

د. ضياء مرهبي
0597627168

Unit 4

Measures of Location- Averages and Percentiles

نوع من
الاصناف
مجموعة جمع
المعلومات
وتصنيفها
وعرضها
لوصف
خواص البيانات

Descriptive statistics is a type of statistics that is concerned with gathering, classification, presentation, and the collection of summarized values to describe the group characteristics of data. It is a fundamental statistics that provides some data needed in inferential statistics.

Objectives:

- State the definition of terms commonly encountered in descriptive statistics.
- Construct a frequency table containing data and complete information.
- Use formulas in descriptive statistics correctly.
- Give simple analyses or description about data obtained.

Topics:

- Frequency Distribution
- Percent
- Rank
- Measures of Central Tendency
 - Mean
 - Median
 - Mode
- Weighted Value and Weighted Mean
- Percentiles

لا يخل النشر او التصوير
د. ضياء
0597627168

Part I. FREQUENCY DISTRIBUTION

قيمة توضح عدد الحالات أو عدد المرات التي تظهر فيها
الخاصة

Frequency is the value showing the number of cases or number of times a case (or cases) appear. A frequency distribution is a way of classifying statistical data that allows comparisons of the results of each category. It is the arrangement of statistical data (Encarta Dictionary). Frequency distribution is applicable if N is equal to or greater than 30 (Calmorin and Calmorin, 1997).

طريقة لتقسيم البيانات الاصلية ← مقارنة النتائج

Example of frequency:

If five persons were found to have elevated blood sugar as shown by the result of blood chemistry, then report or write that as 5.

Sometimes several information are given to form a group of related data. To facilitate and interpretation, they are shown in a frequency table.

Example of frequency distribution in a frequency table

Common Laboratory Examination Undertaken in Hospital "E"
for the Month of April, 2013.

Laboratory Exam	Frequency	Percent	Rank
Urinalysis	47	40.9	4
Fecalysis	32	27.8	3
Sputum test	11	9.6	1
Biopsy	25	21.7	2
Total	115	100.0	

↖ N

You will notice that the table gives complete information about the common laboratory examination conducted for a certain period. Apparently, the item with the lowest value or frequency is Rank 1. As the frequency increases, the rank also increases.

To change the frequency to percentage, use the following formula or rule:

$$\% = \frac{f}{N} \times 100$$

where: % = is the percentage value

f is the frequency; N is the total of frequencies

النسبة -
مجموع النسب = 100

Example: frequency (f) = 47 Total (N) = 115

$$\frac{47 \times 100}{115} = 40.86 \text{ or } 40.9\% \text{ but you are not going to put the \% sign in the table}$$

Take note that after changing the frequency to percent, the total should be 100 or 100%. If you get a lower or higher total, it is important for you to review your

calculations particularly when you have decimal numbers and you rounded-off the numbers.

Important: Consider the following when you round-off decimal numbers.

لو طانة لقيمة
بعد لفاصلة
اعل 0.5
الذي

A. If you want a whole number:
.1 to .4 = these would be 0 in value or you have to drop them.
Example: 29.3 = 29

.5 to .9 = increase the whole number to 1
Example: 43.6 = 44

لو طانة القيمة بعد لفاصلة
مكبر عدد اربعا 5
اجبرها

B. If 1 decimal place is needed:
.01 to .44 = Just take the first decimal number (or the 10th place)
Example: 34.21 = 34.2

.45 above = increase the first decimal place to 1
Example; 67.09 = 67.1

لو بدي خانة
واحدة بعد لفاصلة
الخانة الثانية
عكس اهنفزاو
اجبرها

لا يجل النشر او التصوير
د. حنان
0597627168

C. If 2 decimal places are needed:
Retain the first 2 decimal places if the third decimal place is less than 5
Example; 65.873 = 65.87

Increase the second decimal place if the third decimal place is 5 or higher
Example: 312.319 = 312.32

نفسا

Part II. MEASURES OF CENTRAL TENDENCY

مياى التزعة المركزية

For every group of scores, there are those ones which are noticed as "most". It is these scores which commonly appear, obtained or grouped. Thus, this is called central tendency which simply serves as the index in the clustering of the X-value towards the central value. There are 3 measures of central tendency. These are the mean, median and mode.

