BIRZEIT UNIVERSITY

Mathematics Department

Second Semester 2016/2017

STAT2311 - Statistics 1

Instructor : Hani Kabajah

final exam Study guide

prepared by : Maryam Shaheen

Chapter 1

Definitions :

Statistics : Collection of methods for planning experiments, obtaining data, and then organizing, summarizing, presenting, analyzing, interpreting, and drawing conclusions.

Variable : Characteristic or attribute that can assume different values

Random Variable : A variable whose values are determined by chance.

Population : All subjects possessing a common characteristic that is being studied.

Sample : A subgroup or subset of the population.

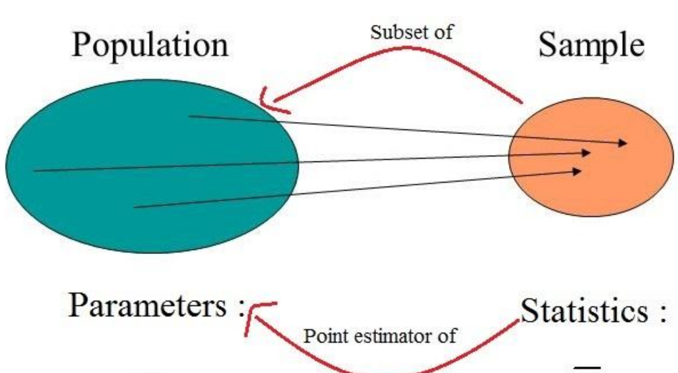
Parameter : Characteristic or measure obtained from a population.

Statistic (not to be confused with Statistics) : Characteristic or measure obtained from a sample.

Descriptive Statistics : Collection, organization, summarization, and presentation of data.

Inferential Statistics : Generalizing from samples to populations using probabilities. Performing hypothesis testing, determining relationships between variables, and making predictions.

Inference Overview



	Parameters	Statistics
Mean:	μ	$\overline{x}$
Standard Deviation:	σ	s
Proportion:	π	p

by Maryam Shaheen

Data Types :

Qualitative Variables : Variables which assume non-numerical values.

Quantitative Variables : Variables which assume numerical values.

Discrete Variables : Variables which assume a finite or countable number of possible values. Usually obtained by counting.

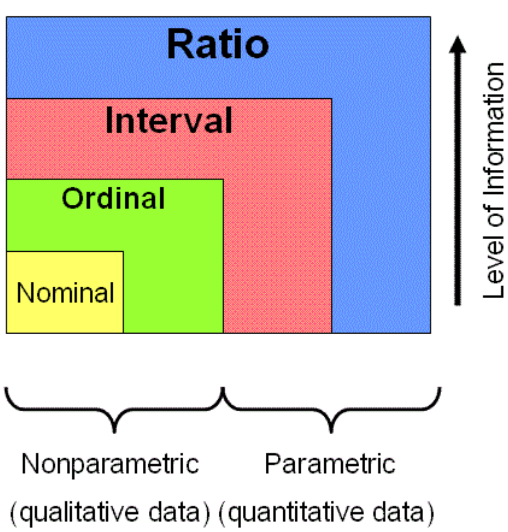
Continuous Variables : Variables which assume an infinite number of possible values. Usually obtained by measurement.

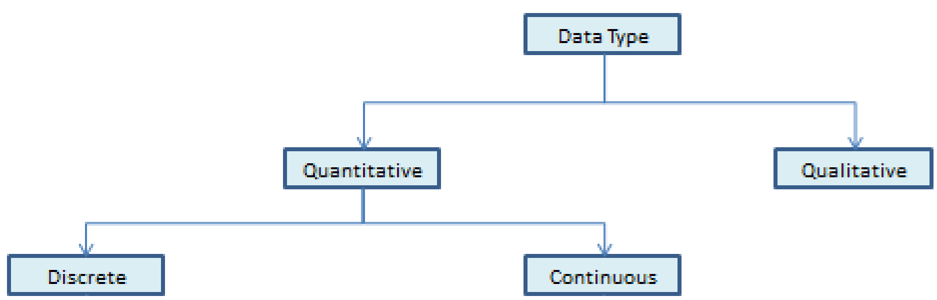
Nominal Scale : Scale of measurement which classifies data into mutually exclusive, all inclusive categories in which no order or ranking can be imposed on the data.

Ordinal Scale: Scale of measurement which classifies data into categories that can be ranked. Differences between the ranks do not exist.

Interval Scale: Scale of measurement which classifies data that can be ranked and differences are meaningful. However, there is no meaningful zero, so ratios are meaningless.

Ratio Scale: Scale of measurement which classifies data that can be ranked, differences are meaningful, and there is a true zero. True ratios exist between the different units of measure.





by Maryam Shaheen

Sampling Types :

1) Non-Probability Sampling Techniques :

Convenience Sampling : is probably the most common of all sampling techniques. With convenience sampling, the samples are selected because they are accessible to the researcher. Subjects are chosen simply because they are easy to recruit. This technique is considered easiest, cheapest and least time consuming.

Quota Sampling : is a non-probability sampling technique wherein the researcher ensures equal or proportionate representation of subjects depending on which trait is considered as basis of the quota.

Judgmental Sampling : is more commonly known as purposive sampling. In this type of sampling, subjects are chosen to be part of the sample with a specific purpose in mind. With judgmental sampling, the researcher believes that some subjects are more fit for the research compared to other individuals. This is the reason why they are purposively chosen as subjects.

Snowball Sampling : is usually done when there is a very small population size. In this type of sampling, the researcher asks the initial subject to identify another potential subject who also meets the criteria of the research. The downside of using a snowball sample is that it is hardly representative of the population.

2) Probability Sampling Techniques : (Chapter 7)

