الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة محمد البشير الإبراهيمي برج بوعريريج

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

UNIVERSITE MOHAMED EL BACHIR EL IBRAHIMI
BORDJ BOU ARRERIDJ

محاضرات في مقياس تحليل قواعد البيانات

مدعمة بسلاسل وتمارين الأعمال الموجهة

مطبوعة محكمة موجهة لطلبة السنة الأولى ماستر إدارة مالية

إعـداد:

د: ميهوب مسعود

سلسلة المطبوعات الجامعية المحكمة

تقديم:

إن دراسة الإحصاء هو في الواقع دراسة لمنهج من مناهج البحث العلمي، وذلك لا يعني أنه ليس علما قائما بذاته، بل هو علم له قوانينه وقواعده الرياضية الخاصة التي تساهم مساهمة فعالة في إمداد مختلف العلوم الأخر ى بالمعلومات الضرورية لضمان توسعها وزيادة فعاليتها في خدمة المجتمعات البشرية.

إن اتساع مجال الظواهر جعل من الصعوبة بمكان التحكم فيها والإحاطة بمختلف جوانبها، فكانت الضرورة إلى استحداث الإستدلال الإحصائي كأحد محاور علم الإحصاء بمفهومه الحديث، فعادة عند دراسة مجتمع ما فإنه قلما نستطيع جمع بيانات عن كل مفرداته وذلك لمجموعة من الإعتبارات منها: طبيعة المجتمع نفسه، التكلفة والوقت والجهد. فلا نستطيع مثلا حصر جميع الأسماك في الماء ودراسة خصائصها وعاداتها....، فالإستدلال الإحصائي يحاول تجاوز حدود هذه الإشكالية عبر در اسة عينة من ذلك المجتمع، وانطلاقا من المعلومات المحصلة من تلك العينة يمكن الإستدلال على خصائص المجتمع ككل.

إن الإستدلال على خصائص المجتمع انطلاقا من الخصائص المدروسة للعينة يطرح إشكالية أخرى تتعلق بتعميم النتائج للجزء على الكل، ففي كثير من الأحيان يمكن أن لا يحمل ذلك الجزء جميع خصائص الكل، وهو ما دفع إلى ضرورة الإعتماد على نظرية الإحتمالات كأساس يستطيع بواسطته الباحث أو متخذ القرار تحديد احتمال الخطأ الممكن الوقوع فيه نتيجة دراسة الجزء بدلا من الكل.

إن استخدام الإستدلال الإحصائي لا بد وأن يرافقه فهم عميق لسلوك الظواهر الاقتصادية وذلك بالحد الذي يمكن من تحليل سلوك تلك الظواهر، فلغة الأرقام لا تكفي وحدها لإطلاق الأحكام وتوقع السلوكيات ما لم تكن هناك خلفية نظرية تساعد على نحو واسع في تحليل تلك البيانات وتفسير المخرجات، وهو صلب القضية التي تسعى إلى تحقيقها مختلف الطرق والأساليب الإحصائية في وقتنا المعاصر الذي يميزه التطور المستمر والتغير الكبير.

انطلاقا مما سبق تأتي هذه المطبوعة لعرض مقياس تحليل البيانات التسويقية والمقررة لطلبة السنة الثالثة تسويق كلية العلوم الاقتصادية والتجارية وعلوم التسيير، نحاول من خلالها تقديم مقررات هذا المقياس بشكل منهجي وبلغة مبسطة دون التعرض للبراهين الرياضية المعقدة لزيادة القدرة الإستيعابية للطلبة وإعطاء أكثر تحكم في تحليل مختلف جوانب مخرجات الإستدلال الإحصائي، مع تدعيمها بأمثلة وتمارين تقترب أكثر من الواقع لزيادة تسهيل الفهم وحتى يتسنى تطبيق المعلومات المحصلة في الواقع العملي.

- **توصيف المقرر وأهدافه:**

يأتي هذا المقياس كمكمل للمقاييس المدروسة للسنة الأولى جذع مشترك والتي تناولت في فصلها الأول الإحصاء الوصفي (إحصاء 1) حيث تم التطرق في البداية إلى مفاهيم عامة حول الإحصاء ثم البيانات الإحصائية وطريقة عرضها وتبويبها والتمثيل البياني لها حسب نوع المتغير، ثم تناول مقاييس النزعة المركزية (الوسط الحسابي، الوسيط، المنوال) ومقاييس التشتت (المدى والتباين والانحراف المعياري والربيعيات)، وفي الختام تطرق الفصل الأول إلى الأشكال (الشكل المتماثل، الإلتواء، التفلطح والتدبدب). أما الفصل الثاني (إحصاء 2) فقد تناول الإحتمالات والتوزيعات الإحتمالية سواء للمتغير العشوائي المنفصل أو المتصل وهو ما شكل قاعدة يـ ـبنى عليها الفصل الثالث (إحصاء 3) الذي يأخذ في الحسبان المكتسبات السابقة لمحاولة التعرف على خصائص المجتمعات الإحصائية انطلاقا من بيانات عينات مسحوبة من تلك المجتمعات مع إخضاع النتائج المحصلة إلى مبادئ الإحتمالات حتى يمكن تحديد إحتمالية القبول أو الرفض لتلك النتائج مما يضفي عليها المزيد من الدقة ويـ ـمكن من الإعتماد عليها واتخاذ قرارات سليمة إلى حد كبير، ليأتي الفصل الرابع (تحليل البيانات التسويقية) والخاص بتحليل البيانات والمخرجات المحصلة في إطار الإستدلال الإحصائي المدروس في الفصول السابقة.

لكي تتحقق الاستفادة فقد تم تقديم المواضيع بلغة بسيطة، كما تم التقيد بالجوانب الشكلية والمنهجية المتعارف عليها من حيث التوثيق باستخدام الهوامش والمحافظة على التسلسل والترابط بين المحاور وأجزائها، إلا أنه للاستيعاب لابد على الطالب أن يكون متحكم في القواعد الأساسية للرياضيات.

تتمثل الأهداف التعليمية لهذا العمل في العناصر التالية: .

- التعرفعلى أنواع البيانات وطرق عرضها و تبويبها، مع التعرض لإستبانة البحث الميداني؛
- تزويد الطالب بآليات اختبارات الفروض الإحصائية؛
- التعرف على طريقة التعامل مع العلاقات بين المتغيرات واحتساب التباينات؛
- تمكين الطالب من القدرة على تحليل المخرجات وفق منظور اقتصادي.

بالنسبة للمضمون: فقد بوب هذا العمل إلى أربع محاور مشفوعة ببعض الأمثلة التطبيقية المحلولة.

المحور الأول:

أنواع البيانات وطرق جمعها وتبويبها

الأهداف التعليمية

- التعرف على أنواع البيانات الإقتصاية؛
- طرق جمع البيانات وتبويبها؛
- منهجية استبانة البحث الميداني.

تمهيد:

إن تفسير سلوك الظواهر الاقتصادية في الواقع العملي يعتبر من أصعب المهام التي يواجهها الباحثون في ميدان العلوم الاجتماعية، ففضلا عن كون هذه الظواهر تحمل في طياتها تناقضات كثيرة ناتجة عن تفاعل مجموعة غير قليلة من المتغيرات المستقلة، فإنه يضاف إلى ذلك أن عديد المتغير ات المفسرة هي في الأصل غير كمية مما يطرح إشكالات أخرى تتعلق أولا بكيفية تكميمها ثم قياس تأثيرها على المتغيرات التابعة.

إن الوضعية السابقة ناتجة بالأساس عن عدم تفسير سلوك الظواهر انطلاقا من متغير مؤثر واحد، لذلك فقد درج الباحثون في دراسة الظواهر الإقتصادية على بناء ما سمي بالمقياس Scale والذي يضم غالبا جميع المؤشرات الموضوعية التي لا تعكس بالضرورة مستوى واحدا من حيث القوة أو الشدة وإن كانت في الوقت ذاته تعبر بصدق عن المفهوم المراد قياسه.

إن عملية بناء المقياس تحتم على الباحث تحديد المؤشرات المكونة لهذا المقياس والتي تعبر بصدق وكفاية عن المفهوم المراد قياسه. فعلى سبيل المثال لو أردنا تصميم مقياس يعبر عن درجة الاستقرار الاقتصادي فلابد من استخدام مؤشرات تدل بصدق عليه مثل متغيرات المربع السحري لكالدور (النمو الاقتصادي، التضخم، البطالة وتوازن ميزان المدفوعات).. وهكذا.

بعد عملية بناء المقاييس تأتي نقطة غاية في الأهمية تتعلق بمدى تأثير كل مؤشر في الظاهرة المراد تحديد مجال تغيرها، وهو ما يستلزم الإستعانة بمختلف الأدوات الإحصائية التي من شأنها توضيح مختلف علاقات التأثير والتأثر، ومن ثمة اتخاذ القرارات واختبار مختلف الفروض التي تصاغ حول شبكة العلاقات التي تحكم المنظومة الدافعة لسلوك الظواهر في الواقع العملي.

1- أنواع البيانات العلوم الإقتصادية:

عموما، تقسم البيانات الإحصائية Statistical Data أو ما يسمى بالمقياس الإحصائي أحيانا Statistical Index إلى عدة أنواع مختلفة سنتناول فيما يلي كل واحدة منها بنبذة مختصرة.

أولا : البيانات الوصفية:

هي بيانات كيفية تصنف فيها البيانات إلى حالات أو صفات أو أسماء مختلفة لا يمكن قياسها عدديا.

مثال1: تصنيف العمال في المؤسسة إلى:

1- ذكور 2- إناث

مثال2: تصنيف عمال أحد المصانع حسب المنطقة التي ينتمون إليها:

1- الوسطى 2- الشرقية

3- الغربية 4- الشمالية

5- الجنوبية 6- أخرى

مثال3: عند دراسة الأسباب (أو العوامل) التي تدفع الطلاب للالتحاق بقسم العلوم الاقتصادية، من الممكن تصنيف تلك الأسباب كما يلي:

1- الرغبة في دراسة العلوم الاقتصادية.

2- المعدل الذي حصل عليه الطالب في مقرر المبادئ.

3- الوظيفة التي يمكن أن يعمل بها خريج القسم.

4- وجود الأصدقاء في هذا التخصص.

5- رغبة الوالدين.

6- أخرى (تذكر).

في الأمثلة السابقة – ومثيلاتها – يعبر عن البيانات بحالات أو صفات أو أنواع، فالجنس – في المثال الأول – إما ذكر أو أنثى، والعامل – في المثال الثاني – أما من الوسطى أو الشرقية أو الغربية أو...... وهكذا.

كذلك نلاحظ على البيانات الوصفية أنها حتى لو أعطيت أرقام – كما في الأمثلة السابقة – فإن هذه الأرقام لا تعني قيما مختلفة، فالرقم (2) لا يعني أنه ضعف الرقم (1) وبالتالي لا يمكن إجراء أية عمليات حسابية عليها (جمع أو طرح أو ضرب وقسمة)، ولا إجراء عمليات مقارنة (أيها أكبر أو أصغر).

ونقطة أخرى جديرة بالملاحظة هنا، ألا وهي أنه على الرغم من الأمثلة السابقة قد صنفت فيها البيانات تبعا لمتغير اسمي واحد فإنه من الممكن تصنيفها في أمثلة أخرى تبعا لأكثر من متغير اسمي كما في المثال التالي:

مثال4: لو أردنا قياس العلاقة بين النوع (ذكر – أنثى) والذوق بالنسبة لمنتج معين (لديه ذوق – ليس لديه ذوق) على مجموعة تتكون من ثلاثين شخصا فإننا في هذه الحالة سنجد أن لدينا أربع مجموعات كما في الجدول أدناه.

توزيع افتراضي للعلاقة بين الذوق لمنتج معين والنوع لثلاثين شخصا

النوع / الذوق	ذكر	أنثى	المجموع
لديه ذوق	7	5	12
ليس لديه ذوق	12	6	18
المجموع	19	11	30

ثانيا : البيانات الترتيبية Ordinal Data

البيانات الترتيبية والتي يطلق عليها أحيانا وصف المقياس الترتيبي Ordinal Data هي البيانات الوصفية التي يمكن ترتيبها تصاعديا أو تنازليا.

مثال1: التقدير الذي يحصل عليه الطالب في مقياس إحصاء3:

1-ضعيف.

2- مقبول.

3- جيد.

4- جيد جدا.

5- ممتاز.

واضح من هذا المثال أن الرقم 5 أعلى في قيمته من باقي الأرقام. فالطالب الذي يحصل على رقم " أكبر " يعني أن معدله " أعلى " وأن مستواه أفضل من الطالب الذي يحصل على رقم أصغر.

مثال2: عند دراسة مدى مشاركة شريحة من المجتمع في النشاطات الاقتصادية يمكن تصنيف مدى المشاركة كما يلي:

1- لا أشارك إطلاقا.

2- أشارك مشاركة ضعيفة (أو قليلة).

3- أشارك مشاركة متوسطة.

4- أشارك كثيرا.

مثال3: عند دراسة مستوى المعرفة الاقتصادية لدى شريحة من المجتمع، من الممكن تقسيم هذه الشريحة إلى ثلاث فئات:

1- مجموعة ضعيفة المعرفة.

2- مجموعة متوسطة المعرفة.

3- مجموعة عالية المعرفة.

في الأمثلة السابقة 1، 2، 3 تدل الأرقام " الأكثر " على قيمة أكبر، فهي تدل على زيادة حجم أو معدل درجة الطالب في المثال 1 وزيادة حجم مشاركته في النشاط الاقتصادي في المثال 2، وزيادة المعرفة الاقتصادية التي يتمتع بها الشخص في المثال 3، والعكس صحيح.

ولكن الملاحظة المهمة هنا أيضا – هو أنه لا يمكن إجراء أية عمليات حسابية على هذه الأرقام، فلا نستطيع أن نقول مثلا أن 2+1= 3 بمعنى أن تقدير (مقبول + ضعيف = جيد) في المثال (5) أو 3-1= 2 أي أن

مستوى المعرفة الاقتصادية في المثال (6) يمكن أن تكون (مجموعة عالية المعرفة – مجموعة ضعيفة المعرفة = مجموعة متوسطة المعرفة)... وهكذا.

كذلك نلاحظ أنه على الرغم من أن هذه البيانات ترتيبية بمعنى أن 2 أكبر من 1، 4 أكبر من 3. إلا أننا لا نعلم بأي قدر (أو بأي مدى) يكون مقدار هذا الكبر. فمثلا واضح أن الذي يشارك مشاركة متوسطة في النشاط الاقتصادي يكون أعلى في الترتيب من الذي يشارك مشاركة ضعيفة، لكن غير واضح (أو غير معروف) بأي قدر يكون أعلى في المشاركة. لذلك فعادة ما تندرج البيانات الترتيبية كذلك تحت ما يسمى " البيانات الوصفية Qualitative Data "

ثالثا : البيانات الكمية Quantitative Date

البيانات الكمية أو ما يطلق عليها أحيانا Interval Data هي البيانات التي يتم التعبير عنها بكميات أو أعداد، وبالتالي يمكن إجراء أي مقارنات بينها ويمكن ترتيبها ومعرفة الفروق بينها ويمكن إجراء أية عمليات حسابية عليها.

مثال1: لدراسة أجور العمال في مصنع معين يمكن كتابة الأجور بشكل تفصيلي بالشكل التالي:

الأعمار: 1000دج، 2000دج، 3000دج،....

أو على شكل فئات أجرية كما يلي:

من 1000 إلى أقل من 2000 دج.

من 2000 إلى أقل من 3000 دج.

من 3000 إلى أقل من 4000 دج،.... وهكذا.

فالأجور يمكن ترتيبها تصاعديا أو تنازليا، ويمكن إجراء المقارنات بينها. فالعامل الذي أجره 2000دج أكبر من العامل الذي أجره 3000 دج، وكذلك يمكن حساب الفرق بينهما بالتحديد (في هذه الحالة الفرق بين أجريهما 1000دج)، فالبيانات الكمية تمكن الباحث من إجراء كل العمليات الحسابية عليها من جمع وطرح وضرب وقسمة.

مثال 2: حجم العمالة لدولة ما عبارة عن بيانات كمية. فإذا أخذنا عينة عشوائية لمجموعة من دول العالم من حيث حجم عمالتها فإن النتيجة يمكن تلخيصها في الجدول.

حجم العمالة لعدد من دول العالم في (الكيلومتر المربع)

الدولة	حجم العمالة (الكيلومتر المربع)
زامبيا	12.6
الجزائر	8.22
ماليز يا	43.73
سنغافوره	4302.92
الدانمارك	118.87
فرنسا	98.66
الولايات المتحدة	24.56
كندا	2.42
الأرجنتين	10.15
البارغواي	8.03

فدولة سنغافورة – حسب ما هو موضح في الجدول أعلاه – تعتبر أكبر أو أعلى دولة من حيث حجم العمالة ، تليها جمهورية الدنمارك، ففرنسا، فالولايات المتحدة الأمريكية، فاتحاد ماليزيا، فزمبيا... وهكذا، كما أن الفرق بين حجم العمالة بين سنغافورة وكندا هو (4300.5) وهكذا يمكن إجراء كل العمليات الحسابية على البيانات الكمية.

ملاحظة هامة جدا: إن المتغير الكمي في العلوم الاقتصادية يجب أن يكون بالأساس متغيرا اقتصاديا، فلا يمكن اعتبار مثلا أن درجة الحرارة متغير اقتصادي رغم أنها عامل مهم في التأثير مثلا في مبيعات المبردات، لذلك فإنها تعتبر متغير نوعي ترتيبي.

مثلا: 1- درجة حرارة 0

2- درجة حرارة 10 درجات..... وهكذا.

وتنقسم البيانات الكمية إلى نوعين أساسيين:

(أ) البيانات الكمية المتقطعة Discrete

وهي البيانات التي تأخذ في الغالب (وليس دائما ً) عددا ً محدودا ً من القيم في مجال تغيرها أو بمعنى أدق تأخذ قيما ً منفصلة، أو متميزة عن بعضها البعض Distinct Values فمثلا ً: عدد عمال مؤسسة يمثل بيانات كمية متقطعة (أو منفصلة) لأن عدد العمال يعبر عنه بأعداد مثل 2، 3، 4، 5،..... 12 مثلا ً ، وعدد العمال مهما بلغ من

كبر سنجد أنهايأخذ عددا ً محدودا ً من القيم ونلاحظ أيضا ً في هذا المثال أن عدد العمال لا يأخذ كسورا ً بمعنى أنه لا يوجد مؤسسة بها 2.04 فرد أو 5.75 فرد.. وهكذا.

(ب) البيانات الكمية المتصلة (المستمرة) Continuous

وهي البيانات التي تأخذ عددا ً غير محدود من القيم داخل نطاق تغيرها، فمثلا ً بيانات أعمار العمال أو صافي دخل الفرد أو... كلها أمثلة للبيانات الكمية المتصلة.

ففي مثال دخل العامل: إذا كان أقل دخل للعامل هو 1000دج وأكبر دخل هو 6000دج فإنه بين هذين الحدين يوجد عدد كبير جدا ً (غير محدود من المداخيل) يتوقف على مدى دقة القياس (دينار ، سنتيم...) وبالطريقة نفسها تكون الأمثلة الأخرى.

بالإضافة إلى ما سبق، تبقى هنا ملاحظة مهمة وهي أن الموضوع أو الظاهرة محل الدراسة أو القياس تسمى " متغير " Variable حيث أنها تتغير من شخص إلى آخر أو من وحدة أو مفردة إلى أخرى، فيقال أن المتغير محل الدراسة هو الدخل أو العمر أو الحالة الاجتماعية، وبالتالي فقد يكون لدينا متغيرا ً وصفيا ً (اسميا ً أو ترتيبيا ً) متغيرا ً كميا ً (متقطعا ً أو متصلا ً).

فالمتغير هو صفة أو عدد لشيء من الممكن له أن يتغير نوعيا ً أو كميا ً سواء بالنسبة للشخص الواحد نفسه أو لمجموعة من الأشخاص، فتصنيف العمال حسب النوع إلى (ذكور – إناث) هو صفة من الممكن اعتمادها والتمييز بينها كمتغير، في حين أن الاختلاف في الدخل الشهري يمكن قياسه كمياً .

والمتغيرات في العلوم الاقتصادية يمكن تقسيمها إلى نوعين:

(أ) المتغيرات التابعة Dependent Variables وهي التي تشمل الصفات أو الحالات التي يريد الباحث تقديم تفسير لها، ومعرفة مدى تأثرها بالمتغيرات المستقلة، فهي المتغيرات أو الظواهر التي تتغير بناء على التغيرات التي تحدث في المتغيرات المستقلة.

(ب) المتغيرات المستقلة Independent Variables وهي المتغيرات التي تتغير أولا ً ، وتؤثر في متغيرات أخرى تابعة لها، فمثلا ً عند دراسة الدخل والإنفاق للأسرة نلاحظ أن الدخل هو المتغير المستقل لأنه هو الذي يتغير أولا ً وأن الإنفاق هو المتغير التابع لأنه يتغير بناء على التغير في الدخل.

2- جمع البيانات وتبويبها:

تمثل البيانات – بكل أنواعها العمود الفقري للإحصاء أو للتحليل الإحصائي عموما ، ولتحليل الظواهر الاقتصادية بصفة خاصة، لذلك فإن جمع البيانات عن الظاهرة محل الدراسة تعتبر المرحلة الأولى من مراحل البحث العلمي بالأسلوب الإحصائي والتي تتوقف عليها كل المراحل التالية في البحث.

فلكي نقوم بدراسة أي ظاهرة أو مشكلة ما من الظواهر الاقتصادية يجب على الباحث أولا ، جمع كل البيانات والمعلومات اللازمة والممكن الحصول عليها والتي تساعد في تحديد المشكلة أو الظاهرة تحديدا ، دقيق ، وتساعد بالتالي في اتخاذ القرارات المناسبة، فالبيانات لا تجمع لذاتها بل تجمع بغرض دراستها وتحليلها واستخلاص النتائج منها لذلك فإنه قد يكون من المناسب أن نتناول موضوع جمع البيانات قبل الدخول بعمق في تحليلها.

2-1- مفاهيم أساسية:

أ- المجتمع الإحصائي Statistical Population :

يعرف المجتمع الإحصائي بأنه كل المفردات التي تجمعها صفات وخصائص عامة مشتركة، أي يجمعها إطار عام واحد "كما أن هناك حدود واضحة وقاطعة من حيث الزمان والمكان لهذه المفردات بحيث يكون معروفا ، تماما ، المفردات التي تنتمي لهذا المجتمع دون أدنى شك" وهذا المفردات قد تكون بشرية وقد تكون غير بشرية كما قد تكون قياسات (أو مشاهدات) لظاهرة ما. فالمجتمع الإحصائي هو تجمع من المفردات تشترك في صفات وخصائص عامة يمكن قياسها.

فالعمال في مؤسسة ما، يشكلون مجتمع في هذه المؤسسة، مثلا ، كما لو يمثل طلاب قسم العلوم الاقتصادية مجتمع قسم الاقتصاد في كليتهم، كذلك يمكن أن يشكل إنتاج أحد المصانع الحربية لسلاح في فترة زمنية معينة - وهي في هذه الحالة مفردات غير بشرية - مجتمع إنتاج هذا المصنع في هذه الفترة.

من هنا نستطيع القول بأن المجتمع الإ حصائي يتميز بمجموعة من الصفات أهمها ما يلي:

1- الشمول: أي تشمل جميع المفردات. مثل جميع طلاب جامعة محمد البشير الابراهيمي.

2- الخصائص والصفات العامة الواحدة (أو الإطار العام الواحد): ففي حالة المثال السابق نستطيع القول بأن مجتمع الطلبة تجمعهم خصائص أو صفات عامة واحدة هي كونهم يدرسون بجامعة محمد البشير الابراهيمي.

3- المفردات: قد تكون بشرية (أي أفراد) كما في حالة الطلاب أو غير بشرية (حيوانية أو نباتية أو قياسات لأي ظاهرة).

4- المفردات معرفة تماما ، ومحددة بحدود زمانية ومكانية واضحة لا يشوبها أي شك.

5-قد يكون المجتمع محددا Finite (أي يمكن حصره ومعرفة عدد مفرداته)، وقد يكون غير محدود Infinite (لا يمكن حصره أو لا يمكن معرفة عدد مفرداته).

ونلاحظ هنا أن علم الإحصاء يعتمد على دراسة المجتمعات الإحصائية ككل ولا يهتم كثيرا بدراسة مفردة واحدة بعينها.

ب- معالم المجتمع Parameters of the Population :

معالم المجتمع هي المقاييس الإحصائية التي تصف المجتمع وتحدده والتي يمكن حسابها بالاعتماد على القياس الكمي لمفردات المجتمع، فكما نعلم فإن التحليل الإحصائي أو الدراسات الإحصائية تعتمد في معظمها (إن لم يكن في كلها) على القياس الكمي للظواهر محل الدراسة، فإذا كان موضوع الدراسة مثلا - هو دخول الأسر في إحدى المدن فيكون لدى الباحث دخول جميع الأسر في هذه المدينة، وبالتالي يكون في استطاعة الباحث حساب عدد من المقاييس الإحصائية التي تصف هذا المجتمع (مجتمع الأسر في هذه المدينة) وتعبر بدقة عن جميع مفرداته مثل: - متوسط دخل الأسرة في مدينة ما، والذي يمثل المستوى العام للدخل فيها (وهو الدخل الذي يتركز حوله دخول الأسر بهذه المدينة).

-التباين (أو التشتت) في توزيع هذه الدخول والذي يقيس مدى تشتت دخول الأسر حول متوسط الدخل.

- نسبة الأسر التي يقل دخلها مثلا عن حد معين.

فهذه المقاييس الإحصائية (المتوسط والتباين والنسبة) هي بعض معالم المجتمع والتي يمكن حسابها من القياس الكمي لكل مفردات المجتمع.

ومن ذلك يتضح أن معالم المجتمع تقاس من خلال جميع مفرداته، ونلاحظ أن ذلك قد لا يكون ممكنا في كثير من المجتمعات أو في كثير من الظروف مما يجعل الباحث يفكر في أسلوب أخر بديل لتقدير هذه المعالم وذلك من خلال العينات.

ج- الحصر الشامل Complete Enumeration :

المقصود بأسلوب الحصر الشامل هو شمول الدراسة (أو البحث) لجميع مفردات المجتمع الإحصائي محل الدراسة دون استثناء، وتتركز أهم مزايا الحصر الشامل فيما يلي:

أولا : إعطاء صورة شاملة للمجتمع، حيث يوفر بيانات ومعلومات عن كل مفردات المجتمع.

ثانيا : نتائج الدراسة بالحصر الشامل نهائية وليست في حاجة إلى تعديل أو تعميم.

ثالثا: أنه الأسلوب الوحيد المناسب في بعض الحالات مثل:

1- التعدادات (مثل التعداد العام للسكان، وتعداد القوى العاملة، وتعداد المنشآت الصناعية...) فكلمة تعداد مرادفة للحصر الشامل في الغالب.

2- الحالات التي قد يترتب عليها أضرار كبيرة لو تركت بعض المفردات دون فحص أو دراسة. مثل تطعيم الأطفال ضد أمراض معينة، إذ لابد تطعيم جميع الأطفال الذين يبلغون سناً معينة ضد بعض الأمراض (شلل الأطفال والحصبة)، وكمثال آخر فحص أسطوانات الغاز قبل توزيعها إذ يجب أن تفحص جميع الاسطوانات قبل التوزيع وإلا - لا سمح الله - قد تحدث بعض الأضرار لو تركت بعضها بغير فحص وكانت غير سليمة.

وكما للحصر الشامل مزايا فإن له عيوب تتركز في الوقت والمجهود والتكاليف وخصوصاً كلما كان حجم المجتمع كبيراً فإنه في هذه الحالات يتطلب وقتاً أطول ومجهوداً أكبر وتكاليف أكثر، كما أن هناك بعض الحالات التي قد يصعب فيها (بل قد يستحيل أحياناً) الدراسة عن طريق الحصر الشامل مثل:

(أ) الحالات التي تتلف أو تهلك بسبب الدراسة (مثل فحص الدم، أو فحص إنتاج مزرعة من البيض أو الدجاج،.. يجب في هذه الحالات أخذ عينة وتكون صغيرة بقدر الإمكان وإلا تتلف أو تهلك كل مفردات المجتمع أو يموت الشخص إذا أخذ كل دمه).

(ب) المجتمعات غير المحدودة، في هذه الحالات تستحيل الدراسة عن طريق الحصر الشامل، إذ كيف يكون المجتمع غير محدود وتجري عليه الدراسة بالحصر الشامل ؟ وكمثال على ذلك، إجراء دراسة عن سمك الهامور في البحر الأحمر فالأسلوب الوحيد هو عن طريق عينة وليس الحصر الشامل.

د-العينة Sample :

العينة هي جزء من المجتمع- أي هي جزء من الكل على أن يكون هذا الجزء ممثلاً للكل، بمعنى أنه يجب أن تكون العينة ممثلة للمجتمع المسحوبة منه تمثيلاً صادقاً، أو بمعنى آخر يجب أن تكون خصائص المجتمع بما فيها من فروق واختلافات ظاهرة في العينة بقدر الإمكان، فالعينة يتم اختيارها - عادة - بهدف تعميم النتائج التي يحصل عليها الباحث منها على المجتمع بأكمله بعد ذلك ولذا يجب أن تكون العينة ممثلة للمجتمع المسحوبة منه تمثيلاً صادقاً حتى يتسنى للباحث استخدام بيانات ونتائج العينة في تقدير معالم المجتمع بشكل جيد، فدراسة السلوك الاستهلاكي للأفراد في مجتمع معين يعني أن تكون العينة " أفراداً من المجتمع محل الدراسة وممثلة لهذا المجتمع، وبنفس الطريقة فإن دراسة السلوك الاستهلاكي للدول يعني أن العينة يجب أن تكون "دولاً " وأن تكون ممثلة لهذه الدول.

وتسمى عملية اختيار العينة "بالمعاينة Sampling وسوف نرى أنه يمكن اختيار أكثر من عينة من المجتمع وبطرق مختلفة لكل منها خصائصها، كما أن هناك حالات يفضل فيها استخدام نوع معين من العينات دون الأخر، وسوف نتناول بالتفصيل أهم أنواع العينات ومتى وكيف يتم اختيار كل منها وحدود ذلك:

- **العينة الكبيرة Large Sample** تكون العينة كبيرة إذا كان حجمها أو عدد مفرداتها يساوي (30) مفردة أو أكثر. مع ملاحظة أنه كلما كانت العينة كبيرة بدرجة كافية كلما كان ذلك أفضل، حيث يتمكن الباحث – في الغالب – من استخدام الكثير من أساليب التحليل الكمي والإحصائي، أي كلما كان حجم العينة أكبر من 30 كلما كان ذلك أفضل.

- **العينة الصغيرة Small Sample** تكون العينة صغيرة إذا كان حجمها أو عدد مفرداتها أقل من (30) مفردة، وهناك أساليب تحليل إحصائي خاصة إذا كانت العينة صغيرة.

هـ- إحصاءات العينة Sample Statistics :

إحصاءات العينة هي المقاييس الإحصائية الخاصة بالعينة، أي التي تصف العينة بالاعتماد على القياس الكمي لمفردات العينة مثل:

.الوسط الحسابي للعينة Mean

.تباين العينة (الانحراف المعياري لها) Standard Deviation

.النسبة في العينة (Ratio أو Proporation)

وإحصاءات العينة هي التي تستخدم عادة في الاستدلال على معالم المجتمع، أي هي التي تستخدم في تقدير معالم المجتمع أو في اختبار الفروض حولها.

ونتيجة لما يقدمه أسلوب العينات من مزايا فإنه أصبح واسع الانتشار، وأصبحت له استخدامات كثيرة في مختلف المجالات، وفيما يلي نتناول مزايا العينات وأيضا ً العيوب (أو المحاذير) التي يجب الانتباه إليها عند الدراسة بهذا الأسلوب:

المزايا:

1- توفير الوقت والجهد والتكاليف. فبدلا ً من قيام البحث بدراسة جميع مفردات ظاهرة اقتصادية ما كمجتمع البطالين في دولة معينة مثلا ً يكتفى عادة بدراسة عينة منها.

2- الحصول على معلومات أكثر تفصيلا ً من تلك التي نحصل عليها من مفردات المجتمع لا سيما إذا كان م المجتمع كبيرا ً ، فاختيار عينة يساعد الباحث على تركيز الجهد والدقة في جمع البيانات وبشكل أكثر تفصيلا ً وبالتالي قد تكون نتائج العينة أكثر دقة من النتائج التي يحصل عليها الباحث باستخدام الحصر الشامل، كما أنه

يمكن تقدير الأخطاء التي تتعرض لها العينات باستخدام النظريات الإحصائية الأمر الذي يمكن الباحث من تقدير دقة النتائج التي يحصل عليها باستخدام العينات.

3- هي الأسلوب الوحيد في حالة المجتمعات غير المحدودة حيث يستحيل حصر ودراسة جميع المفردات كدراسة الأسماك في البحر الأحمر مثلا .

4- هي الأسلوب الوحيد المناسب في الحالات التي يترتب على الدراسة أهلاك وإتلاف المفردات، مثل فحص دم مريض، اختبارات الجودة لكثير من المنتجات التي تؤدي إلى عدم صلاحيتها مرة أخرى مثل المعلبات، والمصابيح الكهربائية، إنتاج مزرعة من البيض.. الخ.

5- كذلك فإن العينات تناسب الظواهر التي تتغير بسرعة أي ذات الطبيعية المتغيرة باستمرار، في هذه الحالات يكون الأفضل استخدام العينات بدلا من الحصر الشامل الذي يتطلب وقتا أطول تكون الظاهرة أثناء هذا الوقت تغيرت أكثر من مرة وبالتالي تكون نتائج الحصر الشامل لا تعبر عن الظاهرة عند الانتهاء من الدراسة.

