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 ≤ x ≤ 1 x}

x= { x : 0 ≤ x ≤ 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 :-

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

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

$(A \cup B) = \{1, 2, 3, 4, 6, 7, 8\}$

التقاطع :-

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

مثال :-

إذا كان $A = \{-1, 0, 1, 2, 3\}$ و $B = \{0,2,4,6,\}$ أوجد $A \cap B$

الحل

$(A \cap B) = \{0, 2\}$

المكملة أو المتممة :-

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

مثال

إذا كان $\cup = \{1,2,3,4,5,6,7,8,9,10\}$ 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 :

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

و المجموعة الكلية

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

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

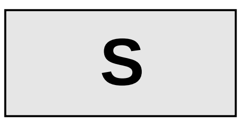
أشكال ِفنْ

VIN Figures

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

١ـ المجموعة الكلية:

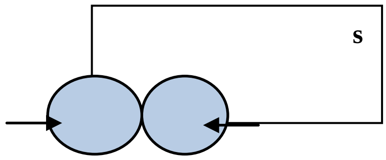
تمثل المجموعات الكلية بمستطيل ويرمز لها بالرمز S



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 ِفنْ 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 :

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

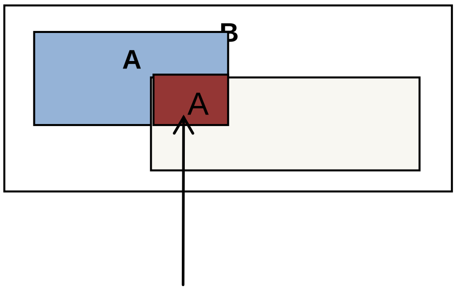
ويعني ذلك توزيع الإتحاد على التقاطع.

- وكذلك هناك خاصية التبديل والتي تعني أن

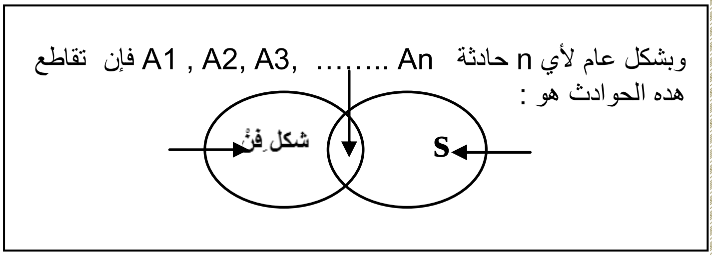
$$(A \cup B) = (B \cup A)$$

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 (B A) أو (A و B) وباستخدام أشكال ِفنْ يكون الجزء المحدد بـ A and B هو الذي يمثل تقاطع الحادثتين :



$$(A \cap B)$$



B

فإن تقاطع هده الحوادث هو : A1 , A2, A3, …….. An حادثة nوبشكل عام لأي

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap A_3 \ldots\ldots \cap_n$$

ويمكن القول أن $\bigcap_{i=1}^{n} A_i$ هو حدث يقع إذا وفقط وقعت كل الحوادث Ai على الأقل وهو 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:

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

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

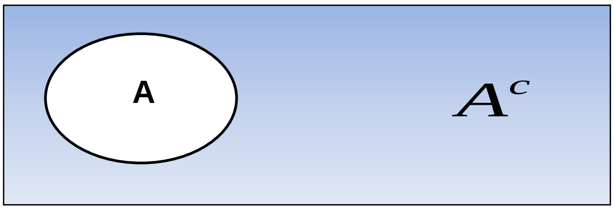
ويعني دلك توزيع التقاطع على الإتحاد

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

٤- الحادثة المتممة Complementary Event

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

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



$\overline{A}$ 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}

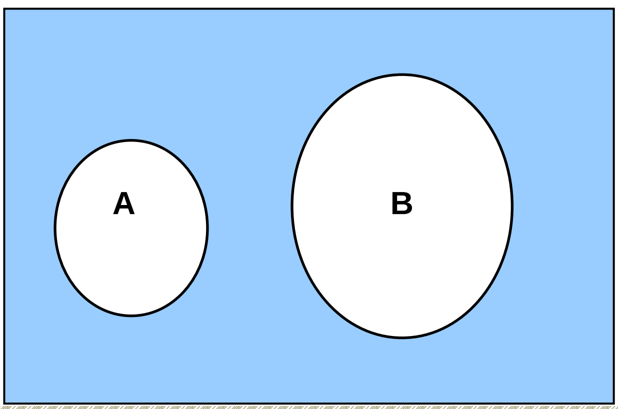
B={1, 2, 3, 6, 8, 11, 12, 14, 16}

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



$$A \cap B = \phi$$

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

بعض العلاقات المهمة :-

$$\overline{B \cup A} = \overline{B} \cap \overline{A} \qquad A \cup A^c = S$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \qquad A \cap A^c = \phi$$

إذا كانت $A \subset B$ فإن:

$$A = A \cap B$$

$$B = A \cup B$$

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

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

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

$$A \cup S = S$$

$$A \cap S = A$$

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

و = ∩
أو = ∪

أمثلة وتمارين

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 { 7 , 5 , 3} = A و {1,2,3,4,5,6,7,8 } = B 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 }

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

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

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

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

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

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

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

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

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:

فراغ العينة هو المجموعة الشاملة التي تحتوي على جميع النتائج الممكنة للتجربة العشوائية ويرمز له

Ω و يطلق عليه الحالات الممكنه

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 تجربة (Experiment) وظهور رقم فردي هو محل اهتمامنا يسمى حادثاً (Event) ومجموعة جميع الحالات الممكنة الظهور تسمى بالفراغ العيني (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 المواتية Favorable 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.

٣ـ أنواع الحوادث :-

أـ الحوادث المتنافية

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

ب-الحوادث المستقلة

حادثين مستقلين إذا كان وقوع إحداهما أو عدم وقوعه لا يؤثر في وقوع الآخر. فمثلاً B أو A يعتبر الحادثين 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

$$P(A)= \frac{\text{عدد العمال العزاب}}{\text{عدد العمال الكلي}} = 23/50= 0.46$$

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

$$P(B)= \frac{\text{عدد العمال المتزوجين}}{\text{عدد العمال الكلي}} = 27/50= 0.54$$

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

$$P(C)= \frac{\text{عدد عمال القسم الأول}}{\text{عدد العمال الكلي}} = 12/50= 0.24$$

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

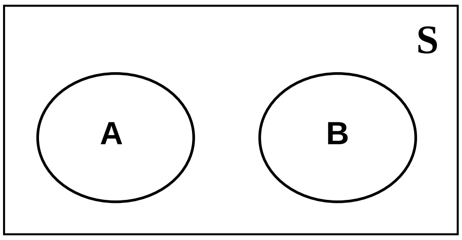
فإذا كان A ، B حادثين متنافيين كما في الشكل (١)

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:

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

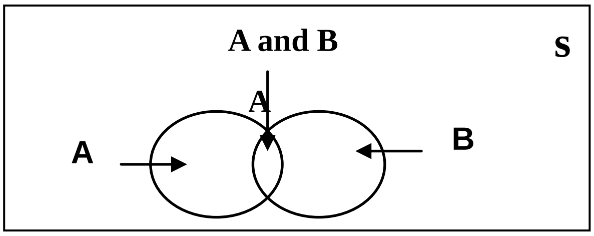
وحيث أن الحصول على رقم زوجي يعني الحصول على رقم ٢ أو رقم ٤ أو رقم ٦ وكلها حوادث متنافي، أي أن:

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 او B)= P (A) + P (B) – P (A 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

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

الاحتــمال 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:

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

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

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

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:

احتمال أن يكون من القسم الأول ومتزوج

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

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

P (B1|B2) =

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:

١. احتمال أن يكون العاملان من القسم الأول؟

٢. احتمال أن يكون العاملان متزوجان؟

٣. احتمال أن يكون للعاملين نفس الحالة الاجتماعية؟

٤. 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:

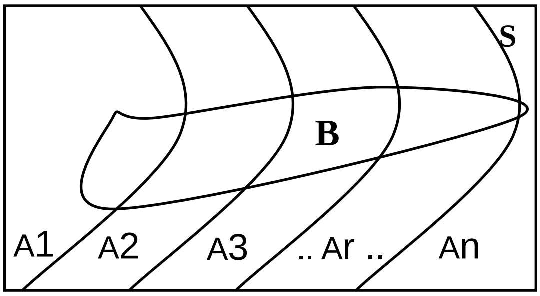
$$\begin{aligned} P(A \cup B \cup C) &= P(A) + (P(B) + (P(C) \\ &= P(A_1A_2) + P(B_1B_2) + P(C_1C_2) \\ &= P(A_1 \times A_2) + P(B_1 \times B_2) + P(C_1 \times C_2) \\ &= \left[\frac{12}{50} \times \frac{12}{50}\right] + \left[\frac{22}{50} \times \frac{22}{50}\right] + \left[\frac{16}{50} \times \frac{16}{50}\right]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

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 :

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



<u>مثال:-</u>

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؟

<u>الحل:</u>

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:

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}=\frac{0.2\times 0.02}{(0.2\times 0.02)+(0.35\times 0.025)+(0.45\times 0.03)}=0.152$$

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

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}=\frac{0.35\times 0.025}{(0.2\times 0.02)+(0.35\times 0.025)+(0.45\times 0.03)}=0.333$$

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 :

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{4}P(A_i)P(B|A_i)}=\frac{0.3\times 0.15}{(0.3\times 0.15)+(0.4\times 0.18)+(0.2\times 0.12)+(0.1\times 0.09)}=0.3$$

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

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{4}P(A_i)P(B|A_i)}=\frac{0.4\times 0.18}{(0.3\times 0.15)+(0.4\times 0.18)+(0.2\times 0.12)+(0.1\times 0.09)}=0.48$$

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 التالى يمثل توزيع موظفى أحد الشركات حسب الحالة الاجتماعية للموظف والمستوى الادارى الذى 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 العشوائي 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 الاحتمالى للمتغير العشوائي 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، $X : \{x = x_1, x_2, \dots, x_n\}$

، 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 $X : \{x = x_1, x_2, \dots, x_n\}$ ، 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.

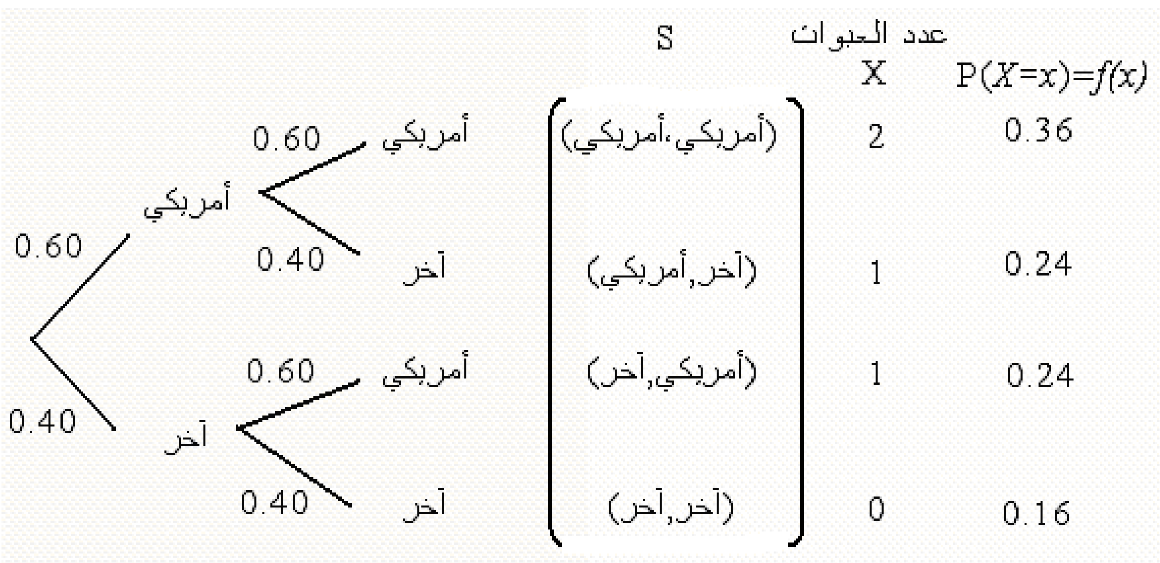
إذا عرف المتغير العشوائي بأنه عدد العبوات المشتراة من التفاح الأمريكي، فأوجد الآتي:

- 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)

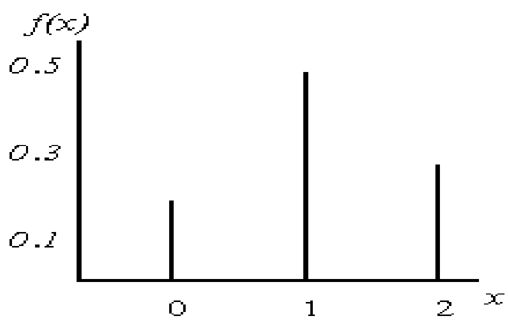
$x=2$ إذا كان العبوتين من النوع الأمريكي، أي إذا كانت نتيجة التجربة (أمريكي ، أمريكي)

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 *(30-40)*:** $\{X = x : 55 < x < 80\}$
- **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:

$$\mu = E(x) = \int_a^b x f(x) dx$$

$$\sigma^2 = E(x^2) - u^2 \ , \ E(x^2) = \int_a^b x^2 f(x) \ dx \qquad (١٣-٨)$$

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 الاحتمالى ثنائى 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: $X : \{x = 0,1,2,\dots,n\}$

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 $\mu = np$

والانحراف 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:

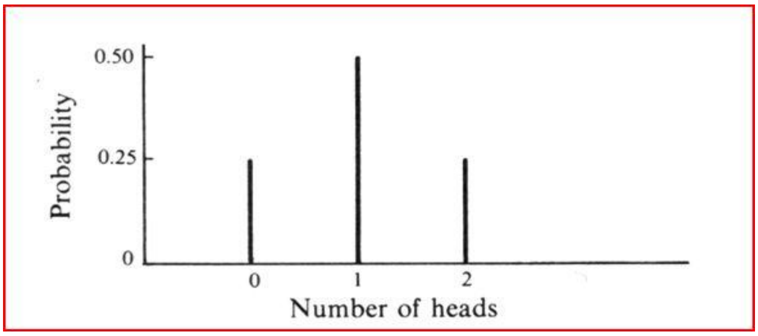
عند رمى عملة متوازنة مرتين فإن النواتج الممكنة هي TT,TH,HT,HH وعلى ذلك فإن :

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

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 :-

١. كون جدول توزيع ثنائي الحدين .
٢. أوجد إحتمال حياة ٤ أشخاص .
٣. أوجد احتمال وفاة ٣ أشخاص .
٤. أوجد احتمال حياة ٣ أشخاص على الاقل .
٥. القيمة المتوقعة (الوسط الحسابي) .

٦. الانحراف المعياري .

الحل :-

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

١- جدول توزيع ثنائي الحدين :-

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

$n = 10,\ p = 1/5,\ q = 4/5;\quad x = 0, 1, 2, \ldots, 10$

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 ≤ 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 ≥ 2

P(X ≥ 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: $x!=x(x-1)(x-2)\ldots3\times2\times1$

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) = $e^{-0.83}\dfrac{0.83^{0}}{0!}$

= 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>

<u>مثال:-</u>

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

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

<u>التوزيع الإحصائى :-</u>

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:

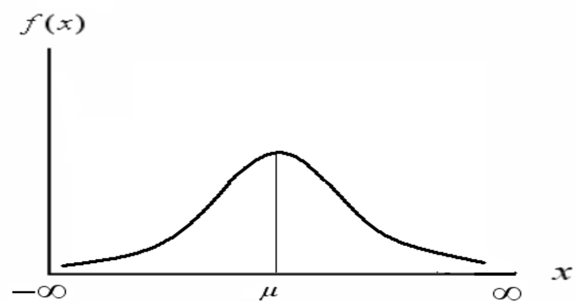
- توزيع جرسي أي يشبه الجرس.
- توزيع متصل
- توزيع متماثل حول الوسط
- الالتواء (الاطراف) والتفلطح (القمة) يساوي صفر.
- يحوي منوال ووسط ووسيط واحد وذات قيم متساوية بمعنى أن الجزء الذي على يمين الوسط مطابق للجزء الايسر
- الذيلين الايمن والايسر يقتربان من الخط الافقي ولكن لا تلامسه
- المساحة الكلية تحت المنحنى 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 الصغـيرة لــ σ تعني أن لدينا جرس طويل مدبب، والقيمة الكبيرة لــ 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 .

التوزيع الطبيعى القياسى (المعيارى) :-

- 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:

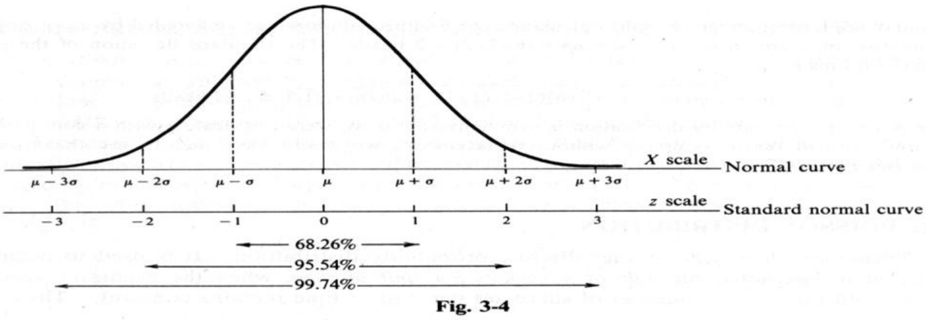


Fig. 3-4

مثال :-

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:-

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

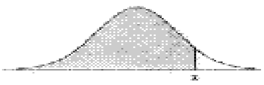
$$z < \frac{160-180}{10}$$

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 التى تقل سرعتها عن ٥٠ ميلا فى 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 التى تزيد سرعتها عن 65 ميلا فى 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 التى تكون سرعتها بين 60 ميلا و 80 فى الساعة :

$$P(60\leq X\leq 77.45)=P(\frac{60-60}{\sqrt{25}}\leq Z\leq\frac{70-60}{\sqrt{25}})$$
$$=P(0\leq Z\leq 2)=P(Z\leq 2)-P(Z\leq 0)$$
$$=(0.9545/2)=0.4772$$

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

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.

٣. 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.

٤. 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:

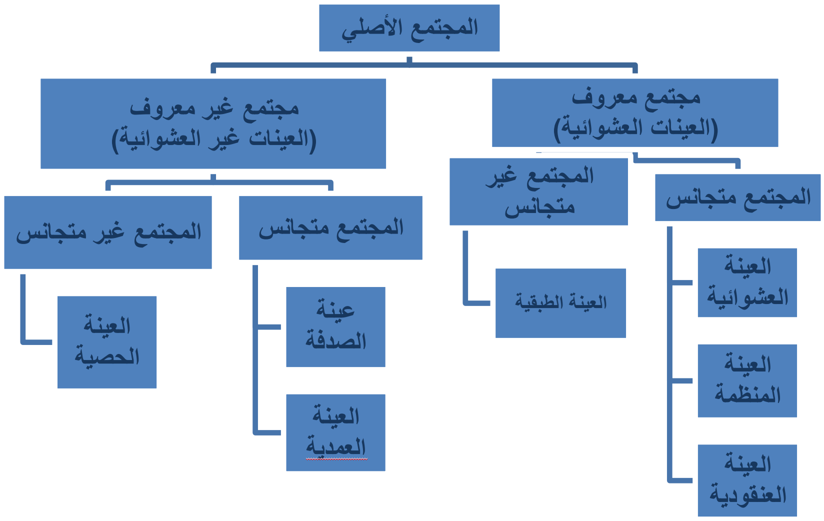
وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما 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 - 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 - 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 الحصية	يقسم المجتمع إلى أجزاء ثم نختار العينة من كل جزء من أجزاء المجتمع وفقا للنسب المحددة

أخطاء البيانات الإحصائية: -

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 العشوائي:

- 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 الحسابي 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 العينة:-

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. فكلما ازداد حجم العينة ازدادت الثقة بتقديرات خصائص المجتمع وصغرت معه حدود الثقة.

(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. وعندما يكون توزيع قيم المجتمع معروفا أو متوقعا أن لا يكون طبيعي حينها يجب اعتماد حجم كبير للعينة. 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). ودرجة الحرية تفضل على حجم العينة في الأحجام الصغيرة للعينة لأنها تقلل من الانحياز في تقدير خصائص مجتمع الدراسة.

العوامل المحددة لحجم العينة:

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. إن الدقة في المعلومات المطلوبة من العينة أهم بكثير من حجم العينة ، 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- التقدير 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 (40 - 6) و (40 + 6) 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 أحسب: $Z\frac{\sigma}{\sqrt{n}}$

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.

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

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

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 أمراً 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 السنوى لمجموع الناخبين في هذه الدولة بدرجة ثقة 95% ؟

الحل : -

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

والمعلومات المعطاة هي :

حجم العينة n = 100

$\bar{X} = 90$

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 : $\hat{\mu} = \overline{X} \pm 1.96 S_{\overline{X}}$

$$= \overline{\mathrm{X}} \pm 1.96 \frac{S}{\sqrt{n}}$$

$$= 100 \pm 1.96 \frac{60}{\sqrt{144}}$$

$$= 100 \pm 1.96(5)$$

$$= 100 \pm 9.8$$

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 :

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

ومنها نجد أن حجم العينة يأخذ الشكل التالي : $n = \dfrac{Z^2 \sigma^2}{e^2}$

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 :

<u>مثال : -</u>

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 %؟

<u>الحل : -</u>

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 %.

<u>مثال : -</u>

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 .**

<u>الحل : -</u>

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 التى 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** $\sigma\overline{x} = \sigma/\sqrt{n}$

شروط توزيع t :

ويمكن تحديد الشروط الثلاثة لاستخدام توزيع t كما يلي:

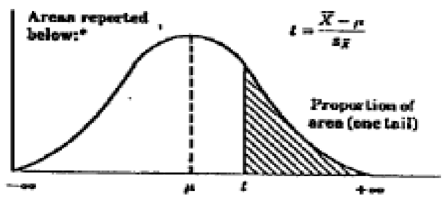
1. **أن يكون المجتمع المسحوبة منه العينة له توزيع طبيعي.**
2. **والانحراف المعياري للمجتمع غير معروف (أو مجهول).**
3. **والعينة صغيرة (حجمها أقل من 30 مفردة).**

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):

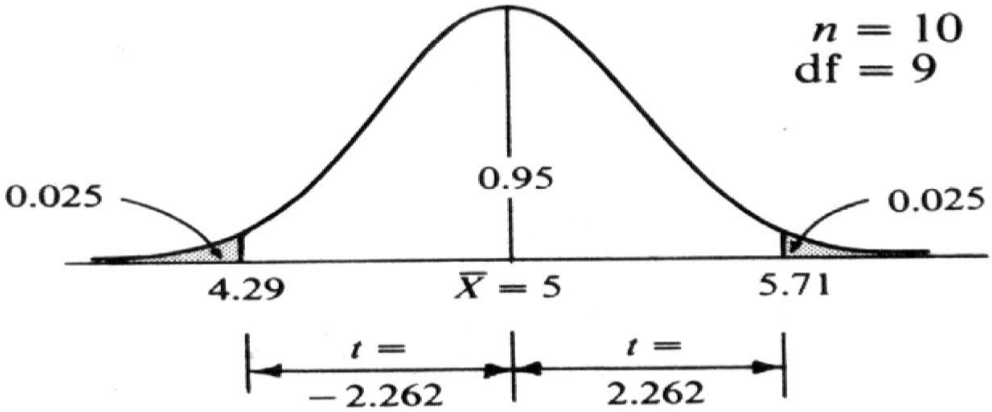


Fig. 4-4

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 :

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

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 المناســب 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 :

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

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

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 %.

المحاضرة (٧)

إختبار الفروض الإحصائية

اختبارات الفروض الاحصائية

Testing Statistical Hypotheses

المقصود بالفروض هنا الفروض الإحصائية statistical hypotheses بمعنى الفروض التي تتعلق بالمجتمع الإحصائي المسحوبة منه العينة، أو توزيع هذا المجتمع أو معالمه كالوسط الحسابي أو النسبة في المجتمع.

والفرض ما هو إلا تخمين أو استنتاج ذكي مبني على حيثيات معقولة أو منطقية ولكنه ليس مبنياً على حسابات دقيقة خاصة بالمجتمع لأننا نفترض أنه لا يمكن دراسة المجتمع بالكامل عن طريق الحصر الشامل بل نحاول استنتاج أو الاستدلال على مقاييس المجتمع باستخدام بيانات ونتائج العينة.

فمثلاً : قد يفترض الباحث أن متوسط الدخل الشهري للفرد في دولة ما هو 200 ريال (بناءً على ما يراه من مستوى المعيشة في هذا البلد وأوضاعه الاقتصادية)، ويحتاج إلى اختبار علمي (إحصائي) لمعرفة مدى صحة هذا الفرض أو قد يفترض باحث آخر أن نسبة الناخبين في إحدى الدوائر الذين يؤيدون مرشحاً معيناً لا تقل عن 30 % وهكذا... والمطلوب هو اختيار مدى صحة هذه الفروض. أي أن يصل الباحث إلى قرار إما بقبول الفرض أو عدم قبوله (أي رفضه) وذلك باحتمال معين. وقبل تناول كيفية إجراء الاختبارات الإحصائية نستعرض أولاً بعض المفاهيم والتعريفات الأساسية اللازمة لهذا الموضوع حتى تكون الصورة أكثر وضوحاً..

مفهوم الاختبارات الإحصائية :-

الفرض العدمي (أو الصفري) The Null Hypothesis

الفرض العدمي هو "الفرض الأساسي المراد اختباره". ويرمز له عادة بالرمز :Ho. هذا الفرض يأخذ – عادة – شكل معادلة أو مساواة. فمثلاً إذا كان الفرض العدمي المراد اختباره هو أن متوسط دخل الفرد في إحدى المناطق هو 200 ريال شهرياً فإن هذا الفرض يكتب بالرموز كما يلي : $Ho : \mu = 200$

ويقرأ بالشكل التالي :

الفرض العدمي هو : أن متوسط دخل الفرد في المنطقة هو 200 ريال شهرياً.

وكمثال آخر : إذا كان الفرض المراد اختباره هو أن نسبة المؤيدين لبرنامج اقتصادي معين بين عمال أحد المصانع هي 30 %، فإن هذا الفرض يكتب بالرموز كما يلي :

$$Ho : P = 0.30$$

ويقرأ بالشكل التالي :

الفرض العدمي هو : أن نسبة المؤيدين للبرنامج الاقتصادي بين عمال المصنع هي 0.30

الفرض البديل : The Alternative Hypothesis

في اختبارات الفروض يتحتم وضع فرض آخر غير الفرض العدمي المراد اختباره يسمى الفرض البديل. وهذا الفرض " هو الذي سيقبل في حالة رفض الفرض العدمي " أي لابد من تحديد فرض آخر بديل في الوقت الذي نحدد فيه الفرض العدمي، وبالتالي فإن الفرض البديل يعرف كما يلي :

"الفرض البديل هو الفرض الآخر الذي سيقبل في حالة رفض الفرض العدمي" ويرمز له عادة بالرمز : H1

<u>والفرض البديل له أهمية كبيرة فى قياس الظواهر الاجتماعية – كما سوف نرى – فهو الذى يحدد نوع الاختبار المستخدم لذلك فهو يأخذ أحد أشكال ثلاثة هى : -</u>

<u>أ- أن يأخذ شكل " لا يساوي ". وفي هذه الحالة نستخدم ما يسمى : اختبار الطرفين</u>

فمثلاً : إذا كان الفرض العدمي هو أن متوسط الدخل الشهري لفئة معينة في المجتمع

هو 200 ريال . $Ho : \mu = 200$

فإن الفرض البديل في هذه الحالة يأخذ الشكل التالي :

بمعنى أن متوسط دخل هذه الفئة من المجتمع " لا يساوي " 200 ريال شهرياً. $H1:\mu \neq 200$

بمعنى أن متوسط دخل هذه الفئة من المجتمع " لا يساوي " 200 ريال شهرياً.

<u>ب- أو أن يأخذ شكل " أكبر من ". وفي هذه الحالة نستخدم ما يسمى " اختبار الطرف الأيمن ".</u>

فمثلاً : قد يكون الفرض البديل كما يلي : $H1 : \mu > 200$

أي أن متوسط الدخل لهذه الفئة من المجتمع أكبر من 200 ريال شهرياً.

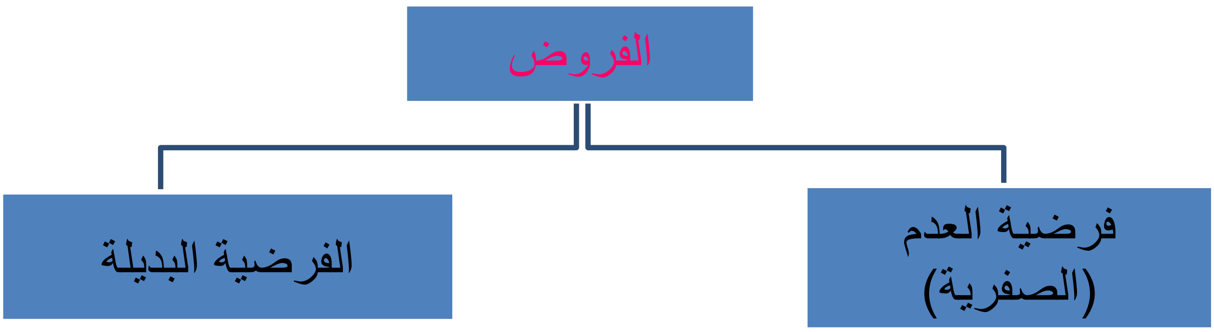
<u>ج- وأخيراً قد يأخذ الفرض البديل شكل " أقل من ".وفي هذه الحالة نستخدم ما يسمى " اختبار الطرف الأيسر".</u>

فمثلاً : قد يكون الفرض البديل هو : $H1 : \mu < 200$

أي أن متوسط الدخل لهذه الفئة من المجتمع أقل من 200 ريال شهرياً.

