

**COMPUTER AND INFORMATION SCIENCE
MANSOURA UNIVERSITY**



CIS Team

YOUR PATH TO SUCCESS

GENERAL

1ST

   @cisteam15

The basic principle of counting

Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of m possible outcomes and if, for each outcome of experiment 1, there are n possible outcomes of experiment 2, then together there are $(m \cdot n)$ possible outcomes of the two experiments.

افترض ان عندي تجربتين التجربه الاولى هيمنتج منها عدد نواتج M وكل M من التجربه الاولى لها عدد اختيارات / نواتج N من التجربه الثانيه فيكون الناتج الكلي للتجربتين يساوي $M \cdot N$

EXAMPLE 2a

A small community consists of 10 women, each of whom has 3 children. If one woman and one of her children are to be chosen as mother and child of the year, how many different choices are possible?

Solution

عندي 10 سيدات بكم طريقه اختار سيده من الـ 10 وكل سيده اختار ليها 3 اطفال
Answer = $10 \cdot 3 = 30$

The generalized basic principle of counting

If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes; and if, for each of these n_1 possible outcomes, there are n_2 possible outcomes of the second experiment; and if, for each of the possible outcomes of the first two experiments, there are n_3 possible outcomes of the third experiment; and if ..., then there is a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes of the r experiments.

هنا تعميم لمبدأ العد السابق فلو عندي عدد من التجارب R ينتج منها عدد من النواتج N_1 من التجربه الاولى و N_2 من التجربه الثانيه وهكذا حتي التجربه R ينتج منها عدد من النواتج N_R فيكون الناتج الكلي هو $(N_1 \cdot N_2 \cdot \dots \cdot N_R)$

EXAMPLE 2d

How many functions defined on n points are possible if each functional value is either 0 or 1?

Solution.

عندي عدد من $(\text{Function} = n)$ وكل Function فيهم ليها احتمالان (0 او 1) فعدد Function اللي هقدر اكونها (2^n)

EXAMPLE 2e

How many different 7-place license plates are possible if the first 3 places are to be occupied by letters and the final 4 by numbers? And how many license plates would be possible if repetition among letters or numbers were prohibited?

Solution.

عندي لوحه من 7 اماكن اول 3 للحروف والحروف 26 حرف باعتبار اني بتعامل بالانجليزيه والارقام 10 ارقام
 (1) مسموح عندي بالتكرار
Answer = $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10$
 (2) مش مسموح عندي التكرار
Answer = $26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7$

PERMUTATIONS

Suppose now that we have n objects. Reasoning similar to that we have just used for the 3 letters then shows that there are $n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1 = n!$

EXAMPLE 3b

A class in probability theory consists of 6 men and 4 women. An examination is given, and the students are ranked according to their performance. Assume that no two students obtain the same score.

(a) How many different rankings are possible?

(b) If the men are ranked just among themselves and the women just among themselves, how many different rankings are possible?

Solution.

عندي شرح لكورس ما حضره 6 رجال و 4 سيدات و تم اجراء اختبار لهم ورتبت درجات الاختبار بحيث مفيش 2 حصلوا علي نفس الدرجة

Answer = 10!

(1) بكم طريقه ممكن ارتب الـ 10 من حيث الدرجات

Answer = 6! * 4!

(2) ازاي ارتبهم من حيث الدرجات بردوا بس الرجال علي حدي والسيدات علي حدي

EXAMPLE 3c

Ms. Jones has 10 books that she is going to put on her bookshelf. Of these, 4 are mathematics books, 3 are chemistry books, 2 are history books, and 1 is a language book. Ms. Jones wants to arrange her books so that all the books dealing with the same subject are together on the shelf. How many different arrangements are possible?

Solution.

الاول ترتيب الكتب = $4! * 3! * 2! * 1!$ بعد كده ترتيب الكتب في الارفف = $4! * 4! * 3! * 2! * 1!$

COMBINATIONS

We define $\binom{n}{r}$, for $r \leq n$, by

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

and say that $\binom{n}{r}$ represents the number of possible combinations of n objects taken r at a time.*

EXAMPLE 4a

A committee of 3 is to be formed from a group of 20 people. How many different committees are possible?

Solution Answer = ${}^{30}C_3 = (20)! / (20-3)! * (3)!$

عايز اختار 3 اشخاص من 20

EXAMPLE 4b From a group of 5 women and 7 men, how many different committees consisting of 2 women and 3 men can be formed? What if 2 of the men are feuding and refuse to serve on the committee together?

Solution.

Answer = $({}^5C_2) * ({}^7C_3) = [5! / (3! * 2!)] * [7! / (4! * 3!)] = 350$

الاول هجيب احتمال الكل

هجين احتمال عمل الرجلان المتخاصمين معا واطرحه من الاحتمالات الكلية لعمل الرجال التي تساوي 35
 $300 = 10 * 30 = 10$ فالناتج الكلي هو $30 = 5 - 35$ هبقي $(^2C_2) * (^5C_1) = 5$

MULTINOMIAL COEFFICIENTS

If $n_1 + n_2 + \dots + n_r = n$, we define $\binom{n}{n_1, n_2, \dots, n_r}$ by

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1! n_2! \dots n_r!}$$

Thus, $\binom{n}{n_1, n_2, \dots, n_r}$ represents the number of possible divisions of n distinct objects into r distinct groups of respective sizes n_1, n_2, \dots, n_r .

EXAMPLE 5a

A police department in a small city consists of 10 officers. If the department policy is to have 5 of the officers patrolling the streets, 2 of the officers working full time at the station, and 3 of the officers on reserve at the station, how many different divisions of the 10 officers into the 3 groups are possible?

Solution.

حيث ان الـ 10 عدد الاحتمالات الكلية والـ 5 عدد احتمالات الضباط المتواجدين بالشارع والـ 2 عدد الضباط المتواجدين دائما بالقسم والـ 3 عدد الضباط الاحتياطي بالقسم

$$\frac{10!}{5! 2! 3!} = 2520$$

EXAMPLE

Using the digits (1, 2, 3, 5) how many 4 digits number can be formed if :-

- 1) The first digit must be 1 and repetition of the digits is allowed.
- 2) The first digit must be 1 and repetition of the digits is not allowed.
- 3) The number must be divisible by 2 and repetition is allowed.
- 4) The number must be divisible by 2 and repetition is not allowed.

Solution.

Answer = $1 * 4 * 4 * 4 = 64$

Answer = $1 * 3 * 2 * 1 = 6$

(1) عندي شرط اول خانه تكون 1 والـ 3 خانات الباقيين مسموح فيهم التكرار
 (2) عندي شرطين اول خانه تكون 1 والـ 3 خانات الباقيين مش مسموح فيهم التكرار
 (3) عندي شرط ان الرقم يكون يقبل القسمة علي الـ 2 فالاحاد بتاعه لازم يبغي زوجي وعندي رقم واحد بس زوجي فالخانه بتاعت الاحاد هيبقي فيها رقم واحد اللي هو الزوجي والـ 3 خانات الباقيين مسموح فيهم التكرار

Answer = $4 * 4 * 4 * 1 = 64$

Answer = $3 * 2 * 1 * 1 = 6$

(4) عندي شرطين انه يقبل القسمة علي الـ 3 خانات الباقيين مش مسموح فيهم التكرار

EXAMPLE

A coin is tossed 3 times what is the total number of all possible outcomes?

Solution.

$$\underline{2^3 = 2 * 2 * 2 = 8}$$

العمله لها وجهان وهختبرها 3 مرات فالنتائج 2 اسس عدد الاختبارات

EXAMPLE

A coin is tossed and a die is rolled what is the total number of all possible outcomes from a coin with a die?

Solution.

$$\underline{2 * 6 = 12}$$

العمله لها احتمالان و حجر نرد له 6 فالنتائج حاصل ضرب احتمالات الـ 2

EXAMPLE

How many choices of 5 digit numbers divisible by 3 using digits (0,1,2,3,4,5) without repeating any digits?

Solution.

الارقام التي تقبل القسمة على الـ 3 هي الارقام التي مجموعها يقبل القسمة على الـ 3 فالارقام التي تقبل من دول هي (1,2,3,4,5) و (0,1,2,4,5) لان مجموعهم 15 و 12 والـ 2 يقبلان القسمة على الـ 3

$$5 * 4 * 3 * 2 * 1 = 120$$

&

$$4 * 4 * 3 * 2 * 1 = 96$$

النتائج الكلي هي مجموع نواتج الاحتمالات معا = $216 = 96 + 120$

CIS Team

Statistics is the science of collecting, organizing, summarizing, and analyzing information to draw conclusions or answer questions. In addition, statistics is about providing a measure of confidence in any conclusions.

ده علم قائم علي 4 حاجات هما (Collecting, organizing, summarizing, and analyzing information) يستخدم للوصول لحل مشكله ما وهدفها الاساسي وصف وفهم المصدر

Population is the entire group of individuals to be studied مجموعته كامله من الافراد هيتم دراستها

An individual is a person or object that is a member of the population being studied.

الشخص او الكائن اللي هيتم دراسته

A sample is a subset of the population that is being studied. جزء من المجموعه اللي هيتم دراستها

Descriptive statistics consist of organizing and summarizing data. Descriptive statistics describe data through numerical summaries, tables, and graphs. الوصف الاحصائي ودا بيتيم عن طريق تنظيم وتلخيص البيانات من خلال جداول وملخصات رقميه

Inferential statistics uses methods that take results from a sample, extends them to the population, and measures the reliability of the result.

الـ 2 دول بيبقوا علي شكل اسئله

A parameter is a numerical summary of a population هنا بيبقي عندي كلمه داله ف السؤال زي all

A statistics is a numerical summary of a sample هنا بيبقي عندي كلمه داله ف السؤال زي sample

The process of statistics

- 1) Identify the research objective معرفه اهداف وابعاد البحث او المشكله اللي عندي
- 2) Collect the data needed to answer the question posed in (1) تجميع البيانات اللازمه لحل المشكله
- 3) Describe the data حساب نسبه البيانات اللي انا جمعتها
- 4) Perform inference بختبر اذا كانت الفروض اللي انا وضعتها صح ولا غلط

الـ 2 بيبقوا علي شكل اسئله

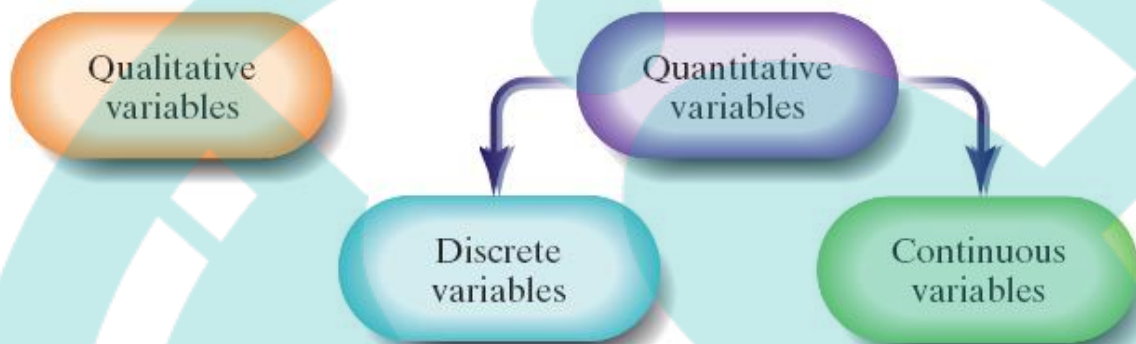
Qualitative or Categorical variables allow for classification of individuals based on some attribute or characteristic. تصنيف بناءً علي بعض الخصائص بيبقي فيه ف السؤال حاجه تدل علي صفه

Quantitative variables provide numerical measures of individuals. Arithmetic operations such as addition and subtraction can be performed on the values of the quantitative variable and provide meaningful results. دي حاجه اقدر اعمل عليها عمليه حسابيه بلاقي كلمه ف السؤال بتشير لعدد او كميه

A discrete variable is a quantitative variable that either has a finite number of possible values or a countable number of possible values. The term "countable" means the values result from counting such as 0, 1, 2, 3, and so on. دا متغير كمي بيبحتوي علي عدد محدود او محسوب من القيم

A continuous variable is a quantitative variable that has an infinite number of possible values it can take on and can be measured to any desired level of accuracy.

