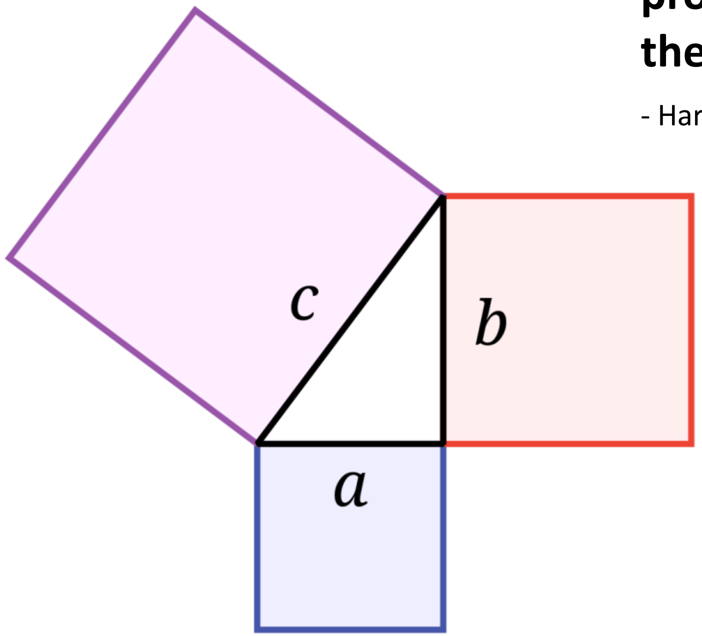
Pythagorean vs. Bayes

Pythagoras
c. 570–500

Thomas Bayes
c. 1701 – 1761

Pythagorean Theorem

The subject of the Pythagorean theorem is any right triangle.



$a^2 + b^2 = c^2$

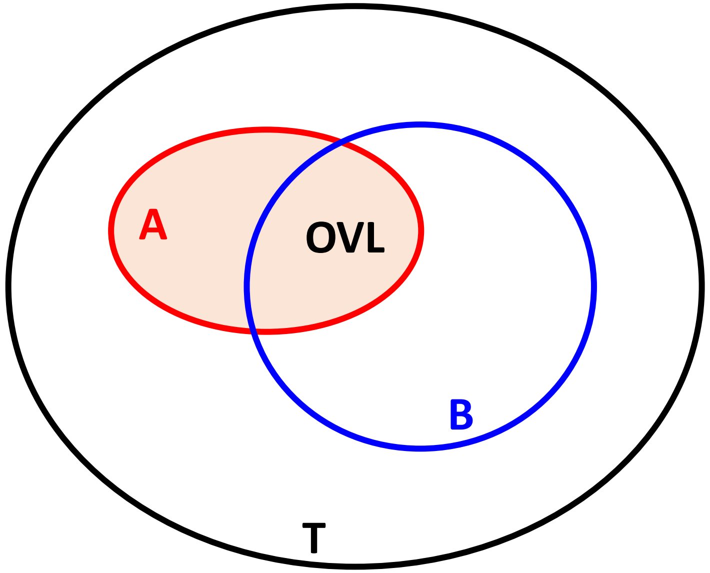
https://en.wikipedia.org/wiki/Pythagorean_theorem

Bayes' theorem is to the theory of probability what the Pythagorean theorem is to geometry.

- Harold Jeffreys (1973) *Scientific Inference*

Bayes’ theorem

The subject of Bayes’ Theorem is any set containing two subsets, where the two subsets partially overlap one another.



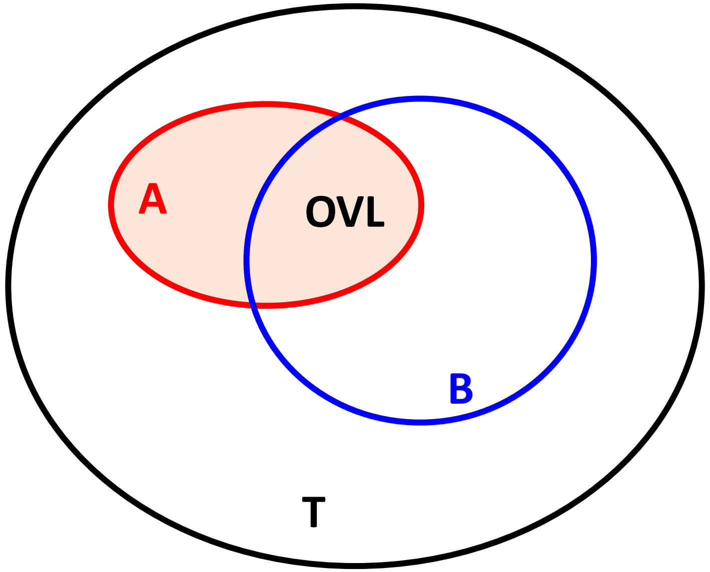
$$(OVL/A) = \frac{(OVL/B) \times (B/T)}{(A/T)}$$

https://theyhavenowine.wordpress.com/2020/11/11
/the-pythagorean-theorem-and-bayes-theorem/

From Bayes' Theorem to Bayesian Inference

Bayes' theorem

The subject of Bayes' Theorem is any set containing two subsets, where the two subsets partially overlap one another.



