# Expectation, Variance, and Standard Deviation of Discrete Random Variable and Covariance and Correlation

From Michael Barron Book (MB)

### 3.3.1 Expectation

*DEFINITION 3.5*

**Expectation** or **expected value** of a random variable $X$ is its mean, the average value.

**Expectation, discrete case**

$$\boxed{\mu = \mathbf{E}(X) = \sum_x x P(x)}$$

### 3.3.2 Expectation of a function

Often we are interested in another variable, $Y$, that is a function of $X$. For example, downloading time depends on the connection speed, profit of a computer store depends on the number of computers sold, and bonus of its manager depends on this profit. Expectation of $Y = g(X)$ is computed by a similar formula,

$$\mathbf{E}\{g(X)\} = \sum_x g(x)P(x). \tag{3.4}$$

**Remark:** Indeed, if $g$ is a one-to-one function, then $Y$ takes each value $y = g(x)$ with probability $P(x)$, and the formula for $\mathbf{E}(Y)$ can be applied directly. If $g$ is not one-to-one, then some values of $g(x)$ will be repeated in (3.4). However, they are still multiplied by the corresponding probabilities. When we add in (3.4), these probabilities are also added, thus each value of $g(x)$ is still multiplied by the probability $P_Y(g(x))$.

### 3.3.3 Properties

The following *linear* properties of expectations follow directly from (3.3) and (3.4). For *any* random variables $X$ and $Y$ and any non-random numbers $a$, $b$, and $c$, we have

**Properties of expectations**

$$\boxed{\begin{array}{lcl} \mathbf{E}(aX+bY+c) & = & a\,\mathbf{E}(X)+b\,\mathbf{E}(Y)+c \\ \text{In particular,} & & \\ \quad\mathbf{E}(X+Y) & = & \mathbf{E}(X)+\mathbf{E}(Y) \\ \quad\mathbf{E}(aX) & = & a\,\mathbf{E}(X) \\ \quad\mathbf{E}(c) & = & c \\ \text{For } \textbf{independent } X \text{ and } Y, & & \\ \quad\mathbf{E}(XY) & = & \mathbf{E}(X)\,\mathbf{E}(Y) \end{array}} \tag{3.5}$$

*DEFINITION 3.6*

**Variance** of a random variable is defined as the expected squared deviation from the mean. For discrete random variables, variance is

$$\sigma^2 = \mathrm{Var}(X) = \mathbf{E}\,(X - \mathbf{E}X)^2 = \sum_x (x - \mu)^2 P(x)$$

