Lecture 13

Covariance, Correlation, and MGFs

Grace Tompkins

Last modified — 07 Jun 2026

Learning Outcomes

By the end of this lecture, students are anticipated to be able to:

- Define covariance and correlation
- Calculate covariances and variances from discrete and continuous distributions
- Define and calculate a moment generating function
- Use the moment generating function to calculate moments

1 Covariance and Correlation

Covariance

If we have two random variables X and Y, we can measure the relationship between them.

📖 **DEFINITION**

The **covariance** between two random variables X and Y is defined by

$$\mathrm{Cov}(X,Y) = \mathbb{E}\Big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\Big].$$

$$= E[XY] - E[X]E[Y]$$

- $\mathrm{Cov}(X,Y)$ is a scalar-valued function of X and Y.
- The covariance measures the linear relationship between X and Y.
- If $\mathrm{Cov}(X,Y) > 0$, then X and Y tend to increase together.

Covariance

We can compute the covariance using the joint PMF/PDF of X and Y.:

$$\operatorname{Cov}(X, Y) = \sum_x \sum_y (x - \mathbb{E}[X])(y - \mathbb{E}[Y]) p_{X,Y}(x, y).$$

(discrete)

$$\operatorname{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathbb{E}[X])(y - \mathbb{E}[Y]) f_{X,Y}(x, y)\, \mathrm{d}x\, \mathrm{d}y.$$

(continuous)

Properties of covariance

$$\mathrm{Cov}(X+d,\ Y+e) = \mathrm{Cov}(X,Y)$$

Linearity

For any $a, b, c \in \mathbb{R}$, and random variables X, Y, and Z,

$$\mathrm{Cov}(aX + bY, cZ) = \overset{a\cdot c}{\cancel{a}}\,\mathrm{Cov}(X, Z) + \overset{b\cdot c}{\cancel{b}}\,\mathrm{Cov}(Y, Z).$$

Easier calculation

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Independence

If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$. The converse is false.

X ind $Y \Rightarrow \mathrm{Cov}(X,Y) = 0$

$\mathrm{Cov}(X,Y) = 0 \not\Rightarrow X$ ind Y

Covariance

✏️ **EXERCISE: HEADS OR TAILS?**

Let X be the number of heads in 5 tosses of a fair coin, and let Y be the number of tails in the same 5 tosses. Find $\text{Cov}(X, Y)$.

Hint: If $Z \sim \text{Binom}(n, \theta)$, then $\text{Var}(Z) = n\theta(1-\theta)$. + use linearity.

X = # heads

Y = # tails = $5 - X$

$X \sim \text{Bin}(n=5, \theta=0.5)$

$n = 5$

$\theta = 0.5$

$$\begin{aligned}
\text{Cov}(X,Y) &= \text{Cov}(X, 5-X) \\
&= \text{Cov}(X, -X) \\
&= -\text{Cov}(X, X) \\
&= -1[E(XX) - E(X)E(X)] \\
&= -1[E(X^2) - E(X)^2] \\
&= -1\,\text{Var}(X) \\
&= -1(5)(0.5)(1-0.5) = \frac{-5}{4}
\end{aligned}$$

Covariance

Variance, Covariance, and Sums

Let X and Y be random variables with finite variances.

try to show this!

- $$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \underline{2\,\mathrm{Cov}(X,Y)}.$$
- If X and Y are independent, then
$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$
- More generally, if $X_1, \ldots, X_n$ are independent random variables with finite variances, then
$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Correlation

Covariance is not a standardized measure of the relationship between X and Y.

For example, if we multiply X by 100, then $\mathrm{Cov}(X, Y)$ will also be multiplied by 100.

📖 DEFINITION

The **correlation** between two random variables X and Y is defined by

$$\rho_{XY} = \mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

standard deviations

- The correlation is a standardized measure of the linear relationship between X and Y.
- We'll see later that $-1 \leq \rho_{XY} \leq 1$.

perfect neg. corr.

perfect positive correlation

previous coin flip ex:

$$\rho_{XY} = \frac{-5/4}{\sqrt{5/4}\sqrt{5/4}} = -1$$

Correlation

✏️ **EXERCISE: TABULAR EXAMPLE**

Let X and Y be discrete random variables with the following joint distribution:

P(X, Y)	Y = 2	Y = 4
X = 1	0.2	0.3
X = 3	0.1	0.4

Hint: $\sigma_X = 1$ and $\sigma_Y = 0.92$.

Calculate the correlation between X and Y.

