

# Joint Distributions

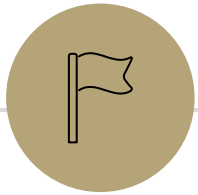
CSE 312 Autumn 25  
Lecture 26-27

# Announcements

This deck is for both today and Wednesday.

Wednesday's CC (CC27) is released already if you want to work ahead.

Midterm Second Chance is next Monday; we'll have a form you need to fill out (so we know how many copies to make).



# Multiple Random Variables

---

# This lecture and next lecture

Somewhat out-of-place content.

When we introduced multiple random variables, we've always had them be independent.

Because it's hard to deal with non-independent random variables.

Today and Wednesday are a crash-course in the toolkit for when you have multiple random variables and they aren't independent.

Going to focus on discrete RVs, we'll talk about continuous at the end.

# Why

Independent random variables are easier to interact with.

But sometimes you **want** to interact with the dependence

ML/Data science takes advantage of dependence: Netflix knows you like movie A; people liking movie A is dependent on people liking movie B, and so recommends you movie B

Random variables might be indicators for specific individual people liking movies, or “if we select a person at random, will they like this movie”

Our examples are artificial/simple; we’re just hoping to get the tools down.

# Joint PMF, support

For two (discrete) random variables  $X, Y$  their joint pmf

$$p_{X,Y}(x, y) = \mathbb{P}(X = x \cap Y = y)$$

When  $X, Y$  are independent then  $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ .

# Examples

Roll a blue die and a red die. Each die is 4-sided. Let  $X$  be the blue die's result and  $Y$  be the red die's result.

Each die (individually) is fair. But not all results are equally likely when looking at them both together.

$$p_{X,Y}(1,2) = 3/16.$$

$p_{X,Y}$	$X=1$	$X=2$	$X=3$	$X=4$
$Y=1$	1/16	1/16	1/16	1/16
$Y=2$	3/16	0	0	1/16
$Y=3$	0	2/16	0	2/16
$Y=4$	0	1/16	3/16	0

# Marginals

What if I just want to talk about  $X$ ?

Well, use the law of total probability:

$$\mathbb{P}(X = k) = \sum_{\text{partition } \{E_i\}} \mathbb{P}(X = k | E_i) \mathbb{P}(E_i)$$

and use  $E_i$  to be possible outcomes for  $Y$  For the dice example

$$\mathbb{P}(X = k) = \sum_{\ell=1}^4 \mathbb{P}(X = k | Y = \ell) \mathbb{P}(Y = \ell)$$

$$= \sum_{\ell=1}^4 \mathbb{P}(X = k \cap Y = \ell)$$

$$p_X(k) = \sum_{\ell=1}^4 p_{X,Y}(k, \ell)$$

$p_X(k)$  is called the “marginal” distribution for  $X$  (we “marginalized out”  $Y$ ) it’s the same pmf we’ve always used; the name comes from being in the margin of the paper when people printed these on paper.



# Marginals

$$p_X(k) = \sum_{\ell=1}^4 p_{X,Y}(k, \ell)$$

So

$$p_X(2) = \frac{1}{16} + 0 + \frac{2}{16} + \frac{1}{16} = \frac{4}{16}$$

$p_{X,Y}$	$X=1$	$X=2$	$X=3$	$X=4$
$Y=1$	1/16	1/16	1/16	1/16
$Y=2$	3/16	0	0	1/16
$Y=3$	0	2/16	0	2/16
$Y=4$	0	1/16	3/16	0

# Different dice

Roll two fair dice independently.  
Let  $U$  be the minimum of the two rolls and  $V$  be the maximum

Are  $U$  and  $V$  independent?

Write the joint distribution in the table

What's  $p_U(z)$ ? (the marginal for  $U$ )

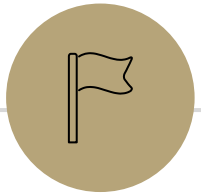
$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$				
$V=2$				
$V=3$				
$V=4$				

# Different dice

Roll two fair dice independently.  
Let  $U$  be the minimum of the two rolls and  $V$  be the maximum

$$p_U(z) = \begin{cases} \frac{7}{16} & \text{if } z = 1 \\ \frac{5}{16} & \text{if } z = 2 \\ \frac{3}{16} & \text{if } z = 3 \\ \frac{1}{16} & \text{if } z = 4 \\ 0 & \text{otherwise} \end{cases}$$

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16



# Expectations and LTE

---

# Joint Expectation

## Expectations of joint functions

For a function  $g(X, Y)$ , the expectation can be written in terms of the joint pmf.

$$\mathbb{E}[g(X, Y)] = \sum_{x \in \Omega_X} \sum_{y \in \Omega_Y} g(x, y) \cdot p_{X,Y}(x, y)$$

This definition hopefully isn't surprising at this point (it's the value of  $g$  times the probability  $g$  takes on that value), but it's good to see.

