- Terms & Definitions
- Bayes Theorem - Simple Example
- Bayes Classifier – SOS Example
- Q. Assignment 1

Tutorial 1
CS-473

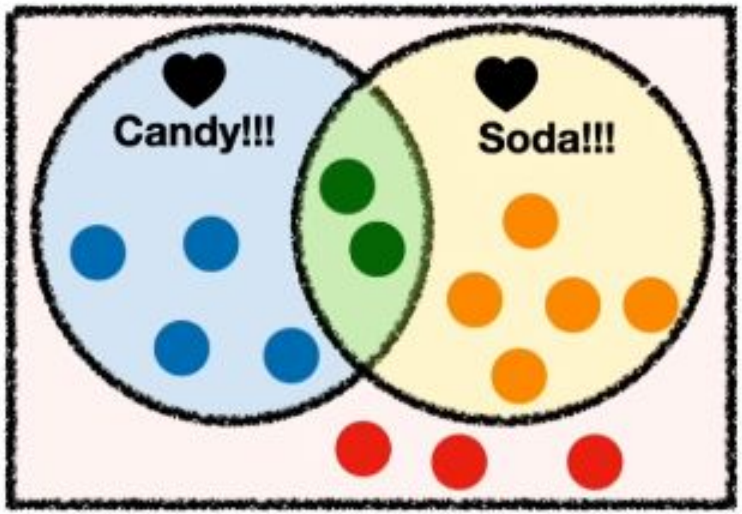
## Why?

P(A): Cloudy
P(B): Raining

P(B|A): Raining given day is cloudy
P(A|B): Cloudy given it is raining

## Probabilities

- P(Loves candy) = 6/14 {**P(A)**}
- P(Loves soda) = 7/14 {**P(B)**}
- P(Loves candy and not soda) = 4/14
- P(Loves soda and not candy)= 5/14
- P(Loves candy and love soda)= 2/14
- P(Doesn't love candy and doesn't love soda)= 3/14




## Conditional Probabilities

- p(loves candy and loves soda | loves soda) = 2/7
- p(loves candy and loves soda | loves candy) = 2/6

**Conditional Probability Formula**

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

Probability of A given B = Probability of A and B / Probability of B

## Bayes theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Bayes' Theorem EXPLAINED with Examples

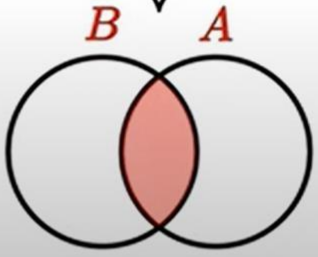
# Bayes' Theorem

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

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

$$P(B \cap A)$$

$$\Downarrow$$




$$P(A) \cdot P(B|A) = \frac{P(A \cap B)}{\cancel{P(A)}} \cdot \cancel{P(A)}$$

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

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

https://www.youtube.com/watch?v=cqTwHnNbc8g

4:01 / 8:02

Gaussian Naive Bayes is a machine learning classification technique based on a probabilistic approach that assumes each class follows a normal distribution. It assumes each parameter has an independent capacity of predicting the output variable.

Why use it?

- Speed: It’s lightning-fast compared to more complex AI models.
- Small Data: It works well even if you don't have millions of examples to train it on.
- Great for Text: It’s the "Gold Standard" for basic sentiment analysis (is this review happy or sad?) and spam detection.

