STA286 Lecture 04

Neil Montgomery

Last edited: 2017-01-18 09:57

real-valued functions with arguments that live inside sample spaces that eventually we will pretty much forget even exist

sample spaces: too general?

The goal for *our* study of probability is to define a mathematical model for a variable in a dataset.

Sample spaces are needed to define the important building block P (probability function), but are not usually convenient.

sample spaces: too general?

The goal for *our* study of probability is to define a mathematical model for a variable in a dataset.

Sample spaces are needed to define the important building block P (probability function), but are not usually convenient.

e.g. six horses in a race, labelled 'A' through 'F'. You bet \$2 on 'D' to win. You will get \$6.75 if 'D' wins.

sample spaces: too general?

The goal for *our* study of probability is to define a mathematical model for a variable in a dataset.

Sample spaces are needed to define the important building block P (probability function), but are not usually convenient.

e.g. six horses in a race, labelled 'A' through 'F'. You bet \$2 on 'D' to win. You will get \$6.75 if 'D' wins.

What is the sample space? Some options:

- all 6! possible orders of finish: $\{DFCEAB, ECAFDB, FEABCD, EFBDAC \ldots\}$.

sample spaces: too general?

The goal for *our* study of probability is to define a mathematical model for a variable in a dataset.

Sample spaces are needed to define the important building block P (probability function), but are not usually convenient.

e.g. six horses in a race, labelled 'A' through 'F'. You bet \$2 on 'D' to win. You will get \$6.75 if 'D' wins.

What is the sample space? Some options:

- all 6! possible orders of finish: $\{DFCEAB, ECAFDB, FEABCD, EFBDAC \ldots\}$.
- all possible combinations of the six horse distances run as functions of time.

sample spaces: too general?

The goal for *our* study of probability is to define a mathematical model for a variable in a dataset.

Sample spaces are needed to define the important building block P (probability function), but are not usually convenient.

e.g. six horses in a race, labelled ‘A’ through ‘F’. You bet \$2 on ‘D’ to win. You will get \$6.75 if ‘D’ wins.

What is the sample space? Some options:

- all 6! possible orders of finish: $\{DFCEAB, ECAFDB, FEABCD, EFBDAC \ldots\}$.
- all possible combinations of the six horse distances run as functions of time.

sample spaces: too general?

The goal for *our* study of probability is to define a mathematical model for a variable in a dataset.

Sample spaces are needed to define the important building block P (probability function), but are not usually convenient.

e.g. six horses in a race, labelled 'A' through 'F'. You bet \$2 on 'D' to win. You will get \$6.75 if 'D' wins.

What is the sample space? Some options:

- all 6! possible orders of finish: $\{DFCEAB, ECAFDB, FEABCD, EFBDAC \ldots\}$.
- all possible combinations of the six horse distances run as functions of time.

But really all you care about is whether you suffer a \$2 loss or a \$4.75 profit, and it turns out to be better just to focus on that.

a new style of function to add to the stable

We will now consider functions that have a sample space S as a domain and range $\mathbb{R}$, e.g.

Element of S	Value in $\mathbb{R}$
DFCEAB	4.75
ECAFDB	-2.00
FEABCD	-2.00
EFBDAC	-2.00
$\vdots$	$\vdots$

a new style of function to add to the stable

We will now consider functions that have a sample space S as a domain and range $\mathbb{R}$, e.g.

Element of S	Value in $\mathbb{R}$
DFCEAB	4.75
ECAFDB	-2.00
FEABCD	-2.00
EFBDAC	-2.00
$\vdots$	$\vdots$

Note the impossibility of doing things like drawing pictures of such functions.

So if all the usual function things like derivatives, local maxima, integrals etc. are not of interest, what do we do with these things?

a new style of function to add to the stable

We will now consider functions that have a sample space S as a domain and range $\mathbb{R}$, e.g.

Element of S	Value in $\mathbb{R}$
DFCEAB	4.75
ECAFDB	-2.00
FEABCD	-2.00
EFBDAC	-2.00
$\vdots$	$\vdots$

Note the impossibility of doing things like drawing pictures of such functions.

So if all the usual function things like derivatives, local maxima, integrals etc. are not of interest, what do we do with these things?

All we care about are the *distributions* of such functions—roughly speaking the possible values in $\mathbb{R}$ and their probabilities.

