Lecture 11

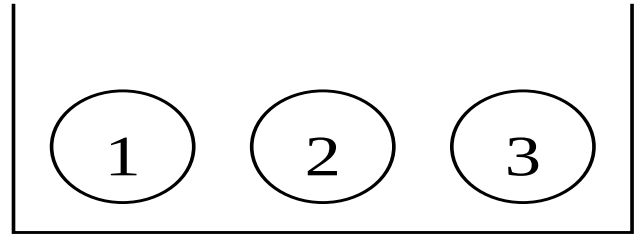
Bivariate distibutions

Marginal distributions

Conditional distributions

Independence

Joint distributions



- Two balls are drawn at random, <u>without</u> replacement:
  X = number on first ball. Y = number on second ball.

		Possible values for X			Dist'n of Y
		1	2	3	(row sums)
Possible	1	0	1/6	1/6	1/3
values	2	1/6	0	1/6	1/3
for Y.	3	1/6	1/6	0	1/3
	Dist'n of X (column sums)	1/3	1/3	1/3	1 (Total sum)

- Another example: A ball is drawn at random. X = number on ball. X fair coins are tossed. Y = number of heads.

$$P\{X = x, Y = y\} \;=\; P\{Y = y \big| X = x\} \times P\{X = x\}.$$

		Possible values for X			Dist'n of Y
		1	2	3	(row sums)
	0	1/6	1/12	1/24	7/24
Possible	1	1/6	2/12	3/24	11/24
values	2	0	1/12	3/24	5/24
for Y.	3	0	0	1/24	1/24
	Dist'n of X (column sums)	1/3	1/3	1/3	1 (Total sum)

Marginal distibutions

The marginal of X

$$
\begin{aligned}
p_X(x) &:= P\{X = x\} \\
&= P\left\{\bigcup_y \{X = x, Y = y\}\right\} \\
&= \sum_y P\{X = x, Y = y\} \\
&= \sum_y p_{XY}(x, y)
\end{aligned}
$$

The marginal of Y

$$
p_Y(y) := P\{Y = y\} = \sum_x p_{XY}(x, y)
$$

- Note: Identical marginals $\not\Rightarrow$ identical rv's.

Example:

Choose a pair at random from the set

$$\{(n, m) : 1 \leq n, m \leq 4,\ n + m \leq 5\}.$$

X = first number, Y = second number.
Find the marginals of X and Y.

$Y \backslash X$	1	2	3	4	Y Marginal
1	1/10	1/10	1/10	1/10	4/10
2	1/10	1/10	1/10	-	3/10
3	1/10	1/10	-	-	2/10
4	1/10	-	-	-	1/10
X Marginal	4/10	3/10	2/10	1/10	1

Conditional distributions

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

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

- Choosing points from $\{n + m \leq 5\}$.

$P\{X = x\|Y = 2\}$					
X	1	2	3	4	
Prob	1/3	1/3	1/3	0	1

- For each y, $p_{X|Y}(\cdot|y)$ is a probability function.

$$\sum_x p_{X|Y}(x|y) \quad = \quad \sum_x \frac{p_{XY}(x,y)}{p_Y(y)} \quad = \quad \frac{1}{p_Y(y)} \sum_x p_{XY}(x,y) \quad = \quad 1.$$

- One of three balls is drawn at random.

  X = number on ball.

  X fair coins are tossed.

  Y = number of heads.

$X \backslash Y$	0	1	2	3	
$P_{XY}(3, y)$	1/24	3/24	3/24	1/24	1/3
$P_{Y\|X}(y\|3)$	1/8	3/8	3/8	1/8	1

Because

$$p_{X|Y}(x|y) = \frac{p_{XY}(x,y)}{p_Y(y)} \quad \Rightarrow \quad p_{XY}(x,y) = p_{X|Y}(x|y) \times p_Y(y)$$

$$p_{Y|X}(y|x) = \frac{p_{XY}(x,y)}{p_X(x)} \quad \Rightarrow \quad p_{XY}(x,y) = p_{Y|X}(y|x) \times p_X(x)$$

Bivariate distributions

Objects of interest

Joint probabilities	$P\{X = x, Y = y\}$	$p_{XY}(x, y)$
Marginal probs	$P\{X = x\}$	$p_X(x) = \sum_y p_{XY}(x, y)$
	$P\{Y = y\}$	$p_Y(y) = \sum_x p_{XY}(x, y)$
Conditional probs	$P\{X = x \mid Y = y\}$	$p_{X\mid Y}(x\mid y) = \frac{p_{XY}(x,y)}{p_Y(y)}$
	$P\{Y = y \mid X = x\}$	$p_{Y\mid X}(y\mid x) = \frac{p_{XY}(x,y)}{p_X(x)}$

Independence

- X and Y are called independent $\iff$

$$P\{X \in A,\ Y \in B\} = P\{X \in A\} \times P\{Y \in B\}$$

for every pair of sets A, B.

