Bayes Everyday video

Book intro

> what is Bayes' theorem and how does it work
>
> is simple and built on elementary math
>
> different names: Bayes' rule, Bayes' theorem, Bayes' formula

used everywhere

> help display google search result
>
> same true to Netflix recommendations
>
> Hedge funds, self-driving cars, search and rescue

Nature

> a simple math formula revolutionalized how understand and deal with uncertainty

> if life is black and white, Bayes' Theorem help to think about the gray areas
>
> how much should we change our confidence in a belief when given new evidence

Examples

> a test for cancer came back positive, what is the probability of having cancer if the test is positive?
>
> what is the probability that the dog likes you given it licks you?
>
> what is the probability the stock prices will fall given interest rate rises?
>
> what is the probability you are truly drunk given the breathalyzer test is positive?

why visual?

> applying the theorem is not intuitive for most people

Visual Aids

> Venn Diagrams
>
> Decision Trees
>
> Letters (T, H for tail and head)
>
> Physical objects (real coins)

A Visual Intro Part 1

Bayes' theorem without formula video

Problem setting

> a box with 2 outcomes, both are equally likely to happen
>
> - p(A) = 1/2, p(B) = 1/2

> a box with 3 outcomes, all of them are equally likely to happen
>
> - p(A) = 1/3, p(B) = 1/3, p(C) = 1/3
>
> focus on first case, two types of cookies, chocolate and peanut
>
> - p(A) = 1/2, p(B) = 1/2
> - P(chocolate) = 3/4, P(peanut butter) = 1/4

Problem proposed

> prob(boxA | chocolate cookie) ?>? prob(boxB | chocolate cookie)
>
> **intuitive:** box A has double the amount of chocolate than box B

Explanation

> New Evidence change the universe
>
> - Evidence: given a chocolate cookie
> - Ignorance: peanut butter cookie (sample space changed)
> - new universe has only 15 chocolate cookies, A has 10, B has 5
>
> prob(boxA | chocolate cookie) = 2/3
>
> prob(boxB | chocolate cookie) =1/3

The Bayes' theorem Fomula video

3 ingredients of the formula

> meaning of each component of the formula
>
> posterior probability = normalized weighted average

solution with formula

> apply formula to the problem above
>
> likelihood
>
> - P(chocolate cookie | box A) = 100% sure = 1 (in Universe of box A)

> prior probability
>
> - $P(A) = P(boxA) = P(boxA|Universe)$ = 50% of universe = 0.5 (in Universe Original)
>
> intersection
>
> - $P(B|A)P(A) = P(A \cap B) = P(A \cap B|Universe)$ = 10/20 = 1/2 (in Universe Original)
>
> probability of data or evidence
>
> - P(B) = P(get a cholocate cookie | Universe) = # chocolate / # universe = 15/20 = 0.75 (in Universe Original)
>
> key intuition
>
> - $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ shift Universe from original to Chocolate cookies to look at intersection
>
> simple calculation process

Solving for one possible outcome with all data provided video

Senario 1: The Flu

Problem Setting

> “You have a headache and sore throat, and you know that people with the flu have the same symptoms roughly 90% of the time. In other words, 90% of people with the flu have the same symptoms you currently have. Does this mean you have the flu?”

Base Rate Fallacy

> **Don't fall into common error trap** to assume
>
> - P(A | B) = P(B | A)
> - P(headache and sore throat | Flu) = P(Flu | headache and sore throat)
> - so P(Flu | headache and sore throat) = 90% (**wrong**)

Posterior Probability

> what is probability you have a flu, given you find out you have a headache, sore throat?
>
> - P(Flu | headache and sore throat) **Posterior probability**

Prior Knowlege

> search in google to find more **prior knowledge** about the flu
>
> - P(Flu | all population in general) = 5%
>
> search in google to find **prior knowledge** about headache and sore throat
>
> - P(headache and sore throat | all population in general) = 20%

Likelihood

> personal belief about relationship between Flu and [headache with sore throat]
>
> - P(headache and sore throat | Flu) = 90% **current belief or likelihood**

Visualizing the problem and solution

> use Venn Diagram to show what we know
>
> $P(A \bigcap B)$ is **not known or given**, but formula does not need it
>
> apply to Bayes' theorem and calculate the answer
>
> many highly educated people don't know about Bayes' theorem

Scenario 2: Breathalyzer

Problem

> you are a police officer, “Around 2 am you randomly pull over a vehicle and have the driver take a breathalyzer test, and the result is positive. You assume the test is accurate and think nothing of it as you process the driver”
>
> are you really correct in assuming so?

Visualizing the solution

> prior knowledge on drunk driving and testing positive in general
>
> visualizing likelihood, causal effect from drunk driving to test result
>
> apply formula to the problem
>
> far cry difference between P(positve | drunk driving) vs P(drunk driving | positive)

Scenario 3: Peacekeeping – A surprise attack

problem

> "the probability of the truck being rigged with a gun given that we were just fired at with heavy firepower."

visualizing the solution

> prior knowledge P(a rebel truck with guns | all rebel trucks) = 40%
>
> prior knowledge or observations P(rebel truck with heavy firepower | all rebels trucks) = 50%
>
> likelihood P(rebel truck with heavy firepower | rebel with gun) = 80%
>
> apply formula to problem and calculate the result
>
> whether or not have a far cry difference depends on how much is updated on the belief

Solving for one possible outcome with No P(B) provided video

> previously 3 ingredients are given, now only 2 ingredients are given, P(B) must be discovered by ourselves
>
> decision tree is a powerful tool to help find P(B)

