STAT 234 Lecture 3A
Bayes Theorem

Yibi Huang
Department of Statistics
University of Chicago

Reminder

- Lab 1 is at **7-8 pm** on **Monday, 10/3** on Zoom,
  See Canvas Announcement for Zoom Link
- Office Hour schedule posted on Canvas
  https://canvas.uchicago.edu/courses/45317/pages/office-hours-and-lab-sessions
- HW1 due Wed 10/5
- HW2 due Fri 10/7
- Additional practice problems on Probability: See

https://canvas.uchicago.edu/courses/45317/pages/lecture-slides-and-videos-10-30-section

Review

- Probability Axioms
- Complementation Rule: $P(A^c) = 1 - P(A)$
- General Addition Rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Conditional Probability: $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$ if $P(B) > 0$.
  - or calculate by restricting the sample space
- General Multiplication Rule: $P(A \cap B) = P(A) \times P(B \mid A)$
- Independence: Events A and B are independent if $P(A|B) = P(A)$
- Multiplication Rule for Independent events:

$$P(A \cap B) = P(A) \times P(B) \quad \text{if } A, B \text{ are indep.}$$

Abuse of the Multiplication Rule

As estimated in 2020, of the U.S. population,

- 2.0% were 85 or older, and
- 49.5% were male.

True or False and explain: $0.495 \times 0.02 \approx 1\%$ of the U.S. population are males and age 85+.

Abuse of the Multiplication Rule

As estimated in 2020, of the U.S. population,

- 2.0% were 85 or older, and
- 49.5% were male.

True or False and explain: $0.495 \times 0.02 \approx 1\%$ of the U.S. population are males and age 85+.

False, Age and Gender are dependent.
In particular, as women on average live longer than men, there are more old women than old men.

Among 85+ year-olds, only 36.5% are male, not 49.5%.
Of the U.S. population in 2020, only

$$P(M \cap 85+) = P(85+)P(M|85+) = 0.02 \times 0.365 \approx 0.0073 = 0.73\%$$

were males age 85+.

Tree Diagrams and Bayes’ Theorem

Example – A Nervous Job Applicant

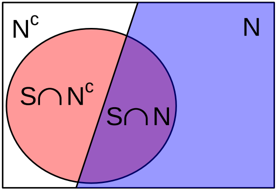
Suppose an job applicant has been invited for an interview.

The probability

- that he is nervous is $P(N) = 0.7$,
- of successful interview given he is nervous is $P(S \mid N) = 0.2$,
- of successful interview given he is not nervous is $P(S \mid N^c) = 0.9$.

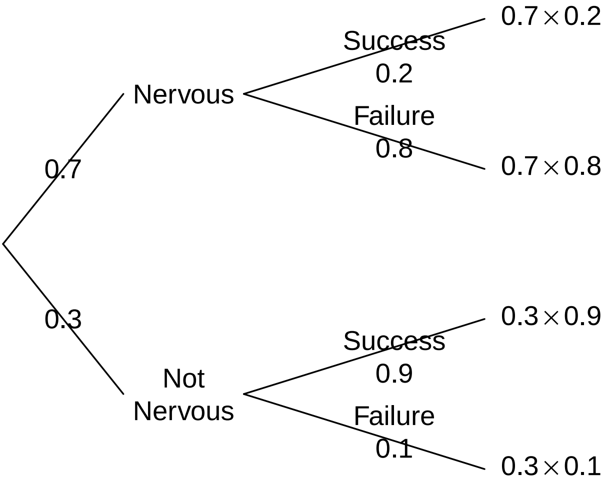
What is the probability that the interview is successful?

$$
\begin{aligned}
P(S) &= P(S \cap N) + P(S \cap N^c) \\
&= P(N)P(S \mid N) + P(N^c)P(S \mid N^c) \\
&= 0.7 \times 0.2 + 0.3 \times 0.9 = 0.41
\end{aligned}
$$



Tree Diagram for the Nervous Job Applicant Example

Another look at the nervous job applicant example:



Nervous Job Applicant Example Continued

Conversely, given the interview is successful, what is the probability that the job applicant is nervous during the interview?

$$
\begin{aligned}
P(N \mid S) &= \frac{P(N \cap S)}{P(S)} \\
&= \frac{P(N \cap S)}{0.41} \qquad \left(\begin{array}{c} \text{where } P(S) = 0.41 \text{ was} \\ \text{found in the previous page} \end{array} \right) \\
&= \frac{P(N)P(S \mid N)}{0.41} \qquad \text{since } P(N \cap S) = P(N)P(S \mid N) \\
&= \frac{0.7 \times 0.2}{0.41} = \frac{14}{41} \approx 0.34.
\end{aligned}
$$

Bayes' Theorem (or Bayes' Rule)

The problem in the previous slide is an example of **Bayes' Theorem**.