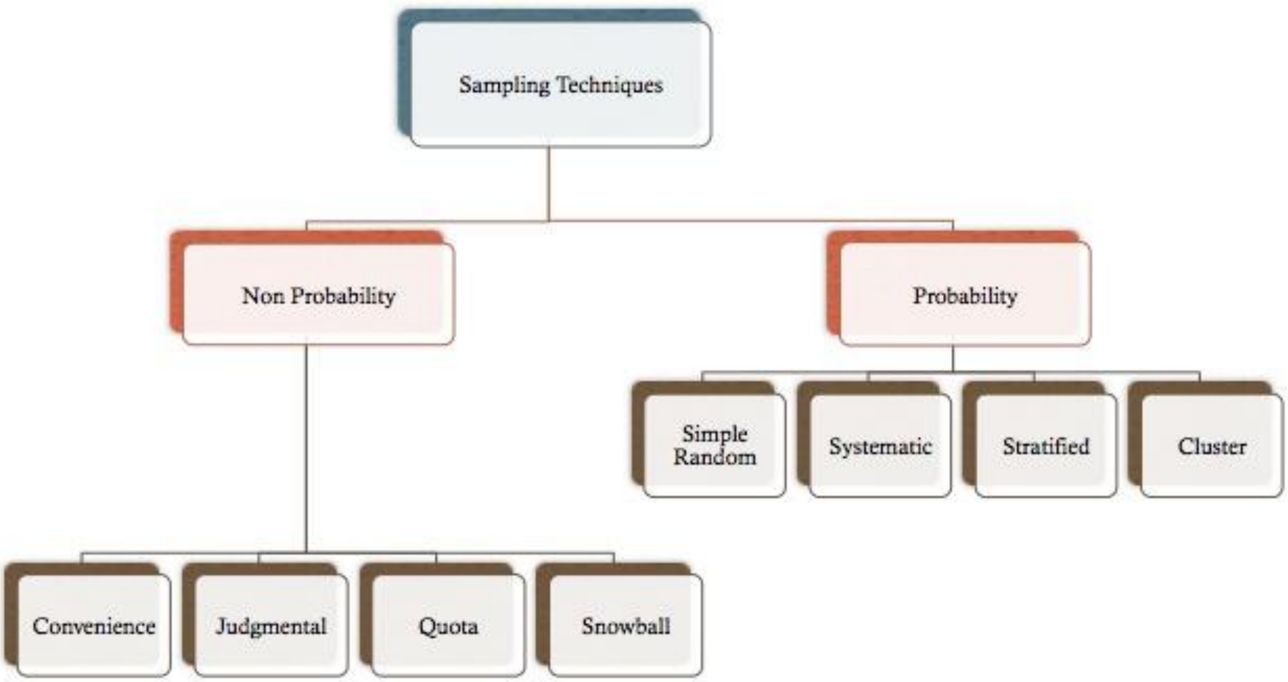
Random Sampling : Sampling in which the data is collected using chance methods or random numbers.

Systematic Sampling : Sampling in which data is obtained by selecting every *k*th object.

Convenience Sampling : Sampling in which data is which is readily available is used.

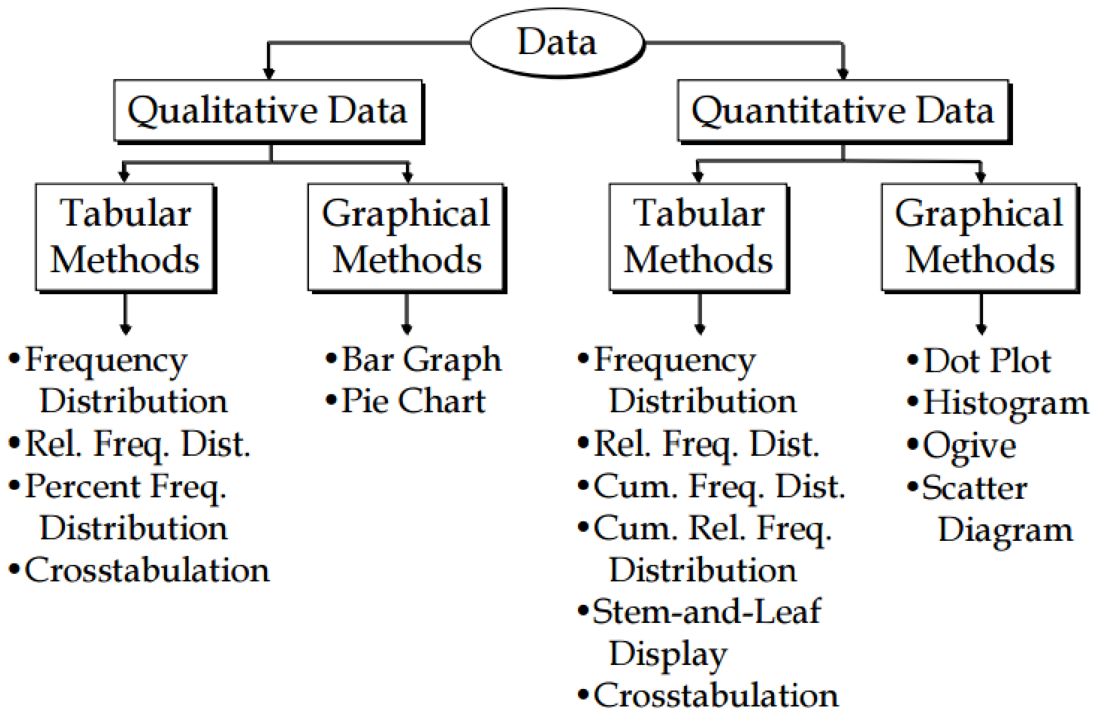
Stratified Sampling : Sampling in which the population is divided into groups (called strata) according to some characteristic. Each of these strata is then sampled using one of the other sampling techniques.

Cluster Sampling : Sampling in which the population is divided into groups (usually geographically). Some of these groups are randomly selected, and then all of the elements in those groups are selected.



Chapter 2

Tabular and Graphical Procedures



1) Qualitative Data :

a) Tabular :

Frequency distribution : is a tabular summary of data showing the frequency (or number) of items in each of several non-overlapping classes.

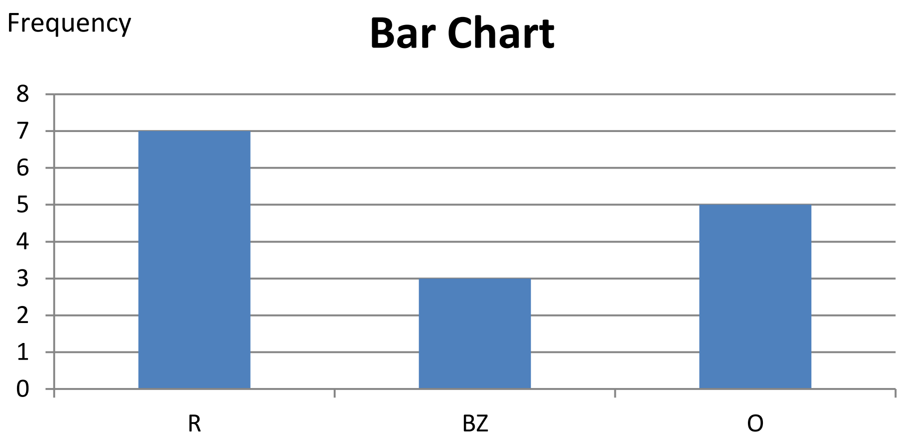
Relative frequency : is the fraction or proportion of the total number of data items belonging to the class.

