CPE412 Pattern Recognition

Week 3

Probability and Statistics (Bayesian Decision Theory)

Dr. Nehad Ramaha,
Computer Engineering Department
Karabük Universities

Probability

Idea of Probability

- Probability is the science of chance behavior
- Chance behavior is unpredictable in the short run but has a regular and predictable pattern in the long run

Randomness

- Random: individual outcomes are uncertain
- But there is a regular distribution of outcomes in a large number of repetitions.
- Example: select any number from a bag of numbers {1,2,3,…,100}

Random Experiment…

… a ***random experiment*** is an action or process that leads to one of several possible outcomes. For example:

- If an experiment has n possible outcomes **[all equally likely to occur]**.

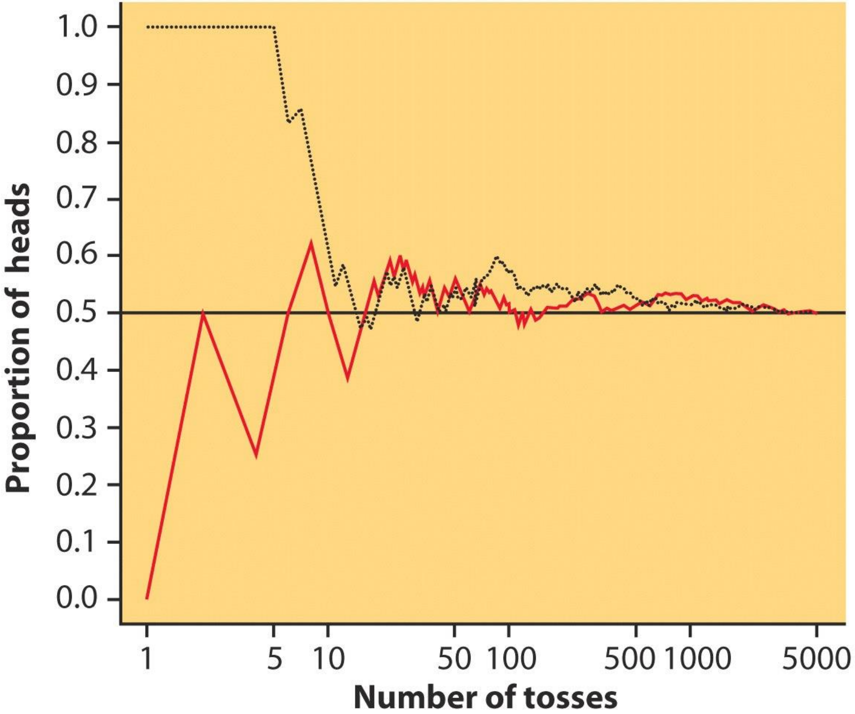
Experiment	Outcomes
Flip a coin	Heads, Tails
Selecting a color ball	Green, red, blue
Rolling a die	1,2,3,4,5,6
Picking a card from a deck	52 cards

Relative-Frequency Probabilities

- *Relative frequency* (proportion of occurrences) of an outcome settles down to one value over the long run. That *one value* is then defined to be the ***probability*** of that outcome.
- Can be determined (or checked) by observing a long series of *independent* trials (empirical data)
  - experience with many samples
  - simulation

