

Probability & Probability Distributions

Outline

1. Introduction to Probability
2. Types of Probability
3. Bayes Theorem and its Applications
4. Binomial Distribution
5. Poisson Distribution
6. Normal distribution

Introduction

- Managers will have to cope with uncertainty in many decision situations. Concepts of probability will help you measure uncertainty and perform associated analyses that are essential in making effective business decisions.

Question

- What is the probability that a man lives for 1000 years?
- Ans – Empirical Probability

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.

3 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



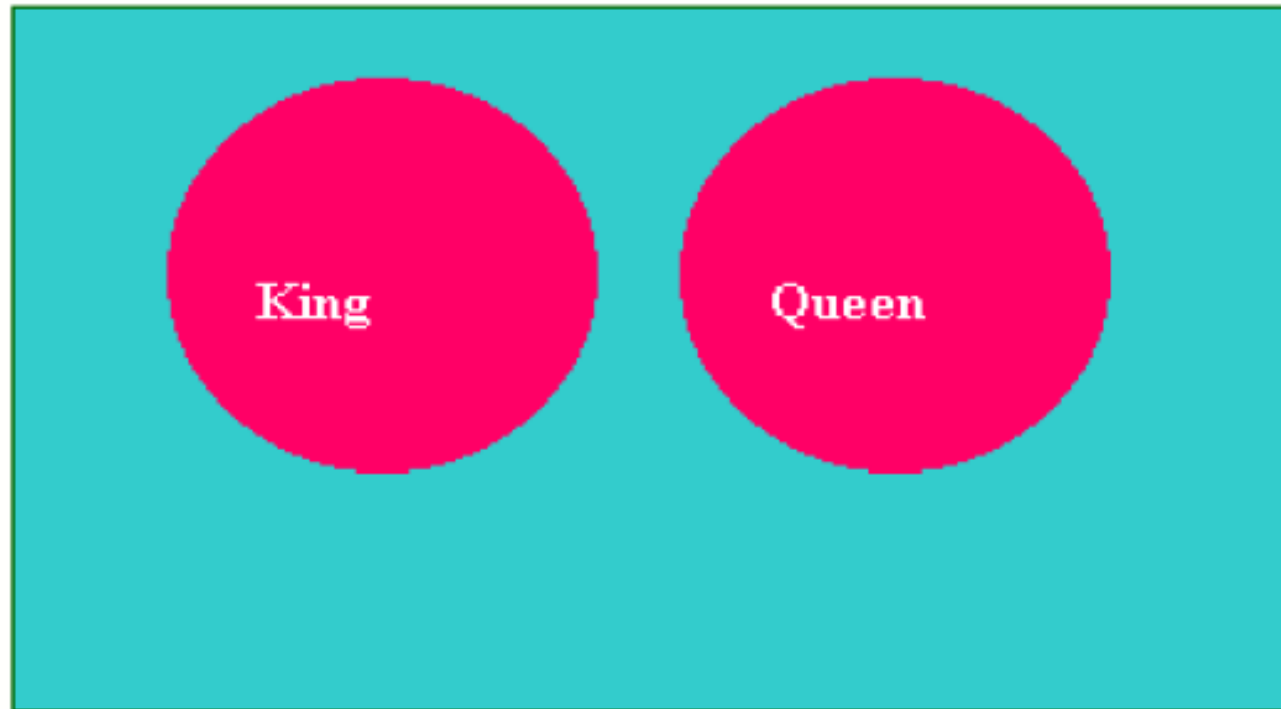
(0% Chance) Impossibility

Types of Probability

- **A Priori Classical Probability**
- **Empirical Probability**
- **Subjective Probability**

