IES302 2011/2 Part I.6 Dr.Prapun

9 Multiple Random Variables

One is often interested not only in individual random variables, but also in relationships between two or more random variables. Furthermore, one often wishes to make inferences about one random variable on the basis of observations of other random variables.

Example 9.1. If the experiment is the testing of a new medicine, the researcher might be interested in cholesterol level, blood pressure, and the glucose level of a test person.

9.1 A Pair of Random Variables

Definition 9.2. If X and Y are random variables, we use the shorthand

$$\begin{aligned}[X \in B, Y \in C] &= [X \in B \text{ and } Y \in C] \\ &= \{\omega \in \Omega : X(\omega) \in B \text{ and } Y(\omega) \in C\} \\ &= \{\omega \in \Omega : X(\omega) \in B\ \} \cap \{\omega \in \Omega : Y(\omega) \in C\} \\ &= [X \in B] \cap [Y \in C].\end{aligned}$$

- Observe that the "," in $[X \in B, Y \in B]$ means "and".

Consequently,

$$\begin{aligned}P[X \in B, Y \in C] &= P[X \in B \text{ and } Y \in C] \\ &= P([X \in B] \cap [Y \in C]).\end{aligned}$$

Similarly, the concept of conditional probability can be straightforwardly applied to random variables via

$$P[X \in B | Y \in C] = P([X \in B] | [Y \in C]) = \frac{P([X \in B] \cap [Y \in C])}{P([Y \in C])} = \frac{P[X \in B, Y \in C]}{P[Y \in C]}.$$

Example 9.3. We also have

$$P[X = x, Y = y] = P[X = x \text{ and } Y = y],$$
$$P[X = x | Y = y] = \frac{P[X = x \text{ and } Y = y]}{P[Y = y]},$$

and

$$P[3 \le X < 4, Y < 1] = P[3 \le X < 4 \text{ and } Y < 1] = P[X \in [3,4) \text{ and } Y \in (-\infty, 1)].$$
$$P[3 \le X < 4 | Y < 1] = \frac{P[3 \le X < 4 \text{ and } Y < 1]}{P[Y < 1]}$$

Definition 9.4. ***Joint pmf***: If X and Y are two discrete random variables (defined on a same sample space with probability measure P), the probability mass function $p_{X,Y}(x, y)$ defined by

$$p_{X,Y}(x, y) = P[X = x, Y = y]$$

is called the ***joint probability mass function*** of X and Y. We can then evaluate $P[(X, Y) \in R]$ by $\sum_{(x,y):(x,y) \in R} p_{X,Y}(x, y)$.

Definition 9.5. The ***joint cdf*** of X and Y is defined by

$$F_{X,Y}(x, y) = P[X \le x, Y \le y].$$

Definition 9.6. The ***conditional pmf*** of X given Y is defined as

$$p_{X|Y}(x|y) = P[X = x | Y = y]$$

which gives

$$p_{X,Y}(x, y) = p_{X|Y}(x|y) p_Y(y) = p_{Y|X}(y|x) p_X(x).$$

Example 9.7. Toss-and-Roll Game:

Step 1 Toss a fair coin. Define X by

$$X = \begin{cases} 1, & \text{if result} = \text{H}, \\ 0, & \text{if result} = \text{T}. \end{cases}$$

Step 2 You have two dice, Dice 1 and Dice 2. Dice 1 is fair. Dice 2 is unfair with $p(1) = p(2) = p(3) = \frac{2}{9}$ and $p(4) = p(5) = p(6) = \frac{1}{9}$.

(i) If $X = 0$, roll Dice 1.

(ii) If $X = 1$, roll Dice 2.

Record the result as Y.

9.8. Suppose at $x = a$, we have $p_X(a) = 0$. Then $p_{X,Y}(a, y) = 0$ for any y.

Similarly, suppose at $y = a$, we have $p_Y(a) = 0$. Then $p_{X,Y}(x, a) = 0$ for any x.

