Stat 394 Probability I

Lecture 5

Richard Li

July 4, 2017

Example

- Toss a coin 10 times and let X be the number of Heads.
- Choose a random point in the unit square $\{(x, y) : 0 \leq x, y \leq 1\}$ and let X be its distance from the origin.
- Choose a random person in a class and let X be the height of the person, in inches.

Random variables

A **random variable** is a number whose value depends upon the outcome of a random experiment.

$S \xrightarrow{X} \{1, 2, 3, 5\}$

A random variable X is a real-valued function on sample space S:

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

Example

Toss 2 coins. $X = \#$ of H's.

$\{HH\}$, $\{HT\}$, $\{TH\}$, $\{TT\}$.

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$

$2 \quad 1 \quad 1 \quad 0$

$P(X=0) = \frac{1}{4}$

$s \in S$

$P(X=x) \longrightarrow P(s)$ where $X(s)=x$

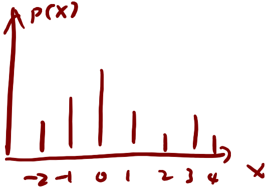
correspond to some events.

Discrete random variables

A **discrete random variable** X has finitely or countably many values $x_i, i = 1, 2, \dots$

$p(x_i) = P(X = x_i)$ with $i = 1, 2, \dots$ is called the **probability mass function** (or p.m.f) of X.

$$\sum_{i=1}^{\infty} p(x_i) = 1$$



Using r.v. to solve probability problems

4 balls are to be randomly selected without replacement from an urn containing 10 balls numbered 1 through 10. If we bet that at least one of the balls that are drawn has a number as large as or larger than 7, what is the probability that we win the bet?

X : largest # selected

$$P(X=i) = \frac{\binom{i-1}{3}}{\binom{10}{4}} \qquad i = 4, 5, 6, \ldots, 10$$

$$P(X \geq 7) = P(X=7) + P(X=8) + P(X=9) + P(X=10)$$

$$= \ldots$$

Bernoulli random variables

(and the probability of success is p)

Suppose an experiment where the outcome can be considered as either a *success* or a *failure*, let $X = 1$ when outcome is success and $X = 0$ when outcome is failure, then X is said to be a **Bernoulli random variable** with parameter p.

p.m.f.

$$\begin{cases} p(0) = 1 - p \\ p(1) = p \end{cases}$$

Def 2.

a r.v. is a Bernoulli r.v if its p.m.f is as on the left.

And we can say $X \sim \text{Bern}(p)$

Binomial random variables

- $X \sim \text{Bern}(p) \iff X \sim \text{Bin}(n, p)$

- Example: Toss a fair coin 3 times, let X denote # of heads $\Rightarrow X \sim \text{Bin}(3, 1/2)$

Suppose for n independent trials, each of which results in a *success* with probability p and *failure* with probability $1 - p$, and let X represent the number of success that occur in the n trials, then X is said to be a **binomial random variable** with parameter (n, p).

$(X \sim \text{Bin}(n, p))$

p.m.f: $p(i) = P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \; i = 0, 1, \cdots, n$

We can check

binomial formula.

$$\sum_{i=0}^{n} p(i) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = (p + (1-p))^n = 1$$

Bernoulli and Binomial distribution

$X \sim Bin(n, p)$

- X is # of success out of n trials with success prob p
- $X = X_1 + X_2 + \cdots + X_n$, $X_i = \begin{cases} 1 & \text{if trial } i \text{ is success.} \\ 0 & \text{otherwise.} \end{cases}$ $i = 1, 2, 3, \ldots, n$.

i.e. $X_1, X_2, \ldots, X_n \overset{i.i.d}{\sim} Bern(p)$

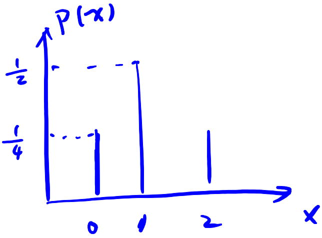
independently and identically distributed

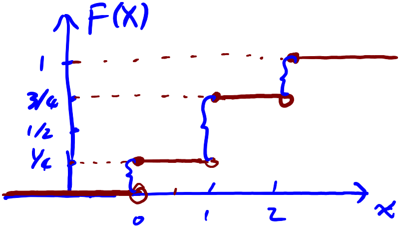
Cumulative distribution function (CDF)

Def.

$$F(x) = P(\underbrace{X \le x}_{\text{event.}}), \quad -\infty < x < \infty$$

- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$





$F(0) = P(X \le 0) = P(X=0) = \frac{1}{4}$

$F(0.5) = P(X \le 0.5) = P(X=0)$

Example

$X \sim Bin(2k-1, p)$

Example

Let X: # of functional comp. in the first $(2k-1)$ comp

A communication system consists of n components, each of which will, independently, function with probability p. The total system will be able to operate effectively if at least one-half of its components function. In general, when is a $(2k+1)$-component system better than a $(2k-1)$-component system?

Ans: when $p > 1/2$

$p_1 = P(\text{2k+1 system is effective}) = P(X \geq k+1) + P(X=k)\cdot(1-(1-p)^2) + P(X=k-1)\cdot p^2$

$\Rightarrow$ at least $k+1$ comp. function

$p_2 = P(\text{2k-1 system is effective}) = P(X \geq k) = P(X=k) + P(X \geq k+1)$

$p_1 - p_2 = -P(X=k)(1-p)^2 + P(X=k-1)\cdot p^2$

$= -\binom{2k-1}{k} p^k (1-p)^{k-1} \cdot (1-p)^2 + \binom{2k-1}{k-1} p^{k-1}(1-p)^k \cdot p^2$

$\binom{2k-1}{k}$

$= \binom{2k-1}{k}\left[-p^k(1-p)^{k+1} + p^{k+1}(1-p)^k\right] = \binom{2k-1}{k} p^k (1-p)^k \left[-(1-p)+p\right]$

$2p-1$

First, why do we want to study random variables

A quote:

"Once you get what a random variable is, it can be hard to explain. Now that I understand what a random variable is, it is difficult to remember what was difficult to understand about it"

First, why do we want to study random variables

But seriously, the way I see the advantage of using random variables

- Describe more complicated events without cumbersome notations
- Focus on the quantities we care about
- Consider problems beyond calculating probabilities
- Handle continuity and infinity more easily

more on 395

Expectation

Variance

...

Expectation

Assume X is a discrete random variable with possible values $x_i, i = 1, 2, \dots$. Then the **expected value**, or **expectation**, of X is

$$E(X) \text{ or } EX = \sum_i x_i p(x_i)$$

X : # H for 2 tosses

$p(0) = 1/4$
$p(1) = 1/2$
$p(2) = 1/4$.

$E(X) = 0 \times 1/4 + 1 \times 1/2 + 2 \times 1/4$
$= 1/2 + 1/2 = 1$

Example

A box contains 5 red balls and 3 blue balls. Two balls are withdrawn randomly, and if they are of the same color, you win \$2, otherwise you lose \$1. What is the amount you expect to win?

X : amount of \$ you walk away with.

$$P(X=2) = \frac{\binom{5}{2}+\binom{3}{2}}{\binom{8}{2}} = \frac{13}{28}$$

$$P(X=-1) = 1-\frac{13}{28} = \frac{15}{28}$$

$$E(X) = 2\times\frac{13}{28} + (-1)\times\frac{15}{28} = \frac{11}{28}$$

Chocolate store

A candy store has M boxes of chocolates. The number of chocolates in each box vary. Suppose for $k = 1, 2, ..., N$, there are n_k boxes containing k chocolates, and $\sum_{k=1}^{N} n_k = M$. The store owner keeps a magic book where all the chocolates labeled $1, 2, ..., \sum_{k=1}^{N} kn_k$ are listed.

The store owner lets you to take one box of chocolate for free, and gives you two options:

1. Take one box without knowing how many chocolates are in the box.
2. Pick one chocolate from the book, and you can take the box containing that chocolate.

Chocolate store

Expectation of function of a r.v.

Assume X is a discrete random variable with possible values $x_i, i = 1, 2, \ldots$. Then the **expected value**, or **expectation**, of $g(X)$, where g is some real-valued function, is

$$E(g(X)) \text{ or } Eg(X) = \sum_i g(x_i)p(x_i)$$

Example

Variance

If X is a random variable, we say the **variance** of X is defined by

$$Var(X) = E((X - E(X))^2)$$

and the square root of the variance is usually called the **standard deviation** of X

Example

A box contains 5 red balls and 3 blue balls. There are two games offered to you. Both games ask you to take one ball out,

- Game 1: if it is a red ball, you win \$2, otherwise you lose \$2.
- Game 2: if it is a red ball, you win \$5, otherwise you lose \$7.

Alternative formula for variance

Expectation and variance under linear transformation

Binomial random variable revisit

Binomial random variable revisit

Newsboy

A newsboy buy papers at 10 cents and sells them at 15 cents. However, he is not allowed to return unsold papers. If his daily demand is a binomial random variable with $n = 10$ and $p = \frac{1}{3}$, approximately how many papers should he purchase so as to maximize his expected profit?

Newsboy

Binomial p.m.f

http://shiny.albany.edu/stat/binomial/

Birthday problem revisit

Birthday problem revisit