

# Lecture 5: Information Channels

## Information Theory

Willard Þór Rafnsson

Reykjavík University, Iceland

09. February, 2009

- Introduction
- Problem Setting
- Role of Redundancy
- Main Result
- Preliminaries
  - Conditional Probability
  - Bayes' Formula
- Information Channels
  - Definition
  - Examples
- Combining Channels
- Epilogue

# Outline

## 1 Introduction

- Problem Setting
- Role of Redundancy
- Main Result

## 2 Preliminaries

- Conditional Probability
- Bayes' Formula

## 3 Information Channels

- Definition
- Examples

## 4 Combining Channels

## 5 Epilogue

### Lecture 5: Information Channels

#### Introduction

Problem Setting  
Role of  
Redundancy  
Main Result

#### Preliminaries

Conditional  
Probability  
Bayes' Formula

#### Information Channels

Definition  
Examples

#### Combining Channels

#### Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

**Problem Setting**

Role of  
Redundancy  
Main Result

Preliminaries

Conditional  
Probability  
Bayes' Formula

Information  
Channels

Definition  
Examples

Combining  
Channels

Epilogue

# Shannon's Goal

Recall Shannon's goals.

*Find the fundamental limits of i) compressing and ii) reliably communicating data.*

So far,

- We have been focusing on *i*).

From now on,

- We will start looking at *ii*)

## Lecture 5: Information Channels

Introduction

**Problem Setting**

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition  
Examples

Combining  
Channels

Epilogue

# Source Coding Problem Setting, Recap

Recall the scenario of transmitting source symbols.

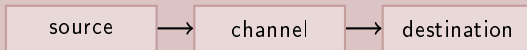


Figure: Communication System

## Lecture 5: Information Channels

Introduction

**Problem Setting**

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Source Coding Problem Setting, Recap

Recall the scenario of transmitting source symbols.

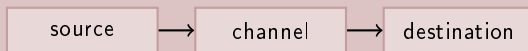


Figure: Communication System

If  $r < q$ , we needed to *encode*  $s_i$  to  $\mathbf{t} = w_i$ .

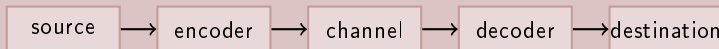


Figure: Communication System with Encoder & Decoder

## Lecture 5: Information Channels

Introduction

**Problem Setting**

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

Bayes' Formula

Information

**Channels**

Definition

Examples

Combining

Channels

Epilogue

# Source Coding Problem Setting, Recap

Recall the scenario of transmitting source symbols.

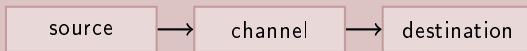


Figure: Communication System

If  $r < q$ , we needed to *encode*  $s_i$  to  $t = w_i$ .

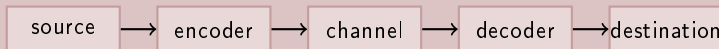


Figure: Communication System with Encoder & Decoder

In this setting, we could get  $L(\mathcal{C})$  arbitrarily close to  $H(\mathcal{S})$ ...



Figure: Communication System with Source Extension

# Source Coding Problem Setting, Recap

Recall the scenario of transmitting source symbols.

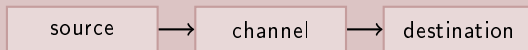


Figure: Communication System

If  $r < q$ , we needed to *encode*  $s_i$  to  $\mathbf{t} = w_i$ .

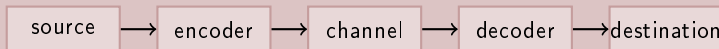


Figure: Communication System with Encoder & Decoder

In this setting, we could get  $L(\mathcal{C})$  arbitrarily close to  $H(\mathcal{S})$ ...



Figure: Communication System with Source Extension

... **provided** the channel **just transfers symbols!**



# Channel Coding Problem Setting

But what if the channel does **more**? Introduces **errors**? Is **unreliable**?

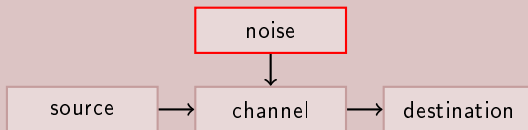


Figure: Communication System

## Lecture 5: Information Channels

Introduction

**Problem Setting**

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information

Channels

Definition

Examples

Combining

Channels

Epilogue

# Channel Coding Problem Setting

But what if the channel does **more**? Introduces **errors**? Is **unreliable**?

