CS6462

Probabilistic and Explainable AI

Lesson 6

Bayes' Theorem

by Emil Vassev

January 24, 2025

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

Probability of event A occured and event B occured

Probability of event A given B has occured

Probability of event B

Calculating Conditional Probability

Definition (recall):

- the *"likelihood of an event occurring, based on the occurrence of another event"*:

  the *conditional probability* ***P(E|F)*** of an event ***E*** given an event ***F***, is the probability that ***E*** occurs given that ***F*** has occurred

$$P(E|F) = \frac{P(E \cap F)}{P(F)}, \text{ iff } P(F) > 0$$

Alternative way to calculate conditional probability:

- one conditional probability can be calculated using the other conditional probability $P(E|F) = P(F|E) * P(E) / P(F)$ – useful when $P(E \cap F)$ is challenging to calculate
- reverse

  $P(F|E) = P(E|F) * P(F) / P(E)$

General multiplication: $P(E \cap F) = P(E|F) * P(F) = P(F|E) * P(E)$

Probability of event A occured and event B occured

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

Probability of event A given B has occured

Probability of event B

Bayes’ Theorem

Bayes’ Theorem:

- principled way of calculating a conditional probability without the composite probability $P(E \cap F)$

$$P(E|F) = \frac{P(F|E)*P(E)}{P(F)}$$

- if we do not have access to the denominator $P(F)$ directly, we can calculate it:

$$P(F) = P(F|E)*P(E) + P(F|E^c)*P(E^c)$$

Partition rule: $P(E) = P(E \cap F) + P(E \cap F^c)$

- new form:

$$P(E|F) = \frac{P(F|E)*P(E)}{P(F|E)*P(E) + P(F|E^c)*P(E^c)}$$

Probability of event A occured and event B occured

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

Probability of event A given B has occured

Probability of event B

Bayes’ Theorem (cont.)

General form of Bayes’ Theorem:

if $\boldsymbol{E_1, \ldots, E_k}$ are mutually exclusive events and $\boldsymbol{E_1 \cup E_2 \ldots E_k = S}$, then

$$\boldsymbol{P(E_i|F) = \frac{P(F|E_i)*P(E_i)}{P(F|E_1)*P(E_1) + \ldots + P(F|E_k)*P(E_k)}}$$

Total probability:
$\boldsymbol{P(F) = P(F|E_1)*P(E_1) + \ldots + P(F|E_k)*P(E_k)}$

Terms:

- P(E|F): Posterior probability
- P(E): Prior probability
- P(F|E): Likelihood probability
- P(F): Evidence probability
- this allows Bayes Theorem to be restated as:

  Posterior = Likelihood * Prior / Evidence

Bayes' Theorem for Independent Events

Probability of event A occured and event B occured

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

Probability of event A given B has occured

Probability of event B

Independent Events E and F:

- if two events ***E*** and ***F*** are independent, then:

$$P(E|F) = P(E)$$

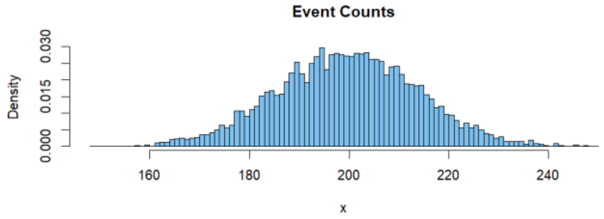
$$P(F|E) = P(F)$$

$$P(F|E^c) = P(F)$$

$$P(E|F) = P(E) = \frac{P(F|E)*P(E)}{P(F|E)*P(E) + P(F|E^c)*P(E^c)} = \frac{P(F)*P(E)}{P(F)*P(E) + P(F)*P(E^c)} = \frac{P(E)}{P(E) + P(E^c)} = P(E)$$

Bayes' Theorem cannot be used for independent events as *we need to determine the total probability and there is no dependency of events.*

Bayes’ Theorem - Example



Example 1:

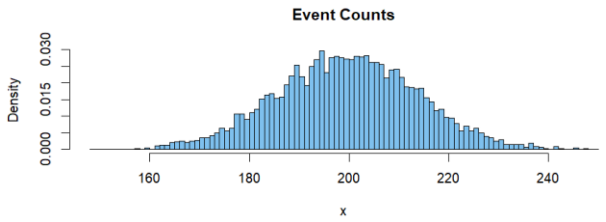
Finding out a patient’s probability of having liver disease if they are an alcoholic. “Being an alcoholic” is the **test** (kind of like a litmus test) for liver disease.

