ملخص التحليل الإحصائي

شيءٌ آخر (أبو فيصل)

إهداء لدفعة ٢٠١٣م وأسأل الله بأنني قد وفقت في شرح وتلخيص هذا المقرر ، وأن يكون فيه خير ومنفعة للجميع ، مع أمنياتي لكم بتحقيق أفضل الدرجات في هذا المقرر.

للأستاذ : محمد الخذيف

المحاضرة الأولى

المجموعات

تعريف المجموعة:

المجموعة ببساطة هي تجمع من الأشياء أو العناصر المحددة تماماً ، وقد تكون هذه الأشياء أعداداً أو أشخاصاً أو أحداثاً أو أي شيء آخر.

ترمز للمجموعات بواسطة حروف كبيرة مثل:

$$A, B, C, \cdots$$

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

$$a, b, c, \cdots$$

الانتماء :

- يستخدم الرمز ∈ "ينتمي إلى" ليبين عناصر المجموعة ، فمثلاً إذا كان العنصر a من ضمن عناصر المجموعة A فإننا نقول أن a ينتمي إلى المجموعة A ويكتب بالصورة $a \in A$
- أما إذا كان a ليس عنصرا من عناصر المجموعة A فإننا نقول أن a لا ينتمي إلى المجموعة A ويكتب بالصورة $a \notin A$

ملاحظة : تعد دراسة المجموعات ذات أهمية كبيرة في دراسة العلاقات والدوال.

أمثلة على المجموعات:

مثال:

A={a, b, c, d}

أي أن المجموعة A تتكون من العناصر a و b و c و d

b ∈ A

أي أن العنصر b ينتمي إلى المجموعة A

f ∉ A

أي أن العنصر f لا ينتمي إلى المجموعة A

هذا المثال بسيط وواضح متى ما كان أحد العناصر من ضمن المجموعة نستخدم الرمز ∈
ومتى ما كان العنصر ليس من عناصر المجموعة نستخدم الرمز ∉

➢ طرق كتابة المجموعات:

١- طريقة العد (سرد العناصر):

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

$$A = \{1,3,5,7\}$$
$$B = \{a, b, c, d\}$$
$$C = \{1,2,3,\cdots\}$$

بحيث لا يتم تكرار العناصر

٢- طريقة القاعدة (الصفة المميزة):

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

A={x: x عدد طبيعي زوجي }

B={x: x كلية بجامعة الملك فيصل}

C={x: x طالب مسجل بالمقرر الحالي}

D={x: 0≤x ≤12 ، x عدد صحيح }

يعني هنا أن X عدد أو اسم أو غيره ويكون مشروط بصفه حيث في A تجد أنه مشروط بأن يكون عدد طبيعي زوجي فـ ٢- ليس من ضمنها و ٣ فردي ليس من ضمنها.
في المجموعة D هناك شرطين أو صفتين بأنها أعداد صحيحه وتتكون من ضمن صفر إلى ١٢

مثال على طرق كتابة المجموعات:

فمن خلال رمي حجر نرد مرتين نستطيع أن نعبر عن الحادثة (الحصول على مجموع يساوي ٧) من خلال التالي:

- طريقة سرد جميع العناصر وبينهما فاصلة كالتالي:

A={(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

عملية تخمين كم مرة ممكن يظهر لنا مجموع الرقم سبعه عند رمي النرد لو جمعنا كل رقم نجد أنها تساوي ٧

- ويمكن أن نعبر عن الحادثة نفسها بطريقة الصفة المميزة وهي كتابة مميزات العناصر بين القوسين {} عوضا عن كتابة العناصر نفسها كالتالي:

A={ (x,y) : x + y = 7 }

إذا المجموعة بشكل عام يمكن أن تكتب بميزة عناصرها بأشكال مختلفة طالما كانت الميزة كافيه لتحديد العناصر بشكل دقيق.

➢ أنواع المجموعات:

١- المجموعة الخالية:

وهي مجموعة الأعداد الصحيحة التي بين العددين ٠،١ مجموعة خالية ، أيضا مجموعة أسماء الأسماك التي تتحدث اللغة العربية مجموعة خالية بالتأكيد. ويرمز للمجموعة الخالية بالرمز ∅ أو بقوسين { }.

A={x: x عدد طبيعي زوجي وفردي }

B={x: x دولة عربية تقع في اوروبا }

لا يوجد عدد صحيح بين صفر وواحد ولا يوجد أسماك تتكلم ولا يوجد دوله عربية في أوروبا إذا تكتب قوسين فارغه وتسمى المجموعة الخالية.

٢- المجموعة المنتهية:

المجموعة التي تكون عناصرها محدودة.

مثال: المجموعات التالية مجموعات منتهية

$$A = \{2,4,6,8\}$$
$$B = \{1,2,3,..,100\}$$
$$C = \{x, y, z, w, u\}$$

٣- المجموعة غير المنتهية:

المجموعة التي تكون عناصرها غير محدودة.

مثال: المجموعات التالية مجموعات غير منتهية

A={x: عدد طبيعي فردي x }

$B=\{10,20,30,....\}$

حيث أن X عدد طبيعي فردي ولا يوجد لها نهاية وفي المجموعة B نلاحظ بأنها إلى ما لانهاية.

٤- المجموعة الشاملة:

هي المجموعة التي تدرس جميع المجموعات باعتبارها مجموعات جزئية منها ، ويرمز لها بالرمز U .

٥- المجموعة الجزئية:

فنقول عن مجموعة A أنها مجموعة جزئية subset من مجموعة B إذا كان كل عنصر ينتمي إلى A ينتمي إلى B

ونعبر عن هذا بكتابة التالي :

- إذا كانت $A\subset B$ وكانت $A\neq B$ قلنا أن A جزئية فعلية proper subset من B أو A محتواه في B أو المجموعة B تحتوي A
- أما إذا كانت A=B فإن كل عنصر ينتمي إلى أحدهما ينتمي للأخرى وبالتالي $A\subset B$ و $B\subset A$

أمثلة:

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

فإن $A\subset B$

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

٦- تساوي المجموعات:

تكون المجموعتان A ، B متساويتان إذا كانت

$$A\subseteq B, B\subseteq A\Rightarrow A=B$$

مثال:

$\{-1,+1\}=\{x:x^2=1\}$

{ x حرف من كلمة سلام : x} ≠ {س، ل، م}

أي مربع عدد يساوي واحد حيث ١ في ١ يساوي ١ و -١ في -١ يساوي ١

المجموعة الثانية غير متساوية لأن سلام من أربعة أحرف

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

مثال:

أي المجموعات التالية متكافئة وأيها متساوية؟

1) $A=\{1,3,5,7\}$, $B=\{3,1,5,7\}$

2) $A=\{0,1,2\}$, $B=\{a,b,c\}$

الحل:

1) $A=B$

2) $A\equiv B$

لاحظ أن في المجموعتين في الرقم واحد العناصر متساوية بينما في المجموعتين في الرقم 2 العناصر مختلفة ولكن عددها واحد إذا متكافئة

➢ العمليات على المجموعات:

١- الاتحاد :

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

مثال :

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

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

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

لاحظ أننا لا نكرر الأعداد المكررة في كلتا المجموعتين.

٢- التقاطع :

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

مثال على ذلك :

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

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

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

٣- المكملة أو المتممة:

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

أي أن

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}

$\bar{A}$ ={2, 4, 6, 20}

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

واضحه المجموعة الكلية هي S وتجدون أن المجموعة A لا يوجد بها بعض الأعداد من المجموعة الكلية نضعها في مجموعه أخرى ونسميها مكملة المجموعة A ونرمز لها بالرمز $\bar{A}$ وكذلك الأمر على المجموعة B

٤- الفرق

إذا كانت مجموعتان A ، B فان A-B يسمى بالفرق وهو مجموعة كل العناصر الموجودة في A وليست في B.

مثال:

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

فإن : $A - B = \{1,2, y\}$

أي العناصر التي في A وليست في B واضحه ☺

مثال شامل للعمليات على المجموعات :

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

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

لمعرفة طريقة الحل راجع ما سبق ☺

أوجد :

1) $A \cup B$ 2) $A \cap B$ 3) $A - B$ 4) $\bar{A}$ 5) $\bar{B}$

الحل :

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

$A \cap B = \{3, x\}$

$A - B = \{1,2, y\}$

$\bar{A} = \{4,5, w, z\}$

$\bar{B} = \{1,2, y, z\}$

تدريبات :

١- 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.

$(i)\ 3\in A$
$(ii)\ 3\notin B$
$(iii)\ x\in A$
$(iv)\ x\in B$
$(v)\ z\notin A$
$(vi)\ z\in B$
$(vii)\ 1\notin A$
$(viii)\ 1\notin B$

طبعاً هذا التدريب أتى على شكل فراغات وأنا قمت بحله حيث تضع ينمتي إلى أو لا ينتمى

٢- اسرد عناصر كل مجموعة من المجموعات التالية ، يمكن استخدام النقط للتعبير عن استمرار سرد عناصر المجموعة عندما يكون بها عدد لانهائي من العناصر.

i. A={x: عدد طبيعي اصغر من ٧ x}
ii. B={x: عدد طبيعي زوجي يقبل القسمة على٢ x}
iii. C={y: حرف من حروف الهجاء المحصورين بين c و h y}
iv. D={x: عدد طبيعي فردي اصغر من ١٧ x}

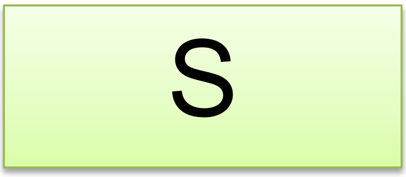
هذا السؤال غير محلول وحله بسيط
١ – ٦
٢ ، ٤ ، ٦ ، .....
نكتب الحروب بين c و h
١٥ ، ١٣ ، ... ، ١

➢ أشكال ِفنْ (VIN Figures).

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

١- المجموعة الشاملة:

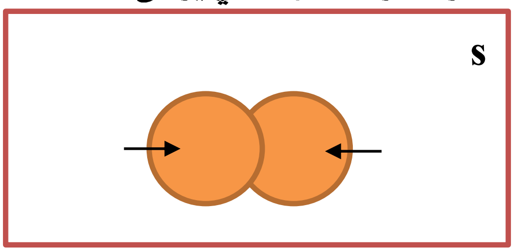
تمثل المجموعات الكلية بمستطيل ويرمز لها بالرمز S



S

٢- اتحاد 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 اتحاد حادثتين ويرمز لها ($A\cup B$) أو (A وأ B) والشكل التالي يوضح ذلك:



شكل فنْ لتمثيل اتحاد حادثتين A و B
($A\cup B$)

وبشكل عام لأي n حادثة A1 , A2, A3, ... An فإن اتحاد هذه 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**

فالاتحاد $\cup$ 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}

نفس ما تم دراسته سابقا في العمليات على المجموعات (الاتحاد)

خواص العمليات الجبرية لاتحاد الحوادث:

- إذا كانت A و B و C ثلاث حوادث فإن :

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

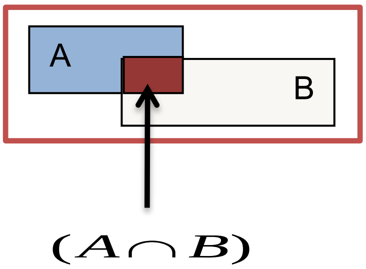
ويعني ذلك توزيع الاتحاد على التقاطع.

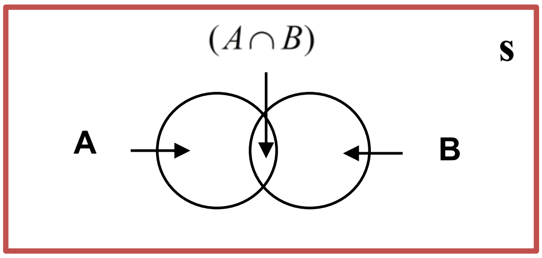
- وكذلك هناك خاصية التبديل والتي تعني أن :

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

٣- **تقاطع الحوادث Events Intersection :**

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





شكل فنْ لتمثيل تقاطع حادثتين A و B

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

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

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.**

فالتقاطع $\cap$ إذاً هو مجموعة العناصر المشتركة بين مجموعتين أو أكثر.

مثال:

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

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

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

نفس ما تم دراسته سابقاً في العمليات على المجموعات (التقاطع)

خواص العمليات الجبرية لتقاطع الحوادث:

- إذا كانت A و B و C ثلاث حوادث فإن :

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

ويعني ذلك توزيع التقاطع على الاتحاد.

- وكذلك هناك خاصية التبديل والتي تعني أن :

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

٤- الحادثة المتممة Complementary Event :

لأي حادثة A فإن متممتها هي الحادثة التي تتضمن كافة العناصر التي لا تنتمي إلى A ويرمز لها بالرمز A^c أو $\overline{A}$ وهو حدث يتألف من جميع عناصر Ω غير المنتمية إلى A وباستخدام أشكال ِفنْ فإن الجزء المظلل يمثل الحادثة المتممة :

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

شكل ِفنْ لتمثيل مكملة الحادثة 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}

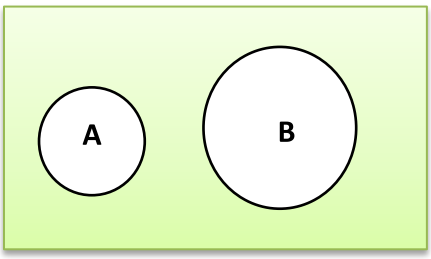
$\overline{B}$ ={4, 5, 7, 9, 10, 13, 15, 17, 18, 19, 20}

نفس ما تم دراسته سابقاً في العمليات على المجموعات (المتممة أو المكملة)

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 $A\cap A^c=\phi$ ، وباستخدام أشكال ِفنْ فإن الحادثتان المنفصلتان يمثلان بالشكل التالي :



$$A\cap B=\phi$$

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

شكل ِفنْ لتمثيل حادثتان متنافيتان A و B

بعض العلاقات المهمة

$\overline{B\cup A}=\overline{B}\cap\overline{A}$	متممة اتحاد مجموعتين يساوي تقاطع متممة كلا المجموعتين	$A\cup A^c=S$	اتحاد أي مجموعة مع متممتها تساوي المجموعة الشاملة
$\overline{B\cap A}=\overline{B}\cup\overline{A}$	متممة تقاطع مجموعتين يساوي اتحاد متممة كلا المجموعتين	$A\cap A^c=\phi$	تقاطع أي مجموعة مع متممتها تساوي المجموعة الخالية
إذا كانت $A\subset B$ فإن : $A=A\cap B$ $B=A\cup B$ $\overline{B}\subset\overline{A}$	عندما نقول أن الـ A جزء من B فإن : الـ A تساوي تقاطع الـ A مع B الـ B تساوي اتحاد الـ A مع B متممة الـ B جزء من متممة الـ B	$\overline{S}=\phi$	متممة المجموعة الشاملة يساوي المجموعة الخالية
		$\overline{\phi}=S$	متممة المجموعة الخالية يساوي المجموعة الشاملة
		$A\cup S=S$	اتحاد أي مجموعة مع المجموعة الشاملة تساوي المجموعة الشاملة
		$A\cap S=A$	تقاطع أي مجموعة مع المجموعة الشاملة تساوي نفس المجموعة
		$A\cap\phi=\phi$	تقاطع أي مجموعة مع المجموعة الخالية يساوي المجموعة الخالية

أسئلة وتمارين :

١- 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 وعدم توفر النوعين الآخرين أي بمعنى : $(A\cap\overline{B}\cap\overline{C})\cup(\overline{A}\cap B\cap\overline{C})\cup(\overline{A}\cap\overline{B}\cap C)$

٢- قذفت قطعة نقود معدنية ثلاث مرات ، أوجد المجموعة الشاملة Ω وعدد عناصرها واكتب الحوادث التالية وعدد عناصر كل منها:

- الحادثة A ظهور صورة في الرمية الأولى.
- الحادثة B ظهور صورة واحدة على الأقل.
- الحادثة C ظهور كتابة في الرمية الأولى وصورة في الرمية الثانية.
- الحادثة $(A\cap B)$
- الحادثة $(A\cup C)$
- الحادثة $(\overline{A}\cup B)$
- الحادثة $(A\cap\overline{B})$
- الحادثة $(\overline{A\cap B})$

الحل :

المجموعة الشاملة Ω يمكن إيجاده من خلال حساب ظهور كل رمية مباشرة على النحو التالي:

Ω= {(HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH), (TTT)}

لقطعة النقد وجهين وجه صورة ووجه كتابة ممكن أن نرمز للصورة بـ H والكتابة بـ T
فنجد في المجموعة الشاملة جميع الاحتمالات عند رمي القطعة ثلاث مرات.

- الحادثة A ظهور صورة في الرمية الأولى:

A= {(HHH), (HHT), (HTH), (HTT)}

هنا نأخذ العناصر التي تبدأ الرمية الأولى بصورة H من المجموعة الشاملة

- الحادثة B ظهور صورة واحدة على الأقل.

B= {(HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH)}

هنا نأخذ العناصر التي بها صورة أو أكثر H من المجموعة الشاملة.

- الحادثة C ظهور كتابة في الرمية الأولى وصورة في الرمية الثانية.

C = {(THH), (THT)}

هنا نأخذ العناصر التي تبدأ الرمية الأولى بكتابة T والرمية الثالثة بصورة H من المجموعة الشاملة.

$(A\cap B)$ = {(HHH), (HHT), (HTH), (HTT)} تقاطع A مع B

$(A\cup C)$ = {(HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH)} اتحاد A مع C

$(\bar{A}\cup\bar{B})$ = {(THH), (THT), (TTH), (TTT)} متممة الـ A اتحاد متممة الـ B

$(A\cap\bar{B})$ = ϕ تقاطع الـ A مع متممة الـ B

$(\overline{A\cap B})$ = { (THH), (THT), (TTH), (TTT)} متممة تقاطع الـ A مع الـ B

المحاضرة الثانية

طرق العد ونظرية الاحتمالات

تعريف الاحتمالات:

- يمكن تعريف الاحتمالات بطرق عديدة غير أن أبسطها

"هو مقياس لإمكانية وقوع حدث (Event) معين"

- 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 كلمة 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 Random Experiment :

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.

نستنتج من ذلك أنه في مثل هذه التجارب تكون النتائج التي نحصل عليها من تكرار للتجربة تتذبذب عشوائيا ومهما حاولنا التحكم بظروف التجربة فإن النتائج المتعاقبة ستتغير بشكل غير منتظم.

٢- الحادث والفراغ العيني:

فراغ العينة : هو المجموعة الشاملة التي تحتوي على جميع النتائج الممكنة للتجربة العشوائية ويرمز له بالرمز Ω ويطلق علية الحالات الممكنة Possible Cases.

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) و مجموعة جميع الحالات الممكنة الظهور تسمى بالفراغ العيني (Sample Space) ويلاحظ أن الحادث قد يكون حالة أو 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.

أمثلة وتمارين :

المثال الأول / أوجد فراغ العينة في كل من التجارب العشوائية التالية:

١- رمي عملة معدنية واحدة.

٢- رمي عملة معدنية مرتين.

٣- رمي حجر نرد مرتين.

الحل :

١- عند رمي عملة معدنية مرة واحدة جميع النتائج الممكن الحصول عليها إما صورة H أو كتابة T، فيكون بالتالي فراغ العينة :

$\Omega = \{H , T\}$

٢- عند رمي عملة معدنية مرتين تكون جميع النتائج الممكن الحصول عليها إما صورة في الرمية الأولى وصورة في الرمية الثانية أو صورة في الرمية الأولى وكتابة في الرمية الثانية أو كتابة في الرمية الأولى وصورة في الثانية أو كتابة في الأولى وكتابة في الثانية فيكون بالتالي فراغ العينة في هذه التجربة :

$\Omega=\{(H,H), (H,T), (T,H), (T,T)\}$

٣- عند رمي حجر نرد مرتين (حجر النرد هو مكعب صغير كتب أو رسم على أوجهه الستة الأرقام من ١ إلى ٦) فتكون جميع النتائج الممكن الحصول عليها إما ظهور رقم ١ في الرمية الأولى ورقم ١ في الرمية الثانية أو رقم ١ في الرمية الأولى ورقم ٢ في الرمية الثانية وهكذا، فيكون بالتالي فراغ العينة في هذه التجربة :

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

كما يمكننا كتابة فراغ العينة في تجربة رمي حجر النرد مرتين على شكل جدول كالتالي:

X,y	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6

المثال الثاني / في تجربة رمي عملة معدنية ثلاث مرات، اكتب فضاء العينة لهذه التجربة وعبر عن الحوادث التالية:

١- الحصول على H (صورة) مرة واحدة.

٢- الحصول على H (صورة) مرتين.

٣- الحصول على H (صورة) ثلاث مرات.

٤- عدم الحصول على H (صورة).

الحل:

❖ عند رمي عملة معدنية ثلاث مرات فيكون بالتالي فراغ العينة :

Ω= {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

١- ويمكن الحصول على الحادثة H (صورة) لمرة واحدة ونرمز لها بالرمز A1 كالتالي:

A1 ={HTT, THT, TTH}

٢- ويمكن الحصول على الحادثة H (صورة) لمرتين ونرمز لها بالرمز A2 كالتالي:

A2 ={HHT, HTH, THH}

٣- - ويمكن الحصول على الحادثة H (صورة) ثلاث مرات ونرمز لها بالرمز A3 كالتالي:

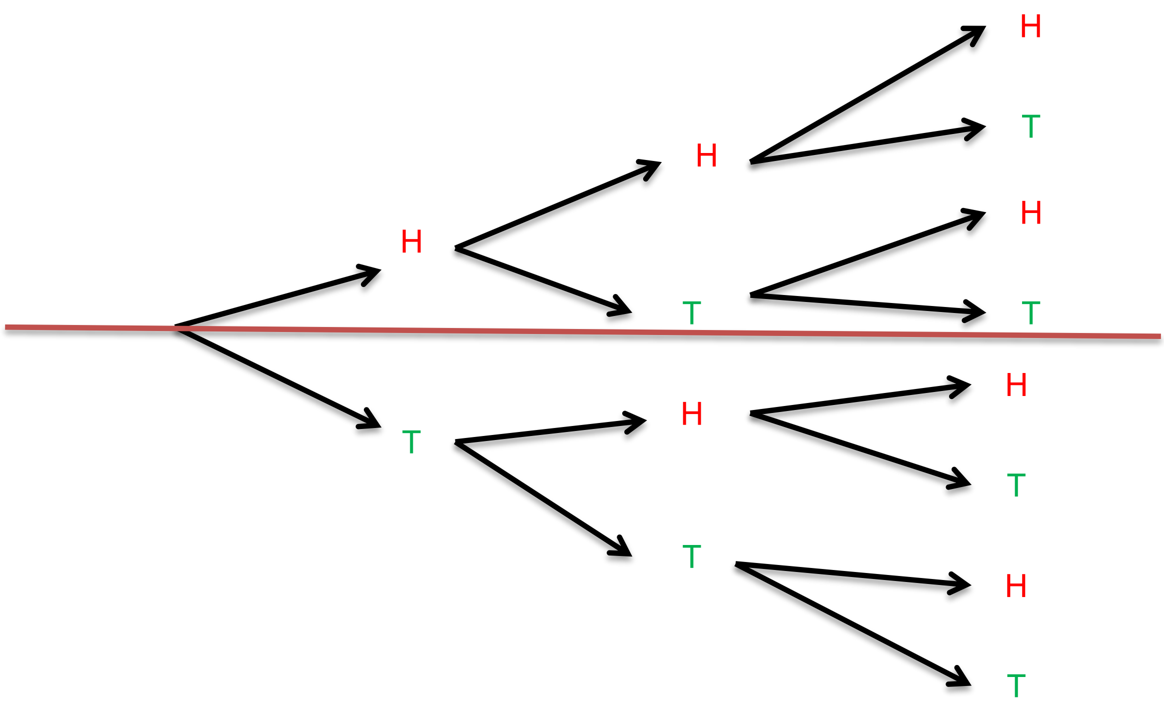
A3 ={HHH}

٤- ويمكن عدم الحصول على الحادثة H (صورة) ونرمز لها بالرمز A4 كالتالي:

A4 ={TTT}

هذا المثال بسيط وواضح حيث نرمز للصورة بـ H ونرمز للكتابة بـ T فعندما نرمي قطعة العملة ثلاث مرات تكون جميع الاحتمالات كما هو في فراغ العينة . وبعد ذلك نتابع النقاط من واحد إلى أربعه فمثلاً إذا رمينها ثلاث مرات يمكن أن نرى صوره في ثلاث احتمالات كما في A1 وهكذا.

ويمكن من خلال استخدام الرسم الشجري معرفة فراغ العينة للمثال السابق (في تجربة رمي عملة معدنية ثلاث مرات) كالتالي:



المثال الثالث /

في طريقك إلى الجامعة توجد إشارتا مرور ، ما هو فضاء العينة لتجربة ذهابك إلى الجامعة؟

الحل:

نفترض أنه عندما تكون الإشارة خضراء نرمز لها بالرمز G وعندما تكون حمراء نرمز لها بالرمز R فيكون بالتالي فضاء العينة كالتالي:

Ω={GG, GR, RG, RR}

هنا أيضا مثال بسيط اشارتين ممكن أجد كلها GG وممكن أجد واحدة خضراء وواحدة حمراء وهكذا تكون الاحتمالات.

المثال الرابع / في تجربة رمي **حجر نرد مرتين** عبر عن الحوادث التالية بدلالة نقاط العينة والصفة المميزة؟

١- A : الحصول على مجموع يساوي ٧
٢- B : الفرق بين العددين الناتجين يساوي القيمة المطلقة ١
٣- C : الحصول على مجموع يساوي ٩ على الأقل
٤- D : الحصول على ١ في الرمية الأولى
٥- E : الحصول على حاصل ضرب يساوي ٦ على الأكثر
٦- F : الحصول على مجموع أقل من أو يساوي ٢

الحل:

يمكننا كتابة فراغ العينة في تجربة رمي حجر النرد مرتين على شكل جدول كالتالي:

X,y	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6

وإذا رمزنا للرمية الأولى بـ x والرمية الثانية بـ y فإنه يمكننا كتابة الحوادث المطلوبة في السؤال على النحو التالي :

١- الحصول على مجموع يساوي 7 :

- بطريقة سرد جميع العنااصر بينهما فاصلة (طريقة نقاط العينة):

A={(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

- بطريقة الصفة المميزة:

A={ (x,y) : x + y = 7 }

٢- الفرق بين العددين الناتجين يساوي القيمة المطلقة 1

- بطريقة سرد جميع العناصر بينهما فاصلة (طريقة نقاط العينة):

B={(1,2), (2,1), (2,3), (3,2), (3,4), (4,3), (4,5), (5,4), (5,6), (6,5)}

- بطريقة الصفة المميزة:

B={ (x,y) : x – y = | 1 | }

٣- الحصول على مجموع يساوي 9 على الأقل :

- بطريقة سرد جميع العناصر بينهما فاصلة (طريقة نقاط العينة):

C={(4,5), (4,6), (5,4), (5,5), (5,6), (6,3), (3,6) , (6,4), (6,5), (6,6)}

- بطريقة الصفة المميزة:

C={ (x,y) : x + y ≥ 9 }

٤- الحصول على الرقم 1 في الرمية الأولى :

- - بطريقة سرد جميع العناصر بينهما فاصلة (طريقة نقاط العينة):

D={(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)}

- – بطريقة الصفة المميزة:

D={ (x,y) : x = 1 }

٥- الحصول على حاصل ضرب يساوي 6 على الأكثر :

- - بطريقة سرد جميع العناصر بينهما فاصلة (طريقة نقاط العينة):

E={(1,1), (1,2), (1,3), (1,4), (1,5), (2,1), (2,2), (2,3), (3,1), (3,2), (4,1), (1,6), (6,1), (5,1)}

- - بطريقة الصفة المميزة:

$E=\{ (x,y) : x * y \le 6 \}$

٦- الحصول على مجموع أقل من أو يساوي 2 :

- - بطريقة سرد جميع العناصر بينهما فاصلة (طريقة نقاط العينة):

F={(1,1)}

- - بطريقة الصفة المميزة:

$F=\{ (x,y) : x + y \le 2 \}$

المثال الرابع / عبر بالكلمات عن كل الحوادث الممثلة بالمجموعات الجزئية التالية من نقاط العينة:

G ={(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}

H ={(1,1), (2,1), (1,2), (3,1), (2,2), (1,3)}

I ={(5,1), (1,5), (6,2), (2,6)}

J ={(1,4), (2,4), (3,4), (4,4), (5,4), (6,4)}

K ={(2,2), (2,4), (2,6), (4,2), (4,4), (4,6), (6,2), (6,4), (6,6)}

الحل :

الحادثة	التعبير بالكلمات عن الحوادث
الحادثة G	تعني الحصول على نفس العدد في الرمية الأولى والرمية الثانية
الحادثة H	تعني الحصول على مجموع رميتين أقل من (٥)
الحادثة I	تعني الحصول على فرق بين الرميتين يساوي (٤)
الحادثة J	تعني الحصول على (٤) في الرمية الثانية
الحادثة K	تعني الحصول على عدد زوجي في كلا الرميتين

٣- **الحالات الممكنة (Possible Cases)**

هي الحالات أو النتائج المختلفة التي يمكن أن تظهر نتيجة لإجراء تجربة معينة ، فمثلاً : عند رمي قطعة عملة تكون نتيجتها صورة أو كتابة.

وعند رمي زهرة نرد تكون نتيجتها ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ فيقال أن عدد الحالات الممكنة ٢ في حالة رمي قطعة العملة و ٦ في حالة رمي زهرة النرد.

٤- **الحالات المواتية (Favorable Cases)**

هي النتائج أو الحالات التي تؤدي إلى تحقيق الحادث الذي هو موضع اهتمامنا ، فإذا كان الحادث هو الحصول على رقم فردي في حالة رمي زهرة النرد فإن الحالات التي تحقق هذا الحادث هي الحصول على ١ أو ٣ أو ٥ ، هذه الحالات الثلاثة تسمى الحالات المواتية.

٥- الحالات المتماثلة (Equally Likely Cases)

إذا كان لدينا عدة كرات معدنية مصنوعة من مادة واحدة متجانسة في الكثافة ولها نفس الوزن والحجم وضعناها في كيس وسحبنا كرة منها بعد خلطها جيداً فإن هذه الكرات تكون حالات متماثلة أي يكون لكل منها نفس النصيب في السحب.

٦- الحوادث المتنافية (Mutually Exclusive Events)

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

٧- الحوادث المستقلة (Independent Events)

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

٨- الحوادث الشاملة (Exhaustive Events)

تسمى الحوادث A ، B ، C ... 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.

➢ طرق العد :

إن من المهم لحساب احتمال حدوث معين في تجربة ما هو أن نعرف عدد مرات حدوث هذا الحادث بالنسبة لعدد الاحتمالات الممكنة لتلك التجربة ، وقد يكون من السهل عد الاحتمالات الممكنة ومرات حدوث ذلك الحادث في بعض التجارب كما في تجربة إلقاء حجر نرد أو قطعة نقد، إلا إنه في كثير من الأحيان يصعب فعل ذلك باستخدام العد لكل احتمال ممكن بعينه ، فيستلزم بالتالي أن نستخدم طرق رياضية لحساب عدد مرات الحدوث بدون الحاجة لمعرفة كل عنصر بالتحديد من عناصر مجموعتي الحادث والمجموعة الشاملة.

١- طريقة الضرب :

إذا كانت التجربة E_1 تحدث n من الطرق ومع كل طريق من هذه الطرق كانت التجربة E_2 تحدث في m من الطرق فإن التجربتين تحدثان معاً في $n \times m$ من الطرق.

مثال (١) :

إذا كان هناك طريقان يمكن أن يستخدمها المسافر من الأحساء إلى الرياض ، و ٣ طرق مختلفة يمكن أن يستخدمها المسافر من الرياض إلى مكة المكرمة.

فإن عدد الطرق التي يمكن أن يستخدمها المسافر من الأحساء إلى مكة المكرمة مروراً بالرياض هي :

- السفر من الأحساء إلى الرياض : E_1 ويحصل في عدد من الطرق مقداره : $n = 2$
- السفر من الرياض إلى مكة المكرمة : E_2 ويحصل في عدد من الطرق مقداره : $m = 3$
- إذا عدد الطرق للسفر من الأحساء إلى الرياض مروراً بالرياض : $n \times m = 2 \times 3 = 6$

مثال (٢) :

إذا فرض أن بإمكان طالب أن يسجل ٣ مقررات هذا الفصل ، بحيث يختار مقرر من قسم المحاسبة من بين ٤ مقررات متاحة ، ويختار مقرر واحد من قسم التمويل من بين ٣ مقررات متاحة ، ومقرر واحد من قسم إدارة الأعمال من بين مقررين متاحين .

عدد الطرق لاختيار هذه المقررات الثلاثة : $4 \times 3 \times 2 = 24$

٢- طريقة الجمع :

إذا كانت تجربتان مانعتين لبعضهما البعض وكانت الأولى تحدث في n من الطرق وكانت الثانية تحدث في m من الطرق فإن واحدة منهما أو الأخرى تحدث في $n + m$ من الطرق.

مثال :

إذا فرض أن طالباً من كلية العلوم الإدارية متاح له أن يسجل مقرر واحد فقط من كلية التربية كمتطلب للتخرج بحيث يختاره حسب اختياره من بين الأقسام العلمية في الكلية المتاحة له ، ما عدد الطرق لاختيار هذا المقرر إذا علمت أن المقررات المتاحة له كالآتي :

م	القسم	عدد المقررات المتاحة
1	الدراسات الإسلامية	3
2	اللغة العربية	4
3	اللغة الإنجليزية	2
4	علم النفس	1

عدد الطرق المتاحة لاختيار مقرر من بين هذه المقررات هو :

$3 + 4 + 2 + 1 = 10$

٣- المضروب :

عدد الطرق التي يمكن أن نرتب بها n من الأشياء ويرمز له بالرمز $n!$

$$n! = n \times (n-1) \times (n-2) \times \dots\dots\dots \times 3 \times 2 \times 1$$

ملاحظة:

$0! = 1$

مضروب الصفر دائماً يساوي واحد

ملاحظة:

$$\begin{aligned} n! &= n \times (n-1)! \\ &= n \times (n-1) \times (n-2)! \\ &= n \times (n-1) \times (n-2) \times (n-3)! \end{aligned}$$

أمثلة (١) :

$$2! = 2 \times 1 = 2$$
$$3! = 3 \times 2 \times 1 = 6$$
$$4! = 4 \times 3 \times 2 \times 1 = 24$$
$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$
$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$
$$\begin{aligned} 5! &= 5 \times 4 \times 3 \times 2 \times 1 \\ &= 5 \times 4! \\ &= 5 \times 4 \times 3! \\ &= 5 \times 4 \times 3 \times 2! \\ &= 120 \end{aligned}$$

في هذا المثال نجد أن مضروب الـ 3 مثلاً يساوي :

$n! = n \times (n-1) \times (n-2)$

$3! = 3 \times (3-1) \times (3-2)$

$3! = 3 \times 2 \times 1 = 6$

وهكذا على البقية ثم فصل لنا مضروب الـ 5

حيث نقدر نقول أن مضروب الـ 5 هو :

$5! = 5 \times 4!$

$= 5 \times 24 = 120$

أمثلة (٢) :

$$\frac{7!}{6!} = \frac{7 \times 6!}{6!} = 7$$
$$\frac{7!}{5!} = \frac{7 \times 6 \times 5!}{5!} = 7 \times 6 = 42$$

في هذا المثال نجد أن مضروب الـ 7 يساوي :

$n! = n \times (n-1)!$

$7! = 7 \times (7-1)!$

$7! = 7 \times 6!$

وهذا كما ظهر لنا بالبسط ويروح مضروب الـ 6 في البسط مع مضروب الـ 6 في المقام يبقى الناتج 7 ، وكذلك الأمر على المثال الأخر.

