Last time

- Expectation

Today

- Applications of Expectation
  - Variance & Standard Deviation
  - Entropy

Next time

- Conditional Prob.
- Bayes Theorem

Variance

Consider heights

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"

3 ways

① $Y = X - E\{X\}$ ✗ neg. & pos. cancel

② $Y = |X - E\{X\}|$ - mean absolute deviation

③ $Y = (X - E\{X\})^2$ - variance

① case 2

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

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

$$= 0$$

②

$$E\{Y\} = |-12''| \cdot 1/3 + |0''| \cdot 1/3 + |+12''| \cdot 1/3$$

$$= 12 \cdot 1/3 + 0 \cdot 1/3 + 12 \cdot 1/3 = 8 \text{ inches}$$

③

$$E\{Y\} = E[(X - E\{X\})^2] = \overset{\text{case 1}}{(-2'')^2 \cdot 1/3 + (0'')^2 \cdot 1/3 + (+2'')^2 \cdot 1/3}$$

$$= 4/3 + 0 + 4/3 = 8/3 \text{ inches}^2$$

variance σ^2

Take square root, get back inches...

$$\sigma^2 = Var(X) = E\{(X - E\{X\})^2\}$$

$$\Rightarrow \quad \sigma = \sqrt{Var(x)} = \sqrt{E[(X - E\{X\})^2]}$$

– back in units originally measured.

e.g.

case 1 : $\sigma^2 = Var(X) = 8/3$ inches^2

$$\sigma = \sqrt{Var(X)} = \sqrt{8/3 \text{ inches}^2} = 1.63''$$

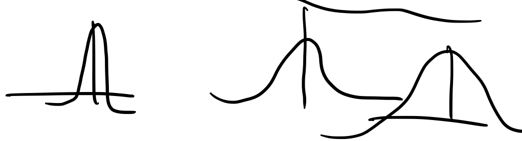
case 3 : ... $\sigma = 19.6''$

Exam

$E\{X\} = 76$

$\sigma = 15$ $\qquad \pm 2/3\,\sigma$ – quartiles

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



Entropy

- Consider 8 letters only $\{A, B, C, D, E, F, G, H\}$
- Need 3-bits

$$\left.\begin{array}{l} A \to 000 \\ B \to 001 \\ \vdots \\ H \to 111 \end{array}\right\} \text{fixed length code}$$

- suppose must encode n- things, need: $\lceil \log_2 n \rceil$
- Efficiency of code is measured in bits-per-character on average, BPC. $BPC = 3$
- variable length code: assign short codes to frequent letters
  long codes to infrequent letters.
- $$BPC = \sum_{w \in \Omega} X(w) \cdot p(w) = \sum_i l_i \cdot p_i$$

l_i: code length for i^{th} letter

p_i: probability of i^{th} letter

Why must we have long codes to make up for short codes?

Analog to PHP: Kraft's Inequality

Example:

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

BBB → 010101

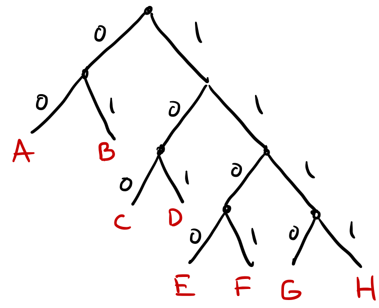
CF → 010101

Suppose we have following skewed distribution

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

(00, 01, 100, 101, 1100, 1101, 1110, 1111)

Code Tree :



ADF → 00101101

00 | 101 | 1101

A D F

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

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

$$= 2.75$$

* 8.33% savings over fixed code of length 3.

	A	B	C	D	E	F	G	H
$P_2 =$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{64}$
	0	10	110	1110	111100	111101	111110	111111

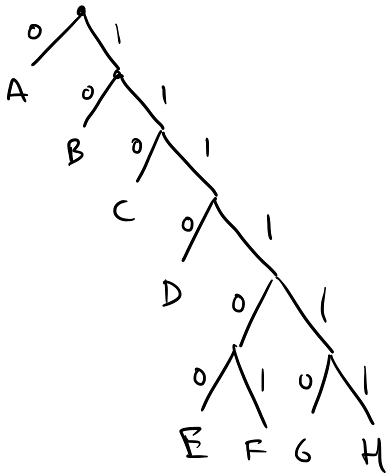
$$P_2 = \left(\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}, \tfrac{1}{64}, \tfrac{1}{64}, \tfrac{1}{64}, \tfrac{1}{64} \right)$$

