PROB & STATS REPORT

Name: Bilal Ahmed Khan

Roll No: K200183   Sec: B

Topic:

Statistical Measures of Data: Mean, Variance, and Standard Deviation.

1. Mean:

a. Definition:

Mean is an essential concept in mathematics and statistics. The mean is the average or the most common value in a collection of numbers.

In statistics, it is a measure of central tendency of a probability distribution along median and mode. It is also referred to as an expected value.

It is a statistical concept that carries a major significance in finance. The concept is used in various financial fields, including but not limited to portfolio management and business valuation.

b. Formula:

There are different ways of measuring the central tendency of a set of values. There are multiple ways to calculate the mean. Here are the two most popular ones:

Arithmetic mean is the total of the sum of all values in a collection of numbers divided by the number of numbers in a collection. It is calculated in the following way:

$$\text{Arithmetic mean} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

In finance, the arithmetic mean may be misleading in the calculations of returns, as it does not consider the effects of volatility and compounding, producing an inflated value for the central point of the distribution.

Geometric mean is an nth root of the product of all numbers in a collection. The formula for the geometric mean is:

$$\text{Geometric mean} = \sqrt[n]{X_1 \times X_2 \times \ldots \times X_n}$$

The geometric mean includes the volatility and compounding effects of returns. Thus, the geometric average provides a more accurate calculation of an average return.

c. Examples

i. Arithmetic Mean:

Jim wants to find a stock for investment. He is a big fan of Apple Inc. He knows that the company has strong financials. However, to ensure that this investment will bring him a substantial return, he has decided to check how the stock performed in the past. He decides to find the average price of Apple's share price for the past five months.

He gathered the monthly company's stock prices from January 2018 to June 2018 and found the monthly returns. The stock prices and returns are summarized in the table below:

	Stock Price	Return (%)	Return (decimal)
December	167.90	N/A	
January	166.11	-1.07%	0.99
February	176.72	6.39%	1.06
March	167.14	-5.42%	0.95
April	164.63	-1.50%	0.98
May	186.15	13.07%	1.13
June	185.11	-0.56%	0.99

The formula used for the calculation would be the following:

$$\text{Arithmetic mean} = \frac{-1.07\% + 6.39\% - 5.42\% - 1.50\% + 13.07\% - 0.56\%}{6} = 1.82\%$$

ii. Geometric Mean:

In order to check the obtained result, Jim has decided to calculate the geometric mean return of Apple's share price. However, it should be calculated not in percentages but in decimal numbers.

The geometric mean is equal to:

$$\text{Geometric mean} = \sqrt[6]{0.99 \times 1.06 \times 0.95 \times 0.98 \times 1.13 \times 0.99} = 1.0164 \text{ or } 1.64\%$$

2. Variance:

a. Definition:

The variance is a measure of variability. It is calculated by taking the average of squared deviations from the mean.

Variance tells you the degree of spread in your data set. The more spread the data, the larger the variance is in relation to the mean.

b. Formula

i. Population Variance:

When you have collected data from every member of the population that you're interested in, you can get an exact value for population variance.

The population variance formula looks like this:

Formula	Explanation
$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$	• σ^2 = population variance • $\sum$ = sum of… • X = each value • μ = population mean • N = number of values in the population

ii. Sample Variance:

When you collect data from a sample, the sample variance is used to make estimates or inference about the population variance.

The sample variance formula looks like this:

Formula	Explanation
$s^2 = \frac{\sum (X - \bar{x})^2}{n - 1}$	• s^2 = sample variance • $\sum$ = sum of… • X = each value • $\bar{x}$ = sample mean • n = number of values in the sample

c. Importance of variance

i. Homogeneity of variance in statistical test:

Variance is important to consider before performing parametric tests. These tests require equal or similar variances, also called homogeneity of variance or homoscedasticity, when comparing different samples.

Uneven variances between samples result in biased and skewed test results. If you have uneven variances across samples, non-parametric tests are more appropriate.

ii. Using variance to assess group differences:

Statistical tests like variance tests or the analysis of variance (ANOVA) use sample variance to assess group differences. They use the variances of the samples to assess whether the populations they come from differ from each other.

3. Standard Deviation:

a. Definition:

Standard deviation is a measure of dispersement in statistics. “Dispersement” tells you how much your data is spread out. Specifically, it shows you how much your data is spread out around the mean or average.

b. Graphical Representation of Standard Deviation:

The bell curve (what statisticians call a “normal distribution “) is commonly seen in statistics as a tool to understand standard deviation.

The following graph of a normal distribution represents a great deal of data in real life. The mean, or average, is represented by the Greek letter μ, in the center. Each segment (colored in dark blue to light blue) represents one standard deviation away from the mean. For example, 2σ means two standard deviations from the mean.

c. Formula:

When you’re running an experiment (or test, or survey), you’re usually working with a sample— a small fraction of the population. The formula to find the standard deviation (s) when working with samples is:

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

The Σ sign in the formula means “to add up”. To solve the formula,

Add the numbers,

Square them,

Then divide.