Litmus test for liver disease:

- ***P(A) = 0.10*** (past data tells you that 10% of patients entering your clinic “<u>have liver disease</u>” - ***A***)
- ***P(B) = 0.05*** (5% of the clinic’s “<u>patients are alcoholics</u>” - ***B***)
- ***P(B|A) = 0.07*** (among those patients diagnosed with liver disease, 7% are alcoholics)
- ***P(A|B)*** = ?

$$P(A|B) = \frac{P(B|A)*P(A)}{P(B)} = \frac{0.07 * 0.1}{0.05} = 0.14 = 14\%$$

Bayes' Theorem – Example (cont.)



Example 2: Testing for Covid-19

Antibody Test for Covid-19:

- ***P(T+|I+) = 0.98*** (sensitivity - positive test ***T+*** for infected ***I+***)
- ***P(T-|I-) = 0.995*** (specificity - negative test ***T-*** for non-infected ***I-***)
- ***P(I+) = 0.0003*** (prevalence)

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

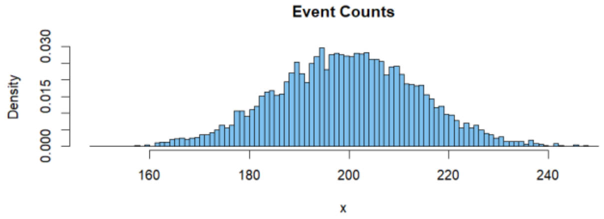
$$P(I+|T+) = \frac{P(T+|I+)*P(I+)}{P(T+|I+)*P(I+) + P(T+|I-)P(I-)} = \frac{0.98 * 0.0003}{0.98 * 0.0003 + 0.005 * 0.9997} = 0.05556$$

Testing on large scale not sensible (too many false positives).

Considering different population with P(I+) = 0.1 (greater risk):

$$P(I+|T+) = \frac{P(T+|I+)*P(I+)}{P(T+|I+)*P(I+) + P(T+|I-)P(I-)} = \frac{0.98 * 0.1}{0.98 * 0.1 + 0.005 * 0.9997} = 0.956$$

Bayes' Theorem – Example (cont.)

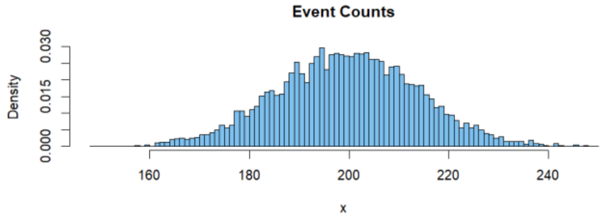


Example 3:

There are two bags. Bag I has 7 red and 4 blue balls and bag II has 5 red and 9 blue balls. We draw a ball at random and it turns out to be red. Determine the probability that the ball was from the bag I using the Bayes' Theorem.

- ***X** – ball is from bag I*
- ***Y** – ball is from bag II*
- ***A** – red ball*
- ***P(X) = P(Y) = ½***
- ***P(A|X) = 7/11, P(A|Y) = 5/14***

$$P(X|A) = \frac{P(A|X)*P(X)}{P(A|X)*P(X) + P(A|Y)*P(Y)} = \frac{7/11*1/2}{7/11*1/2 + 5/14*1/2} = \frac{0.31818181818}{0.49675324675} = 0.64052287581$$