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

$$= 1 \cdot \tfrac{1}{2} + 2 \cdot \tfrac{1}{4} + 3 \cdot \tfrac{1}{8} + 4 \cdot \tfrac{1}{16}$$

$$+ 6 \cdot \tfrac{1}{64} \cdots + 6 \cdot \tfrac{1}{64}$$

$$= 2$$



p_i	l_i
$\frac{1}{2}$	1
$\frac{1}{4}$	2
$\frac{1}{8}$	3
$\frac{1}{16}$	4
$\frac{1}{64}$	6

$$p_i = \tfrac{1}{2}^k$$
$$l_i = k$$
$$\Rightarrow p_i = \tfrac{1}{2}^{l_i}$$
$$\Leftrightarrow 2^{l_i} = \frac{1}{p_i}$$
$$\Leftrightarrow \boxed{l_i = \log_2 \frac{1}{p_i}}$$

Entropy. —

Q: $p = (p_1, p_2, p_3, \ldots\ )$

$$l_i = \log_2 \frac{1}{p_i}$$

what is expected code length?

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

$$= \boxed{\sum_i p_i \cdot \log_2 \frac{1}{p_i}}$$ entropy

$$H(p) = \sum_i p_i \cdot \log_2 \frac{1}{p_i}$$

ACBADAFGAH

$p_A = \frac{4}{10}$

$p_B = 1/10$

$p_C = 1/10$

$p_D = 1/10$

$\vdots$

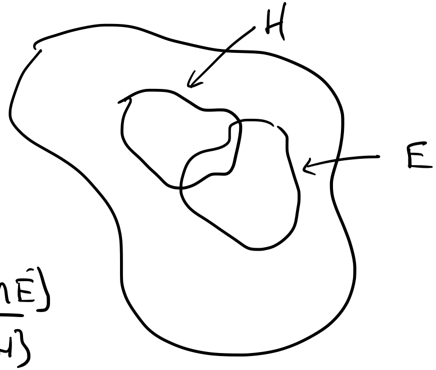
Facts

1. Entropy is the compression limit.
2. Entropy is essential to bound an ability to communicate in presence of noise.
3. Fundamental measure of randomness.
4. Cryptography

Conditional Probability

prob of E given H

$$\Pr\{H \cap E\} = \Pr\{H\} \cdot \underline{\Pr\{E \mid H\}}$$



$$\Rightarrow \Pr\{E \mid H\} = \frac{\Pr\{H \wedge E\}}{\Pr\{H\}}$$

$$= \Pr\{E\} \cdot \Pr\{H \mid E\}$$

$$\Rightarrow \Pr\{H \mid E\} = \frac{\Pr\{H \wedge E\}}{\Pr\{E\}}$$

Bayes Law:

$$\underline{\Pr\{E\}} \cdot \underline{\Pr\{H \mid E\}} = \Pr\{H \wedge E\} = \underline{\Pr\{H\}} \cdot \underline{\Pr\{E \mid H\}}$$

$$\Pr\{H \mid E\} = \frac{\Pr\{E \mid H\} \cdot \Pr\{H\}}{\Pr\{E\}}$$

H = hypothesis: do have Zika?

E = evidence: did blood test come back pos?

CS1800

Discrete Structures

Fall 2017

Lecture 25

11/2/17

Last time

- Finish Expectation
  - variance / standard deviation
- Start Entropy

Today

- Finish Entropy
- Conditional Prob. & Bayes Law

Next time

- Markov Chains & Page Rank

Announcements

Regrades → Kevin Gold

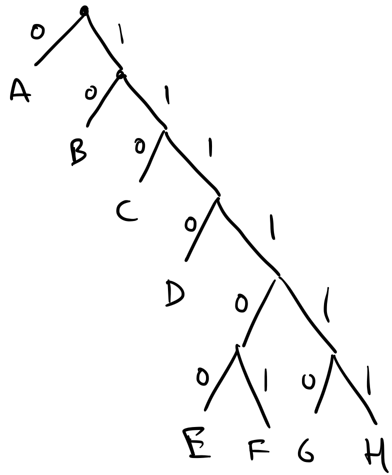
	A	B	C	D	E	F	G	H
$P_2 =$	$(\,1/2,$	$1/4,$	$1/8,$	$1/16,$	$1/64,$	$1/64,$	$1/64,$	$1/64\,)$
	0	10	110	1110	111100	111101	111110	111111

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

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

$$+ 6 \cdot 1/64 \cdots + 6 \cdot 1/64$$

$$= 2$$



p_i	l_i
$1/2$	1
$1/4$	2
$1/8$	3
$1/16$	4
$1/64$	6

$$p_i = 1/2^k$$

$$l_i = k \implies p_i = 1/2^{l_i}$$

$$\iff 2^{l_i} = \frac{1}{p_i}$$

$$\iff \boxed{l_i = \log_2 \frac{1}{p_i}}$$

Entropy.–

Q: $p = (p_1, p_2, p_3, \ldots)$

$$l_i = \log_2 \frac{1}{p_i}$$

What is expected code length?

