Lecture 6

Topics:

- Covariance
- Master formulas 1 & 2
- z-scores

If two random variables X, Y are not independent (i.e. they are dependent) there are many ways to summarize how dependent they are.

The most common is with covariance and correlation.

Definition: If X and Y are two r.v.s then the covariance between X & Y is defined as

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

This number measures how much X and Y "co-vary together"

Remark 1: we will see that cov is very closely related to var. Indeed notice the similarity of the definitions

$$cov(X,Y) = E(X \cdot Y) - E(X)E(Y)$$

$$var(X) = E(X \cdot X) - E(X)E(X)$$

$$\therefore \ cov(X,X) = var(X).$$

Remark 2: We will talk about correlation later but briefly mention that

$$\text{"The correlation btwn } X \text{ \& } Y\text{"} = \frac{cov(X,Y)}{sd(X)\, sd(Y)}$$

which measures how much X looks like $aY+b$ for some non-random a & b

Remark 3: $cov(X,Y)$ is a key component for computing $var(aX+bY)$.

Remark 4:

If X and Y are independent then $cov(X,Y)=0$.

... but if $cov(X,Y)=0$ it does not imply that X and Y are independent!

[here is an analogy:
if it is hot then I'll eat ice cream.
... but if I'm eating icecream that doesn't imply it must be hot out]

Covariance examples

Example: Let X and Y have joint PMF:

Find $cov(X,Y)$

$X \backslash Y$	0	1	$P_X(x)$
0	1/6	0	1/6
1	2/6	2/6	4/6
2	0	1/6	1/6
$P_Y(y)$	3/6	3/6	1

$E(X) = \frac{0}{6} + \frac{4}{6} + \frac{2}{6} = 1$

$E(Y) = \frac{1}{2}$

$E(XY) = 0\cdot 0\cdot \frac{1}{6} + 0\cdot 1\cdot 0 + 1\cdot 0\cdot \frac{2}{6}$
$+ 1\cdot 1\cdot \frac{2}{6} + 2\cdot 0\cdot 0 + 2\cdot 1\cdot 1/6$
$= 4/6$

$cov(X,Y) = \frac{4}{6} - 1\cdot\frac{1}{2} = \frac{1}{6}$

Example Let X, Y have joint PMF

$X \backslash Y$	0	1	
-1	0	1/3	
0	1/3	0	
1	0	1/3	
			1

Are X, Y independent? No

Is $cov(X,Y) = 0$?

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

$= (-1)(1)\frac{1}{3} + 0 + (1)(1)\frac{1}{3}$

$- (0)(\frac{2}{3})$

$= 0.$

So you can have $cov(X,Y) = 0$ even for dependent R.V.s

Master Formulas 1 & 2

Master Formula 1

$$E(aX+bY) = aE(X) + bE(Y).$$

Master Formula 2

$$\text{var}(aX+bY) = a^2\text{var}(X) + b^2\text{var}(Y) + 2ab\,\text{cov}(X,Y).$$

where a,b are non-random numbers and X, Y are random variables.

These two formulas are very general. There are some special cases that are useful to explicity state.

consequence 1:

- $E(aX+b) = aE(X)+b$
- $\text{var}(aX+b) = a^2\text{var}(X)$

consequence 2:

- $E(aX) = aE(X)$
- $\text{var}(aX) = a^2\text{var}(X)$
- $sd(aX) = |a|\,sd(X)$

Consequence 3:

- $E(X+Y) = E(X)+E(Y)$ always
- if X & Y are independendent then
  $\text{Var}(X+Y) = \text{var}(X) + \text{var}(Y)$.
- if X and Y are not independent then
  $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X,Y)$.

Consequence 3:

- $E(X_1+X_2+\cdots+X_n) = E(X_1)+E(X_2)+\cdots+E(X_n)$
- If $X_1,\ldots,X_n$ are independent then
  $\text{var}(X_1+\cdots+X_n) = \text{var}(X_1)+\cdots+\text{var}(X_n)$.

Consequence 4:

If $X_1, X_2, \ldots, X_n$ are independent with $E(X_i)=\mu$ & $\text{var}(X_i)=\sigma^2$ then

$$E\left(\frac{X_1+\cdots+X_n}{n}\right) = \mu$$

$$\text{var}\left(\frac{X_1+\cdots+X_n}{n}\right) = \frac{\sigma^2}{n}$$

Consequen 5:

$$E\left(\frac{X-E(X)}{sd(X)}\right) = 0$$

$$\text{var}\left(\frac{X-E(X)}{sd(X)}\right) = 1$$

Consequence 1 & 2

Tells you how to recompute E & var after changing units of measurment.

