المحاضرات المعتمدة في مقياس الاحصاء 3 للسنة ثانية جذع مشترك

قسم: العلوم تجارية + المالية والمحاسبة

تحت إشراف:

الأستاذ: سحنون فاروق

الأستاذة : جناد مباركة

مقـــــــدمة

مقدمة

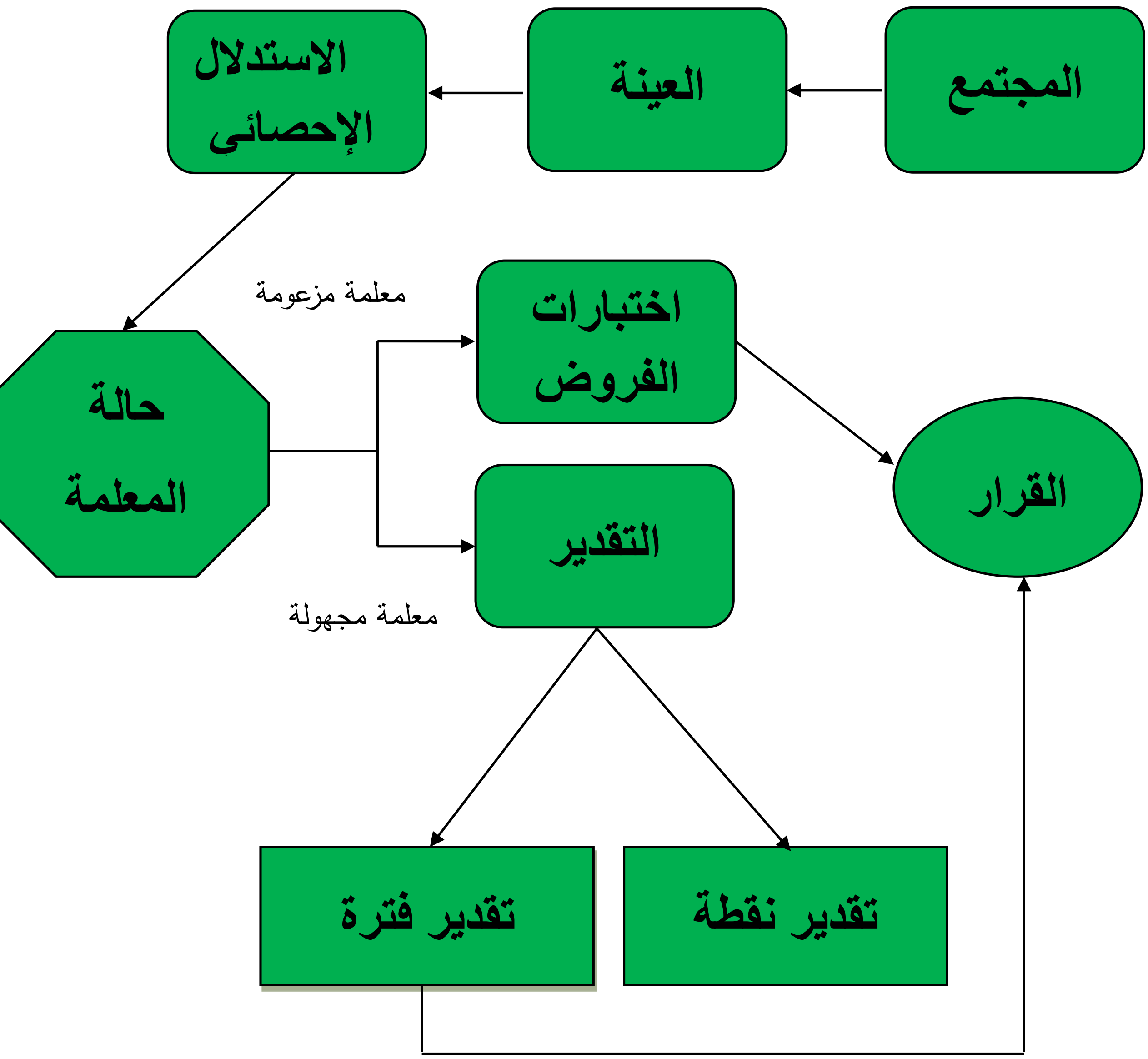
جاءت هذه المحاضرات في إحصاء 3، والموجهة لطلبة السنة الثانية جذع مشترك لقسم المالية والمحاسبة والتي تقع ضمن المنهاج الموضوع للمقرر، تهدف في الأساس إلى مساعدة الطالب على التحكم أكثر في استخدام الأساليب الإحصائية، في تحليل البيانات الخاصة بالظواهر محل الدراسة.

وإن الهدف من هذه المحاضرات هو تقديم المفاهيم الأساسية في مجال الإحصاء الاستدلالي، والذي يعتبر فرع من فروع الإحصاء يشمل كل الأساليب الإحصائية والنظريات القائمة عليها وتطبيقاتها العملية المستخدمة، في تحليل البيانات المتحصل عليها من العينة، وذلك للاستدلال على معالم وخواص المجتمع الذي سحبت منه العينة بكامله وتكون هذه الاستدلالات على شكل تقديرات أو اتخاذ القرارات، ومن الواضح أن مثل هذه العمليات تتطلب من الباحث الإلمام بالأساليب الإحصائية والرياضية، حتى يستطيع أن يختار الأسلوب الإحصائي الذي يناسب بحثه ونوع العينة وعددها ونوع المعطيات التي حصل عليها وينقسم الإحصاء الاستدلالي إلى قسمين:

(1) تقدير معالم المجتمع ويستخدم لتقدير معالم المجتمع إذا كان الهدف هو تحديد قيمة المعلمة المجهولة .

(2) اختبارات فروض بشأن صحة قيم معالم المجتمع وتستخدم بهدف الوصول إلى قرار بشأن رفض أو عدم رفض فرض إحصائي عن معلمة مزعومة.

ويتم ذلك عن طريق سحب عينة أو عينات من المجتمع المراد تقدير معالمه أو إجراء اختبارات فروض بشأنه، كما يتضح من **نموذج الاستدلال الإحصائي** التالي:



ولقد تضمنت هذه المحاضرات أربعة فصول كما يلي :

الفصل الأول: للإطار المفاهيمي؛

الفصل الثاني: توزيع المعاينة؛

الفصل الثالث: تقدير معالم المجتمع؛

الفصل الرابع: اختبار الفرضيات.

الفصـــــــل الأول
الإطار المفاهيمي

1. البيانات الإحصائية

1.1. تعريف البيانات الإحصائية

البيانات الإحصائية هي القياسات أو التعدادات أو قيم المشاهدات للظواهر أو المتغيرات أو التجارب التي يجريها الباحث أو الإحصائي. وتُعرّف أيضا بأنها مجموعة من الأرقام أو الحقائق الرقمية التي تحتاج إلى معالجة وتنظيم أو إعادة تنظيم لكي تتحول إلى معلومات، ويمكن التعبير عنها بأنها المادة الخام، يتم تحويلها إلى معلومات قابلة للاستعمال بالاعتماد على مجموعة من الوسائل والأساليب الإحصائية الكمية. و بصفة عامة، تمثل البيانات الإحصائية في البحث العلمي كل ما يحصل عليه الباحث من حقائق تخص الظاهرة محل الدراسة وتتعلق بمفردات المجتمع المبحوث. وعلى هذا الأساس فهي المادة الرئيسة في أي بحث إحصائي وترتبط دقة البحث والتحليل بمدى توافرها ودقتها وهذا يؤثر بالطبع على مدى أهمية النتائج التي نتوصل إليها وكذا على صحة ما نتخذه من قرارات على أساس هذه النتائج. وتأخذ البيانات في الواقع العملي صيغا وأشكالا مختلفة مثل: الأرقام، المؤشرات الكمية الأشكال، الرسومات البيانية أو مزيج من هذه العناصر.

2.1. أنواع وطبيعة البيانات الإحصائية

يرتبط نوع البيانات بنوع المتغير المدروس، كما تختلف طبيعة البيانات حسب الهدف من استخدامها، لهذا سنحاول من خلال هذه النقطة عرض أنواع وطبيعة البيانات الإحصائية.

1.2.1. أنواع البيانات الإحصائية

تُعبر البيانات عن قيم لمتغير أو أكثر من المتغيرات الإحصائية، ويمثل المتغير الإحصائي خاصية مشتركة أو أكثر تميز عناصر المجتمع المدروس، تختلف باختلاف الظاهرة المدروسة باختلاف الزمان والمكان وغيرها من العوامل. تفيد في تحديد جوانب المجتمع المدروس الذي سيتم تحليله لاحقا، مثل: المستوى التعليمي، الدخل، الحالة الاجتماعية، الطول، الوزن...

إن كل خاصية أو صفة من هذه الصفات تعتبر متغيراً إحصائياً في حد ذاتها، ويرتبط نوع البيانات بنوع المتغير المدروس، وبالتالي فإن أنواع المتغيرات الإحصائية تمثل بدورها نوع البيانات التي يمثلها هذا المتغير.

يأخذ المتغير الإحصائي قيما مختلفة، تمثل النتائج الممكنة للخاصية المدروسة وبالتالي فالخاصية المدروسة تسمح بوصف وتمييز الوحدات الإحصائية عن بعضها البعض، وكل وحدة إحصائية يمكن دراستها أو وصفها وفقا لخاصية أو أكثر ولا يمكن لوحدة إحصائية أن تأخذ أكثر من قيمة لكل خاصية مدروسة.

يرمز للمتغير الإحصائي برموز كبيرة كالحروف اللاتينية **Z،Y،X** أو بعنوان "**Libellé**" يعكس إسم المتغير أو بمجموع مشاهدات تصف مختلف القيم الممكنة للمتغير. ويرمز لقيم المتغير بالحروف اللاتينية الصغيرة **xi**، **yi**، **zi**، حيث يمثل **i** القيم المختلفة التي يأخذها المتغير.

تنقسم المتغيرات الإحصائية إلى نوعين رئيسين وفقا للقيم التي يأخذها هذا المتغير، هما: **المتغيرات النوعية** و**المتغيرات الكمية**، لهذا التصنيف أهمية كبيرة حيث أنه يحدد طبيعة التحليلات الإحصائية الملائمة لهذه المتغيرات. وتقاس المتغيرات الإحصائية وفق مقاييس معينة ويسمح سلَّم القياس بإنجاز العمليات الحسابية.

أ. المتغيرات النوعية (الوصفية)

المتغيرات النوعية: هي المتغيرات أو الظواهر التي لا يمكن قياسها كميا باستخدام وحدات معينة، مثل: الجنس، المستوى التعليمي، مستوى الرضا لزبائن مؤسسة معينة، الحالة الاجتماعية... فهي تُعبر عن حالات، آراء، سلوك، خصائص، صفات أشخاص أو أشياء وبالتالي هي عبارة عن قيم ذات طابع نوعي (تتعلق بالنوعية). سواء أخذت قيم المتغير النوعي ترتيبا معينا أو لم تأخذ فليس لها دلالة حقيقية مثل المتغير الكمي لأنه لا يوجد أي قياس مرجعي. تسمى البيانات التي تعبر عن هذا النوع من المتغيرات **بالبيانات النوعية.**

يقاس هذا النوع من المتغيرات بمقياسين: **المقياس الإسمي** وتسمى المتغيرات في هذه الحالة بالمتغيرات **النوعية الاسمية** و**المقياس الترتيبي** وتسمى في هذه الحالة بالمتغيرات **النوعية الترتيبية**، وهذا وفقا لقابلية قيم هذا المتغير للترتيب من عدمه.

ب. المتغيرات الكمية

المتغيرات الكمية هي الظاهرة أو الصفة التي يمكن قياسها كميا، بمعنى أن هذا النوع من المتغيرات يعكس كميا مدى توافر خاصية معينة.

فالقيم الممكنة للمتغير الكمي قابلة للقياس كميا بأرقام عددية لها خصائص حسابية وباستخدام وحدات قياس محددة (مرجعية)، تمكننا من المقارنة الدقيقة بين قيمتين مختلفتين، فهي عبارة عن أعداد حقيقية (موجبة، سالبة أو معدومة) تتعلق بكميات كالوزن، الطول، الحجم، الدخل...

تسمى البيانات التي تعبر عن مثل هذه الخصائص أو عن هذا النوع من المتغيرات بالبيانات الكمية، ونميز في الواقع بين نوعين من المتغيرات الكمية: **المتغيرات المنفصلة** و**المتغيرات المتصلة**.

- **المتغيرات المنفصلة:** هي المتغيرات التي تأخذ قيما (أعداد) أو وحدات كاملة، مثل: عدد أفراد الأسر عدد أفراد السائحين في منطقة معينة...، فهذه المتغيرات عبارة عن صفة (خاصية) تتغير بوحدات كاملة وليس بأجزاء بمعنى آخر لا يوجد انقطاع أو انفصال بين القيم، وتسمى البيانات في هذه الحالة بالبيانات الكمية المنفصلة.

- **المتغيرات المتصلة:** هي المتغيرات التي تأخذ أي قيمة في مجال معين ولا يوجد أي انفصال أو انقطاع بين القيم ولذلك تسمى قيما متصلة أو مستمرة، مثل خاصية الطول، الوزن...، فقيم هذا المتغير غير محدودة تمثل المجال المعرف حسب امتداد المتغير.

تقاس المتغيرات الكمية بمقياسي: **المقياس الفتري** و**المقياس النسبي**.

جـ. العلاقة بين المتغيرات

التصنيف السابق للمتغيرات أو البيانات الإحصائية لا يعني بالضرورة أنها نوعية أو كمية فأحيانا نجد تداخلا بينهما، حيث يمكن تحويل أي متغير كمي إلى متغير نوعي، فمتغير دخل الأسرة مثلا يمكن استبداله بفئات للدخل وذلك بتقسيم الأسر إلى أصحاب الدخل المرتفع، المتوسط والمتدني. في المقابل يمكن التعبير عن المتغيرات النوعية بقيم كمية مثل متغير العمل، حيث تمثل حالة العمل بقيمة(1) بينما تمثل حالة البطالة بقيمة (0) وغيرها من المتغيرات.

2.2.1. طبيعة البيانات الإحصائية

إن الهدف من جمع البيانات الإحصائية لا يكمن في تحصيلها وتكوين قاعدة بيانات فقط، بل إن الهدف الأساسي من ذلك هو تحليلها وصولا إلى اكتشاف المشكلات المختلفة ومن ثم الوصول إلى النتائج الأساسية لحل هذه المشكلات. على هذا الأساس فإن تبويب ونشر البيانات الإحصائية يجب أن يكون ملائما لعملية استخدامها بهدف التحليل واستنباط النتائج المرجوة.

تنشر البيانات الإحصائية إما بشكل **سلاسل الزمنية** أو **بيانات مقطعية** أو مزيج بينهما وتكون البيانات في هذه الحالة بشكل **بيانات سلسلة قطاعية.**

3.1 أنواع الأساليب الإحصائية

تختلف أساليب جمع البيانات باختلاف الهدف من الدراسة وطبيعة المجتمع المدروس وإمكانات البحث، عموما يمكن تقسيم أساليب جمع البيانات إلى ثلاثة أنواع:

- أسلوب الحصر الشامل؛
- أسلوب الحصر الجزئي؛
- أسلوب المعاينة.

1.3.1.أسلوب الحصر الشامل

الحصر الشامل هو الدراسة الشاملة لجميع وحدات المجتمع الإحصائي، بهدف الحصول على معلومات إحصائية شاملة لخاصية أو أكثر من خواص المجتمع، ومن ثم إجراء التحاليل المنهجية اللازمة.

يُصطلح عليه أيضا بالتعداد، الذي يمثل العملية الإحصائية التي تتناول عدًّا شاملا لوحدات مجتمع معين. من أمثلته التعداد العام للسكان الذي ينظم كل عشر سنوات، التعداد العام الزراعي الصناعي، والتجاري وغيرها من المجالات الأخرى.

يتطلب هذا النوع من الدراسات وقتا طويلا، تكاليف بشرية، مالية ومادية كبيرة لتنفيذها وجهوداً ضخمة لإتمامها، لذلك يتم تنظيمه عادة على فترات متباعدة كما هو الحال في التعداد العام للسكان والذي ينظم على فترات دورية منتظمة، عادة كل عشر سنوات في معظم أقطار العالم.

إن إجراء الحصر الشامل على فترات دورية منتظمة يتيح إمكانية دراسة الفروق وتقييم السياسات والإجراءات والتنبؤ بالتغيرات المستقبلية. ويهدف التعداد العام للسكان إلى معرفة خصائص السكان في فترة زمنية معينة، تمثل هذه الخصائص عدد السكان، معدلات نموهم، توزيعهم الجغرافي خصائصهم الاجتماعية والاقتصادية.

تُستخدم بيانات الحصر الشامل في التخطيط الاقتصادي والاجتماعي لمختلف البرامج على مستوى الدولة وفي الدراسات والأبحاث الإجتماعية والديموغرافية، كما يمثل التعداد القاعدة الأساسية أو الإطار الإحصائي لتنفيذ أنواع أخرى منها: الاستقصاءات بالعينة لتأمين الأنظمة أو الأجهزة الإحصائية للدول.

يتم تنفيذ الحصر الشامل وفق برنامج يتألف من خمس مراحل أساسية هي: مرحلة الأعمال التحضيرية، مرحلة العد، مرحلة تجهيز البيانات، مرحلة تحليل وتقييم البيانات، مرحلة نشر البيانات.

أ. **خصائص الحصر الشامل:** يتميز الحصر الشامل ببعض الخصائص، يمكن إيجازها فيما يلي:

– **العد الفردي:** هو تدوين البيانات من كل فرد من الأفراد المشمولين بالحصر كل على حدة بشكل مستقل تماما عن بيانات غيره، فالوحدة الإحصائية المستهدف حصرها قد تكون أسرة أو فردا أو مساكن أو مؤسسات. تمكن هذه الطريقة من تصنيف الخصائص المختلفة للسكان في جداول يحتوي كل منها على أكثر من خاصية؛

– **الشمول:** يعني ذلك أن التعداد يغطي جميع الأفراد في المناطق المشمولة بالعد دون حذف أو تكرار؛

– **الآنية:** يقصد بالآنية أن عدد الأفراد في المناطق المشمولة بالعد يُسند إلى لحظة زمنية واحدة تسمى لحظة الإسناد الزمني، وغالبا ما تحدد بمنتصف ليلة معينة بحيث تعتبر البيانات المتعلقة بعدد الأفراد وتوزيعهم الجغرافي مرتبطة بهذه اللحظة؛

– **الدورية:** يفصل بين تعداد وآخر فترة تسمى الدورة، ويقصد بالدورية أن التعدادات السكانية تجري بشكل دوري منتظم محدد في الغالب بعشر سنوات كاملة.

ب. أسباب استخدام الحصر الشامل: يُستخدم أسلوب الحصر الشامل للأسباب التالية:

- الرغبة في الحصول على بيانات شاملة عن جميع وحدات المجتمع حيث يعطي صورة كلية للمجتمع المدروس؛
- عدم إمكانية معرفة طبيعة المجتمع المدروس خاصة إذا لم تنفذ عليه بحوث في السابق، في هذه الحالة يكون من الضروري دراسة المجتمع بشكل كلي، أما إذا كانت طبيعته معلومة من حيث وحداته وقابليتها للتلف، نكتفي بدراسة عينة منه لأنه يصعب دراسة هذه المجتمعات بطريقة الحصر الشامل، مثل مجتمع الأسماك أو إنتاج دولة من النفط أو إنتاج مصنع من منتج معين؛

– عدم القدرة على تصميم عينة ممثلة للمجتمع أي تتوافر فيها خصائص المجتمع الأصلي الذي سُحبت منه، لتعميم النتائج على المجتمع الإحصائي المدروس.

جـ. مزايا وعيوب الحصر الشامل: يمكن إيجاز مزايا وعيوب الحصر الشامل في النقاط التالية:

جـ.1. المزايا: تتمثل أهم مزايا الحصر الشامل فيما يلي:

– يعطي صورة شاملة للمجتمع لأنه يوفر معلومات عن كل مفردات المجتمع، وهذا المستوى مهم جدا في عملية التخطيط والتنمية. والمعلومات الناتجة عن هذا الأسلوب تمثل إطارا كاملا للبرامج الإحصائية القادمة التي تستخدم أسلوب المعاينة في جمع المعلومات؛

– نتائج الدراسة للحصر الشامل تكون نهائية ولا تحتاج إلى تعديل أو تعميم؛

– الحصر الشامل هو الأسلوب الوحيد المناسب في بعض الحالات، مثل: التعدادات (التعداد العام للسكان، تعداد القوى العاملة و تعداد المؤسسات الصناعية...)، الحالات التي يترتب عليها أضرار كبيرة لو تركت بعض المفردات دون فحص أو دراسة (مثل فحص أسطوانات الغاز قبل توزيعها أو تطعيم الأطفال ضد أمراض معينة...)؛

– أسلوب الحصر الشامل خال من أخطاء المعاينة.

جـ.2. العيوب: تتمثل أهم عيوب الحصر الشامل فيما يلي:

– يتطلب الحصر الشامل تكاليف باهظة وجهودا ضخمة، تستغرق وقتا طويلا لإتمامها خصوصا إذا كان حجم المجتمع كبيرا. في غالب الأحيان تكون الدراسة مقيدة بوقت محدد، ولهذا فإن أسلوب الحصر الشامل لا يتيح الفرصة لدراسة تفاصيل الظاهرة مما يجعل عملية جمع بيانات شاملة وتفصيلية عن كل مفردة أمرا في غاية الصعوبة؛

– إن الموضوعات التي يغطيها هذا الأسلوب تكون شاملة وتكون الاستمارات عادة طويلة، تتطلب متخصصين وخبراء وتحتاج إلى وقت طويل لإتمامها. على هذا الأساس تتطلب خصوصية الاستمارات عددا كبيرا من الأسئلة، تهدف إلى الحصول على بيانات عامة، وتترك التفاصيل للبحوث والإستقصاءات اللاحقة. كما يحدث في مسوح الأسر، حيث تكون معلومات التعداد السكاني معلومات عامة عن عدد

الأفراد، الحالة الاجتماعية وغيرها، تستخدم فيما بعد في الاستقصاءات التي تهدف إلى جمع بيانات أكثر تفصيل ودقة والتي يصعب جمعها بأسلوب الحصر الشامل؛

– هذا النوع من الدراسات لا يناسب بعض المجتمعات التي يؤدي حصرها كلية إلى تلف مفرداتها والمجتعات ذات الأطر غير المحدودة، حيث تكون الدراسة الشاملة مستحيلة إذ لابد في هذه الحالات التضحية بجزء من المجتمع لإجراء الدراسة عليه بأسلوب المعاينة؛

– ينتج عن الحصر الشامل نوع من الأخطاء يطلق عليه الأخطاء الإجرائية كأخطاء التحيز، أخطاء عدم الإجابة، الإجابات الخاطئة...، هذا النوع من الأخطاء لا يمكن قياسه أو ضبطه بدرجة كافية. ورغم أن النوع نفسه من الأخطاء يتعرض له أسلوب المعاينة إلا نطاقها أقل نسبيا، وهناك فرصة لإمكانية ضبطها أفضل من الحصر الشامل، بالإضافة إلى سهولة اتخاذ التدابير اللازمة لمواجهة الأسباب المؤدية إليها؛

– عندما تتميز البيانات التي يغطيها الحصر الشامل بالتغير مع الزمن، فإن الحصر في هذه الحالة ما هو إلا عينة بالنسبة للزمن؛

– إن إستخدام أسلوب الحصر الشامل لا يمنع إستخدام أسلوب المعاينة في الوقت نفسه، إذ توجد العديد من البرامج الإحصائية التي تتوجه ببعض أسئلتها إلى كل الوحدات التي تنتمي إلى مجتمع معين، في حين تتوجه بأسئلة أخرى إلى جزء من الوحدات التي تنتمي إلى المجتمع نفسه وتحت البرنامج الإحصائي نفسه، كما يحدث عند استخدام عينة وحدات التعداد بهدف دراسة نوعية التباينات التي أنتجت من خلال التعداد وهو ما يسمى بالسيطرة النوعية على البيانات؛

– كما هو معروف أن التعداد العام للسكان يجري على فترات زمنية متباعدة نسبيا، مدة كل عشر سنوات نظرا لمتطلباته المادية والبشرية الكبيرة، هذا يعني أن بياناته لا تعكس واقع الحال للفترة اللاحقة وهذا ما يتم التصدي له باستخدام أسلوب المعاينة لتوفير البيانات بين كل تعدادين سكانيين.

2.3.1.أسلوب الحصر الجزئي

الحصر الجزئي هو حصر لجزء من المجتمع الإحصائي المدروس والذي يمثل الجزء الأكبر للظاهرة أو المشكلة المدروسة، ويستبعد الجزء الآخر من المجتمع نظرا لقلة أهميته أو لتوقع صعوبات في الحصول على بيانات صحيحة من هذا الجزء.

يُستخدم هذا النوع في مجالات متعددة، خاصة لحصر المؤسسات والمصانع الصغيرة والعاملين في الصناعات الحرفية التي يكون عدد وحداتها كبيرا ومساهمتها بالإنتاج قليلة (مثل: صناعة السجاد والنسيج والأحذية...) . وأحسن مثال على ذلك، الإحصاء الصناعي الذي يُستبعد منه كل المؤسسات الصناعية التي يقل عدد العمال فيها عن عشرة مع أن نسبتها تصل في بعض الأحيان إلى حوالي 95% من مجموع المؤسسات إلا أن مساهمتها في الإنتاج لا تزيد عن 20% من مجموع الإنتاج، ولهذا فإن هذا النوع من الحصر يعتبر غير كامل بالنسبة لعدد المؤسسات الصناعية المشمولة ولكنه يحتوي على نسبة كبيرة لكل من العمال و قيمة الإنتاج الصناعي.

عند استخدام هذا الأسلوب نقوم بتقسيم الوحدات الإحصائية إلى وحدات تتركز فيها الظاهرة المدروسة، نقوم بحصرها حصرا شاملا (غالبا ما يكون عددها قليل) وتسمى بالوحدات المحصورة. أما بقية الوحدات فإنها قليلة الأهمية لقلة مساهمتها في الظاهرة (رغم ضخامة عددها أحيانا) لذلك نستغني عنها ولا ندخلها في البحث، ونقوم بتقدير مساهمتها باستخدام إحدى طرق التقدير وتسمى هذه الوحدات بالوحدات المبتورة.

إن عيوب ومزايا هذه الطريقة مشابهة لعيوب ومزايا الحصر الشامل، لهذا نلجأ إلى أسلوب المعاينة لأنه أصبح من أهم الأساليب المستخدمة للحصول على البيانات الإحصائية في الوقت الحاضر.

3.3.1. أسلوب المعاينة

يُعتبر أسلوب المعاينة من أفضل الطرق العلمية المستخدمة في البحوث الإحصائية في المجالات كافة، ونظرا لتطوره فقد أصبح ضرورة علمية لكل الأبحاث مهما كان هدفها لما له من أهمية في قياس مدى مصداقية ودقة النتائج، على هذا الأساس تطور استخدامه تطورا سريعا في شتى الميادين وأصبح يلعب دورا مهما في الكثير من الدراسات النظرية والتجريبية.

يهدف أسلوب المعاينة إلى تقدير المعالم الرئيسة للمجتمع من خلال بيانات أخذت من عينة ممثلة للمجتمع، أي تتوافر فيها خصائص المجتمع الأصلي، وهذا لتخفيض أخطاء المعاينة إلى حدها الأدنى. يرتبط تمثيل العينة للمجتمع بعوامل عديدة كحجم العينة، تباين خصائص المجتمع، طريقة إختيار العينة، الطريقة المعتمدة في تصميم العينة وغيرها من العوامل الأخرى، سنتطرق لهذا الأسلوب بالتفصيل

2. المعاينة الإحصائية كأسلوب لجمع البيانات

إذا كانت طبيعة المجتمع الإحصائي الذي نقوم بدراسته وطبيعة البيانات المطلوبة تفرض على الباحث استخدام أسلوب الحصر الشامل أو أسلوب الحصر الجزئي، فإنه لاعتبارات مادية وفنية وبشرية يرى الباحثون تنفيذ كثير من البحوث باستخدام أسلوب المعاينة.

1.2. مفاهيم أساسية

أسلوب المعاينة هو عملية اختيار جزء من وحدات المجتمع الإحصائي المدروس والذي يُطلق عليه إسم المجتمع المرجعي أو المجتمع الهدف، بطريقة علمية محددة للإستدلال على خواص المجتمع.

يسمى الجزء المختار بالعينة وتسمى عملية الاختيار بالمعاينة وتسمى الطريقة المستخدمة في الاختيار بطريقة المعاينة، ويشترط في العينة أن تكون ممثلة للمجتمع أي تعكس الصورة الحقيقية أو الخواص المميزة له.

يمكن إختيار أكثر من عينة من المجتمع بطرق مختلفة لكل منها خصائص معينة وتوجد حالات يفضل فيها استخدام نوع معين دون الآخر، وللحصول على تقديرات واستدلالات دقيقة عن المجتمع محل الدراسة لابد من الاهتمام بطرق المعاينة حتى نحصل على نتائج دقيقة وقريبة من الواقع يمكن استخدامها. وفيما يلي عرض لأهم المفاهيم المرتبطة بالمعاينة الإحصائية:

1.1.2. المجتمع الإحصائي

المجتمع الإحصائي هو جميع الوحدات أو العناصر التي تشكل مجال دراسة معينة، تجمعها خاصية أو خصائص عامة مشتركة تميزها عن غيرها من المجتمعات، تمثل الظاهرة موضوع الدراسة. يجمعها إطار واحد من حيث الخصائص، الزمان والمكان تمثل هذه الوحدات المجموعة الأساسية أو المجموعة المرجعية. فقد يتكون المجتمع من أوزان مجموعة من الأفراد أو الحيوانات أو عدد السائحين أو الناخبين في دولة ما، إنتاج أحد المصانع لمنتج معين في فترة زمنية معينة أو مباني ومؤسسات وغيرها، حسب مجال وموضوع الدراسة.

2.1.2. الوحدة الإحصائية

الوحدة الإحصائية هي المفردة الأساسية من المفردات التي تشكل المجتمع الإحصائي المدروس والمطلوب جمع البيانات الإحصائية عنها، وفقا لإطار محدد بخاصية وزمان ومكان معين.

3.1.2. العينة

العينة هي مجموعة جزئية من المجتمع الإحصائي ويشترط أن تكون ممثلة للمجتمع تمثيلا دقيق، أي تعكس خصائصه من حيث الحجم وتشتت الوحدات.

يتم اختيار العينة بطريقة معينة لدراسة خصائصها والاستدلال على خصائص المجتمع، بمعنى تقدير معالم المجتمع المدروس من خلال تعميم إحصاءات العينة، باستخدام نظرية المعاينة الإحصائية التي تمكننا من تقييم مدى دقة استنتاجاتنا الإحصائية. والتي ترتبط بمدى تمثيل العينة للمجتمع.

يتشكل حجم العينة من عدد وحدات العينة المسحوبة و يرمز له بـ n. ويمكن تقسيم العينات من حيث الحجم إلى:

- **عينات صغيرة ($n < 30$):** هي العينة التي يكون حجمها أو عدد مفرداتها أقل من 30 مفردة وهذا النوع له أساليب تحليل إحصائي خاصة به.
- **عينات كبيرة ($n \geq 30$):** هي العينات التي يساوي أو يزيد عدد مفرداتها عن 30 مفردة.

ترتبط دقة النتائج بدرجة كبيرة بحجم العينة، فكلما كان الحجم كبيرا كانت النتائج أدق وكان بالإمكان استخدام الكثير من أساليب التحليل الإحصائي.

2.2. مراحل تصميم خطة المعاينة

تصميم خطة المعاينة يحدد مختلف الخطوات أو تفاصيل العمليات التي يتضمنها تنفيذ هذا النوع من الدراسات. فخطة معاينة هي وصف لتفاصيل الإجراءات المنهجية الخاصة بكيفية تنفيذ كل مرحلة من مراحل المعاينة، والتي يمكن توضيحها وعرضها من خلال النقاط التالية:

أ. تحديد مشكلة وهدف الدراسة:

تُعد هذه المرحلة من المراحل الهامة في أي دراسة أيا كان نوعها، فتعريف المشكلة المطروحة وتحديد الهدف تحديدا دقيقا يمكننا من تحديد متطلبات ومستلزمات الدراسة من متغيرات الدراسة ومن طبيعة وحجم البيانات المطلوبة وكذا المجتمع الإحصائي المستهدف، وغيرها من المتطلبات التي تشكل الإطار العام للدراسة أو جوانب الموضوع، والتي تساعدنا في تحقيق الهدف المطلوب والإجابة على إشكاليات الموضوع المدروس.

ب. تحديد وتعريف مجتمع الدراسة:

يجب تعريف وتحديد المجتمع المراد دراسته، بدقة من حيث المحتوى الذي يضمن الخصائص أو المتغيرات التي تستهدفها الدراسة ومن حيث الوحدات، فيجب تعريف هذه الأخيرة بطريقة تحدد إنتماء أو عدم إنتماء هذه الوحدة بوضوح تام، وكذا من حيث الحدود الزمانية والمكانية التي تحدد المجتمع المدروس.

جـ. تحديد البيانات المطلوبة، مصادرها وطريقة جمعها:

تتضمن هذه المرحلة تحديد طبيعة وحجم البيانات المطلوبة للدراسة وفقا لموضوعها وهدفها وفرضياتها. يجب أن تستهدف خطة المعاينة جمع بيانات محددة تخدم متطلبات الدراسة، وأن تأخذ بعين الاعتبار الدراسات السابقة المرتبطة بالموضوع والتي قد تتضمن المعلومات التي نبحث عنها بهدف حصر البيانات الضرورية وتجنب تكرار جمع بيانات تم التطرق إليها في البحوث السابقة كذلك الوقوف على أهم سلبياتها وإيجابياتها والعراقيل التي واجهتها، والإستفادة منها في تحديد بعض المؤشرات الإحصائية. هذا ما يجنبنا تكاليف إضافية ويساعدنا على تطوير جوانب وإشكاليات الدراسة وتصميم أسئلة الإستبيان بما يخدم موضوع وتساؤلات البحث، وكذلك تحديد الطريقة المناسبة لجمع البيانات.