# Expectation of a function of two RVs

What's  $\mathbb{E}[UV]$  for  $U, V$  from the last slide?

# Expectation of a function of two RVs

What's  $\mathbb{E}[UV]$  for  $U, V$  from the last slide?

$$\begin{aligned} & \sum_{u \in \Omega_U} \sum_{v \in \Omega_V} uv \cdot p_{U,V}(u, v) \\ &= 1 \cdot 1 \cdot \frac{1}{16} + 1 \cdot 2 \cdot \frac{2}{16} + 1 \cdot 3 \cdot \frac{2}{16} + 2 \cdot 2 \cdot \frac{1}{16} + 2 \cdot 3 \cdot \frac{2}{16} + 2 \cdot 4 \cdot \frac{2}{16} + \\ & \quad 3 \cdot 3 \cdot \frac{1}{16} + 3 \cdot 4 \cdot \frac{2}{16} + 4 \cdot 4 \cdot \frac{1}{16} \\ &= \frac{92}{16} = \frac{23}{4} = 5.75. \end{aligned}$$

# Conditional Expectation

Waaaaaay back when, we said conditioning on an event creates a new probability space, with all the laws holding.

So we can define things like “conditional expectations” which is the expectation of a random variable in that new probability space.

$$\mathbb{E}[X|E] = \sum_{x \in \Omega_X} x \cdot \mathbb{P}(X = x|E)$$

$$\mathbb{E}[X|Y = y] = \sum_{x \in \Omega_X} x \cdot \mathbb{P}(X = x|Y = y)$$



# Conditional Expectations

All your favorite theorems are still true.

For example, linearity of expectation still holds

$$\mathbb{E}[(aX + bY + c) | E] = a\mathbb{E}[X|E] + b\mathbb{E}[Y|E] + c$$

# Law of Total Expectation

Let  $A_1, A_2, \dots, A_k$  be a partition of the sample space, then

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X|A_i] \mathbb{P}(A_i)$$

Let  $X, Y$  be discrete random variables, then

$$\mathbb{E}[X] = \sum_{y \in \Omega_Y} \mathbb{E}[X|Y = y] \mathbb{P}(Y = y)$$

Similar in form to law of total probability, and the proof goes that way as well.

# LTE

You will flip 2 (independent, fair coins). Call the number of heads  $X$ . Then (independently of the coin flips) draw an exponential random variable  $Y$  from the distribution  $\text{Exp}(X + 1)$ .

What is  $\mathbb{E}[Y]$ ?

# LTE

You will flip 2 (independent, fair coins). Call the number of heads  $X$ . Then (independently of the coin flips) draw an exponential random variable  $Y$  from the distribution  $\text{Exp}(X + 1)$ .

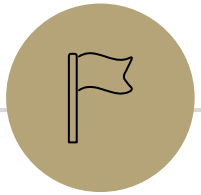
What is  $\mathbb{E}[Y]$ ?

$$\mathbb{E}[Y]$$

$$= \mathbb{E}[Y|X = 0]\mathbb{P}(X = 0) + \mathbb{E}[Y|X = 1]\mathbb{P}(X = 1) + \mathbb{E}[Y|X = 2]\mathbb{P}(X = 2)$$

$$= \mathbb{E}[Y|X = 0] \cdot \frac{1}{4} + \mathbb{E}[Y|X = 1] \cdot \frac{1}{2} + \mathbb{E}[Y|X = 2] \cdot \frac{1}{4}$$

$$= \frac{1}{0+1} \cdot \frac{1}{4} + \frac{1}{1+1} \cdot \frac{1}{2} + \frac{1}{2+1} \cdot \frac{1}{4} = \frac{7}{12}.$$



# Covariance

---

# Covariance

We sometimes want to measure how “intertwined”  $X$  and  $Y$  are – how much knowing about one of them will affect the other.