عيوب العينات:

1- تتوقف نتائج الدراسة بالعينات على مدى تمثيل العينة للمجتمع، فإذا كانت العينة غير ممثلة للمجتمع تمثيلا صادقا كانت النتائج غير دقيقة لذلك ينصح دائما باتباع الأسلوب العلمي السليم عند اختيار العينة.

2- نتائج دراسة العينة قد تكون غير نهائية وقد تكون في حاجة إلى تعميم، وسوف تختلف النتائج باختلاف نوع العينة حتى باتباع الأسلوب العلمي في الاختيار.

3- تتعرض دراسة العينات عموما لنوع من الأخطاء غير الأخطاء العادية تسمى " أخطاء المعاينة " Sampling Errors وهي الأخطاء الناجمة عن الدراسة بالعينة وإن كانت هذه الأخطاء تتميز بأنه يمكن تقديرها والتحكم فيها باتباع الأساليب الإحصائية السليمة عند اختيار العينة.

4- لا تصلح الدراسة للعينات في بعض الحالات التي لو تركت فيها بعض المفردات دون فحص أو دراسة يترتب عليها إلحاق الضرر بالمجتمع أو ببعض مفرداته (التطعيم، اسطوانات الغاز...).

5- في بعض الدراسات التي تحتاج دقة أكثر يتطلب الأمر زيادة حجم العينة الأمر الذي قد لا يكون متاحا عمليا ، أو لا يوفر كثيرا في الوقت والجهد والتكاليف وبالتالي لا تتم الاستفادة من مزايا العينات أو لا يكون هناك فرق كبير بين الدراسة بالعينة والدراسة بالحصر الشامل.

و- أقسام العينات:

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

1- العينات العشوائية:

هي تلك العينات التي يتم اختيار مفرداتها حسب خطة إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها، حيث يتم الإ ختيار 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،و من أهم أنواع العينات العشوائية نذكر ما يلي:

(أ) العينة العشوائية البسيطة: Simple random sample

ويلجأ إليها الباحث في حالة ما إذا كان مجتمع الدراسة ليس كبيرا ، ويحمل قدرا ، من التجانس بين المفردات للصفة أو الصفات موضع الدراسة، والعينة العشوائية البسيطة تستغل فرص متكافئة لمفردات المجتمع للدخول في العينة ولكن المفردات التي تدخل في العينة تكون عن طريق الصدفة البحتة والإ ختيار العشوائي يدويا ، عن طريق بطاقات متماثلة في الحجم واللون أو عن طريق جداول الأعداد العشوائية أو عن طريق الحاسب الآلي، ولكي يتحقق ذلك فإن الأمر يتطلب تحديد مفردات المجتمع تحديدا ، كاملا ، ويكون هذا التحديد على شكل قائمة (أو خريطة) تضم كل مفردات المجتمع وهذه القائمة تسمى الإطار (Frame)ولا يجوز الإ ختيار العشوائي إلا من المفردات التي يضمها الإطار.

(ب) العينة المنتظمة: Systematic sample

اختيار هذه العينة يتطلب وجود إطار للمجتمع كما في حالة العينة العشوائية البسيطة بحيث يعطى لكل مفرده من مفردات المجتمع رقما ، متسلسلا ، داخل الإطار، ثم نختار مفردات العينة من الإطار بحيث يكون الرقم المتسلسل لكل مفردة يبعد بعدا ، ثابتا ، منتظما ، عن رقم المفردة السابقة لها وكذلك رقم المفردة اللاحقة لها فمثلا ، إذا كان لدينا مجتمعا ، حجمه 2000 مفرده ونريد اختيار عينه منتظمة حجمها 100 مفرده فإننا نقسم الإطار إلى فترات منتظمة طول كل فترة $\frac{2000}{100}$ =20 مفردة ومن داخل مفردات الفترة الأولى (1–20) يختار مفرده واحدة عشوائيا ، ولتكن رقم 14 مثلا ، وبناء على رقم تلك المفردة يتحدد باقي مفردات العينة المنتظمة فتكون هي المفردات ذات الأرقام 34 ، 54 ، .... ، 1974، 1994.

والعينة المنتظمة كثيرة الإ ستعمال في التطبيقات العملية لقلة تكاليفها وقلة الأخطاء التي ترتكب في اختيار مفردات العينة فضلا ، عن سهولة إجرائها، ولكن أهم عيوب المعاينة المنتظمة هو عدم صلاحيتها إذا ما وجدت

علاقة قوية مع ترتيب العناصر في القائمة وكان طول الفترة بين عناصر العينة مساويا لطول الدورة أو إحدى مضاعفاتها.

(ج) العينة العشوائية الطبقية: Stratified random sample

ويلجأ إليها الباحث في حالة ما إذا كان مجتمع الدراسة واضحا به فئات (طبقات) بحيث أن التجانس أو التقارب داخل كل طبقة من طبقات مجتمع الدراسة أكبر من التجانس داخل المجتمع ككل (أي أن التشتت داخل المجتمع ككل أكبر من التشتت داخل كل فئة من فئاته على حدى)، في هذه الحالة يجب على الباحث مراعاة أن الطبقة داخل العينة بنفس نسبة وجودها داخل المجتمع (وأحيانا يوضع في الإعتبار عناصر أخرى مثل التشتت داخل الطبقة أو عنصر التكلفة لجمع البيانات عن الطبقة)، بعد أن يتم تحديد عدد المفردات التي يجب سحبها من كل طبقة للدخول في العينة فإن هذه المفردات يتم سحبها عشوائيا من داخل الطبقة ومجموع هذه المفردات تكون العينة الطبقية العشوائية.

(د) العينة متعددة المراحل أو العنقودية: Clustered sample

يلجأ إليها الباحث عندما يكون مجتمع الدراسة كبير جدا ومتناثرا على مساحات شاسعة تكلف الكثير من الوقت والجهد في التنقل بينها عند جمع البيانات، أيضا في حالة عدم وجود إطار يضم جميع مفردات المجتمع فيستحيل الإختيار العشوائي مباشر ة من المجتمع لهذا يلجأ الباحث إلى أخذ العينة على مراحل متعددة متتالية، في المرحلة الأولى يتم تقسيم المجتمع إلى عدد محدد من وحدات المعاينة الكبيرة الحجم ومنها يختار بعضها عشوائيا ثم يتلو ذلك كمرحلة ثانية تقسيم الوحدات المختارة عشوائيا من المرحلة الأولى إلى وحدات أقل منها في الحجم ثم يختار بعضها عشوائيا .. وهكذا تابع مراحل التقسيم والإختيار العشوائي، وعدد هذه المراحل ليس ثابت بل يتوقف على طبيعة مجتمع الدراسة وإمكانيات الباحث .. في المرحلة الأخيرة يصل الباحث إلى وحدات المعاينة التي سيجمع عنها بيانات البحث ويطلق عليها وحدات المعاينة الأولية.

2- العينات غير العشوائية:

هي تلك العينات التي لا تكفل لجميع مفردات المجتمع احتمال ثابت ومحدد للاختيار، وغالبا يتدخل الباحث في عملية الإختيار بصورة أو بأخرى ... ومن أهم أنواع العينات غير العشوائية:

(أ) العينة العمدية أو المقصودة: Purposive sample

يلجأ الباحث إلى هذه الطريقة فيما إذا كان مجتمع الدراسة كبير جدا وكانت إمكانياته لا تسمح له إلا بدراسة عينة حجمها صغير جدا بالنسبة لمجتمع الدراسة، في هذه الحالة يتعمد الباحث اختيار مفردات معينة كعينة لمجتمع الدراسة ويرى بخبرته السابقة أن هذه العينة يمكن أن تعطي تمثيلا مقبولا لمجتمع الدراسة.

مثلا ، إذا أراد باحث دراسة خصائص اقتصادية أو اجتماعية معينه عن ريف دولة ما، وكانت إمكانياته المالية والإدارية لا تسمح له بعينة سوى سكان قرية واحدة، ففي هذه الحالة إذا ما تم اختيار القرية عشوائيا ، من بين آلاف القرى بتلك الدولة فإن الصدفة قد تأتي بقرية بعيدة في خصائصها (من حيث الظاهرة موضوع الدراسة) عن خصائص معظم قرى تلك الدولة ... كأن تأتي بالصدفة قرية ساحلية معظم سكانها من الصيادين أو قرية قريبة من مشروع صناعي ضخم يستوعب في قواه العاملة معظم سكانها.. هذه القرية أو تلك قد يأخذ النمط المعيشي لسكناها طابعا ، خاصا ، نابعا ، عن ظروفها الخاصة بعيدا ، عن النمط المعيشي المعتاد لبقية القرى، لذلك فأي منها لا يمكن أن يعطي تمثيلا ، مقبولا لريف تلك الدولة، لهذا فإن الباحث وعلى ضوء خبراته السابقة يتعمد اختيار قرية معينة يرى أنها – من وجهة نظره الشخصية- يمكن أن تمثل الريف، وهذه الطريقة غير علمية غالبا ، يتم اللجوء إليها في حالة البحوث التمهيدية.

(ب) العينة الحصصية: Quota sample

وهي نوع خاص من العينات غير العشوائية وتستخدم كثيرا ، في معاينة الرأي العام (على سبيل المثال عمليات الإستطلاع للرأي العام التي يقوم بها معهد جالوب قبل إجراء انتخابات الرئاسة في الولايات المتحدة الأمريكية)... في هذه الطريقة يقسم المجتمع موضوع الدراسة إلى طبقات بالنسبة إلى صفات أو خصائص معينة ويتم العمل على تمثيل كل طبقة منها في العينة بنسبة وجودها في المجتمع الأصلي (وعلى سبيل المثال في حالة دراسة الدخل لمنطقة ما ورؤى أن يكون حجم العينة المطلوبة 100 فرد مثلا ، عندما يريد الباحث أن يقوم جامعوا البيانات بالحصول على البيانات من 20 موظفا ، ، 45 من العمال الحرفيين، 35 من ذوي الأعمال الحرة .. وتترك الحرية لجامعي البيانات في اختيار الأفراد المطلوبة فيها حدود المواصفات الموضوعة لكل طبقة من الطبقات المذكورة.

ويتضح أنه رغما ، من أن هذه الطريقة في ظاهرها مماثلة للعينة الطبقية العشوائية.. إلا أنه في الحالة الأخيرة (العينة الطبقية العشوائية) يكون اختيار المفردات عشوائيا ، من داخل كل طبقة ولا يترك لجامع البيانات حرية اختيار المفردات من كل طبقة والذي قد يترتب عليه تميزا ، كبيرا ، .

عموما ، .. يلجأ الباحث إلى العينة الحصصية إذا كان من المرغوب فيها إظهار النتائج في وقت قصير مع التغاضي عما قد يترتب عليه من تميزا كبير أو رؤى عن توافر درجة دقة عالية بتلك النتائج.

3- استبانة البحث الميداني :

إذا قرر الباحث استخدام المصادر الميدانية للحصول على البيانات، أي إذا قرر أن يجمع بياناته بنفسه سواء عن طريق الحصر الشامل أو العينات وكان المجتمع بشريا ، (أي وحدات المعاينة عبارة عن أفراد أو أسر أو أي وحدات تستطيع الإجابة على الأسئلة) فإن على الباحث أن يصمم استبيان أو استقصاء أو ما يسمى " استبانة

البحث " يضع فيها كل الأسئلة التي لو أجاب عليها أفراد المجتمع أو العينة لحصل على البيانات اللازمة للبحث، وبديهي فإن استمارة البحث سوف تختلف من بحث لآخر حسب موضوع البحث وأهدافه وحجمه وطبيعة البيانات المطلوبة، ولكن هناك قواعد (أو ملاحظات) عامة يجب مراعاتها عند تصميم استبانة البحث أياً كان موضوع البحث أو الدراسة.

ويمكن تصنيف هذه القواعد أو الملاحظات إلى ثلاث أنواع هي: الملاحظات الأولية (أو التعريفية) والملاحظات الشكلية، ثم الملاحظات الموضوعية، وفيما يلي نتناول هذه الملاحظات بشيء من التفصيل:

أولاً : الملاحظات الأولية (أو التعريفية):

ونعني بها تلك الملاحظات أو المعلومات التي يوردها (أو يذكرها) الباحث في الصفحة الأولى من الاستبانة وتساعد أفراد المجتمع أو العينة (المبحوثين) على التعرف على البحث وأهدافه ومن الممكن إيجازها في النقاط التالية:

1- ما هو البحث؟ ما موضوعه؟ وما أهدافه وما جدواه ؟

2- من هو الباحث ؟ أو ما هي الجهة القائمة بالبحث أو المشرفة عليه (أو التي تموله)؟

3- التأكيد على سرية البيانات، وأن هذه البيانات لن تستخدم لأغراض أخرى غير هذا البحث، ولن يطلع عليها غير القائمين بالبحث، وفي هذا الصدد تميل معظم البحوث الآن (وإن لم يكن كلها) على ألا يذكر المبحوث اسمه أو ما يشير إلى شخصه عند الإجابة عن الأسئلة بالاستبانة حتى يطمئن تماماًعند الإجابة على أسئلة الاستبانة.

4- توجيه الشكر والتقدير للمبحوثين على اهتمامهم بالإجابة وإضاعة بعض الوقت في قراءة الاستبانة والإجابة عليها.

ثانياً : الملاحظات الشكلية :

ونعني هنا بالملاحظات الشكلية كل ما يتعلق بشكل الاستبانة (أي بالنواحي الشكلية) من حيث الورق وحجم الاستبانة والطباعة ولون الحبر والتصميم (أو الإخراج) الفني للاستمارة من حيث التنظيم والشكل العام لها، فيجب ألا يغفل الباحث كل هذه النواحي الشكلية لأن لها تأثير نفسي (سواء بالسلب أو الإيجاب) على المبحوثين سواء من ناحية الاهتمام أو بالإدلاء بالبيانات الصحيحة وبالسرعة المطلوبة.

ثالثاً : الملاحظات الموضوعية :

ونقصد بالملاحظات الموضوعية كل الملاحظات المتعلقة بموضوع البحث أي بالأسئلة التي بالاستبانة، ويمكن تلخيص أهم هذه الملاحظات فيما يلي:

1- أن تكون الأسئلة واضحة تماما ً لا لبس فيها ولا غموض وأن تكون بلغة المبحوثين (أو لهجتهم). فغموض صيغة السؤال قد يكون مرده عموميته كأن تسأل" هل تشارك في اجتماعات الإدارة العليا في إدارتك ؟ " كما قد يكون مرده إلى عدم تحديد الزمان بالقطع مثل السؤال " هل شاركت في اجتماعات إدارتك العليا؟"، وقد يكون مرده إلى تجاهل تحديد المكان في السؤال " هل هناك فساد مالي ؟".

2- أن تقسم الأسئلة إلى مجموعات متجانسة، وأن تكون مرتبة ترتيبا ً منطقيا ً فالأسئلة التي تتعلق بمتغير الخلفية الثقافية يجب أن تلي بعضها البعض ثم الأسئلة التي تتعلق بالخلفية الاجتماعية.. وهكذا.

3- أن تكون الأسئلة مغلقة بقدر الإمكان (أي محددة تماما ً كأن تكون الإجابة عليها بنعم أو لا)، وفي الأسئلة الأخرى تحدد الإجابات المحتملة ويطلب من المبحوث وضع علامة أمام الإجابة المناسبة.

4- أن لا تكون الأسئلة أكثر من اللازم حتى لا يمل المبحوثون، ولا تكون أقل من اللازم حتى يحصل الباحث على البيانات المطلوبة للبحث.

5- الإبتعاد ما أمكن عن الأسئلة الإيحائية (أي التي توحي للمبحوثين بالإجابة) فصيغة سؤال مثل "أليس من الأفضل الاستعانة بالعمالة العربية بدلا ً من الأجنبية؟ يبدو وكأنه إيحاء بالإجابة ومن المؤكد أن ردة الفعل ستختلف فيما لو كانت صيغته " هل تفضل العمالة العربية ؟، أو ما نوع العمالة التي تفضلها؟.

6- الابتعاد عن الأسئلة التي تتطلب عمليات حسابية معقدة (ومن نوعية هذه الأسئلة هو أن تسأل المبحوث عن معدل إنفاقه الشهري على الأحذية).

7- الابتعاد عن الأسئلة المحرجة أو الحساسة أو التي تتنافى مع عادات وتقاليد وقيم مجتمع الدراسة كأن تسأل المبحوثين في مجتمع كمجتمعنا " هل تصلي ؟ ".

8- الابتعاد عن الأسئلة التي تعتمد الإجابة عليها على ذاكرة المبحوثين، كأن تسأل "اذكر أخر تاريخ لاجتماعات إدارتك العام الماضي؟".

9- أن يوضح في السؤال كل الوحدات أو المقاييس المطلوبة، فمثلا ً: لا يجوز وضع سؤال في الاستبانة كما يلي: (ما دخلك ؟) فهل المقصود هنا بالدخل الراتب فقط (إذا كان موظفا ً) أم الدخل بكل مصادره ؟ وهل المقصود الدخل الشهري أم السنوي أم اليومي ؟ وما العملة التي يحسب بها الدخل هل هي الدينار الجزائري أم الأورو" وهكذا في الأسئلة المشابهة يجب أن توضح كل الوحدات أو المعايير التي يقصدها الباحث.

10- ينصح عادة بوضع بعض الأسئلة الضابطة في الاستبانة بهدف التأكد من صحة البيانات التي يدلي بها المبحوث، أي يجب وضع بعض الأسئلة المهمة (والتي تشكل أساس البحث) بصيغ مختلفة وفي أماكن متفرقة من الاستبانة حتى يتأكد الباحث من مدى صحة الإجابات أو ما إذا كان هناك تناقض فيها أم لا.

وتبقى هنا ملاحظة مهمة وهي أنه بعد أن يأخذ الباحث كل الملاحظات السابقة في الاعتبار عند تصميم استمارة البحث فإنه يتم أولا تحكيمها من طرف خبراء في المجال ثم تجر بتها على عينة صغيرة من المبحوثين حتى يتعرف الباحث على نقاط الضعف فيها أو الأسئلة التي تحتاج إلى تعديل حتى تكون أكثر وضوحا ، ثم بعد ذلك تستخدم الاستبانة للمجتمع (أو العينة) التي تجري عليها الدراسة.

- العرض الجدولي للبيانات أو (تبويب البيانات) :

بعد انتهاء عملية جمع البيانات الميدانية بواسطة استمارة البحث الميداني تبدأ المرحلة الثانية وهي تبويب هذه البيانات أو عرضها جدوليا . فالبيانات المجموعة بواسطة الاستمارات تعتبر بيانات "خام" يصعب دراستها وتحليلها واستخلاص النتائج منها وهي في الاستمارات، لذلك يقوم الباحث بتفريغ هذه البيانات في جداول يتناول كل جدول منها سؤال (أو متغير) أو أكثر بهدف تلخيص هذه البيانات وإعطاء صورة شاملة لكل مفردات البحث، حتى يسهل دراسة هذه البيانات وتحليلها واستخلاص النتائج منها، لأن اهتمام الباحث الإحصائي يكون – عادة- بالمفردات ككل وليس بكل مفردة على حدة: ونتناول فيما يلي تبويب البيانات أو عرضها جدوليا بشيء من الإيجاز.

أولا : البيانات (أو المتغيرات) الوصفية:

ذكرنا فيما سبق أن البيانات الوصفية هي التي يتم التعبير عنها " بصفات " أو "حالات "، أو لا يتم التعبير عنها كميا ، وكأمثلة على ذلك: - الحالة الاجتماعية للعامل - جنسية الطالب - الحالة التعليمية للموظف، ففي هذه الأمثلة يتم التعبير عن المتغيرات بصفات أو حالات، فإذا كان بالاستبانة سؤالا عن الحالة الاجتماعية للعامل فإنه يمكن تفريغ إجابات هذا السؤال في جدول واحد صغير يعطي الصورة الشاملة للحالة الاجتماعية لكل العاملين (محل الدراسة)، فإذا كان عددهم مثلا - 120 عاملا فيمكن أن يكون الجدول كما يلي (مثال افتراضي):

جدول الحالة الاجتماعية للعامل

الحالة الاجتماعية للعامل	أعداد العمال	النسب
أعزب	45	0.375
متزوج	60	0.500
أرمل	9	0.075
مطلق	6	0.050
المجموع	120	1

ملاحظة: حصلنا على النسب في العمود الثالث بقسمة كل من التكرارات على إجمالي عدد الناخبين، فمثلا نسبة العزاب = $^{45}/_{120}$ = 0.375 وهكذا.. كذلك يمكن ضرب هذه النسب في (100) فنحصل على النسب المئوية للحالات الاجتماعية للناخبين. وفي هذه الحالة تكون النسب المئوية على الأخرى ويكون المجموع 100% أو واحد صحيح.

وقد يرغب الباحث في دراسة العلاقة بين متغيرين وصفيين فيتم تبويب بيانات السؤالين في جدول واحد مزدوج، فمثلا إذا كان الباحث بصدد دراسة العلاقة بين الحالات الاجتماعية للعمال والمؤسسة التي يشتغلون بها فيمكن تفريغ بيانات (إجابات) السؤالين في جدول مزدوج (يسمى أحيانا جدول التوافق) كما يلي:

المؤسسة / الحالة الاجتماعية	سوناطراك	كوندور	موظف ببلدية	المجموع
أعزب	10	15	20	45
متزوج	30	18	12	60
أرمل	4	3	2	9
مطلق	2	2	2	6
المجموع	46	38	36	120

فمن الجدول السابق يتبين أن إجمالي العمال (120) منهم (45) عزبا، (60) متزوجا، (9) أرامل، (6) مطلقين. ومن هؤلاء العمال (46) يعملون بسوناطراك، (38) بكوندور، (36) ببلدية.....

أما إذا كنا بصدد دراسة متغير كمي آخر مثل الدخل فإن الوضع سوف يختلف، فإذا كان الدخل هو الدخل الشهري للعاملين بإحدى المؤسسات الجزائرية فإن الدخل يعبر عنه بآلاف الدينارات، كما يمكن أن يحتوي كسورا، لذلك فإنه من الصعوبة غير المفيد أيضا كتابة الدخول كلها بالتفصيل، فإذا كان أقل دخل مثلا هو 1000 دج وأكبر دخل هو أقل من 7000 دج فالمدى الذي تنتشر عليه البيانات كبير جدا ولذا فإننا نلجأ إلى ما يسمى "بالفئات"، حيث نكون فئات للدخل تكون متجانسة بقدر الإمكان ويمكن اعتبار كل منها فئة دخلية واحدة بشرط أن لا يكون هناك تداخل بين هذه الفئات، ولا يكون بينها فجوات.

والشكل المناسب للفئة الذي يحقق هذين الشرطين هو: من... إلى أقل من.... "أي" من رقم معين إلى أقل من رقم آخر.. وهكذا.

والجدول التالي يوضح توزيع العاملين بإحدى المؤسسات حسب فئات الدخل الشهري بالدينار الجزائري.

توزيع تكراري للعاملين حسب فئات الدخل:

فئات الدخل	أعداد العاملين (التكرارات)
1000 – 2000	8
2000 – 3000	15
3000 -4000	20
4000 – 5000	25
5000 – 6000	12
6000 – 7000	10
المجموع	90

ومن هذا الجدول نجد أن هناك 8 عاملين تتراوح دخولهم الشهرية من 1000 إلى أقل من 2000 دج وهذه هي الفئة الأولى، والفئة الثانية تقول: أن هناك 15 من العاملين تتراوح دخولهم الشهرية من 2000 إلى أقل من 3000 دج، وهكذا مع باقي الفئات، وعادة يسمى هذا الجدول " جدول توزيع تكراري Distribution Frequecy حيث يبين كيفية توزيع التكرارات (أعداد العاملين في هذا المثال) على الفئات المختلفة (فئات الدخل في هذا المثال).

- **ملاحظة هامة**: هناك عمليات ضرور ية يقوم بها الباحث قبل العرض الجدولي للبيانات تسمى "عمليات التجهيز"، وهذه العمليات ضرورية لأنها تسهل تصنيف البيانات وتبويبها وذلك استعدادا لتحليلها، وأهم هذه العمليات هي عملية " الترميز " Coding حيث يتم تحويل البيانات -وخاصة الوصفية- إلى إجابات رقمية لتسهيل عملية التصنيف والتبويب والتحليل، فمثلا في بيان " نوع " العامل: هناك حالتان إما ذكر أو أنثى، فيعطى الرقم (1) للذكر والرقم (2) للأنثى، وفي بيان حالته الاجتماعية نعطي الرقم (1) للأعزب، (2) للمتزوج، (3) للأرمل، (4) للمطلق، وهكذا.. لأنه من الأسهل التعامل مع الأرقام بدلا من الإجابات الوصفية وذلك بالنسبة للحاسبات الآلية.

المحور الثاني:

اختبارات الفروض الإحصائية

الأهداف التعليمية

- التذكير باختبار الفروض للوسط الحسابي؛
- التذكير باختبار الفروض للنسبة ؛
- التذكير باختبار الفروض للتباين ؛
- التذكير باختبار الفروض للفروق.

تمهيد:

تعتبر اختبارات الفروض الإحصائية واحدة من أهم التطبيقات التي قدمها علم الإحصاء كحل للمشاكل العلمية المختلفة بشتى فروع العلم، حيث أنه وبرغم أهمية موضوع تقدير المعالم إلا أنه غالبا ما يكون الإهتمام مركز ا ليس على مجرد تقدير المعالم ولكن علي عملية وضع قواعد تمكن من التوصل إلى قرار بقبول أو رفض خاصية، أو بالمعنى الإحصائي فرض عن معالم مجتمع واحد أو أكثر وهذا ما يسمى "اختبارات الفروض الإحصائية"، حيث ومن خلالها يمكن لأي باحث أن يتخذ قرار برفض أو قبول فرض معين أو مجموعة من الفروض المتعلقة بمشكلة معينة موجودة في الحياة العامة.

1- مدخل إلى اختبار الفروض الإحصائية:

من المعروف أن اتخاذ أي قرار لا يتم إلا من خلال اختبارات الفروض الإ حصائية التي تعتمد بدورها على الإ حتمالات وتوزيعات المعاينة وهذا يؤكد أهمية الدور الذي تلعبه نظريه الإ حتمالات في التنبؤ والتخطيط واتخاذ القرارات بالإضافة إلى أهميتها في تقدير معالم المجتمع المجهولة والتي تعتبر أحد اهتمامات الباحثين.

تبدأ مشكله التعرف على معالم المجتمع المجهولة بما يسمى بالإ ستدلال الإحصائي (Statistical Inferences) حيث ينقسم إلى فرعين:

الفرع الأول: يهتم بتقدير (Estimation) معالم المجتمع (تم التطرق إليه في المحور السابق).

- **والفرع الثاني:** يختص بإجراء اختبارات فروض(Testing Hypotheses) تدور حول معالم المجتمع المجهولة.

يرتكز اختبار الفروض الإحصائية أساسا على فكرة أن العينة المسحوبة من المجتمع لا تكون مثالية إلا في حالات نادرة جدا، وبالتالي فإن الفرق بين المعلمة المقدرة من هذه العينة المسحوبة وبين المعلمة المجهولة قد يكون فرقا معنويا Significant غير راجع للصدفة، وبالتالي وباستخدام اختبار الفروض نستطيع أن نحدد وبسهوله هل الفروق بين المعلومات المحسوبة من العينة وبين المعلومات المفروضة لمجتمع معين هي فروق ترجع إلى الصدفة أو أنها فروق حقيقية، وبأسلوب آخر هل هي فروق معنوية أو فروق غير معنوية؟ وبذلك سميت هذه الإ ختبارات كذلك باسم اختبارات المعنوية Test of Significant.

تعتبر الفروض الإحصائية بمثابة اقترا ح عن معالم المجتمع موضوع الدراسة والتي ما زالت غير معلو مة للباحث، وبالتالي فهي حلول ممكنة لمشكلة البحث (افتراضات) يمكن على أساسها تحديد المعلمة المجهولة المذكورة سلفا، ومنه فالفرض ما هو إلا تخمين أو استنتاج ذكي مبني على حيثيات معقولة أو منطقية ولكنه ليس بنيا 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.

فمثلا ً قد يفترض الباحث أن متوسط الدخل الشهري للفرد في دولة ما هو 200 دولار (بناء ً على ما يراه من مستوى المعيشة في هذا البلد وأوضاعه الإقتصادية)، ويحتاج إلى اختبار علمي (إحصائي) لمعرفة مدى صحة هذا الفرض، أي أن يصل الباحث إلى قرار إما بقبول الفرض أو عدم قبوله (أي رفضه) وذلك باحتمال معين، وقبل تناول كيفية إجراء الإختبارات الإحصائية نستعرض أولا ً بعض المفاهيم والتعريفات الأساسية اللازمة لهذا الموضوع حتى تكون الصورة أكثر وضوحا ً .

- 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 دولار 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 دولار شهريا ً .

وليس شرطا ً أن يصاغ الفرض العدمي بالرموز حيث يمكن أن يتم التعبير عنه بدون رموز، فقد يريد الباحث أن يختبر ما إذا كانت هناك علاقة بين الذوق والولاء للمنتج، أو بين المؤهل العلمي ودرجة الوعي الإقتصادي، فقد يصيغ الباحث الفرض العدمي بالشكل التالي (على سبيل المثال): الذوق والولاء للمنتج مستقلان (أي لا توجد علاقة بينهما أو أن العلاقة بينهما منعدمة).

- الفرض البديل: The Alternative Hypothesis

في اختبارات الفروض يتحتم وضع فرض آخر غير الفرض العدمي المراد اختباره يسمى الفرض البديل. وهذا الفرض "هو الذي سيـ ُقبل 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 وبالذات في قياس الظواهر الإقتصادية – كما سوف نرى – فهو الذي يحدد نوع الإختبار المستخدم، لذلك فهو يأخذ أحد الأشكال الثلاثة التالية:

أ- اختبار الطرفين: يأخذ شكل" لا يساوي" ويستخدم في حالة كان الإهتمام في البحث حول وجود فرق بين المعلمة المقدرة في العينة والمعلمة المقترحة (المفترضة) أو لا يوجد فرق، وفي هذه الحالة لا يهم إذا كانت المعلمة المفترضة أكبر أو أصغر من تلك المقدرة، فمثلا ً إذا كان الفرض العدمي هو أن متوسط الدخل الشهري لفئة معينة في المجتمع هو 200 دولار.

$$Ho : \mu = 200$$

فإن الفرض البديل في هذه الحالة يأخذ الشكل التالي:

$$H1 : \mu \neq 200$$

بمعنى أن متوسط دخل هذه الفئة من المجتمع " لا يساوي " 200 دولار شهريا .

ب- اختبار الطرف الأيمن: يأخذ الفرض البديل شكل " أكبر تماما من" ويستخدم في حالة كان تقدير الباحث أن القيمة المقدرة هي أكبر من القيمة المفترضة، فمثلا قد يكون الفرض البديل كما يلي:

$$H1 : \mu > 200$$

أي أن متوسط الدخل لهذه الفئة من المجتمع أكبر من 200 دولار شهريا .

ج- اختبار الطرف الأيسر: يأخذ الفرض البديل شكل " أقل تماما من " ويستخدم في حالة كان تقدير الباحث أن القيمة المقدرة هي أصغر من القيمة المفترضة، فمثلا قد يكون الفرض البديل كما يلي:

$$H1 : \mu < 200$$

أي أن متوسط الدخل لهذه الفئة من المجتمع أقل من 200 دولار شهريا .

والخلاصة أنه لابد للباحث من تحديد الفرض البديل الذي لا يخرج عن أحد الأشكال الثلاثة السابقة، وهذا التحديد مهم جدا قبل الدخول في تفاصيل الإختبار الإحصائي وذلك لأنه هو الذي يحدد نوع الإختبار المستخدم كما سوف نرى.

- الخطأ في اتخاذ القرار:

في 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 القرار 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[1]:

أ- الخطأ من النوع الأول: Type I error

[1] - بالتصرف عن: عبد الحميد عبد المجيد البلداوي، الأساليب الإحصائية التطبيقية، ط2، دار الشروق، 2008، ص282.

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".

ب- الخطأ من النوع الثاني: Type II error

في 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 ، لكن ذلك غير ممكن حيث لا يمكن تصغيرهما معا إلى أدنى حد ممكن، ويبدو أن الطريقة الوحيدة 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.

يعتبر مصطلح " مستوى المعنوية " واحد من أهم المصطلحات المستخدمة في دراسة نظرية اختبارات الفروض، والمقصود بمستوى المعنوية هو " احتمال حدوث الخطأ من النوع الأول " أو نسبة حدوثه أي "احتمال رفض الفرض العدمي بينما هو صحيح ".

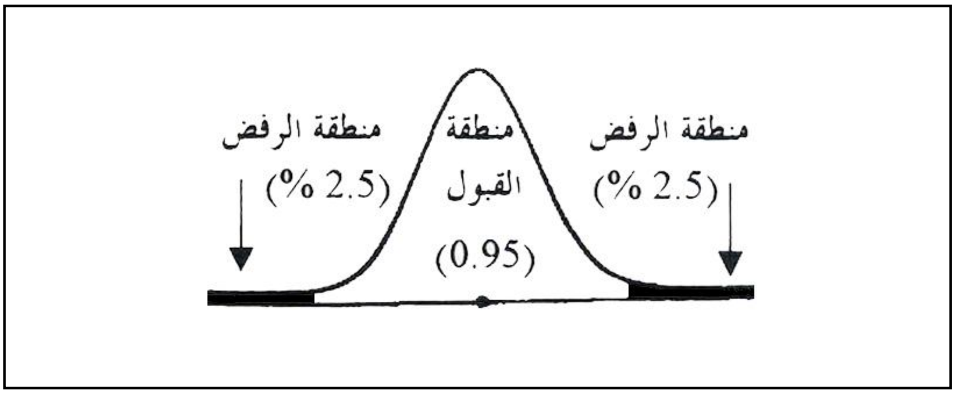
وعادة ما يرمز إلى مستوى المعنوية بالرمز اللاتيني ألفا α وأشهر القيم لمستوى المعنوية هي 5% ولكن ليس هناك ما يمنع من أن يأخذ قيما أخرى.

