1

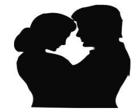
Stat 104: Quantitative Methods

Class 16: Jointly Distributed (Discrete) Random Variables

2

The Relationship Between Two Random Variables

- Previously, we talked about the distribution, mean and variance for a single random variable.
- However, like the concepts of correlation and covariance *for data*, there are similar ideas for random variables.



3

The Joint Distribution Function

- When we deal with two random variables, X and Y, it is convenient to work with *joint probabilities*. We define the joint probability distribution to be

$$P_{X,Y}(x,y) = P(X = x \, and \, Y = y)$$

- As usual, we require that

$$P_{X,Y}(x,y) \geq 0 \text{ for any pairs x, y}$$

$$\sum_{\text{all x,y}} P_{X,Y}(x,y) = 1$$

4

Example: Students and Museums

- Students in a college were classified according to years in school (X) and number of visits to a museum in the last year (0 for no visits, 1 for one visit, 2 for more than one visit). The joint probabilities in the accompanying table were estimated for these random variables.

Number of Visits (Y)	Years in School (X)			
	1	2	3	4
0	0.07	0.05	0.03	0.02
1	0.13	0.11	0.17	0.15
2	0.04	0.04	0.09	0.10

5

Example: More Credit, More Purchases

- The accompanying table shows, for credit card holders with one to three cards, the joint probabilities for number of cards owned (X) and number of credit purchases made in a week (Y).

Number of Cards (X)	Number of Purchases in Week (Y)				
	0	1	2	3	4
1	0.08	0.13	0.09	0.06	0.03
2	0.03	0.08	0.08	0.09	0.07
3	0.01	0.03	0.06	0.08	0.08

6

Marginal Distributions

- Suppose we are interested only in X, yet have to work with the joint distribution of X and Y. We can obtain the *marginal distribution* of X as follows.
- The marginal probabilities of X and Y are given by

$$P_X(x) = \sum_y P_{X,Y}(x,y) \;\; and \;\; P_Y(y) = \sum_x P_{X,Y}(x,y)$$

- As before, the term “marginal” merely describes how the distribution of X can be calculated from the joint distribution of X and another variable Y; row sums (or column sums) are calculated and placed “in the margin”

Example of Marginal Distributions

- Compute the marginal distributions

Number of Cards (X)	Number of Purchases in Week (Y) 0	1	2	3	4
1	0.08	0.13	0.09	0.06	0.03
2	0.03	0.08	0.08	0.09	0.07
3	0.01	0.03	0.06	0.08	0.08

x	P(X=x)
1	
2	
3	

y	P(Y=y)
0	
1	
2	
3	
4	

Independence

- Two random variables X and Y are called **independent** if the events (X=x) and (Y=y) are independent. That is,
- The random variables X and Y are independent if for ***all*** **values** of *x* and *y*:

$$P_{X,Y}(X = x \, and \, Y = y) = P_X(x)P_Y(y)$$

Example

- The approximately 100 million adult Americans in 1985 were roughly classified by education X and age Y as follows

Education (X)	(25-35) 30	(35-55) 45	(55-100) 70	$P_X(x)$
None (0)	.01	.02	.05	.08
Primary (1)	.03	.06	.10	.19
Secondary (2)	.18	.21	.15	.54
College (3)	.07	.08	.04	.19
$P_Y(y)$	.29	.37	.34	

Example (cont)

- What is P(X=3 and Y=30) ? 0.07
- Calculate the marginal probabilities
- Are X and Y independent ?

Take the 0.04 entry. If X and Y are independent, P(X and Y) = P(X)*P(Y)
Is 0.04 = 0.19 * 0.34? -> No
If it was, you'd have to check all others

Test for Independence

Number of Cards (X)	Number of Purchases in Week (Y) 0	1	2	3	4
1	0.08	0.13	0.09	0.06	0.03
2	0.03	0.08	0.08	0.09	0.07
3	0.01	0.03	0.06	0.08	0.08

Test for Independence

- Consider the following table of number of food complaints and service complaints.
- Are the random variables independent?