Definition 9.9. When X and Y take finitely many values (have finite supports), say $x_1, \dots, x_m$ and $y_1, \dots, y_n$, respectively, we can arrange the probabilities $p_{X,Y}(x_i, y_j)$ in the $m \times n$ matrix

$$\begin{bmatrix} p_{X,Y}(x_1, y_1) & p_{X,Y}(x_1, y_2) & \dots & p_{X,Y}(x_1, y_n) \\ p_{X,Y}(x_2, y_1) & p_{X,Y}(x_2, y_2) & \dots & p_{X,Y}(x_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{X,Y}(x_m, y_1) & p_{X,Y}(x_m, y_2) & \dots & p_{X,Y}(x_m, y_n) \end{bmatrix}.$$

- We shall call this matrix the ***joint pmf matrix***.
- The sum of all the entries in the matrix is one.

- The sum of the entries in the ith row is $p_X(x_i)$, and the sum of the entries in the jth column is $p_Y(y_j)$:

$$p_X(x_i) = \sum_{j=1}^{n} p_{X,Y}(x_i, y_j) \tag{17}$$

$$p_Y(y_j) = \sum_{i=1}^{m} p_{X,Y}(x_i, y_j) \tag{18}$$

9.10. *Summary*: From the joint pmf, we can find $p_X(x)$ and $p_Y(y)$ by

$$p_X(x) = \sum_{y} p_{X,Y}(x, y) \tag{19}$$

$$p_Y(y) = \sum_{x} p_{X,Y}(x, y) \tag{20}$$

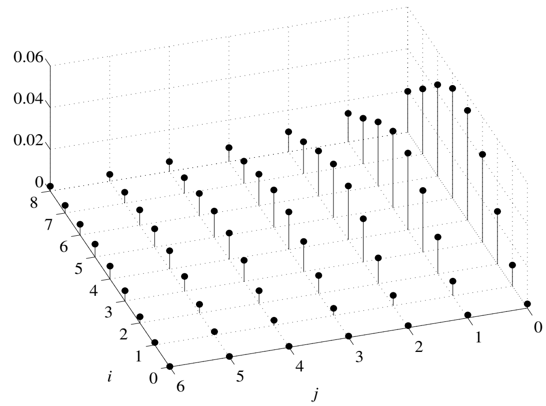

Figure 3: Example of the plot of a joint pmf. [7, Fig. 2.8]

In this setting, $p_X(x)$ and $p_Y(y)$ are call the ***marginal pmfs*** (to distinguish them from the joint one).

In MATLAB, if we define the joint pmf matrix as P_XY, then the marginal pmf (row) vectors p_X and p_Y can be found by

```
p_X = (sum(P_XY,2))'
p_Y = (sum(P_XY,1))
```

Example 9.11. Consider the following joint pmf matrix

Exercise 9.12 (F2011)**.** Random variables X and Y have the following joint pmf

$$p_{X,Y}(x,y) = \begin{cases} c(x+y), & x \in \{1,3\} \text{ and } y \in \{2,4\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Check that $c = 1/20$.

(b) Find $P\left[X^2 + Y^2 = 13\right]$.

(c) Find $p_X(x)$.

(d) Find $\mathbb{E}X$.

(e) Find $p_{Y|X}(y|1)$. Note that your answer should be of the form

$$p_{Y|X}(y|1) = \begin{cases} ?, & y = 2, \\ ?, & y = 4, \\ 0, & \text{otherwise}. \end{cases}$$

(f) Find $p_{Y|X}(y|3)$.

Definition 9.13. Two random variables X and Y are said to be ***identically distributed*** if, for every B, $P[X \in B] = P[Y \in B]$.

9.14. The following statements are equivalent:

(a) Random variables X and Y are ***identically distributed***.

(b) For every B, $P[X \in B] = P[Y \in B]$

(c) $p_X(c) = p_Y(c)$ for all c

(d) $F_X(c) = F_Y(c)$ for all c

Definition 9.15. Two random variables X and Y are said to be ***independent*** if the events $[X \in B]$ and $[Y \in C]$ are independent for all sets B and C.