<u>الخلاصة : الفروض الإحصائية :-</u>

تعتبر الفروض الإحصائية بمثابة اقتراح عن معالم المجتمع موضوع الدراسة، والتي ما زالت غير معلومة للباحث، فهي حلول ممكنة لمشكلة البحث



<u>الخطأ في اتخاذ القرار : -</u>

ففي حالة قبول الباحث لفرضه العدمي، فلا مجال للبحث في الفرض البديل، أما في حالة حدوث العكس بمعنى رفض الفرض العدمي فإنه يتحتم في هذه الحالة قبول الفرض البديل، على أنه من الجدير بالذكر أن الباحث هنا عرضة للوقوع في الخطأ عند اتخاذ قراره بقبول الفرض العدمي أو رفضه، فقد يرفض فرضاً هو في الواقع صحيح، وقد يقبل فرضا هو في الواقع غير صحيح. لذلك فقد تم تصنيف هذه الأخطاء إلى نوعين هما :

<u>الخطأ من النوع الأول : Type I error</u>

الخطأ من النوع الأول هو "رفض الفرض العدمي بينما هو صحيح ". أي أنه على الرغم من أن الفرض العدمي في الواقع صحيح وكان من الواجب قبوله فقد تم أخذ قرار خاطئ برفضه. وباختصار شديد فإن الخطأ من النوع الأول هو : " رفض فرض صحيح".

<u>الخطأ من النوع الثاني : Type II error</u>

وفي المقابل فإن الخطأ من النوع الثاني يعني " قبول الفرض العدمي بينما هو خاطئ " أي أنه على الرغم من أن الفرض العدمي خاطئ وكان من الواجب رفضه فقد تم أخذ قرار خاطئ بقبوله وباختصار شديد فإن الخطأ من النوع الثاني هو " قبول فرض خاطئ ".

وقد يتساءل البعض عند مدى إمكانية تصغير الخطأين معاً ولكن لسوء الحظ لا يمكن تصغيرهما معاً إلى أدنى حد ممكن، ويبدو أن الطريقة الوحيد المتاحة لذلك هي زيادة (أو تكبير) حجم العينة، الأمر الذي قد لا يكون ممكنا في كل الحالات. لذلك فإن الذي يحدث عادة هو تثبيت أحدهما كأن يكون نسبة أو احتمال حدوث الخطأ من النوع الأول ومحاولة تصغير الآخر.

مستوى المعنوية : Level of Significance

والمقصود بمستوى المعنوية هو " احتمال حدوث الخطأ من النوع الأول ". أو نسبة حدوثه " أي احتمال رفض الفرض العدمي بينما هو صحيح ".

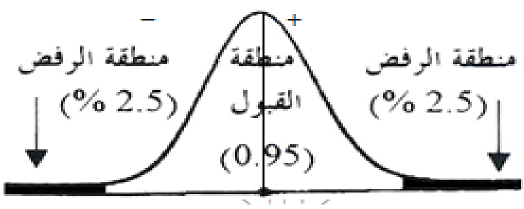
وعادة ما يرمز إلى مستوى المعنوية بالرمز اللاتيني ألفا وأشهر قيمتين لمستوى المعنوية هما 5%، 1 %، ولكن ليس هناك ما يمنع من أن يأخذ قيما أخرى.

ومن الملاحظات المهمة هنا هو أن " مستوى المعنوية " والذي يسمى أحياناً " مستوى الدلالة " هو المكمل لدرجة الثقة " بمعنى أن مجموعهما يساوي 100% أو واحد صحيح. فإذا كانت درجة الثقة 95% فإن مستوى المعنوية يساوي 5%. والعكس صحيح فإذا كان مستوى المعنوية 5% فإن هذا يعني أن درجة الثقة 95 %. ولعل من أهم الملاحظات هنا هو استخدام تعبير "مستوى المعنوية" في حالات اختبارات الفروض، بينما يستخدم مصطلح "درجة أو مستوى الثقة" في حالات التقدير.

والفكرة الأساسية في اختبار الفرض هي تقسيم المساحة تحت المنحنى إلى منطقتين: أحداهما تسمى " منطقة القبول " أي منطقة قبول الفرض العدمي. والأخرى تسمى " منطقة الرفض"، أي منطقة رفض الفرض العدمي والتي تسمى أحيانا " بالمنطقة الحرجة Critical region ". والنقطة الجديرة بالملاحظة هنا هي أن منطقة القبول تمثل درجة الثقة، بينما تمثل منطقة الرفض مستوى المعنوية. وهناك ثلاث حالات مختلفة لمنطقتي القبول والرفض هي :

الأولى : إذا كان الفرض البديل يأخذ شكل " لا يساوي " كأن يكون الفرض في هذه الحالة هو أن متوسط دخل الفرد لا يساوي 200 ريال فإن منطقة الرفض تكون موزعة على طرفي المنحنى بالتساوي، ويسمى الاختبار في هذه الحالة " اختبار الطرفين "، والذي يأخذ الشكل التالي (بافتراض أن α=5%) :

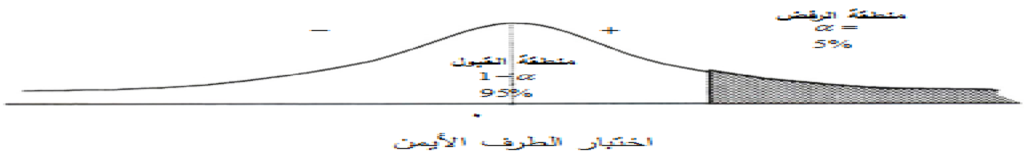
فالفرض العدمي هنا يعني أن متوسط دخل الفرد يساوي 200 ريال شهريا، والفرض البديل في هذه الحالة هو بمعنى أن متوسط دخل الفرد لا يساوي 200 ريال شهرياً. حيث تمثل المنطقة البيضاء غير المظللة منطقة القبول والتي قد تساوي 95% وبالتالي فمنطقة الرفض مقسمة بالتساوي على طرفي المنحنى في هذه الحالة تكون قيمة كل منهما 2.5 %.



اختبار الطرفين

والنتيجة هو أن القرار أيا كان نوعه سيكون بمستوى معنوية 5 % بمعنى أن احتمال أو نسبة الخطأ فيه من النوع الأول تساوي 5 %.

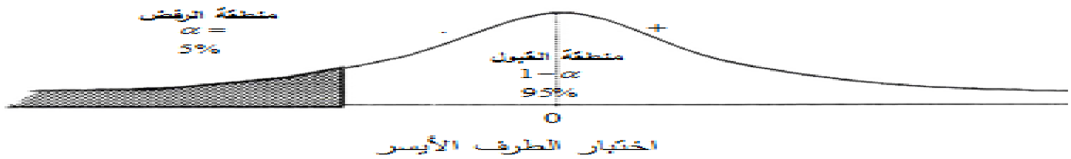
الثانية : إذا كان الفرض البديل يأخذ شكل "أكبر من" فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيمن للمنحنى. ويسمى الاختبار في هذه الحالة اختبار الطرف الأيمن. والذي يأخذ الشكل التالي أدناه :



فالفرض العدمي هنا نفس فرض المثال السابق، بينما الفرض البديل هو H1 : µ > 200

بمعنى أن متوسط دخل الفرد أكبر من 200 ريال شهرياً. وبالتالي فإن مستوى المعنوية والذي يساوي 5% مثلاً مركز في الطرف الأيمن من المنحنى.

الثالثة : إذا كان الفرض البديل يأخذ شكل " أقل من " فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيسر للمنحنى. ويسمى الاختبار في هذه الحالة اختبار الطرف الأيسر. والشكل التالي يوضح ذلك :



مع افتراض ثبات الفرض العدمي كما في المثال السابق، بينما الفرض البديل هوH1:µ<200 بمعنى أن متوسط دخل الفرد أقل من 200 ريال شهرياً، وبالتالي فإن مستوى المعنوية والذي يساوي 5 % مثلاً مركز في الطرف الأيسر من المنحنى.

خطوات الاختبار الإحصائى :

١) وضع الفرض العدمي Ho، والذي يأخذ – عادة – شكل " يساوي " فمثلاً إذا كان المطلوب هو اختبار ما إذا كان متوسط عمر الناخب هو 20 سنة فإن هذا الفرض يصاغ كما يلي : Ho : µ = 20

٢) وضع الفرض البديل H1، والذي يأخذ أحد أشكال ثلاثة إما :

" لا يساوي "

أو " أكبر من "

أو " أقل من "

وبالرموز فإن الفرض البديل قد يأخذ شكل أحد الصيغ التالية :

$$H1: \mu \neq 20$$

$$OR \mu > 20$$

$$OR \mu < 20$$

والفكرة الأساسية في اختبار الفرض هي تقسيم المساحة تحت المنحنى إلى منطقتين: أحداهما تسمى "منطقة القبول" أي منطقة قبول الفرض العدمي. والأخرى تسمى "منطقة الرفض" أي منطقة رفض الفرض العدمي والتي تسمى أحيانا " بالمنطقة الحرجة Critical region ".

والنقطة الجديرة بالملاحظة هنا هي أن منطقة القبول تمثل درجة الثقة، بينما تمثل منطقة الرفض مستوى المعنوية.

مثال (١): -

عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 ريال . كيف يمكن اختبار الفرض الصفري بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72 ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية 5 % إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال .

الحل : -

١- الفرض العدمي : هو أن متوسط المجتمع يساوي 72 وبالرموز: $Ho: \mu = 72$

٢- الفرض البديل : هو أن المتوسط لا يساوي 72 وبالرموز: $H1: \mu \neq 72$

٣- الإحصائية: بما أن العينة كبيرة فإن الإحصائية في حالة اختبار الوسط تأخذ الشكل التالي : $Z_{\bar{X}} = \dfrac{\bar{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

حيث $n = 49, \sigma = 14, \bar{X} = 75, \mu = 72$

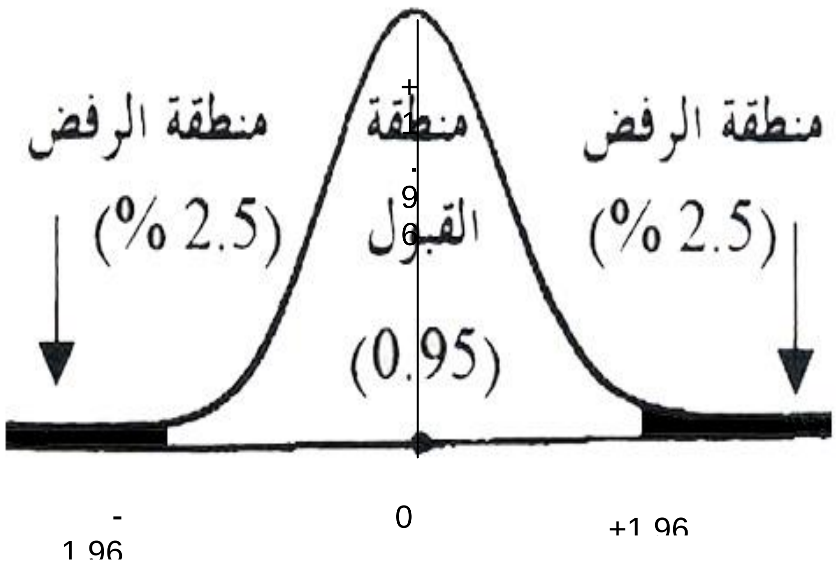
وبالتعويض نحصل على :

$$Z_{\bar{X}} = \frac{75 - 72}{\dfrac{14}{\sqrt{49}}}$$

$$Z_{\bar{X}} = \frac{3}{\dfrac{4}{7}} = \frac{3}{2} = 1.5$$

أي أن قيمة الإحصائية تساوي 1.5

٤- حدود منطقتي القبول والرفض: نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% وبما أن الفرض البديل هو: "لا يساوي" فإن ما يستخدم في هذه الحالة هو اختبار الطرفين كما في الشكل التالي :

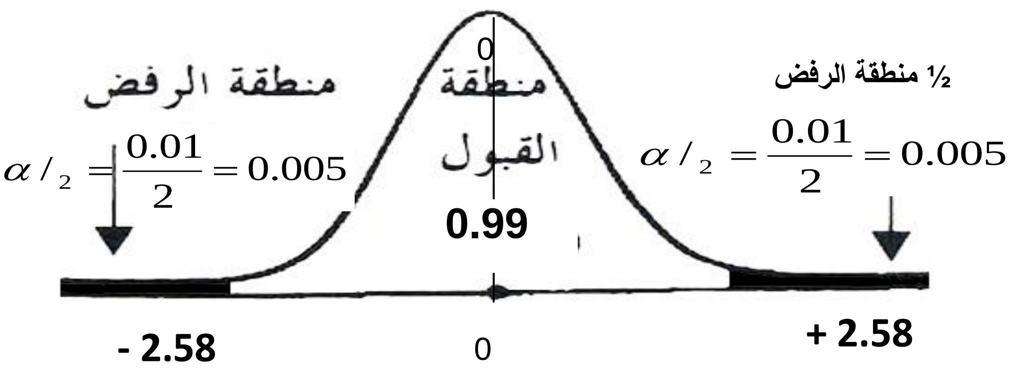


وقد حصلنا على حدود منطقتي القبول والرفض وذلك بقسمة درجة الثقة (المكملة لمستوى المعنوية) والتي تساوي 0.95 على 2 فنحصل على 0.4750 وبالكشف في جدول التوزيع الطبيعي المعياري عن Z التي تقابل المساحة 0.4750 نجد أنها تساوي 1.96 وحيث أنها موزعة على طرفي المنحنى بالتساوي فنضع إشارة موجبة في النصف الأيمن، وإشارة سالبة في النصف الأيسر، أي أن منطقة القبول تبدأ من القيمة ١.٩٦- وتستمر حتى القيمة ١.٩٦ + (أي أن أي قيمة محصورة بين هاتين القيمتين تكون في منطقة القبول، وأي قيمة خارج هذه الحدود تكون في منطقة الرفض).

٥- المقارنة والقرار: وبمقارنة قيمة الإحصائية المحسوبة من الخطوة رقم 3 (والتي تساوي 1.5) بحدود منطقتي القبول والرفض (من الخطوة رقم 4) نجد أنها تقع في منطقة القبول لذلك فإن القرار هو:

قبول الفرض الصفري بأن متوسط دخول الأفراد الأسبوعية في هذه الدولة يساوي 72 دولاراً وذلك بمستوى معنوية 5 %.

لو استخدمنا مستوى معنوية 1% بدلاً من 5% كما في المثال أعلاه فإن حدود منطقتي القبول والرفض تصبح كما يلي :



وبمقارنة قيمة الإحصائية 1.5 بحدود منطقتي القبول والرفض نجد أنها تقع في منطقة القبول أي أن القرار هو نفسه قبول الفرض الصفري ولن يتغير بل يتأكد باستخدام مستوى معنوية 1%.

مثال (٢):-

أفترض أن شركة ترغب في اختبار ما إذا كان يمكنها الادعاء بأن متوسط عمر المصباح من إنتاجها هو 1000 ساعة احتراق. وأنها قامت بأخذ عينة عشوائية حجمها $n = 100$ من إنتاجها فوجدت أن متوسط العينة $X = 980$ ساعة والانحراف المعياري للعينة $s = 80$ ساعة.

فإذا أرادت الشركة القيام بالاختبار عند مستوى معنوية 5%، فعليها القيام بالتالي:

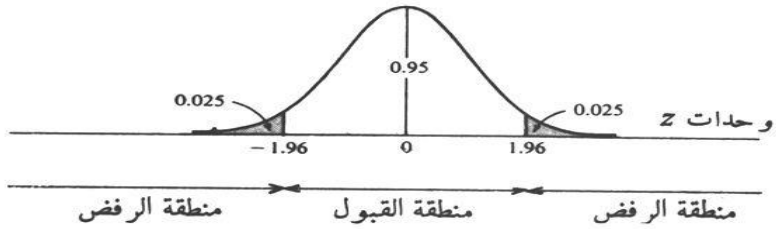
الحل :-

حيث أن μ يمكن أن تساوي أو تزيد عن، أو تقل عن 1,000، فان الشركة يجب أن تضع الفرض الصفري والفرض البديل كالآتي:

$H_1 : \mu \neq 1{,}000 \quad H_0 : \mu = 1{,}000$

وحيث أن $n > 30$ ، فإن توزيع المعاينة للوسط يكون تقريباً طبيعياً (ويمكن استخدام s كتقدير بدلاً من σ) . وتكون منطقة القبول للاختبار عند مستوى المعنوية 5% بين 1.96 تحت التوزيع الطبيعي القياسي وحيث أن منطقة الرفض تقع عند ذيل التوزيع، فإن الاختبار يسمى اختبار ذو ذيلين. وتكون الخطوة الثالثة إيجاد القيمة المناظرة لقيمة X :

$$z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{\bar{X} - \mu_0}{80 / \sqrt{100}} = \frac{-20}{8} = -2.5$$



وحيث أن قيمة z المحسوبة تقع داخل منطقة الرفض، فإن على الشركة أن ترفض الفرضية الصفرية (H_0) أي أن $\mu = 1{,}000$ وتقبل الفرضية البديلة (H_1) أي $\mu \neq 1{,}000$ وذلك عند مستوى معنوية 5%.

مثال (٣): -

ترغب شركة أن تعرف بدرجة ثقة 95% ما إذا كان يمكنها الادعاء بأن صناديق الصابون المسحوق الذي تبيعها تحتوي على اكثر من 500 جرام (حوالي 1.1 رطل) من الصابون. وتعرف الشركة من الخبرة الماضية أن أوزان الصابون بالصناديق تتبع التوزيع الطبيعي. وقد أخذت الشركة عينة عشوائية حجمها $n = 25$ ووجدت أن $X = 520$ جرام و $s = 75$ جرام.

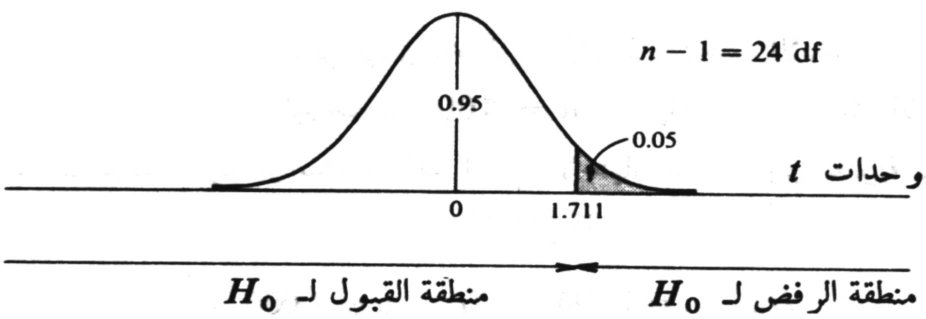
الحل : -

وحيث أن الشركة ترغب في اختبار ما إذا كانت $\mu > 500$، فإن :

$$H_1 : \mu > \quad H_0 : \mu = 500$$

وحيث أن التوزيع طبيعي ، $n < 30$، وكذلك σ غير معلومة، فعلينا أن نستخدم توزيع t (بدرجة حرية $n - 1 = 24$) لتحديد المنطقة الحرجة، أي منطقة الرفض، للاختبار بمستوى معنوية 5%. ونجد ذلك في الجدول المخصص لاختبار t ويعرضها الشكل التالي، ويسمى هذا اختبار الذيل الأيمن. وأخيراً ، حيث أن

وهي تقع داخل منطقة القبول، وتقبل H_0 أى $\mu = 500$ ، عند مستوى معنوية 5% (أو بدرجة ثقة 95%) .



مثال (٤) :-

عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 ريال . كيف يمكن اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72 ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية 5 % إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال .

الحل :-

١- الفرض العدمي : هو أن متوسط المجتمع يساوي 72 وبالرموز : Ho : µ = 72

٢- الفرض البديل : هو أن المتوسط لا يساوي 72 وبالرموز : Ho : µ ≠ 72

٣- الإحصائية : بما أن العينة كبيرة فإن الإحصائية في حالة اختبار الوسط تأخذ الشكل التالي :

$$Z_{\bar{x}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

حيث أن :- $n = 49, \sigma = 14, \bar{X} = 75, \mu = 72$

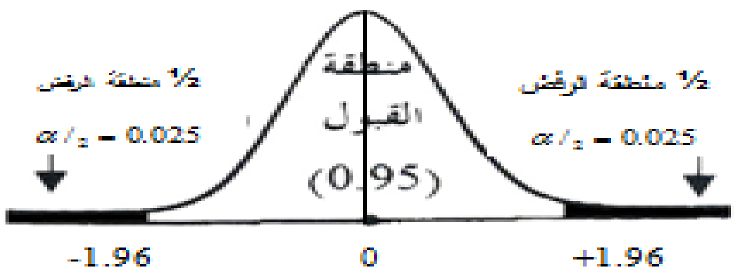
وبالتعويض نحصل على :-

$$Z_{\bar{x}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}}$$

$$Z_{\bar{x}} = \frac{3}{\frac{4}{7}} = \frac{3}{2} = 1.5$$

أي أن قيمة الاحصائي تساوي 1.5.

٤- حدود منطقتي القبول و الرفض : نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% وبما أن الفرض البديل هو : "لا يساوي " فإن ما يستخدم في هذه الحالة هو اختبار الطرفين كما في الشكل التالي :-

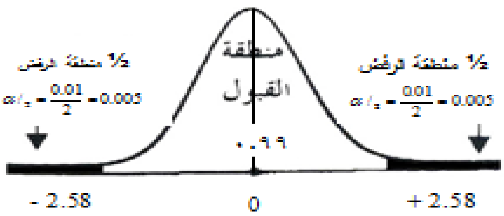


٥- المقارنة والقرار : وبمقارنة قيمة الإحصائية المحسوبة من الخطوة رقم 3 (والتي تساوي 1.5) بحدود منطقتي القبول والرفض (من الخطوة رقم 4) نجد أنها تقع في منطقة القبول لذلك فإن القرار هو :

قبول الفرض العدمي بأن متوسط دخول الأفراد الأسبوعية في هذه الدولة يساوي 72 دولاراً وذلك بمستوى معنوية 5 %.

ملاحظة :

لو استخدمنا مستوى معنوية 1% بدلاً من 5% كما في المثال أعلاه فإن حدود منطقتي القبول والرفض تصبح كما يلي :



وبمقارنة قيمة الإحصائية 1.5 بحدود منطقتي القبول والرفض نجد أنها تقع في منطقة القبول أي أن القرار هو نفسه قبول الفرض العدمي ولن يتغير بل يتأكد باستخدام مستوى معنوية 1%.

مثال (٥):-

يدّعيِ أحد المرشحين في الانتخابات أنه سيحصل على نسبة 70% من أصوات الناخبين عندما تجري الانتخابات. ولاختبار هذا الادعاء تم اختيار عينة عشوائية من الناخبين حجمها 100 ناخب، ووجد أن نسبة من يؤيدون المرشح في العينة هي 60 % اختبر مدى صحة ادعاء المرشح بأن النسبة في المجتمع هي 70 % مقابل الفرض البديل أن النسبة أقل من 70% وذلك بمستوى معنوية 5 %.

الحل :-

١- الفرض العدمي هو أن النسبة في المجتمع (نسبة من يؤيدون المرشح في المجتمع) هي 0.70 أي أن الفرض العدمي هو أن الادعاء صحيح وأن المرشح سيحصل على النسبة التي ادعاها وهي 70 % بالرموز $HO : P = 0.70$

٢- الفرض البديل والمنطقي : في هذه الحالة هو أن النسبة في المجتمع أقل من هذا الادعاء وبالرموز : $H1 : P < 0.70$

٣- الإحصائية : وتأخذ الإحصائية في حالة اختبار النسبة الشكل التالي :

$$Z_{\hat{p}} = \frac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}}$$

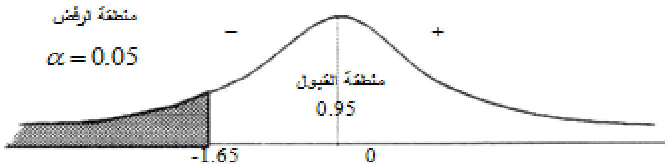
حيث أن :- $n = 100, \hat{P} = 0.60, P = 0.70، 1 - p = 1 - 0.70 = 0.30$

$$Z_{\hat{p}} = \frac{0.60 - 0.70}{\sqrt{\frac{0.70 \times 0.30}{100}}} = \frac{-0.10}{0.046}$$

$$Z_{\hat{p}} = -2.17$$

أي أن قيمة الإحصائية تساوي 2.17 -

٤- حدود منطقتي القبول والرفض نحصل عليها من التوزيع الطبيعي المعياري، حيث مستوى المعنوية $\alpha = 5\%$ وبما أن الفرض البديل هو " أقل من " فنستخدم اختبار الطرف الأيسر.



٥- المقارنة والقرار : وبمقارنة قيمة الإحصائية التي حصلنا عليها في الخطوة رقم (٣) التي تساوي 2.17 - بحدود منطقتي القبول والرفض (من الخطوة رقم ٤) نجد أن قيمة الإحصائية تقع في منطقة الرفض لأن 2.17 - أصغر من - 1.65 فإن القرار هو :

<u>رفض الفرض العدمي بادعاء المرشح بأن نسبة مؤيديه في المجتمع هي 70 % وقبول الفرض البديل بأن النسبة أقل من 70 % وذلك بمستوى معنوية 5 % (أي أن احتمال الخطأ في هذا القرار لا يتعدى 5 %).</u>

مثال (٦):-

البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما :

حيث
$$\overline{X}_1 = 35, \overline{X}_2 = 29, n_2 = 80, n_1 = 100$$

اختبر الفرض العدمي : أن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5% مقابل الفرض البديل أنهما غير متساويين إذا علمت أن :

$$\sigma_1^2 = 60, \sigma_2^2 = 32$$

١- الفرض العدمي أن المتوسطين متساويان وبالرموز : $Ho: \mu_1 = \mu_2$

٢- الفرض البديل أن المتوسطين غير متساويين وبالرموز : $HI: \mu_1 \neq \mu_2$

٣- الإحصائية : تأخذ الشكل التالي :

$$Z_{x1-x2} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

وبالتعويض عن :- $n_1 = 100, n_2 = 80, \overline{X}_1 = 35, \overline{X}_2 = 29, \sigma_1^2 = 60, \sigma_2^2 = 32$

نحصل على :-

$$Z_{\overline{X}1-\overline{X}2} = \frac{35-29}{\sqrt{\frac{60}{100}+\frac{32}{80}}}$$

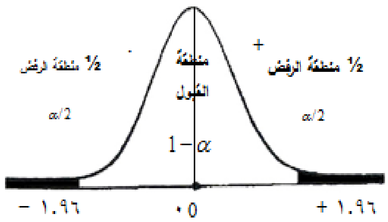
$$= \frac{60}{\sqrt{0.60+0.40}}$$

$$= \frac{6}{\sqrt{1}} = 6$$

أي أن قيمة إحصائي الاختبار تساوي 6 .

٤- حدود منطقتي القبول والرفض التي نحصل عليها من جدول التوزيع الطبيعي Z لأن العينات كبيرة، والاختبار هو اختبار الطرفين (لأن الفرض البديل لا يسـاوي)

ومستوى المعنوية المطلوب هو % 5.



أي أن منطقة القبول تبدأ من 1.96- إلى 1.96+ ومنطقة الرفض هي القيم التي أصغر من 1.96 - والتي أكبر من 1.96 +.

٥- المقارنة والقرار ولما كانت قيمة الإحصائية (والتي تساوي) 6 تقع في منطقة الرفض فإن القرار هو رفض الفرض العدمي وقبول الفرض البديل بمستوى معنوية 5% أي أننا نرفض الفرض القائل بأن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية وذلك بمستوى معنوية % 5.

المحاضرة (٨)

اختبار الفروض الإحصائية المعلمية

انواع الاختبار (الفروض)

الاختبارات الاحصائية لعينة واحدة One Sample Test

اختبار Z-test :

في كثير من الأحيان لا يمكن معرفة تباين المجتمع الذي سحبت منه العينة ، إلا أنه إذا كان حجم العينة كبيرا ($n > ٣٠$) فإنه يمكن استخدام تباين العينة الكبيرة (S^2) عوضاً عن تباين المجتمع (σ^2) الغير معلوم، وذلك لأن (S^2) مقدر جيد (σ^2) ولأنه لا يتغير كثيراً من عينة لأخرى ما دام حجم العينة كبير، ففي هذه الحالة يمكننا استخدام اختبار (Z) لاختبار الفرضيات الصفرية موضع الدراسة وذلك من خلال المختبر الإحصائي التالي:

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

ويعتبر ذلك مدخل ضروري لفهم اختبار t-test .