Scenario 1: The Flu

Problem

> prior probability
>
> - Google tells 5% of population will get Flu every year P(Flu | all) = 5%
>
> likelihood probabilities
>
> - P(positive | Flu) = 75% (correctly predict with Flu 75% of time)
> - P(positive | No Flu) = 20% (wrong predict with Flu 20% of time)
>
> what to know posterior probability
>
> - P(Flu | positive)

Visualizing the solution

> start with what we know, the most basic node (have two branches)
>
> - P(Flu | over all population)
> - P(No Flu | over all population)
>
> build upon primary branches, produce 4 secondary branches
>
> - branches from each node, sum to 1
> - having 3 nodes so far
>
> meaning of final nodes on the end of secondary branches
>
> - top final node = P(Flu) * P(positive | Flu) = P(Flu ^ Positive)
> - each final node = branches (on each node path) multiple together
> - final node 1 = path 1 = P(Positive ^ Flu)
> - final node 2 = path 2 = P(Negative ^ Flu)
> - final node 3 = path 3 = P(Positive ^ no Flu)
> - final node 4 = path 4 = P(Negative ^ no Flu)
> - all final nodes sum up to 1
>
> How to discover P(B)
>
> - P(B) = path 1 + path 3 = P(positive ^ Flu) + P(positive ^ no Flu)
>
> apply the formula and calc the result

Scenario 2: The Drunk Driver

problem

> prior knowledge
>
> - “Approximately 3 out of every 1000 drivers will drive while drunk. This is .3%.
>
> likelihood P(positive | drunk)
>
> - The breathalyzer test does not always detect a drunk person. This is not 100% like you both previously thought, but 98%.
>
> likelihood P(positive | not drunk)
>
> - 4% of the time breathalyzer tests give a positive result for someone who is not drunk. This is called a false positive.”
>
> goal = Posterior probability = P(drunk | positive)

Visualizing Solution

> decision tree solution
>
> apply the formula

Scenario 3 Peacekeeping – a surprise attack

problem

> prior knowledge: P(rebels) = 100/175
>
> - “There are roughly 100 rebels in the city and 75 coalition troops.”
>
> likelihood: P(positive | rebels) = 65%
>
> - “Local Intel is not always reliable. In your experience it correctly predicts rebels 65% of the time.”
>
> likelihood: P(positive | no rebels) = 15%
>
> - “Intel has been sketchy lately and has incorrectly predicted men as rebels when they are not rebels 15% of the time”
>
> what to know: Posterior probability = P(rebels | intel says positive)

> - “We want to know the probability that the soldiers are rebels given that Intel says they are.”

Visualizing Solution

> apply what we know above to a decision tree
>
> apply the formula

Solving for two possible outcomes with all probability data provided

> given symptoms (sneezing and coughing), previously we want to know P(cold | symptoms)
>
> now, we want to know P(cold | symptoms) vs P(allergy | symptoms)?
>
> so, we will use Bayes' formula twice

Scenario 1: The Flu

problem

> likelihood on Flu or food poison causing symptoms
>
> - P(symptoms | Flu) = 90%, P(symptoms | food poison) = 75%
> - “You have a slight headache and sore throat, and you see that people with the flu have the same symptoms as you 90% of the time. People with food poisoning have the same symptoms 75% of the time”
>
> prior knowledge on Flu or food poison
>
> - P(Flu | all population) = 5%, P(food poison | all population) = 16%
> - “You see that the probability of having the flu is 5%, while the probability of having food poisoning is 16%”
>
> prior knowledge or observations
>
> - P(symptoms = headache and sore throat | all population) = 20%

Visualizing Solutions

> using Venn Diagram twice
>
> apply the formula twice

Advanced usage

Search and Rescue

> location and time - find the lost person at sea
>
> SAROPS paper

spam filtering

> 1998 Microsoft use bayesian filter for spam emails
>
> microsoft paper on Bayesian filter

Driverless Car

> use a Bayesian model similar to hidden markov model for localization
>
> stanford driverless car page

3 steps to think like a Bayesian everyday video

scenario 1: Dating

Non-Visual Approach

> prior knowledge

> - P(a gilr likes you | all girls on date) = 20%
> - P(a gilr not like you | all girls on date) = 80%
> - number 1 = 2 (out of 10)
> - number 2 = 8 (out of 10)
>
> likelihood
>
> - P(laughing and flirting | she likes you) = 90%
> - number 3 = 9 (out of 10)
> - P(laughing and flirting | she not like you) = 10%
> - number 4 = 1 (out of 10)
>
> goal = P(she likes you | laughing and flirting)
>
> - option1 = P(laughing and flirting | she likes you) * P(a girl likes you) = 9 * 2 = 18 = P(laughing and flirting ^ she likes you)
> - option2 = P(laughing and flirting | she not like you) * P(a girl not like you) = 1 * 8= 8 = P(laughing and flirting ^ she not like you)
> - option1 : option2 = 9:4 > 50%

Visual Approach

> prior knowledge visualized
>
> likelihood visualized
>
> calculate option 1 and option 2 to compare

Scenario 2: Can you trust your mechanic?

Non-Visual approach

> prior knowledge
>
> - P(honest mechanics | all mechanics) = 70%
> - Number 1 = 7 (out of 10)
> - P(dishonest mechanics | all mechanics) = 30%
> - Number 2 = 3 (out of 10)
>
> likelihood
>
> - P(bad reviews | honest mechanics) = 30%
> - Number 3 = 3
> - P(bad reviews | dishonest mechanics) = 90%

> - Number 4 = 9
>
> goal
>
> - P(honest mechanics ^ bad reviews) = number 1 * number 3 = 21
> - P(dishonest mechanics ^ bad reviews) = number 2 * number 4 = 27
> - 27 > 21 = given bad reviews, more likely to run into a dishonest mechanics

Visual approach

> prior knowledge visualized
>
> likelihood visualized
>
> calculation visualized