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

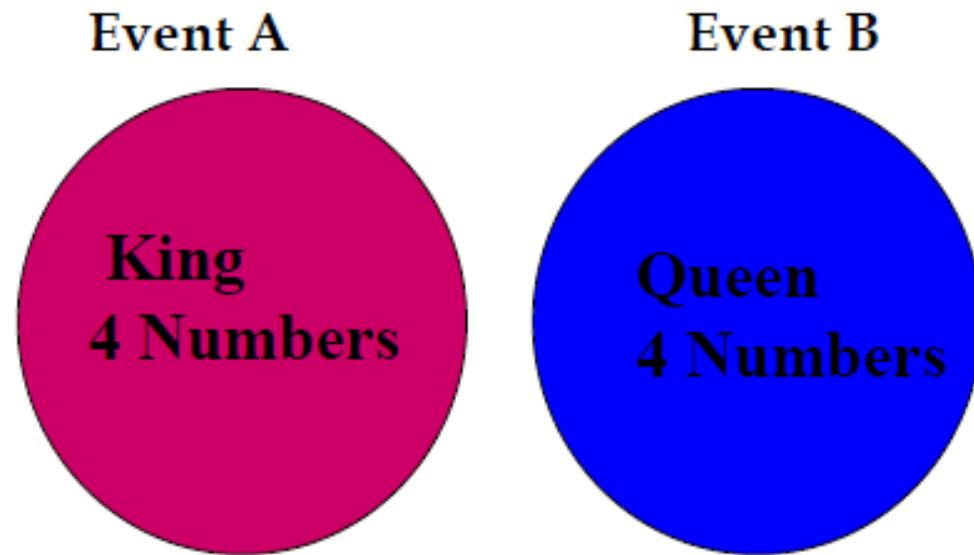
both A and B simultaneously occur.

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:



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)

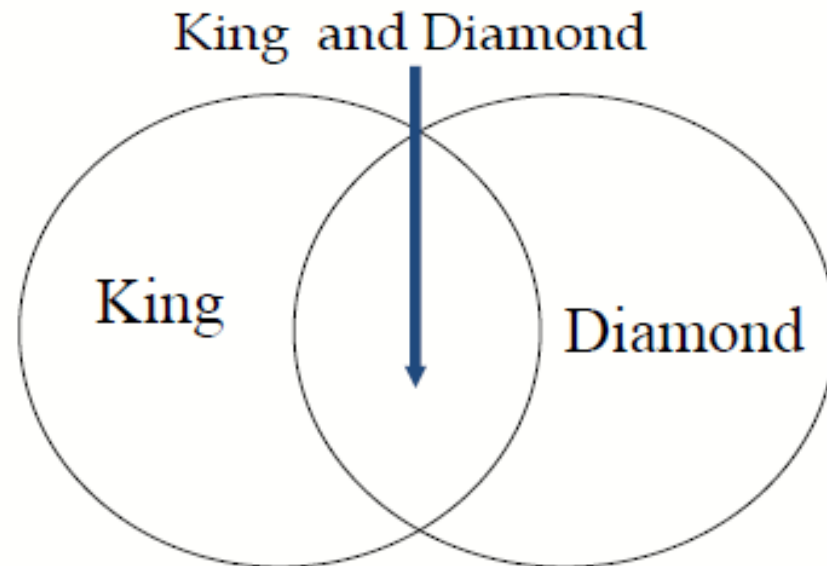
INSTITUTE OF MANAGEMENT, CHEN

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

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

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

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

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 $P(A \cap B) = P(B).P(A/B)$ if B has already happened.

Multiplication Rule

INSTITUTE OF MANAGEMENT, CHE

Events are not independent-Example

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)

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

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 Rs 10 Lakhs	Income of Rs. ≥ 10 lakhs	Total
Buyer of Car	38	42	80
Non-Buyer	82	38	120
Total	120	80	200

- a) What is the probability that a randomly selected family is a buyer of the car?
- b) What is the probability that a randomly selected family is both a buyer of car and belonging to income of Rs. 10 lakhs and above?
- c) A family selected at random is found to be belonging to income of Rs 10 lakhs 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 Rs. 10 lakhs and above?

- $42/200 = 0.21$.

c) A family selected at random is found to be belonging to income of Rs 10 lakhs 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 Rs. 10 lakhs and above.

Marginal Probability - Problem

- Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). 20% of the cars have both.

- What is the probability that a car has a CD player, given that it has AC ?

i.e., we want to find $P(\text{CD} \mid \text{AC})$

Solution

	CD	No CD	
AC	20	50	70
No AC	20	10	30
	40	60	100

$$P(\text{CD}/\text{AC}) = 20/70$$

Bayes' Theorem

- Bayes' Theorem is used to revise previously calculated probabilities based on new information.
- Developed by Thomas Bayes in the 18th Century.
- It is an extension of conditional probability.

Bayes' Theorem

Given a hypothesis H and evidence E , Bayes' theorem states that the relationship between the probability of the hypothesis $P(H)$ before getting the evidence and the probability $P(H | E)$ of the hypothesis after getting the evidence is

$$P(H | E) = \frac{P(E | H)}{P(E)} P(H).$$

Many modern machine learning techniques rely on Bayes' theorem. For instance, spam filters use Bayesian updating to determine whether an email is real or spam, given the words in the email. Additionally, many specific techniques in statistics, such as calculating p-values or interpreting medical results, are best described in terms of how they contribute to updating hypotheses using Bayes' theorem.