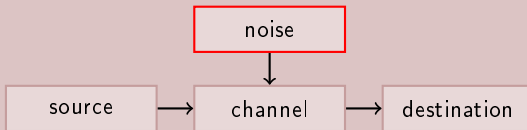


Figure: Communication System

This is a *common problem* in practice.

- Space-probe, sending data over long distances w. a weak power supply; messages get distorted.
- Talking through a crowd.
- Making data on CDs resistant to a few scratches.
- Mobile phone networks, the internet, ...
- Brain-Computer-Interfacing; making sense of spikes from several neighbouring neurons in the motor cortex.

# Outline

## 1 Introduction

- Problem Setting
- **Role of Redundancy**
- Main Result

## 2 Preliminaries

- Conditional Probability
- Bayes' Formula

## 3 Information Channels

- Definition
- Examples

## 4 Combining Channels

## 5 Epilogue

### Lecture 5: Information Channels

Introduction

Problem Setting

**Role of  
Redundancy**

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# We Cannot Eliminate Redundancy

## Example (Biased Die, with errors)

### Lecture 5: Information Channels

Introduction

Problem Setting

**Role of  
Redundancy**

Main Result

Preliminaries

Conditional

Probability

Bayes' Formula

Information

Channels

Definition

Examples

Combining

Channels

Epilogue

## Example (Biased Die, with errors)

$$S = (6, 5, 4, 3, 2, 1),$$

$$\mathcal{C} = (0, 101, 1000, 1001, 110, 111).$$

Any (infinite) sequence of bits *uniquely decodable*. For instance,

$$\mathcal{C}^{*-1}(10101010101 \dots) = 56565 \dots$$

But what if our channel is **unreliable**? If we maliciously flip just 1 bit, we don't just get an invalid code symbol sequence; we get a *different code word sequence*!

$$\mathcal{C}^{*-1}(10001010101 \dots) = 4565 \dots$$

This is the *price of eliminating redundancy*.

Introduction
Problem Setting
Role of Redundancy
Main Result
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

## Example (Biased Die, with errors)

$$S = (6, 5, 4, 3, 2, 1),$$

$$\mathcal{C} = (0, 101, 1000, 1001, 110, 111).$$

Any (infinite) sequence of bits *uniquely decodable*. For instance,

$$\mathcal{C}^{*-1}(10101010101 \dots) = 56565 \dots$$

But what if our channel is **unreliable**? If we maliciously flip just 1 bit, we don't just get an invalid code symbol sequence; we get a *different code word sequence*!

$$\mathcal{C}^{*-1}(10001010101 \dots) = 4565 \dots$$

This is the *price of eliminating redundancy*.

## Example (Biased Die, block code)

$$\mathcal{C} = (110, 101, 100, 011, 010, 001),$$

we can ask for a “re-transmission” if ever we see 000 or 111.

Introduction
Problem Setting
<b>Role of Redundancy</b>
Main Result
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# Outline

## 1 Introduction

- Problem Setting
- Role of Redundancy
- **Main Result**

## 2 Preliminaries

- Conditional Probability
- Bayes' Formula

## 3 Information Channels

- Definition
- Examples

## 4 Combining Channels

## 5 Epilogue

### Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

**Main Result**

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

**Main Result**

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue



# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

**Main Result**

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ .

### Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ . Say on average,  $\frac{1}{3}$  of each code symbol gets distorted.

### Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ . Say on average,  $\frac{1}{3}$  of each code symbol gets distorted. Then we can *send each word 3 times* to achieve better error correction:  $\mathcal{C}' : h \mapsto 000, t \mapsto 111$ .

## Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

## How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

### Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ . Say on average,  $\frac{1}{3}$  of each code symbol gets distorted. Then we can *send each word 3 times* to achieve better error correction:  $\mathcal{C}' : h \mapsto 000, t \mapsto 111$ . Pick *majority* symbol as original code word. So, 010 is assumed to be 0, that is, *heads*.

#### Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ . Say on average,  $\frac{1}{3}$  of each code symbol gets distorted. Then we can *send each word 3 times* to achieve better error correction:  $\mathcal{C}' : h \mapsto 000, t \mapsto 111$ . Pick *majority* symbol as original code word. So, 010 is assumed to be 0, that is, *heads*.

Now we can *always* detect an error, *provided* no more than  $\frac{1}{3}$  of digits in a word gets distorted. To improve correctness, we increase  $n$ : nr. of re-sends.

## Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ . Say on average,  $\frac{1}{3}$  of each code symbol gets distorted. Then we can *send each word 3 times* to achieve better error correction:  $\mathcal{C}' : h \mapsto 000, t \mapsto 111$ . Pick *majority* symbol as original code word. So, 010 is assumed to be 0, that is, *heads*.

Now we can *always* detect an error, *provided* no more than  $\frac{1}{3}$  of digits in a word gets distorted. To improve correctness, we increase  $n$ : nr. of re-sends.

- **Problem:** Naïve approach ruins transmission rate.  
 $R \rightarrow 0$  as  $n \rightarrow \infty$ .

## Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# How Good Error-Correction Can We Achieve?

- Our block code could only *sometimes* detect a single bit of error per code word.
- **Naïve approach:** Detect more errors by using *repetition codes*.

## Example (Coin Toss)

We have a code  $\mathcal{C} : h \mapsto 0, t \mapsto 1$ . Say on average,  $\frac{1}{3}$  of each code symbol gets distorted. Then we can *send each word 3 times* to achieve better error correction:  $\mathcal{C}' : h \mapsto 000, t \mapsto 111$ . Pick *majority* symbol as original code word. So, 010 is assumed to be 0, that is, *heads*.

Now we can *always* detect an error, *provided* no more than  $\frac{1}{3}$  of digits in a word gets distorted. To improve correctness, we increase  $n$ : nr. of re-sends.

- **Problem:** Naïve approach ruins transmission rate.  
 $R \rightarrow 0$  as  $n \rightarrow \infty$ .
- **Shannon's Fundamental Theorem:** There are techniques where  
 $R \rightarrow C$  as  $n \rightarrow \infty$ . ( $C > 0$ )

## Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
<b>Main Result</b>
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue



# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

**Preliminaries**

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

**Conditional  
Probability**

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Probability, Recap

Recall the basic definitions of probability:

## Definition (Sample Space, Event, Probability Measure)

Let  $S$  be a *sample space*<sup>a</sup>, and  $A \subseteq S$  an *event*.  $\mathcal{P}(S)$  is the set of all events of  $S$ . A *probability measure* is a function  $\Pr : \mathcal{P}(S) \rightarrow [0, 1]$  mapping events to probabilities, satisfying  $\Pr(S) = 1$ <sup>b</sup>, and for all disjoint events  $A_i \subseteq S$  it holds that

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i).$$

<sup>a</sup>All possible outcomes of a random experiment.

<sup>b</sup>The *certain event*.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

**Conditional  
Probability**

Bayes' Formula

Information  
Channels

Definition  
Examples

Combining  
Channels

Epilogue

# Probability, Recap

Recall the basic definitions of probability:

## Definition (Sample Space, Event, Probability Measure)

Let  $S$  be a *sample space*<sup>a</sup>, and  $A \subseteq S$  an *event*.  $\mathcal{P}(S)$  is the set of all events of  $S$ . A *probability measure* is a function  $\Pr : \mathcal{P}(S) \rightarrow [0, 1]$  mapping events to probabilities, satisfying  $\Pr(S) = 1$ <sup>b</sup>, and for all disjoint events  $A_i \subseteq S$  it holds that

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i).$$

<sup>a</sup>All possible outcomes of a random experiment.

<sup>b</sup>The *certain event*.

It is helpful to consider  $S$  a rectangular area of size  $1 \times 1$ . Then  $A$  is a part of this area. For instance, for  $S = \{1, 2, \dots, 6\}$ , with  $\Pr(\{i\}) = \frac{1}{6}$  for each  $i \in S$ , if  $A = \{2\}$ , then  $\Pr(A)$  is  $\frac{1}{6}$ th of the image to the right.



Figure: Sample Space  $S$  of 1d6

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

**Conditional  
Probability**

Bayes' Formula

Information

Channels

Definition

Examples

Combining

Channels

Epilogue

# Conditional Probability

## Definition (Conditional Probability)

Let  $B$  be an event s.t.  $\Pr(B) > 0$ . For any  $A$ , the *conditional probability* of  $A$  given  $B$  is

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

## Lecture 5: Information Channels

[Introduction](#)[Problem Setting](#)[Role of  
Redundancy](#)[Main Result](#)[Preliminaries](#)[Conditional  
Probability](#)[Bayes' Formula](#)[Information  
Channels](#)[Definition](#)[Examples](#)[Combining  
Channels](#)[Epilogue](#)

# Conditional Probability

## Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
Main Result
Preliminaries
<b>Conditional Probability</b>
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

### Definition (Conditional Probability)

Let  $B$  be an event s.t.  $\Pr(B) > 0$ . For any  $A$ , the *conditional probability* of  $A$  given  $B$  is

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

Let  $B$  contain even numbers from 1 to 6. That is,  $B = \{2, 4, 6\}$ . Give that *we know something* about an event (that it is even), the probability of  $\Pr(A)$ , for  $A = \{2\}$ , changes.

