CS1800
Discrete Structures
Fall 2019

Lecture 12
10/15/19

Last time

- Conditional Probability
- Bayes Law
- Markov Chains

Today

- Finish M.C.
  - iteration method
- Finish Expectation & Variance
  - linearity of expectation
  - variance & standard deviation
- Entropy

Next time

- Review

Guess an answer for stationary distribution

$B_0 = 1/3$

$M_0 = 1/3$

$S_0 = 1/3$

① $B = .7 \cdot B + .3 \cdot M + .3 \cdot S$

② $M = 0.2 \cdot B + 0.6 M + 0.2 \cdot S$

③ $S = 0.1 \cdot B + 0.1 M + 0.5 \cdot S$

Try it

new guess

$B_1 = .7 \times 1/3 + .3 \times 1/3 + .3 \times 1/3 = .433$

$M_1 = .2 \times 1/3 + .6 \times 1/3 + .2 \times 1/3 = .333$

$S_1 = .1 \times 1/3 + .1 \times 1/3 + .5 \times 1/3 = .233$

① 1/2

② 1/3

③ 1/6

$B_2 = .7 \times .433 + .3 \times .333 + .3 \times .233 = .473\overline{3}$

$M_1 = \quad \ldots \quad = .33\overline{3}$

$S_1 = \quad \ldots \quad = .193\overline{3}$

new guess

More on Expectation

Example: · Roll two fair 6-sided die

· Let X = sum of die faces

· Q: $E[X]$?

die 2

die 1	1	2	3	4	5	6
1	(1,1)	(1,2)				
2	(2,1)			(2,4)		
3						
4						
5						
6						(6,6)

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

e.g. $X((2,4)) \to 6$

$$E[X] = \sum_{x} x \cdot \Pr[X=x]$$

$$\left\{\begin{array}{ll} X=2 & \Pr[X=2] = 1/36 \\ X=3 & \Pr[X=3] = 2/36 \\ X=4 & \Pr[X=4] = 3/36 \\ & \vdots \\ X=12 & \Pr[X=12] = 1/36 \end{array}\right.$$

$$E[X] = \sum_{x} x \cdot \Pr[X=x]$$

$$= 2 \cdot 1/36 + 3 \cdot 2/36 + \cdots + 12 \cdot 1/36 = 7$$

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot p(\omega)$$

$$= \sum_{\omega \in \Omega} X(\omega) \cdot 1/36$$

$$= 1/36 \cdot \sum_{\omega \in \Omega} X(\omega)$$

$$= 1/36 \cdot [\text{sum of table}]$$

$$\vdots$$

$$= 7$$

Linearity of Expectation

Let X_1 = r.v. for first die roll

Let X_2 = r.v. for second die roll

Let $X = X_1 + X_2$

$$E[X] = E[X_1 + X_2] \overset{*}{=} E[X_1] + E[X_2] = 3.5 + 3.5 = 7$$

Variance & standard deviation

Case 1	case 2	case 3
4'10" 5' 5'2"	4' 5' 6'	3' 5' 7'
$E[X_1] = 5'$	$E[X_2] = 5'$	$E[X_3] = 5'$

How to measure "variability"

case 2

① $E[Y_1] = \sum_{\omega \in \Omega} Y(\omega) \cdot P(\omega)$

$= -12'' \cdot 1/3 + 0'' \cdot 1/3 + (+12'') \cdot 1/3$

$= 0''$

② $E[Y_2] = |-12''| \cdot 1/3 + |0''| \cdot 1/3 + |12''| \cdot 1/3$

$= 12 \cdot 1/3 + 0 \cdot 1/3 + 12 \cdot 1/3 = 8''$

③ $E[Y_3] = (-12'')^2 \cdot 1/3 + (0'')^2 \cdot 1/3 + (12'')^2 \cdot 1/3$

$= 96 \text{ in}^2$

3 ways

① $Y_1 = X - E[X]$ × neg & pos cancel

② $Y_2 = |X - E[X]|$ - mean absolute deviation

③ $Y_3 = (X - E[X])^2$ - variance

$\Rightarrow$ take square root, get standard deviation $\sqrt{96 \text{ in}^2} = 9.8$ in

$$\sigma^2 = Var(X) = E\{(X - E\{X\})^2\}$$ — <u>variance</u>

$$\Rightarrow \sigma = \sqrt{Var(X)} = \sqrt{E\{(X - E\{X\})^2\}}$$ — <u>standard deviation</u>