د. تحديد طريقة المعاينة:

في هذه المرحلة يتم تحديد أسلوب المعاينة الذي سيستخدم لاختيار وحدات العينة. تنقسم أساليب المعاينة إلى مجموعتين أساسيتين هما: المعاينة الاحتمالية (العشوائية) والمعاينة غير الاحتمالية (غير العشوائية) ولكل منهما أنواع مختلفة وإجراءات خاصة بها. لهذا فإن إختيار طريقة المعاينة يتم وفقا لأهداف البحث وقيوده ويعتمد ذلك على عوامل متعددة، كالوقت والتكاليف المخصصة للبحث، مدى إمكانية الحصول على المعلومات ومدى خبرة وتكوين الباحث أو الإحصائي فهي تخضع لخبرته ورأيه الشخصي.

هـ. تحديد الإطار:

الإطار هو قائمة تحتوي على جميع الوحدات التي تشكل المجتمع الإحصائي المدروس و بالتالي هو الوسيلة التي يعتمد عليها الباحث في سحب وحدات العينة، لذلك فهو يختلف وفقا لطبيعة الدراسة ونوع المعاينة.

كما يُعتبر الإطار أيضا، الوعاء لتشكيل كل العينات الممكنة ويمكننا من خلاله معاينة المجتمع بدقة.

يلعب الإطار دورا أساسيا في تحديد طريقة المعاينة المناسبة للمجتمع محل الدراسة، وهو يخص المعاينة الاحتمالية فقط دون المعاينة غير الاحتمالية، هذه الأخيرة تعتمد أساسا على خبرة الباحث وتقديره الشخصي والهدف من الدراسة ولا تستخدم النظريات الإحصائية.

و. تحديد درجة الدقة وحجم العينة:

تخضع نتائج المعاينة إلى أخطاء المعاينة وهي الأخطاء الناتجة عن جزئية المعلومات وأخطاء القياس. وترتبط دقة المعلومات طرديا بحجم العينة والذي يرتبط بدوره طرديا بتكاليف الدراسة، لهذا تستلزم الدراسات التي تستخدم أسلوب المعاينة تحديد حجم العينة بما يتناسب مع درجة الدقة المطلوبة والميزانية المتاحة، أي تحديد نسبة الخطأ المسموح به في قبول النتائج وتعميمها على المجتمع الكلي.

يعتمد تحديد حجم العينة ودرجة الدقة المطلوبة على خبرة الباحث أو الإحصائي وعلى الدراسات السابقة وكذلك على تطبيق بعض القواعد الإحصائية الخاصة بنظرية المعاينة، وهنا يبرز دور الإحصائي في حل هذه المشكلة، بمعنى تحديد حجم العينة الذي يحقق التوافق بين دقة المعلومات وتكاليف الدراسة.

ز. إجراء اختبار مسبق:

يفضل إجراء اختبار أولي لعينة صغيرة من المجتمع المدروس لبعض الجوانب، خاصة فيما يتعلق بتعديل الاستمارة الإحصائية، فرغم إلتزام الباحث بكل متطلبات الإستمارة الجيدة وحرصه على صياغتها بشكل مناسب إلا أنه لا بد من إختبارها، حتى يمكن إكتشاف نقاط الضعف والعمل على تصحيحها. يفيد هذا الإختبار في تقييم القيمة العملية للإستمارة كما أنه يفيد في تسهيل الكثير من الصعوبات التي سيواجهها الباحث عند إجراء البحث.

ح. العمل الميداني (جمع البيانات من وحدات العينة):

يتم في بداية هذه المرحلة تدريب وتكوين القائمين بالعمل الميداني لتجنب الكثير من أخطاء القياس، حيث ترتبط دقة النتائج كثيرا بجودة هذه المرحلة. لذلك يجب الاهتمام بإعداد وتكوين القائمين بالعمل الميداني ووضع الخطط الملائمة لمعالجة حالات عدم الاستجابة أو عدم الحصول على البيانات من بعض وحدات العينة، ثم يتم بعد ذلك التنفيذ الفعلي لخطة المعاينة بجمع البيانات من مفردات العينة المختارة وهي مرحلة هامة لأنها عرضة لكثير من أخطاء القياس.

ط. مراجعة وترميز البيانات:

يقوم الباحث أو فريق العمل في هذه المرحلة بفرز ومراجعة البيانات، ثم ترميزها ومعالجتها آليا باستخدام برامج إحصائية معينة بهدف الوصول إلى النتائج ودراستها.

ي. تحليل البيانات وتقدير معالم المجتمع:

بعد الحصول على نتائج المعاينة والمتمثلة في إحصاءات العينة يتم دراستها وتحليلها بهدف الاستدلال على خصائص المجتمع، وهو أحد أهم فروع نظرية المعاينة، والذي يشمل فرعين التقدير وإختبار الفرضيات، بهدف الوصول إلى نتائج تمكن من تفسير الظاهرة المدروسة بدرجة معينة من الدقة.

3.2. أنواع المعاينة الإحصائية

تصنف المعاينة بناء على كيفية سحب العينة إلى نوعين رئيسين هما المعاينة العشوائية (الاحتمالية) والمعاينة غير العشوائية (غير الاحتمالية)، تسمى العينات في النوع الأول بالعينات العشوائية أو الإحتمالية، بينما يسمى النوع الثاني بالعينات غير العشوائية. وينقسم كل نوع بدوره إلى أنواع متعددة، يختلف تصنيفها من مرجع لآخر.

1.3.2. المعاينة العشوائية (الإحتمالية)

يعتمد هذا النوع من المعاينة على نظرية الاحتمالات، بحيث يكون لكل مفردة فرصة أو احتمال معلوم، للظهور في العينة، فكلمة "عشوائية" لا تعني أنها عينة سحبت بطريقة اعتباطية بل تعني إتاحة فرصة الاختيار لجميع الوحدات أو عناصر الظاهرة المدروسة.

تعتمد المعاينة العشوائية أيضا على توفر إطار للمعاينة، وعلى العكس من المعاينة غير الاحتمالية فإن هذا النوع من المعاينة يسمح بتقييم مدى دقة النتائج، حيث أنه لا يسمح فقط بتقدير معالم المجتمع ولكن أيضا بقياس الخطأ المحتمل في النتائج والناتج عن الدراسة الجزئية بدلا من الحصر الشامل.

ونميز عموما بين أربع طرق للمعاينة العشوائية (الإحتمالية)، وهي:

- المعاينة العشوائية البسيطة؛
- المعاينة العشوائية المنتظمة؛
- المعاينة العشوائية الطبقية؛
- المعاينة العشوائية العنقودية.

أ. المعاينة العشوائية البسيطة

هي طريقة اختيار عينة مكونة من n وحدة من بين N وحدة من وحدات المجتمع محل الدراسة، بحيث يكون لكل عينة من العينات الممكنة ذات الحجم n احتمال متساو في الظهور. يتم اختيار الوحدات الإحصائية على أساس تكافؤ الفرص، بمعنى أن لكل وحدة من وحدات المعاينة نفس الاحتمال لاختيارها في العينة. وتعتمد إجراءات سحب العينة العشوائية البسيطة على وجود إطار للعينة ونميز هنا بين نوعين من السحب:

– **السحب مع الإرجاع:** بمعنى سحب وحدات المعاينة مع إرجاع الوحدة المسحوبة في كل مرة، في هذه الحالة تسمى العينة بالعينة المستقلة، بمعنى أن الوحدة يمكن أن تظهر أكثر من مرة في العينة وبالتالي عدد العينات الممكنة في هذه الحالة ومع مراعاة ترتيب الوحدات هو: N^n.

– **السحب دون إرجاع:** يتم فيه سحب وحدات المعاينة دون إرجاع الوحدة المسحوبة، وتسمى العينة في هذه الحالة بالعينة غير المستقلة، بمعنى أن الوحدة المسحوبة لا تظهر إلا مرة واحدة في العينة وبالتالي عدد العينات الممكنة في هذه الحالة (مع عدم مراعاة الترتيب) هو: C_N^n، حيث:

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

إن المعاينة العشوائية البسيطة عن طريق السحب مع الإرجاع نادرة الاستعمال في الواقع العملي، لأنه لا معنى للحصول على الوحدة نفسها مرتين في العينة، كما أن السحب دون إرجاع يعطي تقديرات أكثر دقة، بالإضافة إلى أن التباين المرتبط بالعينة غير المستقلة يكون دائما أقل من المرتبط بالعينة المستقلة.

وفيما يلي الخطوات المتبعة في سحب عينة عشوائية بسيطة:

– إعداد قوائم تتضمن جميع عناصر المجتمع (الإطار)؛

– ترقيم جميع وحدات المجتمع بأرقام متسلسلة؛

– تحديد حجم العينة المطلوب سحبها؛

– اختيار وحدات العينة وذلك باستعمال إحدى الطرق الثلاثة التالية:

- طريقة القرعة؛

- طريقة جداول الأرقام العشوائية؛
- طريقة توليد الأرقام العشوائية بالحاسب الآلي.

يعاب على العينة العشوائية البسيطة أنها تنطوي على إجراءات طويلة ومملة، خصوصا إذا كانت العينة كبيرة، لأن الاختيار العشوائي يحتاج إلى وقت وجهد كبيرين ولتسهيل هذه العملية يمكن استخدام أسلوب المعاينة المنتظمة والذي يعتبر واسع الانتشار وكثير الاستعمال في التطبيقات العملية.

ب. المعاينة العشوائية المنتظمة

تشير تسمية هذا النوع من العينات إلى أنه يتبع أسلوبا منتظما لإختيار وحدات المجتمع دون إستخدام الأرقام العشوائية أو طرق أخرى، وهي ليست عشوائية بشكل كلي لأن فيها نوع من الإنتظام، راجع إلى تركيب وحدات العينة في إطار تصنيف أو نظام معين ولهذا سميت بالمنتظمة أو النظامية. تمتاز هذه الطريقة بتوفير كثير من الوقت والجهد وتعتبر أكثر كفاءة، من المعاينة العشوائية البسيطة، خاصة إذا كان حجم المجتمع كبيرا.

يتم إختيار العينة المنتظمة من مجتمع متجانس، وتمتاز بأنها تعتمد على العشوائية في تحديد العنصر الأول من العينة الذي يعتمد عليه تحديد باقي عناصر العينة، حيث تنتشر لتشمل المجتمع ككل فهي بذلك تكون ممثلة له.

يتم اختيار وحدات العينة من الإطار بطريقة منتظمة مرتبة وفقا لمتتالية حسابية أساسها k والذي يسمى أيضا مدى المعاينة، وذلك وفقا للخطوات التالية:

- ترقيم جميع وحدات المجتمع من 1 إلى N؛
- تحديد حجم العينة المطلوبة وليكن n؛
- حساب مدى العينة (أساس المتتالية) k والذي يمثل مسافة الاختيار بين عنصرين متتاليين ويحسب وفقا للقاعدة التالية:

$$K=\frac{N}{n}=\frac{1}{f}\quad\Rightarrow f=\frac{n}{N}=\frac{1}{K}$$

حيث:

N: حجم المجتمع؛

n: حجم العينة؛

f: نسبة أو كسر المعاينة.

– تحديد أول رقم عشوائي بين 1 و k فنحصل على رقم وحدة المعاينة الأولى ولتكن (i)، ثم بعد ذلك تحديد باقي الوحدات وفقا لأساس المتتالية k فتكون الوحدات التالية على التوالي:

$$i+k,\quad i+2k,\quad i+3k,\quad ...,\quad i+(n-1)k$$

يبين الجدول وحدات العينة في المجتمع والتي تم اختيارها بطريقة منتظمة، حيث أن أول رقم عشوائي هو i $(1 \le i \le k)$:

جدول رقم 01: ترتيب وحدات العينة المنتظمة

الوحدة في العينة	1	2	3	4	…	n
رقم الوحدة في المجتمع	I	i+k	i+2k	i+3k	…	i+(n−1)k

المصدر: سليمان محمد طشطوش، **أساسيات المعاينة الإحصائية**،دار الشروق، ط 1،،عمان،2001، ص: 94.

من أهم عيوب هذه الطريقة، أنه إذا مجتمع الدراسة يعكس اتجاهات دورية للظاهرة موضوع البحث وكان مدى المعاينة k مساويا لطول الدورة أو إحدى مضاعفاتها، فإن هذه المعاينة تكون غير ممثلة للمجتمع المدروس وبالتالي تعطي تقديرات متحيزة.

جـ. المعاينة العشوائية الطبقية

في كثير من الأحيان يكون المجتمع الإحصائي غير متجانس وذلك من حيث خصائص مفرداته بمعنى تباين أو اختلاف في الخاصية أو الخصائص المدروسة بين الوحدات، كالتباين الحاصل في دخول الأفراد والمستوى الثقافي وغير ذلك، مما يؤثر على تمثيل العينة للمجتمع وبالتالي على دقة النتائج. في هذه الحالة يكون استخدام أسلوب المعاينة الطبقية. هو المناسب لأنه يهدف إلى تصميم عينة ممثلة لكافة الطبقات التي يتكون منها المجتمع المدروس، مما يؤدي إلى تقليل أخطاء المعاينة دون زيادة حجم العينة.

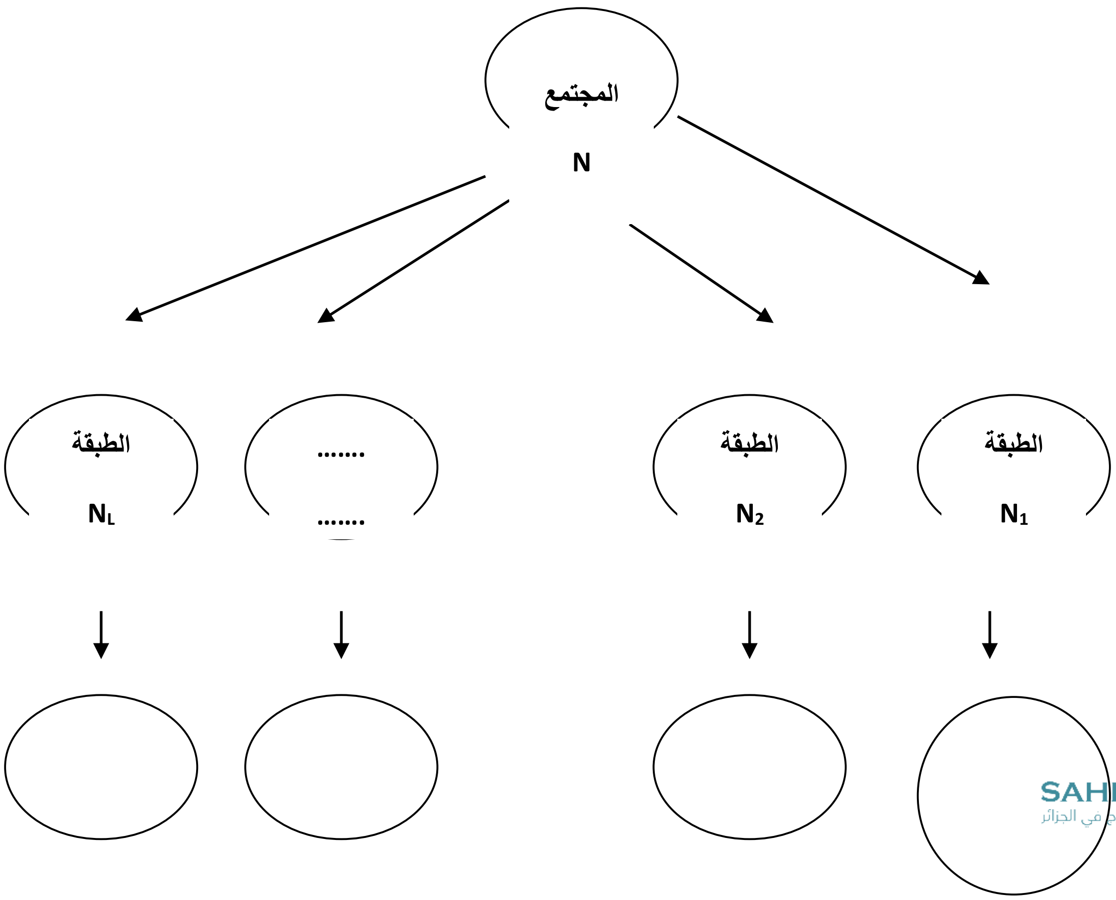
يُقسِم هذا الأسلوب المجتمع المدروس إلى مجموعات جزئية غير متقاطعة ومتجانسة نسبة للمتغير قيد الدراسة، تسمى هذه المجموعات الجزئية بالطبقات، ثم يتم إجراء معاينة عشوائية بسيطة على كل طبقة، وتعامل كل طبقة وكأنها مجتمع مستقل تسحب منه عينة عشوائية بسيطة ذات حجم معين.

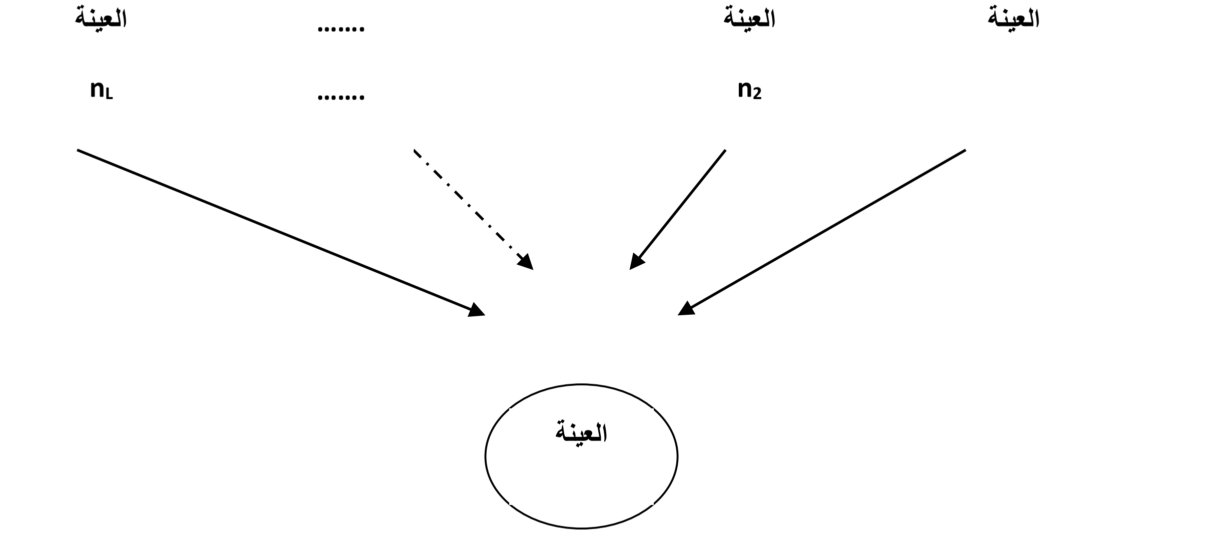
يجب أن يحقق تقسيم المجتمع التجانس بين مفردات كل طبقة لهذا يتطلب هذا النوع من المعاينة معرفة مسبقة بتركيبة المجتمع المدروس، كما أن إستخدامه تحت شروط معينة يزيد من دقة التقديرات الخاصة بمعالم المجتمع.

لابد من معرفة حجم كل طبقة في المعاينة العشوائية الطبقية، كما أن إختيار عينة من كل طبقة يستلزم وجود إطار لكل طبقة ويتم هذا الأسلوب وفقا للخطوات التالية:

- تحديد حجم العينة الإجمالي المطلوب اختياره من مجتمع البحث؛
- تقسيم المجتمع إلى طبقات متجانسة ويعتمد هذا التقسيم على التقدير الشخصي وخبرة الباحث؛
- توزيع مفردات العينة على طبقات مجتمع البحث لتمثيل هذه الطبقات؛
- سحب عدد مفردات العينة من كل طبقة وفقا لطريقة العينة العشوائية البسيطة.

يبين الشكل رقم 01: كيفية سحب عينة عشوائية طبقية من المجتمع:





يظهر من خلال الشكل أن المجتمع المكون من N وحدة قد تم تقسيمه إلى L من المجتمعات الجزئية والتي تسمى بالطبقات وكل طبقة h مكونة من الوحدات N_i حيث:

$$h = 1,2,...L$$
$$N = N_1 + N_2 + ...N_L$$

بعد ذلك يتم إختيار عينة عشوائية بسيطة من كل طبقة، وبالتالي فإن حجم العينة التي نختارها من الطبقة i هي n_i، بحيث يكون الحجم الكلي للعينة الطبقية هو:

$$n = n_1 + n_2 + ...n_L$$

لهذا النوع من العينات عيب واحد، وهو صعوبة عمل إطار للمجتمع إذا كانت الطبقات كبيرة الحجم.

د. العينة العنقودية

في بعض الحالات نجد أن وحدات بعض المجتمعات تشكل تجمعات عادة ما تكون مشابهة إلى حد كبير للخاصية المدروسة مثل المدن، الشوارع، المناطق، الجامعات...، هذه التجمعات تسمى العناقيد ويتم اللجوء إلى هذه الطريقة إذا كان المجتمع كبيرا جدا.

يعتمد هذا النوع من العينات على تجزئة مجتمع الدراسة إلى مجموعات (عناقيد) وذلك وفقا لخاصية معينة كما هو الحال في العينة الطبقية، بعدها يتم الاختيار العشوائي لعينة الدراسة والمتمثلة في بعض هذه العناقيد، كعينة عشوائية بسيطة ثم ندرس أفراد كل منها، وفي هذه الحالة تسمى عينة عنقودية من مرحلة واحدة. أما إذا قمنا باختيار عينة عشوائية بسيطة من الأفراد داخل كل عنقود اخترناه في المرحلة

الأولى فتسمى عينة عنقودية من مرحلتين، عندها تسمى العناقيد بوحدات معاينة أولية والمفردات داخل العناقيد تسمى وحدات معاينة ثانوية.

يمكن أن تكون العينة العنقودية مكونة من عدة مراحل وتسمى في هذه الحالة عينة عنقودية متعددة المراحل. وتوجد عدة عوامل يجب مراعاتها عند استخدام العينة العنقودية من بينها ما يلي:

- يجب أن تكون العناقيد معرفة بدقة وكل مفردة من مجتمع الدراسة يجب أن تنتمي لمجموعة أو عنقود واحد فقط؛
- يجب أن يكون عدد المفردات في العنقود معروفا؛
- يجب اختيار العناقيد عشوائيا لتقليل خطأ العينة؛
- يجب مراعاة التوازن في حجم العناقيد لتقليل خطأ العينة.

العيب الوحيد لهذا النوع أن درجة دقة النتائج عادة ما تكون منخفضة.

2.3.2. المعاينة غير العشوائية (غير الاحتمالية)

في هذا النوع يتم اختيار الوحدات بطريقة غير عشوائية، حيث يتعمد الباحث إختيار وحدات معينة لإدخالها في العينة على إعتبار أنها تمثل المجتمع المدروس تمثيلا جيدا، وذلك حسب رأيه وخبرته ومن ثم إجراء الدراسة على العينة المختارة. ومن أكثر الطرق إستخداما في المعاينة غير العشوائية ما يلي:

- عينة الحصص؛
- العينة الميسرة؛
- العينة الحكمية (القصدية).

أ. عينة الحصص

طريقة الحصص، هي أكثر طرق المعاينة غير الاحتمالية استعمالا، خاصة في التحقيقات الإجتماعية والاقتصادية (دراسات السوق، سبر الآراء...)، تقوم هذه الطريقة على اختيار عدة خصائص للمجتمع، بحيث ترتبط بموضوع البحث وتسمى هذه الخصائص بمتغيرات المراقبة.

يتم الحصول على الحصص من خلال ضرب نسبة المعاينة في القيم المناسبة لمتغيرات المراقبة، ثم تترك عملية اختيار العناصر من مجتمع البحث للمقابل الميداني، ويعتمد المقابل في اختياره على الخصائص المحددة لموضوع الدراسة، ومن أمثلة الخصائص المرتبطة بموضوع البحث: السن، المهنة، الحالة الاجتماعية...

لا تحتاج طريقة الحصص إلى استخدام إطار المعاينة وهي الميزة المناسبة للحالات التي لا تتوافر على إطار المعاينة وفي الدراسات التي تتميز بالسرية الإحصائية. كما أن تكلفة المعاينة بطريقة الحصص تكون أقل من تكلفة المعاينة الاحتمالية بسبب انخفاض تكلفة التنقل لاستجواب الوحدة الإحصائية التي يتم اختيارها قصدا، عكس المعاينة الاحتمالية التي تتطلب ربما تنقلات متعددة وبالتالي تكاليف أكبر.

يعاب على هذه الطريقة عدم إمكانية حساب دقة التقدير المحصل عليه حيث أن حرية اختيار الوحدات يؤدي إلى إرتكاب أخطاء التحيز، كما أن صعوبة متابعة العمل الميداني يؤدي إلى تمثيل أقل للوحدات المدروسة.

ب. العينة الميسَّرة

يتم إختيار هذا النوع من العينات بصفة عرضية أي عن طريق الصدفة، تتميز إجراءاتها بالبساطة حيث يقوم الباحث باختيار وحدات العينة التي يمكن الاتصال بها بسهولة وبأقل جهد وتكلفة ممكنة. ومن عيوب، العينات الميسَّرة أنها لا تضمن توافر الوحدات على بعض الخصائص المطلوبة وبالتالي لا يمكن اعتبارها ممثلة للمجتمع ويصعب تعميم نتائجها على بقية أفراد مجتمع الدراسة.

كثيرا ما تستخدم العينة المسيرة في الدراسات الاستطلاعية بهدف جمع أكبر عدد من البيانات وخاصة عندما تُستخدَم كمقدمة لدراسات لاحقة تعتمد على عينات إحتمالية.

جـ. العينة الحكمية (القصدية)

يتم إختيار مفردات هذه العينة حسب ما يراه الباحث مناسبا، على أساس أنها تمثل مجتمع الدراسة أو ذات خبرة في الميدان وبالتالي تخدم أغراض الدراسة. وعادة ما يتم اختيار هذا النوع من العينات عندما يكون حجم العينة صغيرا جدا، حيث أنها تكون أكثر مصداقية من العينات الاحتمالية.

4.2. حجم العينة ومصادر أخطاء المعاينة

إن السؤال الأساسي الذي يجب أن يطرحه الباحث هو: "كم عدد الأشخاص الواجب إخضاعهم للدراسة حتى تكون العينة ممثلة للمجتمع المدروس؟". على هذا الأساس سنتطرق من خلال هذه النقطة إلى العوامل التي يمكن أن تؤثر في تحديد حجم العينة، ثم نتطرق إلى تحديد حجم العينة حسب أهم أنواع المعاينة، وأيضا مصادر الأخطاء في دراسات العينة.

1.4.2. العوامل المؤثرة في تحديد حجم العينة

يجب الموازنة بين عاملين رئيسيين عند تحديد حجم العينة هما:

– الدقة في تمثيل المجتمع وهذا يتطلب الزيادة في حجم العينة، ويقلل من خطأ المعاينة؛

– تكلفة الإستقصاء والتي ترتبط طرديا بحجم العينة.

من هذا التناقض بين العاملين يظهر مدى تعقد قرار حجم العينة، إلا أن بعض الباحثين قد توصلوا إلى أن هناك عوامل رئيسة تؤثر في قرار حجم العينة، وهي:

- **درجة الدقة المطلوبة في النتائج:** ترتبط دقة النتائج طرديا بحجم العينة والذي يرتبط بدوره عكسيا مع خطأ المعاينة، لهذا وبهدف الحصول على نتائج دقيقة يجب على الباحث إختيار حجم عينة كبير؛

- **درجة تجانس المجتمع المدروس:** كلما كان المجتمع غير متجانس من حيث الخصائص المدروسة، تطلب أن يكون حجم العينة أكبر نسبيا والعكس صحيح؛

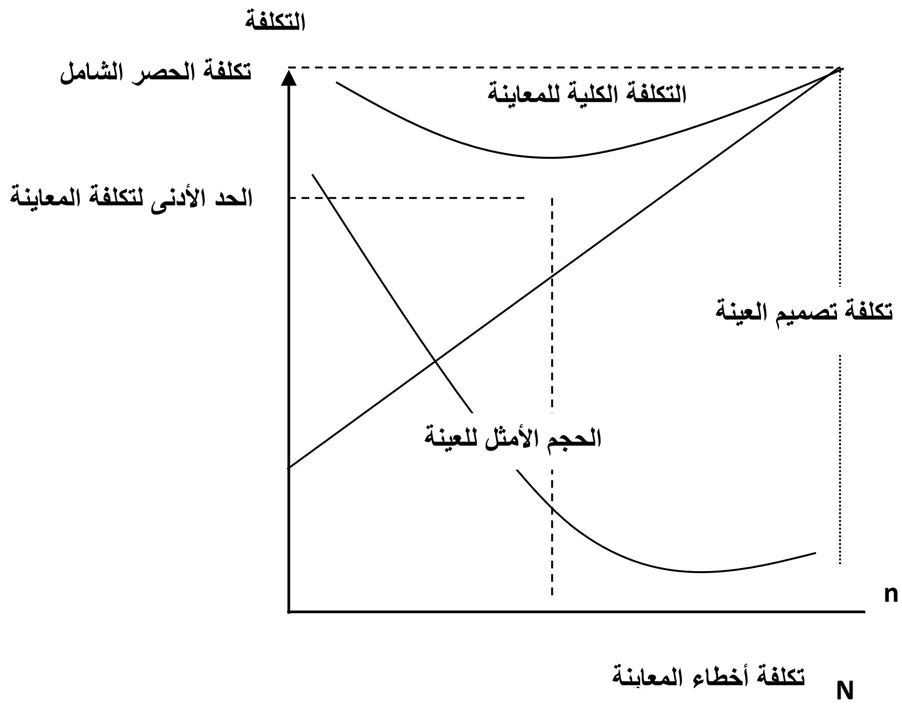
- **حجم المجتمع المدروس:** يجب الأخذ بعين الإعتبار أنه كلما كبر حجم المجتمع، كان من الأفضل زيادة حجم العينة، لكن يبقى هذا الارتباط بين حجم المجتمع وحجم العينة محدودا، فإذا كان المجتمع متجانس، تكفي عينة صغيرة لتمثيله؛

- **القيود الداخلية للدراسة:** تتمثل في تكاليف الدراسة (يجب أن تكون في حدود الميزانية المخصصة) وكذا الوقت المتاح للدراسة.

بصفة عامة، للحصول على نتائج دقيقة يجب على الباحث إعتماد حجم عينة كبير، هذا بدوره يؤدي إلى تخفيض الخطر المرتبط بخطأ المعاينة، في المقابل ترتفع تكاليف الدراسة. وبالتالي فإن حجم العينة الأمثل هو الحجم الذي يسمح بتحقيق التوازن بين خطر إرتكاب أخطاء المعاينة والتكاليف المرتبطة بها(عدم

دقة المعلومات) وتكلفة المعاينة في حد ذاتها،بمعنى حجم العينة الأمثل هو الذي يحقق الحد الأدنى للتكلفة الإجمالية للمعاينة وفي الوقت نفسه يخفض تكلفة أخطاء المعاينة كما يوضحه الشكل:

شكل رقم 02: الحجم الأمثل للعينة



2.4.2. تحديد حجم العينة

إن طريقة تحديد حجم العينة تختلف على حسب الأهداف المسطرة للدراسة، وهامش الخطأ الأعظمي المقبول من قبل الباحث، وكذا طريقة إختيار العينة المتبعة. وفي هذا الصدد نقترح الصيغ الأكثر استعمالا في الدراسات الإحصائية:

أ. تحديد حجم العينة العشوائية البسيطة

في حالة العينة العشوائية البسيطة، فإن حجم العينة يمكن حسابه بدلالة هامش الخطأ المقبول من طرف الباحث وهناك حالتين:

– **تحديد حجم العينة في حالة النسبة:** يتحدد حجم العينة في حالة ما إذا كانت القيم محل القياس أو الدراسة عبارة عن نسبة كما يلي:

$$n = \frac{(1{,}96)^2 \times (p) \times (1-p)}{(e)^2}$$

حيث أن:

n:حجم العينة؛

p: نسبة مفردات مجتمع الدراسة التي تتوافر فيها الخاصية محل الدراسة؛

e: الخطأ الأعظمي المقبول.

– تحديد حجم العينة في حالة المتوسط:

عوض تقييم النسبة، مثل: "كم عدد المدخنين الذين يدخنون علبتي سيجارة في اليوم"، فإن الدراسة يمكن أن تأخذ كهدف "حساب متوسط عدد علب السيجارة المستعملة في يوم واحد من قبل المستهلك". وهنا يتحدد حجم العينة كما يلي:

$$n = \left[\frac{1{,}96.s}{e}\right]^2$$

حيث أن :

n: حجم العينة؛

s: الانحراف المعياري للعينة؛

e: الخطأ الأعظمي المقبول.

ب. تحديد حجم العينة العشوائية الطبقية

في حالة العينة العشوائية الطبقية فإن حجم العينة في حالة النسبة يمكن حسابه حسب الصيغة التالية:

$$n = \frac{\sum_{i=1}^{L} N_i^2 \frac{p_i(1-p_i)}{w_i}}{N^2 D + \sum_{i=1}^{L} N_i p_i (1-p_i)}$$

حيث أن:

n:حجم العينة؛

p_i: نسبة مفردات المجتمع المتوافرة على الخاصية المدروسة والمنتمية إلى الطبقة i؛

w_i: وزن الطبقة i (ويحسب بقسمة حجم الطبقة i على حجم المجتمع)؛

N_i: حجم الطبقة i؛

N: حجم المجتمع؛

D: ويساوي مربع الخطأ الأعظمي المقبول مقسوما على مربع مجال الثقة.

ج. تحديد حجم العينة بدلالة الميزانية المتاحة

تقوم بعض المراكز المتخصصة في دراسات السوق بالاعتماد على حجم الميزانية المخصصة للدراسة لتحديد حجم العينة وتحسب كما يلي:

$$\boxed{n = \frac{B}{C}}$$

حيث أن :

n:حجم العينة؛

B: الميزانية المخصصة للدراسة؛

C: التكلفة المتوسطة لاستقصاء شخص واحد.

إلا أن هذه الطريقة لتحديد حجم العينة لا تعتمد على أي أساس أو مبرر علمي أو موضوعي وبالتالي فإنه أسلوب غير مقبول من الناحية العلمية رغم إتباعه من قبل بعض مراكز الدراسات.

3.4.3. مصادر الأخطاء في دراسات العينة

رغم الاحتياطات التي يأخذها الباحث بعين الاعتبار في إجراءات المعاينة، تكون نتائج الدراسة عرضة لأخطاء يمكن تصنيفها كما يلي:

أ. أخطاء المعاينة:

يتعلق خطأ المعاينة بمدى تمثيل العينة للمجتمع، يحدث أثناء تصميم العينة، ناتج عن جزئية المعلومات بمعنى أنه يرتبط أساسا بحجم العينة(حجم العينة غير كافي)، والسحب العشوائي للوحدات وينقسم إلى: الخطأ العشوائي والخطأ الناتج عن تباين المجتمع.

