Stat 260 Lecture Notes
Set 19 - Expected Value and Covariance For Joint Distributions

For a single discrete random variable X we saw that
$E(g(X)) = \sum_x g(x) \cdot f(x)$. $E(\sqrt{X}) = \sum \sqrt{x} \cdot f(x)$

For discrete random variables X and Y, we have that

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

where $p(x,y)$ is the joint pmf of X and Y.

Example 1: Suppose X and Y are discrete random variables with the joint pmf given below. Find $E(XY)$.

	$p(x,y)$	Y: 0	10	20	
X	5	0.15	0.30	0.15	0.60
	10	0.20	0.10	0.10	0.40
		0.35	0.40	0.25	

X \ Y: $X \cdot Y$	0	10	20
5	0	50	100
10	0	100	200

$$E(X \cdot Y) = \sum_x \sum_y x \cdot y \cdot p(x,y)$$
$$= 0(0.15) + (50)(.30) + 100(0.15) + (0)(0.2) + (100)(.10) + 200(0.10)$$
$$= 60$$

In general:
$E(XY) \neq E(X) \cdot E(Y)$

Rule: If X and Y are independent random variables, then $E(XY) = E(X) \cdot E(Y)$. $E(X \cdot Y) = \sum_x \sum_y x \cdot y \cdot p(x,y) = \sum_x \sum_y x \cdot y \cdot P(X=x) \cdot P(Y=y)$

The rule above comes from the fact that if X and Y are independent, then $p(x,y) = P(X = x) \cdot P(Y = y)$.

$\sum_x x \cdot P(X=x) \cdot \sum_y y \cdot P(Y=y)$

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

For a single discrete random variable X we saw that $V(X) = E((X-\mu)^2) = E(X^2) - [E(X)]^2$.

For discrete random variables X and Y, we have that $Cov(X,Y)$ $= E((X-\mu_X)(Y-\mu_Y)) = E(X\cdot Y) - E(X)\cdot E(Y)$. The value $Cov(X,Y)$ is called the **covariance** of X and Y.

↳ tells us about a linear relationship

Using the rule from the bottom of the last page, this means that if X and Y are indpendent, then $Cov(X,Y) = E(X\cdot Y) - E(X)\cdot E(Y) = E(X)\cdot E(Y) - E(X)\cdot E(Y) = 0$.

Rule: This doesn't work the other way around. If you have that $Cov(X,Y) = 0$, this doesn't necessarily mean that X and Y are independent.

Example 2: Suppose X and Y are discrete random variables with the joint pmf given below. Find $Cov(X,Y)$.

	$p(x,y)$	Y: 0	10	20	
X	5	0.15	0.30	0.15	0.60
	10	0.20	0.10	0.10	0.40
		0.35	0.40	0.25	

From ex. 1: $E(X\cdot Y) = 60$

For X: marginal pmf for X

X	5	10
$f(x)$	0.6	0.4

$E(X) = \sum x\cdot f(x)$
$= (5)(0.6) + (10)(0.4) = 7$

For y:

Y	0	10	20
$f(y)$	0.35	0.40	0.25

$E(Y) = \sum y\cdot f(y) = (0)(0.35) + (10)(0.40) + (20)(0.25) = 9$

$Cov(X,Y) = E(X\cdot Y) - E(X)\cdot E(Y)$
$= 60 - (9)(7)$
$= -3$

*we can get a neg. covarience (can be +, −, or 0) r: Sample Correlation coefficent ρ: population correlation coefficent

The **correlation coefficient** ρ is defined as

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Example 3: Use the X and Y variables with the pmf from Examples 1 and 2. Find the correlation coefficient ρ.