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، ولعل من أهم الملاحظات هنا هو استخدام تعبير "مستوى المعنوية" في حالات اختبارات الفروض، بينما يستخدم مصطلح "درجة أو مستوى الثقة" في حالات التقدير.

و الفكرة الأساسية في اختبار الفرض هي تقسيم المساحة تحت المنحنى إلى منطقتين: إحداهما 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:

أ- في حالة اختبار الطرفين: أي إذا كان الفرض البديل يأخذ شكل " لا يساوي" كأن يكون الفرض في هذه الحالة هو أن متوسط دخل الفرد لا يساوي 200 دولارا ، فإن منطقة الرفض تكون موزعة على طرفي المنحنى بالتساوي، والذي يأخذ الشكل التالي (بافتراض أن $\alpha = 5\%$):

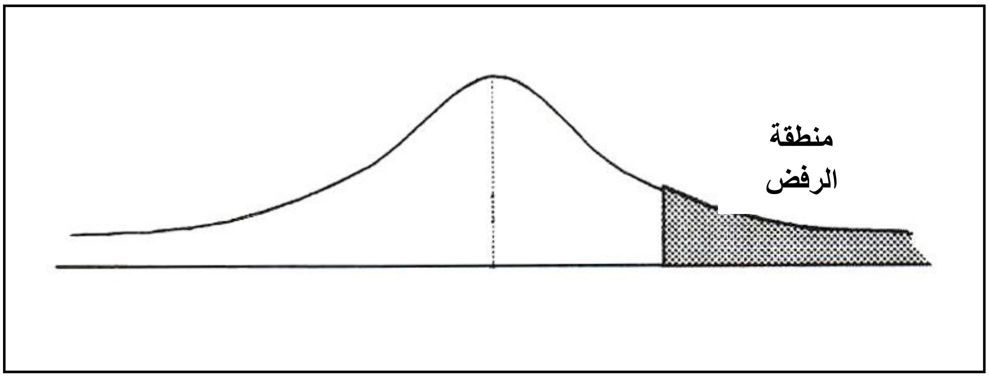
منطقتي الرفض ومنطقة القبول في حالة اختبار الطرفين



فالفرض العدمي هنا $H_0: \mu = 200$ يعني أن متوسط دخل الفرد يساوي 200 دولار شهريا، والفرض البديل في هذه الحالة هو $H_1: \mu \neq 200$ بمعنى أن متوسط دخل الفرد لا يساوي 200 دولار شهريا ، حيث تمثل المنطقة البيضاء غير المظللة منطقة القبول والتي تساوي 95% وبالتالي فمنطقة الرفض مقسمة بالتساوي على طرفي المنحنى حيث تكون في هذه الحالة 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%.

ب- في حالة اختبار الطرف الأيمن: إذا كان الفرض البديل يأخذ شكل "أكبر تماما من" فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيمن للمنحنى والذي يأخذ الشكل التالي:

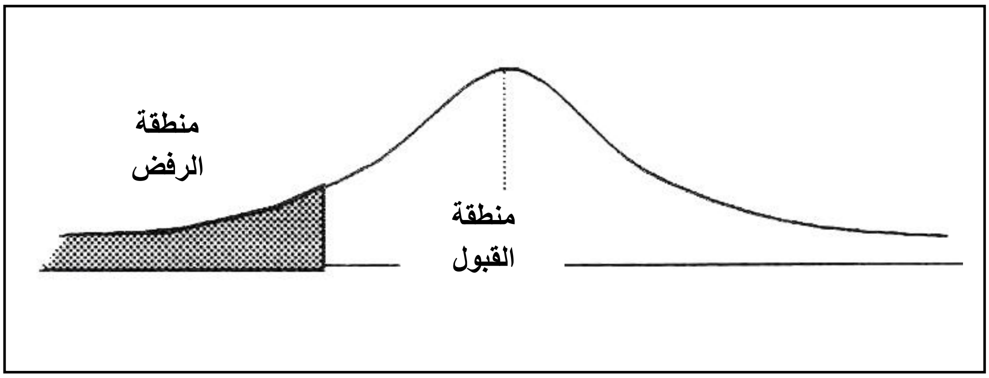
منطقة الرفض ومنطقة القبول في حالة اختبار الطرف الأيمن



فالفرض العدمي هنا نفس فرض المثال السابق، بينما الفرض البديل هو $H_1: \mu > 200$ بمعنى أن متوسط دخل الفرد أكبر من 200 دولار شهريا وبالتالي فإن مستوى المعنوية والذي يساوي مثلا 5% مركز في الطرف الأيمن من المنحنى.

ج- في حالة اختبار الطرف الأيسر: إذا كان الفرض البديل يأخذ شكل "أقل تماما من" فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيسر للمنحنى والذي يأخذ الشكل التالي:

منطقة الرفض ومنطقة القبول في حالة اختبار الطرف الأيسر



مع افتراض ثبات الفرض العدمي كما في المثال السابق، بينما الفرض البديل هو $H_1: \mu < 200$ بمعنى أن متوسط دخل الفرد أقل من 200 دولار شهريا ، وبالتالي فإن مستوى المعنوية والذي يساوي مثلا 5% مركز في الطرف الأيسر من المنحنى.

وسوف نتناول فيما يلي خطوات الإختبار الإحصائي بشيء من التفصيل.

- خطوات الاختبار الإحصائي:

يمكن تلخيص خطوات الإختبار الإحصائي في أربع خطوات كما يلي:

أ- صياغة الفرضيات: والتي تأخذ – عادة – الشكل التالي:

في حالة اختبار طرفين $\begin{cases} H_0: \vartheta = \vartheta_0 \\ H_1: \vartheta \neq \vartheta_0 \end{cases}$

في حالة اختبار الطرف الأيمن $\begin{cases} H_0: \vartheta = \vartheta_0 \\ H_1: \vartheta > \vartheta_0 \end{cases}$

في حالة اختبار الطرف الأيسر $\begin{cases} H_0: \vartheta = \vartheta_0 \\ H_1: \vartheta < \vartheta_0 \end{cases}$

والذي يحدد شكل الفرض البديل هو مدى اقتناع الباحث بذلك أو مدى توفر المعلومات الأولية، فمثلاً إذا كانت وجهة نظر الباحث أن متوسط دخل الفرد لا يمكن أن يقل عن 200 دولار فإنه يختار الفرض البديل "أكبر تماما من" والعكس صحيح إذا كان يعتقد أن متوسط دخل الفرد لا يزيد عن 200 دولار فإنه يختار الفرض البديل " أقل تماما من " أما إذا لم يكن لديه أي تصور أو أي معلومات فإنه يختار الفرض البديل "لا يساوي".

ب- إحصائية الإختبار (المحسوبة): وهي الإحصائية التي يتم حسابها من بيانات العينة بافتراض أن الفرض العدمي صحيح، ويتوقف شكل الإحصائية على العوامل التالية:

- توزيع المجتمع، وهل هو طبيعي أم لا، وهل تباينه معروف أم لا.
- حجم العينة، وهل هو كبير أم صغير.
- الفرض العدمي المراد اختباره وهل هو عن الوسط أو النسبة أو التباين أو الارتباط... الخ.
- الفرض البديل وهل الإختبار ذو طرفين أو طرف أيمن أو أيسر.

والفكرة الأساسية (غالباً) في إحصائية الاختبار هي: حساب الفرق بين قيمة المعلمة التي نفترضها للمجتمع (في الفرض العدمي) والقيمة المقابلة لها في العينة أي التابع الإحصائي، ثم نقسم (أو ننسب) هذا الفرق إلى الخطأ المعياري للتابع الإحصائي، فمثلاً إذا كان الإختبار عن الوسط الحسابي فإنه يتم حساب الفرق بين قيمة الوسط الحسابي للمجتمع التي نفترضها وقيمة الوسط الحسابي للعينة، ثم نقسم هذا الفرق على الخطأ المعياري للوسط، فلو أراد الباحث اختبار فرضية أن متوسط دخل الفرد في دولة ما هو مثلاً 200 دولار وللتأكد من مدى صحة هذه الفرضية فإنه عادة ما تسحب عينة عشوائية من المجتمع، ولنفرض أن متوسط دخل الفرد في هذه العينة كان 202 دولار، فالفرق هنا هو 2 دولار وهو فرق صغير بين الإفتراض والعينة الحقيقية فالباحث عادة ما يميل إلى قبول فرضه العدمي، أما إذا كان متوسط دخل الفرد في العينة مثلاً هو 250دو لار فالفرق هنا كبير بين الفرض والعينة، ولذا فإن احتمال رفض الفرض العدمي هو احتمال كبير نظراً لكبر الفرق بين قيمة الفرض والقيمة المحصلة من العينة، من هنا نستطيع القول بأن إحصائية الإختبار تعتمد على حساب الفرق بين قيمة الوسط المفترض وقيمة متوسط العينة.

هنا قد يثور تساؤل عن المعيار الذي يستطيع من خلاله الباحث الحكم على هذا الفرق ومدى كبره أو صغره، والإجابة الإحصائية عليه تتم من خلال قسمة هذا الفرق على الخطأ المعياري للوسط ثم مقارنة خارج القسمة بالقيمة الجدولية أو ما يسمى بحدود منطقتي القبول والرفض كما سوف نلاحظه لاحقاً.

ج- إيجاد القيمة الجدولية: مثل إحصائية الإختبار فإن القيمة الجدولية تتوقف على توزيع المجتمع وهل هو طبيعي أم لا وهل تباينه معروف أم لا ، كما تتوقف على حجم العينة وهل هو كبير أم صغير، يضاف إلى ذلك

الفرض العدمي المراد اختباره وهل هو عن الوسط أو النسبة أو التباين أو الإرتباط... الخ، بالإضافة إلى الفرض البديل وهل هو ذو طرفين أو ذو طرف أيسر أو أيمن (وهي الحالات التي نتناولها لاحقا).

د- المقارنة واتخاذ القرار: يتم قبول فرض العدم إذا كانت القيمة الجدولية أكبر من القيمة المحسوبة في حالتي اختبار الطرفين أو اختبار الطرف الأيمن، لكن في حالة اختبار الطرف الأيسر فإنه إذا كانت القيمة الجدولية أكبر من القيمة المحسوبة فإننا نقبل الفرض البديل.

2- اختبار الفروض للوسط الحسابي μ:

عند دراستنا لمقياس "الإحصاء 3" توصلنا إلى أن الوسط الحسابي يتبع أحد التوزيعين إما الطبيعي أو ستيودنت وذلك حسب الحالتين التاليتين:

نستخدم التوزيع الطبيعي إذا توفر أحد الشرطين:

أ-تباين المجتمع معلوم وتوزيعه طبيعي (لا يهم حجم العينة).

ب- تباين المجتمع معلوم وتوزيع المجتمع غير معلوم ولكن حجم العينة كبير $n \geq 30$ وهو منطوق نظرية النهاية المركزية.

نستخدم توزيع ستيودنت في حالة عدم توفر شروط استخدام التوزيع الطبيعي المذكورة سابقا.

وعليه فإننا نميز الحالات التالية:

2-1- اختبار الفروض للوسط الحسابي للمجتمع μ في حالة σ^2 معلوم و30 ≤ n (أو توزيع المجتمع طبيعي):

علمنا مما سبق في حال توفر الشروط المذكورة فإن المتغير العشوائي $Z = \frac{\bar{X}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}$ سيتوزع توزيعا طبيعيا معياريا، وعلى هذا الأساس فقد تمت كتابة مجال الثقة μ عند مستوى معنوية α (مستوى ثقة $1-\alpha$) بالعلاقة التالية:

$$\bar{X} - {}^{\sigma}/_{\sqrt{n}}\left(z_{1-{}^{\alpha}/_{2}}\right) \leq \mu \leq \bar{X} + {}^{\sigma}/_{\sqrt{n}}\left(z_{1-{}^{\alpha}/_{2}}\right)$$

فإذا افترض الباحث أن قيمة الوسط الحسابي للمجتمع هي μ_0 انطلاقا من معطيات أولية متوفرة، فالمطلوب منه هو اختبار ذلك الفرض عند مستوى معنوية α، وهذا يعني تحديد مجال الثقة للمتوسط ثم ملاحظة هل القيمة μ_0 تنتمي إلى المجال المقدر فيتم قبول الفرض H_0 $(\mu = \mu_0)$ أو لا تنتمي فيتم رفض الفرض العدمي وقبول الفرض البديل، أو بمعنى آخر هل المتغير العشوائي $Z_C = \frac{\bar{X}-\mu_0}{{}^{\sigma}/_{\sqrt{n}}}$ ينتمي إلى المجال $\left[-z_{1-{}^{\alpha}/_{2}} \;,\; z_{1-{}^{\alpha}/_{2}}\right]$ والإجابة على هذا التساؤل تحمل في طياتها اختبار الفرضيات الإحصائية كما يلي:

أ-اختبار الطرفين: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

ثانيا: حساب الإحصائية:

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

هذه الإحصائية التي سيتم مقارنتها مع القيمة الدنيا للمجال أي مع $-z_{1-\alpha/2}$ في حالة كانت $\bar{X}$ أقل من μ_0 (لاحظ أن المقدار $\sigma/\sqrt{n}$ دائما موجب)، أو مع $z_{1-\alpha/2}$ في حالة كانت $\bar{X}$ أكبر من μ_0، لذلك يمكن رياضيا أخذ القيمة المطلقة للإحصائية Z_C ومقارنتاها مع القيمة العليا للمجال أي $z_{1-\alpha/2}$ (القيمة المطلقة للقيمة الجدولية التي يأتي بيانها لاحقا)، لذلك فإن الإحصائية Z_C في اختبار الطرفين تأخذ الشكل التالي:

$$\boldsymbol{Z_C = \left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|}$$

ثالثا: إيجاد القيمة الجدولية: باستخدام جدول التوزيع الطبيعي يمكن استخراج القيمة الجدولية.

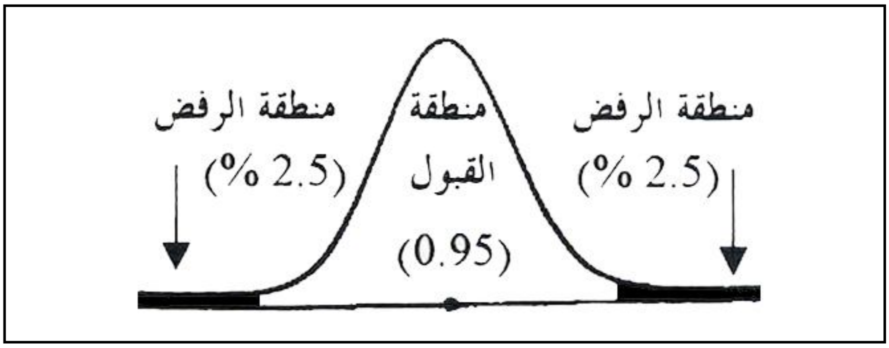
$$\boldsymbol{Z_{tab} = z_{1-\alpha/2}}$$

وفي علم الإحصاء عادة ما يتم استخدام مستوى المعنوية $\alpha = 5\%$ والتي تكون عندها $Z_{tab} = z_{0.975} = 1.96$

رابعا: المقارنة واتخاذ القرار:

إذا كانت Z_C أكبر من Z_{tab} فمعنى ذلك أن الإحصائية Z_C تقع خارج مجال الثقة $\left[-z_{1-\alpha/2}\ ,\ z_{1-\alpha/2}\right]$ وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 عند مستوى معنوية α.

إذا كانت Z_C أقل من أو يساوي Z_{tab} فمعنى ذلك أن الإحصائية Z_C تقع داخل مجال الثقة $\left[-z_{1-\alpha/2}\ ,\ z_{1-\alpha/2}\right]$ وبالتالي يتم رفض الفرض البديل H_1 وقبول الفرض العدمي H_0 عند مستوى معنوية α.



مثال 1:

يعتقد مدير الإنتاج في مصنع كوندور لتركيب الثلاجات بأن متوسط الإنتاج الأسبوعي هو 1000 ثلاجة، وللتأكد من ذلك قام باختيار عينة عشوائية حجمها 36 يوم إنتاج من مجتمع موزعا طبيعيا بمتوسط حسابي μ وتباين 324، وقد لوحظ أن متوسط الإنتاج المحسوب للعينة بلغ 1200 ثلاجة.

المطلوب: هل تؤيد بيانات العينة اعتقاد مدير الإنتاج عند مستوى معنوية $\alpha = 5\%$ ؟.

الحل:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 1000 \\ H_1: \mu \neq 1000 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \left| \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \right| = \left| \frac{1200 - 1000}{\sqrt{\frac{324}{36}}} \right| = 66.67$$

ج- إيجاد القيمة الجدولية:

$$Z_{tab} = z_{1-\alpha/2} = z_{0.975} = 1.96$$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_C أكبر من Z_{tab} وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 بمعنى أن متوسط الإنتاج الأسبوعي ليس 1000 ثلاجة أسبوعيا كما يعتقد مدير الإنتاج عند مستوى معنوية 5%.

ب- اختبار الطرف الأيمن: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$$

ثانيا حساب الإحصائية:

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

لاحظ عدم استخدام القيمة المطلقة نتيجة أن المقارنة ستتم مع طرف فقط وليس طرفين.

ثالثا: إيجاد القيمة الجدولية: باستخدام جدول التوزيع الطبيعي يمكن استخراج القيمة الجدولية.

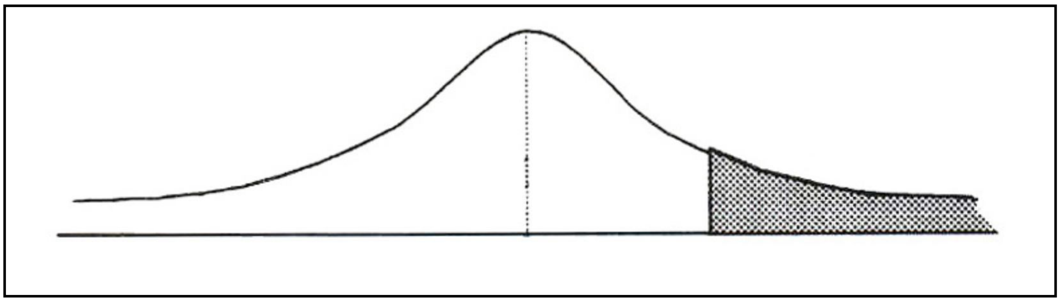
$$Z_{tab} = z_{1-\alpha}$$

لاحظ تمركز احتمالية الخطأ في الجانب الأيمن فقط.

وعند استخدام مستوى المعنوية $\alpha = 5\%$ تكون عندها $Z_{tab} = z_{0.95} = 1.645$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت Z_C أكبر من Z_{tab} فمعنى ذلك أن يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 عند مستوى معنوية α.
- إذا كانت Z_C أقل من أو يساوي Z_{tab} فمعنى ذلك أن يتم ر فض الفرض البديل H_1 وقبول الفرض العدمي H_0 عند مستوى معنوية α.



مثال 2:

ادعت شركة مساهمة أن متوسط مبيعاتها اليومية هي أكبر من 4000 دينار، ولغرض اختبار هذا الإدعاء من قبل المساهمين سحبت عينة من مبيعات الشركة اليومية لـ 49 يوم السابقة ووجدت أن متوسط المبيعات هو 4180، إذا علمت أن تباين المجتمع (المبيعات اليومية) هو 93000 فهل إدعاء الشركة صحيح عند مستوى معنوية 5%؟.

الحل:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 4000 \\ H_1: \mu > 4000 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{4180 - 4000}{\sqrt{\frac{93000}{49}}} = 4.13$$

ج- إيجاد القيمة الجدولية:

$$Z_{tab} = z_{1-\alpha} = z_{0.95} = 1.645$$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_C أكبر من Z_{tab} وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 بمعنى أن ادعاء الشركة صحيح وأن المبيعات اليومية هي أكبر من 4000 دينار عند مستوى معنوية 5%.

ج- اختبار الطرف الأيسر: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$$

ثانيا: حساب الإحصائية:

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

لاحظ عدم استخدام القيمة المطلقة نتيجة أن المقارنة ستتم مع طرف فقط وليس طرفين.

ثالثا: إيجاد القيمة الجدولية: باستخدام جدول التوزيع الطبيعي يمكن استخراج القيمة الجدولية.

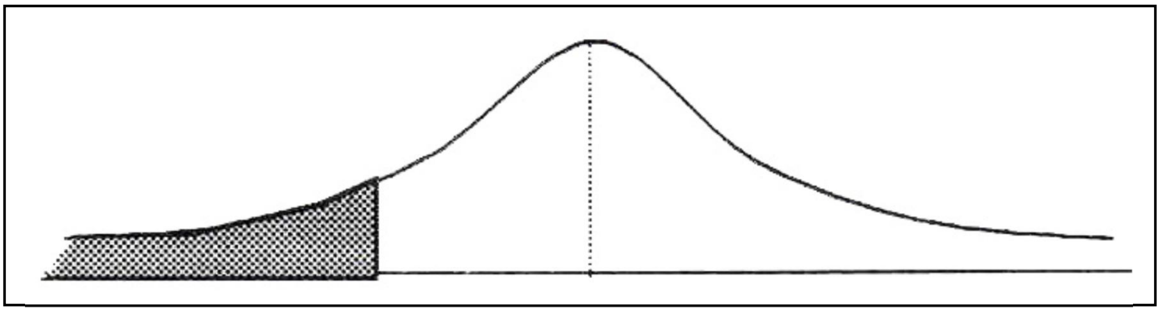
$$\boldsymbol{Z_{tab} = -z_{1-\alpha}}$$

لاحظ تمركز احتمالية الخطأ في الجانب الأيسر (القيمة الجدولية تكون بالسالب).

وعند استخدام مستوى المعنوية $\alpha = 5\%$ تكون عندها $Z_{tab} = -z_{0.95} = -1.645$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت Z_C أكبر من أو يساوي Z_{tab} فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α (لاحظ تغير التفسير مقارنة باختبار الطرفين أو اختبار الطرف الأيمن لأنه عند التعامل مع القيم السالبة فالقيمة الكبرى لـ Z_C تخرجه من مجال الخطأ).
- إذا كانت Z_C أقل من Z_{tab} فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية α.



مثال 3:

أخذت عينة بطريقة عشوائية من 100 كيس إسمنت من مصنع الإسمنت ووجد أن متوسط وزن الكيس هو 48 كلغ، فهل يمكن استنتاج أن متوسط الإنتاج هو أقل من المتوسط العام المفترض لوزن الكيس وهو 50 كلغ إذا علمت أن الإنحراف المعياري $\sigma = 5$ عند مستوى المعنوية $\alpha = 5\%$؟

الحل:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 50 \\ H_1: \mu < 50 \end{cases}$$

ب- إحصائية الإختبار: $Z_C = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{48-50}{5/\sqrt{100}} = -4$

ج- إيجاد القيمة الجدولية: $Z_{tab} = -z_{1-\alpha} = -z_{0.95} = -1.645$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_{tab} أكبر من Z_C وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0

بمعنى أن متوسط الإنتاج هو أقل من الوزن المفترض عند مستوى معنوية 5%.

2-2- اختبار الفروض للوسط الحسابي للمجتمع μ في حالة σ^2 مجهول أو n < 30:

علمنا فيما سبق أنه إذا كان حجم العينة صغيرا (n<30) أو كان تباين المجتمع σ^2 مجهولا، واستخدامنا تباين العينة S^2 كتقدير لذلك التباين المجهول σ^2، حيث أن

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

فإننا سنحصل على متغير عشوائي آخر هو المتغير t (عوضا عن Z) حيث: $t = \frac{\bar{X}-\mu_{\bar{x}}}{\sigma\bar{x}} = \frac{\bar{X}-\mu}{s/\sqrt{n}}$

والتوزيع الإحتمالي للمتغير العشوائي t هو توزيع ستيودنت بدرجة حرية v=n-1 ومستوى معنوية α (مستوى ثقة $1-\alpha$)، وعلى هذا الأساس فإن اختبار الفرضيات يكون بالشكل التالي:

أ- اختبار الطرفين: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

ثانيا: حساب الإحصائية:

$$t_C = \left| \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \right|$$

ثالثا: إيجاد القيمة الجدولية: باستخدام جدول توزيع ستيودنت يمكن استخراج القيمة الجدولية.

$$t_{tab} = t_{tab} = t_{\alpha/2,n-1}$$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت t_C أكبر من t_{tab} فمعنى ذلك أن الإحصائية t_C تقع خارج مجال الثقة $\left[-t_{\alpha/2,n-1} , t_{\alpha/2,n-1}\right]$ وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 عند مستوى معنوية α.

- إذا كانت t_C أقل من أو يساوي t_{tab} فمعنى ذلك أن الإحصائية t_C تقع داخل مجال الثقة $\left[-t_{\alpha/2\,,\mathrm{n-1}}\ ,\ t_{\alpha/2\,,\mathrm{n-1}}\right]$ وبالتالي يتم رفض الفرض البديل H_1 وقبول الفرض العدمي H_0 عند مستوى معنوية α.

ب- اختبار الطرف الأيمن: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$$

ثانيا: حساب الإحصائية:

$$t_C = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

لاحظ عدم استخدام القيمة المطلقة نتيجة أن المقارنة ستتم مع طرف فقط وليس طرفين.

ثالثا: إيجاد القيمة الجدولية: باستخدام جدول توزيع ستيودنت يمكن استخراج القيمة الجدولية.

$$\boldsymbol{t_{tab} = t_{\alpha\,,\mathrm{n-1}}}$$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت t_C أكبر من t_{tab} فمعنى ذلك أن يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 عند مستوى معنوية α.
- إذا كانت t_C أقل من أو يساوي t_{tab} فمعنى ذلك أن يتم رفض الفرض البديل H_1 وقبول الفرض العدمي H_0 عند مستوى معنوية α.

ج- اختبار الطرف الأيسر: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$$

ثانيا: حساب الإحصائية:

$$t_C = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

ثالثا: إيجاد القيمة الجدولية: باستخدام جدول توزيع ستيودنت يمكن استخراج القيمة الجدولية.

$$\boldsymbol{t_{tab} = -t_{\alpha\,,\mathrm{n-1}}}$$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت t_C أكبر من أو يساوي t_{tab} فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α (لاحظ تغير التفسير مقارنة باختبار الطرفين أو اختبار الطرف الأيمن لأنه عند التعامل مع القيم السالبة فالقيمة الكبرى لـ t_C تخرجه من مجال الخطأ).

- إذا كانت t_C أقل من t_{tab} فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفر ض البديل H_1 عند مستوى معنوية α.

ملاحظة هامة جدا: في حال كان تباين المجتمع مجهولا فإننا نستخدم توزيع ستيودنت مهما كان حجم العينة، حيث أن قاعدة التقارب بين التوزيعين ستيودنت والطبيعي تكون عندما يؤول حجم العينة إلى ما لا نهاية.

مثال 1:

أخذت عينة عشوائية من علامات 25 طالب في مادة الإحصاء لقسم العلوم الإقتصادية ووجد أن متوسط العلامات هو 12 من 20، فهل يمكن الإدعاء أن تحصيل الطلاب في هذا المقياس جيد وأنه يتجاوز معدل النجاح 10من 20 إذا علمت أن الإنحراف المعياري للعينة $S = 3$ وذلك عند مستوى المعنوية $\alpha = 5\%$؟.

- إذا كان متوسط علامات الطلاب في العينة يساوي 9 فهل يمكن الإدعاء أن المعدل العام هو أقل من معدل النجاح؟.
- إذا كان متوسط علامات الطلاب في العينة يساوي 10.5 فهل فهل يمكن الإدعاء أن المعدل العام يساوي معدل النجاح؟.

الحل:

البيان	اختبار الطرفين $(\bar{x} = 10.5)$	اختبار الطرف الأيمن $(\bar{x} = 12)$	اختبار الطرف الأيسر $(\bar{x} = 9)$
1- صياغة الفرضيات	$\begin{cases} H_0: \mu = 10 \\ H_1: \mu \neq 10 \end{cases}$	$\begin{cases} H_0: \mu = 10 \\ H_1: \mu > 10 \end{cases}$	$\begin{cases} H_0: \mu = 10 \\ H_1: \mu < 10 \end{cases}$
2- إحصائية الإختبار	$t_C = \left\lvert \frac{10.5 - 10}{3/\sqrt{25}} \right\rvert$ =0.83	$t_C = \frac{12 - 10}{3/\sqrt{25}}$ =3.33	$t_C = \frac{9 - 10}{3/\sqrt{25}}$ = - 1.67
3- القيمة الجدوليـة	$t_{tab} = t_{0.025,24}$ =2.064	$t_{tab} = t_{0.05,24}$ =1.711	$t_{tab} = -t_{0.05,24}$ = - 1.711
4- المقارنة واتخاذ القرار	$t_C < t_{tab}$ نقبل الفرض العدمي H_0 وبالتالي يمكن الإدعاء أن المعدل العام يساوي معدل النجاح عند مستوى معنوية 5%	$t_C > t_{tab}$ نقبل الفرض البديل H_1 وبالتالي يمكن الإدعاء أن المعدل العام أكبر من معدل النجاح عند مستوى معنوية 5%	$t_C > t_{tab}$ نقبل الفرض العدمي H_0 وبالتالي لا يمكن الإدعاء أن المعدل العام أقل من معدل النجاح عند مستوى معنوية 5%

3- اختبار الفروض للنسبة P:

علمنا عند تعرضنا لتوزيعات المعاينة أن توزيع المعاينة للنسبة P يقترب من التوزيع الطبيعي عندما يكون حجم العينة كبيرا (np≥5 et nq≥5)، وبنفس الطريقة التي أجريت بها الإختبارات حول المتوسط الحسابي فإن اختبارات النسبة تعطى كما يلي:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: P = P_0 \\ H_1: P \neq P_0 \end{cases}$	$\begin{cases} H_0: P = P_0 \\ H_1: P > P_0 \end{cases}$	$\begin{cases} H_0: P = P_0 \\ H_1: P < P_0 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\| \frac{p' - p_0}{\sqrt{\frac{p_0 . q_0}{n}}} \right\|$ لاحظ تغيير σ_p، في المقام بـ σ_{p_0}	$z_C = \frac{p' - p_0}{\sqrt{\frac{p_0 . q_0}{n}}}$	$z_C = \frac{p' - p_0}{\sqrt{\frac{p_0 . q_0}{n}}}$
3- القيمة الجدولية	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

مثال 2:

سحبت عينة عشوائية من مصنع للمصابيح الكهربائية تحتوي على 50 مصباح ووجد فيها 4 مصابيح تالفة، فهل يمكن الإدعاء أن نسبة المصابيح التالفة في الإنتاج الكلي للمصنع تتجاوز 5% وذلك عند مستوى المعنوية $\alpha = 5\%$؟.

-إذا عدد المصابيح التالفة هو 3 فهل يمكن الإدعاء أن نسبة المصابيح في الإنتاج الكلي للمصنع تساوي 5% ؟.

-إذا عدد المصابيح التالفة هو 2 فهل يمكن الإدعاء أن نسبة المصابيح في الإنتاج الكلي للمصنع أقل من 5% ؟.