دا بردوا متغير كمي بس بيحتوي علي عدد لا نهائي من القيم التي يمكن اتخاذها



An observational study measures the value of the response variable without attempting to influence the value of either the response or explanatory variables.

هنا بعمل مشاهدته وملاحظته علي شئ معين من غير ماثري انا عليه

designed experiment If a researcher assigns the individuals in a study to a certain group, intentionally changes the value of the explanatory variable, and then records the value of the response variable for each group.

هنا بجيب شئ ما اثر انا عليه الاول وبعدين اشوف نتيجته التاثير اللي انا عملته ده

Random sampling is the process of using chance to select individuals from a population to be included in the sample.

عملية اختيار فرد من العينة اللي عندي عشوائي

Sample method the method of choose sample

طرق اختيار العينة

A sample of size n from a population of size N is obtained through **simple random sampling**

If every possible sample of size n has an equally likely chance of occurring. The sample is then called **a simple random sample**.

Sample of size — **(n) دي رمز الـ** **

Population of size — **(N) دي رمز الـ** **

A stratified sample is one obtained by separating the population into homogeneous, non-overlapping groups called *strata*, and then obtaining a simple random sample from each stratum.

بقسم المجموعات لمجموعات متجانسه غير متداخله تسمى طبقه وبعدها اختار منها عينه عشوائيه بسيطه

A systematic sample is obtained by selecting every k^{th} individual from the population. The first individual selected is a random number between 1 and k .

بختار فرد من الافراد اللي عندي والافراد دي كل واحد فيهم ليه رقم عشوائي بين 1 و K

STEPS IN SYSTEMATIC SAMPLING, POPULATION SIZE KNOWN

Step 1: Determine N .

Step 2: Determine n .

Step 3: $K = N/n$.

Step 4: select a number between 1 and k and call this number (p).

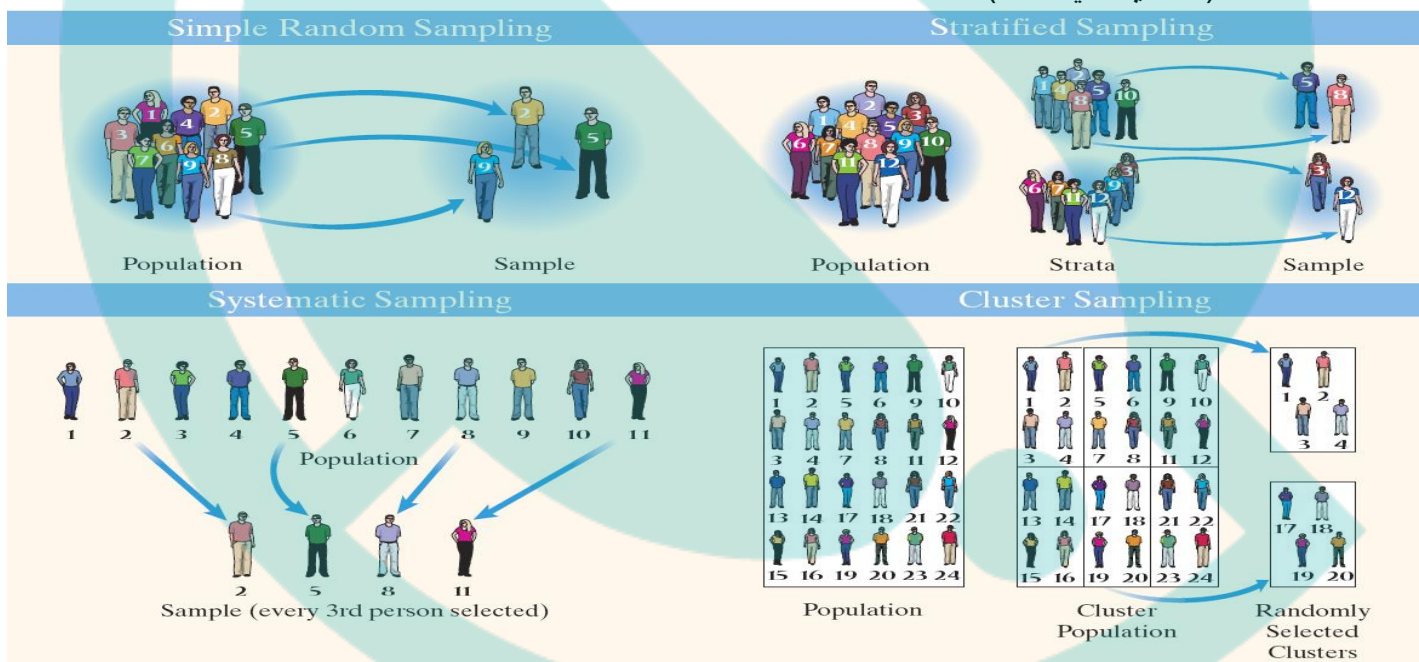
Step 5: The sample will consist of the following individuals: $[p, p + k, p + 2k, \dots, p + (n - 1)k]$

A cluster sample is obtained by selecting all individuals within a randomly selected collection or group of individuals.

بختار العينة هنا عن طريق تقسيم الـ Population لمجموعات وبختار اختيار عشوائي من المجموعات دا

A convenience sample is one in which the individuals in the sample are easily obtained.

دي عينة الراحة بختار العينة بتاعتي زي ما نا عايز ** اي دراسه بتختار النوع ده في اختيار العينة بشكل عام بتبقي النتائج بتاعتها غير موفقه (تحتوي علي خطأ) لذلك يجب اعاده النظر اليها بتشكك شديد



If the results of the sample are not representative of the population, then the sample has **bias**.

الانحياز في اختيار العينة بتاعتي لهدف ما

Three Sources of Bias

1. Sampling Bias
2. Nonresponse Bias
3. Response Bias

الانحياز في اختيار العينة

عدم الاستجابة في الاختيار

انحياز الشخص نفسه في اختيار العينة

Sampling bias means that the technique used to obtain the individuals to be in the sample

tend to favor one part of the population over another. الانحياز في اختيار العينة للحصول علي افراد تميل

لاختيار الاختيارات التي اريدها انا

Nonresponse bias exists when individuals selected to be in the sample who do not respond to the survey have different opinions from those who do. هنا بيتتم عدم الاستجابة لبعض الاختيارات من قبل بعض الأشخاص اللي اختارتهم في العينه بتاعتي

Response bias exists when the answers on a survey do not reflect the true feelings of the respondent.

هنا بيتتم الانحياز في الاختيار من قبل الشخص اللي عامل الاستبيان يعني الشخص بيختار اختيار غير اللي هو عايزه

Types of Response Bias

1. Interviewer error اخطاء الانتريفير
2. Misrepresented answers التعبير الخطا في الاجابه
3. Words used in survey question وجود كلمه في السؤال لها اكثر من معني
4. Order of the questions or words within the question ترتيب كلمات السؤال بشكل معين

Blinding refers to nondisclosure of the treatment an experimental unit is receiving.

A single-blind experiment is one in which the experimental unit (or subject) does **not know** which treatment he or she is receiving.

A double-blind experiment is one in which neither the experimental unit nor the researcher in contact with the experimental unit **knows** which treatment the experimental unit is receiving.

Steps in Conducting an Experiment

Step 1: Identify the problem to be solved. تحديد المشكله اللي عايز احلها

Step 2: Determine the factors that affect the response variable. تحديد العوامل اللي هتأثر علي المتغيرات.

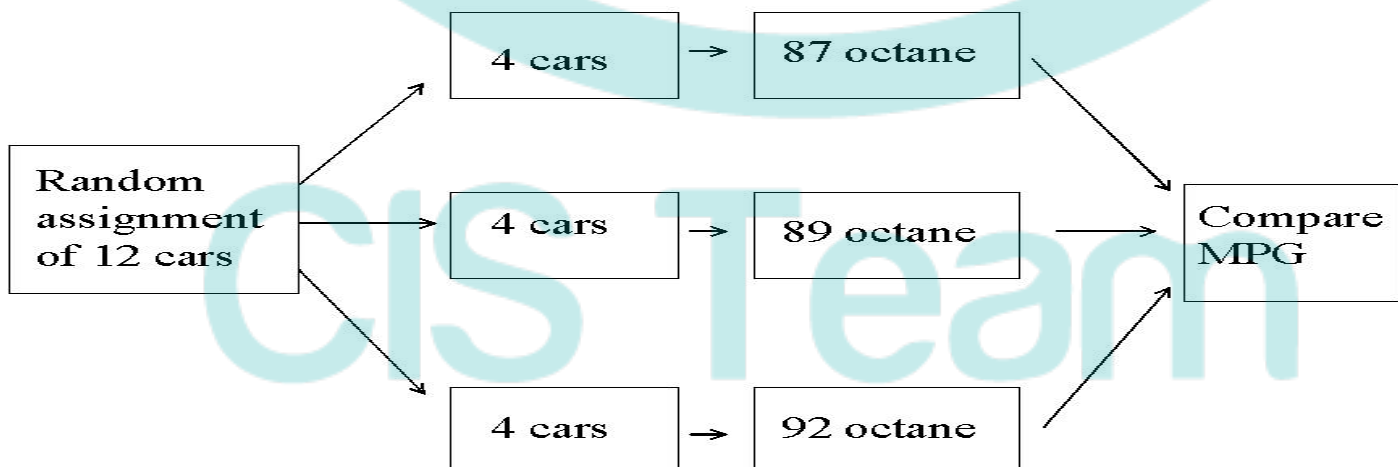
Step 3: Determine the number of experimental units. تحديد عدد وحدات التجربه

Step 4: Determine the level of the predictor variables تحديد مستوي المتغيرات اللي انا متوقعها

Step 5: Conduct the Experiment اجراء التجربه علي هذه المتغيرات

Step 6: Test the claim. اختبار النتائج عن طريق مقارنه بالنتائج اللي انا كنت متوقعها

Completely Randomized Design



EXAMPLE If we know from CS department data that the average age of all CS students is 76 in C# exam. Identify whether this value a parameter or statistic?

Parameter

عندي هنا كلمه داله وهي all

EXAMPLE In a sample of 200 CS departments, 50% of the students were 22 years old or younger. Identify whether this value a parameter or statistic?

Statistic

عندي هنا كلمه داله وهي sample

EXAMPLE In the following determine whether the given description correspond to an observation study or a designed experiment

a) The centers for disease control obtains current flu data by polling 3000 people this month

Observation study

b) Patients are given Lipitor to determine whether this drug has the effect of lowering cholesterol

Designed experiment

EXAMPLE

Researcher Elisabeth Kvaavik and others studied factors that affect the eating habits of adults in their mid-thirties. Classify each of the following variables considered in the study as qualitative or quantitative.

a) Nationality.

Qualitative

لان الجنسيه تشير لصفه

b) Number of children.

Quantitative

لان عدد الاطفال كميه اقدر اعدھا

c) Household income in the previous year.

Quantitative

d) Level of education.

Qualitative

e) Daily intake of whole grains (measured in grams per day).

Quantitative

EXAMPLE

Obtaining a Simple Random Sample

The 110th Congress of the United States had 435 members in the House of Representatives. Explain how to conduct a simple random sample of 5 members to attend a Presidential luncheon. Then obtain the sample.

Answer

Put the members in alphabetical order. Number the members from 1 - 435.

نضع الاعضاء حسب الترتيب الابجدي ثم نقوم بترقيمهم من 1 الي 435

EXAMPLE*Obtaining a Stratified Sample*

In 2008, the United States Senate had 49 Republicans, 49 Democrats, and 2 Independents. The president wants to have a luncheon with 4 Republicans, 4 Democrats and 1 other. Obtain a stratified sample in order to select members who will attend the luncheon.