It is helpful to consider  $B$  a *new sample space*. Then it is easier to see that  $\Pr(A | B) = \frac{1}{3}$ . Also, that  $\Pr(A' | B) = 0$  for  $A' = \{3\}$ . (since  $A' \cap B = \emptyset$ )

1	3	5
2	4	6

Figure: S of 1d6

2	4	6
---	---	---

Figure: B of 1d6

# Independent Events

## Definition (Independence)

Let  $A, B \subseteq S$  be events. If

$$\Pr(A \cap B) = \Pr(A) \Pr(B),$$

then  $A, B$  are said to be *independent*.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

**Conditional  
Probability**

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Independent Events

## Definition (Independence)

Let  $A, B \subseteq S$  be events. If

$$\Pr(A \cap B) = \Pr(A) \Pr(B),$$

then  $A, B$  are said to be *independent*.

Note that

$$\begin{aligned} \Pr(A \cap B) &= \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A) \Pr(B)}{\Pr(B)} \\ &= \Pr(A). \end{aligned}$$

Interpretation:  $A$ 's portion of  $S$  is *the same* as  $A$ 's portion of  $B$ . So when we focus only on  $B = \{2, 4, 6\}$  when we wish to know  $\Pr(A)$  for  $A = \{1, 2\}$ , we are none the wiser<sup>a</sup>.

---

<sup>a</sup>we know the result is not 1, but the chance of getting a desired result is still the same

1	3	5
2	4	6

Figure:  $S$  of 1d6

2	4	6
---	---	---

Figure:  $B$  of 1d6

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

**Conditional  
Probability**

Bayes' Formula

Information  
Channels

Definition  
Examples

Combining  
Channels

Epilogue



# Independence of Multiple Samplings

**Recall:** Sampling  $S = \{1, \dots, 6\}$  twice. That is, sampling  $S^2$ .

If the dice are independent,  $S^2$  can be illustrated this way  $\longrightarrow$

Let  $A = \text{"Get 6 in 2nd throw"}$ .

You can get 36 different results:  $(1, 1), (1, 2), \dots, (6, 6)$ . 6 of these are 6 in the 2nd throw.

$$\text{So } \Pr(A) = \frac{6}{36} = \frac{1}{6}.$$

If you get a 4 in the 1st throw, only  $\frac{1}{6}$  of the sample space, and the potential 6's, remain.

$$\text{So } \Pr(A) = \frac{1}{6}.$$

1	2	1	2	1	2
3	1	3	2	3	3
5	6	5	6	5	6
1	2	1	2	1	2
3	4	3	5	3	6
5	6	5	6	5	6

Figure:  $S^2$  of 2d6; red letters: 1st result.

1	2
3	4
5	6

Figure:  $S^2$  of 2d6, given 1st result was 4

## Lecture 5: Information Channels

Introduction
Problem Setting
Role of Redundancy
Main Result
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

# Conditional Probability As A Matrix

For two independent, *unbiased* dice<sup>1</sup>:

$$M = P_{i,j} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

If we *replace* the die with a biased one in the second throw when we roll less than 4,

$$M = P_{i,j} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/2 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/2 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/2 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \end{pmatrix}$$

Note, the sum of each **row** is 1. To get a *joint probability* matrix, we scale each row with the probability associated with it.

<sup>1</sup>i: 1st throw, j: 2nd throw

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

**Bayes' Formula**

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Law of Total Probability

Assume  $S$  is (countably) infinite<sup>2</sup>.

## Definition (Partition)

$\{B_i\}_{i \in \mathbb{N}}$  is a partition of  $S$  if

- a)  $\forall i \in \mathbb{N}. \Pr(B_i) > 0$
- b)  $\forall i, j \in \mathbb{N}. i \neq j \implies B_i \cap B_j = \emptyset$
- c)  $S = \bigcup_{i=1}^{\infty} B_i$

## Theorem (Law of Total Probability)

Let  $\{B_i\}_{i \in \mathbb{N}}$  be a partition of  $S$ . Then

$$\forall A \subseteq S. \Pr(A) = \sum_{i=1}^{\infty} \Pr(A \mid B_i) \Pr(B_i)$$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

**Bayes' Formula**

Information  
Channels

Definition  
Examples

Combining  
Channels

Epilogue

<sup>2</sup>The results we see are easily specialised to the finite case

# Law of Total Probability

Assume  $S$  is (countably) infinite<sup>2</sup>.