		Number of Service Complaints (X) 0	1	2	3	
	0	0.0216	0.0456	0.0408	0.012	0.12
Number of Food	1	0.0522	0.1102	0.0986	0.029	0.29
Complaints (Y)	2	0.0756	0.1596	0.1428	0.042	0.42
	3	0.0306	0.0646	0.0578	0.017	0.17
		0.18	0.38	0.34	0.1	

Checking like prev slide, here all the cells do match up

Conditional Distributions

- Let X and Y be jointly distributed random variables. Then the **conditional distribution** of X given Y is given by

$$P_{X|Y}(X=x|Y=y)=\frac{P_{X,Y}(x,y)}{P_Y(y)}$$

- Note that for a given y value, P(X=x|Y=y) is a probability distribution. That is, for any y

$$\sum_{all\ x\ values} P(X=x|Y=y)=1$$

Example: Does money make you happy?

Happiness (Y)

Salary (X)	0	1	2	$P_X(x)$
2.5	.03	.12	.07	.22
7.5	.02	.13	.11	.26
12.5	.01	.13	.14	.28
17.5	.01	.09	.14	.24
$P_Y(y)$	.07	.47	.46	1.0

Example (cont)

- Given the fact that you're very happy (i.e. Y=2), what is the *conditional distribution* of your salary ?
- That is, we want to compute P(X|Y=2)

x	P(X=x|Y=2)
2.50	.07/.26=.15
7.50	.11/.46=.24
12.50	.14/.46=.305
17.50	.14/.46=.305

Conditional distribution of salary

x	P(X=x)
2.50	0.22
7.50	0.26
12.50	0.28
17.50	0.24

Unconditional distribution of salary

Since the two are different, X and Y are not independent

Example: Condition on 3 Service complaints

		Number of Service Complaints (X)				
		0	1	2	3	
	0	0.0216	0.0456	0.0408	0.012	0.12
Number of Food	1	0.0522	0.1102	0.0986	0.029	0.29
Complaints (Y)	2	0.0756	0.1596	0.1428	0.042	0.42
	3	0.0306	0.0646	0.0578	0.017	0.17
		0.18	0.38	0.34	0.1	

y	P(Y=y|X=3)
0	.012/.1=.12
1	.029/.1=.29
2	.042/.1=.42
3	.017/.1=.17

y	P(Y=y)
0	0.12
1	0.29
2	0.42
3	0.17

Conditional and unconditional are the same -> Indpendent

Conditional Expectation

Mean of a sub-group

- One useful application of conditional distributions is in calculating conditional expectations. You will see a lot more of this when we get to regression analysis.
- The basic idea is that given a conditional distribution, we can also calculate a conditional expectation:

$$E(X|Y=y)=\sum_{all\ x\ values} xP(X=x|Y=y)$$

Given Y is some value, what's the expecation

If X and Y were independent, conditional expectation would be same as overall expectation

Example

- What is the expected salary for someone who is depressed?
- We need to compute E(X|Y=0). How do we do this?

x	P(X=x|Y=0)
2.5	.03/.07=.43
7.5	.02/.07=.29
12.5	.01/.07=.14
17.5	.01/.07=.14

$$E(X|Y=0)=\Sigma x_i P(X=x_i|Y=0)$$

=2.5(.43)+7.5(.29)+12.5(.14)+17.5(.14) = 7.45

Note that E(X|Y=2) = 11.325. Any conclusions ?

Combining Random Variables

19

- IF X and Y are independent, it is easy to combine random variables.
- That is, IF X and Y are independent,
- E(X+Y)=E(X)+E(Y) [actually always true]
- Var(X+Y)=Var(X)+Var(Y)
- To understand how to calculate E(X+Y) and Var(X+Y) for all scenarios, we need to first introduce the concepts of covariance and correlation for random variables.

Covariance and Correlation

20

- The variance of a random variable is a measure of its variability, and the covariance of two random variables is a measure of their joint variability.
- The covariance is a measure of the *linear association* of two random variables. Its sign reflects the direction of the association; if the variables tend to move in the same direction the covariance is positive. If the variables tend to move in opposite directions the covariance is negative.
- Similar ideas to what we saw with data several classes ago.

Covariance is a pain to calculate

21

- The covariance is a bit of a pain to calculate.

