2: Conditional Probability, and Independence

PSTAT 120A: Summer 2022

Ethan P. Marzban

June 22, 2022

University of California, Santa Barbara

Announcements

Where We've Been

- Axioms of Probability; Probability Measure $\mathbb{P}$
- Probability Space $(\Omega, \mathcal{F}, \mathbb{P})$
- Classical DeQfinition of Probability
- Probability Rules (e.g. Complement Rule, Set Difference Rule, etc.)

Conditional Probability

Leadup

- Given an event A, the quantity $\mathbb{P}(A)$ represents our beliefs on the event A.
- Suppose we get some more information in the form of another event B.
- How, if at all, do ou beliefs on A Change?

- As an example: suppose we want to estimate the chance of rain. In the absence of any information, we might say that the chance of rain tomorrow is 50%.
- But, we know that it is summer, in Santa Barara; thus, we intuitively feel that the true chance of rain should probably be lower than 50%.

Proposition

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an event $B \in \mathcal{F}$ such that $\mathbb{P}(B) \neq 0$, the probability measure $\mathbb{P}_B : \mathcal{F} \to \mathbb{R}$ defined by

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

is a valid probability measure.

- I won't prove this, but the proof is quite straightforward and a very good exercise in applying the axioms of probability!
- Often times, instead of writing $\mathbb{P}_B(A)$ we will write $\mathbb{P}(A \mid B)$, read "the probability of A **given** B."
- $\mathbb{P}(A \mid B)$ represents an **updating** of our beliefs on A, in the presence of B.
  - Sometimes read "if B, then A."

Example

Suppose I randomly select a number from the set $[|1:100|]$ (this is a shorthand notation for $\{1, 2, \cdots, 100\}$). Define the events A and B as follows:

$$A := \{\text{the number I selected was strictly greater than 50}\}$$
$$B := \{\text{the number I selected was a multiple of 5}\}$$

- Because the selection is done "randomly," we can use the classical definition of probability.
  - There are 50 numbers greater than 50 (that are in the set $[|1:100|]$), meaning $\mathbb{P}(A) = 50/100 = 1/2$.
  - There are 20 multiples of 5 in the set $[|1:100|]$, meaning $\mathbb{P}(B) = 20/100 = 1/5$.
- Additionally, $A \cap B$ represents the event "the number I selected was both greater than 50 and a multiple of 5." There are 10 multiples of 5 that are greater than 50; therefore $\mathbb{P}(A \cap B) = 10/100 = 1/10$.
- Thus, putting everything together,

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/10}{1/5} = \frac{5}{10} = \frac{1}{2}$$

Multiplication Rule

- Our notion of conditional probability gives us a way of computing probabilities of intersections: since
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$
we can multiply both sides by $\mathbb{P}(B)$ to obtain:

Formula: The Multiplication Rule

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and two events $A, B \in \mathcal{F}$ with $\mathbb{P}(B) \neq 0$,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$$

- As an example: if A and B are two events with $\mathbb{P}(A) = 2/5$ and $\mathbb{P}(B \mid A) = 1/4$, then $\mathbb{P}(A \cap B) = \mathbb{P}(B \mid A) \cdot \mathbb{P}(A) = (1/4)(2/5) = 1/10$

Example

A recent survey at the *Isla Vista Co-Op* revealed that 50% of customers buy bread. Of those customers who buy bread, 20% buy cheese.

- **Always define notation first!** Let B denote "customer buys bread" and C denote "customer buys cheese." Then the problm tells us

$$\mathbb{P}(B) = 0.5; \quad \mathbb{P}(C \mid B) = 0.2$$

- We seek $\mathbb{P}(B \cap C)$. Since $\mathbb{P}(B \cap C) = \mathbb{P}(C \mid B) \cdot \mathbb{P}(B)$, we conclude that the proportion of customers who buy bread and cheese is

$$(0.2) \cdot (0.5) = \boxed{10\%}$$

Partitions (Again?)

- Now that we have the multiplication rule, we can derive a very useful formula.
- Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an event $A \in \mathcal{F}$.
- Consider another event $B \in \mathcal{F}$, and say we want to compute $\mathbb{P}(A)$.
- It is either the case that A happened along with B, or it happened along with not-B. That is,
$$A = [A \cap B] \cup [A \cap B^{\complement}]$$
- Taking the probability of both sides, and invoking the third axiom of probability, we find
$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^{\complement})$$

Partitions (Again?)