Algebraic form

$$(OVL/A) = \frac{(OVL/B) \times (B/T)}{(A/T)}$$

https://theyhavenowine.wordpress.com/2020/11/11/the-pythagorean-theorem-and-bayes-theorem/

ratio = probability

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

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

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

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

Bayesian inference

It derives the **posterior** probability as a consequence of two antecedents: a **prior** probability and a "**likelihood function**" derived from a statistical model for the observed data. – wikipedia.org

$$\underbrace{P(X \mid d)}_{\text{Posterior}} = \frac{\overbrace{P(d \mid X)}^{\text{Likelihood}} * \overbrace{P(X)}^{\text{Prior}}}{P(d)}$$

Let d be Data or Observation.

A guessing game with blindfolded

www.vectorstock.com

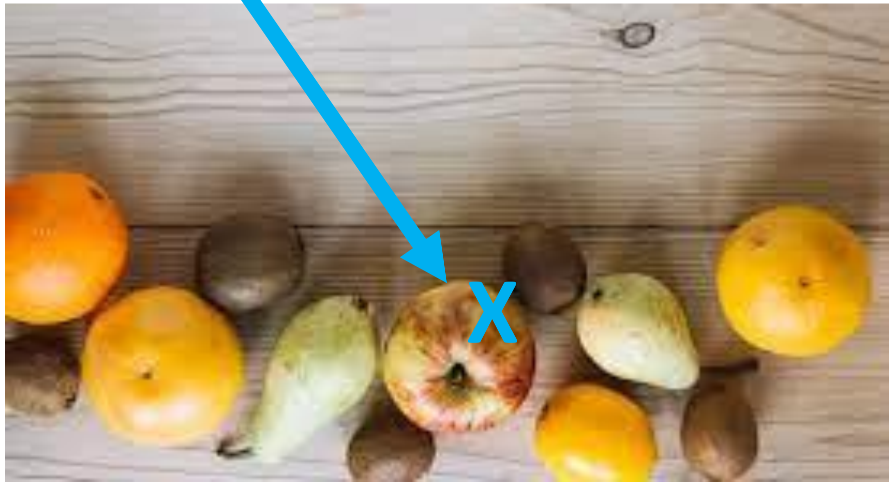
You (blind-folded) were led to a room in which lots of fruits were placed on a tabletop.

You:

Given a stick to touch an item.

You hear what your friends are saying about the touched item and guess what kind of fruit it is.

T_0 You touched the item X



www.freepik.com

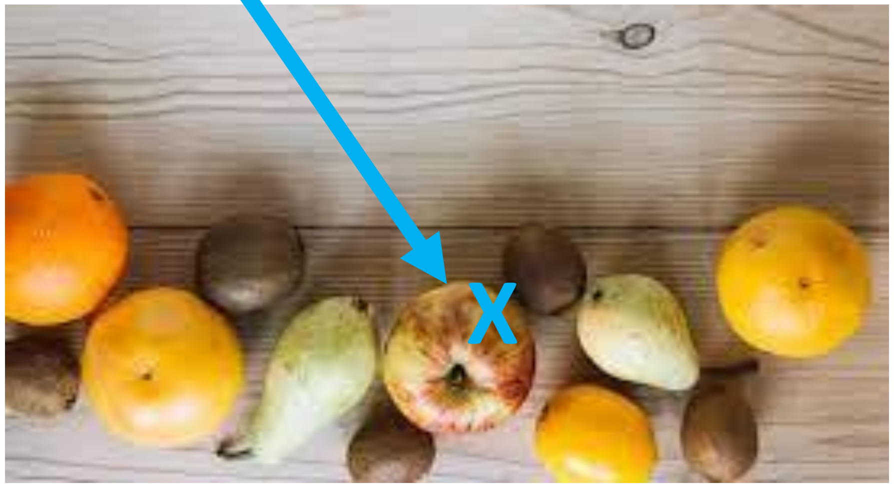
A guessing game with blindfolded

You(blind-folded) were led to a room in which lots of fruits were placed on a tabletop.

You:
Given a stick to touch an item.
You hear what your friends are saying about the touched item and guess what kind of fruit it is.

T_0 **You touched the item X.**

T_2 **You ponder if it is an apple.**



Friends:

