Expectation and Variance

Expectation

3.3.1 Expectation

DEFINITION 3.5

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

We know that X can take different values with different probabilities. For this reason, its average value is *not* just the average of all its values. Rather, it is a *weighted average*.

Example

Example 3.7. Consider a variable that takes values 0 and 1 with probabilities $P(0) = P(1) = 0.5$. That is,

$$X = \begin{cases} 0 & \text{with probability } 1/2 \\ 1 & \text{with probability } 1/2 \end{cases}$$

Observing this variable many times, we shall see $X = 0$ about 50% of times and $X = 1$ about 50% of times. The average value of X will then be close to 0.5, so it is reasonable to have $\mathbf{E}(X) = 0.5$. $\Diamond$

Expectation of a Function

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,

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

Properties of Expectations

Properties of expectations

$$\mathbf{E}(aX + bY + c) = a\,\mathbf{E}(X) + b\,\mathbf{E}(Y) + c$$

In particular,

$$\begin{aligned} \mathbf{E}(X+Y) &= \mathbf{E}(X) + \mathbf{E}(Y) \\ \mathbf{E}(aX) &= a\,\mathbf{E}(X) \\ \mathbf{E}(c) &= c \end{aligned}$$

For **independent** X and Y,

$$\mathbf{E}(XY) = \mathbf{E}(X)\,\mathbf{E}(Y)$$

PROOF: The first property follows from the Addition Rule (3.2). For any X and Y,

$$\begin{aligned}
\mathbf{E}(aX+bY+c) &= \sum_x \sum_y (ax+by+c) P_{(X,Y)}(x,y) \\
&= \sum_x ax \sum_y P_{(X,Y)}(x,y) + \sum_y by \sum_x P_{(X,Y)}(x,y) + c \sum_x \sum_y P_{(X,Y)}(x,y) \\
&= a \sum_x x P_X(x) + b \sum_y y P_Y(y) + c.
\end{aligned}$$

The next three equalities are special cases. To prove the last property, we recall that $P_{(X,Y)}(x,y) = P_X(x)P_Y(y)$ for independent X and Y, and therefore,

$$\mathbf{E}(XY) = \sum_x \sum_y (xy) P_X(x) P_Y(y) = \sum_x x P_X(x) \sum_y y P_Y(y) = \mathbf{E}(X)\,\mathbf{E}(Y). \qquad \square$$

Remark: The last property in (3.5) holds for some dependent variables too, hence it cannot be used to verify independence of X and Y.

Variance and Standard Deviation

3.3.4 Variance and standard deviation

Expectation shows where the average value of a random variable is located, or where the variable is *expected* to be, plus or minus some error. How large could this "error" be, and how much can a variable *vary* around its expectation? Let us introduce some measures of variability.

Variability???? – Is it a good thing or bad thing

Example 3.10. Here is a rather artificial but illustrative scenario. Consider two users. One receives either 48 or 52 e-mail messages per day, with a 50-50% chance of each. The other receives either 0 or 100 e-mails, also with a 50-50% chance. What is a common feature of these two distributions, and how are they different?

We see that both users receive the same average number of e-mails:

$$\mathbf{E}(X) = \mathbf{E}(Y) = 50.$$

However, in the first case, the actual number of e-mails is always close to 50, whereas it always differs from it by 50 in the second case. The first random variable, X, is more stable; it has *low variability*. The second variable, Y, has *high variability*. $\Diamond$

Variance Discussion

DEFINITION 3.6

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

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

Standard Deviation? Why should we care about standard deviation rather than Variance

DEFINITION 3.7

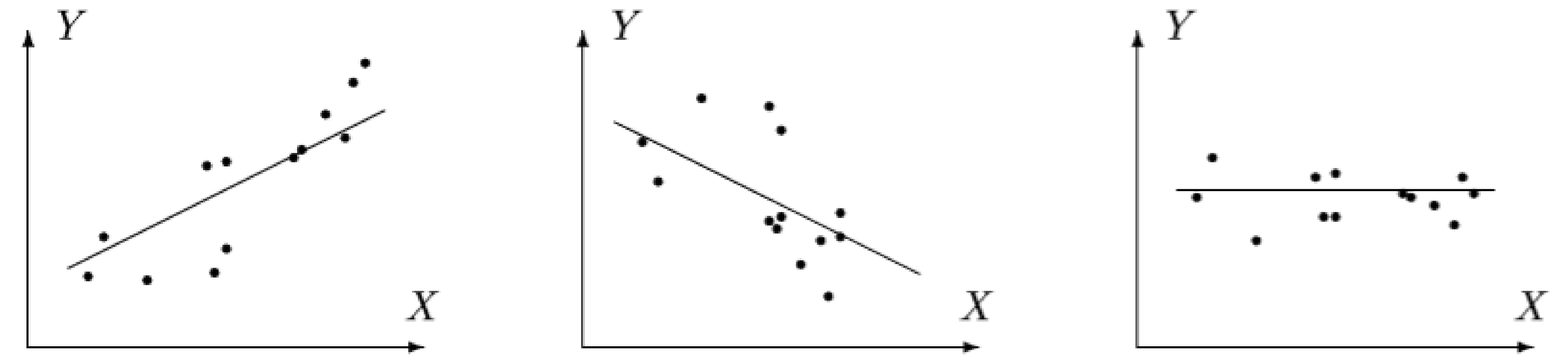
Standard deviation is a square root of variance,