٤- التباديل :

هي ترتيب جميع عناصر أو جزء من عناصر أي مجموعة ويرمز له بالرمز : P

والتباديل يأخذ صوراً مختلفة يمكن تصنيفها كالآتي :

أ- ترتيب n من الأشياء المميزة مأخوذة سوياً (جميعها) :

$$P(n,n) = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

مثال (١) :

ترتيب الحروف :

$$a, b, c$$
$$abc, acb, bac, bca, cab, cba$$
$$p(n,n) = n! = 3 \times 2 \times 1 = 6$$

هنا قلنا أن :

$P(n,n) = n!$

مباشرة نحسب n! وذلك مثل ما درسناه سابقاً في المضروب.

مثال (٢) :

ترتيب ستة أشخاص على ستة كراسي :

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$

هنا قلنا أن :

$P(n,n) = n!$

مباشرة نحسب n! (6!) وذلك مثل ما درسناه سابقاً في المضروب.

الحل بالألة الحاسبة: (طريقة التباديل)

للمثال الأول : ندخل الرقم 3 ثم Shift ثم علامة الضرب ثم 2 ثم = يطلع لنا الناتج 6

للمثال الثاني : ندخل الرقم 6 ثم Shift ثم علامة الضرب ثم 6 ثم = يطلع لنا الناتج 720

ب- ترتيب n من الأشياء المميزة مأخوذة r في كل مرة حيث : $r \le n$

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

مثال (١) :

بكم طريقة يمكن ترتيب حرفين من الحروف : $a,b,c,d.e$

$$P(5,2)=\frac{5!}{(5-2)!}=\frac{5!}{3!}=\frac{5\times4\times\cancel{3!}}{\cancel{3!}}=5\times4=20$$

عدد الحروف 5 وطلب في السؤال ترتيب حرفين فعوضنا بهذه الطريقة وطبقنا القانون.

طريقة أخرى :

$$P(5,2)=5\times4=20$$

هنا نحل بطريقة أسهل بدون استخدام القانون حيث نضرب الخمسة في العدد 1-5 لأن r = 2

ولتوضيح ذلك أكثر :

$$P(100,3)=100\times99\times98=970200$$

$$P(32,4)=32\times31\times30\times29=863040$$

وهنا نفس الشيء مرة الـ r = 3 ومرة تساوي 4 وذلك للتوضيح وجرب حسابها بالقانون ☺

مثال (٢) :

ترتيب 4 من الأشخاص على 6 كراسي في خط أفقي :

$$P(6,4)=\frac{6!}{2!}=\frac{6\times5\times4\times3\times2\times1}{2\times1}=360$$

وبالطريقة الأسهل ☺ $P(6,4)=6\times5\times4\times3=360$

ج- ترتيب n من الأشياء التي من بينها n_1 عنصراً متماثلاً ، و n_2 عنصراً متماثلاً و n_3 عنصراً متماثلاً ... الخ :

$$\frac{n!}{n_1!\times n_2!\times n_3!\ldots}$$

مثال :

بكم طريقة يمكن ترتيب كلمة : ***Statistics***

$$\frac{10!}{3!\times3!\times2!\times1!\times1!}=\frac{10\times9\times8\times7\times6\times5\times4\times3!}{3!\times3!\times2!\times1!\times1!}$$

$$=10\times9\times8\times7\times6\times5\times2=50400$$

لاحظ كلمة ***Statistics*** بها أحرف مكرره لذلك تجد وضعنا في المقام أن S تكرر 3 مرات وكذلك الـ t أما الـ i تكرر مرتين و الـ a والـ c لم تتكرر فعبرنا عنها بـ !1 ، وعدد حروف الكلمة 10 لذلك وضعنا البسط !10

ولو كانت كلمة أخرى من عشرة أحرف لا يوجد بها أحرف مكررة نحسب !10 فقط.

٥- **السحب مع الإرجاع :**

مثال (١) :

بكم طريقة يمكن ترتيب رقم مكون من عشر خانات من بين الأعداد : $0,1,2,3,4,5,6,7,8,9$

الخانة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الخيارات الممكنة	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠	١٠

$$10^{10}=10,000,000,000$$

مثال (٢) :

بكم طريقة يمكن ترتيب رقم مكون من ثلاث خانات من بين الأعداد : $0,1,2,3,4,5,6,7,8,9$

الخانة	١	٢	٣
الخيارات الممكنة	١٠	١٠	١٠

$$10^3 = 1000$$

٦- السحب بدون إرجاع :

نفس طريقة المضروب ، حيث لا تتكرر الأعداد

مثال (١) :

بكم طريقة يمكن ترتيب رقم مكون من عشر خانات من بين الأعداد : $0,1,2,3,4,5,6,7,8,9$

الخانة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الخيارات الممكنة	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١

$$10! = 10\times9\times8\times7\times6\times5\times4\times3\times2\times1 = 3628800$$

مثال (٢) :

بكم طريقة يمكن ترتيب رقم مكون من ثلاث خانات من بين الأعداد : $0,1,2,3,4,5,6,7,8,9$

الخانة	١	٢	٣
الخيارات الممكنة	١٠	٩	٨

$$P(10,3) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{10\times9\times8\times\cancel{7!}}{\cancel{7!}} = 10\times9\times8 = 720$$

٧- التوافيق :

هي الطرق التي نختارها عدداً معيناً من عناصر مجموعة معينة دون النظر إلى الترتيب ، ويرمز له بالرمز : $C(n,r)$

ويتم استخدام العلاقة التالية :

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

مثال (١) :

ما عدد الطرق التي نختارها حرفين من الحروف : a,b,c

الحل : ab, ac, bc

نلاحظ هنا بأنه ذكر في السؤال نختار حرفين
في المعادلة نعوض عدد الحروف n وهو ثلاثة
والاختيار r و حرفين 2

$$C(n,r) = \binom{3}{2} = \frac{3!}{2!\times(3-2)!} = \frac{3\times\cancel{2!}}{\cancel{2!}\times1!} = 3$$

الحل بالآلة الحاسبة: (طريقة التوافيق)

للمثال الأول : ندخل الرقم 3 ثم Shift ثم علامة القسمة ثم 2 ثم = يطلع لنا الناتج 3

مثال (٢) :

بكم طريقة يمكن اختيار **4** طلاب من بين **20** طالباً لا عفاءهم من دخول الاختبار ؟ :

الحل :

$$C(20,4) == \frac{20!}{4!\times 16!} = \frac{20\times 19\times 18\times 17\times \cancel{16!}}{4!\times \cancel{16!}} = \frac{116280}{24} = 4845$$

الحل بالألة الحاسبة: (طريقة التوافيق)

للمثال الثاني : ندخل الرقم **20** ثم **Shift** ثم **علامة القسمة** ثم **4** ثم **=** يطلع لنا الناتج **4845**

مثال (٣) : يحل بالألة بنفس الطريقة السابقة ☺

إذا فرض أن طالباً يحق له تسجيل **5** مقررات هذا الفصل من بين **8** مقررات متاحة له ، فبكم طريقة يمكنه اختيار هذه المقررات الخمسة ؟

الحل :

$$C(8,5) == \frac{8!}{5!\times 3!} = \frac{8\times 7\times 6\times \cancel{5!}}{\cancel{5!}\times 3!} = \frac{336}{3\times 2\times 1} = 56$$

ملاحظات :

$\binom{n}{r} = \binom{n}{n-r}$ ، $\binom{n}{0} = 1$ ، $\binom{n}{n} = 1$ ، $\binom{n}{1} = n$

مثال (٤) :

قيمة كلاً مما يلي :

- $\binom{12}{1} = 12$
- $\binom{20}{20} = 1$
- $\binom{23}{23} = 1$
- $\binom{9}{3} = \frac{9!}{3!\times 6!} = \frac{9!}{6!\times 3!} = \binom{9}{6}$

عدد اختيار عنصر من 12 = 1
عدد اختيار 20 عنصر من 20 = 1
وهكذا

عدد اختيار ثلاث من بين تسعة يساوي عدد اختيارسته من بين تسعة.
والسبب أن الفرق بين تسعة وثلاثة يساوي سته ، والفرق بين تسعة وسته يساوي ثلاثة.

المحاضرة الثالثة

نظرية الاحتمالات

مقدمه :

ترتبط كلمة احتمال دائما بذكر حدثا ما ، فاحتمال وقوع حدث معين هو نسبة وقوع هذا الحدث في الأجل الطويل ، فعندما نقول إن احتمال الحصول على وحدة معيبة من إنتاج إحدى الآلات هو 0.08 ، فإننا نعني بذلك أن 8 في المئة من إنتاج هذه الآلة في الأجل الطويل سيكون معيبا ، ونحن لا نستطيع ضمان وجود نسبة معينة من الوحدات المعيبة بالضبط في أي 100 وحدة من إنتاج هذه الآلة ، ولكننا نتوقع أن نجد نسبة معينة في المئة من إنتاج هذه الآلة معيبا إذا فحصنا عددا كبيرا وكافيا من إنتاجها.

وللاحتمالات تعريفات عدة سنتعرض لها فيما يلي:

- **التعريف النسبي للاحتمالات Rational Probability Definition :**

عند إجراء تجربة عدد N من المرات وكان A حدثا عشوائيا متعلقا بهذه التجربة فإن التكرار النسبي يعرف على أنه حاصل قسمة عدد مرات حدوث الحادثة A مقسوما على عدد مرات حدوث التجربة أو بمعنى : $f_N(A)=\frac{N_A}{N_\Omega}$ أي أن التكرار النسبي لـ A يساوي تكرار A مقسوما على التكرار الكلي.

- $0\le f_N(A)\le 1$ التكرار النسبي لـ A أكبر من أو يساوي (صفر) وأصغر من أو يساوي (١)
- $f_N(A)=1$ إذا وفقط إذا وقع الحدث A في N مرة أجريت فيها التجربة.
- $f_N(A)=0$ إذا وفقط إذا لم يقع الحدث A في N مرة أجريت فيها التجربة.
- إذا كان A و B حادثتان متنافيتان فإن : $f_N(A\cup B)=f_N(A)+f_N(B)$
- إن التكرار النسبي $f_N(A)$ لحادثة A يأخذ قيمة ثابتة إذا زاد عدد محاولات التجربة عن عدد معين ، ويكون في العادة العدد كبير ، وهذا ما نسميه احتمال وقوع الحادثة A .

مثال :

إذا أخذنا التجربة العشوائية رمي قطعة نقود عدة مرات وتسجيل عدد مرات ظهور الحادثة الحصول على صورة H وكررنا التجربة لعدد من المرات سجلت النتائج في الجدول التالي:

عدد الرميات الكلي N	عدد مرات ظهور الصورة H	التكرار النسبي
30	12	12 / 30
50	20	20 / 50
80	38	0.475
100	49	0.49
300	150	0.5
500	250	0.5
1000	500	0.5
1500	750	0.5

نلاحظ : أنه كلما زاد عدد الرميات N فإن التكرار النسبي يختلف اختلافا بسيطا حتى تستقر عند قيمة معينة 0.5 وهذا يوضح لنا تعريف الاحتمال وفق مفهوم التكرار النسبي.

➢ التعريف التقليدي للاحتمالات Classical Probability Definition :

لأي حادثة A فإن احتمال حدوثها يمثل نسبة عدد حالات ظهورها إلى عدد حالات ظهور فراغ العينة الكلي أي بمعنى:

$$P(A)=\frac{N_A}{N_\Omega}$$

أي أن :
$$\frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}}$$

وللتعريف التقليدي للاحتمال عدد من المسلمات وهي :

- **قيمة أي احتمال أكبر من أو يساوي صفر ، بمعنى أنه لأي حادثة A فإن :** $P(A)\geq 0$
- **قيمة احتمال فراغ العينة يساوي واحد صحيح :** $P(\Omega)=1$
- **إذا كانت A1 و A2 حادثين متنافيين أو منفصلين بمعنى :** $A_1\cap A_2=\phi$
فإن : $P(A_1\cup A_2)=P(A_1)+P(A_2)$
- **ويمكن القول بشكل عام لأي n حادثة منفصلة :**

$$P(A_1\cup A_2\cup A_3\cup .....\cup A_n)=P(A_1)+P(A_2)+P(A_3)...+P(A_n)$$

مثال (١):

رمي حجر نرد مرة 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**

لو رجعنا لما درسناه بالإحصاء التكرار النسبي
نجد بأنه نفسه التكرار على المجموع الكلي

الحل:

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

P(A=5)=1/6
P(A=2,4, 6)=3/6
P(A>2)=4/6
P(A<7)=6/6=1
P(A=7)=0/6=0

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

مثال (٢):

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

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	٥	٧	١٢
القسم الثاني	٨	١٤	٢٢
القسم الثالث	١٠	٦	١٦
المجموع	٢٣	٢٧	٥٠

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.

تذكرون التكرار النسبي في الإحصاء في الإدارة نفسه بالضبط التكرار تقسيم مجموع التكرار.

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 الاحتمال المطلوب:

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

- نفرض أن الحادثة C أن يكون العامل من القسم الأول أي أن C = {أن يكون العامل من القسم الأول} فيكون الاحتمال المطلوب:

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

- نفرض أن الحادثة D أن يكون العامل من القسم الأول أو الثاني أي أن D = {أن يكون العامل من القسم الأول أو الثاني} فيكون الاحتمال المطلوب:

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

- نفرض أن الحادثة E أن يكون العامل من القسم الأول و أعزب أي أن E = {أن يكون العامل من القسم الأول وأعزب} فيكون الاحتمال المطلوب:

$$P(D)=\frac{\text{عدد عمال القسم الأول وعزاب}}{\text{عدد العمال الكلي}} = 5/50=0.1$$

بدهيات الاحتمال :

١- احتمال وقوع الحادث الأكيد يساوي واحداً: $P(S)=1$

مثلاً : احتمال ظهور رقم من أرقام النرد الستة عند رمية.

٢- احتمال وقوع الحادث المستحيل يساوي صفراً $P(\phi)=0$

مثلاً : احتمال ظهور رقم سبعة عند رمي حجر النرد.

٣- إذا كان الحادث A مجموعة جزئية من الفضاء العيني S فإن: $0 \le P(A) \le 1$

قيمة الاحتمال دائماً تكون محصورة ما بين واحد وصفر.

٤- إذا كان A و B حادثين منفصلين بمعنى : $P(A \cap B)=0$

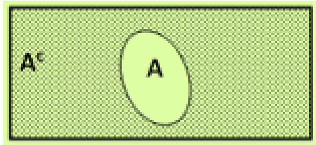
فإن :

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

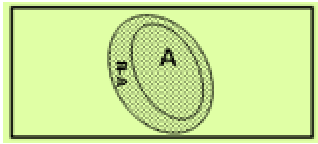
نظريات الاحتمال :

قد يأتي في الاختبار صورة ويطلب العلاقة أو النظرية الخاصة بهذه الصورة وأيضاً لا تنسوا أشكال فن في المحاضرة الأولى ☺

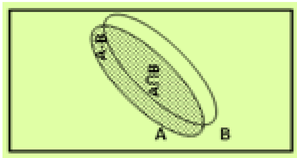
١- إذا كانت A^c متممة الحادث A فإن : $P(A^c) = 1 - P(A)$



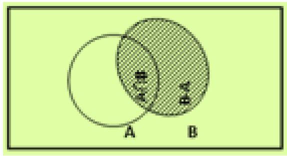
٢- إذا كان $A \subset B$ فإن : $P(A) \le P(B)$



٣- إذا كان **A** و **B** أي حادثتين فإن : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



٤- إذا كان **A** و **B** أي حادثين فإن : $P(B - A) = P(B) - P(A \cap B)$



مثال (١) :

أجري امتحانان في مادة الإحصاء على 100 طالب فنجح في الامتحان الأول 60 طالباً ونجح في الامتحان الثاني 40 طالباً ونجح في الامتحانين معاً 20 طالباً ، أوجد احتمال نجاح الطلاب في الامتحان الأول واحتمال نجاح الطلاب في الامتحان الثاني واحتمال نجاح الطلاب في الامتحانين معاً ، ثم أوجد احتمال نجاح الطلاب في الامتحان الأول أو الامتحان الثاني.

أولاً نفرض :

- نجاح الطلاب في الامتحان الأول = **A**
- نجاح الطلاب في الامتحان الثاني = **B**
- نجاح الطلاب في الامتحانين معاً = $A \cap B$
- نجاح الطلاب في أحد الامتحانين = $A \cup B$

هذا المثال واضح لو أخذنا أول سؤال نجاح الطلاب في الامتحان الأول عددهم ستين نقسمه على العدد الكلي للطلاب اللي هو 100 يعطينا الناتج وهكذا

الحل :

- $P(A) = \dfrac{60}{100}$
- $P(B) = \dfrac{40}{100}$
- $P(A \cap B) = \dfrac{20}{100}$

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

$$= \frac{60}{100} + \frac{40}{100} - \frac{20}{100}$$

$$= \frac{80}{100}$$

مثال (٢) :

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

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	٥	٧	١٢
القسم الثاني	٨	١٤	٢٢
القسم الثالث	١٠	٦	١٦
المجموع	٢٣	٢٧	٥٠

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

نعوض مباشرة من الجدول ونقسم على المجموع الكلي ٥٠ ودائماً اقرأ السؤال صحيح هل هو مشروط أم لا حيث سـ ندرس ذلك لاحقاً في الاحتمال الشرطي

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 u 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 u A2) =P(A1)+P(A2)-P(A1 ∩ 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 u A2) =P(A1)+P(A2)-P(A1 ∩ A2) = (16/50)+(23/50)-(10/50) = 29/50= 0.58

الاحتـمال الشرطي Conditional Probability

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

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

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

مثال (١) :

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}

هنا نقسم مباشرة احتمال نجاحه في مقرر الاحصاء ومقرر الرياضيات على احتمال نجاحه في مقرر الرياضيات يعطينا احتمال نجاحه في الاحصاء.

وبذلك يكون:

$P(A2)=0.64$

$P(A1 \cap A2)=0.32$

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

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

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

مثال (٢) :

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

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	٥	٧	١٢
القسم الثاني	٨	١٤	٢٢
القسم الثالث	١٠	٦	١٦
المجموع	٢٣	٢٧	٥٠

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

١- احسب احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج؟

٢- 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}

$B1$= {أن يكون العامل من القسم الثالث}

$B2$= {أن يكون العامل أعزب}

في هذا المثال نعوض مباشرة من الجدول ونقسم على المجموع الكلي ٥٠ ودائماً اقرأ السؤال صحيح هل هو مشروط أم لا.

فيكون بالتالي:

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

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

$$P(A1|A2) = \frac{P(A_1 \cap A_2)}{P(A_2)} = \frac{\frac{7}{50}}{\frac{27}{50}} = \frac{7}{27} = 0.259$$

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:

$$\frac{\text{احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث}}{\text{احتمال أن يكون من القسم الثالث}}$$

$$P(B1|B2) = \frac{P(B_1 \cap B_2)}{P(B_2)} = \frac{\frac{10}{50}}{\frac{16}{50}} = \frac{10}{16} = 0.625$$

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

➢ **ضرب الاحتمالات:**

قانون ضرب الاحتمالات:

من قانون الاحتمال الشرطي نستنتج:

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

نستنتج أن:

$$P(A \cap B) = P(B) \times P(A \setminus B)$$
$$P(A \cap B) = P(A) \times P(A \setminus B)$$

ولو كانت ثلاثة حوادث تكون كالتالي:

$$P(A \cap B \cap C) = P(A) \times (B \setminus A) \times (C \setminus A, B)$$

مثال (١):

إذا كان الجدول التالي يمثل الاحتمال لرغبة زبون على الشراء من محل تجاري:

الرغبة \ القدرة	عنده قدرة الشراء	ليس عنده قدرة الشراء
عنده رغبة الشراء	0.3	0.1
ليس عنده رغبة الشراء	0.2	0.4

١- ما احتمال أن يكون لدى هذا الزبون رغبة الشراء؟

$$P(A) = 0.3 + 0.1 = 0.4$$

٢- ما احتمال أن يكون قادراً على الشراء إذا كان يرغب في الشراء؟

$$P(B \setminus A) = \frac{P(A \cap B)}{P(A)} = \frac{0.3}{0.4} = 0.75$$

مثال (٢):

إذا فرض أن مركزاً لتحليل الأسواق المالية يعتقد أنه سوف يكون هناك ارتفاع عام في القيمة السوقية باحتمالية 60% وأنه في حال حصل ذلك فإن احتمالية أن تحقق محفظة البركة المالية أرباحاً كبيرة هي 85% ، فأوجد احتمال أن تحقق أن يحصل ارتفاع عام وأن تحقق المحفظة المذكورة أرباحاً كبيرة.

الحل :

نفرض أن :

- ارتفاع عام في القيمة السوقية = A
- تحقيق محفظة البركة المالية أرباحاً كبيرة في حال حصل ارتفاع عام = $B \setminus A$
- حصول ارتفاع عام وتحقيق المحفظة أرباحاً كبيرة = $A \cap B$

طبعاً هنا تعويض مباشرة من السؤال والـ 100 اللي هي النسبة 100%
ولا ننسى موضوع الشرط \

$$P(A \cap B) = P(A) \times P(B \setminus A)$$

$$= \frac{60}{100} \times \frac{85}{100} = \frac{5100}{10000} = 51\%$$

المحاضرة الرابعة

تابع نظرية الاحتمالات

➢ استقلال الحوادث

يقال عن حادثين A و B انهما مستقلان إذا حدوث أحدهما لا يعتمد على حدوث الآخر والمعنى الرياضي لذلك هو:

$$P(A \setminus B) = P(A)$$
$$P(B \setminus A) = P(B)$$

وبالتالي فإن احتمال وقوعهما معاً مساوٍ لحاصل ضرب احتمال كل منهما ونكتب رياضياً :

$$P(A \cap B) = P(B) \times P(A \setminus B)$$
$$= P(B) \times P(A)$$

وكذلك :

$$P(A \cap B) = P(A) \times P(B \setminus A)$$
$$= P(A) \times P(B)$$

مثال :

إذا كان A و B حادثين في S بحيث أن :

$$P(A) = 0.5 \qquad ،P(B) = 0.6 \qquad P(A \cup B) = 0.8$$

هل A و B حادثان مستقلان ؟

الحل :

أولاً : من المعلوم أنه لأي حادثتين A و B فإن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$0.8 \quad = 0.5 + 0.6 - P(A \cap B)$$
$$P(A \cap B) = 0.5 + 0.6 - 0.8$$
$$= 0.3$$

ثانياً: نحسب حاصل ضرب احتمالي وقوع الحادثتين :

$$P(A) \times P(B) = 0.5 \times 0.6 = 0.3$$

وبما أن العلاقة التالية تتحقق :

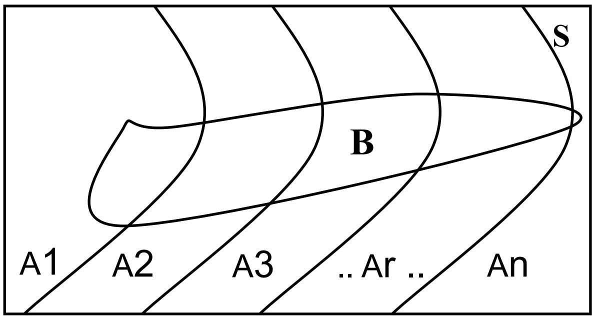
$$P(A \cap B) = P(A) \times P(B)$$

إذا الحادثان A و B مستقلان.

➢ نظريــة بايــز (Bayes' Theorem)

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

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



مثال:

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

١- أن تكون القطعة المعيبة من إنتاج الآلة الأولى؟

٢- أن تكون القطعة المعيبة من إنتاج الآلة الثانية؟

الحل:

نفرض أن :

A1= {إنتاج الآلة الأولى}	**P(A1)=0.2**
A2= {إنتاج الآلة الثانية}	**P(A2)=0.35**
A3= {إنتاج الآلة الثالثة}	**P(A3)=0.45**
B= {إنتاج سلعة معينة}	

فيكون بالتالي:

P(B|A1)=0.02

P(B|A2)=0.025

P(B|A3)=0.03

إذا أن 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:

نفرض أن

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

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

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

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

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

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

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

المحاضرة الخامسة

المتغيرات العشوائية المنفصلة

والتوزيعات الاحتمالية المنفصلة

مقدمة :

عند دراسة تجربة عشوائية قد يكون اهتماما متوجها إلى نتائج تلك التجربة بذاتها كما في الفصل السابق وقد يكون الاهتمام متوجها إلى قيم عددية مرتبطة بالنتائج وهذه القيم تسمى بالمتغير العشوائي ، فمثلاً / في تجربة إلقاء حجر نرد مرتين متتاليتين عندما يكون الحادث A هو ظهور الصورة مرة واحدة على الأقل فإن فضاء العينة والحادث A واحتماله هم على التوالي :

$$S = \{HH, HT, TH, TT\}$$
$$A = \{HH, HT, TH\}$$
$$P(A) = \frac{3}{4}$$

أما في حالة المتغير العشوائي فيكون الاهتمام متوجها على سبيل المثال إلى عدد مرات ظهور الصورة ، فتكون القيم العددية التي يأخذها المتغير العشوائي وليكن X في المثال السابق هي :

$$X = \{2,1,0\}$$

➢ المتغير العشوائي

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

المتغيرات العشوائية المتقطعة (المنفصلة):

المتغير العشوائي المنفصل هو الذي يأخذ قيم بينية و متباعدة ، 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 العشوائي المنفصل هو كل قيمة من قيم المتغير العشوائي كنتيجة لعد الأشياء.

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.
- وهكذا..... الأمثلة كثيرة.

التوزيع الاحتمالي 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 التكراري النسبي للقيم التي يمكن أن يأخذها المتغير.

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

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

شروط التوزيع الاحتمالي :

$$1)\quad 0\le P(x_i)\le 1$$

$$2)\quad \sum_{i=1}^{n} P(x_i)=1$$

مثال :

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.

المطلوب:

كون فراغ العينة.

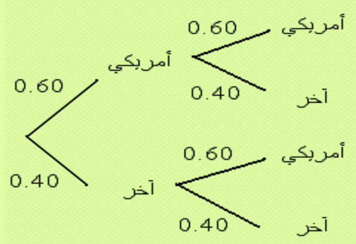
إذا عرف المتغير العشوائي بأنه عدد العبوات المشتراة من التفاح الأمريكي ، فأوجد الآتي:

- 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:



S	عدد العبوات X	$P(X=x)=f(x)$
(أمريكي،أمريكي)	2	0.36
(آخر,أمريكي)	1	0.24
(أمريكي,آخر)	1	0.24
(آخر,آخر)	0	0.16

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 ، أي 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)

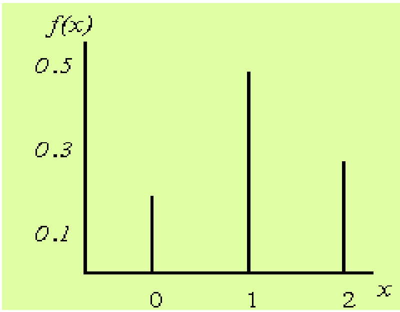
$x=2$ إذا كان العبوتين من النوع الأمريكي ، أي إذا كانت نتيجة التجربة (أمريكي، أمريكي)

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	(0.6 × 0.4) + (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:

$$\mu = \sum x_i f(x_i)$$

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

$$\sigma^2 = \sum (x_i - \mu)^2 f(x_i)$$
$$= \sum x_i^2 f(x_i) - \mu$$

مثال:

في المثال السابق احسب ما يلي:

١) الوسط الحسابي لعدد العبوات المشتراة من النوع الأمريكي.

٢) احسب الانحراف المعياري لعدد العبوات المشتراة من النوع الأمريكي.

٣) أوجد معامل الاختلاف النسبي.

الحل:

١) الوسط الحسابي لعدد العبوات من النوع الأمريكي:

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

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

إذا الوسط الحسابي هو:

$$\mu = \sum x_i \ f(x_i) = 1.20$$

٢) لحساب 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:

$$\sigma = \sqrt{\sigma^2} = \sqrt{0.48} = 0.693$$

٣) معامل الاختلاف النسبي هو:

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

❖ وكما أوضحنا أن المتغير العشوائي المنفصل (مقارنة بالمتصل) هو المتغير الذى يمكن أن يأخذ قيما محدودة ومتميزة ، وتسمى مجموعة كل القيم الممكنة للمتغير العشوائي واحتمالاتها المناظرة بالتوزيع الاحتمالي ، ويكون مجموع الاحتمالات = 1

الحل بالألة الحاسبة: نوجد الوسط الحسابي ثم الانحراف المعياري ثم التباين للمثال السابق (بيانات مبوبة) نتبع التالي ابتداء من اليمين:

(shift) ثم (Mode) ثم (سهم تحت) ثم (4: STAT) ثم (1:ON) ثم (shift) ثم (1) ثم (2: Data) ثم ندخل أرقام x_i كالتالي ابتداء من الرقم 0 في الجدول (0=1=2=) ثم (سهم يمين) ثم (سهم تحت) ثم ندخل أرقام $f(x_i)$ كالتالي ابتداء من الرقم 0.16 (0.16=0.48=0.72=)

ثم (AC) ثم (shift) ثم (1) ثم (4: Var) ثم (2: $\bar{x}$) ثم = تطلع لنا النتيجة 1.2

لا زالت البيانات مخزنه في الألة نحصل على الإنحراف المعياري كالتالي:

(shift)) ثم (1) ثم (4: Var) ثم (3: σ X) ثم = تطلع لنا النتيجة 0.693

والتباين : مباشرة نقوم بتربيع الناتج من خلال x^2 ويطلع لنا الناتج 0.48

➢ التوزيعات الاحتمالية الخاصة

١- توزيع ذي الحدين:

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

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

(استجابة للدواء ، أو عدم استجابة)

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

(الوحدة إما أن تكون سليمة ، أو تكون معيبة)

- عند إلقاء قطعة عملة، لها نتيجتان :

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

- نتيجة الطالب في الاختبار:

(نجاح ، رسوب)

- استخدام المزارع لبرنامج معين في الزراعة :

(يستخدم ، أو لا يستخدم)

شكل التوزيع الاحتمالي ثنائي الحدين:

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

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

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

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

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

فبالتالي يمكن حساب الاحتمال من خلال المعادلة التالية:

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

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

حيث $n!$ (وتقرأ " مضروب n ") = $0!=1,\ n.(n-1).(n-2)\ldots3.2.1$

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.

$p>0.5$ **الاحتمال أكبر من 0.5**

مثال:

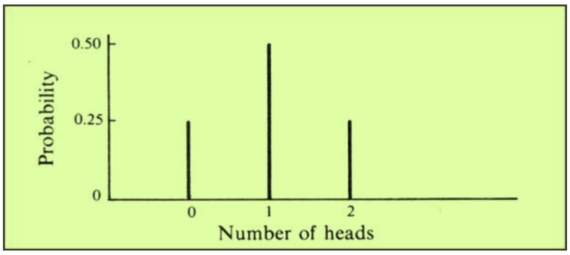
عند رمي عملة متوازنة مرتين فإن النواتج الممكنة هي TT,TH,HT,HH وإذاً :

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

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:

باستخدام توزيع ذي الحدين يمكننا إيجاد احتمال الحصول على 4 صور في 6 رميات لعملة متوازنة كالآتي:

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

$$P(4)=\frac{6!}{4!(6-4)!}(\frac{1}{2})^{4}(\frac{1}{2})^{2}=\frac{6\times5\times4\times3\times2\times1}{4\times3\times2\times1\times2\times1}(\frac{1}{16})(\frac{1}{4})=15(\frac{1}{64})=\frac{15}{64}\cong 0.23$$

لأنها قطعة متزنة فاحتمال النجاح 50% (1/2) واحتمال الفشل 50% (1/2) فعوضنا بهذا الشكل.

إن عدد الصور المتوقع في ست رميات هو:

$$\mu=np=(6)(1/2)=3$$

ويكون الانحراف المعياري للتوزيع الاحتمالي لست رميات هو:

$$\sigma=\sqrt{np(1-p)}=\sqrt{(6)(1/2)(1/2)}=\sqrt{6/4}=\sqrt{1.5}\cong1.22heads$$

٢- 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 :

$$P(x)=\frac{e^{-\mu}\mu^{x}}{x!} \quad , \quad x=0,1,2,\ldots$$

x	العدد المعين من النجاحات.
$P(x)$	احتمال عدد x من النجاحات
e	أساس نظام اللوغاريتمات الطبيعي وتوجد في بعض الآلات الحاسبة، وقيمتها هي: $e=2.718$ تقريبا، ويمكن حساب قيمتها باستخدام الآلة الحاسبة.
$x!$	مضروب العدد x ويساوي: $x!=x(x-1)(x-2)..3\times2\times1$

❖ ويكثر استخدام هذا التوزيع في الحالات التي تقع فيها الأحداث وفقا لمعدلات زمنية ، وكذلك في حالة الأحداث نادرة الوقوع، ومن أمثلة ذلك:

- عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر $X:\{x=0,1,2,...\}$
- عدد مرات ري نوع معين من المحاصيل الزراعية خلال الموسم $X:\{x=0,1,2,...\}$
- عدد العملاء الذين يتم خدمتهم البنكية لكل 10 دقائق $X:\{x=0,1,2,...\}$
- عدد مرات زيارة المريض للطبيب كل سنة $X:\{x=0,1,2,...\}$
- عدد مرات تناول الأسرة للحوم الحمراء خلال الأسبوع $X:\{x=0,1,2,...\}$
- عدد أخطاء الطباعة لكل صفحة من صفحات الكتاب $X:\{x=0,1,2,...\}$

دائما يكون كمي منفصل أي أرقام بدون كسور عكس الكمي المتصل الذي يوجد به كسور.