1. **Mean (or average)**. This is probably the most common way of describing central tendency. It is done by adding up all the values and then divide the total by the number of values.

Fundamentals of Biostatistics -34-

مقدرة لوصف
التزعة المركزية

مجم القيم
عدد القيم

3

Illustration: Ungrouped data:

84
79
93

$$\frac{84 + 79 + 93}{3} = 85.33$$

Total = 256

$256 \div 3 = 85.33 \rightarrow \text{Mean}$

(You can also calculate the Mean using the Excel).

لقد لمعونة
في تصنيف
القيم
أهم شي بين
القيم لصاندي
أوتشاركي

2. **Median.** This is the score found exactly at the middle of a set of values. A simple way for looking the median is to list down all the scores either from highest to lowest or from lowest to highest. The score located at the center of the samples is the median. If there are 2 values at the center, add them then divide it by 2. The resulting value is the median.

Illustration: 2 3 4 7 9
Median

لقد لعدد رقم واحد فقط
لقد لعدد رقمين خدي فتوسطهم

16 18 19 20 22 23 26 27

Middle numbers

$$20 + 22 = 42 \div 2 = 21 \rightarrow \text{Median}$$

the number repeated many

ملاحظة هنام الرصبي /

$$\text{odd} = \frac{n+1}{2} \rightarrow 3^{\text{rd}} \text{ one.}$$

$$\text{even} = \left(\frac{n}{2}, \frac{n}{2} + 1 \right) \rightarrow \text{مستوسطهم}$$

3. **Mode.** This is the score found as the most frequently occurring in a set of scores. An easy way to determine the mode is to arrange the scores in either ascending or descending order.

لقد لعدد الأكثر
تكراراً

Illustration: 44 37 37 32 45 30 37 42 45 33 36 39

Arrange them: 30 32 33 36 37 37 37 39 42 44 45

Mode

Important:

When mean is calculated through Microsoft Excel, it is easier. After encoding each value in a horizontal or column, click "fx". A box for statistical tools will appear. Click "AVERAGE" and the result will appear. Do not click "SUM" anymore if it is not required. This is true to both ungrouped and grouped data. However, the computer does not show anymore how the data are grouped.

Descriptive statistics about weighted values that results to **weighted mean** uses the **Likert-point scale**. The points are expressed in weights and arbitrary descriptions used depending upon the situation or condition where it is applied. It could be progressive or regressive when it is one directional; or integeric when descriptions contain both positive and negative descriptions. It is also known as two directional.

Illustration:

one directional

Progressive/Regressive

Weight - Arbitrary Descriptions

- 5 - Excellent
- 4 - Very Good
- 3 - Good
- 2 - Fair
- 1 - Poor

Integeric

positive and negative descrip.

Weight - Arbitrary Descriptions

- 5 - Strongly Agree
- 4 - Agree
- 3 - Uncertain
- 2 - Disagree
- 1 - Strongly Disagree

The weights in a Likert's-point scale is not always 5. It can be 3 to any number provided an arbitrary description is given. To facilitate interpretation, each weight (point) or a certain distance should have equivalent description. Here are some examples.

The laboratory department of Hospital Z released the following report in 2010 regarding laboratory examinations requested by different wards in that hospital.

27 * 5
51 * 4
19 * 3

امتنى $\sum Fx$ مع عدد مراتب
عدد مراتب وزن عدد مراتب في موقعا صلا /

Laboratory Examination	Always 5	Often 4	Sometimes 3	Rarely 2	Never 1	$(\sum fx)$	W_m	Rank	Interpretation
Urinalysis	27	51	19	13	10	432	3.60	2	Often
Fecalysis	15	28	42	30	5	378	3.15	1	Sometimes
Chest X-ray	29	47	30	7	7	444	3.70	3	Often
Blood test	77	36	4	2	1	546	4.55	4	Always
Overall Mean							3.75		Often

Some simple analyses that could be derived from the values are:

1. Blood test appears to be "Always" a laboratory examination requested while fecalysis is "Sometimes" requested by the different wards in Hospital Z.
2. Generally, the different wards in Hospital Z " Often" request for certain laboratory examinations.