- FACT: X and Y (discrete) are independent $\iff$ for every pair x, y

$$p_{XY}(x, y) = p_X(x) \times p_Y(y)$$

$\Rightarrow$ Take $A = \{x\}$ and $B = \{y\}$.

$\Leftarrow$ Consider the example $A = \{x_i\}_{i=1}^n$ and $B = \{y_i\}_{i=1}^m$, all points <u>distinct</u>. Then

$$
\begin{aligned}
P\{X \in A,\ Y \in B\} &= P\left\{\bigcup_{i,j}(X = x_i,\ Y = y_j)\right\} \\
&= \sum_{i,j} P\{X = x_i,\ Y = y_j\} \\
&= \sum_{i,j} P\{X = x_i\}\, P\{Y = y_j\} \\
&= \sum_{i} P\{X = x_i\} \times \sum_{j} P\{Y = y_j\} \\
&= P\{X \in A\} \times P\{Y \in B\}
\end{aligned}
$$

and we need consider no more cases!

- Note: Lack of independence is now <u>very</u> easy to check, since we need find only <u>one</u> pair (x, y) for which

$$p_{XY}(x, y) \neq p_X(x) \times p_Y(y).$$

Facts about independence

- Zeroes in the table: If there is a x with $p_X(x) > 0$, and a y with $p_Y(y) > 0$, but $p_{XY}(x, y) = 0$, then X and Y are dependent.

- X and Y independent
  $\iff$ $p_{X|Y}(x|y)$ does not depend on y
  $\iff$ $p_{Y|X}(y|x)$ does not depend on x.

$$p_{X|Y}(x|y) = \frac{p_{XY}(x, y)}{p_Y(y)} = \frac{p_X(x) \times p_Y(y)}{p_Y(y)} = p_X(x)$$

Example

- A box contains R red and B blue balls. $R + B = N$.
- K balls are drawn <u>without</u> replacement, and X are red.
- Another K are drawn, without replacement. Y are red.
- What is the distribution of Y? (Assume $2K < \min(R, B)$.)

(1) $$p_X(x) = \frac{\binom{R}{x}\binom{B}{K-x}}{\binom{B+R}{K}} \qquad 0 \le x \le K \wedge R$$

(2) $$p_{Y|X}(y|x) = \frac{\binom{R-x}{y}\binom{B-(K-x)}{K-y}}{\binom{B+R-K}{K}} \qquad 0 \le y \le R - x$$

(3) $$p_{XY}(x, y) = \frac{\binom{R}{x}\binom{B}{K-x}\binom{R-x}{y}\binom{B-(K-x)}{K-y}}{\binom{B+R}{K}\binom{B+R-K}{K}} \qquad 0 \le x, y \le K, \quad x + y \le R$$

(4) $$p_Y(y) = \sum_x p_{XY}(x, y) \qquad 0 \le y \le K \wedge R$$

A quick check

$$p_{XY}(x,y) = \frac{\binom{R}{x}\binom{B}{K-x}\binom{R-x}{y}\binom{B-(K-x)}{K-y}}{\binom{B+R}{K}\binom{B+R-K}{K}}$$

- Test case: $B = R = K$. By substitution

$$p_{XY}(x,y) = \frac{\binom{K}{x}\binom{K}{K-x}\binom{K-x}{y}\binom{K-(K-x)}{K-y}}{\binom{K+K}{K}\binom{K+K-K}{K}}$$

$$= \frac{\binom{K}{x}\binom{K}{y}}{\binom{2K}{K}} = \frac{\binom{K}{x}^2}{\binom{2K}{K}}, \text{ since } x+y=K.$$

- But if $B = R = K$, since $x + y = K$

$$p_{XY}(x,y) = \frac{\#\text{ of ways to have } x \text{ reds in 1st draw and } y \text{ in 2nd}}{\#\text{ of possible draws}}$$

$$= \frac{\binom{K}{x}\binom{K}{y}}{\binom{2K}{K}}$$