COMP2610 / COMP6261 Information Theory

Lecture 2: First Steps and Basic Probability

Australian
National
University

July 24, 2018

Outline

A General Communication System

Source : The information source that generat communicated

Encoder : Operates on the message to produce transmission

Channel : The medium used to transmit the signal

Decoder : Reconstructs the message from the signal

Destination : The entity that receives the message

Communication over Noisy Channels

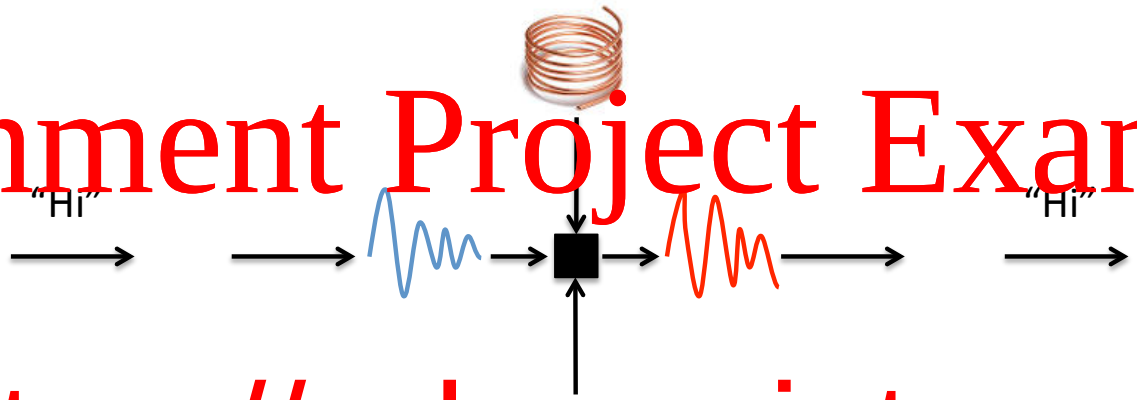
A channel is some medium for transmitting messages.

A noisy ch
sender's

The Prob

"The fundamental problem of communication is th
point either exactly or approximately a message se
(Claude Shannon, 1948)

Example: Telephone Network



Source : Aditya

Encoder : Telephone handset

Channel : Analogue telephone line

Decoder : Telephone handset

Destination : Mark

Examples of Noisy Communication Channels

Other examples of noisy channels:

- A radio communication link
- VDSL NBN connection
- The inte
- Reproducing cells
- A magnetic hard disk drive
  - Channel does not need to involve physical movement

What would the other components be for each of these channels?

Uncertainty in Communication

We cannot avoid uncertainty when

1. Dealing with noise (imperfections) in the channel

2. "Compressing" the messages (compare a h
of a manuscript with the typed text that capture
transcript of a spoken utterance)

Channel Noise

A noisy ch

Thus, rec

How to model, quantify, and mitigate this uncertain

Message Compression – I

Cover and Thomas, Example 1.1.2

- Suppose we'd like to relay the outcome of an 8 horse race to a friend
  - We wish to convey one of `A, B, ..., H`
- Sup
sch

$$\begin{aligned} \texttt{B} &\rightarrow \texttt{0} \\ \texttt{C} &\rightarrow \texttt{0} \\ &\vdots \\ \texttt{H} &\rightarrow \texttt{111} \end{aligned}$$

Message Compression – II

Cover and Thomas, Example 1.1.2

- Now say the probabilities of the horses winning are $(1/2, 1/4, 1/8, 1/16, 1/64, 1/64, 1/64, 1/64)$
- Encoding messages based on their probability of the being chosen will g

$$\begin{aligned}
\mathtt{C} &\to \mathtt{110} \\
\mathtt{D} &\to \mathtt{111} \\
\mathtt{E} &\to \mathtt{111} \\
\mathtt{F} &\to \mathtt{111100} \\
\mathtt{G} &\to \mathtt{111101} \\
\mathtt{H} &\to \mathtt{111111}
\end{aligned}$$

What is “Information”?

