ورشة عمل مادة

WWW.OKFU.ORG

التحليل الاحصائي

الدفعة الماسية

ملخص التحليل الاحصائي
(١٤٣٦هـ)

د/احمد فرحان .

اعداد اخوكم (MOHaMED_KFU)

دعواتكم أسال الله لي ولكم التوفيق والنجاح

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.

$A = \{a, b, c, d \dots\}$

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:

$b \in A$ $\qquad$ $\cancel{A \in a}$

$f \notin A$

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

العمليات على المجموعات

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.

علاقة الانتماء بين عنصر ومجموعة فقط

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

$A = \{ X : X$ دولة عضو دول مجلس التعاون $\}$

حيث

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 وتكون 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:** $\emptyset$ ، $\{ \}$

وهي مجموعة الأعداد الصحيحة التي بين العددين ١،٠ 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 المجموعه 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:** U الشاملة S

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 $B \subset A$**

$\not\subset$ غير جزئية

$\subset$ جزئية

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.**

ملاحظة علاقة جزئية تستخدم لتوضيح العلاقة بين مجموعة ومجموعة

٦- **تساوي المجموعات:**

تكون المجموعتان A ، B متساويتان إذا كانت

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

مثال:

$x^2 = 1$

$x = \pm 1$

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

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

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

الحل:

1) $A = B$

2) $A \equiv B$ تكافؤ

> لاحظ أن في المجموعتين في الرقم واحد العناصر متساوية بينما في المجموعتين في الرقم 2 العناصر مختلفة ولكن عددها واحد اذا 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.

مجموع العناصر بدون تكرار

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}

المشترك فقط

٣- المكملة أو المتممة:

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

أي أن

$\overline{A} = S - A$

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}

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

٤- الفرق (الطرح) نأخذ العناصر الموجودة في المجموعة الأولى وليست الموجودة في الثانية

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) $\overline{A}$ 5) $\overline{B}$

الحل: ={1,2,3,x,y,4,5,w} {3,x} {1,2,y} {4,5,w,z} {1,2,y,z}

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

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: ٧ عدد طبيعي اصغر من x} = A {1, 2, 3, 4, 5, 6}

ii. B={x: ٢ عدد طبيعي زوجي يقبل القسمة على x} = {2, 4, 6, ...}

iii. C={y: h و c حرف من حروف الهجاء المحصورين بين y} = {d, e, f, g}

iv. D={x: ١٧ عدد طبيعي فردي اصغر من x} = {1, 3, 5, 7, 9, 11, 13, 15}

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

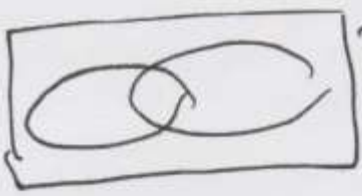
أشكال فن (VIN Figures).

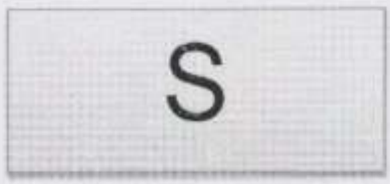
يمكن تمثيل المجموعات والعمليات المختلفة عليها من خلال استعمال اشكال هندسية تسمى أشكال فن وذلك وفق ما يلي:

١- المجموعة الشاملة:

شاملة

تمثل المجموعات الكلية بمستطيل ويرمز لها بالرمز S

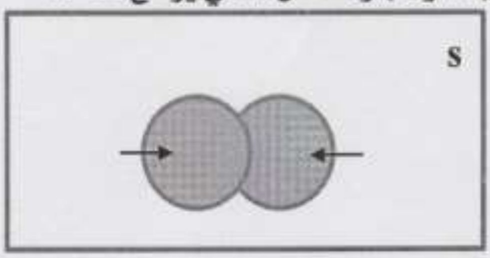




S

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 أو B) والشكل التالي يوضح ذلك:



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

وبشكل عام لأي n حادثة An .... ,A1 , A2, A3 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 ∪ 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 :

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

توزيع الاتحاد على التقاطع

ويعني ذلك توزيع الاتحاد على التقاطع.

- وكذلك هناك خاصية التبديل والتي تعني أن :

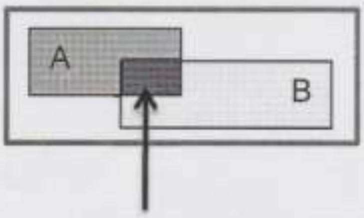
$$(A \cup B) = (B \cup A)$$

$A \cap (B \cup C)$

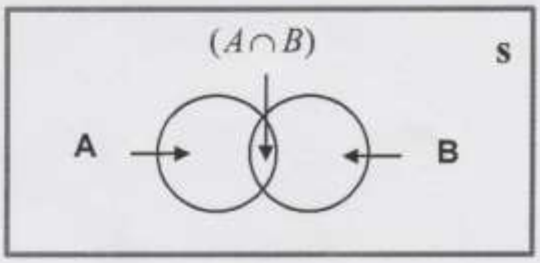
$= (A \cap B) \cup (A \cap C)$

٢- 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 فن يكون الجزء المحدد بـ A and B هو الذي يمثل تقاطع الحادثتين :



$(A \cap B)$



شكل فن لتمثيل تقاطع حادثتين A و B

وبشكل عام لأي n حادثة An .... ,A1 , A2, A3 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 ∩ 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 :

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

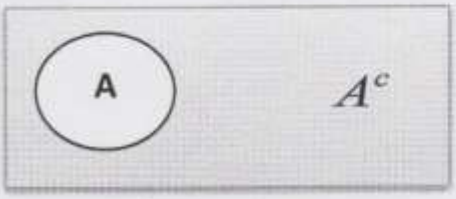
ويعني ذلك توزيع التقاطع على الاتحاد.

- 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 فن فإن الجزء المظلل يمثل الحادثة المتممة :



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

شكل فن 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}

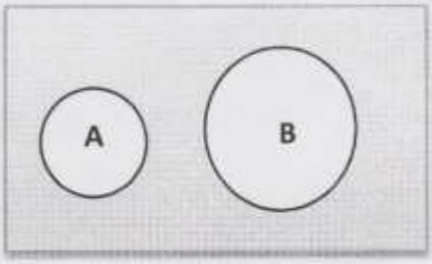
$\overline{B}$ ={4, 5, 7, 9, 10, 13, 15, 17, 18, 19, 20}

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 فن فإن الحادثتان المنفصلتان يمثلان بالشكل التالي :**



$$A\cap B=\phi$$

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

شكل فن 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 $(\overline{A}\cup\overline{B})$
- 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:

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.

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

- 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.

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

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

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

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

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

المحاضرة (٢)

نظرية 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 ككلمة مرادفة لبعض الكلمات الأخرى مثل: ممكن، غالبا، مؤكد، مستحيل ...

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

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

١. التجربة العشوائية 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.

٢. الحادثة والفراغ العيني:

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

افترض أننا نقوم بإجراء تجربة ما كرمي زهرة النرد مثلا 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 حادثا (Event) ومجموعة جميع الحالات الممكنة الظهور تسمى بالفراغ العيني (Sample Space) ويلاحظ أن الحادث قد يكون حالة او اكثر من الفراغ العيني.

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

٣. أنواع الحوادث :

صح

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

أ. الحوادث المتنافية (Mutually Exclusive Events)

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

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

ب. الحوادث المستقلة (Independent Events)

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

جـ. الحوادث الشاملة (Exhaustive Events)

تسمى الحوادث A ، B ، C ... حوادث شاملة في تجربة ما إذا كان لابد من حدوث إحداها عند إجراء التجربة.

مثال:

رمي حجر نرد مرد واحدة ، أحسب التالي:

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

- احتمال الحصول على رقم ٥ $= \frac{1}{6}$
- احتمال الحصول على رقم زوجي $= \frac{3}{6}$
- احتمال الحصول على رقم أكبر من ٢ $= \frac{4}{6}$
- احتمال الحصول على رقم أقل من ٧ $= \frac{6}{6} = 1$
- احتمال الحصول على رقم ٧ $\frac{0}{6} = 0$

الحل:

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

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

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

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

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

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

$\cup$ أو إتحاد

$\cap$ تقاطع و

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

مثال:

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

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

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

- أن يكون أعزبا $\frac{23}{50}$
- أن يكون متزوجا $\frac{17}{50}$
- أن يكون من القسم الأول $\frac{12}{50}$
- أن يكون من القسم الأول أو الثاني

$= \frac{12}{50} + \frac{22}{50} = \frac{34}{50}$

- أن يكون من القسم الأول وأعزب

$= \frac{12}{50} + \frac{23}{50} - \frac{5}{50} = \frac{30}{50}$

$\Leftarrow$ أن يكون من القسم الأول و أعزب $= \frac{5}{50}$

$\Leftarrow$ أن يكون من القسم الأول و الثاني $= \frac{0}{50} = 0$

أ. في حالة كون الحوادث متنافية

- إذا كانت الحوادث A1 , A2 , A3 .... حوادث متنافية بمعنى أن حدوث أحدها يؤدي إلى استحالة حدوث أي من الحوادث الأخرى وبالتالي فإن احتمال حدوث هذه الحوادث معا يكون معدوما فإن احتمال وقوع أي حادث من الحوادث المتنافية يساوي مجموع احتمالات وقوع هذه الحوادث

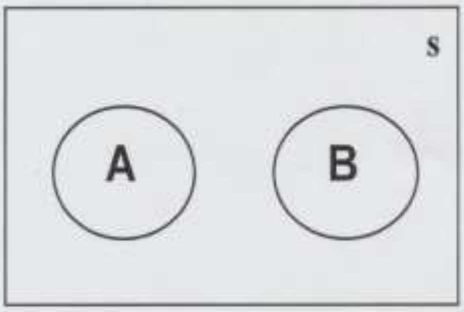
فإذا كان A ، B حادثين متنافيين كما في الشكل (١)

فإن $P_{(A \text{ أو } B)} = P(A) + P(B)$

يرمز أيضا لاحتمال وقوع أحد الحادثين بالرمز

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

الشكل (١)



حوادث متنافية

مثال:

رمي حجر نرد مرة واحدة ، احسب: احتمال الحصول على رقم ٥ أو ٦

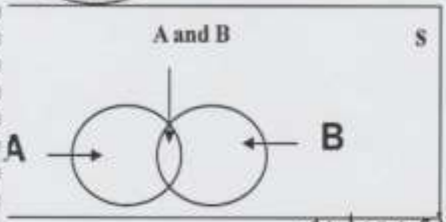
الحل:

$\frac{1}{6} + \frac{1}{6} = \frac{2}{6}$

ب- في حالة كون الحوادث غير متنافية

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

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



حوادث غير متنافية

مثال:-

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

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

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

- احتمال أن يكون العامل من القسم الأول أو الثاني. $P(A \cup B) = \frac{12}{50} + \frac{22}{50} =$
- احتمال أن يكون العامل متزوجا أو من القسم الأول $\frac{27}{50} + \frac{12}{50} - \frac{7}{50} =$
- احتمال أن يكون العامل من القسم الثالث أو أعزب $\frac{16}{50} + \frac{23}{50} - \frac{10}{50} =$

الاحتـــمال الشرطي Conditional Probability

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

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

مثال:

A تقاطع B A

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:

$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{.32}{.64}$

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

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الذي يعمل به: اختير عامل من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

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

① $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{7}{50}}{\frac{27}{50}} = .25$

② $= \frac{\frac{10}{50}}{\frac{16}{50}} = .62$

ضـــرب الاحتـــمالات

إن احتمال حدوث حادثين مستقلين أو أكثر معا يساوي حاصل ضرب احتمال حدوث كل واحد من هذه الحوادث ببعضها بعضا حيث أن نتيجة السحب إذا تكرر أكثر من مرة الأول لا تؤثر على نتيجة السحب الثاني وهذا ما يسمى سحب بإرجاع أو إحلال أو إعادة.فإذا كان لدينا الحادثين المستقلين A1 و A2 فإن احتمال حدوثهما معا هو:

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

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

$= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$P(A \cap A) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

مثال:

إذا رمينا قطعة نقود مرة واحدة ، احسب الاحتمالات التالية:

- أن تكون الأولى صورة (B) والثانية كتابة (A)
- أن تكون كلتاهما صورة.

مثال:

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

اختير عاملان من الجدول السابق بطريقة عشوائية،

احسب الاحتمالات التالية

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

١. احتمال أن يكون العاملان من القسم الأول؟
٢. احتمال أن يكون العاملان متزوجان؟
٣. احتمال أن يكون للعاملين نفس الحالة الاجتماعية؟
٤. احتمال أن يكون العاملان من القسم نفسه؟

الحل:

١ـ احتمال أن يكون العاملان من القسم الأول أي العامل الأول من القسم الأول والثاني من القسم الأول

$= P(A) \cdot P(A)$

$\frac{12}{50} \cdot \frac{12}{50} = 0{,}057$

٢ـ احتمال أن يكون العاملان متزوجان ،

$= \frac{27}{50} \cdot \frac{27}{50} = 0{,}29$

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

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

$$= P(A_1 A_2) + P(B_1 B_2)$$

$$= P(A_1 \times A_2) + P(B_1 \times B_2)$$

$$= \left[\frac{27}{50} \times \frac{27}{50}\right] + \left[\frac{23}{50} \times \frac{23}{50}\right] = 0.5032$$

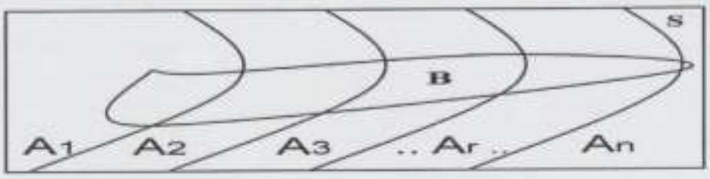
٤ـ احتمال أن يكون العاملان من القسم نفسه يعني أن يكون العاملان كلاهما من القسم الأول (الحادثة A) أو أن يكون كلاهما من القسم الثاني (الحادثة B) أو أن يكون كلاهما من القسم الثالث (الحادثة C) فإن:

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

نظريــــة 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:ـ

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:

١ـ أن تكون القطعة المعيبة من إنتاج الآلة الأولى؟

٢ـ أن تكون القطعة المعيبة من إنتاج الآلة الثانية؟

مصنع

آلة ٣   آلة ٢   آلة ١

٪٤٥   ٪٣٥   ٪٢٠

٪٣   ٪٢,٥   ٪٢   نسبة المعيب

$$=\frac{2\% \cdot 20\%}{(2\%\cdot 20\%)+(2,5\%\cdot 35\%)+(3\%\cdot 45\%)}$$

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

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

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

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

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

المحاضرة (٣)
المتغيرات العشوائية

المتغير العشوائي Random Variable:

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

وينقسم المتغير العشوائي إلى قسمين هما:

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

أولا : المتغيرات العشوائية المتقطعة (المنفصلة): دائماً يسبقها كلمة عدد

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

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

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

التوزيع الاحتمالي للمتغير العشوائي المنفصل:

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

مثـال :

إذا كانت نسبة مبيعات أحد المراكز التجارية من التفاح الأمريكي 0.60، بينما يكون نسبة مبيعاته من الأنواع الأخرى للتفاح 0.40، اشترى أحد العملاء عبوتين . والمطلوب: كون فراغ العينة. إذا عرف المتغير العشوائي بأنه عدد العبوات المشتراة من التفاح الأمريكي، فأوجد الآتي:

- التوزيع الاحتمالي للمتغير العشوائي .

