الجمهـورية الجزائــرية الديمقـراطية الشعبيـــة
République Algérienne Démocratique et Populaire
وزارة التعليـم العــالي و البحــث العلمـي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جــامعة محــمد خيضــر – بسكرة –
كـلية العلـوم الاقتصـادية و التجـارية و علـوم التسييـر

قسم العلوم المحاسبية

محاضرات وتطبيقات في مقياس:
إحصاء 3

موجهة لطلبة السنة الثانية ليسانس علوم مالية ومحاسبية
السداسي الثالث

إعداد الأستاذة:

- ***د/ دريدي أحلام***
أستاذة محاضرة أ

المــوسم الجـامعي:2021-2022

فهرس المحتويات

تمهيد

هذه المطبوعة هي عبارة عن محاضرات وتطبيقات في مقياس إحصاء 3 والموجهة لطلبة السنة الثانية حيث يهدف هذا المقياس إلى تعريف الطلبة بمواضيع مختلفة في الإحصاء، هذا العلم الذي تختلف الرؤية له فالناس العاديون يرون الإحصاء مجرد مجموعة جداول والعديد من الأرقام والأشكال البيانية، أما الطلبة فيتعرفون على أنواعه المختلفة كالإحصاء الوصفي والإحصاء الإستدلالي والإحصاء التطبيقي وكل أنواع الإحصاء لا يستغني عنها الطالب مستقبلا في مختلف مجالات المعرفة وإتخاذ القرار بخصوص المشكلات التي سيتم دراستها

بناءا على ما تقدم فهذه المطبوعة هي مجهود تدريس هذا المقياس بعد الاطلاع على العديد من المراجع وإختيار أهم المواضيع التي تهم الطالب وترسخ المفاهيم المبدئية التي درسها في السنوات السابقة،

ويتضمن برنامج المقياس أربعة محاور رئيسية حيث تعرضنا في المحور الأول إلى مفاهيم أساسية في حول الإحصاء وأهم الإختبارات الإحصائية مع التعرف على نظرية النهاية المركزية التي سنحتاجها في المحاور اللاحقة، ثم في المحور الثاني تعرفنا على توزيع المعاينة حيث تم شرح توزيع المعاينة للوسط الحسابي للمجتمع والفرق بين متوسطين، ونسبة المجتمع وكذلك الفرق بين نسبتين، وتباين المجتمع ونسبة تباينين ، ثم في المحور الثالث تعرفنا على نظرية التقدير الإحصائي حيث بداية تطرقنا إلى التقدير بنقطة ثم التقدير بفترة لمعلة من معلمات المجتمع، حيث تعرفنا على كيفية تحديد فترة الثقة للمتوسط الحسابي والفرق بين متوسطين، والنسبة والفرق بين نسبتين ، وتباين المجتمع ونسبة تباينين، وفي المحور الرابع الأخير تطربقنا إلى إختبار الفرضيات حيث تعرفنا على مفاهيم أساسية في إختبار الفروض ثم تم شرح إختبار الفروض لوسط المجتمع والفرق بين متوسطين، والنسبة والفرق بين نسبتين ، وتباين المجتمع ونسبة تباينين، وقد دعمت المحاور الأربعة بالعديد من الأمثلة وفي نهاية كل محور تم تقديم سلسلة تمارين شاملة

وقد حاولنا تقديم محاور هذه المطبوعة بطريقة وشرح بسيط بغية تمكين الطالب من فهم هذه المواضيع المهمة التي ستخدمه في مختلف أبحاثه مستقبلا

الدكتورة: دريدي أحلام

المحور الأول:

مفاهيم أساسية حول الإحصاء وأهم الإختبارات

و التوزيعات الإحصائية

أولا: مفاهيم أساسية حول علم الإحصاء.

أستخدم الإحصاء منذ القدم وبداية الإنسان فعندما كان الإنسان يحصي أفراد عائلته وعشيرته أو ما يمتلكه من مواشي بالفطرة فهو تطبيق لمعنى الإحصاء وكان يعبر عن ذلك في صور، لذلك سمي "الإحصاء الصوري" وعندما بدأ باستخدام الأرقام سمي "الإحصاء العددي" وقد استخدم البابليون والآشريون والفراعنة الإحصاء العددي في عد الجند والأسلحة والمعدات الحربية وكذلك بالنسبة لإحصاء الإنتاج الزراعي وتخمين الضرائب، وسمي "الحساب السياسي أو حساب الدولة" عندما أستخدم الإحصاء في جمع المعلومات المتعلقة بشؤون الدولة، حيث كلمة "statistics" تعني ذلك في أوروبا، وسمي "بعلم الأوساط أو المعدلات" عندما أستخدمت كمؤشرات إحصائية كما أطلق عليه "علم الأعداد الكبيرة" واستمر الحال حتى منتصف القرن السادس عشر وظهور "الإحصاء الاستدلالي" والذي يعتمد على نظرية الاحتمالات كنتيجة لانتشار لعب الورق في أوروبا إذ اتجه اللاعبون إلى علماء الرياضيات لإعطائهم معلومات حول فرص ربحهم أو خسارتهم المتوقعة ومن أشهر العلماء [Bernoulli, Fermat, Lebinitz, Pascal] وهم مؤسسي نظرية الاحتمالات، ومع ظهور الآلة الحاسبة واستخدام أساليب جديدة في التعامل مع البيانات والمعلومات الإحصائية وتحليلها تطور "الإحصاء الاستدلالي" إلى "الإحصاء الاستنتاجي" وأصبح الحصر الشامل أو التعداد العام للسكان أحد المؤشرات الأساسية في وضع الخطط والبرامج التنموية المهمة.[1]

[1]– خالد أحمد فرحان المشهداني، مبادئ الإحصاء والاحتمالات (الأسس النظرية والعملية)، دار الأيام للنشر والتوزيع، عمان، الأردن، 2015، ص، ص: 17، 19.

1- تعريف علم الإحصاء:

يمكن تعريف علم الإحصاء بأنه:

- "جمع وعرض وتحليل وتفسير البيانات"،
- كما يمكن تعريفه: "علم التقدير والاحتمالات"[1]
- ويعرف أيضا بأنه: "العلم أو مجموعة القواعد والطرق والنظريات التي تهتم بجانب البيانات وتبويبها أو عرضها جدوليا أو بيانيا ثم تحليلها وتفسيرها وإجراء المقارنات واستنتاج العلاقات بهدف استخدامها في اتخاذ القرارات المناسبة".[2]
- كما يعرف بأنه: "العلم الذي يبحث في جمع البيانات وعرضها وتحليلها وتبويبها واستخدام النتائج في التنبؤ أو التقرير أو التحقق".

أي أن هذا التعريف أشار إلى:

(1) المشاهدة والملاحظة: فالباحث يشاهد ويلاحظ ما يحدث ويجمع الحقائق المتعلقة بالمشكلة.

(2) الفرضية: لتفسير الحقائق المشاهدة، يضع الباحث العديد من الفرضيات.

(3) التنبؤ: يستنتج الباحث من فرضياته بعض الحقائق الجديدة التي يمكن اعتبارها معرفة جديدة يطلق عليها اسم التنبؤ.

(4) التحقق: مرحلة التأكد من صحة الفرضية التي يفسر بها المشكلة.[3]

وفي الأخير يمكن القول أنّ علم الإحصاء هو العلم الذي يهتم بالأعداد وتجميع مختلف البيانات الرقمية وتنظيمها وتصنيفها وعرضها باستخدام مختلف الأدوات المساعدة كالجدول والدوائر النسبية والمدرجات التكرارية... من أجل تحليلها ووضع مختلف الفرضيات من أجل اختبارها والوصول إلى نتائج يمكن أن تقيد متخذ القرار في عمليات التنبؤ مستقبلا.

---

[1]– وحيد مصطفى أحمد: أساسيات علم الإحصاء، دار الكتب للنشر والتوزيع، القاهرة، مصر، ص 15.

[2]– أحمد عودة بن عبد المجيد عودة ومنصور عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 1)، مكتبة الفلاح للنشر والتوزيع، الطبعة 3، الكويت، 2014، ص 16.

[3]– محمد حسين محمد رشيد: الإحصاء الوصفي والتطبيقي والحيوي، دار الصفاء للطباعة والنشر والتوزيع، الطبعة 2، عمان (الأردن)، 2015، ص 15.

2- مراحل علم الإحصاء:

تقسم مراحل علم الإحصاء إلى:

- الجمع: تعتبر الأساس في التحليل الإحصائي فإذا كانت البيانات غير دقيقة فإن الاستنتاج والقرارات المبنية عليها لذلك يجب الدقة في الحصول على البيانات سواء من المصادر المنشورة أو غير المنشورة أو يجمعها في الميدان.
- التنظيم: إذا أخذت البيانات من المصادر المنشورة أو غير المنشورة غالبا ما تكون منظمة أما التي يحصل عليها من الميادين فهي بحاجة إلى تنظيم ويتم ذلك من خلال معالجة المشاكل والتباين في المعلومات وعدم علاقتها بموضوع الدراسة ومن ثم تبويبها على شكل جداول تكرارية.
- التقديم: تقديم البيانات والمعلومات من خلال عرضها بأشكال هندسية أو رسومات بيانية.
- التحليل: أساليب التحاليل كثيرة ومتعددة تمتد في المشاهدة البسيطة إلى الأساليب الرياضية المعقدة والتحليل الإحصائي يركز على البيانات المبوبة في الجداول التكرارية.
- التفسير: تعتبر هذه المرحلة أهم مراحل البحث الإحصائي وتحتاج إلى درجة عالية من المهارة والخبرة وذلك بسبب استخلاص النتائج من البيانات التي تمّ جمعها وتحليلها وفي ضوء ذلك يتم اتخاذ القرارات المناسبة.[1]

3- أنواع الإحصاء:

هناك من يقسمها إلى الإحصاء الوصفي والاستدلالي فقط حيث هناك من يضم الإحصاء الاستنتاجي ضمن الاستدلالي وهناك من يفرقهما، وفيما يلي بعض أنواع الإحصاء:

- الإحصاء الاستقرائي: ظهر بوجود الإنسان في الأرض حتى ظهور الإحصاء الوصفي.
- الإحصاء الوصفي: بدأ هذا العلم مع بدأ الإنسان في استخدام أسلوب العد واستخدام الأرقام واستخراج المتوسطات أو المعدلات بالاعتماد على أساليب التحليل الاستنباطي، أي يهتم بطرق جمع تمثيل وعرض البيانات.

---

[1]- سهيل أحمد سمحان ومحمد حسين الوادي: مبادئ الإحصاء للاقتصاد والعلوم الإدارية، دار الصفاء للنشر والتوزيع، الطبعة 2، عمان- الأردن، 2014، ص ص: 29/28.

- الإحصاء الاستدلالي: ظهر في منتصف القرن السادس عشر عندما ظهر علم الاحتمالات إلى حيز التطبيق العملي.
- الإحصاء الاستنتاجي: ويبدأ عند تحليل البيانات والمعلومات الإحصائية في الوصول إلى استنتاجات والتوصيات وله فرعين التقدير واختبار الفرضيات.[1]

4- بعض المفاهيم المهمة:

الوحدة الإحصائية: هي الوحدة الأساسية لتكوين المجتمع الإحصائي.

الظاهرة الإحصائية: هي الخاصية المدروسة أو المتغير المدروس في المجتمع الإحصائي مثلا: الطول، الوزن، الإنتاج...[2]

المجتمع: هو المجموعة الكلية لمفردات الدراسة سواء كانت أفراد أو أشياء واستخلاص خصائص هذا المجتمع هو الهدف النهائي للدراسة الإحصائية وقد يكون المجتمع محدود أو غير محدود.

العينة: مجموعة جزئية من مفردات المجتمع محل الدراسة يتم اختيارها بحيث تكون ممثلة للمجتمع تمثيل صحيح.[3]

5- أنواع المتغيرات:

أي ظاهرة تظهر اختلاف بين مفرداتها ويرمز له ب X مثلا وتنقسم المتغيرات إلى:

- متغيرات وصفية أو نوعية: هي تلك الظواهر أو الصفات التي لا يمكن قياسها بالأرقام مثل لون العيون (أزرق، أسود)، الحالة الاجتماعية (غني، فقير)، الجنس (ذكر، أنثى).
- متغيرات كمية: هي الظواهر أو الصفات التي يمكن قياسها بالأرقام مثل (الطول، الوزن) وتنقسم إلى:

---

[1]- خالد أحمد فرحان المشهداني، مرجع سابق، ص، ص: 20/19.

[2]- جيلاطو جيلالي: الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الطبعة 9، الجزائر، 2012، ص: 5.

[3]- قسم الإحصاء بجامعة الملك عبد العزيز: مبادئ الإحصاء للتخصصات النظرية الإدارية والإنسانية، دار خوارزم العلمية للنشر والتوزيع، الطبعة 9، جدة، السعودية، 2015، ص 12.

أ- متغيرات مستمرة أو متصلة: وهي المتغيرات التي تأخذ المشاهدة أو المفردة فيها أي قيمة رقمية في مدى معين كوحدات الزمن أو الوزن وبصورة عامة كل البيانات التي تقاس تعتبر بيانات متغير مستمر.

ب- المتغيرات الغير مستمرة المنفصلة: وهي المتغيرات التي تأخذ فيها المشاهدة أو المنفردة قيما متباعدة أو متقطعة كعدد أبناء الأسرة وبصورة عامة كل البيانات التي نحصل عليها من العد تعتبر ببيانات لمتغير منفصل.[1]

6- أنواع العينات:

يمكن تقسيم العينات إلى مجموعتين العينات الغير عشوائية (الغير الاحتمالية)، العينات العشوائية (الاحتمالية).

النوع الأول: العينات الغير عشوائية (الغير الاحتمالية): هي العينات التي لا تحقق مبدأ العشوائية عند اختيار مفرداتها ونذكر نوعين منها:

- العينة الغرضية (العمدية): هي العينة التي يتم فيها اختبار أفضل أو أكفأ العناصر أو المفردات القادرة أكثر من غيرها على تحقيق الغرض الذي اختيرت من أجله فمثلا عند اختيار طلبة لتمثيل الكلية في مسابقة حفظ القرآن سيتم اختيار الأفضل ولا يمكن أن نختار عشوائيا.
- العينة الحصصية: حيث يتم تحديد حصة أو نصيب لكل فرد من مساعدي الباحث من مفردات العينة على أن تترك للمساعدة الحرية الكاملة في اختبار حصته أو نصيبه من العينة.

مثال: إذا كان المطلوب اختيار عينة حصصية من طلاب إحدى الكليات حجمها 100 طالب ويساعد الباحث ثلاثة مساعدين وتم تحديد حصتهم كما يلي:

الأول: حصته من العينة 30 طالب.

الثاني: حصته من العينة 60 طالب.

الثالث: حصته من العينة 10 طلبة.

هنا في هذه الحالة تترك الحرية الكاملة لكل مساعد في اختيار العدد المحدد له بالطريقة التي يراها مناسبة.

[1]- خالد أحمد فرحان المشهداني، مرجع سابق، ص، ص: 24/23.

النوع الثاني: العينات العشوائية: وتعتبر الأهم لأنه من خلالها يمكن الاستدلال على المجتمع الإحصائي المسحوبة منه وتوجد عدة أنواع وسنحاول الإجابة على سؤالين: متى؟ وكيف يتم اختيار كل منها مع ذكر مثال:

أ- العينة العشوائية البسيطة: تكون هذه العينة مناسبة إذا كان المجتمع محل الدراسة صغيرا ومتجانس ويتم اختيار مفرداتها عن طريق الكيس المثالي أو جداول الأرقام العشوائية.

مثال: اختيار عينة عشوائية من تلاميذ السنة الأولى ابتدائي ومن إحدى المدارس، فالمجتمع هنا هو تلاميذ السنة الأولى بالمدرسة وهو مجتمع صغير ومتجانس ومعنى متجانس هو أن كل التلاميذ يدرسون المقررات نفسها، ونسحب العينة ونرقمها.

ب- العينة الطبقية: هنا تكون العينة مناسبة إذا كان المجتمع غير متجانس أو مجتمع طبقي والمقصود بالمجتمع الطبقي أو غير متجانس أنه يتكون من عدة أقسام أو أجزاء أو مجموعات.

مثال: كلية بها 1000 طالب وتتكون من ثلاث أقسام وعدد الطلاب بهذه الأقسام على الترتيب 500، 300، 200، والعينة الطبقية المطلوبة 10 % من المجتمع أي حجمها 100 طالب، والجدول التالي يمثل حجم المجتمع والعينة ونسبة كل طبقة في المجتمع والعينة.

القسم / المعلومات	الأول	الثاني	الثالث	المجموع
حجم المجتمع (عدد الطلبة بالكلية)	500	300	200	1000
حجم العينة (عدد الطلبة بالعينة)	50	30	20	100
نسبة حجم الطلبة	% 50	% 30	% 20	

نلاحظ أن الاختيار من كل قسم بما يتناسب وحجم هذا القسم بالكلية أي يتم اختيار 50 % من القسم الأول و30 % من الثاني و20 % الثالث بطريقة عشوائية بسيطة لأن كل قسم هو مجتمع متجانس، وبهذه الطريقة نضمن أن كل طبقات المجتمع تم تمثيلها بالعينة وقد تم تمثيل كل طبقة بالعينة بالنسبة نفسها الموجودة بالمجتمع وأن الاختيار من كل طبقة يتم بطريقة عشوائية بسيطة.

ج- العينة المتعددة المراحل: ويتم اختيار هذا النوع إذا كان المجتمع الإحصائي كبيرا ومنتشرا على مساحات جغرافية شاسعة ويتم الاختيار على عدة مراحل على أن يكون الاختيار في كل مرحلة بطريقة عشوائية.

مثال: اختيار عينة عشوائية من تلاميذ السنة الخامسة في ولاية بسكرة في هذه الحالة المجتمع كبير ومنتشر، حيث في المرحلة الأولى نقسم الولاية إلى دوائر ونختار عينة عشوائية من دوائر الولاية وفي المرحلة الثانية نقسم هذه الدوائر إلى بلديات ثم نختار عينة عشوائية من هذه البلديات، وفي المرحلة الثالثة نقوم بحصر هذه المدراس ونختار عينة عشوائية وفي المرحلة الرابعة نقوم بحصر عدد الطلاب بالسنة الخامسة بهذه المدارس التي تم اختيار هذه العينة عبر أربع مراحل ويكون الاختيار في كل مرحلة بطريقة عشوائية، أي أن هناك أربع مراحل ونلاحظ أن المفردة كل مرحلة تختلف وهي تسمى وحدة المعاينة حيث (دائرة، بلدية، مدرسة، تلميذ سنة خامسة) حيث العينة في النهاية تحتوي تلاميذ السنة الخامسة.

د- العينة المنتظمة: ويتم اختيار هذه العينة وفق الخطوات التالية:

- اختيار وحدة المعاينة K والتي تسمى نسبة المعاينة $K = \frac{N}{n}$
- اختيار رقم عشوائي بين 1-K ليكون رقم العينة الأولى.
- نضيف K بالتعاقب على رقم العينة الأولى وصولا إلى آخر وحدة معاينة.

مثال: مصنع به 100 عامل مرقمون بالتسلسل من 1 إلى 100 ونريد اختيار عينة عشوائية من عمال هذا المصنع للعمل في دوام ليلي قدرها 20 % أي 200 عامل وبقسمة 100 على 20 نجد 5، ثم نختار رقم عشوائي من 1 إلى 5 فمثلا نختار 4 وفي كل مرة نضيف 5 حتى نصل إلى 20 مفردة أي: [4، 9، 14، 19، 24، 29، 34، 39، 44، 49، 54، 59، 64، 69، 74، 79، 84، 89، 94، 99].

أي أن المفردة الأولى فقط تم اختيارها بطريقة عشوائية ثم تم اختيار باقي المفردات بطريقة منتظمة.

ه- العينة العنقودية: عينة العناقيد أو عينة المساحة أهم مثال عنها هو اختيار عينة عنقودية من مناطق أو أحياء إحدى المدن فالأحياء هي حبات العنقود وبكل حي هناك كمختلف المستويات الاجتماعية بمعنى أن الأحياء متجانسة وندرس جميع المفردات العائلات داخل الحي الذي تم اختياره والفرق بينها وبين الطبقية أن الطبقية تدرس جميع الطبقات التي تدخل في الدراسة على أن يتم اختيار

عينة عشوائية في كل طبقة أما العنقودية يتم اختيار عينة عشوائية من الطبقات (الأحياء) على أن يتم دراسة جميع المفردات داخل الطبقة (الحي) المختار. [1]

❋ طريقة العينة المعيارية: هي تلك العينة التي تتفق مع المجتمع الاحصائي من حيث مقاييسه الإحصائية كالوسيط، الانحراف المعياري ونختار مثل هذه العينات بطريقة تتابعية مثل تقدير نجاح عملية معينة فهنا سنختار المرضى الذين هم بحاجة لعملية فتقدر نسبة النجاح لأول عشرة مرضى ثم لأول 20 مريض حتى تستقر النسبة وبعدها نعمم النتيجة وهذه العينة تختار بدقة وبعناية وبشكل تتابعي.[2]

7- طرق جمع البيانات الإحصائية:

تجمع البيانات بإحدى الطرق التالية:

أ: طريقة المسح الشامل: تجمع البيانات الإحصائية من جميع أفراد المجتمع دون استثناء وتمتاز هذه الطريقة بأنها تعطي صورة مفصلة عن جميع أفراد المجتمع الإحصائي وبأنها لا تحتوي أخطاء سببها استثناء بعض عناصر المجتمع الإحصائي.

ب: طريقة العينة: هنا تجمع البيانات من مجموعة جزئية من المجتمع الإحصائي وتعمم النتيجة من الجزء إلى الكل لكن هناك حالات يتعذر فيها المسح الشامل فنستخدم طريقة العينة منها:

- فساد عناصر المجتمع الإحصائي: إذا أردنا فحص دم مريض مستحيل أخذ دمه كله لأنه يؤدي إلى الوفاة فنأخذ عينة فقط.
- عندما لا تتوافر جميع عناصر المجتمع الإحصائي: إذا أردنا دراسة كميات الأمطار وأثرها على كمية إنتاج الحبوب مثل: 1995.
- الجهد والوقت والتكاليف: لأن كل دراسة إحصائية مرتبطة بهذه الظروف.
- يحتاج المسح الشامل إلى عدد كبير من الأشخاص: لجمع البيانات الإحصائية وينتج عن ذلك أخطاء متعددة منها الفروق الفردية لذلك أخذ عينة بواسطة عدد قليل من المختصين سبب من أسباب تقليل الأخطاء وبالتالي تؤدي إلى نتائج أكثر دقة.

[1]– أحمد عودة بن عبد المجيد عودة، منصور عبد الرحمن القاضي، الإحصاء الوصفي والاستدلالي (الجزء 1)، مرجع سابق، ص ص: 41/34.

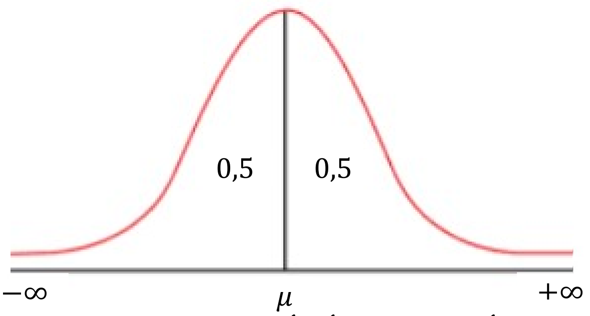
[2]– محمد حسين محمد رشيد، مرجع سابق، ص 25.

- عندما يكون المجتمع متصلا: كأن تكون مجموعة عناصر المجتمع غير قابلة للعد كمخزون الغاز الطبيعي لمعرفته يجب تنقيب كل أراضي الدولة لكن هذا غير ممكن عمليا لهذا تؤخذ عينة من تلك الأراضي وتجرى عملية التنقيب فيها.[1]

ثانيا: التوزيع الطبيعي Z.

يعتبر التوزيع الطبيعي أو المعتدل من أشهر التوزيعات الاحتمالية لأنه يصف معظم الظواهر الطبيعية مثل درجات الحرارة، الوزن، درجات الطلبة، وهو يعرف أيضا باسم توزيع جاوس (Gauss)، كما أنه يستخدم كتقريب لكثير من التوزيعات الاحتمالية، وهذه الميزة تجعله مؤهلا لاستخدامات واسعة النطاق في مجالات عديدة من فروع الإحصاء والاستدلال الإحصائي لكثير من البحوث.[2]

ويمكن تعريف التوزيع الطبيعي بأنه: توزيع طبيعي احتمالي مستمر متماثل على جانبيه عند وسطه وشكله يشبه الجرس، ومركز التوزيع الطبيعي هو وسطهμ. والشكل التالي يمثل المنحنى الطبيعي:



ويمكن أن يكون شكل الجرس أكثر تسطحا أو أكثر طولا اعتمادا على درجة تشتت قيم المتغير العشوائي عن مركز التوزيع، كما تمثل المساحة تحت المنحنى الطبيعي احتمالا وتساوي المساحة الكلية الواحدة.[3]

1- خواص التوزيع الطبيعي:

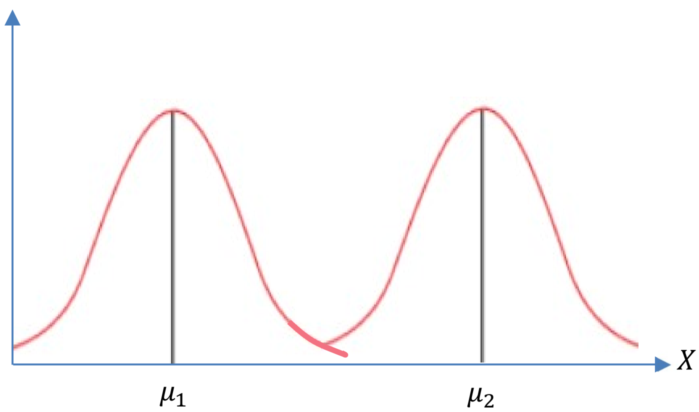
ويمكن عرض أهم خواص التوزيع الطبيعي في:

1- محمد حسين محمد رشيد، مرجع سابق، ص ص: 22/21.

2- قسم الإحصاء جامعة الملك عبد العزيز، مرجع سابق، ص 223.

3- برنارد تايلور الثالث: مقدمة في علم الإدارة (الجزء الثاني)، ترجمة سرور علي إبراهيم وآخرون، دار المريخ للنشر، الرياض، السعودية، ص، ص: 668/667.

- التوزيع الطبيعي متماثل حول العمود المقام على الوسط μ وشكله يشبه الجرس.
- للتوزيع الطبيعي قمة واحدة وبذلك له منوال واحد ينطبق على الوسط.
- يتقارب طرفا منحنى التوزيع الطبيعي من الصفر عندما تقترب من $+\infty$ و$-\infty$.
- المساحة الواقعة تحت التوزيع الطبيعي تساوي الواحد.
- الوسط الحسابي يساوي الوسيط المنوال.
- المساحة على يمين الوسط تساوي المساحة على اليسار وتساوي 0,5.
- إذا تحركت μ إلى اليمين أو اليسار ينتقل مركز التوزيع ولا يتغير شكل التوزيع، أما إذا تغيرت γ وبقيت μ نفسها فإن تشتت وتباعد المنحنى حول المركز يقل كلما صغرت γ، أما إذا تغيرت μ و γ فإن مركز التوزيع يتغير وتباعد منحناه حول المركز يتغير كذلك، والأشكال التالية توضع مختلف الحالات حيث الشكل الأول يوضح تغير الوسط الحسابي فإن منحنى التوزيع يتحرك يمينا أو يسارا ولكن شكل التوزيع لا يتغير.



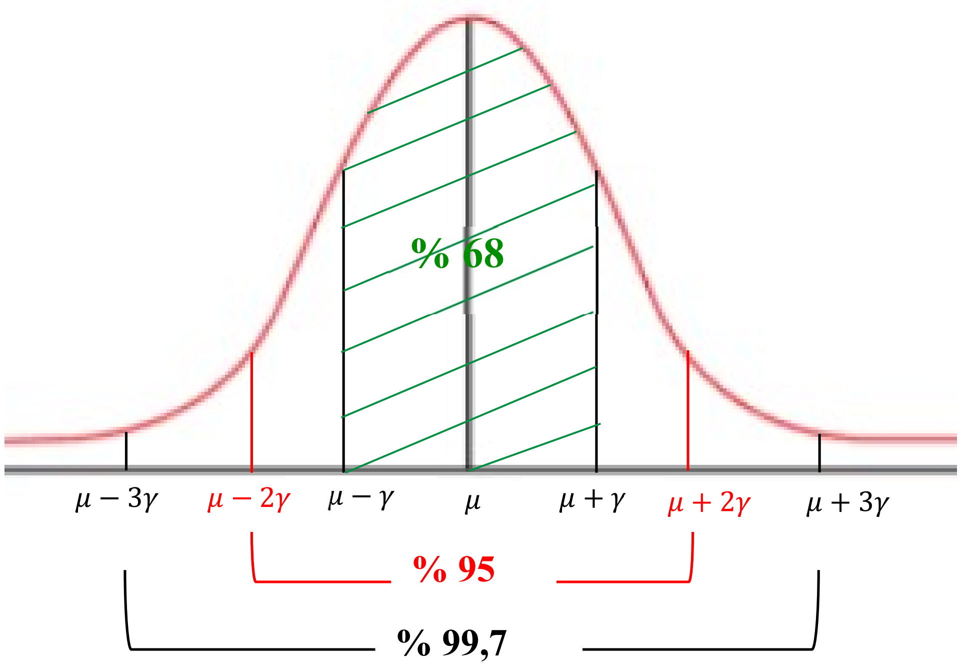
2-القانون التجريبي للتوزيع الطبيعي:

تنقسم المساحة تحت المنحنى الطبيعي حسب القانون التجريبي [Emprical Rule] إلى ثلاثة أقسام كما هو موضح في الجدول التالي:

<table>
<tr><td rowspan="1">المساحة التقريبية</td><td colspan="2">تنحصر بين $[\mu - \gamma k, \mu + \gamma k]$</td></tr>
<tr><td>تحت المنحنى</td><td>الطبيعي</td><td>الطبيعي القياسي</td></tr>
</table>

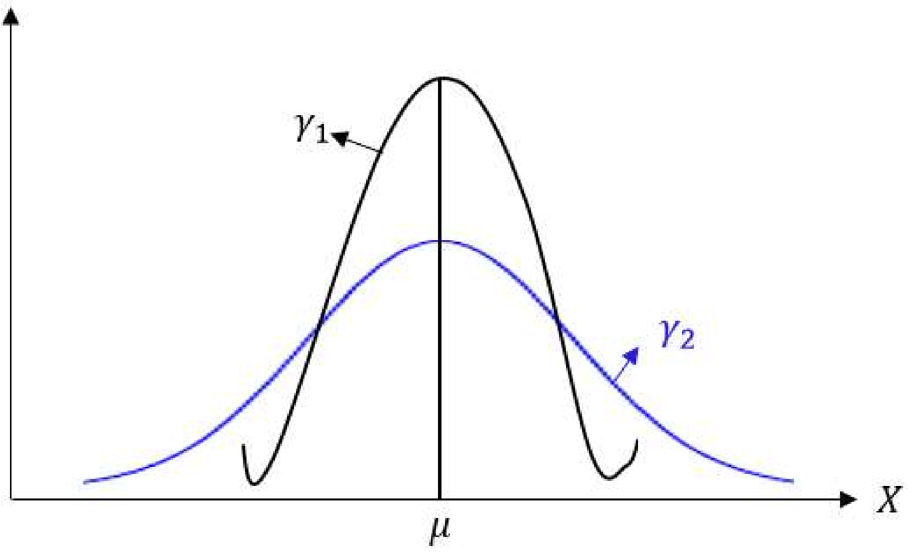
68 % من المساحة الكلية	$(\mu-\gamma,\mu+\gamma)$	$(-1,1)$
95 % من المساحة الكلية	$(\mu-2\gamma,\mu+2\gamma)$	$(-2,2)$
99.7 % من المساحة الكلية	$(\mu-3\gamma,\mu+3\gamma)$	$(-3,3)$

ويمكن توضيح هذه النسب من خلال المساحات تحت المنحنى.[1]



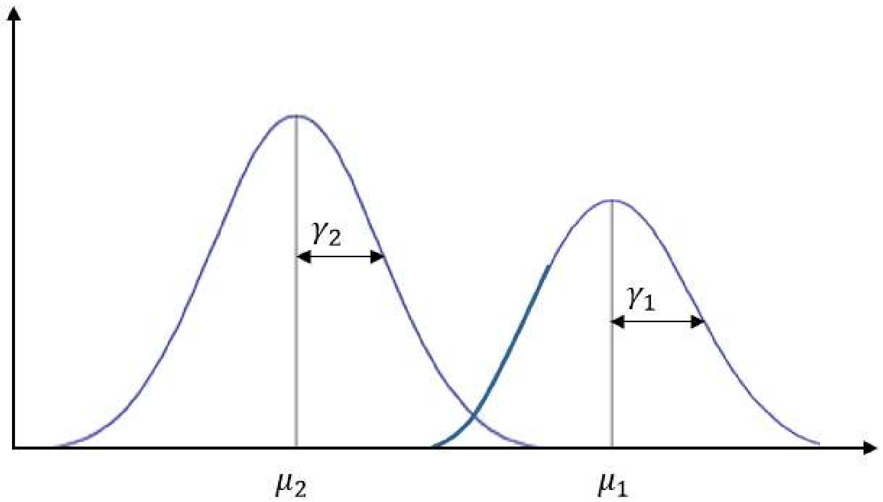
[1]– قسم الإحصاء جامعة الملك عبد العزيز، مرجع سابق، ص، ص: 226، 227.

والشكل الثاني إذا تغير الانحراف المعياري وبقي الوسط ثابتا فإن تشتت وتباعد المنحنى حول المركز يقل كلما صغرت γ



$\gamma_2 > \gamma_1$

أما الشكل الثالث يظهر لنا إذا تغيرت γ و μ فإن مركز التوزيع يتغير وتباعد منحناه حول المركز يتغير كذلك.[1]



$\gamma_2 < \gamma_1$ و $\mu_2 < \mu_1$

3- التوزيع الطبيعي المعياري:

[1]- محمد حسين محمد رشيد، مرجع سابق، ص، ص: 229/227.

يمكن القول أنه إذا كان X متغير عشوائي طبيعي متوسطه μ وانحرافه المعياري γ فإن الصيغة:

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

تحدد العلاقة بين X وZ، حيث Z هي متغير عشوائي طبيعي متوسطه الصفر وانحرافه المعياري الواحد أي أن المتغير العشوائي القياسي Z له توزيع طبيعي وحيد متوسطه μ=0 وانحرافه المعياري γ=1 هذا التوزيع يعرف بالتوزيع الطبيعي المعياري أو القياسي مثلا إذا كان X متغير عشوائي طبيعي متوسطه $\mu = 10$، $\gamma = 2$

لحساب احتمال $P(X \leq 13{,}5)$ باستخدام العلاقة نجد قيمة Z التي تناظر X

$$Z = \frac{13{,}5 - 10}{2} = 1{,}75$$

أي أصبح احتمال $P(X \leq 13.5)$

هو نفسه احتمال $P(Z \leq 17.5)$ [1]

4- المساحات في التوزيع الطبيعي المعياري:

من أجل حساب مختلف المساحات سنعتمد على جدول التوزيع الطبيعي المعياري من النوع $P(Z < a)$

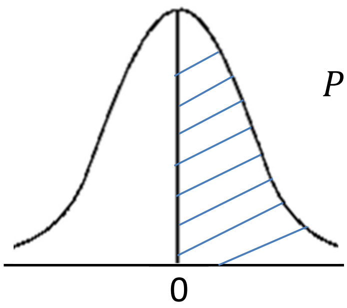
حيث سنعتمد أيضا على خاصية التناظر لأن أي قيمة سالبة أو مساحة في النصف السالب يمكن أن نناظرها للنصف الموجد وهذه أهم المساحات التي يمكن أن يجدها الطالب في مختلف التمارين حيث

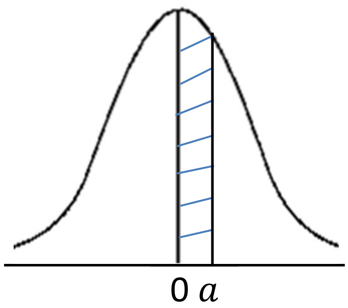
$$\begin{pmatrix} a_1 & \text{قيمة موجبة} \\ -a_2 & \text{قيمة سالبة} \end{pmatrix}:$$

---

[1]– جورج كانافوس ودون ميلر : الإحصاء للتجاريين مدخل حديث ترجمة: سلطان محمد عبد الحميد ومحمد توفيق البلقيني، دار المريخ، الرياض، السعودية، 2004، ص، ص: 220، 221.

