الدرس الأول

"Basic Concepts"

* population: is the largest collection of entities (elements or individuals) in which we are interested at particular time and about which we want to draw some conclusions (N = population size).
* Sample: is a part of population (n = sample size).
* Variable: characteristic to be measured (number or words) on the elements (in population or in sample).
* types of Variables
  - quantitative (by numbers)
    - Discrete (Jumps or gaps or specific)
    - Continuous (no gaps or within interval)
  - qualitative (by words)
    - Nominal (not order)
    - Ordinal (order)

* data: raw material of statistics.
* statistics: field of study concerned with.
* descriptive statistics: collection, organization, summarization, describe, and analysis of data.
* inferential statistics: reach decisions, inference, and conclusions about a large body of data (population) by only part of the data (sample) is observed.

الدرس الثاني

جزء الأول

"Frequency Distribution"

* Ordered array or set: listing of values in order from the smallest to the largest values with frequency.

class interval	true class interval	mid point or mid interval
[lower(1) - upper(1)]	[true lower(1) - true upper(1))	$\text{mid}(1) = \frac{\text{lower}(1) + \text{upper}(1)}{2} = \frac{\text{true lower}(1) + \text{true upper}(1)}{2}$
+w +w	=	+w
[lower(2) - upper(2)]	[true lower(2) - true upper(2))	$\text{mid}(2) = \frac{\text{lower}(2) + \text{upper}(2)}{2} = \frac{\text{true lower}(2) + \text{true upper}(2)}{2}$
+w +w	=	+w
[lower(3) - upper(3)]	[true lower(3) - true upper(3))	$\text{mid}(3) = \frac{\text{lower}(3) + \text{upper}(3)}{2} = \frac{\text{true lower}(3) + \text{true upper}(3)}{2}$

The method to find the following:

* Class interval ⇒ true class interval:

$d = \text{lower}(2) - \text{upper}(1) = \text{lower}(3) - \text{upper}(2)$ (gaps) ;

$\text{true lower}(1) = \text{lower}(1) - \frac{d}{2}$;

$\text{true upper}(1) = \text{upper}(1) + \frac{d}{2}$;

=

$\text{true lower}(2) = \text{lower}(2) - \frac{d}{2}$;

$\text{true upper}(2) = \text{upper}(2) + \frac{d}{2}$;

=

$\text{true lower}(3) = \text{lower}(3) - \frac{d}{2}$;

$\text{true upper}(3) = \text{upper}(3) + \frac{d}{2}$.

* true class interval $\Rightarrow$ class interval (where d or $\frac{d}{2}$ is known):

$$\begin{cases} \text{lower}(1) = \text{true lower}(1) + \frac{d}{2} \\ \text{upper}(1) = \text{true upper}(1) - \frac{d}{2} \end{cases}$$

$$\begin{cases} \text{lower}(2) = \text{true lower}(2) + \frac{d}{2} \\ \text{upper}(2) = \text{true upper}(2) - \frac{d}{2} \end{cases}$$

$$\begin{cases} \text{lower}(3) = \text{true lower}(3) + \frac{d}{2} \\ \text{upper}(3) = \text{true upper}(3) - \frac{d}{2} \end{cases}$$

* midpoint $\Rightarrow$ true class interval:

$$\begin{aligned} w = \text{width} &= \text{lower}(2) - \text{lower}(1) = \text{lower}(3) - \text{lower}(2) \\ &= \text{upper}(2) - \text{upper}(1) = \text{upper}(3) - \text{upper}(2) \end{aligned}$$ by using the class interval

$$\begin{aligned} &= \text{true upper}(1) - \text{true lower}(1) \\ &= \text{true upper}(2) - \text{true lower}(2) \\ &= \text{true upper}(3) - \text{true lower}(3) \end{aligned}$$ by using the true class interval

$$= \text{mid}(2) - \text{mid}(1) = \text{mid}(3) - \text{mid}(2)$$ by using the mid point

$$\begin{cases} \text{true lower}(1) = \text{mid}(1) - \frac{w}{2} \\ \text{true upper}(1) = \text{mid}(1) + \frac{w}{2} \end{cases}$$

$$\begin{cases} \text{true lower}(2) = \text{mid}(2) - \frac{w}{2} \\ \text{true upper}(2) = \text{mid}(2) + \frac{w}{2} \end{cases}$$

$$\begin{cases} \text{true lower}(3) = \text{mid}(3) - \frac{w}{2} \\ \text{true upper}(3) = \text{mid}(3) + \frac{w}{2} \end{cases}$$

true upper(1) = true lower(2); true upper(2) = true lower(3)

الجزء الثاني

* Frequency ≅ freq.
* Relative ≅ r. (relative ≅ propotion)
* Percentage ≅ per. (percentage ≅ percent)
* Cumulative ≅ cum.