جدول التوزيع الاحتمالي لعدد العبوات المشتراة من التفاح الأمريكي

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

الوسط الحسابي والتباين للمتغير العشوائي المنفصل:

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

$$\boxed{\mu = \sum x_i \; f(x_i)}$$

وأما التباين ويرمز له بالرمز (سيجما)، فيحسب بتطبيق المعادلة التالية:

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

مثـال:

في المثال السابق احسب ما يلي:

الوسط $\mu = 0 \cdot 0.16 + 1 \cdot 0.48 + 2 \cdot 0.36 = 1.20$

للتباين $\sigma^2 = 0^2 \cdot 0.16 + 1^2 \cdot 0.48 + 2^2 \cdot 0.36 - 1.20^2$

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

باستخدام الحاسبة

الحـــــــــــل:

- الوسط الحسابي لعدد العبوات من النوع الأمريكي:

MODE → 3 → 1 → البيانات → AC

Shift → 1 → 4 → الوسط = 2 =
→ الانحراف = 3 = → تربيع = تباين

- لحساب الوسط الحسابي والانحراف المعياري يتم استخدام المعادلة الخاصة بذلك وهذا يتطلب تكوين جدول يشمل المجاميع التالية وذلك كما يلي:

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

إذا الوسط الحسابي هو:

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

ولحساب الانحراف المعياري يجب أولا حساب التباين وهو:

$$\sigma^2 = \sum x_i^2 \ f(x_i) - \mu^2 = 1.92 - (1.20)^2 = 0.48$$

إذا الانحراف المعياري قيمته هي:

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

معامل الاختلاف النسبي هو: = $\frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times ١٠٠$

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

ثانيا : المتغيرات العشوائية المستمرة Continuous Random Variables

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

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

الوسط الحسابي والتباين للمتغير العشوائي المستمر:

معادلة الوسط والتباين يمكن كتابها كما يلي:

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

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

التوزيعات الاحتمالية للمتغيرات المنفصلة:-

أ- توزيع ذي الحدين:

يستخدم هذا التوزيع في الحالات التي يكون للظاهرة محل الدراسة نتيجتان فقط متنافيتان، النتيجة محل الاهتمام وتسمى بحالة النجاح(p)، والأخرى تسمى بحالة الفشل(q)، 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

- النتيجة الأخرى · حالة فشل · وتتم باحتمال ثابت أيضا

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

- هناك ناتجان ممكنان فقط ومتنافيان لكل محاولة.
- المحاولات وعددها n مستقلة عن بعضها البعض.
- احتمال وقوع الحدث المعين في كل محاولة (النجاح) P ثابت ولا يتغير من محاولة لأخرى.
- **فبالتالي يمكن حساب الاحتمال من خلال المعادلة التالية:**

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

- حيث n! (وتقرأ · مضروب n) = n. (n-1). (n-2)...3.2.1 ،0! =1.
- ويكون متوسط توزيع ذي الحدين

الحد ابو 1

$$\mu = np$$

التي

- وانحراف المعياري

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

الانحراف $= \sqrt{n.p.q}$ → q

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.

تمرين:

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

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

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

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

تمرين

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

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

2 $P = ,80 \qquad q = ,20$

$p(x=3) = C^{n}_{x} \cdot p^{x} \cdot q^{n-x}$

$\boxed{\text{shift} \div} \rightarrow C^{4}_{3} \cdot ,80^{3} \cdot ,20^{4-3} = ,40$

$C^{4}_{1} \cdot ,80^{1} \cdot ,20^{4-1} = ,025$

3 $p(x=1) =$

4 $p(x \geq 2) = p(x=2) + p(x=3) + p(x=4)$

$C^{4}_{2} \cdot ,80^{2} \cdot ,20^{2} + ,40 + C^{4}_{4} \cdot ,80^{4} \cdot ,20^{4-4} = ,96$

5 الوسط $= 4 \cdot ,80 = 3,2$

6 الانحراف $= \sqrt{4 \cdot ,80 \cdot ,20} = ,8$

تمارين مراجعة

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

أ- $(A\cup B)\cap(A\cup C)$	ب- $(A\cup B)\cup(A\cup C)$
ج- $(A\cap B)\cap(A\cap C)$	د- لا شئ مما سبق

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

أ- $(A\cup B)\cap(A\cup C)$	ب- $(A\cap B)\cup(A\cap C)$
ج- $(A\cap B)\cap(A\cap C)$	د- لا شئ مما سبق

٣- يراد شراء ثلاث أنواع من الكتب الدراسية A و B و C فإن :-

وحرف [و] يعني تقاطع

أ- $A\cup B\cup C$	ب- $A\cap B\cap C$
ج- $A\cup B\cap C$	د- لا شئ مما سبق

٤- عدم توافر الكتب الدراسية الثلاثة يرمز لها بالرمز :-

عدم كونه $\bar{A}$

أ- $A\cup B\cup C$	ب- $\bar{A}\cap\bar{B}\cap\bar{C}$
ج- $A\cup B\cap C$	د- لا شئ مما سبق

٥- توافر نوع واحد من الكتب الدراسية على الأقل A أو B أو C أو كلها يرمز لها بالرمز :-

أ- $A\cup B\cup C$	ب- $A\cap\bar{B}\cap\bar{C}$
ج- $\bar{A}\cap B\cap C$	د- لا شئ مما سبق

٦- توافر الكتاب الدراسي A فقط يمكن الرمز له بالرمز :

أ- $A\cup B\cup C$	ب- $A\cap\bar{B}\cap\bar{C}$
ج- $A\cap B\cap C$	د- لا شئ مما سبق

٧- توافر نوع واحد فقط من الكتب الدراسية يمكن الرمز له بالرمز :-

أ- $\bar{A}\cup\bar{B}\cup\bar{C}$	ب- $A\cap\bar{B}\cap\bar{C}$
ج- $(A\cap\bar{B}\cap\bar{C})\cup(\bar{A}\cap B\cap\bar{C})\cup(\bar{A}\cap\bar{B}\cap C)$	د- لا شئ مما سبق

٨- الجدول التالي يمثل توزيع مجموعة من الطلاب والطالبات حسب التخصص الدقيق بكلية إدارة الأعمال :-
تم اختيار احد الدارسين من الجدول السابق بطريقة عشوائية ، أحسب الاحتمالات التالية :-

	طلاب A	طالبات B	المجموع
محاسبة A	١٠	١٤	٢٤
نظم	١٦	٢٨	٤٤
ادارة B	٢٠	١٢	٣٢
المجموع	٤٦	٥٤	١٠٠

احتمال أن يكون طالب :-

$\frac{46}{100} = ,46$

أ- ٠,٥٤	ب- ٠,٤٦
ج- ٠,٢٤	د- لا شئ مما سبق

٩- احتمال أن تكون طالبة :-

$\frac{54}{100} = ,54$

أ- ٠,٥٤	ب- ٠,٤٦
ج- ٠,٢٤	د- لا شئ مما سبق

١٠- احتمال أن يكون من قسم المحاسبة :-

$\frac{24}{100} = ,24$

أ- ٠,٥٤	ب- ٠,٤٦
ج- ٠,٢٤	د- لا شئ مما سبق

١١- احتمال أن يكون من قسم المحاسبة وطالب :-

تمارين مراجعة

أ- ٠,١٠	ب- ٠,٤٦
ج- ٠,٢٤	د- لا شئ مما سبق

$\frac{10}{100} = 0,10$

١٢- أن يكون طالبة أو من قسم المحاسبة : -

أ- ٠,٦٤	ب- ٠,٧٨
ج- ٠,٥٤	د- لا شئ مما سبق

$\frac{54}{100} + \frac{24}{100} - \frac{14}{100} =$

١٣ أن يكون من قسم الإدارة أو طالب :-

أ- ٠,٧٨	ب- ٠,٣٢
ج- ٠,٥٨	د- لا شئ مما سبق

$\frac{32}{100} + \frac{46}{100} - \frac{20}{100}$

١٤- احتمال أن يكون من قسم المحاسبة بشرط أن تكون طالبة : -

أ- $\frac{7}{27}$	ب- $\frac{24}{100}$
ج- $\frac{54}{100}$	د- لا شئ مما سبق

$P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{\frac{14}{100}}{\frac{54}{100}} = \frac{7}{27}$

١٥- احتمال أن يكون طالب بشرط أنه من من قسم الإدارة : -

أ- $\frac{32}{100}$	ب- $\frac{5}{8}$
ج- $\frac{20}{100}$	د- لا شئ مما سبق

$\frac{\frac{20}{100}}{\frac{32}{100}} =$

إذا علمت أن :

مصنع لإنتاج لعب الأطفال يمتلك ثلاث آلات A و B و C تنتج الآلة الأولى ٢٥% من الإنتاج والآلة الثانية ٤٠% من الإنتاج والباقي من إنتاج الآلة الثالثة فإذا كانت نسبة المعيب في الآلات الثلاثة على الترتيب هو ٣% و ٤% ٦% ، سحبت وحدة واحدة عشوائياً من إنتاج المصنع . احسب الإحتمالات التالية :

C	B	A
35%	40%	25%
6%	4%	3%
94%	96%	97%

١٦ احتمال أن تكون الوحدة المسحوبة معيبة :-

أ- ٠,٢٥×٠,٩٧+٠,٤٠×٠,٩٦+٠,٣٥×٠,٩٤	ب- ٠,٢٥×٠,٠٣+٠,٤٠×٠,٠٤+٠,٣٥×٠,٠٦
ج- ٠,٧٥×٠,٠٣+٠,٦٠×٠,٠٤+٠,٦٥×٠,٠٦	د- لا شئ مما سبق

$= 0,054$

١٧- احتمال أن تكون الوحدة المسحوبة جيدة :-

أ- ٠,٢٥×٠,٩٧+٠,٤٠×٠,٩٦+٠,٣٥×٠,٩٤	ب- ٠,٢٥×٠,٠٣+٠,٤٠×٠,٠٤+٠,٣٥×٠,٠٦
ج- ٠,٧٥×٠,٠٣+٠,٦٠×٠,٠٤+٠,٦٥×٠,٠٦	د- لا شئ مما سبق

١٨- احتمال أن تكون الوحدة المسحوبة معيبة ومن انتاج الآلة الثالثة :-

أ- $\frac{0.94\times0.35}{0.25\times0.97+0.40\times0.96+0.35\times0.94}$	ب- $\frac{0.40\times0.04}{0.25\times0.03+0.40\times0.04+0.35\times0.06}$
ج- $\frac{0.06\times0.35}{0.25\times0.03+0.40\times0.04+0.35\times0.06}$	د- لا شئ مما سبق

إذا كان إذا علمت أنه :

" أحد المصانع وجد أنه من بين كل ١٠٠٠ وحدة هناك ١٥٠ وحدة معيبة ،أخذت عينة مكونة من خمس وحدات ، فإذا علمت أن هذه الظاهرة تتبع التوزيع ثنائي الحدين " أوجد الإحتمالات التالية :-

$q = \frac{150}{1000} = 0,15$

$p = 0,85$

١٩- احتمال أن تكون الوحدات المختارة كلها سليمة :- $P(X=5) = C^5_5\, 0,85^5\, 0,15^0$

أ- ٠,٥٥٦٣	ب- ٠,٤٤٣٧
ج- ٠,٨٣٥٢	د- لا شئ مما سبق

٢٠ احتمال وجود وحدة على الأكثر معيبة :-

$P(X\leq 1) = P(X=0) + P(X=1)$

$= C^5_0\, 0,15^0\, 0,85^5 + C^5_1\, 0,15^1\, 0,85^4$

أ- ٠,٣٩١٥	ب- ٠,٤٤٣٧
ج- ٠,٨٣٥٢	د- ٠,٠٠٢٢٢٧

٢١ احتمال وجود وحدتان معيبتان على الأقل :-

$P(X\geq 2) = P(X=2) + P(X=3) + P(X=4) + P(X=5)$

أ- ٠,١٦٤٨	ب- ٠,٨٥٠٠
ج- ٠,٨٣٥٢	د- ٠,٩٩٧٧٧

المحاضرة (٤)

تابع التوزيعات الإحتمالية

$P(x) = C^{n}_{x} \cdot p^{x} \cdot q^{n-x}$

تابع توزيع ثنائي الحدين :-

مثال : 1

إذا كان احتمال إصابة الهدف لشخص ما هو $\frac{1}{5}$ أتيحت له فرصة الرماية في ١٠ محاولات : المطلوب

$\boxed{p = \frac{1}{5}}$ ، $\boxed{q = 1 - \frac{1}{5} = \frac{4}{5}}$

$\boxed{n = 10}$

- (1) ما هو إحتمال إصابة الهدف مرتين على الأكثر. $P(X \le 2) = x=0 , x=1 , x=2$
- (2) إحتمال إصابة الهدف مرة واحدة

(1) $P(X \le 2) = C^{10}_{0} \cdot (\frac{1}{5})^{0} \cdot (\frac{4}{5})^{10-0} + C^{10}_{1} (\frac{1}{5})^{1} (\frac{4}{5})^{9} + C^{10}_{2} \cdot (\frac{1}{5})^{2} \cdot (\frac{4}{5})^{8}$

$,107 + ,26 + ,301 = ,668$

1. احتمال إصابة الهدف مرتين على الأكثر :-

(2) $P(X=1) = C^{10}_{1} \cdot (\frac{1}{5})^{1} \cdot (\frac{4}{5})^{9}$

$= ,26$

أي احتمال x = 0 or x = 1 or x = 2

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

$$= \binom{10}{0}(1/5)^0(4/5)^{10} + \binom{10}{1}(1/5)^1(4/5)^9 + \binom{10}{2}(1/5)^2(4/5)^8$$

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

2. احتمال إصابة الهدف مرة واحدة

أي احتمال x = 1

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

مثال :- 2

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

الحل

احتمال النجاح (ظهور صورة) p = 0.5

احتمال الفشل (ظهور كتابة) q = 0.5

عدد الرميات المستقلة n = 3

X متغير عشوائي يمثل عدد الصور يأخذ القيم 0, 1, 2, 3 ويكون له توزيع ذي الحدين

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

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

$$p(X=1) = \binom{3}{1}(0.5)^1(0.5)^{3-1} = \frac{3!}{1!(3-1)!}(1/2)(1/2)^2$$

$$p(X=2) = \binom{3}{2}(0.5)^2(0.5)^{3-2} = \frac{3!}{2!(3-2)!}(1/2)^2(1/2)^1$$

$$= \frac{3\times2\times1}{2\times1\times1} \times \frac{1}{2} \times \frac{1}{4} = \frac{3}{8}$$

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

$$= \frac{3\times2\times1}{3\times2\times1\times0!} \times \frac{1}{8} \times 1 = \frac{1}{8}$$

التوقع $E(X) = np = 3(0.5) = 1.5$

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

الموضوع: تابع حل ك

حل الكويز

المعطيات

$p = ,5$

$q = ,5$

$n = 3$

$$p(x) = C^{n}_{x} \cdot p^{x} \cdot q^{n-x}$$

∴ إحتمال لا بد أن يتراوح $0 \le p \le 1$

مجموع الاحتمالات يجب $= 1$

المطلوب: أ) التوزيع الاحتمالي

$x = 0 \;,\; x = 1 \;,\; x = 2 \;,\; x = 3$

$$p(x=0) = C^{3}_{0} \cdot ,5^{0} \cdot ,5^{3} = ,125$$

$$p(x=1) = C^{3}_{1} \cdot ,5^{1} \cdot ,5^{2} = ,375$$

$$p(x=2) = C^{3}_{2} \cdot ,5^{2} \cdot ,5^{1} = ,375$$

$$p(x=3) = C^{3}_{3} \cdot ,5^{3} \cdot ,5^{0} = ,125$$

$\overline{\qquad 1 \qquad}$

مجموع الاحتمالات $= 1 \longleftarrow$

ب) أوجد التوقع

$$\mu = n \cdot p = 3 \cdot ,5 = 1,5$$

ج) أوجد التباين

$$\sigma^{2} = n \cdot p \cdot q$$
$$= 3 \cdot ,5 \cdot ,5 = ,75$$

د) الانحراف المعياري = جذر التباين

$$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{,75} = ,86$$

مثال ٣:-

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

١. الوحدات المختارة كلها سليمة

٢. على الأكثر توجد واحدة معيبة — يستخدم علامة أصغر من أو يساوي ١

٣. على الأقل توجد وحدتان معيبتان — يستخدم علامة أكبر من أو يساوي ٢

٤. القيمة المتوقعة والتباين للوحدات المعيبة.

الحل

الأحداث النادرة

يستخدم عندما n كبير

توزيع بواسون

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

حيث : x = العدد المعين من النجاحات.

$P(x)$ = احتمال عدد x من النجاحات.

e = أساس نظام اللوغاريتمات الطبيعي وتوجد في بعض الآلات الحاسبة، وقيمتها هي: $e = 2.718$ تقريبا، ويمكن حساب قيمتها باستخدام الآلة الحاسبة.