شكل التوزيع الاحتمالي البواسوني :

إذا كان متوسط عدد مرات وقوع حادث وفقا لمعدل زمني معين هو μ ، وكان المتغير العشوائي X يعبر عن عدد مرات وقوع الحادث وفقا لهذا المعدل ، فإن مدي المتغير العشوائي X هو: $X:\{x=0,1,2,...\}$ ، وهذا المدى عبارة عن فئة مفتوحة من اليمين ، فإن الاحتمال $P(X=x)=f(x)$ والذي يعبر عن احتمال وقوع الحادث عدد من المرات وفقا لهذا المعدل.

مثال (١):

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:

١) ما نوع المتغير العشوائي؟
٢) اكتب شكل دالة الاحتمال لهذا المتغير.
٣) احسب الاحتمالات التالية:
- **احتمال أن الأسرة تستهلك وحدتين خلال الشهر؟**
- **احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر؟**

٤) احسب الوسط الحسابي ، والانحراف المعياري لعدد الوحدات المستهلكة.
٥) حدد شكل التوزيع.

الحل:

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

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

---

٢) شكل دالة الاحتمال:

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

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

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

---

٣) حساب الاحتمالات:

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

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

عوضنا كما أتى في السؤال وعوضنا عن e بـ 2.718 كما وضحنا سابقاً.

الحل بالآلة الحاسبة: (حساب e^{-3}) ودائماً يكون الأس بالسالب

نبدأ أولاً بـ ALPHA ثم $x10^x$ ثم x فوقفها مربع أبيض ثم 3- ثم = يطلع لنا الناتج 0.0498

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

حسبنا e^{-3} حيث تساوي 0.0498 ونأخذها عامل مشترك ونقوم بحساب كلاً من p(2) , p(3) .... إلخ كما حسبنا p(2) سابقا .

$$P(X \le 3) = p(3) + p(2) + p(1) + p(0)$$

$$= \left[\frac{3^3}{3!} + \frac{3^2}{2!} + \frac{3^1}{1!} + \frac{3^0}{0!}\right]\left[\frac{0.0498}{1}\right]$$

$$= [0.0498]\left(\frac{27}{6} + \frac{9}{2} + \frac{3}{1} + \frac{1}{1}\right) = 0.0498(13) = 0.6474$$

---

٤) **حساب الوسط الحسابي ، والانحراف المعياري لعدد حالات الاستجابة:**

الوسط الحسابي (μ) في حالة التوزيع البواسون هو معلمة معطاة هي:

ذكر بالسؤال بأن المتوسط يساوي 3

$$\mu = 3$$

في هذا التوزيع ، فإن التباين يساوي الوسط الحسابي:

التباين في هذا التوزيع على وجه الخصوص يساوي المتوسط

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 في بداية هذه المحاضرة ، وهو:

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

---

٥) **تحديد شكل التوزيع:**

دائما توزيع بواسون موجب الالتواء

مثال (٢):

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

$$P(X = 2) = \frac{e^{-\mu}\mu^x}{x!}$$

$$= \frac{e^{-5}5^x}{x!} = \;, \quad x = 0,1,2,\ldots$$

$$= \frac{(0.00674)(25)}{(2)(1)} = 0.08425$$

المحاضرة السادسة

المتغيرات العشوائية المتصلة

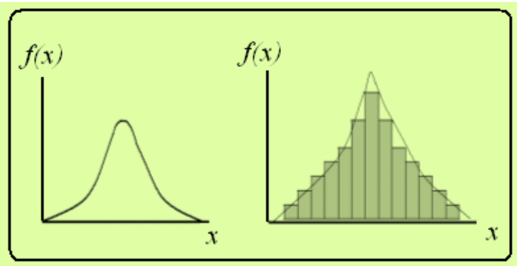
والتوزيعات الاحتمالية المتصلة

➢ **المتغيرات العشوائية المستمرة Continuous Random Variables**

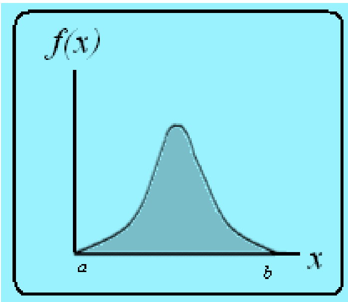
المتغير العشوائي المستمر هو الذي يأخذ قيما متصلة ، ويأخذ عدد لانهائي من القيم الممكنة له داخل مجاله ، فإذا كان متغير عشوائي مستمر ، ويقع في المدى *(a,b)*، أي أن: $\{X = x : a < x < b\}$ فإن للمتغير *X* عدد لانهائي من القيم تقع بين الحدين الأدنى والأعلى *(a,b)*، ومن الأمثلة على المتغيرات الكمية المستمرة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

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

عند تمثيل بيانات المتغير الكمي المستمر في شكل مدرج تكراري نسبي ، نجد أن شكل هذا المدرج هو أقرب وصف لمنحنى التوزيع الاحتمالي للمتغير المستمر ، وكلما ضاقت الفترات بين مراكز الفئات ، يمكن الحصول على رسم دقيق للمنحنى الخاص بدالة احتمال المتغير المستمر ، كما هو مبين بالشكل التالي:



والمساحة أسفل المنحنى تعبر عن مجموع الاحتمالات الكلية ، ولذا تساوي هذه المساحة الواحد الصحيح ، وتسمى الدالة *f(x)* بدالة كثافة الاحتمال *Probability Distribution Function(p.d.f)* ، وبفرض المتغير العشوائي المستمر يقع في المدى: $X = \{x : a \prec x \prec b\}$ وأن منحنى هذه الدالة يأخذ الصورة التالية:



الوسط الحسابي والتباين للمتغير العشوائي المستمر:

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

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

$$\sigma^2 = E(x^2) - \mu^2 \ , \ E(x^2) = \int_a^b x^2 f(x) \ dx$$

يعنى التوزيع الإحصائي : 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:

١) التوزيع الطبيعي
٢) التوزيع الطبيعي (القياسي) المعياري
٣) 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**

١) 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.

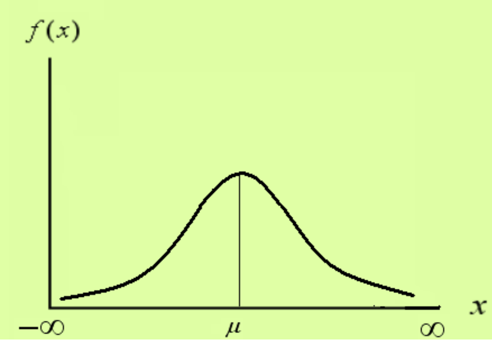
خصائص التوزيع الطبيعي:

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 لـ σ تعني أن لدينا جرس طويل مدبب ، والقيمة الكبيرة لـ 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 اعتبارهذه 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 **.**

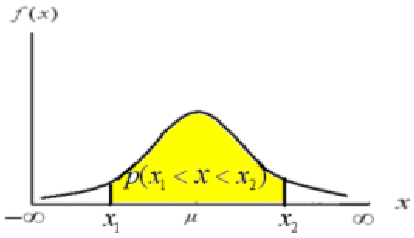
شكل دالة كثافة الاحتمال:

إذا كان لدينا توزيع طبيعي ذو وسط حسابي μ **وانحراف معياري** σ **فإن معادلة منحنى دالة كثافة الاحتمال تكون على الصورة التالية:**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad , -\infty < x < \infty \quad , \quad \pi = 22/7$$

كيفية حساب الاحتمالات:

بفرض أن الاحتمال المطلوب حسابه هو $p(x_1 < X < x_2)$ **وهذا الاحتمال يحدد بالمساحة التالية:**



وحيث أن هذا التوزيع من التوزيعات المستمرة ، فإن هذه المساحة (الاحتمال) تحسب بإيجاد التكامل التالي:

$$p(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

وهذا التكامل يصعب حسابه، ومن ثم لجأ الإحصائيين إلى عمل تحويلة رياضية Transform ، يمكن استخدام توزيعها الاحتمالي في حساب مثل هذه الاحتمالات، وهذه التحويلة هي:

$$z = \left(\frac{x-\mu}{\sigma}\right)$$

ويعرف المتغير الجديد بـ z وهو المتغير الطبيعي القياسي Standard Normal Variable ، أو المعياري.

---

٢) التوزيع الطبيعي القياسي (المعياري) :

هو توزيع طبيعي وسطه الحسابي 0 وانحرافه المعياري 1 (أي أن $\sigma = 1, \mu = 0$)

ويمكن تحويل أي توزيع طبيعي (بوحدات x) إلى توزيع طبيعي قياسي (بوحدات z) ، وتحت هذه الشروط ، فإن 68.26% من المساحة (الاحتمال) تحت المنحنى الطبيعي القياسي تقع بين إحداثيين رأسيين يبعدان بمقدار انحراف معياري واحد عن الوسط الحسابي (أي داخل $\mu \pm 1\sigma$)، 95.54% تقع بين $\mu \pm 2\sigma$ ، 99.74% تقع بين $\mu \pm 3\sigma$

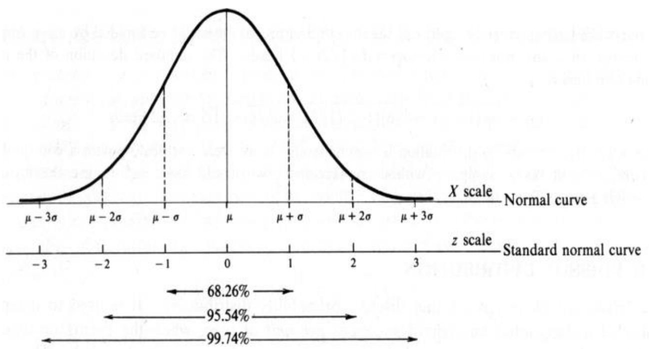
ولإيجاد الاحتمالات (المساحات) في مسائل تحتوى على التوزيع الطبيعي ، فإننا نحول أولا قيم *X* إلى قيم *Z* المناظرة لها ، من خلال المعادلة التالية:

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

ثم نكشف عن قيمة *Z* في الجداول المخصصة لذلك ، ويعطى هذا قيمة الجزء من المساحة (الاحتمال) تحت المنحنى بين قيمة الوسط الحسابي وقيمة *Z*

- 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:



العلاقة بين التوزيع الطبيعي والتوزيع الطبيعي القياسي

هنا نجد شرح سهل وواضح لاستخراج قيمة Z من الجدول الإحصائي

فالمساحة (الاحتمال) تحت المنحنى الطبيعي المعياري *Z*=0 و *Z*=1.96 نحصل عليها مقابلة للقيمة 1.96 في جدول التوزيع الطبيعي ، ففي عمود *Z* نبدأ بالقيمة 1.9 ونتحرك في الصف المناظر لها حتى نصل إلى العمود المعنون 06.، وتكون القيمة التي نحصل عليها 0.9750 .

ويعنى هذا أن 97.50% من المساحة الكلية (1 أو 100%) تحت المنحنى تقع بين *Z*=0, *Z*=1.96 المساحة المظللة في الشكل فوق الجدول) ، ولأن التوزيع متماثل ، فإن المساحة بين Z=0 , Z=-1.96 (ليس مدرجة في الجدول) هي أيضاً 0.9750 أو 97.50%

مثال (١):

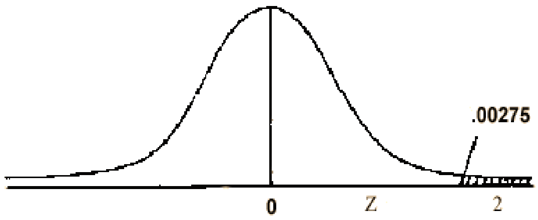
احتمال أن تكون قيمة *Z* أكبر من 2 :

> هنا حلنا كان بطريقة أخرى حل بها الدكتور بشكل أوضح
> ولابد أن نعلم من خلال الجدول الإحصائي أن:
> $P(Z \le 0)=0.5000=50\%$

الحل:

من الجدول الإحصائي نجد أن *Z* أقل من 2 = 0.9772 اذن احتمال أن تكون قيمة *Z* أكبر من 2 هي :

$1 - P(Z \le 2) = 0.0228$



مثال (٢) :

- احتمال أن تقع *Z* بين صفر و 0.5.
- احتمال أن تقع *Z* بين 0.5 و 0.5-.

> نأخذ الأول ونستخدم الجدول الإحصائي
> $P(0 \le Z \le 0.5) = P(Z \le 0.5) - P(Z \le 0)$
> $= 0.6915 - 0.5000$
> $=0.1915$

الحل:

من الرسم المساحة المظللة بين 0 و 0.5 تمثل احتمال أن تقع *Z* بين (0 و 0.5) ، والمساحة المظللة إلى شمال (0 ويمين - 0.5) هي احتمال أن تقع *Z* في الفترة (0 , 0.5-) .

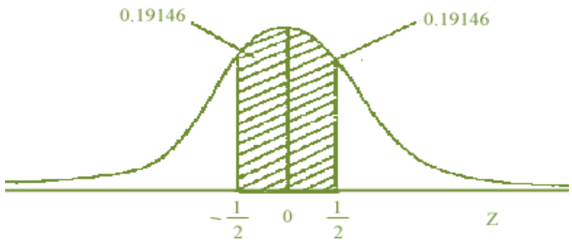
واحتمال أن تقع *Z* في الفترة (0 , 05) والمساحة المقابلة لقيمة *Z* = 0.1915

كذلك احتمال أن تقع *Z* في الفترة (0, 0.5-) = 0.1915

اذاً احتمال أن تقع *Z* في الفترة :

(0.5 و 0.5-) = 2 × 0.1915 = 0.383

وهي تتمثل بالمساحة المظللة في الرسم التالي:



مثال (٣) :

قامت إحدى الشركات بإجراء اختبار للمتقدمين لشغل بعض الوظائف الشاغرة بها ، فإذا علمت أن درجات هذا الاختبار تتبع توزيعا معتدلا وسطه الحسابي 500 وانحرافه المعياري 100 درجة وأن أحد الممتحنين قد اختير عشوائيا.

ما هو احتمال أن تكون درجة المتقدم أكبر من 700؟

الحل:

إذا كانت *X* تمثل أي درجة لأي ممتحن ، فإن *X* تتبع توزيعا معتدلا وسطه الحسابي 500 درجة وانحرافه المعياري 100 درجه ، وباستخدام المعادلة الخاصة بالدرجة المعيارية ، نجد أن القيمة المعيارية *Z* للقيمة 700 هي :

$$Z = \frac{X-\mu}{\sigma} = \frac{700-500}{100} = \frac{200}{100} = +2$$

> ولأنها موجب
> فهذا يعني أن
> $P(Z \ge 2)$

> القيمة المعيارية تمت دراستها في الإحصاء في الإدارة وسهلة فقط نقوم بالتعويض بالأرقام في المعادلة ☺

❖ وبالتالي يمكننا صياغة السؤال السابق كما يلي : إذا اختير أحد الممتحنين عشوائيا ، فما هو احتمال أن تزيد درجته عن الوسط الحسابي بأكثر من انحرافين معياريين؟

للإجابة على هذا السؤال فإننا نستخدم الشكل التالي:

ويمكن أن نحسبها بهذه الطريقة أبسط من خلال الجدول

$P(Z \geq 2) = 1 - P(Z \geq 2)$

$= 1- 0.9772 = 0.0228$

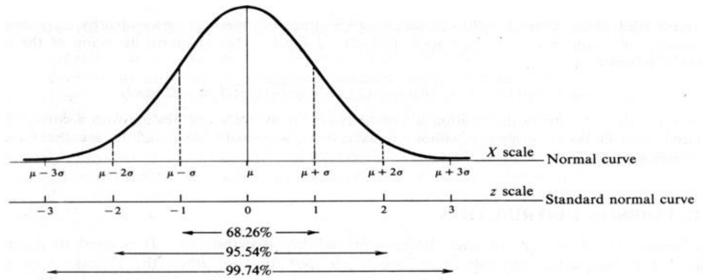


Fig. 3-4

والذي يبين أن المساحة تحت المنحنى المحصورة بين انحرافين معياريين من الوسط الحسابي 95.45% وبالتالي تكون المساحة المتبقية من المنحنى أي مساحة طرفي المنحنى هي (1-0.9545=0.0455) ، ونتيجة لتماثل المنحنى حول وسطه الحسابي ، فإن مساحة الطرف الأيمن للمنحنى تساوي مساحة طرفه الأيسر ، أي تساوي (0.0455/2=0.02275).

لذا فإن المساحة تحت المنحنى على يمين ($\mu+2\sigma$) من الوسط الحسابي (أي على يمين $+2\sigma$) تساوي 0.02275 وهي قيمة احتمال أن تكون درجة الشخص الذي اختير عشوائيا أكبر من 700

الحل بالآلة الحاسبة: (حساب التوزيع الاحتمالي) في حالة ذكر لنا احتمال أن تزيد أو أكبر من

Mode بعد ذلك 3: STAT ثم 1: 1-VAR ثم AC ثم SHIFT ثم 1 ثم 5:Distr ثم 3: R(ثم ندخل القيم كالتالي:

R((700-500)÷100)=0.02275

كيفية استخدام جدول توزيع الاحتمالات المتجمعة للمتغير العشوائي Z

وبمعرفة القيمة المعيارية Z يمكننا أن نحصل على احتمالات أي متغير عشوائي معتدل ، والتعبير Z <+2 يعني أن القيمة المشاهدة تقع على مسافة أقل من $+2\sigma$ على يمين الوسط الحسابي ، أيضا فإن التعبير -1< Z <+3 يعني أن القيمة المشاهدة تقع بين $\mu-1\sigma$ و $\mu+3\sigma$

ومن الواضح نه لا يمكن استخدام الشكل السابق لتحديد الاحتمالات المطلوبة بسهوله كافية ، لذا يستخدم جدول توزيع الاحتمالات المتجمعة للمتغير العشوائي Z لإيجاد الاحتمالات المطلوبة ، ويعطي العمود الأول بيسار الجدول مع الصف العلوي قيم Z المختلفة إلى رقمين عشريين فقط ، والرقم الأول بالعمود الأول على يسار الجدول هو 0.0 والرقم الأول بالصف العلوي من الجدول هو 0.00 ومجموع هذين الرقمين يعطينا القيمة المعيارية *Z=0.00* والاحتمال المتجمع المناظر هو *0.5000* أي أن *P(Z > 0.000)=0.5000* وهذه بطبيعة الحال نتيجة منطقية لأن توزيع Z متماثل حول وسطه الحسابي وهو الصفر ، وبالتالي لا يوجد أي احتمال متجمع بالجدول قيمته أقل من *0.5000*

في هذا المثال نجد شرح موضح بشكل أفضل لطريقة استخدام الجدول وحساب قيمة Z ودائماً جدول الـ Z يحسب المساحة على يسار المنحنى

مثال (١) :

أوجد احتمال أن Z أقل من ($<$) 1.64

الحل:

الجدول التالي جزء من جدول الاحتمالات المتجمعة للتوزيع المعتدل المعياري

Z	.00	.01	.02	.03	.04	.05
1.6					0.9495	

ويتم الحصول على القيمة المعيارية Z بجمع القيمتين المناسبتين الموجودتين بالصف العلوي والعمود الأول بيسار الجدول ، ويحوي العمود الأول من جهة اليسار على قيم تصل إلى رقم عشري واحد فقط ، بينما يحوي الصف العلوي على الرقم المئوي.

فالاحتمال المتجمع المناظر للقيمة 1.64 يوجد أمام الصف 1.6 وتحت العمود 0.04 (لاحظ أن 1.6 + 0.04 = 1.64) وهي قيمة Z المطلوب إيجاد الاحتمال المتجمع عندها ، وهذا الاحتمال هو 0.9495 ، أي أن $P(Z<1.64)=0.9495$ وهذا هو الاحتمال المتجمع للمتغير Z من ($-\infty$) إلى 1.64

والجدول التالي يوضح ذلك:

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

مثال (٢):

أوجد أن احتمال أن Z أكبر من ($>$) 1.64 .

الحل:

إن مجموع الاحتمالات المتجمعة لأي متغير عشوائي يساوي (١) ، وحيث أن المساحة الكلية تحت منحنى أي متغير عشوائي مستمر تمثل مجموع الاحتمالات ، لذا فإن هذه المساحة تساوي (١) لذا فإن :

$$P(Z>1.64) = 1 - P(Z<1.64) = 1 - 0.9495 = 0.0505$$

مثال (٣):

أوجد المساحة تحت المنحنى المعتدل المعياري على يمين $Z=-1.65$.

الحل:

المنحنى المعتدل كما أوضحنا منحنى متماثل حول الصفر ، وبالتالي فإن المساحة تحت المنحنى على يمين 1.65- تساوي المساحة تحت المنحنى على يسار 1.65 ، أي أن $(P(Z>-1.65) = P(Z<1.65))$

وباستخدام جدول التوزيع الطبيعي نجد أن $P(Z<1.65) = 0.9505$ أي أن الاحتمال المتجمع من 1.65- إلى $+\infty$

أي أن:

$$P(Z<1.65) = P(Z>-1.65) = 0.9505$$

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:

١) نسبة السيارات التي تقل سرعتها عن 50 ميلا في الساعة .
٢) نسبة السيارات التي تزيد سرعتها عن 65 ميلا في الساعة.
٣) نسبة السيارات التي تكون سرعتها بين 60 ميلا و 77.45 ميلا في الساعة .
٤) عدد السيارات التي تكون سرعتها بين 60 ميلا و 77.45 ميلا من بين 10000 سيارة .

الحل:

هذا المثال واضح والأهم أن نعلم اننا نريد الانحراف المعياري لذلك نأخذ جذر التباين المعطى لما في السؤال 25

١) نسبة السيارات التي تقل سرعتها عن 50 ميلا في الساعة :

$$P(X<50) = P(\frac{x-\mu}{\sigma} < \frac{50-60}{\sqrt{25}}) = P(Z<-2) = 1-0.9772 = 0.0228$$

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

$$P(X>65) = P(\frac{x-\mu}{\sigma} > \frac{65-60}{\sqrt{25}}) = P(Z>1) = 1-P(Z\le 1) = 1-0.8413 = 0.1587$$

٣) نسبة السيارات التي تكون سرعتها بين 60 ميلا و 77.45 في الساعة :

$$P(60 \le X \le 77.45) = P(\frac{60-60}{\sqrt{25}} \le Z \le \frac{77.45-60}{\sqrt{25}})$$
$$= P(0 \le Z \le 3.49) = P(Z \le 3.49) - P(Z \le 0)$$
$$= 0.9998 - 0.5000 = 0.4998$$

٤) عدد السيارات المتوقع سرعتها بين 60 ميلا و 77.45 ميلا من بين 10000 سيارة :

$$10000 \times (0.4998) = 4998$$

❖ أشار الدكتور لمن أحب الاطلاع على أمثلة أكثر عليه العودة للكتاب صفحة 150 إلى 155

٣) توزيع t ستيودنت :

توجد عائلة أخرى من المتغيرات العشوائية المتصلة المستخدمة في الإحصاء الاستدلالي وهي مجموعة المتغيرات العشوائية t ويعتبر وليم جوست w.s. Gosset هو أول من درس تلك المتغيرات حيث سجل نتائجه عام ١٩٠٨ تحت اسم مستعارهو student ولذلك يسمى توزيع t في بعض الأحيان بتوزيع ستيودنت.

ويرمز لهذه العائلة من التوزيعات بالرموز(t1,t2,t3 ......tdf) كما يرمز لدرجات حريتها بالرمز V حرف إغريقي ينطق نيو) وهي تأخذ القيم (1, 2, 3, ... , df)

الفرق بين توزيع t والتوزيع الطبيعي:

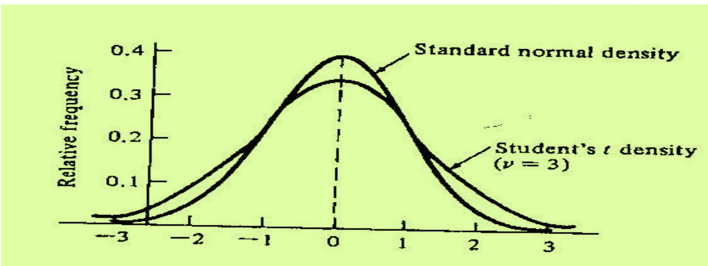
يختلف المتغير العشوائي t عن المتغير العشوائي الاعتدالي ، حيث يتحدد المتغير العشوائي الاعتدالي بمعلمين هما الانحراف المعياري والمتوسط ، بينما يتحدد المتغير العشوائي t بمعلم واحد فقط هو درجة الحرية.

ولاشتقاق المتغير العشوائي t من المتغير العشوائي (الطبيعي) الاعتدالي ، فإن ذلك يتطلب معرفة قيمة المتوسط μ للمتغير العشوائي الاعتدالي ، بينما لا نحتاج إلي معرفة انحرافه المعياري.

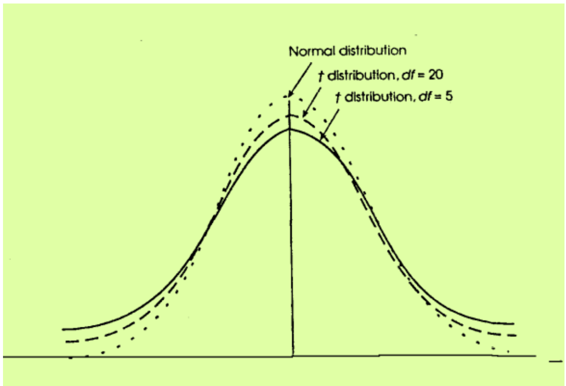
ولنفرض أن قيمة متغير العشوائي الاعتدالي التي تم ملاحظتها n من المرات $(n>=z)$وأن هذه الملاحظات البالغ عددها n تكون عينة متوسطها μ وانحرافها المعياري S وحسبنا قيم المتغير العشوائي t باستخدام الصيغة التالية :

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

وتتحدد درجة حرية المتغير العشوائي t بأنها تساوي (n-1) كذلك فإنه لكل قيم n نجد أن توزيع t له قيمة واحدة عند النقطة صفر ، وهو توزيع متماثل يقل تدريجيا كلما اتجهنا ناحيتي الذيلين الأيمن والأيسر ، وهذا ما يوضحه الشكل التالي :



ونلاحظ من الشكل السابق ان توزيع t يشبه توزيع z فيما عدا أنه أكثر انتشارا diffuse لأنه أكثر كثافة عند الذيلين وخاصة عندما تكون n صغيرة ، اما إذا كانت n كبيرة فإن توزيع t يكون أقل انتشارا وأكثر قربا من شكل توزيع Z ، وبزيادة درجات الحرية يقترب توزيع t من التوزيع الاعتدالي ، وهذا ما يوضحه الشكل:



خصائص توزيع t :

- متوسط المتغير العشوائي t يساوي صفر لكل درجات الحرية (n-1) ، وهذا يعني أن $\mu = 0$
- الانحراف المعياري للمتغير العشوائي t لدرجات حرية أكبر من اثنين يساوي :

$$\sigma = \frac{s}{S-2}$$

حيث df هي درجة حرية المتغير العشوائي t .

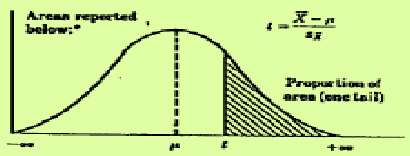
ويتبين من المعادلة السابقة للانحراف المعياري أنه كلما زادت درجات حرية المتغير العشوائي t بحيث تصل إلي 30 فأكثر ، فإن الانحراف المعياري يقترب من الواحد الصحيح ، وبصفة عامة فإن الانحراف المعياري لتوزيع t يساوي 1.035 أو أقل.

ولذلك فإن التوزيع الاحتمالي للمتغير يكون قريبا جدا من التوزيع الاحتمالي للمتغير العشوائي Z وبصفة خاصة عندما تكون df>30 وفي هذه الحالة نستخدم جدول Z للإجابة على الأسئلة الاحتمالية حول المتغير العشوائية t .

الجدول أدناه يعطي قيمة t ∝

المقابلة للمساحة المظللة وقيمتها ∝

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
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 :

احسب القيمة الحرجة (نقطة القطع) بتوزيع t لدرجات حرية 8 ومستوى الدلالة 10. (الاحتمال بالذيل الأيمن)

الحل :

بالبحث في الجدول توزيع t عند درجات 8 والعمود الخاص بمستوى الدلالة 10. نجد أن القيمة عند تقاطع الصف و العمود تساوي 1.397

$$P(t_8 \geq 1.397) = .10$$

نستخدم الجدول الإحصائي لإيجاد قيمة t

$$P(-1.397 \leq t_8 \geq 1.397) = .80$$

وهنا تكون الـ t محصورة ما بين 10% و 10% ويتبقى لدينا 80%

❖ أشار الدكتور لمن أحب الاطلاع على أمثلة أكثر عليه العودة للكتاب صفحة 164 إلى 165

المحاضرة السابعة

توزيعات المعاينة

الجزء الأول

مقدمة :

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**) سليما ، ينبغي أن يبنى على **عينة ممثلة للمجتمع** ، ويمكن تحقيق ذلك **بالمعاينة العشوائية**، حيث يكون لكل مفردة في المجتمع فرصة متكافئة للدخول في العينة .

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 إعطاء 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 يتطلب وفرة في الوقت والمال والمجهود الفني وتزداد هذه المتطلبات وتتضاعف كلما ازداد حجم المجتمع (عدد أفراد المجتمع) ، وهذا الأسلوب لا يتبع عادة إلا في حالة التعدادات التي تجريها الدول وتدعمها بإمكانيات ضخمة 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 ، أي أن أسلوب العينة يقصد به دراسة خصائص المجتمع من خلال دراسة عينه مسحوبة منه ، ونجاح هذا الأسلوب يعتمد على أن تحمل العينة أقصى درجة من دقة التمثيل للمجتمع المسحوبة منه.

بعض مزايا أسلوب المعاينة:

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا 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:

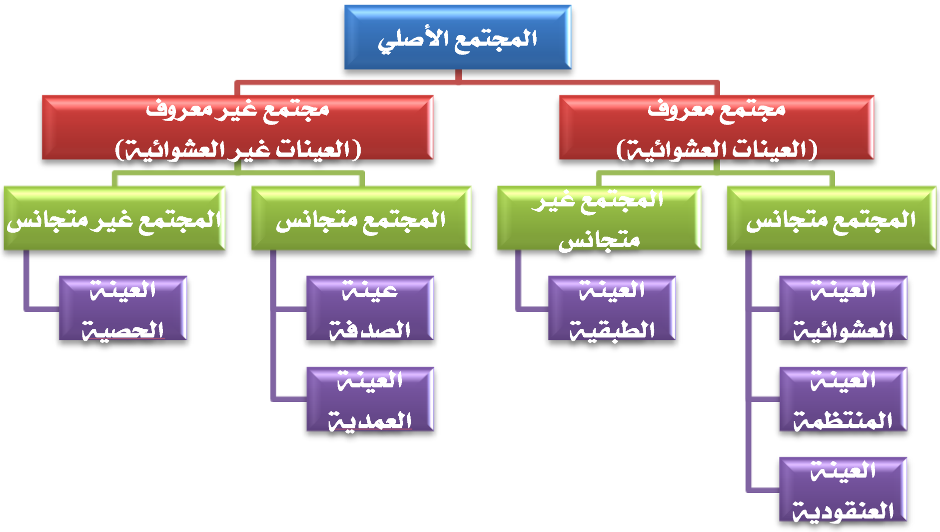
وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار ، والعينات العشوائية إذا ما 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:



١) 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 منها
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.

٢) 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.

كيفية التقليل من أخطاء التميز أوالتحيز:

- اختيار جميع وحدات العينة عشوائياً باستخدام إحدى طرق الاختيار العشوائي.
- عدم استبدال أية وحدة تم اختيارها بوحدة أخرى.
- تدريب الباحثين بشكل جيد على جمع البيانات والتقيد بالتعليمات.

---

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 تحصل لها فرصة أن تدخل في العينة وهذا الخطأ يسمى بخطأ المعاينة العشوائي أو خطأ الصدفة.

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).

المحاضرة الثامنة

توزيعات المعاينة

الجزء الثاني

مقدمة

اعتاد البعض على معاملة القيم التي يحصل عليها من العينة وكأنها قيم مجتمعها ، وهذه المعاملة غير دقيقة ، فلكي يستدل على خصائص مجتمع الدراسة من خلال العينة ، لا بد من الأخذ في عين الاعتبار:

1) **ما هو المقياس الذي يود الباحث أن يستدل عليه من خلال العينة.**
2) **الحجم المناسب للعينة.**
3) **خصائص كل من المجتمع والعينة.**

- **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 $\overline{X}$ ، أيضاً للانحراف المعياري للمجتمع بالرمز σ بينما يرمز للانحراف المعياري للعينة بالرمز S وهكذا.

- **وتعتبر كل إحصاءه بمثابة تقدير أو قيمة تقديرية لمعلمة المجتمع المناظرة ، فيكون المتوسط الحسابي المحسوب من بيانات العينة تقدير لمعلمة المجتمع المناظرة وهي المتوسط الحسابي المسحوب منه هذه العينة وهكذا.**
- **إن حساب قيمة المتوسط الحسابي من بيانات العينة ليس هدفاً في حد ذاته ولكن وسيلة للتعرف على المتوسط الحسابي للمجتمع موضوع الدراسة ، وهكذا الحال بالنسبة لباقي المقاييس الإحصائية التي تحسب من العينة.**

إن الهدف من اخذ العينة هو معرفة خصائص مجتمعها ، فاخذ العينات ليس القصد منه العينة لذاتها بل المجتمع الذي أخذت منه ، فالعينة وسيلة وليست الهدف.

وتقدم العينات تقديرات لخصائص مجتمعها ، وهذه التقديرات تدور حول المتوسط الحقيقي لمجتمع الدراسة ، أي أن متوسط العينة هو ليس متوسط مجتمعها، بل قيمة تمثل العينة ذاتها ، وتعتمد في تقدير القيمة المحتملة لمتوسط المجتمع وفق حدود معينة للثقة.

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.

توزيع المعاينة:

وهو ذلك التوزيع التكراري لأحد التوابع الإحصائية المحسوب من بيانات العينات العشوائية ذات الحجم الواحد والتي يمكن سحبها من مجتمع إحصائي واحد.

نفرض أننا أخذنا عينه حجمها n من مجتمع ما ، ثم حسبنا بعض المقاييس الإحصائية مثل المتوسط الحسابي ، التباين ، ... فإن كل مقياس من هذه المقاييس يعتبر متغير عشوائي في ذاته يختلف من عينه إلى أخرى – هذا المتغير العشوائي يخضع لتوزيع معين – هذا التوزيع يسمى بتوزيع العينة.

فمثلاً: نقول إن توزيع المعاينة للمتوسط الحسابي وهو عبارة عن توزيع جميع المتوسطات الحسابية للعينات المأخوذة من نفس هذا المجتمع ذو الحجم n ، وكذلك فإن توزيع المعاينة للتباين هو توزيع جميع التباينات المحسوبة من عينات لها نفس الحجم n ومأخوذة من نفس المجتمع ، وهكذا ...