Exp : Data on place of residence of sample of students (Ramallah, Birzeit , Other)
Data : R, O, O, BZ, R, O, BZ, R, O, R, R, BZ, O, R, R

Place of residence	Tally	Frequency	Relative frequency
R	𝍸 ||	7	7/15 = 0.46
BZ	|||	3	3/15 = 0.20
O	𝍸	5	5/15 = 0.33
Total		15	1

b) Graphical:

Bar Chart : is a graphical device for depicting qualitative data.



2) Quantitative Data :

a) Tabular :

Cumulative frequency distribution : shows the *number* of items with values less than or equal to the upper limit of each class..

Class Width = $\frac{\text{Largest Data Value} - \text{Smallest Data Value}}{\text{Number of Classes}} = \frac{33-12}{5} = 4.1 \rightarrow$ round up = 5

Number of Classes is preferred to be between 5 and 6 classes

Exp : Summarize the following Data :

Data : 12, 14, 19, 18, 15, 15, 18, 17, 20, 27, 22, 23, 22, 21, 33, 28, 14, 18, 16, 13

Class	Tally	Frequency	Relative freq.	Cumulative freq.
10 - 14	||||	4	4/20 = 0.2	<= 14 → 4
15 - 19	||| ~~||||~~	8	8/20 = 0.4	<= 19 → 12
20 - 24	~~||||~~	5	5/20 = 0.25	<= 24 → 14
25 - 29	||	2	2/20 = 0.1	<= 29 → 19
30 - 34	|	1	1/20 = 0.05	<= 34 → 20
Total		20	1	

Stem and leaf Display : shows both the *rank order* and *shape of the distribution* of the data.

Exp: Construct a stem and leaf display for the following data

Data : 70, 72, 75, 64, 58, 83, 83, 80, 82, 76, 75, 68, 63, 57, 57, 78, 85, 72

Stem = 10

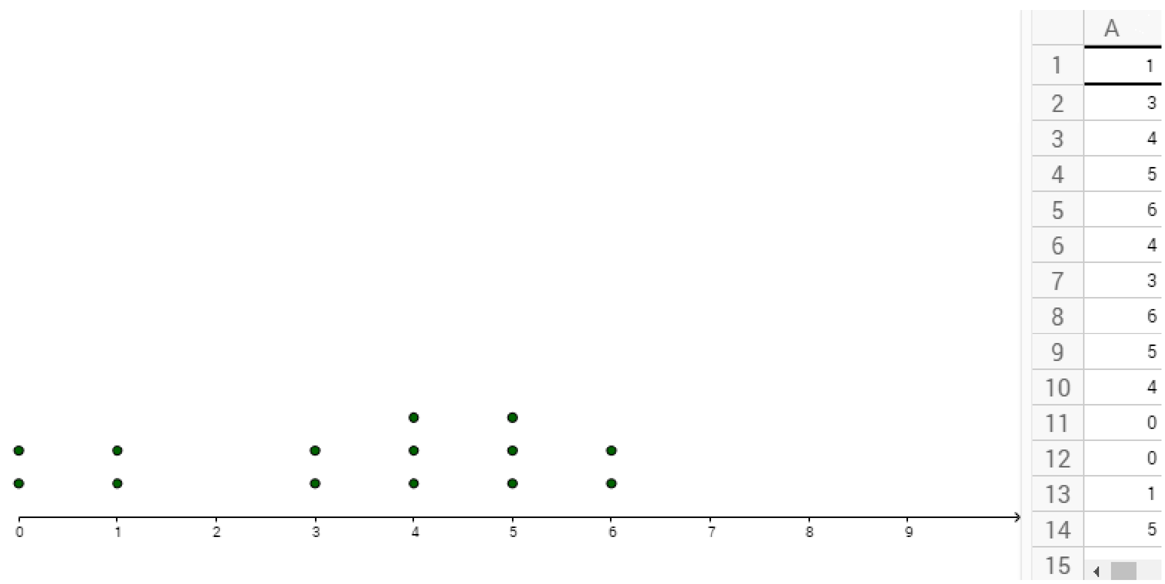
Leaf = 1

Stem	Leaf
5	**8 7**
6	**4 8 3**
7	**0 2 5 6 5 8 2**
8	**3 0 2 5**

b) Graphical:

Dot Plot : is One of the simplest graphical summaries of data.

Exp :

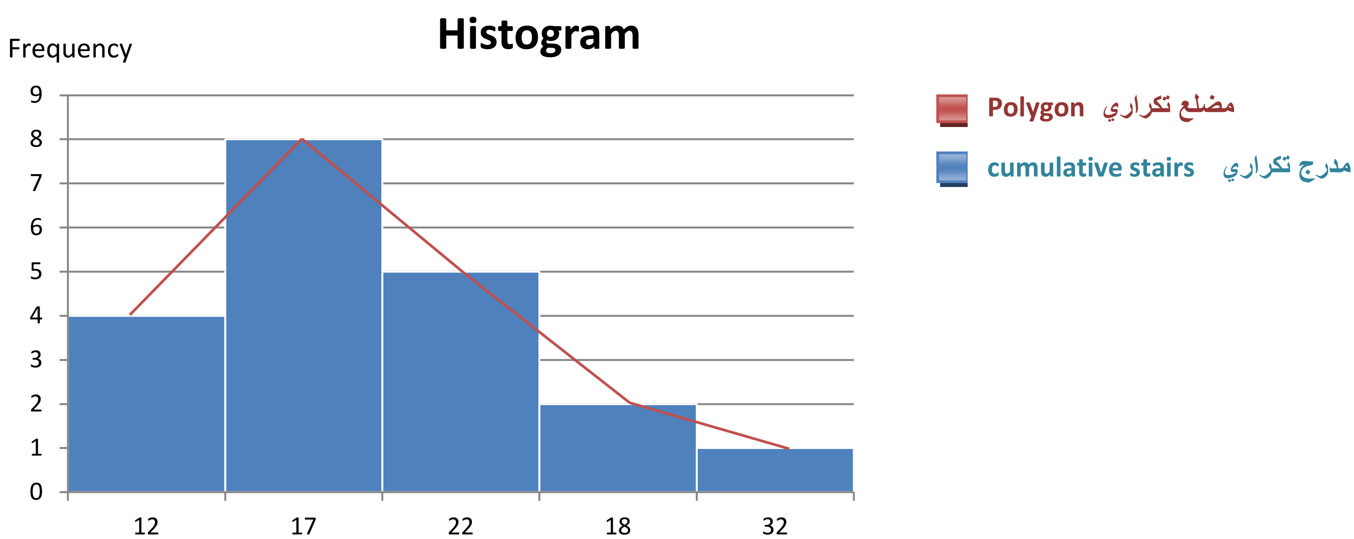


Histogram : is Another common graphical presentation of quantitative data

$$\text{Histogram} = \frac{\text{Max of The Class} + \text{min of The Class}}{2}$$