Relative-Frequency Probabilities

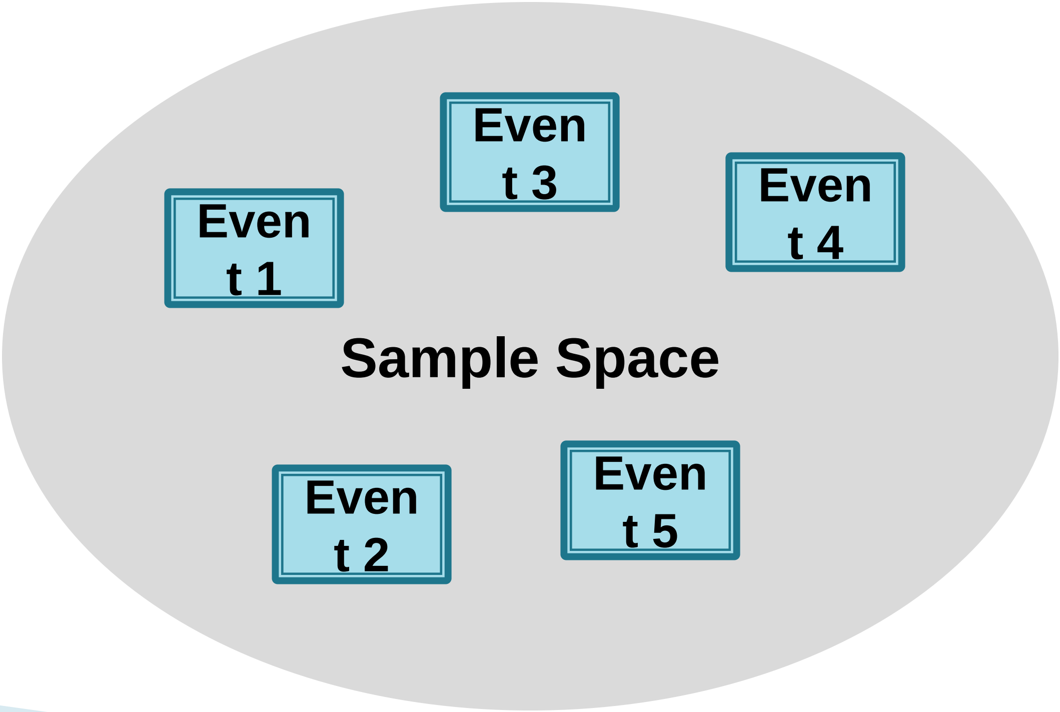
Coin flipping:



Probability Models

- The sample space S of a random phenomenon is the set of all possible outcomes.
- An event is an outcome or a set of outcomes (subset of the sample space).
- A probability model is a mathematical description of long-run regularity consisting of a <u>sample space S</u> and a <u>way of assigning probabilities</u> to events.

Sample Space and Events



Example

Rolling an odd number={2,4,6 }

Rolling an even number={2,4,6}

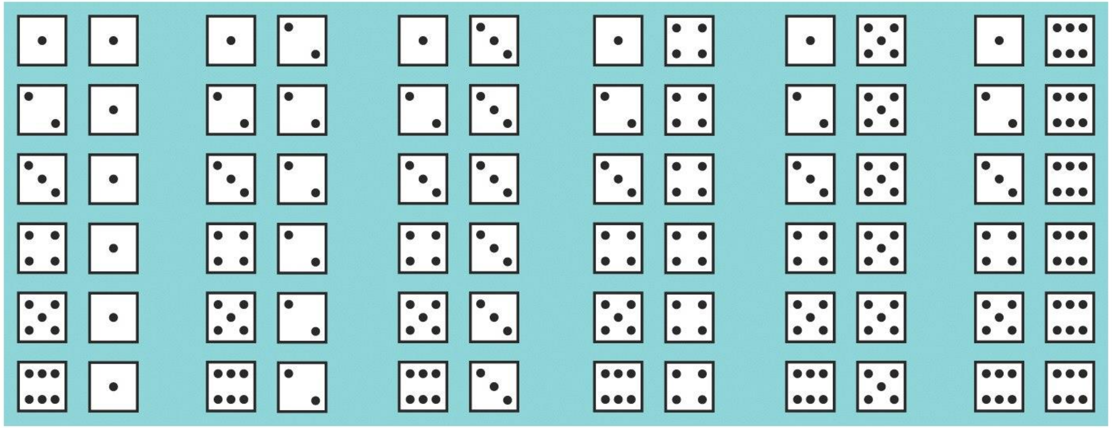
Sample Space ={1,2,3,4,5,6}

Rolling a prime number={2,3,5}

Probability Model for Two Dice

Random phenomenon: roll pair of fair dice.

Sample space:

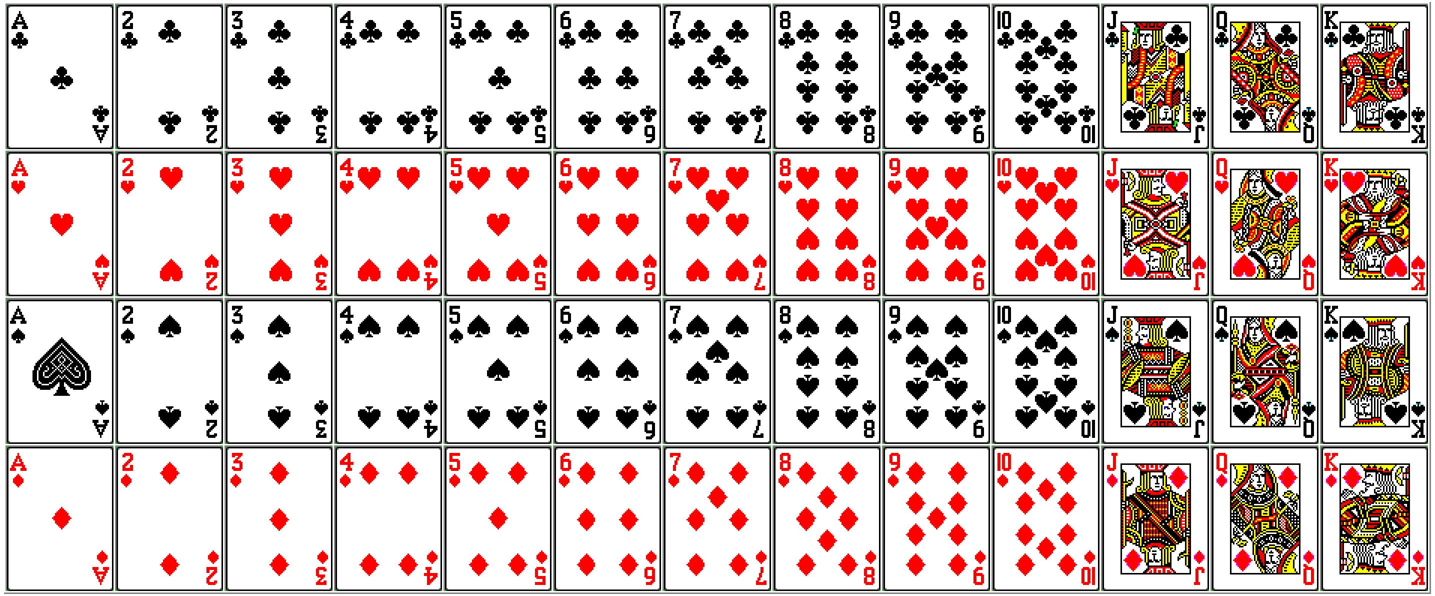


Event: rolling even numbers on both dice

Probability Model for 52 card deck

Random phenomenon: Arrange 52 card deck in a zigzag way

Sample space:



Event: pick an ace

Probability

What is a <u>PROBABILITY</u>?

- Probability is the chance that some event will happen

- It is the ratio of the number of ways a certain event can occur to the number of possible outcomes

Probability

What is a PROBABILITY?

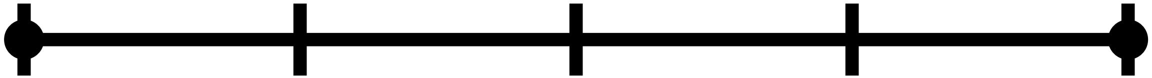
$$P(\text{event}) = \frac{\text{number of favorable outcomes}}{\text{number of possible outcomes}}$$

Examples that use Probability:

(1) Dice, (2) Spinners, (3) Coins, (4) Deck of Cards, (5) Evens/Odds, (6) Alphabet, etc.

Probability

What is a PROBABILITY?



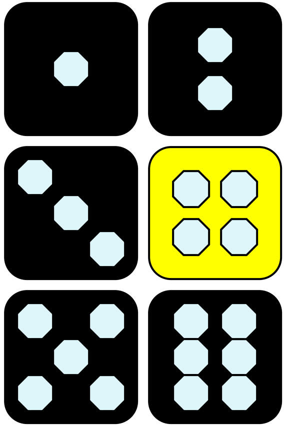
0	¼ or .25	½ or .5	¾ or .75	1
Impossible Certain	Not Very Likely	Equally Likely	Somewhat Likely	

Probability of Simple Events

<u>Example 1</u>: Roll a dice.

