More Joint Distributions

CSE 312 Winter 25
Lecture 19

Announcements

Ed's interface for the coding question is broken at the moment.

The underlying code is fine, but it won't import a library we need.

Gradescope can run the code properly.

Robbie has a ticket with Ed support to get our environment fixed; in the meantime, the coding question will be due with HW7 not HW6 in case you weren't able to work on it. That's just the coding question; 1-3 (including the math for distinct elements) still due Wednesday.

Robbie is traveling at the end of this week (at SIGCSE)

TAs will guest lecture on Wednesday/Friday.

Robbie will have access to email but will be slower.

Multiple Random Variables

Analogues for continuous

Everything we saw today has a continuous version.

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

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

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

$p_{X,Y}(x,y) = p_{Y,X}(y,x)$

Find these values

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

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

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

$p_{U,V}$	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 $x \cdot$ (cond)density(x)

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

Analogues for continuous

Everything we saw today has a continuous version.

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

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

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!

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

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

$\mathbb{P}(Y = y)$ 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 $\Omega_X \times \Omega_Y$.

$f_{x,y}(x,y) > 0$

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

Often easier to verify <u>dependence</u> 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:

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

LTE:

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

LTP:

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

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

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

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

Covariance

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

If X, Y go in the same direction

$X > E[X]$ $X - E[X] > 0$ $> +$
$Y > E[Y]$ $Y - E[Y] > 0$

$X < E[X]$ $X - E[X] < 0$ $> +$
$Y < E[Y]$ $Y - E[Y] < 0$

If X, Y go in the opposite directions

$X > E[X]$ $X - E[X] > 0$
$Y < E[Y]$ $Y - E[Y] < 0$

$X < E[X]$ < 0
> 0

Covariance

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

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 $\mathrm{Var}(X+Y)$?

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

Covariance

$Y^2 = \begin{cases} 1^2 & \text{w/p } p \\ (-1)^2 & \text{w/p } 1-p \end{cases}$

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.

Is $\mathrm{Cov}(X, Y) \neq 0$

What is $\mathrm{Var}(X+Y)$? $\mathrm{Cov}(X, Y)$?

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

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

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

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

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

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, indentically 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, indentically distributed to X

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

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

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

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

Covariance, Another example

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

Let $Y = X$ (Y **<u>is</u>** X, not an iid copy, literally the same experiment)

Let $Z = -X$

Let W be an independent Bernoulli, indentically distributed to X

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

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

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

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

Covariance, Another example

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

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

Let $Z = -X$

Let W be an independent Bernoulli, indentically distributed to X

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

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

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

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

A Few Notes

Covariance is an un-normalized number.

It measures both how intertwined X, Y are and in some sense how much X, Y vary in the first place (if you multiply both X, Y by 2, the strength of the relationship intuitively is the same, but covariance increases).

If you want just the strength of the relationship, you probably want the "correlation coefficient": $\frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$ always between -1 and 1.

Covariance directly measures only "linear" relationships; if Y depends on X^2, the covariance might not be as high as you expect.

If dealing with real data, look at a plot to see if you should be looking for a linear relationship in the first place.

A Continuous-ish Example

Recall from Friday: 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 $\text{Exp}(X+1)$.

Let's find the PDF of Y.

Let $g_\lambda(x) = \lambda e^{-\lambda x}$, i.e. density for Exp($\lambda$) (for $x \geq 0$)

$f_Y(y) = g_1(y) \cdot \frac{1}{4} + g_2(y) \cdot \frac{1}{2} + g_3(y) \cdot \frac{1}{4}$

$f_Y(y) = \frac{1}{4} e^{-y} + \frac{1}{2} \cdot 2 \cdot e^{-2y} + \frac{1}{4} \cdot 3e^{-3y}$ (for $y \geq 0$)

Notice this isn't an exponential random variable!

A Continuous-ish Example

Now we can check that expectation…

$$\mathbb{E}[Y] = \int_0^\infty y\left(\frac{1}{4}e^{-y} + \frac{1}{2}\cdot 2\cdot e^{-2y} + \frac{1}{4}\cdot 3e^{-3y}\right)\mathrm{d}y$$

$$= \int_0^\infty y\cdot\frac{1}{4}e^{-y}dy + \int_0^\infty y\,e^{-2y}dy + \int_0^\infty y\frac{1}{4}\cdot 3e^{-3y}dy$$

Integral of ye^{-y} will be 1, since that's the expectation of Exp(1)

$$= \frac{1}{4}\cdot 1 + \frac{1}{2}\int_0^\infty 2y\,e^{-2y}dy + \frac{1}{4}\cdot\int_0^\infty y\,3e^{-3y}dy$$

Setup for same trick, Exp(2), Exp(3)

$$= \frac{1}{4}\cdot 1 + \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{3} = \frac{1}{4} + \frac{1}{4} + \frac{1}{12} = \frac{7}{12}$$