If  $X$  turns out “big” how likely is it that  $Y$  will be “big” how much do they “vary together”

## Covariance

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

## Covariance

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

If  $X, Y$  go in the same direction

If  $X, Y$  go in the opposite directions

# Covariance

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

That's consistent with our previous knowledge for independent variables. (for  $X, Y$  independent,  $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ ).

You and your friend are playing a game, you flip a coin: if heads you pay your friend a dollar, if tails they pay you a dollar. Let  $X$  be your profit and  $Y$  be your friend's profit.

What is  $\text{Var}(X + Y)$ ?

Before you calculate, make a prediction. What should it be?



# Covariance

You and your friend are playing a game, you flip a coin: if heads you pay your friend a dollar, if tails they pay you a dollar. Let  $X$  be your profit and  $Y$  be your friend's profit.

What is  $\text{Var}(X + Y)$ ?

$$\text{Var}(X) = \text{Var}(Y) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 1 - 0^2 = 1$$

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

$$\mathbb{E}[XY] = \frac{1}{2} \cdot (-1 \cdot 1) + \frac{1}{2} (1 \cdot -1) = -1$$

$$\text{Cov}(X, Y) = -1 - 0 \cdot 0 = -1.$$

$$\text{Var}(X + Y) = 1 + 1 + 2 \cdot -1 = 0$$

# Covariance, Another example

Let  $X$  be a Bernoulli RV with probability  $p$  of success.

Let  $Y = X$  ( $Y$  is  $X$ , not an iid copy, literally the same experiment)

Let  $Z = -X$

Let  $W$  be an independent Bernoulli, identically distributed to  $X$

Find

$\text{Cov}(X, Y)$ ,  $\text{Cov}(X, Z)$ ,  $\text{Cov}(X, W)$

# Covariance, Another example

Let  $X$  be a Bernoulli RV with probability  $p$  of success.

Let  $Y = X$  ( $Y$  is  $X$ , not an iid copy, literally the same experiment)

Let  $Z = -X$

Let  $W$  be an independent Bernoulli, identically distributed to  $X$

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

$$= (1 \cdot 1 \cdot p + 0 \cdot 0 \cdot [1 - p]) - p \cdot p$$

$$= p - p^2 = p(1 - p)$$

Hey, that's the variance of  $X$ . This is a pattern:  $\text{Cov}(X, X) = \text{Var}(X)$

# Covariance, Another example

Let  $X$  be a Bernoulli RV with probability  $p$  of success.

Let  $Y = X$  ( $Y$  is  $X$ , not an iid copy, literally the same experiment)

Let  $Z = -X$

Let  $W$  be an independent Bernoulli, identically distributed to  $X$

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

$$= (1 \cdot -1 \cdot p + 0 \cdot -0 \cdot [1 - p]) - (p \cdot [-p])$$

$$= -p - [-p^2] = -p(1 - p)$$

General pattern:  $\text{Cov}(X, -Y) = -\text{Cov}(X, Y)$

# Covariance, Another example

Let  $X$  be a Bernoulli RV with probability  $p$  of success.

Let  $Y = X$  ( $Y$  is  $X$ , not an iid copy, literally the same experiment)

Let  $Z = -X$

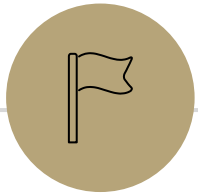
Let  $W$  be an independent Bernoulli, identically distributed to  $X$

$$\text{Cov}(X, W) = \mathbb{E}[XW] - \mathbb{E}[X]\mathbb{E}[W]$$

$$= (1 \cdot 1 \cdot p^2 + 1 \cdot 0 \cdot p[1 - p] + 0 \cdot 1 \cdot [1 - p]p + 0 \cdot 0 \cdot [1 - p]^2) - (p \cdot [p])$$

$$= (p^2) - p^2 = 0$$

General pattern: if  $X, Y$  independent  $\text{Cov}(X, Y) = 0$



# Conditional Distributions

---

# Different dice

Roll two fair dice independently.  
Let  $U$  be the minimum of the two rolls and  $V$  be the maximum

What is  $\mathbb{P}(U = 2|V = 3)$ ?