مثال على اختبار Z :

إذا كان متوسط استهلاك الفرد السعودي من الدجاج حسب تقارير وزارة الصحة هو (١٢) كيلوجرام بانحراف معياري (٦) كيلوجرامات لفترة السبعينات الميلادية. أجرى أحد الباحثين دراسة في عام ٢٠٠٣م من عينة قوامها (٤٩) فرداً ووجد أن متوسط الاستهلاك للفرد هو (١٤) كيلوجرام. هل تشير الدراسة الحالية أن متوسط الاستهلاك أرتفع عما عليه في السبعينات.

الحل:

١) فرض العدم والفرض البديل.

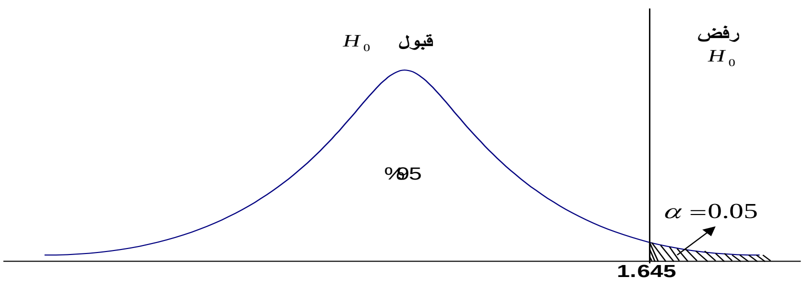
فرض العدم: $H_0: \mu=12$

الفرض البديل: $H_1: \mu>12$

٢) مستوى الدلالة =(0.05):

٣) إحصائية الاختبار (Z): $Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{14-12}{6/\sqrt{49}} = 2.33$

٤) تحديد قيمة Z المعيارية من الجدول عند مستوى دلالة (٠.٠٥)، نحتاج لتحديد قيمة Zα التي تقع علي اليمين وتساوي ١.٦٤٥ (أنظر الشكل التالي):



٥) بما أن القيمة المحسوبة أكبر من القيمة النظرية المستخرجة من الجدول كما يبين الشكل، فإنها تقع في منطقة الرفض. وبذلك نرفض فرض العدم حيث أن البيانات المتوفرة تقدم دليلاً كافياً على أن متوسط استهلاك الفرد من لحوم الدواجن في الوقت الحالي قد اختلف بمستوى معنوي أو ذو دلالة عما علية في سبعينات القرن الماضي.

اختبار t-test :

ولكن إذا كان حجم العينة صغيراً (n < ٣٠) فإن قيمة (S^2) تتغير كثيراً من عينة إلى أخرى وبالتالي لا يمكننا هنا أن نستخدم اختبـار (Z)، مما دفع كثيراً من الإحصائيين للبحث عن البديل المناسب .

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

مثال على اختبار t :

لو كانت لدينا عينة عشوائية تتكون من ٢٥٠ طالب وجد أن الوسط الحسابي لأطوال طلاب العينة ١٥٥.٩٥سم، والانحراف المعياري = ٢.٩٤ سم، علما بأن الوسط الحسابي لأطوال طلاب الجامعة يبلغ ١٥٨ سم، اختبر أهمية الفرق المعنوي بين الوسط الحسابي لأطوال طلاب العينة والوسط الحسابي لأطـوال طلاب الجامعة .

الحل :

سيتم اختبار الفرضيات التالية :

الفرضية الصفرية : لا توجد فروق ذات دلالة إحصائية بين متوسط أطوال الطلاب في العينة ومتوسط أطوال الطلاب في الجامعة

(μ = μ₀)

الفرضية البديلة : توجد فروق ذات دلالة إحصائية بين متوسط أطوال الطلاب في العينة ومتوسط أطوال الطلاب في الجامعة

(μ   μ0)

مستوى الدلالة : α = ٠.٠٥

منطقة الرفض : قيمة (ت) الجدولية عند مستوى دلالة α ٠.٠٥ودرجات حرية ٢٤٩ = ١.٩٦٠

المختبر الإحصائى :

= ١٥٥.٩٥سم ، = ٢٥٠ طالب ، = ٢.٩٤ سم

= ١٥٨سم

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}} = \frac{155.95 - 158}{2.94/\sqrt{250}} = -11.006$$

القرار :

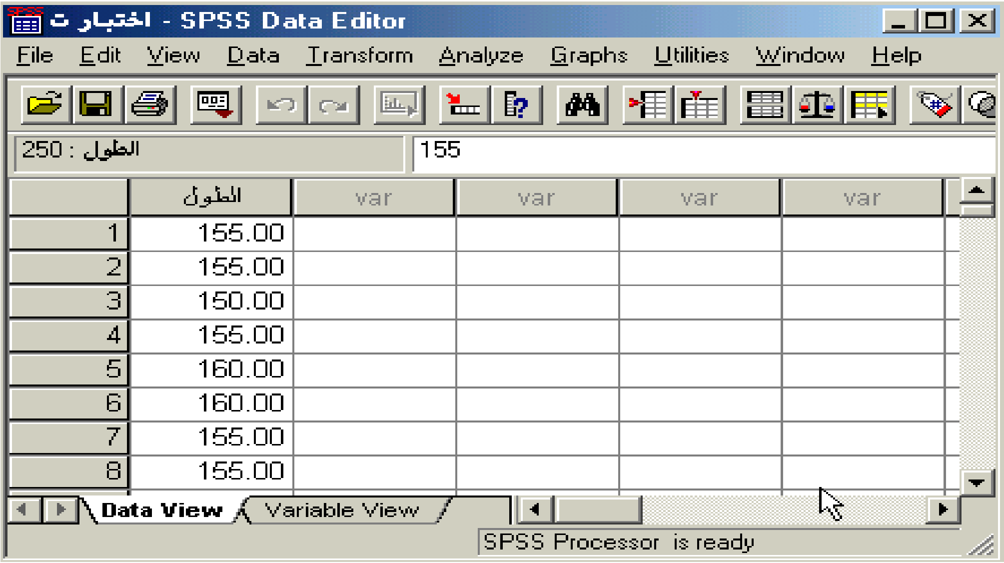
∵ قيمة ت المحسوبة (- ١١.٠٠٦) أكبر من قيمة ت المجدولـة (١.٩٦) عند مستوى دلالة α = ٠.٠٥ .

∴ نرفض الفرضية الصفرية ونقبل البديلة .

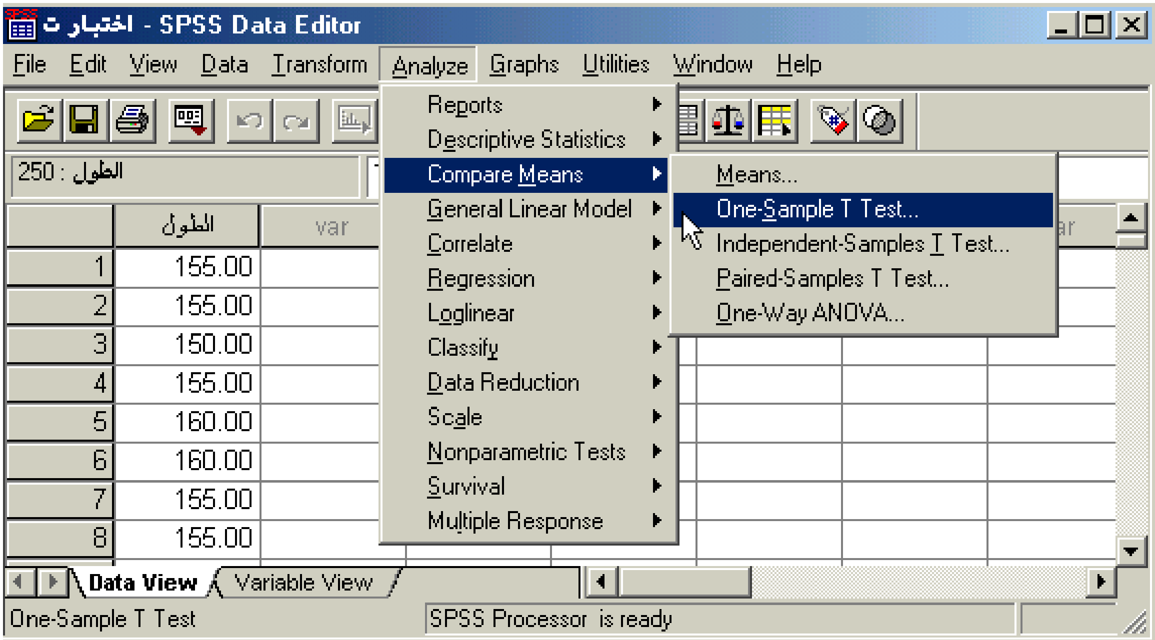
أي أنه توجد فروق ذات دلالة إحصائية بين الوسط الحسابي للعينة والوسط الحسابي لمجتمـع البحث .

لغرض حساب قيمة (ت) لنفس المثال السابق من خلال استخدام برنامج الـ SPSS اتبع الخطوات التالية :

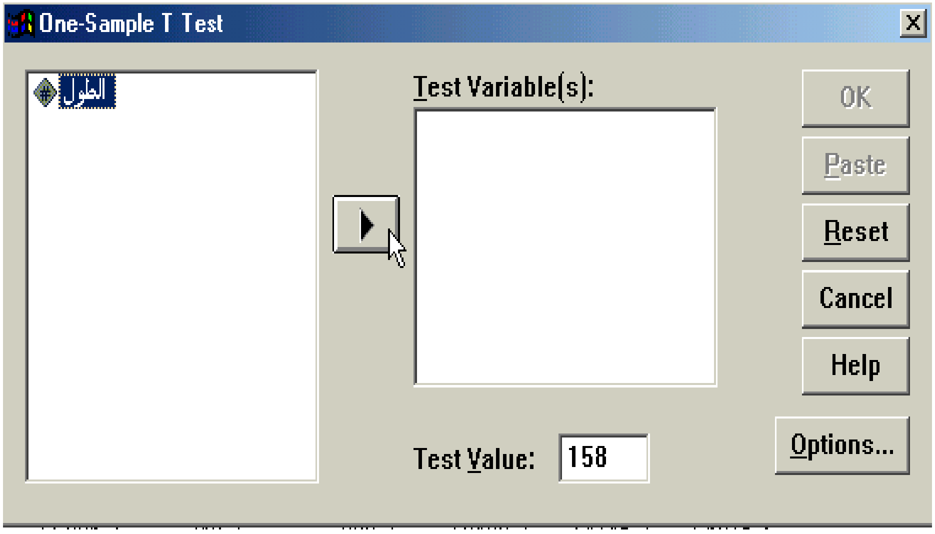
✓ قم بإدخال البيانات المراد تحليلها من خلال شاشة تحرير البيانات Data Editor بالطريقة المناسبة كالتالي :



✓ من القائمة "تحليل" Analyze اختر الأمر "مقارنة المتوسطات" Compare Means فتظهر قائمة أوامر فرعية اختر منها "اختبار (ت) لعينة واحدة" One-Sample T Test كالتالي:



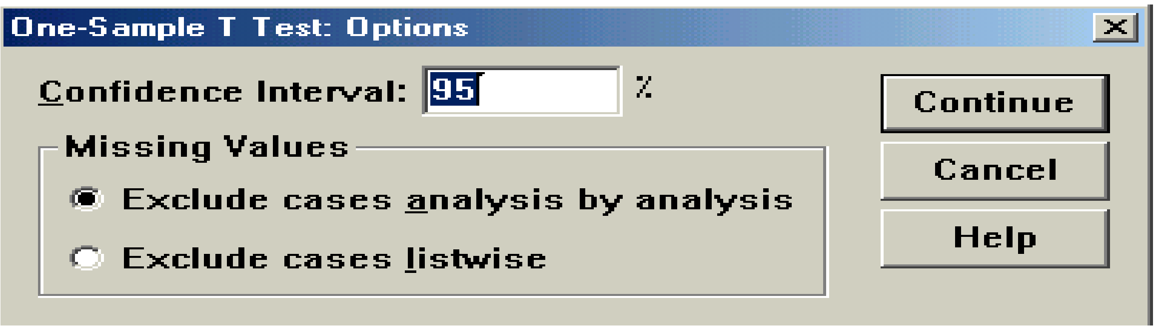
- ✓ بعد اختيار الأمر "اختبار (ت) لعينة واحدة" One-Sample T Test سوف يظهر لك صندوق الحوار التالي :



- ✓ من قائمة المتغيرات في الجهة اليسرى من صندوق الحوار انقر نقرا مزدوجا على المتغير "الطول" (أو أنقر على السهم الذي يظهر في صندوق الحوار بعد التظليل على المتغير المرغوب نقله إلى الجهة الأخرى) ستلاحظ انتقاله مباشرة في المستطيل "متغيرات الاختبار" Test Variable(s).

- ✓ في الحقل الخاص بـ "القيمة المختبرة" Test Value أكتب القيمة التي تريد أن تقارن بها متوسط العينة موضع الدراسة (في هذا المثال يتم كتابة الرقم ١٥٨ والذي يمثل متوسط أطوال الطلاب في الجامعة) .

- ✓ قم بالنقر على زر "خيارات" Options في الجهة السفلية اليمنى من صندوق الحوار السابق وذلك عند الرغبة في تغيير قيمة "فترة الثقة" Confidence Interval حيث يظهر لك صندوق الحوار التالي والذي يتيح إمكانية تغيير فترة الثقة المختبرة (بشكل تلقائي سوف تظهر القيمة ٩٥%) ، وبعد الانتهاء من التعديل على هذا الصندوق الحواري أنقر على زر "استمرار" Continue .



- ✓ أنقر بعد ذلك على زر "موافق" OK سيؤدي ذلك إلى تنفيذ الاختبار، وستلاحظ ظهور النتائج في شاشة المخرجات كالتالي :

➡ T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
الطول	250	155.9520	2.9422	.1861

One-Sample Test

	Test Value = 158					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
الطول	-11.006	249	.000	-2.0480	-2.4145	-1.6815

يتضح من النتائج أن قيمة (ت) المحسوبة t-test = -١١.٠٠٦ ، ودرجات الحرية df = ٢٤٩ ، وقيمة Sig. (2-tailed) =٠.٠٠٠ ، وبما أن قيمة Sig. (2-tailed) في الجدول (٠.٠٠٠) أصغر من قيمة α = ٠.٠٥ فإننا بالتالي نرفض الفرضية الصفرية، أي أنه توجد فروق ذات دلالة إحصائية بين متوسط أطوال العينة ومتوسط أطوال طلاب الجامعة .

الاختبارات الاحصائية لعينتين مستقلتين Independent Samples t-test

مثال :-

أراد باحث أن يعرف أثر استخدام نظم مساندة القرارات على كفاءة القرارات التي تتخذها الإدارة بمساعدة تلك النظم، فوزع ٥٠ مديرا لمنشآت صناعية عشوائيا في مجموعتين، ثم عين أحدهما بطريقة عشوائية لتكون مجموعة تجريبية والأخرى ضابطة، وفي نهاية التجربة وزع على المجموعتان استقصاء يقيس درجة فاعلية القرار وكفاءته عندما يتم اتخاذه باستخدام نظم مساندة القرارات بدلا من الطريقة التقليدية فكانت النتائج كما يلي:

المجموعة الضابطة	المجموعة التجريبية
= ٢٥	= ٢٥
= ٦.٠	= ٧.٦٠
= ١.٧٨	= ٢.٢٧

فهل تدل هذه البيانات على أن أداء المجموعة التجريبية كان أفضل من أداء المجموعة الضابطة عند مستوى α = ٠.٠٥ ؟

الحل :-

سيتم اختبار الفرضيات التالية :

الفرضية الصفرية : لا توجد فروق ذات دلالة إحصائية بين متوسط المجموعة التجريبية ومتوسط المجموعة الضابطة ($\mu_1=\mu_2$).

<u>الفرضية البديلة :</u> توجد فروق ذات دلالة إحصائية بين متوسط المجموعة التجريبية ومتوسط المجموعة الضابطة لصالح المجموعة التجريبية ($\mu_1>\mu_2$)

<u>مستوى الدلالة :</u> α = ٠.٠٥

<u>منطقة الرفض :</u> ٠.٠ قيمة مستوى الدلالة α = ٠.٠٥ والاختبار بذيل واحد ، ودرجات الحرية = ٢٥ + ٢٥ - ٢ = ٤٨ ، بذلك تكون قيمة (ت) الجدولية = ١.٦٨

<u>المختبر الإحصائى :</u>

$$t=\frac{\overline{X}_1-\overline{X}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

ولتطبيق هذه العلاقة يلزمنا حساب قيمة الانحراف المعياري (S) من خلال العلاقة التالية:

$$S^2=\frac{[(n_1-1)(S_1^2)]+[(n_2-1)(S_2^2)]}{(n_1+n_2)-2}$$

إذا التباين يساوي:

$$S^2=\frac{[(25-1)(2.27)^2]+[(25-1)(1.78)^2]}{(25+25)-2}=4.16$$

إذن الإنحراف المعياري يساوي :

$$S=\sqrt{S^2}=\sqrt{4.16}=2.04$$

ثم نحسب قيمة (ت) من خلال تطبيق العلاقة التالية :

$$t=\frac{\overline{X}_1-\overline{X}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{7.60-6.0}{2.04\sqrt{\frac{1}{25}+\frac{1}{25}}}=2.77$$

<u>القرار :</u>

٠.٠ قيمة (t) المحسوبة (٢.٧٧) أكبر من قيمـــة (ت) المجدولة (١.٦٨) عند مستوى دلالة α = ٠.٠٥ .

∴ نرفض الفرضية الصفرية ونقبل البديلة

أي أن المجموعة التي خضعت للتجربة يصبح أداؤهم أفضل في عملية اتخاذ القرار من الذين لم يخضعون للتجربة وذلك عند مستوى دلالة ٠.٠٥α .

الاختبارات الاحصائية لعينتين غير مستقلتين (العينات المرتبطة)

Paired Samples t-test

مثال :-

أراد باحث أن يعرف أثر برنامج التدريب الصيفي في الميدان على أداء الطلاب وتحصيلهم في كلية العلوم الإدارية، ولغرض تحقيق ذلك قام الباحث باختبار الطلاب قبل وبعد البرنامج التدريبي، ولكون نفس الطلاب أخذوا الاختبارين، فإن الباحث يتوقع معامل ارتباط موجب بين تحصيل الطلبة في كلا القياسين. ولغرض اختبار مدى دلالة الفروق بين الاختبار القبلي والاختبار البعدي، لابد على الباحث أن يتأكد من قيمة الارتباط بين الاختبارين والتي كانت r = 0.46 ، وقد كانت النتائج التي تم التوصل إليها كما يلي :

الإختبار القبلي	الإختبار البعدي
= ١٠٠	= ١٠٠
= ٥٤.٢٨	= ٥٨.٦٦
= ٤٩	= ٦٤

فهل تدل هذه البيانات على أن أداء الطلاب التحصيلي في الكلية بعد أخذ البرنامج التدريبي كان أفضل من أدائهم قبل أخذ البرنامج التدريبي عند مستوى α = ٠.٠٥ ؟

الحل :

سيتم اختبار الفرضيات التالية :

الفرضية الصفرية : لا توجد فروق ذات دلالة إحصائية بين متوسط تحصيل الطلاب قبل وبعد البرنامج التدريبي ($\mu_1=\mu_2$).

الفرضية البديلة : توجد فروق ذات دلالة إحصائية بين متوسط تحصيل الطلاب قبل وبعد البرنامج التدريبي ($\mu_1 \quad \mu_2$)

مستوى الدلالة : α = ٠.٠٥

منطقة الرفض : قيمة مستوى الدلالة α = ٠.٠٥ والاختبار بذيلين، ودرجات الحرية = ١٠٠-١= ٩٩ ، بذلك تكون قيمة (ت) الجدولية = ١.٩٨٠

المختبر الإحصائي :

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} - 2r\left(\frac{S_1}{\sqrt{n_1}}\right)\left(\frac{S_2}{\sqrt{n_2}}\right)}}$$

إذا قيمة (ت) تساوي :

$$t = \frac{58.66 - 54.28}{\sqrt{\frac{64.0}{100} + \frac{49.0}{100} - 2(0.46)\left(\frac{8}{\sqrt{100}}\right)\left(\frac{7}{\sqrt{100}}\right)}} = 5.57$$

في هذه المعادلة ليس هناك مانع من الابتداء بـ X_1 أو X_2 في الترتيب ، لأن الإشارة ليس لها أي تأثير على النتيجة المتحصلة

القرار :

قيمة (ت) المحسوبة (٥.٥٧) أكبر من قيمـــة (ت) المجدولة (١.٩٨٠) .عند مستوى دلالة α = ٠.٠٥ .

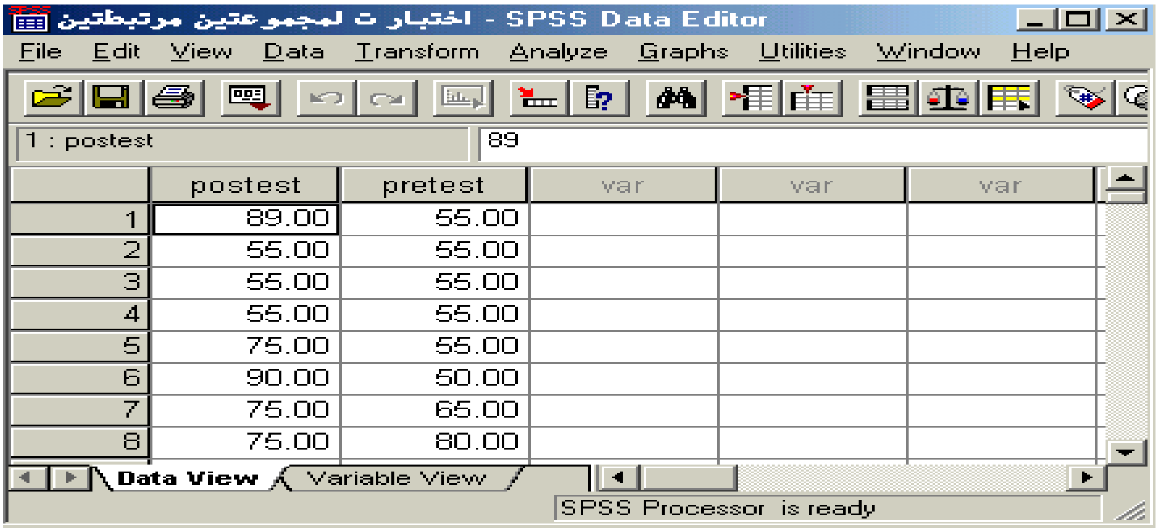
∴ نرفض الفرضية الصفرية ونقبل البديلة، أي أن للبرنامج التدريبي تأثير إيجابي على تحصيل الطلاب وأدائهم في الكلية وذلك عند مستوى دلالة α.

حساب اختبار (ت) لعينتين غير مستقلتين (العينات المرتبطة)

Paired Samples T-Test من خلال الـ SPSS

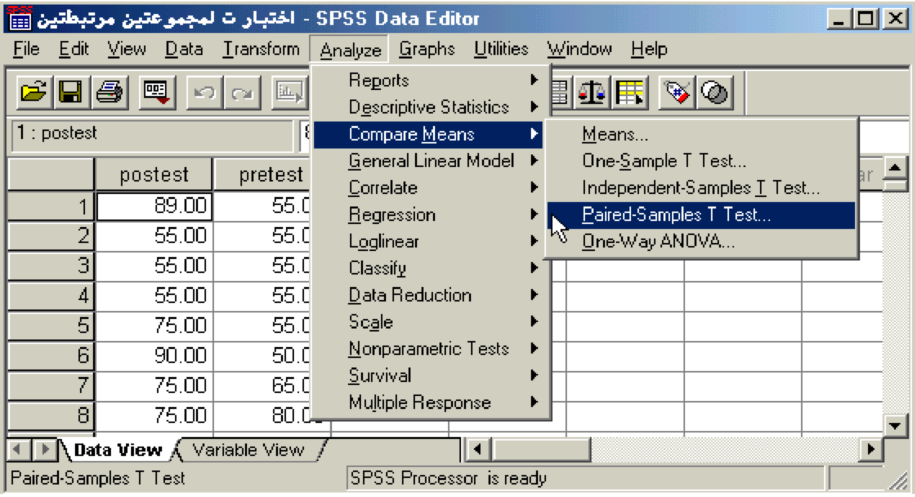
لغرض حساب قيمة (ت) لنفس المثال السابق من خلال استخدام برنامج الـ SPSS نتبع الخطوات التالية :

✓ قم بإدخال البيانات المراد تحليلها من خلال شاشة تحرير البيانات Data Editor بالطريقة المناسبة كالتالي :

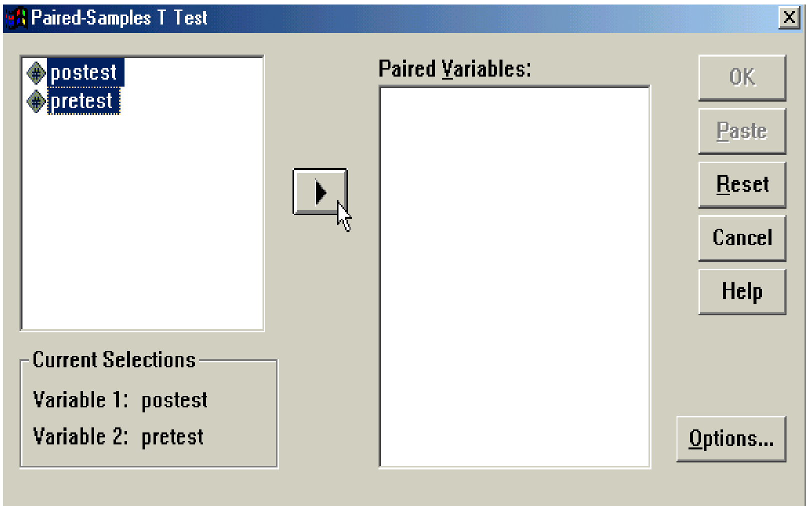


لاحظ أنه تم إدخال البيانات بطريقة مختلفة عن ما تم اتباعه في حالة العينتين المستقلتين، هنا لابد من إدخال بيانات كل متغير في عمود منفصل عن الآخر، وقد تم إعطاء كل متغير اسم مختلف عن الآخر posttest و pretest

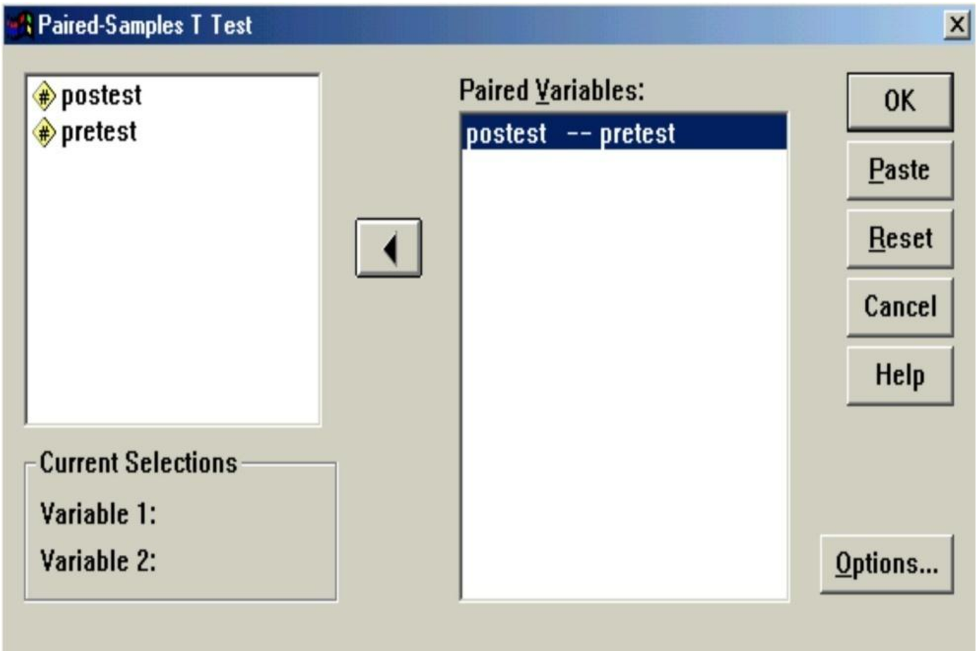
✓ من القائمة "تحليل" Analyze اختر الأمر "مقارنة المتوسطات" Compare Means فتظهر قائمة أوامر فرعية اختر منها "اختبار (ت) للعينات المرتبطة" Paired-Samples T-Testكالتالي :



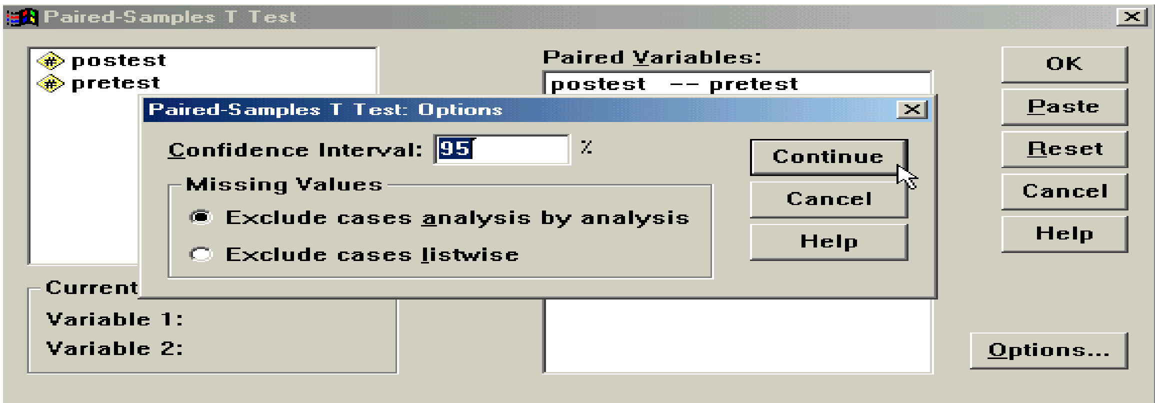
✓ بعد اختيار الأمر "اختبار (ت) للعينات المرتبطة" Paired-Samples T-Test سوف يظهر لك صندوق الحوار التالي :



✓ من قائمة المتغيرات في الجهة اليسرى من صندوق الحوار حدد المتغيرين المرتبطين مع بعضها لتحليلها كأزواج، ونقلها إلى المستطيل الخاص بـ "المتغيرات الزوجية" Paired Variables (سوف تلاحظ أثناء التحديد ظهور اسم المتغير الأول واسم المتغير الثاني بعد كل عملية تحديد في المربع اسفل قائمة المتغيرات)، ثم بعد ذلك أنقر على السهم الذي يظهر مقابل المستطيل الخاص بـ "متغيرات الاختبار" ، ستلاحظ انتقال المتغير مباشرة في المستطيل "المتغيرات الزوجية" Paired Variable(s)، كرر نفس الإجراء مع المتغيرات الزوجية الأخرى والمراد تحليلها



✓ أنقر على زر "خيارات" Options في الجهة السفلية اليمنى من صندوق الحوار السابق وذلك عند الرغبة في تغيير قيمة "فترة الثقة" Confidence Interval حيث يظهر لك صندوق الحوار التالي والذي يتيح إمكانية تغيير فترة الثقة المختبرة (بشكل تلقائي سوف تظهر القيمة ٩٥%) ، وبعد الانتهاء من التعديل على هذا الصندوق الحواري أنقر على زر "استمرار" Continue .



