Math-UA.233: Theory of Probability
Lecture 13

Tim Austin

tim@cims.nyu.edu
cims.nyu.edu/∼tim

Recap of some previous facts about RVs

If X is a RV for some experiment, then there are events that can be described in terms of X. They have the form

$$\{X \in A\} = \{X \text{ takes a value in } A\},$$

where A is a chosen set of real numbers.

If X is discrete with possible values $x_1, x_2, \dots,$ then we can find the probability using the PMF of X:

$$\boxed{P(X \in A) = \sum_{i \text{ such that } x_i \in A} p_X(x_i).}$$

For more general RVs we can't use the PMF this way.

For the general case, the only tool we've met so far is the CDF

$$F_X(a) = P(X \leqslant a), \quad a \in \mathbb{R}.$$

We can deduce many other probabilities from the CDF, by describing more general events in terms of the events $\{X \leqslant a\}$.

Example

$$P(2.5 < X \leqslant 7.5) = F(7.5) - F(2.5)$$

and

$$P(1 < X \leqslant 2 \ \ \textit{or} \ \ 3 < X \leqslant 4) = [F(4) - F(3)] + [F(2) - F(1)].$$

Example (Adapted from Ross E.g. 5.7b)

Let X be a RV whose CDF is

$$F(x) = \begin{cases} 0 & \text{if } x \leqslant -1 \\ (1+x)/2 & \text{if } -1 < x < 1 \\ 1 & \text{if } x \geqslant 1. \end{cases}$$

Find $P(X^2 > 1/4)$.

MATHEMATICAL REMARK

Can we really obtain the probability

$$P(X \in A)$$

for *any* subset $A \subseteq \mathbb{R}$ using only the CDF together with the axioms of probability?

Actually, NO, not always.

For some RVs and some crazy choices of A, it is *theoretically impossible* to even make sense of these probability values.

MATHEMATICAL REMARK, contd.

The fix: restrict to certain 'nice' sets A, called *measurable sets*, for which everything works OK.

Mathematically, this means that the function

$$P'(A) = P(X \in A)$$

is *not* a set function defined on the whole collection

$$\{\text{subsets of } \mathbb{R}\},$$

but only on the subcollection of measurable sets.

Defining and studying such sets and set functions is a deep and difficult subject called *measure theory* — no more about that here.

Happily, all the sets we ever meet in practice are measurable.

Continuous random variables (Ross, Chapter 5)

Sometimes we want to model a ‘random quantity’ which can take an uncountable (that is, ‘un-listable’) set of possible values.

Examples:

- Time that the next train arrives at a station,
- The lifetime of a transistor.

For both of these, it would be natural to use an RV that can take any value in $[0, \infty)$.

So discrete RVs aren’t enough to handle these examples. We now introduce another special class, this time based on the use of calculus.

Definition

*A RV X is **continuous** if there is an integrable function f on the real line such that*

$$\boxed{P(a \leqslant X \leqslant b) = \int_a^b f(x)\,dx}$$

for any real numbers $a < b$, or with $a = -\infty$ or $b = +\infty$.

*Then f is the **probability density function** or '**PDF**' of X.*

(This assumption may not tell us about $P(X \in A)$ for arbitrary sets A, as discussed above. But we're not going to worry about that.)

In particular, the CDF of X is given by

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

That is, X is continuous if its CDF is the definite integral of another function, the PDF. Note: this definite integral is a continuous function of a, and this is why they're called 'continuous' RVs (roughly; conventions vary a bit).

Therefore, by the Fund. Thm. of Calculus (see Ross p179):

$$\boxed{F'(a) = \frac{d}{da}\int_{-\infty}^{a} f(x)\,dx = f(a)}$$

(if f is continuous at a).

So the CDF and the PDF are related by differentiation/integration. Each determines the other.

INTUITION: consider a very short interval $[x, x + dx]$ near the point x. Then the probability that X lands in this interval is

$$\boxed{P(x \leqslant X \leqslant x + dx) = \int_x^{x+dx} f(y)\, dy \approx f(x)dx}$$

(assuming f is continuous at x).

So $f(x)$ is *NOT* a probability. It is the 'density' of the probability $P(x \leqslant X \leqslant x + dx)$ *relative to* the length of the interval dx.

CONSEQUENCE:

$$\boxed{f(x) \geqslant 0 \quad \text{for all } x},$$

because $P(x \leqslant X \leqslant x + dx) \geqslant 0$ and $dx > 0$.

BUT, $f(x)$ can be bigger than 1, because $P(x \leqslant X \leqslant x + dx)$ can be large *relative to dx*.

Two more basic consequences of the definition (Ross pp176-7):

We must have

$$\int_{-\infty}^{\infty} f(x)\,dx = P(-\infty < X < \infty) = P(S) = 1.$$

If f doesn't satisfy this, then it cannot be the PDF of a RV.

