Mathematics for Computer Science

MIT 6.042J/18.062J

Introduction to Random Variables

Guess the Bigger Number

Team 1:

- Write different integers between 0 and 7 on two pieces of paper
- Show to Team 2 face down

Team 2:

- Expose one paper and look at number
- Either *stick* or *switch* to other number

Team 2 wins if gets **larger** number

Strategy for Team 2

- pick a paper to expose giving each paper equal probability.
- if exposed number is "small" then switch, otherwise stick. That is

switch if $\leq$ threshold Z where

Z is a random integer, $0 \leq Z \leq 6$.

Analysis of Team 2 Strategy

Case M: low $\leq Z <$ high

Team 2 wins in this case, so

Pr{Team 2 wins | M} = 1

and $\Pr\{M\} = \frac{1}{7}$

Analysis of Team 2 Strategy

Case H: high $\leq Z$

Team 2 will switch, so wins iff low card gets exposed

$\Pr\{\text{Team 2 wins} \mid H\} = \frac{1}{2}$

Analysis of Team 2 Strategy

Case L: $Z <$ low

Team 2 will stick, so wins iff high card gets exposed

$\Pr\{\text{Team 2 wins} \mid L\} = \frac{1}{2}$

Analysis of Team 2 Strategy

So 1/7 of time, sure win.

Rest of time, win 1/2, so

Pr{Team 2 wins} =

$$\frac{1}{7}\cdot 1+\left(1-\frac{1}{7}\right)\cdot\frac{1}{2}=\frac{4}{7}$$

Analysis of Team 2 Strategy

Does not matter what Team 1 does!

Team 1 Strategy

...& Team 1 can play so

$$\Pr\{\text{Team 2 wins}\}\leq\frac{4}{7}$$

whatever Team 2 does

Random Variables

Informally: an RV is a number produced by a *random process:*

- threshold variable Z
- number of larger card
- number of smaller card
- number of exposed card

Random Variables

Informally: an RV is a number produced by a *random process:*

- #hours to next system crash
- #faulty chips in production run
- avg # faulty chips in many runs
- #heads in n coin flips

Intro to Random Variables

Example: Flip three fair coins

C ::= # heads (**C**ount)

$$M ::= \begin{cases}1 & \text{if all Match,}\\ 0 & \text{otherwise.}\end{cases}$$

Intro to Random Variables

Specify events using values of variables

- [C = 1] is event "exactly 1 head"

  Pr{C = 1} = 3/8
- Pr{$C \geq 1$} = 7/8
- Pr{$C \cdot M > 0$} = Pr{M>0 and C>0}

  = Pr{all heads} = 1/8

What is a Random Variable?

Formally,

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

Sample space (S) (usually) ($\mathbb{R}$)

Independent Variables

random variables R,S are independent iff

[$R = a$], [$S = b$]

are independent events for all a, b

Independent Variables

alternate version:

Pr{$R = a$ AND $S = b$} = Pr{$R = a$} · Pr{$S = b$}

Binomial Random Variable

$B_{n,p}$::= # heads in n independent flips.

Coin may be biased. So 2 parameters

n ::= # flips, p ::= Pr{head}.

C is binomial for 3 flips: C is $B_{3,1/2}$

for n=5, p=2/3

Pr{HHTTH} =

Pr{H} · Pr{H} · Pr{T} · Pr{T} · Pr{H}

(by independence)

Binomial Random Variable

$B_{n,p}$::= # heads in n independent flips.

Coin may be biased. So 2 parameters

n ::= # flips, p ::= Pr{head}.

C is binomial for 3 flips: C is $B_{3,1/2}$

for n=5, p=2/3

Pr{HHTTH} = $\left(\frac{2}{3}\right)^3 \cdot \left(\frac{1}{3}\right)^2$

$\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3}$

Binomial Random Variable

$B_{n,p}$ ::= # heads in n independent flips.
Coin may be biased. So 2 parameters
n ::= # flips, p ::= Pr{head}.

Pr{each sequence w/i H's, n-i T's} =

$$p^i(1-p)^{n-i}$$

Binomial Random Variable

$B_{n,p}$ ::= # heads in n independent flips.
Coin may be biased. So 2 parameters
n ::= # flips, p ::= Pr{head}.

Pr{get i H's, n-i T's} =

$$\binom{n}{i}p^i(1-p)^{n-i}$$

Binomial Random Variable

$B_{n,p}$ ::= # heads in n independent flips.
Coin may be biased. So 2 parameters
n ::= # flips, p ::= Pr{head}.

Pr{ $B_{n,p} = i$ } =

$$\binom{n}{i}p^i(1-p)^{n-i}$$

Density & Distribution

The Probability Density Function
of random variable R,

$$PDF_R(a) ::= \Pr\{R = a\}$$

so

$$PDF_{B_{n,p}}(i) = \binom{n}{i}p^i(1-p)^{n-i}$$

Uniform Distribution

R is uniform iff PDF_R is constant.
R ::= outcome of fair die roll.
Pr{R=1} = Pr{R=2} = ··· = Pr{R=6} = 1/6
S ::= 4-digit lottery number
Pr{S = 0000} = Pr{S = 0001} = ···
= Pr{S = 9999} = 1/10000

Team Problems

Problems
2 — 4