Exp :

Class	Tally	Frequency	Histogram
10 - 14	||||	4	10 + 14 / 2 = 12
15 - 19	||| 𝍸	8	15 + 19 / 2 = 17
20 - 24	𝍸	5	20 + 24 / 2 = 22
25 - 29	||	2	25 + 29 / 2 = 18
30 - 34	|	1	30 + 34 / 2 = 32
Total		20	

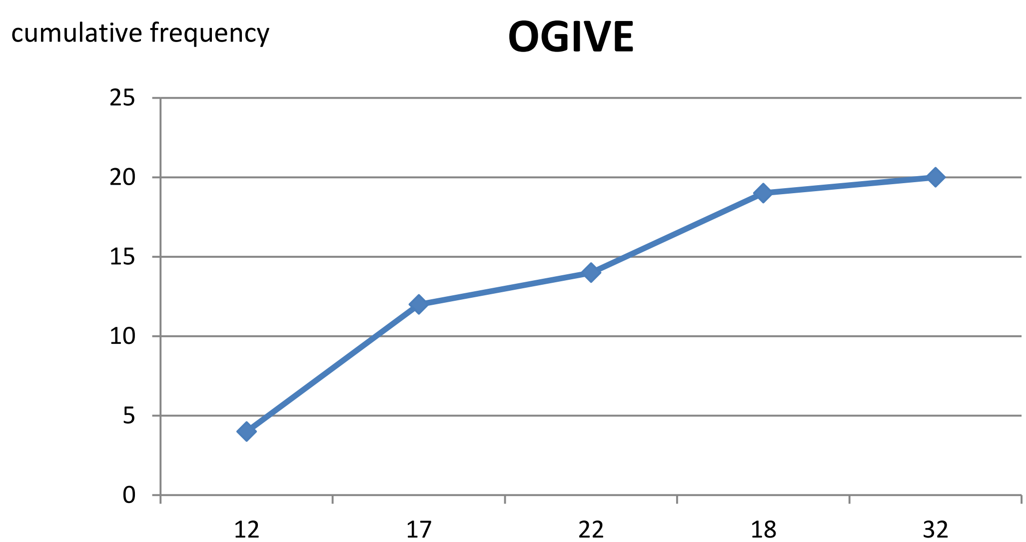


by Maryam Shaheen

OGIVE : is a graph of a cumulative distribution

Exp :

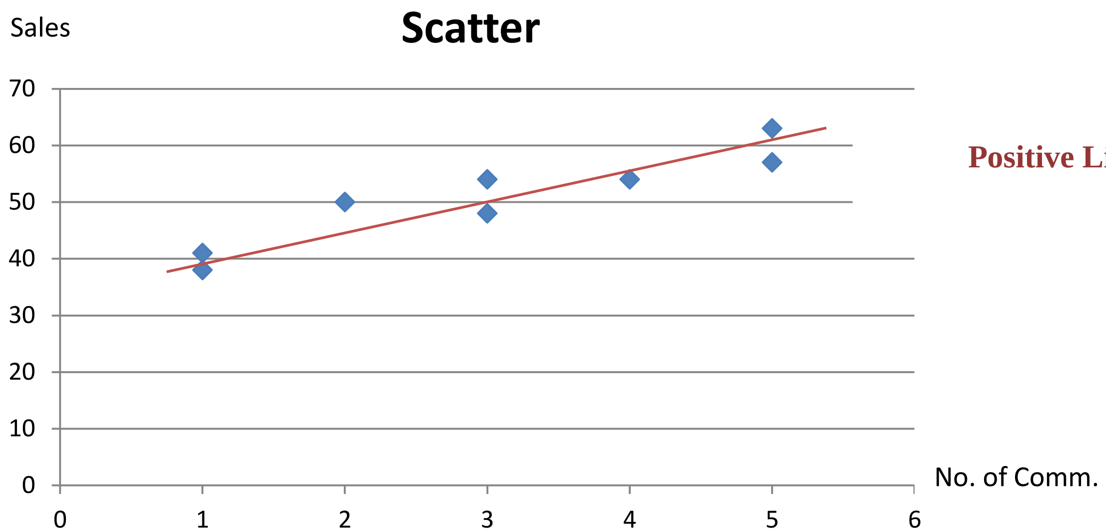
Class	Tally	Frequency	Cumulative freq.
10 - 14	ǀǀǀǀ	4	<= 14 → 4
15 - 19	ǀǀǀ 卌	8	<= 19 → 12
20 - 24	卌	5	<= 24 → 14
25 - 29	ǀǀ	2	<= 29 → 19
30 - 34	ǀ	1	<= 34 → 20
Total		20	



Salter Plot : is a graphical presentation of the relationship between two quantitative variables.

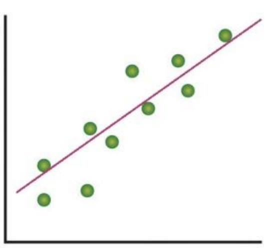
Exp :

Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
No. of Comm.	**2**	**5**	**1**	**3**	**4**	**1**	**5**	**3**	**4**	**2**
Sales ($100)	**50**	**57**	**41**	**54**	**54**	**38**	**63**	**48**	**59**	**46**

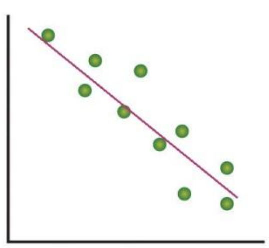


Positive Linear Relation Ship

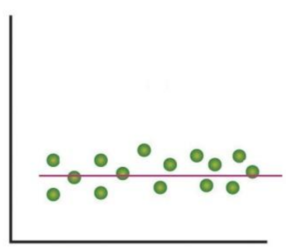
Types of Relation Ships :



