Math-UA.233: Theory of Probability
Lecture 16

Tim Austin

Joint distributions of RVs (Ross Chap 6)

At this point, we have learned a lot about RVs, and met several important examples.

But we have always studied one RV at a time, or occasionally studied one RV which is a function of another.

Now we will study *several random variables at once*.

We will mostly focus on pairs of RVs for simplicity, but larger collections behave similarly.

Suppose we have an experiment with sample space S and probability values given by P.

If X and Y are two RVs, then there are many events we can express in terms of them, such as

- $\{X \leqslant 4, Y \leqslant 7\}$
- $\{X^2 + Y^2 \leqslant 1\}$
- $\{X < Y\}$.

Worth emphasizing: X and Y are RVs for the *same experiment*: that is, they are two functions on the *same sample space* S. But they may describe different features of the outcome.

It's nice to think of the pair (X, Y) as a *random vector*. Mathematically, it is a function

$$S \longrightarrow \mathbb{R}^2 \ : \ s \mapsto (X(s), Y(s)).$$

Now the general form of an event defined in terms of X and Y is

$$\{(X, Y) \in A\},$$

where A is some region (that is, subset) of two-dimensional space $\mathbb{R}^2$.

Previous examples:

- $\{X \leqslant a, Y \leqslant b\}$ is the event that (X, Y) lands in the quadrant $(-\infty, a] \times (-\infty, b]$
- $\{X^2 + Y^2 \leqslant 1\}$ is the event that (X, Y) lands inside or on a unit circle around the origin
- $\{X < Y\}$ is the event that (X, Y) lands in the set

  $$A = \{(x, y) : \ x < y\},$$

  which is the open half-plane above the main diagonal in $\mathbb{R}^2$.

In general, we can describe the probabilities of such events using the following.

Definition

The ***joint cumulative distribution function*** *(‘**joint CDF**’) of X and Y is the function*

$$F(a, b) = P(X \leqslant a,\ Y \leqslant b) = P\big((X, Y) \in (-\infty, a] \times (-\infty, b]\big)$$

for $a, b \in \mathbb{R}$.

Thus, the joint CDF is a function $\mathbb{R}^2 \longrightarrow [0, 1]$. It generalizes the CDF of a single RV.

Now the probabilities of other events can be found in terms of this function. Most basic example:

Proposition (Ross eqn. (5.1.2))

For any real values $a_1 < a_2$ *and* $b_1 < b_2$, *we have*

$$\begin{aligned} P(a_1 < X \leqslant a_2,\ b_1 < Y \leqslant b_2) \\ = F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1). \end{aligned}$$

IDEA: Careful use of axiom 3, or re-arrangement of inclusion-exclusion formula.

Another important consequence of:

$$F(a, b) = P(X \leqslant a,\ Y \leqslant b).$$

Sending $a \longrightarrow \infty$, we get

$$\boxed{\lim_{a \longrightarrow \infty} F(a, b) = P(X < \infty,\ Y \leqslant b) = P(Y \leqslant b) = F_Y(b),}$$

(intuitive, but I'm ignoring the issue of why the limit exists, etc).

Similarly,

$$\boxed{\lim_{b \longrightarrow \infty} F(a, b) = F_X(a).}$$

So the joint CDF determines the two individual CDFs.

But for more general regions, using the joint CDF gets very tricky. For example, to find

$$P(X^2 + Y^2 \leqslant 1)$$

in terms of F, we have to approximate the unit circle with a union of many small boxes, and take a limit as the boxes get smaller and smaller. Other regions can be even more complicated.

For this reason we mostly focus on two special classes of RV which are more manageable: discrete and continuous, just like for single RVs.

Multiple discrete RVs

Definition

*If X and Y are two discrete RVs, then their **joint PMF** is the function*

$$p(a,b) = P(X = a,\ Y = b),$$

defined for all real numbers a and b.

Of course, $p(a,b)$ is zero unless a is one of the possible values of X and b is one of the possible values of Y.

Many examples are already quite familiar.

Example

Two dice are rolled, and X and Y denote the minimum and maximum of the values shown, respectively. Find their joint PMF.

GOOD PRESENTATION: Draw a table!

OBSERVATION FROM THIS EXAMPLE: If X has possible values $x_1, x_2, \ldots$ and Y has possible values $y_1, y_2, \ldots$, there may be some pairs (x_i, y_j) which are *not* possible for (X, Y).

The basic facts about single PMFs have analogs here:

- $p(a, b) \geqslant 0$ for all a and b
- If X has possible values $x_1, x_2, \ldots$ and Y has possible values $y_1, y_2, \ldots$, then

$$\sum_{\text{all } i \text{ and } j} p(x_i, y_j) = 1$$

- For any region $A \subseteq \mathbb{R}^2$,

$$P((X, Y) \in A) = \sum_{i,j \text{ such that } (x_i, y_j) \in A} p(x_i, y_j).$$

In particular, the joint CDF is

$$F(x, y) = \sum_{i \text{ such that } x_i \leqslant x \text{ and } j \text{ such that } y_j \leqslant y} p(x_i, y_j).$$

In particular:

Proposition (Ross p221)

If X and Y are discrete with joint PMF p, then the PMF of X by itself is

$$\boxed{p_X(x_i) = P(X = x_i, Y = \mathit{anything}) = \sum_j p(x_i, y_j),}$$

and similarly

$$\boxed{p_Y(y_j) = \sum_i p(x_i, y_j).}$$

So p completely determines the two single-RV PMFs p_X and p_Y. They are sometimes called the **marginals** of p.

Example (Ross E.g. 6.1a)

An urn contains 3 red, 4 white and 5 blue balls. 3 are randomly selected. Let X and Y be the numbers of red and white balls we select. Find their joint and individual PMFs.

