**MATH 320: Probability**

# Lecture 6: Bayes' Theorem

Chapter 1: Probability (1.5)

## Law of total probability

Motivation

- Example: A company has three assembly lines. The first line, the second line and the third produce 30%, 50% and 20% of productions, respectively. unconditional

  Additionally, 1 of 100 productions is defective in Line 1; 2 of 100 productions are defective in Line 2; 3 of 100 productions are defective in Line 3. Conditional

  We want to know the probability of defectives produced in the company.

- *STRATEGY*:

  (a) Define all (unconditional) events given in the problem and find their probabilities.

  If conditional events are given, define them using unconditional events.

  $L_1$ = Line 1 $\rightarrow P(L_1) = 0.3$ D = Defective $\rightarrow P(D|L_1) = 0.01$

  $L_2$ = Line 2 $\rightarrow P(L_2) = 0.5$ $P(D|L_2) = 0.02$

  $L_3$ = Line 3 $\rightarrow P(L_3) = 0.2$ $P(D|L_3) = 0.03$

  (b) Find the event of interest. Try to express the event of interest as a composition (union) of the given events.

  Often it is desirable to form compositions mutually exclusive or independent events.

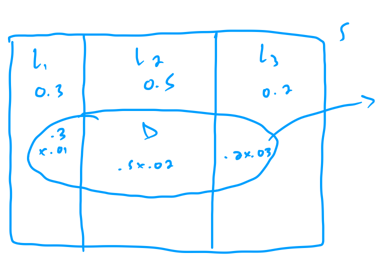
  $$P(D) = P[(D \cap L_1) \cup (D \cap L_2) \cup (D \cap L_3)]$$
  $$= P(D \cap L_1) + P(D \cap L_2) + P(D \cap L_3)$$ skip to here

  (c) Use the general multiplication rule to find the probability for the event of interest.

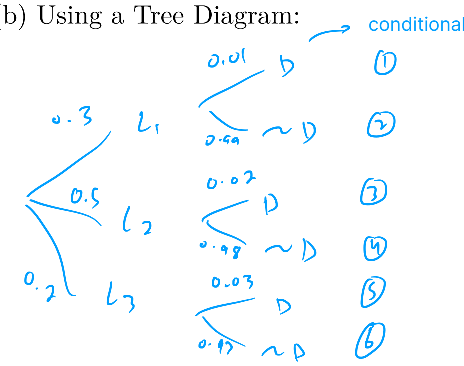
  $$P(D) = \underbrace{P(L_1)P(D|L_1)}_{1} + \underbrace{P(L_2)P(D|L_2)}_{3} + \underbrace{P(L_3)P(D|L_3)}_{5}$$
  $$= 0.3(0.01) + 0.5(0.02) + 0.2(0.03)$$
  $$= 0.019$$

- Visualizing scenario:

  (a) Using a Venn Diagram:

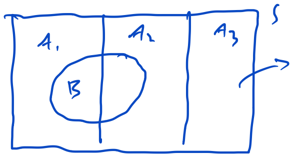



  (b) Using a Tree Diagram:



Law of total probability

- The **law of total probability** says that a partition $(A_1, \ldots, A_n)$ of the sample space will lead to a partition of any event $B$ into mutually exclusive pieces.




Some can be $\emptyset$

$\Rightarrow P(\,) = 0$

$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)$

Then we can write $P(B)$ as the sum of the probabilities of those pieces. Note that an event $A$ and its complement $\sim A$ always partition $S$.