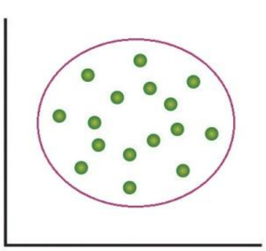
Positive linear
$y = ax + b$
$a > 0$



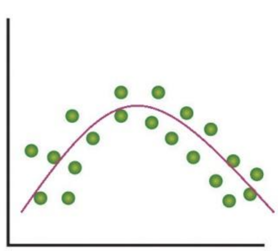
Negative linear
$y = ax + b$
$a < 0$



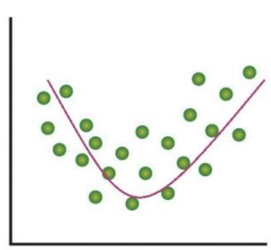
Cinstant Linear
$y = ax + b$
$a = 0$



Independent
No Relation Ship



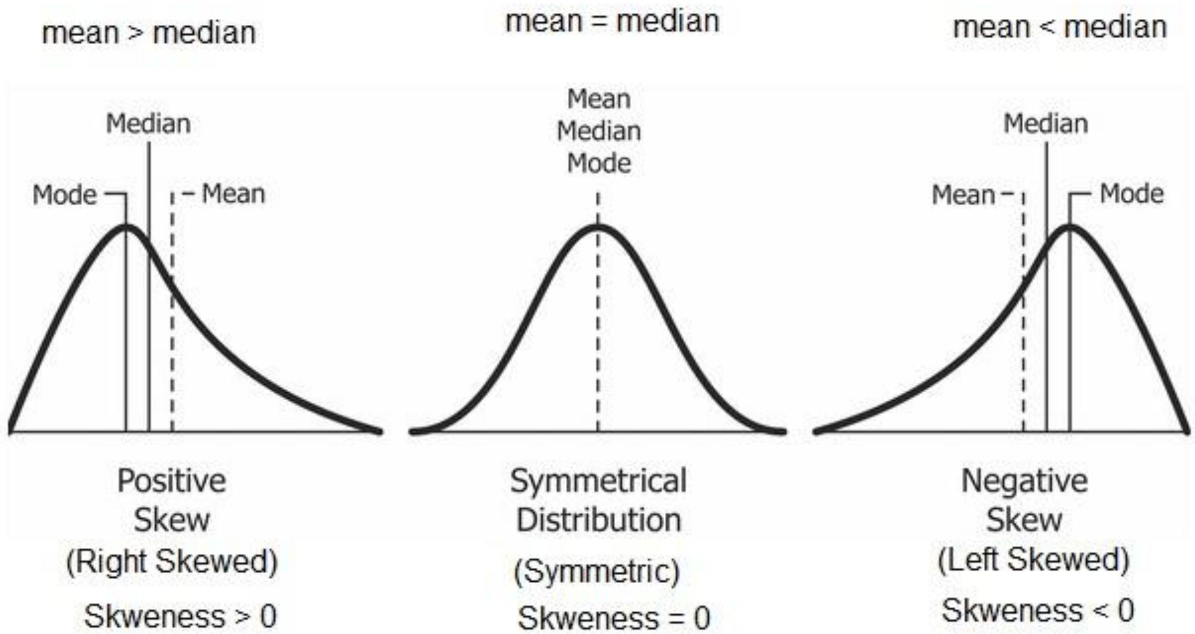
Curvilinear



Curvilinear

by Maryam Shaheen

$$\textbf{Skewness :}\ \frac{n}{(\text{n - 1})(\text{n - 2})}\sum_{i=1}^{n}\left(\frac{Xi-\overline{X}}{S}\right)$$



by Maryam Shaheen

Chapter 3

- **Mean :**

Population Mean: $\mu = \frac{\sum(x)}{N}$

Sample Mean: $\bar{x} = \frac{\sum(x)}{n}$

Frequency Distribution: $\bar{x} = \frac{\sum(xf)}{\sum f}$

- **Median :**

if **n is odd** : Median = the value in the middle (after data is Sorted)

if **n is even** : Median = average of the two values in the middle (after data is Sorted)

- **Mod =** The Value with Highest Frequency

- **Percentile :**

Step 1 : Sort the Data

Step 2 : find the index : i **=** $\frac{P}{100} * n$

if i (**non- integer**) → we round it up

if i (**integer**) →we take the average → ($\frac{(i)+(i+1)}{2}$)

- **Quartile :**

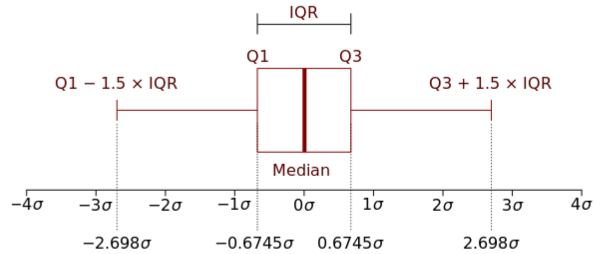
Quartiles are specific percentiles

First Quartile = 25th Percentile

Second Quartile = 50th Percentile = Median

Third Quartile = 75th Percentile

- **Range :** max - min
- **interquartile range (IQR) =** Q3 - Q1
- **Box Plot :**
  Upper Fence : Q3 + 1.5 (IQR)
  Lower Fence : Q1 - 1.5 (IQR)



- **Five Number Summary :** min, Q1, Q2 (median) , Q3, **max**
- **Extreme Outliers**
  Extreme outliers are any data values which lie more than 3.0 times the interquartile range below the first quartile or above the third quartile. x is an extreme outlier if ...
  x < Q1 - 3(IQR) or x > Q3 + 3(IQR)

- **Mild Outliers**
  Mild outliers are any data values which lie between 1.5 times and 3.0 times the interquartile range below the first quartile or above the third quartile. x is a mild outlier if ...
  Q1 - 3(IQR) <= x < Q1 - 1.5(IQR) or Q1 + 1. 5(IQR) < x <= Q3 + 3(IQR)

- **variance:**
  population Variance : $\sigma^2 = \frac{\sum (x-\mu)^2}{N}$
  Sample Variance : $s^2 = \frac{\sum (x-\bar{x})^2}{n-1}$

- **Standard Deviation :**
  Population Standard Deviation : $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x-\mu)^2}{N}}$
  Sample Standard Deviation : $s = \sqrt{s^2} = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$

Coefficient of Variance : CV = $\frac{S}{\bar{X}}$ *100%
if CV > 20% , then the data has high variation

by Maryam Shaheen

- **Z - Score :**

Population : $Z_i = \frac{Xi - \mu}{\sigma}$

Sample : $Z_i = \frac{Xi - \overline{X}}{S}$

Applications on Z - Score :

1) Chebyshev's Theorem :

The proportion of the values that fall within **K** standard deviations of the mean will be **at least** 1 - $\frac{1}{K^2}$, where **K** > 1.

Within **K** standard deviations" interprets as the interval: ($\overline{X}$ - KS - $\overline{X}$ + KS)

At least 75% of the data values must be within Z = 2 standard deviations of the mean.

interval : ($\overline{X}$ - 25 - $\overline{X}$ + 25)

At least 89% of the data values must be within Z = 3 standard deviations of the mean.

interval : ($\overline{X}$ - 35 - $\overline{X}$ + 35)

At least 94% of the data values must be within Z = 4 standard deviations of the mean.

interval : ($\overline{X}$ - 45 - $\overline{X}$ + 45)

Chebyshev's Theorem is true for any sample set, not matter what the distribution.

