Section 3.4 Variance

Thursday, September 8, 2016 1:11 PM

Section 3.4 Variance

(a) Compute the expected number of gallons produced during a one-hour period.

(b) The cost (in hundreds of dollars) to produce Y gallons is given by the cost function $C(Y) = 3 + 12Y + 2Y^2$. What is the expected cost in a one-hour period?

SOLUTION: (a) The expected value of Y is

$$E(Y) = \sum_{y \in R} y p_Y(y) = 0(0.2) + 1(0.3) + 2(0.3) + 3(0.2) = 1.5.$$

That is, we would expect 1.5 gallons of the toxic chemical to be produced per hour. For (b), we first compute $E(Y^2)$:

$$E(Y^2) = \sum_{y \in R} y^2 p_Y(y) = 0^2(0.2) + 1^2(0.3) + 2^2(0.3) + 3^2(0.2) = 3.3.$$

Finally,

$$\begin{aligned} E[C(Y)] &= E(3 + 12Y + 2Y^2) \\ &= 3 + 12E(Y) + 2E(Y^2) = 3 + 12(1.5) + 2(3.3) = 27.6. \end{aligned}$$

The expected hourly cost is \$2,760.00. □

3.4 Variance

TERMINOLOGY: Let Y be a discrete random variable with pmf $p_Y(y)$, support R, and expected value $E(Y) = \mu$. The **variance** of Y is given by

$$\sigma^2 \equiv V(Y) \equiv E[(Y-\mu)^2] = \sum_{y \in R} \underbrace{(y-\mu)^2}_{\geq 0} \underbrace{p_Y(y)}_{\geq 0}. \quad \geq 0$$

The **standard deviation** of Y is given by the positive square root of the variance; i.e.,

$$\sigma = \sqrt{\sigma^2} = \sqrt{V(Y)}.$$

FACTS: The variance σ^2 satisfies the following:

(a) $\sigma^2 \geq 0$. ✓

(b) $\sigma^2 = 0$ if and only if the random variable Y has a **degenerate distribution**; i.e., all the probability mass is located at one support point. $V(c) = 0$

(c) The larger (smaller) σ^2 is, the more (less) spread in the possible values of Y about the mean $\mu = E(Y)$.

(d) σ^2 is measured in $(\text{units})^2$ and σ is measured in the original units.

VARIANCE COMPUTING FORMULA: Let Y be a random variable with (finite) mean $E(Y) = \mu$. Then

$$V(Y) = E[(Y-\mu)^2] = E(Y^2) - [E(Y)]^2.$$

Proof. Expand the $(Y-\mu)^2$ term and distribute the expectation operator as follows:

$$\begin{aligned} E[(Y-\mu)^2] &= E(Y^2 - 2\mu Y + \mu^2) = E(Y^2) + E(-2\mu Y) + E(\mu^2) \\ &= E(Y^2) - 2\mu E(Y) + \mu^2 \\ &= E(Y^2) - 2\mu^2 + \mu^2 \\ &= E(Y^2) - \mu^2. \ \square \end{aligned}$$

Example 3.9. *The discrete uniform distribution*. Suppose that the random variable X has pmf

$$p_X(x) = \begin{cases} 1/m, & x = 1, 2, ..., m \\ 0, & \text{otherwise}, \end{cases}$$

$\mu = E[X]$

where m is a positive integer larger than 1. Find the variance of X.

$V[X] = E[X^2] - \mu^2$

SOLUTION. We find $\sigma^2 = V(X)$ using the variance computing formula. In Example 3.7, we computed

$$\mu = E(X) = \frac{m+1}{2}.$$

$\sum_{x=1}^{m} x^2 = \frac{m(m+1)(2m+1)}{6}$

We first find $E(X^2)$:

$$E(X^2) = \sum_{x \in R} x^2 p_X(x) = \sum_{x=1}^{m} x^2 \left(\frac{1}{m}\right) = \frac{1}{m}\sum_{x=1}^{m} x^2 = \frac{1}{m}\left[\frac{m(m+1)(2m+1)}{6}\right] = \frac{(m+1)(2m+1)}{6}.$$

$\sum_{x=1}^{m} x^3 = \left(\frac{m(m+1)}{2}\right)^2$

$E(X^3) = \sum_{x \in R} x^3 P_X(x) = \sum_{x=1}^{m} x^3 \left(\frac{1}{m}\right) = \frac{1}{m}\sum_{x=1}^{m} x^3 = \frac{1}{m} \times \left(\frac{m(m+1)}{2}\right)^2$

$= \frac{m(m+1)^2}{4}$

We have used the well-known fact that $\sum_{x=1}^{m} x^2 = m(m+1)(2m+1)/6$; this can be proven by induction. The variance of X is equal to

$$\begin{aligned} \sigma^2 &= E(X^2) - [E(X)]^2 \\ &= \frac{(m+1)(2m+1)}{6} - \left(\frac{m+1}{2}\right)^2 = \frac{m^2-1}{12}. \ \square \end{aligned}$$

EXERCISE: Find $\sigma^2 = V(Y)$ in Examples 3.5 and 3.8 (notes).

$(a+bE(Y))^2$

$=$

IMPORTANT RESULT: Let Y be a random variable (not necessarily a discrete random variable). Suppose that a and b are fixed constants. Then

$E(a+bY) = a+bE(Y)$

$$V(a+bY) = b^2V(Y).$$

$E((a+bY)^2) - [E(a+bY)]^2$

$= E[a^2 + 2abY + b^2Y^2] - (a^2 + 2abE(Y) + b^2[E(Y)]^2)$

REMARK: Taking $b = 0$ above, we see that $V(a) = 0$, for any constant a. This makes sense intuitively. The variance is a measure of variability for a random variable; a constant (such as a) does not vary. Also, by taking $a = 0$, we see that $V(bY) = b^2V(Y)$.

$= \cancel{a^2} + \cancel{2abE(Y)} + b^2E(Y^2)$

$- \cancel{a^2} - \cancel{2abE(Y)} - b^2(E(Y))^2$

$= b^2(E(Y^2) - (E(Y))^2)$

$= b^2 V(Y)$

3.5 Moment generating functions

TERMINOLOGY: Let Y be a discrete random variable with pmf $p_Y(y)$ and support R. The **moment generating function (mgf)** for Y, denoted by $m_Y(t)$, is given by

$$m_Y(t) = E(e^{tY}) = \sum_{y \in R} e^{ty} p_Y(y),$$

provided $E(e^{tY}) < \infty$ for all t in an open neighborhood about 0; i.e., there exists some $h > 0$ such that $E(e^{tY}) < \infty$ for all $t \in (-h, h)$. If $E(e^{tY})$ does not exist in an open neighborhood of 0, we say that the moment generating function does not exist.

TERMINOLOGY: We call $\mu'_k \equiv E(Y^k)$ the k**th moment** of the random variable Y:

$E(Y)$	1st moment (mean!)
$E(Y^2)$	2nd moment
$E(Y^3)$	3rd moment
$E(Y^4)$	4th moment
$\vdots$	$\vdots$