$Cov(X,Y) = E(XY) - E(X)E(Y)$

$E(XY) = (1)(2)(0.2) + (3)(2)(0.1) + (1)(4)(0.3) + (3)(4)(0.4) = 7$

$E(X) = (1)(0.2+0.3) + (3)(0.5) = 2$

$E(Y) = (2)(0.2+0.1) + (4)(0.3+0.4) = 3.4$

Correlation

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y} = \frac{7 - 2(3.4)}{(1)(0.92)}$$

$$= \frac{0.2}{0.92} = 0.22$$

2 Moment generating functions

The Moment Generating Function

Important

We will not cover/discuss probability generating functions or characteristic functions in this course. We will not discuss:

- (Beginning of 3.1) $r_X(t) = \mathbb{E}[t^X]$, the probability generating function of a random variable X.
- (Section 3.1.4) $c_X(t) = \mathbb{E}[e^{itX}]$, the characteristic function

$m_X(t)$

The Moment Generating Function

📖 **DEFINITION**

The **moment generating function** (MGF) of a random variable X is defined by

$$m_X(t) = \mathbb{E}[e^{tX}].$$

- The MGF is a scalar-valued function of t.

The Moment Generating Function

💡 **EXAMPLE**

Last lecture, we saw that $\mathbb{E}[\exp(tX)] = \boxed{\left(1 - \frac{t}{\lambda}\right)^{-\alpha}}$ when $X \sim \mathrm{Gam}(\alpha, \lambda)$. (*see "Exercise: More Gamma Expectations"*). This was actually the moment generating function!

$$m_X(t) = \mathbb{E}[e^{tX}] = \left(1 - \frac{t}{\lambda}\right)^{-\alpha}, \quad t < \lambda.$$

- $m_X(t)$ depends on the parameters α and λ of the distribution of X.
- It also depends on t, which is a free variable that we can choose.
- This seems like a weird object to care about, but it turns out to be very useful.

Using the MGF to Compute Moments

THEOREM

If X is a random variable with MGF $m_X(t)$, and there exists $s > 0$ such that, for all $t \in (-s, s)$, $m_X(t) < \infty$.

Then for any integer $k \geq 1$,

$$\mathbb{E}[X^k] = m_X^{(k)}(0) = \frac{\mathrm{d}^k}{\mathrm{d}t^k} m_X(t)\Big|_{t=0}.$$

- $\mathbb{E}[X^k]$ is called the k-th moment of X.
- The MGF is called the "moment generating function" because we can use it to compute the moments of X.
- Specifically, the first moment is $\mathbb{E}[X] = m'_X(0)$, and the second moment is $\mathbb{E}[X^2] = m''_X(0)$.

E(X): Find the MGF, take first derivative wrt t, set t=0.

Using the MGF to Compute Moments

we solved for this in a prev lecture!

💡 EXAMPLE

Let $X \sim \text{Gam}(\alpha, \lambda)$ with MGF $m_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha}$. Find $\text{Var}(X)$.

$$E(X) = m_X'(0) = \frac{d}{dt}\left(1 - \frac{t}{\lambda}\right)^{-\alpha}\Big|_{t=0} = \frac{\alpha}{\lambda}\left(1 - \frac{t}{\lambda}\right)^{-\alpha-1}\Big|_{t=0} = \frac{\alpha}{\lambda}$$

$$E(X^2) = m_X''(0) = \frac{d}{dt}\left(\frac{\alpha}{\lambda}\left(1 - \frac{t}{\lambda}\right)^{-\alpha-1}\right)\Big|_{t=0} = \frac{\alpha(\alpha+1)}{\lambda^2}\left(1 - \frac{t}{\lambda}\right)^{-\alpha-2}\Big|_{t=0} = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}$$

Using the MGF to Compute Moments

Using the MGF to Compute Moments

✏️ **EXERCISE: MEAN AND VARIANCE OF NORMAL**

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then,

$$m_X(t) = \mathbb{E}[e^{tX}] = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right).$$

Find $\mathbb{E}[X]$ and $\mathrm{Var}(X)$ using the MGF of X.

$$E(X) = m_X'(t=0) = (\mu + \sigma^2 t) e^{(\mu t + \sigma^2/2\, t^2)} \Big|_{t=0} = \mu$$

$$E(X^2) = (\sigma^2 + (\mu + \sigma^2 t^2)^2) e^{\mu t + \frac{1}{2}\sigma^2 t^2} \Big|_{t=0} = \sigma^2 + \mu^2$$

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = (\sigma^2 + \mu^2) - \mu^2 = \sigma^2$$

Using the MGF to Compute Moments

Sums of Independent Random Variables

📐 **THEOREM**

Let X and Y be independent random variables with MGFs $m_X(t)$ and $m_Y(t)$, respectively.

