

Statistical Learning - Probability and Distributions

Proprietary content. ©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited

This file is meant for personal use by jmanuel.perez@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Probability – Meaning & Concepts

- **Probability** refers to chance or likelihood of a particular event-taking place.
- An **event** is an outcome of an experiment.
- An **experiment** is a process that is performed to understand and observe possible outcomes.
- Set of all outcomes of an experiment is called the **sample space**.

Example

- In a manufacturing unit three parts from the assembly are selected. You are observing whether they are defective or non-defective. Determine
 - a) The sample space.
 - b) The event of getting at least two defective parts.

Solution

a) Let S = Sample Space. It is pictured as under:



D = Defective

G = Non-Defective

b) Let E denote the event of getting at least two defective parts. This implies that E will contain two defectives, and three defectives. Looking at the sample space diagram above, $E = \{GDD, DGD, DDG, DDD\}$. It is easy to see that E is a part of S and commonly called as a subset of S . Hence an event is always a subset of the sample space.

Definition of Probability

Probability of an event A is defined as the ratio of two numbers m and n. In symbols

$$P(A) = \frac{m}{n}$$

where m= number of ways that are favorable to the occurrence of A and n= the total number of outcomes of the experiment (all possible outcomes)

Please note that P (A) is always ≥ 0 and always ≤ 1 .

P (A) is a pure number.

Extreme Values of Probability

The range within which probability of an event lies can be best understood by the following diagram. The glass shows three stages- Empty, half-full, and full to explain the properties of probability.