But p_X and p_Y don't completely determine p.

Example

Flip two independent fair coins. Let X and Y be the indicator variables that the first and second coins give H, respectively.

After tossing the first coin, you shout the outcome to your friend in the next room. She can't hear you clearly, and mis-hears you with probability p. Let Z be the indicator variable that she thinks you said H.

Find the individual PMFs of X, Y and Z and the joint PMFs of (X, Y) and (X, Z).

It's easy to generalize the previous ideas to more than two discrete RVs. Here's one of the most important examples.

Example (Ross E.g. 6.1f)

Suppose that n independent experiments are performed, all of which have r possible outcomes with individual probabilities p_1, $\dots$, p_r, $\sum_{i=1}^{r} p_i = 1$. Let X_i be the number of the experiments which result in outcome i. Then the discrete RVs $X_1, \dots, X_r$ all have possible values $0, 1, \dots, n$, and they have joint PMF

$$P(X_1 = n_1, \dots, X_r = n_r) = \begin{cases} \binom{n}{n_1, \dots, n_r} p_1^{n_1} \cdots p_r^{n_r} & \text{if } \sum_i n_i = n \\ 0 & \text{otherwise.} \end{cases}$$

That last joint PMF again:

$$P(X_1 = n_1, \ldots, X_r = n_r) = \begin{cases} \binom{n}{n_1,\ldots,n_r} p_1^{n_1} \cdots p_r^{n_r} & \text{if } \sum_i n_i = n \\ 0 & \text{otherwise.} \end{cases}$$

Marginals: each X_i is binom(n, p_i). Moreover, $X_1 + X_2$ is binom$(n, p_1 + p_2)$, $X_2 + X_3 + X_5$ is binom$(n, p_2 + p_3 + p_5)$, etc.

A tuple of discrete RVs $(X_1, \ldots, X_r)$ with this joint PMF are called **multinomial**$(p_1, \ldots, p_r)$. They are a generalization of binomial RVs. They arise very often in statistical sampling problems.

Jointly continuous RVs

The two-dimensional analog of continuous RVs works like this:

Definition

*The pair of RVs (X, Y) is **jointly continuous** if there is an integrable function $f(x, y)$ of two real variables such that, for any 'nice' region A of the plane, we have*

$$P((X, Y) \in A) = \iint_A f(x, y)\, dx\, dy.$$

*The function f is the **joint PDF** of (X, Y).*

(What's a 'nice' region? One for which the above integral is defined, as studied in multi-variable calculus. For instance, any region bounded by a smooth curve is 'nice'. We won't worry about other regions in this course.)

Once again, there are many analogs to the case of a single RV.

First, if (X, Y) are jointly continuous, then their joint CDF is

$$\boxed{F(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)\, dx\, dy.}$$

Therefore

$$\boxed{\frac{\partial^2 F(a,b)}{\partial a \partial b} = f(a,b),}$$

by applying the Fundamental Theorem of Calculus once in each variable.

So the joint CDF and joint PDF determine each other by differentiation/integration.

INTUITIVE INTERPRETATION:

$$P(x < X \leqslant x + dx,\ y < Y \leqslant y + dy) \approx f(x, y)\, dx\, dy$$

So $f \geqslant 0$, but it can be bigger than 1: it's a *density*, not an actual *probability*.

RULE OF THUMB BETWEEN DISCRETE AND CONTINUOUS:

$$\text{(discrete)} \quad p(x, y) \leftrightarrow f(x, y)\, dx\, dy \quad \text{(jointly continuous)}.$$

Also, if A is a two-dimensional region whose *area equals zero*, such as a line segment, or one-dimensional curve, then

$$\boxed{P((X, Y) \in A) = 0}$$

because any two-dimensional integral over A must be zero.

MATHEMATICAL SUBTLETY:

There can be continuous RVs X and Y which are not *jointly* continuous.

Example: let X be Unif$(0, 1)$, and let $Y = X^2$. Then Y is also continuous — we've computed its PDF previously.

But (X, Y) always lies on the one-dimensional curve

$$C \ : \ (t, t^2) \quad \text{for } 0 \leqslant t \leqslant 1.$$

The two-dimensional *area* of C is zero. So X, Y cannot have a *joint* PDF.

So we often have to explicitly assume that our two RVs are jointly continuous.

Using the rule of thumb, we obtain continuous analogs of results for joint PMFs. First example:

Proposition (Ross p224)

If X, Y are jointly continuous with joint PDF f, then X itself is continuous, and its PDF is

$$\boxed{f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy,}$$

*and similarly for f_Y. Call f_X and f_Y the **marginal PDFs** of f.*

INTUITIVE REASON:

$$\begin{aligned} f(x)\, dx &\approx P(X \in [x, x + dx]) = P(X \in [x, x + dx], Y \in \mathbb{R}) \\ &= \int_{-\infty}^{\infty} \left[\int_{x}^{x+dx} f(x', y)\, dx' \right] dy \approx \left[\int_{-\infty}^{\infty} f(x, y)\, dy \right] dx \end{aligned}$$

Example (Ross E.g. 6.1c)

The joint PDF of X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{if } x, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find (a) $P(X > 1, Y < 1)$, (b) $P(X < Y)$.

Example (Pishro-Nik E.g. 5.18)

Let X and Y be jointly continuous with joint PDF

$$f(x,y) = \begin{cases} cx^2y & \text{if } 0 \leqslant y \leqslant x \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) *Sketch the region where $f(x,y) > 0$.*

(b) *Find the constant c.*

(c) *Find the marginal PDFs f_X and f_Y.*

(d) *Find $P(Y \leqslant \frac{1}{2}X)$.*

(e) *Find $P(Y \leqslant \frac{1}{4}X \mid Y \leqslant \frac{1}{2}X)$.*