Reporter – Sue says:
"The shape of the touched item is **round**." ***T_1***

A guessing game with blindfolded

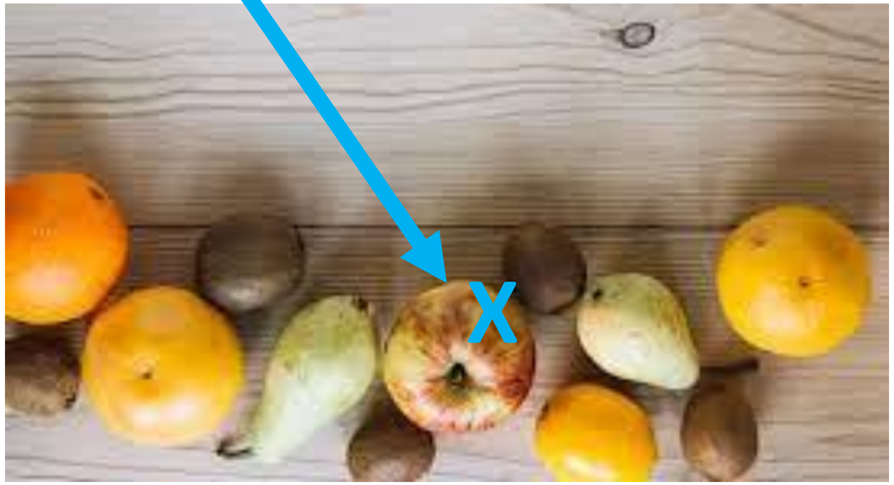
You(blind-folded) were led to a room in which lots of fruits were placed on a tabletop.

You:
Given a stick to touch an item.
You hear what your friends are saying about the touched item and guess what kind of fruit it is.

T_0 **You touched the item X.**

T_2 **You ponder if it is an apple.**

T_4 **You almost say "It is an apple".**



Friends:

Reporter – Sue
"The shape of the touched item is round." T_1

Helper 1 – Joe chimes in (reading your mind): T_3
"**All apples** are round."

A guessing game with blindfolded

You(blind-folded) were led to a room in which lots of fruits were placed on a tabletop.

You:
Given a stick to touch an item.
You hear what your friends are saying about the touched item and guess what kind of fruit it is.

Friends:

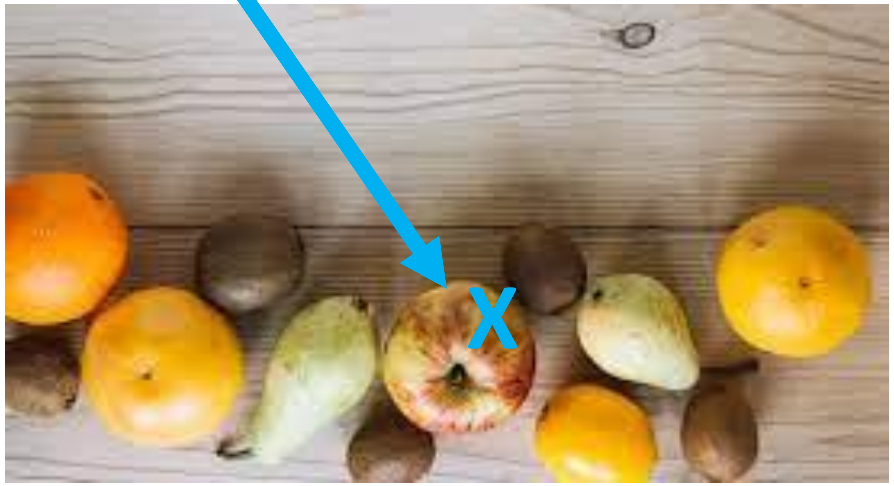
Reporter – Sue
"The shape of the touched item is round." T_1

T_0 **You touched the item X.**

T_2 **You ponder if it is an apple.**

Helper 1 – Joe chimes in. T_3
"All apples are round."

T_4 **You almost say "It is an apple".**



Helper 2 – Mary warns (reading your mind): T_5
"There are **other fruits** that are round!"

T_6 **You now try to be smart by thinking:**

(i) what other fruits are round; (ii) what other fruits could be on the table.

Do you agree with this scenario?
Quite a bit of inferencing involved.

A guessing game—mapped to Bayesian Thinking

You(blind-folded) were led to a room in which lots of fruits were placed on a tabletop.

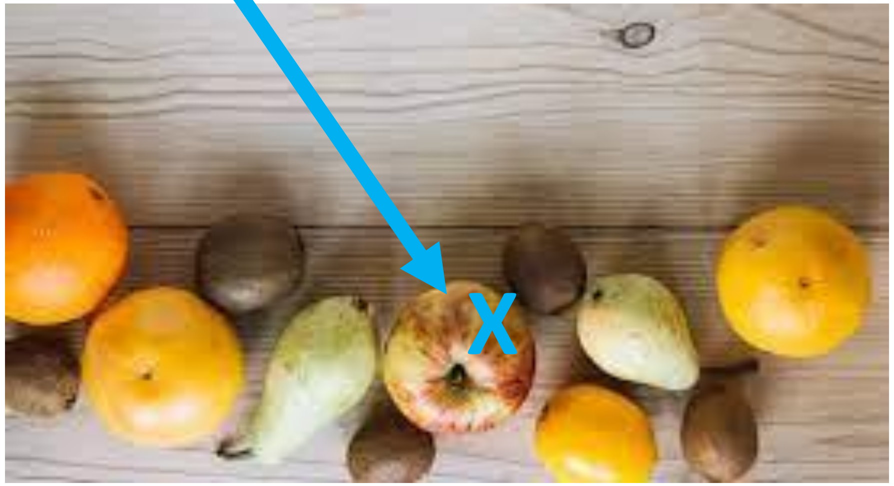
You:
Given a stick to touch an item.
You hear what your friends are saying about the touched item and guess what kind of fruit it is.