سرها الصفحة السابقة

Procedure in Determining the Weighted Mean Value (Wm)

عدد عدد كل صنف
في الجدول
المجموع
 $\sum fx$

1. Determine the number of cases (N) per row or horizontally or in the manner the weights are presented.
2. Get the partial products of the frequencies and weight (fx).
3. Add the partial products to get the total product ($\sum fx$).
4. Divide the total product by the number of cases (N) to get the weighted mean value.
5. Give the appropriate description about the weighted mean value obtained.

The formula is: $Wm = \frac{\sum fx}{n}$

where: Wm is the weighted mean value

Σ is the summation sign

f is the frequency

X is the weight

N is the number of cases

Σfx is the sum of the partial products

لا يخلو النشر او التصوير
د. حنان
0597627168

The way to do it in the calculator is: $(27 \times 5) + (51 \times 4) + (19 \times 3) + (13 \times 2) + (10 \times 1) = \sum fx \div N$

Part IV PERCENTILES

Averages or mean value discussed so far are measures of central value, therefore, they locate the center or midpoint of a distribution. It may also be of interest to locate other points in the range. Percentiles do that.

Percentiles are values in a series of observations arranged in ascending order of magnitude which divide the distribution into 100 equal parts. Accordingly, 10th percentile should have 10% observations to the left and 90% to the right.

For example, the 25th percentile is that value of a variable such that 25% of the observations are less than that value and 75% of the observations are greater.

Fundamentals of Biostatistics -37-

مجموعة من البيانات
المرتبة تصاعدياً
حيث تقسم
الموزع
إلى 100 جزء

على جدول المثال

25th يعني 25% أقل من
أقل من 75% أعلى من

Calculation of percentile:

الترتيب →

The Pth percentile of a sample of n observations is that **value** of the variable with **rank $(P / 100)(1 + n)$** . If this rank is **not** an integer, it is rounded to the nearest half rank.

Example: Compute the 10th, 90th and 85th percentiles for the following birth weight data (Measured in gram):

3200, 2951, 2830, 2759, 3215, 3219, 2581, 2130, 3216,

Solution:

First arrange data in increased (or decreased) order

2130, 2581, 2759, 2830, 2951, 3200, 3215, 3216, 3219

لا يَحِلُّ النِّشْرُ أو التَّصْوِيرُ
د. جنان
0597627168

الأردن

The 10th percentile (or decile) is that value with rank $(10/100)(1 + 9) = 1$, so we take the value with rank 1, which is 2130 gram.

The 90th percentile is the value with rank $(90/100)(1 + 9) = 9$, so we take the value with rank 9, which is 3219 g.

The 85th percentile is the value with rank $(85/100)(1 + 9) = 8.5$ so we take the value halfway between the rank 8 and rank 9, which is $(1/2)(3216 + 3219) = 3217$ g.

Unit 5

Variability and its Measures

The previous lesson deals with measures of central tendency. This lesson is focused on measures of dispersion.

Objectives:

- To determine the type of the variability in the health sciences
- To state the definition of the different measures of dispersion.
- To calculate range, variance, standard deviation and standard of variation correctly.
- To give some simple analyses about the result of calculations obtained.

Topics:

- Types of Variability (dispersion)
- Measure of Variability (dispersion)
 - Range
 - Variance
 - Standard Deviation
 - Coefficient of Variation

لا يَحِلُّ النِّشْرُ او التَّصْوِيرُ
د. حنان
0597627168

Part I. Types of Variability

There are three main types of variability:

الحيوي
الطبيعي
التجريبية

- Biological variability
- Real variability
- Experimental variability.

أنواع الاختلاف

1. Biological variability

Individuals differ when compared as regards, sex, class and other attributes but the difference noted Variability and its Measures may be small and is said to occur by chance. Such a difference or variability is called **biological variability**

Example:

- Difference observed in the mean blood pressures of two persons.
- A small variation in birth rates of two cities.