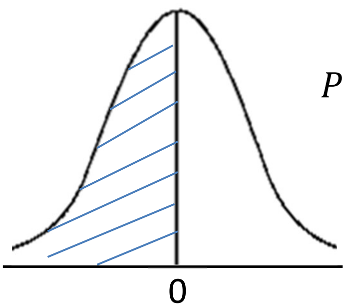
$P(Z \geq 0) = 0{,}5$

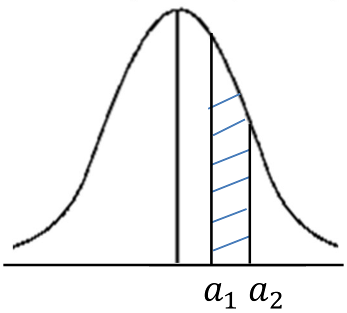




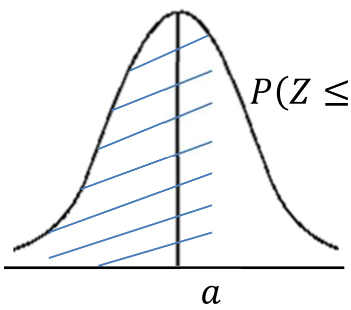
$P(0 \leq Z \leq a) = P(Z \leq a) - 0{,}5$

$P(Z \leq 0) = 0{,}5$

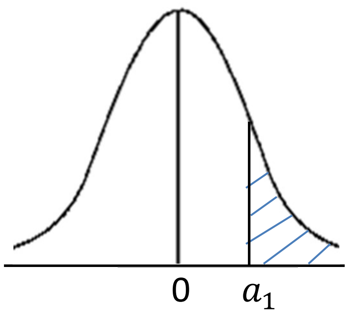




$P(a_1 \leq Z \leq a_2) = P(Z \leq a_2) - P(Z \leq a_1)$

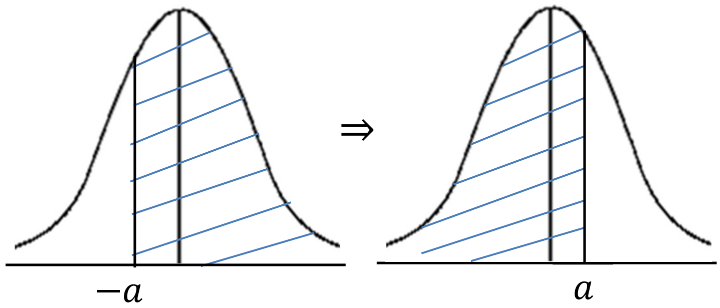


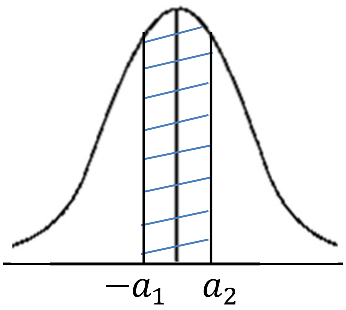
$P(Z \leq a) =$ الجدول



$P(Z \geq a_1) = 1 - P(Z \leq a_1)$

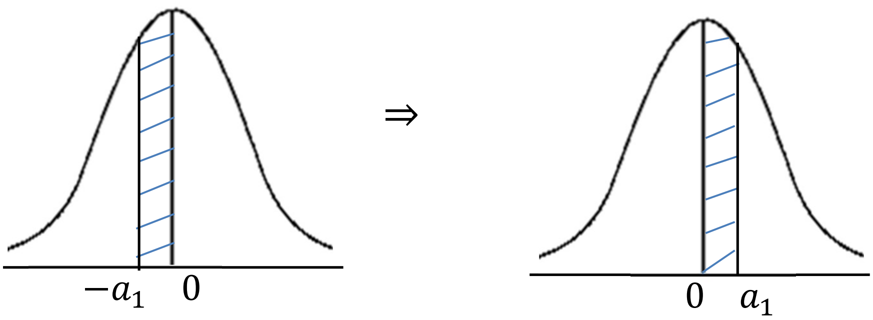
$P(Z \geq -a) = P(Z \leq a)$





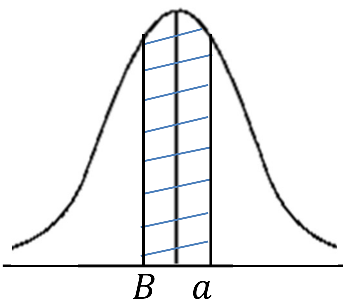
$$P(-a_1 \leq Z \leq a_2) = [P((Z \leq a_2) - 0{,}5) + P((Z \leq a_1) - 0{,}5)]$$

الموجبة



$$P(-a_1 \leq Z \leq 0) = P(0 \leq Z \leq a_1) = P(Z \leq a_1) - 0{,}5$$

عبارة عن مجموع مساحتين مساحة في الجهة الموجبة ومساحة في الجهة السالبة نناظرها للجهة الموجبة



$$P(0 \leq Z \leq a) + P(0 \leq Z \leq B)$$

وبعد عرض مختلف المساحات سنقدم المثال التالي:

مثال: إذا كان دخل مجموعة أسر في ولاية بسكرة يتبع التوزيع الطبيعي بمتوسط $(\mu = 100)$ وانحراف معياري $(\gamma = 10)$ أوجد:

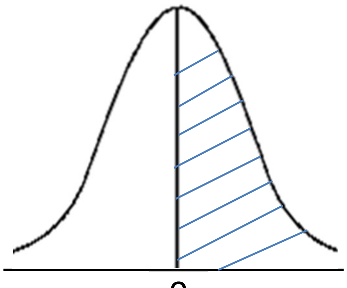
1) احتمال الحصول على دخل أسرة أكبر من 100؟
2) الحصول على دخل أسرة أقل من 80؟
3) احتمال الحصول على دخل أسرة على الأكثر 115؟
4) احتمال الحصول على دخل أسرة على الأقل 104؟
5) احتمال الحصول على دخل أسرة أكبر من 85؟
6) احتمال الحصول على دخل أسرة محصور بين 84 و95؟
7) احتمال الحصول على دخل أسرة أكبر من 120؟
8) احتمال الحصول على دخل أسرة بين 85 و120؟

الحل:

احتمال P(X≥ 100)

$$Z = \frac{X-\mu}{\gamma} = \frac{100-100}{10} = 0$$

$$P(X \geq 100) \Rightarrow P(Z \geq 0) = 0.5$$



2) احتمال $P(X \leq 80) \ldots\ldots (1)$

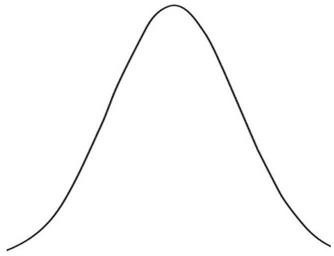
$$Z = \frac{80-100}{10} = -2$$

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

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

$$= 1 - 0.9772$$

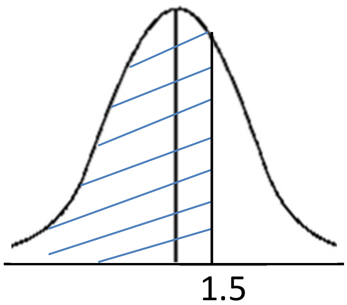
$$= 0.0228$$



3) احتمال $P(X \leq 115) \dots\dots (1)$

$$Z = \frac{115 - 100}{10} = 1.5$$

$$(1) \Rightarrow P(Z \leq 1.5) = 0.9332$$



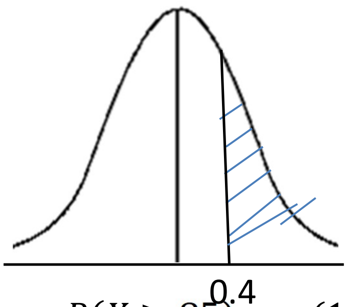
4) حساب احتمال $P(X \geq 104) \dots\dots (1)$

$$Z = \frac{104 - 100}{10} = 0.4$$

$$(1) \Rightarrow P(Z \geq 0.4) = 1 - P(Z \leq 0.4)$$

$$= 1 - 0.6554$$

$$= 0.3446$$

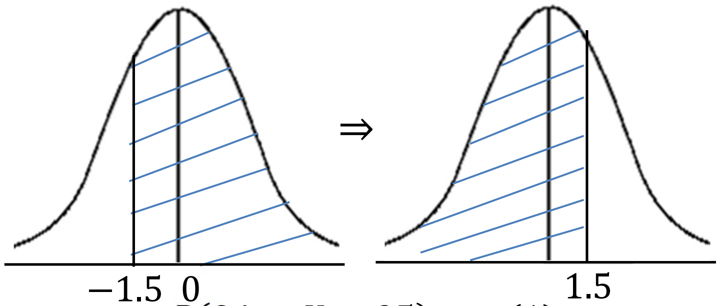


5) احتمال $P(X \geq 85) \dots\dots (1)$

$$Z = \frac{85 - 100}{10} = -1.5$$

$$(1) \Longrightarrow P(Z \geq -1.5) = P(Z \leq 1.5)$$

$$= 0.9332$$



6) حساب احتمال $P(84 \leq X \leq 95) \dots \dots (1)$

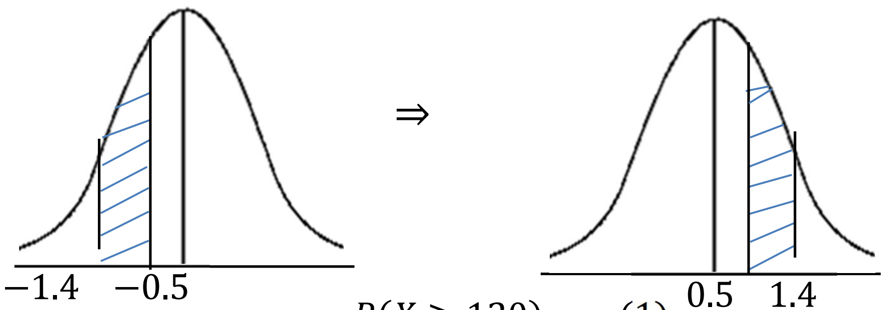
$$Z_1 = \frac{84 - 100}{10} = -1.4$$

$$Z = \frac{95 - 100}{10} = -0.5$$

$$(1) \Longrightarrow P(-1.4 \leq Z \leq -0.5)$$

$$= P(Z \leq 1.4) - P(Z \leq 0.5)$$

$$= 0.9192 - 0.6915 = 0.2277$$



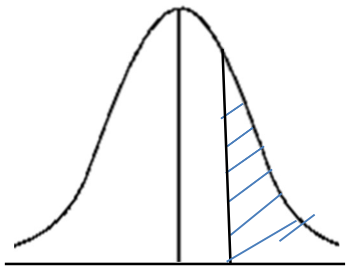
7) حساب $P(X \geq 120) \dots \dots (1)$

$$Z = \frac{120 - 100}{10} = 2$$

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

$$= 1 - 0.9772$$

$$= 0.0228$$



8) حساب احتمال $P(85 \leq X \leq 120) \dots \dots (1)$

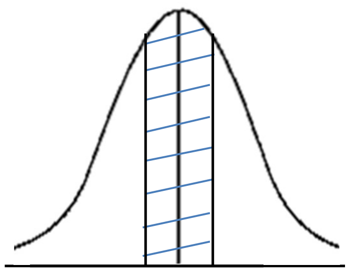
$$Z = \frac{85 - 100}{10} = -1.5$$

$$Z = \frac{120 - 100}{10} = 2$$

$$(1) \Longrightarrow P(-1.5 \leq Z \leq 2)$$

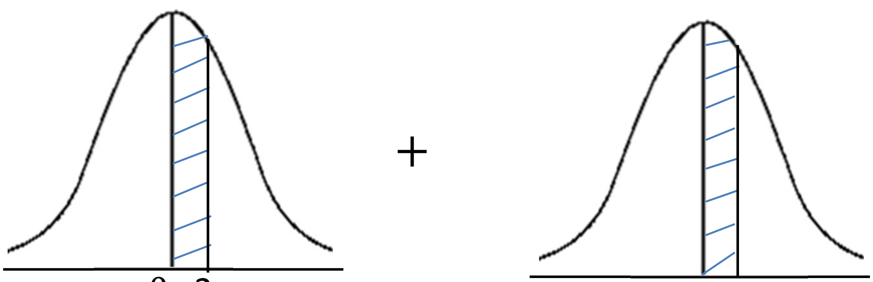
$$P(0 \leq Z \leq 2) + P(0 \leq Z \leq 1.5$$
$$= (0.9772 - 0.5) + (0.9332 - 0.5)$$

$$= 0.9304$$



هذا الشكل نحسب المساحة الموجبة ثم نناظر المساحة التي في الجهة السالبة إلى الجهة الموجبة

أي يصبح الشكل كما يلي:



ثالثا: توزيع t (student)

يعود الفضل في ظهور هذا الاختبار للعالم william sealy gosset، والملقب بـ student ويستعمل هذا الاختبار لحساب دلالة الفروق بين المتوسطات المرتبطة وغير المرتبطة للعينات المتساوية وغير المتساوية وتوجد مجموعة من النماذج لاختبار "T" ولكل نموذج، مجال استخدامه ومهما كان النموذج، لابد من فحص توفر الشروط التالية قبل تطبيقه:

- أن يكون توزيع العينتين اعتداليا.
- أن يكون حجم العينتين متقاربا.
- ألا يقل حجم العينتين عن 30 فردا.
- أن تكون العينتان متجانستان.[1]

منحنى التوزيع t يشبه شكل الجرس وهو أحادي المنوال حيث له قمة تقابل $(t=0)$ وهو متماثل حول العمود المقام على $(t=0)$ ويشبه شكله شكل التوزيع الطبيعي المعياري إلا أنه أكثر انخفاضا منه بالإضافة إلى تقارب طرفيه من الصفر عندما $t\to-\infty$ و $t\to+\infty$ أبطأ في تقارب طرفي التوزيع الطبيعي المعياري وعندما يزداد عدد درجات الحرية فإن توزيع t يقترب من التوزيع الطبيعي المعياري.

وتحسب الاحتمالات تحت توزيع t بحساب المساحات تحت منحنى ذلك التوزيع بعد معرفة درجات الحرية له ويوجد جداول خاصة لهذه المساحات حيث تسجل درجات الحرية في العمود الأيسر وعلى الخط الأفقي تسجل المساحات، أما الأعداد داخل الجدول فتمثل قيم t أما إذا أردنا إيجاد المساحة الواقعة إلى يسار قيمة t تحت توزيع t ذي درجات حرية معينة فإننا ننظر في عمود درجات الحرية ونجد أقرب عدد للقيمة المعطاة، ثم نجد المساحة إلى يسار تلك القيمة من الخط الأفقي المقابل لذلك العدد.[2]

1- <u>حالات استخدام اختبار T</u>: [3]

---

[1]– محمد بوعلاق: الموجه في الإحصاء الوصفي والاستدلالي في العلوم النفسية والتربوية والاجتماعية، الطبعة 2، دار الأمل للطباعة والنشر والتوزيع، 2012، ص 141.

[2]– محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء، ديوان المطبوعات الجامعية، الجزائر، 1984، ص 111.

[3]– محمد بوعلاق، نفس المرجع، ص 141-150.

أ- اختبار T لعينة واحدة (أو لعينتين مرتبطتين أو متشابهتين)

هو أحد استخدامات اختبار T والغرض منه هو اختبار فرضية حول متوسطي عينة واحدة، ويستخدم:

- عندما تكون للباحث مجموعة من الأفراد يلاحظها في وضعيتين مختلفتين.
- عندما تكون لدى الباحث عينة واحدة يطبق عليها اختبارا قبليا، واختبارا بعديا، وغالبا ما تحدث هذه الوضعية في المنهج التجريبي.
- عندما تكون لدى الباحث عينتان مختلفتان، ولكنهما متشابهتين في بعض الخصائص، غير أنه في مثل هذه الحالة، يجب على الباحث أن يتأكد من أن هذه الخصائص ترتبط ارتباطا وثيقا بالمتغير التابع.

وصيغة هذا النموذج: $T_0 = \frac{\bar{D}}{s\bar{D}}$

$\bar{D}$ متوسط الفرق بين الدرجات في الوضعية الأولى ودرجاتها في الوضعية الثانية.

S الانحراف المعياري للعينة ويحسب بالمعادلة التالية: $\bar{D} = \frac{\sum D}{n}$

D الفرق بين الدرجات

n عدد أفراد العينة

$$S\bar{D} = \frac{SD}{\sqrt{n}}$$

$$SD = \sqrt{\frac{n\sum D^2 - (SD)^2}{n(n-1)}}$$

تقدر قيمة درجة الحرية في حالة استخدام T لعينة واحدة كما يلي: dF=n−1

dF درجة الحرية

مثال: أراد باحث تجريب فعالية دواء يعالج الاكتئاب فاختيار عينة تتكون من 10 مرضى، وقاس درجة الاكتئاب قبل تجريب الدواء عليها، ثم قاس درجة الاكتئاب بعد إعطائهم الدواء لمدة معينة، وافترض الباحث ما يلي:

لا يوجد اختلاف بين درجة الاكتئاب قبل تناول وبعد تناول الدواء، وجاءت نتائج التجربة كما يلي:

القياس القبلي	القياس البعدي	الفروق	D	مربع انحراف الفروق SD
7	10	3	1	1
3	5	2	0	0
7	6	1−	3−	9
5	8	2	0	0
8	10	2	0	0
4	6	2	0	0
5	7	2	0	0
2	8	6	4	16
3	6	3	1	1
6	5	1−	3−	9
مج= 50 المتوسط= 5	مج= 70 المتوسط= 7	مج= 20 المتوسط= 2	0	36

$$T_0 = \frac{\bar{D}}{\sqrt{\frac{\sum SD}{n(n-1)}}}$$

$$T_0 = \frac{2}{\sqrt{\frac{36}{10(10-1)}}} = 3.16$$

درجة الحرية dF=n−1=10−1=9

T المجدولة عند مستوى 0.05 بدرجة حرية 9 هي 2.26 أي أن T المحسوبة 3.16 أكبر من T الجدولية وبالتالي فإن الفرق دال إحصائيا بين القياس القبلي والقياس البعدي، فالدواء المجرب ذو فعالية.

ب− اختبار T لعينتين مستقلتين وغير متساويتين في الحجم $n_1 \# n_2$

يخضع هذا النموذج لنفس الشروط التي يتطلبها أي اختبار معلمي وهو يقوم على مقارنة متوسطي عينتين اختيرتا بطريقة عشوائية ولهما تباينين متساويين وتكون العينتان مستقلتان عندما لا يكون الفرد الواحد موجودا في كلتا العينتين، مثل اختبار علامات فوجين من الطلبة في امتحان واحد، ويقوم هذا النموذج على التحقق من الافتراضات التالية:

$$H_0 \Longrightarrow \bar{X}_1 - \bar{X}_2 = 0$$

$$H_2 \Rightarrow \bar{X}_1 - \bar{X}_2 \neq 0$$

يحسب اختبار T لهذا النموذج بالمعادلة التالية:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

حيث أن:

$\bar{X}_1$ متوسط العينة الأولى.

$\bar{X}_2$ متوسط العينة الثانية.

S_1^2 تباين العينة الأولى.

S_2^2 تباين العينة الثانية.

n_1 عدد أفراد العينة الأولى.

n_2 عدد أفراد العينة الثانية.

$dF = n_1 + n_2 - 2$ درجة الحرية للمقارنة بين عينتين مستقلتين.

مثال: افترض باحث ما يلي: "لا يوجد اختلاف بين الذكور والإناث في تحصيل الإحصاء"، وبعد إجراء العمليات الحسابية تحصل على ما يلي:

الإناث	الذكور
$N_2=4$	$N_1=5$
$\bar{X}_2 = 10.7667$	$\bar{X}_1 = 10.9625$
$S_2^2 = 3.073$	$S_1^2 = 3.239$

نلاحظ أن حجم $n_1 \# n_2$ إذا لا بد من حساب التجانس ولاختبار التجانس نستخدم اختبار فيشر (F) وإذا كانت قيمة F المحسوبة أقل من F المجدولة نقول أن العينتين متجانستين.

1. نحسب النسبة (F) $\frac{s_1^2}{s_2^2} = \frac{3.239}{3.073} = 1.05$

2. نحسب درجة الحرية للتباين الكبير S_1^2: n−1=5−1=4 ودرجة حرية التباين الأصغر S_2^2: n−1=4−1=3

3. من خلال جدول (F) عند درجة حرية التباين الكبير S_1^2 وهي 4 ودرجة حرية التباين الصغير S_2^2 هي 3 نجد قيمة (F)=9.12 عند مستوى دلالة 0.05، وبما أن F المجدولة أكبر من F المحسوبة (1.05) فإننا نستنتج أن العينتين متجانستين، ثم نطبق اختبار T لعينتين مستقلتين ومتجانستين وغير متساويتين في الحجم.

$$T = \frac{10.9625 - 10.7667}{\sqrt{\frac{(5-1)3.239 + (4-1)3.073}{5+4-2}\left[\frac{1}{5} + \frac{1}{4}\right]}} = 0.16$$

درجة الحرية $n_1+n_2-2=5+4-2=7$.

من خلال جدول ستودنت T نجد القيمة المقابلة لدرجة الحرية 7 عند مستوى دلالة 0.05 لاختبار الطرفين (لأن الباحث لم يحدد اتجاه الاختلاف في فرضتيه) = 2.36 وبما أن T المحسوبة 0.16 أقل من القيمة المجدولة فإننا نقبل الفرضية الصفرية التي تقول: "لا يوجد اختلاف بين الذكور والإناث في تحصيل الإحصاء"، وهو ما يفيد أن الفروق في تحصيل وحدة الإحصاء بين الذكور والإناث لا يرجع إلى عامل الجنس، فالجنس لا يؤثر على تحصيل الإحصاء بنسبة خطأ في النتيجة تساوي 5 % ودرجة دقة 95 %.

ت− اختبار T لعينتين مستقلتين وغير متجانستين $(S_1^2 \neq S_2^2)$ وغير متساويتين في الحجم (n_1#n_2)

للتأكد من فرضية تدعي عدم وجود فروق دالة بين عينتين مستقلتين وغير متساوتين في الحجم وغير متجانستين، ينبغي تطبيق المعادلة الآتية:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

الخطوات:

1− نحسب قيمة T بالمعادلة السابقة.

2- نحدد مستوى الدلالة 0.05.

3- نحسب درجات حرية العينة الأولى ودرجة حرية العينة الثانية.

4- نحسب قيمة T_1 للعينة الأولى المقابلة لدرجة الحرية الأولى.

5- نحسب قيمة T_2 للعينة الثانية المقابلة لدرجة الحرية الثانية.

6- نحسب قيمة 'T الفروق باستخدام كل من T_1 و T_2 باستخدام المعادلة التالية:

$$'T = \frac{T_1[\frac{S_1^2}{n_1}] + T_2[\frac{S_2^2}{n_2}]}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

نقارن بين قيمتي T و 'T فإذا كانت T أكبر أو يساوي 'T عند مستوى الدلالة المحدد، دل ذلك على وجود فروق دالة إحصائيا بين متوسطي درجات المجموعتين، أما إذا كانت T أصغر من 'T دل ذلك على عدم وجود فروق دالة بين متوسطي الدرجات.

مثال: في دراسة للعدوانية عند مجموعتين من الريفيين والمدن وجدت باحثة البيانات التالية:

الريف	المدن
N_1=10	N_2=20
=206	=16
=28.42	=6.72

نلاحظ أن قيمة T المحسوبة 2.58 أكبر من 'T (2.24) عند مستوى الدلالة المحدد وهو 0.05 وبالتالي يوجد فرق دال احصائيا بين متوسطي درجات العدوانية لصالح سكان الريف باعتبار أن متوسطهم هو الأكبر.

و- اختبار T لعينتين مستقلتين ومتساويتين في الحجم (n_1=n_2)

في حالة وجود عينتين متساويتين في الحجم وغير مرتبطتين ينبغي تطبيق المعادلة التالية:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2 + s_2^2}{N - 1}}}$$

N : n_1=n_2

dF درجة الحرية

مثال: في اختبار لقياس الذكاء المتعدد على مجموعة تتكون كل واحدة منهما من 31 طالب وطالبة، حيث انطلق باحث من فرضية تقول: "لا يوجد فرق بين الطلبة والطالبات فيما يخص الذكاء المتعدد"

مجموعة الطلبة	مجموعة الطالبات
N_2=15	N_1=15
$\overline{X_2} = 15.81$	$\overline{X_1} = 23.63$
$S_2 = 2.62$	$S_1 = 3.62$

الحل:

1- نحسب قيمة F لفحص التجانس $\frac{28.42}{6.72} = 4.23$

2- درجة حرية التباين الكبير $n_1-1=9$، ودرجة حرية التباين الصغير $n_2-1=19$

3- من خلال جدول (F) عند درجة حرية التباين الكبير (9) ودرجة حرية التباين الصغير (19) نجد أن القيمة الجدولية F=2.42 عند مستوى الدلالة 0.05 وهذا يدل على عدم تجانس العينتين مما يدعو إلى تطبيق المعادلة الخاصة بهذا النموذج:

$$T = \frac{20.6 - 16}{\sqrt{\frac{28.42}{10} + \frac{6.72}{20}}}$$

$$T = 2.85$$

إذا حددنا مستوى الدلالة 0.05 فإن T_1 للمجموعة الأولى المقابلة لدرجات حرية 9 عند مستوى 0.05 يساوي المقابلة لدرجات حرية 19 عند مستوى 0.05 يساوي 2.093

$$'T = \frac{T_1[\frac{S_1^2}{n_1}] + T_2[\frac{S_2^2}{n_2}]}{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$'T = \frac{2.262\left[\frac{28.42}{10}\right] + 2.093\left[\frac{6.72}{20}\right]}{\frac{28.42}{10} + \frac{6.72}{20}} = 2.24$$

الحل:

$$T = \frac{23.63 - 15.81}{\sqrt{\frac{(3.62)^2 + (2.62)^2}{15-1}}}$$

$$T = 6.55$$

بما أن درجة الحرية 29=1−(15)2، وقيمتها المقابلة في جدول T عند مستوى الدلالة 0.05 وعند اختبار الطرفين تساوي 2.045، وبالتالي فالفرضية الصفرية مرفوضة لأن القيمة المحسوبة أكبر من القيمة المجدولة والفرق لصالح مجموعة الإناث لأنها تحصلت على متوسط أكبر من متوسط الذكور.

أمثلة حول إيجاد قيمة T أو المساحة:

1) أوجد قيمة t بدرجة الحرية 10 ومساحة 0.95
2) أوجد قيمة t بدرجة الحرية 21 ومساحة 0.99
3) أوجد المساحة الواقعة إلى يسار t=2.13 بدرجة الحرية 15

الحل:

1- T=1.812

2- T=2.518

3- بالنظر لدرجة الحرية 15 والسير أفقيا نجد أقرب عدد للقيمة المعطاة 2.131 وبالصعود عموديا نلاحظ أن المساحة 0.975. [1]

رابعا: توزيع X^2 (كاي تربيع)

يعتبر هذا التوزيع من التوزيعات الاحتمالية المتصلة المهمة وله تطبيقات عديدة في الاستدلال الاحصائي وخصوصا على البيانات الوصفية وأيضا يستخدم في تقدير فترة الثقة واختبارات الفروض لتباين المجتمع الاحصائي γ^2.

يرتبط توزيع X^2 بدرجة كبيرة بعدد درجات الحرية وهو توزيع متصل ذا قيمة واحدة مثل التوزيع الطبيعي (Z) وتوزيع ستودنت (T) لكنهما يختلفان في كونهما متماثلان حول وسطهما الحسابي (μ) ويعتبر

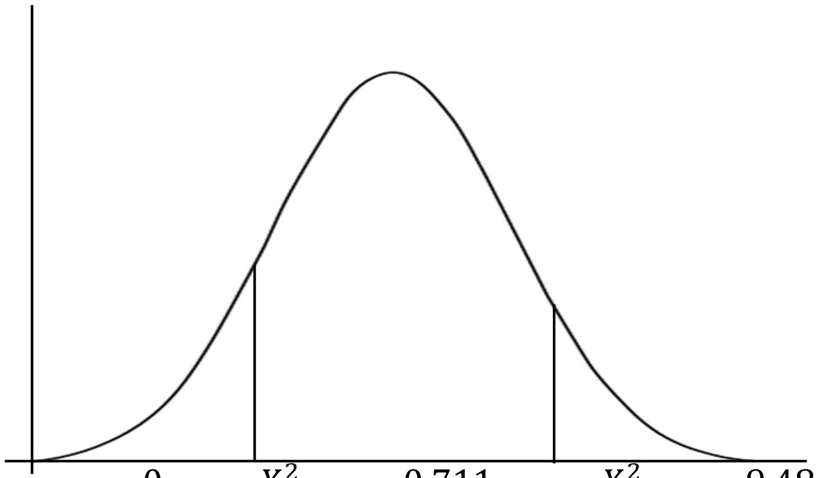
[1]– محمد صبيحي أبو صالح، عدنان محمد عوض، مرجع سابق، ص، ص: 112/111.

توزيع (X^2) توزيعا ملتويا جهة اليمين وخصوصا عندما تكون درجات الحرية صغيرة، وكلما زادت درجات الحرية كلما قل الالتواء واقترب التوزيع إلى التماثل واقترب أكثر للتوزيع الطبيعي أي أن توزيع X^2 يؤول إلى التوزيع الطبيعي عندما تؤول درجة الحرية إلى مالا نهاية.

وقد تم إعداد جداول جاهزة تعطي أهم احتمالات توزيعX^2، والجدول التالي يمثل جزء من جدولX^2:

X^2 / V	$X^2.995$	...	$X^2.95$	...	$X^2.05$	...	$X^2.01$	$X^2.005$
4	0.207		0.711				13.227	14.860

وقيمX^2 دائما أكبر أو تساوي الصفر، والشكل التالي يمثل قيمتين من قيمX^2 بدرجات حرية 4:



من خلال الشكل نجد المساحة على يمين$X^2_{0.05,4}$ (9.488) تساوي 0.05 بينما نجد المساحة على يسار $X^2_{0.95,4}$(0.711) تساوي 0.95.

ومثلا$X^2_{0.05,4}$ تمثل قيمةX^2 المعيارية بدرجات حرية تساوي 4 والتي تساوي المساحة على يمينها 0.05 ومن خلال الجدول نجد: $X^2_{0.05,4} = 9.488$.

والنظريات التالية مهمة في استخدام X^2.

- النظرية الأولى مهمة في اختبار الفروض حول تباين المجتمع γ^2:

إذا كانS^2 يمثل التباين لعينة عشوائية حجمها n مسحوبة من مجتمع طبيعي تباينه γ^2 فإن:

$$X^2 = \frac{(n-1)S2}{\delta 2} \sim X^2_{n-1}$$

أي أن الإحصائيةX^2 تتبع توزيع γ^2بدرجة حرية (n−1).

- النظرية الثانية (مهمة في الاختبار المطابقة، جودة التوفيق والاستقلال والتجانس):
- إذا كان حجم العينة n كبيرا فإن:

$$X^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{ei} \sim X^2_v$$

o_iالتكرارات المشاهدة الفعلية

e_iالتكرارات المتوقعة

k عدد الخلايا أو أزواج التكرارات المشاهدة والمتوقعة

v درجة الحرية.[1]

خامسا: توزيع F

يعتبر من التوزيعات الاحتمالية، الهامة التي تستعمل في اختبار الفرضيات.

إذا كان توزيع الكثافة الاحتمالي للمتغير العشوائي F معطى بالمعادلة:

$$f(F) = \frac{cF^{(v_1-2)/2}}{(v_2 + v_1F)^{(v_1+v_2)/2}}\ ,F > 0$$

هذا التوزيع يسمى F ويعبر عنه بالرمز $F(v_2, v_1)$

حيثv_2, v_1 هي درجات الحرية وC ثابت يعتمد علىv_1وv_2 ويعين بحيث تصبح المساحة تحت منحنى التوزيع تساوي 1، v_2 درجة حرية المقام و v_1 درجة حرية البسط وكلما ازدادت درجات الحريةv_2, v_1 يقترب توزيع F من التوزيع الطبيعي.

[1]– أحمد عودة بن عبد المجيد عودة، منصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي، الجزء (2) الإحصاء الاستدلالي، مكتبة الفلاح للنشر والتوزيع، الطبعة 3، 2014، ص ص: 215–219.

مثال: $F[\lambda; v1; v2] \leftarrow F[0.975\ ;9.7]=3.68$

$F[0.025\ ;11.10]=0.283$

- إيجاد المساحة إلى يسار F=3 و $7v_1=$، $20v_2=$

نلاحظ أن أقرب قيمة في الجدول لـ3 هي 3.01 فنجد $F[0.975;\ 7.20]=3.01$ أي المساحة 0.975.

قراءة جدول F في حالة بعض القيم غير الموضوعة

هناك بعض القيم غير موضوعة في الجداول مثل$F[0.05; v_1, v_2]$ أو $[0.01; v_1, v_2]$ ولإيجاد هذه القيم نستخدم القاعدة: $F[\lambda; v_1, v_2] = \frac{1}{F[1-\lambda; v_2, v_1]}$

نلاحظ أن المقام هو القيمة المتممة للمساحة المطلوبة مع تبديل درجات الحرية فيما بينها.

مثال:[1]

$$F[0.05; 10,7] = \frac{1}{F[0.95;7,10]} = \frac{1}{3.14} = 0.3181)$$

$$F[0.01; 5,10] = \frac{1}{F[0.99;10,5]} = \frac{1}{10.1} = 0.0992)$$

سادسا: توزيع ذو الحدينThe Binomial Distribution

في هذا الجزء ندرس مجموعة التجارب التي تقابل الشروط التالية:

1. التجربة عموما يمكن أن توصف بأنها متابعة من n من التجارب المتماثلة التي تسمى محاولات.
2. نتيجتان محتملتان لكل محاولة، نشير إلى النتيجة الواحدة كنجاح والأخرى كفشل.
3. احتمالات النتيجتين لا تتغيران من محاولة واحدة إلى الأخرى.
4. إن المحاولات مستقلة (بمعنى أن نتيجة المحاولة الواحدة لا تؤثر على نتيجة أي محاولة أخرى).

[1]– محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام SPSS، دار المسيرة للنشر والتوزيع، الطبعة 6، عمان–الأردن، 2012، ص، ص: 169/167.

التجارب التي تحقق شروط 2، 3، 4 يقال أنها تتولد بعملية Bernoulli بالإضافة إلى ذلك، إذا تحقق شرط 1 (هناك n من المحاولات المتماثلة)، يكون عندنا تجربة ذو حدين. من الأهمية أن المتغير العشوائي المتقطع المناظر تجربة ذو حدين يكون عدد النتائج الناجحة في n من المحاولات. إذا افترضنا أن x يشير إلى قيمة هذا المتغير العشوائي، إذن x يمكن أن يأخذ قيمة من 0، 1، 2، 3، ....، n معتمدا على عدد حالات النجاح في n من المحاولات. توزيع الاحتمال المرتبط بهذا المتغير العشوائي يسمى توزيع احتمال ذو الحدين.

في الحالات التي يكون فيها التوزيع ذو الحدين قابلا للتطبيق، فإن الصيغة الرياضية لحساب احتمال أي قيمة للمتغير العشوائي تكون دالة احتمال ذو الحدين:

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \qquad x = 0,1,\dots\dots,n$$

حيث:

N= عدد الحالات.

P= احتمال حالات النجاح في المحاولة واحدة.

X= عدد حالات النجاح في n من المحاولات.

F(x)= احتمال x من حالات النجاح في n من المحاولات.

التعبير n! المشار إليه كمضروب n يعرف كما يلي:

كمثال؛ 4! =(4) (3) (2) (1)=24، أيضا من التعريف، الحالة الخاصة من مضروب صفر تكون 0! =1.

مثال محل ملابس Nastie The Nasike Clothing Store Problème

لكي نوضح توزيع ذو الحدين، دعنا ندرس تجربة العملاء الذين يدخلون محل ملابس Nastke . لنجعل المشكلة صغيرة نسبيا، نحدد التجربة بالعملاء الثلاثة القادمين. إذا استند مدير المحل على الخبرة، يقدر بأن احتمال عميل واحد يقوم بعملية الشراء 0.30. ما الاحتمال أن بالضبط اثنان من العملاء الثلاثة القادمين يقومون بعملية الشراء؟

نرغب أولا أن تعرض بأن العملاء الثلاثة يدخلون محل الملابس ونقرر ما إذا كان اتخاذ قرار الشراء يمكن أن يعتبر تجربة ذات حدين، لتأكيد المتطلبات الأربعة لتجربة ذات الحدين، نلاحظ التالي:

1. يمكن أن توصف التجربة كمتتابعة من ثلاث محاولات متماثلة. محاولة واحدة لكل عميل من العملاء الثلاثة الذين سيدخلون المحل.

2. نتيجتان يتخذ العميل قرار الشراء (نجاح) أو لا يتخذ العميل قرار الشراء (فشل) – محتملتان لكل محاولة.
3. احتمالات الشراء (0.30) وعدم الشراء (0.70) النتائج تفترض أن تكون نفسها مع كل العملاء.
4. قرار شراء كل عميل مستقل عن قرار شراء العملاء الآخرين.

وبالتالي، إذا عرفنا المتغير العشوائي x كعدد العملاء الذين يتخذون قرار الشراء (مثال، عدد النجاح في المحاولات الثلاث)، نتحقق من متطلبات توزيع ذو الحدين.

مع n=3 محاولات واحتمال شراء p=0.30 لكل عميل، نستخدم المعادلة السابقة لكي نحسب احتمال أن عميلين يقومان بالشراء. هذا الاحتمال يشار إليه f(2) ويكون:

$$f(2) = \frac{3!}{2!\,1!}(0.30)^2(0.70)^1$$
$$= \frac{3\times 2\times 1}{2\times 1\times 1}(0.30)^2(0.70)^1 = 0.189$$

بالمثل، احتمال لا يوجد عملاء يقومون بالشراء، يرمز إليه f(0) ويكون:

$$f(0) = \frac{3!}{0!\,1!}(0.30)^0(0.70)^3$$
$$= \frac{3\times 2\times 1}{1\times 3\times 2\times 1}(0.30)^0(0.70)^3 = 0.343$$

وأيضا، يمكن أن تستخدم المعادلة لكي تعرض احتمال أن واحد يقوم بالشراء واحتمال أن ثلاثة يقومون بالشراء مشتريات يكون f(1)=0.441 وf(3)=0.027 على التوالي الجدول والشكل يلخص احتمالات توزيع ذو الحدين لمثال محل ملابس Nastke.

