Systems and Signals 344

2026

Lecture 3

Total Probability, Bayes' Theorem and Independence

Sections 1.4-1.6

Recap

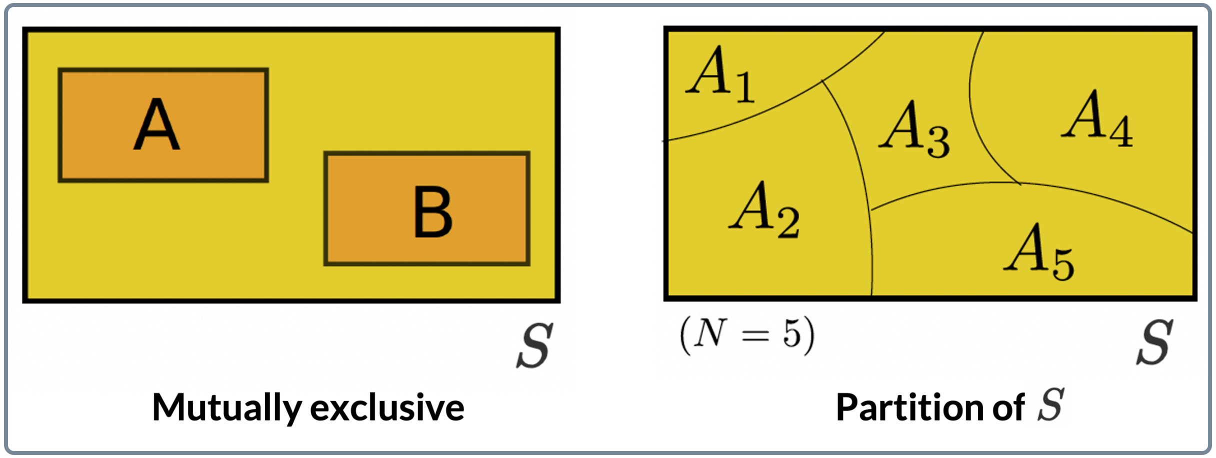
The conditional probability of the event A given the occurrence of the event B is

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

Recap

- A and B are **disjoint** (or **mutually exclusive**) if their intersection is the empty set, i.e. when $A \cap B = \emptyset$
- A collection of sets $A_1, \ldots, A_n$ is a **partition** of S if their union is S and their intersection is always the empty set, i.e.

$$\bigcup_{i=1}^{n} A_i = S \quad \text{and} \quad A_i \cap A_j = \emptyset \ \forall \ i \neq j$$



Independence

Events A and B are **independent** when the occurrence of one does not affect the probability of occurrence of the other, i.e.

$$P(A|B) = P(A) \textbf{ and } P(B|A) = P(B)$$

But from the definition of conditional probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$ and hence

$$P(A \cap B) = P(A) \cdot P(B)$$

when A and B are independent

Independence

Events A and B are **independent** if and only if

$$P(A \cap B) = P(A) \cdot P(B)$$

This is a **necessary and sufficient** condition for two **mutually independent** events.

Independence

A collection $A_1, \ldots, A_n$ is **mutually independent** if, for every subset of two or more events,

$$P\left(\bigcap_{r=1}^{k} A_{i_r}\right) = \prod_{r=1}^{k} P(A_{i_r}), \qquad 2 \leq k \leq n.$$

For the full collection, the same condition can be written more explicitly as

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2) \cdots P(A_n).$$

Note that showing statistical independence for all possible pairs of events is <u>not</u> sufficient to show independence between all n. The equation above must also hold true.

Independence

A_1, A_2, and A_3 are **mutually independent** if and only if

1. A_1 and A_2 are independent
2. A_2 and A_3 are independent
3. A_1 and A_3 are independent
4. $P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$

Example 1.20

In an experiment with equiprobable outcomes, the partition is $S = \{1, 2, 3, 4\}$. $\mathrm{P}[s] = 1/4$ for all $s \in S$. Are the events $A_1 = \{1, 3, 4\}$, $A_2 = \{2, 3, 4\}$, and $A_3 = \varnothing$ mutually independent?