Fundamentals of Biostatistics -44-

-1-

تضاد لها نفس العمر ولكن أطول تنميف (احتمالات) مزرية

Classification of biological variability

- Individual variability:** One student's height is 160 cm and another of the same age is 170 cm
- Periodical variability:** The same individuals show variations in temperature, pulse rate, blood pressure, WBC count, blood sugar, urea, cholesterol, etc. at different times of the day
- Class, group or category variability:** Height, weight, blood pressure, etc. vary from class to class depending on age, sex, caste, social status or nature of work
- Sampling variability:** there will be variability from one sample to another sample

نفس الأعمار
كذلك عندهم تقريبا
عند قياس حرارة مثلا
في أوقات مختلفة
من اليوم

الطول والوزن
في الفئط
تختلف مع كلاس
لكلاس اعتمادا
على العمر وطب
و الحالة الاجتماعية

الاختلاف مع عينة لأخرى

2. Real variability

For example the difference in the incidence of cancer among smokers and non-smokers may be due to **excessive** smoking and not due to **chance**. **Higher coronary incidence may be associated with high cholesterol in blood, tension, obesity and such other factors.** Standard error of mean and proportion help to verify and solve such problems. **If the variability is found to be real, the offending external factor has to be found and duly dealt with.**

الفرق هنا
سبب وليس
بالصدفة.

حل مع حقيقة ← العامل المسبب يجب التعامل معه.

لو الاختلاف

3. Experimental Variability

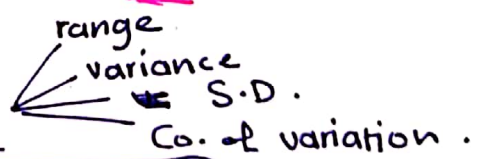
Error or difference or variation may be due to materials, methods, procedures employed in the study or defects in the techniques involved in the experiment.

They are of 3 types: **observer, instrumental and sampling.**

الاختلاف
سبب المواد
أو الطرق أو
الاجراء أو
التقنية.

Part II. MEASURES OF VARIABILITY (DISPERSION)

The spread of the values around the central tendency is called dispersion or variability. **Measures of dispersion serve as index of spread of x-values away from the central value.** There are 4 common measures of dispersion. These are range, variance, standard deviation and coefficient of variation.



1. **Range.** This is the **difference from the highest and lowest scores.**

Example: 62 = highest score
- 35 = lowest score
27 = **RANGE**

→ الفرق بين
أعلى قيمة و
أقل قيمة

معايير التشتت
تعمل كمؤشر
لانتشار القيم
ويجدها علم إحصائية
الأحاديث

لا يخل النشر أو التصوير
د. حنان
0597627168

تغير نسبة بعد الرقم عن المتوسط

2. **Variance** The variance measures how far each number in the set is from the mean. This is obtained by squaring each discrepancy after which sum up the values then divide it by the number of cases or scores (N-1).

Example:

$$V = \frac{\sum(x-M)^2}{N-1} = \frac{120}{12-1} = \frac{120}{11} = 10.91$$

تصنيف
→
أزلي لبيانات
مركزة أو ممتدة
حول المتوسط

3. **Standard deviation** The standard deviation measures how concentrated (or scattered) the data are around the mean; the more concentrated, the smaller the standard deviation. It is a value that describes the homogeneity or heterogeneity of a set of data being analyzed. It is computed by taking the square root of the variance.

Example:

$$s.d. = \sqrt{\frac{\sum(xi-M)^2}{N-1}} = \sqrt{\frac{120}{12-1}} = \sqrt{10.91} = 3.30$$

تصنيف
تجانس أو عدم
تجانس لبيانات

↓ SD → ↑ concen.
↑ SD → ↑ Scattering

Interpretation of sd values:

or
3 - below = homogenous
above 3 = heterogeneous

4. **Coefficient of Variation**. This is the ratio of the standard deviation to the mean. (The higher the coefficient of variation, the greater the level of dispersion around the mean). It is generally expressed as a percentage. The lower the value of the coefficient of variation, the more precise the estimate. This is obtained by dividing the value of standard deviation (sd) by the mean and multiplied by 100. The value is expressed in percent (%).

$$\text{Example: } CV = \frac{sd}{Mean} \times 100 = \frac{3.30}{8} = 41.28 \%$$

↑ Co. of variation → ↑ dispersion.
↓ " " → ↑ precision

لا يحل النشر أو التصوير
د. حنان
0597627168

Unit 6

Normal Distribution and Standard Normal Distribution

لوفي كبرية

من البيانات

الكلمة

يجب الجمع

في جدول

freq. distribution .

Large quantitative data, collected about any variable has to be summarized by making a frequency distribution table as explained in Unit 3. This frequency distribution describes how the data are distributed around the mean . Further analysis of data is determined primarily by the shape of this distribution.

