University of Dhaka

Institute of Business Administration (IBA)

Master of Business Administration (MBA)

K501: *Quantative Analysis for Business Decision*

Last Updated: Wednesday 14th May, 2025

Author Details

Name	:	Md Hasibul Islam
Student ID	:	201-67-011
Program	:	Master of Business Administration (MBA)
Institute	:	Institute of Business Administration (IBA)
University	:	University of Dhaka
Email	:	hasiee8004@gmail.com
LinkedIn	:	linkedin.com/in/hasib009
GitHub	:	github.com/HasibRockie
Website	:	hasibrockie.github.io

Contents

Statistics and Its Fundamental Concepts

21-04-25 Monday

Statistics is the science of collecting, organizing, analyzing, interpreting, and presenting data to support decision-making and problem-solving in business and other domains. It provides a quantitative foundation for managerial decisions by offering meaningful insights from data.

Types of Statistics:

1. **Descriptive Statistics:** This involves methods of organizing, summarizing, and displaying data.
   - Examples: Mean, Median, Mode, Standard Deviation, Frequency tables, Pie charts, Histograms.
   - Use Case: A retail manager summarizes last month's sales performance using a bar chart and average daily revenue.
2. **Inferential Statistics:** This refers to techniques for making generalizations from a sample to a population using probability theory.
   - Examples: Hypothesis testing, Confidence intervals, Regression analysis.
   - Use Case: A pharmaceutical company tests a new drug on a sample group to infer its effectiveness on the broader population.

Types of Variables:

- **Qualitative (Categorical) Variables:** Represent categories or labels.
  - Examples: Gender, Brand name, Type of customer (new/returning).
- **Quantitative Variables:** Represent numeric values.
  - **Discrete Variables:** Countable values (e.g., Number of employees).
  - **Continuous Variables:** Measurable and can take any value within a range (e.g., Sales revenue, Temperature).

Levels of Measurement:

Other Key Concepts:

- **Population:** The entire group of individuals or instances about whom we hope to learn.
- **Sample:** A subset of the population, selected for analysis.
- **Parameter:** A numerical summary or measure that describes a characteristic of a population (e.g., population mean μ).
- **Statistic:** A numerical summary derived from a sample (e.g., sample mean $\bar{x}$). Statistics are used to estimate parameters.

Characteristic	**Nominal**	**Ordinal**	**Interval**	**Ratio**
Definition	Categorical data without any order	Categorical data with a logical order	Numeric data with equal intervals, no true zero	Numeric data with equal intervals and a true zero
Nature of Data	Labels or names	Ordered categories	Quantitative	Quantitative
Mathematical Operations	Equality only	Comparisons (>, <)	Addition, subtraction	All mathematical operations
Meaningful Zero	No	No	No	Yes
Can Calculate Mean?	No	No (median preferred)	Yes	Yes
Examples	Gender, Blood Type, Product Type	Socioeconomic Status, Education Level	Temperature (Celsius/Fahrenheit), IQ Score	Height, Weight, Age, Sales Revenue
Applicable Statistics	Mode, Frequency	Mode, Median, Percentile	Mean, SD, Correlation	All descriptive and inferential statistics
Distance between values is meaningful?	No	Not always	Yes	Yes
Has absolute zero?	No	No	No	Yes

Table 1: Comparison of Levels of Measurement

Classification of Variables with Examples

28-04-25 Monday

Table 2: Classification of Variables — Qualitative vs Quantitative, Discrete vs Continuous

Variable Type	**Discrete Examples**	**Continuous Examples**
Qualitative	Shirt size (S, M, L) Product category (A, B, C) Number of children category (None, One, Two+) Room type (Single, Double) Education level (High School, UG, PG)	Skin tone spectrum Customer feedback scale Dialect variation Shade of color preferences Accent variation
Quantitative	Number of cars owned Number of transactions Exam scores (out of 100) Number of employees Number of visits	Height (cm) Weight (kg) Income ($) Temperature (°C) Time spent (hours)