## Definition (Partition)

$\{B_i\}_{i \in \mathbb{N}}$  is a partition of  $S$  if

- a)  $\forall i \in \mathbb{N}. \Pr(B_i) > 0$
- b)  $\forall i, j \in \mathbb{N}. i \neq j \implies B_i \cap B_j = \emptyset$
- c)  $S = \bigcup_{i=1}^{\infty} B_i$

## Theorem (Law of Total Probability)

Let  $\{B_i\}_{i \in \mathbb{N}}$  be a partition of  $S$ . Then

$$\forall A \subseteq S. \Pr(A) = \sum_{i=1}^{\infty} \Pr(A \mid B_i) \Pr(B_i)$$

## Proof.

$A = A \cap S = \bigcup_{i=1}^{\infty} (A \cap B_i)$ .  $\{(A \cap B_i)\}_{i \in \mathbb{N}}$  is a sequence of pairwise disjoint sets. Thus  $\Pr(A) = \sum_{i=1}^{\infty} \Pr(A \cap B_i) = \sum_{i=1}^{\infty} \Pr(A \mid B_i) \Pr(B_i)$ .  $\square$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

**Bayes' Formula**

Information

Channels

Definition

Examples

Combining

Channels

Epilogue

<sup>2</sup>The results we see are easily specialised to the finite case

# Bayes' Formula

## Theorem

Let  $\{B_i\}_{i \in \mathbb{N}}$  be a partition of  $S$ . Then

$$\forall A \subseteq S; \Pr(A) > 0. \forall i \in \mathbb{N}. \Pr(B_i | A) = \frac{\Pr(A | B_i) \Pr(B_i)}{\sum_{j=1}^{\infty} \Pr(A | B_j) \Pr(B_j)}$$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

**Bayes' Formula**

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Bayes' Formula

## Theorem

Let  $\{B_i\}_{i \in \mathbb{N}}$  be a partition of  $S$ . Then

$$\forall A \subseteq S; \Pr(A) > 0. \forall i \in \mathbb{N}. \Pr(B_i | A) = \frac{\Pr(A | B_i) \Pr(B_i)}{\sum_{j=1}^{\infty} \Pr(A | B_j) \Pr(B_j)}$$

## Proof.

From definition of conditional probability,

$$\begin{aligned} \Pr(B_i \cap A) &= \Pr(A \cap B_i) \\ \Pr(B_i | A) \Pr(A) &= \Pr(A | B_i) \Pr(B_i) \\ \Pr(B_i | A) &= \frac{\Pr(A | B_i) \Pr(B_i)}{\Pr(A)} \end{aligned}$$

Result follows from law of total probability. □

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

**Bayes' Formula**

Information

Channels

Definition

Examples

Combining

Channels

Epilogue

## Bayes' Formula, Simplified

The formula is often simplified as

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

which can be reversed

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

Thus, Bayes' formula expresses *exactly* how two events relate.

### Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

**Bayes' Formula**

Information  
Channels

Definition  
Examples

Combining  
Channels

Epilogue



## Bayes' Formula, Simplified

The formula is often simplified as

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

which can be reversed

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

Thus, Bayes' formula expresses *exactly* how two events relate.

At last,

### Definition (Joint Probability)

We let  $\Pr(A \cap B)$  be the *joint probability* of  $A$  and  $B$ .

We have (from above, and conditional probability),

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

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

**Bayes' Formula**

Information

Channels

Definition

Examples

Combining

Channels

Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

**Information  
Channels**

Definition

Examples

Combining  
Channels

Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - **Definition**
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

**Definition**

Examples

Combining  
Channels

Epilogue

# Formalising The Channel

Now we formally define the *noisy channel*.