Answer

- 1) Put the 49 Republicans Senates in alphabetical order. Number the members from 1 – 49 and select 4 Senates of them using a sample random sample of them
نضع الـ 49 اللي في الحزب الجمهوري في ترتيب ابجدي ثم نقوم بترقيمهم من 1 الي 49 ثم نختار منهم 4 عشوائي
- 2) Put the 49 Democrats Senates in alphabetical order. Number the members from 1 – 49 and select 4 Senates of them using a sample random sample of them
نضع الـ 49 اللي في الحزب الديموقراطي في ترتيب ابجدي ثم نقوم بترقيمهم من 1 الي 49 ثم نختار منهم 4 عشوائي
- 3) Put the 2 Independents Senates in alphabetical order. Number the members from 1 – 2 and select one of them using a sample random sample of them
نضع الـ 2 المستقلين في ترتيب ابجدي ثم نقوم بترقيمهم من 1 الي 2 ثم نختار منهم 1 عشوائي

EXAMPLE*A Matched-Pairs Design*

Xylitol has proven effective in preventing dental caries (cavities) when included in food or gum. A total of 75 Peruvian children were given milk with and without Xylitol and were asked to evaluate the taste of each. The researchers measured the children's ratings of the two types of milk.

- (a) What is the response variable in this experiment? **Rating**
التقييم , يعني طعم الحليب ببالماده ومن غير الماده
- (b) Think of some of the factors in the study. Which are controlled? Which factor is manipulated? **Age and gender of the children; Milk with and without Xylitol is the factor that was manipulated**
الحليب اللي عندي فيه الماده اللي هتقلل التسوس ولا لا لان الماده دي هي العامل اللي انا بتلاعب فيه
- (c) What are the treatments? How many treatments are there? **Milk with Xylitol and milk without xylitol; 2**
الحليب فيه الماده , والحليب من غير الماده / يعني حالتين

EXAMPLE

A student has to answer 10 question choosing at least 4 from each of part A to B if there are 6 question in part A and 7 in part B in how many ways can the student choose 10 questions?

Answer

الشخص هيختار 10 اسئله يجاوب عليهم بحيث يختار علي الاقل 4 من A و 4 من B ايضا ومسموح للشخص ان يختار اكثر من 4 ولكن تحت شرطين (1) الا يزيد العدد عن 10 (2) الا يتعدى الرقم الموجود في الجزء الذي يختار منه سواء كان 6 او 7

$$[({}^6C_4) ({}^7C_6)] + [({}^6C_5) ({}^7C_5)] + [({}^6C_6) ({}^7C_4)] = 245$$

EXAMPLE

In how many number are there between 99 and 1000 having 7 in the units place?

Answer

في الاعداد من 99 الي 1000 عندي كام رقم يحتوي في خانه ما وليكن الاحاد علي الرقم 7

From 99 to 200 = 10

From 201 to 300 = 10

From 301 to 400 = 10

From 401 to 500 = 10

From 501 to 600 = 10

From 601 to 700 = 10

From 701 to 800 = 10

From 801 to 900 = 10

From 901 to 1000 = 10

دا معناه عندي في الفتره من 99 الي 200 عندي 10 ارقام الاحاد بتاعهم فيه 7 وكذلك في باقي الارقام فالمجموع هيساوي مجموع نواتج كل فتره من الفترات اللي عندي دي

$$\text{Total} = 10 + 10 + 10 + 10 + 10 + 10 + 10 + 10 + 10 = 90$$

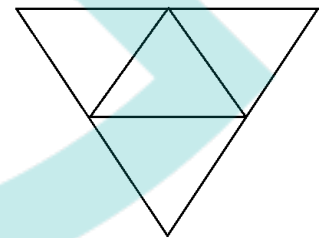
EXAMPLE

In how many ways can this diagram be colored subject to the following conditions

1) Each of the smaller triangles is to be painted with

One of 3 colors [red, blue and green]

2) Now 2 adjacent regions have the same color

**Answer**

ازاي اللون الشكل دا بالـ 3 اللون اللي عندي بحيث ان المثلث الاصغر هيتلون بلون ما من الـ 3 والمثلثات الباقية بنفس الـ 3 اللون بس بشرط كمان ان مفيش 2 جنب بعض يكون ليهم نفس اللون

$$3 * 2 * 2 * 2 = 24$$

الـ 3 دي احتمالات المثلث الاصغر وبعد كده كل مثلث هيبقي ليه احتمال لون من 2 عشان مفيش 2 يكونوا نفس اللون

Sample Mean

يرمز له بالرمز "X bar"

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Example 3.1

- The average fuel efficiencies, in miles per gallon, of cars sold in the United States in the years 1999 to 2003 were 28.2, 28.3, 28.4, 28.5, and 29.0. Find the sample mean of this set of data.

- Solution** The sample mean \bar{x} is the average of the five data values. Thus,

$$\bar{x} = \frac{28.2 + 28.3 + 28.4 + 28.5 + 29.0}{5} = \frac{142.4}{5} = 28.48$$

بجمع كل القيم اللي عندي واقسمها علي عددها

Example 3.2

- The winning scores in the U.S. Masters Golf Tournament in the years from 1981 to 1990 were as follows: 280, 284, 280, 277, 282, 279, 285, 281, 283, 278

Find the sample mean of these winning scores.

- Solution**

لما بيبقي عندي اكثر من 5 قيم بحل بطريقه اسهل شويه هختار قيمه من القيم اللي عندي واطرح منها كل القيم اللي عندي بما فيهم القيمه اللي انا مختارها واكتب ناتج الطرح بنفس الاشاره بتاعته (2- , 3 , 1 , 5 , 1- , 2 , 3- , 0 , 4 , 0) وارجع اجمع النواتج دي واقسمها علي عددها واسميها اسم ما

$$\bar{y} = \frac{0 + 4 + 0 - 3 + 2 - 1 + 5 + 1 + 3 - 2}{10} = \frac{9}{10}$$

هاخذ الناتج اللي طلع عندي دا واجمعه علي القيمه اللي كن تبطرح منها والناتج هو (X bar)

Example 3.3

- The number of suits sold daily by a women's boutique for the past 6 days has been arranged in the following frequency table. What is the sample mean?

Value	Frequency
3	2
4	1
5	3

- Solution**

هضرب كل قيمه في ال Frequency بتاعتها واجعها كلها واقسمها علي مجموع ال Frequency

$$\bar{x} = \frac{3 \times 2 + 4 \times 1 + 5 \times 3}{6} = \frac{25}{6}$$

Deviation

The differences between each of the data values and the sample mean

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x} = n\bar{x} - n\bar{x} = 0$$

Prove that

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x} = n\bar{x} - n\bar{x} = 0$$

Solution:

$$\therefore \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\therefore \sum_{i=1}^n x_i = x_i \sum_{i=1}^n 1$$

$$\therefore \sum_{i=1}^n \bar{x} = \bar{x} \sum_{i=1}^n 1$$

الـ \bar{x} مفيش تأثير عليها فهطلعها بره الـ summation

$$\therefore \sum_{i=1}^n 1 = n$$

$$\therefore \bar{x} \sum_{i=1}^n 1 = n\bar{x}$$

$$n\bar{x} - n\bar{x} = 0$$

Example 3.5 In Example 3.1: The average fuel efficiencies, in miles per gallon, of cars sold in the United States in the years 1999 to 2003 were 28.2, 28.3, 28.4, 28.5, 29.0 find the deviation?

Solution

الـ ($\bar{x} = 28.48$) زي ماطلعتها في المثال الاول وبعد كذا هأخذ كل قيمه عندي اطرحها من 28.48

$$x_1 - \bar{x} = 28.2 - 28.48 = -0.28$$

$$x_2 - \bar{x} = 28.3 - 28.48 = -0.18$$

$$x_3 - \bar{x} = 28.4 - 28.48 = -0.08$$

$$x_4 - \bar{x} = 28.5 - 28.48 = 0.02$$

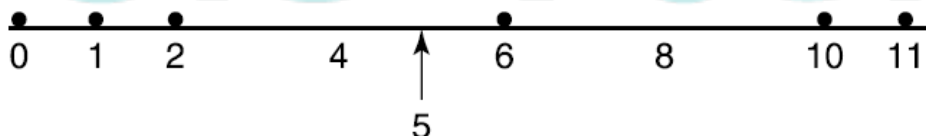
$$x_5 - \bar{x} = 29.0 - 28.48 = 0.52$$

As a check, we note that the sum of the deviations is

$$-0.28 - 0.18 - 0.08 + 0.02 + 0.52 = 0$$

Center of Gravity The sample mean is a balancing point

- For example The center of gravity of 0, 1, 2, 6, 10, 11 is $(0 + 1 + 2 + 6 + 10 + 11)/6 = 30/6 = 5$



Sample Median Order the data values from smallest to largest.

القيم اللي عندي بتبقى Odd or Even

(1) لو القيمه دي Odd فالـ Sample Median هيساوي $n/2$ دا بعد ترتيب القيم اللي عندي تصاعديا طبعاً

(2) لو القيمه دي Even فالـ Sample Median هيساوي $[(n/2) + ((n/2) + 1)]/2$ يعني مجموع القيمتان علي 2

Example 3.6 The following data represent the number of weeks it took seven individuals to obtain their driver's licenses. Find the sample median. 2, 110, 5, 7, 6, 7, 3

Solution

2, 3, 5, 6, 7, 7, 110

هرتب القيم اللي عندي

القيم اللي عندي عددها 7 يعني فردي فهيبقي $7/2 = 3.5$ هقربها لاقرّب عدد صحيح يليها = 4 يبقى الناتج = القيمه 4 يعني = 6

Sample Percentiles

The sample $100p$ percentile is that data value having the property that at least $100p$ percent of the data are less than or equal to it and at least $100(1 - p)$ percent of the data values are greater than or equal to it.

- **Note** that the sample median is the sample 50th percentile. $p = 0.50$

Example 3.9 Which data value is the sample 90th percentile when the sample size is

(a) 8

(b) 16

(c) 100?

Solution (a) $0.9 \times 8 = 7.2 = 8$

بضرب قيمه الـ sample percentile * sample size والناتج لو كان عشري بقربه لاقرّب عدد صحيح يليه

(b) $0.9 \times 16 = 14.4 = 15$

(c) $0.9 \times 100 = 90$

Quartile

The sample 25th percentile is called the first quartile (q_1)

The sample 50th percentile is called the median or the second quartile (q_2)

The sample 75th percentile is called the third quartile (q_3)

- **Quartile range** = $q_3 - q_1$

Example 3.11 Find the sample quartiles for the following 18 data values, which represent the ordered values of a sample of scores from a league bowling tournament 122, 126, 133, 140, 145, 145, 149, 150, 157, 162, 166, 175, 177, 177, 183, 188, 199, 212

- **Solution**

الاول هرتب القيم اللي عندي هو عندي في المثال دي عطهالي مرتبه جاهزه فحسب (q_1, q_2, q_3) علي طول

◦ $q_1 = 0.25 \times 18 = 4.5 = 5$ which is 145.

◦ $q_2 = 0.50 \times 18 = 9$ which is $(157 + 162) / 2 = 159.5$

◦ $q_3 = 0.75 \times 18 = 13.5 = 14$ which is 177.

Exercise Seventy-five values are arranged in increasing order. How would you determine the sample?

(a) 80th percentile (b) 60th percentile (c) Quartile range of this data set?

Solution

- a) $= 0.8 * 75 = 60$
- b) $= 0.6 * 75 = 45$
- c) $q1 = 0.25 * 75 = 18.7 = 19$
- $q3 = 0.75 * 75 = 56.2 = 57$
- Quartile range $= 57 - 19 = 38$

Sample mode the data value that occurs most frequently in the data set

Example 3.12 The following are the sizes of the last 8 dresses sold at a women's boutique 8, 10, 6, 4, 10, 12, 14, 10 What is the sample mode?

