الجمهوريــة الجزائريــة الديمقراطيــة الشعبيــة

République Algérienne Démocratique et Populaire

و البـحـث الـعـلـمــي وزارة الـتعـلـيـم الـعـالي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة مصطفى اسطمبولي

معسكر

Université MUSTAPHA Stambouli

Mascara

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

قسم: علوم التسيير

ملزمة دروس موجهة لطلبة السنة الثانية علوم التسيير

في مقياس :

احصاء 03

إعداد الأستاذ: بشرول فيصل

السنة الجامعية: 2023/2022

لقد تعددت في الوقت الحاضر مساهمات الإحصاء في شتى فروع العلوم حيث أصبح علما قائما بذاته يخدم جميع العلوم الأخرى بما فيها علم الاقتصاد، و يمكن الإعتماد على الاحصاء كأسلوب فعال في حل مشكلة ما عند توافر البيانات والمعلومات والمؤشرات الإحصائية، حيث له عدة وظائف منها تحليل البيانات والمعلومات، تحديد الفرضيات المناسبة، إجراء الاختبارات الإحصائية و إتخاذ القرار الاحصائي المناسب، والأهم هو استخلاص النتائج.

تعالج هذه المطبوعة جانب من جوانب الاحصاء بفروعه المتشعبة والواسعة، حيث تتضمن هذه المطبوعة محاضرات و تمارين في مقياس الاحصاء 03 ، وهي موجهة أساسا لطلبة السنة الثانية علوم التسيير، كما يمكن أن يستفيد منها طلبة باقي التخصصات في طوري الليسانس وحتى الماستر، ويعتبر مقياس الاحصاء 03 تكملة للمعارف المكتسبة في مقياس الاحصاء 01 و الاحصاء 02 اذ يشترط الالمام بهما بصورة حسنة حتى يمكن استيعاب برنامج مقياس احصاء 03 بكل سهولة.

لقد كانت هذه المطبوعة عبارة عن حصيلة لتجربة سنوات في تدريس مقياس الإحصاء 03 ، حيث حاولنا الاستفادة من هذه التجربة المتواضعة في تقديم محاضرات وتمارين مصاغة بطريقة ميسرة ومفهومة بعيدة عن التعقيد بالاعتماد على مراجع فيها معلومات و أمثلة واضحة مع اجراء تعديلات عليها عند الحاجة للتبسيط أكثر، وكان الهدف دوما هو ايصال المعلومة في صورة بسيطة و نقل المبادئ الأساسية اللازمة التي تضمنها هذا المقياس الى الطلبة دون التركيز على البراهين الرياضية المعقدة، بحيث يتم مراعات مستويات أغلبية الطلاب الذين يجدون صعوبة في استيعاب المقاييس التقنية بما فيها الاحصاء.

تم تقديم هذه المطبوعة ضمن ثلاث فصول أساسية وهي توزيعات المعاينة، نظرية التقدير (التقدير الاحصائي) و اختبار الفرضيات حيث تعتبر هذه الفصول الحجر الأساس في الاحصاء الاستدلالي مع اعطاء مجموعة من الأمثلة ضمن كل محور حتى يمكن استيعاب ما تم التطرق اليه من نظريات، قوانين ومفاهيم، ففي الفصل الأول : توزيعات المعاينة تطرقنا من خلاله الى بعض المفاهيم الاحصائية ثم تطرقنا الى توزيعات المعاينة بأشكالها الأساسية وهي توزيع المعاينة للمتوسط الحسابي، للنسبة، للتباين وكذلك للفروقات، في الفصل الثاني: في نظرية التقدير تناولنا أنواع التقدير بداية بالتقدير النقطي، ثم التقدير بمجال الثقة ، في الفصل الثالث والأخير عرضنا خلاله اختبار الفرضيات الخاصة بثلاث معالم للمجتمع الاحصائي وهي المتوسط الحسابي، النسبة و التباين، وقد تم في نهاية كل فصل عرض مجموعة من التمارين المحلولة والخاصة بكل عنصر تم التطرق اليه في فصول هذه المطبوعة .

الفصل الأول: توزيعات المعاينة

تمهيد

تلعب توزيعات العينات دورًا مهمًا للغاية في التحليل الإحصائي واتخاذ القرار . ، لأن العينة عبارة عن مجموعة من المتغيرات العشوائية X_1، . . .، X_n ، يتبع ذلك أن إحصائية العينة هي دالة للعينة و هي أيضًا عشوائية. نسمي التوزيع الاحتمالي لإحصائية العينة بتوزيع المعاينة، توفر توزيعات المعاينة الرابط بين نظرية الاحتمالات والاستدلال الإحصائي، و تعد القدرة على تحديد توزيع الإحصائية جزءًا مهمًا في بناء وتقييم الإجراءات الإحصائية. من المهم ملاحظة أن هناك فرقًا بين توزيع المجتمع الذي أخذت منه العينة وتوزيع إحصائية العينة. بشكل عام لدى المجتمع الاحصائي توزيع يسمى توزيع المجتمع وهو عادة غير معروف ، في حين أن إحصائية العينة لها توزيع المعاينة ، والذي عادة ما يكون مختلفًا عن توزيع المجتمع.

1.مفاهيم احصائية

1.1/*علم الاحصاء*

يعرف علم الاحصاء بأنه ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع وتنظيم وتلخيص وعرض وتحليل البيانات وذلك للوصول الى نتائج مقبولة وقرارات سليمة، ينقسم علم الاحصاء الى قسمين هما الاحصاء الوصفي والاحصاء الاستدلالي.

1.1.1/ *الاحصاء الوصفي* : هو عبارة عن الطرق الخاصة بتنظيم وتلخيص المعلومات و تشمل الطرق الوصفية للإحصاء على التوزيعات التكرارية (الجداول التكرارية)، ورسوم بيانية وطرق حساب مقاييس النزعة المركزية ومقاييس التشتت ومختلف القياسات الأخرى.

2.1.1/ *الاحصاء الاستدلالي*: هو عبارة عن الطرق العلمية التي تعمل على الاستدلال عن معالم المجتمع بناءا على المعلومات التي تم الحصول عليها من العينة المأخوذة منه، وذلك وفقا للطرق الاحصائية المعلومة. ويطلق عليه الإحصاء الاستدلالي Inferential ، الإحصاء الاستنباطي Inductive ، ويعتمد على مجموعة من النظريات الإحصائية من أبرزها نظرية الاحتمالات ونظرية المعاينة، اذ يعتبران حلقة الوصل بين الإحصاء الوصفي والاحصاء الاستدلالي .

ويتضمن الاحصاء الاستدلالي دراسة عنصرين أساسيين هما :

-***التقدير الاحصائي*** **(Statistical Estimation)** هو عملية استنتاج أو تقدير معلمة ما من معالم المجتمع بالإحصائية المناظرة لها في العينة. فمثلا تقدير متوسط المجتمع (μ)بمتوسط العينة $(\bar{X})$.

- ***اختبار الفرضيات*(Hypotheses Tests)** : هو تحديد قرار على أساسه نقبل أو نرفض فرضية ما حول معلمة من معالم المجتمع، وذلك على ضوء المعلومات التي تم تجميعها من العينة المسحوبة من ذات المجتمع.

2.1/ *المجتمع الاحصائي و العينة*

- ***تعريف المجتمع الاحصائي***: يعرف على أنه مجموعة من الأفراد محل الدراسة و التي تشترك في خصائص مشتركة ويقسم المجتمع الإحصائي الى نوعين: مجتمع احصائي محدود(منتهي) يشمل على عدد محدود من الأفراد و مجتمع غير محدود يكون فيه عدد الأفراد غير منتهي.

- ***تعريف العينة*** : يمكن تعريف العينة بأنها ذلك الجزء الذي يتم اختياره من المجتمع بهدف تعميم نتائجه على المجتمع ككل ولهذا يجب أن تكون العينة ممثلة بشمل جيد للمجتمع المدروس بصورة صادقة حتى يمكن استخدام بياناتها في ايجاد تقديرات جيدة لمعالم المجتمع. و يكون أحيانا من الصعب ملاحظة بيانات جميع أفراد المجتمع لما لذلك من **تكلفة ووقت ومال،** أو أنه يستحيل تحديد جميع بيانات أفراد المجتمع مثل فحص جميع دم المرضى وعلى ذلك يتم اللجوء الى اختيار جزء من المجتمع يسمى عينة.

3.1/ *المعلمة والاحصائية*

- ***المعلمة***: هي خاصية تميز المجتمع ككل مثل متوسط الدخل الشهري للأسر في دولة معينة، نسبة القطع المعيبة في الانتاج لإحدى السلع، متوسط طول الطلاب في مدرسة ما...، كما تعرف معالم المجتمع بأنها مقاييس احصائية (المتوسط، التباين ...) تصف المجتمع وتحدده ويمكن حسابها من خلال القياس الكمي لجميع مفردات المجتمع.

- ***الاحصائية*** : هي خاصية تميز العينة مثل متوسط الدخل الشهري لعينة مكونة من 100 أسرة في دولة ما، متوسط الطول لعينة من 50 طالب في مدرسة ما ...

بعبارة أخرى احصاءات العينة تمثل تلك المقاييس والتقديرات لمعالم المجتمع ويتم الحصول عليها من خلال استخدام قياسات مفردات العينة مثل الوسط الحسابي، الانحراف المعياري، مدى العينة ...

4.1/ *أنواع العينات*

تقسم العينات الى قسمين رئيسيين هما العينات الاحتمالية والعينات الغير احتمالية.

1.4.1/ *العينات الاحتمالية (العشوائية)* **Probability Samples**

يتم سحب مفردات هذه العينة على أساس قانون الاحتمالات بشكل متتالي وباحتمال معروف مما يميز هذا النوع من العينات أنها تعطي نسبة الخطأ في التقديرات الناتجة عنها وتعطي مجال الثقة في التقديرات . كما أنه ليس من الضروري أن يكون هذا الاحتمال متساوي لكل وحدة في المجتمع وهذا الاختلاف في الاحتمال يكون مفيدا في توفير دقة أكبر للتقديرات المتحصل عليها في العينة وذلك في حالة المجتمعات غير المتجانسة (مثل العينات العشوائية الطبقية) ، يوجد عدة انواع من العينات الاحتمالية بحيث أن استخدام كل منها يتوقف على طبيعة مجتمع الدراسة والهدف من الدراسة والامكانيات المتاحة ومن أنواعها العينة العشوائية البسيطة Simple Random Sample ، العينة المنتظمة Systematic Sample ، العينة الطبقية Stratified Sample والعينة العنقودية Cluster Sample.

2.4.1/ *العينات غير احتمالية (الغير عشوائية)* **(Non-Probability Samples)**

من خلال هذا النوع من العينات يتم اختيار مفردات العينة بناءا على معايير خاصة يضعها الباحث دون أن يتقيد بقوانين الاحتمالات، يطبق هذا النوع من العينات غالبا في مسوح استطلاعات الرأي .

5.1/ *أساليب اختيار العينة*: يمكن أن نميز بين أسلوبين لاختيار العينات هما :

- ***المعاينة بالارجاع*** (السحب بالإرجاع) : عند اختيار مفردة من المجتمع فانه يتم اعادته للمجتمع ليتم اختيار المفردة الثانية بحيث هناك احتمال أن تظهر نفس المفردة أو غيرها .

-***المعاينة بدون ارجاع***(السحب بدون إرجاع): عند اختيار المفردة الأولى من المجتمع لا يتم اعادتها للمجتمع بحث يتم اختيار من المفردات ما تبقى في المجتمع وهكذا، وعمليا يتم اعتماد هذا الأسلوب في المعاينة .

6.1/ أساليب المعاينة

1.6.1/ أساليب المعاينة الاحتمالية (العشوائية)

1.1.6.1/ أسلوب المعاينة العشوائية البسيطة

يعتبر هذا الأسلوب من أبسط الطرق وأكثرها انتشارا في أساليب المعاينة، يمتاز هذا الأسلوب بأنه يعطي لكل وحدة من وحدات المعاينة الموجودة في المجتمع احتمال متساوي للاختيار أو الظهور في العينة، لاستخدام هذا الأسلوب يشترط في وحدات المعاينة للمجتمع المستهدف في الدراسة أن تكون متجانسة بالنسبة للصفة المدروسة (التباين بين وحدات المعاينة في المجتمع للصفة المدروسة يكون قليل نسبيا).

من بين الطرق المستخدمة في اختيار العينات العشوائية البسيطة طريقة القرعة، جدول الأرقام العشوائية أو باستخدام الحاسوب عبر برامج احصائية متخصصة مثل برنامج SPSS بحيث يتم اعطاء رقم متسلسل لكافة وحدات المجتمع الخاضعة للمعاينة.

من عيوب استخدام المعاينة العشوائية البسيطة نذكر :

- ان استخدام العينة العشوائية البسيطة لا يضمن ان تكون ممثلة لصفة المجتمع المدروس في حالة ما اذا كانت وحدات المجتمع غير متجانسة في الصفة المدروسة .
- اذا كان حجم مجتمع الدراسة كبيرا فان هناك حاجة لجهد كبير لتهيئة اطار المجتمع خاصة اذا لم يتم استخدام الحاسب الآلي.
- تكاليف جمع المعطيات من وحدات العينة تكون عالية عادة اذا انت وحدات العينة موزعة على مناطق جغرافية واسعة ومتباعدة ويتم معالجة هذا الاشكال باستخدام انواع أخرى من العينات .

مثال 01

ليكن لدينا مجتمع متكون من خمسة مفردات N=5 هي : A,B,C,D,E حيث نرغب في اختيار عينة عشوائية حجمها n=2

- أزواج المفردات التي يمكن اختيارها كعينة بسيطة (بدون ارجاع) (توفيقة لاختيار عنصرين من بين خمسة عناصر من المجتمع) هي كالتالي: AB, AC, AD, AE, BC, BD, BE, CD, CE, DE

ان عدد العينات الممكن اختيارها ذات الحجم n=2 تساوي 10 يعطى بالعلاقة التالية $C_5^2=10$

أما الاحتمال المتساوي لاختيار كل عينة فيحسب بالعلاقة: $\frac{1}{\frac{n}{N}}=\frac{1}{C_5^2}=\frac{1}{10}$

لكل من هذه العينات نفس الاحتمال في الظهور ويساوي $1/10$ ، كما انه لكل وحدة في المجتمع نفس الاحتمال للظهور في عينة الدراسة ويساوي $1/2$

– نقوم بكتابة كل زوج من هذه المفردات على ورقة صغيرة ثم تخلط هذه الأوراق ثم نسحب ورقة واحدة بحيث تحوي على زوج من المفردات تعتبر عينة بسيطة، وتتطلب هذه الطريقة حصر جميع العينات (زوج المفردات) التي يمكن سحبها من المجتمع وكتابة كل عينة على ورقة، اذا كان حجم المجتمع محل الدراسة صغيرا فان هذه الطريقة يمكن تطبيقها بسهولة وخلافا لذلك اذا كان حجم المجتمع كبيرا فان اختيارها يتطلب جهدا كبيرا .

2.1.6.1/ *أسلوب المعاينة العشوائية المنتظمة*

يمتاز هذا الأسلوب بالسهولة والبساطة في التطبيق، يقوم هذا النوع من المعاينة على اختيار المفردة الأولى عشوائيا ثم تتحدد باقي مفردات العينة تلقائيا بإضافة مقدار ثابت يسمى طول الفترة الى رقم المفردة الأولى فنحصل على رقم المفردة الثانية ويتكرر كما أن هذه الطريقة تضمن انتشار عينة الدراسة على أكبر مساحة من المجتمع لأن أسلوب السحب يتم بانتظام متسلسل.

يعتبر أسلوب المعاينة المنتظمة الخطية الأسلوب الأكثر شيوعا في العينات المنتظمة ويلخص فيما يلي :

– لدينا مجتمع يتكون من N مفردة وحجم العينة المطلوب سحبها هو n

– ليكن المقدار k الذي يعرف بأنه فترة الانتظام حيث $k = N/n$.

– نختار رقم عشوائي يقع بين 1 و k يسمى هذا الرقم برقم البداية العشوائية ويرمز له بـ l، يكون الرقم الأول في العينة هو l ، الرقم الثاني في العينة l+k والثالث هو l+2k وهكذا .....

من أهم مزايا العينة المنتظمة هو سهولة سحب العينات وتوزيع العينة على المجتمع بشكل جيد وتعتبر العينة المنتظمة فعالة مقارنة بالعينة العشوائية البسيطة لكثير من المجتمعات وخاصة إذا كان ميل الصفة المدروسة في المجتمع خطي.

أما عيوبها ففي حالة وجود صفة دورية في المجتمع قد يؤدي الى التحيز في العينات المسحوبة وفي التقدير ، فمثلا ضمن مجموعة من الأسر تم سحب عينة فرد واحد من كل أسرة وكان رقم البداية هو 1 وفترة العينة هي 2 وكان الأفراد الذين تسلسلهم 3 هم المختارين في العينة، سنلاحظ أن

جميع أفراد العينة هم ذوي الترتيب الثالث في الأسرة، واذا كان ترتيب الأفراد في الأسرة هو الأب، الأم،الابن، الابنة فسوف نلاحظ أن العينة جميعها "أبناء" مما يؤدي الى تحيز في التقديرات المطلوبة.

وتعتبر أسلوب المعاينة العشوائية المنتظمة بديلا للمعاينة العشوائية البسيطة للأسباب التالية:

- تتميز تطبيق المعاينة المنتظمة بالسهولة عند الاختيار مما يجعل أخطاء الاختيار الممكن حدوثها أقل من غيرها في باقي أنواع المعاينة خاصة في حالة عدم توفر اطار جيد للمجتمع .
- يمكن للمعاينة المنتظمة أن توفر معلومات أكثر من المعاينة العشوائية البسيطة بالنسبة لتكلفة وحدة المعاينة.
- قد يكون حجم المجتمع غير معروف بدقة أو حتى بصورة تقريبية مما يجعل من الصعوبة اختيار المعاينة العشوائية البسيطة في المقابل يمكن استخدام المعاينة العشوائية المنتظمة بسهولة .

مثال 02

تحتوي قائمة أحد البنوك التجارية على 1500 حساب جاري للزبائن ، ويرغب باحث احصائي في اختيار عينة منتظمة من 15 حسابا لمراجعتها، فما هي أرقام الحسابات الجارية التي يمكن اختيارها للمراجعة.

لدينا n=15 N=1500 , $k = N/n = 1500/15 = 100$

نختار رقما عشوائيا من جداول الأرقام العشوائية من بين الأرقام 01 الى غاية 100 ، وليكن على سبيل المثال الرقم العشوائي الأول هو 72 فيكون الحساب الجاري رقم 72 هو الحساب الأول في العينة المطلوبة

الحساب الجاري الثاني : 100+72 = 172 ، الحساب الجاري الثالث : (100)2+ 72=272 ، الى غاية الحساب الجاري الخامس عشر: (100)14 + 72=1472

3.1.6.1/*المعاينة العشوائية الطبقية*

نظرا لصعوبة تحقق التجانس في العديد من المجتمعات(حالة المعاينة العشوائية البسيطة) فانه يتم اللجوء الى تطبيق أسلوب المعاينة العشوائية الطبقية بحيث يقسم المجتمع الى عدد من المجموعات غير المتداخلة بحيث كل مجموعة تكون متجانسة للصفة المدروسة وتسمى طبقة وذلك بهدف الحصول على نتائج أكثر دقة، فعلى سبيل المثال عند دراسة متوسط دخل الأسرة يمكن تقسيم المجتمع الى ريف وحضر يوجب مراعات الدقة عند اجراء ما يلي ما يلي:

- تكوين الطبقات.
- عدد الطبقات المراد عملها.
- حجم العينة في كل طبقة.

تحليل البيانات لتصميم العينة الطبقية.

تقوم المعاينة العشوائية الطبقية على سحب عينة عشوائية بسيطة من كل مجموعة (طبقة) على حدى ثم تقدر معالم المجتمع المختلفة من خلال جميع العينات العشوائية البسيطة التي تم اختيارها من جميع الطبقات (المجموعات).

يمكن تلخيص خطوات اختيار **العينة العشوائية الطبقية** فيما يلي:

الخطوة الأولى : يتم فيها تقسيم المجتمع غير المتجانس الى مجتمعات صغيرة حجم كل منها هو $N_1, N_2, N_3 \dots N_k$ تكون كل منها متجانسة في الصفة المدروسة مثلا ان تكون متجانسة في صفة العمر أو الدخل ...بحيث يتحقق : $\mathrm{N} = N_1 + N_2 + N_3 + \cdots + N_k$

الخطوة الثانية : يتم فيها اختيار عينة عشوائية بسيطة من كل طبقة $n_1, n_2, n_3 \dots n_k$ بحيث تكون العينة المختارة من الطبقات المختلفة تمثل العينة العشوائية الطبقية بمعنى أن $n = n_1 + n_2 + n_3 + \cdots + n_k$:

- طرق اختيار عدد وحدات العينة التي يتم سحبها من كل طبقة :

✓ **طريقة الاختيار التناسب Oproportional allocation Method**

من خلال هذه الطريقة فان حجم العينة العشوائية المأخوذة من كل طبقة الى حجم العينة النهائي يكون متناسبا لحجم تلك الطبقة مع الحجم الكلي للمجتمع ويمكن صياغة ذلك بالعلاقة:

$$W_i = \frac{N_i}{N} = \frac{n_i}{n}$$

W_i : نسبة حجم العينة العشوائية رقم i الى حجم العينة الكلي يكون متناسبا لحجم تلك الطبقة مع الحجم الكلي للمجتمع

$$n_i = n\left({N_i}/{N}\right)$$

N : حجم المجتمع الكلي

n : حجم العينة الكلي

مثال03 : بافتراض أن لدينا مجتمع احصائي مكون من 25 أسرة بحيث أن المصروفات الشهرية بالدينار لكل من هذه الأسر وكانت البيانات على النحو التالي :

الأسرة	1	2	3	4	5	6	7	8	9	10	11	12	13
المصاريف	48	43	44	19	16	14	18	12	17	15	10	46	42
الأسرة	14	15	16	17	18	19	20	21	22	23	24	25	-
المصاريف	38	45	41	40	50	32	23	30	29	24	26	24	-

المطلوب : سحب عينة عشوائية طبقية تتكون من 08 أسر باستخدام طريقة الاختيار التناسب

الحل:

من خلال ملاحظة البيانات نجد انه يمكن تقسيمها الى ثلاث طبقات وهي :

الطبقة الأولى $N_1 = 8$: 10 ,15,17,12,18,14,16,19

الطبقة الثانية $N_2 = 7$: 32,23,24,26,27,29,30

الطبقة الثالثة $N_3 = 10$: 50,40,41,45,38,42,46,44,43,48

باستخدام الصيغة $n_i = n\left({N_i}/{} \right)$ نجد حجم كل عينة مختارة من كل طبقة كما يلي :

$$n_1 = 8\left({8}/{25}\right) = 2.56 \cong 3$$
$$n_2 = 8\left({7}/{25}\right) = 2.24 \cong 2$$
$$n_3 = 8\left({10}/{25}\right) = 3.2 \cong 3$$

وكخطوة أخيرة فانه يتم استخدام احدى طرق المعاينة العشوائية البسيطة مثلا الجداول العشوائية وفق الخطوات السابقة فنتحصل على وحدات العينة التي تظهر من كل طبقة على النحو التالي :

n_1 : 10, 17 ,14

n_2 : 23, 27

n_3 : 44, 41, 38

وبالتالي فان وحدات العينة هي : 14, 17, 10, 23, 27, 44, 41 ,38

- ✓ **طريقة الاختيار العشوائي(الاختيار الأمثل لـ Nymen)** : تقوم هذه الطريقة على أساس تقليل التباين وبافتراض أن تكاليف اختيار الوحدة متساوية فان صيغة علاقة هذه الطريقة تعطى كما يلي:

$$n_i = n = \frac{N_i S_i}{\sum N_i S_i}$$

n : تمثل حجم العينة الطبقية

S_i : الانحراف المعياري

4.1.6.1/ *المعاينة العشوائية العنقودية*

في حالة ارتفاع تكلفة الحصول على اطار يضم جميع مفردات المجتمع أو/و ترتفع تكلفة مقابلة مفردات المجتمع نظرا لبعد المسافة بين تلك المفردات فان المعاينة العنقودية تصبح الأقل تكلفة مقارنة بالمعاينة العشوائية البسيطة أو الطبقية نظرا لأن مفردات المجتمع التي سيتم معاينتها متلاصقة ومتجاورة في تجمعات وفي مجموعات مما يجعل تكلفة الانتقال من مفردة لأخرى أقل بكثير من تكلفة الانتقال من مفردة لأخرى في حالة المعاينة العشوائية البسيطة . فاذا كانت مفردات المجتمع على هيئة تجمعات (عناقيد) Clusters فنجد أن كل تجمع أو عنقود يحتوي على العديد من مفردات المجتمع و يمكن المعاينة من هذا المجتمع باختيار عدة عناقيد عشوائيا من بين كل المجموعات أو التجمعات التي يتكون منها المجتمع ثم ندرس جميع المفردات التي توجد داخل هذه المجموعات التي تم اختيارها عشوائيا، وبالتالي فان المعاينة العنقودية هي عينة عشوائية بسيطة من مجموعات أو عناقيد المجتمع حيث كل عنقود أو مجموعة يتكون من عدد من المفردات تمثلا تجمعا.

2.6.1/ *أساليب المعاينة غير احتمالية (غير عشوائية)*

في بعض الحالات يكون تطبيق المعاينة الاحتمالية غير ممكن أو غير مرغوب فيه، وبالتالي يتم استخدام المعاينة غير الاحتمالية، و من طرق اختيار العينات غير احتمالية نذكر:

1.2.6.1/ *المعاينة التحكمية Judgment Sampling* ، حيث أن الباحث له سلطة مباشرة أو غير مباشرة في عملية اختيار مفردات العينة، وتكون العينة التحكمية مفيدة في حالة ما اذا رأى الباحث أن بعض مفردات المجتمع تكون أفضل من مفردات أخرى في تمثيل مجتمع الدراسة.

ان المعاينة التحكمية تكون مفيدة وبنتائج ايجابية في حال ما كان صاحب القرار ذا حكم في عملية اختيار مفردات العينة

2.2.6.1/ *المعاينة الحصصية* Quota Samling : تستخدم في استقصاء آراء الناس فعلى سبيل المثال تعطى حصص محددة للمقابلين من الأشخاص ذو مستويات اجتماعية مختلفة، مجموعات من العمر، وفي مناطق مختلف ...، تستخدم مبادئ العينات الطبقية وهي أقل تكلفة من المعاينة العشوائية.

3.2.6.1/ *معاينة التناسب* **Convenience sampling:** يصممها المستجوب لأسباب عملية، بشكل عام يتم دعوة أشخاص في الطريق عند الخروج من صندوق الدفع للمتاجر...، انها الأقل تكلفة والأسرع من بين جميع تقنيات المعاينة ولكن تمتاز بالتحيز في الاختيار وعدم التمثيل.

7.1/ *تعريف توزيعات المعاينة*

نفترض أنه لدينا مجتمع مكون من مفردات تتبع توزيعا احتماليا معينا ونقوم بسحب عينة ذات حجم معين (n) من مجتمع حجمه (N) فيكون أمامنا عدد كبير من العينات التي يمكن سحبها. نفترض أننا قمنا بسحب عينة حجمها (n) وقمنا بحساب المتوسط الحسابي ثم سحبنا عينة ثانية لها نفس الحجم (n) وحسبنا منها أيضا المتوسط الحسابي ، ثم عينة ثالثة (حجمها n)،...الخ.وهكذا بالنسبة لجميع العينات التي يمكن سحبها من هذا المجتمع. فنجد أمامنا عدد كبير من القيم لنفس المقياس (المتوسط الحسابي) تختلف من عينة لأخرى، وهذا المقياس يعتبر متغير عشوائي يأخذ قيما مختلفة ويخضع لتوزيع معين يسمى **توزيع المعاينة للمتوسط الحسابي** (Sampling Distribution of the sample Mean) وهو بمثابة توزيع جميع المتوسطات الحسابية للعينات المأخوذة من نفس هذا المجتمع ذات الحجم (N) .

بعبارة أخرى فان توزيع المعاينة هو التوزيع الاحتمالي لجميع القيم الممكنة لأحد المقاييس الإحصائية الوصفية التي نحصل عليها عند سحب جميع العينات العشوائية المتساوي الحجم والمسحوبة من نفس المجتمع وبنفس الكيفية. وكذلك بالنسبة للتباين هو توزيع جميع التباينات المحسوبة من عينات لها نفس الحجم (n) ومسحوبة من نفس المجتمع.

ملاحظة : لإيجاد عدد العينات المسحوبة من المجتمع حجمه N هناك حالتين حالة السحب بالإرجاع وحالة السحب بدون ارجاع.

أ- في حالة **السحب بالإرجاع** فإن عدد العينات الممكنة يساوي : N^n

ب- في حالة **السحب بدون ارجاع** فإن هناك حالتين:

- **في حالة الترتيب مهم:** عدد العينات هو(عبارة عن ترتيبة) : $A_N^n = \binom{n}{N} = \frac{N!}{(N-n)!}$
- **في حالة الترتيب غير مهم:** عدد العينات هو(عبارة عن توفيقة): $C_N^n = \binom{n}{N} = \frac{N!}{(N-n)!n!}$

***مثال04*:**

بافتراض لدينا مجتمع احصائي مكون من ثلاث أعداد **N=3** وهي {2،3،4} في حالة المعاينة بالإرجاع أوجد :

عدد العينات الممكن تكوينها عند سحب عينة عشوائية بحجم **n=2** من هذا المجتمع مع تحديد العينات الممكنة؟

- في حالة السحب بالإرجاع (المعاينة بالإرجاع) فان عدد العينات الممكن اختيارها ذات الحجم
- n = 2 و المسحوبة من مجتمع حجمه N=5 هو: $N^n=3*3=3^2=9$ ؟
- العينات الممكن اختيارها هي : (2,2)، (2,3)، (2,4)، (3,2)، (3,3)، (3,4)، (4,2)، (4,3)، (4,4)

2/ توزيع المعاينة للمتوسط الحسابي

يتضمن هذا الجزء **متوسط** توزيعات المعاينة للمتوسطات الحسابية، **تباين** توزيع المعاينة للمتوسطات، طبيعة توزيع المعاينة للمتوسط الحسابي (طبيعة التوزيع الاحتمالي) في حالة التي تكون فيها المعاينة بالارجاع أو بدون ارجاع وتباين المجتمع معلوم أو مجهول .