✓ أنقر بعد ذلك على زر "موافق" OK سيؤدي ذلك إلى تنفيذ الاختبار، وستلاحظ ظهور النتائج في شاشة المخرجات كالتالي :

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	POSTEST	58.6600	100	8.0000	.8000
	PRETEST	54.2800	100	7.0000	.7001

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	POSTEST & PRETEST	100	.458	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	POSTEST - PRETEST	4.3800	7.8570	.7857	2.8210	5.9390	5.575	99	.000

نلاحظ أن برنامج الـ SPSS قام مباشرة بحساب الإحصاءات الأساسية للبيانات مثل المتوسط الحسابي للمتغير Posttest (٥٨.٦٦٠) والانحراف المعياري لنفس المتغير (٨.٠٠) ، أما المتغير Pretest فقد كان المتوسط الحسابي (٥٤.٢٨٠) والانحراف المعياري (٧.٠٠) . بالإضافة إلى ذلك تم حساب معامل ارتباط بيرسون للمتغيرات موضع الدراسة Paired Sample Correlation وقد كانت قيمته (٠.٤٥٨) .

ثم بعد ذلك قام البرنامج بحساب قيمة (ت) للمتغيرات موضع الدراسة في الجدول المعنون بـ "اختبار العينات المرتبطة" Paired Sample Test ، ومن هذه النتائج نلاحظ أن قيمة (ت) المحسوبة t-test = ٥.٥٧٥ ، ودرجات الحرية df = ٩٩ ، وقيمة Sig. (2-tailed) =٠.٠٠٠ ، وبما أن قيمة Sig. (2-tailed) في الجدول (٠.٠٠٠) أصغر من قيمة α = ٠.٠٥ فإننا بالتالي نرفض الفرضية الصفرية ، أي أن أداء الطلاب في الكلية بعد أخذ البرنامج التدريبي كان أفضل من أدائهم قبل أخذ البرنامج التدريبي عند مستوى α = ٠.٠٥ .

المحاضرة (٩)

اختبار الفروض الإحصائية المعلمية

تحليل التباين الأحادي One Way ANOVA

تحليل التباين الأحادى (مستوى واحد) :-

هو طريقة لاختبار معنوية الفرق بين المتوسطات لعدة عينات بمقارنة واحدة، ويعرف أيضاً بطريقة تؤدي لتقسيم الاختلافات الكلية لمجموعة من المشاهدات التجريبية لعدة أجزاء للتعرف على مصدر الاختلاف بينها ولذا فالهدف هنا فحص تباين المجتمع لمعرفة مدى تساوى متوسطات المجتمع .

ولكن لا بد من تحقيق ثلاثة أمور قبل استخدامه وهي:

١) العينات عشوائية ومستقلة.

٢) مجتمعات هذه العينات كلاً لها توزيع طبيعي.

٣) تساوي تباين المجتمعات التي أخذت منها العينات العشوائية المستقلة.

ولتوضيح ما سبق بمقارنة متوسطات ثلاث مجتمعات باستخدام ثلاث عينات (تحقق فيها الشروط الثلاثة السابقة) موضحة بالجدول الآتي:

مثال (١) : -

إذا كان لدينا ثلاث منتجات لإحدى الشركات الصناعية ، وتم تقييمها من قبل مجموعة من المستهلكين وحصلنا على النتائج التالية :

المنتج (١) X_1	المنتج (٢) X_2	المنتج (٣) X_3
٧	٤	٢
١٠	٦	٢
١٠	٧	٣
١١	٩	٧
١٢	٩	٦
٥٠	٣٥	٢٠

المطلوب: هل هناك فروق ذات دلالة بين المنتجات الثلاثة ؟

الحل :

- ❑ **وضع فرض العدم والفرض البديل.**

صياغة الفرضية الصفرية كالتالي:

H0: ١μ = ٢μ = ٣μ

في حين تفترض الفرضية البديلة التالي :

متوسطان على الأقل غير متساويين :HA

- **تحديد مستوى الدلالة ( ): وتحدد مستويات المعنوية سلفاً وهي عادة 0.05 أو 0.01**
- **حساب إحصائية الاختبار (F) وذلك من خلال اتباع الخطوت التالية:**
- **وضع فرض العدم والفرض البديل.**

صياغة الفرضية الصفرية كالتالي:

H0: ١μ = ٢μ = ٣μ

في حين تفترض الفرضية البديلة التالي :

متوسطان على الأقل غير متساويين :HA

- **تحديد مستوى الدلالة (α): وتحدد مستويات المعنوية سلفاً وهي عادة 0.05 أو 0.01**
- **حساب إحصائية الاختبار (F) وذلك من خلال اتباع الخطوت التالية:**

المنتج (١) X_1		المنتج (٢) X_2		المنتج (٣) X_3	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
٧	٤٩	٤	١٦	٢	٤
١٠	١٠٠	٦	٣٦	٢	٤
١٠	١٠٠	٧	٤٩	٣	٩
١١	١٢١	٩	٨١	٧	٤٩
١٢	١٤٤	٩	٨١	٦	٣٦
٥٠	٥١٤	٣٥	٢٦٣	٢٠	١٠٢

- ✓ **المتوسط الحسابي لـ** $\bar{X_1}$ $\bar{X} = \frac{50}{5} = 10$
- ✓ **المتوسط الحسابي لـ** $\bar{X_2}$ $\bar{X} = \frac{35}{5} = 7$
- ✓ **المتوسط الحسابي لـ** $\bar{X_3}$ $\bar{X} = \frac{20}{5} = 4$

١- مجموع المربعات الكلي Total Sum of Squares =

حيث ng تعني عدد أفراد المجموعة المحددة

k تعني عدد المجموعات موضع الدراسة

٢- مجموع المربعات بين المجموعات Between Sum of Squares = $Beween..groups..mean..square = \frac{90}{2} = 45$

$$Between..SS = \sum \frac{\left(\sum X_g\right)^2}{n_g} - \frac{\left(\sum X\right)^2}{(n_g)(k)} = \frac{(50)^2}{5} + \frac{(35)^2}{5} + \frac{(20)^2}{5} - \frac{(105)^2}{15} = 90$$

✓ مجموع المربعات داخل المجموعات Within Sum of Squares =

٣- مجموع المربعات داخل المجموعات *Within SS*=

= مجموع المربعات الكلي – مجموع المربعات بين المجموعات

= ١٤٤-٩٠=٥٤

✓ نحسب درجات الحرية:

درجات الحرية بين المجموعات Between groups

degrees of freedom

$$(K-1) = 3-1 = 2$$

درجات الحرية داخل المجموعات Within groups

degrees of freedom

$$(n-K) = 15-3 = 12$$

درجات الحرية الكلية Total degrees of freedom

$$(n-1) = 15-1 = 14$$

✓ التباين بين المجموعات أو ما يسمى متوسط المربعات بين المجموعات Between mean square =

$$Beween..groups..mean..square = \frac{Between..SS}{K-1}$$

$$Total..SS = \sum X^2 - \frac{\left(\sum X\right)^2}{(n_g)(k)} = 879 - \frac{(105)^2}{15} = 144$$

✓ التباين داخل المجموعات أو ما يسمى متوسط المربعات داخل المجموعات Within mean square =

$$Within..groups..mean..square = \frac{Within..SS}{(n-K)}$$

$$Within..groups..mean..square = \frac{54}{12} = 4.5$$

✓ قيمة F =
$$F = \frac{Between..groups..mean..square}{Within..groups..mean..square} = \frac{45}{4.5} = 10$$

✓ نقوم بعد ذلك بتفريغ ما تم الحصول عليه من معلومات في جدول تحليل التباين كالتالي :

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات Means	قيمة F
بين المجموعات Between groups	٩٠	٢	٤٥	١٠
داخل المجموعات Within groups	٥٤	١٢	٤.٥	
الكلي (المجموع) Total	١٤٤	١٤		

بالرجوع إلى جدول توزيع *F* نجد أن القيمة الحرجة لـ *F* بدرجات حرية للبسط تساوي ٢ ودرجات حرية للمقام تساوي ١٢ وباستخدام مستوى= ٠.٠٥ نجد أن القيمة الحرجة تساوي ٣.٨٨،

وحيث أن القيمة المحسوبة لـ *F* = ١٠ وهي بالتالي أكبر من القيمة الحرجة المجدولة، نستنتج أن الفرضية الصفرية تكون مرفوضة، أي يوجد اختلاف بين متوسطي مجتمعين على الأقل من المجتمعات التي قيّمة من المستهلكين ولمعرفة بين أي من المنتجات تكون الفروق ينبغي علينا اللجوء إلى أسلوب المقارنات المتعددة Multiple Comparisons .

مثال (٢) : -

قام أحد الباحثين بتفريغ ما تم الحصول عليه من معلومات في جدول تحليل التباين كالتالي :

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات Means	قيمة F
بين المجموعات Between groups	200	10	.....	
داخل المجموعات Within groups	......	......	....	....
الكلي (المجموع) Total	250	20		

<u>**قيمة إحصائي الاختبار F تساوى :-**</u>

(أ)10

(ب)5

(ج)80

<u>**(د) لا شيء مما سبق**</u>

<u>**من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض (إذا علمت أن قيمة Fالجدولية تساوى 6.88) يمكن :-**</u>

(أ) قبول الفرض البديل .

<u>**(ب) قبول الفرض العدمي .**</u>

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

<u>**مثال (٣) : -**</u>

إذا كانت لدينا البيانات التالية والتي تمثل بيانات أربع مجموعات تم استطلاع آراؤها حول موضوع ما:

A	B	C	D
8	4	5	6
7	3	3	5
9	6	4	6
5	5	5	4
6	2	3	3
7	7	2	4
42	27	22	28

المطلوب:

هل هناك فروق بين آراء هذه المجموعات الأربع ولصالح من هذا الفرق؟

الحل :-

- وضع فرض العدم والفرض البديل.

صياغة الفرضية الصفرية كالتالي:

H0: µ1 = µ2 = µ3 = µ4

في حين تفترض الفرضية البديلة التالي :

متوسطان على الأقل غير متساويين :HA

- **تحديد مستوى الدلالة (α): وتحدد مستويات المعنوية سلفاً وهي عادة إما 0.05 أو 0.01 وليكون α = 0.05**
- **حساب إحصائية الاختبار (F) وذلك من خلال اتباع الخطوت التالية:**

✓ **نجمع قيم كل متغير للحصول على $\sum X$ لكل متغير .**

✓ **نربع كل درجة في كل متغير للحصول على X^2 لكل متغير .**

✓ **نجمع قيم مربع كل درجة للحصول على $\sum X^2$ لكل متغير .**

✓ **نربع مجموع كل متغير للحصول على $(\sum X)^2$ لكل متغير .**

✓ **نحسب متوسط كل متغير من خلال العلاقة : $\overline{X} = \frac{\sum x}{n}$**

	A	A^2	B	B^2	C	C^2	D	D^2	
	8	64	4	16	5	25	6	36	
	7	49	3	9	3	9	5	25	
	9	81	6	36	4	16	6	36	
	5	25	5	25	5	25	4	16	
	6	36	2	4	3	9	3	9	
	7	49	7	49	2	4	4	16	
	42		27		22		28		119
		304		139		88		138	669
	1764		729		484		784		3761
	42/6= 7		27/6 =4.5		22/6 =3.7		28/6 =4.7		

١- مجموع المربعات الكلي Total Sum of Squares =

$$Total..SS = \sum X^2 - \frac{\left(\sum X\right)^2}{(n_g)(k)} = 669 - \frac{(119)^2}{24} = 669 - 590.042 = 78.958$$

٢- مجموع المربعات بين المجموعات =

$$\text{Between SS} = \sum\left[\sum \frac{\left(\sum X_g\right)^2}{n_g}\right] - \frac{\left(\sum X\right)^2}{n.k}$$

$$= \frac{(42)^2 + (27)^2 + (22)^2 + (28)^2}{6} - 590.042 = 36.8$$

٣- مجموع المربعات داخل المجموعات *Within SS*=

= مجموع المربعات الكلي – مجموع المربعات بين المجموعات

= ٧٨.٩ – ٣٦.٨ = ٤٢.١

ثم نقوم بعمل جدول تحليل التباين كالتالي:

Source of Variance	SS	d f	MS	F	$f_{0.05,3,20}$
بين المجموعات	SST – SSW = 36.8	K – 1 = 4 – 1 = 3	12.264	5.818	3.10
داخل المجموعات	42.1	K(n – 1) = 4(6 – 1) = 20	2.108		
TOTAL	78.958	Kn – 1 = 24 – 1 = 23			

قيمة F المحسوبة أكبر من قيمة F الجدولية (٥.٨١٨ > ٣.١) فالمتوسطات غير متساوية

أي قبول الفرض البديل القائل بعدم تساوي المتوسطات .

مثال (٤) : -

للمقارنة بين أربعة أنواع من القمح (A ، B ، C ، D) بغرض تعميم أفضلها في الزراعة تم زراعة كل نوع في حوض تجريبي بحيث كانت الاحواض الاربعة متماثلة في الخصوبة ونوع التربة وكمية الاسمدة المستعملة ومنسوب المياة ودرجة الحرارة وكانت انتاجية الفدان من كل نوع كما يلي :-

D	C	B	A
2	7	6	1
6	8	5	3
4	6	7	5

والمطلوب :-

إختبار معنوية الفروق بين الانواع الاربعة من القمح بدرجة ثقة 95% .

الحل :-

	A	A^2	B	B^2	C	C^2	D	D^2	
	1	1	6	36	7	49	2	4	
	3	9	5	25	8	64	6	36	
	5	25	7	49	6	36	4	16	
	9		18		21		12		60
		35		110		149		56	350
	81		324		441		144		990
	9 /3 =3		18/3=6		21/3=7		12/3=4		

Ho : $\mu_A = \mu_B = \mu_C = \mu_D$

H1 : $\mu_A \neq \mu_B \neq \mu_C \neq \mu_D$

عدد المجموعات k = 4

عدد مفردات المجموعة الواحدة n = 3

العدد الكلى للمفردات = 3 × 4 = 12

١- مجموع المربعات الكلي Total Sum of Squares =

$$Total..SS = \sum X^2 - \frac{\left(\sum X\right)^2}{(n_g)(k)} = 350 - \frac{(60)^2}{12} = 350 - 300 = 50$$

٢-مجموع المربعات بين المجموعات =

$$\text{Between SS} = \sum\left[\sum \frac{\left(\sum X_g\right)^2}{n_g}\right] - \frac{(\sum X)^2}{n.k}$$

$$= \frac{(9)^2 + (18)^2 + (21)^2 + (12)^2}{3} - 300 = 30$$

البرنامج الأول	البرنامج الثاني	البرنامج الثالث

٣- مجموع المربعات داخل المجموعات *Within SS*=

= مجموع المربعات الكلي – مجموع المربعات بين المجموعات

= ٥٠ - ٣٠ = ٢٠

ثم نقوم بعمل جدول تحليل التباين كالتالي:

Source of Variance	SS	d f	MS	F	$f_{0.05,3,8}$
بين المجموعات	SST – SSW = 30	K – 1 = 4 – 1 = 3	10	4	4.07
داخل المجموعات	20	K(n – 1) = 4(3 – 1) = 8	2.5		
TOTAL	50	Kn – 1 = 12 – 1 = 11			

قيمة Fالمحسوبة أقل من قيمة F الجدولية (٤.٠٧ < ٤) فالمتوسطات متساوية

أي قبول الفرض العدمي القائل بتساوي المتوسطات ، وعلى ذلك فلا توجد فروق معنوية بين متوسطات إنتاجية الفدان للأنواع الأربعة من القمح والفروق الموجودة بينهم ترجع إلى عوامل الصدفة .

مثال (٥) : -

طبقت ثلاثة برامج مختلفة للتدريب على ثلاث مجموعات من اللاعبين : الأولى تضم أربعة أفراد و الثانية تضم ستة أفراد و الثالثة تضم خمسة أفراد و في نهاية فترة التدريب أجرى لهم اختبار وكان عدد الأهداف المسجلة لكل لاعب كما يلي :-

اختبر ما إذا كانت هناك فروقاً معنوية بين

برامج التدريب الثلاث بمستوى معنوية 5% .

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

الحل : -

$H_o : \mu_1 = \mu_2 = \mu_3$

$H1 : \mu_1 \neq \mu_2 \neq \mu_3$

عدد المجموعات k = 3

العدد الكلى للمفردات = ٤ + ٦ + ٥ = 15

	A	A^2	B	B^2	C	C^2	
	5	25	6	36	7	49	
	6	36	2	4	6	36	
	7	49	4	16	8	64	
	6	36	5	25	9	81	
			3	9	5	25	
			4	16			
	24		24		35		83
		146		106		255	507

	576		576		1225		2377
	24 /4 =6		24/6=4		35/5=7		

١- مجموع المربعات الكلي Total Sum of Squares =

$$Total..SS = \sum X^2 - \frac{\left(\sum X\right)^2}{(n_g)(k)} = 507 - \frac{(83)^2}{15} = 507 - 459.267 = 47.73$$

٢-مجموع المربعات بين المجموعات =

$$\text{Between SS} = \sum\left[\sum \frac{\left(\sum X_g\right)^2}{n_g}\right] - \frac{\left(\sum X\right)^2}{n.k}$$

$$= \left(\frac{(24)^2}{4} + \frac{(24)^2}{6} + \frac{(35)^2}{5}\right) - 459.267 = 25.73$$

٣- مجموع المربعات داخل المجموعات ***Within SS***=

= مجموع المربعات الكلي – مجموع المربعات بين المجموعات

= ٤٧.٧٣ – ٢٥.٧٣ = ٢٢

ثم نقوم بعمل جدول تحليل التباين كالتالي:

Source of Variance	SS	d f	MS	F	$f_{0.05,2,12}$
بين المجموعات	25.73	2	12.865	7.018	3.89
داخل المجموعات	22	12	1.833		
TOTAL	47.73	14			

قيمة Fالمحسوبة أكبر من قيمة F الجدولية (٧.١٨ > ٣.٨٩) فالمتوسطات غير متساوية

أي قبول الفرض البديل القائل بعدم تساوي المتوسطات مما يعنى وجود فروق معنوية بين برامج التدريب الثلاثة وذلك بدرجة ثقة 95%.

مثال (٦) : -

في تجربة لمقارنة ٣ مجموعات تحتوى كل منها على ٥ مفردات حصلنا على النتائج التالية :-

مجموع المربعات الكلي = ١٧٦

مجموع المربعات بين المجموعات = ١٠٤

فالمطلوب :-

١- اختبار معنوية الفروق بين متوسطات المجتمعات التي سحبت منها العينات بمستوى معنوية 5% .

٢- إذا أظهر الاختبار وجود فروق معنوية بين المتوسطات فالمطلوب تحليل معنوية هذه الفروق .

الحل :-

$$Ho : \mu_A = \mu_B = \mu_C = \mu_D$$

$$H1 : \mu_A \neq \mu_B \neq \mu_C \neq \mu_D$$

عدد المجموعات k = 3

عدد مفردات المجموعة الواحدة n = 5

العدد الكلى للمفردات = 3 × 5 = 15

مجموع المربعات الكلي = ١٧٦

مجموع المربعات بين المجموعات = ١٠٤

مجموع المربعات داخل المجموعات = ١٧٦ – ١٠٤ = ٧٢

ثم نقوم بعمل جدول تحليل التباين كالتالي:

Source of Variance	SS	d f	MS	F	$f_{0.05,2,12}$
بين المجموعات	104	2	52	8.67	3.89
داخل المجموعات	72	12	6		
TOTAL	176	14			

قيمة F المحسوبة أكبر من قيمة F الجدولية (٨.٦٧ > ٣.٨٩) فالمتوسطات غير متساوية

أي قبول الفرض البديل القائل بعدم تساوي المتوسطات .

المحاضرة (١٠)

إختبار الفروض الإحصائية المعلمية

معامل الارتباط: هو تعبير يشير إلى المقياس الإحصائي الذي يدل على مقدار العلاقة بين المتغيرات سلبية كانت أم إيجابية، وتتراوح قيمته بين الارتباط الموجب التام (+١) وبين الارتباط السالب التام (-١) .

العلاقة الطردية بين المتغيرات: هو تعبير يشير إلى تزايد المتغيرين المستقل والتابع معا، فإذا كانت الإنتاجية مرتفعة، ومستوى الجودة مرتفع، يقال حينئذ أن بينهما ارتباط موجب، وأعلى درجة تمثله هي (+١) .

العلاقة العكسية بين المتغيرات: هو تعبير يشير إلى تزايد في متغير يقابله تناقص في المتغير الآخر، فإذا كانت الإنتاجية منخفضة ومستوى الجودة مرتفع، يقال حينئذ أن بينهما ارتباط سالب، وأعلى درجة تمثله هي (-١) .

الإرتباط الجزئي Partial Correlation

هو عبارة عن مقياس لقوة واتجاه الارتباط بين متغيرين كميين بعد استبعاد اثر متغير كمي ثالث ، حيث يلاحظ انه بالرغم من ان قيمة معامل الارتباط بيرسون قد تكون كبيرة ولكن لا يمكن الاعتماد عليها لكونه يعتمد في قياسه على متغيرين فقط ، فقد يوجد متغير ثالث يؤثر في المتغيرين ولهذا برزت اهمية معامل الارتباط الجزئي.

فمثلا يمكن قياس قوة الارتباط بين مستوى الطلبة في الجامعات والبيئة الجامعية بعد استبعاد عدد ساعات الدراسة لكل طالب .

ويتم حساب الإرتباط الجزئي من خلال حساب الإرتباطات الثنائية بين متغيرات الدراسة (على الباحث أن يستخدم معامل الإرتباط المناسب لعدد العينة ولطبيعة توزيع المتغيرات).

أي أن بإمكان الباحث استخدام معامل ارتباط بيرسون أو معامل ارتباط سبيرمان أو غير ذلك من معاملات الارتباط تبعا كما ذكر لطبيعة توزيع متغيرات الدراسة.

مثال:

أراد باحث أن يدرس العلاقة بين التحصيل الدراسي والغياب لدى مجموعة من الطلبة، ومن المعروف أنه إلى جانب الغياب فإن طريقة التدريس للطالب تؤثر في تحصيله الدراسي أيضا، فإذا استطاع الباحث أن يضبط هذا المتغير (المتغير الخاص بطريقة التدريس) أثناء إجرائه للتجربة، ويختار الطلبة من بين الذين يتعلمون بطريقة تدريس واحدة فإنه يكون بذلك قد عزل تأثير هذا المتغير.

أما إذا لم يستطع الباحث اختيار الطلبة من الذين يخضعون لطريقة تدريس واحدة، وكان الطلبة يتلقون تدريسهم وفقا لطرق تدريس مختلفة، فإنه بذلك يكون في حاجة لمعامل الإرتباط الجزئي لكي يعزل تأثير متغير طريقة التدريس في العلاقة بين التحصيل الدراسي والغياب، والبيانات التالية توضح هذا المثال:

الطلبة	الغياب (١)	التحصيل (٢)	طريقة التدريس (٣)
١	٧٠	١٥	١٣
٢	١١٠	١٣	٢٠
٣	١٢٠	١١	٥٥
٤	٩٥	١٣	٨٠
٥	١٠٥	٠٨	٠٦

المطلوب:

١. **حساب معامل الارتباط بين الغياب والتحصيل مع تثبيت طريقة التدريس؟**

الحل:

لغرض حساب معامل الارتباط بين الغياب والتحصيل مع تثبيت طريقة التدريس لا بد من حساب معاملات الإرتباط بين المتغيرات الثلاثة السابقة كالتالي:

- **معامل ارتباط بيرسون بين الغياب والتحصيل الدراسي ونرمز له ١.٢ أي معامل الإرتباط بين المتغير (١) والمتغير (٢)**
- **معامل ارتباط بيرسون بين الغياب وطريقة التدريس ونرمز له ١.٣ أي معامل الإرتباط بين المتغير (١) والمتغير (٣)**
- **معامل ارتباط بيرسون بين التحصيل الدراسي وطريقة التدريس ونرمز له ٢.٣ أي معامل الإرتباط بين المتغير (٢) والمتغير (٣)**

ويتم حساب معامل ارتباط بيرسون من خلال العلاقـة التالية:

$$r = \frac{\sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{n}}{\sqrt{\left(\sum X^2 - \frac{\left(\sum X\right)^2}{n}\right)\left(\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right)}}$$

حيث :

تعني مجموع حاصل ضرب كل قيمة من X في Y.

تعني مجموع قيم المتغير X .

تعني مجموع قيم المتغير Y .

تعني مجموع مربع قيم المتغير X .

تعني مربع مجموع قيم المتغير X .

تعني مجموع مربع قيم المتغير Y .

تعني مربع مجموع قيم المتغير Y .

عدد قيم الدراسة (عدد الأزواج المطلوب حساب الارتباط بينها) .

معامل ارتباط بيرسون بين الغياب والتحصيل الدراسي ونرمز له $r_{١.٢}$

الطلبة	الغياب (١) X	التحصيل الدراسي (٢) Y	X^2	Y^2	XY
١	٧٠	١٥	٤٩٠٠	٢٢٥	١٠٥٠
٢	١١٠	١٣	١٢١٠٠	١٦٩	١٤٣٠
٣	١٢٠	١١	١٤٤٠٠	١٢١	١٣٢٠
٤	٩٥	١٣	٩٠٢٥	١٦٩	١٢٣٥
٥	١٠٥	٨	١١٠٢٥	٦٤	٨٤٠
المجموع	٥٠٠	٦٠	٥١٤٥٠	٧٤٨	٥٨٧٥

$$r_{1.2}=\frac{5875-\frac{(500)(60)}{5}}{\left[\sqrt{(51450)-\frac{(500)^2}{5}}\right]\left[\sqrt{(748)-\frac{(60)^2}{5}}\right]}=\frac{5875-6000}{\left(\sqrt{51450-50000}\right)\left(\sqrt{748-720}\right)}$$

$$=\frac{-125}{\left(\sqrt{1450}\right)\left(\sqrt{28}\right)}=\frac{-125}{(38.08)(5.292)}=\frac{-125}{201.519}=-0.620$$

معامل ارتباط بيرسون بين الغياب وطريقة التدريس

ونرمز له $r_{١.٣}$

XY	Y^2	X^2	طريقة التدريس (٣) Y	الغياب (١) X	الطلبة
٩١٠	١١٦٩	٤٩٠٠	١٣	٧٠	١
٢٢٠٠	٤٠٠	١٢١٠٠	٢٠	١١٠	٢
٦٦٠٠	٣٠٢٥	١٤٤٠٠	٥٥	١٢٠	٣
٧٦٠٠	٦٤٠٠	٩٠٢٥	٨٠	٩٥	٤
٦٣٠	٣٦	١١٠٢٥	٦	١٠٥	٥
١٧٩٤٠	١٠٠٣٠	٥١٤٥٠	١٧٤	٥٠٠	المجموع

$$r_{1.3}=\frac{17940-\frac{(500)(174)}{5}}{\left[\sqrt{(51450)-\frac{(500)^2}{5}}\right]\left[\sqrt{(10030)-\frac{(174)^2}{5}}\right]}=\frac{17940-17400}{\left(\sqrt{51450-50000}\right)\left(\sqrt{10030-6055.2}\right)}$$

$$=\frac{540}{\left(\sqrt{1450}\right)\left(\sqrt{3974.8}\right)}=\frac{540}{(38.08)(63.046)}=\frac{540}{2400.73}=+0.225$$

معامل ارتباط بيرسون بين التحصيل الدراسي وطريقة التدريس

ونرمز له $r_{٢.٣}$

XY	Y^2	X^2	طريقة التدريس (٣) Y	التحصيل الدراسي (٢) X	الطلبة
١٩٥	١٦٩	٢٢٥	١٣	١٥	١
٢٦٠	٤٠٠	١٦٩	٢٠	١٣	٢
٦٠٥	٣٠٢٥	١٢١	٥٥	١١	٣
١٠٤٠	٦٤٠٠	١٦٩	٨٠	١٣	٤
٤٨	٣٦	٦٤	٦	٨	٥
٢١٤٨	١٠٠٣٠	٧٤٨	١٧٤	٦٠	المجموع

$$r_{2.3}=\frac{2148-\frac{(60)(174)}{5}}{\left[\sqrt{(748)-\frac{(60)^2}{5}}\right]\left[\sqrt{(10030)-\frac{(174)^2}{5}}\right]}=\frac{2148-2088}{\left(\sqrt{748-720}\right)\left(\sqrt{10030-6055.2}\right)}$$

$$=\frac{60}{\left(\sqrt{28}\right)\left(\sqrt{3974.8}\right)}=\frac{60}{(5.292)(63.046)}=\frac{60}{333.639}=+0.179$$

بعد حساب معامل الارتباط الثنائي المناسب نقوم بعدها بتطبيق قانون معامل الارتباط الجزئي كالتالي:

$$r_{1.2.3} = \frac{(r_{1.2}) - [(r_{1.3})(r_{2.3})]}{\sqrt{[1-(r_{1.3})^2][1-(r_{2.3})^2]}}$$

$$r_{1.2.3} = \frac{(-.620) - [(.225)(.179)]}{\sqrt{[1-(.225)^2][1-(.179)^2]}}$$

$$= \frac{(-.620) - (.0402)}{\sqrt{(1-0.0506)(1-0.032)}}$$

$$= \frac{-0.662}{\sqrt{(0.9494)(0.968)}} = \frac{-0.662}{\sqrt{0.919}}$$

$$= \frac{-0.660}{0.9586} = -0.689$$

مثال :-

يقوم أحد الباحثين بدراسة العلاقة بين ثلاث من الظواهر و هي A و B و C و وجد أن الارتباط بين كل من الظاهرة الأولى A و الظاهرة الثانية B يساوي (0.62) و الارتباط بين الظاهرة الأولى و الثالثة يساوي -) 0.225) و الارتباط بين كل من الظاهرة الثانية و الثالثة تساوي (0.179) ، فالمطلوب تقدير قيمة الارتباط الجزئي بين كل من هذه الظواهر .