$\div n$ $\quad\quad$ $\times 100\%$

freq.	cum. freq.	r. freq.	cum. r. freq.	per. freq.	cum. per. freq.
	n		1		100%
Total = n	~~XXX~~	Total = 1	~~XXX~~	Total = 100%	~~XXX~~

$\div n$ $\quad\quad$ $\times 100\%$

we have another Method to find the following :

* Cumulative frequency :

  cum. freq. of 1^{st} class interval = freq.

  cum. freq. of any class interval = freq. + cum. freq. of preceding class interval

* Cumulative relative frequency :

  cum. r. freq. of 1^{st} class interval = r. freq.

  cum. r. freq. of any class interval = r. freq. + cum. r. freq. of preceding class interval

* cumulative percentage frequancy :

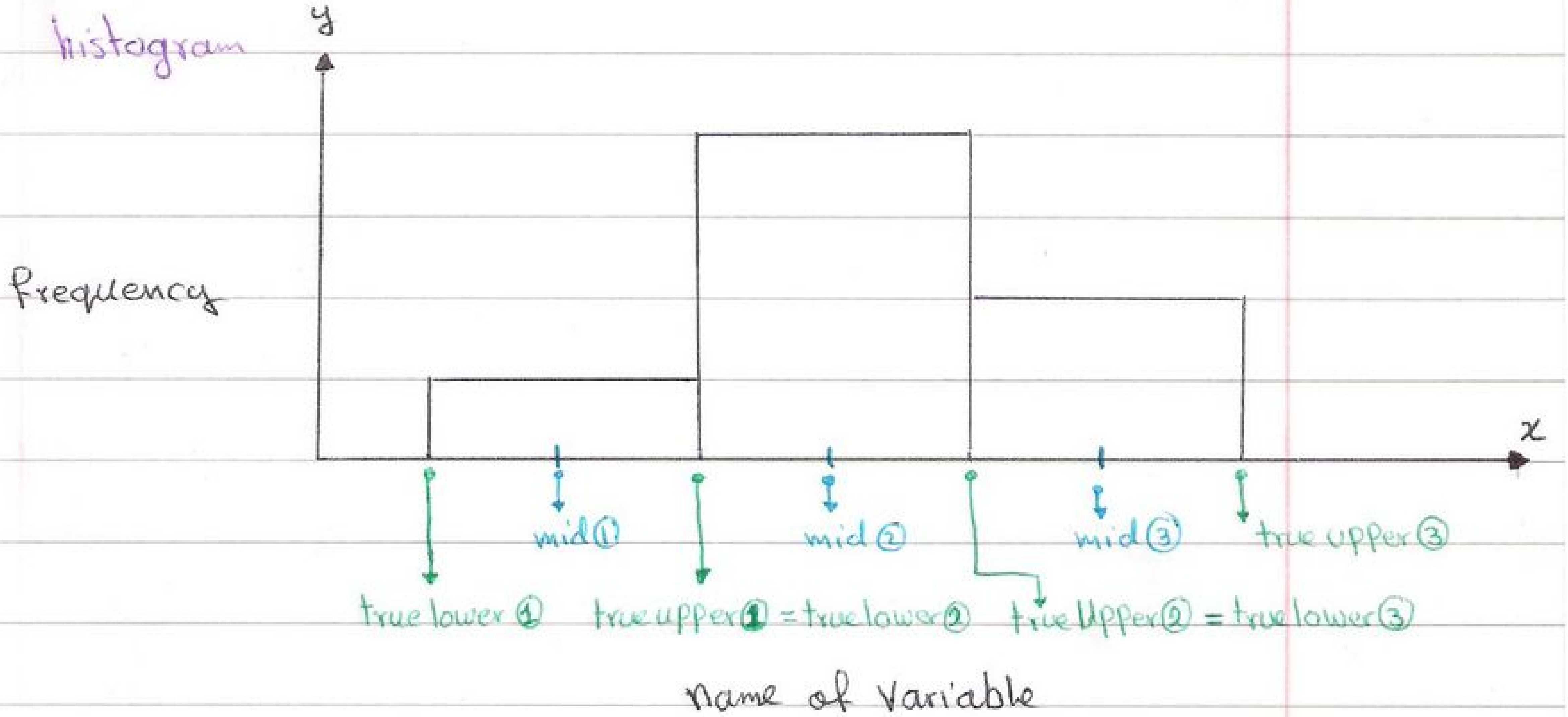
  cum. per. freq. of 1^{st} class interval = per. freq.