1.2/ متوسط وتباين توزيع المعاينة للمتوسط الحسابي

مثال 05:

من خلال المثال السابق في حالة المعاينة بالارجاع أوجد:

1. المتوسط الحسابي μ و الانحراف المعياري σ لمتغير الدراسة في المجتمع الاحصائي ؟
2. المتوسط الحسابي (التوقع الرياضي) لتوزيع المعاينة للمتوسط الحسابي ثم قارن بين $E(\bar{X})$ و μ ؟
3. تباين توزيعات المعاينة للمتوسط الحسابي (تباين المتوسط الحسابي) $\sigma^2_{\bar{X}}$ ثم قارنه مع σ^2/n ؟

الحل :

1. المتوسط والانحراف المعياري في مجتمع الدراسة

$$\mu = \frac{\sum X_i}{N} = \frac{2+3+4}{3} = 3$$

$$Var(X_i) = \sigma^2 = \frac{\sum (X_i - \mu)^2}{N}$$

$$Var(X_i) = \frac{1}{3}[(2-3)^2 + (3-3)^2 + (4-3)^2] = 2/3$$

$$\sigma = \sqrt{Var(X_i)} = \sqrt{2/3}$$

2. متوسط (التوقع الرياضي) توزيع المعاينة للمتوسط الحسابي $E(\bar{X})$

أولا يجب حساب مختلف قيم المتوسط الحسابي وفقا لمختلف العينات الممكن اختيارها (**عمليا نختار عينة واحدة فقط)**

العينات الممكنة (السحب الارجاع)	(2,2)	(3,2)	(4,2)	(2,3)	(3,3)	(4,3)	(2,4)	(3,4)	(4,4)
المتوسطات الممكنة للعينة ($\bar{X}_i$)	2	2.5	3	2.5	3	3.5	3	3.5	4

$$E(\bar{X}) = \sum \bar{X}_i . P(\bar{X}_i) = \frac{\sum \bar{X}_i}{9} = \frac{2+2.5+3+2.5+3+3.5+3+3.5+4}{9} = 3$$

$$E(\bar{X}) = \mu = \mathbf{3}$$

3. تباين توزيع المعاينة للمتوسط الحسابي $\sigma_{\bar{X}}^2$

$$\mathrm{Var}(\bar{X}) = \boldsymbol{\sigma_{\bar{X}}^2} = \mathrm{E}(\bar{X}_i - E(\bar{X}))^2 = \mathrm{E}(\bar{X}_i - \mu)^2$$

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

$$\mathrm{Var}(\bar{X}_i) = \frac{1}{9}[\sum(2-3)^2 + (2.5-3)^2 + (3-3)^2 + \cdots + (4-3)^2] = \frac{3}{9} = \frac{1}{3}$$

$$\boldsymbol{\sigma_{\bar{X}}^2} = \frac{1}{3}$$

لدينا مما سبق الت**باين في المجتمع:** $Var(\quad_i) = \sigma^2 = 2/3$

من خلال ما سبق نلاحظ أن : $\sigma^2/n = \frac{2/3}{2} = \frac{1}{3} = \boldsymbol{\sigma_{\bar{X}}^2}$

نستنتج أن : $\boldsymbol{\sigma_{\bar{X}}^2} = \boldsymbol{\frac{\sigma^2}{n}}$

<u>مثال 06</u>

نفس المثال السابق في **<u>حالة المعاينة بدون ارجاع</u>** (مع العلم أن الترتيب غير مهم) أوجد ما يلي :

1. العينات الممكن اختيارها (الحالات الكلية) ؟
2. متوسط توزيع المعاينة للمتوسط الحسابي ثم قارن بين $E(\bar{X})$ و μ ؟
3. تباين توزيعات المعاينة للمتوسط الحسابي (تباين المتوسط الحسابي) $\bar{X}$ ثم قارنه مع $\frac{\sigma^2}{n}\frac{(N-1)}{(N-1)}$ ؟

الحل:

1. العينات الممكنة ذات الحجم n = 2 المسحوبة من مجتمع حجمه N=5 بدون ارجاع مع العلم أن الترتيب غير مهم هي عبارة عن توفيقة لاختيار مفردتين من بين خمسة أفراد و عددها هو:

$$C_3^2 = \frac{3!}{(3-2)! \times 2!} = \frac{3 \times 2!}{1! \times 2!} = 3 \qquad C_5^2 = 3$$

2. متوسط توزيع المعاينة للمتوسط الحسابي

أولا يجب حساب مختلف قيم المتوسط الحسابي وفقا لمختلف العينات الممكن اختيارها .

العينات الممكنة (السحب بدون ارجاع)	(3,2)	(4,2)	(4,3)
المتوسطات الممكنة للعينة ($\bar{\ }_i$)	2.5	3	3.5

$$E(\bar{X}) = \sum \bar{X}_i . P(\bar{X}_i) = \frac{\sum \bar{x}_i}{3} = \frac{2.5+3.5+3}{3} = 3$$

$$E(\bar{X}) = \mu = \mathbf{3}$$

3. تباين توزيع المعاينة للمتوسط الحسابي

$$\mathrm{Var}(\bar{\mathrm{X}}_i) = \boldsymbol{\sigma_{\bar{X}}^2} = \mathrm{E}(\bar{X}_i - E(\bar{X}))^2 = \mathrm{E}(\bar{X}_i - \mu)^2$$
$$= \frac{1}{n}\sum(\bar{X}_i - \mu))^2$$
$$\mathrm{Var}(\bar{\mathrm{X}}_i) = \frac{1}{3}[\sum(2.5-3)^2 + (3-3)^2 + (3.5-3)^2] = \frac{0.5}{3} = \frac{1}{6}$$
$$\boldsymbol{\sigma_{\bar{X}}^2} = \frac{1}{6}$$

لدينا مما سبق **التباين في المجتمع** : $Var(\boldsymbol{\bar{X}_i}) = \sigma^2 = 2/3$

المقارنة بين : $\boldsymbol{\sigma_{\bar{X}}^2}$ و $\frac{\sigma^2}{n}\frac{(N-n)}{(N-1)}$

$$\frac{\sigma^2}{n}\frac{(N-n)}{(N-1)} = \frac{2/3}{2}\frac{(3-2)}{(3-1)} = \frac{1}{6} = \sigma_{\bar{X}}^2$$

نستنتج أن : $\boldsymbol{\sigma_{\bar{}}^2} = \frac{\boldsymbol{\sigma^2}}{\boldsymbol{n}}\frac{\boldsymbol{(N-1)}}{\boldsymbol{(N-1)}}$

نظرية 01: لتكن X_i متغيرة عشوائية بحجم n و متوسط $\bar{X}$ مسحوبة من مجتمع احصائي بمتوسط $\boldsymbol{\mu}$ وتباين $\boldsymbol{\sigma^2}$ ، في حالة المجتمع محدود (السحب **بالإرجاع**) فإن :

$$\boldsymbol{E(\bar{X}_i) = \mu}$$
$$\boldsymbol{\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}}$$

البرهان: لتكن Xi قيم المتغيرة الأصليةX فيكون لدينا:

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

$$\sigma_{\bar{X}}^2 = V\mathrm{ar}(\bar{X}) = Var\left(\frac{\sum_i^n X_i}{n}\right) = \frac{1}{n^2}Var\left(\sum_i^n X_i\right) = \frac{1}{n^2}\sum_i^n Var(X_i)$$
$$= \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

نظرية 02: لتكن X_i متغيرة عشوائية بحجم n مسحوبة من مجتمع احصائي، في حالة المجتمع غير محدود (المعاينة بدون إرجاع) :

$$\boldsymbol{E(\bar{X}_i) = \mu}$$

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

$\left(\frac{N-n}{N-1}\right)$: معامل التصحيح المحدود للمجتمع **finite population correction factor**

- **ملاحظة : نستخدم معامل الارجاع** $\left(\frac{N-n}{N-1}\right)$ اذا تحقق الشرط التالي: $\frac{n}{N} \geq 0.05$ (في هذه الحالة تكون المعاينة بدون ارجاع) .

2.2/ دراسة توزيع المعاينة للمتوسط الحسابي للعينة

لتكن X متغيرة عشوائية في مجتمع احصائي بمتوسط μ وتباين σ^2 حيث أن قيم X_i مستقلة فيما بينها، سيكون لدينا عدة حالات حسب معلومية طبيعة التوزيع الاحتمالي للمتغيرة المدروسة و التباين في المجتمع وحجم العينة n وبافتراض **حالة السحب بالارجاع** **(أو المجتمع غير محدود)** نوجز الحالات التالية :

1.2.2/ حالة التوزيع الطبيعي للمجتمع و التباين σ^2 معلوم في المجتمع

إذا كانت المتغيرة X_i في مجتمع الدراسة تتبع **التوزيع الطبيعي** بمتوسط μ والتباين σ^2 **معلوم** فإن متوسط توزيع المعاينة للمتوسط الحسابي ذو التوقع الرياضي (المتوسط) $(\bar{X}) = \mu$ والتباين $\sigma^2_{\bar{x}}$ يتبع التوزيع الطبيعي مهما كان حجم العينة : $\bar{X} \rightsquigarrow N(\mu, \sigma^2_{\bar{x}})$

ان المتغيرة المعيارية $Z \rightsquigarrow N(0,1)$ (بمتوسط 0 وانحراف معياري يساوي 1) والتي تمثل إحصائية التوزيع الطبيعي المعياري المحسوبة انطلاقا من عينة الدراسة تعطى بالعلاقة :

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \rightsquigarrow N(0,1)$$

ملاحظة : لحساب الاحتمالات للمتغير العشوائي $\bar{X}$ يتم تحويله من خلال إيجاد القيم المعيارية له Z والاستعانة بالجدول الاحصائي الخاص بالتوزيع الطبيعي المعياري.

مثال07:

لدينا مجتمع احصائي غير محدود تخضع فيه المتغيرة العشوائية x للتوزيع الطبيعي بمتوسط حسابي $\mu = 76$ وانحراف معياري $\sigma = 20$، تم اختيار من هذا المجتمع الاحصائي عينة عشوائية حجمها n=16 .

المطلوب :

1. أحسب التوقع الرياضي $E(\bar{X})$ و الانحراف المعياري $\sigma_{\bar{X}}$ للمتوسط الحسابي للعينة ؟
2. أوجد احتمال أن يكون المتوسط الحسابي للعينة أكبر من 83.5 بمعنى $Pr(\bar{X} \geq 83.5)$ ؟
3. أوجد احتمال أن يكون المتوسط الحسابي للعينة محصورا بين70و80 بمعنى $Pr(70 \leq \bar{X} \leq 80)$؟

الحل :

لدينا : σ^2 **معلوم** ، المتغيرة X في المجتمع تتبع **التوزيع الطبيعي** وبالتالي توزيع المعاينة للمتوسط الحسابي يتبع أيضا التوزيع الطبيعي : $\bar{X} \rightsquigarrow N(\mu, \sigma^2_{\bar{X}})$

1. حساب $E(\bar{X})$ و $\sigma_{\bar{X}}$

$$E(\bar{X}) = \mu = 76$$

المجتمع غير محدود و بالتالي : $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$

$$\sigma^2_{\bar{X}} = \frac{(20)^2}{16} = 25$$

$$\sigma_{\bar{X}} = \sqrt{25} = 5$$

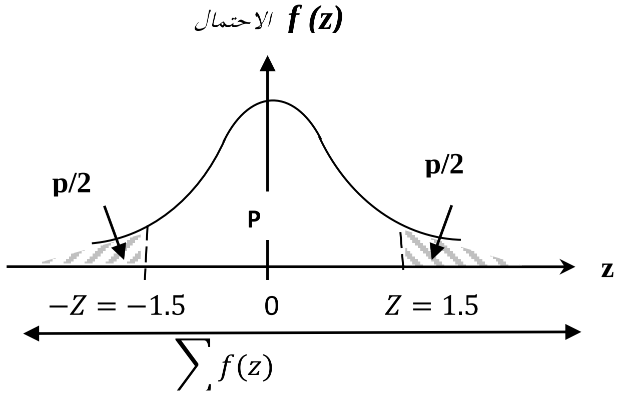
2. حساب $P(\bar{X} \geq 83.5)$

نحسب أولا قيمة المتغير العشوائي Z حيث : $Z = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}} = \frac{83.5-76}{5} = 1.5$

$$P(\bar{X} \geq 83.5) = P\left(Z \geq \frac{83.5-\mu}{\sigma_{\bar{X}}}\right) =$$

$$= P\left(Z \geq \frac{83.5-76}{5}\right) = P(Z \geq 1.5)$$

نعلم أن التوزيع الطبيعي يكون متناظر وبالتالي:



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

من الجدول الاحصائي للتوزيع الطبيعي نوجد قيم الاحتمالات الموافقة لقيم Z :

$$= 1 - 0.93319 = 0.06681$$

$$P(\bar{X} \geq 83.5) = 0.06681$$

3. حساب $P(70 \leq \bar{X} \leq 80)$

$$P(70 \leq \bar{X} \leq 80) = P(Z_1 \leq Z \leq Z_2)$$
$$P\left(\frac{70-\mu}{\sigma_{\bar{X}}} \leq Z \leq \frac{80-\mu}{\sigma_{\bar{X}}}\right) = P\left(\frac{70-76}{5} \leq Z \leq \frac{80-76}{5}\right)$$
$$= P(-1.2 \leq Z \leq 0.8) = P(Z \leq 0.8) - P(Z \leq -1.2) \ldots (01)$$

نعلم أن التوزيع الطبيعي يكون متناظرا وبالتالي يكون لدينا :

$$P(Z \leq -1.2) = P(Z \geq 1.2) = 1 - P(Z \leq 1.2) \ldots (2)$$

بتعويض (02) في (01) نجد :

$$P(Z \leq 0.8) - P(Z \leq -1.2) = P(Z \leq 0.8) - [1 - P(Z \leq 1.2)]$$

من الجدول الاحصائي للتوزيع الطبيعي نوجد قيم الاحتمالات الموافقة لقيم Z :

$$= 0.78814 - [1 - 0.88493] = 0.67307$$

$$P(70 \leq \bar{X} \leq 80) = 0.67307$$

2.2.2/ *حالة التوزيع الطبيعي في المجتمع، والتباين σ^2 مجهول في المجتمع*

إذا كانت X في المجتمع تتبع **التوزيع الطبيعي** و **σ^2 مجهول** ، فتكون الاحصائية $T = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}}$ تتبع توزيع ستيودنت (T−Student) بدرجة حرية (v = n−1)

$$\boldsymbol{\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{S^2}{n}}$$

بما أن **التباين في المجتمع مجهول** فنقوم بتقديره **بالتباين الغير متحيز في العينة** $\boldsymbol{S^2}$ حيث :

$$\boldsymbol{S^2 = \mathrm{Var}(X_i) = \frac{1}{n-1}\sum(X_i - \bar{X})^2}$$

<u>**ملاحظة**</u> : في حالة ما يكون حجم العينة أكبر أو يساوي **30** فان توزيع ستيودنت يؤول الى التوزيع الطبيعي (خاصية تقريب التوزيعات الاحتمالية) وبالتالي فان المتغيرة العشوائية $\bar{X}$ تتبع التوزيع الطبيعي .

3.2.2/ *حالة المجتمع لا يتبع التوزيع الطبيعي*

نميز ثلاث حالات :

1.3.2.2/ *حجم العينة* 30≤n

اذا كان المجتمع الذي تسحب منه العينة العشوائية لا تتبع فيه المتغيرة العشوائية التوزيع الطبيعي فان متوسط توزيع المعاينة للمتوسط الحسابي يتبع التوزيع الطبيعي حسب نظرية النهاية المركزية عندما يكون حجم العينة n كبيرا (**30≤n**) بمعنى :

$$\bar{X} \rightsquigarrow N(\mu, \sigma^2/_n)$$

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

نظرية: (نظرية النهاية المركزية Central Limit Theory)

اذا كان لدينا مجتمع احصائي بمتوسط حسابي μ وتباين σ^2 ليس بالضرورة تكون فيه المتغيرة العشوائية تتبع التوزيع الطبيعي (دالة كثافة الاحتمال ليس بالضرورة طبيعية) نسحب من هذا المجتمع عينة حجمها n فحسب **نظرية النهاية المركزية** فان المتغيرة العشوائية $\bar{X}$ تؤول الى التوزيع الطبيعي عندما يكون حجم العينة كبيرا (حدده بعض الاحصائيين بـ **30≤n**).

مثال08 :

بافتراض لدينا مجتمع احصائي يحتوي على **N=900** مفردة بمتوسط حسابي $\mathbf{\mu = 20}$ و انحراف معياري $\boldsymbol{\sigma = 12}$ ، نختار عينة عشوائية ذات الحجم **n=36** .

المطلوب : في حالة **المعاينة بالارجاع** :

1. أحسب المتوسط الحسابي $\boldsymbol{E(\bar{X})}$ والانحراف المعياري $\boldsymbol{\sigma_{\bar{X}}}$ للمتوسط الحسابي للعينة ؟
2. أوجد طبيعة التوزيع الاحتمالي للمتغيرة العشوائية $\boldsymbol{\bar{X}}$ ؟
3. أحسب احتمال أن تكون $\boldsymbol{\bar{X}}$ محصورة بين 18 و 24 بمعنى $\boldsymbol{P(18 \le \bar{X} \le 24)}$ ؟

الحل :

1. حساب $E(\bar{X})$ و $\sigma_{\bar{X}}$ ؟

$$E(\bar{X}) = \mu = 20$$

- نعتبر أن المجتمع الاحصائي محدودا اذا تحقق الشرط : $\boldsymbol{\frac{n}{N} \ge 0.05}$

من خلال المعطيات نجد أن : $\frac{n}{N}=\frac{36}{900}=0.04<0.05$ وبالتالي فالمجتمع الاحصائي يعتبر غير محدود (المعاينة بالارجاع) اذا : $\sigma_{\bar{X}}^2=\frac{\sigma^2}{n}$

$$\boldsymbol{\sigma_{\bar{X}}^2}=\frac{(12)^2}{36}=4$$
$$\sigma_{\bar{X}}=\sqrt{4}=2$$

2. طبيعة التوزيع الاحتمالي للمتغيرة العشوائية $\bar{\boldsymbol{X}}$

بما أن : $n=36>30$ اذا حسب نظرية النهاية المركزية فان : $\bar{\boldsymbol{X}}\rightsquigarrow \boldsymbol{N}(\boldsymbol{\mu},\boldsymbol{\sigma^2}/\boldsymbol{n})$

3. حساب الاحتمال $\boldsymbol{P(18\le\bar{X}\ \le)}$

$$P(18\le\bar{X}\ \le 24)=P(Z_1\le Z\le Z_2)$$
$$=\ P(\frac{18-\mu}{\sigma_{\bar{X}}}<Z<\frac{24-\mu}{\sigma_{\bar{X}}})$$
$$=P(\frac{18-20}{2}\le Z\le\frac{24-20}{2})=P(-1\le Z\le 2)$$
$$=P(Z\le 2)-P(Z\le -1)\ \ldots\ (1)$$

نعلم أن **التوزيع الطبيعي يكون متناظر** : $P(Z\le -1)=P(Z\ge 1)=\ \ldots(02)$ $[1-P(Z\le 1)]$

بتعويض (02) في (01) نجد :

$$P(Z\le 2)-P(Z\le -1)=P(Z\le 2)-[1-P(Z\le 1)]$$

$$=0.97725-[1-0.84134]=0.81859$$
$$P(18\le\bar{X}\le 24)=0.81859$$

2.3.2.2/***حجم العينة n<30 والتباين في المجتمع $\boldsymbol{\sigma^2}$ معلوم***

إذا كان المجتمع الذي سحبت منه العينة **ليس بالضرورة طبيعي** و**التباين** σ^2 **معلوم و** حجم العينة المسحوبة **n<30** فإننا نقوم بتطبيق متراجحة **Bienaymé Tchebytchev** والتي تنص على أنه مهما يكن (k) فإن : $\forall k>0:\quad P(|\bar{x}-\mu|\le k\sigma_{\bar{X}})\ge 1-\frac{1}{k^2}$

3.3.2.2/***حجم العينة n<30 والتباين في المجتمع $\boldsymbol{\sigma^2}$ مجهول*** : في هذه الحالة لا يوجد حل، الا بافتراض طبيعية توزيع X في المجتمع.

3/ توزيع المعاينة للفرق بين متوسطي عينتين مستقلتين

1.3/ التوقع الرياضي و التباين للفرق بين متوسطي عينتين مستقلتين

ليكن لدينا مجتمعين احصائيين، بافتراض أن المتغيرة العشوائية المدروسة في كل منهما **تتبع التوزيع الطبيعي** وأن المجتمعين الاحصائيين غير محدودين (أو **المعاينة بالارجاع**) ، المجتمع الأول حجمه N1 بمتوسط μ_1 وتباين σ_1^2 والمجتمع الثاني حجمه N2 بمتوسط μ_2 وتباين σ_2^2، نسحب من كل مجتمع عينة عشوائية حيث العينتين المسحوبتين مستقلتين، العينة الأولى مسحوبة من المجتمع الأول بحجم n_1 بمتوسط $\bar{X}_1$ وتباين المتوسط $\sigma_{\bar{X}_1}^2$، والعينة الثانية مسحوبة من المجتمع الثاني بحجم n_2 بمتوسط $\bar{X}_2$ وتباين المتوسط $\sigma_{\bar{X}_2}^2$.

- **التوقع الرياضي للفرق بين متوسطي العينتين:**

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

- **تباين الفرق بين متوسطي العينتين :**

$$Var(\bar{X}_1 - \bar{X}_2) = Var(\bar{X}_1) + Var(\bar{X}_2) - 2Cov(\bar{X}_1, \bar{X}_2)$$

في حالة العينات المستقلة فان : $Cov(\bar{X}_1, \bar{X}_2) = 0$ و بافتراض أن المجتمعين غير محدودين (أو المعاينة بالإرجاع) :

$$Var(\bar{X}_1 - \bar{X}_2) = \sigma_{\bar{X}_1-\bar{X}_2}^2 = Var(\bar{X}_1) + Var(\bar{X}_2)$$

$$\boldsymbol{\sigma_{\bar{X}_1-\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ملاحظة : في حالة المعاينة بدون ارجاع فان تباين توزيع المعاينة للفرق بين متوسطي عينتين (مستقلتين) يساوي :

$$\sigma_{\bar{X}_1-\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1}\left(\frac{N_1 - n_1}{N_1 - 1}\right) + \frac{\sigma_2^2}{n_2}\left(\frac{N_2 - n_2}{N_2 - 1}\right)$$

2.3 / *دراسة توزيع المعاينة للفرق بين متوسطي عينتين مستقلتين*

1.2.3/ *التباين معلوم في مجتمعين موزعين توزيعا طبيعيا*

في هذه الحالة فان توزيع المعاينة للفرق بين المتوسطين يؤول الى التوزيع الطبيعي مهما يكن حجم العينيتين ونكتب :

$$(\bar{X}_1 - \bar{X}_2) \rightsquigarrow N(E(\bar{X}_1 - \bar{X}_2), \sigma^2_{\bar{X}_1-\bar{X}_2})$$

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

تعطى إحصائية التوزيع الطبيعي Z (المتغيرة المعيارية Z تتبع التوزيع الطبيعي المعياري بمتوسط 0 وتباين 1) بالعلاقة التالية :

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}} \rightsquigarrow N(0,1)$$

مثال 09:

بافتراض أن **المعاينة بالارجاع،** سحبت عينة عشوائية من مجتمع احصائي أول $X_1 \rightsquigarrow N(30;25)$ حجمها $n_1 = 30$ وعينة عشوائية ثانية مسحوبة من مجتمع ثاني $X_2 \rightsquigarrow N(20;16)$ حجمها $n_2 = 35$

المطلوب: أحسب الاحتمال $P(\bar{X}_1 - \bar{X}_2 \leq 12)$ ؟

الحل:

$\mu_1 = 30$ $\sigma_1^2 = 25$ $n_1 = 30$

$\mu_2 = 20$ $\sigma_2^2 = 16$ $n_2 = 35$

بما أن: متغير الدراسة في كلا المجتمعين يتبع التوزيع الطبيعي والتباين معلوم , $X_1 \rightsquigarrow N(\mu_1, \sigma_1^2)$

$X_2 \rightsquigarrow N(\mu_2, \sigma_2^2)$

فان : $(\bar{X}_1 - \bar{X}_2) \rightsquigarrow N(E(\bar{X}_1 - \bar{X}_2), \sigma^2_{\bar{X}_1-\bar{X}_2})$

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2 = 30 - 20 = 10$$

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{25}{30} + \frac{16}{35} = 1.29$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = 1.135$$

$$(\bar{X}_1 - \bar{X}_2) \rightsquigarrow N(10; 1.29)$$

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}} \rightsquigarrow N(0,1)$$

$$Pr(\bar{X}_1 - \bar{X}_2 \leq 10) = Pr\left(Z \leq \frac{12 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}\right)$$

باستخدام الجدول الاحصائي للتوزيع الطبيعي المعياري نجد :

$$= Pr\left(Z \leq \frac{12-10}{1.135}\right) = Pr(Z \leq 1.762) = 0.9608$$

اذا : $Pr(\overline{X}_1 - \overline{X}_2 \leq 12) = 0.9608$

2.2.3/ التباين مجهول في مجتمعين غير موزعين توزيعا طبيعيا وحجم العينتين كبير

في حالة ما التباين مجهول في مجتمعين غير موزعين توزيعا طبيعيا وحجم العينتين كبير $n_1, n_2 \geq 30$ ، فانه وفقا لنظرية النهاية المركزية فان توزيع المعاينة للفرق بين المتوسطين يقترب من التوزيع الطبيعي بمعنى :

$$(\overline{X}_1 - \overline{X}_2) \rightsquigarrow N(\mu_1 - \mu_2, \sigma^2_{\bar{X}_1-\bar{X}_2})$$

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

حيث : S_1^2 ، S_2^2 تمثلان التباين الغير متحيز في العينة الأولى والثانية على التوالي (كل منهما يعتبر مقدر للتباين في المجتمع)

$$S_2^2 = \frac{1}{n_2-1}\sum(X_{i2} - \bar{x}_1)^2 \ , \qquad S_1^2 = \frac{1}{n_1-1}\sum(X_{i1} - \bar{x}_1)^2$$

3.2.3/حالة التباين مجهول في مجتمعين موزعين توزيعا طبيعيا

نميز حالتين :

➢ ***الحالة الأولى***: $\boldsymbol{\sigma_1^2 = \sigma_2^2}$

التباين في المجتمعين مجهول ومتساوي ويتم تقديرهما من خلال **التباين المشترك** $\hat{\boldsymbol{\sigma}}$ ويصبح:

$\sigma_1^2 = \sigma_2^2 = \hat{\sigma}^2$ ، و تعطى علاقة **التباين المشترك** كما يلي:

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

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{\hat{\sigma}^2}{n_1} + \frac{\hat{\sigma}^2}{n_2} = \hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

تعطى صيغة الإحصائية **T** التي تخضع لتوزيع ستيودنت بدرجة حرية $v = n_1+n_2-2$ على النحو التالي:

$$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\overline{X}_1-\overline{X}_2}} \rightsquigarrow t(n_1+n_2-2)$$

مثال10:

بافتراض أنه تم سحب عينتين عشوائيتين من مجتمعين موزعين توزيعا طبيعيا بتباين مجهول ولكن متساوي $\sigma_1^2 = \sigma_2^2$، المتوسط الحسابي في المجتمع الأول $\mu_1 = 600$ حيث تم سحب عينة عشوائية أولى من هذا المجتمع بحجم $n_1 = 20$ و بانحراف معياري $S_1 = 30$، أما المجتمع الثاني فالمتوسط الحسابي فيه يساوي $\mu_2 = 400$ حيث تم سحب عينة عشوائية من هذا المجتمع بحجم $n_2 = 18$ بانحراف معياري $S_2 = 25$.

المطلوب: أحسب الاحتمال ($P(\bar{X}_1 - \bar{X}_2 \geq 215$ ؟

الحل :

لدينا التباين مجهول ومتساوي في المجتمعين حيث $n_1 = 18$، $n_2 = 20$ و ، اذا توزيع العينية يؤول الى توزيع ستيودنت و تعطى صيغة إحصائية **T** لتوزيع ستيودنت بدرجة حرية $v = n_1 + n_2 - 2$ كما يلي:

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow t(n_1 + n_2 - 2)$$

$$T = \frac{\bar{X}_1 - \bar{X}_2 - 200}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow t(36)$$

$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$\hat{\sigma}^2 = \frac{(20-1)30^2 + (18-1)25^2}{20 + 18 - 2} = 770.138$$

$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = 770.138 \left(\frac{1}{20} + \frac{1}{18}\right) = 81.292$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{81.292} = 9.016$$

$$P(\bar{X}_1 - \bar{X}_2 \geq 215) = P\left(T \geq \frac{215 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right)$$

$$= P\left(T \geq \frac{215 - 200}{9.016}\right)$$

$$P(T \geq 1.663)$$

من خلال الجدول الاحصائي للتوزيع الاحتمالي لستيودنت نجد:

$$P(T \geq 1.663) = 0.05$$

- **الحالة الثانية : $\boldsymbol{\sigma_1^2 \neq \sigma_2^2}$**

تعطى صيغة الإحصائية **T** التي تخضع لتوزيع ستيودنت بدرجة حرية **v** على النحو التالي:

$$T = \frac{\overline{\mathrm{X}}_1 - \overline{\mathrm{X}}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow \mathrm{t}(v)$$

حيث :

$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

تعطى علاقة درجة الحرية بالصيغة التالية :

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(S_2^2/n_2\right)^2}{n_2 - 1}}$$

4/ *توزيع المعاينة للنسبة*

1.4/ *تذكير بالتوزيعات الاحتمالية المتقطعة (توزيع برنولي، توزيع ذو الحدين)*

تنقسم التوزيعات الاحتمالية الى قسمين أساسيين القسم الأول: هي التوزيعات الاحتمالية المستمرة خاصة بالمتغيرات العشوائية المستمرة مثل التوزيع الطبيعي، توزيع ستيودنت، توزيع فيشر وتوزيع مربع كاي...، أما القسم الثاني فهي التوزيعات الاحتمالية المتقطعة خاصة بالمتغيرات العشوائية المتقطعة مثل توزيع ثنائي الحدين، توزيع بواسون...

تستخدم التوزيعات ***الاحتمالية المتقطعة*** في العديد من المسائل في مجال التسيير الصناعي، التجاري وفي الادارة ومن أكثرها شيوعا التوزيع ثنائي الحدين، توزيع بواسون، التوزيع الهندسي الزائد، التوزيع الهندسي المتعدد...

1.1.4/ ***توزيع برنولي*** **Distribution bernoulli** لتكن لدينا تجربة عشوائية تحتمل نتيجتين فقط النجاح أو الفشل، X متغيرة عشوائية تأخذ القيمة 1عند تحقق نتيجة النجاح، والقيمة 0 في الحالة المعاكسة (أي عند تحقق نتيجة الفشل).

p:احتمال النجاح (تحقق الظاهرة محل الدراسة) **q** : احتمال الفشل (عدم تحقق الظاهرة محل الدراسة)

$P(X=1)=p$ $\qquad$ $P(X=0)=q$ $\qquad$ **X**=$\{0,1\}$ $\qquad$ **p**+**q**=1

خصائص توزيع برنولي : $\qquad X \rightsquigarrow B(1, \mathrm{p})$

$$E(X)=p \quad , \quad Var(X)=\sigma^2=\mathrm{pq}$$

X:متغير عشوائي يسمى توزيع يمثل نجاح أو فشل التجربة العشوائية ويسمى بمتغير برنولي وتوزيعه الاحتمالي هو توزيع برنولي

2.1.4/ ***توزيع ذو الحدين The Binomial Distribution***

في حالة التجارب العشوائية المستقلة التي تجرى عدة مرات بمعنى تجربة برنولي مكررة n مرة بحيث خلال كل محاولة نتحصل على نتيجتين فقط هما اما النجاح أو الفشل باحتمال ثابت لكل منها في كل مرة تعاد التجربة.

ليكن X متغير عشوائي يمثل عدد مرات النجاح يتبع توزيع احتمالي معين يسمى توزيع ذو الحدين .

احتمال النجاح هو **P** واحتمال الفشل هو **q**=$(\boldsymbol{1-p})$

n : تمثل عدد مرات تكرار التجربة العشوائية (عدد المحاولات)

احتمال المتغير العشوائي X الذي يتبع توزيع ذي الحدين يعطى بالعلاقة التالية :

$$\boldsymbol{P(X)=c_x^n p^x(1-p)^{n-x} \quad ; \quad x=0,1,2\ldots n}$$

X : عدد مرات النجاح **n,p** : معلمتي هذا التوزيع (نسبة النجاح وعدد مرات التكرار على التوالي)

ونعبر عن أن X تتبع توزيع ثنائي الحدين : $X \rightsquigarrow B(n,\mathrm{p})$

المتوسط والانحراف المعياري للمتغيرة العشوائية لثنائي الحدين بالمعلمتين n و p

ونعبر عن أن X تتبع توزيع ثنائي الحدين : $X \rightsquigarrow B(n,\mathrm{p})$

$$E(X)=np \quad ; \quad \sigma^2=npq \quad ; q=1-p \ ,$$

- ***مبادئ توزيع ثنائي الحدين***

- يوجد محاولة مكررة
- كل محاولة تتكون من نتيجتين فقط هما النجاح أو الفشل
- كل محاولة مستقلة عن المحاولة الأخرى
- احتمال النجاحp (أو الفشل) ثابت من محاولة الى أخرى ولا يتغير

- **تقريب توزيع ثنائي الحدين الى التوزيع الطبيعي**

- اذا كان لدينا X متغير عشوائي يتبع توزيع ثنائي الحدين $X \rightsquigarrow B(n,\mathrm{p})$ فان توزيع ثنائي الحدين يقترب من التوزيع الطبيعي وفقا لنظرية النهاية المركزية اذا كان حجم العينة كبير ، ويكون كبيرا كفاية اذا كان اذا تحقق الشرط التالي :

$$\mathrm{np}\geq 5 \ , \ \mathrm{n}(1-\mathrm{p})\geq 5$$

هناك قاعدة تقريب أخرى اذا كان : $\mathrm{n}(1-\mathrm{p})\geq 10$, $\mathrm{np}\geq 10$ أو , $\mathrm{np}(1-\mathrm{p})\geq 10$

تعطى احصائية التوزيع الطبيعي Z (المتغيرة المعيارية) بمتوسط 0 وتباين 1 كما يلي:

$$Z=\frac{X-E(X)}{\sigma}=\frac{X-np}{\sigma} \rightsquigarrow N(0,1)$$

لتقريب توزيع ذو الحدين (ثنائي الحدين) وهو توزيع متقطع الى التوزيع الطبيعي وهو توزيع مستمر فانه يجب ادراج معامل التصحيح المستمر correction factor for continuity كما يلي :

- ✓ التقريب في حالة الاحتمال $\boldsymbol{P(X\leq a)}$ **أو** $\boldsymbol{P(X>a)}$: فان معامل التصحيح يكون على الشكل $(a+0.5)$

$$P(X>a)=P\left(Z>\frac{(a+0.5)-np}{\sigma}\right)$$

$$P(X\leq a)=P\left(Z\leq\frac{(a+0.5)-np}{\sigma}\right)$$

✓ التقريب في حالة الاحتمال $Pr(X < a)$ أو $(X \geq a)$ فان معامل التصحيح يساوي: $(a - 0.5)$

$$P(X < a) = P\left(Z < \frac{(a-0.5)-np}{\sigma}\right)$$
$$P(X \geq a) = P\left(Z \geq \frac{(a-0.5)-np}{\sigma}\right)$$

التقريب في حالة الاحتمال$P(a \leq X \leq b)$ فان معامل التصحيح يكون على الشكل التالي :

$$P\left(\frac{(a-0.5)-np}{\sigma} \leq Z \leq \frac{(b+0.5)-np}{\sigma}\right)$$

- اذا كان k عددا صحيحا موجبا بين 0 و n فان :

$$P(X = k) = P(k - 0.5 \leq X \leq k + 0.5)$$
$$= P\left(\frac{k-0.5-np}{\sigma} \leq Z \leq\right)\frac{k+0.5-np}{\sigma}$$

وجد ياتس Yates أن تقريب توزيع ثنائي الحدين (توزيع متقطع) الى التوزيع الطبيعي (توزيع مستمر) يكون دقيقا وواقعيا اذا تم اجراء تصحيح على قيمة المتغيرة العشوائية x بإضافة أو طرح القيمة 0.5 منها، ويسمى بتصحيح ياتس أو معامل التصحيح، والذي يعد ضروريا في الواقع عند حساب الاحتمال $P(X = x)$.