Example:

Randomly choose a student in class. Let X be their height in inches. Suppose

$$E(X) = 65.2 \text{ inches}$$
$$sd(X) = 5.1 \text{ inches}$$

Let $Y = 2.54X$ = height in cm.

Find $E(Y)$ & $sd(Y)$.

$$E(Y) = E(2.54X) \overset{MF1}{=} 2.54\,E(X) = 2.54(65.2) = 165.608 \text{ cm}$$

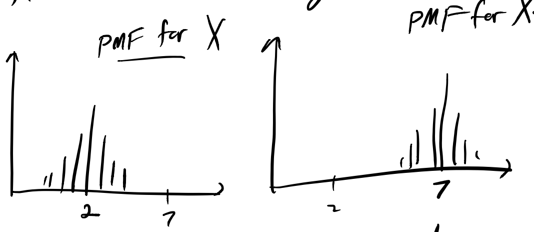
$$var(Y) \overset{MF2}{=} var(2.54X) = (2.54)^2 var(X) = (2.54)^2(5.1)^2 = 167.8 \text{ cm}^2$$

$$\therefore sd(Y) = \sqrt{167.8} \text{ cm.}$$

Note: It should make sense that

$$var(X+5) = var(X)$$

since $X+5$ is just a shifted version of X. So the variability doesn't change

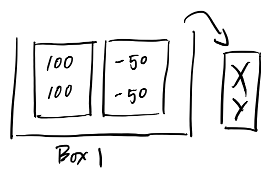


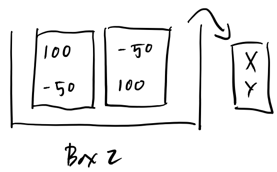
same variability but different expected value

Consequence 3 allows one to study portfolio risk.

Example:

Two investments





	expected profit $E(X+Y)$	Risk $sd(X+Y)$
Box 1	\$50	$sd(X+Y) > 0$.
Box 2	\$50	\$0

Both boxes have

$$E(X+Y) = E(X) + E(Y) = \left(\tfrac{100}{2} - \tfrac{50}{2}\right) + \left(\tfrac{100}{2} - \tfrac{50}{2}\right) = \$50$$

... but Box 1 is more risky.

The reason is that the dependency in Box 1 makes X, Y move up or down in unison.

The covariance btwn X & Y, quantifies exactly how risk is effected by dependence

To use $cov(X,Y)$ to see the impact on risk use:

$$\therefore var(X+Y) = var(X) + var(Y) + 2\underbrace{cov(X,Y)}$$

small cov ⇒ small risk
large cov ⇒ large risk

From Box 1 we have

$$E(XY) = 100 \cdot 100 \cdot \frac{1}{2} + (-50)(-50)\frac{1}{2}$$
$$= 6250$$
$$\therefore \quad cov(X,Y) = 6250 - (25)^2$$
$$= 5625$$

From Box 2 we have

$$E(XY) = 100(-50)\frac{1}{2} + 100(-50)\frac{1}{2}$$
$$= -5000$$
$$\therefore \quad cov(X,Y) = -5000 - (25)^2$$
$$= -5625$$

From both boxes

$$var(X) = var(Y)$$
$$= E(X^2) - (E(X))^2$$
$$= \frac{100^2}{2} + \frac{(-50)^2}{2} - (25)^2$$
$$= 5625$$

From Box 1:

$$var(X+Y) = var(X) + var(Y) + 2\,cov(X,Y)$$
$$= 2(5625) + 2(5625)$$
$$= 22{,}500$$
$$sd(X+Y) = \$150$$

From Box 2

$$var(X+Y) = 2(5625) + 2(-5625)$$
$$= 0$$
$$sd(X+Y) = \$0.$$

This explains the subprime Mortgage crash & why people invest in Gold.

Mortgage-backed securities:

$$\text{total return} = X_1 + X_2 + \cdots + X_{100,000}$$

Returns on individual Mortgages. each one has a chance of defaulting.

Previously assumed X_i's are nearly independent but bank practices introduced subtle dependence...

they can default in unison

$\Longrightarrow$ made total return risky.

Gold is typically a hedging investment i.e. often has negative cov with the market so adding it to a portfolio reduces risk.

Example

Allocate a fund of \$100 btwn two investments with future payouts X & Y (per \$1)

Let a = amount allocated to X
b = amount allocated to Y.

if $a=30$ & $b=70$ then your betting \$30 on X & \$70 on Y & Your total payout is

$$30X + 70Y = aX + bY$$

Suppose $sd(X) = 25$, $sd(Y) = 10$, $cov(X,Y) = -175$, $E(X) = E(Y)$

$$\therefore \quad sd(aX+bY) = \sqrt{a^2 var(X) + b^2 var(Y) + 2ab\, cov(X,Y)}$$
$$= \begin{cases} \$563.47 & \text{if } a=30 \text{ \& } b=70 \\ \$1000 & \text{if } a=0 \text{ \& } b=100 \end{cases}$$

Consequence 4 will be very important when studying random sampling in statistics.

