Expectations

Expectations. (See also Hays, Appendix B; Harnett, ch. 3).

A. The expected value of a random variable is the arithmetic mean of that variable, i.e. $E(X) = \mu$. As Hays notes, the idea of the expectation of a random variable began with probability theory in games of chance. Gamblers wanted to know their expected long-run winnings (or losings) if they played a game repeatedly. This term has been retained in mathematical statistics to mean the long-run average for any random variable over an indefinite number of trials or samplings.

B. Discrete case: The expected value of a discrete random variable, X, is found by multiplying each X-value by its probability and then summing over all values of the random variable. That is, if X is discrete,

$$E(X) = \sum_{All\ X} xp(x) = \mu_X$$

C. Continuous case: For a continuous variable X ranging over all the real numbers, the expectation is defined by

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

D. Variance of X: The variance of a random variable X is defined as the expected (average) squared deviation of the values of this random variable about their mean. That is,

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

In the discrete case, this is equivalent to

$$V(X) = \sigma^2 = \sum_{\text{All X}} (x - \mu)^2 P(x)$$

E. Standard deviation of X: The standard deviation is the positive square root of the variance, i.e.

$$SD(X) = \sigma = \sqrt{\sigma^2}$$

F. Examples.

1. Hayes (p. 96) gives the probability distribution for the number of spots appearing on two fair dice. Find the mean and variance of that distribution.

x	**p(x)**	**xp(x)**	**$(x - \mu_x)^2$**	**$(x - \mu_x)^2 p(x)$**
2	1/36	2/36	25	25/36
3	2/36	6/36	16	32/36
4	3/36	12/36	9	27/36
5	4/36	20/36	4	16/36
6	5/36	30/36	1	5/36
7	6/36	42/36	0	0
8	5/36	40/36	1	5/36
9	4/36	36/36	4	16/36
10	3/36	30/36	9	27/36
11	2/36	22/36	16	32/36
12	1/36	12/36	25	25/36

Σ xp(x) = 252/36 = 7 = μ_x. The variance σ^2 = 210/36 = 35/6 = 5 5/6. (NOTE: There is a simpler solution to this problem, which takes advantage of the independence of the two tosses.)

2. Consider our earlier coin tossing experiment. If we toss a coin three times, how many times do we expect it to come up heads? And, what is the variance of this distribution?

x	**p(x)**	**xp(x)**	**$(x - \mu_x)^2$**	**$(x - \mu_x)^2 p(x)$**
0	1/8	0	2.25	2.25/8
1	3/8	3/8	0.25	0.75/8
2	3/8	6/8	0.25	0.75/8
3	1/8	3/8	2.25	2.25/8

Σ xp(x) = 1.5. So (not surprisingly) if we toss a coin three times, we expect 1.5 heads. And, the variance = 6/8 = 3/4.

G. EXPECTATION RULES AND DEFINITIONS. a, b are any given constants. X, Y are random variables. The following apply. [NOTE: we'll use a few of these now and others will come in handy throughout the semester.]

1. **$E(X) = \mu_x = \Sigma\, xp(x)$ (discrete case)**

2. **$E(g(X)) = \Sigma\, g(x)p(x) = \mu_{g(X)}$ (discrete case)**

NOTE: g(X) is some function of X. So, for example, if X is discrete and $g(X) = X^2$, then $E(X^2) = \Sigma\, x^2p(x)$.

3. **$V(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2 = \sigma^2_X$**

4. **$E(a) = a$**

That is, the expectation of a constant is the constant, e.g. E(7) = 7

5. **$E(aX) = a * E(X)$**

e.g. if you multiple every value by 2, the expectation doubles.

6. **$E(a \pm X) = a \pm E(X)$**

e.g. if you add 7 to every case, the expectation will increase by 7

7a. **$E(a \pm bX) = a \pm bE(X)$**

7b. **$E[(a \pm X) * b] = (a \pm E(X)) * b$**

8. **$E(X + Y) = E(X) + E(Y)$.** (The expectation of a sum = the sum of the expectations. This rule extends as you would expect it to when there are more than 2 random variables, e.g. $E(X + Y + Z) = E(X) + E(Y) + E(Z)$)

9. **If X and Y are independent,**

   $E(XY) = E(X)E(Y)$. (This rule extends as you would expect it to for more than 2 random variables, e.g. $E(XYZ)=E(X)E(Y)E(Z)$.)

10. **$COV(X,Y) = E[(X - E(X)) * (Y - E(Y)] = E(XY) - E(X)E(Y)$**

    Question: What is COV(X,X)?

11. **If X and Y are independent,**

    $COV(X,Y) = 0$. (However, if $COV(X,Y) = 0$, this does not necessarily mean that X and Y are independent.)

12. **$V(a) = 0$**

A constant does not vary, so the variance of a constant is 0, e.g. V(7) = 0.

13. **$V(a \pm X) = V(X)$**

Adding a constant to a variable does not change its variance.