9.16. The following statements are equivalent:

(a) Random variables X and Y are ***independent***.

(b) $[X \in B] \perp\!\!\!\perp [Y \in C]$ for all B, C.

(c) $P[X \in B, Y \in C] = P[X \in B] \times P[Y \in C]$ for all B, C.

(d) $p_{X,Y}(x, y) = p_X(x) \times p_Y(y)$ for all x, y.

(e) $F_{X,Y}(x, y) = F_X(x) \times F_Y(y)$ for all x, y.

Definition 9.17. Two random variables X and Y are said to be ***independent and identically distributed (i.i.d.)*** if X and Y are both independent and identically distributed.

Example 9.18. Suppose the pmf of a random variable X is given by

$$p_X(x) = \begin{cases} 1/4, & x = 3, \\ \alpha, & x = 4, \\ 0, & \text{otherwise.} \end{cases}$$

Let Y be another random variable. Assume that X and Y are i.i.d.

Find

(a) α,

(b) the pmf of Y, and

(c) the joint pmf of X and Y.

Example 9.19. Consider a pair of random variables X and Y whose joint pmf is given by

$$p_{X,Y}(x,y) = \begin{cases} 1/15, & x=3, y=1, \\ 2/15, & x=4, y=1, \\ 4/15, & x=3, y=3, \\ \beta, & x=4, y=3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Are X and Y identically distributed?

(b) Are X and Y independent?

9.2 Extending the Definitions to Multiple RVs

Definition 9.20. Joint pmf:

$$p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P\left[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right].$$

Joint cdf:

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P\left[X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n\right].$$

Definition 9.21. ***Identically distributed*** random variables: The following statements are equivalent.

(a) Random variables $X_1, X_2, \ldots$ are ***identically distributed***

(b) For every B, $P\left[X_j \in B\right]$ does not depend on j.

(c) $p_{X_i}(x) = p_{X_j}(x)$ for all x, i, j.

(d) $F_{X_i}(x) = F_{X_j}(x)$ for all x, i, j.

Definition 9.22. ***Independence*** among finite number of random variables: The following statements are equivalent.

(a) $X_1, X_2, \ldots, X_n$ are ***independent***

(b) $[X_1 \in B_1], [X_2 \in B_2], \ldots, [X_n \in B_n]$ are independent, for all $B_1, B_2, \ldots, B_n$.

(c) $P\left[X_i \in B_i, \forall i\right] = \prod_{i=1}^{n} P\left[X_i \in B_i\right]$, for all $B_1, B_2, \ldots, B_n$.

(d) $p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$ for all $x_1, x_2, \ldots, x_n$.

(e) $F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$ for all $x_1, x_2, \ldots, x_n$.

Example 9.23. Toss a coin n times. For the ith toss, let

$$X_i = \begin{cases} 1, & \text{if H happens on the } i\text{th toss,} \\ 0, & \text{if T happens on the } i\text{th toss.} \end{cases}$$

We then have a collection of i.i.d. random variables $X_1, X_2, X_3, \ldots, X_n$.

Example 9.24. Roll a dice n times. Let N_i be the result of the ith roll. We then have another collection of i.i.d. random variables $N_1, N_2, N_3, \ldots, N_n$.

Example 9.25. Let X_1 be the result of tossing a coin. Set $X_2 = X_3 = \cdots = X_n = X_1$.

9.26. If $X_1, X_2, \ldots, X_n$ are independent, then so is any subcollection of them.

9.27. For i.i.d. $X_i \sim$ Bernoulli(p), $Y = X_1 + X_2 + \cdots + X_n$ is $\mathcal{B}(n, p)$.

Definition 9.28. A ***pairwise independent*** collection of random variables is a collection of random variables any two of which are independent.

(a) Any collection of (mutually) independent random variables is pairwise independent

(b) Some pairwise independent collections are not independent. See Example (9.29).