T_0 **You touched the item X.**

T_2 **You ponder if it is an apple.**

T_4 **You almost say "It is an apple".**

$P(X_{apple} | S_{round}) = 1$



T_6 **You now try to be smart by thinking:**

(i) what other fruits are round; (ii) what other fruits could be on the table.

$P(X_i)$

Friends:

Reporter – Sue says:
"The shape of the touched item is round." ***T_1***

Helper 1 – Joe warns (reading your mind): ***T_3***
"All apples are round."

$P(S_{round} | X_{apple}) = 1$

Helper 2 – Mary chimes in (reading your mind): ***T_5***
"There are **other fruits** that are round!"

$P(X_i | S_{round})$

The key to Bayesian Thinking is "Problem Modeling"!

What is X? What is d?

$$P(X \mid d) \overset{?}{\propto} P(X)$$

Posterior

Prior

$$\overset{?}{\propto} P(d \mid X)$$

Likelihood

$$P(X \mid d) = \frac{P(X) * P(d \mid X)}{P(d)}$$

Bayesian Thinking – Some observations

Posterior ? Prior Likelihood

$$P(X \mid d) \propto P(X) * P(d \mid X)$$

Obs 1

$$P(X \mid d) = \frac{P(X) * P(d \mid X)}{P(d)}$$

Denominator P(d) is needed if you wish to make P(X|d) range over (0, 1), otherwise not needed.

Obs 2

HW4 Part III

$$\frac{P(X_i \mid d)}{P(X_j \mid d)} = \frac{\dfrac{P(X_i) * P(d \mid X_i)}{\cancel{P(d)}}}{\dfrac{P(X_j) * P(d \mid X_j)}{\cancel{P(d)}}}$$

If you were to compare $P(X_i|d)$ vs. $P(X_j|d)$, i.e., compare if $P(X_i|d) > P(X_j|d)$ or $P(X_i|d) < P(X_j|d)$, you only need to calculate the ratio.

Would round fruit be apple or pear?

Obs 3 $P(d)$?

Often P(d) is unknown or hard to estimate.

How many different shapes needed for fruit and what are the proportion for each shape in the fruit world?

We do Bayesian thinking all the time without knowing P(d).

The key in Bayesian Thinking is in "Problem Modeling"!

Posterior ? Prior Likelihood

$$P(X \mid d) \propto P(X) * P(d \mid X) \qquad P(X \mid d) = \frac{P(X) * P(d \mid X)}{P(d)}$$

Example 1. A librarian or a salesperson?

You saw a person *x* who is shy, and you are asked to guess if *x* is a librarian or a salesperson.
Let C_1 = librarian, C_2 = sales person.

$$P(\text{librarian} \mid x) > P(\text{sales} \mid x) \text{ or}$$
$$P(\text{librarian} \mid x) < P(\text{sales} \mid x)$$

You will try to check what are the **likelihood** of shy person from the two groups → helps the decision.

$$P(C_i \mid x) \propto P(x \mid C_i)$$

Do you agree? Basically, you will try to determine *i* such that

$$\underset{i}{argmax}\, P(C_i \mid x)$$

$$P(x \mid \text{librarian}) = 0.9 \quad \text{and} \quad P(x \mid \text{sales}) = 0.2$$

You decide *i* should be librarian. → Would this type of reasoning correct? Anything wrong with it?

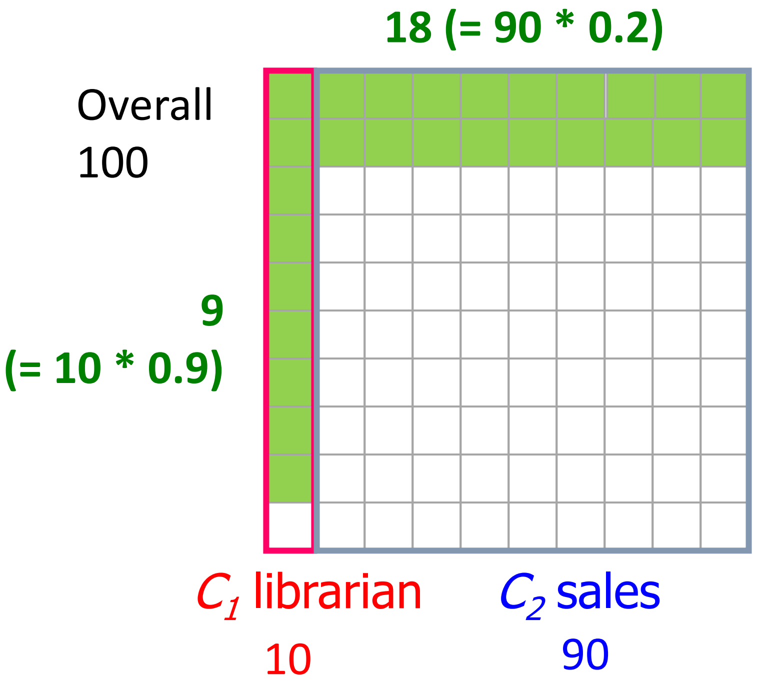
Example 1. A librarian or a salesperson?

