EXAMPLE 20 Let the three-dimensional random variable (X_1, X_2, X_3) have the density

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = 8x_1x_2x_3 I_{(0,1)}(x_1)I_{(0,1)}(x_2)I_{(0,1)}(x_3).$$

Suppose we want to find (i) $\mathscr{E}[3X_1 + 2X_2 + 6X_3]$, (ii) $\mathscr{E}[X_1X_2X_3]$, and (iii) $\mathscr{E}[X_1X_2]$. For (i) we have $g(x_1, x_2, x_3) = 3x_1 + 2x_2 + 6x_3$ and obtain

$$\mathscr{E}[g(X_1, X_2, X_3)] = \mathscr{E}[3X_1 + 2X_2 + 6X_3]$$
$$= \int_0^1 \int_0^1 \int_0^1 (3x_1 + 2x_2 + 6x_3)8x_1x_2x_3\, dx_1\, dx_2\, dx_3 = \tfrac{22}{3}.$$

For (ii), we get

$$\mathscr{E}[X_1X_2X_3] = \int_0^1 \int_0^1 \int_0^1 8x_1^2x_2^2x_3^2\, dx_1\, dx_2\, dx_3 = \tfrac{8}{27},$$

and for (iii) we get $\mathscr{E}[X_1X_2] = \frac{4}{9}$. ////

The following remark, the proof of which is left to the reader, displays a property of joint expectation. It is a generalization of (ii) in Theorem 3 of Chap. II.

Remark $\mathscr{E}\left[\sum_1^m c_i g_i(X_1, \ldots, X_k)\right] = \sum_1^m c_i \mathscr{E}[g_i(X_1, \ldots, X_k)]$ for constants $c_1, c_2, \ldots, c_m$. ////

4.2 Covariance and Correlation Coefficient

Definition 19 Covariance Let X and Y be any two random variables defined on the same probability space. The *covariance* of X and Y, denoted by cov $[X, Y]$ or $\sigma_{X,Y}$, is defined as

$$\text{cov}\,[X, Y] = \mathscr{E}[(X - \mu_X)(Y - \mu_Y)] \tag{21}$$

provided that the indicated expectation exists. ////

Definition 20 Correlation coefficient The *correlation coefficient*, denoted by $\rho[X, Y]$ or $\rho_{X,Y}$, of random variables X and Y is defined to be

$$\rho_{X,Y} = \frac{\text{cov}\,[X, Y]}{\sigma_X \sigma_Y} \tag{22}$$

provided that cov $[X, Y]$, σ_X, and σ_Y exist, and $\sigma_X > 0$ and $\sigma_Y > 0$. ////

Remark $\text{cov}[X, Y] = \mathscr{E}[(X - \mu_X)(Y - \mu_Y)] = \mathscr{E}[XY] - \mu_X \mu_Y$.

PROOF
$$\begin{aligned}\mathscr{E}[(X - \mu_X)(Y - \mu_Y)] &= \mathscr{E}[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \\ &= \mathscr{E}[XY] - \mu_X \mathscr{E}[Y] - \mu_Y \mathscr{E}[X] + \mu_X \mu_Y \\ &= \mathscr{E}[XY] - \mu_X \mu_Y. \qquad ////\end{aligned}$$

EXAMPLE 21 Find $\rho_{X,Y}$ for X, the number on the first, and Y, the larger of the two numbers, in the experiment of tossing two tetrahedra. We would expect that $\rho_{X,Y}$ is positive since when X is large, Y tends to be large too. We calculated $\mathscr{E}[XY]$, $\mathscr{E}[X]$, and $\mathscr{E}[Y]$ in Example 18 and obtained $\mathscr{E}[XY] = \frac{135}{16}$, $\mathscr{E}[X] = \frac{5}{2}$, and $\mathscr{E}[Y] = \frac{50}{16}$. Thus $\text{cov}[X, Y] = \frac{135}{16} - \frac{5}{2} \cdot \frac{50}{16} = \frac{10}{16}$. Now $\mathscr{E}[X^2] = \frac{30}{4}$ and $\mathscr{E}[Y^2] = \frac{170}{16}$; hence $\text{var}[X] = \frac{5}{4}$ and $\text{var}[Y] = \frac{55}{64}$. So,

$$\rho_{X,Y} = \frac{\frac{10}{16}}{\sqrt{\frac{5}{4}}\sqrt{\frac{55}{64}}} = \frac{2}{\sqrt{11}}. \qquad ////$$

EXAMPLE 22 Find $\rho_{X,Y}$ for X and Y if $f_{X,Y}(x, y) = (x + y)I_{(0,1)}(x)I_{(0,1)}(y)$. We saw that $\mathscr{E}[XY] = \frac{1}{3}$ and $\mathscr{E}[X] = \mathscr{E}[Y] = \frac{7}{12}$ in Example 19. Now $\mathscr{E}[X^2] = \mathscr{E}[Y^2] = \frac{5}{12}$; hence $\text{var}[X] = \text{var}[Y] = \frac{11}{144}$. Finally

$$\rho_{X,Y} = \frac{\frac{1}{3} - \frac{49}{144}}{\frac{11}{144}} = -\frac{1}{11}.$$

4.3 Conditional Expectations

In the following chapters we shall have occasion to find the expected value of random variables in conditional distributions, or the expected value of one random variable given the value of another.

