الجمهورية الجزائرية الديمقراطية الشعبية

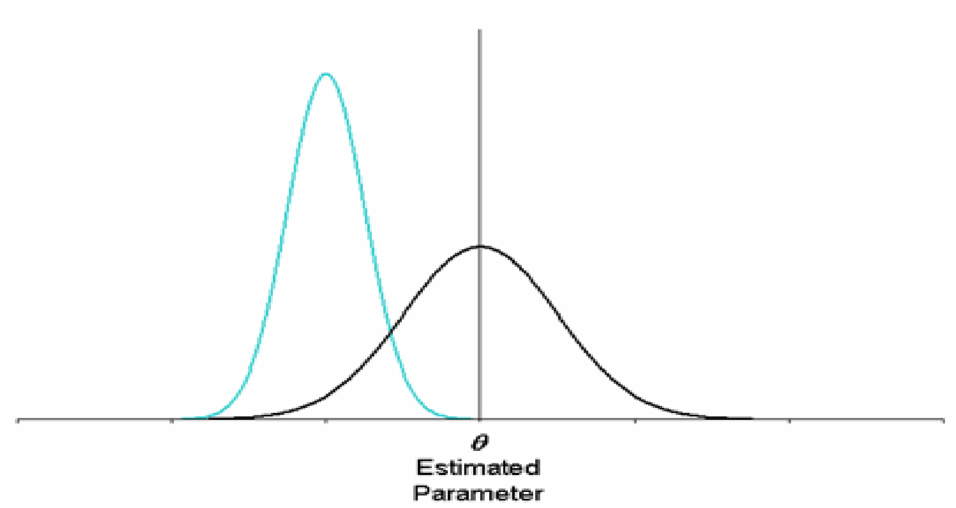
وزارة التعليم العالي والبحث العلمي

جامعة 20 أوت 1955

مطبوعة دروس في مادة:

الاحصاء 03

موجهة لطلبة: السنة ثانية علوم اقتصادية، مالية، تجارية، وعلوم التسيير.



من اعداد الدكتورة:

بوالشعور شريفة.

السنة الجامعية: 2018/2017.

تقديم عام:

هذه المطبوعة هي عبارة عن محاضرات في الإحصاء (03) أو الاحصاء التطبيقي، حيث أنه يمكن تقسيم علم الاحصاء إلى احصاء وصفي واحصاء استدلالي، هذا الأخير تتدرج ضمنه نظرية الاحتمالات، والاحصاء التطبيقي، وحسب مقرر التدريس المحدد من قبل وزارة التعليم العالي الجزائرية، فيتم التدرج في تدريس علم الاحصاء على ثلاث مراحل كما يلي: الاحصاء الوصفي "الذي يختص بتلخيص وتوصيف مجموعة من البيانات"[1] ويسمى ب الاحصاء (01)، والذي يتم تدريسه خلال الفصل الأول لطلبة السنة أولى جدع مشترك، والاحصاء (02) والذي يُعني بنظرية الاحتمالات، والتي تعتبر عنصرا أساسيا في الاستدلال الاحصائي، حيث تساعد نظرية الاحتمالات على تفسير وتحليل المعطيات بشكل أفضل، ويدرس هذا المقياس خلال الفصل الثاني لطلبة السنة أولى جدع مشترك، وأخيرا الاحصاء (03) أو الاحصاء التطبيقي وهو موجه لتدريس طلبة السنة ثانية (علوم اقتصادية، علوم التسيير، علوم تجارية، والعلوم المالية) خلال الفصل الأول.

وتمثل هذه المطبوعة ثمرة خبرة عدة سنوات في تدريس مقياس احصاء (03) لطلبة السنة ثانية بقسم العلوم الاقتصادية في جامعة سكيكدة، وفي محاولة لاستيفاء البرنامج المحدد تم تقسيم المطبوعة إلى الفصول التالية:

فصل تمهيدي: خصص هذا الفصل لمجموعة من المفاهيم النظرية المتعلقة بالمقياس. وعلى اعتبار أن هناك ثلاث أنواع من الأساليب الاحصائية مستخدمة في الاستدلال الاحصائي وهي: توزيعات المعاينة، تقدير فترات الثقة، واختبار الفرضيات، فقد تم تخصيص:

الفصل الأول: حول نظرية توزيع المعاينة، ويعتبر هذا الفصل حجر الأساس الذي ترتكز عليه باقي الفصول.

ثم **الفصل الثاني**: حول نظرية التقدير، وتحديد فترات الثقة، حيث تستخدم مجموعة من البيانات لتقدير مدى أو مجال لمعلمة من معالم المجتمع المجهولة.

أما **الفصل الثالث**: حول اختبارات الفرضيات الاحصائية، ويتم من خلال صياغة فرضية ما حول معلمة من معالم المجتمع، ثم استخدام عينة لاتخاذ قرار بخصوص صحة الفرضية من عدمها.

---

[1]- دومينيك سالفاتور، **"الاحصاء والاقتصاد القياسي"**، الطبعة الخامسة العربية، الدار الدولية للاستثمارات الثقافية، القاهرة، مصر، 2001، ص 01.

فهرس المحتويات:

الفصل التمهيدي: مدخل للإحصاء الاستدلالي.

مقدمة:

تزايدت أهمية علم الاحصاء في المجتمع المعاصر نتيجة الأهمية التي تكتسيها عمليات الاستقصاء التي تقوم بها المؤسسات الاقتصادية ومنظمات الأعمال، وذلك لتحقيق عدة أهداف كمعرفة توجهات المستهلكين، وأذواقهم.....إلخ. فالإحصاء هو علم استخدام البيانات لفهم العالم من حولنا. والأدوات الإحصائية تساعد على الاجابة على تساؤلات حول معالم التوزيع المجهولة. مثلا: ما هو متوسط مداخيل السكان؟ ماهي نسبة انتشار ظاهرة معينة في المجتمع؟ هل تختلف الدخول بين مجتمع الرجال ومجتمع النساء، وإذا كانت الاجابة بنعم، ما هو مقدار الاختلاف.......إلخ؟

يمكن ان تكون الاجابة على هذه التساؤلات من خلال دراسة كل فرد من أفراد مجتمع الدراسة (مثلا: قياس الدخل لكل فرد في المجتمع، أو قياس نسبة انتشار ظاهرة معينة أو مرض ما بين أفراد المجتمع...إلخ)، وهو ما يعرف **بالمسح الشامل**، إلا أن هذا النوع من الدراسات سيكون جد مكلف، ويتطلب مجهودات كبيرة، وفترة زمنية طويلة جدا. ومن الأمثلة الموجودة في الواقع العملي على المسح الشامل هو احصاء السكان الذي تقوم به بعض الدول كل بضعة سنوات مثلا: كل عشر سنوات. ويعتبر هذا النوع من الاحصاء جد مكلف، كما أن تحليل هذا الكم الهائل من البيانات يتطلب عشرات السنين. وهنا تتجلي أهمية الاحصاء في امكانية دراسة مجتمع ما من خلال اختيار عينة عشوائية منه، عوضا عن اجراء مسح شامل للمجتمع، يمكن القول أن 1000 فرد من المجتمع يشكلون عينة مختارة بطريقة عشوائية. وباستخدام الأساليب الاحصائية، وتطبيقها على العينة سنتمكن من التوصل إلى نتيجة تجريبية، تساعد على رسم استدلال احصائي حول معالم المجتمع الكلي.

ويمكن تقسيم علم الاحصاء إلى قسمين هما

الاحصاء الوصفي (Descriptive statistics): ويتضمن الطرق الإحصائية المستخدمة في جمع البيانات والمعلومات عن ظاهرة معينة او مجموعة ظواهر وكيفية تنظيم وتصنيف وتبويب هذه البيانات مع إمكانية عرضها في جداول ورسوم بيانية وحساب بعض المؤشرات الاحصائية.[1]

الاحصاء الاستدلالي (Statistical Inference): يشير الاستدلال الإحصائي إلى استنتاج معلومات حول المجتمع انطلاقا من جزء يكون مختارا من ذات المجتمع بطريقة علمية ما يعرف باسم العينة، وعليه فالاستدلال الاحصائي يعتمد على استخدام بيانات هذه العينة في التوصل إلى نتائج، يمكن تعميمها على مجتمع الدراسة.

[1]– عماد توما كرش، ولاء أحمد القزاز، وفاء يونس حمودي، **"علم الاحصاء"**، وزارة التعليم العالي والبحث العلمي، هيئة التعليم التقني، العراق، 2014، ص 10.

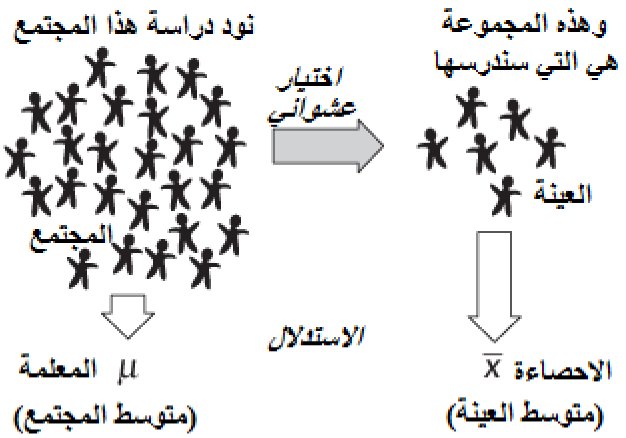
I/ 1. مفاهيم احصائية عامة:

I/ 1. 1. المجتمع والعينة.

يشير **المجتمع** إلى كافة العناصر المكونة لمجموعة من الأشياء التي يراد إخضاعها لتجربة (أو دراسة) ما[1]. وعليه المجتمع يمثل كافة القياسات، القيم، أو المفردات وليس الأفراد، أو الأشياء التي تم قياسها (مجتمع الأوزان، مجتمع آراء الناخبين....)، كما أن المجتمع قد يكون محدود (أي ممكن حصر عدد مفرداته، مثلا: مجموع الناخبين المسجلين في منطقة ما) أو غير محدود (من الصعب حصر عدد مفرداته، مثلا: نسبة تركيز الكبريت في نبع مائي) . ونرمز عادة لحجم المجتمع ب N.

أما **العينة** فهي مجموعة جزئية من المجتمع، أي أنها نموذج يشمل جزء من مفردات المجتمع الأصلي، ويرمز لحجم العينة ب n، وعادة ما تكون العينة محدودة الحجم. وهذه المجموعة الجزئية (العينة) تغنينا عن دراسة كل مفردات المجتمع خاصة في حالة المجتمعات الغير محدودة والتي يصعب حصر مفرداتها، والشكل التالي يقدم مثالا يوضح كيفية استخدام احصاءه العينة (متوسط العينة $\bar{X}$) للاستدلال حول معلمة المجتمع (متوسط المجتمع μ).

الشكل رقم (I-1): الاستدلال الاحصائي حول المجتمع باستخدام العينة.



المصدر: http://www.cliffsnotes.com/study-guides/statistics/sampling/populations-samples-parameters-and-statistics; 21/01/2018 a 20:40

---

[1] - عبد الرزاق بني هاني، **"الاقتصاد القياسي: المبادئ الرياضية والاحصائية"**، الطبعة الأولى، دار وائل للنشر، عمان، الأردن، 2014، ص 233.

1 /I. 2. تعريف المعلمة (Parameter).

يقصد بمعالم المجتمع مجموعة من الخصائص مثل: المتوسط μ، الانحراف المعياري σ،...إلخ. ومن خصائص المجتمع أيضا طبيعة توزيعه الاحتمالي $f(x)$ كأن يكون طبيعيا أو غيره.

1 /I. 3. تعرف الإحصائية (Statistics).

الإحصاءه أو الاحصائية هي متغير عشوائي أنشئ انطلاقا من مشاهدات العينة وصمم لتقدير معلمة محل الاهتمام[1].

وعليه يمكن القول أن المعلمة تعبر عن خاصية من خصائص المجتمع، بينما تمثل الاحصائية خاصية من خصائص العينة.

I /2. تعريف المعاينة:

هي عملية اختيار (*Process*) عدد كاف من عناصر المجتمع، بحيث يتمكن الباحث من خلال دراسة العينة المختارة وفهم خصائصها من تعميم هذه الخصائص على عناصر المجتمع الأصلي، ولابد أن نتذكر دوما بأن ناتج عملية المعاينة هو العينة (*Sample*) المرغوب بها.[2]

3 /I. أنواع العينات:

هناك عدة أنواع من العينات تختلف باختلاف طريقة المعاينة (طريقة سحبها)، ووفقا لدرجة دقتها وتمثيلها للمجتمع الأصلي، وعموما نميز بين عينات احتمالية، وعينات غير احتمالية.

1 .3 /I. المعاينة الاحتمالية:

نقول عن عينة أنها احتمالية إذا كانت طريقة اختيار العينة عشوائية، أي أنها المعاينة التي تعطي لكل مفردة من مفردات المجتمع احتمال معين (معروف) لأن تكون في العينة، وهذه العينة من أكثر العينات استعمالا وتمتاز بكونها ممثلة للمجتمع الاحصائي بشكل جيد، وهذا النوع بدوره ينقسم إلى معاينة غير مقيدة ومعاينة مقيدة.

أولا: معاينة غير مقيدة: أي المعاينة التي لا تستند إلى أي قيد أو شرط، وتسمح المعاينة الغير مقيدة بالحصول على **عينة عشوائية بسيطة.**

---

[1]- Kairat Mynbaev ; "**Companion for \Statistics for Business and Economics**", by Paul Newbold, William L. Carlson and Betty Thorne", MPRA Paper No. 23069, posted 6. June 2010 16:00, Online at http://mpra.ub.uni-muenchen.de/23069/. P81.

[2] - فايز جمعة النجار، نبيل جمعة النجار، ماجد راضي الزعبي، **"أساليب البحث العلمي: منظور تطبيقي"**، الطبعة الثانية، دار الحامد للنشر والتوزيع، الأردن، 2010، ص 111.

أ. **العينة العشوائية البسيطة:**

نقول عن عينة أنها عشوائية إذا كان لكل مفردة في المجتمع نفس الاحتمال لأن تكون في العينة، وهذه العينة من أكثر العينات استعمالا وتخص المجتمعات المتجانسة (قد يصعب تحقيق ذلك في الواقع)، أي تلك التي لها خصائص متشابهة أو متقاربة، مثلا: مجتمع الطلبة متجانس من حيث السن. ويمكن الحصول على عينة عشوائية عن طريق القرعة مثلا. ونميز بين نوعين من المعاينة العشوائية باختلاف طريقة السحب: **معاينة نفادية، ومعاينة غير نفادية.**

- **العينة الغير نفاديه:** عندما يكون السحب بالإرجاع حيث يمكن أن تظهر المفردة أكثر من مرة في العينة، وتسمى هذه العينة غير نفاديه لأن تكرار العملية لا يؤدي إلى تقليص عدد المفردات في المجتمع. وفي هذه الحالة نستخدم القانون التالي لإيجاد عدد العينات الممكنة N^n.
- **العينة النفادية** تسمى المعاينة بدون ارجاع معاينة نفاديه حيث لا يمكن أن تظهر المفردة في العينة أكثر من مرة واحدة. وفي هذه الحالة نستخدم القانون التالي لاستخراج العينات الممكنة: $C_N^n = \frac{N!}{(N-n)!n!}$

مثال: ليكن لدينا المجتمع التالي {1، 3، 5}، أحسب عدد العينات ذات الحجم $n=2$ التي يمكن استخراجها ثم حددها: 1) في حالة السحب بالإرجاع، 2) في حالة السحب بدون ارجاع.

الحل:

1) استخراج كل العينات الممكنة في حالة السحب بالإرجاع:

$N^n = 3^2 = 9$، العينات الممكنة هي 9 عينات.

(1،1)، (3 ،3)، (5 ،5)، (3 ،1)، (5 ،1)، (1 ،3)، (5 ،3)، (1 ،5)، (3 ،5).

2) استخراج كل العينات الممكنة في حالة السحب بدون إرجاع:

$$C_N^n = \frac{N!}{(N-n)!\,n!} \Rightarrow C_3^2 = \frac{3!}{(3-2)!\,2!} = \frac{3 \times 2 \times 1}{1 \times 2} = \boxed{\text{3 عينات}}$$

(3 ،1)، (5 ،1)، (5 ،3).

إلا أنه في بعض الحالات التي يكون فيها حجم المجتمع غير منته، أو حجم N و n كبير جدا، يصعب تكون العينات العشوائية بالطريقة السابقة، لذلك يلجأ الباحثون والاحصائيون إلى استخدام ما يعرف بالأرقام العشوائية أو الجداول العشوائية (أنظر الملحق). ولاستخدام هذه الجداول للحصول على عينة حجمها n من مجتمع عدد مشاهداته N، فإنه يتم ترقيم عناصر المجتمع من 1 إلى N، فإنه يتم بطريقة

عشوائية اختيار أي جدول من جداول الأرقم العشوائية، ثم نختار أحد الأرقام في الجدول بطريقة عشوائية أيضا، ثم يتم التحرك بشكل عمودي (في العمود) أو أفقي (في الصف)، وذلك لاختيار مفردات العينة بشكل عشوائي وبالتالي نحصل على عينة عشوائية بسيطة.

مثال: على فرض أن هناك شركة ما تمارس نشاطها الانتاجي على مستوى الجزائر العاصمة، وتود التعريف بمنتجها الجديد وتسويقه بطريقة عشوائية في مختلف مناطق الوطن، فقامت باختيار عينة مكونة من 24 ولاية على المستوى الوطني، وذلك باستخدام جدول الأعداد العشوائية لتحديد الولايات المعنية بتسويق منتوجها.

الحل: من المعروف أن اجمالي الولايات في الجزائر يقدر ب 48 ولاية حيث أن كل ولاية تحمل ترقيم معين من 1 إلى 48.

باستخدام أعداد الجداول العشوائية قام المسؤول عن التسويق في هذه الشركة باختيار القيم التالية:

{44، 39، 26، 19، 40، 15، 12، 29، 38، 23، 36، 45، 30، 31، 08، 34، 41، 42، 16، 05، 48، 06، 02، 04}

وبالتالي نكون قد حصلنا على عينة من 24 ولاية بطريقة عشوائية.

(عين الدفلى، الوادي، المدية، سطيف، خنشلة، تيزي وزو، تبسة، معسكر، تسمسيلت، عنابة، الطارف، النعامة، ورقلة، وهران، بشار، برج بوعريريج، سوق أهراس، تيبازة، الجزائر العاصمة، باتنة، غيليزان، بجاية، الشلف، أم البواقي).

ثانيا: المعاينة غير المقيد:

وهي المعاينة التي تتم تحت بعض القيود، والمعاينة الاحتمالية المقيدة تسمح بالحصول على 4 أنواع من العينات كما يلي:

أ. العينة العنقودية:

وهي نوع من أنواع المعاينة العشوائية البسيطة، إلا أنه في هذه الطريقة يقع الاختيار على مجموعات (عناقيد) حيث يقسم المجتمع إلى مجموعات مستقلة عن بعضها البعض كل مجموعة تضم عددا من أفراد العينة بدلا من اختيار الأفراد مباشرة, أي أن وحدة الاختيار هو المجموعة وليس الفرد.

مثال: فإذا أراد باحث أن يدرس نسبة نجاح طلاب جامعة ما فبدلا من أن يختار عينته من خلال قائمة تضم جميع الطلاب فإنه يختار عينة من الكليات ثم يدرس هؤلاء الطلاب الموجودين داخل هذه الكليات.

وأحيانا للحصول على هذا النوع من العينات يجب المرور على مرحلتين أو أكثر من مرحلة. **مثلا**: بالعودة إلى المثال السابق يمكن أن يختار الباحث في المرحلة الأولى عددا من الكليات كعناقيد ودراسة معدلات النجاح فيها وهنا تكون **عينة عنقودية بمرحلة واحدة.** أما اذا سحبنا من كل كلية مجموعة من الأقسام عشوائيا فتكون عندها **عينة عنقودية بمرحلتين** واذا تابعنا السحب و أخذنا عشوائيا من كل قسم مجموعة من التخصصات سنحصل على **عينة عشوائية متعددة المراحل.**

ب. **العينة الطبقية**:

تخص المجتمعات غير المتجانسة أي المجتمعات المتكونة من عدة فئات اجتماعية -مهنية وهذا بناء على عدة اعتبارات منها: مستوى الدخل، مستوى الانفاق، المستوى التعليمي، ...إلخ. إن العينة الطبقية هي عبارة عن مجموعة من العينات العشوائية.[1]

مثال: عن العينة الطبقية:

ليكن المجتمع الأصلي هو كلية العلوم الاقتصادية، التجارية وعلوم التسيير المتكونة من 1000 طالب، وحجم العينة الطلوبة هو 200 طالب، فاذا كانت شرائح أو طبقات هذا المجتمع هي مختلف الأقسام المكونة للكلية، والتي عددها أربعة أقسام (قسم الاقتصاد +قسم التسيير +قسم العلوم التجارية +قسم العلوم المالية)،فإنه سيتم أخذ 50 طالب من كل قسم (200÷4)، وبذلك نحصل على عينة طبقية.

أما في حالة **العينة الطبقية التناسبية** فإنه لا يتم تحديد حجم العينة على أساس متساوي من كل طبقة من طبقات المجتمع، وحتى تكون العينة أكثر مصداقية وأكثر تمثيلا للمجتمع بحيث أن العدد المختار من طبقة يجب أن يتناسب مع حجمها الفعلي، وتمثيلها داخل المجتمع، لذلك فإن اختيار مفردات عينة طبقية تناسبية يمر بالمراحل التالية:

1. تحديد طبقات المجتمع حسب متغير أو متغيرات
2. تحديد حجم كل طبقة في المجتمع

---

[1]– جلاطو جيلالي، "**الاحصاء: مع تمارين ومسائل محلولة**"، الطبعة الخامسة، ديوان المطبوعات الجامعية، الجزائر، 2005، ص 06.

3. تحديد العدد المناسب من كل طبقة بحيث يكون هناك تناسب بين ما نختاره من هذه الطبقة للبحث مع حجمها في المجتمع
4. اختيار العدد المحدد من كل طبقة بالطريقة العشوائية

والمثال التالي يوضح طريقة تكوين العينة الطبقية التناسبية:

على فرض أن المجتمع الأصلي هو (1000) مشاهدة وكان تمثيلهم كما يلي: (طلبة الاقتصاد 340+ طلبة التسيير 260+ طلبة التجارة 220+ طلبة المالية 180)=1000 مشاهدة، واذا كان حجم العينة المراد اختيارها هو (200) مشاهدة، فإن تمثيلهم في العينة الطبقية التناسبية سيكون كالآتي:

1000 حجم المجتمع÷ 200 حجم العينة = 5 (هذا الرقم يمثل أساس التقييم)، وستكون العينة كما يلي:

1. طلبة الاقتصاد: 340÷5=68
2. طلبة التسيير: 260÷5= 52
3. طلبة التجارة: 220÷5= 44
4. طلبة المالية: 180÷5= 36

200 =36+44+52+68

ت. **العينة المنتظمة العشوائية**:

يتم اختيار أو سحب العينة المنتظمة من عناصر مجتمع يكون مرتبا بشكل تصاعدي أو تنازلي، ويتم اختيار مفردات هذه العينة بعد تحديد ما يلي:

أولا: فترة السحب أو الزيادة المنتظمة (k)، ونقصد بذلك المسافة الفاصلة بين كل اختيار وآخر ونحصل عليها بقسمة حجم المجتمع الكلي على حجم العينة المراد تكوينها $\left(k = \frac{N}{n}\right)$.

ثانيا: نقطة البداية، وتشير إلى النقطة التي تنطلق أو تبدء بها العينة، وهي قيمة نختارها عشوائية حيث تكون محصورة بين 1 وطول فترة السحب k (الزيادة المنتظمة).

وتسمى عينة منتظمة لأنه يتم اختيار مسافة ثابتة بين كل رقم والرقم الذي يليه، ويمكن توضيح الطريقة التي يتم بها اختيار مفردات العينة المنتظمة من خلال المثال التالي:

اذا كان المجتمع الاحصائي يتكون من 35 مفردة والمطلوب الحصول على عينة حجمها 5 فيكون توزيع الوحدات الكلية للأصلية للمجتمع كما يلي:

حجم المجتمع N÷ حجم العينة n = 35/5=7 وعلى هذا الأساس يتحدد الرقم الأول للعينة حيث يكون محصورا بين 1 و 7 مثلا: ليكن الرقم 2، ثم نبدأ بتوزيع العينة على باقي المفردات كما يلي: العنصر الأول في العينة هو الرقم 2، والرقم الثاني سيكون 9 (7+2=9)، والرقم الثالث هو 16، والرقم الرابع هو 23، والعنصر الأخير في العينة سيكون 30.

كما يمكن أن تكون نقطة انطلاق العينة 1 أو 3 أو....أو 7 وبالتالي يمكن أن نشكل العينات الموضحة في الجدول أدناه.

العينة 01	العينة 02	العينة 03	العينة 04	العينة 05	العينة 06	العينة 07
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35

I/ 3. 2. المعاينة غير الاحتمالية:

المعاينة غير الاحتمالية هي التي يتم فيها اختيار مفردات العينة بطريقة غير عشوائية، أي أنها معاينة تعتمد على معايير شخصية وغير موضوعية، ومن أهم طرق اختيار العينات الغير احتمالية ما يلي:

أ. **العينة الميسرة**:

يعتبر هذا النوع من أسهل أنواع المعاينة، حيث يتم اختيار مفردات العينة حسب ما هو متوفر أو حسب ما تيسر، كأن يقوم الباحث باختيار الأفراد الذين يتم مقابلتهم بشكل عرضي أو صدفة في مكان ما وفترة زمنية معينة. ويعاب على هذا النوع افتراض تجانس جميع مفردات الدراسة.

مثال: اذا أراد صاحب بنك خاص التعرف على مدى رضى العملاء على الخدمات المقدمة من طرف الموظفين في البنك، فما عليه إلا أن يسأل أول 40 عميل يصادفهم عند البوابة الرئيسية للبنك ليتعرف على آرائهم في أداء الموظفين في ذلك البنك.

ب. **العينة العمدية أو الغرضية (القصدية)**:

اختيار العينة بشكل متعمد يعتقد الباحث مسبقا بان مفردات هذه العينة هي خير من يمثل مجتمع الدراسة[1]. ويمكن اعتبار باقي طرق اختيار العينة التالية على أنها أنواعاً فرعية من العينة القصدية (عينة الحصة، وكرة الثلج).

[1] - عماد توما كرش، ولاء أحمد القزاز، وفاء يونس حمودي، مرجع سبق ذكره، ص 20.

ت. العينة الحصصية (العينة الطبقية التناسبية)[1]:

تستخدم العينة الحصصية في مقابلات المعاينة وهي غير عشوائية تمام، وتقوم على افتراض أن العينة تمثل المجتمع وأن التغير بالنسبة لمتغيرات العينة الحصصية هي نفسها بالنسبة لمتغيرات المجتمع، لذا فإن العينة الحصصية هي نوع من العينة العشوائية الطبقية ولكنها تختار أفراد الطبقة بطريق غير عشوائي، اذ تعتمد على تقسيم المجتمع إلى مجموعات خاصة، ثم حساب حصة كل مجموعة اعتمادا على علاقتها بالبيانات المتوفرة وحجم المجتمع ثم الحصول على تلك الحصة بأيسر الطرق (Saunders et al, 2007, 226)

وتستخدم العينة الحصصية عندما يكون هناك صفات محددة يجب أن تؤخذ مسبقا بالاعتبار في العينة مثل: (الجنس، الوظيفة، التوزيع الجغرافي)، اذا لابد في هذه الحالة من توزيع العينة بالحصة على المجتمع لتمثل التنوع بداخله.