It will be incorrect if you just do

$$\underset{i}{argmax}\, P\,(\,x\,|\,C_i\,)$$

since you think $P\,(C_i\,|\,x\,) \propto P\,(\,x\,|\,C_i\,)$

Why? Consider the case of 100 people where 10 are librarians and 90 are sales, of which the distributions for shy and not shy are 0.9 and 0.2, respectively.



The chance x being a librarian is 1/3 (9 out of 27) while chance being a salesperson is 2/3 (18 out of 27).
→ You need to take into account **marginal probabilities**, i.e.,

$$P\,(C_i\,|\,x\,) \propto P\,(\,x\,|\,C_i\,) * P\,(\,C_i\,)$$

*0.9 * 0.1 = 0.09*

*0.2 * 0.9 = 0.18*

And you should bet that x is more likely a salesperson.

Example 2 – Drug testing

What is the probability that a random person tested positive is really a cannabis user?

Cannabis use testing is 90% sensitive, meaning the true positive rate (TPR)=0.90.

→ 90% success in identifying cannabis users.

The test is also 80% specific, meaning true negative rate (TNR)=0.80.

→ 80% success in identifying non-use for non-users, but also generating 20% false positives, or false positive rate (FPR)=0.20, for non-users.

Assuming 0.05 prevalence, meaning 5% of people use cannabis.

Frequency box visualization

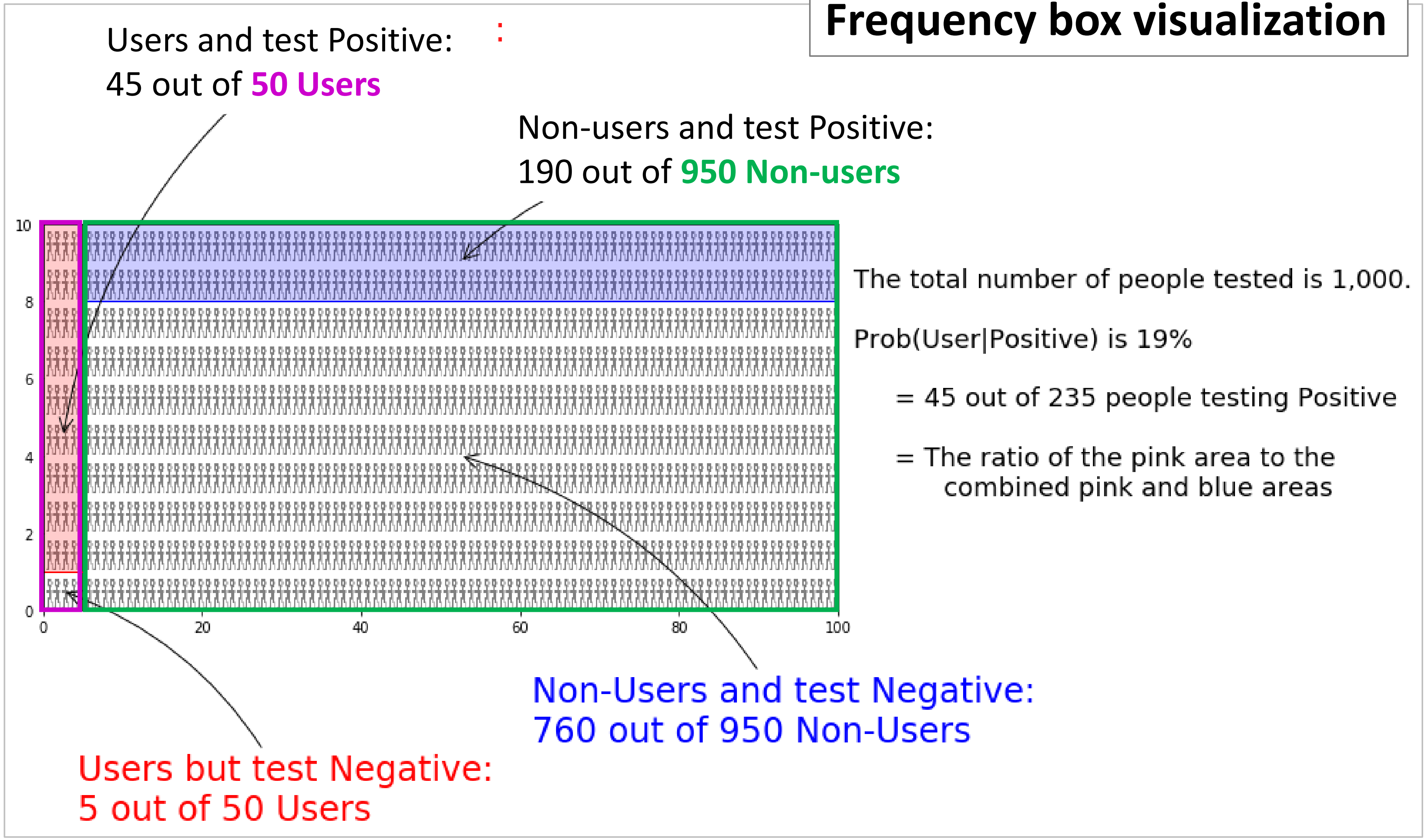


Figure 1: Using a frequency box to show visually by comparison of areas

https://en.wikipedia.org/wiki/Bayes%27_theorem

Example 2 – Drug testing

What is the probability that a random person tested positive is really a cannabis user?

