3.3 Expected Values

The Expected Value of X

Definition

Let X be a discrete rv with set of possible values D and pmf $p(x)$. The **expected value** or **mean value** of X, denoted by $E(X)$ or μ_X or just μ, is

$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)$$

Example 16

Consider a university having 15,000 students and let X = of courses for which a randomly selected student is registered. The pmf of X follows.

x	1	2	3	4	5	6	7
$p(x)$	.01	.03	.13	.25	.39	.17	.02
Number registered	150	450	1950	3750	5850	2550	300

$$\mu = 1 \cdot p(1) + 2 \cdot p(2) + \ldots + 7 \cdot p(7)$$

$$= (1)(.01) + 2(.03) + \ldots + (7)(.02)$$

$$= .01 + .06 + .39 + 1.00 + 1.95 + 1.02 + .14$$

$$= 4.57$$

Example 16

cont'd

If we think of the population as consisting of the *X* values 1, 2, . . . , 7, then μ = 4.57 is the population mean.

In the sequel, we will often refer to μ as the *population mean* rather than the mean of *X* in the population.

Notice that μ here is not 4, the ordinary average of 1, . . . , 7, because the distribution puts more weight on 4, 5, and 6 than on other *X* values.

The Expected Value of a Function

The Expected Value of a Function

Sometimes interest will focus on the expected value of some function $h(X)$ rather than on just $E(X)$.

Proposition
If the rv X has a set of possible values D and pmf $p(x)$, then the expected value of any function $h(X)$, denoted by $E[h(X)]$ or $\mu_{h(X)}$, is computed by

$$E[h(X)] = \sum_{D} h(x) \cdot p(x)$$

That is, $E[h(X)]$ is computed in the same way that $E(X)$ itself is, except that $h(x)$ is substituted in place of x.

Example 23

A computer store has purchased three computers of a certain type at \$500 apiece. It will sell them for \$1000 apiece.

The manufacturer has agreed to repurchase any computers still unsold after a specified period at \$200 apiece.

Let X denote the number of computers sold, and suppose that $p(0) = .1$, $p(1) = .2$, $p(2) = .3$ and $p(3) = .4$.

Example 23

cont'd

With $h(X)$ denoting the profit associated with selling X units, the given information implies that

$$h(X) = \text{revenue} - \text{cost}$$

$$= 1000X + 200(3 - X) - 1500$$

$$= 800X - 900$$

The expected profit is then

$$E[h(X)] = h(0) \cdot p(0) + h(1) \cdot p(1) + h(2) \cdot p(2) + h(3) \cdot p(3)$$

$$= (-900)(.1) + (-100)(.2) + (700)(.3) + (1500)(.4)$$

$$= \$700$$

Rules of Expected Value

The $h(X)$ function of interest is quite frequently a linear function $aX + b$. In this case, $E[h(X)]$ is easily computed from $E(X)$.

Proposition

$E(aX + b) = a \cdot E(X) + b$

(Or, using alternative notation, $\mu_{aX + b} = a \cdot \mu_X + b$)

To paraphrase, the expected value of a linear function equals the linear function evaluated at the expected value $E(X)$. Since $h(X)$ in Example 23 is linear and $E(X) = 2$, $E[h(x)] = 800(2) - 900 = \$700$, as before.

The Variance of *X*

The Variance of X

Definition

Let X have pmf $p(x)$ and expected value μ. Then the **variance** of X, denoted by $V(X)$ or σ^2_X , or just σ^2, is

$$V(X) = \sum_D (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$$

The **standard deviation** (SD) of X is

$$\sigma_X = \sqrt{\sigma_X^2}$$

The Variance of X

The quantity $h(X) = (X - \mu)^2$ is the squared deviation of X from its mean, and σ^2 is the expected squared deviation—i.e., the weighted average of squared deviations, where the weights are probabilities from the distribution.

If most of the probability distribution is close to μ, then σ^2 will be relatively small.

However, if there are x values far from μ that have large $p(x)$, then σ^2 will be quite large.

Very roughly σ can be interpreted as the size of a representative deviation from the mean value μ.

The Variance of *X*

So if $\sigma = 10$, then in a long sequence of observed *X* values, some will deviate from μ by more than 10 while others will be closer to the mean than that—a typical deviation from the mean will be something on the order of 10.

Example 24

A library has an upper limit of 6 on the number of videos that can be checked out to an individual at one time. Consider only those who check out videos, and let *X* denote the number of videos checked out to a randomly selected individual. The pmf of *X* is as follows:

x	1	2	3	4	5	6
$p(x)$	.30	.25	.15	.05	.10	.15

The expected value of *X* is easily seen to be $\mu = 2.85$.

Example 24

cont'd

The variance of *X* is then

$$V(X) = \sigma^2 = \sum_{x=1}^{6} (x - 2.85)^2 \cdot p(x)$$

$$= (1 – 2.85)^2(.30) + (2 – 2.85)^2(.25) + ... + (6 – 2.85)^2(.15) = 3.2275$$

The standard deviation of *X* is $\sigma = \sqrt{3.2275} = 1.800$.

The Variance of X

When the pmf $p(x)$ specifies a mathematical model for the distribution of population values, both σ^2 and σ measure the spread of values in the population; σ^2 is the population variance, and σ is the population standard deviation.

A Shortcut Formula for σ^2

The number of arithmetic operations necessary to compute σ^2 can be reduced by using an alternative formula.

Proposition

$$V(X) = \sigma^2 = \left[\sum_D x^2 \cdot p(x)\right] - \mu^2 = E(X^2) - [E(X)]^2$$

In using this formula, $E(X^2)$ is computed first without any subtraction; then $E(X)$ is computed, squared, and subtracted (once) from $E(X^2)$.

Rules of Variance

The variance of $h(X)$ is the expected value of the squared difference between $h(X)$ and its expected value:

$$V[h(X)] = \sigma^2_{h(X)} = \sum_D \{h(x) - E[h(X)]\}^2 \cdot p(x) \qquad \textbf{(3.13)}$$

When $h(X) = aX + b$, a linear function,

$$h(x) - E[h(X)] = ax + b - (a\mu + b) = a(x - \mu)$$

Substituting this into (3.13) gives a simple relationship between $V[h(X)]$ and $V(X)$:

Rules of Variance

Proposition

$V(aX + b) = \sigma^2_{aX+b} = a^2 \cdot \sigma^2_X$ and $\sigma_{aX+b} = |a| \cdot \sigma_X$

In particular,

$\sigma_{aX} = |a| \cdot \sigma_X$, $\sigma_{X+b} = \sigma_X$

(3.14)

The absolute value is necessary because *a* might be negative, yet a standard deviation cannot be.

Usually multiplication by *a* corresponds to a change in the unit of measurement (e.g., kg to lb or dollars to euros).

Rules of Variance

According to the first relation in (3.14), the sd in the new unit is the original sd multiplied by the conversion factor.

The second relation says that adding or subtracting a constant does not impact variability; it just rigidly shifts the distribution to the right or left.

Example 26

In the computer sales scenario of Example 23, $E(X) = 2$ and

$$E(X^2) = (0)^2(.1) + (1)^2(.2) + (2)^2(.3) + (3)^2(.4) = 5$$

so, $V(X) = 5 – (2)^2 = 1$. The profit function $h(X) = 800X – 900$ then has variance $(800)^2 \cdot V(X) = (640{,}000)(1) = 640{,}000$ and standard deviation 800.