الحل:

البيان	اختبار الطرفين 3 مصابيح تالفة	اختبار الطرف الأيمن 4 مصابيح تالفة	اختبار الطرف الأيسر 2 مصباح تالف
1- صياغة الفرضيات	$\begin{cases} H_0: P = 0.05 \\ H_1: P \neq 0.05 \end{cases}$	$\begin{cases} H_0: P = 0.05 \\ H_1: P > 0.05 \end{cases}$	$\begin{cases} H_0: P = 0.05 \\ H_1: P < 0.05 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\| \frac{0.06 - 0.05}{\sqrt{\frac{0.05 . 0.95}{50}}} \right\|$ =0.3244	$Z_C = \frac{0.08 - 0.05}{\sqrt{\frac{0.05 . 0.95}{50}}}$ =0.9733	$Z_C = \frac{0.04 - 0.05}{\sqrt{\frac{0.05 . 0.95}{50}}}$ = - 0.3244
3- القيمة الجدولية	$Z_{tab} = z_{0.975}$ =1.96	$Z_{tab} = z_{0.95}$ =1.645	$Z_{tab} = -z_{0.95}$ = -1.645

4- المقارنة واتخاذ القرار	$Z_C < Z_{tab}$ نقبل الفرض العدمي H_0 وبالتالي لا يمكن الإدعاء أن نسبة المصابيح التالفة في الإنتاج الكلي للمصنع تتجاوز 5% عند مستوى معنوية 5%	$Z_C < Z_{tab}$ نقبل الفرض العدمي H_0 وبالتالي يمكن الإدعاء أن نسبة المصابيح التالفة في الإنتاج الكلي للمصنع تساوي 5% عند مستوى معنوية 5%	$Z_C > Z_{tab}$ نقبل الفرض العدمي H_0 وبالتالي لا يمكن الإدعاء أن نسبة المصابيح التالفة في الإنتاج الكلي للمصنع أقل من 5% عند مستوى معنوية 5%

4- اختبار الفروض للتباين σ^2:

علمنا فيما سبق أنه إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بتباين معلوم σ^2 وسحبنا منه عينة عشوائية حجمها n، وحسبنا قيم المتغير العشوائي x^2 والذي يعطى بالصيغة التالية:

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

فالتوزيع الإحتمالي لهذا المتغير يسمى توزيع كاي تربيع x^2 بمستوى معنوية α ودرجة حرية v=n-1، بحيث:

$$P\left(x^2{}_{1-\alpha/2,n-1} \le x^2 \le x^2{}_{\alpha/2,n-1}\right)=1-\alpha$$

وبتعويض X^2 يكون الإحتمال كالتالي:

$$P\left(x^2{}_{1-\alpha/2,n-1} \le \frac{(n-1)S^2}{\sigma^2} \le x^2{}_{\alpha/2,n-1}\right)=1-\alpha$$

وعليه فإن:

$$x^2{}_{1-\alpha/2,n-1} \le \frac{(n-1)S^2}{\sigma^2} \le x^2{}_{\alpha/2,n-1}$$

لكن إذا كان تباين المجتمع σ^2 مجهول، وكانت هناك قيمة افتراضية $\sigma^2{}_0$ فإن اختبار هذا الفرض يكون حسب الحالات التالية:

أ- اختبار الطرفين: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

$$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 \neq \sigma^2{}_0 \end{cases}$$

ثانيا: حساب الإحصائية:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2{}_0}$$

ثالثا: إيجاد القيمة الجدولية: من خلال الصيغة المطروحة سابقا:

$$x^2{}_{1-\alpha/2,n-1} \le \frac{(n-1)S^2}{\sigma^2} \le x^2{}_{\alpha/2,n-1}$$

فحتى نقبل فرض العدم H_0 $(\sigma^2 = \sigma^2_0)$ يجب أن تكون القيمة $x^2{}_C = \frac{(n-1)S^2}{\sigma^2_0}$ داخل المجال:

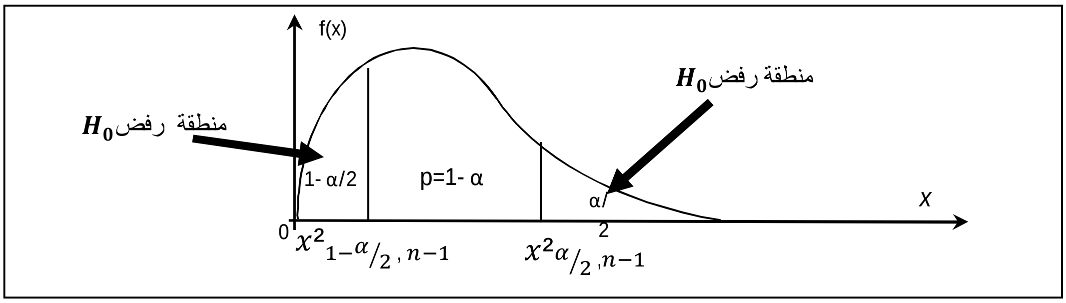
$\left[x^2{}_{1-\alpha/2, n-1},\ x^2{}_{\alpha/2, n-1}\right]$ لذلك فإن القيمة الجدولية تصبح مجال.

$$x^2{}_{tab} = \left[x^2{}_{1-\alpha/2, n-1},\ x^2{}_{\alpha/2, n-1}\right]$$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت قيمة $x^2{}_C$ داخل المجال $x^2{}_{tab}$ فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α.
- إذا كانت قيمة $x^2{}_C$ خارج المجال $x^2{}_{tab}$ فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية α.

منطقتي رفض الفرض العدمي إلى يسارويمين منحنى توزيع كاي تربيع



مثال 1:

اعتادت إحدى معامل إنتاج مسحوق الغسيل على إنتاج علبة بوزن 2000غ بانحراف معياري مقداره 5غ (تباين 25غ)، ولغرض قياس كفاءة الإنتاج من حيث أن وزن العلبة لا يزال ضمن القياسات المحددة وأن الإنحراف المعياري لا يزال 5غ سحبت عينة عشوائية بحجم 31 علبة وجد أن متوسط وزنها هو 1990غ بانحراف معياري قدره 7غ، فهل توجد فروق معنوية في إنتاج المعمل عما كان عليه سابقا من حيث التباين في وزن العلبة المنتجة عند مستوى معنوية 5%؟.

الحل:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \sigma^2 = \sigma^2_0 \\ H_1: \sigma^2 \neq \sigma^2_0 \end{cases} \Rightarrow \begin{cases} H_0: \sigma^2 = 25 \\ H_1: \sigma^2 \neq 25 \end{cases}$$

ب- إحصائية الإختبار:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2_0} = \frac{(31)\,49}{25} = 60.76$$

ج- إيجاد القيمة الجدولية:

$$x^2_{tab} = \left[x^2_{1-\alpha/2\ ,n-1}\ ,\ \ x^2_{\alpha/2\ ,n-1}\right] = \left[x^2_{0.975\ ,30}\ ,\ \ x^2_{0.025\ ,30}\right] = [16.791\ ,\ \ 46.979\ ]$$

د- المقارنة واتخاذ القرار:

نلاحظ أن قيمة $x^2{}_C$ خارج المجال x^2_{tab} فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية 5%، بمعنى أن الإنحراف المعياري اختلف مقارنة بالفترة السابقة.

ب- اختبار الطرف الأيمن: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

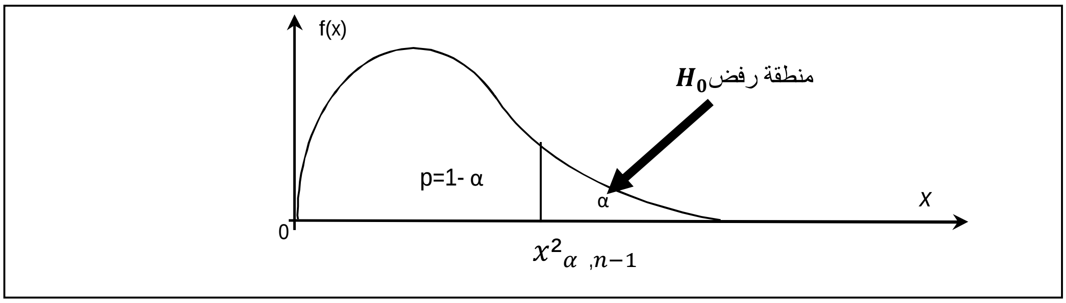
$$\begin{cases} H_0 : \sigma^2 = \sigma^2{}_0 \\ H_1 : \sigma^2 > \sigma^2{}_0 \end{cases}$$

ثانيا: حساب الإحصائية:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2{}_0}$$

ثالثا: إيجاد القيمة الجدولية: في هذه الحالة تتمركز منطقة الرفض يمين منحنى كاي تربيع وهو ما يظهره الشكل التالي:

منطقة رفض الفرض العدمي إلى يمين منحنى توزيع كاي تربيع



وبالتالي فإن القيمة الجدولية تعطى كما يلي:

$$x^2_{tab} = x^2{}_{\alpha\ ,n-1}$$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت قيمة $x^2{}_C$ أكبر من x^2_{tab} فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية α.
- إذا كانت قيمة $x^2{}_C$ أقل من x^2_{tab} فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α.

مثال 2:

إذا علمت بأن إنتاج الأبقار للحليب في مزرعة (A) يتوزع طبيعيا بتباين قدره 9 لتر، ولغرض زيادة الإنتاج تم اتباع نمط جديد من التغذية للأبقار (خليط جديد من الأعلاف) ووجد في عينة مكونة من 25 بقرة أن تباين إنتاجهم هو 12 لتر، فهل تؤيد بيانات العينة الإعتقاد بأن تباين إنتاج الحليب لأبقار المزرعة قد ارتفع نتيجة استخدام النمط الجديد من الأعلاف وذلك عند مستوى معنوية 5%؟.

الحل:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 > \sigma^2{}_0 \end{cases} \Rightarrow \begin{cases} H_0: \sigma^2 = 9 \\ H_1: \sigma^2 > 9 \end{cases}$$

ب- إحصائية الإختبار:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2{}_0} = \frac{(24)\,12}{9} = 32$$

ج- إيجاد القيمة الجدولية:

$$x^2{}_{tab} = x^2{}_{\alpha\ ,n-1} = x^2{}_{0.05\ ,24} = 36.415$$

د-المقارنة و اتخاذ القرار:

نلاحظ أن قيمة $x^2{}_C$ أقل من $x^2{}_{tab}$ فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية 5%، بمعنى أن النمط الجديد من التغذية لم يرفع من تباين إنتاج الحليب.

ج- اختبار الطرف الأيسر: وفيه يتم اتباع المراحل التالية:

أولا صياغة الفرضيات:

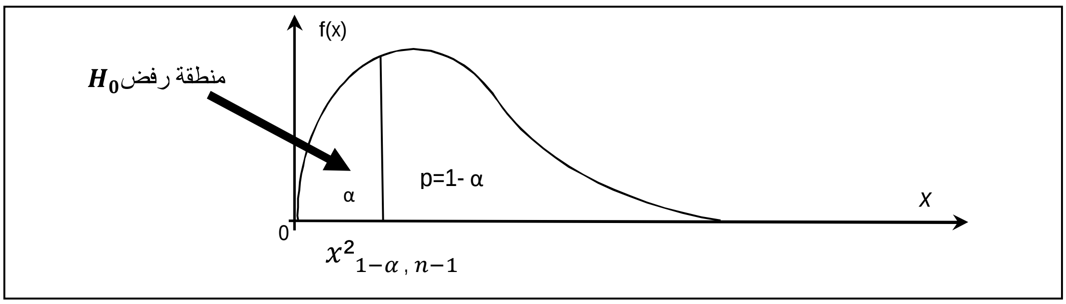
$$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 < \sigma^2{}_0 \end{cases}$$

ثانيا: حساب الإحصائية:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2{}_0}$$

ثالثا: إيجاد القيمة الجدولية: في هذه الحالة تتمركز منطقة الرفض يسار منحنى كاي تربيع وهو ما يظهره الشكل التالي:

منطقة رفض الفرض العدمي إلى يسار منحنى توزيع كاي تربيع



وبالتالي فإن القيمة الجدولية تعطى كما يلي:

$$x^2{}_{tab} = x^2{}_{1-\alpha\ ,n-1}$$

رابعا: المقارنة واتخاذ القرار:

- إذا كانت قيمة $x^2{}_C$ أكبر من $x^2{}_{tab}$ فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α.

- إذا كانت قيمة $x^2{}_C$ أقل من $x^2{}_{tab}$ فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية α.

مثال 3:

إذا علمت بأن علامات الإمتحان النهائي لطلبة الباكالوريا في مادة الرياضيات تتوزع طبيعيا بتباين قدره 25، إذا اتبعت طريقة جديدة في تدريس هذه المادة ويعتقد أن هذه الطريقة ستقلل من تباين علامات الطلبة، ولإختبار هذا الإعتقاد سحبت عينة عشوائية تشمل 25 طالب وتم تدريسهم بالطريقة الجديدة وأجري لهم امتحان فكان تباين علاماتهم 9، فهل تؤيد بيانات العينة الإعتقاد بأن الطريقة الجديدة تقلل تباين علامات الطلبة وذلك عند مستوى معنوية 5%؟.

الحل:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 < \sigma^2{}_0 \end{cases} \Rightarrow \begin{cases} H_0: \sigma^2 = 25 \\ H_1: \sigma^2 < 25 \end{cases}$$

ب- إحصائية الإختبار:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2{}_0} = \frac{(24)\ 9}{25} = 8.64$$

ج- إيجاد القيمة الجدولية:

$$x^2{}_{tab} = x^2{}_{1-\alpha\ ,n-1} = x^2{}_{0.95\ ,24} = 14.848$$

د- المقارنة واتخاذ القرار:

نلاحظ أن قيمة $x^2{}_C$ أقل من $x^2{}_{tab}$ فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية 5%، بمعنى أن النمط الجديد من التدريس قلل من تباين علامات الطلبة.

5- اختبار الفروض للفروق:

كما ذكرنا سابقا ففي بعض الأحيان نكون بحاجة إلى إجراء مقارنة بين مجتمعين مختلفين سواء فيما يخص المتوسط أو النسبة أو التباين، هذه المقارنات التي تستند إلى فرضيات نكون بحاجة على قبولها أو نفيها وهو ما يمكن أن نطلق عليه اختبارات الفروض للفروق.

5-1- اختبار الفروض للفرق بين وسطين:

وفيه نميز الحالتين التاليتين:

5-1-1-اختبار الفروض للفرق بين وسطين في حالة تباين المجتمعين $\sigma^2{}_1$ و $\sigma^2{}_2$ معلومين و $n_1, n_2 \geq 30$:

إذا كان لدينا مجتمعان يتوزعان توزيعا ما بوسط حسابي μ_1 وتباين معلوم $\sigma^2{}_1$ للأول ومتوسط حسابي μ_2 وتباين معلوم $\sigma^2{}_2$ للثاني، وأردنا وضع فروض إحصائية حول الفرق بين متوسطي هذين المجتمعين $(\mu_1 - \mu_2)$ وسحبنا لذلك عينتين عشوائيتين من المجتمعين حجمهما أكبر من 30 (حتى تتحقق شروط نظرية النهاية المركزية)، فستكون الفروض الإحصائية في إحدى الصور التالية:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\lvert\frac{\mu_{\bar{X}_1-\bar{X}_2}}{\sigma_{\bar{X}_1-\bar{X}_2}}\right\rvert = \left\lvert\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma^2{}_1}{n1}+\frac{\sigma^2{}_2}{n2}}}\right\rvert$	$Z_C = \frac{\mu_{\bar{X}_1-\bar{X}_2}}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma^2{}_1}{n1}+\frac{\sigma^2{}_2}{n2}}}$	$Z_C = \frac{\mu_{\bar{X}_1-\bar{X}_2}}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma^2{}_1}{n1}+\frac{\sigma^2{}_2}{n2}}}$
3- القيمة الجدولية	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – *والعكس صحيح*-	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – *والعكس صحيح*-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – *والعكس صحيح*-

مثال 1:

يرغب مشتري أن يقرر عند مستوى دلالة 5% أي صنف يشتري من صنفين لهما نفس السعر من مصابيح كهربائية، أخذ هذا المشتري عينة عشوائية من كل صنف حجمها 100 مصباح ووجد أن مصابيح الصنف الأول تستمر دون احتراق في المتوسط 980 ساعة بينما تستمر مصابيح الصنف الثاني في المتوسط 1000 ساعة، فأي الصنفين من المصابيح يجب شراؤه إذا علمت أن $\sigma^2{}_1 = 25$ وأن $\sigma^2{}_2 = 36$.

الحل:

في هذه الحالة نقوم بإجراء إختبار الطرفين فإذا تحقق الفرض العدم فإن الصنفين متساويين لذلك فإنه يشتري أي صنف، أما إذا تحقق الفرض البديل والتي يكون فيها الصنفين مختلفين فإنه ينتقل إلى المرحلة الثانية والتي يقوم فيها بإجراء أحد الإختبارين سواء اختبار الطرف الأيمن أو الأيسر، فإذا أجرى مثلا اختبار الطرف الأيمن في المرحلة الثانية وقبل بالفرض البديل فإنه يشتري الصنف الأول وإذا تحقق الفرض العدمي فبالضرورة سيتحقق الفرض البديل في اختبار الطرف الأيسر ويشتري الصنف الثاني، وهذا ما تبينه الإختبارات التالية:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\| \frac{980-1000}{\sqrt{\frac{25}{100}+\frac{36}{100}}} \right\|$ =25.61	$Z_C = \frac{980-1000}{\sqrt{\frac{25}{100}+\frac{36}{100}}}$ = - 25.61	$Z_C = \frac{980-1000}{\sqrt{\frac{25}{100}+\frac{36}{100}}}$ = - 25.61
3- القيمة الجدولية	$Z_{tab} = 1.96$	$Z_{tab} = 1.645$	$Z_{tab} = -1.645$
4- المقارنة واتخاذ القرار	$Z_C > Z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية 5% وبالتالي فالصنفين غير متساويين وننتقل للمرحلة الثانية	$Z_C < Z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية 5% – وبالتالي فالصنف الأول ليس الأفضل	$Z_C < Z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية 5% – وبالتالي فالصنف الثاني هو الأفضل

إذن فإن المشتري يختار الصنف الثاني عند مستوى دلالة 5%.

2-1-5- اختبار الفروض للفرق بين وسطين في حالة تباين المجتمعين $\sigma^2{}_1$ و $\sigma^2{}_2$ مجهولين أو $n_1, n_2 < 30$

في حال كان تبايني المجتمعين مجهولين وتم تقديرهما انطلاقا من تبايني العينتين $S1^2$ و$S2^2$ (بغض النظر عن حجمي العينتين) أو كان التوزيعين الإحتماليين للمجتمعين مجهولين وحجمي العينتين أقل من 30 مشاهدة فإنه يتم استخدام توزيع ستيودنت عوض التوزيع الطبيعي وذلك وفق الحالات التالية:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$
2- إحصائية الإختبار	$t_C = \left\| \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2{}_1}{n1}+\frac{S^2{}_2}{n2}}} \right\|$	$t_C = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2{}_1}{n1}+\frac{S^2{}_2}{n2}}}$	$t_C = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2{}_1}{n1}+\frac{S^2{}_2}{n2}}}$
3- القيمة الجدولية	$t_{tab} = t_{\alpha/2, n_1+n_1-2}$	$t_{tab} = t_{\alpha, n_1+n_1-2}$	$t_{tab} = -t_{\alpha, n_1+n_1-2}$
4- المقارنة واتخاذ القرار	$t_C \leq t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$t_C \leq t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$t_C < t_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

مثال 2:

انطلاقا من المثال السابق وبافتراض أن تبايني المجتمعين مجهولين حيث تم تقديرهما بتبايني العينتين أن $S^2{}_1 = 25$ و$S^2{}_2 = 36$ وأن حجمي العينتين هو 51 لكل منهما.

الحل:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$
2- إحصائية الإختبار	$t_C = \left\lvert \frac{980-1000}{\sqrt{\frac{25}{51}+\frac{36}{51}}} \right\rvert$ =18.29	$t_C = \frac{980-1000}{\sqrt{\frac{25}{51}+\frac{36}{51}}}$ = - 18.29	$t_C = \frac{980-1000}{\sqrt{\frac{25}{51}+\frac{36}{51}}}$ = - 18.29
3- القيمة الجدولية	$t_{tab} = t_{\alpha/2, n_1+n_1-2}$ = $t_{0.025,100}$ =1.984	$t_{tab} = t_{\alpha, n_1+n_1-2}$ = $t_{0.05,100}$ =1.660	$t_{tab} = -t_{\alpha, n_1+n_1-2}$ =- $t_{0.05,100}$ =-1.660
4- المقارنة واتخاذ القرار	$t_C > t_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية 5%وبالتالي فالصنفين غير متساويين وننتقل للمرحلة الثانية	$t_C < t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية 5%– وبالتالي فالصنف الأول ليس الأفضل	$t_C < t_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية 5%– وبالتالي فالصنف الثاني هو الأفضل

إذن فإن المشتري يختار الصنف الثاني عند مستوى دلالة 5%.

2-5- اختبار الفروض للفرق بين نسبتين في حالة $n_1, n_2 \geq 30$:

إذا كان لدينا مجتمعان، وكانت نسبة ظاهرة معينة في المجتمع الأول P_1 ونسبة هذه الظاهرة في المجتمع الثاني P_2، وأردنا وضع فروض إحصائية حول الفرق بين النسبتين $(P_1 - P_2)$ وسحبنا لذلك عينتين عشوائيتين من المجتمعين حجمهما أكبر من 30، فستكون الفروض الإحصائية في إحدى الصور التالية:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \neq P_2 \end{cases}$	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 > P_2 \end{cases}$	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 < P_2 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\lvert \frac{\mu_{p'_1-p'_2}}{\sigma_{p'_1-p'_2}} \right\rvert = \left\lvert \frac{p'_1-p'_2}{\sqrt{\frac{p'_1q'_1}{n1}+\frac{p'_2q'_2}{n2}}} \right\rvert$	$Z_C = \frac{\mu_{p'_1-p'_2}}{\sigma_{p'_1-p'_2}} = \frac{p'_1-p'_2}{\sqrt{\frac{p'_1q'_1}{n1}+\frac{p'_2q'_2}{n2}}}$	$Z_C = \frac{\mu_{p'_1-p'_2}}{\sigma_{p'_1-p'_2}} = \frac{p'_1-p'_2}{\sqrt{\frac{p'_1q'_1}{n1}+\frac{p'_2q'_2}{n2}}}$
3- القيمة الجدولية	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$Z_C \leq Z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α –والعكس صحيح-	$Z_C \leq Z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α –والعكس صحيح-	$Z_C < Z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α –والعكس صحيح-

مثال 1: ترغب شركة أن تختار عند مستوى معنوية 5% ما إذا كانت نسبة القبول من المكونات الإلكترونية (أي الرقاقات الإلكترونية الصالحة) لمورد أجنبي P_1 أفضل من تلك التي يوفرها المورد المحلي P_2 أو العكس، ولأجل ذلك أخذت الشركة عينة عشوائية من شحنة كل مورد ووجدت أن $P'_1 = 0.9$ و $P'_2 = 0.7$ من عينات ذات الحجم: $n_1 = 100$ و $n_2 = 80$، فأي نسبة قبول هي الأفضل.

الحل:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \neq P_2 \end{cases}$	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 > P_2 \end{cases}$	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 < P_2 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\lvert \frac{0.9-0.7}{\sqrt{\frac{0.9\times0.1}{100}+\frac{0.7\times0.3}{80}}} \right\rvert = 3.37$	$Z_C = \frac{0.9-0.7}{\sqrt{\frac{0.9\times0.1}{100}+\frac{0.7\times0.3}{80}}} = 3.37$	$Z_C = \frac{0.9-0.7}{\sqrt{\frac{0.9\times0.1}{100}+\frac{0.7\times0.3}{80}}} = 3.37$
3- القيمة الجدولية	$Z_{tab} = 1.96$	$Z_{tab} = 1.645$	$Z_{tab} = -1.645$
4- المقارنة واتخاذ القرار	$Z_C > Z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية 5% وبالتالي فالنسبتين غير متساويين وننتقل للمرحلة الثانية	$Z_C > Z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية 5% – وبالتالي فنسبة القبول للمورد الأجنبي أفضل	$Z_C > Z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية 5% – وبالتالي فنسبة القبول للمورد الأجنبي ليست أقل من تلك الموردة محليا

إذن فإن الشركة تختار المكونات الإلكترونية للمورد الأجنبي عند مستوى دلالة 5%.

3-5- اختبار الفروض للنسبة بين تباينين:

إذا كان لدينا مجتمعان يتوزع كل منهما توزيعا طبيعيا، الأول تباينه $\sigma^2{}_1$ والثاني $\sigma^2{}_2$، وأردنا إجراء اختبار خاص بمقارنة تلك التباينات، فإن إحصائية الإختبار المناسبة في هذه الحالة هي المتغير العشوائي F والذي يعرف كما يلي:

$F = \frac{S^2{}_1/\sigma_1{}^2}{S^2{}_2/\sigma_2{}^2}$ والذي يخضع إلى توزيع فيشر بمستوى معنوية α ودرجتي حريتين $v_1=n_1-1$ و $v_2=n_2-1$، وبذلك فقد تم تحديد مجال الثقة للنسبة بين تباينين كما يلي:

$$F_{1-\alpha/2,(v1,v2)} \leq \frac{S^2{}_1/\sigma_1{}^2}{S^2{}_2/\sigma_2{}^2} \leq F_{\alpha/2,(v1,v2)}$$

ومنه: $$F_{1-\alpha/2,(v1,v2)} \leq \frac{\sigma_2{}^2}{\sigma_1{}^2} \times \frac{S^2{}_1}{S^2{}_2} \leq F_{\alpha/2,(v1,v2)}$$

ونتيجة لكون فرضية العدم: $H_0: \sigma^2{}_1 = \sigma^2{}_2$ فمعنى ذلك أن $H_0: \frac{\sigma_2{}^2}{\sigma_1{}^2} = 1$ وهو ما يعني أنه لقبول فرضية العدم يجب أن تكون النسبة $\frac{S^2{}_1}{S^2{}_2}$ داخل المجال $\left[F_{1-\alpha/2,(v1,v2)}\ ,\ F_{\alpha/2,(v1,v2)}\right]$، ومن هذا المنطلق يمكن إجراء اختبارات الفروض للنسبة بين تباينين وفق الحالات التالية:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 \neq \sigma^2{}_2 \end{cases}$	$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 > \sigma^2{}_2 \end{cases}$	$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 < \sigma^2{}_2 \end{cases}$
2- إحصائية الإختبار	$F_C = \frac{S^2{}_1}{S^2{}_2}$	$F_C = \frac{S^2{}_1}{S^2{}_2}$	$F_C = \frac{S^2{}_1}{S^2{}_2}$
3- القيمة الجدولية	$F_{tab} = \left[F_{1-\alpha/2\,,(v1,v2)}\,,F_{\alpha/2\,,(v1,v2)}\right]$	$F_{tab} = F_{\alpha\;,(v1,v2)}$	$F_{tab} = F_{1-\alpha\;,(v1,v2)}$
4- المقارنة واتخاذ القرار	إذا كانت F_C داخل مجال F_{tab} نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C \le z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

مثال:

إذا علمت أن مجتمع علامات الطالبات ومجتمع علامات الطلبة في جامعة برج بوعريريج يتبعان التوزيع الطبيعي، وسحبنا عينة عشوائية تشمل علامات 25 طالبة، ومن علامات الطلبة عينة عشوائية تشمل 21، ووجد أن تباين علامات عينة الطالبات يساوي 16 وأن تباين علامات الطلبة يساوي 9، اختبر ما إذا كان هناك فرق بين تباين مجتمع علامات الطالبات والطلبة عند مستوى معنوية 5% ؟.

الحل:

في هذه الحالة فالإختبار هو اختبار الطرفين (يوجد فرق أو لا يوجد فرق)، لكن نستعين بإختبارات الطرف الأيمن والأيسر لزيادة التوضيح.

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 \neq \sigma^2{}_2 \end{cases}$	$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 > \sigma^2{}_2 \end{cases}$	$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 < \sigma^2{}_2 \end{cases}$
2- إحصائية الإختبار	$F_C = \frac{S^2{}_1}{S^2{}_2} = \frac{16}{9} = 1.78$	$F_C = \frac{S^2{}_1}{S^2{}_2} = \frac{16}{9} = 1.78$	$F_C = \frac{S^2{}_1}{S^2{}_2} = \frac{16}{9} = 1.78$
3- القيمة الجدولية	$F_{tab} = \left[F_{0.975\,,(24,20)}\,,F_{0.025\,,(24,20)}\right]$ =[0.43 ,2.41]	$F_{tab} = F_{0.05\;,(24,20)}$ =2.08	$F_{tab} = F_{0.95\;,(24,20)}$ =0 .49
4- المقارنة واتخاذ القرار	F_C داخل مجال F_{tab} نقبل الفرض العدمي H_0 – معناه التباينين متساويين عند مستوى معنوية 5%	$F_C \le F_{tab}$ نقبل الفرض العدمي H_0 معناه التباينين متساويين عند مستوى معنوية 5%	$z_C > z_{tab}$ نقبل الفرض العدمي H_0 معناه التباينين متساويين عند مستوى معنوية 5%

نلاحظ إثبات مختلف الإختبارات لصحة فرضية العدم عند مستوى معنوية 5%.

ملخص اختبار الفروض الإحصائية

- اختبار الفروض للوسط الحسابي للمجتمع μ في حالة σ^2 معلوم و$n \geq 30$:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$	$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$	$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\lvert \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right\rvert$	$Z_C = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$Z_C = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$
3- القيمة الجدوليـــة	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

- اختبار الفروض للوسط الحسابي للمجتمع μ في حالة σ^2 مجهول أو$n < 30$

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$	$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$	$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$
2- إحصائية الإختبار	$t_C = \left\lvert \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right\rvert$ **لاحظ تغيير σ بـ S**	$t_C = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$t_C = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$
3- القيمة الجدوليـــة	$t_{tab} = t_{\alpha/2, n-1}$	$t_{tab} = t_{\alpha, n-1}$ **لاحظ تركز احتمال الخطأ في الجانب الأيمن**	$t_{tab} = -t_{\alpha, n-1}$ **لاحظ أن القيمة الجدولية مسبوقة بإشارة (-)**
4- المقارنة واتخاذ القرار	**$t_C \leq t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α** – والعكس صحيح-	**$t_C \leq t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α** – والعكس صحيح-	**$t_C < t_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α** – والعكس صحيح-

- اختبار الفروض للنسبة P:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: P = P_0 \\ H_1: P \neq P_0 \end{cases}$	$\begin{cases} H_0: P = P_0 \\ H_1: P > P_0 \end{cases}$	$\begin{cases} H_0: P = P_0 \\ H_1: P < P_0 \end{cases}$
2- إحصائية الإختبار	$z_C = \left\lvert \dfrac{p' - p_0}{\sqrt{\dfrac{p_0 \cdot q_0}{n}}} \right\rvert$ لاحظ تغيير $\sigma_{p'}$ في المقام بـ σ_{p_0}	$z_C = \dfrac{p' - p_0}{\sqrt{\dfrac{p_0 \cdot q_0}{n}}}$	$z_C = \dfrac{p' - p_0}{\sqrt{\dfrac{p_0 \cdot q_0}{n}}}$
3- القيمة الجدولية	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$z_C \le z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α - والعكس صحيح-	$z_C \le z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α - والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α - والعكس صحيح-

- اختبار الفروض للتباين σ^2:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 \neq \sigma^2{}_0 \end{cases}$	$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 > \sigma^2{}_0 \end{cases}$	$\begin{cases} H_0: \sigma^2 = \sigma^2{}_0 \\ H_1: \sigma^2 < \sigma^2{}_0 \end{cases}$
2- إحصائية الإختبار	$x^2{}_C = \dfrac{(n-1)S^2}{\sigma^2{}_0}$	$x^2{}_C = \dfrac{(n-1)S^2}{\sigma^2{}_0}$	$x^2{}_C = \dfrac{(n-1)S^2}{\sigma^2{}_0}$
3- القيمة الجدولية	$x^2{}_{tab} = \left[x^2{}_{1-\alpha/2,\,n-1} \,,\; x^2{}_{\alpha/2,\,n-1}\right]$	$x^2{}_{tab} = x^2{}_{\alpha\,,n-1}$	$x^2{}_{tab} = x^2{}_{1-\alpha\,,n-1}$
4- المقارنة واتخاذ القرار	إذا كانت قيمة $x^2{}_C$ داخل المجال $x^2{}_{tab}$ فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α-	إذا كانت قيمة $x^2{}_C$ أكبر من $x^2{}_{tab}$ فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1 عند مستوى معنوية α- والعكس صحيح-	إذا كانت قيمة $x^2{}_C$ أكبر من $x^2{}_{tab}$ فمعنى ذلك أن يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 عند مستوى معنوية α - والعكس صحيح-

- اختبار الفروض للفرق بين وسطين في حالة تباين المجتمعين $\sigma^2{}_1$ و $\sigma^2{}_2$ معلومين و$n_1, n_2 \geq 30$:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\lvert \frac{\mu_{\bar{X}_1-\bar{X}_2}}{\sigma_{\bar{X}_1-\bar{X}_2}} \right\rvert = \left\lvert \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma^2{}_1}{n1}+\frac{\sigma^2{}_2}{n2}}} \right\rvert$	$Z_C = \frac{\mu_{\bar{X}_1-\bar{X}_2}}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma^2{}_1}{n1}+\frac{\sigma^2{}_2}{n2}}}$	$Z_C = \frac{\mu_{\bar{X}_1-\bar{X}_2}}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma^2{}_1}{n1}+\frac{\sigma^2{}_2}{n2}}}$
3- القيمة الجدولية	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

- اختبار الفروض للفرق بين وسطين في حالة تباين المجتمعين $\sigma^2{}_1$ و $\sigma^2{}_2$ مجهولين أو$n_1, n_2 < 30$

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$
2- إحصائية الإختبار	$t_C = \left\lvert \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S^2{}_1}{n1}+\frac{S^2{}_2}{n2}}} \right\rvert$	$t_C = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S^2{}_1}{n1}+\frac{S^2{}_2}{n2}}}$	$t_C = \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S^2{}_1}{n1}+\frac{S^2{}_2}{n2}}}$
3- القيمة الجدولية	$t_{tab} = t_{\alpha/2, n_1+n_1-2}$	$t_{tab} = t_{\alpha, n_1+n_1-2}$	$t_{tab} = -t_{\alpha, n_1+n_1-2}$
4- المقارنة واتخاذ القرار	$t_C \leq t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$t_C \leq t_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$t_C < t_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

- اختبار الفروض للفرق بين نسبتين في حالة $n_1, n_2 \geq 30$:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \neq P_2 \end{cases}$	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 > P_2 \end{cases}$	$\begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 < P_2 \end{cases}$
2- إحصائية الإختبار	$Z_C = \left\lvert\frac{\mu_{p'_1-p'_2}}{\sigma_{p'_1-p'_2}}\right\rvert = \left\lvert\frac{p'_1-p'_2}{\sqrt{\frac{p'_1q'_1}{n1}+\frac{p'_2q'_2}{n2}}}\right\rvert$	$Z_C = \frac{\mu_{p'_1-p'_2}}{\sigma_{p'_1-p'_2}} = \frac{p'_1-p'_2}{\sqrt{\frac{p'_1q'_1}{n1}+\frac{p'_2q'_2}{n2}}}$	$Z_C = \frac{\mu_{p'_1-p'_2}}{\sigma_{p'_1-p'_2}} = \frac{p'_1-p'_2}{\sqrt{\frac{p'_1q'_1}{n1}+\frac{p'_2q'_2}{n2}}}$
3- القيمة الجدولية	$Z_{tab} = z_{1-\alpha/2}$	$Z_{tab} = z_{1-\alpha}$	$Z_{tab} = -z_{1-\alpha}$
4- المقارنة واتخاذ القرار	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

- اختبار الفروض للنسبة بين تباينين:

البيان	اختبار الطرفين	اختبار الطرف الأيمن	اختبار الطرف الأيسر
1- صياغة الفرضيات	$\begin{cases} H_0: \sigma^2_1 = \sigma^2_2 \\ H_1: \sigma^2_1 \neq \sigma^2_2 \end{cases}$	$\begin{cases} H_0: \sigma^2_1 = \sigma^2_2 \\ H_1: \sigma^2_1 > \sigma^2_2 \end{cases}$	$\begin{cases} H_0: \sigma^2_1 = \sigma^2_2 \\ H_1: \sigma^2_1 < \sigma^2_2 \end{cases}$
2- إحصائية الإختبار	$F_C = \frac{S^2_1}{S^2_2}$	$F_C = \frac{S^2_1}{S^2_2}$	$F_C = \frac{S^2_1}{S^2_2}$
3- القيمة الجدولية	$F_{tab} = \left[F_{1-\alpha/2,(v1,v2)}, F_{\alpha/2,(v1,v2)}\right]$	$F_{tab} = F_{\alpha,(v1,v2)}$	$F_{tab} = F_{1-\alpha,(v1,v2)}$
4- المقارنة واتخاذ القرار	إذا كانت F_C داخل مجال F_{tab} نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C \leq z_{tab}$ نقبل الفرض العدمي H_0 عند مستوى معنوية α – والعكس صحيح-	$z_C < z_{tab}$ نقبل الفرض البديل H_1 عند مستوى معنوية α – والعكس صحيح-

السلسلة رقم: 01-

التمرين الأول: (اختبار ثنائي الإتجاه للمتوسط)

تنتج شركة مصابيح كهربائية، وترغب في معرفة إذا ما كان يمكنها الإدعاء بأن مصابيحها الكهربائية تستمر 1000 ساعة دون احتراق، أخذت عينة عشوائية من هذه المصابيح حجمها 101 مصباح ووجدت أن المتوسط الحسابي لحياة المصابيح هو 980 ساعة بانحراف معياري 80 ساعة، إختبر ما إذا كان إدعاء الشركة صحيحا عند $\alpha = 0.05$

- أجري نفس الإختبار في حال كان الإنحراف المعياري للمجتمع $\sigma = 81$.