$$\frac{\mathbb{P}(U=2 \cap V=3)}{\mathbb{P}(V=3)} = \frac{2/16}{5/16} = \frac{2}{5}$$

$$p_{U|V}(2|3) = \frac{2}{5}$$

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

# Different dice

Find these values

$$p_{V|U}(2|1) =$$

$$p_{U|V}(1|2) =$$

$$p_{U|V}(4|1) =$$

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16



# Different dice

Find these values

$$p_{V|U}(2|1) = \frac{p_{V,U}(2,1)}{p_U(1)} = \frac{2/16}{7/16} = \frac{2}{7}$$

$$p_{U|V}(1|2) = \frac{p_{U,V}(1,2)}{p_V(2)} = \frac{2/16}{3/16} = \frac{2}{3}$$

$$p_{U|V}(4|1) = \frac{p_{U,V}(4,1)}{p_V(1)} = \frac{0}{1/16} = 0$$

$p_{U,V}$	$U=1$	$U=2$	$U=3$	$U=4$
$V=1$	1/16	0	0	0
$V=2$	2/16	1/16	0	0
$V=3$	2/16	2/16	1/16	0
$V=4$	2/16	2/16	2/16	1/16

# What about the continuous versions?

In the continuous case, everything is still a density function, not a mass function.

Joint density

Marginal density

Conditional density

Expectations, conditional expectations integrate  $x \cdot (\text{cond})\text{density}(x)$

You aren't getting a probability, you're getting a density; have to integrate to get a value.

# Analogue for continuous

Everything we saw today has a continuous version.

There are “no surprises”– replace pmf with pdf and sums with integrals.

	Discrete	Continuous
<b>Joint PMF/PDF</b>	$p_{X,Y}(x, y) = P(X = x, Y = y)$	$f_{X,Y}(x, y) \neq P(X = x, Y = y)$
<b>Joint CDF</b>	$F_{X,Y}(x, y) = \sum_{t \leq x} \sum_{s \leq y} p_{X,Y}(t, s)$	$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$
<b>Normalization</b>	$\sum_x \sum_y p_{X,Y}(x, y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
<b>Marginal PMF/PDF</b>	$p_X(x) = \sum_y p_{X,Y}(x, y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
<b>Expectation</b>	$E[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y)$	$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$
<b>Conditional PMF/PDF</b>	$p_{X Y}(x   y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$	$f_{X Y}(x   y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$
<b>Conditional Expectation</b>	$E[X   Y = y] = \sum_x x p_{X Y}(x   y)$	$E[X   Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x   y) dx$
<b>Independence</b>	$\forall x, y, p_{X,Y}(x, y) = p_X(x) p_Y(y)$	$\forall x, y, f_{X,Y}(x, y) = f_X(x) f_Y(y)$

# Conditioning on probability 0

We said for discrete spaces, when  $\mathbb{P}(B) = 0$ ,  $\mathbb{P}(A|B)$  is undefined

How can you condition on something that doesn't happen?

Also, how can you have  $\mathbb{P}(B)$  in the denominator?

For continuous spaces, we have to use densities to avoid the problem, but we can avoid the problem with densities!

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$\mathbb{P}(Y = y)$  is 0, but the density might not be 0 there so this expression can be defined (and it works!).

If density is 0 for  $Y = y$ , the conditional density is undefined there.

# A note on independence

The definition of independence says  $X, Y$  independent if and only if  $p_{X,Y}(x, y) = p_X(x)p_Y(y)$  or  $f_{X,Y}(x, y) = f_X(x)f_Y(y)$  (as appropriate)

There's often a nice shortcut. If  $X, Y$  are independent then joint support of  $X, Y$  (denoted  $\Omega_{X,Y}$ ) must be  $\Omega_X \times \Omega_Y$ .

Joint support is  $\{(x, y): p_{X,Y}(x, y) > 0\}$ .

Often easier to verify dependence when those are different (especially in the continuous case).

But note this is a single implication not an if-and-only-if.

# Continuous definitions and theorems

Conditional expectation:

$$\mathbb{E}[X|Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y) \, dx$$

LTE:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \mathbb{E}[X|Y = y] \cdot f_Y(y) \, dy$$

LTP:

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A|X = x) \cdot f_X(x) \, dx$$

$X$  is continuous; integrating over all values for  $X$  gives the full space