وزارة التعليم العالي والبحث العلمي

جامعة زيان عاشور – الجلفة –

كلية العلوم الإقتصادية والتجارة وعلوم التسيير

قسم علوم التسيير

محاضرات وتمارين في:

الإحصاء -03-

موجهة لطلبة السنة الثانية ليسانس

شعبة علوم التسيير

الإعداد:

الدكتور: المختار بن سكري

أستاذ محاضر - ب -

كلية العلوم الإقتصادية والعلوم التجارية وعلوم التسيير

جامعة الجلفة - الجزائر -

2023/2022

الفهرس

الفصل الثالث: اختبار الفرضيات

هذه المطبوعة هي عبارة عن محاضرات في مقياس الإحصاء (03) (كما يعرف بعدة مسميات تختلف باختلاف الجامعات في الوطن العربي منها :الإحصاء التطبيقي، الاحصاء الرياضي، الاحصاء التحليلي او الاستنتاجي) ، حيث أنه يمكن تقسيم علم الإحصاء إلى احصاء وصفي واحصاء استدلالي، هذا الأخير تندرج ضمنه نظرية الاحتمالات، والإحصاء التطبيقي، وحسب مقرر التدريس المحدد من قبل وزارة التعليم العالي الجزائرية، فيتم التدرج في تدريس علم الإحصاء على ثلاث مراحل كما يلي:

1. الإحصاء الوصفي: وهو علم يهتم بتلخيص وتوصيف مجموعة من البيانات "ويسمى ب الإحصاء (01) والذي يتم تدريسه خلال السداسي الأول لطلبة السنة أولى جدع مشترك،
2. والإحصاء (02): والذي يعني بنظرية الاحتمالات، والتي تعتبر عنصرا أساسيا في الاستدلال الإحصائي، حيث تساعد نظرية الاحتمالات على تفسير وتحليل المعطيات بشكل أفضل، ويدرس هذا المقياس خلال السداسي الثاني لطلبة السنة أولى جدع مشترك،
3. وأخيرا الإحصاء (03): وهو موجه لتدريس طلبة السنة ثانية علوم اقتصادية، علوم التسيير، علوم تجارية، والعلوم المالية خلال السداسي الثالث.

تمثل هذه المطبوعة ثمرة خبرة عدة سنوات في تدريس مقياس الإحصاء (03) لطلبة السنة ثانية بقسم علوم التسيير في جامعة زيان عاشور بالجلفة، وفي محاولة منا الاستيفاء للبرنامج المحدد تم تقسيم المطبوعة إلى الفصول التالية:

- **الفصل الأول**: **نظرية المعاينة**، ويعتبر هذا الفصل حجر الأساس الذي ترتكز عليه باقي الفصول حيث تستعمل النتائج المتحصل عليها فيه مباشرة.
- **الفصل الثاني**: **نظرية التقدير**، حيث تستخدم مجموعة من البيانات لتقدير (التقدير النقطي او التقدير بمجال ثقة) لمعلمة من معالم المجتمع المجهولة بالاعتماد على إحصاءة العينة المقابلة لها.
- **الفصل الثالث**: **اختبار الفرضيات**، ويتم من خلال صياغة فرضية ما حول معلمة من معالم المجتمع، ثم استخدام عينة لإتخاذ قرار بخصوص صحة هذه الفرضية من عدمها.

تمهيد

تهتم نظرية العينات بطريقة اختيار العينات واستخدام النتائج التي يتم الحصول عليها من العينة لتقدير خصائص المجتمع الأساسية (The parameters) مثل المتوسط، نسبة المجتمع، حجم المجتمع أو تباين المجتمع...الخ وتقييم هذه التقديرات. ونلجأ إلى دراسة العينة بدلا من المجتمع لعدة أسباب منها: إقتصاد الوقت وبعض من الجهد والمال.

إنّ الهدف من نظرية المعاينة هو تحديد طرق المعاينة التي تحقق الخصائص الرئيسية التالية:

- أن تكون العينة ممثلة تمثيلا جيدا لمجتمع الدراسة، أي تحمل نفس خصائص المجتمع؛
- أن تكون التقديرات التي نحصل عليها لمعالم المجتمع من بيانات العينة دقيقة ويمكن قياس مصداقيتها؛
- أن تكون تكلفة اختيار العينة أقل ما يمكن.

سنتناول في هذا الفصل بعض المفاهيم الإحصائية بالإضافة إلى التوزيعات الاحتمالية المتعلقة بموضوع نظرية أخذ العينات والتي سيحتاجها الطالب في دراسته لهذا الفصل الدراسي وبقية الفصول الأخرى، كما سنتناول توزيع العينات للمتوسطات وتوزيع العينات للتباين وتوزيع العينات للنسبة، الاختلافات والمجاميع.

أولا: مفاهيم إحصائية

من الضروري قبل الدخول في التفاصيل أن نتطرق إلى التعريف ببعض المصطلحات الإحصائية والتمييز بين المفاهيم مثل المجتمع الإحصائي والعينة، المعاينة بالإرجاع وبدون إرجاع، العينة العشوائية، معالم المجتمع، إحصاءات العينة وخطأ المعاينة.

1-1- المجتمع الإحصائي

تستخدم لفظة مجتمع في علم الإحصاء لتعنى مجموعة من المفردات أو العناصر تهم الباحث وتشترك في خاصية أو أكثر. ومفردات المجتمع تسمى وحدات أولية (Units Elementary) وتمتلك الوحدات الأولية خصائص (Characteristics) وهذه الخصائص قد تكون نوعية(Qualitative) أو كمية (Quantitative).

<u>مثال</u>: إذا كان الهدف من الدراسة هو التعرف على أعمار مجموعة من الطلبة بجامعة الجلفة في فترة ما، فإن المجتمع الإحصائي هو جميع الطلبة المسجلون بجامعة الجلفة في تلك الفترة.

2-1- العينة Sample

تعرّف العينة على أنها جزء من المجتمع حيث تتكون من عدد محدود من المفردات أو المشاهدات(أي أنها مجموعة جزئية من المجتمع).

<u>مثال</u>: إذا كان عدد الطلبة الذين تقدموا لامتحان مقياس الرياضيات 150 طالب، فالمجتمع في هذه الحالة هو علامات 150 طالب في مقياس الرياضيات. أخذنا علامات 15 طالب فقط من بين الذين تقدموا لهذا الامتحان نقول أنه تم اختيار عينة من علامات 15 طالب من مجتمع علامات مقياس الرياضيات.

3-1- المعاينة بالإرجاع (Sampling with replacement) والمعاينة بدون إرجاع (without replacement)

عندما يكون السحب بالإرجاع يعني المفردة التي تم سحبها بمكن أن تظهر أكثر من مرة في العينة، في هذه الحالة تسمى المعاينة بالمعاينة بالإرجاع. والعكس نسمي المعاينة بدون إرجاع إذا لم يكن السحب بدون إرجاع. فلو أردنا سحب عينة من المجتمع حجمها n وكان حجم المجتمع N وكان السحب بالارجاع فإن عدد الطرق الممكنة لسحب عينة هو N^n. أما إذا كان السحب بدون ارجاع فعدد الطرق الممكنة هو $C_N^n = \frac{N!}{(N-n)!\,n!}$.

1-4- العينات العشوائية Random Sample

يتم اختيار أفراد العينة من المجتمع بطريقة غير متحيزة بحيث نضمن لكل فرد من المجتمع نفس الإمكانية في الظهور في العينة، وهذا يضمن إمكانية إخضاع هذا النوع من العينات للقوانين الاحتمالية. ومن أنواع العينات العشوائية نجد:

أ- العينات العشوائية البسيطة Simple Random Sample

وهي من أبسط أنواع العينات العشوائية، وتستخدم في حالة تجانس أفراد المجتمع محل الدراسة في الظاهرة المدروسة ومعرفة جميع أفراده. توجد عدة طرق لاختيار الأفراد منها:

الطريقة الأولى: من خلال إجراء القرعة، إذ يتم ترقيمهم ثم اختيار الأفراد بسحب أرقام الأفراد الداخلين في العينة بطريقة غير متحيزة وذلك بكتابة الأرقام على أوراق متشابهة وخلطها جيدا ثم سحب العدد المطلوب.

الطريقة الثانية: وتعتبر أفضل من الطريقة الأولى وهي استخدام جداول الأرقام العشوائية ولا سيما حين يكون حجم العينة كبيرا. ويتكون جدول الأرقام العشوائية من مجموعة الأعداد المكونة من خمس خانات مرتبة في صفوف وأعمدة، وتحتوي على مجموعة الأرقام من 0 إلى 9 بنسب متساوية، وأنّ اختيار رقم من هذا الجدول يكافئ سحب ورقة بشكل عشوائي من مجموعة الأوراق المخلوطة جيدا التي تحمل الأرقام من 0 إلى 9، والتمرين التالي يوضح طريقة استخدام جداول الأرقام العشوائية.

مثال (1-1): نريد اختيار عينة عشوائية بسيطة مكونة من 10 أشخاص من الذين زاروا المعرض الدولي خلال السنة الماضية. فإذا افترضنا أن عددهم 4000 شخص، يحمل كل اسم واحد منهم رقما من 1 إلى 1000.

تبدأ العملية باختيار صفحة من صفحات جداول الأرقام العشوائية، ومن تلك الصفحة المختارة نحدد بشكل عشوائي سطرا وعمودا. نبدأ في عملية الاختيار من العدد الذي تقاطع عنده السطر مع العمود المختارين، ونختار منه أول أربعة خانات من اليسار (لأن العدد 4000 مكون من أربعة خانات)، ثم ننتقل إلى عدد ثان وثالث بترتيب نحدده مسبقا سواء إلى اليمين أو اليسار أو الأعلى أو الأسفل، ونسجل قائمة الأعداد التي تم اختيارها، ثم نحذف كل عدد يزيد عن 4000.

بإفتراض أن رقم السطر الذي اخترناه هو 10 ورقم العمود هو 3 نجد العدد 03811، نأخذ منه أول أربع أرقام 3811 وبتحديد جهة الانتقال إلى اليسار مثلا وتابعنا في قراءة الأعداد سنسجل (3811، 8342، 7863، 2743، 1547، 8250، 8140، 8470، 4364، 9797، 3498، 5837، 8821، 6426، 0496، 4843، 8360، 1252، 9134، 8931، 9538، 1160، 9411، 4659، 8914...). فتكون العينة مكونة من الأشخاص الذين أرقامهم (3811، 2743، 1547، 3498، 0496، 1252، 1160، 2707، 2026، 1767)، وذلك بعد إهمال الأعداد التي تزيد عن 4000.

الطريقة الثالثة: باستخدام البرمجيات، حيث توجد العديد من البرمجيات التي تسمح لنا آنيا بالحصول على الأعداد العشوائية، حيث يكفي إدخال بعض المعلومات لها، والنقر للحصول على العينة العشوائية المرغوبة. من هذه البرمجيات نجد ... :، Minitab، SPSS، Excel

ب-العينة العشوائية المنتظمة Systematic Random Sample

تستخدم هذه العينة عندما يكون المجتمع متجانسا ومرتبا وفقا لصفة معينة، ونريد أن نختار فردا من كل عدد محدد من الأفراد (على سبيل المثال 8). فنختار الفرد الأول بطريقة عشوائية، فإذا افترضنا أن رقمه 4 نختار الفرد الثاني الذي رقمه (8+4=12) والفرد الثالث الذي رقمه (12+8=20) وهكذا إلى غاية اكتمال العدد 8 المطلوب، ويعتمد اختيار كل فرد على حجم العينة وحجم المجتمع.

ترجع تسميتها بالعينة العشوائية المنتظمة لتميزها بانتظام الفترة بين المفردات المختارة، حيث ينتج عن العينة العشوائية المنتظمة توزيع منتظم لأفراد العينة.

ت-العينة العشوائية الطبقية Stratified Random Sample

تستخدم في حالة المجتمعات غير المتجانسة، وتعتمد هذه العملية على تقسيم المجتمع إلى مجموعات متجانسة تسمى طبقات، ومن ثم إختيار عينة عشوائية بسيطة من كل طبقة. فعلى سبيل المثال إذا درسنا متوسط إنفاق الأسر الجزائرية على المشروبات الغازية في شهر رمضان، نتوقع تجانس معدلات الإنفاق ضمن كل ولاية، في حين تختلف فيما بين الولايات. لذا يتم توزيع الأسر على طبقات حسب الولايات، ونختار عينة عشوائية بسيطة من كل طبقة (ولاية) حجمها يتناسب مع حجم أسرها نسبة إلى حجم السكان في الجزائر، وندرس متوسط الإنفاق للأسر التي تم اختيارها بغية ضمان تمثيل الطبقات الصغيرة.

تجدر الإشارة الى أنه توجد ثلاثة طرق لتحديد حجم العينة: (1) التوزيع المتساوي، (2) التوزيع المتناسب، (3) التوزيع الأمثل.

ث-العينة العشوائية العنقودية Cluster Random Sample

وهي الطريقة التي بموجبها يقسم المجتمع إلى مجموعات صغيرة منفصلة عن بعضها البعض تسمى عناقيد، بحيث أن كل عنصر أو فرد من المجتمع ينتمي إلى عنقود وعنقود واحد. بعد هذه العملية يتم اختيار عينة من العناقيد على أساس المعاينة العشوائية البسيطة. فعلى سبيل المثال إذا أردنا معرفة مستوى الرفاه لسكان ولاية ما من ولايات الجزائر، نقوم بتكوين عناقيد بالطريقة التالية: تقسيم الولاية إلى مناطق ثم تقسيم المناطق إلى أحياء وباستعمال المعاينة العشوائية البسيطة نقوم بإختيار بعض الأحياء. حيث تكون العينة النهائية مكونة من جميع الأفراد المنتمين للعناقيد المختارة.

يتميز أسلوب العينة العشوائية العنقودية عن أسلوب العينة الطبقية في كون أفراد العنقود غير متجانسين وأن كل عنقود هو صورة مصغرة للمجتمع الأصلي.

1-5- العينات غير العشوائية Nonrandom Samples

وهي تلك العينات التي لا تكفل لجميع مفردات المجتمع احتمال ثابت ومحدد للاختيار، وغالبا يتدخل الباحث في عملية الاختيار بصورة أو بأخرى ما يعرض البيانات الإحصائية عند جمعها الى نوعين من الأخطاء:

- **خطأ التحيز**: وهو خطأ ناتج عن مصادر متعددة، منها أخطاء في تصميم البحث، التجربة، أخطاء أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المتجمعة.
- **خطأ المعاينة**: يطلق عليه أحيانا خطأ الصدفة وهو الخطأ الناتج عن فروق الصدفة بين مفردات المجتمع التي دخلت العينة وبين تلك المفردات التي لم تشأ الصدفة أن تدخل العينة.

تجد الإشارة الى أن هذا النوع من العينات تكون عينات متحيزة ولا يمكن من خلالها تعميم نتائجها على المجتمع، من بين هذه العينات نجد:

أ- عينة الصدفة أو المصادفة Accidental Sample

عينة الصدفة تعني ليس للباحث أثر في اختيارها. فعلى سبيل المثال، إذا أجرى الطبيب دراسة على مجموعة من المرضى يزورونه تكون هذه المجموعة عبارة عن عينة اختارتها الصدفة وليس للطبيب دخل في اختيارها.

ب- العينة الحصصية Quota Sample

هنا يقوم الباحث بتقسيم المجتمع إلى طبقات، ثم يختار من كل طبقة فئة صغيرة ممثلة له يختارها حسب معيار ما وليس عشوائيا. ويمكن الاختلاف بين أسلوب العينة العشوائية الطبقية والعينة الحصصية في أسلوب إختيار أفراد العينة في كل طبقة فقط.

ت- العينة العمدية (القصدية) Purposive Sample

في هذا النوع من العينات يتم تحديد مفردات العينة بواسطة الباحث حسب ما يراه مناسبا لتحقيق هدف معين. فقد يختار أحد المخترعين(مهندس، طبيب أو إقتصادي) في مجال إختصاصه مجموعة معينة من الباحثين يختارهم كما يريد لمعرفة رأيهم في اختراعه الجديد(محرك، دواء أو نظرية).

1-6- معالم المجتمع Parameters

يقصد بمعالم المجتمع مجموعة من خصائصه مثل المتوسط، التباين، النسبة،...، يسمى أي مقياس إحصائي تم حسابه للمجتمع بالمعلمة (Parameter) وعادة ما يرمز لمعالم المجتمع بالحروف الإغريقية، فالمتوسط الحسابي في المجتمع يرمز له بالرمز μ (ميو)، والتباين في المجتمع بالرمز σ^2 (سيجما مربع)، والنسبة في المجتمع بالرمز **P**. ومن خصائص المجتمع أيضا توزيعه الإحتمالي.

1-7- إحصائيات العينة Sample statistic

يسمى أي مقياس إحصائي تم حسابه للعينة بـ "إحصائية" statistic وعادة ما يرمز للمتوسط الحسابي في العينة بالرمز $\bar{X}$، وللتباين في العينة بـ S^2 أو $\sigma^2_{\bar{x}}$. وتجدر الإشارة إلى أن إحصائيات العينة تكون دائما عبارة عن متغير عشوائي (random variable) عكس معالم المجتمع التي تكون دائما ثابتة.

1-8- خطأ المعاينة (الخطأ العشوائي) Sampling error

عند إختيار عينة عشوائية من المجتمع قد تكون إحصائية العينة مساوية لمعلمة المجتمع، أو قد تكون أصغر أو أكبر منها. يسمى الفرق بين إحصائية العينة ومعلمة المجتمع بخطأ المعاينة أو الخطأ العشوائي.

ثانيا: بعض التوزيعات الاحتمالية الهامة في نظرية المعاينة

في نظرية العينات، تلعب التوزيعات الإحتمالية دورا كبيرا في التطبيقات الإحصائية. ويمكن تقسيم التوزيعات الإحصائية إلى مجموعتين:

- التوزيعات الاحتمالية المتقطعة أو المنفصلة.
- التوزيعات الاحتمالية المستمرة.

سوف نكتفي هنا بالإشارة إلى بعض التوزيعات الهامة التي لها أهمية خاصة في نظرية العينات (ويمكن للطالب الرجوع إلى كتب الإحصاء للتعرف على هذه التوزيعات وخصائصها أكثر)، وهي: توزيع ذي الحدين، التوزيع الطبيعي، توزيع ستودنت، توزيع مربع كاي وتوزيع فيشر.

2-1- توزيع ذي الحدين Binomial Distribution :

تم اكتشاف هذا التوزيع في نهاية القرن السابع عشر من طرف العالم الرياضي الشهير برنولي **(Bernouli)**. تقوم فكرة هذا التوزيع على كيفية حساب الاحتمالات الخاصة بالتجارب التي تكون نتيجتها إما النجاح أو الفشل أي أنه عبارة عن تجربة برنولي مكررة n مرة. فإذا قمنا بإجراء تجربة معينة وكانت نتيجتها هي النجاح أو وقوع الحادث المعين باحتمال قدره **P** والفشل (عدم وقوع الحادث) باحتمال **(1-P = q)**، فإن هذه

التجربة تسمى بتجربة أو توزيع برنولي. وإذا قمنا بتكرار هذه التجربة **n** مرة فإننا نحصل على ما يعرف بتوزيع تجارب برنولي المتكررة أو توزيع ذي الحدين **(Binomial)**. وفي هذه الحالة إذا كان المتغير العشوائي **X** يمثل عدد مرات النجاح عند إجراء **n** مرة من التجارب فإن المتغير العشوائي **X** يأخذ القيمة **x**، حيث عدد مرات النجاح **x** تأخذ القيم **x = 1, 2, 3…, n** (أما عدد مرات الفشل فيأخذ القيم **x-n**) باحتمال قدره **P(X = x)** حيث:

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

أي:

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

وتعتمد دالة الاحتمال لتوزيع ذي الحدين على معلمتين هما **n** و **P** ويرمز لها اختصارا بـ $X\sim B(n,P)$. ويمكن حساب القيمة المتوقعة والتباين لهذا التوزيع بحيث:

القيمة المتوقعة $$E(x)=n.P$$

التباين $$\sigma_x^2=n.P.(1-P)=n.p.q$$

وتجدر الإشارة إلى أن توزيع ذي الحدين له استخدامات عديدة في الحياة العملية مما يجعله من بين أكثر التوزيعات الإحصائية أهمية ومن أكثرها شيوعا، ومن هذه التجارب نجد على سبيل المثال نسبة البطالة في المجتمع، نسبة الأمية في المجتمع، نسبة الإنتاج المعيب في مصنع، نجاح أو فشل الطلبة في الامتحانات...الخ.

مثال (1- 2):

إذا كان معدل نجاح عقار جديد لمعالجة داء الصرع يساوي **0.8**، وقد تم تجريبه على **20** شخص مريض بداء الصرع، المطلوب حساب احتمال:

- شفاء **18** مريضا منهم.
- شفاء **18** مريضا منهم على الأقل.

الحل:

عند إستعمال العقار الجديد نحن أمام تجربة ثنائية تتمثل في إستجابة المريض أو عدم إستجابته، وتم تكرار هذه التجربة بـ **20** مرة بطريقة مستقلة عن بعضها البعض، وبالتالي الظاهرة المدروسة تتبع توزيع ثنائي الحدين (أو توزيع Bernouli مكرر بـ **20** مرة) بالمعالم الاتية: **n = 20** و **P =0.8**

أي:

$$X \sim B(n,\ P)$$
$$X \sim B(20,\ 0.8)$$

❖ احتمال شفاء **18** مريضا من بين **20** مريض يعني:

$$P(X=18)=\frac{20!}{18!\,2!}(0.8)^{18}(0.2)^{3}=\frac{20.19.18!}{18!\,2!}(0.8)^{18}(0.2)^{2}=0.1369$$

❖ احتمال شفاء **18** مريضا منهم على الأقل يعني:

$$P(X\geq 18)\ =\ P(X=18)\ +\ P(X+19)\ +\ P(X=20)$$
$$=\frac{20!}{18!\,2!}(0.8)^{18}(0.2)^{2}+\frac{20!}{19!\,1!}(0.8)^{19}(0.2)^{1}+\frac{20!}{20!\,0!}(0.8)^{20}(0.2)^{0}=0.2061$$

2-2- التوزيع الطبيعي Normal Distribution

يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية وأكثرها استخداما، ذلك أن التوزيع الطبيعي يستخدم بكثرة في وصف الكثير من الظواهر الطبيعية والاقتصادية والاجتماعية (المناح، الفلك والطب...) التي نصادفها في حياتنا اليومية بالإضافة إلى أنه يمكن تقريب العديد من التوزيعات الإحتمالية له وأيضا إشتقاق البعض الاخر منه. ومن خصائصه أن المنحنى البياني له شكل جرسي (متماثل)، أو كما نعبر عنه إحصائيا شكل التوزيع الطبيعي متناظر.

فإذا كان X متغير عشوائي يتبع التوزيع الطبيعي بمتوسط حسابي μ وتباين σ^2 فإن دالة الكثافة الاحتمالية له تكون بالشكل التالي:

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

وإختصارا يكتب التوزيع الطبيعي بالعلاقة التالية:

$$X \sim N(\mu, \sigma^2)$$

وتقرأ: X تتبع التوزيع الطبيعي بمتوسط حسابي μ وتباين σ^2.

تعطى دالة التوزيع المتراكم للتوزيع الطبيعي كما يلي:

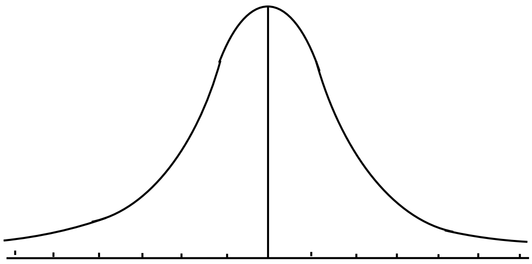
وهي دالة التوزيع المتراكم او ما يعرف بـ **fonction de répartition** $P(X\leq x)=F(x)=\int_{-\infty}^{x}f(x)dx$

وبحساب التكامل ينتج:

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \sigma\sqrt{2\pi}$$

حيث المساحة تحت منحنى دالة التوزيع المتراكم للتوزيع الطبيعي تساوي الواحد تماما.

الرسم التالي يوضح منحنى التوزيع الطبيعي:



من خلال هذا المنحنى يتضح أن التوزيع متماثل (symmetric) حول المتوسط الحسابي **μ** وأن أكثر المشاهدات تقع حول المتوسط الحسابي وأقلها في الطرفين.

ونظرا لصعوبة حساب التكامل السابق تم عمل جداول جاهزة لحساب الاحتمالات عند قيم مختلفة لـ **μ** و $\boldsymbol{\sigma^2}$ ، وذلك بعد تحويل أي توزيع طبيعي إلى توزيع طبيعي معياري باستخدام القيمة المعيارية Z.

نظرية (1-1): إذا كان المتغير العشوائي X يتبع التوزيع الطبيعي بمتوسط **μ** و تباين $\boldsymbol{\sigma^2}$، أي: $X\sim N(\mu,\sigma^2)$

فإن المتغير العشوائي **Z** يتبع التوزيع الطبيعي المعياري بمتوسط $\mu = 0$ وتباين $\sigma^2 = 1$، أي :

$.Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

حسب النظرية **السابقة**، تكتب دالة الكثافة الاحتمالية لـ **Z** كما يلي:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad ,(-\infty < z < +\infty)$$

في حين دالة التوزيع المتراكم تكون بالصيغة الأتية:

$$P(Z \le z) = F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}} dz$$

نرمز لدالة الكثافة الاحتمالية في هذه الحالة بـ $\varphi(z)$ بدلا من **f(z)** ولدالة التوزيع بـ $\emptyset(z)$ بدلا من **F(z)**. وبالتالي إذا كان $Z\sim N(0,1)$ فإن كثافته $\varphi(z)$ ودالة توزيعه $\emptyset(z)$.

إن حساب قيم دالة التوزيع للمتغير الطبيعي المعياري يعتمد على جدول يعطي قيم دالة التوزيع $\emptyset(\mathbf{z})$ ابتداء من الصفر وبفاصل 0.01 بين كل قيمة والقيمة التي تليها (أنظر الملحق 2). وقد وضعت قيم z ابتداء من 3.4- وبفاصل 0.1 في العمود الأيسر، ووضعت القيم العشرية الثانية من قيم z في السطر الرأسي. أما القيم الاحتمالية و التي تمثل المساحات (أي قيم الدالة $\emptyset(\mathbf{z})$) فقد وضعت في صلب الجدول.

ملاحظة:

يمكن حساب الاحتمال للقيم السالبة بإستعمال الخاصية : $\emptyset(-z) = 1 - \emptyset(\mathbf{z})$

لتوضيح استخدام جدول التوزيع الطبيعي المعياري نأخذ المثال التالي:

مثال (3-1):

لتكن المتغيرة العشوائية X تتبع التوزيع الطبيعي بمتوسط $\mathbf{\mu = 5}$ وتباين $\boldsymbol{\sigma^2 = 9}$ أي: $\boldsymbol{X \sim N(5,9)}$

المطلوب: حساب الاحتمالات التالية

$\boldsymbol{P(X<6)}$ ، $\boldsymbol{P(4.21<X<5.05)}$ ، $\boldsymbol{P(Z<-2.47)}$ ، $\boldsymbol{P(Z<1.23)}$ ، $\boldsymbol{P(1.13<Z<1.89)}$

الحل: هنا نميز بين حالتين حسب المتغير X أو Z

- بالنسبة للمتغير العشوائي X، يجب أن نحول المتغير العشوائي X إلى القيمة Z المعيارية كما يلي:

$$\boldsymbol{X \sim N(\mu,\sigma^2) \Rightarrow Z = \frac{X-\mu}{\sqrt{\sigma^2}} = \frac{X-5}{\sqrt{9}} \sim N(0,1)}$$

ومنه:

$$P(X \prec 6) = P(\frac{X-\mu}{\delta} \prec \frac{6-5}{3}) = P(Z \prec 1/3) = \phi(0.33) = 0.6293$$

حيث يتم استخراج القيمة من جدول التوزيع الطبيعي المعياري (الملحق **2**)، وتمثل تقاطع السطر **0.3** مع العمود **0.03**(وذلك لأن العدد 0.33 عبارة عن مجموع القيمتين 0.3 و 0.03، حيث نقطة تقاطع السطر مع العمود تمثل قيمة الإحتمال).

$$P(4.21 \prec X \prec 5.05) = P(\frac{4.21-5}{3} \prec \frac{X-\mu}{\delta} \prec \frac{5.05-5}{3}) = P(-0.79/3 \prec Z \prec 0.5/3) = \phi(0.17) - \phi(-0.26)$$
$$= \phi(0.17) - \left(1 - \phi(0.26)\right) = 0.5675 - 1 + 0.6026 = 0.1701$$

يتم استخراج القيم من جدول التوزيع الطبيعي المعياري بالنسبة للقيم الموجبة مباشرة مثل: $\phi(0.17) = 0.5675$ وهي تمثل تقاطع السطر **0,10** مع العمود **0,07**. أما القيمة $\phi(-0.26)$ فيتم استخراجها أيضا من جدول التوزيع الطبيعي المعياري ولكن للقيم السالبة وهي تمثل تقاطع السطر **0,2−** مع العمود **0,06-** **أو باستخدام الخاصية التالية:** $\emptyset(-0,26) = 1 - \emptyset(0,26)$.

- بالنسبة لقيم الاحتمال الخاصة بالمتغير العشوائي **Z**، فتستخرج بشكل مباشر باستخدام جدول التوزيع الطبيعي المعياري (الملحق **2**) للقيم السالبة أو باستخدام الخاصية سابقة الذكر حيث نجد:

$$P(Z \prec -2.47) = 0.0068$$

وهذه القيمة هي عبارة عن تقاطع السطر 2.4- مع العمود 0.07- (بالنسبة لجدول القيم السالبة).

وبنفس الطريقة نجد:

$$P(Z \prec 1.23) = 0.8907$$

$$P(1.13 \prec Z \prec 1.89) = \phi(1.89) - \phi(1.13) = 0.9706 - 0.8708 = 0.0998$$

كما يمكننا أيضا استخدام جدول دالة التوزيع الطبيعي المعياري بشكل عكسي، كما هو موضح في التمرين التالي:

مثال (4-1):

إذا علمت أن: $P(Z < a) = 0.9788$ و $P(Z \succ b) = 0.2546$، أوجد قيمة كل من **a** و **b**.

الحل:

- باستعمال جدول التوزيع الطبيعي المعياري نجد أن القيمة 2.03 تقابل قيمة الاحتمال 0.9788، أي أن قيمة **a = 2.03** (*العمود 4 والسطر 19*) .
- بالنسبة لقيمة ***b*** نحوّل أولا العلاقة من أكبر إلى أصغر كما يلي:

$$P(Z \succ b) = 0.2546 \Leftrightarrow P(Z \prec b) = 1 - 0.2546 = 0.7454$$

بالبحث في جدول التوزيع الطبيعي المعياري عن قيمة الاحتمال 0.7454، نجد أن b = 0.66 وهي تقع بين العمود 7 والسطر 7 وبالتالي قيمة ***b*** تساوي 0.66 .

2-3- توزيع مربع كاي χ^2

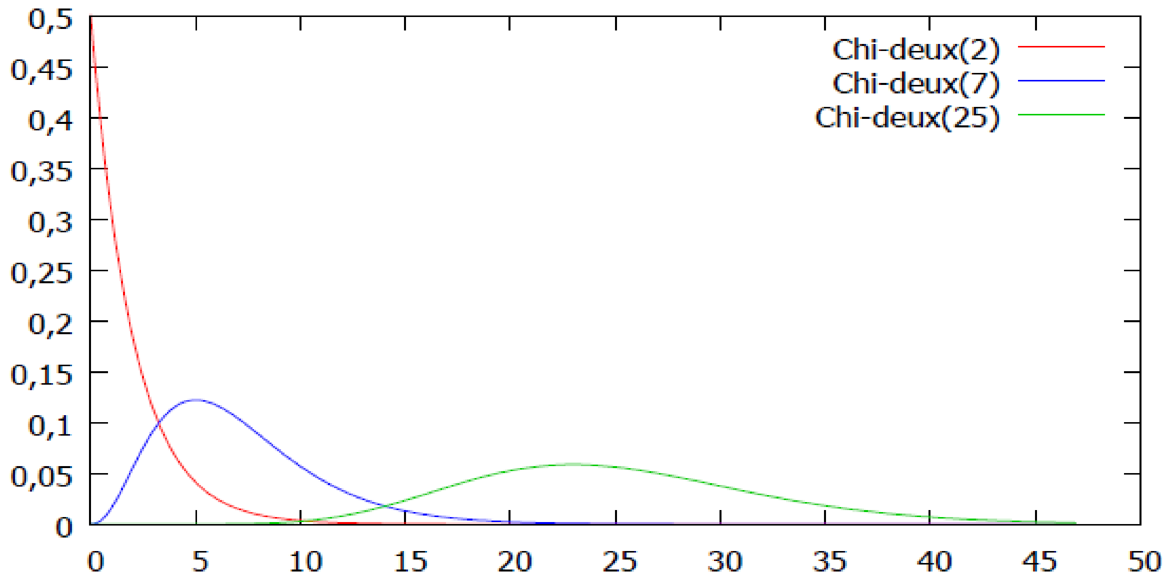
نقول عن المتغير العشوائي X أنه يتبع توزيع مربع كاي χ^2 بدرجة حرية ϑ (**vartheta**) إذا كانت دالة كثافته الاحتمالية على الشكل التالي:

$$f(x) = \frac{1}{\Gamma\left(\frac{\vartheta}{2}\right)(2)^{\frac{\vartheta}{2}}}.x^{\frac{\vartheta}{2}-1}.e^{-\frac{x}{2}} \qquad ,x \geq 0$$

توزيع مربع كاي يعتبر كحالة خاصة من توزيع جاما (*Gamma*) بالمعلمتين: $\alpha = \frac{\vartheta}{2}$ و $\beta = 2$.