2) Empirical Rule :

The empirical rule is only valid for bell-shaped (normal) distributions.

Interval : ($\overline{X}$ - ZS - $\overline{X}$ + ZS)

Approximately 68% of the data values fall within Z =1 standard deviation of the mean.

Approximately 95% of the data values fall within Z =2 standard deviations of the mean.

Approximately 99.7% of the data values fall within Z =3 standard deviations of the mean.

- **Covariance :**

Population Covariance : $\sigma_{xy} = \dfrac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$

Sample Covariance : $s_{xy} = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$

- **Correlation Coefficient :**

Population Correlation Coefficient : $\rho_{xy} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$, $\sigma_x = \dfrac{\sum (\sigma_{xi} - \overline{\sigma_x})}{N}$, $\sigma_y = \dfrac{\sum (\sigma_{yi} - \overline{\sigma_y})}{N}$

Sample Correlation Coefficient : $r_{xy} = \dfrac{s_{xy}}{s_x s_y}$, $s_x = \dfrac{\sum (x_i - \bar{x})}{n-1}$, $s_y = \dfrac{\sum (y_i - \bar{y})}{n-1}$

Weighted Mean : $\bar{x} = \dfrac{\sum (\omega_i - x_i)}{\omega_i}$

Section 12.2

Regression Line : $\hat{y_i} = ax_i + b$
Error summation : $\sum_{i=1}^{n}(y_i - \hat{y_i})^2$ = min!
Q(a , b) = $\sum_{i=1}^{n}(y_i - ax_i - b)^2$

$$\frac{dQ}{da} = 0 \quad , \quad \frac{dQ}{db} = 0$$

$$\boldsymbol{a} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \rightarrow \boldsymbol{a} = \sum \frac{x_i y_i - n\bar{x}\bar{y}}{(n-1)\, s^2 x}$$

b = $\bar{y} - a\bar{x}$

Coefficient of determination = $r^2 = {r_{xy}}^2$
if $\boldsymbol{I_0} < \boldsymbol{r}$ **,** then the **relation is Strong**

by Maryam Shaheen

Chapter 4

OverView :

Exp1 : How many outcomes for a 5-digit **binary** number?

$(2)(2)(2)(2)(2) = 2^5 = 32$

Exp2 : How many outcomes for 10-digit number ?

$(10)\ (10)\ (10)\ (10)\ (10)\ (10)\ (10)\ (10)\ (10)\ (10)\ (10) = 10^{10}$

Exp3: How many outcomes for a 10-digit number with constant first 3-digits "059" ?

$(10)\ (10)\ (10)\ (10)\ (10)\ (10)\ (10) = 10^7$

- **The number of Combinations of n objects taken r at time :**

$$C_r^n = \frac{n!}{r!(n-r)!}$$

- **The number of Permutations of n objects taken r at time : (Order is important)**

$$P_r^n = \frac{n!}{(n-r)!}$$

n! = (n)(n-1)(n-2)....... (2)(1)
1! = 1
0! = 1

Probability :

1) Classical Method : P(outcome) = $\frac{1}{numberOfOutcomes}$ (if outcomes are equally likely)

2) Relative frequency Method : P(outcome) = $\frac{NumberOfTimesOutcomesOccured}{numberOfOutcomes}$

EVENTS :

Event : is a subset of S

Exp : S = {1,2,4,5,6}

Event 1 : E1 = {2,4,6}
Event 2 : E2 = {3,4,5,6}

Union : $E1 \cup E2$ = E1 OR E2 = {2,3,4,5,6}
Intersection : $E1 \cap E2$ = E1 AND E2 = {4,6}

Complement : $E1'$ = S - E1 = {1,3,5}

Rules :

Addition Law : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Complement Law : $P(A') = 1 - P(A)$
Multiplication Law : $P(A \cap B) = P(A) . P(B|A)$, $P(A \cap B) = P(B) . P(A|B)$
P(ϕ) = 0

if $P(A \cap B) = 0$ **OR** $A \cap B = \phi$, Then they are **Mutually Exclusive (Disjoint)**

Conditional Probability :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad , P(B|A) = \frac{P(A \cap B)}{P(A)}$$

if P(A|B) = P(A) OR , P (B|A) = P(B) , then the two events are **Independent**

if A, B are **Independent** , then $P(A \cap B) = P(A) . P(B)$

Bayes' Theorem :

$$P(A|B) = \frac{P(B|A).P(A)}{P(B|A).P(A) + P(B|A').P(A')}$$

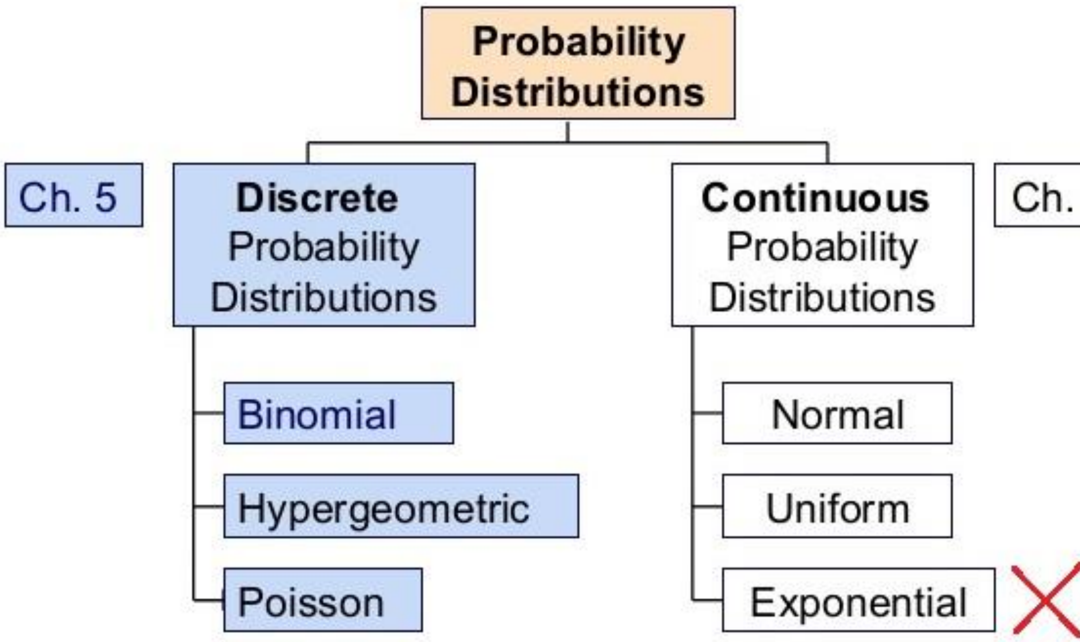
$$P(A_1|B) = \frac{P(B|A_1).P(A_1)}{P(B|A_1).P(A_1) + P(B|A_2).P(A_2)} = \frac{P(B|A_1)P(A_1)}{P(B)}$$

$$P(A_2|B) = \frac{P(B|A_2).P(A_2)}{P(B|A_1).P(A_1) + P(B|A_2).P(A_2)} = \frac{P(B|A_2)P(A_2)}{P(B)}$$

by Maryam Shaheen

Chapter 5 & 6

OverView :