مثل: اذا أردنا توجيه سؤال معين إلى مجموعة من العمال والعاملات في الشركة المتحدة، وتبين أن العمالة في الشركة تتكون من 30% من العمال الذكور، بينما نسبة 70% من العمالة اناث، وتقرر أن يكون اجمالي العينة 10 أشخاص. فإننا سنوجه السؤال إلى 3 عمال ذكور، و7 عاملات إناث تتم مواجهتهم في ظروف مريحة وبصورة كيفية دون الاعتماد على الأسلوب العشوائي، ليصبح مجموع العينة (3+ 7 = 10) أشخاص.

ث. كرة الثلج:

تستخدم عينة كرة الثلج غالبا عندما يكون من الصعب الوصول إلى أفراد المجتمع المرغوب دراسته، وبالتالي عدم وجود قائمة معلومة ومحددة يمكن استخدامها لتحديد مفردات العينة، وعليه فإن الباحث يبدأ بعينة صغيرة ميسرة، ثم تبدأ العينة بالكبر شيئا فشيئا كأن يتم اختيار شخص يستوفي المواصفات الموضوعة للاختيار ضمن العينة ثم يطلب منه أن يقترح آخرين بنفس المواصفات.

مثل: متعاطي المخدرات أو ضحايا العنف أو المشردين ...إلخ حيث لا توجد قوائم محددة تحتوي أسماء هاته الفئات. وعليه فطريقة الحصول على العينة تبدأ بالتعرف على عضو منهم وبعد إجراء الدراسة معه واطمئنانه بالأمر يطلب منه الباحث مساعدته في إيصاله بأفراد من جماعته يحملون نفس الخواص. وهكذا تتضاعف العينة وتكبر كما تكبر كرة الثلج.

[1] - فايز جمعة النجار، نبيل جمعة النجار، ماجد راضي الزعبي، ص 119.

تمارين حول الفصل (السلسلة الأولى).

التمرين 1:

1. ماهي المعلمة، وما هي الاحصائية وما الفرق بينهما؟
2. لماذا نلجأ للعينة بدل دراسة المجتمع كاملا؟
3. ما هي العينة التي تمنح لكل مفردة من مفردات المجتمع نفس احتمال الظهور في العينة؟
4. ما هو الفرق بين العينة الطبقية، والعينة الحصصية؟

التمرين 2:

في مصنع ما يتم اختيار مجموعة من العمال عشوائيا كل سنة بهدف ارسالهم لأداء مناسك العمرة، اذا علمت أن عدد عمال المصنع يقدر ب 537 عامل، استخرج عينة المستفيدين المكونة من 8 عمال وذلك باستخدام الجداول العشوائية.

التمرين 3:

باستعمال المعاينة المنتظمة العشوائية، اشتق عينة مكونة من 10 طلاب من بين 90 طالبا متفوقا في جامعة سكيكدة لتمثيلها في مسابقة وطنية.

التمرين 4:

عين عدد المنشآت التي يجب سحبها من كل طبقة، إذا أردنا سحب عينة عشوائية طبقية متساوية حجمها n=30 من مجتمع يتكون من 3 طبقات وبه 100 منشأة، وتتوزع الطبقات بالشكل التالي:

الطبقة	صناعات صغرى	صناعات صغيرة	صناعات متوسطة	مجموع المجتمع
عدد العناصر في كل طبقة	50	20	30	100

التمرين 5:

بهدف دراسة قطاع الاتصالات استخرج عينة طبقية تناسبية مكونة من 200 من أصل 21000 زبون في قطاع الاتصالات المقسم إلى ثلاث طبقات حسب الجدول التالي :

الطبقة	موبيليس	جيزي	أوريدو	مجموع المجتمع
سكيكدة	2000	4000	5000	11000
قسنطينة	4000	3000	3000	10000

الفصل الأول: توزيعات المعاينة.

مقدمة:

تعتبر نظرية المعاينة بمثابة حلقة وصل بين نظرية الاحتمالات والاستدلال الاحصائي حول المجتمع، وتهتم نظرية المعاينة بدراسة العلاقة بين العينة والمجتمع الذي أخذت منه، حيث تختصر البيات في التعبير عنها بمعلمات (احصاءه) على غرار متوسط العينة والانحراف المعياري للعينة...الخ، وعند القيام بمعاينة عشوائية سنتحصل على قيم مختلفة لهذه الاحصائية التي نكون بصدد دراستها، وبالتالي تصبح هذه الاحصائية متغيرا عشوائيا يخضع لقوانين نظرية الاحتمالات.

وعموما يمكن اعتبار هذا الفصل (توزيعات المعاينة) بمثابة مفتاح لفهم الاستدلال الاحصائي، حيث سيتم من خلاله التطرق إلى مجموعة من النظريات التي تبرز لنا طبيعة العلاقة الرياضية والاحصائية بين معالم وخصائص المجتمع (متوسط المجتمع ب μ (حرف اغرقي *mu*) بينما يرمز للانحراف المعياري للمجتمع ب σ (*sigma*)، P نسبة صفة معينة في المجتمع) ومعالم وخصائص العينة ($\bar{X}$ متوسط العينة، و S انحرافها المعياري، $\hat{P}$ نسبة صفة معينة في العينة).

1 /Π. توزيعات المعاينة:

يشير الاستدلال الاحصائي إلى استنتاج معالم وخصائص المجتمع. حيث تختصر البيانات باستخدام الاحصاء على غرار متوسط العينة والانحراف المعياري للعينة. عندما نحصل على البيانات من خلال معاينة عشوائية فإن الرابط بين نظرية الاحتمالات وتكوين البيانات هو **توزيع المعاينة** للإحصائيات. توزيع المعاينة يبرز إلى أي مدى تتغير الإحصاء عند تكرار السحب وتكوين البيانات.

ويمكن إعطاء تعريف شامل لتوزيع المعاينة: على أنه التوزيع الاحتمالي الذي يحدد الاحتمالات الممكنة لقيم إحصائية العينة (*Agresti & Finlay* 1997)، وتوزيع المعاينة لإحصائية ما يعتمد على عدد مشاهدات حجمها n وهو التوزيع الاحتمالي لهذه الاحصائية الناتج عن تكرار سحب العينات ذات الحجم n، كل مرة نخلص إلى قيمة ما لهذه الاحصائية.[1]

مثلا: على فرض أن هذه الإحصائية التي نريد تكوين توزيعها الاحتمالي هي الوسط الحسابي، وليكن لدينا مجتمع ما عدد مفرداته N مفردة، ويتبع توزيعا احتماليا معينا، على فرض أننا سحبنا عينة عشوائية حجمها n من هذا المجتمع، ثم قمنا بحساب وسطها الحسابي $\bar{x}_1$، ثم سحبنا عينة عشوائية

---

[1]- Jarkko Isotalo, “**Basics of Statistics**”; Basics of Statistics course held in University of Tampere, Finland ; p 46.

ثانية حجمها n من ذات المجتمع، ثم قمنا بحساب وسطها الحسابي $\bar{x}_2$، ...وهكذا، أي إذا قمنا بتكرار السحب، وسحبنا كل العينات الممكن سحبها من هذا المجتمع فإنه سيتوفر لدينا عدد كبير من قيم المتوسطات الحسابية للعينات، هذه القيم ل $\bar{x}$ تشكل مجتمعا آخر عدد مفرداته أكبر بكثير من مفردات المجتمع الأصلي، وعلى ذلك يمكن النظر لهذا المقياس (الوسط الحسابي) على أنه متغير عشوائي يأخذ قيما مختلفة، ويتبع توزيعا احتماليا معينا قد يختلف عن توزيع المجتمع الأصلي. ويسمى هذا المجتمع الجديد **بمجتمع توزيع المعاينة للوسط الحسابي.**

وعليه فإن توزيع المعاينة لمقياس ما أو احصائية ما (وسط حسابي، نسبة صفة معينة، تباين...) ببساطة هو نوع من التوزيع الاحتمالي لمجتمع هذا المقياس. وتوزيع المعاينة لا يشير إلى المشاهدات الفردية ولكن يشير لقيم إحصاءه ما محسوبة من هذه المشاهدات، عينة بعد عينة.

2 /II. أنواع توزيعات المعاينة:

2. 1 /II. توزيع المعاينة للمتوسطات.

كما سبق ذكره فإن سحب عينات متكرر من مجتمع ما، وحساب متوسط كل عينة سيسمح لنا بالحصول على مجتمع من المتوسطات $\bar{X}$، ويسمى هذا المجتمع الجديد **"بتوزيع المعاينة للمتوسطات"**. وتجدر الاشارة إلى أنه لتوزيع المعاينة للمتوسطات وسط يرمز له ب $\mu_{\bar{X}}$، وتباين $\sigma_{\bar{X}}^2$.

ويمكن الحصول على متوسط توزيع المعاينة للمتوسطات كما يلي[1]:

$$\mu_{\bar{X}} = E(\bar{X}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$
$$= \frac{E(X_1 + X_2 + \cdots + X_n)}{n}$$
$$= \frac{E(X_1) + E(X_2) + \cdots + E(X_n)}{n}$$
$$= \frac{\mu X_1 + \mu X_2 + \cdots + \mu X_n}{n}$$
$$= \frac{\mu + \mu + \cdots + \mu}{n}$$
$$= \mu$$

[1] Jarkko Isotalo, Op.Cit; p 47.

وهو متوسط ($\bar{X}$) توزيع المعاينة للمتوسط، أي أن القيمة المتوقعة لمتوسط العينة تساوي الوسط الحقيقي للمجتمع μ.

ومن النتيجة السابقة يمكننا اشتقاق تباين المتوسط $\bar{X}$ (تباين توزيع المعاينة للمتوسط):

$$\left(\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

وأن: $$Var[X_i]=\sigma^2_i$$

$$Var[\bar{X}] = Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$$

وحسب خواص التباين: $Var(a.X)= a^2Var(X)$ *فإن:*

$$(1/n)^2 \sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2$$

$$= \frac{{\sigma_1}^2 + {\sigma_2}^2 + \cdots + {\sigma_n}^2}{n^2}$$

$$Var[\bar{X}] = \frac{\sigma^2}{n}$$

ومنه فإن الانحراف المعياري لتوزيع المعاينة للمتوسط $\bar{X}$ هو $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$. والانحراف المعياري لتوزيع المعاينة للمتوسط $\bar{X}$ يدعى أيضا **الخطأ المعياري** (*standard error*) ل $\bar{X}$ (الخطأ المعياري لتوزيع المعاينة للمتوسط).

أ. المعاينة الغير النفاذية (السحب بالإرجاع).

مثال: ليكن لدينا مجتمع احصائي مكون من 5 سلع، أوزانها ب (كغ) هي: {11، 8، 6، 3، 2}، نستخرج جميع العينات الممكنة ذات الحجم n=2، والتي يمكن استخراجها من هذا المجتمع بصورة عشوائية و**بالإرجاع.**

1. أحسب وسط عناصر المجتمع μ.
2. أحسب عدد العينات التي يمكن استخراجها ثم حددها.
3. أحسب وسط كل عينة.

4. أحسب الوسط الحسابي لعناصر المجتمع الجديد (مجتمع المتوسطات).

5. أحسب تباين عناصر المجتمع الجديد (مجتمع المتوسطات) $\sigma_{\bar{X}}^2$.

6. ماذا تلاحظ.

الحل:

1. حساب وسط عناصر المجتمع:

$$\mu = \frac{\sum X_i}{N} = \frac{2+3+6+8+11}{5} = \boxed{6}$$

2. استخراج كل العينات التي يمكن استخراجها:

بما أن السحب بالإرجاع فإن عدد العينات الممكنة يساوي 25 عينة: $N^n = 5^2 = 25$

(2 ،11)	(3 ،11)	(6 ،11)	(8 ،11)	(11 ،11)
(2 ،8)	(3 ،8)	(6 ،8)	(11 ،8)	(8 ،8)
(2 ،6)	(3 ،6)	(8 ،6)	(11 ،6)	(6 ،6)
(2 ،3)	(3،6)	(8 ،3)	(11 ،3)	(3 ،3)
(3 ،2)	(6 ،2)	(8 ،2)	(11 ،2)	(2 ،2)

3. حساب الأوساط الحسابية للعينات الممكنة:

$\bar{X}_5 = 6.5$	$\bar{X}_4 = 7$	$\bar{X}_3 = 8.5$	$\bar{X}_2 = 9.5$	$\bar{X}_1 = 11$
$\bar{X}_{10} = 5$	$\bar{X}_9 = 5.5$	$\bar{X}_8 = 7$	$\bar{X}_7 = 9.5$	$\bar{X}_6 = 8$
$\bar{X}_{15} = 4$	$\bar{X}_{14} = 4.5$	$\bar{X}_{13} = 7$	$\bar{X}_{12} = 8.5$	$\bar{X}_{11} = 6$
$\bar{X}_{20} = 2.5$	$\bar{X}_{19} = 4.5$	$\bar{X}_{18} = 5.5$	$\bar{X}_{17} = 7$	$\bar{X}_{16} = 3$
$\bar{X}_{25} = 2.5$	$\bar{X}_{24} = 4$	$\bar{X}_{23} = 5$	$\bar{X}_{22} = 6.5$	$\bar{X}_{21} = 2$

والمجتمع الجديد مجتمع الأوساط الحسابية للعينات الممكنة يعرف **بمجتمع توزيع المعاينة للوسط الحسابي** في حالة السحب بإرجاع (معاينة غير نفادية)**.**

4. حساب الوسط الحسابي للمجتمع الجديد (مجتمع توزيع المعاينة للوسط الحسابي).

$$E(\bar{X}) = \frac{\sum \bar{X}_i}{n} = \frac{150}{25} = \boxed{6}$$

5. حساب تباين توزيع المعاينة للمتوسطات. (المعاينة غير نفادية)

$$\sigma_{\bar{X}}^2 = \frac{\sum \bar{X}_i}{n} - [E(\bar{X})]^2 = \frac{1035}{25} - 6^2 = 5.4 \rightarrow \sigma_{\bar{X}} = \sqrt{5.4} = 2.32$$

6. الملاحظة: يلاحظ أن: $E(\bar{X}) = \mu = 6$.

كما يلاحظ أن $\sigma_{\bar{X}}^2 = \frac{10.8}{2} = 5.4 \overset{\text{أي أن}}{\Longrightarrow} \sigma_{\bar{X}} = \frac{3.28}{\sqrt{2}} = 2.32$

نظرية 1: إذا سحبنا عينات عشوائية حجمها n من مجتمع ذو متوسط μ وانحراف معياري σ، فان توزيع المعاينة للمتوسطات $\bar{X}$ في حالة **المعاينة الغير نفاديه** (السحب بإرجاع) يكون له الخصائص التالية:

توقع (متوسط) $E(\bar{X})$ مساوي لمتوسط المجتمع، ونكتب: $E(\bar{X}) = \mu_{\bar{X}} = \mu$

وانحراف معياري (خطأ معياري) $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

ب. المعاينة النفادية (السحب بدون ارجاع).

مثال: أوجد نفس مطالب المثال الأول في حالة السحب بدون ارجاع.

الحل:

1. استخراج كل العينات التي يمكن استخراجها:

بما أن السحب بدون إرجاع فإن عدد العينات الممكنة يساوي 10 عينة:

$$C_N^n = \frac{N!}{(N-n)!n!}$$

$$C_5^2 = \frac{5!}{(5-2)!\,2!} = \frac{5!\,4!\,3!}{3!\,2!\,1!} = \frac{20}{2} = \boxed{10}$$

$(11, 8)$ $(11, 6)$ $(11, 3)$ $(11, 2)$

$(8, 6)$ $(8, 3)$ $(8, 2)$

$(6, 3)$ $(6, 2)$

$(3,2)$

2. حساب الأوساط الحسابية للعينات الممكنة (مجتمع المتوسطات)

$\bar{X}_1=9.5$ $\bar{X}_2=8.5$ $\bar{X}_3=7$ $\bar{X}_4=6.5$

$\bar{X}_5=7$ $\bar{X}_6=5.5$ $\bar{X}_7=5$

$\bar{X}_8=4.5$ $\bar{X}_9=4$

$\bar{X}_{10}=2.5$

مجتمع الأوساط الحسابية للعينات الممكنة **(مجتمع توزيع المعاينة للوسط الحسابي)** في حالة السحب بدون إرجاع (معاينة نفادية**).**

3. حساب القيمة المتوقعة للوسط الحسابي لمجتمع توزيع المعاينة للوسط الحسابي (م. نفادية)

$$E(\bar{X}) = \frac{\sum \bar{X}_i}{n} = \frac{60}{10} = 6 = \mu$$

4. حساب تباين توزيع المعاينة للمتوسطات (معاينة نفادية).

$$\sigma_{\bar{X}}^2 = \frac{\sum \bar{X}_i}{n} - [E(\bar{X})]^2 = \frac{400.5}{10} - 6^2 = 4.05 \rightarrow \sigma_{\bar{X}} = \sqrt{4.05} = 2.01$$

نلاحظ أن:

$$\sigma_{\bar{X}}^2 = \frac{10.8}{2} . \frac{5-2}{5-1} = 4.05 \overset{\text{أي أن}}{\Longrightarrow} \sigma_{\bar{X}} = \frac{3.28}{\sqrt{2}} . \sqrt{\frac{5-2}{5-1}} = 2.01$$

نظرية 2: إذا سحبنا عينات عشوائية حجمها n من مجتمع ذو متوسط μ وانحراف معياري σ، فان توزيع المعاينة للمتوسطات $\bar{X}$ في حالة **المعاينة النفادية** (السحب بدون إرجاع) له الخصائص التالية:

توقع (متوسط) $E(\bar{X})$ مساوي لمتوسط المجتمع، ونكتب: $E(\bar{X})=\mu_{\bar{X}}=\mu$

وانحراف معياري (خطأ معياري) $\sigma_X = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$.

ملاحظة: تجدر الاشارة إلى أن النسبة $\frac{N-n}{N-1}$ تسمى **معامل الارجاع أو معامل التصحيح**، وعادة يتم استخدامها عند حساب التباين (أو الانحراف المعياري) في حالة السحب بدون ارجاع (معاينة نفادية) أو اذا كان حجم المجتمع N صغير أو منته، حيث أن N صغيرة مقارنة ب n $(0.05 \le \frac{n}{N})$.

وكلما كان حجم المجتمع كبيرا أو غير منته حيث أن N كبيرة مقارنة ب n ($\frac{n}{N} < 0.05$). أو في حالة المعاينة الغير نفاديه (السحب بالإرجاع)، فإن قيمة النسبة $\frac{N-n}{N-1}$ تقترب من الواحد الصحيح.

ت. طبيعة توزيع مجتمع توزيع المعاينة للمتوسطات:

يعتمد شكل ونوع التوزيع الاحتمالي لمجتمع المتوسطات الحسابية للعينات العشوائية على توزيع المجتمع الأصلي الذي اختيرت منه هذه العينات العشوائية. أي أنه في حالة ما إذا كان المجتمع موزعا توزيعا طبيعيا(*) بمتوسط μ وتباين σ^2 ونكتب $X \sim N(\mu, \sigma^2)$، فإن توزيع المعاينة للمتوسط أيضا يتبع التوزيع الطبيعي بمتوسط μ وانحراف معياري $\frac{\sigma}{\sqrt{n}}$. ونكتب $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

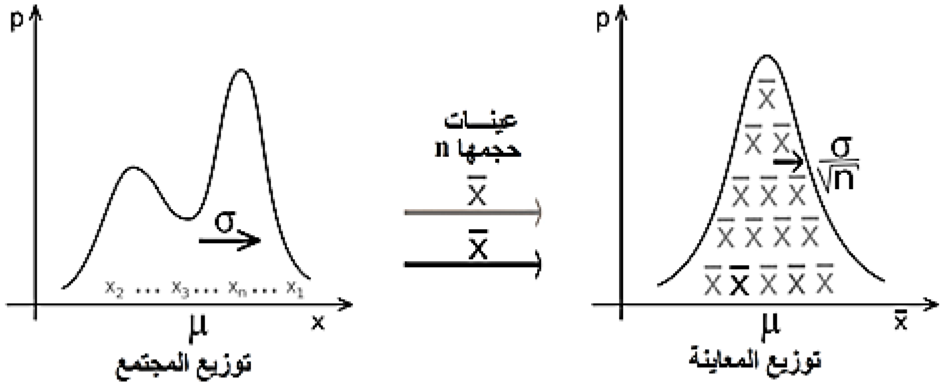
أما في حالة عدم معرفة طبيعة توزيع المجتمع الأصلي الذي سحبت منه العينات العشوائية، فإن التوزيع النهائي لعينة عشوائية سيؤول إلى التوزيع الطبيعي كلما ارتفع حجم العينة n. وهو ما تشير إليه نظرية النهاية المركزية.

نظرية 3: النهاية المركزية. تنص نظرية النهاية المركزية على أنه في كل متتالية من محاولات السحب المكررة من مجتمع ذو متوسط μ وتباين σ^2 لكن ليس بالضرورة موزع توزيعا طبيعيا، فإن متوسط العينة المعياري $z = \frac{\bar{X}-\mu}{\sigma_X/\sqrt{n}}$ يؤول إلى التوزيع الطبيعي المعياري وذلك كلما زاد عدد المحاولات أي كلما كان n كبيرا ($n \geq 30$) ونكتب: $z \sim N(0, 1)$.

ملاحظة: المتغير $z = \frac{\bar{X}-\mu}{\sigma_X/\sqrt{n}}$ يسمى المتغير المتمركز المختصر.

(*)– التوزيع الطبيعي يعرف أيضا بتوزيع غوس (Gauss Distribution)، أول من اكتشف هذا التوزيع العالم دي موافر De Moiver عام 1733، ومن بعده العالم غوس Gauss عام 1809.

الشكل رقم (II.1): التمثيل البياني لتوزيع المعاينة للمتوسطات حسب نظرية النهاية المركزية.



المصدر: Mathieu ROUAUD ; (janvier 2017): “Probabilités, statistiques et analyses multicritères”, p11.

مثال:

مجتمع مكون من 12000 عنصر بوسط 100 وانحراف معياري 60 أوجد الوسط والانحراف المعياري لتوزيع المعاينة للمتوسط عندما يكون حجم العينة: أ- 100 ، ب- 900؟

- أحسب احتمال أن يكون $\bar{X}$ محصورا بين 98 و 102 في الحالتين.

الحل:

ايجاد توزيع المعاينة للمتوسط.

أ. الحالة 1: حجم العينة $n=100$.

$$N=12000, \qquad \mu=100, \qquad \sigma_x=60.$$

1) القيمة المتوقعة لمتوسط توزيع المعاينة للمتوسط $E(\bar{X})=\mu_{\bar{X}}=\mu=100$

2) قبل حساب الانحراف المعياري لتوزيع المعاينة للمتوسط لابد من اختبار النسبة n/N

$n/N=100/12000=0.0083<0.05$ إذا نهمل معامل التصحيح.

$$\sigma_{\bar{X}}=\frac{\sigma}{\sqrt{n}}=\frac{60}{\sqrt{100}}=\boxed{6}$$

ونكتب: $\bar{X}\sim N(100, 6^2)$.

ب. الحالة 2: حجم العينة $n=900$.

1) القيمة المتوقعة لمتوسط توزيع المعاينة للمتوسط $E(\bar{X})=\mu_{\bar{X}}=\mu=100$

2) قبل حساب الانحراف المعياري لتوزيع المعاينة للمتوسط لابد من اختبار النسبة n/N

$n/N=900/12000=0.075>0.05$ إذا نستخدم معامل التصحيح.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{60}{\sqrt{900}}\sqrt{\frac{12000-900}{12000-1}} = \boxed{1.92}$$

ونكتب: $\bar{X} \sim N(100, 1.92^2)$.

2. إيجاد احتمال أن يكون $\bar{X}$ محصورا بين 98 و 102.

أ. <u>الحالة 1</u>: حجم العينة $n=100$.

$$P(98 \leq \bar{X} \leq 102) \xrightarrow{\text{نحولها إلى قيم معيارية}} P\left(\frac{98-100}{6} \leq \frac{\bar{X}-\mu}{\sigma_X} \leq \frac{102-100}{6}\right)$$

$$P(-0.33 \leq Z \leq 0.33) = P(Z \leq 0.33) - P(Z < -0.33)$$

$$P(Z \leq 0.33) - [1 - P(Z < 0.33)] = 2F(0.33) - 1$$

$$=2(0.6293)-1=0.2586$$

ب. <u>الحالة 2</u>: حجم العينة $n=900$.

$$P(98 \leq \bar{X} \leq 102) \xrightarrow{\text{نحولها إلى قيم معيارية}} P\left(\frac{98-100}{1.92} \leq \frac{\bar{X}-\mu}{\sigma_X} \leq \frac{102-100}{1.92}\right)$$

$$P(-1.04 \leq Z \leq 1.04) = P(Z \leq 1.04) - P(Z < -1.04)$$

$$P(Z \leq 1.04) - [1 - P(Z < 1.04)] = 2F(1.04) - 1$$

$$=2(0.8508)-1=0.7016$$

II/ 2. 2. توزيع المعاينة للنسبة.

في بعض الأحيان يبدو من الضروري تقدير او اتخاذ قرار يتعلق بنسبة عناصر أو مفردات المجتمع الاحصائي التي تحمل صفة معينة، فمثلا نسبة المصابين بمرض معين في مجتمع ما، أو نسبة العاطلين عن العمل، أو نسبة التلف في الانتاج لسلعة ما...إلخ، إن إحصاءه العينة التي عادة ما تستخدم لإعطاء معلومات عن نسبة مفردات المجتمع الاحصائي التي تحمل صفة معينة يطلق عليه تسمية **معاينة النسبة**[1].

ونقصد بالنسبة في المجتمع الكسر $P=\frac{N_a}{N}$، حيث أن N تمثل حجم المجتمع، و N_a عدد المفردات التي تتحقق فيها صفة ما.

ونقصد بالنسبة في العينة $\hat{P}=\frac{n_a}{n}$، حيث أن n تمثل حجم العينة، و n_a عدد مفردات العينة التي تحقق نفس الصفة.

نظرية 4: إذا كانت P نسبة صفة معينة في مجتمع ما وسحبنا عينات عشوائية حجم كل منها n، فإن مختلف النسب $\hat{P}_i$ ($i=1 ;.... ;n$)، تشكل توزيع المعاينة للنسبة، ويتميز التوزيع النظري لمعاينة النسبة $\hat{P}$ بالخصائص التالية:

توقع (متوسط توزيع المعاينة للنسبة) $E(\hat{P})=\mu_{\hat{P}}=P$.