For noise correction and message compression, we will need to quantify the information contained in a message

Roughly

- Un

- The receiver’s uncertainty is reduced on seei

The Case for Probability

We run into the notion of uncertainty when trying to pin down:

1. How to deal with machine noise
2. Ho
3. Wh

To make progress, we need to formalise uncertain

We will do this using probability theory

We now commence our *review* of probability; this will be hard going if you have not met it before!

Probability: Example

Quantification and Manipulation of Uncertainty (Bishop, PRML, 2006)

Probability: Example

Quantification and Manipulation of Uncertainty (Bishop, PRML, 2006)

1. Pick a box at random

Probability: Example

Quantification and Manipulation of Uncertainty (Bishop, PRML, 2006)

1. Pick a box at random
2. From the selected box, pick a fruit (apple or orange) at random

Probability: Example

Quantification and Manipulation of Uncertainty (Bishop, PRML, 2006)

1. Pick a box at random
2. From the selected box, pick a fruit (apple or orange) at random
3. Ob

Probability: Example

Quantification and Manipulation of Uncertainty (Bishop, PRML, 2006) — Cont'd

- Identity of the box is a random variable $\in \{\ ,\ \}$
- Identity of the fruit is a random variable $F \in \{a, o\}$

Probability: Example

Quantification and Manipulation of Uncertainty (Bishop, PRML, 2006) — Cont'd

- Identity of the box is a random variable $\in \{\ ,\ \}$
- Identity of the fruit is a random variable $F \in \{a, o\}$

Probability of an event: Proportion of times it happens out of a large number of trials

Probability: Example

Suppose we repeat the selection process many tim
picking up the blue box 60% of the time and the red box 40% of the time

- $p(B = r) = 0.4$, $p(B = b) = 0.6$

Probability: Basic Properties

By definition, $0 \leq p(B = b) \leq 1$ and $0 \leq p(B = r) \leq 1$

Outcomes are mutually exclusive:

Outcomes are jointly exhaustive:

$$\begin{aligned} p(B = r \text{ OR } B = b) &= p(B = r) + p(B \qquad b) \\ &= p(B = r) + p(B = b) \\ &= 1 \end{aligned}$$

- What is the probability of picking the red box, and an apple within that box?

- Wh

- Given that we selected a red box, what is the pro an apple?

We can answer these and more complex questions by using the *rules of probability*.

Joint Probability

What is the probability of selecting the red box and selecting an apple?

Joint Probability

What is the probability of selecting the red box and selecting an apple?

Joint Probability of a Set of Events

The proportion of times these events happened together out of the total number o

Joint Probability

What is the probability of selecting the red box and selecting an apple?

Joint Probability of a Set of Events

The proportion of times these events happened together out of the total number o

If we repea of the trials we sa

$$p(B = r \text{ AND } F = a) = p$$

$$= \frac{\quad}{1}$$

$$= 0.1$$

Marginal Probability

What is the probability of an apple being picked, regardless of the box we selected?

Marginal Probability

What is the probability of an apple being picked, regardless of the box we selected?

Marginal Probability of an Event

The proportion of times that this event happened out of the total number of trials.

Marginal Probability

What is the probability of an apple being picked, regardless of the box we selected?

Marginal Probability of an Event

The proportion of times that this event happened out of the total number of trials.

Remem

Say that in 45 of the trials, we selected a blue box and an a

So, irrespective of B, an apple was selected 45 + 1

$$p(F = a) = \frac{55}{100} = \frac{11}{20}$$

Conditional Probability

What is the probability of an apple being picked up, given that a red box was selected?

Conditional Probability

What is the probability of an apple being picked up, given that a red box was selected?

Conditional Probability of an Event

The conditional probability of an event X with respect to an event Y is the proportio pens.

Conditional Probability

What is the probability of an apple being picked up, given that a red box was selected?

Conditional Probability of an Event

The conditional probability of an event X with respect to an event Y is the proportio pens.

The trials apple wa

We selected a red box and an apple 10 out of 100 times

We selected a red box (regardless of the fruit) 40 out of 1

$$p(F = a \,\texttt{GIVEN}\, B = r) = p(F = a|B = r) = \frac{10}{40} = \frac{1}{4}$$

Conditional Probability

What is the probability of an apple being picked up, given that a red box was selected?