Then the MGF of $X + Y$ is given by

$$\begin{aligned} m_{X+Y}(t) &= \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}e^{tY}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] \qquad X \text{ and } Y \text{ are independent} \\ &= m_X(t)m_Y(t). \end{aligned}$$

* only true if X, Y are independent *

Sums of Independent Random Variables

- We saw before how to find the PMF/PDF of $X + Y$ using convolution. The MGF gives us an alternative way to find the distribution of $X + Y$.
- If $X_1, \ldots, X_n$ are independent random variables with common MGF $m_X(t)$, then the MGF of $S_n = \sum_{i=1}^{n} X_i$ is given by

$$m_{S_n}(t) = (m_X(t))^n.$$

- If X has MGF $m_X(t)$, then $aX + b$ has MGF $e^{bt} m_X(at)$ for any $a, b \in \mathbb{R}$.

can show this by looking at

$m_Y(t)$ where $Y = aX + b$

Uniqueness of the MGF

THEOREM

If X and Y are random variables with MGFs $m_X(t)$ and $m_Y(t)$, respectively, and there exists $s > 0$ such that for all $t \in (-s, s)$, $m_X(t) = m_Y(t) < \infty$, then X and Y have the same distribution.

This is a very important result, as it allows us to identify the distribution of a random variable by finding its MGF.

Using These Theorems Together

✏️ **EXERCISE: SUMS OF NORMALS**

Let $X_1, \ldots, X_n$ be independent identically distributed (i.i.d.) random variables with $X_i \sim \mathcal{N}(\mu, \sigma^2)$ for all i.

Note that $m_{X_i}(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$ for $i = 1, \ldots, n$.

Find the distribution of $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ using MGFs.

hint: try to get the MGF of $\overline{X}$ into a familiar form.

Using These Theorems Together

From slide 23:

MGF for $\frac{1}{n}X_i = m_X(\frac{1}{n}t)$

MGF for $\sum_{i=1}^{n} \frac{1}{n}X_i = \prod_{i=1}^{n} m_X(\frac{1}{n}t)$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \frac{1}{n}X_i$$

$$m_{\bar{X}}(t) = \prod_{i=1}^{n} m_{X_i}\left(\frac{1}{n}t\right) = \prod_{i=1}^{n} \exp\left(\mu\frac{t}{n} + \frac{1}{2}\sigma^2\left(\frac{t}{n}\right)^2\right)$$

$$= \exp\left(\mu t + \frac{1}{2}\frac{\sigma^2 t^2}{n}\right) = \exp\left(\mu t + \frac{1}{2}\left(\frac{\sigma^2}{n}\right)t^2\right)$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Functions of Non-Independent Random Variables

The previous theorems assume independence. Consider the scenario where X and Y are **not** independent.

📌 **COROLLARY**

Let X and Y be random variables that are not necessarily independent.

Then the MGF of $h(X, Y)$ is given by

$$m_{h(X,Y)}(t) = \mathbb{E}[e^{t \times h(X,Y)}]$$

For example, if we were interested in the MGF of $X + Y$, we would solve $m_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}]$ directly.

Functions of Non-Independent Random Variables

✏️ **EXERCISE: NONINDEPENDENT SUMS**

Let X and Y have joint pmf:

(X,Y)	0	1
0	0.1	0.4
1	0.4	0.1

a. Find the MGF of $Z = X + Y$.

b. Find $\mathbb{E}[Z]$ using the MGF.

$$m_{x+y}(t) = E\left(e^{t(x+y)}\right) = \sum_x \sum_y e^{t(x+y)} p(x,y)$$

Functions of Non-Independent Random Variables

$$m_{X+Y}(t) = E\left(e^{t(X+Y)}\right) = \sum_x \sum_y e^{t(x+y)} p(x,y)$$

$$= e^{t(0+0)}(0.1) + e^{t(0+1)}(0.4) + e^{t(1+0)}(0.4) + e^{t(1+1)}(0.1)$$

$$= 0.1 + 0.8e^{t} + 0.1e^{2t}$$

$$m_{X+Y}'(t) = 0.8e^{t} + 0.2e^{2t}$$

$$E(X+Y) = m_{X+Y}'(t=0) = 0.8e^{0} + 0.2e^{2(0)} = 1$$

To Do

- Work on Assignment 3, due Wednesday June 10, 11:59pm on Gradescope.
- Read Chapter 3.5 and 3.6 before next class.