ملاحظة : اذا كان حجم العينة كبير جدا فانه يتم اهمال معامل التصحيح، في الأمثلة الموالية نعتبر أن حجم العينة كبير وبالتالي سنهمل معامل تصحيح ياتس (معامل التصحيح).

<u>مثال11</u>:

وفقا للجداول التي قدمها المركز الوطني للإحصاءات الصحية في الإحصاءات الحيوية للولايات المتحدة، هناك احتمال بنسبة **80٪** تقريبا أن يكون الشخص البالغ من العمر 20 عاما أن يكون على قيد الحياة في سن 65 عاما، بافتراض أنه تم اختيار ثلاثة أشخاص من عمر 20 عاما بشكل عشوائي.

<u>المطلوب</u>

1. حدد التوزيع الاحتمالي لعدد من هم على قيد الحياة عند عمر 65 عاما.

أوجد احتمال أن يكون عدد الأشخاص على قيد الحياة في العمر 65 سنة سيكون :

- اثنان بالضبط **P(X=2)**
- على الأكثر واحد **P(X≤1)**
- على الأقل واحد **P(X≥1)**

الحل

- يمثل النجاح هو أن الشخص البالغ من العمر 20 عامًا سيكون على قيد الحياة في سن 65
- لنفترض أن X تشير إلى عدد الأشخاص (من بين ثلاثة) الذين هم على قيد الحياة عند سن 65 (عدد الناجحين)
- إن احتمال أن يكون الشخص الحالي البالغ من العمر 20 عاما على قيد الحياة في سن 65 هو 80٪، أي p =0.8
- عدد المحاولات هو عدد الأشخاص في الدراسة ، وهو ثلاثة لذا فان n=3

1. تحديد التوزيع الاحتمالي لعدد من هم على قيد الحياة عند عمر 65 أي صيغة احتمال ثنائي الحدين :

$$\mathbf{P}(X=x) = \binom{n}{x} P^X (1-p)^{n-x}, \;\; X = 0,1,2 \ldots . n$$

من أجل n=3 و p=0.8 فان :

$$P(X=x) = \binom{3}{x} 0.8^X (0.2)^{3-x}, \;\; X = 0,1,2 \ldots . n$$

X : متغير عشوائي لثنائي الحدين

- ✓ **حساب احتمال أن يكون شخصان على قيد الحياة في سن 65 من بين الأشخاص الثلاثة المختارين عشوائيا P(X=2)**

من أجل 2 = *X* : $P(X=2) = \binom{3}{2} 0.8^2 (0.2)^{3-2} = 0.384$

باحتمال 0.384 أن اثنين من الأشخاص الثلاثة المختارين عشوائيا سيكونون على قيد الحياة في سن 65.

- ✓ **حساب احتمال أن يكون شخص واحد على الأكثر على قيد الحياة في سن 65 من بين الأشخاص الثلاثة المختارين عشوائيا P(X≤1)**

$$P(X\leq 1)=P(X=0)+ P(X=1)$$

$$= \binom{3}{0} 0.8^0 (0.2)^{3-0} + \binom{3}{1} 0.8^1 (0.2)^{3-2}$$

$$= 0.008 + 0.096 = 0.104$$

باحتمال 0.104 يوجد شخصًا واحدًا على الأكثر من الأشخاص الثلاثة المختارين سيكونون على قيد الحياة في سن 65.

- ✓ **حساب احتمال أن يكون شخص واحد على الأقل على قيد الحياة في سن 65 من بين الأشخاص الثلاثة المختارين عشوائيا $P(X\geq 1)$**

$$P(X\geq 1)= P(X=1) +P(X=2)+ P(X=3)$$

أو نكتب بعلاقة أسهل :

$$P(X\geq 1)=1- P(X<1)$$

$$=1- P(X=0)$$

$$=1-0.008=0.992$$

باحتمال0.992 يوجد شخص واحد على الأقل من الأشخاص الثلاثة المختارين سيكونون على قيد الحياة في سن 65.

2.4/ *التعريف بتوزيع المعاينة للنسبة*

لتكن X متغيرة عشوائية تتبع توزيع ذو الحدين (التوزيع الثنائي) $\boldsymbol{X} \rightsquigarrow \boldsymbol{B}(N;p)$ ، يتم رصد خاصية معينة تدرس في مجتمع الدراسة يرمز لها بالرمز p مثلا نسبة الناجحين، نسبة القطع السليمة من انتاج مصنع معين، نسبة المرضى بمرض معين.... حيث أن النسبة p في المجتمع (نسبة النجاح) تساوي

$$p = \frac{\text{عدد المفردات التي تحقق خاصية الدراسة في المجتمع}}{\text{حجم المجتمع}} = \frac{X}{N}$$

في حالة ما تم أخذ جميع العينات الممكنة ذات الحجم n من مجتمع احصائي حجمه N ، فانه يمكن أن تختلف النسب $\hat{p}$ بين مختلف العينات وهذا ما يعني أن النسبة $\hat{p}$ في العينة تتبع سلوك متغير عشوائي لها توزيع احتمالي معين يسمى **توزيع المعاينة للنسبة** حيث أن:

$$\hat{p} = \frac{\text{عدد المفردات التي تحقق خاصية الدراسة في العينة}}{\text{حجم العينة}} = \frac{X}{n}$$

$\hat{p}$ تمثل نسبة تحقق خاصية الدراسة في العينة (نسبة النجاح في العينة)

$$E(\hat{p}) = p \quad ; \quad Var(p) = \sigma_{\hat{p}} = \frac{pq}{n} \quad ; \quad q = 1-p$$

يقترب توزيع ثنائي الحدين للمتغيرة العشوائية x من التوزيع الطبيعي وفقا لنظرية النهاية المركزية اذا كان حجم العينة كبير ، ويكون كبيرا كفاية اذا كان : $np\geq 5 \quad , \quad n(1-p)\geq 5$

$$\boldsymbol{X} \rightsquigarrow \boldsymbol{N}(np;\sigma)$$
$$\hat{\boldsymbol{p}} \rightsquigarrow \boldsymbol{N}(\text{p};\sigma_{\hat{p}})$$

$$Z=\frac{\widehat{\boldsymbol{p}}-E(\widehat{\boldsymbol{p}})}{\sigma_{\widehat{\boldsymbol{p}}}}=\frac{\widehat{\boldsymbol{p}}-p}{\sigma_{\widehat{\boldsymbol{p}}}}\rightsquigarrow N(0,1)$$

بما أن $\widehat{\boldsymbol{p}}$ هي مجرد تابع للمتغيرة العشوائي لثنائي الحدين x فان التصحيح يتم باضافة أو طرح معامل التصحيح $\frac{\mathbf{1}}{(\mathbf{2n})}$

نعلم أن المتغير العشوائية لثنائي الحدين X (بالمعلمتين n و p) تمثل عدد النجاحات ضمن n محاولة حيث خلال كل محاولة تأخذ المتغيرة العشوائية X القيمة 0 عند النجاح أو القيمة 1 عند الفشل :

البرهان الرياضي

$$\boldsymbol{E}(\widehat{\boldsymbol{p}})=E\left(\frac{X}{n}\right)=\frac{1}{n}E(X)=\frac{1}{n}np=p$$

$\widehat{\boldsymbol{p}}=\frac{X}{n}$ تمثل نسبة النجاح في العينة المختارة (يعبر كذلك عن متوسط العينة)

$$\boldsymbol{Var}(\widehat{\boldsymbol{p}})=\sigma_{\widehat{\boldsymbol{p}}}^2=Var\left(\frac{X}{n}\right)=\frac{1}{n^2}Var(X)=\frac{1}{n^2}npq=\frac{pq}{n}$$

<u>مثال12</u> : عند القاء قطعة نقدية 30 مرة، أحسب احتمال ان تكون نسبة مرات الحصول على الصورة أقل أو يساوي 50% .

<u>الحل</u> :

$$\widehat{\boldsymbol{p}}\rightsquigarrow \boldsymbol{N}(\mathbf{p};\sigma_{\widehat{\boldsymbol{p}}})$$

$$P(\widehat{\boldsymbol{p}}\ \leq 0.5)=P\left(\widehat{\boldsymbol{p}}\leq 0.5+\frac{\mathbf{1}}{(\mathbf{2n})}\right)=P(\widehat{\boldsymbol{p}}\leq 0.516)$$

$$\widehat{\boldsymbol{p}}=\frac{X}{n}=\frac{X}{30}$$

$$P(\widehat{\boldsymbol{p}}\ \leq 0.5)=P\left(\frac{X}{30}\ \leq 0.5\right)=P(X\ \leq 15)$$

<u>الاحتمال</u> $P(\widehat{\boldsymbol{p}}\leq 0.5)$ **هو في الأصل** $P(\boldsymbol{X}\leq 15)$

$$Z=\frac{b-p}{\sigma_{\widehat{\boldsymbol{p}}}}$$

$$\sigma_{\widehat{\boldsymbol{p}}}^2=\frac{pq}{n}=\frac{0.5(0.5)}{30}=0.0083$$

$$\sigma_{\widehat{\boldsymbol{p}}}=0.0911$$

لأن X متغيرة عشوائية متقطعة فانه وجب إدخال معامل التصحيح واستبدال الاحتمال السابق بالاحتمال $P(X\ \leq 15.5)$

ويقابله $P\left(\widehat{\boldsymbol{p}}\geq 0.5+\frac{\mathbf{1}}{(\mathbf{2n})}\right)$ ليصبح الاحتمال المطلوب **كما يلي** :

$$P(\widehat{\boldsymbol{p}} \leq 0.516) = P\left(Z \leq \frac{0.516 - P}{\sigma_{\widehat{\boldsymbol{p}}}}\right)$$

$$P\left(Z \leq \frac{0.516 - 0.5}{0.0911}\right) = P(Z \leq 0.1756) = 0.56749$$

مثال13:

في مصنع ينتج عادة 25% من علب العصير كبيرة الحجم (01 لتر) ، سحب من هذا المصنع عينة عشوائية من 2200 عبوة تبين أن منها 500 عبوة كبيرة الحجم.

المطلوب :

1. أحسب نسبة العبوات الكبيرة الحجم في العينة المختارة ؟
2. أوجد التوزيع الاحتمالي لنسبة العبوات الكبيرة الحجم في العينة $\widehat{\boldsymbol{p}}$ ؟
3. أحسب احتمال أن ينتج المصنع من العبوات الكبيرة الحجم النسبة أقل أو يساوي 26% خلال فترة الدراسة ؟

الحل :

المتغير العشوائي X في مجتمع الدراسة يتبع توزيع ثنائي الحدين بنسبة نجاح p=0.25

1. نسبة العبوات الكبيرة الحجم في العينة المختارة

$$\widehat{\boldsymbol{p}} = \left(\frac{X}{n}\right) = \frac{500}{2200} = 0.2272$$

$$\widehat{\boldsymbol{p}} = 22.72\%$$

2. في حالة حجم العينة كبير فانه يمكن اهمال معامل التصحيح عند تقريب توزيع ثنائي الحدين (توزيع متقطع) الى التوزيع الطبيعي (توزيع احتمالي مستمر)

بما أن حجم العينة كبير فانه وفقا لنظرية النهاية المركزية بعد تحقق الشرطين :

np=550>5 , n(1−p)=1650>5

فان التوزيع الاحتمالي للمتغيرة $\widehat{\boldsymbol{p}}$ تقترب من التوزيع الطبيعي ونكتب :

$$\hat{p} \rightsquigarrow N(\mathrm{p}; \sigma_{\hat{p}})$$

$$Z = \frac{\hat{p} - p}{\sigma_{\hat{p}}}$$

$$\sigma_{\hat{p}}^2 = \frac{pq}{n} = \frac{(0.25)(0.75)}{2200} = 8.522 \times 10^{-5}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.25)(0.75)}{2200}} = 0.0092$$

بإهمال معامل التصحيح لأن حجم العينة كبير :

$$(\hat{\boldsymbol{p}} \le 0.26) = P\left(Z \le \frac{0.26 - p}{\sigma_{\hat{\boldsymbol{p}}}}\right)$$

$$P\left(Z \le \frac{0.26 - 0.25}{0.0092}\right)$$

$$= P(Z \le 1.08) = 0.8599$$

مثال14:

اذا علمت أنه من خلال دراسات سابقة في مدينة معينة وجد أن 24% من زبائن بريد الجزائر يفضلون سحب أموالهم عن طريق بطاقة السحب عبر الشبابيك ، فاذا تم اختيار عينة عشوائية من 400 زبون

المطلوب : أوجد احتمال أن يكون :

- ✓ على الأكثر 30% من هؤلاء الزبائن في العينة يفضلون سحب أموالهم عبر بطاقة السحب ؟
- ✓ على الأكثر 28% من هؤلاء الزبائن في العينة يفضلون سحب أموالهم عبر بطاقة السحب ؟

الحل :

بما أن حجم العينة كبير فانه وفقا لنظرية النهاية المركزية بعد تحقق الشرطين :

$$np=100>5 \quad , \quad n(1-p)=300>5$$

فان التوزيع الاحتمالي للمتغيرة $\hat{p}$ يقترب من التوزيع الطبيعي ونكتب : $\hat{p} \rightsquigarrow N(\mathrm{p}; \sigma_{\hat{p}}^2)$

$$\sigma_{\hat{\boldsymbol{p}}}^2 = \frac{pq}{n} = \frac{(0.24)(0.75)}{400} = \mathbf{0,000456}$$

$$\sigma_{\hat{\boldsymbol{p}}} = 0.021354$$

$$Z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} \rightsquigarrow N(0; 1)$$

- ✓ حساب احتمال على الأكثر 30% من هؤلاء الزبائن في العينة يفضلون سحب أموالهم عبر بطاقة السحب

$$(\hat{\boldsymbol{p}} \le 0.30) = P\left(Z \le \frac{0.30 - p}{\sigma_{\hat{\boldsymbol{p}}}}\right)$$

$$P\left(Z \le \frac{0.30 - 0.24}{0.021354}\right)$$

$$= P(Z \leq 2.80) =$$

$$= P(Z \leq 1.08) = 0.9974$$

حساب احتمال على الأكثر 28% من هؤلاء الزبائن في العينة يفضلون

سحب أموالهم عبر بطاقة السحب

$$(\hat{\boldsymbol{p}} \ \geq 0.28) = P\left(Z \geq \frac{0.28 - p}{\sigma_{\hat{\boldsymbol{p}}}}\right)$$

$$P\left(Z \geq \frac{0.28 - 0.24}{0.021354}\right)$$

$$= P(Z \geq 1.87) = 1 - P(Z \leq 1.87)$$

$$= 1 - 0.96926 = 0,03074$$

5/ توزيع المعاينة للفرق بين نسبتي عينتين مستقلتين

بافتراض سحب عينة عشوائية حجمها n_1 من مجتمع احصائي $X_1 \rightsquigarrow B(N_1; p_1)$ بنسبة نجاح (أو تحقق الظاهرة محل الدراسة) p_1 ، وسحب عينة عشوائية ثانية حجمها n_2 من مجتمع احصائي ثاني $X_1 \rightsquigarrow B(N_2; p_2)$ بنسبة نجاح p_2 ، وكانت العينتين المسحوبتين مستقلتين ، تمثل $\hat{p}_1 - \hat{p}_2$ متغيرا عشوائيا للفرق بين نسبتي العينتين حيث يعطى كل من :

التوقع الرياضي للفرق بين نسبتي العينتين :

$$E(\hat{p}_1 - \hat{p}_2) = E(\hat{p}_1) - E(\hat{p}_2) = p_1 - p_2$$

التباين للفرق بين نسبتي العينتين :

$$Var(\hat{p}_1 - \hat{p}_2) = \sigma^2_{(\hat{p}_1 - \hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2}$$

$$= \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

يقترب توزيع ثنائي الحدين للمتغيرة العشوائية $\hat{p}_1 - \hat{p}_2$ من التوزيع الطبيعي وفقا لنظرية النهاية المركزية اذا كان حجم العينة كبير ، ويكون كبيرا كفاية اذا كان :

$$n_1 p_1 \geq 5 \quad n_2 p_2 \geq 5 \quad , \quad n_1\ q_1 \geq 5 \quad n_2\ q_2 \geq 5$$

$$(\hat{p}_1 - \hat{p}_2) \rightsquigarrow N(p_1 - p_2; \sigma_{\hat{p}_1 - \hat{p}_2})$$

وتعطى صيغة المتغيرة المعيارية **Z** التي تتبع التوزيع الطبيعي المعياري كما يلي :

$$\mathrm{Z} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \rightsquigarrow N(1; 0)$$

<u>مثال15</u>

اذا علمت أن شركة تنتج أجهزة الكترونية تشتري قطع غيار من موردين اثنين، حيث أنها تشتري قطع من المورد الأول و ترفض الغير صالحة منها بنسبة 8% بسبب عيوب تقنية، بينما القطع التي تشتريها من المورد الثاني ترفض منها ما نسبته 5% و أيضا بسبب عيوب تقنية، اذا كانت الشركة تشتري كل أسبوع 150 قطعة من المورد الأول و 300 قطعة من المورد الثاني.

المطلوب:أحسب احتمال أن يكون الفرق بين نسبتي الرفض في العينتين العشوائيتين أقل أو يساوي7% ؟

الحل:

$\hat{p}_1$: نسبة القطع المرفوضة في العينة العشوائية الأولى المختارة من المورد الأول

$\hat{p}_2$: نسبة القطع المرفوضة في العينة العشوائية المختارة من المورد الثاني

نستنتج أولا التوزيع الاحتمالي للفرق بين نسبتي العينتين

بما أن : $n_1p_1 = 12 > 5 \qquad n_2p_2 = 15 > 5 \qquad , \quad n_1 \; q_1 = 138 > 5 \qquad n_2$
$q_2 = 285 > 5$

و الملاحظ ان القطع التي تشتريها الشركة من الموردين الاثنين تتضمن عينات عشوائية، حيث أن حجم العينتين كبير كفاية فوفقا لنظرية النهاية المركزية فان توزيع ثنائي الحدين للمتغيرة العشوائية $\boldsymbol{\hat{p}_1 - \hat{p}_2}$ يقترب من التوزيع الطبيعي : $(\boldsymbol{\hat{p}_1 - \hat{p}_2}) \rightsquigarrow \boldsymbol{N(p_1 - p_2; \sigma_{\hat{p}_1 - \hat{p}_2})}$

حيث : $\mathrm{Z} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \rightsquigarrow N(1;0)$

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2 = 0.08 - 0.05 = 0.03$$

$$\sigma^2_{(\hat{p}_1 - \hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2}$$

$$= \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

$$\frac{(0.08)(0.92)}{150} + \frac{(0.05)(0.95)}{300} = 0.000649$$

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{0.0.000649} = 0.025475$$

نفترض أن حجم العينتين كبير وبالتالي نهمل معامل التصحيح للانتقال من توزيع ثنائي الحدين الى التوزيع الطبيعي .

$$P(\boldsymbol{\hat{p}_1 - \hat{p}_2} \geq 0.07)$$

$$= P\left[\mathrm{Z} \leq \left(\frac{\mathbf{0.07} - (\boldsymbol{p_1 - p_2})}{\boldsymbol{\sigma_{\hat{p}_1 - \hat{p}_2}}}\right)\right]$$

$$= P\left(\mathrm{Z} \geq \frac{\mathbf{0.07} - 0.03}{0.025475}\right)$$

$$= P(\mathrm{Z} \leq 1.570166) = 0.94179$$

6./ توزيع المعاينة للتباين

1.6/ تذكير حول التوزيع الاحتمالي لمربع كاي

لهذا التوزيع الاحتمالي عدة استخدامات في مجال الاحصاء الاستنتاجي منها اختبارات جودة المطابقة، والتجانس، الاستقلالية و التباين ...

<u>نظرية 01</u>: اذا كانت $X_1, X_2 \dots X_v$ متغيرات عشوائية مستقلة فيما بينها تتبع التوزيع الطبيعي المعياري $X_i \rightsquigarrow N(0;1)$ فان المتغيرة العشوائية $X = \sum_{i=1}^{v} X_i^2 = X_1^2 + X_2^2 + \cdots + X_v^2$ تتبع **توزيع مربع كاي** χ_v^2 بدرجة حرية v ونكتب :

$$X \sim \chi_v^2$$

تعطى دالة كثافة الاحتمال لهذا التوزيع بدرجة حرية v بالصيغة التالية :

$$f(X) = \frac{1}{(2)^{\frac{v}{2}}\,\Gamma\left(\frac{v}{2}\right)}(X)^{\frac{v}{2}-1}(e)^{-\frac{X}{2}} \quad , \qquad x > 0$$

ان هذا التوزيع هو حالة خاصة من توزيع جاما بالمعلمتين : $\beta = 2\ ,\ \alpha = \frac{d}{2}$

يمكن ايجاد قيم هذا التوزيع باستخدام الجدول الخاص بهذا التوزيع

حيث : $\Gamma(\alpha)$: هي الدالة قاما

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}\, e^{-t} dt \qquad , \alpha > 0$$

يمكن ايجاد قيم هذا التوزيع باستخدام الجدول الخاص بهذا التوزيع

اذا كان لدينا X متغير عشوائي حيث $\boldsymbol{X \rightsquigarrow N(\mu, \sigma^2)}$ فان : $Z = \frac{\bar{X}-\mu}{\sigma} \rightsquigarrow N(0,1)$

نعرف التوزيع الاحتمالي لمربعات Z بانه توزيع مربع كاي بدرجة حرية d=1 ويرمز لها بالرمز χ^2 (حيث درجة الحرية هي معلمة هذا التوزيع) ونعبر عن ذلك بالعلاقة التالية :

$$Z^2 = \left(\frac{\bar{X}-\mu}{\sigma}\right)^2 \rightsquigarrow {\chi^2}_{(1)}$$

نظرية02 :اذا كانت $U_1, U_2, U_3 \dots U_n$ متغيرات عشوائية مستقلة تتبع توزيع مربع كاي بدرجات حرية $d_1, d_2, d_3 \dots d_n$: على الترتيب فان:

$$\sum_{i=1}^{n} U_i \sim \chi_{(d)}^2$$

حيث : $d = \sum_{i=1}^{n} d_i$

- **خصائص هذا التوزيع الاحتمالي**

حسب تعريف توزيع مربع كاي $X_i \rightsquigarrow N(0;1)$ لدينا :

$$E(X_i^2) = Var(X_i) = 1, \quad ou\ E\left(\sum_{i=1}^{v} X_i^2\right) = E(X) = v$$

$$Var(X_i^2) = E(X_i^4) - \left(E(X_i^2)\right)^2 = \mu_4 - 1\ ;\ \mu_4 = 3$$

$$Var(X_i^2) = 2$$

$$Var\left(\sum_{i=1}^{v} X_i^2\right) = Var(X) = 2v$$

اذا نستنتج حسب تعريف توزيع مربع كاي $X_i \rightsquigarrow N(0;1)$ بان :

$$E(X) = v$$

$$Var(X) = 2v$$

***مثال16*:**

اذا كانت χ_v^2 متغيرة عشوائية تتبع توزيع مربع كاي بدرجة حرية v ، من خلال الجدول الاحصائي لتوزيع مربع كاي $\chi_{p,v}^2$ في العمود الأول لدرجة الحرية **v** والصف الأول للاحتمال **p** نجد احصائية مربع كاي :

- ✓ $Pr(\chi_{v=15}^2 > a) = 0.95$

عند درجة الحرية v=15 و P=0.95 نجد :

$$Pr(\chi_{15}^2 > 7.261) = 0.95$$

- ✓ $Pr(\chi_{11}^2 > b) = 0.05$

عند درجة الحرية v=11 و P=0.95 نجد :

$$Pr(\chi_{v=11}^2 > 19.675) = 0.95$$

2.6/ التعريف بتوزيع المعاينة للتباين

نظرية 03: اذا تم اختيار عينة عشوائية $X_1, X_2, \dots X_n$ بحجم n مسحوبة بالارجاع من مجتمع طبيعي $X \rightsquigarrow N(\mu;\sigma^2)$ فان:

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

ملاحظة : نعلم أن : $\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

من العلاقة السابقة لدينا $\left(\tilde{S}^2\right) = \frac{n-1}{n}\sigma^2$ وبالتالي نستنتج أن تباين في العينة $\tilde{S}^2$ هو مقدر متحيز لتباين المجتمع σ^2 لأن : $E\left(\tilde{S}^2\right) \neq \sigma^2$

يمكن كتابة العلاقة السابقة على النحو التالي :

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

نقول أن $\left(\frac{\tilde{S}^2 n}{n-1}\right)$ بمعنى S^2 بأنه مقدر غير متحيز لتباين المجتمع لأن : $E(S^2) = \sigma^2$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

التوقع والتباين للمتغيرة العشوائية S^2

$$E(S^2) = \sigma^2$$

$$Var(S^2) = \frac{2\sigma^4}{n-1}$$

<u>نظرية</u> : اذا كانت X متغيرة عشوائية في مجتمع احصائي محدود (أو المعاينة بدون ارجاع) و $\tilde{S}^2$ متغيرة عشوائية تمثل تباين عينة مسحوبة من ذات المجتمع فان:

$$E\left(\tilde{S}^2\right) = \sigma^2\left(\frac{n-1}{n}\right)\left(\frac{N}{N-1}\right)$$

ملاحظة :

- **التوقع الرياضي لـ** $\tilde{S}^2$

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

من أجل قيمة n كبيرة فان التوقع الرياضي لـ $\tilde{S}^2$ له قيمة النهاية هي σ^2

- **تباين** $\tilde{S}^2$

$$Var\left(\tilde{S}^2\right) = \frac{n-1}{n^3}\left[(n-1)\mu_4 - (n-3)\sigma^4\right]$$

μ_4 : العزم المركزي من الدرجة 4 للمتغير x

من أجل قيمة n كبيرة فان تباين $\tilde{S}^2$ له قيمة نهاية هي : $\frac{(\mu_4 - \sigma^4)}{n}$

<u>ملاحظة:</u>

من أجل $\tilde{S}^2$ وفقا للنظرية المركزية المحدودة فان المتغير العشوائي :

$$Z = \frac{\tilde{S}^2 - \left(\frac{n-1}{n}\sigma^2\right)}{\sqrt{Var\tilde{S}^2}}$$

لما يكون حجم العينة n يتجه للانهاية فان توزيع هذا المتغير يقترب الى متغير يتبع التوزيع الطبيعي المعياري $\boldsymbol{N}(\mathbf{0},1)$ ، نأخذ نهاية التوقع الرياضي والتباين من أجل قيمة n كبيرة تصبح العلاقة السابقة على الشكل :

$$Z = \frac{\tilde{S}^2 - \sigma^2}{\sqrt{(\mu_4 - \sigma^4)/n}} \rightsquigarrow N(0;1)$$

يقترب من التوزيع الطبيعي بمتوسط 0 وتباين يساوي 1

3.6/ طبيعة التوزيع الاحتمالي لتباين العينة

نظرية : اذا تم اختيار عينات عشوائية $X_1, X_2, \dots X_n$ بحجم n مسحوبة بالارجاع من مجتمع طبيعي بمتوسط μ وتباين σ^2 $\quad X \rightsquigarrow N(\mu;\sigma^2)$ فان :

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

المتغير العشوائي $\frac{(n-1)S^2}{\sigma^2}$ يتبع توزيع مربع كاي بدرجة حرية v= $n-1$

$\bar{X}\,, s^2$: متغيرين عشوائيين مستقلين

يمكن من خلال ماسبق استنتاج أن 200 :

$$\frac{(n-1)S^2}{\sigma^2} = \frac{n\tilde{S}^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \rightsquigarrow {\chi^2}_{(n-1)}$$

<u>مثال17:</u>

لتكن لدينا في مجتمع احصائي المتغيرة العشوائية x حيث $\boldsymbol{X} \rightsquigarrow \mathbf{N}(\boldsymbol{\mu};\mathbf{9})$ ، تم سحب من هذا المجتمع عينة عشوائية بحجم **n=21**

المطلوب: أحسب احتمال أن يزيد تباين العينة $\boldsymbol{S^2}$ عن القيمة **15** حيث : $\boldsymbol{S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}}$

الحل :

لدينا $\sigma^2 = 09$: n=21 المجتمع الاحصائي موزع توزيعا طبيعيا

$$Pr(S^2 \geq 15) = Pr\left(\frac{(n-1)S^2}{\sigma^2} \geq \frac{20(14)}{9}\right)$$

$$= Pr(\chi^2_{11} < 31.111) = 0.05$$

عند درجة الحرية v=20 نلاحظ من خلال الجدول الاحصائي أن أقرب قيمة لـ 31.111 هي 31.410 وتوافق الاحتمال 0.05 .

مثال18

لتكن لدينا X متغيرة عشوائية تتبع التوزيع الطبيعي بتباين 46 في مجتمع احصائي ما حيث :

$X \rightsquigarrow N(\mu; 46)$ نسحب من هذا المجتمع عينة عشوائية حجمها n=12

المطلوب :

1. ماهو احتمال أن يكون تباين العينة أكبرمن القيمة 10.9 $Pr(S^2 \geq 10.9)$ ؟
2. ماهو احتمال أن يكون تباين العينة أقل من القيمة 16 $Pr(S^2 < 16)$ ؟

الحل

1. حساب الاحتمال $Pr(S^2 \geq 10.9)$

$$Pr(S^2 \geq 10.9) = Pr\left(\frac{(n-1)S^2}{\sigma^2} \geq \frac{(12-1)10.9}{46}\right)$$
$$= Pr(\chi^2_{11} > 2.606) = 0.995$$

2. حساب الاحتمال $Pr(S^2 < 16)$

$$Pr(S^2 < 16) = Pr\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(12-1)16}{46}\right)$$
$$= Pr(\chi^2_{11} < 3.826) = 1 - Pr(\chi^2_{11} \geq 3.826) = 1 - 0.975 = 0.025$$

تمارين محلولة :

التمرين رقم 01

من خلال الجداول الاحصائية للتوزيعات الاحتمالية عند درجة الحرية v أوجد قيم الاحتمالات ***P*** وقيم إحصائية توزيع ستيودنت **t** والتوزيع الطبيعي **z** :

$P\left(T_{(v=12)} \le \ldots\right) = 0{,}9$	$P\left(T_{(v=7)} \ge -1.895\right) = \ldots\ldots$	$P\left(T_{(v=10)} \le -1.093\right) = \ldots\ldots$
$P(Z \le \ldots\ldots) = 0.9632$	$P(Z \le 1.71) = \ldots\ldots.$	$P(Z \le -1.42) = \ldots\ldots.$

الحل :

✓ $$P\left(T_{(v=10)} \le -1.093\right) = P\left(T_{(v=10)} \ge 1.093\right) = 0.15$$

✓ $$P\left(T_{(v=7)} \ge -1.895\right) = P\left(T_{(v=7)} \le 1.895\right) = 1 - P\left(T_{(v=7)} \ge 1.895\right) = 1 - 0.05 = 0.95$$

✓ $$P\left(T_{(v=12)} \le \ldots\right) = 0{,}9$$
$$1 - P\left(T_{(v=12)} \ge \ldots\right) = 0{,}9 \Rightarrow P\left(T_{(v=12)} \ge \ldots\right) = 0.1$$
$$P\left(T_{(v=12)} \ge 1.356\right) = 0.1$$

✓ $$P(Z \le -1.42) = P(Z \ge 1.42) = 1 - P(Z \le 1.42) = 1 - 0.9222 = 0.0778$$

✓ $$P(Z \le 1.71) = 0.9563$$

✓ $$P(Z \le 1.79) = 0.9632$$

التمرين رقم 02

بافتراض أن مجتمع احصائي يحتوي على نقاط خمسة طلاب **N=5** **(1 ، 3 ، 5، 7، 9)**، نختار عشوائيا من هذا المجتمع جميع العينات الممكنة ذات الحجم **n=2**.