إذا أخذنا عينات متكررة من مجتمع ما وقمنا بحساب متوسط كل عينة ، فإننا نجد أن معظم هذه المتوسطات تختلف عن بعضها البعض، ويسمى التوزيع الاحتمالي لمتوسطات العينات "**توزيع المعاينة للوسط** "

توزيعات المعاينة:

نظرية (1):

إذا كان X يخضع لتوزيع وسطه μ وتباينه σ^2 ، وكان $\overline{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع فإن :

$$E(\overline{X}) = \mu_{\overline{X}} = \mu$$

$$Var(\overline{X}) = \sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$$

المقصود أن وسط العينة مساوي لوسط المجتمع بينما التباين لمتوسط العينات مساوي لتباين المجتمع مقسوم على n

مثال:

إذا كان لدينا مجتمع إحصائي مؤلف من العاملين في منشأة صغيرة حجمه $N=4$ مفردات ، ومكون من القيم {0,2,4,6}

في هذا المثال نحاول إثبات مصداقية هذه النظرية ، هنا استخرنا الوسط الحسابي ثم التباين بقوانينها المعروفة لدينا سابقاً

$$\mu = \frac{0+2+4+6}{4} = 3$$

$$\sigma^2 = \frac{(0-3)^2+(2-3)^2+(4-3)^2+(6-3)^2}{4} = 5$$

نفرض أننا سحبنا جميع العينات الممكنة مع الإعادة ذات الحجم $n=2$ ثم حسبنا متوسطاتها.

هنا نأخذ عينات ثم نقوم بحساب متوسطاتها

عدد العينات الممكن سحبها مع الإعادة يعطى بالعلاقة:

$N^n = 4^2 = 16$

وان متوسطات العينات العشوائية المسحوبة تتأرجح بين (0 ، 6) أنظر الجدول التالي:

رقم العينة	العينة		المتوسط	رقم العينة	العينة		المتوسط
1	0	0	0	9	4	0	2
2	0	2	1	10	4	2	3
3	0	4	2	11	4	4	4
4	0	6	3	12	4	6	5
5	2	0	1	13	6	0	3
6	2	2	2	14	6	2	4
7	2	4	3	15	6	4	5
8	2	6	4	16	6	6	6

وان الجدول الاحتمالي لتوزيع معاينة الأوساط الحسابية $\overline{X}_i$ (لاحظ من الجدول السابق المتوسط 1 تكرر مرتين وهكذا)

المتوسط	0	1	2	3	4	5	6
(P)	1/16	2/16	3/16	4/16	3/16	2/16	1/16

ولو رسمنا المدرج التكراري ، نلاحظ أن توزيع المعاينة للأوساط الحسابية للعينات يمكن أن يقترب وبشكل جيد من منحنى التوزيع الطبيعي والذي يرمز لمتوسطه بـ μ_x وتشتته بـ :

$$\mu_{\bar{x}} = \sum \overline{X}_i P = 0.\left(1/16\right) + 1.\left(2/16\right) + 2.\left(3/16\right) + 3.\left(4/16\right)$$

$$+4.\left(3/16\right) + 5.\left(2/16\right) + 6.\left(1/16\right) = 3$$

من الجدول السابق هنا نضرب المتوسط في الكسر من الصف (P) ويعطينا الناتج وهذا ما يثبت لنا صحة التوقع السابق.

وهي نفس قيمة μ

إذاً:

$$E(\overline{X}) = \mu_{\overline{X}} = \mu$$

نظرية (2):

إذا كانت $X_1, X_2, X_3, \ldots \ldots X_n$ عينة عشوائية من مجتمع طبيعي وسطة μ وتباينه σ^2 ، وكان $\overline{X}$ يمثل الوسط الحسابي للعينة ذات الحجم *n* والمسحوبة من هذا المجتمع فإن $\overline{X}$ يتبع التوزيع الطبيعي بوسط μ وانحراف معياري $\sigma / \sqrt{n}$ أي أن :

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

❖ يخضع لتوزيع طبيعي معياري.

مثال:

أخذت عينة عشوائية حجمها 9 من الأطفال حديثي الولادة في أحد المستشفيات ، فإذا علم أن وزن الطفل حديث الولادة يخضع لتوزيع طبيعي 2,900 غرام والانحراف المعياري 600 غرام.

١) أوجد معدل وتباين والانحراف المعياري للوسط الحسابي لأوزان الأطفال في العينة ؟

$$X \sim N(2900, (600)^2)$$

$$\overline{X} \sim N\left(2900, \frac{(600)^2}{n}\right)$$

$$\mu_{\overline{X}} = \mu \quad , \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

٢) أوجد احتمال أن الوسط الحسابي لأوزان الأطفال في العينة يزيد عن 3100 غرام ؟

$$P(\overline{X} > 3100) = P\left(\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} > \frac{3100 - \mu}{\sigma / \sqrt{n}}\right)$$

$$= P\left(Z > \frac{3100 - 2900}{600 / \sqrt{9}}\right)$$

$$= P\left(Z > \frac{200}{200}\right)$$

$$= P(Z > 1)$$

$$= 1 - P(Z < 1) = 1 - 0.8413 = 0.1587$$

هنا ما لون بالأحمر عوضنا عنه بطريقة رياضية.

الكسر الأول $\frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$ عوضنا عنه بـ Z كما في الصيغة في نظرية 2

الجدول الطبيعي يحسب المساحات عندما تكون أصغر من فنحن نحصل على متممتها وهي $1 - P(Z < 1)$

نحصل على قيمتها من الجدول Z حيث أنها تكون عند 1.00 وتطلع 0.8413

٣) أوجد احتمال أن الوسط الحسابي لأوزان الأطفال في العينة يقع بين 2700 و 3200 غرام ؟

$$P(2700<\overline{X}<3200)=P\left(\frac{2700-\mu}{\sigma/\sqrt{n}}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<\frac{3200-\mu}{\sigma/\sqrt{n}}\right)$$

$$=P\left(\frac{2700-2900}{600/\sqrt{9}}<Z<\frac{3200-2900}{600/\sqrt{9}}\right)$$

نفس الفكرة السابقة ولكن هنا حصرنها ما بين 2700 و 3200 كما هو مطلوب.

$$=P\left(\frac{-200}{200}<Z<\frac{300}{200}\right)$$

$$=P(-1<Z<1.5)$$

$$=P(Z<1.5)+P(Z<1)-1$$

$$=0.9332+0.8413-1=0.7745$$

هنا نستخرج القيم من الجدول Z حيث عند 1.00 سبق طلعناها وعند 1.50 ولأنه يوجد اجزاء من ميه فنكون مباشرة عند 00. تطلع 0.9332

تمرين : تخضع علامات الطلاب في أحد المقررات لتوزيع طبيعي وسطه 65 وانحرافه المعياري 18 ، أخذت عينة عشوائية حجمها 36 ، احسب احتمال أن يزيد وسط علامات العينة على 74 ؟

الحل:

$$X\sim N(65,(18)^2)\ ,\ \ \overline{X}\sim N(65,\frac{(600)^2}{2})\ ,\ \ n=36$$

$$P(\overline{X}>74)=P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}>\frac{74-\mu}{\sigma/\sqrt{n}}\right)$$

$$=P\left(Z>\frac{74-65}{18/\sqrt{36}}\right)$$

نفس ما سوينا في الخطوة رقم 2 في المثال السابق

$$=P\left(Z>\frac{9}{3}\right)$$

$$=P(Z>3)$$

$$=1-P(Z<3)=1-0.9987=0.0013$$

نظرية (3): النهاية المركزية (تقارب التوزيعات).

إذا كانت $X_1, X_2, X_3, \dots\dots X_n$ عينة عشوائية من مجتمع طبيعي وسطه μ وتباينه σ^2 ، وكان $\overline{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع فإن $\overline{X}$ يقترب من التوزيع الطبيعي بوسط μ وانحراف معياري $\sigma/\sqrt{n}$ كلما كبرت n $(n\geq 30)$ أي أن :

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

❖ يخضع لتوزيع طبيعي معياري.

مثال:

تخضع أوزان علب سائل غسيل الصحون من نوع معين لتوزيع معدله 1,000 غرام ، وانحرافه المعياري 80 غرام ، إذا أخذت عينة عشوائية حجمها 48 علبة.

يختلف عن النظرية واحد في أنه لم ينص على كلمة توزيع طبيعي

١) أوجد معدل وتباين والانحراف المعياري للوسط الحسابي لأوزان الأطفال في العينة ؟

حجم العينة كبير $(n = 48 \geq 30)$ ، المعدل والانحراف المعياري للمجتمع معلومة ، ولذلك فشروط نظرية (3) متحققة أي أن :

$$\bar{X} \sim N\left(1000, \frac{(80)^2}{48}\right)$$

$$\mu_{\bar{X}} = \mu = 1000 \quad , \quad \sigma^2{}_{\bar{X}} = \frac{\sigma^2}{n} = \frac{(80)^2}{48} = \frac{6400}{48} \approx 133.33$$

$$\sigma_{\bar{X}} = \sqrt{\sigma^2{}_{\bar{X}}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} = \frac{80}{\sqrt{48}} \approx \sqrt{133.33} = 11.55$$

٢) ما احتمال أن يزيد الوسط الحسابي لأوزان العلب في العينة عن 1072 غرام ؟

نفس طريقة حلنا السابق.

$$P(\bar{X} > 1072) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{1072-1000}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z > \frac{1072-1000}{80/\sqrt{48}}\right)$$
$$= P\left(Z > \frac{72}{11.55}\right)$$
$$= P(Z > 6.23) = 1 - P(Z < 6.23) \approx 1 - 1 = 0$$

٣) ما احتمال أن يقل الوسط الحسابي عن 980 غرام ؟

هنا لأنها سالبة نأخذ متممتها وهي نفس متمم العدد الطبيعي.

$$P(\bar{X} < 980) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{980-\mu}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z < \frac{980-1000}{80/\sqrt{48}}\right)$$
$$\approx P(Z < -1.73) = 1 - P(Z < 1.73) = 1 - 0.9582 = 0.0418$$

تمرين: تخضع أوزان عبوات أحد مبيدات الحشرات المنزلية لتوزيع وسطه 135 غرام وانحرافه المعياري 14 غرام ، إذا قررت وزارة التموين رفض كل صندوق من هذه العبوات إذا نقص وزنه عن 6.24 كغم، فما نسبة الصناديق المرفوضة ، علما بأن عدد العبوات في كل صندوق 48 عبوة ؟

$$P(Y < 6240) = P\left(\frac{\sum X_1}{48} < \frac{6240}{48}\right) = P(X < 130)$$
$$= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{130-\mu}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z < \frac{130-135}{14/\sqrt{48}}\right)$$
$$\approx P(Z < -2.47) = 1 - P(Z < 2.47) = 1 - 0.9932 = 0.0068$$

نظرية (4):

إذا كانت $X_1, X_2, X_3, \dots \dots X_n$ عينة عشوائية من مجتمع طبيعي وسطة μ معلوم وتباينه σ^2 غير معلوم ، وكان $\overline{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع وكان S الانحراف المعياري لهذه العينة فإن:

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

❖ يخضع لتوزيع **t** بدرجة حرية $n-1$ ، ويكتب : $T \sim t_{n-1}$

مثال:

إذا كانت درجات طلاب التحليل الإحصائي تأخذ شكل التوزيع الطبيعي بوسط حسابي مقداره **70 درجة** ، وأخذت عينة حجمها **9 طلاب** ، ووجد أن الانحراف المعياري لعلاماتهم **11 درجات** ، احسب احتمال أن يزيد وسط درجاتهم عن **75 درجة** ؟

المتغير العشوائي X يتبع توزيع طبيعي بوسط 70 وتباين مجهول ونكتب ذلك : $X \sim N(70, \sigma^2)$

ثم سحب عينة حجمها 9 ، وانحراف معياري هذه العينة معلوم وهو 11 ، ولذلك فإن : $T \sim t_8$

حيث :

$$T = \frac{\overline{X} - 70}{11 / \sqrt{9}}$$

المطلوب : احسب احتمال أن يزيد وسط درجاتهم عن **75 درجة** ؟

$$\begin{aligned} P(\overline{X} > 75) &= P\left(\frac{\overline{X} - \mu}{S / \sqrt{n}} > \frac{75 - \mu}{S / \sqrt{n}}\right) \\ &= P\left(T > \frac{75 - 70}{11 / \sqrt{9}}\right) \qquad , T \sim t_8 \\ &= P\left(T > \frac{5}{11/3}\right) \\ &= P\left(T > \frac{15}{11}\right) \\ &\approx P(T > 1.363) \approx P(T > 1.397) \approx 10\% \end{aligned}$$

عند درجة حرية 8 من الجدول t نقرب 1.363 إلى 1.397 وتطلع 10%

تمرين: إذا كانت ساعات المذاكرة الأسبوعية للطلاب الجامعيين في إحدى الدول تأخذ شكل التوزيع الطبيعي بوسط حسابي مقداره **20 ساعة** ، أخذت عينة حجمها **25 طالباً** ، ووجد ان الانحراف المعياري لعدد ساعات مذاكرتهم الاسبوعية **8 ساعات** ، احسب احتمال أن يقل وسط عدد ساعات مذاكرتهم الاسبوعية عن **18 ساعة** ؟

يمكن حل هذا المثال بنفس الطريقة السابقة

أشار الأستاذ في المحتوى إلى أن هناك توزيعات معاينة أخرى ، مثل توزيع المعاينة للنسبة ، وتوزيع المعاينة لتباين العينة ، وتوزيع المعاينة للفرق بين وسطين ، وتوزيع المعاينة للنسبة بين تباين عينتين.

يمكن الرجوع إليها في الكتاب المقرر(ص 188 – 194).

المحاضرة التاسعة

التقدير (الجزء الأول)

التقدير

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

ويكون الاهتمام بتقدير معالم مجتمع ما على أساس التوزيع الإحصائي الذي يتبع له ذلك المجتمع ، فقد يكون الاهتمام متوجها نحو تقدير الوسط الحسابي للمجتمع عندما يكون المجتمع يتبع توزيعا طبيعيا مثلا ، ويكون الاهتمام متوجها نحو تقدير النسبة عندما يكون المجتمع يتبع توزيع ذي الحدين على سبيل المثال وهكذا.

ويجب أن يراعى عند التقدير ألا يكون التقدير متحيزا.

وهناك نوعان (أو أسلوبان) للتقدير:

الأول: تقدير النقطة (أو القيمة الواحدة).

التقدير بنقطة هو عبارة عن عدد واحد ، ويكون هذا التقدير بنقطة غير متحيز إذا كانت القيمة المتوقعة للإحصاءه المناظرة عند تكرار المعاينة العشوائية مساوية لمعلمة المجتمع.

فمثلا/ $\overline{X}$ هي تقدير (بنقطة) غير متحيز للمعلمة μ ، أما الانحراف المعياري S للعينة فهو تقدير غير متحيز للمعلمة σ ، والنسبة في العينة p هي تقدير غير متحيز للمعلمة P (وهي نسبة المفردات التي لها خاصية معينة في المجتمع كله).

فالتقدير بنقطة يعني بالتالي 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.

الثاني: تقدير الفترة (أو فترة التقدير أو الثقة).

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.

وتتميز تقديرات الفترة بالإضافة إلى أنها (١) تحتوي على عدد كبير جداً من القيم ، بأنه (٢) يمكن حساب احتمال أن يكون التقدير صحيحا ، وبالتالي فإنه يمكن معرفة مدى دقة التقديرات ، لذا فإن فترات التقدير تسمى أيضاً " فترات الثقة Confidence intervals " لأن هذه الفترات تعتمد في تكوينها الإحصائي على درجات أو مستويات ثقة

معينة Confidence Levels مثل 95% أو 99 % وغيرها ، بمعنى أن احتمال أن تكون فترة التقدير صحيحة هو 0.95 أو 0.99 وهكذا...

مثلاً: لو كان متوسط أعمار الناخبين يتراوح ما بين 46 و 34 سنة ، ودرجة الثقة هي 95% فإن هذا معناه أنه لو تكررت التجربة مائة مرة ، فإن التقدير سيكون محصوراً بين هذين الرقمين في 95 مرة من الحالات (أي احتمال أن يكون صحيحاً هو 95%).

التقدير الإحصائي لفترة الثقة

أولاً: التقدير الإحصائي لفترة الثقة للوسط الحسابي عند الظروف التالية:

- عينة إحصائية مأخوذة من توزيع طبيعي.
- σ^2 (تباينه) معلومة.

ذكرنا في المحاضرة السابقة (في توزيعات المعاينة) النظرية التالية:

إذا كانت $X_1, X_2, X_3, \dots \dots X_n$ عينة عشوائية من مجتمع طبيعي وسطه μ وتباينه σ^2 ، وكان $\bar{X}$ يمثل الوسط الحسابي للعينة ذات الحجم *n* والمسحوبة من هذا المجتمع فإن $\bar{X}$ يتبع التوزيع الطبيعي بوسط μ وانحراف معياري $\sigma / \sqrt{n}$ أي أن :

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

- يخضع لتوزيع طبيعي معياري.

ومن جدول التوزيع الطبيعي المعياري نجد أن :

$$P(-z < Z < z) = 95\%$$

$$P\left(-z < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < z\right) = 95\%$$

$$P\left(-z \times \sigma / \sqrt{n} < \bar{X} - \mu < z \times \sigma / \sqrt{n}\right) = 95\%$$

$$P\left(\bar{X} - z \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z \times \frac{\sigma}{\sqrt{n}}\right) = 95\%$$

ومعنى المقدار الأخير هو أن القيمة التقديرية للوسط الحسابي للمجتمع تتراوح بين هذين المقدارين الأعلى منهما والأدنى ، وأن مستوى ثقتنا بهذا التقدير تساوي 95% ولذا فإننا نكتب $\hat{\mu}$ (أي القيمة المقدرة لوسط المجتمع) بدلاً من μ

$$\hat{\mu} = \bar{X} \pm z \times \sigma / \sqrt{n}$$

$$\left(\bar{X} - z \times \frac{\sigma}{\sqrt{n}}, \bar{X} + z \times \frac{\sigma}{\sqrt{n}}\right)$$

حيث أن :

μ	الوسط الحسابي للمجتمع.
$\bar{X}$	الوسط الحسابي للعينة.
σ	الانحراف المعياري للمجتمع.
n	حجم العينة
z	معامل الثقة المناظر لمستوى (درجة) الثقة $\left(z_{1-\frac{\alpha}{2}}\right)$

حسب معامل الثقة

إذا أردنا حساب معامل الثقة المناظر لمستوى الثقة المراد حساب فترة الثقة عنده ، فيتم ذلك كما يلي:

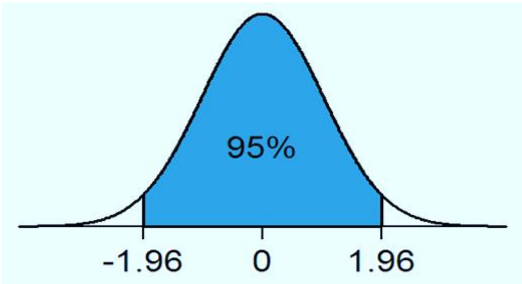
إذا كان مستوى الثقة يساوي 95% $(1-\alpha=95\%)$

فإن مستوى عدم الثقة (وهو ما يسمى بمستوى المعنوية) يساوي 5% $(\alpha=5\%)$

وبالتالي فإن $\frac{\alpha}{2}$ تساوي 2.5% $(\frac{\alpha}{2}=2.5\%)$

أي أن : $(1-\frac{\alpha}{2}=1-2.5\%)=97.5\%$

فنبحث في جدول **Z** عن النقطة التي تكون عندها قيمة الاحتمال مساوي للقيمة **0.9750** هذه القيمة هي **1.96** $(z_{1-\frac{\alpha}{2}}=1.96)$ ويمكن ملاحظة ذلك من خلال الرسم التالي:



وباختصار ينبغي حفظ معاملات الثقة لمستويات الثقة الأكثر استعمالا وهي كالآتي/

عند مستوى ثقة **90%** تكون قيمة معامل الثقة **1.64** $(z_{1-\frac{\alpha}{2}}=1.64)$

عند مستوى ثقة **95%** تكون قيمة معامل الثقة **1.96** $(z_{1-\frac{\alpha}{2}}=1.96)$

عند مستوى ثقة **99%** تكون قيمة معامل الثقة **2.58** $(z_{1-\frac{\alpha}{2}}=2.58)$

هذي تحفظ مثل اسمك لأن الأسئلة في هذا الخصوص غالباً ما تكون عليها فمثلاً عند مستوى ثقة **90%** تكون قيمة معامل الثقة **1.64**

ونلخص ما سبق بإيراد النظرية التالية:

نظرية (1)

إذا أخذت عينة عشوائية حجمها **n** من مجتمع طبيعي $N(\mu,\sigma^2)$ وكانت σ^2 معلومة فإن فترة ثقة % $(1-\alpha)$ للمعلمة μ هي :

$$(\bar{X}-z_{1-\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}},\ \bar{X}+z_{1-\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}})$$

نطبق النظرية واحد عندما يكون مجتمع طبيعي والتباين معلوم

مثال:

أخذت عينة عشوائية حجمها **n = 16** من مجتمع طبيعي $N(\mu,9)$ فوجد أن $\bar{X}=11.3$ أوجد فترة ثقة **95%** للمعلمة المجهولة μ ؟

فترة الثقة 95% وهي مقابلة لمعامل الثقة 1.96 كما وضحنا سابقاً ونعوض تعويض عادي.

المجتمع طبيعي وتباينه معلوم وقيمة وسطة الحسابي للعينة $\bar{X}=11.3$

إذا فترة ثقة 95% هي:

$$(\bar{X}-z\times\frac{\sigma}{\sqrt{n}},\bar{X}+z\times\frac{\sigma}{\sqrt{n}})=(11.3-1.96\times\frac{3}{4}\ \ ,\ \ 11.3+1.96\times\frac{3}{4})$$

$$=(11.3-1.47\ \ ,\ \ 11.3+1.47)$$

$$=(9.83\ \ ,\ \ 12.77)$$

التباين تسعة إذا الانحراف المعياري جذره ويساوي ثلاثة.

تمرين: عينة عشوائية حجمها **n = 25** من مجتمع طبيعي $\sigma = 4$ إذا كان معدل العينة $\overline{X} = 60$ أوجد فترة ثقة **99%** لوسط المجتمع μ ؟

$$(\overline{X} - z \times \frac{\sigma}{\sqrt{n}}, \overline{X} + z \times \frac{\sigma}{\sqrt{n}}) = (60 - 2.58 \times \frac{4}{5} \;,\; 60 + 2.58 \times \frac{4}{5})$$
$$= (60 - 2.064 \;,\; 60 + 2.064)$$
$$= (57.936 \;,\; 62.064)$$

نفس المثال السابق وتذكر مره تكون بالسالب ومره بالموجب كما المثال السابق.

إذاً الوسط الحسابي للمجتمع له قيمة معينة مقدره بين $(57.936 \;,\; 62.064)$ ونحن واثقين من هذا التقدير بنسبة 99% ،

وكذلك الأمر على المثال السابق ونسبة 1% بأنها ليس بينها.

ثانياً: التقدير الإحصائي لفترة الثقة للوسط الحسابي عند الظروف التالية:

- حجم العينة كبير ($n \geq 30$).
- σ^2 (تباينة) معلومة.

نظرية (2)

إذا أخذت عينة عشوائية حجمها **n** من مجتمع وسطة μ وتباينه σ^2 معلومة فإن فترة ثقة % $(1-\alpha)$ للمعلمة μ هي تقريباً :

$$(\overline{X} - z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}})$$

إذا كانت **n** كبيرة ($n \geq 30$).

مثال:

أخذت عينة عشوائية حجمها **n = 100** من مجتمع تباينه $\sigma^2 = 25$ وجد أن $\overline{X} = 52$

١) أوجد فترة ثقة **90%** للمعلمة المجهولة μ ؟

الحل:

التباين معلوم وحجم العينة كبير وقيمة الوسط الحسابي للعينة $\overline{X} = 52$ **فنطبق النظرية (2)** لتقدير فترة الثقة 90% لوسط المجتمع المجهول μ كالتالي:

$$(\overline{X} - z \times \frac{\sigma}{\sqrt{n}}, \overline{X} + z \times \frac{\sigma}{\sqrt{n}}) = (52 - 1.64 \times \frac{5}{10} \;,\; 52 + 1.64 \times \frac{5}{10})$$
$$= (52 - 0.82 \;,\; 52 + 0.82)$$
$$= (51.18 \;,\; 52.82)$$

تحل بنفس طريقة النظرية واحد.

٢) أوجد فترة ثقة **99%** للمعلمة المجهولة μ ؟

$$(\overline{X} - z \times \frac{\sigma}{\sqrt{n}}, \overline{X} + z \times \frac{\sigma}{\sqrt{n}}) = (52 - 2.58 \times \frac{5}{10} \;,\; 52 + 2.58 \times \frac{5}{10})$$
$$= (52 - 1.29 \;,\; 52 + 1.29)$$
$$= (51.71 \;,\; 53.29)$$

ثالثاً: التقدير الإحصائي لفترة الثقة للوسط الحسابي عند الظروف التالية:

- عينة احصائية مأخوذة من توزيع طبيعي.
- حجم العينة صغير ($n < 30$).
- σ^2 (تباينة) مجهولة.

هنا اجتمعت عندنا سلبيتين صغيرة وتباين مجهول لذلك نستخدم توزيع t والانحراف المعياري S

نظرية (3)

إذا أخذت عينة عشوائية حجمها n من مجتمع طبيعي μ ، فإن فترة ثقة % $(1-\alpha)$ للمعلمة μ هي تقريباً :

$$(\bar{X} - t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}})$$

عند درجة حرية ($n - 1$).

مثال:

أخذت عينة عشوائية حجمها $n = 15$ من مجتمع طبيعي فوجد أن $\bar{X} = 17.4$ ، $S = 2.1$ ، أوجد فترة ثقة 95% للمعلمة المجهولة μ ؟

$$(\bar{X} - t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}})$$
$$= (17.4 - 2.145 \times \frac{2.1}{\sqrt{15}} \quad , \quad 17.4 + 2.145 \times \frac{2.1}{\sqrt{15}})$$
$$= (17.4 - 1.16 \quad , \quad 17.4 + 1.16)$$
$$= (16.2 \quad , \quad 18.56)$$

تعويض عادي ولكن نستخرج قيمة فترة الثقة من الجدول عند 95% إما أقول 5% الباقي أقسمه على 2 يبقى 0.025 وابحث في الجدول عند درجة حرية 15- 1=14 ومعامل الثقة المناظر له هو 0.025 وتلتقي عند 2.145

أو في أسفل الجدول محدد نسب مباشرة احدد درجة الحرية 14 وأذهب إلى النسبة لأرى أين تلتقي

تمرين: أخذت عينة عشوائية حجمها $n = 15$ من مجتمع طبيعي ، وجد أن $\bar{X} = 17.4$ ، $S = 2.1$ أوجد فترة ثقة 90% للمعلمة المجهولة μ ؟

$$(\bar{X} - t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}})$$
$$= (17.4 - 1.761 \times \frac{2.1}{\sqrt{15}} \quad , \quad 17.4 + 1.761 \times \frac{2.1}{\sqrt{15}})$$
$$= (17.4 - 0.95 \quad , \quad 17.4 + 0.95)$$
$$= (16.45 \quad , \quad 18.35)$$

ممكن أن تأتي بهذا الشكل زائد وناقص مرة واحدة وكذلك الأمر على الأمثلة السابقة.

$$(\bar{X} \pm t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}})$$
$$= (17.4 \pm 1.761 \times \frac{2.1}{\sqrt{15}}$$
$$= (17.4 \pm 0.95$$
$$= (16.45 \quad , \quad 18.35)$$

ملاحظة من الأستاذ في آخر المحاضرة:

هناك فترات ثقة لمعالم أخرى ، مثل فترة الثقة للنسبة ، وفترة الثقة للتباين ، وفترة الثقة للفرق بين وسطين ، وفترة الثقة للنسبة بين تباين عينتين ، وسوف نكتفي بذكر فترة الثقة للنسبة، وطريقة تحديد حجم العينة المناسب عند تقدير الفترات، بالإضافة إلى إيراد نبذة مختصرة عن تقدير الفترات باستخدام برنامج SPSS، وهذا سوف يكون موضوع المحاضرة العاشرة بمشيئة الله.

يمكن الرجوع لهذه الموضوعات في الكتاب المقرر(ص ٢١٩ – ٢٣٢).

كما يمكن الاطلاع على (أسئلة موضوعية ٩) والتي سوف يتم تنزيلها في مجلد المحاضرة التاسعة في نظام (بلاك بورد)

المحاضرة العاشرة

التقدير (الجزء الثاني)

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 والاجتماعية ...الخ ، ويختلف تحديد حجم العينة باختلاف الهدف من التقدير.

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

$$\hat{\mu} = \bar{X} \pm z \times \sigma / \sqrt{n}$$

ومنها نجد أن حجم العينة يأخذ الشكل التالي:

$$n = \frac{z^2 \times \sigma^2}{e^2}$$

حيث أن :

e	أقصى خطأ مسموح به في تقدير الوسط.
σ^2	تباين المجتمع.
n	حجم العينة.
z	معامل الثقة المناظر لمستوى (درجة) الثقة $\left(z_{1-\frac{\alpha}{2}} \right)$

مثال:

إذا كانت دخول الأفراد اليومية في إحدى دول العالم النامية تتبع التوزيع الطبيعي بانحراف معياري $\sigma = 15$ دولاراً ، فما هو حجم العينة المناسب لتقدير متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ اليومي 5 دولارات ، وذلك بدرجة ثقة 99% ؟

من تشوف كلمة خطأ في السؤال نطبق على هذه المعادلة وتعويض عادي.

العطيات : $e = 5$ ، $\sigma = 15$ ، $(1-\alpha)\% = 99\%$

المطلوب : تحديد حجم العينة المناسب لتقدير الوسط الحسابي للمجتمع (n) ؟

الحل:

$$n = \frac{z^2 \times \sigma^2}{e^2} = \frac{(2.58)^2 \times (15)^2}{(5)^2} = \frac{6.6564 \times 225}{25} = 59.9076 \approx 60$$

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

تمرين: يرغب أحد مدراء إحدى المصانع في تقدير متوسط عد الدقائق التي يأخذها العمال لإنجاز عملية صناعية معينة بحيث لا يتعدى الخطأ في تقدير متوسط الأداء ± 3 دقائق ، وبدرجة ثقة 90% ، ويعلم المدير من خبرته الماضية أن الانحراف المعياري يساوي 15 دقيقة ، ولكنه يرد بداية أن يحدد حجم العينة (n) التي يختارها الإجراء هذا التقدير.

العطيات : $e = 3$ ، $\sigma = 15$ ، $(1-\alpha)\% = 90\%$

المطلوب : تحديد حجم العينة المناسب لتقدير الوسط الحسابي للمجتمع (n) ؟

نفس المثال السابق بتغير الأرقام ودرجة الثقة.

$$n = \frac{z^2 \times \sigma^2}{e^2} = \frac{(1.64)^2 \times (15)^2}{(3)^2} = \frac{2.6896 \times 225}{9} = 67.24 \approx 67$$

فترة تقدير النسبة للمجتمع (فترة الثقة للنسبة):

إن تقدير النسبة في المجتمع تعتبر من الحالات المهمة لقياس الظواهر الإنسانية المختلفة ، 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.

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

١) حساب النسبة في العينة $\hat{P}$

٢) حساب الخطأ المعياري للنسبة والتي تساوي في هذه الحالة:

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

٣) ضرب الخطأ المعياري للنسبة في معامل الثقة المناسب Z (حسب درجة الثقة المطلوبة) والتي نحصل عليها من جدول التوزيع الطبيعي المعياري (أو من الجدول الذي يحوي أهم درجات ومعاملات الثقة والذي ذكرناه انفا) أي نحسب:

$$Z \times \sigma_{\hat{p}} = Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

٤) للحصول على الحد الأدنى لتقدير النسبة نطرح حاصل الضرب (السابق) من نسبة العينة $\hat{P}$ ، وللحصول على الحد الأعلى نجمع حاصل الضرب مع النسبة في العينة $\hat{P}$ فنحصل على فترة تقدير النسبة ، وبالتالي فإن فترة تقدير النسبة تكون في شكلها النهائي كما يلي:

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

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%

هنا حسبنا النسبة في بعض الأسئلة تجي جاهزة محسوبة ☺

المعطيات :

☜ حجم العينة $n = 144$ ، عدد المؤيدين في العينة $\hat{p} = \frac{60}{144} \approx 0.42$

☜ درجة الثقة $(1-\alpha)\% = 95\%$ مما يعني أن معامل الثقة المناظر لهذه الدرجة هو (1.96)

المطلوب : تقدير نسبة المؤيدين لهذا المرشح في المدينة (P) ؟

الحل:

من يقول تقدير نسبة نستخدم هذه الصيغة.

$$P = \hat{p} \pm \left(Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right) = 0.42 \pm \left(1.96 \times \sqrt{\frac{0.42 \times 0.58}{144}} \right)$$

$$= 0.42 \pm (1.96 \times 0.0411)$$

$$\approx 0.42 \pm (0.08)$$

وبالتالي تكون تقدير نسبة المؤيدين لهذا المرشح في المدينة (P) بدرجة ثقة 95% هي: [0.34 , 0.50]

تابع المثال السابق/ 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 فرصته في الفوز كمرشح قد لا تكون كبيرة وذلك بدرجة ثقة 95% بمعنى أن هذا الحكم لا تتجاوز نسبة الخطأ فيه 5%.

تمرين: لإيجاد فترة ثقة 90% لنسبة المدخنين بين طلبة إحدى الجامعات قام باحث بمقابلة عينة عشوائية من 100 طالب ، فوجد أن 30 طالباً يدخنون ، أوجد فترة الثقة المطلوبة ؟

العطيات :

- حجم العينة $n = 100$ ، عدد المؤيدين في العينة $\hat{p} = \dfrac{30}{100} \approx 0.30$
- درجة الثقة $(1-\alpha)\% = 90\%$ مما يعني أن معامل الثقة المناظر لهذه الدرجة هو (1.64)

المطلوب : تقدير نسبة المدخنين في هذه الجامعة (*P*) ؟

الحل:

$$P = \hat{p} \pm \left(Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right) = 0.30 \pm \left(1.64 \times \sqrt{\frac{0.30 \times 0.70}{100}} \right)$$

$$= 0.30 \pm (1.64 \times 0.046)$$

$$\approx 0.30 \pm (0.08)$$

وبالتالي تكون تقدير نسبة المدخنين في الجامعة (*P*) بدرجة ثقة 90% هي: [0.22 , 0.38]

هذه المحاضرة طويلة نوعاً ما لذلك
أنصح بمتابعة شرحها مع الدكتور
لأنها تحتاج لتركيز وفهم جيد ☹

المحاضرة الحادية عشر

اختبار الفروض الإحصائية

تعتبر اختبارات الفروض الإحصائية واحدة من أهم التطبيقات التي قدمها علم الإحصاء كحل للمشاكل العلمية المختلفة بشتى فروع العلم ورغم أهمية موضوع تقدير المعالم إلا انه غالبا ما يكون الاهتمام مركز ليس علي مجرد تقدير المعالم ولكن علي عملية وضع قواعد تمكن من التوصل إلي قرار بقبول أو رفض خاصية أو بالمعني الإحصائي فرض عن معالم مجتمع واحد أو أكثر وهذا ما يسمي اختبارات الفروض الإحصائية ، ومن خلالها يمكن لأي شخص أن يتخذ قرار برفض أو قبول فرض معين أو مجموعة من الفروض المتعلقة بمشكلة معينة موجودة في الحياة العامة.