ودرجة الحرية ϑ تدل على معلمة توزيع مربع كاي، ونكتب: $X \sim \chi^2_{(\vartheta)}$

الشكل البياني التالي يوضح دالة التوزيع لمربع كاي عند قيم مختلفة لـ ϑ.



يتضح من التمثيل البياني الخاص بمربع كاي أنّ الالتواء موجب ويقترب من التماثل كلما كبرت درجة الحرية.

بعبارة أخرى، كلما كان حجم العينة كبيرا $n \geq 30$ يقترب التوزيع من التوزيع الطبيعي.

سنذكر هنا بعض النظريات الهامة الخاصة بهذا التوزيع التي سنستخدمها في تطبيقاتنا ولكن بدون برهان.

نظرية (2-1): إذا كان المتغير العشوائي $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ فإن $Z^2 = (\frac{X-\mu}{\sigma})^2 \sim \chi^2_{(1)}$.

نظرية (3-1): إذا كانت $X_1, X_2, \dots, X_n$ متغيرات عشوائية مستقلة وكل متغيرة تتبع التوزيع الطبيعي المعياري $N(0, 1)$، فإن: $\sum_{i=1}^{n} X_i^2 \sim \chi^2_n$.

نظرية (4-1): لتكن $U_1, U_2, \dots, U_n$ متغيرات عشوائية مستقلة تتبع توزيع مربع كاي بدرجات حرية $\vartheta_1, \vartheta_2, \dots, \vartheta_n$ على التوالي، فإن: $\sum_{i=1}^{n} U_i \sim \chi^2_{(\vartheta)}$.

نظرية (5-1): إذا كانت مجموعة من العينات العشوائية ذات الحجم n وتباين S^2 من مجتمع يتوزع طبيعيا بمتوسط حسابي μ و تباين σ^2، فإن: $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$. **حيث:** $S^2 = \frac{\sum(X_i-\bar{X})^2}{n-1}$.

حيث قيمة المتغير العشوائي الذي يتبع توزيع مربع كاي يعبر عنها بالرمز $\chi^2_{(\alpha,\vartheta)}$ وهي القيمة التي يقع على يمينها مساحة α عن منحنى مربع كاي بدرجة حرية ϑ ويتم حساب قيم هذا المتغير العشوائي والاحتمالات المرافقة له باللجوء الى جدول مربع كاي الموضح في الملحق 3، ولتوضح كيفية استخدام الجدول الخاص بالتوزيع نأخذ المثال الأتي:

مثال (5-1): أوجد القيم $\chi^2_{(0.025,20)}$، $\chi^2_{(0.05,5)}$، $\chi^2_{(0.999,7)}$

الحل:

لإيجاد قيمة $\chi^2_{(0.025,20)}$ من السطر العلوي الخاص بالمساحات والتي تمثل الاحتمالات نختار القيمة **0.025** ، ونختار من العمود الأيسر الخاص بدرجات الحرية ϑ القيمة **20**. وبتقاطع السطر مع العمود نحصل على القيمة المطلوبة، وتكون $\chi^2_{(0.025,20)} = 34.17$.

نتبع نفس الخطوات السابقة للحصول على $\chi^2_{(0.05,5)} = 16.75$ وأيضا $\chi^2_{(0.999,7)} = 1.69$.

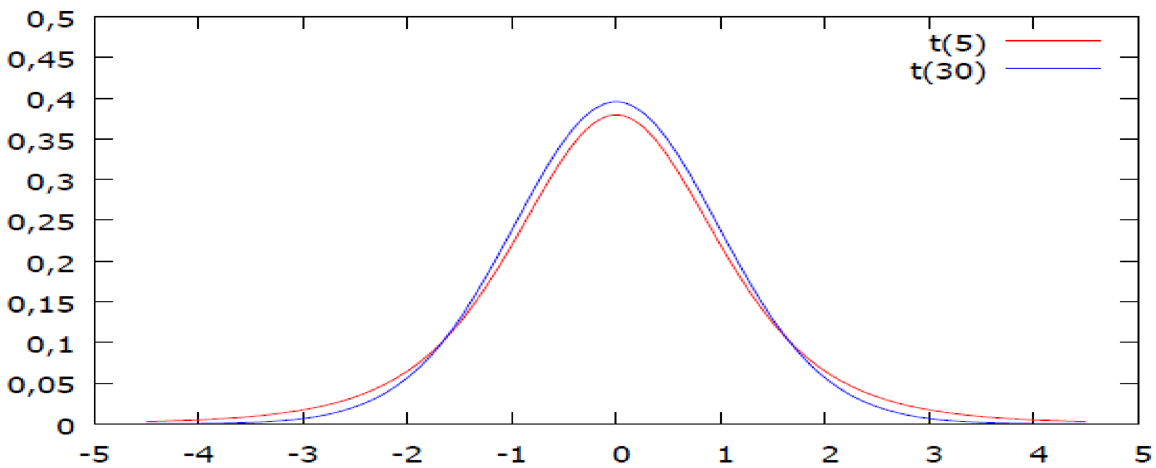
2-4- توزيع ستودنت t

إذا كانت دالة الكثافة الاحتمالية للمتغير العشوائي t على الصورة:

$$f(t) = \frac{\Gamma\left(\frac{\vartheta+1}{2}\right)}{\Gamma\left(\frac{\vartheta}{2}\right)\sqrt{\vartheta\pi}}\left(1+\frac{t^2}{\vartheta}\right)^{-\frac{n+1}{2}} \qquad , -\infty < t < +\infty$$

فإن هذا التوزيع يسمى توزيع **(Student t-distribution)** يرمز له إختصارا بـ **t** حيث: ϑ تمثل درجات الحرية وهي في نفس الوقت معلمة التوزيع، وتظهر أهمية توزيع **t** عند سحب العينات الصغيرة الحجم $(n < 30)$ من مجتمع مجهول التباين.

يتشابه التوزيع **t** مع التوزيع الطبيعي القياسي من حيث الشكل الجرسي إلاّ أنه أكثر انخفاضا منه، وعند تزايد درجات الحرية فإن التوزيع **t** يقترب من التوزيع الطبيعي القياسي. والشكل التالي يوضح ذلك.



من بين النظريات الهامة والمتعلقة بالتوزيع t والتي نتطرق إليها دون البرهان عليها:

نظرية (1-6): إذا كان المتغيران العشوائيان المستقلان Y و Z، حيث أن Y يتوزع توزيعا طبيعيا قياسيا $Y \sim N(0,1)$ و Z يتوزع توزيع مربع كاي $Z \sim \chi^2_\vartheta$ فإن المتغير العشوائي $T = \frac{Y}{\sqrt{Z/\vartheta}} \sim t_\vartheta$.

نظرية (1-7): إذا أخذت مجموعة من العينات العشوائية صغيرة الحجم n وبتباين S^2 من مجتمع يتوزع طبيعيا بمتوسط حسابي μ وتباين σ^2، فإن: $T = \frac{\bar{X}-\mu}{\sqrt{s^2/n}} \sim t(n-1)$

الرمز $t_{(\alpha,\vartheta)}$ يستعمل للتعبير عن قيمة المتغير العشوائي، وهي القيمة التي يقع على يمينها مساحة α على منحنى توزيع t بدرجة حرية ϑ. يمكننا حساب الاحتمالات المرافقة لقيم هذا التوزيع باستخدام جدول خاص يسمى جدول توزيع ستودنت الموضح بالملحق 4.

خاصية:

التوزيع t متماثل حول الصفر، لذا يمكننا الاعتماد على الخاصية: $t_{(\alpha,\vartheta)} = -t_{(1-\alpha,\vartheta)}$، **وذلك لإختلاف الجداول الإحصائية.**

مثال (1-6): أوجد قيم t في الحالات التالية: $t_{(0.25,14)}$، $t_{(0.99,7)}$، $t_{(0.995,10)}$

الحل:

لإيجاد قيمة $t_{(0.25,14)}$ نختار من العمود الأيسر الخاص بدرجات الحرية القيمة **14** ومن السطر العلوي الخاص بالاحتمالات القيمة **0.25**. وبتقاطع السطر مع العمود نحصل على القيمة المطلوبة، وبالتالي يكون

$t_{(0.25,14)}=2.145$

لإيجاد قيمة $t_{(0.95,7)}$ نلاحظ أن قيمة $\alpha = 0.99$ غير متوفرة في جدول توزيع ستودنت، وبذلك سوف نستعمل تماثل المنحنى حول الصفر فنحصل على:

$$t_{(0.99,7)} = -t_{(0.01,7)} = -2.998$$

وبنفس الطريقة نستخرج قيمة $t_{(0.995,10)}$ حيث:

$$t_{(0.995,10)} = -t_{(1-0.995,10)} = -t_{(0.005,10)} = -3.169$$

2-5- توزيع فيشر (F)

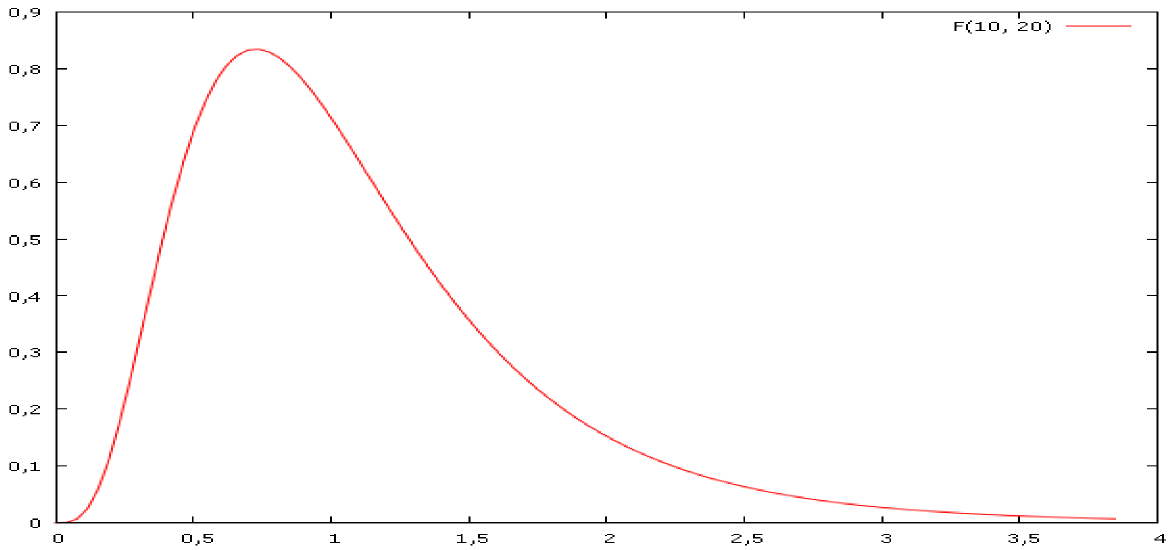
إذا كان المتغيران العشوائيان المستقلان X_1 و X_2 يتوزعان توزيع مربع كاي بدرجتي حرية ϑ_1 و ϑ_2 على التوالي ($X_1 \sim \chi^2_{\vartheta_1}$ و $X_2 \sim \chi^2_{\vartheta_2}$) فإن المتغير X المعرف بالصيغة الآتية: $X = \frac{X_1/\vartheta_1}{X_2/\vartheta_2}$ له دالة الكثافة الاحتمالية التالية:

$$g(X) = \frac{\Gamma\left(\frac{\vartheta_1+\vartheta_2}{2}\right)}{\Gamma\left(\frac{\vartheta_1}{2}\right)\Gamma\left(\frac{\vartheta_2}{2}\right)} (\vartheta_1)^{\frac{\vartheta_1}{2}} (\vartheta_2)^{\frac{\vartheta_2}{2}} X^{\left(\frac{\vartheta_1}{2}\right)-1} (\vartheta_1+\vartheta_2 X)^{\frac{\vartheta_1+\vartheta_2}{2}} \quad , X > 0$$

وعليه نقول بأن المتغير X يتوزع توزيع فيشر بدرجة حرية ϑ_1 و ϑ_2 ويعبّر عنه بالشكل التالي:

$$X \sim F_{(\vartheta_1,\vartheta_2)}$$

ويقترب شكل توزيع فيشر من الشكل التالي:



ونعبّر بالرمز $F_{(\alpha,\vartheta_1,\vartheta_2)}$ عن قيمة المتغير العشوائي الذي يتبع توزيع فيشر، وهي تمثل القيمة التي يقع على يمينها مساحة α على منحنى توزيع F بدرجة حرية ϑ_1 في البسط و ϑ_2 في المقام.

والجدول (5) الموضح في الملاحق حسب قيم α : $\{0.01 ، 0.025 ، 0.05 ، 0.10 ، 0.25\}$، خاص بقيم المتغير العشوائي الذي يتبع توزيع فيشر $F_{(\alpha,\vartheta_1,\vartheta_2)}$.

مثال (1- 7): أوجد القيم التالية: $F_{(0.05,7,12)}$، $F_{(0.025,8,9)}$، $F_{(0.10,20,10)}$، $F_{(0.01,3,12)}$

الحل:

لإيجاد قيمة $F_{(0.05,7,12)}$ نستخدم جدول توزيع فيشر الخاص بقيمة $\alpha = 0.05$، ومن السطر في أعلى الجدول الخاص بدرجات حرية البسط نختار القيمة 7 ومن العمود الخاص بدرجات حرية المقام نختار القيمة 12.

وبتقاطع السطر مع العمود نجد أن قيمة $F_{(0.05,7,12)} = 2.91$.

وباستخدام نفس الطريقة وبتغيير قيم α نحصل على $F_{(0.025,8,9)} = 4.1$ و $F_{(0.10,20,10)} = 2.2$ و $F_{(0.01,3,12)} = 5.95$

وفي حالة عدم توفر بعض القيم في جداول فيشر لإحدى قيم α:

$$\{0.01 ، 0.025 ، 0.05 ، 0.10 ، 0.25\}$$

ومن أجل استخراج قيم المتغير العشوائي في هذه الحالة، يمكننا الإستعانة بالنظرية التالية:

نظرية (1-8): بفرض أن $n_1, n_2 > 0$ فإن:

$$F(1-\alpha, n_1-1, n_2-1) = \frac{1}{F(\alpha, n_2-1, n_1-1)}$$

مثال (1- 8): أوجد القيم الأتية: $F_{(0.99,4,5)}$، $F_{(0.95,5,4)}$

الحل:

في مثالنا هذا قيم $\alpha = 0.99$، $\alpha = 0.95$ في جداول فيشر غير متوفرة ، لذلك سنستعين بتطبيق النظرية (1-8) كما يلي:

$$F_{(0.99,4,5)} = \frac{1}{F(0.01,5,4)} = \frac{1}{15.52} = 0.064$$

$$F_{(0.95,5,4)} = \frac{1}{F(0.05,4,5)} = \frac{1}{5.19} = 0.192$$

ثالثا: توزيعات المعاينة Sampling Distributions

عند اختيار عينة ذات الحجم n من مجتمع ما، وحساب إحصائية لها مثل المتوسط الحسابي أو التباين...، فإن قيم هذه الأخيرة سوف تختلف من عينة إلى أخرى حيث يصبح عبارة عن متغير عشوائي يتبع توزيع معين، يعرف باسم توزيع المعاينة.

توزيع المعاينة للإحصائية هو عبارة عن التوزيع الإحتمالي لهذه الإحصائية المحسوبة لكل العينات الممكنة والتي لها نفس الحجم n **والمأخوذة من المجتمع الإحصائي قيد الدراسة مهما كان حجمه ومهما كانت طريقة السحب.**

3-1- توزيع المعاينة للمتوسطات Sampling Distribution of Means

3-1-1- توزيع المعاينة للمتوسط الحسابي $\bar{X}$ عند المعاينة من مجتمع طبيعي

إذا سحبنا عينة عشوائية وحسبنا متوسطها الحسابي سوف نجده قيمة ثابتة، أما إذا سحبنا عدة عينات (بعض العينات) بنفس الحجم من نفس المجتمع فمن المتوقع أن يأخذ المتوسط الحسابي قيما مختلفة في هذه العينات، أما إذا قمنا بسحب كل العينات الممكنة والتي لها نفس الحجم n وحسبنا المتوسطات الحسابية لهذه

العينات، فإن المتوسط الحسابي العام $\bar{X}$ لمتوسطات هذه العينات يجب أن يساوي تماما المتوسط الحسابي للمجتمع μ، ويسمى توزيع المتوسطات الحسابية بتوزيع المعاينة للمتوسط الحسابي.

فيما يلي سوف نذكّر بنظرية جدّ مهمة بدون برهان والمتعلقة بهذا التوزيع مع استعراض الحالات المختلفة وكيفية حساب الاحتمالات لهذا المتغير العشوائي.

نظرية (1- 9): توزيع المعاينة للمتوسط الحسابي $\bar{X}$ بالنسبة لعينات عشوائية ذات الحجم n مأخوذة من مجتمع يتوزع توزيعا طبيعيا متوسطه μ وتباينه $\sigma_{\bar{X}}^2$ يكون مساويا لمتوسط المجتمع μ (سواء كان السحب بالإرجاع أو بدون إرجاع)، أما تباينه $\sigma_{\bar{X}}^2$ فيكون:

في حالة السحب بالإرجاع (بالإعادة) $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$

في حالة السحب بدون إرجاع (بدون إعادة) $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \cdot \left(\frac{N-n}{N-1}\right)$

ملاحظات:

- N يمثل حجم المجتمع وقد يكون محدود الحجم (منته) أو غير محدود (غير منته)، فإذا كان حجم المجتمع محدود وكان السحب بالإرجاع، فيعتبر المجتمع هنا غير منته أو غير محدود الحجم.
- القيمة $\left(\frac{N-n}{N-1}\right)$ تسمى بمعامل التصحيح أو الإرجاع.
- بصفة عامة نهمل معامل الإرجاع في حالة تحقق الشرط الاتي:

$$\frac{n}{N} < 0.05$$

مثال (1- 9):

تم سحب عينة عشوائية حجمها 2 مشاهدة من مجتمع حجمه 3 وعناصره هي: $\{2, 4, 6\}$، والمطلوب أثبت ما يلي:

- $E(\bar{X}) = \mu$ في حالة السحب بالإرجاع ثم بدون إرجاع.
- $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$ في حالة السحب بالإرجاع.
- $\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} \cdot \left(\frac{N-n}{N-1}\right)$ في حالة السحب بدون إرجاع.

الحل:

في حالة السحب بالإرجاع

نحدد أولا عدد العينات التي يمكن سحبها من المجتمع ثم نحسب متوسط توزيع المعاينة للمتوسط الحسابي $\bar{X}$ وتباينه $\sigma^2_{\bar{X}}$.

عدد العينات التي يمكن سحبها في هذه الحالة هو $N^n = 3^2 = 9$. والجدول التالي يوضح مختلف العينات ومتوسطاتها وكذا احتمالاتها.

العينات الممكنة	$\bar{x}_i$	$P(\bar{x}_i)$	$\bar{x}_i P(\bar{x}_i)$	$\bar{x}_i^2 P(\bar{x}_i)$	$(\bar{x}_i - \mu_{\bar{X}})^2 P(\bar{x}_i)$
(2.2)	2	$1/9$	$2/9$	$4/9$	$4/9$
(2.4)	3	$1/9$	$3/9$	$9/9$	$1/9$
(2.6)	4	$1/9$	$4/9$	$16/9$	$0/9$
(4.2)	3	$1/9$	$3/9$	$9/9$	$1/9$
(4.4)	4	$1/9$	$4/9$	$16/9$	$0/9$
(4.6)	5	$1/9$	$5/9$	$25/9$	$1/9$
(6.2)	4	$1/9$	$4/9$	$16/9$	$0/9$
(6.4)	5	$1/9$	$5/9$	$25/9$	$1/9$
(6.6)	6	$1/9$	$6/9$	$36/9$	$4/9$
المجموع	-	1	$36/9 = 4$	$156/9 = 52/3$	$12/9 = 4/3$

وبالتالي:

متوسط المجتمع
$$\mu_X = \frac{2+4+6}{3} = 4$$

متوسط العينة
$$\mu_{\bar{X}} = E(\bar{X}) = \sum \bar{x}_i P(\bar{x}_i) = 4$$

تباين المجتمع
$$\sigma^2{}_X = \sum \frac{(X-\mu_X)^2}{N} = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3} = 8/3$$

تباين العينة
$$\sigma^2_{\bar{X}} = E(\bar{X}^2) - \mu^2_{\bar{X}} = 156/9 - 4^2 = 156/9 - 144/9 = 12/9 = 4/3$$

بطريقة أخرى تباين العينة :
$$\sigma^2_{\bar{X}} = \sum (\bar{x}_i - \mu_X)^2 P(\bar{x}_i) = 4/3$$

بمقارنة نتائج الجدول أعلاه نجد أنّ : $\mu_{\bar{X}} = \mu$ كما أنّ: $\sigma^2_{\bar{X}} = \frac{\sigma^2_X}{n}$.

ويمكن أن نلاحظ وبكل سهولة أن متوسط المتغير $\bar{x}$ يساوي دائما متوسط المجتمع μ الذي أخذنا منه العينة، أما تباينه (تباين المتغير $\bar{x}$) فيتوقف على حجم العينة ويكون مساو لتباين المجتمع الأصلي مقسوما على حجم العينة n، والنتيجة:

كلما كان حجم العينة أكبر ، كلما كان الانحراف المعياري لـ $\bar{x}$ أصغر ، وبالتالي أقرب قيمة لـ $\bar{x}$ سوف تكون أقرب كذلك لـ $\mu_{\bar{X}}$ وهكذا فإن $\bar{x}$ يمكن استخدامها كتقدير لـ μ.

في حالة السحب بدون إرجاع

عدد العينات التي يمكن سحبها في هذه الحالة هو $3 = \frac{3!}{(3-2)!2!} = \frac{N!}{(N-n)!n!}$. والجدول التالي يوضح مختلف العينات ومتوسطاتها وكذا احتمالاتها.

العينات الممكنة	$\bar{x}_i$	$P(\bar{x}_i)$	$\bar{x}_i P(\bar{x}_i)$	$\bar{x}_i^2 P(\bar{x}_i)$
(2.4)	3	31/	33/	3/9
(2.6)	4	31/	3/4	3/16
(4.6)	5	31/	3/5	3/25
المجموع	–	-	4	3/50

وبالتالي:

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

متوسط العينة
$$\mu_{\bar{X}} = E(\bar{X}) = \sum \bar{x}_i P(\bar{x}_i) = 4$$

تباين المجتمع
$$\sigma^2 = \sum \frac{(X-\mu_X)^2}{N} = \frac{(2-4)^2+(4-4)^2+(6-4)^2}{3} = 8/3$$

تباين العينة
$$\sigma_{\bar{X}}^2 = E(\bar{X}^2) - \mu_{\bar{X}}^2 = \frac{50}{3} - 4^2 = \frac{50}{3} - \frac{48}{3} = \frac{2}{3}$$

بطريقة أخرى تباين العينة يساوي
$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{8}{6}\left(\frac{3-2}{3-1}\right) = 2/3$$

بمقارنة نتائج الجدول أعلاه نجد أنه في حالة السحب بدون إرجاع يكون :

$\mu_{\bar{X}} = \mu$ كما أن $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$

فيما يلي نستعرض الحالات المختلفة لإيجاد توزيع المعاينة للمتوسط الحسابي $\bar{X}$.

❖**الحالة الأولى: تباين المجتمع الطبيعي σ^2 معلوم**

عندما يكون مجتمع موزع توزيعا طبيعيا بمتوسط μ وتباين σ^2 معلوم أي: $X\sim N(\mu,\sigma^2)$ فإنه وبغض النظر عن حجم العينة n التي تسحب من هذا المجتمع، سوف يكون توزيع المتوسط الحسابي للعينة هو: $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$.

ولحساب الاحتمالات للمتغير العشوائي $\bar{X}$ يتم إيجاد القيم المعيارية له (أي تحويل قيمته إلى Z المعيارية) حيث:

$$Z=\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

مثال (1- 10):

ليكن $X\sim N(12,9)$ ، أوجد التوزيع الاحتمالي لمتوسط عينة عشوائية حجمها **30** سحبت من المجتمع **X**، ثم أحسب قيمة الاحتمال $P(\bar{x}\le 13)$.

الحل:

بما أنّ $X\sim N(12,9)$ فإن $\bar{X}\sim N(12,\frac{9}{30})$

ولحساب قيمة الاحتمال $P(\bar{x}\le 13)$ نحول $\bar{x}$ إلى Z المعيارية: $Z=\frac{\bar{x}-12}{\sqrt{9/30}}$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{x}\le 13)=P\left(Z\le\frac{13-12}{\sqrt{9/30}}\right)=P(Z\le 1.83)=\emptyset(1.83)=0.9664$$

❖ **الحالة الثانية: تباين المجتمع الطبيعي σ^2 مجهول**

في هذه الحالة نستبدل تباين المجتمع σ^2 بتباين العينة S^2 المسحوبة منه. ويجب أن نميز بين حالتين:

- **في حالة العينات الكبيرة** $(n\ge 30)$:

يكون توزيع المتوسط الحسابي للعينة هو:

$$\bar{X}\sim N(\mu,\frac{S^2}{n})$$

ويكون التوزيع الطبيعي المعياري هو:

$$Z=\frac{\bar{X}-\mu}{\sqrt{S^2/n}}\sim N(0,1)$$

علما أنّ تباين العينة S^2 يحسب بالعلاقة التالية: $S^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$

مثال (1- 11):

أوجد التوزيع الاحتمالي لمتوسط عينة عشوائية حجمها *30* مشاهدة من المجتمع X وتباينها يساوي *14* ، حيث: $X \sim N(12, \sigma^2)$ ، ثم أحسب قيمة الاحتمال $P(\bar{x} \leq 13)$.

الحل:

تباين المجتمع الطبيعي مجهول وحجم العينة كبير، أي أن توزيع المتوسط الحسابي للعينة يتبع التوزيع الطبيعي ويكون:

$$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\right) = N\left(12, \frac{14}{30}\right)$$

وبالتالي يكون التوزيع الطبيعي المعياري هو:

$$Z = \frac{\bar{X} - 12}{\sqrt{14/30}} \sim N(0,1)$$

ومنه الاحتمال المطلوب يحسب كما يلي:

$$P(\bar{x} \leq 13) = P\left(Z \leq \frac{13-12}{\sqrt{14/30}}\right) = P(Z \leq 1.46) = \emptyset(1.46) = 0.9279$$

- **في حالة العينات الصغيرة** $(n < 30)$:

توزيع المتوسط الحسابي للعينة في هذه الحالة سوف يتبع توزيع ستودنت بدرجة حرية $n-1$:

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1)$$

مثال (1- 12):

أوجد التوزيع الاحتمالي لمتوسط عينة عشوائية حجمها **25** مشاهدة من المجتمع X وتباينها يساوي **9** حيث: $X \sim N(10, \sigma^2)$، ثم أو جد قيمة الاحتمال $P(\bar{x} \leq 12)$.

الحل:

لدينا تباين المجتمع الطبيعي مجهول وحجم العينة صغير، فإن توزيع المتوسط الحسابي للعينة يتبع توزيع ستودنت بدرجة حرية $(n-1=19)$:

$$T=\frac{\bar{X}-10}{\sqrt{9/25}}\sim t(24)$$

ومنه الاحتمال المطلوب يصبح:

$$P(\bar{x}\leq 12)=P\left(T\leq\frac{12-10}{\sqrt{9/25}}\right)=P(T\leq 3.33)=0.999$$

❖ **الحالة الثالثة: حجم العينة كبير $(n\geq 30)$ والسحب بدون إرجاع من مجتمع محدود حجمه N.**

في هذه الحالة ندخل معامل التصحيح $\left(\frac{N-n}{N-1}\right)$ على تباين العينة. ويجب أيضا أن نميز بين حالتين:

- **تباين المجتمع الطبيعي معلوم (σ^2 معلوم)**

في هذه الحالة يكون توزيع متوسط العينة هو:

$$\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)\right)$$

أما التوزيع الطبيعي المعياري فيكون:

$$Z=\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)}}\sim N(0,1)$$

مثال (1- 13):

تم سحب عينة عشوائية بدون إرجاع حجمها **40** من مجتمع يتوزع توزيعا طبيعيا حجمه **400** بمتوسط $\mu=20$ وتباين $\sigma^2=25$. والمطلوب أوجد توزيع المتوسط الحسابي للعينة ثم أحسب $P(\bar{x}>19)$.

الحل:

من المعطيات نجد أن تباين المجتمع الطبيعي معلوم والسحب بدون إرجاع، وبذلك يكون توزيع المتوسط الحسابي للعينة هو:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)\right) = N\left(20, \frac{25}{40}\left(\frac{400-40}{400-1}\right)\right) = N(20, 0.56)$$

أي التوزيع المعياري يصبح:

$$Z = \frac{\bar{X}-20}{\sqrt{0.56}} \sim N(0,1)$$

ومنه الاحتمال المطلوب هو:

$$P(\bar{x} > 19) = P\left(Z > \frac{19-20}{\sqrt{0.56}}\right) = P(Z > -1.34) = 1 - P(Z \leq -1.34) = 1 - \emptyset(-1.34)$$
$$= 1 - (1 - \emptyset(1.34)) = \emptyset(1.34) = 0.9099$$

- **تباين المجتمع الطبيعي مجهول (σ^2 مجهول)**

في هذه الحالة نستبدل تباين المجتمع بتباين العينة S^2، ويصبح توزيع متوسط العينة كما يلي:

$$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\left(\frac{N-n}{N-1}\right)\right)$$

أي أن التوزيع الطبيعي المعياري يصبح:

$$Z = \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}\left(\frac{N-n}{N-1}\right)}} \sim N(0,1)$$

مثال (1- 14):

تم سحب عينة عشوائية بدون إرجاع حجمها **40** من مجتمع يتوزع توزيعا طبيعيا حجمه **400** بمتوسط $\mu = 20$. والمطلوب، إذا كان تباين العينة $S^2 = 36$ أوجد توزيع المتوسط الحسابي للعينة ثم أحسب $P(\bar{x} > 19)$.

الحل:

لدينا تباين المجتمع الطبيعي مجهول والسحب بدون إرجاع، نستعين في هذه الحالة بتباين العينة كتقدير لتباين المجتمع، وبذلك يصبح توزيع المتوسط الحسابي للعينة هو:

$$\bar{X} \sim N\left(\mu, \frac{S^2}{n}\left(\frac{N-n}{N-1}\right)\right) = N\left(20, \frac{36}{40}\left(\frac{400-40}{400-1}\right)\right) = N(20, 0.81)$$

وبالتالي التوزيع المعياري يكون:

$$Z = \frac{\bar{X}-20}{\sqrt{0.81}} \sim N(0,1)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{x} > 19) = P\left(Z > \frac{19-20}{\sqrt{0.81}}\right) = P(Z > -1.11) = 1 - P(Z \leq -1.11) = 1 - \emptyset(-1.11)$$
$$= \emptyset(1.11) = 0.8665$$

3-1-2- توزيع المعاينة للمتوسط $\bar{X}$ عند المعاينة من مجتمع غير طبيعي (نظرية النهاية المركزية)

أحيانا يكون المجتمع لا يتوزع طبيعيا، فهناك مجتمعات مجهولة التوزيع (توزيعها غير معروف) وفي هذه الحالة نكون أمام التساؤل التالي:

إذا كان لدينا مجتمع متوسطه μ وتباينه σ^2، وأخذت منه جميع العينات العشوائية ذات الحجم n فما هو توزيع المتوسط الحسابي $\bar{X}$ مع إمكانية عدم تحقق التوزيع الطبيعي للمجتمع؟

هنا وللإجابة على هذا السؤال، نستعين بنظرية النهاية المركزية (نظرية التقارب) بدون برهان.