14. **$V(a \pm bX) = b^2 * V(X) = \sigma^2_{bX}$** [Proof is below]

15. **$V(X \pm Y) = V(X) + V(Y) \pm 2\, COV(X,Y) = \sigma^2_{X \pm Y}$**

16. **If X and Y are independent, $V(X \pm Y) = V(X) + V(Y)$**

    However, it is generally NOT TRUE that $V(XY) = V(X)V(Y)$

PROBLEMS: HINT. Keep in mind that μ_X and σ_X are constants.

1. Prove that $V(X) = E[(X - \mu_X)^2] = E(X^2) - \mu_X^2$. HINT: Rules 4, 5, and 8 are especially helpful here.

Solution.

Equation	Explanation
$E[(X - \mu_X)^2] =$	Original Formula for the variance.
$E(X^2 - 2X\mu_X + \mu_X^2) =$	Expand the square
$E(X^2) - E(2\mu_X X) + E(\mu_X^2) =$	Rule 8: $E(X + Y) = E(X) + E(Y)$. That is, the expectation of a sum = Sum of the expectations
$E(X^2) - 2\mu_X E(X) + \mu_X^2 =$	Rule 5: $E(aX) = a * E(X)$, i.e. Expectation of a constant times a variable = The constant times the expectation of the variable; and Rule 4: $E(a) = a$, i.e. Expectation of a constant = the constant
$E(X^2) - u_X^2$	Remember that $E(X) = \mu_X$, hence $2\mu_X E(X) = 2\mu_X^2$. QED.

2. Prove that $V(aX) = a^2 * V(X)$. HINT: Rules 3 and 5 are especially helpful.

Solution. Let $Y = aX$. Then,

Equation	Explanation
$V(Y) = E(Y^2) - E(Y)^2 =$	Rule 3: $V(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2 = \sigma^2_X$, i.e. Definition of the variance
$E(a^2 X^2) - E(aX)^2 =$	Substitute for Y. Since $Y = aX$, $Y^2 = a^2X^2$
$a^2 E(X^2) - a^2 E(X)^2 =$	Rule 5: $E(aX) = a * E(X)$, i.e. Expectation of a constant times a variable = The constant times the expectation of the variable
$a^2(E(X^2) - E(X)^2) =$	Factor out a^2
$a^2 V(X)$	Rule 3: Definition of the variance, i.e. $V(X) = E(X^2) - E(X)^2$. QED.

3. Let $Z = (X - \mu_X)/\sigma_X$. Find E(Z) and V(Z). HINT: Apply rules 7b and 14.

Solution. In this problem, $a = -\mu_X$, $b = 1/\sigma_X$.

Equation	Explanation
$E(Z) = E\left(\frac{X - \mu_X}{\sigma_X}\right) =$	Definition of Z
$\frac{E(X) - \mu_X}{\sigma_X} =$	Rule 7b: $E[(a \pm X) * b] = (a \pm E(X)) * b$. Remember, $a = -\mu_X$, $b = 1/\sigma_X$.
0	Remember $E(X) = \mu_X$, so the numerator = 0. QED

Intuitively, the above makes sense; subtract the mean from every case and the new mean becomes zero. Now, for the variance,

Equation	Explanation
$V(Z) = V\left(\frac{X - \mu_X}{\sigma_X}\right) =$	Definition of Z
$\frac{1}{\sigma_X^2} * V(X) =$	Rule 14: $V(a \pm bX) = b^2 * V(X) = \sigma^2_{bX}$. Remember, $b = 1/\sigma_X$
1	Remember, $V(X) = \sigma_X^2$, hence σ_X^2 appears in both the numerator and denominator. QED.

NOTE: This is called a z-score transformation. As we will see, such a transformation is extremely useful. Note that, if Z = 1, the score is one standard deviation above the mean.

4. Use a different method than the one presented earlier for finding the mean and variance for the number of heads obtained in 3 coin tosses.

Solution. Let $X_1 = 1$ if the first coin toss comes up heads, 0 otherwise. In this case, $X_1^2 = X_1$ (since $1^2 = 1$ and $0^2 = 0$). Let X_2 and X_3 be the corresponding random variables for the second and third tosses. This question is asking you to find $E(X_1 + X_2 + X_3)$ and $V(X_1 + X_2 + X_3)$. Note that

Formula	Explanation
$E(X_1) = E(X_2) = E(X_3) = .5$	Each coin has a 50% chance of a heads
$E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = .5 + .5 + .5 = 1.5.$	Rule 8: $E(X + Y) = E(X) + E(Y)$, i.e. Expectation of a sum = Sum of the Expectations
$V(X_1) = V(X_2) = V(X_3) = E(X_1^2) - E(X_1)^2 = .5 - .25 = .25.$	Rule 3: Definition of the variance. Since $X_1^2 = X_1$, $E(X_1^2) = E(X_1) = .5$; and $E(X_1)^2 = .5^2 = .25$
$V(X_1 + X_2 + X_3) = V(X_1) + V(X_2) + V(X_3) = .25 + .25 + .25 = .75$	Rule 16: If X and Y are independent, $V(X \pm Y) = V(X) + V(Y)$