Independence

- If a set of events is statistically independent, then any one event is also statistically independent of unions, intersections and/or complements of the remaining events.
  - For example, if A and B are independent, then A and $\bar{B}$ are also independent.
- **Independent and mutually exclusive are *not* synonyms**
  - Independence: $P(A \cap B) = P(A) \cdot P(B)$
  - Mutually exclusive: $P(A \cap B) = 0$

  $\Rightarrow$ Cannot be both! (unless $P(A) = 0$ or $P(B) = 0$)

Side Note

- You can only *add* probabilities directly for the *union* ($\cup$) of events that are **mutually exclusive**:
  - $P(A \cup B) = P(A) + P(B)$ when $A \cap B = \emptyset$.
  - In general, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- You can only *multiply* probabilities directly for the *intersection* ($\cap$) of events that are **independent**:
  - $P(A \cap B) = P(A)P(B)$ when A and B are independent.
  - In general, $P(A \cap B) = P(A)P(B \mid A)$ when $P(A) > 0$.

Partitions and the Law of Total Probability

*A **partition** divides the sample space into mutually exclusive sets.*

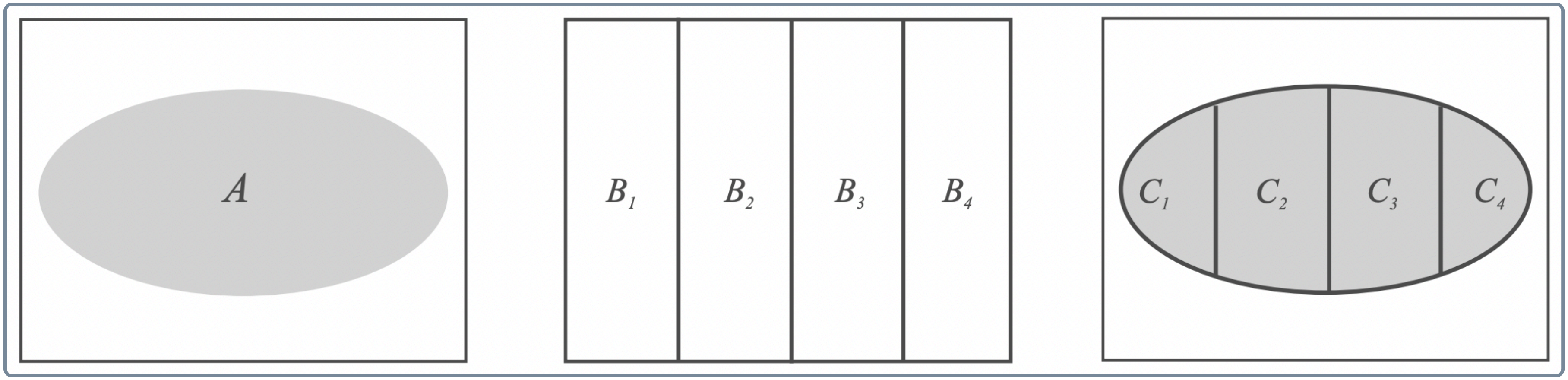
*The **law of total probability** expresses the probability of an event as the sum of the probabilities of outcomes that are in the separate sets of a partition.*

Partitions and the Law of Total Probability

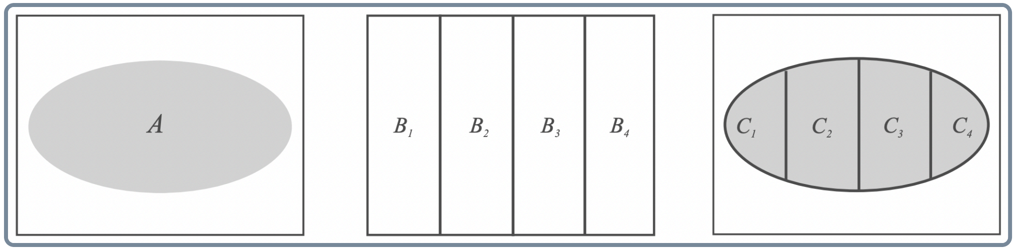
For a partition $B = \{B_1, B_2, \ldots\}$ *and any event* A *in the sample space (i.e.* $A \subseteq S$*), let* $C_i = A \cap B_i$.

For $i \neq j$*, the events* C_i *and* C_j *are mutually exclusive and*

$$A = C_1 \cup C_2 \cup \ldots$$



Partitions and the Law of Total Probability



We do this when $P(A \cap B_i)$ is easy to calculate but $P(A)$ is hard to calculate.