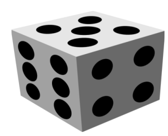
$$\sigma_{XY} = \sum_{i=1}^{N} [x_i - E(X)][(y_i - E(Y)] \, P(x_i, y_i)$$

- Three interesting facts: (1) Cov(X,X)=Var(X),
  (2) **if X and Y are independent**, Cov(X,Y)=0,
  (3) Cov(X,Y)=E(XY)-E(X)E(Y) (alternative formula)

Example

22

- Consider a 6 sided die and let X be the number on the top of the die and Y be the number on the bottom.
- Does anyone know the relationship between X and Y?



X+Y=7

Calculate the Covariance

23

- We will use the formula Cov(X,Y)=E(XY)-E(X)E(Y)
- For a die E(X)=E(Y)=3.5
- We need to find E(XY)

Probability	*X*	*Y*	*XY*	Prob × *XY*
1/6	1	6	6	6/6 = 1
1/6	2	5	10	10/6 = 5/3
1/6	3	4	12	12/6 = 2
1/6	4	3	12	12/6 = 2
1/6	5	2	10	10/6 = 5/3
1/6	6	1	6	6/6 = 1
				$E(XY)$ = sum = $9\frac{1}{3}$ = 9.333

- So Cov(X,Y)=9.33-(3.5)(3.5) = -2.91
- The covariance is negative because larger values of X are associated with smaller values of Y.

Covariance Example

24

- X=consumer satisfaction
- Y=number years living in a town

	x				
y	*1*	*2*	*3*	*4*	Total
1	.04	.14	.23	.07	.48
2	.07	.17	.23	.05	.52
Total	.11	.31	.46	.12	1

Covariance Example (cont)

- From the probability table we can calculate:

$$\mu_X = \sum_{i=1}^{4} X_i P(X_i) = 1(.11) + 2(.31) + 3(.46) + 4(.12) = 2.59$$

$$\mu_Y = \sum_{j=1}^{2} Y_j P(Y_j) = 1(.48) + 2(.52) = 1.52$$

$$\begin{aligned} E(XY) &= \sum_{j=1}^{m}\sum_{i=1}^{n} (X_i Y_j) P(X_i Y_j) \\ &= (1)(1)(.04) + (1)(2)(.14) + (1)(3)(.23) + (1)(4)(.07) + (2)(1)(.07) \\ &\quad + (2)(2)(.17) + (2)(3)(.23) + (2)(4)(.05) \\ &= 3.89 \end{aligned}$$

$$\boxed{\mathrm{Cov}(X,Y) = 3.89 - (2.59)(1.52) = -0.05}$$

Example: Stocks and Bonds (skip)

Stocks (S)

Treasury Bills (T)	-10%	0	10%	20%	$P_T(t)$
6%	0	0	.10	.10	
8%	0	.10	.30	.20	
10%	.10	.10	0	0	
$P_S(s)$					

Example using the longer formula for covariance

$$E(S) = -10(.1) + 0(.2) + 10(.4) + 20(.3) = 9$$

$$E(T) = 6(.2) + 8(.6) + 10(.2) = 8$$

$$\begin{aligned} Cov(S,T) &= (-10-9)(10-8)(.1) \\ &\quad + (0-9)(8-8)(.1) + (0-9)(10-8)(.1) \\ &\quad + (10-9)(6-8)(.1) + (10-9)(8-8)(.3) \\ &\quad + (20-9)(6-8)(.1) + (20-9)(8-8)(.2) \\ &= -9.1 \end{aligned}$$

The Covariance Matrix

- Sometimes the covariance between random variables is presented in a table, or matrix of the following form:

	X1	X2	X3
X1	Var(X1)	Cov(X1,X2)	Cov(X1,X3)
X2	Cov(X2,X1)	Var(X2)	Cov(X2,X3)
X3	Cov(X3,X1)	Cov(X3,X2)	Var(X3)

this is called a covariance matrix

Correlation: Covariance Rescaled

- Covariance can indicate whether X and Y have a positive, negative, or zero relation. Yet, it turns out not to be a good measure of association since it depends on the units of measurement.
- To eliminate this difficulty, we define the correlation:

$$\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

- This is a dimensionless measure of association.
- The correlation is always between -1 and 1, with 1 indicating a perfect positive linear relationship, -1 a perfect negative linear relationship and 0 no linear relationship between X and Y.