ب. أخطاء القياس:

هي الأخطاء غير الاحتمالية، وتسمى أيضا أخطاء عدم المعاينة، وتضم أخطاء الملاحظة(العمل الميداني، معالجة البيانات) وأخطاء عدم الملاحظة (عدم الإجابة، عدم التغطية).

.

الفصل الثاني
توزيع المعاينة

يعد فهم توزيعات المعاينة مدخلا مهما لفهم موضوع الإحصاء الاستدلالي حيث يستخدم توزيع إحصاءات العينة بشكل واسع في موضوع اختبار الفرضيات الإحصائية.

فتوزيع المعاينة هو التوزيع الاحتمالي لأي إحصائية (مقدرة) تحسب قيمتها من كل العينات العشوائية ذات الحجم المتساوي والممكن سحبها من المجتمع، ويعرف توزيع المعاينة بأنه عبارة عن " توزيع كافة القيم المحتملة التي يمكن إفتراضها بإحصاءة ما محسوبة على أساس عينات مختلفة بنفس الحجم اختيرت عشوائيا من نفس المجتمع ".

يعتمد شكل ومعالم توزيع المعاينة على طبيعة المجتمع وحجمه (معلوم أو مجهول غير محدود) وعلى طريقة سحب العينات (بالإرجاع أو دون إرجاع).

وفي ضوء ما تقدم، سيتم تسليط الضوء على بعض توزيعات المعاينة الشائعة بشيء من التفصيل.

1- توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$

العينات المختلفة غالبا ما تكون قيم متوسطاتها الحسابية مختلفة، أي أن الوسط الحسابي للعينة $\bar{X}$، كأي إحصائية يتغير من عينة إلى أخرى، ولا نستطيع معرفة الوسط الحسابي لأي عينة عشوائية قبل سحبها، أي أن الوسط الحسابي للعينة هو متغير عشوائي وبالتالي له توزيع احتمالي. وكما ذكرنا يطلق على التوزيع الاحتمالي لأي إحصائية توزيع معاينة، وبالتالي يسمى التوزيع الاحتمالي للوسط الحسابي للعينة، توزيع المعاينة للوسط الحسابي للعينة، والذي يمكن تعريفه كما يلي:

تعريف توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$: إذا كان لدينا مجتمع محدود حجمه N ووسطه الحسابي μ وتبيانه σ^2, وسحبنا منه كل العينات العشوائية البسيطة الممكنة المتساوية في الحجم وليكن حجمها n، وحسبنا الوسط الحسابي $\bar{X}$ لكل عينة، فالتوزيع الاحتمالي للمتغير العشوائي $\bar{X}$ يُسمى توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$.

وسنجد أنه أياً كان التوزيع الاحتمالي للمجتمع، فهناك علاقة بين الوسط الحسابي لتوزيع المعاينة $E(\bar{X})$ والوسط الحسابي للمجتمع μ، وعلاقة بين تباين توزيع المعاينة $V(\bar{X})$ وتباين المجتمع σ^2، وهذه العلاقات موضحة فيما يلي:

1- الوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة يساوى الوسط الحسابي للمجتمع، أي أن: $E(\overline{X}) = \mu$

2- العلاقة بين تباين توزيع المعاينة للوسط الحسابي للعينة وتباين المجتمع، تعتمد على حجم المجتمع محدوداً أم غير محدود، وعلى الطريقة التي تم بها سحب العينات من المجتمع، فهل تم السحب مع الإرجاع، أي إرجاع المفردة المسحوبة إلى المجتمع قبل سحب المفردة التي تليها، أم تم السحب مع عدم الإرجاع، أي سحب المفردة دون إرجاع المفردة التي سُحبت قبلها، وذلك كما هو موضح فيما يلي:

أ- إذا كان حجم المجتمع غير محدود أو كبيراً جدا أو محدوداً (صغيراً) ولكن عملية سحب مفردات العينة تمت مع الإرجاع، أو تمت مع عدم الإرجاع ولكن حجم العينة صغير جدا نسبة لحجم المجتمع، وذلك عندما يكون $\frac{n}{N} \leq 0.05$ ، فإن تباين توزيع المعاينة للوسط الحسابي للعينة يساوي تباين المجتمع مقسوما على حجم العينة n،

أي أن: $V(\overline{X}) = \frac{\sigma^2}{n}$

ب- إذا كان حجم المجتمع محدوداً (صغيراً) وعملية سحب المفردات من العينة تمت مع عدم الإرجاع و $\frac{n}{N} > 0.05$ ، فإن العلاقة ستأخذ الصورة التالية:

$$V(\overline{X}) = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$$

تلخص النظرية التالية هذه العلاقات:

نظرية: إذا كان لدينا مجتمع محدود حجمه N، وسطه الحسابي μ وتباينه σ^2 ، وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم n ، فإن الوسط الحسابي $E(\overline{X})$ ، والتباين $V(\overline{X})$ لتوزيع المعاينة للوسط الحسابي للعينة يرتبط بالوسط الحسابي وتباين المجتمع بالعلاقات التالية:

$$E(\overline{X}) = \mu$$

- ✓ في حالة السحب مع الإرجاع ، أو مع عدم الإرجاع و $\frac{n}{N} \leq 0.05$ فإن: $V(\overline{X}) = \frac{\sigma^2}{n}$
- ✓ في حالة السحب مع عدم الإرجاع و $\frac{n}{N} > 0.05$ فإن: $V(\overline{X}) = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

نلاحظ من هذه العلاقات، أننا نستطيع حساب تباين توزيع المعاينة للوسط الحسابي للعينة من الصيغة $V(\bar{X}) = \frac{\sigma^2}{n}$ بغض النظر ما إذا تمت عملية السحب مع الإرجاع أو مع عدم الإرجاع، وذلك عندما يكون حجم المجتمع كبيراً جداً، أو عندما تكون نسبة حجم العينة إلى حجم المجتمع (n/N) أقل من أو تساوي 0.05، لأن في هذه الحالة يقترب معامل التصحيح $(\frac{N-n}{N-1})$ من الواحد الصحيح. أما إذا كان حجم المجتمع محدوداً والنسبة (n/N) أكبر من 0.05، فيجب أن نهتم عندئذ بعملية السحب، هل تمت مع الإرجاع أو مع عدم الإرجاع.

ملاحظة:

إذا كان σ^2 مجهول و $n < 30$ فإن: $$V(\bar{X}) = \frac{S^2}{n-1}$$

حيث: $$S^2 = \frac{\sum(X_i - \bar{X})}{n-1}$$

وللحصول على توزيع المعاينة للوسط الحسابي للعينة، يجب معرفة كل العينات العشوائية الممكن سحبها من المجتمع لحساب متوسطاتها، وقد علمنا أن عدد العينات العشوائية ذات الحجم n الممكن سحبها من مجتمع محدود حجمه N ، يحدد كما يلي:

1- إذا كان السحب مع الإرجاع، يكون عدد كل العينات الممكن سحبها يساوي N^n .

2- إذا كان السحب مع عدم الإرجاع، فتوجد طريقتان لتحديد عدد العينات الممكن سحبها وهما:

أ–تُعطى أهمية لترتيب المفردات داخل العينة، وفي هذه الحالة نحدد عدد العينات الممكن سحبها كمايلي:

$$_{N}P_{n} = \frac{N!}{(N-n)!}$$

ب–إذا لم تُعطي أهمية لترتيب المفردات داخل العينة، وفي هذه الحالة نحدد عدد العينات بحساب التوافيق والذي نرمز له كما يلي:

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

بمقارنة القانونين في أ، ب، نلاحظ أن العدد الكلي للعينات مع عدم إهمال الترتيب تساوي (n!) مضروبا في العدد الكلي للعينات عند إهمال الترتيب، لأن في حالة عدم إهمال الترتيب تتكرر العينة المحتوية على

نفس القيم $(n!)$ من المرات. وبالتالي يكون تكرار أي قيمة للوسط الحسابي للعينة عند عدم إهمال الترتيب، يساوي $(n!)$ مضروبا في تكرار نفس القيمة عند إهمال الترتيب، ولذلك يكون احتمال أي قيمة للوسط الحسابي في حالة عدم إهمال الترتيب هو نفسه احتمالها عند إهمال الترتيب، أي سنحصل على نفس توزيع المعاينة سواء أهملنا الترتيب أو لم نهمله، وحيث إن عند إهمال الترتيب سنتعامل مع عدد أقل من العينات، فسنهمل دائما الترتيب في حالة السحب مع عدم الإرجاع لتسهيل العمليات الحسابية.

مثال: بفرض أن أحد المجتمعات يتكون من 4 أفراد حسب العمر $\{54, 32, 18, 16\}$

- أوجد وسط وتباين المجتمع.
- عند سحب عنصرين من المجتمع أحسب $E(\bar{X})$ و $V(\bar{X})$ إذا كانت العينة عشوائية بسيطة مع الإعادة ، بدون إعادة.
- قارن بين كل هذه النتائج.

الحل:

1- حساب وسط وتباين المجتمع:

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{16+18+32+54}{4}$$

$$\mu = 30$$

$$V(\bar{X}) = \sigma^2 = \frac{1}{N}.\sum_{i=1}^{N}(X_i - \mu)^2$$

$$\sigma^2 = 230$$

2- حساب $E(\bar{X})$ و $V(\bar{X})$ في كل حالة من الحالات التالية:

1.2. عينة عشوائية بسيطة مع الإعادة

$[\bar{X}_i - E(\bar{X})]^2$	$\bar{X}_i$	العينة
196	16	16−16

16−18	17	169
16−32	24	36
16−54	35	25
18−18	18	144
18−16	17	169
18−32	25	25
18−54	36	36
32−32	32	4
32−16	24	36
32−18	25	25
32−54	43	169
54−54	54	576
54−16	35	25
54−18	36	36
54−32	43	169
Σ	480	1840

$$E(\overline{X}) = \frac{1}{16}\sum_{i=1}^{16}\overline{X}_i$$

$$= \frac{480}{16} \rightarrow \quad E(\overline{X}) = 30$$

$$V(\overline{X}) = \frac{1}{16}\sum_{i=1}^{16}(\overline{X}_i - E(\overline{X}))^2$$

$$= \frac{1840}{16} \rightarrow V(\overline{X}) = 115$$

نلاحظ أنّ:

- $E(\overline{X}) = \mu = 30$
- $\frac{\sigma^2}{n} = \frac{230}{2} = 115 = V(\overline{X})$

.2.2 عينة عشوائية بسيطة بدون إعادة

العينة	$\overline{X}_i$	$[\overline{X}_i - E(\overline{X})]^2$
16−18	17	169
16−32	24	36
16−54	35	25
18−16	17	169
18−32	25	25
18−54	36	36
32−16	24	36
32−18	25	25
32−54	43	169
54−16	35	25
54−18	36	36
54−32	43	169
$\sum$	360	920

$$E(\overline{X}) = \frac{1}{12}\sum_{i=1}^{12}\overline{X}_i$$

$$= \frac{360}{12} \rightarrow \quad E(\overline{X}) = 30$$

$$V(\overline{X}) = \frac{1}{12}\sum_{i=1}^{12}(\overline{X}_i - E(\overline{X}))^2$$

$$= \frac{920}{12} \rightarrow V(\overline{X}) = 76.67$$

نلاحظ أنّ:

- $E(\overline{X}) = \mu = 30$
- $\frac{\sigma^2}{n}.\frac{N-n}{N-1} = \frac{230}{2} \times \frac{4-2}{4-1}$
- $\frac{\sigma^2}{n}.\frac{N-n}{N-1} = 76.67 = V(\overline{X})$

3- المقارنة بين النتائج :

$$E(\overline{X}) \qquad = E(\overline{X}) \qquad = \mu \qquad\qquad (1)$$

عينة
عشوائية
بسيطة مع

عينة
عشوائية
بسيطة مع

$$V(\overline{X}) \qquad > \qquad V(\overline{X}) \qquad (2)$$

عينة عشوائية بسيطة
مع
إعادة

عينة عشوائية بسيطة
بدون
إعادة

ومنه $\overline{X}$ في حالة العينة العشوائية البسيطة بدون إرجاع أكثر كفاءة من $\overline{X}$ في حالة العينة العشوائية بالإرجاع

(3) عدد العينات المشكلة في حالة العينة العشوائية البسيطة بدون إعادة (12) وحالة العينة العشوائية البسيطة مع الإعادة (16).

2. حالات مختلفة لإيجاد توزيع الوسط الحسابي للعينة

1.2 إذا كان تباين المجتمع معلوم

إذا كان $X\sim N(\mu,\sigma^2)$ وسُحبت منه عينة حجمها n توزيع الوسط الحسابي للعينة هو: $\bar{X}\sim N(\mu,V(\bar{X}))$

ولحساب الاحتمالات للمتغير العشوائي $\bar{X}$ يتم إيجاد القيم المعيارية له حيث:

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

مثال: إذا كان $X\sim N(9,4)$، فما هو التوزيع الاحتمالي لمتوسط عينة عشوائية بحجم 25 من مجتمع X، أحسب $P(\bar{X}>10)$.

الحل:

حيث أن $X\sim N(9,4)$ إذن $\bar{X}\sim N(9,\frac{4}{25})$.

ولحساب الاحتمال المطلوب لاحظ أن $Z=\frac{\bar{X}-9}{2/\sqrt{25}}=\frac{\bar{X}-9}{0.4}\sim N(0,1)$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{X}>10)=P\left(\frac{\bar{X}-9}{2/\sqrt{25}}=\frac{10-9}{2/5}\right)=P(Z>2.5)$$

$$=1-P(Z>2.5)=1-0.9938=0.0062$$

2.2 إذا كان تباين المجتمع الطبيعي مجهول

في هذه الحالة نقدر تباين المجتمع بتباين العينة المسحوبة منه ونفرق هنا بين حالتين:

1.2.2 عندما يكون حجم العينة $n\geq 30$

يكون توزيع الوسط الحسابي للعينة هو:

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

حيث S^2 هو تباين العينة ويحسب من العلاقة: $S^2 = \frac{\sum(X_i-\overline{X})^2}{n-1}$ ويكون التوزيع المعياري المقابل هو:

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

مثال: إذا كان $X \sim N(7, \sigma^2)$، فما التوزيع الاحتمالي لمتوسط عينة عشوائية بحجم 36 من مجتمع X تباينها 9، ثم أحسب $P(\overline{X} > 8)$.

الحل: حيث أن تباين المجتمع مجهول فإن: $\overline{X} \sim N(\mu, \frac{S^2}{n}) = N(7, \frac{9}{16})$.

ولحساب الاحتمال المطلوب لاحظ أن $Z=\frac{\overline{X}-7}{0.5}=\frac{\overline{X}-7}{3/\sqrt{36}} \sim N(0,1)$

وبالتالي يكون الاحتمال المطلوب:

$$P(\overline{X} > 8) = P\left(Z > \frac{8-7}{0.5}\right) = 1 - P(Z < 2) = 1 - 0.9772 = 0.0228$$

2.2.2 عندما يكون حجم العينة $n < 30$

يكون توزيع الوسط الحسابي للعينة هو:

$$T = \frac{\overline{X} - \mu}{s/\sqrt{n-1}} \sim t(n-1)$$

مثال: إذا كان $X \sim N(8.5, \sigma^2)$ أوجد التوزيع الاحتمالي لوسط العينة {9,10,9,8,8,7,8,5,8} ثم أحسب $P(7 < \overline{X} < 9)$

الحل: بما أن تباين مجتمع X مجهول وحجم العينة n=9 وهو أقل من 30 يكون توزيع الوسط الحسابي للعينة هو $\overline{X} \sim t(8)$.

ولحساب الاحتمال المطلوب نجد أولا تباين العينة:

$$\overline{X} = \frac{\sum X}{n} = \frac{72}{9} = 8$$

$$S^2 = \frac{\sum(X-8)^2}{8} = 2$$

وبالتالي يكون التوزيع :

$$T=\frac{\bar{X}-8.5}{1.14/\sqrt{8}}\sim t(8)$$

ويكون الاحتمال المطلوب:

$$P(7<\bar{X}<9)=P\left(\frac{7-8.5}{1.14/\sqrt{8}}<\frac{\bar{X}-8.5}{1.14/\sqrt{8}}<\frac{9-8.5}{1.14/\sqrt{8}}\right)$$

$$=P(-3<T<1)$$

$$=P(T<1)-P(T<-3)$$

وحيث أن التوزيع متماثل حول الصفر، إذن $P(T<-3)=P(T>3)$ ومن قوانين الاحتمالات نعلم أن: $P(T<1)=1-P(T>1)$ لذلك يكون المطلوب

$$P(7<\bar{X}<9)=1-P(T>1)-P(T>3)$$

$$=1-0.1-0.01=0.89$$

3.2 إذا كان المجتمع طبيعي ذو حجم محدود وكان حجم العينة $n\geq 30$ وكان السحب بدون إرجاع

1.3.2عند معلومية σ^2 يكون توزيع الوسط الحسابي للعينة هو:

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

مثال: من مجتمع بتوزيع طبيعي حجمه 200 مشاهدة وتباينه $\sigma^2=64$ تم سحب عينة عشوائية بدون إرجاع بحجم 50 مشاهدة ، فما هو توزيع الوسط الحسابي للعينة؟ ثم أحسب $P_r(\bar{X}>32)$ عندما $\mu=30$.

الحل: حيث أن تباين المجتمع معلوم فإن توزيع الوسط الحسابي للعينة هو:

$$\bar{X}\sim N\left(30.\frac{64}{50}\times\frac{200-50}{200-1}\right)=N(30.0.96)$$

$$Z=\frac{\bar{X}-30}{\sqrt{0.96}}\sim N(0,1)$$

ويكون الاحتمال المطلوب هو

$$P(\bar{X} > 32) = P\left(Z > \frac{32-30}{\sqrt{0.96}}\right) = P(Z > 2.02)$$

$$= 1 - P(Z < 2.02) = 1 - 0.9783 = 0.0217$$

2.3.2 عند مجهولية σ^2 يكون توزيع الوسط الحسابي للعينة هو:

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

حيث تم استبدال تباين العينة S^2 بدلاً من تباين المجتمع σ^2 في الحالة السابقة.

مثال: أعد حل المثال السابق إذا كان σ^2 مجهول وكان تباين العينة $S^2 = 49$.

الحل:

حيث أن تباين المجتمع مجهول فإن توزيع الوسط الحسابي للعينة هو:

$$\bar{X} \sim N\left(30 . \frac{49}{50} \times \frac{200-50}{200-1}\right) = N(30,0.74)$$

وبالتالي

$$Z = \frac{\bar{X} - 30}{\sqrt{0.74}} \sim N(0,1)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{X} > 32) = P\left(Z > \frac{32-30}{\sqrt{0.74}}\right) = P(Z > 2.32)$$

$$= 1 - P(Z < 2.32) = 1 - 0.9898 = 0.0102$$

ملاحظة: عند سحب عينة من مجتمع توزيعه مجهول أو يتبع توزيعا غير التوزيع الطبيعي، فإن توزيع الوسط الحسابي للعينة التي حجمها $n \geq 30$ يقرب إلى التوزيع الطبيعي وذلك حسب نظرية النهاية المركزية، حيث يكون:

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

مثال: تم سحب عينة عشوائية بحجم 36 مشاهدة من مجتمع مجهول التوزيع، يبلغ الوسط الحسابي للعينة 10 بتباين 12 فما هو التوزيع الاحتمالي لمتوسط قيم العينة.

الحل: مباشرة من الملاحظة السابقة يكون :

$$\overline{X} \sim N\left(10.\frac{12}{36}\right)$$

3-توزيع الفرق بين متوسطي عينتين

إذا كان $X_1 \sim N(\mu_1, \sigma_1^2\ )$ وسحبت من مجتمعه عينة حجمها n_1 و $X_2 \sim N(\mu_2, \sigma_2^2)$ وسحبت من مجتمعه عينة حجمها n_2 وكان المجتمعين مستقلين وعرفنا المتغيرين العشوائيين $\overline{X_1}$ و $\overline{X_2}$ فإن : $\overline{X_1} - \overline{X_2}$ يمثل متغيراً عشوائيا للفرق بين متوسطي العينتين، نحصل على الوسط الحسابي والتباين له وذلك كالتالي:

$$E(\overline{X_1} - \overline{X_2}) = E(\overline{X_1}) - E(\overline{X_2}) = \mu_1 - \mu_2$$

$$V(\overline{X_1} - \overline{X_2}) = V(\overline{X_1}) + V(\overline{X_2}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

وسنتحدث الآن عن حالات مختلفة لإيجاد توزيع الفرق بين وسطي العينتين.

وهنا نفرق بين ثلاث حالات:

1.3 عند معلومية σ_1^2 و σ_2^2

$$\overline{X_1} - \overline{X_2} \sim N(\mu_1 - \mu_2.\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

وبالتالي:

$$Z = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

مثال: إذا كان $X_1 \sim N(30,25)$ وسحبت من مجتمعه عينة حجمها 30 مشاهدة و $X_2 \sim N(20,16)$ وسحبت من مجتمعه عينة حجمها 35 مشاهدة أوجد توزيع الفرق بين متوسطي العينتين. ثم أحسب الاحتمال: $P(\overline{X_1} - \overline{X_2} < 12)$.

الحل: بالتطبيق المباشر للعلاقة

$$\overline{X_1} - \overline{X_2} \sim N(\mu_1 - \mu_2 . \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

نحصل على

$$\overline{X_1} - \overline{X_2} \sim N\left(30 - 20 . \frac{25}{30} + \frac{16}{35}\right) = N(10, 1.29)$$

وبالتالي فإن

$$Z = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{1.29}} = \frac{(\overline{X_1} - \overline{X_2}) - 10}{\sqrt{1.29}}$$

فيكون الاحتمال المطلوب هو:

$$P(\overline{X_1} - \overline{X_2} < 12) = P\left(\frac{(\overline{X_1} - \overline{X_2}) - 10}{\sqrt{1.29}} < \frac{12 - 10}{\sqrt{1.29}}\right)$$

$$= P\left(Z < \frac{12 - 10}{\sqrt{1.29}}\right)$$

$$= P(Z < 1.76) = 0.9608$$

2.3عند مجهولية σ_1^2 و σ_2^2 وحجم العينتين كبير أي $n_2, n_1 \geq 30$

$$\overline{X_1} - \overline{X_2} \sim N(\mu_1 - \mu_2 . \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})$$

مثال: معمل ينتج 700 كغ من العجائن كمعدل يومي، سحبت منه عينة عشوائية تمثل إنتاج 40 يوماً من يبلغ وسطها الحسابي 740 كغ بانحراف معياري 40 كغ. معمل آخر ينتج 500 كغ كمعدل يومي، سحبت منه عينة عشوائية تمثل إنتاج 35 يوماً يبلغ وسطها الحسابي 480 كغ بانحراف معياري 20 كغ. فما هو التوزيع الاحتمالي للفرق بين متوسطي العينتين، ثم أحسب الاحتمال التالي:

$$P(180 < \overline{X_1} - \overline{X_2} < 210)$$

الحل:

لدينا من المسألة $S_1^2 = 1600$ و $S_2^2 = 400$ وحيث أن $n_2 = 35, n_1 = 40$ وكليهما أكبر من 30

إذن يكون توزيع الفرق بين وسطي العينتين هو :

$$\overline{X_1} - \overline{X_2} \sim N(700 - 500.\frac{1600}{40} + \frac{400}{35})$$

$$\overline{X_1} - \overline{X_2} \sim N(200, 51.43)$$

وبالتالي:

$$Z = \frac{(\overline{X_1} - \overline{X_2}) - 200}{\sqrt{51.43}} \sim N(0,1)$$

لذلك يكون الاحتمال المطلوب هو:

$$P(180 < \overline{X_1} - \overline{X_2} < 210) = P\left(\frac{180-200}{\sqrt{51.43}} < \frac{(\overline{X_1} - \overline{X_2}) - 200}{\sqrt{51.43}} < \frac{210-200}{\sqrt{51.43}}\right)$$

$$= P(-2.79 < Z < 1.39) = P(Z < 1.39) - P(Z < -2.79)$$

$$= 0.9177 - 0.0029 = 0.9148$$

ملاحظة: التوزيع أعلاه يصلح أيضا عند السحب من مجتمعين غير طبيعيين أو مجهولي التوزيع بشرط كبر حجم العينة، وذلك استنادا لنظرية النهاية المركزية.

3.3 عند مجهولية σ_1^2 و σ_2^2 وحجم العينتين صغير، أي $n_2 < 30, n_1 < 30$

إذا سحبنا من المجتمع الأول الذي يخضع لتوزيع طبيعي وسطه μ_1 وتباينه σ_1^2 عينة صغيرة n_1، وسحبنا من مجتمع طبيعي ثاني وسطه μ_2 وتباينه σ_2^2 عينة n_2 أيضا صغيرة، فإن المتغير العشوائي $\overline{X_1} - \overline{X_2}$ يتبع توزيع ستودينت أي t حسب الكيفية التالية:

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = t_{n_1+n_2-2}$$

مثال: عند سحب عينتين مستقلتين من مجتمعين طبيعيين بحيث $(\sigma_1 = \sigma_2$ و $\mu_1 = \mu_2)$

كانت النتائج على النحو التالي:

$$n_1 = 10;\ n_2 = 14;\ S_1^2 = 16.4;\ S_2^2 = 22.5$$

والمطلوب: أحسب احتمال $P(\overline{X_1} - \overline{X_2}) < 2.5$

الحل:

- بما أن حجم n_1 و n_2 أقل من 30 نستعمل العلاقة التالية لتوزيع t.

$$\frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1S_1^2 + n_2S_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = t_{n_1+n_2-2}$$

وبما أن: $\mu_1 = \mu_2$ أي $\mu_1 - \mu_2 = 0$ ونكتب من جديد :

$$P(\overline{X_1} - \overline{X_2}) > 2.5 = P\left(\frac{(\overline{X_1} - \overline{X_2})}{\sqrt{\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > \frac{2.5}{\sqrt{\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\right)$$

$$= P\left(t_{n_1+n_2-2} > \frac{2.5}{\sqrt{\frac{14.(16.4) + (10).22.5}{22}\left(\frac{1}{14} + \frac{1}{10}\right)}}\right)$$

$$= P(t_{22} > 1.87) = 1 - P(t \leq 1.87)$$

$$=1-0.961$$

$$=0.039$$

4. توزيع النسبة للعينة The sampling distribution of proportion

في المجتمعات الإحصائية التي تتبع توزيع ذي الحدين يتم رصد نسبة الخاصية معينة ندرسها في المجتمع الإحصائي ونرمز لها بالرمز P، مثل نسبة المدخنين ونسبة الإنتاج التالف مصنع معين ...إلخ، حيث يتم حساب النسبة بقسمة مجموع مفردات الخاصية X على حجم المجتمع N أي أن $P=\frac{X}{N}$ فلو أخذنا جميع العينات التي حجمها n من مجتمع حجمه N قد تختلف النسب لهذه العينات، مما يعنى أن

سلوك هذه النسب هو سلوك متغير عشوائي وبالتالي هذا المتغير العشوائي له توزيع احتمالي مستمدا من توزيع المجتمع الذي سحبت منه العينات وسنرمز لنسبة الظاهرة في العينة بالرمز $\widehat{P}$ وبعدد مفردات الخاصية فيها x فتكون نسبة الخاصية في العينة $\widehat{P} = \frac{x}{n}$.

إن هذه النسبة تمثل متغير عشوائي له توزيع احتمالي يسمى توزيع النسبة حيث يكون وسطه الحسابي:

$E(\widehat{P}) = P$

ويكون تباينه هو : $V(\widehat{P}) = \frac{P(1-P)}{n}$

وفي حالة العينات الكبيرة $n \geq 30$ (وهي موضوع دراستنا) يقترب التوزيع للتوزيع الطبيعي وذلك استنادا لنظرية النهاية المركزية ونعبر عن ذلك بالصيغة:

$$Z = \frac{\widehat{P} - P}{\sqrt{V(\widehat{P})}} \sim N(0,1)$$

مثال: مصنع ينتج عادة 25% عبوات كبيرة الحجم، سحبت من إنتاجه عينة حجمها 2200 عبوة تبين أن منها 500 عبوة كبيرة الحجم، أوجد توزيع نسبة العبوات الكبيرة ثم احسب احتمال أن المصنع ينتج أقل من 26% من العبوات الكبيرة في فترة إجراء البحث.

الحل: لاحظ أن المجتمع هنا هو مجتمع ذي الحدين بنسبة نجاح P=0.25، ولاحظ أن حجم العينة كبير وبحساب وسط التوزيع وتباينه نجد أن:

$E(\widehat{P}) = P = 0{,}25$

$$V(\widehat{P}) = \frac{P(1-P)}{n} = \frac{0.25 \times 0.75}{500} = 0.0004$$

$$\widehat{P} \sim N(0.25, 0.0004)$$

$$Z = \frac{\widehat{P} - P}{\sqrt{V(\widehat{P})}} = \frac{\widehat{P} - 0.25}{0.02} \sim N(0,1)$$

$$P(\widehat{P} < 0.26) = P\left(Z < \frac{0.26 - 0.25}{0.02}\right) = P(x < 0.5) = 0.6915$$

ملاحظة: في حالة مجهولية P نعوض بدلا عنها بقيمة $\widehat{P}$.

5.توزيع الفرق بين نسبتين

بفرض أن $X_1 \sim N(N_1, P_1)$ وسحبت من مجتمعه عينة حجمها $n_1 \geq 30$ و $X_2 \sim N(N_2, P_2)$ وسحبت من مجتمعه عينة حجمها $n_2 \geq 30$ وكان المجتمعين مستقلين فإن: $\widehat{P}_1 - \widehat{P}_2$ يمثل متغيرا عشوائيا للفرق بين نسبتي العينتين.

6. توزيع المعاينة لتباين العينة S^2:

نظرية: إذا كانت $(X_1، X_2، X_3 \ldots \ldots X_n)$ عينة عشوائية مسحوبة من مجتمع له توزيع طبيعي وسطه μ وتباينه σ^2، وكانت S^2 تباين العينة فإن المتغير K^2 حيث: $K^2 = \frac{(n-1)S^2}{\sigma^2}$

له توزيع كاي تربيع بدرجات حرية $v = n - 1$

مثال: تقوم آلة أوتوماتيكية بقطع أسلاك معدنية لإستخدامها في صنع قضبان التلحيم، نعتبر أن أطوال القضبان المنتجة هي متغير عشوائي يتبع التوزيع الطبيعي بمتوسط قدره 30 سم وانحراف معياري 0.5 سم. إذا سحبنا عينة من 25 قضيب من انتاج هذه الآلة. ما هو احتمال أن يكون تباينها أكبر من 0.38 سم، **الحل:**

$$\mathbf{p}(\mathbf{s^2} > 0.38) = ?$$

نعلم أن:

$$\frac{(\mathbf{n-1}).\mathbf{s^2}}{\boldsymbol{\sigma^2}} \to \mathbf{K^2_{n-1}}$$

و منه $\mathbf{p}(\mathbf{s^2} > 0.38) = \mathbf{p}\left(\frac{(\mathbf{n-1}).\mathbf{s^2}}{\boldsymbol{\sigma^2}} > \frac{\mathbf{24.0,38}}{\mathbf{0,5^2}}\right)$

$$= \mathbf{p}(\mathbf{K^2_{24}} > 38) = \mathbf{1} - \mathbf{p}(\mathbf{K^2_{24}} \leq \mathbf{38})$$

من جدول كاي تربيع نجد:

القيمة الجدولية 38 محصورة بين القيمتين 36.4 التي تقابل الاحتمال 0.95 و 39.4 التي تقابل الاحتمال 0.975، وبعد الحساب الاستكمالي نجد أن الاحتمال المقابل للقيمة الجدولية 38 هو 0.963 وهذا طبعا عند درجة حرية 24. وبذلك يكون:

$$\mathbf{p}(\mathbf{s^2} > 0.38) = \mathbf{1} - \mathbf{0.963} = \mathbf{0.037}$$

7. توزيع النسبة بين تباين عينتين:

يعتبر هذا التوزيع من التوزيعات الهامة التي تبحث في تجانس المجتمعات ونلجأ لحساب النسب بين التباينات وليس الفرق بينها لسهولة دراسة النسب وتفسيرها.

فإذا سحبت عينة حجمها n_1 وتباينها S_1^2 من مجتمع $N(\mu_1, \sigma_1^2)$ وعينة أخرى حجمها n_2 وتباينها S_2^2 من مجتمع $N(\mu_2, \sigma_2^2)$ مستقل عن المجتمع الأول فإن:

$$\frac{(n_1-1)s_1^2}{\sigma_1^2} \sim X^2_{(n_1-1)} \quad \text{،} \quad \frac{(n_1-1)s_2^2}{\sigma_2^2} \sim X^2_{(n_2-1)}$$

وبالتالي:

$$\frac{\frac{(n_1-1)s_1^2}{\sigma_1^2} \Big/ (n_1-1)}{\frac{(n_2-1)s_2^2}{\sigma_2^2} \Big/ (n_2-1)} \sim F(n_1-1, n_2-2)$$

بعد تبسيط الطرف الأيسر نحصل على

$$F = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \sim F(n_1-1, n_2-2)$$

وإذا تساوى تبايني المجتمعين فإن:

$$\frac{s_1^2}{s_2^2} \sim F(n_1-1, n_2-2)$$

مثال: سحبت عينة حجمها 13 من مجتمع طبيعي تباينه 9، وسحبت عينة أخرى حجمها 21من مجتمع طبيعي تباينه 25 مستقل عن المجتمع الأول. أوجد احتمال النسبة بين تبايني العينتين أقل من 0.8.

الحل: حيث أن:

$$F = \frac{s_1^2/9}{s_2^2/25} \sim F(12,20)$$

إذن

$$P\left(\frac{s_1^2}{s_2^2} < 0.8\right) = P\left(\frac{s_1^2/9}{s_2^2/25} < 0.8\left(\frac{25}{9}\right)\right)$$

$$= P(F < 2.22) = 1 - P(F > 2.22) = 1 - 0.05 = 0.95$$

تمارين السلسلة الثانية

أسئلة نظرية:

1- ما المقصود بالمجتمع الإحصائي، العينة و المعاينة، وما هي أهم أنواعها؟

2- ما هي الخطوات الأساسية لتصميم العينة؟

3- ما هي العلاقة بين أخطاء المعاينة العشوائية وكل من حجم العينة و تباين المجتمع؟

4- ما هي معايير جودة التقدير مع الشرح كل معيار؟

التمرين الأول: إذا أردنا أن نأخذ عينة منتظمة من طلاب جامعة سطيف الذين يستعملون النقل الجامعي، من أجل التعرف على أسباب استعمال هذا النوع و مدى رضا الطلبة عن وسائل النقل الممنوحة لهم، إذا كان حجم المجتمع في هذه الدراسة هو 40000 طالب و نرغب في إختبار عينة حجمها 100 طالب، فكيف يتم ذلك.