“random variables”

Sadly these functions have a terrible name.

Definition: A *random variable* is a real-valued function of a sample space.

More examples:

Toss to first head: Count the number of tosses of a coin until the first `H` appears.

Element of S	Value in $\mathbb{R}$
H	1
TH	2
TTH	3
TTTH	4
TTTTH	5
$\vdots$	$\vdots$

examples of random variables

See if a product is defective: Select an item at random from a factory. See if it is defective. Define the following function:

Element of S	Value in $\mathbb{R}$
Defective	1
Not Defective	0

examples of random variables

See if a product is defective: Select an item at random from a factory. See if it is defective. Define the following function:

Element of S	Value in $\mathbb{R}$
Defective	1
Not Defective	0

Seems like a stupid example, but this is merely an instance of one of the most important random variables of all. To generalize:

"Bernoulli trial" (book: *Bernoulli random variable*) Observe a random process to see if an event A occurred. The *indicator function* I_A which takes on the value 1 if A occurred and 0 otherwise is called a Bernoulli trial.

more examples of random variables

Bus stop Busses arrive at a stop at 10 minute intervals. You arrive at the bus stop at a random time. Observe the amount of time you have to wait for the next bus.

more examples of random variables

Bus stop Busses arrive at a stop at 10 minute intervals. You arrive at the bus stop at a random time. Observe the amount of time you have to wait for the next bus.

This random variable is of a fundamentally different character from the other examples. This random variable could take on *any* real number between 0 and 10 (using minutes as the unit of measure.)

more examples of random variables

Bus stop Busses arrive at a stop at 10 minute intervals. You arrive at the bus stop at a random time. Observe the amount of time you have to wait for the next bus.

This random variable is of a fundamentally different character from the other examples. This random variable could take on *any* real number between 0 and 10 (using minutes as the unit of measure.)

Failure time(s) Observe a pump until its bearing fails. Observe a pump until its seal fails. Observe a pump until either component fails.

These random variables could take on any positive real number.

notation and naming conventions

Random variables are given names that tend to be capital Roman letters near the end of the alphabet.

$$X : S \to \mathbb{R}$$

notation and naming conventions

Random variables are given names that tend to be capital Roman letters near the end of the alphabet.

$$X : S \to \mathbb{R}$$

Usual names: X, Y, Z, X_1, X_2, X_3, etc.

(Just like in calculus: f, g, h, f_1, f_2, etc.)

notation and naming conventions

Random variables are given names that tend to be capital Roman letters near the end of the alphabet.

$$X : S \to \mathbb{R}$$

Usual names: X, Y, Z, X_1, X_2, X_3, etc.

(Just like in calculus: f, g, h, f_1, f_2, etc.)

In calculus one tends to conflate f (the function name) and $f(x)$ (the value at a generic x)

In probability we just use the function name, i.e. X.

notation and naming conventions

Random variables are given names that tend to be capital Roman letters near the end of the alphabet.

$$X : S \to \mathbb{R}$$

Usual names: X, Y, Z, X_1, X_2, X_3, etc.

(Just like in calculus: f, g, h, f_1, f_2, etc.)

In calculus one tends to conflate f (the function name) and $f(x)$ (the value at a generic x)

In probability we just use the function name, i.e. X.

Is notation important? If you like to watch the world burn put this on a calculus exam:

$$\int_0^{2\pi} sin(f)\, df$$

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	$\{TH, TTH, TTTH\}$

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	$\{TH, TTH, TTTH\}$
$X \in [2, 4)$	

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	$\{TH, TTH, TTTH\}$
$X \in [2, 4)$	$\{TH, TTH\}$

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	$\{TH, TTH, TTTH\}$
$X \in [2, 4)$	$\{TH, TTH\}$
$X \in (-9, -3.2]$	

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	$\{TH, TTH, TTTH\}$
$X \in [2, 4)$	$\{TH, TTH\}$
$X \in (-9, -3.2]$	$\emptyset$

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2,4]$	$\{TH, TTH, TTTH\}$
$X \in [2,4)$	$\{TH, TTH\}$
$X \in (-9,-3.2]$	$\emptyset$
$X \in [0.5,1.3) \cup (\pi, \pi+1)$	

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2, 4]$	$\{TH, TTH, TTTH\}$
$X \in [2, 4)$	$\{TH, TTH\}$
$X \in (-9, -3.2]$	$\emptyset$
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\{H, TTTH\}$
$X \in [0, \infty)$	

values of random variables imply events

For any random variable X, any subset of the real line you could think of implies an event.