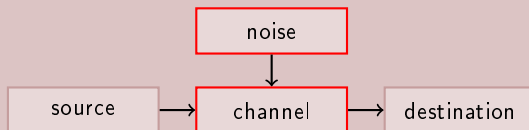


Figure: Communication System

Channel:  $\Gamma$

Input to  $\Gamma$ : a *source*  $\mathcal{A}$

- finite alphabet  $\mathcal{A} = \{a_1, \dots, a_r\}$ ,
- probability  $p_i$  associated to  $a_i$ , for each  $i$ ,  
s.t.  $0 \leq p_i \leq 1$  and  $\sum_{i=1}^r p_i = 1$

Output of  $\Gamma$ : a *source*  $\mathcal{B}$

- finite alphabet  $\mathcal{B} = \{b_1, \dots, b_s\}$ ,
- probability  $q_j$  associated to  $b_j$ , for each  $j$ ,  
s.t.  $0 \leq q_j \leq 1$  and  $\sum_{j=1}^s q_j = 1$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

Bayes' Formula

Information

Channels

**Definition**

Examples

Combining

Channels

Epilogue

# Definition

**Note:** it is tempting to consider  $\Gamma$  a function  $\Gamma: A \rightarrow B$  s.t.  $\Gamma(a) = b$ .

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

**Definition**

Examples

Combining  
Channels

Epilogue

---

<sup>3</sup>Book uses ' $\alpha$ '. I found it inconsistent with prior established habit.

<sup>4</sup>Book uses ' $b$ '.

# Definition

**Note:** it is tempting to consider  $\Gamma$  a function  $\Gamma: A \rightarrow B$  s.t.  $\Gamma(a) = b$ .

- Do not! At one point,  $\Gamma(a) = b$ , and later,  $\Gamma(a) = b'$ , due to *noise*.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

**Definition**

Examples

Combining  
Channels

Epilogue

---

<sup>3</sup>Book uses ' $\alpha$ '. I found it inconsistent with prior established habit.

<sup>4</sup>Book uses ' $b$ '.

# Definition

**Note:** it is tempting to consider  $\Gamma$  a function  $\Gamma: A \rightarrow B$  s.t.  $\Gamma(a) = b$ .

- Do not! At one point,  $\Gamma(a) = b$ , and later,  $\Gamma(a) = b'$ , due to *noise*.

**Input Symbol**<sup>3</sup>: random variable  $X$ , with  $\mathcal{X} = A$ ,  
and  $p_X(a_i) = \Pr(X = a_i) = p_i$ .

**Output Symbol**<sup>4</sup>: random variable  $Y$ , with  $\mathcal{Y} = B$ ,  
and  $p_Y(b_j) = \Pr(Y = b_j) = q_j$ .

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

**Definition**

Examples

Combining  
Channels

Epilogue

<sup>3</sup>Book uses ' $\alpha$ '. I found it inconsistent with prior established habit.

<sup>4</sup>Book uses ' $b$ '.

## Definition

**Note:** it is tempting to consider  $\Gamma$  a function  $\Gamma: A \rightarrow B$  s.t.  $\Gamma(a) = b$ .

- Do not! At one point,  $\Gamma(a) = b$ , and later,  $\Gamma(a) = b'$ , due to *noise*.

**Input Symbol<sup>3</sup>:** random variable  $X$ , with  $\mathcal{X} = A$ ,  
and  $p_X(a_i) = \Pr(X = a_i) = p_i$ .

**Output Symbol<sup>4</sup>:** random variable  $Y$ , with  $\mathcal{Y} = B$ ,  
and  $p_Y(b_j) = \Pr(Y = b_j) = q_j$ .

## Definition (Discrete Memoryless Channel)

A discrete memoryless channel  $\Gamma$  is a triple

$$\Gamma = (\mathcal{X}, p(b | a), \mathcal{Y}),$$

where  $X$  is the input variable,  $Y$  the output variable, and where  $Y$  depends only on  $X$  with the conditional probability  $p(b | a) = \Pr(Y = b | X = a)$ .

Recall that  $p(b | a)$  can be expressed as a matrix...

<sup>3</sup>Book uses 'a'. I found it inconsistent with prior established habit.

<sup>4</sup>Book uses 'b'.



# Perspectives

For  $\Gamma = (\mathcal{X}, p(b | a), \mathcal{Y})$ ,

**Forward probabilities:** the matrix  $P_{ij}$  for  $p$ .

$$P_{ij} = \Pr(Y = b_j | X = a_i) = \Pr(b_j | a_i) = p(b_j | a_i)$$

POV: Sender, guessing output.

**Backward probabilities:** the matrix  $Q_{ij}$  defined by

$$Q_{ij} = \Pr(X = a_i | Y = b_j) = \Pr(a_i | b_j)$$

POV: Receiver, guessing input.