Knowing $P(B \mid A), P(B \mid A^c)$, and $P(A)$, is there a way to know $P(A \mid B)$?

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

Bayes' Theorem (or Bayes' Rule)

The problem in the previous slide is an example of **Bayes' Theorem**.

Knowing $P(B \mid A), P(B \mid A^c)$, and $P(A)$, is there a way to know $P(A \mid B)$?

$$
\begin{aligned}
P(A \mid B) &= \frac{P(A \cap B)}{P(B)} \\
&= \frac{P(A)P(B \mid A)}{P(B)} \quad \text{since } P(A \cap B) = P(A)P(B \mid A)
\end{aligned}
$$

Bayes' Theorem (or Bayes' Rule)

The problem in the previous slide is an example of **Bayes' Theorem**.

Knowing $P(B \mid A), P(B \mid A^c)$, and $P(A)$, is there a way to know $P(A \mid B)$?

$$\begin{aligned} P(A \mid B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(A)P(B \mid A)}{P(B)} \quad \text{since } P(A \cap B) = P(A)P(B \mid A) \\ &= \frac{P(A)P(B \mid A)}{P(B \cap A) + P(B \cap A^c)} \quad \text{since } B = (B \cap A) \cup (B \cap A^c) \end{aligned}$$

Bayes' Theorem (or Bayes' Rule)

The problem in the previous slide is an example of **Bayes' Theorem**.

Knowing $P(B \mid A), P(B \mid A^c)$, and $P(A)$, is there a way to know $P(A \mid B)$?

$$\begin{aligned}
P(A \mid B) &= \frac{P(A \cap B)}{P(B)} \\
&= \frac{P(A)P(B \mid A)}{P(B)} \quad \text{since } P(A \cap B) = P(A)P(B \mid A) \\
&= \frac{P(A)P(B \mid A)}{P(B \cap A) + P(B \cap A^c)} \quad \text{since } B = (B \cap A) \cup (B \cap A^c) \\
&= \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A^c)P(B \mid A^c)}
\end{aligned}$$

Medical Testing

A common application of Bayes' rule is in medical testing

- Let D denote the event that an individual has the disease that we are testing for

Medical Testing

A common application of Bayes' rule is in medical testing

- Let D denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*, and $T-$ denote the event that the test result is *negative*

Medical Testing

A common application of Bayes' rule is in medical testing

- Let D denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*, and $T-$ denote the event that the test result is *negative*
- $P(T+|D)$ is called the *sensitivity* of the test

Medical Testing

A common application of Bayes' rule is in medical testing

- Let D denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*, and $T-$ denote the event that the test result is *negative*
- $P(T+|D)$ is called the *sensitivity* of the test
- $P(T-|D^c)$ is called the *specificity* of the test

Medical Testing

A common application of Bayes' rule is in medical testing

- Let D denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*, and $T-$ denote the event that the test result is *negative*
- $P(T+|D)$ is called the *sensitivity* of the test
- $P(T-|D^c)$ is called the *specificity* of the test
- Ideally, we hope $P(T+|D)$ and $P(T-|D^c)$ both are equal to 1. However, diagnostic tests are not perfect. They may give false positives and false negatives.

Enzyme Immunoassay Test for HIV

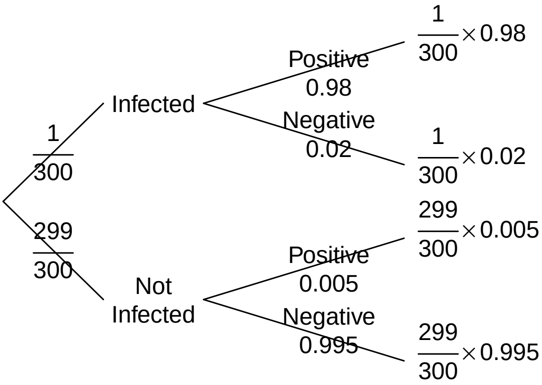
- $P(T+|D) = 0.98$ (sensitivity - positive for infected)
- $P(T-|D^c) = 0.995$ (specificity - negative for not infected)
- $P(D) = 1/300$ (prevalence of HIV in USA)

What is the probability that the tested person is infected if the test was positive?

$$
\begin{aligned}
P(D|T+) &= \frac{P(D)P(T+|D)}{P(D)P(T+|D) + P(D^c)P(T+|D^c)} \\
&= \frac{1/300 \times 0.98}{(1/300) \times 0.98 + (299/300) \times 0.005} \\
&= 39.6\%
\end{aligned}
$$

This test is not confirmatory. Need to confirm by a second test.