وخطأ معاينة (الانحراف المعياري لتوزيع المعاينة للنسبة) $\sigma_{\hat{P}}=\sqrt{\frac{pq}{n}}=\sqrt{\frac{p(1-p)}{n}}$

نكتب: $\approx N\hat{P}\ (p,\sigma_{\hat{P}}^2)$ (في حالة $n\geq 30$، أو في حالة مجتمع طبيعي)

ملاحظة: إذا كانت المعاينة نفادية، أو حجم المجتمع محدود ($n>0.05N$)، فإن $\sigma_{\hat{P}}^2=\frac{pq}{n}.(\frac{N-n}{N-1})$

أ. طبيعة توزيع المعاينة للنسبة

بصفة عامة يمكن القول بأنه كلما زاد حجم العينة ($n\sim\infty$) كلما اقترب توزيع المعاينة للنسبة ($\hat{p}$) من التوزيع الطبيعي حسب نظرية النهاية المركزية.

[1] - علي عبد السلام العماري، علي حسين العجيلي، "**الاحصاء والاحتمالات: النظرية والتطبيق**"، شركة النشر ELGA، فاليتا، مالطا، 2000، ص 429.

ونكتب: $z = \frac{\hat{P}-\mu_p}{\sigma_{\hat{P}}} \sim N(0,1)$.

والقاعدة العامة لتطبيق هذه النظرية تقول لكي يكون التقريب باستخدام التوزيع الطبيعي مقبولا يجب أن يتحقق ما يلي:

$$\begin{cases} n \geq 30 \\ np > 5 \\ nq > 5 \end{cases}$$

مثال:

سحبنا عينة عشوائية حجمها 31 من نقاط الطلبة في امتحان الاحصاء الرياضي فكانت كالتالي:

{19.5، 11.5، 14، 11، 16.5، 18، 7.5، 3.5، 2.5، 2.5، 8.5، 10.5، 8.5، 2، 5، 3، 2.5، 1، 2.5، 10، 6، 6، 15، 6، 8.5، 5.5، 1.5، 12، 3، 18.5، 10.5}

1. أحسب $\hat{P}$ نسبة الطلبة الحاصلين على المعدل في العينة.
2. أوجد التوزيع العيني لنسبة الطلبة الحاصلين على المعدل في العينة من مجموع الطلبة البالغ 300.

الحل:

1. حساب $\hat{P}$ نسبة الطلبة الحاصلين على المعدل في العينة.

لدينا 12 علامة أكبر أو تساوي 10 ($12=n_a$) $\Leftarrow$ $\hat{p}=n_a/n=12/31=0.387$

2. التوزيع العيني لنسبة الطلبة الحاصلين على المعدل في العينة (علامة أكبر أو تساوي 10).

القيمة المتوقعة لنسبة الطلبة الحاصلين على المعدل في العينة: $E(\hat{P}) = \mu_{\hat{P}} = P = 0.38$

لحساب تباين توزيع المعاينة للنسبة (خطأ المعاينة) لابد من اختبار النسبة $n/N = 31/300 = 0.103> 0.05$ $\Leftarrow$ نستخدم معامل التصحيح لحساب تباين توزيع المعاينة للنسبة.

$$\sigma_{\hat{P}}^2 = \frac{pq}{n} \cdot \frac{N-n}{N-1} = \frac{p(1-p)}{n} \cdot \frac{N-n}{N-1} = \frac{0.387(1-.387)}{31} \cdot \frac{300-31}{300-1} = \boxed{0.0069}$$

ونكتب: $\hat{P} \sim N(P, \sigma_{\hat{P}}^2) >\Rightarrow \hat{P} \sim N(0.387, 0.0069)$

2 /II. 3. توزيع المعاينة للفرق للفروق والمجاميع.

في كثير من الأبحاث ينصب التركيز على دراسة مجتمعين وليس مجتمع واحد فقط، ولذلك يتم دراسة الفروق أو المجموع بين احصائيتين (متوسط، نسبة ...) محسوبيتين من عينتين مسحوبتين من مجتمعين مختلفين، فإذا كانت لدينا $\{X_1, X_2,..., X_n\}$ تشكل عينة عشوائية حجمها n من مجتمع احصائي ما ذو متوسط μ_X، وانحراف معياري σ_X، وكانت $\{Y_1, Y_2,..., Y_n\}$ تشكل عينة عشوائية حجمها m من مجتمع احصائي ثان ذو متوسط μ_Y، وانحراف معياري σ_Y، نحسب لكل عينة مسحوبة من المجتمع الأول الاحصائية S_X، ونحسب نفس الاحصائية في كل عينة مسحوبة من المجتمع الثاني S_Y. إن الفرق $S_X - S_Y$ يشكل بدوره متغيرة عشوائية لها المتوسط والتباين التاليين:

$$\mu_{Sx-SY} = \mu_{Sx} - \mu_{SY} \qquad \sigma^2_{Sx-SY} = \sigma^2_{Sx} + \sigma^2_{SY}$$

إذا كان الاهتمام هو على مجموع الاحصائيتين بدلا من الفرق بينهما فإن:

$$\mu_{Sx+SY} = \mu_{Sx} + \mu_{SY} \qquad \sigma^2_{Sx+SY} = \sigma^2_{Sx} + \sigma^2_{SY}$$

وذلك بشرط استقلال العينات عن بعضها البعض وبلغة أخرى يشترط استقلال المتغيرات S_X، S_Y .

أ. توزيع المعاينة للفرق بين متوسطي عينتين مستقلتين:

إذا كانت الاحصائية S_X والاحصائية S_Y تمثلان **متوسطي** عينتين مسحوبتين من مجتمعين مستقلين طبيعيين معلومي التباين $\sigma^2_{\bar{X}}$ و $\sigma^2_{\bar{Y}}$ على التوالي فإن:

الوسط الحسابي لتوزيع المعاينة للفرق بين وسطي عينتين هو الفرق بين متوسطي المجتمعين ونكتب:

$$\mu_{\bar{X}-\bar{Y}} = \mu_{\bar{X}} - \mu_{\bar{Y}} = \mu_X - \mu_Y$$

وتباين توزيع المعاينة للفرق بين متوسطي عينتين يكتب كما يلي: $\sigma^2_{\bar{X}-\bar{Y}} = \frac{\sigma^2_{\bar{X}}}{n} + \frac{\sigma^2_{\bar{Y}}}{m}$

نظرية 5: في حالة $m, n \geq 30$، فإن توزيع المعاينة للفرق بين متوسطين يتوزع تقريبا وفق التوزيع الطبيعي، ونكتب:

$$(\bar{X} - \bar{Y}) \sim N\left(\mu_X - \mu_Y, \sigma^2_{\bar{X}-\bar{Y}}\right)$$

وعليه فإن توزيع المتغير المعياري $(\bar{X} - \bar{Y})$ يقترب من التوزيع الطبيعي المعياري ونكتب:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_{\bar{X}} - \mu_{\bar{Y}})}{\sqrt{\frac{\sigma^2_{\bar{X}}}{n} + \frac{\sigma^2_{\bar{Y}}}{m}}} \sim N(0,1)$$

مثال:

ليكن لدينا المجتمع 1: $\Omega_x=\{3, 4,5\}$، والمجتمع 2: $\Omega_y=\{0, 1,2\}$

تحقق من أن: $\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y$. $\qquad \mu_{x-y} = \mu_x - \mu_y$;

ثم تحقق أن: $\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y$. $\qquad \mu_{x+y} = \mu_x + \mu_y$;

الحل:

حساب متوسط المجتمع الأول: $\mu_x=\Sigma X/N = 3+4+5/3 =4$

حساب تباين المجتمع الأول: $\sigma^2_X=(\Sigma X^2/N)-(\mu_x)^2= (9+46+25/3)-16 = \boxed{0.66}$

حساب متوسط المجتمع الثاني: $\mu_Y =\Sigma Y/M = 0+1+2/3 =1$

حساب تباين المجتمع الثاني: $\sigma^2_Y =(\Sigma Y^2/M)-(\mu_Y)^2= (0+1+4/3)-1 = \boxed{0.66}$

توزيع المعاينة للفرق بين متوسطين:

مجتمع الفروق: Ω_X-Ω_Y: {5, 4, 3, 4, 3, 2, 3, 2, 1}

حساب متوسط المجتمع الجديد: $\mu_{X-Y}=5+4+3+4+3+2+3+2+1/9 = \mathbf{27/9 =3}$

تباين المجتمع الجديد: $\sigma^2_{X-Y} =(25+16+9+16+9+4+9+4+1/9)-9= 93/9-9 = \boxed{1.33}$

وعليه فإن توزيع المعاينة للفرق بين متوسطين يكتب بالشكل: $(\bar{X}-\bar{Y})\sim N(3, 1.33)$

التحقق من أن: $\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y$. $\qquad \mu_{x-y} = \mu_x - \mu_y$;

$\sigma^2_{x-y} = 0.66+0.66\simeq 1.33 \qquad \mu_{x-y} = 4-1=\mathbf{3}$;

توزيع المعاينة لمجموع متوسطين:

مجتمع المجموع: Ω_X-Ω_Y: {5, 6, 7, 4, 5, 6, 3, 4, 5}

حساب متوسط المجتمع الجديد: $\mu_{X+Y}=5+6+7+4+5+6+3+4+5/9 = \mathbf{45/9 =5}$

تباين المجتمع الجديد: $\sigma^2_{X+Y} =(25+36+49+16+25+36+9+16+25/9)-25= \boxed{1.33}$

وعليه فإن توزيع المعاينة للمجموع بين متوسطين يكتب بالشكل: $(\bar{X}+\bar{Y})\sim N(5, 1.33)$

التحقق من أن: $\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y$. $\mu_{x+y} = \mu_x + \mu_y$;

$\sigma^2_{x+y} = 0.66+0.66 \simeq 1.33$ $\mu_{x+y} = 4 + 1 = \mathbf{5}$;

ملاحظة: اذا كان المجتمعين مجهولي التباين فإننا نميّز بين حالتين:

أولا: اذا كان حجم العينتين كبير (n و $m \geq 30$)، فإن توزيع المعاينة للفرق بين متوسطي عينتين يتبع التوزيع الطبيعي بمتوسط $\mu_X - \mu_Y$ وتباين: $\sigma^2_{\bar{X}-\bar{Y}} = \frac{S_X^2}{n} + \frac{S_Y^2}{m}$، ويكون التوزيع الاحتمالي كالتالي:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_{\bar{X}} - \mu_{\bar{Y}})}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}} \sim N(0,1)$$

ثانيا: اذا كان حجم احدى العينتين أو كلاهما صغير (أقل من 30)، والمجتمعين موزعين توزيعا طبيعيا فإن توزيع المعاينة للفرق بين متوسطي عينتين يتبع التوزيع الطبيعي بمتوسط $\mu_X - \mu_Y$ وتباين: $\sigma^2_{\bar{X}-\bar{Y}} = \frac{S^2}{n} + \frac{S^2}{m}$، ويكون التوزيع الاحتمالي كالتالي:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_{\bar{X}} - \mu_{\bar{Y}})}{\sqrt{\frac{S^2}{n} + \frac{S^2}{m}}} \sim N(0,1)$$

حيث أن S^2 تدعى التباين المشترك لتبايني العينتين S_X^2 و S_Y^2 وتحسب كما يلي:

$$\boxed{S^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}$$

ب. توزيع المعاينة للفرق بين نسبتي عينتين:

اذا كان اهتمامنا منصبا حول دراسة توزيع المعاينة للفرق بين نسبتي عينتين، وكانت الاحصائية S_X والاحصائية S_Y تمثلان **نسبتي** عينتين مسحوبتين من مجتمعين مستقلين طبيعيين ، وكان حجم العينتين كبيرا ($m, n \geq 30$)، فإن:

متوسط توزيع المعاينة لنسبة عينتين هو: $\mu_{\hat{P}_X - \hat{P}_Y} = \mu_{\hat{P}_X} - \mu_{\hat{P}_Y} = P_X - P_Y$

وتباين هذا التوزيع يكتب كما يلي: $\sigma^2_{\hat{P}_X - \hat{P}_Y} = \frac{P_X(1-P_X)}{n} + \frac{P_Y(1-P_Y)}{m}$

نظرية 6: توزيع المعاينة للفرق بين نسبتي عينتين حجمهما كبير($m,\ n\geq 30$) ومسحوبتين من مجتمعين مستقلين، فإن توزيع المعاينة للفرق بين **نسبتي عينتين** يتبع التوزيع الطبيعي تقريبا، ونكتب:

$$\left(\hat{P}_X - \hat{P}_Y\right) \sim N(\mu_{PX-PY}, \sigma^2_{PX-PY})$$

وعليه فإن توزيع المتغير المعياري للفرق بين نسبتي عينين يقترب من التوزيع الطبيعي المعياري ونكتب:

$$Z = \frac{\left(\hat{P}_X - \hat{P}_Y\right) - \left(\mu_{\hat{P}_X} - \mu_{\hat{P}_Y}\right)}{\sqrt{\sigma^2_{\hat{P}_X - \hat{P}_Y}}} \sim N(0,1)$$

مثال:

اذا علمت أن نسبة البطالة في منطقة ما تمثل 30**%** بالنسبة للإناث، و20**%** بالنسبة للذكور، تم سحب عينتين عشوائيتين من مجتمع الذكور ومجتمع الاناث حجمها $n_1=n_2=100$.

- أحسب احتمال أن الفرق بين نسبة البطالة بين الاناث ونسبة البطالة بين الذكور أكبر من 20%.

الحل:

بما أن حجم العينتين كبير $n_1 = n_2 = 100 > 30$ و

$$\begin{cases} n_1 . P_1 = 100(0.3) = 30 > 5 \\ n_1 . q_1 = 100(0.7) = 70 > 5 \end{cases} \text{و}$$

$$\begin{cases} n_2 . P_2 = 100(0.2) = 20 > 5 \\ n_2 . q_2 = 100(0.8) = 80 > 5 \end{cases}$$

وعليه فإن توزيع المعاينة للفرق بين نسبتي العينتين P_1-P_2 يقترب من التوزيع الطبيعي بمتوسط:

$$\mu_{\hat{P}_1 - \hat{P}_2} = \mu_{\hat{P}_1} - \mu_{\hat{P}_2} = 0.3 - 0.2 = \boxed{0.1}$$

وخطأ معياري:

$$\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = \sqrt{\frac{0.3 \times 07}{100} + \frac{0.2 \times 0.8}{100}} = \boxed{0.061}$$

حساب احتمال أن يكون الفرق بين نسبة البطالة بين الاناث ونسبة البطالة بين الذكور أكبر من 20%.

$$P\left[\left(\hat{P}_1-\hat{P}_2\right)>0.2\right] \Rightarrow P\left[\frac{\left(\hat{P}_1-\hat{P}_2\right)-\left(\mu_{\hat{P}_1}-\mu_{\hat{P}_2}\right)}{\sigma_{\hat{P}_1-\hat{P}_2}}>\frac{0.2-0.1}{0.061}\right]$$

$$=P(Z>1.64)=1-P(Z\leq 1.64)$$

$$=1-F(1.64)=1-0.9495=\boxed{0.0505}$$

II/ 2. 4. توزيع المعاينة للتباين.

انطلاقا من تعريف توزيع المعاينة فإنه يمكن الحصول على توزيع المعاينة للتباين من خلال سحب كل العينات الممكنة ذات الحجم n وحساب كل التباينات الممكنة لكل العينات الممكنة، وبالتالي نحصل على مجتمع جديد يعرف بتوزيع المعاينة للتباين، يتبع توزيعا يعرف بتوزيع **كاي مربع**.

بالعودة إلى المسألة رقم 01، (مجتمع احصائي مكون من 5 سلع، أوزانها ب (كغ): 11، 8، 6، 3، 2) بعد أن تم استخراج كل العينات الممكنة ذات الحجم n=2، في حالة السحب بالإرجاع،

1. أحسب تباين المجتمع σ^2_X.
2. أحسب تباين كل عينة S_i^2.
3. أحسب القيمة المتوقعة لعناصر المجتمع الجديد (مجتمع التباينات).
4. قارن بين تباين المجتمع، والقيمة المتوقعة لتباين العينة، ماذا تلاحظ؟

أحسب نفس المطالب السابقة في حالة السحب بدون إرجاع.

الحل:

1. حساب تباين المجتمع:

$$\sigma_X^2=E(X^2)-[E(X)]^2=\frac{\sum X_i^2}{N}-\mu^2.$$

$$\sigma_X^2=\frac{234}{5}-6^2=10,8\ \Rightarrow \sigma_X=\sqrt{10.8}=3.28.$$

أ. **السحب بالارجاع:**

التباينات الممكنة S_i^2				
20.25	16	6.25	2.25	0
9	6.25	1	2.25	0
4	2.25	1	6.25	0
0.25	2.25	6.25	16	0
0.25	4	9	20.25	0

$E(S^2)=(\sum_i S^2_i)/n$

$= 135/25 = 5.4$

$E(S^2) = 5.4=10.8/2$

$=\sigma_x^2(n-1/n)$

ب. السحب بدون إرجاع:

$E(S^2) = (\sum_i S^2_i)/n$

$= 67.5/10 = 6.75$

$E(S^2) = 6.75 = 10.8/2(1/2)(5/4)$

$= \sigma_x^2(n-1/n)(N/N-1)$

التباينات الممكنة S_i^2			
20.25	16	6.25	2.25
9	6.25	1	
4	2.25		
0.25			

نظرية 7: اذا كان X المتغير العشوائي من مجتمع ما و S^2 متغيرة عشوائية تمثل تباين عينة مسحوبة من ذات المجتمع فإن القيمة المتوقعة لتباين العينة يعطى كالتالي:

(1) السحب بدون ارجاع
(أ) إذا كان حجم المجتمع غير منته
(ب) إذا كانت نسبة حجم العينة إلى المجتمع $0.05 > \left(\frac{n}{N}\right)$
(2) السحب بالارجاع

$$\mu_{S^2} = \sigma_X^2\left(\frac{n-1}{n}\right)$$

إذا كان السحب بدون ارجاع أو
إذا كان حجم المجتمع N صغير أو منته
أي أن $0.05 \leq \left(\frac{n}{N}\right)$

$$\mu_{S^2} = \sigma_X^2\left(\frac{n-1}{n}\right)\left(\frac{N-n}{N}\right)$$

ملاحظة:

تجدر الاشارة إلى أنه كلما زاد حجم المجتمع N فإن قيمة معامل التصحيح ستؤول إلى الواحد $[\left(\frac{N-n}{N}\right) \approx 1]$، كما أنه كلما كان حجم العينة كبيرا ($n \geq 30$) فإن $E(S) \simeq \sigma^2$.

ت. طبيعة توزيع المعاينة للتباين.

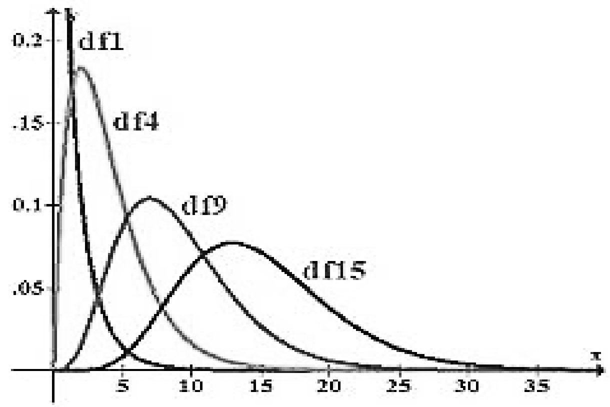
توزيع كاي مربع ك² (χ^2 chi squared)

اذا كان لدينا $X_1 ; X_2 ; \ldots\ldots ; X_n$ متغيرات عشوائية مستقلة، موزعة طبيعيا بمتوسط يساوي 0 وتباين ثابت يساوي 1، ونكتب بالشكل $X_i \sim N(0, 1)$

وليكن Z_i يمثل مجموع مربعات المتغير المعياري X_i ونكتب: $Z_i = \sum_{i=1}^{n} x_i^2$ ، وبذلك نحصل على متغير جديد (Z_i) يتبع توزيعا احتماليا يسمى توزيع **كاي مربع**، ويرمز له بالرمز χ_v^2.

تعريف: مما سبق يمكن تعريف توزيع χ^2 على أنه توزيع مجموع المربعات ل n من المتغيرات المستقلة التي تتبع التوزيع الطبيعي المعياري[1]. ونقول Z *يتبع* توزيع χ_{n-1}^2 بدرجة حرية *d.f* (*degree of freedom*)، بحيث يشير المؤشر v إلى درجة الحرية (*v=n-1*). ونكتب $Z \sim \chi_v^2$.

الشكل رقم (**2.Π**): التمثيل البياني لتوزيع كاي مربع χ^2 (حسب درجة الحرية).



نظرية 8: انطلاقا من عينة عشوائية حجمها n من المشاهدات المستقلة مسحوبة من مجتمع طبيعي بتباين σ_X^2 فإن:

$$\chi_v^2 = \frac{(n-1)S_X^2}{\sigma_X^2}$$

تتبع توزيع كاي مربع (*chi-square*) بدرجة حرية *v=n-1*.

مثال:

على فرض أن علامات الطلاب لعينة مكونة من 25 علامة لمادة الإحصاء موزعة طبيعيا كما يلي: $X \sim N(20, 8)$، أحسب احتمال أن يكون تباين علامات الطلاب في العينة أقل من 6.

الحل:

[1]- G.S.Maddala, “**Introduction to Econometrics**”, third Edition, JOHN WILEY & SONS, LTD, New York, July 2002, P20.

حساب احتمال أن يكون تباين نقاط الطلاب أقل من 6:

$P(S^2<6)= P(nS^2/\sigma^2<(n-1)\hat{s}^2/ \sigma^2)$

$=P(\chi^2_{n-1}<(24)6/8) =P(\chi^2_{24}<18)=1- P(\chi^2_{24}\geq 18)$

من جدول كاي مربع نستخرج قيمة الاحتمال المقابل لدرجة الحرية وعليه فإن:

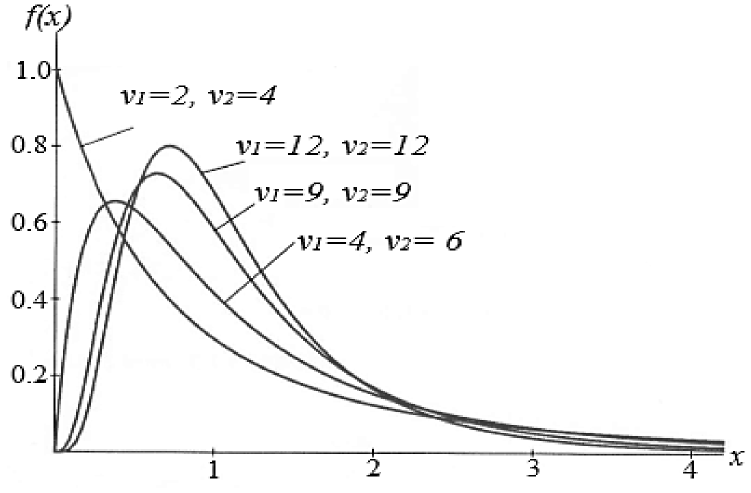
$P(S^2<6)= P(\chi^2_{24}<18) =1- P(\chi^2_{24}\geq 18)=1-0.8=\mathbf{0.2}$

5 .2 /Π. توزيع المعاينة لنسبة تبايني عينتين.

اذا كان لدينا $X \sim \chi^2_{v1}$ و $Y \sim \chi^2_{v2}$ وX وY مستقلان عن بعضهما، $Z=\frac{(X/n)}{(Y/m)}$ يتبع توزيع F (*Fisher*) بدرجتي حرية v_1 و v_2 ($v_1= n-1$ و $v_2=m-1$). ونكتب: $Z \sim F_{v_1 v_2}$

ويشير v_1 إلى درجة حرية البسط، و v_2 يشير إلى درجة حرية المقام. وعليه توزيع فيشر هو توزيع نسبة متغيرين مستقلين يتبعان توزيع χ^2 بدرجتي حرية v_1 و v_2. وهذه الدرجات هي التي تحدد شكل التوزيع، حيث أن توزيع فيشر ملتوي نحو اليمين وتقل درجة الالتواء كلما زادت درجات أحدى درجات الحرية أو كلاهما حسب ما هو موضح في الشكل التالي:

الشكل رقم (**3.Π**): التمثيل البياني لتوزي فيشر **F** (حسب درجة الحرية).



ومما سبق نستنتج أن توزيع المعاينة لنسبة تبايني عينتين مستقلتين يتبع توزيع فيشر، ويمكن صياغة النظرية التالي:

نظرية 9: ان نسبة تباين عينتين مستقلتين حجمهما على الترتيب n , m، مسحوبتين من مجتمعين طبيعيين تبايناهما $\sigma^2{}_X$, $\sigma^2{}_Y$ على التوالي تكتب كما يلي:

$$F=\frac{\left[\frac{S_X^2 n}{n-1}\right]\frac{1}{\sigma_X^2}}{\left[\frac{S_Y^2 m}{m-1}\right]\frac{1}{\sigma_Y^2}}=\frac{\hat{S}_X^2/\sigma_X^2}{\hat{S}_Y^2/\sigma_Y^2}\to F_{v_1,v_2}$$

مثال[1] . عينتين حجمهما 8 و10 مسحوبتين من مجتمعين طبيعيين تبايناهما على التوالي 20 و36. ما احتمال أن يكون تباين الأولى أكبر من ضعف تباين الثانية؟

$P(S^2{}_1>2\ S^2{}_2)=P(S^2{}_1/\ S^2{}_2>2)=$

$$P\left(\frac{S_1^2\left[\frac{n_1}{n_1-1}\right]\frac{1}{\sigma_1^2}}{S_2^2\left[\frac{n_2}{n_2-1}\right]\frac{1}{\sigma_2^2}}>2\frac{\left[\frac{n_1}{n_1-1}\right]\frac{1}{\sigma_1^2}}{\left[\frac{n_2}{n_2-1}\right]\frac{1}{\sigma_2^2}}\right)=P\left(\frac{S_1^2\left[\frac{n_1}{n_1-1}\right]\frac{1}{\sigma_1^2}}{S_2^2\left[\frac{n_2}{n_2-1}\right]\frac{1}{\sigma_2^2}}>2\frac{\left[\frac{8}{7}\right]\frac{1}{20}}{\left[\frac{10}{9}\right]\frac{1}{36}}\right)$$

$= P(F_{7,9} > 3.7)$

من الجدول نجد $0.05 > P(F_{7,9} > 3.7) > 0.01$ و في الحقيقة $P(F_{7,9} > 3.7) = 0.036$

[1]- بوعبد الله صالح، **"محاضرات الاحصاء الرياضي"**، لطلبة كلية العلوم الاقتصادية، 2005-2006، ص 10.

تمارين حول الفصل (السلسلة 02).

التمرين 01:

مجتمع مكون من 5 موظفين كان عدد مرات تغيبهم عن العمل خلال السنوات الثلاث الماضية هي: $\{ x_1=2, \quad x_2=0, \quad x_3=1, \quad x_4=4, \quad x_5=3 \}$

a) أوجد متوسط μ وتباين σ^2 المجتمع؟
b) اوجد توزيع المعاينة للوسط الحسابي للعينات ذات الحجم $n=3$ إذا كان السحب بدون إرجاع؟
c) احسب متوسط وتباين توزيع المعاينة للوسط الحسابي؟
d) تحقق من العلاقة التي تربط متوسط وتباين المجتمع بمتوسط وتباين المعاينة للوسط الحسابي؟

التمرين 2:

افترض أن المجتمع يتكون من 900 عنصر بوسط حسابي 20 وحدة، وانحراف معياري 12 وحدة.