Table 2: Classification of Variables: Qualitative vs Quantitative and Discrete vs Continuous

Table 3: Classification by Levels of Measurement — Nominal, Ordinal, Interval, Ratio with Discrete and Continuous Types

Level of Measurement	Discrete Examples	Continuous Examples
Nominal	• Jersey number • Postal code • Nationality • Product ID • Car model	• Color shade • Accent pattern • Logo design variation • Pattern of speech • Ink density
Ordinal	• Customer rating (1–5 stars) • Survey rank (Strongly disagree to Agree) • Academic grade (A, B, C) • Job level (Junior, Mid, Senior) • Market tier (Low, Mid, High)	• Satisfaction level on 0–10 scale • Credit score bands • Health condition severity • Employee performance level • Risk tolerance scale
Interval	• Test scores (e.g., IQ, SAT) • Temperature recorded hourly • Credit scores in discrete brackets • Year of birth • Calendar dates	• Temperature (°C or °F) • Time of day (without AM/PM) • Financial index points • Sound intensity • Wind speed variation
Ratio	• Number of products sold • Number of goals scored • Number of books owned • Defect counts in production • Visitors per day	• Income • Distance traveled • Weight • Time • Age

Table 3: 50 Examples Categorized by Level of Measurement and Variable Type

Classification of Interval and Ratio Data

05-05-25 Monday

Table 1: Examples of Interval vs Ratio Data

Interval Examples	Ratio Examples
Temperature (°C or °F)	Height (cm)
IQ scores	Weight (kg)
Calendar years (e.g., 1990, 2000)	Income ($)
Time of day (without AM/PM)	Distance (miles, km)
Standardized test scores (e.g., SAT)	Age (years)
Body temperature (°C or °F)	Speed (km/h, m/s)
Elevation (altitude above sea level)	Amount of money in a bank account ($)
Scores on a thermometer (not Celsius)	Number of products produced
Electrical voltage (in volts)	Number of hours worked
Sea level pressure in hPa	Time spent on an activity (minutes, hours)
Temperature difference (relative to a baseline)	Volume of water in a tank (liters)
Time intervals (measured in minutes, seconds)	Weight of an object (kg)
Age difference between two people (in years)	Height of a building (meters)
Test scores in degrees (e.g., temperature degrees in a test)	Distance traveled (km)
Time elapsed (counted in minutes or hours)	Duration of an event (seconds, hours)
Yearly income (adjusted for inflation)	Car mileage (km per liter)
Currency fluctuation in a day	Number of children in a family
Dates in history (e.g., 1776, 2000)	Average monthly rent ($)
Income tax (as a percentage of income)	Calories burned (kcal)
Temperature increase over a day	Amount of rainfall (cm)
Time difference in hours (e.g., time zones)	Number of cups of coffee consumed
Survey scores (e.g., 1 to 10 scale)	Blood pressure level (mm Hg)
Duration in seconds (e.g., a stopwatch)	Quantity of goods sold
Difference in height between two points on Earth	Quantity of raw material processed
Height of a mountain (in meters)	Hours of sleep per night

Table 4: Examples of Interval and Ratio Data

Table 2: Continuous vs Discrete Examples for Interval and Ratio Data

Continuous Examples	**Discrete Examples**
Temperature (°C or °F)	Number of children in a family
Height (cm)	Number of cars owned
Weight (kg)	Number of transactions made
Age (years)	Number of goals scored in a game
Distance (miles, km)	Number of students in a class
Speed (km/h, m/s)	Number of employees in a company
Time spent on a task (minutes, hours)	Number of books in a library
Amount of money in a bank account ($)	Number of visitors to a website
Volume of water in a tank (liters)	Number of rooms in a house
Duration of an event (seconds, minutes)	Number of items sold in a day
Calories burned (kcal)	Number of people in a line
Amount of rainfall (cm)	Number of phone calls received
Number of cups of coffee consumed	Number of hours worked per week
Blood pressure level (mm Hg)	Number of students passing an exam
Height of a building (meters)	Number of cars in a parking lot
Hours of sleep per night	Number of pets owned
Monthly rent ($)	Number of emails sent in a day
Total income for the year ($)	Number of tickets sold for an event
Survey scores (e.g., 1 to 10 scale)	Number of products produced in a factory
Time intervals (measured in minutes)	Number of patients in a hospital
Test scores (standardized)	Number of questions answered correctly on a quiz
Time difference between locations (e.g., time zone differences)	Number of units produced per hour
Total calories consumed in a day	Number of visitors to a museum
Survey feedback score (1 to 10)	Number of cups in a cupboard
Time elapsed in hours (e.g., work hours)	Number of toys in a box