$Cov(X,Y) = -3$ from example 2

$V(X) = ?$

$V(Y) = ?$

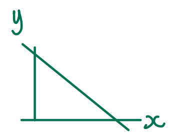
$V(X) = E(X^2) - (E(X))^2 = 55 - 7^2 = 6$

$V(Y) = E(Y^2) - (E(Y))^2 = 140 - 9^2 = 59$

$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)} \cdot \sqrt{V(Y)}} = \frac{-3}{\sqrt{6} \cdot \sqrt{59}} = -0.1594$

neg. linear relationship & weak linear relationship it is not close to -1

* same interpretation as before (how strong or weak linear relationship)



The correlation coefficient ρ is the population version of the sample correlation coefficient r we saw back in Set 3. That is, $\rho = Corr(X,Y)$ measures how much of a linear relationship X and Y have. We follow all the same rules we saw with r back in Set 3.

Rule: $-1 \leq \rho \leq 1$

Rule: If $\rho = \pm 1$, then there is a perfect linear relationship.

Going back to our expected value and variance rules, we follow the same shortcut rules as before:

↱ for expected value, we can pull out constants

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

have to square the constants

- $V(aX+bY+c) = V(aX+bY) = V(aX)+V(bY) = a^2V(X)+b^2V(Y)$, but only when X and Y are independent.

↳ for varience, adding a constant won't change the varience

For any discrete random variables X and Y (independent or not) we have that

$$V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$$

This can be generalized to

$$V(aX + bY + c) = V(aX + bY) = a^2V(X) + b^2V(Y) + 2abCov(X, Y)$$

For even more random variables X_1, X_2, ..., X_n we can generalize $V(a_1X_1 + a_2X_2 + \ldots + a_nX_n)$ to

$$\boxed{a_1^2V(X_1) + a_2^2V(X_2) + \ldots + a_n^2V(X_n) + 2\sum_{i<j} a_ia_jCov(X_i, X_j)}$$

on formula sheet

The last summation there says to work over all pairs of variables X_i and X_j where $i \neq j$ and count the $Cov(X_i, X_j)$ once. (That is, we don't count both $Cov(X_i, X_j)$ and $Cov(X_j, X_i)$ as they would be the same value.) So if we wanted to find $V(a_1X_1 + a_2X_2 + a_3X_3)$ we would calculate this as

$$a_1^2V(X_1) + a_2^2V(X_2) + a_3^2V(X_3)$$
$$+2a_1a_2Cov(X_1, X_2) + 2a_1a_3Cov(X_1, X_3) + 2a_2a_3Cov(X_2, X_3)$$

Note that when X_i and X_j are independent we have that $Cov(X_i, X_j) = 0$. So this generalized calculation of $V(a_1X_1 + a_2X_2 + \ldots + a_nX_n)$ agrees with our rule from before that $V(a_1X_1 + a_2X_2 + \ldots + a_nX_n) = a_1^2V(X_1) + a_2^2V(X_2) + \ldots + a_n^2V(X_n)$ when X_1, X_2, ..., X_n are all independent.

Let's practice one more expected value calculation.

equally likely

Example 4: Suppose we roll two fair 4-sided dice, and record the value of the absolute difference between the two dice. Find the expected value of the number that we would record.

X = # on die 1 record $|X-Y|$

Y = # on die 2

want to find $E(|X-Y|) = \sum_x \sum_y |x-y| \cdot p(x,y)$

joint pmf

$p(x,y)$	Y: 1	2	3	4
X: 1	1/16	1/16	1/16	1/16
2	⋮	⋮	⋮	⋮
3	⋮	⋮		⋮
4	1/16	1/16	1/16	1/16

$\|x-y\|$	Y: 1	2	3	4
x: 1	0	1	2	0
2	1	0	1	2
3	2	1	0	1
4	3	2	1	0

$E(|X-Y|) = (0)(1/16) + (1)(1/16) + (2)(1/16) + (3)(1/16) + \dots$ adding all the cases and prob. together

$= 20(1/16) = \frac{20}{16} = 5/4 = 1.25$

On average, the #s on the two dice differ by 1.25