نظرية (1- 10): إذا سحبت عينة حجمها n من مجتمع إحصائي متوسطه μ وتباينه σ^2، فإنه بغض النظر عن توزيع المجتمع يكون توزيع المتوسط الحسابي للعينة $\bar{X}$ هو بالتقريب يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu$ وتباين $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ وبالتالي فإن المتغير Z له التوزيع الطبيعي المعياري:

$$Z = \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

ملاحظة:

تظهر أهمية هذه النظرية في تبين أن توزيع المعاينة للمتوسط الحسابي يتبع التوزيع الطبيعي إذا كان حجم العينة كبيرا $(n \geq 30)$ بغض النظر عن شكل المجتمع. أما إذا كان حجم العينة صغيرا $(n < 30)$

فالتقريب من التوزيع الطبيعي يكون مقبولا فقط عندما يكون توزيع المجتمع المسحوبة منه تلك العينة طبيعيا (تباين المجتمع معلوم).

مثال (1- 15):

أوجد توزيع المتوسط الحسابي لعينة حجمها $n = 90$ سحبت من مجتمع مجهول التوزيع بمتوسط $\mu = 40$ وتباين $\sigma^2 = 4$.

الحل:

لدينا حجم العينة $n = 90 > 30$ ، ومنه توزيع المتوسط الحسابي للعينة $\bar{X}$ يكون طبيعيا بمتوسط $\mu_{\bar{x}} = 40$ وتباين $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{4}{90}$ وبالتالي:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N(40, \frac{4}{90})$$

في حين التوزيع الطبيعي المعياري يصبح:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} = \frac{\bar{X} - 40}{\sqrt{4/90}} \sim N(0, 1)$$

3-2- توزيع المعاينة للفرق بين متوسطي عينتين Sampling Distribution of the Differences between Two Sample Means

يجب أن نفرق بين العينتين المستقلتين والعينتين المرتبطتين، قبل تحديد توزيع المعاينة للفرق بين متوسطي عينتين.

3-2-1- توزيع المعاينة للفرق بين متوسطي عينتين مستقلتين.

نظرية (1-11): إذا كان لدينا مجتمعان طبيعيان مستقلان بحيث $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$ فإن توزيع المعاينة للفرق بين متوسطي عينتين $(\bar{X}_1 - \bar{X}_2)$ بحجم n_1 و n_2 مسحوبتين من هذين المجتمعين هو توزيع طبيعي بحيث:

- $\mu_{\bar{x}_1 - \bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2$
- $\sigma^2_{\bar{x}_1 - \bar{x}_2} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

مع مراعاة الحالات 03 الأتية:

- **σ_1^2 و σ_2^2 معلومين:**

يكون توزيع الفرق بين متوسطي عينتين هو:

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

ومنه:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

مثال (1- 16):

تم سحب عينتين حجمهما 81 و 49 على التوالي من مجتمعين طبيعيين مستقلين حيث: $X_1 \sim N(30,16)$ و $X_2 \sim N(28,9)$ من هذين المجتمعين. المطلوب: أوجد توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{X}_1 - \bar{X}_2)$، ثم أحسب إحتمال الفرق أقل من 3.

الحل:

لدينا تبايني المجتمعين الطبيعيين معلومين، ومنه توزيع الفرق بين متوسطي عينتين يعطى بالعلاقة التالية:

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

بالتعويض العددي نجد:

$$(\bar{X}_1 - \bar{X}_2) \sim N\left(30 - 28, \frac{16}{81} + \frac{9}{49}\right) = N(2, \frac{38}{100})$$

ومنه القيمة المعيارية تصبح:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 2}{\sqrt{\frac{38}{100}}} \sim N(0,1)$$

أي أن قيمة الاحتمال المطلوب هي:

$$P(\bar{x}_1 - \bar{x}_2 \leq 3) = P\left(Z \leq \frac{3-2}{\sqrt{\frac{38}{100}}}\right) = P(Z \leq 1.62) = \emptyset(1.62) = 0.9474$$

❖ **σ_1^2 و σ_2^2 مجهولين و n_1 و n_2 كبيرين:**

في هذه الحالة توزيع الفرق بين متوسطي العينتين يتبع التوزيع الطبيعي، مع إستبدال σ_1^2 و σ_2^2 بـ تبايني العينتين S_1^2 و S_2^2 كتقدير لهما.

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})$$

وتصبح القيمة المعيارية كما يلي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$$

مثال (1- 17):

أوجد توزيع المعاينة للفرق بين متوسطي عينتين سحبتا من المجتمعين التاليين: $X_1 \sim N(15, \sigma_1^2)$ و $X_2 \sim N(10, \sigma_2^2)$، حيث حجم العينتين هو 50 و 40 على التوالي، وتباينهما $S_1^2 = 36$ و $S_2^2 = 16$ على الترتيب. ما إحتمال الفرق بين المتوسطين أقل من 2؟.

الحل:

حجم العينتين كبير وتبايني المجتمعين الطبيعيين مجهولين، وبالتالي يصبح توزيع الفرق بين متوسطي عينتين هو:

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})$$

بالتعويض العددي نجد:

$$(\bar{X}_1 - \bar{X}_2) \sim N\left(15 - 10, \frac{36}{50} + \frac{16}{40}\right) = N(5, 1.12)$$

أي أن القيمة المعيارية هي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 5}{\sqrt{1.12}} \sim N(0,1)$$

و قيمة الاحتمال المطلوبة فهي:

$$P(\bar{x}_1 - \bar{x}_2 \le 2) = P\left(Z \le \frac{2-5}{\sqrt{1.12}}\right) = P(Z \le -2.83) = \emptyset(-2.83) = 1 - \emptyset(2.83)$$
$$= 1 - 0.9977 = 0.0023$$

❖ σ_1^2 و σ_2^2 **مجهولين و n_1 و n_2 صغيرين أو أحدهما صغير:**

توزيع الفرق بين متوسطي العينتين $(\bar{X}_1 - \bar{X}_2)$ في هذه الحالة يتبع توزيع ستودنت بدرجة حرية $(\vartheta = n_1 + n_2 - 2)$، أي:

$$(\bar{X}_1 - \bar{X}_2) \sim t(n_1 + n_2 - 2)$$

في حين إحصائية ستودنت تكون:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}} \sim t(n_1 + n_2 - 2)$$

مثال (1- 18):

سحبت عينتين عشوائيتين حجمها على التوالي **9** و**33** وكان تباينهما يساوي **36** و25 على الترتيب من مجتمعين يتوزعان طبيعيا بمتوسط يساوي **65** و**59 تواليا.** ما هو إحتمال الفرق بين المتوسطين يكون أكبر من 12.

الحل:

لدينا

$$\mu_1 = 65 \qquad S_1^2 = 36 \qquad n_1 = 9$$

$$\mu_2 = 59 \qquad S_2^2 = 25 \qquad n_2 = 33$$

وبما أنّ $n_1 = 9 < 30$ وتباين المجتمعين مجهول، إذن توزيع الفرق بين متوسطي العينتين هو:

$$(\bar{X}_1 - \bar{X}_2) \sim t(n_1 + n_2 - 2)$$

$$(\bar{X}_1 - \bar{X}_2) \sim t(40)$$

ومنه فإن:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}} = \frac{(\bar{X}_1 - \bar{X}_2) - 6}{\sqrt{\left(\frac{36}{9} + \frac{25}{33}\right)}} \sim t(40)$$

أخيرا الاحتمال المطلوب يكون:

$$P(\bar{x}_1-\bar{x}_2>12)=P\left(t>\frac{12-6}{\sqrt{\left(\frac{36}{9}+\frac{25}{33}\right)}}\right)=P(t>2.7)=0.005$$

ملاحظة01: هناك حالة خاصة عندما يكون $\sigma_1^2=\sigma_2^2=\sigma^2$

عند توفر الشرطين: حجم العينتين يكون صغير أو أحدهما صغير، العلم أنّ $\sigma_1^2=\sigma_2^2=\sigma^2$ فإننا نقدر σ^2 بـ S_p^2 حيث S_p^2 تحسب بالعلاقة التالية:

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1-1)+(n_2-1)}=\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}$$

وتصبح إحصائية ستودنت كما يلي:

$$T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\sim t(n_1+n_2-2)$$

ملاحظة02: تساوي تبايني المجتمعين σ_1^2 و σ_2^2 ، يعني أن المجتمعين متجانسين وذلك وفقا لدراسات سابقة.

مثال (1- 19):

سحبت عينتين عشوائيتين من مجتمعين طبيعيين تباينهما مجهولين حيث: $X_1\sim N(30,\sigma_1^2)$ و $X_2\sim N(28,\sigma_2^2)$، بافتراض تجانس تبايني المجتمعين و علما أن حجم العينة الأولى يساوي **12** بتباين يساوي **04** وحجم العينة الثانية يساوي **10** بتباين يساوي **5**.

أوجد توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{X}_1-\bar{X}_2)$، ثم احسب إحتمال الفرق أقل من 2.

الحل: من المعطيات لدينا:

$\mu_1=30$ $\quad S_1^2=4$ $\quad n_1=12$

$\mu_2=28$ $\quad S_2^2=5$ $\quad n_2=10$

ولدينا تبايني المجتمعين مجهولين ومتساويين وحجم كل من العينتين أقل من **30** مشاهدة، فإن توزيع الفرق بين متوسطي العينتين هو:

$$(\bar{X}_1-\bar{X}_2)\sim t(n_1+n_2-2)$$

بالتعويض العددي نجد أن:

$$(\bar{X}_1-\bar{X}_2)\sim t(20)$$

نقوم بحساب تقدير تباين المجتمعين σ^2 بـ S_P^2 حيث:

$$S_P^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}=\frac{(11)(4)+(9)(5)}{20}=4.45$$

ومنه:

$$T=\frac{(\bar{X}_1-\bar{X}_2)-2}{\sqrt{4.45(\frac{1}{12}+\frac{1}{10})}}\sim t(20)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{x}_1-\bar{x}_2<2)=P\left(t<\frac{2-2}{\sqrt{4.45(\frac{1}{12}+\frac{1}{10})}}\right)=P(t<0)=0.5$$

3-2-2- توزيع المعاينة للفرق بين متوسطي عينتين مرتبطتين.

في كثير من الأحيان نلجأ في بعض الأحيان إلى سحب عينتين مرتبطتين، ويتوقف الأمر هنا على الظاهرة المدروسة. فمثلا لقياس فعالية منتوج جديد من الأدوية على عينة من المرضى يتم قياس مستوى الحالة الصحية للمرضى قبل تناول الدواء وبعد تناوله، حيث تكون قيم العينة الأولى مقترنة بقيم العينة الثانية أي أن قيم المشاهدات تكون عبارة عن ثنائيات من القيم، فتكون القراءة قبل وبعد تناول الدواء تشكل عينتين مقترنتين أي مرتبطتين.

ولإيجاد توزيع للفرق بين متوسطي المجتمعين الذين سحبت منهما العينتين في هذه الحالة، يتم إيجاد توزيع لمتوسط واحد لهما. وفيما يلي توضيح للعملية كيف تتم.

إذا كان لدينا مجتمعين مرتبطين كل منهما يتوزع طبيعيا أي: $X\sim N(\mu_1,\sigma_1^2)$ و $Y\sim N(\mu_2,\sigma_2^2)$ وسحبت من كل مجتمع عينة حجمها n $\{X_1,X_2,\dots\dots,X_n\}$ و $\{Y_1,Y_2,\dots\dots,Y_n\}$ على الترتيب، بحيث تمثل X_i و Y_i القيمتين المقترنتين في العينتين، إذا كانت الفروق بين قيم العينتين المتناظرة هي $\{D_1,D_2\dots,D_n\}$ حيث $D_i=X_i-Y_i$ لجميع قيم $i=1,2,\dots,n$، فإن $\{D_1,D_2\dots,D_n\}$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على أنها عينة عشوائية من مجتمع طبيعي متوسطه $\mu_D=\mu_1-\mu_2$ وتباينه σ_D^2، ولنفرض أن متوسط العينة μ_d و تباينها S_d^2.

ونميز بين الحالتين الأتيتين (التوزيع في حالة العينات الكبيرة والعينات الصغيرة):

- **الحالة الأولى: العينات الكبيرة** $(n \geq 30)$

يكون توزيع المتوسط الحسابي هو:

$$\bar{D} \sim N(\mu_D, \frac{\sigma_D^2}{n})$$

وتعطى القيمة المعيارية له بالعلاقة الاتية:

$$Z = \frac{\bar{D} - \mu_D}{\sqrt{\frac{\sigma_D^2}{n}}} \sim N(0,1)$$

وفي حالة ما إذا كان تباين المجتمعين مجهول يستبدل بتباين العينتين.

- **الحالة الثانية: العينات الصغيرة** $(n < 30)$

يكون توزيع المتوسط الحسابي هو:

$$\bar{D} \sim t(n-1)$$

وقيمته المعيارية تعطى بالعلاقة:

$$T = \frac{\bar{D} - \mu_D}{\sqrt{\frac{S_D^2}{n}}} \sim t(n-1)$$

مثال (1- 20):

تم إختيار عينة من موظفي مؤسسة ما بحجم **10** أشخاص، تم إخضاعهم لإختبار الذكاء وبعدها ادخلوا دورة لتحسين سرعة البديهة لمدة ستة أشهر، وبعدها تم إخضاعهم مرة ثانية لإختبار الذكاء. الجدول أدناه يبين الدرجات التي حصلوا عليها قبل وبعد الدورة. المطلوب أوجد توزيع المتوسط الحسابي للعينة ثم أحسب احتمال أن يكون الفرق قبل أخذ الدورة وبعدها أكبر من **26**.

قبل الدورة X_i	135	140	128	125	130	136	137	140	132	138
بعد الدورة Y_i	138	150	134	125	131	136	141	144	133	139

الحل:

نحسب أولا الفروق بين قبل/بعد الخضوع للدورة التحسينية كما يلي:

قبل الدورة X_i	135	140	128	125	130	136	137	140	132	138
بعد الدورة Y_i	138	150	134	125	131	136	141	144	133	139
الفروق D_i	3	10	6	0	1	0	4	4	1	1

نقوم بحساب التوقع والتباين أي توزيع المعاينة للفرق بين العينتين كما يلي:

$$\mu_D = \frac{\sum D_i}{n} = 3$$

$$S_D^2 = \frac{\sum(D_i - \mu_D)^2}{n-1} = 10$$

وبذلك تكون القيمة المعيارية على الشكل الأتي:

$$T = \frac{\bar{D} - \mu_D}{\sqrt{\frac{S_D^2}{n}}} = \frac{\bar{D} - 3}{\sqrt{\frac{10}{10}}} \sim t(9)$$

ومنه الاحتمال المطلوب هو:

$$P(\bar{D} > 7) = P\left(T > \frac{7-3}{\sqrt{1}}\right) = P(T > 4) = 0.005$$

3-3- توزيع المعاينة للنسبة في العينة Sampling Distribution of the Sample Proportions

عندما يمكن لكل مشاهدة في المجتمع أن تحمل خاصية من الخاصيتين 0 أو 1، نعم أو لا، فشل أو نجاح، مدخن أو غير مدخن، أمي أو غير أمي...الخ. فإننا نسمي كل مشاهدة من هذا المجتمع بتجربة برنولي. وعندما نأخذ عينة عشوائية من هذا المجتمع فإن المعاينة في هذه الحالة تكون من مجتمع ذي الحدين (تجربة برنولي مكررة عدة مرات).

وينصب اهتمامنا في هذه الحالة على نسبة ظاهرة معينة في المجتمع، ويرمز لها بـ P . **حيث** يتم الحصول عليها بقسمة مجموع مفردات الخاصية المدروسة X (عدد الراسبين في المجتمع مثلا) على حجم المجتمع N كما يلي $\left(P = \frac{X}{N}\right)$. أما حساب النسبة في العينة لخاصية ما (عدد الراسبين في العينة مثلا) والتي يرمز لها بـ $\bar{P}$ فيتم الحصول عليها بقسمة مجموع مفردات الخاصية المدروسة على حجم العينة كما يلي $\bar{P} = \frac{x}{n}$.

لو أخذنا جميع العينات التي حجمها n مجتمع حجمه N لوجدنا أن النسب لهذه العينات تختلف من عينة لأخرى، مما يعني أن هذه النسب أصبحت عبارة عن متغير عشوائي، هذا المتغير له توزيع احتمالي مستمدا من توزيع المجتمع الذي سحبت منه هذه العينات يسمى توزيع النسبة في العينة.

عندما يكون حجم العينة كبيرا ($n \geq 30$) فإن توزيع المعاينة لنسبة العينة سيكون تقريبا طبيعيا (بالإعتماد على نظرية النهاية المركزية) بمتوسط يساوي:

- $\mu_{\hat{P}} = P$

وتباين يساوي:

- $\sigma^2_{\hat{P}} = \frac{P(1-P)}{n} = \frac{PQ}{n}$

أي أن توزيع النسبة في العينة هو:

$$\hat{P} \sim N\left(P, \frac{P(1-P)}{n}\right)$$

ولحساب الإحتمالات المتعلقة بالنسبة P نستخدم العلاقة التالية:

$$Z = \frac{\hat{P} - P}{\sqrt{P(1-P)/_n}} \sim N(0,1)$$

ملاحظة:

1. حتى نتمكن من استخدام التقريب من التوزيع ثنائي الحد إلى التوزيع الطبيعي بشكل صحيح، يجب أن يكون كل من nP و $n(1-P)$ أكبر من 5 (وهذا في حالة حجم العينة n كبير بما فيه الكفاية).
2. في حالة السحب بدون إرجاع نضرب تباين العينة في معامل التصحيح.
3. في حالة مجهولية P نستبدلها بـ $\hat{P}$.

مثال (1- 21):

من سجلات الجامعة للسنوات السابقة لـ 200 طالب سنة ثانية لقسم علوم التسيير، 150 طالب نجح في الدورة العادية في مادة الإحصاء 03 ، إذا سحبنا عينة من **36** طالب. أوجد توزيع نسبة الطلبة الناجحين في العينة ثم أحسب احتمال أن تكون نسبة الناجحين في العينة أقل من **75%**.

الحل:

$$X = 150 \qquad N = 200 \qquad n = 36$$

من معطيات المثال، يتبين لنا أن نسبة الطلبة الناجحين في المجتمع تساوي: $P = \frac{150}{200} = 0.75$

ومنه توزيع المعاينة لنسبة الطلبة الناجحين يحدد وفق قيم التوقع والتباين، حيث:

- $\mu_{\hat{P}} = P = \frac{X}{N} = 0.75$
- $\sigma^2_{\hat{P}} = \frac{P(1-P)}{n} = \frac{(0.75)(0.25)}{36} = 0,005$

وبذلك يكون توزيع المعاينة للنسبة في العينة هو:

$$\hat{P} \sim N(0.75\,,0.005)$$

وبالتالي القيمة المعيارية هي:

$$Z = \frac{\hat{P} - P}{\sqrt{P(1-P)/_n}} = \frac{\hat{P} - 0.75}{\sqrt{0.005}} \sim N(0,1)$$

في حين قيمة الاحتمال المطلوبة تساوي:

$$P(\hat{P} < 0.75) = P\left(Z < \frac{0.75 - 0.75}{\sqrt{0.005}}\right) = P(Z < 0) = \emptyset(0) = 0.5$$

3- 4- توزيع المعاينة للفرق بين نسبتي عينتين Sampling Distribution of the Differences between Two Sample Proportions

عند القيام بدراسة مقارنة بين ظاهرة معينة في مجتمعين مستقلين أي مختلفين، فإننا نقوم بالاستدلال على المعلمة $P_1 - P_2$ أي أننا نستنتجها باستخدام الفرق بين نسبتي العينتين العشوائيتين المسحوبتين من هذين المجتمعين وذلك بإستخدام الإحصائية $\widehat{P_1} - \widehat{P_2}$.

__نظرية__: إذا سحبنا عينة كبيرة حجمها n_1 من مجتمع يتبع التوزيع ثنائي الحد $X_1 \sim b(n_1, P_1)$ وسحبنا عينة كبيرة أخرى n_2 من مجتمع آخر مستقل عن المجتمع الأول ويتبع التوزيع ثنائي الحد $X_2 \sim b(n_2, P_2)$، فإن الفرق بين النسبتين في العينتين $(\widehat{P_1} - \widehat{P_2})$ هو متغير عشوائي يخضع تقريبا للتوزيع الطبيعي بمتوسط:

$$\mu_{\hat{P}_1-\hat{P}_2} = \mu_{\hat{P}_1} - \mu_{\hat{P}_2} = P_1 - P_2$$

وتباينه:

$$\sigma^2_{\hat{P}_1-\hat{P}_2} = \sigma^2_{\hat{P}_1} + \sigma^2_{\hat{P}_2} = \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}$$

ويكون توزيع المعاينة للفرق بين نسبتي العينتين كما يلي:

$$(\widehat{P_1} - \widehat{P_2}) \sim N\left(P_1 - P_2, \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}\right)$$

ومنه القيمة المعيارية :

$$Z = \frac{(\widehat{P_1} - \widehat{P_2}) - P_1 - P_2}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} \sim N(0,1)$$

مثال (1- 22):

عند إجراء امتحان الاحصاء 03 لطلبة السنة الثانية علوم التسيير تحصلنا على النتائج التالية:

الطالبات	$n_1=150$	$P_1=0.64$
الطلبة	$n_2=120$	$P_2=0.60$

أوجد توزيع المعاينة للفرق بين نسبتي العينتين ثم أحسب احتمال أن تكون نسبة الناجحات في عينة الطالبات أقل من نسبة الناجحين في عينة الطلبة بمقدار 0.06 ؟

الحل:

$$P_1 = 0.64 \qquad P_2 = 0.60 \qquad n_1 = 150 \qquad n_2 = 120$$

توزيع المعاينة للفرق بين نسبتي العينتين هو:

- $\mu_{\hat{P}_1-\hat{P}_2} = \mu_{\hat{P}_1} - \mu_{\hat{P}_2} = P_1 - P_2 = 0.64 - 0.60 = 0.04$
- $\sigma^2_{\hat{P}_1-\hat{P}_2} = \sigma^2_{\hat{P}_1} + \sigma^2_{\hat{P}_2} = \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2} = \frac{0.64(0.36)}{150} + \frac{0.60(0.40)}{120} = 0.0035$

$$(\widehat{P_1} - \widehat{P_2}) \sim N(0.04, 0.0035)$$

ومنه:

$$Z = \frac{(\widehat{P_1} - \widehat{P_2}) - 0.04}{\sqrt{0.0035}} \sim N(0,1)$$

ومنه قيمة الاحتمال المطلوبة تساوي:

$$P(\widehat{P}_1 - \widehat{P}_2 < 0.06) = P\left(Z < \frac{0.06 - 0.04}{\sqrt{0.0035}}\right) = P(Z < 5.66) = \emptyset(5.66) = 1$$

3-5- توزيع المعاينة للتباين Sampling Distribution of Variance

يستخدم تباين العينة S^2 للإستدلال عن تباين المجتمع حيث *نحتاج* في كثير من الأحيان في العمل الإحصائي إلى معرفة توزيع تباين العينة. وسنقدم هذا التوزيع في حالة المعاينة من توزيع طبيعي.

نميز بين القيمتين التالتين لتباين العينة وذلك حسب حجم العينة أكبر أو أقل من 30 على الترتيب:

$$S^2 = \frac{\sum(X_i - \overline{X})^2}{n} \qquad\qquad S^2 = \frac{\sum(X_i - \overline{X})^2}{n-1}$$

إذا أخذنا كل العينات العشوائية الممكنة والتي حجمها n من مجتمع موزع توزيعا طبيعيا متوسطه μ وتباينه σ^2 أي $X \sim N(\mu, \sigma^2)$ وكان S^2 هو تباين العينة فإن $\frac{(n-1)S^2}{\sigma^2}$ يتبع التوزيع مربع كاي بدرجة حرية $n-1$ أي $\frac{(n-1)S^2}{\sigma^2} = \frac{nS^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

<u>**ملاحظة**</u>: **في حالة المعاينة بدون إرجاع و** $\frac{n}{N} > 0.05$ **نستعمل معامل التصحيح** $\frac{N-n}{N-1}$

مثال (1- 23):

سحبت عينة عشوائية حجمها **10** من مجتمع طبيعي إنحرافه المعياري يساوي 4، ما هو احتمال أن يكون تباين العينة S^2 أقل من 6.

الحل:

$$n = 10 \qquad\qquad \sigma^2 = 16$$

توزيع المعاينة لتباين العينة هو:

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

نحسب درجة الحرية : V=n-1=10-1=9

أي أن قيمة الاحتمال المطلوبة هي:

$$P(S^2 < 6) = P\left(\frac{10S^2}{16} < \frac{10 * 6}{16}\right) = P(\chi^2 < 3.75)$$

بالبحث عن هذه القيمة في جدول مربع كاي عند درجة حرية 9، لا نجدها. في الحقيقة هي محصورة بين قيمتين هما: 3.325 و 4.168 وذلك عند الاحتمالين: 0.05 و 0.1 على الترتيب.

$$\begin{array}{ccc} 3.325 & 3.75 & 4.168 \\ 0.05 & \alpha & 0.1 \end{array}$$

$$4.168-3.325 \rightarrow 0.1-0.05$$

$$3.75-3.325 \rightarrow \alpha-0.05$$

$$0.843 \rightarrow 0.05$$

$$0.425 \rightarrow \alpha-0.05 \qquad \alpha=\frac{0.425*0.05}{0.843}+0.05=0.075$$

$$P(S^2<6)=P\left(\frac{10S^2}{16}<\frac{10*6}{16}\right)=P(\chi^2<3.75)=0.075$$

3- 6- توزيع المعاينة للنسبة بين تبايني عينتين

للمقارنة بين تبايني مجتمعين فإننا نحتاج النسبة بين تبايني عينتين مأخوذتين من هذين المجتمعين، ونظرا لسهولة الحسابات والتفسير نلجأ لحساب النسب بين التباينات وليس الفرق بينها. وسنتطرق إلى توزيع هذه النسبة في حالة المعاينة من مجتمعين مستقلين يتوزعان طبيعيا.

إذا سحبنا عينة حجمها n_1 من مجتمع يتوزع طبيعيا $N(\mu_1,\sigma_1^2)$ وكان تباينها S_1^2، وسحبنا عينة أخرى حجمها n_2 من مجتمع آخر يتوزع طبيعيا مستقل عن الأول $N(\mu_2,\sigma_2^2)$ وكان تباينها S_2^2 فإن:

$$\frac{(n_1-1)S_1^2}{\sigma_1^2}\sim\chi^2_{(n_1-1)} \quad , \quad \frac{(n_2-1)S_2^2}{\sigma_2^2}\sim\chi^2_{(n_2-1)}$$

وبحسب توزيع فيشر ومن أجل مجتمعين مستقلين تكون الإحصاءة:

$$f=\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}\Big/(n_1-1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2}\Big/(n_2-1)}\sim F(n_1-1,n_2-1)$$

ومنه:

$$f=\frac{S_1^2}{S_2^2}.\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1-1,n_2-1)$$

$$F=\frac{\frac{\chi_1^2}{(n_1-1)}}{\frac{\chi_2^2}{(n_2-1)}}=\frac{\frac{\frac{(n_1-1)\hat{S}_1^2}{\sigma_1^2}}{(n_1-1)}}{\frac{\frac{(n_2-1)\hat{S}_2^2}{\sigma_2^2}}{(n_2-1)}}=\frac{\frac{\hat{S}_1^2}{\sigma_1^2}}{\frac{\hat{S}_2^2}{\sigma_2^2}}=\frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}\sim F_{n_1-1,n_2-1}$$

حيث:

مثال (1- 24):

نسحب عينتين حجمهما 8 و 10، من مجتمعين طبيعيين تباينهما على التوالي 20 و 36.

ما هو احتمال أن يكون $P\left(\frac{s_1^2}{s_2^2}>7.56\right)$؟

الحل:

لدينا المعطيات التالية.

$n_1=8$ $\sigma_1^2=20$ $n_2=10$ $\sigma_2^2=36$

توزيع المعاينة للنسبة بين تبايني العينتين هو:

$$f=\frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}=\frac{S_1^2}{S_2^2}\cdot\frac{20}{36}\sim F(7,9)$$

حيث: $n_1-1=8-1=7$ $n_2-1=10-1=9$

ومنه قيمة الاحتمال هي:

$$P\left(\frac{S_1^2}{S_2^2}>7.56\right)=P\left(\frac{S_1^2}{S_2^2}\frac{20}{36}>7.56\frac{20}{36}\right)=P\ (f\ >4.2\ )=0.025$$

تمارين مقترحة

التمرين الأول:

إذا كانت X متغير عشوائي يتبع التوزيع الطبيعي بمتوسط 50 وبانحراف معياري 5، المطلوب:

- أوجد القيم الاحتمالية التالية:

$P(45 \leq x \leq 50)$ $P(55 \leq x \leq 60)$ $P(50 \leq x \leq 55)$

- أوجد قيمة **a** و **b** بحيث:

- $P(Z < a) = 0.579260$, $P(Z < -b) = 0.011304$

التمرين الثاني:

إذا كان المتغير العشوائي X يتبع التوزيع الطبيعي حيث، متوسطه يساوي 3 وانحراف معياري قدره 4.

- ما هو احتمال أن يقع X بين القيمتين 3 و 5؟
- ما هو احتمال أن يكون X اكبر من 1؟

التمرين الثالث:

أوجد القيم الجدولية التالية:

$$\chi^2_{(0.995,5)} \quad , \quad \chi^2_{(0.1,15)} \quad , \quad t_{(0.975,3)} \quad , t_{(0.01,10)} \quad , \quad F_{(0.05,2,8)} \quad , F_{(0.975,2,8)}$$

التمرين الرابع:

- إذا كانت المدة المتوسطة لوصول طلبة مدينة عين وسارة إلى الجامعة هي 45 دقيقة، بانحراف معياري 12 دقيقة، نختار عينة من الطلبة حجمها 50 طالب.

أوجد توزيع المعاينة للمتوسط؟ ما هو احتمال أن ينحرف توزيع المعاينة للمتوسط بـ 1 دقيقة عن متوسط المجتمع؟

- لتكن البيانات التالية:

المجتمع	متوسط المجتمع	تباين المجتمع
1	**40**	**18**
2	**35**	**25**

تم سحب عينتين مستقلتين من المجتمعين حيث: $n_1 = 12$ و $n_2 = 10$، أوجد توزيع المعاينة للفرق بين متوسطي العينتين إذا علمت أن كلا المجتمعين يتوزعان طبيعيا؟

التمرين الخامس:

إذا كانت نسبة الإناث في قسم العلوم الاقتصادية هي **70 %** وفي قسم علوم التسيير هي **90 %**، وسحبنا عينة حجمها **60** من قسم العلوم الاقتصادية وعينة أخرى حجمها **100** من قسم علوم التسيير، أوجد توزيع الفرق بين نسبتي العينتين ؟ ما احتمال أن تزيد نسبة النجاح في قسم العلوم الاقتصادية عن نسبة النجاح في قسم علوم التسيير بمقدار **10% على الاكثر ؟**.

التمرين السادس:

إذا علمت أن المجتمع التالي يتبع التوزيع الطبيعي، (1، 3، 5، 7، 9)، تم سحب جميع العينات الممكنة ذات الحجم $n = 2$.

- أحسب متوسط المجتمع وتباينه؟
- أوجد كل العينات الممكنة في حالة السحب بالارجاع وبدون ارجاع؟
- أحسب القيمة المتوقعة لتباين العينة S^2 في الحالتين؟ قارنها بتباين المجتمع؟

التمرين السابع:

إذا كانت $\boldsymbol{X \sim N(\mu, 70)}$ وتم سحب عينة حجمها **11** فما هو احتمال أن يكون تباين العينة أقل من أو يساوي **82.9 ؟**.