For any fixed $a \in \mathbb{R}$ we have

$$P(X = a) = P(a \leqslant X \leqslant a) = \int_a^a f(x)\,dx = 0.$$

CONSEQUENCE: a RV *cannot* be both continuous and discrete!

Example (Ross E.g. 5.1a)

Let X be a continuous RV whose PDF is

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{for } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

for some real number C.

(a) *What is the value of C?*

(b) *Find $P(X > 1)$.*

Example (Ross E.g. 5.1b)

The lifetime (in hours) of a device is a continuous RV X whose PDF is

$$f(x) = \begin{cases} \lambda e^{-x/100} & \text{if } x \geqslant 0 \\ 0 & \text{if } x < 0 \end{cases}$$

for some real parameter λ. What is the probability that

(a) *it will last between 50 and 150 hours?*

(b) *it will function for less than 100 hours?*

THE RULE OF THUMB

The theory of continuous RVs is a lot like the theory of discrete RVs if you keep in mind the analogy

$$\text{(for discrete RVs)} \qquad p(x) \leftrightarrow f(x)dx \qquad \text{(for continuous RVs)}$$

(where p is a P<u>M</u>F and f is a P<u>D</u>F).

It's important to keep track of that 'dx' on the right-hand side: f is a *density*, not a *probability*.

As a result, expressions involving *sums* for discrete RVs turn into *integrals* for continuous RVs. First example:

$$\underbrace{\sum_x p(x)}_{\text{discrete}} = 1 = \underbrace{\int_{-\infty}^{\infty} f(x)\, dx}_{\text{continuous}}$$

Uniform RVs (Ross Sec 5.3)

These are the simplest continuous RVs, but still very important.

Definition

*Given real numbers $c < d$, a RV X is **uniform over** (c, d) (or just '**Unif**(c, d)') if it is continuous and*

$$f(x) = \begin{cases} \frac{1}{d-c} & \text{if } c \leqslant x \leqslant d \\ 0 & \text{otherwise} \end{cases}$$

NOTE: it's enough to say that f has to be constant inside (c, d) and zero outside. Then the constant is forced to equal $1/(d-c)$ by the condition $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

Basic consequences (Ross p184):

If $c \leqslant a < b \leqslant d$ then

$$P(a \leqslant X \leqslant b) = \frac{b-a}{d-c}.$$

The CDF is given by

$$F(a) = \int_{-\infty}^{a} f(x)\,dx = \begin{cases} 0 & \text{if } a \leqslant c \\ \frac{a-c}{d-c} & \text{if } c < a < d \\ 1 & \text{if } a \geqslant d. \end{cases}$$

Example (Ross E.g. 5.3b)

Let X be Unif$(0, 10)$, and find (a) $P(X < 3)$ and (b) $P(3 < X < 8)$.

Example (Ross E.g. 5.3c)

Buses arrive at a stop every 15 minutes starting at 7am. A passenger arrives at a random time which is uniformly distributed between 7 and 7:30am. Find the probability that she waits

(a) *less than 5 minutes for a bus,*

(b) *more than 10 minutes for a bus.*

Expectation and Variance of Continuous RVs (Ross Sec 5.2)

Recall that if Y is a discrete RV with PMF p, then we defined

$$E[Y] = \sum_y yp(y).$$

We make the analogous definition for continuous RVs.

Remember the rule of thumb: replace $p(y)$ with $f(x)dx$, and replace the *sum* with an *integral*.

Definition (Ross p178)

*If X is a continuous RV with PDF f, then its **expectation** (or **expected value** or **mean**) is the number*

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$$

(provided the improper integral $\int_{-\infty}^{\infty} |x| f(x)\, dx$ is well-defined and finite).

Example (Ross E.g. 5.2a)

Find $E[X]$ when the PDF of X is

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

The theory of expectation for continuous RVs is analogous to the theory for discrete RVs.

Recall that if X is any RV, and $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a function, then we can form the new RV $g(X)$.

Unfortunately, if X is continuous, it can happen that $g(X)$ is *not* continuous. Even if it is, it can be tricky to find the PDF of $g(X)$ in terms of that of X. But we can always use the following.

Theorem (LOTUS for continuous RVs; Ross Prop 2.1)

If X is continuous with PDF f, and $g : \mathbb{R} \longrightarrow \mathbb{R}$, then

$$\boxed{E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx}$$

(provided the integral is well-defined).

See Ross p181 for a proof (with a lot of cheating).

Example (Ross E.g. 5.2b)

Let X be Unif $(0, 1)$. Find $E[e^X]$.

You can do this either by first finding the PDF of e^X (more on that next time), or by using LOTUS (much quicker).

Example (Ross E.g. 5.2c)

A stick of length 1 is split into two pieces at a random point which is uniformly distributed along its length. A red dot is marked on the stick at distance p from one end. Find the expected length of the piece which has the red dot.