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

"3 choose x"

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

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

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

expand the cubic expression to verify the identity

To find the mean $E(X)$ *(1st moment about the origin):*

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

CHAIN RULE! $= 3\cdot\frac{1}{8}(1+e^t)^2\cdot e^t$

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

$\frac{3}{8}\cdot 4 = \frac{12}{8} = \frac{3}{2} = 1.5 = \mu$

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

To find the 2^{nd} moment about the origin, we need the 2^{nd} derivative of the moment generating function.

$$\frac{d}{dt} M_X(t) = \frac{3}{8}(1+e^t)^2 \cdot e^t \quad (1^{st} \text{ derivative})$$

$$\frac{d}{dt}\left(\frac{3}{8}\underbrace{(1+e^t)^2}_{f} \cdot \underbrace{e^t}_{g}\right) \quad (2^{nd} \text{ derivative})$$

$$= \frac{3}{8}\left[\underbrace{(1+e^t)^2}_{f} \cdot \underbrace{e^t}_{g'} + \left(\underbrace{2(1+e^t)^1 \cdot e^t}_{f'}\right)\underbrace{e^t}_{g}\right]$$

$$= \frac{3}{8}(1+e^t)^2 \cdot e^t + \frac{6}{8}(1+e^t) \cdot e^{2t}$$

We need to evaluate this expression at $t=0$.

$$\left[\frac{3}{8}(1+e^t)^2 \cdot e^t + \frac{6}{8}(1+e^t) \cdot e^{2t}\right]\Bigg|_{t=0}$$

$$= \frac{3}{8}(1+e^0)^2 e^0 + \frac{6}{8}(1+e^0) \cdot e^{2 \cdot 0}$$

$$= \frac{3}{8} \cdot 2^2 \cdot 1 + \frac{6}{8} \cdot 2 \cdot 1 = \frac{12}{8} + \frac{12}{8} = \underbrace{3}_{E(X^2)}$$

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

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

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

$$= \frac{6}{8}(1+e^t)e^{2t} + \frac{3}{8}(1+e^t)^2 e^t\Big|_{t=0}$$

$$= 3.$$

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

$3 - (1.5)^2$

$= 3 - 2.25 = 0.75$

$= \frac{3}{4}$

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.

$M_X(t)$: moment generating function of X

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

p.146 Example

1.

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

2.

$$M_{bX}(t) = M_X(bt)$$

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

discrete

joint probability distribution

The following is a special case of this:

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

μ'_r : r^{th} moment about the origin

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

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

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

discrete: $\mu_x = \mu'_{1,0} = E(X^1Y^0) = \sum_x \sum_y x^1 y^0 \cdot f(x,y) = \sum_x \sum_y x \cdot f(x,y) = E(X)$

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 $(X-\mu_X)^r(Y-\mu_Y)^s$:

mean of X   mean of Y

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:

discrete case

mean of X

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

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

mean of Y

continuous case:

$\mu_x = E(X) = \int\int x \cdot f(x,y)\, dx\, dy$

$\mu_y = E(Y) = \int\int y \cdot f(x,y)\, dx\, dy$

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

(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

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

$\mu_{1,1} = \mathrm{cov}(X,Y)$

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

		x = 0	1	2
y	0	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{3}{36}$
	1	$\frac{8}{36}$	$\frac{6}{36}$	
	2	$\frac{1}{36}$		

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

$= E(X^1 Y^1) - \mu_X \cdot \mu_Y$

We start by finding μ_X and μ_Y:

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

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

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

		x = 0	1	2		
y	0	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{3}{36}$	$\frac{21}{36}$	$= h(0)$
	1	$\frac{8}{36}$	$\frac{6}{36}$		$\frac{14}{36}$	$= h(1)$
	2	$\frac{1}{36}$			$\frac{1}{36}$	$= h(2)$
		$\frac{15}{36}$	$\frac{18}{36}$	$\frac{3}{36}$		
		$= g(0)$	$= g(1)$	$= g(2)$		

marginal distribution for Y

marginal distribution for X

$$\mu_x = \sum_x x \cdot g(x) \qquad \mu_y = \sum_y y \cdot h(y)$$

Thus, $\mu_X = \cancel{0 \cdot \frac{15}{36}} + 1 \cdot \frac{18}{36} + 2 \cdot \frac{3}{36} = \frac{24}{36} = \frac{2}{3}$

$\mu_Y = \cancel{0 \cdot \frac{21}{36}} + 1 \cdot \frac{14}{36} + 2 \cdot \frac{1}{36} = \frac{16}{36} = \frac{4}{9}$

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

$$\begin{aligned}
E(X^1Y^1) &= \sum_{x=0}^{2} \sum_{y=0}^{2} (x^1y^1) \cdot f(x,y) \\
&= \cancel{(0 \cdot 0) \cdot \frac{6}{36}} + \cancel{(0 \cdot 1) \cdot \frac{8}{36}} + \cancel{(0 \cdot 2) \cdot \frac{1}{36}} \\
&+ \cancel{(1 \cdot 0) \cdot \frac{12}{36}} + (1 \cdot 1) \cdot \frac{6}{36} + \cancel{(2 \cdot 0) \cdot \frac{3}{36}} = \frac{6}{36} = \frac{1}{6}
\end{aligned}$$

So we have $\sigma_{XY} = \overset{E(XY)}{\frac{6}{36}} - \underbrace{(\tfrac{24}{36})}_{\mu_X}\underbrace{(\tfrac{16}{36})}_{\mu_Y} = -\frac{7}{54} \approx -0.1296.$

σ_{XY}: cov(X, Y)