Variance can also be computed as

$$\mathrm{Var}(X) = \mathbf{E}(X^2) - \mu^2. \tag{3.6}$$

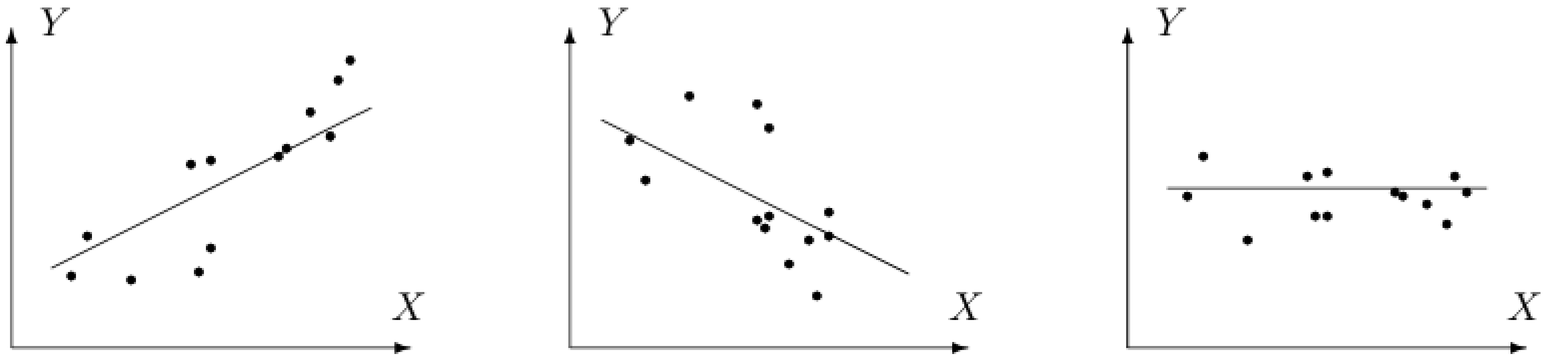
*DEFINITION 3.7*

**Standard deviation** is a square root of variance,

$$\sigma = \mathrm{Std}(X) = \sqrt{\mathrm{Var}(X)}$$

### 3.3.5 Covariance and correlation

Expectation, variance, and standard deviation characterize the distribution of a single random variable. Now we introduce measures of *association* of two random variables.




(a) $\mathrm{Cov}(X, Y) > 0$ (b) $\mathrm{Cov}(X, Y) < 0$ (c) $\mathrm{Cov}(X, Y) = 0$

FIGURE 3.5: *Positive, negative, and zero covariance.*

*DEFINITION 3.8*

Covariance $\sigma_{XY} = \mathrm{Cov}(X, Y)$ is defined as

$$\begin{aligned} \mathrm{Cov}(X, Y) &= \mathbf{E}\{(X - \mathbf{E}X)(Y - \mathbf{E}Y)\} \\ &= \mathbf{E}(XY) - \mathbf{E}(X)\,\mathbf{E}(Y) \end{aligned}$$

It summarizes interrelation of two random variables.

*DEFINITION 3.9*

**Correlation coefficient** between variables $X$ and $Y$ is defined as

$$\rho = \frac{\mathrm{Cov}(X, Y)}{(\,\mathrm{Std}X)(\,\mathrm{Std}Y)}$$

$$-1 \le \rho \le 1,$$

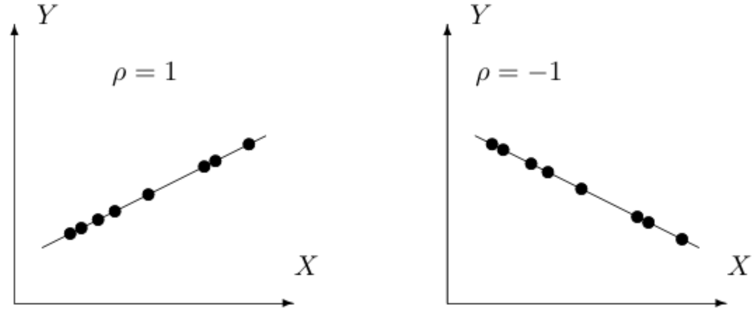



FIGURE 3.6: *Perfect correlation:* $\rho = \pm 1$.

**Example 3.6.** A program consists of two modules. The number of errors, $X$, in the first module and the number of errors, $Y$, in the second module have the joint distribution, $P(0,0) = P(0,1) = P(1,0) = 0.2$, $P(1,1) = P(1,2) = P(1,3) = 0.1$, $P(0,2) = P(0,3) = 0.05$. Find (a) the marginal distributions of $X$ and $Y$, (b) the probability of no errors in the first module, and (c) the distribution of the total number of errors in the program. Also, (d) find out if errors in the two modules occur independently.

Solution. It is convenient to organize the joint pmf of $X$ and $Y$ in a table. Adding rowwise and columnwise, we get the marginal pmfs,

| $P_{(X,Y)}(x,y)$ | | $y$ | | | | $P_X(x)$ |
|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | |
| $x$ | 0 | 0.20 | 0.20 | 0.05 | 0.05 | 0.50 |
| | 1 | 0.20 | 0.10 | 0.10 | 0.10 | 0.50 |
| $P_Y(y)$ | | 0.40 | 0.30 | 0.15 | 0.15 | 1.00 |