- وبصفه عامة فإن قبول أو رفض أي قرار يجب أن يمر بعدة مراحل وهي:

١) سحب عينة من المجتمع بحيث تكون ممثله للمجتمع (عشوائية).
٢) تجميع البيانات المتعلقة بالمشكلة من العينة.
٣) تطبيق قواعد معينة لاختبار الفروض الموضوعة عن طريق الباحث وهي مشكله تحتاج لخبرة إحصائية.
٤) اتخاذ القرار بناء على ما توصل إليه الباحث من نتائج.

اختبارات الفروض الاحصائية Testing Statistical Hypotheses:

من المعروف ان اتخاذ أي قرار لا يتم الا من خلال اختبارات الفروض الاحصائية التي تعتمد بدورها كما سبق على الاحتمالات وتوزيعات المعاينة وهذا يؤكد اهمية الدور الذى تلعبه نظريه الاحتمالات في التنبؤ والتخطيط واتخاذ القرارات بالإضافة الى اهميتها في تقدير معالم المجتمع المجهولة والتي تعتبر احد اهتمامات الباحثين.

تبدأ مشكله التعرف على معالم المجتمع المجهولة بما يسمى بالاستدلال الاحصائي (Statistical Inferences)

حيث ينقسم الى فرعين /

الفرع الأول/ يهتم بتقدير (Estimation) معالم المجتمع.

الفرع الأخر / يختص بإجراء اختبارات فروض (Testing Hypothesis) تدور حول معالم المجتمع المجهولة.

الاستدلال الإحصائي يتم باستخدام عينة عشوائية مسحوبه من المجتمع وذلك لاستحالة التعامل مع المجتمع ككل ، فالإحصاءات التحليلية قدمت القوانين التي سهلت هذه العملية وجعلتها تتم بأقل الاخطاء الممكنة.

اختبارات الفروض ترتكز اساسا على فكرة ان هناك عينه اخرى غير مسحوبه من المجتمع المسحوب منه العينات فان الفرق بين الوسط الحسابي المحسوب من هذه العينة وبين المعلمة المجهولة قد يكون فرقا معنويا Significant غير راجع للصدفة ، واشتقت اسمها منها حيث عن طريقها نستطيع ان نحدد وبسهوله هل الفروق بين المعلومات المحسوبة من العينة وبين المعلومات المفروضة لمجتمع معين فرقا يرجع الى الصدفة ام فرق حقيقي ، وبأسلوب أخر هل هو فرق معنوي او فرق غير معنوي ؟ وبذلك سميت هذه الاختبارات باسم الاختبارات المعنوي Test of Significant

مفهوم الاختبارات الإحصائية المعلمية واللامعلمية

الاختبارات الإحصائية قد تدور حول معالم المجتمع المجهولة **مثل الفروض المتعلقة بالوسط الحسابي ، النسبة ، التباين ، معامل الارتباط ، ...** وفي هذه الحالة يطلق على هذه الاختبارات اسم **الاختبارات المعلميه Parametric Tests**

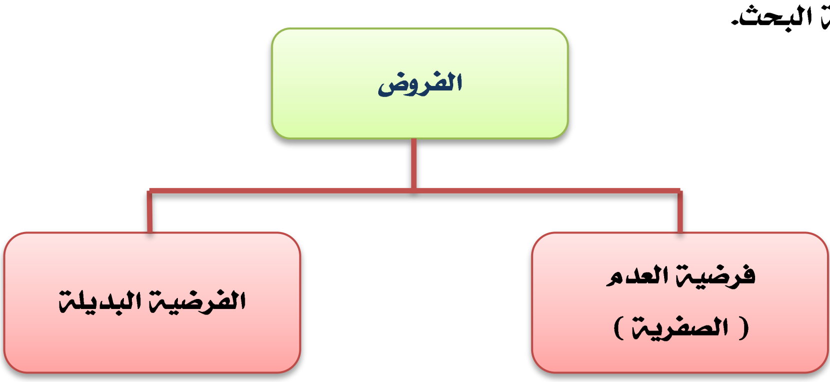
وقد تكون فروضا لا تتعلق بمعالم المجتمع ولكن تتعلق بأشياء اخرى قد تكون وصفية **مثل العلاقة بين التعليم والتدخين ، خضوع نتائج معينه لنظرية معينه ، العلاقة بين لون العينين ولون الشعر ، ...** وفي هذه الحالة يسمى الاختبار باسم **الاختبار اللامعلمي Non Parametric Test**

فالإحصاء المعلمي: أساليب إحصائية تتطلب افتراضات **محددة** عن طبيعة التوزيعات الاحتمالية للمجتمع.

أما الإحصاء اللامعلمي: أساليب إحصائية تتطلب افتراضات **أقل** عن طبيعة التوزيعات الاحتمالية للمجتمع.

الفروض الإحصائية

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



☜ **الفرضية الصفرية (فرضية العدم) H_0 (Null Hypothesis) :**

هي الفرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما ، وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها ، وخلاف ذلك نقبلها وتعني كلمة Nul انه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).

☜ **الفرضية البديلة (Ha) Alternative Hypothesis :**

هي الفرضية التي يضعها الباحث كبديل عن فرضية العدم و نقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

مستوى الدلالة الإحصائية ومنطقة الرفض:

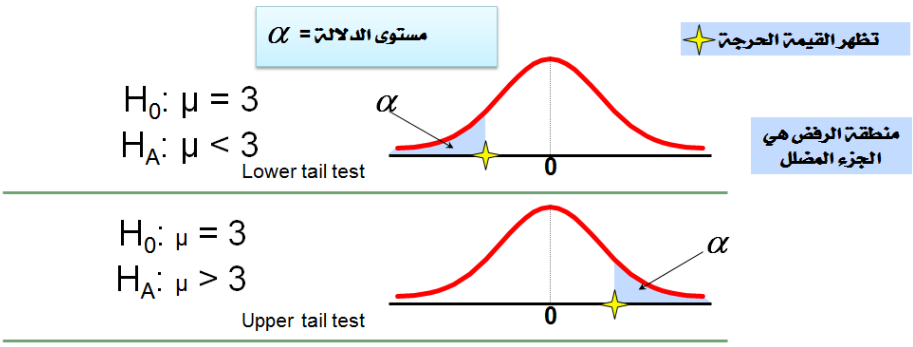
عندما نقبل الفرضية الصفرية (فرضية العدم) فإننا نقبلها بنسبة دقة **90% أو 95% أو 99%** أو غير ذلك وتسمى مستويات الدلالة أو الثقة Significance Levels أي يوجد نسبة خطأ معين في قبولنا للفرضية المبدئية بمعنى أننا نقبل صحة الفرضية الصفرية وهي خاطئة وهذا الخطأ هو α ويسمى مستوى المعنوية .

أي إذا كان مستوى الثقة 95% $(1-\alpha)$ فإن مستوى المعنوية α تساوي 95% وهي عبارة عن مساحة المنطقة التي تقع تحت منحنى التوزيع والتي تمثل منطقة الرفض ، وتكون أما على صورة ذيل واحد جهة اليمين أو اليسار أو ذيلين متساويين في المساحة واحد جهة اليمين والثاني جهة اليسار.

- **اختبار الفرضيات من طرف واحد:**

هو الاختبار الذي تبين فيه الفرضية البديلة بأن معلمة المجتمع اكبر أو اصغر من معلمة المجتمع المفترضة.

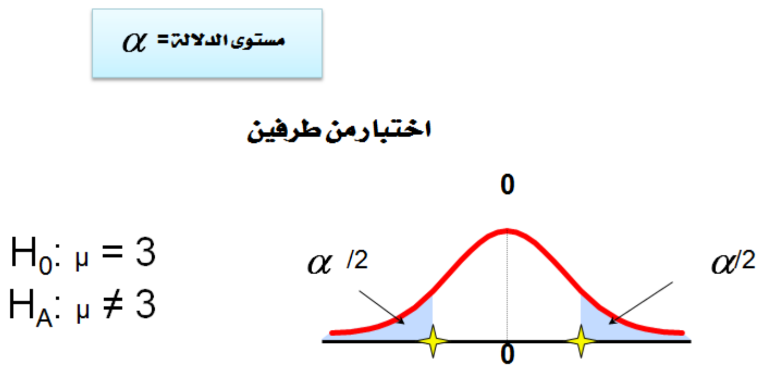
مستوى الدلالة الإحصائية ومنطقة الرفض من طرف واحد



- **اختبار الفرضيات من طرفين:**

هو الاختبار الذي لا تبين فيه الفرضية البديلة بأن معلمة المجتمع اكبر أو اصغر من معلمة المجتمع المفترضة بل مجرد أنها تختلف.

مستوى الدلالة الإحصائية ومنطقة الرفض من طرفين



اختبار الفروض عن خصائص المجتمع

إن اختبار الفروض عن خصائص المجتمع (مثل μ و σ) هو جانب أساسي أخر من جوانب الاستدلال والتحليل الإحصائي ، وفي اختبار الفروض نبدأ بعمل فرض ما عن خاصية المجتمع غير المعلومة ، ثم نأخذ عينة عشوائية من المجتمع ، وعلى أساس الخاصية المناظرة في العينة ، أما أن نقبل وإما أن نرفض الفرض بدرجة ثقة محددة.

وفي اختبار الفروض يمكن أن نرتكب نوعين من الخطأ:

الخطأ من النوع الأول Type I error: 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". ويرمز له بالرمز α .

الخطأ من النوع الثاني Type II error: 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 ". ويرمز له بالرمز β .

ويمكن أن نمثل ذلك في الجدول التالي:

الفرضية / القرار	صحيحة (H_0)	خاطئة (Ha)
قبول (HO)	صواب	خطأ ٢ بيتا (B)
رفض (HO)	خطأ ١ ألفا (α)	صواب

١) فرضية صحيحة نتائج العينة تؤيد صحتها. (قبول صواب)

٢) فرضية صحيحة نتائج العينة غير مؤيدة لصحتها. (رفض خطأ) وهذا يعطينا خطأ من النوع الأول ألفا (α) ويمكن أن يقلل برفع مستوى الدلالة

٣) فرضية خاطئة نتائج تؤيد صحتها (قبول خطأ) وهذا يعطينا خطأ من النوع الثاني بيتا (B) ويمكن أن يقلل بزيادة حجم العينة.

٤) فرضية خاطئة نتائج غير مؤيدة صحتها (رفض صواب)

مستوى المعنوية Level of Significance

يعتبر مصطلح "مستوى المعنوية" واحدا من أهم المصطلحات المستخدمة في دراسة نظرية اختبارات الفروض ، والمقصود بمستوى المعنوية α 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 α وأشهر قيمتين لمستوى المعنوية α 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.

ويمكننا ضبط أو تحديد احتمال ارتكاب خطأ من النوع الأول α. ولكن إذا خفضنا α فسوف نضطر إلى قبول احتمال أكبر لارتكاب خطأ من النوع الثاني β ، اللهم إلا إذا رفعنا حجم العينة.

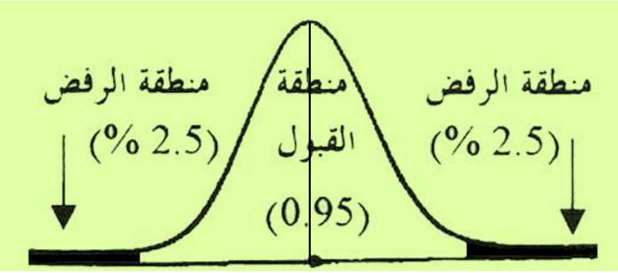
وتمسى α مستوى المعنوية ، و α- 1 مستوى الثقة للاختبار.

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.

هناك ثلاث حالات مختلفة لمنطقتي القبول والرفض (كما أوضحنا ذلك من قبل عند الحديث عن أنواع اختبارات الفروض) وهي :

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

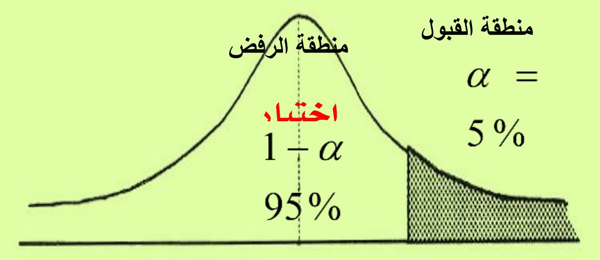


فالفرض الصفري هنا يعني أن متوسط دخل الفرد يساوي 200 دولار شهريا ، والفرض البديل في هذه الحالة هو بمعنى أن متوسط دخل الفرد لا يساوي 200 دولار شهرياً.

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%.

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

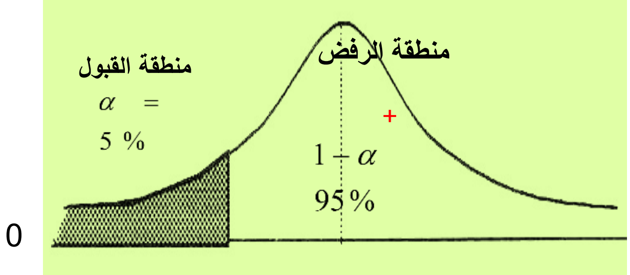
الثانية: إذا كان الفرض البديل يأخذ شكل "أكبر من" فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيمن للمنحنى ، ويسمى الاختبار في هذه الحالة اختبار الطرف الأيمن ، والذي يأخذ الشكل التالي:



منطقة+ القبول

فالفرض الصفري هنا نفس فرض المثال السابق ، بينما الفرض البديل هو بمعنى أن متوسط دخل الفرد أكبر من 200 دولاراً شهرياً ، وبالتالي فإن مستوى المعنوية والذي يساوي مثلاً 5% 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 الصفري كما في المثال السابق ، بينما الفرض البديل هو بمعنى أن متوسط دخل الفرد أقل من 200 دولار شهرياً ، وبالتالي فإن مستوى المعنوية والذي يساوي مثلاً 5 % مركز في الطرف الأيسر من المنحنى.

خطوات إجراء الاختبار الإحصائي

الاختبار الإحصائي قد يكون متعلقا بعينة واحدة او عينتين او اكثر وقد يكون اختباراً معلمياً او غير معلمياً ويجب أن يمر الاختبار أياً كان نوعه بعدة خطوات وهذه الخطوات يمكن إيجازها في التالي:

١) **صياغة الفرض الصفري H_0 :**

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:

$$Ho: \mu = 20$$

٢) **صياغة 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:

$$H1: \mu < 20 \qquad H1: \mu > 20 \qquad H1: \mu \neq 20$$

والذي يحدد شكل الفرض البديل هو مدى اقتناع الباحث بذلك أو مدى توفر المعلومات الأولية.

فمثلاً: إذا كانت وجهة نظر الباحث أن متوسط عمر الناخب **لا يمكن أن يقل عن 20 سنة** فإنه يختار الفرض البديل "**أكبر من**" والعكس صحيح إذا كان يعتقد أن متوسط عمر الناخب **لا يزيد عن 20 سنة** فإنه يختار الفرض البديل "**أقل من**" أما إذا لم يكن **لديه أي تصور أو أي معلومات** فإنه يختار الفرض البديل "**لا يساوي**"

٣) **إحصائية الاختبار:**

وهي الإحصائية التي يتم حسابها من بيانات العينة بافتراض أن الفرض العدمي صحيح.

ويتوقف شكل الإحصائية على العوامل التالية:

أ- توزيع المجتمع ، وهل هو طبيعي أم لا ، وهل تباينه معروف أم لا.

ب- حجم العينة ، وهل هو كبير أم صغير.

ج- الفرض العدمي المراد اختباره وهل هو عن الوسط أو النسبة أو التباين أو الارتباط... الخ.

والفكرة الأساسية (غالباً) في إحصائية الاختبار هي: حساب الفرق بين قيمة المعلمة التي نفترضها للمجتمع (في الفرض العدمي) والقيمة المقابلة لها في العينة أي التابع الإحصائي ، ثم نقسم (أو ننسب) هذا الفرق إلى الخطأ المعياري للتابع الإحصائي.

فمثلاً: إذا كان الاختبار عن الوسط الحسابي فإنه يتم حساب الفرق بين قيمة الوسط الحسابي للمجتمع التي نفترضها وقيمة الوسط الحسابي للعينة ، ثم نقسم هذا الفرق على الخطأ المعياري للوسط ، وهكذا مع باقي الإحصائيات.

أ) بافتراض أن المجتمع الإحصائي المسحوبة منه العينة هو مجتمع طبيعي وانحرافه المعياري معروف ، (أو) أن العينة كبيرة بدرجة كافية فإن إحصائية الاختبار والتي نرمز لها بالرمز $Z_{\bar{X}}$

الإحصائية في حالة اختبار الوسط للعينات الكبيرة:

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

❖ لاحظ أن البسط هو الفرق بين متوسطي المجتمع والعينة ، والمقام هو الخطأ المعياري للوسط ، ومن الناحية العملية فإن الانحراف المعياري للمجتمع عـادة ما يكون غير معروف ولكن طالما أن العينة كبيرة بدرجة كافية فإنه يمكن استخدام الانحراف المعياري للعينة S بدلا من الانحراف المعياري للمجتمع σ .

ب) أما في حالة العينات الصغيرة وذلك عندما يكون المجتمع طبيعياً وانحرافه المعياري غير معروف فإن الإحصائية تأخذ الشكل التالي :

الإحصائية في حالة اختبار الوسط للعينات الصغيرة :

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

والتي لها توزيع t بدرجات حرية $n - 1$

٤) تحديد منطقتي القبول والرفض: وذلك بناءً على الجداول الإحصائية والتي تعتمد على:

أ- توزيع المعاينة (وهل هو طبيعي أو t أو...)

ب- والفرض البديل (وهل هو لا يساوي أو أكبر من أو أقل من أي هل يستخدم اختبار الطرفين أو الطرف الأيمن أو الأيسر).

ج- ومستوى المعنوية (وهل هو 1% أو 5% أو غير ذلك).

٥) المقارنة والقرار:

بمعنى أن نقارن قيمة الإحصائية (المحسوبة من الخطوة الثالثة) بحدود منطقتي القبول والرفض (والتي حددناها في الخطوة الرابعة) ، فإذا وقعت قيمة الإحصائية داخل منطقة القبول فإن القرار هو: قبول الفرض الصفري. أما إذا وقعت قيمة الإحصائية في منطقة الرفض فإن القرار هو: رفض الفرض الصفري ، وفي هذه الحالة نقبل الفرض البديل ، مع ملاحظة أن القرار مرتبط بمستوى المعنوية المحدد ، بمعنى أن القرار قد يتغير إذا تغير مستوى المعنوية المستخدم (وفي بعض الحالات قد لا يتغير القرار، فهذا يتوقف على قيمة الإحصائية وما إذا كانت تقع في منطقة القبول أو منطقة الرفض).

أي أنه توجد طريقتين لاتخاذ القرار في الاختبارات الاحصائية وهي:

(i) حساب احصاء الاختبار ومقارنته بقيمة جدوليه وتحدد القيمة الجدولية بناء على نوع الاختبار ذو طرف واحد One Tail Test أو ذو طرفين Two Tail Test

(ii) حساب ما يسمى بالقيمة الاحتمالية P-value ويرمز لها في بعض البرامج الإحصائية بالرمز .Sig فاذا كان الاختبار ذو طرف واحد تقارن .Sig بالقيمة α لكن اذا كان الاختبار ذو طرفين تقارن بالقيمة α/2

مثال:

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

الحل :

١- **الفرض العدمي:** هو أن متوسط المجتمع يساوي 72 وبالرموز:

$$Ho: \mu = 72$$

٢- **الفرض البديل:** هو أن المتوسط لا يساوي 72 وبالرموز:

$$HI: \mu \neq 72$$

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

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

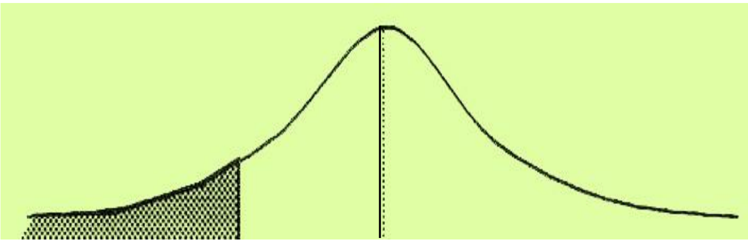
حيث أن: $n = 49 \quad , \quad \sigma = 14 \quad , \quad \bar{X} = 75 \quad , \quad \mu = 72$

وبالتعويض نحصل على:

$$Z_{\bar{X}} = \frac{75-72}{\frac{14}{\sqrt{49}}} = \frac{3}{\frac{14}{7}} = \frac{3}{2} = 1.5$$

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 :

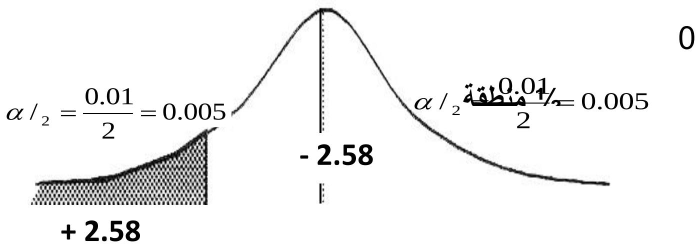


وقد حصلنا على حدود منطقتي القبول والرفض وذلك بقسمة درجة الثقة (**المكملة لمستوى المعنوية**) والتي تساوي 0.95 على 2 فنحصل على 0.4750 وبالكشف في جدول التوزيع الطبيعي المعياري عن Z التي تقابل المساحة 0.4750 نجد أنها تساوي 1.96 وحيث أنها موزعة على طرفي المنحنى بالتساوي فنضع إشارة موجبة في النصف الأيمن ،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

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:

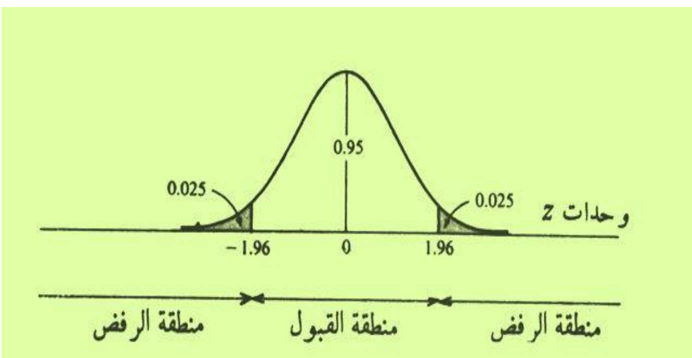
الحل:

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

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

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:

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 $\bar{X} = 520$ جرام و $S = 75$ جرام.

الحل:

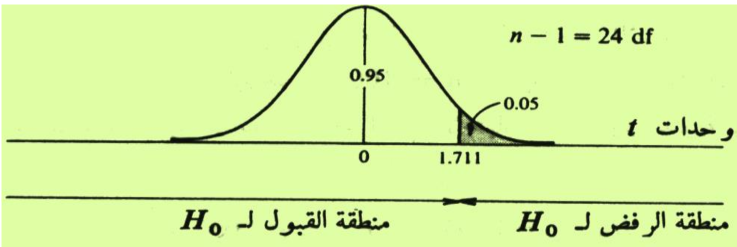
حيث أن الشركة ترغب في اختبار ما إذا كانت $\mu > 500$ ، فإن :

$H_0 : \mu = 500$ ، $H_1 : \mu > 500$

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:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{520 - 500}{75/\sqrt{25}} = \frac{20}{15} = 1.33$$

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



المحاضرة الثانية عشر

اختبار الفروض الإحصائية المعلمية

الجزء الأول

يهدف الباحث من اختبار الفرضيات حول المتوسط مثلا إلى اتخاذ قرار حول ما إذا كانت هذه الفرضية مقبولة أم مرفوضة ، ويتم ذلك من خلال استخدام مختبر إحصائي مناسب ، والمختبر الإحصائي هو متغير عشوائي ذو توزيع احتمالي يصف العلاقة بين القيم النظرية للمعلم والقيم المحسوبة من العينة ، وفي العادة تقارن قيمة المختبر الإحصائي المحسوب من العينة مع قيمته المستخرجة من توزيعه الاحتمالي (باستخدام جداول خاصة) ومنها نتخذ القرار برفض أو قبول الفرضية الصفرية .

أنواع الاختبار (الفروض)

١) أنواع الاختبار (الفروض) في حالة عينة واحدة:

بفرض اننا سوف نرمز للمعلمة المجهولة بالرمز μ ونريد اختبار الفرض القائل أن قيمتها تساوى μ_0 سيكون فرض العدم على الصورة التالية: $H_0 : \mu = \mu_0$

وسيكون الفرض البديل في حالة الاختبار ذو طرف واحد: $H_a : \mu < \mu_0 \quad or \quad H_a : \mu > \mu_0$

وسيكون الفرض البديل في حالة الاختبار ذو طرفين: $H_a : \mu \neq \mu_0$

---

٢) انواع الاختبار (الفروض) في حالة عينتين:

بفرض اننا سوف نرمز للمعلمة المجهولة بالرمز μ ونريد اختبار الفرض القائل ان قيمتها متساوية في المجتمعين سيكون فرض العدم على الصورة التالية: $H_0 : \mu_1 = \mu_2$

وسيكون الفرض البديل في حالة الاختبار ذو طرف واحد: $H_a : \mu_1 < \mu_2 \quad or \quad H_a : \mu_1 > \mu_2$

وسيكون الفرض البديل في حالة الاختبار ذو طرفين: $H_a : \mu_1 \neq \mu_2$

---

٣) انواع الاختبار (الفروض) في حالة أكثر من عينتين:

بفرض اننا سوف نرمز للمعلمة المجهولة بالرمز μ ونريد اختبار الفرض القائل ان قيمته متساوية في المجتمعات التي عددها r سيكون فرض العدم على الصورة التالية:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_r$$

وسيكون الفرض البديل:

$$H_a : \text{at least two are different}$$

١) 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 ($n > 30$) فإنه يمكن استخدام تباين العينة الكبيرة (S^2) عوضاً عن تباين المجتمع (σ^2) الغير معلوم ، وذلك لأن (S^2) مقدر جيد ($\sigma 2$) 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 .

خطوات اختبار Z :

- وضع فرض العدم والفرض البديل.

مثال: ينتج مصنع دقيق قمح في عبوات زنة العبوة (**2.5**) كيلوجرام ، فإن فرضية العدم هي:

$$H_0 : \mu = 2.5$$

- في حين يأخذ الفرض البديل عدة أشكال حسب طبيعة الاختبار:

$H_a : \mu \neq 2.5$	الفرض البديل أن متوسط المجتمع لا يساوي القيمة الافتراضية بغض النظر عن كون الاختلاف زيادة أو نقصاً.
$H_a : \mu > 2.5$	الفرض البديل أن متوسط المجتمع أكبر من القيمة الافتراضية.
$H_a : \mu < 2.5$	الفرض البديل أن متوسط المجتمع أصغر من القيمة الافتراضية.

- تحديد مستوى الدلالة (α): وتحدد مستويات المعنوية سلفاً وهي عادة **0.05 أو 5%**
- حساب إحصائية الاختبار (Z) حيث:

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

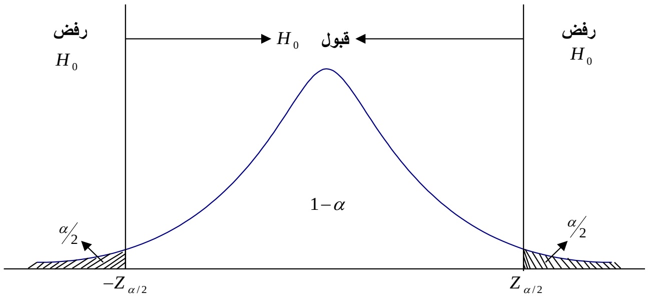
حيث أن :

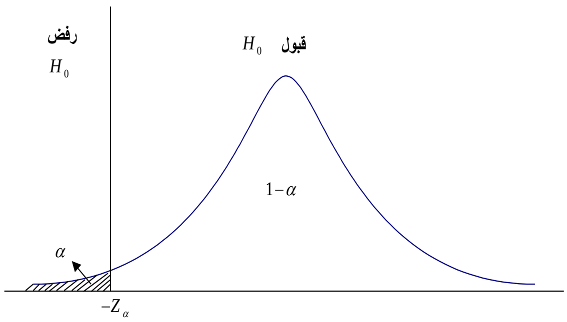
$\bar{X}$	الوسط الحسابي للعينة.
μ_0	القيمة الفرضية للوسط الحسابي للمجتمع.
α	مستوى الدلالة.
σ	الانحراف المعياري للمجتمع.
n	حجم العينة.

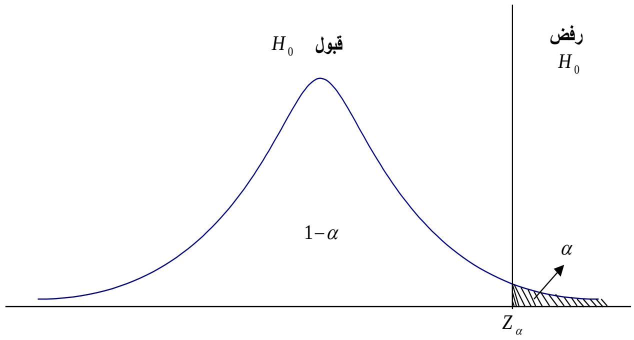
- اتخاذ قرار حول بيانات العينة من خلال مقارنة قيمة إحصائية الاختبار (المحسوبة) بالقيمة النظرية (الجدولية) للتوزيع الطبيعي المعياري عند مستوى معنوية محدد (Z_α). **وفيما يلي قاعدة القرار لرفض فرض العدم :**

الفرض البديل	قاعدة القرار: رفض فرض العدم إذا
$H_1 : \mu \neq \mu_0$	إذا كانت القيمة المطلقة لـ Z أكبر من قيمة Z النظرية عند مستوى معنوية.
$H_1 : \mu > \mu_0$	إذا كانت قيمة Z أكبر من قيمة Z النظرية عند مستوى معنوية.
$H_1 : \mu < \mu_0$	إذا كانت قيمة Z أقل من قيمة Z النظرية عند مستوى معنوية.

والقاعدة السابقة توضحها الأشكال التالية:







مثال على اختبار Z :

إذا كان متوسط استهلاك الفرد السعودي من الدجاج حسب تقارير وزارة الصحة هو (12) كيلوجرام بانحراف معياري (6) كيلوجرامات لفترة السبعينات الميلادية ، أجرى أحد الباحثين دراسة في عام ٢٠٠٣م من عينة قوامها (49) فرداً ووجد أن متوسط الاستهلاك للفرد هو (14) كيلوجرام ، هل تشير الدراسة الحالية أن متوسط الاستهلاك أرتفع عما عليه في السبعينات.

الحل:

١) فرض العدم والفرض البديل.

فرض العدم: $H_o: \mu=12$

الفرض البديل: $H_1: \mu>12$

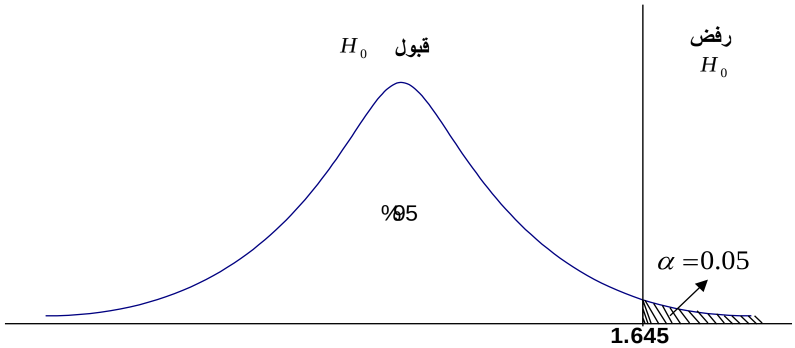
٢) مستوى الدلالة = (0.05)

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 (0.05) ، نحتاج لتحديد قيمة Z_α التي تقع على اليمين وتساوي 1.645

(أنظر الشكل التالي):



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

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 ارتفع بمستوى معنوي أو ذو دلالة عما علية في سبعينات القرن الماضي.

➢ اختبار t-test :

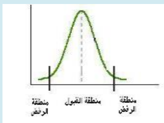
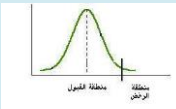
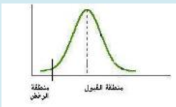
ولكن إذا كان حجم العينة صغيراً (30 > n) فإن قيمة (S^2) تتغير كثيراً من عينة إلى أخرى وبالتالي لا يمكننا هنا أن نستخدم اختبار(Z)، مما دفع كثيراً من الإحصائيين للبحث عن البديل المناسب .

ففي سنة ١٩٠٨م استطاع العالم الايرلندي وليم كوسيت W.S. Gosset من نشر بحث تحت اسم مستعار بسبب ظروف خاصة هو (استيودنت، Student) استطاع من خلاله أن يشتق معادلة للتوزيع الاحتمالي (t) الذي قيمته هي :

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

وهذا الاختبار يشبه اختبار التوزيع الطبيعي (Z) ، بيد أنه يختلف عنه في تضمنه للانحراف المعياري (S) للعينة بدلاً من الانحراف المعياري (σ) للمجتمع .

يستخدم هذا الصنف من اختبار(t) للحكم على معنوية الفروق بين متوسط عينة ومتوسط المجتمع الذي سحبت منه ، ولغرض توضيح ذلك إليك هذا العرض الموجز لخطوات اختبار (t) حول متوسط حسابي واحد على افتراض أن تباين المجتمع σ^2 غير معلوم وأن حجم العينة صغير.