إذا كنا ندرس أي اختلاف عن مثال Nastke مثل 10 عملاء بدلا من 3 عملاء يدخلون المحل. كمثال، الاحتمال أن 4 من 10 عملاء يقومون بالشراء يكون:

$$f(4) = \frac{10!}{4!\,6!}(0.30)^0(0.70)^6 = 0.2001$$

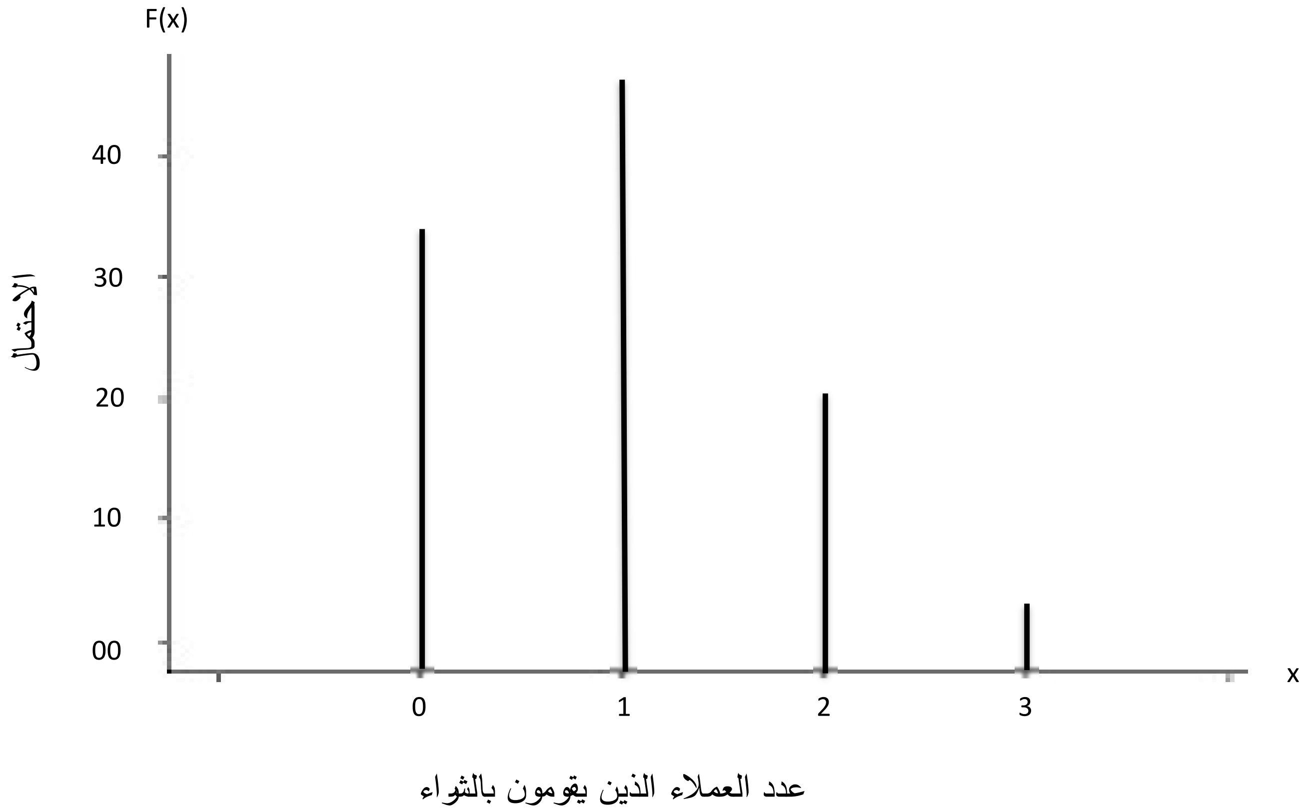
في هذه التجربة ذات الحدين n=10، p=0.30، x=4

جدول: التوزيع الاحتمالي لعدد العملاء الذين يقومون بالشراء

F(x)	X
0.343	0
0.441	1
0.189	2

0.027	3
1.000	الإجمالي

شكل التوزيع الاحتمالي لمشكلة محل ملابس NASTKE



القيمة المتوقعة والتباين لتوزيع ذو الحدين The Expected Value and Variance for the Binomial Distribution

تحسب القيمة المتوقعة أو العدد المتوقع من العملاء الذين يقومون بالشراء

$$\mu = \sum xf(x) = 0(0.343) + 1(0.441) + 2(0.189) + 3(0.027) = 0.9$$

نلاحظ أننا يمكن أن نحصل على هذه القيمة المتوقعة نفسها ببساطة بضرب n (عدد المحاولات) في p (احتمال النجاح في أي محاولة واحدة).

$$Np=3(0.30)=0.9$$

كحالة خاصة لتوزيع احتمال ذو حدين، القيمة المتوقعة للمتغير العشوائي تعطn كالتالي:

$$\mu = np$$

افترض أنه خلال الشهر القادم محل ملابس Nastke يتوقع 1000 عميل يمكن أن يدخلوا المخزن. ما العدد المتوقع للعملاء الذين سيقومون بالشراء؟ باستخدام المعادلة تكون الإجابة $\mu = np = (1000)(0.3) = 300$. لزيادة العدد المتوقع للمبيعات، Nastke يجب أن يغري عملاءه أكثر أن يدخلوا المحل و/ أو يزيد احتمال أن أي عميل فردي سيقوم بالشراء بعد الدخول.

كحالة خاصة من توزيع ذو الحدين التباين للمتغير العشوائي يكون:

$$\sigma^2 = np(1-p)$$

لمثال محل الملابس Nastke لثلاثة عملاء. التباين والانحراف المعياري لعدد العملاء الذين يقومون بالشراء يكونا:[1]

$$\sigma^2 = np(1-p) = 3(0.3)(0.7) = 0.63$$
$$\sigma = \sqrt{0.63} = 0.79$$

سابعا: نظرية النهاية المركزية

إذا كان لدينا مجتمع غير محدود مفرداته (X) تتبع توزيعا احتماليا متوسطه μ وانحرافه المعياري γ سحبنا من هذا المجتمع عينات عشوائية حجم كل منها (n) وكانت (n) كبيرة الحجم (n≥ 30) فإن الوسط الحسابي لهذه العينات X يتبع تقريبا توزيعا طبيعيا له الخصائص التالية:

$$\mu_{\bar{X}} = \mu_X \text{ و } \gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}}$$

وبالتالي فعلاقة تحويل X إلى Z هي:

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\gamma_{\bar{X}}} \Rightarrow \frac{\bar{X} - \mu}{\frac{\gamma}{\sqrt{n}}}$$

مثال: إذا كان راتب الموظف يتبع التوزيع الطبيعي متوسطه μ 170 وانحرافه γ=8.

1) اخترنا موظفا عشوائيا فما احتمال أن يقل أجره عن (166ون).
2) سحبت عينة في 64 موظف فما هو احتمال أن يكون متوسط أجورهم أكبر من 172ون؟

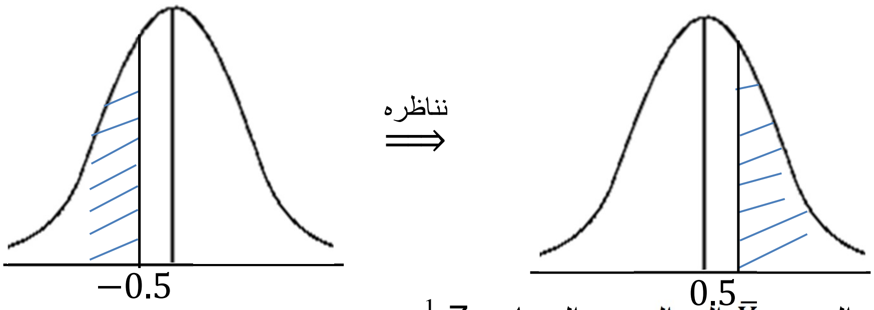
الحل: 1) نحول قيمة X إلى قيمة معيارية:

---

[1] – دفيد أندرسون وآخرون: الأساليب الكمية في الإدارة، ترجمة: محمد توفيق البلقيني ومرفت طلعت المحلاوي، دار المريخ للنشر والتوزيع، المملكة العربية السعودية، 2006، ص، ص: 92-95.

$$Z = \frac{X-\mu}{\gamma} = \frac{X-170}{8} \Rightarrow \frac{166-170}{8} = -0.5$$

$$\rho(X \leq 166) \Rightarrow \rho(z \leq -0.5) = 1 - \rho(z \leq -0.5) = 0.3085$$



2- نحول المتغير $\bar{X}$ إلى المتغير المعياري Z:[1]

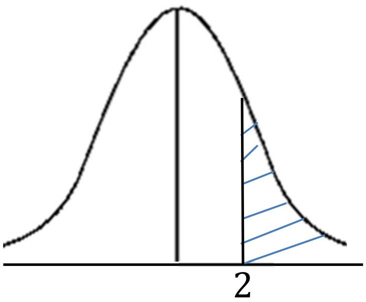
$$Z = \frac{\bar{X}-\mu_{\bar{X}}}{\gamma_{\bar{X}}}$$

$$\mu_{\bar{X}} = \mu_X = 170$$

و $\gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}} = \frac{8}{\sqrt{64}} = 1$

$$z = \frac{172-170}{1} = 2$$

$$\rho(\bar{X} \geq 172) \Rightarrow \rho(z \geq 2) = 1 - \rho(z \leq 2) = 1 - 0.9772 = 0.0228$$



[1]- قسم الإحصاء جامعة الملك عبد العزيز، مرجع سابق، ص، ص: 240/239.

ثامنا: سلسلة تمارين حول المحور الأول.

- **التمرين الأول:** قامت إحدى شركات الإلكترونيات بعمل دراسة على عدد الغسالات التي تم انتاجها فوجدت أنها تتبع التوزيع الطبيعي بعدد متوقع 150 وانحراف معياري 15.

المطلوب:

1- أحسب احتمال أن عدد الغسالات أقل من 150.

2- أحسب احتمال أن عدد الغسالات أقل من 130.

3- أحسب احتمال أن يكون عدد الغسالات محصور بين:

أ- 160 و190 غسالة.

ب- 120 و180 غسالة.

ت- 120 و140 غسالة.

- **التمرين الثاني:** يخضع معامل الذكاء للطلبة المسجلين في كلية العلوم الاقتصادية للتوزيع الطبيعي ذو الوسط $\mu = 100$ والانحراف المعياري 12،

المطلوب:

1- ماهي نسبة الطلبة التي يفوق معدل ذكائهم 100؟

2- ماهي نسبة الطلبة التي يقل معدل ذكائهم 130؟

3- ماهي نسبة الطلبة المحصور معدل ذكائهم بين 90 و140؟

- **التمرين الثالث:** إذا كانت نقاط طلبة السنة الثانية علوم تسيير تخضع للتوزيع الطبيعي بمعدل 60 وتباين 44 مع العلم أن إدارة القسم تمنح جوائز تقديرية للمتفوقين الأوائل بنسبة 3%.

ماهي أقل نقطة ستحصل على جائزة تقديرية من إدارة القسم؟

- **التمرين الرابع:**

أوجد قيم T لكل مما يأتي :

T[0.995,18] و T[0.975,20]

T[0.95,10] و T[0.05,8]

- **التمرين الخامس:**

أوجد لكل مما يأتي:

F[0.01,10.5] و F[0.05,10.8]

F[0.05,12.7] و F[0.10,15.15]

- **التمرين السادس:**

أوجد قيمة X^2 :

$X^2[0.025,15]$ و $X^2[0.975, 15]$

$X^2[0.05,10]$ و $X^2[0.99, 10]$

- **التمرين السابع:** تخضع معدلات الطلبة في كلية العلوم الاقتصادية للتوزيع الطبيعي ذو الوسط 60 وتباين 49، وكان عدد الطلبة 500 طالب.

المطلوب: ما هو عدد الطلبة الحاصلين على:

أ- أعلى من 65.

ب- أقل من 80.

ت- محصور بين 50 و70.

- **التمرين الثامن:**

1. في توزيع T بدرجات حرية 10 أوجد:

$$\rho(t < -2.5258)\rho(t > 1.5)$$

2. في توزيع X^2 بدرجات حرية 15 أوجد:

$$\rho(X^2 < 27.587) \qquad \rho(X^2 < 5.697$$

المحور الثاني:

توزيعات المعاينة

أولا: مدخل توزيعات المعاينة.

عند اختيار عينة وحساب إحصائية لها وليكن الوسط الحسابي$\bar{X}$ فإن قيمة الوسط الحسابي سوف تختلف عند سحب عينة ثانية وثالثة...، أي سوف تختلف قيمته من عينة إلى أخرى والسؤال هو: ماذا يكون الوضع لو تم سحب كل العينات الممكنة؟ وما هو التوزيع الاحتمالي الذي تأخذه الإحصائية المحسوبة لكل هذه العينات والذي يسمى (توزيع المعاينة الإحصائية).

ويمكن تعريف توزيع المعاينة (التابع الإحصائي) هو التوزيع الاحتمالي لهذه الإحصائية المحسوبة لكل العينات الممكنة والتي حجمها n والمأخوذة من المجتمع الإحصائي المدروس أيا كان حجمه وأيا كانت طريقة السحب.[1]

توزيع المعاينة: هو توزيع احتمالي لجميع القيم الممكنة لإحصاء ما في كل العينات العشوائية الممكنة والتي لها نفس الحجم،

أما الانحراف المعياري أو الخطأ المعياري لقيم الإحصاء الممكنة في كل ما في العينات العشوائية الممكنة من نفس الحجم.[2]

ويمكن تقديم بعض المفاهيم المهمة في توزيعات المعاينة:

- معلمات المجتمع: هي المقاييس الإحصائية الخاصة بالمجتمع وعادة تكون مجهولة كالمتوسط الحسابي μ والانحراف المعياري γ.
- إحصاءات العينة: هي المقاييس الإحصائية الخاصة بالعينة وهي تستخدم عادة في الاستدلال على معالم المجتمع، أي هي التي تستخدم في تقرير معالم المجتمع أو في اختبار الفروض حولها كما أنها (إحصاءات العينة) تصف العينة بالاعتماد على القياس الكمي لمفردات العينة مثل (المتوسط الحسابي للعينة، تباين العينة، النسبة في العينة).[3]
- المعاينة بالإرجاع والمعاينة بدون إرجاع: نميز المعاينة بالإعادة عن المعاينة بدون إعادة وذلك بأنها مسحوبة في مجتمع منتهي ومجتمع غير منتهي فإذا كانت المعاينة مأخوذة من مجتمع حجمه N فإن حجم العينة المسحوبة لن يتجاوز حجم المجتمع N.

---

[1] - أحمد عودة بن عبد المجيد، منصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 147-148.

[2] - جورج كانافوس ودون ميلر، مرجع سابق، ص271.

[3] - محمد بوعلاق، مرجع سابق، ص16.

أما إذا كانت العينة المسحوبة في مجتمع غير منتهي العدد فإنه يمكن سحب عينة بأي حجم نشاء مهما كان عددها ومن ناحية أخرى إذا كانت العينة مسحوبة من مجتمع منتهي وكان السحب بطريقة الإعادة فإن هذا المجتمع يفترض على أنه مجتمع غير منتهي.[1]

أي أن المجتمع المنتهي يتكون من عدد محدود أو منته من المشاهدات (مثل عدد طلاب إحدى الجامعات في عام دراسي معين)، أما المجتمع غير المنتهي فهو يتكون من عدد غير محدود من المشاهدات مثل عدد الأسماك.[2]

خطأ المعاينة: يعني يمكن أن يكون هناك فرق بين القيمة المقدرة من ناحية والقيمة الفعلية لمعلمه المجتمع من ناحية أخرى هذا لأننا نستخدم عينة فقط وليس المجتمع كله كبيانات مدخلات لتقديرنا، وليست التقديرات جيدة ودقيقة بالضرورة ولا يمكن حساب خطأ المعاينة عمليا لأننا لا نعرف القيمة الحقيقية لمعلمة المجتمع.[3]

- بعض العلاقات المهمة:[4]

وسط العينة	$\bar{X}=\frac{\sum_{i=1}^{n} xi}{n}$
تباين العينة	$S^2=\frac{\sum_{i=1}^{n}(xi-\bar{X})2}{n}$
تباين العينة الغير متحيز	$S^2=\sum_{i=1}^{n}\frac{(xi-\bar{x})2}{n-1}$ $S^2=\frac{n}{n-1}\gamma^2$ مجتمع محدود $S^2=\gamma^2$ كبير جدا Nمجتمع غير محدود أو حجمه
الانحراف المعياري	$S=\sqrt{s^2}$

أي أن متوسط المجتمع $\mu=\frac{\sum Xi}{N}$

تباين المجتمع $\mu\frac{\sum(Xi-\mu)2}{N}$

[1] - عزام صبري، الإحصاء الرياضي، دار صفاء للنشر والتوزيع، عمان-الأردن، 2010، 233.
[2] - أحمد عودة بن عبد المجيد عودة، منصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص 146.
[3] - إدوارد مينيكا وزوريانا كورزيجا: الإحصاء في الإدارة مع التطبيق على الحاسوب الالي، (الكتاب الأول)، ترجمة سرور علي إبراهيم سرور، دار المريخ للنشر والتوزيع، الرياض، السعودية، ص: 439.
[4] - برنارد تايلور الثالث، مرجع سابق، ص 674.

ثانيا: توزيع المعاينة للمتوسط$\bar{X}$ (للوسط الحسابي)

قبل تحديد توزيع المعاينة للوسط الحسابي للعينة$\bar{X}$ نحتاج لدراسة بعض المعلمات التي يمكن أن نحتاجها لوصف هذا التوزيع منها المتوسط والتباين وعلاقة هذه الإحصاءات بمعلمات المجتمع متوسطه وتباينه وسيتم الشرح من خلال هذا المثال.

مثال عن السحب بالإرجاع وبدون إرجاع:

إذا كان لدينا المجتمع الإحصائي التالي: 1,2,6

أي أن حجم المجتمع N=3 سحبنا جميع العينات الممكنة التي حجمها n=2.

متوسط المجتمع: $\mu = \frac{1+2+6}{3} = 3$

وتباين المجتمع $\gamma^2 = \frac{1}{n}\sum_{i=1}^{N} X_1^2 - \mu^2$

$\gamma^2 = 4.67 = \frac{41}{3} - \frac{27}{3} = \frac{1+4+36}{3} - 9$

$\gamma = \sqrt{4.67} = 2.16$

1) حساب المتوسط والانحراف المعياري في السحب بدون إرجاع لتوزيع المعاينة:

عدد العينات الممكنة: $\frac{3!}{2!} = \frac{N!}{n!} = 3$

والعينات هي: (2،6)، (1،6)، (1،2)

متوسطاتها هي: 4، 3.5، 1.5

وإذا كانت المعاينة عشوائية بسيطة فإن كل عينة من العينات السابقة لها احتمال مساوي قدره $\frac{1}{3}$

وبالتالي توزيع المعاينة للمتوسط $\bar{X}$ بالشكل التالي:

$\bar{X}$	1.5	3.5	4	Σ
$F(\bar{X})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1
$\bar{X}F(\bar{X})$	$\frac{1.5}{3}$	$\frac{3.5}{3}$	$\frac{4}{3}$	$\frac{9}{3} = 3$
$\bar{X}^2F(\bar{X})$	$\frac{2.25}{3}$	$\frac{12.2}{3}$	$\frac{16}{3}$	$\frac{30.5}{3}=10.17$

وبالتالي نجد أن:

$$\mu_{\bar{X}} = E(\bar{X}) = \sum_{i=1}^{3} \bar{X}_i\, F(\bar{X}_i)$$

$$E(\bar{X}) = \left(\frac{1}{3}1.5\right) + \left(\frac{1}{3}3.5\right) + \left(\frac{1}{3}4\right)$$

$$E(\bar{X}) = \frac{1}{3}(1.5 + 3.5 + 4) = \left(\frac{1}{3}\right)(9) = 3 = \mu$$

$$\gamma_{\bar{X}}^2 = \sum_{i=1}^{3} \bar{X}_i^2 \,.\, F(\bar{X}_i) - (\mu_{\bar{X}})^2$$

$$E(\bar{X}^2) = \frac{1}{3}(2.25 + 12.25 + 16) = \left(\frac{30}{3}\right) = 10.17$$

$$\gamma_{\bar{X}}^2 = 10.17 - (3)^2 = 1.17$$

ونستنتج أنه يمكن اثبات ذلك رياضيا:

$$E(\bar{X}) = \mu_{\bar{X}} = \mu = 3 \quad \text{و} \quad \gamma_{\bar{X}}^2 = \frac{\gamma^2}{n}\,\frac{N-n}{N-1} \Longrightarrow \gamma_X = \frac{\gamma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$\gamma_{\bar{X}} = \frac{\gamma^2}{n}\frac{N-n}{N-1} = \frac{4.67}{2}\,\frac{3-2}{3-1} = 2.335 * \frac{1}{2} = 1.17$$

2) حساب المتوسط والانحراف المعياري في السحب بالإرجاع عدد العينات الممكنة: $N^n = 3^2 = 9$

والعينات هي: (6.6)، (6.2)، (6.1)، (2.6)، (2.2)، (2.1)، (1.6)، (1.2)، (1.1)

ومتوسطاتها: 6، 4، 3.5، 4، 2، 1.5، 3.5، 1.5، 1.

وتوزيع المعاينة للمتوسط$\bar{X}$ بالشكل التالي:

$\bar{X}$	1	1.5	2	3.5	4	6	Σ
F$(\overline{X})$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	1
$\bar{X}$F$(\overline{X})$	$\frac{1}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{7}{9}$	$\frac{8}{9}$	$\frac{6}{9}$	$\frac{27}{9} = 3$
$\bar{X}^2$F$(\overline{X})$	$\frac{1}{9}$	$\frac{4.5}{9}$	$\frac{4}{9}$	$\frac{24.5}{9}$	$\frac{32}{9}$	$\frac{36}{9}$	$\frac{102}{9}$

نلاحظ ما يلي:

[1] $E(\bar{X}) = \mu_{\bar{X}} = \sum_{i=1}^{N} \bar{X}.F(\bar{X}_i) = 3 = \mu$

$$\gamma_{\bar{X}}^2 = \sum_{i=1}^{N} \bar{X}_i^2\, F(\bar{X}_i) - [E(\bar{X})]^2 = \frac{102}{9} - (\frac{9}{3})^2 = \frac{21}{9} = 2.33$$

$$\gamma_{\bar{X}}^2 = \frac{\gamma^2}{n} = \frac{4.67}{2} = 2.33$$

[1] - أحمد عودة بن عبد المجيد، منصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 148-151.

بعد عرض المثال يمكن شرح توزيع المعاينة للمتوسط $(\bar{X})$ من خلال افتراض أن لدينا مجتمع مفرداته تتبع توزيعا احتماليا معينا، اخترنا عينة عشوائية أولى حجمها n مفردة من هذا المجتمع وتم حساب وسطها الحسابي فكان $_1\bar{X}$، ثم اخترنا عينة عشوائية ثانية لها نفس الحجم n مفردة وحسبنا وسطها الحسابي $_2\bar{X}$ ثم اخترنا عينة عشوائية ثالثة لها نفس الحجم ومتوسطها $_3\bar{X}$، وتم تكرار هذه العملية لجميع العينات التي يمكن سحبها من هذا المجتمع أي سيتوفر لدينا عدد كبير من المتوسطات الحسابية للعينات $(_1\bar{X}$، $_2\bar{X}....$ $_n\bar{X})$ وليست كلها متساوية، وهذه القيم ستكون مجتمعا آخر جديد عدد مفرداته أكبر من مفردات المجتمع الأصلي ولذلك يمكن القول أن الوسط الحسابي متغير عشوائي يأخذ قيما مختلفة، ويتبع توزيعا احتماليا معينا قد يختلف عن توزيع المجتمع الأصلي، ويسمى هذا المجتمع الجديد بمجتمع "المتوسطات الحسابية" أو "توزيع المعاينة للوسط الحسابي" وبهذا يمكن أن نعرف توزيع المعاينة للوسط الحسابي بأنه التوزيع الاحتمالي لمجتمع هذا المقياس. [1]

أي أن توزيع المعاينة للوسط هو عبارة عن أخذ عينات متكررة من المجتمع الأصلي وقمنا بحساب المتوسط لكل عينة فإننا سنجد أن هذه المتوسطات مختلفة عن بعضها البعض، أي هو توزيع احتمالي لمتوسطات العينات، وتوزيع المعاينة للوسط له أيضا وسط يعبر عنه بالرمز $\mu_{\bar{X}}$ وانحراف معياري أو خطأ معياري $\gamma_{\bar{X}}$.

ويمكن عرض نظريتين تربطان بين توزيع المعاينة للوسط والمجتمع الأصلي.

النظرية الأولى: إذا أخذنا عينات متكررة حجمها n من مجتمع ما:

1) $\mu_{\bar{x}} = \mu_x$

2) $\gamma_{\bar{x}} = \frac{\gamma}{\sqrt{n}}$

$$\gamma_{\bar{x}} = \frac{\gamma}{\sqrt{n}}\sqrt{\frac{N-n}{N-n}}$$

تستخدم العلاقة 2.ب للمجتمعات المحدودة ذات الحجم N عندما تكون $[n \geq 0.05N]$ أي حالة سحب بدون إرجاع.

حيث $\sqrt{\frac{N-n}{N-n}}$ هو معامل التصحيح.

[1] - قسم الإحصاء جامعة الملك عبد العزيز، مرجع سابق، ص 238.

النظرية الثانية: مع تزايد حجم العينات (أي عندما $n\to+\infty$) فإن توزيع المعاينة للمتوسط يقترب من التوزيع الطبيعي بصرف النظر عن شكل المجتمع الأصلي ويعتبر التقريب جيدا عندما تكون ($n\geq 30$) هذه هي نظرية النهاية المركزية التي أشرنا إليها سابقا.[1]

النظرية الثالثة:

1) إذا كانت قيمة الانحراف المعياري في المجتمع معلومة وكان:

أ) توزيع المجتمع هو التوزيع الطبيعي.

ب) توزيع المجتمع ليس طبيعي ولكن حجم العينة ($n\geq 30$) فإن توزيع المعاينة يتبع Z هو تقريبا التوزيع الطبيعي $Z=\frac{\bar{x}-\mu_{\bar{x}}}{\gamma}$

2) إذا كانت قيمة الانحراف المعياري في المجتمع غير معلومة فإن:

أ) توزيع المجتمع هو التوزيع الطبيعي.

ب) حجم العينة ($n< 30$) توزيع المعاينة يتبع T

حيث $T=\frac{\bar{x}-\mu}{s\sqrt{n}}$ هو تقريبا توزيع T بدرجات حرية ($n-1$).[2]

النظرية الرابعة:

إذا كانت γ^2 مجهولة فإن ما نستطيع فعله هو أن نستبدل الانحراف المعياري للمجتمع بالانحراف المعياري للعينة S وتصبح:

1) يخضع 'Z إذا كان: $n\geq 30 \Leftarrow Z=\frac{\bar{x}-\mu}{s\sqrt{n}}$

2) يخضع لتوزيع T إذا كان ($n< 30$)[3]

1- توزيع المعاينة للوسط الحسابي $\bar{X}$ عندما يكون المجتمع له توزيع طبيعي:

إذا أخذت العينة العشوائية $X_1, X_2.....X_n$ وكان $\bar{X}$ الوسط الحسابي لهذه العينة، فما توزيع $\bar{X}$؟

قبل الإجابة عن هذا السؤال على الطالب أن يلاحظ أن في مرحلة التخطيط لدراسة عينة ما وقبل جمع البيانات أو تسجيل قيم المشاهدات يكون بإمكانك التعبير عن الوسط الحسابي $\bar{X}$ بدلالة الاحتمالات فقط لأن الوسط حينذاك يكون عبارة عن مجموع متغيرات عشوائية مقسومة على عدد ثابت n، ولذلك فقبل إجراء التجربة وتسجيل قيم المشاهدات يكون الإحصاء $\bar{X}$ متغيرا عشوائيا، أما بعد التجربة فإن قيمته تتعين

[1] - دومينيك سالفاتور: ملخصات شوم نظريات ومسائل في الإحصاء والاقتصاد القياسي، ترجمة سعدية حافظ منتصر، الدار الدولية للاستثمارات الثقافية،القاهرة، 1997، ص67.

[2] - جورج كانافوس ودون ميلر، مرجع سابق، ص 291.

[3] - محمد صبحي أبو صالح، عدنان محمد عوض: مقدمة في الإحصاء، مرجع سابق، ص143.

من العينة المشاهدة وبما أن $\bar{X}$ متغير عشوائي فإنه يكون له توزيع احتمالي يعرف بتوزيع المعاينة للوسط وهو التوزيع الاحتمالي لجميع الأوساط $\bar{X}$ لجميع العينات العشوائية البسيطة ذات حجم معين والتي يمكن أخذها من المجتمع تحت الدراسة، والسؤال الآن هو إذا أخذت جميع العينات العشوائية البسيطة ذات الحجم n من مجتمع إحصائي فمت هو التوزيع الذي يخضع له الإحصاء $\bar{X}$؟

وللإجابة على السؤال سنعرض النظرية التالية:

إذا كانت $X_1, X_2 .....X_n$ عينة عشوائية من مجتمع طبيعي وسطه (μ) وتباينه γ^2 فإن توزيع $\bar{X}$ يكون التوزيع الطبيعي ذو الوسط μ والتباين $\frac{\gamma^2}{n}$ أي أن المتغير العشوائي $Z=\frac{\bar{X}-\mu}{\frac{\gamma}{\sqrt{n}}}$ يخضع للتوزيع الطبيعي المعياري. [1]

مثال: مصنع ينتج كراسي ترتكز على قاعدة دائرية اعتمادا على التجارب السابقة فإن مفتش الرقابة على العملية الإنتاجية مقتنع بما يلي:

1- متوسط قطر القاعدة الدائرية 5 سم.

2- الانحراف المعياري لها 0.005 سم.

3- توزيع العملية الإنتاجية هو التوزيع الطبيعي.

يهتم الفاحص بالمحافظة على متوسط قطر العملية الإنتاجية عند 5 سم، ولتحقيق ذلك تسحب عينات عشوائية بصفة دورية حجم كل منها 9 كراسي وذلك في محاولة الاكتشاف أيه انحرافات عن الأرقام الطبيعية المشار إليها.

المطلوب:

1- حدد توزيع المعاينة.

2- بفرض أن الفاحص سحب عينة عشوائية من 9 كراسي, وقيست أقطار قاعدتها ووجد أن $\bar{X}=$ 5.004 سم، ماهي إمكانية احتمال أن متوسط القطر في تلك العينة العشوائية سيكون على الأقل 5.004 سم على فرض أن متوسط العملية باقيا عند 5 سم والانحراف المعياري 0.005 سم؟

3- ما هو حجم العينة التي يجب سحبها لتحقيق خطأ معياري ل $\bar{X}$ يساوي 0.001.

4- لماذا يفضل أن يكون الخطأ المعياري لـ$\bar{X}$ يساوي 0.001 سم، على أن يكون الخطأ المعياري كما حصلت عليه في السؤال الأول.

[1] - محمد صبحي أبو صالح، عدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص183.

الحل:

1. بما أن توزيع العملية الإنتاجية طبيعي له $\mu = 5$،$\gamma = 0.005$ فإن توزيع المعاينة لـ$\bar{X}$ يكون أيضا طبيعي بمتوسط $\mu_X = 5$ وخطأ معياري عند n=9:

$$\gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}} = \frac{0.005}{\sqrt{9}} = 0.001667$$

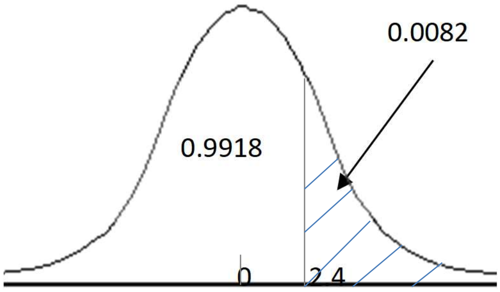
2. هذا السؤال يقع في صميم الاستنتاج الإحصائي، نتيجة العينة التي حصلنا عليها هي: $\bar{X} = 5.004$، ما هي فرصة احتمال وقوع مثل هذه النتيجة اذا كان هذا الاحتمال كبير، فهذا يعني أن هناك سببا ضعيفا لكي نشك في وقوع انحراف عن متوسط العملية الإنتاجية، وبالتالي فإن أي تغيير في النظام الحالي للعملية الإنتاجية يعد عبئا على المصنع. من ناحية أخرى إذا كان هذا الاحتمال صغيرا فمن الممكن أن يكون هناك سببا مقنعا للتصديق بوقوع انحراف عن متوسط العملية الإنتاجية.

حيث أن توزيع المعاينة لـ $\bar{X}$ هو الطبيعي بمتوسط $\mu_{\bar{X}} = 5$ وانحراف معياري $\gamma_{\bar{X}} = 0.001667$، فإنه يمكن تحديد الاحتمال المطلوب بأن قيمة $\bar{X}$ على الأقل 5.004 وذلك بتحويل القيمة 5,004 إلى قيمة معيارية Z المناظرة. أي لحساب $\rho(\bar{X}) \geq 5.004$ يجب حساب:

$$Z = \frac{5.004 - 5}{0.001667} = 2.4$$

$$\rho(\bar{X}) \geq 5.004 \Rightarrow \rho(z \geq 2.4) = 1 - \rho(z < 2.4) = 1 - 0.9918 = 0.0082$$

من الواضح أن احتمال قدره أقل من 1% يعتبر صغيرا جدا، لذا فهناك سبب مقنع للتصديق بأن الانحراف عن $\mu = 5$ فقد حدث، ولكن يجب أن نكون حذرين قبل أن نقرر بأن هناك حاجة لإجراء عملا تصحيحيا على العملية الإنتاجية.



3. حجم العينة المطلوب يتحدد ببساطة بمساواة الخطأ المعياري لـ$\bar{X}$ بالقيمة المطلوبة له والحل بالنسبة إلى n أي:

$$\gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}} \Rightarrow \frac{0.005}{\sqrt{n}} = 0.001$$
$$\sqrt{n} = \frac{0.005}{0.001} = 5 \Rightarrow n = 25$$

4. الخطأ المعياري 0.001 هو أصغر من 0.001667 وهو الخطأ المعياري في السؤال الأول فإذا كان الخطأ المعياري هو 0.001 فإن الاستنتاج اعتمادا على $\bar{X}$ يكون أكثر موثوقية، فمثلا نعيد حساب السؤال الثاني باستخدام $\gamma_{\bar{X}} = 0.001$ نجد أن احتمال أن يكون متوسط العينة العشوائية من العملية الإنتاجية الطبيعية، يكون أكبر من 5.004 هو:

$$\rho(\bar{X} \geq 5.004)$$
$$Z = \frac{5.004 - 5}{0.001} = 4$$
$$\rho(\bar{X} \geq 5.004) \Longrightarrow \rho(Z \geq 4) = 1 - \rho(z < 4) = 1 - 0.9999 = 0.001$$

وهو احتمال ضئيل جدا عن الذي حصلنا عليه من قبل، وبالتالي فإن القرار بأن متوسط العملية الانتاجية قد انحرف عن المعالم المحددة لها أصبح الآن أكثر قناعة. [1]

2-<u>توزيع المعاينة عندما يكون المجتمع له توزيع غير طبيعي (نظرية النهاية المركزية)</u>

في كثير من الحالات، لا نستطيع تعيين هوية توزيع المجتمع وبالتالي لا يمكن تحديد توزيع المعاينة $\bar{X}$ مع ذلك فقد تمكن علماء الإحصاء من إثبات أن توزيع المعاينة لـ$\bar{X}$ هو تقريبا التوزيع الطبيعي في حالة العينات ذات الأحجام الكبيرة أيا كان توزيع المجتمع هذه النتيجة الحاسمة في الاستنتاج الإحصائي تعرف باسم "نظرية النهاية المركزية" والتي تنص على أنه كلما زاد حجم العينة، كلما اقترب توزيع المعاينة من التوزيع الطبيعي بغض النظر عن توزيع المجتمع. [2]

نظرية: إذا أخذت عينة عشوائية حجمها n من مجتمع إحصائي وسطه μ وتباينه γ^2، فإن توزيع الوسط الحسابي للعينة $\bar{X}$ يقترب من التوزيع الطبيعي ذي الوسط الحسابي $\mu_{\bar{X}}$ والتباين $\frac{\gamma^2}{n}$ كلما كبرت n وبعبارة أخرى فإن: $\frac{\bar{X}-\mu}{\frac{\gamma}{\sqrt{n}}}$ يقترب من التوزيع الطبيعي المعياري كلما كبرت n، وعند التطبيق $n \geq 30$ تعتبر كبيرة بشكل كافي، من أجل استعمال النظرية.

مثال: تخضع أوزان على سائل غسل الصحون من نوع معين لتوزيع معدله 1000غ وانحرافه المعياري 80غ، إذا أخذت عينة عشوائية حجمها 48 علبة:

[1] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 281-282.
[2] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 283.

1- فما هو المتوسط والتباين والانحراف المعياري للوسط الحسابي $\bar{X}$ لأوزان العلب في العينة؟

2- ما احتمال أن يزيد الوسط الحسابي لأوزان العلب في العينة عن 1072غ؟

3- ما احتمال أن يقل الوسط الحسابي عن 980غ؟

الحل: بما أن (n=48>30) تعتبر كبيرة ومتوسط وتباين المجتمع معلومة وبالتالي فشروط النظرية محققة ويكون:

1. $\mu_{\bar{X}} = 1000g$

$$\gamma_{\bar{X}} = \frac{\gamma_X}{\sqrt{n}} \Rightarrow \frac{80}{\sqrt{48}} = 11.58g$$

2. $\rho(\bar{X} > 1072)$

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\gamma_{\bar{X}}} = \frac{1072 - 1000}{11.58} = 6.22$$

$$\rho(\bar{X} > 1072) \Rightarrow \rho(Z > 6.22) = 1 - \rho(Z < 6.22) = 1 - 1 = 0$$

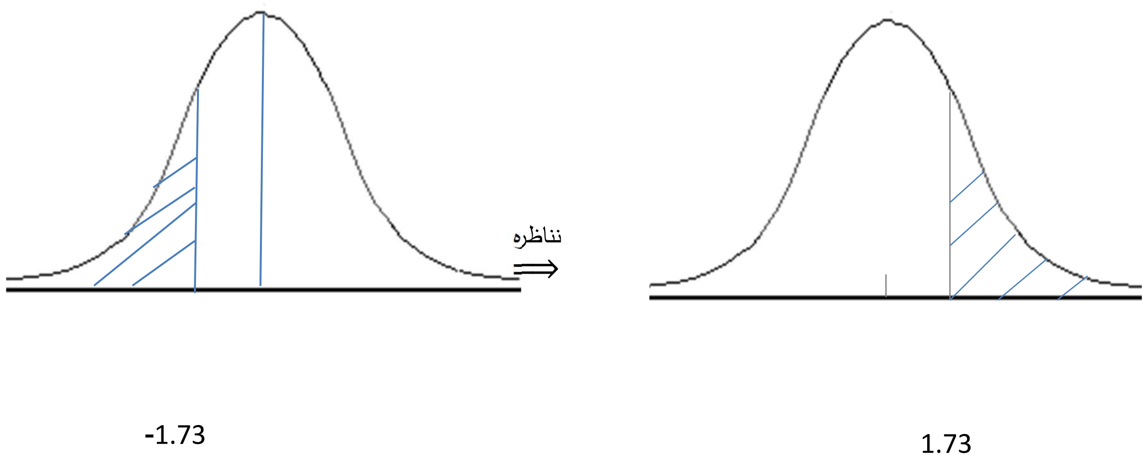
3.

$$\rho(\bar{X} > 980) =$$

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\gamma_{\bar{X}}} = \frac{980 - 1000}{11.58} = -1.73$$

$$\rho(\bar{X} < 980) \Rightarrow \rho(Z < -1.73)$$

$$\rho(Z < -1.73) = \rho(Z > 1.73) = 1 - \rho(Z < 1.73) = 0.0418$$



مثال: تخضع أوزان العبوات لأحد أنواع المبيدات لتوزيع وسطه 135غ وانحرافه المعياري 14غ، إذ قررت وزارة التموين رفض كل صندوق من هذه العبوات إذا نقص وزنه عن 6.24غ، فما نسبة الصناديق المرفوضة، علما بأن عدد العبوات في كل صندوق 48 عبوة؟

الحل: المطلوب هو حساب احتمال أن يتقص وزن الصندوق عن 6.24غ أي 6240غ فإذا افترضنا أوزان العلب في الصندوق ما هي: $X_1, X_2 \dots X_{48}$ فالمطلوب إيجاد:

$$\rho\left(\sum_{i=1}^{48} X_i < 6240\right) = \rho\left(\frac{\sum X_i}{48} < \frac{6240}{48}\right.$$

$$\rho(\bar{X} < 130)$$

وبما أن شروط نظرية النهاية المركزية محققة حيث μ و γ^2 معلومان وحجم العينة أكبر من 30 فإن بإمكاننا تطبيق النظرية باعتبارها تخضع للتوزيع الطبيعي المعياري

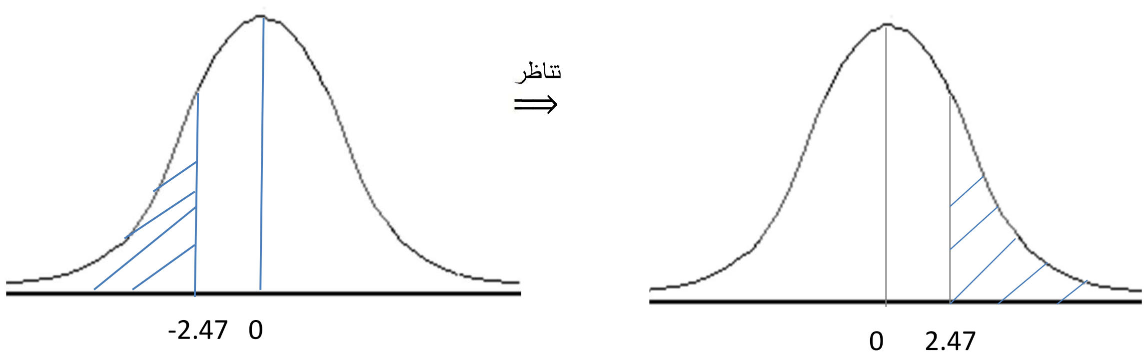
$$Z = \frac{\bar{X} - 135}{\frac{14}{\sqrt{48}}}$$

حساب

$$\rho(\bar{X} < 130)$$

$$Z = \frac{130 - 135}{\frac{14}{\sqrt{48}}} = -2.47$$

$$\rho(\bar{X} < 130) \Longrightarrow \rho(Z < -2.47) = \rho(Z > 2.47) = 1 - \rho(Z < 2.47) = 0.0068 \approx 0.007$$



أي أن الوزارة ترفض 7 بالألف من الصناديق تقريبا من خلال المثالين المدروسين نلاحظ أن تباين المجتمع γ^2 كان معلوما ولذلك أمكننا تطبيق النظريات المتعلقة بالوسط $\bar{X}$ في حالة المعاينة من مجتمع طبيعي أو الحالة التي فيها حجم العينة كبير.