**Example 3.11.** Continuing Example 3.6, we compute

| $x$ | $P_X(x)$ | $xP_X(x)$ | $x - \mathbf{E}X$ | $(x - \mathbf{E}X)^2 P_X(x)$ |
|---|---|---|---|---|
| 0 | 0.5 | 0 | $-0.5$ | 0.125 |
| 1 | 0.5 | 0.5 | 0.5 | 0.125 |
| $\mu_X = 0.5$ | | | $\sigma_X^2 = 0.25$ | |

and (using the second method of computing variances)

| $y$ | $P_Y(y)$ | $yP_Y(y)$ | $y^2$ | $y^2 P_Y(y)$ |
|---|---|---|---|---|
| 0 | 0.4 | 0 | 0 | 0 |
| 1 | 0.3 | 0.3 | 1 | 0.3 |
| 2 | 0.15 | 0.3 | 4 | 0.6 |
| 3 | 0.15 | 0.45 | 9 | 1.35 |
| $\mu_Y = 1.05$ | | | $\mathbf{E}(Y^2) = 2.25$ | |

Result: $\mathrm{Var}(X) = 0.25$, $\mathrm{Var}(Y) = 2.25 - 1.05^2 = 1.1475$, $\mathrm{Std}(X) = \sqrt{0.25} = 0.5$, and $\mathrm{Std}(Y) = \sqrt{1.1475} = 1.0712$.

Also,

$$\mathbf{E}(XY) = \sum_x \sum_y xyP(x,y) = (1)(1)(0.1) + (1)(2)(0.1) + (1)(3)(0.1) = 0.6$$

(the other five terms in this sum are 0). Therefore,

$$\mathrm{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\,\mathbf{E}(Y) = 0.6 - (0.5)(1.05) = 0.075$$

and

$$\rho = \frac{\mathrm{Cov}(X,Y)}{(\mathrm{Std}X)(\mathrm{Std}Y)} = \frac{0.075}{(0.5)(1.0712)} = 0.1400.$$

Thus, the numbers of errors in two modules are *positively and not very strongly correlated.*
◊

## Exercises

**3.1.** A computer virus is trying to corrupt two files. The first file will be corrupted with probability 0.4. Independently of it, the second file will be corrupted with probability 0.3.

(a) Compute the probability mass function (pmf) of $X$, the number of corrupted files.

(b) Draw a graph of its cumulative distribution function (cdf).

**3.2.** Every day, the number of network blackouts has a distribution (probability mass function)

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(x)$ | 0.7 | 0.2 | 0.1 |

A small internet trading company estimates that each network blackout results in a \$500 loss. Compute expectation and variance of this company's daily loss due to blackouts.

**3.3.** There is one error in one of five blocks of a program. To find the error, we test three randomly selected blocks. Let $X$ be the number of errors in these three blocks. Compute $\mathbf{E}(X)$ and $\operatorname{Var}(X)$.

**3.4.** Tossing a fair die is an experiment that can result in any integer number from 1 to 6 with equal probabilities. Let $X$ be the number of dots on the top face of a die. Compute $\mathbf{E}(X)$ and $\operatorname{Var}(X)$.

**3.6.** A computer program contains one error. In order to find the error, we split the program into 6 blocks and test two of them, selected at random. Let $X$ be the number of errors in these blocks. Compute $\mathbf{E}(X)$.

**3.7.** The number of home runs scored by a certain team in one baseball game is a random variable with the distribution

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(x)$ | 0.4 | 0.4 | 0.2 |

The team plays 2 games. The number of home runs scored in one game is independent of the number of home runs in the other game. Let $Y$ be the *total* number of home runs. Find $\mathbf{E}(Y)$ and $\mathrm{Var}(Y)$.

**3.8.** A computer user tries to recall her password. She knows it can be one of 4 possible passwords. She tries her passwords until she finds the right one. Let $X$ be the number of wrong passwords she uses before she finds the right one. Find $\mathbf{E}(X)$ and $\mathrm{Var}(X)$.

**3.16.** The number of hardware failures, $X$, and the number of software failures, $Y$, on any day in a small computer lab have the joint distribution $P(x, y)$, where $P(0, 0) = 0.6$, $P(0, 1) = 0.1$, $P(1, 0) = 0.1$, $P(1, 1) = 0.2$. Based on this information,

(a) Are $X$ and $Y$ (hardware and software failures) independent?

(b) Compute $\mathbf{E}(X + Y)$, i.e., the expected total number of failures during 1 day.

Exercise questions mentioned in updated outline which was shared on GCR