Probability Theory

Applications for Data Science
Module 2: Conditional Probability

Anne Dougherty

February 6, 2021

TABLE OF CONTENTS

Learning Goals

Learning Goals for Module 2

In this module, we'll learn about conditional probability and Bayes formula. At the end of this module, learners should be able to:

- Explain the concept of conditional probability.
- Calculate probabilities using conditioning and Bayes Theorem.
- **Explain the concepts of independence and mutually exclusive events and provide examples.**

Independence

Two events are **independent** if knowing the outcome of one event does not change the probability of the other.

Examples:

- Flip a two-sided coin repeatedly. Knowing the outcome of one flip does not change the probability of the next.
- Roll a dice repeatedly.
- What about polling? What if you ask two randomly selected people about their political affiliation? What if the two people are friends?

If the 2 people are randomly chosen, knowing one person's affiliation shouldn't tell you anything about the other's affiliation.
This is not necessarily true if they are friends

Definition

Two events, A and B, are **independent** if $\underbrace{P(A|B)}_{\text{posterior}} = \underbrace{P(A)}_{\text{prior}}$, or equivalently, if $P(B|A) = P(B)$.

so finding new information about event B doesn't change the prob of A.

Recall:

$$P(A) \underset{\substack{\uparrow \\ \text{if } A \,\&\, B \text{ are indep.}}}{=} P(A|B) = \frac{P(A \cap B)}{P(B)}$$

then, if A and B are independent, we get the multiplication rule for independent events:

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

· This is for 2 events to be indep.
What if we have n events.

Definition Events $A_1, \cdots, A_n$ are **mutually independent** if for every k ($k = 2, 3, \ldots n$) and every subset of indices $i_1, i_2, \ldots, i_k$:

$$P(\underbrace{A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}}_{\text{every possible grouping of the events.}}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

Use the definition of independence in two ways: or any number

- We can use the definition to show two events A and B are (or are not) independent. To do this, we calculate $P(A)$, $P(B)$, and $P(A \cap B)$ to check if $P(A \cap B) = P(A)P(B)$.
- If we know two events are independent, we can find the probability of their intersection.

Example 1

Example: Roll a six-sided dice twice. Recall,
$S = \{(i,j) \mid i,j \in \{1,2,3,4,5,6\}\}$, $|S| = 36$ and each of the 36 outcomes of S is equally likely.

Let E be the event that the sum is 7.
Let F be the event that the first roll is a 4.
Let G be the event that the second roll is a 3.
What can you say about the independence of E, F and G?

$P(E) = P(\{16, 25, 34, 43, 52, 61\}) = \frac{1}{6}$

Also find, $P(F) = P(G) = \frac{1}{6}$

$P(E \cap F) = P(\{43\}) = \frac{1}{36} = P(E)P(F)$

$P(E \cap G) = P(\{43\}) = P(F \cap G) = \frac{1}{36}$

So, any pair of E, F + G are indep.

Mutual indep? $P(E \cap F \cap G) = P(\{43\}) = \frac{1}{36} \neq P(E)P(F)P(G)$

So not mutually indep. Think about this: If you know 2 of the events has occurred, then the 3rd one did too.

Example 2

In a school of 1200 students, 250 are juniors, 150 students are taking a statistics course, and 40 students are juniors and also taking statistics. One student is selected at random from the entire school. Let J be the event the selected student is a junior. Let S be the event that the selected student is taking statistics.

If the randomly chosen student is a junior, then what is the probability that they are also taking stats? Are J and S independent?

$P(J)=\frac{250}{1200}$, $\underbrace{P(S)=\frac{150}{1200}}_{.125}$ prior, $P(S\cap J)=\frac{40}{1200}$

J = junior

S = stats class

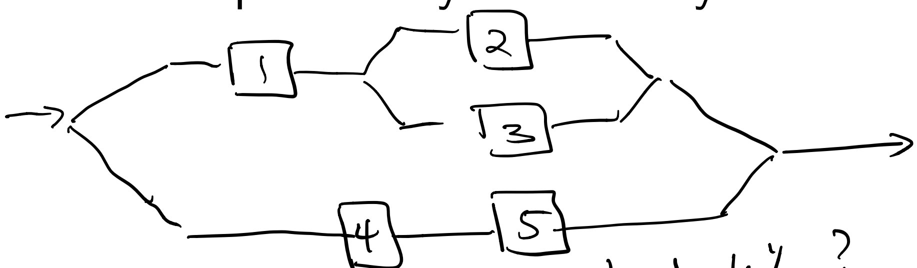
① $P(S|J)=\frac{P(S\cap J)}{P(J)}=\frac{40/1200}{250/1200}=\frac{40}{250}=\underbrace{\frac{4}{25}}_{\text{posterior}}=.16$

② $P(S|J)\neq P(S)$ so not indep.

(Also, note: $P(S\cap J)=\underbrace{\frac{40}{1200}}_{.0\bar{3}}\neq P(S)P(J)=\underbrace{\frac{250}{1200}\cdot\frac{150}{1200}}_{\sim .026}$

Example 3

Suppose you have a system of components as in the diagram. Let A_i be the event that the i^{th} component works and assume $P(A_i) = .9$ for $i = 1, 2, 3, 4, 5$. Assume the components work independently of each other. For the system to work, you need a path of working components from the start to the finish. Find the probability that the system works.



What does sample space look like?

$S = \{ (x_1 x_2 x_3 x_4 x_5) \mid x_i = 0$ if ith component works, $x_i = 1$ if ith component doesn't work $\}$

$|S| = 2^5 = 32$ but each element is not equally likely

For example: $P((0000)) = (.1)^5$, $P(\{11111\}) = (.9)^5$

Example 3 - continued

Recall: $P(A\cup B) = P(A)+P(B) - P(A\cap B)$

$P(A\cup B\cup C) = P(A)+P(B)+P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A\cap B\cap C)$

$$P(\text{system works}) = P\Big(\underbrace{(A_1\cap A_2)}_{11***} \cup \underbrace{(A_1\cap A_3)}_{1*1**} \cup \underbrace{(A_4\cap A_5)}_{***11}\Big)$$

$$= P(A_1\cap A_2) + P(A_1\cap A_3) + P(A_4\cap A_5) - P(A_1\cap A_2\cap A_4\cap A_5) - P(A_1\cap A_3\cap A_4\cap A_5)$$
$$- P(A_1\cap A_2\cap A_3) - P(A_1\cap A_2\cap A_3) + P(A_1\cap A_2\cap A_3\cap A_4\cap A_5)$$

$$= 3(.9)^2 - (.9)^3 - 2(.9)^4 + (.9)^5 = \underbrace{.97929}_{\text{overall prob.}}$$

So $\underbrace{P(A_1\cap A_2)}_{\text{prob system works with components 1+2}} = (.9)^2 = .81$

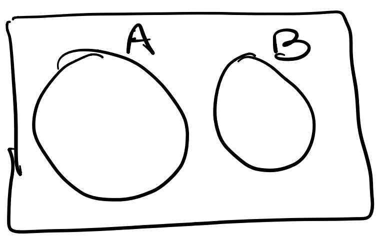
redundancy is key.

Extend: prob. of various components might not be the same, etc. But basic idea is the same.

One final question: Suppose you know two events A and B are mutually exclusive, that is, $A \cap B = \emptyset$. Are A and B independent?

You might think yes since if A & B are mutually exclusive then knowing the prob of one doesn't influence the other.

But:



$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 0$$

If we know B has occurred, then A cannot occur.