والسؤال الذي يطرح نفسه هو: ماذا لو كان التباين γ^2 مجهولا؟ أي بماذا سنعوض γ^2 في حالة $n\geq 30$ سنعوض γ^2 ب S^2 تباين العينة أما في حالة $n< 30$ سندرسها في العناصر اللاحقة. [1]

3–<u>توزيع المعاينة لـ $\bar{X}$ عندما يكون الانحراف المعياري للمجتمع مجهول</u>

في كثير من التطبيقات المتعلقة باستعمال الوسط الحسابي$\bar{X}$ لا يكون الانحراف المعياري للمجتمع معروفا، ولذلك إذا كان حجم العينة $(n\geq 30)$ فإننا نستعمل الانحراف المعياري للعينة (S) بدلا من γ ويكون المقدار $\frac{\bar{X}-\mu_{\bar{X}}}{\frac{s}{\sqrt{n}}}$ خاضعا لتوزيع يمكن تقريبه بالتوزيع الطبيعي المعياري ما دامت $(n\geq 30)$، أما في الحالات التي تكون فيها $(n< 30)$ فإن الانحرافات المعيارية للعينة تكون ذات تغيير كبير لدرجة أن S لا تكون تقديرا موثوقا للانحراف المعياري γ وبالتالي فإن $\frac{\bar{X}-\mu_{\bar{X}}}{\frac{s}{\sqrt{n}}}$ يخضع لتوزيع لا يمكن تقريبه للتوزيع الطبيعي المعياري.

نظرية: إذا أخذت عينات عشوائية من توزيع طبيعي وسطه μ وتباينه γ^2 غير معلوم وإذا كان $\bar{X}$ الوسط الحسابي لعينة حجمها $(n< 30)$ وكان S الانحراف المعياري لهذه العينة فإن المتغير:

$$T=\frac{\bar{X}-\mu_{\bar{X}}}{\frac{s}{\sqrt{n}}}$$

حيث يخضع للتوزيع T بدرجة حرية $V=n-1$. [2]

مثال: وكالة لحماية البيئة حددت متوسطا لمعدل الأميال/ جالون على الطرق السريعة قدره 45 وذلك لنوع معين من السيارات. اشترت منظمة مستقلة للمستهلكين إحدى هذه السيارات واختبرتها للتحقق من المعدل وتم ذلك بقيادة السيارة لمسافة 100 ميل في 25 رحلة مختلفة وسجلت القيم الفعلية للأميال المقطوعة لكل جالون في كل رحلة من خلال 25 رحلة بعد حساب المتوسط والانحراف المعياري فكانا على التوالي 43.5، 2.5 ميل/ جالون.

هناك اعتقاد أن التوزيع الفعلي للأميال/ جالون على الطريق السريع لهذا النوع من السيارات يقترب من التوزيع الطبيعي.

المطلوب:

[1] - محمد صبحي أبو صالح، عدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 185-187.

[2] - محمد صبحي أبو صالح، عدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 187-188.

1- على افتراض أن معدل الوكالة 45 ميل/ جالون متحقق لهذه السيارة، أوجد احتمال أن متوسط الأميال/ جالون في العينة العشوائية المكونة من 25 رحلة يجب أن يكون 43.5 أو أقل؟

2- اعتمادا على بيانات العينة الحالية، هل هناك سببا مقنعا للمنظمة لكي تشك في أن معدل الوكالة متحققا لهذه السيارة؟

الحل:

1- حجم العينة n=25، أعطت النتائج: $\bar{X} = 43.5$، S=2.5 ميل/ جالون، لتحديد ما إذا كانت معلومات العينة هذه تؤكد المعدل المعطى من الوكالة، يجب أن نعتمد على الاحتمال، بمعنى أننا نرغب في الإجابة على السؤال التالي:

إذا كانت μ حقيقة تساوي 45 ميل/ جالون، ما هو احتمال أنه بالصدفة وحدها مشاهدة قيمة $\bar{X}$ تساوي 43.5 ميل/ جالون أو أقل؟ وحيث أن الانحراف المعياري في المجتمع غير معلوم، علينا حساب قيمة T التي تناظر $\bar{X}$=43.5، أي:

$$T = \frac{\bar{X} - \mu_{\bar{X}}}{\frac{s}{\sqrt{n}}} = \frac{43.5 - 45}{\frac{2.5}{\sqrt{25}}} = -3$$

وهي قيمة التوزيع T عند درجة حرية v=25−1=24

أي:

$$\rho(\bar{X} \leq 43.5) = \rho(T_{24} \leq -3) = 0.005$$

2- اعتمادا على الإجابة في السؤال الأول من البديهي الشك في معدل الوكالة، ومع ذلك لا نستطيع لوم الوكالة على معدلها المرتفع وغير المناسب، قبل إجراء العديد من الأبحاث الإضافية، لأن التعارض ربما يكون نتيجة للفروق بين طريقتي القياس للأميال في وكالة حماية البيئة ومنظمة المستهلكين. [1]

ثالثا: توزيع المعاينة للفرق $(\bar{X}_1 - \bar{X}_2)$

نفرض أننا نرغب في مقارنة متوسطي مجتمعين أو عمليتين مستقلتين سنرمز للمتوسطات بالرموز μ_1، μ_2 والانحرافات المعيارية للمجتمعات بالرموز γ_1، γ_2 وبعد سحب عينتين عشوائيتين مستقلتين أحجامها n_1، n_2 من المجتمعين، غالبا ما يكون الهدف هو تحديد ما إذا كان من الممكن اعتبار متوسطات تلك المجتمعات متساوية أم لا، في هذه الحالة يكون من المناسب رياضيا أن نعتبر الفرق $\mu_1 - \mu_2$ على أنه

[1] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 292.

المعلمة محل الاهتمام فإذا كان الفرق $\mu_1 - \mu_2$ يساوي الصفر فإن متوسطي المجتمعين μ_1، μ_2 يكونان متساويان.

حيث أن $\bar{X}_1$ هي أفضل إحصاءة تستخدم للاستنتاج حول μ_1 وأن $\bar{X}_2$ هي أفضل إحصاءة استنتاج حول μ_2، وأفضل إحصاءة للاستنتاج حول الفرق $(\mu_1 - \mu_2)$ هو $(\bar{X}_1 - \bar{X}_2)$ ويعتبر $\bar{X}_1 - \bar{X}_2$ إحصاءة لأنه فرق بين إحصائيتين أي أن العينتين العشوائيتين تعطي القيم $\bar{X}_1$، $\bar{X}_2$ والفرق هو إحصاءة ناتجة من معاينة عشوائية وهو غير متحيز وله أصغر خطأ معياري.

- القيمة المتوقعة للفرق بين متغيرين عشوائيين هي الفرق بين توقعهما

$$\mu_{(\bar{X}_1 - \bar{X}_2)} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

- لتقدير الخطأ المعياري للفرق $\bar{X}_1 - \bar{X}_2$ يلزمنا تحديد تباينه، حيث إذا كان المتغيرين العشوائيين المستقلين فإن تباين الفرق يساوي مجموع تباينهما

$$\gamma^2_{(\bar{X}_1 - \bar{X}_2)} = \gamma^2_{\bar{X}_1} + \gamma^2_{\bar{X}_2} = \frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}$$

والانحراف المعياري للفرق

$$\gamma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}$$

والمعنى العام للصيغ هو نفسه في حالة مجتمع واحد إذا استطعنا سرد قائمة بكل العينات العشوائية المستقلة ذات الحجم n_1،n_2 من المجتمعين، فإنه يمكن تكوين قائمة بالفروق المتناظرة للنواتج $\bar{X}_1 - \bar{X}_2$ هذه القيم تتجه لأن تتجمع حول الفرق $\mu_1 - \mu_2$ بعضها سيكون أعلى منها والآخر يقل عنها ولكن بصفة عامة.

معظمها يتمركز حول $\mu_1 - \mu_2$، بالإضافة إلى ذلك كلما زاد حجم العينة أو كلاهما تناقص الخطأ المعياري للفرق $\bar{X}_1 - \bar{X}_2$ وبالتالي تحسنت دقة الاحصاءة كمقدر للفرق $\mu_1 - \mu_2$ في الحقيقة إذا تساوى حجم العينات فإن الخطأ المعياري يتناسب عكسيا مع الجذر التربيعي لحجم العينة المشترك وهذه نفس الخاصية في حالة المجتمع الواحد. [1]

1- <u>توزيع المعاينة للفرق $\bar{X}_1 - \bar{X}_2$ عندما تكون γ_1، γ_2 معلومة</u>

[1] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 272-273.

إذا أخذت العينات العشوائية من توزيعات طبيعية مستقلة عن بعضها فإن توزيع الفرق بين الوسيطين يعطى بالنظرية التالية:

إذا أخذت عينة عشوائية حجمها n_1 من توزيع طبيعي معدله وتباينه وأخذت عين عشوائية حجمها n_2 من مجتمع طبيعي معدله μ_2 وتباينه γ_2^2 ومستقل عن المجتمع الأول ورمزنا للوسط الحسابي للعينة الأولى بالرمز $\bar{X}_1$ والوسط الحسابي للعينة الثانية $\bar{X}_2$ فإن توزيع $\bar{X}_1 - \bar{X}_2$ يكون التوزيع الطبيعي ذا المعدل $\mu_1 - \mu_2$ والتباين $\frac{\gamma_1}{n_1} + \frac{\gamma_2}{n_2}$ أي أن:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_{\bar{X}_1} - \mu_{\bar{X}_2})}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

يخضع للتوزيع الطبيعي المعياري N(0.1)

مثال: إذا كانت رواتب المعلمين في القطاع العام تخضع للتوزيع الطبيعي معدله 230 وانحرافه المعياري (36 ون) ورواتب المعلمين في المدارس الخاصة تخضع لتوزيع طبيعي معدله 180ون وانحرافه المعياري 40ون أخذت عينة عشوائية من المعلمين في القطاع العام حجمها 16 معلما وسطها الحسابي $\bar{X}_1$، وعينة عشوائية من معلمي المدارس الخاصة حجمها 10 وسطها $\bar{X}_2$، أوجد احتمال أن يزيد $\bar{X}_1$ عن $\bar{X}_2$ بمقدار 60

الحل: حساب احتمال $\rho(\bar{X}_1 - \bar{X}_2 \geq 60)$

بما أن شروط النظرية السابقة متوفرة فإن: [1]

$$\rho(\bar{X}_1 - \bar{X}_2 \geq 60)$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

$$Z = \frac{60 - (230 - 180)}{\sqrt{\frac{(36)^2}{16} + \frac{(40)^2}{10}}} = \frac{10}{15.52} = 0.64$$

$$\rho(\bar{X}_1 - \bar{X}_2 \geq 60) = \rho(Z \geq 0.64) = 1 - \rho(Z < 0.64) = 1 - 0.7389 = 0.2611$$

2-توزيع المعاينة للفرق $\bar{X}_1 - \bar{X}_2$ عندما تكون γ_1، γ_2 مجهولة

[1] - محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 192-193.

أ- إذا كان تباين المجتمعين متساويين $\gamma_1^2 = \gamma_2^2$ لكنهما مجهولين:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_c\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

يخضع التوزيع T بدرجة حرية $n_1 + n_2 - 2$ حيث S_c^2 هو التباين المجمع للعينتين معا والمعرف بالعلاقة التالية: [1]

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

ب- إذا كان تباين المجتمعين مجهولة وغير متساوية $\gamma_1^2 \neq \gamma_2^2$

إذا كان افتراض تساوي التباينات يبدو غير مقبولا، فإنه يمكن تقدير الانحراف المعياري لـ $\bar{X}_1 - \bar{X}_2$ باستخدام تباينات العينات S_1^2، S_2^2 بمعنى أن الانحراف المعياري للفرق $\bar{X}_1 - \bar{X}_2$ يتحدد باستخدام:

$$\gamma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

نجد أن الإحصاءة المناسبة تعطى بالعلاقة:

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

توزيع المعاينة للإحصاء T تتبع التوزيع T بدرجة حرية لها الصيغة المركبة التالية: [2]

$$V = \frac{[\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}]^2}{\frac{[\frac{S_1^2}{n_1}]^2}{n_1} + \frac{[\frac{S_2^2}{n_2}]^2}{n_2}}$$

مثال:

سحبت عينة عشوائية حجمها 14 و$\bar{X}_1$ وسطها الحسابي وتباينها 6، من مجتمع طبيعي وسطه 26 وتباينه مجهول، وسحبت عينة أخرى وسطها $\bar{X}_2$ من مجتمع مستقل عن الأول وسطه 24 وتباينه مجهول، وكان حجم العينة 22 وتباينها 8.

---

[1] - محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء، مرجع سابق، ص145.

[2] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 376.

المطلوب: على افتراض أن تباين المجتمعين متساويين فما هو احتمال أن يكون الفرق بين متوسطي العينتين أقل من 4؟

الحل: بما أن $n_1 = 14$ و $n_2 = 22$ وأقل من 30 وتباين المجتمعين متساوي ومجهول أما تباين العينتين

$$S_1^2 = 6 \text{ , } S_2^2 = 8$$

وأوساط المجتمعات

$$\mu_1 = 26 \text{ , } \mu_2 = 24$$

بما أن تباين المجتمعين متساوي ومجهول فإننا نستخدم التوزيع T بدرجة حرية

$$n_1 + n_2 - 2$$
$$14 + 22 - 2 = 34$$
$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_c\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$
$$S_c^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(14 - 1)6 + (22 - 1)8}{34} = 7.24$$
$$S_c = 2.69$$
$$T = \frac{4 - (26 - 24)}{2.69\sqrt{\frac{1}{14} + \frac{1}{22}}} = 2.13$$
$$\rho(\bar{X}_1 - \bar{X}_2 < 4) = \rho(t_{24} < 2.13)$$

ملخص قوانين توزيع المعاينة للوسط الحسابي $\bar{X}$ والفرق بين وسطين $\bar{X}_1 - \bar{X}_2$

1. القوانين المتعلقة بالوسط الحسابي $\bar{X}$

$\gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ المجتمع محدود والسحب بدون إرجاع $n \geq 0.05N$	1) $\mu_{\bar{X}} = \mu_X$ 2) $\gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}}$ حالة مجتمع غير محدود والسحب بالإرجاع $n < 0.05N$ 3) توزيع المعاينة $Z = \frac{\bar{X} - \mu_{\bar{X}}}{\gamma_{\bar{X}}}$ • حجم العينة $n \geq 30$ • $n < 30$ والانحراف المعياري في المجتمع γ_X معلوم

$T=\frac{\bar{X}-\mu_{\bar{X}}}{\frac{S}{\sqrt{n}}}$ بدرجة حرية (n−1) حجم العينة$(n<30)$ والانحراف المعياري في المجتمع γ_X مجهول	

2. القوانين المتعلقة بالفرق بين وسطين $\bar{X}_1-\bar{X}_2$

1) $\mu_{\bar{X}_1-\bar{X}_2}=\mu_{\bar{X}_1}-\mu_{\bar{X}_2}=\mu_1-\mu_2$

2) $\gamma_{\bar{X}_1-\bar{X}_2}=\sqrt{\frac{\gamma_{\bar{X}_1}}{n_1}+\frac{\gamma_{\bar{X}_2}}{n_2}}$

3) توزيع المعاينة للفرق $\bar{X}_1-\bar{X}_2$ يخضع للتوزيع:

أ- حالة $\gamma_1^2=\gamma_2^2$ ومجهولين

$$T=\frac{(\bar{X}_1-\bar{X}_2)-\mu_{\bar{X}_1-\bar{X}_2}}{S_C\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

حيث:

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

توزيع T بدرجة حرية $V=n_1+n_2-2$

ب-حالة $\gamma_1^2\neq\gamma_2^2$ ومجهولين

$$T=\frac{(\bar{X}_1-\bar{X}_2)-\mu_{\bar{X}_1-\bar{X}_2}}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

يتبع توزيع T بدرجة حرية لها صيغة مركبة:

$$V=\frac{[\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}]^2}{\frac{[\frac{S_1^2}{n_1}]^2}{n_1}+\frac{[\frac{S_2^2}{n_2}]^2}{n_2}}$$

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-\mu_{\bar{X}_1-\bar{X}_2}}{\gamma_{\bar{X}_1-\bar{X}_2}}$$

شرط أن تكون γ_1، γ_2 معلومة.

رابعا: توزيع المعاينة لنسبة العينة $\bar{\rho}$

إذا كانت قيمة كل عنصر من مجتمع 0 أو 1 أي لا أو نعم أي فشل أو نجاح فإننا نسمي كل مشاهدة من هذا المجتمع "تجربة بيرنولي"، وعندما نأخذ عينة عشوائية من هذا المجتمع فإن المعاينة في هذه الحالة تكون من مجتمع بيرنولي وينصب اهتمامنا في هذه الحالة على نسبة النجاح في العينة تختلف نسبة النجاح في العينة من عينة إلى أخرى، فلو كان احتمال النجاح (الحصول على 1) في أي تجربة من مجتمع بيرنولي يساوي ρ وأخذت عينة عشوائية حجمها n من هذا المجتمع وعبرت عن عدد النجاح بالرمز X فإن نسبة النجاح في العينة تكون $\rho = \frac{X}{n}$

وإذ أخذت عينة عشوائية أخرى من هذا المجتمع فإن نسبة النجاح في العينة تتغير ولذلك فمن الواضح أن نسبة النجاح في العينة $\bar{\rho}$ متغير عشوائي وللتعرف أكثر فهم توزيع نسبة النجاح في العينة ندرس المثال التالي:

يخضع إنتاج مصنع مصابيح كهربائية للفحص الدوري فإذا وقعت أطوال حياة المصابيح ضمن حدود معينة أعتبر المصابيح صالحة للاستعمال، وإذا خرجت عن تلك الحدود اعتبرت المصابيح غير صالحة للاستعمال فإذا فحصت n من المصابيح ووجدت X منها صالحة فإن نسبة الإنتاج الصالح تكون $\rho = \frac{X}{n}$، وإذا كررت هذه التجربة عدة مرات تغيرت قيمة $\bar{\rho}$، وهذا يعني أن $\bar{\rho}$ متغير عشوائي والسؤال هنا ما هو توزيع الاحصاءة $\bar{\rho}$ وماهي خصائصه؟

للإجابة عن هذه الأسئلة قارن $\bar{\rho}$ بالوسط الحسابي للعينة ذات الحجم n المأخوذة من مجتمع قيمة كل فرد فيه إما 1 أو 0 ولتوضيح ذلك عرف المتغير العشوائي X_1 على النحو التالي:

إذا كان المصباح الذي رقمه A صالحا نضع: X_i =1

إذا كان المصباح الذي رقمه i غير صالح نضع: X_i =0

حيث $i = 1,2, \dots n$

نلاحظ $\bar{X} = \frac{\text{عدد المصابيح الصالحة}}{n}$

$$X = \frac{\sum_{i=1}^{n} X_i}{n}$$

وهذا هو $\bar{\rho}$، ومن هنا فإن نسبة النجاح $\bar{\rho}$ هي وسط حسابي لعينة ذات حجم n مأخوذة من مجتمع بيرنولي أي ذو الحدين الذي معلمتاه (1ρ،) والذي يأخذ القيمة 1 باحتمال ρ والقيمة 0 باحتمال $q = 1 - \rho$

نلاحظ أن معدل مجتمع ذو الحدين (1ρ،) هو ρ =ρ .1 وتباينه ρq 1، وأن التوقع للوسط الحسابي يساوي معدل المجتمع والتباين للوسط الحسابي يساوي تباين المجتمع مقسوما على حجم العينة، لذلك فإن توقع $\bar{\rho}$ يكون ρ وتباين $\bar{\rho}$ يكون $\frac{\rho q}{n}$ وبتطبيق نظرية النهاية المركزية يقترب توزيع $\bar{\rho}$ من التوزيع الطبيعي عندما يزداد حجم n.

نظرية:

إذا أخذت عينة عشوائية حجمها n من مجتمع إحصائي يخضع لتوزيع بيرنولي أي ذو الحدين (1ρ،) وكانت نسبة النجاح المتوفرة في العينة $\bar{\rho}$ فإن توزيع $\bar{\rho}$ يقترب من التوزيع الطبيعي الذي وسطه ρ وتباينه $\frac{\rho q}{n}$ كلما كبرت n ‏(n≥ 30) هنا عندما نقول احتمال النجاح في المجتمع ذو الحدين هوρ فإن كلمة نجاح هنا تعني حدوث الظاهرة التي اصطلحنا أن حدوثها هو نجاح فمثلا قد يكون النجاح يعني حدوث الوفاة في عملية جراحية وربما الحصول على تقدير جيد جدا، أي أن ρ تفسر على أنها نسبة الأفراد الذين لهم صفة معينة في مجتمع ما، $\bar{\rho}$ تفسر على ظانها نسبة الأفراد الذين لهم هذه الصفة في عينة أخذت من ذلك المجتمع.[1]

ملخص لأهم علاقات توزيع المعاينة للنسبة $\bar{\rho}$

$$\bar{\rho} = \frac{X}{n}$$

$$q = 1 - \rho$$

$$\mu_{\bar{\rho}} = \rho$$

عملية سحب بالإرجاع $(n \leq 0.05N)$: $\gamma_{\bar{\rho}} = \sqrt{\frac{\rho q}{n}}$

عملية السحب بدون إرجاع أو المجتمع محدود $(n \leq 0.05N)$

$$\gamma_{\bar{\rho}} = \sqrt{\frac{\rho q}{n}}\sqrt{\frac{N-n}{N-1}}$$

إذا كانت n كبيرة بدرجة كافية فإن توزيع المعاينة للنسبة يتتبع تقريبا التوزيع الطبيعي المعياري

$$Z = \frac{\bar{\rho} - \mu_{\bar{\rho}}}{\gamma_{\bar{\rho}}}$$

مثال: مصنع به 100 عامل منهم 20 يدخنون سحبت كل العينات الممكنة من العمال والتي حجمها 36 مع الإرجاع، أحسب الوسط الحسابي والخطأ المعياري لتوزيع المعاينة لنسبة المدخنين.

[1] - محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء مبادئ وتحليل استخدام spss، مرجع سابق، ص، ص: 188-189.

الحل: نسبة العمال المدخنين بالمصنع $\rho = \frac{20}{100} = 0.20$

معناه نسبة غير المدخنين $q = 1 - \rho = 1 - 0.2 = 0.8$

- الوسط الحسابي لتوزيع المعاينة $\mu_{\bar{\rho}} = \rho = 0.20$
- الخطأ المعياري لنسبة المدخنين[1]

$$\gamma_{\bar{\rho}} = \sqrt{\frac{\rho q}{n}} = \sqrt{\frac{0.20 * 0.80}{36}} = 0.067$$

مثال: أخذت عينة عشوائية حجمها 400 من مجتمع نسبة النجاح فيه ρ = 0.6، إذا كان ρ نسبة النجاح في العينة، أوجد الاحتمال $\rho(0.5 \leq \bar{\rho} \leq 0.72)$

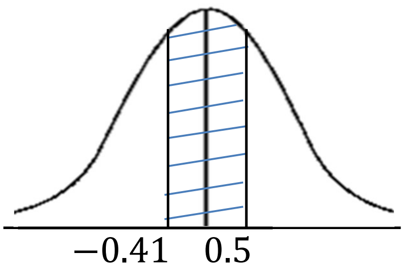
الحل:

بما أن حجم العينة كبير n> 30 فإنه يمكن تقريبا توزيع $Z = \frac{\bar{\rho}-\mu_{\bar{\rho}}}{\gamma_{\bar{\rho}}}$ بالتوزيع الطبيعي المعياري

وعليه يكون: $Z_{\bar{\rho}} = \frac{0.5-0.6}{\sqrt{\frac{0.6\ 0.4}{400}}} = -0.41$

$$Z_{\bar{\rho}} = \frac{0.72 - 0.6}{\sqrt{\frac{0.6\ 0.4}{400}}} = 0.5$$

$$\rho(0.5 \leq \rho \leq 0.72) = \rho(-0.41 \leq z \leq 0.5) = [\rho(z \leq 0.5) - 0.5] + [\rho(z \leq 0.41) - 0.5] = 0.3506$$ [2]



[1] - أحمد عودة بن عبد المجيد، منصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 156-157.

[2] - محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء مبادئ وتحليل استخدام spss، مرجع سابق، ص 190.

خامسا: توزيع المعاينة للفرق بين نسبتين $(\bar{\rho}_1 - \bar{\rho}_2)$

نظرية: إذا أخذت عينتان مستقلتان حجمهما n_1، n_2، من مجتمعين يخضع الأول لتوزيع ذي الحدين $(\rho_1 ، 1)$ والثاني لتوزيع ذي الحدين $(\rho_2 ، 1)$ فإن الفرق بين النسبتين في العينتين $\bar{\rho}_1 - \bar{\rho}_2$ يخضع تقريبا لتوزيع طبيعي وسطه $(\rho_1 - \rho_2)$ وتباينه $\frac{\rho_2 q_2}{n_2}+\frac{\rho_1 q_1}{n_1}$ وذلك عندما تكون n_1، n_2 كبيرتان.[1]

مثال: إذا كانت نسبة النجاح في الامتحان في فوج البنات ρ هي 0.7 وكانت نسبة النجاح في نفس الامتحان في فوج الذكور 0.65، أخذت عينة عشوائية حجمها $n_1 = 70$ زعينة عشوائية ثانية من الفوج الثاني $n_2 = 35$ فما هو احتمال أن تزيد نسبة النجاح في فوج الإناث عن نسبة النجاح في فوج الذكور بمقدار 0.01 على الأكثر؟

الحل: حساب الاحتمال $\rho(\bar{\rho}_1 - \bar{\rho}_2 \leq 0.01)$

$$Z_{(\bar{\rho}_1-\bar{\rho}_2)} = \frac{(\bar{\rho}_1 - \bar{\rho}_2) - \mu_{\bar{\rho}_1-\bar{\rho}_2}}{\gamma_{\bar{\rho}_1-\bar{\rho}_2}} = \frac{(\bar{\rho}_1 - \bar{\rho}_2) - \mu_{\bar{\rho}_1-\bar{\rho}_2}}{\sqrt{\frac{\rho_1 q_1}{n_1} + \frac{\rho_2 q_2}{n_2}}}$$

$$Z_{(\bar{\rho}_1-\bar{\rho}_2)} = \frac{0.01 - (0.70 - 0.65)}{\sqrt{(\frac{0.7 * 0.3}{70}) + (\frac{0.65 * 0.35}{35})}} = 0.51$$

$$\rho(\bar{\rho}_1 - \bar{\rho}_2 \leq 0.01) = \rho(Z \leq 0.51) = 0.6950$$

ملخص علاقات توزيع المعاينة للفرق بين نسبتين $\bar{\rho}_1 - \bar{\rho}_2$

$$\mu_{\bar{\rho}_1-\bar{\rho}_2} = \rho_1 - \rho_2$$

- إذا كان المجتمع غير محدود والسحب بالإرجاع و$n_1 \leq 0.005N_1$ و $n_2 \leq 0.005N_2$ فإن
$$\gamma_{\bar{\rho}_1-\bar{\rho}_2} = \sqrt{\frac{\rho_1 q_1}{n_1} + \frac{\rho_2 q_2}{n_2}}$$
- إذا كان المجتمع محدود والسحب بدون إرجاع وكان كلاهما أو أحدهما أقل (حجم المجتمع 0.5X)

[1] - محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء، مرجع سابق، ص 146.

$$\gamma_{\bar{\rho}_1-\bar{\rho}_2}=\sqrt{(\frac{\rho_1 q_1}{n_1}\frac{N_1-n_1}{N_1-1})+(\frac{\rho_2 q_2}{n_2}\frac{N_2-n_2}{N_2-1})}$$

- توزيع المعاينة للفرق بين نسبتين يتبع توزيع قريبا من التوزيع الطبيعي:

$$Z_{(\bar{\rho}_1-\bar{\rho}_2)}=\frac{(\bar{\rho}_1-\bar{\rho}_2)-\mu_{\bar{\rho}_1-\bar{\rho}_2}}{\gamma_{\bar{\rho}_1-\bar{\rho}_2}}$$

سادسا: توزيع المعاينة لتباين العينةS^2.

تباين العينةS^2 هو أفضل احصاءة يمكن استخدامها للاستدلال عن تباين المجتمع γ^2، حيث S^2 تقييس الاختلافات ومن ثم فهي تدل على التشتت أو الانتشار بين القيم في عينة عشوائية، وبما أن التباين من المقاييس الهامة مثله مثل مقاييس النزعة المركزية، أي أن أهمية S^2 للاستدلال على γ^2 تضاهي أهمية $\bar{X}$ للاستدلال فمثلا في معظم التطبيقات العملية خاصة بمجال الإنتاج يكون تخفيض الاختلافات بين وحدات المنتج النهائي ذو أهمية بالنسبة للإدارة، فمثلا في عملية تصنيع القضبان الحديدية إذا كان طول القضيب يختلف كثيرا عن الطول النمطي فإنه لن يستخدم وبالتالي يكون من المهم التأكد ليس فقط أن تكون القضبان المنتجة ذات أطوال صحيحة في المتوسط ولكن أيضا تكون الاختلافات في الأطوال صغيرة بدرجة كافية حتى تكون القضبان كلها ذات أطوال مطابقة للمواصفات. [1]

نظرية: إذا كانت$X_1,X_2\ldots\ldots X_n$ عينة عشوائية من توزيع طبيعي $N(\mu,\gamma^2)$، وكان S^2 تباين هذه العينة وكانت γ^2 معلومة فإن $X^2=\frac{(n-1)S^2}{\gamma^2}$

يخضع لتوزيع كأي تربيع على (n−1) من درجات الحرية.

مثال: إذا كانت $X_1,X_2\ldots\ldots X_{10}$ عينة عشوائية من توزيع طبيعي $N(\mu,9)$ تباينها S^2 فأوجد النقطة c بحيث: $\rho(S^2\le c)=0.95$

الحل: $\rho(S^2\le c)=\rho\left(\frac{n-1S^2}{\gamma^2}\le\frac{c(n-1)}{\gamma^2}\right)=\rho(X_9^2\le c)=0.95$

باستخدام الجدول نجد $c=16.9190$

[1] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 301.

سابعا: توزيع المعاينة للنسبة بين تبايني عينتين $\frac{s_1^2}{s_2^2}$

غالبا ما تظهر حاجتنا للمقارنة بين تباينان مجتمعين أو عمليتين فمثلا في كثير من العمليات التصنيعية يكون التركيز على اختلاف العملية الإنتاجية أكثر أهمية من التركيز على متوسط العملية، فيما يتعلق بالاستنتاج الإحصائي حول تباينات مجتمعين γ_1^2، γ_2^2 يكون من المناسب رياضيا أن نستخدم النسبة $\frac{\gamma_1^2}{\gamma_2^2}$ كأساس لهذا الاستنتاج فإذا كانت هذه النسبة تساوي الواحد فهذا يعني تساوي تباينات المجتمعين، جدير بالذكر أن أفضل إحصاءة لـ$\frac{\gamma_1^2}{\gamma_2^2}$ هو $\frac{s_1^2}{s_2^2}$ أي نسبة تباينات العينتين والاستنتاج الاحصائي المتعلق بمقارنة تباينات مجتمعين يعتمد على توزيع فيشر F، حيث توزيع فيشر F هو دالة في درجات الحرية مثل توزيع ستودنت T وتوزيع كاي تربيع X^2 لكن يختلف توزيع فيشر بأن له درجتي حرية مقترنة بتباين العينة s_1^2 في البسط و s_2^2 في المقام حيث $V_1 = n_1 - 1$ و $V_2 = n_2 - 1$ ويتحدد توزيع F تماما بدلالة درجات الحرية ولا يتوقف على أي معالم أخرى، وأي متغير عشوائي يتبع توزيع فيشر لا يمكن أن يأخذ قيما سالبة [1]

نظرية: ليكن s_1^2 تباين العينة $X_1, X_2 \ldots\ldots X_n$ التي تخضع لتوزيع طبيعي $N(\mu_{19}\gamma_1^2)$، و s_2^2 تباين العينة $y_1, y_2 \ldots\ldots y_{n_2}$ الخاضعة لتوزيع طبيعي $N(\mu_2, \gamma_2^2)$ والمستقل عن التوزيع الأول فإن:

$$\frac{\gamma_2^2}{\gamma_1^2} = \frac{\frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(y - \bar{y})^2}{\frac{1}{n_2 - 1}\sum_{i=1}^{n^2}(y - \bar{y})^2} = \frac{S_1^2}{S_2^2}\frac{\gamma_2^2}{\gamma_1^2}$$

يخضع لتوزيع فيشر F بدرجات حرية $(n_1 - 1, n_2 - 1)$

مثال: افرض أن $X_1, X_2 \ldots\ldots X_{21}$ عينة عشوائية من توزيع طبيعي $N(\mu, \gamma^2)$ وأن $y_1, y_2 \ldots\ldots y_{15}$ عينة عشوائية من توزيع طبيعي $N(\mu, \gamma^2)$ مستقلة عن التوزيع الأول أوجد العدد c بحيث أن: $\rho\left(\frac{S_1^2}{S_2^2} \geq c\right) = 0.05$

[1] - جورج كانافوس، دون ميلر، مرجع سابق، ص، ص: 403-404.

الحل: نلاحظ أن $\frac{s_1^2}{s_2^2}$ يخضع لتوزيع F على درجات الحرية (20،14) من خلال الجدول لنجد c=2.37 .[1]

[1] - محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء، مرجع سابق، ص، ص: 147-148.

ثامنا: سلسلة تمارين حول توزيع المعاينة.

- **التمرين الأول:** إذا كان X متغير عشوائي وسطه 180 وانحراف المعياري 12، إذا علمت أن عينة حجمها 420 أخذت من هذا التوزيع.

المطلوب:

- ما هو احتمال أن يزيد وسط العينة عن 185؟
- ما هو احتمال ألا يتجاوز وسط العينة 200؟

- **التمرين الثاني:** مصنع لإنتاج أكياس السميد، كانت كمية الإنتاج 28000 كيس سميد إذا علمت أن هذا المجتمع يتبع التوزيع الطبيعي بمتوسط قدره 50 كغ وتباين 16 كغ، قمنا بسحب 60 عينة حجم كل منها 20 كيس.

المطلوب:

- حدد الوسط الحسابي والانحراف المعياري لتوزيع المعاينة؟
- ما هو عدد العينات التي تتوقع أن يكون الوسط الحسابي فيها محصور بين 60 كغ و70 كغ؟
- ما هو احتمال أن يفوق متوسط العينة 55 كغ؟

- **التمرين الثالث:** إذا كان X متغير عشوائي يمثل أوزان أكياس الاسمنت في مصنع البركة، وإذا علمت أن X خاضع للتوزيع الطبيعي وسطه 50 كغ وأخذت عينة حجمها 10 أكياس من إنتاج المصنع ووجد أن انحرافها المعياري يساوي 1 كغ.

المطلوب: أحسب احتمال أن يزيد وسط العينة عن 53 كغ؟

- **التمرين الرابع:** إذا علمت أن مجتمع إحصائي يتكون من 900 فرد بوسط حسابي 20 وحدة وانحراف معياري يساوي 12 وحدة.

المطلوب:

1- أوجد الوسط الحسابي والانحراف المعياري لتوزيع المعاينة في حالة سحب عينة حجمها 36 ثم في حالة سحب عينة حجمها 64.

2- ما هو احتمال أن يقع وسط العينة في حالة سحب عينة حجمها 36 بين 18 و24؟

3- ما هو احتمال أن يزيد وسط العينة عن 22 في حالة سحب عينة حجمها 64؟

- **التمرين الخامس:** إذا كان متوسط وزن أكياس الطحين في المصنع A هو 14 كغ والانحراف المعياري هو 4، أما المصنع B فمتوسط الوزن هو 12 كغ والانحراف المعياري هو 2. إذا سحبت عينة عشوائية حجمها 100 كيس طحين من كل مصنع وتم اختبارها.

المطلوب: أوجد احتمال أن يزيد متوسط وزن أكياس الطحين في المصنع A عن متوسط وزن أكياس الطحين في المصنع B بـ2 كغ؟

- **التمرين السادس:** أخذت عينة عشوائية من طالبات كلية العلوم الاقتصادية حجمها 120 طالبة فكانت نسبة النجاح $\rho = 0.3$.