Table 5: Continuous vs Discrete Examples for Interval and Ratio Data

Frequency Distribution and Mean Classifications

10-04-25 Thursday

Frequency Distribution and Parameters

Frequency Distribution: A frequency distribution is a table that displays the frequency (i.e., how often each value appears) of different values or categories in a dataset. It helps to organize and summarize data, providing insight into the underlying patterns.

Parameters of Frequency Distribution:

- **k (Number of Classes):** The number of categories or intervals that data will be grouped into. It is typically chosen based on the range and size of the dataset.
- **N (Total Number of Data Points):** The total number of observations or data points in the dataset.
- **i (Class Width):** The width of each class interval, calculated using the formula:
$$i = \frac{\text{Max} - \text{Min}}{k}$$
where Max and Min represent the maximum and minimum values in the dataset, respectively.
- The rule for selecting k is: $2^k > N$, where k is the number of classes, and N is the total number of observations.

Example: Suppose you have a dataset of 50 data points ($N = 50$), and the range of the data is from 10 to 100.

To determine the number of classes (k) and the class width (i):

$$2^k > 50 \quad \Rightarrow \quad k = 6 \quad (\text{as } 2^6 = 64 > 50)$$

$$i = \frac{100 - 10}{6} = 15$$

Thus, the number of classes k will be 6, and each class will have a width of 15.

Mean

1. Arithmetic Mean: The arithmetic mean, commonly referred to as the average, is the sum of all values in the dataset divided by the number of values. It is the most widely used measure of central tendency.

$$\text{Arithmetic Mean}(\bar{x}) = \frac{\sum x}{N}$$

Where: - x represents each data point, - N is the number of data points.

Example: For the dataset: 2, 4, 6, 8, 10,

$$\bar{x} = \frac{2 + 4 + 6 + 8 + 10}{5} = \frac{30}{5} = 6$$

2. Geometric Mean: The geometric mean is the nth root of the product of all values in a dataset, where n is the number of values. The geometric mean is often used when comparing values that have different ranges, such as growth rates.

$$\text{Geometric Mean}(GM) = \left(\prod_{i=1}^{N} x_i\right)^{\frac{1}{N}}$$

Where: - x_i represents each data point, - N is the number of data points.

Example: For the dataset: 1, 4, 16, 64,

$$GM = (1 \times 4 \times 16 \times 64)^{\frac{1}{4}} = 1024^{\frac{1}{4}} = 8$$

3. Weighted Mean: The weighted mean is a variation of the arithmetic mean where each data point is assigned a weight that reflects its importance or frequency. The weighted mean is particularly useful when some values in the dataset are more significant than others.

$$\text{Weighted Mean} = \frac{\sum w_i x_i}{\sum w_i}$$

Where: - w_i is the weight associated with data point x_i, - x_i is the value of the data point, - N is the total number of data points.

Example: For the dataset: (value, weight): $(2,3), (4,2), (6,5)$,

$$\text{Weighted Mean} = \frac{(3 \times 2) + (2 \times 4) + (5 \times 6)}{3 + 2 + 5} = \frac{6 + 8 + 30}{10} = \frac{44}{10} = 4.4$$

Measures of Location and Distribution Types

10-04-25 Thursday

Mode, Median, and Common Measures of Location

1. Mode:
The mode is the value that appears most frequently in a data set. It is the only measure of central tendency that can be used for nominal data. A dataset may have no mode, one mode (unimodal), or multiple modes (bimodal or multimodal).