التمرين الثاني: (اختبار الطرف الأيسر للمتوسط)

للتحقق من شكاوى المستهلكين بخصوص نقص تعبئة علب المنظف حجم 500غ، أخذت مصلحة الجودة عينة من 101 علبة، فكان الوزن المتوسط في العينة 497غ. إذا كان $\sigma = 3g$. هل تؤيد بيانات العينة دعاوى المستهلكين بمستوى معنوية $\alpha = 0.05$. أجري نفس الإختبار في حال كان تباين المجتمع مجهول وأن $s=2.5g$.

التمرين الثالث: (اختبار النسبة)

في مؤسسة ما، 20% من العمال أجورهم أقل من الحد الأدنى المضمون، أخذت عينة عشوائية حجمها 30 شخص، ووجد أن نسبة العمال الذين أجورهم أقل من الحد الأدنى المضمون 34%، هل توجد فروق معنوية بين النسبة المشاهدة والنسبة النظرية عند مستوى دلالة $\alpha = 0.05$.

- في حالة كانت النسبة داخل العينة 15%. هل يمكن الإدعاء بأن نسبة العمال (الذين أجورهم أقل من الحد الأدنى المضمون) أقل من 20% عند مستوى دلالة $\alpha = 0.05$.

التمرين الرابع: (اختبار الفرق بين وسطين)

يرغب مدير مؤسسة أن يحدد عند مستوى معنوية 5% إذا ما كان الأجر بالساعة للعمال متساويا في مدينتين، أخذ عينتين عشوائيتين من المدينتين حجمهما على التوالي 45 عاملا و57 عامل، ووجد أن المتوسط الحسابي للعينة الأولى يساوي 6دج للساعة و5.4 دج/سا للعينة الثانية، بانحرافات معيارية على التوالي 2دج و1.8دج.

- اختبر إذا ما كانت هناك فروق معنوية بين الأجور في المدينتين.

التمرين الخامس: (اختبار التباين)

من الخبرة الماضية وجدنا أن الإنحراف المعياري لمجتمع ما هو 0.25، أخذنا عينة عشوائية حجمها 20 ووجدنا أن انحرافها المعياري يساوي 0.32، هل لهذا الإرتفاع في الإنحراف دلالة عند مستوى معنوية 0.05.

التمرين السادس: (اختبار النسبة بين تباينين)

يعطي أستاذ دروسه لمجموعتين من الطلبة، حجمهما على التوالي 25 و21 طالب، بعد إجراء الإمتحان وجد أن التباين للمجموعتين هو على التوالي 9 و12، هل يوجد فروق عند معنوية 5% بين التباينين.

التمرين السابع:

آلة صناعية مختصة في إنتاج قطع حديدية دائرية متوسط قطرها 25 ملم بانحراف معياري 1.5 ملم. بعد مدة أصاب الآلة عطل وتم تصليحه، أخذت عينة عشوائية من منتوج الآلة بعد تصليحها وتم قياس الأقطار المنتجة فكانت النتائج كمايلي (بالملم): 22، 23، 21، 25، 26، 22، 26، 21.

- هل يمكن الجزم بأن الآلة مازالت تعمل كما كانت سابقا عند معنوية 5% ؟.

الحل:

التمرين الأول:

لدينا المعطيات التالية: $\mu_0 = 1000$، $\bar{x} = 980$، $s = 80$، $n = 101$

على اعتبار أن تباين المجتمع مجهول فالتوزيع المناسب هو توزيع ستيودنت.

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 1000 \\ H_1: \mu \neq 1000 \end{cases}$$

ب- إحصائية الإختبار:

$$t_C = \left| \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \right| = \left| \frac{980 - 1000}{80/\sqrt{101}} \right| = 2.51$$

ج- إيجاد القيمة الجدولية:

$$t_{tab} = t_{\alpha/2, n-1} = t_{0.025, 100} = 1.984$$

د- المقارنة واتخاذ القرار:

نلاحظ أن t_C أكبر من t_{tab} وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 بمعنى أنه لا يمكن الادعاء أن المصابيح تستمر 1000 ساعة دون احتراق عند مستوى معنوية 5%.

- في حال كان الإنحراف المعياري للمجتمع $\sigma = 81$ وعلى اعتبار أن حجم العينة يتجاوز 30 مشاهدة فالتوزيع المناسب هو التوزيع الطبيعي.

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 1000 \\ H_1: \mu \neq 1000 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| = \left|\frac{980 - 1000}{81/\sqrt{101}}\right| = 2.48$$

ج- إيجاد القيمة الجدولية:

$$Z_{tab} = z_{1-\alpha/2} = z_{0.975} = 1.96$$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_C أكبر من Z_{tab} وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 بمعنى أنه لا يمكن الادعاء أن المصابيح تستمر 1000 ساعة دون احتراق عند مستوى معنوية 5%.

التمرين الثاني:

لدينا المعطيات التالية: $n = 101$ ،$\sigma = 3$ ،$\bar{x} = 497$ ،$\mu_0 = 500$

الإنحراف المعياري للمجتمع معلوم وحجم العينة يتجاوز 30 مشاهدة فالتوزيع المناسب هو التوزيع الطبيعي.

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 500 \\ H_1: \mu < 500 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{497 - 500}{3/\sqrt{101}} = -10.05$$

ج- إيجاد القيمة الجدولية:

$$Z_{tab} = -z_{1-\alpha} = -z_{0.95} = -1.645$$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_C أقل من Z_{tab} وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 بمعنى أن دعاوي المستهلكين مؤسسة عند مستوى معنوية 5%.

- في حال كان تباين المجتمع مجهول وأن S=2.5g.

التوزيع المناسب هو توزيع ستيودنت.

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases} \Rightarrow \begin{cases} H_0: \mu = 500 \\ H_1: \mu < 500 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{497 - 500}{2.5/\sqrt{101}} = -12.06$$

ج- إيجاد القيمة الجدولية:

$$t_{tab} = -t_{\alpha\,,n-1} = -t_{0.05\,,100} = -1.660$$

د- المقارنة واتخاذ القرار:

نلاحظ أن t_C أقل من t_{tab} وبالتالي يتم قبول الفرض البديل H_1 ورفض الفرض العدمي H_0 بمعنى أن دعاوي المستهلكين مؤسسة عند مستوى معنوية 5%.

التمرين الثالث:

لدينا المعطيات التالية: $P_0 = 0.2$ ،$p' = 0.34$ ،$n = 30$

نلاحظ أن حجم العينة يتجاوز 30 مشاهدة فالتوزيع المناسب هو التوزيع الطبيعي.

أ- صياغة الفرضيات:

$$\begin{cases} H_0: P = P_0 \\ H_1: P \neq P_0 \end{cases} \Rightarrow \begin{cases} H_0: P = 0.2 \\ H_1: P \neq 0.2 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \left| \frac{p' - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}} \right| = \left| \frac{0.34 - 0.2}{\sqrt{\frac{0.2 \cdot 0.8}{30}}} \right| = 1.92$$

ج- إيجاد القيمة الجدولية:

$$Z_{tab} = z_{1-\alpha/2} = z_{0.975} = 1.96$$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_C أقل من Z_{tab} وبالتالي يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 بمعنى أنه لا توجد فروق معنوية بين النسبة المشاهدة والنسبة النظرية عند مستوى معنوية 5%.

- في حالة كانت النسبة داخل العينة 15%.

الإختبار هو اختبار الطرف الأيسر:

أ- صياغة الفرضيات:

$$\begin{cases} H_0: P = P_0 \\ H_1: P < P_0 \end{cases} \Rightarrow \begin{cases} H_0: P = 0.2 \\ H_1: P < 0.2 \end{cases}$$

ب- إحصائية الإختبار:

$$Z_C = \frac{p' - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}} = \frac{0.15 - 0.2}{\sqrt{\frac{0.2 \cdot 0.8}{30}}} = -0.68$$

ج- إيجاد القيمة الجدولية:

$$Z_{tab} = -z_{1-\alpha} = -z_{0.95} = -1.645$$

د- المقارنة واتخاذ القرار:

نلاحظ أن Z_C أكبر من Z_{tab} وبالتالي يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 بمعنى أنه لا يمكن الإدعاء بأن نسبة العمال (الذين أجورهم أقل من الحد الأدنى المضمون) أقل من 20% عند مستوى معنوية 5%.

التمرين الرابع:

لدينا المعطيات التالية: $\bar{x}_1 = 6$ ، $\bar{x}_2 = 5.4$ ، $S_1 = 2$ ، $S_2 = 1.8$ $n_1 = 45$ ، $n_2 = 57$

وباعتبار أن تبايني المجتمعين مجهولين فالتوزيع المناسب هو توزيع ستيودنت.

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

ب- إحصائية الإختبار:

$$t_C = \left| \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2_1}{n1} + \frac{S^2_2}{n2}}} \right| = \left| \frac{6 - 5.4}{\sqrt{\frac{2^2}{45} + \frac{1.8^2}{57}}} \right| = 1.57$$

ج- إيجاد القيمة الجدولية:

$$t_{tab} = t_{\alpha/2, n_1 + n_1 - 2} = t_{0.025, 100} = 1.984$$

د- المقارنة واتخاذ القرار:

نلاحظ أن t_C أقل من t_{tab} وبالتالي يتم قبول الفرض العدمي H_0 ورفض الفرض البديل H_1 بمعنى أنه لا توجد فروق معنوية بين الأجور في المدينتين عند مستوى معنوية 5%.

التمرين الخامس:

من الخبرة الماضية وجدنا أن الإنحراف المعياري لمجتمع ما هو 0.25، أخذنا عينة عشوائية حجمها 20 ووجدنا أن انحرافها المعياري يساوي 0.32، هل لهذا الإرتفاع في الإنحراف دلالة عند مستوى معنوية 0.05.

لدينا المعطيات التالية: $\sigma_0 = 0.25$ $S = 0.32$ $n = 20$

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \sigma^2 = \sigma^2_0 \\ H_1: \sigma^2 > \sigma^2_0 \end{cases} \Rightarrow \begin{cases} H_0: \sigma^2 = 0.25 \\ H_1: \sigma^2 > 0.25 \end{cases}$$

ب- إحصائية الإختبار:

$$x^2{}_C = \frac{(n-1)S^2}{\sigma^2{}_0} = \frac{(19)\ 0.32^2}{0.25^2} = 31.13$$

ج- إيجاد القيمة الجدولية:

$$x^2{}_{tab} = x^2{}_{\alpha\ ,n-1} = x^2{}_{0.05\ ,19} = 30.144$$

د- المقارنة واتخاذ القرار:

نلاحظ أن قيمة $x^2{}_C$ أكبر من $x^2{}_{tab}$ فمعنى ذلك أن يتم رفض الفرض العدمي H_0 وقبول الفرض البديل H_1، بمعنى أن لهذا الإرتفاع في الإنحراف دلالة عند مستوى معنوية 5%.

التمرين السادس:

لدينا المعطيات التالية: $S^2{}_1 = 9$، $S^2{}_2 = 5.4$، $n_1 = 25$، $n_2 = 21$

أ- صياغة الفرضيات:

$$\begin{cases} H_0: \sigma^2{}_1 = \sigma^2{}_2 \\ H_1: \sigma^2{}_1 \neq \sigma^2{}_2 \end{cases}$$

ب- إحصائية الإختبار:

$$F_C = \frac{S^2{}_1}{S^2{}_2} = \frac{9}{12} = 0.75$$

ج- إيجاد القيمة الجدولية:

$$F_{tab} = [F_{0.975\ ,(24,20)}\ ,F_{0.025\ ,(24,20)}] = [0.43\ ,2.41]$$

د- المقارنة واتخاذ القرار:

نلاحظ أن F_C داخل مجال F_{tab} نقبل الفرض العدمي H_0– معناه التباينين متساويين عند مستوى معنوية 5% (لا توجد فروق بين التباينين).

التمرين السابع:

1- اختبار المتوسط:

- صياغة الفرضيات:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases} \qquad \begin{cases} H_0: \mu = 25 \\ H_1: \mu \neq 25 \end{cases}$$

- إحصائية الإختبار:

$$t_{cal} = \left|\frac{m-u_0}{\sigma_m}\right| = \left|\frac{m-u_0}{s/\sqrt{n-1}}\right|$$

حساب $\bar{x}$ وS:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{22+23+\cdots.26+21}{8} = 23.25$$

$$s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} = \frac{(22-23.25)^2+(23-23.25)^2+\cdots+(21-23.25)^2}{7} = 4.50$$

$$s = \sqrt{4.50} = 2.12$$

ومنه:

$$t_{cal} = \left|\frac{\bar{x} - u_0}{s/\sqrt{n-1}}\right| = \left|\frac{23.25 - 25}{2.12/\sqrt{8}}\right| = 2.33$$

- القيمة الجدولية:

$$t_{tab} = t_{\alpha/2\ ,n-1} = t_{0.025\ ,7} = 2.365$$

- المقارنة والقرار:

مما سبق نجد أن: $t_{tab} < t_{cal}$ وعليه نقبل H_0 ونرفض H_1، وعليه فمتوسط الأقطار بقي مساويا لـ25 ملم مستوى معنوية 5%.

2- اختبار التباين:

- صياغة الفرضيات: $\begin{cases} H_0: \alpha^2 = \alpha_0{}^2 \\ H_1: \alpha^2 \neq \alpha_0{}^2 \end{cases}$ $\begin{cases} H_0: \alpha^2 = 2.25 \\ H_1: \alpha^2 \neq 2.25 \end{cases}$

- إحصائية الإختبار:

$$\chi^2{}_{cal} = \frac{(n-1)\ s^2}{\alpha_0{}^2} = \frac{7 \times 4.50}{2.25} = 14$$

- القيمة الجدولية:

$$x^2{}_{tab} = \left[x^2{}_{1-\alpha/2\ ,n-1}\ ,\ x^2{}_{\alpha/2\ ,n-1}\right] = \left[x^2{}_{0.975\ ,7}\ ,\ x^2{}_{0.025\ ,7}\right] = [1.690\ ,16.013]$$

- المقارنة والقرار:

مما سبق نجد أن: $\chi^2{}_{cal}$ تنتمي إلى المجال المحدد لـ $\chi^2{}_{tab}$ وعليه نقبل H_0 ونرفض H_1، وعليه فالتباين للأقطار بقي مساويا لـ2.25 ملم مستوى معنوية 5%.

القرار العام: يمكن الجزم بأن الآلة مازالت تعمل كما كانت عند مستوى معنوية 5% لأن المتوسط والإنحراف المعياري بقيا على حالهما.

تمارين مقترحة:

التمرين الأول:

اختيرت عينة عشوائية حجمها 64 مشاهدة من مجتمع يخضع للتوزيع الطبيعي بوسط حسابي μ وتباين $\sigma^2 = 25$، وقد لوحظ أن متوسط مشاهدات العينة بلغ $\bar{x} = 20$.

المطلوب: هل تعتقد أن متوسط المجتمع يقل عن 23 عند مستوى معنوية 5%.

التمرين الثاني:

ادعت إدارة إحدى المستشفيات أن متوسط وقت انتظار المريض لتلقي خدمة الطبيب تقل عن 20 دقيقة، وللتحقق من ذلك اختيرت عينة عشوائية من المرضى المراجعين حجمها 64 مريض وتبين أن متوسط وقت انتظار المريض هو 22 دقيقة بتباين قدره 100 دقيقة.

المطلوب: اختبر صحة ادعاء إدارة المستشفى عند مستوى معنوية 5%.

التمرين الثالث:

للمقارنة بين الأجور اليومية للعمال في الشركتين (أ و ب) في مدينة برج بوعريريج، والتي تخضع الأجور فيها للتوزيع الطبيعي $N(\mu_1, 16)$ و $N(\mu_2, 9)$ على الترتيب، اختيرت عينة عشوائية من عمال الشركة (أ) بحجم 64 عامل وتبين أن متوسط أجورهم بلغ 5 دج، واختيرت عينة عشوائية من عمال الشركة (ب) بحجم 36 عامل وتبين أن متوسط أجورهم بلغ 3 دج.

المطلوب: هل يمكن استنتاج أن متوسط الأجور اليومية للعمال في الشركة (أ) يختلف عن متوسط الأجور اليومية للعمال في الشركة (ب) عند مستوى معنوية 5%.

التمرين الرابع:

للمقارنة بين رواتب أعضاء هيئة التدريس في جامعتي برج بوعريريج وفيلادلفيا، اختيرت عينة عشوائية من أساتذة جامعة برج بوعريريج بحجم 50 أستاذ وتبين أن متوسط رواتبهم بلغ 50000 دج وانحراف معياري قدره 10000دج، واختيرت عينة عشوائية من أساتذة جامعة فيلادلفيا بحجم 60 أستاذ وتبين أن متوسط رواتبهم بلغ 60000 دج وانحراف معياري قدره 20000دج.

المطلوب: هل يمكن استنتاج أن متوسط رواتب هيئة التدريس في جامعة برج بوعريريج يقل عن رواتب نظرائهم في جامعة فيلادلفيا عند مستوى معنوية 5%.

التمرين الخامس:

ادعت إحدى الشركات المتخصصة في صناعة الحقائب الجلدية بأن ما لا يقل عن 97% من انتاجها مطابق لمواصفات الجودة، قامت مصلحة مراقبة جودة الإنتاج في المؤسسة العامة التي تنتمي إليها الشركة

المذكورة باختيار عينة عشوائية حجمها 200 حقيبة وبعد فحصها تبين أن عدد الحقائب المطابقة للمواصفات بلغ 180 حقيبة.

المطلوب: هل ادعاء الشركة صحيح عند مستوى معنوية 5%.

التمرين السادس:

اختيرت عينة عشوائية حجمها 700 من طلبة جامعة برج بوعريريج، مكونة من 360 طالب و340 طالبة، وقد لوحظ أن عدد الطلاب الراسبين كان 50 طالب، وعدد الطالبات الراسبات هو 30 طالبة.

المطلوب: هل يمكن استنتاج أن نسبة الطلبة الذكور الراسبين أعلى من نسبة الطالبات الراسبات عند مستوى معنوية 5%.

التمرين السابع:

اختيرت عينة عشوائية حجمها 40 مشاهدة وتم حساب الإنحراف المعياري لها ووجد أنه يساوي 15، فهل يمكن الإدعاء أن تباين المجتمع هو 230 عند معنوية 5%.

التمرين الثامن:

للمقارنة بين الإنحرافات المعيارية للإنفاق الشهري على إحدى السلع الإستهلاكية في المدينتين (أ و ب)، اختيرت عينة عشوائية من الأسر في المدينة (أ) حجمها 60 أسرة وتبين أن الإنحراف المعياري للإنفاق الشهري فيها 12دج، واختيرت عينة عشوائية من الأسر في المدينة (ب) حجمها 80 أسرة وتبين أن الإنحراف المعياري للإنفاق الشهري فيها 15دج.

المطلوب: هل يمكن استنتاج أن هاتين المدينتين متجانستين في نمط إنفاقها على تلك السلعة عند مستوى معنوية 5%.

المحور الثالث:

اختبارات الإستقلالية والتوافق

الأهداف التعليمية

- التعرف على اختبار التجانس؛
- اختبار الإستقلال أو الإقتران ؛
- تحليل التباين (التصنيف الأحادي للبيانات)؛

تمهيد:

يعتبر توزيعي x^2 و F من أشهر وأهم الأدوات الإحصائية المستخدمة في تحليل الظواهر الاقتصادية سواء الوصفية منها أو غير الوصفية، لذا فغالبا ما لا تخلو الدراسات والأبحاث الاقتصادية التي تنتهج الأسلوب الكمي أو السلوكي من تطبيق أو استخدام هاذين الأسلوبين في التحليل الإحصائي، وهو ما نحاول معالجته في هذه الجزئية من الدروس المقدمة.

1- استخدام اختبار x^2 في التحليل الإحصائي:

الفكرة الأساسية من استخدام هذا الأسلوب هي مقارنة التكرارات الفعلية أو المشاهدة بالمتوقعة، فمثلا عند رمي قطعة نقود (وليكن مائة مرة) بافتراض أن هذه القطعة سليمة، فإنه من المتوقع أن نصف هذه الرميات صورة والنصف الآخر كتابة، إلا أنه من الناحية العملية من النادر أن نحصل على النصف (تماما) صورة والنصف الآخر كتابة، وبالتالي كلما قل الفرق بينهما فإن ذلك يعني أن هناك تجانسا بين الصورة والكتابة- والعكس صحيح كلما كبر الفرق بينهما كلما كان ذلك مؤشرا على أن وجهي القطعة غير متجانسين، وللتأكد من ذلك يستخدم اختبار يسمى " باختبار التجانس ".

فعندما تتوفر بيانات عن الظاهرة محل الدراسة في شكل تكرارات (تسمى التكرارات المشاهدة Observed Frequencies) فإن مقارنة هذه التكرارات بما هو متوقع يمكننا من التوصل إلى بعض خصائص المجتمع محل الدراسة، كذلك فعند دراسة العلاقة بين مستوى التعليم في المجتمع ودرجة الوعي الاقتصادي مثلا فإنه بمقارنة التكرارات الفعلية بالتكرارات المتوقعة يمكننا من التوصل إلى نتيجة فيما إذا كانت هاتان الظاهرتان مستقلتين أم أن هناك علاقة أو ارتباط فيما بينهما.

وهكذا فقد تكون البيانات مصنفة تبعا لمتغير واحد أو لمتغيرين أو ظاهرتين ففي مثل هذه الحالات – التي تعتمد على تحليل التكرارات المشاهدة ومقارنتها بالتكرارات المتوقعة - فإن التوزيع الذي يسمى " مربع كاي " (Chi-square dist) هو أداة التحليل الإ حصائي المناسبة لمثل هذه الحالات.

ومن مزايا هذا الأسلوب في التحليل – أنه اختبار غير معلمي Non-Parametric بمعنى أنه لا يعتمد على طبيعة التوزيعات التي تتبعها البيانات علاوة على وجود جداول إحصائية للتوزيع تعطي القيم المختلفة عند مختلف درجات الحرية، ولمزيد من الإيضاح نسوق المثال التالي :

افترض أننا نريد قياس مردودية العمال، ثم قمنا بإجراء هذه الدراسة على مجموعة (عددها n) من العمال، وكانت التكرارات المشاهدة لمردودية العمال (بافتراض أن المردودية مقسمة إلى عدد k من الدرجات) هي:

$O_1, O_2, \dots, O_K$ والتكرارات المتوقعة هي : $e_1, e_2, \dots, e_K$

فإنه كلما كانت الفروق بين التكرارات المشـــاهـدة والتكرارات المتوقعة طفيفة كلما كنا نميل إلى قبول الفرض بأن التكرارات المشـــاهـدة تتفق مع التكرارات المتوقعة- والعكس صـحيح- أي كلما كانت الفروق بين التكرارات المشـــاهـدة والمتوقعة كبيرة كان هـذا دليلا على ضـعـف التطابق وبالتالي رفض النموذج أو التوزيع المفترض، **والمقياس** الذي يحدد إلى أي مدى تتفق التكرارات المشاهدة مع التكرارات المتوقعة هو الذي يسمى كا² أو X^2 يأخذ الشكل التالي :

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i}$$

بدرجات حرية تساوي $K - 1$ حيث O_i ترمز للتكرار المشـاهد في الخلية (أو الفئة) رقم i، e_i ترمز للتكرار المتوقع في الخلية (أو الفئة) رقم i , k ترمز لعدد الخلايا (أو الفئات أو التقسـيمات). فلو تم على سـبيل المثال تقسـيم مردودية العمال إلى أربعة أقسـام هي: عالي جدا – عال – متوسـط – منخفض) فإن K في هذه الحالة = 4.

ونلاحظ أن X^2 تؤول إلى الصفر إذا ما كان هناك تطابق كامل بين التكرارات المشاهدة والتكرارات المتوقعة. وأن قيمة X^2 تكبر كلما كبرت الفروق بينهما، بمعنى أن زيادة قيمة X^2 تعني ضـعـف التطابق بين التكرارات المشـاهدة والمتوقعة وبالتالي يكون احتمال رفض الفرض العدمي كبيرا في هذه الحالات والعكس صـحيح بطبيعة الحال. ويشـترط ألا يقل أي تكرار متوقع عن 5 تكرارات في كل خلية حتى يكون اسـتخدام هذا النوع من التحليل مناسبا، ومن التطبيقات المعروفة لتوزيع X^2:

1– اختبار التجانس (أو جودة التوفيق).

2 – اختبار الاستقلال.

أ- اختبار التجانس (اختبار جودة التوافق) :

يعتبر اختبار التجانس – أو التماثل – كما سـبق وأن أشـرنا أحد التطبيقات المهمة لتوزيع X^2 فلو أراد الباحث اختبار ما إذا كانت لمجموعة من الاعلانات الأفضلية نفسها أو (التجانس أو التساوي في الأفضلية) لدى المسـتهلكين فإن المطلوب في مثل هذه الحالات هو اختبار فرض التجانس (أو التماثل) بين الفئات أو الخلايا أو التقسيمات المختلفة (الاعلانات في هذا المثال) وذلك مقابل الفرض البديل أنها غير متجانسة.

وتكون خطوات اختبار التجانس كما يلي :-

1 – الفرض العدمي : هو فرض التجانس (أو التماثل).

2 – الفرض البديل : هو عدم التجانس.

3 – الإحصائية : وتأخذ الإحصائية الشكل التالي :

$$X^2=\sum_{I}^{K}\frac{(O-e)^2}{e}$$

والتي لها توزيع كا^2 بدرجات حرية K–1 حيث:

K- هي عدد الخلايا أو الأقسام أو الفئات أو الاعلانات...

- O ترمز للتكرارات المشاهدة

- e ترمز للتكرارات المتوقعة

مع ملاحظة أن مجموع التكرارات المشاهدة يساوي مجموع التكرارات المتوقعة، أي أن :

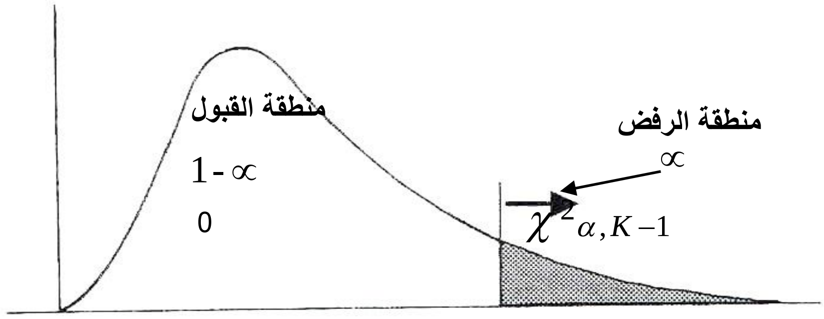
$$\sum O=\sum e$$

<u>ويتم حساب التكرارات المتوقعة بقسمة مجموع التكرارات المشاهدة على عدد الخلايا K</u>، فإذا كان عدد الاعلانات هي 4، وعدد المستهلكين (المستهدفين الذين يتابعون الاعلان) هم 100 مستهلك، فإن التكرارات المتوقعة لكل البرامج هي $\frac{100}{4}=25$ أي نعيد توزيع مجموع التكرارات المشاهدة بالتساوي (أو بالتماثل) بين الفئات أو الخلايا المختلفة. ثم نحسب الفروق بين التكرارات المشاهدة والمتوقعة، ثم نربع هذه الفروق ونقسم هذه المربعات على التكرارات المتوقعة فنحصل على قيمة الإحصائية. ويمكن تنظيم الحل أو الحصول على قيمة الإحصائية كما في الجدول التالي :

الخلايا أو الأقسام	التكرارات المشاهدة **O**	التكرارات المتوقعة **e**	الفرق بين التكرارات المشاهدة والتكرارات المتوقعة (o-e)	*مربعات الفروق* $(o-e)^2$	$\frac{(O-e)^2}{e}$
1	O_1	e_1	(O_1-e_1)	$(O_1-e_1)^2$	$(O_1-e_1)^2/e_1$
2	O_2	e_2	(O_2-e_2)	$(O_2-e_2)^2$	$(O_2-e_2)^2/e_2$
K	O_k	e_k	(O_k-e_k)	$(O_k-e_k)^2$	
المجموع	$\sum O$	$\sum e$	Zero		$X^2=\sum\frac{(O-e)^2}{e}$

ومجموع العمود الأخير يمثل قيمة الإحصائية χ^2.

<u>4 – حدود منطقتي القبول والرفض :</u> يستخدم هنا اختبار الطرف الأيمن لتوزيع كا^2 بدرجات حرية K-1 عند مستوى المعنوية ∝ كما في الشكل التالي :



ونلاحظ ما يلي :

أ – أن توزيع X^2 **توزيع غير متماثل**، بل هو **توزيع ملتو** ناحية اليمين (موجب الالتواء).

ب – أنه غير معرف في المنطقة السالبة، أو بمعنى آخر فإنه معرف فقط في المنطقة الموجبة والتي تبدأ من **الصفر**.

ج – أن **منطقة القبول** تبدأ من الصفر وحتى القيمة $X^2_{\alpha,K-1}$ والتي نحصل عليها من جدول كا^2 عند مستوى معنوية يساوي α ودرجات حرية تساوي $K-1$.

د – أن **منطقة الرفض** تشمل القيم التي أكبر من $X^2_{\alpha,K-1}$ (اختبار الطرف الأيمن).

5 – المقارنة والقرار: حيث تتم مقارنة قيمة الإحصائية (المحسوبة من الخطوة الثالثة) بحدود منطقتي القبول والرفض، فإذا وقعت قيمة الإحصائية في منطقة القبول فإن القرار يكون قبول الفرض العدمي (أي قبول فرض التجانس)، والعكس إذا وقعت قيمة الإحصائية في منطقة الرفض يكون القرار هو رفض الفرض العدمي وقبول الفرض البديل. أي قبول فرض عدم التجانس وذلك بمستوى معنوية يساوي α.

مثال1:

عند سؤال 400 مشاهد عن أفضل برنامج تليفزيوني لديهم من بين أربعة برامج مشهورة بغرض الإشهار في فواصل إحدى البرامج، كانت إجاباتهم كما يلي :

البرنامج	الأول	الثاني	الثالث	الرابع	المجموع
عدد المشاهدين	120	150	80	50	400

اختبر ما إذا كانت للبرامج الأربعة الأفضلية نفسها لدى المشاهدين (أي اختبر فرض التجانس بين البرامج بمستوى معنوية 1 %، أو بتعبير آخر: اختبر ما إذا كان هناك تجانس (أو تماثل) في الأفضلية بالنسبة للبرامج لدي المشاهدين.

الحل:

1 – الفرض العدمي: التجانس بين البرامج (أو أن للبرامج المختلفة الأفضلية نفسها لدى المشاهدين).

2 – الفرض البديل: عدم التجانس بين البرامج (ليس للبرامج نفس الأفضلية لدى المشاهدين).

3 – الإحصائية : وهي : $X^2 = \sum_{1}^{K} \frac{(O-e)^2}{e}$

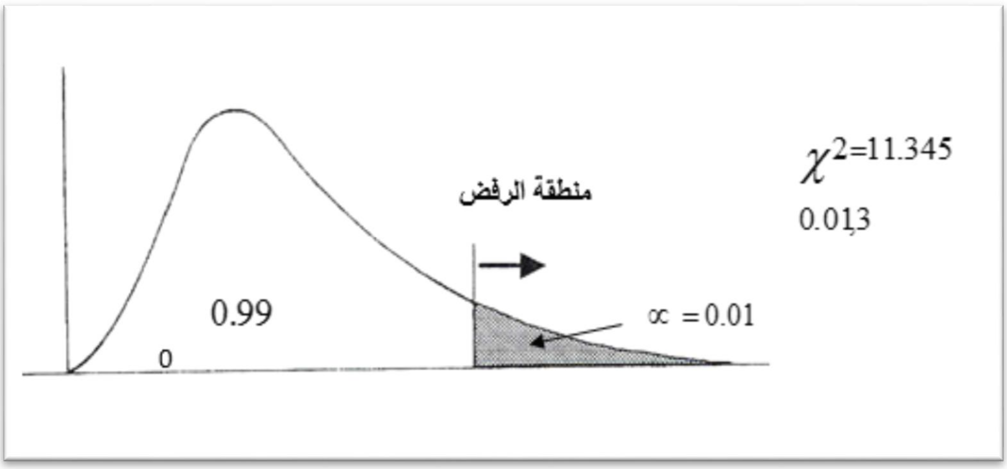
ونحصل عليها كما في الجدول التالي :

البرامج	عدد المشاهدين O	$e = \frac{400}{4}$	O – e	$(O-e)^2$	$(O-e)^2 / e$
1	120	100	20	400	4
2	150	100	50	2500	25
3	80	100	- 20	400	4
4	50	100	- 50	2500	25
المجموع	400	400	Zero		$X^2 = 58$

أي أن قيمة الإحصائية تساوي (58). لاحظ أننا حصلنا على التكرارات المتوقعة بقسمة 400 على 4 أي بقسمة مجموع التكرارات المشاهدة (مجموع المشاهدين) على عدد البرامج.