المطلوب: حدد توزيع المعاينة للنسبة، وما هي قيمة التوقع والخطأ المعياري في هذا التوزيع؟

- **التمرين السابع:** إذا كانت نسبة الطالبات في قسم علوم التسيير 0.7 ونسبة الطالبات في قسم العلوم التجارية هي 0.3 وأخذت عينة عشوائية من قسم علوم التسيير حجمها 220 طالبة، ومن قسم العلوم التجارية عينة حجمها 120 طالبة، إذا افترضنا أن $\bar{\rho}_1$ نسبة الطالبات في قسم علوم التسيير، و$\bar{\rho}_2$ نسبة الطالبات في قسم العلوم التجارية.

المطلوب: أحسب احتمال:

- $\rho(\bar{\rho}_1 - \bar{\rho}_2 \leq 0.2)$
- $\rho(\bar{\rho}_1 - \bar{\rho}_2 \geq 0.1)$

- **التمرين الثامن:** قمنا بسحب 24 طالب من كلية العلوم الاقتصادية فكان وسط معدلاتهم 14، وكان تباين هذه العينة هو 6، كما قمنا بسحب عينة حجمها 18 طالب من كلية العلوم والتكنولوجيا وسطها الحسابي 11 وتباينها 4.

المطلوب: أوجد احتمال أن يكون الفرق بين المتوسطين، أكبر من 5، وذلك مع افتراض أن تباين المجتمعين مجهولين وغير متساويين؟ ثم بافتراض أن تباين المجتمعين مجهولين ومتساويين؟

- **التمرين التاسع:** ينتج أحد المخابر الكيميائية نوعا من المحاليل يحتوي على مادة فعالة ويجب أن تكون هذه المادة محددة بشكل دقيق، ولدراسة مدى دقة المخبر في إضافة هذه الكمية لكل محلول قام المسؤول عن المخبر بتحليل عينة من 24 محلول.

المطلوب: أحسب احتمال ألا يزيد الانحراف المعياري لكمية المادة في هذه المحاليل عن 1.20 مع علما بأن الانحراف المعياري لوزن هذه المادة في إنتاج المخبر ككل يساوي 1.10 مع؟

المحور الثالث:

التقدير الإحصائي

أولا: مدخل للتقدير الإحصائي

كل شخص يقوم بعملية التقدير في حياته اليومية فمثلا عندما تريد أن تعبر الشارع فإنك تقوم بعملية تقدير سريعة لما يلي: (سرعة السيارات القادمة، المسافة بينك وبين السيارات، عرض الشارع)، وبناءا على ما سبق تقرر أحد القرارات التالية: (الانتظار، العبور مشيا، العبور جريا)، وهكذا في علم الإدارة والاقتصاد مثلا فإن العاملين في هذه المجالات يستخدمون التقدير كثيرا مثل تقدير العرض والطلب لسلعة ما، تقدير نسبة العاملين المتأخرين، نسبة تعطل الآلات، متوسط المبيعات الشهرية الانحراف المعياري لأطوال الطلبة، وللحصول على هذه التقديرات نسحب عينة عشوائية من مجتمع الدراسة ونحسب منها الاحصائيات المطلوبة ثم نستخدمها لتقدير معلمات المجتمع كتقدير متوسط المجتمع μ أو نسبة المجتمع P أو الانحراف المعياري γ.[1]

1- تعريفات مهمة

- التقدير النقطي: بالاستعانة بقيم المشاهدات من العينة نستطيع أن نبحث عن عدد بحيث يكون هذا العدد قريب من المجتمع ويمثله تمثيلا سليما وهذا العدد هو التقدير النقطي، وتعتبر المشاهدات بأنها متغيرات عشوائية أيضا وكما تعرف أيضا بأن دوال هذه المتغيرات هي إحصاءات وتقديرات للمعلمات والهدف هو إيجاد إحصاءة (تقدير) قريب جدا من قيمة المقدر (المعلمة).
- فترة الثقة: إذا كان أحد ثلاث نقاط على الأقل هي متغير عشوائي فإنه يطلق عليها فترات الثقة وبدل أن تجد تقدير نقطي لمعلمة المجتمع سنجد حدود عليا ودنيا وسنجعل هذه الحدود متغيرات عشوائية حيث أن معلمة المجتمع تقع ضمن الحدين.[2]

أي يمكن التفرقة بينهما بأن التقدير بنقطة هو تقدير معلمة المجتمع بقيمة إحصائية مقابلة لها من العينة (مثل تقدير متوسط المجتمع μ بمتوسط العينة $\bar{X}$)، ويمتاز التقدير النقطي بدقته ولكن يعاب عليه أن احتمال الخطأ فيه كبير، أما التقدير بفترة هو تقدير معلمة المجتمع بمدى من القيم مثل تقدير متوسط المجتمع μ بالفترة $\bar{X} \pm 5$ ويمتاز تقدير الفترة بأنه يضم بين حديه الأدنى والأعلى عددا غير محدود من القيم كما يمكن حساب احتمال صحة التقدير أو درجة الثقة فيه، لذلك تسمى فترات التقدير بفترات الثقة.

1- أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص 161.

2- عزام صبري، مرجع سابق، ص، ص: 255-273.

- المقدر: هو إحصائية العينة التي تستخدم لتقدير معلمة المجتمع، مثل $\bar{X}$ يعتبر مقدر لمتوسط المجتمع μ.
- التقدير: هو قيمة رقمية معينة للمقدر مثل $\bar{X} = 5$ يستخدم كتقدير لمتوسط المجتمع μ، أي أن التقدير حالة خاصة من المقدر، وبالتالي فإن التقدير قد يختلف من عينة لأخرى باستخدام المقدر نفسه.[1]

2- خواص المقدر الجيد

تستخدم القيمة المحسوبة من العينة العشوائية لتقدير القيمة المناظرة لها في المجتمع يتم تقديرها بإحصاءة العينة، ويكون المقدر (الاحصاءة) مقدرا جيدا إذا كان التوزيع الاحتمالي لهذا التقدير يتركز حول المعلمة المجهولة المراد تقديرها حيث يتميز المقدر الجيد بالخصائص الآتية وهي نفسها الخصائص التي يمكن باستخدامها معرفة (هل التوزيع الاحتمالي للتقدير يتركز حول المعلمة المراد تقديرها أم لا؟).[2]

وهذه الخصائص هي:

أ- الانحياز: لبعض المقدرات في المتوسط أخطاء معاينة صفرا بينما في المتوسط أخطاء معاينة سالبة أو موجبة يسمى متوسط خطأ المعاينة للمقدر انحياز المقدر ومن أهم يكون المقدر غير منحاز وإلا فإنه يميل إلى إعطاء تقديرات إما أكبر أو أقل من القيمة الحقيقية للمجتمع.

ومن أمثلة المقدرات الغير منحازة

- أي متوسط مرجح للمشاهدات يكون مقدرا غير منحاز لوسط المجتمع.
- تباين العينة مقدر غير منحاز لتباين المجتمع.

أما أمثلة المقدرات المنحاز

- تكون أكبر مشاهدة في العينة مقدرة منحازا لوسط المجتمع لأنها متوسط أكبر أو يساوي وسط المجتمع، ونفس الشيء مع أقل مشاهدة في العينة حيث تكون أقل أو يساوي وسط المجتمع.
- يكون مدى العينة مقدرا منحازا لمدى المجتمع لأنه دائما يكون أقل من أو يساوي مدى المجتمع.

---

[1]- أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء، 2)، مرجع سابق، ص162.

[2]- أحمد السيد عامر: الإحصاء الوصفي والتحليلي، دار الفجر للنشر والتوزيع، القاهرة، مصر، مصر، 2007، ص209.

- تكون أكبر مشاهدة في العينة مقدرا منحازا لأكبر قيمة في المجتمع لأنها دائما تكون أقل من أو تساوي أكبر قيمة في المجتمع.

كما يجب التركيز فعند حساب تباين العينة S^2 في المقام نستخدم (n−1) بدل n أي نستخدم التباين الغير منحازا ونضبطه لجعله غير منحاز.

ب- الكفاءة والاتساق

من الأفضل دائما اختيار مقدر يكون لتوزيعه أقل تباين عمليا نفضل اختيار المقدرات الغير منحازة ولها تباين صغير.

أما المقدر يكون مسبقا إذا اقترب كل من انحيازه وتباينه من الصفر مع زيادة حجم العينة أي كلما زاد حجم العينة زادت دقة التقدير فمثلا أفرض أن مكتب تعيين الجامعة وجد الموارد لأخذ عينة من 8 بدلا من 5 من الخريجين الجدد ويريد أن يستخدم هذه المقدرات فقط التي تميل إلى 3 مشاهدات إضافية منها إلى تحسين دقة التقدير، حيث يمكن أن يكون الثلاثة خريجين التاليين غير ممثلين بصورة كبيرة للمجتمع الأصلي مما يزيد خطأ المعاينة، إلا أن في المتوسط يجب أن تحسن المشاهدات الإضافية التقدير إذا كان المقدر متسقا بل أكثر إذا استخدم المقدر المتسق كل الخريجين في العينة فيجب أن تكون القيمة المقدرة لمعلمة المجتمع مساوية للقيمة الحقيقية لمعلمة المجتمع، وفيما يلي بعض الأمثلة للمقدرات المتسقة:

- يكون وسط العينة مقدرا متسقا لوسط المجتمع.
- يكون تباين العينة مقدرا متسقا لتباين المجتمع.
- يكون الانحراف المعياري للعينة مقدرا متسقا للانحراف المعياري للمجتمع.

بعد عرضنا لهذه الخواص المرغوب فيها في المقدرات عدم الانحياز والتباين الصغير والاتساق يمكن القول أنه بصفة عامة يكون وسط المجتمع والانحراف المعياري للمجتمع معلمات المجتمع الأكثر أهمية في تقديرها ويكون وسط العينة $\bar{X}$ المقدر المفضل لوسط المجتمع μ لأنه غير منحاز ومتسق وأكثر من هذا من كل المقدرات غير المنحازة لـμ، يكون التوزيع المعاينة $\bar{X}$ أقل تباين ممكن لحجم العينة الثابت ويكون الانحراف المعياري للعينة S المقدر المفضل للانحراف المعياري للمجتمع γ فهو مقدر غير متحيز ومتسق لـγ ومن كل مقدراته γ المتسقة يكون له أقل تباين يمكن لحجم العينة الثابت لذلك فلأسباب عدم الانحياز والاتساق والكفاءة يجب أن يستخدم $\bar{X}$ وS كمقدري نقطة لكل من μ و γ.

وفي الأخير وبعد عرض هذه الخواص يمكن تلخيصها في:

- الانحياز: إذا كان وسط توزيع المعاينة للمقدر يساوي قيمة معلمة المجتمع التي يحدث لها تقدير فيسمى المقدر عندئذ غير متحيز، وفي المتوسط تساوي قيمة المقدر الغير متحيز قيمة معلمة المجتمع.
- الكفاءة: بالنسبة لحجم العينة الثابت يجب ألا يكون توزيع المعاينة كثير التشتت.
- الاتساق: مع زيادة حجم العينة يقترب انحياز المقدر وتباينه من الصفر لذلك يكون من الأكثر ترجيحا للقيمة المقدرة أن تكون أقرب إلى قيمة معلمة المجتمع.[1]

ثانيا: التقدير النقطي

كما قلنا سابقا أن التقدير النقطي هو تقدير إحدى معلمات المجتمع بنقطة وذلك بإعطاء قيمة واحدة لتلك المعلمة، ومن أبسط طرق إيجاد تقدير معلمة مجهولة إذا أعطيت عينة من المشاهدات من المجتمع الأم الذي يعتمد توزيعه على هذه المعلمة ما يسمى بطريقة العزوم وتعتمد هذه الطريقة على مساواة العزم الرياضي بالعزم الإحصائي المقابل لهذه العينة وبوجه عام إذا كانت المعلمة هي عزم التوزيع فإن تقدير هذه المعلمة يكون عزم العينة المناظر.

مثال:

إذا كانت $X_1, X_2 \dots . X_n$ عينة عشوائية من مجتمع طبيعي $N(\mu, \gamma^2)$، فلإيجاد التقدير النقطي بطريقة العزوم μ, γ^2 نجد العزمين الأول والثاني للعينة $\bar{X}$ و $\frac{\sum_{i=1}^{n} X_i^2}{n}$، نجد العزمين الأول والثاني للمجتمع هما μ و $E(X^2)$ من المعادلة:

$$\gamma^2 = E(X^2) - \mu^2$$

إذا:

$$E(X^2) = \gamma^2 + \mu^2$$

والآن نضع المساواة بين كل عزمين من نفس الرتبة أي: $\mu = \bar{X}$

$$\frac{\sum X_i^2}{n} = \gamma^2 + \mu^2$$

وبحل هاتين المعادلتين نجد:

$$\hat{\mu} = \bar{X}$$

$$\hat{\gamma}^2 = \frac{\sum X_i^2}{n} - \bar{X}^2$$

[1]– إدوارد مينيكا وزوريانا كورزيجا، مرجع سابق، ص، ص: 442-446.

وهما التقديرات النقطيان بطريقة العزوم للمعلمات μ، γ^2

مثال: أخذت عينة عشوائية من مجتمع طبيعي $N(\mu,\gamma^2)$ فكانت قيمها 3، 1، 0، 4، 5، 3، 6، 4، 7، 8، 2، −3، 2، 0

أوجد تقديرا لمعدل المجتمع μ وتقدير التباين γ^2

الحل: التقدير بطريقة العزوم للمعلمة μ هو $\bar{X}$ وللتباين هو $\hat{\gamma}^2$

$$\bar{X}=\frac{3+1+0+4+5+3+6+4+7+8+2-3+2+0}{14}=\frac{42}{14}=3$$

أي تقدر μ بالقيمة 3.

وتقدير γ^2 بطريقة العزوم هو $\hat{\gamma}^2$ الذي قيمته:

$$\hat{\gamma}^2=\frac{\sum X^2-X^2}{n}$$
$$=\frac{3^2+1^2+0^2+4^2+5^2+3^2+6^2+4^2+7^2+8^2+2^2+(-3)^2+2^2+0^2}{14}$$
$$=-3^2=\frac{242}{14}-9=8.29$$

ولكن لأسباب إحصائية نستعمل S^2

$$S^2=\frac{1}{13}[(3-3)^2+(1-3)^2+(0-3)^2+(4-3)^2+(5-3)^2+(3-3)^2$$
$$+(6-3)^2+(4-3)^2+(7-3)^2+(8-3)^2+(2-3)^2$$
$$+(-3-3)^2+(2-3)^2+(0-3)^2=\frac{116}{13}=8.29$$

أي تقدر قيمة γ^2 بالقيمة 8.29.[1]

1. تقدير النسبة بنقطة

من المنطقي أن نقدر نسبة وجود ظاهرة في مجتمع ما بنسبة وجود تلك الظاهرة في عينة عشوائية تؤخذ من ذلك المجتمع، حيث إذا أردت أن تقدر نسبة العائلات التي تمتلك سيارة فبإمكانك أن تختار عينة عشوائية وتحسب نسبة العائلات التي تمتلك سيارة وتستعمل النسبة في العينة كتقدير نقطي للنسبة في المجتمع.

---

[1]– محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 206–208.

هذه التجربة تبين أنه إذا كانت نسبة النجاح في تجربة ذات الحدين P يكون بالإمكان تقدير P كما يلي:

خذ عينة عشوائية حجمها n وافرض أن عدد النجاحات في هذه العينة X يمكن استعمال $\bar{P} = \frac{X}{n}$ وهي نفس أخذ عينة عشوائية من مجتمع برنولي $b(1,P)$ ولتكن العينة $X_1, X_2 \ldots . X_n$ عندها يكون $X = \sum X_i$ هو عدد النجاحات في العينة ويكون $\bar{P} = \frac{X}{n} = \frac{\sum X_i}{n}$

X عدد النجاحات في العينة.

مثال: لتقدير نسبة المدخنين بين طلبة جامعة بسكرة قمت كباحث بمقابلة عينة عشوائية حجمها 100 طالب فوجدت أن 30 طالب يدخنون، فما هي نسبة المدخنين في جامعة بسكرة؟

الحل: نسبة المدخنين في العينة $0.3 = \bar{P} = \frac{30}{100}$ أي نقدر نسبة المدخنين في الجامعة بنسبة المدخنين في العينة وهي 0.3 (أي التقدير النقطي لنسبة P في المجتمع هو 0.3).[1]

ويمكن بعد التطرق للتقدير النقطي وعرض طريقة العزوم عرض الجدول التالي لأفضل تقدير نقطة لبعض معلمات المجتمع والخطأ المعياري للتقدير.

المعلمة المراد تقديرها		أفضل تقدير نقطة	الرمز	الخطأ المعياري للتباين
الوسط الحسابي للمجتمع	μ	الوسط الحسابي للعينة	$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$	$-\gamma_{\bar{X}} = \frac{\gamma}{\sqrt{n}}$
النسبة في المجتمع	P	النسبة في العينة	$\bar{P} = \frac{a}{n}$	$\gamma_{\bar{P}} = \sqrt{\frac{Pq}{n}}$
الفرق بين متوسطي مجتمعين	$\mu_1 - \mu_2$	الفرق بين متوسطي العينتين	$\bar{X}_1 - \bar{X}_2$	$\gamma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n}}$
الفرق بين نسبتي المجتمع	$P_1 - P_2$	الفرق بين نسبتي العينتين	$\bar{P}_1 - \bar{P}_2$	$\gamma_{\bar{P}_1 - \bar{P}_2} = \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$
تباين المجتمع	γ^2	تباين العينة S^2	$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$	$\gamma_{\bar{X}}^2 = \gamma^2 \sqrt{\frac{2}{n-1}}$

[1] – محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 222، 223.

ملاحظة: في الخطأ المعياري هناك حالات مثل توزيع المعاينة (مجتمع محدود، سحب بدون إرجاع، $n >$ $0.05N$) يجب إضافة معامل التصحيح $\sqrt{\frac{N-n}{n-1}}$.[1]

ثالثا: التقدير بفترة الثقة

بعد تطرقنا إلى التقدير بنقطة والذي نادرا ما يساوي المعلمة التي نرغب في تقديرها، لذلك فإننا نحدد فترة تحتوي على مجموعة من القيم تتضمن فيما بينها قيمة معلمة المجتمع، وتسمى هذه الفترة بتقدير فترة الثقة واحتمال وقوع المعلمة في هذه الفترة يسمى درجة الثقة، حيث نرمز لدرجة الثقة بالرمز $(1-\propto)$ ومكمل هذه القيمة يسمى مستوى المعنوية ويرمز له بالرمز $\propto$ فمثلا إذا كانت درجة الثقة 95 % فمستوى المعنوية 0.005.[2]

عند دراستنا سابقا للتوزيع الطبيعي وجدنا أن معظم البيانات تتركز حول متوسط التوزيع μ أي وجدنا احتمال تركز البيانات كما يلي:

- $P(\mu-\gamma \leq X \leq \mu+\gamma) = 0.6827 = 68.27\ \%$
- $P(\mu-2\gamma \leq X \leq \mu+2\gamma) = 0.9545 = 95.45\ \%$
- $P(\mu-3\gamma \leq X \leq \mu+3\gamma) = 0.9973 = 99.73\ \%$

أي أننا على ثقة مقدارها 95.45% أن قيمة المتغير X حيث $X \sim N(\mu,\gamma)$ تقع في الفترة $(\mu-2\gamma, \mu+2\gamma)$

ونجد أيضا فترات الثقة للمتغير العشوائي X:

- $P(\mu-1.64\gamma \leq X \leq \mu+1.64\gamma) = 0.90 = 90\ \%$
- $P(\mu-1.96\gamma \leq X \leq \mu+1.96\gamma) = 0.95 = 95\ \%$
- $P(\mu-2.58\gamma \leq X \leq \mu+2.58\gamma) = 0.99 = 99\ \%$

نظرية:

- إذا كان لدينا متغير عشوائي X حيث $X \sim N(\mu,\gamma)$ فإنه يكفي معرفة قيمة μ,γ للحصول على فترة ثقة للمتغير العشوائي X بالشكل التالي:

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص: 163.

[2]– قسم الإحصاء بجامعة الملك عبد العزيز، مرجع سابق، ص: 252.

$$P\left(\mu - Z_{\frac{\alpha}{2}}.\gamma \leq X \leq \mu + Z_{\frac{\alpha}{2}}.\gamma\right) = 1-\alpha$$

- معامل γ في فترة الثقة هو الذي يحدد درجة الثقة ولذلك يسمى معامل الثقة، وفيما يلي أشهر درجات الثقة ومعاملاتها للتوزيع الطبيعي.

درجة الثقة $1-\alpha$	معامل الثقة $Z_{\frac{\alpha}{2}}$
90 %	1.64
95 %	1.96
99 %	2.58

وعند استخدام درجات ثقة أخرى نقوم باستخراج معامل الثقة من جدول التوزيع الطبيعي.[1]

1. <u>أهم المصطلحات المقترنة بفترات الثقة</u>

- فترة الثقة: تتكون من فترة تقدير للمعلمة مصحوبة بمستوى الثقة يبين أن تلك الفترة تحتوي على قيمة المعلمة.
- حدود الثقة: هي الحدود العليا والدنيا لفترة الثقة.
- مستوى الثقة: هو التكرار النسبي في المدى الطويل وبه تعطى العينات العشوائية فترة الثقة تحتوي على قيمة المعلمة المطلوب تقديرها.
- هامش خطأ المعاينة: يصف الدقة في المقدر وهو أقصى كمية من الخطأ يمكن أن تحدث في المقدر عند مستوى ثقة معين، واتساع فترة الثقة يضاعف الهامش في خطأ المعاينة.
- القيم الجزئية: يتم اختيارها من توزيع المعاينة لتحقيق مستوى الثقة المطلوب.

ويمكن تلخيص الهيكل العام لفترات الثقة كما يلي: [2]

الهيكل العام لفترات الثقة

التقدير بنقطة $\pm$ (قيمة جزئية $\times$ الخطأ المعياري للتقدير بنقطة)

التقدير بنقطة $\pm$ هامش خطأ المعاينة.

---

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 163–164.

[2]– جورج كانافوس ودون ميلر، مرجع سابق، ص: 322.

2. تقدير فترة الثقة لـ μ باستخدام التوزيع الطبيعي: Z

نفرض أن المعلمة المجهولة التي نريد تقديرها هي الوسط الحسابي للمجتمع (μ) فنسحب لهذا الغرض عينة عشوائية بسيطة، ونفترض أيضا أن المجتمع موزع حسب القانون الطبيعي بوسط قدره (μ) وانحراف معياري قدره (γ)، فإن مجال الثقة للوسط يكتب على النحو التالي:

$$\mu \in [\bar{X} - Z_{1-\frac{\alpha}{2}}.\gamma_{\bar{X}} \; ; \; \bar{X} + Z_{1-\frac{\alpha}{2}}.\gamma_{\bar{X}}]$$

أو:

$$\mu \in [\bar{X} \pm Z_{1-\frac{\alpha}{2}}.\gamma_{\bar{X}}]$$

حيث $Z_{1-\frac{\alpha}{2}}$ القيمة الجدولية الموجودة في جدول القانون الطبيعي المعياري والمناسبة لاحتمال قدر $(1-\alpha)$ مع العلم:

$$\gamma_{\bar{X}} = \frac{\gamma_X}{\sqrt{n}} \quad \text{أو} \quad \gamma_{\bar{X}} = \frac{\gamma_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

ملاحظات مهمة:

أ- حدد مجال الثقة على أساس أن المجتمع يتبع القانون الطبيعي ويجب مراعاة ذلك.

ب-إن طول مجال الثقة يكبر كلما كبرت قيمة الانحراف المعياري أو درجة الثقة أو الإثنين معا وأيضا كلما صغر حجم العينة.

ج- التعليق عن النتيجة إن المجال بمقدار $\%(1-\alpha)$ ثقة يعني أننا أخذنا عينات عشوائية متكررة من مجتمع وأن 95 % من مجالات الثقة هذه تحتوي على الوسط الحقيقي (μ) المجهول وبما أن المجال الذي تم حسابه هو إحدى هذه المجالات فإننا نقبل بمخاطرة قدرها α % بأننا على خطأ.[1]

[1]- موساوي عبد النور وبركان يوسف: الإحصاء (الجزء 2)، دار العلوم للنشر والتوزيع، عنابة الجزائر، 2010، ص، ص: 131-132.

أي إذا كان لدينا مجتمع إحصائي طبيعي، وسحبنا منه عينة عشوائية حجمها n فإن توزيع المعاينة لمتوسط العينة $\bar{X}$ هو $X \sim N(\mu, \gamma)$، نجد:

$$P\left(\mu - Z_{\frac{\alpha}{2}}.\gamma \leq \bar{X} \leq \mu + Z_{\frac{\alpha}{2}}.\gamma\right) = 1-\alpha$$

بإضافة $(-\mu, \gamma_{\bar{X}})$ للطرفين نجد أن:

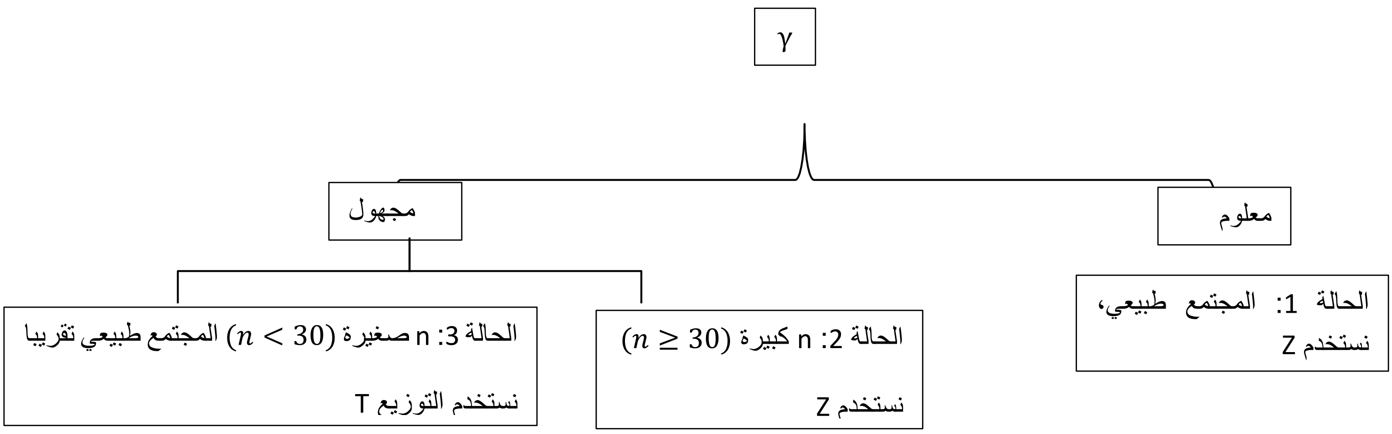
$$P(-\bar{X} - Z_{\frac{\alpha}{2}}.\gamma_{\bar{X}} \leq -\mu \leq -\bar{X} + Z_{\frac{\alpha}{2}}.\gamma_{\bar{X}}) = 1-\alpha$$

وبضرب الطرفين في (-1) وإعادة الترتيب نجد أن:

$$P(\bar{X} - Z_{\frac{\alpha}{2}}.\gamma_{\bar{X}} \leq \mu \leq \bar{X} + Z_{\frac{\alpha}{2}}.\gamma_{\bar{X}}) = 1-\alpha$$

وتمثل فترة الثقة لمتوسط المجتمع μ بدرجة $(1-\alpha)\%$ ويمكن كتابتها بالصيغة المختصرة: $\mu = \bar{X} \pm Z_{\frac{\alpha}{2}}.\gamma_{\bar{X}}$

نلاحظ أن فترة الثقة تعتمد على توزيع المعاينة الإحصائية $\bar{X}$ وهو بدوره يعتمد على الانحراف المعياري للمجتمع وحجم العينة العشوائية n كما هو موضح في الشكل:



وسنتطرق فيما يلي للحالة الأولى والثانية لأنها تتبع التوزيع الطبيعي بينما الحالة الثالثة سنتطرق لها في العنصر الموالي:

- الحالة الأولى: المجتمع طبيعي، γ معلوم، n كبيرة أو صغيرة في هذه الحالة X يتبع التوزيع الطبيعي بمتوسط μ وانحراف معياري $\gamma_{\bar{X}}$ أي أن: $X \sim N(\mu, \gamma_{\bar{X}})$

وبالتالي فترة الثقة لمتوسط المجتمع تكون: $\mu = \bar{X} \pm Z_{\frac{\alpha}{2}} . \gamma_{\bar{X}}$

- الحالة الثانية: γ مجهول و n كبيرة $(n \geq 30)$ بتطبيق نظرية النهاية المركزية حيث n كبيرة فإن $\bar{X}$ يتبع التوزيع الطبيعي وبالتالي تكون فترة الثقة لـ μ بدرجة ثقة $(1-\alpha)\%$ بالشكل التالي:

$$\mu = \bar{X} \pm Z_{\frac{\alpha}{2}} . \gamma_{\bar{X}}$$

حيث: γ مجهول و S هو الانحراف المعياري للعينة أي:

$$\gamma_{\bar{X}} = \frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \quad \text{أو} \quad \gamma_{\bar{X}} = \frac{S}{\sqrt{n}}.$$ [1]

من خلال ما سبق نلاحظ أن هناك أربع صفات أساسية تميز فترة الثقة:

- تكون $\bar{X}$ نقطة المنتصف لفترة ثقة الطرفين لكل مستوى ثقة.
- كلما ازداد مستوى الثقة $(1-\alpha)$ تزداد $Z_{\frac{\alpha}{2}}$ وبالتالي يزداد عرض فترة الثقة وتكون الفترة التي لها معامل ثقة أصغر محتواه داخل الفترة لها مستوى ثقة أكبر وكما يمكن أن نتوقع يجب أن تكون فترة الثقة 95 % محتواه داخل فترة الثقة 99 %.
- مع زيادة γ يقل عرض فترة الثقة.
- مع زيادة n يقل عرض فترة الثقة، كما يمكن أن نتوقع إذا سحبت عينة كبيرة جدا فيجب أن نكون قادرين على إيجاد وسط المجتمع في حدود فترة صغيرة وفي الأخير يمكن تلخيص كيف تغير حساسية فترة ثقة الطرفين قيم إحصائيات العينة:

أ- وسط العينة: على افتراض وجود خطأ في القيمة المحسوبة لوسط العينة، نظرا لأن $\bar{X}$ هي نقطة المنتصف لفترة ثقة الطرفين فيكون تأثير التصحيح في قيمة وسط العينة عبارة عن تحويل نقطة منتصف الفترة إلى القيمة الصحيحة فقط ويظل طول فترة الثقة كما هو دون تغيير.

مثال: على افتراض أن فترة ثقة الطرفين 95 % لوسط المجتمع كانت [50,80] يمكننا استنتاج قيمة وسط العينة التي استخلصت منها هذه الفترة تساوي 66 لأنها منتصف الفترة، فإذا اكتشفنا أن وسط العينة

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 163-164.

قد حسب خطأ وكانت قيمته الفعلية 67 بدل 66 فتصبح فترة الثقة % 95 الناتجة [53,81] أي أن الطول لم يتغير.

ب- الانحراف المعياري للعينة: على افتراض اكتشاف خطأ في الانحراف المعياري للعينة، إذا كانت القيمة الصحيحة للانحراف المعياري أكبر من القيمة الاصلية فيزداد على ذلك طول فترة الثقة بالتناسب والعكس إذا كانت القيمة الصحيحة أصغر.

مثال: بافتراض أن قيمة S الصحيحة كانت أكبر بـ 25% لكن $\bar{X}$ بقى كما هو 66، أي تزداد لذلك الكمية بنسبة 25% وتصبح النهاية العليا الصحيحة لفترة الثقة:

$$66+(1.25)(80-66)=66+17.5=83.5$$

وبالتالي تصبح نقطة النهاية الأقل الصحيحة: $66-17.5=48.5$

أي فترة الثقة الصحيحة [48.5,83.5] أكبر بنسبة 25% من فترة الثقة [52,80]

ج- معامل الثقة $(1-\propto)$: على افتراض أن اختيار قيمة معامل الثقة $(1-\propto)$ انخفض، أو ما يكافئ هذا أن قيمة $\propto$ الجديدة أكبر من قيمة $\propto$ الأصلية في هذه الحالة تصبح فترة الثقة الجديدة فئة فرعية مناسبة من فترات الثقة الأصلية يمكن رؤية بملاحظة أن مع زيادة $\propto$ تقل قيمة $t_{(n-1,\frac{\propto}{2})}$ (لأي عدد درجات الحرية)، وبالتالي يقل حاصل الضرب أيضا $t_{(n-1,\frac{\propto}{2})}.\gamma_{\bar{X}}$ أيضا مما يتسبب في احتواء فترة الثقة الجديدة داخل فترة الثقة الأصلية.

د- حجم العينة n: على افتراض أننا اكتشفنا أن حجم العينة n لم يكن صحيحا وأن قيمة n الصحيحة أكبر من n الاصلية في هذه الحالة ستكون فترة الثقة الصحيحة محتواة في فترة ثقة الطرفين الأصلية ويمكن رؤية هذا عن طريق مشاهدة أنه إذا ازدادت n فإن $\gamma_{\bar{X}}$ يقل وإذا ازدادت n يقل $t_{(n-1,\frac{\propto}{2})}$ وينتج عن هذا أن حاصل الضرب $t_{(n-1,\frac{\propto}{2})}.\gamma_{\bar{X}}$ يقل أيضا، لذلك إذا ازدادت n تصبح فترة الثقة المصححة أقل ويمكن أن تصبح محتواة في فترة الثقة الأصلية ويكون لهذا معنى

فقط لأن حجم العينة الأكبر يشمل أن لدينا معلومات اكبر وبناءا على ذلك نكون قادرين على تضييق فترة الثقة.[1]

ويمكن تقديم صيغة حجم العينة n الذي يناظر هامش خطأ المعاينة مرغوب فيه أو بمعنى مكافئ يناظر فترة ثقة للمتوسط μ ذات اتساع أو مدى معين لتحقيق فترة ثقة $(1-\propto)\ 100\%$ للمتوسط μ عند اتساع أو مدى معين قدره 2E حيث E هامش خطأ المعاينة المرغوب فيه فإننا نحدد حجم العينة باستخدام الصيغة التالية:

$$n = [\frac{Z_{1-\frac{\propto}{2}} . \gamma_{\bar{X}}}{E}]^2$$ [2]

أمثلة حول تقدير فترة الثقة لـ$\bar{X}$ باستخدام التوزيع الطبيعي

المثال الأول: يمكن اعتبار الإنتاج اليومي للدقيق في إحدى المطاحن كمتغير عشوائي يتبع القانون الطبيعي بانحراف معياري قدره 1000كغ تم القياس بطريقة عشوائية ومستقلة إنتاج 25 يوما وقدرت الكمية المتوسطة ب 4290كغ.

المطلوب:

1- قدر بنقطة الوسط الحسابي.

2- إيجاد مجال الثقة بمستوى ثقة قدره 95%.

3- إيجاد مجال الثقة عندما تكون المخاطرة قدرها 1%.

الحل:

[1]- إدوارد مينيكاو زوريانا كوزريجا، مرجع سابق، ص، ص: 492-497.

[2]- جورج كانافوس ودون ميلر، مرجع سابق، ص 226.

1- التقدير بنقطة للوسط الحسابي:

$\mu_{\bar{X}} = \mu = 4290kg$ و $\gamma_{\bar{X}} = \frac{\gamma_X}{\sqrt{n}} = \frac{1000}{\sqrt{25}} = 200kg$

2- إيجاد مجال الثقة بمستوى ثقة 95%:

$$1-\propto = 0.95 \Rightarrow 1-\frac{\propto}{2} = 0.975 \Rightarrow Z_{0.975} = 1.96$$

مجال الثقة:

$$\mu \in [\bar{X} - Z_{1-\frac{\propto}{2}}\frac{\gamma}{\sqrt{n}}\ ;\ \bar{X} + Z_{1-\frac{\propto}{2}}\frac{\gamma}{\sqrt{n}}]$$

$$\mu \in [4200 - (1.96)(200)\ ;\ 4200 + (1.96)(200)]$$

$$\mu \in [3900\ ;\ 4680]$$

أو يحسب: $\bar{X} \pm Z_{1-\frac{\propto}{2}}\frac{\gamma}{\sqrt{n}} = 4290 \pm 390kg$

3- إيجاد مجال الثقة عندما تكون المخاطرة قدرها 1% $(\propto = 1\%)$

$$\propto = 0.01 \Rightarrow 1-\frac{\propto}{2} = 0.995 \Rightarrow Z_{0.995} = 2.58$$

$$\bar{X} \pm Z_{0.995}\left(\frac{\gamma}{\sqrt{n}}\right) = 4290 \pm 2.58(200) = 4290 \pm 596$$

$\mu \in [3774\ ;\ 4806]$ [1]

[1]- موساوي عبد النور وبركان يوسف، مرجع سابق، ص، ص: 132-133.

المثال الثاني: أخذت عينة عشوائية حجمها 144 بوسط مقداره 100 وانحراف معياري مقداره 60 من مجتمع حجمها 1000، فترة 95% ثقة لوسط المجتمع غير المعلوم.

$$Z_{1-\frac{\propto}{2}} = 1.96\ (1-\propto= 0.95 \Rightarrow 1-\frac{\propto}{2} = 0.975)$$

بما أن $n > 30$ حيث:

$$\mu \in \bar{X} \pm 1.96\gamma_{\bar{X}}$$

بما أن $N > 0.05N$ نجد: $\mu \in \bar{X} \pm 1.96\frac{\gamma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

وباستخدام S كتقدير γ:

$$100 \pm 1.96\frac{6}{\sqrt{144}}\sqrt{\frac{1000-100}{1000-1}}$$

$$100 \pm 1.96(5)(0.95)$$

$$100 \pm 9.31$$

$$\mu \in [90.89\ ;\ 109,11]$$

المثال الثالث: يرغب مدير في تقدير متوسط الدقائق التي يأخذها العمال لإنجاز عملية صناعية في حدود (± 3) دقيقة وبدرجة ثقة 90% ويعلم المدير من خبرته الماضية أن الانحراف المعياري γ هو 15 دقيقة والحد الأدنى لحجم العينة المطلوب $n > 30$ يمكن إيجاده كالآتي: $Z = \frac{\bar{X}-\mu}{\gamma_{\bar{X}}} \Rightarrow Z\gamma_{\bar{X}} = \bar{X} - \mu$

بما أن خطأ التقدير $\bar{X} - \mu = 3$

$$1.64\frac{\gamma}{\sqrt{n}} = \bar{X} - \mu$$

$$1.64\frac{15}{\sqrt{n}} = 3$$

$$1.64\frac{15}{3} = \sqrt{n} \Longrightarrow n = 67.24$$

بالتقريب n=68.[1]

المثال الرابع: أراد عميد إحدى كليات أن يحدد حجم العينة اللازمة لحساب متوسط الزمن اللازم لانتقال الطلبة من مدرج إلى آخر وذلك ألا يزيد الخطأ في هذا المتوسط عن 0.3 دقيقة وهذا بدرجة ثقة 95% على أساس أن الانحراف المعياري γ من دراسات مماثلة يساوي 1.5 دقيقة.

