]$$

$$E(X \pm Y) = E(X) \pm E(Y) \qquad [10]$$

if X and Y are independent: $E(XY) = E(X)E(Y)$ [11]

$$\mathrm{var}(X \pm k) = \mathrm{var}(X) \qquad [12]$$

$$\mathrm{var}(kX) = k^2\,\mathrm{var}(X) \qquad [13]$$

if X and Y are independent: $\mathrm{var}(X \pm Y) = \mathrm{var}(X) + \mathrm{var}(Y)$ [14]

if X and Y are *not* independent: $\mathrm{var}(X \pm Y) = \mathrm{var}(X) + \mathrm{var}(Y) \pm 2\,\mathrm{cov}(X,Y)$ [15]

where $\mathrm{cov}(X,Y) = E[(X - E(X))(Y - E(Y))]$ (covariance of X and Y)

A key piece of advice is that you use equations [12]-[15] to work with variances, rather than equations [6], [7] with [8]-[11], unless an exam question constrains you. By doing so, you will save a *lot* of time.

Some examples:

To prove the expectation of the mean of a sample equals the mean of the population:

$$E(\bar{x}) = E\left(\sum_{i=1}^{n} x_i / n\right) \overset{[10]}{=} \sum_{i=1}^{n} E(x_i / n) \overset{[9]}{=} \sum_{i=1}^{n} E(x_i)/n = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n} nE(X) = E(X)$$

To prove the variance of the mean of a sample of size n equals the variance of individuals in the population, divided by *n*:

$$\mathrm{var}(\bar{x}) = \mathrm{var}\left(\sum_{i=1}^{n} x_i / n\right) \overset{[14]}{=} \sum_{i=1}^{n} \mathrm{var}(x_i / n) \overset{[13]}{=} \sum_{i=1}^{n} \mathrm{var}(x_i)/n^2 = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{var}(x_i)$$
$$= \frac{1}{n^2} n\,\mathrm{var}(X) = \frac{\mathrm{var}(X)}{n}$$

To prove that [7] above follows from [6]:

$$\mathrm{var}(X) = E(X - E(X))^2 = E\left(X^2 - 2X\,E(X) + [E(X)]^2\right) \overset{[10]}{=} E(X^2) - E(2E(X)X) + E([E(X)]^2)$$
$$\overset{[9]}{=} E(X^2) - 2E(X)E(X) + [E(X)]^2 = E(X^2) - [E(X)]^2$$