**Bayesian Classifier**

$$P(C_i|x) = \frac{p(x|C_i)P(C_i)}{p(x)}$$

**Assignment 1:**
- C0: no heart disease
- C1: heart disease

**Naive Bayesian Classifier**

$$P(x \mid C) = \prod_{i=1}^{k} P(x_i \mid C)$$

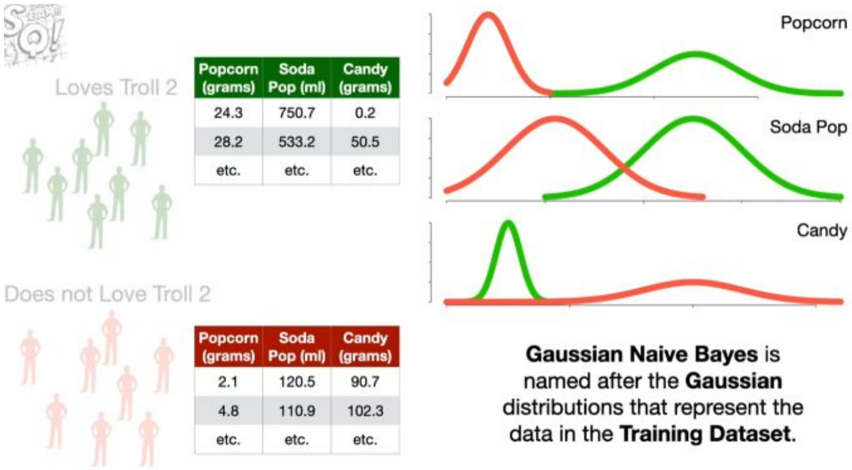
$$P(x \mid C) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}\right)$$

**Full Bayesian Classifier**

$$P(x \mid C) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

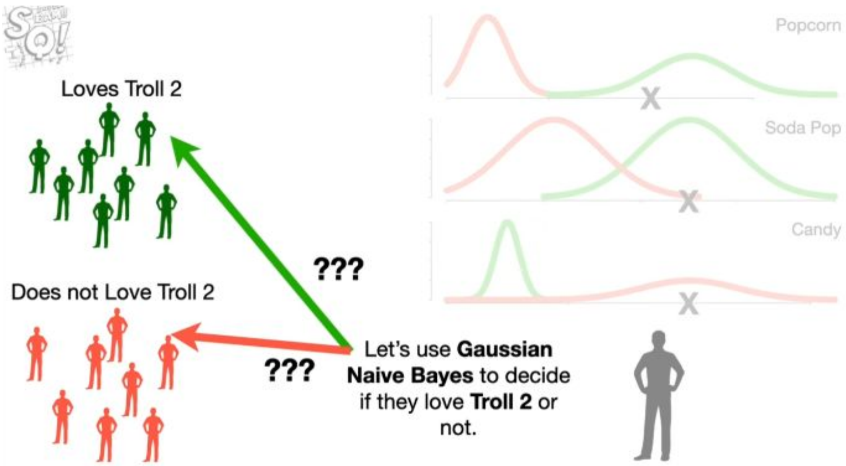
For assignment 1, we have 3 variables:
- The Vector **x**: For each individual in the dataset, **x** is a vector containing three values.
- The Mean **μ**: This is also a 3-dimensional vector representing the average values for a specific class (Heart Disease or No Heart Disease).
- The Covariance Matrix **Σ**: This is a 3x3 matrix that describes how those three variables relate to each other within that class.

# Gaussian Naive Bayes Classifier




| Popcorn (grams) | Soda Pop (ml) | Candy (grams) |
|---|---|---|
| 24.3 | 750.7 | 0.2 |
| 28.2 | 533.2 | 50.5 |
| etc. | etc. | etc. |

| Popcorn (grams) | Soda Pop (ml) | Candy (grams) |
|---|---|---|
| 2.1 | 120.5 | 90.7 |
| 4.8 | 110.9 | 102.3 |
| etc. | etc. | etc. |



# Gaussian Naive Bayes Classifier for Continuous Data

- Continuous values associated with each feature are assumed to be distributed based on Gaussian distribution.
- Likelihood of the features is assumed to be Gaussian.

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| Male | 6 | 180 | 12 |
| Male | 5.92 | 190 | 11 |
| Male | 5.58 | 170 | 12 |
| Male | 5.92 | 165 | 10 |
| Female | 5 | 100 | 6 |
| Female | 5.5 | 150 | 8 |
| Female | 5.42 | 130 | 7 |
| Female | 5.75 | 150 | 9 |

# Gaussian Naive Bayes Classifier for Continuous Data

- Continuous values associated with each feature are assumed to be distributed based on Gaussian distribution.
- Likelihood of the features is assumed to be Gaussian.

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| Male | 6 | 180 | 12 |
| Male | 5.92 | 190 | 11 |
| Male | 5.58 | 170 | 12 |
| Male | 5.92 | 165 | 10 |
| Female | 5 | 100 | 6 |
| Female | 5.5 | 150 | 8 |
| Female | 5.42 | 130 | 7 |
| Female | 5.75 | 150 | 9 |

- Example:
- Classify whether a given person's datapoint is a male or a female.
- The features include height, weight, and foot size.