الحل: لتحديد الحجم المناسب للعينة نطبق العلاقة التالية:

$$n \geq [\frac{Z_{\frac{\propto}{2}}}{E}]^2 \Leftrightarrow n \geq [\frac{1.96 \times 1.5}{0.3}]^2$$

$$n \geq 96.04$$

$n \approx 96 \approx$ مفردة

المثال الخامس: قامت مجموعة من الخبراء بتقدير متوسط الزمن اللازم لتركيب آلة معينة في صناعة كبيرة وذلك استنادا على عينة عشوائية حجمها 150 ومن المعلومات المتوفرة لدى تلك المجموعة أن الانحراف المعياري γ يساوي 6.2 دقيقة.

[1]– دومينيك سالفاتور، مرجع سابق، ص، ص: 68–69.

المطلوب: أحسب الحد الأقصى للخطأ لهذا التقدير بدرجة ثقة 99%.

الحل:

$$\text{الخطأ} = Z_{\frac{\propto}{2}}.\frac{\gamma}{\sqrt{n}} = 2.575\frac{6.2}{\sqrt{150}} = 1.30$$

وبذلك يمكن القول بأن الخطأ الأقصى هو 1.3 بدرجة ثقة 99%.[1]

3. <u>تقدير فترة الثقة لـ μ باستخدام التوزيع T ‏(γ مجهولة و $n < 30$)</u>

في معظم التطبيقات الإحصائية عمليا تكون قيمة γ مجهولة بفرض أننا مهتمين بتقدير فترة ثقة للمتوسط μ وبعينة عشوائية n من المشاهدات من مجتمع توزيعه قريبا من التوزيع الطبيعي وبانحراف معياري γ غير معلوم وبما أن توزيع المعاينة للإحصاءة:

$$T = \frac{\bar{X}-\mu}{\gamma_{\bar{X}}} \qquad \gamma_{\bar{X}} = \frac{S}{\sqrt{n}} \text{ أو } \gamma_{\bar{X}} = \frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

هو توزيع T بدرجات حرية $(n-1)$، طريقة تحديد فترة الثقة لـμ عندما تكون γ مجهولة هي صورة مطابقة للطريقة المتبعة عندما تكون γ معلومة، وذلك بإحلال قيم T محل قيم Z والخطأ المعياري لـ $\bar{X}$ يكون بـS بدل من γ المجهولة، بصفة عامة فترة الثقة$(1-\propto)\%$ 100 للمتوسط μ عندما تكون γ مجهولة هي:

$$\epsilon\mu\bar{X} \pm T_{1-\frac{\propto}{2}, n-1} \cdot \gamma_{\bar{X}}$$

[1]– محمد قريشي: الاحتمالات والإحصاء التطبيقي، الطبعة 2، دار علي بن زيد للطباعة والنشر، بسكرة، الجزائر، 2020، ص، ص: 111-112.

حيث: هامش خطأ المعاينة هو: $T_{1-\frac{\alpha}{2},n-1}.\gamma_{\bar{X}}$

$\pm T_{1-\frac{\alpha}{2},n-1}$ القيم الجزئية المناظرة لتوزيع T بدرجة حرية (n-1).[1]

المثال الأول: أخذت عينة عشوائية حجمها 20 من مجتمع موزع $N\sim(\mu,\gamma)$، وأعطيت النتائج التالية $\bar{X} = 17 \quad S^2 = 36$

المطلوب: قدر بمستوى ثقة 90% متوسط هذا المجتمع.

الحل: $1-\frac{\alpha}{2} = 0.95 \Rightarrow 1-\alpha = 90\% \Rightarrow \alpha = 0.10 \Rightarrow \frac{\alpha}{2} = 0.05$

$$n-1 = 20-1 = 19 \Rightarrow T_{0.95,19} = 1.729$$

مجال الثقة: $\bar{X} \pm T_{1-\frac{\alpha}{2},n-1}\left(\frac{S}{\sqrt{n}}\right) = 17 \pm (1.729)\left(\frac{6}{\sqrt{20}}\right) = 17 \pm 2.32$

$\epsilon\mu[14.68,19,32]$.[2]

4. فترة الثقة للفرق بين متوسطي مجتمعين $(\mu_1-\mu_2)$

1.4 عندما يكون تباين المجتمعين γ_1^2 و γ_2^2 معلومين

نحتاج في بعض الدراسات الإحصائية إلى مقارنة متوسطين حسابين لمجتمعين ومعرفة الفرق بينهما $(\mu_1 - \mu_2)$، فإذا كان المجتمع الأول يتوزع توزيعا طبيعيا بمتوسط μ_1 وتباين γ_1^2 معلوم، والمجتمع الثاني يتوزع توزيعا طبيعيا بمتوسط μ_2 وتباين γ_2^2 معلوم أيضا، وسحبنا من المجتمع الأول عينة عشوائية حجمها n_1

[1]- جورج كانافوس ودون ميلر، مرجع سابق، ص327.

[2]- موساوي عبد النور وبركان يوسف، مرجع سابق، ص 143.

ومن المجتمع الثاني عينة عشوائية حجمها n_2 وكانت عينتين مستقلتين فقد علمنا سابقا أن توزيع المعاينة للفرق بين متوسطي العينتين $(\bar{X}_1 - \bar{X}_2)$ توزيعا طبيعيا بمتوسط وتباين قدرهما على التوالي:

$$\gamma^2_{\bar{X}_1-\bar{X}_2} = \frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2} \quad .\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2$$

وبالتالي فإن المتغير العشوائي Z حيث:

$$Z = \frac{[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)]}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

هو متغير عشوائي وبالتالي نصل إلى النظرية التالية.

نظرية: إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 تم سحبها من مجتمع طبيعي متوسطه μ_1 وتباينه γ_1^2 معلوم، وكان $\bar{X}_2$ هو الوسط الحسابي لعينة عشوائية حجمها n_2 مستقلة عن الأولى تم سحبها من مجتمع طبيعي أخرى متوسطه μ_2 وتباينه γ_2^2 معلوم فإن فترة الثقة $100\%(1-\propto)$ للفرق بين متوسطي المجتمعين $(\mu_1 - \mu_2)$ بالعلاقة التالية:

$$\left[(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\propto}{2}}\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}} \text{ و } (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\propto}{2}}\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}\right].$$ [1]

<u>2.4 عندما يكون تباين المجتمعين γ_1^2، γ_2^2 مجهولين و n_1، n_2 كبيرة $(n_2 \geq 30\ ;\ n_1 \geq 30)$</u>

نستخدم فترة الثقة كالسابقة لكن باستخدام S_1^2، S_2^2 بدلا γ_1^2، γ_2^2 المجهولين، وتعطى فترة الثقة بالعلاقة التالية:

---

[1]– محمد قريشي، مرجع سابق، ص، ص: 177-118.

$$\left[(\bar{X}_1-\bar{X}_2)-Z_{\frac{\propto}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\ \text{و}\ (\bar{X}_1-\bar{X}_2)+Z_{\frac{\propto}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right]$$

مثال:

تملك إحدى الشركات مصنعين للمرطبات في ولايات بسكرة وباتنة وتحاول الشركة دوما أن تحافظ على مستوى متماثل للإنتاج في المصنعين وحتى تعرف مدى تماثل المصنعين في الإنتاج سحبت عينة عشوائية حجمها 36 علبة من كل مصنع بسكرة ومصنع باتنة فوجد أن متوسط محتوى العلبة في المصنعين هو: 235مل و322مل بانحراف معياري 4مل و3مل على التوالي:

أوجد فترة ثقة للفرق بين متوسطي المجتمعين بدرجة ثقة 88% في الحالتين التاليتين:

أ- $(\gamma_1=3\ \text{و}\ \gamma_2=\sqrt{7})$

ب- $(\gamma_1,\gamma_2$ مجهولان$)$

الحل:

أ. من جدول التوزيع الطبيعي نجد أن $Z_{\frac{\propto}{2}}$ والتي تقابل درجة الحرية 88% هي 1.55 وبالتالي:

$$\mu_1-\mu_2=(\bar{X}_1-\bar{X}_2)\pm Z_{\frac{\propto}{2}}\sqrt{\frac{\gamma_1^2}{n_1}+\frac{\gamma_2^2}{n_2}}=(325-322)\pm 1.55\sqrt{\frac{9}{36}+\frac{7}{36}}$$

$$=3\pm 1$$

$$\mu_1-\mu_2\epsilon[2,4]$$

نحن على ثقة مقدارها 88% أن متوسط محتوى علبة مصنع بسكرة يزيد عن متوسط علبة مصنع باتنة بما لا يقل عن 2مل ولا يزيد عن 4مل.

ب. بما أن العينتين كبيرتين $(n_2 \geq 30\ ;\ n_1 \geq 30)$ و (γ_2, γ_1) مجهولين فإننا سنستخدم S_1^2 و S_2^2 بدل من γ_1^2 و γ_2^2 وتكون فترة الثقة كما يلي:

$$\mu_1 - \mu_2 = (\bar{X}_1 - \bar{X}_2) \pm Z_{\frac{\propto}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$\mu_1 - \mu_2 = 3 \pm 1.55 \sqrt{\frac{16}{36} + \frac{9}{36}}$$

$$(\mu_1 - \mu_2) \epsilon [1.7,4.3]$$

نحن على ثقة مقدارها 88% أن متوسط محتو علبة بسكرة يزيد عن متوسط علبة مصنع باتنة بما لا يقل على 1.7مل ولا يزيد عن 4.3مل.[1]

<u>3.4 عندما يكون تبايني المجتمعين مجهولين ومتساويين والعينتين مستقلتين وصغيرتين:</u>

إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 سحبت من مجتمع طبيعي وسطه μ_1 وتباينه γ_1^2، وكان $\bar{X}_2$ الوسط الحسابي لعينة عشوائية حجمها n_2 سحبت من مجتمع طبيعي آخر مستقل عن الأول وسطه μ_2 وتباينه γ_2^2 وكان حجم العينتين صغير وتباين المجتمعين مجهولين ومتساويين فإن فترة الثقة $(1-\propto)$ 100% للفرق بين متوسطي المجتمعين $(\mu_1 - \mu_2)$ تعطى بالعلاقة:

$$(\bar{X}_1 - \bar{X}_2) - T_{\left(\frac{\propto}{2}, V\right)} \sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \text{ و } (\bar{X}_1 - \bar{X}_2) + T_{\left(\frac{\propto}{2}, V\right)} \sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

---

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 170–172.

حيث درجة الحرية: $V = n_1 + n_2 - 2$

مثال: استخدمت طريقتين لإنتاج سلعة معينة وتوضيح البيانات التالية الوقت بالدقائق المستغرق في إنتاج الوحدة بالنسبة لـ6 وحدات انتجت بالطريقة الأولى، و5 وحدات أنتجت بالطريقة الثانية، حيث X_1 يمثل الوقت المستغرق عند استعمال الطريقة الأولى، وX_2 يمثل الوقت المستغرق عند استعمال الطريقة الثانية:

X_1	40	40	50	60	60	50
X_2	40	45	55	58	62	/

فإذا علمت أن الوقت المستغرق في الإنتاج بالطريقتين يتوزع توزيعا طبيعيا بتباينين متساويين، أوجد فترة ثقة 95% للفرق بين متوسطي الوقت المستغرق بالطريقتين $(\mu_1 - \mu_2)$

الحل: بما أن الوقت المستغرق في الإنتاج بالطريقتين يتوزع توزيعا طبيعيا بتباينين مجهولين ومتساويين والعينتين صغيرتين ومستقلتان لأن كل عينة خاصة بآلة مستقلة عن آلة أخرى، إذا فترة الثقة المناسبة في هذه الحالة:

$$\left[(\bar{X}_1 - \bar{X}_2) - T_{\left(\frac{\propto}{2}, V\right)}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \text{ و } (\bar{X}_1 - \bar{X}_2) + T_{\left(\frac{\propto}{2}, V\right)}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right]$$

لإيجاد هذه الفترة يجب حساب $\bar{X}_1$، $\bar{X}_2$، S_1^2، S_2^2، S_p^2

$$\bar{X}_1 = \frac{\sum X_{1i}}{n_1} = \frac{300}{6} = 50$$

$$\bar{X}_2 = \frac{\sum X_{2i}}{n_2} = \frac{260}{5} = 52$$

$$S_1^2 = \frac{\sum(X_{1i} - \bar{X}_1)^2}{n-1} = \frac{400}{5} = 80$$

$$S_2^2 = \frac{\sum(X_{2i} - \bar{X}_2)^2}{n_2-1} = \frac{338}{4} = 84.5$$

$$S_p^2 = \frac{(n-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{(6-1)+(5-1)84.5}{6+5-2} = 82$$

وبما أن: $1-\propto = 0.95$ فإن $\frac{\propto}{2} = 0.025$

ولدينا: $V = n_1 + n_2 - 2 = 6 + 5 - 2 = 9$

$$t_{(\frac{\propto}{2},V)} = t_{(0.025,9)} = 2.262$$

ومنه فترة الثقة:

$$(\bar{X}_1 - \bar{X}_2) \pm t_{(\frac{\propto}{2},V)}\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = (50-52) \pm 2.262\sqrt{82\left(\frac{1}{6}+\frac{1}{5}\right)}$$

$$= [10.40 \text{ و } -14.40]$$

نحن نثق بدرجة ثقة 95% بأن الفرق الحقيقي بين متوسط الوقت المستغرق لإنتاج الوحدة بالطريقة الأولى ومتوسط الوقت المستغرق لإنتاج الوحدة بالطريقة الثانية يقع بين القيمتين −14.40 دقيقة و10.40 دقيقة.

4.4 عندما يكون تبايني المجتمعين مجهولين وغير متساويين والعينتين مستقلتين وصغيرتين:

إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 سحبت من مجتمع طبيعي وسطه μ_1 وتباينه γ_1^2، وكان $\bar{X}_2$ هو الوسط الحسابي لعينة عشوائية حجمها n_2 من مجتمع طبيعي آخر مستقل عن

المجتمع الأول وسطه μ_2 وتباينه γ_2^2 وكان حجم العينتين صغير وتبايني المجتمعين مجهولين وغير متساويين فإن فترة الثقة $100\%(1-\propto)$ للفرق بين متوسطي المجتمعين $(\mu_1 - \mu_2)$ تعطى العلاقة:

$$\left[(\bar{X}_1 - \bar{X}_2) - T_{\left(\frac{\propto}{2},V\right)}\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \text{ و } (\bar{X}_1 - \bar{X}_2) + T_{\left(\frac{\propto}{2},V\right)}\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right]$$

حيث V درجة الحرية صيغتها مركبة كما يلي:

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

مثال: البيانات التالية خاصة بعينتين مستقلتين مسحوبتين من مجتمعين يتوزعان طبيعيا بتباينين غير متساويين:

العينة الأولى	العينة الثانية
$n_1 = 4$	$n_2 = 5$
$\bar{X}_1 = 1$	$\bar{X}_2 = 1.2$
$S_1^2 = 0.0547$	$S_2^2 = 0.0991$

المطلوب: أوجد فترة ثقة 90% للفرق بين متوسطي المجتمعين.

الحل: بما أن المجتمعين يتوزعان طبيعيا بتباينين غير متساويين والعينتان مستقلتان فإن فترة الثقة المطلوبة هي: $(\bar{X}_1 - \bar{X}_2) \pm T_{\left(\frac{\propto}{2},V\right)}\sqrt{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)}$

لدينا: $\frac{\propto}{2} = 0.05 \Leftarrow 1-\propto = 0.90$

$$V = \left[\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{(n_1-1)}+\frac{\left(\frac{S_1^2}{n_1}\right)^2}{(n_2-1)}}\right] = \left[\frac{(\frac{0.0547}{4}+\frac{0.0991}{5})^2}{\frac{\left(\frac{0.0547}{4}\right)^2}{(4-1)}+\frac{\left(\frac{0.0991}{5}\right)^2}{(5-1)}}\right] = 6.99 = 7$$

وعليه $T_{\left(\frac{\propto}{2},V\right)} = T_{(0.05,7)} = 1.895$

وبالتعويض في فترة الثقة نجد:

$$(\bar{X}_1-\bar{X}_2) \pm T_{\left(\frac{\propto}{2},V\right)}\sqrt{\left(\frac{S_1^2}{n_1},\frac{S_2^2}{n_2}\right)} = (1-1.2) \pm 1.895\sqrt{\left(\frac{0.0547}{4}+\frac{0.0991}{5}\right)}$$
$$= [-0.5468\,;0.1468]$$

فترة الثقة للفرق بين متوسطي المجتمعين $(\mu_1-\mu_2)$ عند مستوى الثقة 90% هي بالتقريب: -[0.55,0.15].[1]

5. تقدير فترة الثقة لنسبة المجتمع P:

إن تقدير النسبة بفترة هو عبارة عن إيجاد تقدير نقطي لنسبة النجاح في المجتمع P ثم إيجاد توزيع المعاينة لذلك المقدار واستعمال هذه المعلومات لإيجاد فترة ذات معامل ثقة معين تحصر نسبة النجاح P بداخلها. إذا كانت n كبيرة فالتوزيع يقترب من التوزيع الطبيعي المعياري: $Z=\frac{\bar{P}-P}{\sqrt{\frac{Pq}{n}}}$

ويمكن وضع العبارة الاحتمالية التالية:

[1]– محمد قريشي، مرجع سابق، ص، ص: 120–127.

$$\mathrm{P}\left(-Z_{1-\frac{\propto}{2}} < Z < Z_{1-\frac{\propto}{2}}\right) = 1{-}\propto$$

حيث: $Z = \frac{\bar{P}-P}{\sqrt{\frac{Pq}{n}}}$ تخضع للتوزيع الطبيعي المعياري تقريبا $X, \bar{P} = \frac{X}{n}$ عدد النجاحات في العينة

التي حجمها n أي يصبح من الصعب استخدام العبارة الاحتمالية السابقة لإيجاد فترة الثقة للنسبة P وذلك

لأن P في المقام غير معلومة ولذلك نستعمل $\bar{P} = \frac{X}{n}$ بدلا من P في المقام فنحصل على فترة ثقة

$(1{-}\propto)100\%$ التقريبية للنسبة $\bar{P}$ وهي:

$$\bar{P} - Z_{1-\frac{\propto}{2}}\sqrt{\frac{\bar{P}\bar{q}}{n}} < P < \bar{P} + Z_{1-\frac{\propto}{2}}\sqrt{\frac{\bar{P}\bar{q}}{n}}$$

ملاحظة: ونضيف معامل التصحيح إذا كانت $n \geq 0.05N$ والسحب بدون إرجاع.[1]

المثال الأول: أخذت عينة عشوائية من 900 شخص بإحدى القرى فوجد أن عدد الذين يحملون فصيلة دم B في هذه العينة هو 50 شخص، أوجد نسبة 99% فترة ثقة لنسبة الأشخاص الذين يحملون فصيلة الدم B في هذه القرية.

الحل:

$$\bar{P} = \frac{X}{n} = \frac{50}{900} = 0.06$$

$$Z_{1-\frac{\propto}{2}} = 2.58 \Leftarrow 1{-}\propto = 99\%$$

[1]– محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء ومبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 224–225.

فترة الثقة للنسبة:

$$\bar{P} \pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{P}\bar{q}}{n}} = 0.06 \pm 2.58\sqrt{\frac{(0.06 \times 0.04)}{900}}$$

$$P \in [0.04 \text{ و } 0.08]$$

نحن واثقون بنسبة 99% بأن نسبة الأشخاص الذين يحملون فصيلة الدم B في هذه القرية تقع بين النسبتين 0.04 والنسبة 0.08.[1]

المثال الثاني: في عينة عشوائية حجمها 100 عامل في مصنع به 1200 عامل، وجد أن 70 عامل يفضلون الاشتراك في نظام المعاشات كأفراد بدلا من الاشتراك في مشروع معاشات خاص بالشركة، أوجد فترة الثقة 95% لنسبة العاملين الذين يفضلون مشروعات معاشات فردية.

الحل:

$$\bar{P} = \frac{X}{n} = \frac{70}{100} = 0.7$$

$$1-\alpha = 95\% \Longrightarrow Z_{1-\frac{\alpha}{2}} = 1.96$$

لدينا: $n_q > 5$ و $n > 30$ أو $n > 0.05N$ و $100 > 0.05(1200)$

$$100 > 60$$

أي نستخدم معامل التصحيح وتصبح فترة الثقة:

[1]– أحمد السيد عامر، مرجع سابق، ص216.

$$P \in P \pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{P}\bar{q}}{n}}\sqrt{\frac{N-n}{N-1}}$$

$$P = 0.7 \pm 1.96\sqrt{\frac{(0.7\times 0.3)}{100}}\sqrt{\frac{1200-100}{1200-1}}$$

$$= 0.7 \pm 0.09$$

$$P \epsilon [0.61, 0.79]$$

نحن واثقون بدرجة ثقة 95% أن نسبة كل العاملين الذين يفضلون مشروعات معاشات فردية تقع بين النسبة 0.61 والنسبة 0.79.[1]

تحديد حجم العينة المناسب لتقدير نسبة المجتمع.

من الجدير بالذكر أ، P غالبا تكون مجهولة وإذا كانت هناك معلومات سابقا من دراسات مماثلة فبالإمكان استخدام قيمة P المعروفة من الدراسات السابقة، أما إذا لم تكن هنالك أي فكرة عن قيمة P فإننا نأخذ بمبدأ أسوء الأوضاع وهو أ، تكون قيمة $P = \frac{1}{2}$ لأن ذلك يؤدي إلى أكبر خطأ معياري للمقدار، ولهذا إذا تم تحديد المقدار الأكبر المسموح به للخطأ في تقدير P يمكن حساب حجم العينة اللازمة لتحقيق ذلك الحد، حيث تكون:

$$n \geq \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 . P . q$$

إذا كانت P معلومة من دراسات سابقة، أما إذا كانت مجهولة فإن:

[1]– دومينيك سالفاتور، مرجع سابق، ص 90.

$$n \geq \frac{1}{4}(\frac{Z_{\frac{\propto}{2}}}{E})^2$$

مثال: نريد القيام بدراسة طبية لتقدير نسبة المواطنين الذين يعانون من مشاكل في النظر، كم شخصا يجب فحصهم كي نكون واثقين بنسبة 98% أن الخطأ في تقدير هذه النسبة لا يزيد عن 0.05 في الحالتين التاليتين:

1- إذا لم يكن لدينا أية معلومات سابقة عن هذه النسبة.

2- إذا كنا نعلم من دراسات سابقة أن هذه النسبة تكون حوالي 0.3.

الحل:

لدينا: $1-\propto= 0.98 \Longrightarrow \frac{\propto}{2} = 0.01 \Longrightarrow Z_{\frac{\propto}{2}} = 2.33$

الحالة الأولى: $E = 0.05$ و $P = \frac{1}{2}$

$$n \geq \frac{1}{4}(\frac{Z_{\frac{\propto}{2}}}{E})^2$$

$$n \geq \frac{1}{4}(\frac{2.33}{0.05})^2$$

$$n \geq 542.89 \Longrightarrow n = 543$$

الحالة الثانية:

$E = 0.05$ و $P = 0.3$

$$n \geq \frac{1}{4}(\frac{Z_{\frac{\propto}{2}}}{E})^{2}\hat{}Pq$$

$$n \geq \frac{1}{4}(\frac{2.33}{0.05})^2(0.3)(0.7)$$

$$n \geq 456.0276 \Rightarrow n = 456$$

6. فترة الثقة للفرق بين نسبتي المجتمع $(P_1 - P_2)$

بنفس الطريقة التي وجدنا بها فترة الثقة لنسبة المجتمع فإن إيجاد فترة الثقة $(1-\propto)100\%$ للفرق بين نسبتي المجتمعين $(p_1 - p_2)$، حيث إذا أخذت عينة عشوائية حجمها n_1 من توزيع ذي الحدين $B(1, p_1)$، ثم أخذت عينة عشوائية مستقلة عن الأولى حجمها n_2 من توزيع آخر ذي الحدين $B(1, p_2)$ وكان حجم العينتين كبير فإن فترة الثقة $(1-\propto)100\%$ للفرق بين نسبتي المجتمعين $(p_1 - p_2)$ هي:

$$(\bar{P}_1 - \bar{P}_2) \pm Z_{\frac{\propto}{2}}\sqrt{\frac{\bar{P}_1\bar{q}_1}{n_1} + \frac{\bar{P}_2\bar{q}_2}{n_2}}$$

المثال الأول: سحبت عينتان عشوائيتان مستقلتان مع الإرجاع الأولى تحتوي على 120 وحدة منتجة بالآلة (أ) ووجدنا بها 6 وحدات معيبة، والثانية تحتوي على 200 وحدة منتجة بالآلة (ب) ووجدنا بها 9 وحدات معيبة.

المطلوب: قدر الفرق بين نسبة الوحدات المعيبة في الإنتاج الكلي للآلة (أ) ونسبة الوحدات المعيبة في الإنتاج الكلي للآلة (ب) باستخدام مستوى ثقة 95%.

الحل: بما أن العينتين عشوائيتين ومستقلتين وحجم العينتين كبير $n_1, n_2 \geq 30$ فإن فترة الثقة:

$$(\bar{P}_1 - \bar{P}_2) \pm Z_{\frac{\propto}{2}}\sqrt{\frac{\bar{P}_1\bar{q}_1}{n_1} + \frac{\bar{P}_2\bar{q}_2}{n_2}}$$

$$\bar{P}_1 = \frac{6}{120} = 0.05$$

$$\bar{P}_2 = \frac{9}{200} = 0.045$$

$$1-\propto = 0.95 \Rightarrow \frac{\propto}{2} = 0.025 \Rightarrow Z_{\frac{\propto}{2}} = Z_{0.025} = 1.96$$

فترة الثقة:

$$0.005 \pm 1.96\sqrt{\frac{(0.05)(0.95)}{120} + \frac{(0.045)(0.955)}{200}}$$

$$P_1 - P_2 \epsilon [-0.0434 \text{ و } 0.0534]$$

أي نحن نثق بدرجة 95% أن الفرق بين نسبة الوحدات المعيبة في إنتاج الآلة (أ) ونسبة الوحدات المعيبة في إنتاج الآلة (ب) يقع بين النسبتين 4.43%- و5.34%.[1]

المثال الثاني: قامت إدارة التعليم في إحدى المناطق يسحب عينتين عشوائيتين مستقلتين من مدرسة (أ) و (ب) وحجمهما 100، 81 طالب على التوالي فوجد أن عدد الناجحين 80، 61 على التوالي، أوجد ما يلي:

أ- فترة ثقة لنسبة الناجحين في المدرسة (أ) بدرجة ثقة 95%.

ب- فترة ثقة لنسبة الراسبين في المدرسة (ب) بدرجة ثقة 80%.

ج- فترة ثقة للفرق بين نسبتي الناجحين في المدرستين بدرجة ثقة 99%.

[1]- محمد قريشي، مرجع سابق، ص، ص: 139-141.

الحل:

$\bar{P}_2 = \frac{61}{81} = 0.75$ و $\bar{P}_1 = \frac{80}{100} = 0.8$

$\bar{q}_2 = 0.25$ و

$$\bar{q}_2 = 0.2$$

أ- $$P_1 \epsilon \bar{P}_1 \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{p}_1 \bar{q}_1}{n_1}} = 0.8 \pm 1.96 \sqrt{\frac{0.8 \times 0.2}{100}} = 0.8 \pm 0.08$$

$P_1 \epsilon [0.72, 0.88]$

نحن واثقون بدرجة 95% أن نسبة الناجحين في المدرسة (أ) تقع في المجال 72% ولا تزيد عن 88%.

ب- $$1 - \alpha = 0.8 \Rightarrow \alpha = 0.20 \Rightarrow \frac{\alpha}{2} = 0.10$$

$$Z_{\frac{\alpha}{2}} = Z_{0.10} = 1.28$$

فترة الثقة:

$$q_2 \epsilon \bar{q}_2 \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{P}_2 \bar{q}_2}{n_2}} = 0.25 \pm 1.28 \sqrt{\frac{0.75 \times 0.25}{81}} = 0.25 \pm 0.06$$

$$q_1 \epsilon [0.19, 0.31]$$

نحن واثقون بنسبة 80% أن نسبة الراسبين في المدرسة ب لا تقل عن 19% ولا تزيد عن 31%.

د- $P_1 - P_2 \epsilon(\bar{P}_1 - \bar{P}_2) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{P}_1\bar{q}_1}{n_1} + \frac{\bar{P}_2\bar{q}_2}{n_2}} = (0.8 - 0.75) \pm$

$$2.58\sqrt{\frac{0.8\times0.2}{100} + \frac{0.75\times0.25}{81}} = 0.05 \pm (2.58 \times 0.063) = 0.05 \pm 0.16$$

$$(P_1 - P_2)\epsilon[-0.11,0.21]$$

نحن واثقون بدرجة 95% أن الفرق بين نسبتي النجاح في المدرستين لا يقل عن 11% ولا يزيد عن 21%.[1]

7. فترة الثقة للتباين:

من خلال محور توزيع المعاينة عرفنا أنه إذا كانت $X_1, X_2 \ldots . X_n$ عينة عشوائية من مجتمع طبيعي $N(\mu, \gamma^2)$ فإن $\frac{S^2(n-1)}{\gamma^2\gamma}$ تخضع لتوزيع X^2 بدرجات حرية (n−1) حيث S^2 هو تباين العينة لهذا لإيجاد فترة ثقة $100\%(1-\propto)$ للتباين γ^2 تجد من جداول X^2 أن النقطتين $X^2[\frac{\propto}{2}, n-1]$ و$X^2[1-\frac{\propto}{2}, n-1]$ تحصران بينهما مساحة $(1-\propto)$ تحت توزيع X^2 بدرجة حرية (n−1) أي أن:

$$P(X^2_{\left[1-\frac{\propto}{2},n-1\right]} \leq \frac{(n-1)S^2}{\gamma^2} \leq X^2_{\left[\frac{\propto}{2},n-1\right]} = 1-\propto$$

وبالتبسيط الجبري خذ مقلوب كل حد مع تغيير اتجاه المتباينة ثم نضرب $(n-1)S^2$ نجد:

$$P\left[\frac{(n-1)S^2}{X^2_{\left[1-\frac{\propto}{2},n-1\right]}} \leq \gamma^2 \leq \frac{(n-1)S^2}{X^2_{\left[\frac{\propto}{2},n-1\right]}}\right] = 1-\propto$$

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 2)، مرجع سابق، ص، ص: 147، 176.

نظرية: إذا كانت $X_1, X_2 \ldots\ldots X_n$ عينة عشوائية من توزيع طبيعي $N(\mu, \gamma^2)$ فإن فترة $100\%(1-\propto)$ ثقة للتباين γ^2 هي:

$$P\left[\frac{(n-1)S^2}{X^2_{\left[1-\frac{\propto}{2}, n-1\right]}} \leq \gamma^2 \leq \frac{(n-1)S^2}{X^2_{\left[\frac{\propto}{2}, n-1\right]}}\right]$$

حيث S^2 هي تباين العينة $S^2 = \frac{\sum(X-\bar{X})^2}{(n-1)}$

$$S^2 = \frac{\sum X^2 - n\bar{X}^2}{n-1}$$

وتجد فترة الثقة للانحراف المعياري بإيجاد الجذر التربيعي لجذور حدود فترة الثقة للتباين.

مثال: عينة عشوائية حجمها 20 أخذت من مجتمع طبيعي $N(\mu, \gamma^2)$ فأعطت تباين $S^2 = 16$، أوجد فترة الثقة 95% للتباين γ^2؟

الحل: $1-\propto = 0.95 \Longrightarrow \frac{\propto}{2} = 0.975 \Longrightarrow 1 - \frac{\propto}{2} = 0.025$

$$V = n - 1 = 20 - 1 = 19$$

$$X^2 = [0.025, 19] = 8.907$$

$$X^2 = [0.025, 19] = 8.907$$

فترة الثقة: $\left[\frac{19\times16}{8.907}\right.$ و $\left.\frac{19\times16}{32.852}\right]$

[34.13 و 9.33]

أما فترة ثقة 95% للانحراف المعياري γ هي: [$\sqrt{34.13}$ و $\sqrt{9.33}$]

8. فترة الثقة للنسبة بين تباينين:

من خلال محور توزيع المعاينة فإن $\frac{\frac{S_1^2}{S_2^2}}{\frac{\gamma_1^2}{\gamma_2^2}}$ يخضع للتوزيع F_{n_1-1,n_2-1} لهذا إذا أردت إيجاد فترة

$(1-\propto)100\%$ ثقة للنسبة $\frac{\gamma_1^2}{\gamma_2^2}$ فيجب إيجاد الاحصائيات W, L بحيث:

$$P\left(L \leq \frac{\gamma_1^2}{\gamma_2^2} \leq \mu\right) = 1-\propto$$

وباستخدام توزيع F نحصل على النظرية.

نظرية: إذا كانت $X_1, X_2 \ldots\ldots X_n$ عينة عشوائية من $N(\mu_1, \gamma_1^2)$ وكانت $y_1, y_2 \ldots\ldots y_n$ عينة عشوائية من $N(\mu_2, \gamma_2^2)$ مستقل عن المجتمع الأول فإن فترة الثقة $(1-\propto)100\%$ للنسبة $\frac{\gamma_1^2}{\gamma_2^2}$ هي:

$$\frac{S_1^2}{S_2^2}\frac{1}{F[1-\frac{\propto}{2}, n_1-1, n_2-1]} < \frac{\gamma_1^2}{\gamma_2^2} < \frac{S_1^2}{S_2^2}\frac{1}{F[1-\frac{\propto}{2}, n_2-1, n_1-1]}$$

مثال: أخذت عينة عشوائية حجمها $n_1 = 9$ من مجتمع $N(\mu_1, \gamma_1^2)$ فأعطت تباين $S_1^2 = 18$،
وأخذت عينة عشوائية $n_2 = 11$ من مجتمع $N(\mu_2, \gamma_2^2)$ مستقل عن الأول فأعطت التباين $S_2^2 = 12$.

أوجد فترة الثقة 90% للنسبة $\frac{\gamma_1^2}{\gamma_2^2}$

الحل: من جدول F نجد: F[0.95,8.10]=3.07

F[0.95,10.8]=3.35

فترة الثقة: $\frac{18}{12}.\frac{1}{3.07} < \frac{\gamma_1^2}{\gamma_2^2} < \frac{18}{12}3.35$

$$0.488 < \frac{\gamma_1^2}{\gamma_2^2} < 5.02$$

فترة الثقة: 90% للنسبة $\frac{\gamma_1^2}{\gamma_2^2}$ هي [0.488, 5.02].[1]

[1]– محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 229، 232.

رابعا: سلسلة تمارين حول التقدير الإحصائي.

- **التمرين الأول:** لمعرفة نسبة المواطنين الذين يشربون المشروبات الغازية في ولاية بسكرة أخذت عينة حجمها 2000 مواطن فوجد أن 1200 يحبون هذه المشروبات.

المطلوب: ما هو التقدير النقطي لنسبة المواطنين الذين يحبون المشروبات الغازية؟

- **التمرين الثاني:** تسبب المقاعد الخالية لشركات الطيران في خسارة لمصدر الدخل، بفرض إحدى شركات الطيران الكبرى أرادت تقدير عدد المقاعد الخالية لكل رحلة خلال العام الماضي، ولهذا الغرض تم اختيار عشوائي لعدد 225 رحلة طيران وتسجيل عدد المقاعد الخالية في كل رحلة من هذه العينة وكان الوسط الحسابي والانحراف المعياري لعدد المقاعد الخالية في هذه العينة هما $\bar{X} = 11.6$ مقعد S=4.1 مقعد.

المطلوب: قدر μ (عدد المقاعد الخالية للرحلة خلال العام الماضي) باستخدام فترة ثقة 90%.

- **التمرين الثالث:** أخذت عينة عشوائية حجمها 9 من مجتمع طبيعي فكان وسطها الحسابي 12 وانحرافها المعياري 0.6.

المطلوب: أوجد فترة الثقة 90% لمعدل المجتمع μ

- **التمرين الرابع:** أوجد حجم العينة n التي يجب أخذها من توزيع طبيعي $N(\mu, 25)$ للحصول على فترة ثقة للوسط 90% وبطول 4.

- **التمرين الخامس:** إذا كان مراقب الجودة في أحد مصانع الحديد يختار عينة عشوائية كل شهر لاختبار قوة تحمل القضبان المنتجة فإذا سحب عينة تتكون من 16 قضيبا، فوجد أن متوسط قوة تحملها (6200 وحدة تحمل) بانحراف معياري (180 وحدة تحمل)، أوجد فترة الثقة لمتوسط قوة

تحمل القضبان المنتجة بدرجة ثقة 95% في الحالتين التاليتين وبافتراض أن قوة تحمل القضبان تتبع التوزيع الطبيعي:

1- الحالة الأولى: الانحراف المعياري للمجتمع معلوم (γ معلوم=200)

2- الحالة الثانية: γ مجهول.

- **التمرين السادس:** أجري امتحان في مادة الرياضيات لمجموعتين مستقلتين من الطلبة الأولى تشمل 80 طالب والثانية 60 طالبة وكانت نتائج هذا الامتحان للمجموعتين كما يلي:

مجموعة الطلبة n_1	مجموعة الطالبات n_2
$\bar{X}_1 = 75$ $S_1^2 = 81$	$\bar{X}_2 = 68$ $S_2^2 = 49$

المطلوب: بافتراض أن تبايني المجتمعين غير متساويين، أوجد فترة الثقة للفرق بين متوسطي نقاط الطلبة والطالبات $(\mu_1 - \mu_2)$ باستخدام مستوى ثقة 95%.

- **التمرين السابع:** لمعرفة نسبة الطالبات القاطنات بالأحياء الجامعية في كلية العلوم الاقتصادية تم اختيار عينة مكونة من 1000 طالبة فوجدنا عدد الطالبات المقيمات في الإقامات الجامعية هي 200 طالبة، أوجد 95% فترة ثقة لنسبة الطالبات القاطنات بالأحياء الجامعية في كلية العلوم الاقتصادية؟

- **التمرين الثامن:** قامت باحثة بسحب عينتين عشوائيتين مستقلتين من قسم الاقتصاد وقسم علوم التسيير حجمهما 110، 90 على التوالي، فوجدت عدد الناجحين 85، 65 على التوالي.

المطلوب: أوجد ما يلي:

أ) فترة ثقة لنسبة الناجحين في قسم الاقتصاد بدرجة ثقة 95%.

ب) فترة ثقة لنسبة الراسبين في قسم علوم التسيير بدرجة ثقة 80%.