Tree Diagram for the HIV Test

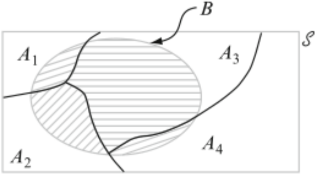


$$P(D|T+) = \frac{(1/300) \times 0.98}{(1/300) \times 0.98 + (299/300) \times 0.005}$$

Bayes' Theorem for 3 or More Cases

- In the 2 examples above, we split the sample space into 2 parts A or A^c (nervous or not nervous, infected or not infected)
- In some cases, we need to calculate $P(B)$ by splitting it into several parts, using the *law of total probability*:

Suppose $A_1, A_2, \ldots, A_k$ are disjoint and $A_1 \cup A_2 \cup \cdots \cup A_k = S$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$, then



$$
\begin{aligned}
P(B) &= P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_k) \\
&= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_k)P(B|A_k).
\end{aligned}
$$

Using the *law of total probability*, Bayes Theorem becomes

$$
P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_K)P(B|A_K)}
$$

Exercise 58 on p.83 of MMSA

At a gas station,

- 40% of the customers use regular gas (A_1),
- 35% use mid-grade gas (A_2), and
- 25% use premium gas (A_3).

Moreover,

- of those customers using regular gas, only 30% fill their tanks;
- of those using mid-grade, 60% fill their tanks;
- of those using premium, 50% fill their tanks.

Let B denote the event that the next customer fills the tank.

Exercise 58 on p.83 of MMSA

At a gas station,

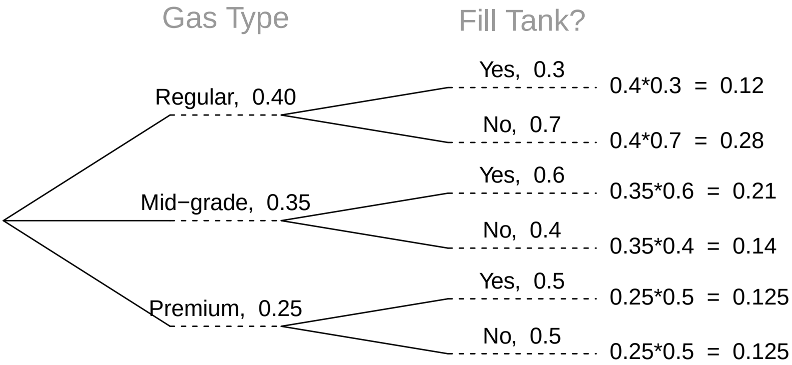
- 40% of the customers use regular gas (A_1),
- 35% use mid-grade gas (A_2), and
- 25% use premium gas (A_3).

Moreover,

- of those customers using regular gas, only 30% fill their tanks; $P(B|A_1) = 0.3$
- of those using mid-grade, 60% fill their tanks; $P(B|A_2) = 0.6$
- of those using premium, 50% fill their tanks. $P(B|A_3) = 0.5$

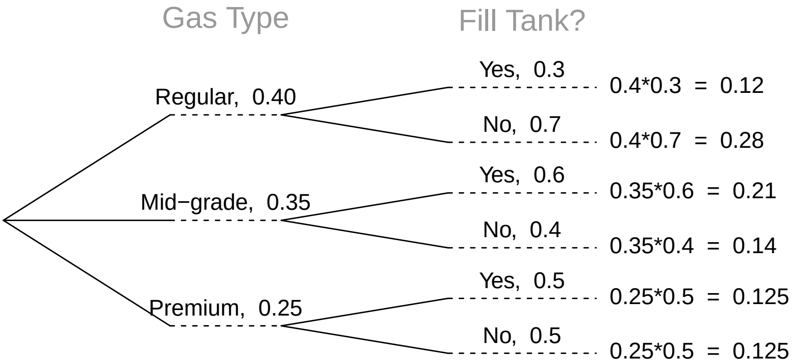
Let B denote the event that the next customer fills the tank.

Exercise 58 on p.83 of MMSA — Tree Diagram



Q1: What is the probability that the next customer request premium gas and fill the tank.