هذا التوزيع
يسمى قد تسمى لبيانات متوزعة حول المتوسط
المعروفة .

Objectives:

1. To explain and understand the probability distributions of a random variable.
2. To find probabilities associated with a normal probability distribution (i.e., z distribution) using the standard normal table.
3. Apply the basic knowledge of normal distribution to solve problems.

Topics:

1. Probability definition and Properties.
2. Random Variable and Probability Distribution.
3. Normal probability distribution and Properties.
4. Using Table of standardized normal distribution.

لا يعلو النشر او التصوير
د. خندان
0597627168

1. Probability definition and Properties

Definition: probability is the numerical assessment of the likelihood of the occurrence of an event in the experiment that is repeated on the same conditions.

The probability close to **zero**, $P(A) \sim 0$ means that the event has a small chance of occurrence and does not necessarily mean it will not occur.

$P(A)=1$, means that the event will occur with certainty

Example: There is a 90% probability of a fading pain two days after surgery.

- Probabilities can express as decimal subdivision or fraction then the probability **is between 0 and 1 values**.
- It can be expressed as a proportion, then the probability values are between 0.00 and 100.00 percent

The probability is a value between 0 and 1, i.e. **not** less than 0 and not **greater** than 1

الاحتمالية هي
تقييم عددي
لاحتمالية
حدث ما
في تجربة تكرر
تحت الظروف

احتمالية
صغير
صغيره
وهذا
لا يعني أنه
لديك !!!

الاحتمالية (0-1)

من أصغر من صفر
ولا أكبر من 1

جميع النتائج
المحتملة لتجربة
محصية
الحدث
جزء من
فضاء
النتائج

فضاء
النتائج

Definition: the **sample space** is the set of alle possible outcomes of an experiment denoted by Ω (omega, Greek letter)

Definition: the **Event is** a subset of the sample space, denoted by A, B, C, ... or E1, E2,

Definition: Probability of any event is given by the formula

$$P(A) = \frac{N(A)}{N(\Omega)} = \frac{\text{عدد مرات حدوث A}}{\text{Total number of outcomes in the samplespaces}} = \frac{\text{عدد مرات النتائج كلها}}{\text{Total number of outcomes in the samplespaces}}$$

Basic examples

- Example 1:** when a fair coin is tossed twice, or two fair coins are tossed once then the sample space is: $\Omega = \{HH, HT, TH, TT\}$. There are 4 possible outcomes in this experiment. Each outcome is equally likely to occur $P(HH) = 1/4, P(HT) = 1/4, P(TH) = 1/4, P(TT) = 1/4$.
- Example:** calculate the probability that you get at least one Head in this experiment: $P(\{HH, HT, TH\}) = 3/4$
- Example:** calculate the pr. of getting no head: $P(TT) = 1/4$.
- Example:** pr. of getting T and H: $P(\{TH, HT\}) = 2/4 = 1/2$
- Example:** suppose the possible drugs prescribed by dentists that may affect smell or taste ampicillin, codeine, tetracycline. Calculate the probability of the event $A = \{\text{codeine}\}$.
Solution: firstly the sample space is $\Omega = \{\text{ampicilline, codeine, tetracycline}\}$. $A = \{\text{codeine}\}$, then $P(A) = 1/3$

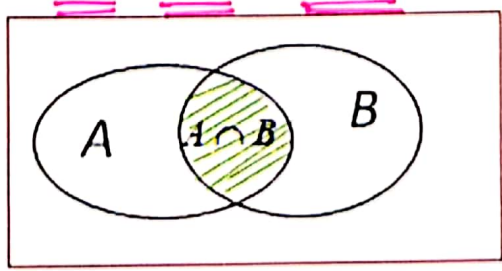
لا يعلَن النشر أو التصوير
د. حنان
0597627168