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

$$= \boxed{\sum_i p_i \cdot \log_2 \frac{1}{p_i}}$$ entropy

$$H(p) = \sum_i p_i \cdot \log_2 \frac{1}{p_i}$$

ACBADAFGAH

$p_A = \frac{4}{10}$

$p_B = 1/10$

$p_C = 1/10$

$p_D = 1/10$

$\vdots$

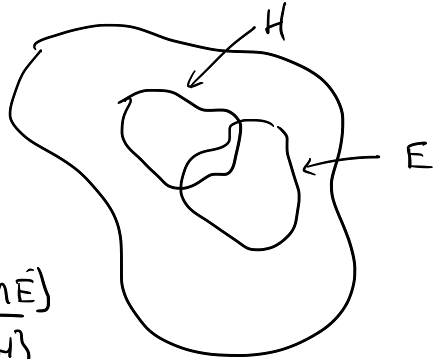
Facts

1. Entropy is the compression limit.
2. Entropy is essential to bound an ability to communicate in presence of noise.
3. Fundamental measure of randomness.
4. Cryptography

Conditional Probability

prob of E given H

$$\Pr\{H \cap E\} = \Pr\{H\} \cdot \underline{\Pr\{E \mid H\}}$$



$$\Rightarrow \Pr\{E \mid H\} = \frac{\Pr\{H \wedge E\}}{\Pr\{H\}}$$

$$= \Pr\{E\} \cdot \Pr\{H \mid E\}$$

$$\Rightarrow \Pr\{H \mid E\} = \frac{\Pr\{H \wedge E\}}{\Pr\{E\}}$$

Bayes Law:

$$\underline{\Pr\{E\}} \cdot \underline{\Pr\{H \mid E\}} = \Pr\{H \wedge E\} = \underline{\Pr\{H\}} \cdot \underline{\Pr\{E \mid H\}}$$

$$\Pr\{H \mid E\} = \frac{\Pr\{E \mid H\} \cdot \Pr\{H\}}{\Pr\{E\}}$$

H = hypothesis: do have Zika?

E = evidence: did blood test come back pos?

CS1800
Discrete Structures
Fall 2017

Lecture 26
11/6/17

Last time

- Finish Entropy
- Start Cond. Prob.
  - Bayes Law

Today

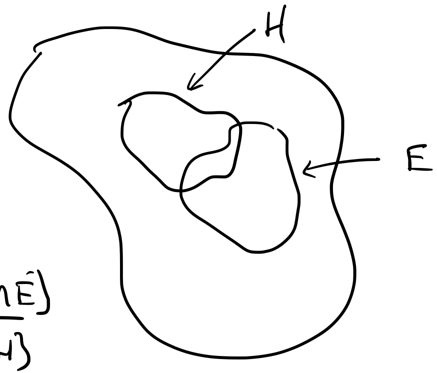
- Finish Cond. Prob. & Bayes Law
- Start Markov Chains & Page Rank

Next time

- Finish MC & PR

- Start Module 4: Algorithms & Analysis

Conditional Probability



prob of E given H

$$\Pr\{H \cap E\} = \Pr\{H\} \cdot \underline{\Pr\{E \mid H\}}$$

$$\Rightarrow \Pr\{E \mid H\} = \frac{\Pr\{H \wedge E\}}{\Pr\{H\}}$$

$$= \Pr\{E\} \cdot \Pr\{H \mid E\}$$

$$\Rightarrow \Pr\{H \mid E\} = \frac{\Pr\{H \wedge E\}}{\Pr\{E\}}$$

Independence

$$\Pr\{E \mid H\} = \Pr\{E\}$$

$$\Pr\{H \mid E\} = \Pr\{H\}$$

$$\Rightarrow \Pr\{H \cap E\} = \Pr\{H\} \cdot \Pr\{E \mid H\} = \Pr\{H\} \cdot \Pr\{E\}$$

Bayes Law:

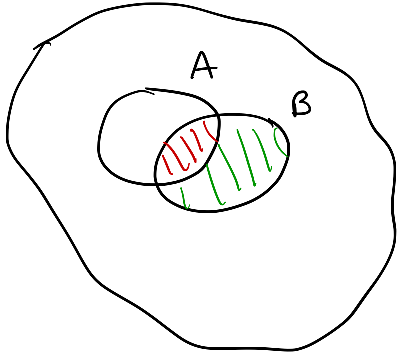
$$\underline{\Pr\{E\}} \cdot \underline{\Pr\{H \mid E\}} = \Pr\{H \wedge E\} = \underline{\Pr\{H\}} \cdot \underline{\Pr\{E \mid H\}}$$

$$\Pr\{H \mid E\} = \frac{\Pr\{E \mid H\} \cdot \Pr\{H\}}{\Pr\{E\}}$$

H = hypothesis: do have Zika?

E = evidence: did blood test come back pos?