خطوات الاختبار	اختبار ذو طرفين	اختبار ذو طرف واحد	
		طرف يمين	طرف يسار
١- الفرضية الصفرية H_0 ٢- الفرضية البديلة H_1	$\mu = \mu_0$ $\mu \neq \mu_0$	$\mu = \mu_0$ $\mu > \mu_0$	$\mu = \mu_0$ $\mu < \mu_0$
٣- مستوى الدلالة	α		
٤- منطقة الرفض	ا ت ا ≤ ت (α ،df) 	ت ≤ ت (α،df) 	ت ≥ ت (α،df) 
٥- المختبر الإحصائي	$t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ هذا المختبر الإحصائي يستخدم لتوضيح أهمية الفروق بين الوسط الحسابي للعينة والوسط الحسابي للمجتمع		
٦- القرار	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [(α $\frac{1}{2}$ ،df) ت] الجدولية أي أن: ا ت ا ≤ ت (α $\frac{1}{2}$ ،df)	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [(α ،df) ت] الجدولية أي أن: ت ≤ ت (α،df)	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [(α،df) ت-] الجدولية أي أن: ت ≥ - ت (α،df)

مثال على اختبار t :

لو كانت لدينا عينة عشوائية تتكون من 250 طالب وجد أن الوسط الحسابي لأطوال طلاب العينة 155.95 سم، والانحراف المعياري = 2.94 سم، علما بأن الوسط الحسابي لأطوال طلاب الجامعة يبلغ 158 سم، اختبر أهمية الفرق المعنوي بين الوسط الحسابي لأطوال طلاب العينة والوسط الحسابي لأطوال 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:

$$(\mu = \mu_0)$$

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

$$(\mu \neq \mu_0)$$

مستوى الدلالة : α = 0.05

منطقة الرفض : قيمة (ت) الجدولية عند مستوى دلالة α = 0.05 ودرجات حرية 249 = 1.960

المختبر الإحصائي :

$\bar{X}$ = 155.95 سم ، n = 250 طالب ، S = 2.94 سم ، μ = 158 سم

وبالتعويض في المعادلة التالية:

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

القرار:

قيمة t المحسوبة (-11.006) أصغر من قيمة *t* المجدولة (-1.96) عند مستوى دلالة α = 0.05

أو نقول أن القيمة المطلقة للقيمة المحسوبة أكبر من القيمة المطلقة للقيمة الجدولية.

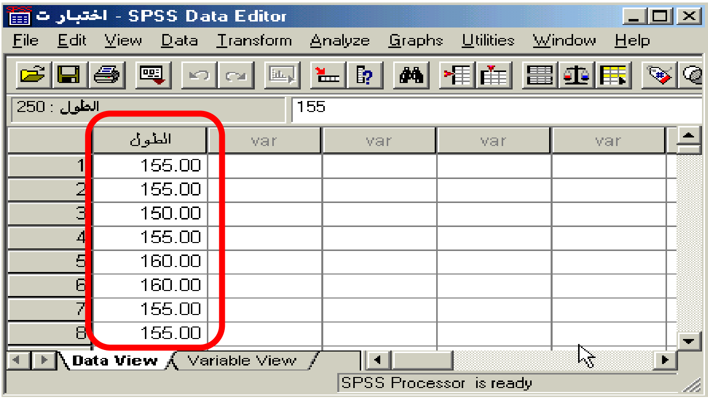
∴ نرفض الفرضية الصفية ونقبل البديلة .

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

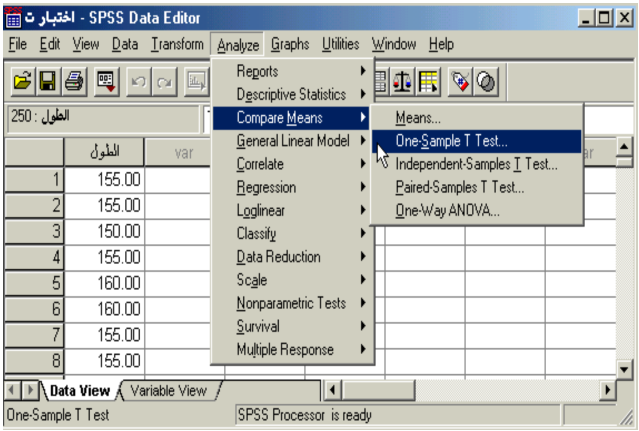
حساب اختبار (ت) لعينة واحدة One Sample T-Test من خلال الـ SPSS

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

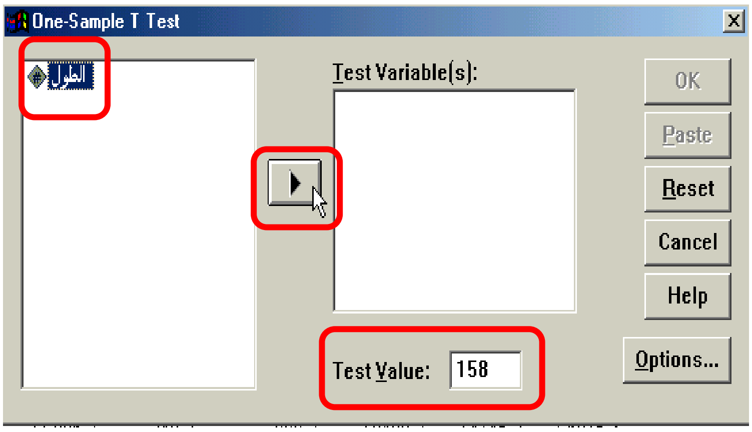
✓ قم بإدخال البيانات المراد تحليلها من خلال شاشة تحرير البيانات Data Editor بالطريقة المناسبة كالتالي :



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



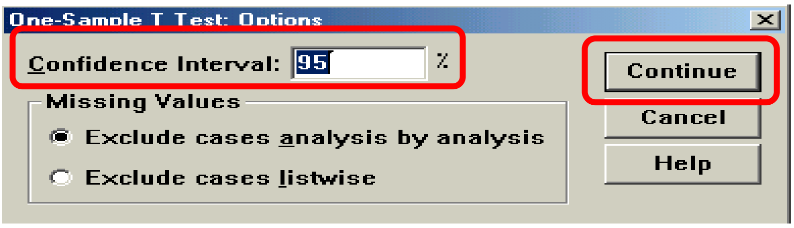
✓ بعد اختيار الأمر "اختبار (ت) لعينة واحدة" One-Sample T Test سوف يظهر لك صندوق الحوار التالي :



✓ 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 158 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 95%) ، 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 :

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها ولديك المعرفة عن القبول والرفض.

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

يتضح من النتائج أن قيمة (ت) المحسوبة t-test = -11.006 ، ودرجات الحرية df = 249 ، وقيمة Sig. (2-tailed) = 0.000 ، وبما أن قيمة Sig. (2-tailed) في الجدول (0.000) أصغر من قيمة α = 0.05 فإننا بالتالي نرفض الفرضية الصفرية ، أي أنه توجد فروق ذات دلالة إحصائية بين متوسط أطوال العينة ومتوسط أطوال طلاب الجامعة .

---

٢) الاختبارات الاحصائية لعينتين مستقلتين Independent Samples t-test

يستخدم هذا النوع من اختبار (ت) للحكم على معنوية Significance الفروق بين متوسطي عينتين غير مرتبطتين Independent ، ولغرض توضيح ذلك إليك هذا العرض الموجز لخطوات اختبار (ت) حول متوسطين على افتراض أن تباين المجتمع σ_1^2 و σ_2^2 غير معلومين ولكنهما متجانسين وأن حجم العينة صغير.

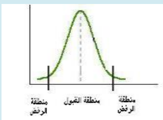
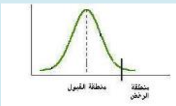
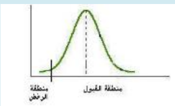
ويستند هذا الاختبار إلى توفر عدد من الافتراضات وهذه الافتراضات هي :

- مستوى القياس: يشترط لاستخدام هذا الاختبار أن تكون البيانات فئوية (فتريه) أو نسبية.
- أن يكون حجم العينة صغيراً: يقتضي هذا الافتراض أن يكون حجم العينة أقل من (30) وأكبر من (5) (وإذا كان حجم العينة أكبر من 30 فلا بأس من ذلك) ، ولتجنب الخطأ نستعيض عن (σ) بالمعلمة (S) ، ويجب

كذلك أن يكون الفرق بين حجم عينتي البحث صغيراً ، لأنه كلما زاد الفرق بين حجم العينتين أثر ذلك على قيمة (t) المحسوبة.

- **التوزيع الطبيعي:** ويقضي هذا الافتراض أن المشاهدات (N1) في المجتمع الأول تتخذ شكل التوزيع الطبيعي لوسط يساوي (μ_1) وكذلك الأمر بالنسبة للمشاهدات (N2) في المجتمع الثاني يفترض فيها أن تتخذ شكل التوزيع الطبيعي لوسط يساوي (μ_2) ، وإن مخالفة هذا الافتراض ليس لها تبعات تذكر.
- **تجانس التباين في المجتمعين:** وبموجب هذا الافتراض يكون لتباين المشاهدات في كل من المجتمعين نفس القيمة (σ^2) وبذلك تكون القيمة المتوقعة للتباين في كل العينتين مساوية للمقدار (σ^2) أي يكون كل من (S_1^2 ، S_2^2) تقديراً مستقلاً لنفس المقدار (σ^2)
- **الاستقلالية:** ويقتضي هذا الافتراض أن (n_1) من المشاهدات قد تم الحصول عليها عشوائيا من المجتمع الأول بشكل مستقل عن (n_2) من المشاهدات والتي تم الحصول عليها عشوائياً من المجتمع الثاني ، وإن الاستقلالية هنا لا تعني استقلالية البيانات بين المجتمعين فقط ، بل تعني استقلالية المشاهدات ضمن المجتمع الواحد أيضاً (مثل عملية تطبيق اختبار قبلي واختبار بعدي على مجموعة واحدة).

والجدول التالي يوضح خطوات اختبار الفرق بين متوسطين باستخدام المختبر الإحصائي (ت) على افتراض أن تباين المجتمع σ_1^2 و σ_2^2 غير معلومين ولكنهما متجانسين ، وأن كل من n_1 و n_2 صغير.

خطوات الاختبار	اختبار ذو طرفين	اختبار ذو طرف واحد	
		طرف يمين	طرف يسار
١- الفرضية الصفرية H0 ٢- الفرضية البديلة H1	$\mu = \mu_0$ $\mu_0 \neq \mu$	$\mu = \mu_0$ $\mu > \mu_0$	$\mu = \mu_0$ $\mu < \mu_0$
٣- مستوى الدلالة	α		
٤- منطقة الرفض	ا ت ا ≤ ت (df، $\alpha\frac{1}{2}$) 	ت ≤ ت (df، α) 	ت ≥ ت (df، α) 
٥- المختبر الإحصائي	$t = \frac{\overline{X}_1 - \overline{X}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ هذا المختبر الإحصائي يستخدم لتوضيح أهمية الفروق بين متوسطين للعينات المستقلة وعند تجانس التباين		
٦- القرار	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [ت (df، $\alpha\frac{1}{2}$)] الجدولية أي أن: ا ت ا ≤ ت (df، $\alpha\frac{1}{2}$)	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [ت (df، α)] الجدولية أي أن: ت ≤ ت (df، α)	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [-ت (df، α)] الجدولية أي أن: ت ≥ - ت (df، α)

ولتوضيح ما ورد في الجدول السابق دعنا نتناول هذا المثال :

أراد باحث أن يعرف أثر استخدام نظم مساندة القرارات على كفاءة القرارات التي تتخذها الإدارة بمساعدة تلك النظم ، فوزع 50 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
25 = n_2	25 = n_1
6 = $\overline{X}_2$	7.60 = $\overline{X}_1$
1.78 = S_2^2	2.27 = S_1^2

فهل تدل هذه البيانات على أن أداء المجموعة التجريبية كان أفضل من أداء المجموعة الضابطة عند مستوى α = 0.05 ؟

الحل :

سيتم اختبار الفرضيات التالية :

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

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

مستوى الدلالة : α = 0.05

منطقة الرفض : ٠.٠ قيمة مستوى الدلالة α = 0.05 والاختبار بذيل واحد ، ودرجات الحرية = 25 + 25 - 2 = 48 ، بذلك تكون قيمة (ت) الجدولية = 1.68

المختبر الإحصائي :

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

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

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

إذا التباين يساوي:

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

إذاً الانحراف المعياري يساوي :

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

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

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

القرار:

قيمة (t) المحسوبة (2.77) أكبر من قيمـة (ت) المجدولة (1.68) عند مستوى دلالة $\alpha = 0.05$.

∴ 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 $\alpha = 0.05$

❖ أما في حالة عدم تجانس التباين فإننا نستخدم طريقة ويلتش Weltch لحساب قيمة (t) وذلك من خلال تطبيق العلاقة:

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

وللتحقق من تجانس التباين يتم ذلك من خلال تطبيق العلاقة التالية:

$$F = \frac{S_g^2}{S_l^2}$$

حيث أن:

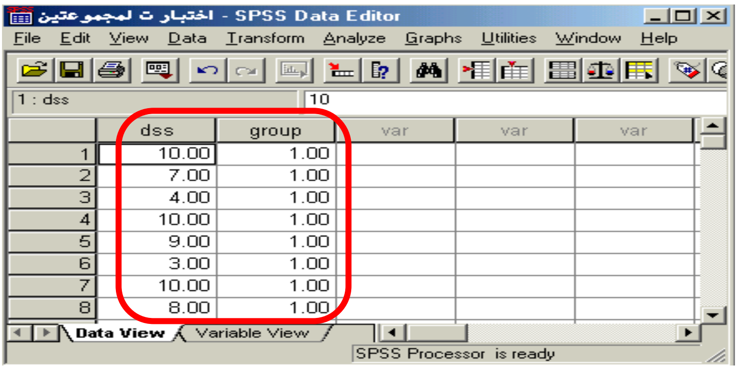
S_g^2	تعني التباين الأكبر.
S_l^2	تعني التباين الأصغر.

ومن ثم مقارنة قيمة (F) المحسوبة (لقياس تجانس التباين) بقيمة (F) المجدولة عند درجة حرية ($1-n_1$ و $1-n_2$)

حساب اختبار(ت) للعينات المستقلة Independent Samples T-Test من خلال الـ SPSS

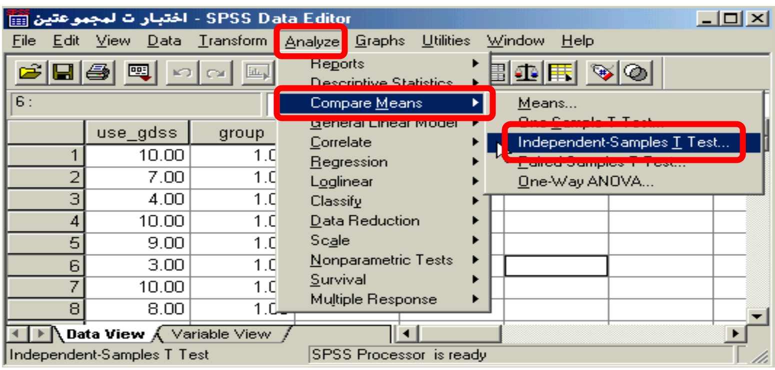
لغرض حساب قيمة (ت) لنفس المثال السابق من خلال استخدام برنامج الـ SPSS نتبع الخطوات التالية:

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

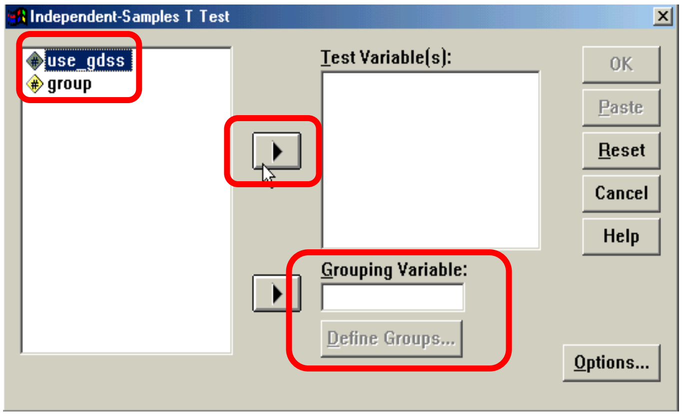


لاحظ أنه تم إدخال النتائج المتحصل عليها تحت متغير واحد باسم dss ، وتم إنشاء متغير آخر بمسمى group ليحوي رمز للمجموعة التجريبية والمجموعة الضابطة .

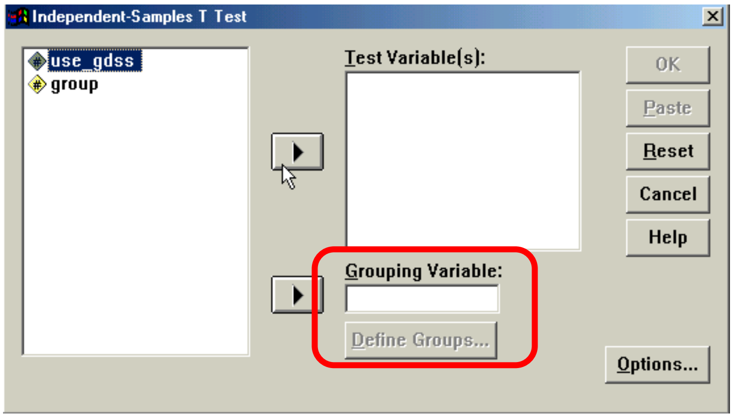
✓ من القائمة "تحليل" Analyze اختر الأمر "مقارنة المتوسطات" Compare Means فتظهر قائمة أوامر فرعية اختر منها "اختبار(ت) للعينات المستقلة" Independent Samples T-Test كالتالي:



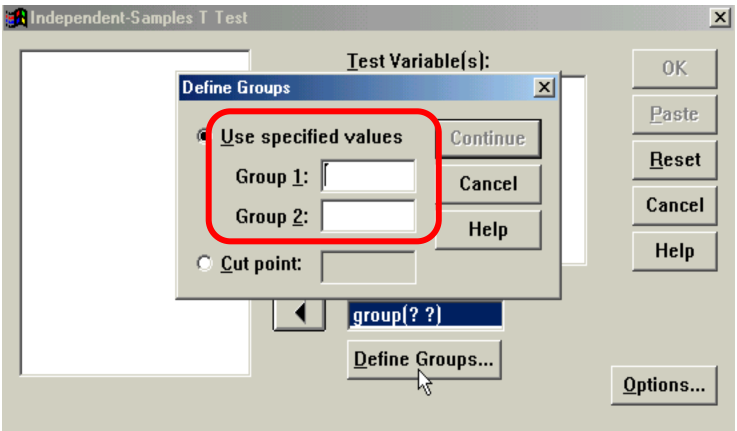
✓ بعد اختيار الأمر "اختبار (ت) للعينات المستقلة" Independent Samples T-Test سوف يظهر لك صندوق الحوار التالي :



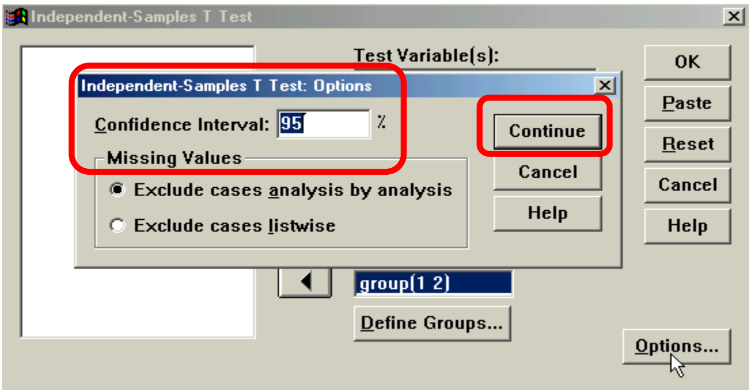
✓ من قائمة المتغيرات في الجهة اليسرى من صندوق الحوار حدد المتغير المراد نقله إلى المستطيل الخاص بـ "متغيرات الاختبار" Test Variable(s) ومن ثم أنقر على السهم الذي يظهر مقابل المستطيل الخاص بـ "متغيرات الاختبار" ، ستلاحظ انتقال المتغير مباشرة في المستطيل "متغيرات الاختبار" Test Variable(s) ، واعمل نفس الشيء مع المتغير group لنقله إلى الحقل الخاص بـ "مجاميع المتغيرات" Grouping Variable .



✓ أنقر على زر "تعريف المجموعات" Define Groups في أسفل صندوق الحوار السابق ، سيؤدي ذلك إلى فتح صندوق حوار صغير يتيح الفرصة للمستخدم لتعريف قيم المجموعة الأولى (والتي يمثلها الرقم ١) وقيم المجموعة الثانية (والتي يمثلها الرقم ٢) .



✓ 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 95%) ، 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

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها ولديك المعرفة عن القبول والرفض.

Group Statistics

	GROUP	N	Mean	Std. Deviation	Std. Error Mean
USE_GDSS	USE DSS	25	7.6000	2.2730	.4546
	NOT USE DSS	25	6.0000	1.7795	.3559

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. E Differe
USE_GDSS	Equal variances assumed	1.095	.301	2.771	48	.008	1.6000	.5
	Equal variances not assumed			2.771	45.386	.008	1.6000	.5

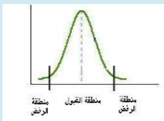
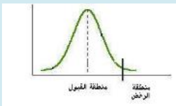
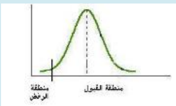
يتضح من النتائج أن قيمة (F) = 1.095 ومستوى دلالتها 0.301 وهذه القيمة أكبر من 0.05 ، مما يدل على أنها غير دالة (وهذا يعني أن هناك تجانس بين تباين المجموعتين)، وهذا يدفعنا إلى قراءة نتائج اختبار (ت) المقابلة للعبارة "افتراض تساوي التباين" Equal variances assumed ، من هذه النتائج نلاحظ أن قيمة (ت) المحسوبة t-test = 2.771 ، ودرجات الحرية df = 48 ، وقيمة Sig. (2-tailed) =0.008، وبما أن قيمة Sig. (2-tailed) في الجدول (0.008) أصغر من قيمة α = 0.05

فإننا بالتالي نرفض الفرضية الصفرية ، أي أنه توجد فروق ذات دلالة إحصائية بين متوسط المجموعة التجريبية ومتوسط المجموعة الضابطة لصالح المجموعة الضابطة (وذلك بسبب حصولها على متوسط حسابي أكبر = 7.60).

الاختبارات الاحصائية لعينتين غير مستقلتين (العينات المرتبطة) Paried Samples t-test

يستخدم هذا النوع للحكم على دلالة الفروق ومعنويتها Significance بين متوسطي عينتين مرتبطتين Correlated Data ، مثل اختبار دلالة الفروق بين متوسط أداء الموظفين قبل التدريب وبعد التدريب ، ولغرض توضيح ذلك إليك هذا العرض الموجز لخطوات اختبار (ت) حول متوسطين مرتبطين على افتراض أن تباين المجتمع و غير معلومين وأن حجم العينة صغير .

والجدول التالي يوضح خطوات اختبار الفرق بين متوسطين مرتبطين باستخدام المختبر الإحصائي (ت) على افتراض أن تباين المجتمع σ_1^2 و σ_2^2 غير معلومين ، وأن حجم العينة صغير.

خطوات الاختبار	اختبار ذو طرفين	اختبار ذو طرف واحد	
		طرف يمين	طرف يسار
١- الفرضية الصفرية H0 ٢- الفرضية البديلة H1	μ = μ0 μ0 ≠ μ	μ = μ0 μ > μ0	μ = μ0 μ < μ0
٣- مستوى الدلالة	α		
٤- منطقة الرفض	ا ت ا ≤ ت ($\frac{1}{2}$ α ،df) 	ت ≤ ت (α،df) 	ت ≥ ت (α،df) 
٥- المختبر الإحصائي	$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} - 2r\left(\frac{S_1}{\sqrt{n_1}}\right)\left(\frac{S_2}{\sqrt{n_2}}\right)}}$ هذا المختبر الإحصائي يستخدم لتوضيح أهمية الفروق بين متوسطين للعينات المرتبطة		
٦- القرار	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [ت ($\frac{1}{2}$ α،df)] الجدولية أي أن: ا ت ا ≤ ت ($\frac{1}{2}$ α،df)	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [ت (α ،df)] الجدولية أي أن: ت ≤ ت (α،df)	أرفض الفرضية الصفرية إذا كانت قيمة المختبر الإحصائي (ت) المحسوبة تساوي أو أكبر من قيمة [-ت (α،df)] الجدولية أي أن: ت ≥ - ت (α،df)

ولتوضيح ما ورد في الجدول السابق دعنا نتناول هذا المثال :

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 ، لا بد على الباحث أن يتأكد من قيمة الارتباط بين الاختبارين والتي كانت $r = 0.46$ ، وقد كانت النتائج التي تم التوصل إليها كما يلي :

الاختبار القبلي	الاختبار البعدي
n_1 = 100	n_2 = 100
$\overline{X}_1$ = 54.28	$\overline{X}_2$ = 58.66
S_1^2 = 49	S_2^2 = 64

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 = 0.05 ؟

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 ($\mu_2 \neq \mu_1$)

مستوى الدلالة : 0.05 = α

منطقة الرفض : قيمة مستوى الدلالة 0.05 = α والاختبار بذيلين ، ودرجات الحرية = 100 – 1 = 99 ، بذلك تكون قيمة (ت) الجدولية = 1.980

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 (5.57) أكبر من قيمة (ت) المجدولة (1.980) .عند مستوى دلالة 0.05 = α .

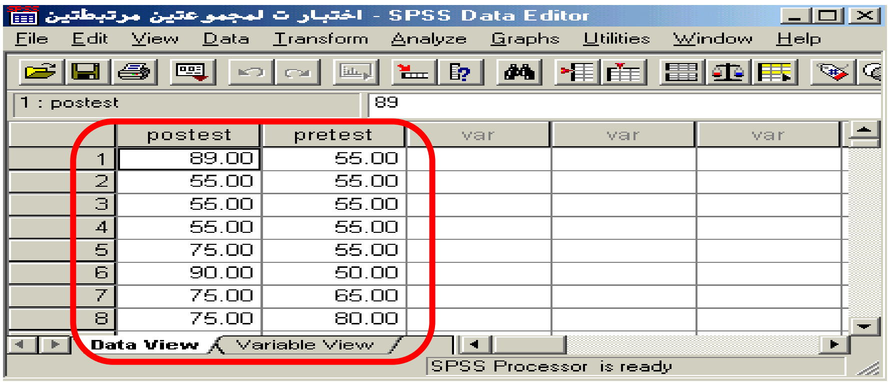
∴ نرفض الفرضية الصفرية ونقبل البديلة.

أي أن للبرنامج التدريبي تأثير إيجابي على تحصيل الطلاب وأدائهم في الكلية وذلك عند مستوى دلالة 0.05 = α

حساب اختبار(ت) لعينتين غير مستقلتين (العينات المرتبطة) Paired Samples T-Test من خلال الـ SPSS

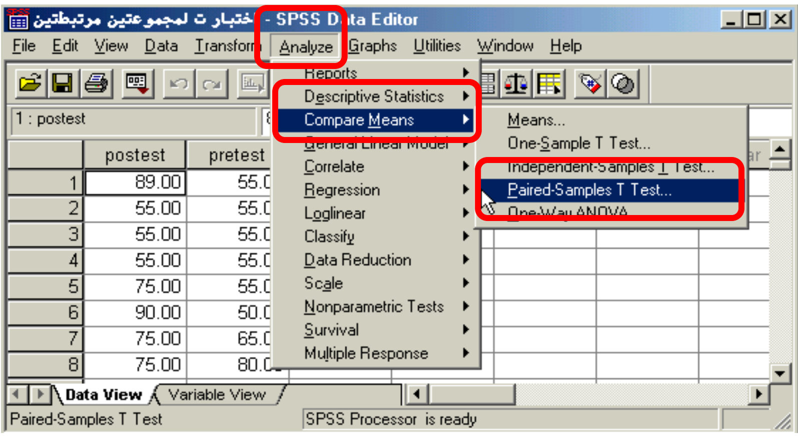
لغرض حساب قيمة (ت) لنفس المثال السابق من خلال استخدام برنامج الـ SPSS نتبع الخطوات التالية :

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

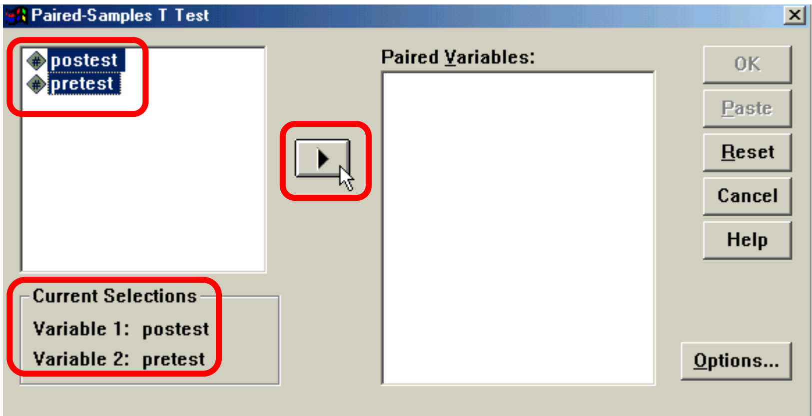


لاحظ أنه تم إدخال البيانات بطريقة مختلفة عن ما تم اتباعه في حالة العينتين المستقلتين ، هنا لابد من إدخال بيانات كل متغير في عمود منفصل عن الآخر ، وقد تم إعطاء كل متغير اسم مختلف عن الآخر posttest و pretest

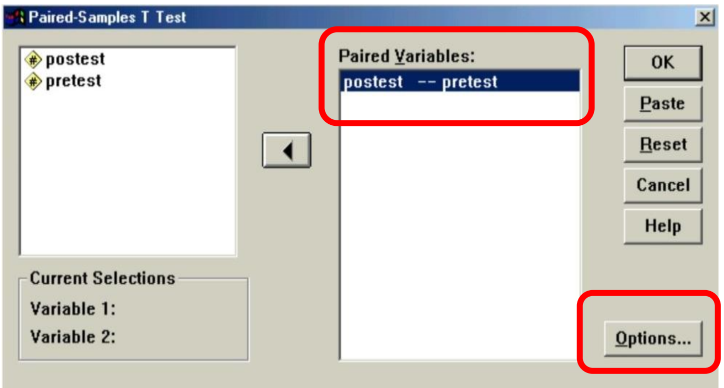
✓ من القائمة "تحليل" Analyze اختر الأمر "مقارنة المتوسطات" Compare Means فتظهر قائمة أوامر فرعية اختر منها "اختبار (ت) للعينات المرتبطة" Paired-Samples T-Test كالتالي :



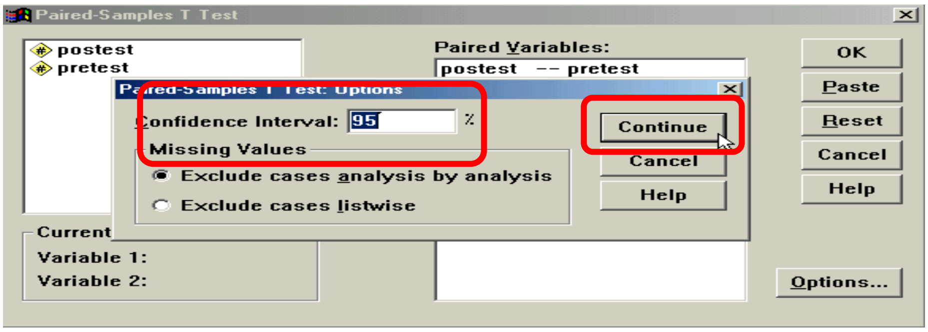
✓ بعد اختيار الأمر "اختبار (ت) للعينات المرتبطة " Paired-Samples T-Test سوف يظهر لك صندوق الحوار التالي :



✓ 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 95%) ، 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

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها ولديك المعرفة عن القبول والرفض.

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 (58.6600) والانحراف المعياري لنفس المتغير (8.00) ، أما المتغير Pretest فقد كان المتوسط الحسابي (54.2800) والانحراف المعياري (7.00) ، بالإضافة إلى ذلك تم حساب معامل ارتباط بيرسون للمتغيرات موضع الدراسة Paired Sample Correlation وقد كانت قيمته (0.458) .

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 = 5.575 ، ودرجات الحرية df = 99 ، وقيمة Sig. (2-tailed) =0.000 ، وبما أن قيمة Sig. (2-tailed) في الجدول (0.000) أصغر من قيمة α = 0.05 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 = 0.05 .

المحاضرة الثالثة عشر

اختبار الفروض الإحصائية المعلمية

الجزء الثاني

٣) الاختبارات الاحصائية لأكثر من عينتين مستقلتين

ناقشنا في المحاضرة السابقة طرق الاستدلال الإحصائي عن متوسط المجتمع والفرق بين متوسطين ، وسنناقش في هذه المحاضرة طرق الاستدلال الإحصائي للفرق بين ثلاث متوسطات أو أكثر وذلك من خلال توزيع F.

سمي توزيع F بهذا الاسم تخليدا للعالم رونالد فيشر R.A. Fisher الذي يعتبر أول من اشتق هذا التوزيع ووصفه وذلك في العشرينات والثلاثينات من القرن العشرين لذلك تعرف أحيانا بتحليل فيشر للتباين.

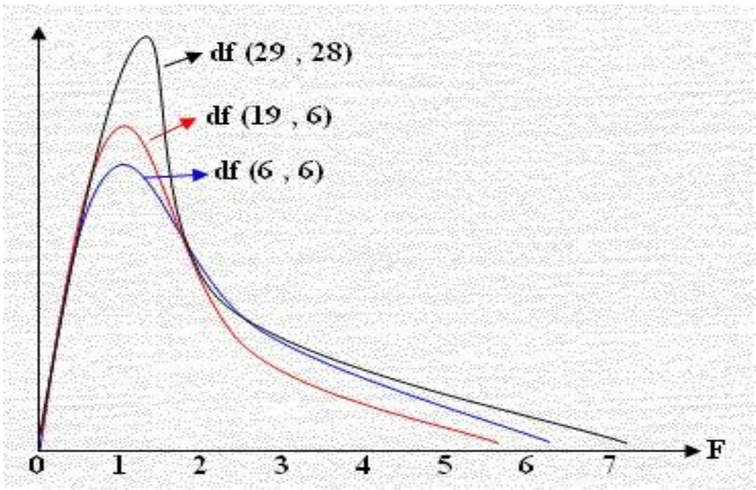
ويستخدم توزيع F أساسا لاختبار تساوي تبايني مجتمعين ، ومن المثير للانتباه ملاحظة أن اختبار تساوي التباينين يستخدم لاختبار تساوي ثلاث متوسطات أو أكثر.

وتسمى طريقة الاستدلال الاحصائي عن تساوي ثلاث متوسطات أو أكثر بتحليل التباين Analysis of Variance .

وتوزيع F عبارة عن مجموعة من المنحنيات التكرارية يتميز كل منها عن الآخر برقمين لدرجات الحرية أحدهما يمثل درجة حرية للبسط والآخر درجة حرية للمقام.

وقيمة F هي قيمة توضح نسبة التباين Variance ratio لعينتين والرمز F إشارة إلي العالم Fisher الذي قام بعمل هذا الاختبار والمعروف باختبار F وقد قام العالم Snedecor بحساب جداول خاصة لتوزيع F وفيها درجات الحرية التي في أعلى الجدول تخص البسط أما درجات الحرية على العمود الجانبي فتخص المقام.