- أوجد الوسط الحسابي والانحراف المعياري لوسط عينة حجمها أ) $n = 36$. ب) $n = 64$.
- أحسب احتمال أن يكون $\bar{X}$ محصورا بين 18 و 22 في الحالتين أ و ب.

التمرين 3:

اذا كانت أجور العاملين في مصنع للسيارات تتبع التوزيع الطبيعي بمتوسط قدره $\mu = 100$ ألف دينار وانحراف معياري قدره $\sigma = 75$ ألف دينار، تم اختيار عينة عشوائية من 25 عاملا وبدون ارجاع.

1. اذا كان عدد العمال العاملين في المصنع هو 60000 عامل، ما هو عدد العمال الذين تزيد أجورهم عن 130 ألف دينار في المجتمع.
2. أوجد احتمال أن يكون متوسط الأجور في العينة أقل من 80 ألف دينار.
3. أوجد احتمال أن يكون متوسط الأجور في العينة محصور بين 70 و130 ألف دينار.

التمرين 4:

أجريت دراسة على مجتمع للمراهقين حجمه 70 مراهقا، فأثبتت النتائج أن 28 مراهقا يعانون من البدانة، اذا تم أخذ عينة عشوائية ل 10 مراهقين من هذا المجتمع:

1. أوجد توزيع المعاينة لنسبة المراهقين الذين يعانون من البدانة.
2. أوجد احتمال أن تكون نسبة البدانة في العينة أقل من 35%.
3. أوجد احتمال ان تكون نسبة البدانة بين المراهقين في العينة محصورة بين 35 و 45%.

التمرين 5:

إذا علمت أن وزن علبة طماطم ينتجها مصنع ما تتبع التوزيع الطبيعي ومعرفة بالشكل التالي $X_1 \sim N(850,100)$، بينما كان وزن علب الطماطم في مصنع آخر معرف بالشكل التالي $X_2 \sim N(835,60)$.

- ما هو التوزيع العيني للفرق بين متوسطي العينتين الأولى من المجتمع الأول حجمها 50، والثانية من مجتمع 2 حجمها 30؟
- ما هو احتمال أن لا يزيد الفرق بين المتوسطين عن 15؟

التمرين 6:

تقدر نسبة نجاح الطالبات في جامعة ما ب 80%، بينما تقدر نسبة نجاح الطلبة ب 72%، تم سحب عينتين من مجتمعي الطالبات والطلبة 100 و80 على التوالي.

1. أوجد توزيع المعاينة للفرق بين نسبة نجاح الطالبات ونسبة نجاح الطلاب.
2. ما هو التوزيع التقريبي للفرق بين النسبتين.
3. ما احتمال أن تكون نسبة نجاح الطالبات أكبر من نسبة نجاح الطلاب ب 15%؟

التمرين 7:

تتوزع ساعات إضاءة لنوع من المصابيح طبيعيا بمتوسط μ، وتباين σ^2. اختيرت عينة عشوائية حجمها n=25 مصباح من خط انتاج المصنع، إذا علمت أن الانحراف المعياري لساعات الاضاءة لجميع المصابيح المنتجة من هذا النوع تقدر ب $\sigma= 10$.

1. أوجد احتمال أن يكون تباين العينة أكبر من 80.
2. أوجد احتمال أن يكون تباين العينة أقل من 90.
3. أوجد احتمال أن يكون تباين العينة محصور بين 80 و 120.

التمرين 8:

اذا علمت أن تباين عينة مسحوبة من مجتمع أول تساوي $S_1^2=0.20$، وتباين عينة أخرى مسحوبة من مجتمع ثان تقدر ب $S_2^2=0.14$، وكان حجم العينة الأولى $n_1=25$، وحجم العينة الثانية $n_2=30$. على فرض أن تباين المجتمعين متساويين $\sigma_1= \sigma_2$. أحسب احتمال أن يكون تباين العينة الأولى أكبر من تباين العينة الثانية.

الفصل الثاني: التقدير.

مقدمة:

يتكون المجتمع الاحصائي من مجموعة من المفردات التي تكون محل اهتمام الباحث، ويتم دراسة المجتمع من خلال معالمه أو خصائصه، على غرار متوسط المجتمع μ والانحراف المعياري σ، ونسبة صفة معينة في المجتمع P. وبدلا من دراسة جميع مفردات المجتمع فإنه يتم اختيار عينة تمثله (عينة ممثلة *échantillon représentatif*) عادة ما تكون عينة عشوائية، ثم يتم دراسة مفردات هذه العينة وحساب معالمها (متوسط العينة $\bar{X}$، الانحراف المعياري للعينة S، ونسبة صفة معينة في العينة $\hat{P}$)، ويتم استخدام هذه المعالم في عملية الاستدلال الاحصائي لمعالم المجتمع التي عادة ما تكون المجهولة.

وخلال هذا الفصل سيتم ابراز كيفية استخدام بيانات العينة **لتقدير (***Estimation***) معالم** المجتمع المجهولة. وتسمى احصائية العينة المستخدمة في تقدير معلمة المجتمع **بالمقدر (*Estimator*)**، بينما تسمى قيمته بالتقدير. ونميز بين نوعين من التقدير، التقدير النقطي (*point estimation*). والتقدير بمجال (*interval estimation*).

III. 1. تعريف التقدير.

التقدير هو استنتاج معلمات المجتمع المجهولة بالاعتماد على خصائص عينة مأخوذة من هذا المجتمع. حيث أن العينة تعتبر صورة مصغرة من المجتمع فإنه يتم اللجوء إلى حساب ما يقابل معالم المجتمع في العينة من مقاييس هي على التوالي $\bar{X}$، δ_x، $\hat{p}$، وذلك من بيانات العينة، ويمكن استخدام هذه المقاييس كتقدير للمعلمة المناظرة لها في المجتمع (σ، μ، P). **مثلا**: يمكن استخدام متوسط العينة $\bar{x}$ لتقدير متوسط المجتمع μ. ونميّز بين نوعين من التقدير: **التقدير النقطي**، و**التقدير بمجال.**

III. 2. التقدير النقطي.

III. 2. 1 تعريف التقدير النقطي.

وهو تقدير معلمة من معالم المجتمع المجهولة بقيمة واحدة أو بعدد واحد، مثلا: اذا قدرنا متوسط علامات الطلاب في مادة ما ب 13.5، نكون قد قدرنا علامات الطلاب نقطيا. أو كأن نقول أن نسبة البدانة في عينة عشوائية من المراهقين تقدر ب 25 % نكون قد قدرنا نسبة البدانة في مجتمع ما نقطيا.

وعليه يمكن القول أن متوسط العينة $\bar{X}$ مقدرا لوسط المجتمع μ، وقيمة مفردة للمتوسط $\bar{X}$ هو تقدير بنقطة لوسط المجتمع μ. وبالمثل فإنه يمكن استخدام الانحراف المعياري للعينة S كمقدر للانحراف المعياري للمجتمع σ، والقيمة المفردة للانحراف المعياري S كتقدير بنقطة للانحراف المعياري للمجتمع σ. وكذلك يمكن استخدام النسبة في العينة $\hat{P}$ كمقدر للنسبة P في المجتمع، والقيمة المفردة للنسبة $\hat{P}$ يمكن استخدامها كتقدير للنسبة P.

III. 2. 2 طرق التقدير النقطي.

هناك مجموعة من الطرق يمكن اعتمادها من أجل ايجاد المقدر النقطي من بينها:

- طريقة المعقولية العظمى.
- طريقة العزوم.
- طريقة المربعات الصغرى.
- طريقة بايز.

III. 2. 3 خواص التقدير.

يجب أن يتميز المقدر النقطي بمجموعة من الخصائص حتى نعتبره مقدرا جيدا لمعلمة المجتمع المجهولة، وهذه الخصائص هي:

أ. خاصية عدم التحيّز:

يكون المقدر غير متحيز إذا أعطى توزيع المعاينة النظري النتائج عن المعاينة العشوائية المتكررة من المجتمع احصائية مساوية لمعلمة المجتمع.

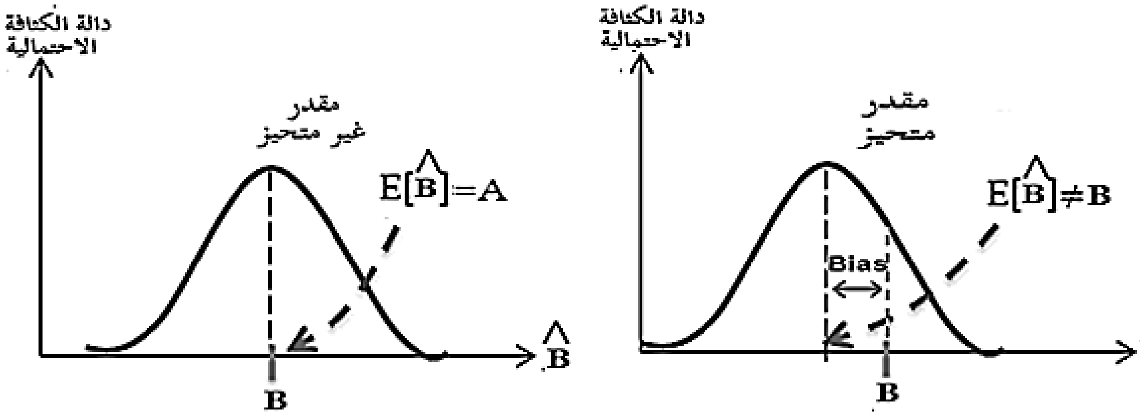
أي نقصد بعدم التحيّز أن يكون الفرق بين القيمة المتوقعة للمقدر $E(\hat{B})$، وقيمة المعلمة الحقيقية للمجتمع الاحصائي (B) يساوي الصفر[1]. $E(\hat{B}) - B = 0$

وبعبارة أخرى يعتبر المقدر غير متحيّز اذا كانت القيمة المتوقعة مساوية لمعلمة المجتمع موضع التقدير $E(\hat{B}) = B$.

وفي حالة المقدر المتحيز فإن الفرق $E(\hat{B}) - B$ = مقدار التحيز ($E(\hat{B}) - B \neq 0$).

[1] - حسين علي البخيت، سحر فتح الله، **"الاقتصاد القياسي"**، دار اليازوري للنشر والتوزيع، عمان، الأردن، 2009، ص 64.

الشكل رقم (III.1): التمثيل البياني لصفة عدم التحيز.



أمثلة:

$\bar{x}$: هو مقدر نقطي غير متحيز ل μ، إذا كان $E(\bar{x}) = \mu$.

$\hat{p}$: هو مقدر نقطي غير متحيز ل P، إذا كان $E(\hat{p}) = P$.

S^2: هو مقدر نقطي غير متحيز ل σ^2، إذا كان $E(S^2) = \sigma^2$. (حيث أن $E(S^2) = \sigma^2(n/n-1)$ ويتحقق ذلك في حالة العينات الكبيرة حيث تؤول n/n-1 إلى 1).

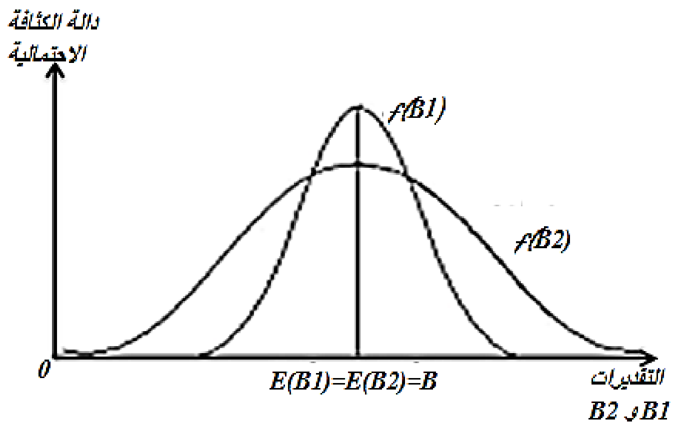
ب. خاصية الكفاءة:

تتعلق كفاءة مقدر ما بمقدار التباين لتوزيع المعاينة للإحصائية، أي أنه اذا كان لدينا مقدران B_1 و B_2، غير متحيزان فإن المقياس الذي يستخدمه الاحصائيون للمفاضلة بين المقدران هو اختيار ذلك التقدير الذي يعطي **أقل تباين ممكن**، أي نقول عن المقدر $\hat{B}$ ل B إذا تحقق ما يلي:

1) $E(\hat{B}_1) = B.$
2) $Var(\hat{B}_1) < Var(\hat{B}_2).$

ويمكن توضيح ذلك من خلال المنحنى التالي:

الشكل رقم (III.2): التمثيل البياني لصفة الكفاءة (أقل تباين).



مثال: لكل من توزيعي المعاينة للمتوسط والوسيط نفس المتوسط هو متوسط المجتمع μ، ولكن يعتبر المتوسط$\bar{x}$ مقدرا أكثر كفاءة لمتوسط المجتمع μ من الوسيط Me لأن:

تباين توزيع المعاينة للمتوسطات: $\delta_{\bar{x}} = V(\bar{x}) = {\delta_x^2}/{n}$.

بينما تباين توزيع المعاينة للوسيط:

$$\delta_{Me} = V(Me) = {\pi.\delta_x^2}/{2n} = \left({\delta_x^2}/{n}\right) \times \left({3.14159}/{2}\right).$$

ومنه فان: $\mathrm{Var}(\bar{X}) \times \frac{\pi}{2}\mathrm{Var}(Me)=$، وبالتالي فإن: $\bar{x}$ هو أفضل مقدر نقطي غير متحيز كفئ ل μ.

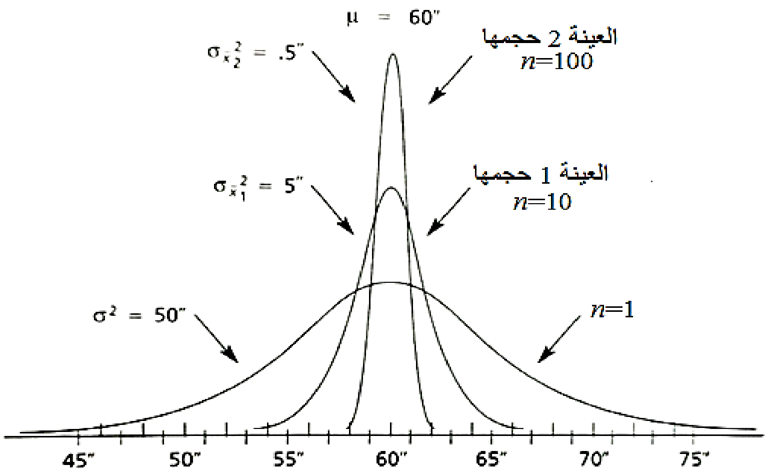
ت. خاصية التقارب:

في كثير من الأحيان لا يمكن الحصول على مقدرات تتميز بالخصائص المرغوبة في العينات الصغيرة، على غرار عدم التحيز أو الكفاءة، وفي هذه الحالات عادة ما يبحث عن هذه الخائص المرغوبة في العينات الكبيرة[1]. ونقول عن مقدر ما أنه متقارب إذا كان يؤول إلى قيمة المعلمة المقدرة عندما يؤول حجم العينة إلى مالا نهاية **مثال**: $\bar{x}$ مقدر نقطي متقارب لمتوسط المجتمع μ لأن:

$$\mathrm{E}(\bar{x}) = \mu \qquad\qquad \delta_{\bar{X}}^2 = V(\bar{x}) = {\delta_x^2}/{n} \xrightarrow[n\to+\infty]{} 0$$

ويمكن تمثيل خاصية االتقارب من خلال المنحنى التالي:

الشكل رقم (III.3): التمثيل البياني لخاصية التقارب.



---

[1]- G.S.Maddala, Op.Cit,P24.

III. 3. التقدير بمجال.

نادرا ما يتساوى تقدير النقطة والمعلمة المراد تقديرها، لذلك يتم بدلا من ذلك تحديد فترة أو مجال. ويشير التقدير بفترة إلى مدى يبين مجالا من القيم التي يحتمل أن تقع القيمة الفعلية المجهولة ضمنها[1]، أي تقدير معلمة المجتمع بنقطتين يحددان مجال لقيمة المعلمة. مثال: يكون تقدير بمجال اذا قلنا أن متوسط علامات الطلاب تساوي 2.5±13.5 أي أن متوسط علامات الطلاب في المجتمع يتراوح بين 11 و 16. ومجال الثقة (فترة الثقة *Confedence interval*) معرف بحدي الثقة، الحد الأدنى الذي يرمز له بالرمز *L* (*Lower limit*)، والحد الأعلى الذي يرمز له بالرمز *U* (*Upper limit*). وبشكل عام لتحديد فترة الثقة لأي معلمة كانت، ولتكن *B* لابد من ايجاد هذين الحدين، والصيغة الرياضية لتقدير حدود الثقة هي:

$$P(L \leq B \leq U)$$

$$= P(\hat{b}+ME \leq B \leq \hat{b}+ME)$$

$$=1- \alpha$$

حيث أن: L: الحد الأدنى لمجال الثقة.

U: الحد الأعلى لمجال الثقة.

$1- \alpha$: معامل الثقة حيث أن $(1- \alpha)$.100% هي درجة أو مستوى الثقة، وتجدر الاشارة إلى أنه كلما زاد هذا المقدرا كلما كانت فترة الثقة أكثر طولا. وأن تكون فترة الثقة أكثر طولا لا يعني بالضرورة أن تكون أفضل، وذلك لأنه اذا زاد طول مجال أو فترة التقدير عن حد معين تصبح الفائدة العلمية لها ذون جدى.

ME: (*margin of error*) ويسمى هامش الخطأ وهي مقدار أو قيمة تأخذ في الحسبان تقلبات معاينة المقدر $\hat{b}$ ومعامل الثقة $(1-\alpha)$.

α: احتمال الخطأ ويسمى أيضا مستوى المعنوية.

مثال:

على فرض أن متوسط علامات الطلاب في مادة ما ينتمي إلى المجال [11، 16] بمستوى معنوية 5%

أي أن: $\alpha =$ 5% مستوى المعنوية، أو احتمال الخطأ.

$1- \alpha =$ 95% مستوى الثقة.

11، و 16: حدود الثقة (11 الحد الأدنى، 16 الحد الأعلى)

[11، 16]: مجال الثقة، أو فترة الثقة.

[1] - طالب محمد عوض، **"مقدمة في الاقتصاد القياسي"**، منشورات الجامعة الأردنية، عمان، الأردن، 2000، ص 131

III. 2. 1. مجال الثقة للمتوسط.

نقطيا يقدر متوسط المجتمع μ من خلال الاحصائية المقابلة في العينة $\bar{X}$، ومن أجل انشاء مجال الثقة حول متوسط المجتمع نميز بين حالتين:

أ. تقدير متوسط المجتمع μ باستخدام التوزيع الطبيعي (حالة σ معروف).

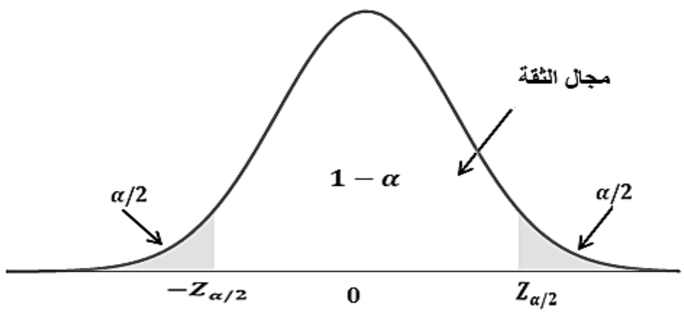
بهدف تقدير مجال الثقة لمتوسط المجتمع المجهول μ، نسحب عينة عشوائية حجمها n. وعلى فرض أن المجتمع الذي سحبت منه العينة موزع طبيعيا وتباينه σ^2 معلوم. أو في حالة العينات الكبيرة ($n\geq30$) حسب نظرية النهاية المركزية فإن $\bar{X}$ يتبع التوزيع الطبيعي، ونكتب:

$$\bar{X} \sim N(\mu, \sigma_{\bar{X}})$$

أي أن: $\frac{\bar{X}-\mu}{\sigma_{\bar{X}}} \sim N(0,1)$

ومنه يمكن تمثيل هذا التوزيع بيانيا كما يلي:

الشكل رقم (**III.4**): التوزيع Z.



من الشكل أعلاه: $P(-z_{\alpha/2}\leq z\leq + z_{\alpha/2})= 1- \alpha$،

وبما أن: $z = \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}}$

ومنه فإن:

$$P\left(-z_{\alpha/2} \leq \frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}} \leq +z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-z_{\alpha/2}\frac{\sigma_X}{\sqrt{n}} \leq \bar{X}-\mu_x \leq z_{\alpha/2}\frac{\sigma_X}{\sqrt{n}}\right) = 1-\alpha$$

وعليه يمكن تكوين فترة الثقة لمتوسط (حول متوسط) المجتمع 100(1-α) μ% كما يلي:

$$\boxed{\bar{X} - z_{1-\alpha/2}\frac{\sigma_X}{\sqrt{n}} \leq \mu_X \leq \bar{X} + z_{1-\alpha/2}\frac{\sigma_X}{\sqrt{n}}}$$

ويسمى المقدار $ME = z_{1-\alpha/2}\frac{\sigma_X}{\sqrt{n}}$ بهامش الخطأ (*Margin error*)، ويمكن كتابة مجال ثقة لمتوسط المجتمع كما يلي:

$$\boxed{P(\bar{X} - ME \leq \mu_X \leq \bar{X} + ME) = 1 - \alpha}$$

ملاحظة 1:

في حالة تباين المجتمع مجهول وحجم العينة n≥30، سنستخدم الانحراف المعياري للعينة كتقدير للانحراف المعياري المجتمع σ، وستكون فترة الثقة لمتوسط المجتمع μ كالتالي:

$\bar{X} \pm Z_{1-\frac{\alpha}{2}}.\frac{S}{\sqrt{n}}$ (في حالة حساب الانحراف المعياري بالقانون التالي: $S = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$)، أو

$\bar{X} \pm Z_{1-\frac{\alpha}{2}}.\frac{S}{\sqrt{n-1}}$ (في حالة حساب الانحراف المعياري بالقانون التالي: $S = \sqrt{\frac{\sum X_i^2}{n} - (\bar{X})^2}$).

كما أنه في حالة المعاينة نفادية، أو المجتمع محدود (*n>0.05N*)، فإن مجال الثقة يكتب بالشكل التالي:

$$\bar{X} \pm Z_{1-\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}}$$

ملاحظة 2[1]:

- حدد مجال الثقة على أساس أن المجتمع يتبع القانون الطبيعي ويجب مراعاة ذلك.
- إن طول مجال الثقة يكبر كلما كبرت قيمة الانحراف المعياري أو درجة الثقة أو الاثنين معا، وأيضا كلما صغر حجم العينة.
- التعليق على النتيجة: إن المجال بمقدار 100(*1-α*)% ثقة يعني أننا أخذنا عينات عشوائية متكررة من مجتمع وأن 95% (مثلا في حالة α= 5%) 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] - موساوي عبد النور، بركان يوسف، "**الاحصاء 2**"، دار العلوم للنشر والتوزيع، عنابة، الجزائر، 2010، ص 131.

مثال: مجتمع موزع طبيعيا انحرافه المعياري $\sigma = 3.75$، ومتوسطه μ، انطلاقا من عينة عشوائية من ذات المجتمع حجمها $n=15$، ومجموع مفرداتها $\Sigma X_i = 2400$، قدر فترة الثقة لمتوسط المجتمع μ المجهول عند مستوى معنوية $\alpha = 5\%$.

الحل:

العينة مسحوبة من مجتمع طبيعي تباينة معلوم $\sigma = 3.75$، نستخدم التوزيع الطبيعي لتقدير فترة الثقة للمتوسط:

$$IC_1 = [\bar{X} \pm Z_{1-\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}}] \ ;$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{2400}{15} = \boxed{160} \qquad Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.925} = \boxed{1.96}$$

$$\left[160 - (1.96).\left(\frac{3.75}{\sqrt{15}}\right);\ \ 160 + (1.96).\left(\frac{3.75}{\sqrt{15}}\right)\right]$$

$[158; 162]$ أي أن $158 < \mu < 162$ عند مستوى ثقة 95%.

ب. تقدير مجال الثقة لمتوسط المجتمع حالة σ مجهول (*t-Distribution*)

في الحقيقة عندما يكون متوسط المجتمع μ مجهولا فإن تباين المجتمع σ^2 أيضا يكون مجهولا، وعندما يكون المجتمع موزعا توزيعا طبيعيا و لكن σ غير معلومة وحجم العينة n صغير ($n<30$)، فإنه لا يمكننا استخدام التوزيع الطبيعي لتحديد مجال الثقة لمتوسط المجتمع μ المجهول، ولكن يمكننا استخدام توزيع ستيودنت *Student* ويرمز له ب t (*t-Distribution*).

تعريف توزيع t (*t-Distribution*)[*]:

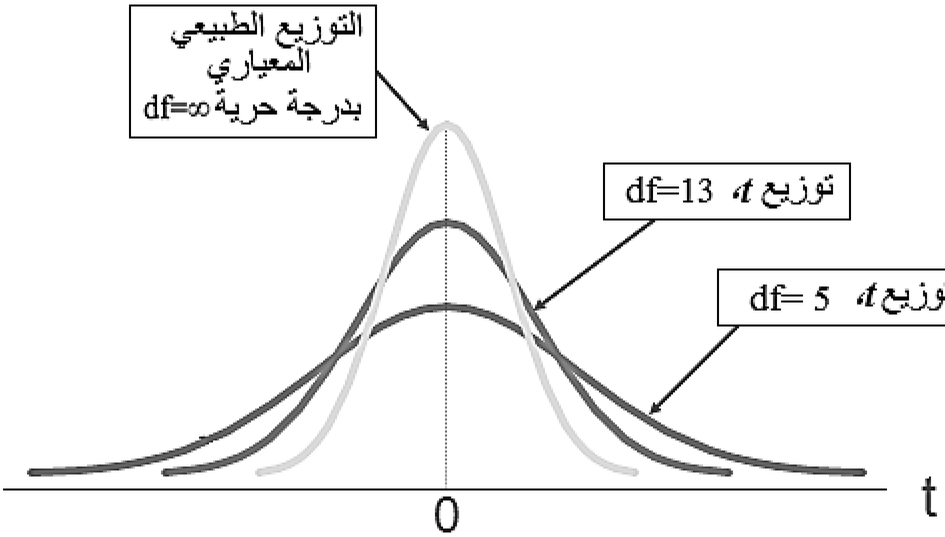
يعتبر هذا التوزيع من أهم التوزيعات الاحتمالية المستخدمة في التقدير واختبار الفرضيات، وخاصة عندما يكون حجم العينة **صغير** (***n<30***)، و الانحراف المعياري للمجتمع غير معلوم (**σ مجهول**).

توزيع t لا يمكن تمثيله بمنحى واحد، بل بعدد من المنحنيات التي يتميز كل منها عن الآخر بعدد درجات الحرية، فتزداد تفرطحا وتشتتا بالمقارنة بالمنحنى الطبيعي كلما نقص حجم العينة، وعلى

[*]– اكتشف هذا التوزيع من قبل العالم الألماني William Sealy Gosset، ونشره في مقال سنة 1908 تحت اسم مستعار توزيع طالب (Student Distribution).