Sensitivity: **TPR=0.90**. 90% success in identifying cannabis users (**45**/**50**).

Specificity: TNR=0.80. 80% specific, meaning true negative rate (**760**/**950**). (FPR)=0.20

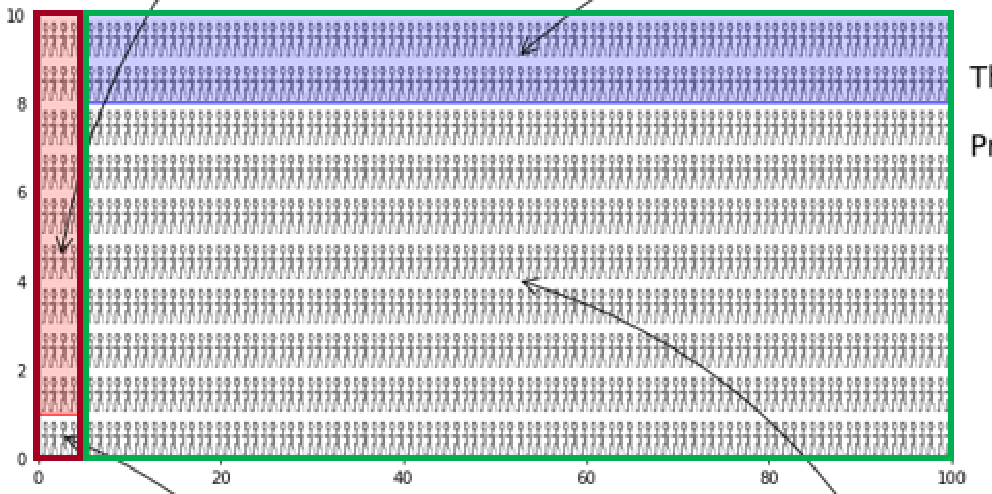
Prior: 0.05 prevalence
5% of people use cannabis.

		Test				
		yes	no			
	User	**45**	5	**50**	**0.9**	0.05
	Non-User	190	**760**	950	**0.2**	0.95
		235	765	**1000**		
		0.191				
	P(U)	0.05	prior			
	P(Y|U)	**0.9**	likelihood			
	P(U) update	0.045	likelihood updates prior			
	P(Y|NU)	**0.2**	false positive			
	1 -P(U)	**0.95**				
	DN 2nd Term	**0.19**				
	P(Y)	**0.235**	denominator summed			
	P(U|Y)	**0.191**	posterior			

Frequency box visualization

TP **Users** and test **Positive**: **45** out of **50 Users**

Non-users and test **Positive**: **190** out of **950 Non-users** FP



The total number of people tested is 1,000.

Prob(User|Positive) is 19%

= 45 out of 235 people testing Positive

= The ratio of the pink area to the combined pink and blue areas

FN **Users** but test **Negative**: 5 out of **50 Users**

Non-users and test **Negative**: 760 out of **950 Non-users** TN

https://en.wikipedia.org/wiki/Bayes%27_theorem

$$P(U \mid Y) = \frac{P(U) * P(Y \mid U)}{P(Y)} = \frac{0.05 * 0.9}{0.235} = \frac{0.045}{0.235} = 0.1915$$

