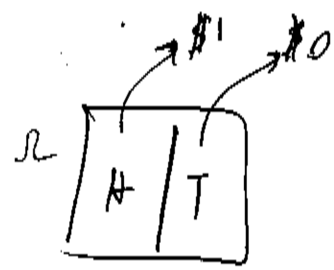
Lecture 3/10/15

Random Variable Theory: just the vocabulary (nothing hard)

Prob: used to make calculations of chance/odds

What if you want to model an outcome's result? There is an outcome for each result...



$P(H) = \frac{1}{2}$ and get \$1

$P(T) = \frac{1}{2}$ " \$0

This can be modeled with a random variable (r.v.)

$$X \sim \begin{cases} \$1 & \text{w.p. } \frac{1}{2} \\ \$0 & \text{w.p. } \frac{1}{2} \end{cases}$$

H,T not here!

Capital letter "distrib" resulting outcome chance of getting it

$P(X=\$1) = \frac{1}{2}$

$P(X=\$0) = \frac{1}{2}$

H,T not here because we don't care if they happen, because we just see the resulting outcomes

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

an outcome you put in — a real number out

$f : \mathbb{R} \to \mathbb{R}$ (domain, range)

$X(\omega) <^{x_1}_{x_2}$ not allowed

$X(H) = \$1, \quad X(T) = \$0$

Remember P is a set function $P : 2^{\Omega} \to [0,1]$

How is it defined for X?

$P(X = \$1) := P(\{\omega : X(\omega) = \$1\}) = P(\{H\}) = \frac{|\{H\}|}{|\{H,T\}|} = \frac{1}{2}$

Rule: $\sum_{x \in Supp} P(X = x) = 1$

wrong but okay for now

How many ω's give us this outcome?

Process: $\omega \in \Omega$ chosen $\Rightarrow X(\omega) \Rightarrow \#$ (dumb switchboard)

all we see is \$1, \$0, \$1, \$1, ... pop out

$Supp[X] := \{X(\omega) : \omega \in \Omega\} \subseteq \mathbb{R}$ always

$= \{x : P(X = x) > 0\}$

"all results that can happen"

$X(Red) = \$1$

$X(Green) = \$0$

$\Rightarrow X \sim \begin{cases} \$1 & \text{wp } \frac{1}{2} \\ \$0 & \text{wp } \frac{1}{2} \end{cases}$

There are many Ω's that can make the same r.v.

$\Rightarrow$ We don't care about the underlying Ω.

This X is called Bernoulli

$X \sim \text{Bernoulli}\left(\frac{1}{2}\right) := \begin{cases} 1 & \text{wp } \frac{1}{2} \\ 0 & \text{wp } \frac{1}{2} \end{cases}$ What is $Supp[X]$? $= \{0, 1\} \subseteq \mathbb{R}$

Generally,

$X \sim \text{Bernoulli}(p) := \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$ a parameter

$X \sim Deg(c) = \{ c \text{ w.p. } 1$

Degenerate r.v.'s are technically r.v.'s but not interesting.

What are the valid values of p?

$p \in (0,1)$ the parameter space

Why not $p \in \{0,1\}$?

If $|Supp(X)| \le |\mathbb{N}|$ i.e. finite or cntbly. infinite,

X is called a "discrete r.v."

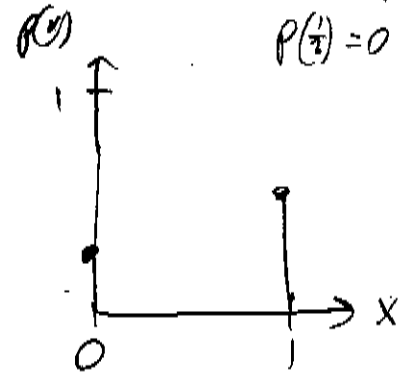
let $p(x) := P(X=x)$ called the prob. mass function (PMF)

Rules: ① $p(x) \in [0,1]$ $\forall x \in \mathbb{R}$ ← they're prob's (of course) ② $\sum_{x \in Supp(X)} p(x) = 1$