Conditional Probability of an Event

The conditional probability of an event X with respect to an event Y is the proportio pens.

The trials apple wa

We selected a red box and an apple 10 out of 100 times

We selected a red box (regardless of the fruit) 40 out of 1

$$p(F = a \,\texttt{GIVEN}\, B = r) = p(F = a|B = r) = \frac{10}{40} = \frac{1}{4}$$

Can we write this in terms of the joint and marginal probabilities?

Joint, Marginal and Conditional Probabilities:

A More General Formulation (1)

Consider the more general case of two random variables:

$X \in \{x_1, x_2, \ldots, x_M\}$ and $Y \in \{y_1, y_2, \ldots, y_L\}$

N : Total number of trials

Joint, Marginal and Conditional Probabilities:

A More General Formulation (1)

Consider the more general case of two random variables:

$X \in \{x_1, x_2, \ldots, x_M\}$ and $Y \in \{y_1, y_2, \ldots, y_L\}$

N : Total number of trials

n_{ij} : $\sharp(X = x_i, Y = y_j)$ = $\sharp$ of times that x_i and y_i happen

Joint, Marginal and Conditional Probabilities:

A More General Formulation (1)

Consider the more general case of two random variables:

$X \in \{x_1, x_2, \ldots, x_M\}$ and $Y \in \{y_1, y_2, \ldots, y_L\}$

N : Total number of trials

n_{ij} : $\sharp(X = x_i, Y = y_j)$ = $\sharp$ of times that x_i and y_i happen

c_i : $\sharp(X = x_i) = \sum_j n_{ij}$ = $\sharp$ of times that x_i happens

Joint, Marginal and Conditional Probabilities:

A More General Formulation (1)

Consider the more general case of two random variables:

$X \in \{x_1, x_2, \ldots, x_M\}$ and $Y \in \{y_1, y_2, \ldots, y_L\}$

N : Total number of trials

n_{ij} : $\sharp(X = x_i, Y = y_j)$ = $\sharp$ of times that x_i and y_i happen

c_i : $\sharp(X = x_i) = \sum_j n_{ij}$ = $\sharp$ of times that x_i happens

r_j : $\sharp(Y = y_j) = \sum_i n_{ij}$ = $\sharp$ of times that y_j happens

Joint, Marginal and Conditional Probabilities:

A More General Formulation (2)

By definiti

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$

$$p(X = x_i) = \frac{c_i}{N}$$

$$p(Y = y_j) = \frac{r_j}{N} \text{ (Marginal)}$$

$$p(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i} \text{ (Conditional)}$$

Joint, Marginal and Conditional Probabilities:

A More General Formulation (1)

Bins and fr

Red	30	

Verify the computations from previous section wit

Joint, Marginal and Conditional Probabilities:

A More General Formulation (3)

Observe

$$p(X = x_i) = \frac{\sum_j n_{ij}}{N}$$

$$= \sum_j p(X$$

$$p(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i} = \frac{n_{ij}}{N} \Big/ \frac{c_i}{N}$$

$$= p(X = x_i, Y = y_j)/p(X = x_i)$$

The Rules of Probability

Sum Rule / Marginalization :

$$p(X = x_i) = \sum_j p(X = x_i, Y = y_j)$$

The Rules of Probability

Sum Rule / Marginalization :

$$p(X = x_i) = \sum_j p(X = x_i, Y = y_j)$$

Product R

$$= x_i)$$

The Rules of Probability

Sum Rule / Marginalization :

$$p(X = x_i) = \sum_j p(X = x_i, Y = y_j)$$

Product R

$$= x_i)$$

and by symmetry:

$$p(X = y_j, X = x_i) = p(X =$$

The Rules of Probability

Sum Rule / Marginalization :

$$p(X = x_i) = \sum_j p(X = x_i, Y = y_j)$$

Product R

$$= x_i)$$

and by symmetry:

$$P(Y = y_j, X = x_i) = p(X =$$

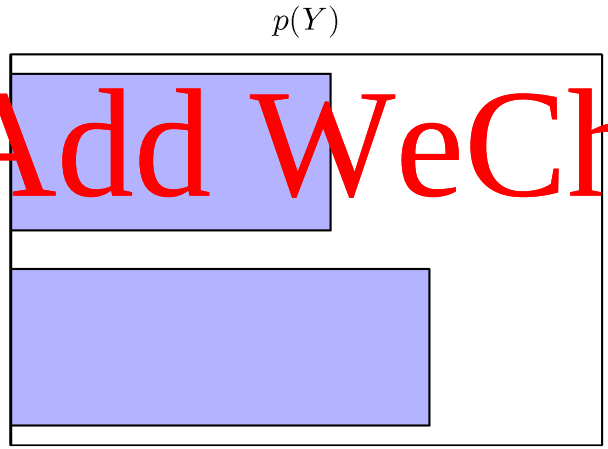
Therefore:

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j) = \sum_j P(X = x_i | Y = y_j) P(Y = y_j)$$

An Illustration of a Distribution over Two Variables

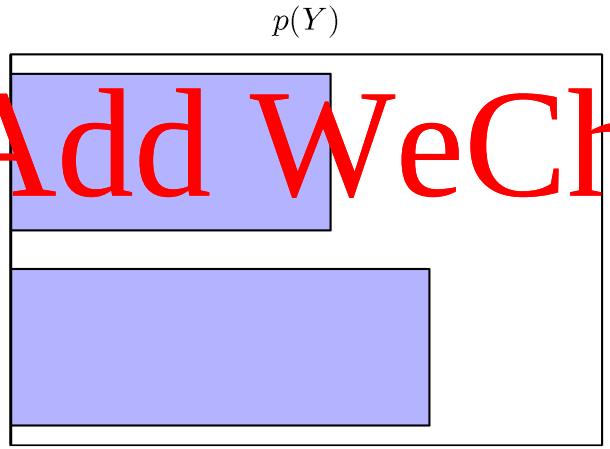
An Illustration of a Distribution over Two Variables

An Illustration of a Distribution over Two Variables



marginal

An Illustration of a Distribution over Two Variables



marginal

conditional

Joint, Marginal and Conditional Probabilities:

An even More General Formulation

Given D random variables $X_1, \ldots, X_D$

Joint, Marginal and Conditional Probabilities:

An even More General Formulation

Given D random variables $X_1, \ldots, X_D$

Chain Ru

$$p(X_1, X_2) = p(X_1)p(X_2|X_1)$$

Joint, Marginal and Conditional Probabilities:

An even More General Formulation

Given D random variables $X_1, \ldots, X_D$

Chain Ru

$$p(X_1, X_2) = p(X_1)p(X_2|X_1)$$

$$p(X_1, X_2, X_3) = p(X_1, X_2)p(X_3|X_1, X_2)$$

$$, X_2)$$

Joint, Marginal and Conditional Probabilities:

An even More General Formulation

Given D random variables $X_1, \ldots, X_D$

Chain Ru

$$p(X_1, X_2) = p(X_1)p(X_2|X_1)$$

$$p(X_1, X_2, X_3) = p(X_1, X_2)p(X_3|X_1, X_2)$$

$$p(X_1, \ldots, X_D) = p(X_1)p(X_2|X_1)p(X_3|X_2, X_1)\ldots p(X_D|X_1, \ldots, X_{D-1})$$

Summary

- General architecture for communication systems

- Wh

- Probability theory: joint, marginal and condit

- Reading: Mackay § 2.1 and § 2.2; Bishop § 1.2

Exercise

Coming Back to our Original Example

Given: $p(B = r) = 0.4$, $p(B = b) = 0.6$

Assume the fruit are selected uniformly from each box

- Write down all conditional probabilities $P(F|B)$
- Evaluate the overall probabilities $P(F)$

Next time

- More on joint, marginal and conditional distributions

- Wh

- What, if anything, does $p(X = x|Y = y) = p(Y = y|X = x)$?

Next time

- Sign-up for tutorials open at 9am wednesday 25 July. Will offer **three** tuto

- Class rep.