# Gaussian Naive Bayes Classifier for Continuous Data

- Continuous values associated with each feature are assumed to be distributed based on Gaussian distribution.
- Likelihood of the features is assumed to be Gaussian.

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| Male | 6 | 180 | 12 |
| Male | 5.92 | 190 | 11 |
| Male | 5.58 | 170 | 12 |
| Male | 5.92 | 165 | 10 |
| Female | 5 | 100 | 6 |
| Female | 5.5 | 150 | 8 |
| Female | 5.42 | 130 | 7 |
| Female | 5.75 | 150 | 9 |

- Example:
- Classify whether a given person's datapoint is a male or a female.
- The features include height, weight, and foot size.

- P(Male) = 4/8 = 0.5 ; P(Female) = 4/8 = 0.5

# Gaussian Naive Bayes Classifier for Continuous Data

- Continuous values associated with each feature are assumed to be distributed based on Gaussian distribution.
- Likelihood of the features is assumed to be Gaussian.

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| Male | 6 | 180 | 12 |
| Male | 5.92 | 190 | 11 |
| Male | 5.58 | 170 | 12 |
| Male | 5.92 | 165 | 10 |
| Female | 5 | 100 | 6 |
| Female | 5.5 | 150 | 8 |
| Female | 5.42 | 130 | 7 |
| Female | 5.75 | 150 | 9 |

- Example:
- Classify whether a given person's datapoint is a male or a female.
- The features include height, weight, and foot size.

- P(Male) = 4/8 = 0.5 ; P(Female) = 4/8 = 0.5

**Male:**

- Mean(height) = 5.855
- Variance(Height) = $\frac{\Sigma(x_i-\overline{x})^2}{n-1} = \frac{(6-5.855)^2+(5.92-5.855)^2+(5.58-5.855)^2+(5.92-5.855)^2}{4-1}$ =0.035055
- Calculate for all features

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

# Gaussian Naive Bayes Classifier for Continuous Data

- Continuous values associated with each feature are assumed to be distributed based on Gaussian distribution.
- Likelihood of the features is assumed to be Gaussian.

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| Male | 6 | 180 | 12 |
| Male | 5.92 | 190 | 11 |
| Male | 5.58 | 170 | 12 |
| Male | 5.92 | 165 | 10 |
| Female | 5 | 100 | 6 |
| Female | 5.5 | 150 | 8 |
| Female | 5.42 | 130 | 7 |
| Female | 5.75 | 150 | 9 |

- Example:
- Classify whether a given person's datapoint is a male or a female.
- The features include height, weight, and foot size.

- P(Male) = 4/8 = 0.5 ; P(Female) = 4/8 = 0.5

**Male:**

- Mean(height) = 5.855
- Variance(Height) = $\frac{\Sigma(x_i-\overline{x})^2}{n-1} = \frac{(6-5.855)^2+(5.92-5.855)^2+(5.58-5.855)^2+(5.92-5.855)^2}{4-1}$ =0.035055
- Calculate for all features

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

# Gaussian Naive Bayes Classifier for Continuous Data

- Continuous values associated with each feature are assumed to be distributed based on Gaussian distribution.
- Likelihood of the features is assumed to be Gaussian.
- So, conditional probability is given by: $P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}}$

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| Male | 6 | 180 | 12 |
| Male | 5.92 | 190 | 11 |
| Male | 5.58 | 170 | 12 |
| Male | 5.92 | 165 | 10 |
| Female | 5 | 100 | 6 |
| Female | 5.5 | 150 | 8 |
| Female | 5.42 | 130 | 7 |
| Female | 5.75 | 150 | 9 |

- Example:
- Classify whether a given person's datapoint is a male or a female.
- The features include height, weight, and foot size.

- P(Male) = 4/8 = 0.5 ; P(Female) = 4/8 = 0.5

**Male:**

- Mean(height) = 5.855
- Variance(Height) = $\frac{\Sigma(x_i-\overline{x})^2}{n-1} = \frac{(6-5.855)^2+(5.92-5.855)^2+(5.58-5.855)^2+(5.92-5.855)^2}{4-1}$ =0.035055
- Calculate for all features

