Stats Cheat Sheet

Standard score → z score; Raw score → z score

$$Z = \frac{X - M}{\sigma}$$

- X → Raw Score
- M → Mean
- σ → Standard Deviation

mean (grouped data)

$$\overline{X} = \frac{\Sigma f X}{N}$$

- $\overline{X}$ → Mean
- f → Frequency
- X → Value or Midpoint
  - ↳ $= \left(\frac{\text{Lower} + \text{upper boundary}}{2}\right)$
- N (Σf) → # of items in the set

mean (ungrouped data)

$$\overline{X} = \frac{\Sigma X}{N}$$

- $\overline{X}$ → Mean
- X → Value or Midpoint
- N → Total # of items in the set

Median (ungrouped data)

When value of n is an odd number

$$\text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{item}$$

n → Total # of items in the set.

When value of n is an even number

$$\text{Median} = \left[\frac{\frac{n}{2}^{th} + \left(\frac{n}{2}+1\right)^{th}}{2}\right] \text{item}$$

n → Total # of items in the set.

Median (grouped data)

$$\text{Median} = L + \left[\frac{\frac{n}{2} - cf}{f}\right] i$$

- L → Lower Limit of median class − 0.5
- n → number of observations (Σf)
- cf → cumulative frequency of class preceding the median class
- f → frequency of median class
- i → class size

Mean Deviation

$$MD = X - \overline{X}$$

Average Deviation (ungrouped)

$$AD = \frac{\Sigma |x|}{N}$$

$x = |X - \overline{X}|$ (MD)

(always consider all values as positive)

Average Deviation (grouped)

$$AD = \frac{\Sigma |f x|}{N}$$

- f → frequency
- $x = |X - \overline{X}|$ (MD) (always consider all values as positive)
- X → Midpoint of class Interval

Standard Deviation (ungrouped data)

$$SD(\sigma) = \sqrt{\frac{\Sigma x^2}{N}}$$

- x → Mean Deviation
- $x = X - \overline{X}$

Standard Deviation (grouped data)

$$SD(\sigma) = \sqrt{\frac{\Sigma f x^2}{N}}$$

- x → Mean Deviation
- $x = X - \overline{X}$
- f → frequency

Quartile Deviation (QD)

$$Q_D = \frac{Q_3 - Q_1}{2}$$

$$Q_3 = l + \left(\frac{\frac{75}{100}n - cf}{f}\right) i$$

$$Q_1 = l + \left(\frac{\frac{25}{100}n - cf}{f}\right) i$$

Hypothesis Testing

Standard Error of Mean

Step 1: calculate Standard deviation

> Large Sample (N>30)

> For ungrouped data

$$\sigma = \sqrt{\frac{\Sigma (X - \overline{X})^2}{N}}$$

> For Grouped data

$$\sigma = \sqrt{\frac{\Sigma f (X - \overline{X})^2}{N}}$$

> Small Sample

> ungrouped data

$$\sigma = \sqrt{\frac{\Sigma (X - \overline{X})^2}{N-1}}$$

> Grouped data

$$\sigma = \sqrt{\frac{\Sigma f (X - \overline{X})^2}{N-1}}$$

Step 2: calculate standard error of mean

$$SE_M = \frac{\sigma}{\sqrt{N}}$$

Step 3: calculate Z_T (tabulated) [95% or 99%]

$$Z_{T95} = 1.96$$

$$Z_{T99} = 2.58$$

Step 4: calculate Confidence interval

> Upper limit $= M + (SE_M \times Z_T)$

> Lower Limit $= M - (SE_M \times Z_T)$

- M → mean
- SE_M → Standard error of mean
- Z_T → Z value from table

Step 5: calculate Z value obtained

Z_O = Upper limit − Lower limit

Step 6:

If $Z_O \geq Z_T$ Reject Null hypothesis (H_0)

If $Z_O < Z_T$ accept Null hypothesis (H_0)

Z-test for hypothesis testing (Two Groups)

$$Z = \frac{\text{Mean difference}}{\text{Combined standard Error of mean}} = \frac{M_X - M_Y}{SE_{M_{XY}}}$$

→ Step 6

- M_X → Mean of group X $\left(\frac{\Sigma X}{N}\right)$
- M_Y → Mean of group Y $\left(\frac{\Sigma Y}{N}\right)$

$$SE_{M_{XY}} = \sqrt{\frac{SD_X^2}{N_X} + \frac{SD_Y^2}{N_Y}}$$

$$SD_X = \sqrt{\frac{\Sigma (X - \overline{X})^2}{N_X}} \qquad SD_Y = \sqrt{\frac{\Sigma (Y - \overline{Y})^2}{N_Y}}$$

T-test for hypothesis testing (Two Groups) [Independent Sample]

1. Draw this table

X	Y	$x (X - \overline{X})$	x^2	$y (Y - \overline{Y})$	y^2

2. Calculate mean of Group X & Y i.e M_X & M_Y
3. Calculate deviation of Group X & Y i.e $(X - \overline{X})$ & $(Y - \overline{Y})$
4. Calculate Combined S D

$$\sigma = \sqrt{\frac{\Sigma x^2 + \Sigma y^2}{N - k}}$$

- N → Total no. of samples in all groups
- k → No. of groups

5. Calculate $SE_{md} = \sigma \sqrt{\frac{N_1 + N_2}{N_1 \times N_2}}$
6. calculate value of $t = \frac{M_X - M_Y}{SE_{md}}$
7. If $T_{obtained} \geq T_{table}$ → H_0 rejected

   If $T_{obtained} < T_{table}$ → H_0 accepted