4 – حدود منطقتي القبول والرفض : اختبار الطرف الأيمن لتوزيع X^2 حيث مستوى المعنوية هو $\propto = 0.01$ ودرجات حرية تساوي K – 1 وحيث أن عدد البرامج أربعة أي K = 4 فإن درجات الحرية هي K –1 أي تساوي 1 – 4 أي تساوي **3** ومن جدول X^2 نجد أن :

$X^2_{0.01,3} = 11.345$ وبالتالي فإن منطقتي القبول والرفض كما في الشكل التالي :



أي أن منطقة القبول تبدأ من **الصفر** وحتى 11.345 وتكون منطقة الرفض للقيم التي أكبر من 11.345.

5 – المقارنة والقرار : وحيث أن قيمة الإحصائية (من الخطوة الثالثة والتي تساوي 58) أكبر بكثير من 11.345 أي تقع في منطقة الرفض فإن القرار **هو : رفض الفرض العدمي.**

(أي رفض فرض التجانس بين البرامج الأربعة) **وقبول الفرض البديل** بأن البرامج **غير متجانسة** أي **ليس لها الأفضلية نفسها** لدى المشاهدين وذلك بمستوى معنوية 1% (أي أن احتمال الخطأ في هذا القرار لا يتعدى 1%).

مثال 2:

وجد محل تجاري من خبرته السابقة أن 30% من التليفزيونات المباعة من الحجم الصغير، 40% من الحجم المتوسط، 30% من الحجم الكبير، لتحديد حجم المخزون الواجب الاحتفاظ به من كل نوع، أخذ المدير عينة عشوائية من 100 من المبيعات الحديثة للتلفزيون فوجد أن 20 من النوع الصغير أي 20%، و40 من النوع المتوسط (40%) و40 من النوع الكبير، باستخدام معنوية 5%، يختبر المدير الفرض أن نمط المبيعات الماضي لا يزال سائدا.

الحجم	نسبة المبيعات O	E	O – e	$(O-e)^2$	$(O – e)^2 / e$
الصغير	20	30	-10	100	3.33
المتوسط	40	40	0	0	0
الكبير	40	30	10	100	3.33
المجموع	100	100	Zero		$X^2 = 6.66$

وبالمقارنة مع القيمة الجدولية بمعنوية 5% ودرجة K-1=3-1=2 نجد القيمة الجدولية 5.991 نجد أن كاي المحسوبة أكبر من الجدولية لذلك فالقرار نرفض H_0 ونقبل الفرض البديل أي أن النمط القديم من البيع ليس سائدا.

ب- اختبار الإستقلال أو معامل الاقتران : Coefficient of Association

يستخدم معامل الاقتران لقياس العلاقة بين ظاهرتين تنقسم كل منهما إلى قسمين (أو صفتين) فقط، وتكون البيانات موضوعة في جدول مزدوج. ويسمى الجدول في هذه الحالة " جدول الاقتران حيث توضع إحدى الظاهرتين أفقيا ، والأخرى رأسيا ، ويكون الشكل العام لجدول الاقتران كما يلي :

الجدول رقم 05: جدول الاقتران 2X2

الظاهرة الثانية / الظاهرة الأولى	الخاصة (1)	الخاصة (2)
الخاصية (1)	A	B
الخاصية (2)	C	D

نلاحظ أن الجدول أعلاه يتكون من أربع خلايا، حيث تمثل الحروف، A,B,C,D تكرارات هذه الخلايا.

فالحرف A يمثل تكرار أو عدد المفردات التي تجمع بين الخاصية رقم (1) لكل من الظاهرتين، وهكذا مع باقي التكرارات.

مثال:

جمع تاجر سيارات البيانات الموضحة في الجدول أدناه عن عدد السيارات الأجنبية والمحلية التي يشتريها عملاء أعمارهم تحت سن 30 سنة، والتي يشتريها عملاء أعمارهم 30 سنة أو أكثر، وذلك لإختبار ما إذا كان نوع السيارة المشتراة (أجنبية أو محلية) مستقبلا عن سن المشتري باستخدام معنوية 5%. (هل للسن دور في تحديد نوع السيارة المشتراة) أي أن **H_0 هي أن السن ليس عاملا لتحديد نوع السيارة، والفرض البديل أنه عامل.**

الإجمالي	نوع السيارة		السن
	أجنبية	محلية	
70	40	30	تحت 30 سنة
100	80	20	30 سنة أو أكثر
170	120	50	إجمالي

التكرار المتوقع e في حالة الإقتران (كما في المثال) فإنه يعطي بالعلاقة:

$$e = \frac{\sum r . \sum c}{n}$$

حيث أن:

r: مجموع العمود في جدول الإقتران

C: مجموع الصف

n: الإجمالي

ونجد بالنسبة للسيارة المحلية $e = \frac{50 .70}{170} \approx 21$

ونجد بالنسبة للسيارة الأجنبية $e = \frac{70 .120}{170} \approx 49$

وتكون قيم التكرارات المتوقعة الثلاثة الباقية بالطرح من مجموع الأعمدة والصفوف، وعليه يكون جدول التكرارات المتوقعة كمايلي:

الإجمالي	نوع السيارة		السن
	أجنبية	محلية	
70	49	21	تحت 30 سنة
100	71	29	30 سنة أو أكثر
170	120	50	إجمالي

وبالتالي نحصل على الإحصائية كالتالي:

	التكرار الفعلي O	التكرار المتوقع e	O – e	$(O-e)^2$	$(O – e)^2 / e$
	30	21	9	81	3.86
	40	49	-9	81	1.65
	20	29	-9	81	2.79
	80	71	9	81	1.14
المجموع	170	170	Zero		$X^2 = 9.44$

وللمقارنة مع القيمة الجدولية فإن درجة الحرية K تعطى بالعلاقة التالية:

$$K= (r-1) . (c-1) = 1. 1=1$$

وبالمقارنة مع القيمة الجدولية بمعنوية 5% ودرجة K=1 نجد القيمة الجدولية 3.841 نجد أن كاي المحسوبة أكبر من الجدولية لذلك فالقرار نرفض H_0 ونقبل الفرض البديل أي أن السن عامل في تحديد نوع السيارة المشتراة، حيث يميل الكبار إلى السيارات الأجنبية.(يوجد ارتباط بين السن ونوع السيارة).

2- تحليل التباين (التصنيف الأحادي للبيانات):

ذكرنا عند الحديث عن اختيار الفرق بين وسطين حسابيين أن الباحث قد يرغب في اختبار ما إذا كان متوسط الدخل في إحدى الدول يساوي متوسط الدخل في دولة أخرى. أو إجراء اختبار عما إذا كان متوسط عمر العامل في إحدى المؤسسات يساوي متوسط عمر عامل في مؤسسة أخرى. أي أن الباحث قد يرغب في إجراء اختبار ما إذا كان متوسط مجتمع يساوي متوسط مجتمع آخر.

ويكون الفرض العدمي في هذه الحالات هو : $HO \ : \mu_1 \ = \ \mu_2$

ولكن في كثير من الحالات قد يرغب الباحث في إجراء اختبار عن تساوي ثلاث متوسطات أو أكثر. فقد يرغب في اختبار ما إذا كانت متوسطات الدخول في أربع دول متساوية أم لا، أو ما إذا كانت متوسطات أعمار العمال في ست مؤسسات متساوية أم أن هناك اختلافات بينها. ففي المثال الأول يكون الفرض العدمي هو :

$$HO : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

مقابل الفرض البديل بعدم تساوي بعض هذه المتوسطات (اثنان على الأقل غير متساويين) وفي المثال الثاني يكون الفرض العدمي هو :

$$HO \ : \mu_1 \ = \mu_2 \ = \mu_3 \ = \mu_4 \ = \mu_5 \ = \mu_6$$

مقابل الفرض البديل بعدم تساوي بعض هذه المتوسطات (اثنان على الأقل غير متساويين).

في مثل هذه الحالات لا نأخذ كل اثنين من المتوسطات على حدة ونجري اختبار الفرق بين وسطين لهما لأن هذا سوف يستغرق وقتا أطول ومجهودا أكبر، والأهم من ذلك أن احتمال أخذ قرار صحيح سوف يقل أو ينخفض، ويزيد بالتالي كثيرا - احتمال الخطأ أو احتمال اتخاذ قرار غير سليم، ولتوضيح ذلك نأخذ المثال الأول الخاص

بمقارنة المتوسط لأربع دول: فإذا أخذنا كل دولتين على حده فإنه يلزم إجراء هذا الاختبار ست مرات، فإذا كان مستوى المعنوية المستخدم في كل اختبار هو 0.05 (أو أن درجة الثقة هي 0.95)فإن احتمال الوصول إلى القرار الصحيح للاختبارات الستة معاً يساوي $(0.95)^6$ أي 0.95 مضروبة في نفسها ست مرات أي يساوي 0.745 ومعنى ذلك أن احتمال اتخاذ القرار الصحيح سوف ينخفض من 0.95 إلى 0.735 .فقط وبمعنى آخر فإن احتمال الخطأ في اتخاذ القرار الصحيح سوف يرتفع من مجرد 0.05 فقط إلى 0.265 والذي يساوي (1 - 0.735) وهو احتمال كبير للخطأ عند اتخاذ القرار.

وكلما زاد عدد المتوسطات كلما زاد احتمال الخطأ وقل احتمال اتخاذ قرار صحيح ففي المثال الثاني الخاص بمقارنة متوسطات أعمار العمال في ست مناطق فإنه يلزم إجراء الاختبار 15 مرة وبالتالي سوف ينخفض احتمال اتخاذ قرار صحيح في الخمسة عشر اختبار معاً من 0.95 إلى $(0.95)^{15}$ أي إلى 0.46 فقط وبالتالي يرتفع احتمال الخطأ في اتخاذ القرار الصحيح من مجرد 0.05 إلى 0.54 والذي يساوي (1-0.46) وهو احتمال كبير جداً للخطأ في اتخاذ القرار.

لذلك كان لابد من التفكير في أسلوب آخر بديل يوفر الوقت والمجهود وفي الوقت نفسه لا يقلل احتمال اتخاذ القرار الصحيح أو يكبر احتمال الخطأ في اتخاذ القرار، هذا الأسلوب هو الذي يسمى " تحليل التباين" والذي يختبر ما إذا كانت المتوسطات كلها متساوية مرة واحدة دون أخذهم اثنين اثنين ودون أن ينخفض احتمال اتخاذ قرار صحيح أو يزيد احتمال الخطأ عند اتخاذه. وهو الذي يسمى اختصاراً ANOVA أي Analysis of Variance.

ويعتمد هذا الأسلوب من أساليب التحليل الإحصائي على ما يعرف باختبار F (فيشر) والذي يعتمد أساساً على تحليل التباين. ونعلم من الفصول الأولى من الكتاب أن التباين ما هو إلا متوسط مربعات انحرافات القيم عن وسطها الحسابي. أي أن التباين يعتمد أساساً على مجموع مربعات ثم القسمة على عدد المشاهدات. ويعتمد أسلوب تحليل التباين على تقسيم مجموع المربعات الكلي إلى أقسام فيمثل كل منهما أو يقيس أحد مصادر التغير أو الاختلاف Source of Variation يمثل أحدهما مثلاً – التغير بسبب المعاملات (أو المجتمعات) المختلفة ، ويمثل آخر التغير بسبب الأخطاء ثم تعرف الإحصائية (أو الاختبار) F بأنها خارج قسمة التباين بسبب المعاملات على التباين بسبب الأخطاء. وهكذا. أي أنه يتم حساب التباين بسبب المعاملات، والتباين بسبب الأخطاء فيحصل على قيمة F المحسوبة وبمقارنة هذه القيمة بالقيمة الجدولية F نصل إلى قرار إما بقبول الفرض العدمي أو عدم قبوله عند مستوى المعنوية المطلوب. ولتحليل التباين تطبيقات كثيرة في مختلف المجالات :

وسوف نتناول هنا أبسط حالة لتحليل التباين وهي التي تسمى التصنيف الأحادي One - Way Classification مع العلم بأن هناك حالات أخرى كثيرة لتحليل التباين منها على سبيل المثال التصنيف الأحادي في حالة اختلاف أحجام العينات وتحليل التباين الثنائي Two-way Analysis لكننا لن نتعرض في هذا الجزء سوى لأبسط حالة وهي حالة التصنيف الأحادي بافتراض تساوي أحجام العينات.

يعتبر التصنيف الأحادي أبسط أنواع تحليل التباين، حيث تصنف المشاهدات إلى عدة مجموعات على أساس متغير واحد أو خاصية واحدة.

والافتراضات الأساسية لهذا التحليل هي ما يلي :

1- نفترض أن عدد المجتمعات K وكلها جميعاً مستقلة.

2- كلها جميعاً تخضع للتوزيع الطبيعي بمتوسطات تساوي $\mu_k ,..., \mu_2 , \mu_1$

3- لها جميعاً التباين نفسه σ^2 أي أن التباين لكل المجتمعات ثابت ويساوي σ^2

4- يتم سحب عينة عشوائية من كل من هذه المجتمعات وأن أحجام هذه العينات كلها متساوية وتساوي n ويمكن بكل بساطة افتراض عدم تساوي أحجام العينات ولن يختلف أسلوب التحليل على الإطلاق إلا في أشياء بسيطة جداً .

وتصنف البيانات عادة في هذا التحليل على النحو التالي:

المجتمع أو المعاملة	المشاهدات		
	1	**2**	**n**
1	Y11	y12	y1n
2	Y21	y22	y2n
:	:		
k	Yk1	yk2	ykn

حيث يمثل الصف الأول مشاهدات العينة الأولى أي المسحوبة من المجتمع الأول، ويمثل الصف الثاني مشاهدات العينة المسحوبة من المجتمع الثاني،.. وهكذا يمثل الصف الأخير مشاهدات العينة المسحوبة من المجتمع الأخير رقم K.

كما تكون خطوات الاختبار كما يلي:

1- الفرض العدمي: هو أن متوسطات هذه المجتمعات متساوية وبالرموز $HO : \mu_1 = \mu_2 = ... = \mu_k$

2- الفرض البديل : هو أن بعض هذه المتوسطات غير متساوية (أو: يوجد متوسطان على الأقل غير متساويين).

3- إحصائية الاختبار : في هذه الحالة يرمز لها بالرمز F وتأخذ الشكل التالي: $F = \frac{S_R^2}{S_E^2}$

والتي لها توزيع F بدرجات حرية للبسط K-1 وللمقام K. (n-1) حيث : S_R^2 هو التباين بسبب الظاهرة أو المتغير أو المعاملات، S_E^2 هو التباين بسبب الخطأ.

ويمكن الحصول على الإحصائية F بتنظيم الحسابات في جدول يسمى " جدول تحليل التباين " AVOVA TABLE كما يلي :

جدول رقم 06: جدول تحليل التباين (ANOVA)

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	(الإحصائية) المحسوبة
بسبب المعاملات	SSR	K- 1	$S_R^2 = SSR/K-1$	$F = \frac{S_R^2}{S_E^2}$
بسبب الخطأ	SSE	K (n-1)	$S_E^2 = SSE/K(n-1)$	
الكلي	SST	Nk-1		

وسوف نوضح من المثال كيفية حساب المقادير الثلاثة SSE, SSR , SST

5- حدود منطقتي القبول والرفض : ويتم الحصول عليها من جدول توزيع F بدرجات حرية للبسط K-1 وللمقام

k(n- 1)). (اختبار الطرف الأيمن).

5- المقارنة والقرار: إذا كانت قيمة F المحسوبة من تحليل التباين أقل من قيمة F الجدولية نقبل الفرض العدمي بتساوي المتوسطات والعكس صحيح.

مثال 1:

البيانات التالية تمثل دخول أربع عينات عشوائية من العمال سحبت من أربع مدن مستقلة (نفترض أن لها توزيعات طبيعية بمتوسطات $\mu 4, \mu 3, \mu 2, \mu 1$ وتباين مشترك يساوي σ^2):

المدن k	المشاهدات						المجموع
	1	2	3	4	5	6	
الأولى 1	20	21	25	28	30	26	150
الثانية 2	23	22	27	20	26	20	138

الثالثة 3	19	20	21	28	20	18	126
الرابعة 4	24	29	30	28	27	24	162

والمطلوب اختبار الفرض العدمي بأن متوسطات دخول العمال من المدن الأربع متساوية، أي أن المطلوب بالرموز هو :

$$HO: \mu1 = \mu2 = \mu3 = \mu4$$

وذلك بمستوى معنوية 5%

الحل :

تكون خطوات الحل كما يلي :

1- الفرض العدمي :

$$HO: \mu1 = \mu2 = \mu3 = \mu4$$

2الفرض البديل : أن بعض هذه المتوسطات غير متساوٍ (اثنان على الأقل غير متساويين).

3- الإحصائية : وهي في هذه الحالة $F = \frac{S_R^2}{S_E^2}$

وتكون الحسابات التفصيلية لتحليل التباين كما يلي :

حيث K= 4 n= 6

أ) متوسطات الصفوف (المدن) :

متوسط الصف الأول : 150/6 = 25

متوسط الصف الثاني : 138/ 6 = 23

متوسط الصف الثالث : 126/ 6 = 21

متوسط الصف الرابع : 162/ 6 = 27

ب) المتوسط الكلي : $\frac{150+138+126+162}{24} = \frac{576}{24} = 24$

ج) مجموع المربعات الكلي:

$$SST = (20^2 + 21^2 + 25^2 + 28^2 + 30^2 + 26^2 + 23^2 + \ldots + 27^2 + 24^2) - 6x4\,(24)^2$$

$$= 14160 - 6x4\,x24x24$$

$$= 14160 - 13824$$

$$SST = 336$$

د) مجموع المربعات للصفوف (المدن) :

$$SSR = 6(25^2 + 23^2 + 21^2 + 27^2) - 6x4x(24)^2$$
$$= 13944 - 13824$$
$$SSR = 120$$

هـ) مجموع مربعات الخطأ :

$$SSE = SST - SSR$$
$$= 336 - 120$$
$$SSE = 216$$

و) ثم نكون جدول تحليل التباين كما يلي :

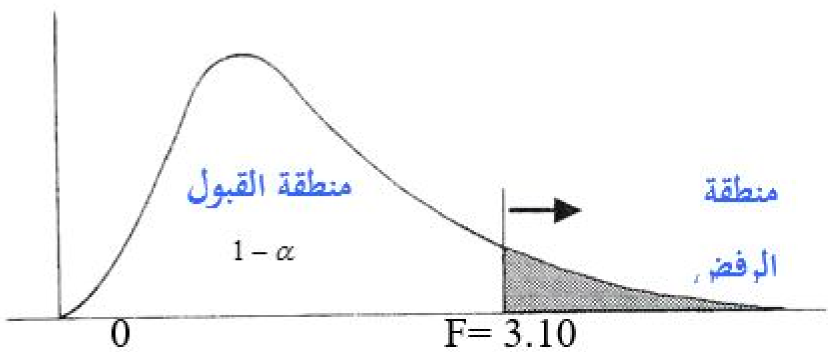
مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	الإحصائية F
بسبب المعاملات	SSR= 120	3	SR2 = 120/ 3 = 40	$F = \frac{40}{10.8} = 3.7$
بسبب الخطأ	SSE= 216	20	SE2= 216/20= 10.8	
الكلي	SST= 336	23		

أي أن قيمة الإحصائية (أو F المحسوبة) هي 3.7

$$F = \frac{40}{10.8} = 3.7$$

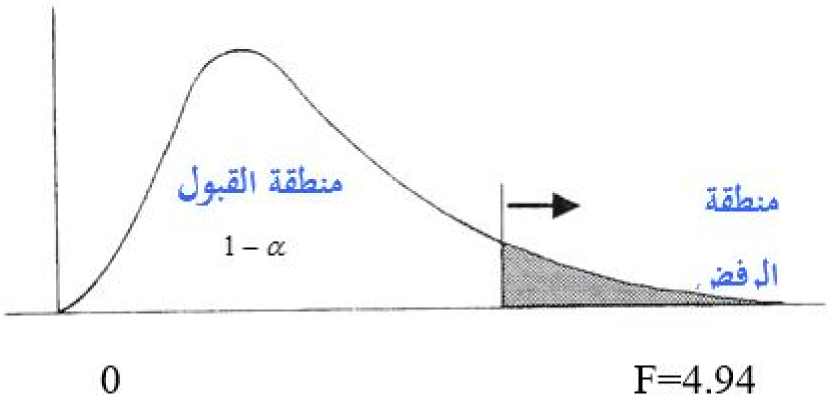
4- حدود منطقتي القبول والرفض : من جدول توزيع F وعند مستوى معنوية 5% وبدرجات حرية 3 للبسط، 20 للمقام نجد أن F الجدولية تساوي 3.10

ويمكن توضيح ذلك بالرسم كما يلي :



5- المقارنة والقرار: وحيث أن قيمة الإحصائية المسحوبة والتي تساوي 3.7 أكبر من القيمة الجدولية فإنها تقع في منطقة الرفض وبالتالي فإن القرار هو رفض الفرض العدمي بتساوي متوسطات دخول العمال في المدن الأربع وذلك بمستوى معنوية 5%

أما إذا استخدمنا مستوى معنوية 1% فإن قيمة F من الجدول تصبح 4.94 أي تصبح حدود منطقتي القبول والرفض هي



ملاحظة هامة جدا: في حالة عدم تساوي أحجام العينات يجب مراعاة حجم العينة لكل مجتمع وهو ما يمكن استكشافه في المثال التالي:

مثال 2:

البيانات التالية تخص عينات عشوائية مستقلة لعبوة علب عصير بالميلمتر مكعب معبئة بواسطة أربع آلات مستخدمة في مصنع للمشروبات:

الآلة k	المشاهدات						المجموع
	1	2	3	4	5	6	
الأولى 1	201,08	200,12	201,1	199,89	199,91	202,1	1204
الثانية 2	198,92	199,75	199,05				597,7
الثالثة 3	200,01	200,2	199,84	199,98	200,03		1000.1
الرابعة 4	200,3	201,01	201,22	199,99			802,5

الحل:

تكون خطوات الحل كما يلي :

1- الفرض العدمي :

$$HO: \mu 1 = \mu 2 = \mu 3 = \mu 4$$

2الفرض البديل : أن بعض هذه المتوسطات غير متساو ٍ (اثنان على الأقل غير متساويين).

3- الإحصائية : وهي في هذه الحالة $F = \dfrac{S_R^2}{S_E^2}$

وتكون الحسابات التفصيلية لتحليل التباين كما يلي :

حيث K= 4

أ) متوسطات الصفوف (المدن) :

متوسط الصف الأول : $1204/6 = 200.7$

متوسط الصف الثاني : $597.7/ 3 = 199.24$

متوسط الصف الثالث : $1000.1/ 5 = 200.012$

متوسط الصف الرابع : $802.5/ 4 = 200.63$

ب) المتوسط الكلي : $\frac{1204+597.7+1000.1+802.5}{18} = 200.25$

ج) مجموع المربعات الكلي:

$$SST = (201.08^2 + \ldots\ldots + 199.99^2) - 18\,(200.25)^2$$

$$SST = 10.493$$

د) مجموع المربعات للصفوف (المدن) :

$$SSR = 6(200.7-200.25)^2+3(199.24-200.25)^2+5(200.12-200.25)^2+4(200.63-200.25)^2=5.136$$

ه) مجموع مربعات الخطأ :

$$SSE = SST - SSR$$

$$= 10.493 - 5.136$$

$$SSE = 5.357$$

و) ثم نكون جدول تحليل التباين كما يلي :

الإحصائية F	متوسط المربعات	درجات الحرية	مجموع المربعات	مصدر التغير
$F = 4.5$	$SR^2 = 1.71$	3	SSR= 5.136	بسبب المعاملات
	$SE^2 = 0.38$	14	SSE= 5.357	بسبب الخطأ
		23	SST= 10.493	الكلي

4- حدود منطقتي القبول والرفض : من جدول توزيع F وعند مستوى معنوية 5% وبدرجات حرية 3 للبسط، 14 للمقام نجد أن F الجدولية تساوي 3.34

5- المقارنة والقرار: وحيث أن قيمة الإحصائية المسحوبة والتي تساوي 4,5 أكبر من القيمة الجدولية فإنها تقع في منطقة الرفض وبالتالي فإن القرار هو رفض الفرض العدمي بتساوي متوسطات دخول العمال في المدن الأربع وذلك بمستوى معنوية 5% .

السلسلة رقم: 02-

التمرين الأول: (اختبار التجانس)

كان متوسط درجات الحرارة المسجلة من قبل الدوائر الجوية في 6 عواصم عربية خلال شهر جويلية كما يلي :

العواصم	بغداد	الجزائر	دمشق	طرابلس	بيروت	القاهرة	المجموع
درجات الحرارة	43	35	32	39	32	41	222

هل يمكن الإستدلال بأن درجات الحرارة في العواصم الست كانت متجانسة خلال شهر جويلية بمستوى معنوية 5%، أو بتعبير آخر: اختبر ما إذا كان هناك تجانس (أو تماثل) في درجات الحرارة بالنسبة للعواصم.

التمرين الثاني: (اختبار الإستقلال)

بغرض اختبار ما إذا كانت علامة السيارة المشتراة (تويوتا، رونو، بيجو) مستقبلا يعتمد على دخل المشتري، جمع تاجر سيارات البيانات الموضحة بالجدول أدناه عن عدد السيارات المشتراة من كل علامة لعملاء دخلهم الشهري أقل من 40000 دج، والتي يشتريها عملاء دخلهم الشهري بين 40000-60000 دج، والمشتراة من عملاء دخلهم الشهري أكبر من 60000 دج.

الإجمالي	علامة السيارة			الدخل الشهري
	بيجو	رونو	تويوتا	
80	20	50	10	أقل من 40000 دج
100	30	40	30	بين 40000-60000 دج
150	40	50	60	أكبر من 60000 دج
330	90	140	100	الإجمالي

اختبر ما إذا كان للدخل الشهري دور في تحديد علامة السيارة المشتراة وذلك باستخدام معنوية 5%؟.

التمرين الثالث: (تحليل التباين)

لدراسة الاستقرار الإقتصادي لأربع دول نفطية وبالأخذ بمؤشر نصيب الفرد من الناتج الداخلي الخام، أخذت بيانات 6 سنوات كما هو مبين في الجدول الموالي، علما أن البيانات موزعة توزيعا طبيعيا ولها تباين متساو.

البلدان	السنوات						الإجمالي
	2013	2014	2015	2016	2017	2018	
الجزائر	8	14	13	11	10	10	66
العراق	16	11	11	13	9	12	72
ليبيا	15	12	9	9	11	10	66
إيران	13	10	11	13	15	16	78

اختبر إذا كان نصيب الفرد من الناتج الداخلي الخام متساويا بين الدول أم لا عند مستوى معنوية $\alpha = 0.05$.

الحل :

✓ **التمرين الأول:**

1 – الفرض العدمي : التجانس بيندرجات الحر ارة.

2 – الفرض البديل : عدم التجانس بين درجات الحرارة.

3 – الإحصائية : وهي : $X^2 = \sum_{I}^{K} \frac{(O-e)^2}{e}$

ونحصل عليها كما في الجدول التالي :

العواصم	درجات الحرارة O	$e = \frac{222}{6}$	O – e	$(O-e)^2$	$(O-e)^2$ / e
بغداد	43	37	6	36	0,972973
الجزائر	35	37	-2	4	0,108108
دمشق	32	37	-5	25	0,675676
طرابلس	39	37	2	4	0,108108
بيروت	32	37	-5	25	0,675676
القاهرة	41	37	4	16	0,432432
المجموع	222	222	Zero		$X^2 = 2.97$

4 – حدود منطقتي القبول والرفض : اختبار الطرف الأيمن لتوزيع X^2 حيث مســتوى المعنوية هو $\propto = 0.05$ ودرجات حرية تساوي K – 1 وحيث أن عدد العواصم K = 6 فإن درجات الحرية هي K –1 أي تساوي 5 ومن جدول X^2 نجد أن :

$X^2_{0.05,5} = 11.07$

5 – المقارنة والقرار : وحيث أن قيمة الإحصـــائية أقل من المجدولة أي تقع في منطقة القبول فإن القرار **هو :**

قبول الفرض العدمي (أي رفض فرض التجانس بين درجات الحرارة) وذلك بمستوى معنوية 5%.

✓ **التمرين الثاني:**

1 – الفرض العدمي : الدخل ليس عاملا لتحديد علامة السيارة المشتراة.

2 – الفرض البديل : الدخل عامل لتحديد علامة السيارة المشتراة.

3 – الإحصائية : وهي : $X^2 = \sum_{I}^{K} \frac{(O-e)^2}{e}$

وتكون قيم التكرارات المتوقعة كما يلي:

الإجمالي	علامة السيارة			الدخل الشهري
	بيجو	رونو	تويوتا	
80	22	34	24	أقل من 40000 دج
100	27	42	31	بين 40000-60000 دج
150	41	64	45	أكبر من 60000 دج
330	90	140	100	الإجمالي

وبالتالي نحصل على الإحصائية كالتالي:

$(O-e)^2 / e$	$(O-e)^2$	$O-e$	التكرار المتوقع e	التكرار الفعلي O	
19,6	196	14	24	10	
0,033333	1	1	31	30	
3,75	225	-15	45	60	
5,12	256	-16	34	50	
0,1	4	2	42	40	
3,92	196	14	64	50	
0,2	4	2	22	20	
0,3	9	-3	27	30	
0,025	1	1	41	40	
$X^2 = 33.05$		Zero	330	330	المجموع

وللمقارنة مع القيمة الجدولية فإن درجة الحرية K تعطى بالعلاقة التالية:

$$K= (r-1) . (c-1) = 2. 2=4$$

وبالمقارنة مع القيمة الجدولية بمعنوية 5% ودرجة K=4 نجد القيمة الجدولية 9.488 نجد أن كاي المحسوبة أكبر من الجدولية لذلك فالقرار نرفض H_0 ونقبل الفرض البديل أي أن الدخل عامل في تحديد علامة السيارة المشتراة بمعنوية 5%.

✓ **التمرين الثالث:**

تكون خطوات الحل كما يلي :

1- الفرض العدمي :

$$HO: \mu 1 = \mu 2 = \mu 3 = \mu 4$$

2- الفرض البديل : أن بعض هذه المتوسطات غير متساو ٍ (اثنان على الأقل غير متساويين).

3- الإحصائية : وهي في هذه الحالة $F = \frac{S_R^2}{S_E^2}$

وتكون الحسابات التفصيلية لتحليل التباين كما يلي :

حيث K= 4

أ) متوسطات الصفوف (البلدان) :

متوسط الصف الأول : $66/6 = 11$

متوسط الصف الثاني : $72/6 = 12$

متوسط الصف الثالث : $66/6 = 11$

متوسط الصف الرابع : $78/6 = 13$

ب) المتوسط الكلي : $\frac{66+72+66+78}{24} = 11.75$

ج) مجموع المربعات الكلي:

$$SST = (8 + \ldots\ldots + 10)^2 - 24\,(11.75)^2 = 120.5$$

$$SST = 120.5$$

د) مجموع المربعات للصفوف (المدن) :

$$SSR = 6(11-11.75)^2+6(12-11.75)^2+6(11-11.75)^2+6(13-11.75)^2=16.5$$

ه) مجموع مربعات الخطأ :

$$SSE = SST - SSR$$
$$= 120.5 - 16.5$$
$$SSE = 104$$

و) ثم نكون جدول تحليل التباين كما يلي :

الإحصائية F	متوسط المربعات	درجات الحرية	مجموع المربعات	مصدر التغير
$F = 1.06$	$S_R^2 = 5.5$	3	SSR= 16.5	بسبب المعاملات
	$S_E^2 = 5.2$	20	SSE= 104	بسبب الخطأ
		23	SST= 120.5	الكلي

4- حدود منطقتي القبول والرفض مع العلم أن: من جدول توزيع F وعند مستوى معنوية 5% وبدرجات حرية 3 للبسط، 20 للمقام نجد أن F الجدولية تساوي 3.10.

5- المقارنة والقرار : وحيث أن قيمة الإحصائية المسحوبة أقل من القيمة الجدولية فإنها تقع في منطقة القبول وبالتالي فإن القرار هو قبول الفرض العدمي بتساوي المتوسطات عند مستوى معنوية 5% .

