Example 4.2.9 *Let X be a discrete random variable with distribution $f(x) = \frac{1}{8}\binom{3}{x}$ for $x = 0, 1, 2, 3$.*

The moment generating function for X is $M_X(t) = \sum_{x=0}^{3} e^{tx} \cdot \left(\frac{1}{8}\binom{3}{x}\right)$

$$= \frac{1}{8}\left(e^0\binom{3}{0} + e^t\binom{3}{1} + e^{2t}\binom{3}{2} + e^{3t}\binom{3}{3}\right) = \frac{1}{8}(1+e^t)^3$$

To find the mean (1st moment about the origin):

$$\frac{d}{dt}M_X(t)\bigg|_{t=0} = \frac{d}{dt}\frac{1}{8}(1+e^t)^3\bigg|_{t=0}$$

$$= \frac{3}{8}(1+e^t)^2e^t\bigg|_{t=0} = \frac{3}{8}2^2 = \frac{3}{2}$$

The second moment about the origin, $E(X^2)$: $\frac{d^2}{dt^2} M_X(t)\Big|_{t=0}$

$$
\begin{aligned}
&= \frac{d^2}{dt^2}\frac{1}{8}(1+e^t)^3\Big|_{t=0} \\
&= \frac{d}{dt}\frac{3}{8}(1+e^t)^2 e^t\Big|_{t=0} \\
&= \frac{6}{8}(1+e^t)e^{2t} + \frac{3}{8}(1+e^t)^2 e^t\Big|_{t=0} \\
&= 3.
\end{aligned}
$$

These two could now be used to find the variance, $\sigma^2 = E(X^2) - \mu^2$.

Properties of Moment Generating Functions:

One advantage of knowing the moment generating function for a random variable is that it can be used to find the moment generating function for related random variable via the following theorem.

Theorem 4.2.10 *If $M_X(t)$ is the moment generating function for X and a and b are nonzero constants, then*

1.

$$M_{X+a}(t) = e^{at} \cdot M_X(t)$$

2.

$$M_{bX}(t) = e^{at} \cdot M_X(bt)$$

3.

$$M_{\frac{X+a}{b}}(t) = e^{\frac{a}{b}t} \cdot M_X(\frac{t}{b})$$

Product Moments about the Origin:

We have already discussed the expected value of a bivariate function $g(X, Y)$, where $E(g(X, Y)) = \sum_x \sum_y g(x, y) \cdot f(x, y)$.

The following is a special case of this:

The r**th and** s**th product moment about the origin** of X and Y, denoted by $\mu'_{r,s}$, is the expected value of $X^r Y^s$:

$$\text{discrete:} \qquad \mu'_{r,s} = E(X^r Y^s) = \sum_x \sum_y x^r y^s \cdot f(x, y)$$

$$\text{continuous:} \qquad \mu'_{r,s} = E(X^r Y^s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^r y^s \cdot f(x, y)\, dx\, dy$$

for $r = 0, 1, 2, \dots$, $s = 0, 1, 2, \dots$.

Note that $\mu'_{1,0} = E(X)$ which we will denote by μ_X and $\mu'_{0,1} = E(Y)$ which we will denote by μ_Y.

Product Moments about the Means:

Similarly, the r**th and** s**th product moment about the mean** of X and Y, denoted by $\mu_{r,s}$, is the expected value of $(X-\mu_X)^r(Y-\mu_Y)^s$:

discrete:

$$\begin{aligned}\mu_{r,s} &= E((X-\mu_X)^r(Y-\mu_Y)^s)\\ &= \sum_x \sum_y (x-\mu_X)^r(y-\mu_Y)^s \cdot f(x,y)\end{aligned}$$

continuous:

$$\begin{aligned}\mu_{r,s} &= E((X-\mu_X)^r(Y-\mu_Y)^s)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)^r(y-\mu_Y)^s \cdot f(x,y)\ dx\ dy\end{aligned}$$

4.2.2 Covariance

The 1st and 1st product moment about the means of X and Y, i.e. $\mu_{1,1}$, is called the **covariance** of X and Y. It is commonly denoted σ_{XY}, $\text{cov}(X, Y)$ or $C(X, Y)$.

Summary:

$$\mu_X = E(X) = \sum_x \sum_y x \cdot f(x, y)$$

$$\mu_Y = E(Y) = \sum_x \sum_y y \cdot f(x, y)$$

$$\text{cov}(X, Y) = E((X - \mu_X)(Y - \mu_Y)) = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) \cdot f(x, y)$$

(discrete case shown here, continuous case is analogous)

The covariance of describes the relationship between X and Y.

If there is a high probability that large X values and large Y values appear together, then the covariance is positive (or small X with small Y). On the other hand large/small X values occurring with small/large Y values is more likely, then the covariance will be negative.

The following is a a very useful formula for the covariance.

Theorem 4.2.11

$$\sigma_{XY} = \mu'_{1,1} - \mu_X \mu_Y$$

Example 4.2.12 *In the caplet example, find the covariance of X and Y.*

$$\begin{array}{cc|ccc|}
 & & & x & \\
 & & 0 & 1 & 2 \\
\hline
 & 0 & \frac{6}{36} & \frac{12}{36} & \frac{3}{36} \\
\text{y} & 1 & \frac{8}{36} & \frac{6}{36} & \\
 & 2 & \frac{1}{36} & & \\
\hline
\end{array}$$

We start by finding μ_X and μ_Y:

$$\mu_X = E(X) = \sum_x \sum_y x \cdot f(x,y) = \sum_x x \sum_y \cdot f(x,y) = \sum_x x \cdot g(x)$$

$$\mu_Y = E(Y) = \sum_x \sum_y y \cdot f(x,y) = \sum_y y \sum_x \cdot f(x,y) = \sum_y y \cdot h(y)$$

where $g(x)$ and $h(y)$ are the marginal distributions of x and y respectively.

			x		
		0	1	2	
	0	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{3}{36}$	$\frac{21}{36}$
y	1	$\frac{8}{36}$	$\frac{6}{36}$		$\frac{14}{36}$
	2	$\frac{1}{36}$			$\frac{1}{36}$
		$\frac{15}{36}$	$\frac{18}{36}$	$\frac{3}{36}$	

Thus, $\mu_X = 0 \cdot \frac{15}{36} + \dots$

$\mu_Y =$

To use the formula $\sigma_{XY} = \mu'_{1,1} - \mu_X \mu_Y$, we need $\mu'_{1,1} = E(XY)$:

$$
\begin{aligned}
&= \sum_{x=0}^{2} \sum_{y=0}^{2} (xy) \cdot f(x,y) \\
&= (0 \cdot 0) \cdot \frac{6}{36} + (0 \cdot 1) \cdot \frac{8}{36} + \dots
\end{aligned}
$$

So we have $\sigma_{XY} = \frac{6}{36} - (\frac{24}{36})(\frac{16}{36}) = -\frac{7}{54} \approx -0.1296$.

Covariance and Independence:

Recall that joint random variables X and Y are independent if and only if $f(x, y) = g(x) \cdot h(y)$; i.e. their joint distribution is the product of the marginal distributions.

The following theorem is a consequence of this.

Theorem 4.2.13 *If X and Y are independent, then*

$$E(XY) = E(X) \cdot E(Y)$$

and $\sigma_{XY} = 0$.

This theorem is only a one-way implication as seen in the next example.