Example 9.29. Let suppose X, Y, and Z have the following joint probability distribution: $p_{X,Y,Z}(x, y, z) = \frac{1}{4}$ for $(x, y, z) \in \{(0,0,0), (0,1,1), (1,0,1), (1,1,0)\}$. This, for example, can be constructed by starting with independent X and Y that are Bernoulli-$\frac{1}{2}$. Then set $Z = X \oplus Y = X + Y \mod 2$.

(a) X, Y, Z are pairwise independent.

(b) X, Y, Z are not independent.

9.3 Function of Discrete Random Variables

9.30. For discrete random variable X, the pmf of a derived random variable $Y = g(X)$ is given by

$$p_Y(y) = \sum_{x: g(x)=y} p_X(x).$$

Similarly, for discrete random variables X and Y, the pmf of a derived random variable $Z = g(X,Y)$ is given by

$$p_Z(z) = \sum_{(x,y): g(x,y)=z} p_{X,Y}(x,y).$$

Example 9.31. Suppose the joint pmf of X and Y is given by

$$p_{X,Y}(x,y) = \begin{cases} 1/15, & x=0, y=0, \\ 2/15, & x=1, y=0, \\ 4/15, & x=0, y=1, \\ 8/15, & x=1, y=1, \\ 0, & \text{otherwise.} \end{cases}$$

Let $Z = X + Y$. Find the pmf of Z.

9.4 Expectation of function of discrete random variables

9.32. Suppose X is a discrete random variable.

$$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x).$$

Similarly,

$$\mathbb{E}[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y).$$

These are called the **law/rule of the lazy statistician** (LOTUS) [20, Thm 3.6 p 48],[7, p. 149] because it is so much easier to use the above formula than to first find the pmf of $g(X)$ or $g(X,Y)$. It is also called **substitution rule** [19, p 271].

	Discrete
$P[X \in B]$	$\sum_{x \in B} p_X(x)$
$P[(X,Y) \in R]$	$\sum_{(x,y):(x,y) \in R} p_{X,Y}(x,y)$
Joint to Marginal: (Law of Total Prob.)	$p_X(x) = \sum_y p_{X,Y}(x,y)$ $p_Y(y) = \sum_x p_{X,Y}(x,y)$
$P[X > Y]$	$\sum_x \sum_{y:\, y<x} p_{X,Y}(x,y)$ $= \sum_y \sum_{x:\, x>y} p_{X,Y}(x,y)$
$P[X = Y]$	$\sum_x p_{X,Y}(x,x)$
$X \perp\!\!\!\perp Y$	$p_{X,Y}(x,y) = p_X(x) p_Y(y)$
Conditional	$p_{X\|Y}(x\|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$
$\mathbb{E}[g(X,Y)]$	$\sum_x \sum_y g(x,y) p_{X,Y}(x,y)$

Table 4: Joint pmf: A Summary

9.33. $\mathbb{E}[\cdot]$ is a **linear** operator: $\mathbb{E}[aX + bY] = a\mathbb{E}X + b\mathbb{E}Y$.

(a) Homogeneous: $\mathbb{E}[cX] = c\mathbb{E}X$

(b) Additive: $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$

(c) Extension: $\mathbb{E}\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} c_i \mathbb{E} X_i$.

Example 9.34. Recall from 9.27 that when i.i.d. $X_i \sim \text{Bernoulli}(p)$, $Y = X_1 + X_2 + \cdots X_n$ is $\mathcal{B}(n,p)$. Also, from Example 8.47, we have $\mathbb{E}X_i = p$. Hence,

$$\mathbb{E}Y = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] = \sum_{i=1}^{n} p = np.$$

Therefore, the expectation of a binomial random variable with parameters n and p is np.

Example 9.35. A binary communication link has bit-error probability p. What is the expected number of bit errors in a transmission of n bits?

Theorem 9.36 (Expectation and Independence)**.** Two random variables X and Y are independent if and only if

$$\mathbb{E}\left[h(X)g(Y)\right] = \mathbb{E}\left[h(X)\right]\mathbb{E}\left[g(Y)\right]$$

for all functions h and g.