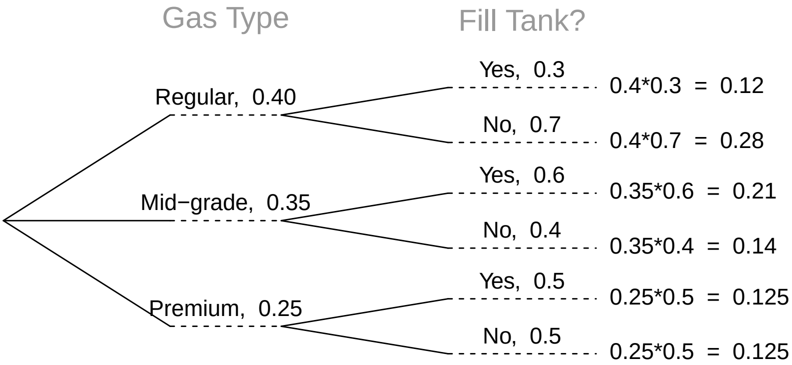
Exercise 58 on p.83 of MMSA — Tree Diagram



Q1: What is the probability that the next customer request premium gas and fill the tank.

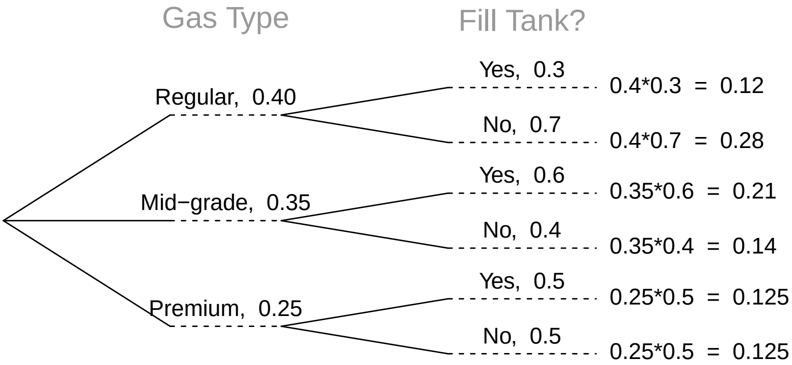
$$P(A_3 \cap B) = P(A_3)P(B|A_3) = 0.25 \times 0.5 = 0.125.$$

Exercise 58 on p.83 of MMSA — Tree Diagram



Q2: What is the probability that the next customer fills the tank.

Exercise 58 on p.83 of MMSA — Tree Diagram

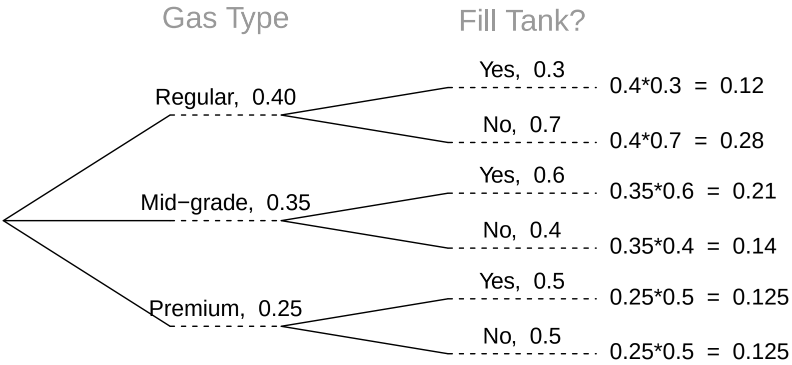


Q2: What is the probability that the next customer fills the tank.

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3)$$

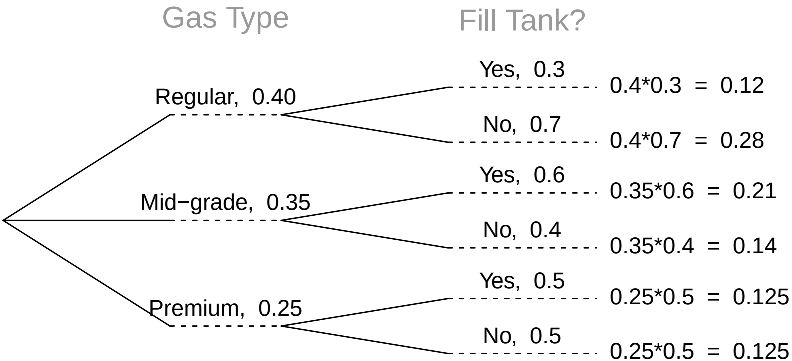
$$= 0.4 \times 0.3 + 0.35 \times 0.6 + 0.25 \times 0.5 = 0.455$$

Exercise 58 on p.83 of MMSA — Tree Diagram



Q3: If the next customer fills the tank, what is the probability that premium gas is requested?

Exercise 58 on p.83 of MMSA — Tree Diagram



Q3: If the next customer fills the tank, what is the probability that premium gas is requested?

$$P(A_3|B) = \frac{P(A_3 \cap B)}{P(B)} = \frac{0.125}{0.455} \approx 0.275.$$

Draw a tree diagram!

Don’t try to memorize the formula of Bayes’ Theorem.