Examples:

Toss to first head: Let X be the number of tosses.

X taking on any of these values	Implies this event
$X \in \{1\}$	$\{H\}$
$X \in [2,4]$	$\{TH, TTH, TTTH\}$
$X \in [2,4)$	$\{TH, TTH\}$
$X \in (-9,-3.2]$	$\emptyset$
$X \in [0.5, 1.3) \cup (\pi, \pi+1)$	$\{H, TTTH\}$
$X \in [0, \infty)$	S

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	$\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	$\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$
$X \in (-9, -3.2]$	

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	$\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$
$X \in (-9, -3.2]$	0

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2,4]$	$\frac{1}{4}+\frac{1}{8}+\frac{1}{16}=\frac{7}{16}$
$X \in [2,4)$	$\frac{1}{4}+\frac{1}{8}=\frac{3}{8}$
$X \in (-9,-3.2]$	0
$X \in [0.5,1.3) \cup (\pi,\pi+1)$	

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	$\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{1}{2} + \frac{1}{8} = \frac{5}{8}$

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	$\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{1}{2} + \frac{1}{8} = \frac{5}{8}$
$X \in [0, \infty)$	

distribution of a random variable

The distribution of a random variable is the mapping between values of X and their probabilities of the implied events.

(Over-)simplified: the values, and their probabilities.

Examples:

X taking on any of these values	Probability
$X \in \{1\}$	$\frac{1}{2}$
$X \in [2, 4]$	$\frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}$
$X \in [2, 4)$	$\frac{1}{4} + \frac{1}{8} = \frac{3}{8}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{1}{2} + \frac{1}{8} = \frac{5}{8}$
$X \in [0, \infty)$	1

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2,4]$	$\frac{2}{10}$
$X \in [3,5]$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2,4]$	$\frac{2}{10}$
$X \in [3,5]$	$\frac{2}{10}$

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes “uniformly”

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$
$X \in [2, 2.1]$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$
$X \in [2, 2.1]$	0.01

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$
$X \in [2, 2.1]$	0.01
$X \in [2, 2.001]$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$
$X \in [2, 2.1]$	0.01
$X \in [2, 2.001]$	0.0001

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$
$X \in [2, 2.1]$	0.01
$X \in [2, 2.001]$	0.0001
$X \in \{2\}$	

distribution in the bus stop example

X is the amount of time you wait for the bus, with the idea that the bus could come at any random time in the next 10 minutes "uniformly"

X taking on any of these values	Probability
$X \in [2, 4]$	$\frac{2}{10}$
$X \in [3, 5]$	$\frac{2}{10}$
$X \in (-9, -3.2]$	0
$X \in [0.5, 1.3) \cup (\pi, \pi + 1)$	$\frac{0.8}{10} + \frac{\pi}{10}$
$X \in [2, 2.1]$	0.01
$X \in [2, 2.001]$	0.0001
$X \in \{2\}$	0

more notation

Here's how we will actually write the probability statements:

Inconvenient	Usual Notation
$X \in [2, 4]$	$P(2 \leq X \leq 4) = \frac{2}{10}$
$X \in [3, 5]$	$P(3 \leq X \leq 5) = \frac{2}{10}$
$X \in (-9, -3.2]$	$P(-9 < X \leq -3.2) = 0$
$X \in \{2\}$	$P(X = 2) = 0$

distributions and their representations

When you know the distribution of a random variable, you know everything *(as far as probability theory goes.)*

distributions and their representations

When you know the distribution of a random variable, you know everything *(as far as probability theory goes.)*

It's clear from the examples that "distribution" is a complex object, so we'll need convenient representations for them. Here is the first one.

distributions and their representations

When you know the distribution of a random variable, you know everything *(as far as probability theory goes.)*

It's clear from the examples that "distribution" is a complex object, so we'll need convenient representations for them. Here is the first one.

("Theorem":)the distribution of any random variable X is completely determined by the following function $F : \mathbb{R} \to \mathbb{R}$, defined at every $x \in \mathbb{R}$:

$$F(x) = P(X \leq x).$$

F is called the *cumulative distribution function* of X, or cdf for short.