التمرين الثامن:

ما هو احتمال أن يأخذ المتغير F قيمة أكبر من 0.36 إذا كان لدينا: $n_1 = 16$ و $n_2 = 20$

تمهيد

غالبا ما يحتوي المجتمع الإحصائي محل الدراسة على معالم(مؤشرات) مجهولة مثل المتوسط الحسابي للمجتمع μ، تباين المجتمع σ^2، النسبة في المجتمع P...، لأنّ معرفتها يتطلب توفر جميع البيانات الخاصة بكل أفراد المجتمع، ويرغب الباحثون في تقدير هذه المعالم من البيانات المأخوذة من العينات العشوائية وذلك عن طريق حساب بعض الإحصاءات بالإعتماد على بيانات العينة وإعتبارها كتقديرات لهذه المعالم. كأن يستخدم متوسط العينة كتقدير لمتوسط المجتمع.

وتتم عملية التقدير الاحصائي بطريقتين:

- **التقدير النقطي Point estimation**.
- **التقدير بمجال ثقة Confidence interval estimation**.

وفي ما يلي عرض لهذين الطريقتين بشيء من التفصيل.

أولا: التقدير النقطي Point estimation

رأينا في الفصل الأول أنّ إحصائيات العينة تستخدم كتقدير لمعالم المجتمع، فمثلا يمكن إعطاء تقديرات لمعالم مجتمع انطلاقا من إحصائيات عينة عشوائية مسحوبة منه مثل: متوسط العينة $\bar{X}$ المحسوب من عينة حجمها n هو تقدير لمتوسط المجتمع μ، وتباين العينة S^2 هو تقدير لتباين المجتمع σ^2، نسبة العينة $\widehat{P}$ حيث $\widehat{P} = \frac{x}{n}$ هو تقدير لنسبة المجتمع P في توزيع برنولي. وينطبق الأمر على باقي الحالات مثل حالة عينتين نجد الفرق بين متوسطي عينتين هو تقدير للفرق بين متوسطي المجتمعين، والنسبة بين تبايني عينتين هو تقدير للنسبة بين تبايني مجتمعين ...وهكذا. يسمى حينها هذا المقدر بالمقدر النقطي و يرمز عموما للمقدر النقطي بالرمز $\widehat{\theta}$.

في ما يلي نتناول شروط التقدير الجيد و أنواع التقدير وبعض نظرياته.

تعريف (2- 1):

المقدِّر (Estimator) عبارة عن إحصائية (Statistic)، بينما التقدير (Estimation) عبارة عن قيمة (value) المقدِّر (الإحصائية)، بعبارة أخرى يمكننا القول أن المقدِّر هو إحصاءة أو دالة تعتمد على المشاهدات في حين أن التقدير هو قيمة هذه الدالة عند التعويض بقيم المشاهدات.

مثال (2- 1):

- $\bar{X}$ عبارة عن **مقدِّر** لـ

تجدر الإشارة إلى أن التقدير يختلف من عينة لأخرى وذلك لإختلاف قيم المشاهدات من عينة لأخرى، لهذا علينا ألاّ نتوقع من المقدِّر أن يعطينا معالم المجتمع بدون أخطاء.

خواص المقدِّر الجيد Properties of Good Estimators

نقول عن المقدر (Estimator) أنه مقدر جيد وفعال إذا توفر على الخواص التالية:

1.عدم التحيز 2.الإتساق 3.الكفاءة 4.الكفاية

1- عدم التحيز Unbiasedness

يعتبر المقدر غير متحيز إذا كان توقعه الرياضي يساوي قيمة المعلمة الحقيقية للمجتمع أي : $E(\hat{\theta}) = \theta$

أما إذا كان $E(\hat{\theta}) = \theta + a$ فإن المقدار a يسمى مقدار التحيز أي الخطأ.

مثال (2- 3):

أثبت أن $\bar{X}$ مقدر غير متحيز لمتوسط مجتمع يتبع التوزيع الطبيعي

الجواب:

المقدر $\bar{X}$ مقدر غير متحيز لمتوسط مجتمع يتبع التوزيع الطبيعي μ، وأيضا القيمة $\hat{S}^2 = \frac{\sum(X_i-\bar{X})^2}{n-1}$ مقدرغير متحيز لـ σ^2 لأنّ:

- $E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \left(\frac{\sum E(X_i)}{n}\right) = \frac{1}{n}(E(X_1) + E(X_2) + \cdots + E(X_n)) = \frac{(\mu+\mu+\cdots+\mu)}{n} = \frac{n\mu}{n} = \mu$
- $E(\hat{S}^2) = E\left(\frac{\sum(X_i-\bar{X})^2}{n-1}\right) = \frac{1}{n-1}E(\sum[(X_i-\mu) - (\bar{X}-\mu)]^2)$

$$= \frac{1}{n-1}E(\sum(X_i-\mu)^2 - 2(\bar{X}-\mu)\sum(X_i-\mu) + \sum(\bar{X}-\mu)^2)$$

$$= \frac{1}{n-1}E\left(\sum(X_i-\mu)^2 - n(\bar{X}-\mu)^2\right)$$

$$= \frac{1}{n-1}\left(\sum E(X_i-\mu)^2 - nE(\bar{X}-\mu)^2\right)$$

$$= \frac{1}{n-1}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \frac{(n-1)\,\sigma^2}{n-1} = \sigma^2$$

وبنفس الطريقة يمكن إثبات أن $S^2 = \frac{\sum(X_i-\bar{X})^2}{n}$مقدر متحيز لـ σ^2، بحيث:

$$E(S^2) = E\left(\frac{\sum(X_i-\bar{X})^2}{n}\right) = \frac{1}{n}E\left(\sum[(X_i-\mu) - (\bar{X}-\mu)]^2\right)$$

نعلم أن:

$$\left(x_i - \bar{X}\right) = \left(x_i - \mu\right) - \left(\bar{X} - \mu\right)$$

$$E(S^2) = \frac{1}{n}E\left(\sum(X_i-\mu)^2 - 2(\bar{X}-\mu)\sum(X_i-\mu) + \sum(\bar{X}-\mu)^2\right)$$

$$E(S^2) = \frac{\sum_{i=1}^{n} E(x_i-\mu)^2}{n} + \frac{\sum_{i=1}^{n} E(\bar{X}-\mu)^2}{n} - 2E\left[\frac{(\bar{X}-\mu)\sum_{i=1}^{n}(x_i-\mu)}{n}\right]$$

ونعلم أن: $\sum_{i=1}^{n}(x_i-\mu) = n(\bar{X}-\mu)$

$$E(S^2) = \frac{\sum_{i=1}^{n} E(x_i-\mu)^2}{n} + E(\bar{X}-\mu)^2 - 2E(\bar{X}-\mu)^2$$

ومنه: $E(S^2) = E(x_i-\mu)^2 - E(\bar{X}-\mu)^2$

أي أن:

$$E(S^2) = V(X) - V(\bar{X}) = \delta^2 - \delta^2/n = (\frac{n-1}{n})\delta^2$$

نلاحظ ببساطة أن S^2 مقدر متحيز لتباين المجتمع δ^2.

2- الاتساق Consistency

نقول أن المقدر $\hat{\theta}$ مقدرا متسقا للمعلمة θ إذا كان:

$$\lim_{n\to\infty} V(\hat{\theta}) = 0 \qquad , \quad E(\hat{\theta}) = \theta$$

مثال (2- 4):

مجتمع يتوزع توزيعا طبيعيا، سحبت منه عينة عشوائية.

أثبت أن $\bar{X}$ هو تقدير متسق.

الحل:

في المثال السابق تم إثبات أن المتوسط الحسابي للعينة $\bar{X}$ هو مقدر غير متحيز للمتوسط الحسابي للمجتمع μ

أي أن: $E(\bar{X}) = \mu$ وتباينه: $\delta_{\bar{x}}^2 = \delta^2/n$

المتوسط الحسابي للعينة $\bar{X}$ هو مقدر متسق للوسط الحسابي للمجتمع μ ، لأنه تقدير غير متحيز وتباينه يقترب من الصفر أي يؤول إلى الصفر كلما زاد حجم العينة حيث: $\lim_{n\to\infty}\delta_{\bar{x}}^2 = \lim_{n\to\infty}\delta^2/n = 0$.

3- الكفاءة Efficiency

إذا كان لدينا أكثر من مقدر لمعلمة ما، فإن المقدر الكفؤ $\widehat{\theta}$ هو المقدر الذي يكون له أصغر تباين من بين المقدرات الأخرى لنفس المعلمة θ.

إذا كان لدينا المقدران $\widehat{\theta}_1$ و$\widehat{\theta}_2$ غير متحيزان للمعلمة المجهولة θ، أي أنّ:

$$E(\widehat{\theta}_1) = \theta \qquad , \quad E(\widehat{\theta}_2) = \theta$$

فإن المقدر $\widehat{\theta}_1$ يكون أكثر كفاءة من $\widehat{\theta}_2$ ، **إذا كان** $\sigma^2_{\widehat{\theta}_1} < \sigma^2_{\widehat{\theta}_2}$.

مثال (2- 5):

بالرجوع إلى توزيعات المعاينة للمتوسط في حالة السحب بالإرجاع وبدون إرجاع نجد أن: $E(\overline{X}) = \mu$

أي من المقدرين للمتوسط يمثل المقدر الكفؤ؟

الحل: في حالة السحب بالإرجاع نجد أن تباين المتوسط يساوي $\delta^2_{\overline{X}} = \delta^2/n$، و في حالة كان السحب بدون إرجاع فإن تباين المتوسط يساوي $\delta^2_{\overline{X}} = \frac{\delta^2}{n}\frac{N-n}{N-1}$.

يمكن ببساطة ملاحظة أن تباين المتوسط في حالة السحب بدون إرجاع أقل من تباين المتوسط في حالة السحب بالإرجاع، لأن القيمة $0 \prec \frac{N-n}{N-1} \prec 1$. وبالتالي يصبح متوسط العينة تقدير كفء في حالة السحب بدون إرجاع.

ملاحظة:

عادة ما تستخدم النسبة $\frac{\sigma^2_{\widehat{\theta}_1}}{\sigma^2_{\widehat{\theta}_2}}$ كمقياس لتحديد أي من المقدرين الأكفأ، فإذا كانت هذه النسبة $1 < \frac{\sigma^2_{\widehat{\theta}_1}}{\sigma^2_{\widehat{\theta}_2}}$ فإنّ $\widehat{\theta}_1$ هو المقدر الأكفأ، أما إذا كانت النسبة $1 < \frac{\sigma^2_{\widehat{\theta}_1}}{\sigma^2_{\widehat{\theta}_2}}$ فإن $\widehat{\theta}_2$ هو المقدّر الأكفأ.

4- الكفاية Suffisance

يقال أن المقدر $\widehat{\theta}$ مقدرا كافيا للمعلمة المجهولة θ إذا استخدمت جميع البيانات الموجودة في العينة والخاصة بحساب المعلمة المراد تقديرها. وهذا يعني أنه بمعرفة $\widehat{\theta}$ **فإن المعلومات المتبقية في العينة لا تفيد في معرفة** θ. لأن المقدر الكافي يستخدم كل المعلومات الموجودة في العينة والخاصة بتقدير المعلمة.

مثال (2- 6): أي المقدرات من بين مقاييس النزعة المركزية يعتبر مقدرا كافيا؟

الجواب: من بين مقاييس النزعة المركزية الثلاث: المتوسط الحسابي، الوسيط والمنوال نجد أن المتوسط الحسابي يأخذ بعين الاعتبار كل القيم عند حسابه، ولهذا يعتبر مقدرا كافيا لمعلمة المجتمع μ. على عكس الوسيط الذي لا يستخدم في حسابه سوى قيم المفردات التي تقع في وسط العينة، كذلك الأمر بالنسبة للمنوال حيث يتم الاعتماد على القيم الأكثر تكرارا عند حسابه.

ملاحظة: يتضح مما سبق، أنه إذغ توفرت الخصائص الأربعة (عدم التحيز، الإتساق، الكفاءة والكفاية) في التقدير المحسوب من بيانات العينة فإن التقدير هو تقدير جيد.

ثانيا: التقدير بمجالات الثقة Confidence interval estimation

رأينا سابقا أنه إذا قدرت معلمة المجتمع بقيمة واحدة فهذا يسمى بتقدير المعلمة بنقطة واحدة، وهذا التقدير نادراً ما يساوي المعلمة التي نرغب في تقديرها.

عموما لا يمكن توقع الحصول على تقدير لمعلمة المجتمع بدون خطأ مهما كان التقدير جيدا، لذلك فإننا نحدد فترة تحتوي على مجموعة من القيم تتضمن فيما بينها قيمة معلمة المجتمع، وتسمى هذه الفترة بتقدير فترة الثقة، وإحتمال وقوع المعلمة في هذه الفترة يسمي درجة الثقة وذلك كما يلي:

$$\hat{\theta}_L \leq \hat{\theta} \leq \hat{\theta}_U$$

حيث:

$(\hat{\theta}_L \ , \ \hat{\theta}_U)$: يسمى مجال الثقة أو فترة الثقة للمعلمة θ.

$\hat{\theta}_L$: الحد الأدنى لفترة الثقة .

$\hat{\theta}_U$: الحد الأعلى لفترة الثقة.

ويتم الإعتماد على بيانات العينة في حسابهما، كما نعتمد على الإحصائية المستخدمة كأفضل مقدر بالقيمة للمعلمة θ، وعلى توزيع المعاينة لهذا المقدر وعلى خطئه المعياري، وعلى حجم العينة، وعلى معامل الثقة المرغوب فيه $(1-\alpha)$.

وحيث أن درجة الثقة محصورة بين الصفر والواحد، أي: $0 < 1-\alpha < 1$ فإن ذلك يعني:

$$P(\hat{\theta}_L \leq \hat{\theta} \leq \hat{\theta}_U) = 1-\alpha$$

ملاحظات مهمة:

- نرمز لدرجة الثقة بالرمز $(1-\alpha)$، ومكمل هذه القيمة يسمى مستوى المعنوية ويرمز له بالرمز α.
- غالبا ما نستخدم القيم الأكثر شيوعا لمستوى المعنوية α وهي: **1%، 5% و10%**. ويتوقف الاختيار في ما بينها على حسب نوعية الظاهرة المدروسة.

- الإنحراف المعياري لتوزيع المعاينة للمقدر يسمى بالخطأ المعياري.

2-1- مجال الثقة لمتوسط المجتمع μ Confidence Interval on the Mean

2-1-1- عندما يكون تباين المجتمع σ^2 معلوما Variance Known

توصلنا سابقا إلى أنه:

- إذا كان لدينا مجتمع إحصائي متوسطه μ وتباينه σ^2 بحيث $n \geq 30$ فإنه بغض النظر عن توزيع المجتمع من خلال (نظرية النهاية المركزية: الفصل الأول، النظرية 1.10) أنه إذا سحبنا منه عينة حجمها n فإن توزيع المعاينة للمتوسط الحسابي للعينة $\bar{X}$ هو : $\overline{X} \sim N(\mu, \delta/\sqrt{n})$
- عندما يكون مجتمع موزع توزيعا طبيعيا بمتوسط μ وتباين σ^2 معلوم أي: $X \sim N(\mu, \sigma^2)$ فإنه وبغض النظر عن حجم العينة n التي تسحب من هذا المجتمع، فإن توزيع المعاينة للمتوسط الحسابي للعينة $\bar{X}$ هو الآخر يتبع التوزيع الطبيعي أي : $\overline{X} \sim N(\mu, \delta/\sqrt{n})$.

في الحالات المذكورة سابقا تعطى القيمة المعيارية Z للمتوسط $\bar{X}$ كما يلي :

$$Z = \frac{\overline{X} - \mu}{\sqrt{\delta^2 / n}} = \frac{\overline{X} - \mu}{\delta / \sqrt{n}} = \frac{\sqrt{n} * (\overline{X} - \mu)}{\delta}$$

وبالإستعانة بالقيمة المعيارية، نعرّف فترة الثقة بأنها الفترة التي تحقق:

$$P\left(-Z_{1-\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة الموجودة بين القوسين، بحيث نجعل الحد الأوسط للمتباينة يحتوي على المعلمة المجهولة فقط، ويتم ذلك بضرب الأطراف الثلاثة للمتباينة في $\sqrt{\sigma^2/n}$ نحصل على ما يلي:

$$P\left(-Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n} \leq \bar{X} - \mu \leq Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right) = 1 - \alpha$$

وعند طرح $\bar{X}$ من الأطراف الثلاثة للمتباينة نجد أن :

$$P\left(-\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n} \leq -\mu \leq -\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right) = 1 - \alpha$$

بضرب الأطراف الثلاثة للمتباينة في (1-) نحصل على:

$$P\left(\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n} \geq \mu \geq \bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right) = 1 - \alpha$$

نقوم بإعادة ترتيب المتباينة نجد أن:

$$P\left(\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n} \le \mu \le \bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right) = 1-\alpha$$

وتمثل هذه العلاقة الأخيرة فترة ثقة لمتوسط المجتمع μ (المجهول) بإحتمال يساوي $(1-\alpha)$:

$$\left(\bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right) \text{ و } \left(\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right)$$

يمكن كتابة فترة الثقة السابقة على الشكل التالي: $\mu = \overline{X} \pm Z_{1-\alpha/2} * \frac{\delta}{\sqrt{n}}$ أو على الشكل:

$$\left[\overline{X} - Z_{1-\alpha/2} * \frac{\delta}{\sqrt{n}},\quad \overline{X} + Z_{1-\alpha/2} * \frac{\delta}{\sqrt{n}}\right]$$

ملاحظات مهمة:

- **المقدار** $Z_{1-\alpha/2}$ يسمى معامل الثقة، وهو يستخرج من جدول التوزيع الطبيعي المعياري.
- يسمى المقدار $Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}$ بحد الخطأ، وهو يعني أنّ متوسط العينة $\bar{X}$ يختلف عن متوسط المجتمع μ بمقدار $Z_{1-\frac{\alpha}{2}}$ مضروبا في الخطأ المعياري $\sigma_{\bar{X}} = \sqrt{\sigma^2/n}$، يرمز له عادة بالرمز ϵ.
- في حالة كان المجتمع محدود (السحب بدون إرجاع أو $\frac{n}{N} \ge 0.05$)، نقوم بإضافة معامل التصحيح $\left(\frac{N-n}{N-1}\right)$ في حساب تباين العينة.
- عند حساب حجم العينة التي ينبغي أخذها بحيث لا يتجاوز حد الخطأ في التقدير لـ μ بـ $\bar{X}$ المقدار ϵ والذي يتم تحديده مسبقا من طرف الباحث، نعتمد على حل المتباينة $\epsilon \ge Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}$ حيث:

$$\epsilon \ge Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n} \Rightarrow \epsilon^2 \ge \left(Z_{1-\frac{\alpha}{2}}\sqrt{\sigma^2/n}\right)^2 \Rightarrow \epsilon^2 \ge (Z_{1-\frac{\alpha}{2}})^2 \,.\, \sigma^2/n \Rightarrow n \ge (Z_{1-\frac{\alpha}{2}})^2 \,.\, \sigma^2/\epsilon^2.$$

- عندما يكون تباين المجتمع مجهول نستبدله بتباين العينة.

مثال (2- 6):

سحب عينة عشوائية حجمها يساوي 64 ومتوسطها يساوي 10 من مجتمع يتبع التوزيع الطبيعي انحرافه المعياري يساوي 4، أوجد فترة ثقة بنسبة 95% لمتوسط المجتمع.

الحل:

من معطيات المثال لدينا:

$\sigma = 4$ $n = 64$ $\bar{X} = 10$ $1 - \alpha = 95\%$

حجم العينة $n = 64 > 30$ وتباين المجتمع الطبيعي معلوم، أي أننا نستخدم التوزيع الطبيعي في إيجاد فترة الثقة لمتوسط المجتمع μ عند مستوى ثقة $1 - \alpha = 95\%$.

ومنه:

$$\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975 \Rightarrow Z_{1-\frac{\alpha}{2}} = Z_{0.975}$$

بالبحث في جدول التوزيع الطبيعي المعياري نجد $Z_{0.975} = 1.96$

أي أن فترة الثقة لـ μ هي:

$$\bar{X} - Z_{1-\alpha/2} * \frac{\delta}{\sqrt{n}} \le \mu \le \bar{X} + Z_{1-\alpha/2} * \frac{\delta}{\sqrt{n}}$$

$$10 - 1.96 * \frac{4}{\sqrt{64}} \le \mu \le 10 + 1.96 * \frac{4}{\sqrt{64}} \Leftrightarrow 10 - 1.96 * \frac{4}{8} \le \mu \le 10 + 1.96 * \frac{4}{8}$$

$$10 - 0.98 \le \mu \le 10 + 0.98 \Leftrightarrow 9.02 \le \mu \le 10.98$$

ومنه يقع متوسط المجتمع μ بين القيمتين 9.02 و10.98 وذلك باحتمال قدره **95%.**

2-1-2- عندما يكون تباين المجتمع σ^2 مجهولا Variance Unknown

في هذه الحالة نهتم بحجم العينة حيث نميز بين حالتين: العينات الكبيرة ($n \ge 30$) أو العينات الصغيرة ($n < 30$).

- **في حالة العينات الكبيرة** ($n \ge 30$):

نستخدم تباين العينة كتقدير لتباين المجتمع، ويبقى التوزيع توزيع طبيعي. وبذلك يكون مجال الثقة لمتوسط المجتمع μ على النحو الأتي:

$$\bar{X} - Z_{1-\frac{\alpha}{2}} * \frac{S}{\sqrt{n}} \le \mu \le \bar{X} + Z_{1-\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}$$

مثال (2- 7): أخذت عينة حجمها 49 بمتوسط يسلوي **14**وتباين يسلوي **16** من مجتمع طبيعي التوزيع بتباين مجهول. المطلوب: أوجد فترة ثقة لمتوسط المجتمع عند مسقوى معنوية تسلوي **5%.**

الحل:

لدينا المعطيات:

$\alpha = 5\%$ $S^2 = 16$ $\bar{X} = 14$ $n = 49$ $\sigma^2 = ?$

تباين المجتمع الطبيعي مجهول، حجم العينة $n = 49 > 30$ لذا سوف نستخدم تباين العينة كتقدير لتباين المجتمع ونستخدم التوزيع الطبيعي لإيجاد فترة الثقة لمتوسط المجتمع μ عند مستوى معنوية $\alpha = 5\%$.

وحيث أنّ:

$$1 - \alpha = 0.95 \qquad \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1 - \frac{\alpha}{2} = 0.975 \Rightarrow Z_{1-\frac{\alpha}{2}} = Z_{0.975}$$

بالإستعانة بجدول التوزيع الطبيعي المعياري نجد أن : $Z_{0.975} = 1.96$

ومنه تكون فترة الثقة لـ μ هي:

$$\bar{X} - Z_{1-\frac{\alpha}{2}}\sqrt{s^2/n} \leq \mu \leq \bar{X} + Z_{1-\frac{\alpha}{2}}\sqrt{s^2/n}$$
$$14 - 1,96\sqrt{16/49} \leq \mu \leq 14 + 1,96\sqrt{16/49}$$
$$12.9 \leq \mu \leq 15.1$$

- **في حالة العينات الصغيرة** $(n < 30)$:

نستبدل تباين المجتمع بتباين العينة، والتوزيع هو توزيع ستودنت بدرجة حرية $(n - 1)$. وبذلك يكون مجال الثقة لمتوسط المجتمع μ كما يلي:

$$\bar{X} - t_{(1-\frac{\alpha}{2},n-1)}\sqrt{s^2/n} \leq \mu \leq \bar{X} + t_{(1-\frac{\alpha}{2},n-1)}\sqrt{s^2/n}$$

مثال (2- 8):

تم سحب عينة حجمها **25** بمتوسط يساوي **18** وتباين يساوي **36** من مجتمع طبيعي التوزيع بتباين مجهول.

المطلوب: أوجد مجال الثقة لمتوسط المجتمع عند مسوى معنوية تساوي **10%**.

الحل:

من معطيات المثال لدينا:

$\alpha = 10\%$ $S^2 = 36$ $\bar{X} = 18$ $n = 25$ $\sigma^2 = ?$

وبما أنّ تباين المجتمع الطبيعي مجهول وحجم العينة $n = 25 < 30$ فإننا سوف نستبدل تباين المجتمع بتباين العينة ونستخدم توزيع ستودنت في إيجاد مجال الثقة لمتوسط المجتمع μ عند مستوى معنوية $\alpha = 10\%$.

وحيث أنّ:

$$\alpha = 0.10 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow 1 - \frac{\alpha}{2} = 0.95 \Rightarrow t_{(1-\frac{\alpha}{2},n-1)} = t_{(0.95,24)}$$

من جدول توزيع ستودنت نجد $t_{(0.95,24)} = 1.711$

ويكون مجال الثقة لـ μ هو:

$$\bar{X} - t_{(1-\frac{\alpha}{2},n-1)}\sqrt{s^2/n} \leq \mu \leq \bar{X} + t_{(1-\frac{\alpha}{2},n-1)}\sqrt{s^2/n}$$

$$18 - 1.711\sqrt{36/25} \leq \mu \leq 18 + 1.711\sqrt{36/25}$$

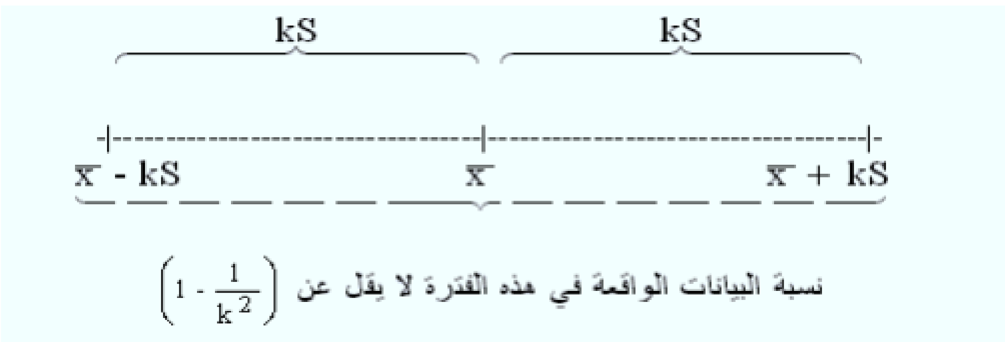
$$15.9 \leq \mu \leq 20.1$$

3-1-3- التقدير بمجال ثقة لمتوسط باستخدام نظرية تشيبتشيف Chebyshev/Confidence Interval

تعتبر نظرية تشيبيشيف من النظريات المفيدة وذات أهمية، إذ أنها تعطينا حد أدنى لنسبة البيانات الواقعة في مجال معين عند معرفة متوسطها وانحرافها المعياري دون الحاجة إلى معرفة البيانات الأصلية أو طبيعة التوزيع الذي أخذت منها العينة.

<u>نظرية</u>: حسب متراجحة تشيبيتشيف، إذا كان لدينا عينة من البيانات متوسطها $\overline{X}$ وانحرافها المعياري S، فإن المتباينة التالية تكون محققة: $P(|\overline{X} - \mu| \leq ks) \geq 1 - \frac{1}{k^2}$. أي أن نسبة البيانات الواقعة في المجال $\left[\overline{X} - ks; \overline{X} + ks\right]$ لا يقل عن $1 - \frac{1}{k^2}$، حيث: $k \succ 1$.

الشكل التالي يوضح فكرة نظرية تشيبيشيف.



مثال (2- 9): ما هي نسبة البيانات الواقعة في المجال $[-4;18]$، إذا علمت أن متوسط هذه البيانات يسوي 7 تباينها يسوي 25؟

<u>الحل</u>: من معطيات المثال نجد أن توزيع المجتمع مجهول، انحرافه المعياري مجهول وحجم العينة أقل من 30.

$$\begin{cases} \overline{X} - k*s = 18 \\ \overline{X} - k*s = -4 \end{cases} \Leftrightarrow \begin{cases} 7 + k*5 = 18 \\ 7 - k*5 = -4 \end{cases} \Leftrightarrow k*5 = 11 \Leftrightarrow k = \frac{11}{5} = 2.2$$

ومنه بتعويض قيمة K نجد: $1 - \frac{1}{k^2} = 1 - \frac{1}{(11/5)^2} = 0.79$

إذا نسبة البيانات الواقعة في المجال المعطى في المثال لا تقل عن 79 %.

2-2- التقدير بمجال الثقة للفرق بين متوسطي مجتمعين *مستقلين* Confidence Interval on the difference in Two Means

نميز بين حالتين:

2-2-1- عندما يكون تباين المجتمعين معلومين (σ_2^2 و σ_1^2 معلومين) Variances Known

رأينا فيما سبق، النظرية (1-11) ووفق توزيع المعاينة للفرق بين متوسطي مجتمعين طبيعيين:

- $\boldsymbol{\mu_{\bar{X}_1-\bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2}$
- $\boldsymbol{\sigma^2_{\bar{X}_1-\bar{X}_2} = \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

أي أن:

$$(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

ومنه القيمة المعيارية :

$$Z = \frac{(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0.1)$$

ونعلم أن فترة الثقة كما يلي:

$$P(-Z_{1-\frac{\alpha}{2}} \leq Z \leq Z_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

بالتعويض عن قيمة Z نجد:

$$P(-Z_{1-\frac{\alpha}{2}} \leq \frac{(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq Z_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة، نجد:

$$P[(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq (\mu_1 - \mu_2) \leq (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}] = 1 - \alpha$$

يمكن كتابة فترة الثقة للفرق بين متوسطي المجتمعين كما يلي:

$$(\mu_1-\mu_2)=\ (\bar{X}_1-\bar{X}_2)\pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

مثال (2- 9):

تمتلك إحدى الشوكات مزرعتين لإنتاج البرتقال في البليدة والعاصمة، أردنا مقارنة متوسط كميات الإنتاج لشجيرات البرتقال في مزرعة **البليدة** بكميات الإنتاج لشجيرات البرتقال في مزرعة **العاصمة**، ولهذا الغرض سحبت عينة عشوائية من شجيرات البرتقال من مزرعة **البليدة** حجمها **25** شجيرة، وسحبت عينة عشوائية من مزرعة **العاصمة** حجمها **32** شجيرة، **فوجدنا** متوسط كمية الإنتاج لهما هو: **50** كلغ و **40** كلغ على التوالي. إذا كانت كميات الإنتاج لجميع شجيرات البرتقال في المزرعتين تتوزع طبيعيا بتباين يساوي **16** كلغ، يساوي **9** كلغ على التوالي، أوجد فترة ثقة للفرق بين متوسطي كميات الإنتاج في المزرعتين عند مسوى ثقة **90%**.

الحل:

حسب معطيات المثال لدينا: مزرعة البليدة تمثل المجتمع 1، مزرعة العاصمة تمثل المجتمع 2.

$\alpha=10\%$ $\quad\sigma_1^2=16$ $\quad\bar{X}_1=50$ $\quad n_1=25$

$1-\alpha=90\%$ $\quad\sigma_2^2=9$ $\quad\bar{X}_2=40$ $\quad n_2=32$

المجتمعان يتوزعان طبيعيا وهما مستقلان وتباينهما معلومان، أي أن فترة الثقة لـ $\mu_1-\mu_2$ هو:

$$(\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ \leq(\mu_1-\mu_2)\leq\ (\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

بالتعويض نجد:

$$10-1.64\sqrt{\frac{16}{25}+\frac{9}{32}}\ \leq(\mu_A-\mu_B)\leq\ 10+1.64\sqrt{\frac{16}{25}+\frac{9}{32}}$$

$$8.4\ \leq(\mu_1-\mu_2)\leq 11.6$$

2-2-2- عندما يكون تباين المجتمعين مجهولين (σ_2^2 و σ_1^2 مجهولين) Variances Unknown

في هذه الحالة نستخدم تباين العينتين S_2^2 و S_1^2 كتقدير لتباين المجتمعين σ_2^2 و σ_1^2. ونميز بين حالتين:

2-2-2-1- عندما يكون σ_2^2 و σ_1^2 مجهولين وغير متساويين

- **في حالة العينات الكبيرة أي $n_1, n_2 \geq 30$:** يكون مجال الثقة للفرق بين متوسطي مجتمعين كالتالي:

$$(\bar{X}_1 - \bar{X}_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \leq (\mu_1 - \mu_2) \leq (\bar{X}_1 - \bar{X}_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

- **في حالة العينات الصغيرة أي $n_1, n_2 < 30$ أو أحدهما صغير:** يكون مجال الثقة للفرق بين متوسطي مجتمعين كالتالي:

$$(\bar{X}_1 - \bar{X}_2) - t_{1-\frac{\alpha}{2}, n-2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \leq (\mu_1 - \mu_2) \leq (\bar{X}_1 - \bar{X}_2) + t_{1-\frac{\alpha}{2}, n-2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

حيث:

$$n = n_1 + n_2$$

مثال (2- 10):

من مجتمعين طبيعيين مستقلين بتباينين غير متساويين تم سحب عينة عشوائية حجمها **50** مشاهدة من المجتمع الأول بمتوسط يسوي **20** وتباين يسوي **25**، كما تم سحب عينة أخوى حجمها **60** مشاهدة من المجتمع الثاني بمتوسط يسوي **15** وتباين يسوي **18**. والمطلوب، عند مسوى ثقة 99%:

1- أوجد مجال الثقة للفوق بين متوسطي المجتمعين.