أ/في حالة المعاينة بالارجاع أوجد :

1. عدد العينات الممكن اختيارها ؟
2. المتوسط الحسابي $\boldsymbol{\mu}$ و الانحراف المعياري $\boldsymbol{\sigma}$ لمتغير الدراسة في المجتمع الاحصائي ؟
3. التوقع الرياضي (المتوسط) لتوزيع المعاينة للمتوسط الحسابي $\overline{\boldsymbol{X}}$ ثم قارن بين $\boldsymbol{E}(\overline{\boldsymbol{X}})$ و µ ؟
4. تباين توزيعات المعاينة للمتوسط الحسابي (تباين المتوسط الحسابي) $\boldsymbol{\sigma^2_{\overline{X}}}$ ثم قارنه مع $\boldsymbol{\sigma^2/n}$ ؟

ب/في حالة المعاينة بدون ارجاع (مع العلم أن الترتيب غير مهم) أوجد ما يلي :

1. عدد العينات الممكن اختيارها ؟
2. متوسط توزيع المعاينة للمتوسط الحسابي $E(\bar{X})$ ثم قارنه مع **μ** ؟
3. تباين توزيعات المعاينة للمتوسط الحسابي $\sigma_{\bar{X}}^2$ ثم قارنه مع $\frac{\sigma^2}{n}\frac{(N-n)}{(N-1)}$ ؟

الحل:

أ. في حالة المعاينة بالارجاع

1. إيجاد عدد العينات الممكن اختيارها ؟

في حالة السحب بالإرجاع (المعاينة بالإرجاع) فان العينات الممكن اختيارها ذات الحجم n = 2 و المسحوبة من مجتمع حجمه N=5 عددها هو: $N^n=5*5=5^2=25$ ؟

2. المتوسط الحسابي μ و الانحراف المعياري σ لمتغير الدراسة في المجتمع الاحصائي

$$\mu = \frac{\sum X_i}{N} = \frac{1+3+5+7+9}{5} = \frac{25}{5} = 5$$

$$Var(X_i) = \sigma^2 = \frac{\sum (X_i-\mu)^2}{N}$$

$$Var(X_i) = \frac{1}{5}[(1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2] = 8$$

$$\sigma = \sqrt{Var(X_i)} = \sqrt{8} = 2.82$$

3. إيجاد التوقع الرياضي (المتوسط) لتوزيع المعاينة للمتوسط الحسابي $E(\bar{X})$ ثم مقارنته مع μ

نقوم بإيجاد قيم المتوسط الحسابي الذي يوافق جميع العينات الممكن اختيارها ذات الحجم 2 والتي عددها 25 كما يلي :

X \ X	1	3	5	7	9
1	1	2	3	4	5
3	2	3	4	5	6
5	3	4	5	6	7
7	4	5	6	7	8
9	5	6	7	8	9

$\bar{X}_i$	1	2	3	4	5	6	7	8	9
$P(\bar{X}_i)$	1/25	2/25	3/25	4/25	5/25	4/25	3/25	2/25	1/25

$$E(\bar{X}) = \sum \bar{X}_i . P(\bar{X}_i) = 1\left(\frac{1}{25}\right) + 2\left(\frac{2}{25}\right) + \cdots + 9\left(\frac{1}{25}\right) = 5$$

المقارنة : $E(\bar{X}) = \mu = \mathbf{5}$

في حالة المعاينة بالارجاع نستنتج أن المتوسط الحسابي هو مقدر غير متحيز لمتوسط المجتمع.

4. إيجاد تباين توزيع المعاينة للمتوسط الحسابي $\boldsymbol{\sigma^2_{\bar{X}}}$ و مقارنته مع $\boldsymbol{\sigma^2/n}$

$$\text{Var}(\bar{X}) = \boldsymbol{\sigma^2_{\bar{X}}} = \text{E}(\bar{X} - E(\bar{X}))^2 = \text{E}(\bar{X} - \mu)^2$$

$$\sum(\bar{X}_i - \mu)^2 P(\bar{X}_i)$$

$(\bar{X}_i - \mu)^2$	16	9	4	1	0	1	4	9	16
$(\bar{X}_i - \mu)^2 P(\bar{X}_i)$	16/25	18/25	12/25	4/25	0	4/25	12/25	18/25	16/25

$$\boldsymbol{\sigma^2_{\bar{X}}} = \left(\frac{16}{25}\right) + \left(\frac{18}{25}\right) + \cdots + \left(\frac{16}{25}\right) = 5 = \frac{100}{25} = 4$$

لدينا مما سبق **تباين المجتمع:** $Var(X_i) = \sigma^2 = 8$

$$\sigma^2/n = 8/2 = 4$$

اذا نستنتج أنه في حالة المعاينة بالإرجاع (أو المجتمع غير محدود) : $\boldsymbol{\sigma^2_{\bar{X}}} = \sigma^2/n$

<u>ب- في حالة المعاينة بدون إرجاع</u>

1. إيجاد عدد العينات الممكن اختيارها

العينات الممكنة ذات الحجم n = 2 المسحوبة من مجتمع حجمه N=5 هي عبارة عن توفيقة لاختيار مفردتين من بين خمسة أفراد و عددها هو:

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

$$C_5^2 = 10$$

2. متوسط توزيع المعاينة للمتوسط الحسابي $\boldsymbol{E(\bar{X})}$ و مقارنته مع μ

نقوم بإيجاد قيم المتوسط الحسابي الذي يوافق جميع العينات الممكن اختيارها ذات الحجم 2 والتي عددها 10:

X \ X	1	3	5	7	9
1		2	3	4	5
3			4	5	6
5				6	7
7					8
9					

$\bar{X}_i$	2	3	4	5	6	7	8
$P(\bar{X}_i)$	1/10	1/10	2/10	2/10	2/10	1/10	1/10

$$E(\bar{X}) = \sum \bar{X}_i . P(\bar{X}_i) = 2\left(\frac{1}{10}\right) + 3\left(\frac{1}{10}\right) + \cdots + 8\left(\frac{1}{10}\right) = 5$$

المقارنة : $E(\bar{X}) = \mu = \mathbf{5}$

مثلما تم التوصل اليه سابقا في حالة المعاينة بالإرجاع، فانه في حالة المعاينة بدون ارجاع يكون المتوسط الحسابي مقدر غير متحيز لمتوسط المجتمع .

3. تباين توزيعات المعاينة للمتوسط الحسابي $\boldsymbol{\sigma^2_{\bar{X}}}$ ثم مقارنته مع $\boldsymbol{\frac{\sigma^2}{n}\frac{(N-n)}{(N-1)}}$

$$\mathrm{Var}(\bar{X}_i) = \boldsymbol{\sigma^2_{\bar{X}}} = \sum (\bar{X}_i - \mu)^2 P(\bar{X}_i)$$

$(\bar{X}_i - \mu)^2$	9	4	1	0	1	4	9	
$(\bar{X}_i - \mu)^2 P(\bar{X}_i)$	9/10	4/10	2/10	0	2/10	4/10	9/10	

$$\boldsymbol{\sigma^2_{\bar{X}}} == \left(\frac{9}{10}\right) + \left(\frac{4}{10}\right) + \cdots + \left(\frac{9}{10}\right) = 3$$

$$\boldsymbol{\sigma^2_{\bar{X}}} = 3$$

المقارنة بين : $\sigma^2_{\bar{X}}$ و $\frac{\sigma^2}{n}\frac{(N-1)}{(N-1)}$

$$\frac{\sigma^2}{n}\frac{(N-1)}{(N-1)} = \frac{8}{2}\frac{(5-2)}{(5-1)} = 3$$

نستنتج أنه في حالة المعاينة بدون ارجاع:

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}\frac{(N-1)}{(N-1)}$$

التمرين رقم 03

متوسط النقاط لمقياس الرياضيات لـ **300** طالب سنة أولى جامعي يساوي **9.8** ، الانحراف المعياري هو **3.68**

المطلوب :في حالة المعاينة **بدون إرجاع** أوجد احتمال أن يكون المتوسط الحسابي من أجل عينة عشوائية لنقاط **40** طالب:

1- محصور بين **10** و **11.5** ؟ **2-** أقل من **10** ؟

الحل :

لدينا $\boldsymbol{\sigma}$ معلومة و $X \rightsquigarrow$ ؟ وبالتالي وفقا لنظرية النهاية المركزية في حالة العينات الكبيرة $n = 40 > 30$ فان التوزيع الاحتمالي لمتوسط العينة يقترب من التوزيع الطبيعي ونكتب :

$$\overline{\boldsymbol{X}} \rightsquigarrow \mathrm{N}(\mu; \sigma_{\overline{X}}^2)$$

- نعلم أن المجتمع يكون محدود اذا تحقق الشرط التالي : $\frac{n}{N} > 0.05$

من المعطيات السابقة نجد :

$$\frac{n}{N} = \frac{40}{300} = 0.133 > 0.05$$

وبالتالي فان المجتمع يعتبر محدودا.

$$\sigma_{\overline{X}} = \sqrt{\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)} = \sqrt{\frac{3.68^2}{40}\left(\frac{300-40}{300-1}\right)} = \sqrt{0.2944} = 0.5425$$

1. حساب احتمال $\boldsymbol{Pr(10 \leq \overline{X} \leq 11.5)}$

$$\boldsymbol{Pr(10 \leq \overline{X} \leq 11.5) = Pr\left(\frac{10-\mu}{\sigma_{\overline{X}}} \leq Z \leq \frac{11.5-\mu}{\sigma_{\overline{X}}}\right)}$$

$$= \boldsymbol{Pr}\left(\frac{10-9.8}{0.5425} \leq \boldsymbol{Z} \leq \frac{11.5-9.8}{0.5425}\right)$$

$$= Pr(0.3686 \leq Z \leq 3.1336)$$

$$Pr(Z \leq 3.1336) - Pr(Z \leq 0.368)$$

$$= 0.9991 - 0.6405 = 0.3586$$

2. حساب احتمال $\boldsymbol{Pr(\bar{X} \leq 10)}$

$$Pr(\bar{X} \leq 10) = Pr\left(Z \leq \frac{10-9.8}{0.5425}\right)$$

$$= Pr(Z \leq 0.3686) = 0.6405$$

التمرين رقم 04

في احدى الدول يتبع متغير الدخل العائلي X التوزيع الطبيعي بمتوسط حسابي 16000 دج $\mu =$ وانحراف معياري 2000 دج $\sigma =$ ، تم اختيار عينة عشوائية حجمها n=25 ، و باعتبار أن المجتمع الاحصائي غير محدود (المعاينة بالارجاع) :

1. أوجد التوزيع الاحتمالي، المتوسط والانحراف المعياري للمتوسط الحسابي للعينة $\bar{X}$ ؟
2. ما هو احتمال أن يكون متوسط الدخل العائلي في العينة العشوائية :

- أقل من 15000دج
- أكبر من 16400دج ؟
- محصورا بين 15000 دج و16400دج ؟

الحل :

1. **ايجاد التوزيع الاحتمالي، المتوسط والانحراف المعياري للمتوسط الحسابي للعينة $\bar{X}$**

لدينا $\boldsymbol{\sigma}$ معلومة و $X \rightsquigarrow \mathrm{N}(16000; 2000^2)$ وبالتالي فان :

$$\bar{\boldsymbol{X}} \rightsquigarrow \mathrm{N}(\mu; \sigma_{\bar{X}}^2)$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2000}{\sqrt{25}} = 400$$

$$\bar{\boldsymbol{X}} \rightsquigarrow \mathrm{N}(16000; 400^2)$$

2. حساب احتمال $\boldsymbol{Pr\left(\bar{X} \leq 15000\right)}$

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \rightsquigarrow \mathrm{N}(1; 0)$$

$$Pr\left(\bar{X} \leq 15000\right) = Pr\left(Z \leq \frac{1500 - \mu}{\sigma_{\bar{X}}}\right)$$

$$= Pr\left(Z \leq \frac{1500 - 16000}{400}\right)$$

$$= Pr(Z \leq -2.5) = Pr(Z \geq 2.5) = 1 - Pr(Z \leq 2.5)$$

$$= 1 - 0.99379 = 0.0062$$

3. **حساب احتمال $\boldsymbol{Pr\left(\bar{X} > 16400\right)}$**

$$Pr\left(\bar{X} > 16400\right) = Pr\left(Z > \frac{16400 - 16000}{400}\right)$$

$$= Pr(Z > 1) = 1 - Pr(Z \leq 1)$$

$$= 1 - 0.8413 = 0.1587$$

4. حساب احتمال $Pr(15000 \leq \bar{X} \leq 16400)$

$$Pr(15000 \leq \bar{X} \leq 16400)$$

$$Pr\left(\frac{1500 - 16000}{400} \leq Z \leq \frac{16400 - 16000}{400}\right)$$

$$= Pr(-2.5 \leq Z \leq 1)$$

$$Pr(Z \leq 1) - Pr(Z \leq -2.5)$$

$$= 0.8413 - 0.0062 = 0.8351$$

التمرين رقم 05

لتكن لدينا X متغيرة عشوائية تتبع التوزيع الطبيعي في مجتمع احصائي حيث :

$X \rightsquigarrow N(8{,}5\ ;\sigma^2)$

المطلوب :

1. أوجد توزيع المعاينة لمتوسط عينة عشوائية بحجم $n=\ 9$ حيث عناصر العينة هي
$(9\ ;10\ ;9\ ;8\ ;8\ ;7\ ;8\ ;5\ ;8\ )$

2. أحسب $P(7 \leq \bar{X} \leq 9)$ ؟

الحل :

1. ايجاد توزيع المعاينة لمتوسط عينة عشوائية

لدينا $\boldsymbol{\sigma}$ مجهولة و $X \rightsquigarrow N(8{,}5\ ;\sigma^2)$ فان التوزيع الاحتمالي في العينة يؤول الى توزيع ستيوزدنت بدرجة حرية $v=n\text{-}1$ وتعطى احصائية ستيودنت بالصيغة التالية :

$$T = \frac{\bar{X} - E(\bar{X})}{\sigma_{\bar{X}}} \rightsquigarrow \mathrm{t}_{(n-1)}$$

$$T = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \rightsquigarrow \mathrm{t}_{(n-1)}$$

متوسط العينة يساوي :

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

نحسب تباين العينة الغير متحيز S^2 (يعتبر كمقدر غير متحيز لتباين المجتمع σ^2)

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{\sum[(8-8)^2 + (5-8)^2 + \cdots + (9-8)^2]}{9-1} = \frac{16}{8}$$

$$S^2 = 2$$

$$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{2}{9}} = 0.4714$$

$$\bar{\boldsymbol{X}} \rightsquigarrow N(8{,}5\ ;0.4714^2)$$

.2 حساب احتمال $\boldsymbol{Pr(15000 \leq \bar{X} \leq 16400)}$

$$Pr(7 \leq \bar{X} \leq 9) = Pr\left(\frac{7-8.5}{0.4714} \leq T \leq \frac{9-8.5}{0.4714}\right)$$
$$= Pr(-3.1820 \leq T \leq 1.060)$$
$$= Pr(T \leq 1.060) - Pr(T \leq -3.1820)$$

نعلم أن :

$$Pr(T \leq 1.060) = 1 - Pr(T \geq 1.060)$$

ان توزيع ستيودنت هو توزيع متناظر (التوزيع الطبيعي أيضا متناظر) و بالتالي :

$$Pr(T \leq -3.1820) = Pr(T \geq 3.1820)$$

نعوض فيما سبق فنجد :

$$Pr(T \leq 1.060) - Pr(T \leq -3.1820)$$
$$= 1 - Pr(T \geq 1.060) - Pr(T \geq 3.1820)$$
$$= 1 - 0.15 - 0.005 = 0.845$$

اذا : $\boldsymbol{Pr(7 \leq \bar{X} \leq 9) = 0.845}$

التمرين رقم 06

ليكن لدينا مجتمع احصائي $X_1 \rightsquigarrow N(350;\sigma_1^2)$ سحبت منه عينة عشوائية حجمها $n_1 = 20$ بتباين $S_1^2 = 09$ ، و مجتمع ثاني $X_2 \rightsquigarrow N(300;\sigma_2^2)$ مسحوبة منه عينة عشوائية حجمها $n_2 = 15$ بتباين $S_1^2 = 16$

المطلوب :

1. أوجد توزيع المعاينة للفرق بين متوسطي العينتين ؟
2. أحسب الاحتمال $Pr(\bar{X}_1 - \bar{X}_2 \geq 52.5)$ ؟

الحل :

.1 ايجاد توزيع المعاينة للفرق بين متوسطي العينتين

لدينا التباين في المجتمعين الاحصائيين σ_2^2 ,σ_1^2 مجهول و غير متساوي ، توزيع العينية اذا يتبع التوزيع الاحتمالي لستيودنت حيث تعطى صيغة إحصائية **T** لتوزيع ستيودنت بدرجة حرية v كما يلي:

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow t(v)$$

حيث :

$$\sigma^2_{\bar{X}_1-\bar{X}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}=\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}$$

$$\sigma_{\bar{X}_1-\bar{X}_2}=\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}=\sqrt{\frac{9}{20}+\frac{16}{15}}=1.2315$$

تعطى علاقة درجة الحرية بالصيغة التالية :

$$v=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left({S_1^2}/{n_1}\right)^2}{n_1-1}+\frac{\left({S_2^2}/{n_2}\right)^2}{n_2-1}}$$

$$v=\frac{\left(\frac{9}{20}+\frac{16}{15}\right)^2}{\frac{\left({9}/{20}\right)^2}{20-1}+\frac{\left({16}/{15}\right)^2}{15-1}}=25.252\cong 25$$

$$T=\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}\rightsquigarrow t(v)$$

2. حساب الاحتمال $Pr(\bar{X}_1-\bar{X}_2\geq 52.5)$

$$Pr(\bar{X}_1-\bar{X}_2\geq 52.5)=Pr\left(T\geq\frac{52.5-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}\right)$$

$$=Pr\left(T\geq\frac{52.5-(350-300)}{1.2315}\right)$$

$$=Pr(T\geq 2.0300)=0.025$$

عند درجة الحرية $v=25$ احصائية ستيودنت الأقرب الى القيمة 2.03 هي القيمة 2.0595 وتوافق الاحتمال $p=0.025$

التمرين رقم 07

سحبت عينتين عشوائيتين من شركتين مختلفتين لإنتاج العتاد الزراعي $n_1 = 36$ ، $n_2 = 49$ ، الأجور المدفوعة من قبل الشركتين تتبع التوزيع الطبيعي في مجتمعي الدراسة، المتوسط الحسابي للأجور في الشركة الأولى $\mu_1 = 230$ دينار وبانحراف معياري $\sigma_1 = 36$ دينار ، المتوسط الحسابي للأجور في الشركة الثانية $\mu_2 = 180$ دينار، وبانحراف معياري $\sigma_2 = 40$.

المطلوب:أحسب احتمال أن يكون الفرق بين متوسطي الأجور للعينتين أكبر أو يساوي 60 ؟

الحل :

لدينا التباين في المجتمعين σ_2^2, σ_1^2 معلوم وحجم العينتين كبير $n_1, n_2 > 30$ ، وفقا لنظرية النهاية المركزية فان الفرق بين متوسطي الأجور للعينتين يتبع التوزيع الاحتمالي الطبيعي حيث تعطى صيغة إحصائية التوزيع الطبيعي كما يلي:

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow N(0;1)$$

Z : متغيرة معيارية تتبع التوزيع الطبيعي المعياري بمتوسط 0 وتباين يساوي 1

$$(\bar{X}_1 - \bar{X}_2) \rightsquigarrow N\left(\mu_1 - \mu_2; \sigma^2_{\bar{X}_1 - \bar{X}_2}\right)$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{36^2}{36} + \frac{40^2}{49}} = 8.2857$$

$$Pr(\bar{X}_1 - \bar{X}_2 \geq 60) = Pr\left(Z \geq \frac{60 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right)$$

$$= Pr\left(Z \geq \frac{60 - (230 - 180)}{8.2857}\right)$$

$$Pr(Z \geq 1.2086) = 1 - Pr(Z \leq 1.2086) = 1 - 0.8849 = 0.1151$$

$$Pr(\bar{X}_1 - \bar{X}_2 \geq 60) = 0.1151$$

التمرين رقم 08

مصنعين A و B لإنتاج الأنابيب البلاستيكية (وحدة القياس بالسنتيمتر)، متوسط طول الأنابيب في المصنع A يساوي $\mu_1 = 62.00$ تم سحب عينة عشوائية أولى من هذا المصنع $n_1 = 5$ بتباين $S_1^2 = 0.34$ ، أما المصنع B فقد بلغ متوسط طول الأنابيب فيه $\mu_2 = 61.45$ حيث تم سحب

عينة عشوائية من هذا المصنع $n_2 = 4$ بتباين $S_2^2 = 0.213$ ، نفترض أن قياس الأنابيب في كلا المصنعين يتبع التوزيع الطبيعي و بتباين متساوي $\sigma_1^2 = \sigma_2^2$.

المطلوب : **أحسب الاحتمال** $\boldsymbol{P(\bar{X}_1 - \bar{X}_2 \geq 0.95)}$ ؟

الحل :

قبل حساب الاحتمال يجب معرفة طبيعة توزيع المعاينة للفرق بين متوسطي العينتين (طبيعة التوزيع الاحتمالي)

لدينا: $n_1 = 5$, $S_1^2 = 0.34$ $\quad n_2 = 0.213$, $S_2^2 = 4$

- التباين في المجتمعين متساوي ومجهول $\boldsymbol{\sigma_1^2 = \sigma_2^2}$

- في المصنع **A** $\boldsymbol{X \rightsquigarrow N(62, \sigma_1^2)}$ في المصنع **B** $\boldsymbol{X \rightsquigarrow N(61.45, \sigma_2^2)}$

اذا فالتوزيع الاحتمالي في العينتين يتبع توزيع ستيودنت بدرجة حرية $v = n_1 + n_2 - 2$ ، و تعطى صيغة الإحصائية **T** التي تخضع لتوزيع ستيودنت على النحو التالي:

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow t(n_1 + n_2 - 2)$$

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightsquigarrow t(7)$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(5-1)0.34 + (4-1)0.213}{5 + 4 - 2} = 0.2855$$

$$Pr(\bar{X}_1 - \bar{X}_2) \geq 0.95 = Pr\left(T \geq \frac{0.95 - (\mu_1 - \mu_2)}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}\right)$$

$$= Pr\left(T \geq \frac{0.95 - (62 - 61.45)}{\sqrt{0.2855 \left(\frac{1}{5} + \frac{1}{4}\right) =}}\right)$$

$$= Pr\left(T \geq \frac{0.4}{0.3584}\right) = Pr(T \geq 1.118) = 0.15$$

التمرين رقم 09

بلغ وزن أحد المواد الغذائية بأحد المصانع في المتوسط 700 كغ ، سحبت عينة عشوائية من هذا المصنع من 40 علبة حيث بلغ المتوسط الحسابي 740 كغ وبانحراف معياري 40 كغ، مصنع ثاني بلغ فيه متوسط وزن نفس المنتوج 500 كغ سحبت منه عينة عشوائية من 35 علبة بمتوسط حسابي 480 كغ وبانحراف معياري 20 كغ.

المطلوب:

1. أوجد توزيع المعاينة للفرق بين متوسطي العينتين (التوزيع الاحتمالي، التوقع الرياضي، الانحراف المعياري للفرق بين متوسطي العينتين) ؟
2. أحسب الاحتمال $Pr(180 \leq \overline{X}_1 - \overline{X}_2 \leq 210)$؟

الحل :

1. ايجاد توزيع المعاينة للفرق بين متوسطي العينتين :

لدينا التباين في كلا المجتمعين مجهول، أما التباين في العينتين معلوم : $S_2^2 = 20^2,\ \ S_1^2 = 40^2$

$n_1 = 40 > 30$، $n_2 = 35$ اذا حسب نظرية النهاية المركزية فان توزيع المعاينة للفرق بين المتوسطين يتبع التوزيع الطبيعي :

$$(\overline{X}_1 - \overline{X}_2) \rightsquigarrow N(E(\overline{X}_1 - \overline{X}_2), \sigma^2_{\bar{X}_1 - \bar{X}_2})$$
$$(\overline{X}_1 - \overline{X}_2) \rightsquigarrow N(\mu_1 - \mu_2, \sigma^2_{\bar{X}_1 - \bar{X}_2})$$
$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2 = 700 - 500 = 200$$
$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = \frac{40^2}{40} + \frac{20^2}{35} = 51.4285$$
$$(\overline{X}_1 - \overline{X}_2) \rightsquigarrow N(200;\ 51.428)$$
$$\sigma_{\bar{X}_1 - \bar{X}_2} = 7.1713$$

2. حساب الاحتمال $Pr(180 \leq \overline{X}_1 - \overline{X}_2 \leq 210)$؟

$$Z = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow N(0,1)$$
$$Pr(180 \leq \overline{X}_1 - \overline{X}_2 \leq 210) = Pr\left(\frac{180 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \leq Z \leq \frac{210 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right)$$
$$= Pr\left(\frac{180 - 200}{7.1713} \leq Z \leq \frac{210 - 200}{7.1713}\right)$$
$$= Pr(-2.7888 \leq Z \leq 1.3944)$$

$$= Pr(Z \leq 1.3944) - Pr(Z \leq -2.7888)$$

$$= Pr(Z \leq 1.3944) - [1 - Pr(Z \leq 2.7888)]$$

$$= 0.9177 - [1 - 0.9972] = 0.9149$$

اذا : $Pr(180 \leq \overline{X}_1 - \overline{X}_2 \leq 210) = 0.9149$

التمرين رقم 10

من خلال دراسات سابقة معلوم أن أجور المعلمين لقطاع التربية في المدارس الخاصة والمدارس العمومية يخضع للتوزيع الطبيعي بمتوسط أجور للمعلمين في المدارس الخاصة يساوي 36000دج وبمتوسط أجور في المدارس العمومية يساوي 34550 دج، و بغرض اجراء دراسة بحثية حول القدرة الشرائية ومدى وجود فروق بين أجور المعلمين في القطاعين العام والخاص فانه في تم اختيار عينة عشوائية من المجتمع الأول من المدارس الخاصة من 9 معلمين بانحراف معياري يساوي 600 دج ، و اختيار عينة عشوائية ثانية من المجتمع الثاني من المدارس العمومية من 16 معلم بانحراف معياري يساوي 800 دج .

المطلوب: باعتبار أن مجتمعي المدارس العمومية والمدارس الخاصة مستقلين و بتباين متساوي $\sigma_1^2 = \sigma_2^2$

1. احسب احتمال أن يكون متوسط أجور المدارس العمومية في العينة العشوائية المختارة أقل أو يساوي 34290 دج ؟
2. احسب احتمال أن يكون الفرق بين متوسط الأجور في عينتي المدارس الخاصة و المدارس العمومية أكبر أو يساوي 1975 دج بمعنى: $P(\bar{X}_1 - \bar{X}_2) \geq 1975$ ؟

<u>الحل :</u>

لدينا :

$S_2 = 800$, $n_2 = 16$ $S_1 = 600$, $n_1 = 9$

التباين في المجتمعين متساوي ومجهول $\boldsymbol{\sigma_1^2 = \sigma_2^2}$

في مجتمع المدارس الخاصة

$$X \rightsquigarrow N(36000, \sigma_1^2)$$

في مجتمع المدارس العمومية

$$X \rightsquigarrow N(34550, \sigma_2^2)$$

⇐ تعطى صيغة الإحصائية T التي تخضع لتوزيع ستيودنت على النحو التالي:

$$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\hat{\sigma}^2/\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \rightsquigarrow t(n_1 + n_2 - 2)$$

$$T = \frac{\overline{X}_2 - \mu_2}{\sigma_{\overline{X}_2}} \rightsquigarrow t(n_2 - 1)$$

✓ حساب الاحتمال $\boldsymbol{P(\bar{X}_2 \leq 34290)}$

$$\sigma_{\bar{X}_2} = \sigma_2/\sqrt{n} = S_2/\sqrt{n_2} = {}^{800}/\sqrt{16} = 200$$

$$Pr(\bar{X}_2 \leq 34290) = Pr\left(T \leq \frac{34290 - \mu_2}{\sigma_{\bar{X}_2}}\right) = Pr\left(T \leq \frac{34290 - 34550}{200}\right)$$
$$= Pr(T \leq -1.3) = Pr(T \geq 1.3) = 0.1$$

✓ حساب الاحتمال $\boldsymbol{Pr(\bar{X}_1 - \bar{X}_2) \geq 1975}$

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \rightsquigarrow t(n_1 + n_2 - 2)$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(9 - 1)600^2 + (16 - 1)600^2}{9 + 16 - 2}$$
$$= 542608.6956$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{542608.6956\left(\frac{1}{9} + \frac{1}{16}\right)} = \sqrt{94202.8958} = 306.9249$$

$$Pr(\bar{X}_1 - \bar{X}_2) \geq 1975 = Pr\left(T \geq \frac{1975 - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right)$$

$$Pr\left(T \geq \frac{\bar{X}_1 - \bar{X}_2 - (36000 - 34550)}{306.9249}\right)$$

$$= Pr\left(T \geq \frac{1975 - 1450}{306.9249}\right) = Pr(T \geq 1.7104) = 0.05$$
$$Pr(\bar{X}_1 - \bar{X}_2) \geq 1975 = 0.05$$

التمرين رقم 11

لتكن لدينا متغيرة عشوائية X تتبع توزيع ذو الحدين حيث $X \rightsquigarrow B(4; 0.2)$

المطلوب : أوجد ما يلي :

1. صيغة التوزيع الاحتمالي للمتغيرة العشوائية X ؟
2. $P(x = 2)$ ؟
3. $P(x \geq 1)$ ؟

الحل :

لدينا : p=0.2 n=4

1. ايجاد التوزيع الاحتمالي للمتغيرة العشوائية X $P(X)$

$$P(X) = c_n^X p^X (1-p)^{n-X}\ (X) = c_4^X (0.2)^X (0.8)^{4-X}\ \ ;\ \ x = 0,1,2,4$$

2. ايجاد $P(X=2)$

$$P(2) = c_4^2 (0.2)^2 (0.8)^{4-2} \ = 0.1536$$

3. ايجاد $P(X \geq 1)$

$$P(X \geq 1) = P(1) + P(2) + P(3) + P(4)$$

أو بصيغة أسهل كما يلي :

$$P(X \geq 1) = 1 - P(X=0) = 1 - c_4^0 (0.2)^0 (0.8)^{4-0} \ = 0.5904$$

التمرين رقم 12

بافتراض القاء قطعة نقود 15 مرة ، احتمال ظهور الصورة يساوي 0.4 ، و ليكن المتغير العشوائي X الذي يمثل عدد الصور التي يمكن أن تظهر.