المحور الرابع:

تحليل الإرتباط

الأهداف التعليمية

- **التعرف على أنواع العلاقة بين المتغيرين؛**
- **شكل الانتشار** Scatter Diagram ؛
- **معامل بيرسون للارتباط الخطي البسيط واختبار معنويته ؛**
- معامل ارتباط الرتب Rank Correlation واختبار معنويته؛

تمهيد :

تناولنا في الفصول السابقة دراسة متغير واحد أو ظاهرة واحدة من حيث قياس وحساب متوسط هذه الظاهرة وكذلك حساب مقياس لتشتها. ولكن في الحياة العملية كثيرا ما يحتاج الباحث لدراسة العلاقة بين ظاهرتين (أو متغيرين) لمعرفة مدى الارتباط بينهما ونوع هذا الارتباط. فقد يريد الباحث معرفة ما إذا كان هناك علاقة بين التدخين والإصابة بمرض في الرئة، أو بين درجة تعليم الشخص ومستوى دخله. أو بين الحالة التعليمية والحالة الاجتماعية للناخب. وكما نرى فإنه يمكن أن نذكر الكثير بين الأمثلة في مختلف المجالات بل قد يرغب الباحث في دراسة العلاقة بين أكثر من متغيرين في وقت واحد. فمثلا قد يريد الباحث معرفة تأثير درجة التعليم ومستوى الدخل وحجم الأسرة على درجة الوعي السياسي للشخص. في هذا المثال يريد الباحث معرفة تأثير ثلاثة متغيرات على متغير رابع. وفي هذا الكتاب سوف نركز على دراسة العلاقة بين **متغيرين اثنين فقط** وهو ما يعرف **بالارتباط " البسيط "** Simple Correlation. بينما الحالات التي تتناول الدراسة فيها **أكثر من متغيرين** تعرف **بالارتباط المتعدد** Multiple Correlation وهي – كما ذكرنا – خارج نطاق هذا الكتاب.

1-أنواع العلاقة بين المتغيرين:

إذا كان المتغيران يتغيران معا في الاتجاه نفسه بمعنى أنه إذا زاد أو نقص أحدهما، زاد أو نقص الأخر، فإن العلاقة بينهما تكون طردية والارتباط بينهما يكون موجبا. مثال ذلك العلاقة بين زيادة حجم الطبقة الوسطى في المجتمع وزيادة الاستقرار السياسي.

وإذا كان المتغيران يتغيران معا ولكن في عكس الاتجاه بمعنى أنه إذا زاد أحدهما نقص الأخر، أو إذا نقص أحدهما زاد الأخر، فإن العلاقة بينهما تكون عكسية والارتباط بينهما يكون سالبا. مثال العلاقة بين تدني مستوى الفرد التعليمي ودرجة الوعي الاجتماعي.

وتختلف العلاقات بين الظواهر من حيث القوة. فقد تكون العلاقة قوية جدا (أو حتى تامة)، وقد تكون متوسطة، أو ضعيفة أو منعدمة تماما. وفي هذا الفصل سوف نتناول بالتفصيل كيفية حساب الارتباط بين متغيرين سواء كان المتغيران كميين أو وصفيين (ترتيبيين أو اسميين)، أو أحدهما كميا والأخر وصفيا.

2-شكل الانتشار Scatter Diagram:

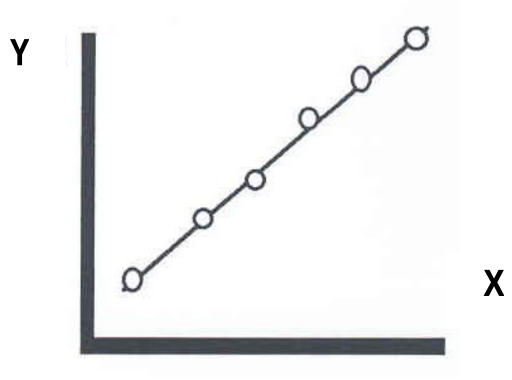
هناك وسيلة مبدئية يعرف الباحث من خلالها نوع الارتباط بين المتغيرين وما إذا كان الارتباط قويا وضعيفا أو منعدما، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة. هذه الوسيلة هي " شكل الانتشار " والتي تصلح إذا كان المتغيران كميين. وجدير بالذكر أن هذه وسيلة مبدئية تساعد فقط في معرفة نوع الارتباط ولا تعتبر بديلا عن الطرق الإحصائية التي سوف نتناولها بالتفصيل في هذا الفصل.

والمقصـــود بشكل الانتشـــار هو تمثيل قيم الظاهرتين بيانيا على المحورين، المتغير الأول X على المحور الأفقي، والمتغير الثاني Y على المحور الرأســي، حيث يتم تمثيل كل زوج Pair من القيم بنقطة، فنحصـــل على شـكل يمثل كيفية انتشـار القيم على المسـتوى، وهو الذي يسـمى شـكل الانتشـار. وطريقة انتشـار القيم تدل على وجود أو عدم وجود علاقة بين المتغيرين ومدى قوتها ونوعها. فإذا كانت تتوزع بشـــكل منتظم دل ذلك على وجود علاقة (يمكن اســـتنتاجها)، أما إذا كانت النقط مبعثرة ولا تنتشـــر حســـب نظام معين دل ذلك على عدم وجود علاقة بين المتغيرين أو أن العلاقة بينهما ضعيفة. والأشكال التالية تظهر بعض أشكال الانتشار المعروفة :

الشكل الأول :

إذا وقعت جميع النقاط على خط مستقيم، دل ذلك على أن العلاقة بينهما خطية وأنها ثابتة أو تامة. وهذه تمثل أقوى أنواع الارتباط بين المتغيرين " ارتباط تام ". فإذا كانت العلاقة طردية فإن " الارتباط طردي تام " كما في الشكل الأول (أ). ومثاله العلاقة بين الكمية المشتراة من سلعة والمبلغ المدفوع لشراء هذه الكمية.

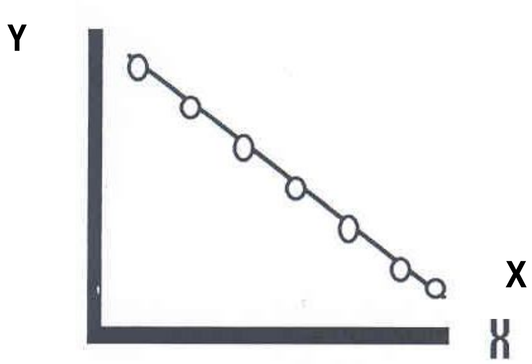
أما إذا كانت العلاقة عكسية (وجميع النقاط تقع على خط مستقيم واحد) فإن " الارتباط عكسي تام " كما في الشكل الأول (ب). ومثال على ذلك العلاقة بين السرعة والزمن.



الشكل الأول (أ)

ارتباط طردي تام

(موجب)



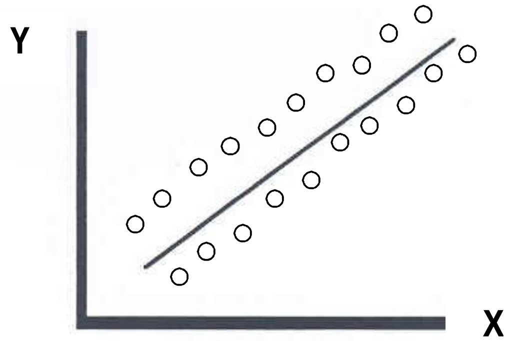
الشكل الأول (ب)

ارتباط عكسي تام

(سالب)

الشكل الثاني :

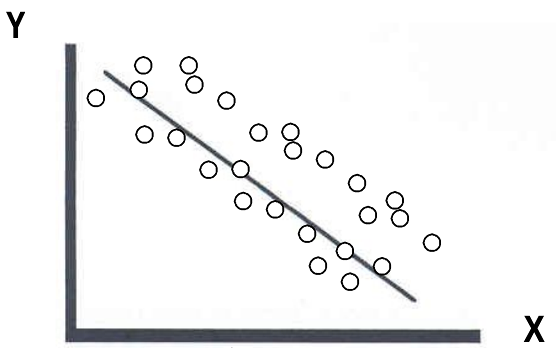
أما إذا كانت النقاط تأخذ شـــكل خط مســـتقيم ولكن لا تقع جميعها على الخط قيل أن العلاقة خطية (موجبة أو سالبة) كما في الشكل الثاني أ، ب.



الشكل الثاني (أ)

ارتباط موجب قوي

(ارتباط خطي طردي)



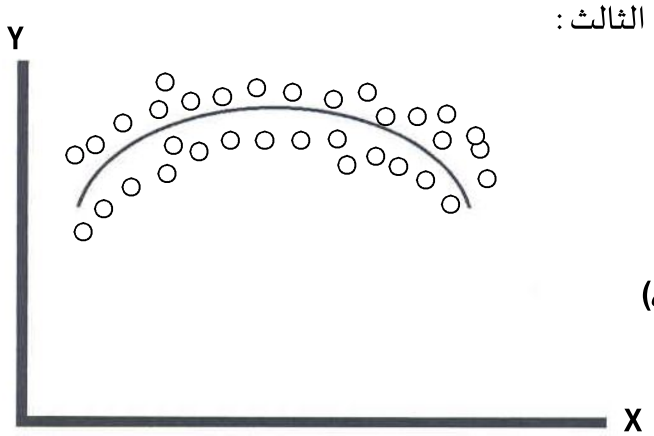
الشكل الثاني (ب)

ارتباط سالب قوي

(ارتباط خطي عكسي)

الشكل الثالث :

وإذا كانت العلاقة تأخذ شـــكل منحنى فإن الارتباط لا يكون خطيا ، "ارتباط غير خطي" Non Linear Correlation كما في الشكل الثالث :

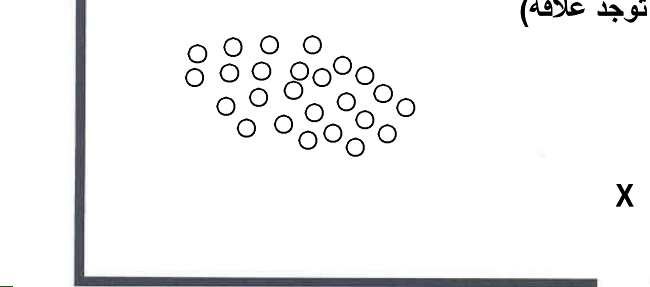


الشكل الثالث

(ارتباط غير خطي)

الشكل الرابع :

أما إذا كانت النقاط تتبعثر بدون نظام معين فإن ذلك يدل على عدم وجود علاقة بين المتغيرين (أو أن العلاقة بينهما ضعيفة جدا ،) كالعلاقة مثلا ، بين دخل الشخص وطوله كما في الشكل الرابع :



الشكل الرابع

(لا توجد علاقة)

3-معامل الارتباط Correlation Coefficient:

يقاس الارتباط بين متغرين بمقياس إحصائي يسمى " معامل الارتباط " ويعكس هذا المقياس درجة أو قوة العلاقة بين المتغيرين واتجاه هذه العلاقة. وتنحصر قيمة معامل الارتباط بين + 1، - 1. فإذا كانت قيمة معامل الارتباط تساوي + 1 فمعنى ذلك أن الارتباط بين المتغيرين طردي تام، وهو أقوى أنواع الارتباط الطردي بين متغيرين. وإذا كانت قيمة معامل الارتباط تساوي – 1 فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام، وهو أقوى أنواع الارتباط العكسي بين متغيرين. وإذا كانت قيمة معامل الارتباط تساوي صفر، فمعنى ذلك أنه لا يوجد ارتباط بين المتغيرين. وكلما اقتربت قيمة معامل الارتباط من + 1 أو – 1 كلما كان الارتباط قويا ً، وكلما اقترب من الصفر كلما كان الارتباط ضعيفا ً.

والخلاصة : أنه كلما كانت العلاقة قوية بين المتغيرين كلما اقترب معامل الارتباط من + 1 أو – 1 فإذا وصلت قيمة المعامل إلى + 1 أو – 1 كان الارتباط تاما ً بين المتغيرين. وأنه كلما كانت العلاقة ضعيفة بين المتغيرين كلما اقترب معامل الارتباط من الصفر فإذا وصلت قيمة المعامل إلى الصفر كان الارتباط منعدما ً بين المتغيرين. ومعنى ذلك أيضا ً أنه لا يوجد ارتباط بين متغيرين تكون قيمة المعامل فيه أكبر من + 1 ولا أصغر من – 1.

وسنبدأ بقياس العلاقة بين متغيرين كميين، ثم متغيرين ترتيبين، وأخيرا ً متغيرين أسميين.

3-1-معامل بيرسون للارتباط الخطي البسيط :

عند دراسة العلاقة بين متغيرين نجد مجموعة من الأسئلة التي وجب الإجابة عليها: هل يرتبط هاذان المتغيران بعلاقة طردية أو عكسية؟، أو بمعنى آخر هل ترتبط القيم الكبيرة لأحد المتغيرين بالقيم الأكبر للمتغير الآخر؟، أو أن القيم الأكبر للمتغير الأول عادة ما تكون مصاحبة للقيم الأصغر للمتغير الثاني؟، أو أنه لا يوجد ارتباط بين القيم للمتغيرين بمعنى أنه في حالات تصاحب القيم الكبيرة للمتغير الأول القيم الكبيرة للمتغير الثاني وفي حالات أخرى تصاحب القيم الكبيرة نظيرتها من القيم الصغيرة؟.

إن المعلمة التي تجيب على الأسئلة السابقة تسمى "التباين المشترك" أو "التغاير"، وبالنسبة لمتغير ين X و Y وسطيهما الحسابيين $E(x)$ و $E(y)$ فيعرف التغاير بالعلاقة الرياضية التالية:

$$cov(x,y) = E[(x_i - \bar{x})\ (y_i - \bar{y})]$$

وبهذا المعنى فالتباين المشترك (أو التغاير) هو القيمة المتوقعة لحاصل ضرب $(x_i - \bar{x})$ في $(y_i - \bar{y})$، فإذا كان هذا التغاير موجبا $cov(x,y)\ > 0$ فمعناه أن القيم الأكبر لـ $(x_i - \bar{x})$ تقابلها في العادة (ليس بالضرورة كل القيم) القيم الأكبر من $(y_i - \bar{y})$ والعكس صحيح، وبديهيا فإن الفكرة وببساطة أن القيمة الموجبة لـ $cov(x,y)\ > 0$ تستوجب أن يكون $(y_i - \bar{y})$ و $(x_i - \bar{x})$ موجبين معا أو سالبين معا، ولهذا نقول إذا كان X أكبر من متوسطه فإن فإن y لا بد أون يكون كذلك أكبر من متوسطه

لتحقق العلاقة الطردية وكذلك الحال في إذا كانا سالبين معا. وفي حالة كانت القيم الموجبة لـ $(x_i - \bar{x})$ تقابلها القيم السالبة لـ $(y_i - \bar{y})$ فإن العلاقة تكون عكسية.

في الحالة الوسطية التي يكون فيها $cov(x,y) = 0$ أو تقترب من 0 تكون قيم X الأكبر من متوسطها تقابلها قيم y الأصغر والأكبر من متوسطها، حيث تكون هذه الوضعية في حال كان المتغيرين مستقلين أو أنه تربط بينهما علاقة غير خطية.

مثال1

لتكن البيانات التالية والخاصة بالعلاقة بين القمح المنتج (Y) وكمية السماد المستعملة (X).

X	6	10	12	14	16	18	22	24	26	32
Y	40	44	46	48	52	58	60	68	74	80

والمطلوب حساب التباين المشترك (التغاير) للمتغيرين؟

لإحتساب التغاير للمتغيرين X وy نستعين بالجدول الموالي والمبين للحسابات التفصيلية:

X	Y	$(y_i - \bar{y})$	$(x_i - \bar{x})$	$(x_i - \bar{x})\,(y_i - \bar{y})$
6	40	-17	-12	204
10	44	-13	-8	104
12	46	-11	-6	66
14	48	-9	-4	36
16	52	-5	-2	10
18	58	1	0	0
22	60	3	4	12
24	68	11	6	66
26	74	17	8	136
32	80	23	14	322
Σ				956

حيث أن: $\bar{x} = 18$ و $\bar{y} = 57$

وبالرجوع إلى العلاقة:

$$cov(x,y) = E[(x_i - \bar{x})\,(y_i - \bar{y})]$$

نجد أن:

$$cov(x,y) = \frac{[(x_i - \bar{x})\,(y_i - \bar{y})]}{n}$$

$$cov(x,y) = \frac{956}{10} = 95.6$$

بالنظر إلى جدول الحسابات التفصيلية يمكن ملاحظة أن القيم السالبة من X تقابلها القيم السالبة من Y والعكس صحيح مما يوحي بعلاقة طردية بين المتغيرين، كما أن حجم القيم بالقيمة المطلقة (من الأكبر للأصغر) للمتغير X تقابلها بنفس التدرج تقريبا القيم الطلقة للمتغير Y مما يوحي بعلاقة قوية.

إن قيمة معامل التغاير والمقدرة بـ 95.6 لا يمكن من خلالها تحديد على وجه الدقة حجم العلاقة بين المتغيرين، وهو ما يستوجب التحول نحو القيم المعيارية.

إن تحويل قيمة التغاير إلى قيمة معيارية يعبر في حقيقة الأمر عن معامل الإرتباط لبيرسون، فاختلاف قيم التغاير باختلاف قيم المتغيرات يجعل من الصعوبة تحديد حجم العلاقات الرابطة بين المتغيرات مما يستوجب التحويل للقيم المعيارية أو ما يسمى بمعاملات الإرتباط.

بالرجوع إلى المثال السابق يمكن التحول من تحويل قيمة التغاير إلى قيمة معيارية كما يلي:

$$r_{x.y} = \frac{\sum[(x_i-\bar{x})\,(y_i-\bar{y})]/n}{\sqrt{\sum(x_i-\bar{x})^2/n}\cdot\sqrt{\sum(y_i-\bar{y})^2/n}}$$

وحيث أن: $$\delta_x = \sqrt{\sum(x_i-\bar{x})^2/n}$$

و $$\delta_y = \sqrt{\sum(y_i-\bar{y})^2/n}$$

والتي تمثل الإنحرافات المعيارية لكل من X وy.

وعليه يمكن الكتابة:

$$r_{x.y} = \frac{cov(x,y)}{\delta_x.\delta_y}$$

وهي جوهر التحول من التغاير إلى الإرتباط الخطي أو بتعبير آخر التحويل المعياري للتغاير.

وبالرجوع إلى المثال السابق نجد أن معامل الإرتباط:

$$r_{x.y} = \frac{95.6}{12.78\ .\ 7.59} = 0.985$$

وهو ما يبين العلاقة القوية جدا بين المتغير X وy والتي تقدر بـ 98.5%.

مثال 2: (صيغة ثانية لمعامل بيرسون)

نحاول استخدام الصيغة $$r = \frac{n(\sum xy)-(\sum x)(\sum y)}{\sqrt{n\sum x^2-(\sum x)^2}\ \sqrt{n\sum y^2-(\sum y)^2}}$$

البيانات التالية تمثل أعمار ثمانية من العمال ودخولهم اليومية بالدولار، والمطلوب حساب معامل بيرسون للارتباط الخطي بين الأعمار والدخول.

الأعمار x : 25 32 29 43 38 51 47 35

الدخول y : 10 18 15 35 40 62 100 50

الحل :

لحساب معامل بيرسون للارتباط الخطي يلزم حساب المجاميع:

$\sum x, \sum y, \sum xy, \sum x^2, \sum y^2$ لذلك يتم تنظيم حساب هذه المجاميع كما في الجدول التالي:

الأعمار x	الدخول y	xy	x^2	y^2
25	10	250	625	100
32	18	576	1024	324
29	15	435	841	225
43	35	1505	1849	1225
38	40	1520	1444	1600
51	62	3162	2601	3844
47	100	4700	2209	10000
35	50	1750	1225	2500
300	330	13898	11818	19818

ثم نطبق في الصيغة المختصرة رقم (2) لمعامل الارتباط حيث n = 8 :

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

$$= \frac{8(13898) - (300)(330)}{\sqrt{8(11818) - (300)^2}\sqrt{8(19818) - (330)^2}}$$

$$= \frac{111184 - 99000}{\sqrt{94544 - 9000}\sqrt{158544 - 108900}}$$

$$= \frac{12184}{\sqrt{4544}\sqrt{49644}}$$

$$= \frac{12184}{15019.6}$$

$$r = 0.81$$

أي أن معامل بيرســون للارتباط الخطي بين أعمار العمال ودخولهم اليومية يســاوي 0.81 وهو ارتباط طردي (لأن إشارته موجبة) وقوى (لأنه قريب من الواحد الصحيح). بمعنى أخر، إن هناك علاقة طردية قوية بين عمر العامل ودخله مقدارها 81 %. فمع زيادة عمر العامل يزيد دخله، والعكس صحيح (الخبرة).

ملاحظة مهمة :

من خواص معامل بيرسون للارتباط الخطي أنه لا يتأثر بالعمليات الحسابية التي تجري على المتغيرين y , x. بمعنى أنه لا يتأثر بالطرح (أو الجمع)، ولا بالقسمة (أو الضرب). أي إذا طرحنا (أو جمعنا) قيمة معينة من كل قيم x وقيمة أخرى من كل قيم y، أو قسمنا (أو ضربنا) قيم x على قيمة معينة وكل قيم y على قيمة أخرى فإن قيمة معامل الارتباط لا تتغير أي نحصل على القيمة نفسها.

اختبار معنوية الارتباط Significance Of Correlation Coefficient :

إذا كانت قيمة معامل ارتباط العينة r قريبة من + 1 أو – 1 فإن هناك علاقة خطية قوية بين المتغيرين، وإذا كانت r = 0 فإنه لا توجد علاقة خطية بينهما، أما إذا كانت قيم r متوســطة 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. وسوف نميز بين حالتين:

الأولى: إذا كان المطلوب هو اختبار ما إذا كان هناك ارتباط بين المتغيرين أم لا أي إذا كان المطلوب هو اختبار الفرض العدمي : R = Zero <u>**أي أنه لا يوجد**</u> ارتباط حقيقي بين المتغيرين في المجتمع، مقابل الفرض البديل : $R \neq zero$ أي يوجد ارتباط حقيقي بين المتغيرين في المجتمع.

الثانية: إذا كان المطلوب هو اختبار أن R أي معامل ارتباط المجتمع يســاوي قيمة معينة (غير الصــفر) ولتكن RO، أي أن الفرض العدمي هو أن : R = RO مقابل الفرض البديل : R # RO (أي لا يســاوي هذه القيمة)، أو R > Ro (أي أكبر من هذه القيمة)، أو R < RO (أي أقل من هذه القيمة).

أولا : 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 يســاوي $\sqrt{\frac{1-r^2}{n-2}}$ وذلك بدرجات حرية n – 2. وبالتالي تكون خطوات اختبار أن معامل ارتباط المجتمع يساوي صفر كما يلي:

<u>1– الفرض العدمي :</u> أن معامل ارتباط المجتمع يساوي صفر، أي لا يوجد ارتباط بين المتغيرين. وبالرموز :

$$Ho : R = o$$

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

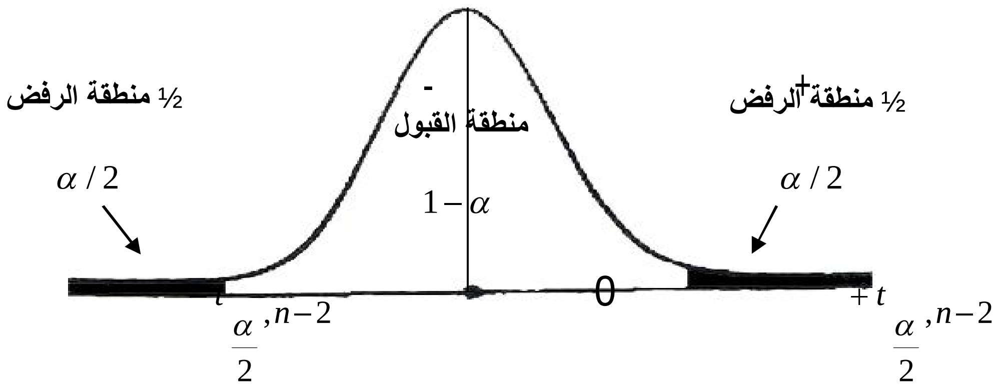
$$H_1 : R \# 0$$

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

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

والتي لها توزيع t بدرجات حرية $n-2$.

4 – حدود منطقتي القبول والرفض : والتي نحصل عليها من جدول t لمستوى معنوية يساوي $\frac{\alpha}{2}$ ودرجات حرية تساوي $n-2$ (اختبار الطرفين) :



- المقارنة والقرار: حيث نقارن قيمة إحصائية الاختبار (المحسوبة في الخطوة رقم 3) بحدود منطقتي القبول والرفض (من الخطوة رقم 4). 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 α.

مثال:

اختبر معنوية معامل الارتباط المحسوب في المثال رقم (2) وذلك بمستوى معنوية 5%.

الحل :

من المثال رقم (2) نجد أن :

$r = 0.81$ ، $n = 8$

وتكون خطوات اختبار معنوية الارتباط كما يلي :

1 – الفرض العدمي :

$$Ho : R = o$$

2 – الفرض البديل :

$$H1 : R \# o$$

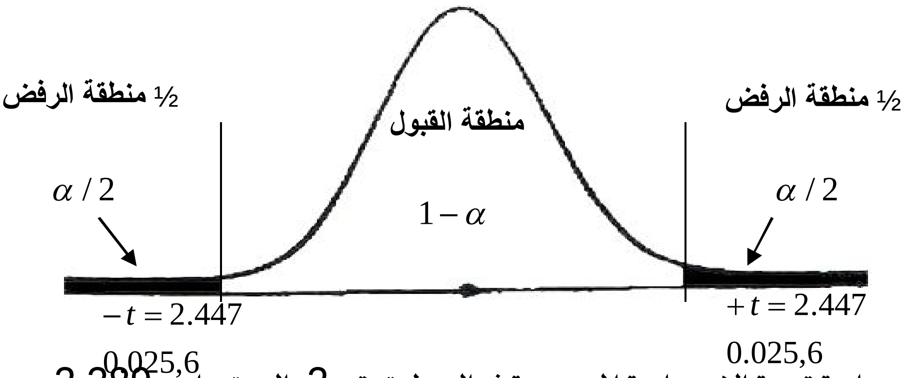
3 – الإحصائية :

$$t=\frac{r}{\sqrt{\frac{1-r^2}{n-2}}}=\frac{0.81}{\sqrt{\frac{1-(0.81)^2}{8-2}}}=\frac{0.81}{\sqrt{\frac{0.3439}{6}}}=\frac{0.81}{0.239}$$

$$t=3.389$$

4 – حدود منطقتي القبول والرفض :

من جدول t حيث مستوى المعنوية $\alpha=0.05$, $\frac{\alpha}{2}=0.025$ ودرجات الحرية تساوي (n – 2 = 8 – 2=6) نجد أن قيمة t تساوي 2.447 وتكون حدود منطقتي القبول والرفض كما يلي:

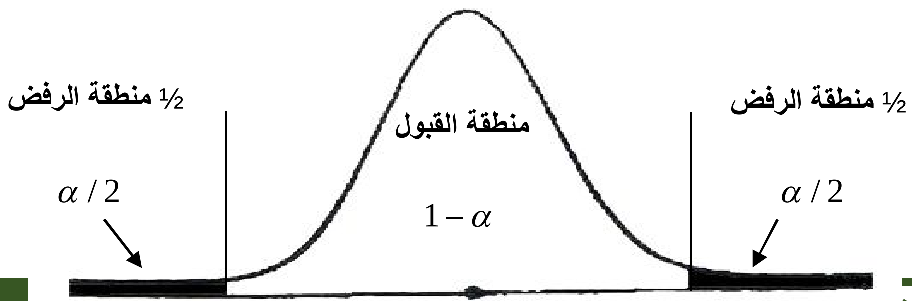


5 – المقارنة والقرار: بمقارنة قيمة الإحصائية المحسوبة في الخطوة رقم 3 والتي تساوي 3.389 بحدود منطقتي القبول والرفض (أو قيم t الجدولية في الخطوة رقم 4) نجد أنها تقع في منطقة الرفض (حيث أنها أكبر من 2.447) لذلك فإن القرار هو : **رفض الفرض العدمي وقبول الفرض البديل**. **أي رفض أن معامل الارتباط يساوي صفر**. وقبول أن معامل الارتباط **لا** يساوي صفر أي **يوجد ارتباط** بين المتغيرين (أعمار العمال ودخولهم اليومية) وذلك بمستوى معنوية 5 %.

ملاحظة مهمة :

إذا تغير مستوى المعنوية إلى 1% بدلا من 5 % فإن حدود منطقتي القبول والرفض تصبح كما يلي (من جدول t حيث $\alpha = 1\%$ ، ودرجات الحرية = 6):

حيث: $\alpha=0.01$ ، $\frac{\alpha}{2}=0.005$ فإن $t_{0.005,6}=3.707$



=0.005 =0.005

$$-t_{0.005,6}=3.707 \qquad +t_{0.005,6}=3.707$$

وبالتالي فإن قيمة الإحصائية 3.389 سوف تقع في منطقة القبول (حيث أنها أقل من 3.707 وهذا يعني أن القرار هو : قبول الفرض العدمي " أي قبول أن معامل الار تباط يساوي صفر، أي لا يوجد ارتباط خطي بين أعمار الناخبين ودخولهم اليومية في المجتمع، وذلك بمستوى معنوية 1% أي قد يتغير القرار إذا تغير مستوى المعنوية.

ثانيا : اختبار أن معامل ارتباط المجتمع يساوي قيمة معينة غير الصفر: (*)

خطوات الاختبار:

1– <u>الفرض العدمي :</u> H0 : R = Ro

أي أن الفرض العدمي هو أن معامل ارتباط المجتمع يساوي قيمة معينة (غير الصفر) ولتكن Ro.

2– <u>الفرض البديل :</u>

H1 : R $\neq$ Ro
or R > Ro
or R < Ro

أي أن الفرض البديل قد يأخذ أحد أشكال ثلاثة هي :

أ – أما معامل ارتباط المجتمع لا يساوي هذه القيمة Ro.

ب – أو أنه أكبر من هذه القيمة.

ج – أو أنه أصغر منها.

3– <u>إحصائية الاختبار:</u> ستكون الإحصائية في هذه الحالة Z والتي تأخذ الشكل التالي:

$$Z=\frac{U-\mu_u}{\sigma_u}$$

والتي لها توزع طبيعي معياري حيث :

(*) يمكن الإشارة فقط إلى هذه الحالة دون الدخول في التفاصيل.

$$u=\frac{1}{2}\log e\left(\frac{1+r}{1-r}\right)\approx 1.15\log_{10}\left(\frac{1+r}{1-r}\right)$$

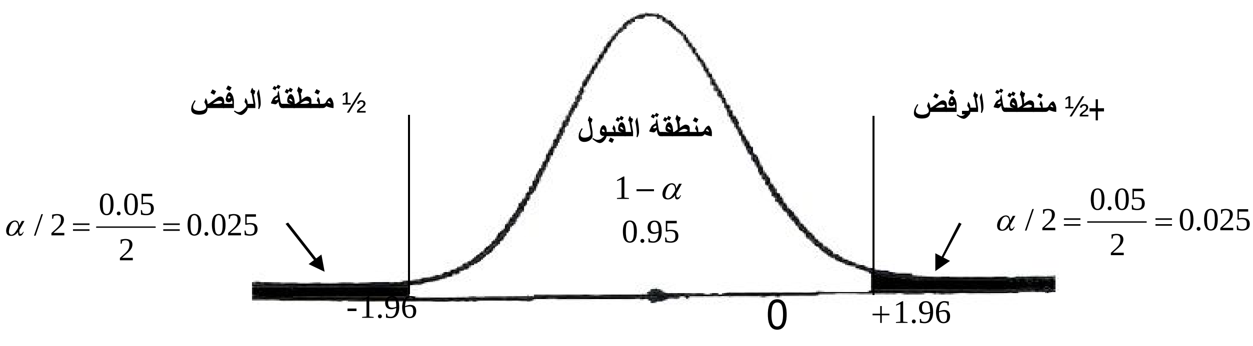
$$\mu_u=\frac{1}{2}\log e\left(\frac{1+Ro}{1-Ro}\right)\approx 1.15\log_{10}\left(\frac{1+Ro}{1-Ro}\right)$$

$$\sigma_u=\frac{1}{\sqrt{n-3}}$$

- أي يتم أولا حساب كل من u, μ_u, σ_u حسب العلاقات السابقة مع ملاحظة أن log e يعني اللوغاريتم للأساس الطبيعي e بينما تعني العلاقة أنها تساوي تقريبا وأن log 10 هو اللوغاريتم المعروف للأساس 10.

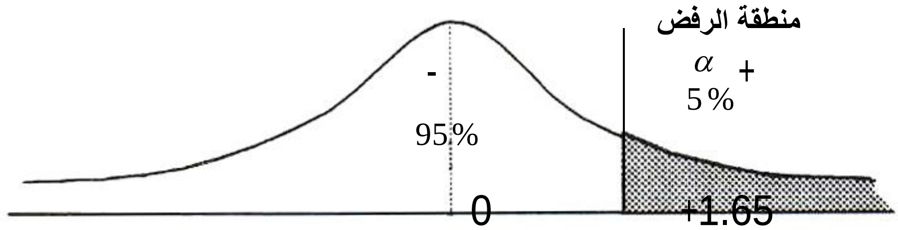
4 – <u>حدود منطقتي القبول والرفض</u>: سنحصل عليها من توزيع Z (الطبيعي المعياري) حسب مستوى المعنوية المطلوب وعلى حسب الفرض البديل.

* فإذا كان الفرض البديل **<u>" لا يساوي "</u>** نستخدم **<u>اختبار الطرفين</u>** حيث تتوزع منطقة الرفض (أو مستوى المعنوية) على طرفي المنحنى بالتساوي كما يلي: (بافتراض أن مستوى المعنوية 5% مثلا).