الحل :

بعد حساب معامل الارتباط الثنائي المناسب نقوم بعدها بتطبيق قانون معامل الارتباط الجزئي كالتالي:

$$r_{1.2.3} = \frac{(r_{1.2}) - [(r_{1.3})(r_{2.3})]}{\sqrt{[1-(r_{1.3})^2][1-(r_{2.3})^2]}}$$

$r_{1.2} = \mathbf{0.62}$

$r_{1.3} = \mathbf{-0.225}$

$r_{2.3} = \mathbf{0.179}$

$$r_{1.2.3} = \frac{(.620) - [(-.225)(.179)]}{\sqrt{[1-(-.225)^2][1-(.179)^2]}}$$

$$= \frac{(.620) + (.0402)}{\sqrt{(1-0.0506)(1-0.032)}}$$

$$= \frac{0.662}{\sqrt{(0.9494)(0.968)}} = \frac{0.662}{\sqrt{0.919}}$$

$$= \frac{0.660}{0.9586} = 0.689$$

اختبار معنوية معامل الارتباط Significance Of Correlation Coefficient

إذا كانت قيمة معامل ارتباط العينة r قريبة من + ١ أو – ١ فإن هناك علاقة خطية قوية بين المتغيرين، وإذا كانت r = 0 فإنه لا توجد علاقة خطية بينهما، أما إذا كانت قيم r متوسطة فإنه يجب اختبار معنوية (أو دلالة) معامل ارتباط العينة، وهل هناك ارتباط حقيقي بين المتغيرين في المجتمع، أم أن الارتباط بينهما زائف وغير حقيقي.

وفيما يلي نتناول بالتفصيل اختبار معنوية معامل ارتباط المجتمع والذي نرمز له بالرمز R.

اختبار أن معامل ارتباط المجتمع يساوي الصفر:

بافتراض أن المجتمع له توزيع طبيعي فإن معامل ارتباط العينة r يكون له توزيع t بوسط حسابي يساوي R وانحراف معياري يساوي $\sqrt{\frac{1-r^2}{n-2}}$ وذلك بدرجات حرية n – 2. وبالتالي تكون خطوات اختبار أن معامل ارتباط المجتمع يساوي صفر كما يلي

١ – الفرض العدمي : أن معامل ارتباط المجتمع يساوي صفر، أي لا يوجد ارتباط بين المتغيرين. وبالرموز :

$$H_0 : R = 0$$

٢ – الفرض البديل : معامل ارتباط المجتمع لا يساوي صفر، أي ويجد ارتباط بين المتغيرين، وبالرموز :

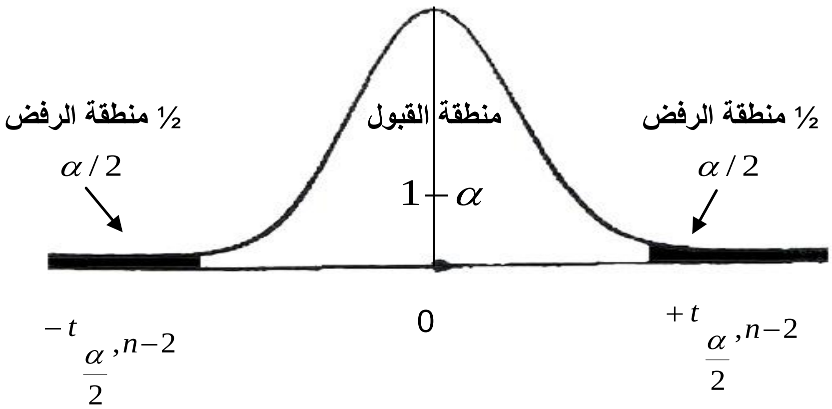
$$H_A : R \neq 0$$

٣ – إحصائية الاختبار : ستكون إحصائية الاختبار في هذه الحالة هي t والتي تأخذ الشكل التالي :

$$T = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

والتي لها توزيع t بدرجات حرية n – 2.

٤ – حدود منطقتي القبول والرفض : والتي نحصل عليها من جدول t لمستوى معنوية يساوي ودرجات حرية تساوي n – 2 (اختبار الطرفين) :



0

- **المقارنة والقرار :** حيث نقارن قيمة إحصائية الاختبار (المحسوبة في الخطوة رقم ٣) بحدود منطقتي القبول والرفض (من الخطوة رقم ٤). فإذا وقعت قيمة الإحصائية في منطقة القبول فإن القرار هو قبول الفرض العدمي بأن R = 0 أي لا يوجد ارتباط بين المتغيرين والعكس إذا وقعت قيمة الإحصائية في منطقة الرفض فإن القرار هو رفض الفرض العدمي، وفي هذه الحالة نقبل الفرض البديل بأن هناك ارتباط بين المتغيرين وذلك بمستوى معنوية يساوي .

مثال:

اختبر معنوية معامل الارتباط لوكان لدينا البيانات التالية :

n = 10 ، r = 0.91

وذلك بمستوى معنوية % 5.

الحل :

لوكان لدينا البيانات التالية :

n = 10 ، r = 0.91

وتكون خطوات اختبار معنوية الارتباط كما يلي :

١ – الفرض العدمي : $H_0 : R = 0$

٢ – الفرض البديل : $H_A : R \neq 0$

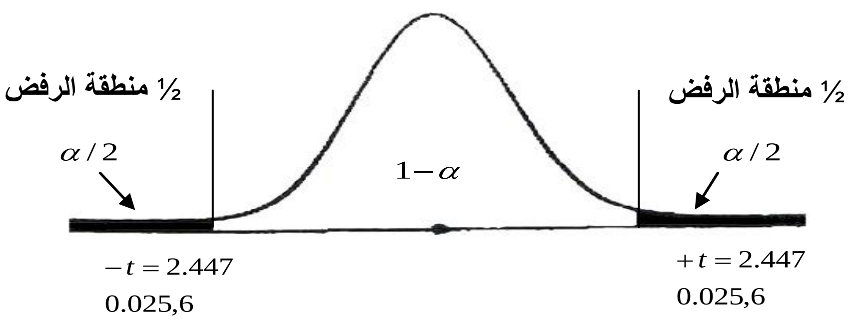
– الإحصائية :

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.91}{\sqrt{\frac{1-(0.91)^2}{10-2}}} = \frac{0.91}{\sqrt{\frac{0.1719}{8}}} = \frac{0.91}{\sqrt{0.0215}} = \frac{0.91}{0.1466} = 6.208$$

إذا : $t = 6.208$

٤ – حدود منطقتي القبول والرفض:

من جدول t حيث مستوى المعنوية $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$ **ودرجات الحرية تساوي (n – 2 = 10 – 2=8 نجد أن قيمة t تساوي ٢.٤٤٧ وتكون حدود منطقتي القبول والرفض كما يلي:**



٥ – المقارنة والقرار : بمقارنة قيمة الإحصائية المحسوبة في الخطوة رقم ٣ والتي تساوي **6.2074** بحدود منطقتي القبول والرفض (أو قيم t الجدولية في الخطوة رقم ٤) نجد أنها تقع في منطقة الرفض (حيث أنها أكبر من **2.447**) لذلك فإن القرار هو : رفض الفرض العدمي وقبول الفرض البديل. أي رفض أن معامل الارتباط يساوي صفر. وقبول أن معامل الارتباط لا يساوي صفر أي يوجد ارتباط بين المتغيرين (أعمار الناخبين و دخولهم اليومية) وذلك بمستوى معنوية 5 %.

مثال :-

" أن معامل الإرتباط بين ثلاث ظواهر إقتصادية قد بلغت (r = 0.21) و كان عدد المفردات التى تم دراستها (n = 10)، وقد رغب الباحث فى دراسة معنوية الارتباط و ذلك بمستوى 5% "

(1) قيمة إحصائي الاختبار t تساوي :-

(أ) 0.6075

(ب) -0.6075

(ج) 6.208

(د) لا شيء مما سبق

(٢) إذا علمت أن حدود منطقتي القبول و الرفض هي (2,447- , 2.447)فعلى ذلك يمكن :-

(أ) قبول الفرض العدمي.

(ب) رفض الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

الحل :

لوكان لدينا البيانات التالية :

n = 10 ، r = 0.21

وتكون خطوات اختبار معنوية الارتباط كما يلي :

١ – الفرض العدمي : $H_0 : R = 0$

٢ – الفرض البديل : $H_A : R \neq 0$

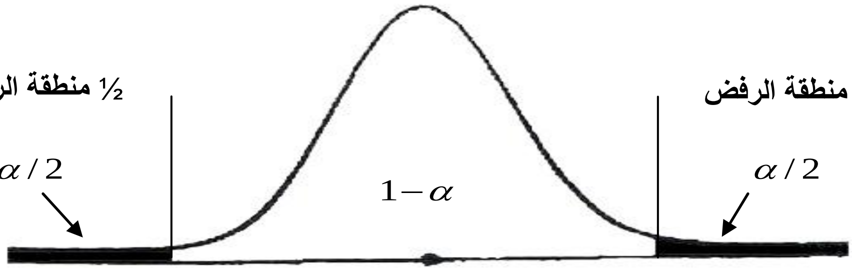
٣ – إحصائي الاختبار :

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.21}{\sqrt{\frac{1-(0.21)^2}{10-2}}} = 0.6075$$

إذا : $t = 0.675$

٤ – حدود منطقتي القبول والرفض:

من جدول t حيث مستوى المعنوية $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$ ودرجات الحرية تساوي (n – 2 = 10 – 2=8)
نجد أن قيمة t تساوي ٢.٤٤٧ وتكون حدود منطقتي القبول والرفض كما يلي:



٥ – المقارنة والقرار : بمقارنة قيمة الإحصائية المحسوبة في الخطوة رقم ٣ والتي تساوي 0.6075 بحدود منطقتي القبول والرفض (أو قيم t الجدولية في الخطوة رقم ٤) نجد أنها تقع في منطقة القبول (حيث أنها أقل من 2.447) لذلك فإن القرار هو : قبول الفرض العدمي. أي قبول الفرض القائل أن معامل الارتباط يساوي صفر. أي لا يوجد ارتباط بين المتغيرين وذلك بمستوى معنوية 5 %.

تمرين واجب :-

إذا علمت أنه :-

" أن معامل الإرتباط بين ثلاث ظواهر إقتصادية قد بلغت (r = 0.91) و كان عدد المفردات التي تم دراستها (n = 10)، وقد رغب الباحث في دراسة معنوية الارتباط و ذلك بمستوى 5% "

(١) قيمة إحصائي الاختبار t تساوي :-

(أ) 0.6208

(ب) 0.6208-

(ج) 6.208

(د) لا شيء مما سبق

(٢) إذا علمت أن حدود منطقتي القبول و الرفض هي (2,447- , 2.447)فعلى ذلك يمكن :-

(أ) قبول الفرض العدمي.

(ب) رفض الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

المحاضرة (١١)

الاختبارات الاحصائية اللامعلمية

١- إختبارات جودة التوفيق :-

إذا كان لدينا مجموعة من المفردات لعينة مأخوذة من مجتمع الدراسة و نرغب في التعرف على التوزيع الاحتمالي لهذه البيانات ، ويتم ذلك عن طريق ما يسمى بتوفيق المنحنيات حيث نبدأ بافتراض توزيع إحتمالى نظري يمكن أن تخضع له البيانات ، ويتوقف على بعض المقاييس الاحصائية الهامة للبيانات مثل الوسط الحسابي و الوسيط والانحراف المعياري ، ثم تستخدم بيانات العينة و التوزيع الاحتمالي المفترض في تقدير معالم هذا التوزيع والذى يفيد بدوره في الحصول على الاحتمالات ومن ثم التكرارات المتوقعة .

ثم تستخدم بيانات العينة و التوزيع الاحتمالي المفترض في تقدير معالم هذا التوزيع و الذي يفيد بدوره في الحصول على الاحتمالات ومن ثم التكرارات المتوقعة .

ومن ثم فإن جودة التوفيق هو اختبار إحصائي يمكن بإستخدامه معرفة هل التوزيع أو المنحنى الاحصائي النظري الذي تم توفيقه باستخدام بيانات العينة المأخوذه من المجتمع الأصلي يمثل تمثيلاً جيدا توزيع المتغير محل الدراسة في هذا المجتمع أم لا ؟ أو بمعنى أخر هل هناك إختلاف بين التوزيع الاحتمالي النظري الذي تم توفيقه و توزيع العينة ؟

إختبار كا٢ لجودة التوفيق :-

يستخدم أختبار كا٢ لاختبار ما اذا كانت بيانات العينة تتبع توزيع إحتمالى نظري معين أو للتأكد من صحة فرض معين ، ويتم ذلك من خلال معرفة ما إذا كان هناك فروق معنوية بين التكرارات الفعلية و التكرارات المتوقعة ، فكلما كان هذا الفرق صغير كلما اقترب التوزيعان الفعلي و النظري .

أولاً : إختبار كا٢ لتوفيق التوزيعات الاحتمالية النظرية أو أى توزيع آخر غير محدد الصيغة

مثال (١) :-

الجدول التالي يبين توزيع ٢٠٠ طالب بكلية العلوم الإدارية و التخطيط بجامعة الملك فيصل حسب المعدل التراكمي للطالب :-

المعدل التراكمي	0 -	1 -	2 -	3 -	4 - 5	المجموع
عدد الطلاب	28	35	53	45	39	200

و المطلوب : توفيق توزيع منتظم يوضح توزيع الطلاب حسب المعدل التراكمي وإختبار جودة التوفيق بدرجة ثقة 95% .

الحل :-

H_0 : توزيع الطلاب بكلية حسب فئات المعدل التراكمي يتبع التوزيع المنتظم .

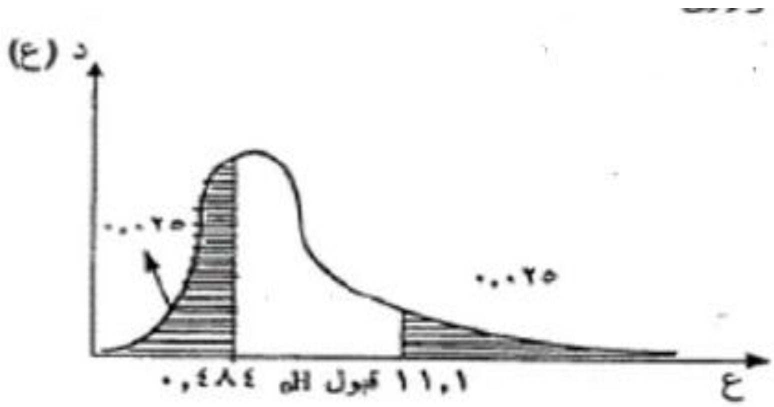
H_1 : توزيع الطلاب بكلية حسب فئات المعدل التراكمي لا يتبع التوزيع المنتظم .

حيث أن هناك خمس فئات للمعدل التراكمي فيتم توزيع الطلاب على الخمس فئات بالتساوي و لكل فئة تتكرار متوقع يساوي مجموع التكرارات على خمسة (40 = 200/5) ، كما يتضح من الجدول التالي :-

فئات المعدل التراكمي	التكرارات المشاهدة ش	التكرارات المتوقعة ت	^{2}(ش – ت)	^{2}(ش – ت) ت
0 -	28	40	144	3.6
1 -	35	40	25	0.625
2 -	53	40	169	4.225
3 -	45	40	25	0.625
4 - 5	39	40	1	0.025
المجموع	**200**	**200**		**9.1**

إذاً كا ^{2} المحسوبة = 9.1

درجات الحرية = (عدد الفئات – ١)
= (٥-١) = 4
و عند مستوى معنوية 5% فإن قيمتي كا ^{2} الجدولية هما (0.484 , 11.1) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-
و حيث أن قيمة كا ^{2} المحسوبة تقع في منطقة القبول لذلك نقبل الفرض العدمى و هو ما يعني أن منحنى التوزيع المنتظم يعتبر توفيق جيد لتوزيع طلاب الكلية حسب فئات المعدل التراكمي .



مثال (٢) :-

الجدول التالي يبين توزيع ١٠٠ موظف من موظفي إحدى الشركات حسب فئات الدخل الشهري (بالريال) :-

فئات الدخل الشهرى	100 -	200 -	300 -	400 -	500 - 600	المجموع
عدد الموظفين	12	15	22	35	16	100

و المطلوب : توفيق توزيع منتظم يوضح توزيع الموظفين حسب الدخل الشهري وإختبار جودة التوفيق بدرجة ثقة 95% .

الحل :-

H_0 : توزيع الموظفين حسب فئات الدخل الشهرى يتبع التوزيع المنتظم .

H_1 : توزيع الموظفين حسب فئات الدخل الشهرى لا يتبع التوزيع المنتظم .

حيث أن هناك خمس فئات للدخل الشهري فيتم توزيع الموظفين على الخمس فئات بالتساوي و لكل فئة تكرار متوقع يساوي مجموع التكرارات على خمسة (20 = 100/5) ، كما يتضح من الجدول التالي :-

فئات الدخل الشهري	التكرارات المشاهدة ش	التكرارات المتوقعة ت	(ش – ت)٢	(ش – ت)٢ / ت
100 -	12	20	64	3.2
200 -	15	20	25	1.25
300 -	22	20	4	0.2
400 -	35	20	225	11.25
500 -600	16	20	16	0.8
المجموع	**100**	**100**		**16.7**

إذاً كا ٢ المحسوبة = **16.7**

درجات الحرية = (عدد الفئات – ١)

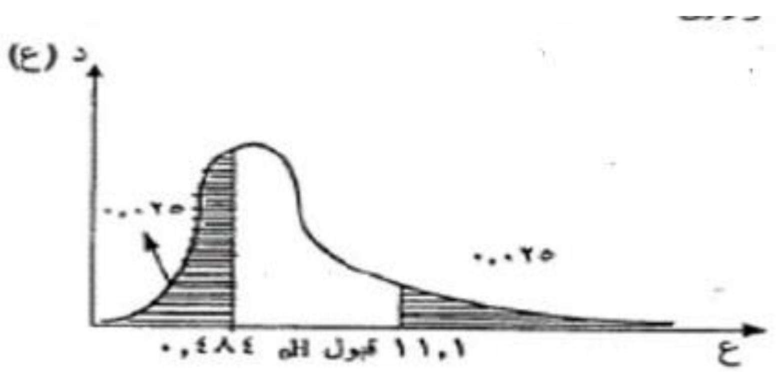
= (٥-١) = 4

و عند مستوى معنوية 5% فإن قيمتي كا ٢ الجدولية هما (0.484 , 11.1) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-

و حيث أن قيمة كا ٢ المحسوبة تقع في منطقة الرفض

لذلك نرفض الفرض العدمى و نقبل الفرض البديل و هو

ما يعني أن منحنى التوزيع المنتظم لا يعتبر توفيق جيد لتوزيع الدخل الشهري للموظفين .



مثال (٣) :-

قامت إحدى شركات الادوية بتوريد ١٠٠ كرتونه مصل الحمة الشوكية لأحد المستشفيات كل كرتونة تحتوى على ٣٠ زجاجة مصل و لوحظ توزيع عدد زجاجات المصل المكسورة بالكرتونة وكان كما يلي :-

عدد الزجاجات المكسورة بالكرتونه	0	1	2	3	4	5	المجموع
عدد الكراتين	22	28	35	10	3	2	100

و المطلوب : توفيق دالة إحتمال توزيع ذات الحدين لعدد زجاجات المصل المكسورة بالكرتونة في الشركة واختبار جودة التوفيق عند درجة الثقة 95% .

الحل :-

دالة التوزيع الاحتمال لتوزيع ذات الحدين تتوقف على معلمتين n و p أي عدد الفئات و الاحتمال :-

أولاً عدد الفئات تساوى 5 أي أن n = 5 .

ثانياً : الاحتمال :-

نحسب المتوسط أولاً
$$\mu = \frac{0\times 22 + 1\times 28 + 2\times 35 + 3\times 10 + 4\times 3 + 5\times 2}{100} = 1.5$$

لا تنسى أن :- $\mu = n\,p$

$$1.5 = 5 \times P$$

$$P = 0.3$$

H_0 : عدد زجاجات المصل المكسورة بالكرتونة الواحدة يتبع التوزيع ثنائي الحدين بالمعلمتين n = 5 , p = 0.3.

H_1 : عدد زجاجات المصل المكسورة بالكرتونة الواحدة لا تتبع التوزيع ثنائي الحدين بالمعلمتين n = 5 , p = 0.3.

و من خلال الاعتماد على معلمات التوزيع ثنائي الحدين يمكن تكوين جدول توزيع ثنائي الحدين ، كما يتضح من الجدول التالي :-

عدد الزجاجات المكسورة	الاحتمال	التكرار المتوقع
0	$^5c_0 \times (0.3)^0 \times (0.7)^5 = 0.1681$	16.81
1	$^5c_1 \times (0.3)^1 \times (0.7)^4 = 0.3602$	36.02
2	$^5c_2 \times (0.3)^2 \times (0.7)^3 = 0.3087$	30.87
3	$^5c_3 \times (0.3)^3 \times (0.7)^2 = 0.1323$	13.23
4	$^5c_4 \times (0.3)^4 \times (0.7)^1 = 0.0284$	2.84
5	$^5c_5 \times (0.3)^5 \times (0.7)^0 = 0.0024$	0.24
المجموع	**1**	**100**

لاحظ أن التكرار المتوقع = الاحتمال × عدد الكراتين ١٠٠

ولان إختبار كا٢ يشترط ألا يقل التكرار المتوقع لاي خلية عن ٥ ، لذلك سيتم دمج الخلايا الثلاثة الاخيرة لكي يصبح التكرار المتوقع لهم معا أكبر من أو يساوي ٥ كما يتضح من الجدول التالي :-

عدد الزجاجات المكسورة	التكرارات المشاهدة ش	التكرارات المتوقعة ت	(ش – ت)٢	(ش – ت)٢ / ت
0	22	16.81	26.94	1.60
1	28	36.02	64.32	1.79
2	35	30.87	17.06	0.55
3-5	15	16.31	1.72	0.11
المجموع	100	100		4.05

إذاً كا ٢ المحسوبة = 4.05

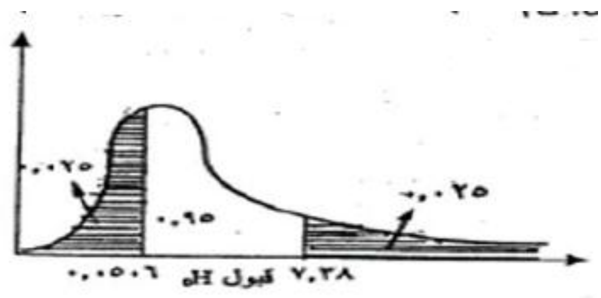
درجات الحرية = عدد الخلايا بعد الدمج – عدد المعلمات =

= ٤ – ٢ = ٢

و عند مستوى معنوية 5% فإن قيمتي كا ٢ الجدولية هما (0.0506 , 7.38) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-

و حيث أن قيمة كا ٢ المحسوبة تقع في منطقة القبول

لذلك نقبل الفرض العدمى و هو ما يعني أن منحنى التوزيع ثنائي الحدين يعتبر توفيق جيد لتوزيع عدد الزجاجات حسب الزجاجات المكسورة .



مثال (٤) :-

قامت إحدى المطاعم بتوريد ٢٠٧ صندوق لأحد المستشفيات كل صندوق يحتوى على ٦٠ زجاجة مياه و لوحظ أن توزيع عدد زجاجات المياه المكسورة بالكرتونة كان كما يلي :-

عدد زجاجات المياه المكسورة بالكرتونه	0	1	2	3	4	5	المجموع
عدد الكراتين	40	50	72	29	9	7	207

و المطلوب : توفيق دالة إحتمال توزيع ذات الحدين لعدد زجاجات المياه المكسورة بالكرتونة واختبار جودة التوفيق عند درجة الثقة 95% .

دالة التوزيع الاحتمال لتوزيع ذات الحدين تتوقف على معلمتين n و p أي عدد الفئات و الاحتمال :-

أولاً عدد الفئات تساوى 5 أي أن n = 5 .

ثانياً : الاحتمال :-

نحسب المتوسط أولاً $\mu = \dfrac{0 \times 40 + 1 \times 50 + 2 \times 72 + 3 \times 29 + 4 \times 9 + 5 \times 7}{207} = 1.7$

لا تنسى أن :- $\mu = n\,p$

$1.7 = 5 \times P$

$P = 0.34$

H_0 : عدد زجاجات المياه المكسورة بالكرتونة الواحدة يتبع التوزيع ثنائى الحدين بالمعلمتين n = 5 , p = 0.34.

H_1 : عدد زجاجات المياه المكسورة بالكرتونة الواحدة لا تتبع التوزيع ثنائى الحدين بالمعلمتين n = 5 , p = 0.34.

و من خلال الاعتماد على معلمات التوزيع ثنائي الحدين يمكن تكوين جدول توزيع ثنائي الحدين ، كما يتضح من الجدول التالي :-

عدد الزجاجات المكسورة	الاحتمال	التكرار المتوقع
0	$^5c_0 \times (0.34)^0 \times (0.66)^5 = 0.1252$	25.9
1	$^5c_1 \times (0.34)^1 \times (0.66)^4 = 0.3226$	66.77
2	$^5c_2 \times (0.34)^2 \times (0.66)^3 = 0.3323$	68.795
3	$^5c_3 \times (0.34)^3 \times (0.66)^2 = 0.1712$	35.44
4	$^5c_4 \times (0.34)^4 \times (0.66)^1 = 0.0441$	9.128
5	$^5c_5 \times (0.34)^5 \times (0.66)^0 = 0.0045$	0.94
المجموع	**1**	**207**

لاحظ أن التكرار المتوقع = الاحتمال × عدد الكراتين ٢٠٧

ولان إختبار كا٢ يشترط ألا يقل التكرار المتوقع لاي خلية عن ٥ ، لذلك سيتم دمج الخلايا الاثنين الاخيرة لكي يصبح التكرار المتوقع لهم معا أكبر من أو يساوي ٥ كما يتضح من الجدول التالي :-

عدد الزجاجات المكسورة	التكرارات المشاهدة ش	التكرارات المتوقعة ت	(ش – ت)٢	(ش – ت)٢ / ت
0	40	25.9	198.81	7.676
1	50	66.77	281.2329	4.211
2	72	68.795	10.27	0.149
3	29	35.44	41.47	1.17
4-5	16	10.07	35.16	3.49
المجموع	207	207		16.7

إذاً كا ٢ المحسوبة = 16.7

درجات الحرية = عدد الخلايا بعد الدمج – عدد المعلمات =

= ٥ – ٢ = ٣

و عند مستوى معنوية 5% فإن قيمتي كا ٢ الجدولية هما (0.216 , 9.346) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-

و حيث أن قيمة كا ٢ المحسوبة تقع في منطقة الرفض لذلك نقبل الفرض البديل و هو ما يعني أن منحنى التوزيع ثنائي الحدين لا يعتبر توفيق جيد لتوزيع عدد الزجاجات حسب الزجاجات المكسورة .