✓ **المطلوب**: باستخدام علاقة توزيع ذو الحدين ثم باستخدام التقريب بالتوزيع الطبيعي أحسب الاحتمال $P(X=4)$ ؟

الحل :

p=0.4. ; n =15 $X \rightsquigarrow B(15;0.4)$

✓ **باستخدام علاقة توزيع ذو الحدين :**

-حساب الاحتمال: $P(X=4)$

$$P(X) = \ c_X^{15} (0.4)^X (0.6)^{15-X} \quad , \quad X = 0,1,2, \dots 15$$

$$\boldsymbol{P(X=4)} = c_4^{15} (0.4)^4 (0.6)^{15-4} = \frac{15!}{(15-4)!\,4!} (0.4)^4 (0.6)^{11} = 0.1267$$

✓ باستخدام **تقريب توزيع ذو الحدين الى التوزيع الطبيعي فان :**

$E(X) = np = (15)(0.4) = 6$ و $\sigma^2 = npq = (15)(0.4)(0.6) = 3.6 \ ,\sigma =$ 1.8973

$$P(X=4) \approx P(4-0.5 \leq X \leq 4+0.5)$$
$$P\left(\frac{(4-0.5)-np}{\sigma} \leq Z \leq \frac{(4+0.5)-np}{\sigma}\right)$$

$$P\left(\frac{3.5-6}{1.897} \leq Z \leq \frac{4.5-6}{1.897}\right)$$

$$P(-1.3176 \leq Z \leq -0.7905)$$

$$P(Z \leq -0.790) - P(Z \leq -1.318)$$

=(1−0.78524)−(1−0.90490)

=0.11966

التمرين رقم 13

في مصنع معلوم أن 90 % من القطع المنتجة سليمة ، فاذا تم اختيار عشوائيا 03 قطع منتجة من هذا المصنع.

المطلوب : فأحسب احتمال أن تكون في العينة المختارة على الأكثر قطعة سليمة واحدة ؟

الحل :

حساب احتمال أن تكون في العينة المختارة على الأكثر قطعة سليمة واحدة $Pr(X \leq 1)$:

لدينا :

P =0.9 : نسبة القطع السليمة في المصنع (نسبة النجاح في المجتمع)

n=3 : عدد القطع المختارة عشوائيا من المصنع (حجم العينة أو عدد المحاولات)

X : تمثل عدد القطع السليمة في العينة المختارة (عدد النجاحات) حيث $\boldsymbol{X = (0, 1, 2, 3)}$

احتمال المتغير العشوائي X الذي يتبع توزيع ذو الحدين $X \rightsquigarrow B(n, \mathrm{p})$ يعطى بالعلاقة التالية :

$$\boldsymbol{P(X) = c_X^n p^X (1-p)^{n-X} = c_X^3 (0.9)^X (0.1)^{3-X}}$$

$$Pr(X \leq 1) = Pr(X = 0) + Pr(X = 1)\ c_0^3 (0.9)^0 (0.1)^{3-0} + c_1^3 (0.9)^1 (0.1)^{3-1}$$

$$= 0.001 + 0.027 = 0.028$$

التمرين رقم 14

اذا علمت أن 35% من حوادث المرور في مدينة ما خلال سنة معينة ناتجة عن الافراط في السرعة لسائقي المركبات، فاذا تم اختيار عينة عشوائية من 100 حادث مرور مسجل بأحد المراكز المختصة بالمدينة.

المطلوب : أوجد:

1. أوجد التوزيع الاحتمالي، التوقع الرياضي و الانحراف المعياري لنسبة الحوادث الناتجة عن السرعة المفرطة في العينة ؟

2. احتمال أن تكون في العينة المختارة نسبة 45 % من هذه الحوادث أو أقل نتيجة للسرعة المفرطة $P(\hat{p} \leq 0.45)$ ؟

3. احتمال أن تكون في العينة المختارة نسبة ما بين 30% و 40 % من هذه الحوادث نتيجة للسرعة المفرطة $P(0.3 \leq \hat{p} \leq 0.4)$ ؟

<u>الحل :</u>

1. **التوزيع الاحتمالي، التوقع الرياضي والتباين لنسبة العبوات الكبيرة الحجم في العينة $\hat{\boldsymbol{p}}$ ؟**

بما أن حجم العينة كبير فانه وفقا لنظرية النهاية المركزية بعد تحقق الشرطين :

np=35>5 , n(1−p)=65>5

فان التوزيع الاحتمالي للمتغيرة $\hat{\boldsymbol{p}}$ يقترب من التوزيع الطبيعي ونكتب: $\hat{\boldsymbol{p}} \rightsquigarrow \boldsymbol{N}(\mathbf{p}; \sigma^2_{\hat{\boldsymbol{p}}})$

$$\sigma^2_{\hat{\boldsymbol{p}}} = \frac{pq}{n} = \frac{(0.35)(0.65)}{100} = 0.002275$$
$$\sigma_{\hat{\boldsymbol{p}}} = 0.04769$$

2. **حساب احتمال $\boldsymbol{Pr(\hat{p} \leq 0.45)}$**

$$Pr(\hat{p} \leq 0.45) == P\left(\mathrm{Z} \leq \frac{0.45 - p}{\sigma_{\hat{\boldsymbol{p}}}}\right)$$
$$Pr\left(\mathrm{Z} \leq \frac{0.45 - 0.35}{0.0476}\right)$$
$$= Pr(\mathrm{Z} \leq 2.1) = 0.9821$$

3. **حساب احتمال $\boldsymbol{P(0.3 \leq \hat{p} \leq 0.4)}$**

$$P(0.3 \leq \hat{\boldsymbol{p}} \leq 0.4)$$
$$= Pr\left(\frac{0.3 - p}{\sigma_{\hat{\boldsymbol{p}}}} \leq \mathrm{Z} \leq \frac{0.4 - p}{\sigma_{\hat{\boldsymbol{p}}}}\right)$$
$$= Pr\left(\frac{0.3 - 0.35}{0.0476} \leq \mathrm{Z} \leq \frac{0.4 - 0.35}{0.0476}\right)$$
$$= Pr(-1.05 \leq \mathrm{Z} \leq 1.05)$$
$$= Pr(\mathrm{Z} \leq 1.05) - P(\mathrm{Z} \leq -1.05)$$
$$= 0.8531 - (1 - 0.8531) = 0.7062$$

التمرين رقم 15

تريد مؤسسة انتاجية تقدير نسبة الأجهزة الالكترونية الفاسدة خلال فترة زمنية معينة ضمن اجمالي 800 وحدة منتجة ، ولتحقيق هذا الغرض تم اختيار عينة عشوائية n=60 وحدة منتجة من بين اجمالي الانتاج ومن المعلوم أن نسبة القطع الفاسدة في المؤسسة تساوي 10% .

المطلوب :

- أحسب احتمال أن تكون نسبة الأجهزة الفاسدة في عينة عشوائية مختارة أكبر من 20% ؟

الحل :

حساب احتمال أن تكون نسبة الأجهزة الفاسدة في عينة عشوائية مختارة أكبر من 20%

بما أن حجم العينة كبير فانه وفقا لنظرية النهاية المركزية بعد تحقق الشرطين :

np=35>5 , n(1−p)=65>5

فان التوزيع الاحتمالي للمتغيرة $\hat{\boldsymbol{p}}$ يقترب من التوزيع الطبيعي ونكتب: $\hat{\boldsymbol{p}} \rightsquigarrow \boldsymbol{N}(\mathbf{p};\sigma_{\hat{\boldsymbol{p}}})$

يكون المجتمع الاحصائي محدود اذا تحقق الشرط التالي : $\frac{n}{N} > 0.05$

من المعطيات السابقة نجد :

$$\frac{n}{N} = \frac{60}{800} = 0.1 > 0.05$$

وبالتالي فان المجتمع الاحصائي محدود.

$$\sigma_{\hat{\boldsymbol{p}}} = \sqrt{\frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)} = \sqrt{\frac{(0.1)(1-0.1)}{60}\left(\frac{800-60}{800-1}\right)} = \sqrt{0.0013}$$
$$= 0.0387$$

$$\sigma_{\hat{\boldsymbol{p}}} = 0.0360$$

$$Pr(\hat{p} > 0.2) = Pr\left(\mathrm{Z} \leq \frac{0.45-p}{\sigma_{\hat{\boldsymbol{p}}}}\right)$$

$$Pr\left(\mathrm{Z} > \frac{0.2-0.1}{0.036}\right)$$

$$= Pr(\mathrm{Z} > 2.77) = 1 - P(\mathrm{Z} \leq 2.77)$$

$$= 1 - 0.9972 = 0.0028$$

التمرين رقم 16

اذا كانت نسبة العمال الراضين عن ظروف العمل في مصنع A ينتج منتوج الحليب هو 80% وكانت نسبة العمال الراضين عن ظروف العمل في مصنع آخر B ينتج نفس المنتوج هو 75% ، لأغراض بحثية قام باحث احصائي باختيار عينة عشوائية من 70 عامل من المصنع A و اختيار عينة عشوائية أخرى من المصنع B من 35 عامل.

المطلوب :

1. أوجد أحتمال أن تزيد أو تساوي نسبة العمال الراضيين عن الظروف المهنية للعمل عن 75 % في العينة المختارة من المصنع A ؟
2. أوجد عدد العينات التي تكون فيها نسبة العمال الراضيين عن الظروف المهنية للعمل تزيد أو تساوي نسبة 75 % مع العلم أن اجمالي عدد العينات الممكن اختيارها من المصنع A يساوي 351 ؟
3. أوجد احتمال أن يكون الفرق بين نسبتي العمال الراضيين عن الظروف المهنية للعمل في العينة المختارة من المصنع A و العينة المختارة من المصنع B أكبر أو يساوي 10% ؟

الحل:

لدينا : $p_1 = 0.8$ $\quad q_1 = 0.2$ $\quad n_1 = 70$ $\quad p_2 = 0.75$

$q_1 = 0.25$ $\quad n_2 = 35$

بما أن : $n_2p_2 = 26.25 > 5$ $\quad n_1p_1 = 56 > 5\ ,\ n_1\ q_1 = 14 > 5$

$n_2\ q_2 = 8.75 > 5$;

وكان حجم العينتين كبير فوفقا لنظرية النهاية المركزية فان توزيع ثنائي الحدين يقترب من التوزيع الطبيعي وبالتالي:

$$\hat{p}_1 \rightsquigarrow N(p_1; \sigma^2_{\hat{p}_1})$$

$$Z = \frac{\hat{p}_1 - p_1}{\sigma_{\hat{p}_1}} \rightsquigarrow N(1;0)$$

$$(\hat{p}_1 - \hat{p}_2) \rightsquigarrow N(p_1 - p_2; \sigma^2_{(\hat{p}_1 - \hat{p}_2)})$$

$$Z = \frac{(\hat{\boldsymbol{p}}_1 - \hat{\boldsymbol{p}}_2) - (\boldsymbol{p}_1 - \boldsymbol{p}_2)}{\boldsymbol{\sigma}_{\hat{\boldsymbol{p}}_1 - \hat{\boldsymbol{p}}_2}} \rightsquigarrow N(1;0)$$

1.ايجاد احتمال أن تزيد أو تساوي نسبة العمال الراضيين عن الظروف المهنية للعمل عن 75 % في العينة المختارة من المصنع A :

$$\sigma_{\hat{p}_1} = \sqrt{p_1q_1/n_1}\ \sqrt{(0.8)(0.2)/70} = \sqrt{0.0022} = 0.04690$$

$$Pr(\hat{p}_1 \geq 0.75) = Pr\left(\text{Z} \geq \frac{0.75 - p_1}{\sigma_{\hat{p}_1}}\right) = Pr\left(\text{Z} \geq \frac{0.75 - 0.8}{0.04690}\right)$$

$$= Pr(\text{Z} \geq -1.0660) = Pr(\text{Z} \leq 1.0660) = 0.8554$$

2. ايجاد عدد العينات التي تكون فيها نسبة العمال الراضيين عن الظروف المهنية للعمل تزيد أو تساوي نسبة 75 % مع العلم أن اجمالي عدد العينات الممكن اختيارها من المصنع A يساوي 351:

$$(0.8554)(0.8554) = 300.2454 \cong 300$$

اذا عدد العينات الي تحقق الاحتمال السابق تساوي 300 عينة

3. ايجاد احتمال أن يكون الفرق بين نسبتي العمال الراضيين عن الظروف المهنية للعمل في العينة المختارة من المصنع A و العينة المختارة من المصنع B أكبر أو يساوي 10% :

$$\sigma^2_{(\hat{p}_1-\hat{p}_2)} = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} = \sqrt{\frac{(0.8)(0.2)}{70} + \frac{(0.75)(0.25)}{35}} = \sqrt{0.0022 + 0.0053} = \sqrt{0.0075}$$

$$= 0.0866$$

$$Pr(\hat{p}_1 - \hat{p}_2 \geq 0.1) = Pr\left[\text{Z} \geq \left(\frac{0.1 - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}}\right)\right] = Pr\left[\text{Z} \geq \left(\frac{0.1 - (0.8 - 0.75)}{0.0866}\right)\right]$$

$$= Pr(\text{Z} \geq 0.577) = 1 - Pr(\text{Z} \leq 0.577) = 1 - 7156 = 0.2844$$

التمرين رقم 17

بلغت نسبة المتوفقين في مؤسستين جامعيتين A وB 30% و 20% على التوالي، تم سحب عينتين عشوائيتين من المؤسستين السابقتين بحجم $\boldsymbol{n_1} = 100$ و $\boldsymbol{n_2} = 200$

المطلوب :

1. استنتج التوزيع الاحتمالي للفرق بين نسبتي العينتين $\hat{p}_1 - \hat{p}_2$ ؟
2. أحسب التوقع الرياضي$E(\hat{p}_1 - \hat{p}_2)$ و كذلك الانحراف المعياري $\sigma^2_{(\hat{p}_1-\hat{p}_2)}$ للفرق بين نسبتي التفوق في العينتين؟
3. أحسب احتمال أن يكون الفرق بين نسبتي التفوق في العينتين أكبر أو يساوي 6% بمعنى : $P(\hat{p}_1 - \hat{p}_2 \geq 0.06)$؟

الحل

1. استنتاج التوزيع الاحتمالي للفرق بين نسبتي العينتين

بما أن : $n_1p_1 = 30 > 5 \quad n_2p_2 = 40 > 5 \quad , \; n_1\, q_1 = 70 > 5 \quad n_2$
$q_2 = 160 > 5$

وكان حجم العينتين كبير فوفقا لنظرية النهاية المركزية فان توزيع ثنائي الحدين للمتغيرة العشوائيةx

يقترب من التوزيع الطبيعي: $\boldsymbol{(\hat{p}_1 - \hat{p}_2) \rightsquigarrow N(p_1 - p_2; \sigma_{\hat{p}_1 - \hat{p}_2})}$

حيث : $\mathrm{Z} = \frac{\boldsymbol{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}}{\boldsymbol{\sigma_{\hat{p}_1 - \hat{p}_2}}} \rightsquigarrow N(1; 0)$

2. حساب $\boldsymbol{E(\hat{p}_1 - \hat{p}_2)}$ ، $\boldsymbol{\sigma^2_{(\hat{p}_1 - \hat{p}_2)}}$

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2 = 0.3 - 0.2 = 0.1$$
$$\sigma^2_{(\hat{p}_1 - \hat{p}_2)} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2}$$
$$= \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}$$
$$\frac{(0.3)(0.7)}{100} + \frac{(0.2)(0.8)}{200} = 0.0029$$
$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{0.0029} = 0.0538$$

3. حساب الاحتمال $P(\hat{p}_1 - \hat{p}_2 \geq 0.06)$

نفترض أن حجم العينتين كبير وبالتالي نهمل معامل التصحيح لانتقال من توزيع ثنائي الحدين الى التوزيع الطبيعي

$$P(\hat{p}_1 - \hat{p}_2 \geq 0.06)$$
$$= P\left[\mathrm{Z} \geq \left(\frac{0.06 - (p_1 - p_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}}\right)\right]$$
$$= P\left[\mathrm{Z} \geq \left(\frac{0.06 - 0.1}{0.0538}\right)\right]$$
$$= P[\mathrm{Z} \geq -0.7434] = P[\mathrm{Z} \leq 0.7434] = 0.77035$$

التمرين رقم 18

لتكن لدينا χ_v^2 متغيرة عشوائية تتبع توزيع مربع كاي بدرجة حرية v، فمن خلال الجدول الاحصائي لتوزيع مربع كاي، أكمل الجدول الموالي :

$Pr(\chi_{v=5}^2 \geq \cdots) = 0.01$	$Pr(\chi_{20}^2 \geq \cdots) = 0.05$	$Pr(\chi_{20}^2 < 12.443) = \cdots$
$Pr(\chi_{v=28}^2 \geq \cdots) = 0.90$	$Pr(\chi_{20}^2 < \cdots) = 0.05$	

الحل :

من خلال الجدول الاحصائي للتوزيع الاحتمالي لمربع كاي نجد :

$$Pr(\chi_{20}^2 < 12.443) = 1 - Pr(\chi^2 \geq 12.443) = 1 - 0.9 = 0.10$$

$$Pr(\chi_{20}^2 \geq 31.4104) = 0.05$$

$$Pr(\chi_{v=5}^2 \geq \ 15.0862\ ) = 0.01$$

$$Pr(\chi_{20}^2 < \cdots) = 0.051 - Pr(\chi_{20}^2 \geq \cdots.) = 0.05Pr(\chi_{20}^2 \geq \cdots) = 0.95$$

$$Pr(\chi_{20}^2 \geq 10.8508) = 0.95$$

$$Pr(\chi_{v=28}^2 \geq 18.9392) = 0.90$$

التمرين رقم 19

يرغب باحث في شركة لإنتاج علب العصير في معرفة التباين (التغير) في الكمية المعبأة في علب بالعصير، ولتحقيق هذا الغرض تم اختيار عينة عشوائية بسيطة **n=25** علبة عصير من مخرجات انتاج الشركة في يوم معين، من خلال دراسات سابقة معلوم أن كمية التعبئة **X** التي يتم تفريغها في العلب تتبع التوزيع الطبيعي بمتوسط μ وتباين $\sigma^2 = 2$

المطلوب أوجد قيمة b حيث : $Pr(S^2 \geq \mathrm{b}) = 0.05$ حيث : $s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$

الحل:

نعلم أن :

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

وبالتالي :

$$Pr(S^2 \geq \mathrm{b}) = Pr\left(\frac{(n-1)S^2}{\sigma^2} \geq \frac{\mathrm{b}(n-1)}{\sigma^2}\right)$$

$$= Pr\left({\chi^2}_{(n-1)} \geq \frac{24\mathrm{b}}{2}\right) = Pr\left({\chi^2}_{(n-1)} \geq 12\mathrm{b}\right)$$

$$= Pr\left({\chi^2}_{(24)} \geq 12\mathrm{b}\right) = 0.05$$

من الجدول الاحصائي لتوزيع مربع كاي عند درجة الحرية v=24 والاحتمال p=0.05 نجد :

$$Pr\left({\chi^2}_{(24)} \geq 36.42\right) = 0.05$$
$$12\text{b} = 36.42 \implies \text{b} = 3.03$$

التمرين رقم 20

في مصنع لانتاج مواد التنظيف معلوم أن كمية التعبئة **X** التي يتم تفريغها في العبوات البلاستيكية تتبع التوزيع الطبيعي بمتوسط μ وتباين $\boldsymbol{\sigma^2 = 0.8}$ (الوحدة المليلتر)، تم اختيار عينة عشوائية بسيطة **n=10** عبوة تنظيف من مخرجات الانتاج في يوم معين .

المطلوب :

- أوجد قيم b_{2_1} و b_2 التي تحقق مجال الاحتمال التالي لتباين العينة الغير متحيز S^2 :

$$\mathbf{Pr(b_1 \leq \boldsymbol{S} \leq b_2) \leq 0.9} \quad ؟$$

الحل:

من المعلوم أن : $\frac{(n-1)S^2}{\sigma^2} \sim {\chi^2}_{(n-1)}$

وبالتالي :

$$\Pr(\text{b}_1 \leq S \leq \text{b}_2) \leq 0.9$$

حد المعنوية : $1 - \alpha = 0.9 \implies \alpha = 0.1$

$$\Pr\left({\chi^2}_{(n-1),1-\frac{\alpha}{2}} \leq \chi^2 \leq {\chi^2}_{(n-1),\frac{\alpha}{2}}\right) \leq 0.9 \quad ....(1)$$
$$\Pr\left(\frac{(n-1)S^2}{\sigma^2}\text{b}_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq \frac{(n-1)S^2}{\sigma^2}\text{b}_2\right) \leq 0.9......(2)$$

من خل الجدول الاحصائي لمربع كاي عند درجة الحرية $v = (n-1) = 9$ نبحث عن قيمة احصائيتي مربع كاي على اليمين وعلى اليسار حيث أن توزيع مربع كاي غيرمتناظر (على خلاف كل من التوزيع الطبيعي وتوزيع ستيودنت المتناظرين) فنجد :

$${\chi^2}_{(n-1),1-\frac{\alpha}{2}} = {\chi^2}_{9;0.95} = 3.33$$
$${\chi^2}_{(n-1),\frac{\alpha}{2}} = {\chi^2}_{9;0.05} = 16.92$$
$$\Pr\left(\frac{(n-1)S^2}{\sigma^2}\text{b}_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq \frac{(n-1)S^2}{\sigma^2}\text{b}_2\right) \leq 0.9$$

بمطابقة (1) و (2) نجد :

$$\frac{(n-1)S^2}{\sigma^2}\mathrm{b}_1 = \chi^2{}_{(n-1),1-\frac{\alpha}{2}}$$

$$\frac{(10-1)\mathrm{b}_1}{0.8} = 11.25\mathrm{b}_1 = 3.33$$

$$\mathbf{b_1 = 0.296}$$

$$\frac{(n-1)S^2}{\sigma^2}\mathrm{b}_1 = \chi^2{}_{(n-1),1-\frac{\alpha}{2}}$$

$$\frac{(10-1)\mathrm{b}_2}{0.8} = 11.25\mathrm{b}_2 = 16.92$$

$$\mathbf{b_2 = 1.504}$$

اذا :

$$\Pr(\mathrm{b}_1 \le S \le \mathrm{b}_2) \le 0.9$$

$$= \Pr(0.296 \le S \le 1.504) \le 0.9$$

الفصل الثاني: نظرية التقدير

تمهيد

يتمثل أحد الجوانب المهمة للاستدلال الإحصائي في الحصول على تقديرات موثوقة لخصائص المجتمع الاحصائي من عينة مأخوذة من هذا المجتمع، إنها مشكلة قرار تتعلق بمعلمات المجتمع مثل المتوسط الحسابيμ ، التباين أو الانحراف المعياري و النسبة p لخاصية قابلة للقياس، نظرا لأن العينة لا يمكن إلا أن تعطي معلومات جزئية عن المجتمع، فإن التقديرات التي يتم الحصول عليها بهذه الطريقة سينتج عنها حتما أخطاء يجب تقليلها إلى أدنى حد ممكن وبالتالي فان تقدير المعلمة يعني إعطاء قيمة تقريبية لهذه المعلمة وهذا انطلاقا من النتائج التي يتم الحصول عليها من عينة عشوائية مأخوذة من المجتمع.

يعرف **التقدير الاحصائي** على أنه عملية استنتاج أو تقدير معلمة من معالم المجتمع بالاحصائية المناظرة لها في العينة مثلا تقدير متوسط المجتمع μ بمتوسط العينة $\bar{X}$. نظرية المعاينة تسمح بالحصول على معلومات حول مجتمع ما من خلال معطيات مأخوذة من العينة عشوائية، يمكن تقدير معلمة هذا المجتمع بطريقتين هما التقدير النقطي والتقدير بمجال ثقة (فترة ثقة)

1/ *التقدير النقطي*

1.1/ *ماهية التقدير النقطي*

يتضمن تقدير معلمة من معالم المجتمع (المتوسط، التباين، النسبة....) بنقطة واحدة أو قيمة واحدة محددة، فمثلا نحسب من بيانات العينة قيمة المتوسط الحسابي $\bar{X}$ كتقدير نقطي لمتوسط المجتمع ، كذلك تباين العينة S^2 هو مقدر نقطي لتباين المجتمع σ^2 ، وبالمثل بالنسبة للنسبة p فهي مقدر نقطي للنسبة في المجتمع p .

2.1/ *الخصائص المرغوبة للمقدرات*

حسب نوع طريقة التقدير المستخدمة يمكن أن يكون هناك عدة احصائيات (مقدرات) في العينة الاحصائية لنفس معلمة المجتمع مثل المتوسط، المنوال، السؤال هو أي من هذه الاحصائيات أو المقدرات تمثل معلمة المجتمع بشكل جيد، هناك بعض الخصائص المرغوبة لهذه المقدرات تأخذ بعين الاعتبار لاختيار افضل المقدرات وهي :

1.2.1/ *عدم التحيز* ***Unbiased Estimators***

من المرغوب احصائيا أن تكون القيمة المتوقعة للمقدر $\hat{\theta}$ تساوي القيمة الحقيقية لمعملة المجتمع θ ، ويسمى المقدر بالمقدر غير متحيز بمعنى أن : $E(\hat{\theta}) = \theta$

اضافة الى ماسبق يعطى تحيز $\hat{\theta}$ بالعلاقة التالية : $B = E\left(\hat{\theta}\right) - \theta$

لاحظ أن التحيز ليس سوى القيمة المتوقعة للخطأ (العشوائي) ، وبالتالي فإن المقدر غير متحيز إذا كان التحيز يساوي صفرًا لجميع قيم θ. يحدث التحيز عندما لا تمثل العينة بدقة المجتمع الذي تنتمي إليه العينة. من المهم ملاحظة أنه من أجل التحقق مما إذا كانت المقدر غير متحيز فليس من الضروري معرفة قيمة المعلمة الحقيقية θ .

2.2.1/ الاتساق Consistency (التقارب convergeance)

نقول عن المقدر $\hat{\theta}$ أنه مقدر متسق للمعلمة θ من أجل أي $\varepsilon > 0$ اذا كان :

$$\lim_{n\to\infty} Pr\left[\left|\hat{\theta}_n - \theta\right| \le \varepsilon\right] = 1$$

أو بصيغة مساوية

$$\lim_{n\to\infty} Pr\left[\left|\hat{\theta}_1 - \theta\right| > \varepsilon\right] = 1$$

العبارة أن $\hat{\theta}_n$ مقدر متسق للمعلمة θ تكافئ $\hat{\theta}_n$ مقدر متسق تقارب θ باحتمال، أي يجب أن يكون لمقدر العينة احتمالية عالية للاقتراب من قيمة المجتمع θ من أجل حجم عينة كبير n

- هناك شرط كافي حتى يكون المقدر الغير متحيز $\hat{\theta}_n$ متسقا للمعلمة θ و هو أن يكون :

$$\lim_{n\to\infty} Var\left(\hat{\theta}_n\right) = 0$$

انه من المهم ملاحظة أنه ليس من الضروري أن يكون المقدر المتسق غير متحيز، وبالتالي هذه النتيجة ليست شرطا ضروريا.

يمكن تلخيص اجراءات اختبار الاتساق في النقاط التالية:

1. التأكد من أن المقدر θ متحيز أم لا
2. حساب $Var\left(\hat{\theta}\right)$ و تحيز المقدر $\hat{\theta}$ أي $B\left(\hat{\theta}_n\right)$
3. مقدر غير متحيز يكون متسق اذا كان $\lim_{n\to\infty} Var\left(\hat{\theta}_n\right) = 0$
4. مقدر متحيز يكون متسقا اذا كان :

$$\lim_{n\to\infty} Var\left(\hat{\theta}_n\right) = 0$$
$$\lim_{n\to\infty} B\left(\hat{\theta}_n\right) = 0$$

<u>مثال :</u>

نعلم أن متوسط العينة $\bar{X}$ هو مقدر غير متحيز لمتوسط المجتمع μ بحيث أن $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

$$\lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

نقول أن $\bar{X}$ مقدر متسق لـ μ

3.2.1/ الفعالية (الكفاءة) Efficiency

يمكن أن يكون هناك أكثر من مقدر واحد غير متحيز للمعلمة θ بحيث أن الأقل تباينًا مرغوب فيه. هنا نقدم مفهوم الكفاءة الذي يعتمد على مقارنة التباينات بين مختلف المقدرات غير المتحيزة. إذا كان هناك مقدران غير متحيزين ، فمن المستحسن أن يكون هناك مقدر ذو تباين أصغر.

ليكن $\hat{\theta}_1$ و $\hat{\theta}_2$ مقدرين غير متحيزين لمعلمة المجتمع $\hat{\theta}$ ، ان كفاءة $\hat{\theta}_1$ بالنسبة لـ $\hat{\theta}_2$ هي النسبة :

$$e(\hat{\theta}_1,\hat{\theta}_2)=\frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$$

اذا كان : $e(\hat{\theta}_1,\hat{\theta}_2)<1$ نستنتج أن : $\hat{\theta}_2$ أكثر فعالية من $\hat{\theta}_1$

أما اذا كان : $e(\hat{\theta}_1,\hat{\theta}_2)>1$ فان : 1 أكثر فعالية من $\hat{\theta}_2$ ولذلك من بين جميع المقدرات غير متحيزة ، وبالتالي فالمقدر الأكثر فعالية هو المفضل .

2/ التقدير بمجال الثقة (فترة الثقة) : Confidence Interval

السلبية الأساسية للمقدر النقطي أنه لا يأخذ بعين الاعتبار عدم اليقين على عكس ذلك فان التقدير بمجال الثقة يأخذ بعين الاعتبار عدم اليقين حول قيمة معلمة المجتمع، لا يجب فهم أن مجال الثقة يقوم على أساس متغيرات عشوائية واحتمال حيث أن تحقق هذا المجال ما هو الا مجرد جزء محدد بقيمتين حقيقيتين وبالتالي يجب أن تجنب على سبيل المثال الكتابة التالية :

$$Pr(1.2\le\theta\le1.5)=0.95$$

ليس لها معنى لأن θ ليس ما غير عشوائي وبالتالي لا يوجد سبب لاستخدام الاحتمال في مجال الثقة لان المعلمة θ ثابتة وهذا ما يميز هذا المنهج عن المنهج البايزي.

$$IC_{0.95}[1.2\le\theta\le1.5]$$

1.2/ تعريف مجال الثقة

هو تقدير معلمة المجتمع المجهولة بين حدين حد أعلى وحد أدنى باحتمال معين حيث أن هذه الحدود هي متغيرات عشوائية وبالتالي فان فترة الثقة هي متغير عشوائي وبالتالي فان احتمال وقوع المعلمة ضمن فترة الثقة المطلوبة (ضمن الحدين) باحتمال $1-\alpha$ يعرف على الشكل التالي :

$$Pr(\hat{\theta}_1\le\theta\le\hat{\theta}_2)=1-\alpha \qquad ;\hat{\theta}_1<\hat{\theta}_2$$

$\hat{\theta}_2-\hat{\theta}_1$: طول فترة الثقة

يطلق على هذين الحدين بحدي الثقة أو مجال الثقة أو فترة الثقة ، واحتمال وقوع التقدير بين هذين الحدين باحتمال درجة الثقة (احتمال الدقة) $1-\alpha$ والاحتمال المتمم لاحتمال درجة الثقة يسمة مستوى

الدلالة الاحصائية (المعنوزية الاحصائية) ويرمز له بالرمز α ، ومع افتراض أن العينة العشوائية تتبع التوزيع الطبيعي فان مستوى الثقة $1-\alpha$ يعبر عن المساحة المحصورة تحت منحنى التوزيع الطبيعي .

2.2/ مستوى المعنوية (الدلالة الاحصائية): هو عبارة عن احتمال رفض الفرضية الصفرية بشرط أنها صحيحة بمعنى أنها تمثل نسبة الخطأ المسموح به في اتخاذ القرار في نتائج الاختبارات الاحصائية وعادة ما نرمز له بالرمز α وبذلك يكون مقدار الثقة في النتائج هو $1-\alpha$ فمثلا اذا كان مستوى المعنوية الاحصائية هو 0.05 فان مستوى الثقة يساوي 0.95. عمليا نختار القيمة $1-\alpha$ كبيرة كفاية مثلا 0.95 وبهذا يكون هناك احتمال كبير بأن يتضمن مجال الحد العشوائي

$[\hat{\theta}_1, \hat{\theta}_2]$ يحتوي على القيمة الحقيقية للمعلمة ، من جهة أخرى يمكن القول أن ضمن العينات الممكنة فانه بنسبة لا تقل عن $1-\alpha$ يوجد مجال ثقة يحتوي على المعلمة الحقيقية للمجتمع θ بافتراض أن العينة العشوائية تتوزع طبيعيا، ولتكن عبارة عن حد المعنوية (الدلالة الإحصائية) فان مستوى الثقة (درجة الثقة) تساوي $1-\alpha$ وتعبر عن المساحة المحصورة تحت المنحنى الطبيعي .