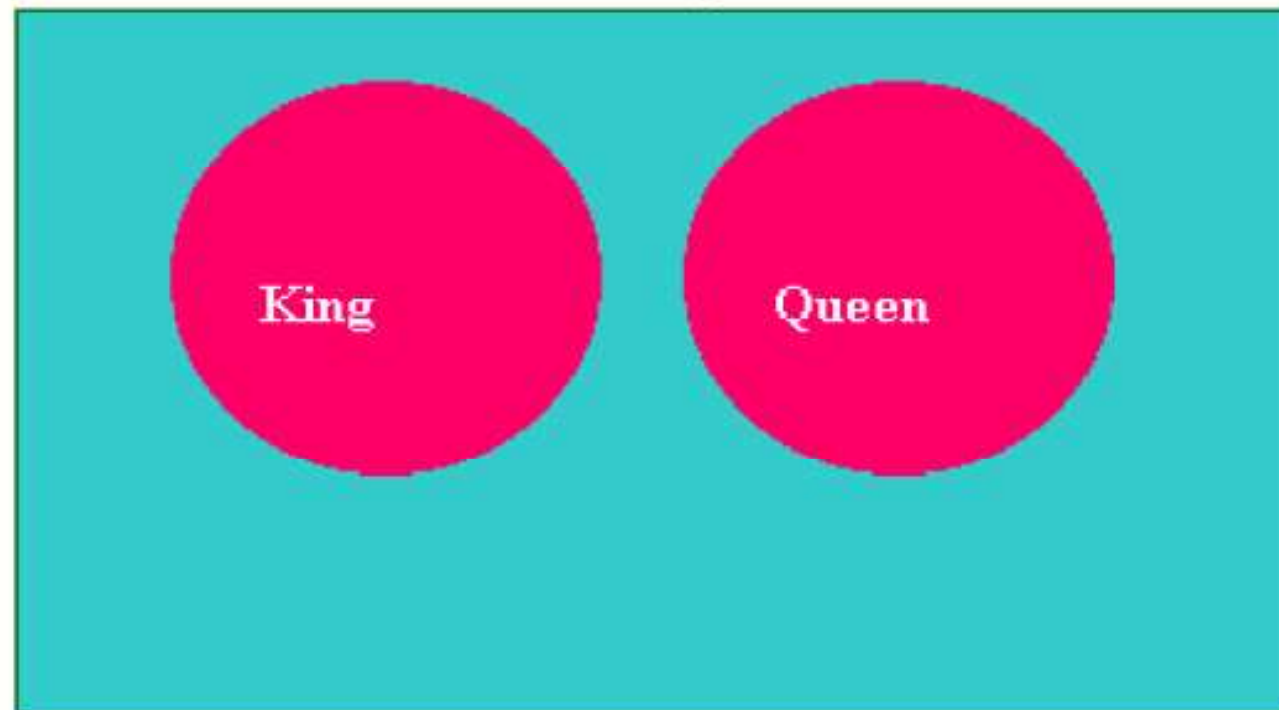
100% Chance or Certainty

(50% Chance) Equally Likely

(0% Chance) Impossibility

Mutually Exclusive Events

Two events A and B are said to be mutually exclusive if the occurrence of A precludes the occurrence of B. For example, from a well shuffled pack of cards, if you pick up one card at random and would like to know whether it is a King or a Queen. The selected card will be either a King or a Queen. It cannot be both a King and a Queen. If King occurs, Queen will not occur and Queen occurs, King will not occur.



Independent Events

- Two events A and B are said to be independent if the occurrence of A is in no way influenced by the occurrence of B. Likewise occurrence of B is in no way influenced by the occurrence of A.

Rules for Computing Probability

1) Addition Rule -Mutually Exclusive Events

$$P(A \cup B) = P(A) + P(B)$$

This rule says that the probability of the union of A and B is determined by adding the probability of the events A and B.

Here the symbol $A \cup B$ is called A union B meaning A occurs, or B occurs or both A and B simultaneously occur. When A and B are mutually exclusive, A and B cannot simultaneously occur.

Rules for Computing Probability

2) Addition Rule -Events are not Mutually Exclusive

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

This rule says that the probability of the union of A and B is determined by adding the probability of the events A and B and then subtracting the probability of the intersection of the events A and B.

The symbol $A \cap B$ is called A intersection B meaning

both A and B simultaneously occur.

Example for Addition Rules

- From a pack of well-shuffled cards, a card is picked up at random.
 - 1) What is the probability that the selected card is a King or a Queen?
 - 2) What is the probability that the selected card is a King or a Diamond?

Solution to Part 1)

Look at the Diagram:

Event A



Event B



Let A = getting a King

Let B = getting a Queen

There are 4 kings and there are 4 Queens. The events are clearly mutually exclusive. Applying the formula $P(A \cup B) = P(A) + P(B)$
 $= 4/52 + 4/52 = 8/52 = 2/13$

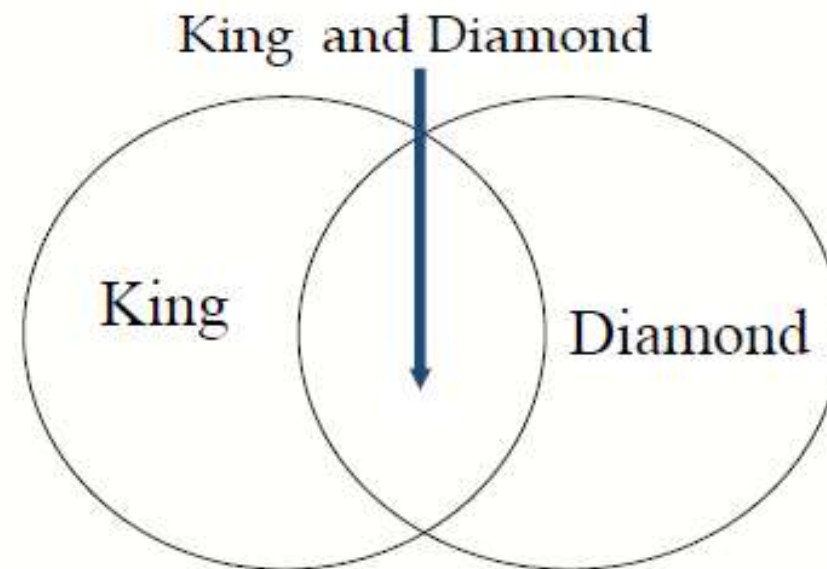
Solution to part 2)