Solution

The sample mode is 10

Sample mode هنا المقصود القيمه الاكثر تكرارا ولو كل القيم عدد مرات تكرارها متساوي يبقي لا يوجد

Sample Variance and Sample Standard Deviation

Definition The quantity s , defined by

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

is called the sample standard deviation.

Definition

The sample variance, call it s^2 , of the data set x_1, \dots, x_n having sample mean $\bar{x} = (\sum_{i=1}^n x_i) / n$ is defined by

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Example 3.15 Find the sample variance of data set A. **A: 1, 2, 5, 6, 6**

Solution

It is determined as follows:

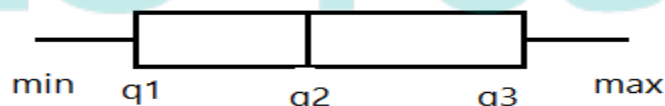
x_i	1	2	5	6	6
\bar{x}	4	4	4	4	4
$x_i - \bar{x}$	-3	-2	1	2	2
$(x_i - \bar{x})^2$	9	4	1	4	4

Hence, for data set A,

$$s^2 = \frac{9 + 4 + 1 + 4 + 4}{4} = 5.5$$

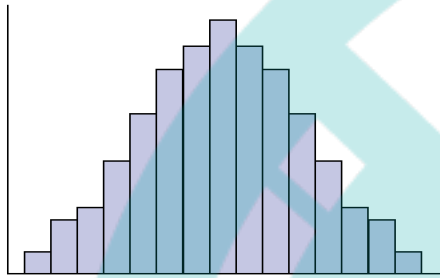
A Box Plot is often used to plot some of the summarizing statistics of a data set.

هنا بجيب 5 حاجات الاول الـ (Max, Min, q_1 , q_2 , q_3) وبعد كده هكتبهم عالشكل كما هو موضح

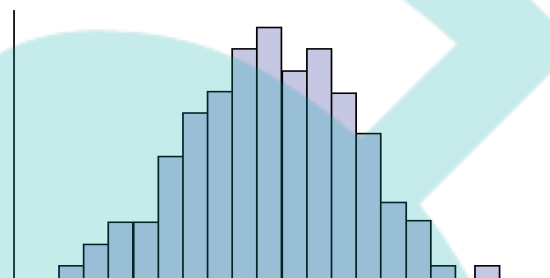


Normal Data Sets and the Empirical Rule

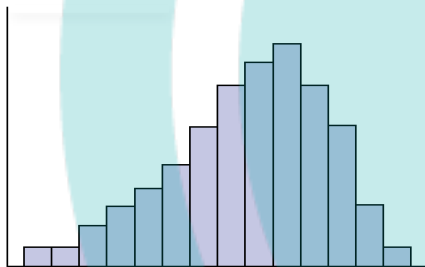
الـ Normal Data بتتمثل بـ histogram وفي الـ histogram بيكون الـ [Main = Mode = Middle]



Normal

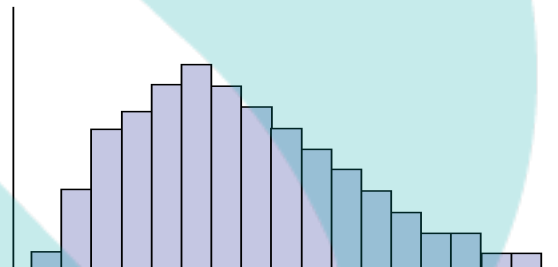


Approximately Normal



Skewed to the left

Mean > Median



Skewed to the right

Median > Mean

Empirical Rule

If a data set is approximately normal with sample mean \bar{x} and sample standard deviation s , then the following are true.

1. Approximately 68 percent of the observations lie within

$$\bar{x} \pm s$$

2. Approximately 95 percent of the observations lie within

$$\bar{x} \pm 2s$$

3. Approximately 99.7 percent of the observations lie within

$$\bar{x} \pm 3s$$

68 % دي يعني القيم المحصوره في الفتره $[x^- - s ; x^- + s]$ ممكن يكون فيه اختلاف

في القيمه 68 دي فهي تتراوح بين 67 و 69

Example 3.20

The scores of 25 students on a history examination are listed on the following stem-and-leaf plot.

9	0, 0, 4
8	3, 4, 4, 6, 6, 9
7	0, 0, 3, 5, 5, 8, 9
6	2, 2, 4, 5, 7
5	0, 3, 5, 8

By standing this figure on its side, we can see that
The corresponding histogram is approximately normal.
Use it to assess the empirical rule.

Solution

$$\bar{X} = 73.68$$

&

$$S = 12.80$$

$$1) \bar{X} - S = 73.68 - 12.80 = 60.88$$

$$\bar{X} + s = 73.68 + 12.80 = 86.48$$

لما هشوف القيم المحصوره بين الـ 60.88 والـ 86.48 هلاقي عندي 17 قيمه لما هاجي بحسب نسبتهم من 25 هيبقي
(17/25) * 100 = 68 وكده اتأكدت ان نص القاعده صحيح

$$2) \bar{X} - 2S = 73.68 - 12.80 = 48.08$$

$$\bar{X} + 2s = 73.68 + 2 * 12.80 = 99.28$$

$$3) \bar{X} - 3S = 73.68 - 12.80 = 35.28$$

$$\bar{X} + 3s = 73.68 + 3 * 12.80 = 112.08$$

الـ Stem _ and leaf plot عشان اعمله اول حاجه برتب الـ data اللي عندي تصاعدي بعد كذا كل رقم عندي بطلع منه حاجتين الـ Steam وده بيمثل خانه العشرات والـ Leaf وده بيمثل خانه الاحاد بيبقي بالشكل دا

Stem	9	0, 0, 4	Leaf
	8	3, 4, 4, 6, 6, 9	
	7	0, 0, 3, 5, 5, 8, 9	
	6	2, 2, 4, 5, 7	
	5	0, 3, 5, 8	

- These data represent grades of 25 students:

9	0, 0, 4
8	3, 4, 4, 6, 9
7	0, 0, 3, 5, 5, 8, 9
6	2, 2, 4, 5, 7
5	0, 3, 5, 8

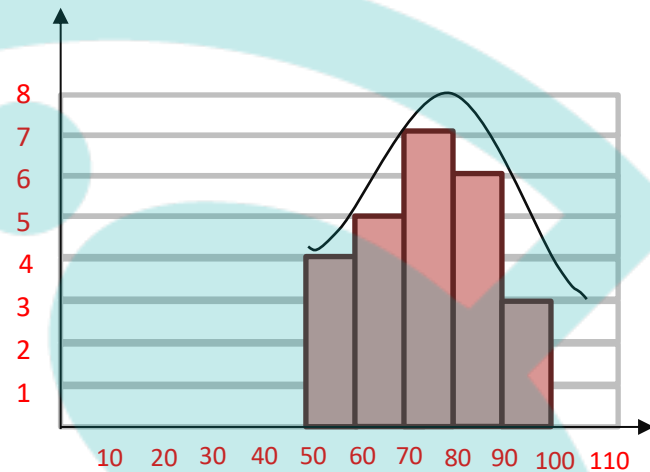
- Plot histogram those data.

طالب فالسؤال إني أرسم Histogram لدرجات الطلاب, المفروض إن الـ Histogram عبارته عن علاقه بين حاجتين, طيب احنا فالسؤال معندناش بيانات غير درجات الطلاب, فعشان نرسم الـ Histogram هنضطر نجيب الـ Frequency لدرجات الطلاب عشان يكون فيه حاجتين نعمل ما بينهم علاقه لتمثيلهم ف صورته.

Histogram

Solution:

Grade	Frequency
50 → 59	4
60 → 69	5
70 → 79	7
80 → 89	6
90 → 99	3



Bimodal A data set whose histogram has two local peaks is said to be bimodal.

Sample Correlation Coefficient

كلما زادت قيمة X زادت قيمة Y وكلما قلت قيمة X قلت قيمة Y
 بكتب القيم كالتالي $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ وبمثل القيم دي بـ Scatter diagram

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

لو الـ $(r > 0)$ يبقى الـ Sample data عندي Positively يعني طردي
 لو الـ $(r < 0)$ يبقى الـ Sample data عندي Negatively يعني عكسي
 الـ r بتبقى دايما محصوره بين $(-1, 1)$

A value of $|r| = 1$ means that there is a **perfect linear relation**.

A value of $|r| = 0.8$ means that the linear relation is **relatively strong**.

A value of $|r| = 0.3$ means that the linear relation is **relatively weak**.

Example 3.22

This gives the new set of data pairs:

Find r ??

	i		
	1	2	3
x_i	4.3	1.9	0
y_i	0	0.9	2.6

Solution

$$\bar{x} = \frac{4.3 + 1.9 + 0}{3} = 2.0667 \quad \bar{y} = \frac{0 + 0.9 + 2.6}{3} = 1.1667$$

$$\sum_{i=1}^3 x_i y_i = (1.9)(0.9) = 1.71 \quad \sum_{i=1}^3 x_i^2 = (4.3)^2 + (1.9)^2 = 22.10$$

$$\sum_{i=1}^3 y_i^2 = (0.9)^2 + (2.6)^2 = 7.57$$

$$r = \frac{1.71 - 3(2.0667)(1.1667)}{\sqrt{[22.10 - 3(2.0667)^2][7.57 - 3(1.1667)^2]}} = -0.97$$

That is very negative correlation

Qultier Value

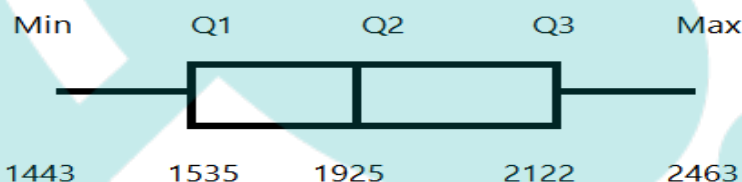
Those value that lie outside the interval $[Q1 - 1.5 * IQR, Q3 + 1.5 * IQR]$: $IQR = Q3 - Q1$

Example For the data [2009,1915,2011,2463,2180,1925,1482,2122,1542,1443,1535] would a value 300 be considered an outlier?

Solution

هنا هجيب 5 حاجات الـ [Min, Max, Q1, Q2, Q3]

Min = 1443 Max = 2463 Q1 = 1535 Q2 = 1925 Q3 = 2122



$$IQR = Q3 - Q1 = 2100 - 1535 = 587$$

$$\text{Outlier value } [Q1 - 1.5 * IQR, Q3 + 1.5 * IQR] = [654.5, 3002.5]$$

300 is not find in this range so, this is not Qultier value

Confusion matrix / Crosstabs

- Confusion Matrix (An Error Matrix): A specific table layout that allows visualization of the performance of an algorithm
- It is a special kind of contingency table, with two dimensions ('actual' and 'predicted'), and identical sets of "classes" in both dimensions

Actual class / predicted class	Disease = Yes	Disease = No	Total
Disease = Yes	TP	FN	TP + FN
Disease = No	FP	TN	FP + TN
Total	TP + FP	FN + TN	TP + FN + FP + TN

- TP = True Positives
- TN = True Negatives

- FP = False Positives
- FN = False Negatives

- **Precision** = $TP / (TP + FP)$
- **Recall / Sensitivity** = $TP / (TP + FN)$
- **Specificity** = $TN / (TN + FP)$
- **Accuracy** = $(TP + TN) / (TP + TN + FP + FN)$
- **F-Score** = $(2 * Precision * Recall) / (Precision + Recall)$

- Below is a table of 40 cotinine level of smokers:

0	1	1	3	17	32	35	44	48	86
87	103	112	121	123	130	131	149	164	167
173	173	198	208	210	222	227	234	245	250
253	265	266	277	284	289	290	313	477	491

- Find of the percentile corresponding to the cotinine level of 149?

