# Statistics Practical

Abdullah Al Mahmud

2024-06-05

# Contents

# Determination of Arithmetic Mean, Combined Arithmetic Mean, Geometrci Mean, and Harmonic Mean from Grouped and Ungrouped Data.

# Construction of Histogram and Ogive

## Problem

Given below are daily wages of 30 workers in an agency.

**515, 833, 938, 511, 960, 968, 542, 842, 767, 694, 674, 955, 675, 972, 501, 987, 708, 846, 568, 721, 592, 867, 644, 966, 663, 551, 746, 942, 760, 601**

Draw a Histogram and an Ogive from the data and interpret.

## Solution

### Theory

A histogram is constructed from a frequency distribution with continuous class intervals. The frequencies corresponding to different classes are shown on the axis as bars, leaving no space or gap between the bars.

# Determination of Quartiles, Deciles, and Percentiles

## Problem

Given below are temperatures (in degree Celsius) of a city in 30 random days in a year.

33.86, 34.83, 35.59, 31.66, 26.31, 26.90, 33.10, 26.52, 35.17, 25.21, 25.28, 28.38, 29.62, 30.69, 32.72, 30.00, 30.14, 27.97, 28.45, 35.93, 33.34, 29.07, 34.00, 27.55, 34.03, 33.76, 29.48, 31.24, 33.79, 33.41

Find the quartiles, 4th and 7th Decile, and 35th & 87th Percentiles from the data and interpret.

## Solution

### Theory

Determination of First Quartile

Let n = no. of observations

If n = odd

Location of first quartile = $\frac{n+1}{4}$th item

Location of 2nd quartile = $\frac{2(n+1)}{4}$th item

Other quantiles are determined in the same way.

If n = even

Location of first quartile = $\frac{\frac{n}{4}th+\frac{(n+1)}{4}th}{2}$

Location of 2nd quartile = $\frac{\frac{2n}{4}th+\frac{2(n+1)}{4}th}{2}$

Other quantiles are determined similarly.

The general formula to find the ith quantile, which divides the dataset into k parts.

$$Q_i = \frac{i(n+1)}{k}th; \text{if n is odd}$$

$$Q_i = \frac{\frac{in}{k}th + (\frac{in}{k} + 1)th}{2}; \text{if n is even}$$

# Solution of Different Types of Problems Using Histogram and Ogive

## Problem

The following are the scores of 80 applicants in a screening test given by an institution.

| Score | 45-49 | 50-54 | 55-59 | 60-64 | 65-69 | 70-74 | 75-79 | 80-84 | 84-89 | 90-94 |
|---|---|---|---|---|---|---|---|---|---|---|
| Frequency | 3 | 5 | 11 | 15 | 20 | 8 | 7 | 5 | 4 | 2 |

i. Draw an Ogive and therefrom find $Q3, D_4, and P_{47}$ and interpret.
ii. Construct a histogram from the data and find Mode.

# Determination of First Four Moments from Data

## Problem

Temperature (in degree ceslsius) in Sylhet in first 10 days in May are given below:

29, 31, 30, 32, 30, 31, 28, 29, 34, 33

Find the first four central moments and the corresponding raw moments around 5.

## Solution

### Theory

rth raw moment around a, $\mu'_r(a) = \frac{\sum_{i=1}^{n}(x_i - a)^r}{n}$

If we denote $d_i = x_i - a$, then

$\mu'_r(a) = \frac{\sum d^r}{n}$

rth central moment, $\mu_r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^r}{n}$

#### Conversion from raw moments to central moments

We can use the binomial expansion.

For the first moment, we use $(a+b)^1 = a + b$

For the second moment, We use $(a+b)^2 = a^2 + 2ab + b^2$

Similarly the subsequent moments are used to convert the origin moments. We can get the coefficients of binomial expansion from the pascal triangle.

To convert an origin `a` to another origin `k`, i.e. to convert $\mu'_r(a) to \mu'_r(k)$

we let $a^r = \mu'_r(a)$ and $b = a - k$

Thus, to, for example, convert $\mu_2(5)$ to $\mu'_2(8)$ we have

$b = a - k = 5 - 8 = -3$

$\mu_2'(8) = \mu_2'(5) + 2\mu_1'(5)b + b^2$

**Calculation**

Make a table

**Central Moments**

We find using the conversion

# Determination of Skewness and Kurtosis and Their Types

## Problem

A shrimp producer wanted to get an insight into his shrimp production. To do so, he randomly collected weights of different shrimps in his farm.

| Weight of shrimp (gm) | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 |
|---|---|---|---|---|---|
| Frequency | 5 | 8 | 10 | 9 | 4 |

a. Estimate skewness and kurtosis of the data and interpret.

## Solution

### Theory

Coefficient of skewness, $\gamma_1 = \frac{\mu_3}{\sqrt{\mu_2^3}}$

Coefficient of kurtosis, $\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3$

# Construction of Box & Whisker Plot and Five Numbers Summary And Analysis of Their Properties

## Problem

In the asteroid belt in the Solar System, there are estimated to be between 1.1 and 1.9 million objects with a radius above 500 m. The radii of the 25 largest bodies are given below.

990, 980, 975, 924, 831, 824, 820, 780, 750, 731
700, 700, 675, 658, 653, 609, 570, 515, 500, 466
450, 432, 409, 400, 390

a. Display the data on a box and whisker plot and explain.
b. Determine the five number summary and explain.

## Solution

### Theory

To construct the Box and Whisker plot, we need to find the following values.

$Median = \frac{n+1}{2}th$ term [since n is odd]

First quartile, $Q_1 = \frac{n+1}{4}th$ term

Third quartile, $Q_3 = \frac{3(n+1)}{4}th$ term

We also need to find the inner and outer fences using the Interquartile range (IQR)

$IQR = Q_3 - Q_1$

Inner fence: $Q_1 - 1.5 \times IQR$ and $Q_3 + 1.5 \times IQR$

Outer fence: $Q_1 - 3 \times IQR$ and $Q_3 + 3 \times IQR$

# Determination of General Trend of Time Series and Prediction

# Exercises