STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-09-26 14:52

random variables - recap

the *functions* so far

1. Probabilitity measure: $P : \mathcal{A} \longrightarrow \mathbb{R}$ and satisfies the three axioms. In general no "picture" possible, because its domain is a collection of events.

2. **Random variable** $X : S \longrightarrow \mathbb{R}$. In general no "picture" possible, because its domain is a sample space. We care about: *its distribution*.

3. Cumulative distribution function F for the random variable X. Defined as $F(x) = P(X \leq x)$. A picture is possible, and does give some information of limited use.

recall "Example 1" "Example 2" and "Example 3"

Example 1: Toss a coin. $S = \{H, T\}$. $X = 1$ if H appears and $X = -1$ if T appears.

Example 2: Toss a coin repeatedly until H appears. $S = \{H, TH, TTH, \ldots\}$. X_2 is the number of tosses required.

Example 3: Pick a real number "uniformly" between 0 and 1. $S = (0, 1)$. X_3 is the identity function. So:

$$F_{X_3}(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

defining properties of a CDF

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. I'll do the middle property (see next slide for an extra detail). The other two are left for you.

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

note — a sequence definition of function limit

There is the general notion of limit: $\lim_{x \to a} f(a) = L$.

Also, a function f has a left-hand limit L^- at a (notation: $\lim_{x \to a-} f(x) = L^-$) if:

$$\lim_{n \to \infty} f(x_n) = L^- \text{ for all non-decreasing sequences } x_n \nearrow a$$

Also, a function f has a right-hand limit L^+ at a (notation: $\lim_{x \to a+} f(x) = L^+$) if:

$$\lim_{n \to \infty} f(x_n) = L^+ \text{ for all non-increasing sequences } x_n \searrow a$$

Fact: If $L^- = L^+ = L$, then $\lim_{x \to a} f(x) = L$.

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).

defining properties of a pmf

A function $p(x)$ is a pmf if:

$$p(x) \geq 0$$

and

$$\sum_{\{x\,:\,p(x)>0\}} p(x) = 1$$

important concept: pmf and cdf are "equivalent"

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

Proof: (surprisingly fussy!)

some important discrete random variables with special named distributions

the Bernoulli(p) distributions - fundamental building blocks

If a random variable X 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).

The phrase "X has a Bernoulli distribution with parameter p" will be abbreviated as:

$$X \sim \text{Bernoulli}(p)$$

"indentically distributed"

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$\left(\frac{1}{2}\right)$, but they are not the same *function*.

We say X_1 and X_2 are "identically distributed", and therefore the same as far as probability is concerned.

Bernoulli(p) pmf and cdf

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

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

Often used as a model for an "experiment" or other random process that either produces an event A of interest, or it doesn't. If A is some event we can define the useful *indicator function*:

$$I_A = \begin{cases} 1 & \text{when } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

I_A will have a Bernouilli distribution with parameter $p = P(A)$.

thank you for subscribing to Bernoulli facts!

Sometimes we'll say q instead of $1 - p$.

It's not clear why the outcomes 0 and 1 are important and useful. Why not just focus on the underlying random outcome?

It starts with the idea of *Bernoulli process*, which I'll introduce now but cannot completely describe until later.

Bernoulli process

Consider an experiment with an event of interest A. Sometimes (unfortunately!), when A occurs we call it a "success". Otherwise, it's a "failure".

Suppose we replicate the experiment. Maybe a finite number n times, or maybe indefinitely.

Each experiment repetition is independent of all the others. (What does this mean?)

If we let the result of the ith replication be $X_i = I_A$, then $X_1, X_2, \ldots$ is called a Bernoulli process (or "sequence of Bernoulli trials").

the Binomial distributions

Stop a Bernoulli(p) process after n trials. Count the number of "successes", or 1's.

This is a random variable. Call it X.
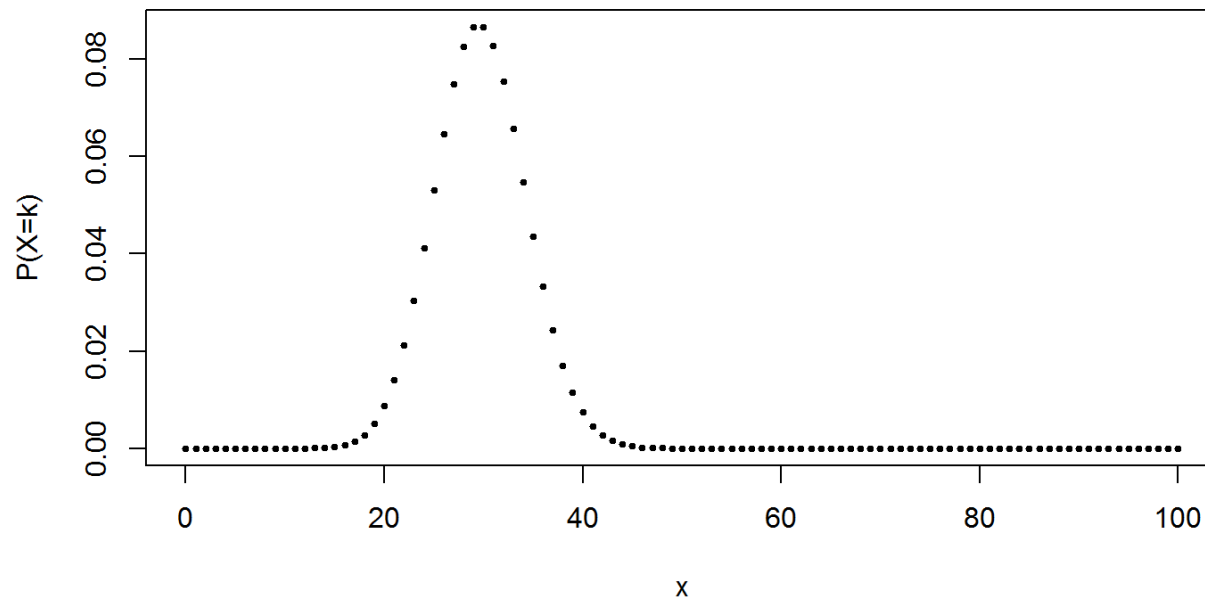
What is the p.m.f. of X?

$$p(k) = P(X = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & x \in \{0, \ldots, n\} \\ 0 & \text{otherwise} \end{cases}$$

Is this a valid pmf? Yes. ("Binomial Theorem")

We say X has a Binomial(n, p) distribution, or $X \sim \text{Binomial}(n, p)$.

a few examples…

The probability that someone is HIV+ given that their ELISA test comes back positive is 0.2971. Suppose we have 100 people with a positive ELISA test. How might one visualize the distribution of the number of people who are HIV+?

…a few examples

The extreme cases have already been considered.

$$P(X = n)$$

$$P(X = 0)$$

$$1 - P(X = n)$$

$$1 - P(X = 0)$$

the geometric distributions

Consider a Bernoulli(p) process. Count the number of trials until the first "success".

This is a random variable. Call it X.

What is the p.m.f. of X?

$$p(k) = P(X = k) = \begin{cases} (1-p)^{k-1}p & x \in \{1, 2, 3, \ldots\} \\ 0 & \text{otherwise} \end{cases}$$

Is this a valid pmf? Yes.

We say X has a geometric distribution with paramter p, or $X \sim \text{Geometric}(p)$.

the "negative binomial" distributions

Consider a Bernoulli(p) process. Count the number of trials until the rth "success".

This is a random variable. Call it X.

What is the p.m.f. of X?

$$p(k) = P(X = k) = \begin{cases} \binom{k-1}{r-1} p^k (1-p)^{k-r} & x \in \{r, r+1, r+2, \ldots\} \\ 0 & \text{otherwise} \end{cases}$$

Is this a valid pmf? Yes (a little obscure to figure out)

We say X has a negative binomial distribution with paramters p and r, or $X \sim \text{NegBin}(p, r)$.