وتوزيع F هو توزيع ملتو جهة اليمين بمعلمتين تتمثلان بدرجتي حرية (البسط ، المقام) وهما k – 1 للبسط ، n – k للمقام حيث n مجموع أحجام العينات و k هي عدد المجموعات موضع الدراسة.



أهم استخدامات توزيع F (ف) هي :

- تقدير فترة الثقة لـ σ_1^2/σ_2^2
- اختبار فرضيات حول تساوي تباينين أي : $H_0:\sigma_1^2=\sigma_2^2$
- اختبار فرضيات حول تساوي أكثر من متوسطين أي : $H_0:\mu_1=\mu_2=\mu_3=....$

تحليل التباين الأحادي One Way ANOVA

تحليل التباين هو عملية يقصد بها تقسيم مجموع مربعات الانحرافات عن المتوسط الحسابي إلي مكوناته إرجاع كل من هذه المكونات إلي مسبباتها ، وطريقة تحليل التباين تفيد في مقارنة عدد من المعاملات يزيد عن أثنين كما تمتاز طريقة تحليل التباين بأنه يمكن فيها استعمال كل البيانات المأخوذة من التجربة في حساب قيمة واحدة للانحراف القياسي يمكن بها مقارنة المجموعات أو المعاملات التجريبية.

فهي مجموعة من النماذج الإحصائية (statistical model) مع إجراءات مرافقة لهذه النماذج تمكن من مقارنة المتوسطات لمجتمعات إحصائية مختلفة عن طريق تقسيم التباين variance الكلي الملاحظ بينهم إلى أجزاء مختلفة.

تتلخص طريقة تحليل التباين في:

١- حساب المجموع الكلي لمربعات انحرافات كل المفردات في التجربة عن المتوسط العام.
٢- تقسيم هذا المجموع الكلي لمربعات الانحرافات Total Sum Squares إلي مكوناته طبقا للمصادر المسببة لها والذي يختلف عددها طبقا للتصميم المستعمل في التجربة.
٣- تقسم درجات الحرية الكلية طبقا للمصادر السابقة أيضا.
٤- تدون النتائج في جدول يسمي جدول تحليل التباين ANOVA ترتب فيه مصادر الاختلافات حسب التصميم الإحصائي المستعمل ويسهل هذا الجدول عمل اختبار معنوية المعاملات.

تحليل التباين الأحادي (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.

الافتراضات الاساسية لاختبار تحليل التباين:

يستند اختبار تحليل التباين إلى توفر عدد من الافتراضات ، ومن هذه الافتراضات ما يلي :

١- مستوى القياس :

يشترط لاستخدام هذا الاختبار أن تكون البيانات فترية (فئوية) أو نسبية.

٢- حجم العينة :

يقتضي هذا الافتراض أن يكون حجم العينة كبيراً .

٣- التوزيع الطبيعي للمجتمع الإحصائي :

يقتضي هذا الافتراض أن تكون المشاهدات في كل مجتمع من المجتمعات موزعة بشكل طبيعي ، ولكن يرى الإحصائيون أن اختبار (F) لا يتأثر كثيراً بعدم توفر هذا الشرط وذلك عندما يكون حجم الخلية كبيراً والتوزيع ليس طبيعياً.

٤- تجانس التباين :

أي أن يكون للمجتمعات في مستويات المعالجة المختلفة نفس التباين (S^2) بالرغم من أن لها بالطبع أوساطاً مختلفة.

مثال :

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

المنتج (1) X_1	المنتج (2) X_2	المنتج (3) X_3
7	4	2
10	6	2
10	7	3
11	9	7
12	9	6
50	35	20

المطلوب: هل هناك فروق ذات دلالة بين المنتجات الثلاثة ؟

الحل :

لكون لدينا ثلاث متغيرات فترية ، ولرغبة الشركة معرفة الفروق بين هذه المتغيرات موضع الدراسة ، فإن أنسب أسلوب إحصائي هنا هو تحليل التباين الأحادي One Way ANOVA ، ولغرض حساب تحليل التباين الأحادي، **علينا اتباع الخطوات 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 $\left(\sum X\right)^2$ لكل متغير.
- نحسب متوسط كل متغير من خلال العلاقة : $\overline{X} = \frac{\sum x}{n}$
- نحسب مجموع المربعات الكلي Total Sum of Squares وذلك من خلال العلاقة التالية :

$$Total..SS = \sum X^2 - \frac{\left(\sum X\right)^2}{n}$$

حيث أن :

n	تعني مجموع أعداد الأفراد في جميع المجموعات.

أو يمكن حساب مجموع المربعات الكلي من خلال العلاقة التالية :

Total SS = Between SS + Within SS

- نحسب مجموع المربعات بين المجموعات Between Sum of Squares وذلك من خلال العلاقة التالية :

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

حيث أن :

n_g	تعني عدد الأفراد في كل مجموعة.
n	تعني مجموع أعداد الأفراد في جميع المجموعات.
$\sum \frac{\left(\sum X_g\right)^2}{n_g}$	تعني مربع مجموع قيم كل مجموعة مقسوما على عدد أفراد تلك المجموعة.
$\frac{\left(\sum X\right)^2}{n}$	تعني مربع مجموع قيم كل المجموعات مقسوما على مجموع أعداد الأفراد في جميع المجموعات.

- نحسب مجموع المربعات داخل المجموعات Within Sum of Squares وذلك من خلال العلاقة التالية :

$$Within..SS=\sum\left[\sum X_g^2-\frac{\left(\sum X_g\right)^2}{n_g}\right]$$

أو يمكن حساب مجموع المربعات داخل المجموعات من خلال العلاقة التالية :

Within SS = Total SS - Between SS

- نحسب درجات الحرية بين المجموعات Between groups degrees of freedom من خلال العلاقة التالية :

$K-1$ حيث *K* تعني عدد المجموعات .

ودرجات الحرية داخل المجموعات Within groups degrees of freedom من خلال العلاقة التالية :

$n-K$ حيث *n* تعني عدد الأفراد أو الاستجابات في المجموعات موضع الدراسة ، و *K* تعني عدد المجموعات .

ودرجات الحرية الكلية Total degrees of freedom من خلال العلاقة التالية :

$n-1$ حيث *n* تعني عدد الأفراد أو الاستجابات في المجموعات موضع الدراسة .

- نحسب التباين بين المجموعات أو ما يسمى متوسط المربعات بين المجموعات Between mean square وذلك من خلال العلاقة التالية:

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

- نحسب التباين داخل المجموعات أو ما يسمى متوسط المربعات داخل المجموعات Within mean square وذلك من خلال العلاقة التالية :

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

- نحسب قيمة *F* من خلال العلاقة التالية :

$$F=\frac{Between..groups..mean..square}{Within..groups..mean..square}$$

- نقارن بعد ذلك قيمة *F* المحسوبة بقيمة *F* المجدولة لاتخاذ القرار المناسب اتجاه الفرضية موضع الدراسة.

❖ نقوم الآن بتطبيق جميع الخطوات السابقة لحساب تحليل التباين الأحادي على بيانات المثال السابق وذلك حتى يسهل علينا استخلاص بعض القيم المطلوبة لحساب هذا النوع من الاختبارات الإحصائية وتطبيق المعادلات السابقة .

المنتج (1) X_1		المنتج (2) X_2		المنتج (3) X_3	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
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
50	514	35	263	20	102

❑ وضع فرض العدم والفرض البديل.

صياغة الفرضية الصفرية كالتالي:

H0: μ١ = μ٢ = μ٣

في حين تفترض الفرضية البديلة التالي:

HA: متوسطان على الأقل غير متساويين

❑ تحديد مستوى الدلالة (α): وتحدد مستويات المعنوية سلفاً وهي عادة 0.05 أو 0.01

❑ حساب إحصائية الاختبار (*F*) وذلك من خلال اتباع الخطوات التالية:

✓ المتوسط الحسابي لـ X_1 = $\overline{X} = \frac{50}{5} = 10$

✓ المتوسط الحسابي لـ X_2 = $\overline{X} = \frac{35}{5} = 7$

✓ المتوسط الحسابي لـ X_3 = $\overline{X} = \frac{20}{5} = 4$

✓ مجموع المربعات الكلي Total Sum of Squares =

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

حيث أن :

n_g	تعني عدد أفراد المجموعة المحددة.
k	تعني عدد المجموعات موضع الدراسة.

✓ مجموع المربعات بين المجموعات Between Sum of Squares =

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

✓ مجموع المربعات داخل المجموعات Within Sum of Squares =

$$Within..SS = \sum \left[\sum X_g^2 - \frac{(\sum X_g)^2}{n_g} \right]$$

$$\sum x_1^2 = 514 - \frac{(50)^2}{5} = 14$$

$$\sum x_2^2 = 263 - \frac{(35)^2}{5} = 18$$

$$\sum x_3^2 = 102 - \frac{(20)^2}{5} = 22$$

نقوم بعد ذلك بجمع نواتج هذه المعادلات لنحصل على مجموع المربعات داخل المجموعات كالتالي :

Within sum of squares = 14 + 18 + 22 = 54

✓ 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 =

$$Beween..groups..mean..square=\frac{Between..SS}{K-1}=\frac{90}{2}=45$$

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

$$Within..groups..mean..square=\frac{Within..SS}{(n-K)}=\frac{54}{12}=4.5$$

✓ 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	90	2	45	10
داخل المجموعات Within groups	54	12	4.5	
الكلي (المجموع) Total	144	14		

وبالرجوع إلى جدول توزيع *F* نجد أن القيمة الحرجة لـ *F* بدرجات حرية للبسط تساوي 2 ودرجات حرية للمقام تساوي 12 وباستخدام مستوى= 0.05 نجد أن القيمة الحرجة تساوي 3.89 .

وحيث أن القيمة المحسوبة لـ *F* = 10 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

Table of F-statistics P=0.05

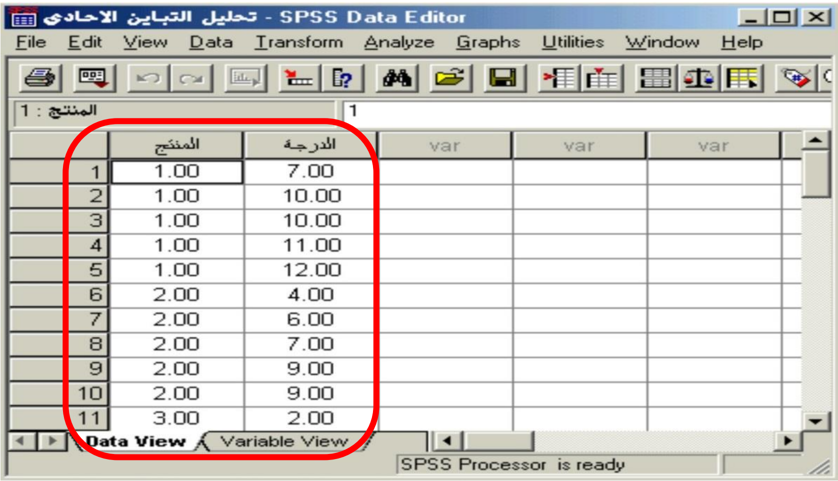
df2 \df1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66	8.65
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80	5.79
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56	4.54
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87	3.86
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44	3.43
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15	3.13
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94	2.92
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77	2.75
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65	2.63
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54	2.52
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46	2.44

حساب تحليل التباين الأحادي من خلال برنامج الـ SPSS

One Way Analysis of Variance

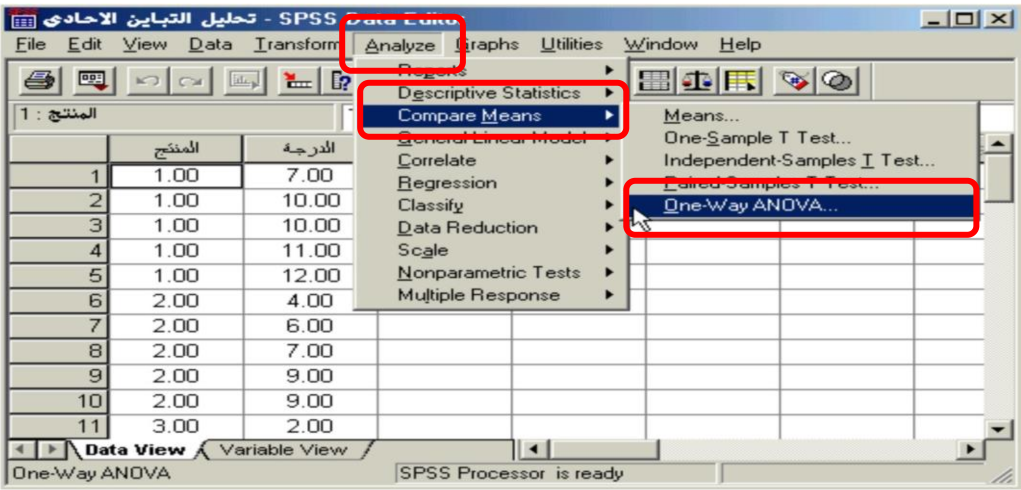
لغرض حساب قيمة تحليل التباين الأحادي One Way Analysis of Variance لنفس المثال السابق من خلال استخدام برنامج الـ SPSS نتبع الخطوات التالية :

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

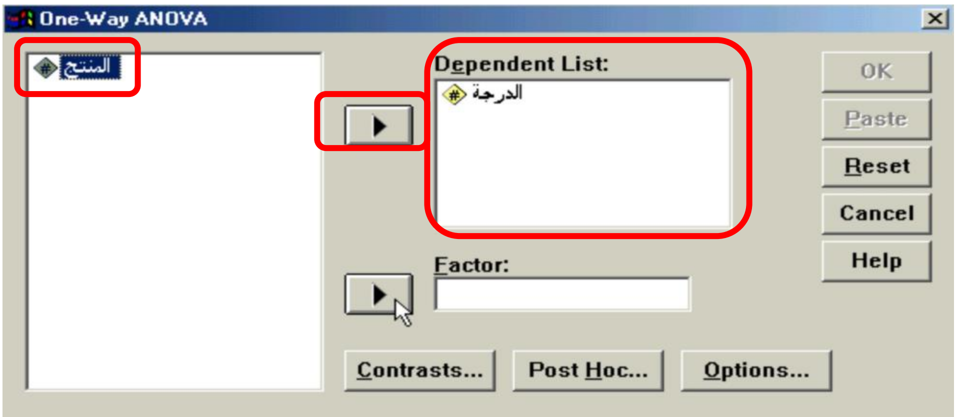


لاحظ أنه تم إدخال البيانات بطريقة مناسبة للتحليل الذي تم اختياره ، حيث أدخلت مستويات المتغير المستقل في عمود وأطلق عليه اسم "المنتج"، وأدخلت درجات التقييم للمنتج تحت عمود آخر أطلق عليه اسم "الدرجة" .

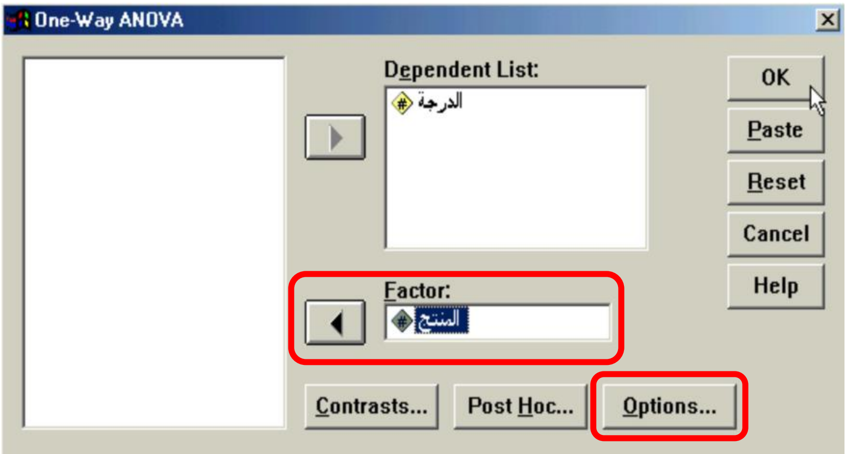
✓ من القائمة "تحليل" Analyze اختر الأمر "مقارنة المتوسطات" Compare Means فتظهر قائمة أوامر فرعية اختر منها "تحليل التباين الأحادي" One-Way ANOVA كالتالي :



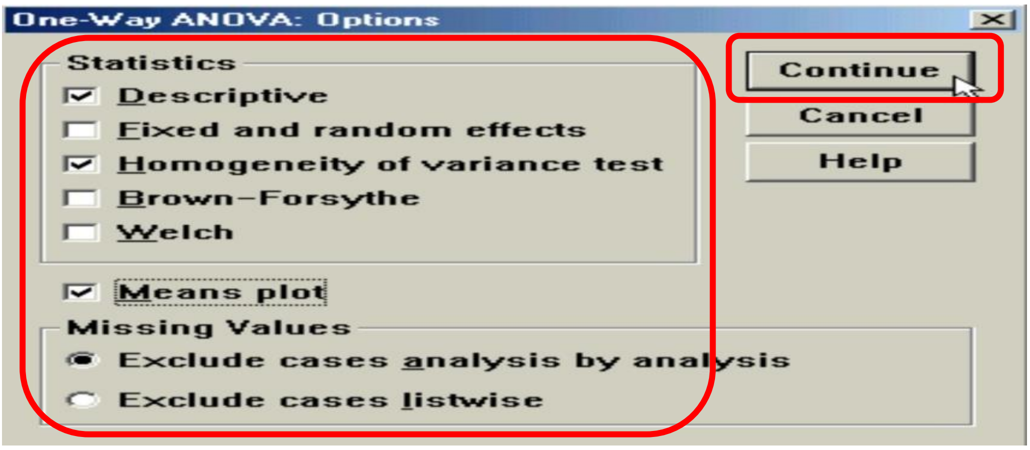
✓ بعد اختيار الأمر "تحليل التباين الأحادي" One-Way ANOVA سوف يظهر لك صندوق الحوار التالي :



✓ من قائمة المتغيرات في الجهة اليسرى من صندوق الحوار حدد المتغير المستقل والمتغير التابع المراد إجراء تحليل التباين الأحادي لها ، ونقلها إلى المستطيل الخاص بـ "المتغيرات التابعة" Dependent List من خلال النقر على السهم الذي يظهر مقابل المستطيل الخاص بـ "متغيرات الاختبار" ، ستلاحظ انتقال المتغير مباشرة في المستطيل "المتغيرات التابعة" Dependent List (في هذا المثال المتغير التابع هو "الدرجة") ، كرر نفس الإجراء مع المتغير المستقل Independent Variable وقم بنقله إلى المستطيل الخاص بـ "العامل" Factor (في هذا المثال المتغير المستقل هو "المنتج") كما يبدوا ذلك في الشكل التالي :



✓ أنقر على زر "خيارات" Options في الجهة السفلية اليمنى من صندوق الحوار السابق وذلك عند الرغبة في حساب الخصائص الأساسية للمتغيرات موضع الدراسة Statistics ، وعند الرغبة في عرض المتوسطات من خلال رسم بياني Means Plot ، وكذلك كيفية التعامل مع القيم المفقودة Missing Values ، وبعد الانتهاء من التعديل على هذا الصندوق الحواري أنقر على زر "استمرار" Continue .



✓ أنقر على زر "المقارنات البعدية المتعددة" Post Hoc في الجهة السفلية من صندوق الحوار السابق وذلك عند الرغبة في حساب المقارنات البعدية بين متوسطات المتغيرات موضع الدراسة والكشف عن مواقع الفروق وذلك في حالة كون قيمة F ذات دلالة إحصائية ، وهناك العديد من الأساليب الإحصائية المعدة لهذا الغرض اشهرها:

١) طريقة شيفيه Scheffe :

وتستخدم هذه الطريقة في إجراء جميع المقارنات بين الأوساط وهي الطريقة المفضلة في حالة كون حجوم الخلايا غير متساوية أو عند الرغبة في إجراء مقارنات معقدة كأن نقارن ثلاث مجتمعات بمجتمع واحد ، أو مجتمعين مقابل مجتمعين أو غيرها من مثل هذه المقارنات .

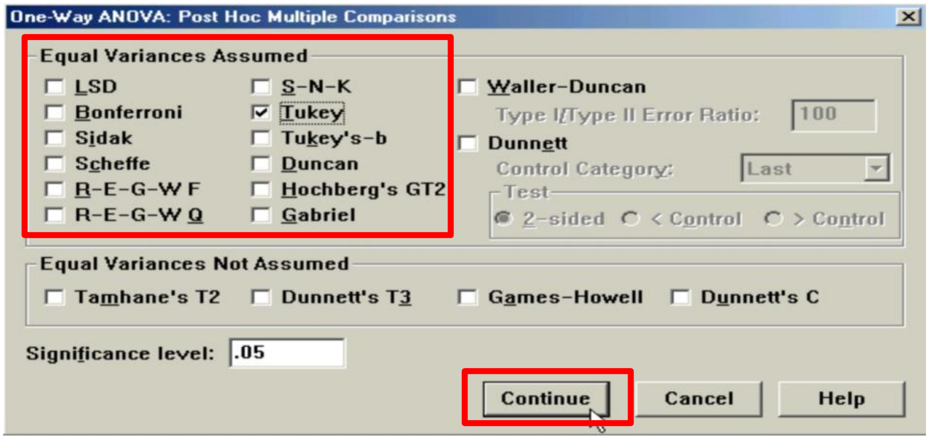
٢) طريقة توكي Tukey :

وتستخدم هذه الطريقة لمقارنة جميع الأزواج الممكنة للأوساط موضع الدراسة سواء كانت حجوم الخلايا متساوية أو غير متساوية (في حالة عدم تساوي حجوم الخلايا يستخدم الوسط التوافقي لحجم الخلية) ، ويعتبر هذا الاختبار أدق من اختبار شيفيه Scheffe لمقارنة أزواج الأوساط.

٣) طريقة نيومن-كولز Newman-Keuls (S-N-K) :

وتفيد هذه الطريقة في المقارنة بين أزواج الأوساط فقط، وهي تستند كما هي الحال في طريقة توكي على توزيع مدى ستيودنتايز Studentize range ، وهي طريقة جيدة وقوية للكشف عن الفروق بين الأوساط في حالة تساوي حجوم الخلايا أو عدم تساويها (في حالة عدم تساوي حجوم الخلايا يستخدم الوسط التوافقي لحجم الخلية كما هو الحال في اختبار توكي) ، ويعتبر هذا الاختبار (نيومن-كولز) أدق الاختبارات البعدية للكشف عن الفروق بين أزواج الأوساط ، يليه اختبار توكي ثم بعد ذلك اختبار شفيه .

وفي المثال الحالي تم اختيار طريقة توكي Tukey للمقارنة البعدية بين أزواج الأوساط (ويمكن اختيار أكثر من طريقة في وقت واحد) كما يبدوا ذلك في الشكل التالي:



وبعد الانتهاء من التعديل على هذا الصندوق الحواري أنقر على زر "استمرار" Continue ، وستنتقل إلى صندوق الحوار الرئيسي، ثم أنقر بعد ذلك على زر "موافق" OK في صندوق الحوار الرئيسي سيؤدي ذلك إلى تنفيذ الاختبار ، وستلاحظ ظهور النتائج في شاشة المخرجات كالتالي:

Oneway

Descriptives

VAR00001

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
1.00	5	10.0000	1.87083	.83666	7.6771	12.3229	7.00	12.00
2.00	5	7.0000	2.12132	.94868	4.3660	9.6340	4.00	9.00
3.00	5	4.0000	2.34521	1.04881	1.0880	6.9120	2.00	7.00
Total	15	7.0000	3.20713	.82808	5.2239	8.7761	2.00	12.00

Test of Homogeneity of Variances

VAR00001

Levene Statistic	df1	df2	Sig.
.686	2	12	.522

ANOVA

VAR00001

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	90.000	2	45.000	10.000	.003
Within Groups	54.000	12	4.500		
Total	144.000	14			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: VAR00001
Tukey HSD

(I) VAR00002	(J) VAR00002	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	3.00000	1.34164	.105	-.5793	6.5793
	3.00	6.00000*	1.34164	.002	2.4207	9.5793
2.00	1.00	-3.00000	1.34164	.105	-6.5793	.5793
	3.00	3.00000	1.34164	.105	-.5793	6.5793
3.00	1.00	-6.00000*	1.34164	.002	-9.5793	-2.4207
	2.00	-3.00000	1.34164	.105	-6.5793	.5793

*. The mean difference is significant at the 0.05 level.

نلاحظ أن برنامج الـ SPSS قام مباشرة بحساب الإحصاءات الأساسية للبيانات مثل المتوسط الحسابي والانحراف المعياري وغيرها من الإحصاءات ذات العلاقة.

ثم بعد ذلك قام البرنامج بحساب قيمة (F) للمتغيرات موضع الدراسة في الجدول المعنون بـ ANOVA ، ومن هذه النتائج نلاحظ أن قيمة (F) المحسوبة = 10 ، ودرجات الحرية df (2 ، 12) ، والقيمة الحرجة Sig.=0.003 ، وبما أن القيمة الحرجة لـ F Sig. في الجدول (0.003) أصغر من قيمة α = 0.05

فإننا 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 قمنا بحساب اختبار المقارنات البعدية Post Hoc Comparisons لتحديد هذه الفروق ، وقد أظهرت النتائج وجود فروق ذات دلالة إحصائية بين منتج (1) والذي متوسطه 10 ومنتج (3) والذي متوسطه 4 وذلك لصالح منتج (1).

المحاضرة الرابعة عشر

اختبار الفروض الإحصائية اللامعلمية

الطرق الإحصائية اللامعلمية Nonparametric Methods

تتطلب معظم التحليلات تحديد بعض الافتراضات أو الشروط حول المجتمع أو المجتمعات التي اختيرت منها العينة أو العينات. ففي كثير من الحالات يتم افتراض أن المجتمعات موضع الدراسة تتصف بالتالي:

- افتراض أن المجتمعات موضع الدراسة تتبع توزيعاً طبيعياً.
- افتراض أن تباينات هذه المجتمعات معلومة.
- افتراض أن تباينات هذه المجتمعات غير معلومة ولكنها متساوية.
- افتراض أن العينات المختارة مستقلة.

وحيث أنه توجد مواقف أو حالات كثيرة يكون من الصعب التأكد من تحقق هذه الافتراضات ، أو يكون هناك شك في تحققها ، وحيث أننا نواجه في كثير من الأحيان بيانات واقعية يصعب فيها التعرف على صيغة التوزيع الاحتمالي الذي تتبعه ، لذلك فقد طور الإحصائيين أساليباً وطرقاً إحصائية بديلة وهذه الطرق تتصف بالتالي:

- لا تتطلب افتراضات كثيرة.
- لا تتطلب معرفة صيغة التوزيع الاحتمالي للمجتمعات التي تختار منها العينات.

ومن هنا نشأت الطرق اللامعلمية.

وهذه الطرق بالإضافة إلى أنه يمكن استخدامها تحت شروط وافتراضات عامة فإنها غالباً لا تحتاج إلى مجهود في العمليات الحسابية ، كما انه يمكن التعامل معها لمتغيرات منفصلة ومتغيرات متصلة على السواء ، ولهذه الأسباب أصبحت الطرق اللامعلمية مرغوبة بكثرة.

طالما أن الاختبارات اللابارامترية (اللامعلمية) تتطلب هذا العدد القليل من الافتراضات حول البيانات ، فلماذا لا نستخدمها في كل الحالات ؟

إن الميزة السيئة للاختبارات اللابارامترية هي أنها غير جيدة عادة لإيجاد الفروقات عندما يكون هناك فروقات في المجتمع ، وعندما تكون الافتراضات من أجل الاختبارات البارامترية محققة ، بمعنى آخر الاختبارات اللابارامترية غير قوية كاختبارات تفترض توزيعا طبيعيا ، الاختبارات البارامترية ، ذلك بسبب أن الاختبارات اللابارامترية تتجاهل بعض المعلومات المتوفرة ، فعلى سبيل المثال في اختبار ويلكوكسن نستبدل قيم البيانات برتبها.

بشكل عام: إذا كانت افتراضات اختبار بارامتري مقنعة فيجب أن نستخدم اختبارات بارامترية للتحليل لأنها أكثر قوة ، وقد رأينا أن العديد من هذه الاختبارات يمكن أن تقوم بانتهاك الافتراض إلى حد معقول ، أي أنها قوية robust ، الإجرائيات اللابارامترية أكثر نفعًا من أجل العينات الصغيرة ، عندما يكون هناك ابتعاد ملموس عن الافتراضات المطلوبة ، وهي أيضًا مفيدة عندما يكون هناك قيم حدودية ، حيث أن الحالات المتطرفة لن تؤثر على النتائج بقدر التأثير الناتج في حال استخدمنا اختبارات معتمدة على إحصائية بسيطة كالمتوسط مثلا.

بعض الأمثلة للاختبارات الاحصائية اللامعلمية Some Examples of Nonparametric Methods

اختبار مان وتني Mann – Whitney U :

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 في البحوث عندما يكون المتغير التابع من المستوى الرتبى بدلاً من الدرجات الأصلية ، كما يمكن استخدام هذا الاختبار إذا كانت المتغيرات من المستوى الفترى أو النسبي ولكنها لا تفي بشروط اختبار النسبة التائية مثل عدم اعتدالية التوزيع أو اختلاف التباين بين المجموعتين اختلافاً كبيراً.

نفرض أن لدينا عينتين مستقلتين الأولى حجمها n_1 والثانية حجمها n_2 تم اختيارهما من مجتمعين متصلين ومتماثلين الأول متوسطه μ_1 والثاني متوسطه μ_2

والمطلوب:

اختبار الفروض التالية:

الفرض العدمي:

$$H_0 : \mu_1 = \mu_2$$

في حالة تساوي متوسطي المجتمعين يكون شكل الفرضية البديلة:

الفرض البديل :

$$H_A : \mu_1 \neq \mu_2$$

في حالة عدم تساوي متوسطي المجتمعين (وجود اختلاف معنوي بين متوسطي المجتمعين) يكون شكل الفرضية البديلة:

الفرض البديل :

$$\text{Or } H_A : \mu_1 < \mu_2$$
$$\text{Or } H_A : \mu_1 > \mu_2$$

حساب اختبار مان وتني Mann – Whitney U من خلال برنامج SPSS

مثال:

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

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

10	14	7	8	16
3	7	15	14	7

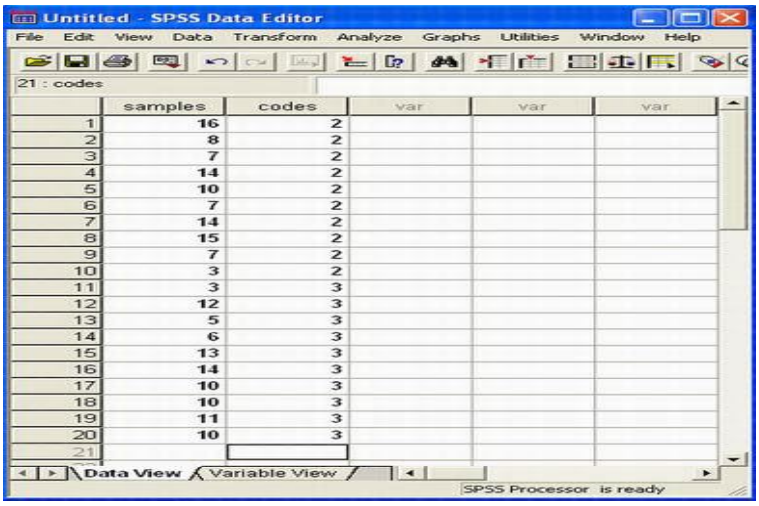
(٢) درجات مادة المحاسبة بكلية إدارة الأعمال جامعة الدمام:

13	6	5	12	3
10	11	10	10	14

المطلوب:

باستخدام اختبار مان – ويتني: اختبر هل هناك اختلاف في متوسط درجات مادة المحاسبة بين جامعة الملك فيصل وجامعة الدمام وذلك عند مستوى معنوية 5% .

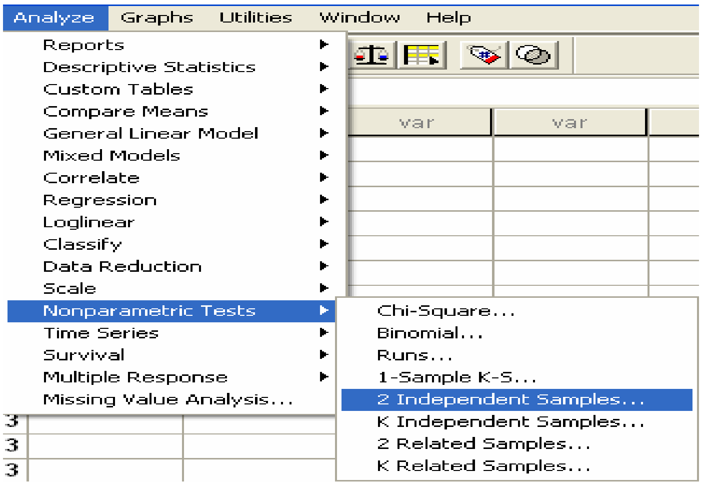
أولا: ندخل البيانات كالتالي:



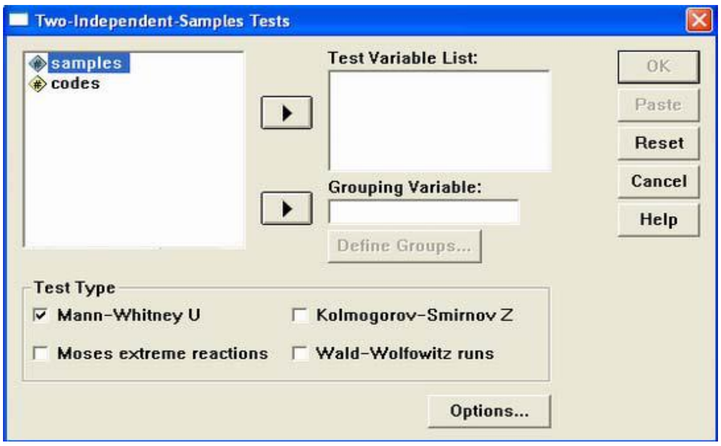
ملاحظة: في هذا التدريب نحن بصدد إدخال بيانات لعينات مستقلة ، لذا تم إدخال جميع المشاهدات في عمود ، والترميز الخاصة بالعينات في عمود آخر وذلك من خلال إعطاء الرقم (2) لبيانات العينة الأولى و (3) لبيانات العينة الثانية.