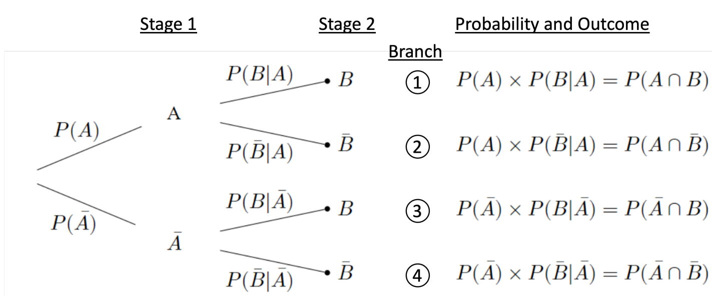
- Definition: **Law of total probability**

  Let $B$ be an event. If $A_1, \ldots, A_n$ partition the sample space, then

$$
\begin{aligned}
P(B) &= P\left[\bigcup_{i=1}^{n} (A_i \cap B)\right] && \langle \text{partition} \rangle \\
&= P(A_1 \cap B) + \cdots + P(A_n \cap B) && \langle \text{disjoint} \rangle \\
&= P(A_1)\, P(B|A_1) + \cdots + P(A_n)\, P(B|A_n) && \langle \text{multiplication rule} \rangle \\
&= \sum_{i=1}^{n} P(A_i)\, P(B|A_i)
\end{aligned}
$$

- Using the general tree diagram below, we can summarize

  Law of total probability $= P(\text{Second stage event}) = \sum$ Branches of interest




$P(B) = ① + ③$

$P(A|B) = \dfrac{①}{① + ③}$

## Bayes' Theorem

Bayes' Theorem

- Using the same terminology, we can summarize

  Bayes' Theorem $= P(\text{First stage event} \mid \text{Second stage event}) = \dfrac{\text{Main branch of interest}}{\sum \text{All branches of interest}}$

- In essence, Bayes' Theorem reverses the natural order of the tree for the conditional probability of interest.

- Continuing example:

  Find the probability that a defective product was made in Line 1.

$$P(L_1 \mid D) = \frac{P(L_1 \cap D)}{P(D)} = \frac{1}{1 + 3 + 5}$$

$$= \frac{0.3(0.01)}{0.3(0.01) + 0.5(0.02) + 0.7(0.03)}$$

$$\approx 0.158$$

- Definition: **Bayes' Theorem**

  Let $B$ be an event. If $A_1, \ldots, A_n$ partition the sample space, then

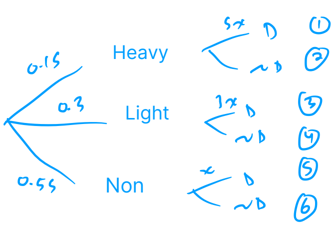
$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)}$$

$$= \frac{P(A_i)\,P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)\,P(B \mid A_j)}$$

law of total probability

- Example: At the beginning of a certain study of a group of persons, 15% were classified as heavy smokers, 30% as light smokers, and 55% as nonsmokers. In the five year study, it was determined that the death rates of the heavy and light smokers were five and three times that of the nonsmokers, respectively.

  A randomly selected participant died over the five-year period; calculate the probability that the participant was a nonsmoker.




$$P(\text{non} \mid D) = \frac{P(\text{non} \cap D)}{P(D)}$$

$$= \frac{5}{1 + 3 + 5}$$

$$= \frac{0.55\cancel{x}}{0.15(5\cancel{x}) + 0.3(3\cancel{x}) + 0.55(\cancel{x})}$$

$$= 0.25$$

Bayes' Theorem from another perspective

- Bayes' Theorem is all about changing probabilities based on new evidence.
- In a previous example, we drew a tree diagram about testing for the presence of a disease and the result of the test. We used the following events:

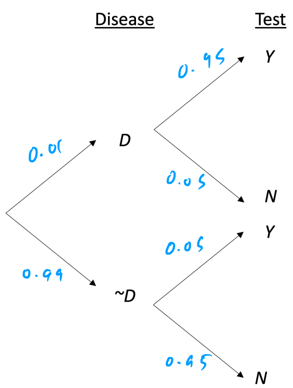
$D$ = the person tested has the disease
$\sim D$ = the person tested does not have the disease
$Y$ = the test is positive
$N$ = the test is negative

Lets now consider a disease test that is "95% accurate", which can be defined as follows:

a. If you have the disease, 95% chance of a positive test.

b. If you do not have the disease, 95% chance of a negative test.

