STAT 511 su 2020 Lec 11 slides

Joint and marginal distributions

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Time to consider more than just one rv:

- X on $\mathcal{S}$ with support $\mathcal{X}$
- Y on $\mathcal{S}$ with support $\mathcal{Y}$

How do X and Y behave *together* as a pair (X, Y)?

Let

$$\mathcal{X} \times \mathcal{Y} = \{(x, y) : x \in \mathcal{X}, y \in \mathcal{Y}\} \subset \mathbb{R} \times \mathbb{R}.$$

Probabilities about a pair of rvs

For any $A \in \mathcal{E}_{\mathcal{X}\times\mathcal{Y}}$ we write

$$P_{X,Y}((X, Y) \in A) = P(\{s \in \mathcal{S} : (X(s), Y(s)) \in A\}),$$

and call $P_{X,Y}(\cdot)$ the *joint probability distribution* of (X, Y).

Exercise: Let $X =$ sum of two dice rolls, $Y =$ max.

Tabulate $P((X, Y) = (x, y))$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

For reference:

$$\mathcal{S} = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\}$$

$\mathcal{X}$ \ $\mathcal{Y}$	1	2	3	4	5	6
2	1/36					
3		2/36				
4		1/36	2/36			
5			2/36	2/36		
6			1/36	2/36	2/36	
7				2/36	2/36	2/36
8				1/36	2/36	2/36
9					2/36	2/36
10					1/36	2/36
11						2/36
12						1/36

Joint pmf of two discrete random variables

For two discrete rvs X and Y, the *joint pmf* of the rv pair (X, Y) is the function

$$p(x, y) = P((X, Y) = (x, y)) \text{ for all } (x, y) \in \mathbb{R}^2$$

- Use to compute probabilities: For any set $A \in \mathcal{X} \times \mathcal{Y}$

$$P((X, Y) \in A) = \sum_{(x,y) \in A} p(x, y).$$

- Use to compute expected values: For any function $g : \mathbb{R}^2 \to \mathbb{R}$

$$\mathbb{E}g(X, Y) = \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} g(x, y)p(x, y).$$

- Note (i) $p(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$ and (ii) $\sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p(x, y) = 1$.

Exercise: Let X = sum of two dice rolls, Y = max. Find

1. $P(X = 2Y)$
2. $\mathbb{E}[X/Y]$

To focus on just X or just Y of (X, Y), find the *marginal distribution* of X or Y.

Theorem (obtaining marginal pmfs from the joint pmf)

Let (X, Y) be a pair of rvs with supports $\mathcal{X}$ and $\mathcal{Y}$ and joint pmf p. Then the

- *pmf of X is*
$$p_X(x) = \sum_{y \in \mathcal{Y}} p(x, y) \text{ for all } x \in \mathbb{R}$$
- *pmf of Y is*
$$p_Y(y) = \sum_{x \in \mathcal{X}} p(x, y) \text{ for all } y \in \mathbb{R}.$$

The pmfs p_X and p_Y are called the *marginal pmfs* of X and Y.

Exercise: Let X = sum of two dice rolls, Y = max.

Tabulate the marginal distributions of X and Y.

For X = sum of two dice rolls, Y = max, table gives joint and marginal distrib's.

		$\mathcal{Y}$						
		1	2	3	4	5	6	
	2	1/36						1/36
	3		2/36					2/36
	4		1/36	2/36				3/36
	5			2/36	2/36			4/36
$\mathcal{X}$	6			1/36	2/36	2/36		5/36
	7				2/36	2/36	2/36	6/36
	8				1/36	2/36	2/36	5/36
	9					2/36	2/36	4/36
	10					1/36	2/36	3/36
	11						2/36	2/36
	12						1/36	1/36
		1/36	3/36	5/36	7/36	9/36	11/36	

Joint pdf of two continuous random variables

For two continuous rvs X and Y with support $\mathcal{X}$ and $\mathcal{Y}$, the *joint pdf* of (X, Y) is the function $f : \mathbb{R}^2 \to [0, \infty)$ which satisfies

$$P((X, Y) \in A) = \iint_A f(x, y) dxdy$$

for any set $A \in \mathcal{E}_{\mathcal{X}\times\mathcal{Y}}$.

In the above, $\mathcal{E}_{\mathcal{X}\times\mathcal{Y}}$ is a collection of sets of interest in $\mathcal{X} \times \mathcal{Y}$.

Notation $\iint_A$ denotes integration over all $(x, y) \in A$.

So $P((X, Y) \in A)$ is "volume" under f over the region A.

Note (i) $f(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$ and (ii) $\int_{\mathbb{R}^2} f(x, y) = 1$.

Exercise: Let (U, V) be a pair of rvs with joint pdf given by

$$f(u, v) = 6(v - u)\mathbf{1}(0 < u < v < 1).$$

1. Show that $f(u, v)$ is a legitimate joint pdf.
2. Find $P(U + V \leq 1)$.
3. Find $\mathbb{E}[U/V]$.
4. Find $\mathbb{E}[(U + V)/2]$.

Theorem (obtaining marginal pdfs from the joint pdf)

Let X and Y be continuous rvs such that (X, Y) has joint pdf f. Then the

- *pdf of X is*

$$f_X(x) = \int_{\mathbb{R}} f(x, y)dy \text{ for all } x \in \mathbb{R}.$$

- *pdf of Y is*

$$f_Y(y) = \int_{\mathbb{R}} f(x, y)dx \text{ for all } y \in \mathbb{R}.$$

The pdfs f_X and f_Y are called the *marginal pdfs* of X and Y.

Exercise: Let (X, Y) have joint pdf given by

$$f(x, y) = \frac{6}{5}[1 - (x - y)^2] \cdot \mathbf{1}(0 < x < 1, 0 < y < 1).$$

Find the marginal pdfs of X and Y.

Exercise: Let (X, Y) have joint pdf given by

$$f(x, y) = y(1 - x)^y e^{-y} \cdot \mathbf{1}(0 < x < 1, 0 < y < \infty).$$

Find the marginal pdfs of X and Y.

Joint cumulative distribution function

For a rv pair (X, Y), the *joint cdf* F of (X, Y) is defined as

$$F(x, y) = P(X \leq x, Y \leq y) \text{ for all } (x, y) \in \mathbb{R}^2$$

- For (X, Y) a pair of discrete rvs with support $\mathcal{X}$ and $\mathcal{Y}$ and joint pmf p,

$$F(x, y) = \sum_{\{t_1 \in \mathcal{X} : t_1 \leq x\}} \sum_{\{t_2 \in \mathcal{Y} : t_2 \leq y\}} p(t_1, t_2) \text{ for all } (x, y) \in \mathbb{R}^2.$$

- For (X, Y) a pair of continuous rvs with support $\mathcal{X}$ and $\mathcal{Y}$ and joint pdf f,

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t_1, t_2) dt_1 dt_2 \text{ for all } (x, y) \in \mathbb{R}^2.$$

Exercise: Let $(U, V) \sim f(u, v) = 6(v - u)\mathbf{1}(0 < u < v < 1)$.

Find the joint cdf $F(u, v)$ of for all $(u, v) \in \mathbb{R}^2$.