ت) فترة ثقة للفرق بين نسبتي الناجحين في القسمين بدرجة ثقة 99%.

- **التمرين التاسع:** عينة عشوائية حجمها 15 أخذت من مجتمع طبيعي$N(\mu,\gamma^2)$ فأعطت التباين $S^2 = 9$

أوجد فترة ثقة 99% للتباين γ^2.

المحور الرابع:

إختبار الفرضيات

أولا: تعريفات مهمة:

إن أحد فروع الإحصاء الاستنتاجي هو اختبار الفرضيات في كثير الأحيان لا نكتفي بتقدير معلمة المجتمع بأن نعطيها قيمة معينة أو نبني لها فترة ثقة معينة بل نحتاج إلى اتخاذ قرار حول صحة فرضية معينة أو عدم صحتها أي أننا نحتاج إلى اختبار الفرضيات المتعلقة بمعلمات المجتمع ففي موضوع التقدير لا يكون لديك معلومات عن المعلمة أو المعلمات التي تريد تقديرها أي انك تستعمل المعلومات من العينة لتحديد تقديرات المعالم التي أنت بصدد دراستها، أما موضوع اختبار الفرضيات فيكون هناك ادعاء أو افتراض أن المعلمة أو المعالم لها قيم محددة وانت تحتاج أن نختبر هذا الادعاء أي تستعمل البيانات من العينات للحكم أن هذه البيانات تعطي الأدلة كافية حول صحة ذلك الادعاء.

1- <u>الفرضية الإحصائية والفرضية الصفرية والفرضية البديلة:</u>

الفرضية الإحصائية هي كل عبارة عن إحدى معالم المجتمع أو عدة معالم تكون قابلة للاختبار وبالتالي تكون صحتها أو عدم صحتها بحاجة إلى قرار، من هنا فإن الفرضية الإحصائية تتعلق بعبارة عن إحدى المعلمات مثل وسط المجتمع أو نسبة النجاح أو التباين أو عدة معالم مثل المقارنة بين معلمتين أو أكثر.

في معظم الأحيان هناك نوعان من الفرضيات في المسألة الواحدة النوع الأول هو الفرضية الصفرية أو الابتدائية وهي الفرضية التي تبنى على أمل أن يتخذ قرار بعدم صحتها، ونصطلح الآن على اعتبار أي فرضية نود اختبارها صفرية ونعبر عنها بالرمز H_0، وإن رفض الفرضية الصفرية يؤدي إلى قبول فرضية أخرى تسمى الفرضية البديلة ونعبر عنها بالرمز H_1، وعندما تصاغ الفرضية الصفرية المتعلقة بمعلمة مجتمع معينا بشكل يعين قيمة محددة لتلك المعلمة فإنها تسمى فرضية بسيطة.

مثال: يكتب أحد المصانع في نشرة مواصفات المصابيح الكهربائية (النيون) التي ينتجها أن معدل عمر المصباح 1000 ساعة، أردت اختبار هذا الادعاء. فكيف تكتب الفرضية الصفرية والفرضية البديلة وأيهما فرضية بسيطة.

الحل: افترض أن معدل عمر جميع المصابيح التي ينتجها المصنع هو μ، إذا الفرضية الصفرية هي $H_0: \mu = 1000$ وهي فرضية بسيطة، أما الفرضية البديلة فتعتمد على الحالة المتوقعة التي تريد إجراء الاختبار من أجلها فمثلا إذا كنت تريد اختبار H_0 بغرض شراء مصابيح من المصنع فإن الفرضية البديلة $H_0: \mu > 1000$ وفي هذه الحالة الفرضية البديلة لم تعين قيمة محددة للمعلمة μ بل سمحت بفترة القيم هي جميع القيم أكبر من 1000.

2-الخطأ من النوع الأول والخطأ من النوع الثاني:

كل قرار يبنى على نتائج عينة يكون معرضا للخطأ حيث أن مثل هذه القرارات تبنى على متغير عشوائي أو إحصاء مثل $\bar{X}$، $\bar{P}$ وغيرها، وفي اختبار الفرضيات هناك حالتان بالنسبة للفرضية الصفرية:

- أن تكون الفرضية الصفرية صحيحة وإما أن تكون غير صحيحة، وهناك أيضا نوعان من القرارات التي يتخذها الإحصائي بصدد الفرضية الصفرية هما رفض الفرضية الصفرية أو عدم رفضها.
- إذا كانت H_0 غير صحيحة وكان القرار عدم رفضها فمعنى ذلك الوقوع في خطأ أيضا وهو يسمى "الخطأ من النوع الثاني" أي يحدث الخطأ من النوع الأول إذا رفضت الفرضية الصفرية وهي في الحقيقة صحيحة، أما الخطأ الثاني يحدث إذا لم ترفض الفرضية الصفرية وهي في الحقيقة غير صحيحة.

		حالة المعلمة	
		H_0 صحيحة H_1 غير صحيحة	H_0 غير صحيحة H_1 صحيحة
قرار الإحصاء	أقبل H_0	قرار صحيح	خطأ من النوع الثاني
	أرفض H_0	خطأ من النوع الأول	قرار صحيح

من خلال التعاريف السابقة نستطيع حساب احتمالات حدوث الأخطاء من النوع الأول والثاني كما يلي:

- احتمال الخطأ من النوع الأول ونعبر عنه بالرمز $\propto$ هو: (H_0 صحيحة، رفض الفرضية H_0)
$\propto = P$

أي $\propto$ تساوي احتمال رفض الفرضية إذا علم أن H_0 صحيحة.

- احتمال الخطأ من النوع الثاني ونعبر عنه بالرمز B هو: (H_1 صحيحة/ عدم رفض الفرضية
B=P (H_1

أي B تساوي احتمال عدم رفض للفرضية H_0 إذا علم أن H_1 صحيحة.

- عند حساب الاحتمالات السابقة نلاحظ أن الحادث رفض الفرضية H_0 يعني وقوع إحصاء الاختبار في منطقة الرفض وأن العبارة H_0 صحيحة تعني استعمال قيمة المعلمة المحددة لك في الفرضية الصفرية.

- إن الحادث عدم رفض الفرضية H_0 يعني وقوع إحصاء الاختبار في متممة منطقة الرفض أي متممة المنطقة الحرجة.
- العبارة H_1 صحيحة تعني وجوب إعطاء قيمة معنية للمعلمة تحت الاختبار على أن تكون هذه القيمة من القيم المسموح بها تحت الفرضية H_1.[1]
- وبصفة عامة يمكن توضيح الخصائص التالية للعلاقة بين الخطأين:
- يرتبط الخطأ من النوع الأول (∝) ارتباطا عكسيا مع الخطأ من النوع الثاني (B) أي أن انخفاض احتمال أحدهما يؤدي إلى زيادة احتمال الآخر.
- زيادة حجم العينة يؤدي إلى تناقص كلا النوعين من الخطأ.
- إذا كان الفرض العدمي H_0 غير صحيح، فإن قيمة الخطأ من النوع الثاني (B) تكون أكبر ما يمكن عندما تكون القيمة الحقيقية للمعلمة تكاد تتطابق مع القيمة الافتراضية لها والعكس صحيح عندما يكون الفرق بين القيمتين الحقيقية والافتراضية، فإن قيمة الخطأ من النوع الثاني (B) تقل.

3-مستوى المعنوية أو مستوى الدلالة

في اختبار فرضية معينة فإن أقصى احتمال والذي يمكن أن نتحصل به خطأ من النوع الأول يسمى "مستوى المعنوية للاختبار"، هذا الاحتمال يرمز له بالرمز (∝) ويحدد بشكل عام قبل سحب أي عينة لكي لا تتأثر النتائج التي حصلنا عليها في اختبارنا، ومن الناحية العملية فإننا نستخدم عادة مستوى المعنوية 0.05 أو 0.01، حيث مثلا 0.05 أو 5% تعني أن هناك حوالي 5 فرص من أصل 100 فرصة أننا سوف نرفض الفروض وهي صحيحة، بمعنى أننا سنكون واثقين بنسبة 95% أننا سنتخذ القرار الصحيح، وبالمقابل فإنه من الممكن أن نكون على خطأ باحتمال 0.05.[2]

4-إحصائية الاختبار: هي متغير عشوائي يجب أن يكون توزيعه الاحتمالي معلوما عندما يكون فرض العدم H_0 صحيحا ونحسب قيمته من بيانات العينة ونستخدم قيمة إحصائية الاختبار المحسوبة من بيانات العينة العشوائية المسحوبة من المجتمع محل الدراسة والتي يطلق عليها القيمة المشاهدة لإحصائية الاختبار لاتخاذ القرار برفض أو عدم رفض فرض العدم H_0 فيتم

[1]– محمد صبحي أبو صالح، عدنان محمد عوض، مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 233-235

[2]– محمد قريشي، مرجع سابق، ص: 189.

تقسيم كل النتائج الممكن الحصول عليها أي كل القيم التي يمكن أن تأخذها إحصائية الاختبار لجزئين غير متداخلين أحدهما للنتائج التي إذا ظهرت نقبل فرض العدم ويسمى منطقة القبول، والآخر للنتائج التي إذا ظهرت نرفض فرض العدم وتسمى منطقة الرفض.[1]

5- <u>أنواع اختبار الفروض</u>

إذا أردنا اختبار فرض العدم $H_0: \mathrm{F} = F_0$ حيث F إحدى معلمات المجتمع فإنه يكون لدينا ثلاث أنواع من الاختبارات بحسب صيغة الفرض البديل H_0 كما يلي:

- اختبار ذو طرفين (ذيلين): وذلك عندما تكون منطقة الرفض موزعة على طرفين بالتساوي وذلك يحدث عندما يكون الفرض البديل "لا يساوي" أي: $H_1: \mathrm{F} \neq F_0$
- اختبار ذو طرف أيمن (علوي): وذلك عندما تكون منطقة الرفض على يمين التوزيع أي مركزة في الطرف الأيمن للمنحنى وذلك يحدث عندما يكون الفرض البديل "أكبر من " أي: $H_1: \mathrm{F} > F_0$
- اختبار ذو طرف أيسر (سفلي): ذلك عندما تكون منطقة الرفض على يسار التوزيع أي مركزه بالكامل في الطرف الأيسر للمنحنى وذلك يحدث عندما يكون الفرض البديل "أقل من" أي: $H_1: \mathrm{F} < F_0$

6- <u>توضيح كيفية تحديد الفرض البديل</u>

يمثل الفرض البديل خطوة مهمة في اختبارات الفروض لأنه هو الذي يحدد نوع الاختبار المستخدم وعادة يعطي الفرض البديل في السؤال، ولكن ماذا عن الوضع في الحياة العملية حيث أن الباحث هو الذي عليه اختيار الشكل المناسب للفرض البديل (إما يساوي أو أكبر من أو أقل من) وبصفة عامة نقول إذا لم يكن لدى الباحث انطباع أو اتجاه معين نحو الزيادة أو النقصان فإن الفرض البديل يكون في الغالب (لا يساوي)، أي يستخدم الباحث في هذه الحالات اختبار ذو طرفين، أما إذا كان لدى الباحث انطباع أو اتجاه نحو الزيادة مثل أن يقوم مصنع بدورة تدريبية للعمال على الطرق الحديثة للإنتاج بهدف زيادة متوسط الإنتاجية، أو باتباع أسلوب جديد بهدف زيادة نسبة التوزيع أو المبيعات فإن الفرض البديل في مثل هذه الحالات يكون أكبر من ويستخدم الباحث اختبار ذو الطرف الأيمن، وكذلك إذا كان لدى الباحث انطباع أو اتجاه نحو النقصان (أو التخفيض) كان يتبع المصنع طريقة جديدة لتخفيض متوسط التكاليف أو تتبع

---

[1]- أحمد السيد عامر، مرجع سابق، ص: 247.

محطة الخدمة أسلوب جديد لتقليل متوسط وقت الانتظار فإن الغرض البديل يكون أقل من ويستخدم الباحث اختبار ذو طرف أيسر.

7-خطوات الاختبار الاحصائي:

يمكن تلخيص خطوات الاختبار الاحصائي فيما يلي:

- الخطوة الأولى: وضع أو تحديد فرض العدم H_0 والذي يأخذ عادة شكل معادلة أو مساواة وتتعلق بمعلمات المجتمع.
- الخطوة الثانية: وضع أو تحديد الفرض البديل H_1 والذي يأخذ أحد أشكال ثلاثة إما لا يساوي أو أكبر من، أقل من وهو الذي يحدد نوع الاختبار فنستخدم إما اختبار ذو طرفين أو اختبار ذو طرف أيمن أو اختبار ذو طرف أيسر.
- الخطوة الثالثة: حساب إحصائيو الاختبار وهي z أو T وذلك حسب الحالات المدروسة.
- الخطوة الرابعة: تحديد مستوى المعنوية $\propto$.
- الخطوة الخامسة: تحديد المنطقة الحرجة (منطقة الرفض والمتبقي هو منطقة القبول) وذلك بتحديد قيم z أو T المعيارية أو الجدولية التي بناءا عليها نحدد المنطقة الحرجة (منطقة الرفض) وكما ذكرنا سابقا فإن الفرض البديل هو الذي يحدد منطقة الرفض (طرفين أو طرف أيمن أو طرف أيسر).
- الخطوة السادسة: اتخاذ القرار الإحصائي: وذلك بمقارنة قيم z أو T الجدولية، فإذا وقعت z أو T الحسابية في منطقة الرفض فإن القرار هو رفض H_0 وقبول H_1 (رفض فرض العدم وقبول الفرض البديل، أما إذا وقعت z أو T الحسابية في منطقة القبول فإن القرار هو قبول H_0 ورفض H_1.[1]

[1]- أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي، الإحصاء الوصفي والاستدلالي، (الجزء 2)، مرجع سابق، ص، ص: 185-187.

ثانيا: اختبار الفروض حول متوسط المجتمع μ

في اختبار الفروض حول متوسط المجتمع μ سنميز بين ثلاث حالات سنتناولها بالتفصيل لاحقا وهي كما يلي:

الفروض لـ μ

1- <u>الحالة الأولى: اختبار الفروض لمتوسط المجتمع μ لما γ معلوم</u>

إذا كانت $(X_1, X_2 \dots \dots X_n)$ عينة عشوائية مسحوبة من مجتمع يتوزع توزيعا طبيعيا وسطه μ وانحرافه المعياري γ معلوم فإن إحصاء الاختبار المناسب هو $Z = \frac{\bar{X}-\mu_0}{\frac{\gamma}{\sqrt{n}}}$ وأردنا اختبار الفرضية الصفرية التالية: $H_0: \mu = \mu_0$ مقابل الفرضية البديلة:

- $H_1: \mu \neq \mu_0$ فإننا نرفض H_0 عند مستوى الدلالة $(\propto)$ إذا كانت:

$Z < -Z_{\frac{\propto}{2}}$ أو $Z > Z_{\frac{\propto}{2}}$

- $H_1: \mu > \mu_0$ فإننا نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت: $Z > Z_{\propto}$

- $H_1: \mu < \mu_0$ فإننا نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت: $Z < -Z_{\propto}$.[1]

ويمكن تحديد الخطوات الرسمية لاختبار الفروض لمتوسط المجتمع بـ:

أ) افترض أن μ تساوي قيمة افتراضية μ_0 ويمكن تمثيل ذلك بالعبارة $H_0: \mu = \mu_0$ ويسمى الفرض العدمي وتكون الفروض البديلة: $H_1: \mu \neq \mu_0$ أو $H_1: \mu > \mu_0$ أو $H_1: \mu < \mu_0$ وذلك حسب المسألة.

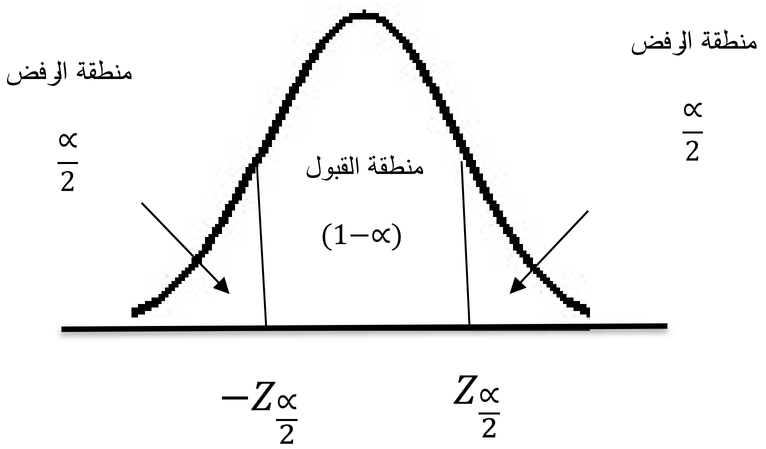
ب) تحديد مستوى معنوية للاختبار $\propto$ وعادة يكون (5% أو 1%) وعرف منطقة القبول والرفض للاختبار باستخدام التوزيع الملائم.

ج) خذ عينة عشوائية من المجتمع وأحسب $\bar{X}$ فإذا وقعت $\bar{X}$ داخل منطقة القبول يتم قبول H_0 ورفض H_1.[2]

ويتم تحديد المنطقة الحرجة (منطقة الرفض) كما يلي:

أ) $H_1: \mu \neq \mu_0$ منطقة الرفض تكون بالصورة التالية نرفض H_0 عندما تكون $Z < -Z_{\frac{\propto}{2}}$ أو $Z > Z_{\frac{\propto}{2}}$

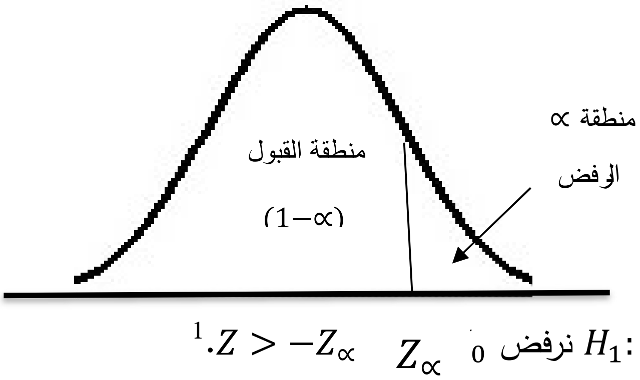
$Z_{\frac{\propto}{2}}$ نسميها Z المعيارية أو الجدولية



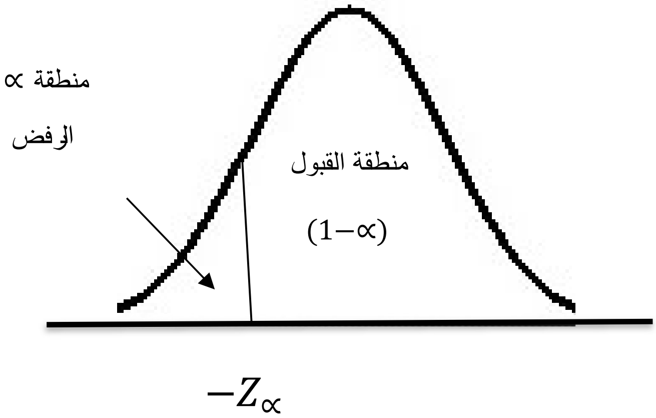
[1] محمد قريشي، مرجع سابق، ص: 190.

[2] دومينيك سالفاتور، مرجع سابق، ص: 99.

ب) $H_1: \mu > \mu_0$ نرفض H_0 عندما $Z > Z_\propto$



ج) $H_1: \mu < \mu_0$ نرفض H_0 $Z > -Z_\propto$.[1]



2-<u>الحالة الثانية: اختبار الفروض حول متوسط المجتمع μ، عندما يكون المجتمع طبيعي ذو γ مجهول n كبيرة $(n \geq 30)$</u>

إذا تم أخذ عينة عشوائية كبيرة الحجم من مجتمع يتوزع توزيعا طبيعيا وسطه μ وانحرافه γ مجهول فإن إحصاء الاختبار المناسب هو: $Z = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$

وأردنا اختبار الفرضية الصفرية $H_0: \mu = \mu_0$ مقابل الفرضية البديلة:

[1] – أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي، الإحصاء الوصفي والاستدلالي، (الجزء 2)، مرجع سابق، ص، ص: 189-190.

- $H_1: \mu \neq \mu_0$ فإننا نرفض H_0 عند مستوى الدلالة $(\propto)$ إذا كانت:

$Z < -Z_{\frac{\propto}{2}}$ أو $Z > Z_{\frac{\propto}{2}}$

- $H_1: \mu > \mu_0$ فإننا نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت: $Z > Z_{\propto}$
- $H_1: \mu < \mu_0$ فإننا نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت: $Z < -Z_{\propto}$.[1]

أمثلة حول الحالة الأولى والثانية:

المثال الأول: البيانات التالية تعطي الدخل الإجمالي كل سنة لعدد 36 شخص مختارين عشوائيا.

6.5	10.5	12.7	13.8	13.2	11.4
5.5	8	9.6	9.1	9	8.5
4.8	7.3	8.4	8.7	7.3	7.4
5.6	6.8	6.9	6.8	6.1	6.5
4.0	6.4	6.4	8	6.6	6.2
4.7	7.4	8	8.3	7.6	6.7

$$\propto = 0.05$$

$$n = 36$$

$$\sum_{i=1}^{36} X_i = 280.7$$

$$\sum_{i=1}^{36} X_2^2 = 2368.75$$

على أساس بيانات العينة هل يمكن أن تستنتج أن متوسط الدخل للشخص هو 10 وحدات نقدية في السنة؟

الحل: يجب اختبار الفرض التالي:

[1]– محمد قريشي، مرجع سابق، ص: 192.

$$H_0: \mu = 10$$

$$H_1: \mu \neq 10$$

بما أن n=36 نستخدم $Z = \frac{\bar{X}-\mu_0}{\frac{\gamma}{\sqrt{n}}}$

نحسب $\bar{X}$ و γ:

$$\bar{X} = \frac{\sum_{i=1}^{36} X_i}{n} = \frac{280.7}{36} = 7.8$$

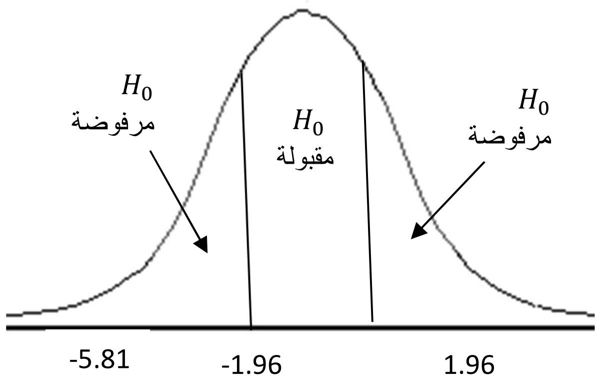
$$\gamma^2 = \frac{1}{35}\left[2368.75 - \frac{(280.7)^2}{36}\right] = \frac{180.07}{35} = 5.14$$

$$\gamma = \sqrt{\gamma^2} = \sqrt{5.14} = 2.27$$

$$Z = \frac{7.8 - 10}{\frac{2.27}{\sqrt{36}}} = -5.81$$

قيمة Z الجدولية عند $\propto = 0.05$ لاختبار الذيلين هي 1.96 حيث $Z < -1.96$ نرفض H_0 وهذا يعني أن متوسط الدخل السنوي يكون أقل من 10 وحدات نقدية.[1]

المثال الثاني: إذا كان متوسط الزيادة في أجور العاملين في إحدى المؤسسات عام 2016 هو (360ون) وفي عام 2020 أخذت عينة عشوائية حجمها 64 عامل في هذه المؤسسة فوجد أن الوسط الحسابي للزيادة في أجورهم (400ون) والانحراف المعياري (80ون) هل يدل ذلك على أن متوسط الزيادة في أجور العاملين في سنة 2020 قد اختلف عن متوسط الزيادة في الأجور عام 2016؟ وذلك عند مستوى معنوية $\propto = 0.05$



[1]– وحيد مصطفى أحمد، مرجع سابق، ص: 234.

الحل:

$$\propto = 0.05 \Longrightarrow Z_{\frac{\propto}{2}} = 1.96$$

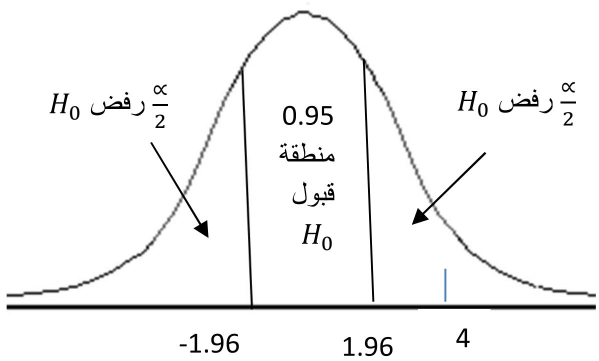
$$S = 80\,, \bar{X} = 400\,, n = 64$$

يجب اختبار الفرض التالي:

$$H_0: \mu = 360$$

$$H_1: \mu \neq 360$$

بما أن $n \geq 30$ نستخدم: $Z = Z = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{40}{\frac{80}{\sqrt{64}}} = 4$



نرفض H_0 ونقبل H_1 أي أن متوسط الزيادة في الأجور لعام 2020 قد اختلف عن متوسط الزيادة في الأجور عام 2016 وذلك عند مستوى $\propto = 0.05$.[1]

المثال الثالث: إذا كان عدد الطلاب المتفوقين بأحد الأقسام يتبع التوزيع الطبيعي بمتوسط μ وتباين $\gamma^2 = 16$ طالب، وإذا كان متوسط عدد الطلبة المتفوقين يزيد عن 4 طلاب، تم أخذ عينة من 10 طلاب من هذا القسم فكان متوسط عددهم $\bar{X} = 4.5$، فاختبر هذا الفرض عند مستوى معنوية $\propto = 0.05$.

الحل: يجب اختبار الفرض التالي:

$H_0: \mu = 4$

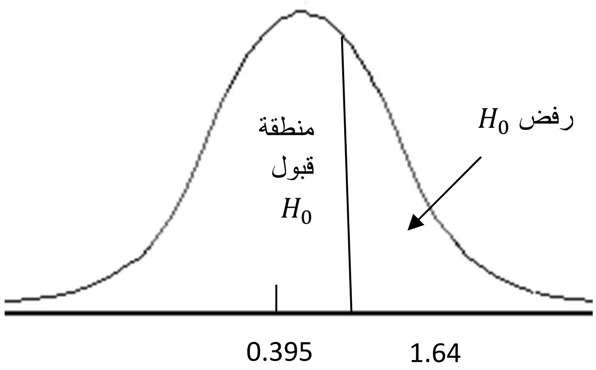
$H_1: \mu > 4$

[1]– قسم الإحصاء بجامعة الملك عبد العزيز، مرجع سابق، ص، ص: 259–260.

$\bar{X} = 4.5$ و $\gamma^2 = 16$ و n=10 و $\propto= 0.05 \Longrightarrow Z_{0.95} = 1.64$

$$Z = \frac{\bar{X} - \mu_0}{\frac{\gamma}{\sqrt{n}}} = \frac{4.5 - 4}{\frac{4}{\sqrt{10}}} = 0.395$$

نقبل H_0 ونرفض H_1 عند مستوى المعنوية $\propto= 0.05$.[1]



3-الحالة الثالثة: اختبار الفروض حول متوسط المجتمع µ عندما يكون المجتمع طبيعي ذو انحراف معياري، γ مجهول، و n صغيرة $(n < 30)$

نلاحظ في هذه الحالة أنها تختلف عن الحالة الأولى حيث أن الانحراف المعياري γ مجهول، وتختلف عن الحالة الثانية أن n صغيرة $(n < 30)$ وبالتالي لا تتوفر شروط تطبيق إحصاءة الاختيار Z، أي نستخدم التوزيع T، بدرجة حرية (n−1) إذا كانت العينة التي حجمها n قد أخذت من توزيع طبيعي $N(\mu, \gamma)$ حيث $\bar{X}$ الوسط الحسابي، S الانحراف المعياري للعينة إذا إحصاء الاختبار المناسب هو: $T = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}$

وخطوات الاختبار حول µ هي: $H_0: \mu = \mu_0$

مقابل الفرض البديل:

- $H_1: \mu > \mu_0$، النقاط الحرجة $T_\propto = T_{(1-\propto, n-1)}$ ومنطقة الرفض هي رفض H_0 إذا كان $T > T_\propto$.

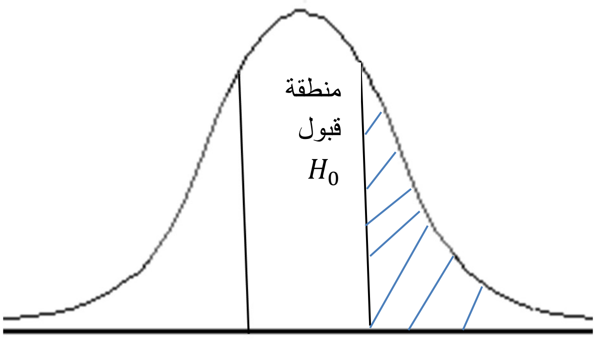
[1]– أحمد السيد عامر، مرجع سابق، ص: 252.

- $H_1: \mu < \mu_0$ النقطة الحرجة هي $T_\propto = T_{(1-\propto,n-1)}$ ومنطقة الرفض نرفض H_0 إذا كان $T > -T_\propto$.
- $H_1: \mu \neq \mu_0$ فتكون نقطتان حرجتان $T_{\frac{\propto}{2}}$، $-T_{\frac{\propto}{2}}$ حيث $T_{\frac{\propto}{2}} = T_{(1-\frac{\propto}{2},n-1)}$، ومنطقة الرفض هي أرفض H_0 إذا كان: $T < -T_{\frac{\propto}{2}}$ أو $T > T_{\frac{\propto}{2}}$ بعبارة أخرى $|T| > T_{\frac{\propto}{2}}$.[1]

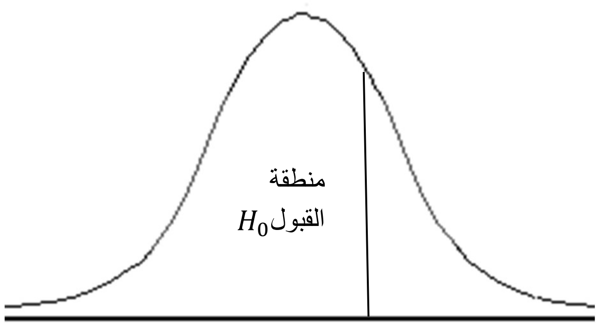
ويمكن تحديد المنطقة الحرجة (منطقة الرفض) كما يلي:

أ. $H_1: \mu \neq \mu_0$ منطقة الرفض تكون بالصورة التالية:

ترفض H_0 عندما $T > T_{\frac{\propto}{2},n-1}$ أو $T < T_{\frac{\propto}{2},n-1}$

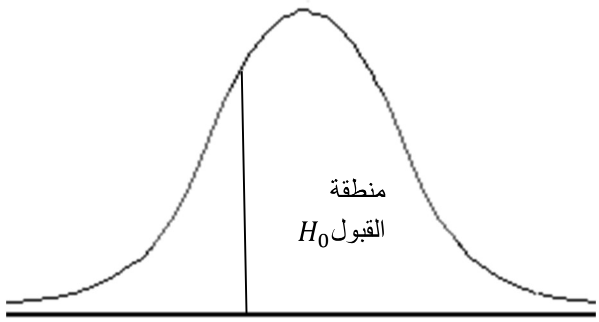


ب. $H_1: \mu > \mu_0$ نرفض H_0 عندما $T > T_{\propto,n-1}$



[1] – محمد صبحي أبو صالح، عدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 244–245.

ت. $H_1: \mu < \mu_0$ نرفض H_0 عندما $T > -T_{\propto}$.[1]



المثال الأول: ترغب شركة أن تعرف بدرجة ثقة 95% ما إذا كان يمكنها الادعاء بأن صناديق الصابون المسحوق الذي تبيعه تحتوي على أكثر من 500غ (حوالي 1.1 رطل) من الصابون وتعرف الشركة من الخبرة الماضية أن أوزان الصابون بالصناديق تتبع التوزيع الطبيعي، وقد أخذت الشركة عينة عشوائية n=25 ووجدت $\bar{X} = 520$ و S=75غ، حيث أن الشركة ترغب في اختبار ما إذا كانت μ > 500؟

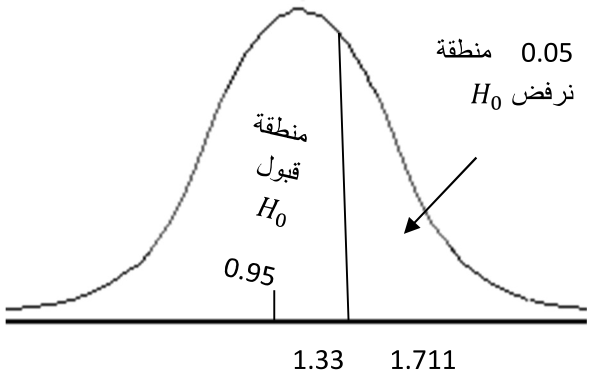
الحل: سنختبر الفرض: $H_0: \mu = 500$

$H_1: \mu > 500$

بما أن التوزيع طبيعي و $n < 30$، γ غير معلومة سنستخدم التوزيع T بدرجة حرية n−1=25−1=24

$$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{520 - 500}{\frac{75}{\sqrt{25}}} = \frac{20}{15} = 1.33$$

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي، الإحصاء الوصفي والاستدلالي، (الجزء 2)، مرجع سابق، ص، ص: 190−191.



بما أن 1.33 تقع في منطقة القبول، نقبل H_0، $\mu = 500$ عند مستوى المعنوية $\propto = 0.05$.[1]

ثالثا: اختبار الفروض حول الفرق بين وسطي مجتمعين $\mu_1 - \mu_2$

إذا كان لدينا مجتمعان لهما متوسطان μ_1،μ_2 وانحراف معياري γ_1،γ_2 على التوالي، سحبنا منهما عينتان عشوائيتان مستقلتان حجمهما n_1،n_2 بمتوسط $\bar{X}_1$،$\bar{X}_2$ وانحراف معياري S_1،S_2 على التوالي اختبر فرض العدم: $H_0: \mu_1 - \mu_2 = D_0$

وغالبا ما يكون D_0 المقدار الثابت مساويا للصفر فيكون فرض العدم H_0 أن المتوسطين متساويين $(\mu_1 = \mu_2)$.[2]

[1]– دومينيك سالفاتور، مرجع سابق، ص100.

[2]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي، الإحصاء الوصفي والاستدلالي، (الجزء 2)، مرجع سابق، ص، ص: 194.

اختبار الفروض $\mu_1 = \mu_2$

الحالة الأولى:

المجتمع طبيعي أو غير طبيعي γ_2، γ_1 معلومة

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

الحالة الثانية:

المجتمع طبيعي γ_2، γ_1 مجهولة

n_2، n_1 كبيرة n_2، $n_1 \geq 30$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

الحالة الثالثة:

المجتمع طبيعي γ_2، γ_1 مجهولة $\gamma_2 = \gamma_1$

n_2، n_1 صغيرة n_2، $n_1 < 30$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

أو تكتب: $S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

الحالة الرابعة:

المجتمع طبيعي γ_2، γ_1 مجهولة $\gamma_2 = \gamma_1$

n_2، n_1 صغيرة n_2، $n_1 < 30$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

1- الحالة الأولى: اختبار الفروض حول الفرق بين وسطي المجتمعين $(\mu_1 - \mu_2)$ لما تكون γ_2، γ_1 معلومة.

إذا أخذت عينة عشوائية حجمها n_1 من مجتمع $N(\mu_1, \gamma_1)$ وأخذت عينة عشوائية ثانية n_2 من مجتمع $N(\mu_2, \gamma_2)$ وهي مستقلة عن الأولى، وكان γ_2، γ_1 معلومتين فإن إحصاء الاختبار للفرضية $H_0: \mu_1 - \mu_2 = D_0$ هو:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

حيث $\bar{X}_1$،$\bar{X}_2$ هما وسطا العينتين على التوالي، ومن هذه النظرية نرى أن خطوات اختبار الفرضية $H_0: \mu_1 - \mu_2 = D_0$ هي نفسها الخطوات السابقة، وأغلب المراجع يفترض $D_0 = 0$ أي يتم وضع $H_0: \mu_1 - \mu_2$، أما الفرضيات البديلة هي:

- $H_1: (\mu_1 - \mu_2) \neq D_0$، نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت $Z > Z_{\frac{\propto}{2}}$ أو $Z <$ $-Z_{\frac{\propto}{2}}$.
- $H_1: (\mu_1 - \mu_2) > D_0$، نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت $Z > Z_{\propto}$.
- $H_1: (\mu_1 - \mu_2) < D_0$، نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت $Z < -Z_{\propto}$.

مثال: أخذت عينة عشوائية حجمها 22 من المجتمع $N(\mu_1; 110)$ فأعطت $\bar{X}_1 = 83$، وأخذت عينة ثابتة مستقلة عن الأولى حجمها 27 من المجتمع $N(\mu_2; 81)$ فأعطت الوسط الحسابي $\bar{X}_2 = 69$.

أ) اختبر $H_0: \mu_1 = \mu_2$ مقابل $H_1: \mu_1 > \mu_2$ عند مستوى دلالة $\propto = 0.05$.

ب) اختبر الفرضية $H_0: \mu_1 - \mu_2 = 10$ مقابل $H_1: \mu_1 - \mu_2 > 10$ عند مستوى الدلالة $\propto =$ 0.05.

الحل:

أ- اختبار الفرضية $H_0: \mu_1 - \mu_2 = 0$ أو $H_0: \mu_1 = \mu_2$

$$H_1: \mu_1 > \mu_2$$

إحصاء الاختبار:

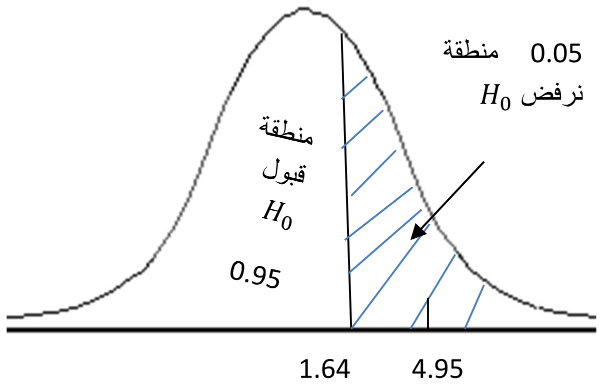
$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

بما أن الاختبار ذو طرف واحد.

$$\propto = 0.05 \Longrightarrow Z_{0.95} = 1.645$$

$$Z = \frac{(83 - 69) - 0}{\sqrt{\frac{110}{22} + \frac{81}{27}}} = \frac{14}{\sqrt{8}} = 4.95$$

بما أن $1.645 < 4.95$ تقع في منطقة رفض H_0. فإننا نرفض H_0 ونقبل H_1 أي $\mu_1 > \mu_2$.