Discrete Probability Distributions : (Chapter 5)

- **Probability function f(x) :**

1 - $0 <= f(x) <= 1$
2 - $\sum f(x) = 1$

- ## **Random Variables :**

Expected Value of X = E(x) = $\mu_x = \sum [xf(x)]$

f(x) = P = $\frac{1}{n}$ → $\mu_x = \frac{\sum x}{n}$ (Classical)

Variance = Var(x) = $\sigma_x^2 = \sum [(x-\mu_x)^2 . f(x)]$ = $\sum [(x^2 f(x) - \mu_x^2)]$

Standard deviation = $\sqrt{var(x)} = \sigma_x = \sqrt{\sum [(x-\mu_x)^2 . f(x)]} = \sqrt{\sum [(x^2 f(x) - \mu_x^2)]}$

- ## **Binomial :**

Probability of x occurs n trails : p(x) = f(x) = $\binom{n}{x} p^x (1-p)^{n-x}$

Expected Value of X = E(x) = $\mu_x = n.p$

Variance = Var(x) = $\sigma_x^2 = n.p(1-p)$

Standard deviation = $\sqrt{var(x)} = \sigma_x = \sqrt{n.p(1-p)}$

- ## **Poison :**

Probability of x occurs in an interval : p(x) = f(x) = $\frac{\mu_x e^{-\mu}}{x!}$

$\mu = \sigma^2$

by Maryam Shaheen

- **Hypergeometric :**

Probability of x successes in a sample of n size : p(x) = f(x) = $\dfrac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \dfrac{C_n^r C_{n-x}^{N-r}}{C_n^N}$

Expected Value of X = E(x) = $\mu_x = n.p = n.\left(\dfrac{r}{N}\right)$

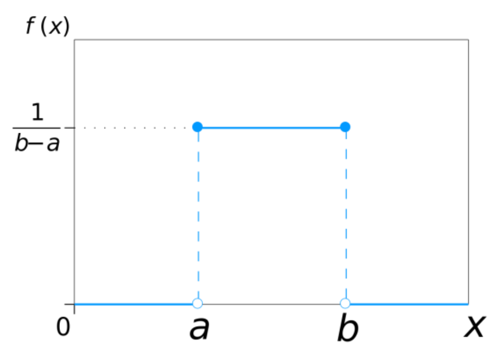
Variance = Var(x) = $\sigma_x^2 = n.\left(\dfrac{r}{N}\right)\left(1-\left(\dfrac{r}{N}\right)\right)\left(\dfrac{N-r}{N-1}\right)$

Standard deviation = $\sqrt{var(x)} = \sigma_x = \sqrt{n.\left(\dfrac{r}{N}\right)\left(1-\left(\dfrac{r}{N}\right)\right)\left(\dfrac{N-r}{N-1}\right)}$

Continuous Probability Distributions : (chapter 6)

- **Uniform :**

probability of x : $f(x) = \begin{cases} \dfrac{1}{b-a}, a <= x <= b \\ 0 \end{cases}$



Expected Value of X = E(x) = $= \dfrac{a+b}{2}$

Variance = Var(x) = $\sigma_x^2 = \dfrac{(b-a)^2}{12}$

Standard deviation = $\sqrt{var(x)} = \sigma_x = \sqrt{\dfrac{(b-a)^2}{12}}$

- **Normal :**

probability of x : $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, - ∞ <x < ∞

Z - Score : $Z = \dfrac{x-\mu}{\sigma}$

if **n.p >= 5 and n(1 - p) >= 5** , then **Binomial probs. can be approximated by Normal probs**. in this case : $\mu = n.p$

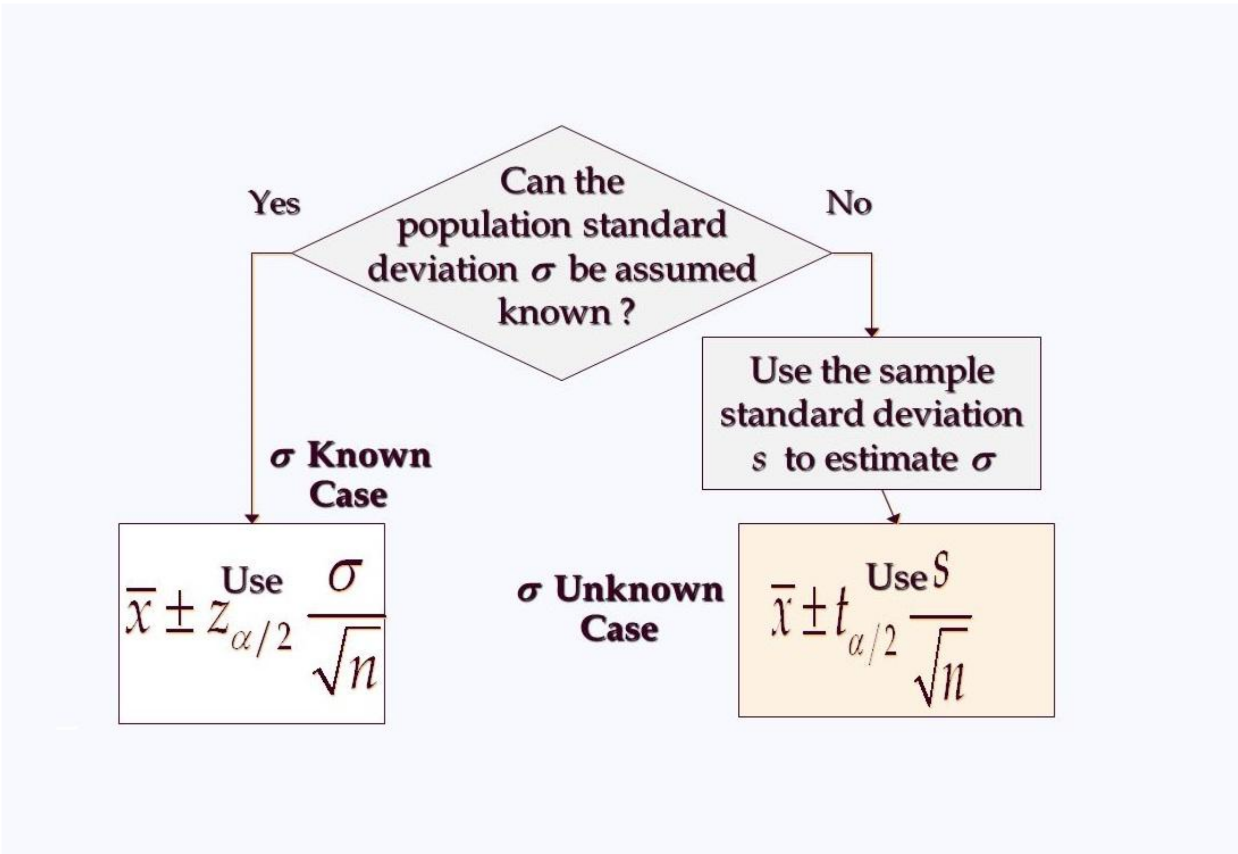
$\sigma^2 = n.p(1-p)$

$\sigma = \sqrt{n.p(1-p)}$

by Maryam Shaheen

Chapter 8

OverView :