2- أعد الإجابة على السؤال الأول علما أن حجم العينة الأولى يسوي **28** والثانية **24**.

الحل:

- **الحالة الأولى:**

من المثال لدينا المعطيات التالية:

$\alpha = 1\%$	$S_1^2 = 25$	$\bar{X}_1 = 20$	$n_1 = 50$
$1 - \alpha = 99\%$	$S_2^2 = 18$	$\bar{X}_2 = 15$	$n_2 = 60$

نحن في حالة العينات الكبيرة لأن حجم العينتين كبير، تباين المجتمعين مجهول، نستخدم تباين العينتين كتقدير لتباين المجتمعين. يصبح مجال الثقة للفرق بين متوسطي المجتمعين هو:

$$(\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\ \leq(\mu_1-\mu_2)\leq\ (\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

بالتعويض العددي نجد:

$$5-2.58\sqrt{\frac{25}{50}+\frac{18}{60}}\ \leq(\mu_1-\mu_2)\leq\ 5+2.58\sqrt{\frac{25}{50}+\frac{18}{60}}$$

فترة الثقة هي: $2.7\ \leq(\mu_1-\mu_2)\leq\ 7.3$

- **الحالة الثانية عندما حجم العينة الأولى يساوي 28 والثانية 24**

تصبح معطيات المثال (2- 10) كما يلي:

$n_1=28$ $\quad\bar{X}_1=20$ $\quad S_1^2=25$ $\quad\alpha=1\%$

$n_2=24$ $\quad\bar{X}_2=15$ $\quad S_2^2=18$ $\quad 1-\alpha=$

99%هنا نحن في حالة العينات الصغيرة لأن حجم العينتين صغير، تباين المجتمعين مجهول،

نستخدم تباين العينتين كتقدير لتباين المجتمعين. يصبح مجال الثقة للفرق بين متوسطي المجتمعين هو:

$$(\bar{X}_1-\bar{X}_2)-t_{1-\frac{\alpha}{2},n-2}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\ \leq(\mu_1-\mu_2)\leq\ (\bar{X}_1-\bar{X}_2)+t_{1-\frac{\alpha}{2},n-2}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

بالتعويض نجد:

$$5-2,678\sqrt{\frac{25}{28}+\frac{18}{24}}\ \leq(\mu_1-\mu_2)\leq\ 5+2,678\sqrt{\frac{25}{28}+\frac{18}{24}}$$

فترة الثقة هي: $1.6\ \leq(\mu_1-\mu_2)\leq\ 8.4$

2-2-2-2- عندما يكون σ_2^2 و σ_1^2 مجهولين و متساويين أي $\delta^2=\delta_1^2=\delta_2^2$

إذا كان لدينا مجتمعين طبيعيين $X_1\sim N(\mu_1,\sigma_1^2)$ حيث μ_1 و σ_1^2 مجهولان، و $X_2\sim N(\mu_2,\sigma_2^2)$ حيث μ_2 و σ_2^2 مجهولان أيضا، وتم سحب عينتين عشوائيتين مستقلتين من هذين المجتمعين حجمهما n_1 و n_2 على التوالي، وكان تباين المجتمع الأول وتباين المجتمع الثاني مجهولين ومتساويين أي أن $\delta^2=\delta_1^2=\delta_2^2$. وبالاعتماد على توزيعات المعاينة نعلم أن أفضل مقدر لتباين المجتمع هو تباين العينة المسحوبة منه، ومنه فإن أفضل مقدر لـ σ_1^2 و σ_2^2 هو S_1^2 و S_2^2 على التوالي، لذلك نقدّر التباينين بنفس المقدّر (التباين المشترك)، وهو عبارة عن الوسط المرجّح لتباين العينة الأولى وتباين العينة الثانية، ويرمز له بالرمز S_P^2، حيث:

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

وباستبدال قيمة كل من σ_1^2 و σ_2^2 بـ S_P^2 في قيمة الإحصائية $Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ نحصل على متغير

عشوائي آخر يتبع التوزيع الإحتمالي ستودنت بدرجة حرية n_1+n_2-2، يرمز له بـ T حيث :

$$T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_P\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\sim t(\vartheta=n_1+n_2-2)$$

وبالتعويض عن قيمة T نجد مجال الثقة للفرق بين المتوسطين $(\mu_1-\mu_2)$ في هذه الحالة هو:

$$(\bar{X}_1-\bar{X}_2)-t_{1-\frac{\alpha}{2},\vartheta}*S_P\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\leq(\mu_1-\mu_2)\leq(\bar{X}_1-\bar{X}_2)+t_{1-\frac{\alpha}{2},\vartheta}*S_P\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

ملاحظة هامة:

إن لم يتوفر شرط التوزع الطبيعي للمجتمعين وفي حالة العينات الصغيرة أي: $n_1,n_2\geq 30$، وطبقا لنظرية النهاية المركزية نستطيع استخدام مجال الثقة الأخير للفرق بين المتوسطين $(\mu_1-\mu_2)$.

مثال (2- 11):

أخذت عينة عشوائية حجمها 12 فكان متوسطها 25وانحوافها المعياري 4 من مجتمع طبيعي،وأخذت عينة عشوائية أخرى حجمها 8 فكان متوسطها 35وانحوافها المعياري 3.7 من مجتمع طبيعي مستقل عن المجتمع السابق. أوجد مجال الثقة للفرق بين متوسطي المجتمعين عند مسوى الثقة 95% إذا علمت أن تبايني المجتمعين مجهولين وهما متساويين.

الحل: بما أن المجتمعين يتوزعان طبيعيا بتباينين مجهولين ومتساويين، العينتين مستقلتين وحجمهما أقل من 30 فإن مجال الثقة في هذه الحالة يعطي بالعلاقة التالية:

$$(\bar{X}_1-\bar{X}_2)-t_{1-\frac{\alpha}{2},\vartheta}*S_P\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\leq(\mu_1-\mu_2)\leq(\bar{X}_1-\bar{X}_2)+t_{1-\frac{\alpha}{2},\vartheta}*S_P\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

أولا: نقوم بحساب التباين المشترك S_P^2.

$$S_P^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{(n_1-1)+(n_2-1)}=\frac{(12-1)\,16+(8-1)\,14}{(12-1)+(8-1)}=15.22$$

ولدينا:

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1-\frac{\alpha}{2} = 0.975 \Rightarrow t_{1-\frac{\alpha}{2},\vartheta} = t_{0.975,18} = 2.101$$

ثانيا: فترة الثقة هي

$$(25-35) - 2.101 * 5.02\sqrt{\left(\frac{1}{12}+\frac{1}{8}\right)} \le (\mu_1-\mu_2) \le (25-35) + 2.101 * 5.02\sqrt{\left(\frac{1}{12}+\frac{1}{8}\right)}$$

$$-14.8 \le (\mu_1-\mu_2) \le -5.2$$

2-3- التقدير بمجال الثقة للفرق بين متوسطي مجتمعين مرتبطين

هناك حالات كثيرة ترتبط فيها البيانات المتعلقة بالعينتين أي أن العينتين مرتبطتين، هنا نميز بين حالتين:

- **في حالة العينات الكبيرة** $n \ge 30$:

في هذه الحالة تكون المتغيرة: $Z = \frac{\bar{D}-\mu_D}{\sqrt{S_D^2/n}}$ ، ويعطى مجال الثقة بالعلاقة الأتية:

$$P(\bar{D} - Z_{1-\frac{\alpha}{2}}\sqrt{S_D^2/n} \le \mu_1-\mu_2 \le \bar{D} + Z_{1-\frac{\alpha}{2}}\sqrt{S_D^2/n}) = 1-\alpha$$

- **في حالة العينات الصغيرة** $n < 30$:

في هذه الحالة تكون المتغيرة: $T = \frac{\bar{D}-\mu_D}{\sqrt{S_D^2/n}} \sim t(n-1)$ ، ويعطى مجال الثقة بالعلاقة الأتية:

$$P(\bar{D} - t_{1-\frac{\alpha}{2},n-1}\sqrt{S_D^2/n} \le \mu_1-\mu_2 \le \bar{D} + t_{1-\frac{\alpha}{2},n-1}\sqrt{S_D^2/n}) = 1-\alpha$$

مثال (2- 12): تم إجراء إمتحان لعينة مكونة من **10** معلمين لمعرفة مدى تحكمهم في أدوات التعليم عن بعد وسجلت علاماتهم، ثم أخضعوا لدورة تكوينية لمدة شهر ليتم بعدها إعادة الامتحان عقب إنتهاء الدورة، فكانت النتائج كالتالي:

قبل الدورة X_i	13	10	12	9	6	8	11	14	7	6
بعد الدورة Y_i	16	15	16	15	11	12	9	16	12	10

والمطلوب: أوجد مجال الثقة لفرق متوسطي الوزنين عند مستوى معنوية تساوي 5%.

الحل:

إعتمادا على بيانات المثال يمكن إيجاد الفروق كما يلي:

قبل الامتناع X_i	13	10	12	9	6	8	11	14	7	6
بعد الامتناع Y_i	16	15	16	15	11	12	13	16	12	10
الفروق D_i	3-	5-	4-	6-	5-	4-	2-	2-	5-	4-

ومنه يمكن حساب المتوسط والتباين للفروق كما يلي:

$$\bar{\mathbf{D}} = \frac{\sum D_i}{n} = \frac{-40}{10} = -4$$

$$S_D^2 = \frac{\sum(D_i - \mu_D)^2}{n-1} = \frac{16}{9} = 1.78$$

وبذلك نحصل على:

$$T = \frac{\bar{D} - \mu_D}{\sqrt{\frac{S_D^2}{n}}} = \frac{\bar{D} - (-4)}{\sqrt{\frac{1.78}{10}}} \sim t(9)$$

باستعمال جدول توزيع ستودنت نجد أن:

$$t_{(1-\frac{\alpha}{2},n-1)} = t_{(0.975,9)} = 2.262$$

وعليه مجال الثقة المطلوب هو:

$$P(\bar{D} - t_{1-\frac{\alpha}{2},n-1}\sqrt{S_D^2/n} \leq \mu_1 - \mu_2 \leq \bar{D} + t_{1-\frac{\alpha}{2},n-1}\sqrt{S_D^2/n}) = 1 - \alpha$$

$$-4 - 2.262\sqrt{1.78/10} \leq \mu_1 - \mu_2 \leq -4 + 2.262\sqrt{1.78/10}$$

$$-4.95 \leq \mu_1 - \mu_2 \leq -3.05$$

2- 4- التقدير بمجال الثقة لنسبة المجتمع P Confidence Interval on a Proportion

راينا سابقا في الفصل الخاص بتوزيعات المعاينة، أنّ توزيع المعاينة لنسبة العينة $\widehat{P}$ يقترب من التوزيع الطبيعي كلما كان حجم العينة n كبيرا، وتحقق الشرط $np \geq 5$ و $nq \geq 5$ ، وذلك بمتوسط حسابي $\mu_{\widehat{P}} = P$ وتباين $\sigma_{\widehat{P}}^2 = \frac{P.q}{n}$.

قيمة المتغير Z الذي يتبع التوزيع الطبيعي المعياري تكون على الشكل:

$$Z = \frac{\widehat{P} - P}{\sqrt{P.q/n}}$$

وبما أنّ نسبة المجتمع P غير معلومة، وهي التي نريد تقدير فترة الثقة لها، فلا يمكن حساب التباين $\sigma_{\widehat{P}}^2 = \frac{P.q}{n}$ لذا نستخدم نسبة العينة $\widehat{P}$ المعلومة، ويصبح مقدّر التباين هو: $\widehat{\sigma}_{\widehat{P}}^2 = \frac{\widehat{P}.\widehat{q}}{n}$.

وتصبح قيمة المتغير Z **بعد التعويض بقيمة التباين الجديد** كما يلي:

$$Z = \frac{\widehat{P} - P}{\sqrt{\widehat{P}.\widehat{q}/n}}$$

ونعرّف فترة الثقة للنسبة بأنها الفترة التي تحقق ما يلي:

$$P\left(-Z_{1-\frac{\alpha}{2}} \leq \frac{\widehat{P} - P}{\sqrt{\widehat{P}.\widehat{q}/n}} \leq Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

وبإجراء العمليات الجبرية على الحدود الثلاثة للمتباينات تكون فترة الثقة هي:

$$P(\widehat{P} - Z_{1-\frac{\alpha}{2}}\sqrt{\widehat{P}.\widehat{q}/n} \leq P \leq \widehat{P} + Z_{1-\frac{\alpha}{2}}\sqrt{\widehat{P}.\widehat{q}/n}) = 1 - \alpha$$

أي أن:

$$\widehat{P} - Z_{1-\frac{\alpha}{2}}\sqrt{\widehat{P}.\widehat{q}/n} \leq P \leq \widehat{P} + Z_{1-\frac{\alpha}{2}}\sqrt{\widehat{P}.\widehat{q}/n}$$

ملاحظة هامة:

إذ كان المجتمع محدود والسحب بدون إرجاع، بمعنى تحقق الشرط $\frac{n}{N} \geq 0,05$، نستخدم معامل التصحيح.

وتعطى فترة الثقة على الشكل:

$$P(\widehat{P} - Z_{1-\frac{\alpha}{2}}\sqrt{\widehat{P}.\widehat{q}/n}\sqrt{N-n/_{N-1}} \leq P \leq \widehat{P} + Z_{1-\frac{\alpha}{2}}\sqrt{\widehat{P}.\widehat{q}/n}\sqrt{N-n/_{N-1}}) = 1 - \alpha$$

مثال (2- 13):

اختونا عينة عشوائية مكونة من **80** مطعم فيولاية الجلفة، فوجدنا **76** مطعم يمتلك أجهزة تكييف (Climatiseur). المطلوب: قدّر نسبة المطاعم التي تمتلك أجهزة تكييف فيولاية الجلفة باستخدام مستوى ثقة 90%.

الحل:

من المثال لدينا المعطيات التالية:

$$\alpha = 10\% \qquad 1 - \alpha = 90\% \qquad \widehat{P} = \frac{76}{80} = 0.95 \qquad n = 80$$

معامل الثقة:

$$\alpha = 10\% \Rightarrow \frac{\alpha}{2} = 0,05 \Rightarrow 1 - \frac{\alpha}{2} = 0,95 \Rightarrow Z_{0.95} = 1.64$$

فترة الثقة هي:

$$\hat{P} - Z_{1-\frac{\alpha}{2}}\sqrt{\hat{P}.\hat{q}/_{n}} \leq P \leq \hat{P} + Z_{1-\frac{\alpha}{2}}\sqrt{\hat{P}.\hat{q}/_{n}}$$

بالتعويض العددي نجد:

$$0,95 - 1,64\sqrt{\frac{(0,95).(0,05)}{80}} \leq \boldsymbol{P} \leq 0,95 + 1,64\sqrt{\frac{(0,95).(0,05)}{80}}$$

$$0,95 - 1,64\sqrt{\frac{(0,95).(0,05)}{80}} \leq \boldsymbol{P} \leq 0,95 + 1,64\sqrt{\frac{(0,95).(0,05)}{80}}$$

$$0,91 \leq \boldsymbol{P} \leq 0,99$$

2- 5- التقدير بمجال ثقة للفرق بين نسبتين Confidence Interval on the Difference in the Two Proportions

من بين نتائج توزيع المعاينة للفرق بين نسبتي عينتين، راينا أنه إذا سحبنا عينة كبيرة n_1 من مجتمع يخضع لتوزيع ذي الحدين $X_1 \sim B(n_1, P_1)$، وسحبنا عينة كبيرة أخرى n_2 من مجتمع آخر مستقل عن المجتمع الأول ويخضع لتوزيع ذي الحدين $X_2 \sim B(n_2, P_2)$، فإن الفرق بين النسبتين في العينتين $(\widehat{P_1} - \widehat{P_2})$ هو متغير عشوائي يخضع تقريبا للتوزيع الطبيعي بالخصائص التالية:

$$\mu_{\hat{P}_1-\hat{P}_2} = \mu_{\hat{P}_1} - \mu_{\hat{P}_2} = P_1 - P_2$$

وتباين

$$\sigma^2{}_{\hat{P}_1-\hat{P}_2} = \sigma^2{}_{\hat{P}_1} + \sigma^2{}_{\hat{P}_2} = \frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2} = \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}$$

ومنه المتغيرة:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} \sim N(0,1)$$

نظرا أنّ نسبة المجتمعين P_1 وP_2 مجهولتين، نستخدم $\sigma^2{}_{\hat{P}_1-\hat{P}_2}$ كتقدير لـ $\hat{\sigma}^2_{P_1-P_2}$.

وعليه تصبح قيمة المتغير Z هي:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - P_1 - P_2}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \sim N(0,1)$$

وبالتالي مجال الثقة للفرق بين النسبتين $(P_1 - P_2)$ يكون مساو إلى:

$$(\widehat{P_1} - \widehat{P_2}) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\hat{q}_1}{n_1} + \frac{\hat{P}_2\hat{q}_2}{n_2}} \leq P_1 - P_2 \leq (\widehat{P_1} - \widehat{P_2}) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\hat{q}_1}{n_1} + \frac{\hat{P}_2\hat{q}_2}{n_2}}$$

ملاحظة هامة:

إذ كان المجتمع محدود والسحب بدون إرجاع، بمعنى تحقق الشرط $\frac{n_1}{N_1} \geq 0,05$ و$\frac{n_2}{N_2} \geq 0,05$ يجب استعمال معامل التصحيح عند حساب فترة الثقة.

مثال (2- 14):

قامت إدارة كلية الاقتصاد بسحب عينتين عشوائيتين مستقلتين من قسم التسيير وقسم التجارة. فإذا كان حجم العينة الأولى **30** طالب، نجح منهم **16** طالب، وكان حجم العينة الثانية **40** طالب، نجح منهم **13** طالب.

المطلوب: أوجد فترة ثقة للفرق بين نسبتي الناجحين في القسمين عند مستوى معنوية **1%**.

الحل:

من المعطيات لدينا:

$$\boldsymbol{n_1 = 30} \qquad \hat{P}_1 = \frac{16}{30} = 0,53 \qquad \alpha = 1\%$$

$$\boldsymbol{n_2 = 40} \qquad \hat{P}_2 = \frac{13}{40} = 0,33$$

ومعامل الثقة هو:

$$\alpha = 1\% \Rightarrow 1 - \frac{\alpha}{2} = 0,995 \Rightarrow Z_{0.995} = 2.58$$

فترة الثقة للفرق بين النسبتين $(P_1 - P_2)$ يساوي إلى:

$$(\hat{P}_1 - \hat{P}_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\hat{q}_1}{n_1} + \frac{\hat{P}_2\hat{q}_2}{n_2}} \leq P_1 - P_2 \leq (\hat{P}_1 - \hat{P}_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\hat{q}_1}{n_1} + \frac{\hat{P}_2\hat{q}_2}{n_2}}$$

بالتعويض العددي نجد:

$$0,21 - 2,58\sqrt{\frac{(0,53)(0,47)}{30} + \frac{(0,33)(0,67)}{40}} \leq P_1 - P_2$$

$$\leq 0,21 + 2,58\sqrt{\frac{(0,53)(0,47)}{30} + \frac{(0,33)(0,67)}{40}}$$

$$-0.0951 \leq P_1 - P_2 \leq 0.5118$$

2- 6- التقدير بمجال الثقة لتباين المجتمع σ^2 Confidence Interval on the Variance

لإيجاد فترة ثقة لتباين المجتمع نحن بحاجة إلى التوزيع كاي مربع، حيث تكلمنا سابقا في توزيع المعاينة للتباين أنه إذا كان لدينا مجتمع طبيعي $X \sim N(\mu, \sigma^2)$ وكان S^2 هو تباين العينة، فإن $\frac{(n-1)S^2}{\sigma^2}$ يتبع التوزيع كاي مربع بدرجة حرية $n-1$ أي $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

ونميز بين الحالتين التاليتين عند كتابة فترة الثقة لتباين المجتمع:

- **الحالة الأولى متوسط المجتمع μ معلوم.**

يتم في هذه الحالة حساب تباين العينة بالاعتماد على متوسط المجتمع حيث: $$S^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n}$$

وتصبح فترة الثقة للتباين كما يلي:

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

- **الحالة الثانية متوسط المجتمع μ غير معلوم.**

في هذه الحالة يتم استخدام تباين العينة كتقدير لتباين المجتمع حيث: $$\hat{S}^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

وتصبح فترة الثقة للتباين كما يلي:

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

مثال (2- 15):

تم سحب عينة حجمها ***19*** من مجتمع يتوزع طبيعيا، فوجد أن تباينها يساوي **16**، المطلوب: أوجد مجال الثقة لتباين المجتمع عند مستوى معنوية ***10%***.

الحل:

من معطيات هذا المثال لدينا:

$\alpha = 10\%$ $\qquad S^2 = 16$ $\qquad n = 19$

من جدول توزيع مربع كاي نجد:

$$\chi^2_{(\frac{\alpha}{2},n-1)} = \chi^2_{0.05,18} = \mathbf{28.87} \qquad , \qquad \chi^2_{(1-\frac{\alpha}{2},n-1)} = \chi^2_{0.95,19} = \mathbf{9.39}$$

بالتعويض العددي نجد فترة الثقة:

$$\frac{18 \times 16}{28.87} \le \sigma^2 \le \frac{18 \times 16}{9.39}$$
$$\mathbf{9.98 \le \sigma^2 \le 30.67}$$

2-7- التقدير بمجال ثقة للنسبة بين تبايني مجتمعين Confidence Interval on the Ratio of Variances

عند المقارنة بين تبايني مجتمعين نعتمد على النسبة بين التباينيين لعينتين مأخوذتين منهما، حيث: إذا كان لدينا مجتمع $X_1 \sim N(\mu_1, \sigma_1^2)$ سحبت منه عينة عشوائية حجمها n_1، وكان $X_2 \sim N(\mu_2, \sigma_2^2)$ مجتمع ثاني سحبت منه عينة عشوائية حجمها n_2 تباينهما S_1^2 و S_2^2 على الترتيب فإنه ومن خلال توزيعات المعاينة في الفصل الأول لدينا:

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(\nu 1, \nu 2)$$

حيث: $\nu 1 = n_1 - 1$ و $\nu 2 = n_2 - 1$

نعرّف فترة الثقة للنسبة بين تبايني مجتمعين بالفترة التي تحقق:

$$P\left(F\left(1-\frac{\alpha}{2}, n_1 - 1, n_2 - 1\right) \le \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \le F\left(\frac{\alpha}{2}, n_1 - 1, n_2 - 1\right)\right) = 1 - \alpha$$

بحل هذه المتباينة بالنسبة لـ $\frac{\sigma_2^2}{\sigma_1^2}$ نجد أن فترة الثقة للنسبة بين تبايني المجتمعين كما يلي:

$$P\left(\frac{S_1^2}{S_2^2} \frac{1}{F\left(\frac{\alpha}{2}, n_1 - 1, n_2 - 1\right)} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2} F\left(\frac{\alpha}{2}, n_2 - 1, n_1 - 1\right)\right) = 1 - \alpha$$

مثال (2- 16):

تم سحب عينة عشوائية حجمها **20** بتباين يساوي **16** من مجتمع يتوزع طبيعيا، وسحبت عينة ثانية حجمها **16** بتباين يساوي **12** من مجتمع آخر يتوزع طبيعيا ومستقل عن المجتمع الأول. أوجد مجال الثقة للنسبة بين تبايني المجتمعين وذلك عند مستوى معنوية تساوي **5%**.

الحل:

لدينا المعطيات التالية:

$\boldsymbol{n_1 = 20}$ $\boldsymbol{S_1^2 = 16}$ $\boldsymbol{n_2 = 16}$ $\boldsymbol{S_2^2 = 12}$ $\boldsymbol{\alpha = 5\%}$

من جدول التوزيع فيشر عند α=0.025 نجد:

$$\frac{\mathbf{1}}{\boldsymbol{F\left(\frac{\alpha}{2}, n_2 - 1, n_1 - 1\right)}} = \frac{1}{F(0.025,15,19)} = \frac{1}{2.77} = \mathbf{0.3610}$$

$$\boldsymbol{F\left(\frac{\alpha}{2}, n_1 - 1, n_2 - 1\right)} = F(0.025,19,15) = \mathbf{2.62}$$

بالتعويض في مجال الثقة نجد:

$$\frac{S_2^2}{S_1^2}.\frac{1}{F\left(\frac{\alpha}{2}, n_2 - 1, n_1 - 1\right)} \leq .\frac{\sigma_2^2}{\sigma_1^2} \leq \frac{S_2^2}{S_1^2}.F\left(\frac{\alpha}{2}, n_1 - 1, n_2 - 1\right)$$

$$\frac{12}{16}.(0.3610) \leq .\frac{\sigma_2^2}{\sigma_1^2} \leq \frac{12}{16}.(2.62)$$

$$0.2708 \leq .\frac{\sigma_2^2}{\sigma_1^2} \leq 1.965$$

تمارين مقترحة

التمرين الأول:

- عرّف المقدر النقطي، وماذا يقصد بالمقدّر والتقدير؟.
- ماهي شروط المقدر الجيد؟

التمرين الثاني:

ليكن X1, X2 عينة من مجتمع إحصائي يتوزع طبيعيا حيث: (2, μ)N ~ X.

ولتكن المقدرات التالية للمتوسط معرفة بالعلاقات الاتية:

T1=1/2X1+1/2X2, T2=2/3X1+1/3X2, T3= X

أحسب الأمل الرياضي لكل من T1, T2, T3 ؟ ماذا تستنتج؟ أي المقدرات أكفأ؟

التمرين الثالث:

لتكن المتغيرتين:

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

أحسب ما يلي: E(S2) , E(S2)، أيهما يمثل تقدير غير متحيز للتباين(تباين المجتمع)؟

التمرين الرابع:

مصنع لإنتاج المصابيح الكهربائية، اختير من انتاجه عينة حجمها 100 مصباح وذلك لتقييم جودة الانتاج، حيث حدد مدير المصنع معيار الجودة ان يتراوح عمر المصباح بين 1000 و1100 .إذا كان المتوسط الحسابي لعمر المصباح في العينة 1200 ساعة وانحرافه المعياري 250 ساعة، قدر بدرجة ثقة 95% متوسط عمر المصباح من انتاج المصنع كله؟

التمرين الخامس:

إذا كان عدد الأسر ذات الدخل المحدود لعينة من 500 أسرة هو 200 اسرة.

قدر نسبة الأسر ذات الدخل المحدود بفترة ثقة عند 95% ؟

التمرين السادس:

1. لتقدير التفرقة في الاجور بين الرجال والنساء وجد لعينتين 80 رجل و 60 امراة أن متوسطي الدخول هما على الترتيب 20600 بانحراف معياري 30 و 19700 بانحراف معياري 25.

المطلوب: قدر الفرق بين الدخلين بفترة ثقة عند 95% باعتبار التوزيع الطبيعي؟

2. بافتراض ان نسبة الرجال ذوي الدخول اكثر من 10000 هي 40% وللنساء تساوي 30%.

المطلوب: قدر الفرق بين النسبتين عند درجة الثقة 95%؟.

التمرين السابع:

أوجد مجال ثقة لتقدير تباين مجتمع طبيعي عند مستوى ثقة 95% سحبت منه عينة حجمها 10 مفردات وتباينها $S^2 = 117.2$.

التمرين الثامن:

سحبت عينتين من مجتمعين مستقلين وتحصلنا على النتائج التالية:

حجم العينة الأولى $n_1 = 6$	تباين العينة الأولى $S_1^2 = 77.9$
حجم العينة الثانية $n_2 = 8$	تباين العينة الثانية $S_2^2 = 194.2$

أوجد مجال للثقة لـ S_1^2 / S_2^2 عند مستوى معنوية 5%؟

تمهيد

ينقسم الإحصاء إلى قسمين، إحصاء وصفي وإحصاء استنتاجي (استنباطي)، هذا الأخير أي الإحصاء الاستنتاجي بدوره ينقسم إلى قسمين وهما: التقدير واختبار الفرضيات الإحصائية. تحدثنا في الفصول السابقة عن الاستدلال الإحصائي عن طريق تقدير فترة ثقة. وقد استخدمنا بيانات العينة لتقدير قيمة معلمة المجتمع حيث أعطى التقدير قيمتان للفترة (مجال الثقة) والتي يقدر حدوث معلمة المجتمع ضمنها. وقد كان الافتراض بأن العينة هي مصدر كل المعلومات المتعلقة بالمعلمة مضمنا في إنشاء تقدير الفترة من عينة عشوائية. يقدم هذا الفصل طريقة أخرى للاستدلال الإحصائي حول معلمة المجتمع. وتسمى هذه الطريقة للاستدلال الإحصائي اختبار الفرضية.

الغرض من مثل هذا النوع من الاستدلال الإحصائي هو تحديد ما إذا كانت نتائج العينة تساند أي تدعم وتؤكد اعتقاد معين أو فرضية محددة حول قيمة معلمة المجتمع التي يحددها الباحث. ولهذا اختبار الفرضيات الإحصائية يعتبر أهم جزء في نظرية اتخاذ القرار، لذلك نحدد بدقة ماذا نعني بالفرضية الإحصائية والاختبارات المتعلقة بها.

في هذا الفصل سوف نقوم بدراسة العناصر الضرورية للقيام باختبار الفرضية.

أولا: المفاهيم المتعلقة باختبار الفرضيات

1-1- الفرضية الإحصائية Statistical Hypothesis:

الفرضية الإحصائية هي مقولة أو إدعاء تتعلق بالمجتمع الإحصائي تحتمل الصحة أو الخطأ وتنقسم الى قسمين:

- فروض حول معالم المجتمع Parametric
- فروض حول شكل تابع التوزيع Non Parametric

وإثبات صحة الفرضية من عدم صحتها لا يمكننا معرفته بدقة إلاّ إذا تفحصنا المجتمع الإحصائي بأكمله، وهذا صعب جدا خاصة إذا كان المجتمع كبير وغير محدود، لذا نختار عينة عشوائية من هذا المجتمع، ونستخدم المعلومات الموجودة فيها لنقرر ما إذا كانت الفرضية صحيحة أم خاطئة. إنّ الوصول إلى هذا الهدف لا يتم بصورة عشوائية وإنما يعتمد على طرق إحصائية علمية تسمى باختبار الفرضيات.

يتمثل اختبار الفرضيات بصياغة فرضية إحصائية ما، ونحن نهدف إلى رفضها أو إلغائها، هذه الفرضية تسمى بفرضية العدم (الفرضية الصفرية null hypothesis) ويرمز لها بالرمز H_0 ثم نقوم بصياغة فرضية جديدة أخرى تسمى بالفرضية البديلة (alternative hypothesis) ويرمز لها بالرمز H_1.

وتوجد ثلاث **(03)** صيغ ممكنة لاختبار الفرضيات وتكتب على الشكل التالي:

$$\begin{cases}\mathbf{H_0}: \mu \leq \mu_0 \\ \mathbf{H_1}: \mu > \mu_0\end{cases} , \quad \begin{cases}\mathbf{H_0}: \mu \geq \mu_0 \\ \mathbf{H_1}: \mu < \mu_0\end{cases} , \quad \begin{cases}\mathbf{H_0}: \mu = \mu_0 \\ \mathbf{H_1}: \mu \neq \mu_0\end{cases}$$

وفي حالة النسب نعوّض μ بـ P و μ_0 بـ P_0.

وتشير كل صورة للفرضية البديلة إلى جهة الاختبار ما إذا كان من جهة واحدة أو من جهتين. فالفرضية $\mathbf{H_1}: \mu \neq \mu_0$ تعني أنّ متوسط المجتمع لا يساوي μ_0 وإنما أكبر منها أو أصغر منها وهي تعبّر عن أنّ شكل الاختبار من جهتين، بينما الفرضية $\mathbf{H_1}: \mu < \mu_0$ تعني أنّ متوسط المجتمع أصغر من μ_0 وهي تعبّر عن أنّ شكل الاختبار من جهة اليسار، في حين أنّ الفرضية $\mathbf{H_1}: \mu > \mu_0$ تعني أنّ متوسط المجتمع أكبر من μ_0 بمعنى أنّ شكل الاختبار من جهة اليمين.

نشير في هذا الصدد إلى أنّ المساواة تظهر فقط في فرضية العدم، كما أنّ الشيء الذي نريد تصديقه يتم وضعه في الفرضية البديلة H_1.