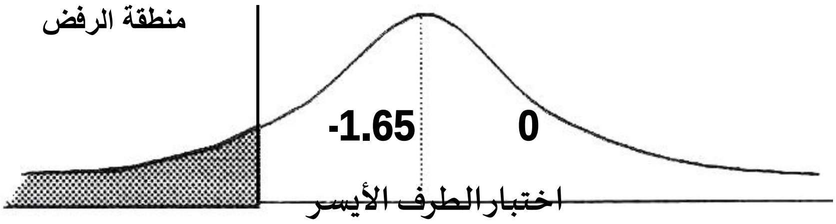
اختبار الطرفين

* أما إذا كان الفرض البديل "أكبر من" فإننا نستخدم اختبار الطرف الأيمن حيث تتركز منطقة الرفض (أو مستوى المعنوية) في الطرف الأيمن للمنحنى كما يلي **(وذلك بافتراض أن مستوى المعنوية 5%):**



اختبار الطرف الأيمن

* وإذا كان الفرض البديل "أقل من" فإننا نستخدم اختبار الطرف الأيسر حيث تتركز منطقة الرفض (أو مستوى المعنوية) في الطرف الأيسر للمنحنى كما يلي **(وذلك بافتراض أن مستوى المعنوية 5%):**



5 – <u>المقارنة والقرار :</u> حيث نقارن قيمة الإحصائية المحسوبة من الخطوة رقم 3 بحدود منطقتي القبول والرفض (أي قيم Z من الجدول). فإذا وقعت في منطقة القبول نقبل الفرض العدمي، والعكس، إذا وقعت في منطقة الرفض نرفض الفرض العدمي ونقبل البديل.

<u>مثال (3)</u>

في المثال رقم (1) الخاص بمعامل الارتباط بين أعمار عينة من الناخبين ودخولهم اليومية والذي يساوي 0.81 اختبر الفرض العدمي أن معامل الارتباط في المجتمع (بين أعمار الناخبين ودخولهم اليومية) يساوي 0.90 أي اختبر أن : Ho: R=0.90 مقابل الفرض البديل أن معامل الارتباط في المجتمع أقل من ذلك، أي أن : H1 : R < 0.90 وذلك بمستوى معنوية 5 %.

<u>الحل :</u>

1 – <u>الفرض العدمي :</u> أن معامل الارتباط الخطي في المجتمع بين اعمار الناخبين ودخولهم اليومية يساوي 0.90

أي أن :

$$Ho : R = 0.90$$

2 – <u>الفرض البديل :</u> أن معامل الارتباط الخطي في المجتمع بين أعمار الناخبين ودخولهم اليومية أقل من 0.90

أي أن :

$$H1 : R < 0.90$$

3 – <u>إحصائية الاختبار:</u>

$$Z = \frac{u - \mu_u}{\sigma_u}$$

ويتم حساب u , μ_u , σ_u كما يلي :

$$u=\frac{1}{2}\log_{e}\left(\frac{1+r}{1-r}\right)=1.15\ \log_{10}\left(\frac{1+r}{1-r}\right)$$

$$=1.15\log 10\left(\frac{1+0.81}{1-0.81}\right)=1.15\log_{10}\frac{1.81}{0.19}$$

$$=1.15\log_{10}9.53\quad=1.15\,(0.9789)$$

$$u=1.125769$$

$$\mu_u=\frac{1}{2}\log_{e}\left(\frac{1+Ro}{1-Ro}\right)=1.15\ \log_{10}\left(\frac{1+Ro}{1-Ro}\right)$$

$$\mu_u=1.15\log_{10}\left(\frac{1+0.90}{1-0.9}\right)=1.15\log_{10}\frac{1.9}{0.1}$$

$$=1.15\ (1.27875)$$

$$\mu_u=1.47056$$

$$\sigma_u=\frac{1}{\sqrt{n-3}}$$

$$=\frac{1}{\sqrt{8-3}}=\frac{1}{\sqrt{5}}=\frac{1}{2.236}$$

$$\sigma_u=0.44723$$

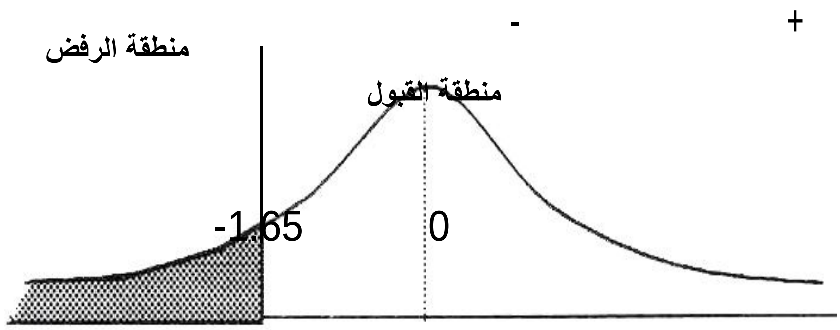
وبالتعويض في الإحصائية نحصل على :

$$Z=\frac{u-\mu_u}{\sigma_u}$$

$$=\frac{1.12576-1.47056}{0.44723}$$

$$Z=-0.77$$

4 – <u>حدود منطقتي القبول والرفض</u> : من جدول التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% واستخدام اختبار الطرف الأيسر (لأن الفرض البديل أقل من) تكون منطقتي القبول والرفض كما يلي :



ومن الرسم نجد أن منطقة القبول تشمل النصف الموجب كله (الأيمن) من المنحنى، وكذلك من الصفر وحتى 1.65 – بينما منطقة الرفض تشمل القيم التي أقل من 1.65 –.

5 – <u>المقارنة والقرار</u>: وحيث أن قيمة الإحصائية والتي تساوي 0.77 – تقع في منطقة القبول فإن القرار هو قبول الفرض العدمي بأن معامل الارتباط الخطي في المجتمع بين أعمار الناخبين ودخولهم اليومية يساوي 0.90 وذلك بمستوى معنوية 5 %.

2-3-معامل ارتباط الرتب Rank Correlation :

قد يرغب الباحث في حســـاب معامل الارتباط بين رتب المتغيرين وليس بين القيم ذاتها، فقد يكون المتغيران **وصــفيين ترتيبين Ordinal أو أن يكون أحد المتغيرين كميا ، بينما الأخر وصــفيا ، ترتيبيا** ، أو أن يكون **المتغيران كميين** ويكون اهتمام البـاحـث منصــبا على الرتب أكثر من القيم. ففي انتخـابـات مجلس الشــيوخ أو النواب الأكبر مثلا ، يعتبر المرشح الأول هو من حصل على أعلى الأصوات بغض النظر عن عددها، والذي يحصل على عدد أصوات أقل منه مباشرة هو الثاني.. وهكذا.

فإذا كانت رتب المتغيرين تســير في الاتجاه نفســه : بمعنى أن الرتب الأعلى للمتغير الأول تناظرها رتب أعلى للمتغير الثاني كانت العلاقة طردية بينهما. وإذا كانت الرتب الأعلى للمتغير الأول تناظرها رتب أدنى للمتغير الثاني كانت العلاقة بينهما عكســية. ففي مثالنا الســابق عن العلاقة بين دخل الناخب وعمره، كان الناخب الأكبر عمرا (بصــفة عامة) هو الأعلى دخلا ، فالواضــح أن العلاقة بينهما طردية، أما إذا كان الناخب الأكبر عمرا (بصــفة عامة) هو الأقل مشاركة في العمل السياسي فإننا في هذه الحالة نكون أمام علاقة عكسية.

ولحساب معامل ارتباط الرتب هناك طرق مختلفة أهمها معاملي سبيرمان وكيندال.

- معامل سبيرمان لارتباط الرتب : Spearman rank Correlation Coefficient

لحســاب معامل ســبيرمان لارتباط الرتب نقوم بترتيب كل من المتغيرين ترتيبا تصــاعديا أو تنازليا (أما تصاعديا لكلا المتغيرين أو تنازليا لكليهما). وفي حالة الترتيب التصاعدي تأخذ أقل قيمة من قيم المتغير الرتبة رقم 1، والقيمة الأعلى منها مباشرة الرتبة رقم 2 وهكذا (بالنسبة لكل من المتغيرين). أما في حالة الترتيب التنازلي تأخذ أكبر قيمة من قيم المتغير الرتبة رقم 1، والقيمة الأقل منها مبـاشــرة الرتبـة رقم 2 وهكـذا (بـالنســبـة لكل من المتغيرين). وعند تســاوي قيمتين (أو أكثر) من قيم المتغير نعطي كل قيمة رتبة مختلفة (كما لو كانت القيم غير متساوية) ثم نحسب متوسط هذه الرتب، ويعطى هذا المتوسط لكل من هذه القيم المتساوية.

وبعد ترتيب المتغيرين نحســب الفروق بين رتب كل من المتغيرين (ونرمز للفروق بالرمز d) ثم نقوم بتربيع هذه الفروق ونحصــل على مجموعها أي نحصــل على $\sum d^2$ ثم نعوض في معامل ســبيرمان لارتباط الرتب والذي يأخذ الشكل التالي :

$$r_s = 1 - \frac{6(\sum d^2)}{n(n^2-1)}$$

حيث : $\sum d^2$ هو مجموع مربعات الفروق بين رتب المتغيرين، n هي عدد أزواج القيم.

مما سبق نستطيع إجمال بعضا من الملاحظات فيما يلي :

1 – مجموع الفروق بين الرتب يساوي صفر.

2 – أن قيمة معامل ارتباط الرتب تنحصر بين – 1، + 1 فإذا كانت الرتبة رقم 1 للمتغير الأول تناظرها الرتبة 1 للمتغير الثاني، والرتبة 2 للمتغير الأول تناظرها الرتبة رقم 2 للمتغير الثاني، وهكذا.. فإن معامل ارتباط الرتب يساوي + 1 (ارتباط طردي تام بين الرتب). وإذا كانت الرتبة رقم 1 (أقل رتبة) للمتغير الأول تناظرها أعلى رتبة للمتغير الثاني وهكذا.. فإن معامل ارتباط الرتب يساوي – 1 (ارتباط عكسي تام بين الرتب).

3 – كذلك نلاحظ أن مجموع الرتب لكل من المتغيرين تساوي $\frac{n(n+1)}{2}$.

مثال (4):

البيانات التالية تمثل إجابات عينة من سبعة أشخاص حول برامج الضمان الاجتماعي، ومدى ملاءمتها لحاجات الناس.

السؤال الأول	جيدة	مقبولة	ممتازة	جيدة	جيدة جداً	مقبولة	جيدة
السؤال الثاني	جيدة جداً	مقبولة	جيدة جداً	جيدة	جيدة	جيدة	ممتازة

والمطلوب حساب معامل سبيرمان لارتباط الرتب بين هذين السؤالين ؟

الحل :

تنظم الحل في الجدول التالي مع ملاحظة ما يلي :

1 – بالنسبة للسؤال الأول، فإن التقدير الأعلى سيحصل على الرتبة رقم 1 والأقل منه مباشرة سيحصل على الرتبة رقم 2 وهكذا.. أي أن الترتيب تنازلي. ونكرر العمل نفسه مع السؤال الثاني.

2 – عند حول إجابتين أو أكثر على التقدير نفسه نعطي لكل إجابة مبدئيا رتبة **كما لو كانوا مختلفين** ثم نحسب متوسط هذه الرتب، وهذا المتوسط هو الذي يعطى لكل إجابة.

3 – ثم نحسب الفروق بين رتب السؤالين ونرمز لها بالرمز d ثم نربع هذه الفروق فنحصل على d^2 ونعوض في القانون عن $\sum d^2$ مع ملاحظة أن n = 7.

السؤال الأول X	السؤال الثاني Y	رتب X	رتب Y	d الفرق بين الرتب	d^2 مربعات الفرق
جيدة	جيدة جداً	4	2.5	1.5	2.25
مقبولة	مقبولة	6.5	7	- 0.05	0.25
ممتازة	جيدة جداً	1	2.5	- 1.5	2.25
جيدة	جيدة	4	5	- 1.0	1.00

ة جداً	جيدة	2	5	- 3.0	9.00
مقبولة	جيدة	6.5	5	1.5	2.25
جيدة	ممتازة	4	1	3.0	9.00
المجموع				**Zero**	**26.0**

$$r_s = 1 - \frac{6(Sd^2)}{n(n^2-1)} = 1 - \frac{6(\sum d^2)}{n(n^2-1)} = 1 - \frac{6(26)}{7(49-1)}$$

$$= 1 - \frac{156}{336} = 1 - 0.46$$

$$r_s = 0.54$$

وهذا يعني أن الارتباط بين إجابات المبحوثين بالنسبة للسؤالين هو ارتباط طردي متوسط. وبالتالي فليس بالضرورة أن يكون رأي المجيبين في برامج الضمان الاجتماعي تعني ملاءمتها لحاجات الناس.

مثال (5)

البيانات التالية تمثل أعداد الساعات التي ذاكرها عشرة طلاب والدرجات التي حصلوا عليها في امتحان أحد المقررات :

عدد الساعات X	10	6	12	14	11	6	19	16	3	9
الدرجات y	60	48	83	76	74	58	98	89	37	69

أحسب معامل سبيرمان لارتباط الرتب.

الحل :

كما في المثال السابق ننظم الحل في الجدول التالي مع ملاحظة أن n = 10

عدد الساعات X	الدرجات Y	رتب X	رتب Y	الفروق d	d^2
10	60	6	7	- 1	1.00
6	48	8.5	9	- 0.5	0.05
12	83	4	3	1	1.00
14	76	3	4	- 1	1.00
11	74	5	5	0	0
6	58	8.5	8	0.5	0.25
19	98	1	1	0	0
16	89	2	2	0	0
3	37	10	10	0	0
9	69	7	6	1	1.00

المجموع				Zero	4.50

وبالتعويض في القانون حيث $\sum d^2 = 4.5$ ، $n = 10$ نحصل على :

$$r_s = 1 - \frac{6(\sum d^2)}{n(n^2 - 1)} = 1 - \frac{6(4.5)}{10(1-99)} = 1 - \frac{27}{990} = 1 - 0.027$$

$$r_s = 0.973$$

مما يعني أننا أمام علاقة طردية قوية بين المتغيرين. فكلما زادت عدد الساعات التي يدرسها الطالب في هذا المثال، كلما زادت درجاته في الامتحانات قوة وذلك بما نسبته 97 %.

اختبار معنوية ارتباط الرتب:

عند اختبار الفرض العدمي بعدم وجود ارتباط رتب بين المتغيرين لسنا في حاجة لوضع أي شروط عن طبيعة المجتمع المسحوبة منه العينة.

وتحت الفرض العدمي بعدم وجود ارتباط فإن توزيع المعاينة للمعامل يكون له متوسط يساوي صفر وانحراف معياري يساوي : $\sigma\, rs = \frac{1}{\sqrt{n-1}}$ وأن هذا التوزيع يكون لـقريبا – توزيع طبيعي فإن خطوات الاختبار تكون كما يلي :

1 – الفرض العدمي : لا يوجد ارتباط بين المتغيرين **(أو معامل الارتباط يساوي الصفر):**

$$Ho : R = Zero$$

2 – الفرض البديل : يوجد ارتباط بين المتغيرين **(أو معامل الارتباط لا يساوي الصفر):**

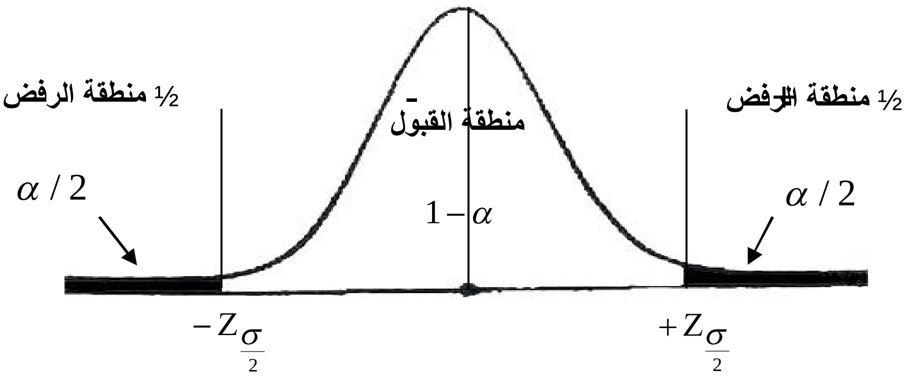
$$H1 : R \# Zero$$

3 – الإحصائية : والتي تكتبلختصارا – كما يلي :

$$Z = r_s \sqrt{n-1}$$

والتي لها توزيع طبيعي معياري.

4 – حدود منطقتي القبول والرفض : (اختبار الطرفين للتوزيع الطبيعي)



5 – المقارنة والقرار: حيث نقارن قيمة الإحصـــائية بحدود منطقتي القبول والرفض. فإذا وقعت في منطقة القبول نقبل الفرض العدمي والعكس صحيح.

مثال (6)

لبيانات المثال السابق رقم (5) حيث $n = 10$، $r_s = 0.973$

اختبر الفرض العدمي بعدم وجود ارتباط بين عدد الساعات التي يذاكرها الطالب والدرجات التي يحصل عليها في الامتحان وذلك بمستوى معنوية 1 %.

الحل:

1 – الفرض العدمي: لا يوجد ارتباط بين المتغيرين. أي أن :

$$Ho : R = Zero$$

2 – الفرض البديل: يوجد ارتباط بينهما. أي أن :

$$H1 = R \# Zero$$

3 – الإحصائية :

$$Z = r_s \sqrt{n - 1} = 0.973 \sqrt{10 - 1} = 0.973 \sqrt{9}$$
$$Z = 2.919$$

4 – حدود منطقتي القبول والرفض :

توزيع Z، واختبار الطرفين، ومستوى المعنوية 1%

$$Z_{1-\alpha/2} = 2.58$$

5 – المقارنة والقرار: وحيث أن قيمة الإحصـائية (2.919) تقع في منطقة الرفض (أكبر من 2.58) فإن القرار هو رفض الفرض العدمي وقبول الفرض البديل بأن هناك ارتباط بين المتغيرين وذلك بمســتوى معنوية 1%.

السلسلة رقم: 03-

التمرين الأول:

البيانات التالية تمثل حجم المبيعات(Y) لمؤسسة A ومصاريف الدعاية (X) خلال الفترة 2005-2012.

السنوات	2005	2006	2007	2008	2009	2010	2011	2012
yحجم المبيعات	10	18	15	35	40	62	100	50
Xمصاريف الدعاية	25	32	29	43	38	51	47	35

والمطلوب: - أحسب معامل بيرسون للارتباط الخطي بين حجم المبيعات ومصاريف الدعاية مع تفسير شكل العلاقة؟

- اختبر معنوية معامل الإرتباط عند مستوى معنوية 5%؟.

التمرين الثاني:

البيانات التالية تمثل متوسط المبيعات الشهرية للمشروبات الغازية للشركة (A) ومتوسط درجات الحرارة الشهرية خلال سنة 2017.

ديسمبر	نوفمبر	أكتوبر	سبتمبر	أوت	جويلية	جوان	ماي	أفريل	مارس	بفري ي	جانفي	الشهر
7	10	15	30	31	35	32	20	15	10	7	5	درجات الحرارة X
15	30	40	55	65	70	60	50	40	30	20	10	حجم المبيعات y

- أحسب معامل سبيرمان لإرتباط الرتب؟.
- قدم تفسيرا للنتائج المستخلصة.
- اختبر معنوية معامل الإرتباط عند مستوى معنوية 5%؟.

الحل:

✓ **التمرين الأول:**

1- حساب معامل بيرسون للارتباط الخطي

لإحتساب معامل الإرتباط للمتغيرين X وy نستعين بالجدول الموالي والمبين للحسابات التفصيلية:

السنوات	Y	X			$(y_i-\bar{y})$	$(x_i-\bar{x})$	$(x_i-\bar{x})(y_i-\bar{y})$	$(Y-\bar{Y})^2$	$(X-\bar{X})^2$
2005	10	25	41,25	37,5	-31,25	-12,5	390,625	976,5625	156,25
2006	18	32	41,25	37,5	-23,25	-5,5	127,875	540,5625	30,25
2007	15	29	41,25	37,5	-26,25	-8,5	223,125	689,0625	72,25

2008	35	43	41,25	37,5	-6,25	5,5	-34,375	39,0625	30,25
2009	40	38	41,25	37,5	-1,25	0,5	-0,625	1,5625	0,25
2010	62	51	41,25	37,5	20,75	13,5	280,125	430,5625	182,25
2011	100	47	41,25	37,5	58,75	9,5	558,125	3451,5625	90,25
2012	50	35	41,25	37,5	8,75	-2,5	-21,875	76,5625	6,25
Σ	330	300					1523	6205,5	568

حيث أن: $\bar{x} = 37.5$ و $\bar{y} = 41.25$

وبالرجوع إلى العلاقة:

$$r_{x.y} = \frac{cov(x,y)}{\delta_x . \delta_y}$$

$$cov(x,y) = E[(x_i - \bar{x})\,(y_i - \bar{y})]$$

نجد أن:

$$r_{x.y} = 0.8112$$

التفسير: وجود علاقة طردية قوية بين مصاريف الدعاية وحجم المبيعات تقدر بـ 81,12%.

- **اختبار معنوية معامل الإرتباط:** وتكون خطوات اختبار معنوية الارتباط كما يلي :

1 – الفرض العدمي :

$$Ho : R = o$$

2 – الفرض البديل :

$$H1 \; : R \# o$$

3 – الإحصائية :

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.81}{\sqrt{\frac{1-(0.81)^2}{8-2}}} = \frac{0.81}{\sqrt{\frac{0.3439}{6}}} = \frac{0.81}{0.239}$$

$$t = 3.389$$

4 – حدود منطقتي القبول والرفض :

من جدول t حيث مستوى المعنوية $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$ ودرجات الحرية تساوي 6=2 – 8 = 2 – n))

نجد أن قيمة t تساوي 2.447 .

5 – المقارنة والقرار: بمقارنة قيمة الإحصائية المحسوبة في الخطوة رقم 3 والتي تساوي 3.389 بحدود منطقتي القبول والرفض (أو قيم t الجدولية في الخطوة رقم 4) 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 وذلك بمستوى معنوية 5%.

✓ **التمرين الثاني:**

ننظم الحل في الجدول التالي مع ملاحظة أن n = 12

درجات الحرارة	حجم المبيعات	رتب X	رتب Y	الفروق d	d^2
X	Y				
5	10	1	1	0	0
7	20	2,5	2	0,5	0,25
10	30	4,5	4,5	0	0
15	40	6	6,5	-0,5	0,25
20	50	8	8	0	0
32	60	11	11	0	0
35	70	12	12	0	0
31	65	10	10	0	0
30	55	9	9	0	0
15	40	7	6,5	0,5	0,25
10	30	4,5	4,5	0	0
7	15	2,5	3	-0,5	0,25
المجموع				**0**	**1**

وبالتعويض في القانون حيث $\sum d^2 = 4.5$ ، n = 10 نحصل على :

$$r_s = 1 - \frac{6(\sum d^2)}{n(n^2-1)} = 1 - \frac{6(1)}{12(144-1)}$$

$$r_s = 0.9965$$

مما يعني أننا أمام علاقة طردية قوية بين المتغيرين. فكلما زادت درجات الحرارة كلما زاد حجم وذلك بما نسبته 99.65%.

اختبار معنوية معامل ارتباط الرتب:

1 – الفرض العدمي : لا يوجد ارتباط بين المتغيرين. أي أن :

Ho : R = Zero

2 – الفرض البديل : يوجد ارتباط بينهما. أي أن :

$$H1 = R \# Zero$$

3 – الإحصائية :

$$Z = r_s \sqrt{n-1} = 0.9965 \sqrt{12-1}$$
$$Z = 3.305$$

4 – القيمة الجدولية :

$$Z_{1-\alpha/2} = 1.96$$

5 – المقارنة والقرار : وحيث أن قيمة الإحصائية تقع في منطقة الرفض (أكبر من المجدولة) فإن القرار هو ر فض الفرض العدمي و قبول الفرض البديل بأن هناك ارتباط بين المتغيرين وذلك بمستوى معنوية 5%.

المراجع:

أولا- المراجع باللغة العربية:

1- محمد عبد العال النعيمي ومزهر شعبان العاني، الأساليب الإحصائية باستخدام حزمة MATLAB، دار وائل، عمان، 2008.

2- إبراهيم علي إبراهيم عبد ربه وناشد محمود عبد السلام، مبادئ التحاليل الإحصائية (بين النظرية والتطبيق)، دار المطبوعات الجامعية، مصر، 2008.

3- وليد إسماعيل السيفو وآخرون، أساسيات الأساليب الإحصائية للأعمال، دار زمزم، عمان، 2010.

4- يحي سعد زغلول، الإحصاء التطبيقي، الدار الجامعية، بيروت، 1988.

5- إبراهيم محمد البطاينة، مبادئ الإحصاء، دار المسيرة، عمان، 2010.

6- مراد كمال عوض، أساسيات الإحصاء، دار البداية، عمان، 2010.

7- عبد الله صالح، محاضرات في الإحصاء الرياضي، مطبوعة مقدمة لطلبة العلوم الاقتصادية، جامعة المسيلة.

8- فتحي عبد العزيز أبو راضي، مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، مصر، 2013.

9- امتثال محمد حسن ومحمد علي محمد أحمد، مبادئ الإستدلال الإحصائي، الدار الجامعية، مصر، 2002.

10- ياسر أحمد السيد، الإحصاء التطبيقي، مكتبة بستان المعرفة، مصر، 2009.

11- سالم عيسى بدر وعماد غصاب عبابنة، مبادئ الإحصاء الوصفي والإستدلالي، دار المسيرة، عمان، 2010.

12- عبد الحميد عبد المجيد البلداوي، الأساليب الإحصائية التطبيقية، ط2، دار الشروق، 2008.

13- حسن ياسين طعمة، الإختبارات الإحصائية، دار الصفاء، عمان، 2015.

14- ثائر فيصل شاهر، اختبار الفرضيات الإ حصائية، دار الحامد، عمان، 2013.

15- عزام صبري، الإحصاء الرياضي، دار الصفاء، ط2، عمان، 2014.

16- نجاة رشيد الكيخيا، أساسيات الإستنتاج الإحصائي، دار المريخ، 2007.

17- معتوق أمحمد، الإحصاء الرياضي والنماذج الإحصائية، ديوان المطبوعات الجامعية، الجزائر، 2015.

18- دومينيك سالفاتور، نظريات ومسائل في الإحصاء والإقتصاد القياسي، ترجمة سعدية حافظ منتصر، الدار الدولية للنشر والتوزيع، مصر، 2011.

19- طه حسين الزبيدي، مبادئ الإحصاء، دار المنهل، لبنان، 2013.

ثانيا- المراجع باللغات الأجنبية:

1- Jean-Jacques Droesbeke, Eléments De Statistique, OPU, Alger, 1988.

2- Wayne A.Fuller, Sampling Statistics, Wiley Publication, New Jersey, 2009.

3- Bernard Verlant et Geneviève Saint-Pierre, Statistiques et Probabilités, Berti éditions, Alger, 2008.

4- David R.Anderson et Dennis J.Sweeney, Statistique pour L'économie et la Gestion, Traduction de la 7eme édition américaine par Claire Borsenberger, de boeck supérieure, Paris, 2015.

5- Steven J. Janke, Frederick Tinsley, Introduction to Linear Models and Statistical Inference, Wiley Interscience Publication, New Jersey, 2005.

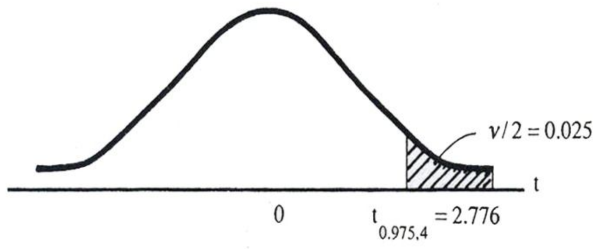
6- Nicholas P. Cheremisinoff, Louise Ferrante, Practical Statistics for Engineers and Scientists, Technomic Publishing Company,1987.

الجداول الإحصائية

Areas Under The Standard Normal Curve

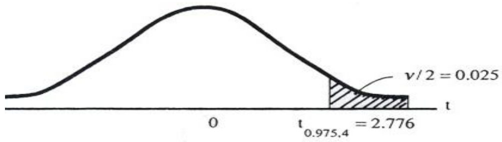
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
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
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Values of – t



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**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
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
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
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
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
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
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
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										

– χ^2 Values of



v	$\alpha = 0.995$	$\alpha = 0.99$	$\alpha = 0.975$	$\alpha = 0.95$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
1	.0000393	.000157	.000982	.00393	3.841	5.024	6.635	7.879
2	.0100	.0201	.0506	.103	5.991	7.378	9.210	10.597
3	.0717	.155	.216	.352	7.815	9.348	11.345	12.838
4	.207	.297	.484	.711	9.488	11.143	13.277	14.860
5	5.214	.554	.831	1.145	11.070	12.832	15.086	16.750
6	.676	.872	1.237	1.635	12.592	14.449	16.812	18.548
7	.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	14.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

Values of $F_{0.05}$

v_1 Degrees of freedom for numerator درجات حرية البسط

	1	2	3	4	5	6	7	8	9	10	12	15	20	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	563
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.71
10	4.96	4.01	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	1.95
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.00

ملاحظة:لإيجاد قيم $F0.95(V1,V2)$ نستخدم العلاقة التالية $F0.95(V1,V2) = {}^{1}/_{F0.05\,(V2,V1)}$

Values of $F_{0.025}$

/	df_1=1	2,00	3,00	4,00	5,00	6,00	7,00	8,00	9,00	10,00	12,00	15,00	20,00
df_2=1	647.79	799.5	864.16	899.58	921.85	937.11	948.22	956.66	963.28	968.63	976.71	984.87	993.10
2,00	38.51	39.00	39.16	39.25	39.298	39.33	39.35	39.373	39.387	39.398	39.41	39.43	39.45
3,00	17.4434	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.4731	14.4189	14.3366	14.2527	14.1674
4,00	12.2179	10.65	9.9792	9.6045	9.3645	9.1973	9.0741	8.9796	8.9047	8.8439	8.7512	8.6565	8.5599
5,00	10.0070	8.4336	7.7636	7.3879	7.1464	6.9777	6.8531	6.7572	6.6811	6.6192	6.5245	6.4277	6.3286
6,00	8.8131	7.2599	6.5988	6.2272	5.9876	5.8198	5.6955	5.5996	5.5234	5.4613	5.3662	5.2687	5.1684
7,00	8.0727	6.5415	5.8898	5.5226	5.2852	5.1186	4.9949	4.8993	4.8232	4.7611	4.6658	4.5678	4.4667
8,00	7.5709	6.0595	5.4160	5.0526	4.8173	4.6517	4.5286	4.4333	4.3572	4.2951	4.1997	4.1012	3.9995
9,00	7.2093	5.7147	5.0781	4.7181	4.4844	4.3197	4.1970	4.1020	4.0260	3.9639	3.8682	3.7694	3.6669
10,00	6.9367	5.4564	4.8256	4.4683	4.2361	4.0721	3.9498	3.8549	3.7790	3.7168	3.6209	3.5217	3.4185
11,00	6.7241	5.2559	4.6300	4.2751	4.0440	3.8807	3.7586	3.6638	3.5879	3.5257	3.4296	3.3299	3.2261
12,00	6.5538	5.0959	4.4742	4.1212	3.8911	3.7283	3.6065	3.5118	3.4358	3.3736	3.2773	3.1772	3.0728
13,00	6.4143	4.9653	4.3472	3.9959	3.7667	3.6043	3.4827	3.3880	3.3120	3.2497	3.1532	3.0527	2.9477
14,00	6.2979	4.8567	4.2417	3.8919	3.6634	3.5014	3.3799	3.2853	3.2093	3.1469	3.0502	2.9493	2.8437
15,00	6.1995	4.7650	4.1528	3.8043	3.5764	3.4147	3.2934	3.1987	3.1227	3.0602	2.9633	2.8621	2.7559
16,00	6.1151	4.6867	4.0768	3.7294	3.5021	3.3406	3.2194	3.1248	3.0488	2.9862	2.8890	2.7875	2.6808
17,00	6.0420	4.6189	4.0112	3.6648	3.4379	3.2767	3.1556	3.0610	2.9849	2.9222	2.8249	2.7230	2.6158
18,00	5.9781	4.5597	3.9539	3.6083	3.3820	3.2209	3.0999	3.0053	2.9291	2.8664	2.7689	2.6667	2.5590
19,00	5.9216	4.5075	3.9034	3.5587	3.3327	3.1718	3.0509	2.9563	2.8801	2.8172	2.7196	2.6171	2.5089
20,00	5.8715	4.4613	3.8587	3.5147	3.2891	3.1283	3.0074	2.9128	2.8365	2.7737	2.6758	2.5731	2.4645
21,00	5.8266	4.4199	3.8188	3.4754	3.2501	3.0895	2.9686	2.8740	2.7977	2.7348	2.6368	2.5338	2.4247
22,00	5.7863	4.3828	3.7829	3.4401	3.2151	3.0546	2.9338	2.8392	2.7628	2.6998	2.6017	2.4984	2.3890
23,00	5.7498	4.3492	3.7505	3.4083	3.1835	3.0232	2.9023	2.8077	2.7313	2.6682	2.5699	2.4665	2.3567
24,00	5.7166	4.3187	3.7211	3.3794	3.1548	2.9946	2.8738	2.7791	2.7027	2.6396	2.5411	2.4374	2.3273
25,00	5.6864	4.2909	3.6943	3.3530	3.1287	2.9685	2.8478	2.7531	2.6766	2.6135	2.5149	2.4110	2.3005
30,00	5.5675	4.1821	3.5894	3.2499	3.0265	2.8667	2.7460	2.6513	2.5746	2.5112	2.4120	2.3072	2.1952
40,00	5.4239	4.0510	3.4633	3.1261	2.9037	2.7444	2.6238	2.5289	2.4519	2.3882	2.2882	2.1819	2.0677
60,00	5.2856	3.9253	3.3425	3.0077	2.7863	2.6274	2.5068	2.4117	2.3344	2.2702	2.1692	2.0613	1.9445
120,00	5.1523	3.8046	3.2269	2.8943	2.6740	2.5154	2.3948	2.2994	2.2217	2.1570	2.0548	1.9450	1.8249
∞	5.0239	3.6889	3.1161	2.7858	2.5665	2.4082	2.2875	2.1918	2.1136	2.0483	1.9447	1.8326	1.7085

ملاحظة: لإيجاد قيم $F0.975(V1,V2)$ نستخدم العلاقة التالية $F0.975(V1,V2) = {}^{1}/_{F0.025\,(V2,V1)}$