INSTITUTE OF MANAGEMENT, CHEN

Look at the Diagram:

There are totally 52 cards in a pack out of which 4 are Kings and 13 are Diamonds. Let A= getting a King and B= getting a Diamond. The two events here are not mutually exclusive because you can have a card, which is both a King and a Diamond called King Diamond.

$$P(K \cup D) = P(K) + P(D) - P(K \cap D)$$
$$= 4/52 + 13/52 - 1/52 = 16/52 = 4/13$$



Multiplication Rule

Independent Events

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

This rule says when the two events A and B are independent, the probability of the simultaneous occurrence of A and B (also known as probability of intersection of A and B) equals the product of the probability of A and the probability of B. Of course this rule can be extended to more than two events.

Multiplication Rule

Independent Events-Example

Example:

The probability that you will get an A grade in Quantitative Methods is 0.7. The probability that you will get an A grade in Marketing is 0.5. Assuming these two courses are independent, compute the probability that you will get an A grade in both these subjects.

Solution:

Let A = getting A grade in Quantitative Methods

Let B = getting A grade in Marketing

It is given that A and B are independent.

$$P(A \cap B) = P(A).P(B) = 0.7.0.5 = 0.35.$$

Multiplication Rule

Events are not independent

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

This rule says that the probability of the intersection of the events A and B equals the product of the probability of A and the probability of B given that A has happened or known to you. This is symbolized in the second term of the above expression as $P(B/A)$. $P(B/A)$ is called the conditional probability of B given the fact that A has happened.

We can also write $P(A \cap B) = P(B) \cdot P(A/B)$ if B has already happened.

Multiplication Rule

Events are not independent-Example

INSTITUTE OF MANAGEMENT, CHE

From a pack of cards, 2 cards are drawn in succession one after the other. After every draw, the selected card is not replaced. What is the probability that in both the draws you will get Spades?

Solution:

Let A = getting Spade in the first draw

Let B = getting spade in the second draw.

The cards are not replaced.

This situation requires the use of conditional probability.

$P(A) = 13/52$ (There are 13 Spades and 52 cards in a pack)

$P(B/A) = 12/51$ (There are 12 Spades and 51 cards because the first card selected is not replaced after the first draw)

$$P(A \cap B) = P(A).P(B/A) = (13/52).(12/51) = 156/2652 = 1/17.$$

rohibited

Marginal Probability

- Contingency table consists of rows and columns of two attributes at different levels with frequencies or numbers in each of the cells. It is a matrix of frequencies assigned to rows and columns.
- The term marginal is used to indicate that the probabilities are calculated using a contingency table (also called joint probability table).

Marginal Probability - Example

A survey involving 200 families was conducted. Information regarding family income per year and whether the family buys a car are given in the following table.

Family	Income below \$ 20K	Income of >= \$ 20K	Total
Buyer of Car	38	42	80
Non-Buyer	82	38	120
Total	120	80	200

- What is the probability that a randomly selected family is a buyer of the car?
- What is the probability that a randomly selected family is both a buyer of car and belonging to income of \$ 20000 and above?
- A family selected at random is found to be belonging to income of \$ 20K and above. What is the probability that this family is buyer of car?

Solution

a) What is the probability that a randomly selected family is a buyer of the Car?

- $80/200 = 0.40$.

b) What is the probability that a randomly selected family is both a buyer of car and belonging to income of \$20 k and above?

- $42/200 = 0.21$.

c) A family selected at random is found to be belonging to income of \$ 20K and above. What is the probability that this family is buyer of car?

- $42/80 = 0.525$. Note this is a case of conditional probability of buyer given income is \$20 k and above.