التمرين الثاني: ليكن مجتمعا مؤلفا من 12 دخل قيمة كل منها موضحة في الجدول الموالي:

i	1	2	3	4	5	6	7	8	9	10	11	12
D(i)	100	1100	600	150	1000	620	1050	680	520	170	130	650

- نريد أن نقدر متوسط الدخل من خلال سحب عينة طبقية مكونة من 9 مداخيل فكيف يتم ذلك؟
- ما هي أهم المميزات الخاصة بالعينة الطبقية؟

التمرين الثالث: لدينا مجتمع إحصائي حجمه N=4 مكون من القيم التالية: 3،5،7،8

1- أحسب المتوسط الحسابي للمجتمع μ وتباينه V(X) ؟

2- إستخرج جميع العينات العشوائية ذات الحجم n=2 الممكن سحبها مع الإعادة.

3- أحسب المتوسط الحسابي في كل عينة.

4- أثبت أن التوقع الرياضي لمتوسطات هذه العينات يساوي وسط المجتمع.

5- أجب عن جميع الأسئلة في حالة السحب دون إرجاع.

التمرين الرابع: بفرض أن أحد المجتمعات يتكون من 4 أفراد حسب العمر $\{16 ، 18 ، 32 ، 54\}$

أ- أوجد وسط و تباين المجتمع .

ب- عند سحب عنصرين من المجتمع أحسب $E(\overline{X})$ و $V(\overline{X})$ إذا كانت العينة عشوائية بسيطة مع الإعادة، بدون إعادة، و العينة الطبقية.

ت- قارن بين كل هذه النتائج؟

التمرين الخامس: ليكن لدينا مجتمع من 6 أسر في مكان ما، دخل كل أسرة كالآتي: 38 / 39 / 43 42 / 41 / 40

أ- حدد جميع العينات الممكن سحبها بحجم 3 عناصر في العينة الواحدة و ذلك بدون إعادة.

ب- أحسب متوسط الدخل في كل عينة.

ت- برهن على أن التوقع الرياضي لمتوسطات هذه العينات يساوي وسط المجتمع.

التمرين السادس: الإنحراف المعياري لمجتمع حجمه 200000 وحدة يساوي 800، أما وسطه الحسابي 30 سحبت منه عينة عشوائية حجمها 400.

المطلوب:

أ- أحسب الوسط الحسابي و الإنحراف المعياري لتوزيع المعاينة في حالة السحب مع الإرجاع؟

ب- أوجد احتمال أن يكون الوسط الحسابي للعينة أكبر من أو يساوي 75؟

التمرين السابع: مصنع ينتج كراسي ترتكز على قاعدة دائرية، إعتمادا على التجارب السابقة فإن مفتش الرقابة على العملية الإنتاجية مقتنع أن متوسط قطر الدائرة 5 سم و الإنحراف المعياري له 0.005 سم و توزيع العملية الإنتاجية هو التوزيع الطبيعي، يهتم الفاحص بالمحافظة على متوسط قطر العملية الإنتاجية عند 5 سم، ولتحقيق ذلك تسحب عينات عشوائية بصفة دورية حجم كل منها 9 كراسي و ذلك في محاولة لإكتشاف أية انحراف على الأرقام الطبيعية المشار إليها.

المطلوب:

أ- حدد توزيع المعاينة لـ: $\overline{X}$؟

ب- بفرض أن الفاحص سحب عينة عشوائية من 9 كراسي و قيست أقطار قاعدتها ووجد أن سم$\overline{X} = 5.004$ ، ما هي إمكانية (إحتمال) أن متوسط القطر في تلك العينة العشوائية سيكون على الأقل 5.004 سم على فرض أن متوسط العملية باقي عند 5 سم والإنحراف المعياري هو 0.005 سم؟

ت- ما هو حجم العينة التي يجب سحبها لتحقيق خطأ معياري لـ: $\overline{X}$ يساوي 0.001.

التمرين الثامن: بفرض أن n=16، أوجد القيمة الجزئية T و التي لها الإحتمالات 0.95 و 0.025 على التوالي.

التمرين التاسع: كمية المبيعات اليومية بالوحدات في المتوسط تساوي 210 وحدة/يوم ، وبإنحراف معياري 30 وحدة/يوم، خلال فترة مدتها 36 يوم.

المطلوب: ما هو إحتمال أن المبيعات اليومية في المتوسط تساوي 200 وحدة أو أكثر؟

التمرين العاشر: مقاول بنايات كبيرة قرر شراء كميات من مصابيح الإضاءة عالية القوة من صاحب مصنع معين، صاحب المصنع أكد للمقاول أن هذه المصابيح لها متوسط عمر 1000 ساعة بإنحراف معياري 80 ساعة، المقاول قرر شراء مصابيح صاحب المصنع إذا كان متوسط العمر لعينة عشوائية 64 من المصابيح 1010 ساعة على الأقل.

المطلوب: في ظل هذا الشرط، ما هو إحتمال أن المقاول سوف يشتري هذه المصابيح من هذا المصنع؟

التمرين الحادي عشر: في مصنع لإنتاج الخيوط، متوسط إنتاج الخيوط فيها 82.5 أخذت عينة عشوائية مكونة من 25 خيطا، فوجد أن الإنحراف المعياري لها هو 15

المطلوب: ما هو إحتمال أن يكون متوسط قوة الخيوط في العينة على الأقل 90؟

التمرين الثاني عشر: إذا كانت معدلات الطلبة في السنة الأولى في إحدى الجامعات تتبع التوزيع الطبيعي بمتوسط 60 أخذت عينة عشوائية مؤلفة من 16 طالب فوجد أن الإنحراف المعياري لها هو 12.

المطلوب: أوجد إحتمال أن يزيد متوسط معدل الطلبة سنة أولى في إحدى الجامعات في العينة عن 66.3؟

التمرين الثالث عشر: إذا كان متوسط العمر الإنتاجي لمصابيح كهربائية ينتجها المصنع A هو 1400 ساعة و إنحرافها المعياري هو 200 ساعة، بينما التي ينتجها المصنع B فمتوسط عمرها الإنتاجي هو 1200 ساعة وإنحرافها المعياري هو 100 ساعة، إذا سحبت عينة عشوائية حجمها 125 مصباح من كل مصنع.

المطلوب: أوجد إحتمال أن يزيد متوسط العمر الإنتاجي لمصابيح المصنع A عن متوسط العمر الإنتاجي لمصابيح المصنع B بمقدار 250 ساعة؟

التمرين الرابع عشر: تنتج شركة (أ) دواء السعال الخاص بالأطفال، متوسط فترة صلاحيته هو 910 ساعة، أما شركة (ب) فتنتج دواء السعال الخاص بالأطفال متوسط فترة صلاحيته هو 925 ساعة. إذا تم سحب عينة عشوائية حجمها 16 عبوة من إنتاج الشركة (أ) بإنحراف معياري هو 8 ساعات ، ثم بعد ذلك أخذت عينة عشوائية أخرى مستقلة عن الأولى حجمها 10 عبوات من إنتاج الشركة (ب) فكان إنحرافها المعيار هو 15 ساعة. إذا كان تبايني المجتمعين مجهولين ومتساويين.

المطلوب: أوجد إحتمال أن يكون متوسط فترة صلاحية دواء شركة (أ) أكبر من متوسط فترة صلاحية دواء شركة (ب).

التمرين الخامس عشر: إذا تم سحب عينتان عشوائيتان مستقلتان حجمهما 15 و 12 وتباينهما 45 و 72 على التوالي لمجتمعين إحصائيين متوسطهما 38 و 33 عل التوالي.

المطلوب: أوجد إحتمال أن يكون الفرق بين هاتين العينتين 10 أو أقل؟

التمرين السادس عشر: مجتمع إحصائي معين يتوزع توزيعا طبيعيا إنحرافه المعياري 5، سحبت منه عينة عشوائية حجمها 4 وحدات.

المطلوب: أوجد إحتمال أن تباين العينة سوف يكون 2.5 أو أقل؟

التمرين السابع عشر: أخذت عينة عشوائية حجمها 11 من مجتمع إحصائي معين، و أخذت عينة عشوائية أخرى مستقلة عن الأولى حجمها 17 من مجتمع آخر مستقل عن الأول.

المطلوب: أوجد إحتمال أن نسبة تباين العينتين يكون على أقل 3.69؟

التمرين الثامن عشر: يدرس في إحدى المدارس 600 تلميذا وتلميذة منهم 240 ذكور فإذا سحبنا عينة عشوائية من هذه المدرسة تشمل 55 تلميذا وتلميذة.

المطلوب: ما هو إحتمال أن تكون نسبة الذكور في هذه المدرسة أكثر من 50% ؟

التمرين التاسع عشر: إذا كان 17% من منتج معين غير صالح للاستعمال، اشترى زبون 230 وحدة من هذا المنتج.

المطلوب: ما هو إحتمال أن 15% أو أكثر من الوحدات المشترات في العينة تكون غير صالحة؟

التمرين العشرون: إذا كانت نسبة النجاح من الطالبات في امتحان اللغة العربية 64% و نسبة الناجحين من الطلبة الذكور في نفس الإمتحان 60%. و تم إختيار عينتين مستقلتين الأولى حجمها 120 طالبة و الثانية حجمها 150 طالبا من الذين اشتركوا في هذا الإمتحان.

المطلوب: ما هو إحتمال أن تكون نسبة الناجحات في عينة الطالبات أكبر من نسبة الناجحين في عينة الطلبة بمقدار 6% أو أكثر؟

الفصل الثالث
تقدير معالم المجتمع

يقصد بمعالم المجتمع مجموعة من خصائصه كالمتوسط والتباين (أو الانحراف المعياري) والنسبة حيث يتم حساب هذه المعالم عند القيام بالحصر الشامل بشكل دقيق. ولكن غالبا ما تكون هذه المعالم مجهولة لذلك نقوم بتقديرها من بيانات العينة تقديرا لنظيرتها من معالم المجتمع، حيث يتم:

*تقدير المتوسط الحسابي للمجتمع μ انطلاقا من متوسط العينة $\bar{x}$ ؛

*تقدير الانحراف المعياري للمجتمع σ انطلاقا من الانحراف المعياري للعينة S؛

*تقدير نسبة المجتمع p انطلاقا من نسبة العينة f.

والمقصود بالتقدير هو تقدير معالم المجتمع الاحصائي (أو التوزيع الاحتمالي) والتي غالبا ما تكون مجهولة ويكون المطلوب هو الحصول على تقديرات لها من بيانات العينة فقد يكون المطلوب تقدير متوسط دخل الدولة، أو تقدير متوسط عمر جهاز ما ...الخ.

مع العلم أن التقدير الإحصائي نوعان:

1-التقدير النقطي؛

2-التقدير بمجال.

1- التقدير النقطي

عندما نقدر معلمة المجتمع بقيمة واحدة تحسب من بيانات العينة فإن هذه القيمة تسمى **تقدير نقطي** لهذه المعلمة والتي تحسب اعتمادا على **المقدر النقطي** لهذه المعلمة وهو الدالة أو الصيغة الرياضية التي نقدر بها معلمة المجتمع، بينما التقدير هو القيمة العددية.

مثال: إذا كانت المعلمة المجهولة هي متوسط المجتمع μ فإن الإحصائية $\bar{x}$ (متوسط العينة) هي مقدر نقطي لـ μ ونكتب:

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

إذا سحبنا عينة عشوائية من هذا المجتمع وحسبنا متوسطها الحسابي ووجدنا أن:

$$\bar{x} = 27$$

فإن القيمة 27 تعتبر تقدير نقطي لمتوسط المجتمع ونكتب:

$$\hat{\mu} = 27$$

μ: معلمة المجتمع؛

$\overline{x}$: المقدر النقطي؛

27: التقدير النقطي.

1-1 خصائص المقدر الجيد(معايير اختبار المقدرات النقطية):

غالبا ما نجد أكثر من مقدر يمكن استخدام قيمته كتقدير لمعلمة المجتمع المجهولة، لذلك نحتاج إلى معايير يمكن الاستعانة بها لإختيار المقدر الذي يعتبر أفضل من غيره لتقدير معلمة المجتمع، وتتلخص هذه المعايير في ثلاثة معايير أساسية هي:

أ-عدم التحيز:

نقول عن المقدر $\hat{\theta}$ أنه غير متحيز للمعلمة θ إذا كان التوقع الرياضي لـ $\hat{\theta}$ يساوي المقدار نفسه ونكتب رياضيا:

$$\mathbf{E(\hat{\theta}) = \theta}$$

أما إذا كان:

$$\mathbf{E(\hat{\theta}) \neq \theta}$$

فإننا نقول عن $\hat{\theta}$ أنه مقدر متحيز للمعلمة θ أي أنه لا يمثل المعلمة تمثيلا صحيحا ومقدار التحيز هو b حيث:

$$\mathbf{b = E\left(\hat{\theta}\right) - \theta}$$

مثال:

نفرض أن x متغير عشوائي حيث:

$$E(x_i) = \mu \quad و \quad V(x_i) = \sigma^2$$

يكون $\overline{x}$ مقدر غير متحيز لـ μ إذا كان:

$$\mathbf{E(\overline{x}) = \mu}$$

لدينا:

$$\overline{x} = \frac{\sum x_i}{n}$$

ومنه:

$$E(\overline{x}) = E\left(\frac{\sum x_i}{n}\right) = \frac{1}{n}.\sum E(x_i) \ldots\ldots\ldots..(1)$$

كما نعلم أن:

$$E(x_i) = \mu$$

وعليه تصبح العلاقة (1) كما يلي:

$$E(\overline{x}) = \frac{1}{n}.n.\mu = \mu$$

أي أن:

$$\mathbf{E(\overline{x}) = \mu}$$

وبذلك يمكن القول أن وسط العينة $\overline{x}$ هو مقدر **غير متحيز** لوسط المجتمع μ.

ب- التقارب(التماسك):

نقول عن المقدر $\hat{\theta}$ أنه متقارب إذا كانت قيمته تؤول إلى قيمة المعلمة

إلى ما لا نهاية أو إلى حجم n كلما زاد حجم العينة، أي أن تباينه يؤول إلى الصفر عندما تؤول θ أي:Nالمجتمع

$$\left\{\boldsymbol{\hat{\theta}} \xrightarrow[\mathbf{n \to \infty (N)}]{} \boldsymbol{\theta}\right.$$

أي:

$$\left\{\mathbf{V(\hat{\theta})} \xrightarrow[\mathbf{n \to \infty (N)}]{} \mathbf{0}\right.$$

مثال: يكون $\bar{x}$ مقدر مقارب لـ μ أو متماسك إذا كان:

$$\left\{\bar{x} \xrightarrow[n\to\infty(N)]{} \mu\right.$$

أي:

$$\left\{V(\bar{x}) \xrightarrow[n\to\infty(N)]{} 0\right.$$

لدينا:

$$V(\bar{x}) = V\left(\frac{\sum x_i}{n}\right) = \frac{1}{n^2}.n.V(x_i) \dots\dots\dots (2)$$

ونعلم أن:

$$V(x_i) = \sigma^2$$

وعليه تصبح العلاقة (2) كما يلي:

$$V(\bar{x}) = \frac{1}{n^2}.n.\sigma^2 = \boldsymbol{\frac{\sigma^2}{n}}$$

نلاحظ أن المقدار:

$$\boldsymbol{\frac{\sigma^2}{n}}$$

يؤول فعلا إلى الصفر عندما يؤول n إلى ∞ (N) ومنه يمكن القول أن متوسط العينة هو مقدر **مقارب** لـ وسط المجتمع μ أي أنه **متماسك**.

ج–الفعالية (الكفاءة):

إذا كان $\hat{\theta}_1$ و $\hat{\theta}_2$ مقدرين غير متحيزين للمعلمة θ وكان تباين $\hat{\theta}_1$ أقل من تباين $\hat{\theta}_2$ أي:

$$V(\widehat{\theta 1}) < V(\widehat{\theta 2})$$

فإننا نقول أن المقدر $\hat{\theta}_1$ أكثر كفاءة من المقدر $\hat{\theta}_2$ أي أن المقدر الفعال هو المقدر الأقل تباين.

مثال: ليكن لدينا x_1، x_2، x_3 عينة عشوائية مسحوبة من مجتمع غير محدود حيث:

$$E(x_i) = \mu \quad \text{و} \quad V(x_i) = \sigma^2$$

المطلوب: أثبت أن المقدر المعطى بالعلاقة التالية:

$$\hat{\mu} = \frac{x_1 + 2x_2 + 3x_3}{6}$$

هو مقدر غير متحيز لـ μ إلا أنه أقل كفاءة من المقدر المعطى بالعلاقة:

$$\bar{x} = \frac{x_1 + x_2 + x_3}{3}$$

الحل:

1- يكون $\hat{\mu}$ مقدر غير متحيز لـ μ إذا كان:

$$E(\hat{\mu}) = \mu$$

$$E(\hat{\mu}) = E(\frac{x_1 + 2x_2 + 3x_3}{6})$$

$$= \frac{1}{6}E(x_1 + x_2 + x_3)$$

$$= \frac{1}{6}[E(x_1) + 2E(x_2) + 3E(x_3)]$$

$$= \frac{1}{6}(\mu + 2\mu + 3\mu) = \frac{6\mu}{6} = \mu$$

أي:

$$\mathbf{E(\hat{\mu}) = \mu}$$

ومنه $\hat{\mu}$ هو فعلا مقدر **غير متحيز** لـ μ.

2- يكون $\hat{\mu}$ أقل كفاءة من $\bar{x}$ إذا كان:

$$\mathbf{V(\hat{\mu})} > V(\bar{\mathbf{x}})$$

لدينا:

$$V(\hat{\mu}) = V\left(\frac{x_1 + 2x_2 + 3x_3}{6}\right)$$

$$= \frac{1}{36}[V(x_1) + 4V(x_2) + 9V(x_3)]$$

$$= \frac{1}{36}[\sigma^2 + 4\sigma^2 + 9\sigma^2] = \frac{14\sigma^2}{36} = \mathbf{0,38\sigma^2}$$

$$V(\bar{x}) = V\left(\frac{x_1 + x_2 + x_3}{3}\right)$$

$$= \frac{1}{9}[V(x_1) + V(x_2) + V(x_3)]$$

$$= \frac{1}{9}[\sigma^2 + \sigma^2 + \sigma^2] = \frac{3\sigma^2}{9} = \mathbf{0,33\sigma^2}$$

نلاحظ أن:

$$V(\hat{\mu}) > V(\bar{x})$$

وعليه فإن $\hat{\mu}$ أقل كفاءة (فعالية) من $\bar{x}$.

وبما أن فعالية المقدرات تزيد كلما كان **تباينها أقل** فإن فعالية $\bar{x}$ تبقى مرتبطة بحجم العينة n حيث كلما زادت قيمة n كان $\bar{x}$ أكثر فعالية.

كما يعتبر $\bar{x}$ المقدر الأكثر فعالية من المقدرات الأخرى لـ μ (الوسيط والمنوال)لأنه الأقل تباين.

ملاحظة1: في التطبيقات العملية عادة ما تكون قيمة **σ^2 مجهولة** وبالتالي يتم الاعتماد على تباين العينة وفق العلاقة التالية:

$$V(\bar{x}) = \frac{\sigma^2}{n} = \frac{\mathbf{s^2}}{\mathbf{n-1}}$$

حيث يعطى تباين العينة بالعلاقة:

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

ملاحظة2: لما يكون السحب مع الإرجاع فإنه لا يهم إذا كان المجتمع محدود أو غير محدود.

ب–حالة عينة عشوائية بدون إعادة: (ESSR)

إذا تم السحب بدون إرجاع من مجتمع محدود فإن:

***عدم التحيز:**

في هذه الحالة أيضا فإن العلاقة:

$$\mathbf{E(\bar{x}) = \mu}$$

محققة مما يعني أن $\bar{x}$ مقدر غير متحيز لـ μ .

***التقارب (التماسك):**

يكون $\bar{x}$ مقدر مقارب لـ μ أو متماسك إذا كان:

$$\left\{\bar{x} \xrightarrow[n\to\infty(N)]{} \mu\right.$$

أي:

$$\left\{V(\bar{x}) \xrightarrow[n\to\infty(N)]{} 0\right.$$

وبما أن السحب هنا بدون إرجاع فإن:

$$\left\{V(\bar{x}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}\right.$$

وهذا المقدار واضح أنه يؤول إلى الصفر كلما اقترب حجم العينة n من حجم المجتمع N (∞).

ومنه فإن $\bar{x}$ مقدر **مقارب لـ μ (متماسك)**.

***الفعالية:**

لدينا:

$$V(\bar{x}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

وبالتالي فإن فعالية $\bar{x}$ تبقى مرتبطة بحجم العينة n حيث كلما زادت قيمة n كان $\bar{x}$ أكثر فعالية.

ملاحظة 1: نعلم أنه عادة ما تكون قيمة σ^2 **مجهولة** وبالتالي نعتمد على تباين العينة وفق العلاقة التالية:

$$V(\bar{x}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1} = \frac{\mathbf{s^2}}{\mathbf{n-1}} \cdot \frac{\mathbf{N-n}}{\mathbf{N-1}}$$

ملاحظة 2: نلاحظ أن:

$$\frac{\sigma^2}{n} \cdot \frac{N-n}{N-1} < \frac{\sigma^2}{n}$$

أي:

$$V(\bar{x})_{ESSR} < V(\bar{x})_{ESAR}$$

ومنه فإن $\bar{x}$ في حالة السحب بدون إعادة أكثر فعالية من $\bar{x}$ في حالة السحب مع الإعادة. بمعنى أن الخطأ في اختيار العينة يكون أقل عند السحب بدون إعادة منه عند السحب مع الإعادة.

ملاحظة 3: في حالة N كبير فإن المقدار :

$$\frac{N-n}{N-1}$$

يصبح يساوي تقريبا:

$$\frac{N-n}{N}$$

أي:

$$\frac{\mathbf{N-n}}{\mathbf{N-1}} = \frac{\mathbf{N-n}}{\mathbf{N}} = \mathbf{1} - \frac{\mathbf{n}}{\mathbf{N}} = \mathbf{1-f}$$

حيث f هو معدل العينة.

2- بناء مجالات الثقة "التقدير بمجال"

بما أننا لا نتوقع أن تكون قيمة المقدر مساوية للمعلمة المجهولة مهما كان المقدر جيد فيتم اللجوء إلى تحديد مجال يحتوي على مجموعة من القيم تتضمن فيما بينها قيمة معلمة المجتمع المجهولة θ يسمى هذا المجال **بمجال الثقة أو فترة الثقة**. حيث نسبة الحظوظ أو احتمال أن تقع معلمة المجتمع θ في هذا المجال تسمى درجة الثقة أو مستوى الثقة وبذلك نكون قد **قدرنا معلمة المجتمع بمجال** ونكتب:

$$IC(\theta)_{1-\alpha} = [T_1; T_2]$$

أي:

$$p(\theta\epsilon[T_1;T_2]) = 1-\alpha$$

حيث:

$T_2; T_1$:هما على التوالي الحد الأدنى والحد الأعلى لمجـــــــال الثقة، مع :

$$T_1 < T_2$$

$1-\alpha$: مستوى الثقة أو درجة الثقة؛

α : احتمال الخطأ أو مستوى المعنوية أو درجة المخاطرة والذي يعبر عن احتمال عدم احتواء مجال الثقة على القيمة الحقيقية لـθوالتي عادة ما تكون من اختيار الباحث حسب درجة الدقة التي يبحث عنها.

1.2 مجال الثقة للمتوسط الحسابي للمجتمع μ:

1.1.2 إذا كان حجم العينة كبير $(n \geq 30)$:

نعلم أنه في هذه الحالة $\bar{x}$ يتبع تقريبا التوزيع الطبيعي حيث:

$$Z = \frac{\bar{x}-E(\bar{x})}{\sqrt{V(\bar{x})}} \rightarrow N(0;1)$$

لبناء مجال الثقة لـ μ ممكن أن ننطلق من مفهوم مجال الثقة كما يلي:

$$p(\mu\,\epsilon\,IC) = 1-\alpha$$

$$\Leftrightarrow p(\mu\,\epsilon[T_1;T_2]) = 1-\alpha$$

$$\Leftrightarrow p(T_1 \leq \mu \leq T_2) = 1-\alpha$$

ينبغي إذن تحديد حدود مجال الثقة T_1و T_2 إنطلاقا من حصر Z بين قيمتين حيث:

$$Z = \frac{\bar{x}-\mu}{\sqrt{V(\bar{x})}}$$

(لأن$\mu = E(\bar{x})$ في كل الحالات).

لدينا Z محصور بين $-z_{1-\frac{\alpha}{2}}$ و $z_{\frac{\alpha}{2}}$ ومنه:

$$p\left(z_{\frac{\alpha}{2}} \leq z \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

ومن خاصية التناظر نستنتج أن:

$$z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$$

ومنه يصبح لدينا:

$$p\left(-z_{1-\frac{\alpha}{2}} \leq z \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

وبتعويض قيمة z بما تساويه نجد:

$$p\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu}{\sqrt{V(\bar{x})}} \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

وبضرب أطراف المتراجحة في $\sqrt{V(\bar{x})}$نجد:

$$p\left(-z_{1-\frac{\alpha}{2}}.\sqrt{V(\bar{x})} \leq \bar{x}-\mu \leq z_{1-\frac{\alpha}{2}}.\sqrt{V\bar{x}}\right) = 1-\alpha$$

من أطراف المتراجحة نجد:$\bar{x}$ ثم بطرح قيمة

$p\left(-\bar{x}-z_{1-\frac{\alpha}{2}}.\sqrt{V(\bar{x})} \leq -\mu \leq -\bar{x}+z_{1-\frac{\alpha}{2}}.\sqrt{V\bar{x}}\right) = 1-\alpha$ نضرب الآن أطراف المتراجحة في -1 فيتغير اتجاهها وتصبح كما يلي:

$$p\left(\bar{x}-z_{1-\frac{\alpha}{2}}.\sqrt{V(\bar{x})} \leq \mu \leq \bar{x}+z_{1-\frac{\alpha}{2}}.\sqrt{V(\bar{x})}\right) = 1-\alpha$$

هذا يعني أن مجال الثقة لـ μ هو:

$$\mathbf{IC(\mu)_{(1-\alpha)\%} = \left[\bar{x}-z_{1-\frac{\alpha}{2}}.\sqrt{V(\bar{x})}\,;\bar{x}+z_{1-\frac{\alpha}{2}}.\sqrt{V(\bar{x})}\right]}$$

حيث:

معلوم فإن: σ^2إذا كان

$$IC(\mu)_{(1-\alpha)\%} = \left[\bar{x}-z_{1-\frac{\alpha}{2}}.\frac{\boldsymbol{\sigma}}{\sqrt{\mathbf{n}}}\,;\bar{x}+z_{1-\frac{\alpha}{2}}.\frac{\boldsymbol{\sigma}}{\sqrt{\mathbf{n}}}\right]$$

وإذا كان σ^2 مجهول فإن:

$$IC(\mu)_{(1-\alpha)\%} = \left[\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\mathbf{s}}{\sqrt{\mathbf{n-1}}} ; \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\mathbf{s}}{\sqrt{\mathbf{n-1}}}\right]$$

$$IC(\mu)_{(1-\alpha)\%} = \left[\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\mathbf{s}}}{\sqrt{\mathbf{n}}} ; \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\mathbf{s}}}{\sqrt{\mathbf{n}}}\right]$$

ملاحظة: يبين الجدول التالي درجات الثقة الشائعة واحتمالات الخطأ وقيم$z_{1-\frac{\alpha}{2}}$الموافقة لها:

مستوى الثقة $\mathbf{1-\alpha}$	**0.90** **90%**	**0.95** **95%**	**0.99** **99%**
احتمال الخطأ α	10% =0.10	0.05 = 5%	0.01 = 1%
$\mathbf{z_{1-\frac{\alpha}{2}}}$	1.645	1.96	2.58
أي	$z_{0.95} = 1.645$	$z_{0.975} = 1.96$	$z_{0.995} = 2.58$

مثال: إذا كانت أعمار المصابيح الكهربائية التي تنتجها أحد المصانع تتبع التوزيع الطبيعي بانحراف معياري قدره 250 ساعة. اختيرت عينة من انتاج هذا المصنع حجمها 100 مصباح وكان متوسط عمر المصباح في هذه العينة هو 1200 ساعة.

- قدر بمجال (حدد مجال الثقة لـ) متوسط عمر المصباح في المصنع كله بدرجة ثقة **95%**.

الحل: لدينا:

$$\mu = ? ; \quad (1-\alpha) = 95\%; \quad \bar{x} = 1200; \quad \sigma^2 = 250^2$$

$$n = 100 \geq 30$$

أي أن $\bar{x}$ يتبع التوزيع الطبيعي، ومنه مجال الثقة لـ μ يعطى بالصيغة التالية:

$$\mathbf{IC(\mu)_{(95)\%}} = \left[\bar{\mathbf{x}} - \mathbf{z}_{1-\frac{\alpha}{2}} \cdot \frac{\boldsymbol{\sigma}}{\sqrt{\mathbf{n}}} ; \bar{\mathbf{x}} + \mathbf{z}_{1-\frac{\alpha}{2}} \cdot \frac{\boldsymbol{\sigma}}{\sqrt{\mathbf{n}}}\right]$$

حيث:

$$\frac{\sigma}{\sqrt{n}} = \frac{250}{\sqrt{100}} = 25\,; \quad Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$$

ومنه:

$$IC(\mu)_{95\%} = [1200 - 1{,}96 * 25; 1200 + 1{,}96 * 25]$$

$$= [\mathbf{1151}; \mathbf{1249}]$$

أي أن متوسط عمر المصباح من انتاج المصنع كله يتراوح بين 1151 و 1249 ساعة بمستوى ثقة 95%.

2.1.2 إذا كان حجم العينة صغير $n < 30$:

نعلم أنه إذا كان $n < 30$ و $\mathbf{x}$ يتوزع طبيعيا مع $\boldsymbol{\sigma^2}$ مجهول فإن :

$$\mathbf{T} = \frac{\bar{\mathbf{x}} - \mathbf{E}(\bar{\mathbf{x}})}{\sqrt{\mathbf{V}(\bar{\mathbf{x}})}} \to \mathbf{t_{n-1}}$$

وبذلك يكون مجال الثقة لـ μ في هذه الحالة كما يلي:

$$\mathbf{IC(\mu)_{(1-\alpha)\%}} = \left[\bar{\mathbf{x}} - \mathbf{t}_{\mathbf{1-\frac{\alpha}{2};n-1}} \cdot \frac{\mathbf{s}}{\sqrt{\mathbf{n-1}}}\,; \bar{\mathbf{x}} + \mathbf{t}_{\mathbf{1-\frac{\alpha}{2};n-1}} \cdot \frac{\mathbf{s}}{\sqrt{\mathbf{n-1}}}\right]$$

$$\mathbf{IC(\mu)_{(1-\alpha)\%}} = \left[\bar{\mathbf{x}} - \mathbf{t}_{\mathbf{1-\frac{\alpha}{2};n-1}} \cdot \frac{\hat{\mathbf{s}}}{\sqrt{\mathbf{n}}}\,; \bar{\mathbf{x}} + \mathbf{t}_{\mathbf{1-\frac{\alpha}{2};n-1}} \cdot \frac{\hat{\mathbf{s}}}{\sqrt{\mathbf{n}}}\right]$$

مثال: إذا كانت أجور العمال لإحدى الشركات تتبع التوزيع الطبيعي وتم اختيار عينة من 25 عامل فوجد أن متوسط أجرهم الشهري 4250 ون والانحراف المعياري 400 ون.

- أوجد مجال الثقة (قدر بمجال)لمتوسط الأجر الشهري لعمال المصنع كله بدرجة ثقة 95%.

الحل: لدينا:

$$\mu = ?\,; \ \ (1-\alpha) = 95\%; \ \ \bar{x} = 4250; S = 400$$

$$n = 25 < 30; \quad x \to N(\mu, \sigma^2); \quad \sigma^2 = ?\left(\text{مجهول}\right)$$

في هذه الحالة نستخدم توزيع ستيودنت ويكون مجال الثقة لـ μ كما يلي:

$$\mathbf{IC(\mu)_{(95)\%}} = \left[\bar{\mathbf{x}} - \mathbf{t}_{\mathbf{1-\frac{\alpha}{2};n-1}} \cdot \frac{\mathbf{s}}{\sqrt{\mathbf{n-1}}}\,; \bar{\mathbf{x}} + \mathbf{t}_{\mathbf{1-\frac{\alpha}{2};n-1}} \cdot \frac{\mathbf{s}}{\sqrt{\mathbf{n-1}}}\right]$$

$$t_{1-\frac{\alpha}{2};n-1} = t_{1-\frac{0.05}{2};25-1} = t_{0.975;24} = 2.12$$

$$IC(\mu)_{(95)\%} = \left[4250 - 2.06.\frac{400}{\sqrt{24}}\ ; 4250 + 2.06.\frac{400}{\sqrt{24}}\right]$$

$$= [\mathbf{4081,8; 4418,9}]$$

أي أننا نستطيع أن نقول بمستوى ثقة 95% أن متوسط الأجر الشهري لعمال المصنع كله يتراوح بين 4081.8 ون و 4418.9 ون.