عندي بيانات مكونه من 40 قيمه و عايز أحسب النسبة المئوية ل Level 10 و هي 149.

$$\text{Percentile} = \frac{\text{مكان الرقم في الجدول}}{\text{عدد القيم}} \times 100$$

Solution:

$$\text{Percentile} = \frac{18}{40} \times 100 = 45$$

- Find the level that correspond t p_{20} (the 20th percentile)

هنا هو اداني percentile وعايز Level

Solution:

$$20 = \frac{X}{40} * 100 \rightarrow \frac{X}{40} = 0.2 \rightarrow X = 8 \quad \text{Then } p_{20} = 44$$

ده ال Level
رقم 8 و قيمته ف الجدول = 44

A sample of 36 male patients yields the following data:

7	1 , 2 , 4 , 5
6	0 , 1 , 2 , 2 , 3 , 4 , 5 , 7
5	0 , 1 , 2 , 3 , 3 , 4 , 4 , 4 , 5 , 6 , 7 , 8 , 9
4	1 , 2 , 2 , 3 , 4 , 5 , 7 , 8 , 9
3	7 , 9

1) Determine \bar{x} (mean) and S (standard deviation)?

Solution: ● $\bar{x} = \frac{2025}{36} = 56.25$ ● $S = \sqrt{s^2}$ $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$

$$S = \sqrt{\frac{(71-56.25)^2 + (72-56.25)^2 + (74-56.25)^2 + (75-56.25)^2 + (60-56.25)^2 + (61-56.25)^2 + (62-56.25)^2 + \dots}{35}}$$

$$= 16.766$$

2) From the shape of the steam_and_leaf plot, what percentage of data value would you expected to be?

في السؤال ده عايز نسبه القيمه المتوقعه للبيانات بين الفترات التاليه:

- a) Between $\bar{x} - s$ and $\bar{x} + s$? 68%
 b) Between $\bar{x} - 2s$ and $\bar{x} + 2s$? 95%
 c) Between $\bar{x} - 3s$ and $\bar{x} + 3s$? 99.7%

3) Find the actual percentage for the intervals given into: هنا عايز النسب الفعلية للقيم ف الفترات دي**a) Between $\bar{x} - s$ and $\bar{x} + s$?**

$$\bar{x} - s = 56.25 - (16.766) = 39.48 \quad \&\& \quad \bar{x} + s = 56.25 + (16.766) = 73.1$$

* الـ 39.48 و 73.1 دول قيم مش النسبه إللي طالباها في السؤال عشان أجيب النسبه هشوف عدد القيم إللي بتقع بين الـ 39.48 و 73.1 و أقسم عدد القيم إللي هطلعها على عدد البيانات الكليه و أضرب في 100.

$$\text{النسبه الفعلية} = 100 * \frac{32}{36} = 88.89\%$$

b) Between $\bar{x} - 2s$ and $\bar{x} + 2s$?

$$\bar{x} - 2s = 56.25 - 2(16.766) = 22.718 \quad \&\& \quad \bar{x} + 2s = 56.25 + 2(16.766) = 89.782$$

$$\text{النسبه الفعلية} = 100 * \frac{36}{36} = 100\%$$

c) Between $\bar{x} - 3s$ and $\bar{x} + 3s$?

طبعا دي من غير مالحسبها ان عارف انها بتبقي في نطاق اكبر من $\bar{x} - 2s$ يبقى النسبه الفعلية بتاعتها = 100%

Probability Theory: is the study of the mathematical rules that govern random events.

Random Event: is an event in which we don't know the outcome without observing it.

Probability tells us: what we can say about such events , given our assumptions about the possible outcomes.

Examples:**1) Tossing 2 dice the sample space consist of the 36 point**

$$S = \Omega = \{ (i, j) : i, j = 1, 2, 3, 4, 5, 6 \}$$

حيث ا هي ناتج الرمي الاول والـ ز هي ناتج الرمي الثانيه

$$\text{Number of sample space} = |\Omega| = 36$$

2) If the experiment consists of measuring (in hours) the life time of a transistor, then the sample space consists of all non negative real numbers

$$S = \Omega = \{ x : 0 \leq x < \infty \}$$

حيث الـ x عمر الترانزستور والـ ∞ بتمثل الفتره اللتي هيتلف فيها

Event: a subset of sample space.

--Roll with a 6 – sided die.

حجر نرد له 6 اوجه

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

دي احتمالات ظهور عدد زوجي $E = \{2, 4, 6\}$ && دي معناها احتمال ظهور الرقم 3 $E = \{3\}$

De Morgan's laws (generalized)

$$(\bigcup_{i=1}^n E_i)^c = \bigcap_{i=1}^n E_i^c \quad \&\& \quad (\bigcap_{i=1}^n E_i)^c = \bigcup_{i=1}^n E_i^c$$
 Complement الـ C دي ترمز الي

Probability P (E): a measure of the likelihood of an event E occurring.

**if E_i is the (i^{th}) experimental result and $P(E_i)$ is considered it is probability of occurring hence.

$$0 \leq P(E_i) = \frac{\text{Number of ways that an event } E_i \text{ can occur}}{\text{Number of sample space}} \leq 1 \quad \text{for all (i).}$$

Axioms of probability

- $0 \leq P(E_i) \leq 1$
- $P(E_1) + P(E_2) + \dots + P(E_n) = 1$
- $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$
- $p(A) = 0$
 $\{\} = \phi$
- $p(A) = 1$
- $p(A) = 60\%$

مسلّمات الاحتمالات

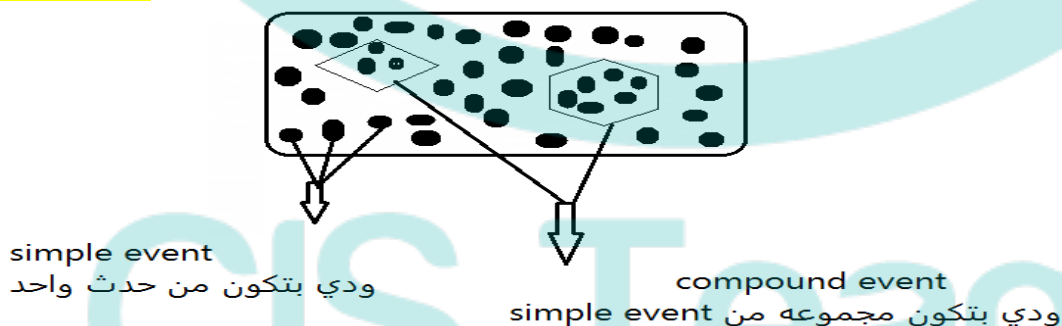
- اي حدث محصور بين 0 و 1
مجموع الاحتمالات = 1
اتحاد مجموع الاحداث = مجموع هذه الاحداث
دي معناها ان الحدث دا مستحيل حدوثه
دا معناه ان مفيش عندي outcome
دي معناها ان الحث دا مؤكد حدوثه
دي معناها ان احتمال تحقق هذا الحث = 60%

Flip twice: probability of flipping (H, T) = $(1/2) * (1/2) = (1/4)$

[H: يعني صوره && T: يعني كتابه]

Flip 3 times: probability of (H, H, and H) = $(1/2) * (1/2) * (1/2) = (1/8)$

Venn diagram: A graphical representation of sample space and curtain event.



■ $P(A) + P(A^c) = 1$

■ $A \cap B$ = Intersection of A and B

$A \cap B = \phi$ if (A and B) are disjoint

يعني الحدثان متنافيان مفيش بينهم تشابه

Addition rule:

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B) = P(A) + P(B)$: (A, B) are disjoint

Examples:

- The probability of rolling an event number and a number greater than 3 A = rolling an even number B = rolling a number greater than 3.

Solution:

$$A = \{2, 4, 6\} \quad \&\& \quad B = \{4, 5, 6\} \quad \&\& \quad A \cap B = \{4, 6\} \quad \&\& \quad A \cup B = \{2, 4, 5, 6\}$$

$$P(A \cap B) = 2/6 = 1/3. \quad \&\& \quad P(A \cup B) = 4/6 = 2/3.$$

- If 3 balls are “randomly drawn” from a bowl containing 6 white and 5 black balls, what is the probability that one of the balls is white and the other two black?

Solution:

انا عندي 11 كره هختار منهم 3 هاحسب الاول عدد الاحتمالات الكلي $990 = 9 * 10 * 11$
 عندي 3 طرق هاختار بيهم الكرات بحيث اختار كره بيضاء و 2 سوداء

1. $6 * 5 * 4 = 120$
2. $5 * 6 * 4 = 120$
3. $5 * 4 * 6 = 120$

اول كره بيضاء ثم 2 سوداء

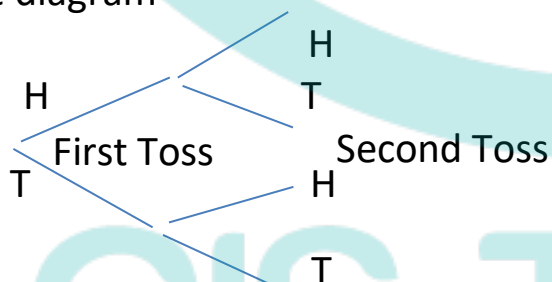
اول كره سوداء ثم بيضاء ثم سوداء

اول 2 سوداء ثم كره بيضاء

بعد كذا هجيب الاحتمال من خلال جمع احتمال الـ 3 طرق واقسمه علي عدد الاحتمالات الكلي

$$\text{Probability} = \frac{120 + 120 + 120}{990} = \frac{4}{11}$$

- **An experiment** : is a process that leads to well-defined results called outcomes
- **An out comes** : is the result of single trial of a probability experiment
- **Probability** : the chance that event will occur
 - Equally likely : يعني فرص الظهور كلها متساوية
 - Tree diagram



← First Toss يعني الرمي الاول ويكون الـ Outcomes منها (H, T)

- كل event ينتج منه (H, T)
- عدد الاحتمالات 2^n حيث الـ n بتساوي عدد الـ event

- Complement of an event A (A')

All possible events that aren't part of event A

كل الـ event الممكنة الى غير موجوده في A

- Joint probability

The probability of an occurrence of two or more event

الاحتمالات المشتركة بين احتمالان أو اكثر

Number of all outcome satisfying (A and B)

$$P(A \text{ and } B) = \frac{\text{Number of all outcome satisfying (A and B)}}{\text{The total number of sample space}}$$

The total number of sample space

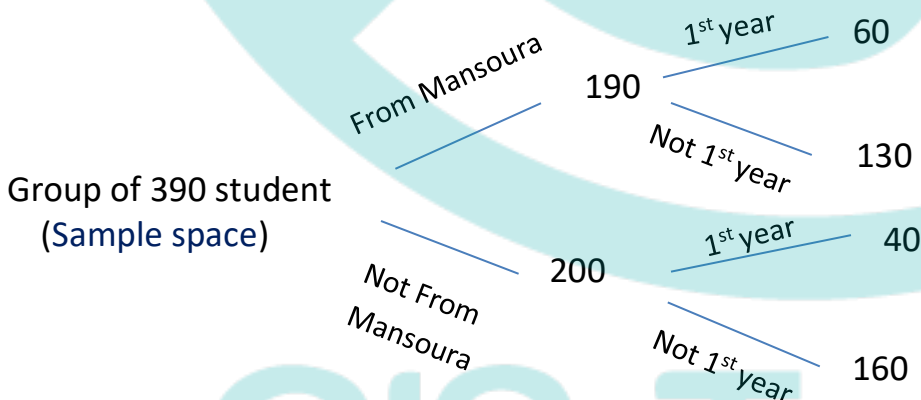
- Contingency tables

Event	C	D	Total
A	$P(A \text{ and } C)$	$P(A \text{ and } D)$	$P(A)$
B	$P(B \text{ and } C)$	$P(B \text{ and } D)$	$P(B)$
Total	$P(C)$	$P(D)$	Sample space

	1 st year	Not 1 st year	Total
Mansoura	60	130	190
Not from Mansoura	40	160	200
Total	100	290	390