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

μ = المتوسط

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

إذا كان للمتغير X توزيع بواسون فإن

الإحتمال : $$p(x) = e^{-\lambda} \times \frac{\lambda^x}{x!}$$

التوقع $E(X) = \lambda = np$

التباين $Var(X) = \lambda$

الموضوع: تابع ص ٢٢

مثال ٣

المعطيات

$,15 = \frac{150}{1000} = q =$ الخطأ

$,85 = 1-,15 = p =$ السليم

$n = 5$

① $P(X=5) = C^{5}_{5} \cdot ,85^{5} \cdot ,15^{5-5} = ,44$

② $P(X \leq 1) = (x=0) ، (x=1)$

$C^{5}_{0} \cdot ,85^{0} \cdot ,15^{5-0} + C^{5}_{1} \cdot ,85^{1} \cdot ,15^{5-1}$

$= ,0022$

علماً بأن هذه تستخدم علامة أصغر من أو يساوي

أي على الأقل تستخدم علامة أكبر من أو يساوي

③ $P(X \geq 2) = \cancel{(x=2)} ، (x=3) ، (x=4) ، x=5)$

$= 1 - (x=0 ، x=1)$

من أصغر

$= 1 - ,0022 = ,9971$

مجموع الاحتمالات = 1

المطلوب في السؤال $2 ، 3 ، 4 ، 5$

الحل الصحيح ← $1 - [P(x=0) + P(x=1)]$

④ الوسط المتوقع

$\mu = n \cdot p$

$= 5 \cdot ,85 = 4,25$

⑤ التباين

$\sigma^2 = n \cdot p \cdot q = 5 \cdot ,85 \cdot ,15 = ,63$

⑥ الانحراف المعياري

$\sigma = \sqrt{n \cdot p \cdot q} = \sqrt{,63} = ,79$

p = ٣٠,

مثال ٤

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

n = 350 ← توزيع بواسون

١) وجود قطعة معيبة
٢) وجود قطعتان معيبتان
٣) عدم وجود أية قطع معيبة
٤) وجود على الأكثر وحدتان معيبتان

الحل

عملية سحب العينة تمثل سلسلة عددها n=350

واحتمال أن تكون القطعة معيبة (النجاح) p=0.003

واضح n كبيرة و p صغيرة

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

بفرض أن X يمثل عدد القطع المعيبة في العينة له توزيع بواسون

x = 0, 1, 2, ...

١) وجود قطعة معيبة في العينة

$$p(X=1) = e^{-1.05}\frac{1.05^1}{1!} = (0.3499)(1.05) = 0.367$$

٢) وجود قطعتان معيبتان في العينة

$$p(X=2) = e^{-1.05}\frac{1.05^2}{2!} = (0.3499)(0.55125) = 0.193$$

٣) عدم وجود أي قطع معيبة في العينة

$$p(X=0) = e^{-1.05}\frac{1.05^0}{0!} = 0.350$$

٤) وجود على الأكثر وحدتان معيبتان

$$P(X \leq 2) = p(X=0) + p(X=1) + p(X=2)$$
$$= 0.350 + 0.367 + 0.193$$
$$= 0.91$$

التوزيع الإحصائي :-

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

ويرتبط التوزيع الاحصائي عادة بنوعين من البيانات المتصلة والمنفصلة، ويناسب النوع المنفصل المقاييس الاسمية والرتبية، وهناك بعض المقياس المنفصلة ثنائية أي انه لا يوجد بها الا قيميتين، وهي لا تسمي توزيعات طبيعية وانما تسمى توزيعات ثنائية، ومن أهم مقاييس التوزيعات المنفصلة مقياس ذو الحدين وذلك عائد لان الاجابة على المقياس الاسمي اما نعم أو لا، ولذلك غالبا ما يرمز لها في الحاسب بصفر (غياب الصفة) ذكور – لا) أو ١ (وجود الصفة) اناث – نعم).

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

التوزيعات الاحتمالية للمتغيرات المتصلة:

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

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

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

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

الموضوع: مثال ٤ ص ٢٤

بواسون

اعتبرت نجاح لأنه جميع المطلوب عليه	$p = 0.003$
نتوجه نحو توزيع بواسون لأن حجم العينة كبير	$n = 350$

المطلوب:

(1) وجود ورقة معيبة

1. $p(X=1)$

$$= e^{-\lambda} \cdot \frac{\lambda^{x}}{x!}$$

$$\lambda = n.p$$

$$= 350 \times 0.003 = 1.05$$

$$\therefore p(X=1) = e^{-1.05} \cdot \frac{1.05^{1}}{1!} = 0.36$$

وجود ورقتان معيبتان

2. $p(X=2)$

$$= e^{-1.05} \cdot \frac{1.05^{2}}{2!} = 0.19$$

3. $p(X=0) = e^{-1.05} \cdot \frac{1.05^{0}}{0!} = 0.34$

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

$$= 0.34 + 0.36 + 0.19 = 0.89$$

5. $p(X \ge 2)$

المطلوب وجود ورقتان وجدتا معيبتان.

سؤال امتحان

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

$$1 - [0.34 + 0.36] = 0.3$$

استخدام الآلة

$$\text{shift ln } -1.05 \times \frac{1.05^{2}}{2 \text{ shift } x^{-1}} =$$

التوزيع الطبيعي

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

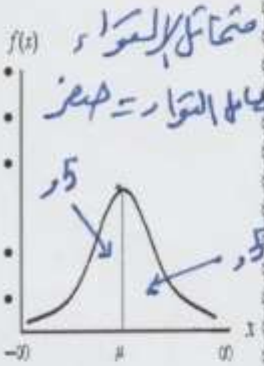
والتوزيع الطبيعي هو توزيع احتمالي متصل، وهو جرسي الشكل ومتماثل حول الوسط الحسابي، ويمتد إلى مالا نهاية في الاتجاهين، ولكن معظم المساحة (الاحتمال) تتركز حول الوسط الحسابي.

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 تماما لأي توزيع طبيعي خاصة إذا علمنا الوسط الحسابي والانحراف المعياري لهذا التوزيع.
- تدل قيمة على مكان مركز الجرس، كما تدل على كيفية الانتشار.
- القيمة الصغيرة لـ تعني أن لدينا جرس طويل مدبب، والقيمة الكبيرة لـ تعني أن الجرس قصير ومفرطح.

والشكل المقابل يوضح ذلك:



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

معالم هذا التوزيع:

توجد معلمتين لهذا التوزيع هما :

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

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

التوزيع الطبيعي القياسي (المعياري): - هام جداً بل حفظ

- 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

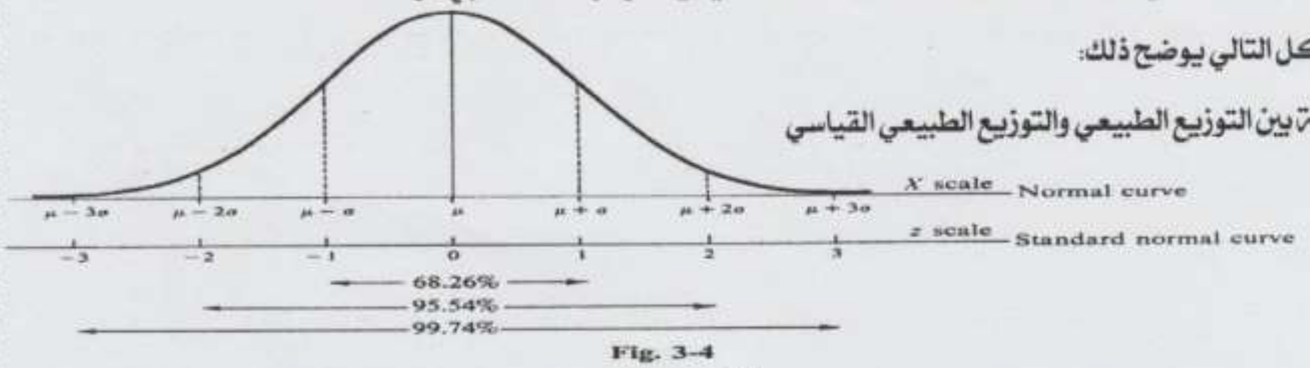


Fig. 3-4

الإنحراف المعياري	الإحتمال
±١	٠,٦٨٢٧
±٢	٠,٩٥٤٥
±٣	٠,٩٩٧٣

هام

القانون المستخدم لتحويل التوزيع الطبيعي إلى قياسي :

القيمة المراد تحويلها للتوزيع الطبيعي القياسي

وسط المجتمع

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

الانحراف المعياري

مثال :-

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

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

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

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

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

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

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

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

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

تابع

الموضوع: التوزيع الطبيعي

حل مثال ح ٢٧

المعطيات:

$\mu = 180$

$\sigma = 10$

المطلوب

(١)

$P(170 < X < 190)$

$Z_{170} = \frac{170-180}{10} = \frac{-10}{10} = -1$

$Z_{190} = \frac{190-180}{10} = \frac{10}{10} = 1$

$P(-1 < Z < 1) = ,68 = 68\%$

(٢)

$P(160 < X < 200)$

$Z_{160} = \frac{160-180}{10} = \frac{-20}{10} = -2$

$Z_{200} = \frac{200-180}{10} = \frac{20}{10} = 2$

$P(-2 < Z < 2) = ,95 = 95\%$

(٣)

$P(150 < X < 210)$

$Z_{150} = \frac{150-180}{10} = \frac{-30}{10} = -3$

$Z_{210} = \frac{210-180}{10} = \frac{30}{10} = 3$

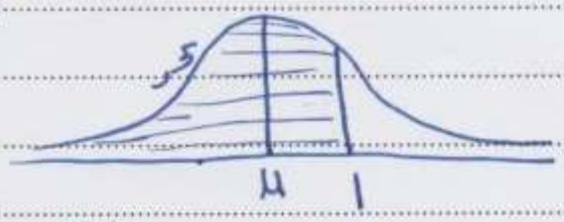
$P(-3 < Z < 3) = ,99 = 99\%$

الموضوع: تابع تمرين التوزيع الطبيعي

(٤) احتمال أن يكون الطالب أقل من ~~١٨٠~~ ١٩٠

$P(X < 190)$

$$Z_{190} = \frac{190-180}{10} = 1$$



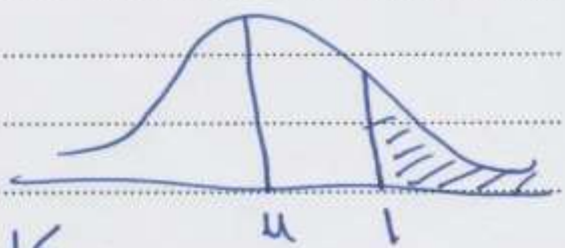
$P(Z < 1)$

$$= .5 + \frac{.68}{2} = .84$$

(٥) احتمال أن يكون الطالب أكبر من ١٩٠

$P(X > 190)$

$$Z_{190} = \frac{190-180}{10} = 1$$



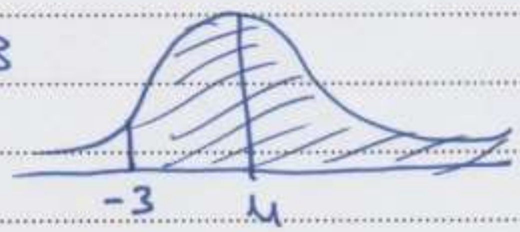
$P(Z > 1)$

$$= .5 - \frac{.68}{2} = .16$$

(٦) احتمال أن يكون طول الطالب أكبر من ١٥٠ سم

$P(X > 150)$

$$Z_{150} = \frac{150-180}{10} = -3$$



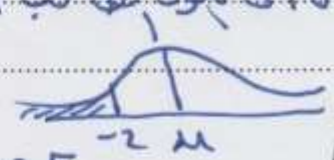
$P(Z > -3)$

$$= \frac{.99}{2} + .5 = .995$$

(٧) احتمال أن يكون الطالب أقل من ١٦٠

$P(X < 160)$

$$Z_{160} = \frac{160-180}{10} = -2$$



$$P(Z < -2) = .5 - \frac{.95}{2} = .025$$

استخدامات التوزيع الطبيعي القياسي:-

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

مثال:

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

- نسبة السيارات التي تقل سرعتها عن 50 ميلا في الساعة.
- نسبة السيارات التي تزيد سرعتها عن 65 ميلا في الساعة.
- نسبة السيارات التي تكون سرعتها بين 60 ميلا و 70 ميلا في الساعة.
- عدد السيارات التي تكون سرعتها بين 60 ميلا و ~~77.5~~ ميلا من بين 10000 سيارة .

70

الموضوع: مثال ص ٣٠

المعطيات

$\mu = 60$

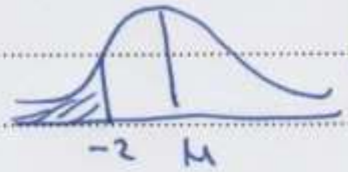
التباين $= 25$

الانحراف $\sigma = \sqrt{25} = 5$

1. $P(X < 50)$

$$Z_{50} = \frac{50-60}{5} = \frac{-10}{5} = -2$$

$P(Z < -2)$



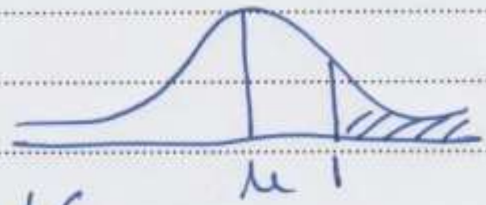
$$= ,5 - \frac{,95}{2}$$

$$= ,025$$

2. $P(X > 65)$

$$Z_{65} = \frac{65-60}{5} = 1$$

$P(Z > 1)$

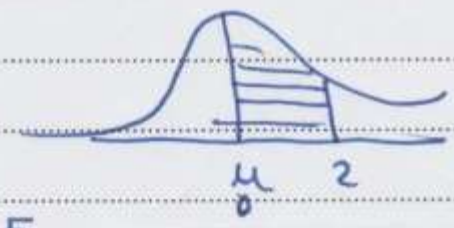


$$= ,5 - \frac{,68}{2} = ,16$$

3. $P(60 < X < 70)$

$$Z_{60} = \frac{60-60}{5} = 0$$

$$Z_{70} = \frac{70-60}{5} = 2$$



$$P(0 < Z < 2) = \frac{,95}{2} = ,475 = 47,5\%$$

4. $P(60 < X < 70)$

المطلوب عدد

$$= ,475 \times 10,000 = 4750$$

ناتج النسبة × صاحب المعطى!

المحاضرة (٥)

توزيعات المعاينة

الاستدلال الإحصائي :-

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

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

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

العينة العشوائية :-

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

فمثلا نستعين بعينه مسحوبة من المجتمع لتقدير معالم هذا المجتمع مثل متوسطة أو تباينه أو غير ذلك. أو أعطاء عينه من المرضى بارتفاع الضغط، مثلا 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 اختصارا 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 محدودا 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

أساليب جمع البيانات :-

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

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

بعض مزايا أسلوب المعاينة:-

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا عدديه 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 بدلا 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. أيضا 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 مثلا 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:

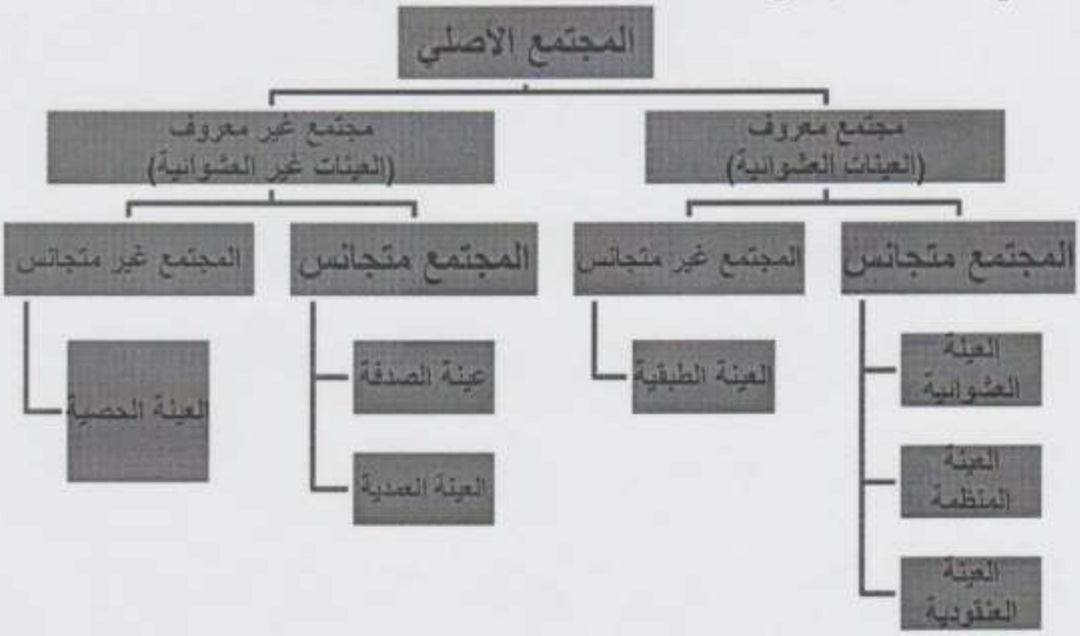
١. العينات العشوائية:

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، وغالبا يتدخل الباحث في عملية الاختيار بصورة أو بأخرى.