الجدول رقم (01): قيم احصائية التوزيع الطبيعي المجدولة الموافق لبعض مستويات المعنوية وفترات الثقة

مستوى الثقة	%80	%90	%95	%95,45	%98	%99
حد المعنوية α	%20	%10	%5	%4 .55	%2	%1
$Z_{\alpha/2}$	1.28	1.645	1.96	2	2.33	2.58

3.2/ تقدير مجال الثقة للمتوسط الحسابي للمجتمع

1.3.2/ مجال الثقة في حالة التوزيع الطبيعي

من أجل إيجاد مجال الثقة في حالة التوزيع الطبيعي فانه يفترض ما يلي :

- العينة عشوائية
- التوزيع الطبيعي في المجتمع أو أن تكون العينة المختارة كبيرة الحجم
- التباين في المجتمع معلوم (أو يكون مقدرا بتباين العينة)

يمكن تلخيص خطوات إيجاد مجال الثقة فيا يلي :

- ✓ الخطوة الأولى : من أجل مستوى ثقة $1-\alpha$ نستخدم الجدول الاحصائي لايجاد قيمة إحصائية التوزيع الطبيعي $Z_{\alpha/2}$
- ✓ الخطوة الثانية : مجال الثقة لمتوسط المجتمع هو :

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

$\bar{X}$: المتوسط الحسابي في العينة المختارة

✔ الخطوة الثالثة : تفسير مجال الثقة

نقول أنه بنسبة $(1-\alpha)\%$ من بين اجمالي العينات ذات الحجم n فيها متوسط حسابي له مجال الثقة السابق .

بافتراض لدينا مجتمع احصائي حجمه N وتم سحب جميع العينات العشوائية بحجم متساوي ، المتوسط الحسابي في المجتمع µ و تباين σ^2 معلوم وبالتالي فان الحدود التي يقع ضمنها المتوسط الحسابي للمجتمع يمكن صياغته على الشكل التالي:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$Pr\left(Z_{1-\alpha/2} \leq T \leq Z_{1-\alpha/2}\right) = 1-\alpha$$

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

بضرب طرفي المتباينة في $\frac{\sigma}{\sqrt{n}}$ نجد :

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

بطرح $\bar{X}$ من جميع أطراف المتباينة ينتج:

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

بضرب جميع أطراف المتباينة في العدد (-1) نجد :

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

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

يمكن كتابة مجال الثقة السابقة على الشكل :

$$\bar{X} \mp Z_{1-(\frac{\alpha}{2})}\frac{\sigma}{\sqrt{n}}$$

أو على الشكل :

$$IC_{1-\alpha}\left[\bar{X}-Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X}+Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

مثال 01:

تم اختيار عينة عشوائية من 225 وحدة من منتج غذائي في مصنع ما حيث وجد أن متوسط تكلفة الإنتاج في العينة تساوي168دج، ومن المعلوم أن الانحراف المعياري لتكلفة الإنتاج في المصنع30= σ.

المطلوب : تقدير المتوسط الحسابي لتكلفة الإنتاج في المصنع عند مستوى ثقة 95% ، 95.45% .

الحل:

لدينا :

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

عند مستوى ثقة 95.45%

$$1-\alpha = 0.9545 \Rightarrow \alpha = 0.0455$$

$$Z_{1-\alpha/2} = Z_{1-\frac{0.045}{2}} = Z_{1-\frac{0.045}{2}} = Z_{0.97725} = 2$$

$$IC_{0.9545}\left[168 - 2\frac{30}{\sqrt{225}}, 168 + 2\frac{30}{\sqrt{225}}\right]$$

$$IC_{0.9545}[164\ ,172]$$

يشير مجال الثقة أعلاه الى أن : متوسط المجتمع يقع ضمن الحدين 164 و 172 بمعنى آخر أننا واثقون بنسبة 95,45% بأن متوسط المجتمع لن يقل عن 164 دينار ولن يزيد عن 172 وحدة .

عند مستوى ثقة 99%

$$1-\alpha = 0.99 \Rightarrow \alpha = 0.01$$

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

$$IC_{0.9545}\left[168 - 2.58\frac{30}{\sqrt{225}}, 168 + 2.58\frac{30}{\sqrt{225}}\right]$$

$$IC_{0.9545}[162.84\ \ ,173.17]$$

أي أننا واثقون بنسبة 99% بأن متوسط المجتمع لن يقل عن 162.84 دينار ولن يزيد عن 173.17 وحدة .

مثال 02:

بهدف اجراء دراسة علمية قام مكتب الاحصاء في احدى الدول بجمع معلومات حول أعمار الأفراد من القوى العاملة المدنية حيث تم اختيار عينة عشوائية من 50 عاملا ووجد أن المتوسط الحسابي للأعمار يساوي 36.4 سنة .

المطلوب: عند مستوى ثقة 95% أوجد مجال الثقة لمتوسط عمر جميع القوى العاملة بافتراض أن الانحراف المعياري في المجتمع يساوي 12.1 سنة

الحل :

$$1-\alpha = 0.95 \Rightarrow \alpha = 1-0.95 = 0.05$$

من خلال الجدول الاحصائي للتوزيع الطبيعي نجد أن :

$$Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$$

مجال الثقة يعطى بالعلاقة التالية :

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

$$IC_{0.95}\left[36.4 - 1.96\frac{12.1}{\sqrt{50}}, 36.4 + 1.96\frac{12.1}{\sqrt{50}}\right]$$

$$IC_{0.95}[33 \quad , \quad 39.8]$$

التحليل: نحن على ثقة بنسبة 95% أن متوسط سن كل أفراد القوى العاملة المدنية محصور بين 33 سنة و 39.8 سنة

***مثال03*:**

مصنع في مدينة ما ينتج مصابيح كهربائية، تم سحب عينة عشوائية من 200 مصباح اذا وجد المتوسط الحسابي $\bar{X} = 501.2$ ، من معطيات سابقة وجد أن التباين في المجتمع $\sigma^2 = 16$

المطلوب : أوجد فترة الثقة لمتوسط عمر المصابيح المنتجة من قبل المصنع بمستوى ثقة 95% ؟

الحل :

عند مستوى معنوية 5% نجد قيمة احصائية التوزيع الطبيعي $Z_{\frac{\alpha}{2}} = Z_{1-\frac{\alpha}{2}} = Z_{0.975}$

لدينا:

$$Pr\left(\bar{X} - Z_{1-\alpha/2}\sigma_{\bar{X}} \leq \mu \leq \bar{X} + Z_{1-\alpha/2}\sigma_{\bar{X}}\right) = 1-\alpha$$

$$IC_{0.95}\left[501.2-1.96\frac{4}{\sqrt{100}},501.2+1.96\frac{4}{\sqrt{100}}\right]$$
$$IC_{0.95}[500.416\quad,\qquad 501.984]$$

2.3.2/ مجال الثقة في حالة توزيع ستيودنت

في حالة عدم معلومية σ^2 من مجتمع موزع توزيعا طبيعيا فان هذه العينة تتوزع توزيع ستيودنت t بدرجة حرية $n-1$ وعليه فان المتغير العشوائي.

$$T=\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}\backsim t_v$$
$$\sigma_{\bar{X}}=\frac{S}{\sqrt{n}}$$
$$Pr\left(-T_{\frac{\alpha}{2}}\leq T\leq T_{\frac{\alpha}{2}}\right)=1-\alpha$$
$$Pr\left(\bar{X}-\frac{S}{\sqrt{n}}T_{\frac{\alpha}{2};(n-1)}\leq\mu\leq\bar{X}+\frac{S}{\sqrt{n}}T_{\frac{\alpha}{2};(n-1)}\right)=1-\alpha$$

<u>مثال04</u> :

في مصنع لانتاج القمح اللين تم اختيار عينة عشوائية من 25 كيس فوجد أن متوسط الوزن في العينة يساوي 45 كلغ والانحراف المعياري يساوي 4 كلغ.

المطلوب : أوجد فترة الثقة للمتوسط الحسابي لانتاج المصنع من القمح μ ؟

1. عند مستوى ثقة 95%
2. عند مستوى ثقة 99% ؟

<u>الحل</u>:

درجة الحرية n−1=25−1=24

$$T_{\frac{\alpha}{2};(n-1)}=T_{0.025;(24)}=2.064$$
$$Pr\left(\bar{X}-\frac{S}{\sqrt{n}}T_{\frac{\alpha}{2};(n-1)}\leq\mu\leq\bar{X}+\frac{S}{\sqrt{n}}T_{\frac{\alpha}{2};(n-1)}\right)=1-\alpha$$

1. عند مستوى ثقة 95%

$$IC_{0.95}\left[45-2.064\frac{4}{\sqrt{25}},45+2.064\frac{4}{\sqrt{25}}\right]$$
$$IC_{0.95}[43.3488\quad,46.6512]$$

2. عند مستوى ثقة 99%

$$T_{\frac{\alpha}{2};(n-1)} = T_{0.05;(24)} = 2.797$$

$$IC_{0.99}\left[45 - 2.797\frac{4}{\sqrt{25}}, 45 + 2.797\frac{4}{\sqrt{25}}\right]$$
$$IC_{0.99}[42.7624 \quad , 47.2376]$$

4.2/ مجال الثقة للفرق بين متوسطي عينتني مستقلتين

1.4.2/ في حالة التوزيع الطبيعي (أو/و حجم العينتين كبير)

$$Pr\left(Z_{1-\alpha/2} \le Z \le Z_{1-\alpha/2}\right) = 1-\alpha$$

$$Pr\left(Z_{1-\alpha/2} \le \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} \le Z_{1-\alpha/2}\right) = 1-\alpha$$

$$Pr\left(-Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2} \le (\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2) \le Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2}\right) = 1-\alpha$$

$$Pr\left(-(\bar{X}_1 - \bar{X}_2) - Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2} \le -(\bar{X}_1 - \bar{X}_2) - (\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2) \le Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2}\right) = 1-\alpha$$

بضرب جميع الأطراف في (-1) ثم اعادة ترتيب المتراجحة فنحصل على العلاقة التالية :

$$Pr\left(-(\bar{X}_1 - \bar{X}_2) - Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2} \le -(\mu_1 - \mu_2) \le -(\bar{X}_1 - \bar{X}_2) + Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2}\right) = 1-\alpha$$

$$Pr\left((\bar{X}_1 - \bar{X}_2) - Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2} \le (\mu_1 - \mu_2) \le (\bar{X}_1 - \bar{X}_2) + Z_{1-\alpha/2}\sigma_{\bar{X}_1-\bar{X}_2}\right) = 1-\alpha$$

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

$$IC_{1-\alpha}\left[(\bar{X}_1 - \bar{X}_2) - Z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\bar{X}_1 - \bar{X}_2) + Z_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

مثال05 :

في امتحان تعليم السياقة كانت قيم متوسط النقاط لعينة عشوائية من 50 امراة و 75 رجل كما يلي :

المجموعة الأولى (النساء)	المجموعة الثانية (رجال)
$n_1 = 50$	$n_2 = 75$
$\bar{X}_1 = 76$	$\bar{X}_2 = 82$
$\sigma_1 = 6$	$\sigma_1 = 8$

المطلوب: ايجاد فترة الثقة للفرق بين متوسطي نقاط النساء والرجال لهذا المجتمع بمستوى ثقة96% ؟

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

$$IC_{0.96}\left[(76-82)-2.06\sqrt{\frac{6^2}{50}+\frac{8^2}{75}}\ ,(76-82)+2.06\sqrt{\frac{6^2}{50}+\frac{8^2}{75}}\right]$$

$$IC_{0.96}[-6-2.06(1.25)\ ,-6+2.06(1.25)]$$

مجال الثقة المطلوب هو : $IC_{0.96}[-8.575,-3.425]$

2.4.2/ مجال الثقة للفرق بين متوسطي عينتني مستقلتين في حالة توزيع ستيودنت

عند درجة حرية $v=n_1+n_2-2$ تعطى إحصائية توزيع ستيودنت **T** على النحو التالي:

$$T=\frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}\rightsquigarrow t(n_1+n_2-2)$$

$$Pr\left(T_{1-\alpha/2}\le T\le T_{1-\alpha/2}\right)=1-\alpha$$

$$Pr\left(-T_{\frac{\alpha}{2},n_1+n_2-2}\le\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}}\le T_{\frac{\alpha}{2},n_1+n_2-2}\right)=1-\alpha$$

$$Pr\left(T_{\frac{\alpha}{2},n_1+n_2-2}\le(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)\le T_{\frac{\alpha}{2},n_1+n_2-2}\right)=1-\alpha$$

$$Pr\left(-(\overline{X}_1-\overline{X}_2)-T_{\frac{\alpha}{2},n_1+n_2-2}\sigma_{\bar{X}_1-\bar{X}_2}\le-(\overline{X}_1-\overline{X}_2)-(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)\right.$$
$$\left.\le T_{\frac{\alpha}{2},n_1+n_2-2}\sigma_{\bar{X}_1-\bar{X}_2}\right)=1-\alpha$$

بضرب جميع الأطراف في (-1) ثم اعادة ترتيب المتراجحة فنحصل على العلاقة التالية :

$$Pr\left((\overline{X}_1-\overline{X}_2)-T_{\frac{\alpha}{2},n_1+n_2-2}\sigma_{\bar{X}_1-\bar{X}_2}\le(\mu_1-\mu_2)\le(\overline{X}_1-\overline{X}_2)+T_{\frac{\alpha}{2},n_1+n_2-2}\sigma_{\bar{X}_1-\bar{X}_2}\right)=1-\alpha$$

اذا كان التباين في المجتمعين مجهول ومتساوي تعطى صيغته على الشكل :

$$\sigma^2_{\bar{X}_1-\bar{X}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}=\widehat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

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

$$Pr\left((\overline{X}_1-\overline{X}_2)-T_{(\frac{\alpha}{2},n_1+n_2-2)}\sqrt{\widehat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\le(\mu_1-\mu_2)\le(\overline{X}_1-\overline{X}_2)+T_{(\frac{\alpha}{2},n_1+n_2-2)}\sqrt{\widehat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)$$
$$=1-\alpha$$

مثال 06:

قام باحث في جامعة معسكر بعد صدور نتائج امتحان مادة الاحصاء باختيار عينة عشوائية من 14 طالب (ة) من قسم علوم التسيير و 75 طالب(ة) من قسم العلوم الاقتصادية وكانت البيانات المتحصل عليها على النحو التالي :

قسم علوم التسيير	قسم العلوم الاقتصادية
$n_1 = 14$ $\overline{X}_1 = 72$ $S_1 = 5$	$n_2 = 12$ $\overline{X}_2 = 75$ $S_2 = 6$

المطلوب : ايجاد فترة الثقة للفرق بين متوسطي النقاط في مادة الاحصاء بين قسم علوم التسيير و قسم العلوم الاقتصادية عند مستوى ثقة 98% :

$$\hat{\sigma}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{(14-1)5^2 + (14-1)6^2}{14+12-2} = 30.04$$

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{\hat{\sigma}^2}{n_1} + \frac{\hat{\sigma}^2}{n_2} = \hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$Pr\left((\overline{X}_1 - \overline{X}_2) - T_{(\frac{\alpha}{2}, n_1+n_2-2)}\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \le (\mu_1 - \mu_2) \le (\overline{X}_1 - \overline{X}_2) + T_{(\frac{\alpha}{2}, n_1+n_2-2)}\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \right) = 1-\alpha$$

$$IC_{0.98}\left[(72-75) - 2.485\sqrt{30.04\left(\frac{1}{14}+\frac{1}{12}\right)}, (72-75) + 2.485\sqrt{30.04\left(\frac{1}{14}+\frac{1}{12}\right)} = 0.98\right]$$

$$IC_{0.98}[-9.468, 3.468]$$

5.2/ مجال الثقة للنسبة

لنعتبر توزيع ذو الحدين بالمعلمة p ، ولتكن X عدد النجاحات خلال n محاولة ، $\hat{p}$ يعتبر مقدر للمعلمة p حيث $\hat{p}=\frac{X}{n}$ ، يمكن تقدير نسبة تحقق الظاهرة أو الصفة المدروسة في المجتمع الاحصائي p الذي سحبت منه جميع العينات المتساوية الحجم من خلال انشاء حدي الثقة حول قيمة إحصائية العينة على النحو التالي:

$$Pr\left(\hat{p}-\sigma_{\hat{p}}Z_{\frac{\alpha}{2}}\leq p\leq \hat{p}+\sigma_{\hat{p}}Z_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$Pr\left(\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\leq \hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)=1-\alpha$$

أو بصيغة أخرى يمكن كتابة مجال الثقة كما يلي:

$$\hat{p}\mp\sigma_{\hat{p}}Z_{\frac{\alpha}{2}}$$

في حالة ما تكون النسبة p في المجتمع غير معلومة فانه يتم تعويضها بنسبة العينة $\hat{p}$

$$\sigma_{\hat{p}}=\sqrt{\frac{p(1-p)}{n}}=\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

مثال07:

قام باحث احصائي بعمل استطلاع رأي حول مدى رضى الزبائن حول منتوج غذائي جديد لمؤسسة إنتاجية، ولتحقيق هذا الغرض قام هذا الباحث باختيار عينة عشوائية من 100 زبون من بين اجمالي زبائن هذه المؤسسة، فوجد أن 55% من أفراد هذه العينة راضون عن جودة هذا المنتوج .

المطلوب : عند مستوى ثقة 99 % أوجد مجال الثقة لنسبة المقتنعين بهذا المجتمع من بين جميع زبائن هذه المؤسسة التجارية .

الحل :

$$Pr\left(\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\leq p\leq \hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)=1-\alpha$$

$$IC_{0.99}\left[0.55-2.58\sqrt{\frac{0.55(0.45)}{100}},0.55+2.58\sqrt{\frac{0.55(0.45)}{100}}\right]$$

$$IC_{0.99}[0.55-0.1283\ \ ,0.55+0.1283]$$

$$IC_{0.99}[0.4217\ \ ,0.6783]$$

عند مستوى ثقة 99% (احتمال 0.99) نحن واثقون أن نسبة زبائن هذه المؤسسة الراضين عن جودة هذا المنتوج محصورة بين 42.17% و 67.83%

مثال08:

ترغب هيئة استطلاع للرأي العام في تقدير النسبة المتوقعة للراضين عن منتوج معين في حدود بين ∓ 0.06 من النسبة الحقيقية للمجتمع بين الناخبين .

المطلوب : عند مستوى ثقة 90% ماهو حجم العينة الأدنى اذا كانت استطلاعات الرأي في عينة عشوائية مختارة تشير الى أن نسبة المصوتين لهذا المرشح هي 0.3 ؟

الحل :

لدينا مجال الثقة لنسبة الراضين عن الخدمة المقدمة للهذه المؤسسة :

$$Pr\left(\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} Z_{1-\frac{\alpha}{2}} \leq p \leq \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

نعلم أن :

$$Z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

$$Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \hat{p} - p$$

بتربيع الطرفين :

$$(\hat{p} - p)^2 = \left(Z_{1-\frac{\alpha}{2}}\right)^2 \frac{\hat{p}(1-\hat{p})}{n}$$

$$n = \frac{Z_{1-\frac{\alpha}{2}}{}^2 \hat{p}(1-\hat{p})}{(\hat{p} - p)^2} = \frac{(1.64^2)(0.3)(0.7)}{(0.06)^2} = 156.89$$

$$n \cong 157$$

6.2/ مجال الثقة للفرق بين نسبتي مجتمعين مستقلين

$$Pr\left((\hat{p}_1-\hat{p}_2)-Z_{1-\frac{\alpha}{2}}\ \sigma_{\hat{p}_1-\hat{p}_2} \leq (p_1-p_1) \leq (\hat{p}_1-\hat{p}_2)-Z_{1-\frac{\alpha}{2}}\ \sigma_{\hat{p}_1-\hat{p}_2}\right)=1-\alpha$$

$$\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}$$

$$Pr\left((\hat{p}_1-\hat{p}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}} \leq (p_1-p_2) \leq (\hat{p}_1-\hat{p}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}\right)=1-\alpha$$

في حالة ما تكون النسبتين p_2, p_1 في المجتمع غير معلومتين فانه يتم تعويضهما بنسبتي العينتين $\hat{p}_1,\ \hat{p}_2$

$$\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}$$

مثال09 :

خلال *سنة* 2022 في موسم الحج خلال سنة معينة قام باحث احصائي باختيار عينة عشوائية من600 زبون من بين زبائن وكالة سياحية *A* و 400 زبون من وكالة سياحية *B* ، حيث وجد في العينة الأولى أن 300 زبون لم تعترضهم عراقيل صعبة أثناء موسم الحج مع الوكالة *A* بينما في العينة الثانية المختارة من زبائن الوكالة *B* وجد أن 100 زبون لم تعترضهم عراقيل صعبة أثناء موسم الحج مع هذه الوكالة السياحية .

المطلوب: أوجد مجال الثقة للفرق بين نسبتي الزبائن الذين لم تعترضهم عراقيل في الوكالتين الأولى والثانية على التوالي عند مستوى *95.5%* .

الحل : نوجد أولا نسبة الزبائن الذين لم تعترضهم عراقيل في العينة الأولى والثانية

$$Z_{1-\frac{0.045}{2}}=Z_{0.9775}=2.98 \qquad \hat{p}_1=\frac{300}{600}=50\% \qquad \hat{p}_2=\frac{100}{400}=25\%$$

$$Pr\left((\hat{p}_1-\hat{p}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}} \leq (p_1-p_1) \leq (\hat{p}_1-\hat{p}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}\right)=1-\alpha$$

$$IC_{0.955}\left[(0.5-0.25)-2.98\sqrt{\frac{(0.5)(0.5)}{600}+\frac{(0.25)(0.75)}{300}},(0.5-0.25)+2.98\sqrt{\frac{(0.5)(0.5)}{600}+\frac{(0.25)(0.75)}{300}}\right]$$

$$IC_{0.955}[0.1904,0.3096]$$

باحتمال 0.955 (بمستوى ثقة *95.5%*) نحن واثقون أن الفرق بين نسبتي الزبائن الذين لم تعترضهم عراقيل من الوكالتين أثناء موسم الحج محصورة بين 19.04% و 30.96%.

7.2/ مجال الثقة للتباين

يمكن التمييز بين حالتين وفقا لما اذا كانت قيمة μ مجهولة أو معلومة

1.7.2/ حالة المتوسط في المجتمع الاحصائي μ مجهول

لتكن $X_1, X_2 \dots X_n$ بيانات متغيرة عشوائية تتبع التوزيع الطبيعي بمتوسط μ وتباين σ^2 حيث كل منهما مجهول، و نعلم أن :

$$\frac{(n-1)s^2}{\sigma^2} \rightsquigarrow {\chi^2}_{(n-1)}$$

تتبع توزيع مربع كاي بدرجة حرية $(n-1)$ ، يمكننا اذا ايجاد الحدين ${\chi^2}_L$ و ${\chi^2}_U$ لمجال الثقة للتباين في المجتمع كما يلي :

$$Pr\left({\chi^2}_L \leq \frac{(n-1)s^2}{\sigma^2} \leq {\chi^2}_U\right) = 1-\alpha$$

$$Pr\left(\frac{(n-1)s^2}{{\chi^2}_U} \leq \sigma^2 \leq \frac{(n-1)s^2}{{\chi^2}_L}\right) = 1-\alpha$$

من أجل التسهيل نضع في الجهة اليمنى ${\chi^2}_U = \chi^2_{\alpha/2}$ وفي الجهة اليسرى ${\chi^2}_L = \chi^2_{1-(\frac{\alpha}{2})}$

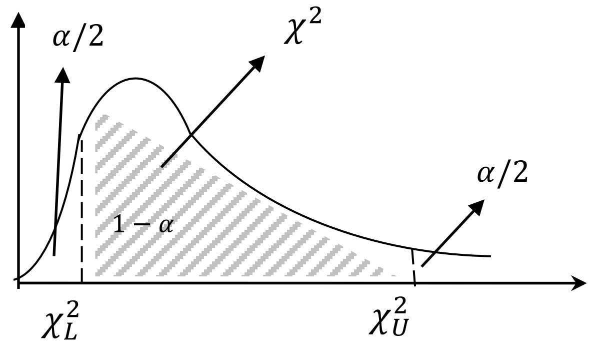
باستخدام جدول توزيع مربع كاي (هو توزيع احتمالي غير متناظر) يمكن ايجاد قيم $\chi^2_{1-(\frac{\alpha}{2})}$ و $\chi^2_{\alpha/2}$

نظرية : اذا كانت $\overline{X}$ المتوسط الحسابي و S الخطأ المعياري(الانحراف المعياري) لعينة عشوائية بحجم n من مجتمع طبيعي اذا :

$$Pr\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-(\frac{\alpha}{2})}}\right) = 1-\alpha$$

لدينا ثقة بـ $(1-\alpha)100\%$ أن تباين المجتمع σ^2 يقع في المجال $\frac{(n-1)s^2}{\chi^2_{1-(\alpha/2)}}$ و $\frac{(n-1)s^2}{\chi^2_{\alpha/2}}$

الشكل رقم *(01)*: منحنى توزيع مربع كاي بمساحة متساوية على جانبي مجال الثقة



المصدر: *KandethodyM.Ramachandran, Chris P.Tsokos(2009), Mathematical Statistics with Applications, Elsevier Academic Press P316*

بافتراض أن المجتمع الاحصائي موزع طبيعيا يمكن ايجاز مراحل ايجاد مجال الثقة لتباين المجتمع فيمايلي :

- ✓ حساب متوسط وتباين العينة $\bar{X}$ و s^2
- ✓ ايجاد حدي احصائية مربع كاي من جدول التوزيع الاحتمالي $\chi^2{}_U = \chi^2_{\alpha/2}$ و $\chi^2{}_L = \chi^2_{1-(\frac{\alpha}{2})}$

وذلك عند درجة الحرية $v = n-1$

- ✓ حساب مجال الثقة بنسبة $100\%(1-\alpha)$ لتباين المجتمع σ^2 وهو $\frac{(n-1)s^2}{\chi^2_{\alpha/2}}$, $\frac{(n-1)s^2}{\chi^2_{1-(\frac{\alpha}{2})}}$

مثال10 :

من مجتمع طبيعي تم اختيار عينة عشوائية حجمها $n=21$ بانحراف معياري يساوي 09.

المطلوب : عند مستوى ثقة 90% حدد مجال الثقة لتباين المجتمع σ^2 ؟

الحل:

لدينا $n=21$ و $s^2 = 9^2 = 81$

عند درجة الحرية $v=n-1=20$ نجد احصائية مربع كاي المجدولة $\chi^2_{\alpha/2} = \chi^2_{0.05} = 31.4104$ و

$\chi^2_{1-(\frac{\alpha}{2})} = \chi^2_{0.95} = 10.8508$

يعطى مجال الثقة عند مستوى ثقة 90% كما يلي :

$$Pr\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-(\frac{\alpha}{2})}}\right) = 1-\alpha$$

$$IC_{0.90}\left[\frac{(20-1)81}{31.4104} \quad , \frac{(20-1)81}{10.8508}\right]$$

$$IC_{0.90}[1.575 \quad , 149.298]$$

نحن على ثقة بنسبة 90% أن تباين المجتمع محصور في المجال أعلاه

ملاحظة : من أجل ايجاد مجال الثقة للانحراف المعياري في المجتمع σ عند نفس مستوى الثقة السابق، فإننا نأخذ الجذر التربيعي لمجال الثقة الأخير المتحصل عليه .

2.7.2/ حالة المتوسط في المجتمع الاحصائي $\boldsymbol{\mu}$ معلوم

لتكن$\boldsymbol{X_1, X_2 \dots X_n}$ بيانات عينة عشوائية مسحوبة من مجتمع طبيعي $N(\boldsymbol{\mu}, \sigma^2)$ ، حيث $Z_i = \frac{(X_i-\mu)}{\sigma}$ مستقلة عن بعضها البعض موزعة توزع طبيعي معياري بحيث:

$$\sum_{i=1}^{n}(Z_i)^2 = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \rightsquigarrow {\chi^2}_{(n)}$$

نعلم أن :

اذا كان : $X_1, X_2 \dots X_n$ مستقلة فيما بينها وتتبع التوزيع الطبيعي المعياري فان : $\sum_{i=1}^{n}(X_i)^2 \rightsquigarrow {\chi^2}_{(n)}$

مجال الثقة في حالة $\boldsymbol{\mu}$ معلوم يعطى على الشكل التالي :

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

$$Pr\left(\chi^2_{1-\alpha/2,(n)} \le \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \le \chi^2_{\alpha/2,(n)}\right) = 1-\alpha$$

$$Pr\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2,(n)}} \le \sigma^2 \le \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2,(n)}}\right) = 1-\alpha$$

تمارين محلولـــــة

التمرين رقم 01

لتكن X متغيرة عشوائية تتبع التوزيع الطبيعي ذو المتوسط μ والانحراف المعياري σ ، نستخرج عينتين مستقلتين من المجتمع حيث $\bar{X}_2$ ، $\bar{X}_2$ متوسطيهما على التوالي ، ليكن لدينا المقدرين التاليين لمتوسط المجتمع μ :

$$\hat{\mu}_1 = (\bar{X}_1 + \bar{X}_2)/2$$

$$\hat{\mu}_2 = (n_1 \bar{X}_1 + n_2 \bar{X}_2) / (n_1 + n_2)$$

المطلوب:

1. من بين المقدرين السابقين أيهما غير متحيز ؟
2. أحسب تباين كل مقدر وقارن بينهما من حيث الفعالية ؟
3. من خلال ما سبق استنتج أحسن مقدر من بين المقدرين السابقين ؟

الحل :

1. تبيان أي من المقدرين μ_1 أو μ_2 غير متحيز ؟

$$E(\hat{\mu}_1) = E[\bar{X}_1 + \bar{X}_2)/2] = \frac{1}{2}E[\bar{X}_1 + \bar{X}_2)] = \frac{1}{2}[E(\bar{X}_1) + E(\bar{X}_2)]$$

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

$$E(\hat{\mu}_2) =$$

$$E[(n_1 \bar{X}_1 + n_2 \bar{X}_2)/(n_1 + n_2)] = \frac{1}{(n_1 + n_2)}E[n_1 \bar{X}_1 + n_2 \bar{X}_2)] = \frac{1}{(n_1 + n_2)}[n_1 E(\bar{X}_1) + n_2 E(\bar{X}_2)]$$

$$= \frac{1}{(n_1 + n_2)}[n_1\mu + n_2\mu] = \mu$$

نستنتج أن μ_1 أو μ_2 هما مقدرين غير متحيزين لمتوسط المجتمع μ

2. ايجاد تباين كل مقدر ثم المقارنة بينهما من حيث الفعالية :

$$Var(\hat{\mu}_1) = Var[\bar{X}_1 + \bar{X}_2)/2] = \left(\frac{1}{2}\right)^2 Var[\bar{X}_1 + \bar{X}_2)]$$

$$= \frac{1}{4}[Var(\bar{X}_1) + Var(\bar{X}_2)]$$

$$\frac{1}{4}\left[\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right] = \frac{\sigma^2(n_1 + n_2)}{4 n_1 n_2}$$

$$Var(\hat{\mu}_2) = Var\left[\frac{(n_1 \bar{X}_1 + n_2 \bar{X}_2)}{n_1 + n_2}\right]$$

$$= \left(\frac{1}{n_1 + n_2}\right)^2 [(n_1)^2 Var\,\bar{X}_1 + (n_2)^2\, Var\,\bar{X}_2)\,]$$

$$= \left(\frac{1}{n_1 + n_2}\right)^2 \left[(n_1)^2 \frac{\sigma^2}{n_1} + (n_2)^2 \frac{\sigma^2}{n_2}\right]$$

$$= \left(\frac{1}{n_1 + n_2}\right)^2 [\sigma^2(\,n_1 + n_2)\,] = \sigma^2 \left(\frac{1}{n_1 + n_2}\right)$$

$$\frac{Var(\hat{\mu}_2)}{Var(\hat{\mu}_1)} = \frac{\sigma^2 \left(\frac{1}{n_1 + n_2}\right)}{\frac{\sigma^2(\,n_1 + n_2)}{4\,n_1\,n_2}} = \frac{4\,n_1\,n_2}{(\,n_1 + n_2)^2}$$

من أجل $n_1 = n_2$ و $n_1, n_2 > 1$ فان : $\frac{Var(\hat{\mu}_2)}{Var(\hat{\mu}_1)} < 1$

وبالتالي :

$$Var(\hat{\mu}_1) > Var(\hat{\mu}_2)$$

بما أن المقدر $\hat{\mu}_2$ له التباين الأصغر فهو اذا المقدر الأكثر فعالية مقارنة بالمقدر $\hat{\mu}_1$

3. استنتاج أي من المقدرين السابقين هو الأحسن

بما أن $\hat{\mu}_2$ هو مقدر فعال وغير متحيز وفي المقابل $\hat{\mu}_1$ هو مقدر متحيز فقط فان $\hat{\mu}_2$ يعتبر أحسن مقدر

التمرين رقم 02

أجريت دراسة حول قيمة الودائع الادخارية في أحد البنوك التجارية، حيث تم اختيار عينة عشوائية من 20 حساب بنكي خلال فترة زمنية محددة، فكان متوسط الادخار هو 300 وحدة نقدية والانحراف المعياري في العينة يساوي 100 وحدة نقدية .

