6.02
Introduction to EECS II: Digital Communication Systems
Müjdat Çetin
Recitation 01 – Sept. 5, 2013

Announcements

- Sign-up sheet
- Spots available in 11am recitations. Interested? E-mail Eduardo at sverdlin@mit.edu

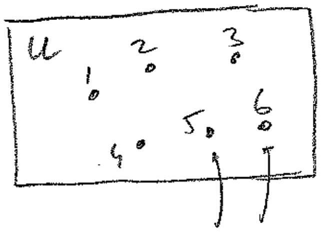
Today

- 6.02 Recitations: Basics, administration, objectives
- Review of some basic probabilistic concepts
- Information
- Entropy
- Variable length binary coding – Huffman coding

- Mujdat Cetin, 32-D616, mcetin@mit.edu, Office hours TBD
  - Visiting professor on sabbatical here
  - Worked at MIT from 2001-2005
  - Taught 6.011 at MIT before
- TA:
  - 10-11 Max Dunitz?
  - 11-12 Elaine Han
- Recitations -- Tue/Thu 10-11 / 11-12
  - will build upon some concepts covered in lectures
  - will go through illustrative examples of concepts covered in lectures
  - attend lectures, reviews (at least) the lecture slides before recitation
  - please be interactive, ask and answer questions
  - please come to my office hours for technical questions or any concerns you might have about this course
- 6.02 overview
  - Intro to EECS through digital communications – very appropriate, because
    -digital communications contains many interesting problems on signals and systems and computer science
    -digital communications is ubiquitous these days!
  - Pieces of the course: bits, signals, packets
    -Bits: information, coding
    -Signals: transmission in a physical channel
    -Packets: network communications
  - Digital comms: Discrete set of messages rather than waveforms
    -remember the Morse alphabet example from lecture
    -Shannon: source emitting i.i.d. symbols
- Start with a review of some basic probabilistic concepts

Probability Model (probabilistic experiment)

Ex/ Roll of a 6-sided die



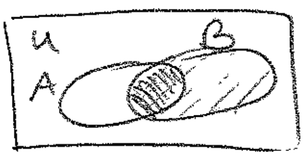
elementary outcomes

- U: universe, sample space
  set of all possible outcomes
  $\{1, 2, 3, 4, 5, 6\}$
- Events: subsets of U.
  Event A has occurred if the outcome of the experiment lies in the set A.
  e.g. – The outcome is ≤ 3
  – The outcome is odd
- Probability measure.
  non-negative, sums to 1.
  Fair die: $p(e_i) = \frac{1}{6}$ $\quad \forall e_i \in U$.
  (e_i: elementary outcome)

Events form a algebra of sets
Consider events A, B

- Union: $A \cup B$ / A or B / $A + B$
- Intersection: $A \cap B$ / A and B / AB

- Conditional probability $P(A|B)$: prob. of event A given that B occurred.



$$P(A|B) = \frac{P(AB)}{P(B)} \quad (*)$$

When do you think two events are independent?

- Independence of two events: whether one event has occurred or not does not change the probability of the other event

$$P(A|B) = P(A)$$

From $(*)$ this is equivalent to $P(AB) = P(A)P(B)$

independent events intersection probabilities multiply.

Intuition

- A: it rains B: I get wet - dependent
- A: it's Monday B: I get wet - independent

EX Roll of a 6-sided die

A: outcome ≤ 2
B: outcome odd } Indep?

$P(A) = \frac{1}{3}$ $P(B) = \frac{1}{2}$ $P(A|B) = \frac{1}{3}$

$P(AB) = \frac{1}{6}$

$P(A|B) = P(A)$
$P(AB) = P(A)P(B)$
Yes, indep

What if A: outcome ≤ 3 $\rightarrow$ A, B not indep.

EX Roll of a 6-sided die

A: outcome odd
B: outcome > 5) clearly dependent..

Note $P(AB) = 0$ — mutually exclusive events.

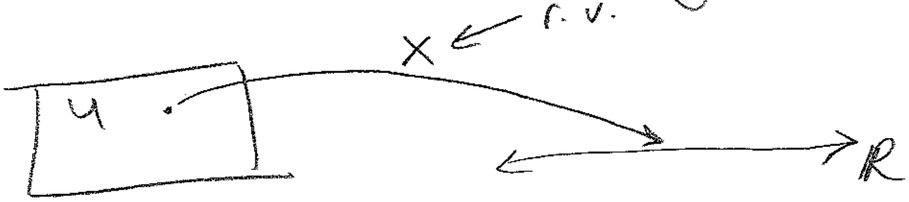
$P(A+B) = \frac{2}{3} = P(A) + P(B)$

mutually exclusive events union: probabilities add

In general, for any two events A, B:

$P(A+B) = P(A) + P(B) - P(AB)$

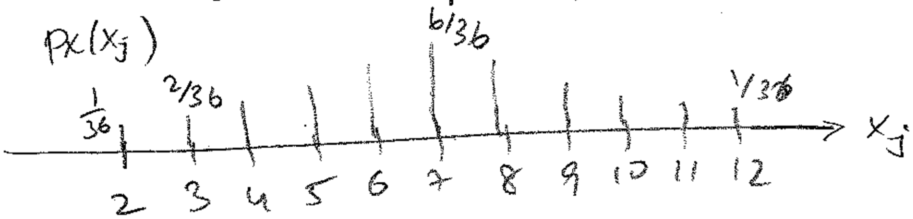
Random variable : A random variable is a function on the outcome of a random experiment that assigns a number to each elementary outcome in the universe.



Ex Roll 2 dice : X : Sum of two numbers showing up.

Can write a probability mass fcn.

$$P(X = x_j) = p_X(x_j)$$



The expected value or mean value of a r.v.

$$E(X) = \sum_j x_j P(X = x_j)$$

For the experiment above $E(X) = 7$

Information

- Shannon: 1948 paper "A Mathematical Theory of Communication"
  1949 book "The Mathematical Theory of Communication"

- Information obtained by being told the outcome s_i of a probabilistic experiment S:

$$I(S = s_i) = \log_2\left(\frac{1}{p_S(s_i)}\right) \quad \text{bits.}$$

Measure of the uncertainty associated with this outcome prior to its being announced.

Less probable outcome/event happens
→ more information revealed

Information as degree of surprise.

EX Deck of 52 cards. – Drawing a card

- How much info. do you get if I tell you it's a face card?

Uncertainty before I tell you anything:

$$\log_2(52) = 5.700 \text{ bits.}$$

Uncertainty after I tell you it's a face card:

4 suits x 3 face cards/suit = 12 face cards

$$\log_2(12) = 3.585 \text{ bits}$$

Info I have provided: $\log_2(52) - \log_2(12)$

$$= \log_2\left(\frac{52}{12}\right) = 2.115 \text{ bits.}$$

- What if I tell you it's a face card & a spade

Note: indep events

$P(\text{face \& spade}) = \frac{3}{52} \Rightarrow \text{Info} = \log_2\left(\frac{52}{3}\right) = 4.115$ bits.

Note: Info. provided by occurrence of two indep. events: add info provided by each. $\underbrace{\log_2\left(\frac{52}{12}\right)}_{\text{face}} + \underbrace{\log_2(4)}_{\text{spades}}$

Entropy

The entropy $H(S)$ of random variable S is the <u>expected</u> information obtained by learning the outcome of S.

$$H(S) = \sum_{i}^{N} p_S(s_i)\, I(S=s_i) = \sum_{i}^{N} p_S(s_i) \log_2\left(\frac{1}{p_S(s_i)}\right) \text{ bits}$$

- Measure of uncertainty of the random variable
  - Maximal uncertainty when $p_S(s_i) = \frac{1}{N}$ $\forall i$.
    $\rightarrow H(S) = \log_2 N$.
- Average amount of information that must be delivered to resolve the uncertainty of a random variable (source)
- Lower bound on the no. of binary digits needed to encode messages generated by a probabilistic source

Higher entropy → more difficult to compress.

EX Image coding – Consider images with 4 graylevels:
"B" – background
F_1, F_2, F_3 – foreground pixel intensities.

Simple binary coding:

symbol	codeword
B	00
F_1	01
F_2	10
F_3	11

We use 2 binary digits per symbol (pixel).

- Can we do better? (use smaller-length codewords on average?) Yes, if we exploit any knowledge about the source.
- "Most pixels in images are background pixels".
- Let's say we are told that with prob. 0.8 we have a background pixel, and different foreground intensities are equiprobable.

Intuition: use shorter codewords for more frequent symbols.

Consider

symbol	codeword
B	0
F_1	10
F_2	110
F_3	111

variable-length coding

Average codeword length:

$$0.8 \times 1 + \frac{0.2}{3} \times 2 + \frac{0.2}{3} \times 3 + \frac{0.2}{3} \times 3 = 1.333 \text{ binary digits.}$$

So, we can do better than simple binary coding!

Entropy provides lower bound on the no. of binary digits to be used on average to encode messages from a particular source. In this case:

$$H = 0.8 \log_2\left(\frac{1}{0.8}\right) + 3 \cdot \frac{0.2}{3} \log_2\left(\frac{3}{0.2}\right) = 1.0389 \text{ bits}$$

How can we design the minimum average length code for a particular source?