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



أ- العينات الاحتمالية:

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

ب- العينات غير الاحتمالية:

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

أخطاء البيانات الإحصائية:-

هام جداً

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:

١ـ 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

٢ـ خطأ المعاينة العشوائية Random Sampling Error

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

كيف نقلل من خطأ المعاينة العشوائي:

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

المعالم والإحصاءات:-

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

فالمقاييس الإحصائية التي تحسب من بيانات مجتمع الدراسة بأكمله يطلق عليها معالم المجتمع أما المقاييس الإحصائية التي تحسب من بيانات عينه مسحوبة من مجتمع الدراسة فيطلق عليها إحصاءات (Statistics)

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

الخطأ المعياري:

بمعرفة قيمة الانحراف المعياري لقيم العينة يمكن تقدير

- قيمة الخطأ المعياري في الانحراف المعياري للعينة مع معرفة حجم المجتمع:

خطأ معياري

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

حيث ان: (N) = حجم مجتمع العينة و (n) = حجم العينة

وفق هذه المعادلة تؤخذ نسبة العينة إلى مجتمعها، وكلما كبرت هذه النسبة تحسن تمثيلها لمجتمع الدراسة.

- أما عندما يكون حجم مجتمع العينة مجهولا، حينها تعتمد المعادلة الآتية:-

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

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

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

مستوى الثقة وحدودها:

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

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

قيمة الخطأ = ١ - ٦٨ و

- مستوى ثقة إحصائية قدره (٦٨.٢٦٪) أو باحتمالية قدرها (٠.٦٨٢٧) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و (+ و-) درجة واحدة من الخطأ المعياري.
- مستوى ثقة إحصائية قدرها (٩٥٪)، أو باحتمالية (٠.٩٥) يقع متوسط مجتمع الدراسة بين متوسط متوسطات العينات و (+ و −) درجتان من الخطأ المعياري تقريبا. ٠.٠٥
- مستوى ثقة إحصائية قدرها (٩٩٪) أو باحتمالية قدرها (٠.٩٩) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و (+ و −) ثلاث درجات من قيمة الخطأ المعياري تقريبا. ٠.١
- وتسمى هذه بمستويات الثقة Confidence Level و يعبر عنها بإشارة النسبة المئوية (٪) بان تكون التقديرات صحيحة أو باحتمالية (٠.٠١) أو (٠.٠٥) أن تكون خاطئة.

حجم العينة:-

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

1. التوزيع الطبيعي للقيم
2. توزيع (ت) للقيم

القانون المستخدم : $n=\dfrac{Z^2\sigma^2}{(e)^2}$ سيتم استخدام هذا لاحقا

(أ) التوزيع الطبيعي للقيم: حجم العينة ٣٠ فأكثر نستخدم التوزيع الطبيعي

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

(ب) توزيع (ت) للقيم :

من الضروري اخذ الحذر عندما يكون حجم العينة صغيرا اقل من (٣٠) وذلك لأنها تتطلب إجراءات خاصة عند التحليل لهذا يستخدم في الدراسة توزيع (ت) . فعندما يكون حجم العينة أكثر من (٣٠) يتجه توزيع قيمها نحو التوزيع الطبيعي وبغض النظر عن التوزيع الحقيقي لقيم مجتمع الدراسة. وبالنسبة للعينات الصغيرة الحجم فان توزيع قيمها يتأثر بطبيعة توزيع قيم المجتمع المأخوذة منه. وعندما يكون توزيع قيم المجتمع معروفا أو متوقعا أن لا يكون طبيعي حينها يجب اعتماد حجم كبير للعينة. أما إذا كان توزيع قيم المجتمع طبيعيا عندها يمكن اخذ عينات بحجم صغير ويعتمد توزيع (ت) (T) في التحليل و المقارنة.

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

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

العوامل المحددة لحجم العينة:

- **درجة التباين في خصائص مجتمع الدراسة:** يلعب التباين في خصائص مجتمع الدراسة دورا مهما في تحديد درجة دقة نتائج العينة، فكلما كان التباين كبيرا تطلب الأمر زيادة حجم العينة ليكون تمثيلها للتباين في المجتمع صحيحا.
- **طريقة التحليل المعتمدة:** عند إقرار حجم العينة، من الضروري تحديد الحجم الأصغر المقبول للعينة في المجاميع الثانوية ضمن مجتمع الدراسة ، إذ أن بعض الاختبارات الإحصائية تتطلب عددا معينا كحد أدنى لكل فئة أو صنف لتكون النتائج ذات معنى.
- **حجم المعلومات المطلوبة:** فكلما كانت المعلومات المطلوبة من العينة (الواحدة) كثيرة وتفصيلية كان حجم العينة صغيرا، ما لم يكن المشروع البحثي كبيرا وتتوفر له المصادر البشرية والمادية اللازمة. إن الدقة في المعلومات المطلوبة من العينة أهم بكثير من حجم العينة ، فحجم العينة لا يتحدد بحجم مجتمع الدراسة فقط، بل وبالدقة المتوخاة والتفاصيل المطلوبة
- **المصادر المالية والبشرية المتوفرة:** تتطلب الدراسة الميدانية توفر مصادر مالية وبشرية لتغطية تكاليفها التي تكون في الغالب باهظة لتأثيراتها على تحديد حجم منطقة الدراسة، مجتمع الدراسة وبالتالي حجم العينة. إن مضاعفة حجم العينة يتطلب زيادة في كمية المصادر المالية والجهد البشري
- **حدود الثقة في تقديرات خصائص مجتمع الدراسة:** لزيادة الدقة في النتائج يعمد البعض إلى تقليص حدود الثقة (المدى الذي يفترض أن يقع ضمنه المعدل المتوقع للمجتمع). إن إنقاص حدود الثقة من (٦٪) إلى (٤٪) يتطلب زيادة حجم العينة بنسبة (٢٢٥٪)، وكلما كان المدى كبيرا كان حجم العينة صغيرا، والعكس صحيح.
- **حالات الإخفاق وعدم الاستجابة:** العامل الآخر الذي يحدد حجم العينة هي حالات الإخفاق في الحصول على المعلومات وعدم الاستجابة أو المعلومات غير الوافية

المحاضرة (٦)

التقدير الإحصائي

التقدير:-

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

الانحراف المعياري للمجتمع

حجم العينة

١) التقدير بنقطة :-

التقدير بنقطة يعني أن نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة. بحساب الخطأ المعياري = $\frac{\sigma}{\sqrt{n}}$

فمثلا 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 معينا كتقدير لهذه النسبة في المجتمع نكون حصلنا على تقدير نقطة للنسبة في مجتمع الناخبين.

٢) التقدير بفترة:-

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

مثلا:

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

تقدير الوسط الحسابي للمجتمع :-

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

$$\hat{\mu} = \bar{X} \pm Z\frac{S}{\sqrt{n}}$$

مع ملاحظة أنه إذا كان الانحراف المعياري للمجتمع غير معروف. وهو غالبا ما يحدث في الواقع – فيمكن استخدام الانحراف المعياري للعينة S بدلا منه طالما كان حجم العينة كبيرا بدرجة كافية وتصبح فترة تقدير الوسط الحسابي للمجتمع كما يلي : ولإيضاح هذه النقطة بشيء من التفصيل نأخذ المثال التالي :

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

مثال :- 1

لو أردنا معرفة متوسط الدخل اليومي لمجموعة من الناخبين في دولة ما، فإن ذلك يبدو أمرا صعبا من الناحية العملية نظرا لكبر حجم مجتمع الناخبين، إضافة إلى طول الوقت والتكاليف. لذا فإن الأسلوب العلمي المتبع في حالة كهذه هو اختيار عينة عشوائية نستطيع من خلال معرفة نتائجها لتقدير متوسط دخول الناخبين في هذه الدولة. فلو سحبت عينة عشوائية من مجموع مجتمع الناخبين في دولة ما حجمها 100 ناخب، فإذا كان الوسط الحسابي والانحراف المعياري للدخل السنوي للناخبين بالعينة هما على الترتيب 90 ألف ريال و 25 ألف ريال.

المطلوب:

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

الحل :-

$$\hat{\mu} = \bar{X} \pm Z\frac{S}{\sqrt{n}}$$

المعطيات

$n = 100$

$\bar{X} = 90$

$S = 25$

$Z = 1,96$

من الجدول

$$\mu = 90 \pm 1,96 \cdot \frac{25}{\sqrt{100}}$$

$$90 \pm 4,9$$

الحد الأدنى للتقدير

$90 - 4,9$

$= 85,1$

الحد الأعلى للتقدير

$90 + 4,9$

$= 94,9$

مثال :- 2

أخذت عينة عشوائية حجمها 144 بوسط مقداره 100 وانحراف معياري مقداره 60 وبالتالي فإن فترة تقدير الوسط الحسابي للمجتمع بدرجة ثقة 95٪ هي : ثم أوجد حجم العينة عندما يكون الخطأ المسموح به = 5

المعطيات

$n = 144$

$\bar{X} = 100$

$S = 60$

$Z = 1,96$

$$\mu = \bar{X} \pm Z \cdot \frac{S}{\sqrt{n}}$$

$$= 100 \pm 1,96 \cdot \frac{60}{\sqrt{144}}$$

$$100 \pm 9,8$$

$90,2$ = الأدنى

$109,8$ = الأعلى

$$n = \frac{Z^2\sigma^2}{e^2} = \frac{1,96^2 \cdot 60^2}{5^2} = 553,19$$

الخطأ

مثال :- 3

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

الحل :-

$\sigma = 15$

$n = ??$

$e = 5$

$Z = 2,58$

$$n = \frac{Z^2 \cdot \sigma^2}{e^2} = \frac{2,58^2 \cdot 15^2}{5^2} = 59,9$$

مثال :-

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

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

الحل :-

$$n = \frac{1.68^2 . 15^2}{3^2} = 70.56$$

$e = \pm 3$

$Z = 1.68$

$\sigma = 15$

فترات الثقة للمتوسط باستخدام توزيع t :-

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

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

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

درجات الحرية:

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

وكمثال مبسط لشرح فكرة درجات الحرية نفترض أن لدينا 3 قيم واشترطنا أن مجموع القيم يساوي 10 فإن لدى الباحث في هذه الحالة حرية في اختيار الرقم الأول (وليكن 2) والثاني (وليكن 3) لذلك فإن قيمة الثالثة لابد وأن تكون (5) بالتالي نستطيع القول بأن درجة الحرية المتاحة لدى الباحث هي (2) أي 1 - 3 = 2 أي أن درجات الحرية في هذه الحالة هي : n - 1 حيث n تساوي حجم العينة (والتي تساوي في المثال السابق 3) والرقم (1) والذي طرحناه يعني الشرط الذي يحتم أن مجموع القيم = 10 وبصفة عامة إذا كان عدد القيود k فإن درجات الحرية تساوي n - k

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

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

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

شروط توزيع t :

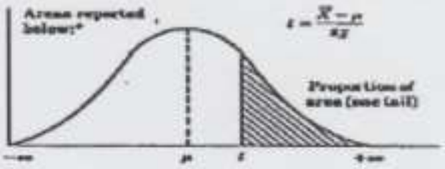
ويمكن تحديد الشروط الثلاثة لاستخدام توزيع t كما يلي:

١. أن يكون المجتمع المسحوبة منه العينة له توزيع طبيعي.
٢. والانحراف المعياري للمجتمع غير معروف (أو مجهول).
٣. والعينة صغيرة (حجمها أقل من 30 مفردة).

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	316.31	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										

df

مثال :-

العينة أقل من ٢٠ والمجتمع يتبع التوزيع الطبيعي :

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

المطلوب :

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

الحل:

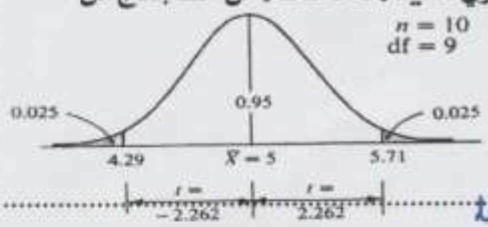


Fig. 4.4

$\mu = \bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$

$= 5 \pm 2{,}28 \cdot \frac{1}{\sqrt{10}}$

n = 10

$\bar{x} = 5$

s = 1

مستوى المعنوية = ,05

df = 10 − 1 = 9

t = 2,28

أدنى

5 − ,72

= 4,28

أعلى

5 + ,72

= 5,72

40

تقدير فترة النسبة للمجتمع

(فترة الثقة للنسبة)

إن تقدير النسبة في المجتمع تعتبر من الحالات المهمة لقياس الظواهر السياسية، وبالذات الوصفية منها كقياس اتجاهات الرأي العام، وقياس نسبة قتلى الحروب، ونسبة الدول التي أوفت بالتزاماتها في المنظمات الدولية أو الإقليمية... وغيرها

ونظرا لأنه من الصعوبة بمكان في كثير من الأحيان حساب هذه النسبة مباشرة من المجتمع، فإننا غالبا ما نلجأ لتقدير هذه النسبة من عينة عشوائية مسحوبة من هذا المجتمع.

خطوات تقدير النسبة في المجتمع:

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

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

مثال :-

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

الحل:

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

$n = 144$

$\bar{P} = \frac{60}{144} = ,41$

$Z = 1,96$

$$P = ,41 \pm 1,96 \cdot \sqrt{\frac{,41(1-,41)}{144}}$$

$$= ,41 \pm ,08$$

ادنى: $,33$

الأعلى: $,49$

ملاحظات:

* عند التقدير لفترة وكان حجم العينة أكبر من ٣٠ نستخدم تقدير بـ Z

$$\mu = \bar{x} \pm Z \cdot \frac{s}{\sqrt{n}}$$

* عند التقدير لفترة وكان حجم العينة أقل من ٣٠ نستخدم t

$$\mu = \bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$$

* عند التقدير لنسبة نستخدم القانون

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

الموضوع: التقدير الإحصائي.

التقدير بفترة:

1) إذا كان حجم العينة أكبر من ٣٠

$$\mu = \bar{x} \pm Z \cdot \frac{s}{\sqrt{n}}$$

متوسط المجتمع

وسط العينة

من الجدول

الانحراف المعياري

حجم العينة

* قانون حساب حجم العينة مثال ٣

من الجدول

الانحراف المعياري

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

الخطأ المسموح به

2) إذا كان حجم العينة أصغر من ٣٠

$$\mu = \bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$$

مستوى المعنوية

هو الثقة المكملة لدرجة ١٠٠ (طرح الثقة إلى ١٠٠)

$df = n - 1$

درجة الحرية

المحاضرة (٧)

إختبار الفروض الإحصائية

اختبارات الفروض الاحصائية

Testing Statistical Hypotheses

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

والفرض ما هو إلا تخمين أو استنتاج ذكي مبني على حيثيات معقولة أو منطقية ولكنه ليس مبنيا على حسابات دقيقة خاصة بالمجتمع لأننا نفترض أنه لا يمكن دراسة المجتمع بالكامل عن طريق الحصر الشامل بل نحاول استنتاج أو الاستدلال على مقاييس المجتمع باستخدام بيانات ونتائج العينة. **وقبل تناول كيفية إجراء الاختبارات الإحصائية نستعرض أولا بعض المفاهيم والتعريفات الأساسية اللازمة لهذا الموضوع حتى تكون الصورة أكثر وضوحا..**

مفهوم الاختبارات الإحصائية :-

الفرض العدمي (أو الصفري) The Null Hypothesis

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

$Ho : \mu = 200$

ويقرأ بالشكل التالي : الفرض العدمي هو : أن متوسط دخل الفرد في المنطقة هو 200 ريال شهريا.