(back in original units)

Claim: $E[(X - E[X])^2] = E\{X^2\} - (E\{X\})^2$

example: case 2 $E(X) = 5' = 60''$

$$E[X^2] = \frac{(48'')^2 + (60'')^2 + (72'')^2}{3}$$

claim: $3696 - 60^2 = 3696 - 3600 = \underline{\underline{96}}$

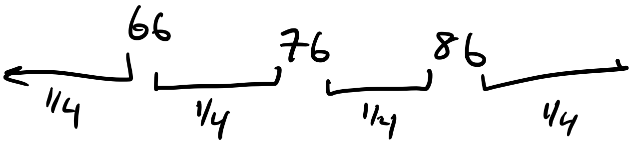
Exam scores

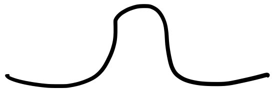
$E\{X\} = 76$

$\sigma = 15$

$2/3\,\sigma = 10$ pts

$\pm\, 2/3\,\sigma$ — breaks scores into quartiles





Entropy

- Consider 8 letters only $\{\overbrace{A, B, C, D, E, F, G, H}^{8 \text{ letters}}\}$

  000 001 010 011 100 101 110 111

- needs 3-bits

  $A \to 000$
  $B \to 001$
  $\vdots$
  $H \to 111$

  $FAD = 101\ 000\ 011$

- suppose had to encode n- things need: $\lceil \log_2 n \rceil$

  k bits, can represent 2^k things

  $$n = 2^k \Rightarrow k = \log_2 n$$

- Efficiency of code is measured in bits-per-character used on average, BPC   BPC = 3

- Variable length code: assign short codes to frequent letters

  longer codes to infrequent letters

✓ why do we need longer codes?

E.g.

A	B	C	D	E	F	G	H
00	01	010	011	100	101	110	111

BBB 010101

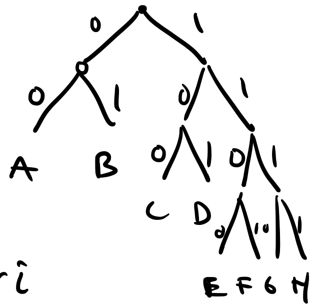
CF 010101

E.g.

$$\begin{array}{ccccccccc} & A & B & C & D & E & F & G & H \\ P_1 = (& 1/4, & 1/4, & 1/8, & 1/8, & 1/16, & 1/16, & 1/16, & 1/16 \;) \end{array}$$

Code: 00, 01, 100, 101, 1100, 1101, 1110, 1111

Code tree:



ADF → 001011101

00|101|1101|

A D F

$$BPC = \sum_{w \in \Omega} X(w) \cdot p(w) = \sum_i l_i \cdot p_i$$

← prob of letter i

↑ length of code for letter i

$$= 1/4 \cdot 2 + 1/4 \cdot 2 + 1/8 \cdot 3 + 1/8 \cdot 3 + 1/16 \cdot 4 + \cdots + 1/16 \cdot 4$$

$$= 2.75$$

☆ 8.33% savings over a fixed code of length 3

$$P_2 = \left(\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}, \tfrac{1}{64}, \ldots, \tfrac{1}{64}\right)$$

$$0 \quad 10 \quad 110 \quad 1110 \quad 111100 \quad \ldots \quad 111111$$

$$BPC = \sum_i l_i \cdot p_i$$

$$= 1 \cdot \tfrac{1}{2} + 2 \cdot \tfrac{1}{4} + 3 \cdot \tfrac{1}{8} + \ldots \quad 6 \cdot \tfrac{1}{64}$$

$$= 2$$

p_i	l_i
$1/2$	1
$1/4$	2
$1/8$	3
$\vdots$	$\vdots$

$$p_i = 1/2^k$$
$$l_i = k$$

$$\Rightarrow$$

$$p_i = 1/2^{l_i}$$
$$2^{l_i} = 1/p_i$$
$$l_i = \log_2 \frac{1}{p_i}$$

$$BPC = \sum_i l_i \cdot p_i = \boxed{\sum_i p_i \cdot \log_2 \frac{1}{p_i}}$$ **Entropy**