1-2- الخطأ من النوع الأول والخطأ من النوع الثاني Type one and Type Two Errors

قد يحدث أحيانا أن نرفض الفرضية الصفرية وهي صحيحة أو نقبلها وهي خاطئة، وبذلك نكون أمام نوعين من الأخطاء:

1. **الخطأ من النوع الأول Type one Error** : هو الخطأ الناتج عن رفض فرضية العدم $\mathbf{H_0}$ وهي في الواقع صحيحة ونرمز له بالرمز $\boldsymbol{\alpha}$. وتسمى $\boldsymbol{\alpha}$ بمستوى المعنوية، ويحدد هذا المستوى من قبل الباحث حسب الدراسة وحسب الدقة المطلوبة. وتؤخذ قيمة $\boldsymbol{\alpha}$ في الحسابات كاملة إذا كان الاختبار في اتجاه واحد سواء من جهة اليمين أو من جهة اليسار وتقسّم لنصفين إذا كان شكل الاختبار في اتجاهين. وتحسب قيمة $\boldsymbol{\alpha}$ من العلاقة:

$$\boldsymbol{\alpha} = P(H_0 \text{ رفض} / \mathbf{H_1} \text{صحيحة})$$

ونقرأ: احتمال رفض الفرضية H_0 علما أنّ H_1 صحيحة.

2. **الخطأ من النوع الثاني Type Two Error** : هو الخطأ الناتج عن قبول فرضية العدم $\mathbf{H_0}$ وهي في الواقع خاطئة ونرمز له بالرمز $\boldsymbol{\beta}$. وتحسب قيمة $\boldsymbol{\beta}$ بالعلاقة التالية:

$$\boldsymbol{\beta} = P(H_0 \text{ قبول} / \mathbf{H_0} \text{خاطئة})$$

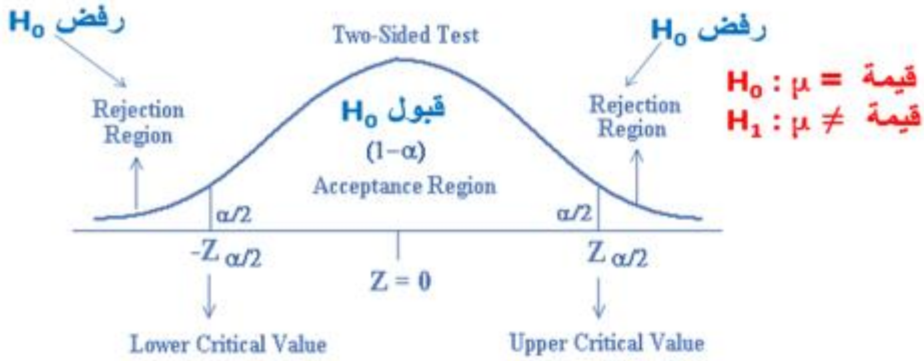
<u>ملاحظة هامة</u>: أخذ القرار برفض فرضية العدم وهي فرضية خاطئة فهذا يسمى بقوة الاختبار، ويرمز له بالرمز $(1-\beta)$ أي أن: احتمال (رفض H_0 / H_0 خاطئ) $= 1-\beta$

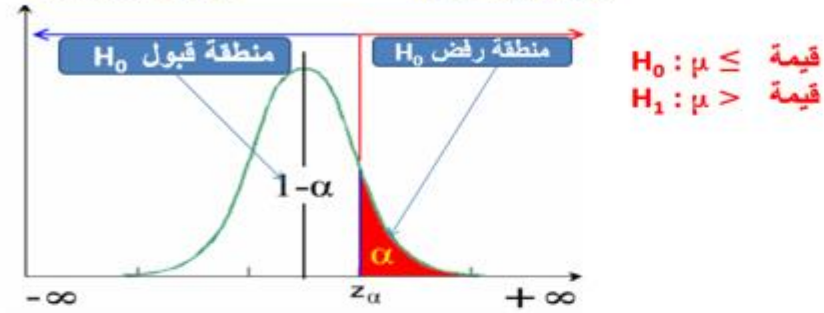
1-3- المنطقة الحرجة والقيمة الحرجة المعيارية:

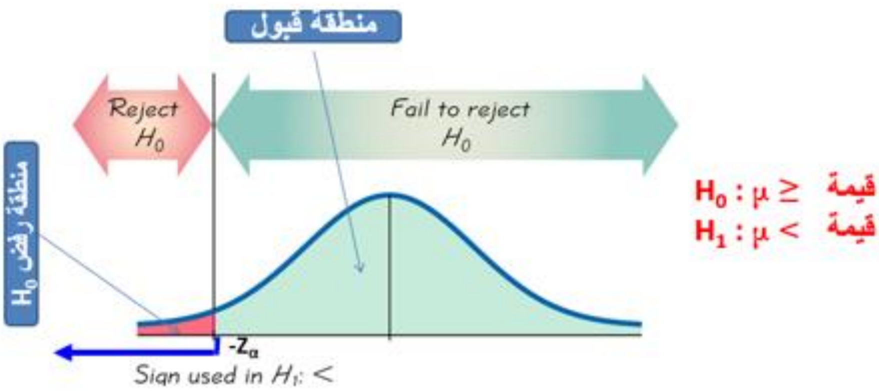
1-3-1- **المنطقة الحرجة:** تعرف بأنها المنطقة الواقعة تحت المنحني والتي تضم جميع القيم الممكنة أي المساحة التي تقع أسفل منحنى التوزيع الاحتمالي المستخدم في عملية التحليل الإحصائي، وهي تعني المنطقة التي اذا وقعت فيها القيمة المحسوبة ترفض الفرضية الصفرية. اما اذا وقعت القيمة المحسوبة خارجها فأنه يتم قبول الفرضية الصفرية، وتسمى أيضا بمنطقة الرفض (rejection region) ويحدد قيمتها مستوى المعنوية $\boldsymbol{\alpha}$.

منطقة الرفض اما ان تقع على جهة واحدة من التوزيع في حالة الاختبار ذو إتجاه أو ذيل واحد او تقع على جهتي التوزيع في الاختبار ذو إتجاهين أو ذيلين.

الشكل البياني أدناه يوضح هذه المفاهيم حسب نوع الاختبار.







ويتم رفض أو قبول H_0 حسب نوع الاختبار، والقرار يمكن أن يتخذ بطريقتين مختلفتين كما يلي:

➢ **الطريقة الأولى: مقارنة القيمة المحسوبة للإحصائية مع القيمة الجدولية لها:**

1- إذا كان الاختبار من جهتين أي $H_1: \mu \neq \mu_0$، نرفض H_0 إذا كانت القيمة المحسوبة للاختبار تقع في منطقة الرفض، ويعبر عنه رياضيا بـ: $-Z_{\alpha/2} < Z_C < Z_{\alpha/2}$، حيث تكون قيمة إحصائية الاختبار أكبر من القيمة الجدولية الموجبة أو أصغر من القيمة الجدولية السالبة أي $\begin{matrix} Z_C > Z_{\alpha/2} \\ Z_C < -Z_{\alpha/2} \end{matrix}$.

2- إذا كان الاختبار من جهة واحدة نحو اليمين أي $H_1: \mu > \mu_0$ نرفض فرضية العدم H_0 إذا كانت قيمة إحصائية الاختبار أكبر من القيمة الجدولية أي أننا نقبل فرض العدم إذا تحققت المعادلة : $Z_C < Z_\alpha$ أو نرفض فرض العدم إذا تحققت المعادلة : $Z_C > Z_\alpha$.

-3 إذا كان الاختبار من جهة واحدة نحو اليسار أي $\mathbf{H_1: \mu < \mu_0}$ نرفض فرضية العدم H_0 إذا كانت قيمة إحصائية الاختبار أصغر من القيمة الجدولية أي نقبل فرض العدم إذا تحققت المعادلة : $Z_C > -Z_\alpha$ أو نرفض فرض العدم إذا تحققت المعادلة : $Z_C < -Z_\alpha$.

➢ **الطريقة الثانية: باستخدام المعنوية المحسوبة p- value**

المعنوية المحسوبة p- value عبارة عن احتمال يحسب لقيمة الإحصائية، وهي تتأثر في حسابها بالتوزيع المستخدم في الاختبار وصيغة الفرضية البديلة وذلك كما يلي:

1– إذا كان الاختبار من جهتين أي $\mathbf{H_1: \mu \neq \mu_0}$ فإن:

$$p - value = 2.P(\text{قيمة الاحصائية إذا كانت موجبة} < \text{المتغير العشوائي})$$

أو

$$p - value = 2.P(\text{قيمة الاحصائية إذا كانت سالبة} > \text{المتغير العشوائي})$$

والقرار : نرفض H_0 إذا كانت **p- value** أقل من مستوى المعنوية ونقبلها إذا كانت **p- value** أكبر من أو تساوي مستوى المعنوية (level of significant).

2– إذا كان الاختبار من جهة واحدة من جهة اليمين فإن:

$$p - value = P(\text{قيمة الاحصائية} < \text{المتغير العشوائي})$$

والقرار : نرفض H_0 إذا كانت **p- value** أقل من مستوى المعنوية ونقبلها إذا كانت **p- value** أكبر من أو تساوي مستوى المعنوية.

3– إذا كان الاختبار من جهة واحدة من جهة اليسار فإن:

$$p - value = P(\text{قيمة الاحصائية} > \text{المتغير العشوائي})$$

والقرار : نرفض H_0 إذا كانت **p- value** أقل من مستوى المعنوية ونقبلها إذا كانت **p- value** أكبر من أو تساوي مستوى المعنوية.

ملاحظة هامة: تعتبر الطريقة الثانية عمليا أكثر إستعمالا وأسهل من الطريقة الأولى، حيث أنه في كل الحالات يتم قبول فرضية العدم H_0 إذا كانت $p - value \geq \alpha$ وتقبل الفرضية البديلة إذا كانت $p - value < \alpha$.

1-3-2- القيمة الحرجة أو المعيارية حسب مستويات α

يمكننا تلخيص قيمة Z سواء كان الاختبار في اتجاهين أو اتجاه واحد في الجدول التالي:

α	في اتجاهين $Z_{\alpha/2}$	اتجاه واحد يمين $+Z_{\alpha}$	اتجاه واحد يسار $-Z_{\alpha}$
10%	± 1,65	1.28	−1.28
5%	± 1,96	1.65	−1.65
1%	± 2,58	2.33	−2.33

1-4-خطوات اختبارات الفرضيات الاحصائية

توجد خمس (05) خطوات لاختبار أية فرضية احصائية وهي:

1. تحديد الفرضية الصفرية والفرضية البديلة لها(**Null Hypothesis and Alternative Hypothesis**)
2. تحديد احصائية الاختبار وحساب قيمتها **Calculation of the Test Statistic**.
3. استخراج القيمة أو القيم الجدولية لاحصائية الاختبار.
4. مقارنة قيمة الاحصائية المحسوبة مع القيمة أو القيم الجدولية لها
5. اتخاذ القرار أي كتابة الإستنتاج.

ثانيا: اختبار الفرضيات الإحصائية

2-1- اختبار الفرضيات للمتوسط الحسابي للمجتمع μ

نميز بين حالتين:

2-1-1- عندما يكون تباين المجتمع σ^2 معلوما

نستخدم في هذه الحالة التوزيع الطبيعي، وتكون إحصائية الاختبار هي:

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

مثال (3-1): تم سحب عينة عشوائية حجمها **100** من مجتمع يتوزع توزيعا طبيعيا بتباين يساوي **49**، فوجد أن متوسطها يساوي **76**. المطلوب، اختبر الفرضيات التالية عند مستوى معنوية **5%**:

1. متوسط المجتمع يساوي **74**.
2. متوسط المجتمع يزيد عن **74**.
3. متوسط المجتمع يقل عن **74**.

الحل:

- **بالنسبة للحالة الأولى:** لدينا من معطيات المثال مايلي:

$\boldsymbol{n = 100} > 30$ $\quad \boldsymbol{\bar{X} = 76}$ $\quad \boldsymbol{\sigma^2 = 49}$ $\quad \boldsymbol{\alpha} = 5\%.$

شكل الاختبار هو:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{\mu = 74} \\ \mathbf{H_1}: \boldsymbol{\mu \neq 74} \end{cases}$$

وحيث أنّ الاختبار من جهتين (جانبين) ومستوى المعنوية $\boldsymbol{\alpha = 5\%}$ فإن القيم الحرجة (الجدولية) سوف تكون:

$$\boldsymbol{-Z_{1-\frac{\alpha}{2}} = -Z_{1-\frac{0.05}{2}} = -Z_{0.9750} = -1.96} \qquad \boldsymbol{Z_{1-\frac{\alpha}{2}} = -Z_{0.9750} = 1.96}$$

أما قيمة إحصائية الاختبار فقيمتها تكون مساوية إلى:

$$\boldsymbol{Z} = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} = \frac{76 - 74}{\sqrt{49/100}} = \boldsymbol{2.86}$$

والقرار: يمكن أن نتخذه بطريقتين مختلفتين:

➢ الطريقة الأولى:

قيمة $\boldsymbol{Z}$ المحسوبة أكبر من القيم الجدولية أي أن $2.86 < 1.96$ ، ومنه نرفض الفرضية H_0 (ونقبل الفرضية H_1) وبالتالي متوسط المجتمع لا يساوي 74.

➢ الطريقة الثانية:

الاختبار من جهتين أي $\mathrm{H_1}: \mu \neq \mu_0$ وبالتالي:

$$\boldsymbol{p-value = 2.P(Z > 2.86)} = 2[1 - P(Z \leq 2.86)] = 2\,[1 - 0.9979] = \boldsymbol{0.0042}$$

والقرار: نرفض H_0 لأنّ **p- value** أقل من مستوى المعنوية 5%. وبالتالي متوسط المجتمع لا يساوي 74.

- **بالنسبة للحالة الثانية**

يكون شكل الاختبار هو:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{\mu \leq 74} \\ \mathbf{H_1}: \boldsymbol{\mu} > 74 \end{cases}$$

وحيث أنّ الاختبار من جهة اليمين ومستوى المعنوية $\boldsymbol{\alpha = 5\%}$ فإن القيمة الحرجة (الجدولية) سوف تكون:

$$\boldsymbol{Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.64}$$

أما **قيمة إحصائية الاختبار** فقيمتها تكون مساوية إلى:

$$\boldsymbol{Z} = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/_n}} = \frac{76 - 74}{\sqrt{49/_{100}}} = \mathbf{2.86}$$

القرار:

➢ الطريقة الأولى:

الاختبار من جهة واحدة نحو اليمين أي $\mathbf{H_1}: \boldsymbol{\mu} > \boldsymbol{\mu_0}$ وبالتالي نرفض $\boldsymbol{H_0}$ لأنّ قيمة إحصائية الاختبار أكبر من القيمة الجدولية. بمعنى متوسط المجتمع لا يساوي 74.

➢ الطريقة الثانية:

الاختبار من جهة واحدة من جهة اليمين وبالتالي:

$$\boldsymbol{p - value = P(\mathrm{Z} > 2.86) = 1 - P(\mathrm{Z} \leq 2.86) = 1 - 0.9979 = 0.0021}$$

والقرار: نرفض $\boldsymbol{H_0}$ لأنّ **p- value** أقل من مستوى المعنوية 5%. بمعنى متوسط المجتمع لا يساوي 74.

- **الحالة الثالثة:**

يكون شكل الاختبار هو:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{\mu} \geq \mathbf{74} \\ \mathbf{H_1}: \boldsymbol{\mu} < 74 \end{cases}$$

وحيث أنّ الاختبار من جهة اليسار ومستوى المعنوية $\boldsymbol{\alpha} = \mathbf{5\%}$ فإن القيمة الحرجة (الجدولية) سوف تكون:

$$\boldsymbol{-Z_{1-\alpha} = -1.64}$$

أما قيمة إحصائية الاختبار فقيمتها تكون مساوية إلى:

$$\boldsymbol{Z} = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/_n}} = \frac{76 - 74}{\sqrt{49/_{100}}} = \mathbf{2.86}$$

والقرار:

➢ الطريقة الأولى:

الاختبار من جهة واحدة نحو اليسار أي $\mathbf{H_1}: \boldsymbol{\mu} < \boldsymbol{\mu_0}$ وبالتالي نقبل $\boldsymbol{H_0}$ لأنّ قيمة إحصائية الاختبار أكبر من القيمة الجدولية.

➢ الطريقة الثانية:

الاختبار من جهة واحدة من جهة اليسار وبالتالي:

$$\boldsymbol{p - value = P(\mathrm{Z} < 2.86) = 0.9979}$$

والقرار: نقبل H_0 لأنّ p- value أكبر من مستوى المعنوية 5%.

2-1-2- عندما يكون تباين المجتمع σ^2 مجهولا

في هذه الحالة نهتم بحجم العينة أي العينات الكبيرة $(n \geq 30)$ أو العينات الصغيرة $(n < 30)$.

- **في حالة العينات الكبيرة** $(n \geq 30)$:

نستخدم تباين العينة كتقدير لتباين المجتمع ، مع التوزيع الطبيعي. ومنه إحصائية الاختبار هي:

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} \sim N(0,1)$$

مثال (3- 2):

من سجلات الوفيات لإحدى البلديات لولاية الجلفة أخذت عينة من **100** شخص متوفي فوجد أن متوسط العمر لهؤلاء كان ***72*** سنة بانحراف معياري قدره ***9.8*** سنة. اختبر صحة الفوض القائل بأن متوسط العمر أكبر من **70** سنة وذلك عند مستوى معنوية **5%**.

الحل:

$n = 100 > 30 \qquad \bar{X} = 72 \qquad S = 9.8 \qquad \alpha = 5\%.$

يكون <u>شكل الاختبار</u> هو:

$$\begin{cases} H_0: \mu \leq 70 \\ H_1: \mu > 70 \end{cases}$$

وحيث أنّ الاختبار من جهة اليمين ومستوى المعنوية $\alpha = 5\%$ فإن <u>القيمة الحرجة (الجدولية)</u> سوف تكون:

$$Z_{1-\alpha} = 1.64$$

أما <u>قيمة إحصائية الاختبار</u> فقيمتها تكون مساوية إلى:

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} = \frac{72 - 70}{\sqrt{96.04/100}} = 2.04$$

<u>القرار:</u>

➢ <u>الطريقة الأولى:</u>

الاختبار من جهة واحدة نحو اليمين أي $H_1: \mu > \mu_0$ وبالتالي نرفض H_0 لأنّ قيمة إحصائية الاختبار أكبر من القيمة الجدولية أي أن $2.04 > 1.64$. بمعنى متوسط عمر المتوفي يزيد عن 70 سنة.

➢ الطريقة الثانية:

الاختبار من جهة واحدة من جهة اليمين وبالتالي:

$$p-value = P(Z > 2.04) = 1 - P(Z \leq 2.04) = 1 - 0.9793 = 0.0217$$

والقرار: نرفض H_0 لأنّ **p- value** أقل من مستوى المعنوية 5%. بمعنى متوسط عمر المتوفي يزيد عن 70 سنة.

- **في حالة العينات الصغيرة** ($n < 30$):

نستخدم تباين العينة كتقدير لتباين المجتمع ، والتوزيع الإحتمالي هو توزيع ستودنت بدرجة حرية $(n-1)$. ومنه إحصائية الاختبار هي:

$$T = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} \sim t(n-1)$$

مثال (3-3): من سجلات الوفيات لإحدى البلديات لولاية الجلفة أخذت عينة من **25** شخص متوفي، فوجد أن متوسط العمر لهؤلاء كان ***72*** سنة بانحراف معياري قدره ***9.8*** سنة. اختبر صحة الفرض القائل بأن متوسط العمر أكبر من **70** سنة وذلك عند مستوى معنوية **5%**.

الحل:

$n = 25 < 30$ $\bar{X} = 72$ $S = 9.8$ $\alpha = 5\%.$

يكون شكل الاختبار هو:

$$\begin{cases} H_0: \mu \leq 70 \\ H_1: \mu > 70 \end{cases}$$

وحيث أنّ الاختبار من جهة اليمين ومستوى المعنوية **$\alpha = 5\%$** فإن القيمة الحرجة (الجدولية) سوف تكون:

$$t_{(1-\alpha, n-1)} = t_{(0.95,24)} = 1.711$$

أما قيمة إحصائية الاختبار فقيمتها تكون مساوية إلى:

$$t = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} = \frac{72-70}{\sqrt{96.04/25}} = 1.020$$

القرار: الطريقة الأولى: الاختبار من جهة واحدة نحو اليمين أي $H_1: \mu > \mu_0$ وبالتالي نقبل H_0 لأنّ قيمة إحصائية الاختبار اصغر من القيمة الجدولية أي أن $1.020 < 1.711$. بمعنى متوسط عمر المتوفي لا يزيد عن 70 سنة.

➢ الطريقة الثانية:

الاختبار من جهة واحدة من جهة اليمين وبالتالي:

$$p - value = P(t > 1.020) = 0.20$$

والقرار: نقبل H_0 لأنّ **p- value** أكبر من مستوى المعنوية 5%. بمعنى متوسط عمر المتوفي لا يزيد عن 70 سنة.

2-2- اختبار الفرضيات للفرق بين متوسطي مجتمعين طبيعيين مستقلين

إذا كان لدينا مجتمعان يتوزعان توزيعا طبيعيا بحيث $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$، وأردنا وضع فرضيات إحصائية للفرق بين متوسطي المجتمعين $(\mu_1 - \mu_2)$ فستكون في إحدى الأشكال التالية:

➢ اختبار الفرضيات من جهتين (طرفين): $\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{cases}$

➢ اختبار الفرضيات من جهة اليمين: $\begin{cases} H_0: \mu_1 - \mu_2 \leq 0 \\ H_1: \mu_1 - \mu_2 > 0 \end{cases}$

➢ اختبار الفرضيات من جهة اليسار: $\begin{cases} H_0: \mu_1 - \mu_2 \geq 0 \\ H_1: \mu_1 - \mu_2 < 0 \end{cases}$

والقرار في قبول أو رفض الفرضية الصفرية H_0 يمكن أن تتخذ بطريقتين مختلفين:

➢ **الطريقة الأولى: مقارنة القيمة المحسوبة مع القيمة الجدولية لها:**

1- إذا كان الاختبار من جهتين أي $H_1: \mu_1 - \mu_2 \neq 0$، نرفض H_0 إذا كانت القيمة المحسوبة للاختبار تقع في منطقة الرفض، أي تكون قيمة إحصائية الاختبار أكبر من القيمة الجدولية الموجبة أو أصغر من القيمة الجدولية السالبة.

2- إذا كان الاختبار من جهة واحدة نحو اليمين أي $H_1: \mu_1 - \mu_2 > 0$ نرفض H_0 إذا كانت قيمة إحصائية الاختبار أكبر من القيمة الجدولية.

3- إذا كان الاختبار من جهة واحدة نحو اليسار أي $H_1: \mu_1 - \mu_2 < 0$ نرفض H_0 إذا كانت قيمة إحصائية الاختبار أصغر من القيمة الجدولية.

➢ **الطريقة الثانية: باستخدام المعنوية المحسوبة p- value**

1- إذا كان الاختبار من جهتين أي $H_1: \mu_1 - \mu_2 \neq 0$ فإن:

$$p - value = 2.P(\text{المتغير العشوائي} > \text{قيمة الاحصائية إذا كانت موجبة})$$

أو

$$p - value = 2.P(\text{قيمة الاحصائية إذا كانت سالبة} > \text{المتغير العشوائي})$$

والقرار : نرفض H_0 إذا كانت **p- value** أقل من مستوى المعنوية ونقبلها إذا كانت **p- value** أكبر من أو تساوي مستوى المعنوية.

2- إذا كان الاختبار من جهة اليمين فإن:

$$p - value = P(\text{قيمة الاحصائية} > \text{المتغير العشوائي})$$

والقرار : نرفض H_0 إذا كانت **p- value** أقل من مستوى المعنوية ونقبلها إذا كانت **p- value** أكبر من أو تساوي مستوى المعنوية.

3- إذا كان الاختبار من جهة اليسار فإن:

$$p - value = P(\text{قيمة الاحصائية} > \text{المتغير العشوائي})$$

والقرار : نرفض H_0 إذا كانت **p- value** أقل من مستوى المعنوية ونقبلها إذا كانت **p- value** أكبر من أو تساوي مستوى المعنوية.

سنتطرق تبعا الى الحالات المختلفة عند اختبار الفرضيات للفرق بين متوسطي مجتمعين طبيعيين مستقلين:

2-2-1- حالة تباين المجتمعين معلوم $\left(\sigma_2^2\right.$ و $\left.\sigma_1^2\right)$ معلومين

في هذه الحالة إحصائية الاختبار هي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0.1)$$

مثال (3-4):

سحبت عينة عشوائية حجمها **30** من مجتمع طبيعي تباينه 25 فوجد أن متوسطها الحسابي يسلوي **14**، وسحبت عينة عشوائية أخوى حجمها **25** من مجتمع طبيعي أخر تباينه 16 فوجد أن متوسطها الحسابي يسلوي **12**. المطلوب: هل يوجد فوق معنوي بين متوسطي المجتمعين عند مسقوى معنوية يسلوي **10%** ؟.

الحل:

من المعطيات لدينا:

$n_1 = 30$ $\bar{X}_1 = 14$ $\sigma_1^2 = 25$ $\alpha = 10\%.$

$n_2 = 25$ $\qquad \bar{X}_2 = 12$ $\qquad \sigma_2^2 = 16$

شكل الاختبار المطلوب هو:

$$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{cases}$$

بما أنّ الاختبار ذو جانبين ومستوى المعنوية $\alpha = 10\%$ فإن القيم الحرجة (الجدولية) سوف تكون:

$$Z_{1-\frac{\alpha}{2}} = Z_{0.950} = 1.64 \qquad -Z_{1-\frac{\alpha}{2}} = -Z_{1-\frac{0.1}{2}} = -Z_{0.950} = -1.64$$

أما قيمة إحصائية الاختبار فقيمتها تكون مساوية إلى:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{2 - 0}{\sqrt{\frac{25}{30} + \frac{16}{25}}} = 1.65$$

القرار: يمكن أن يتخذ القرار بطريقتين مختلفتين:

➢ الطريقة الأولى:

قيمة Z المحسوبة أكبر من القيم الجدولية أي أن $1.65 > 1.64$ ، ومنه نرفض الفرضية H_0 (ونقبل الفرضية H_1) أي أنه يوجد فرق بين متوسطي المجتمعين، وبالتالي المجتمعين ليس لهما نفس المتوسط.

➢ الطريقة الثانية:

الاختبار من جهتين أي $H_1: \mu_1 - \mu_2 \neq 0$ أي أن:

$$p - value = 2.P(Z > 1.65) = 2[1 - P(Z \leq 1.65)] = 2\,[1 - (0.9505)] = 0.099$$

ومنه نرفض H_0 لأنّ **p- value** أقل من مستوى المعنوية 10%. وبالتالي يوجد فرق بين متوسطي المجتمعين.

2-2-2- حالة تباين المجتمعين مجهول (σ_2^2 وσ_1^2 مجهولين)

نميز بين حالتين مع إستخدام تباين العينتين S_2^2 وS_1^2 كتقدير لتباين المجتمعين σ_2^2 وσ_1^2:

2-2-2-1- عندما يكون σ_2^2 وσ_1^2 مجهولين وغير متساويين

- **في حالة العينات الكبيرة** $n_1, n_2 \geq 30$، تكون إحصائية الاختبار في هذه الحالة هي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0.1)$$

مثال (3- 5):

في امتحان لمادة الاقتصاد الكلي تقدم **80** طالبة و**60** طالبا فكان متوسط علامات الطالبات هو **14** بانحراف معياري يساوي **2** نقطة، بينما كان متوسط علامات الطلبة الذكور هو **12** بانحراف معياري يساوي **1.5** نقطة. المطلوب: هل يوجد فرق بين مستوى الطالبات والطلاب (مع العلم أن تباين علامات الطلاب في الجامعة ككل يختلف عن تباين علامات الطالبات) عند مستوى معنوية تساوي **5%** ؟.

الحل:

$\boldsymbol{n_1} = \mathbf{80}$ $\quad \boldsymbol{\bar{X}_1} = \mathbf{14}$ $\quad \boldsymbol{S_1} = \mathbf{2}$ $\quad \boldsymbol{\alpha} = 5\%.$

$\boldsymbol{n_2} = \mathbf{60}$ $\quad \boldsymbol{\bar{X}_2} = \mathbf{12}$ $\quad \boldsymbol{S_2} = \mathbf{1.5}$

بإتباع الخطوات المعروفة نجد:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{\mu_1} - \boldsymbol{\mu_2} = \mathbf{0} \\ \mathbf{H_1}: \boldsymbol{\mu_1} - \boldsymbol{\mu_2} \neq \mathbf{0} \end{cases}$$

وحيث أنّ مستوى المعنوية $\boldsymbol{\alpha} = \mathbf{10\%}$ فإن القيم الحرجة (الجدولية) سوف تكون:

$$-\boldsymbol{Z}_{1-\frac{\alpha}{2}} = -\mathbf{1.96} \qquad \boldsymbol{Z}_{1-\frac{\alpha}{2}} = \boldsymbol{Z}_{1-\frac{0.05}{2}} = \boldsymbol{Z}_{0.975} = \mathbf{1.96}$$

أما قيمة إحصائية الاختبار فقيمتها تكون مساوية إلى:

$$\boldsymbol{Z} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{2-0}{\sqrt{\frac{2^2}{80} + \frac{1.5^2}{60}}} = \frac{2-0}{\sqrt{\frac{4}{80} + \frac{2.25}{60}}} = \mathbf{6.76}$$

القرار:

➢ الطريقة الأولى:

بما أن قيمة $\boldsymbol{Z}$ المحسوبة أكبر من القيم الجدولية أي أن $\mathbf{6.76 > 1.96}$ ، نرفض الفرضية H_0 (ونقبل الفرضية H_1) ومنه يوجد فرق بين مستوى الطالبات والطلاب لوجود فرق بين متوسطي المجتمعين.

الطريقة الثانية:

الاختبار من جهتين أي $\mathbf{H_1}: \boldsymbol{\mu_1} - \boldsymbol{\mu_2} \neq \mathbf{0}$ وبالتالي:

$$\boldsymbol{p-value} = \mathbf{2}.\boldsymbol{P}(\mathbf{Z} > \mathbf{6.76}) = 2[1 - P(Z \leq 6.76)] = 2\,[1-1] = \mathbf{0.00}$$

ومنه يتم رفض فرضية العدم H_0 لأنّ **p- value** أقل من مستوى المعنوية 5%. وهو ما يعني وجود فرق بين مستوى الطالبات و الطلاب.

- **حالة العينات الصغيرة** $n_1, n_2 < 30$ **(أو أحدهما صغير)**، في هذه الحالة تكون إحصائية الاختبار هي:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t(n_1 + n_2 - 2)$$

مثال (3-6):

بإستعمال معطيات المثال (3-5) السابق، مع افتراض أنّ حجم العينة الأولى يسلوي **20** وحجم العينة الثانية يسلوي **18**.

الحل:

$n_1 = 20$ $\bar{X}_1 = 14$ $S_1 = 2$ $\alpha = 5\%.$

$n_2 = 18$ $\bar{X}_2 = 12$ $S_2 = 1.5$

بإتباع الخطوات السابقة نجد أن:

$$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{cases}$$

مستوى المعنوية $\alpha = 10\%$ فإن القيم الحرجة (الجدولية) في هذه الحالة تكون:

$$-t_{(1-\frac{\alpha}{2}, n-2)} = -t_{(0.975,36)} = -2.0281 \quad , \quad t_{(1-\frac{\alpha}{2}, n-2)} = t_{(0.975,36)} = 2.0281$$

وقيمة إحصائية الاختبار هي:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{2 - 0}{\sqrt{\frac{4}{20} + \frac{2.25}{18}}} = 3.51$$

القرار:

➢ الطريقة الأولى:

قيمة T المحسوبة أكبر من القيم الجدولية أي أن $3.51 > 2.0281$ ، ومنه نرفض الفرضية H_0 (ونقبل الفرضية H_1) أي أنه يوجد فرق بين مستوى الطالبات والطلبة لوجود فرق بين متوسطي المجتمعين.

➢ الطريقة الثانية:

الاختبار من جهتين أي $H_1: \mu_1 - \mu_2 \neq 0$ وبالتالي قيمة الاحتمالية المشاهدة هي:

$$p - value = 2.P(T > 3.51) = 2\,[0.0005] = 0.001$$

نرفض فرضية العدم H_0 لأنّ **p- value** أقل من مستوى المعنوية 5%. وهذا دليل على وجود فرق بين مستوى الطلاب والطالبات.