Dice Example Again

- X top of dice, Y= bottom of dice
- E(X)=3.5 and Var(X) = 2.91 (same for Y)
- We found earlier that Cov(X,Y) = -2.91
- Then the correlation is

$$\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} = \frac{-2.91}{\sqrt{2.91}\sqrt{2.91}} = -1$$

- This makes sense since X= 7-Y (perfect negative relationship)

Combinations of Random Variables

- If X and Y are independent

$$E(X+Y) = E(X) + E(Y)$$
$$Var(X+Y) = Var(X) + Var(Y)$$

- If X and Y are not independent

$$E(X+Y) = E(X) + E(Y)$$
$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

- The most general case

$$E((a+bX)+(c+dY)) = a + bE(X) + c + dE(Y)$$
$$Var((a+bX)+(c+dY)) = b^2 Var(X) + d^2 Var(Y) + 2bdCov(X,Y)$$

Example: Stock Prices

- Suppose your portfolio consists of $1000 under your mattress, 5 shares of Stock X and 10 shares of stock Y.
- The value (wealth) of your portfolio is then

$$W=1000 + 5X + 10Y$$

Joint probability table

- Suppose next months prices of the two stocks are modeled as the following joint probability table

		Stock Y Price			
		40	50	60	70
Stock	45	0.24	0.003333	0.003333	0.003333
X	50	0.003333	0.24	0.003333	0.003333
Price	55	0.003333	0.003333	0.24	0.003333
	60	0.003333	0.003333	0.003333	0.24

Find all the expected value and variances

- It may be shown (do this yourself-good practice!)
- E(X) = 53 and E(Y) = 55
- Var(X) = 31.3 and Var(Y) = 125
- COV(X,Y) = 59.17

Find E(W) and Var(W)

- Using

$$E((a+bX)+(c+dY)) = a+bE(X)+c+dE(Y)$$
$$Var((a+bX)+(c+dY)) = b^2Var(X)+d^2Var(Y)+2bdCov(X,Y)$$

$$\begin{aligned} E(W) &= 1000+5E(X)+10E(Y) \\ &= 1000+5(53)+10(55) \\ &= 1815 \end{aligned}$$

$$\begin{aligned} Var(W) &= 5^2Var(X)+10^2Var(Y)+2(5)(10)Cov(X,Y) \\ &= 25(31.3)+100(125)+100(59.17) \\ &= 19199.5 \end{aligned}$$

Summary: Combinations of Random Variables

- If X and Y are independent

$$E(X+Y) = E(X)+E(Y)$$
$$Var(X+Y) = Var(X)+Var(Y)$$

- If X and Y are not independent

$$E(X+Y) = E(X)+E(Y)$$
$$Var(X+Y) = Var(X)+Var(Y)+2Cov(X,Y)$$

- The most general case

$$E((a+bX)+(c+dY)) = a+bE(X)+c+dE(Y)$$
$$Var((a+bX)+(c+dY)) = b^2Var(X)+d^2Var(Y)+2bdCov(X,Y)$$

Example: Sums of Normals

- If X_1 and X_2 are each normally distributed

$$X_i \sim N(\mu_i, \sigma_i^2)$$

- Then the sum is normally distributed

$$aX_1 + bX_2 \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\sigma_{12})$$

- This rules holds for combining any number of normal random variables.

Example

- Suppose two rats A and B have been trained to navigate a large maze.
- X = Time of run for rat A X~N(80, 10^2)
- Y = Time of run for rat B Y~N(78, 13^2)
- On any given day what is the probability that rat A runs the maze faster than rat B?

P(X<Y)

Example

- Let D = X - Y be the difference in times of rats A and B
- If rat A is faster than rat B then D < 0 so we want to find P(D < 0).
- From rule for combining normals we know

D = X - Y ~ N(80 – 78,10^2 + 13^2) = N(2,269)

Means add/subtract. But Var(X-Y) = Var(X+Y)

Example

- Standardization

$$P(D<0) = P\left(\frac{D-2}{\sqrt{269}} < \frac{0-2}{\sqrt{269}}\right)$$
$$= P(Z < -0.122)$$
$$= 0.4514$$

Things you should know

- ❑Conditional Distributions
- ❑Conditional Expectation
- ❑Covariance and Correlation
- ❑Expectation and variance of a sum