العكس من ذلك يقترب توزيع ستيودنت من شكل المنحنى الطبيعي كلما كبر حجم العينة (عدد درجات الحرية)[1]. والشكل التالي يوضح خصائص هذا التوزيع:

الشكل رقم (III. 5): التمثيل البياني لتوزيع t *مقارنة* بتوزيع z.



الخاصية 1: نلاحظ أن توزيع t مثل توزيع z أيضا متناظر حول 0 أي $E(t_n)=0$.

الخاصية 2: يختلف التمثيل البياني ل t_n عن التمثيل البياني ل z في كون ذيلي توزيع ستيودنت أعرض من التمثيل البياني للتوزيع الطبيعي. ويشتركان في أن كلاهما التمثيل البياني لدالة كثافتهما يتخذ شكل جرسي، والمساحة تحت المنحنى الجرسي تساوي الواحد، إلا أن منحنى توزيع T أكثر تفرطحا من توزيع Z، ويعود ذلك لكون تباين T أكبر من تباين Z.

الخاصية الثالثة: عندما يؤول n إلى ما لانهاية $(n \to \infty)$، فإن t_n يقترب من z. وعادة يتم تقريب $t_{(n)}$ ب $N(0,1)$ إذا كان $30\leq n$

وعند استخدام توزيع t لتقدير مجال الثقة ل μ، يتم تعويض معلمة المجتمع σ المجهولة بمقدرها الانحراف المعياري للعينة S. ومن العلاقات السابقة للحصول على المتغير المعياري يمكن كتابة:

$$T = \frac{\bar{X} - \mu_X}{S_X/\sqrt{n}} \sim t_{n-1}$$

أي أن التوزيع الاحتمالي للإحصائية يتبع توزيع t بدرجة حرية $v=(n-1)$، وعليه فإن:

---

[1] - السعدي رجال، **"نظرية الاحتمالات: لكل التخصصات-لكل المستويات دروس وتمارين"**، الجزء الثاني، ديوان المطبوعات الجامعية، الجزائر، 1995، ص 245.

$$P\left(-t_{\left(\frac{\alpha}{2};n-1\right)} \leq T \leq +t_{\left(\frac{\alpha}{2};n-1\right)}\right) = 1-\alpha$$

وبالتعويض بقيمة T نجد:

$$P\left(-t_{\left(\frac{\alpha}{2};n-1\right)} \leq \frac{\bar{X}-\mu_X}{S_X/\sqrt{n}} \leq +t_{\left(\frac{\alpha}{2};n-1\right)}\right) = 1-\alpha$$

ومنه فإن:

$$P\left(\bar{X}-t_{\left(\frac{\alpha}{2};n-1\right)}\frac{s_X}{\sqrt{n}} \leq \mu_x \leq \bar{X}+t_{\left(\frac{\alpha}{2};n-1\right)}\frac{s_X}{\sqrt{n}}\right) = 1-\alpha$$

وعليه يمكن كتابة مجال الثقة لمتوسط المجتمع μ في حالة العينات الصغيرة باحتمال قدره $(1\text{-}\alpha)$ كالتالي:

$$\boxed{= P\left(\bar{X}-t_{\left(\frac{\alpha}{2},v\right)}\frac{s_X}{\sqrt{n}} \leq \mu_X \leq \bar{X}+t_{\left(\frac{\alpha}{2},v\right)}\frac{s_X}{\sqrt{n}}\right).}$$

مثال:

بهدف تقدير متوسط مجتمع ما μ المجهول، من ذات المجتمع تم سحب عينة عشوائية حجمها n=10، متوسطها $\bar{X}$= 5، وانحرافها المعياري S= 1. أوجد فترة الثقة حول متوسط المجتمع μ عند مستوى معنوية 5%.

الحل:

بما أن تباين المجتمع مجهول و $n = 10 < 30$ نستخدم توزيع t لتقدير فترة الثقة للمتوسط:

$$IC = \left[\bar{X} \pm t_{v,\frac{\alpha}{2}}.\frac{S}{\sqrt{n}}\right]$$

$$t_{n-1;\frac{\alpha}{2}} = t_{10-1;\frac{0.05}{2}} = t_{09;0.025} = \boxed{2.262}$$

$$\left[5-(2.262).\left(\frac{1}{\sqrt{10}}\right);\ \ 5+(2.262).\left(\frac{1}{\sqrt{10}}\right)\right]$$

$[4.28; 5.71]$ أي أن μ محصور بين القيمتين 4.28 و 5.71 عند مستوى ثقة 95%

III. 2. 2. مجال الثقة للنسبة.

كما سبق ذكره في المحور السابق (توزيعات المعاينة للنسبة) فإن نسبة صفة معينة في العينة$\hat{P}$ ، والمعرفة كما يلي: $E(\hat{P}) = P$,$V(\hat{P}) = \frac{P(1-P)}{n}$. تعتبر كمقدر نقطي لنسبة صفة معينة في المجتمع P المجهولة، وحسب نظرية النهاية المركزية فإن القانون الاحتمالي للإحصائية يقترب من الطبيعية اذا كان $\boldsymbol{n}$ كبيرا بما فيه الكفاية ($30 \leq n$) (تحقق شروط تقريب التوزيع الثنائي من التوزيع الطبيعي):

$$\frac{\hat{P} - E(\hat{P})}{\sigma_{\hat{P}}} = \frac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}} \sim N(0,1).$$

وبالتالي يمكن كتابة مجال الثقة لنسبة صفة معينة في المجتمع $\boldsymbol{P}$ المجهولة كما يلي:

$$P\left(\hat{P} - z_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}} \leq P \leq \hat{P} + z_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}}\right) = 1-\alpha$$

. والمشكلة في الصيغة أعلاه هي أن الصفة $\boldsymbol{p}$ غير معروفة، وبالتالي يتم تعويضها بمقدرها في العينة $\boldsymbol{\hat{P}}$ ونحصل على:

$$P\left(\hat{P} - z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \leq P \leq \hat{P} + z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right) = 1-\alpha$$

وعليه يمكن تكوين فترة الثقة (حول النسبة) لنسبة صفة معية في المجتمع P 100(1-α)% كما يلي:

$$\boxed{\hat{P} - z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \leq P \leq \hat{P} + z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}}$$

كما يمكن كتابة مجال الثقة حول نسبة المجتمع كما يلي:

$$\boxed{P(\hat{P} - ME \leq P \leq \hat{P} + ME) = 1-\alpha}$$

حيث أن هامش الخطأ: $ME = z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$، في حالة المجتمع غير محدود أو المعاينة غير نفادية والعينات الكبيرة ($0.05>n/N$)

بينما: $ME = z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\sqrt{\frac{N-n}{N-1}}$، في حالة حجم المجتمع N محدود، أو المعاينة نفادية ($n\geq 0.05N$).

مثال[1]:

مخزن يتكون من 10000 قطعة. لتقييم عدد القطع المعيبة في المخزن تم سحب عشوائيا 400 قطعة، وجد أن 45 قطعة معيبة. أعط مجال الثقة بنسبة 99% لإجمالي القطع المعيبة.

الحل:

نسبة القطع المعيبة في العينة: $\hat{P} = \frac{\hat{P}_a}{\hat{P}} = \frac{45}{400} = \boxed{0.1125}$

تباين النسبة في العينة (خطأ المعاينة): $\left(0.05 > 0.04 = \frac{400}{10000} = \frac{n}{N}\right) \xLeftarrow{\text{نهمل معامل الارجاع}}$

$$\sigma^2_{\hat{P}} = \frac{\hat{P}.(1-\hat{P})}{n} = \frac{(0.1125)(0.8875)}{400} = \boxed{0.00025}$$

ونكتب: $\hat{P} \sim N(0.4; 0.00025)$

وعليه فإن مجال الثقة لنسبة المجتمع: $IC = \left[\hat{P} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right]$

$(1-\alpha)=99\% \Leftarrow \alpha=1\%$ وبالتالي فإن: $z_{1-\alpha/2}=z_{1-0.01/2}=z_{0.995}=2.58$

$$0.1125 - (2.58)\sqrt{0.00025} \leq P \leq 0.1125 + (2.58)\sqrt{0.00025}$$

[0.071; 0.15]أي أن P نسبة القطع المعيبة في المجتمع محصورة بين القيمتين 0.071 و 0.15 عند مستوى ثقة 99%.

---

[1]- Michel Lejeune, '**Statistique la théorie et ses applications**', deuxième édition, Springer Verlag France, Paris, 2010, P 165.

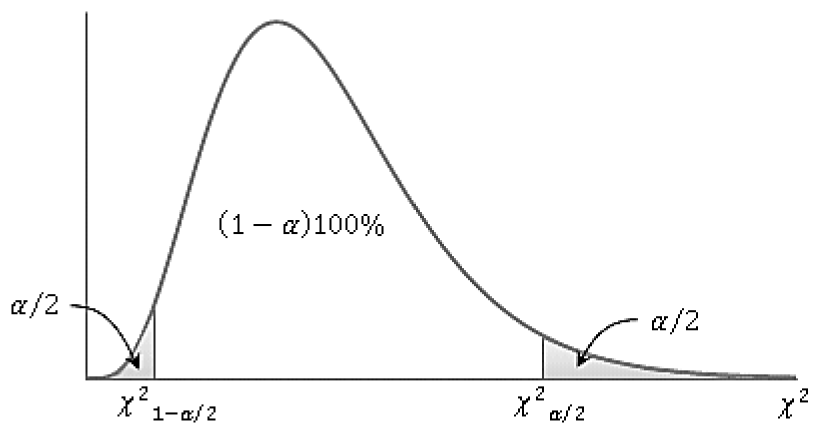
III. 2. 3. مجال الثقة للتباين.

نفرض أن المجتمع يتبع التوزيع الطبيعي، وبما أن المتغيرة $\frac{S^2}{\sigma^2}$ تتبع التوزيع ك (χ^2) بدرجة حرية (n-1) فانه لتقدير تباين المجتمع لمجتمع بمجال ثقة نستخدم الخاصية:

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

ومن التمثيل البياني لتوزيع (χ^2) التالي يمكننا استخراج فترة الثقة للتباين المجتمع المجهول:

الشكل رقم (**III. 6**): التمثيل البياني لتوزيع كاي مربع χ^2 بدرجة حرية v.



من الشكل أعلاه فإن:

$$P\left(x^2{}_{1-\left(\frac{\alpha}{2}\right)}^{(n-1)} \leq x^2 \leq x^2{}_{\frac{\alpha}{2}}^{(n-1)}\right) = 1-\alpha$$

ونكتب:

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

حيث أن $x^2{}_{\frac{\alpha}{2}}^{(n-1)}$، تمثل قيمة كاي، والتي يمكن استخراجها من الجدول انطلاقا من قيمة α ودرجة الحرية، وبحصر σ^2، نحصل على المجال التالي:

$$\boxed{\frac{(n-1)S^2}{x^2{}_{1-\left(\frac{\alpha}{2}\right)}^{(n-1)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{x^2{}_{\frac{\alpha}{2}}^{(n-1)}}}$$

ويمكن استنباط فترة الثقة **للانحراف المعياري σ** من مجال الثقة للتباين كالتالي:

$$\boxed{\frac{\sqrt{(n-1)}s}{\sqrt{x^2{}^{(n-1)}_{1-\left(\frac{\alpha}{2}\right)}}} \leq \sigma \leq \frac{\sqrt{(n-1)}s}{\sqrt{x^2{}^{(n-1)}_{\frac{\alpha}{2}}}}}$$

مثال[1]:

لغرض تقدير الانحراف المعياري لأوزان طلبة احدى الجامعات تم اختيار عينة عشوائية متكونة من 16 طالبا فوجد أن الانحراف المعياري للعينة (2.4كغ).

المطلوب: استخدم اختبار (χ^2) لتقدير الانحراف المعياري لأوزان طلبة الجامعة ضمن حدود ثقة مقدارها 99%.

الحل:

من جدول توزيع χ^2 (*chi squared*) نحصل على:

$$\begin{cases} v = 15;\ \alpha = 0.01 \Rightarrow \chi^2_{v;\frac{\alpha}{2}} = \chi^2_{15;0.005} = 4.6 \\ v = 15;\ \alpha = 0.01 \Rightarrow \chi^2_{v;1-\frac{\alpha}{2}} = \chi^2_{15;0.995} = 32.8 \end{cases}$$

وعليه فإن حدود الثقة للانحراف المعياري لأوزان الطلبة كما يلي:

$$\frac{\sqrt{(n-1)}s}{\sqrt{x^2{}^{(n-1)}_{1-\left(\frac{\alpha}{2}\right)}}} \leq \sigma \leq \frac{\sqrt{(n-1)}s}{\sqrt{x^2{}^{(n-1)}_{\frac{\alpha}{2}}}}$$

$$\frac{\sqrt{15}(2.4)}{\sqrt{x^2{}^{(15)}_{0.995}}} \leq \sigma \leq \frac{\sqrt{15}(2.4)}{\sqrt{x^2{}^{(15)}_{0.005}}} = \frac{\sqrt{15}(2.4)}{\sqrt{32.8}} \leq \sigma \leq \frac{\sqrt{15}(2.4)}{\sqrt{4.6}}$$

$$1.62 \leq \sigma \leq 4.33$$

أي أن الانحراف المعياري لمجتمع أوزان الطلبة محصور بين القيمتين 1.62 و 4.33 كغ عند مستوى معنوية 1%.

---

[1] - عدنان كريم نجم الدين، **"الاحصاء للاقتصاد والادارة"**، الطبعة الأولى، دار وائل للنشر، عمان، الأردن، 2000، ص 181.

III. 2. 4. الاستدلال الاحصائي حول مجتمعين.

تم التطرق أعلاه إلى الاستدلال الاحصائي (التقدير) حول مجتمع واحد (تقدير المتوسط μ، النسبة P،أو تباين σ^2 لمجتمع واحد). إلا أنه في كثير من الأحيان يتطلب الأمر مقارنة بين مجتمعين، ما يستدعي قياس الفرق بين متوسطي مجتمعين، أو تقدير الفرق بين نسبتي صفة معينة في مجتمعين مستقلين...إلخ، وفيما يلي سيتم تناول التقدير حول مجتمعين مستقلين.

أ. فترة الثقة للفرق بين متوسطي عينتين من مجتمعين طبيعيين

يمكن إيجاد فترة ثقة للفرق بين متوسطي مجتمعين من خلال سحب عينة عشوائية $\{X_1, X_2,..., X_n\}$ حجمها n من مجتمع ما موزع طبيعيا بمتوسط μ_X، وتباين σ_X^2، ونسحب عينة عشوائية أخرى حجمها m $\{Y_1, Y_2,..., Y_n\}$ من مجتمع ثاني موزع أيضا طبيعيا، بمتوسط μ_Y، وتباين σ_Y^2. فإذا كانت $\bar{X}$ تمثل متوسط العينة الأولى، و $\bar{Y}$ تمثل متوسط العينة الثانية، يمكن كتابة فترة الثقة للفرق بين متوسطي المجتمعين كما يلي:

$$\boxed{IC_{\mu_X-\mu_Y} = [(\bar{X}-\bar{Y}) \pm ME]}$$

ولحساب هامش الخطأ ME سنواجه الحالات التالية:

- تبايني المجتمعين σ_X^2، و σ_Y^2 معلومين.
- تبايني المجتمعين σ_X^2، و σ_Y^2 مجهولين و $\sigma^2 \neq \sigma_Y^2 \neq \sigma_X^2$.
- تبايني المجتمعين σ_X^2، و σ_Y^2 مجهولين و $\sigma^2 = \sigma_Y^2 = \sigma_X^2$.

الحالة 1: تبايني المجتمعين σ_1، و σ_2 معلومين:

في حالة ما اذا تم سحب عينتين مستقلتين من مجتمعين طبيعيين (أو أن طبيعة توزيع المجتمع غير معلومة وحجم العينتين كبير بالشكل الكافي حسب نظرية النهاية المركزية)، معلومي التباين، وحسب توزيع المعاينة للفرق بين متوسطين فإن المعلمة $D = (\bar{X}-\bar{Y})$ معرفة بالشكل التالي:

$$D \sim N(\mu_D, \sigma_D^2) \Rightarrow D \sim N\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$$

أي أن:

$$Z = \frac{(\bar{X}-\bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim N(0,1)$$

وعليه فإن فترة الثقة 100(1-α)% للفرق بين متوسطي عينتين مستقلتين (μ_X - μ_Y) يكتب كما يلي:

$$\boxed{(\bar{X}-\bar{Y}) - z_{1-\alpha/2}\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}} \leq \mu_X - \mu_Y \leq (\bar{X}-\bar{Y}) + z_{1-\alpha/2}\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}}$$

مثال:

في دراسة لمعرفة متوسط استهلاك الكهرباء في بعض مدن الجمهورية، تم سحب عينة من مدينة (أ) حجمها 200 أسرة فوجد أن متوسط استهلاك الأسرة 150 كيلو وات شهريا بانحراف معياري 50 كيلو وات، وسحبت عينة من مدينة (ب) حجمها 100 أسرة فوجد أن متوسط استهلاك الأسرة 90 كيلو وات شهريا بانحراف معياري 30 كيلو وات.

المطلوب: تقدير الفرق بين متوسطي استهلاك الأسرة من الكهرباء شهريا بين المدينتين وبدرجة ثقة 95%.

الحل:

(أ): $n = 200$ ، $\bar{X}_1 = 150$ ، $S_1 = 50$.

(ب): $m = 100$ ، $\bar{X}_2 = 90$ ، $S_2 = 30$.

n، $m \geq 30$ ⇐ نستخدم القيمة المعيارية Z

$$IC: \left[(\bar{X}_1-\bar{X}_2) \pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}}\right]$$

$$= \left[(150-90) \pm 196\sqrt{\frac{(50)^2}{200}+\frac{(30)^2}{100}}\right] \Rightarrow IC: [50.9, 69.1]$$

ومعنى ذلك أن الفرق بين متوسط استهلاك الأسرة من الكهرباء شهريا في المدينة (أ) ومتوسط استهلاك الأسرة في المدينة (ب) يتراوح بين 51 و 69 كيلو وات عند مستوى معنوية 5%.

الحالة 2: تبايني المجتمعين σ_X^2، و σ_Y^2 مجهولين، حجم العينات كبير ($30\leq n, m$).

في حالة سحب عينتين مستقلتين من مجتمعين طبيعيين، مجهولي التباين، سنستخدم تباين العينتين S_X^2، و S_Y^2، وفي حالة ما اذا كان حجم العنتين n و m كبير **($30\leq n, m$)** فإن:

$$(\bar{X}-\bar{Y})-ME<\mu_X-\mu_Y<(\bar{X}-\bar{Y})+ME$$

حيث أن:

$$ME=z_{1-\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}$$

وبالتالي فإن فترة الثقة $100(1-\alpha)$% للفرق بين متوسطي مجتمعين مجهولي التباين و**($30\leq n, m$)** هي:

$$\boxed{(\bar{X}-\bar{Y})-z_{1-\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}\leq\mu_X-\mu_Y\leq(\bar{X}-\bar{Y})+z_{1-\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}}$$

من المحتمل أن يكون تباين المجتمعين متساويين $\sigma^2=\sigma_Y^2=\sigma_X^2$، وبالتالي بدلا من استخدام الانحراف المعياري للعينتين لحساب الانحراف المعياري للفرق، سيتم حساب انحراف معياري مشترك (أو تباين مشترك).

$$\boxed{S^2=\frac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2}}$$

وفي حالة العينات الكبيرة ($n+m-2\geq 30$) فإنه من الملائم استخدام الإحصاءه Z كما يلي:

$$Z=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{S^2}{n}+\frac{S^2}{m}}}$$

وعليه فإن فترة الثقة $100(1-\alpha)$% للفرق بين متوسطي عينتين مستقلتين (μ_X - μ_Y) يكتب كما يلي:

$$\boxed{(\bar{X}-\bar{Y})-z_{1-\alpha/2}.S\sqrt{\frac{1}{n}+\frac{1}{m}}\leq\mu_X-\mu_Y\leq(\bar{X}-\bar{Y})+z_{1-\alpha/2}.S\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

مثال[1]:

تم اجراء دراسة جامعية لتحديد ما اذا كانت ملكية السيارة محددة لتحقيق الانجاز الأكاديمي. بنيت الدراسة على عينتين عشوائيتين من 20 طالبا. كان متوسط المعدلات التراكمية للطلاب ال 20 الذين لا يملكون سيارات هو 2.75 بتباين يعادل 0.36، بينما متوسط المعدلات التراكمية للطلاب ال 20 الذين يملكون سيارات 2.51 بتباين 0.40. أوجد فترة ثقة بنسبة 90% للفرق بين متوسطي المعدلات التراكمية للطلاب مالكي السيارات مقابل الطلاب الذين لا يملكون سيارات.

افترض أن معدلات الطلاب التراكمية تتبع توزيعا طبيعيا وأن تبايني المعدلات التراكمية للمجتمعين متساوية.

الحل:

بما أن الانحراف المعياري للمجتمعين (تبايني المجتمعين) غير معروفين ولكن يفترض أنها متساوية وأن $n+m-2 \geq 30$ (38>30)، وعليه يجب أن نستخدم الصيغة التالية:

$$(\bar{X}-\bar{Y}) \pm z_{1-\alpha/2} \cdot \sqrt{\frac{S^2}{n}+\frac{S^2}{m}}$$

بما أن: $\alpha = 10\% \Leftarrow Z_{1-\alpha/2} = Z_{0.95} = 1.65$.

حساب الانحراف المعياري المشترك للعينتين:

$$S=\sqrt{\frac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2}}=\sqrt{\frac{19(0.36)+19(0.40)}{20+20-2}}=\boxed{0.6164}$$

حساب حدي فترة الثقة:

$$IC:\left[(2.75-2.51) \pm 1.65(0.6164)\sqrt{\frac{1}{20}+\frac{1}{20}}\right] \Rightarrow IC:[-0.081, 0.561]$$

[1] - فؤاد عبد الله العواد، **"مبادئ التحليل الاحصائي"**، 2015، ص ص 91-92، موجود على الموقع: http://fac.ksu.edu.sa/sites/default/files/mbdy_lthlyl_lhsyy35.pdf. تم زيارة الموقع بتاريخ 2018/02/20، الساعة 22:00.

إننا واثقون بنسبة 90% بأن الفترة من -0.081 إلى 0.561 تتضمن الفرق الحقيقي بين متوسطي المعدلات التراكمية للمجموعتين.

التفسير العلمي: بما أن الحد الأعلى موجب والحد الأدنى سالب، فإن القيمة (صفر) متضمنة في الفترة. وحيث أن الفرق صفر محصور ضمن هذين الحدين، فمن الممكن أنه لا يوجد فرق بين متوسطي المعدلات التراكمية للمجموعتين.

الحالة 3: تبايني المجتمعين σ_X^2، و σ_Y^2 مجهولين وحجم العينات صغير ($30>n, m$).

وفي حالة حجم العنتين n و m صغير **($30>n, m$)** فإننا سنستخدم التوزيع t لتقدير فترة الثقة:

ونكتب: $t = \frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{S_D}$، ونميز أيضا بين حالتين:

تبايني المجتمعين σ_X^2، و σ_Y^2 مجهولين و $\sigma^2 \neq \sigma_Y^2 \neq \sigma_X^2$

عليه فإن $ME = t_{n+m-2,\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}$

وبالتالي فإن فترة الثقة (1-α)100% للفرق بين متوسطي مجتمعين في هذه الحالة هي:

$$\boxed{(\bar{X}-\bar{Y}) - t_{v,\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}} \leq \mu_X-\mu_Y \leq (\bar{X}-\bar{Y}) + t_{v,\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}}$$

وفي حالة تبايني المجتمعين σ_X^2، و σ_Y^2 مجهولين وحجم العينات صغير و $\boldsymbol{\sigma^2 = \sigma_Y^2 = \sigma_X^2}$.

فإن: $ME = t_{n+m-2,\alpha/2}.S\sqrt{\frac{1}{n}+\frac{1}{m}}$

وبالتالي فإن فترة الثقة (1-α)100% للفرق بين متوسطي مجتمعين في هذه الحالة هي:

$$\boxed{(\bar{X}-\bar{Y}) - t_{v,\frac{\alpha}{2}}.S\sqrt{\frac{1}{n}+\frac{1}{m}} \leq \mu_X-\mu_Y \leq (\bar{X}-\bar{Y}) + t_{v,\frac{\alpha}{2}}.S\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

مثال:

نرغب بتقدير أرباح المستهلك الناجمة عن استخدامه وقود جديد لسيارته. اثبت اختبار مخبري أجري على 20 محرك لنفس النوع. 10 منها تم تزويدها بوقود تقليدي، خلال فترة معينة استهلكت ما يقدر ب 11.65 لتر بانحراف معياري 0.21 لتر. وبالنسبة للعشر محركات الباقية تم استخدام الوقود الجديد ولوحظ أن متوسط الاستهلاك قدر ب 10.8 لتر بانحراف معياري 0.18 لتر. على فرض أن العينتين مسحوبتين من مجتمعين موزعين طبيعيا اعط تقدير لمجال الثقة بنسبة 90% للفرق بين متوسط الربح الذي يحققه المستهلك من استخدامه للوقود الجديد والوقود التقليدي.

الحل:

لدينا عينتين مسحوبتين من مجتمعين طبيعيين مجهولي التباين، $n=m=10<30 \Leftarrow$ نستخدم توزيع t لتقدير فترة الثقة للفرق بين المتوسطين. والمجال المستخدم في التقدير هو:

$$(\bar{X}-\bar{Y})-t_{v,\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}} \leq \mu_X-\mu_Y \leq (\bar{X}-\bar{Y})+t_{v,\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}$$

$$\alpha=10\%,\ n=m=10 \Rightarrow t_{v,\alpha/2}=t_{n+m-2,\alpha/2}=t_{10+10-2,\ 0.05}=t_{18,0.05}=1.734$$

$$IC=\left[(\bar{X}-\bar{Y}) \pm t_{v,\alpha/2}\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}\right]$$

$$=\left[(11.65-10.8) \pm (1.734)\sqrt{\frac{(0.21)^2}{10}+\frac{(0.18)^2}{10}}\right] \Rightarrow IC:[0.7,\ 1]$$

وعليه فإن متوسط الفرق بين استهلاك الوقود الجديد والوقود التقليدي محصور بين 0.7 و 1 لتر عند مستوى ثقة 90%.