Sample Space

- من خلال الـ Contingency tables اعمل الـ Decision Tree



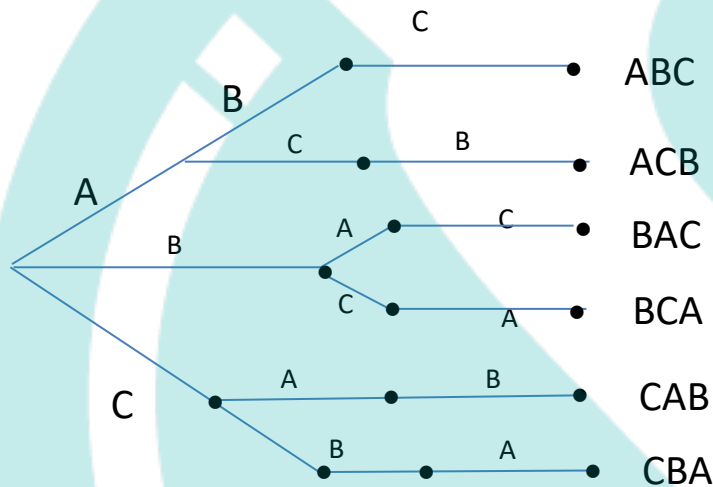
EX 1:

	1 st year	Not 1 st year	Total
Mansoura	200	50	250
Not from Mansoura	150	100	250
Total	350	150	500

$$P(A) = 250/500 = \frac{1}{2}, \quad P(C) = 350/500 = \frac{7}{10}$$

$$P(B) = 250/500 = \frac{1}{2}, \quad P(D) = 150/500 = \frac{3}{10}$$

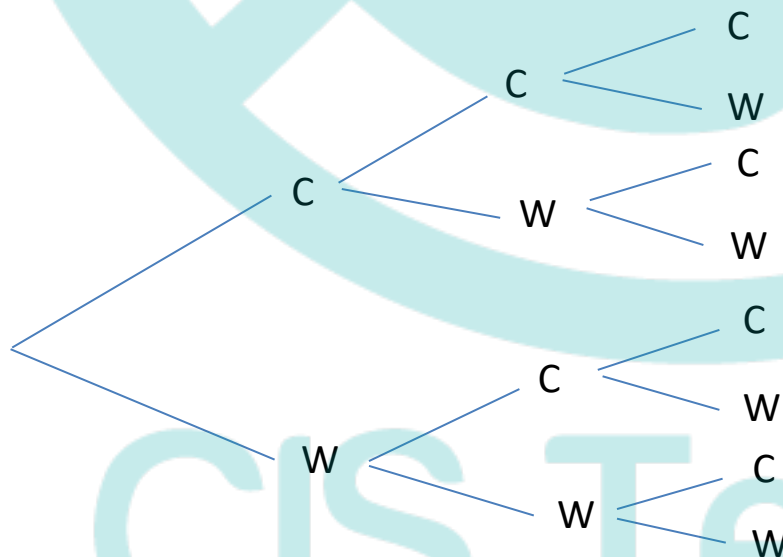
- $P(A \text{ and } C) = 200/500 = \frac{2}{5}$ && $P(A \text{ and } D) = 50/500 = \frac{1}{10}$
- $P(B \text{ and } C) = 150/500 = \frac{3}{10}$ && $P(B \text{ and } D) = 100/500 = \frac{1}{5}$
- ❖ Order three different English letter A,B and C



- نفس فكرة مبدأ العد (Counting)

EX: for a question there is only two answer either correct (C), wrong (W)

1) What is the probability of getting at least one correct answer when asking 3 questions ??



- هنا عايز يكون في 1 صح (C) على الأقل يعني مش يكون الـ 3 غلط (W) $P = 7/8$

2) What is the probability that student is answering at least 2 question correctly

- هنا يعني يجاوب على الأقل 2 صح يعني ممكن 2 أو 3 $P = 4/8$

3) What is the probability that student is answering at most 2 question correctly

- هنا يعني على الأكثر 2 صح يعني ممكن 2 أو 1 أو صفر بمعنى آخر مفيش (CCC)

$$P = 7/8$$

4) Answering at least 1 correct and 1 wrong

- هنا يعني لازم يكون على الأقل 1 صح و 1 غلط يعني مفيش (CCC) ولا في (WWW)

$$P = 6/8$$

5) Exactly one correct and one wrong

- هنا حدث مستحيل لان عايز 1 فقط صح و واحد فقط غلط وانا لازم يبقى عندي 3 كذا هما 2 فقط

$$P = 0$$

Conditional probability: is used to determine the probability of one given the occurrence of another related event

$$P(A/B) = \frac{\text{Number of outcomes in } (A \cap B)}{\text{Number of outcomes in } B} = \frac{P(A \cap B)}{P(B)} : P(B) > 0$$

❖ Conditional probability of (A) given (C) is $\Rightarrow P(A/C) = \frac{P(A \cap C)}{P(C)}$

❖ Multiplication rule $\Rightarrow P(A \cap C) = P(A/C) * P(C)$

EX: تابع مثال 1 اللي في صفحه 23

1) Find P (Student in the first year / student from Mansoura city)

هنا عايز احتمال ان يكون الطالب في سنه اولي بمعلوميه انه من المنصوره

$$= \frac{(\text{Student in the first year and from Mansoura})}{(\text{student from Mansoura})} = \frac{(\frac{2}{5})}{(\frac{1}{2})} = \frac{4}{5}$$

2) P (Student not from Mansoura / Student in the first year)

هنا عايز احتمال ان الطالب دا يكون مش من المنصوره بمعلوميه انه في سنه اولي

$$= \frac{(\frac{3}{10})}{(\frac{7}{10})} = \frac{3}{7}$$

EX: If a coin was throwing 3 times, what is the probability of A given B where:

A (2 or more tosses resulted in Heads)

B (the 1st toss resulted in Heads)

Solution:

$\Omega = (HHH, HHT, HTH, HTT, THH, THT, TTH, TTT)$

$$P(A) = \frac{4}{8}$$

$$P(B) = \frac{4}{8}$$

دا احتمال ان يكون اكثر من 2 H

دا احتمال ان تكون الاولى H

$$P(A \text{ and } B) = \frac{3}{8}$$

دا احتمالان يكون عندي 2 (H) بشرط الاول (H) فالـ (THH) مش هتبقى معنا لانها فقدت شرط ان الاول يكون (H)

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\left(\frac{3}{8}\right)}{\left(\frac{4}{8}\right)} = \frac{3}{4}$$

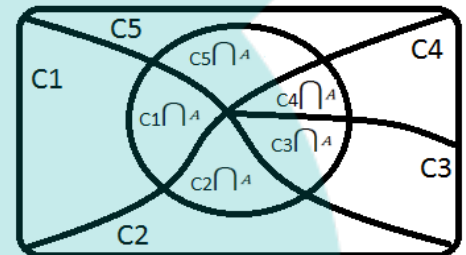
Total Probability and Bayes rule

The law of total probability suppose (c_1, c_2, \dots, c_m) are disjoint

Events such that $\Omega = c_1 \cup c_2 \cup \dots \cup c_m$

The probability of an arbitrary event A can be expressed as:

$$P(A) = P(A \cap c_1) + P(A \cap c_2) + \dots + P(A \cap c_m)$$



ϕ = تقاطع الاحداث

&&

اتحاد الاحداث اللي عندي = فضاء العينه

$$P(A) = P(A|c_1).P(c_1) + P(A|c_2).P(c_2) + \dots + P(A|c_m).P(c_m) : P(A \cap c) = P(A|c).P(c)$$

Special case:

$C_1: B$ && $C_2: B^C$ (مكملة الـ B)

$$P(A) = P(A|B).P(B) + P(A|B^C).P(B^C)$$

لو الـ (M=2) ودي بتكون ابسط حاله

Bayes Rule

Bayes: Theorems enable you. Knowing just a little more than the probability of (A) given (B) to find the probability of (B) given (A)

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

هطبق قاعده الـ Multiplication Rule علي البسط وقاعده الـ Total probability علي المقام

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

دا لو كان عندي حدثين بس

تعميم النظرية

$$P(B|A) = \frac{P(A|C_1).P(C_1)}{P(A|C_1).P(C_1) + \dots + P(A|C_I).P(C_I)}$$

Example 3.20:

Fifty-two percent of the students at a certain college are females. Five percent of the students in this college are majoring in computer science. Two percent of the students are women majoring in computer science. If a student is selected at random, find the conditional probability that

(a) The student is female given that the student is majoring in computer science

(b) This student is majoring in computer science given that the student is female

Solution:

a) هنا بيقولي كليه نسبه البنات فيها = 52% و 5% من الي في الكليه متخصصين في قسم علوم الحاسب و 2% من البنات متخصصين في قسم علوم الحاسب ايضا

Student is female (F) && majoring in computer science (C)

$$P(F) = 0.25 \quad \&\& \quad P(C) = 0.05 \quad \&\& \quad P(F \cap C) = 0.02$$

$$P(F|C) = P(F \cap C) / P(C) = (0.02) / (0.05) = 0.40$$

$$b) P(C|F) = \frac{P(C \cap F)}{P(F)} = \frac{0.02}{0.25} = 0.08$$

Example

A high school student is anxiously waiting to receive mail telling her whether she has been accepted to a certain college. She estimates that the conditional probabilities of receiving notification on each day of next week, given that she is accepted and that she is rejected, are as follows:

Day	P(Mail accepted(A))	P(Mail rejected(R))
Monday (M)	0.15	0.05
Tuesday (T)	0.20	0.10
Wednesday (W)	0.25	0.10
Thursday (Th)	0.15	0.15
Friday (F)	0.10	0.20

➤ She estimates that her probability of being accepted is 0.6.

(a) What is the probability that she receives mail on Monday?

(b) What is the conditional probability that she received mail on Tuesday given that she does not receive mail on Monday?

(c) If there is no mail through Wednesday, what is the conditional probability that she will be accepted?

(d) What is the conditional probability that she will be accepted if mail comes on Thursday?

(e) What is the conditional probability that she will be accepted if no mail arrives that week?

Solution:

طالبه في الثانويه تنتظر رساله لتخبرها ما اذا كان تم قبولها في كليه ما او لا واحتمال ارسال هذه الرساله سيكون في يوم من ايام الاسبوع المقبل بمعلوماته ما اذا تم قبولها او لا

$$P(\text{accepted}) = P(A) = 0.6 \quad \&\& \quad P(\text{rejected}) = P(R) = 1 - 0.6 = 0.4$$

a) عايز احتمال استلام الميل يوم الاثنين

$$P(M) = P(M|A) \cdot P(A) + P(M|R) \cdot P(R) = (0.15 \cdot 0.6) + (0.05 \cdot 0.4) = 0.11$$

b) عايز احتمال استلام الميل يوم الثلاثاء بمعلوماته عدم استلامه يوم الاثنين

$$P(T|M^c) = \frac{P(T \cap M^c)}{P(M^c)} = \frac{P(T|A) \cdot P(A) + P(T|R) \cdot P(R)}{1 - P(M)} = \frac{(0.2 \cdot 0.6) + (0.1 \cdot 0.4)}{0.89} = \frac{0.16}{0.89}$$

c) استلمته لو الميل لم تستلمه حتي يوم الاربعاء يعني لم تستلمه (الاثنين, الثلاثاء, الاربعاء) يبقي هجيب احتمال انها بمعلوميه انها لم تستلمه حتي يوم الاربعاء لكنها اتقبلت

$$P(A|M^C T^C W^C) = \frac{P(M^C T^C W^C \cap A)}{P(M^C T^C W^C)} = \frac{P(M^C T^C W^C | A) \cdot P(A)}{P(M^C T^C W^C)}$$

$$= \frac{P(M^C T^C W^C | A) \cdot P(A)}{P(M^C T^C W^C | A) \cdot P(A) + P(M^C T^C W^C | A^C) \cdot P(A^C)}$$