Definition 21 Conditional expectation Let (X, Y) be a two-dimensional random variable and $g(\cdot,\cdot)$, a function of two variables. The *conditional expectation* of $g(X, Y)$ given $X = x$, denoted by $\mathscr{E}[g(X, Y) | X = x]$, is defined to be

$$\mathscr{E}[g(X, Y) | X = x] = \int_{-\infty}^{\infty} g(x, y) f_{Y|X}(y|x)\, dy \qquad (23)$$

if (X, Y) are jointly continuous, and

$$\mathscr{E}[g(X,Y)\,|\,X=x] = \sum g(x, y_j) f_{Y|X}(y_j|x) \qquad (24)$$

if (X, Y) are jointly discrete, where the summation is over all possible values of Y. ////

In particular, if $g(x, y) = y$, we have defined $\mathscr{E}[Y|X = x] = \mathscr{E}[Y|x]$. $\mathscr{E}[Y|x]$ and $\mathscr{E}[g(X, Y)|x]$ are functions of x. Note that this definition can be generalized to more than two dimensions. For example, let $(X_1, \ldots, X_k, Y_1, \ldots, Y_m)$ be a $(k + m)$-dimensional continuous random variable with density $f_{X_1,\ldots,X_k,Y_1,\ldots,Y_m}(x_1, \ldots, x_k, y_1, \ldots, y_m)$; then

$$\begin{aligned}\mathscr{E}[g(X_1, \ldots, X_k, Y_1, \ldots, Y_m)|x_1, \ldots, x_k] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_k, y_1, \ldots, y_m) \\ &\quad \times f_{Y_1,\ldots,Y_m|X_1,\ldots,X_k}(y_1, \ldots, y_m|x_1, \ldots, x_k)\, dy_1 \ldots dy_m. \qquad ////\end{aligned}$$

EXAMPLE 23 In the experiment of tossing two tetrahedra with X, the number on the first, and Y, the larger of the two numbers, we found that

$$f_{Y|X}(y|2) = \begin{cases} \frac{1}{2} & \text{for } y = 2 \\ \frac{1}{4} & \text{for } y = 3 \\ \frac{1}{4} & \text{for } y = 4 \end{cases}$$

in Example 9. Hence $\mathscr{E}[Y|X = 2] = \sum y f_{Y|X}(y|X = 2) = 2\cdot\frac{1}{2} + 3\cdot\frac{1}{4} + 4\cdot\frac{1}{4}$

EXAMPLE 24 For $f_{X,Y}(x, y) = (x + y)I_{(0,1)}(x)I_{(0,1)}(y)$, we found that

$$f_{Y|X}(y|x) = \frac{x+y}{x+\frac{1}{2}} I_{(0,1)}(y)$$

for $0 < x < 1$ in Example 12. Hence

$$\mathscr{E}[Y|X = x] = \int_0^1 y \frac{x+y}{x+\frac{1}{2}}\, dy = \frac{1}{x+\frac{1}{2}}\left(\frac{x}{2} + \frac{1}{3}\right)$$

for $0 < x < 1$. ////

As we stated above, $\mathscr{E}[g(Y)|x]$ is, in general, a function of x. Let us denote it by $h(x)$; that is, $h(x) = \mathscr{E}[g(Y)|x]$. Now we can evaluate the expectation of $h(X)$, a function of X, and will have $\mathscr{E}[h(X)] = \mathscr{E}[\mathscr{E}[g(Y)|X]]$. This gives us

This gives us

$$\begin{aligned}
\mathscr{E}[\mathscr{E}[g(Y)\,|\,X]] = \mathscr{E}[h(X)] &= \int_{-\infty}^{\infty} h(x) f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} \mathscr{E}[g(Y)\,|\,x] f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(y) f_{Y|X}(y\,|\,x)\,dy\right] f_X(x)\,dx \\
&= \int_{-\infty}^{-\infty} \int_{-\infty}^{\infty} g(y) f_{Y|X}(y\,|\,x) f_X(x)\,dy\,dx
\end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y) f_{X,Y}(x, y)\,dy\,dx \\
&= \mathscr{E}[g(Y)].
\end{aligned}$$

Thus we have proved for jointly continuous random variables X and Y (the proof for X and Y jointly discrete is similar) the following simple yet very useful theorem.

Theorem 6 Let (X, Y) be a two-dimensional random variable; then

$$\mathscr{E}[g(Y)] = \mathscr{E}[\mathscr{E}[g(Y)\,|\,X]], \qquad (25)$$

and in particular

$$\mathscr{E}[Y] = \mathscr{E}[\mathscr{E}[Y\,|\,X]]. \qquad (26)$$

////

Definition 22 **Regression curve** $\mathscr{E}[Y\,|\,X = x]$ is called the *regression curve* of Y on x. It is also denoted by $\mu_{Y|X=x} = \mu_{Y|x}$. ////