2.2 تقدير الفرق بين وسطي مجتمعين مستقلين بفترة:

وفيما يلي سنجد فترة الثقة للفرق بين وسطي مجتمعين طبيعيين مستقليين في حالات مختلفة:

2. 1.2−عند معلومية تباين مجتمعين σ_1^2 و σ_2^2

إذا كان $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$ فإن:

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

ومن النظرية نجد أن:

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

وتكون فترة ثقة $(1-\alpha)\%$ كالتالي:

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

حيث بحل المتباينة بالنسبة إلى $\mu_1 - \mu_2$ نجد أن:

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

مثال: سحبت عينتان من مجتمعين طبيعيين مستقلين وكانت عينة تمثل سرعة إنجاز تمرين معين والعينتان هما:

$$X_1: 20,18,22,22,18,20,25,25$$

$$X_2: 25,22,24,23,25,30,30,27,20$$

أوجد فترة ثقة 95% للفرق بين إنجازي المجموعتين في حالة تباين المجتمع الأول 6 وتباين المجتمع الثاني 12

الحل: نحسب في البداية وسطي العينتين

$$\overline{X}_1 = \frac{\sum x_i}{n} = \frac{170}{8} = 21.25$$

$$\overline{X}_2 = \frac{\sum x_i}{n} = \frac{226}{9} = 25.11$$

معلوم لدينا $\sigma_1^2 = 6$ و $\sigma_2^2 = 12$ وحجم العينات $n_1 = 8$ و $n_2 = 9$ وحيث أن $1 - \alpha = 0.95$

إذن $\alpha = 0.05$ ومنها $\frac{\alpha}{2} = 0.025$ وبالتالي $Z_{\frac{\alpha}{2}} = 1.96$ وذلك من الجدول التوزيع الطبيعي المعياري:

فتكون فترة الثقة كما يلي:

$$\left(\overline{X}_1 - \overline{X}_2\right) - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < \left(\overline{X}_1 - \overline{X}_2\right) + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$-3.86 - 1.96\sqrt{\frac{6}{8} + \frac{12}{9}} < \mu_1 - \mu_2 < -3.86 + 1.96\sqrt{\frac{6}{8} + \frac{12}{9}}$$

$$-6.69 < \mu_1 - \mu_2 < -1.03$$

2. 2.2. عند مجهولية تباين المجتمعين σ_1^2 و σ_2^2

1.2.2.2 إذا كان $n_1 . n_2 \geq 30$

في هذه الحالة نقدر تباين كل مجتمع بتباين عينة عشوائية تسحب منه ونستبدل تباين المجتمعين σ_1^2, σ_2^2 بتباين S_1^2 و S_2^2 في فترة ثقة للحالة السابقة تصبح فترة الثقة لهذه الحالة كما يلي:

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

مثال: أوجد فترة الثقة 98% للفرق بين وسطي مجتمعين طبيعيين مستقلين إذا علمت أن: حجم العينة الأولى 160 ووسطها 81.2 والانحراف المعياري لها 7.6 وحجم العينة الثانية 90 ووسطها 76.4 والانحراف لها 8.2 .

الحل:

$$n_1=160\,,\quad \bar{X}_1=81.2\ ,\quad S_1^2=7.6$$

$$n_2=90\,,\quad \bar{X}_2=76.4\ ,\quad S_2^2=8.2$$

ولاحظ أن$\alpha=0.02$ ومن الجدول للتوزيع الطبيعي نجد أن $Z_{\frac{\alpha}{2}}=2.33$ فتكون فترة الثقة المطلوبة كما يلي:

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

$$4.8-2.33\sqrt{\frac{7.6}{160}+\frac{8.2}{90}}<\mu_1-\mu_2<4.8+2.33\sqrt{\frac{7.6}{160}+\frac{8.2}{90}}$$

$$3.933<\mu_1-\mu_2<5.567$$

2.2.2.2 إذا كان$n_1.n_2<30$ (أو أحدهما أقل من 30) وكان تباينا المجتمعين غير متساو فإن فترة الثقة:

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

مثال: سحبت عينتان من مجتمعين طبيعيين مستقلين وكانت عينة تمثل سرعة إنجاز تمرين معين والعينتان هما:

$$X_1: 20,18,22,22,18,20,25,25$$

$$X_2: 25,22,24,23,25,30,30,27,20$$

أوجد فترة ثقة 95% للفرق بين إنجازي المجموعتين في حالة تباين المجتمعين مجهولين

الحل: معلوم من المثال السابق أن $\overline{X}_1 = 21.25$ و $\overline{X}_2 = 25.11$ وحجوم العينات هو $n_1 = 8$ و $n_2 = 9$ وهما أقل من 30 ومن t نجد أن $t_{(0.025,15)} = 2.131$.

بحساب تباين العينتين نجد أن $S_1^2 = 7.618$ و $S_2^2 = 11.628$ وبالتالي تكون فترة الثقة الثقة المطلوبة هي:

$$(\overline{X}_1 - \overline{X}_2) - t_{\left(\frac{\alpha}{2}n-1\right)}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + t_{\left(\frac{\alpha}{2}n-1\right)}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$-3.86 - 2.131\sqrt{\frac{7.618}{8} + \frac{11.628}{9}} < \mu_1 - \mu_2 < -3.86 + 2.131\sqrt{\frac{7.618}{8} + \frac{11.628}{9}}$$

$$-7.05 < \mu_1 - \mu_2 < -0.67$$

2. 3: فترة الثقة لنسبة المجتمع P:

لقد علمنا عند تعرضنا لتوزيعات المعاينة أن توزيع المعاينة لنسبة العينة $\widehat{P}$ يقترب من التوزيع الطبيعي عندما يكون حجم العينة n كبيراً، وذلك بوسط حسابي وتباين قدرهما على التوالي كما يلي:

$$E(\widehat{P}) = P \quad ، V(\widehat{P}) = \frac{Pq}{n}$$

إذن:

$$Z = \frac{\hat{P} - P}{\sqrt{\frac{Pq}{n}}}$$

حيث يتبع المتغير Z تقريبا التوزيع الطبيعي المعياري.

بما أن نسبة المجتمع P مجهولة، وهي التي نرغب في تقديرها بإيجاد فترة الثقة، فلا نستطيع حساب تباين توزيع المعاينة لنسبة العينة $\left(V(\widehat{P}) = pq/n\right)$ ولكننا سنقدره باستخدام أفضل مقدّر بالقيمة لنسبة المجتمع المجهولة P ، وهو نسبة العينة $\widehat{P}$، وعند استخدام مقدّر التباين وهو:

$$\hat{q} = 1 - \widehat{P} \qquad \text{حيث} \qquad V(\widehat{P}) = \frac{\widehat{P}\hat{q}}{n}$$

في الصيغة السابقة سيتوزع المتغير الجديد توزيعا قريبا من التوزيع الطبيعي المعياري، عندما يكون حجم العينة n كبيرا، إذن:

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

بإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة، سنحصل على الصورة التالية للاحتمال :

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

إذن: فترة الثقة لنسبة المجتمع P عندما يكون حجم العينة كبيراً، وعند مستوى ثقة $(1-\alpha)\%100$ هي:

$$\left(\widehat{P}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}\widehat{q}}{n}} \quad , \quad \widehat{P}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}\widehat{q}}{n}}\right)$$

وبالطبع إذا كان السحب مع عدم الإرجاع و $\frac{n}{N}>0.05$ ، يجب استعمال معامل التصحيح.

مثال: إذا اخترنا عشوائيا 500 طالب من طلبة جامعة سطيف 1، ووجدنا أن 180 منهم يملكون سيارة، فقدر نسبة الطلبة الذين يملكون سيارة في جامعة سطيف 1، ككل باستخدام مستوى ثقة 99%.

الحل: بما أن حجم العينة n كبير، فنستطيع استخدام فترة الثقة لنسبة المجتمع وهي كما يلي:

$$\left(\widehat{P}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}\widehat{q}}{n}} \quad , \quad \widehat{P}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{P}\widehat{q}}{n}}\right)$$

ولن نستخدم معامل التصحيح، لأن حجم المجتمع كبير، وبما أن:

$$(1-\alpha)\%100=99\% \quad ، \quad \widehat{P}=\frac{180}{500}=0.36 \ ، \quad n=500$$

إذن:

$$-z_{\frac{\alpha}{2}}=z_{0.005}=2.58 \ ، \ 0.005=\alpha/2 \ ، \quad 0.01=\alpha \ ، 0.99=1-\alpha$$

وبالتعويض في الحد الأدنى والحد الأعلى لفترة الثقة نحصل على:

الحد الأدنى لفترة الثقة:

$$\hat{P}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}} = 0.36-(2.58)\sqrt{\frac{(0.36)(0.64)}{500}} = 0.36-0.0554 = 0.3046$$

الحد الأعلى للفترة:

$$\hat{P}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}} = 0.36+(2.58)\sqrt{\frac{(0.36)(0.64)}{500}} = 0.36+0.0554 = 0.4154$$

إذن فترة الثقة لنسبة المجتمع P ، أي لنسبة الطلبة الذين يملكون سيارة في جامعة سطيف1، عند مستوى ثقة 99% هي:

$$(\%30.46 \quad , \quad \%41.54)$$

فنستطيع القول بثقة 99%بأن نسبة الطلبة الذين يملكون سيارة في جامعة سطيف 1، تقع بين النسبتين 30.46 % و 41.54 %.

مثال: من المسح الاقتصادي والاجتماعي الصادر عن الهيئة الوطنية للمعلومات والتوثيق في الجزائر لسنة 2002-2003، كان عدد الأسر التي شملتها العينة 11111 أسرة، منها 2566 أسرة ذات دخل منخفض، قدّر نسبة الأسر ذات الدخل المنخفض في الجزائر سنة 2002-2003 وذلك باستخدام فترة ثقة 95% .

الحل: بما أن حجم العينة n كبير، فنستطيع استخدام فترة الثقة لنسبة المجتمع وهي كما يلي:

$$\left(\hat{P}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}} \quad , \quad \hat{P}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}}\right)$$

ولن نستخدم معامل التصحيح، لأن حجم المجتمع كبير، وبما أن :

$$(1-\alpha)\%100 = 95\% \quad ، \quad \hat{P} = \frac{2566}{11111} = 0.2309 \text{ ، } n=11111$$

إذن:

$$1-\alpha = 0.95 \text{ ،}\alpha = 0{,}05 \text{ ، } Z_{\alpha/2} = Z_{.025} = 1.96$$

وبالتعويض في الحد الأدنى والحد الأعلى لفترة الثقة نحصل على:

الحد الأدنى لفترة الثقة:

$$\widehat{P}-Z_{\alpha/2}\sqrt{\frac{\widehat{P}\widehat{q}}{n}}=0.2309-(1.96)\sqrt{\frac{(0.2309)(0.7691)}{11111}}$$

$$=0.2309-0.0078=0.2231$$

الحد الأعلى لفترة الثقة:

$$\widehat{P}+Z_{\alpha/2}\sqrt{\frac{\widehat{P}\widehat{q}}{n}}=0.2309+(1.96)\sqrt{\frac{(0.2309)(0.7691)}{11111}}$$

$$=0.2309+0.0078=0.2387$$

إذن فترة الثقة لنسبة المجتمع P، أي لنسبة الأسر التي دخلها منخفض في الجزائر سنة 2002–2003 عند مستوى ثقة 95% هي:

(%22.31 , %23.87)

فنستطيع القول بثقة 95% بأن نسبة الأسر التي دخلها منخفض في الجزائر سنة 2002 – 2003 تقع بين النسبتين %22.31 و %23.87.

2 .4 فترة الثقة للفرق بين نسبتي مجتمعين (P_1-P_2):

بما أن توزيع المعاينة للفرق بين نسبتي عينتين عشوائيتين مستقلتين $(\widehat{P}_1-\widehat{P}_2)$

هو توزيع قريب من التوزيع الطبيعي إذا كانت $n_1 \geq 30$ ، $n_2 \geq 30$ وذلك بوسط حسابي

$E(\widehat{P}_1-\widehat{P}_2)$ وتباين $V(\widehat{P}_1-\widehat{P}_2)$ ، حيث :

$$E(\widehat{P}_1-\widehat{P}_2)=P_1-P_2 ،\quad V(\widehat{P}_1-\widehat{P}_2)=\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}$$

إذن المتغير العشوائي Z ، حيث:

$$Z=\frac{(\widehat{P}_1-\widehat{P}_2)-(P_1-P_2)}{\sqrt{\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}}}$$

ستوزع توزيعا قريبا من التوزيع الطبيعي المعياري.

بما أن نسبة المجتمع الأول P_1 مجهولة، ونسبة المجتمع P_2 مجهولة، فلا نستطيع حساب الخطأ المعياري لتوزيع المعاينة للفرق بين نسبتي العينتين$(\widehat{P}_1 - \widehat{P}_2)$، ولكننا سنقدره باستخدام أفضل مقدِّر بالقيمة لنسبة المجتمع الأول المجهول P_1 وهو نسبة العينة المسحوبة منه $\widehat{P}_1$ ، وأفضل مقدِّر بالقيمة لنسبة المجتمع الثاني المجهولة P_2 وهو نسبة العينة المسحوبة منه $\widehat{P}_2$ ، وعند استخدام مقدِّر الخطأ المعياري في الصيغة سيتوزع المتغير الجديد كذلك توزيعا قريبا من التوزيع الطبيعي المعياري لأن $n_1 \geq 30$ ، $n_2 \geq 30$ إذن :

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

بإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة سنحصل على الصورة التالية للاحتمال :

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

إذن:

فترة الثقة للفرق بين نسبتي مجتمعين $(P_1 - P_2)$ عندما يكون $n_1 \geq 30$ ،$n_2 \geq 30$ وعند مستوى ثقة $100\%(1-\alpha)$ هي :

$$\left[(\widehat{P}_1 - \widehat{P}_2) - Z_{\alpha/2}\sqrt{\frac{\widehat{P}_1\hat{q}_1}{n_1} + \frac{\widehat{P}_2\hat{q}_2}{n_2}},\qquad (\widehat{P}_1 - \widehat{P}_2) + Z_{\alpha/2}\sqrt{\frac{\widehat{P}_1\hat{q}_1}{n_1} + \frac{\widehat{P}_2\hat{q}_2}{n_2}}\right]$$

مثال: سُحبت عينتان عشوائيتان مستقلتان مع الإرجاع، الأولى تحتوي على 120 وحدة منتجة بالآلة (أ) ووجدنا بها 6 وحدات بها عيوب، والثانية تحتوي على 200 وحدة منتجة بالآلة (ب) ووجدنا بها 9 وحدات

بها عيوب، فقدّر الفرق بين نسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (أ) ونسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (ب)، وذلك باستخدام مستوى ثقة 95%.

الحل: بما أن العينتين العشوائيتين مستقلتان، $n_1 = 120 > 30$ و $n_2 = 200 > 30$ إذن نستطيع استعمال فترة الثقة وهي:

$$\left[(\widehat{P}_1 - \widehat{P}_2) - Z_{\alpha/2}\sqrt{\frac{\widehat{P}_1\widehat{q}_1}{n_1} + \frac{\widehat{P}_2\widehat{q}_2}{n_2}}, \quad (\widehat{P}_1 - \widehat{P}_2) + Z_{\alpha/2}\sqrt{\frac{\widehat{P}_1\widehat{q}_1}{n_1} + \frac{\widehat{P}_2\widehat{q}_2}{n_2}}\right]$$

ومن البيانات المتوفرة لدينا نستطيع حساب ما يلي:

نسبة الوحدات التي بها عيوب لعينة إنتاج الآلة (أ) $\widehat{P}_1 = \frac{6}{120} = 0.05$

نسبة الوحدات التي بها عيوب لعينة إنتاج الآلة (ب) $\widehat{P}_2 = \frac{9}{200} = 0.045$

وبما أن: $95\% = (1-\alpha)\%100$

إذن:

$0.95 = 1 - \alpha$ ،$1.96 = Z_{.025} = Z_{\alpha/2}$، $0.025=\alpha/2$

وبالتعويض في الحد الأدنى والحد الأعلى لفترة الثقة نحصل على:

الحد الأدنى لفترة الثقة:

$$(\widehat{P}_1 - \widehat{P}_2) - Z\alpha_{/2}\sqrt{\frac{\widehat{P}_1\widehat{q}_1}{n_1} + \frac{\widehat{P}_2\widehat{q}_2}{n_2}}$$

$$= (0.05 - 0.045) - (1.96)\sqrt{\frac{(0.05)(0.95)}{120} + \frac{(0.45)(0.955)}{200}} = -0.0434$$

الحد الأعلى لفترة الثقة:

$$(\widehat{P}_1 - \widehat{P}_2) + Z\alpha_{/2}\sqrt{\frac{\widehat{P}_1\widehat{q}_1}{n_1} + \frac{\widehat{P}_2\widehat{q}_2}{n_2}}$$

$$= (0.05 - 0.045) + (1.96)\sqrt{\frac{(0.05)(0.95)}{120} + \frac{(0.45)(0.955)}{200}} = 0.0534$$

إذن فترة الثقة للفرق بين نسبتي المجتمعين $(P_1 - P_2)$ عند مستوى ثقة 95% :

$(\%-4.34, \%5.34)$

أي بثقة قدرها 95% نستطيع القول بأن الفرق بين نسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (أ) ونسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (ب) يقع بين النسبتين %−4.34 و %5.34 .

5.2 فترة الثقة لتباين المجتمع σ^2:

علمنا أنه إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بتباين معلوم σ^2 ، وسحبنا منه عينة عشوائية حجمها n، فإن المتغير العشوائي التالي يطلق عليه x^2 ، حيث:

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

والتوزيع الاحتمالي لهذا المتغير يُسمى توزيع x^2 (توزيع مربع كاي) بدرجة حرية v حيث $v = n-1$.

من جدول x^2 نستطيع الحصول على قيمتين للمتغير x^2 عند درجة حرية معينة، بحيث تكون المساحة على يمين القيمة الكبرى تساوي المساحة على يسار القيمة الصغرى تساوي $^{\alpha}/_{2}$، وإذا رمزنا لكل قيمة باستعمال المساحة التي على يمينها سنرمز للقيمة الكبرى بالرمز $x^2_{(\alpha/2)}$ وللقيمة الصغرى بالرمز $x^2_{(1-\alpha/2)}$ وحيث إن المساحة الكلية تحت منحنى توزيع x^2 تساوي الواحد الصحيح، إذن المساحة بين هاتين القيمتين تساوي $(1-\alpha)$، ونعبِّر عن هذه المساحة باستخدام الاحتمال كما يلي:

$$P(x^2_{(1-\alpha/2)}(v) \leq x^2 \leq x^2_{(\alpha/2)}(v)) = 1 - \alpha$$

وبالتعويض عن x^2 سيأخذ الاحتمال الصورة التالية :

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

وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة، سنحصل على الصورة التالية للاحتمال:

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

إذن: **فترة الثقة لتباين المجتمع σ^2 عند مستوى ثقة $(1-\alpha)100\%$ هي:**

$$\boxed{\left(\frac{(n-1)S^2}{x^2_{(\alpha/2)}(v)} \; , \; \frac{(n-1)S^2}{x^2_{(1-\alpha/2)}(v)}\right)}$$

فترة الثقة للانحراف المعياري للمجتمع عند مستوى ثقة $(1-\alpha)100\%$ هي:

$$\boxed{\left(\sqrt{\frac{(n-1)S^2}{x^2_{(\alpha/2)}(v)}} \; , \; \sqrt{\frac{(n-1)S^2}{x^2_{(1-\alpha/2)}(v)}}\right)}$$

مثال:

إذا علمت أن أوزان الطلبة تتوزع توزيعا طبيعيا، وسحبنا عينة عشوائية منهم تحتوي على 10 طلاب، وكانت أوزانهم كما يلي:

69 ، 67 ، 49 ، 81 ، 38 ، 55 ، 65 ، 50 ، 72 ، 44

قدّر التباين والانحراف المعياري لأوزان الطلبة وذلك باستخدام فترة عند مستوى ثقة 95% .

الحل: المجتمع محل الدراسة يتكون من أوزان الطلبة ، إن هذا المجتمع يتبع توزيعا طبيعيا، إذن فترة الثقة لتباين أوزان ستكون هي:

$$\left(\frac{(n-1)S^2}{x^2_{(\alpha/2)}(v)} \; , \; \frac{(n-1)S^2}{x^2_{(1-\alpha/2)}(v)}\right)$$

حيث: $n=10$ ، $v=n-1=10-1=9$ ، $(1-\alpha)100\% = \%95$

إذن: $1-\alpha = 0.95$ ، $\alpha = 0.05$ ، $\alpha/2 = 0.025$ ،

$19.023 = x^2_{0.025}(9) = x^2_{(\alpha/2)}(v)$ ، $2.70 = x^2_{0.975}(9) = x^2_{(1-\alpha/2)}(v)$ ، أما تباين

العينة S^2 نحسبه بعد حساب الوسط الحسابي للعينة حيث:

$$\overline{X} = \frac{\sum x}{n} = \frac{590}{10} = 59$$

$$S^2 = \frac{\sum(x-\overline{x})^2}{n-1} = \frac{(69-59)^2 + (67-59)^2 + \ldots\ldots\ldots + (44-59)^2}{10-1}$$
$$= 190.67$$

وبالتالي فإن الحد الأدنى لفترة: $\frac{(n-1)S^2}{x^2_{(\alpha/2)}(v)} = \frac{(10-1)(190.67)}{19.023} = 90.21$

والحد الأعلى لفترة الثقة: $\frac{(n-1)S^2}{x^2_{(1-\alpha/2)}(v)} = \frac{(10-1)(190.67)}{2.70} = 635.56$

إذن فترة الثقة لتباين المجتمع σ^2 عند مستوى ثقة 95% هي: $(90.21, 635.56)$

أي نستطيع القول بثقة قدرها 95% بأن تباين أوزان الطلبة يقع بين القيمتين 90.21 و 635.56 .

بأخذ الجذر التربيعي للحدين نحصل على حدي فترة الثقة للانحراف المعياري، وبالتالي فإن فترة الثقة للانحراف المعياري لأوزان الطلبة عند مستوى ثقة 95% هي:(25.21 , 9.5)

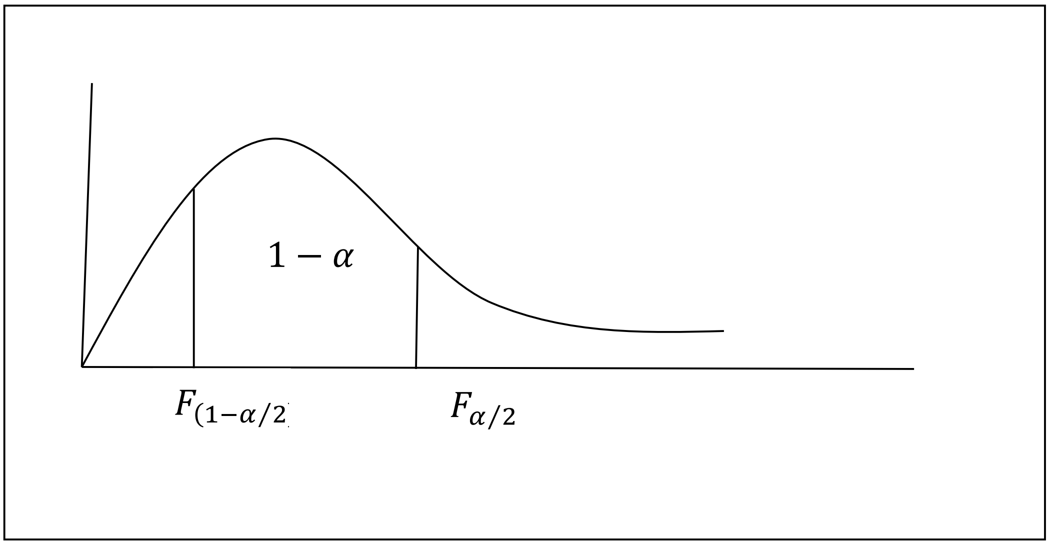
6.2 **فترة الثقة للنسبة بين تبايني مجتمعين** $\frac{\sigma_1^2}{\sigma_2^2}$: إذا كان لدينا مجتمعان يتوزع كل منهما توزيعا طبيعيا، وسحبنا من المجتمع الأول الذي تباينه σ_1^2 كل العينات العشوائية الممكنة ذات الحجم n_1 ، وحسبنا منها قيم المتغير S_1^2 ، وسحبنا من المجتمع الثاني الذي تباينه σ_2^2 كل العينات العشوائية الممكنة ذات الحجم n_2 ، وحسبنا منها قيم المتغير S_2^2 ، وكانت العينات المسحوبة من المجتمع الأول مستقلة عن العينات المسحوبة من المجتمع الثاني، فإن المتغير العشوائي F يمكن تعريفه كما يلي:

$$F = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

من الجدول الإحصائي نستطيع الحصول على قيمتين للمتغير F عند درجتي معينتين للحرية v_1 ، v_2 حيث v_1 درجة حرية البسط و v_2 درجة حرية المقام، بحيث تكون المساحة على يمين القيمة الكبرى تساوي المساحة على يسار القيمة الصغرى تساوي $\alpha/2$، وإذا رمزنا لكل قيمة باستعمال المساحة التي على يمينها سنرمز للقيمة الكبرى بالرمز $F_{\alpha/2}$ ، وللقيمة الصغرى بالرمز $F_{(1-\alpha/2)}$ ، وحيث أن المساحة الكلية تحت

منحنى توزيع F تساوي الواحد الصحيح، إذن المساحة بين هاتين القيمتين تساوي $(1-\alpha)$ ونعبر عن هذه المساحة باستخدام الاحتمال كما يلي:

$$P(F_{(1-\alpha/2)}(v_1,v_2) \leq F \leq F_{\alpha/2}(v_1,v_2)) = 1-\alpha$$



وبالتعويض عن F سيأخذ الاحتمال الصورة التالية:

$$P(F_{(1-\alpha/2)}(v_1,v_2) \leq \frac{S_1^2}{S_2^2}.\frac{\sigma_2^2}{\sigma_1^2} \leq F_{\alpha/2}(v_1,v_2)) = 1-\alpha$$

بإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة سنحصل على الصورة التالية للاحتمال:

$$P\left(\frac{S_1^2/S_2^2}{F_{(\alpha/2)}(v_1,v_2)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2/S_2^2}{F_{(1-\alpha/2)}(v_1,v_2)}\right) = 1-\alpha$$

إذن :

فترة الثقة للنسبة بين تبايني مجتمعين $\frac{\sigma_1^2}{\sigma_2^2}$ عند مستوى ثقة $(1-\alpha)100\%$

هي: $\left(\frac{S_1^2/S_2^2}{F_{(1-\alpha/2)}(v_1,v_2)} \text{ ، } \frac{S_1^2/S_2^2}{F_{(\alpha/2)}(v_1,v_2)}\right)$

يجب الانتباه إلى أن هذه الفترة لا تُستخدم إلا إذا توفرت الفروض التالية:

- يتوزع المجتمعان توزيعا طبيعيا.

- العينة المسحوبة من المجتمع الأول مستقلة عن العينة المسحوبة من المجتمع الثاني.

مثال: إذا علمت أن مجتمع أطوال الطالبات ومجتمع أطوال الطلبة في جامعة سطيف 1 يتبع توزيعا طبيعيا، وسحبنا من الطالبات عينة عشوائية تشمل 25 طالبة، ومن الطلبة عينة عشوائية تشمل 21 طالبا وكانت العينتان مستقلتين، ووجدنا أن تباين أطوال عينة الطالبات يساوي 64، وتباين عينة أطوال الطلبة يساوي 36، أوجد فترة الثقة لنسبة تباين مجتمع أطوال الطالبات إلى تباين مجتمع أطوال الطلبة، وذلك باستخدام مستوى ثقة 95%.

الحل: إذا رمزنا لتباين مجتمع أطوال الطالبات بالرمز σ_1^2، ورمزنا لتباين مجتمع أطوال الطلبة بالرمز σ_2^2، إذن المطلوب هو فترة الثقة للنسبة بين تبايني المجتمعين (σ_1^2/σ_2^2) عند مستوى ثقة 95%، بما أن المجتمعين يتوزعان توزيعا طبيعيا، والعينتين مستقلتان، إذن فترة الثقة عند مستوى ثقة $(1-\alpha)100\%$ هي:

$$\left(\frac{S_1^2/S_2^2}{F_{(\alpha/2)}(v_1,v_2)} \text{ ، } \frac{S_1^2/S_2^2}{F_{(1-\alpha/2)}(v_1,v_2)}\right)$$

حيث:

$$v_2 = n_2 - 1 = 21 - 1 = 20 \,, v_1 = n_1 - 1 = 25 - 1 = 24$$

وبما أن: $0.025 = \alpha/2 \;,\; 0.05 = \alpha \;,\; 0.95 = 1-\alpha$

إذن: $F_{(\alpha/2)}(v_1,v_2) = F_{0.025}(24,20) = 2.41$

$$F_{(1-\alpha/2)}(v_1,v_2) = F_{0.975}(24{,}20) = \frac{1}{F_{.025}(24{,}20)} = \frac{1}{2.33} = 0.43$$

إذن الحد الأدنى لفترة الثقة:

$$\frac{S_1^2/S_2^2}{F_{(\alpha/2)}(v_1,v_2)} = \frac{64/36}{2.41} = 0.74$$

والحد الأعلى لفترة الثقة:

$$\frac{S_1^2/S_2^2}{F_{(1-\alpha/2)}(v_1,v_2)} = \frac{64/36}{0.43} = 4.13$$

إذن فترة الثقة لنسبة تباين مجتمع أطوال الطالبات إلى تباين مجتمع أطوال الطلبة (σ_1^2/σ_2^2) في جامعة سطيف1 عند مستوى ثقة 95% هي: (4.13 ، 0.74)

أي نثق بمقدار 95% بأن النسبة بين تباين أطوال كل الطلبة وتباين أطوال كل الطالبات في هذه الجامعة يقع بين 0.73 و 4.13 .

وبأخذ الجذر التربيعي للحدين نحصل على حدي فترة الثقة عند مستوى ثقة 95% للنسبة بين الانحراف المعياري لأطوال كل الطلبة، والانحراف المعياري لأطوال كل الطالبات σ_1/σ_2 وهي: 2.03، 0.86

أي نثق بمقدار 95% بأن النسبة بين الانحراف المعياري لأطوال كل الطلبة والانحراف المعياري لأطوال كل الطالبات في هذه الجامعة يقع بين 0.86 و 2.03.

7.2. خطأ المعاينة في التقدير وحجم العينة اللازم لعدم تجاوز خطأ معين:

إن تقدير أي معلمة من معالم المجتمع $(\partial^2، P، \mu)$ بواسطة $(S^2, f, \overline{X})$ يترتب عليه خطأ يدعى خطأ المعاينة وسببه إجراء الدراسة الإحصائية على عينة بدلا من أن تكون شاملة. ورياضيا يمثل خطأ المعاينة القيمة المطلقة للفرق بين قيمة المقدر والقيمة الحقيقية للمعلمة أي: $|\theta - \hat{\theta}|$

1.7.2. حالة المتوسط الحسابي للمجتمع μ

نعلم أنه عند مستوى ثقة $(1-\alpha)$ % فإن مجال الثقة للمتوسط الحسابي في حالة ما إذا كان تباين المجتمع معلوم هو:

$$IC(\mu)_{(1-\propto)\%} = \left(\overline{X} - Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}} , \overline{X} + Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}}\right)$$

أي أن:

$$\overline{X} - Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}} \leq \mu \leq \overline{X} + Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}}$$

ومنه:

$$-Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}} \leq \mu - \overline{X} \leq +Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}}$$

وبذلك يكون:

$$|\overline{X}-\mu| \leq +Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}}$$

أي أن خطأ المعاينة في تقدير μ باستخدام $\overline{X}$ هو على الأكثر المقدار:

$Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}}$ ونسمي هذا المقدار بالحد الأعلى لخطأ المعاينة في تقدير μ ونرمز له ب E حيث:

$$E = Z_{1-\frac{\propto}{2}}.\frac{\partial}{\sqrt{n}}$$

وانطلاقا من صيغة E يمكن حساب حجم العينة n حيث نقوم بتربيع الطرفين فنجد :

$$E^2 = (Z_{1-\frac{\propto}{2}})^2.\frac{\partial^2}{n}$$

ومنه :

$$n = (Z_{1-\frac{\propto}{2}})^2.\frac{\partial^2}{E^2}$$

وهو حجم العينة المطلوب.

ملاحظة1: هناك علاقة عكسية بين n و E حيث كلما زاد حجم العينة كلما كان خطأ المعاينة ضعيف وكانت الدراسة الإحصائية أكثر دقة.

ملاحظة2: يتأثر تحديد حجم العينة بعدة عوامل منها: حجم المجتمع المراد دراسته، مستوى الثقة، الانحراف المعياري.

ملاحظة3: إن تحديد حجم العينة مسبقا في المعاينة الإحصائية يترتب عليه توفير الوقت اللازم والتحكم في التكاليف والحصول على نتائج جيدة وفي وقت قصير.

2.7.2. حالة النسبة P للمجتمع:

عند تقدير نسبة المجتمع P بواسطة f فإن الحد الأعلى لخطأ المعاينة في تقدير P عند مستوى ثقة $(1-\alpha)$ % هو:

$$E = Z_{1-\frac{\propto}{2}} . \sqrt{\frac{f(1-f)}{n}}$$

أما حجم العينة اللازم فهو:

$$n = (Z_{1-\frac{\propto}{2}})^2 . \frac{f(1-f)}{E^2}$$

مثال: في مصنع لإنتاج السميد ومشتقاته تقوم آلة بملأ أكياس الفرينة ذات الوزن 1كغ بغرض معرفة فيما إذا كانت هذه الآلة تقوم بعملية الملأ بكيفية جيدة نقوم بسحب عينة عشوائية ذات حجم معين ونقدّر متوسط وزن الأكياس التي تملأها الآلة انطلاقا من هذه العينة بافتراض أن وزن الأكياس عبارة عن متغير عشوائي يتبع التوزيع الطبيعي بانحراف معياري يساوي 15غ. فما هو حجم العينة التي يجب أخذها حتى نكون واثقين بنسبة 95% أن الخطأ في تقدير متوسط وزن الأكياس التي تملأها هذه الآلة لن يزيد عن 5غ.

الحل:

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

$$\sigma^2 = 15 \,, 1-\alpha = 0.95 \,, E = 5$$

$$E^2 = (Z_{1-\frac{\propto}{2}})^2 . \frac{\partial^2}{n}$$

$$n = (Z_{1-\frac{\propto}{2}})^2 . \frac{\partial^2}{E^2}$$

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

$$n = (1.96)^2 . \frac{15^2}{5^2} = 34.57 = 35$$

الفصل الرابع
إختبار الفرضيات

إن الغرض من اختبارات الفروض، هو وصول الباحث إلى قرار بخصوص فرض معين حول معلمة المجتمع، وما يجب أن يوصى به مع إحتمال الوقوع في خطأ يحدده هذا الباحث. ويمكن إجراء ذلك من خلال معرفة توزيع معاينة المقدر Estimator الخاص بمعلمة المجتمع، ومن خلال هذا التوزيع يمكن تحديد مناطق رفض وقبول فرض العدم (وهو ما يود الباحث أن يثبت ضده) أسفل منحى التوزيع.

1. مفاهيم أساسية:

1.1 الفرضية الإحصائية: عبارة عن تخمين معين حول خصائص مجتمع ما مثل قيمة معلمة معينة و أحيانا حول أكثر من مجتمع مثل تساوي متوسط مجتمعين، قد يكون هذا التخمين صحيحا أو خاطئا.

2.1 إختبار الفرضية: عبارة عن أسلوب إحصائي هدفه وضع قاعدة قرار نتمكن على ضوئها من الإختيار بين فرضيتين إحصائيتين إعتمادا على بيانات العينة.

