STAT 712 fa 2022 Lec 8 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:

- $X : \Omega \to \mathcal{X}$
- $Y : \Omega \to \mathcal{Y}$

Consider behavior of X and Y *together* as a pair $(X, Y) : \Omega \to \mathbb{R}^2$.

Probabilities about a pair of rvs

For a probability space $(\Omega, \mathcal{B}, P)$ and a pair of rvs $(X, Y) : \Omega \to \mathbb{R}^2$, we write

$$P_{X,Y}((X, Y) \in A) = P(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in A\}) \quad \text{for any } A \in \mathcal{B}(\mathbb{R}^2),$$

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

We tend to omit X, Y from the subscript of $P_{X,Y}(\cdot)$.

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

Table of $P((X, Y) = (x, y))$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ is

		$\mathcal{Y}$					
		1	2	3	4	5	6
$\mathcal{X}$	1	1/36	2/36	2/36	2/36	2/36	2/36
	2		1/36	2/36	2/36	2/36	2/36
	3			1/36	2/36	2/36	2/36
	4				1/36	2/36	2/36
	5					1/36	2/36
	6						1/36

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

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 $A \in \mathcal{X} \times \mathcal{Y}$

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

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

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

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

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

Theorem (Get marginal pmfs from the joint)

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

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

Exercise: Let $X =$ min of two dice rolls, $Y =$ max. Get marginal pmfs.

Joint pdf of two continuous random variables

The *joint pdf* of a pair of continuous rvs (X, Y) is fun. $f : \mathbb{R}^2 \to [0, \infty)$ satisfying

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

for any set $A \in \mathcal{B}(\mathbb{R}^2)$.

The 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.

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

$$\mathbb{E}g(X, Y) = \iint_{\mathbb{R}^2} g(x, y)f(x, y)dxdy.$$

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

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 given by* $f_X(x) = \int_{\mathbb{R}} f(x,y)dy$ *for all* $x \in \mathbb{R}$.
- *The pdf of Y is given by* $f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$ *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

$$f(x, y) = \frac{1}{x^3} e^{-1/x} e^{-y/x} \cdot \mathbf{1}(x > 0, y > 0).$$

1. Find $P(X > Y)$.
2. 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 joint pdf f,

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

- We also have

$$\frac{\partial^2}{\partial x \partial y} F(x, y) = f(x, y), \text{ for } x, y \text{ at which } f(x, y) \text{ is continuous.}$$

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

Joint moment generating function

The *joint mgf* $M_{X_1,X_2}(t_1, t_2)$ of a pair of rvs (X_1, X_2) is defined as

$$M_{X_1,X_2}(t_1, t_2) = \mathbb{E}e^{t_1 X_1 + t_2 X_2}$$

provided the expectation is finite for all t_1, t_2 in some open neighborhoods of 0.

- We can get the marginal mgfs of X_1 and X_2 as

$$M_{X_1}(t_1) = M_{X_1,X_2}(t_1, 0)$$
$$M_{X_2}(t_2) = M_{X_1,X_2}(0, t_2).$$

- Also $\mathbb{E}X_k = \frac{\partial}{\partial t_k} M_{X_1,X_2}(t_1, t_2)\Big|_{t_1=t_2=0}$ for $k = 1, 2$.

Exercise: Let $(X_1, X_2) \sim f(x_1, x_2) = e^{-x_1} \cdot \mathbf{1}(0 < x_2 < x_1 < \infty)$.

1. Find the joint mgf of (X_1, X_2).
2. Give the marginal distributions of X_1 and X_2.