

Statistical and Mathematical Methods for Data Analysis

Dr. Faisal Bukhari

**Punjab University College of Information Technology
(PUCIT)**

Textbooks

- ❑ **Probability & Statistics for Engineers & Scientists**, Ninth Edition, Ronald E. Walpole, Raymond H. Myer
- ❑ **Elementary Statistics: Picturing the World**, 6th Edition, Ron Larson and Betsy Farber
- ❑ **Elementary Statistics**, 13th Edition, Mario F. Triola

Reference books

- ❑ **Probability Demystified**, Allan G. Bluman
- ❑ **Schaum's Outline of Probability and Statistics**
- ❑ **MATLAB Primer**, Seventh Edition
- ❑ **MATLAB Demystified** by McMahan, David

References

Readings for these lecture notes:

- ❑ **Probability & Statistics for Engineers & Scientists**, Ninth edition, Ronald E. Walpole, Raymond H. Myer

These notes contain material from the above book.

The Multinomial Distribution [1]

- ❑ Recall that for a probability experiment to be binomial, two outcomes are necessary. But if each trial of a probability experiment has **more than two outcomes, a distribution that can be used to describe the experiment is called a multinomial distribution.**
- ❑ In addition, there must be a **fixed number of independent** trials, and the probability for each success must remain the **same** for each trial.

The Multinomial Distribution [2]

A short version of the multinomial formula for **three outcomes** is given next. If X consists of events **E_1 , E_2 , and E_3** , which have corresponding probabilities of **p_1 , p_2 , and p_3** of occurring, where x_1 is the number of times **E_1** will occur, **x_2** is the number of times E_2 will occur, and x_3 is the number of times E_3 will occur, then the probability of X is

$$f(x_1, x_2, x_3; p_1, p_2, p_3) = \frac{n!}{x_1! \times x_2! \times x_3!} \times p_1^{x_1} \times p_2^{x_2} \times p_3^{x_3}$$

$$\sum_{i=1}^3 x_i = n, \sum_{i=1}^3 p_i = 1$$

where $x_1 + x_2 + x_3 = n$, $p_1 + p_2 + p_3 = 1$

The Multinomial Distribution [3]

Example: In a large city, **60%** of the workers drive to work, **30%** take the bus, and **10%** take the train. If **5** workers are selected at random, find the probability that **2** will drive, **2** will take the bus, and **1** will take the train.

Solution:

Let

x_1 = No of workers, who **drive to work = 2**

x_2 = No of workers, who **take bus to work = 2**

x_3 = No of workers, who **take train to work = 1**

p_1 = Probability of workers, who **drive to work = 0.60**

p_2 = Probability of workers, who **take bus to work = 0.30**

p_3 = Probability of workers, who **take train to work = 0.10**

Cont.

$$f(x_1, x_2, x_3; p_1, p_2, p_3) = \frac{n!}{x_1! \times x_2! \times x_3!} \times p_1^{x_1} \times p_2^{x_2} \times p_3^{x_3}$$

where $x_1 + x_2 + x_3 = n$, $p_1 + p_2 + p_3 = 1$

$$\begin{aligned} f(2, 2, 1; 0.60, 0.30, 0.1) &= \frac{5!}{2! \times 2! \times 1!} (0.6)^2 (0.3)^2 (0.1)^1 \\ &= 0.0972 \end{aligned}$$

The Multinomial Distribution [4]

Example: A box contains 5 red balls, 3 blue balls, and 2 white balls. If 4 balls are selected with replacement, find the probability of getting 2 red balls, one blue ball, and one white ball.

Solution:

Let

x_1 = No of **red balls** = 2

x_2 = No of **blue balls** = 1

x_3 = No of **white balls** = 1

p_1 = Probability of **red balls** = $5/10 = 0.50$

p_2 = Probability of **blue balls** = $3/10 = 0.30$

p_3 = Probability of **white balls** = $2/10 = 0.20$

$$f(x_1, x_2, x_3; p_1, p_2, p_3) = \frac{n!}{x_1! \times x_2! \times x_3!} \times p_1^{x_1} \times p_2^{x_2} \times p_3^{x_3}$$

where $x_1 + x_2 + x_3 = n$, $p_1 + p_2 + p_3 = 1$

Cont.

$$\begin{aligned} f(2, 1, 1; 0.50, 0.30, 0.20) &= \frac{4!}{2! \times 1! \times 1!} (0.5)^2 (0.3)^1 \\ &\quad (0.2)^1 \\ &= 0.18 \end{aligned}$$

Multinomial Experiments and the Multinomial Distribution [1]

The binomial experiment becomes a **multinomial experiment** if we let each trial have more than **two possible outcomes**.

□ The classification of a manufactured product as being **light, heavy, or acceptable** and the recording of accidents at a certain intersection according to the day of the week constitute **multinomial experiments**.

Multinomial Experiments and the Multinomial Distribution [2]

- ❑ The drawing of a card from a deck *with replacement* is also a multinomial experiment if the **4 suits** are the outcomes of interest.

Multinomial Experiments and the Multinomial Distribution [3]

In general, if a given trial can result in any one of k possible outcomes E_1, E_2, \dots, E_k with probabilities p_1, p_2, \dots, p_k , then the **multinomial distribution** will give the probability that E_1 occurs x_1 times, E_2 occurs x_2 times, \dots , and E_k occurs x_k times in n independent trials,

Multinomial Experiments and the Multinomial Distribution [4]

□ where $x_1 + x_2 + \cdots + x_k = n$.

We shall denote this joint probability distribution by

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n).$$

□ Clearly, $p_1 + p_2 + \cdots + p_k = 1$, since the result of each trial must be one of the k possible outcomes

Multinomial Distribution

If a given trial can result in the k outcomes E_1, E_2, \dots, E_k with probabilities p_1, p_2, \dots, p_k then the probability distribution of the random variables X_1, X_2, \dots, X_k representing the number of occurrences for E_1, E_2, \dots, E_k in n independent trials, is,

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n) = \frac{n!}{x_1! \times x_2! \times \dots \times x_k!} \times p_1^{x_1} \times p_2^{x_2} \times \dots \times p_k^{x_k}$$

$$\sum_{i=1}^k x_i = n, \sum_{i=1}^k p_i = 1$$

Example: The complexity of arrivals and departures of planes at an airport is such that computer simulation is often used to model the “ideal” conditions. For a certain airport with **three runways**, it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet:

Runway 1: $p_1 = 2/9 = 0.2222$ (or 22.22%)

Runway 2: $p_2 = 1/6 = 0.1667$ (or 16.6667%)

Runway 3: $p_3 = 11/18 = 0.6111$ (or 61.1111%)

What is the probability that **6** randomly arriving airplanes are distributed in the following fashion?

Runway 1: 2 airplanes

Runway 2: 1 airplane

Runway 3: 3 airplanes

Solution: Using the multinomial distribution, we have

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n) = \frac{n!}{x_1! \times x_2! \dots \times x_k!} \times p_1^{x_1} \times p_2^{x_2} \times \dots \times p_k^{x_k}$$

$$f(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6) = \frac{6!}{2! \times 1! \times 3!} \left(\frac{2}{9}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{11}{18}\right)^3$$

$$= 0.1127.$$