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

# Gaussian Naive Bayes Classifier for Continuous Data

- Classify the new datapoint

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| ? | 6 | 130 | 8 |

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

# Gaussian Naive Bayes Classifier for Continuous Data

- Classify the new datapoint

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| ? | 6 | 130 | 8 |

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

- $P(\text{male}) = \dfrac{P(\text{male}) * P(\text{H}|\text{male})\, P(\text{w}\,|\text{male})\, P(\text{F}|\text{male})}{\text{Marginal probability or Evidence}}$
- $P(\text{female}) = \dfrac{P(\text{female}) * P(\text{H}|\text{female})\, P(\text{w}\,|\text{female})\, P(\text{F}|\text{female})}{\text{Marginal probability or Evidence}}$

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}}$$

# Gaussian Naive Bayes Classifier for Continuous Data

- Classify the new datapoint

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| ? | 6 | 130 | 8 |

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

- $\text{P(male)} = \dfrac{\text{P(male)} * \text{P(H|male) P(W|male) P(F|male)}}{\text{Marginal probability or Evidence}}$
- $\text{P(female)} = \dfrac{\text{P(female)} * \text{P(H|female) P(W|female) P(F|female)}}{\text{Marginal probability or Evidence}}$

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}}$$

- The evidence (normalizing constant) is the sum of the posteriors equals one.
- evidence=P(male)*P(ht|male)*P(wt|male)*P(footsize|male)+P(female)*P(ht|female)*P(wt|female)*P(foot size|female)
- The evidence may be ignored since it is a positive constant. (Normal distributions are always positive)

# Gaussian Naive Bayes Classifier for Continuous Data

- Classify the new datapoint

| Gender | Height (ft) | Weight (lbs) | Foot size (inch) |
|---|---|---|---|
| ? | 6 | 130 | 8 |

| Gender | Mean (height) | Variance (Height) | Mean (weight) | Variance (Weight) | Mean (Footsize) | Variance (Footsize) |
|---|---|---|---|---|---|---|
| Male | 5.855 | 0.035 | 176.25 | 122.92 | 11.25 | 0.91667 |
| Female | 5.4175 | 0.097225 | 132.5 | 558.33 | 7.5 | 1.6667 |

- $P(\text{male}) = \dfrac{P(\text{male}) * P(\text{H}|\text{male})\, P(\text{w}\,|\text{male})\, P(\text{F}|\text{male})}{\text{Marginal probability or Evidence}}$
- $P(\text{female}) = \dfrac{P(\text{female}) * P(\text{H}|\text{female})\, P(\text{w}\,|\text{female})\, P(\text{F}|\text{female})}{\text{Marginal probability or Evidence}}$

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}}$$

- The evidence (normalizing constant) is the sum of the posteriors equals one.
- evidence=P(male)*P(ht|male)*P(wt|male)*P(footsize|male)+P(female)*P(ht|female)*P(wt|female)*P(foot size|female)
- The evidence may be ignored since it is a positive constant. (Normal distributions are always positive)

- $P(H|M) = \dfrac{1}{\sqrt{2*3.14*0.035033}} * e^{-\frac{(6-5.855)^2}{2*0.035033}} = 1.5789$ P(H|F) = 2.2356e$^{-1}$
- P(W|M) = 5.9881e$^{-6}$ P(W|F)=1.6789e$^{-2}$
- P(Foot|M) = 1.3112e$^{-3}$ P(Foot|F) = 2.8669e$^{-1}$

- $P(\text{male}) = \dfrac{P(\text{male}) * P(\text{H}|\text{male})\, P(\text{w}\,|\text{male})\, P(\text{F}|\text{male})}{\text{Marginal probability or Evidence}}$=0.5*1.5789*5.9881e$^{-6}$*1.3112e$^{-3}$=6.1984e-9
- $P(\text{female}) = \dfrac{P(\text{female}) * P(\text{H}|\text{female})\, P(\text{w}\,|\text{female})\, P(\text{F}|\text{female})}{\text{Marginal probability or Evidence}}$=0.5*2.2346e$^{-1}$*1.6789e$^{-2}$*2.8669e$^{-1}$=**5.377e$^{-4}$**
- **P(female)** > P(male)

## Useful sources

1. Bayes theorem, the geometry of changing beliefs (3Blue1Brown).
   https://youtu.be/HZGCoVF3YvM?si=qNLJ51E-AKJrQpbh

2. Gaussian Naive Bayes, Clearly Explained!!! (StatQuest)
   https://youtu.be/H3EjCKtlVog?si=7dAPkznMnVqiZ_Qs

**Good luck!**