Bayes' Theorem

$$P(B_i | A) = \frac{P(A | B_i)P(B_i)}{P(A | B_1)P(B_1) + P(A | B_2)P(B_2) + \cdots + P(A | B_k)P(B_k)}$$

- where:

B_i = i^{th} event of k mutually exclusive and collectively exhaustive events

A = new event that might impact $P(B_i)$

Bayes' Theorem - Problem

- A drilling company has estimated a 40% chance of striking oil for their new well.
- A detailed test has been scheduled for more information. Historically, 60% of successful wells have had detailed tests, and 20% of unsuccessful wells have had detailed tests.
- Given that this well has been scheduled for a detailed test, what is the probability that the well will be successful?

Solution

	Success	No Success	
DT	24	12	36
No DT	16	48	64
	40	60	100

$$P(S/DT) = 24/36 = 0.67$$

What is a Probability Distribution

- In precise terms, a **probability distribution** is a total listing of the various values the random variable can take along with the corresponding probability of each value. A real life example could be the pattern of distribution of the machine breakdowns in a manufacturing unit.
- The random variable in this example would be the various values the machine breakdowns could assume.
- The probability corresponding to each value of the breakdown is the relative frequency of occurrence of the breakdown.
- The probability distribution for this example is constructed by the actual breakdown pattern observed over a period of time. Statisticians use the term “observed distribution” of breakdowns.

Binomial Distribution

- The Binomial Distribution is a widely used probability distribution of a discrete random variable.
- It plays a major role in **quality control** and **quality assurance** function. Manufacturing units do use the binomial distribution for **defective** analysis.
- Reducing the number of defectives using the proportion defective control chart (p chart) is an accepted practice in manufacturing organizations.
- Binomial distribution is also being used in **service organizations** like banks, and insurance corporations to get an idea of the proportion customers who are satisfied with the service quality.

Defective vs Defects

- In a 1000 lines Code Program, if n number of lines have problems, it is called defective
- However, the type of problem in each problematic line is called a defect
- Binomial deals with defective analysis, while poisson deals with the number of defects
- A computer is said to be defective. Binomial
- The type of defects in each defective computer. Poisson

Defective vs Defects

- Preparing a Project Report
- Number of Defective pages in the Report – Binomial (Defective/Non Defective)
- Number of defects per page – Poisson
- We could say - Without Poisson , there is no Binomial
- Binomial deals with % defective
- Poisson deals with number of defects per item
- Number of people arriving at an atm - Poisson

Defective vs Defects

- Bank – Dissatisfied Customers percentage is Binomial
- Number of complaints made by each customer – Poisson
- Ashok Leyland chooses Leather covers
- Defective covers – Binomial
- Types of Defects in each cover – black spots, thread cuts etc - Poisson

Conditions for Applying Binomial Distribution (Bernoulli Process)

- Trials are independent and random.
- There are fixed number of trials (n trials).
- There are only two outcomes of the trial designated as *success* or *failure*.
- The probability of success is uniform through out the n trials

Binomial Probability Function

Under the conditions of a Bernoulli process,

The probability of getting x successes out of n trials is indeed the definition of a Binomial Distribution. The Binomial Probability Function is given by the following expression

$$P(x) = \binom{n}{x} P^x (1 - P)^{n-x}$$

Where $P(x)$ is the probability of getting x successes in n trials

$$\binom{n}{x} \text{ is the number of ways in which } x \text{ successes can take place out of } n \text{ trials}$$
$$= \frac{n!}{x! (n - x)!}$$

P is the probability of success, which is the same through out the n trials.

P is the parameter of the Binomial distribution

x can take values $0, 1, 2, \dots, n$

Example for Binomial Distribution

- A bank issues credit cards to customers under the scheme of Master Card. Based on the past data, the bank has found out that 60% of all accounts pay on time following the bill. If a sample of 7 accounts is selected at random from the current database, construct the Binomial Probability Distribution of accounts paying on time.
- Work Out Using Python

Mean and Standard Deviation of the Binomial Distribution

The mean μ of the Binomial Distribution is given by $\mu = E(x) = np$

The Standard Deviation σ is given by

$$\sigma = \sqrt{np(1-p)}$$

For the example problem in the previous two slides,
Mean = $7 \times 0.6 = 4.2$.

$$\text{Standard Deviation} = \sqrt{4.2(1-0.60)} = 1.30$$