2-2-2-2- عندما يكون σ_2^2 و σ_1^2 مجهولين و متساويين

نستبدل قيمة كل من σ_1^2 و σ_2^2 بـ التباين المشترك S_P^2. وقيمة إحصائية الاختبار تكون هي:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_P^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t(\vartheta = n_1 + n_2 - 2)$$

مثال (3- 7):عينتان عشوائيتان من مجتمعين طبيعيين، حيث:

$$\bar{X}_1 = 5.38 \qquad S_1 = 1.59 \qquad n_1 = 10$$

$$\bar{X}_2 = 32.1 \qquad S_2 = 0.83 \qquad n_2 = 12$$

$$S_P = 1.33 \qquad \alpha = 5\%$$

المطلوب: هل يوجد فوق معوي بين متوسطي المجتمعين؟.

الحل:

لدينا شكل الاختبار هو:

$$\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{cases}$$

الاختبار ذو جانبين ومستوى المعنوية $\alpha = 5\%$ فإن القيم الحرجة (الجدولية):

$$t_{1-\frac{\alpha}{2},\vartheta} = t_{0.975,20} = 1.725 \qquad -t_{1-\frac{\alpha}{2},\vartheta} = t_{0.975,20} = -1.725$$

أما قيمة إحصائية الاختبار فقيمتها تكون مساوية إلى:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_P^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(5.92 - 5.38) - 0}{\sqrt{1.77\left(\frac{1}{10} + \frac{1}{12}\right)}} = 1.02$$

القرار:

- الطريقة الأولى:

قيمة T المحسوبة تقع بين القيمتين الجدولتين أي في منطقة القبول، ومنه لا يوجد فرق بين متوسطي المجتمعين.

➢ الطريقة الثانية:

الاختبار من جهتين أي $H_1: \mu_1 - \mu_2 \neq 0$ ومنه:

$$p - value = 2.P(T > 1.02) = 0.90$$

نقبل فرضية العدم H_0 لأنّ p- value أكبر من مستوى المعنوية 5%. وهذا يعني عدم وجود فرق بين متوسطي المجتمعين.

2-3- اختبار الفرضيات للفرق بين متوسطي مجتمعين مرتبطين

متوسط مجتمع الفروق μ_D يساوي إلى فرق متوسطي المجتمعين $\mu_1 - \mu_2$ أي أن مسألة اختبار الفرضيات ستكون مطابقة لأحد الأشكال التالية:

$$\begin{cases} H_0: \mu_D \leq 0 \\ H_1: \mu_D > 0 \end{cases} , \quad \begin{cases} H_0: \mu_D \geq 0 \\ H_1: \mu_D < 0 \end{cases} , \quad \begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D \neq 0 \end{cases}$$

ونميز بين حالتين:

- **في حالة العينات الكبيرة $n \geq 30$** : إحصائية الاختبار هي:

$$Z = \frac{\bar{D} - \mu_D}{\sqrt{S_D^2/n}}$$

مثال (3- 8):

تم سحب عينتين عشوائيتين حجم كل منهما يساوي إلى **30** من مجتمعين طبيعيين مرتبطين وكان مجموع الفروق بين القيم المتناظرة للعينتين هو **180** والانحراف المعياري للفروق يساوي 9. المطلوب: اختبر فرضية أن المجتمعين لهما نفس المتوسط عند مسوى معنوية تساوي **5%**.

الحل:

$$\alpha = 5\%. \qquad S_D = 9 \qquad \sum D_i = 180 \qquad n = 30$$

يكون <u>شكل الاختبار</u> هو:

$$\begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D \neq 0 \end{cases}$$

الاختبار ذو جانبين ومستوى المعنوية $\alpha = 5\%$ فإن <u>القيم الحرجة (الجدولية)</u> تكون:

$$Z_{1-\frac{\alpha}{2}} = 1.96 \qquad -Z_{1-\frac{\alpha}{2}} = -1.96$$

نحسب قيمة متوسط الفروق: $\bar{D} = \frac{180}{30} = 6$

قيمة إحصائية الاختبار تساوي:

$$Z = \frac{\bar{D} - \mu_D}{\sqrt{S_D^2/n}} = \frac{6 - 0}{\sqrt{9^2/30}} = 3.65$$

القرار:

- الطريقة الأولى:

قيمة Z المحسوبة في منطقة الرفض لأنها اكبر من القيمة الجدولية 1.96 أي أن:

$3.65 \geq 1.96$ ، ومنه نرفض فرضية العدم H_0 أي أنه وجد فرق بين متوسطي المجتمعين ومنه متوسطي المجتمعين مختلفين.

- الطريقة الثانية:

الاختبار من جهتين أي $H_1: \mu_D \neq 0$ وبالتالي:

$$p - value = 2.P(Z > 3.65) = 2[1 - P(Z \leq 3.65)] = 2\,[1 - 1] = 0$$

وعليه نرفض فرضية العدم H_0 لأنّ **p- value** أقل من مستوى المعنوية 10%. وهذا يعني وجود فرق بين متوسطي المجتمعين.

- **في حالة العينات الصغيرة** $n < 30$، إحصائية الاختبار هي:

$$T = \frac{\bar{D} - \mu_D}{\sqrt{S_D^2/n}} \sim t(n - 1)$$

مثال (3- 9):

البيانات التالية تتعلق بنقاط مجموعة من الطلبة خضعوا لدورة تكوينية في مادة الإنجليزية:

قبل الدورة	14	14	13	14	12	14	13	12	13
بعد الدورة	15	13	16	16	15	17	13	16	17

المطلوب: اختبر وجود تحسن في مستوى الطلبة عند إجراء الدورة التكوينية عند مسوى معنوية 5%؟

الحل:

شكل الاختبار في هذه الحالة هو:

$$\begin{cases} H_0: \mu_D \leq 0 \\ H_1: \mu_D > 0 \end{cases}$$

عينة الفروق، المتوسط والتباين كما يلي:

قبل الدورة	14	14	13	14	12	14	13	12	13	**المجموع**
بعد الدورة	15	13	16	16	15	17	13	16	17	/
d_i	**-1**	**1**	**-3**	**-2**	**-3**	**-3**	**0**	**-4**	**-4**	19-
$(d_i-\bar{d})^2$	**1**	**1**	**9**	**4**	**9**	**9**	**0**	**16**	**16**	**65**

وتكون:

$$\bar{D} = \frac{\sum D_i}{n} = \frac{-19}{9} = -2.1$$

$$S_D^2 = \frac{\sum(D_i - \mu_D)^2}{n-1} = \frac{24.9}{8} = 3.1$$

وبذلك نحصل على:

$$T = \frac{\bar{D} - \mu_D}{\sqrt{\frac{S_D^2}{n}}} = \frac{-2.1 - 0}{\sqrt{\frac{3.1}{8}}} = -3.58$$

بالبحث في جدول توزيع ستودنت عند درجة حرية تساوي 9 نجد أن:

$$t_{(1-\alpha,n-1)} = t_{(0.95,9)} = 1.833$$

القرار:

- الطريقة الأولى:

قيمة T المحسوبة تقع في منطقة القبول لأنها أقل من القيمة الجدولية ومنه نقبل فرضية العدم H_0 وبالتالي لا يوجد تحسن عند تناول الدواء الجديد.

- الطريقة الثانية:

الاختبار من جهة اليمين أي $H_1: \mu_D > 0$ وبالتالي:

$$p-value = P(T > -3.58) = [1 - P(T > 3.58)] = [1 - (0.005)] = 0.995$$

ومنه نقبل الفرضية H_0 لأنّ **p- value** أكبر من مستوى المعنوية 5%. أي لا يوجد تحسن في مستوى الطلبة.

4-2- اختبار الفرضيات للنسبة Testing A Hypothesis About A population Proportion P

هناك ثلاث صيغ ممكنة لاختبار الفرضيات وتكتب على الشكل التالي:

$$\begin{cases} H_0: P \leq P_0 \\ H_1: P > P_0 \end{cases} \quad , \quad \begin{cases} H_0: P \geq P_0 \\ H_1: P < P_0 \end{cases} \quad , \quad \begin{cases} H_0: P = P_0 \\ H_1: P \neq P_0 \end{cases}$$

وتكون إحصائية الاختبار هي:

$$Z = \frac{\hat{P} - P_0}{\sqrt{P_0 . q_0 / n}} \sim N(0,1)$$

مثال (3- 10):

اذا كانت نسبة المدخنين في إحدى المدن عام 2000 هي % 28.8 وفي عام 2010 اختيرت عينة من سكان هذه المدينة حجمها 1283 شخصا فكان من بينهم 320 شخصا من المدخنين . فهل تدل هذه النتائج على انخفاض نسبة المدخنين بين عامي 2000 , 2010 مستخدما مسوى معنوية α=0.05 ؟.

الحل:

لدينا المعطيات التالية:

$$\hat{P} = \frac{320}{1283} = 24.9\% \qquad P = 28.8\% \qquad n = 1280$$

يكون شكل الاختبار هو:

$$\begin{cases} H_0: P \geq 0.288 \\ H_1: P < 0.288 \end{cases}$$

الاختبار من جهة اليسار، التوزيع المستخدم هو التوزيع الطبيعي ومستوى المعنوية $\alpha = 5\%$ فإن القيمة الحرجة تكون: $-Z_{1-\alpha} = -1.64$

أما قيمة إحصائية الاختبار فهي مساوية إلى:

$$Z = \frac{0.249 - 0.288}{\sqrt{(0.288).(0.712) / 1280}} = -3.9$$

القرار:

- الطريقة الأولى:

قيمة Z المحسوبة أقل من القيمة الجدولية لها، أي أنها وقعت في منطقة الرفض ومنه نرفض الفرضية H_0. أي أن نسبة المدخنين في انخفاض بين السنتين 2000 و2010.

- الطريقة الثانية:

الاختبار من جهة اليسار أي $H_1: P < 0.288$ ومنه:

$$p - value = P(Z < -3.9) = 0$$

وبالتالي نرفض الفرضية H_0 لأنّ **p- value** أقل من مستوى المعنوية 5%. وبالتالي فرضية نسبة التدخين في انخفاض مقبولة.

5-2- اختبار الفرضيات للفرق بين نسبتين $(P_1 - P_2)$ Significance of A Difference between Two Proportions

إذا أردنا اختبار فرضية فروق بين نسبتي مجتمعين فإن صياغة الفرضية تكون وفق ثلاث صيغ ممكنة وهي:

$$\begin{cases} H_0: P_1 \le P_2 \\ H_1: P_1 > P_2 \end{cases} \quad , \quad \begin{cases} H_0: P_1 \ge P_2 \\ H_1: P_1 < P_2 \end{cases} \quad , \quad \begin{cases} H_0: P_1 = P_2 \\ H_1: P_1 \ne P_2 \end{cases}$$

وتعطى إحصائية الاختبار بالعلاقة التالية:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \sim N(0,1)$$

من أجل n_1 وn_2 أكبر من 30، لدينا:

$$\sigma_{\hat{P}_2} = \sqrt{\frac{\hat{P}_2 * (1-\hat{P}_2)}{n_2}} \qquad \hat{P}_2 \sim N(P_2, \sigma_{\hat{P}_2}) \qquad \sigma_{\hat{P}_1} = \sqrt{\frac{\hat{P}_1 * (1-\hat{P}_1)}{n_1}} \qquad \hat{P}_1 \sim N(P_1, \sigma_{\hat{P}_1})$$

توزيع المعاينة للفرق يتبع التوزيع الطبيعي ونكتب: $(\hat{P}_1 - \hat{P}_2) \sim N((P_1 - P_2), \sigma_{\hat{P}_1 - \hat{P}_2})$

بما أن الاختبار يتم تحت افتراض عدم وجود اختلاف بين النسبتين $(P_1 - P_2)$ أي أننا في هذه الحالة نتعامل مع نسبة واحدة وليست نسبتين وبالتالي لا بد من حساب النسبة المشتركة والتي سوف نرمز لها بالرمز:

$$\bar{P} = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1 + n_2} \quad \text{، ومنه :} \quad \sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{\bar{P}(1-\bar{P})}{n_1} + \frac{\bar{P}(1-\bar{P})}{n_2}}$$

ملاحظة هامة: $\sigma_{\hat{P}_1 - \hat{P}_2}$ هنا تختلف عن $\sigma_{\hat{P}_1 - \hat{P}_2}$ الموجودة في فترات الثقة.

مثال (3- 11):

في دراسة تمت عام 2000 وجد أن من بين 3000 أسرة تم اختيارهم بطريقة عشوائية، وجد أن 1350 أسرة يمتلكون سيارة خاصة. وفي عام 2020 تم اختيار عينة من 3000 أسرة وجد أن 1650 أسرة تمتلك سيارة خاصة. هل تدل هذه البيانات أن نسبة من يمتلكون سيارة خاصة قد اختلفت في العشرين سنة الماضية مستخدما مستوى معنوية $\alpha = 5\%$ ؟

الحل: لدينا

$$n_1 = 3000 \qquad x_1 = 1350 \qquad \hat{P}_1 = \frac{x_1}{n_1} = \frac{1350}{3000} = 0.45 \qquad \alpha = 1\%.$$

$$n_2 = 3000 \qquad x_2 = 1650 \qquad \hat{P}_2 = \frac{x_2}{n_2} = \frac{1650}{3000} = 0.55$$

$$\bar{P} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{1350 + 1650}{3000 + 3000} = 0.50$$

نريد اختبار الفرضية:

$$\mathrm{H_0}: P_1 = P_2$$
$$\mathrm{H_1}: \mathrm{P_1} \neq \mathrm{P_2}$$

لحساب قيمة إحصاءة الإختبار نقوم اولا بحساب الانحراف المشترك:

$$\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{\bar{P}(1-\bar{P})}{n_1} + \frac{\bar{P}(1-\bar{P})}{n_2}} = \sqrt{\frac{0.5(1-0.5)}{3000} + \frac{0.5(1-0.5)}{3000}} = 0.013$$

ومنه قيمة إحصائية الاختبار هي:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}} = \frac{(0.45 - 0.55) - 0}{0.013} = -7.69$$

الاختبار من جهتين والتوزيع توزيع طبيعي، ومستوى المعنوية تساوي **5%** وبذلك تكون القيم الحرجة هي:

$$-Z_{1-\frac{\alpha}{2}} = -1.96 \qquad Z_{1-\frac{\alpha}{2}} = 1.96$$

القرار:

قيمة Z المحسوبة تقع خارج القيمتين الجدوليتين أي انها تقع في منطقة الرفض، ومنه نرفض الفرضية H_0 ومنه يوجد فرق بين نسبتي امتلاك الأسر لسيارة خاصة خلال العشرين سنة الماضية.

6-2- اختبار الفرضيات لتباين مجتمع يتوزع توزيعا طبيعيا (σ^2)

من خلال ما تطرقنا له خلال الفصلين السابقين: إذا كان لدينا مجتمع طبيعي أي $X \sim N(\mu, \sigma^2)$ وكان S^2 هو تباين العينة فإن $\frac{(n-1)S^2}{\sigma^2}$ تتبع التوزيع كاي مربع بدرجة حرية $n-1$ ومنه $\chi = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$.

لاختبار الفرضيات حول تباين المجتمع σ^2 توجد ثلاث حالات ممكنة هي:

$$\begin{cases} \mathrm{H_0}: \sigma^2 = \sigma_0^2 \\ \mathrm{H_1}: \sigma^2 \neq \sigma_0^2 \end{cases} , \quad \begin{cases} \mathrm{H_0}: \sigma^2 \geq \sigma_0^2 \\ \mathrm{H_1}: \sigma^2 < \sigma_0^2 \end{cases} , \quad \begin{cases} \mathrm{H_0}: \sigma^2 \leq \sigma_0^2 \\ \mathrm{H_1}: \sigma^2 > \sigma_0^2 \end{cases}$$

وبهذا تكون إحصائية الاختبار هي:

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

مثال (3- 12):

معمل للأدوية ينتج نوعا من العقار يحتوي مادة فعالة بكمية محددة ودقيقة، وبغية دراسة مدى دقة المصنع في إضافة هذه المادة لحبات العقار، قام المسؤولون بتحليل عينة حجمها 30 حبة، فكان الانحراف المعياري لكمية المادة الفعالة يساوي 1.3ملغ.

اختبر الفرضية ^{2}

الحل:

من معطيات المثال لدينا:

$\sigma_0^2 = 1.5$ $\quad n = 30$ $\quad S = 1.3$ $\quad \alpha = 5\%$

الاختبار المطلوب هو للفرضية:

$$H_0: \sigma^2 \geq 1.5$$
$$H_1: \sigma^2 < 1.5$$

الاختبار أحادي الجانب وهو من الجانب الأيسر، من جدول توزيع كاي مربع عند مستوى معنوية $\alpha = 5\%$ نجد القيمة الحرجة تساوي:

$$\chi^2_{(1-\alpha,n-1)} = \chi^2_{(0.95,29)} = 17.7$$

وإحصائية الاختبار هي:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{29 \times 1.69}{1.5} = 32.67$$

القرار:

➢ الطريقة الأولى:

قيمة χ^2 المحسوبة **32.67** أكبر من القيمة الجدولية لها **17.7**، ومنه نقبل فرضية العدم H_0 ومنه تباين المجتمع ليس اقل من **1.5**.

➢ الطريقة الثانية:

الاختبار من جهة اليسار أي $H_1: \sigma^2 < 1.5$ ومنه: $p-value = P(\chi^2 < 32.67) = 0.10$

أي أننا نقبل فرضية العدم H_0 لأنّ **p- value** أكبر من مستوى المعنوية 5%.

6-2- اختبار الفرضيات حول نسبة التباين $\left(\frac{\sigma_1^2}{\sigma_2^2}\right)$

في هذه الحالة تكون صيغ اختبار الفرضيات كالتالي:

$$\begin{cases} H_0: \sigma_1^2 \leq \sigma_2^2 \\ H_1: \sigma_1^2 > \sigma_2^2 \end{cases} , \quad \begin{cases} H_0: \sigma_1^2 \geq \sigma_2^2 \\ H_1: \sigma_1^2 < \sigma_2^2 \end{cases} , \quad \begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

أو بعبارة أخرى:

$$\begin{cases} H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \\ H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1 \end{cases} , \quad \begin{cases} H_0: \frac{\sigma_1^2}{\sigma_2^2} \geq 1 \\ H_1: \frac{\sigma_1^2}{\sigma_2^2} < 1 \end{cases} , \quad \begin{cases} H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{cases}$$

لإختبار الفرضيات حول نسبة تبايني مجتمعين مستقلين يتوزعان توزيعا طبيعيا، فإن إحصائية الاختبار حسب توزيعات المعاينة للنسبة بين التباينيين تعطى بالعلاقة:

$$f = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1 - 1, n_2 - 1)$$

مثال (3- 13):

إذا علمت أن مجتمع أطوال الطلبة و مجتمع أطوال الطالبات في إحدى الكليات يتبع التوزيع الطبيعي، وسحبنا عينة من الطلبة حجمها **20** ومن الطالبات حجمها **25** علما أنّ العينتين مستقلتين، ووجدنا أن تباين عينة الطلبة يسوي **49** تباين عينة الطالبات يسوي **36**. المطلوب: هل يوجد فوق بين تباين مجتمع أطوال الطلبة وتباين مجتمع أطوال الطالبات عند مسوى معنوية تسوي **5%** ؟.

الحل:

من المثال لدينا المعطيات الأتية:

$\alpha = 5\%$, $S_1^2 = 49$, $S_1^2 = 36$, $n_1 = 20$, $n_2 = 25$

نريد اختبار الفرضية التالية:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \leftrightarrow \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \leftrightarrow \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{cases}$$

إحصائية الاختبار هي:

$$f = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} = \frac{49}{36} \cdot 1 = 1.36$$

وحيث أنّ الاختبار من جهتين (جانبين) ومستوى المعنوية $\alpha = 5\%$ فإن القيم الحرجة (الجدولية) عند درجتي حرية $n_1 - 1 = 19$ و $n_2 - 1 = 24$ سوف تكون:

$$F_{(\frac{\alpha}{2}, n_1-1, n_2-1)} = F_{(0.025,19,24)} = 2.33$$

$$F_{(1-\frac{\alpha}{2}, n_1-1, n_2-1)} = F_{(1-\frac{0.05}{2}, 20-1, 25-1)} = F_{(0.975,19,24)} = \frac{1}{F_{(0.025,24,19)}} = \frac{1}{2.45} = 0.408$$

القرار:

- الطريقة الأولى:

قيمة f المحسوبة تقع بين القيمتين المجدولتين، أي نقبل فرضية العدم H_0 ومنه لا يوجد فرق بين تباين مجتمع أطوال الطلبة وتباين مجتمع أطوال الطالبات.

- الطريقة الثانية:

الاختبار من جهتين أي $\mathbf{H_1}: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$ وبالتالي:

$$p-value = 2.P(\mathrm{F}_{(19,24)} > 1.36) = \mathrm{F}\ (2.33) - \mathrm{F}\ (0.408) = 0.975 - 0.025 = 0.95$$

نقبل فرضية العدم H_0 لأنّ **p- value** أكبر من مستوى المعنوية 5%.

ثالثا: الاختبارات الخاصة بتحليل الانحدار والارتباط

3-1- اختبار معامل الانحدار الخطي البسيط

تحليل الانحدار عبارة عن أسلوب يختص بدراسة العلاقة بين متغيرين أو أكثر، وكيفية إيجاد معادلة جبرية تمثل هذه العلاقة رياضيا أفضل تمثيل، وذلك لاستخدامها في التقدير أو التنبؤ بمتوسط أحد المتغيرات. ويرمز للمتغير التابع أو المتأثر عادة بالرمز $\boldsymbol{Y}$ بينما يرمز للمتغير المستقل (المفسر أو المؤثر) بـ $\boldsymbol{X}$. بعبارة أخرى، تحليل الانحدار يقصد به الدراسة الخاصة بالتقدير، والتنبؤ بمتوسط أحد المتغيرات إذا علمنا قيم باقي المتغيرات المؤثرة فيه.

ويعبّر رياضيا عن انحدار $\boldsymbol{Y}$ على $\boldsymbol{X}$ بأنّ المتغير $\boldsymbol{Y}$ دالة للمتغير $\boldsymbol{X}$ ويرمز لذلك كما يلي:

$$\boldsymbol{Y = f(X) = a + b.X + \varepsilon}$$

حيث $\boldsymbol{X}$ هو المتغير المستقل و $\boldsymbol{Y}$ هو المتغير التابع، وتسمى $\boldsymbol{a}$ و $\boldsymbol{b}$ بالمعالم، ويعبّر عنهما كما يلي:

$\boldsymbol{a}$: ثابت خط الانحدار على محور $\boldsymbol{Y}$.

b: ميل خط انحدار Y على X. وتلعب دورا مهما في التحليل الإحصائي، حيث أنها تعبّر عن معدل التغير في Y بالنسبة لـ X. فإذا كانت X تمثل كمية الأكسجين المستنشق وكانت Y تمثل عدد ضربات القلب فإنّ b هي معدل الزيادة في ضربات القلب بالنسبة لكميات الأكسجين المستنشق.

$\varepsilon \sim N(0, \sigma^2)$: تعبّر عن الخطأ في التقدير بمتوسط يساوي 0 وتباين ثابت يساوي σ^2.

ويتم تقدير المعلمتين a و b بطريقة المربعات الصغرى العادية والمعروفة اختصارا باللغة الفرنسية بـ MCO.

يتم تقدير b من العلاقة:

$$\hat{b} = \frac{SS_{XY}}{SS_X} = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}$$

حيث SS تعني مجموع المربعات، ومجموع المربعات عبارة عن قياس للتباينات أو للانحرافات بالنسبة للمتوسط.

وتتبع المعلمة المقدرة $\hat{b}$ التوزيع الطبيعي بمتوسط يساوي b وتباين يساوي $\frac{\sigma^2}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}$. واختصارا يعبّر عن توزيع $\hat{b}$ كما يلي:

$$\hat{b} \sim N\left(b, \frac{\sigma^2}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}\right)$$

وفي حالة مجهولية σ^2 تستبدل بتباين العينة المسحوبة منه S^2.

وتكون صيغ اختبار الفرضيات حول المعلمة b كالتالي:

$$\begin{cases} H_0: b \le b_0 \\ H_1: b > b_0 \end{cases} \quad , \quad \begin{cases} H_0: b \ge b_0 \\ H_1: b < b_0 \end{cases} \quad , \quad \begin{cases} H_0: b = b_0 \\ H_1: b \ne b_0 \end{cases}$$

وإحصائية الاختبار المستخدمة تتبع توزيع ستودنت بدرجة حرية $(n-2)$، وتسمى أيضا باختبار المعلمة الواحدة.

$$T = \frac{\hat{b} - b_0}{\sqrt{\frac{S^2}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}}} \sim t_{(n-2)}$$

حيث:

$$S^2 = \frac{SSE}{n-2}$$

$$SS_E = SS_Y - \hat{b}\, SS_{XY}$$

$$SS_Y = \sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2$$

ملاحظة:

SSE تعني مجموع مربعات الأخطاء.

SS_Y تعني المجموع الكلي للمربعات.

$\hat{b}SS_{XY}$ تعني مجموع مربعات القيم المفسرة أو المشروحة.

مثال (3- 14):

البيانات التالية في الجدول **(01)** أدناه، تبين كمية الإنتاج ومقدار التكاليف لليوم الواحد لأحد المصانع.

المطلوب: اختبر فرضية ما إذا كانت $b = 0$ وذلك عند مستوى معنوية 1 %.

X_i	10	11	14	15	20	25	46	50	59	70
Y_i	10	10	12	12	13	13	19	15	16	20

الحل:

لغرض اختبار الفرضية:

$$\begin{cases} H_0: b = b_0 \\ H_1: b \neq b_0 \end{cases}$$

نقدّر أولا $\hat{b}$ من خلال العلاقة:

$$\hat{b} = \frac{SS_{XY}}{SS_X} = \frac{\sum_{i=1}^{n} X_iY_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}$$

وبالتالي نقوم بالحسابات التالية:

Y_i^2	X_i^2	X_iY_i	Y_i	X_i
100	100	100	10	10
100	121	110	10	11
144	196	168	12	14
144	225	180	12	15
169	400	260	13	20
169	625	325	13	25
361	2116	874	19	46
225	2500	750	15	50
256	3481	944	16	59
400	4900	1400	20	70
2068	**14664**	**5111**	**140**	**320**

وبالتالي:

$$\bar{\boldsymbol{X}} = \frac{\sum \mathrm{X_i}}{\mathrm{n}} = \frac{320}{10} = \mathbf{32} \qquad \bar{\boldsymbol{Y}} = \frac{\sum \mathrm{Y_i}}{\mathrm{n}} = \frac{140}{10} = \mathbf{14}$$

$$\boldsymbol{SS_Y} = \sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2 = 2068 - 10(14)^2 = \mathbf{108}$$

$$\boldsymbol{SS_X} = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2 = 14664 - 10(32)^2 = \mathbf{4424}$$

$$\boldsymbol{SS_{XY}} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{X_i Y_i} - \mathrm{n\overline{X}\overline{Y}} = 5111 - 10(32)(14) = \mathbf{631}$$

$$\hat{\boldsymbol{b}} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{X_i Y_i} - \mathrm{n\overline{X}\overline{Y}}}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{X_i^2} - \mathrm{n\overline{X}^2}} = \frac{631}{4424} = \mathbf{0.14}$$

$$\boldsymbol{SS_E} = SS_Y - \hat{b}\, SS_{XY} = 108 - 0.14(631) = \mathbf{18}$$

$$\boldsymbol{S^2} = \frac{SS_E}{n-2} = \frac{18}{8} = \mathbf{2.25}$$

شكل اختبار الفرضيات هو اختبار من جانبين:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{b} = \boldsymbol{b_0} \\ \mathbf{H_1}: \boldsymbol{b} \neq \boldsymbol{b_0} \end{cases}$$

التوزيع المستخدم هو توزيع ستودنت بدرجة حرية تساوي 8، وحيث مستوى المعنوية $\alpha = 1\%$ فإن القيم الحرجة القيم نجدها عند تقاطع السطر 8 مع العمود 0.025 في جدول توزيع ستودنت، ومنه:

$$-\boldsymbol{t}_{(\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}},\boldsymbol{n}-\mathbf{2})} = -\boldsymbol{t}_{(\mathbf{0.975,8})} = \mathbf{-0.23} \qquad \boldsymbol{t}_{(\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}},\boldsymbol{n}-\mathbf{2})} = \boldsymbol{t}_{(\mathbf{0.975,8})} = \mathbf{0.23}$$

قيمة إحصائية الاختبار تساوي:

$$\mathbf{T} = \frac{\hat{\mathrm{b}} - 0}{\sqrt{\frac{S^2}{\mathrm{SS_X}}}} = \frac{0.14 - 0}{\sqrt{\frac{2.25}{4424}}} = \mathbf{6.21}$$

القرار:

➢ الطريقة الأولى:

قيمة $\boldsymbol{T}$ المحسوبة 6.21 تقع في منطقة الرفض أي خارج المجال ***(-0.23 ;0.23)***، ومنه نرفض فرضية العدم H_0 بمعنى $\boldsymbol{b} = \mathbf{0}$، أي أن قيمة المعامل تختلف عن الصفر.

➢ الطريقة الثانية:

الاختبار من جهتين أي $\mathbf{H_1: b \neq 0}$ ومنه:

$$\boldsymbol{p-value} = 2.P(\mathrm{T} > 6.21) = 0.002$$

نقبل **الفرضية البديلة** H_1 لأنّ **p- value** أقل من مستوى المعنوية 1%. أي أنّ $\boldsymbol{b \neq 0}$.

3-2- اختبار ثابت خط الانحدار الخطي البسيط

ثابت خط الانحدار الخطي البسيط هو قيمة متوسط توزيع $\boldsymbol{Y}$ عندما تكون $\boldsymbol{X}$ صفرا وذلك وفق المعادلة التالية:

$$\boldsymbol{E(Y/\,0) = a}$$

أما في الحالات التي يستحيل فيها في الواقع أن تكون قيمة X مساوية إلى الصفر، فإن القيمة $\boldsymbol{a}$ لن يكون لها أي معنى في حد ذاتها.

التوقع بقيمة المتغير $\boldsymbol{Y}$ عند قيمة معينة للمتغير $\boldsymbol{X}$ ولتكن مثلا $\boldsymbol{X_0}$ يعطى بالعلاقة التالية:

$$\boldsymbol{E(Y/X_0) = a + b\,X_0}$$

ويقدر هذا التوقع بالمقدار

$$\boldsymbol{E(\widehat{Y}/X_0) = \widehat{a} + \widehat{b}\,X_0}$$

ويكون:

$$\boldsymbol{E(\widehat{Y}/X_0) \sim N\left(a + bX_0 \quad , \sigma^2\left[\frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}\right]\right)}$$

وبالتالي:

$$\boldsymbol{\widehat{a} \sim N\left(a\,, \sigma^2\left[\frac{1}{n} + \frac{(\overline{X})^2}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}\right]\right)}$$

بعبارة أخرى:

$$\boldsymbol{\widehat{a} \sim N\left(a\,, \sigma^2\left[\frac{1}{n} + \frac{(\overline{X})^2}{SS_X}\right]\right)}$$

وفي حالة مجهولية $\boldsymbol{\sigma^2}$ نستخدم تباين العينة المسحوبة منه $\boldsymbol{S^2}$. وقيمة $\boldsymbol{\widehat{a}}$ تحسب بالعلاقة $\boldsymbol{\hat{a} = \overline{Y} - \hat{b}\overline{X}}$.

ومنه اختبار الفرضيات حول المعلمة $\boldsymbol{a}$ يأخذ الأشكال التالية:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{a \leq a_0} \\ \mathbf{H_1}: \boldsymbol{a > a_0} \end{cases} \quad , \quad \begin{cases} \mathbf{H_0}: \boldsymbol{a \geq a_0} \\ \mathbf{H_1}: \boldsymbol{a < a_0} \end{cases} \quad , \quad \begin{cases} \mathbf{H_0}: \boldsymbol{a = a_0} \\ \mathbf{H_1}: \boldsymbol{a \neq a_0} \end{cases}$$

إحصائية الاختبار المستخدمة تتبع توزيع ستودنت بدرجة حرية $(n-2)$ **حيث:**

$$\mathbf{T} = \frac{\hat{\boldsymbol{a}} - \boldsymbol{a_0}}{\sqrt{\mathbf{S}^2\left[\frac{\mathbf{1}}{\boldsymbol{n}} + \frac{(\bar{\boldsymbol{X}})^2}{\boldsymbol{SS}_{\mathbf{X}}}\right]}} \sim \mathbf{t}_{(\mathbf{n-2})}$$

مثال (3- 15):

بإستعمال معطيات المثال السابق اختبر معنوية ثابت خط الانحدار الخطي البسيط عند مسوى معنوية تسلوي 1%.

