

Stats 117 Lecture 13

Joint Distributions

Prof. John Duchi

Autumn 2026

13-1

Organization

- ▶ two genes: how are X and Y related?
- ▶ joint pmf and independence
 - ▶ linkage and recombination
- ▶ marginal distributions
- ▶ conditional distributions

13-2

Two genes

both parents carry $AaBb$, recessive alleles a and b of two genes

- ▶ X = number of dominant A alleles child inherits
- ▶ Y = number of dominant B alleles child inherits
- ▶ $X \sim$
- ▶ $Y \sim$

question: relationship between X and Y ?

13-3

Joint pmf

Definition

The *joint pmf* of random variables X and Y is

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

- ▶ $p_{X,Y}(x,y) \geq 0$ and $\sum_x \sum_y p_{X,Y}(x,y) = 1$

table:

13-4

Independent random variables

Definition

Discrete random variables X and Y are *independent* if

$$P(X = x, Y = y) = P(X = x) P(Y = y) \text{ for all } x, y,$$

i.e., $p_{X,Y}(x, y) = p_X(x) p_Y(y)$.

- events $\{X = x\}$ and $\{Y = y\}$ independent for every x and y

13-5

Independent assortment

genes on different chromosomes $\Rightarrow X$ and Y independent

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

	$x = 0$	$x = 1$	$x = 2$
$y = 0$			
$y = 1$			
$y = 2$			

13-6

Smoking and cancer

$$X = \begin{cases} 1 & \text{if smoker} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad Y = \begin{cases} 1 & \text{if lung cancer} \\ 0 & \text{otherwise.} \end{cases}$$

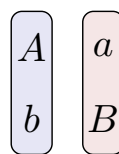
	$x = 0$	$x = 1$
$y = 0$	$72/100$	$20/100$
$y = 1$	$3/100$	$5/100$

► not independent

13-7

Linkage

genes on the *same* chromosome: each parent has



and passes one chromosome (Ab or aB) to the child, each with probability $\frac{1}{2}$, independently

► possible outcomes:

13-8

Linkage: joint pmf

	$x = 0$	$x = 1$	$x = 2$
$y = 0$			
$y = 1$			
$y = 2$			

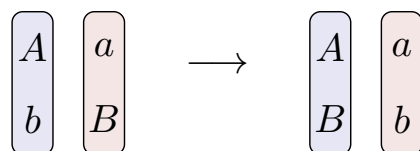
► relationship between X and Y :

► independent?

13-9

Recombination (fill this in yourself)

each parent's chromosomes *recombine* with probability .1, independently



law of total probability: condition on number of recombinations

$$p_{X,Y}(1,0) =$$

13-10

Recombination (fill this in yourself): computing the joint
given one recombination, child equally likely to get

$$(AB, Ab), (AB, aB), (ab, Ab), (ab, aB)$$

► $P(X = 1, Y = 0 \mid \text{no recombination}) =$

► $P(X = 1, Y = 0 \mid \text{one recombination}) =$

► $P(X = 1, Y = 0 \mid \text{two recombinations}) =$

13-11

Three joint pmfs

independent assortment				linkage			
	0	1	2		0	1	2
0	.0625	.125	.0625	0	0	0	.25
1	.125	.25	.125	1	0	.5	0
2	.0625	.125	.0625	2	.25	0	0

recombination			
	0	1	2
0	.0025	.045	.2025
1	.045	.41	.045
2	.2025	.045	.0025

► funny thing: all consistent with $X, Y \sim \text{Bin}(2, \frac{1}{2})$

13-12

Marginal distributions

Definition

The *marginal pmfs* of X and Y are

$$p_X(x) = \sum_y p_{X,Y}(x, y), \quad p_Y(y) = \sum_x p_{X,Y}(x, y).$$

why: events $\{Y = y\}$ partition Ω

13-13

Smoking marginals

	$x = 0$	$x = 1$	$p_Y(y)$
$y = 0$	$72/100$	$20/100$	
$y = 1$	$3/100$	$5/100$	
$p_X(x)$			

► $X \sim$

► $Y \sim$

13-14

Marginals do not determine joint

many joint pmfs consistent with

$$X \sim \text{Bernoulli}\left(\frac{1}{4}\right) \quad Y \sim \text{Bernoulli}\left(\frac{2}{25}\right)$$

- ▶ joint $\Rightarrow$ marginals (sum)
- ▶ marginals $\nRightarrow$ joint
- ▶ exception: X and Y independent

13-15

Conditional distributions

Definition

The *conditional pmf* of Y given $X = x$ is

$$p_{Y|X}(y | x) := P(Y = y | X = x) = \frac{p_{X,Y}(x, y)}{p_X(x)}$$

for x with $p_X(x) > 0$.

- ▶ same for $p_{X|Y}(x | y)$
- ▶ check: $\sum_y p_{Y|X}(y | x) = 1$

13-16

Smoking conditionals

	$x = 0$	$x = 1$
$y = 0$	$72/100$	$20/100$
$y = 1$	$3/100$	$5/100$

$$P(Y = 1 \mid X = 1) =$$

$$P(Y = 1 \mid X = 0) =$$

13-17

Independence and conditionals

Proposition

X and Y are independent if and only if

$$p_{Y|X}(y \mid x) = p_Y(y)$$

for all x (with $p_X(x) > 0$) and y .

- learning X tells you nothing about Y

why:

13-18