What is the probability of rolling a 4?

$$P(event) = \frac{number\ of\ favorable\ outcomes}{number\ of\ possible\ outcomes}$$



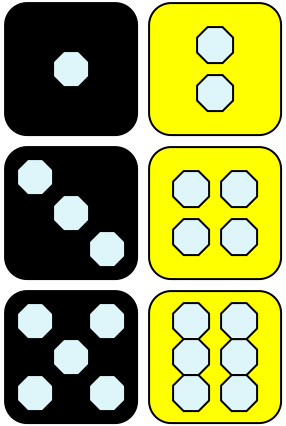
$$P(rolling\ a\ 4) = \frac{1}{6}$$

The probability of rolling a 4 is 1 out of 6

Probability of Simple Events

<u>Example 2</u>: Roll a dice.
What is the probability of rolling an even number?

$$P(even\ \#) = \frac{\#\ favorable\ outcomes}{\#possible\ outcomes} = \frac{3}{6}$$



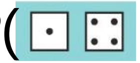
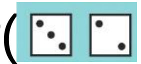
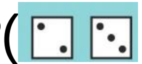
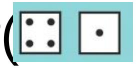
The probability of rolling an even number is 3 out of 6.

Probability of Simple Events

Example 3: Roll a dice.

Random phenomenon: roll pair of fair dice and count the number of pips on the up-faces.

Find the probability of rolling a 5.

P(roll a 5) = P() + P() + P() + P()

= 1/36 + 1/36 + 1/36 + 1/36

= 4/36

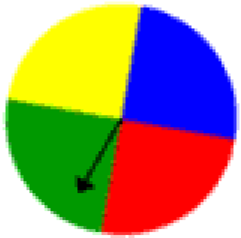
= 0.111

Probability of Simple Events

Example 4: Spinners.

What is the probability of spinning green?

$$P(green) = \frac{\#\ favorable\ outcomes}{\#possible\ outcomes} = \frac{1}{4}$$



The probability of spinning green is 1 out of 4

Probability of Simple Events

Example 5: Flip a coin.

What is the probability of flipping a tail?

$$P(Head) = \frac{\#\ favorable\ outcomes}{\#possible\ outcomes} = \frac{1}{2}$$

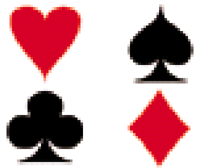
The probability of spinning green is 1 out of 2

Probability of Simple Events

<u>Example 6</u>: Deck of Cards.

- What is the probability of picking a heart?

$$P(Heart) = \frac{\boldsymbol{\#\,favorable\ outcomes}}{\boldsymbol{\#possible\ outcomes}} = \frac{\mathbf{13}}{\mathbf{52}} = \frac{\mathbf{1}}{\mathbf{4}}$$



The probability of picking a heart is 1 out of 4

- What is the probability of picking a non heart?

$$P(non - Heart) = \frac{\boldsymbol{\#\,favorable\ outcomes}}{\boldsymbol{\#possible\ outcomes}} = \frac{\mathbf{39}}{\mathbf{52}} = \frac{\mathbf{3}}{\mathbf{4}}$$

The probability of picking a heart is 3 out of 4

Probability of Simple Events

Key Concepts:

- Probability is the chance that some event will happen

- It is the *ratio* of the number of ways a certain even can occur to the total number of possible outcomes

Complementary Events

- The complement of an event E is the set of all outcomes in a sample space that are not included in event E.

- The complement of an event E is denoted by $E'\, or\ \bar{E}$

Properties of Probability:

$$0 \le P(E) \le 1$$

$$P(E) + P(\bar{E}) = 1$$

$$P(E) = 1 - P(\bar{E})$$

$$P(\bar{E}) = 1 - P(E)$$

Complementary Events

- **Example I:** A sequence of 5 bits is randomly generated. What is the probability that at least one of these bits is zero?
- **Solution:** There are $2^5 = 32$ possible outcomes of generating such a sequence.

Define event E as **at least one of the bits is zeros**

Then event $\bar{E}$, **“none of the bits is zero”**, includes only one

of these outcomes, namely the sequence 11111.

Therefore, p($\bar{E}$) = 1/32.

Now p(E) can easily be computed as

p(E) = 1 – p($\bar{E}$) = 1 – 1/32 = 31/32.

Conditional Probability

- We talk about conditional probability when the probability of one event depends on whether or not another event has occurred.
- e.g. There are 2 red and 3 blue counters in a bag and, without looking, we take out one counter and do not replace it.
- The probability of a 2^{nd} counter taken from the bag being red depends on whether the 1^{st} was red or blue.
- Conditional probability problems can be solved by considering the individual possibilities or by using a table, a Venn diagram, a tree diagram or a formula.

Notation

***P*(A|B) means**

"the probability that event A occurs $\boxed{\text{given}}$ that B has occurred". This is conditional probability.

Example

e.g. 1. The following table gives data on the type of car, grouped by petrol consumption, owned by 100 people.