3.1 أنواع الفروض الإحصائية: تنقسم الفروض إلى قسمين هما:

1.3.1 الفرض البديل (البحثي): H_aأو H_1 Alternative Hypotheses: وهو ما يود الباحث أن يثبت صحته، ويوصي به في كثير من الأحوال.

2.3.1 الفرض العدم: Null Hypotheses Ho : وهو ما يود الباحث أن يثبت ضده وهناك ثلاث اتجاهات لصياغة الفروض البديلة وهي:

الفرض البديل (البحثي): H_a	الفرض العدم: H_0
$H_a: \mu \neq \mu_0$	$H_0: \mu = \mu_0$
$H_a: \mu > \mu_0$	$H_0: \mu \leq \mu_0$
$H_a: \mu < \mu_0$	$H_0: \mu \geq \mu_0$

حيث أن μ_0 هو قيمة متوسط المجتمع تحت صحة الفرض العدم

ويلاحظ أن الفرض العدم مصاحب دائما بعلامة = لذا يمكن كتابة الفرض العدم على الصورة:

$$H_0: \mu = \mu_0$$

4.1 مستوى المعنوية أو مستوى الإحتمال:

وهي درجة الإحتمال الذي نرفض به فرضية العدم H_0 عندما تكون صحيحة أو هو إحتمال الوقوع في الخطأ من النوع الأول ويرمز له بالرمز ∝ ، وهي يحددها الباحث لنفسه منذ البداية ومعظم العلوم التطبيقية نختار ∝ مساوية 1% أو 5% على الأكثر.

5.1 أخطاء إختبار الفرضيات

إن إتخاذ القرار ينطوي على نوعين من الأخطاء في أي من إختبارات الفروض الإحصائية وهذا يقودنا إلى التعريف التالي:

1) **الخطأ من النوع الأول:** (error of type I) في إختبار الفروض هو الخطا الذي يحدث عند رفض فرضية العدم عندما تكون صحيحة. و الإحتمال الذي يحدث نتيجة الخطأ الأول نرمز له بالرمز ∝.

2) **الخطأ من النوع الثاني:** (error of type II) في إختبار الفروض هو الخطا الذي يحدث عند قبول (عدم رفض) فرضية العدم عندما تكون خاطئة. و الإحتمال الذي يحدث نتيجة الخطأ الثاني نرمز له بالرمز β.

والجدول التالي يوضح هذين النوعين من الأخطاء:

الجدول رقم 02: أنواع أخطاء الإختبارات

القرار الإحصائي / الحالة الحقيقية	رفض (H_0)	قبول (H_0)
صحيحة (H_0)	قرار خاطئ (الخطأ من النوع الأول)	قرار صحيح
خاطئة (H_0)	قرار صحيح	قرار خاطئ (الخطأ من النوع الثاني)

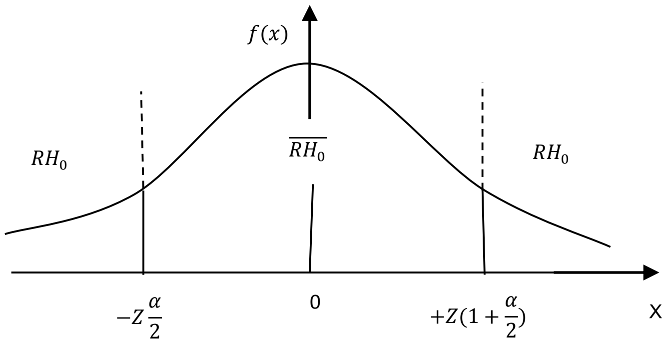
6.1 أنواع الفرضيات

1.6.1 إختبار ذو جانبين (الطرفين): في هذا النوع من الإختبار نأخذ قيمة واحدة بالنسبة للفرضية المبدئية، على عكس الفرضية البديلة التي تأخذ جميع القيم التي تختلف على هذه القيمة سواء بزيادة أو نقصان، فعلى سبيل المثال إذا كان إختبارنا يتعلق بمتوسط المجتمع μ فإننا نقول أن الإختبار ذو جانبين إذا كان إهتمامنا منصبا على تغيرات μ في كلا الإتجاهين ($<$ ، $>$) من قيمة معينة μ_0 ونصيغ المسألة الكيفية التالية:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

وتكون منطقة الرفض موزعة بالتساوي على حدي التوزيع والشكل التالي يوضح ذلك:



حيث تمثل $\overline{RH_0}$ منطقة عدم الرفض لـ: H_0

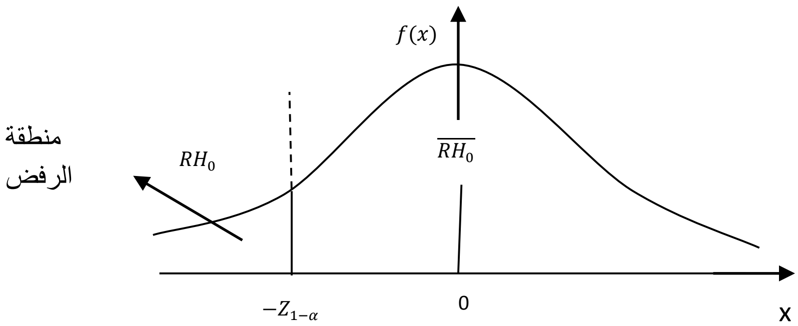
RH_0 تمثل الرفض لـ : H_0

2.6.1 **إختبار الطرف الأيسر:** إذا كان إهتمامنا منصبا على تغيرات μ مثلا في إتجاه واحد من قيمة معينة μ_0 وكان إهتمامنا بالتغير من جهة اليسرى فإن المسألة تكون على النحو التالي:

$$H_0: \mu \geq \mu_0$$

$$H_1: \mu < \mu_0$$

حيث: بيانيا تكون الصورة التالية:



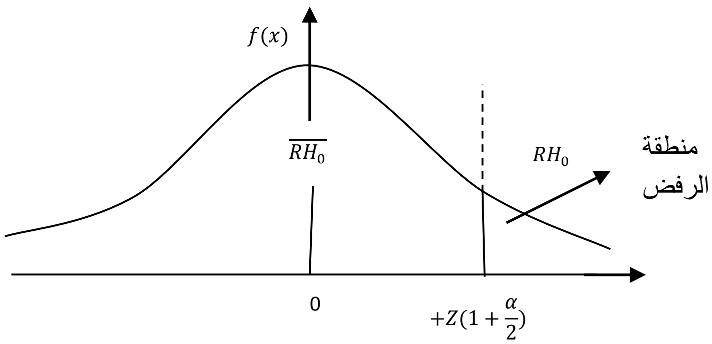
3.6.1 إختبار الطرف الأيمن:

إذا كان إهتمامنا منصبا على تغيرات μ مثلا في إتجاه واحد من قيمة معينة μ_0 وكان إهتمامنا بالتغير من جهة اليمنى فإن المسألة تكون على النحو التالي:

$$H_0: \mu \leq \mu_0$$

$$H_1: \mu > \mu_0$$

وبيانيا يمثل هذا النوع على الشكل التالي:



7.1 الخطوات المتبعة في إختبار الفرضيات

فيما يلي أهم الخطوات المتبعة في إختبار الفرضيات الإحصائية، نوجزها بما يأتي:

- تحديد الفرضية العدمية (H_0) و الفرضية البديلة (H_1)، مع مراعات تحديد نوع الفرضية (H_1) كأن تكون من جانب واحد أو جانبين؛
- تحديد مستوى المعنوية (α) ، كأن يكون (0.05 أو 0.01)؛
- تقدير المؤشر الإحصائي المطلوب إختباره؛
- حساب قيمة إحصاءة الإختبار ولتكن(v)؛

- تحديد القيمة الحرجة (v_α) أو $(v_{\alpha/2})$، إعتمادا على مستوى المعنوية (α)، ونوع الفرضية البديلة (H_1)، ودرجات الحرية (.df) في حالة توزيعات المعانية؛

- إتخاذ القرار بشأن رفض أو قبول الفرضية العدمية (H_0)، بعد مقارنة قيمة إحصاءة الإختبار المحسوبة (v)،مع القيم الحرجة (v_α) أو $(v_{\alpha/2})$ حسب نوع الفرضية البديلة (H_1)، ومستوى المعنوية (α)، فإذا كانت:

أ- قيمة إحصاءة الإختبار المحسوبة (v) تقع في منطقة رفض الفرضية العدمية (H_0) ، فإن ذلك يدل ذلك على رفض الفرضية العدمية (H_0) ، وقبول الفرضية البديلة(H_1)؛

ب-أما إذا وقعت قيمة إحصاءة الإختبار المحسوبة (v)،في منطقة قبول الفرضية العدمية(H_0)، فإن ذلك يدل على قبول الفرضية العدمية (H_0)، ورفض الفرضية البديلة (H_1).

- **الإستنتاج:** في هذه الفقرة يقوم الباحث بتثبيت إستنتاجاته حول معلمات المجتمع التي تم إختبارها في ضوء معطيات العينة التي اختيرت من مجتمع الدراسة.

2.إختبارات المتوسط الحسابي للمجتمع :

1.2 **اختبار الطرفين أو ذو جانبين:** نفرض أننا نريد أن نختبر فرضية أن قيمة متوسط المجتمع هي قيمة معينة μ_0 ، أي :

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

1.1.2 في حالة $n \geq 30$:

لنفرض أنه تم سحب عينة حجمها $n \geq 30$ من مجتمع يتوزع بمتوسط μ ونريد إختبار الفرضية السابقة.

نعلم مما سبق أن أفضل مقدر لمتوسط المجتمع هو متوسط العينة$\overline{X}$ و نعلم أن :

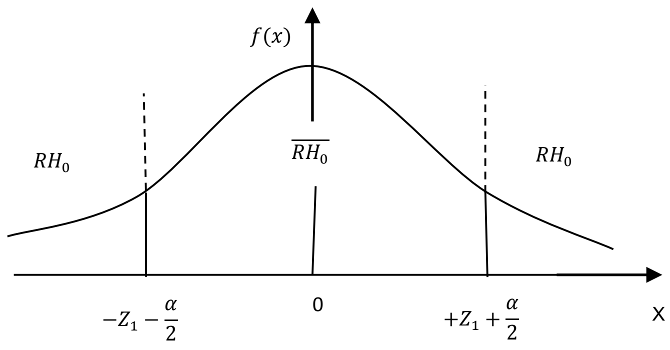
$$Z = \frac{\overline{X} - \mu_0}{\sqrt{V(\overline{X})}} \rightarrow N(0.1)$$

عند مستوى دلالة α فإن قرار رفض أو عدم رفض الفرضية المبدئية H_0 كما هو موضح في الشكل يكون وفقا لقاعدة إتخاذ القرار التالية:

$$\begin{cases} Z = \dfrac{\overline{X} - \mu_0}{\sqrt{V(\overline{X})}} \in \left[-Z_{1-\frac{\alpha}{2}};\ +Z_{1-\frac{\alpha}{2}}\right] \rightarrow \overline{RH_0} \\ Z = \dfrac{\overline{X} - \mu_0}{\sqrt{V(\overline{X})}} \notin \left[-Z_{1-\frac{\alpha}{2}};\ +Z_{1-\frac{\alpha}{2}}\right] \rightarrow RH_0 \end{cases}$$

$$\left.\begin{array}{ll} \sqrt{V(\overline{X})} = \sqrt{\dfrac{\partial^2}{n}} & \text{إذا كان تباين المجتمع معلوم} \\ \sqrt{V(\overline{X})} = \begin{cases} \sqrt{\dfrac{S^2}{n-1}} \\ \dfrac{\hat{s}}{\sqrt{n}} \end{cases} & \text{إذا كان تباين المجتمع مجهول} \end{array}\right\}$$

$$\begin{cases} s^2 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \\ \hat{s}^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \end{cases}$$



$\overline{RH_0}$ تمثل منطقة عدم الرفض لـ H_0.

RH_0 تمثل منطقة الرفض لـ H_0.

مثال: الفترة المعيشية لجهاز كهربائي هي 1750 ساعة مع إنحراف معياري قدره 120 ساعة. وقد تم الحصول على هذه النتائج عن طريق عينة عشوائية حجمها 101 جهاز إذا كان μ يمثل الفترة الوسطية المعيشية للأجهزة المنتجة في المؤسسة.

اختبر الفرضية التالية عند $\alpha = 5\%$

$$H_0: \mu = 1600$$

$$H_1: \mu \neq 1600$$

الحل: لدينا مجتمع غير محدود ولدينا : n=101

نعلم أنه في حالة عينة كبيرة فإن قاعدة إتخاذ القرار تكون على النحو التالي :

$$\begin{cases} Z = \dfrac{\overline{X} - \mu_0}{\sqrt{V(\overline{X})}} \in \left[-Z_{1-\frac{\alpha}{2}};\ +Z_{1-\frac{\alpha}{2}}\right] \rightarrow \overline{RH_0} \\ Z = \dfrac{\overline{X} - \mu_0}{\sqrt{V(\overline{X})}} \notin \left[-Z_{1-\frac{\alpha}{2}};\ +Z_{1-\frac{\alpha}{2}}\right] \rightarrow RH_0 \end{cases}$$

$$\mu_0 = 1600$$

$$\overline{x} = 1750$$

من جدول التوزيع الطبيعي نجد: $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$

$$\sqrt{V(\overline{X})} = \frac{s}{\sqrt{n-1}} = \frac{120}{\sqrt{101-1}} = 12$$

$$Z = \frac{1750 - 1600}{12} = 12.5$$

$$\left[-z_{1-\frac{\alpha}{2}};\ +z_{1-\frac{\alpha}{2}}\right] = [-1.96; 1.96]$$

$$12.5 \notin [-1.96; 1.96]$$

$$Z = \frac{\overline{X} - \mu_0}{\sqrt{V(\overline{X})}} \notin \left[-Z_{1-\frac{\alpha}{2}};\ +Z_{1-\frac{\alpha}{2}}\right] \rightarrow RH_0$$

أي رفض H_0 عند مستوى المعنوية $\alpha = 5\%$ أي أننا لا نقبل أن الفترة الوسطية المعيشية للأجهزة الكهربائية هي 1600 سا.

2.1.2 في حالة $n < 30$:

لنفرض أننا سحبنا عينة عشوائية حجمها $\mathbf{n} < 30$ من مجتمع غير محدود يتوزع توزيع طبيعي بمتوسط $\boldsymbol{\mu}$ وتباين $\boldsymbol{\partial}^2$ مجهولين.

نعلم أن:

$$T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{s^2}{n-1}}} = \frac{\bar{X} - \mu_0}{\frac{\hat{s}}{\sqrt{n}}} \rightarrow t_{n-1}$$

عند مستوى دلالة α فإن قرار رفض أو عدم رفض الفرضية المبدئية H_0 يكون وفقا لقاعدة إتخاذ القرار التالية :

$$\begin{cases} T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{s^2}{n-1}}} = \frac{\bar{X} - \mu_0}{\frac{\hat{s}}{\sqrt{n}}} \in \left[-t_{n-1;1-\frac{\propto}{2}};\ +t_{n-1;1-\frac{\propto}{2}}\right] \rightarrow \overline{RH_0} \\ T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{s^2}{n-1}}} = \frac{\bar{X} - \mu_0}{\frac{\hat{s}}{\sqrt{n}}} \notin \left[-t_{n-1;1-\frac{\propto}{2}} + t_{n-1;1-\frac{\propto}{2}}\right] \rightarrow RH_0 \end{cases}$$

مثال: قمنا بسحب عينة حجمها $n = 17$ من مجتمع غير محدود يتوزع توزيع طبيعي بمتوسط وتباين مجهولين فوجدنا أن $s = 2$, $\bar{X} = 11$

عند مستوى المعنوية $\alpha = 5\%$ إختبر صحة الفرضية التالية:

$$H_0: \mu = 10$$

$$H_1: \mu \neq 10$$

الحل: لدينا مجتمع غير محدود ولدينا:

$$\boldsymbol{\alpha = 5\%} \text{ ، } n = 17 \text{ ، } \boldsymbol{\bar{X} = 11} \text{ ، } \boldsymbol{s = 2}$$

نعلم أنه في حالة عينة صغيرة فإن قاعدة إتخاذ القرار تكون كما يلي:

$$\begin{cases} T = \dfrac{\overline{X} - \mu_0}{\sqrt{\dfrac{s^2}{n-1}}} = \dfrac{\overline{X} - \mu_0}{\dfrac{\hat{s}}{\sqrt{n}}} \in \left[-t_{n-1;1-\frac{\propto}{2}};\ +t_{n-1;1-\frac{\propto}{2}}\right] \rightarrow \overline{RH_0} \\ T = \dfrac{\overline{X} - \mu_0}{\sqrt{\dfrac{s^2}{n-1}}} = \dfrac{\overline{X} - \mu_0}{\dfrac{\hat{s}}{\sqrt{n}}} \notin \left[-t_{n-1;1-\frac{\propto}{2}} + t_{n-1;1-\frac{\propto}{2}}\right] \rightarrow RH_0 \end{cases}$$

$$\mu_0 \in \left[\overline{X} - t_{n-1;1-\frac{\propto}{2}} \frac{s}{\sqrt{n-1}};\ \overline{X} + t_{n-1;1-\frac{\propto}{2}} \frac{s}{\sqrt{n-1}}\right] \rightarrow \overline{RH_0}$$

من جدول توزيع ستيودنت نجد :

$$t_{n-1;1-\frac{\propto}{2}} = t_{16;0.975} = 2.12$$

$$\frac{s}{\sqrt{n-1}} = \frac{2}{\sqrt{17-1}} = 0.5$$

$$T = \frac{\overline{X} - \mu_0}{\sqrt{\frac{s^2}{n-1}}} = \frac{11-10}{0.5} = 2$$

$$\left[-t_{n-1;1-\frac{\propto}{2}};\ t_{n-1;1-\frac{\propto}{2}}\right] = [-2.12\ ;\ +2.12]$$

بما أن $2 \in [-2.12\ ;\ +2.12]$ فإننا لانرفض H_0

أي نعتبر أن متوسط المجتمع يساوي 10.

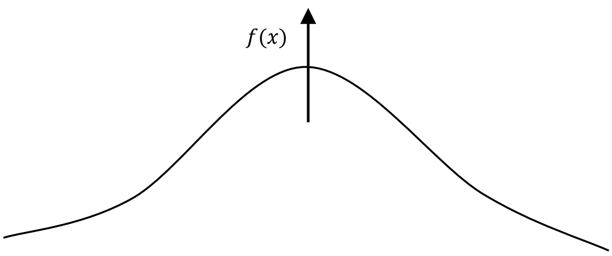
2.2 اختبار ذو طرف أيسر:

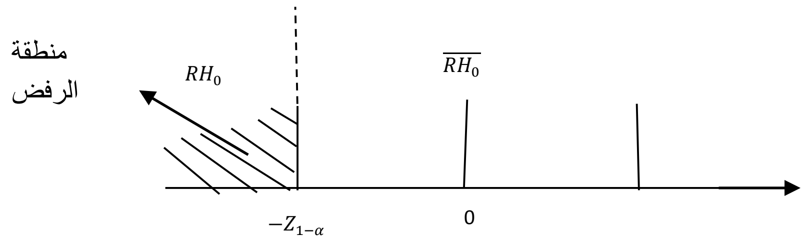
لنفرض أنه تم سحب عينة حجمها $n \geq 30$ من مجتمع يتوزع بمتوسط μ ونريد إختبار الفرضية :

$$H_0: \mu \geq \mu_0$$

$$H_1: \mu < \mu_0$$

إذا كان الإختبار ذو طرف أيسر فإن منطقتي الرفض والقبول تكونان كما هو موضح في الشكل:





عند مستوى دلالة α فإن قرار رفض أو عدم رفض الفرضية المبدئية H_0: يكون وفقا لقاعدة اتخاذ القرار التالية :

$$Z=\frac{\bar{X}-\mu_0}{\sqrt{V(\bar{X})}}\geq -Z_{1-\alpha}\rightarrow \overline{RH_0}$$

$$Z=\frac{\bar{X}-\mu_0}{\sqrt{V(\bar{X})}}< -Z_{1-\alpha}\rightarrow RH_0$$

ملاحظة: إذا كانت $n<30$ فإن قاعدة إتخاذ القرار تكون كما يلي :

$$T=\frac{\bar{X}-\mu_0}{\sqrt{\frac{s^2}{n-1}}}=\frac{\bar{X}-\mu_0}{\frac{\hat{s}}{\sqrt{n}}}\geq -t_{n-1;1-\propto}\rightarrow \overline{RH_0}$$

$$\mathbf{T}=\frac{\bar{\mathbf{X}}-\boldsymbol{\mu}_0}{\sqrt{\frac{\mathbf{s}^2}{\mathbf{n}-\mathbf{1}}}}=\frac{\bar{\mathbf{X}}-\boldsymbol{\mu}_0}{\frac{\hat{\mathbf{s}}}{\sqrt{\mathbf{n}}}}<-\mathbf{t}_{\mathbf{n-1;1}-\propto}\rightarrow RH_0$$

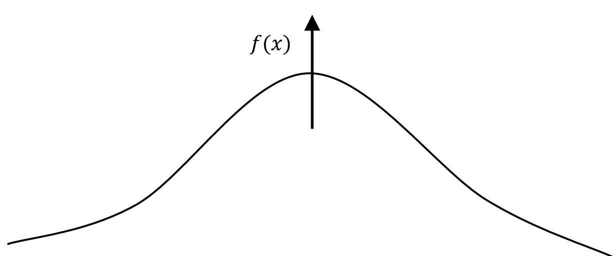
3.2 إختبار ذو طرف أيمن :

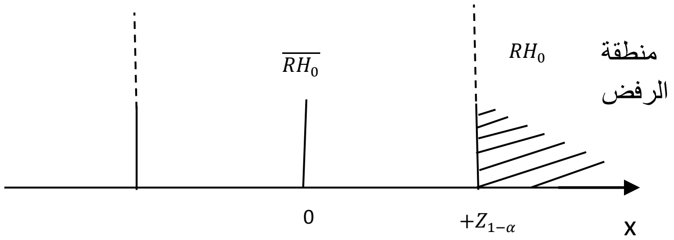
لنفرض أنه تم سحب عينة حجمها $n\geq 30$ من مجتمع يتوزع بمتوسط μ ونريد إختبار الفرضية :

$$H_0: \mu\leq\mu_0$$

$$H_1: \mu>\mu_0$$

إذا كان الإختبار ذو طرف أيمن فإن منطقتي الرفض والقبول تكونان كما هو موضح في الشكل :





عند مستوى دلالة α فإن قرار رفض أو عدم رفض الفرضية المبدئية H_0: يكون وفقا لقاعدة اتخاذ القرار التالية :

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{V(\bar{X})}} \leq +Z_{1-\alpha} \rightarrow \overline{RH_0}$$

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{V(\bar{X})}} > +Z_{1-\alpha} \rightarrow RH_0$$

ملاحظة:

إذا كانت $n < 30$ فإن قاعدة إتخاذ القرار تكون كما يلي :

$$\begin{cases} T = \dfrac{\bar{X} - \mu_0}{\sqrt{\dfrac{s^2}{n-1}}} = \dfrac{\bar{X} - \mu_0}{\dfrac{\hat{s}}{\sqrt{n}}} \leq +t_{n-1;1-\propto} \rightarrow \overline{RH_0} \\ T = \dfrac{\bar{X} - \mu_0}{\sqrt{\dfrac{s^2}{n-1}}} = \dfrac{\bar{X} - \mu_0}{\dfrac{\hat{s}}{\sqrt{n}}} > +t_{n-1;1-\propto} \rightarrow RH_0 \end{cases}$$

4.2 إختبار الفرق بين وسطي مجتمعين: $(\mu_1 - \mu_2)$

بفرض أنه لدينا n_1 و n_2 عينتان مأخوذتان من مجتمعين مستقلين و يخضعان لتوزيع الطبيعي حيث وسط تباين المجتمع الأول هما على التوالي μ_1 و σ_1^2 معلوم و المجتمع الثاني وسطه وسطه μ_2 وتباينه σ_2^2 أيضا معروف، و نرغب إجراء الإختبار للفرق بين $\mu_1 - \mu_2$ لإجراء ذلك نضع المسألة على النحو التالي:

$$H_1: \mu_1 \neq \mu_2 \Leftrightarrow H_0: \mu_1 = \mu_2$$

أي:

$$H_1: \mu_1 - \mu_2 \neq 0 \Leftrightarrow H_0: \mu_1 - \mu_2 = 0$$

نعلم أن توزيع الفرق بين وسطي العينات $\bar{x}_1 - \bar{x}_2$يتبع التوزيع الطبيعي بمتوسط $\mu_1 - \mu_2$و تباين $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ ، ويكون عندئذ معادلة أتخاذ القرار على النحو التالي:

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

عند مستوى دلالة معينة أي α معطاة نتبنى قاعدة القرار التالية:

$$\begin{cases} |\bar{x}_1 - \bar{x}_2| \le z_{\frac{\sigma}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \Rightarrow \overline{RH_0} \\ |\bar{x}_1 - \bar{x}_2| > z_{\frac{\sigma}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \Rightarrow RH_0 \end{cases}$$

ملاحظة1: في حالة σ_1^2 و σ_2^2 مجهولتين، فإن التباينين المحسوبين من العينتين s_1^2 و s_2^2 يعتبرا تقديرين لتبايني المجتمعين، حيث تتغير قاعدة اتخاذ القرار و تصبح:

$$\begin{cases} |\bar{x}_1 - \bar{x}_2| \le z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 2}} \Rightarrow \overline{RH_0} \\ |\bar{x}_1 - \bar{x}_2| > z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 2}} \Rightarrow RH_0 \end{cases}$$

ملاحظة2: هذا النوع من الإختبار يخص حالة العينتين الكبيرتين أي $n_1 \ge 30$ و $n_2 \ge 30$ ، أما إذا كان n_1 و n_2 أقل من 30 وحدة و كان المجتمعين لهما نفس التباين أي $\sigma_1^2 = \sigma_2^2$ ولكنهما مجهولين، فإننا نعلم أن دالة الإختيار هي:

$$L = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = t_{n_1 + n_2 - 2}$$

فعند مستوى الدلالة α لم نرفض H_0إذا كانت:

$$L = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \leq t_{n_1+n_2-2;\, 1-\frac{\alpha}{2}}$$

ونرفض H_0 إذا كان المقدار L أكبر من $t_{n_1+n_2-2;\, 1-\frac{\alpha}{2}}$

وهناك كيفية ثانية لقاعدة إتخاذ القرار التي يمكن كتابتها على النحو التالي :

$$\begin{cases} |\bar{x}_1 - \bar{x}_2| \leq t_{n_1+n_2-2;\, 1-\frac{\alpha}{2}} \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \Rightarrow \overline{RH_0} \\ \text{في حالة العكس} \Rightarrow RH_0 \end{cases}$$

أما إذا كان إختيار ذو جانب واحد أي إذا كانت المسألة:

$H_1: \mu_1 < \mu_2 \Leftrightarrow H_0: \mu_1 \begin{Bmatrix} = \\ \geq \end{Bmatrix} \mu_2$:أو

$H_1: \mu_1 > \mu_2 \Leftrightarrow H_0: \mu_1 \begin{Bmatrix} = \\ \leq \end{Bmatrix} \mu_2$

فعند مستوى الدلالة α لم نرفض H_0 إذا كانت:

$$L = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \leq t_{n_1+n_2-2;1-\alpha}$$

و نرفض H_0 ، إذا كان L أكبر من $t_{n_1+n_2-2;1-\alpha}$

مثال: لمعرفة مدى تأثير فعالية نوع من الأسمدة على المردود الفلاحي، اختيرت 24 قطعة أرضية تحتوي على نفس المساحة، تم إستعمل هذا النوع في نصف الأراضي، فيما يخص الأراضي التي سمدت كان المردود الوسطي في حدود5.1 طن مع إنحراف معياري قدره 0.36 طن. أما الأراضي الباقية غير المسمدة، كان المردود الوسطي 4.8 طن مع إنحراف معياري 0.4 طن .

المطلوب: إختبر صحة الفرضية $H_0: \mu_1 \geq \mu_2 \;\Leftrightarrow\; H_1: \mu_1 < \mu_2$ عند مستوى دلالة $\alpha = 1\%$ ؟

الحل:

$$n_1 = 12\, ;\, \bar{x}_1 = 5.1\, ;\, s_1 = 0.36$$

$$n_2 = 12\,; \overline{x}_2 = 48\,; s_2 = 0.4; \alpha = 0.01$$

قاعدة إتخاذ القرار تكون على النحو التالي :

$$\begin{cases} |\overline{x}_1 - \overline{x}_2| \le t_{n_1+n_2-2;1-\alpha}\sqrt{\dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)} \Rightarrow \overline{RH_0} \\ \text{في حالة العكس} \Rightarrow RH_0 \end{cases}$$

$$|\overline{x}_1 - \overline{x}_2| = |5.1 - 4.8| = 0.3$$

$$t_{n_1+n_2-2;1-\alpha} = t_{22;0.99} = 2.508$$

$$\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{\frac{12(0.36)^2 + 12(0.4)^2}{22}\left(\frac{1}{12}+\frac{1}{12}\right)} = 0.162$$

$$L = 2.508.(0.162) = 0.4062$$

بما أن: $|\overline{x}_1 - \overline{x}_2| = 0.3 < L = 0.40620$

القرار يكون بعدم رفض H_0 أي معناه أن إختلاف المردود الفلاحي لا يعود إلى مجرد الصدفة بل هنالك تأثير فعلي لهذا النوع من الأسمدة عل المردود الفلاحي.

مثال 2 :تم إجراء إمتحان لمجموعة من الطلبة الذين ينتمون إلى فرعين مختلفين، حث يحتوي الفرع الأول على 40 طالب، على عكس الفرع الثاني الذي فيه 50 طالب. كانت النتائج على النحو التالي :

$$s_2 = 7 \;\; ; \;\; \overline{x}_2 = 78 \;\; ; \;\; s_1 = 8 \;\; ; \;\; \overline{x_1} = 74$$

المطلوب: هل هنالك فرق ملموس بين النتائج الوسطية لكلا الفرعين عند مستوى الدلالة 5%؟

الحل:

$$n_1 = 40\,; \overline{x}_1 = 74\,; s_1 = 8$$

$$n_2 = 50\,; \overline{x}_2 = 78\,; s_2 = 7; \alpha = 5\%$$

$$H_1: \mu_1 \neq \mu_2 \;\Leftrightarrow\; H_0: \mu_1 = \mu_2$$

نتبع القاعدة التالية فيما يخص اتخاذ القرار:

$$\begin{cases} |\bar{x}_1 - \bar{x}_2| \leq Z_{\frac{\alpha}{2}}\sqrt{\dfrac{s_1^2}{n_1-1} + \dfrac{s_2^2}{n_2-1}} \Rightarrow \overline{RH_0} \\ \text{في حالة العكس} \Rightarrow RH_0 \end{cases}$$

$$|\bar{x}_1 - \bar{x}_2| = |74 - 78| = 4$$

$$L = Z_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1-1} + \frac{s_2^2}{n_2-1}} = 1.96\sqrt{\frac{64}{39} + \frac{49}{49}} = 3.19$$

بما أن: $|\bar{x}_1 - \bar{x}_2| = 4 > L = 3.19$

القرار يكون رفض H_0أي معناه يوجد فرق ملموس بين النتائج الوسطية بين الفرعين.

3.إختبار نسبةP:

إذا كان الإختبار ذو الطرفين كما في الفرضية التالية :

$$H_0: P = P_0$$

$$H_1: P \neq P_0$$

فعند مستوى دلالة α يكون قرار رفض أو عدم رفض الفرضية المبدئية H_0 وفقا لقاعدة إتخاذ القرار التالية:

$$\begin{cases} Z = \dfrac{f - P_0}{\sqrt{V(f)}} \in \left[-Z_{1-\frac{\alpha}{2}}; +Z_{1-\frac{\alpha}{2}}\right] \rightarrow \overline{RH_0} \\ Z = \dfrac{f - P_0}{\sqrt{V(f)}} \notin \left[-Z_{1-\frac{\alpha}{2}}; +Z_{1-\frac{\alpha}{2}}\right] \rightarrow \overline{RH_0} \end{cases}$$

مع العلم أن :

$$V(f) = \frac{P_0(1 - P_0)}{n}$$

ملاحظة1:

إذا كان الإختبار ذو الطرف أيسر فعند المستوى دلالة معينة α يبنى القرار حسب قاعدة اتخاذ القرار التالية:

$$\begin{cases} Z = \dfrac{f - P_0}{\sqrt{V(f)}} \geq -Z_{1-\alpha} \rightarrow \overline{RH_0} \\ Z = \dfrac{f - P_0}{\sqrt{V(f)}} < -Z_{1-\alpha} \rightarrow RH_0 \end{cases}$$

ملاحظة2: إذا كان الإختبار ذو الطرف أيمن فعند المستوى دلالة معينة α يبنى القرار حسب قاعدة اتخاذ القرار التالية:

$$\begin{cases} Z = \dfrac{f - P_0}{\sqrt{V(f)}} \leq +Z_{1-\alpha} \rightarrow \overline{RH_0} \\ Z = \dfrac{f - P_0}{\sqrt{V(f)}} > +Z_{1-\alpha} \rightarrow RH_0 \end{cases}$$

مثال: أعلنت مؤسسة أن 95% من منتجاتها صالحة للإستعمال طبقا للمعايير المعلول بها عالميا، عند دراسة لعينة حجمها 200 إتضح أن 18 منتوج غير مقبول.

المطلوب: إختبر الفرضية التالية عند مستوى دلالة 5%:

$$H_0 : P = 0.06$$

$$H_1 : P \neq 0.06$$

الحل:

لدينا: $\alpha = 0.05$ $\qquad n = 200$

$$r = 18 \rightarrow f = \frac{r}{n} \rightarrow \frac{18}{200} \rightarrow f = 0.09$$

قاعدة إتخاذ القرار هي:

$$\begin{cases} Z = \dfrac{f - P_0}{\sqrt{V(f)}} \in \left[-Z_{1-\frac{\propto}{2}}; +Z_{1-\frac{\propto}{2}}\right] \rightarrow \overline{RH_0} \\ Z = \dfrac{f - P_0}{\sqrt{V(f)}} \notin \left[-Z_{1-\frac{\propto}{2}}; +Z_{1-\frac{\propto}{2}}\right] \rightarrow RH_0 \end{cases}$$

لدينا:

$$Z_{1-\frac{\propto}{2}} = Z_{0.025} = Z_{0.975} = 1.96$$

$$\sqrt{\frac{P_0(1-P_0)}{n}} = \sqrt{\frac{0.06 \times 0.94}{200}} = 0.017$$

ومنه:

$$Z = \frac{f - P_0}{\sqrt{V(f)}} = \frac{0.09 - 0.06}{0.017} = 1.7647$$

$$\left[-Z_{1-\frac{\alpha}{2}}; +Z_{1-\frac{\alpha}{2}}\right] = [-1.96; +1.96]$$

و بما أن

$$1.7647 \in [-1.96; +1.96]$$

فإننا لا نرفض H_0، أي $\overline{RH_0}$ عند مستوى المعنوي 5% أي أننا نقبل أن نسبة المنتوج غير المقبول 6%.