**Joint probabilities:** the matrix  $R_{ij}$  defined by

$$R_{ij} = \Pr(X = a_i, Y = b_j) = \Pr(a_i, b_j)$$

POV: Observer, guessing both.

**The Relation:** Bayes' Formula:  $Q_{ij} = \frac{p_i}{q_j} P_{ij}$ . From:

$$p_i P_{ij} = \Pr(a_i) \Pr(b_j | a_i) = \Pr(a_i, b_j) = \Pr(b_j) \Pr(a_i | b_j) = q_j Q_{ij}$$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

Bayes' Formula

Information

Channels

**Definition**

Examples

Combining

Channels

Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 **Information Channels**
  - Definition
  - **Examples**
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

**Examples**

Combining  
Channels

Epilogue

# Binary Symmetric Channel

$$A = B = \mathbb{Z}_2 = \{0, 1\}$$

$P$ , where  $0 \leq P \leq 1$ , is the probability that a symbol  $a$  is correctly transmitted. Let  $\bar{P} = 1 - P$ .

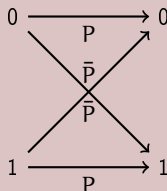


Figure: Binary Symmetric Channel

Let  $a_0 = 0$ ,  $a_1 = 1$ ,  $b_0 = 0$ , and  $b_1 = 1$ .

$$P_{ij} = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} = \begin{pmatrix} P & \bar{P} \\ \bar{P} & P \end{pmatrix}$$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

**Examples**

Combining  
Channels

Epilogue

$$P_{ij} = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} = \begin{pmatrix} P & \bar{P} \\ \bar{P} & P \end{pmatrix}$$

Given  $(p_0, p_1) = (p, \bar{p})$ , we get  $(q_0, q_1) = (pP + \bar{p}\bar{P}, p\bar{P} + \bar{p}P) = (q, \bar{q})$ :

$$(q, \bar{q}) = (p, \bar{p}) \begin{pmatrix} P & \bar{P} \\ \bar{P} & P \end{pmatrix}.$$

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

**Examples**

Combining  
Channels

Epilogue

$$P_{ij} = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} = \begin{pmatrix} P & \bar{P} \\ \bar{P} & P \end{pmatrix}$$

Given  $(p_0, p_1) = (p, \bar{p})$ , we get  $(q_0, q_1) = (pP + \bar{p}\bar{P}, p\bar{P} + \bar{p}P) = (q, \bar{q})$ :

$$(q, \bar{q}) = (p, \bar{p}) \begin{pmatrix} P & \bar{P} \\ \bar{P} & P \end{pmatrix}.$$

Inserting into Bayes' Formula gives

$$Q_{00} = \frac{pP}{pP + \bar{p}\bar{P}}$$

$$Q_{10} = \frac{\bar{p}\bar{P}}{pP + \bar{p}\bar{P}}$$

$$Q_{01} = \Pr(X = 0 \mid Y = 1) = \frac{\Pr(X = 0) \Pr(Y = 1 \mid X = 0)}{\sum_{k=0}^{r-1} \Pr(X = k) \Pr(Y = 1 \mid X = k)}$$

$$= \frac{p\bar{P}}{p\bar{P} + \bar{p}P}$$

$$Q_{11} = \frac{\bar{p}P}{p\bar{P} + \bar{p}P}$$

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

Bayes' Formula

Information

Channels

Definition

**Examples**

Combining

Channels

Epilogue

## Example (A Binary Symmetric Channel)

$P = 0.8$  (quite reliable), and  $p = 0.9$  (almost always *send* 0s). So

$$P_{ij} = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$$

and

$$(p, \bar{p}) = (0.9, 0.1).$$

Thus

$$(q, \bar{q}) = (p, \bar{p}) \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} = (0.74, 0.26).$$

Nr. of 0s *received* **fewer** than those sent.

$$Q_{00} = \frac{p_0 P_{00}}{q_0} = \frac{0.9 * 0.8}{0.74} \approx 0.973, \quad Q_{10} = \frac{p_1 P_{10}}{q_0} = \frac{0.1 * 0.2}{0.74} \approx 0.027,$$

$$Q_{01} = \frac{p_0 P_{01}}{q_0} = \frac{0.9 * 0.2}{0.26} \approx 0.692, \quad Q_{11} = \frac{p_1 P_{11}}{q_0} = \frac{0.1 * 0.8}{0.26} \approx 0.308.$$

Interesting: If 1 is received, usually, 0 was sent!