	Low	Medium	High	Total
Male	12	33	7	
Female	23	21	4	
				100

One person is selected at random.

L is the event "the person owns a low rated car"

Example

e.g. 1. The following table gives data on the type of car, grouped by petrol consumption, owned by 100 people.

	Low	Medium	High	Total
Male	12	33	7	
Female	23	21	4	
				100

One person is selected at random.
L is the event "the person owns a low rated car"
F is the event "a female is chosen".

e.g. 1. The following table gives data on the type of car, grouped by petrol consumption, owned by 100 people.

	Low	Medium	High	Total
Male	**12**	**33**	**7**	
Female	**23**	**21**	**4**	
				100

One person is selected at random.

L is the event “the person owns a low rated car”

F is the event “a female is chosen”.

e.g. 1. The following table gives data on the type of car, grouped by petrol consumption, owned by 100 people.

	Low	Medium	High	Total
Male	**12**	**33**	**7**	
Female	**23**	**21**	**4**	
				100

One person is selected at random.

L is the event "the person owns a low rated car"

F is the event "a female is chosen".

Find (i) $P(\mathrm{L})$ (ii) $P(\mathrm{F} \text{ and } \mathrm{L})$ (iii) $P(\mathrm{F} \mid \mathrm{L})$

We need to be careful which row or column we look at.

Solution:

	Low	Medium	High	Total
Male	12	33	7	
Female	23	21	4	
	35			100

Find (i) P(L) (ii) P(F and L) (iii) P(F| L)

$$\text{(i) } P(\text{L}) = \frac{\cancel{35}^{\,7}}{\cancel{100}_{\,20}} = \frac{7}{20}$$

Solution:

	Low	Medium	High	Total
Male	12	33	7	
Female	23	21	4	
				100

Find (i) $P(\text{L})$ (ii) $P(\text{F and L})$ (iii) $P(\text{F} \mid \text{L})$

(i) $P(\text{L}) = \dfrac{\cancel{35}^{\,7}}{\cancel{100}_{\,20}} = \dfrac{7}{20}$

(ii) $P(\text{F and L}) = \dfrac{23}{100}$

The probability of selecting a female with a low rated car.

Solution:

	Low	Medium	High	Total
Male	12	33	7	
Female	23	21	4	
	35			100

Find (i) $P(\mathrm{L})$ (ii) $P(\mathrm{F} \text{ and } \mathrm{L})$ (iii) $P(\mathrm{F}|\mathrm{L})$

(i) $P(\mathrm{L}) = \dfrac{\cancel{35}^{\,7}}{\cancel{100}_{\,20}} = \dfrac{7}{20}$

(ii) $P(\mathrm{F} \text{ and } \mathrm{L}) = \dfrac{23}{100}$

(iii) $P(\mathrm{F}|\mathrm{L}) = \dfrac{23}{35}$

We must be careful with the denominators in (ii) and (iii). Here we are given the car is low rated. We want the total of that column.

Solution:

	Low	Medium	High	Total
Male	12	33	7	
Female	23	21	4	
				100

Find (i) $P(\mathrm{L})$ (ii) $P(\mathrm{F\ and\ L})$ (iii) $P(\mathrm{F} \mid \mathrm{L})$

(i) $P(\mathrm{L}) = \dfrac{\cancel{35}^{\,7}}{\cancel{100}_{\,20}} = \dfrac{7}{20}$

(ii) $P(\mathrm{F\ and\ L}) = \dfrac{23}{100}$

(iii) $P(\mathrm{F} \mid \mathrm{L}) = \dfrac{23}{35}$

Notice that

$P(\mathrm{L}) \times P(\mathrm{F} \mid \mathrm{L}) = \dfrac{{}^{1}\cancel{7}}{20} \times \dfrac{23}{\cancel{35}_{\,5}} = \dfrac{23}{100}$

$= P(\mathrm{F\ and\ L})$

So, $P(\mathrm{F\ and\ L}) = P(\mathrm{F} \mid \mathrm{L}) \times P(\mathrm{L})$

Conditional Probability

$$\boxed{P(\text{F and L}) = P(\text{F}|\text{L}) \times P(\text{L})}$$

- This result can be used to help solve harder conditional probability problems.