  cum. per. freq. of any class interval = per. freq. + cum. per. freq. of preceding class interval

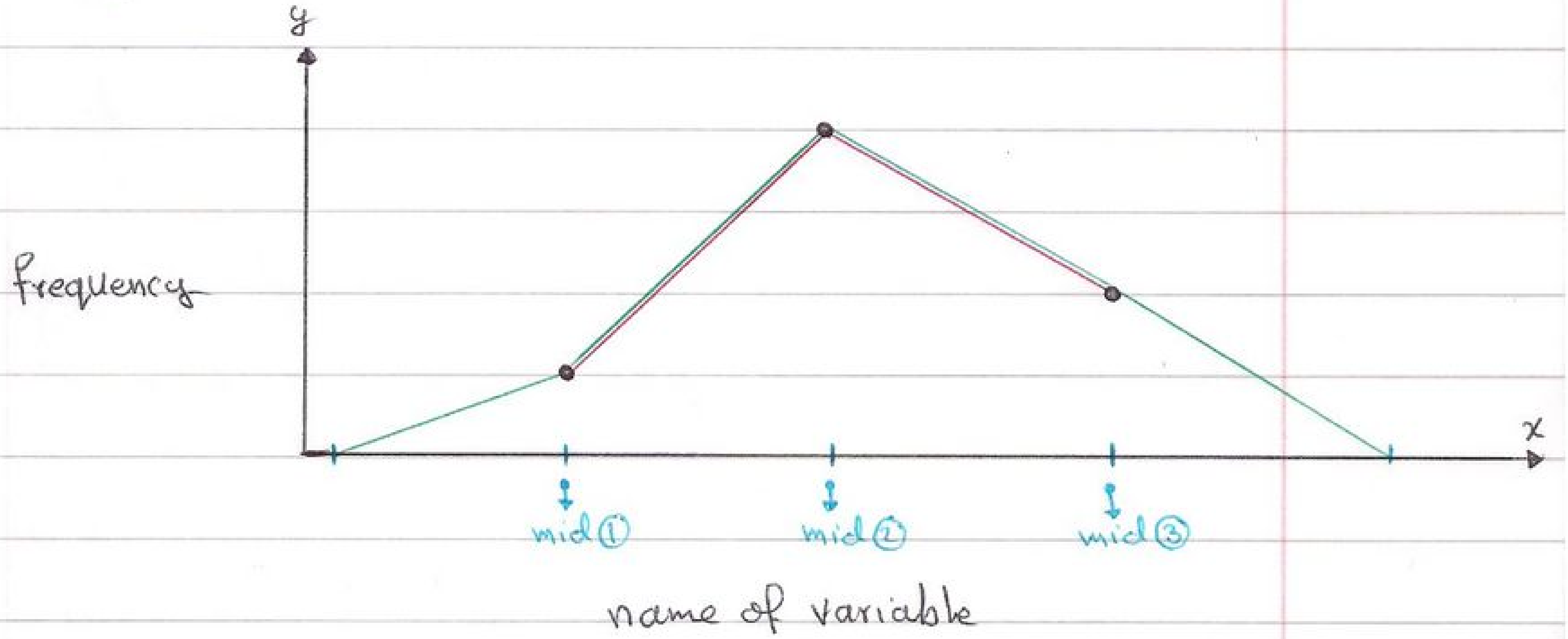
جزء الثاني

* displaying the frequency distribution:

histogram



polgon



= polgon closed

= polgon open

الدرس الثالث

الجزء الأول

"Descriptive Statistics"

measures of

Central Tendency (Location) ← → Dispersion (Variation)

* mean (unit)
* median (unit)
* mode (unit)

* range (unit)
* Variance = (standard deviation)^2 (unit)^2
* standard deviation = $\sqrt{\text{variance}}$ (unit)
* coefficient of variation (unit-less)
C.V.