➤ $P(M^C T^C W^C) = 1 - P(M.T.W) = 1 - (0.15 + 0.20 + 0.25) = 0.4$

$$= \frac{0.4 \cdot 0.6}{(0.4 \cdot 0.6) + (0.75 \cdot 0.4)} = \frac{12}{27}$$

d) هنا عايز احتمال قبولها لو الميل وصل يوم الخميس

$$P(A|Th) = \frac{P(Th|A) \cdot P(A)}{P(Th)} = \frac{0.15 \cdot 0.6}{(0.15 \cdot 0.6) + (0.15 \cdot 0.4)} = \frac{3}{5}$$

e) هنا عايز احتمال قبولها في حاله عدم ارسال ميل خلال الاسبوع

$$P(A|\text{no mail}) = \frac{P(\text{no mail}|A) \cdot P(A)}{P(\text{no mail})} = \frac{0.15 \cdot 0.6}{(0.15 \cdot 0.6) + (0.4 \cdot 0.4)} = \frac{9}{25}$$

• A card is drawn from a deck of cards, find the probability of:

- a) Getting the 3 diamond
- b) Getting queen
- c) Getting a king
- d) Getting a face card
- e) Getting an ace (A)
- f) Getting a diamond

**قبل ما نحل السؤال لازم نكون شويه معلومات عن ورق الكوتشينه

- then, Sample space = 52

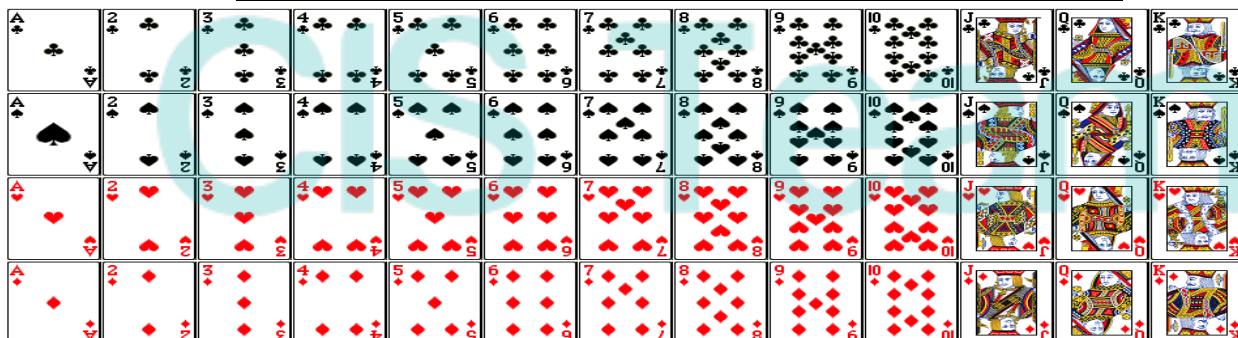
(1) ورق الكوتشينه مكون من 52
(2) الـ 52 ورقه مقسمين لـ 13 نوع و كل نوع له 4 أشكال
الـ 13 نوع هما:

A 2 3 4 5 6 7 8 9 10 JOKER QUEEN KING

الـ 4 أشكال هما:



Diamond Club Spade Heart



Solution:

a) Getting the 3 of Diamond

$$P(E) = \frac{1}{52}$$

احتمال ظهور ورقه عليها 3 diamond

b) Getting queen

$$P(E) = \frac{4}{52}$$

c) Getting king

$$P(E) = \frac{4}{52}$$

d) Getting a Diamond

$$P(E) = \frac{13}{52}$$

e) Getting an Ace (ace = 1)

$$P(E) = \frac{4}{52}$$

f) Getting a face card

(King, queen, joker) احتماليه ظهور ورقه عليها صورته, ورق الكوتشينه بيحتوي على 3 صور

$$E = 3 \times 4 = 12$$

$$P(E) = \frac{12}{52}$$

و كل صورته ليها 4 ورقات

Independence: An event A is called independent B if $\longrightarrow P(A/B) = P(A)$ ➤ A independent of B $\iff P(A \cap B) = P(A) * P(B)$

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

EX: Two coins are flipped, and all 4 outcomes are assumed to be equally likely الاحداث متساويهIf \underline{E} is the event that the first coin lands on heads and \underline{F} the event that the second lands on tails, then \underline{E} and \underline{F} are independent, since $P(EF) = P(\{(H,T)\}) = 1/4$, whereas $P(E) = P(\{(H,H),(H,T)\}) = 1/2$ and $P(F) = P(\{(H,T),(T,T)\}) = 1/2$. $P(E)$ and $P(F)$ are independent or not?**Solution:**

$$P(E) = 1/2 \quad \&\& \quad P(F) = 1/2 \quad \&\& \quad P(E) * P(F) = 1/2 * 1/2 = 1/4 \quad \&\& \quad P(EF) = 1/4$$

$$P(E) * P(F) = P(EF) = 1/4 \quad \longrightarrow \quad P(E) \text{ and } P(F) \text{ are independent}$$

3 events (E, F, and G) are said to be independent if:

$$1. P(E.F.G) = P(E) * P(F) * P(G)$$

$$2. P(E.F) = P(E) * P(F)$$

$$3. P(E.G) = P(E) * P(G)$$

$$4. P(F.G) = P(F) * P(G)$$

لازم الـ 4 شروط يتحققوا

Random Variables**Random Variables:** a real – valued numeric outcome from an experiment**Discrete Random Variables:** can take only a finite (or count ably infinite) set of outcome**Continuous Random Variables:** can take an any value along a continues**Probability Distribution:** my be table, or formula, or graph, that describes random variable values, and its corresponding probability

EX: When 10 coins are to be used

X = the number of Heads

X = the total of 2 dice rolls

$$x=0, 1, 2, 3, \dots, 10$$

$$x=2, 4, 6, \dots, 12$$

اي رقم من (0:12) هو عبارته عن مجموع الرقمين اللي ظهوروا في الرمي

➤ Probability Mass Function:

$$P(y) = P(Y=y), \quad \text{where } P(y) \geq 0$$

(for all y)

$$\sum_{all y} P(y) = 1$$

➤ Cumulative Distribution Function (CDF):

$$F(y) = P(Y \leq y), \quad \text{where } F(b) = P(Y \leq b)$$

$$\sum_{y=-\infty}^b P(y)$$

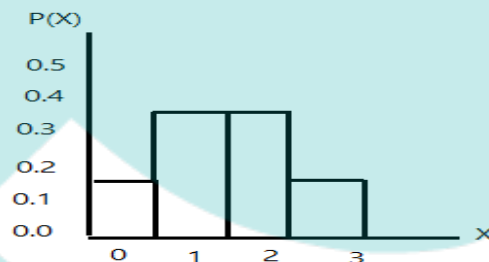
EX: Toss a fair coin 3 times let X = number of Tails

احتمال ظهور الاحداث	X
TTT	1/8
TTH	1/8
THT	1/8
HTT	1/8
TTH	1/8
HTH	1/8
HHT	1/8
HHH	1/8

X	P(x)
0	1/8
1	3/8
2	3/8
3	1/8

$$1 = (1/8 + 3/8 + 3/8 + 1/8) = P(X) \quad \text{مجموع احتمالات}$$

The histogram



The Cumulative Distribution Function (CDF):

EX: $p(1) = \frac{1}{4}$ $p(2) = \frac{1}{2}$ $p(3) = \frac{1}{8}$ $p(4) = \frac{1}{8}$

then its cumulative distribution function is

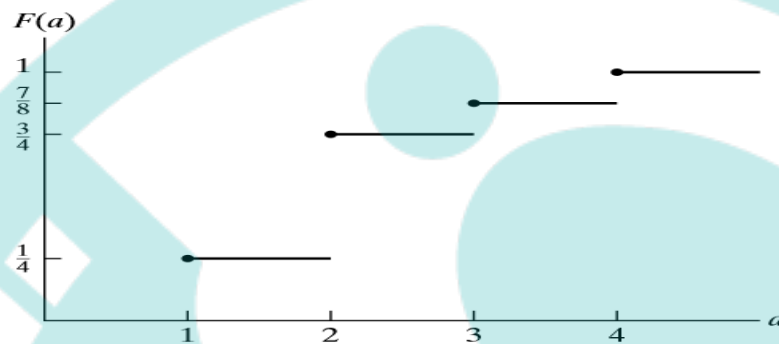
$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases}$$

الـ 0 دا ثابت ببدا بيه في كل مره

$$P(1) + P(2) = (3/4) \quad \text{والـ}$$

$$P(1) = (1/4)$$

$$1 = (1) = \text{مجموع الاحتمالات اللي عندي ولازم مجموع الاحتمالات}$$



EX: Three balls are randomly chosen from an urn containing 3 white, 3 red, and 5 black balls. Suppose that we win \$1 for each white ball selected and lose \$1 for each red ball selected. If we let X denote our total winnings from the experiment, then X is a random variable taking on the possible values $(0, \pm 1, \pm 2, \pm 3)$ with respective probabilities

Solution:

Ball	W	R	B
balls	3	3	5
wins	+1	-1	0

** عندي هنا الـ $(-1) = (+1)$ ← لان عدد الكرات الحمراء = عدد الكرات البيضاء = 3

** انما لو الـ 2 كان اعددهم مختلفه كنت هحسب احتمال الـ $(+1)$ لواحده والـ (-1) لواحده

$$P\{X = 0\} = \frac{\binom{5}{3} + \binom{3}{1}\binom{3}{1}\binom{5}{1}}{\binom{11}{3}} = \frac{55}{165}$$

$$P\{X = 1\} = P\{X = -1\} = \frac{\binom{3}{1}\binom{5}{2} + \binom{3}{2}\binom{3}{1}}{\binom{11}{3}} = \frac{39}{165}$$

$$P\{X = 2\} = P\{X = -2\} = \frac{\binom{3}{2}\binom{5}{1}}{\binom{11}{3}} = \frac{15}{165}$$

$$P\{X = 3\} = P\{X = -3\} = \frac{\binom{3}{3}}{\binom{11}{3}} = \frac{1}{165}$$

CIS Team

Expected Values

- mean = $\mu = E(y) = \sum_{all\ y} y * P(y)$
- variance = $\sigma^2 = V(y) = E[(y - \mu)^2] = E[y^2] - \mu^2$

➤ **EX:** Find $E[X]$, where X is the outcome when we roll a fair die.

Solution: هنا هرمي حجر نرد يعني عندي 6 احتمالات والاحتمالات اللي عندي كلها متساويه

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$$

$$E[X] = \sum_{all\ x} X * P(X) = 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6) = 7/2$$

➤ **EX:** Calculate $Var(X)$ if X represents the outcome when a fair die is rolled.

