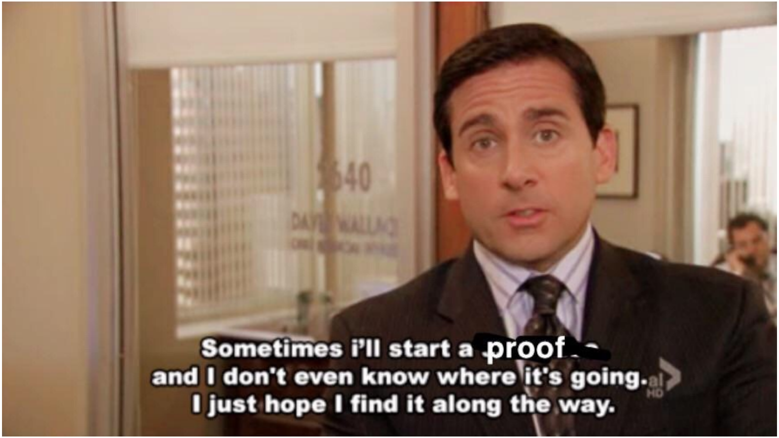
STAT 230 SECTION 2 LECTURE 11



Today's Agenda

Last time:

- Discrete random variables and probability function

Today (Lec 11, 05/27):

- Cumulative distribution function
- Discrete uniform distribution

5. DISCRETE RANDOM VARIABLES

Definition

A **random variable** is a function that maps from the sample space $S$to the set of real numbers $\mathbb{R}$. In other words, we say X is a random variable if

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

Often "random variable" is abbreviated with "RV".

Definition

The values that a random variable takes is called the **range** of the random variable. We often denote the range of a random variable X by $X(S)$.

Definition

- We say that a random variable is **discrete** if its range is a discrete subset of $\mathbb{R}$ (i.e., a finite or a countably infinite set).
- A random variable is **continuous** if its range is an interval that is a subset of $\mathbb{R}$ (e.g. $[0, 1], (0, \infty), \mathbb{R}$).

Definition

The **probability (mass) function** of a discrete random variable X is the function

$$f_X(x) = P(X = x).$$

We use the short hand $P(X = x)$ to denote $P(\{\omega \in S : X(\omega) = x\})$. The definition f is valid for all x, but its value is zero when x is outside the range of random variable X.

Any probability function f satisfied

a)
$$0 \leq f_X(x) \leq 1 \quad \text{for all } x$$

b)
$$\sum_{x \in X(S)} f_X(x) = 1$$

Now we have a way to assign probability to specific values of a random variable.

However, we often encounter situations like “at least” or “at most”. What do we do then?

Definition

The **cumulative distribution function** (cdf) of a random variable X is

$$F_X(x) = P(X \leq x), \;\; x \in \mathbb{R}.$$

Again we use the short hand notation

$$P(X \leq x) = P(\{\omega \in S : X(\omega) \leq x\}).$$

If X is discrete with probability function f_X, then

$$F_X(x) = P(X \leq x) = \sum_{y\colon y \leq x} f_X(y)$$

Properties of the cdf

Let F_X be a cdf. Then,

a) $0 \leq F_X(x) \leq 1$

b) $F_X(x) \leq F_X(y)$ for $x < y$

c) $\lim_{x \to -\infty} F_X(x) = 0$, and $\lim_{x \to \infty} F_X(x) = 1$.

Example

Suppose a fair six sided die is rolled, and let X denote the number of dots on the top face of the die. Compute and graph $F_X(x)$.

Example

Suppose students A,B and C each independently answer a question on a test. The probability of getting the correct answer is 0.9 for A, 0.7 for B and 0.4 for C. Let X denote the number of people who get the answer correct.

a) Compute the probability function of X
b) Compute the cdf of X.
c) Compute the probability that at least one person gets the answer correct.

Example

Suppose that N balls labelled $1, 2, \ldots, N$ are placed in a box, and n balls ($n \leq N$) are randomly selected without replacement. Define the random variable $X =$ largest number selected. Find the probability function of X.

A notation

Two random variables X and Y are said to **have the same distribution** if $F_X(x) = F_Y(x)$ for all $x \in \mathbb{R}$. We denote this by

$$X \sim Y$$

Definition

Suppose the range of X is $\{a, a+1, \ldots, b\}$, where $a, b \in \mathbb{Z}$, and suppose all values are equally likely. Then we say that X has a **discrete uniform distribution** on $\{a, a+1, \ldots, b\}$, shorthand: $X \sim U[a, b]$.

If $X \sim U[a, b]$, then its probability function is defined as

$$P(X = x) = \frac{1}{b - a + 1}, \quad \in \{a, a+1, \ldots, b\},$$

and if $x \notin [a, b]$, then $P(X = x) = 0$.

Examples of discrete uniform random variables

Experiment	X	Distribution
Roll a 6 sided die	# showing on die	$U[1, 6]$
Draw a number between 1 and 50	# Drawn	$U[1, 50]$

Example

Suppose $X \sim U[a, b]$. Compute $F_X(x)$, and verify that $F_X(x)$ is indeed a valid cdf.

Example

Suppose that $X \sim U[1,6]$ and $Y \sim U[1,20]$. Which of the following are true?

a) $P(X > 3) < P(Y > 10)$

b) $P(X > 3) = P(Y > 10)$

c) $P(X > 3) > P(Y > 10)$