- Let's generalize this further. Suppose we have a partition $\{B_i\}_{i=1}^{\infty}$ of Ω. Then:
  - Either A happened along with B_1,
  - ... or B_2,
  - ... or B_3,
  - and so on and so forth.
- Therefore,

$$A = \bigcup_{i=1}^{\infty} (A \cap B_i)$$

and, taking the probability of both sides,

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

- Since $\mathbb{P}(A \cap B_i) = \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)$, we can rewrite this as:

Formula: The Law of Total Probability

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)$$

Example: On the Chalkboard

In *Gauchoville*, motherboards are manufactured by three companies (called A, B, and C). 20% of motherboards manufactured in factory A are defective; 30% of those manufactured in factory B are defective, and 10% of those manufactured in factory C are defective. If a motherboard is selected at random, what is the probability that it is defective?

Leadup

- Let's go back to our definition of conditional probability:

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

- Note that $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$.
- By the multiplication rule, $\mathbb{P}(B \cap A) = \mathbb{P}(B \mid A) \cdot \mathbb{P}(B)$.
- Hence, we have derived the following result:

Formula: Bayes' Theorem

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

- Colloquially, Bayes' Rule gives us a way of "reversing the order" of a conditional. This is especially useful when we have some sort of temporality.
- Oftentimes, we will use the Law of Total Probability in the denominator of Bayes' Rule.

Example

Let's go back to our motherboard example. Given that a randomly selected board was defective, what is the probability that it came from Factory A?

Independence

- Recall that $\mathbb{P}(A)$ represents our beliefs on an event A.
- Additionally, $\mathbb{P}(A \mid B)$ represents our updated beliefs on A, in the presence of B.
- What if $\mathbb{P}(A \mid B) = \mathbb{P}(A)$? In other words, our beliefs about A are completely unchanged by B.
- That is, A and B are *unaffected* by each other... they are **independent** of each other!

Definition: Independence

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and two events $A, B \in \mathcal{F}$, we say that A and B are **independent** (notated $A \perp B$) if $\mathbb{P}(A \mid B) = \mathbb{P}(A)$, or, equivalently, if $\mathbb{P}(B \mid A) = \mathbb{P}(B)$.

An equivalent condition for independence is $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Example

Suppose A and B are events with $\mathbb{P}(A) = 0.2$, $\mathbb{P}(B) = 0.3$, and $\mathbb{P}(A \cap B) = 0.1$. Are A and B independent?

- No, because $\mathbb{P}(A \cap B) = 0.1 \neq 0.2 \cdot 0.3 = \mathbb{P}(A) \cdot \mathbb{P}(B)$

Independence of Multiple Events

Definition: Independence of n Events

We say that a sequence of events $A_1, \ldots, A_n$ are **independent** (or **mutually independent**) if, for *every* subsequence $A_{i_1}, \ldots, A_{i_k}$, with $2 \leq k \leq n$ and $1 \leq i_1 < i_1 < \cdots < i_k \leq n$, we have

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \times \cdots \times \mathbb{P}(A_{i_k})$$

Independence of 4 events:

- $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$
- $\mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C)$
- $\mathbb{P}(A \cap D) = \mathbb{P}(A) \cdot \mathbb{P}(D)$
- $\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C)$
- $\mathbb{P}(B \cap D) = \mathbb{P}(B) \cdot \mathbb{P}(D)$
- $\mathbb{P}(C \cap D) = \mathbb{P}(C) \cdot \mathbb{P}(D)$

two-way intersections

three-way intersections

- $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$
- $\mathbb{P}(A \cap B \cap D) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(D)$
- $\mathbb{P}(A \cap C \cap D) = \mathbb{P}(A) \cdot \mathbb{P}(C) \cdot \mathbb{P}(D)$
- $\mathbb{P}(B \cap C \cap D) = \mathbb{P}(B) \cdot \mathbb{P}(C) \cdot \mathbb{P}(D)$

- $\mathbb{P}(A \cap B \cap C \cap D) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) \cdot \mathbb{P}(D)$

four-way intersections

Pairwise Independence

- Independence is a very strong condition!
- There exists a weaker form of independence:

Definition: Pairwise Independence

A sequence of events $A_1, A_2, \cdots$ is said to be **pairwise independent** if $A_i \perp A_j$ for any $i \neq j$.

- Note that independence implied pairwise independence, but not vice-versa.