Solution: $Var(X) = E[X^2] - \mu^2$

$$E[X^2] = \sum_{all\ x} X^2 * P(X) = 1^2(1/6) + 2^2(1/6) + 3^2(1/6) + 4^2(1/6) + 5^2(1/6) + 6^2(1/6) = 91/6$$

$$\mu^2 = (7/2)^2 \quad \text{من المثال السابق} \quad \longrightarrow \quad Var(X) = E[X^2] - \mu^2 = 91/6 - (7/2)^2 = 35/12$$

➤ **EX:** If the number of patients in the ER in anywhere given hour is considered ar.v(X) with the probability distribution as:

X	10	11	12	13	14
P(X)	0.4	0.2	0.2	0.1	0.1

Find the expected value of (X)

Solution: $E[X] = \sum_{i=1}^5 x_i * P(x) = 10(0.4) + 11(0.2) + 12(0.2) + 13(0.1) + 14(0.1) = 11.3$

Bernoulli Distribution

$$P_X(1) = P(X=1) = P(\text{Success})$$

and

لو كنت بعمل تجربه واحده فقط والتجربه ليها 2 outcome
 $P_X(0) = P(X=0) = 1 - P(\text{Failure})$

Binomial Distribution

$$P_X(K) = P(X=K) = \binom{n}{K} P^K * (1-P)^{n-K}$$

لو كنت بعمل اكثر من تجربه بس التجربه برود ليها 2 outcome
 for $K = 0, 1, 2, \dots$

: n \longrightarrow number of trials عدد مرات اجراء التجربه

&& K \longrightarrow successes outcomes

$$\mu = E[X] = n * P$$

$$Var(X) = n * p(1 - p)$$

EX: X is a binomial random variable with parameters ($n=5$ $p=1/2$)

Solution: $P(X=K) = \binom{n}{K} P^K * (1-P)^{n-K}$

$$P(X=0) = \binom{5}{0} * (1/2)^0 * (1/2)^5 = 1/32$$

$$P(X=1) = \binom{5}{1} * (1/2)^1 * (1/2)^4 = 5/32$$

$$P(X=2) = \binom{5}{2} * (1/2)^2 * (1/2)^3 = 10/32$$

$$P(X=3) = \binom{5}{3} * (1/2)^3 * (1/2)^2 = 10/32$$

$$P(X=4) = \binom{5}{4} * (1/2)^4 * (1/2)^1 = 5/32$$

$$P(X=5) = \binom{5}{5} * (1/2)^5 * (1/2)^0 = 1/32$$

EX: Toss a coin 20 times. Find the probability of obtaining 2 or fewer heads

Solution: هنا هو عايز احتمال ان يظهر 2 head علي الاقل يعني $x = 0, 1, 2$ ← $(0, 1, 2)$

$$P = 1/2 \quad X = 0, 1, 2 \quad n = 20$$

$$P(X=0) = \binom{20}{0} * (1/2)^0 * (1/2)^{20} = 1.9 * 10^{-7}$$

$$P(X=1) = \binom{20}{1} * (1/2)^1 * (1/2)^{19} = 1.9 * 10^{-5}$$

$$P(X=2) = \binom{20}{2} * (1/2)^2 * (1/2)^{18} = 1.8 * 10^{-4}$$

$$P(X \leq 2) = 1.9 * 10^{-7} + 1.9 * 10^{-5} + 1.8 * 10^{-4}$$

EX: If the probability of purchasing a defective computer 0.02 what is the probability of purchasing 2 defective computers in a lot of 10

Solution: لو هشتري 10 اجهزه ماهو احتمال ان يكون فيهم جهازين بهما عطل
مع العلم ان احتمال ان يكون الجهاز به عطل $0.02 = P$

$$X = 2 \quad n = 10 \quad P = 0.02$$

$$P(X=2) = \frac{n!}{X!(n-X)!} P^X (1-P)^{n-X} = \frac{10!}{2!(10-2)!} (0.02)^2 (1-0.02)^{10-2} = 0.01531$$

Poisson Distribution

$$P(i) = P(X=i) = \frac{e^{-\lambda} * \lambda^i}{i!} \quad : i = 0, 1, 2, \dots$$

دا بستخدمه لما يكون عندي مده او فتره زمنييه

λ → The expected number of successes

n → number of successes → $e = 2.71828\dots$

EX: What is the probability of 5 customers arriving in 5 minutes when the mean (expected value) is 4?

Solution: $K = 5 \quad \lambda = 4 \quad P(5) = P(X=5) = \frac{e^{-4} * 4^5}{5!}$

EX: For a 400 page book. Find the probability that there are 5 or less typos. If the expected value of typos is 6?

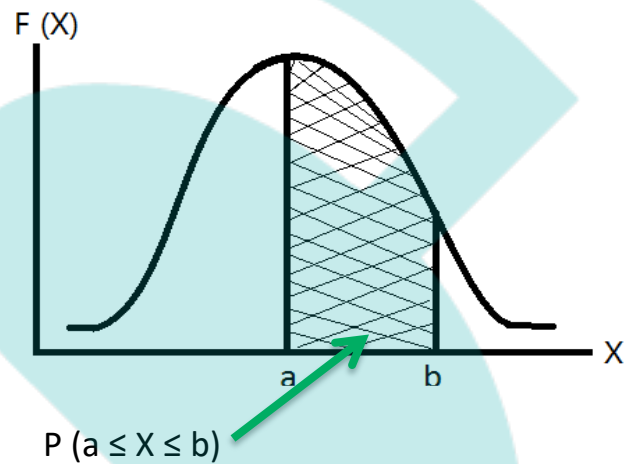
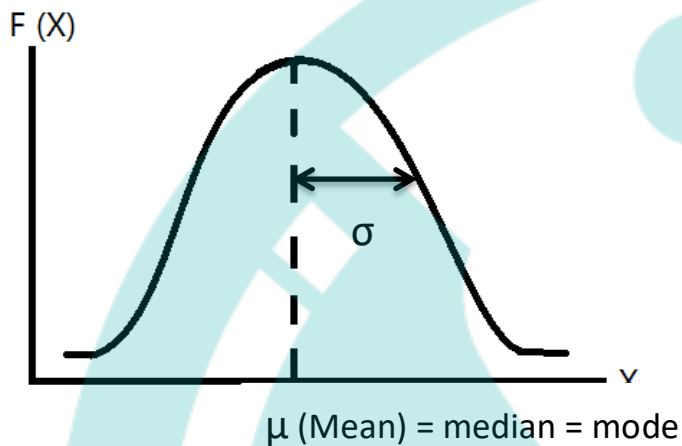
Solution: $K = 5 \quad \lambda = 6$

$$P(X \leq 5) = P(0) + P(1) + \dots + P(5) = \frac{e^{-6} * 6^0}{0!} + \frac{e^{-6} * 6^1}{1!} + \dots + \frac{e^{-6} * 6^5}{5!} = 0.445$$

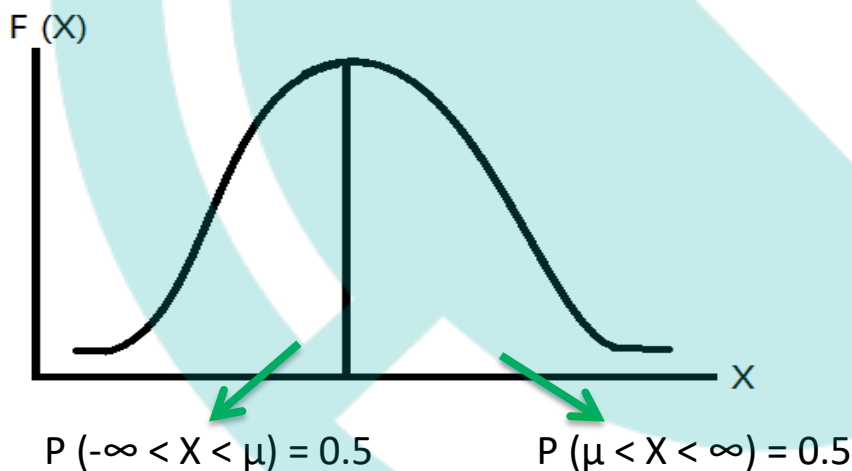
EX: Suppose that the probability that an item produced by a certain machine will be defective is 0.1. Find the probability that a sample of 10 items will contain at most 1 defective item.

Solution: عندي اله بتنتج منتج ما لو افترضت ان احتمال يكون المنتج اللي طالع منها معاب 0.1 لو انتجت 10 منتجات ماهو احتمال ان يكون منهم علي الاكثر منتج معاب يعني ممكن يكون (0 او 1)

$$= \binom{10}{0} (0.1)^0 * (0.9)^{10} + \binom{10}{1} (0.1)^1 * (0.9)^9 = 0.7361$$



الاحتمال = المنطقة المظللة تحت المنحني
= التكامل المحدود لهذه المنطقة



Standard Normal Distribution (Z):

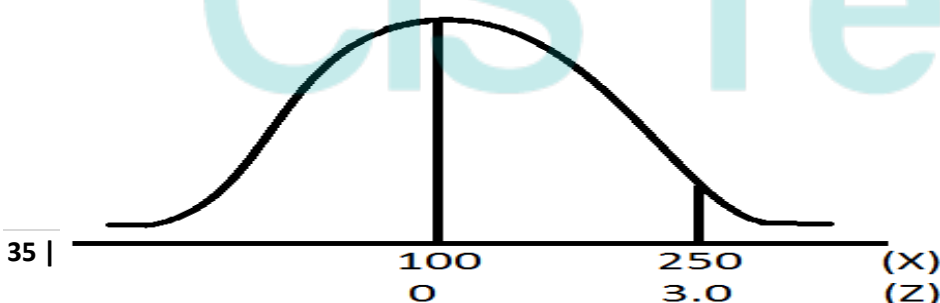
The special case of a normal distribution where ($\mu = 0$ $\sigma = 1$) $Z = \frac{X - \mu}{\sigma}$

EX: If x is distribution normally with mean of 100 and standard deviation of 50 the (Z) value for $x = 250$?

Solution:

$$Z = \frac{X - \mu}{\sigma} = \frac{250 - 100}{50} = 3.0$$

$\mu = 100$ $\sigma = 50$



➤ Counting $(n_1 \cdot n_2 \cdots n_r) : r \rightarrow$ number of experiments

➤ PERMUTATIONS $n (n - 1) (n - 2) \cdots 3 \cdot 2 \cdot 1 = n!$

➤ COMBINATIONS $\frac{n!}{(n-r)! \cdot r!}$

➤ MULTINOMIAL COEFFICIENTS $\frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_r!}$

➤ Sample Mean

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

➤ Deviation

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x} = n\bar{x} - n\bar{x} = 0$$

➤ Quartile range $q_3 - q_1$

➤ Sample Variance and Sample Standard Deviation

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

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

➤ Empirical Rule

1. Approximately 68 percent of the observations lie within

$$\bar{x} \pm s$$

2. Approximately 95 percent of the observations lie within

$$\bar{x} \pm 2s$$

3. Approximately 99.7 percent of the observations lie within

$$\bar{x} \pm 3s$$

➤ Sample Correlation Coefficient

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

➤ Confusion matrix / Crosstabs

- Precision = $TP / (TP + FP)$
- Recall / Sensitivity = $TP / (TP + FN)$

- Specificity = $TN / (TN + FP)$
- Accuracy = $(TP + TN) / (TP + TN + FP + FN)$
- F-Score = $(2 * Precision * Recall) / (Precision + Recall)$

$$\text{Percentile} = \frac{\text{مكان الرقم في الجدول}}{\text{عدد القيم}} \times 100$$

➤ Probability $P(E)$:

$$0 \leq P(E_i) = \frac{\text{Number of ways that an event } E_i \text{ can occur}}{\text{Number of sample space}} \leq 1 \quad \text{for all (i).}$$

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

➤ $P(A \cup B) = P(A) + P(B) \quad : (A, B) \text{ are disjoint}$

Number of all outcome satisfying (A and B)

➤ $P(A \text{ and } B) = \frac{\text{Number of all outcome satisfying (A and B)}}{\text{The total number of sample space}}$

➤ Conditional probability:

$$P(A/B) = \frac{\text{Number of outcomes in } (A \cap B)}{\text{Number of outcomes in } B} = \frac{P(A \cap B)}{P(B)} \quad : P(B) > 0$$

➤ Bayes Rule

$$P(B|A) = \frac{P(A|C_I).P(C_I)}{P(A|C_1).P(C_1) + \dots + P(A|C_I).P(C_I)}$$

➤ Expected Values

- mean = $\mu = E(y) = \sum_{all y} y * P(y)$

- variance = $\sigma^2 = V(y) = E[(y - \mu)^2] = E[y^2] - \mu^2$

➤ Bernoulli Distribution

$$P_X(1) = P(x=1) = P(\text{Success}) \quad \text{and} \quad P_X(0) = P(x=0) = 1-P(\text{Failure})$$

➤ Binomial Distribution

$$P_X(K) = P(X=K) = \binom{n}{K} P^K * (1-P)^{n-K} \quad \text{for } K = 0, 1, 2, \dots$$

➤ Standard Normal Distribution (Z):

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

Self-confidence is the first secret to success so be confident in yourself ... Good luck ♥

With best wishes from CIS_Team