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

$Ho : P = 0.30$

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

الفرض البديل : The Alternative Hypothesis

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

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

والفرض البديل له أهمية كبيرة في قياس الظواهر الاجتماعية – كما سوف نرى – فهو الذي يحدد نوع الاختبار المستخدم لذلك فهو يأخذ أحد أشكال ثلاثة هي :-

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

فمثلا : إذا كان الفرض العدمي هو أن متوسط الدخل الشهري لفئة معينة في المجتمع هو 200 ريال. $Ho : \mu = 200$

فإن الفرض البديل في هذه الحالة يأخذ الشكل التالي : بمعنى أن متوسط دخل هذه الفئة من المجتمع "لا يساوي" 200 ريال شهريا. $H1:\mu\neq200$ بمعنى أن متوسط دخل هذه الفئة من المجتمع "لا يساوي" 200 ريال شهريا.

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

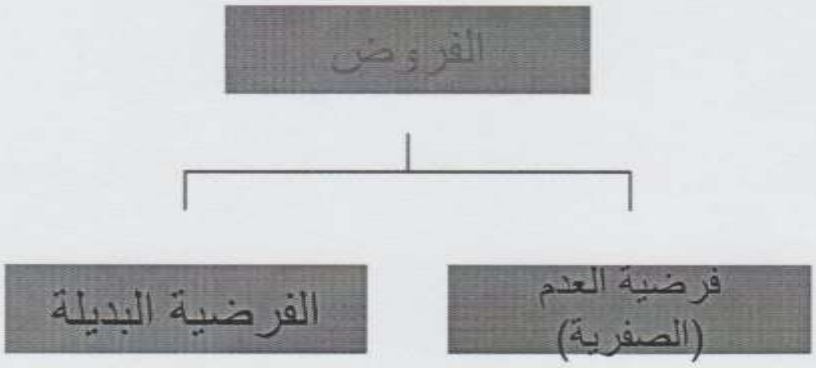
فمثلا : قد يكون الفرض البديل كما يلي : $H1 : \mu > 200$ أي أن متوسط الدخل لهذه الفئة من المجتمع أكبر من 200 ريال شهريا.

ج. وأخيرا قد يأخذ الفرض البديل شكل " أقل من ". وفي هذه الحالة نستخدم ما يسمى " اختبار الطرف الأيسر "

فمثلا : قد يكون الفرض البديل هو : $H1 : \mu < 200$ أي أن متوسط الدخل لهذه الفئة من المجتمع أقل من 200 ريال شهريا.

الخلاصة : الفروض الإحصائية :-

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



الخطأ في اتخاذ القرار :-

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

القرار	الموقف الفعلي	
	H_0 صحيحة	H_0 خطأ
أقبل H_0	قرار صحيح الاحتمال أنه على الأقل يساوي $1-\alpha$	خطأ من النوع II الاحتمال هو قيمة β تعتمد على عوامل عديدة
ارفض H_0	خطأ من النوع I الاحتمال على أكثر تقديريساوي α	قرار صحيح الاحتمال يساوي $1-\beta$

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".

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".

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

مستوى المعنوية : Level of Significance

والمقصود بمستوى المعنوية هو "احتمال حدوث الخطأ من النوع الأول" أو نسبة حدوثه "أي احتمال رفض الفرض العدمي بينما هو صحيح". وعادة ما يرمز إلى مستوى المعنوية بالرمز اللاتيني ألفا وأشهر قيمتين لمستوى المعنوية هما 5%، 1%، ولكن ليس هناك ما يمنع من أن يأخذ قيما أخرى.

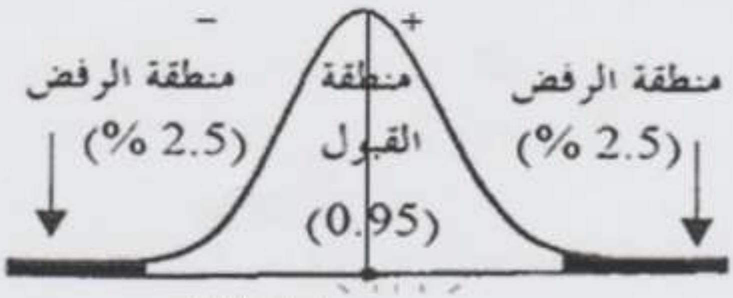
ومن الملاحظات المهمة هنا هو أن "مستوى المعنوية" والذي يسمى أحيانا "مستوى الدلالة" هو المكمل لدرجة الثقة "بمعنى أن مجموعهما يساوي 100% أو واحد صحيح. فإذا كانت درجة الثقة 95% فإن مستوى المعنوية يساوي 5%. والعكس صحيح فإذا كان مستوى المعنوية 5% فإن هذا يعني أن درجة الثقة 95 %. ولعل من أهم الملاحظات هنا هو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

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

وهناك ثلاث حالات مختلفة لمنطقتي القبول والرفض هي :

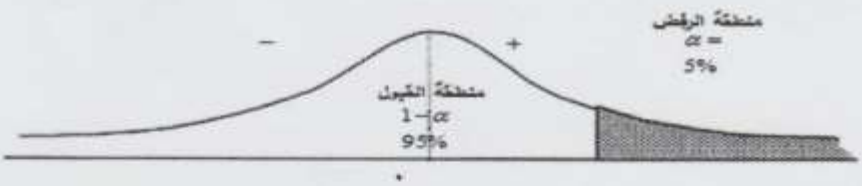
الأولى : إذا كان الفرض البديل يأخذ شكل "لا يساوي" كأن يكون الفرض في هذه الحالة هو أن متوسط دخل الفرد لا يساوي 200 ريال فإن منطقة الرفض تكون موزعة على طرفي المنحنى بالتساوي، ويسمى الاختبار في هذه الحالة "اختبار الطرفين"، والذي يأخذ الشكل التالي (بافتراض أن $\alpha=5\%$) : فالفرض العدمي هنا يعني أن متوسط دخل الفرد يساوي 200 ريال شهريا، والفرض البديل في هذه الحالة هو بمعنى أن متوسط دخل الفرد لا يساوي 200 ريال شهريا. 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 %.



اختبار الطرفين

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

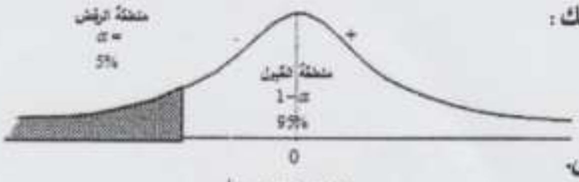


اختبار الطرف الأيمن

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

بمعنى أن متوسط دخل الفرد أكبر من 200 ريال شهريا. وبالتالي فإن مستوى المعنوية والذي يساوي مثلا 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 العدمي كما في المثال السابق، بينما الفرض البديل هو $H1:\mu<200$ بمعنى أن متوسط دخل الفرد أقل من 200 ريال شهريا، وبالتالي فإن مستوى المعنوية والذي يساوي مثلا 5 % مركز في الطرف الأيسر من المنحنى.



اختبار الطرف الأيسر

المحاضرة السابعة والثامنة

الموضوع: إختبار الفروض الإحصائية

الخطوات:

1 الفروض:

← فرض العدم $H_0: \mu = \mu_0$

← الفرض البديل $H_1: \mu \neq \mu_0$ صياغة إختلاف

$\mu > \mu_0$ كلمة أكبر/أفضل

$\mu < \mu_0$ كلمة أصغر/أقل

2 نوجد قيمة الإحصائية.

← Z الإحصائية ← عندما حجم العينة $n \geq 30$

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

← t الإحصائية ← عندما حجم العينة $n < 30$

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

3 نوجد قيمة الإحصائية الجدولية ← بناء على المستخدم بالخطوة 2

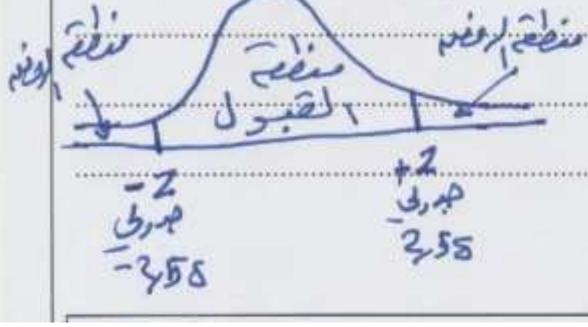
Z يتم حسابها من الجدول المعطى سابقا

t الجدولية

← df درجة الحرية = $n - 1$

← α مستوى المعنوية

4 القرار



← Z الإحصائية وقعت في منطقة القبول

أي تقبل فرض العدم ونرفض البديل

← Z الإحصائية وقعت في منطقة الرفض

تقبل البديل ونرفض العدم.

خطوات الاختبار الإحصائي :

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

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

"لا يساوي"

أو "أكبر من"

أو "أقل من"

وبالرموز فإن الفرض البديل قد يأخذ شكل أحد الصيغ المقابلة :

$$H1: \mu \neq 20$$

$$OR \mu > 20$$

$$OR \mu < 20$$

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

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

مثال (١):- عينة عشوائية حجمها 49 شخصا اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 ريال. كيف يمكن اختبار الفرض الصفري بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72 ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية % 5 إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال.

الحل:-

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

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

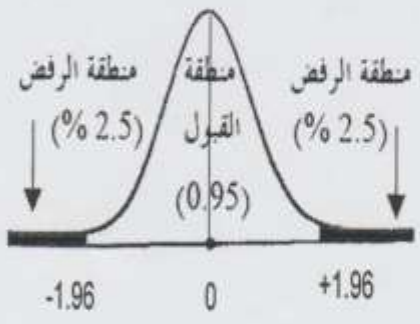
٣- الإحصائية: بما أن العينة كبيرة فإن الإحصائية في حالة اختبار الوسط تأخذ الشكل التالي :

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

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

وبالتعويض نحصل على : $Z_{\bar{X}} = \frac{75-72}{\frac{14}{\sqrt{49}}}$

أي أن قيمة الإحصائية تساوي 1.5 $Z_{\bar{X}} = \frac{3}{\frac{4}{7}} = \frac{3}{2} = 1.5$



٤- حدود منطقتي القبول والرفض: نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% وبما أن الفرض البديل هو: "لا يساوي" فإن ما يستخدم في هذه الحالة هو اختبار الطرفين كما في الشكل التالي : وقد حصلنا على حدود منطقتي القبول والرفض وذلك بقسمة درجة الثقة (المكملة لمستوى المعنوية) والتي تساوي 0.95 على 2 فنحصل على 0.4750 وبالكشف في جدول التوزيع الطبيعي المعياري عن Z التي تقابل المساحة 0.4750 نجد أنها تساوي 1.96 وحيث أنها موزعة على طرفي المنحنى بالتساوي فنضع إشارة موجبة في النصف الأيمن، وإشارة سالبة في النصف الأيسر، أي أن منطقة القبول تبدأ من القيمة ١,٩٦- وتستمر حتى القيمة ١,٩٦ + (أي أن أي قيمة محصورة بين هاتين القيمتين تكون في منطقة القبول، وأي قيمة خارج هذه الحدود تكون في منطقة الرفض).

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

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

الموضوع: تمرين مثال لحل المحاضرة السابعة

المعطيات

$n = 49$	حجم العينة
$\bar{x} = 75$	متوسط العينة
$\mu_0 = 72$	متوسط المجتمع
$\alpha = 5\%$	مستوى المعنوية
$\sigma = 14$	الانحراف المعياري

الحل

1) الفروض:

* فرض العدم: $H_0: \mu = 72$
* الفرض البديل: $H_1: \mu \neq 72$

2) قيمة الإحصائي

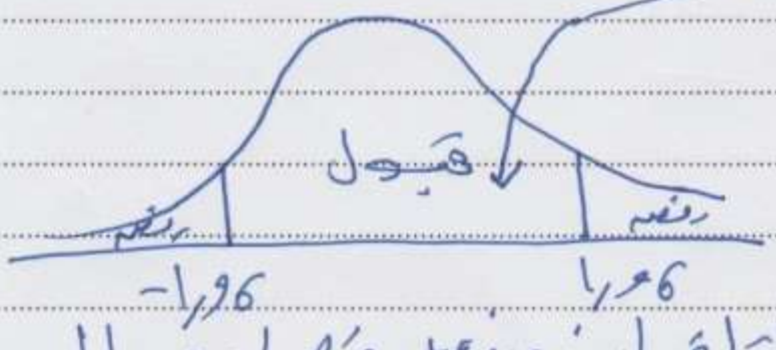
نستخدم اختبار Z لأن حجم العينة أكبر من ٣٠.

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}} = 1,5$$

3) نوجد Z الجدولية:

مستوى المعنوية 5% ⟸ درجة الثقة $= 95\%$ ⟸ $Z = 1,96$

4) القرار



∴ Z الإحصائية تقع في منطقة القبول أي نقبل فرض العدم H_0 ونرفض H_1

المحاضرة (٨)

اختبار الفروض الإحصائية المعلمية

انواع الاختبار (الفروض)

الاختبارات الاحصائية لعينة واحدة

اختبار Z-test :

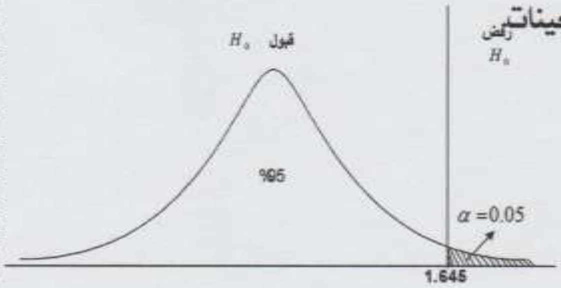
في كثير من الأحيان لا يمكن معرفة تباين المجتمع الذي سحبت منه العينة، إلا أنه إذا كان حجم العينة كبيرا ($n > 30$) فإنه يمكن استخدام تباين العينة الكبيرة (S^2) عوضا عن تباين المجتمع (σ^2) الغير معلوم، وذلك لأن (S^2) مقدر جيد (σ^2) ولأنه لا يتغير كثيرا من عينة لأخرى ما دام حجم العينة كبير، ففي هذه الحالة يمكننا استخدام اختبار (Z) لاختبار الفرضيات الصفرية موضع الدراسة وذلك من خلال المختبر الإحصائي التالي:

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

ويعتبر ذلك مدخل ضروري لفهم اختبار t-test .

مثال

إذا كان متوسط استهلاك الفرد السعودي من الدجاج حسب تقارير وزارة الصحة هو (١٢) كيلوجرام بانحراف معياري (٦) كيلوجرامات لفترة السبعينات الميلادية، أجرى أحد الباحثين دراسة في عام ٢٠٠٢م من عينة قوامها (٤٩) فردا ووجد أن متوسط الاستهلاك للفرد هو (١٤) كيلوجرام هل تشير الدراسة الحالية أن متوسط الاستهلاك أرتفع عما عليه في السبعينات؟



١) فرض العدم والفرض البديل.

فرض العدم: H_0: $\mu = 12$

الفرض البديل: H_1: $\mu > 12$

٢) مستوى الدلالة =(0.05):

٣) إحصائية الاختبار (Z):

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{14 - 12}{6 / \sqrt{49}} = 2.33$$

٤) تحديد قيمة Z المعيارية من الجدول عند مستوى دلالة (٠,٠٥)، نحتاج لتحديد قيمة $Z\alpha$ التي تقع علي اليمين وتساوي ١,٦٤٥ (أنظر الشكل المقابل):

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 دليلا كافيا على أن متوسط استهلاك الفرد من لحوم الدواجن في الوقت الحالي قد اختلف بمستوى معنوي أو ذو دلالة عما علية في سبعينات القرن الماضي.