ب-نختبر الفرضية $H_0: \mu_1 - \mu_2 = 10$

$$H_1: \mu_1 - \mu_2 > 10$$

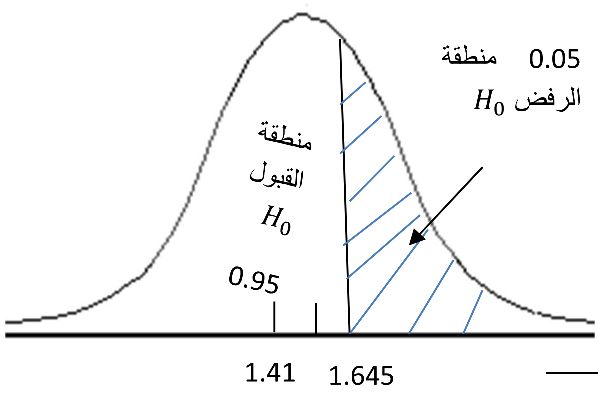
إحصاء الاختبار

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\gamma_1^2}{n_1} + \frac{\gamma_2^2}{n_2}}}$$

$$Z = \frac{(83 - 69) - 0}{\sqrt{\frac{110}{22} + \frac{81}{27}}} = \frac{4}{\sqrt{8}} = 1.41$$

بما أن $\propto = 0.05 \Longrightarrow Z_{0.95} = 1.645$

بما أن $1.41 < 1.645$ فإننا نقبل H_0 ونرفض H_1 أي نقبل الفرضية $H_0: \mu_1 - \mu_2 = 10$.[1]



[1]– محمد صبحي أبو صالح وعدنان محمد عوض: معدمه في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 251-253.

2-الحالة الثانية: اختبار حول الفرق بين وسطي مجتمعين $(\mu_1 - \mu_2)$ عندما يكون γ_1،γ_2 مجهولين وحجمي العينتين كبير بدرجة كافية $n_1, n_2 \geq 30$:

إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 كبير تم سحبها من مجتمع متوسطه μ_1 و γ_1 مجهول وكان $\bar{X}_2$ هو الوسط الحسابي لعينة عشوائية أخرى حجمها n_2 كبير أيضا مستقلة عن الأولى تم سحبها من مجتمع μ_2 و γ_2 مجهول (انحرافه المعياري) فإن إحصاء الاختبار المناسب هو Z:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

مثال: أخذت عينة عشوائية من مجتمع معين وكان حجمها 50 ووسطها الحسابي $\bar{X}_1 = 57.5$ وانحرافها المعياري $\gamma = 6.2$، ثم أخذت عينة عشوائية أخرى مستقلة عن الأولى من مجتمع آخر وكان حجمها 60 ووسطها الحسابي 54.4 وانحرافها المعياري 10.6.

المطلوب: هل نستطيع أن نستنتج أن متوسط المجتمع الأول أكبر من متوسط المجتمع الثاني عند مستوى الدلالة $(\propto = 0.05)$.

الحل: نريد اختبار $H_0: \mu_1 - \mu_2 = 0$

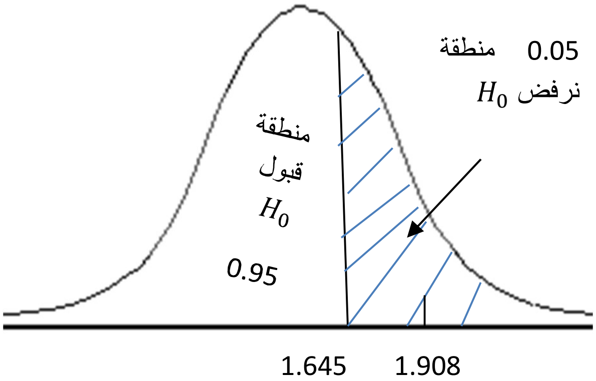
$$H_1: \mu_1 - \mu_2 > 0$$

$$\propto = 0.05 \Longrightarrow Z_{0.95} = 1.645$$

نحسب إحصاءة الاختبار Z بما أن $n_1, n_2 \geq 30$:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(57.5 - 54.4) - 0}{\sqrt{\frac{(6.2)^2}{50} + \frac{(10.6)^2}{60}}} = 1.908$$

بما أن $1.908 > 1.645$ فإننا نرفض H_0 ونقبل H_1 عند مستوى دلالة $(\propto = 0.05)$ نستنتج أن متوسط المجتمع الأول أكبر من متوسط المجتمع الثاني عند $\propto = 0.05$.[1]



3- الحالة الثالثة: اختبار الفروض حول الفرق بين وسطي المجتمعين $(\mu_1 - \mu_2)$ عندما يكون المجتمعين مجهولين $\gamma_1 = \gamma_2$ ،γ_1 ،γ_2 والعينتين صغيرتين ومستقلتين.

أخذت العينة العشوائية التي حجمها n_1 من مجتمع $N(\mu_1, \gamma_1)$ فأعطت الوسط الحسابي $\bar{X}_1$ والتباين S_1^2، وأخذت العينة العشوائية التي حجمها n_2 من مجتمع $N(\mu_2, \gamma_2)$ مستقل عن الأول فأعطت وسط الحسابي $\bar{X}_2$ والتباين S_2^2 فإذا كان $\gamma_1^2 = \gamma_2^2$ وكانت n_1 و n_2 صغيرتين يكون إحصاء اختبار فرضية العدم $H_0: \mu_1 - \mu_2 = D_0$ أو $H_0: \mu_1 = \mu_2$ هو:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

بدرجة حرية $(n_1 + n_2 - 2)$ حيث:

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

بسبب صغر حجم n_1 و n_2 لا نستطيع استعمال نظرية النهاية المركزية.

والفرضيات البديلة:

[1]- محمد قريشي، مرجع سابق، ص، ص: 101-102.

- $H_1: (\mu_1 - \mu_2) \neq D_0$، نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت $T > T_{(\frac{\propto}{2},V)}$ أو $T < -T_{(\frac{\propto}{2},V)}$.
- $H_1: (\mu_1 - \mu_2) > D_0$، نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت $T > T_{(\propto,V)}$.
- $H_1: (\mu_1 - \mu_2) < D_0$، نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت $T < -T_{(\propto,V)}$

مثال: أخذت عينة عشوائية حجمها 9 من أوزان الأطفال الذكور حديثي الولادة في أحد المستشفيات فأعطت $\bar{X} = 3.2$ كغ، $S_1 = 1.3$ كغ، وأخذت عينة عشوائية حجمها 15 من أوزان الإناث حديثات الولادة من نفس المستشفى فكان كغ $\bar{X}_2 = 2.8$، كغ$S_2 = 1.2$ على فرض أن كلا من أوزان الذكور وأوزان الإناث يخضع لتوزيع طبيعي ذي التباين نفسه اختبر الفرضية:

$H_0: \mu_1 - \mu_2 = 0$ مقابل $H_1: \mu_1 - \mu_2 > 0$ عند $\propto = 0.05$.

الحل: سنختبر الفرضية

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 > 0$$

$\propto = 0.05$ و $T_{(0.95,22)} = 1.717$

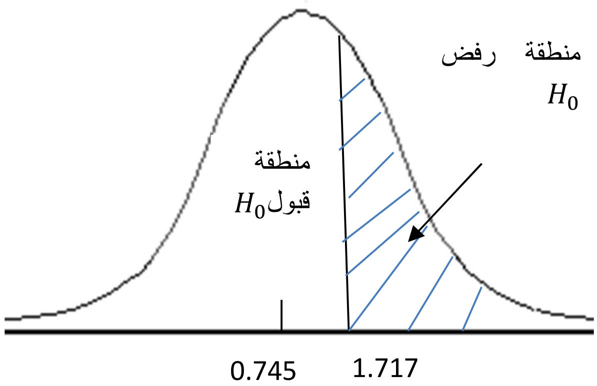
بعد تحديد النقطة الحرجة 1.717 وبما أن الاختبار ذو طرف واحد ستكون منطقة رفض H_0 إذا كان $T > 1.717$.

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ و } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$S_p^2 = \frac{[8(1.3)^2] + [14(1.2)^2]}{9 + 15 - 2} = \frac{33.68}{22} = 1.53$$

$$S_p = \sqrt{S_p^2} = \sqrt{1.53} = 1.24$$

$$T = \frac{(3.2 - 2.8) - 0}{1.24\sqrt{\frac{1}{9} + \frac{1}{15}}} = \frac{0.4}{0.523} = 0.745$$



بما أن $1.717 > 0.745$ أي تقع في منطقة قبول H_0 أي نقبل H_0 ونرفض H_1 أي لا يوجد دلالة كافية أن معدل أوزان الأطفال الذكر حديثي الولادة أكبر من معدل أوزان الإناث حديثات الولادة.[1]

4-الحالة الرابعة: اختبار الفروض حول الفرق بين وسطي المجتمعين $(\mu_1 - \mu_2)$ عندما تكون γ_1،γ_2 مجهولين و $\gamma_1 \neq \gamma_2$ والعينتين صغيرتين ومستقلتين.

إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 صغير سحبت من مجتمع طبيعي وسطه μ_1 وانحرافه γ وكان $\bar{X}_2$ هو الوسط الحسابي لعينة عشوائية أخرى مستقلة عن الأولى حجمها n_2 صغير سحبت من مجتمع طبيعي وسطه μ_2 وانحرافه γ_2، وكان تبايني المجتمعين مجهولين وغير متساويين فإن إحصاء الاختبار المناسب هو: $T = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$

وتوزيعه هو التوزيع T بدرجة حرية بالصيغة المركبة التالية:

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

سنختبر الفرضية الصفرية

$H_0: \mu_1 - \mu_2 = D_0$

1- محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 255-256.

$H_1: \mu_1 - \mu_2 \neq D_0$ أو $H_1: \mu_1 - \mu_2 > D_0$ أو $H_1: \mu_1 - \mu_2 < D_0$

مثال: لنفرض أننا سحبنا عينتين عشوائيتين مستقلتين من مجتمعين لكل منهما توزيع طبيعي فوجدنا النتائج:

$\bar{X}_1 = 32$ و $n_1 = 7$ و $S_1^2 = 4.47$

$\bar{X}_2 = 30.2$ و $n_2 = 6$ و $S_2^2 = 0.652$

المطلوب: اختبر الفرضية:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

عند مستوى المعنوية $\propto = 0.05$ وكان تباين المجتمعين مجهولين وغير متساويين.

الحل: سنختبر الفرضية:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

ولحساب احصاءة الاختبار يجب حساب V أولا:

$$V = \frac{(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{S_2^2}{n_2})^2}{n_2 - 1}} = V = \frac{(\frac{4.470}{7} + \frac{0.652}{6})^2}{\frac{(\frac{4.47}{7})^2}{7-1} + \frac{(\frac{0.652}{6})^2}{6-1}} = 8$$

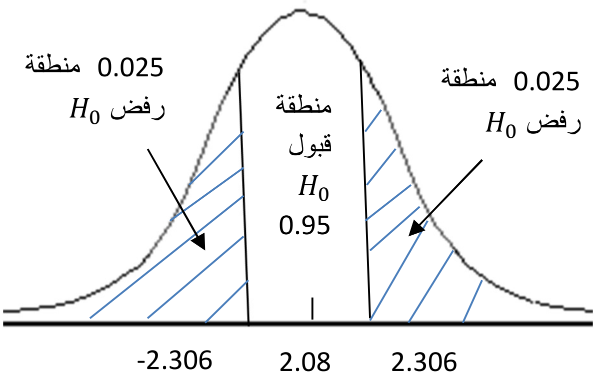
وبما أن الاختبار ذو طرفين سنحدد القيمتين الحرجتين:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = T = \frac{(32 - 30.2) - 0}{\sqrt{\frac{4.47}{7} + \frac{0.652}{6}}} = 2.08$$

$$\propto = 0.05 \Longrightarrow -T_{\frac{\propto}{2},V} = -T_{0.025,8} = -2.306 \Longrightarrow T_{\frac{\propto}{2},V} = -T_{0.025,8} = 2.306$$

نلاحظ أن $2.306 > 2.08$ أي قيمة T تقع في منطقة قبول H_0 أي نقبل H_0 ونرفض H_1 ومنه لا توجد فروق معنوية بين متوسطي المجتمعين عند مستوى المعنوية 0.05.[1]

[1]– محمد قريشي، مرجع سابق، ص، ص: 206–209.



رابعا: اختبار الفرضيات حول النسبة P في المجتمع

إذا كانت $X_1, X_2 \ldots \ldots X_n$ عينة عشوائية من توزيع ذي الحدين وكانت n كبيرة، وأردنا اختبار $H_0: P = P_0$ مقابل:

- $H_1: P \neq P_0$، نرفض H_0 عند مستوى دلالة $\propto$ إذا كانت:

$$Z = \frac{\bar{P}-P_0}{\sqrt{\frac{P_0 q_0}{n}}} < -Z_{\frac{\propto}{2}} \text{ أو } Z = \frac{\bar{P}-P_0}{\sqrt{\frac{P_0 q_0}{n}}} > Z_{\frac{\propto}{2}}$$

- $H_1: P < P_0$ نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت:

$$Z = \frac{\bar{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} < -Z_{\propto}$$

- $H_1: P > P_0$ نرفض H_0 عند مستوى الدلالة $\propto$ إذا كانت:

$$Z = \frac{\bar{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} > Z_{\propto}$$

مثال: لاختبار حدس الإدارة أن 60% من الموظفين ستفضل نظام المنح الجديد، تسحب عينة من 150 موظف، ويؤخذ رأيها إذا كانوا يفضلونه أم لا، وكان 55 موظف من 150 موظف يفضل نظام المنح الجديد عند مستوى الدلالة $\propto = 0.01$.

الحل: سنختبر الفرضية:

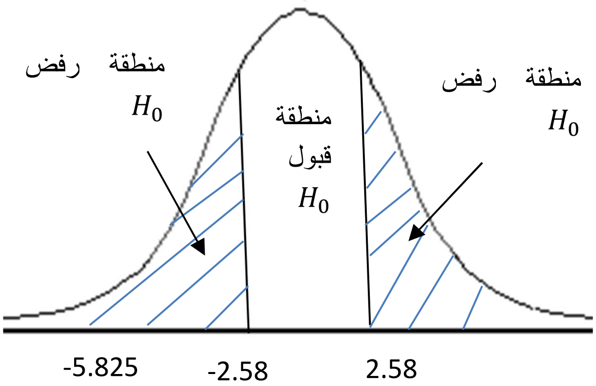
$$H_0: P = 0.60$$
$$H_1: P_1 \neq 0.60$$

بما أن $\propto = 0.01$ والفرضية البديلة ذو طرفين فإن القيم الحرجة هي $Z_{\frac{\propto}{2}} = 2.58$ و $-Z_{\frac{\propto}{2}} = -2.58$

$$\bar{P} = \frac{55}{150} = 0.367$$

$$Z = \frac{\bar{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} = \frac{0.367 - 0.6}{\sqrt{\frac{0.6 \times 0.4}{150}}} = -5.825$$

عند $(\propto= 0.01)$ نجد $Z < -2.58$ لأن قيمتها -5.825 وقعت في منطقة رفض H_0، وهذا يعني قبول الفرضية البديلة H_1 أي أن 60% من الموظفين لا تفضل نظام المنح الجديد.[1]



خامسا: اختبار الفروض حول الفرق بين نسبتي المجتمعين.

إذا أخذت عينتين عشوائيتين مستقلتين من توزيع ذي الحدين وكان حجمهما n_1،n_2 كبير بدرجة كافية فتكون الفرضية الصفرية: $H_0: P_1 - P_2 = D_0$ وبالتالي إحصاء الاختبار: $Z = \frac{(\bar{P}_1 - \bar{P}_2) - D_0}{\sqrt{\frac{\bar{P}_1 \bar{q}_1}{n_1} + \frac{\bar{P}_2 \bar{q}_2}{n_2}}}$

وما يجب الإشارة إليه هو أنه إذا كانت $D_0 = 0$ في الفرضية الصفرية فهذا يعني $P_1 = P_2$ فعندئذ يكون إحصاء الاختبار: $Z = \frac{(\bar{P}_1 - \bar{P}_2)}{\sqrt{\frac{\bar{P}\bar{q}}{n_1} + \frac{\bar{P}\bar{q}}{n_2}}}$

$$\bar{P} = \frac{n_1 \bar{P}_1 + n_2 \bar{P}_2}{n_1 + n_2} = \frac{n_1 \frac{X_1}{n_1} + n_2 \frac{X_2}{n_2}}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2}$$

وتشير $\bar{P}$ المتوسط المرجح لنسبتي العينتين $\bar{P}_1$،$\bar{P}_2$ ويطلق عليه أيضا النسبة التجميعية.

والفرضية البديلة هي:

$H_1: (P_1 - P_2) < 0$ أو $H_1: (P_1 - P_2) > 0$ أو $H_1: (P_1 - P_2) \neq 0$

[1]– وحيد مصطفى أحمد، مرجع سابق، ص، ص: 236-237.

مثال: من أجل المقارنة بين نسبة المدخنين في الفئة العمرية 18−25 سنة مع الفئة العمرية (26−30 سنة)، أخذت عينة عشوائية حجمها 200 من الفئة الأولى فوجد أن 80 منهم يدخنون وأخذت عينة عشوائية من الفئة العمرية الثانية وحجمها 100 فوجد أن 52 منهم يدخنون.

المطلوب: اختبر الفرضية $H_0: P_1 = P_2$ مقابل $H_1: P_1 < P_2$ عند مستوى الدلالة $(\propto = 0.01)$.

الحل:

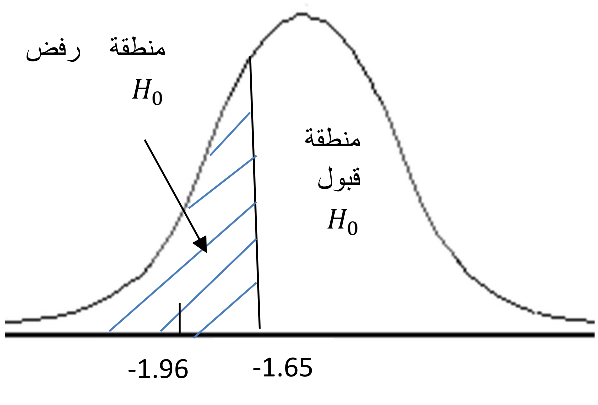
$$H_0: P_1 = P_2$$
$$H_1: P_1 < P_2$$

نلاحظ أن الفرضية البديلة ذات طرف واحد أيسر وعند $\propto = 0.01$ نجد القيمة الحرجة: $-Z_{0.05} = -1.65$ ومن أجل إيجاد إحصاءة الاختبار:

$$Z = \frac{(\bar{P}_1 - \bar{P}_2)}{\sqrt{\frac{\bar{P}\bar{q}}{n_1} + \frac{\bar{P}\bar{q}}{n_2}}}$$

يجب إيجاد $\bar{P}_2$،$\bar{P}_1$، $\bar{P}$

$$\bar{P}_1 = \frac{X_1}{n_1} = \frac{80}{200} = 0.40$$
$$\bar{P}_2 = \frac{X_2}{n_2} = \frac{52}{100} = 0.52$$
$$\bar{P} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{80 + 52}{200 + 100} = 0.44$$
$$Z = \frac{0.4 - 0.52}{\sqrt{\frac{(0.44 \times 0.56)}{200} + \frac{(0.44 \times 0.56)}{100}}} = -1.96$$



نلاحظ أن $-1.65 > -1.96$ أي أن قيمة Z تقع في منطقة رفض الفرضية H_0، لذلك نرفض H_0 ونقبل H_1 عند مستوى المعنوية $\propto = 0.05$ ومنه نجد أن نسبة المدخنين في الفئة العمرية الأولى أقل من نسبة المدخنين في الفئة العمرية الثانية.[1]

سادسا: اختبار الفروض حول تباين المجتمع γ^2

إذا كان لدينا مجتمع له توزيع طبيعي ورغبنا باختبار الفرض:

$H_0: \gamma^2 = \gamma_0^2$ و $H_1: \gamma^2 \neq \gamma_0^2$ أو $H_1: \gamma^2 > \gamma_0^2$ أو $H_1: \gamma^2 < \gamma_0^2$

فإننا نستخدم إحصاءة الاختبار $X^2 = \frac{(n-1)S^2}{\gamma_0^2}$

مثال: مدير مصنع لبطاريات السيارات يزعم أن عمر البطاريات المنتجة يتبع التوزيع الطبيعي بانحراف معياري 0.9 سنة، سحبنا عينة عشوائية حجمها n=10 من هذه البطاريات فوجدنا الانحراف المعياري لها S=1.2

اختبر الفرض:

$$H_0: \gamma = 0.9$$
$$H_1: \gamma > 0.9$$

بمستوى معنوية $\propto = 0.05$ ثم $\propto = 0.01$

الحل: نختبر الفرضية

$H_0: \gamma = 0.9$ و $H_1: \gamma > 0.9$

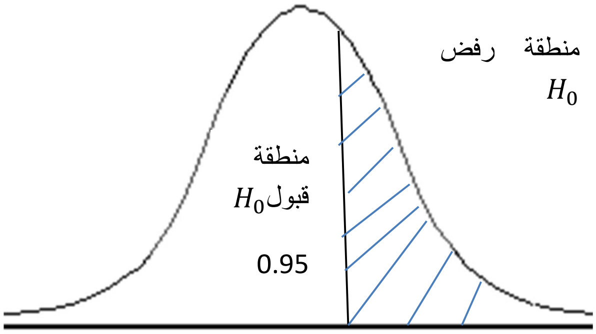
$H_0: \gamma^2 = 0.81$ و $H_1: \gamma^2 > 0.81$

نجد قيمة إحصاءة الاختبار X^2 الحسابية:

$$X^2 = \frac{(n-1)S^2}{\gamma_0^2} = \frac{9 \times 1.44}{0.81} = 16$$

نجد X^2 المعيارية بدرجة حرية n−1=9 عند $\propto = 0.05$، $X^2_{0.05,9} = 16.9$

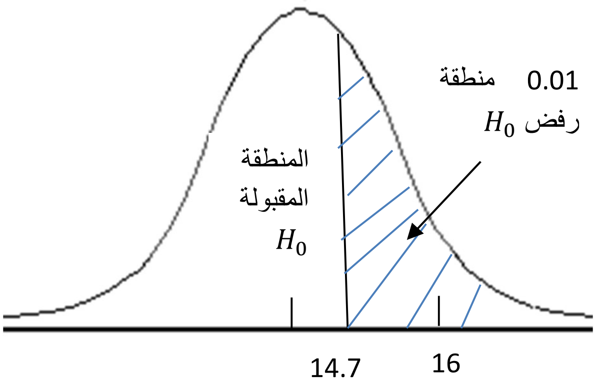
[1] محمد قريشي، مرجع سابق، ص، ص: 220-223.



عند مستوى معنوية $\propto= 0.05$ نجد القيمة المعيارية أكبر من القيمة المحسوبة أي نقبل H_0 ونرفض H_1 أي نقبل زعم المدير بأن الانحراف المعياري لعمر البطاريات يساوي 0.9 سنة عند مستوى معنوية $\propto= 0.05$.

- أما عند مستوى معنوية $\propto= 0.01$ فنجد X^2 المعيارية: $X^2_{0.01,9} = 14.7$

نرفض H_0 ونقبل H_1 بمستوى معنوية 0.01 ونقول أن الانحراف المعياري لعمر البطاريات المنتجة في هذا المصنع يزيد عن 0.9 سنة وذلك عند مستوى معنوية $\propto= 0.01$.[1]



سابعا: اختبار الفروض حول النسبة بين تبايني المجتمعين $\frac{\gamma_1^2}{\gamma_2^2}$

إذا أخذت عينة عشوائية حجمها n_1 من مجتمع طبيعي $N(\mu_1,\gamma_1^2)$ وأخذت عينة مستقلة عن الأولى حجمها n_2 من مجتمع طبيعي $N(\mu_2,\gamma_2^2)$ وكان S_1^2 تباين العينة الأولى، S_2^2 تباين العينة الثانية

[1]– أحمد عودة بن عبد المجيد عودة ومنصور بن عبد الرحمن القاضي، الإحصاء الوصفي والاستدلالي، (الجزء 2)، مرجع سابق، ص، ص:222–223.

فإن إحصاءة الاختبار للفرضية $H_0: \gamma_1^2 = \gamma_2^2$ هو $\frac{S_1^2}{S_1^2}$ الذي يخضع تحت فرض H_0 صحيحة لتوزيع F ذي درجات الحرية $(n_2 - 1)$ في البسط و$(n_1 - 1)$ في المقام أما خطوات اجراء الاختبار هي:

- الفرضية الصفرية والبديلة: $H_0: \gamma_1^2 = \gamma_2^2$

$H_1: \gamma_1^2 < \gamma_2^2$ أو $H_1: \gamma_1^2 > \gamma_2^2$ أو $H_1: \gamma_1^2 \neq \gamma_2^2$

- تحديد مستوى الدلالة $\propto$.
- إحصاء الاختبار تحت فرض H_0 صحيحة فإن $\frac{S_1^2}{S_1^2}$ يخضع لتوزيع $F_{(n_2-1, n_1-1)}$
- منطقة الرفض هناك ثلاث حالات هي:

أ. نرفض H_0 إذا كان: $\frac{S_1^2}{S_1^2} < F_{(\propto, n_2-1, n_1-1)}$

ب. نرفض H_0 إذا كان: $\frac{S_1^2}{S_1^2} < F_{(1-\propto, n_2-1, n_1-1)}$

ت. نرفض H_0 إذا كان $\frac{S_1^2}{S_1^2}$ خارج: $F_{(1-\propto, n_2-1, n_1-1)} > \frac{S_1^2}{S_1^2}$ و $\frac{S_1^2}{S_1^2} < F_{(\propto, n_2-1, n_1-1)}$

- نحسب $\frac{S_1^2}{S_1^2}$
- المقارنة: إذا وقعت $\frac{S_1^2}{S_1^2}$ في منطقة الرفض حسب الحالة المعنية نرفض H_0.

ملاحظة: $F_{(\frac{\propto}{2}, n_2-1, n_1-1)}$ غير مجدولة في جداول F عندما تكون $\propto$ صغيرة وهي الحالات المستعملة دائما ولذلك إما أن نحسب هذه القيمة من المعادلة: $F_{(\frac{\propto}{2}, n_2-1, n_1-1)} = \frac{1}{F_{(1-\frac{\propto}{2}, n_2-1, n_1-1)}}$.[1]

وما يجب الإشارة اليه أيضا أن:

- $F_{(\frac{\propto}{2}, V_1, V_2)}$ هي نقطة على المحور الأفقي لتوزيع فيشر ذي درجتي الحرية V_1, V_2 والتي يقع إلى يمينها مساحة $\frac{\propto}{2}$.
- $F_{(1-\frac{\propto}{2}, V_1, V_2)}$ هي نقطة على المحور الأفقي لتوزيع فيشر ذي درجتي الحرية V_1, V_2 والتي يقع إلى يمينها مساحة $(1 - \frac{\propto}{2})$.
- $F_{(\propto, V_1, V_2)}$ هي نقطة على المحور الأفقي لتوزيع فيشر ذي درجتي حرية V_1, V_2 والتي يقع إلى يمينها مساحة $\propto$.

[1]– محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام spss، مرجع سابق، ص، ص: 265–266.

- $F_{(1-\propto,V_1,V_2)}$ هي نقطة على المحور الأفقي لتوزيع فيشر ذي درجتي حرية V_1,V_2 والتي يقع إلى يمينها مساحة $(1-\propto)$.

مثال: إذا علمت أن مجتمع طول الطالبات، ومجتمع طول الطلبة في جامعة بسكرة يتوزع توزيعا طبيعيا، وسحبنا من المجتمع الأول عينة عشوائية 25 طالبة، ومن المجتمع الثاني عينة حجمها 21 طالبا، وكانت العينتان مستقلتان ووجدنا أن تباين الطول لعينة الطالبات يساوي 64، وتباين الطول لعينة الطلبة الذكور يساوي 36.

المطلوب: اختبر ما إذا كان هناك فرق بين تباين مجتمع طول الطالبات وتباين طول الطلبة الذكور وذلك باستخدام مستوى المعنوية $\propto = 0.05$.

الحل: سنختبر الفرضية:

$$H_0: \gamma_1^2 = \gamma_2^2$$
$$H_1: \gamma_1^2 \neq \gamma_2^2$$

بما أن الفرضية البديلة لا تحدد اتجاها واحدا فإن الاختبار المناسب ذو الطرفين والقيم الحرجة هي:

$$F_{(\frac{\propto}{2},V_1,V_2)} = F_{(0.025,24,20)} = 2.40$$

$$F_{(1-\frac{\propto}{2},V_1,V_2)} = F_{(0.975,24,20)} = 0.43$$

نرفض H_0 عند مستوى الدلالة $\propto = 0.05$ إذا كانت $F > 2.40$ أو $F < 0.43$

ومن ثم فإن منطقة قبول H_0 تقع بين القيمتين

سنقوم بحساب إحصاء الاختبار: $F = \frac{S_1^2}{s_2^2}\frac{\gamma_2^2}{\gamma_1^2} = \frac{64}{36} = 1.78$

نلاحظ أن قيمة إحصاء الاختبار F تقع في منطقة قبول H_0 عند مستوى الدلالة $\propto = 0.05$ ومنه الاختبار ليس له معنوية إحصائية أي أن تباين مجتمع طول الطالبات يساوي تباين مجتمع طول الطلبة ولا يوجد فرق حقيقي بينهما والفرق الظاهر بين تبايني العينتين هو فرق ليس ذو أهمية وسببه خطأ الصدفة.[1]

---

[1]- محمد قريشي، مرجع سابق، ص، ص: 213-215.

ثامنا: سلسلة تمارين حول اختبار الفروض.

- **التمرين الأول:** مصنع لإنتاج وتعبئة التمور أردت اختبار آلة من آلات التعبئة والتي هي مصممة لملأ صندوق $\mu = 120$، وأراد المصنع أن يتأكد من سلامة الآلة ومن أجل ذلك تم سحب عينة حجمها 100 صندوق، ووجد أن الوسط الحسابي للعينة $\bar{X} = 118.5$ والانحراف المعياري S=5.

المطلوب: اختبر الفرض القائل بأن متوسط التعبئة يختلف عن 120كغ وذلك عند مستوى المعنوية $\propto = 0.01$.

- **التمرين الثاني:** تخضع أعداد حبات التفاح على شجرة التفاح في بستان كبير لتوزيع طبيعي وسطه 150 حبة بدأ مالك البستان استعمال نوع جديد من السماد وأراد أن يختبر ما إذا زاد الإنتاج، ومن أجل ذلك أخذ عينة حجمها n=64 شجرة، فوجد أن الوسط الحسابي لأعداد الحبات في العينة 156 بانحراف معياري 12 حبة.

المطلوب: هل تشير هذه البيانات إلى زيادة في الإنتاج عند مستوى الدلالة $\propto = 0.05$.

- **التمرين الثالث:** في قسم علوم التسيير تم سحب عينة حجمها n=9 طلاب وتم حساب الوسط الحسابي والانحراف المعياري لمواظبة حضور الطلبة المحاضرات في مادة الإحصاء في العينة فوجد $\bar{X} = 0.8$ و $\gamma = 3$

اختبر الفرض القائل أن متوسط حضور الطالب للمحاضرات في الإحصاء في المجتمع المسحوب منه العينة أقل من 6.5 بدرجة ثقة 95%.

- **التمرين الرابع:** إذا كان لدينا نوعين من المنتجات الكهرو منزليين سحبنا عينة عشوائية من النوع الأول حجمها 36 جهاز ومن النوع الثاني نفس الحجم، فوجدنا أن: س/$85 = \bar{X}_1$ و $\gamma = 4$ و$\bar{X}_2 = 81$ و $\gamma_2^2 = 25$.

المطلوب:

أ. هل نستطيع أن نستنتج بمستوى معنوية $\propto = 0.05$ أن متوسط عمر النوع الأول يزيد عن متوسط النوع الثاني بساعتين على الأكثر مع العلم أن $\gamma_1 = \gamma_2 = 4$

ب. هل نستطيع أن نستنتج بمستوى معنوية $\propto = 0.05$ أن الفرق بين عمر النوعين هو 2سا.

- **التمرين الخامس:** لنفرض أنه تم سحب عينتين عشوائيتين من مجتمعين طبيعيين فكانت نتائجهما:

$$\bar{X}_1 = 30 \text{ و } n_1 = 16 \text{ و } S_1^2 = 9$$

$$\bar{X}_2 = 24 \text{ و } n_2 = 9 \text{ و } S_2^2 = 4$$

المطلوب: عند مستوى المعنوية $\propto = 0.05$ اختبر الفرضية

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

أ- حالة تبايني المجتمعين مجهولين ومتساويين.

ب- حالة تبايني المجتمعين مجهولين وغير متساويين.

- **التمرين السادس:** تظهر دراسة أن 50% من الطلاب في السنة الأولى يفضلون مادة الرياضيات، ومن أجل التأكد تم سحب عينة في 120 طالب من أجل أخذ رأيهم حول إذا كانوا يفضلون مادة الرياضيات أولا، وكان 60 طالب من بين 120 طالب يفضل مادة الرياضيات.

اختبر الفرضية عند مستوى دلالة ($\propto = 0.05$).

$$H_0: P = 0.50 \text{ و } H_1: P < 0.50$$

- **التمرين السابع:** عينة من 400 عائلة من مدينة بسكرة، وعينة من 500 عائلة من ولاية باتنة اختبرت عشوائيا، قام باحث بدراسة لمعرفة عدد العائلات التي تمتلك الانترنت، حيث في مدينة بسكرة وجد أن 48 من 400 تمتلكها، بينما 120 من 500 في باتنة تمتلك الانترنت، اختبر فرضية ما إذا كانت نسبة مالكي الانترنت في مدينة بسكرة يساوي النسبة في ولاية باتنة عند مستوى دلالة ($\propto = 0.05$).

$$H_0: P_1 = P_2$$
$$H_1: P_1 \neq P_2$$

- **التمرين الثامن:** سحبنا عينة حجمها 15 طالبة من قسم علوم التسيير فوجدنا أن الانحراف المعياري لنقاط هاته الطالبات هو 5 نقاط إذا علمت أن نقاط الطالبات تتبع التوزيع الطبيعي فاختبر الفرضية عند مستوى المعنوية 0.05 ثم 0.01:

$$H_0: \gamma = 8$$
$$H_1: \gamma \neq 8$$

- **التمرين التاسع:** بفرض أن لدينا البيانات التالية:

$n_1 = 9$ و $n_2 = 16$ و $S_1^2 = 25$ و $S_2^2 = 16$

اختبر الفرضية:

$$H_0: \gamma_1^2 \neq \gamma_2^2$$
$$H_1: \gamma_1^2 \neq \gamma_2^2$$

عند مستوى دلالة $\propto = 0.01$.

قائمة المراجع

قائمة المراجع

1. أحمد السيد عامر: الإحصاء الوصفي والتحليلي، دار الفجر للنشر والتوزيع، القاهرة، مصر، مصر، 2007
2. أحمد عودة بن عبد المجيد عودة ومنصور عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي (الجزء 1)، مكتبة الفلاح للنشر والتوزيع، الطبعة 3، الكويت، 2014
3. أحمد عودة بن عبد المجيد عودة، منصور بن عبد الرحمن القاضي: الإحصاء الوصفي والاستدلالي، الجزء (2) الإحصاء الاستدلالي، مكتبة الفلاح للنشر والتوزيع، الطبعة 3، 2014،
4. إدوارد مينيكا وزوريانا كورزيجا: الإحصاء في الإدارة مع التطبيق على الحاسوب الالي، (الكتاب الأول)، ترجمة سرور علي إبراهيم سرور، دار المريخ للنشر والتوزيع، الرياض، السعودية
5. برنارد تايلور الثالث: مقدمة في علم الإدارة (الجزء الثاني)، ترجمة سرور علي إبراهيم وآخرون، دار المريخ للنشر، الرياض، السعودية
6. جورج كانافوس ودون ميلر: الإحصاء للتجاريين مدخل حديث ترجمة: سلطان محمد عبد الحميد ومحمد توفيق البلقيني، دار المريخ، الرياض، السعودية، 2004
7. جيلاطو جيلالي: الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الطبعة 9، الجزائر، 2012
8. خالد أحمد فرحان المشهداني، مبادئ الإحصاء والاحتمالات (الأسس النظرية والعملية)، دار الأيام للنشر والتوزيع، عمان، الأردن، 2015
9. دفيد أندرسون وآخرون: الأساليب الكمية في الإدارة، ترجمة: محمد توفيق البلقيني ومرفت طلعت المحلاوي، دار المريخ للنشر والتوزيع، المملكة العربية السعودية، 2006
10. دومينيك سالفاتور: ملخصات شوم نظريات ومسائل في الإحصاء والاقتصاد القياسي، ترجمة سعدية حافظ منتصر، الدار الدولية للاستثمارات الثقافية،القاهرة، 1997
11. سهيل أحمد سمحان ومحمد حسين الوادي: مبادئ الإحصاء للاقتصاد والعلوم الإدارية، دار الصفاء للنشر والتوزيع، الطبعة 2، عمان– الأردن، 2014
12. قسم الإحصاء بجامعة الملك عبد العزيز: مبادئ الإحصاء للتخصصات النظرية الإدارية والإنسانية، دار خوارزم العلمية للنشر والتوزيع، الطبعة 9، جدة، السعودية، 2015

13. محمد بوعلاق: الموجه في الإحصاء الوصفي والاستدلالي في العلوم النفسية والتربوية والاجتماعية، الطبعة 2، دار الأمل للطباعة والنشر والتوزيع، 2012

14. محمد حسين محمد رشيد: الإحصاء الوصفي والتطبيقي والحيوي، دار الصفاء للطباعة والنشر والتوزيع، الطبعة 2، عمان (الأردن)، 2015

15. محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء مبادئ وتحليل باستخدام SPSS، دار المسيرة للنشر والتوزيع، الطبعة 6، عمان–الأردن، 2012

16. محمد صبحي أبو صالح وعدنان محمد عوض: مقدمة في الإحصاء، ديوان المطبوعات الجامعية، الجزائر، 1984

17. محمد قريشي: الاحتمالات والإحصاء التطبيقي، الطبعة 2، دار علي بن زيد للطباعة والنشر، بسكرة، الجزائر، 2020

18. موساوي عبد النور وبركان يوسف: الإحصاء (الجزء 2)، دار العلوم للنشر والتوزيع، عنابة الجزائر، 2010،

19. وحيد مصطفى أحمد: أساسيات علم الإحصاء، دار الكتب للنشر والتوزيع، القاهرة، مصر