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

**Examples**Combining  
Channels

Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of

Redundancy

Main Result

Preliminaries

Conditional

Probability

Bayes' Formula

Information

Channels

Definition

Examples

**Combining  
Channels**

Epilogue

# Sum, Product, and Composition

Let  $M, M'$  be the channel matrices for  $\Gamma, \Gamma'$ .

$\Gamma + \Gamma'$ :  $A, A'$  and  $B, B'$  must be disjoint. Then

$$M'' = \begin{pmatrix} M & O \\ O & M' \end{pmatrix}.$$

Consider this a sort of “distribution”.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

**Combining  
Channels**

Epilogue

---

<sup>5</sup>Kronecker product of  $M$  and  $M''$ .



## Sum, Product, and Composition

Let  $M, M'$  be the channel matrices for  $\Gamma, \Gamma'$ .

$\Gamma + \Gamma'$ :  $A, A'$  and  $B, B'$  must be disjoint. Then

$$M'' = \begin{pmatrix} M & O \\ O & M' \end{pmatrix}.$$

Consider this a sort of “distribution”.

$\Gamma \times \Gamma'$ : Simultaneously send through  $\Gamma$  and  $\Gamma'$ . Sending *pairs* of input symbols. If  $M = (P_{ij})_{r \times s}$  and  $M'$  is a  $r' \times s'$  matrix, then

$$M'' = M \otimes M' = \begin{pmatrix} P_{11}M' & \cdots & P_{1s}M' \\ \vdots & \ddots & \vdots \\ P_{r1}M' & \cdots & P_{rs}M' \end{pmatrix},$$

a  $rr' \times ss'$  matrix<sup>5</sup>. Consider this a sort of “parallelism”.

Introduction
Problem Setting
Role of Redundancy
Main Result
Preliminaries
Conditional Probability
Bayes' Formula
Information Channels
Definition
Examples
Combining Channels
Epilogue

<sup>5</sup>Kronecker product of  $M$  and  $M''$ .

## Sum, Product, and Composition

Let  $M, M'$  be the channel matrices for  $\Gamma, \Gamma'$ .

$\Gamma + \Gamma'$ :  $A, A'$  and  $B, B'$  must be disjoint. Then

$$M'' = \begin{pmatrix} M & O \\ O & M' \end{pmatrix}.$$

Consider this a sort of “distribution”.

$\Gamma \times \Gamma'$ : Simultaneously send through  $\Gamma$  and  $\Gamma'$ . Sending *pairs* of input symbols. If  $M = (P_{ij})_{r \times s}$  and  $M'$  is a  $r' \times s'$  matrix, then

$$M'' = M \otimes M' = \begin{pmatrix} P_{11}M' & \cdots & P_{1s}M' \\ \vdots & \ddots & \vdots \\ P_{r1}M' & \cdots & P_{rs}M' \end{pmatrix},$$

a  $rr' \times ss'$  matrix<sup>5</sup>. Consider this a sort of “parallelism”.

$\Gamma \circ \Gamma'$ : Output alphabet of  $\Gamma$  must be the input alphabet of  $\Gamma'$ . Then

$$M'' = M \times M'.$$

Consider this a “series” of channels.

<sup>5</sup>Kronecker product of  $M$  and  $M'$ .

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information

Channels

Definition

Examples

Combining  
Channels

Epilogue

# Outline

- 1 Introduction
  - Problem Setting
  - Role of Redundancy
  - Main Result
- 2 Preliminaries
  - Conditional Probability
  - Bayes' Formula
- 3 Information Channels
  - Definition
  - Examples
- 4 Combining Channels
- 5 Epilogue

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue

# Summary

We have been introduced to the issue of channel coding.

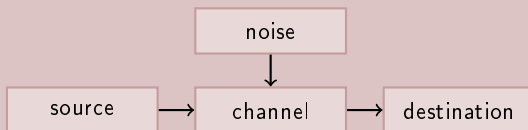


Figure: Communication System

We have seen the *Binary Symmetric Channel*; the channel which will receive most attention for the remainder of the course.

We have seen that *redundancy* plays a *key role* in the error correctability of a code; without it, errors *cannot* be corrected.

We shall look at *entropy* in this new setting, next session.

## Lecture 5: Information Channels

Introduction

Problem Setting

Role of  
Redundancy

Main Result

Preliminaries

Conditional  
Probability

Bayes' Formula

Information  
Channels

Definition

Examples

Combining  
Channels

Epilogue