Basic properties of probability

a) The **union** of the events A and B is given by the formula

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

probability of
A or B

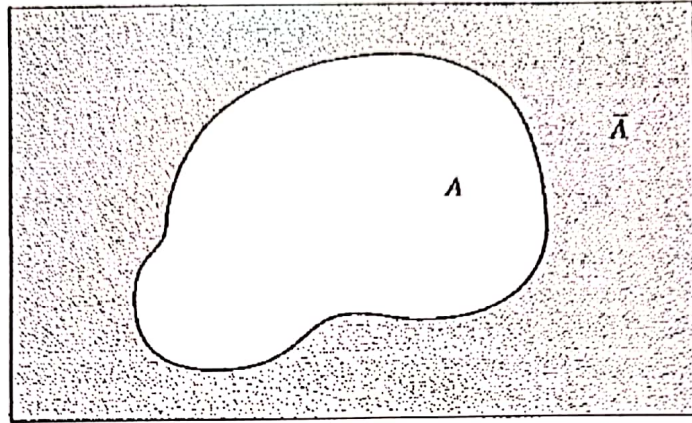


OR $\rightarrow \cup$ (+)

مكمل

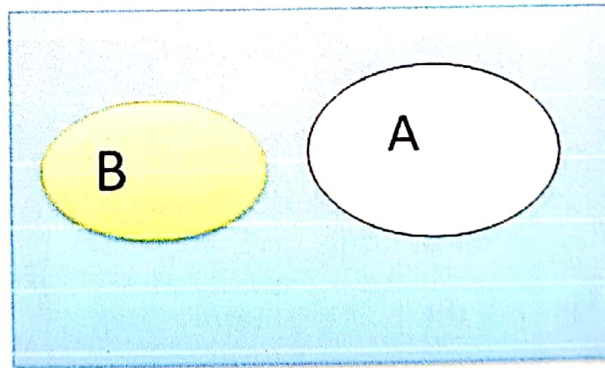
It is the probability of A or B or both
b) The **complement** of an event A is the event not A denoted by \bar{A}

$$P(\bar{A}) = 1 - P(A)$$



عندما لا
تتعا لهما
أر لا يجتمع
سواء
نقول عليهما
Mutual exclusive

c) **Mutual exclusion**, if two events A and B cannot happen at the same time i.e., we say that A and B are mutually exclusive events



In this case $P(A \cup B) = P(A) + P(B)$, since $P(A \cap B) = 0$

الاحتمالية
المشترطة

d) **Conditional Probability**
The conditional probability of A given B is defined as

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Example: A symptom of myocardial infarction in a patient (Male) is a high level of cholesterol in the blood (lets say of more than 320 mg per 100 ml). Let B event expresses the high level of cholesterol, and A, that the patient is exposed to myocardial infarction, and if

$$P(A/B) = 0.28 \quad \& \quad P(B) = 0.14$$

MI

كولسترول

and $\rightarrow \cap$ (متراب)

what is the probability that a patient is exposed to myocardial infarction ^A and ^B has a high cholesterol level?

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A/B) \cdot P(B) = (0.28) \cdot (0.14) = 0.0392$$

لا يخل النشر او التصوير
د. عثمان
0597627168

e) **Independence**
If $P(A/B) = P(A)$, then A is independent of B.

That mean the probability of occurrence of B does not affect the probability of occurrence of A .
equivalently

$$P(A \cap B) = P(A) \cdot P(B) \Rightarrow A \text{ and } B \text{ are independent}$$

Example:
Table (I): Distribution of 60 patients at the chest department of Jazan hospital in May 2017 according to smoking & lung cancer

Smoking	Lung cancer				Total	
	positive		negative		No.	%
	No.	%	No.	%		
Smoker	15	65.2	8	34.8	23	100
Non smoker	5	13.5	32	86.5	37	100
Total	20	33.3	40	66.7	60	100

Assume we draw a person at random from this table.
 What is the probability that this person will be a smoker?
 $P(\text{smoker}) = \text{Number of smoker} / \text{Total number of patients} = 23/60 = 0.38$
 What is the pr. that the person will be non smoker if you know that his test result was negative lung cancer.
 $P(\text{Non smoker/neg. cancer}) = P(\text{N.s and n.l.c.} / \text{total of lung c}) = 32/40 = 0.8$

2. **Random Variable and Probability Distribution**

المقياس العشوائي التوزيع الاحتمالي

نتیجہ ممکنہ طور پر عشوائی

A **random variable**, usually written as X , is a variable whose possible values are **numerical outcomes of a random phenomenon**.

There are two types of **random variables**, **discrete** and **continuous**.

Example: for discrete random variable: Suppose a variable X can take the values 1, 2, 3, or 4.