4.إختبار الفرق بين نسبتين لمجتمعين $(p_1 - p_2)$:

نفرض أنه لدينا عينتان n_2 و n_1 كبيرتان مستقلتان عن بعضها البعض، وإذا كانت p_1 تمثل النسبة في المجتمع الأول و p_2 النسبة في المجتمع الثاني و $\hat{p}_2$ و $\hat{p}_1$ النسبتين في العينتين فإن الفرق بين $\hat{p}_1 - \hat{p}_2$ يخضع لتوزيع الطبيعي ذو الوسط $p_1 - p_2$ و تباين $\frac{P_2q_2}{n_1} + \frac{P_1q_1}{n_1}$ بحيث نكتب:

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}} = N(0.1)$$

وإذا أردنا عند مستوى الدلالة α أن نختبر:

$$H_1: p_1 \neq p_2 \Leftrightarrow H_0: p_1 = p_2$$

أو : $H_1: p_1 - p_2 \neq 0 \Leftrightarrow H_0: p_1 - p_2 = 0$

فإذا كانت:

$$|\hat{p}_1 - \hat{p}_2| \leq Z_{\frac{\alpha}{2}}\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}$$

لم نرفض H_0 أي $\overline{RH_0}$ أما إذا كانت:

$$|\hat{p}_1 - \hat{p}_2| > Z_{\frac{\alpha}{2}}\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}$$

فالقرار يكون رفض H_0 أي : RH_0

عند إختيار ذو جانب واحد للفرق بين نسبتين نكتب:

$$H_1: P_1 > P_2 \Leftrightarrow H_0: P_1 \begin{Bmatrix} = \\ \leq \end{Bmatrix} P_2$$

أو : $H_1: P_1 < P_2 \Leftrightarrow H_0: P_1 \begin{Bmatrix} = \\ \geq \end{Bmatrix} P_2$

فإننا لا نرفض $H_0\left(\overline{RH_0}\right)$ على مستوى الدلالة إذا كانت:

$$|\hat{p}_1 - \hat{p}_2| \leq Z_{\alpha}\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}$$

و نرفض H_0 أي (RH_0) عند

$$|p_1 - p_2| > Z_{\alpha}\sqrt{\frac{P_1q_1}{n_1} + \frac{P_2q_2}{n_2}}$$

مثال: عند دراسة لمجموعتين تحتوي كل واحدة منهما على 1200 فرد مصابين بمرض الإسهال. أعطي نوع من الدواء للمجموعة الأولى ونوع ثاني للمجموعة الثانية، وجد أن في المجموعة الأولى قد شفي 196 مريضا، على عكس المجموعة الثانية شفي فيها 135 مصابا.

المطلوب: إختبر الفرضية أن نوع الدواء الأول يساعد على الشفاء من الإسهال بإستخدام مستوى الدلالة 5%.

الحل:

$$n_1 = n_2 = 1200 \text{ ، } r_1 = 196 \text{ ،} r_2 = 135\text{،}\ \alpha = 0.05$$

نختبر المسألة التالية:

$$H_1: P_1 > P_2 \Leftrightarrow H_1: P_1 \leq P_2$$

$$\hat{p}_1 = \frac{r_1}{n_1} = \frac{196}{1200} = 0.1633$$

$$\hat{p}_2 = \frac{r_2}{n_2} = \frac{135}{1200} = 0.1125$$

$$|\hat{p}_1 - \hat{p}_2| = |0.1633 - 0.1125| = 0.0508$$

$$Z_\alpha \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 1.645\sqrt{\frac{(0.1633)(0.8367)}{1200} + \frac{(0.1125)(0.8875)}{1200}} = 0.023$$

وبما أن: $|\hat{p}_1 - \hat{p}_2| = 0.0508 > 0.023$

القرار يكون برفض H_0 أي RH_0 ونقبل H_1 الذي ينص أن الدواء الأول هو الذي يساعد أكثر في شفاء المصابين بالإسهال.

مثال2: عند سحب عينتان مستقلتان كل منهما تحتوي 400 فردا، فوجد أن نسبة المدخنين في العينتين 0.7 و 0.6 على الترتيب.

المطلوب: عند مستوى دلالة 5% إختبر:

أ: $H_1: P_1 \neq P_2 \Leftrightarrow H_0: P_1 = P_2$

ب: $H_1: P_1 > P_2 \Leftrightarrow H_0: P_1 \leq P_2$

الحل:

$$n_1 = n_2 = 400، \hat{p}_1 = 0.07، \hat{p}_2 = 0.6، \alpha = 0.05$$

أ: $\mathbf{H_1: P_1 \neq P_2 \Leftrightarrow H_0: P_1 = P_2}$:

$$|\hat{p}_1 - \hat{p}_2| = |0.7 - 0.6| = 0.1$$

$$Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = 1.96\sqrt{\frac{(0.7)(0.3)}{400} + \frac{(0.6)(0.4)}{400}} = 1.96 \times 0.034 = 0.0666$$

وبما أن: $|\hat{p}_1 - \hat{p}_2| = 0.1 > 0.0666$

نرفض H_0 ونقبل الفرضية البديلة أي معناه أنه يوجد إختلاف ملموس بين المجموعتين.

ب: $H_1: P_1 > P_2 \Leftrightarrow H_0: P_1 \leq P_2$

إذا كانت :

$$L = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}} \leq [-Z_\alpha; +Z_\alpha]$$

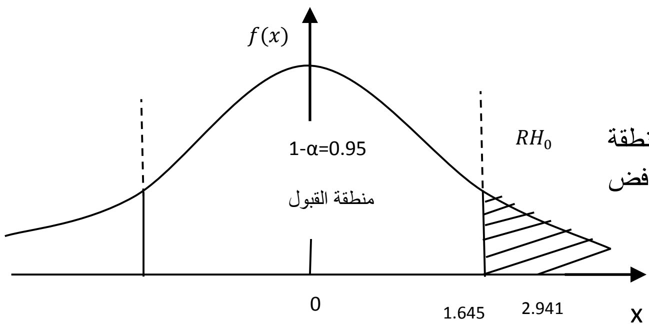
فإننا لا نرفض H_0 عند مستوى الدلالة 5%، وإذا كانت L أكبر من $[-Z_\alpha; +Z_\alpha]$ نرفض H_0.

$$\frac{|0.7 - 0.6|}{\sqrt{\frac{(0.7)(0.3)}{400} + \frac{(0.6)(0.4)}{400}}} = \frac{0.1}{0.034} = 2.941$$

بما أن:

$$2.941 \notin [-1.645, 1.645]$$

القرار يكون أيضا الرفض لـ: H_0 وقبول H_1 أي معناه نسبة المدخنين في المجتمع الأول تفوق نسبة المدخنين في المجتمع الثاني.



5. إختبار لتباين المجت

نعلم من الفصل السابق أن مجال الثقة للتباين المجتمع σ^2 يكتب على النحو التالي:

$$\left[\frac{ns^2}{X^2_{n-1;1-\frac{\alpha}{2}}}; \frac{ns^2}{X^2_{n-1;\frac{\alpha}{2}}}\right] = 1 - \alpha$$

وهذا يعني أن قيمة σ^2 تقع بين حدي هذا المجال الذي يساوي $1 - \alpha$ لنختبر

$$H_1: \sigma^2 \neq \sigma_0^2 \Leftrightarrow H_0: \sigma^2 = \sigma_0^2$$

بحيث σ_0^2 تمثل قيمة إفتراضية لتباين المجتمع ويتم اختبارها. عند مستوى دلالة معين أي α معطاة، قاعدة اتخاذ القرار تكون حسب إحدى الكيفيتين المعروضتين على النحو التالي:

$$\begin{cases} \dfrac{ns^2}{\sigma^2} \in \left[X^2_{n-1;\frac{\alpha}{2}};\ X^2_{n-1;1-\frac{\alpha}{2}}\right] \Leftrightarrow \overline{RH_0} \\ \text{في حالة العكس} \Leftrightarrow RH_0 \end{cases}$$

أما الكيفية الثانية و التي تعتمد على مجال الثقة تكتب:

$$\sigma_0^2 \in \left[\frac{ns^2}{X^2_{n-1;1-\frac{\alpha}{2}}};\frac{ns^2}{X^2_{n-1;\frac{\alpha}{2}}}\right] \Leftrightarrow \overline{RH_0}$$

ملاحظة: في حالة إختبار ذو جانب واحد بالنسبة للتباين المجتمع تصيغ المسألة على النحو التالي:

$$H_1: \sigma^2 < \sigma_0^2 \ \Leftrightarrow\ H_0: \sigma^2 \begin{Bmatrix} = \\ \geq \end{Bmatrix} \sigma_0^2$$

$$H_1: \sigma^2 > \sigma_0^2 \ \Leftrightarrow\ H_0: \sigma^2 \begin{Bmatrix} = \\ \leq \end{Bmatrix} \sigma_0^2$$

أما قاعدة إتخاذ القرار تكون:

$$\begin{cases} \dfrac{ns^2}{\sigma^2} \in \left[X^2_{n-1;\alpha};\ X^2_{n-1;1-\alpha}\right] \Leftrightarrow \overline{RH_0} \\ \text{في حالة العكس} \Leftrightarrow RH_0 \end{cases}$$

أو الكيفية الثانية التي تعتمد على مجال الثقة:

$$\begin{cases} \sigma_0^2 \in \left[\dfrac{ns^2}{X^2_{n-1;1-\alpha}};\dfrac{ns^2}{X^2_{n-1;\alpha}}\right] \Leftrightarrow \overline{RH_0} \\ \text{في حالة العكس} \Leftrightarrow RH_0 \end{cases}$$

مثال: ينتج مصنع للأغذية نوع من المعلبات، قام المصنع بمراقبة دورية لعينة تحتوي على 16 نوع من الأغذية ووجد أن الإنحراف المعياري يساوي 5.5 .

-إستعمل هذه المعلومات لإختبار:

أ: $H_0: \sigma^2 = 20 \ \Leftrightarrow\ H_1: \sigma^2 \neq 20$

ب: $H_1: \sigma^2 < 40 \Leftrightarrow H_0: \sigma^2 \geq 400$

عند مستوى الدلالة يساوي 5%.

الحل:

$$n = 20 \; ; \; s^2 = 30.25 \; ; \; \alpha = 5\% \; ;$$

$$H_1: \sigma^2 \neq 20 \Leftrightarrow H_0: \sigma^2 = 20$$

الحل بالكيفية الأولى:

$$\begin{cases} \dfrac{ns^2}{\sigma_0^2} \in \left[X^2_{n-1;\frac{\alpha}{2}};\ X^2_{n-1;1-\frac{\alpha}{2}}\right] \Leftrightarrow \overline{RH_0} \\ \text{في حالة العكس} \Leftrightarrow RH_0 \end{cases}$$

$$\frac{ns^2}{\sigma_0^2} = \frac{16 \times 30.25}{20} = 24$$

$$X^2_{n-1;\frac{\alpha}{2}} = X^2_{15;0.025} = 27.49$$

$$X^2_{n-1;1-\frac{\alpha}{2}} = X^2_{15;0.9750} = 6.26$$

و بما أن: $\frac{ns^2}{\sigma_0^2} = 24 \in [6.26;\ 27.49]$

القرار يكون عدم رفض H_0 أي : $\overline{RH_0}$

الكيفية الثانية:المشتقة من مجال الثقة:

$$\sigma_0^2 = 20$$

نعلم أن كتابة المجال تكون:

$$\left[\frac{ns^2}{X^2_{n-1;\frac{\alpha}{2}}};\ \frac{ns^2}{X^2_{n-1;1-\frac{\alpha}{2}}}\right]$$

$$\left[\frac{16 \times 30.25}{27.19};\ \frac{16 \times 30.25}{6.26}\right]$$

بما أن: $\sigma_0^2 = 20 \in [17.61;\ 77.32]$

إذن القرار يكون دائما عدم رفض H_0 أي : $\overline{RH_0}$

ب: $H_0:\ \sigma^2 \geq 40 \Leftrightarrow H_1: \sigma^2 < 40$:

يلاحظ هنا أن نوع الإختيار هو إختيار ذو جانب واحد .

نحسب:

$$\frac{ns^2}{\sigma_0^2} = \frac{16 \times 30.25}{40} = 12.1$$

$$X^2_{n-1;\alpha} = X^2_{n-1;0.05} = X^2_{15;0.05} = 25$$

$$X^2_{n-1;1-\alpha} = X^2_{n-1;0.95} = X^2_{15;0.95} = 7.26$$

بما أن: $\frac{ns^2}{\sigma_0^2} = 12.1 \in [7.26;\ 25]$

القرار يكون عدم رفض H_0 أي : $\overline{RH_0}$

نؤكد النتيجة عندما نعتمد الكيفية الثانية:

$$\sigma_0^2 = 40$$

المجال:

$$\left[\frac{16 \times 30.25}{25};\ \frac{16 \times 30.25}{7.26}\right]$$

بما أن: $\sigma_0^2 = 40 \in [19.36; 66.66]$

إذن القرار يكون دائما بعدم رفض H_0 أي : $\overline{RH_0}$

6. إختبار تباينين:

إذا كان لدينا عينة n_1 من توزيع طبيعي $N(n_1, \sigma_1^2)$ و عينة n_2 من توزيع طبيعي $N(n_2, \sigma_2^2)$ مستقلة عن الأولى، و نرغب في وضع الإختبار التالي:

$$H_1: \sigma_1^2 \neq \sigma_2^2 \Leftrightarrow H_0: \sigma_1^2 = \sigma_2^2$$

أي:

$$H_1: \sigma_1^2 - \sigma_2^2 \neq 0 \Leftrightarrow H_0: \sigma_1^2 - \sigma_2^2 = 0$$

عند مستوى دلالة معين، أي معطاة لم نرفض H_0 أي $\overline{RH_0}$

$$\frac{s_1^2}{s_2^2} \in \left[F_{\frac{\alpha}{2};n_1-1,n_2-2}; F_{1-\frac{\alpha}{2};n_1-1,n_2-2}\right] \Leftrightarrow \overline{RH_0}$$

و نرفض H_0 أي RH_0 في حالة العكس.

فيما يخص إختبار ذو جانب واحد نكتب الفرضيات على الكيفية التالية:

$$H_1: \sigma_1^2 < \sigma_2^2 \Leftrightarrow H_0: \sigma_1^2 \left\{\begin{matrix} = \\ \geq \end{matrix}\right\} \sigma_2^2$$

$$H_1: \sigma_1^2 > \sigma_2^2 \Leftrightarrow H_0: \sigma_1^2 \left\{\begin{matrix} = \\ \leq \end{matrix}\right\} \sigma_2^2$$

فإننا لا نرفض H_0 أي $\overline{RH_0}$ على مستوى الدلالة α إذا كانت:

$$\frac{s_1^2}{s_2^2} \in \left[F_{\alpha;n_{1-1},n_{2-1}}; F_{1-\alpha;n_{1-1},n_{2-1}}\right]$$

ونرفض H_0 إذا $\frac{s_1^2}{s_2^2}$لا تنتمي إلى مجال .

مثال: أخذت عينة عشوائية حجمها 21 فكان تباينها 62 من مجتمع له توزيع طبيعي، أخذت عينة عشوائية أخرى مستقلة عن الأولى حجمها 16 فكان تباينها 32 من مجتمع له توزيع طبيعي.

المطلوب: هل هناك فرق بين تبايني المجتمعين عند مستوى الدلالة 0.05؟

الحل:

لدينا
$$\begin{matrix} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{matrix}$$

نلاحظ بأن الفرضية البديلة ذات طرفين وقيمها الحرجة هي:

$$F_{\frac{\alpha}{2};n_{1-1},n_{2-1}} = F_{(0.025;20.15)} = 2.76$$
$$F_{1-\frac{\alpha}{2};n_{1-1},n_{2-1}} = F_{(0.975;20.15)} = 0.39$$

$$F = \frac{s_1^2}{s_2^2} \times \frac{\sigma_1^2}{\sigma_2^2} = \frac{62}{32} = 1.937$$

بالمقارنة نجد أن: $2.76 > 1.937 > 0.39$ إذن نقبل H_0 ونرفض H_1 عند مستوى الدلالة 0.05 ونقول بأن تبايني المجتمعين متساويين عند مستوى المعنوية 0.05.

تمارين السلسلة الرابعة

التمرين الأول: تدعي إدارة أحد المصانع لصناعة نوع معين من الأنابيب المعدنية بأن الوسط الحسابي لطول قطر الأنابيب المصنعة في هذا المصنع مطابق للمواصفات و يساوي 2 سم، وللتأكد من صحة قوله سحبت عينة عشوائية من الإنتاج الكلي تحتوي على 35 أنبوب معدني فكانت أطول أقطارها كما يلي:

2.1	1.89	1.94	1.97	1.99	1.95	1.85	1.84	1.9	2
1.9	2.02	2.04	2.05	1.9	1.95	1.96	1.97	2	2.11
1.9	1.9	1.89	2.12	2.09	2.07	2.06	1.98	1.94	1.93
					2.03	1.98	1.92	2.1	2.08

حيث: $\sum_{i=1}^{35}(xi-\bar{x})^2 = 2.7335$

إذا علمت أن أطول أقطار الأنابيب تتوزع طبيعيا. اختبر صحة إدعاء إدارة هذا المصنع عند مستوى معنوية 0.05.

التمرين الثاني: يدعي مدير مصنع لإنتاج الدقيق أن متوسط أوزان أكياسه هو 25 كلغ أو أكثر و لكن توجد شكاوى من الزبائن مفادها أن متوسط أوزان الأكياس أقل من 25كغ، لاختبار هذه الشكوى سحبت عينة من 25 كيس من الإنتاج ووجد أن متوسط أوزانها 24.75 كغ وتباينها 0.49 كغ. فإذا كان وزن كيس الدقيق هو متغير عشوائي يتبع التوزيع الطبيعي.

إختبر صحة شكوى الزبائن عندى مستوى المعنوية $\alpha = 0.1$

التمرين الثالث: أظهرت سجلات مديرية الأمن العام لولاية سطيف أن نسبة مستعملي حزام الأمان في السيارات (قبل تشريع إلزام الإستعمال) هي55%. وبعد صدور تشريع الإلزام، اختبرت عينة عشوائية حجمها 100 سائق، فوجد أن 72 منهم يستعملون الحزام. اختبر عندى مستوى معنوية 5%ثم 1% فيما إذا كان صدور التشريع قد زاد نسبة المستعملين لحزام الأمان. ماذا تستنتج؟

التمرين الرابع: قامت إحدى الشركات باستيراد شحنة كبيرة من الأجهزة الكهرومنزلية وقد تعهدت الشركة المصدرة للأجهزة بأن لا تزيد نسبة الأجهزة المعيبة في الشحنة عن 3%، تم اختيار عينة عشوائية من 50 جهاز من هذه الشحنة فتبين وجود 2 جهاز معيب.

1- قدر بمجال نسبة الأجهزة المعيبة في الشحنة بثقة قدرها 99%؟

2- ما هو مقدار الخطأ المحتمل ارتكابه في تقدير نسبة الأجهزة المعيبة في الشحنة بثقة قدرها 99%؟

3- ما هو حجم العينة اللازم إذا أردنا تحسين دقة النتائج المحصل عليها في السؤال الأول بـ40%؟

4- هل يمكن القول عند مستوى معنوية 0.01 أن الشركة المصدرة للأجهزة الكهرومنزلية قد التزمت بتعهدها.

التمرين الخامس: تلقت مؤسسة شحنة من الإطارات المطاطية للسيارات و ترغب في تقدير متوسط مدة حياة هذه الإطارات لأجل ذلك تم سحب عينة مكونة من 25 إطار مطاطي فتم التوصل إلى أن مدة حياتها مجتمعة تساوي 987000 كلم علما أن $\sum_{i=1}^{25}(x_i-\bar{x})^2 = 153600$. بافتراض أن مدة حياة الإطارات عبارة عن متغير عشوائي يتبع التوزيع الطبيعي.

1- قدر بمجال متوسط مدة حياة هذه الإطارات عند مستوى دلالة 5%.

2- إذا كان $\mu = 42000$ أحسب الاحتمال التالي: $P(\bar{X} > 42000)$.

3- تدعي المؤسسة المصنعة لهذه الإطارات أن متوسط مدة حياة الإطارات المطاطية في الشحنة هي 42000 كلم، اختبر مدى صحة هذا الادعاء عند مستوى ثقة 95%؟

التمرين السادس: لنفرض أن مجتمعا من المرضى بأحد المراكز الاستشفائية بالجزائر يتكون من 250 مريض، اخذت عينة عشوائية تتكون من 50 مريض، وجد من بينهم 10 مرضى من المدخنين.

1- تقدير نسبة المدخنين في هذا المجتمع.

2- إذا كانت نسبة المدخنين في هذا المجتمع هي 30% أحسب احتمال أن تكون نسبة المدخنين في العينة على الأكثر 40%.

3- احسب حجم العينة اللازم حتى لا يتجاوز خطأ المعاينة في تقدير P=0.1 وذلك عند $\alpha = 5\%$

4- نفرض أنه عند مستوى دلالة $\alpha = 5\%$ تمت صياغة المسألة التالية:

$$H_0: P \geq 0.5 \quad , \quad H_1: P < 0.5$$

التمرين السابع: تم استصلاح عدد كبير من القطع الزراعية ولمعرفة مستوى إنتاجيتها من القمح بعد عملية الاستصلاح، اختيرت عينة عشوائية من 30 قطعة تمت زراعتها بالقمح و كانت إنتاجيتها بالطن كما يلي:

7.2	6	5	5	4.3	7	6.5	6.3	8	5.2
6	6	7.3	6.2	5	5.3	8.2	7	6	7
6.3	5	7.2	6.4	5	8	5.6	7.4	5.5	6.6

علما أن: $\sum_{i=1}^{30}(x_i - \bar{x})^2 = 30.915$ و المتغير X يتوزع توزيعا طبيعيا.

1- أوجد التقدير النقطي لتباين إنتاجية القطع الأرضية المستصلحة؟

2- أوجد مجال الثقة لتباين إنتاجية القطع الزراعية بمستوى ثقة قدره: 95%.

3- إذا كان $\mu = 7$ احسب احتمال أن يفوق متوسط إنتاجية القطعة 7.4 طن.

التمرين الثامن: أخذت عينة عشوائية حجمها 10 مصابيح من النوع الأول فكان متوسطها الحسابي هو 630 ساعة إنارة من مجتمع له توزيع طبيعي وسطه μ_1 و إنحرافه المعياري 20 ساعة إنارة، وأخذت عينة عشوائية أخرى مستقلة عن الأولى حجمها 16 مصابيح من النوع الثاني فكان متوسطها الحسابي هو 650 ساعة إنارة، وذلك من مجتمع آخر يتبع التوزيع الطبيعي وسطه μ_2 و إنحرافه المعياري 40 ساعة إنارة،

المطلوب: إذا علمت أن المجتمعان يتبعان التوزيع الطبيعي، هل يمكن القول أن متوسط عمر المصابيح من النوع الأول أقل من متوسط عمر المصابيح من النوع الثاني عند مستوى دلالة 5%؟

التمرين التاسع: أخذت أربعة قراءات بجهاز معين فكانت كما يلي: 59، 55 ، 51 ، 51

المطلوب:اختبر الفرضية الصفرية $H_0: \sigma^2 = 0.7$ في المقابل الفرضية البديلة $H_0: P > 0.7$ وذلك عند مستوى دلالة 1%؟

التمرين العاشر: تعاقدت شركة لصناعة السيارات لشراء مدخرات كهربائية من أحد المصانع يدعي أن تباين منتجاته 0.6 عام، و عند ورود الشحنة سحبت عينة عشوائية من 5 مدخرات فكان تباينها 0.8 عام.

المطلوب: هل تقبل إدعاء الشركة عند مستوى الدلالة 0.05؟

التمرين الحادي عشر: سحبت عينتين عشوائيتين حجمهما على التوالي: 10 ، 25 من مجتمعين و كان الانحراف المعياري للعينتين هو: 15 ، 20 على الترتيب.

المطلوب: اختبر الفرضية التالية:

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_0: \sigma_1^2 > \sigma_2^2$$

وذلك عند مستوى دلالة 0.01؟

التمرين الثاني عشر: سحبنا من مجتمعين طبيعيين عينتين حجم الأولى $n_1 = 51$ وحجم الثانية $n_2 = 67$ و مستقلتين، وجدنا النتائج التالية.

$$S_1^2 = 13 \;;\; S_2^2 = 11 \;;\; \overline{X_1} = 101 \;;\; \overline{X_2} = 96$$

كيف يمكن إجراء اختبار تساوي تبايني المجتمعين بمستوى معنوية 0.05؟

التمرين الثالث عشـر: أخذت 400 قطعة من إنتاج الآلة(أ) فتبين أن 20 قطعة منها مخالفة للمواصـفات، و أخذت عينة أخرى من 400 قطعة من إنتاج الآلة(ب) فتبين أن 16 قطعة منها مخالفة للمواصفات.

المطلوب: هل نستطيع أن نستنتج أن هناك فرقا بين نسبة الإنتاج المخالفة للمواصفات للآلتين عند مستوى دلالة 0.05؟

التمرين الرابع عشر: في استطلاع للرأي العام لولايتين حول مرشح سياسي تبين أن 20% من بين 60 شخصا في الولاية (أ) يفضلون هذا المترشح، بينما في الولاية (ب) وجد 30% من بين 90 شخصا يفضلون هذا المترشح.

المطلوب: هل يمكن القول بأن نسبة الأشخاص المؤيدين للمترشح السياسي في الولاية (أ) أقل من الولاية (ب) عند مستوى دلالة 0.05؟

حل السلسلة رقم 04

التمرين الأول:

اختبار صحة إدعاء إدارة هذا المصنع:

1) صياغة الفرضيات:

$$\begin{cases} H_0 : \mu = 2 \\ H_1 : \mu \neq 2 \end{cases}$$

2) اختبار الفرضيات :

بما أن $n = 35 \geq 30$ والاختبار ذو جانبين فان قاعدة اتخاذ القرار تكون :

$$\begin{cases} z = \dfrac{\overline{X} - \mu_0}{\sqrt{\dfrac{S^2}{n-1}}} \in \left[-z_{1-\frac{\alpha}{2}}; z_{1-\frac{\alpha}{2}}\right] \Rightarrow \overline{RH_0} \\ z = \dfrac{\overline{X} - \mu_0}{\sqrt{\dfrac{S^2}{n-1}}} \notin \left[-z_{1-\frac{\alpha}{2}}; z_{1-\frac{\alpha}{2}}\right] \Rightarrow RH_0 \end{cases}$$

لدينا:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{69.32}{35} = 1.98$$

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})}{n} = \frac{2.7335}{35} = 0.0781$$

$$z = \frac{1.98 - 2}{\sqrt{\dfrac{0.0781}{34}}} = -0.4173$$

$$z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$$

بما أن: $-0.4173 \in [-1.96, +1.96] \Rightarrow \overline{RH_0}$

أي أن إدعاء إدارة هذا المصنع عند $\alpha = 0.05$ صحيح

التمرين الثاني:

اختبار صحة شكوى الزبائن عند $\alpha = 0.01$

1) صياغة الفرضيات :

$$\begin{cases} H_0 : \mu \geq 25 \\ H_1 : \mu < 25 \end{cases}$$

2) اختبار الفرضيات :

بما أن $n = 25$ و الإختبار ذو جانب أيسر فإن قاعدة اتخاذ القرار تكون:

$$\begin{cases} T = \dfrac{\overline{X} - \mu_0}{\sqrt{\dfrac{S^2}{n-1}}} \geq -t_{1-\frac{\alpha}{2},n-1} \Rightarrow \overline{RH_0} \\ T = \dfrac{\overline{X} - \mu_0}{\sqrt{\dfrac{S^2}{n-1}}} < -t_{1-\frac{\alpha}{2},n-1} \Rightarrow RH_0 \end{cases}$$

لدينا:

$$T=\frac{24.75-25}{\sqrt{\frac{0.49}{24}}}\geq -1.7496$$

$$t_{1-\alpha,n-1}=t_{1-0.01,25-1}=t_{0.99,24}=2.492$$

ومنه $-t_{1-\alpha,n-1}=-1.32$

بما أن: $-1.7496<-1.32$

فإن قرار الرفض $H_0(RH_0)$ أي أن شكوى الزبائن صحيحة.

التمرين الثالث :

1) صياغة الفرضيات:

$$H_0: P\leq 0.55$$
$$H_1: P>0.55$$

2) اختبار الفرضيات$\alpha=0.05\%$ بما ان الاختبار ذو جانب ايمن فان قاعدة اتخاذ القرار تكون كمايلي :

$$\begin{cases} Z=\frac{f-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}\leq z_{1-\alpha}\Rightarrow \overline{RH_0} \\ Z=\frac{f-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}>z_{1-\alpha}\Rightarrow RH_0 \end{cases}$$

لدينا:

$$f=\frac{72}{100}=0.72$$

$$Z=\frac{0.72-0.55}{\sqrt{\frac{(0.55)(0.45)}{100}}}=3.41$$

$$z_{1-\alpha}=z_{1-0.05}=z_{0.95}=1.6449$$

بما أن $3.41>1.6449\Rightarrow RH_0$

أي أن صدور التشريع زاد نسبة المستعملين لحزام الأمان عند$\alpha=5\%$

اختبار الفرضيات عند: $\alpha=1\%$

لدينا نفس قاعدة الاختبار السابقة حيث:

$$Z=3.41$$
$$z_{1-\alpha}=z_{1-0.01}=z_{0.99}=2.3263$$

بما أن: $3.41>2.3263$

فان القرار هو رفض$H_0(RH_0)$

و عليه فصدور التشريع زاد نسبة المستعملين لحزام الأمان.

حل التمرين الثامن:

لدينا:

$$n_1 = 10\,; \overline{X}_1 = 630\,; \sigma_1 = 20$$
$$\alpha = 5\%\,; n_2 = 16\,; \overline{X}_2 = 650\,; \sigma_2 = 40$$

ونريد اختيار:

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

نلاحظ أن الفرضية البديلة لها اتجاه واحد(طرف أيسر) وقيمتها الحرجة هي:

$$-Z_\alpha = -Z_{0.05} = -1.65$$

$$P\left(\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq \frac{630 - 650}{\sqrt{\frac{(20)^2}{10} + \frac{(40)^2}{16}}}\right) = -1.69$$

بالمقارنة نجد أن : $-1.69 < -1.65$ إذن نرفض H_0 ونقبل H_1 عند مستوى الدلالة 0.05 ونقول أن متوسط عمر المصابيح من النوع الأول اقل من متوسط عمر المصابيح من النوع الثاني عند نفس مستوى المعنوية السابق.

حل التمرين التاسع :

لدينا :

$$n = 3 \quad , \qquad \alpha = 0.01$$
$$\begin{cases} H_0 : \sigma^2 = 0.7 \\ H_1 : \sigma^2 > 0.7 \end{cases}$$

نلاحظ ان الفرضية البديلة ذات طرف واحد قيمتها الحرجة هي :

$$X^2{}_{(\alpha,n)} = X^2{}_{(0.01.3)} = 0.11$$

نقوم الآن بحساب إحصاء الاختبار :

$$k^2 = \frac{(n-1)s^2}{\sigma^2}$$

أولا نحسب $\overline{X}$ و s^2 :

$$\overline{X} = 54 \quad ; \quad s^2 = \frac{44}{3}$$

$$k^2 = \frac{(4-1)\frac{44}{3}}{(0.7)} = 62.857$$

بما أن $62.857 > 0.11$ فإننا نرفض الفرضية H_0 ونقبل الفرضية البديلة H_1 نستنتج ان قيمة التباين اكبر من 0.7 عند مستوى المعنوية 0.01.

حل التمرين العاشر :

نريد اختبار ما يلي :

$$n = 4 \quad , \qquad \alpha = 0.05$$

$$\begin{cases} H_0 : \sigma^2 = 0.6 \\ H_1 : \sigma^2 \neq 0.6 \end{cases}$$

نلاحظ أن الفرضية البديلة ذات طرفين و قيمتها الحرجة و هي :

$$k^2{}_{(\frac{\alpha}{2},n)} = k^2{}_{(0.025.4)} = 0.48$$

$$k^2{}_{(1-\frac{\alpha}{2},n)} = k^2{}_{(0.975.4)} = 11.15$$

$$k^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(5-1)(0.8)}{0.6} = 5.333$$

بما أن $0.48 < 5.333 < 11.15$ فاننا نقبل الفرضية الصفرية H_0 ونرفض الفرضية البديلة H_1 عند مستوى الدلالة 0.05 وبالتالي نقبل ادعاء الشركة عند نفس مستوى المعنوية السابق.

حل التمرين الحادي عشر:

لدينا :

$$n_1 = 10 \quad , V_2 = 25 \quad , \qquad \alpha = 0.05$$

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 > \sigma_2^2 \end{cases}$$

الفرضية البديلة لها طرف واحد أيمن و قيمتها حرجة هي:

$$F_{(\alpha,n_1,n_2)} = F_{(0.01,9,24)} = 3.26$$

$$F = \frac{s_1^2}{s_2^2} \cdot \frac{\sigma_1^2}{\sigma_2^2} = \frac{225}{400} = 0.562$$

بما أن $0.562 < 3.26$ فاننا نقبل H_0 ونرفض H_1 عند مستوى الدلالة 0.01 وبالتالي فان تبايني المجتمعين متساووين عند مستوى المعنوية 0.01.