اختبار t-test :

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

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

مثال على اختبار t:

لو كانت لدينا عينة عشوائية تتكون من ٢٥ طالب وجد أن الوسط الحسابي لأطوال طلاب العينة ١٥٥,٩٥سم، والانحراف المعياري = ٢,٩٤ سم، علما بأن الوسط الحسابي لأطوال طلاب الجامعة يبلغ ١٥٨ سم، اختبر أهمية الفرق المعنوي بين الوسط الحسابي لأطوال طلاب العينة والوسط الحسابي لأطوال طلاب الجامعة.

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 ($\mu \neq \mu_0$)

الموضوع: تمرين على المحاضرة ٨

مثال على اختبار t :

الحل

$n = 25$

$\bar{x} = 155,95$

$s = 2,94$

$\mu = 158$

① الفروض

فرض العدم : $H_0 : \mu = 158$

الفرض البديل : $H_1 : \mu \neq 158$

② إحصاء الاختبار : نستخدم اختبار t حيث حجم العينة أقل من 30

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{155,95 - 158}{\frac{2,94}{\sqrt{25}}} = -3,48$$

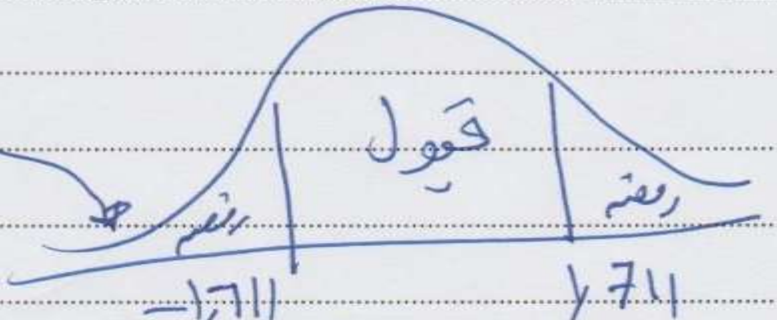
③ نوجد t الجدولية :

مستوى المعنوية بفرضين 5%

درجة حرية $df = n-1 = 25-1 = 24$

$t_{\text{الجدولية}} = 1,711$

④ القرار



رفض فرض العدم وقبول الفرض البديل H_1

مستوى الدلالة: α = ٠,٠٥

منطقة الرفض: قيمة (ت) الجدولية عند مستوى دلالة α ٠,٠٥ ودرجات حرية ٢٤٩ = ١,٩٦٠

المختبر الإحصائي:

$\overline{X}$ = ١٥٥,٩٥سم، n =٢٥ طالب، S = ٢,٩٤ سم

μ = ١٥٨سم

$$t=\frac{\overline{X}-\mu}{S/\sqrt{n}} = \frac{155.95-158}{2.94/\sqrt{25}} = -3,48$$

القرار:

∴ قيمة ت المحسوبة (-١١,[illegible]) أكبر من قيمة ت المجدولة (١,٩٦) عند مستوى دلالة α = ٠,٠٥.

∴ نرفض الفرضية الصفرية ونقبل البديلة.

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

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

- ✓ قم بإدخال البيانات المراد تحليلها من خلال شاشة تحرير البيانات Data Editor بالطريقة المناسبة كالتالي:
- ✓ أنقر بعد ذلك على زر موافق OK سيؤدي ذلك إلى تنفيذ الاختبار، وستلاحظ ظهور النتائج في شاشة المخرجات كالتالي:

➡ T-Test

نبحث عن sig

One-Sample Statistics

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

95-100

= ٠,05

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 = -١١,٠٠٦، ودرجات الحرية df = ٢٤٩، وقيمة Sig. (2-tailed) =٠,٠٠٠، وبما أن قيمة Sig. (2-tailed) في الجدول (٠,٠٠٠) أصغر من قيمة α = ٠,٠٥ فإننا بالتالي نرفض الفرضية الصفرية، أي أنه توجد فروق ذات دلالة إحصائية بين متوسط أطوال العينة ومتوسط أطوال طلاب الجامعة.

قيمة sig ← 000

القاعدة للجدول المحسوب

أكبر

مستوى المعنوية > sig

رفض الفرض العدم الصفري

وقبول الفرض البديل

أصغر

sig > مستوى المعنوية

قبول الفرض العدمي الصفري

ورفض الفرض البديل

الاختبارات الاحصائية لعينتين مستقلتين

مثال :-

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

المجموعة التجريبية	المجموعة الضابطة
n_1 = ٢٥	n_2 = ٢٥
$\overline{X_1}$ = ٧,٦٠	$\overline{X_2}$ = ٦,٠٠
S_1 = ٢,٢٧	S_2 = ١,٧٨

فهل تدل هذه البيانات على أن أداء المجموعة التجريبية كان أفضل من أداء المجموعة الضابطة عند مستوى α = ٠,٠٥ ؟

أكبر

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 = ٠,٠٥.

منطقة الرفض : ∴ قيمة مستوى الدلالة α = ٠,٠٥ والاختبار بذيل واحد ، ودرجات الحرية = ٢٥ + ٢٥ - ٢ = ٤٨ ، بذلك تكون قيمة (ت) الجدولية = ١,٦٨

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:

إذا التباين يساوي:

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

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

إذن الإنحراف المعياري يساوي : $S=\sqrt{S^2}=\sqrt{4.16}=2.04$

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

$$t=\frac{\overline{X_1}-\overline{X_2}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{7.60-6.0}{2.04\sqrt{\frac{1}{25}+\frac{1}{25}}}=2.77$$

القرار :

∴ قيمة (t) المحسوبة (٢,٧٧) أكبر من قيمة (ت) المجدولة (١,٦٨) عند مستوى دلالة α = ٠,٠٥ .

∴ نرفض الفرضية الصفرية ونقبل البديلة

أي أن المجموعة التي خضعت للتجربة يصبح أداؤهم أفضل في عملية اتخاذ القرار من الذين لم يخضعون للتجربة وذلك عند مستوى دلالة α ٠,٠٥ .

المعطيات :

الضابطة	التجريبية
$n_2 = 25$	$n_1 = 25$
$\bar{x}_2 = 6,0$	$\bar{x}_1 = 7,60$
$s_2 = 1,78$	$s_1 = 2,27$

مستوى المعنوية $a = 0.05$

الحل :

[1] الفروض الإحصائية

* فرض العدم

$$H_0 : \mu = \mu_0$$
$$\bar{x}_1 = \bar{x}_2$$

* الفرض البديل

$$H_1 : \bar{x}_1 > \bar{x}_2$$

[2] إحصاء الاختبار :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

[illegible]

$$S^2 = \frac{\left[(n_1-1)s_1^2\right] + \left[(n_2-1)s_2^2\right]}{(n_1+n_2)-2}$$

$$S^2 = \frac{\left[(25-1)\cdot 2,27^2\right] + \left[(25-1)\cdot 1,78^2\right]}{(25+25)-2} = 4,16$$

$$s = \sqrt{4,16} = 2,039$$

$$t = \frac{7,60 - 6,0}{2,039 \cdot \sqrt{\frac{1}{25} + \frac{1}{25}}} = 2,77$$

③ إيجاد t الجدولية:

$df = n_1 + n_2 - 2$

$= 25 + 25 - 2 = 48 = df$

$.05 = \alpha$

$t_{الجدولية} = 1.684$

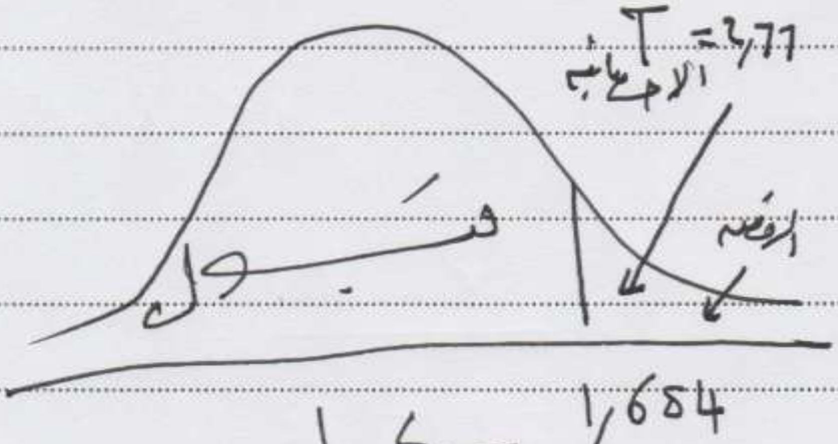
الاختبار من طرف واحد فقط

لأن الفرض البديل $H_1: \bar{X}_1 > \bar{X}_2$

ملاحظة: وكذلك إذا كانت $H_1: \bar{X}_1 < \bar{X}_2$

أما إذا كان $H_1: \bar{X}_1 \neq \bar{X}_2$

فإن الاختبار من طرفين.



④ القرار

$\therefore$ T الإحصائية أكبر من T الجدولية

أي تقع في منطقة الرفض

$\therefore$ نرفض العدم H_0 ونقبل البديل H_1

أي أن $\bar{X}_1 > \bar{X}_2$

الاختبارات الاحصائية لعينتين غير مستقلتين (العينات المرتبطة)

مثال :-

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

الإختبار القبلي	الإختبار البعدي
١٠٠ = n_1	١٠٠ = n_2
٥٤,٢٨ = $\overline{X_1}$	٥٨,٦٦ = $\overline{X_2}$
٤٩ = S_1^2	٦٤ = S_2^2

فهل تدل هذه البيانات على أن أداء الطلاب التحصيلي في الكلية بعد أخذ البرنامج التدريبي كان أفضل من أدائهم قبل أخذ البرنامج التدريبي عند مستوى α = ٠,٠٥ ؟

الحل :

سيتم اختبار الفرضيات التالية :

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

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

مستوى الدلالة : α = ٠,٠٥

منطقة الرفض : قيمة مستوى الدلالة α = ٠,٠٥ والاختبار بذيلين، ودرجات الحرية = ١٠٠-١= ٩٩ ، بذلك تكون قيمة (ت) الجدولية = ١,٩٨٠

المختبر الإحصائي :

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} - 2r\left(\frac{S_1}{\sqrt{n_1}}\right)\left(\frac{S_2}{\sqrt{n_2}}\right)}}$$

إذا قيمة (ت) تساوي :

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

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

القرار :

قيمة (ت) المحسوبة (٥,٥٧) أكبر من قيمة (ت) المجدولة (١,٩٨٠) .عند مستوى دلالة α = ٠,٠٥ .

∴ نرفض الفرضية الصفرية ونقبل البديلة، أي أن للبرنامج التدريبي تأثير إيجابي على تحصيل الطلاب وأدائهم في الكلية وذلك عند مستوى دلالة ٠.٥α.

حساب اختبار (ت) لعينتين غير مستقلتين (العينات المرتبطة) : Paired Samples T-Test من خلال الـ SPSS

T-Test

الوسط الحسابي

الانحراف المعياري

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

مستوى المعنوية ٠٫٠٥

Paired Samples Test

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

نلاحظ أن برنامج الـ SPSS قام مباشرة بحساب الإحصاءات الأساسية للبيانات مثل المتوسط الحسابي للمتغير Posttest (٥٨,٦٦٠) والانحراف المعياري لنفس المتغير (٨,٠٠) ، أما المتغير Pretest فقد كان المتوسط الحسابي (٥٤,٢٨٠) والانحراف المعياري (٧,٠٠) . بالإضافة إلى ذلك تم حساب معامل ارتباط بيرسون للمتغيرات موضع الدراسة Paired Sample Correlation وقد كانت قيمته (٠,٤٥٨) . ثم بعد ذلك قام البرنامج بحساب قيمة (ت) للمتغيرات موضع الدراسة في الجدول المعنون بـ اختبار العينات المرتبطة Paired Sample Test ، ومن هذه النتائج نلاحظ أن قيمة (ت) المحسوبة t-test = ٥,٥٧٥ ، ودرجات الحرية df = ٩٩ ، وقيمة Sig. (2-tailed) = ٠,٠٠٠ ، وبما أن قيمة Sig. (2-tailed) في الجدول (٠,٠٠٠) أصغر من قيمة α = ٠,٠٥ فإننا بالتالي نرفض الفرضية الصفرية ، أي أن أداء الطلاب في الكلية بعد أخذ البرنامج التدريبي كان أفضل من أدائهم قبل أخذ البرنامج التدريبي عند مستوى α = ٠,٠٥ .

الموضوع: تمرين [الثبات المرتبط] ص ٥٥

الاختبار القبلي — الاختبار البعدي

$n_1 = 100$ — $n_2 = 100$

$\bar{X}_1 = 54,28$ — $\bar{X}_2 = 58,66$

$S_1^2 = 49$ — $S_2^2 = 64$ إذاً في عينتان

$\alpha = ,05$

$r = ,46$

الحل: يلزم هنا استخدام اختبار t رغم أن حجم العينة أكبر من ٣٠

السبب هو أن الانحراف المعياري المتاح

هو للعينة وليس للمجتمع

S — σ

١) الفروض:

$H_0: \bar{X}_1 = \bar{X}_2 \Rightarrow$ فرض العدم

$H_1: \bar{X}_1 > \bar{X}_2$ الفرض البديل

٢) إحصاء الاختبار

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

$$t = \frac{54,28 - 58,66}{\sqrt{\frac{49}{100} + \frac{64}{100} - 2\cdot,46\left(\frac{7}{\sqrt{100}}\right)\left(\frac{8}{\sqrt{100}}\right)}} = -5,58$$

$7 = \sqrt{49} = S^2$ جذر التباين

الموضوع: تابع

ص ٥٥

③ t الجدولية

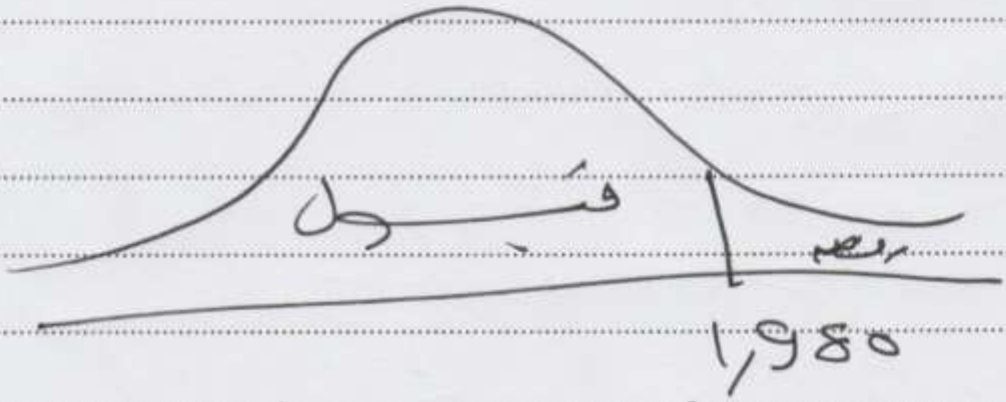
df = ١٠٠ - ١ = 99

α = 0.05

والبحث عن جدول t مرمن واحد.