- **When σ is Known :**

  Interval Estimate of μ : $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

  margin of error : E = $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

  confidence coefficient = 1 - α

- **When σ is Unknown :**

  Interval Estimate of μ : $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

  margin of error : E = $t_{\alpha/2} \frac{s}{\sqrt{n}}$

  confidence coefficient = 1 - α

  degrees of freedom : n - 1

- **Sample Size for an Interval Estimate of a Population Mean :**

  Necessary Sample Size : $n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2}$, E = $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

- **interval estimate of a population proportion =** $\bar{p} \pm z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

margin of error : E = $z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

confidence coefficient = 1 - α

Necessary Sample Size : $n = \frac{\left(Z_{\alpha/2}\right)^2 \bar{p}\left(1-\bar{p}\right)}{E^2}$

However, $\bar{p}$ will not be known until after we have selected the sample. We will use the planning value p* for $\bar{p}$ $\rightarrow$ $n = \frac{\left(Z_{\alpha/2}\right)^2 p^*\left(1-p^*\right)}{E^2}$

Chapter 9

Hypothesis testing can be used to determine whether a statement about the value of a population parameter should or should not be rejected.

The null hypothesis : H_0, is a tentative assumption about a population parameter.

The alternative hypothesis : H_a**, is** the opposite of what is stated in the null hypothesis.

Types of Error :

	Decision	
	Accept H_0	Reject H_0
H_0 (true)	Correct decision	Type I error (α error)
H_0 (false)	Type II error (β error)	Correct decision

A Type I error : is rejecting H_0 when it is true.

A Type II error : is accepting H_0 when it is false.

Hypothesis Test of population mean

- **When σ is Known :**

Test Type	Upper Tail Test	Lower Tail Test	Two Tailed Test
Hypotheses	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Given	$\bar{x}, \sigma, n, \alpha, \mu_0$		
Test Statistics	$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$		
P -value	Area in the Upper tail	Area in the Lower tail	Area in the Two tails
Rejection Rule (p- value approach)	Reject H_0 if p-value $\leq \alpha$		
Critical value(s)	z_α	$-z_\alpha$	$-z_{\alpha/2}$ and $z_{\alpha/2}$
Rejection Rule (Critical value approach)	Reject H_0 if $z \geq z_\alpha$	Reject H_0 if $z \leq -z_\alpha$	Reject H_0 if $z \geq z_{\alpha/2}$ OR $z \leq -z_{\alpha/2}$

- **When σ is Unknown :**

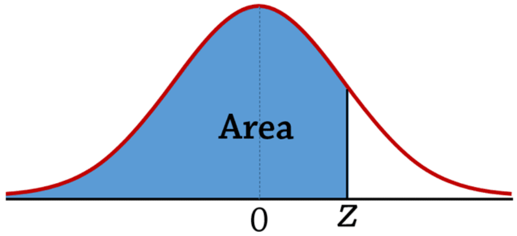
Test Type	Upper Tail Test	Lower Tail Test	Two Tailed Test
Hypotheses	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Given	$\bar{x}, \sigma, n, \alpha, \mu_0$		
Test Statistics	$t = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$ $df = n - 1$		
P -value	Area in the Upper tail	Area in the Lower tail	Area in the Two tails
Rejection Rule (p- value approach)	Reject H_0 if p-value $\leq \alpha$		
Critical value(s)	t_α	$-t_\alpha$	$-t_{\alpha/2}$ and $t_{\alpha/2}$
Rejection Rule (Critical value approach)	Reject H_0 if $t \geq t_\alpha$	Reject H_0 if $t \leq -t_\alpha$	Reject H_0 if $t \geq t_{\alpha/2}$ OR $t \leq -t_{\alpha/2}$

by Maryam Shaheen

Hypothesis Test of population Proportion

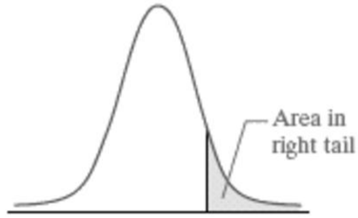
Test Type	Upper Tail Test	Lower Tail Test	Two Tailed Test
Hypotheses	$H_0: p \le p_0$ $H_a: p > p_0$	$H_0: p \ge p_0$ $H_a: p < p_0$	$H_0: p = p_0$ $H_a: p \ne p_0$
Given	$\overline{p}, n, \alpha, p_0$		
Test Statistics	$z = \dfrac{\overline{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$		
P -value	Area in the Upper tail	Area in the Lower tail	Area in the Two tails
Rejection Rule (p- value approach)	Reject H_0 if p-value $\le \alpha$		
Critical value(s)	z_α	$-z_\alpha$	$-z_{\alpha/2}$ and $z_{\alpha/2}$
Rejection Rule (Critical value approach)	Reject H_0 if $z \ge z_\alpha$	Reject H_0 if $z \le -z_\alpha$	Reject H_0 if $z \ge z_{\alpha/2}$ OR $z \le -z_{\alpha/2}$

by Maryam Shaheen



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

by Maryam Shaheen



Conf. Level	50%	80%	90%	95%	98%	99%
One Tail	0.250	0.100	0.050	0.025	0.010	0.005
Two Tail	0.500	0.200	0.100	0.050	0.020	0.010
df = 1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
50	0.679	1.299	1.676	2.009	2.403	2.678
60	0.679	1.296	1.671	2.000	2.390	2.660
70	0.678	1.294	1.667	1.994	2.381	2.648
80	0.678	1.292	1.664	1.990	2.374	2.639
90	0.677	1.291	1.662	1.987	2.368	2.632
100	0.677	1.290	1.660	1.984	2.364	2.626
z	0.674	1.282	1.645	1.960	2.326	2.576

by Maryam Shaheen