↑

Something must happen kinda like $P(\Omega) = 1$

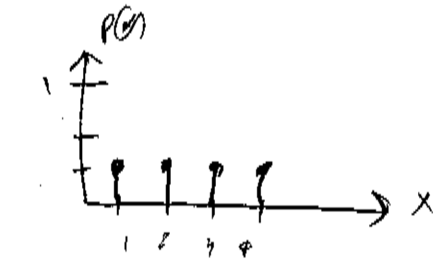
$X \sim Bernoulli(p)$ ① $p(0) = 1-p \in [0,1]$, $p(1) = p \in [0,1]$, $p(\frac{1}{2}) = 0$ (not in $Supp(X)$) ② $\sum_{x \in Supp(X)} p(x) = \sum_{x=0}^{1} p(x) = \underbrace{(1-p) + p}_{p(0)+p(1)} = 1$ ✓



Plot of PMF

$X \sim Unif(\{1,2,3,4\}) := \begin{cases} 1 & \text{w.p. } \frac{1}{4} \\ 2 & \frac{1}{4} \\ 3 & \frac{1}{4} \\ 4 & \frac{1}{4} \end{cases}$ ← $\frac{1}{|Supp(X)|}$ each elem is "equally likely"

"Discrete Uniform Distr"



Parameter space $A \subset \mathbb{R}$ s.t. $|A| < |\mathbb{N}|$

$|A| \ge 2$

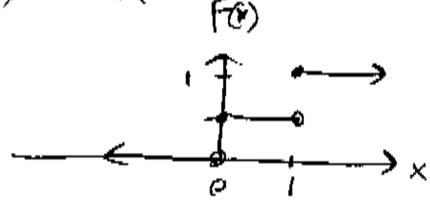
$X \sim Rademacher := \begin{cases} -1 & \text{w.p. } \frac{1}{2} \\ 1 & \text{w.p. } \frac{1}{2} \end{cases}$ Why is this called the "random walk"?

No parameters!

"distribution function" or "cumulative distribution function"

$F(x) := P(X \le x),\ x \in \mathbb{R}$

$X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$



$F(-32) = 0$ defined everywhere

$F(17) = 1$

$F(0) = \frac{1}{2},\ F\left(\frac{1}{7}\right) = \frac{1}{2}, \ldots$

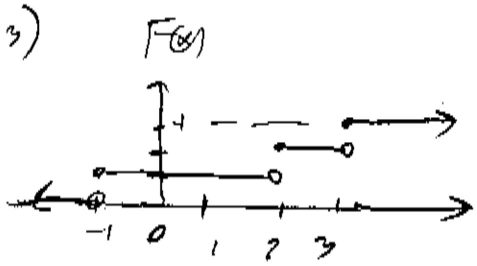
Rules

① $\lim_{x \to \infty} F(x) = 1$ ② $\lim_{x \to -\infty} F(x) = 0$ ③ $F(x) \in [0,1]\ \forall x$ ④ $\forall x,y \in \mathbb{R}$ $x \le y \Rightarrow F(x) \le F(y)$

monotonically increasing

consequences

$X \sim \text{Unif}(-1, 2, 3)$



"Sampling w/o replacement"

10 cards, 4R, 6B

$P(\text{getting } 2 R \text{ on a draw of } 3) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}$

$P(\text{getting } x\ R \text{ on '' } 3) = \frac{\binom{4}{x}\binom{6}{3-x}}{\binom{10}{3}}$ STOP

K: # of total successes

N-K: " failures

N: total successes + failures

n: sample size

X be the r.v. representing the # of R drawn

$X \sim \text{Hyper}(n, K, N) := \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} = p(x)$

$X \sim \text{Hypergeometric}(n=3, K=4, N=10)$ $\quad \text{Supp}[X] = \{0,1,2,3\}$

$F(x)$? HARD...

Parameter space $N \in \{1,2,\ldots\},\ K \in \{1,2,\ldots,N\},\ n \in \{1,2,\ldots,K\}$

(Note $N \ne 0$)

$N=0 \Rightarrow K=0 \Rightarrow n=0 \quad X \sim \text{Deg}(0)$

$N=10, K=10 \Rightarrow X \sim \text{Deg}(n)$