Bayes Law



$$\Pr(B|A)\cdot\Pr(A) = \Pr(A\cap B) = \Pr(A|B)\cdot\Pr(B)$$

$$\Pr(A|B) = \frac{\Pr(B|A)\cdot\Pr(A)}{\Pr(B)}$$

$$= \frac{\Pr(B|A)\cdot\Pr(A)}{\Pr\{B|A\}\cdot\Pr(A) + \Pr\{B|\bar{A}\}\cdot\Pr\{\bar{A}\}}$$

what is $\Pr(B)$?

$$p(B) = \quad + $$

$$= p(B|A)\cdot p(A) + p(B|\bar{A})\cdot p(\bar{A})$$

$$\Pr\{H|E\} = \frac{\Pr\{E|H\}\cdot\Pr\{H\}}{\Pr\{E|H\}\cdot\Pr(H\} + \Pr\{E|\bar{H}\}\cdot\Pr(\bar{H}\}}$$

Example: Zika in FL 2016

- Prevalence of Zika in S. FL. $p(\text{Zika}) = 10^{-5}$ (1 in 100,000)
- accuracy of blood test is 99 %

i.e. $p(\text{pos. test} \mid \text{Zika}) = 0.99$ ← test subjects who had Zika

$p(\text{pos. test} \mid \text{no Zika}) = 0.01$ ← control group who don't have Zika

- You test <u>positive</u> : what it is the chance that you have Zika?

$\Rightarrow$ <u>Not</u> $p(\text{pos. test} \mid \text{Zika}) = 0.99$

$\Rightarrow$ you want $p(\text{Zika} \mid \text{pos. test})$!

$$p(\text{zika} \mid \text{pos test}) = \frac{p(\text{pos. test} \mid \text{zika}) \cdot p(\text{zika})}{p(\text{pos test})}$$

$$= \frac{p(\text{pos. test} \mid \text{zika}) \cdot p(\text{zika})}{p(\text{pos test} \mid \text{zika}) \cdot p(\text{zika}) + p(\text{pos test} \mid \text{not zika}) \cdot p(\text{not zika})}$$

$$= \frac{0.99 \times 10^{-5}}{0.99 \times 10^{-5} + 0.01 \times (1 - 10^{-5})}$$

$$= \frac{0.0000099}{0.0000099 + 0.0099999}$$

$$\approx 0.00099$$

$$\approx 0.1\%$$ i.e. only 1 in 1,000 !

Seems wildly counter intuitive, but...

10,000,000 people in FL. 10^7

w/ zika : $10^7 \cdot 10^{-5} = 10^2 = 100$

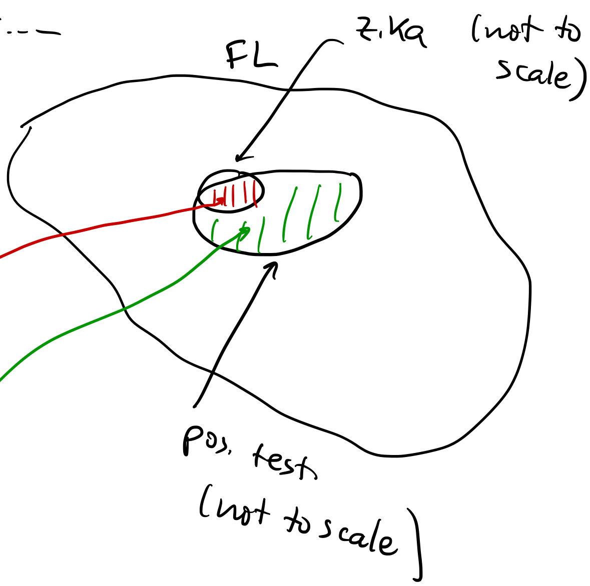
w/o zika : 9,999,900

---

test pos w/ zika : $100 \cdot 0.99 = 99$

test pos. w/o zika : $9{,}999{,}900 \times 0.01$

$= 99{,}999$



∴ If you test pos.,

you're about 1,000 times more likely to be among 99,999 who <u>don't</u> have zika but test pos. than among the 99 who do have zika and test pos.

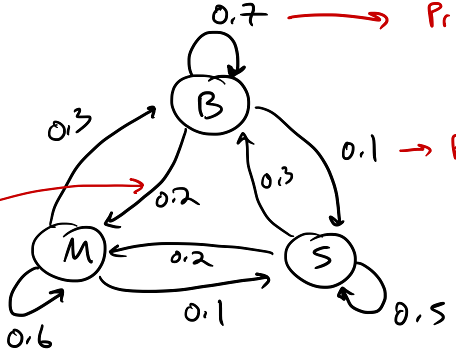
Markov chains

B: Bertucci's

M: Margaritas

S: Sato

Pr{M|B}



Pr{B|B}

Pr{S|B}

3 reds ⇒ must add to 1