- In other words, X and Y are independent if and only if for every pair of functions h and g, the expectation of the product $h(X)g(Y)$ is equal to the product of the individual expectations.

- One special case is that

$$X \perp\!\!\!\perp Y \quad \text{implies} \quad \mathbb{E}\left[XY\right] = \mathbb{E}X \times \mathbb{E}Y. \tag{21}$$

  However, independence means more than this property. In other words, having $\mathbb{E}\left[XY\right] = (\mathbb{E}X)(\mathbb{E}Y)$ does not necessarily imply $X \perp\!\!\!\perp Y$. See Example 9.47.

9.37. It is useful to incorporate what we have just learned about independence into the definition that we already have.

The following statements are equivalent:

(a) Random variables X and Y are ***independent***.

(b) $[X \in B] \perp\!\!\!\perp [Y \in C]$ for all B, C.

(c) $P[X \in B, Y \in C] = P[X \in B] \times P[Y \in C]$ for all B, C.

(d) $p_{X,Y}(x, y) = p_X(x) \times p_Y(y)$ for all x, y.

(e) $F_{X,Y}(x, y) = F_X(x) \times F_Y(y)$ for all x, y.

(f)

Exercise 9.38 (F2011)**.** Suppose X and Y are i.i.d. with $\mathbb{E}X = \mathbb{E}Y = 1$ and $\operatorname{Var} X = \operatorname{Var} Y = 2$. Find $\operatorname{Var}[XY]$.

9.39. To quantify the amount of dependence between two random variables, we may calculate their ***mutual information***. This quantity is crucial in the study of digital communications and information theory. However, in introductory probability class (and introductory communication class), it is traditionally omitted.

9.5 Linear Dependence

Definition 9.40. Given two random variables X and Y, we may calculate the following quantities:

(a) **Correlation**: $\mathbb{E}[XY]$.

(b) **Covariance**: $\operatorname{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$.

(c) **Correlation coefficient**: $\rho_{X,Y} = \frac{\operatorname{Cov}[X,Y]}{\sigma_X \sigma_Y}$

Exercise 9.41 (F2011)**.** Continue from Exercise 9.12.

(a) Find $\mathbb{E}[XY]$.

(b) Check that $\operatorname{Cov}[X, Y] = -\frac{1}{25}$.

9.42. $\text{Cov}[X,Y] = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y$

- Note that $\text{Var}\, X = \text{Cov}[X,X]$.

9.43. $\text{Var}[X+Y] = \text{Var}\, X + \text{Var}\, Y + 2\text{Cov}[X,Y]$

Definition 9.44. X and Y are said to be ***uncorrelated*** if and only if $\text{Cov}[X,Y] = 0$.

9.45. The following statements are equivalent:

(a) X and Y are ***uncorrelated***.

(b) $\text{Cov}[X,Y] = 0$.

(c) $\mathbb{E}[XY] = \mathbb{E}X\mathbb{E}Y$.

(d)

9.46. If $X \perp\!\!\!\perp Y$, then $\text{Cov}[X,Y] = 0$. The converse is not true. Being uncorrelated does not imply independence.

Example 9.47. Let X be uniform on $\{\pm 1, \pm 2\}$ and $Y = |X|$.

Example 9.48. Suppose two fair dice are tossed. Denote by the random variable V_1 the number appearing on the first die and by the random variable V_2 the number appearing on the second die. Let $X = V_1 + V_2$ and $Y = V_1 - V_2$.

(a) X and Y are not independent.

(b) $\mathbb{E}[XY] = \mathbb{E}X\mathbb{E}Y$.