Further, suppose only 1% of the population actually have this disease (aka prevalence).




given → $P(D) = 0.01$ prior probability

→ $P(Y|D) = 0.95$, $P(N|\sim D) = 0.95$ evidence

want → $P(Y)$, $P(N)$ $= ?$ Total probabilities

→ $P(D|Y)$, $P(\sim D|Y)$, $P(D|N)$, $P(\sim D|N)$ $= ?$ posterior probabilities

- Terminology:
  - Prior probability: Original unconditional probabilities
  - Evidence: Conditional probability given the prior information
  - Posterior probability: Prior probability conditioned on the new evidence

- Some calculations in context: The good and the bad of Bayes' Theorem
  1. Find the probability of testing positive (total probability).

     $P(Y) = (1) + (3)$

     $= P(D, Y) + P(\sim D, Y)$

     $= 0.01(0.95) + 0.99(0.05)$

     $= 0.059$

  2. Lets' solve for the probability of having the disease *given that you test positive.*

     (a) For a randomly selected person from the population, we had our original prior probability of having the disease, $P(D) = 0.01$.

     (We don't know if they do or don't have the disease, it remains unknown).

     (b) Then this person got tested, and tested positive; this is our evidence.

     Intuitively, this likelihood of the person having the disease should increase; we are adjusting the prior probability upwards based on the new evidence.

(c) Now we can calculate this new posterior probability.

$$P(D|Y) = \frac{①}{① + ③} = \frac{P(D,Y)}{P(Y)} = \frac{0.01(0.95)}{0.059} = 0.161$$

3. Suppose you know that someone has tested positive for this disease. What is the probability that the person does not actually have the disease?

$$P(\sim D|Y) = 1 - P(D|Y) = 1 - 0.161 = 0.839$$

$$= \frac{③}{① + ③} = \frac{P(\sim D,Y)}{P(Y)} = \frac{0.99(0.05)}{0.059} = 0.839$$

- The practical information here is interesting.
  1. The "95% accurate" test will classify 5.9% of the population as positives, compared to the true prevalence of $P(D) = 0.01$.
  2. This is the good side of Bayes' Theorem! By updating our prior probability with the new evidence, we drastically increased our information about this person having the disease.

     $P(D) = 0.01 \rightarrow P(D|Y) = 0.161$
  3. 83.9% of the individuals who tested positive will actually not have the disease.

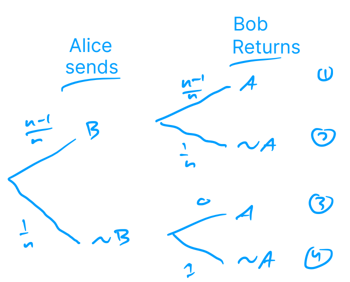
     This percentage depends heavily on the prevalence, for example if $P(D) = 0.1 \rightarrow P(\sim D \mid Y) =$ 32.1%; and if $P(D) = 0.001 \rightarrow P(\sim D \mid Y) =$ 98.1%.

Final example

given

- Alice writes to Bob and does not receive an answer. Assuming that one letter in $n$ is lost in the mail, find the probability that Bob received the letter. It is to be assumed that Bob would have answered the letter if he had received it.

  Let $A$ = Alice receives letter from Bob and $B$ = Bob receives letter from Alice.




$$P(B|\sim A) = \frac{P(B,\sim A)}{P(\sim A)} = \frac{②}{② + ④} = \frac{\frac{n-1}{n}\left(\frac{1}{n}\right)}{\frac{n-1}{n}\left(\frac{1}{n}\right) + \frac{1}{n}(1)} = \frac{n-1}{(n-1)+n} = \frac{n-1}{2n-1}$$