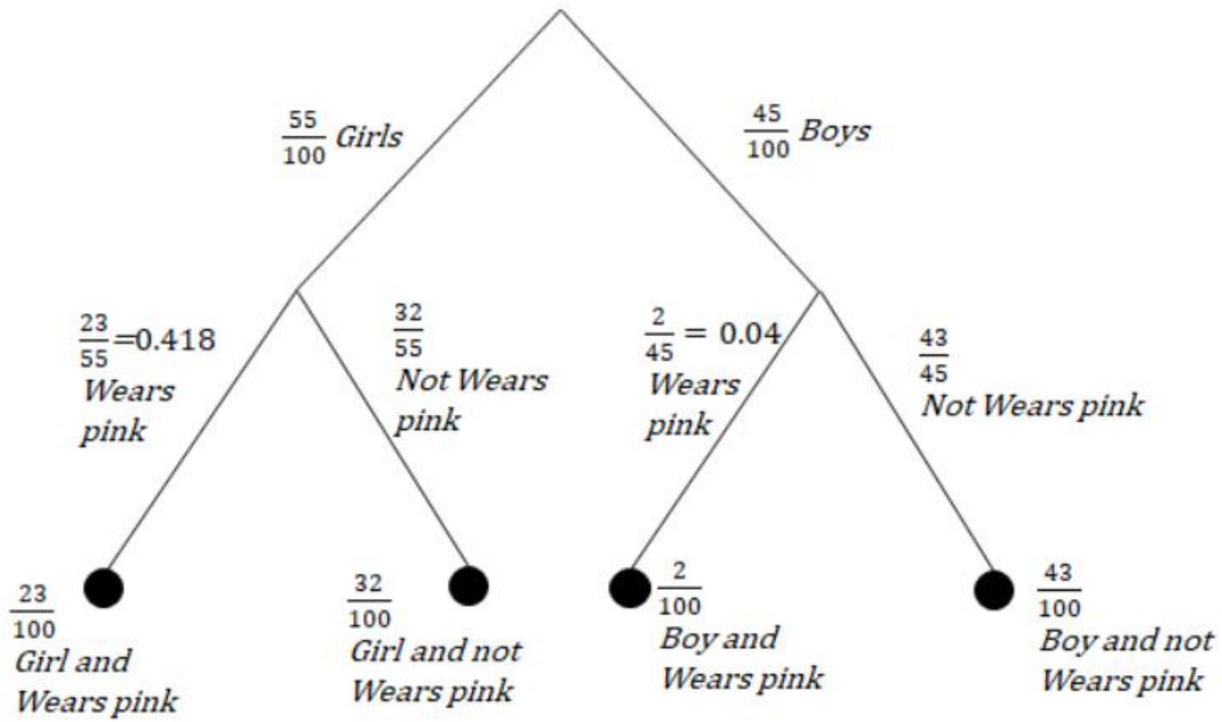


Graphical Representation

Example 4:

Calculate the probability that the student is a girl and she wears pink, **P(Wearing pink ∩ Girl)** or **P(Girl ∩ Wearing pink).** Girls are 55 out 100. Girls wearing pink are 23. Boys wearing pink are 2.

Tree Diagram for Bayesian Probability:



conditional probability of she wears pink given that the student is a gir

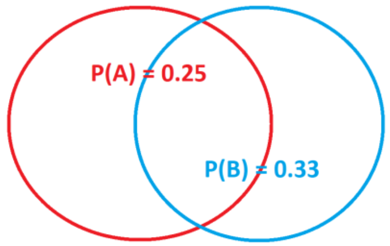
$$P(Wearing\ pink|\ Girl) = \frac{count(Wearing\ pink \cap Girl)}{count(Girl)}$$

$$= \frac{23}{55}$$

$$= \frac{count(Wearing\ pink \cap Girl)/100}{count(Girl)/100}$$

$$= \frac{P(Wearing\ pink \cap Girl)}{P(Girl)}$$

Summary



Conditional Probability:

- the ”*likelihood of an event occurring, based on the occurrence of another event*”

Bayes' Theorem

- principled way of calculating a conditional probability without the composite probability $P(E \cap F)$

$$P(E|F) = \frac{P(F|E)*P(E)}{P(F)}$$

- if we do not have access to the denominator $P(F)$ directly, we can calculate it:

$$P(F) = P(F|E)*P(E) + P(F|E^c)*P(E^c)$$

- new form:

$$P(E|F) = \frac{P(F|E)*P(E)}{P(F|E)*P(E) + P(F|E^c)*P(E^c)}$$

Graphical representation of Bayesian probabilities

Next Lesson – Model-Based Reasoning

Thank You!

Questions?