أ. فترة الثقة للفرق بين نسبتي صفتين

فيما يلي طريقة إيجاد فترة ثقة حول الفرق بين نسبتي مجتمعين. إذا قمنا بسحب عينة عشوائية حجمها n $\{X_1, X_2,..., X_n\}$ من مجتمع ما موزع طبيعيا نسبة صفة معينة فيه P_X مجهولة، مقدرها النقطي نسبة هذه الصفة في العينة $\widehat{P}_X$، إذا كانت $\boldsymbol{n}$ كبيرة بما فيه الكفاية فإنه حسب نظرية النهاية المركزية فإن:

$$\widehat{P}_X \sim N\left(P_X, \frac{P_X(1-P_X)}{n}\right).$$

ونسحب عينة عشوائية أخرى $\{Y_1, Y_2,..., Y_n\}$ حجمهاm من مجتمع ثاني موزع أيضا طبيعيا، نسبة صفة معينة فيه P_Y مجهولة، مقدرها النقطي نسبة هذه الصفة في العينة التي سحبت من هذا المجتمع $\widehat{P}_Y$، إذا كانت $\boldsymbol{m}$ كبيرة بما فيه الكفاية فإنه حسب نظرية النهاية المركزية فإن:

$$\widehat{P}_Y \sim N\left(P_Y, \frac{P_Y(1-P_Y)}{m}\right).$$

وكما سبق ذكره في المحور السابق (توزيعات المعاينة لفرق نسبتين)، فإنه اذا كانت العينتين المسحوبتين مستقلتين، فإن فرق النسبتين $(P_X - P_Y)$ يمثل متغير عشوائي جديد والمعرفة ب:

متوسط: $\mu_{P_X-P_Y} = P_X - P_Y$، وتباين : $\sigma^2_{P_X-P_Y} = \frac{P_X(1-P_X)}{n} + \frac{P_Y(1-P_Y)}{m}$

حيث أن: $$Z = \frac{(\widehat{P}_X-\widehat{P}_Y)-(P_X-P_Y)}{\sqrt{\frac{P_X(1-P_X)}{n}+\frac{P_Y(1-P_Y)}{m}}} \sim N(0,1)$$

وبالتالي يمكن كتابة مجال الثقة لنسبة صفة معينة في المجتمع $\boldsymbol{P}$ المجهولة كما يلي:

$$P\left(-z_{1-\alpha/2} \leq \frac{\left(\widehat{P}_X-\widehat{P}_Y\right)-(P_X-P_Y)}{\sqrt{\frac{P_X(1-P_X)}{n}+\frac{P_Y(1-P_Y)}{m}}} \leq +z_{1-\alpha/2}\right) = 1-\alpha$$

$$P\left(\left(\widehat{P}_X-\widehat{P}_Y\right) - z_{\frac{\alpha}{2}}.\sigma_{\widehat{P}_X-\widehat{P}_Y} \leq (P_X-P_Y) \leq \left(\widehat{P}_X-\widehat{P}_Y\right) + z_{\frac{\alpha}{2}}.\sigma_{\widehat{P}_X-\widehat{P}_Y}\right) = 1-\alpha$$

وعليه يمكن تكوين فترة الثقة 100(1-α)% لتقدير الفرق $(P_X - P_Y)$ كما يلي:

$$\boxed{\left(\hat{P}_X-\hat{P}_Y\right)-EM\leq\left(P_X-P_Y\right)\leq\left(\hat{P}_X-\hat{P}_Y\right)+EM}$$

حيث أن: $EM=Z_{1-\alpha/2}\sqrt{\frac{P_X(1-P_X)}{n}+\frac{P_Y(1-P_Y)}{m}}$

مثال:

نرغب بتقدير الفرق بين نسب القطع التالفة في عمليتي انتاج مختلفتين. ولهذا تم سحب 800 قطعة عشوائيا خلال عملية الانتاج الأولى. وبعد اختبارها تم ايجاد أن 92 من بينها تالفة.

وبنفس الطريقة تم سحب 1000 قطعة خلال عملية الانتاج الثانية ووجد بأن 86 قطعة من بينها تالفة.

اعط تقدير لفترة الثقة للفرق بين نسبتي القطع التالفة في عمليتي الانتاج عند مستوى معنوية 10%.

الحل:

نسبة القطع التالفة في العينة الأولى: $\hat{P}_1=\frac{\hat{P}_{a1}}{\hat{P}_1}=\frac{92}{800}=\boxed{0.115}$

نسبة القطع التالفة في العينة الثانية: $\hat{P}_2=\frac{\hat{P}_{a2}}{\hat{P}_2}=\frac{86}{1000}=\boxed{0.086}$

$$\sigma_{\hat{P}_1-\hat{P}_2}=\sqrt{\frac{\hat{p}_1.\hat{q}_1}{n}+\frac{\hat{p}_2.\hat{q}_2}{m}}=\sqrt{\frac{(0.115)(0.885)}{800}+\frac{(0.086)(0.914)}{1000}}=\boxed{0.014}$$

وبما أن حجم العينات كبير بالشكل الكافي، وعليه فإن مجال الثقة للفرق بين النسبتين كما يلي:

$$IC=\left[(\hat{P}_1-\hat{P}_2)\pm Z_{1-\alpha/2}\sqrt{\frac{\hat{p}_1.\hat{q}_1}{n}+\frac{\hat{p}_2.\hat{q}_2}{m}}\right]$$

$(1-\alpha)=90\%\Leftarrow\alpha=10\%$ وبالتالي فإن: $Z_{1-\alpha/2}=Z_{1-0.1/2}=Z_{0.95}=1.65$

$$0.029-(1.65)(0.014)\leq P\leq 0.029+(1.65)(0.014)$$

$$IC:[0.0059;0.0521]$$

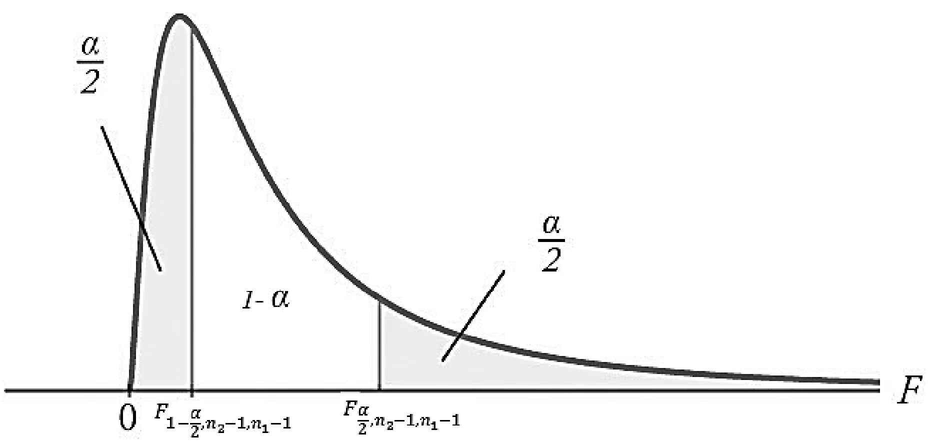
أي أن P نسبة القطع التالفة في عمليتي الانتاج محصورة بين القيمتين 0.0059 و0.0521 عند مستوى ثقة 90%.

ب. فترة الثقة لنسبة تباينين:

يمكن المقارنة بين مجتمعين من خلال المقارنة بين تباينهما، وتكون هذه المقارنة حسب ما تم التطرق إليه سابقا في توزيعات المعاينة من خلال نسبة تبايني هذين المجتمعين المراد المقارنة بينهما، إلا أنه عادة ما تكون تباينات المجتمعات مجهولة، وفيما يلي سيتم ابراز كيفية استخراج فترة الثقة لنسبة تبايني مجتمعين.

وكما هو معلوم (حسب الفصل السابق)، فإن الإحصائية $F = \frac{\hat{S}_Y^2/\sigma_X^2}{\hat{S}_Y^2/\sigma_Y^2}$ تتبع توزيع فيشر بدرجتي حرية$v_1=n-1$، و $v_2=m-1$، ويمكن استخلاص فترة الثقة لنسبة تبايني مجتمعين انطلاقا من شكل توزيع فيشر:

الشكل رقم (7 .III): التمثيل البياني لتوزيع فيشر.



من الشكل أعلاه فإن:

$$P\left(\mathrm{f}_{1-\frac{\alpha}{2},n-1,m-1} \leq F \leq \mathrm{f}_{\frac{\alpha}{2},n-1,m-1}\right) = 1-\alpha$$

$$P\left(\mathrm{f}_{1-\frac{\alpha}{2},n-1,m-1} \leq \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \leq \mathrm{f}_{\frac{\alpha}{2},n-1,m-1}\right) = 1-\alpha$$

ومنه فإن:

$$P\left(\frac{S_X^2/S_Y^2}{\mathrm{f}_{\frac{\alpha}{2},n-1,m-1}} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{S_X^2/S_Y^2}{\mathrm{f}_{1-\frac{\alpha}{2},n-1,m-1}}\right) = 1-\alpha$$

وعليه يمكن كتابة مجال الثقة لنسبة تباينين $\frac{\sigma_X^2}{\sigma_Y^2}$ باحتمال قدره $(1\text{-}\alpha)$ كالتالي:

$$\boxed{\frac{S_X^2/S_Y^2}{\mathrm{f}_{\frac{\alpha}{2},n-1,m-1}} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{S_X^2/S_Y^2}{\mathrm{f}_{1-\frac{\alpha}{2},n-1,m-1}}}$$

مثال[1]:

في قياس لمحتويات 6 صناديق من الخبيز تم تعبئتها باستخدام الآلة (A) كانت $S_1^2= 0.1754$، وفي قياس لمحتويات 11 صندوق من الخبيز تم تعبئتها باستخدام الآلة (B) فكانت $S_2^2= 0.2704$، وعلى افتراض أن الكمية المعبأة تتبع التوزيع الطبيعي لكل آلة أوجد 95% فترة ثقة لتقدير $\frac{\sigma_1^2}{\sigma_2^2}$ ؟

الحل:

بما أن $(1\text{-}\alpha)=0.95 \Leftarrow 1\text{-}\alpha/2=0.975 \Leftarrow \alpha/2= 0.025 \Leftarrow \alpha=0.05$، ومن جدول F نجد أن:

$$\mathrm{f}_{0.025;5;10} = 4.24 \text{ و } \mathrm{f}_{0.975;5;10} = \frac{1}{\mathrm{f}_{0.25;10;5}} = \frac{1}{6.62} = 0.151$$

ومنه فإن:

$$\frac{0.1754/0.2704}{4.24} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{0.1754/0.2704}{0.151}$$

حيث أن فترة الثقة تتضمن القيمة 1، وعليه لا يوجد اعتقاد بأن هناك فرق ما بين σ_1^2 و σ_2^2.

[1] - علي عبد السلام العماري، علي حسين العجيلي، مرجع سبق ذكره، ص 492.

تمارين حول الفصل (السلسلة 03):

التمرين 01:

سحبت عينة عشوائية حجمها 25، متوسطها 4290 من مجتمع موزع طبيعيا وانحرافه المعياري 1000. قدر متوسط المجتمع μ بثقة مقدارها : أ- 90%   ب- 95 %   ج- 99 % ؟

التمرين 02:

أعطى تحليل نسبة النيترات (ملغ/1ل) ل 10 قوارير ماء من نبع ماء القيم التالية: 3.52 3.55
3.52 3.44 3.62 3.64 3.56 3.67 3.56 3.61
قدر مجال الثقة عند مستوى ثقة 95% للانحراف المعياري لنسبة النيترات في قوارير الماء.

التمرين 03:

فيما يلي توزيع 100 موظف، اختيروا عشوائيا من إحدى الشركات الكبرى، حسب الزيادة السنوية التي حصلوا عليها في الراتب:

الزيادة	20-30	30-40	40-50	50-60	60-70	70-80
عدد الموظفين	10	15	30	22	14	9

قدر متوسط الزيادة في الراتب للموظف في الشركة كلها بمستوى ثقة 95%.

التمرين 04:

مخزن يتكون من 10000 قطعة. لتقييم عدد القطع المعيبة في المخزن تم سحب عشوائيا 400 قطعة، وجد أن 45 قطعة معيبة. أعط مجال الثقة بنسبة 99% لإجمالي القطع المعيبة.

التمرين 05:

تم سحب عينة عشوائية حجمها 49 من الانتاج السنوي لأحد مناجم الذهب يتوزع وفق التوزيع الطبيعي، فوجد أن متوسط انتاج هذا المنجم يساوي 150 طن وانحراف معياري يساوي 20 طن. بينما تم سحب عينة أخرى حجمها 36 من الانتاج السنوي لمنجم آخر يتوزع وفق التوزيع الطبيعي، فوجد أن متوسط الانتاج يساوي 125 طن وبانحراف معياري يساوي 25 طن قدر فترة الثقة للفرق بين متوسط انتاج المنجم الأول والمنجم الثاني عند مستوى معنوية 5%.

التمرين 06:

اجري امتحان في الإحصاء لمجموعتين من الطلبة من جامعتين مختلفتين، فإذا نجح في الامتحان 39 طالب من المجموعة الأولى والمتكونة من 65 طالب ونجح 24 طالب من المجموعة الثانية المتكونة من 60 طالب.

أوجد تقدير الفترة للفرق بين نسبتي الطلبة الناجحين في المجموعتين بمعامل ثقة 90%؟

التمرين 07:

نريد تقدير تباين مردودية سماد يستخدم لزراعة القمح. تم أخذ عينات من 12 قطعة ارض، فكانت المردودية بالطن للهكتار كالاتي: 8.3 8.4 7.8 7.6 8.2 8.4 8.5 7.9 8.2 7.8 8.4 7.7.

اعط مجال الثقة 95% لتباين المردودية للسماد.

التمرين 08:

ابرزت احدى الدراسات حول مجموعة مكونة من 11 مريضا من مرضى السكري أن الانحراف المعياري لتركيز الجلوكوز في الدم يقدر ب2.41 غ، بينما كان الانحراف المعياري لتركيز الجلوكوز في الدم لمجموعة أخرى مستقلة عن المجموعة الأولى مكونة من 4 مرضى يقدر 1.87 غ

كون فترة الثقة بنسبة 95% لنسبة تبايني تركيز الجلوكوز في الدم لمجتمعي المرضى.

الفصل الثالث: اختبار الفرضيات.

مقدمة:

انطلاقا من توزيعات المعاينة نعلم أن قيمة معلمة ما تختلف من عينة إلى أخرى، ويرجع ذلك إلى الطبيعية العشوائية لسحب العينات الاحصائية، وقد تمكنا من خلال فصل التقدير وتقدير فترات الثقة من الحصول على تقديرات لمعالم المجتمع المجهولة انطلاقا من بيانات العينة، ويبقى التساؤل المطروح هو مدى مصداقية القيم المقدرة في تمثيلها لقيم معالم المجتمع، وبهدف التأكد من توافق خصائص المجتمع مع التقديرات المستنتجة من بيانات العينة سيتم التطرق إلى ما يعرف باختبارات الفروض الاحصائية.

وتعتبر اختبارات الفروض الإحصائية (*Testing Statistical Hypotheses*) واحدة من أهم التطبيقات التي قدمها علم الإحصاء، إذ تمثل منهجية تساعد على اتخاذ القرارات في الاحصاء الاستدلالي. حيث أن الهدف من اجراء أي اختبار هو اتخاذ قرار على اساس بيانات عينة عشوائية من مجتمع ما، ما اذا كانت معلمة من هذا المجتمع تتوافق مع تنبؤات أو توقعات معينة تسمى الفرضية.

IV/ 1. مفاهيم عامة حول اختبار الفرضيات الاحصائية

IV/ 1. 1. تعريف اختبار الفرضيات

يعتبر اختبار الفروض أحد أساليب الاحصاء الاستدلالي الذي يستخدم فيه بيانات العينة المسحوبة من مجتمع الدراسة لاتخاذ قرارات أو اصدار أحكام حول هذا المجتمع. وعليه فإن اختبار الفرضيات الاحصائية هو اختبار صحة وصدقية القيمة (القيم) التي تأخذها المعلمة (المعلمات) التي تم تقديرها، وعادة ما تجرى الاختبارات حول قيمة محددة، أو حول النطاق (الفترة *Interval*) التي وقعت ضمنها أو خارجها[1]. وفي اختبار الفرضيات نفترض احتمالين، يمثلان **"الفرضيات"** حول طبيعة الحالة. وتنقسم اختبارات الفروض الاحصائية إلى قسمين:

أولا: اختبارات الفروض الاحصائية المعلمية (*parametric*): أي فرضيات حول معالم المجتمع، حيث يكون شكل توزيع بيانات المجتمع معلوم.

[1] - عبد الرزاق بني هاني، مرجع سبق ذكره، ص 377.

ثانيا: اختبارات الفروض الاحصائية اللامعلمية (*Nonparametric*): في بعض الأحيان يكون من الصعوبة بما كان تحديد شكل توزيع بيانات المجتمع، وفي هذه الحالة يتم استخدام الفروض الاحصائية اللامعلمية والتي لا تتطلب معرفة شكل التوزيع الذي تتبعه بيانات المجتمع.

وخلال هذا الفصل سيتم التركيز على النوع الأول فقط (اختبارات الفروض الاحصائية المعلمية).

IV/ 1. 2. تعريف الفرضية الاحصائية

الفرضية الاحصائية هي عبارة عن ادعاء أو تخمين معين حول معلمة من معالم المجتمع، وهذا الفرض قابل لأن يكون صحيحا أو غير صحيح، ولا يمكن التأكد من ذلك إلا من خلال اختباره احصائيا.

IV/ 1. 3. أنواع الفرضيات

يجب أن تكون الفرضيات متكاملة بحيث تشمل كل النتائج الممكنة (تشمل كل حوادث) وتبادلية بحيث تلغي احداهما الأخرى. مثل "المتهم مذنب" و "المتهم بريء".

أ. **الفرضية الصفرية (فرضية العدم) *The Null hypothesis*.**

وهي الفرضية الأساسية المراد اختبارها، ويرمز لها بالرمز H_0، وهي فرضية حول معلمة المجتمع التي نجري عليها الاختبار باستخدام بيانات العينة، ويشير فرض العدم إلى عدم وجود فرق بين معلمة المجتمع والعينة المسحوبة منه.

مثلا: على فرض أننا نود اجراء الاختبار الإحصائي حول متوسط أجور العمال في مصنع ما، حيث يدعي مسؤول في هذا المصنع أن متوسط أجور العاملين فيه تقدر ب 30000 دج، وللتأكد من صحة هذا الادعاء يتم اختيار عينة عشوائية وحساب متوسط أجور العمال في العينة، وفي هذه الحالة يتم صياغة الفرضية الصفرية (أو فرض العدم) التالية:

$$H_0: \mu = 30000.$$

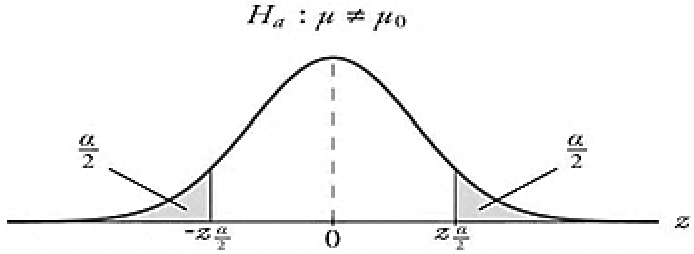
ب. **الفرضية البديلة *The Alternative hypothesis*.**

وهي خلاف فرضية العدم، أي أنه عند رفض الفرضية الصفرية (H_0) تكون الفرضية البديلة صحيحة ويرمز لها بالرمز H_a. وتجدر الاشارة بأن الفرض البديل هو الذي يحدد نوع الاختبار، حيث نميز بين نوعين من الاختبارات، اختبار من طرفين، واختبار من طرف واحد.

أولا: في حالة اختبار من طرفين (*Two Tails*) يتم صياغة H_0، وH_a كالتالي:

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

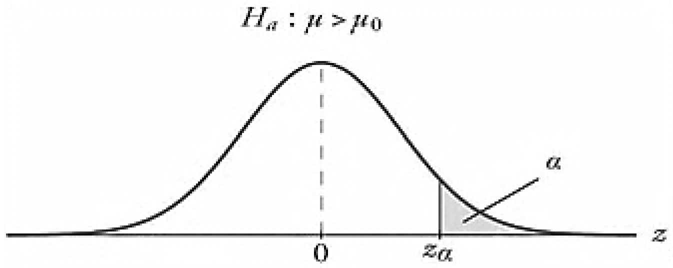
وبالعودة للمثال السابق، الفرض البديل هو: H_a:μ≠30000. ويمكن تمثيله بيانيا كما يلي:



وتسمى المنطقة المحصورة بين $-z_{\alpha/2}$ و $z_{\alpha/2}$ منطقة القبول ($-z_{\alpha/2}<z<z_{\alpha/2}$)، بينما المناطق المظللة تسمى منطقتي الرفض ($z<-z_{\alpha/2}$) و ($z_{\alpha/2}<z$)، بينما تسمى $-z_{\alpha/2}$ و $z_{\alpha/2}$ القيم الحرجة.

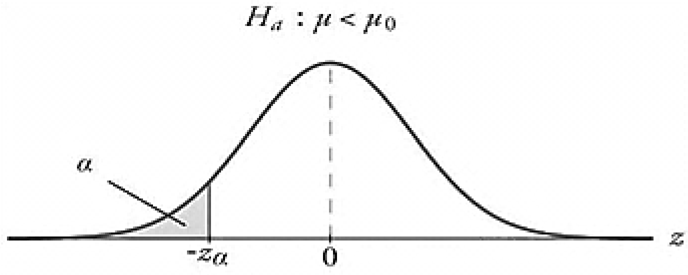
ثانيا: في حالة اختبار من جانب واحد (*One Tail*)، وتكون جهة الرفض إما من اليمين أو من اليسار.

1/ اختبار من طرف واحد، من الجهة اليمنى، وحسب المثال السابق، يصاغ الفرض البديل كما يلي: H_a:μ>30000. وتمثل بيانيا كما يلي:



حيث أن منطقة القبول في هذه الحالة هي: $z<z_\alpha$.

2/ اختبار من طرف واحد، الجهة اليسرى، وبالعودة للمثال السابق، الفرض البديل هو: H_a:μ<30000. وتمثل بيانيا كما يلي:



حيث أن منطقة القبول في هذه الحالة هي: $z> -z_\alpha$.

IV/ 1. 4. أنواع الأخطاء الاحصائية (عند اتخاذ القرار):

إن الفرض الذي يخضع للاختبار الاحصائي هو الفرض الصفري (فرض العدم) H_0، و يتم اتخاذ قرار قبوله أو رفضه على أساس مستوى المعنوية α.

ويشير مستوى **المعنوية الاحصائية** α(*) (*Statistical Significant*) (ويعرف أيضا باسم مستوى الدلالة) إلى احتمال رفض فرضية العدم (H_0)، بينما تكون في واقع الأمر صحيحة. ويعرف هذا النوع من الخطأ ب **خطأ من النوع الأول** I (*Type I Error*). ويتم تحديد قيمة α من قبل الباحث أو صاحب القرار، حيث أنه كلما رغب في الحصول على قرار أكثر دقة كلما عمد إلى اختيار قيمة ضئيلة ل α.

وهناك نوع آخر من الخطأ يمكن أن يقع فيه متخذ القرار، ويعرف ب **الخطأ من النوع الثاني** II (*Error II Type*)، ويحدث هذا النوع من الخطأ في حالة قبول فرضية العدم (H_0)، بينما تكون في واقع الأمر خاطئة. ويرمز لهذا النوع من الخطأ ب (β). وفيما يلي جدول يلخص علاقة القرار الاحصائي بطبيعة الفرضية:

الجدول رقم (IV. 1): أنواع الأخطاء الاحصائية في اختبارات الفروض.

		نوع الفرضية	
		H_0 صحيحة	H_a صحيحة
القرار	قبول H_0	قرار صحيح	خطأ من النوع II
	رفض H_0	خطأ من النوع I	قرار صحيح

مثال[1]:

دعنا نفترض أننا واجهنا حالتي الطبيعة التاليتين:

القرار	**حالات الطبيعة**	
	جو ماطر	**غير ماطر**
عدم حمل مظلة واقية من المطر.	خطأ (رفض الصحيح) $\alpha=p$	صحيح
حمل مظلة واقية من المطر.	صحيح	خطأ (قبول الخطأ) β *Error* $= p$

(*)– عادة تكون قيم α صغيرة (1%، 5%، 10%...).

[1]– عبد الرزاق بني هاني، مرجع سبق ذكره، ص 379.

في الحالة الأولى، قمنا برفض الصحيح (عدم حمل مظلة وقت المطر) فارتكبنا خطأ من النوع $I(\alpha)$.

في الحالة الثانية، قمنا بقبول الخطأ (حمل مظلة في حال الصحو) فارتكبنا خطأ من النوع الثاني II (β).

IV/ 1. 5. مراحل اختبار الفرضيات واتخاذ القرار.

للتأكد من صحة الادعاءات أو الفرضيات، يتم اختبار مدى صدقيتها من المراحل التالية: ويتم اتخاذ للقرار في الاختبار الإحصائي باستخدام إحصائية الاختبار (*Statistic test*)

- صياغة الفرضيات (H_0، وH_a) حول معلمة (أو خاصية) في مجتمع الدراسة.
- تحديد توزيع المعاينة: إما توزيع طبيعي معياري، أو توزيع t بدرجات حرية (n-1) ...إلخ.
- نحسب إحصائية الاختبار (*Statistic test*): وهي الكمية المحسوبة من مقاييس العينة لتساعدنا في التوصل إلى القرار.
- تحديد المناطق الحرجة: ويكون ذلك بعد استخراج القيم الحرجة من جداول التوزيع، والتي تحدد مناطق الرفض والقبول، ويتم ذلك حسب مستوى المعنوية α.
- نتخذ القرار برفض أو عدم رفض H_0: إذا كانت إحصائية الاختبار تأخذ قيمة على النقيض من فرضية العدم، نرفض هذه الفرضية ونستنتج أن الفرضية البديلة H_a صحيحة.

IV/ 2. تطبيقات اختبارات الفروض الاحصائية.

IV/ 2. 1. اختبار الفروض حول الوسط الحسابي

لاختبار فرضيات حول متوسط المجتمع μ يتم استخدام متوسط العينة $\bar{X}$ لصياغة احصائية الاختبار، وعموما عند اختبار فرضيات حول المتوسط نواجه حالتين:

أ. حالة التباين معلوم (استخدام التوزيع الطبيعي)

في حالة مجتمع يتبع التوزيع الطبيعي ومعلوم التباين، أو في حالة تباين المجتمع مجهول وحجم العينات كبير ($30 \leq n$)، فإن توزيع احصائية الاختبار Z يقترب من التوزيع الطبيعي المعياري.