مثال (٥) :-

اختار أحد الباحثين عينة حجمها n=800 شخصا من أحد المدن، وكان توزيعهم حسب فصيلة الدم كالتالي:

فصيلة الدم	A	B	AB	O
عدد الأشخاص (التكرار المشاهد)	200	150	100	350

هل يتفق هذا التوزيع مع توزيع أفراد مدينةأخرى كان توزيع فصيلة دمهم حسب النسب التالية:

فصيلة الدم	A	B	AB	O
النسب المئوية للأشخاص	25%	15%	15%	45%

استخدم مستوى معنوية $\alpha = 0.05$

الحل :-

الفروض الإحصائية:

H_0 : توزيع فصيلة الدم في العينة يتفق مع التوزيع المناظر للمدينة الأخرى.

H_A : توزيع فصيلة الدم في العينة لا يتفق مع التوزيع المناظر للمدينة الأخرى.

لابد أولاً من الحصول على التكرار المتوقع و ذلك عن طريق تحويل النسب التي حصلنا عليها في التمرين إلى أعداد وذلك بضرب هده النسب في مجموع التكرارات ٨٠٠

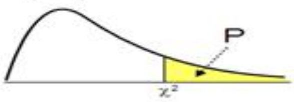
$$E_1 = np_1 = 800(0.25) = 200$$
$$E_2 = np_2 = 800(0.15) = 120$$
$$E_3 = np_3 = 800(0.15) = 120$$
$$E_4 = np_4 = 800(0.45) = 360$$

فصيلة الدم	التكرارات المشاهدة ش	التكرارات المتوقعة ت	(ش – ت)٢	(ش – ت)٢ ÷ ت
A	200	200	0	0
B	150	120	900	7.5
AB	100	120	400	3.33
O	350	360	100	0.2778
المجموع	800	800		11.11

درجات الحرية = ٤ – ١ = ٣

و عند مستوى معنوية 5% فإن قيمتي كا ٢ الجدولية هما (9.346 , 0.216) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-

و حيث أن قيمة كا ٢ المحسوبة تقع في منطقة الرفض لذلك نقبل الفرض البديل و هو ما يعني أن توزيع فصيلة الدم في المدينتين مختلف .



جدول توزيع χ^2

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588
11	2.603	3.816	14.631	17.275	19.675	21.920	22.618	24.725	26.757	29.354	31.264
12	3.074	4.404	15.812	18.549	21.026	23.337	24.054	26.217	28.300	30.957	32.909
13	3.565	5.009	16.985	19.812	22.362	24.736	25.472	27.688	29.819	32.535	34.528
14	4.075	5.629	18.151	21.064	23.685	26.119	26.873	29.141	31.319	34.091	36.123
15	4.601	6.262	19.311	22.307	24.996	27.488	28.259	30.578	32.801	35.628	37.697
16	5.142	6.908	20.465	23.542	26.296	28.845	29.633	32.000	34.267	37.146	39.252
17	5.697	7.564	21.615	24.769	27.587	30.191	30.995	33.409	35.718	38.648	40.790
18	6.265	8.231	22.760	25.989	28.869	31.526	32.346	34.805	37.156	40.136	42.312
19	6.844	8.907	23.900	27.204	30.144	32.852	33.687	36.191	38.582	41.610	43.820
20	7.434	9.591	25.038	28.412	31.410	34.170	35.020	37.566	39.997	43.072	45.315
21	8.034	10.283	26.171	29.615	32.671	35.479	36.343	38.932	41.401	44.522	46.797
22	8.643	10.982	27.301	30.813	33.924	36.781	37.659	40.289	42.796	45.962	48.268
23	9.260	11.689	28.429	32.007	35.172	38.076	38.968	41.638	44.181	47.391	49.728
24	9.886	12.401	29.553	33.196	36.415	39.364	40.270	42.980	45.559	48.812	51.179
25	10.520	13.120	30.675	34.382	37.652	40.646	41.566	44.314	46.928	50.223	52.620
26	11.160	13.844	31.795	35.563	38.885	41.923	42.856	45.642	48.290	51.627	54.052
27	11.808	14.573	32.912	36.741	40.113	43.195	44.140	46.963	49.645	53.023	55.476
28	12.461	15.308	34.027	37.916	41.337	44.461	45.419	48.278	50.993	54.411	56.892
29	13.121	16.047	35.139	39.087	42.557	45.722	46.693	49.588	52.336	55.792	58.301
30	13.787	16.791	36.250	40.256	43.773	46.979	47.962	50.892	53.672	57.167	59.703
31	14.458	17.539	37.359	41.422	44.985	48.232	49.226	52.191	55.003	58.536	61.098

مثال (٦) :-

قام أحد الباحثين بإختبار مدى أتفاق نتائج الطلاب للمعدلات التراكمية مع التوزيع المنتظم و حصل على النتائج التالية :-

فئات المعدل التراكمي	التكرارات المشاهدة ش	التكرارات المتوقعة ت
0 -	56	80
1 -	70	80
2 -	106	80
3 -	90	80
4 - 5	78	80
المجموع	**400**	**400**

المطلوب :-

١- تقدير قيمة كا^{٢} المحسوبة .
٢- إذا علمت أن حدود قيمة كا^{٢} الجدولية هي (11.1 , 0.484) فهل يمكن قبول الفرض العدمي .

فئات المعدل التراكمي	التكرارات المشاهدة ش	التكرارات المتوقعة ت	$(ش - ت)^{٢}$	$\frac{(ش - ت)^{٢}}{ت}$
0 -	56	80	576	7.2
1 -	70	80	100	1.25
2 -	106	80	676	8.45
3 -	90	80	100	1.25
4 - 5	78	80	4	0.05
المجموع	**400**	**400**		**18.2**

إذاً كا ^{٢} المحسوبة = 18.2

درجات الحرية = (عدد الفئات – ١)

= (٥-١) = 4

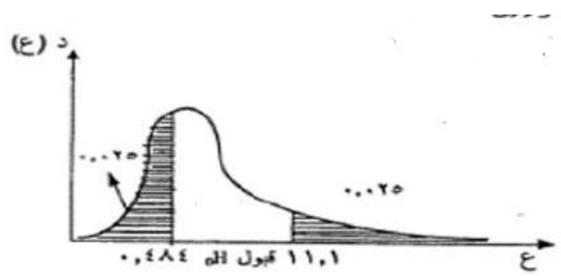
و عند مستوى معنوية 5% فإن قيمتي كا ^{٢} الجدولية هما (11.1 , 0.484) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-

و حيث أن قيمة كا ^{٢} المحسوبة تقع في منطقة الرفض

لذلك نقبل الفرض البديل و هو ما يعني أن منحنى

التوزيع المنتظم يعتبر توفيق جيد لتوزيع طلاب الكلية

حسب فئات المعدل التراكمي .



المحاضرة (١٢)

تابع الاختبارات الاحصائية اللامعلمية

تابع إختبارات جودة التوفيق :-

ثانياً : إختبار كا٢ لإستقلال متغيرين (ظاهرتين)في مجتمع واحد أو تجانس متغير (ظاهرة) ما في عدة مجتمعات

مثال (١) :-

سحبت عينة عشوائية من ١٠٠ فرد من إحدى المدن وتم توزيعهم حسب النوع و مستوى التعليم و كانت بيانتهم كما يلي :-

مستوى التعليم النوع	مؤهل متوسط	مؤهل فوق المتوسط	مؤهل مرتفع	المجموع
ذكر	35	15	10	**60**
أنثى	15	5	20	**40**
المجموع	**50**	**20**	**30**	**100**

المطلوب : إختبر الفرض القائل بوجود علاقة بين نوع الفرد و مستوى التعليم بدرجة ثقة ٩٩%.

الحل :-

Ho : لا يوجد علاقة بين الفرد و مستوى التعليم .

H1 : يوجد علاقة بين الفرد و مستوى التعليم .

و يحسب التكرار المتوقع لكل خلية عن طريق ضرب مجوع الصف في مجموع العمود و القسمة على المجموع و ذلك بالنسبة لكل خلية فمثلاً أول خلية

التكرار المتوقع لاول خلية = $\frac{50 \times 60}{100} = 30$

مستوى التعليم النوع	مؤهل متوسط	مؤهل فوق المتوسط	مؤهل مرتفع	المجموع
ذكر	35	15	10	60
أنثى	15	5	20	40
المجموع	50	20	30	100

مستوى التعليم / النوع	مؤهل متوسط		مؤهل فوق المتوسط		مؤهل مرتفع		المجموع
	ش	ت	ش	ت	ش	ت	
ذكر	35	30	15	12	10	18	60
أنثى	15	20	5	8	20	12	40
المجموع	50		20		30		100

النوع	مستوى التعليم	التكرارات المشاهدة ش	التكرارات المتوقعة ت	$(ش - ت)^2$	$\frac{(ش - ت)^2}{ت}$
ذكر	مؤهل متوسط	35	30	25	0.8333
	مؤهل فوق المتوسط	15	12	9	0.75
	مؤهل مرتفع	10	18	64	3.556
أنثى	مؤهل متوسط	15	20	25	1.25
	مؤهل فوق المتوسط	5	8	9	1.125
	مؤهل مرتفع	20	12	64	5.333
المجموع		100	100		12.8472

إذاً كا ٢ المحسوبة = 12.8472

من جدول توزيع كا٢ و عند درجات الحرية =

= (عدد الصفوف -١)(عدد الاعمدة -١) =

= (٢-١) (٣-١) = ٢

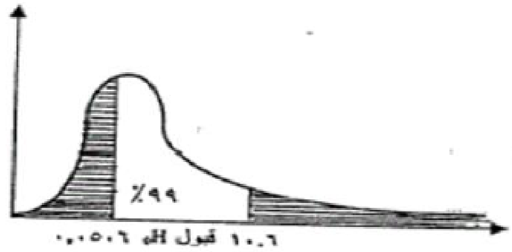
و لمستوى معنوية ١% نجد أن حدود فترة الثقة هي (0.01 , 10.6) كما يتضح من الشكل التالي :-

وحيث أن قيمة كا٢ المحسوبة تقع في منطقة الرفض ، إذا

نرفض الفرض العدمى و نقبل الفرض البديل و القائل

بأنه توجد علاقة بين نوع الفرد و مستوى التعليم بدرجة

ثقة 99% .



مثال (٢) :-

الجدول التالي يبين نتيجة أحد الاختبارات في نهاية دورة تدريبية موحده عقدت لثلاثة أقسام مختلفة بإحدى شركات الغزل و النسيج :-

النتيجة القسم	نجاح	فشل	المجموع
الغزل	65	15	80
النسيج	62	8	70
الطباعة	38	12	50
المجموع	165	35	200

و المطلوب : إختبار ما إذا كانت قدرات المتدربين متقاربة في الاقسام الثلاثة بدرجة ثقة 95% .

Ho : قدرات المتدربين متقاربة في الاقسام الثلاثة .

H1 : قدرات المتدربين غير متقاربة في الاقسام الثلاثة .

نحسب أولاً التكرار المتوقع لكل خلية بنفس الطريقة في المثال السابق ، فمثلاً بالنسبة للخلية الاولى :-

التكرار المتوقع للخلية الاولى :- $\frac{165 \times 80}{200} = 66$

النتيجة القسم	نجاح	فشل	المجموع
الغزل	65	15	80
النسيج	62	8	70
الطباعة	38	12	50
المجموع	165	35	200

القسم	النتيجة	التكرارات المشاهدة ش	التكرارات المتوقعة ت	(ش – ت)٢	(ش – ت)٢ / ت
نجاح	الغزل	65	66	1	0.015152
	النسيج	62	57.75	18.0625	0.312771
	الطباعة	38	41.25	10.5625	0.256061
فشل	الغزل	15	14	1	0.071429
	النسيج	8	12.25	18.0625	1.47449
	الطباعة	12	8.75	10.5625	1.207143
المجموع		**200**	**200**		**3.337**

إذاً كا ٢ المحسوبة = 3.337

و من جدول توزيع كا٢ و عند درجات الحرية =

= (٣-١) (٢-١) = ٢

و لمستوى المعنوية 0.05 فإن حدود كا٢ هي (0.0506 , 7.38)

و حيث أن قيمة كا٢ المحسوبة تقع في منطقة القبول ، فنقبل Ho أي يقبل الفرض الذي يقضي بأن قدرات المتدربين في الاقسام الثلاثة متقاربة عند مستوى المعنوية 5% .

اختبار تباين المجتمع

يستخدم توزيع χ^2 في إجراء العديد من الاختبارات الإحصائية مثل:

الاختبارات المتعلقة بتباين مجتمع ما (وذلك لاختبار المشاكل التي تتطلب اختبار تشتت مجتمع ما)، ويتم ذلك من خلال استخدام المعادلة التالية:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

ويفترض في هذا الاختبار أن العينة مسحوبة من مجتمع معتدل وذلك من خلال مقارنة قيمة χ^2 المحسوبة من المعادلة بالقيمة الحرجة لـ χ^2 والمستخرجة من جداول χ^2 .

<u>مثال (٣) :-</u>

إذا علمت أن تباين قوة مقاومة الكسر للكابلات التي تنتجها إحدى الشركات لاتزيد عن 40000 ، وتستخدم الشركة الآن طريقة إنتاج جديدة يعتقد أنها ستزيد من تباين قوة مقاومة الكابلات للكسر، سحبت عينة عشوائية من عشرة كابلات فوجد تباينها يساوي 50000 .

بافتراض أن قوة مقاومة الكسر للكابلات تتبع التوزيع المعتدل، اختبر الفرض القائل بوجود زيادة معنوية في التباين عند مستوى معنوية α = 0.01 .

<u>الحل :-</u>

- ❑ وضع فرض العدم والفرض البديل.

صياغة الفرضية الصفرية كالتالي:

H0: $\sigma^2 \le 40000$

في حين تفترض الفرضية البديلة التالي :

H_A: $\sigma^2 > 40000$

- ❑ تحديد مستوى الدلالة (α): وهي 0.01 .
- ❑ درجات الحرية = 9 ، فإن قيمة χ^2 المجدولة هي 21.666

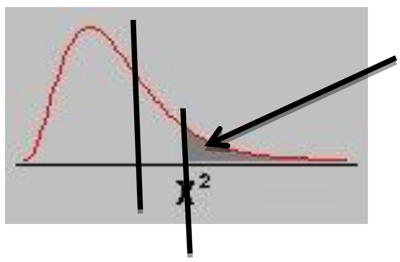
لذا فإن قاعدة القرار هي أن يتم رفض الفرضية الصفرية H0 عندما تكون

$$\chi^2 \ge 21.666$$

وحيث أن قيمة χ^2 لاختبار تباين المجتمع يتم حسابها كالتالي :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{(10-1)50000}{40000} = \frac{(9)50000}{40000} = \frac{450000}{40000} = 11.25$$

وحيث أن قيمة χ^2 المحسوبة أقل من قيمة χ^2 المجدولة، فإننا بالتالي نقبل الفرضية الصفرية H0 عند مستوى دلالة 0.01 وبالتالي يمكننا القول أن بيانات العينة تدل على أن الزيادة الظاهرة في التباين ليست معنوية عند مستوى الدلالة المحدد ، والشكل التالي يوضح ذلك.



مثال (٤) :-

إذا علمت أن تباين درجات الطلاب في جامعة الملك فيصل لا تقل عن ١٠ درجة، وتستخدم الجامعة الآن طريقة جديدة في التدريس يعتقد أنها ستقلل من تباين درجات الطلاب ، سحبت عينة عشوائية من ١٢ طالب فوجد تباينها يساوي ٢٤ .

بافتراض أن درجات الطلاب تتبع التوزيع المعتدل ، اختبر الفرض القائل بإنخفاض معنوية التباين عند مستوى معنوية $\alpha = 0.01$.

الحل :-

- ❑ وضع فرض العدم والفرض البديل.

صياغة الفرضية الصفرية كالتالي:

$$H0: \quad \sigma^2 \geq 10$$

في حين تفترض الفرضية البديلة التالي :

$$H_A: \quad \sigma^2 < 10$$

- **تحديد مستوى الدلالة (α): وهي 0.01 .**
- **درجات الحرية = 11 ، فإن قيمة χ^2 المجدولة هي 24.725**

لذا فإن قاعدة القرار هي أن يتم رفض الفرضية الصفرية H0 عندما تكون

$$\chi^2 \geq 24.725$$

وحيث أن قيمة χ^2 لاختبار تباين المجتمع يتم حسابها كالتالي :

وحيث أن قيمة χ^2 المحسوبة أكبر من قيمة χ^2 المجدولة، فإننا بالتالي نرفض الفرضية الصفرية H0 عند مستوى دلالة 0.01 .

مثال (٥) :-

في دراسة للعلاقة بين التقدير الذي يحصل عليه الطالب في الجامعة وجنسه أخذت عينة من نتائج الطلاب الذكور و الاناث وكانت كما يلي:

أولا: الاناث

راسب	راسب	راسب	راسب	جيد جدا	ممتاز	مقبول	ممتاز
جيد	جيد جدا	جيد جدا	جيد	مقبول	مقبول	مقبول	راسب
مقبول	راسب	مقبول	مقبول	مقبول	راسب	جيد جدا	جيد جدا
جيد	جيد	ممتاز	جيد جدا	ممتاز	جيد	جيد	جيد
					جيد جدا	ممتاز	جيد

ثانيا: الذكور

راسب	جيد	جيد	جيد	راسب	جيد جدا	راسب	جيد جدا
جيد جدا	جيد	راسب	راسب	راسب	راسب	راسب	مقبول
مقبول	ممتاز	ممتاز	راسب	راسب	مقبول	مقبول	ممتاز
ممتاز	جيد	جيد	مقبول	راسب	راسب	جيد	جيد
			جيد جدا	ممتاز	جيد	جيد جدا	ممتاز

والمطلوب:

هل توجد علاقة بين تقدير الطالب وجنسه عند مستوى الدلالة ٠.٠٥ ؟ α =

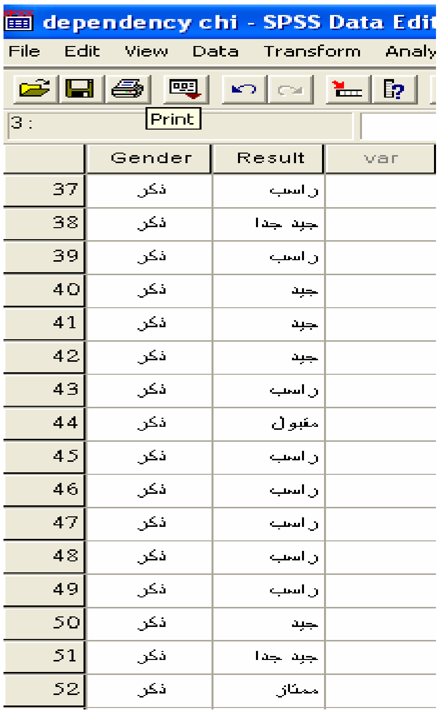
الحل :-

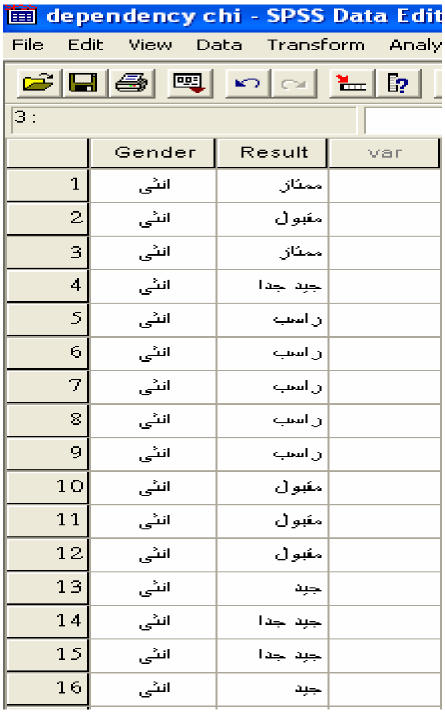
الفرضية الصفرية: تقدير الطالب لا يعتمد على جنسه (متغير الجنس والتقدير مستقلان)

الفرضية البديلة: تقدير الطالب يعتمد على جنسه (توجد علاقة بين جنس الطالب وتقديره)

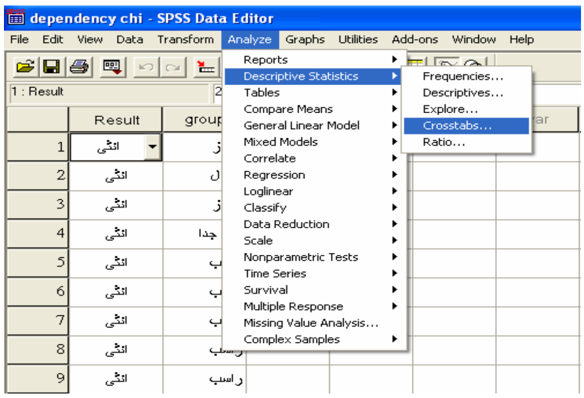
ثم نقوم بتعريف متغيرين نوعيين هما (*Result*) و (*Gender*) في شاشة تعريف المتغيرات بحيث يكون كود متغير (*Result*) هو (٠ = راسب، ١=مقبول، ٢=جيد، ٣ جيد جدًا، ٤=ممتاز) وكود المتغير (*Gender*) هو (١=ذكر، ٢=انثى)

ندخل البيانات كما في الشكل التالي:

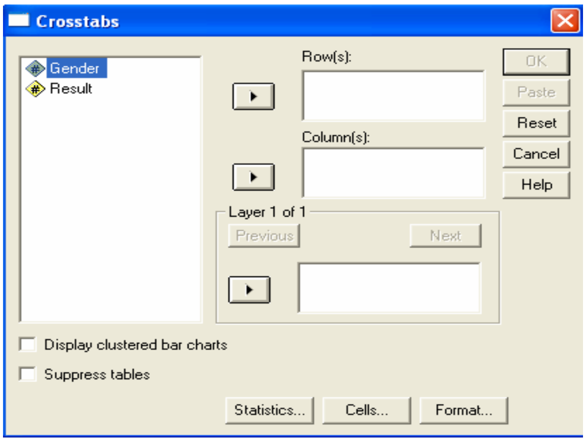




من قائمة التحليل *Analyze* نختار القائمة الفرعية للإحصاءات الوصفية *Descriptive Statistics* ومن ثم نختار الأمر *Cross tabs* كما في الشكل التالي:

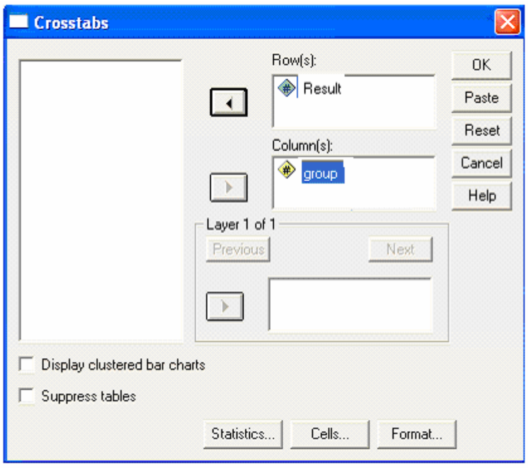


يظهر المربع الحواري التالي:

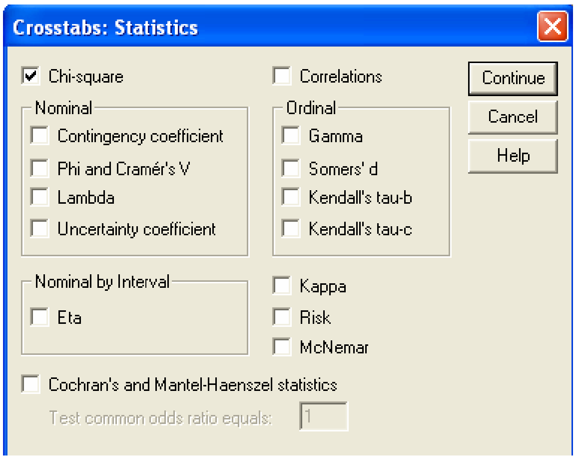


ننقل المتغير ***Result*** لخانة الصفوف ***Rows*** والمتغير ***Gender*** لخانة الأعمدة ***Columns*** باستخدام الأسهم

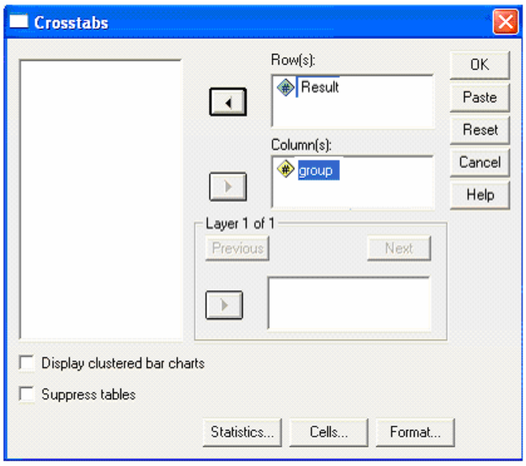
.



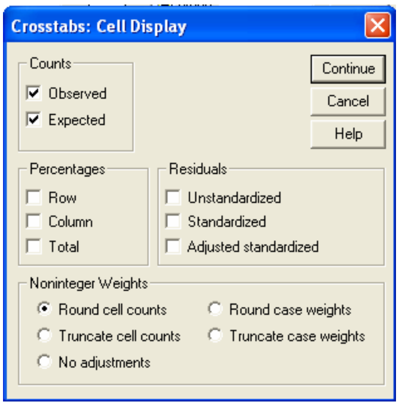
ومن ثم نضغط على ***Statistics*** للحصول على المربع الحواري التالي:



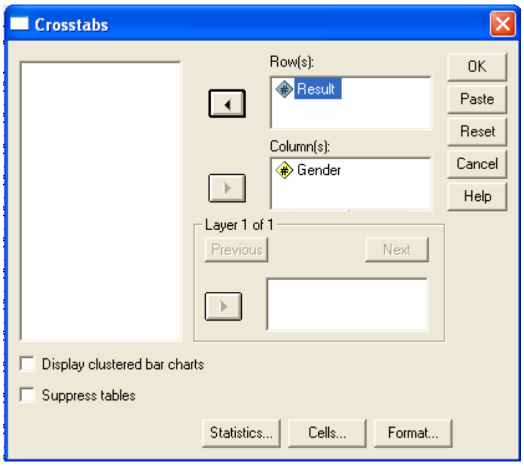
نضع علامة على خانة اختبار مربع كاي ***Chi-Square*** لحساب اختبار الاستقلالية ومن ثم نضغط على ***Continue*** للعودة للمربع الحواري السابق:



لاظهار جدول التوقعات نضغط على زر **Cell** ليظهر المربع الحواري التالي:



نختار الخيار ***Expected*** جدول توقعات ظهور البيانات ومن ثم نضغط ***Continue*** للعودة للمربع الحواري السابق.



نضغط على ***Ok*** للحصول على النتائج.

تتكون نتائج الأمر ***Cross tabulati*** من ثلاثة جداول:

<u>الأول</u> يصف حجم العينات المدخلة ونسب البيانات المفقودة كالتالي:

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
group * Result	72	100.0%	0	.0%	72	100.0%

الجدول الثاني يبين جدول توزيع العينة حسب المتغيرين والقيم المتوقعة حسب اختبار الاستقلالية كالتالي

عدد الذكور الراسبين

group * Result Crosstabulation

			Result		Total
			ذكر	انثى	
group	راسب	Count	12.00	7.00	19
		Expected Count	9.76	9.24	19.0
	مقبول	Count	5.00	8.00	13
		Expected Count	6.68	6.32	13.0
	جيد	Count	9.00	8.00	17
		Expected Count	8.74	8.26	17.0
	جيد جدا	Count	5.00	7.00	12
		Expected Count	6.17	5.83	12.0
	ممتاز	Count	6.00	5.00	11
		Expected Count	5.65	5.35	11.0
Total		Count	37.00	35.00	72
		Expected Count	37.00	35.00	72.0

يبين الجدول الثاني السابق أن عدد البيانات المدخلة ٧٢ ، عدد الذكور ٣٧ (منهم ١٢ راسب وقيمتها المتوقعة ٩.٧٦ ، ٥ مقبول وقيمتها المتوقعة ٦.٦٨ ، ٩ جيد وقيمتها المتوقعة ٨.٧٤ ، ٥ جيدجدا وقيمتها المتوقعة ٦.١٧ ، و ٦ ممتاز وقيمتها المتوقعة ٥.٦٥) والاناث ٣٥ (منهم ٧ راسب وقيمتها المتوقعة ٩.٢٤ ، ٨ مقبول وقيمتها المتوقعة ٦.٣٢ ، ٨ جيد وقيمتها المتوقعة ٨.٢٦ ، ٧ جيدجدا وقيمتها المتوقعة ٥.٨٣ ، و ٥ ممتاز وقيمتها المتوقعة ٥.٣٥)

الجدول الثالث يبين نتيجة اختبار مربع كاي كالتالي:

قيمة الاختبار

درجة الحرية

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.437[a]	4	.656
Likelihood Ratio	2.459	4	.652
Linear-by-Linear Association	.298	1	.585
N of Valid Cases	72		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.35.

يبين الجدول الثالث السابق أن قيمة اختبار مربع كاي هي ٢.٤٣٧ بدرجة حرية مقادرها ٤

يتبين لنا من الجدول أن أقل قيمة لمستوى الدلالة هي ٠.٦٥٦ = Asymp. Sig. (2-sided) وهي اكبر من مستوى الدلالة ٠.٠٠٥ = α وبالتالي لا نستطيع رفض الفرضية الصفرية أي أن تقدير الطالب لا يعتمد على جنسه.