For any event $\boldsymbol{A}$, *and partition* $B = \{B_1, B_2, \dots B_m\}$,

$$P(A) = \sum_{i=1}^{m} P(A \cap B_i)$$

Law of Total Probability

For a partition $\{B_1, B_2, \ldots, B_m\}$ *with* $P(B_i) > 0$ *for all* i

$$P(A) = \sum_{i=1}^{m} P(A|B_i)P(B_i)$$

Example 1.15

A company has a model of email use. It classifies all emails as either long (l), if they are over 10 MB in size, or brief (b). It also observes whether the email is just text (t), has attached images (i), or has an attached video (v). This model implies an experiment in which the procedure is to monitor an email and the observation consists of the type of email, t, i, or v, and the length, l or b. The sample space has six outcomes: $S = \{lt, bt, li, bi, lv, bv\}$. In this problem, each email is classifed in two ways: by length and by type. Using L for the event that an email is long and B for the event that a email is brief, $\{L, B\}$ is a partition. Similarly, the text (T), image (I), and video (V) classification is a partition $\{T, I, V\}$. The sample space can be represented by a table in which the rows and columns are labeled by events and the intersection of each row and column event contains a single outcome. The corresponding table entry is the probability of that outcome. In this case, the table is

$$\begin{array}{c|ccc} & T & I & V \\ \hline L & 0.3 & 0.12 & 0.15 \\ B & 0.2 & 0.08 & 0.15 \end{array} \qquad (1.17)$$

Assume only one class (T, I, V) is assigned at a time.

Example 1.16

A company has three machines B_1, B_2, and B_3 making 1 kΩ resistors. Resistors within 50 Ω of the nominal value are considered acceptable. It has been observed that 80% of the resistors produced by B_1 and 90% of the resistors produced by B_2 are acceptable. The percentage for machine B_3 is 60%. Each hour, machine B_1 produces 3000 resistors, B_2 produces 4000 resistors, and B_3 produces 3000 resistors. All of the resistors are mixed together at random in one bin and packed for shipment. What is the probability that the company ships an acceptable resistor?

Bayes' Theorem

From Conditional Probability and Commutative laws:

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

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

Bayes' Theorem

From Conditional Probability and Commutative laws:

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

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

Bayes' Theorem:

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

Bayes' Theorem

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

- Bayes' theorem is extremely useful for making inferences about phenomena that cannot be observed directly.
    - "reasoning about causes when we observe effects"
- $P(B)$ is referred to as the **prior** probability, since it can be interpreted as the probability of B happening before we know anything about A.
- Similarly, $P(B|A)$ is known as the **posterior** probability, since it can be interpreted as the probability of B after observing A.

Bayes' Theorem

- When $\{B_i\}_{i=1}^m$ partitions S, we know from the Law of Total Probability that: $P(A) = \sum_{i=1}^{m} P(A|B_i) \cdot P(B_i)$
- And so we can reformulate Bayes' rule as:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^m P(A|B_j) \cdot P(B_j)}$$

Example 1.17

In Example 1.16 about a shipment of resistors from the factory, we learned that:

- The probability that a resistor is from machine B_3 is $\mathrm{P}[B_3] = 0.3$.
- The probability that a resistor is *acceptable* — i.e., within 50 Ω of the nominal value — is $\mathrm{P}[A] = 0.78$.
- Given that a resistor is from machine B_3, the conditional probability that it is acceptable is $\mathrm{P}[A|B_3] = 0.6$.

What is the probability that an acceptable resistor comes from machine B_3?

Story Time

Steve is very *shy and withdrawn*, invariably helpful but with very little interest in people or in the world of reality. A *meek and tidy soul*, he has a need for order and structure, and a passion for detail.

Story Time

Steve is very *shy and withdrawn*, invariably helpful but with very little interest in people or in the world of reality. A *meek and tidy soul*, he has a need for order and structure, and a passion for detail.

Which is more likely: Steve is a **librarian** or Steve is a **Farmer**?

Story Time

Steve is very *shy and withdrawn*, invariably helpful but with very little interest in people or in the world of reality. A *meek and tidy soul*, he has a need for order and structure, and a passion for detail.

Which is more likely: Steve is a **librarian** or Steve is a **Farmer**?

See video for explanation:
https://www.youtube.com/watch?v=HZGCoVF3YvM

Bayesian Thinking

- It can be used to reason about anything to which probabilities can be assigned.
- You start with prior belief, then you introduce evidence and you update the probability.