Population		Sample
* $X_1, X_2, \dots, X_N$ * any measure here it called "parameter"	→	* $x_1, x_2, \dots, x_n$ * any measure here it called "statistic"

	Population	sample
mean	$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Variance	$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}$; $\sigma^2 \geq 0$	$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$; $s^2 \geq 0$
standard deviation	$\sigma = \sqrt{\sigma^2}$; $\sigma \geq 0$	$s = \sqrt{s^2}$; $s \geq 0$
size	N	n

جزء الثاني

"استخدام الآلة الحاسبة لإيجاد المقاييس الإحصائية"

① تنظيف الجهاز: نعمل هذه العملية كلما كانت لدينا أي بيانات جديدة أو إرجاع الجهاز لوضع D....

SHIFT + (9) CLR → 3:ALL → = → =

② تحويل الآلة إلى آلة إحصائية وتعبئة البيانات:....

(i) جدول لا يحتوي على عمود التكرار:....

MODE → 3:STAT → 1:VAR → (table below) → AC

	X
1	a =
2	b =
3	c =

(ii) جدول يحتوي على عمود التكرار:....

SHIFT + MODE → 4:STAT → 1:ON

MODE → 3:STAT → 1:VAR → (table below) → AC

	X	FREQ
1	a =	d =
2	b =	e =
3	c =	f =

③ إيجاد المقاييس الإحصائية:....

حساب M أو $\bar{x}$ (متوسط أو وسط حسابي)

SHIFT + (1) STAT → 5:Var → 1:n   2: $\bar{x}$

3: $x\sigma n$ (σx)   4: $x\sigma n$-1 (Sx)   حساب S (الإنحراف المعياري)

حساب σ (الإنحراف المعياري)

→ نختار الرقم حسب المقياس المطلوب → =

ملاحظة: لحساب التباين:....

3 أو 4 → = → x^2 → =

الدرس الرابع
جزء الأول

"Probability"

* Experiment: some procedure or process that we do.
* Sample Space (Ω): set of all possible outcomes of experiment; where $n(\Omega)$ is the number of outcomes (elements) in Ω.
* Event (E): any subset of Ω; where $n(E)$ is the number of outcomes in E.

$E \subseteq \Omega$

$\emptyset \subseteq \Omega$ (impossible event)

$\Omega \subseteq \Omega$ (sure event)

* Equally likely outcomes of experiment: if the outcomes have the same chance of occurrence.
* Probability: measure used to measure the chance of occurrence of event; which is between 0 and 1.
* Probability of event E: $0 \le P(E) = \frac{n(E)}{n(\Omega)} \le 1$

$P(\emptyset) = 0$

$P(\Omega) = 1$

* Operations on events:

1. Union:

- $A \cup B \cong A \text{ or } B \cong A + B$
- $n(A \cup B) = n(A) + n(B) - n(A \cap B) \Leftrightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $A \cup B = \Omega$ or $P(A \cup B) = 1 \Leftrightarrow$ A and B are exhaustive
- $A \cup A' = \Omega$ (exhaustive)
- $P(A \cup B) = \frac{n(A \cup B)}{n(\Omega)}$

② intersection:

- $A \cap B \cong A$ and B
- $P(A \cap B) = \frac{n(A \cap B)}{n(\Omega)}$
- $A \cap B = \emptyset$ or $P(A \cap B) = 0 \Leftrightarrow$ A and B are disjoint (mutually exclusive)
- $A \cap A' = \emptyset$ (disjoint or mutually exclusive)

③ Complement:

- $\bar{A} \cong A^c \cong A' \cong$ not A
- $n(A) + n(A') = n(\Omega) \Leftrightarrow P(A) + P(A') = 1$

* Marginal probability:

① table 2x2:

	B	B^c	
A	$n(A \cap B)$ +	$n(A \cap B^c)$ =	$n(A)$
	+	+	+
A^c	$n(A^c \cap B)$ +	$n(A^c \cap B^c)$ =	$n(A^c)$
	‖	‖	‖
	$n(B)$ +	$n(B^c)$ =	$n(\Omega)$

$\Leftrightarrow$

	B	B^c	
A	$P(A \cap B)$ +	$P(A \cap B^c)$ =	$P(A)$
	+	+	+
A^c	$P(A^c \cap B)$ +	$P(A^c \cap B^c)$ =	$P(A^c)$
	‖	‖	‖
	$P(B)$ +	$P(B^c)$ =	1

الجمع عاموديا أو عرضيا للتقاطع يساوي الأطراف عاموديا أو عرضيا على التوالي ...

② tables of 2×3 ; 3×3 ; 2×4 ; 3×4 , 4×4 , ... : for example if we take

	B_1		B_2		B_3		
A_1	$n(A_1 \cap B_1)$	+	$n(A_1 \cap B_2)$	+	$n(A_1 \cap B_3)$	=	$n(A_1)$
	+		+		+		+
A_2	$n(A_2 \cap B_1)$	+	$n(A_2 \cap B_2)$	+	$n(A_2 \cap B_3)$	=	$n(A_2)$
	+		+		+		+
A_3	$n(A_3 \cap B_1)$	+	$n(A_3 \cap B_2)$	+	$n(A_3 \cap B_3)$	=	$n(A_3)$
	‖		‖		‖		‖
	$n(B_1)$	+	$n(B_2)$	+	$n(B_3)$	=	$n(\Omega)$

$\Longleftrightarrow$

	B_1		B_2		B_3		
A_1	$p(A_1 \cap B_1)$	+	$p(A_1 \cap B_2)$	+	$p(A_1 \cap B_3)$	=	$p(A_1)$
	+		+		+		+
A_2	$p(A_2 \cap B_1)$	+	$p(A_2 \cap B_2)$	+	$p(A_2 \cap B_3)$	=	$p(A_2)$
	+		+		+		+
A_3	$p(A_3 \cap B_1)$	+	$p(A_3 \cap B_2)$	+	$p(A_3 \cap B_3)$	=	$p(A_3)$
	‖		‖		‖		‖
	$p(B_1)$	+	$p(B_2)$	+	$p(B_3)$	=	1

الجمع عامودياً أو عرضياً للتقاطع يساوي الأطراف عامودياً أو عرضياً على التوالي ....

- $A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3 = A_1 \cap A_2 \cap A_3 = \emptyset$;

  $B_1 \cap B_2 = B_1 \cap B_3 = B_2 \cap B_3 = B_1 \cap B_2 \cap B_3 = \emptyset$.
- $A_1' = A_2 \cup A_3$; $A_2' = A_1 \cup A_3$; $A_3' = A_1 \cup A_2$; $B_1' = B_2 \cup B_3$; $B_2' = B_1 \cup B_3$; $B_3' = B_1 \cup B_2$.
- $A_1^c \cap B_1 = (A_2 \cup A_3) \cap B_1 = (A_2 \cap B_1) \cup (A_3 \cap B_1) \Leftrightarrow p(A_1^c \cap B_1) = p(A_2 \cap B_1) + p(A_3 \cap B_1)$

  $A_1 \cap B_1^c = A_1 \cap (B_2 \cup B_3) = (A_1 \cap B_2) \cup (A_1 \cap B_3) \Leftrightarrow p(A_1 \cap B_1^c) = p(A_1 \cap B_2) + p(A_1 \cap B_3)$

وعلى هذا يقاس الأمر لأي حادثتين ....

* conditional probability:

- $A|B \cong A$ given $B \cong A$ knowing B
- $P(A|B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{n(A\cap B)}{n(B)}$

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

- A and B independent $\Leftrightarrow P(A\cap B) = P(A)\,P(B)$

$\Leftrightarrow P(A|B) = P(A)$

$\Leftrightarrow P(B|A) = P(B)$

جزء الثاني

"Bayes' Theorem"

		Disease		
		present (D)	absent (D`)	
test result	+ve (T)	$n(T \cap D)$	$n(T \cap D')$	$n(T)$
	-ve (T`)	$n(T' \cap D)$	$n(T' \cap D')$	$n(T')$
		$n(D)$	$n(D')$	$n(\Omega)$

from the table, we can find:

① false +ve result $P(T|D') = \frac{n(T \cap D')}{n(D')}$

② false -ve result $P(T'|D) = \frac{n(T' \cap D)}{n(D)}$

③ sensitivity of the test $P(T|D) = \frac{n(T \cap D)}{n(D)}$

④ specificity of the test $P(T'|D') = \frac{n(T' \cap D')}{n(D')}$

+ = 1　　+ = 1

Bayes' theorem: if we have $P(D)$ is the relevant disease in general Population which obtained from another independent study ($P(\bar{D}) = 1 - P(D)$), then:

⑤ Predictive value +ve of the test

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D')P(D')}$$

⑥ Predictive value -ve of the test

$$P(D'|T') = \frac{P(T'|D')P(D')}{P(T'|D')P(D') + P(T'|D)P(D)}$$