مثال (٦) :-

الجدول التالي يوضح نتيجة إختبار مربع كاي (كا٢) عند مستوى معنوية 5%:-

Chi-Square Test

	Value	df	Asymp . Sig (2-sided)
Pearson Chi-Square	1.9496	3	.0437
Likelihood Ratio	1.9672	3	.0434
Linear-by- Linear Association	.2384	1	.0390
N of Valid Cases	32		

أجب عن الاسئلة التالية من خلال النتائج الواردة في الجدول السابق :-

(١) قيمة إحصائي الاختبار كا٢ تساوي :-

(أ) 2384.

(ب) 1.9672

(ج) 1.9496

(د) لا شيء مما سبق

(٢) قيمة مستوى الدلاله المحسوبة للإختبار تساوي :-

(أ) 0437.

(ب) 0434.

(ج) 0390.

(د) لا شيء مما سبق

(٣) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل.

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

المحاضرة (١٣)

تابع الاختبارات الاحصائية اللامعلمية

٢-اختبار مان وتني Mann – Whitney

استخدامه:

يعتبر هذا الاختبار بديل لا معلمي للاختبار الخاص بالفرق بين متوسطي مجتمعين والمبني على أساس عينتين مستقلتين أي أن هذا الاختبار بديل لاختبار t لعينتين مستقلتين، بل أنه أفضل منه خاصة إذا كانت العينتان مختارتين من مجتمعين لا يتبعان توزيعاً طبيعياً.

ويعد هذا الاختبار أكثر الاختبارات اللابارامترية استخداماً فى البحوث عندما يكون المتغير التابع من المستوى الرتبى بدلاً من الدرجات الأصلية، كما يمكن استخدام هذا الاختبار إذا كانت المتغيرات من المستوى الفترى أو النسبى ولكنها لا تفى بشروط اختبار النسبة التائية مثل عدم إعتدالية التوزيع أو اختلاف التباين بين المجموعتين اختلافاً كبيراً.

مثال (١) :-

فيما يلى بيان بدرجات مجموعة من الطلاب فى مادة المحاسبة، فى كل من جامعة الملك فيصل وجامعة الدمام:

(١) درجات مادة المحاسبة بكلية إدارة الأعمال جامعة الملك فيصل:

١٦	٨	٧	١٤	١٠
٧	١٤	١٥	٧	٣

(٢) درجات مادة المحاسبة بكلية إدارة الأعمال جامعة الدمام:

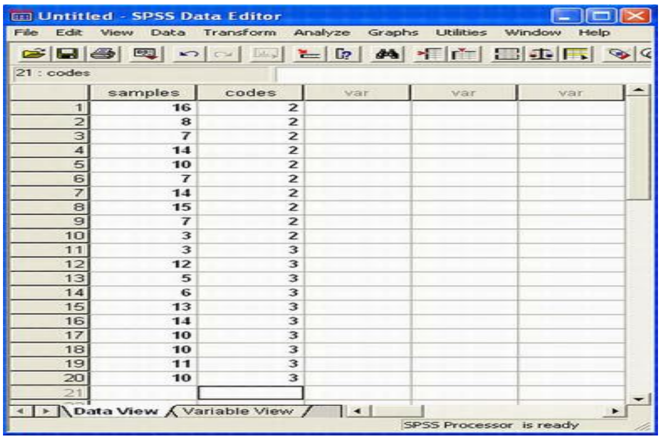
٣	١٢	٥	٦	١٣
١٤	١٠	١٠	١١	١٠

المطلوب:

بإستخدام إختبار مان – ويتنى: إختبر هل هناك إختلاف فى متوسط درجات مادة المحاسبة بين جامعة الملك فيصل وجامعة الدمام وذلك عند مستوى معنوية 5% .

الحل :-

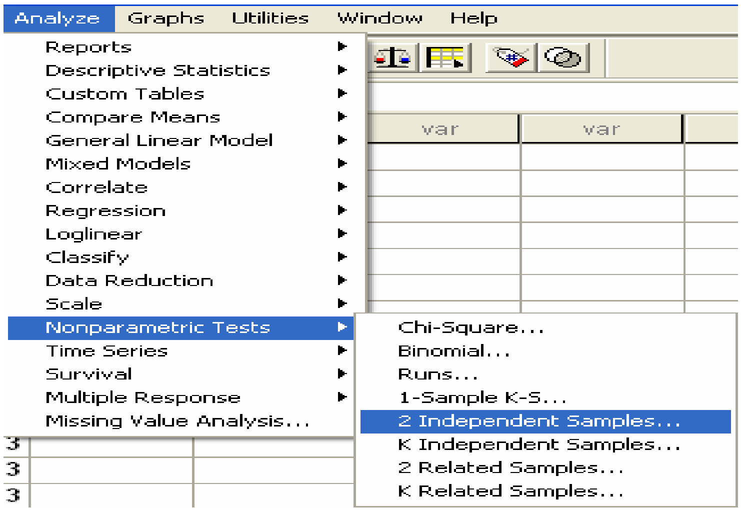
أولا: ندخل البيانات كالتالي:



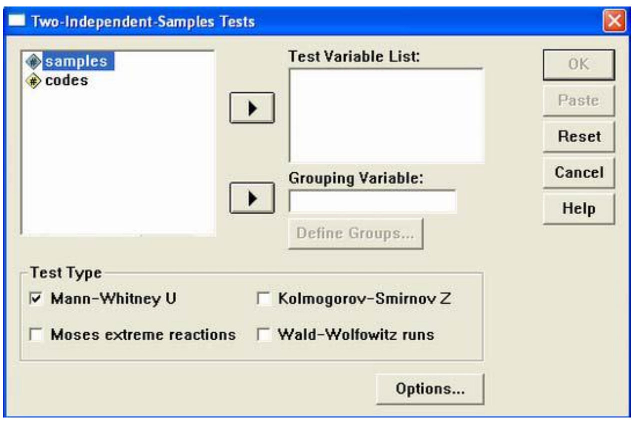
ملاحظة: فى هذا التدريب نحن بصدد إدخال بيانات لعينات مستقلة، لذا تم إدخال جميع المشاهدات فى عمود، ووالترميز الخاصة بالعينات فى عمود آخر وذلك من خلال إعطاء الرقم (٢) لبيانات العينة الأولى و (٣) لبيانات العينة الثانية.

ثانيا: خطوات تنفيذ الاختبار:

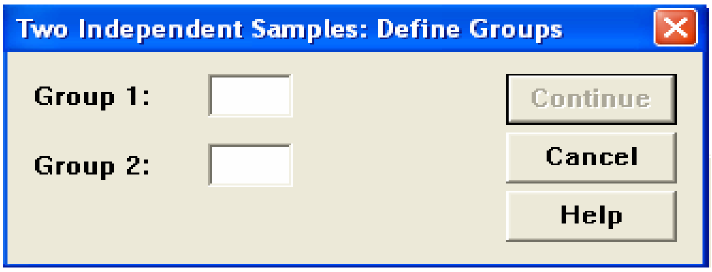
نضغط على قائمة Analyze ومن القائمة الفرعية لـ Nonparametric tests نختار Independent Samples ٢ كما هو موضح بالشكل التالى:



سوف يظهر لنا المربع الحوارى التالى:



انقل المتغير Samples الى المربع الذى بعنوان Test Variable List ، ثم انقل متغير الترميز codes إلى المربع الذي بعنوان Grouping Variable، ثم بعد ذلك اضغط على Define Groups سوف يظهر لنا مربع حوارى جديد كما يلى:



- فى خانة [Group 1] اكتب الرمز الخاص بالعينة الاولى (٢)، وفى خانة [Group 2] اكتب الرمز الخاص بالعينة الثانية (٣)
- ثم اضغط Continue للعودة الى المربع الحوارى السابق
- ثم اضغط Ok سوف تظهر لك نافذة المخرجات الخاصة بهذا الإختبار

Ranks

	CODES	N	Mean Rank	Sum of Ranks
SAMPLES	2	10	11.10	111.00
	3	10	9.90	99.00
	Total	20		

Test Statistics[b]

	SAMPLES
Mann-Whitney U	44.000
Wilcoxon W	99.000
Z	-.457
Asymp. Sig. (2-tailed)	.648
Exact Sig. [2*(1-tailed Sig.)]	.684[a]

يلاحظ من نتائج هذا الإختبار: أن قيمة P.Value تساوى 0.648 وهي أكبر من مستوى المعنوية ٥%
وبالتالى فاننا نقبل الفرض العدمى بأن متوسط درجات مادة المحاسبة فى كلية إدارة الأعمال جامعة الملك فيصل يساوى متوسط درجات مادة المحاسبة فى جامعة الدمام، أى أن الفروق بين الجامعتين غير معنوية.

مثال (٢) :-

" قام أحد الباحثين بمقارنة عينة من مرتبات موظفي القطاع الحكومي من مدينة الرياض بأخرى من مدينة جدة وذلك بصدد الوقوف على ما إذا كان هناك إختلاف في متوسط المرتبات وذلك عند مستوى معنوية 5%، وبإستخدام البرنامج الاحصائي SPSS حصلنا على النتائج التالية :-

Test Statistics

	SAMPLES
Mann-Whitney U	55.000
Wilcoxon W	95.000
Z	-.037
Asymp . Sig . (2-tailed)	.028
Exact Sig .[2*(1-tailed Sig.)]	.034

الحل :-

(١) الاختبار المستخدم لدارسة الفرق بين متوسطى مجتمعين فى هذه الحالة :-

(أ) كا٢ .

(ب) مان وتنى .

(ج) ويلكوكسون .

(د) لا شيء مما سبق

(٢) قيمة إحصائى الاختبار تساوى :-

(أ) -.037

(ب) .028

(ج) .034

(د) لا شيء مما سبق

(٣) من خلال مقارنة قيمة إحصائى الاختبار بقيمة حدود منطقتى القبول والرفض يمكن :-

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق .

٣- إختبار ويلكوكسون Wilcoxon Test

استخدامه:

ويسمى باختبار اشارات الرتب Sign –rank، ويستخدم هذا الاختبار فى تحديد ما إذا كان هناك اختلاف أو فروق بين عينتين مرتبطتين فيما يتعلق بمتغير تابع معين، ويعد بديـلاً لابارامترياً لاختبار T لعينيتين مرتبطين، وتشتمل العينتان على نفس المجموعة من الأفراد يجرى عليهم قياس قبلى Pre test، وقياس بعدى Post test وفى مثل هذه الحالة يكون لكل فرد من أفراد العينة درجتان أحداهما تمثل درجته فى الاختبار القبلى والثانية تمثل درجته فى الاختبار البعدى. ويستخدم مع البيانات العددية فقط دون الاسمية

حتى نحسب اختبار ويلكوكسن يجب اولا أن نجد الفرق بين القيمتين من أجل كل زوج ومن ثم من أجل كافة الحالات التي يكون عندها الفرق غير معدوم، نرتب الفروقات بشكل تصاعدي متجاهلين إشارة الفروقات، ذلك يعني بأن نسند إلى الفرق الصغير في القيمة المطلقة الرتبة ١ ونسند إلى الفرق الصغير التالي الرتبة ٢ وهكذا، أما في حالة الفروقات المتساوية (الحالات المتعادلة) نسند رتبة المتوسط إلى تلك الحالات.

مثال :-

تأثير ممارسة الرياضة على إنقاص الوزن:

الوزن قبل ممارسة الرياضة	الوزن بعد ممارسة الرياضة
٨٥	٨٠
٩٦	٨٥
٨٠	٨٥
٩٥	٨٢
٩٠	٧٥
٨٨	٨٠
١٠٣	٨٤
٩٨	٨٦

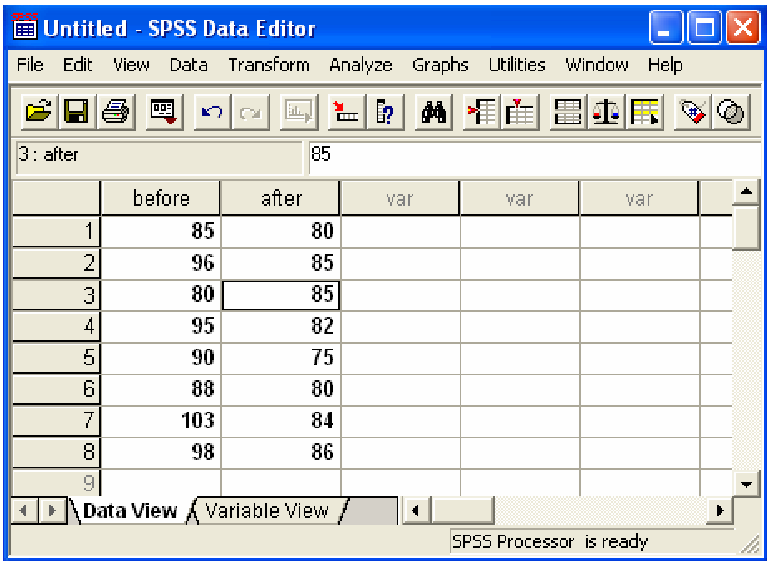
المطلوب:

إختبار هل هناك إختلاف معنوى فى الوزن بسبب ممارسة الرياضة، بإستخدام إختبار ويلكوكسون Wilcoxon عند مستوى معنوية ٥% .

الحل :-

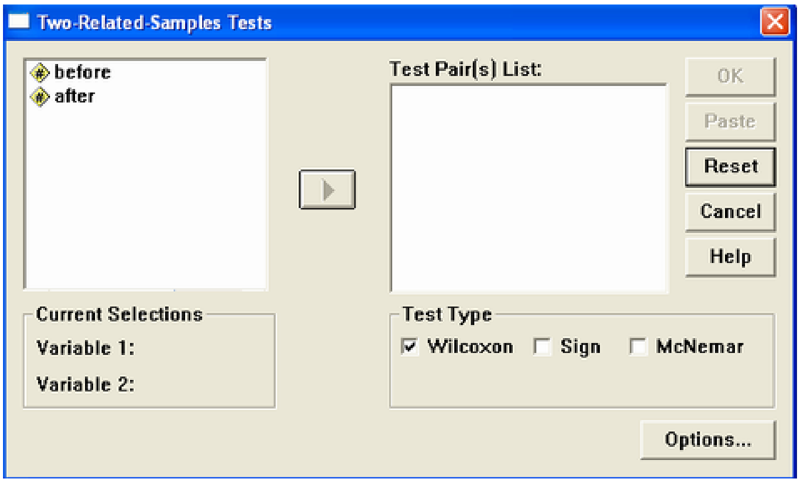
أولا: ندخل البيانات كالتالي:

حيث أننا بصدد عينات غير مستقلة، فإنه سيتم إدخال بيانات كل عينة فى عمود مستقل، كما يلى:



ثانيا: خطوات تنفيذ الاختبار:

نضغط على قائمة Analyze ومن القائمة الفرعية لـ Nonparametric tests نختار Related Samples ٢ كما هو موضح بالشكل التالى:



اضغط بالماوس مرة واحدة على المتغير before ثم على المتغير after (لاحظ أنه قد تم تظليل المتغيرين معًا)، ثم قم بنقل هذين المتغيرين الى المربع الذي بعنوان Test Pair(s) List وذلك من خلال الضغط على السهم الصغير الموجود بين المربعين.

لاحظ فى نفس المربع الحوارى الذى أمامك: أن الإختيار الإفتراضى من جانب البرنامج هو اختبار ويلكوكسن، وهو الإختبار الذى نريده لذا سنتركة كما هو. إضغط Ok ستظهر لك نافذة المخرجات الخاصة بهذا الإختبار كالتالي:

Ranks

		N	Mean Rank	Sum of Ranks
AFTER - BEFORE	Negative Ranks	7[a]	4.93	34.50
	Positive Ranks	1[b]	1.50	1.50
	Ties	0[c]		
	Total	8		

Test Statistics[b]

	AFTER - BEFORE
Z	-2.313[a]
Asymp. Sig. (2-tailed)	.021

قام البرنامج بحساب الفروق في الوزن على أساس التالي:

الفرق = الوزن بعد ممارسة الرياضة – الوزن قبل ممارسة الرياضة

ويلاحظ أيضا: أن متوسط الرتب السالبة (٤.٩٣) أكبر من متوسط الرتب الموجبة (١.٥)، وهذا معناه أن متوسط الوزن قبل ممارسة الرياضة أكبر من متوسط الوزن بعد ممارسة الرياضة (إذا فى غاية الأهمية أن نعرف الترتيب الذى إستخدمه البرنامج للعينتين)

ويلاحظ من نتائج هذا الاختبار أن قيمة P.Value تساوي 0.021 وهي أقل من مستوى المعنوية 5% وبالتالى فإننا نقبل الفرض البديل بأن متوسط الوزن قبل ممارسة الرياضة يختلف معنويًا عن متوسط الوزن بعد ممارسة الرياضة.

مثال :-

إذا علمت أنه :-

" لدراسة تأثير أحد البرامج التدريبية على مجموعة من الطلاب تم إختبار مجموعة من الطلاب قبل البرنامج التدريبي على عينة من ٨ طلاب و إختبار الطلاب بعد الحصول على البرنامج التدريبي ولاختبار هل هناك إختلاف معنوي في مستوى تحصيل الطلاب ، عند مستوى معنوية 5%، أستخدم الباحث البرنامج الاحصائي spss باستخدام إختبار ويلكوكسون Wilcoxonو حصلنا على النتائج التالية :-

Ranks

		N	Mean Rank	Sum of Ranks
AFTER-BEFORE	Negative Ranks	7	2.36	43.50
	Positive Ranks	1	3.54	3.54
	Ties	0		
	Total	8		

Test Statistics

	AFTER-BEFORE
Z	-.313
Asymp . Sig . (2-tailed)	.421

الحل :-

(١) من الجداول السابقة يمكن توضيح أن :-

(أ) مستوى الطلاب قبل الحصول على البرنامج التدريبي أفضل من المستوى بعد الحصول على البرنامج .

(ب) مستوى الطلاب بعد الحصول على البرنامج التدريبي أفضل من المستوى قبل الحصول على البرنامج .

(ج) مستوى الطلاب قبل الحصول على البرنامج التدريبي مساوي لمستوى بعد الحصول على البرنامج .

(د) لا شيء مما سبق

(٢) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق .

٣-اختبار كروسكال واليس Kruskal-Wallis Test

استخدامه:

يعتبر هذا الاختبار بديلاً لامعلميا لاختبار تحليل التباين فى اتجاه واحد، وهو مبني على مجموع الرتب ويستعمل لاختبار الفروق بين ثلاث مجموعات أو أكثر في مثل الحالة الآتية :

مثال:

الجدول التالى يوضح درجات مجموعة من الطلاب فى مادة الإقتصاد فى ثلاث جامعات هى: جامعة الملك فيصل – جامعة الدمام – جامعة الملك سعود:

جامعة الملك فيصل	جامعة الدمام	جامعة الملك سعود
١٣	٤	٥
١٤	٧	٦
١٤	١٠	١٥
١٥	١٢	١٠
١٥	٦	١٤
١٧	١٠	٦
٤	١٣	٦
١٦	١٨	١٢

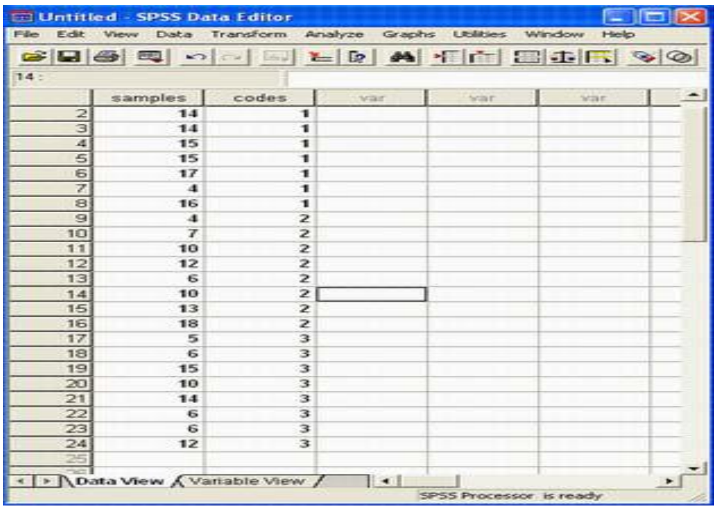
المطلوب:

دراسة مدى وجود إختلاف بين مستوى الطلاب فى الجامعات الثلاثة السابقة بإستخدام إختبار كروسكال-والس، وذلك عند مستوى معنوية 5%

<u>الحل :-</u>

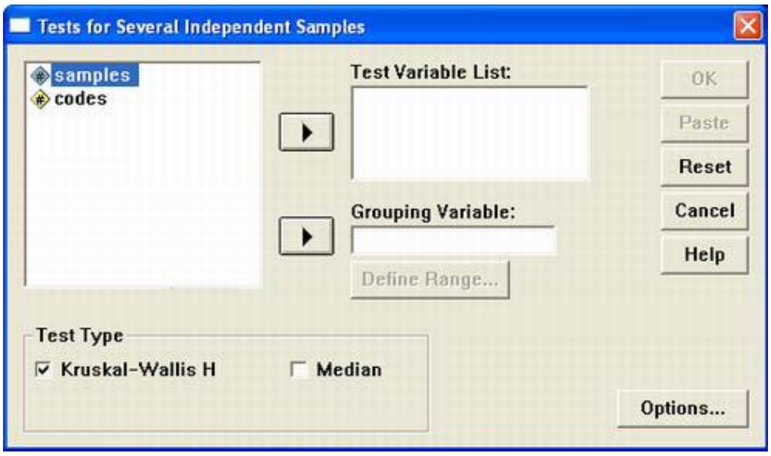
أولا: ندخل البيانات كالتالي:

حيث أننا بصدد ثلاث عينات مستقلة، لذا تم إدخال قيم المشاهدات فى عمود، والرموز الخاصة بالعينات فى عمود اخر، حيث تم إعطاء الرمز (١) لبيانات العينة الأولى، والرمز (٢) لبيانات العينة الثانية، والرمز رقم (٣) لبيانات العينة الثالثة كما يلي:



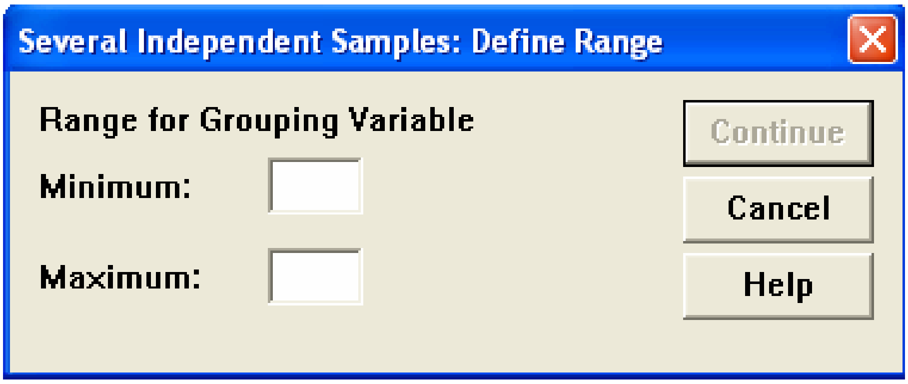
ثانيا: خطوات تنفيذ الاختبار:

نضغط على قائمة Analyze ومن القائمة الفرعية لـ Nonparametric tests نختار independent k Samples كما هو موضح بالشكل التالى:



- انقل المتغير samples الى المربع الذى بعنوان Test Variable List ثم انقل متغير الاكواد codes الى المربع الصغير الذى بعنوان Grouping Variable (لاحظ أن الإختيار الإفتراضى من جانب البرنامج هو إختبار كروسكال – والس)

- اضغط Define Groups سوف يظهر مربع حوارى جديد كما يلى:



- فى خانة Minimum اكتب أصغر الرمز (١) ، وفى خانة Maximum اكتب أكبر الرمز (٣) ، ثم اضغط Continue للعودة الى المربع الحوارى السابق.

- ثم اضغط Ok سوف تظهر لك نافذة المخرجات الخاصة بهذا الإختبار كالتالي:

Ranks

	CODES	N	Mean Rank
SAMPLES	1	8	16.88
	2	8	10.75
	3	8	9.88
	Total	24	

Test Statistics[a,b]

	SAMPLES
Chi-Square	4.706
df	2
Asymp. Sig.	.095

يلاحظ من نتائج هذا الإختبار أن قيمة P.Value تساوى 0.095 وهى أكبر من مستوى المعنوية ٥% ،

وبالتالى فاننا نقبل الفرض العدمى بأن متوسط درجات مادة الإقتصاد فى كلية إدارة الأعمال فى الجامعات الثلاثة متساوى، أي أن الفروق بين الجامعات الثلاثة غير معنوية.

مثال :-

" قام أحد الباحثين بدراسة درجات مجموعة من الطلاب فى مادة التحليل الاحصائي فى ثلاث جامعات هى: جامعة الملك فيصل – جامعة الدمام – جامعة الملك سعود ، وذلك لدراسة مدى وجود إختلاف بين مستوى الطلاب فى الجامعات الثلاثة السابقة بإستخدام إختبار كروسكال- والس، وذلك عند مستوى معنوية 5%، تم الحصول على النتائج التالية باستخدام البرنامج الاحصائى SPSS:-

Test Statistics

	SAMPLES
Ci-Square	.706
df	2
Asymp . Sig .	.025

(١) من الجدول السابق يمكن :-

(أ) قبول الفرض البديل القائل بمعنوية الفروق بين الجامعات الثلاثة .

(ب) قبول الفرض العدمي القائل بأن الفروق بين الجامعات الثلاثة غير معنوية .

(ج) قبول الفرض العدمي القائل بأن الفروق بين الجامعات الثلاثة معنوية .

(د) لا شيء مما سبق .

٤- حساب اختبار كولومجروف سيمرنوف لجودة التوافيق Goodness of Fit - Kolmogorov-Smirnov Test من خلال برنامج SPSS

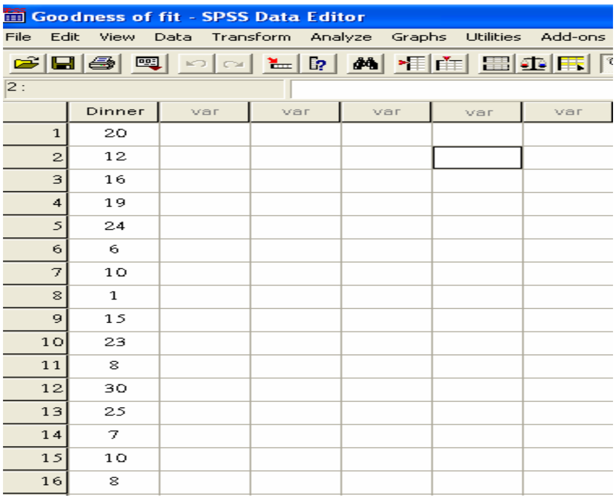
اختبار كولومجروف سيمرنوف لجودة التوافيق

: Goodness of Fit Test - Kolmogorov-Smirnov

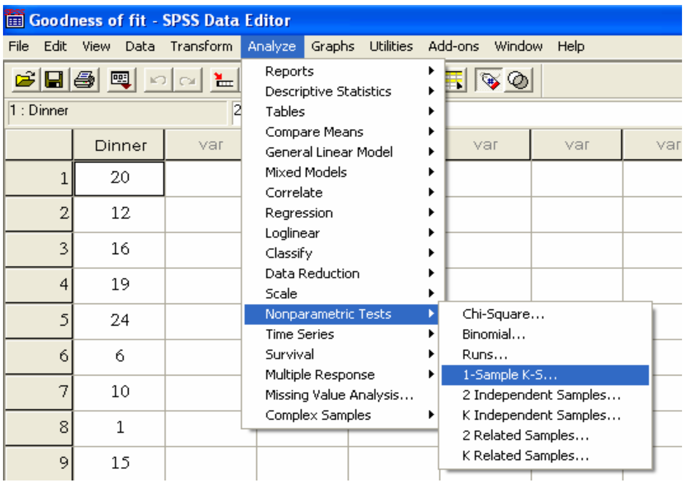
استخدامه:

يستخدم هذا الاختبار لمعرفة إذا ما كانت العينة موضع الاهتمام تتبع توزيعاً احتمالياً معينا ويستخدم عوضاً عن اختبار مربع كاى عندما يكون مجموع التكرارات أقل من ٣٠ أو يكون التكرار المتوقع لأي خلية أقل من خمسة وعملية ضم الخلايا تؤدي إلى فقد كثير من درجات الحرية مما يتعذر معه إجراء الاختبار أو أن تكون عملية الضم غير مناسبة. ويفضل استخدامه أيضاً في حالة كون التوزيع الاحتمالي لمتغير متصل.

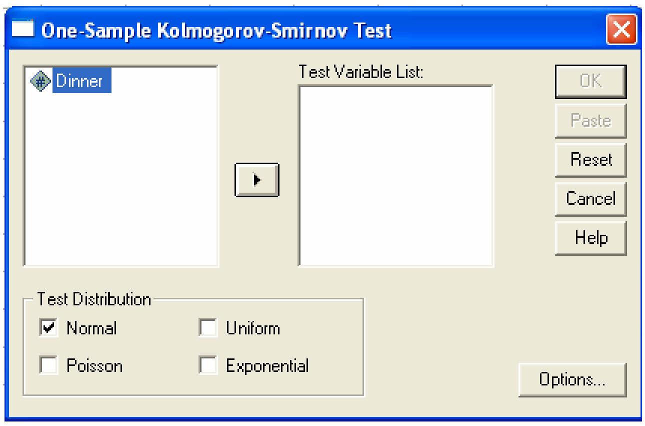
ندخل البيانات في متغير نسميه *Dinner* كما في الشكل التالي:



من قائمة التحليل *Analyze* نختار القائمة الفرعية الاحصاءات الغير بارامترية *Non-Parametric Test* ومن ثم نختار الأمر *1-Sample K-S*



يظهر المربع الحواري التالي:



يمكنك المربع الحواري السابق من اختيار التوزيع الذي تريد اختباره هل هو توزيع طبيعي ***Normal*** أو بواسون ***Poisson*** أو منتظم ***Uniform*** أو أسي ***Exponential*** فنختار التوزيع الطبيعي كما في الشكل أعلاه ونضغط ***Ok*** للحصول على النتائج التالية:

NPar Tests

One-Sample Kolmogorov-Smirnov Test

		Dinner
N		50
Normal Parameters a,b	Mean	15.26
	Std. Deviation	6.782
Most Extreme Differences	Absolute	.081
	Positive	.081
	Negative	-.069
Kolmogorov-Smirnov Z		.573
Asymp. Sig. (2-tailed)		.898

a. Test distribution is Normal.

b. Calculated from data.