The probabilities associated with each outcome are described by the following table:

random variable →

Outcome	1	2	3	4
Probability	0.1	0.3	0.4	0.2

←

The probability that X is **equal to 2 or 3** is the sum of the two probabilities:

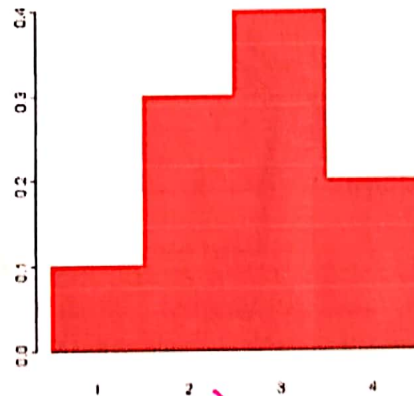
is the sum of the two probabilities:

$$P(X = 2 \text{ or } X = 3) = P(X = 2) + P(X = 3) = 0.3 + 0.4 = 0.7$$

Probability Distribution (histogram)

عمل جدول
الباقي

- This distribution of above table may also be described by the **probability histogram** as



لا يَحِلُّ النِّشْرُ او التَّصْوِيرُ
د. حنان
0597627168

#

- The Area of all bars is equal to 1

3. Normal probability distribution and its properties

Definition: it is a Bell-shaped continuous distribution widely used in statistical inference. It is also frequently called the Gaussian distribution, after the well-known **mathematician Karl Friedrich Gauss** (Figure 3.1).

توزيع منحنى

على خط جرس
يستخدم في الاستدلال
الاحتمالي

Karl Friedrich Gauss (1777–1855)



بكرة

Most of the **biomedical** data follow **normal distribution**.

Example: if we collected data of blood pressure values of a large number of individuals, then the frequency distribution of blood pressure will be a Bell-shaped curve which is symmetrical curve

Anormal Distribution is a theoretical model of the whole population about the central value, where μ is the mean of x and σ is standard deviation of x . When $\mu=0$ and $\sigma=1$ then it is called **standard normal distribution**

population
(parameter) $\left[\begin{array}{l} \mu = \text{mean} \\ \sigma = \text{SD} \end{array} \right.$

Sample .
(statistic) $\left[\begin{array}{l} \bar{x} = \text{mean} \\ s = \text{SD} \end{array} \right.$

Fundamentals of Biostatistics -54-

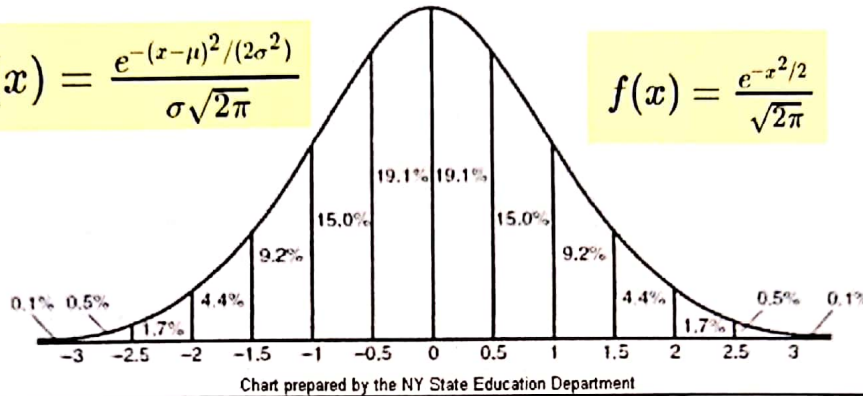
لا يَحِلُّ النِّشْرُ او التَّصْوِيرُ
د. حنان
0597627168

Normal Probability Distribution

Normal Curve
Standard Deviation

$$f(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}$$

$$f(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$



خود انوع الطبيعي

Properties of Normal distribution

1. The normal curve is symmetrical about the mean μ ;
2. The mean is at the middle and divides the area into halves;
3. The total area under the curve is equal to 1;
4. It is completely determined by its mean and standard deviation σ (or variance σ^2)

Note:

In a normal distribution, only 2 parameters are needed, namely μ and σ^2 .