الحل:

من معطيات المثال لدينا:

$$\bar{X} = 32 \quad , \quad \bar{Y} = 14 \quad , \quad SS_Y = 108 \quad , \quad SS_X = 4424 \quad , \quad SS_{XY} = 631$$
$$\hat{b} = 0.14 \quad , \quad SS_E = 18 \quad , \quad S^2 = 2.25$$

ومنه قيمة $\hat{a}$ تكون مساوية إلى:

$$\hat{a} = \bar{\mathrm{Y}} - \hat{\mathrm{b}}\bar{\mathrm{X}} = 14 - 0.1432 \Rightarrow \hat{a} = 9.44$$

وعليه الفرضيات هي:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{a} = \mathbf{0} \\ \mathbf{H_1}: \boldsymbol{a} \neq \mathbf{0} \end{cases}$$

التوزيع المستخدم هو توزيع ستودنت بدرجة حرية تساوي 8، وحيث مستوى المعنوية $\alpha = 1\%$ فإن القيم الحرجة لا تتغير :

$$\boldsymbol{t}_{(1-\frac{\alpha}{2},n-2)} = \boldsymbol{t}_{(0.975,8)} = \mathbf{0.23} \qquad -\boldsymbol{t}_{(1-\frac{\alpha}{2},n-2)} = -\boldsymbol{t}_{(0.975,8)} = \mathbf{-0.23}$$

قيمة إحصائية الاختبار في هذه الحالة هي:

$$\mathrm{T} = \frac{\hat{a} - 0}{\sqrt{\mathrm{S}^2\left[\frac{1}{n} + \frac{(\bar{X})^2}{SS_{\mathrm{X}}}\right]}} = \frac{9.44 - 0}{\sqrt{2.25\left[\frac{1}{10} + \frac{(32)^2}{4424}\right]}} = \mathbf{10.93}$$

القرار: يمكن القيام به بطريقتين.

- الطريقة الأولى:

قيمة $\boldsymbol{T}$ المحسوبة 10.93 تقع في منطقة الرفض أي خارج المجال ***(0.23; 0.23-)***، ومنه نرفض فرضية العدم H_0 ونقبل الفرضية البديلة $\mathbf{H_1}$ بمعنى $\boldsymbol{a} \neq \mathbf{0}$.

- الطريقة الثانية:

الاختبار من جهتين أي $\mathbf{H_1}: \boldsymbol{a} \neq \mathbf{0}$ وبالتالي:

$$\boldsymbol{p-value} = 2.P(\mathrm{T} > 10.93) = 0.001$$

أي أننا نرفض فرضية العدم H_0 لأنّ **p- value** أقل من مستوى المعنوية 1% بعبارة أخرى $\boldsymbol{a} \neq \mathbf{0}$.

3-3- اختبار معامل الارتباط الخطي البسيط

معامل ارتباط بيرسون أو معامل ارتباط حاصل ضرب العزوم $(\boldsymbol{\rho}_{\mathbf{xy}})$ هو مقياس إحصائي لقياس درجة ونوع العلاقة بين متغيرين تربطهما علاقة خطية، وتقع قيمته دائما في المجال $[\mathbf{1}-$ و $\mathbf{1}+]$ حيث يحسب بالعلاقة:

$$\boldsymbol{\rho}_{\mathbf{xy}} = \frac{\boldsymbol{E(XY) - E(X)\,E(Y)}}{\boldsymbol{\sigma}_{\mathbf{x}}\,\boldsymbol{\sigma}_{\mathbf{y}}}$$

ويتم تقدير قيمته من خلال سحب عينة عشوائية حجمها $\boldsymbol{n}$ ويحسب بالعلاقة:

$$\boldsymbol{r_{XY}} = \frac{\sum_{i=1}^{n} \boldsymbol{X_i Y_i - n\bar{X}\bar{Y}}}{\sqrt{\left(\sum_{i=1}^{n} \boldsymbol{X_i^2 - n\bar{X}^2}\right)\left(\sum_{i=1}^{n} \boldsymbol{Y_i^2 - n\bar{Y}^2}\right)}} = \frac{\boldsymbol{SS_{XY}}}{\sqrt{\boldsymbol{SS_X . SS_Y}}}$$

وحيث أن $\widehat{\boldsymbol{b}} = \frac{SS_{XY}}{SS_X}$ فإنه يمكن كتابة معامل ارتباط بيرسون بدلالة $\hat{\mathbf{b}}$ كما يلي:

$$\boldsymbol{r_{XY}} = \sqrt{\frac{\boldsymbol{SS_X}}{\boldsymbol{SS_Y}}}\,\widehat{\boldsymbol{b}}$$

<u>**ملاحظات هامة:**</u>

تجدر الإشارة إلى ما يلي:

1. في نموذج الانحدار الخطي البسيط $(\mathbf{r_{XY}})^2 = \boldsymbol{R^2}$ حيث $\boldsymbol{R^2}$ يسمى بمعامل التحديد.
2. القيمة الموجبة والقريبة من الواحد تدل على أن علاقة الارتباط بين المتغيرين طردية وقوية والعكس صحيح.

بعد حساب معامل الارتباط الخطي للعينة المعطاة، يطرح التساؤل الأتي:

هل قيمة $\boldsymbol{r_{XY}}$ المحسوبة من العلاقات السابقة تدل على وجود علاقة بين المتغيرين في المجتمع الذي سحبت منه العينة؟. بعبارة أخرى، ما هي القيمة التي يجب أن تكون قيمة $\boldsymbol{r_{XY}}$ أكبر منها حتى يكون هناك علاقة ارتباط وأصغر منها حتى تكون علاقة الارتباط ضعيفة، ومن ثمّ لا يوجد ارتباط خطي بين قيم **X** وقيم **Y** ؟.

للإجابة على هذا السؤال نقوم باختبار الفرضيات التالية:

$$\begin{cases} H_0: \rho_{xy} = 0 \\ H_1: \rho_{xy} \neq 0 \end{cases}$$ المتغيران غير مرتبطين خطيا / المتغيران مرتبطان خطيا

$$\begin{cases} H_0: \rho_{xy} \leq 0 \\ H_1: \rho_{xy} > 0 \end{cases}$$ المتغيران غير مرتبطين خطيا / المتغيران مرتبطان خطيا بشكل موجب

$$\begin{cases} H_0: \rho_{xy} \geq 0 \\ H_1: \rho_{xy} < 0 \end{cases}$$ المتغيران غير مرتبطين خطيا / المتغيران مرتبطان خطيا بشكل سالب

احصاءة الاختبار حول معامل الارتباط r_{XY} هي:

$$T = \frac{r_{xy} - \rho_{xy}}{S_{r_{xy}}} = \frac{r_{xy} - 0}{\sqrt{\frac{1-(r_{xy})^2}{n-2}}} = \frac{r_{xy}\sqrt{n-2}}{\sqrt{1-(r_{xy})^2}} \sim t_{(n-2)}$$

حيث:

$S_{r_{xy}} = \sqrt{\frac{1-(r_{xy})^2}{n-2}}$ هو الخطأ المعياري عند اعتبار r_{xy} هو مقدّر لـ ρ_{xy} .

مثال (3- 16):

بإستخدام نفس معطيات المثال السابق، اختبر معنوية معامل الارتباط عند مسوى معنوية تساوي *1*%.

الحل:

$$r_{XY} = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2)(\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2)}} = \frac{SS_{XY}}{\sqrt{SS_X . SS_Y}} = \frac{631}{\sqrt{(4424).(108)}} = \mathbf{0.91}$$

وهي قيمة تقديرية لـ ρ_{xy} .

الفرضية المراد إختبارها هي:

$$\begin{cases} H_0: \rho_{xy} = 0 \\ H_1: \rho_{xy} \neq 0 \end{cases}$$ المتغيران غير مرتبطين خطيا / المتغيران مرتبطان خطيا

قيمة إحصائية الاختبار تساوي:

$$\mathbf{T} = \frac{r_{\text{xy}}\sqrt{\text{n}-2}}{\sqrt{1-\left(\text{r}_{\text{xy}}\right)^2}} = \frac{0.91\sqrt{8}}{\sqrt{1-(0.91)^2}} = \mathbf{6.32}$$

التوزيع المستخدم هو توزيع ستودنت بدرجة حرية تساوي 8، ومستوى المعنوية $\alpha = 1\%$ فإن <u>القيم الحرجة هي</u>:

$$\boldsymbol{t}_{\left(1-\frac{\alpha}{2},n-2\right)} = \boldsymbol{t}_{(0.975,8)} = \mathbf{0.23} \qquad -\boldsymbol{t}_{\left(1-\frac{\alpha}{2},n-2\right)} = -\boldsymbol{t}_{(0.975,8)} = \mathbf{-0.23}$$

<u>**القرار**</u>: يمكن إتخاذ القرار بطريقتين.

- الطريقة الأولى:

قيمة $\boldsymbol{T}$ المحسوبة 6.32 تقع في منطقة الرفض أي خارج المجال***(0.23; 0.23-)***، ومنه نرفض فرضية العدم H_0 أي أن: $\boldsymbol{\rho_{xy} \neq 0}$.

- الطريقة الثانية:

الاختبار من جهتين أي $\mathbf{H_1}: \boldsymbol{\rho_{xy}} \neq \mathbf{0}$ وبالتالي:

$$\boldsymbol{p-value} = 2.P(\text{T} > 6.32) = 0.001$$

أي أننا نرفض فرضية العدم H_0 لأنّ **p- value** أصغر من مستوى المعنوية 1%. أي أنّ: $\boldsymbol{\rho_{xy} \neq 0}$.

تمارين مقترحة

التمرين الأول:

إذا علمنا أن وزن قطعة حلوى من نوع معين لها التوزيع الطبيعي بانحراف معياري 5 غ ، و لدى معاينة عينة تحوي 16 قطعة غذاء من هذا النوع ، تبين أن متوسط وزنها هو 244 غ.
المطلوب: اختبر صحة الفرضية $H_0\text{: } \mu = 250$ ضد الفرضية $H_1\text{: } \mu \neq 250$، وذلك عند مستوى المعنوية $\alpha = 0.05$.

التمرين الثاني:

إذا كان زمن تخثر الدم (بالثانية) لدى عينة عشوائية مكونة من **36** شخصا زاروا مركزا للتبرع بالدم في مدينة معينة كالتالي:

46، 47، 47، 50، 49، 48، 50، 42، 46، 47، 40، 45، 49، 49، 49، 51، 50، 51، 46، 46، 41، 41، 40، 40، 40، 40، 44، 44، 41، 39، 38، 36، 35، 39، 38، 39، 38.

فإذا أدعى مركز التبرع بالدم بأنّ متوسط زمن تخثر الدم هو **40** ثانية وأن تباين زمن تخثر الدم هو **16** ثانية.

المطلوب، اختبر صحة ادعاء مركز التبرع بالدم حول المتوسط عند مستوى معنوية **1%.**

التمرين الثالث:

تدعي شركة لإنتاج البطاريات التي تستخدم في الأجهزة الطبية بأن عمر البطارية من إنتاجها له التوزيع الطبيعي بمتوسط 3 سنوات أخذت عينة عشوائية من إنتاج هذه الشركة تحوي 6 بطاريات ، فكانت أعمارها بالسنوات كما يأتي:

3.5 4.0 0.9 2.9 1.9 3.8

المطلوب: هل نستنتج بأن الشركة تبالغ في ادعائها بالنسبة لمتوسط عمر البطاريات التي تنتجها عند مستوى المعنوية $\alpha=0.01$؟

التمرين الرابع:

كان احتمال ان يصيب احد الرماة الهدف يساوي 0.6، فاذا اعطي مائة طلقة وطلب منه التصويب نحو الهدف وحققت (70) طلقة الهدف، هل يمكن اعتماد هذه النتيجة والقول بأن مستوى الرامي قد تحسن بمستوى ثقة 95%؟.

التمرين الخامس:

تدعي لجنة مختصة أن منتجات شركة صناعة العطور لها مدة استعمال بعد إنتاجها يزيد انحرافها المعياري عن شهرين، ولاختبار هذا الادعاء أخذت عينة حجمها 20 فكان متوسط مدة استعمالها هو 3 سنوات وانحرافها المعياري 3 أشهر.

هل النتائج المتوصل إليها من خلال العينة تفند هذا الادعاء أم تؤيده عند مستوى معنوية 5 % ؟

التمرين السادس:

من سجلات مشفى تبين أنه من بين 1000 رجل دخلوا المشفى كان من بينهم 46 رجلا يعانون مرض القلب، ومن بين 600 إمرأة دخلت المشفى كان من بينهن 18 إمرأة تعاني مرض القلب، هل تقدم هذه المعلومات الإحصائية دلاله كافيه على أن نسبة الإصابة بمرض القلب عند الرجال تساوي نسبة الإصابة بمرض القلب عند النساء عند مستوى معنوية α=0.05 ؟

التمرين السابع:

استخدم المعلومات التالية في اختبار الفرضية $\boldsymbol{H_0: \sigma^2 = 16}$ مقابل $\boldsymbol{H_1: \sigma^2 \neq 16}$:

$$n = 36 \qquad S^2 = 22.54 \qquad \alpha = 1\%$$

التمرين الثامن:

لتكن لديك المعلومات التالية والتي تتعلق بالإنفاق الإستهلاكي وإجمالي الدخل المتاح:

$$n = 10 \qquad \sum_{i=1}^{10} X_i = 170 \qquad \sum_{i=1}^{10} Y_i = 113.5 \qquad \alpha = 5\%$$

$$\sum_{i=1}^{10} {X_i}^2 = 3220 \qquad \sum_{i=1}^{10} {Y_i}^2 = 1383.25 \qquad \sum_{i=1}^{10} X_iY_i = 2103$$

المطلوب، اختبر الفرضيات التالية:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{a \geq a_0} \\ \mathbf{H_1}: \boldsymbol{a < a_0} \end{cases} ، \quad \begin{cases} \mathbf{H_0}: \boldsymbol{b = b_0} \\ \mathbf{H_1}: \boldsymbol{b < b_0} \end{cases} ، \quad \begin{cases} \mathbf{H_0}: \boldsymbol{\rho_{xy} = 0} \\ \mathbf{H_1}: \boldsymbol{\rho_{xy} \neq 0} \end{cases}$$

التمرين التاسع:

البيانات التالية توضح ساعات العمل اليومية لـ 10 عمال في أحد المصانع وكذلك الإنتاج اليومي لكل عامل.

ساعات العمل	10	7	10	5	8	8	6	7	9	10
الإنتاج	11	10	12	6	10	7	9	10	11	10

1. قدر أثر ساعات العمل على الإنتاج؟
2. إختبر الفرضيات التالية:

$$\begin{cases} \mathbf{H_0}: \boldsymbol{a \geq a_0} \\ \mathbf{H_1}: \boldsymbol{a < a_0} \end{cases} ، \quad \begin{cases} \mathbf{H_0}: \boldsymbol{b = b_0} \\ \mathbf{H_1}: \boldsymbol{b < b_0} \end{cases} ، \quad \begin{cases} \mathbf{H_0}: \boldsymbol{\rho_{xy} = 0} \\ \mathbf{H_1}: \boldsymbol{\rho_{xy} \neq 0} \end{cases}$$

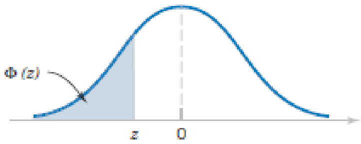
الملحق (1): جدول الأرقام العشوائية

TABLE I - RANDOM DIGITS

36318	75061	37674	26320	75100	10431	20418	19228
91791	76831	58678	87054	31687	93205	43685	19732
44482	66558	37649	08882	90870	12462	41810	01806
26236	33266	66583	60881	97395	20461	36742	02852
04773	12032	51414	82384	38370	00249	80709	72605
12872	14063	93104	78483	72717	68714	18048	25005
48237	41701	73117	33242	42314	83049	21933	92813
72875	38605	29341	80749	80151	33835	52602	79147
26360	64516	17971	48478	09610	04638	17141	09227
55217	13015	72907	00431	45117	33827	92873	02953
97198	12138	53010	94601	15838	16805	61004	43516
57327	38224	29301	31381	38109	34976	65692	98566
99754	31199	92558	68368	04985	51092	37780	40261
76404	86210	11808	12841	45147	97438	60022	12645
98768	04689	87130	79225	08153	84967	64539	79493
99215	84987	28759	19177	14733	24550	28067	68894
63444	21283	07044	92729	37284	13211	37485	10415
95428	33226	55903	31605	43817	22250	03918	46999
39542	71168	57609	91510	77904	74244	50940	31553
59652	50414	31966	87912	87154	12944	49862	96566
95009	27429	72918	08457	78134	48407	26061	58754
66583	62966	12468	20245	14015	04014	35713	03980
75291	71020	17265	41598	64074	64629	63293	53307
37134	54714	02401	63228	26831	19386	15457	17999
88827	09834	11333	68431	31706	26652	04711	34593
05204	30697	44806	96989	68403	85621	45556	35434
99011	14610	40273	09482	62864	01573	82274	81446
94523	97444	59904	16936	39384	97551	09620	63932
89416	52795	10631	09728	68202	20963	02477	55494
34392	96607	17220	51984	10753	76272	50985	97593
55244	70693	25255	40029	23289	48819	07159	60172
74803	97303	88701	51380	73143	98251	78635	27556
41204	47589	78364	38266	94393	70713	53388	79865
61594	26729	58272	81754	14648	77210	12923	53712
19172	08320	20839	13715	10597	17234	39355	74816
75004	86054	41190	10061	19660	03500	68412	57812
65431	16530	05547	10683	88102	30176	84750	10115
55865	07304	47010	43233	57022	52161	82976	47981
26247	18552	29491	33712	32285	64844	69395	41387
34085	58036	09137	47482	06204	24138	24272	16196
58863	96023	88936	51343	70958	96768	74317	27176
27922	28906	55013	26937	48174	04197	36074	65315
22807	10920	26299	23593	64629	57801	10437	43965
33341	77806	12446	15444	49244	47277	11346	15884
12990	23510	68774	48983	20481	59815	67248	17076
86382	48454	65269	91239	45989	45389	54847	77919
12608	18167	84631	94058	82458	15139	76856	86019
64375	74108	93643	09204	98855	59051	56492	11933
62693	35684	72607	23026	37004	32989	24843	01128
61875	23570	75754	29090	40264	80399	47254	40135

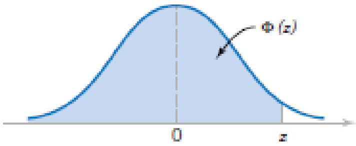
الملحق (2): جدول التوزيع الطبيعي المعياري

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$



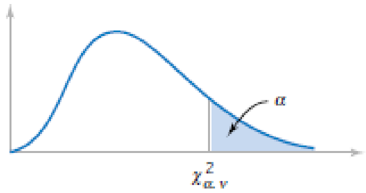
z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.02	−0.01	−0.00
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687
−3.1	0.000711	0.000736	0.000762	0.000789	0.000816	0.000845	0.000874	0.000904	0.000935	0.000968
−3.0	0.001001	0.001035	0.001070	0.001107	0.001144	0.001183	0.001223	0.001264	0.001306	0.001350
−2.9	0.001395	0.001441	0.001489	0.001538	0.001589	0.001641	0.001695	0.001750	0.001807	0.001866
−2.8	0.001926	0.001988	0.002052	0.002118	0.002186	0.002256	0.002327	0.002401	0.002477	0.002555
−2.7	0.002635	0.002718	0.002803	0.002890	0.002980	0.003072	0.003167	0.003264	0.003364	0.003467
−2.6	0.003573	0.003681	0.003793	0.003907	0.004025	0.004145	0.004269	0.004396	0.004527	0.004661
−2.5	0.004799	0.004940	0.005085	0.005234	0.005386	0.005543	0.005703	0.005868	0.006037	0.006210
−2.4	0.006387	0.006569	0.006756	0.006947	0.007143	0.007344	0.007549	0.007760	0.007976	0.008198
−2.3	0.008424	0.008656	0.008894	0.009137	0.009387	0.009642	0.009903	0.010170	0.010444	0.010724
−2.2	0.011011	0.011304	0.011604	0.011911	0.012224	0.012545	0.012874	0.013209	0.013553	0.013903
−2.1	0.014262	0.014629	0.015003	0.015386	0.015778	0.016177	0.016586	0.017003	0.017429	0.017864
−2.0	0.018309	0.018763	0.019226	0.019699	0.020182	0.020675	0.021178	0.021692	0.022216	0.022750
−1.9	0.023295	0.023852	0.024419	0.024998	0.025588	0.026190	0.026803	0.027429	0.028067	0.028717
−1.8	0.029379	0.030054	0.030742	0.031443	0.032157	0.032884	0.033625	0.034379	0.035148	0.035930
−1.7	0.036727	0.037538	0.038364	0.039204	0.040059	0.040929	0.041815	0.042716	0.043633	0.044565
−1.6	0.045514	0.046479	0.047460	0.048457	0.049471	0.050503	0.051551	0.052616	0.053699	0.054799
−1.5	0.055917	0.057053	0.058208	0.059380	0.060571	0.061780	0.063008	0.064256	0.065522	0.066807
−1.4	0.068112	0.069437	0.070781	0.072145	0.073529	0.074934	0.076359	0.077804	0.079270	0.080757
−1.3	0.082264	0.083793	0.085343	0.086915	0.088508	0.090123	0.091759	0.093418	0.095098	0.096801
−1.2	0.098525	0.100273	0.102042	0.103835	0.105650	0.107488	0.109349	0.111233	0.113140	0.115070
−1.1	0.117023	0.119000	0.121001	0.123024	0.125072	0.127143	0.129238	0.131357	0.133500	0.135666
−1.0	0.137857	0.140071	0.142310	0.144572	0.146859	0.149170	0.151505	0.153864	0.156248	0.158655
−0.9	0.161087	0.163543	0.166023	0.168528	0.171056	0.173609	0.176185	0.178786	0.181411	0.184060
−0.8	0.186733	0.189430	0.192150	0.194894	0.197662	0.200454	0.203269	0.206108	0.208970	0.211855
−0.7	0.214764	0.217695	0.220650	0.223627	0.226627	0.229650	0.232695	0.235762	0.238852	0.241964
−0.6	0.245097	0.248252	0.251429	0.254627	0.257846	0.261086	0.264347	0.267629	0.270931	0.274253
−0.5	0.277595	0.280957	0.284339	0.287740	0.291160	0.294599	0.298056	0.301532	0.305026	0.308538
−0.4	0.312067	0.315614	0.319178	0.322758	0.326355	0.329969	0.333598	0.337243	0.340903	0.344578
−0.3	0.348268	0.351973	0.355691	0.359424	0.363169	0.366928	0.370700	0.374484	0.378281	0.382089
−0.2	0.385908	0.389739	0.393580	0.397432	0.401294	0.405165	0.409046	0.412936	0.416834	0.420740
−0.1	0.424655	0.428576	0.432505	0.436441	0.440382	0.444330	0.448283	0.452242	0.456205	0.460172
0.0	0.464144	0.468119	0.472097	0.476078	0.480061	0.484047	0.488033	0.492022	0.496011	0.500000

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$



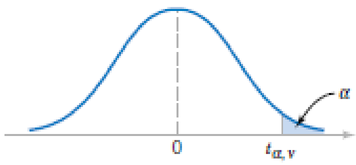
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959071	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965621	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967

الملحق (3): جدول توزيع مربع كاي



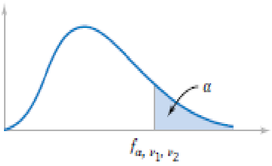
v \ α	.995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.00+	.00+	.00+	.00+	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

الملحق (4): جدول توزيع ستودنت



ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

الملحق (5): جداول توزيع فيشر



$$f_{0.25,v_1,v_2}$$

Degrees of freedom for the numerator (ν_1); Degrees of freedom for the denominator (ν_2)

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.41	9.49	9.58	9.63	9.67	9.71	9.76	9.80	9.85
2	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.41	3.43	3.43	3.44	3.45	3.46	3.47	3.48
3	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.46	2.46	2.46	2.47	2.47	2.47	2.47	2.47
4	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
5	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.88	1.88	1.88	1.88	1.87	1.87	1.87
6	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.76	1.76	1.75	1.75	1.75	1.74	1.74	1.74
7	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.70	1.69	1.68	1.68	1.67	1.67	1.66	1.66	1.65	1.65	1.65
8	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.62	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58
9	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.54	1.53	1.53
10	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.53	1.52	1.52	1.51	1.51	1.50	1.49	1.48
11	1.47	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.47	1.46	1.45
12	1.46	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.45	1.44	1.43	1.42
13	1.45	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.42	1.41	1.40
14	1.44	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.41	1.41	1.40	1.39	1.38
15	1.43	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.46	1.45	1.44	1.43	1.41	1.41	1.40	1.39	1.38	1.37	1.36
16	1.42	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.44	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
17	1.42	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33
18	1.41	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33	1.32
19	1.41	1.49	1.49	1.47	1.46	1.44	1.43	1.42	1.41	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.33	1.32	1.30
20	1.40	1.49	1.48	1.47	1.45	1.44	1.43	1.42	1.41	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31	1.29
21	1.40	1.48	1.48	1.46	1.44	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28
22	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28
23	1.39	1.47	1.47	1.45	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28	1.27
24	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.36	1.35	1.33	1.32	1.31	1.30	1.29	1.28	1.26
25	1.39	1.47	1.46	1.44	1.42	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.29	1.28	1.27	1.25
26	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.35	1.34	1.32	1.31	1.30	1.29	1.28	1.26	1.25
27	1.38	1.46	1.45	1.43	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.33	1.32	1.31	1.30	1.28	1.27	1.26	1.24
28	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.31	1.30	1.29	1.28	1.27	1.25	1.24
29	1.38	1.45	1.45	1.43	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.32	1.31	1.30	1.29	1.27	1.26	1.25	1.23
30	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.34	1.32	1.30	1.29	1.28	1.27	1.26	1.24	1.23
40	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.31	1.30	1.28	1.26	1.25	1.24	1.22	1.21	1.19
60	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.27	1.25	1.24	1.22	1.21	1.19	1.17	1.15
120	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.26	1.24	1.22	1.21	1.19	1.18	1.16	1.13	1.10
∞	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.22	1.19	1.18	1.16	1.14	1.12	1.08	1.00

$$f_{0.10,v_1,v_2}$$

v_2 \ v_1	Degrees of freedom for the numerator (v_1)																	
Degrees of freedom for the denominator (v_2)	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.03	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

$$f_{0.05,v_1,v_2}$$

v_2 \ v_1	Degrees of freedom for the numerator (v_1)																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	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	5.77	5.75	5.72	5.69	5.66	5.63
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	4.53	4.50	4.46	4.43	4.40	4.36
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	3.84	3.81	3.77	3.74	3.70	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	3.41	3.38	3.34	3.30	3.27	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	3.12	3.08	3.04	3.01	2.97	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	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	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	2.61	2.57	2.53	2.49	2.45	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	2.51	2.47	2.43	2.38	2.34	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	2.42	2.38	2.34	2.30	2.25	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	2.35	2.31	2.27	2.22	2.18	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	2.29	2.25	2.20	2.16	2.11	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	2.24	2.19	2.15	2.11	2.06	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	2.19	2.15	2.10	2.06	2.01	1.96
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	2.15	2.11	2.06	2.02	1.97	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	2.11	2.07	2.03	1.98	1.93	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	2.08	2.04	1.99	1.95	1.90	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	2.05	2.01	1.96	1.92	1.87	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	2.03	1.98	1.94	1.89	1.84	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	2.01	1.96	1.91	1.86	1.81	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	1.98	1.94	1.89	1.84	1.79	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	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
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	1.89	1.84	1.79	1.74	1.68	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	1.79	1.74	1.69	1.64	1.58	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	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.55	1.43	1.35	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	1.52	1.46	1.39	1.32	1.22	1.00

Degrees of freedom for the denominator (v_2)

$f_{0.025,v_1,v_2}$

v_2 \ v_1	Degrees of freedom for the numerator (v_1)																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

Degrees of freedom for the denominator (v_2)

$$f_{0.01,v_1,v_2}$$

v_2 \ v_1 (Degrees of freedom for the denominator (v_2))	Degrees of freedom for the numerator (v_1)																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.00	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.46
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.36	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.59
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

1. باللغة العربية

- إمتثال محمد حسن، عادل محمود حلاوة، لبيبة حسب النبي العطار، **مقدمة في أساليب الاستدلال الإحصائي والتنبؤ**، الطبعة الأولى، مكتبة الوفاء القانونية، الإسكندرية، 2012.
- إياد محمد الهوبي، **الإحصاء التطبيقي**، الطبعة الأولى، الكلية الجامعية للعلوم والتكنولوجيا، خان يونس– فلسطين، 2014.
- جورج كانافوس، دون ميلر، **الإحصاء للتجاريين– مدخل حديث**، تعريب سلطان محمد عبد الحميد، دار المريخ، الرياض– السعودية، 2004.
- سمير سليم العبيدي، جمال إبراهيم البياتي، **الإحصاء التطبيقي– طرق اختبارات الفروض وموضوعات أخرى في الاستدلال الإحصائي**، دار شموع الثقافة، بنغازي– ليبيا، 2002.
- عبد الحميد عبد المجيد البلداوي، **الإحصاء للعلوم الإدارية والتطبيقية**، الطبعة الأولى، دار الشروق للنشر والتوزيع، عمان– الأردن، 1997.
- نجاة رشيد الكيخيا، **أساسيات الاستنتاج الإحصائي**، دار المريخ للنشر، المملكة العربية السعودية، 2007.

2. باللغة الأجنبية

- Armitage, P & Berry, G, **Statistical Methods in Medical Research**, Third Edition, Blackwell Scientific Publication, Oxford, 1994.
- Cochran, W. G, **Sampling techniques**, Third Edition, Wiley, New York, 1997.
- Douglas C. Montegomery, George C. Runger, **Applied Statistics and Probability for Engineers**, Third Edition, John Wiley & Sons, Inc, USA, 2002.
- François Dress, **Les Probabilités et Statistique de A a Z- 500 définitions, formules et tests d'hypothèse**, Dunod, Paris, 2013.
- Khaldi Khaled, **Methodes Statistiques- Rappels de cours Exercices corrigés**, Offices Des Publications Universitaires, Alger, 2005.

الدكتور المختار بن سكري،

من مواليد بلدية البيرين (ولاية الجلفة)، في 28 فيفري 1979

دكتوراه علوم في الاقتصاد و الإحصاء التطبيقي، تخصص: طرق كمية

أستاذ محاضر - ب - بجامعة الجلفة - الجزائر -

محتوى المطبوعة:

هذه المطبوعة هي مجموعة محاضرات وتمارين خاصة بمقياس الإحصاء الاستدلالي أو ما يسمى إحصاء 03، وهي موجهة لطلبة السنة الثانية ليسانس علوم التسيير، وهذا من أجل معرفة كل ما له علاقة بالمتغيرات الكمية والنوعية ومعرفة المجتمع والعينة وكل القوانين التوزيعية المستمرة والمتقطعة المستخدمة مما يسمح للطالب سهولة فهم الظواهر الاقتصادية وغيرها ومرونة التعامل مع كامل المعايير الإحصائية المفسرة لها (تعميم النتائج المتحصل عليها من العينة على المجتمع، صياغة الفروض واختيار الأدوات والأساليب المناسبة في حل المشكلات الاقتصادية) .

إن إعداد هذه المطبوعة جاء لتحقيق جملة من الأهداف ونوجزها فيما يلي:

1. تزويد الطالب بمقدمة عامة حول الإحصاء 03 (الإحصاء الاستدلالي).
2. تعريف الطالب بأساليب استخراج العينات من المجتمعات ومن ثم القدرة على استيعاب خصائص العينة والمجتمع الأصلي قيد الدراسة.
3. معرفة كيفية اختيار المعايير المناسبة والعينات لدراسة ومقارنة الظواهر المختلفة.
4. فهم أهم أنواع التوزيعات الإحصائية المستخدمة في الإحصاء 03.
5. إكساب الطالب القدرة على التحليل الإحصائي بتعريفه إتباع الأساليب الإحصائية لغرض الوصول إلى النتائج الدقيقة بأسهل الطرق واقل التكاليف .
6. تقدير المعلمات واختبار الفروض لكل الحالات المختلفة وتفسير نتائج الاختبار وفهم المعنى الاقتصادي لها .