Example

Your friend just fliped a coin 100 times & got 80 heads. Is this unusual?

Let $X_1 = \begin{cases} 1 & \text{if } 1^{st} \text{ flip is heads} \\ 0 & \text{if } 1^{st} \text{ flip is tails} \end{cases}$

$X_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ flip is heads} \\ 0 & \text{if } 2^{nd} \text{ flip is tails} \end{cases}$

$\vdots$

$X_{100} = \begin{cases} 1 & \text{if } 100^{th} \text{ flip is heads} \\ 0 & \text{if } 100^{th} \text{ flip is tails} \end{cases}$

Your friend observed

$$X_1 + X_2 + \cdots + X_{100} = 80$$

Lets Find $E(X_1 + \cdots + X_{100})$ & $sd(X_1 + \cdots + X_{100})$ to see what we would expect.

For each $i = 1, \dots, 100$

x	$p_{X_i}(x)$	$x\,p_{X_i}(x)$	$x^2 p_{X_i}(x)$
0	$\frac{1}{2}$	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	1	$\frac{1}{2}$	$\frac{1}{2}$

$\therefore\ E(X_i) = \frac{1}{2}$

$var(X_i) = \frac{1}{2} - \left(\frac{1}{2}\right)^2$

$= \frac{1}{4}$

$\therefore\ E(X_1 + \cdots + X_{100}) = \underbrace{E(X_1)}_{1/2} + \cdots + \underbrace{E(X_{100})}_{1/2}$

$= \frac{100}{2} = 50$

$var(X_1 + \cdots + X_{100}) = \underbrace{var(X_1)}_{1/4} + \cdots + \underbrace{var(X_{100})}_{1/4}$

$= \frac{100}{4} = 25$

$sd(X_1 + \cdots + X_{100}) = 5$

Your friend got 80 for $X_1 + \cdots + X_{100}$ which is $50 + 30$ ... i.e. 6-times the typical deviation from what we would have predicted.

$\therefore$ either your friend is lying or he has a trick coin!!

Consequence 5 is important for Z-scores

Z-scores are important for identifying unusual outcomes of a random variable.

In the last e.g. we had a R.V. $Y = X_1 + \cdots + X_{100}$ which had

$$E(Y) = 50$$

$$sd(Y) = 5$$

We said "for Y to be 80 it would need to 6-times the typical deviation from what we would predict".

This is the z-score of 75 for Y.

This is E(Y)

This is sd(Y)

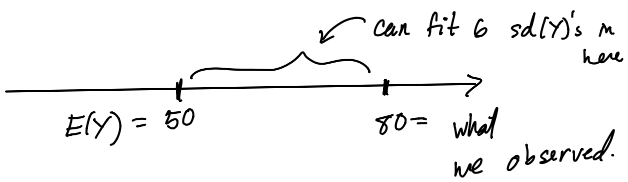
To find the "6-times" part we implicity did the following

$$\text{obs} \rightarrow \frac{80 - E(Y)}{sd(Y)} = 6$$

← subtract what we expect

← divide by the typical deviation

In otherwords you can fit 6 sd(Y)'s between 80 and E(Y)



Z-scores are themselfs R.V.s in the sense that the z-score for Y is

$$Z = \frac{Y - E(Y)}{sd(Y)} = \text{the number of typical deviations } Y \text{ is from } E(Y).$$

is a R.V.

Consequence 5 says

$$E(Z) = 0$$

$$Var(Z) = 1$$

Saying your friend observed

$$Y = 80$$

is the same as saying your friend observed the z-score of Y to be

$$Z = 6$$

Z-scores give a unit free indication if a observation is unusual or not.

A z-score within ± 2 is Normal

A z-score $\approx$ 3 is kinda unlikely

A z-score $\approx$ 4 is rare

A z-score $\approx$ 5 is extremely rare.

Example:

Choose a random person from Cali.

Let X = height in inches.

Suppose $E(X) = 67$ & $sd(X) = 8$

Suppose X is observed to be 69.2

How far is the observed X from $E(X)$ in unit of

- inches : $X - E(X) = 69.2 - 67 = 2.2$
- ft : $\frac{X - E(X)}{12} = \frac{2.2}{12} = 0.183$
- standard deviation:

$$\frac{X - E(X)}{sd(X)} = \frac{2.2}{8} = \underbrace{0.275}_{\text{z-score}}$$