$$variance = \sum f(x - mean)^2 = E(X - mean)^2$$

$$E(X) = mean\ of\ X$$

$$E(X - mean)^2 = varaince\ of\ X = E(X^2) - mean^2$$

$$E(X^2) = \sum x^2 f(x)$$

Example 4.11: Calculate the variance of $g(X) = 2X + 3$, where X is a random variable with probability distribution

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

$$E\big(g(x)\big)^2 - \Big(E\big(g(x)\big)\Big)^2$$

$$E\big(g(x)\big) = \sum g(x)f(x) = \sum(2x + 3)f(x) = 3f(0) + 5f(1) + 7f(2) + 9f(3)$$
$$= 3 * \frac{1}{4} + 5 * \frac{1}{8} + 7 * \frac{1}{2} + 9 * \frac{1}{8} = 6$$

$$E(g(x)^2) = \sum g(x)^2 f(x) = \sum(2x + 3)^2 f(x) = 3^2 f(0) + 5^2 f(1) + 7^2 f(2) + 9^2 f(3)$$
$$= 9 * \frac{1}{4} + 25 * \frac{1}{8} + 49 * \frac{1}{2} + 81 * \frac{1}{8} = a$$

Covariance:

$$\sigma_x \sigma_x = \sigma_x^2 = E(XX) - \mu_x \mu_x = \sum(x - mean)^2 f = \sum(x - mean)(x - mean)f$$

$$\sigma_{xy} = \sum\sum(x - \mu_x)(y - \mu_y)f = E\Big((X - \mu_x)(Y - \mu_y)\Big) = E(XY) - \mu_x \mu_y$$

$$Type\ equation\ here.$$

Example 4.13: Example 3.14 on page 95 describes a situation involving the number of blue refills X and the number of red refills Y. Two refills for a ballpoint pen are selected at random from a certain box, and the following is the joint probability distribution:

	$f(x, y)$	$x = 0$	$x = 1$	$x = 2$	$h(y)$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Find the covariance of X and Y.

$$\sigma_{xy} = E(XY) - \mu_x \mu_y = E(XY) - E(X)E(Y)$$

$$E(X) = \sum\sum x f(x, y) = \sum x g(x) = 0g(0) + 1g(1) + 2g(2) = \frac{15}{28} + 2 * \frac{3}{28} = \frac{21}{28} = \frac{3}{4} = \mu_x$$

$$E(Y) = \sum\sum y f(x,y) = \sum y h(y) = 0h(0) + 1h(1) + 2h(2) = \frac{3}{7} + 2 * \frac{1}{28} = \frac{1}{2} = \mu_y$$

$$E(XY) = \sum\sum xy f(x,y) = \frac{3}{14}$$

$$\sigma_{xy} = E(XY) - \mu_x \mu_y = E(XY) - E(X)E(Y) = \frac{3}{14} - \frac{3}{4} * \frac{1}{2} = -\frac{9}{56}$$

Another example:

A bank operates both a drive-up facility and a walk-up window. On a randomly selected day, let X = the proportion of time that the drive-up facility is in use (at least one customer is being served or waiting to be served) and Y = the proportion of time that the walk-up window is in use. Then the set of possible values for (X, Y) is the rectangle $D = \{(x, y): 0 \le x \le 1, 0 \le y \le 1\}$. Suppose the joint pdf of (X, Y) is given by

$$f(x,y) = \begin{cases} \frac{6}{5}(x + y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find covariance and correlation between variable x and y .

Solution:

$$\sigma_{xy} = E(XY) - E(X)E(Y)$$

$$E(X) = \iint \quad x f(x,y) dx dy = \int \qquad x g(x) dx$$

$$E(Y) = \int_0^1 \quad y h(y) dy$$

$$g(x) = \int_0^1 \quad f(x,y) dy = \int_0^1 \quad \frac{6}{5}(x + y^2) dy = \frac{6}{5}\left(xy + \frac{y^3}{3}\right) (limit\ y\text{: } 0\ to\ 1)$$

$$g(x) = \frac{6}{5}(x + \frac{1}{3})$$

$$h(y) = \int_0^1 \quad f(x,y) dx = \int_0^1 \quad \frac{6}{5}(x + y^2) dx = \frac{6}{5}(\frac{x^2}{2} + xy^2)(limit\ x\text{: } 0\ to\ 1)$$

$$h(y) = \frac{6}{5}(\frac{1}{2} + y^2)$$

$$E(X) = \int_0^1 \quad x g(x) dx = \int_0^1 \quad x * \frac{6}{5}\left(x + \frac{1}{3}\right) dx = \frac{6}{5}\int_0^1 \quad \left(x^2 + \frac{1}{3}x\right) dx = \frac{3}{5}$$

$$E(Y) = \int_0^1 \quad yh(y)dy = \int_0^1 \quad y * \frac{6}{5}\left(\frac{1}{2} + y^2\right)dy = \frac{6}{5}\int_0^1 \quad \left(\frac{y}{2} + y^3\right)dy = \frac{3}{5}$$

$$E(XY) = \int_{x=0}^1 \quad \int_{y=0}^1 \quad xyf(x,y)dydx = \int_{x=0}^1 \quad \int_{y=0}^1 \quad xy * \frac{6}{5}(x + y^2)dydx$$

$$E(XY) = \frac{6}{5}\int_{x=0}^1 \quad xdx \int_{y=0}^1 \quad (xy + y^3)dy = \frac{6}{5}\int_{x=0}^1 \quad xdx\left(\frac{xy^2}{2} + \frac{y^4}{4}\right)(limit\ y{:}\, 0\ to\ 1)$$

$$E(XY) = \frac{6}{5}\int_{x=0}^1 \quad xdx\left(\frac{x}{2} + \frac{1}{4}\right) = \frac{6}{5}\int_0^1 \quad \left(\frac{x^2}{2} + \frac{x}{4}\right)dx = \frac{7}{20}$$

$$\sigma_{xy} = \frac{7}{20} - \frac{3}{5} * \frac{3}{5} = -\frac{1}{100}$$

Correlation:

$$-1 \leq \rho_{xy} \leq 1$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = 0\ (zero\ correlation{:}\ x\ and\ y\ are\ independent)$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = -1\ (perfect\ correlation{:}\ inversly\ proportional)$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = 1(perfect\ correlation{:}\ directly\ proportional)$$

$$\sigma_x^2 = E(X^2) - E(X)^2, \sigma_y^2 = E(Y^2) - E(Y)^2$$

$$E(X^2) = \int_0^1 \quad x^2 g(x)dx = \int_0^1 \quad x^2 * \frac{6}{5}\left(x + \frac{1}{3}\right)dx = \frac{13}{30}$$

$$\sigma_x^2 = E(X^2) - E(X)^2 = \frac{13}{30} - \frac{9}{25}(verify\ answers\ yourself)$$

$$E(Y^2) = \int_0^1 \quad y^2 h(y)dy = \int_0^1 \quad y^2 \frac{6}{5}(\frac{1}{2} + y^2)dy = \frac{11}{25}$$

$$\sigma_y^2 = E(Y^2) - E(Y)^2 = \frac{11}{25} - \frac{9}{25}\ (Verify\ answers\ yourself)$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Summary:

$$E\big(g(x)\big) = \sum g(x)f(x)\ or\ \int g(x)f(x)dx = mean\ of\ g(x)$$

$$\sigma_x^2 = E(X^2) - E(X)^2$$

$$\sigma_{xy} = E(XY) - E(X)E(Y)$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$