Bayesian Decision Theory

- Bayesian Decision Theory is a fundamental statistical approach that quantifies the tradeoffs between various decisions using probabilities and costs that accompany such decisions.
- First, we will assume that all probabilities are known.
- Then, we will study the cases where the probabilistic structure is not completely known.

Bayesian Decision Theory

- Design classifiers to recommend decisions that minimize some total expected "risk".
- The simplest risk is the classification error (i.e., costs are equal).
- Typically, the risk includes the cost associated with different decisions.

Terminology

- State of nature ω (random variable):
  - e.g., $\omega 1$ for sea bass, $\omega 2$ for salmon
- Probabilities $P(\omega 1)$ and $P(\omega 2)$ (priors):
  - e.g., prior knowledge of how likely is to get a seabassor a salmon
- Probability density function $p(x)$ (evidence):
  - e.g., how frequently we will measure a pattern with feature value x (e.g., x corresponds to lightness)

Fish Sorting Example Revisited

- State of nature is a random variable.
- Define w as the type of fish we observe (state of nature, *class*) where
  - $w = w_1$ for sea bass,
  - $w = w_2$ for salmon.
  - $P(w_1)$ is the *a priori probability* that the next fish is a sea bass.
  - $P(w_2)$ is the a priori probability that the next fish is a salmon.

Prior Probabilities

- Prior probabilities reflect our knowledge of how likely each type of fish will appear before we actually see it.
- How can we choose $P(w_1)$ and $P(w_2)$?
  - Set $P(w_1) = P(w_2)$ if they are equiprobable (*uniform priors*).
  - May use different values depending on the fishing area, time of the year, etc.
- Assume there are no other types of fish

$$P(w_1) + P(w_2) = 1$$

(exclusivity and exhaustivity).

Making a Decision

- How can we make a decision with only the prior information?

$$\text{Decide} \quad \begin{cases} w_1 & \text{if } P(w_1) > P(w_2) \\ w_2 & \text{otherwise} \end{cases}$$

- What is the *probability of error* for this decision?

$$P(error) = \min\{P(w_1), P(w_2)\}$$

Class-Conditional Probabilities

- Let's try to improve the decision using the lightness measurement x.
- Let x be a continuous random variable.
- Define $p(x|w_j)$ as the *class-conditional probability density* (probability of x given that the state of nature is w_j for $j = 1, 2$).
- $p(x|w_1)$ and $p(x|w_2)$ describe the difference in lightness between populations of sea bass and salmon.

Class-Conditional Probabilities

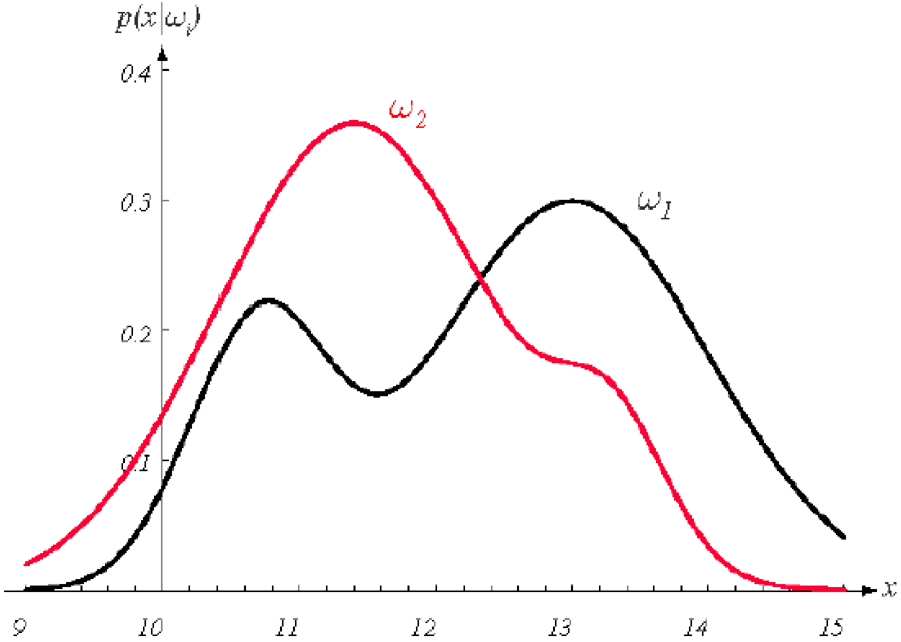


Figure 1: Hypothetical class-conditional probability density functions for two classes.

Thanks ☺