Definition 9.49. ***Correlation coefficient***:

$$\begin{aligned}\rho_{X,Y} &= \frac{\operatorname{Cov}[X,Y]}{\sigma_X \sigma_Y} \\ &= \mathbb{E}\left[\left(\frac{X - \mathbb{E}X}{\sigma_X}\right)\left(\frac{Y - \mathbb{E}Y}{\sigma_Y}\right)\right] = \frac{\mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y}{\sigma_X \sigma_Y}.\end{aligned}$$

- $\rho_{X,Y}$ is dimensionless
- $\rho_{X,X} = 1$
- $\rho_{X,Y} = 0$ if and only if X and Y are uncorrelated.

9.50. Linear Dependence and Cauchy-Schwartz Inequality

(a) If $Y = aX + b$, then $\rho_{X,Y} = \operatorname{sign}(a) = \begin{cases} 1, & a > 0 \\ -1, & a < 0. \end{cases}$

- To be rigorous, we should also require that $\sigma_X > 0$ and $a \neq 0$.

(b) Cauchy-Schwartz Inequality:

$$(\operatorname{Cov}[X,Y])^2 \leq \sigma_X^2 \sigma_Y^2$$

(c) This implies $|\rho_{X,Y}| \leq 1$. In other words, $\rho_{XY} \in [-1, 1]$.

(d) When $\sigma_Y, \sigma_X > 0$, equality occurs if and only if the following conditions holds

$$\begin{aligned}&\equiv \exists a \neq 0 \text{ such that } (X - \mathbb{E}X) = a(Y - \mathbb{E}Y) \\ &\equiv \exists a \neq 0 \text{ and } b \in \mathbb{R} \text{ such that } X = aY + b \\ &\equiv \exists c \neq 0 \text{ and } d \in \mathbb{R} \text{ such that } Y = cX + d \\ &\equiv |\rho_{XY}| = 1\end{aligned}$$

In which case, $|a| = \frac{\sigma_X}{\sigma_Y}$ and $\rho_{XY} = \frac{a}{|a|} = \operatorname{sgn} a$. Hence, ρ_{XY} is used to quantify **linear dependence** between X and Y. The closer $|\rho_{XY}|$ to 1, the higher degree of linear dependence between X and Y.

Example 9.51. [19, Section 5.2.3] Consider an important fact that *investment experience* supports: spreading investments over a variety of funds (diversification) diminishes risk. To illustrate, imagine that the random variable X is the return on every invested dollar in a local fund, and random variable Y is the return on every invested dollar in a foreign fund. Assume that random variables X and Y are i.i.d. with expected value 0.15 and standard deviation 0.12.

If you invest all of your money, say c, in either the local or the foreign fund, your return R would be cX or cY.

- The expected return is $\mathbb{E}R = c\mathbb{E}X = c\mathbb{E}Y = 0.15c$.
- The standard deviation is $c\sigma_X = c\sigma_Y = 0.12c$

Now imagine that your money is equally distributed over the two funds. Then, the return R is $\frac{1}{2}cX + \frac{1}{2}cY$. The expected return is $\mathbb{E}R = \frac{1}{2}c\mathbb{E}X + \frac{1}{2}c\mathbb{E}Y = 0.15c$. Hence, the expected return remains at 15%. However,

$$\operatorname{Var} R = \operatorname{Var}\left[\frac{c}{2}(X+Y)\right] = \frac{c^2}{4}\operatorname{Var} X + \frac{c^2}{4}\operatorname{Var} Y = \frac{c^2}{2} \times 0.12.$$

So, the standard deviation is $\frac{0.12}{\sqrt{2}}c \approx 0.0849c$.

In comparison with the distributions of X and Y, the pmf of $\frac{1}{2}(X+Y)$ is concentrated more around the expected value. The centralization of the distribution as random variables are averaged together is a manifestation of the central limit theorem.

9.52. [19, Section 5.2.3] Example 9.51 is based on the assumption that return rates X and Y are independent from each other. In the world of investment, however, risks are more commonly reduced by combining negatively correlated funds (two funds are negatively correlated when one tends to go up as the other falls).