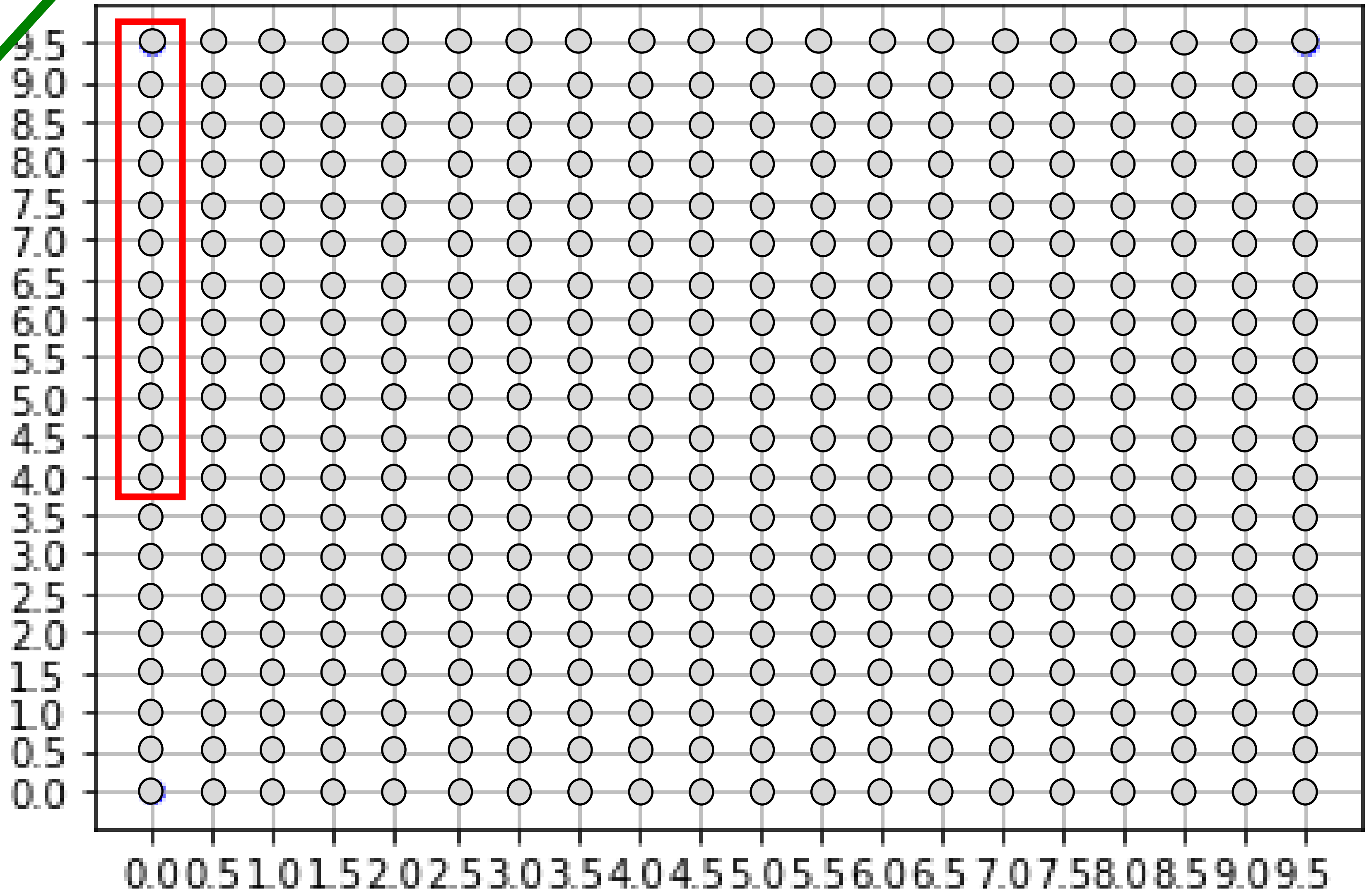
$$P(Y) = P(U) * P(Y \mid U) + ((1 - P(U)) * P(Y \mid NU)) = (0.05 * 0.9) + (0.95 * 0.2) = 0.19$$

HW4 Part III Help

400 dots, not 1000

400 dots = 400 people (Tested population)

		Test				
		yes	no			
	User	**45**	5	**50**	0.9	0.05
	Non-User	190	**760**	950	0.2	0.95
		235	765	**1000**		
		0.191				
	P(U)	0.05	prior			
	P(Y|U)	**0.9**	likelihood			
	P(U) update	0.045	likelihood updates prior			
	P(Y|NU)	**0.2**	false positive			
	1 -P(U)	**0.95**				
	DN 2nd Term	**0.19**				
	P(Y)	**0.235**	denominator summed			
	P(U|Y)	**0.191**	posterior			



You can use BOX (red) to highlight the group you like to emphasize!