ثانيا: خطوات تنفيذ الاختبار:

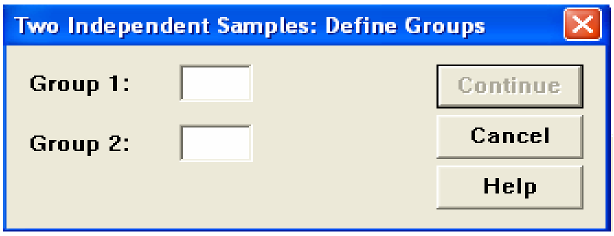
نضغط على قائمة Analyze ومن القائمة الفرعية لـ Nonparametric tests نختار 2 Independent Samples كما هو موضح بالشكل التالي:



سوف يظهر لنا المربع الحوارى التالي:



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:



- في خانة [Group 1] اكتب الرمز الخاص بالعينة الاولى (2)، وفي خانة [Group 2] اكتب الرمز الخاص بالعينة الثانية (3).
- ثم اضغط Continue للعودة الى المربع الحوارى السابق.
- ثم اضغط Ok سوف تظهر لك نافذة المخرجات الخاصة بهذا الاختبار.

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]

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها مع الأخذ في الاعتبار أن الأرقام قد تتغير ولديك المعرفة عن القبول والرفض.

يلاحظ من نتائج هذا الاختبار: أن قيمة P.Value تساوى 0.648 وهي أكبر من مستوى المعنوية 5% وبالتالي فإننا نقبل الفرض العدمى بأن متوسط درجات مادة المحاسبة في كلية إدارة الأعمال جامعة الملك فيصل يساوى متوسط درجات مادة المحاسبة في جامعة الدمام ، أي أن الفروق بين الجامعتين غير معنوية.

اختبار ويلكوكسون Wilcoxon Test :

استخدامه:

ويسمى باختبار اشارات الرتب Sign –rank ، ويستخدم هذا الاختبار في تحديد ما إذا كان هناك اختلاف أو فروق بين عينتين مرتبطتين فيما يتعلق بمتغير تابع معين ، ويعد بديلاً لابارامترياً لاختبار T لعينتين مرتبطين ، وتشتمل العينتان على نفس المجموعة من الأفراد يجرى عليهم قياس قبلي Pre test ، وقياس بعدى Post test وفى مثل هذه الحالة يكون لكل فرد من أفراد العينة درجتان أحداهما تمثل درجته في الاختبار القبلي والثانية تمثل درجته في الاختبار البعدي ، 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.

نفرض أن لدينا عينتين مترابطتين (غير مستقلتين)

والمطلوب:

اختبار الفروض التالية:

الفرض العدمي:

$H_0 : \mu_1 = \mu_2$

في حالة تساوي متوسطي العينتين يكون شكل الفرضية البديلة:

الفرض البديل :

$H_1 : \mu_A \neq \mu_2$

في حالة عدم تساوي متوسطي العينتين (وجود اختلاف معنوي بين متوسطي العينتين) يكون شكل الفرضية البديلة:

الفرض البديل :

Or $H_A : \mu_1 < \mu_2$

Or $H_A : \mu_1 > \mu_2$

حساب اختبار ويلكوكسون Wilcoxon Test من خلال برنامج SPSS

مثال:

تأثير ممارسة الرياضة على إنقاص الوزن:

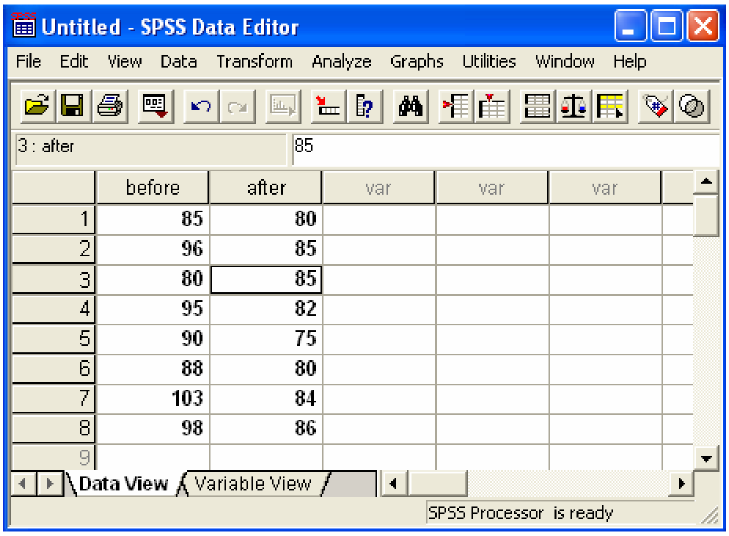
الوزن قبل ممارسة الرياضية	الوزن بعد ممارسة الرياضة
85	80
96	85
80	85
95	82
90	75
88	80
103	84
98	86

المطلوب:

اختبار هل هناك اختلاف معنوي في الوزن بسبب ممارسة الرياضة ، باستخدام اختبار ويلكوكسون Wilcoxon عند مستوى معنوية 5% .

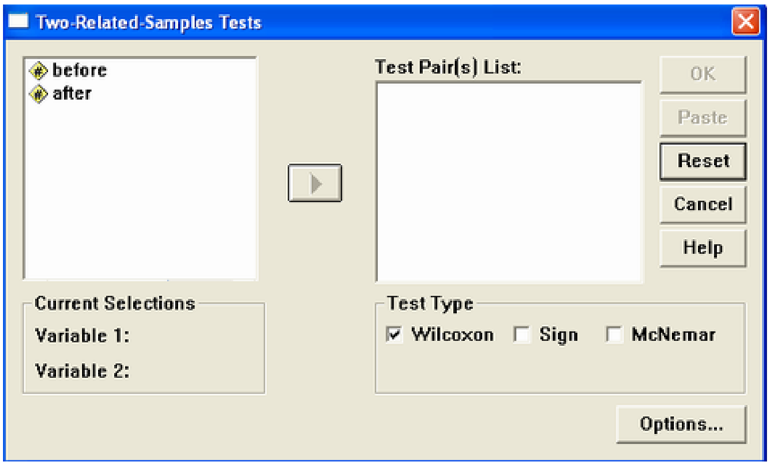
أولا: ندخل البيانات كالتالي:

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



ثانيا: خطوات تنفيذ الاختبار:

نضغط على **قائمة Analyze** ومن القائمة الفرعية لـ **Nonparametric tests** نختار **2 Related Samples** كما هو موضح بالشكل التالي:



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 في نفس المربع الحوارى الذى أمامك: أن الاختيار الافتراضي من جانب البرنامج هو اختبار ويلكوكسن ، وهو الاختبار الذى نريده لذا سنتركه كما هو ، اضغط **Ok** ستظهر لك نافذة المخرجات الخاصة بهذا الاختبار 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

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها مع الأخذ في الاعتبار أن الأرقام قد تتغير ولديك المعرفة عن القبول والرفض.

قام البرنامج بحساب الفروق في الوزن على أساس التالي:

الفرق = الوزن بعد ممارسة الرياضة – الوزن قبل ممارسة الرياضة

ويلاحظ أيضا: أن متوسط الرتب السالبة (**4.93**) أكبر من متوسط الرتب الموجبة (**1.5**) ، وهذا معناه **أن متوسط الوزن قبل ممارسة الرياضة أكبر من متوسط الوزن بعد ممارسة الرياضة** (إذا في غاية الأهمية أن نعرف الترتيب الذى استخدمه البرنامج للعينتين).

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

اختبار كروسكال واليس Kruskal-Wallis Test :

استخدامه:

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

نفرض أن لدينا k عينة عشوائية مستقلة الأولى حجمها n_1 والثانية حجمها n_2 وهكذا ، أي أن العينة الأخيرة حجمها n_k وأن هذه العينات تم اختيارها من مجتمعات متصلة عددها k ومتوسطاتها هي $\mu_1 ,\mu_2 ,...,\mu_k$ على التوالي.

والمطلوب اختبار فرض العدم:

$$H_0 : \mu_1 = \mu_2 = ... = \mu_k$$

أي جميع متوسطات المجتمعات متساوية

الفرض البديل:

ليست جميع متوسطات المجتمعات متساوية : H_1

حساب اختبار كروسكال واليس Kruskal-Wallis Test من خلال برنامج SPSS

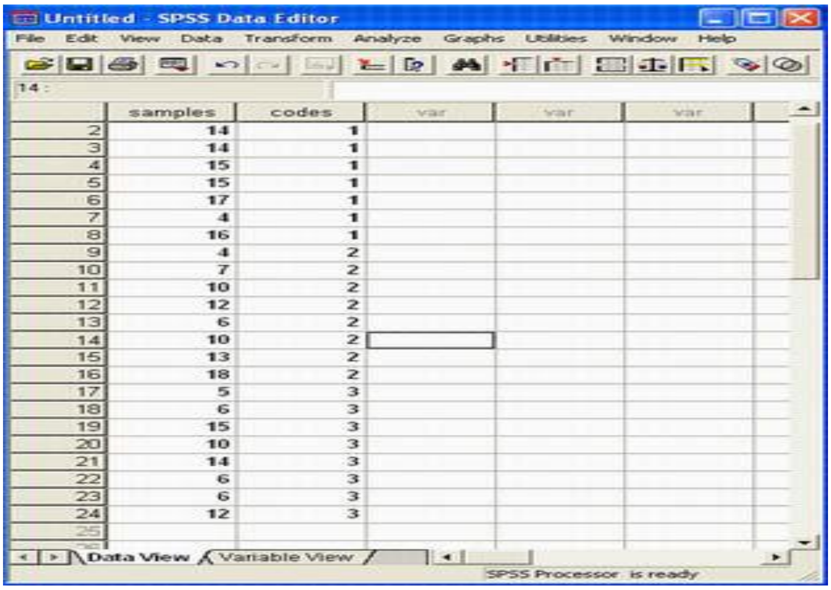
مثال: الجدول التالي يوضح درجات مجموعة من الطلاب في مادة الاقتصاد في ثلاث جامعات هي: جامعة الملك فيصل – جامعة الدمام – جامعة الملك سعود:

جامعة الملك فيصل	جامعة الدمام	جامعة الملك سعود
13	4	5
14	7	6
14	10	15
15	12	10
15	6	14
17	10	6
4	13	6
16	18	12

المطلوب: دراسة مدى وجود اختلاف بين مستوى الطلاب في الجامعات الثلاثة السابقة باستخدام اختبار كروسكال-والس، وذلك عند مستوى معنوية 5%

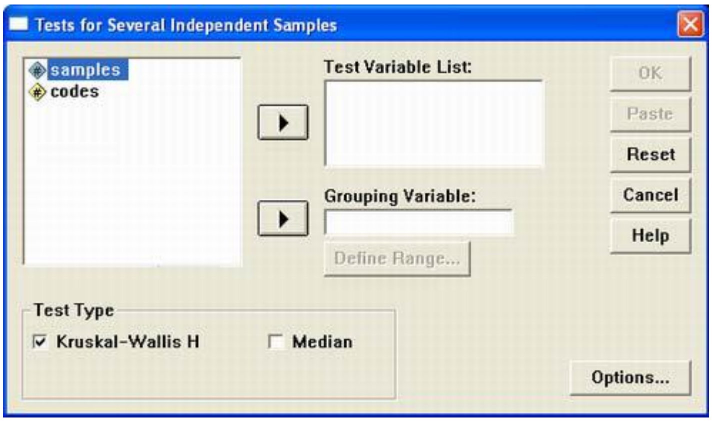
أولا: ندخل البيانات كالتالي:

حيث أننا بصدد ثلاث عينات مستقلة ، لذا تم إدخال قيم المشاهدات في عمود ، والرموز الخاصة بالعينات في عمود اخر ، حيث تم إعطاء الرمز (1) لبيانات العينة الأولى ، والرمز (2) لبيانات العينة الثانية ، والرمز رقم (3) لبيانات العينة الثالثة كما يلي:

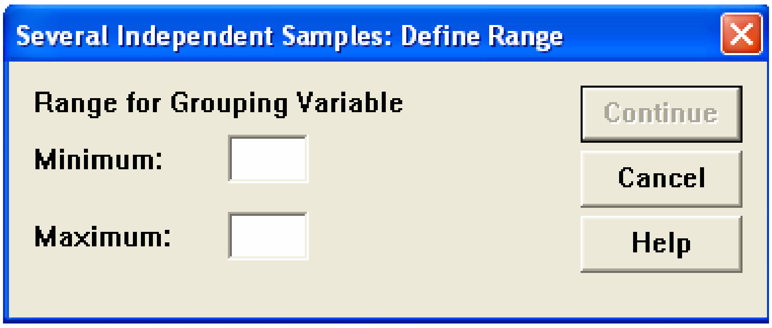


ثانيا: خطوات تنفيذ الاختبار:

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



- انقل المتغير samples الى المربع الذى بعنوان Test Variable List ثم انقل متغير الاكواد codes الى المربع الصغير الذى بعنوان Grouping Variable (لاحظ أن الاختيار الافتراضي من جانب البرنامج هو اختبار كروسكال – والس).
- اضغط Define Groups سوف يظهر مربع حوارى جديد كما يلى:



- فى خانة Minimum اكتب أصغر الرمز (1) ، وفى خانة Maximum اكتب أكبر الرمز (3) ، ثم اضغط Continue للعودة الى المربع الحوارى السابق.
- ثم اضغط Ok سوف تظهر لك نافذة المخرجات الخاصة بهذا الاختبار 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 الاختبار أن قيمة P.Value تساوى 0.095 وهى أكبر من مستوى المعنوية 5% .
وبالتالي فإننا نقبل الفرض العدمي بأن متوسط درجات مادة الاقتصاد في كلية إدارة الأعمال في الجامعات الثلاثة متساوي ، أي أن الفروق بين الجامعات الثلاثة غير معنوية.

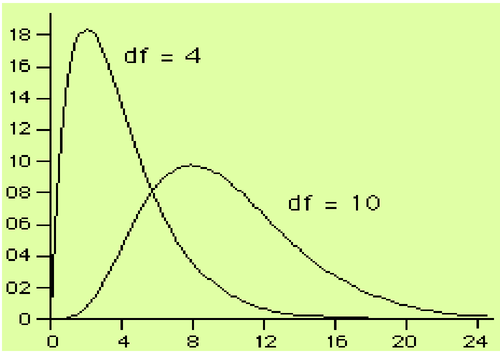
اختبارات الفروض باستخدام توزيع كاي تربيع (χ^2) Test Hypothesis Using Chi-Square Distribution

يعتبر توزيع كاي تربيع χ^2 من التوزيعات الاحتمالية الشائعة الاستخدام حيث توجد له تطبيقات عديدة بدرجة يمكن معها القول أنه يأتي في المرتبة الثانية بعد التوزيع المعتدل من حيث كثرة تطبيقاته.

توزيع كاي تربيع χ^2 :

يعتمد توزيع χ^2 مثل توزيع t اعتمادا كاملا على درجات الحرية ، وعلى الرغم من ذلك يوجد اختلاف رئيس بين التوزيعين حيث نجد أن توزيع t متماثل حول وسطه الحسابي (μ=0)، بينما يعتبر توزيع χ^2 توزيعا ملتويا جهة اليمين (التواء موجب) وخاصة عندما تكون درجات الحرية صغيرة ، وكلما زادت درجات الحرية كلما قل التواء التوزيع واقترب من التماثل.

شكل توزيع كاي تربيع χ^2 :



اختبار مربع كاي للاستقلالية (الإعتمادية) Testing of Independence

كاي تربيع للاستقلالية (Chi-Square test of independency) هو اختبار بسيط يقوم به الباحث لمعرفة ما إذا كان هناك علاقة بين شيئين أو متغيرينK يجرى هذا الاختبار عن طريقة مقارنة قيمة يحددها الباحث مسبقا تعرف بمستوى المعنوية (الفا) بالقيمة المسماة p-Value تحسب من البيانات المتوفرة ، حيث سيتضح عن طريق المقارنة بين القيمتين ما إذا كانت هنالك علاقة بين الاثنين أم لا.

فرضية العدم (Null hypothesis): لا توجد أي علاقة بين المتغيرين ويرمز لهذه الفرضية H_0 والذي يتم افتراض صحته عند القيام بالاختبار.

عند القيام بالاختبار لمتغيرين ، تكتب هذه الفرضية بهذه الطريقة: V_1 مستقل عن V_2 ، حيث V_1 و V_2 تمثل المتغيرين تحت الدراسة ، ويمكن كتابة فرض العدم الإحصائي بالشكل التالي:

H_0: V_1 is independent of V_2

الفرض البديل (Alternative hypothesis): توجد علاقة بين المتغيرين تحت الدراسة ويرمز لهذه الفرضية H_A وتكتب الطريقة التالية: V_1 **غير مستقل** أو يتبع لـ V_2 ، حيث V_1 و V_2 المتغيرين تحت الدراسة ، ويمكن كتابة الفرض البديل بالشكل التالي:

$$H_A: V_1 \text{ is dependent on } V_2$$

حساب اختبار مربع كاي (كا٢) للاستقلالية Test of Independence -Chi Squre من خلال برنامج SPSS

مثال:

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 $\alpha = 0.05$ ؟

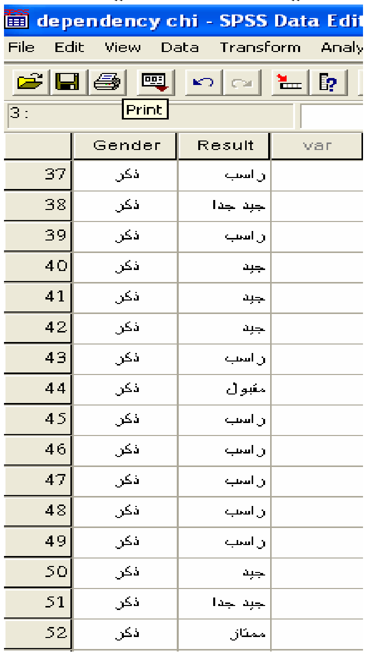
الحل:

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

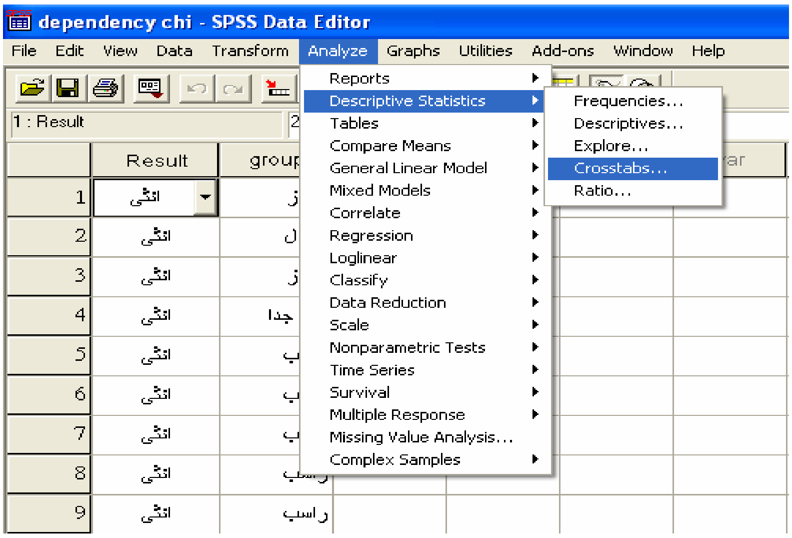
الفرضية البديلة: تقدير الطالب يعتمد على جنسه (توجد علاقة بين جنس الطالب وتقديره).

ثم نقوم بتعريف متغيرين نوعيين هما (*Result*) و (*Gender*) في شاشة تعريف المتغيرات بحيث يكون كود متغير (*Result*) هو (0 = راسب ، 1=مقبول، 2=جيد، 3= جيد جدًا، 4=ممتاز) وكود المتغير (*Gender*) هو (1=ذكر، 2=انثى)

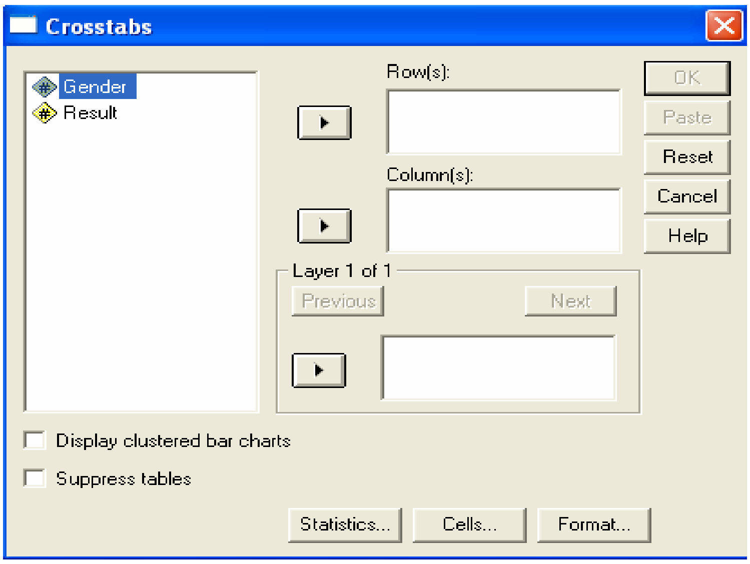
ندخل البيانات كما في الشكل التالي:



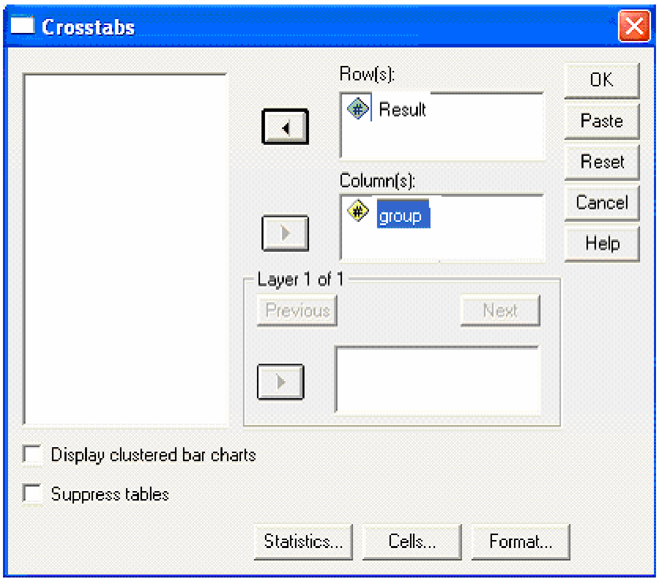
من قائمة التحليل **Analyze** نختار القائمة الفرعية للإحصاءات الوصفية **Descriptive Statistics** ومن ثم نختار الأمر **Cross tabs** كما في الشكل التالي:



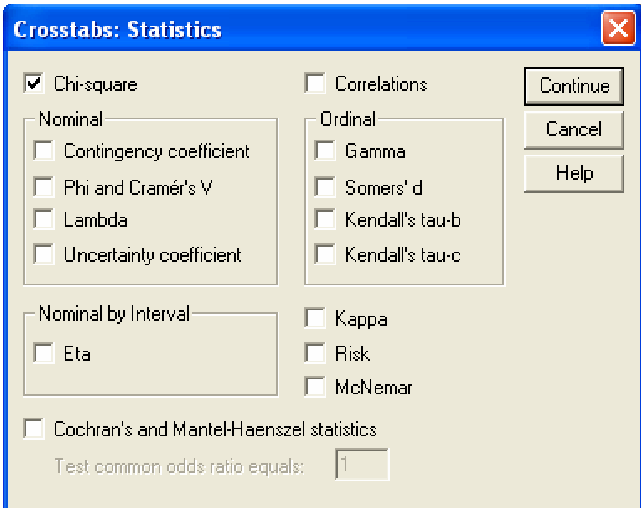
يظهر المربع الحواري التالي:



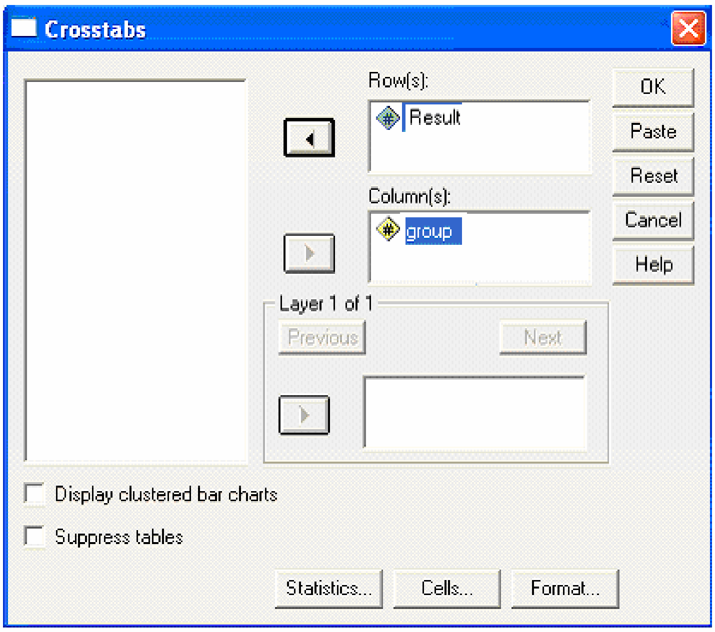
ننقل المتغير **Result** لخانة الصفوف **Rows** والمتغير **Gender** لخانة الأعمدة **Columns** باستخدام الأسهم .



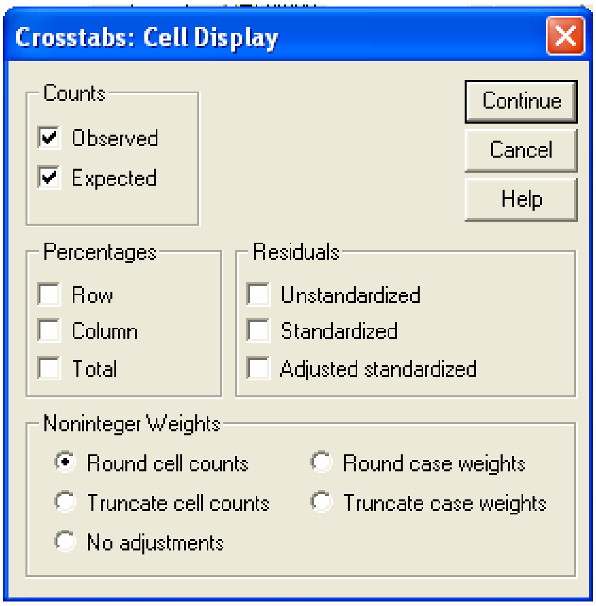
ومن ثم نضغط على **Statistics** للحصول على المربع الحواري التالي:



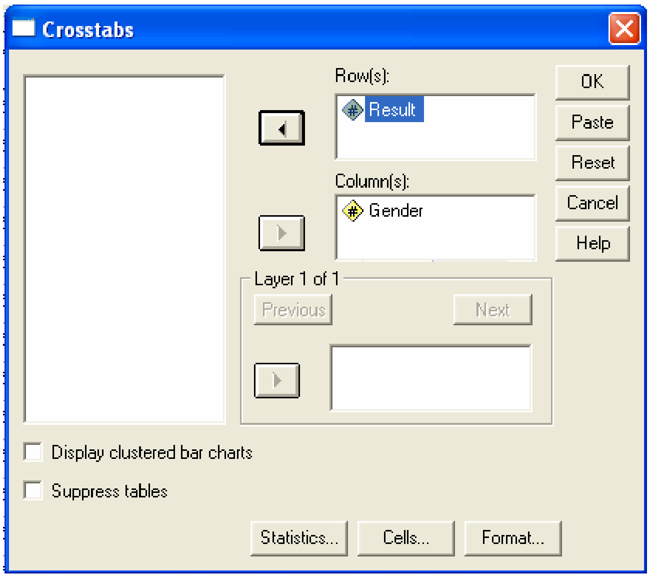
نضع علامة على خانة اختبار مربع كاي **Chi-Square** لحساب اختبار الاستقلالية ومن ثم نضغط على **Continue** للعودة للمربع الحواري السابق:



لإظهار 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:

الجدول الأول: 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

توقع 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 72 ، عدد الذكور 37 (منهم 12 راسب وقيمتها المتوقعة 9.76 ، 5 مقبول وقيمتها المتوقعة 6.68 ، 9 جيد وقيمتها المتوقعة 8.74 ، 5 جيد جدا وقيمتها المتوقعة 6.17 ، و 6 ممتاز وقيمتها المتوقعة 5.65) والاناث 35 (منهم 7 راسب وقيمتها المتوقعة 9.24 ، 8 مقبول وقيمتها المتوقعة 6.32 ، 8 جيد وقيمتها المتوقعة 8.26 ، 7 جيد جدا وقيمتها المتوقعة 5.83 ، و 5 ممتاز وقيمتها المتوقعة 5.35)

الجدول الثالث يبين نتيجة اختبار مربع كاي كالتالي:

قيمة الاختبار

درجة الحرية

مستوى دلالة الاختبار

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 2.437 بدرجة حرية مقادرها 4 يتبين لنا من الجدول أن أقل قيمة لمستوى الدلالة هي Asymp. Sig. (2-sided) = 0.656 وهي اكبر من مستوى الدلالة α = 0.05 وبالتالي لا نستطيع رفض الفرضية الصفرية أي أن تقدير الطالب لا يعتمد على جنسه.

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها مع الأخذ في الاعتبار أن الأرقام قد تتغير ولديك المعرفة عن القبول والرفض.

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 كاي** 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.

ويستخدم لاختبار ما إذا كانت عينتان لهما نفس التوزيع ويعتمد إجراء هذا الاختبار على دالة الاحتمال التجميعي للتكرار المشاهد والمتوقع وبذلك يدور الفرض العدمي والبديل حول هاتين الدالتين وهو كالتالي :

$$H_0 : F_n(X) = F_0(X)$$

$$H_A : F_n(X) \neq F_0(X)$$

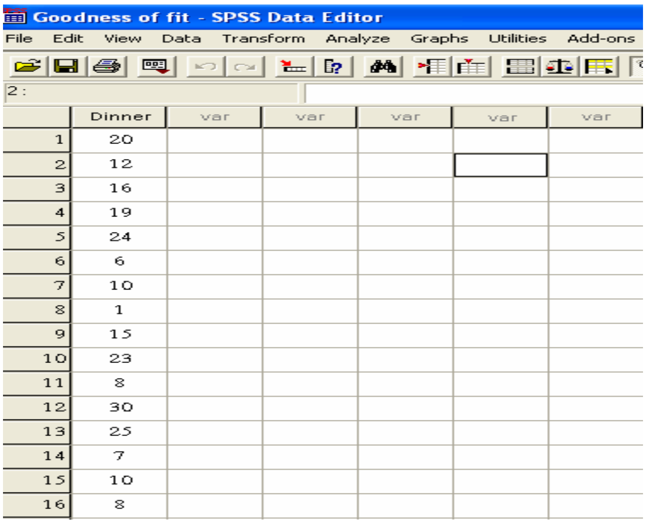
$$F_n(X) < F_0(X)$$

$$F_n(X) > F_0(X)$$

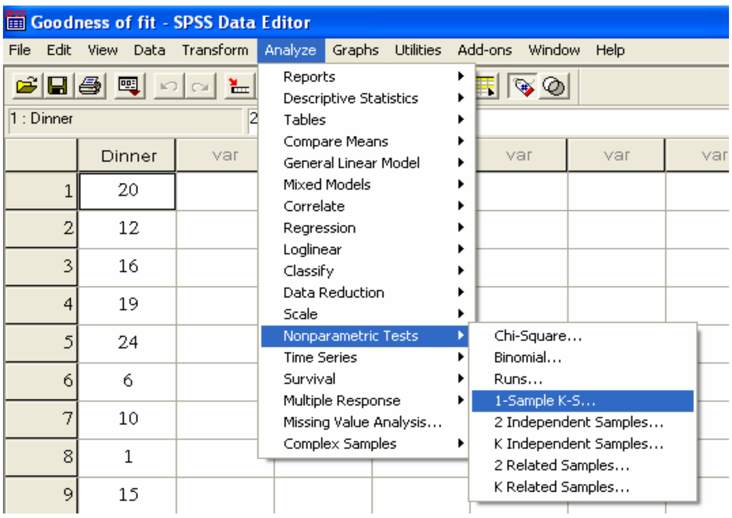
حساب اختبار كولومجروف سيمرنوف لجودة التوافيق

Goodness of Fit Test - Kolmogorov-Smirnov من خلال برنامج SPSS

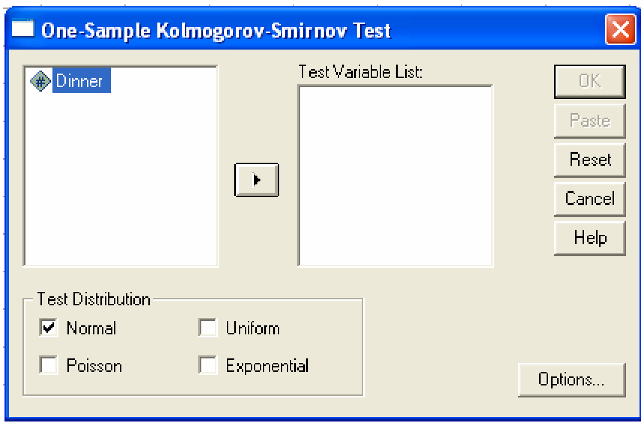
ندخل البيانات في متغير نسميه Dinner كما في الشكل التالي:



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 <u>Normal</u> أو بواسون <u>Poisson</u> أو منتظم <u>Uniform</u> أو أسي <u>Exponential</u> فنختار التوزيع الطبيعي كما في الشكل أعلاه ونضغط <u>Ok</u> 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.

حجم العينة
متوسط البيانات
الانحراف المعياري للبيانات
أكبر فرق بين البيانات ودالة التوزيع الاحتمالية
قيمة اختبار جودة المطابقة
مستوى دلالة الاختبار

مهمة جدا معرفة هذه النتائج لذلك لابد وأن تكون ملم بها مع الأخذ في الاعتبار أن الأرقام قد تتغير ولديك المعرفة عن القبول والرفض.

تبين النتائج أعلاه أن متوسط عدد الزبائن هو <u>15.26</u> بانحراف معياري قدره <u>6.782</u> وأن قيمة اختبار كولموجروف سميرنوف لجودة المطابقة هو <u>0.573</u>

<u>القرار:</u>

يبين الجدول السابق أن قيمة مستوى دلالة الاختبار هي <u>*Asymp. Sig. (2-tailed) =0.898*</u> وهي اكبر من مستوى دلالة الفرضية الصفرية <u>α = 0.05</u> وبالتالي <u>نقبل الفرضية الصفرية</u>، أي أن <u>البيانات تتبع التوزيع الطبيعي</u> وبالتالي نستنتج ان البيانات تتبع التوزيع الطبيعي بمتوسط قدره <u>15.26</u> وانحراف معياري <u>6.782</u> أي <u>X : N (15.26 , 6.782)</u>

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

تم بحمد الله ..

إن أصبت فتوفيق من الله وإن حصل هناك خطأ فمن نفسي والشيطان

وفقنا الله وإياكم ، ودعواتكم الطيبة

أخوكم / شيءٌ آخر (أبو فيصل)