حل التمرين الثالث عشر:

نريد اختبار:

$$\begin{cases} H_0 : \mathrm{P}_1 = \mathrm{P}_2 \\ \mathrm{H}_1 : \mathrm{P}_1 \neq \mathrm{P}_2 \end{cases}$$

بما أن الفرضية البديلة ذات طرفين فإن قيمها الحرجة هي:

$$-Z\alpha_{/2} = -Z_{0.025} = -1.96 \,, \qquad Z\alpha_{/2} = Z_{0.025} = 1.96$$

ولدينا:

$$\overline{\mathrm{P}}_1 = \frac{20}{400} = 0.05 \;;\; \overline{\mathrm{P}}_2 = \frac{16}{400} = 0.04$$

نقوم بحساب دالة الاختبار Z:

$$Z = \frac{\overline{\mathrm{P}}_1 - \overline{\mathrm{P}}_2}{\sqrt{\frac{\overline{\mathrm{p}}\,\overline{\mathrm{q}}}{n_1} + \frac{\overline{\mathrm{p}}\,\overline{\mathrm{q}}}{n_2}}}$$

نحسب أولا $\overline{\mathrm{P}}$ حيث:

$$\overline{\mathrm{p}}=\frac{n_1\overline{\mathrm{P}}_1+n_2\overline{\mathrm{P}}_2}{n_1+n_2}=\frac{400(0.05)+400(0.04)}{400+400}=0.045$$

$$Z=\frac{0.05-0.04}{\sqrt{\frac{(0.045)(0.955)}{400}+\frac{(0.045)(0.955)}{400}}}=0.682$$

نلاحظ أن $0.682<1.96$ أي أن قيمة Z المحسوبة تقع في منطقة قبول الفرضية H_0 إذن نقبل H_0 ونرفض H_1 عند مستوى الدلالة 0.05، وهذا يعني أنه لا يوجد فرقا بين نسبة الإنتاج المخالف للمواصفات للآليتين (أ) و(ب) عند مستوى معنوية 0.05

حل التمرين الرابع عشر:

نريد اختبار:

$$\begin{cases}H_0:\mathrm{P}_1=\mathrm{P}_2\\ \mathrm{H}_1:\mathrm{P}_1<\mathrm{P}_2\end{cases}$$

نلاحظ أن الفرضية البديلة ذات طرف واحد أيسر وقيمتها الحرجة هي:

$$-Z_\alpha=-Z_{0.05}=-1.65$$

والآن نقوم بحساب قيمة دالة الاختبار Z:

$$Z=\frac{\overline{\mathrm{P}}_1-\overline{\mathrm{P}}_2}{\sqrt{\frac{\overline{\mathrm{p}}\,\overline{\mathrm{q}}}{n_1}+\frac{\overline{\mathrm{p}}\,\overline{\mathrm{q}}}{n_2}}}$$

$$\overline{\mathrm{p}}=\frac{n_1\overline{\mathrm{P}}_1+n_2\overline{\mathrm{P}}_2}{n_1+n_2}=\frac{60(0.2)+90(0.3)}{60+90}=0.26$$

$$Z=\frac{0.2-0.3}{\sqrt{\frac{(0.26)(0.74)}{60}+\frac{(0.26)(0.74)}{90}}}=-1.367$$

نلاحظ أن $-1.65<-1.367$ أي أن قيمة Z المحسوبة تقع في منطقة قبول الفرضية الصفرية H_0 إذن نقبل H_0 ونرفض H_1، بمعنى أن نسبة الأشخاص المؤيدين للمترشح السياسي في الولاية (أ) ليست أقل من الولاية (ب) عند مستوى معنوية 0.05.

قائمة المراجع

أولا: المراجع باللغة العربية

1. إبراهيم علي إبراهيم عبد ربه، **مبادئ علم الإحصاء**، الدار الجامعية، الإسكندرية، 2002.
2. أحمد شكري الريماوي وسامي مسعود، **مقدمة في علم الإحصاء الوصفي الإحصاء والتحليلي**، دار حنين، ط 1، عمان، 1998.
3. أموري هادي كاظم وعصام خضير محمود، **طبيعة البيانات الإحصائية وبناء النماذج القياسية**، دار وائل، ط1، عمان،1999.
4. إياد محمد الهوبي، **الإحصاء التطبيقي**، ط1، 2014. وثيقة إنترنت متوفرة على الموقع:www.cst.ps/ebook/applied_statistics/#/12
5. باسم غدير غدير، **العالم الرقمي آلية تحليل البيانات**، دار الرضا للنشر، دمشق، 2003.
6. بوعظم كمال، **الإحصاء الإستدلالي من الجانب النظري والتطبيقي**، دار زهران، عمان، 2008.
7. ثابت عبد الرحمان إدريس، **بحوث التسويق،أساليب القياس والتحليل و اختبار الفروض**، الدار الجامعية ،الإسكندرية، 2005.
8. حسن ياسين طعمة وإيمان حسين حنوش، **الإحصاء الاستدلالي**، دار الصفاء، عمان، 2012.
9. دلال قاضي وآخرون، **الإحصاء للإداريين والاقتصاديين**، دار الحامد، عمان،2005.
10. سليمان محمد طشطوش، **أساسيات الإحصاء الرياضي**، دار اليازوري، عمان، 2012.

11. سليمان محمد طشطوش، **أساسيات المعاينة الإحصائية**،دار الشروق، ط 1،عمان، 2001.
12. شفيق العتوم، **طرق الإحصاء بإستخدام SPSS**، دار المناهج، عمان، 2015.
13. ضياء أحمد القاضي وآخرون، **الإحصاء ونظم المعلومات**،مركز جامعة القاهرة للتعليم المفتوح،القاهرة، 1998.
14. عدنان شهاب حمد ومهدي محسن إسماعيل، **أساليب المعاينة في ميدان التطبيق**، المعهد العربي للتدريب و البحوث الإحصائية، بغداد، 2001.
15. فتحي حمدان وكامل فليفل، **الإحصاء**، دار المناهج، ط1،عمان، 2006.
16. قيس ناجي عبد الجبار، **أصول الإحصاء والطرق الإحصائية**، دار المناهج، ط 1، عمان،2002.
17. محمد جاسم الياسري ومروان عبد المجيد إبراهيم، **الأساليب الإحصائية في مجالات البحوث التربوية**، مؤسسة الوراق، ط1، عمان،2001.
18. محمد حسين محمد رشيد و منى عطا الله الشويلات، **مبادئ الإحصاء والإحتمالات ومعالجنها بإستخدام برنامج spss** ، دار الصفاء، عمان، 2012.
19. محمد عبد العال النعيمي ومؤيد الفضل، **الإحصاء المتقدم في دعم القرار بالتركيز على منظمات الأعمال الإنتاجية**، الوراق للنشر، ط1، عمان، 2007.
20. محمد علي الأطرقجي، **الوسائل التطبيقية في الطرق الإحصائية**، دار الطليعة، بيروت، 1980.
21. مركز الإحصاء الرسمي للسكان لسلطنة عمان[2008]: وثيقة إنترنت متوفرة على الموقع: (تم الإطلاع عليها يوم 2018/05/25) http://www.omancensus.net/new/index.php?lang=ar#
22. المعهد العربي للتدريب والبحوث الإحصائية: **محاضرات حول أساليب جمع البيانات**، وثيقة إنترنت متوفرة على الموقع: www.arab-api.org/cours13/c13 (تم الإطلاع عليها يوم 2018/05/29)
23. نجاة رشيد الكيخيا، **أساسيات الإستنتاج الإحصائي**، دار المريخ، الرياض، 2007.
24. مقيدش نزيهة، **أهمية أسلوب المعاينة في الدراسات الإحصائية دراسة تطبيقية حول الحوكمة في الجامعة الجزائرية من خلال سبر للآراء جامعة فرحات عباس-سطيف-**، مذكرة ماجستير ، جامعة سطيف، 2009-2010.

ثانيا: المراجع باللغة الأجنبية

1. ANSION Guy, **Sondages et statistique**, labor éditions,Bruxelles ,1997.
2. ARDILLY Pascal**, Les techniques de sondage**, edition technip, paris,1994.
3. COCHRAN William-G , **Sampling techniques,** 3rd edition,john willey&sons, NEW YORK, 1977.
4. DODGE Yadolah**, Premiers pas en statistique**, springer verlag,France, 2003.
5. GAUTHY-SINECHAL Martine et VANDERCAMMEN Marc ,**etudes de marchés-methodes et outils,** 2ème édition, de boeck université, bruxelles,2005.
6. GIANNLLONI Jean-luc et VENNETTE Eric , **Etudes de marché,** 2ème édition, vuibert, paris, 2001.
7. GIARD Vincent **, Statistique appliquée à la gestion**, 8ème Edition economica, paris, 2003.
8. GRAIS Bernard , **Méthodes statistiques : techniques statistiques 2** , 3eme édition, 2003.
9. HAMDANI Hocine , **statistique descriptive**, office des publication universitaires, 2001.
10. LESSARD Sabin et MONGA, **statistique concepts et méthodes avec exercices et corrigés**, masson éditeur , canada, 1993 .

11. TILLÉ Yves , **théorie des sondages, Echantillionnage et estimation en populations finies**, dunod, paris, 2001.

الملاحق

دالة التوزيع الاحتمالية للقانون الط... ...ري

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

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	[illegible]	0,5120	0,5160	0,5199	0,5239	0,5279	[illegible]	0,5359
0,1	0,5398	0,5438	[illegible]	0,5517	0,5557	0,5596	0,5636	0,5675	[illegible]	0,5753
0,2	0,5793	0,5832	[illegible]	0,5910	0,5948	0,5987	0,6026	0,6064	[illegible]	0,6141
0,3	0,6179	0,6217	[illegible]	0,6293	0,6331	0,6368	0,6406	0,6443	[illegible]	0,6517
0,4	0,6554	0,6591	[illegible]	0,6664	0,6700	0,6736	0,6772	0,6808	[illegible]	0,6879
0,5	0,6915	0,6950	[illegible]	0,7019	0,7054	0,7088	0,7123	0,7157	[illegible]	0,7224
0,6	0,7257	0,7291	[illegible]	0,7357	0,7389	0,7422	0,7454	0,7486	[illegible]	0,7549
0,7	0,7580	0,7611	[illegible]	0,7673	0,7704	0,7734	0,7764	0,7794	[illegible]	0,7852
0,8	0,7881	0,7910	[illegible]	0,7967	0,7995	0,8023	0,8051	0,8078	[illegible]	0,8133
0,9	0,8159	0,8186	[illegible]	0,8238	0,8264	0,8289	0,8315	0,8340	[illegible]	0,8389
1,0	0,8413	0,8438	[illegible]	0,8485	0,8508	0,8531	0,8554	0,8577	[illegible]	0,8621
1,1	0,8643	0,8665	[illegible]	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	[illegible]	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	[illegible]	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	[illegible]	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	[illegible]	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	[illegible]	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	[illegible]	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	[illegible]	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	[illegible]	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	[illegible]	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	[illegible]	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	[illegible]	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	[illegible]	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	[illegible]	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	[illegible]	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	[illegible]	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	[illegible]	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	[illegible]	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	[illegible]	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

قيم دالة التوزيع من أجل القيم ال...

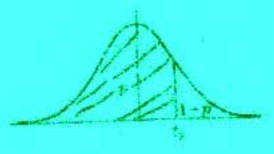
z	3,0	3,1	[illegible]	3,3	3,4	3,5	3,6	3,8	4,0	4,5
F(z)	0,99865	0,999032	[illegible]	[illegible]	0,999663	[illegible]	0,999841	0,999928	0,999968	0,999997

قيم z المختلفة من أجل p معلوم، حيث: $p=F(z)$

p	0,000	0,001	0,002	0,003	0,004	0,005	0,006	0,007	0,008	0,009	0,010	
0,00	∞	3,0902	2.8782	2,7478	2,6521	2,5758	2.5121	2,4573	2,4089	2,3656	2.3263	0,99
0,01	2,3263	2,2904	2.2571	2,2262	2,1973	2.1701	2.1444	2,1201	2,0969	2,0749	2.0537	0,98
0,02	2,0537	2,0335	2,0141	1,9954	1,9774	1,9600	1.9431	1,9268	1,9110	1,8957	1,8808	0,97
0,03	1,8808	1,8663	1.8522	1,8384	1,8250	1,8119	1,7991	1,7866	1,7744	1.7624	1.7507	0,96
0,04	1,7507	1,7392	1.7279	1,7169	1,7060	1.6954	1,6849	1.6747	1,6646	1,6546	1.6449	0,95
0,05	1,6449	1,6352	1.6258	1,6164	1,6072	1,5982	1.5893	1,5805	1,5718	1.5632	1.5548	0,94
0,06	1,5548	1,5464	1.5382	1,5301	1,5220	1,5141	1,5063	1,4985	1,4909	1.4833	1.4758	0,93
0,07	1,4758	1,4684	1.4611	1.4538	1,4466	1.4395	1.4325	1,4255	1,4187	1.4118	1.4051	0,92
0,08	1,4051	1,3984	1.3917	1,3852	1,3787	1,3722	1.3658	1,3595	1,3532	1,3469	1.3408	0,91
0,09	1,3408	1,3346	1.3285	1,3225	1,3165	1.3106	1.3047	1,2988	1,2930	1,2873	1.2816	0,90
0,10	1,2816	1,2759	1.2702	1.2646	1,2591	1,2536	1,2481	1,2426	1.2372	1.2319	1.2265	0,89
0,11	1,2265	1,2212	1.2160	1,2107	1,2065	1.2004	1.1952	1,1901	1,1850	1.1800	1.1750	0,88
0,12	1,1750	1,1700	1.1650	1,1601	1,1552	1.1503	1.1455	1,1407	1,1359	1.1311	1.1264	0,87
0,13	1,1264	1,1217	1.1170	1,1123	1,1077	1.1031	1.0985	1,0939	1,0893	1,0848	1.0803	0,86
0,14	1,0803	1,0758	1.0714	1,0669	1,0625	1.0581	1.0537	1.0494	1,0450	1,0407	1.0364	0,85
0,15	1,0364	1,0322	1.0279	1.0237	1,0194	1.0152	1.0110	1,0069	1,0027	0,9986	0.9945	0,84
0,16	0,9945	0.9904	0.9863	0,9822	0,9782	0.9741	0.9701	0,9661	0,9621	0.9581	0.9542	0,83
0,17	0,9542	0,9502	0.9453	0.9424	0,9385	0,9346	0.9307	0,9269	0,9230	0,9192	0.9154	0,82
0,18	0,9154	0,9116	0.9078	0.9040	0,9002	0.8965	0.8927	0,8890	0,8853	0.8816	0.8779	0,81
0,19	0,8779	0,8742	0.8705	0,8669	0,8633	0.8596	0.8560	0,8524	0,8488	0.8452	0.8416	0,80
0,20	0,8416	0,8381	0.8345	0,8310	0,8274	0,8239	0.8204	0,8169	0,8134	0.8099	0.8064	0,79
0,21	0,8064	0,8030	0.7995	0,7961	0,7926	0.7892	0.7858	0,7824	0,7790	0.7756	0.7722	0,78
0,22	0,7722	0,7688	0.7655	0,7621	0,7588	0.7554	0.7521	0,7488	0,7454	0.7421	0.7388	0,77
0,23	0,7388	0.7356	0.7323	0,7290	0,7257	0.7225	0.7192	0.7160	0,7128	0.7095	0.7063	0,76
0,24	0,7063	0,7031	0.6999	0,6967	0,6935	0.6903	0.6871	0,6840	0,6808	0.6776	0.6745	0,75
0,25	0,6745	0,6713	0.6682	0,6651	0,6620	0.6588	0.6557	0,6526	0,6495	0.6464	0.6433	0,74
0,26	0,6433	0,6403	0.6372	0.6341	0,6311	0.6280	0.6250	0,6219	0,6189	0.6158	0.6128	0,73
0,27	0,6128	0,6098	0.6068	0,6038	0,6008	0.5978	0.5948	0,5918	0,5888	0.5858	0.5828	0,72
0,28	0,5828	0,5799	0.5769	0,5740	0,5710	0.5681	0.5651	0,5622	0,5592	0.5563	0.5534	0,71
0,29	0,5534	0,5505	0.5476	0.5446	0,5417	0,5388	0.5359	0,5330	0,5302	0,5273	0.5244	0,70
0,30	0,5244	0,5215	0.5187	0.5158	0,5129	0.5101	0.5072	0,5044	0,5015	0.4987	0.4959	0,69
0,31	0,4959	0,4930	0.4902	0,4874	0,4845	0.4817	0.4789	0,4761	0,4733	0.4705	0.4677	0,68
0,32	0,4677	0,4649	0.4621	0,4593	0,4565	0.4538	0.4510	0,4482	0,4454	0,4427	0.4399	0,67
0,33	0,4399	0,4372	0.4344	0,4316	0,4289	0.4261	0.4234	0,4207	0,4179	0.4152	0.4125	0,66
0,34	0,4125	0,4097	0.4070	0.4043	0,4016	0,3989	0.3961	0,3934	0,3907	0,3880	0.3853	0,65
0,35	0,3853	0,3826	0.3799	0,3772	0,3745	0.3719	0.3692	0,3665	0,3638	0.3611	0.3585	0,64
0,36	0,3585	0,3558	0.3531	0.3505	0,3478	0.3451	0.3425	0,3398	0,3372	0.3345	0.3319	0,63
0,37	0,3319	0,3292	0,3266	0.3239	0,3213	0.3186	0.3160	0,3134	0,3107	0.3081	0.3055	0,62
0,38	0,3055	0,3029	0.3002	0,2976	0,2950	0,2924	0,2898	0.2871	0,2845	0.2819	0.2793	0,61
0,39	0,2793	0,2767	0,2741	0.2715	0,2689	0.2663	0.2637	0,2611	0,2585	0.2559	0.2533	0,60
0,40	0,2533	0,2508	0.2482	0.2456	0,2430	0.2404	0.2378	0,2353	0,2327	0.2301	0.2275	0,59
0,41	0,2275	0,2260	0,2224	0,2198	0,2173	0.2147	0.2121	0,2096	0,2070	0.2045	0.2019	0,58
0,42	0,2019	0,1993	0.1968	0.1942	0,1917	0,1891	0.1866	0,1840	0,1815	0,1789	0.1764	0,57
0,43	0,1764	0,1738	0.1713	0.1687	0,1662	0.1637	0.1611	0,1586	0,1560	0.1535	0.1510	0,56
0,44	0,1510	0,1484	0.1459	0.1434	0,1408	0.1383	0,1358	0,1332	0,1307	0.1282	0.1257	0,55
0,45	0,1257	0,1231	0.1206	0,1181	0,1156	0.1130	0.1105	0.1080	0,1055	0,1030	0.1004	0,54
0,46	0,1004	0,0979	0,0954	0,0929	0,0904	0.0878	0.0853	0.0828	0,0803	0.0778	0,0753	0,53
0,47	0,0753	0,0728	0.0702	0.0677	0,0652	0.0627	0.0602	0,0577	0,0552	0,0527	0.0502	0,52
0,48	0,0502	0,0476	0.0451	0.0426	0,0401	0.0376	0.0351	0.0326	0,0301	0.0276	0.0251	0,51
0,49	0,0251	0.0226	0.0201	0,0175	0,0150	0.0125	0.0100	0.0075	0,0050	0.0025	0.0000	0,50
	0,010	0,009	0,008	0,007	0,006	0,005	0,004	0,003	0,002	0,001	0,000	p

توزيع ستيودنت

Valeurs des centiles (t_p) pour la distribution t de Student à ν degrés de liberté



ν	$t_{.55}$	$t_{.60}$	$t_{.70}$	$t_{.75}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	,158	,325	,727	1,000	1,376	3,08	6,31	12,71	31,82	63,66
2	[illegible]	[illegible]	[illegible]	,816	1,061	1,89	2,92	4,30	6,96	9,92
3	[illegible]	[illegible]	,584	,765	,978	1,64	[illegible]	[illegible]	4,54	5,84
4	,134	,271	[illegible]	[illegible]	,941	1,53	[illegible]	[illegible]	[illegible]	[illegible]
5	,132	,267	[illegible]	,727	,920	[illegible]	2,02	[illegible]	[illegible]	4,03
6	,131	[illegible]	[illegible]	[illegible]	,906	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
7	,130	[illegible]	[illegible]	,711	,896	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
8	,130	[illegible]	[illegible]	,706	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
9	,129	,261	[illegible]	[illegible]	,883	1,38	[illegible]	2,26	[illegible]	[illegible]
10	,129	,260	,542	,700	,879	1,37	1,81	2,23	2,76	3,17
11	,129	,260	,540	,697	,876	1,36	1,80	2,20	2,72	3,11
12	,128	,259	,539	,696	,873	1,36	1,78	2,18	2,68	3,06
13	,128	,259	,538	,694	,870	1,35	1,77	2,16	2,65	3,01
14	,128	,258	,537	,692	,868	1,34	1,76	2,14	2,62	2,98
15	,128	,258	,536	,691	,866	1,34	1,75	2,13	2,60	2,95
16	,128	,258	,535	,690	,865	1,34	1,75	2,12	2,58	2,92
17	,128	,257	,534	,689	,862	1,33	1,74	2,11	2,57	2,90
18	,127	,257	,534	[illegible]	,862	1,33	1,73	[illegible]	2,55	2,88
19	,127	,257	,533	,688	[illegible]	[illegible]	1,73	[illegible]	2,54	[illegible]
20	,127	,257	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
21	,127	,257	[illegible]	,686	[illegible]	[illegible]	[illegible]	2,07	2,51	[illegible]
22	,127	,256	[illegible]	,686	,858	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
23	,127	,256	[illegible]	,685	,858	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
24	,127	,256	,531	,685	,857	1,32	[illegible]	[illegible]	[illegible]	[illegible]
25	,127	[illegible]	[illegible]	,684	,856	[illegible]	[illegible]	[illegible]	[illegible]	2,78
26	,127	,256	[illegible]	,684	,856	1,32	[illegible]	2,06	2,48	[illegible]
27	,127	,256	[illegible]	[illegible]	[illegible]	[illegible]	1,70	2,05	2,47	[illegible]
28	,127	[illegible]	[illegible]	,683	,855	[illegible]	[illegible]	2,05	[illegible]	[illegible]
29	,127	,256	,530	,683	,854	[illegible]	[illegible]	2,04	2,46	[illegible]
30	,127	,256	,530	,683	,854	1,31	[illegible]	2,04	2,46	2,75
40	,126	[illegible]	[illegible]	[illegible]	,851	1,30	1,68	2,02	2,42	2,70
60	,126	,254	[illegible]	[illegible]	,848	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
120	,126	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1,98	[illegible]	[illegible]
∞	,126	[illegible]	[illegible]	,674	,842	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Source : R.A. Fisher and F. Yates, Statistical Tables for [illegible] Agricultural and Medical Research, Longman Group Ltd, London [illegible] de l'auteur et des [illegible]

جدول توزيع فيشر عند مستوى دلالة1 %

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	[illegible]
13	[illegible]	[illegible]	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	[illegible]	[illegible]	[illegible]	[illegible]
15	[illegible]	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	[illegible]	[illegible]	[illegible]	[illegible]
23	7.88	5.66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.27
26	7.72	5.53	[illegible]	[illegible]	[illegible]	3.59	3.42	3.29	3.18	3.09	[illegible]	2.82	[illegible]	[illegible]	[illegible]	[illegible]	2.33	2.23
27	7.68	5.49	4.60	[illegible]	[illegible]	[illegible]	3.39	3.26	3.15	3.06	2.93	[illegible]	2.63	[illegible]	[illegible]	[illegible]	2.29	2.20
28	7.64	5.45	[illegible]	[illegible]	[illegible]	3.53	3.36	3.23	3.12	3.03	2.90	[illegible]	2.60	[illegible]	[illegible]	[illegible]	2.26	2.17
29	7.60	5.42	4.54	[illegible]	[illegible]	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	[illegible]	2.41	[illegible]	2.23	2.14
30	7.56	5.39	4.51	4.02	[illegible]	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	[illegible]	[illegible]	[illegible]	2.21	2.11
40	7.31	5.18	4.31	[illegible]	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	[illegible]	[illegible]	[illegible]	2.02	1.92
60	7.08	4.98	4.13	[illegible]	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	[illegible]	1.84	1.73
120	6.85	4.79	3.95	[illegible]	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	[illegible]	1.66	1.53

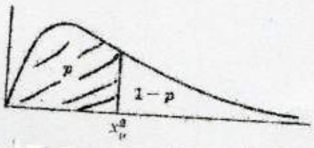
جدول توزيع فيشر عند مستوى دلالة. 5%

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	222	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35

توزيع كاي : χ^2

Appendice E

Valeurs des centiles (χ^2_p) pour la distribution en khi-carré à ν degrés de liberté



ν	$\chi^2_{,005}$	$\chi^2_{,01}$	$\chi^2_{,025}$	$\chi^2_{,05}$	$\chi^2_{,10}$	$\chi^2_{,25}$	$\chi^2_{,50}$	$\chi^2_{,75}$	$\chi^2_{,90}$	$\chi^2_{,95}$	$\chi^2_{,975}$	$\chi^2_{,99}$	$\chi^2_{,995}$	$\chi^2_{,999}$
1	,0000	,0002	,0010	,0039	,0158	,102	,455	1,32	2,71	3,84	5,02	6,63	7,88	10,8
2	,0100	,0201	,0506	,103	,211	,575	1,39	2,77	4,61	5,99	7,38	9,21	10,6	13,8
3	,0717	,115	,216	,352	,584	1,21	2,37	4,11	6,25	7,81	9,35	11,3	12,8	16,3
4	,207	,297	,484	,711	1,06	1,92	3,36	5,39	7,78	9,49	11,1	13,3	14,9	18,5
5	,412	,554	,831	1,15	1,61	2,67	4,35	6,63	9,24	11,1	12,8	15,1	16,7	20,5
6	,676	,872	1,24	1,64	2,20	3,45	5,35	7,84	10,6	12,6	14,4	16,8	18,5	22,5
7	,989	1,24	1,69	2,17	2,83	4,25	6,35	9,04	12,0	14,1	16,0	18,5	20,3	24,3
8	1,34	1,65	2,18	2,73	3,49	5,07	7,34	10,2	13,4	15,5	17,5	20,1	22,0	26,1
9	1,73	2,09	2,70	3,33	4,17	5,90	8,34	11,4	14,7	16,9	19,0	21,7	23,6	27,9
10	2,16	2,56	3,25	3,94	4,87	6,74	9,34	12,5	16,0	18,3	20,5	23,2	25,2	29,6
11	2,60	3,05	3,82	4,57	5,58	7,58	10,3	13,7	17,3	19,7	21,9	24,7	26,8	31,3
12	3,07	3,57	4,40	5,23	6,30	8,44	11,3	14,8	18,5	21,0	23,3	26,2	28,3	32,9
13	3,57	4,11	5,01	5,89	7,04	9,30	12,3	16,0	19,8	22,4	24,7	27,7	29,8	34,5
14	4,07	4,66	5,63	6,57	7,79	10,2	13,3	17,1	21,1	23,7	26,1	29,1	31,3	36,1
15	4,60	5,23	6,26	7,26	8,55	11,0	14,3	18,2	22,3	25,0	27,5	30,6	32,8	37,7
16	5,14	5,81	6,91	7,96	9,31	11,9	15,3	19,4	23,5	26,3	28,8	32,0	34,3	39,2
17	5,70	6,41	7,56	8,67	10,1	12,8	16,3	20,5	24,8	27,6	30,2	33,4	35,7	40,8
18	6,26	7,01	8,23	9,39	10,9	13,7	17,3	21,6	26,0	28,9	31,5	34,8	37,2	42,3
19	6,84	7,63	8,91	10,1	11,7	14,6	18,3	22,7	27,2	30,1	32,9	36,2	38,6	43,8
20	7,43	8,26	9,59	10,9	12,4	15,5	19,3	23,8	28,4	31,4	34,2	37,6	40,0	45,3
21	8,03	8,90	10,3	11,6	13,2	16,3	20,3	24,9	29,6	32,7	35,5	38,9	41,4	46,8
22	8,64	9,54	11,0	12,3	14,0	17,2	21,3	26,0	30,8	33,9	36,8	40,3	42,8	48,3
23	9,26	10,2	11,7	13,1	14,8	18,1	22,3	27,1	32,0	35,2	38,1	41,6	44,2	49,7
24	9,89	10,9	12,4	13,8	15,7	19,0	23,3	28,2	33,2	36,4	39,4	43,0	45,6	51,2
25	10,5	11,5	13,1	14,6	16,5	19,9	24,3	29,3	34,4	37,7	40,6	44,3	46,9	52,6
26	11,2	12,2	13,8	15,4	17,3	20,8	25,3	30,4	35,6	38,9	41,9	45,6	48,3	54,1
27	11,8	12,9	14,6	16,2	18,1	21,7	26,3	31,5	36,7	40,1	43,2	47,0	49,6	55,5
28	12,5	13,6	15,3	16,9	18,9	22,7	27,3	32,6	37,9	41,3	44,5	48,3	51,0	56,9
29	13,1	14,3	16,0	17,7	19,8	23,6	28,3	33,7	39,1	42,6	45,7	49,6	52,3	58,3
30	13,8	15,0	16,8	18,5	20,6	24,5	29,3	34,8	40,3	43,8	47,0	50,9	53,7	59,7
40	20,7	22,2	24,4	26,5	29,1	33,7	39,3	45,6	51,8	55,8	59,3	63,7	66,8	73,4
50	28,0	29,7	32,4	34,8	37,7	42,9	49,3	56,3	63,2	67,5	71,4	76,2	79,5	86,7
60	35,5	37,5	40,5	43,2	46,5	52,3	59,3	67,0	74,4	79,1	83,3	88,4	92,0	99,6
70	43,3	45,4	48,8	51,7	55,3	61,7	69,3	77,6	85,5	90,5	95,0	100	104	112
80	51,2	53,5	57,2	60,4	64,3	71,1	79,3	88,1	96,6	102	107	112	116	125
90	59,2	61,8	65,6	69,1	73,3	80,6	89,3	98,6	108	113	118	124	128	137
100	67,3	70,1	74,2	77,9	82,4	90,1	99,3	109	118	124	130	136	140	149

Source : E.S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. I (1966), table 8, pages 137 et 138. Reproduction autorisée.

العنـــــــــــوان	الصفحة
مقدمة	05
الفصل الأول:...الإطار المفاهيمي	
1. البيانات الإحصائية	09
1.1. تعريف البيانات الإحصائية	09
2.1 أنواع وطبيعة البيانات الإحصائية	09
1.2.1 أنواع البيانات الإحصائي	09
2.2.1 طبيعة البيانات الإحصائية	11
3.1 أنواع الأساليب الإحصائية	12
1.3.1 أسلوب الحصر الشامل	12
2.3.1 أسلوب الحصر الجزئي	15
3.3.1 أسلوب المعاينة	16
2. المعاينة الإحصائية كأسلوب لجمع البيانات	16
1.2. مفاهيم أساسية	17
2.2. مراحل تصميم خطة المعاينة	18
3.2. أنواع المعاينة الإحصائية	21

1.3.2 المعاينة العشوائية (الإحتمالية)	21
2.3.2 المعاينة غير العشوائية (غير الاحتمالية)	27
.4.2 حجم العينة ومصادر أخطاء المعاينة	28
.3 التوزيعات الاحتمالية المتصلة	33
1.3 التوزيع الطبيعي	33
3.3 توزيع كاي تربيع	36
4.3 توزيع ستيودنت	37
5.3 توزيع فيشر	38
الفصل الثاني...توزيع المعاينة	
1.توزيع المعاينة للوسط الحسابي للعينة $\overline{X}$	47
2.حالات مختلفة لإيجاد توزيع الوسط الحسابي للعينة	54
1.2 إذا كان تباين المجتمع معلوم	54
2.2 إذا كان تباين المجتمع الطبيعي مجهول	54
1.2.2 عندما يكون حجم العينة $n \geq 30$	55
2.2.2 عندما يكون حجم العينة $n < 30$	55
3.2 إذا كان المجتمع طبيعي ذو حجم محدود وكان حجم العينة $n \geq 30$ وكان السحب بدون إرجاع	56
1.3.2 عند معلومية σ^2 يكون توزيع الوسط الحسابي للعينة	56
1.3.2 عند مجهولية σ^2 يكون توزيع الوسط الحسابي للعينة	57

3. توزيع الفرق بين متوسطي عينتين	58
1.3 عند معلومية σ_1^2 و σ_2^2	58
2.3 عند مجهولية σ_1^2 و σ_2^2 وحجم العينتين كبير أي $n_2, n_1 \geq 30$	59
3.3 عند مجهولية σ_1^2 و σ_2^2 وحجم العينتين صغير أي $n_2, n_1 < 30$	60
5. توزيع النسبة للعينة	61
5.توزيع الفرق بين نسبتين	63
6. توزيع المعاينة لتباين العينة S^2	63
7.توزيع النسبة بين تباين عينتين	64
الفصل الثالث.. تقدير معالم مجتمع	
1- التقدير النقطي	76
2- بناء مجالات الثقة "التقدير بمجال	83
1.2 مجال الثقة للمتوسط الحسابي للمجتمع μ	84
1.1.2 إذا كان حجم العينة كبير $(n \geq 30)$	84
2.1.2 إذا كان حجم العينة صغير $n < 30$	87
2.2 تقدير الفرق بين وسطي مجتمعين مستقلين بفترة	88
1.2.2 عند معلومية تباين مجتمعين σ_1^2 و σ_2^2	88
2.2.2 عند مجهولية تباين المجتمعين σ_1^2 و σ_2^2	90
1.2.2.2 إذا كان $n_1 . n_2 \geq 30$	90

2.2.2.2 إذا كان$n_1 . n_2 < 30$ (أو أحدهما أقل من 30) وكان تباينا المجتمعين غير متساوي فإن فترة الثقة	90
3.2 فترة الثقة لنسبة المجتمع P	91
2. 4 فترة الثقة للفرق بين نسبتي مجتمعين $(P_1 - P_2)$	94
5.2 فترة الثقة لتباين المجتمع σ^2	97
6.2 فترة الثقة للنسبة بين تبايني مجتمعين $\frac{\sigma_1^2}{\sigma_2^2}$	99
7.2 خطأ المعاينة في التقدير وحجم العينة اللازم لعدم تجاوز خطأ معين	102
الفصل الرابع:..اختبار الفرضيات	
1.مفاهيم أساسية	118
1.1 الفرضية الإحصائية	118
2.1 اختبار الفرضية	118
3.1 أنواع الفروض الإحصائية	118
4.1 مستوى المعنوية	119
5.1 أخطاء إختبار الفرضيات	119
6.1 أنواع الفرضيات	119
7.1 الخطوات المتبعة في اختبار الفرضيات	121
2. اختبارات المتوسط الحسابي للمجتمع	122
1.2 اختبار الطرفين أو ذو جانبين	122
2.2 اختبار ذو طرف أيسر	126

3.2 اختبار ذو طرف أيمن	127
4.2 اختيار الفرق بين وسطي مجتمعين	128
3. اختبار نسبة P	132
4. اختبار الفرق بين نسبتين لمجتمعين $(p_1 - p_2)$	134
5. اختبار لتباين المجتمع	137
6. اختبار تباينين	141
قائمة المراجــع	151
الملاحــق	154