Let g be a function for which the expectation of $g(X)$ is finite. We define the *conditional* expected value of $g(X)$ given $Y = y$ by the formula

$$E[g(X)|Y = y] = \sum_x g(x)p_{X|Y}(x|y) \quad \text{if } p_Y(y) > 0, \tag{2.4}$$

and the conditional mean is not defined at values y for which $p_Y(y) = 0$. The law of total probability for conditional expectation reads

$$E[g(X)] = \sum_y E[g(X)|Y = y]p_Y(y). \tag{2.5}$$

The conditional expected value $E[g(X)|Y = y]$ is a function of the real variable y. If we evaluate this function at the random variable Y, we obtain a random variable that we denote by $E[g(X)|Y]$. The law of total probability in (2.5) now may be written in the form

$$E[g(X)] = E\{E[g(X)|Y]\}. \tag{2.6}$$

Since the conditional expectation of $g(X)$ given $Y = y$ is the expectation with respect to the conditional probability mass function $p_{X|Y}(x|y)$, conditional expectations behave in many ways like ordinary expectations. The following list summarizes some properties of conditional expectations. In this list, with or without affixes, X and Y are jointly distributed random variables; c is a real number; g is a function for which $E[|g(X)|] < \infty$; h is a bounded function; and v is a function of two variables for which $E[|v(X, Y)|] < \infty$. The properties are

1. $E[c_1 g_1(X_1) + c_2 g_2(X_2) | Y = y]$

$$= c_1 E[g_1(X_1)|Y = y] + c_2 E[g_2(X_2)|Y = y]. \tag{2.7}$$

2. if $g \geq 0$, then $E[g(X)|Y = y] \geq 0$. (2.8)

3. $E[v(X, Y)|Y = y] = E[v(X, y)|Y = y]$. (2.9)

4. $E[g(X)|Y = y] = E[g(X)]$ if X and Y are independent. (2.10)

5. $E[g(X)h(Y)|Y = y] = h(y)E[g(X)|Y = y]$. (2.11)

6. $E[g(X)h(Y)] = \sum_y h(y)E[g(X)|Y = y]p_Y(y)$

$$= E\{h(Y)E[g(X)|Y]\}. \tag{2.12}$$

As a consequence of (2.7), (2.11), and (2.12), with either $g \equiv 1$ or $h \equiv 1$, we obtain

$$E[c|Y = y] = c, \tag{2.13}$$

$$E[h(Y)|Y = y] = h(y), \tag{2.14}$$

$$E[g(X)] = \sum_y E[g(X)|Y = y]p_Y(y) = E\{E[g(X)|Y]\}. \tag{2.15}$$