STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-09-21 14:47

real valued functions with arguments that live inside sample spaces

the main focus of this course

We'll use "probability measure" throughout the course, but our main focus will be a different and equally strange object.

Recall that sample space is often arbitrary and difficult or impossible to describe.

It turns out usually we're ultimately interested in a *real number* that is associated with the random outcome, rather than the random outcome itself.

Consider a coin tossing game with $S = \{H, T\}$, which might be repeated, from which a multitude of examples can be invented $\ldots$.

Consider also the notion of picking a real number uniformly from $(0, 1)$ $\ldots$.

Eventually we will not even bother with an underlying S explicitly.

random variable

A *random variable* is a function whose domain is a sample space and whose range is $\mathbb{R}$.

Naming convention: Roman letters near the end of the alphabet X, Y, $X_1, X_2, \ldots$.

Another strange convention - almost always omit the function's "argument".

We will never draw a picture of a random variable, or compute a derivative or an integral of one.

We will instead focus on *the* defining property of a random variable: its *distribution*.

Perversely, we will lack the math to actually define *distribution* rigorously. Informally, the *distribution* of a random variable X is the rule that assigns probabilities to values of X.

assigning probabilities to values of X

As rigorous as we can get is mainly as follows.

A *distribution* is the rule that assigns probabilities that X takes on values in all intervals (closed, open, infinite, whatever), and simple set operations on intervals, e.g.

$$P(X \leq 1) \qquad P(\{X \leq 1\} \cap \{X \geq 1\}) = P(X = 1) \qquad P(X > -15)$$

The actual numbers above aren't important (1 and -15) and often generic statements are made using dummy placeholders like:

$$P(X \leq a) \qquad P(X = u) \qquad P(X > c)$$

With by far the most common being just "little x" as in $P(X = x)$ and $P(X \leq x)$.

complete descriptions of distributions, and other properties

If you know the *distribution* of a random variable, you know *everything* about it.

Most of the rest of this course will be occupied with:

- different ways (all equivalent) to completely and uniquely describe distributions. These ways will *always* be functions (in this course from $\mathbb{R}$ to $\mathbb{R}$)
- examples of random variables so important in practice that their distributions have special names.
- examples of otherwise useless random variables useful as exercises
- other not necessarily unique properties of distributions.

your first complete distribution descriptor

Suppose you have any random variable X. Its distribution can be completely described by the following function:

$$F_X(x) = P(X \leq x)$$

This is called the *cumulative distribution function* for X or *cdf*.

The subscript X is usually omitted unless required for clarity.

The domain is all of $\mathbb{R}$.

Note: that the cdf characterizes a distribution is actually a *theorem* which we lack the tools to prove.

defining properties of *all* cdf no matter what

Theorem: For any r.v. X, its cdf $F(x)$ has the following properties:

$$\lim_{x \to -\infty} F(x) = 0,$$

$$\lim_{x \to \infty} F(x) = 1,$$

and $F(x)$ is *right-continuous*, i.e.

$$\lim_{x \to a+} F(x) = F(a).$$

The proof of this theorem uses The Continuity Theorem and its corollary, and is left as an exercise.

(advanced note: any function with these properties is a cdf for some X)

discrete random variables

a large class of random variables

Discrete random variables take on a finite or countably ("list-able") set of real outcomes.

e.g. the coin toss game, and tossing a coin until the first head appears.

A more convenient complete distribution descriptor is the collection of probabilities of the set of outcomes, called the *probability mass function* or pmf:

$$p(x) = P(X = x)$$

This function is non-zero on the values of X, and formally 0 otherwise (usually just a formality).

pmf and cdf are "equivalent"

Theorem: for any discrete random variable X, the pmf and the cdf can be derived from each other.

Proof: next class

some important discrete random variables with special named distributions

the Bernoulli(p) distributions - fundamental building blocks

If a random variable takes on values 1 and 0 with probabilities p and $1 - p$ (for some fixed $0 < p < 1$), it is said to have a *Bernoulli distribution with parameter p, or Bernoulli(p).

It doesn't really matter what the underlying sample space S actually is:

1. toss a die; $S = \{1, 2, 3, 4, 5, 6\}$; define $X_1(1) = X_1(2) = X_1(3) = 0$ and $X_1(4) = X_1(5) = X_1(6) = 1$
2. flip a coin; $S = \{H, T\}$; define $X_2(H) = 0$ and $X_2(T) = 1$

X_1 and X_2 have the same distribution, Bernoulli$\left(\frac{1}{2}\right)$.

Bernoulli(p) pmf and cdf

$$p(x) = \begin{cases} 1 - p & : \quad x = 0, \\ p & : \quad x = 1 \end{cases} \quad = \quad p^x(1-p)^x \text{ for } x \in \{0, 1\}$$

$$F(x) = P(X \leq x) = \begin{cases} 0 & : \quad x < 0, \\ p & : \quad 0 \leq x < 1 \\ 1 & : \quad x \geq 1 \end{cases}$$