Example:
For the dataset: 3, 3, 4, 5, 6, 6, 6, 7,
- Mode = 6 (appears most frequently).

2. Median:
The median is the middle value when the data is arranged in ascending or descending order. If the number of data points is odd, the median is the middle number. If it is even, the median is the average of the two middle numbers.

Example:
For the dataset: 1, 3, 5, 7, 9,
- Median = 5 (middle value).

For the dataset: 1, 3, 5, 7,
- Median = (3 + 5) / 2 = 4.

3. Other Measures of Location:
In addition to the mode and median, other measures of location include the mean (discussed earlier), quartiles, and percentiles, which divide the data into specific parts: - Quartiles split data into four equal parts. - Percentiles divide the data into 100 equal parts, useful for detailed analysis.

—

Population Mean and Sample Mean

Population Mean:
The population mean (μ) is the average of all the values in the population. It is calculated by summing all the values and dividing by the total number of values in the population.

$$\mu = \frac{\sum x_i}{N}$$

Where: - x_i represents each individual data point, - N is the total number of data points in the population.

Example:
For the population data: 2, 4, 6, 8, 10,

$$\mu = \frac{2+4+6+8+10}{5} = \frac{30}{5} = 6$$

Sample Mean:
The sample mean ($\bar{x}$) is the average of a sample taken from the population. It is calculated in the same way as the population mean but uses the sample data.

$$\bar{x} = \frac{\sum x_i}{n}$$

Where: - x_i represents each individual data point in the sample, - n is the number of data points in the sample.

Example:
For the sample data: 1, 3, 5, 7, 9,

$$\bar{x} = \frac{1+3+5+7+9}{5} = \frac{25}{5} = 5$$

—

Graphs for Symmetric, Negatively Skewed, and Positively Skewed Distributions

The graphs for the three types of distributions can be created using external tools such as Python or pre-generated images. Here's how you would describe and use graphs:

1. **Symmetric Distribution:**
A symmetric distribution has the same shape on both sides of the center. The mean, median, and mode are all at the center. A classic example is the **normal distribution**.

2. **Negatively Skewed Distribution:**
A negatively skewed distribution (or left-skewed distribution) has a longer tail on the left side. The mean is less than the median, and the mode is greater than the median.

3. **Positively Skewed Distribution:**
A positively skewed distribution (or right-skewed distribution) has a longer tail on the right side. The mean is greater than the median, and the mode is less than the median.

To generate the graphs, use the following code in a Python environment or generate similar graphs in other plotting tools:

"'python import numpy as np import matplotlib.pyplot as plt

Create data for symmetric, negatively skewed, and positively skewed distributions np.random.seed(0)

Symmetric (Normal Distribution) symmetric$_d ata = np.random.normal(0, 1, 1000)$

Negatively Skewed (Lognormal Distribution) negative$_s kew_d ata = np.random.lognormal(0, 1, 1000)$

Positively Skewed (Lognormal Distribution) positive$_s kew_d ata = np.random.lognormal(1, 1, 1000)$

Plotting fig, axs = plt.subplots(3, 1, figsize=(6, 12))

Symmetric Distribution axs[0].hist(symmetric$_d ata, bins = 30, edgecolor =' black', alpha = 0.7) axs[0].set_t itle('Sy$

Negatively Skewed Distribution axs[1].hist(negative$_s kew_d ata, bins = 30, edgecolor =' black', alpha = 0.7) axs[1].set_t itle('NegativelySkewedDistribution') axs[1].set_x label('Value') axs[1].set_y label('Frequency')$

Positively Skewed Distribution axs[2].hist(positive$_s kew_d ata, bins = 30, edgecolor =' black', alpha = 0.7) axs[2].se$

plt.tight$_l ayout() plt.show()$

2025-04-23

Wednesday

Topics Covered: