التحليل الإحصائي

د.أحمد فرحان

إعداد :

Ghayda&dody-11

2014-2015

المحاضرة (١)

المجموعات

تعريف المجموعة :-

يمكن تعريف المجموعة على أنها عدد من العناصر بينها صفات مشتركة تكتب بين حاصرتين {   } و تسمى بأحد الحروف الهجائية الكبيرة ..... A , B , C ,

و من الأمثلة على المجموعات $A=\{1,3,5,7,9,.......\}$

الأشياء التي تتكون منها المجموعة تسمى عناصر المجموعة و ترمز للعناصر بواسطة حروف صغير مثل :-

a , b ,c , ...

يستخدم الرمز ϵ "ينتمي إلى " ليبين عناصر المجموعة فمثلاً إذا كان العنصر a من ضمن عناصر المجموعة A فإننا نقول أن a ينتمي إلى المجموعة A و يكتب بالصورة $a \in A$

أما إذا كان a ليس عنصراً من عناصر المجموعة A فإننا نقول أن العنصر a لا ينتمي إلى المجموعة A ويكتب على الصورة $a \notin A$

طريقة كتابة المجموعات :

طريقة العد (سرد العناصر):-

يتم فيها وضع جميع عناصر المجموعة ، أو جزء منها ، بين قوسي المجموعة { } بحيث يفصل بين كل عنصرين بعلامة فاصلة " , " :-

مثال :-

$A = \{1,5,10,15\}$

$B = \{a , b , c , d\}$

$C = \{1 , 2 , 3 .....\}$

(و هي مجموعة منتظمة تسير بنفس الشكل١ ٢ ٣ ٤وهكذا)

$A = \{1, 2 , 3,....,100\}$

(و هي مجموعة مغلقة و لكن المساحة لا تكفي لكتابة من ١ إلى ١٠٠ و سوف نستخدم النقاط للتعبير عن بعض العناصر)

طريقة القاعدة (الصفة المميزة) :-

ويتم فيها وصف المجموعة بذكر صفة يمكن بواسطتها تحديد عناصرها ، أي الصفة التي تحدد ارتباط عناصر المجموعة ،فمثلاً :

A = { x: عدد زوجي X}

B = { x : طالب بمقرر الاحصاء في الادارة X}

C= { x : طالب بنظام التعليم عن بعد x }

D = {x: عدد صحيح $-3 \le x \le 1$ x}

x= { x : $0 \le x \le 12$ عدد صحيح X}

أنواع المجموعات :-

المجموعة الخالية :-

هي المجموعة التي لا تحتوي أي عنصر ويرمز لها بالرمز ϕ (فاي) أو { } .

أمثلة :-

A= { x : عدد طبيعي زوجي و فردي x }

B = { x : دولة عربية تقع في أمريكا الشمالية X }

المجموعة المنتهية :-

المجموعة التي تكون عناصرها محدودة .

مثال :

المجموعات التالية مجموعات منتهية .

A = { 2 , 4 , 6 , 8 }

B = { 1 , 2 , 3 , …….. , 100}

C = { x , y , s, t u }

المجموعة غير المنتهية :-

المجموعة التي تكون عناصرها غير محدودة (وهي المجموعة التي لا يمكن تحديد عناصرها بشكل دقيق)

مثال :

المجموعات التالية مجموعات غير منتهية .

A = { X: عدد طبيعي فردي}

B = { 10, 20 , 30 ,……… }

المجموعة الكلية :-

هي المجموعة التي تدرس جميع المجموعات باعتبارها مجموعات جزئية و يرمز لها بالرمز U .

المجموعة الجزئية :-

تكون المجموعة A جزئية من المجموعة B إذا كانت جميع عناصر A موجودة في B و تكتب على الصورة :-

. $A \subset B$

أمثلة :-

١- إذا كانت $A = \{2 , 4 , 6\}$ و $B = \{1,2,3,4,5,6,7,8\}$ فإن $A \subset B$.

٢- المجموعة المكونة من جميع طلاب التعليم الالكتروني بجامعة الملك فيصل مجموعة جزئية من مجموعة طلاب هذه الجامعة .

تساوى المجموعات :-

تكون المجموعتان A و B متساويتان إذا كانت :-

$A \subseteq B \;,\; B \subseteq A \quad \gg\gg\gg\gg A = B$

أما المجموعتان المتكافئتان فهما المجموعتان اللتان تتساويان في عدد عناصرها وتكتب على الصورة $A \equiv B$

مثال :

أي المجموعات التالية متكافئة وأيها متساوية ؟

1- $A = \{1, 5, 7 , 9\}$, $B = \{9 , 7 ,5 ,1\}$

2- $A= \{2,5,9\}$, $B = \{a , s , d\}$

الحل

1 – $A = B$

2 – $A \equiv B$

العمليات على المجموعات :-

الاتحاد :-

اتحاد المجموعتين A و B $(A \cup B)$ هو مجموعة كل العناصر الموجودة في A أو في B أو في كليهما .

مثال :-

إذا كان $A = \{1 , 2 , 3 , 7\}$ و $B= \{2 , 4 , 6 , 8\}$ أوجد $(A \cup B)$ ؟

الحل

(AUB) =   {1 ,2 , 3, 4, 6, 7, 8}

التقاطع :-

تقاطع المجموعتين A و B (A∩B) هو مجموعة كل العناصر الموجودة في Aو في B معاً أي العناصر المشتركة بين A و B .

مثال :-

إذا كان {3 , 2 , 1 , 0 , 1- } = Aو B = {0,2,4,6,} أوجد A ∩ B

الحل

(A ∩ B) = {0 , 2 }

المكملة أو المتممة :-

يقال أن$\overline{A}$مكملة المجموعة A إذا كانت تحتوي على جميع عناصر المجموعة الكلية ∪ باستثناء عناصر A.

مثال

إذا كان {1,2,3,4,5,6,7,8,9,10 } = ∪و A={2,4,6,8,10} أوجد

الحل

$\overline{A}$ = {1,3,5,7,9}

الفرق :-

إذ كانت مجموعتان A ، B فإن A-B يسمى بالفرق وهو مجموعة كل العناصر الموجودة Aوليست في B .

مثال :-

إذا كانت A={1,2,3,x,y} و B={3,4,5,x,w} أوجد A-B

الحل

A-B = { 1 , 2 , y }

مثال شامل :

A={1,2,3,x,y} و B={3,4,5,x,w}   :- إذا كانت

و المجموعة الكلية

1- A∪B
2- A ∩ B
3- B – A
4- $\overline{A}$
5- $\overline{B}$
6- $\overline{A} \cup \overline{B}$
7- $\overline{A} \cap \overline{B}$
8- $\overline{A}$ ∪ A
9- $\overline{A}$ ∩ A

$U = \{1,2,3,4,5,w,x,y,z\}$

فأوجد :-

الحل

1- $A \cup B = \{1, 2, 3, 4, 5, x, y, w\}$

2- $A \cap B = \{3, X\}$

3- $B - A = \{4, 5, w\}$

4- $\overline{A} = \{4, 5, w, z\}$

5- $\overline{B} = \{1, 2, y, z\}$

6- $\overline{A} \cup \overline{B} = \{1, 2, 4, 5, y, w, z\}$

7- $\overline{A} \cap \overline{B} = \{z\}$

8- $\overline{A} \cup A = U$

9- $\overline{A} \cap A = \{\ \}$

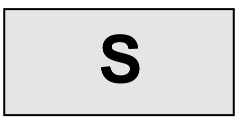
أشكال ِفنْ

VIN Figures

يمكن تمثيل المجموعات والعمليات المختلفة عليها من خلال استعمال اشكال هندسية تسمى أشكال ِفنْ وذلك وفق ما يلي:

١ـ المجموعة الكلية:

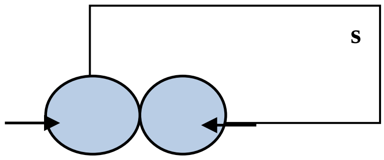
تمثل المجموعات الكلية بمستطيل ويرمز لها بالرمز S



S

٢- إتحاد الحوادث Events Union :

لأي حادثتين A و B فإن الحادثة التي تتضمن كافة العناصر التي تنتمي إلى A أو إلى B أو إلى كليهما معا يطلق عليها إتحاد حادثتين ويرمز لها (A   B) أو (A أو B) والشكل التالي يوضح ذلك:



شكل فِنْ لتمثيل إتحاد حادثتين A و B

$(A \cup B)$

وبشكل عام لأي n حادثة A1 , A2, A3, ... An فإن إتحاد هده الحوادث هو :

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup A_3 ...... \cup_n$$

ويمكن القول أن $\bigcup_{i=1}^{n} A_i$ **هو حدث يقع إذا وقع أحد هذه الحوادث Ai على الأقل وهو ما يطلق عليه جمع الأحداث**

الإتحاد :- يعني اتحاد المجموعتين A و B وهو مجموع العناصر الموجودة في كلتا المجموعتين دون تكرار العناصر.

مثال:

A={1 , 2, -6 , -7}

B={-6, -7, -11}

$(A \cup B)$ **= {1, 2, -6, -7, -11}**

خواص العمليات الجبرية لإتحاد الحوادث:

- **إدا كانت A و B و C ثلاث حوادث فإن :**

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

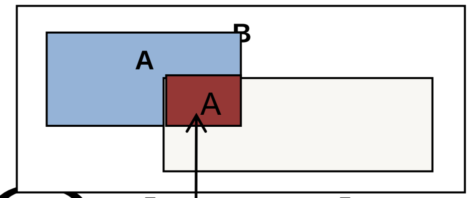
ويعني ذلك توزيع الإتحاد على التقاطع.

- وكذلك هناك خاصية التبديل والتي تعني أن

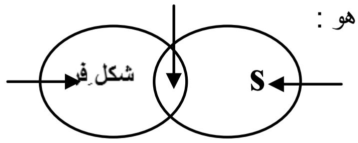
$$(A \cup B) = (B \cup A)$$

٣- تقاطع الحوادث Events Intersection :

لأي حادثتين A و B فإن الحادثة التي تتضمن كافة نقاط العينة التي تنتمي إلى A و B أو إلى كليهما معا في نفس الوقت يطلق عليها تقاطع حادثتين ويرمز لها (A B) أو (A و B) وباستخدام أشكال ِفنْ يكون الجزء المحدد بـ A and B هو الذي يمثل تقاطع الحادثتين :



وبشكل عام لأي n حادثة A1 , A2, A3, …….. An فإن تقاطع هذه الحوادث هو :



$$\bigcap_{i=1} A_i = A_1 \cap A_2$$

$$(A \cap B)$$

ويمكن القول أن $\bigcap_{i=1}^{n} A_i$ هو حدث يقع إذا وفقط وقعت كل الحوادث Ai على الأقل و هو ما يطلق عليه ضرب الحوادث

التقاطع :- إذا هو مجموعة العناصر المشتركة بين مجموعتين أو أكثر.

مثال:

A={1 , 2, -6 , -7}

B={-6, -7, -11}

$(A \cap B)$ **= {-6, -7}**

خواص العمليات الجبرية لتقاطع الحوادث:

- إذا كانت :C و B و A ثلاث حوادث فإن :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

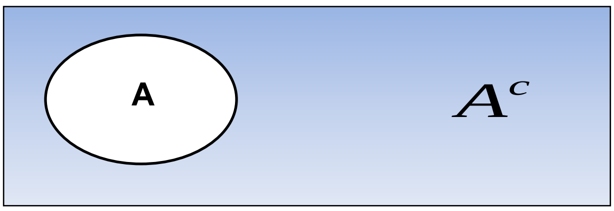
ويعني دلك توزيع التقاطع على الإتحاد

- وكذلك هناك خاصية التبديل والتي تعني $(A \cap B) = (B \cap A)$

٤- الحادثة المتممة Complementary Event

فإن متممتها هي الحادثة التي تتضمن كافة العناصر التي لا تنتمي A ويرمز لها بالرمز A^c **أو** $\bar{A}$

وهو حدث يتألف صر غير المنتمية إلى A وباستخدام أشكال ِفنْ فإن الجزء المظلل يمثل الحادثة المتممة من



$$\bar{A} \quad A^c$$

شكل فن لتمثيل مكملة حادثة A

مثال:

S={1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20}

A={1, 3, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19}

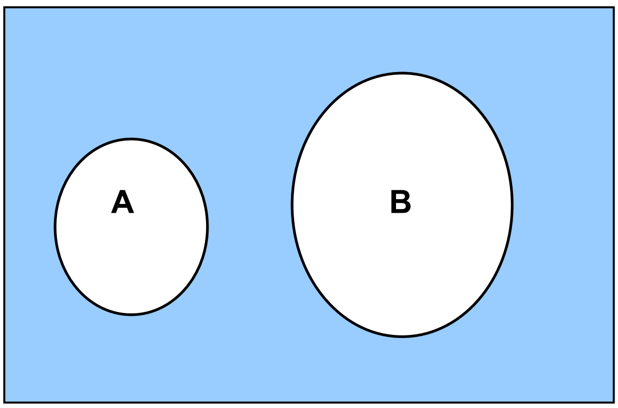
B={1, 2, 3, 6, 8, 11, 12, 14, 16}

$\overline{A}$ **={2, 4, 6, 20}**

$\overline{B}$ **={4, 5, 7, 9, 10, 13, 15, 17, 18, 19, 20}**

٥-الحوادث المتنافية Mutually Execlusive Events :

الحادثتان A و B متنافيتان أو منفصلتان إذا كان تقاطعهما خاليا أي أن $A\cap B=\phi$ **ويمكن القول أيضا أن** $A\cap A^c=\phi$



$$A\cap B=\phi$$

$$A\cap A^c=\phi$$

شكل فنْ لتمثيل حادثتان متنافيتان A و B

بعض العلاقات المهمة :-

$$\overline{B\cup A}=\overline{B}\cap\overline{A}$$

$$A\cup A^{c}=S$$

$$\overline{A\cup B}=\overline{A}\cap\overline{B}$$

$$A\cap A^{c}=\phi$$

إذا كانت $A\subset B$ **فإن:**

$$A=A\cap B$$

$$B=A\cup B$$

$$\overline{B}\subset\overline{A}$$

$$\overline{S}=\phi$$

$$\overline{\phi}=S$$

$$A\cup S=S$$

$$A\cap S=A$$

$$A\cap\phi=\phi$$

أمثلة وتمارين

و = ∩
أو = ∪

يراد شراء ثلاثة أنواع من اللحوم من جزار معين، فإذا رمزنا للحم الدجاج بـ A ولحم الضأن بـ B ، ولحم العجل بـ C فإن :

– توفر أنواع اللحوم الثلاثة يعني توفر لحم A و B و C أي بمعنى

$A\ \cap\ B\ \cap\ C$.

– عدم توفر أي نوع من اللحوم يعني عدم توفر A و B و C أو كلها أي بمعنى :

$$\overline{A}\cap\overline{B}\cap\overline{C}$$

– توفر نوع واحد من اللحوم على الأقل هو توفر A أو B أو C أو كلها أي بمعنى : $A\cup B\cup C$

– توفر نوع A فقط يعني : $A\cap\overline{B}\cap\overline{C}$

– توفر نوع واحد من اللحم يعني إما توفر A وعدم توفر النوعين الآخرين أو توفر B وعدم توفر النوعين الآخرين ، أو توفر C

$$(A\cap\overline{B}\cap\overline{C})\cup(\overline{A}\cap B\cap\overline{C})\cup(\overline{A}\cap\overline{B}\cap C)$$

تمارين :-

١- وضح أي من هذه المجموعات هي مجموعة خالية أو مجموعة منتهية أو مجموعة غير منتهية :-

(a) A ={x: عدد سالب و موجب x }

(b) B = {3 ,6 ,9 ,12 }

(c) C = {x: دولة أوربية تقع في شبة الجزيرة العربية x }

(d) D = { 2 , 4 , 6 , …….,100}

(e) E = { 100 , 200 , 300 ,…….}

(f) F = { w , e , r , t }

٢- إذا كانت A = {3 , 5 , 7 } و B = { 1,2,3,4,5,6,7,8} فهل يمكن القول أن $A \subset B$ ؟

٣- أي المجموعات التالية متكافئة وأيها متساوية ؟

1- A = {5 ,10 ,15 ,20 } , B = {15 ,10 ,5 , 20 }

2- A= { 20,50,70} , B = { k , d , u }

إذا كانت U= {0 , 1, 2, 3,…….,10} وكانت C = { 6, 7,8, 9,10} ، B = { 1 , 2 ,3 ,4 ,5 } ، A = { 2, 4 ,6 ,8 ,10}
فجد ما يلي:-

1- $A \cup B$	5- $A \cap \bar{C}$
2- $A \cap C$	6- $A - (B \cap C)$
3- $\bar{A} \cap \bar{B}$	7- $(\bar{A} \cup B) - C$
4- $B \cup C$	8- $(\overline{\bar{B} \cap \bar{C}})$

حل المثال :-

1- $A \cup B$ = { 1 , 2 , 3 ,4 , 5 ,6 , 8 , 10 }

2- $A \cap C$ = { 6 , 8 , 10 }

3- $\bar{A} \cap \bar{B} = \overline{A \cup B}$ = { 0 , 7 ,9 }

4- $B \cup C$ = { 1 , 2 , 3 , 4 , 5 , 6 , 7 , 8 , 9 ,10 }

5- $A \cap \bar{C} = A - C$ = { 2 , 4 }

6- $A - (B \cap C) = B \cap C = \emptyset$

$A - (B \cap C) = A - (\emptyset) = A$

7- $(\bar{A} \cup B) - C$

$\bar{A} = \{0, 1, 3, 5, 7, 9\}$

$\bar{A} \cup B = \{0, 1, 2, 3, 4, 5, 7, 9\}$

$(\bar{A} \cup B) - C = \{0, 1, 2, 3, 4, 5\}$

8- $(\overline{\bar{B} \cap \bar{C}}) = \overline{(\bar{B} \cup \bar{C})} = B \cup C$

$B \cup C = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

تمارين :-

٤- إذا كانت

$A=\{8,10,12,r,m\}$ و $B=\{4,6,10,o,r\}$

أوجد المجموعة الكلية

ثم أوجد :-

1- $A \cup B$
2- $A \cap B$
3- $B - A$
4- $\bar{A}$
5- $\bar{B}$
6- $\bar{A} \cup \bar{B}$
7- $\bar{A} \cap \bar{B}$
8- $\bar{A} \cup A$
9- $\bar{A} \cap A$

٥- نفترض أن $A = \{3,4,5,x,y\}$ **و** $B = \{4,x,y,z\}$ **ضع الرمز** $\in$ **أو** $\notin$ **في المكان الفارغ لتكون الجملة صحيحة .**

$(i)\ 3 \_\_\_\_\_ A$

$(ii)\ 3 \_\_\_\_\_ B$

$(iii)\ x \_\_\_\_\_ A$

$(iv)\ x \_\_\_\_\_ B$

$(v)\ z \_\_\_\_\_ A$

$(vi)\ z \_\_\_\_\_ B$

$(vii)\ 1 \_\_\_\_\_ A$

$(viii)\ 1 \_\_\_\_\_ B$

المحاضرة (٢)

نظرية الاحتمالات

تعريف الاحتمالات:

يمكن تعريف الاحتمالات بطرق عديدة غير أن أبسطها

معين"(Event)"هو مقياس لامكانية وقوع حدث

وتستعمل كلمة احتمالات بشكل دائم في حياتنا اليومية مثل:

- **احتمال أن ينجح الطالب في مقرر دراسي.**
- **احتمال نزول المطر اليوم**

وفي أحيان أخرى تستخدم كلمة احتمالات ككلمة مرادفة لبعض الكلمات الأخرى مثل: ممكن، غالبا، مؤكد، مستحيل ...

وقد ارتبط المفهوم التقليدي للاحتمال بألعاب الحظ لمدة طويلة، وتختلف درجة إمكانية تحقق أي حادثة من شخص إلى آخر حسب خبرته والمعلومات المتوفرة لديه عن الحادثة

وقد تطور علم الاحتمالات تطورا كبيرا وسريعا وأصبح أساسا لعلم الإحصاء وبحوث العمليات وغيرها.

١ - التجربة العشوائية:-

هي تلك التجربة التي تكون جميع نتائجها معلومة مسبقا ولكن لا يمكن التنبؤ بحدوث أي من هذه النتائج بصفة مؤكدة مثلا:

رمي حجر نرد مرة واحدة يعتبر تجربة عشوائية، حيث نعلم جميع قيم نتائج التجربة وهي إما ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ ولكن لا يمكن تحديد أي الأرقام يظهر إلى الأعلى بصورة مؤكدة قبل إجراء التجربة.

- **رمي عملة معدنية مرة واحدة أو عدد من المرات يعتبر تجربة عشوائية معروف جميع نتائجها قبل أن تبدأ التجربة، ولكن لا يمكن الجزم بظهور أي منها في رمية معينة.**
- **المشاركة في سباق الخيل لحصان معين يعتبر تجربة عشوائية فهو إما أن يفوز أو يخسر أو يتعادل.**

٢-الحادث والفراغ العيني:

<u>فراغ العينة</u> هو المجموعة الشاملة التي تحتوي على جميع النتائج الممكنة للتجربة العشوائية ويرمز له

Ω و يطلق عليه الحالات الممكنة

افترض أننا نقوم بإجراء تجربة ما كرمي زهرة النرد مثلاً ونلاحظ كل النتائج الممكنة وهي ظهور أحد الأوجه الستة ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ ونفترض أننا مهتمون بظهور رقم فردي أي ١ أو ٣ أو ٥ من التجربة . وهكذا فإن عملية رمي الزهرة تسمى <u>تجربة</u> (Experiment) وظهور رقم فردي هو محل اهتمامنا يسمى <u>حادثاً</u> (Event) ومجموعة جميع الحالات الممكنة الظهور تسمى <u>بالفراغ العيني</u> (Sample Space) ويلاحظ أن الحادث قد يكون حالة او اكثر من الفراغ العيني .

الحادثة هي مجموعة جزئية من فراغ العينة وتمثل مجموعة النتائج التي تحقق الحدث وتسمى أيضا الحالات المواتيةFavorable Cases ، فمثلا الحصول على رقم زوجي في تجربة رمي زهرة النرد مرة واحدة تكون الحادثة هي {٢ ، ٤ ، ٦} ، ويمكن أن تحتوي الحادثة على عنصر واحد أو أكثر.

٣- أنواع الحوادث :-

أ- الحوادث المتنافية

يقال عن الحادثين A و B أنهما متنافيان إذا استحال حدوثهما معا. فمثلاً عند رمي عملة معدنية لا يمكن الحصول على وجهين في وقت واحد .

ب-الحوادث المستقلة

يعتبر الحادثين A أو B حادثين مستقلين إذا كان وقوع إحداهما أو عدم وقوعه لا يؤثر في وقوع الآخر. فمثلاً عند رمي قطعة عملة واحدة مرتين متتاليتين فإن نتيجة الرمية الثانية لا تتأثر بنتيجة الأولى

جـ -الحوادث الشاملة

تسمى الحوادث C ، B ، A .... حوادث شاملة في تجربة ما إذا كان لابد من حدوث إحداها عند إجراء التجربة. فمثلاً عند اختيار طالب من الجامعة لمعرفة حالته ما إذا كان مدخنا أو غير مدخن تعتبر هذه الحالات حوادث شاملة لأنه لابد للفرد أن يكون له صفة واحدة من هذه الصفات. كذلك فإن الحصول على العدد ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ عند رمي حجر النرد تعتبر حوادث شاملة لأنه لابد من حدوث إحداها .

مثال

رمي حجر نرد مرد واحدة ، أحسب التالي:

- احتمال الحصول على رقم ٥
- احتمال الحصول على رقم زوجي
- احتمال الحصول على رقم أكبر من ٢
- احتمال الحصول على رقم أقل من ٧
- احتمال الحصول على رقم ٧

الحل:

فراغ العينة لهذه التجربة هو : $\Omega= \{1, 2, 3, 4, 5, 6\}$ فيكون بالتالي الحل كما يلي:

$P_{(A=5)}=\frac{1}{6}$

$P_{(A=2,4,6)}=\frac{3}{6}$

$P_{(A>2)}=\frac{4}{6}$

$P_{(A<7)}=\frac{6}{6}=1$

$P_{(A=7)}=\frac{0}{6}=0$

ويستنتج من ذلك أن أقل قيمة للاحتمال تساوي الصفر ويقال أن الحدث في هذه الحالة (حدث مستحيل) بينما تساوي واحد إذا كان الحدث مؤكد.

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الدي يعمل به:

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختر عامل من الجدول السابق بطريقة عشوائية، ثم احسب الاحتمالات التالية:

- أن يكون أعزبا
- أن يكون متزوجا
- أن يكون من القسم الأول
- أن يكون من القسم الأول أو الثاني
- أن يكون من القسم الأول وأعزب

الحل:

نفرض أن الحادثة A أن يكون العامل أعزب أي A= {أن يكون العامل أعزب} فيكون الاحتمال المطلوب

P(A)= عدد العمال العزاب / عدد العمال الكلي = 23/50= 0.46

نفرض أن الحادثة B أن يكون العامل متزوج أي أن B = {أن يكون العامل متزوج} فيكون الإحتمال المطلوب

P(B)= عدد العمال المتزوجين / عدد العمال الكلي = 27/50= 0.54

نفترض ان الحادثة C أن يكون العامل القسم الأول C أي = { أن يكون العامل من القسم الأول } فيكون الإحتمال المطلوب

P(C)= عدد عمال القسم الأول / عدد العمال الكلي = 12/50= 0.24

نتفرض ان الحادثة D ان يكون العامل القسم الاول او الثاني D أي = { ان يكون العامل من القسم الاول او الثاني { فيكون الاحتمال المطلوب

$$P(D)= \frac{\text{عدد عمال القسم الأول أو الثانى}}{\text{عدد العمال الكلي}} = (12+22)/50=34/50= 0.68$$

نفترض ان الحادثة E ان يكون العامل من القسم الاول و أعزب أي ان E = { أن يكون العامل من القسم الاول و أعزب { فيكون الاحتمال المطلوب

$$P(E)= \frac{\text{عدد عمال القسم الاول و أعزب}}{\text{عدد العمال الكلي}} = 5/50= 0.1$$

جمع الاحتمالات

أ- في حالة كون الحوادث متنافية

إذا كانت الحوادث A1 , A2 , A3 .... حوادث متنافية بمعنى أن حدوث أحدها يؤدي إلى استحالة حدوث أي من الحوادث الأخرى وبالتالي فإن احتمال حدوث هذه الحوادث معاً يكون معدوماً فإن

احتمال وقوع أي حادث من الحوادث المتنافية يساوي مجموع احتمالات وقوع هذه الحوادث

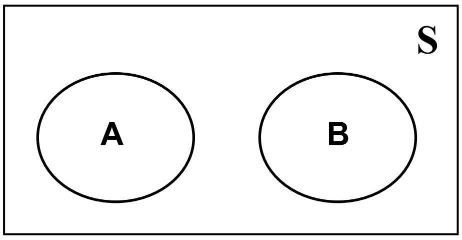
فإذا كان B ، A حادثين متنافيين كما في الشكل (١)

فإن P (A أو B) = P(A) + P(B)

يرمز أيضاً لاحتمال وقوع أحد الحادثين بالرمز

حوادث متنافية

$$P (AUB) = P (A) + P (B)$$



حوادث متنافية

مثال:

رمي حجر نرد مرة واحدة ، احسب:

احتمال الحصول على رقم ٥ أو ٦

احتمال الحصول على رقم زوجي

الحل:

حيث أن الحصول على رقم ٥ أو ٦ حادثتان متنافيتان ، أي أن:

A1={الحصول على الرقم ٥} ، و A2= {الحصول على الرقم ٦} فإن:

$P(A1 \cup A2)=(1/6)+(1/6)=1/3$

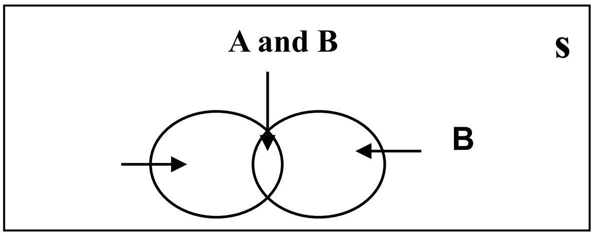
وحيث أن الحصول على رقم زوجي يعني الحصول على رقم ٢ أو رقم ٤ أو رقم ٦ وكلها حوادث متنافي، أي أن:

A1= {الحصول على الرقم ٢} ، و A2= {الحصول على الرقم ٤} ، و A3= {الحصول على الرقم ٦} فإن:

$P(A1 \cup A2 \cup A3)=(1/6)+(1/6)+(1/6) = 1/2$

ب- في حالة كون الحوادث غير متنافية

عند عدم اشتراط تنافي الحادثين A و B يكون المقصود بالحادث (A أو B) وقوع A على انفراد أو وقوع B على انفراد أو وقوع الحادثين A و B معا في وقت واحد كما يتضح من الشكل التالي



حوادث غير متنافية

الآن P (A) + P (B) تمثل مجموع الحالات المواتية للحادث A مضافاً إليها مجموع الحالات المواتية للحادث B ولكن يجب ملاحظة أن كل من الحالات المواتية للحادث A وتلك المواتية للحادث B تتضمن الحالات المواتية لوقوع A و B معاً ، وبهذا فإنه في حالة جمع P (A) و P (B) فإننا نجمع P (A و B) مرتين ، لهذا لابد من طرح P (A و B) مرة واحدة لنحصل على الاحتمال P (A أو B) وهذا هو :

$$P (A \text{ او } B) = P (A) + P (B) – P (A \text{ و } B)$$

أو

$$P (AUB) = P (A) + P(B) – P (A \cap B)$$

مثال:-

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الذي يعمل به:

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختير عامل من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- **احتمال أن يكون العامل من القسم الأول أو الثاني.**
- **احتمال أن يكون العامل متزوجا أو من القسم الأول**
- **احتمال أن يكون العامل من القسم الثالث أو أعزب**

الحل:

نفرض أن الحادثة A1 أن يكون العامل من القسم الأول أي أن A1 = {أن يكون العامل من القسم الأول}

نفرض أن الحادثة A2 أن يكون العامل من القسم الثاني أي أن A2 = {أن يكون العامل من القسم الثاني}

فيكون الاحتمال المطلوب:

$P(A1)=12/50$

$P(A2)=22/50$

$$P(A1 \cup A2) = P(A1)+P(A2) = (12/50)+(22/50) = 34/50 = 0.68$$

نفرض أن الحادثة A1 أن يكون العامل متزوجا أي أن A1 = {أن يكون العامل متزوج}

نفرض أن الحادثة A2 أن يكون العامل من القسم الأول أي أن A2 = {أن يكون العامل من القسم الأول}

فيكون الاحتمال المطلوب:

$P(A1)=27/50$

$P(A2)=12/50$

$$P(A1 \cup A2) = P(A1)+P(A2)-P(A1 \cap A2) = (27/50)+(12/50)-(7/50) = 32/50 = 0.64$$

نفرض أن الحادثة A1 أن يكون العامل من القسم الثالث أي أن A1 = {أن يكون العامل من القسم الثالث}

نفرض أن الحادثة A2 أن يكون العامل أعزب أي أن A2 = {أن يكون العامل أعزب}

فيكون الاحتمال المطلوب:

P(A1)=16/50

P(A2)=23/50

$$P(A1 \cup A2) = P(A1)+P(A2)-P(A1 \cap A2) = (16/50)+(23/50)-(10/50) = 29/50 = 0.58$$

الاحتـمال الشرطي Conditional Probability

إذا كان لدينا الحادثين A1 , A2 وكان P(A2) لا يساوي الصفر فإن الاحتمال الشرطي للحادث A1 بشرط وقوع الحادث A2 يعطي بالمعادلة التالية:

$$P(A1 | A2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

أي أن الاحتمال الشرطي للحادث A1 بشرط وقوع الحادث A2 يساوي حاصل قسمة الاحتمال المركب لـ A1 , A2 على احتمال الحادث A2

مثال:

إذا كان احتمال نجاح طالب في مقرر الرياضيات 0.64 واحتمال نجاحه في مقرر الإحصاء ومقرر الرياضيات معا 0.32 فما هو احتمال نجاحه في مقرر الإحصاء؟ علما بأنه نجح في مقرر الرياضيات.

الحل:

نفرض أن A1={نجاح الطالب في مقرر الإحصاء}

A2={نجاح الطالب في مقرر الرياضيات}

وبذلك يكون:

$$P(A2)=0.64$$

$$P(A1 \cap A2)=0.32$$

ويكون المطلوب في هذه المسألة هو حساب P (A1 | A2) وبتطبيق العلاقة :

$$P(A1 | A2) = \frac{P(A_1 \cap A_2)}{P(A_2)} = \frac{0.32}{0.64} = 0.5$$

إذا احتمال نجاح الطالب في مقرر الإحصاء علما بأنه نجح في مقرر الرياضيات هو 0.5

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الدي يعمل به:

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختير عامل من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- **احسب احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج؟**
- **احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث؟**

الحل :

نفرض أن A1={أن يكون العامل من القسم الأول}

A2={أن يكون العامل متزوج}

B3={أن يكون العامل من القسم الثالث}

B4={أن يكون العامل أعزب}

فيكون بالتالي:

١- احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج هو:

$$\frac{\text{احتمال أن يكون من القسم الأول ومتزوج}}{\text{احتمال أن يكون متزوج}}$$

P (A1|A2) = $$\frac{P(A_1 \cap A_2)}{P(A_2)} = \frac{\frac{7}{50}}{\frac{27}{50}} = \frac{7}{27}$$

إذا احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج هو 0.259

٢- احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث هو:

احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث
―――――――――――――――――――
احتمال أن يكون من القسم الثالث

P (B1|B2) = $\frac{P(B_1 \cap B_2)}{P(B_2)} = \frac{\frac{10}{50}}{\frac{16}{50}} = \frac{10}{16}$

إذا احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث هو 0.625

ضـــرب الاحتـمالات

إن احتمال حدوث حادثين مستقلين أو أكثر معاً يساوي حاصل ضرب احتمال حدوث كل واحد من هذه الحوادث

فمثلا إذا كان لدينا صندوق به ١٠ كرات متماثلة منها ٦ بيضاء و ٤ سوداء وسحبنا كرة من الصندوق فإن احتمال أن تكون بيضاء 6/10 واحتمال أن تكون سوداء 4/10، فإذا أعدنا الكرة إلى الصندوق (ليصبح العدد مكتملا كما كان) وسحبنا الكرة مرة أخرى فإن احتمال أن تكون بيضاء 6/10 واحتمال أن تكون سوداء 4/10 ، فتكرار العملية يؤدي إلى نفس الاحتمال.
ومن هذا نرى أن نتيجة السحب الأول لا تؤثر على نتيجة السحب الثاني وهذا ما يسمى سحب بإرجاع أو إحلال أو إعادة.

فإذا كان لدينا الحادثين المستقلين A1 و A2 فإن احتمال حدوثهما معا هو:

P(A1∩A2) = P(A1) P(A2)

بمعنى أن احتمال وقوع حدثين مستقلين معا يساوي حاصل ضرب احتمال وقوع أي منهما بمفرده في احتمال وقوع الحدث الآخر بمفرده وفي حالة التعميم لـ n فإن :

P(A1 ∩ A2 ∩ A3 ∩ .... An) P (A1) P (A2) P (A3) ...... P(An)

مثال:

إذا رمينا قطعة نقود مرة واحدة ، احسب الاحتمالات التالية:

- أن تكون الأولى صورة والثانية كتابة.
- أن تكون كلتاهما صورة.

الحل:

نفرض أن: A1={ظهور صورة في الرمية الأولى}

A2={ظهور كتابة في الرمية الثانية}

فيكون : $P(A_1) = \frac{1}{2} \quad , P(A_2) = \frac{1}{2}$

حيث أن الحادثتان A1 و A2 مستقلتين فإن احتمال أن تكون الرمية الأولى صورة والثانية كتابة هو :

$$P(A_1 \cap A_2) = P(A_1) \times (P(A_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

نفرض أن: B1={ظهور صورة في الرمية الأولى}

B2={ظهور صورة في الرمية الثانية}

فيكون

$$P(B_1) = \frac{1}{2} \quad , P(B_2) = \frac{1}{2}$$

وحيث أن الحادثتان B1 و B2 مستقلتين فإن احتمال أن تكون الرمية الأولى صورة والثانية صورة هو :

$$P(B_1 \cap B_2) = P(B_1) \times (P(B_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

مما سبق يمكن القول أن الحوادث المعطاة تكون مستقلة عندما تبقى الاحتمالات ثابتة مثل الحوادث:

١- رمي قطع نقود (أو قطعة واحدة عدة مرات)

٢- رمي أحجار نرد (أو حجر نرد عدة مرات)

٣- السحب مع الإرجاع (أو الإعادة)

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الذي يعمل به فإذا سحب عاملان من المصنع مع الإرجاع (أي إرجاع العامل الأول قبل سحب العامل الثاني) احسب :

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16

المجموع	23	27	50

اختير عاملان من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

1. **احتمال أن يكون العاملان من القسم الأول؟**
2. **احتمال أن يكون العاملان متزوجان؟**
3. **احتمال أن يكون للعاملين نفس الحالة الاجتماعية؟**
4. **احتمال أن يكون العاملان من القسم نفسه؟**

الحل:

١- احتمال أن يكون العاملان من القسم الأول يعني أن يكون العامل الأول من القسم الأول (الحادثة A1) والعامل الثاني من القسم الأول (الحادثة A2) وحيث أنهما مستقلان (لأن السحب مع الإرجاع) فإن:

$$P(A_1 \cap A_2) = P(A_1) \times (P(A_2) = \frac{12}{50} \times \frac{12}{50} = \frac{144}{2500} = 0.0576$$

٢- احتمال أن يكون العاملان متزوجان ، يعني أن يكون العامل الأول متزوج (الحادثة B1) والعامل الثاني متزوج (الحادثة B2) وحيث أنهما مستقلان (لأن السحب مع الإرجاع) فإن:

$$P(B_1 \cap B_2) = P(B_1) \times (P(B_2) = \frac{27}{50} \times \frac{27}{50} = \frac{729}{2500} = 0.2916$$

٣- احتمال أن يكون للعاملين نفس الحالة الاجتماعية يعني أن يكون العاملان كلاهما متزوجين (الحادثة A) أو أن يكون كلاهما أعزبين (الحادثة B) فإن:

$$\begin{aligned} P(A \cup B) &= P(A) + (P(B) \\ &= P(A_1A_2) + P(B_1B_2) \\ &= P(A_1 \times A_2) + P(B_1 \times B_2) \\ &= \left[\frac{27}{50} \times \frac{27}{50}\right] + \left[\frac{23}{50} \times \frac{23}{50}\right] = 0.5032 \end{aligned}$$

٤- احتمال أن يكون العاملان من القسم نفسه يعني أن يكون العاملان كلاهما من القسم الأول (الحادثة A) أو أن يكون كلاهما من القسم الثاني (الحادثة B) أو أن يكون كلاهما من القسم الثالث (الحادثة C) فإن:

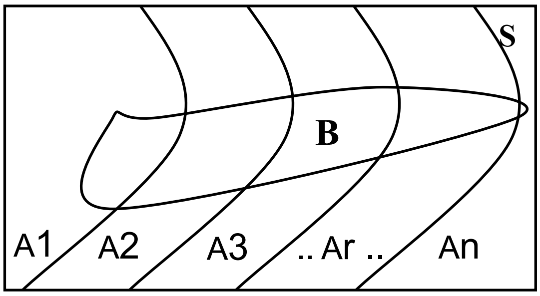
$$\begin{aligned} P(A \cup B \cup C) &= P(A) + (P(B) + (P(C) \\ &= P(A_1A_2) + P(B_1B_2) + P(C_1C_2) \\ &= P(A_1 \times A_2) + P(B_1 \times B_2) + P(C_1 \times C_2) \\ &= \left[\frac{12}{50} \times \frac{12}{50}\right] + \left[\frac{22}{50} \times \frac{22}{50}\right] + \left[\frac{16}{50} \times \frac{16}{50}\right] = 0.3536 \end{aligned}$$

نظريـــة بايــز (Bayes' Theorem)

إذا كانت An .... ,A1, A2 مجموعة أحداث متنافية وكانت احتمالات حدوثها

P(An) .... ,P(A1), P(A2) وإذا كان هناك حدث B يحدث إذا حدث أي من الأحداث المتنافية أنظر الشكل بالأسفل ، فإن احتمال حدوث الحدث Ar بشرط حدوث B هو :

$$P(A_r|B)=\frac{P(A_r)P(B|A_r)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)} \qquad 1\le r\le n$$



مثال:-

مصنع يقوم بإنتاج سلعة معينة به ثلاث آلات، تنتج الآلة الأولى ٢٠% من إجمالي إنتاج السلعة وتنتج الآلة الثانية نسبة ٣٥% والثالثة بنسبة ٤٥ % ، فإذا كانت نسبة الإنتاج المعيب في الثلاث آلات على الترتيب هو ٢% و ٢.٥% و ٣% ، سحبت وحدة عشوائيا من إنتاج المصنع فوجد أنها معيبة، احسب الاحتمالات التالية:

١- أن تكون القطعة المعيبة من إنتاج الآلة الأولى؟

٢- أن تكون القطعة المعيبة من إنتاج الآلة الثانية؟

الحل:

نفرض أن

A1={إنتاج الآلة الأولى} P(A1)=0.2

A2={إنتاج الآلة الثانية} P(A2)=0.35

A3={إنتاج الآلة الثالثة} P(A3)=0.45

B= {إنتاج سلعة معينة}

فيكون بالتالي:

P(B|A1)=0.02

P(B|A2)=0.025

P(B|A3)=0.03

تكون السلعة من إنتاج الآلة الأولى إذا علم - بشرط - أنها معيبة هو:

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}=\frac{0.2\times0.02}{(0.2\times0.02)+(0.35\times0.025)+(0.45\times0.03)}=0.152$$

واحتمال أن تكون السلعة من إنتاج الآلة الثانية إذا علم - بشرط - أنها معيبة هو:

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}=\frac{0.35\times0.025}{(0.2\times0.02)+(0.35\times0.025)+(0.45\times0.03)}=0.333$$

مثال:-

مستشفى به أربعة أقسام، نسب عمال النظافة في هذه الأقسام هي ٣٠% ، ٤٠% ، ٢٠% ، ١٠% على التوالي، إذا كانت نسب العمال المدخنين بهذه الأقسام هي ١٥% ، ١٨% ، ١٢% ، ٩% على التوالي، اختير عامل عشوائيا فوجد أنه مدخن ، احسب الاحتمالات التالية:

١- أن يكون العامل من القسم الأول؟

٢- أن يكون العامل من القسم الثاني؟

٣- أن لا يكون العامل من القسم الأول؟

الحل:

A1={أن يكون العامل من القسم الأول} P(B|A1)=0.15 P(A1)=0.3

A2={أن يكون العامل من القسم الثاني} P(B|A2)=0.18 P(A2)=0.4

A3={أن يكون العامل من القسم الثالث} P(B|A3)=0.12 P(A3)=0.2

A4={أن يكون العامل من القسم الرابع} P(B|A4)=0.09 P(A4)=0.1

احتمال أن يكون العامل من القسم الأول إذا علم – بشرط – أنه مدخن :

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{4}P(A_i)P(B|A_i)}=\frac{0.3\times0.15}{(0.3\times0.15)+(0.4\times0.18)+(0.2\times0.12)+(0.1\times0.09)}=0.3$$

واحتمال أن يكون العامل من القسم الثاني إذا علم – بشرط – أنه مدخن :

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{4}P(A_i)P(B|A_i)}=\frac{0.4\times0.18}{(0.3\times0.15)+(0.4\times0.18)+(0.2\times0.12)+(0.1\times0.09)}=0.48$$

واحتمال أن لا يكون العامل من القسم الأول إذا علم – بشرط – أنه مدخن :

$$P(A_1^c|B)=1-0.3=0.7$$

تمارين واجب :-

١- عرف المصطلحات التالية :-

(التجربة العشوائية – فراغ العينة – الحادث – الحوادث المتنافية – الحوادث المستقلة – الحوادث الشاملة) .

٢- الجدول التالي يمثل توزيع موظفي أحد الشركات حسب الحالة الاجتماعية للموظف والمستوى الاداري الذي يعمل به

الحالة الاجتماعية	أعزب	متزوج	المجموع
مستوى الادارة الدنيا	10	14	24
مستوى الادارة المتوسطة	16	28	44
مستوى الادارة العليا	20	12	32
المجموع	46	54	100

أولاً :-اختر موظف من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- أن يكون أعزبا.
- أن يكون متزوجا .
- أن يكون من مستوى الادارة الدنيا.
- أن يكون من مستوى الادارة الدنيا أو المتوسطة .
- أن يكون من مستوى الادارة الدنيا وأعزب .

ثانياً : اختير موظف من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- احسب احتمال أن يكون موظفي الادارة الدنيا بشرط أنه متزوج؟
- احتمال أن يكون الموظف أعزب بشرط أنه من موظفي الادارة العليا ؟

ثالثاً : تم اختيار ٢ موظف من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- احتمال أن يكون الموظفين من الادارة الدنيا ؟
- احتمال أن يكون الموظفين متزوجان؟
- احتمال أن يكون للموظفين نفس الحالة الاجتماعية؟
- احتمال أن يكون الموظفين من القسم نفسه؟

٣- مصنع يقوم بإنتاج سلعة معينة به ثلاث آلات، تنتج الآلة الأولى ٤٠% من إجمالي إنتاج السلعة وتنتج الآلة الثانية نسبة ٢٥% والباقي من إنتاج الالة الثالثة، فإذا كانت نسبة الإنتاج المعيب في الثلاث آلات على الترتيب هو ٤% و ٣% و ٥.٥% ، سحبت وحدة عشوائيا من إنتاج المصنع فوجد أنها جيدة ، احسب الاحتمالات التالية:

- أن تكون القطعة الجيدة من إنتاج الآلة الأولى؟
- أن تكون القطعة الجيدة من إنتاج الآلة الثانية؟

المحاضرة (٣)

المتغيرات العشوائية

المتغير العشوائي Random Variable:

المتغير العشوائي هو الذي يأخذ قيما حقيقية مختلفة تعبر عن نتائج فراغ العينة، ومن ثم مجال هذا المتغير، يشمل كل القيم الممكنة له، ويكون لكل قيمة من القيم التي يأخذها المتغير احتمال معين .

وينقسم المتغير العشوائى إلى قسمين هما:

- المتغيرات العشوائية المنفصلة Discrete Random Variables
- المتغيرات العشوائية المتصلة(المستمرة) Continuous Random Variables

أولاً : المتغيرات العشوائية المتقطعة (المنفصلة):

المتغير العشوائي المنفصل هو الذي يأخذ قيم بينية، ومتباعدة، ويرمز للمتغير العشوائي بشكل عام بحرف من الحروف الأبجدية الكبيرة....X, Y, Z, ويرمز للقيم التي يأخذها المتغير بالحروف الأبجدية الصغيرة، x, y, z, ...

فالمتغير العشوائى المنفصل هو كل قيمة من قيم المتغير العشوائى ، ومن أمثلة هذه المتغيرات:-

- عدد الأولاد الذكور في الأسرة المكونة من أربع أولاد X، X:{x=0,1,2,3,4}.
- عدد العملاء الذين يتم إنهاء خدمتهم البنكية كل 10 دقائق Y، Y:{y=0,1,2,3,....}.
- عدد مرات استخدام نوع معين من الأسمدة خلال الدورة الزراعية.
- عدد الوحدات التالفة من إنتاج مزرعة معينة تنتج 200 وحدة كل موسم.
- عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر.

- **التوزيع الاحتمالي للمتغير العشوائى المنفصل:**

- التوزيع الاحتمالي، هو الذي يبين احتمالات حدوث القيم التي يمكن يأخذها المتغير، والتي ترتبط باحتمالات النتائج الممكنة في فراغ العينة، وبمعنى آخر هو التكرار النسبي للقيم التي يمكن أن يأخذها المتغير.

- فإذا كان المتغير العشوائي المنفصل X يأخذ القيم، $X : \{x = x_1, x_2, ..., x_n\}$

، وكان $P(X = x_i) = f(x_i)$ هو احتمال أن المتغير العشوائي يأخذ القيمة x_i ، فإنه، يمكن تكوين جدول التوزيع الاحتمالي للمتغير العشوائي X ، وهو جدول مكون من عمودين، الأول به القيم الممكنة للمتغير $X : \{x = x_1, x_2, ..., x_n\}$ ، والثاني به القيم الاحتمالية لهذا المتغير ، أي أن: $P(X = x_i) = f(x_i)$

جدول التوزيع الاحتمالي للمتغير العشوائي المنفصل

x_i	$f(x_i)$
x_1	$f(x_1)$
x_2	$f(x_2)$
$\vdots$	$\vdots$
x_n	$f(x_n)$
Σ	1

وتسمى الدالة $f(x_i)$ بدالة الاحتمال

مثــال :

إذا كانت نسبة مبيعات أحد المراكز التجارية من التفاح الأمريكي 0.60، بينما يكون نسبة مبيعاته من الأنواع الأخرى للتفاح 0.40، اشترى أحد العملاء عبوتين .

والمطلوب:

- **كون فراغ العينة.**

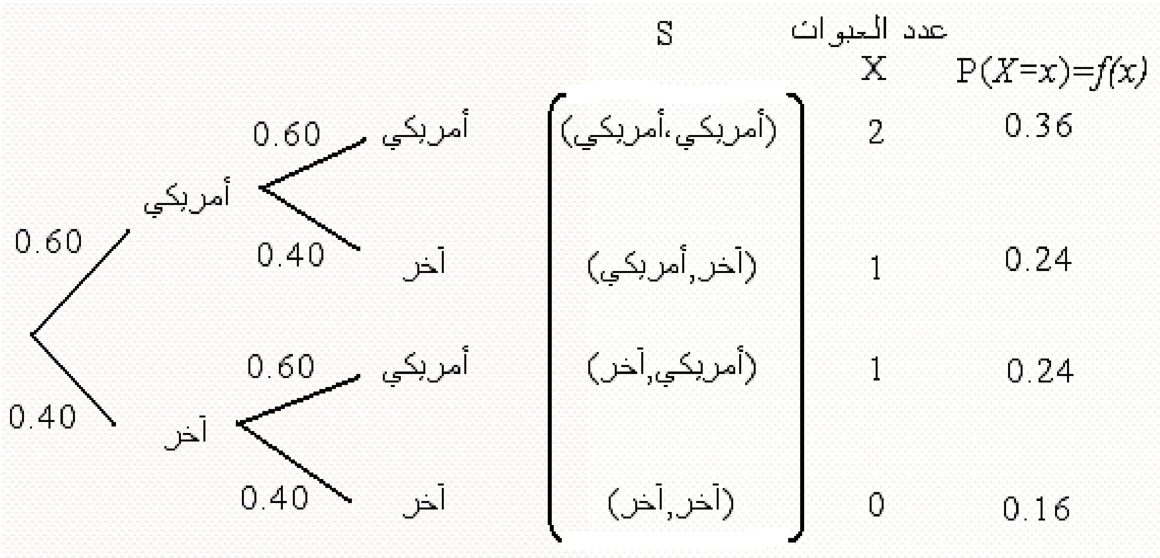
إذا عرف المتغير العشوائي بأنه عدد العبوات المشتراة من التفاح الأمريكي، فأوجد الآتي:

- **التوزيع الاحتمالي للمتغير العشوائي .**
- **ارسم دالة الاحتمال لهذا المتغير.**

الحـل:

تكوين فراغ العينة:

التجربة هنا هو شراء وحدتين من عبوات التفاح، ومن ثم فراغ العينة يتكون من أربع نتائج، هي:



تابع الحل

التوزيع الاحتمالي لعدد العبوات المشتراة من التفاح الأمريكي

من المعلوم أن العميل اشترى عبوتين، وأن المتغير العشوائي هو عدد العبوات المشتراة من التفاح الأمريكي، لذا تكون القيم الممكنة للمتغير العشوائي هي:

$x=0$ **إذا كانت العبوتين من النوع الآخر، أى إذا كانت نتيجة التجربة (آخر، آخر)**

$x=1$ **إذا كان أحد العبوتين من النوع الأمريكي، أي إذا كانت نتيجة التجربة (آخر، أمريكي) أو (أمريكي، آخر)**

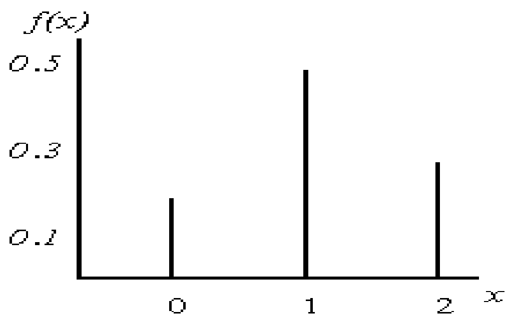
$x=2$ **إذا كان العبوتين من النوع الأمريكي، أي إذا كانت نتيجة التجربة (أمريكي ، أمريكي)**

ومن ثم يأخذ المتغير القيم: $X:\{x=0,1,2\}$ ، ويرتبط احتمالات هذه القيم باحتمالات نتائج التجربة المناظرة لها كما هو مبين أعلاه، ومن ثم يكون التوزيع الاحتمالي للمتغير العشوائي X هو:

جدول التوزيع الاحتمالي لعدد العبوات المشتراة من التفاح الأمريكي

x_i	$f(x_i)$
0	0.4×0.4 = 0.16
1	0.4×0.6+0.6×0.4= 0.48
2	0.6×0.6=0.36
$\sum$	1

رسم دالة الاحتمال $f(x)$:



الوسط الحسابي والتباين للمتغير العشوائي المنفصل:

يرمز للوسط الحسابي للمتغير العشوائي بالرمز (ميو)، ويحسب بتطبيق المعادلة التالية:

$$\boxed{\mu = \sum x_i \, f(x_i)} \qquad (٣-٨)$$

وأما التباين ويرمز له بالرمز σ^2 (سيجما)٢، فيحسب بتطبيق المعادلة التالية:

$$\boxed{\begin{aligned}\sigma^2 &= \sum (x_i - \mu)^2 \ f(x_i) \\ &= \sum x_i^2 \ f(x_i) - \mu^2\end{aligned}} \qquad (٤-٨)$$

مثـال:

في المثال السابق احسب ما يلي:

- الوسط الحسابي لعدد العبوات المشتراة من النوع الأمريكي .
- احسب الانحراف المعياري لعدد العبوات المشتراة من النوع الأمريكي .
- أوجد معامل الاختلاف النسبي .

الحـل:

- الوسط الحسابي لعدد العبوات من النوع الأمريكي:
- لحساب الوسط الحسابي والانحراف المعياري يتم استخدام المعادلة الخاصة بذلك وهذا يتطلب تكوين جدول يشمل المجاميع التالية: $\sum x_i \ f(x_i)$, $\sum x_i^2 \ f(x_i)$ وذلك كما يلي:

x_i	$f(x_i)$	$x_i f(x_i)$	$x_i^2 f(x_i)$
0	0.16	0	0
1	0.48	0.48	0.48
2	0.36	0.72	1.44
$\sum$	1	1.20	1.92

إذا الوسط الحسابي هو: $\mu = \sum x_i \ f(x_i) = 1.20$

ولحساب الانحراف المعياري يجب أولا حساب التباين وهو: $\sigma^2 = \sum x_i^2 \ f(x_i) - \mu^2 = 1.92 - (1.20)^2 = 0.48$

إذا الانحراف المعياري قيمته هي: $\sigma = \sqrt{\sigma^2} = \sqrt{0.48} = 0.693$

معامل الاختلاف النسبي هو:

$$C.V = \frac{\sigma}{\mu} \times 100 = \frac{0.693}{1.2} \times 100 = 57.7$$

ثانياً : المتغيرات العشوائية المستمرة Continuous Random Variables

المتغير العشوائي المستمر هو الذي يأخذ قيما متصلة، ويأخذ عدد لانهائي من القيم الممكنة له داخل مجاله، فإذا كان x متغير عشوائي مستمر، ويقع في المدى (a,b) أي $\{X = x : a < x < b\}$ فإن للمتغير X عدد لانهائي من القيم تقع بين الحدين الأدنى والأعلى (a,b)، ومن الأمثلة على المتغيرات الكمية المستمرة ما يلي:

- كمية الألبان التي تنتجها البقرة في اليوم باللتر: $\{X = x : 10 < x < 40\}$
- المساحة المزروعة بالأعلاف في المملكة بالألف هكتار : $\{X = x : 1000 < x < 15000\}$
- فترة صلاحية حفظ الدجاج المبرد بالأيام : $\{X = x : 1 < x < 5\}$
- وزن الجسم بالكيلوجرام للأعمار من *(40-30)*: $\{X = x : 55 < x < 80\}$
- وهكذا الأمثلة على المتغير الكمي المستمر كثيرة .

الوسط الحسابي والتباين للمتغير العشوائي المستمر:

إذا كانت $f(x)$ هي دالة كثافة الاحتمال للمتغير العشوائي $a < x < b$ فإن معادلة الوسط والتباين يمكن كتابتها كما يلي:

$$\boxed{\begin{array}{l} \mu = E(x) = \int_a^b x f(x) dx \\ \sigma^2 = E(x^2) - u^2 \ , \ E(x^2) = \int_a^b x^2 f(x) \ dx \end{array}} \qquad (١٣-٨)$$

التوزيعات الاحتمالية للمتغيرات المنفصلة:-

أ- توزيع ذي الحدين:

يستخدم هذا التوزيع في الحالات التي يكون للظاهرة محل الدراسة نتيجتان فقط متنافيتان، النتيجة محل الاهتمام وتسمى بحالة النجاح، والأخرى تسمى بحالة الفشل، ومن أمثلة ذلك:

- عند إعطاء مريض نوع معين من الأدوية، لها نتيجتان:

(استجابة للدواء، أو عدم استجابة)

- عند فحص عبوة بداخلها نوع معين من الفاكهة، لها نتيجتان

(الوحدة إما أن تكون سليمة، أو تكون معيبة)

- عند إلقاء قطعة عملة، لها نتيجتان

(ظهور الوجه الذي يحمل الصورة، أو الوجه الذي يحمل الكتابة)

- نتيجة الطالب في الاختبار

(نجاح، رسوب)

- استخدام المزارع لبرنامج معين في الزراعة

(يستخدم، أو لا يستخدم)

شكل التوزيع الاحتمالي ثنائي الحدين:

إذا كررت محاولة من المرات، بحيث أن كل محاولة لها نتيجتان فقط متنافيتان هما:

- النتيجة محل الاهتمام " حالة نجاح " وتتم باحتمال ثابت في كل محاولة هو p
- النتيجة الأخرى " حالة فشل " وتتم باحتمال ثابت أيضا هو $q = 1 - p$

وبافتراض أن هذه المحاولات مستقلة، بمعنى أن نتيجة كل محاولة ليس لها علاقة بنتيجة المحاولة الأخرى، وإذا كان المتغير العشوائي X يعبر عن عدد حالات النجاح "عدد النتائج محل الاهتمام" في الـ n محاولة، فإن مدي المتغير العشوائي X والذي يعبر عن عدد حالات النجاح هو: $X : \{x = 0,1,2,\ldots,n\}$

إذا فتوزيع ذو الحدين هو أحد التوزيعات الاحتمالية المنفصلة، ويستخدم لإيجاد احتمال وقوع حدث معين (نجاح) عدداً من المرات مقداره X من بين n من المحاولات لنفس التجربة (ونرمز لهذا الاحتمال بالرمز P(X) وذلك عندما تتحقق الشروط التالية :

- هناك ناتجان ممكنان فقط ومتنافيان لكل محاولة .
- المحاولات وعددها n مستقلة عن بعضها البعض .
- احتمال وقوع الحدث المعين في كل محاولة (النجاح) P ثابت ولا يتغير من محاولة لأخرى .

فبالتالي يمكن حساب الاحتمال من خلال المعادلة التالية :

$$P(X) = \frac{n!}{X!(n-X)!} P^{x}(1-P)^{n-X}$$

حيث n ! (وتقرأ " مضروب n ") = n. (n-1). (n-2)...3.2.1 ,0! =1.

ويكون متوسط توزيع ذي الحدين $\mu = np$

وانحراف المعياري $\sigma = \sqrt{np(1-p)}$

تحديد شكل التوزيع:

يتحدد شكل التوزيع ثنائي الحدين وفقا لقيمة احتمال النجاح كما يلي:

- إذا كان $p = 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون متماثل.
- إذا كان $p < 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون موجب الالتواء.
- إذا كان $p > 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون سالب الالتواء.

مثـال:

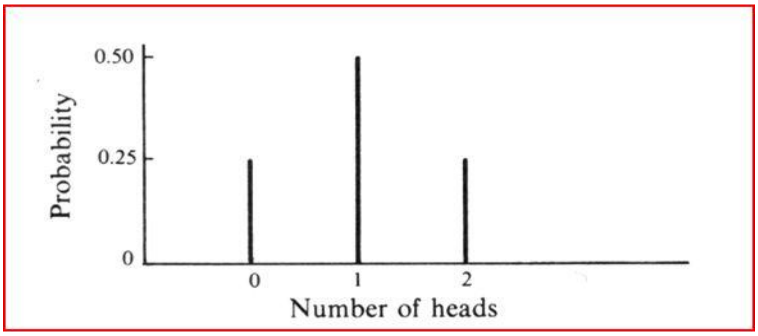
عند رمى عملة متوازنة مرتين فإن النواتج الممكنة هي TT,TH,HT,HH وعلى ذلك فإن :

$P(2H) = \frac{1}{4}$ and $P(1H) = \frac{1}{2}$ $P(0H) = \frac{1}{4}$

وهكذا فإن عدد الصور متغير عشوائي منفصل ، وتمثل مجموعة كل النواتج الممكنة مع احتمالاتها المناظرة توزيعاً احتمالياً منفصلا، أنظر الجدول التالي:

عدد الصور	إمكانية حدوثها	الاحتمال
0	TT	0.25
1	TH, HT	0.50
2	HH	0.25

ويمكن كذلك تمثيل ذلك من خلال الرسم التالي:



مثال :-

إذا كان إحتمال نجاح الطالب في مقرر التحليل الإحصائي ٨٠% تم إختيار ٤ طلاب المطلوب :-

١. كون جدول توزيع ثنائي الحدين .
٢. أوجد إحتمال نجاح ٣ طلاب .
٣. أوجد إحتمال رسوب ٣ طلاب .
٤. أوجد إحتمال نجاح طالبين على الأقل .
٥. القيمة المتوقعة (الوسط الحسابي) .
٦. الانحراف المعياري .

الحل :-

P = 0.80 , (1-P= 0.20) , n=4

١- جدول توزيع ثنائي الحدين :-

الناتج	الاحتمال	عدد الطلاب الراسبين	عدد الطلاب الناجحين
0.0016	$= 4C0 \times (0.80)^0 \times (0.20)^4$	4	0
0.0256	$= 4C1 \times (0.80)^1 \times (0.20)^3$	3	1
0.1536	$= 4C2 \times (0.80)^2 \times (0.20)^2$	2	2
0.4096	$= 4C3 \times (0.80)^3 \times (0.20)^1$	1	3
0.4096	$= 4C4 \times (0.80)^4 \times (0.20)^0$	0	4

٢- أجد إحتمال نجاح ٣ طلاب :-

P(3) = 0.4096

٣- أوجد إحتمال رسوب ٣ طلاب :-

P(1) = 0.0256

٤- أوجد إحتمال نجاح طالبين على الأقل :-

P(2)+P(3)+P(4) = 0.9728

٥- القيمة المتوقعة (الوسط الحسابي) =

$\mu = n \times p = 4 \times 0.80 = 3.2$

٦- الانحراف المعياري =

$$\sigma = \sqrt{n \times p \times (1-p)} = \sqrt{4 \times 0.8 \times 0.2} = 0.8$$

مثال :-

إذا كان إحتمال حياة شخص عند العمر ٣٠ هو ٦٠% تم إختيار ٥ أشخاص عند تمام العمر ٣٠ المطلوب :-

1. **كون جدول توزيع ثنائي الحدين .**
2. **أوجد إحتمال حياة ٤ أشخاص .**
3. **أوجد احتمال وفاة ٣ أشخاص .**
4. **أوجد احتمال حياة ٣ أشخاص على الاقل .**
5. **القيمة المتوقعة (الوسط الحسابي) .**
6. **الانحراف المعياري .**

الحل :-

$P = 0.60$, $(1-P= 0.40)$, $n=5$

<u>١- جدول توزيع ثنائي الحدين :-</u>

الناتج	الاحتمال	عدد الوفيات	عدد الاحياء
0.01024	$= 5C0 \times (0.60)^0 \times (0.40)^5$	5	0
0.0768	$= 5C1 \times (0.60)^1 \times (0.40)^4$	4	1
0.2304	$= 5C2 \times (0.60)^2 \times (0.40)^3$	3	2
0.3456	$= 5C3 \times (0.60)^3 \times (0.40)^2$	2	3
0.2592	$= 5C4 \times (0.60)^4 \times (0.40)^1$	1	4
0.07776	$= 5C4 \times (0.60)^5 \times (0.40)^0$	0	5

٢- أوجد احتمال حياة ٤ أشخاص :-

$P(4) = 0.2592$

٣- أوجد احتمال وفاة ٣ أشخاص :-

$P(2) = 0.2304$

٤- أوجد احتمال حياة ٣ أشخاص على الاقل :-

$P= (p(3) + p(4) + p(5)) = 0.07776+0.2592+0.3456 = 0.68256$

٥- القيمة المتوقعة (الوسط الحسابي) :-

$\mu = n \times p = 5 \times 0.60 = 3$

٦- الانحراف المعياري =

$$\sigma = \sqrt{n \times p \times (1-p)} = \sqrt{5 \times 0.6 \times 0.4} = 1.095445$$

المحاضرة (٤)

تابع التوزيعات الإحتمالية

تابع توزيع ثنائي الحدين :-

مثال :-

إذا كان احتمال إصابة الهدف لشخص ما هو $\frac{1}{5}$ أتيحت له فرصة الرماية في 10 محاولات

1- ما هو احتمال إصابة الهدف مرتين على الأكثر

2- احتمال إصابة الهدف مرة واحدة

الحل

X متغير عشوائي يمثل عدد مرات النجاح فى إصابة الهدف فى 10 محاولات

n = 10, p = 1/5, q= 4/5 ; x = 0, 1, 2, …, 10

1- احتمال إصابة الهدف مرتين على الأكثر:-

أي احتمال x = 0 or x = 1 or x = 2

$$P(X \le 2) = p(X=0) + p(X=1) + p(X=2)$$

$$= \binom{10}{0}(1/5)^0(4/5)^{10} + \binom{10}{1}(1/5)^1(4/5)^9$$

$$+ \binom{10}{2}(1/5)^2(4/5)^8$$

$$= (0.8)^{10} + 2(0.8)^9 + 1.8(0.8)^8 = 0.6778$$

2- احتمال إصابة الهدف مرة واحدة

أي احتمال x = 1

$$P(X=1) = \binom{10}{1}(1/5)^1(4/5)^9$$

مثال :-

ألقيت عملة ثلاث مرات. فإذا كان X يمثل عدد ظهور الصور فأوجد التوزيع الاحتمالي وكذلك التوقع والتباين

الحل

احتمال النجاح (ظهور صورة) p = 0.5

احتمال الفشل (ظهور كتابة) q = 0.5

عدد الرميات المستقلة n = 3

X متغير عشوائي يمثل عدد الصور يأخذ القيم 0, 1, 2, 3

ويكون له توزيع ذي الحدين:

$$p(X=x)=\binom{3}{x}(0.5)^x(0.5)^{3-x}\ ,\quad x=0,1,2,3$$

$$p(X=0)=\binom{3}{0}(0.5)^0(0.5)^{3-0}=\frac{3!}{0!(3-0)!}(1)(1/8)$$
$$=\frac{1}{8}$$

$$p(X=1)=\binom{3}{1}(0.5)^1(0.5)^{3-1}=\frac{3!}{1!(3-1)!}(1/2)(1/2)^2$$
$$=\frac{3\times2\times1}{1\times2\times1}\times\frac{1}{2}\times\frac{1}{4}=\frac{3}{8}$$

$$p(X=2)=\binom{3}{2}(0.5)^2(0.5)^{3-2}=\frac{3!}{2!(3-2)!}(1/2)^2(1/2)^1$$
$$=\frac{3\times2\times1}{2\times1\times1}\times\frac{1}{2}\times\frac{1}{4}=\frac{3}{8}$$

$$p(X=3)=\binom{3}{3}(0.5)^3(0.5)^{3-3}=\frac{3!}{3!(3-3)!}(1/2)^3(1/2)^0$$
$$=\frac{3\times2\times1}{3\times2\times1\times0!}\times\frac{1}{8}\times1=\frac{1}{8}$$

التوقع $E(X) = np = 3(0.5) = 1.5$

التباين $Var(X) = npq = 3(0.5)(0.5) = 0.75$

مثال:-

وجد فى إنتاج أحد المصانع أنه من بين 1000 وحدة إنتاج يوجد 150 وحدة معيبة. أخذت عينة بإرجاع مكونة من 5 وحدات، أوجد الاحتمالات التالية:

١- الوحدات المختارة كلها سليمة

٢- على الأكثر توجد واحدة معيبة

٣- على الأقل توجد وحدتان معيبتان

٤- القيمة المتوقعة و التباين للوحدات المعيبة .

الحل

احتمال النجاح (الحصول على وحدة معيبة) $p = 150/1000 = 0.15$

احتمال الفشل (عدم الحصول على وحدة معيبة) q = 1-p= 1-0.15 =0.85

عدد المحاولات (عينة بإرجاع مكونة من 5 وحدات) n = 5

X متغير عشوائي يمثل عدد الوحدات المعيبة يأخذ القيم 0, 1, 2, 3, 4, 5

ويكون له توزيع ذي الحدين:

$$p(X=x)=\binom{5}{x}(0.15)^{x}(0.85)^{5-x} \ , \quad x=0,1,2,3,4,5$$

١- الوحدات كلها سليمة يعنى أن X = 0

$$p(X=0)=\binom{5}{0}(0.15)^{0}(0.85)^{5-0} = \frac{5!}{0!(5-0)!}(1)(0.85)^{5}=0.4437$$

٢- على الأكثر توجد وحدة معيبة يعنى أن $X \le 1$

P(X≤1) = p(X=0) + p(X=1)

$$p(X\le 1)=\binom{5}{0}(0.15)^{0}(0.85)^{5}+\binom{5}{1}(0.15)^{1}(0.85)^{4}$$

$$=0.4437+\frac{5!}{1!5!}(0.15)(0.522)$$

$$=0.4437+5\times 0.0783=0.4437+0.3915=0.8352$$

$$p(X=x)=\binom{5}{x}(0.15)^{x}(0.85)^{5-x} \ , \quad x=0,1,2,3,4,5$$

٣- على الأقل توجد وحدتان معيبتان يعن أن $X \ge 2$

$P(X \ge 2) = 1 - p(X < 2)$

$= 1 - [p(X = 0) + p(X = 1)]$

$= 1 - 0.8325 = 0.1648$

٤- القيمة المتوقعة و التباين للوحدات المعيبة .

القيمة المتوقعة= $n . P = 5 \times 0.15 = 0.75$

التباين = $n \times p \times (1-p) =$

$= 5 \times 0.15 \times 0.85 = 0.6375$

ب - توزيع بواسون:-

توزيع بواسون هو توزيع احتمالي منفصل آخر يستخدم لتحديد احتمال وقوع عدد معين من النجاحات في وحده الزمن، وذلك عندما تكون الأحداث أو "النجاحات " مستقلة عن بعضها البعض وعندما يبقى متوسط عدد النجاحات ثابتاً لوحدة الزمن . عندئذ :

$$P(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

$$, \quad x = 0,1,2,\ldots$$

حيث : x = العدد المعين من النجاحات.

P (x) =احتمال عدد x من النجاحات.

e = أساس نظام اللوغاريتمات الطبيعي وتوجد في بعض الآلات الحاسبة،

وقيمتها هي: $e = 2.718$ تقريبا، ويمكن حساب قيمتها باستخدام الآلة الحاسبة.

$x!$ = مضروب العدد x " ويساوي: $x! = x(x-1)(x-2)..3\times2\times1$

μ = المتوسط

- يشتق توزيع بواسون من توزيع ذي الحدين عندما يكون :-
  - عدد المحاولات n كبير جدا
  - بينما يكون احتمال النجاح p صغير بحيث تبقى np قيمة ثابتة معتدلة
- يوصف متغيرات عشوائية متقطعة تعبر عن عدد كبير من الحوادث مثل:
  - عدد حوادث السيارات في الشهر داخل مدينة كبيرة
  - عدد الكرات الحمراء في عينة الدم
  - عدد الأخطاء المطبعية في الصفحات المختلفة للكتاب
  - عدد القطع التالفة في الإنتاج الكلى لسلعة معينة
  - توزيع بواسون فإن X إذا كان للمتغير

التوقع $E(X) = \lambda$

التباين $Var(X) = \lambda$

مثال :-

في كمية كبيرة من القطع المصنعة، وكان معلوما أن بها نسبة 0.3% من القطع المعيبة. أخذت منه عينة بإرجاع عشوائية حجمها 350 قطعة. احسب الاحتمالات الآتية:

١) وجود قطعة معيبة

٢) وجود قطعتان معيبتان

٣) عدم وجود أية قطع معيبة

٤) وجود على الأكثر وحدتان معيبتان

الحل

عملية سحب العينة تمثل سلسلة عددها $n=350$

واحتمال أن تكون القطعة معيبة (النجاح) $p=0.003$

واضح n كبيرة و p صغيرة

$$\lambda = np = 350(0.003) = 1.05$$

بفرض أن X يمثل عدد القطع المعيبة فى العينة له توزيع بواسون $p(X=x)=e^{-\lambda}\frac{\lambda^x}{x!}=e^{-1.05}\frac{1.05^x}{x!}$

$x = 0, 1, 2, ...$

١) <u>وجود قطعة معيبة فى العينة</u> $p(X=1)=e^{-1.05}\frac{1.05^1}{1!}=(0.3499)(1.05)=0.367$

٢) <u>وجود قطعتان معيبتان فى العينة</u> $p(X=2)=e^{-1.05}\frac{1.05^2}{2!}=(0.3499)(0.55125)=0.193$

٣) <u>عدم وجود أى قطع معيبة فى العينة</u> $p(X=0)=e^{-1.05}\frac{1.05^0}{0!}=0.350$

٤) <u>وجود على الأكثر وحدتان معيبتان</u> $P(X\le 2) = p(X=0) + p(X=1) + p(X=2)$

$= 0.350 + 0.367 + 0.193$

$= 0.91$

مثال :-

إذا كان عدد الأخطاء المطبعية في كتاب يتكون من 600 صفحة هو 50 خطأ فإذا كانت الأخطاء تتوزع توزيعا عشوائيا. فما احتمال إذا اختيرت 10 صفحات عشوائيا أن لا تحتوى على أخطاء.

الحل

بفرض أن X يمثل عدد الأخطاء في كل صفحة وأن عدد المحاولات (الصفحات) تمثل سلسلة من محاولات برنولي عددها $n = 10$

ونسبة الخطأ (النجاح) هي

$$p=\frac{50}{600}=0.083$$

وعليه فإن :- $\lambda = np = 10(0.083) = 0.83$

وبالتالي فإن لـ X توزيع بواسون:

$$p(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}=e^{-0.83}\frac{0.83^k}{k!}$$

k = 0, 1, 2, 3, …

احتمال أن لا يوجد أخطاء يساوى

$$P(X=0) = e^{-0.83}\frac{0.83^0}{0!}$$

$$= 0.436$$

مثال:-

إذا كان من المعلوم أن عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر تتبع توزيع بواسون بمتوسط 3 وحدات شهريا، إذا عرف المتغير العشوائي x بأنه عدد الوحدات التي تستهلكها الأسرة خلال الشهر من هذه السلعة.

المطلوب:

- ما نوع المتغير العشوائي؟
- اكتب شكل دالة الاحتمال لهذا المتغير.
- احسب الاحتمالات التالية:
  - احتمال أن الأسرة تستهلك وحدتين خلال الشهر؟
  - احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر؟
- احسب الوسط الحسابي، والانحراف المعياري لعدد الوحدات المستهلكة.
- حدد شكل التوزيع.

الحل:-

عدد الوحدات التي تستهلكها الأسرة متغير كمي منفصل، ومدى هذا المتغير في هذه الحالة هو:

$$X:\{x=0,1,2,3,\ldots\}$$

شكل دالة الاحتمال:

بما أن متوسط عدد الوحدات التي تستهلكها الأسرة خلال الشهر هو: $\mu=3$ ، إذا دالة الاحتمال هي:

$$P(x)=\frac{e^{-\mu}\mu^{x}}{x!}$$

$$=\frac{e^{-3}3^{x}}{x!} \quad , \quad x=0,1,2,\ldots$$

حساب الاحتمالات:

حساب احتمال أن أسرة ما تستهلك وحدتين خلال الشهر، p(2)

$$P(2)=\frac{e^{-3}3^{2}}{2!}=\frac{0.0498(9)}{2\times1}=0.22404$$

احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر هو :

$$\begin{aligned} P(X \le 3) &= p(3)+p(2)+p(1)+p(0) \\ &= \left[\frac{3^3}{3!}+\frac{3^2}{2!}+\frac{3^1}{1!}+\frac{3^0}{0!}\right]\left[\frac{0.0498}{1}\right] \\ &= [0.0498]\left(\frac{27}{6}+\frac{9}{2}+\frac{3}{1}+\frac{1}{1}\right) = 0.0498(13) = 0.6474 \end{aligned}$$

حساب الوسط الحسابي، والانحراف المعياري لعدد حالات الاستجابة:

الوسط الحسابي (μ) **في حالة التوزيع بواسون هو معلمة معطاة هي:** $\mu = 3$

في هذا التوزيع، فإن التباين يساوي الوسط الحسابي: أي أن: $\sigma^2 = \mu = 3$

ومن ثم يكون الانحراف المعياري هو: $\sigma = \sqrt{\mu} = \sqrt{3} = 1.732$

ويمكن حساب معامل الاختلاف النسبي، بتطبيق المعادلة التي سبق استخدامها ، وهو:

$$C.V = \frac{\sigma}{\mu}\times 100 = \frac{1.732}{3}\times 100 = 57.7\%$$

تحديد شكل التوزيع:

<u>دائما توزيع بواسون موجب الالتواء</u>

مثال:-

يتلقى قسم شرطة فى المتوسط 5 مكالمات في الساعة فيكون احتمال تلقى مكالمتين في ساعة مختارة عشوائياً هو :

$$\begin{aligned} P(x) &= \frac{e^{-\mu}\mu^x}{x!} \\ &= \frac{e^{-5}5^x}{x!} = \ , \quad x = 0,1,2,\ldots \\ &= \frac{(25)(0.00674)}{(2)(1)} = 0.08425 \end{aligned}$$

التوزيع الإحصائى :-

و هو الشكل الذي تأخذه مجموعة البيانات، وشكل البيانات مهم جدا في تحليلها ووصفها وكخطوة تسبق قرار استخدام أي اسلوب احصائي .

ويرتبط التوزيع الاحصائي عادة بنوعين من البيانات المتصلة والمنفصلة، ويناسب النوع المنفصل المقاييس الاسمية والرتبية ، وهناك بعض المقياس المنفصلة ثنائية أي انه لا يوجد بها الا قيميتين، وهي لا تسمي

توزيعات طبيعية وانما تسمى توزيعات ثنائية ، ومن أهم مقاييس التوزيعات المنفصلة مقياس ذو الحدين وذلك عائد لان الاجابة على المقياس الاسمي اما نعم أو لا ، ولذلك غالبا ما يرمز لها في الحاسب بصفر (غياب الصفة) [ذكور – لا] أو ١ (وجود الصفة) [اناث – نعم] .

أما التوزيعات الاحصائية المتصلة فهي ذات أهمية كبيرة في العلوم الإحصائية وذلك لأن اغلب الاختبارات الاحصائية تتعامل مع هذا النوع من البيانات.

التوزيعات الاحتمالية للمتغيرات المتصلة:

هناك بعض التوزيعات الاحتمالية المتصلة لها دوال كثافة احتمال محددة ومنها:

- التوزيع الطبيعي
- التوزيع الطبيعي (القياسي) المعياري
- توزيع t

وسنقوم في هذه المحاضرة بتناول هذه التوزيعات بشيء من التوضيح والتفصيل:

وكما أوضحنا أن المتغير العشوائي المتصل x هو ذلك المتغير الذي يمكن أن يأخذ عدداً لا نهائياً من القيم المعلومة، واحتمال أن تقع x داخل أي فترة يمثلها مساحة التوزيع الاحتمالي (ويسمى أيضاً دالة الكثافة) داخل هذه الفترة، والمساحة الكلية تحت المنحنى (الاحتمال) تساوى 1

التوزيع الطبيعي

هو أفضل وأكثر التوزيعات الاحتمالية المتصلة استخداماً في النواحي التطبيقية، ومنها الاستدلال الإحصائي شاملا التقدير، واختبارات الفروض، كما أن معظم التوزيعات يمكن تقريبها إلى هذا التوزيع .

والتوزيع الطبيعي هو توزيع احتمالي متصل، وهو جرسي الشكل ومتماثل حول الوسط الحسابي، ويمتد إلى مالا نهاية في الاتجاهين، ولكن معظم المساحة (الاحتمال) تتركز حول الوسط الحسابي .

خصائص التوزيع الطبيعي:

يعتبر التوزيع الطبيعي من أهم أنواع التوزيعات الاحصائية المتصلة ومن خصائصه انه:

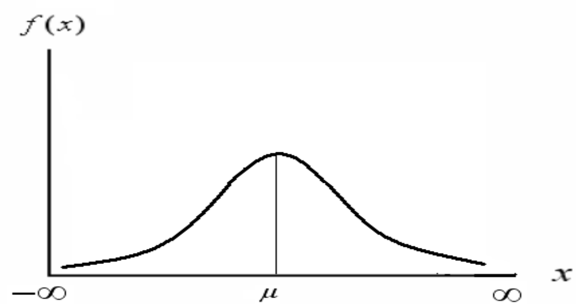
- توزيع جرسي أي يشبه الجرس.
- توزيع متصل
- توزيع متماثل حول الوسط
- الالتواء (الاطراف) والتفلطح (القمة) يساوي صفر.
- يحوي منوال ووسط ووسيط واحد وذات قيم متساوية بمعنى أن الجزء الذي على يمين الوسط مطابق للجزء الايسر
- الذيلين الايمن والايسر يقتربان من الخط الافقي ولكن لا تلامسه
- المساحة الكلية تحت المنحنى تساوي واحد صحيح

منحنى دالة الاحتمال للتوزيع الطبيعي له خاصية شكل الجرس. ويتحدد شكل الجرس تماماً لأي توزيع طبيعي خاصة إذا علمنا الوسط الحسابي μ والانحراف المعياري σ لهذا التوزيع.

تدل قيمة μ على مكان مركز الجرس، كما تدل σ على كيفية الانتشار.

القيمـــة الصغــيرة لـــ σ تعني أن لدينا جرس طويل مدبب، والقيمة الكبيرة لـــ σ تعني أن الجرس قصير ومفرطح.

والشكل التالي يوضح ذلك:



والتوزيع الطبيعي وتطبيقاته الاحصائية ليس موضوعا جديدا بل عرف منذ القرن السابع عشر الميلادي ومن ابرز الدراسات المعروفة تلك الدراسة البريطانية التي اخذت اطوال ٨٥٨٥ من الافراد البريطانيين في القرن التاسع عشر وعمل هذا المنحنى وبالتالي تم اعتبار هذه العينة تمثل التوزيع الطبيعي.

معالم هذا التوزيع:

توجد معلمتين لهذا التوزيع هما :

الوسط الحسابي : $E(x)=\mu$ والتباين : $\text{var}(x)=\sigma^2$

ومن ثم يعبر عن توزيع المتغير بالرموز : $x \sim N(\mu,\sigma^2)$ ويعني ذلك أن المتغير العشوائي x يتبع التوزيع الطبيعي بمتوسط μ ، وتباين σ^2 .

<u>التوزيع الطبيعي القياسي (المعياري) :-</u>

- احتمال وقوع أية مشاهدة على بعد انحراف معياري واحد من الوسط الحسابي هو 0.6827
- احتمال وقوع أي مفردة على بعد إنحرافين معياريين من الوسط الحسابي هو 0.9545
- احتمال وقوع أية مفردة على بعد ثلاثة انحرافات معيارية من الوسط الحسابي هو 0.9973

والشكل التالي يوضح ذلك:

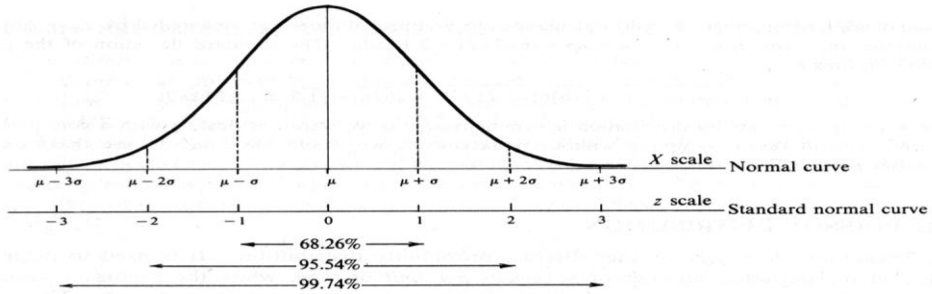


Fig. 3-4

<u>مثال :-</u>

تم دراسة متوسط طول الطالب في كلية إدارة الأعمال هو ١٨٠ سم و ذلك بانحراف معياري ١٠ سم تم اختيار أحد الطالب عشوائياً فإذا علمت أن هذه الظاهرة تتبع التوزيع الطبيعي فأوجد :-

١-احتمال أن ينحصر طول الطالب بين ١٧٠ سم و ١٩٠ سم ((p(170<x<190).

٢-احتمال أن ينحصر طول الطالب بين ١٦٠ سم و ٢٠٠ سم ((p(160<x<200).

٣-احتمال أن ينحصر طول الطالب بين ١٥٠ سم و ٢١٠ سم ((p(150<x<210).

٤- احتمال أن يكون طول الطالب أقل من ١٩٠ سم ((p(x<190).

٥- احتمال أن يكون طول الطالب أكبر من ١٩٠ سم ((p(x>190).

٦-احتمال أن يكون طول الطالب أكبر من ١٥٠ سم ((p(x>150).

٧- احتمال أن يكون طول الطالب أقل من ١٦٠ سم ((p(x<160).

الحل :-

١-احتمال أن ينحصر طول الطالب بين ١٧٠ سم و ١٩٠ سم ((p(170<x<190):-

$$z = \frac{x - \mu}{\sigma}$$

$$\frac{170 - 180}{10} < z < \frac{190 - 180}{10}$$

-1 <z < 1

P =68.26%

٢-احتمال أن ينحصر طول الطالب بين ١٦٠ سم و ٢٠٠ سم ((p(160<x<200) :-

$$z = \frac{x - \mu}{\sigma}$$

$$\frac{160-180}{10} < z < \frac{200-180}{10}$$ -2 < z < 2 P =95.45%

٣-احتمال أن ينحصر طول الطالب بين ١٥٠ سم و ٢١٠ سم ((p(150<x<210):-

$$z = \frac{x - \mu}{\sigma}$$

$$\frac{150 - 180}{10} < z < \frac{210 - 180}{10}$$

-3 <z < 3

P =99.74%

٤- احتمال أن يكون طول الطالب أقل من ١٩٠ سم ((p(x<190) :-

$$z = \frac{x - \mu}{\sigma}$$

$$z < \frac{190 - 180}{10}$$

z < 1

P = (0.6826/2) + 0.5 = 84.13%

٥- احتمال أن يكون طول الطالب أكبر من ١٩٠ سم ((p(x>190):-

$$z = \frac{x - \mu}{\sigma}$$

$$z > \frac{190 - 180}{10}$$

z > 1

P = 0.5- (0.6826/2) = 15.87%

٦-احتمال أن يكون طول الطالب أكبر من ١٥٠ سم ((p(x>150) :-

$$z = \frac{x - \mu}{\sigma}$$

$$z > \frac{150 - 180}{10}$$

z > -3

z < 3

P = (0.9974/2)+0.5 = 99.87%

٧- احتمال أن يكون طول الطالب أقل من ١٦٠ سم ((p(x<160):-

$$z = \frac{x - \mu}{\sigma}$$

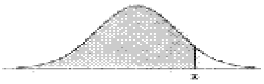
$$z < \frac{160 - 180}{10}$$

z > -2

z > 2

P = 0.5- (0.9545/2) = 0.02275= 2.275 %

Tables of the Normal Distribution



Probability Content from -oo to Z

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

استخدامات التوزيع الطبيعي القياسي:-

يستخدم التوزيع الطبيعي القياسي في التعامل مع الكثير من المشاكل العملية وإيجاد القيم الاحتمالية لها وإليك بعض الأمثلة على ذلك:

مثال:

افترض أن إدارة المرور بالأحساء وضعت جهازا للرادار على طريق الدمام عند مدخل المبرز وذلك لضبط السيارات المسرعة في فترة معينة من اليوم، افترض أن X تمثل السرعة في الساعة للسيارات التي تمر بمدخل المبرز في فترة عمل الرادار، إذا كانت X تتوزع توزيعا معتدلا وسطه الحسابي 60 ميلا وتباينه 25 ميلا، أوجد التالي:

- **نسبة السيارات التي تقل سرعتها عن 50 ميلا في الساعة .**

- نسبة السيارات التي تزيد سرعتها عن 65 ميلا في الساعة.
- نسبة السيارات التي تكون سرعتها بين 60 ميلا و 70 ميلا في الساعة .
- عدد السيارات التي تكون سرعتها بين 60 ميلا و 77.45 ميلا من بين 10000 سيارة .

الحل :-

١- نسبة السيارات التي تقل سرعتها عن ٥٠ ميلا في الساعة :

$$P(X<50)=P(\frac{x-\mu}{\sigma}<\frac{50-60}{\sqrt{25}})=P(Z<-2)=0.5-(0.9545/2)=0.02275$$

٢- نسبة السيارات التي تزيد سرعتها عن 65 ميلا في الساعة :

$$P(X>65)=P(\frac{x-\mu}{\sigma}>\frac{65-60}{\sqrt{25}})=P(Z>1)=1-P(Z\leq 1)=0.5-(0.6826/2)=0.1587$$

٣- نسبة السيارات التي تكون سرعتها بين 60 ميلا و 80 في الساعة :

$$P(60\leq X\leq 77.45)=P(\frac{60-60}{\sqrt{25}}\leq Z\leq\frac{70-60}{\sqrt{25}})$$

$$=P(0\leq Z\leq 2)=P(Z\leq 2)-P(Z\leq 0)$$

٤- عدد السيارات المتوقع سرعتها بين 60 ميلا و 80 ميلا من بين 10000 سيارة :

$$=(0.9545/2)=0.4772$$

10000(0.47725)=4772

ملاحظة مهمة :- الدكتور اعطا تنبيه انه هذه المحاضره مهمة جدا و من اهم المحاضرات ..

المحاضرة (٥)

توزيعات المعاينة

Sampling Distribution's

الاستدلال الإحصائي :-

تهتم نظرية العينات بدراسة العلاقة بين المجتمع والعينات المسحوبة منه فيما يسمى **بالاستدلال الإحصائي** **statistical inference**

يعتبر الاستدلال الإحصائي من أهم الأدوات المساعدة على اتخاذ القرارات في الاقتصاد والأعمال والعلوم، ويشمل الاستدلال الإحصائي اختبار الفرضيات والتقدير.

ولكى يكون التقدير (واختبار الفروض) سليمً، ينبغي أن يبنى على عينة ممثلة للمجتمع، ويمكن تحقيق ذلك بالمعاينة العشوائية، حيث يكون لكل مفردة في المجتمع فرصة متكافئة للدخول في العينة .

العينة العشوائية:-

وهناك عدة طرق لأخذ العينات من المجتمع لاستخدامها في الاستدلال الإحصائي ومن أشهر هذه الطرق هي العينة العشوائية وهي العينة التي تكون لكل مفرده من مفردات المجتمع نفس فرصة الاختيار في العينة.

فمثلاً نستعين بعينه مسحوبة من المجتمع لتقدير معالم هذا المجتمع مثل متوسطة أو تباينه أو غير ذلك. أو أعطاء عينه من المرضى بارتفاع الضغط، مثلاً دواء معين ثم قياس ضغطهم قبل وبعد تناولهم لهذا الدواء لمعرفة ما إذا كان هذا الدواء مفيد في خفض الضغط أم لا.

المجتمع Population

أي مجموعات من المفردات تشترك في صفه أو صفات وتكون موضوع دراسة أو بحث فإن هذه المجموعة يطلق عليها إحصائيا مجتمع الدراسة أو اختصاراً المجتمع Population.

والمجتمع قد يكون مجموعة ما من البشر أو أشجار أنواع معينه من الفاكهة أو الحيوانات الزراعية أو إنتاج دولة ما لسلع معينه خلال فترة زمنية محدده...الخ.

والمجتمع قد يكون محدوداً إذا كان يمكن حصر عدد أفراده مثل سكان مدينة ما أو طلاب مرحلة تعليمية معينة

وقد يكون المجتمع غير محدود (لانهائي) إذا كان لا يمكن حصر عدد أفراده مثل النجوم والكواكب أو الكائنات الحية بمياه المحيطات والأنهار

وعند دراسة صفة ما أو صفات معينه لمجتمع ما فإن البيانات الإحصائية عن تلك الصفة أو الصفات تجمع بأحد أسلوبين:-

أولاً: أسلوب الحصر الشامل (census): وفيه تجمع البيانات عن كل مفرده من مفردات المجتمع، وهذا الأسلوب لا يتبع عادة إلا في حالة التعدادات التي تجريها الدول وتدعمها بإمكانيات ضخمه مثل تعدادات السكان والتعدادات الصناعية والتعدادات الزراعية.

الثاني: أسلوب المعاينة (Sampling method): وفيه يتم جمع البيانات عن جزء من مفردات المجتمع يختار بطريقة أو بأخرى ويطلق عليه عينه (Sample) ثم بعد ذلك يتم تعميم نتائج الدراسة على المجتمع بأكمله.

بعض مزايا أسلوب المعاينة:-

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا عديده منها:

١. يؤدي استخدام العينات العشوائية إلى خفض تكاليف الدراسات الميدانية بسبب صغر حجم العينة بالنسبة إلى حجم المجتمع وهو ما يؤدي إلى تخفيض الأعباء الإدارية والفنية التي تتطلبها أي دراسة ميدانية.

٢. يتحقق وفر واضح في الوقت الذي ينفق في دراسة ميدانية على أساس عينة بدلاً من الحصر الشامل وتتضح أهمية الوقت عندما نقوم بدراسة ظاهرة تتغير بمرور الوقت، فتكون البيانات المجموعة والنتائج وقت ظهورها غير مطابقة لواقع المجتمع وتصبح النتائج ذات قيمه محدودة بعد أن فقدت عنصر المطابقة مع واقع الظاهرة وتوزيعها الحالي في المجتمع.

3. في المجتمعات غير المحدودة (اللانهائية) مثل مجتمع الكائنات الحية في البحار والمحيطات لا يمكن أن تتم الدراسة على أساس الحصر الشامل ولكن لابد وأن تتم الدراسة بأسلوب المعاينة.

4. أيضاً هناك بعض الاختبارات لابد وأن تتم بأسلوب المعاينة لأن إجراء مثل هذه الاختبارات على أساس الحصر الشامل يؤدي إلى تلف المادة المختبرة أو هلاكها.. فاختبار صلاحية شحنه من المفرقعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينه.

أقسام العينات:-

تنقسم العينات عادة إلى قسمين رئيسين وهما عينات عشوائية وعينات غير عشوائية، وفيما يلي تفصيل لكل قسم منها:

١. العينات العشوائية:

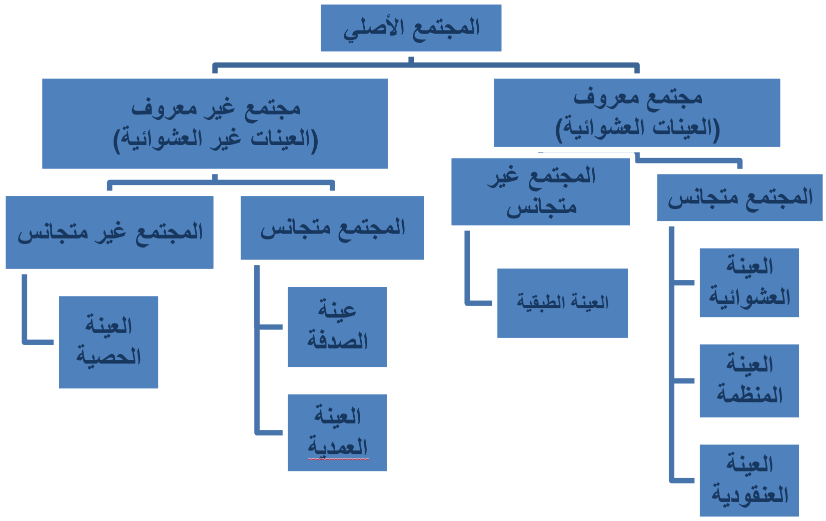
وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما تم اختيارها بالطريقة العلمية السليمة والمناسبة يمكن أن تكفل درجه عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة

٢. العينات غير العشوائية:

وهي تلك العينات التي لا تكفل لجميع مفردات المجتمع احتمال ثابت ومحدد للاختيار، وغالباً يتدخل الباحث في عملية الاختيار بصورة أو بأخرى.

وسيتم فيما يلي استعراض لأهم أنواع العينات العشوائية والعينات غير العشوائية.

أقسام العينات:



أ ـ العينات الاحتمالية:

العينة العشوائية	جميع عناصر المجتمع لها نفس الفرصة في الظهور في العينة
العينة الطبقية	يقسم المجتمع إلى طبقتين على الأقل ثم نختار العينة من كل منهما
العينة المنتظمة	نختار نقطة بداية من المجتمع ثم نختار العنصر الموجود على بعد ثابت من هذه النقطة
العينة العنقودية	يقسم المجتمع إلى مساحات أو أجزاء ثم نختار عشوائيا بعض هذه المساحات، ثم نختار جميع عناصرها بالعينة.

ب ـ العينات غير الاحتمالية:

عينة الصدفة	يتم اختيارها عن طريق الصدفة
العينة العمدية (القصدية)	يتم اختيار أفراد العينة تحت شروط معينة لتحقيق الهدف من التجربة
العينة الحصصية	يقسم المجتمع إلى أجزاء ثم نختار العينة من كل جزء من أجزاء المجتمع وفقا للنسب المحددة

أخطاء البيانات الإحصائية: ـ

تتعرض البيانات الإحصائية التي يتم جمعها إلى نوعين من الأخطاء:

١. **خطأ التميز أو التحيز:** وهو ذلك الخطأ الناتج عن مصادر متعددة، منها أخطاء في تصميم البحث أو التجربة أو أخطاء فنيه أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المتجمعة.. أخطاء التميز تزداد بازدياد الفروق بين الإمكانيات (المادية والفنية) اللازم توافرها لضمان أقصى درجة دقة ممكنه وبين الإمكانيات الفعلية المتاحة للباحث.

٢. **خطأ المعاينة العشوائية أو خطأ الصدفة:** وهو الخطأ الناتج عن فروق الصدفة بين مفردات المجتمع التي دخلت العينة وبين تلك المفردات التي لم تشأ الصدفة أن تدخل العينة

وفيما يلي شرح لهذين الخطأين:

١- **خطأ التميز أو التحيز:**

اذا سحبنا عدة عينات من مجتمع ما وحسبنا المتوسط الحسابي لكل عينة من هذه العينات ثم حسبنا المتوسط الحسابي لهذه المتوسطات فهذا المتوسط يجب أن يساوي المتوسط الحسابي الحقيقي للمجتمع المسحوبة منه هذه العينات، وفي حال وجود فرق بين المتوسطين فإن هذا الفرق يسمى **بخطأ التميز أو التحيز**

أسباب خطأ التميز أوالتحيز:

- الاختيار غير العشوائي للعينة: تعتمد بعض طرق الاختيار للعينة على خاصية معينة كالاعتماد على دليل الهاتف (عند دراسة الدخل والانفاق).
- التحيز المقصود (تعمد إدخال بعض الوحدات)
- استبدال وحدة بوحدة أخرى غير مدرجة ضمن الإطار العام للدراسة

٢- **خطأ المعاينة العشوائية Random Sampling Error**

عند اختيار العينة العشوائية هناك خطأ ينتج عن الاختلاف أو التشتت Variation بين قيم الوحدات التي تتكون منها العينة وتلك الوحدات التي لم تشأ الصدفة أن تدخلها في العينة وهذا الخطأ يسمى **بخطأ المعاينة العشوائي**

كيف نقلل من خطأ المعاينة العشوائى:

- زيادة حجم العينة
- طريقة الاختيار المناسب التي تقلل من اختلاف قيم الوحدات الإحصائية (كالأسلوب الطبقي أو العينة المنتظمة...الخ).

المعالم والإحصاءات:-

اعتاد البعض على معاملة القيم التي يحصل عليها من العينة وكأنها قيم مجتمعها، وهذا خطأ فادح. فلكي يستدل على خصائص مجتمع الدراسة تعتمد معادلات عديدة، ومتنوعة حسب نوع العينة.

فالمقاييس الإحصائية التي تحسب من بيانات مجتمع الدراسة بأكمله يطلق عليها معالم المجتمع (Parameters of population)

أما المقاييس الإحصائية التي تحسب من بيانات عينه مسحوبة من مجتمع الدراسة فيطلق عليها إحصاءات (Statistics)

وللتفرقة بين المعالم والإحصاءات يجب أن نرمز لكل منها برموز تختلف عن رموز الأخرى، على سبيل المثال يرمز للمتوسط الحسـابي للمجتمع بالرمز μ بينما يرمز للمتوسط الحسابي للعينة بالرمز $\bar{X}$ ، أيضاً للانحراف المعياري للمجتمع بالرمز σ بينما يرمز للانحراف المعياري للعينة بالرمز S وهكذا.

الخطأ المعيارى:

بمعرفة قيمة الانحراف المعياري لقيم العينة يمكن تقدير قيمة الخطأ المعياري في الانحراف المعياري للعينة باعتماد المعادلة الآتية:-

$$SE = \left[\frac{S}{\sqrt{n}}\right]\left[\sqrt{1-(\frac{n}{N})}\right]$$

حيث ان:

(N) = حجم مجتمع العينة و (n) = حجم العينة

وفق هذه المعادلة تؤخذ نسبة العينة إلى مجتمعها، وكلما كبرت هذه النسبة تحسن تمثيلها لمجتمع الدراسة، أما عندما يكون حجم مجتمع العينة مجهولا، حينها تعتمد المعادلة الآتية:-

$$SE = \frac{S}{\sqrt{n}}$$

الخطأ المعياري = الانحراف المعياري لقيم العينة / جذر حجم العينة

أما عندما يكون حجم العينة أكثر من (١٠٠) فتعتمد المعادلة أدناه:-

$$SE = \frac{S}{\sqrt{n^2}}$$

الخطأ المعياري = الانحراف المعياري لقيم العينة / جذر (حجم العينة)^2

إن الانحراف المعياري للتوزيع النظري لمتوسطات العينات يقيس خطأ المعاينة ويسمى بالخطأ المعياري للمتوسط، ومن الضروري التذكر دوما أن متوسط المجتمع قيمة محددة تقع ضمن مجال محدد Certain Interval، والباحث غير متأكد من قيمتها، ولكنه يحسب احتمالية وجودها ضمن المجال المحدد وبمستوى ثقة إحصائية معلوم

مستوى الثقة وحدودها:

إذا أخذت جميع العينات المحتملة من مجتمعها فيتوقع أن تكون متوسطات العينات موزعة بالتساوي حول متوسط مجتمع الدراسة. بعبارة أخرى، إن متوسط متوسطات العينات يساوي متوسط مجتمعها.

وتتوزع متوسطات العينات دائما بصورة متماثلة Normal Distribution ، والذي يمتاز رياضيا بالابتعاد بنسب ثابتة عن المتوسط مع كل درجة معيارية، وبالتالي تباينت متوسطات العينات المأخوذة منه فانه يتوقع أن يقع متوسطه وباحتمالية قدرها كالتالي:

- ستوى ثقة إحصائية قدره (٦٨.٢٦%) أو باحتمالية قدرها (٠.٦٨٢٧) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و (+ و -) درجة واحدة من الخطأ المعياري .
- مستوى ثقة إحصائية قدرها (٩٥%)، أو باحتمالية (٠.٩٥) يقع متوسط مجتمع الدراسة بين متوسط متوسطات العينات و (+ و –) درجتان من الخطأ المعياري تقريبا.
- مستوى ثقة إحصائية قدرها (٩٩%) أو باحتمالية قدرها (٠.٩٩) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و (+ و –) ثلاث درجات من قيمة الخطأ المعياري تقريبا.
- وتسمى هذه بمستويات الثقة Confidence Level و يعبر عنها بإشارة النسبة المئوية (%) بان تكون التقديرات صحيحة أو باحتمالية (٠.٠١) أو (٠.٠٥) أن تكون خاطئة.

مثال :-

قام أحد الباحثين في مجال الزراعة بدراسة مائة مزرعة، فوجد أن متوسط مساحة المزرعة الواحدة (٥٣) هكتارا، وبانحراف معياري عن المتوسط بقيمة (٢٦) هكتارا

أحسب حدود الثقة في تقدير متوسط مساحة المزرعة في منطقة الدراسة؟

حجم العينـة:-

إن لحجم العينة أهمية كبيرة في تحديد الثقة بالنتائج، لذا من الضروري أن يسلط الضوء عليه بشيء من التفصيل وحسب التوزيعات المعروفة للقيم، وسيتم هنا تناول نوعين من التوزيعات وهي:

١. التوزيع الطبيعي للقيم
٢. توزيع (ت) للقيم

(أ) التوزيع الطبيعي للقيم:

كلما كبر حجم العينة ازدادت دقة تمثيلها لمجتمعها واقترب توزيع القيم فيها من التوزيع الطبيعي (المتماثل الجانبين) وأصبحت عملية الاستدلال أكثر دقة. وللتوضيح نورد مثالا، إذا أريد معرفة نسبة طلبة قسم الإدارة إلى مجموع طلبة الكلية فان عينة من عشرة طلبة قد لا تفي بالغرض، ولكن عينة من مائة طالب تفي بالغرض حتما. بعبارة أخرى، إن حجم العينة أساسي لإعطاء صورة عن مجتمع الدراسة وليس النسبة المئوية للعينة قياسا بحجم مجتمعها. <u>فكلما ازداد حجم العينة ازدادت الثقة بتقديرات خصائص المجتمع وصغرت معه حدود الثقة.</u>

(ب) توزيع (ت) للقيم :

من الضروري اخذ الحذر عندما يكون حجم العينة صغيرا اقل من (٣٠) وذلك لأنها تتطلب إجراءات خاصة عند التحليل. فعندما يكون حجم العينة أكثر من (٣٠) يتجه توزيع قيمها نحو التوزيع الطبيعي وبغض النظر عن التوزيع الحقيقي لقيم مجتمع الدراسة.

وبالنسبة للعينات الصغيرة الحجم فان توزيع قيمها يتأثر بطبيعة توزيع قيم المجتمع المأخوذة منه. <u>وعندما يكون توزيع قيم المجتمع معروفا أو متوقعا أن لا يكون طبيعي حينها يجب اعتماد حجم كبير للعينة</u>. أما إذا كان توزيع قيم المجتمع طبيعيا عندها يمكن اخذ عينات بحجم صغير ويعتمد توزيع (ت) (T) في التحليل و المقارنة.

يتشابه توزيع قيم (ت) مع شكل الجرس بزيادة الحجم حتى يتطابق معه عندما يتعدى العدد (٣٠) فشكل توزيع (ت) للقيم لا يختلف كثيرا عن التوزيع الطبيعي إلا في الأعداد القليلة، وكلاهما متماثل النصفين لذا يعتمد كبديل له في القيم القليلة العدد.

ولتوزيع (ت) جداول للقيم الحرجة منظمة على شكل اسطر اعتمادا على درجة الحرية التي تقاس بـ (حجم العينة - ١). أما الأعمدة فتمثل درجة الاحتمالية Probability، وتتناقص القيم الحرجة بتزايد درجة الحرية (حجم العينة). <u>ودرجة الحرية تفضل على حجم العينة في الأحجام الصغيرة للعينة لأنها تقلل من الانحياز في تقدير خصائص مجتمع الدراسة.</u>

العوامل المحددة لحجم العينـة:

درجة التباين في خصائص مجتمع الدراسة: يلعب التباين في خصائص مجتمع الدراسة دورا مهما في تحديد درجة دقة نتائج العينة، فكلما كان التباين كبيرا تطلب الأمر زيادة حجم العينة ليكون تمثيلها للتباين في المجتمع صحيحا.

طريقة التحليل المعتمدة: عند إقرار حجم العينة، من الضروري تحديد الحجم الأصغر المقبول للعينة في المجاميع الثانوية ضمن مجتمع الدراسة ، إذ أن بعض الاختبارات الإحصائية تتطلب عددا معينا كحد أدنى لكل فئة أو صنف لتكون النتائج ذات معنى.

حجم المعلومات المطلوبة: فكلما كانت المعلومات المطلوبة من العينة (الواحدة) كثيرة وتفصيلية كان حجم العينة صغيرا، ما لم يكن المشروع البحثي كبيرا وتتوفر له المصادر البشرية والمادية اللازمة. <u>إن الدقة في المعلومات المطلوبة من العينة أهم بكثير من حجم العينة</u> ، فحجم العينة لا يتحدد بحجم مجتمع الدراسة فقط، بل وبالدقة المتوخاة والتفاصيل المطلوبة

المصادر المالية والبشرية المتوفرة: تتطلب الدراسة الميدانية توفر مصادر مالية وبشرية لتغطية تكاليفها التي تكون في الغالب باهظة لتأثيراتها على تحديد حجم منطقة الدراسة، مجتمع الدراسة وبالتالي حجم العينة. إن مضاعفة حجم العينة يتطلب زيادة في كمية المصادر المالية والجهد البشري

حدود الثقة في تقديرات خصائص مجتمع الدراسة: لزيادة الدقة في النتائج يعمد البعض إلى تقليص حدود الثقة (المدى الذي يفترض أن يقع ضمنه المعدل المتوقع للمجتمع). إن إنقاص حدود الثقة من (٦%) إلى (٤%) يتطلب زيادة حجم العينة بنسبة (٢٢٥%)، وكلما كان المدى كبيرا كان حجم العينة صغيرا، والعكس صحيح.
حالات الإخفاق وعدم الاستجابة: العامل الآخر الذي يحدد حجم العينة هي حالات الإخفاق في الحصول على المعلومات وعدم الاستجابة أو المعلومات غير الوافية

المحاضرة (٦)

التقدير الإحصائي

التقدير:-

التقدير هو عملية استنتاج أو تقدير أحد معالم المجتمع (مثل الوسط الحسابي أو الانحراف المعياري) بناءً على بيانات عينة مسحوبة من المجتمع .

وهناك نوعان (أو أسلوبان) للتقدير:

- **الأول تقدير النقطة (أو القيمة الواحدة)**
- **الثاني تقدير الفترة (أو فترة التقدير أو الثقة)**

١- التقدر بنقطة

التقدير بنقطة يعني أن نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة.

فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدولة نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة. وكمثال آخر لو أخذنا نسبة الناخبين في العينة الذين يؤيدون مرشحا معيناً كتقدير لهذه النسبة في المجتمع نكون حصلنا على تقدير نقطة للنسبة في مجتمع الناخبين .

٢- التقدير بفترة:-

أما التقدير بفترة فنحصل من خلاله على مدى Range أو فترة تتحدد بحدين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن فترة التقدير (أو تقدير الفترة) تحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات.

مثلاً:

إذا قدرنا أن الوسط الحسابي لأعمار الناخبين يتراوح بين (6 - 40) و (6 + 40) سنة أي يتراوح بين 34 سنة كحد أدنى و 46 سنة كحد أعلى نكون قد حصلنا على تقدير فترة للوسط الحسابي لأعمار الناخبين في المجتمع - ونلاحظ أن هذه الفترة (34, 46) تحتوي على عدد لا نهائي من الأعمار، بمعنى أن العدد لا يقتصر فقط على الأعداد الصحيحة والتي تشمل السنوات، ولكنها تشمل أيضا كسور السنوات، والأيام والشهور، والساعات.. الخ

تقدير الوسط الحسابي للمجتمع : -

أ – احسب الوسط الحسابي للعينة .

ب- احسب الخطأ المعياري للوسط والذي يساوي: $\frac{\sigma}{\sqrt{n}}$

جـ - أضرب الخطأ المعياري للوسط في معامل الثقة المناسب (أو الدرجة المعيارية) حسب درجة الثقة المطلوبة أي احسب $Z\frac{\sigma}{\sqrt{n}}$

د- فعندما نطرح حاصل الضرب السابق من الوسط الحسابي للعينة نحصل بالتالي على الحد الأدنى لفترة التقدير، وعندما نجمع حاصل الضرب مرة أخرى على الوسط الحسابي للعينة نحصل بالتالي على الحد الأعلى لفترة التقدير.

$$\mu = \bar{x} \pm z\frac{\sigma}{\sqrt{n}}$$

معامل الثقة Z	درجة الثقة
1	68.26%
1.65	**90%**
1.96	**95 %**
2	95.44%
2.58	**99%**

مع ملاحظة أنه إذا كان الانحراف المعياري للمجتمع غير معروف - وهو غالبا ما يحدث في الواقع – فيمكن استخدام الانحراف المعياري للعينة S بدلا منه طالما كان حجم العينة كبيرا بدرجة كافية وتصبح فترة تقدير الوسط الحسابي للمجتمع كما يلي : $\hat{\mu} = \bar{X} \pm Z\frac{S}{\sqrt{n}}$

ولإيضاح هذه النقطة بشيء من التفصيل نأخذ المثال التالي :

مثال : -

لو أردنا معرفة متوسط الدخل اليومي لمجموعة من الناخبين في دولة ما، فإن ذلك يبدو أمرا صعباً من الناحية العملية نظراً لكبر حجم مجتمع الناخبين، إضافة إلى طول الوقت والتكاليف. لذا فإن الأسلوب العلمي المتبع في حالة كهذه هو اختيار عينة عشوائية نستطيع من خلال معرفة نتائجها لتقدير متوسط دخول الناخبين في هذه الدولة.

فلو سحبت عينة عشوائية من مجموع مجتمع الناخبين في دولة ما حجمها 100 ناخب، فإذا كان الوسط الحسابي والانحراف المعياري للدخل السنوي للناخبين بالعينة هما على الترتيب 90 ألف ريال و 25 ألف ريال.

المطلوب:

أوجد فترة تقدير للوسط الحسابي للدخل السنوي لمجموع الناخبين في هذه الدولة بدرجة ثقة 95% ؟

الحل : -

بما أن فترة تقدير الوسط الحسابي للمجتمع هي : $\hat{\mu} = \overline{X} \pm Z \frac{S}{\sqrt{n}}$

والمعلومات المعطاة هي :

حجم العينة n = 100

$$\overline{X} = 90$$

وحيث أن درجة الثقة هي ٩٥% فإن: Z = 1.96 حسب ما هو موضح في الجدول السابق. وبالتالي فإن فترة تقدير الوسط الحسابي للدخل السنوي لمجتمع الناخبين بدرجة ثقة ٩٥% هي :

$$\hat{\mu} = 90 \pm 1.96 \frac{25}{\sqrt{100}}$$

$$= 90 \pm 1.96(2.5)$$

أي أن الوسط الحسابي للدخل السنوي لمجتمع الناخبين يتراوح بين 85.1 ألف ريال كحد أدنى، 94.9 ألف ريال كحد أعلى، وذلك بدرجة ثقة % 95

مثال : -

أخذت عينة عشوائية حجمها 144 بوسط مقداره 100 وانحراف معياري مقداره 60 وبالتالي فإن فترة تقدير الوسط الحسابي للمجتمع بدرجة ثقة ٩٥% هي :

$$\hat{\mu} = \overline{X} \pm 1.96 S_{\overline{X}}$$
$$= \overline{X} \pm 1.96 \frac{S}{\sqrt{n}}$$
$$= 100 \pm 1.96 \frac{60}{\sqrt{144}}$$
$$= 100 \pm 1.96(5)$$
$$= 100 \pm 9.8$$

أى أن $\hat{\mu}$ تقع بين 90.2 , 109.8 بدرجة ثقة 95% . وكثيرا ما تستخدم أيضاً درجات الثقة 90 , 99% وهى مناظرة لقيمة z=1.64 , z=2.58 على الترتيب

تحديد حجم العينة لتقدير الوسط الحسابي للمجتمع :-

يعتبر تحديد حجم العينة المناسب من المشاكل المهمة والشائعة التي تواجه الباحثين في مختلف المجالات، وبالذات عند دراسة الظواهر السياسية والإجتماعية ...الخ ، ويختلف تحديد حجم العينة باختلاف الهدف من التقدير.

فإذا كان المطلوب هو تقدير الوسط الحسابي للمجتمع، فإن فترة تقدير الوسط هي كما سبق وأن أوضحنا :

$$\hat{\mu} = \overline{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

$$n = \frac{Z^2 \sigma^2}{e^2}$$

ومنها نجد أن حجم العينة يأخذ الشكل التالي :

Z = هو معامل الثقة (أو الدرجة المعيارية) المقابل لدرجة الثقة المطلوبة، ونحصل عليا من جدول التوزيع الطبيعي المعياري.

= هو تباين المجتمع (أو هو مربع الانحراف المعياري).

e = هو أقصى خطأ مسموح به في تقدير الوسط، وهو عادة ما يحدده الباحث، وتتوقف قيمته على أهمية الموضوع أو الظاهرة السياسية المراد دراستها، ومدى الدقة المطلوبة في التقدير، ويسمى اختصاراً "الخطأ في تقدير الوسط".

ولتوضيح كيفية تحديد حجم العينة المناسب عند تقدير الوسط الحسابي للمجتمع، نأخذ المثال التالي :

مثال : -

إذا كانت دخول الأفراد اليومية في إحدى دول العالم النامية تتبع التوزيع الطبيعي بانحراف معياري دولاراً، فما هو حجم العينة المناسب لتقدير متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ في تقدير متوسط الدخل اليومي 5 دولارات، وذلك بدرجة ثقة % 99؟

الحل : -

في هذا المثال نجد أن :

درجة الثقة % 99 أي أن : Z = 2.58

أقصى خطأ مسموح به هو 5 دولارات، أي أن : e = 5

والانحراف المعياري للمجتمع : $\sigma = 15$

وبالتعويض بهذه القيم في المعادلة التي تحدد حجم العينة وهي :

$$n = \frac{Z^2 \sigma^2}{(e)^2}$$

فإن حجم العينة مقرباً لأقرب عدد صحيح هو :

$$n = \frac{(2.58)^2 (15)^2}{5^2}$$

$$= \frac{(6.65)(225)}{25}$$

$$= \frac{1496.25}{25} = 59.85 \approx 60$$

أي أنه يجب على الباحث أن يأخذ عينة لا يقل حجمها عن 60 فرداً حتى يكون لديه تقديراً دقيقاً عن متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ في تقديره لمتوسط الدخل عن خمس دولارات، وذلك بدرجة ثقة % 99.

مثال : -

يرغب أحد مدراء إحدى المصانع في تقدير متوسط عدد الدقائق التي يأخذها العمال لإنجاز عملية صناعية معينة بحيث لا يتعدى الخطأ في تقدير متوسط الأداء في حدود ± 3 دقيقة وبدرجة ثقة 90% . ويعلم المدير من خبرته الماضية أن الانحراف المعياري σ هو 15 دقيقة .

الحل : -

في هذا المثال نجد أن :

درجة الثقة ٩٠% أي أن : Z = 1.65

أقصى خطأ مسموح به هو ٣ دقائق، أي أن : e = 3

والانحراف المعياري للمجتمع : $\sigma = 15$

وبالتعويض بهذه القيم في المعادلة التي تحدد حجم العينة وهي : $n = \frac{Z^2 . \sigma^2}{(e)^2}$

فإن حجم العينة مقرباً لأقرب عدد صحيح هو :

$$n = \frac{(1.65)^2 (15)^2}{3^2}$$
$$= \frac{(2.72)(225)}{9}$$
$$= \frac{612}{9} = 68$$

أي أنه يجب على المدير أن يأخذ عينة لا يقل حجمها عن 68 فرداً حتى يكون لديه تقديراً دقيقاً لعدد الدقائق التي يأخذها العمال لإنجاز عملية صناعية معينة بحيث لا يتعدى الخطأ في تقديره لمتوسط الإنجاز عن ثلاث دقائق، وذلك بدرجة ثقة 90 %.

ومما سبق نستنتج أن :-

في حالة تقدير النقطة نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة. فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدولة نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة.

أما في حالة تقدير الفترة أو فترة التقدير فنحصل على مدى Range أو فترة تتحدد بحدين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن فترة التقدير (أو تقدير الفترة) تحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات.

فترات الثقة للمتوسط باستخدام توزيع t :-

تناولنا فيما سبق التقدير الإحصائي للوسط الحسابي للمجتمع في الحالات التي يكون فيها الانحراف المعياري للمجتمع معلوماً، و (أو) أن العينة كبيرة بدرجة كافية.

ولكن إذا كانت العينة صغيرة بمعنى أن حجمها أقل من (30) مفردة، والانحراف المعياري للمجتمع الطبيعي غير معلوم، فإن التوزيع الإحصائي المتبع في مثل هذه الحالات هو ما يطلق عليه " توزيع t "

ولعل الاختلاف الأساسي بين توزيع t والتوزيع الطبيعي هو أن الانحراف المعياري للعينة هو المستخدم في الأول بدلا من الانحراف المعياري للمجتمع في الثاني، وفيما عدا ذلك فالتوزيعان متماثلان وكلما زادت قيمة n كلما اقترب توزيع t من توزيع z ويعتمد توزيع t على ما يعرف بدرجات الحرية DEGREES OF FREEDOM.

درجات الحرية :

تعرف درجات الحرية بأنها عدد المشاهدات المستقلة في العينة والتي تساوي حجم العينة مطروحاً منه عدد القيود أو معالم المجتمع التي يتم تقديرها من بيانات العينة.

وكمثال مبسط لشرح فكرة درجات الحرية نفترض أن لدينا 3 قيم واشترطنا أن مجموع القيم يساوي 10 فإن لدى الباحث في هذه الحالة حرية في اختيار الرقم الأول (وليكن 2) والثاني (وليكن 3) لذلك فإن قيمة الثالثة لابد وأن تكون (5) بالتالي نستطيع القول بأن درجة الحرية المتاحة لدى الباحث هي (2) أي 1 - 3 = 2 أي أن درجات الحرية في هذه الحالة هي :

n - 1

حيث n تساوي حجم العينة (والتي تساوي في المثال السابق 3)

والرقم (1) والذي طرحناه يعني الشرط الذي يحتم أن مجموع القيم = 10

وبصفة عامة إذا كان عدد القيود k فإن درجات الحرية تساوي n - k

وفترة الثقة 95% لوسط المجتمع غير المعلوم عند استخدام توزيع *t* هى:

$$P(\overline{X} - t\frac{s}{\sqrt{n}} < \mu < \overline{X} + t\frac{s}{\sqrt{n}}) = 0.95 \qquad (4.7)$$

حيث تشير *t* الى قيمة *t* التى تقع عندها 2.5% من المساحة الكلية للمنحنى عند كل طرف (عند درجات الحرية المستخدمة)، وتستخدم $s/\sqrt{n}$ بدلا من $\sigma\bar{x} = \sigma/\sqrt{n}$

شروط توزيع t :

ويمكن تحديد الشروط الثلاثة لاستخدام توزيع t كما يلي:

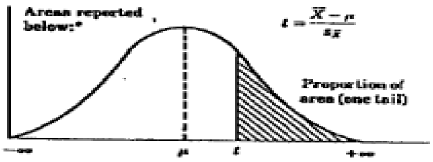
1. **أن يكون المجتمع المسحوبة منه العينة له توزيع طبيعي.**
2. **والانحراف المعياري للمجتمع غير معروف (أو مجهول).**
3. **والعينة صغيرة (حجمها أقل من 30 مفردة).**

٢. جدول توزيع t:

الجدول أدناه يعطي قيمة $t_{\propto}$

المقابلة للمساحة المظللة وقيمتها $\propto$

Proportions of Area
for the t Distributions



df	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898

df	0.10	0.05	0.025	0.01	0.005
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

مثال : -

العينة أقل من ٣٠ و المجتمع يتبع التوزيع الطبيعي :

سحبت عينة عشوائية من n=10 بطارية فلاش متوسطها ٥ ساعات، والانحراف المعياري للعينة s=1 ساعة من خط إنتاج من المعروف أنه ينتج بطاريات عمرها موزع طبقاً للتوزيع الطبيعي .

المطلوب :

إيجاد فترة الـ 95% ثقة للمتوسط غير المعلوم لعمر البطاريات في المجتمع كله.

الحل:

لإيجاد فترة الـ 95% ثقة للمتوسط غير المعلوم لعمر البطاريات فى المجتمع كله، فإننا نوجد أولا قيمة t 0.025 و التى تكون معها 2.5% من المساحة عند الأطراف لدرجات حرية n-1=9 . ونحصل على هذه القيمة من خلال الرجوع إلى جدول t بالتحرك تحت عمود 0.025 حتى درجات حرية 9 والقيمة التى سيتم التحصل عليها هى 2.262 إذن:

$$\hat{\mu} = \overline{X} \pm 2.262 \frac{s}{\sqrt{n}} = 5 \pm 2.262 \frac{1}{\sqrt{10}} \cong 5 \pm 2.262(0.316) \cong 5 \pm 0.71$$

وتقع $\hat{\mu}$ بين 4.29 , 5.71 ساعة بدرجة ثقة 95% (أنظر الشكل التالي):

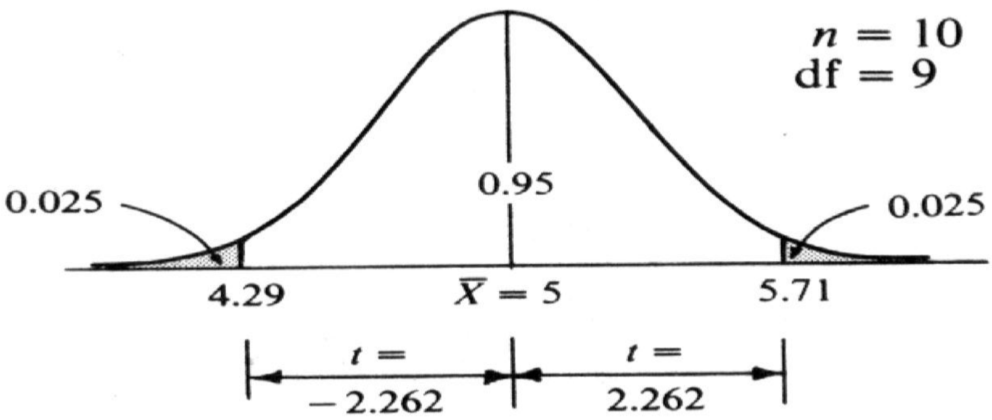


Fig. 4-4

تقدير فترة النسبة للمجتمع

(فترة الثقة للنسبة)

إن تقدير النسبة في المجتمع تعتبر من الحالات المهمة لقياس الظواهر السياسية، وبالذات الوصفية منها كقياس اتجاهات الرأي العام، وقياس نسبة قتلى الحروب، ونسبة الدول التي أوفت بالتزاماتها في المنظمات الدولية أو الإقليمية... وغيرها ونظراً لأنه من الصعوبة بمكان في كثير من الأحيان حساب هذه النسبة مباشرة من المجتمع، فإننا غالبا ما نلجأ لتقدير هذه النسبة من عينة عشوائية مسحوبة من هذا المجتمع.

خطوات تقدير النسبة في المجتمع:

فلو افترضنا أن نسبة المؤيدين للسياسة الاقتصادية التي تنتهجها دولة ما هي P وأن العينة العشوائية كبيرة بدرجة كافية وأن نسبة مؤيدي هذه السياسة في العينة هي $\hat{p}$ فتقدير النسبة في المجتمع تكون كما يلي:

$$P = \hat{P} \pm Z\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

مثال : -

عينة عشوائية حجمها 144 ناخباً سحبت من إحدى المدن فوجد أن عدد المؤيدين في العينة لمرشح معين هو 60 ناخباً، أنشئ فترة تقدير لنسبة المؤيدين لهذا المرشح في المدينة كلها بدرجة ثقة 95 %.

الحل:

نحسب أولاً نسبة المؤيدين للمرشح في العينة $\hat{P}$ التي نحصل عليها بقسمة عدد المؤيدين له على العدد الكلي للعينة (حجم العينة) أي أن :

$$\hat{P} = \frac{60}{144} = 0.42$$

وحيث أن درجة الثقة المطلوبة هي 95 % فإن معامل الثقة المناسب هو: Z = 1.96 وفترة تقدير نسبة المؤيدين لهذا المرشح في المدينة تأخذ الشكل التالي :

$$P = \hat{P} \pm z\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

وبالتعويض عن حجم العينة n = 144

والنسبة في العينة $\hat{P} = 0.42$، $1-\hat{P} = 1-0.42 = 0.58$

ومعامل الثقة Z = 1.96

نحصل بعدها على :

$$P = 0.42 \pm 1.96\sqrt{\frac{0.42 \times 0.58}{144}}$$
$$= 0.42 \pm (1.96)(0.0411)$$
$$= 0.42 \pm 0.08$$

$$\therefore P \begin{cases} 0.34 \\ 0.50 \end{cases}$$

أي أن نسبة المؤيدين للمرشح في المدينة تتراوح بين 0.34 , 0.50 وذلك بدرجة ثقة 95 % ، بمعنى آخر أن نسبة مؤيدي هذا المرشح في هذه المدينة لا تتجاوز 50 % كحد أعلى، وبالتالي ففرصته في الفوز كمرشح قد لا تكون كبيرة وذلك بدرجة ثقة 95 % بمعنى أن هذا الحكم لا تتجاوز نسبة الخطأ فيه 5 %.