$t_{\text{الجدولية}} = ١,980$

④ القرار:



∴ t الإحصائية أكبر وتقع في منطقة الرفض

∴ نرفض العدم ونقبل البديل

1,96 = Z %95

إذا علمت أن" عينة عشوائية حجمها ٤٩ شخصاً اختيرت من أفراد دولة ما ، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو ٧٥ ريال . وترغب في اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي ٧٢ ريال مقابل الفرض البديل أنه لا يساوي ٧٢ وذلك بمستوى معنوية ٥% إذا علمت أن الإنحراف المعياري لدخول الأفراد يساوي ١٤ ريال .
اجتهد في اخراج المعطيات أ/ السيد عرفة : ٠٥٠٧٣٨٠٩٣٠

٧- يمكن صياغة الفرض العدمي والفرض البديل على الشكل :.	
أ- $H_0: \mu = 72$ ، $H_A: \mu \le 72$	ب- $H_0: \mu = 72$ ، $H_A: \mu \ne 72$
ج- $H_0: \mu = 72$ ، $H_A: \mu \ge 72$	د - لا شئ مما سبق
٨- قيمة إحصائي الإختبار في هذه الحالة z تساوي : $Z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}} =$	
أ - ٠,٣	ب - ٠,٧٥
ج - ١,٥	د - لا شئ مما سبق
٩- من خلال مقارنة قيمة إحصائي الإختبار بقيمة حدود منطقتي القبول والرفض يمكن :	
أ - قبول الفرض العدمي	ب - قبول الفرض البديل
ج - عدم قبول أي من الفرضين	د - لا شئ مما سبق

قبول ‎-1,96 ‎1,96

إذا علمت أن" عينة عشوائية حجمها ٤٩ شخصاً اختيرت من أفراد دولة ما ، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو ٧٥ ريال . وترغب في اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي ٧٢ ريال مقابل الفرض البديل أنه لا يساوي ٧٢ وذلك بمستوى معنوية ١% إذا علمت أن الإنحراف المعياري لدخول الأفراد يساوي ١٤ ريال .
اجتهد في اخراج المعطيات أ/ السيد عرفة : ٠٥٠٧٣٨٠٩٣٠

١٠ يمكن صياغة الفرض العدمي والفرض البديل على الشكل :.	
أ- $H_0: \mu = 72$ ، $H_A: \mu \le 72$	ب- $H_0: \mu = 72$ ، $H_A: \mu \ne 72$
ج- $H_0: \mu = 72$ ، $H_A: \mu \ge 72$	د - لا شئ مما سبق
١١- قيمة إحصائي الإختبار في هذه الحالة z تساوي : $Z = \quad = 1,5$	
أ - ٠,٣	ب - ٠,٧٥
ج - ١,٥	د - لا شئ مما سبق
١٢- من خلال مقارنة قيمة إحصائي الإختبار بقيمة حدود منطقتي القبول والرفض يمكن :	
أ - قبول الفرض العدمي	ب - قبول الفرض البديل
ج - عدم قبول أي من الفرضين	د - لا شئ مما سبق

‎-2,58 ‎2,58 %99

إذا كان متوسط استهلاك الفرد السعودي من الدجاج حسب تقارير وزارة الصحة هو ١٢ كيلوجرام بإنحراف معياري ٦ كيلوجرام لفترة التسعينات الميلادية ، أجري أحد الباحثين دراسة في عام ٢٠٠٣ م من عينة قوامها ٤٩ فرداً ووجد أن متوسط الإستهلاك للفرد هو ١٤ كيلوجرام . هل تشير الدراسة الحالية أن متوسط الإستهلاك ارتفع عما عليه في السبعينات .
اجتهد في اخراج المعطيات أ/ السيد عرفة : ٠٥٠٧٣٨٠٩٣٠

١٩- يمكن صياغة الفرض العدمي والفرض البديل على الشكل :.	
أ- $H_0: \mu = 12$ ، $H_A: \mu \le 12$	ب- $H_0: \mu = 12$ ، $H_A: \mu \ne 12$
ج- $H_0: \mu = 12$ ، $H_A: \mu \ge 12$	د - لا شئ مما سبق
٢٠ قيمة إحصائي الإختبار في هذه الحالة z تساوي : $Z = \frac{14 - 12}{\frac{6}{\sqrt{49}}} = 2,33$	
أ - ٢	ب - ٠,٧٥
ج - ٢,٣٣	د - لا شئ مما سبق
٢١ من خلال مقارنة قيمة إحصائي الإختبار بقيمة حدود منطقتي القبول والرفض يمكن :	
أ - قبول الفرض العدمي	ب - قبول الفرض البديل
ج - عدم قبول أي من الفرضين	د - لا شئ مما سبق

1,96

إذا علمت أنه :
" لو كانت لدينا عينة عشوائية تتكون من ٢٥٠ طالب وجد أن الوسط الحسابي لأطوال طلاب العينة ١٥٥,٩٥ سم والإنحراف المعياري = ٢,٩٤ سم ، علماً بأن الوسط الحسابي لأطوال طلاب الجامعة يبلغ ١٥٨ سم ، اختبر أهمية الفرق المعنوي بين الوسط الحسابي لأطوال طلاب العينة واوسط الحسابي لأطوال طلاب الجامعة "
لا يساوي
اجتهد في اخراج المعطيات أ/ السيد عرفة : ٠٥٠٧٣٨٠٩٣٠

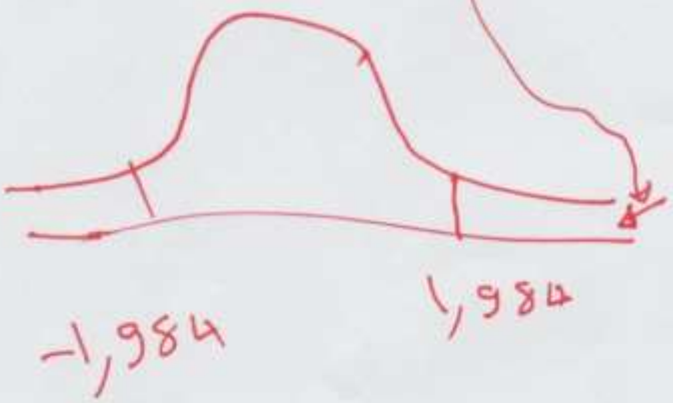
٢٢	يمكن صياغة الفرض العدمي والفرض البديل على الشكل .:
أ- $H_0 : \mu = \mu_0$ ، $H_A : \mu \neq \mu_0$	ب- $H_0 : \mu = \mu_0$ ، $H_A : \mu \leq \mu_0$
ج- $H_0 : \mu = \mu_0$ ، $H_A : \mu \geq \mu_0$	د – لا شئ مما سبق

-٢٣	يسمى إحصائي الإختبار في هذه الحالة :
أ- Z	ب- t
ج- H	د – لا شئ مما سبق

-٢٤	قيمة إحصائي الإختبار في هذه الحالة تساوي :
أ- ٢,٠٥	ب- ٢,٩٤
ج- ١١,٠٢	د – لا شئ مما سبق

$$\frac{155,25-158}{\frac{2,94}{\sqrt{250}}} = 11,02$$

-٢٥	من خلال مقارنة قيمة إحصائي الإختبار بقيمة حدود منطقتي القبول والرفض يمكن :
أ- قبول الرفض العدمي	ب- قبول الفرض البديل
ج- عدم قبول أي من الرفضين	د- لا شئ مما سبق



df = n-1 = 250-1 = 249

a = ,05

t = 1,984

الموضوع: قوانين المحاضرة ٩

البيان	مجموع المربعات	درجات الحرية	التباين
الكلي	= مجموع مربعات X − $\dfrac{(\text{مجموع كل قيم } X)^2}{\text{عدد أفراد المجموعة} \times \text{عدد المجموعات}}$	= إجمالي عدد أفراد المجموعات − 1 $= (K \cdot n) - 1$	—
بين المجموعات	$= \dfrac{(\sum X_1)^2}{n_1} + \dfrac{(\sum X_2)^2}{n_2} + \dfrac{(\sum X_3)^2}{n_3} - \dfrac{(\sum X_1 + \sum X_2 + \sum X_3)^2}{K \cdot n}$	= عدد المجموعات − 1 $= K - 1$	= $\dfrac{\text{مجموع المربعات بين المجموعات}}{\text{درجات الحرية بين المجموعات}}$
داخل المجموعات	= مجموع المربعات الكلي − مجموع المربعات بين المجموعات	= إجمالي عدد أفراد المجموعات − عدد المجموعات $= (n \cdot K) - K$	= $\dfrac{\text{مجموع المربعات داخل المجموعات}}{\text{درجات الحرية داخل المجموعات}}$

$$F = \frac{\text{التباين بين المجموعات}}{\text{التباين داخل المجموعات}}$$

علما بأن:

K = عدد المجموعات

n = عدد أفراد المجموعة الواحدة

المحاضرة (٩)

اختبار الفروض الإحصائية المعلمية

تحليل التباين الأحادي

One Way ANOVA

تحليل التباين الأحادي (مستوى واحد) :-

هو طريقة لاختبار معنوية الفرق بين المتوسطات لعدة عينات بمقارنة واحدة، ويعرف أيضا 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 كلا لها توزيع طبيعي.

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

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

مثال (١) :-

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

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

المطلوب: هل هناك فروق ذات دلالة بين المنتجات الثلاثة ؟

- **الحل : وضع فرض العدم والفرض البديل.**
- **صياغة الفرضية الصفرية كالتالي:** $H0: \mu_1 = \mu_2 = \mu_3$
- **في حين تفترض الفرضية البديلة التالي : متوسطان على الأقل غير متساويين :HA**
- **تحديد مستوى الدلالة (α): وتحدد مستويات المعنوية سلفا وهي عادة 0.05 أو 0.01**

حساب إحصائية الاختبار (F) وذلك من خلال اتباع الخطوات التالية

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

المجاميع

- **المتوسط الحسابي لـ** $\bar{X}_1 = \frac{50}{5} = 10$
- **المتوسط الحسابي لـ** $\bar{X}_2 = \frac{35}{5} = 7$
- **المتوسط الحسابي لـ** $\bar{X}_3 = \frac{20}{5} = 4$

١- مجموع المربعات الكلي Total Sum of Squares -

حيث ng تعني عدد أفراد المجموعة المحددة، K تعني عدد المجموعات موضع الدراسة

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

٢- مجموع المربعات بين المجموعات Between Sum of Squares

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

٣- مجموع المربعات داخل المجموعات Within Sum of Squares -

- **مجموع المربعات داخل المجموعات** *Within SS* **= مجموع المربعات الكلي - مجموع المربعات بين المجموعات**
- **= ١٤٤ - ٩٠ = ٥٤**

٤ـ نحسب درجات الحرية:

أ. درجات الحرية بين المجموعات Between groups

$(K-1)=3-1=2$ degrees of freedom

ب. درجات الحرية داخل المجموعات Within groups

$(n-K)=15-3=12$ degrees of freedom

ج. درجات الحرية الكلية Total degrees of freedom $(n-1)=15-1=14$

٥ـ لحساب قيمة F نتبع ما يلي :

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

$$Beween..groups..mean..square=\frac{Between..SS}{K-1} \quad Beween..groups..mean..square=\frac{90}{2}=45$$

✓ 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 = $\frac{\text{التباين بين المجموعات}}{\text{التباين داخل المجموعات}}$

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

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

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات Means	قيمة F
بين المجموعات Between groups	٩٠	٢	٤٥	١٠
داخل المجموعات Within groups	٥٤	١٢	٤,٥	
الكلي المجموع Total	١٤٤	١٤		

وبالرجوع إلى جدول توزيع F نجد أن القيمة الحرجة لـ F بدرجات حرية للبسط تساوي ٢ ودرجات حرية للمقام تساوي ١٢ وباستخدام مستوى- ٠.٠٥ نجد أن القيمة الحرجة تساوي ٣,٨٨ .

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

الموضوع: المحاضرة الثانية عشر.

تم بقياس الفروض الإحصائية لـ ثلاث متغيرات منفصلة هنا نلجأ إلى اختبار F

مثال: ونتبع خطوات الحل التالي

1) الفروض:

$H_0: X_1 = X_2 = X_3$ ← فرض العدم

H_1: ← الفرض البديل هناك بدائل مختلفة عن بعضها

$X_1 \neq X_2$ أو $X_2 \neq X_3$ أو $X_1 \neq X_3$

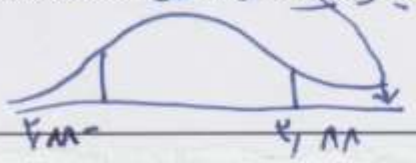
2) نوجد إحصاء الاختبار F كما في الورقة المرفقة:

المصادر	مجموع المربعات	درجات الحرية	التباين
الكلي	$= (١٠٢٤ + ٢٦٣ + ٥١٤) - \frac{(٢٠+٣٥+٥٠)^2}{٣ \times ٥}$ $= ١٤٤$	$= (٥ \times ٣) - ١$ $= ١٥ - ١ = ١٤$	—
بين المجموعات	$\frac{(٥٠)^2}{٥} + \frac{(٣٥)^2}{٥} + \frac{(٢٠)^2}{٥} - \frac{(٢٠+٣٥+٥٠)^2}{٣ \times ٥}$ $= ٩٠$	$= ٣ - ١ = ٢$	$= \frac{٩٠}{٢} = ٤٥$
داخل المجموعات	$= ١٤٤ - ٩٠ = ٥٤$	$= ١٥ - ٣ = ١٢$	$\frac{٥٤}{١٢} = ٤٫٥$

$$F = \frac{45}{4.5} = 10$$

3) F الجدولية دائماً معطاة في السؤال.

4) القرار: نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 أي هناك فروق واختلاف



مثال (٢) :- قام أحد الباحثين بتفريغ ما تم الحصول عليه من معلومات في جدول تحليل التباين كالتالي :

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات Means	قيمة F
بين المجموعات Between groups	200	10	$20 = \frac{200}{10}$	4
داخل المجموعات Within groups	50	10	5	
الكلي (المجموع) Total	250	20		

$= 250 - 200 = 50$

قيمة إحصائي الاختبار F تساوي :-

أ) 10　　　ج) 80　　　4

ب) 5　　　د) لا شيء مما سبق

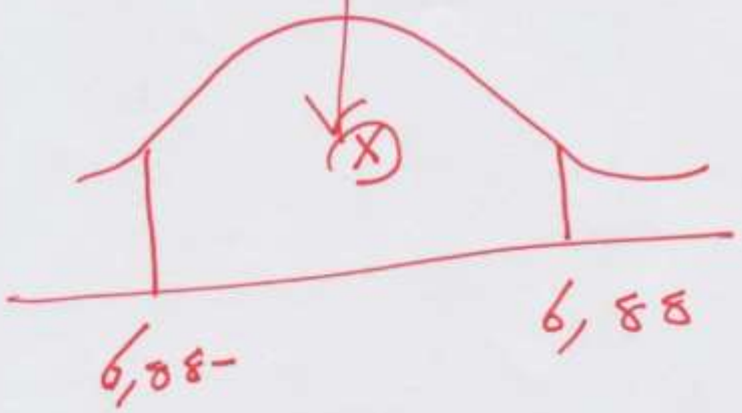
من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض (إذا علمت أن قيمة F الجدولية تساوي 6.88) يمكن :-

أ) قبول الفرض البديل.

ب) قبول الفرض العدمي.

ج) عدم قبول أي من الفرضين.

د) لا شيء مما سبق



المحاضرة (١٠)

إختبار الفروض الإحصائية المعلمية

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

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

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

الإرتباط الجزئي

Partial Correlation

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

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

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

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

مثال:

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

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

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

المطلوب:

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

الحل:

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

- معامل ارتباط بيرسون بين الغياب والتحصيل الدراسي ونرمز له ١,٢ أي معامل الارتباط بين المتغير (١) والمتغير (٢)
- معامل ارتباط بيرسون بين الغياب وطريقة التدريس ونرمز له ١,٣ أي معامل الارتباط بين المتغير (١) والمتغير (٣)
- معامل ارتباط بيرسون بين التحصيل الدراسي وطريقة التدريس ونرمز له ٢,٣ أي معامل الارتباط بين المتغير (٢) والمتغير (٣)
- ويتم حساب معامل ارتباط بيرسون من خلال العلاقــة التالية:

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

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

معامل ارتباط بيرسون بين الغياب والتحصيل الدراسي ونرمز له ر ١,٢

معامل ارتباط بيرسون بين الغياب وطريقة التدريس ونرمز له ر ١,٣

$$r_{1.3}=\frac{5875-\frac{(500)(60)}{5}}{\left[\sqrt{(51450)-\frac{(500)^2}{5}}\right]\left[\sqrt{(748)-\frac{(60)^2}{5}}\right]}=\frac{5875-6000}{\left(\sqrt{51450-50000}\right)\left(\sqrt{748-720}\right)}$$

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

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

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

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

معامل ارتباط بيرسون بين التحصيل الدراسي وطريقة التدريس ونرمز له ر ٢,٣

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

$$r_{2.3}=\frac{2148-\frac{(60)(174)}{5}}{\left[\sqrt{(748)-\frac{(60)^2}{5}}\right]\left[\sqrt{(10030)-\frac{(174)^2}{5}}\right]}=\frac{2148-2088}{\left(\sqrt{748-720}\right)\left(\sqrt{10030-60552}\right)}$$

$$=\frac{60}{\left(\sqrt{28}\right)\left(\sqrt{39748}\right)}=\frac{60}{(5.292)(63.046)}=\frac{60}{333.639}=+0.179$$

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

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

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

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

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

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

الموضوع: المحاضرة العاشرة

[اختبار الفروض الإحصائية المعلمية]

مؤشر معامل الارتباط.