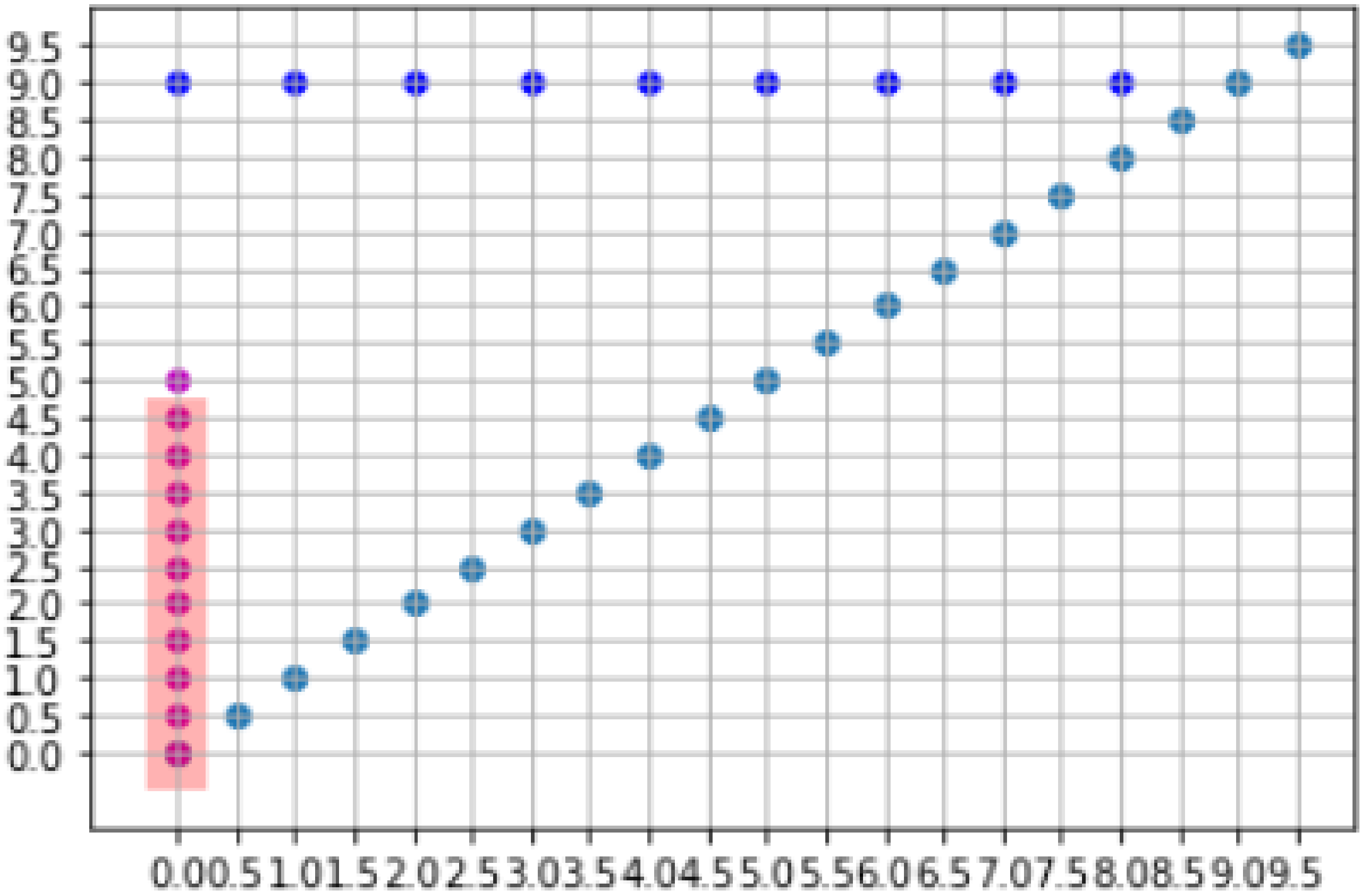
```
# Sample code placing dots on grid and highligting portion of diplayed dots.
import numpy as np
from matplotlib import pyplot as plt
from matplotlib import image as image
from matplotlib.patches import Rectangle

# Create a grid that can place at least 400 (=20x20) dots.
x = np.arange(0, 10, 0.5)
y = np.arange(0, 10, 0.5)

# Place TP/FP dots.
xt = np.array([0, 0,   0, 0,   0, 0,   0, 0,   0, 0,   0])
yt = np.array([0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5])

# Place TN/FN dots.
xfp = np.array([0, 1, 2, 3, 4, 5, 6, 7, 8])
yfp = np.array([9, 9, 9, 9, 9, 9, 9, 9, 9])

# plt.xlabel('x') # putting x-axis labels
# plt.ylabel('y') # putting y-axis labels

fig = plt.figure()
ax = fig.gca()
ax.set_xticks(np.arange(0, 10, 0.5))
ax.set_yticks(np.arange(0, 10, 0.5))
plt.scatter(x, y) # test placing dots.
plt.scatter(xt, yt, color = "m", marker = "o", s = 30)
plt.scatter(xfp, yfp, color = "b", marker = "o", s = 30)
plt.grid()

# Highlighting cases, e.g., TP, FP, TN, or FN.
someX=x[0]
someY=y[0]
width, height = 0.5, 5.25
currentAxis = plt.gca()
currentAxis.add_patch(Rectangle((someX-0.25, someY-0.5), width, height, facecolor='red', alpha=0.3))
plt.show()
```

HW4 Hint

Convention:

Where to place "colored" dots for TP/FP/TN/FN? One color per group.

You can use BOX (red) to highlight the secondary grouping you like to emphasize! (e.g., Cancer group = TP ∪ FP)



The key in Bayesian Thinking is in "Problem Modeling"!

Posterior ? Prior Likelihood

$$P(X \mid d) \propto P(X) * P(d \mid X)$$

$$P(X \mid d) = \frac{P(X) * P(d \mid X)}{P(d)}$$

Example 3. Stock market analysis

"What is the probability of AMZN stock price falling given that the Dow Jones Industrial Average (DJIA) index fell earlier?"

P(AMZN): the probability that AMZN falls.

P(DJIA): the probability that the DJIA fell.

P(AMZN|DJIA) = P(AMZN and DJIA) / P(DJIA)

"the probability that AMZN drops given a DJIA decline is equal to the probability that **AMZN price declines and DJIA declines** over the probability of a decrease in the DJIA index.

$$P(AMZN|DJIA) = \frac{P(AMZN) \times P(DJIA|AMZN)}{P(DJIA)}$$

The formula explains the relationship between the probability of the hypothesis before seeing the evidence that P(AMZN), and the probability of the hypothesis after getting the evidence P(AMZN|DJIA), given a hypothesis for Amazon given evidence in the Dow.

?????

https://www.investopedia.com/terms/b/bayes-theorem.asp