This becomes clear when one considers the following hypothetical situation. Suppose that two stock market outcomes ω_1 and ω_2 are possible, and that each outcome will occur with a probability of $\frac{1}{2}$ Assume that domestic and foreign fund returns X and Y are determined by $X(\omega_1) = Y(\omega_2) = 0.25$ and $X(\omega_2) = Y(\omega_1) = -0.10$. Each of the two funds then has an expected return of 7.5%, with equal probability for actual returns of 25% and .10%. The random variable $Z = \frac{1}{2}(X + Y)$ satisfies $Z(\omega_1) = Z(\omega_2) = 0.075$. In other words, Z is equal to 0.075 with certainty. This means that an investment that is equally divided between the domestic and foreign funds has a guaranteed return of 7.5%.

Example 9.53. The input X and output Y of a system subject to random perturbations are described probabilistically by the following joint pmf matrix:

x \ y	2	4	5
1	0.02	0.10	0.08
3	0.08	0.32	0.40

(a) Evaluate the following quantities.

(i) $\mathbb{E}X$

(ii) $P[X = Y]$

(iii) $P[XY < 6]$

(iv) $\mathbb{E}[(X-3)(Y-2)]$

(v) $\mathbb{E}\left[X(Y^3 - 11Y^2 + 38Y)\right]$

(vi) $\text{Cov}[X, Y]$

(vii) $\rho_{X,Y}$

(b) Calculate the following quantities using what you got from part (a).

(i) $\text{Cov}[3X + 4, 6Y - 7]$

(ii) $\rho_{3X+4,6Y-7}$

(iii) $\text{Cov}[X, 6X - 7]$

(iv) $\rho_{X,6X-7}$

Solution:

(a)

(i) $\mathbb{E}X = 2.6$

(ii) $P[X = Y] = 0$

(iii) $P[XY < 6] = 0.2$

(iv) $\mathbb{E}\left[(X-3)(Y-2)\right] = -0.88$

(v) $\mathbb{E}\left[X(Y^3 - 11Y^2 + 38Y)\right] = 104$

(vi) $\text{Cov}\left[X, Y\right] = 0.032$

(vii) $\rho_{X,Y} = 0.0447$

(b)

(i) Note that

$$\begin{aligned}
\text{Cov}\left[aX+b, cY+d\right] &= \mathbb{E}\left[\left((aX+b) - \mathbb{E}\left[aX+b\right]\right)\left((cY+d) - \mathbb{E}\left[cY+d\right]\right)\right] \\
&= \mathbb{E}\left[\left((aX+b) - (a\mathbb{E}X+b)\right)\left((cY+d) - (c\mathbb{E}Y+d)\right)\right] \\
&= \mathbb{E}\left[(aX - a\mathbb{E}X)(cY - c\mathbb{E}Y)\right] \\
&= ac\mathbb{E}\left[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\right] \\
&= ac\text{Cov}\left[X, Y\right].
\end{aligned}$$

Hence, $\text{Cov}\left[3X+4, 6Y-7\right] = 3 \times 6 \times \text{Cov}\left[X, Y\right] \approx 3 \times 6 \times 0.032 \approx \boxed{0.576}$.

(ii) Note that

$$\begin{aligned}
\rho_{aX+b,cY+d} &= \frac{\text{Cov}\left[aX+b, cY+d\right]}{\sigma_{aX+b}\sigma_{cY+d}} \\
&= \frac{ac\text{Cov}\left[X, Y\right]}{|a|\sigma_X|c|\sigma_Y} = \frac{ac}{|ac|}\rho_{X,Y} = \text{sign}(ac) \times \rho_{X,Y}.
\end{aligned}$$

Hence, $\rho_{3X+4,6Y-7} = \text{sign}(3 \times 4)\rho_{X,Y} = \rho_{X,Y} = \boxed{0.0447}$.

(iii) $\text{Cov}\left[X, 6X-7\right] = 1 \times 6 \times \text{Cov}\left[X, X\right] = 6 \times \text{Var}[X] \approx \boxed{3.84}$.

(iv) $\rho_{X,6X-7} = \text{sign}(1 \times 6) \times \rho_{X,X} = \boxed{1}$.