المطلوب : عند مستوى ثقة 90% أوجد مجال الثقة لمتوسط قيم الادخار في البنك مع العلم أن قيمة الودائع تتبع التوزيع الطبيعي في مجتمع الدراسة ؟

الحل :

في حالة عدم معلومية تباين المجتمع σ^2 من مجتمع موزع توزيعا طبيعيا فان هذه العينة تتوزع توزيع ستيودنت t بدرجة حرية $n-1$ وعليه فان المتغير العشوائي 0

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_v$$

$$\sigma_{\bar{X}} = \frac{S}{\sqrt{n}}$$

من خلال الجدول الاحصائي لتوزيع ستيودنت نجد احصائية ستيودنت المجدول عند درجة الحرية

: v=19

$$T_{\frac{\alpha}{2};(n-1)} = T_{\frac{0.1}{2};(20-1)} = T_{0.05;(19)} = 1.729$$

ويعطى مجال الثقة كما يلي :

$$\mu \in \bar{X} \mp \frac{S}{\sqrt{n}} T_{\frac{\alpha}{2};(n-1)}$$

$$\mu \in 300 \mp \frac{100}{\sqrt{20}}(1.729)$$

$$\mu \in 300 \mp 22.3608(1.729)$$

$$\mu \in 300 \mp 38.6618$$

$$IC_{0.9}[261.3382; 338.6618]$$

التمرين رقم 03

في مصنع لانتاج بطاريات من الحجم المتوسط (مثلا تستخدم في أجهزة التحكم في التلفاز) لأحد العلامات التجارية تم اختيار عينة عشوائية من 09 بطاريات من شحنة معدة للتصدير الى خارج البلد، حيث وجد أن متوسط الاستعمال بالدقيقة هو $\bar{X} = 300$ ، بانحراف معياري S=45

المطلوب : مع العلم ان مدة استعمال هذه البطاريات تتبع التوزيع الطبيعي في مجتمع الدراسة عند مستوى ثقة 90% حدد مجال الثقة لمتوسط استعمال هذه البطاريات لكامل هذه الشحنة ؟

الحل :

بما أن تباين المجتمع σ^2 غير معلوم و المجتمع موزع توزيعا طبيعيا فان هذه العينة تتوزع توزيع ستيودنت t بدرجة حرية $n-1$ وعليه فان المتغير العشوائي

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \backsim t_v$$

$$\sigma_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{45}{\sqrt{9}} = 15$$

$$T_{\frac{\alpha}{2};(n-1)} = T_{\frac{0.1}{2};(9-1)} = T_{0.05;(8)} = 1.86$$

$$\mu \in \bar{X} \mp \frac{S}{\sqrt{n}} T_{\frac{\alpha}{2};(n-1)}$$

$$\mu \in 300 \mp (15)(1.86)$$

$$\mu \in 300 \mp 27.9$$

$$IC_{0.9}[272.1\ ;\ 327.9]$$

التمرين رقم 04

من خلال تجارب سابقة من المعلوم في مصنع أن قيمة الأجور الشهرية للعمال تتبع التوزيع الطبيعي بانحراف معياري يساوي $\sigma = 40$ وحدة نقدية .

المطلوب: ماهو حجم العينة اللازم من الموظفين حتى لا يتجاوز خطأ التقدير (هامش الخطأ)في متوسط الأجور قيمة ∓ 20 وحدة نقدية(انحراف متوسط العينة عن متوسط المجتمع بـ∓ 20) عند مستوى ثقة % 99 ؟

الحل:

$$|\bar{X} - \mu| = 0.3$$

$: \bar{X} - \mu = \mp 20$

نعلم أن :

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \qquad , \ \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad \bar{X} \backsim N(\mu, \sigma_{\bar{X}}^2)$$

$$Z = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}} \quad , \ Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.01}{2}} = Z_{0.995} = 2.58$$

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

$$|\bar{X} - \mu| = \frac{\sigma}{\sqrt{n}} Z_{\frac{\alpha}{2}}$$

$$(\bar{X} - \mu) \ = \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}$$

$$\sqrt{\boldsymbol{n}} = \frac{\boldsymbol{\sigma}\, \boldsymbol{Z}_{\mathbf{1}-\frac{\alpha}{2}}}{(\bar{\boldsymbol{X}} - \boldsymbol{\mu})}$$

$$= \frac{(40)(2.58)}{20} = 5.16$$

$$n = (5.16)^2 = 26.62$$

$$n \cong 27$$

التمرين رقم 05

في مصنع لانتاج زجاجات العصير معلوم أن كمية التعبئة X التي يتم تفريغها في الزجاجات تتبع التوزيع الطبيعي بمتوسط μ وتباين $\sigma^2 = 1$ (الوحدة المليلتر) ، تم اختيار عينة عشوائية بسيطة من n=09 زجاجة عصير من مخرجات الانتاج في يوم معين .

المطلوب : باعتبار أن المجتمع الاحصائي غير محدود :

1. أحسب احتمال أن يكون متوسط العينة على الأقل منحرفا بقيمة0.3 عن المتوسط الحقيقي للمجتمع ؟
2. عند مستوى ثقة 0.95 أوجد مجال الثقة للمتوسط μ حيث يعطى المتوسط الحسابي في العينة $\bar{x} = 100$ ؟
3. أوجد عدد المشاهدات اللازمة حتى ينحرف متوسط العينة بقيمة 0.3 عن المتوسط الحقيقي للمجتمع باحتمال 0.95 ؟

الحل:

1. حساب احتمال انحراف متوسط العينة عن المتوسط الحقيقي للمجتمع بقيمة أقل أو يساوي 0.3

بمعنى ايجاد احتمال أن يكون خطأ التقدير أقل أو يساوي 0,3 : $\boldsymbol{Pr(|\bar{X} - \mu|) \leq 0,3}$

لدينا :

$$X \backsim N(\mu, 1)$$

اذا :

$$\bar{X} \backsim N(\mu, \sigma_{\bar{X}}^2)$$
$$\sigma_{\bar{X}}^2 = \frac{\sigma}{\sqrt{n}}$$
$$Pr(|\bar{X} - \mu|) \leq 0,3$$
$$= Pr(-0,3 \leq \bar{X} - \mu \leq 0,3)$$
$$= Pr\left(\frac{-0,3}{\sigma_{\bar{X}}} \leq \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \leq \frac{0,3}{\sigma_{\bar{X}}}\right)$$
$$= Pr\left(\frac{-0,3}{1/\sqrt{9}} \leq \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \leq \frac{0,3}{1/\sqrt{9}}\right)$$

$$= Pr(-0,9 \leq Z \leq 0,9) =$$
$$= Pr(Z \leq 0,9) - Pr(Z \leq -0,9) = 2Pr(Z \leq 0,9) - 1$$
$$= 2(0.8159) - 1 = 0.6318$$
$$\boldsymbol{Pr(|\bar{X} - \mu|) \leq 0,3 = 0.6318}$$

2. حساب عدد المشاهدات اللازمة حتى ينحرف متوسط العينة بقيمة 0.3 عن المتوسط الحقيقي للمجتمع باحتمال 0.95

نبحث عن حجم العينة اللازم حتى يحقق خطأ التقدير المتراجحة : $Pr\left((|\bar{X}-\mu|) \leq 0{,}3\right) = 0{,}95$

$$Pr(|\bar{X}-\mu|) \leq 0{,}3 = Pr(-0{,}3 \leq \bar{X}-\mu \leq 0{,}3)$$
$$= Pr\left(\frac{-0{,}3}{\sigma/\sqrt{n}} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{0{,}3}{\sigma/\sqrt{n}}\right)$$
$$Pr\left(-Z_{1-\frac{\alpha}{2}} \leq Z \leq Z_{1-\frac{\alpha}{2}}\right) = 0.95$$
$$Pr(-1.96 \leq Z \leq 1.96) = 0.95$$
$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = Z_{1-\frac{\alpha}{2}}$$
$$\sqrt{n} = \frac{\sigma\, Z_{1-\frac{\alpha}{2}}}{(\bar{X}-\mu)} \quad = \frac{(1)(1.96)}{0.3}$$
$$\sqrt{n} = 6.5333$$
$$n = (5.16)^2 = 42.6844$$
$$n \cong 43$$

التمرين رقم 06

في مؤسسة تجارية توظف 1200 عامل ، تم اختيار عينة عشوائية n=100 عامل حيث تم استطلاع رأيهم حول تفضيلهم تمويل ذاتي لتقاعدهم عوضا عن الاشتراك في مخطط التقاعد الذي تشرف عليه أحد شركات التأمين فتبين أن 70 عاملا منهم يقبلون هذه الفكرة .

المطلوب:

1. قدر نقطيا نسبة العمال الذين يفضلون تمويل ذاتي لتقاعدهم عوضا عن الاشتراك في مخطط التقاعد؟
2. عند مستوى ثقة 95 % حدد مجال الثقة لنسبة العمال الذين يفضلون تمويل ذاتي لتقاعدهم ؟

الحل :

1. التقدير النقطي لنسبة العمال الذين يفضلون تمويل ذاتي لتقاعدهم

$$\widehat{P} = 100 = 0.7$$

$\mathbf{\widehat{P} = 0.7}$: يمثل نسبة العمال الذين يفضلون تمويل ذاتي لتقاعدهم في عينة الدراسة ويعتبر مقدر نقطي لنسبة العمال لنسبة العمال الذين يفضلون تمويل ذاتي لتقاعدهم في المؤسسة

2. ايجاد مجال الثقة لنسبة العمال الذين يفضلون تمويل ذاتي لتقاعدهم عند مستوى ثقة 95 %

لدينا:

np=35>5 , n(1−p)=65>5

فان التوزيع الاحتمالي للمتغيرة $\hat{\boldsymbol{p}}$ يقترب من التوزيع الطبيعي ونكتب: $\hat{\boldsymbol{p}} \rightsquigarrow \boldsymbol{N}(\mathbf{p}; \sigma_{\hat{\boldsymbol{p}}})$

$$\sigma_{\hat{\boldsymbol{p}}}^2 = \frac{pq}{n} = \frac{(0.35)(0.65)}{100} = 0.002275$$

$$\sigma_{\hat{\boldsymbol{p}}} = 0.04769$$

بما أن حجم العينة كبير فانه وفقا لنظرية النهاية المركزية بعد تحقق الشرطين :

$$\text{np}=n\hat{\text{p}} = (100)(0.7) = 70 > 5 \text{ , } n(1-\hat{\text{p}}) = \text{n}(1-\text{p})=(100)(0.3) = 30 > 5$$

فان التوزيع الاحتمالي للمتغيرة $\hat{\boldsymbol{p}}$ يقترب من التوزيع الطبيعي ونكتب: $\hat{\boldsymbol{p}} \rightsquigarrow \boldsymbol{N}(\mathbf{p}; \sigma_{\hat{\boldsymbol{p}}})$

بما أن :

$$\frac{n}{N} = \frac{100}{1200} = 0.083 > 0.05$$

فان المعاينة تعتبر بالإرجاع و بالتالي :

$$\sigma_{\hat{\boldsymbol{p}}} = \sqrt{\frac{pq}{n}\left(\frac{N-n}{N-1}\right)} = \sqrt{\frac{(0.7)(0.3)}{100}\left(\frac{1200-100}{1200-1}\right)} = 0.0438$$

$$Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$$

$$Pr\left(\hat{p} - \sigma_{\hat{\boldsymbol{p}}} Z_{\frac{\alpha}{2}} \le p \le \hat{p} + \sigma_{\hat{\boldsymbol{p}}} Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$p \in 0.7 \mp (0.0438)(1.96)$$

$$IC_{0.95}[0.6141\ ;\ 0.7858]$$

بمستوى ثقة 95% هناك ما بين 78.58% و 61.41% يفضلون تسيير تقاعدهم ذاتيا

التمرين رقم 07

تم سحب عينتين عشوائيتين من مصنعين A و B لانتاج الانابيب البلاستيكية الخاصة بالمياه (قطر الأنابيب يقاس بالمليمتر) والبيانات المتحصل عليها ملخصة في ما يلي :

المصنع A : $n_1 = 5$ بيانات العينة : 63.12 ، 63.57 ، 62.81 ، 64.32 ، 63.76

المصنع B : $n_2 = 4$ بيانات العينة : 62.51 ، 63.24 ، 62.31 ، 62.21

المطلوب: عند مستوى ثقة 95% مع العلم ان قياس الأنابيب في المصنعين يتبع التوزيع الطبيعي بتباين مجهول ومتساوي، أوجد مجال الثقة للفرق بين متوسطي طول قطر الأنابيب في المصنعين السابقين ؟

الحل :

المصنع B	المصنع A
$n_2 = 4$ $\bar{x}_1 = 62.5675$ $\sum_{i=1}^{4}(x_{2i}-\bar{x}_2)^2 = 0.6498$ $S_2^2 = \frac{\sum_{i=1}^{4}(x_{1i}-\bar{x}_1)^2}{n_1-1} = \frac{1.3638}{4-1} = 0.2166$	$n_1 = 5$ $\bar{x}_1 = 63.516$ $\sum_{i=1}^{5}(x_{1i}-\bar{x}_1)^2 = 1.3638$ $S_1^2 = \frac{\sum_{i=1}^{5}(x_{1i}-\bar{x}_1)^2}{n_1-1} = \frac{1.3638}{5-1} = 0.3409$

عند درجة حرية $v = n_1+n_2-2$ تعطى إحصائية توزيع ستيودنت **T** على النحو التالي:

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}} \rightsquigarrow t(n_1+n_2-2)$$

$$Pr\left((\bar{X}_1-\bar{X}_2) - T_{\frac{\alpha}{2},n_1+n_2-2}\sigma_{\bar{X}_1-\bar{X}_2} \le (\mu_1-\mu_2) \le (\bar{X}_1-\bar{X}_2)+T_{\frac{\alpha}{2},n_1+n_2-2}\sigma_{\bar{X}_1-\bar{X}_2}\right) = 1-\alpha$$

بما أن التباين في المجتمعين مجهول ومتساوي فان التباين المرجح (التباين المشترك) يعطى بالعلاقة :

$$\hat{\sigma}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{(5-1)0.3409 + (4-1)0.2166}{5+4-2} = 0.2876$$

تعطى صيغة $\sigma^2_{\bar{X}_1-\bar{X}_2}$ على الشكل :

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right) = 0.2876\left(\frac{1}{5}+\frac{1}{4}\right) = 0.1294$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = 0.3597$$

$$T_{\frac{\alpha}{2},n_1+n_2-2} = 2.365$$

$$Pr\left((\bar{X}_1-\bar{X}_2) - T_{(\frac{\alpha}{2},n_1+n_2-2)}\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \le (\mu_1-\mu_2) \le (\bar{X}_1-\bar{X}_2)+T_{(\frac{\alpha}{2},n_1+n_2-2)}\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)$$
$$= 1-\alpha$$

$$(\mu_1-\mu_2) \in 63.516 - 62.5675 \mp (2.365)(0.3597)$$

$$(\mu_1-\mu_2) \in 0.9485 \mp (2.365)(0.8506)$$

$$IC_{0.95}[0.0799\ ;\ 1.7991]$$

التمرين رقم 08

أسفرت نتائج سبر الآراء لمؤسسة استطلاع حول الانتخابات الرئاسية لإحدى الدول لعينة عشوائية مكونة من 900 شخص أن المترشح A سيكسب في الدوائر الانتخابية بـ 468 صوت من أصوات الناخبين في العينة المختارة.

1. أوجد التقدر النقطي لنسبة الناخبين المصوتين للمرشح A ؟
2. أوجد مجال الثقة لنسبة الناخبين المصوتين للمرشح A عند مستوى ثقة 95% ؟
3. أوجد هامش خطأ التقدير $|\hat{p}-p|$ عند مستوى ثقة 95 % ؟
4. كم يجب أن يكون حجم العينة الأدنى اذا أردنا فوز المرشح A بالانتخابات الرئاسية بنسبة أكثر من 50% من الأصوات (مستوى ثقة 95%) ؟

الحل :

1. التقدير النقطي لنسبة الناخبين المصوتين للمرشح A :

$$\hat{P}=\frac{468}{900}=0.52$$

52% من الناخبين مقتنع أن يصوت لصالح المرشح A

2. مجال الثقة لنسبة الناخبين المصوتين للمرشح A عند مستوى ثقة 95% ؟

np= (900)(0.52)= 468>5

n(1-p)=(900)(0.48)= 432>5

بما أن الشرط أعلاه محقق فان حجم العينة n كبير كفاية وبالتالي يمكن تقريب توزيع ذو الحدين الى التوزيع الطبيعي :

$$\hat{\boldsymbol{p}} \rightsquigarrow \boldsymbol{N}(\mathbf{p};\sigma_{\hat{\boldsymbol{p}}}^2)$$

$$Z=\frac{\hat{\boldsymbol{p}}-p}{\sigma_{\hat{\boldsymbol{p}}}}$$

$$\sigma_{\hat{\boldsymbol{p}}}=\sqrt{\frac{pq}{n}}=\sqrt{\frac{(0.52)(0.48)}{900}}=0.016$$

$$Pr\left(\hat{p}-\sigma_{\hat{\boldsymbol{p}}}Z_{1-\frac{\alpha}{2}}\leq p\leq \hat{p}+\sigma_{\hat{\boldsymbol{p}}}Z_{1-\frac{\alpha}{2}}\right)=1-\alpha$$

$$Pr(0.52-(0.016)(1.96)\leq p\leq 0.52+(0.016)(1.96))=1-0.05$$

$$Pr(00.487\leq p\leq 0.552)=0.95$$

3. أوجد هامش خطأ التقدير $|p-\hat{p}|$ عند مستوى ثقة 95 % ؟

$$|p-\hat{p}|=\sigma_{\hat{\boldsymbol{p}}}Z_{\frac{\alpha}{2}}=(0.016)(1.96)=0.032$$

4. ايجاد حجم العينة الأدنى حتى يفوز المرشح A بالانتخابات بنسبة أكثر من 50% من الأصوات (عند مستوى ثقة 95%)

$$p \in \hat{p} \mp \sigma_{\hat{p}} Z_{1-\frac{\alpha}{2}}$$

$$|\hat{p} - p| = \sigma_{\hat{p}} Z_{1-\frac{\alpha}{2}}$$

$$|\hat{p} - p|^2 = \sigma_{\hat{p}}^2 Z_{1-\frac{\alpha}{2}}{}^2$$

$$(\hat{p} - p)^2 = \frac{pq}{n}(Z_{1-\frac{\alpha}{2}})^2$$

$$n = \frac{pq(Z_{1-\frac{\alpha}{2}})^2}{(\hat{p} - p)^2}$$

$$n = \frac{(0.5)(0.5)(1.96)^2}{(0.5 - 0.52)^2} = 2497.15$$

$$n \cong 2497$$

حتى يفوز المرشح A بالانتخابات بنسبة أكثر من 50% من الأصوات عند مستوى ثقة 95% فانه يجب أن يكون حجم العينة على الأقل **2497**

التمرين رقم 09

تمثل البيانات التالية تطور المبيعات لعشر مؤسسات مختارة عشوائيا تنتج نفس المنتوج خلال سنة معينة

116	142	200	318	142	146	160	294	352	360

المطلوب : بافتراض أن بيانات المجتمع موزعة توزيعا طبيعيا

عند مستوى 95% حدد مجال الثقة لتباين المجتمع σ^2 ؟

الحل:

من خلال البيانات اعلاه نحسب كل من المتوسط الحسابي والانحراف المعياري:

$$s = 96.9 \text{ و } \bar{X} = 223$$

من خلال جدول التوزيع الاحتمالي لمربع كاي نجد: $\chi^2_{\alpha/2} = \chi^2_{0.25} = 19.023$ ، $\chi^2_{1-(\frac{\alpha}{2})} = \chi^2_{0.975} = 2.7$

مجال الثقة عند مستوى 95% يعطى على الشكل التالي :

$$Pr\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-(\frac{\alpha}{2})}}\right) = 1 - \alpha$$

$$\sigma^2 \in \left[\frac{(9)(96.9)^2}{19.023} ; \frac{(9)(96.9)^2}{2.7}\right]$$

$$IC_{0.95}[4.442 \quad ; \ 31.299]$$

التمرين رقم 10

لتكن X متغيرة عشوائية تتبع التوزيع الطبيعي $X \rightsquigarrow N(40;\sigma^2)$، من أجل تقدير تباين المجتمع نختار عينة عشوائية بحجم $n=25$ حيث تم التوصل من خلال هذه العينة : $\sum_{i=1}^{n}(X_i-\mu)^2 = 300$.

المطلوب : حدد مجال الثقة لتباين المجتمع عند مستوى ثقة 95% ؟

الحل :

متوسط الحسابي في المجتمع معلوم $\mu = 40$ اذا : $\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \rightsquigarrow \chi^2_{(n)}$

يعطى مجال الثقة على الشكل التالي :

$$Pr\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2,(n)}} \le \sigma^2 \le \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2,(n)}}\right) = 1-\alpha$$

$$\chi^2_{\alpha/2,(n)} = \chi^2_{0.025,(25)} = 40.644$$

$$\chi^2_{1-\alpha/2,(n)} = \chi^2_{0.975,(25)} = 13.12$$

$$IC_{0.95}\left[\frac{300}{40.644} \quad ; \ \frac{300}{13.12}\right]$$

$$IC_{0.95}[7.381 \quad ; \ 22.866]$$

باحتمال 0.95 فان مجال الثقة $[7.381 \ ; \ 22.866]$ يحتوي على القيمة الحقيقية لتباين المجتمع σ^2

الفصل الثالث: اختبار الفرضيات

1/ *تعريف الفرضية الإحصائية*

2/ *تعريف اختبار الفرضيات*

3/ *اختبار الفرضيات للمتوسط الحسابي في حالة مجتمع احصائي طبيعي*

1.3/ في حالة σ^2 معلومة

2.3/ في *حالة* σ^2 *مجهولة*

4/ اختبار الفرضيات للفرق بين متوسطي مجتمعين مستقلين

1.4/ في *حالة تباين المجتمعين معلوم*

2.4/*في حالة تباين المجتمعين مجهول*

5/اختبار الفرضيات للنسبة في المجتمع

6/اختبار الفرضيات للفرق بين نسبتي مجتمعين مستقلين

7/اختبار الفرضيات للتباين في المجتمع

1.7/ *حالة متوسط المجتمع* μ *غير معلوم*

2.7/ *حالة متوسط المجتمع* μ *معلوم*

تمهيد

في الاحصاء الاستنباطي لدينا التقدير و اختبار الفرضيات، يستخدم أسلوب التقدير عند تقدير معلمة من معالم المجتمع المجهولة القيمة بحيث تستخدم بيانات العينة لتقدير هذه القيمة ، أما أسلوب اختبار الفرضيات فيستخدم في حالة ما نكون علم بقيمة هذه المعلمة ونريد التأكد من قيمتها عن طريق اختبار الفرضيات بحيث يتم استخدام بيانات العينة .

يلعب اختبار الفرضيات دورا مهما في تطبيق الإحصائيات على مشاكل الحياة الواقعية، بحيث تستخدم البيانات المأخوذة من العينات لاتخاذ قرارات بشأن التوزيع الاحتمالي غير المعروف للمجتمع أو معلماته.

1/ تعريف الفرضية الإحصائية

الفرضية الإحصائية هي بيان أو صيغة تتعلق بالتوزيع الاحتمالي لمتغير عشوائي أو معلمات للمجتمع. فعلى سبيل المثال. تتمثل إحدى المشكلات الصناعية المهمة في قبول أو رفض الكثير من المنتجات المصنعة، عادة تقوم الشركة المصنعة قبل الإفراج عن كل طلبية كبيرة الحجم للمستهلك ببعض الاختبارات لتحديد ما إذا كانت محتويات الطلبية تتوافق مع المعايير المقبولة. فاذا كانت نسبة العيوب في الطلبية أقل من أو تساوي رقمًا معينًا فسيتم تحرير الطلبية. ففي كثير من الأحيان بدلاً من اختبار كل عنصر في الطلبية، قد نختبر فقط عددًا قليلا من العناصر المختارة عشوائيا من الدفعة ونتخذ قرارات بشأن نسبة العيوب في الدفعة، أي أننا نتخذ القرارات بشأن المجتمع الاحصائي على أساس معلومات العينة. تسمى هذه القرارات بالقرارات الإحصائية. في محاولة للوصول إلى القرارات ، من المفيد إجراء بعض التخمينات الأولية حول المجتمع الاحصائي. تسمى هذه التخمينات بالفرضيات الإحصائية. في بعض الأحيان قد تكون نتائج العينة مختلفة بشكل ملحوظ عن تلك المتوقعة في ظل الفرضية.

يمكن القول أن الفرضية الاحصائية هي كل عبارة تكون صحتها أو عدم صحتها يحتاج الى قرار ويرمز للفرضية بالرمز H وهناك فرضيتان وهما **الفرضية الصفرية** (فرضية العدم) و **الفرضية البديلة،** يرمز لفرضية العدم عادة بالرمز H_0 بحيث يتم اختبار امكانية رفضها على اعتبار أنها صحيحة، أما **الفرضية البديلة** و يرمز لها بالرمز H_1 وهي فرضية مكملة لفرضية العدم حيث يتم قبولها عند رفض فرضية العدم أو رفضها عند قبول فرضية العدم وتصاغ الفرضية الصفرية و الفرضية البديلة بأحد الطرق التالية:

$$\begin{cases} H_0 : \mu \geq \mu_0 \\ H_1 : \mu \prec \mu_0 \end{cases} \qquad \begin{cases} H_0 : \mu \leq \mu_0 \\ H_1 : \mu \succ \mu_0 \end{cases} \qquad \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

تشير الفرضية البديلة الى جهة الاختبار ما اذا كان من جهتين أو جهة واحدة (من اليمين أو من اليسار) فمثلا $H_1 : \mu \succ \mu_0$ تعني أن متوسط المجتمع أكبر من μ_0 أما الفرضية $H_1 : \mu \prec \mu_0$ فتعني أن متوسط المجتمع أصغر من μ_0 وهما تعنيان أن الاختبار أحادي الجانب (من جهة واحدة) من اليمين ومن اليسار على التوالي.

يتطلب تحديد فرضية العدم والفرضية البديلة تحديد مستوى المخاطر التي يمكن القبول بها عند اتخاذ القرار الخاطئ ويعبر عن هذه المخاطر بمستوى المعنوية α وهذا الأخير ما هو الا درجة الاحتمال التي نرفض بها فرضية العدم H_0 عندما تكون صحيحة،

نظرا لأن القرار الاحصائي يعتمد على معطيات العينة ، فنحن عرضة لارتكاب أخطاء في الاختبار الإحصائي فمن المستحيل إثبات صحة الفرضية بنسبة 100٪ من اليقين. هناك نوعان محتملان من الأخطاء.

الجدول رقم(02) : القرار الاحصائي واحتمال الخطأ

القرار الاحصائي	**الحالة الحقيقية لفرضية العدم**	
	H_0 صحيحة	H_0 خاطئة
عدم رفض فرضية H_0	القرار صحيح	الخطأ من النوع الثاني β
رفض فرضية H_0	الخطأ من النوع الأول α	القرار الصحيح

ان اتخاذ القرار الخاص بالفرضيات يقودنا الى نوعين من الأخطاء :

الخطأ من النوع الأول α: يحدث هذا الخطأ عند رفض فرضية العدم عندما تكون في الواقع صحيحة، احتمال الخطأ الأول يسمى مستوى المعنوية (حد المعنوية) وهذا الخطأ ماهو الا عبارة عن مستوى المعنوية α (مستوى الدلالة) والذي يمثل احتمال الخطأ أو المخاطر التي يمكن القبول بها عند اتخاذ القرار الخاطئ حيث يحدد هذا الاحتمال من طرف الباحث، في حالة الاختبار ثنائي الاتجاه تقسم قيمتها الى نصفين أما في حالة اختبار أحادي الاتجاه فتأخذ قيمتها كاملة.

الخطأ من النوع الثاني β: ويحدث عند قبول فرضية العدم عندما تكون خاطئة (بمعنى عندما تكون الفرضية البديلة هي المقبولة)، حيث هناك علاقة عكسية بين حجم الخطأ من النوع الأول وحجم الخطأ من النوع الثاني وللتقليل من كليهما ينبغي زيادة حجم عينة الدراسة.

المنطقة الحرجة (منطقة رفض فرضية العدم) : تعبر عن المساحة الى تقع أسفل منحنى التوزيع الاحتمالي المستخدم (التوزيع الطبيعي أو ستيودنت...) المستخدم في عملية التحليل الاحصائي ويمثل احتمال رفض فرضية العدم H_0 بمعنى تمثل مستوى أو حد المعنوية α وتسمى هذه المنطقة **بمنطقة رفض فرضية العدم**، أما باقي المساحة تحت المنحنى فتمثل **منطقة القبول** بحيث تفصل منطقة القبولة عن منطقة الرفض بقيم التوزيع الاحتمالي المستخدم (الطبيعي أو ستيودنت ..) وتسمى هذه **القيم بالقيم الحرجة** .

2/ تعريف اختبار الفرضيات

الإجراءات التي تمكننا من اتخاذ قرار بشأن قبول أو رفض الفرضيات أو تحديد ما إذا كانت العينات التي تمت ملاحظتها تختلف اختلافًا كبيرًا عن النتائج المتوقعة تسمى اختبارات الفرضيات.

اذا فاختبار الفرضيات هو منهج يؤدي الى تطوير قاعدة قرار تتيح الاختيار بين فرضيتين احصائيتين فرضية العدم و الفرضية البديلة، فاتخاذ قرار ما حول ما اذا كانت الفرضية الصفرية مقبولة أم مرفوضة يتم عن طريق اختبار احصائي وهو عبارة عن متغير عشوائي ذو توزيع احتمالي يصف العلاقة بين القيم النظرية للمعلمة والقيم المحسوبة من العينة، ويتم اتخاذ القرار المتعلق بقبول أو رفض الفرضية الصفرية، وبعد مقارنة الاختبار الاحصائي المحسوب من العينة مع قيمته المجدولة (الحرجة) المستخرجة من جداول خاصة نتمكن من اتخاذ القرار المناسب .

وقد يستخدم اختبار الفرضيات عندما تكون هناك بعض المعلومات المتوفرة عن المجتمع ولكن يود التأكد من دقتها فقد تكون لدى الباحث معرفة بمتوسط المجتمع من دراسات سابقة ولكن بعض التغييرات التي حدثت للمجتمع في وقت لاحق قد غيرت من قيمة متوسط المجتمع ولهذا فمن أجل التأكد من حدوث هذا التغير في المتوسط من عدمه فانه يتم اختبار قيمة المتوسط في المرحلة الجديدة لمعرفة ما اذا كان هو نفسه المتوسط أم قد تغير.

3/ اختبار الفرضيات للمتوسط الحسابي في حالة مجتمع طبيعي

1.3/ اذا كانت σ^2 معلومة

✓ **حالة الاختبار الثنائي الاتجاه** : $H_1 : \mu \neq \mu_0$ vs $H_0 : \mu = \mu_0$

تحت فرضية العدم H_0 يكون لدينا :

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

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

عند حد المعنوية α تحدد منطقة رفض فرضية العدم H_0 كما يلي :

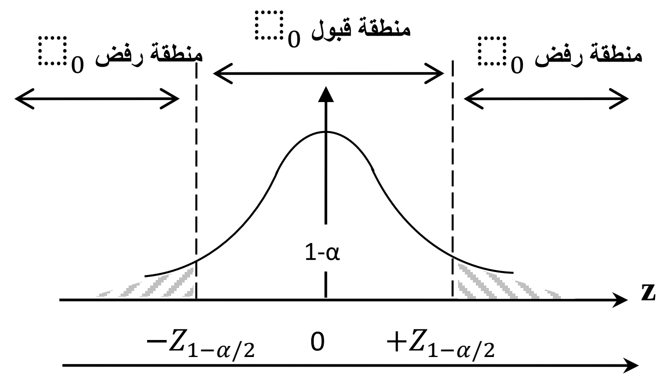
$$Z \notin \left[-Z_{1-\alpha/2}, Z_{1-\alpha/2}\right]$$

أو بصيغة أخرى *فاننا نرفض فرضية العدم* H_0*ونقبل الفرضية البديلة* H_1 اذا كانت $|Z_{cal}| > Z_{1-\alpha/2}$

مثلما يوضحه الشكل الموالي :

عند حد المعنوية α (على سبيل المثال 0.05) يمكن إيجاد إحصائية التوزيع الطبيعي لمجدولة(الحرجة) $Z_{1-\alpha/2}$ من خلال جدول التوزيع الاحتمالي للقانون الطبيعي .