معامل الارتباط: r

-1 ارتباط عكسي — 0 لا يوجد ارتباط — 1 ارتباط طردي

لحساب معامل الارتباط بين الغياب والتحصيل الدراسي مع تشتت طريقة التدريس بتمرين خاص بالمحاضرة العاشرة ص ٦٦.

$\boxed{AC}$ ادخال البيانات $MODE \rightarrow 3 \rightarrow 2$

$shift \rightarrow 1 \rightarrow Reg \rightarrow \overset{r}{3} =$

الحل بالحاسبة

$r_{1,2} = -0.62$

$r_{1,3} = 0.22$

$r_{2,3} = 0.179$

ثم نوجد معامل ارتباط لثلاث متغيرات معاً

$$r_{1,2,3} = \frac{r_{1,2} - [r_{1,3} \cdot r_{2,3}]}{\sqrt{[1-(r_{1,3})^2][1-(r_{2,3})^2]}}$$

$$r_{1,2,3} = \frac{-0.62 - [0.22 \cdot 0.179]}{\sqrt{[1-(0.22)^2][1-(0.179)^2]}} = -0.68$$

ارتباط عكسي

اختبار معنوية معامل الارتباط

Significance Of Correlation Coefficient

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

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

اختبار أن معامل ارتباط المجتمع يساوي الصفر:

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

١ − **القرض العدمي:** أن معامل ارتباط المجتمع يساوي صفر، أي لا يوجد ارتباط بين المتغيرين. وبالرموز: $H_0: R=0$

٢ − **الفرض البديل:** معامل ارتباط المجتمع لا يساوي صفر، أي ويجد ارتباط بين المتغيرين، وبالرموز: $H_A: R\neq 0$

٣ − **إحصائية الاختبار:** ستكون إحصائية الاختبار في هذه الحالة هي t والتي تأخذ الشكل التالي: $T=\frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$

والتي لها توزيع t بدرجات حرية n − 2.

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

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

مثال:

اختبر معنوية معامل الارتباط لو كان لدينا البيانات التالية:

r = 0.91 ، n = 10

وذلك بمستوى معنوية 5 %.

الحل:

لو كان لدينا البيانات التالية: r = 0.91 ، n = 10

وتكون خطوات اختبار معنوية الارتباط كما يلي:

١ − **الفرض العدمي:** $H_0: R=0$

٢ − **الفرض البديل:** $H_A: R\neq 0$

٣ − **الإحصائية:**

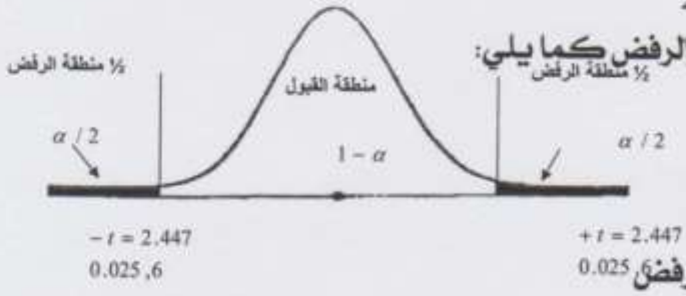
$$t=\frac{r}{\sqrt{\frac{1-r^2}{n-2}}}=\frac{0.91}{\sqrt{\frac{1-(0.91)^2}{10-2}}}=\frac{0.91}{\sqrt{\frac{0.1719}{8}}}=\frac{0.91}{\sqrt{0.0215}}=\frac{0.91}{0.1466}=6.208$$

إذا: $t=6.208$

٤ – حدود منطقتي القبول والرفض:

من جدول t حيث مستوى المعنوية $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$ ودرجات الحرية تساوي (n – 2 = 10 – 2=8)

نجد أن قيمة t تساوي ٢,٤٤٧ وتكون حدود منطقتي القبول والرفض كما يلي:



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

الموضوع: تابع معاملات الارتباط واختبار الفروض

اختبر معنوية معامل الارتباط وفقاً للمعطيات التالية

$r = 0{,}91$

$n = 10$

مستوى المعنوية = 5%

الحل

[1] الفروض:

$H_0 : R = 0$ أي لا يوجد ارتباط ← فرض العدم

$H_1 : R \neq 0$ يوجد ارتباط ← الفرض البديل

[2] إيجاد اختبار الإحصائي:

دائماً نستخدم مع معامل الارتباط

اختبار t

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0{,}91}{\sqrt{\frac{1-0{,}91^2}{10-2}}} = 6{,}20$$

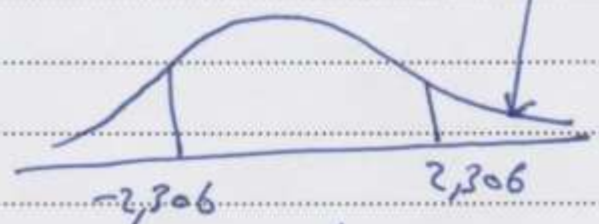
[3] نوجد t الجدولية:

$df = n - 2 = 10 - 2 = 8$ ← df درجة الحرية

$a = 0{,}05$ ← a مستوى المعنوية

$t = 2{,}306$

[4] القرار

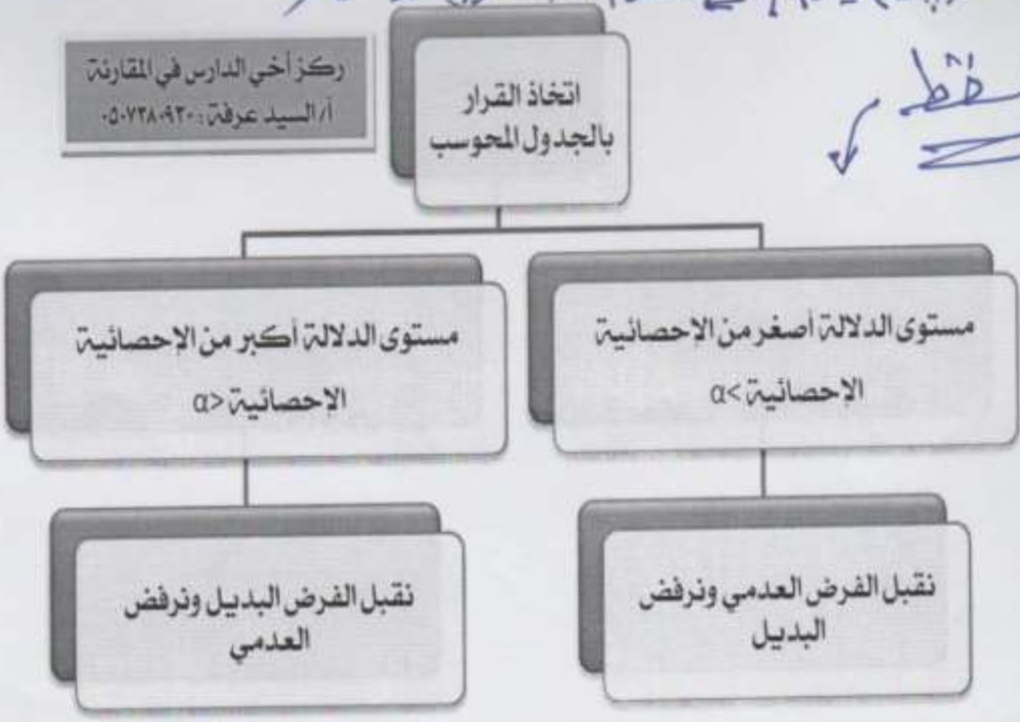


نرفض العدم H_0 ونقبل البديل أي يوجد ارتباط ذو معنوية

درجة الثقة = ٩٥٪ ← مستوى الدلالة (معنوي) = ٠.٠٥

مخطط

ركز أخي الدارس في المقارنة
أ/ السيد عرفة: ٠٥٠٧٣٨٠٩٢٠



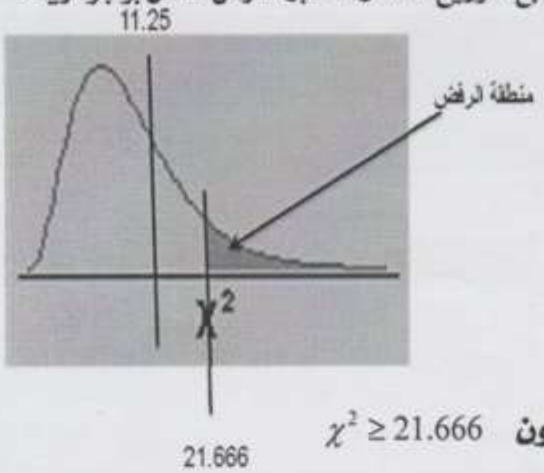
X

المحاضرة (١٢)

تابع الاختبارات الاحصائية اللامعلمية

مثال (٢) :-

إذا علمت أن تباين قوة مقاومة الكسر للكابلات التي تنتجها إحدى الشركات لاتزيد عن 40000 ، وتستخدم الشركة الآن طريقة إنتاج جديدة يعتقد أنها ستزيد من تباين قوة مقاومة الكابلات للكسر، سحبت عينة عشوائية من عشرة كابلات فوجد تباينها يساوي 50000 . بافتراض أن قوة مقاومة الكسر للكابلات تتبع التوزيع المعتدل، اختبر الفرض القائل بوجود زيادة معنوية في التباين عند مستوى معنوية $\alpha = 0.01$.



الحل :-

- وضع فرض العدم والفرض البديل.

صياغة الفرضية الصفرية كالتالي: $H0: \sigma^2 \le 40000$

في حين تفترض الفرضية البديلة التالي : $H_A: \sigma^2 > 40000$

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

لذا فإن قاعدة القرار هي أن يتم رفض الفرضية الصفرية H0 عندما تكون $\chi^2 \ge 21.666$

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

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

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

الجدول الثالث يبين نتيجة اختبار مربع كاي كالتالي:

قيمة الاختبار　　　　درجة الحرية

إجباً مربع كاي ٢　　　　هام

Chi-Square Tests

	Value قيمة	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.437a	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.

مستوى دلالة الاختبار

يبين الجدول السابق أن قيمة اختبار مربع كاي هي ٢,٤٣٧ بدرجة حرية مقدارها ٤

يتبين لنا من الجدول أن أقل قيمة لمستوى الدلالة هي Asymp. Sig. (2-sided) = ٠,٦٥٦ وهي اكبر من مستوى الدلالة ٠,٠٠٥ = α وبالتالي لا نستطيع رفض الفرضية الصفرية أي أن تقدير الطالب لا يعتمد على جنسه.

sig　　　　مستوى المعنوية

,656 > ,05

القرار قبول الفرض العدمي

تمرين :

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

Chi-Square Test

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

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

١) قيمة إحصائي الاختبار كا٢ تساوي :-

أ) .2384

ب) 1.9672

ج) 1.9496

د) لا شيء مما سبق

٢) قيمة مستوى الدلاله المحسوبة للإختبار تساوي :-

أ) .0437

ب) .0434

ج) .0390

د) لا شيء مما سبق

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

أ) قبول الفرض البديل .

ب) قبول الفرض العدمي .

ج) عدم قبول أي من الفرضين .

د) لا شيء مما سبق

sig مستوى المعنوية

,0437 < ,05

تابع الاختبارات الاحصائية اللامعلمية

٢اختبار مان وتني Mann – Whitney

استخدامه:

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

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

مثال (١) :-

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

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

١٠	١٤	٧	٨	١٦
٣	٧	١٥	١٤	٧

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

١٣	٦	٥	١٢	٣
١٠	١١	١٠	١٠	١٤

المطلوب:

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

أولا: ندخل البيانات كالتالي:

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

اختبار مان وتني

هام

sig

,648 > ,05

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]

Ranks

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

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

مثال (٢) :-

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

Test Statistics

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

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

(أ) كا٢ .

(ب) مان وتني .

(ج) ويلكوكسون .

(د) لا شيء مما سبق

sig 2tail

(٢) قيمة إحصائي الاختبار تساوي :-

(أ) -.037

(ب) .028

(ج) .034

(د) لا شيء مما سبق

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

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق .

sig α

$.028 < .05$

٣ـ إختبار ويلكوكسون Wilcoxon Test

استخدامه:

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

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

مثال :ـ تأثير ممارسة الرياضة على إنقاص الوزن:

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

المطلوب:

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

Ranks

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

Test Statistics[b]

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

قام البرنامج بحساب الفروق في الوزن على أساس التالي:

الفرق = الوزن بعد ممارسة الرياضة – الوزن قبل ممارسة الرياضة ويلاحظ أيضا: أن متوسط الرتب السالبة (٤,٩٣) أكبر من متوسط الرتب الموجبة (١,٥)، وهذا معناه أن متوسط الوزن قبل ممارسة الرياضة أكبر من متوسط الوزن بعد ممارسة الرياضة (إذا فى غاية الأهمية أن نعرف الترتيب الذى إستخدمه البرنامج للعينتين)

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

Sig

.021 < .05

قبول البديل

57

مثال :-

إذا علمت أنه :-

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

Ranks

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

Test Statistics

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

sig α

.421 > .05

الحل :-

(١) من الجداول السابقة يمكن توضيح أن :-

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

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

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

(د) لا شيء مما سبق

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

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي . ← إختلاف = أفضل

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق .

٤اختبار كروسكال واليس Kruskal-Wallis Test

استخدامه:

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

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

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

المطلوب:

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

هام

Ranks

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

Test Statistics[a,b]

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

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

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

sig a

$0.095 > 0.05$

قبول الفرض العدمي

مثال :ـ

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

هام

Test Statistics

sig   α

.025 < .05

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

(١) من الجدول السابق يمكن :ـ

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

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

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

(د) لا شيء مما سبق.

٤ـ حساب اختبار كولومجروف سيمرنوف لجودة التوافيق

Goodness of Fit Test ـ Kolmogorov-Smirnov

من خلال برنامج SPSS

استخدامه:

يستخدم هذا الاختبار لمعرفة إذا ما كانت العينة موضع الاهتمام تتبع توزيعا احتماليا معينا ويستخدم عوضا 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 أيضا في حالة كون التوزيع الاحتمالي لمتغير متصل.

NPar Tests

هام

One-Sample Kolmogorov-Smirnov Test

		Dinner	
N		50	حجم العينة
Normal Parameters a,b	Mean	15.26	متوسط البيانات
	Std. Deviation	6.782	الانحراف المعياري للبيانات
Most Extreme Differences	Absolute	.081	أكبر فرق بين البيانات ودالة التوزيع الاحتمالية
	Positive	.081	
	Negative	-.069	
Kolmogorov-Smirnov Z		.573	قيمة اختبار جودة المطابقة
Asymp. Sig. (2-tailed)		.898	مستوى دلالة الاختبار

a. Test distribution is Normal.

b. Calculated from data.

تبين النتائج أعلاه أن متوسط عدد الزبائن هو ١٥.٢٦ بانحراف معياري قدره ٦.٧٨٢ وأن قيمة اختبار كولموجروف سميرنوف لجودة المطابقة هو ٠.٥٧٣.

القرار:

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

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

sig   α

.898 > .05

60

قبول العدم .

الموضوع: الجداول الحوسبة (١١ - ١٢ - ١٣)

هام جداً

sig (2 tail)

a

مستوى الدلالة

مستوى المعنوية

أصغر

قبول الفرض العدمي = sig > a

أكبر

قبول البديل = sig < a

تمت بحمد الله، اسأل الله لي ولكم التوفيق والسداد.