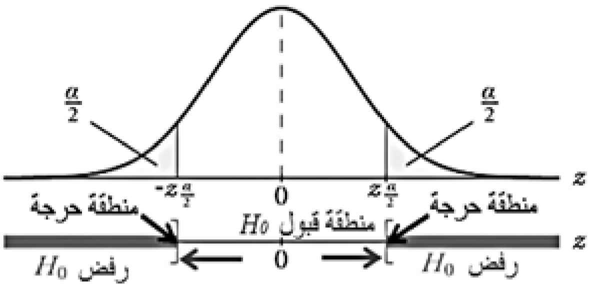
$$Z = \frac{\bar{X} - \mu_0}{\sigma_X} \sim N(0,1)$$

ويتم تحديد القيم الحرجة على أساس α، وتحديد مناطق الرفض والقبول كما يلي:

إذا كان الاختبار من طرفين: نقبل فرض العدم إذا تحققت المعادلة التالية: $-Z_{\alpha/2} < Z_{Stat} < Z_{\alpha/2}$

نرفض فرض العدم إذا تحققت إحدى

المعادلتين: $\left.\begin{matrix} Z_{Stat} > Z\alpha/_{2} \\ Z_{Stat} < -Z\alpha/_{2} \end{matrix}\right\}$



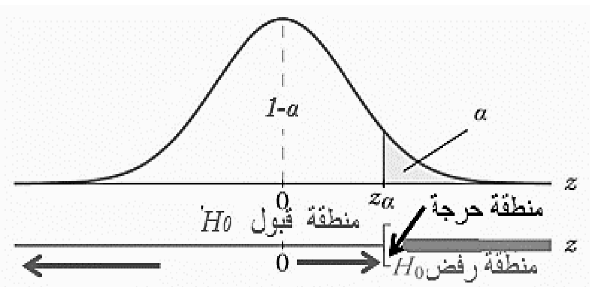
إذا كان الاختبار من طرف واحد الجهة اليمنى:

نقبل فرض العدم إذا تحققت المعادلة:

$Z_{Stat} < Z_{\alpha}$

نرفض فرض العدم إذا تحققت المعادلة:

$Z_{Stat} > Z_{\alpha}$



إذا كان الاختبار من طرف واحد الجهة

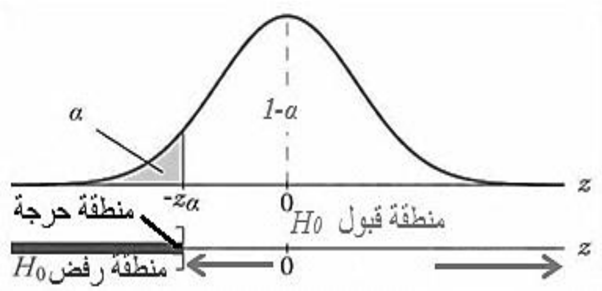
اليسرى :

نقبل فرض العدم إذا تحققت المعادلة:

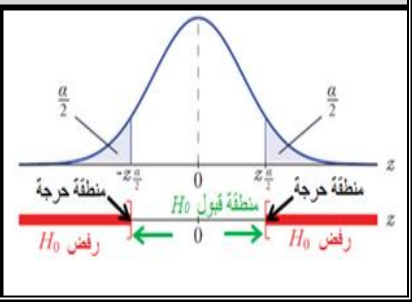
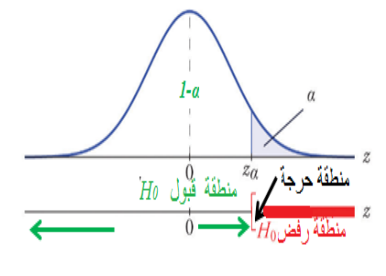
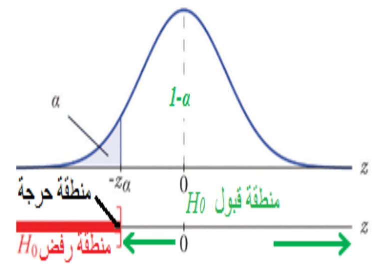
$Z_{Stat} > -Z_{\alpha}$

نرفض فرض العدم إذا تحققت المعادلة:

$Z_{Stat} < -Z_{\alpha}$



ويمكن تلخيص الحالات الثلاث السابقة لأشكال اختبار الفروض في الجدول التالي:

اختبار من طرفين	اختبار من طرف واحد	
	من الجهة اليمنى	من الجهة اليسرى
صيـــــاغة الفرضــــــية		
$\begin{cases}H_0: \mu = \mu_0 \\ H_a: \mu \neq \mu_0\end{cases}$	$\begin{cases}H_0: \mu = \mu_0 \\ H_a: \mu > \mu_0\end{cases}$	$\begin{cases}H_0: \mu = \mu_0 \\ H_a: \mu < \mu_0\end{cases}$
الــــــــقرار		
نرفض H_0 عند مستو معنوية α اذا كانت $Z \geq Z_{\alpha/2}$ أو $Z \leq -Z_{\alpha/2}$	نرفض H_0 عند مستو معنوية α اذا كانت $Z \geq Z_\alpha$	نرفض H_0 عند مستومعنوية α اذا كانت $Z \leq -Z_\alpha$
التمـــــثيل البيـــــاني		
		

مثال:

في حضانة للأطفال تدعى المشرفة أن متوسط وزن الأطفال الذين تتراوح أعمارهم بين (3.5 و4.5) سنوات في الحضانة هو 16 كغ بانحراف معياري نصف كغ. لاختبار صحة ادعاء المشرفة أخذت عينة عشوائية من 36 طفلا وتم اختبارها فوجد أن متوسط أوزان الأطفال خلال هذه المرحلة العمرية في العينة 15.8كغ. على فرض أن مجتمع أوزان الأطفال يتبع التوزيع الطبيعي، هل يمكن تأييد ادعاء المشرفة عند مستوى معنوية 1%.

الحل:

$$\alpha = 0.01 n = 36; \qquad \mu_0 = 16; \qquad \bar{X} = 15.8; \qquad \sigma = 0.5$$

المرحلة الأولى: صياغة فرضيات الدراسة.

فرض العدم: $H_0: \mu = 16$

الفرض البديل: $H_a: \mu \neq 16$

المرحلة الثانية: تحديد طبيعة توزيع المعاينة.

بما أن العينة مسحوبة من مجتمع طبيعي تباينه معلوم، إذا توزيع المعاينة المستخدم هو توزيع Z.

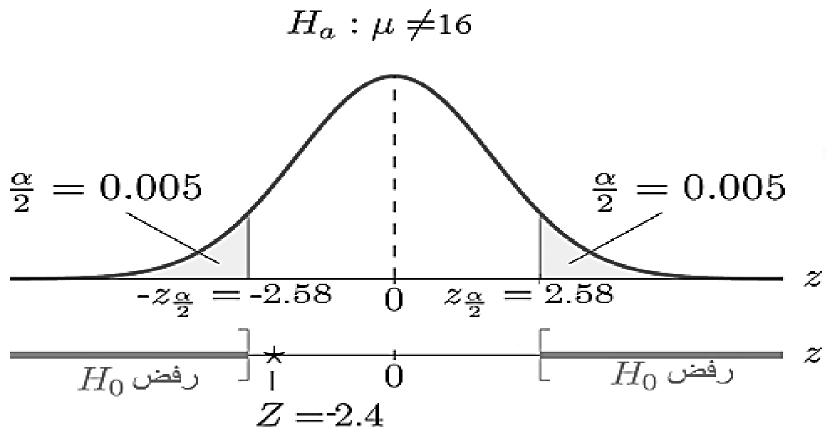
المرحلة الثالثة: حساب احصائية الاختبار.

$$Z_{Stat} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{15.8 - 16}{0.5/\sqrt{36}} = -2.4$$

المرحلة الرابعة: تحديد المناطق الحرجة.

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

ويمكن تمثيل منطقة الرفض، ومنطقة القبول بيانيا كما يلي:



المرحلة الخامسة: اتخاذ القرار.

بما أن القيمة المحسوبة وقعت في منطقة القبول أي أن $-Z_{\alpha/2} < Z_{Stat} < Z_{\alpha/2}$ ، فإن القرار هو :عدم رفض فرض العدم، أي أن ادعاء المشرفة صحيح ولا يوجد هناك فرق معنوي بين المتوسط الحقيقي والمتوسط المدعى عند مستوى معنوية 1%.

ب. اختبار الفروض حول الوسط الحسابي في حالة التباين مجهول.

عندما يكون تباين المجتمع σ^2 **مجهولا و**حجم العينات صغير **(n<30)** نستبدل الاحصائية Z بالإحصائية t.

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

ويتم تحديد القيم الحرجة على أساس α، وتحديد مناطق الرفض والقبول كما يلي:

1) إذا كان الاختبار من طرفين: نقبل فرض العدم إذا تحققت المعادلة التالية: $-t_{n-1;\alpha/2} < T < t_{n-1;\alpha/2}$

نرفض فرض العدم إذا تحققت إحدى المعادلتين: $\left.\begin{matrix} T > t_{n-1} \\ T < -t_{n-1} \end{matrix}\right\}$

2) إذا كان الاختبار من طرف واحد الجهة اليمنى: نقبل فرض العدم إذا تحققت المعادلة: $T < t_{n-1}$

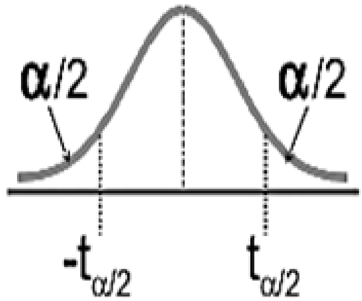
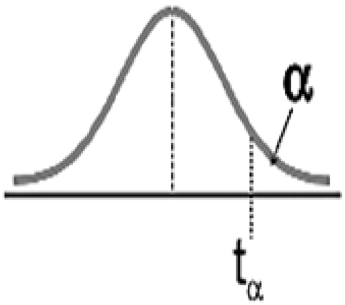
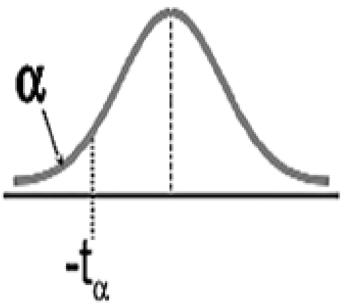
نرفض فرض العدم إذا تحققت المعادلة: $T > t_{n-1}$

3) إذا كان الاختبار من طرف واحد الجهة اليسرى: نقبل فرض العدم إذا تحققت المعادلة:

$$T > -t_{n-1}$$

نرفض فرض العدم إذا تحققت المعادلة: $T < -t_{n-1}$

والجدول التالي يلخص الحالات الثلاث السابقة لأشكال اختبار الفروض حول المتوسط (في حالة σ^2 مجهور و $n <$ 30) الجدول التالي:

اختبار من طرفين	اختبار من طرف واحد	
	من الجهة اليمنى	من الجهة اليسرى
صيـــــاغة الفرضـــــية		
$\begin{cases} H_0: \mu = \mu_0 \\ H_a: \mu \neq \mu_0 \end{cases}$	$\begin{cases} H_0: \mu = \mu_0 \\ H_a: \mu > \mu_0 \end{cases}$	$\begin{cases} H_0: \mu = \mu_0 \\ H_a: \mu < \mu_0 \end{cases}$
القــــــرار		
نرفض H_0 عند مستو معنوية α اذا كانت $T > t_{n-1}$ أو $T < -t_{n-1}$	نرفض H_0 عند مستو معنوية α اذا كانت $T > t_{n-1}$	نرفض H_0 عند مستو معنوية α اذا كانت $T < -t_{n-1}$
التمـــــثيل البيـــــاني		
		

مثال:

ادعى أحد المسؤولين في بنك خاص أن معدل فترة الانتظار في كل شباك لا تزيد عن 15 دقيقة، وبهدف اختبار هذا الادعاء تم دراسة عينة مكونة من 10 شبابيك لخدمة الزبائن على مستوى هذه المؤسسة البنكية فوجد أن معدل الانتظار 16.5 دقيقة بانحراف معياري 4 دقائق، على فرض أن مدة الانتظار لكل الشبابيك في البنك تتبع التوزيع الطبيعي اختبر ادعاء البنك عند مستوى معنوية 0.05.

الحل:

$$\mu_0 = 15, \qquad n = 10, \qquad \bar{X} = 16.5, \qquad S = 4, \qquad \alpha = 5\%.$$

المرحلة الأولى: صياغة فرضيات الدراسة.

فرض العدم: $H_0: \mu = 15$

الفرض البديل: $H_a: \mu > 15$

المرحلة الثانية: تحديد طبيعة توزيع المعاينة.

بما أن العينة مسحوبة من مجتمع طبيعي تباينه مجهول، وحجم العينة صغير ($n=10<30$)⇐ إذا توزيع المعاينة المستخدم هو توزيع t.

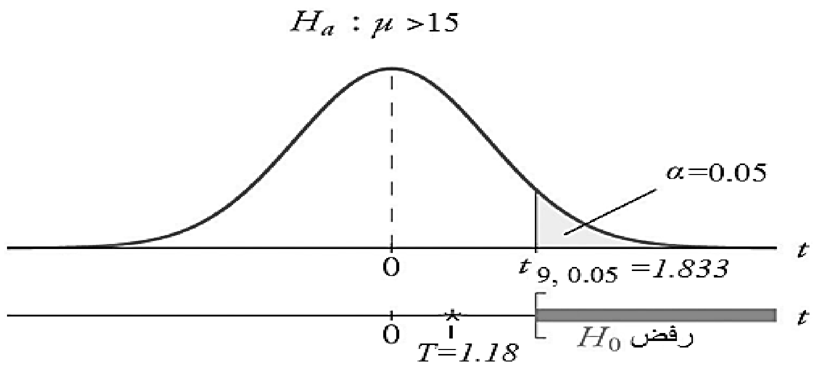
المرحلة الثالثة: حساب احصائية الاختبار.

$$T_{Stat} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{16.5 - 15}{4/\sqrt{10}} = 1.18$$

المرحلة الرابعة: تحديد المناطق الحرجة.

$$\alpha = 0.05 \Rightarrow t_{n-1;\,\alpha} = t_{9;0.05} = 1.833$$

ويمكن تمثيل منطقة الرفض، ومنطقة القبول بيانيا كما يلي:



المرحلة الخامسة: اتخاذ القرار.

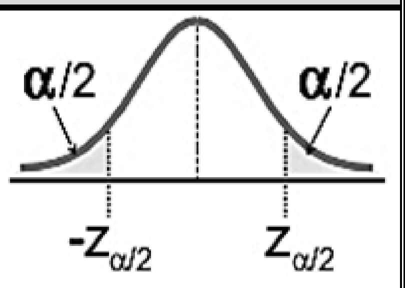
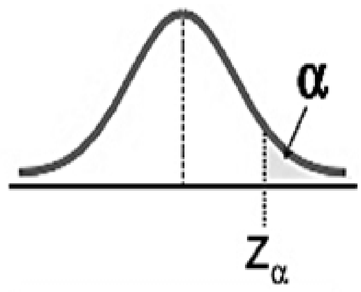
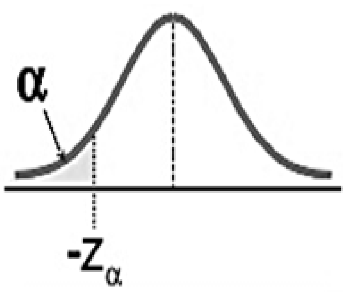
بما أن القيمة المحسوبة وقعت في منطقة القبول أي أن $T_{Stat} < t_{v;\,\alpha}$، فإن القرار هو: عدم رفض فرض العدم، أي أن ادعاء البنك أن متوسط فترة انتظار العملاء على الشبابيك لا يتعدى 15 دقيقة عند مستوى معنوية 5%.

IV/ 2. 2. اختبار الفروض حول النسبة.

يتعلق هذا الاختبار بنسبة مفردات المجتمع التي تتصف بخاصية ما (P)، حيث اذا كان المجتمع الذي نقوم بدراسته يتبع توزيعا متقطعا مثل التوزيع ذي الحدين[1]، وعلى فرض أن P نسبة صفة ما في المجتمع، فإنه سيتم استخدام $\hat{P}$ نسبة نفس الصفة في العينة لبناء فرضيات حول النسبة P، فإذا كان حجم العينة كبيرا بالشكل الكافي $(30\leq n)$ وحسب ما تطرقنا إليه في فصل توزيعات المعاينة، فإن احصائية الاختبار $\frac{\hat{P}-P}{\sqrt{\frac{P(1-P)}{n}}}$ تقترب من التوزيع الطبيعي المعياري،

$$Z = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} \sim N(0,1)$$

والجدول التالي يلخص مختلف أشكال الفروض حول النسبة:

اختبار من طرفين	اختبار من طرف واحد	
	من الجهة اليمنى	من الجهة اليسرى
صيـــــاغة الفرضـــــية		
$\begin{cases} H_0: P = P_0 \\ H_a: P \neq P_0 \end{cases}$	$\begin{cases} H_0: P = P_0 \\ H_a: P > P_0 \end{cases}$	$\begin{cases} H_0: P = P_0 \\ H_a: P < P_0 \end{cases}$
الـــــــقرار		
نرفض H_0 عند مستوى معنوية α اذا كانت $Z\geq Z_{\alpha/2}$ أو $Z\leq -Z_{\alpha/2}$	نرفض H_0 عند مستوى معنوية α اذا كانت $Z\geq Z_{\alpha}$	نرفض H_0 عند مستوى معنوية α اذا كانت $Z\leq -Z_{\alpha}$
التمــــثيل البيــــاني		
		

مثال: من بين 200 شخص وجد أن 160 منهم يستخدمون نوع محدد من منتجات التنظيف. اختبر الفرض بأن نسبة مستخدمي هذا المنتج لا يقل عن 85%، وذلك عند مستوى معنوية 5%.

---

[1] – عدنان بن ماجد البري، محمود محمد هندي، أنور أحمد عبد الله، **"مبادئ الاحصاء والاحتمالات"**، الطبعة الثالثة، النشر والمطابع-جامعة الملك سعود، الرياض، السعودية، 1997، ص 362.

الحل:

$$n = 200; \qquad n_a = 160 \Rightarrow \hat{P} = \frac{n_a}{n} = \frac{160}{200} = 0.80 \qquad P_0 = 0.85$$

المرحلة الأولى: صياغة فرضيات الدراسة.

الفرض العدم: $H_0: P = 0.85$

الفرض البديل: $H_a: P < 0.85$

المرحلة الثانية: تحديد طبيعة توزيع المعاينة.

حجم العينة كبير بما فيه الكفاية ($30<100=n$)، إذا توزيع المعاينة المستخدم هو توزيع *Z*.

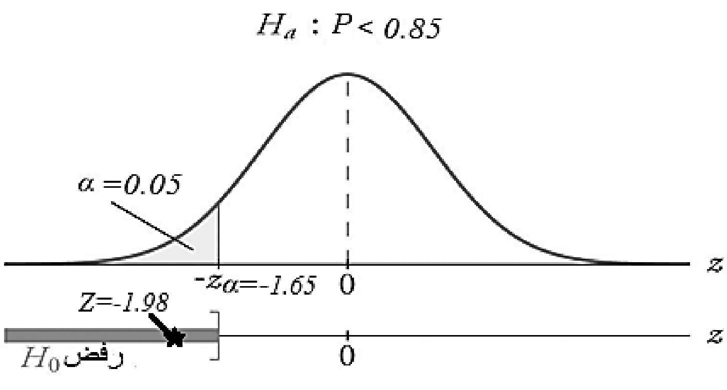
المرحلة الثالثة: حساب احصائية الاختبار.

$$Z_{Stat} = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0(1 - P_0)}{n}}} = \frac{0.80 - 0.85}{\sqrt{\frac{0.85(0.15)}{200}}} = -1.98$$

المرحلة الرابعة: تحديد المناطق الحرجة.

$$\alpha = 0.05 \Rightarrow Z_{1-\alpha} = Z_{0.95} = -1.65$$

ويمكن تمثيل منطقة الرفض، ومنطقة القبول بيانيا كما يلي:



المرحلة الخامسة: اتخاذ القرار.

بما أن القيمة المحسوبة وقعت في منطقة الرفض حيث أن $Z_{Stat}<-Z_\alpha$، فإن القرار هو :رفض فرض العدم، أي أن ادعاء مستخدمي هذا المنتج يفوق 85% غير صحيح عند مستو معنوية 5%.

IV/ 2. 3. اختبار الفروض للفرق بين متوسطين

عند اختبار الفرضيات الاحصائية حول الفرق بين متوسطي مجتمعين مستقلين نواجه الحالتين التاليتين:

أ. <u>الحالة1: حالة التباين معلوم (باستخدام التوزيع الطبيعي)</u>

اذا كان X وY عينتين مسحوبتين من مجتمعين طبيعيين، معلومي التباين، فإن احصائية الاختبار Z تتبع التوزيع الطبيعي،

$$Z=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}}\sim N(0,1)$$

إلا أنه عادة ما تكون σ_X و σ_Y مجهولتين، ولذلك يتم تعويضهما ب S_X و S_Y (الانحراف المعياري للعينتين) كتقدير ل σ_X و σ_Y، وعندما يكون حجم كلا العينتين كبير بالشكل الكافي $(30\leq n\ ;\ m)$ فإن توزيع احصائية الاختبار Z يقترب من التوزيع الطبيعي المعياري ونكتب:

$$Z=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}}\sim N(0,1)$$

اذا لم يكن هناك فرق بين متوسطي المجتمعين فإن الفرق بينهما يكون صفرا، أما إذا كان متوسطا المجتمعين مختلفين فإن الفرق بينهما لا يساوي الصفر وعليه يمكن وعليه يمكن تلخيص مختلف أشكال اختبار الفرضيات حول الفرق بين متوسطي مجتمعين مستقلين (في حالة σ_X و σ_Y معلومين، أو في حالة العينات كبيرة الحجم $(30\leq n\ ;\ m)$ و σ_X و σ_Y مجهولين) كما يلي:

اختبار من طرفين	اختبار من طرف واحد: من الجهة اليمنى	اختبار من طرف واحد: من الجهة اليسرى
صيـــــاغة الفرضـــــية		
$\begin{cases}H_0: \mu_1-\mu_2=0\\ H_a: \mu_1-\mu_2\neq 0\end{cases}$	$\begin{cases}H_0: \mu_1-\mu_2=0\\ H_a: \mu_1-\mu_2>0\end{cases}$	$\begin{cases}H_0: \mu_1-\mu_2=0\\ H_a: \mu_1-\mu_2<0\end{cases}$
الـــــقرار		
نرفض H_0 عند مستو معنوية α اذا كانت $Z\geq Z_{\alpha/2}$ أو $Z\leq -Z_{\alpha/2}$	نرفض H_0 عند مستو معنوية α اذا كانت $Z\geq Z_\alpha$	نرفض H_0 عند مستو معنوية α اذا كانت $Z\leq -Z_\alpha$

مثال[1]:

ترغب مديرة أن تحدد عند مستو معنوية 5% ما اذا كان الأجر بالساعة للعمال نصف المهرة متساويا في مدينتين. لعمل ذلك، فإنها تأخذ عينتين عشوائيتين من الأجر بالساعة من كل مدينة فوجدت أن $\overline{X_1}=6.00\$$، $\overline{X_2}=5.40\$$، $S_1= 2.00$ و $S_2=1.80$، وذلك لعينتين من حجم $n_1=40$، و $n_2=54$.

فإن خطوات اختبار الفرق بين متوسطي الأجر للعمال بين المدينتين كما يلي:

المرحلة الأولى: صياغة فرضيات الدراسة:

$$H_0: \mu_1 = \mu_2 \quad or \quad H_0: \mu_1 - \mu_2 = 0$$
$$H_a: \mu_1 \neq \mu_2 \quad or \quad H_a: \mu_1 - \mu_2 \neq 0$$

المرحلة الثانية: تحديد طبيعة توزيع المعاينة:

العينتين مسحوبتين من مجتمعين مجهولي التباين، وحجم العينات كبير($n_1=40$، و $30<n_2=54$)، وعليه فإن احصائية الاختبار Z تتبع التوزيع الطبيعي.

المرحلة الثالثة: حساب احصائية الاختبار:

بما أن تبايني المجتمعين مجهولين سيتم تعويضهما بتبايني العينتين S_1 و S_2.

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} = \sqrt{\frac{(2)^2}{40}+\frac{(1.8)^2}{54}} = \sqrt{0.1+0.06} = \sqrt{0.16} = \boxed{0.4}$$

$$Z = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{(\bar{X}_1-\bar{X}_2)-0}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{0.6}{0.4} = \boxed{1.5}$$

المرحلة الرابعة: تحديد القيم الحرجة ومناطق الرفض والقبول:

الاختبار ذو ذيلين (ذو طرفين)، وبما أن $\alpha=0.05$، فإن منطقة القبول ل H_0 تقع في حدود ±1.96 تحت المنحنى الطبيعي القياسي.

[1]ــ دومينيك سالفاتور، مرجع سبق ذكره، ص 101.

المرحلة الخامسة: اتخاذ القرار.

وحيث أن قيمة Z المحسوبة تقع داخل منطقة القبول، فإننا نقبل H_0، أي أن $\mu_1=\mu_2$ عند مستو معنوية %5

ب. <u>الحالة 2: حالة تبايني المجتمعين مجهولين وحجم العينات صغير $(30>n\ ;\ m)$</u>، في هذه الحالة فإن احصائية الاختبار تتبع التوزيع t بدرجة حرية $(n+m-2)$

حيث أنه في حالة ما اذا كان σ_1 و σ_2 مجهولين، وحسب ما تم توضيحه سابقا فإن احصائية الاختبار تكتب كما يلي:

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}}\sim t_{n+m-2}$$

ملاحظة: في حالة ما اذا كان تباني المجتمعين مجهولين ومتساويين $(\sigma^2=\sigma^2_1=\sigma^2_2)$ فإننا:

$S=\sqrt{\frac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2}}$ ، وعليه يمكن كتابة احصائية اختبار الفرضيات حول الفرق بين متوسطين في هذه الحالة كما يلي:

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{S\sqrt{\frac{1}{n}+\frac{1}{m}}}\sim t_{n+m-2}$$

ويمكن تلخيص مختلف أشكال اختبار الفرضيات حول الفرق بين متوسطي مجتمعين مستقلين في هذه الحالة في الجدول التالي:

اختبار من طرفين	اختبار من طرف واحد	
	من الجهة اليمنى	من الجهة اليسرى
صيـــــاغة الفرضـــــية		
$\begin{cases}H_0:\mu_1-\mu_2=0\\H_a:\mu_1-\mu_2\neq 0\end{cases}$	$\begin{cases}H_0:\mu_1-\mu_2=0\\H_a:\mu_1-\mu_2>0\end{cases}$	$\begin{cases}H_0:\mu_1-\mu_2=0\\H_a:\mu_1-\mu_2<0\end{cases}$
الـــــــقرار		
نرفض H_0 عند مستو معنوية α اذا كانت $T>t_{n-1}$ أو $T<-t_{n-1}$	نرفض H_0 عند مستو معنوية α اذا كانت $T>t_{n-1}$	نرفض H_0 عند مستو معنوية α اذا كانت $T<-t_{n-1}$

مثال: على فرض أن حجم n_1، و n_2 في المثال السابق هو 17 و 25 على التوالي، $\sigma=\sigma_1=\sigma_2$ ما قرارك؟

الحل:

في هذه الحالة بما أن تبايني المجتمعين مجهولين، وn_1 و $30>n_2$ ، فإن توزيع المعاينة للفرق بين المتوسطين يتبع توزيع t.

وبما أن $\sigma_1=\sigma_2=\sigma$ فإن احصاءه الاختبار تكتب كما يلي:

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

حيث أن الانحراف المعياري المشترك σ يساوي:

$$=\sigma=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}=\sqrt{\frac{(17-1)2^2+(25-1)1.8^2}{17+25-2}}=\sqrt{3.544}=\boxed{1.88}$$

وعليه فإن:

$$T_{Stat}=\frac{(\bar{X}_1-\bar{X}_2)-0}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{6-5.4}{1.88\sqrt{\frac{1}{17}+\frac{1}{25}}}=\frac{0.6}{0.59}=\boxed{1.02}$$

$$t_{n_1+n_2-2;\frac{\alpha}{2}}=t_{38;0.025}=\pm 2.021$$

اتخاذ القرار: يلاحظ في هذه الحالة أيضا أن قيمة t المحسوبة تقع داخل منطقة القبول، وعليه فإننا نقبل H_0، أي أن $\mu_1=\mu_2$ عند مستو معنوية 5%.

