Mathematics for Computer Science

MIT 6.042J/18.062J

Introduction to Random Variables

 November 15, 2002

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 number
– Either *stick*, or *switch* to other number
– Expose other number
– win if end with *larger* number

 November 15, 2002

Guess the Bigger Number

Try it out!

 November 15, 2002

Strategy for Team 2

Choose papers with equal probability. If exposed number is "small" then switch; otherwise stick.

"small" means $\leq$ threshold Z.

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

 November 15, 2002

Analysis of Team 2 Strategy

Case 1 ($low \leq Z < high$):

$\Pr\{Z = low\} = 1/7$, so

$\Pr\{\text{Case 1}\} \geq 1/7.$

Team 2 wins in this case, so

$\Pr\{\text{Team 2 wins} \mid \text{Case 1}\} = 1.$

 November 15, 2002

Analysis of Team 2 Strategy

Case 2 ($high \leq Z$):

Team 2 wins iff low card exposed.

$\Pr\{\text{Team 2 wins} \mid \text{Case 2}\} = 1/2.$

 November 15, 2002

Analysis of Team 2 Strategy

Case 3 ($Z < low$):

Team 2 wins iff high card exposed.

Pr{Team 2 wins | Case 3} = 1/2.

Analysis of Team 2 Strategy

At least 1/7 of time, sure win.

Rest of time, 50/50 win, so

Pr{Team 2 wins}≥

$$\frac{1}{7}\cdot 1 + \frac{6}{7}\cdot\frac{1}{2} = \frac{4}{7} > \frac{1}{2}$$

Analysis of Team 2 Strategy

Does not matter what Team 1 does!!

In-class Problem:Team 1 Strategy

How can Team 1 guarantee

Pr{Team 2 wins} ≤ 4/7

no matter what Team 2 does?

Prove it.

Random Variables

Examples of *random variables:*

- Number of larger card
- Number of smaller card
- Number of exposed card
- "threshold" number Z

Intuitive definition:

a number produced by

a "random process"

Intro to Random Variables

Example: Flip three fair coins.

C ::= number of heads (**c**ount).

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

Intro to Random Variables

Specify events using values of variables.

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

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

  $= \Pr\{\text{all heads}\} = 1/8$

Density & Distribution

- *probability density function* of *R:*

  $f_R(r) ::= \Pr\{R = r\}$
- *(Cumulative) distribution function:*

  $F_R(r) ::= \Pr\{R \leq r\}$

Uniform Distribution

R is *uniform* iff f_R is constant.

Examples: R_1 ::= outcome of fair die.

$\Pr\{R_1=1\} = \cdots = \Pr\{R_1=6\} = 1/6$

R_2 ::= 4 digit lottery number

$\Pr\{R_2 = 0000\} = \Pr\{R_2 = 0001\} = \cdots =$
$\Pr\{R_2 = 9999\} = 1/10000$

Uniform Distribution

"Threshold" variable Z:

$f_Z(i) ::= \Pr\{Z = i\} = 1/7,$

$F_Z(i) ::= \Pr\{Z \leq i\} = (i+1)/7,$

for $i = 0,1,\ldots,6.$

What is a Random Variable?

Formally, an RV, R, is any function:

$$R: \underbrace{S}_{\text{Sample Space}} \rightarrow \text{Reals}.$$

In-class Problems

Problems 1& 2

(Lecture ended here)

Independent Variables

Random variables R, S are *independent* iff $[R = a]$ and $[S = b]$ are independent *events* for all real a,b.

 November 15, 2002 L11-1.19

Independent Variables

Alternative version: R, S *independent* iff

$$\Pr\{R = a \mid S = b\} = \Pr\{R = a\}.$$

Alternative version 2:

$$\Pr\{R = a \text{ and } S = b\} = \Pr\{R = a\} \cdot \Pr\{S = b\}.$$

 November 15, 2002 L11-1.20

Indicator Variables

Event A has *indicator variable*

$$I_A ::= \begin{cases} 1 & \text{if } A \text{ occurs,} \\ 0 & \text{if } \overline{A} \text{ occurs.} \end{cases}$$

Sanity check:

I_A and I_B are independent iff A and B are independent.

 November 15, 2002 L11-1.21

Independent Variables

Mutual Independence:

$\Pr\{A_1=a_1 \text{ and } A_2=a_2 \ldots\text{and } A_n=a_n\}=$
$\Pr\{A_1=a_1\}\times\Pr\{A_2=a_2\}\times\cdots\times\Pr\{A_n=a_n\}.$

Pairwise Independence:

$$\Pr\{A_i=a_i \text{ and } A_j=a_j\}= \Pr\{A_i=a_i\} \times \Pr\{A_j=a_j\}.$$

all $i \neq j$.

 November 15, 2002 L11-1.22

Independent Variables

Example:

R_i ::= indicator for Head on coin i

$$P ::= R_1 + R_2 + R_3 \pmod 2).$$

Not mutually independent: any 3 determine the 4th one.

 November 15, 2002 L11-1.23

Independent Variables

Example:

R_i ::= indicator for Head on coin i

$$P ::= R_1 + R_2 + R_3 \pmod 2).$$

But 3-way independent, e.g.,

$\Pr\{P = 0 \text{ and } R_1 = 1 \text{ and } R_0 = 0\} =$
$\Pr\{P = 0\}\times\Pr\{R_1 = 1\}\times\Pr\{R_0 = 0\}$

 November 15, 2002 L11-1.24

Independent Variables

Pairwise Independence sufficient for major applications (in later lecture).

Good to know, since pairwise holds in important cases where mutual does not.

In-class Problems

Problem 3

Distributions

Bernoulli variable ::= 0-1 valued (indicator variable)

Example:

M, R_i, (C mod 2) are Bernoulli, but not C.

Binomial Distribution

Binomial distribution is distribution of # heads in n mutually independent flips.

Example: C is binomial for 3 flips.

Coin may be biased. So 2 parameters

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

C has binomial density $f_{3,1/2}$

Binomial Distribution

Alternative definition:

B is binomial (n,p) random variable iff

$$B = X_1 + X_2 + \cdots + X_n$$

for $\{X_i\}$ mutually independent Bernoulli with $\text{pr}\{X_i=1\} = p$.

Binomial Distribution

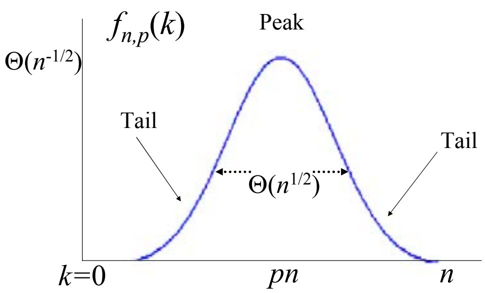
Pr{k Heads in n flips} =

$$\underbrace{f_{n,p}}(k) ::= \binom{n}{k} p^k q^{n-k}$$

binomial density function

$$q ::= 1 - p$$

The Binomial PDF



In-class Problems

Problem 4