5. The % of the data fall within 1, 2 or 3 standard deviation of the mean are

$$[\mu - \sigma, \mu + \sigma] \rightarrow 68.3\%$$

$$[\mu - 2\sigma, \mu + 2\sigma] \rightarrow 95.5\%$$

$$[\mu - 3\sigma, \mu + 3\sigma] \rightarrow 99.7\%$$

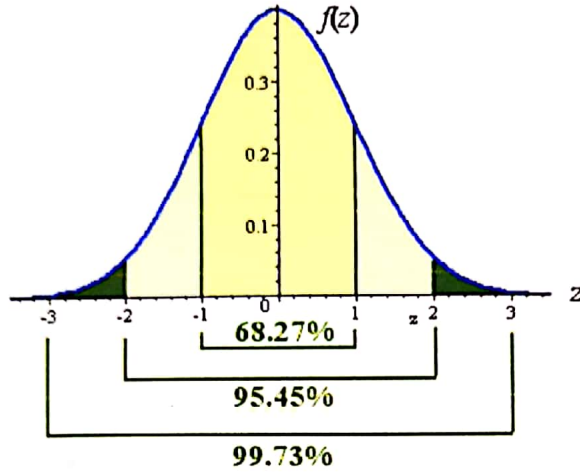
✓

Which is called **The Standard Deviation Rule**

This means: $P(-1 \leq Z \leq +1) = 0.683$ or 68% where z is a standard normal variable
 $P(-2 \leq Z \leq +2) = 0.95$ or 95%

$$P(-3 \leq Z \leq +3) = 0.997 \text{ or } 99.7\%$$

A graph of this standardized (mean 0 and variance 1) normal curve is shown.



Using Table of standardized normal distribution

Example: find $P(Z \leq 1.65) =$

$$P(Z \leq 1.65) = 0.9505$$

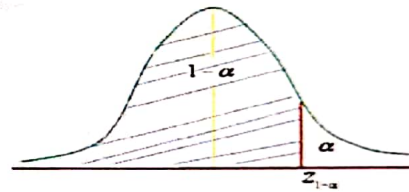


Table D

Table D. Standard Normal Probabilities $P(Z < z)$.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.00	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.10	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.20	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.30	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.40	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.50	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.60	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.70	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.80	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.90	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.00	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.10	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.20	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.30	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.40	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.50	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.60	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545

Example: find the area under the curve

Solution: $P(Z \geq 1.52) =$
 $= 1 - P(z \leq 1.52)$
 $= 1 - 0.9357 = 0.0643$

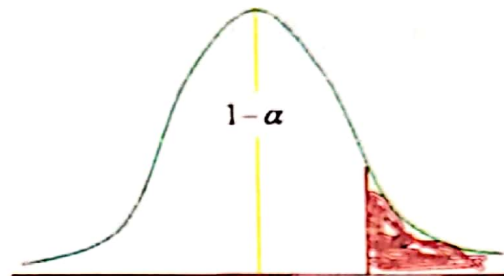


Table D

Table D. Standard Normal Probabilities $P(Z < z)$.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.00	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.10	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.20	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.30	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.40	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.50	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.60	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.70	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.80	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.90	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.00	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.10	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.20	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.30	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.40	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.50	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.60	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545

4. Solving problem using N.D tables

Remark: Transform normal distributed variable to standard normal distributed variable

If X normal distributed and its Mean is M and its standard deviation SD then The variable $Z = \frac{X-M}{SD}$ will be standard normal distributed.

Application: Assuming the normal Heart Rate (H.R) in normal healthy individuals is normally distributed with Mean=70 and Standard deviation =10 beats/min.

Find the probability that a person has a heart rate less than 80 beats/min.

Solution: what know is Mean=70, SD=10 with normal distribution

What want to know $P(X < 80) = ?$.

$$P(x < 80) = P[(x-70)/10 < (80-70)/10] = P(z < 1) = 0.8413$$

$$Z = \frac{80 - 70}{10} = \frac{10}{10} = 1$$

$$P(z < 1) = 0.8413$$

Table D. Standard Norm

Z	.00	.01
0.00	.5000	.5040
0.10	.5398	.5438
0.20	.5793	.5832
0.30	.6179	.6217
0.40	.6554	.6591
0.50	.6915	.6950
0.60	.7257	.7291
0.70	.7580	.7611
0.80	.7881	.7910
0.90	.8159	.8186
1.00	.8413	.8438
1.10	.8643	.8665

لا يَحِلُّ النِّشْرُ او التَّصْوِيرُ
د. حنان
0597627168