IV/ 2. 4. اختبار الفروض للفرق بين نسبتين

إذا كانت X و Y عينتين عشوائيتين مسحوبتين من مجتمعين طبيعيين، حيث تمثل النسبة $\hat{P}_X$ نسبة صفة معينة في العينة X، و$\hat{P}_Y$ نسبة نفس الصفة في العينة Y، **اذا** كان حجم العينتين كبير بما فيه الكفاية ($30\leq n\ ;\ m$)، وحسب الفصل السابق فإن احصائية اختبار الفرق بين النسبتين P_X نسبة صفة معينة

في المجتمع الأول و P_Y نسبة نفس الصفة في المجتمع الثاني تقترب من التوزيع الطبيعي المعياري ونكتب:

$$Z = \frac{(\hat{P}_X - \hat{P}_Y) - (P_X - P_Y)}{\sqrt{\frac{P_X(1-P_X)}{n} + \frac{P_Y(1-P_Y)}{m}}} \sim N(0,1)$$

ملاحظة: تجدر الاشارة إلى أنه في حالة ما اذا كانت $P=P_X=P_Y$، فإن $\frac{n_a}{n}$ و $\frac{m_a}{m}$ سيكون كلاهما تقدير بقيمة واحدة لنفس المعلمة، وبالتالي يمكننا تجميع هذه التقديرات من خلال استخدام مجموع تحقق الحالات التي تتوفر فيها صفة معينة وقسمته على اجمالي عدد المحاولات، أي أنه بدل استخدام $\frac{n_a}{n}$ و $\frac{m_a}{m}$ نستخدم القيمة التالية: $\frac{n_a+m_a}{n+m}$، وتشير هذه الأخيرة إلى ما يعرف ب **"متوسط مرجح للنسبتين"** ونرمز له بالرمز $\acute{P}$، وعليه إذا كان $P=P_X=P_Y$ فإن:

$$\sigma_{\hat{P}_X-\hat{P}_Y} \sqrt{\frac{\acute{P}(1-\acute{P})}{n} + \frac{\acute{P}(1-\acute{P})}{m}}$$

حيث أن التقدير المشترك للنسبتين $\acute{P}$ يساوي:

$$\boxed{\acute{P} = \frac{n\hat{P}_X + m\hat{P}_Y}{n+m}} \quad \text{أو} \quad \boxed{\acute{P} = \frac{na+ma}{n+m}}$$

ويمكن تلخيص مختلف أشكال اختبار الفرضيات حول الفرق بين نسبتين في الجدول التالي:

اختبار من طرفين	اختبار من طرف واحد: من الجهة اليمنى	اختبار من طرف واحد: من الجهة اليسرى
صيـــــاغة الفرضـــــية		
$\begin{cases} H_0: P_X = P_Y \\ H_a: P_X \neq P_Y \end{cases}$	$\begin{cases} H_0: P_X = P_Y \\ H_a: P_X > P_Y \end{cases}$	$\begin{cases} H_0: P_X = P_Y \\ H_a: P_X < P_Y \end{cases}$
الـــــــقرار		
نرفض H_0 عند مستو معنوية α اذا كانت $Z \geq Z_{\alpha/2}$ أو $Z \leq -Z_{\alpha/2}$	نرفض H_0 عند مستو معنوية α اذا كانت $Z \geq Z_\alpha$	نرفض H_0 عند مستو معنوية α اذا كانت $Z \leq -Z_\alpha$

مثال: بهدف معرفة انتشار نسبة البدانة بين المراهقين والمراهقات تم اختيار عينتين عشوائيتين، العينة الأولى من مجتمع المراهقين ووجد أنه من بين 100 مراهق هناك 21 مراهق يعانون من البدانة، ومن بين 100 مراهقة هناك34 مراهقة تعاني البدانة، عند مستوى معنوية 1% اختبر الفرض القائل بأن نسبة البدانة في مجتمع المراهقين أقل من نسبة البدانة في مجتمع المراهقات.

الحل:

$$n = m = 100, \qquad \hat{P}_X = \frac{21}{100} = \boxed{0.21} \qquad \hat{P}_Y = \frac{34}{100} = \boxed{0.34}$$

صياغة فرضيات الدراسة.

الفرض العدم: $H_0: P_X = P_Y$

الفرض البديل: $H_a: P_X > P_Y$

حساب احصائية الاختبار.

$$\acute{P} = \frac{n\hat{P}_X + m\hat{P}_Y}{n+m} = \frac{100(0.21) + 100(0.34)}{100 + 100} = \boxed{0.275}$$

$$\sigma_{\hat{P}_X - \hat{P}_Y} = \sqrt{\frac{0.275(0.725)}{100} + \frac{0.275(0.725)}{100}} = \sqrt{0.004} = \boxed{0.06}$$

وعليه فإن احصاءه الاختبار:

$$Z_{Stat} = \frac{\left(\hat{P}_X - \hat{P}_Y\right) - 0}{\sigma_{\hat{P}_X - \hat{P}_Y}} = \frac{0.21 - 0.34}{0.06} = \frac{-0.13}{0.06} = \boxed{-2.16}$$

منطقة الرفض:

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

القرار:

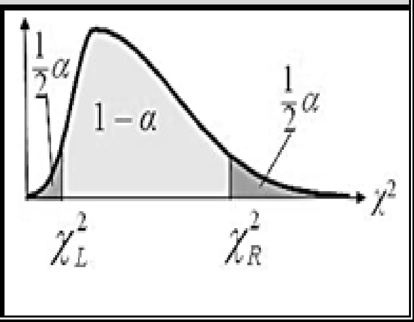
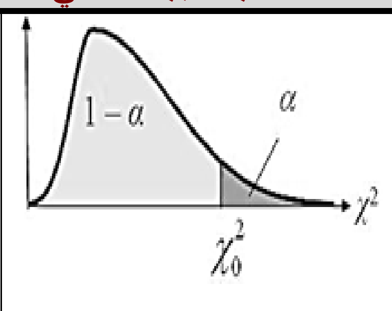
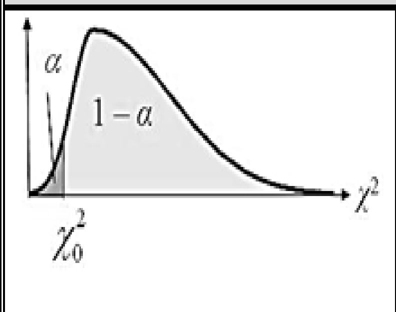
وبما أن القيمة المحسوبة وقعت في منطقة القبول حيث أن $Z_{Stat}<Z_{\alpha}$، فإن القرار هو :قبول فرض العدم، أي عدم وجود فروق معنوية بين نسبة البدانة في مجتمع المراهقين ونسبة البدانة في مجتمع المراهقات، أي رفض الفرض البديل عند مستو معنوية 1%.

IV/ 2. 5. اختبار الفروض حول التباين.

يتعلق هذا الاختبار بتباين المجتمع σ^2، فيما اذا كان يساوي قيمة معينة، ونعلم انطلاقا من فصل توزيعات المعاينة أن الاحصائية $\frac{(n-1)S^2}{\sigma^2}$ تتبع توزيع كاي مربع χ^2، ويمكن أن نكتب احصاءه الاختبار كما يلي:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

ويمكن تلخيص مختلف أشكال اختبار الفروض حول التباين في الجدول التالي:

اختبار من طرفين	اختبار من طرف واحد	
	من الجهة اليمنى	من الجهة اليسرى
صيـــــاغة الفرضـــــية		
$H_0: \sigma^2 = {\sigma^2}_0$ $H_a: \sigma^2 \neq {\sigma^2}_0$	$H_0: \sigma^2 = {\sigma^2}_0$ $H_a: \sigma^2 > {\sigma^2}_0$	$H_0: \sigma^2 = {\sigma^2}_0$ $H_a: \sigma^2 < {\sigma^2}_0$
الـــــــقرار		
نرفض H_0 عند مستوى معنوية α اذا كانت $\chi^2 > \chi^2_{\frac{\alpha}{2};v}$ أو $\chi^2 < \chi^2_{(1-\frac{\alpha}{2});v}$	نرفض H_0 عند مستوى معنوية α اذا كانت $\chi^2 > \chi^2_{\alpha;v}$	نرفض H_0 عند مستوعمعنوية α اذا كانت $\chi^2 < \chi^2_{(1-\alpha);v}$
التمـــــثيل البيـــــاني		
		

مثال:

أدعى أحد المديرين بأن تباين المرتبات الشهرية للموظفين في مؤسسته لا يتجاوز 1440 دج. وللتأكد من ادعائه تم سحب عينة عشوائية من مرتبات 10 موظفين، فوجد بأن تباينها يساوي 1670، على فرض أن توزيع مجتمع المرتبات الشهرية للموظفين في هذه المؤسسة يتبع التوزيع الطبيعي، اختبر صدق ادعاء المدير عند مستوى معنوية 10%.

الحل:

$$n = 10; \qquad S^2 = 1670, \qquad \sigma^2 = 1440 \qquad \alpha = 10\%$$

المرحلة الأولى: صياغة فرضيات الدراسة.

فرض العدم: $H_0: \sigma^2 = 1440$

الفرض البديل: $H_a: \sigma^2 > 1440$

المرحلة الثانية: تحديد طبيعة توزيع المعاينة.

عينة عشوائية مسحوبة من مجتمع موزع طبيعيا، ولاختبار فرضية حول التباين فان احصاءة الاختبار تتبع توزيع كاي مربع χ2. $\chi^2_{Stat} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

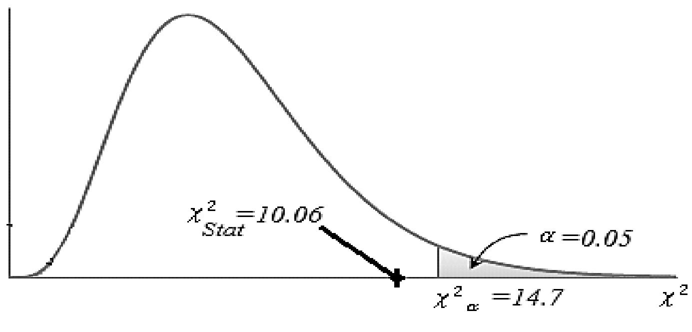
المرحلة الثالثة: حساب احصائية الاختبار.

$$\chi^2_{Stat} = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(10-1)1610}{1440} = 10.06$$

المرحلة الرابعة: تحديد المناطق الحرجة.

$$\alpha = 0.1 \Rightarrow \chi^2_{v,\alpha} = \chi^2_{9,0.1} = 14.7$$

ويمكن تمثيل منطقة الرفض، ومنطقة القبول بيانيا كما يلي:

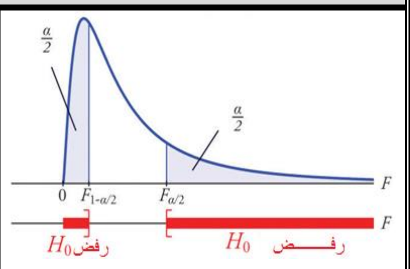
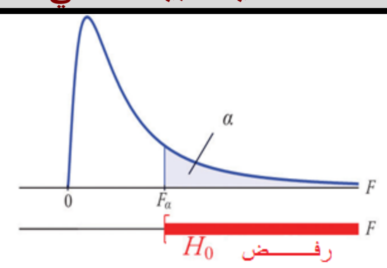
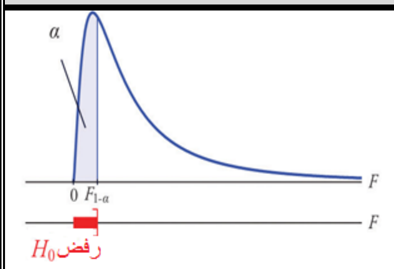


المرحلة الخامسة: اتخاذ القرار.

بما أن القيمة المحسوبة وقعت في منطقة القبول حيث أن $\chi^2_{Stat} < \chi^2_{\alpha;n-1}$، فإن القرار هو :عدم رفض فرض العدم، أي أن ادعاء المدير حول تباين مرتبات الموظفين صحيح عند مستو معنوية 5%.

IV/ 2. 6. اختبار تساوي تباينين.

يتعلق هذا الاختبار بتساوي تبايني مجتمعين مستقلين σ^2_X و σ^2_Y، ولتكن X و Y عينتين عشوائيتين مسحوبتين من مجتمعين مستقلين موزعان توزيعا طبيعيا، ولتكن S^2_X تمثل تباين العينة الأولى، و S^2_Y تمثل تباين العينة الثانية ، ونعلم انطلاقا من الفصل السابق أن توزيع المعاينة للإحصاءه $\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$ يتبع توزيع فيشر F بدرجتي حرية $n-1$ و $m-1$، وعليه يمكن استخدام مئويات توزيع F كقيم حرجة لاختبار الفرضيات أدناه والموضحة في الجدول الذي يلخص مختلف أشكال اختبار الفروض حول تساوي :

<table>
<tr><th rowspan="2">اختبار من طرفين</th><th colspan="2">اختبار من طرف واحد</th></tr>
<tr><th>من الجهة اليمنى</th><th>من الجهة اليسرى</th></tr>
<tr><td colspan="3">صيـــــاغة الفرضـــــية</td></tr>
<tr><td>$H_0: \sigma_1^2 = \sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2} = 1$
$H_a: \sigma_1^2 \neq \sigma_2^2$</td><td>$H_0: \sigma_1^2 = \sigma_2^2$
$H_a: \sigma_1^2 > \sigma_2^2$</td><td>$H_0: \sigma_1^2 = \sigma_2^2$
$H_a: \sigma_1^2 < \sigma_2^2$</td></tr>
<tr><td colspan="3">إحصاءه الاختبار</td></tr>
<tr><td>$(*) F_{Stat} = \frac{L}{M}$</td><td>$F_{Stat} = \frac{S_X^2}{S_Y^2}$</td><td>$F_{Stat} = \frac{S_X^2}{S_Y^2}$</td></tr>
<tr><td colspan="3">الـــــــقرار</td></tr>
<tr><td>نرفض H_0 عند مستو معنوية α
اذا كانت $(**)\ F_{Stat} > F_{\frac{\alpha}{2}, v_1, v_2}$</td><td>نرفض H_0 عند مستو معنوية α
اذا كانت $F_{Stat} > F_{\alpha, n-1, m-1}$</td><td>نرفض H_0 عند مستو معنوية α
اذا كانت $F_{Stat} > F_{\alpha, m-1, n-1}$</td></tr>
<tr><td colspan="3">التمــــثيل البيـــــاني</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

(*) حيث أن: L: يمثل التباين الأكبر، و M : يمثل التباين الأصغر.

(*) حيث أن: v_1: درجة الحرية المصاحبة للتباين الأكبر. v_2: درجة الحرية المصاحبة للتباين الأصغر.

ملاحظة: $F_{1-\alpha;\, v_1;\, v_2} = \frac{1}{F_{\alpha;\, v_2;\, v_1}}$

مثال:

تمت مقارنة مدة صلاحية نوعين من الأجبان الانحراف المعياري للنوع الأول يقدر ب 4 أسابيع، بينما قدر الانحراف المعياري لمدة صلاحية النوع الثاني للجبن ب 5 أسابيع. قمنا بسحب عينتين من النوعين حجمهما $n_2=n_1 = 10$. عند مستوى معنوية 10% اختبر فرض تساوي تباين مدة الصلاحية لهذين النوعين من الجبن.

الحل:

$$n_1 = n_2 = 10; \qquad S_1 = 4, \qquad S_2 = 5 \qquad \alpha = 10\%$$

المرحلة الأولى: صياغة فرضيات الدراسة.

فرض العدم: $H_0: \sigma_1^2 = \sigma_2^2$

الفرض البديل: $H_0: \sigma_1^2 \neq \sigma_2^2$

المرحلة الثانية: تحديد طبيعة توزيع المعاينة.

عينة عشوائية مسحوبة من مجتمع موزع طبيعيا، ولاختبار فرضية حول تبايني مجتمعين فان احصاءة الاختبار تتبع توزيع فيشر . $F_{Stat} = \frac{S_1^2}{S_2^2} \sim F_{v_1,v_2}$

المرحلة الثالثة: حساب احصائية الاختبار.

في هذا التمرين نلاحظ أن $S_2^2 > S_1^2$ وعليه فإن:

$$F_{Stat} = \frac{L}{M} = \frac{S_2^2}{S_1^2} = \frac{25}{16} = 1.5625$$

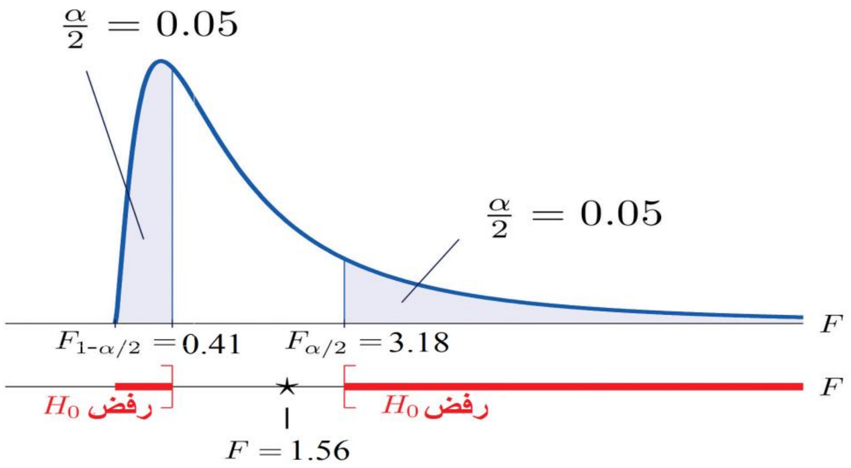
المرحلة الرابعة: تحديد المناطق الحرجة.

بما أن $\alpha=0.1$ و درجات الحرية $v_1 = v_2 = (n_1\text{-}1) = (n_2\text{-}1) = 10\text{-}1=9$، وبالرجوع إلى جدول F نجد:

$$F_{\frac{\alpha}{2},v_2,v_1} = F_{\frac{0.1}{2},9,9} = F_{0.05,9,9} = 3.18$$

$$F_{1-\frac{\alpha}{2},v_1,v_2} = \frac{1}{F_{\frac{\alpha}{2},v_1,v_2}} = \frac{1}{F_{0.95,9,9}} = \frac{1}{2.44} = 0.41$$

ويمكن تمثيل منطقة الرفض، ومنطقة القبول بيانيا كما يلي:



المرحلة الخامسة: اتخاذ القرار.

بما أن القيمة المحسوبة وقعت في منطقة القبول حيث أن $F_{Stat} < F_{\frac{\alpha}{2},v_1,v_2}$، فإن القرار هو :عدم رفض فرض العدم، أي أن ادعاء تساوي تبايني مدة الصلاحية لنوعي الجبن صحيح عند مستو معنوية 10%.

تمارين حول الفصل (السلسلة 04):

التمرين 01:

بهدف دراسة متوسط الأعمار في بلد ما تم تكوين عينة من 100 شخص متوفي، فكان متوسط العمر الافتراضي للأشخاص المتوفيين $\bar{X}$=71.8. ووفقا لدراسات سابقة يفترض أن يساوي الانحراف المعياري للمجتمع σ=8.9. انطلاقا من المعلومات السابقة اختبر الفرض القائل أن متوسط أعمار الأشخاص في هذا البلد أكبر من 70 عند مستوى معنوية 5%.

التمرين 02:

ادعى بعض الطلبة أن متوسط علامات أحد المدرسين هو μ=95. اذا علمنا أن عدد الطلبة كان N= 200، وأن عينة عشوائية من حجم n= 50 قد سحبت وكان متوسط العلامات $\bar{X}$=83 بانحراف معياري قدره S=5، فهل للفرق بين μ و$\bar{X}$ معنوية احصائية على مستوى α= 1 %.

التمرين 03:

ادعى مدير إحدى شركات انتاج العصائر أن حجم كل علبة عصير 1 لتر، من أجل التأكد من مدى مطابقة حجم علب العصير في الشركة ككل تم سحب عينة عشوائية حجمها 16 علبة عصير، ووجد أن متوسط كمية العصير للعلبة الواحدة 0.91 لتر، والانحراف المعياري لها 0.15 لتر، على فرض أن كمية العصير في هذه الشركة تتبع التوزيع الطبيعي اختبر صحة ادعاء هذا المدير عند مستوى معنوية 5.%

التمرين 04:

من بين 1200 شخص تم استجوابهم وجد أن عدد المؤيدين منهم لمرشح للانتخابات البلدية في منطقة ما هو 948. اختبر الفرض القائل أن نسبة المؤيدين لذلك المرشح في المجتمع أكبر من 80%، وذلك عند مستوى معنوية 1%.

التمرين 05:

ذا اختيرت عينة عشوائية من 60 طالب من جامعة خاصة، فوجد أن متوسط ذكائهم 69 درجة وتباين قدره 230 درجة، كذلك تم اختيار عينة عشوائية أخرى من 85 طالب من جامعة حكومية فوجد أن متوسط ذكائهم 74 درجة وتباين قدره 215. اختبر الفرض القائل بأن متوسط ذكاء طالب الجامعة الخاصة أقل من متوسط ذكاء طالب جامعة حكومية وذلك عند مستوى معنوية 5%.

التمرين 07:

ادعى مقاول أن كلفة بناء المتر المربع للدور في المدينة (أ) هي أعلى منها أو تساوي في المدينة (ب)، فسحبت عينتين من الدور من كلا المدينتين وتم حساب كلفة المربع لكل منها فوجد ما يلي:

العينة للمدينة (أ): 372. 380 . 340 . 400 . 305 . 310 .300 .316 .

العينة للمدينة (ب): 391. 306 .405 . 310 . 315 . 350 . 360 . 380 . 305 .370 .355 .318

المطلوب اختبار ادعاء المقاول عند مستوى معنوية مقداره 1 %، علما أن تبايني المجتمعين مجهولين وغير متساويين؟

التمرين 08:

ترغب شركة أن تحدد عند مستوى معنوية 1% ما اذا كانت نسبة المقبول من المكونات الالكترونية لمورد أجنبي P_1، تزيد عنها لمورد محلي P_2. وقد أخذت الشركة عينة عشوائية من شحنة حجمها 100 ووجدت أن نسبة قبول المكونات الالكترونية للمورد الأول تساوي 0.9، بينما أخذت عينة من شحنة المورد المحلي حجمها 80 وكانت نسبة القبول في العينة للمورد المحلي تساوي 0.7، اختبر الفرض القائل بأن نسبة قبول المنتج المورد الأجنبي أكبر من نسبة قبول المنتج المقدم من قبل المنتج المحلي عند مستوى معنوية 1%.

التمرين 09:

اذا كانت البيانات التالية: 150، 140، 120، 150، 140، 125، 140، 130، 125، 140 تمثل أطوال قطع معينة من القماش الذي ينتجه مصنع ما، عند مستوى ثقة 90% اختبر الفرض القائل أن تباين أطوال الأقمشة يساوي 300.

التمرين 10:

تم أخذ عينتين مستقلتين من الرجال والنساء لقياس الزمن المستغرق لإنجاز مهمة معينة فكانت البيانات التالية:

الرجال	10.4	11.6	15.3	14.4	9.7	14.3	13.6	12.9	19.8	16.9
النساء	8.3	12.6	9.6	13.3	10.1	12.7	11.8	14.2	12.9	14.7

a) بافتراض أن الزمن يتبع التوزيع الطبيعي لمجتمعي الرجال والنساء. هل تباين الزمن للرجال يختلف عن تباين الزمن للنساء استخدمي $\alpha=0.05$؟

b) بافتراض أن الزمن يتبع توزيع طبيعي في المجتمعين وبنفس التباين. اختبري الفرض القائل بأن متوسط الزمن للنساء أقل من متوسط الزمن للرجال وذلك عند $\alpha=0.05$؟

c) قدر فترة ثقة 90% للفرق بين متوسطي الزمن للرجال والنساء؟

المراجع باللغة العربية

1. السعدي رجال، **"نظرية الاحتمالات: لكل التخصصات-لكل المستويات دروس وتمارين"**، الجزء الثاني، ديوان المطبوعات الجامعية، الجزائر، 1995.
2. بوعبد الله صالح، "محاضرات الاحصاء الرياضي"، لطلبة كلية العلوم الاقتصادية، 2005-2006.
3. حسين علي البخيت، سحر فتح الله، **"الاقتصاد القياسي"**، دار اليازوري للنشر والتوزيع، عمان، الأردن، 2009.
4. جلاطو جيلالي، **"الاحصاء: مع تمارين ومسائل محلولة"**، الطبعة الخامسة، ديوان المطبوعات الجامعية، الجزائر، 2005.
5. دومينيك سالفاتور، **"الاحصاء والاقتصاد القياسي"**، الطبعة الخامسة العربية، الدار الدولية للاستثمارات الثقافية، القاهرة، مصر، 2001.
6. عبد الرزاق بني هاني، **"الاقتصاد القياسي: المبادئ الرياضية والاحصائية"**، الطبعة الأولى، دار وائل للنشر، عمان، الأردن، 2014.
7. عدنان كريم نجم الدين، **"الاحصاء للاقتصاد والادارة"**، الطبعة الأولى، دار وائل للنشر، عمان، الأردن، 2000.
8. علي عبد السلام العماري، علي حسين العجيلي، **"الاحصاء والاحتمالات: النظرية والتطبيق"**، شركة النشر ELGA، فاليتا، مالطا، 2000.
9. عماد توما كرش، ولاء أحمد القزاز، وفاء يونس حمودي، **"علم الاحصاء"**، وزارة التعليم العالي والبحث العلمي، هيئة التعليم التقني، العراق، 2014.
10. فايز جمعة النجار، نبيل جمعة النجار، ماجد راضي الزعبي، **"أساليب البحث العلمي: منظور تطبيقي"**، الطبعة الثانية، دار الحامد للنشر والتوزيع، الأردن، 2010.
11. طالب محمد عوض، **"مقدمة في الاقتصاد القياسي"**، منشورات الجامعة الأردنية، عمان، الأردن، 2000.
12. موساوي عبد النور، بركان يوسف، **"الاحصاء 2"**، دار العلوم للنشر والتوزيع، عنابة، الجزائر، 2010.

المراجع باللغة الأجنبية

1. G.S.Maddala, “Introduction to Econometrics”, third Edition, JOHN WILEY & SONS, LTD, New York. July 2002.
2. Jarkko Isotalo, “Basics of Statistics”; Basics of Statistics course held in University of Tampere, Finland.
3. Kairat Mynbaev ; “Companion for \Statistics for Business and Economics" by Paul Newbold, William L. Carlson and Betty Thorne”, MPRA Paper No. 23069, posted 6. June 2010 16:00, Online at http://mpra.ub.uni-muenchen.de/23069/.
4. Mathieu ROUAUD ; “Probabilités, statistiques et analyses multicritères”, janvier 2017.
5. Michel Lejeune, ‘**Statistique la théorie et ses applications**’, deuxième édition, Springer Verlag France, Paris, 2010.

المواقع الالكترونية:

1- http://www.cliffsnotes.com/study-guides/statistics/sampling/populations-samples-parameters-and-statistics, visite le 21/01/2018 a 20:40.

2- https://saylordotorg.github.io/text_introductory-statistics/s15-03-f-tests-for-equality-of-two-va.html, 25/01/2018 a 9:18

2. فؤاد عبد الله العواد، "مبادئ التحليل الاحصائي"، 2015، ص ص 91-92، موجود على الموقع:
http://fac.ksu.edu.sa/sites/default/files/mbdy_lthlyl_lhsyy35.pdf. تم زيارة الموقع بتاريخ 2018/02/20، الساعة 22:00.