50 حجم العينه **.898 مستوى دلالة الاختبار**

15.26 متوسط البيانات

6.782 الانحراف المعياري للبيانات

.081 اكبر فرق بين البيانات و دالة التوزيع الاحتمالية

573. قيمة اختبار جودة المطابقة

تبين النتائج أعلاه أن متوسط عدد الزبائن هو ١٥.٢٦ بانحراف معياري قدره ٦.٧٨٢ وأن قيمة اختبار كولموجروف سميرنوف لجودة المطابقة هو

القرار:

يبين الجدول السابق أن قيمة مستوى دلالة الاختبار هي *Asymp. Sig. (2-tailed) = ٠.٨٩٨* وهي اكبر من مستوى دلالة الفرضية الصفرية ٠.٠٥ = α وبالتالي نقبل الفرضية الصفرية، <u>أي أن البيانات تتبع التوزيع الطبيعي</u> وبالتالي نستنتج ان البيانات تتبع التوزيع الطبيعي بمتوسط قدره 15.26 وانحراف معياري 6.782 أي (15.26,6.782) *X* : *N*

وإذا أردنا اختبار أن التوزيع يتبع توزيع بواسون نختار من الشاشة المخصصة لذلك توزيع بواسون وهكذا مع باقي التوزيعات.

المحاضرة (١٤)

تمارين مراجعة

إذا علمت أنه :-

" في دراسة لظاهرة متوسط وزن الاطفال في سن الروضة ، أخذت عينة عشوائية من المجتمع مكونه من 64 طفل فوجد أن الوسط الحسابي لوزن الطفل في هذه العينة هو 20كجم وذلك بإنحراف معياري قدرة 8كجم " :-

(١) إن فترة الثقة للوسط الحسابي للمجتمع بدرجة ثقة ٩٥% هي :-
(أ) (21.65 , 18.35)كجم
(ب) (21.96 , 18.04)كجم
(ج) (22.58 , 17.15)كجم
(د) لا شيء مما سبق

(٢) إن فترة الثقة للوسط الحسابي للمجتمع بدرجة ثقة ٩٠% هي :-
(أ) (21.65 , 18.35)كجم
(ب) (21.96 , 18.04)كجم
(ج) (22.58 , 17.15)كجم
(د) لا شيء مما سبق

(٣) إن فترة الثقة للوسط الحسابي للمجتمع بدرجة ثقة ٩٩% هي :-
(أ) (21.65 , 18.35)كجم
(ب) (21.96 , 18.04)كجم
(ج) (22.58 , 17.15)كجم
(د) لا شيء مما سبق

(٤) " يرغب أحد مديري المدارس الأهلية في تقدير متوسط عدد الوجبات التي يتم صرفها للطلاب في مدرسته خلال الشهر بحيث لا يتعدى الخطأ في تقدير متوسط عدد الوجبات خلال الشهر الواحد عن 5 وجبات و بدرجة ثقة 95% ، ويعلم المدير من خبرته أن الانحراف المعياري هو 10 وجبات " و المطلوب تقدير حجم العينة المطلوب لهذه الدراسة مقرباً الناتج للرقم الاعلي :-
(أ) 11 عينة .
(ب) 16 عينة .
(ج) 33عينة .
(د) لا شيء مما سبق

(٥) " سحبت عينة عشوائية مكونة من 25 طالب من الطلاب الدارسين لمقرر الاحصاء في الادارة فوجد أن متوسط درجاتهم 80 درجة وذلك بانحراف معياري للعينة 5 = s و من المعروف أن درجات الطلاب موزعة طبقاً للتوزيع الطبيعي ، مما سبق يمكن ايجاد حدي الثقة لدرجات الطلاب عند درجة ثقة 95% تساوي :-

درجات الحرية	0.5	0.10	0.05	0.025	0.01	0.005
5	0.000	1.476	2.015	2.571	3.365	4.032
24	0.000	1.318	1.711	2.064	2.492	2.797
25	0.000	1.316	1.708	2.060	2.485	2.787

(أ) (82.060 , 77.94) درجة
(ب) (81.711 , 78.289) درجة
(ج) (82.064 , 77.936) درجة
(د) لا شيء مما سبق

(٦) أن "رفض الفرض العدمي بينما هو صحيح " يسمى.................................. .

(أ) خطأ من النوع الأول .

(ب) خطأ من النوع الثاني .

(ج) الخطأ المعياري .

(د) لا شيء مما سبق

.

إذا علمت أنه :-

" عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 ريال . وترغب في اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية 5 % إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال . "

(٧) **يمكن صياغة الفرض العدمي و الفرض البديل على الشكل :-**

(أ) $Ho: \mu = 72 \ , \ H1: \mu < 72$

(ب) $Ho: \mu = 72 \ , \ H1: \mu > 72$

(ج) $Ho: \mu = 72 \ , \ H1: \mu \neq 72$

(د) لا شيء مما سبق

(٨) **قيمة إحصائي الاختبار في هذه الحالة Z تساوي :-**

(أ) 3

(ب) 0.75

(ج) 1.5

(د) لا شيء مما سبق

(٩) **من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-**

(أ) قبول الفرض العدمي .

(ب) قبول الفرض البديل .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

إذا علمت أنه :-

" عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 ريال . وترغب في اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية 1 % إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال . "

(١٠) **يمكن صياغة الفرض العدمي و الفرض البديل على الشكل :-**

(أ) $Ho: \mu = 72 \ , \ H1: \mu < 72$

(ب) $Ho: \mu = 72 \ , \ H1: \mu > 72$

(ج) $Ho: \mu = 72 \ , \ H1: \mu \neq 72$

(د) لا شيء مما سبق

(١١) **قيمة إحصائي الاختبار في هذه الحالة Z تساوي :-**

(أ) 3

(ب) 0.75

(ج) 1.5

(د) لا شيء مما سبق

(١٢) **من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-**

(أ) قبول الفرض العدمي .

(ب) قبول الفرض البديل .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

إذا علمت أنه :-

"يدّعي أحد المرشحين في الانتخابات أنه سيحصل على نسبة 70% من أصوات الناخبين عندما تجري الانتخابات. ولاختبار هذا الادعاء تم اختيار عينة عشوائية من الناخبين حجمها 100 ناخب، ووجد أن نسبة من يؤيدون المرشح في العينة هي 60 % اختبر مدى صحة ادعاء المرشح بأن النسبة في المجتمع هي 70 % مقابل الفرض البديل أن النسبة أقل من 70% وذلك بمستوى معنوية 5 % "

(١٣) يمكن صياغة الفرض العدمي و الفرض البديل على الشكل :-

(أ) Ho: P = 0.70 , H1: P<0.70
(ب) Ho: P = 0.70 , H1: P >0.70
(ج) Ho: P = 0.70 , H1: P ≠0.70
(د) لا شيء مما سبق

(١٤) قيمة إحصائي الاختبار في هذه الحالة Z تساوي :-

(أ) 0.10
(ب) -0.10
(ج) -2.17
(د) لا شيء مما سبق

(١٥) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض العدمي .
(ب) قبول الفرض البديل .
(ج) عدم قبول أي من الفرضين .
(د) لا شيء مما سبق

إذا علمت أنه :-

"البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما : حيث $\bar{X}_1 = 35 , \bar{X}_2 = 29 , n_1 = 100 , n_2 = 80$ ، اختبر الفرض العدمي : أن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5% مقابل الفرض البديل أنهما غير متساويين إذا علمت أن : $\sigma_1^2 = 60 , \sigma_2^2 = 32$"

(١٦) يمكن صياغة الفرض العدمي و الفرض البديل على الشكل :-

(أ) $Ho: \mu_1 = \mu_2$, $H1: \mu_1 > \mu_2$
(ب) $Ho: \mu_1 = \mu_2$, $H1: \mu_1 < \mu_2$
(ج) $Ho: \mu_1 = \mu_2$, $H1: \mu_1 \neq \mu_2$
(د) لا شيء مما سبق

(١٧) قيمة إحصائي الاختبار في هذه الحالة Z تساوي :-

(أ) 60
(ب) 6
(ج) 0.20
(د) لا شيء مما سبق

(١٨) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض العدمي .
(ب) قبول الفرض البديل .
(ج) عدم قبول أي من الفرضين .
(د) لا شيء مما سبق

إذا علمت أنه :-

"إذا كان متوسط استهلاك الفرد السعودي من الدجاج حسب تقارير وزارة الصحة هو (١٢) كيلوجرام بانحراف معياري (٦) كيلوجرامات لفترة السبعينات الميلادية. أجرى أحد الباحثين دراسة في عام ٢٠٠٣م من عينة قوامها (٤٩) فرداً ووجد أن متوسط الاستهلاك للفرد هو (١٤) كيلوجرام. هل تشير الدراسة الحالية أن متوسط الاستهلاك أرتفع عما عليه في السبعينات."

(١٩) يمكن صياغة الفرض العدمي و الفرض البديل على الشكل :-

(أ) **$Ho: \mu = 12$, $H1: \mu > 12$**

(ب) $Ho: \mu = 12$, $H1: \mu < 12$

(ج) $Ho: \mu = 12$, $H1: \mu \neq 12$

(د) لا شيء مما سبق

(٢٠) قيمة إحصائي الاختبار في هذه الحالة Z تساوي :-

(أ) 2

(ب) **2.33**

(ج) 0.33

(د) لا شيء مما سبق

(٢١) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض العدمي .

(ب) **قبول الفرض البديل .**

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

إذا علمت أنه :-

"لو كانت لدينا عينة عشوائية تتكون من ٢٥٠ طالب وجد أن الوسط الحسابي لأطوال طلاب العينة ١٥٥.٩٥سم، والانحراف المعياري = ٢.٩٤ سم، علما بأن الوسط الحسابي لأطوال طلاب الجامعة يبلغ ١٥٨ سم، اختبر أهمية الفرق المعنوي بين الوسط الحسابي لأطوال طلاب العينة والوسط الحسابي لأطوال طلاب الجامعة."

(٢٢) يمكن صياغة الفرض العدمي و الفرض البديل على الشكل :-

(أ) $Ho: \mu = \mu_0$, $H1: \mu < \mu_0$

(ب) $Ho: \mu = \mu_0$, $H1: \mu > \mu_0$

(ج) **$Ho: \mu = \mu_0$, $H1: \mu \neq \mu_0$**

(د) لا شيء مما سبق

(٢٣) يسمى إحصائي الاختبار في هذه الحالة :-

(أ) Z

(ب) **t**

(ج) H

(د) لا شيء مما سبق

(٢٤) قيمة إحصائي الاختبار في هذه الحالة تساوي :-

(أ) -2.05

(ب) -2.94

(ج) **-11.006**

(د) لا شيء مما سبق

(٢٥) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض العدمي .

(ب) **قبول الفرض البديل .**

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

(٢٦) إذا قدمت إليك النتائج التالية كمخرجات للبرنامج الإحصائي SPSS :-

T – TEST

One –Sample test

	Test Value = 160					
	t	df	Sig.(2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
الطول	-11.006	249	0.000	-2.0480	-2.04145	-1.6815

من خلال الجدول السابق يمكن :-

(أ) قبول الفرض العدمي .

(ب) **قبول الفرض البديل .**

(ج) رفض كل من الفرضين .

(د) لا شيء مما سبق

(٢٧) إذا قدمت إليك النتائج التالية كمخرجات للبرنامج الإحصائي SPSS :-

T – TEST

One –Sample test

	Test Value = 160					
	t	df	Sig.(2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
الطول	-1.006	249	0.060	-2.0480	-2.04145	-1.6815

من خلال الجدول السابق يمكن :-

(أ) قبول الفرض العدمي .
(ب) قبول الفرض البديل .
(ج) رفض كل من الفرضين .
(د) لا شيء مما سبق

إذا علمت أنه :-

"أراد باحث أن يعرف أثر استخدام نظم مساندة القرارات على كفاءة القرارات التي تتخذها الإدارة بمساعدة تلك النظم، فوزع ٥٠ مديرا لمنشآت صناعية عشوائيا في مجموعتين، ثم عين أحدهما بطريقة عشوائية لتكون مجموعة تجريبية والأخرى ضابطة، وفي نهاية التجربة وزع على المجموعتان استقصاء يقيس درجة فاعلية القرار وكفاءته عندما يتم اتخاذه باستخدام نظم مساندة القرارات بدلا من الطريقة التقليدية فكانت النتائج كما يلي:

"

المجموعة الضابطة	المجموعة التجريبية
$n_2 = 25$	$n_1 = 25$
$\overline{X_1}=6$	$\overline{X_1}=7.6$
$S_2^2 = 1.78$	$S_1^2 = 2.27$

وأردنا اختيار ما اذا كان أداء المجموعة التجريبية أفضل من أداء المجموعة الضابطة عند مستوى معنوية ٥% :

(٢٨) يمكن صياغة الفرض العدمي و الفرض البديل على التشكل :-

(أ) $Ho: \mu_1 = \mu_2$, $H1: \mu_1 > \mu_2$
(ب) $Ho: \mu_1 = \mu_2$, $H1: \mu_1 < \mu_2$
(ج) $Ho: \mu_1 = \mu_2$, $H1: \mu_1 = \mu_2$
(د) لا شيء مما سبق

(٢٩) درجات الحرية تساوي :-

(أ) 50
(ب) 49
(ج) 48
(د) لا شيء مما سبق

(٣٠) قيمة الانحراف المعياري S في هذه الحالة تساوي :-

(أ) 2.04
(ب) -2.04
(ج) 2.4
(د) لا شيء مما سبق

(٣١) قيمة إحصائي الاختيار t في هذه الحالة تساوي :-

(أ) -1.6
(ب) 1.6
(ج) 2.77
(د) لا شيء مما سبق

(٣٢) من خلال مقارنة قيمة إحصائي الاختيار بقيمة حدود منطقتي القبول والرفض (إذا علمت أن قيمة t الجدولية تساوي 1.68) يمكن :-

(أ) قبول الفرض العدمي .
(ب) قبول الفرض البديل .
(ج) عدم قبول أي من الفرضين .
(د) لا شيء مما سبق

(٣٣)إذا كانت A, B , C ثلاث حوادث فإن العلاقة $A \cup (B \cap C)$ تساوي :-

(أ) $(A \cup B) \cap (A \cup C)$

(ب) $(A \cap B) \cup (A \cap C)$

(ج) $(A \cup B) \cup (A \cup C)$

(د) لا شيء مما سبق

(٣٤)إذا كانت A, B , C ثلاث حوادث فإن العلاقة $A \cap (B \cup C)$ تساوي :-

(أ) $(A \cup B) \cap (A \cup C)$

(ب) $(A \cap B) \cap (A \cap C)$

(ج) $(A \cap B) \cup (A \cap C)$

(د) لا شيء مما سبق

يراد شراء ثلاث أنواع من الكتب الدراسيةA و b وC فإن :-

(٣٥) توافر أنواع الكتب الدراسيةالثلاثة يرمز لها بالرمز :-

(أ) $A \cup B \cup C$

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

(ج) $A \cap B \cap C$

(د) لا شيء مما سبق

(٣٦) عدم توافر الكتب الدراسيةالثلاثة يرمز لها بالرمز :-

(أ) $A \cup B \cup C$

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

(ج) $A \cap B \cap C$

(د) لا شيء مما سبق

(٣٧)توافر نوع واحد من الكتب الدراسيةعلى الأقل A أو B أو C أو كلها يرمز لها بالرمز :-

(أ) $A \cup B \cup C$

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

(ج) $A \cap B \cap C$

(د) لا شيء مما سبق

(٣٨)توافر الكتاب الدراسي A فقط يمكن الرمز له بالرمز :-

(أ) $A \cup B \cup C$

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

(ج) $\overline{A} \cap B \cap C$

(د) لا شيء مما سبق

(٣٩)توافر نوع واحد فقط من الكتب الدراسيةيمكن الرمز له بالرمز :-

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

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

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

(د) لا شيء مما سبق

الجدول التالي يمثل توزيع مجموعة من الطلاب والطالبات حسب التخصص الدقيق بكلية إدارة الأعمال :-

تم اختيار احد الدارسين من الجدول السابق بطريقة عشوائية ، أحسب الاحتمالات التالية :-

	طلاب	طالبات	المجموع
محاسبة	10	14	24
نظم	16	28	44
إدارة	20	12	32
المجموع	46	54	100

(٤٠)احتمال أن يكون طالب :-

(أ) 0.54

(ب) 0.46

(ج) 0.24

(د) لا شيء مما سبق

(٤١)احتمال أن تكون طالبه :-

(أ) 0.54

(ب) 0.46

(ج) 0.24

(د) لا شيء مما سبق

(٤٢)احتمال أن يكون من قسم المحاسبة :-

(أ) 0.54

(ب) 0.46

(ج) 0.24

(د) لا شيء مما سبق

(٤٣)احتمال أن يكون من قسم المحاسبة وطالب :-

(أ) 0.24

(ب) 0.10

(ج) 0.46

(د) لا شيء مما سبق

(٤٤)أن يكون طالبه أو من قسم المحاسبة :-

(أ) 0.64

(ب) 0.78

(ج) 0.54

(د) لا شيء مما سبق

(٤٥)أن يكون من قسم الإدارة أو طالب :-

(أ) 0.78

(ب) 0.32

(ج) 0.58

(د) لا شيء مما سبق

(٤٦)احتمال أن يكون من قسم المحاسبة بشرط أن

تكون طالبة :-

(أ) $\frac{7}{27}$

(ب) $\frac{24}{100}$

(ج) $\frac{54}{100}$

(د) لا شيء مما سبق

(٤٧)احتمال أن يكون طالب بشرط أنه من

قسم الادارة :-

(أ) $\frac{32}{100}$

(ب) $\frac{5}{8}$

(ج) $\frac{20}{100}$

(د) لا شيء مما سبق

إذا علمت أنه :-

" مصنع لإنتاج لعب الأطفال يمتلك ثلاث آلات A وB وC ، تنتج الآلة الأولى %25 من الإنتاج و الألة الثانية %40 من الإنتاج و الباقي من إنتاج الآلة الثالثة فإذا كانت نسبة المعيب في الآلات الثلاثة على الترتيب هو %3و %4 و %6 ، سحبت وحدة واحدة عشوائياً من إنتاج المصنع " ، احسب الاحتمالات التالية :-

(٤٨) احتمال أن تكون الوحدة المسحوبة معيبة :-

(أ) $0.25\times 0.97 + 0.40 \times 0.96 + 0.35 \times 0.94$

(ب) $\mathbf{0.25\times 0.03 + 0.40 \times 0.04 + 0.35 \times 0.06}$

(ج) $0.75\times 0.03 + 0.60 \times 0.04 + 0.65 \times 0.06$

(د) لا شيء مما سبق

(٤٩) احتمال أن تكون الوحدة المسحوبة جيدة :-

(أ) $\mathbf{0.25\times 0.97 + 0.40 \times 0.96 + 0.35 \times 0.94}$

(ب) $0.25\times 0.03 + 0.40 \times 0.04 + 0.35 \times 0.06$

(ج) $0.75\times 0.03 + 0.60 \times 0.04 + 0.65 \times 0.06$

(د) لا شيء مما سبق

(٥٠) احتمال أن تكون الوحدة معيبة و من إنتاج الآلة الثالثة :-

(أ) $\frac{0.94\times 0.35}{\times 0.97+0.40 \times 0.96+0.35\times 0.94}$

(ب) $\frac{0.40\times 0.04}{0.25\times 0.03+0.40 \times 0.04+0.35\times 0.06}$

(ج) $\frac{0.06\times0.35}{0.25\times0.03+0.40\times0.04+0.35\times0.06}$

(د) لا شيء مما سبق

إذا علمت أنه :-

"أحد المصانع وجد أنه من بين كل 1000 وحدة هناك 150 وحدة معيبة ، أخذت عينة مكونة من خمس وحدات ، فإذا علمت أن هذه الظاهرة تتبع التوزيع ثنائي الحدين " أوجد الاحتمالات التالية :-

(٥١) احتمال أن تكون الوحدات المختارة كلها سليمة :-

(أ) 0.5563

(ب) 0.4437

(ج) 0.8352

(د) لا شيء مما سبق

(٥٢) احتمال وجود وحدة على الأكثر معيبة :-

(أ) 0.4437

(ب) 0.3915

(ج) 0.8352

(د) لا شيء مما سبق

(٥٣) احتمال وجود وحدتان معيبتان على الأقل :-

(أ) 0.8325

(ب) 0.1648

(ج) 0.8500

(د) لا شيء مما سبق

(٥٤) القيمة المتوقعة للتوزيع المعبر عن عدد الوحدات المعيبة :-

(أ) 0.15

(ب) 5

(ج) 0.75

(د) لا شيء مما سبق

(٥٥) قيمة التباين للتوزيع المعبر عن عدد الوحدات المعيبة

(أ) 0.6375

(ب) 0.8536

(ج) 0.7984

(د) لا شيء مما سبق

إذا علمت أنه :-

" 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 "

(٥٦) ما نوع المتغير العشوائي :-

(أ) متغير وصفي .

(ب) متغير كمي متصل .

(ج) متغير كمي منفصل .

(د) لا شيء مما سبق

(٥٧) احتمال أن الأسرة تستهلك وحدتين خلال الشهر يساوي :-

(أ) 0.0498

(ب) 0.2240

(ج) 0.4983

(د) لا شيء مما سبق

(٥٨) احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر :-

(أ) 0.4983

(ب) 0.2240

(ج) 0.6474

(د) لا شيء مما سبق

(٥٩) القيمة المتوقعة للتوزيع السابق :-

(أ) 3

(ب) 9

(ج) 1

(د) لا شيء مما سبق

(٦٠) **قيمة الانحراف المعياري للتوزيع السابق تساوي :-**

(أ) 3

(ب) 1.732

(ج) 0.0498

(د) لا شيء مما سبق

(٦١) **معامل الاختلاف النسبي للتوزيع السابق يساوي :-**

(أ) 100%

(ب) 57.7%

(ج) 90%

(د) لا شيء مما سبق

(٦٢) **شكل التوزيع السابق :-**

(أ) توزيع سالب الالتواء .

(ب) توزيع متماثل .

(ج) توزيع موجب الالتواء .

(د) لا شيء مما سبق

(٦٣) عرف كل من المصطلحات التالية :-

١- أسلوب الحصر الشامل .

٢-أسلوب المعاينة .

٣-العينة العشوائية .

٤-العينة المنتظمة .

٥-العينة العنقودية .

٦-العينة الطبقية .

٧-عينة الصدفة .

٨-العيمنة العمدية .

٩-العينة الحصية .

(٦٤) إذا قدمت إليك النتائج التالية كمخرجات للبرنامج الإحصائي SPSS :-

T – TEST

Paired Samples test

		Paired Difference					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Posttest Pretest	4.3800	7.8570	.7857	.3765	5.9390	0.8546	99	.376

من خلال الجدول السابق يمكن :-

(أ) قبول الفرض العدمى .

(ب) قبول الفرض البديل .

(ج) رفض كل من الفرضين .

(د) لا شيء مما سبق

(٦٥) إذا قدمت إليك النتائج التالية كمخرجات للبرنامج الإحصائي SPSS :-

T – TEST

Paired Samples test

		Paired Difference					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Posttest Pretest	4.3800	7.8570	.7857	2.8210	5.9390	5.575	99	.000

من خلال الجدول السابق يمكن :-

(أ) قبول الفرض العدمى .

(ب) قبول الفرض البديل .

(ج) رفض كل من الفرضين .

(د) لا شيء مما سبق

إذا علمت أنه :-

" إذا كان لدينا ثلاث منتجات لإحدى الشركات الصناعية ، وتم تقييمها من قبل مجموعة من المستهلكين وحصلنا على النتائج التالية (عند مستوى معنوية 5%) : -

	المنتج الاول		المنتج الثاني		المنتج الثالث	
	X_1	X_1^2	X_2	X_2^2	X_1	X_2^2
	7	49	4	16	2	4
	10	100	6	36	2	4
	10	100	7	49	3	9
	11	121	9	81	7	49
	12	144	9	81	6	36
المجموع	50	514	35	263	20	102

(٦٦) مجموع المربعات الكلي يساوي :-

(أ) 879

(ب) 105

(ج) 144

(د) لا شيء مما سبق

(٦٧) مجموع المربعات بين المجموعات يساوي :-

<u>(أ) 90</u>

(ب) 105

(ج) 35

(د) لا شيء مما سبق

(٦٨) مجموع المربعات داخل المجموعات :-

(أ) 22

<u>(ب) 54</u>

(ج) 18

(د) لا شيء مما سبق

(٦٩) درجات الحرية الكلية تساوي :-

(أ) 2

(ب) 12

<u>(ج) 14</u>

(د) لا شيء مما سبق

(٧٠) قيمة إحصائي الاختبار F تساوي :-

(أ) 45

<u>(ب) 10</u>

(ج) 15

(د) لا شيء مما سبق

(٧١) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض (إذا علمت أن قيمة Fالجدولية تساوي 3.88) يمكن :-

<u>(أ) قبول الفرض البديل .</u>

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

إذا علمت أنه :-

قام أحد الباحثين بتفريغ ما تم الحصول عليه من معلومات في جدول تحليل التباين كالتالي (عند مستوى معنوية 5%) :

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات Means	قيمة F
بين المجموعات Between groups	200	5	.....	....
داخل المجموعات Within groups	......	......	....	
الكلي (المجموع) Total	280	15		

(٧٢) قيمة إحصائي الاختبار F تساوي :-

(أ) 10

(ب) 5

(ج) 80

(د) لا شيء مما سبق

(٧٣) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض (إذا علمت أن قيمة Fالجدولية تساوي 7.88) يمكن :-

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

الجدول التالي يوضح نتيجة إختبار مربع كاي (كا٢) عند مستوى معنوية 5% :-

Chi-Square Test

	Value	df	Asymp . Sig (2-sided)
Pearson Chi-Square	1.9496	3	.0437
Likelihood Ratio	1.9672	3	.0434
Linear-by- Linear Association	.2384	1	.0390
N of Valid Cases	32		

أجب عن الاسئلة التالية من خلال النتائج الواردة في الجدول السابق :-

(٧٤) قيمة إحصائي الاختبار كا٢ تساوي :-

(أ) .2384

(ب) 1.9672

(ج) 1.9496

(د) لا شيء مما سبق

(٧٥) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

" قام أحد الباحثين بمقارنة عينة من درجات الطلاب في مادة المحاسبة بكلية إدارة الاعمال جامعة الملك فيصل بأخرى من جامعة الدمام وذلك بصدد الوقوف على ما إذا كان هناك اختلاف في متوسط الدرجات وذلك عند مستوى معنوية 5% ، وباستخدام البرنامج الاحصائي SPSS حصلنا على النتائج التالية :-

Test Statistics

	SAMPLES
Mann-Whitney U	44.000
Wilcoxon W	99.000
Z	-.457
Asymp . Sig . (2-tailed)	.648
Exact Sig .[2*(1-tailed Sig.)]	.684

(٧٦) الاختبار المستخدم لدراسة الفرق بين متوسطي مجتمعين في هذه الحالة :-

(أ) كا٢ .

(ب) مان وتني .

(ج) ويلكوكسون .

(د) لا شيء مما سبق

(٧٧) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق .

(٧٨) " لدراسة تأثير أحد البرامج التدريبية على مجموعة من الطلاب تم إختبار مجموعة من الطلاب قبل البرنامج التدريبي على عينة من ٨ طلاب و إختبار الطلاب بعد الحصول على البرنامج التدريبي ولاختبار هل هناك اختلاف معنوي في مستوى تحصيل الطلاب ، عند مستوى معنوية 5% ، أستخدم الباحث البرنامج الاحصائي spss باستخدام إختبار ويلكوكسون Wilcoxon و حصلنا على النتائج التالية :-

Ranks

		N	Mean Rank	Sum of Ranks
AFTER-BEFORE	Negative Ranks	7	2.36	43.50
	Positive Ranks	1	3.54	3.54
	Ties	0		
	Total	8		

Test Statistics

	AFTER-BEFORE
Z	-.313
Asymp . Sig . (2-tailed)	.421

من الجداول السابقة يمكن توضيح أن :-

(أ) مستوى الطلاب قبل الحصول على البرنامج التدريبي أفضل من المستوى بعد الحصول على البرنامج .

(ب) مستوى الطلاب بعد الحصول على البرنامج التدريبي أفضل من المستوى قبل الحصول على البرنامج .

(ج) مستوى الطلاب قبل الحصول على البرنامج التدريبي مساوي لمستوى بعد الحصول على البرنامج .

(د) لا شيء مما سبق