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

Co-Variance???

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 3.4: *Positive, negative, and zero covariance.*

Co-Variance Contd..

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

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

Correlation Coefficient Discussion

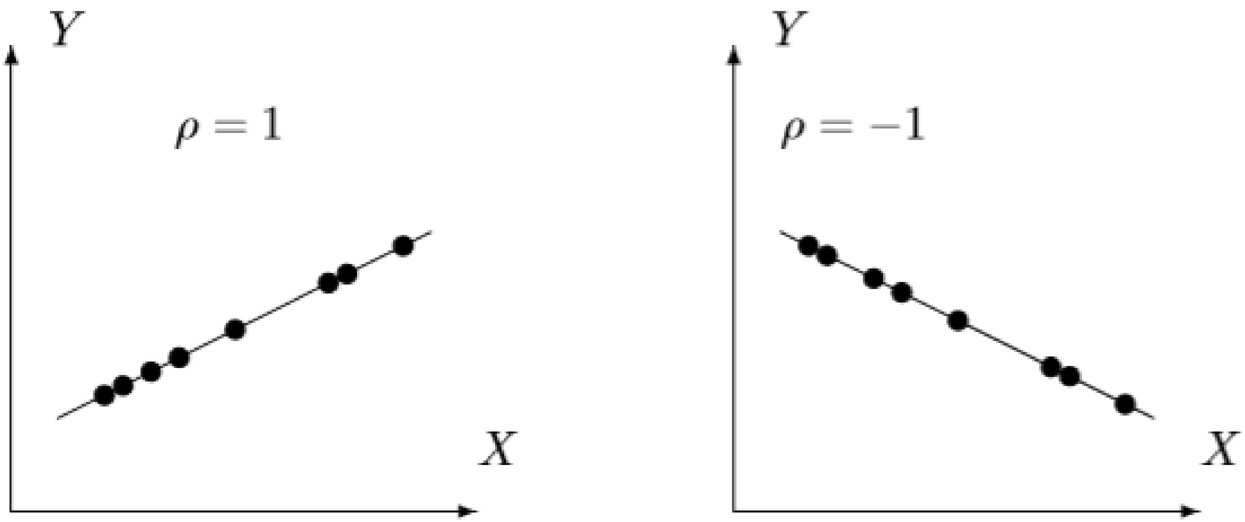


FIGURE 3.5: *Perfect correlation:* $\rho = \pm 1$.

How about $\rho = 0$. What does it mean

Properties of Variance, Co-Variance and Correlation Coefficients

3.3.6 Properties

The following properties of variances, covariances, and correlation coefficients hold for any random variables X, Y, Z, and W and any non-random numbers a, b, c and d.

Properties of variances and covariances

$$\mathrm{Var}(aX + bY + c) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y)$$

$$\begin{aligned} &\mathrm{Cov}(aX + bY, cZ + dW) \\ &\quad = ac\,\mathrm{Cov}(X, Z) + ad\,\mathrm{Cov}(X, W) + bc\,\mathrm{Cov}(Y, Z) + bd\,\mathrm{Cov}(Y, W) \end{aligned}$$

$$\begin{aligned} \mathrm{Cov}(X, Y) &= \mathrm{Cov}(Y, X) \\ \rho(X, Y) &= \rho(Y, X) \end{aligned}$$

In particular,

$$\begin{aligned} \mathrm{Var}(aX + b) &= a^2\,\mathrm{Var}(X) \\ \mathrm{Cov}(aX + b, cY + d) &= ac\,\mathrm{Cov}(X, Y) \\ \rho(aX + b, cY + d) &= \rho(X, Y) \end{aligned} \tag{3.7}$$

For independent X and Y,

$$\begin{aligned} \mathrm{Cov}(X, Y) &= 0 \\ \mathrm{Var}(X + Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) \end{aligned}$$