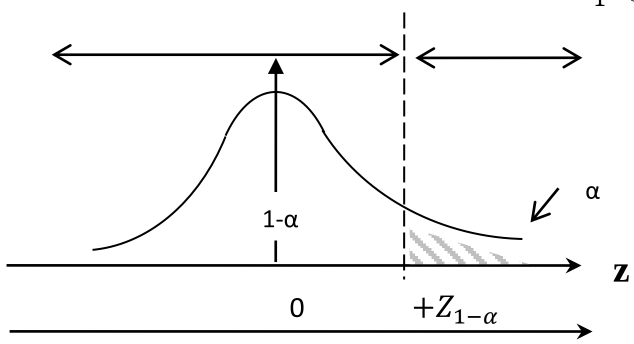
الشكل رقم (02) : منطقة القبول والرفض في حالة الاختبار الثنائي



✓ **حالة اختبار أحادي الاتجاه من اليمين** : $H_0 : \mu \leq \mu_0$ VS $H_1 : \mu \succ \mu_0$

نرفض فرضية العدم H_0 ونقبل الفرضية البديلة اذا كان : H_1 $Z_{1-\alpha} < Z_{cal}$ كما يوضحه الشكل الموالي:

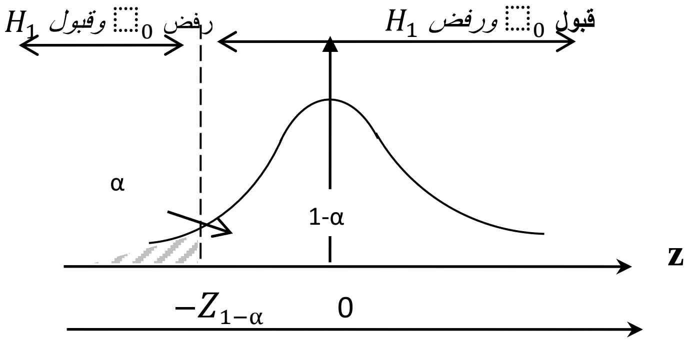
الشكل رقم(03) : منطقة القبول والرفض في حالة الاختبار أحادي الجانب من اليمين



✓ **حالة اختبار أحادي الاتجاه من اليسار :** $H_0 : \mu \geq \mu_0$ VS $H_1 : \mu \prec \mu_0$

نرفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 اذا كان : $-Z_\alpha > Z_{cal}$ كما يوضحه الشكل الموالي :

الشكل رقم (*04*) : منطقة القبول والرفض في حالة الاختبار أحادي الجانب من اليسار



مثال01

تم اختيار عينة عشوائية من 100 كيس اسمنت حيث وجد أن متوسط وزن الاسمنت يساوي49.5 كلغ

المطلوب: عند درجة ثقة 95% اختبر ما اذا كان فعلا متوسط وزن كيس الاسمنت في المصنع يساوي 50 كلغ أم أنه يختلف عن هذه القيمة، مع العلم أن تباين وزن كيس الاسمنت في المصنع يساوي25 كلغ

الحل :

نبدأ أولا بصياغة الفرضيات كما يلي :

$$H_0 : \mu = 50 \quad \text{vs} \quad H_1 : \mu \neq 50$$

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

$$Z_{cal} = \frac{49.8 - 50}{5/\sqrt{100}} = -0.4$$

عند حد *المعنوية* $\alpha = 0.05$ نجد احصائية التوزيع الطبيعي المجدولة $Z_{1-\alpha/2} = Z_{1-0.05/2} =$
$Z_{0.975} = 1.96$

القرار الاحصائي: نلاحظ أن قيمة احصائية التوزيع الطبيعي المحسوبة $Z_{cal} = -0.4$ تقع في منطقة قبول فرضية

العدم $H_0 : \mu = 50$ او بعبارة أخرى نقبل فرضية العدم كون أن $Z_{0.975} = 1.96 > |Z_{cal}| = 0.4$
$Z \in [-1.96, 1.96]$

2.3/ *حالة* σ^2 *مجهولة*

في الحالة العامة والأكثر واقعية لما يكون التباين $\boldsymbol{\sigma^2}$ في المجتمع مجهول

فانه يتم استخدام قيمة توزيع اختبار ستيودنت بدرجة حرية (n-1) حيث: $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \longrightarrow t(n-1)$

حالة الاختبار الثنائي الاتجاه : $H_1 : \mu \neq \mu_0$ vs $H_0 : \mu = \mu_0$

$$Pr\left(T_{1-\alpha/2}^{(n-1)} \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq T_{1-\alpha/2}^{n-1}\right) = 1 - \alpha$$

عند حد المعنوية تحدد منطقة رفض فرضية العدم H_0 كما يلي :

$$T \notin \left[-T_{1-\frac{\alpha}{2}}^{(n-1)}, T_{1-\alpha/2}^{n-1}\right]$$

أو بصيغة أخرى *فاننا نرفض فرضية العدم* H_0 *ونقبل الفرضية البديلة اذا كانت احصائية ستيودنت المحسوبة بالقيمة المطلقة أكبر من احصائية ستيودنت المجدولة* $T_{1-\alpha/2}^{n-1} < |T_{cal}|$

- ✓ **حالة اختبار أحادي الاتجاه من اليمين :** vs $H_0 : \mu \leq \mu_0$ $H_1 : \mu \succ \mu_0$

نرفض فرضية العدم اذا كان : $T_{1-\frac{\alpha}{2}}^{(n-1)} < T_{cal} = \frac{\bar{X}-\mu}{S/\sqrt{n}}$

- ✓ **حالة اختبار أحادي الاتجاه من اليسار :** $H_1 : \mu \prec \mu_0$ $H_0 : \mu \geq \mu_0$

نرفض فرضية العدم اذا كان : $-T_{\frac{\alpha}{2}}^{(n-1)} > T_{cal} = \frac{\bar{X}-\mu}{S/\sqrt{n}}$

ملاحظة :

- ✓ عند حد المعنوية α (على سبيل المثال 0.05) ودرجة الحرية V=n-1 يمكن إيجاد إحصائية توزيع ستيودنت المجدولة(الحرجة) حيث نقارن القيمة المحسوبة مع القيمة المجدولة لتوزيع ستيودنت .

✓ التوزيع الاحتمالي لتوزيع ستيودنت مثلما هو عليه التوزيع الطبيعي هو توزيع متناظر لذلك نأخذ القيمة المحسوبة انطلاقا من عينة الدراسة بالقيمة المطلقة.

مثال02:

تريد مؤسسة انتاجية التأكد من أن علب مادة التنظيف التي تنتجها وزنها هو 500g من المنظف، ومن تجارب سابقة تبين أن متغير وزن العلب يتبع التوزيع الطبيعي، ولتحقيق هذا الغرض اختار أحد خبراء هذه المؤسسة عينة عشوائية من العلب n=9 ووجد أن $\bar{X} = 520$, S=75g

المطلوب: عند حد المعنوية $\alpha = 5\%$ و باعتبار أن المجتمع الاحصائي غير محدود ، اختبر الفرضيات التالية مع توضيح منطقة القبول والرفض لهذه الفرضيات: $\boldsymbol{H_0: \mu \geq 500\ g}$, $\boldsymbol{H_1 : \mu < 500\ g}$ ؟

الحل :

بما أن البيانات في المجتمع تتبع التوزيع الطبيعي والتباين مجهول في المجتمع فان توزيع المعاينة للمتوسط الحسابي يتبع توزيع ستيودنت بدرجة حرية (n−1) بحيث :

$$\sigma_{\bar{x}} = {\sigma}/{\sqrt{n}} = {s}/{\sqrt{n}} = {75}/{\sqrt{9}} = 25$$

تعطى إحصائية توزيع ستيودنت كما يلي :

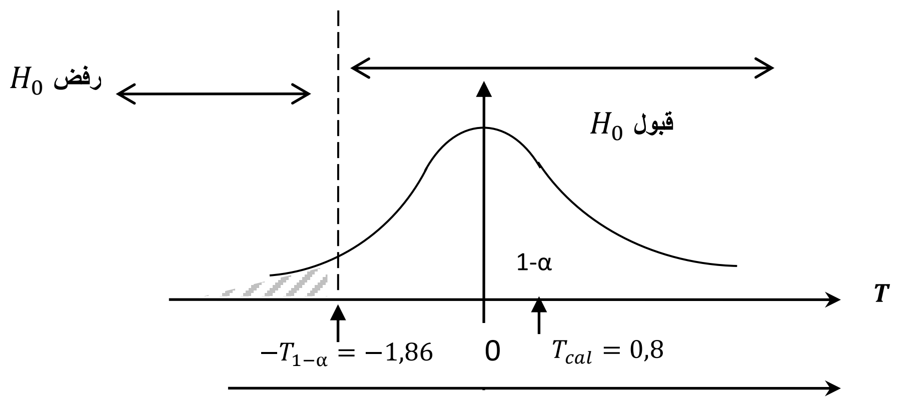
$$\mathrm{T} = \frac{\bar{\mathrm{x}} - \mu}{\sigma_{\bar{x}}} = \frac{520 - 500}{25} = 0{,}8$$

عند حد المعنوية $\alpha = 5\%$ ودرجة الحرية $V=n-1=8$ نجد إحصائية ستيودنت المجدولة $T_{\text{tab}} = \mathrm{T}_{n-1}^{1-\alpha} = -1{,}86$

هذا الاختبار هو اختبار ثنائي الاتجاه و بالمقارنة نجد : $T_{\text{tab}} = -1{,}86 < T_{cal} = 0{,}8$

وبالتالي نقبل فرضية العدم و نرفض الفرضية البديلة .

منطقة القبول والرفض في حالة الاختبار أحادي الجانب من اليسار



4/ اختبار الفرضيات للفرق بين متوسطي مجتمعين مستقلين

لاختبار الفرق في المتوسط الحسابي بين مجتمعين يمكن اتباع نفس الطريقة لاختبار فرضيات المتوسط الحسابي لمجتمع واحد، اختبار الفرق بين متوسطي مجتمعين يتم من خلال سحب عينة عشوائية من كل مجتمع احصائي، ويهدف هذا الاختبار لمعرفة هل يعود الفرق بين متوسطي المجتمعين هو حقيقي أم لا

. فاذا كان لدينا متوسط العينة الأولى $\bar{X}_1$ ومتوسط العينة الثانية $\bar{x}_2$ بتباين $\sigma^2_{\bar{X}_1}$, $\sigma^2_{\bar{X}_2}$ على الترتيب ، وكان حجم العينة الأولى n_1 وحجم العينة الثانية n_2 فان الفرضيات المراد اختبارها تكون على الشكل التالي :

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

التوقع الرياضي للفرق بين المتوسطين:

$$E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$$

تباين الفرق بين المتوسطين :

$$Var(\bar{X}_1 - \bar{X}_2) = Var(\bar{X}_1) + Var(\bar{X}_2) - 2Cov(\bar{X}_1, \bar{X}_2)$$

في حالة العينات المستقلة فان : $Cov(\bar{X}_1, \bar{X}_2) = 0$ و بافتراض أن المجتمعين غير محدودين (أو المعاينة بالإرجاع) :

$$Var(\bar{X}_1 - \bar{X}_2) = \sigma^2_{\bar{X}_1 - \bar{x}_2} = Var(\bar{X}_1) + Var(\bar{X}_1)$$

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

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

1.4/ في حالة تباين المجتمعين معلوم

في حالة توزيع المعاينة للفرق بين المتوسطين يقترب من التوزيع الطبيعي فان إحصائية التوزيع الطبيعي (المتغيرة المعيارية Z تتبع التوزيع الطبيعي الممركز المختصر) تعطى بالعلاقة التالية :

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

القرار : نرفض فرضية العدم و نقبل الفرضية البديلة اذا كان $|Z_{cal}| > Z_{1-\alpha/2}$ ، أما اذا كان $|Z_{cal}| \leq Z_{1-\alpha/2}$ فاننا نقبل فرضية العدم و نرفض الفرضية البديلة

2.4/ في حالة تباين المجتمعين مجهولين

في حالة ما يكون المجتمعين موزعين توزيعا طبيعيا و التباين مجهول ومتساوي في المجتمعين $\boldsymbol{\sigma_1^2} = \boldsymbol{\sigma_2^2} = \widehat{\boldsymbol{S}}^2$ فان الاختبار المستخدم هو اختبار ستيودنت بدرجة حرية $n_1 + n_2 - 2$ والذي تعطى بالعلاقة التالية :

$$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\widehat{\boldsymbol{S}}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightsquigarrow T.Student\,{}^{\alpha/2}_{n_1+n_2-2}$$

حيث :

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

$\widehat{\boldsymbol{\sigma}}^2$: يمثل التباين المشترك (التباين المرجح)

S_1^2, S_2^2 : يمثلان التباين في العينة الأولى و الثانية على التوالي

القرار :

- ✓ اذا كان $|T_{cal}| > T_{\alpha/2}$: نرفض فرضية العدم ونقبل الفرضية البديلة
- ✓ *أما اذا كان* $|T_{cal}| \leq T_{\alpha/2}$ فاننا نقبل فرضية العدم و نرفض الفرضية البديلة

ملاحظة : ان الافتراضات التي يقوم عليها اختبار الفرق بين المتوسطات في حالة مجتمعين مستقلين تشمل ما يلي:

– التوزيع الطبيعي لبيانات المجتمع و اذا كان حجم العينة أكبر من 30 فيمكن الاستغناء عن هذا الشرط.

– العشوائية في اختيار كل عينة واستقلالية العينتين

مثال03:

مؤسسة انتاجية تستهلك مصابيح كهربائية (تعتبر زبون مهم) لنوعين من العلامات، عند مستوى معنوية $\alpha = 5\%$ تم اختيار 100 مصباح من كل علامة تجارية وقد ثبت من أن متوسط مدة صلاحية المصباح المنتج من العلامة الأولى هو $\bar{X}_1 = 980\ h$ و $S_1 = 80\ h$ ، وبالنسبة للعلامة الثانية $\bar{X}_2 = 1010\ h$ ، $S_2 = 120\ h$.

المطلوب : عند حد المعنوية $\alpha = 5\%$ و من خلال المعطيات السابقة بين هل يوجد اختلاف معنوي احصائيا بين مدة صلاحية المصابيح لكل علامة تجارية بمعنى اختبر الفرضية التالية : $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$ ؟

الحل :

بما حجم العينتين أكبر من 30 فوفقا لنظرية النهاية المركزية المحدودة فان توزيع المعاينة للفرق بين المتوسطات يقترب من التوزيع الطبيعي ونكتب :

$$(\bar{X}_1 - \bar{X}_2) \rightsquigarrow N\left(\mu_1 - \mu_2, \sigma_{\bar{x}_1 - \bar{x}_2}\right)$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} \rightsquigarrow N(0,1)$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1 - \bar{x}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$\sigma^2_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{80^2}{100} + \frac{120^2}{100}}$$

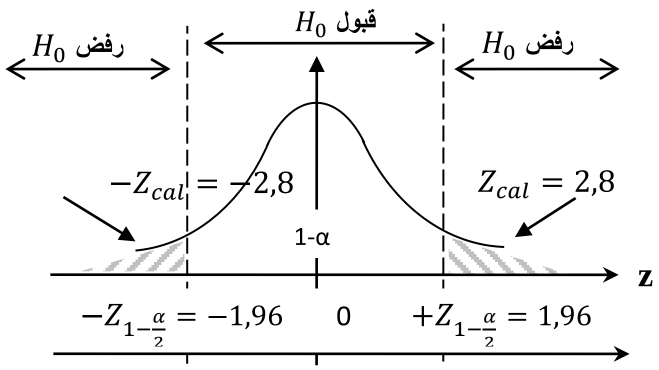
$$\sigma^2_{\bar{x}_1 - \bar{x}_2} = 14{,}42$$

$$Z = \frac{(980 - 1010) - 0}{14{,}42} = -2{,}8$$

بما أن الاختبار ثنائي الاتجاه فانه عند حد المعنوية $\alpha = 0.05$ تعطى إحصائية التوزيع الطبيعي الحرجة (المجدولة) $Z_{tab} = Z_{1-\alpha/2} = 1.96$

بالمقارنة نجد : $Z_{\text{tab}} = 1.96 < |Z_{cal}| = 2{,}8$ وبالتالي نرفض فرضية العدم و نقبل الفرضية البديلة .

منطقة القبول والرفض في حالة الاختبار الثنائي الاتجاه



<u>مثال 04:</u>

أجريت تجربة لمقارنة متوسط أجور الموظفين الذكور والاناث في أحد المؤسسات الخاصة ، ولهذا تم اختيار عينة عشوائية من الموظفين الذكور حجمها $n_1 = 11$ و عينة ثانية عشوائية من الاناث حجمها $n_2 = 16$ حيث أن المتوسط الحسابي للأجور في العينة الأولى هو $\bar{x}_1 = 33000$ و متوسط الأجور في العينة الثانية $\bar{x}_2 = 32500$ ، بانحراف معياري على الترتيب $S_2 = 80,\ S_1 = 100$.

<u>المطلوب</u> : عند مستوى معنوية $\alpha = 0.05$ هل يمكن القول أن متوسط أجور الموظفين الذكور لا يختلف معنويا عن أجور الموظفين الاناث أم لا ؟

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

حجم العينتين أقل من 30 و بالتالي فان الفرق بين المتوسطات يتبع توزيع ستيودنت وتعطى إحصائية هذا التوزيع:

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\widehat{S}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightsquigarrow T_{n_1+n_2-2}^{1-\alpha/2}$$

نقوم بحساب $\widehat{\sigma}^2$ **كما يلي :**

$$\widehat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(11 - 1)100^2 + (16 - 1)80^2}{11 + 16 - 2} = 7840$$

$$\boldsymbol{\widehat{\sigma}^2 = 7840}$$

$$T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\widehat{\boldsymbol{S}}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}=\frac{(33000-32500)-0}{\sqrt{7840\left(\frac{1}{11}+\frac{1}{16}\right)}}=14.41$$

عند حد المعنوية $\alpha=0.05$ و درجة الحرية $v=n_1+n_2-2$ تعطى إحصائية ستيودنت الحرجة (المجدولة) $T_{n_1+n_2-2}^{1-\alpha/2}=2.060$

بمقارنة احصائية التوزيع الطبيعي المحسوبة مع المجدولة نجد أن $T_{\text{tab}}=2.060<|T_{cal}|=14.41$ وبالتالي نرفض فرضية العدم ونقبل الفرضية البديلة بوجود فرق معنوي احصائيا (اختلاف) بين أجور الموظفين الذكور وأجور الموظفين الاناث.

5/ اختبار الفرضيات للنسبة في المجتمع

تطبيقات هذه الاختبارات عديدة مثلا تستخدم لدراسة خاصية ثنائية في مجتمع احصائي (الصفة المدروسة تكون موجودة/غير موجودة) خاصة لمراقبة نوعية أو مراجعة أن معدل الانتاج المعيب (غير السليم) لا يتجاوز قيمة معطاة.

يمكن استخدام تقريب غوس Gaussienne من ايجاد أجل مجال الثقة حيث شرط الاستخدام تحت فرضية العدم تكون قيمة p معطاة و يمكن العمل بالشرط $np_0 \geq 5$ و $n(1-p_0) \geq 5$ ، وبالتالي $\hat{p}$ نسبة النجاح في العينة ، بحيث تحت فرضية العدم H_0 يكون لدينا :

$$\frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \rightsquigarrow N(0\,;1)$$

في حالة الاختبار الثنائي :

$$H_1 : p \neq p_0 \quad vs \quad H_0 : p = p_0$$

نرفض فرضية العدم اذا كان :

$$\frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \notin \left[-z_{1-\alpha/2}, z_{1-\alpha/2}\right]$$

$\hat{p}$: نسبة النجاح المحققة في العينة

<u>مثال05:</u>

من معلوم في مجتمع احصائي قبل صدور تشريع الزامية الحزام الاجباري في المركبات أنه بلغت نسبة مستعملي حزام الأمان في السيارات 80%، لأغراض بحثية قام باحث باختيار عينة عشوائية n=200 من سائقي السيارات وذلك بعد صدور التشريع بمدة كافية حيث وجد منهم أن 170 سائق يستعمل حزام الأمان .

<u>المطلوب</u> : عند مستوى معنوية $\alpha = 0.05$ اختبر ما اذا كان التشريع قد ساهم في زيادة نسبة مستعملي حزام الامان في هذا المجتمع الاحصائي بمعنى :

$$\begin{cases} H_0 : p \leq 0.8 \\ H_1 : p \succ 0.8 \end{cases}$$

الحل :

لدينا n=200

$$\hat{p} = \frac{170}{200} = 0.85$$

ان هذا الاختبار الاحصائي ذو اتجاه واحد(أحادي الاتجاه) نحو اليمين ، اذا $Z_\alpha = 1.64$

نرفض فرضية العدم اذا كان : $Z_\alpha < Z_{cal}$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.85 - 0.8}{\sqrt{\frac{0.8(1-0.8)}{200}}} = 1.8$$

- ✓ نعلم أنه في حالة اختبار أحادي الجانب من اليمين فاننا :

بما أن : $Z_\alpha = 1.64 < Z = 1.8$

فاننا نرفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 وبالتالي نستنتج أن التشريع باجبارية حزام الأمان قد ساهم في زيادة نسبة مستعملي حزام الأمان في هذا المجتمع الاحصائي عما كان عليه قبل صدور هذا التشريع.

6/ اختبار الفرضيات للفرق بين نسبتي مجتمعين مستقلين

المقارنة بين نسبتين في عينتين مستقلتين يتكرر بكثرة بنفس طريقة المقارنة بين متوسطين، فعلى سبيل المثال يستخدم لدراسة فعالية احد العلاجات مقارنة بعلاج اخر، ، عندما يتم تقييم هذه الفعالية بمعيار ثنائي مثلا شفاء أو عدم شفاء المريض.

فيمكننا التقيد بالفرضية التالية: $H_0 : p_1 - p_2 = 0$

مثال06:

صاحب مؤسسة سيارات لنقل الأشخاص يشتري قطع غيار من مصنعين اثنين A،B لانتاج قطع الغيار حيث يدعي أن في المصنع A نسبة انتاج غير مطابق لمعايير السلامة في هذا المصنع أقل منه تمام في المصنع B و بالتالي فان انتاج المصنع A أفضل من حيث النوعية مقارنة بانتاج المصنع B للتحقق من هذا الادعاء قام خبير احصائي باختيار عينة عشوائية من قطع غيار كل مصنع والجدول التالي يلخص البيانات المتحصل عليها :

	المصنع A	**المصنع B**
حجم العينة المختارة	100	100
عدد قطع الغيار الغير المطابقة لمعايير السلامة	9	5

المطلوب : عند مستوى معنوية $\alpha = 0.05$ اختبر ما اذا كانت هذه النتائج تدعم ادعاءات صاحب هذه المؤسسة موضحا الفرضيات المراد اختبارها؟

الحل :

✓ نحسب أولا نسبة قطع الغيار الغير مطابقة لمعايير السلامة في كل عينة

$$\hat{p}_1 = \frac{X_1}{n_1} = \frac{9}{100} = 0.09$$

$$\hat{p}_1 = \frac{X_2}{n_2} = \frac{5}{100} = 0.05$$

الفرضيات

$$H_1 : p_1 < p_2 \quad vs \quad H_0 : p_1 \geq p_2$$

أو بعبارة أخرى : $H_1 : p_1 - p_2 < 0 \,, \quad vs \quad H_0 \colon p_1 - p_2 \geq 0$

$\boldsymbol{p_2, p_1}$: نسبة قطع الغيار الغير مطابقة لمعايير السلامة في المصنع A والمصنع B على التوالي

ان هذا الاختبار الاحصائي ذو اتجاه واحد(أحادي الاتجاه) نحو اليسار اذا: $Z_\alpha = Z_{0.05} = -1.64$

✓ التوزيع الاحتمالي للفرق بين نسبتي العينتين :

يقترب توزيع ثنائي الحدين للمتغيرة العشوائية $\hat{p}_1 - \hat{p}_2$ من التوزيع الطبيعي وفقا لنظرية النهاية المركزية اذا كان حجم العينة كبير ، ويكون كبيرا كفاية اذا كان :

$n_1 p_1 = n_1 \hat{p}_1 = (80)(0.09) = 9 > 5 \quad n_1(1-\hat{p}_1) = (100)(0.91) = 91 > 5$

$n_2 p_2 = n_2 \hat{p}_2 = (100)(0.05)=5,\ n_2(1-\hat{p}_2) = (100)(0.95) = 95 > 5$

اذا: $(\hat{p}_1 - \hat{p}_2) \rightsquigarrow N(p_1 - p_2; \sigma^2_{(\hat{p}_1-\hat{p}_2)})$

وتعطى صيغة المتغيرة المعيارية **Z** التي تتبع التوزيع الطبيعي المعياري كما يلي :

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sigma_{\hat{p}_1-\hat{p}_2}} \rightsquigarrow N(1;0)$$

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

في حالة اختبار أحادي من اليسار فانه يتم رفض فرضية العدم اذا كان :

نرفض فرضية العدم اذا كان : $-Z_\alpha > Z_{cal}$

$$Z = \frac{(0.09 - 0.05) - 0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

$$Z = \frac{(0.09 - 0.05) - 0}{\sqrt{\frac{0.09(0.91)}{100} + \frac{0.05(0.95)}{100}}} = \frac{0.04}{0.0359} = 1.1142$$

بما أن : $Z_\alpha = -1.64 < Z_{cal} = 1.1142$

وبالتالي فاننا نقبل فرضية العدم H_0 ونرفض الفرضية البديلة H_1 ومنه نستنتج أن **ادعاءات صاحب هذه المؤسسة صحيح** بأن في المصنع A نسبة انتاج غير مطابق لمعايير السلامة في هذا المصنع أقل منه تماما في المصنع B و بالتالي فان انتاج المصنع A أفضل من حيث النوعية مقارنة بانتاج المصنع B .

7/ اختبار الفرضيات لتباين المجتمع

لتكن **x** متغيرة عشوائية في مجتمع احصائي تتبع التوزيع طبيعي$N(\mu,\sigma^2)$

1.7/ حالة متوسط المجتمع μ غير معلوم

$$H_0 : \sigma^2 = \sigma_0^2 \quad vs \quad H_1 : \sigma^2 \neq \sigma_0^2$$

نعلم أن :

$$\frac{(n-1)s^2}{\sigma^2} \rightsquigarrow {\chi^2}_{(n-1)}$$

تحت فرضية العدم H_0 :

$$Pr\left({\chi^2}^{(n-1)}_{1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma^2} \leq {\chi^2}^{(n-1)}_{\alpha/2}\right) = 1-\alpha$$

نرفض فرضية العدم H_0 اذا كانت احصائية الاختبار $\frac{(n-1)s^2}{\sigma_0^2}$ لا تقع في مجال القبول التالي :

$$\left[{\chi^2}^{(n-1)}_{1-\alpha/2}, {\chi^2}^{(n-1)}_{\alpha/2}\right]$$

✓ ***في حالة اختبار أحادي الاتجاه من اليمين*** $H_0 : \sigma^2 \leq \sigma_0^2 \quad vs \quad H_1 : \sigma^2 > \sigma_0^2$

نرفض فرضية العدم اذا كان : ${\chi^2}^{(n-1)}_{\alpha} < \frac{(n-1)s^2}{\sigma_0^2}$

✓ ***في حالة اختبار أحادي الاتجاه من اليسار*** $H_0 : \sigma^2 \geq \sigma_0^2 \quad vs \quad H_1 : \sigma^2 < \sigma_0^2$

نرفض فرضية العدم اذا كان : ${\chi^2}^{(n-1)}_{1-\alpha} > \frac{(n-1)s^2}{\sigma_0^2}$

ملاحظة: هذه الاختبارات مقبولة فقط اذا كان التوزيع الاحتمالي في المجتمع هو التوزيع الطبيعي .

2.7/ حالة متوسط المجتمع معلوم

في هذه الحالة نستخدم الاحصائية :

$$\frac{\sum_{i=1}^{n}(X_i-\boldsymbol{\mu})^2}{\sigma^2} \rightsquigarrow {\chi^2}_{(n)}$$

المتغير العشوائي $\frac{\sum_{i=1}^{n}(X_i-\boldsymbol{\mu})^2}{\sigma^2}$ يتبع توزيع مربع كاي بدرجة حرية n

خطوات الاختبار الاحصائي تتم بنفس الطريقة السابقة

<u>مثال07</u>

تنتج شركة أجزاء محرك آلي حيث من المفترض ألا يزيد تباين قطرها عن 0.002 (الأقطار تقاس بالبوصة). تم اختيار عينة عشوائية من عشرة أجزاء حيث وجد التباين يساوي 0.0003 بوصة.

المطلوب: عند مستوى معنوية 5٪ اختبر الفرضية: $H_0 : \sigma^2 = 0.0002$ مقابل

$H_1 : \sigma^2 > 0.0002$ مع العلم أن قياس أجزاء المحرك تتبع التوزيع الطبيعي .

الحل :

الاختبار الاحصائي المستخدم هو توزيع مربع كاي، من خلال معطيات التمرين فهذا الاختبار هو اختبار أحادي الاتجاه نحو اليسار بحيث أننا نرفض فرضية العدم H_0 اذا كانت احصائية الاختبار $\frac{(n-1)s^2}{\sigma_0^2}$ أكبر من قيمة احصائية مربع كاي المجدولة ${\chi^2}_{\alpha}^{(n-1)}$

- ✓ عند مستوى معنوية $\alpha = 0.05$ نجد احصائية مربع كاي المجدولة:

$$ {\chi^2}_{\alpha}^{(n-1)} = {\chi^2}_{0.05}^{(9)} = 16.919 $$

- ✓ $\frac{(n-1)s^2}{\sigma_0^2} = \frac{(10-1)0.0003}{0.0002} = 13.5$

اذا نجد : $\frac{(n-1)s^2}{\sigma_0^2} = 13.5 < 16.919 = {\chi^2}_{\alpha}^{(n-1)}$ وبالتالي لا نفرض فرضية العدم $H_0 : \sigma^2 = 0.0002$ وبالتالي لا توجد أدلة كافية للإشارة إلى أن σ^2 تتجاوز القيمة 0.0002 عند مستوى المعنوية 5٪.

مثال08:

في مصنع للبطاريات عمر البطارية المنتجة لها انحراف معياري يساوي0.9 سنة، تم اختيار عينة عشوائية من 10 بطاريات من هذا المصنع فوجد أن الانحراف المعياري لها يساوي 1.2 سنة .

المطلوب: في رأيك هل تعتقد أن تباين عمر البطارية المنتجة أكبر من 0.81 بمعنى $\sigma^2 > 9$ أم لا وذلك عند مستوى معنوية 5% .

الحل:

الفرضيات هي: $H_0 : \sigma^2 \leq 0.81 \quad vs \quad H_1 : \sigma^2 > 0.81$

حيث أن المتوسط الحسابي غير معلوم فاننا نرفض فرضية العدم اذا كان :

$$ {\chi^2}_{\alpha}^{(n-1)} < \frac{(n-1)s^2}{\sigma_0^2} $$

$$ \frac{(n-1)s^2}{\sigma_0^2} = \frac{(10-1)1.2^2}{0.81} = 16 $$

عند مستوى معنوية 5% نجد احصائية مربع كاي المجدولة : ${\chi^2}_{\alpha}^{(n-1)} = {\chi^2}_{0.05}^{(10-1)} = 16.92$

اذا نجد أن: $\frac{(n-1)s^2}{\sigma_0^2} = 16 < 16.92 = {\chi^2}_{\alpha}^{(n-1)}$

وبالتالي نقبل فرضية العدم $H_0 : \sigma^2 \leq 0.81$ ونرفض الفرضية البديلة $H_1 : \sigma^2 > 0.81$

تمارين محلولة

التمرين رقم 01

استقبلت مصالح مديرية التجارة بإحدى المدن العديد من الشكاوى حول كمية أحد المواد الغذائية الأساسية وهو الحليب لأحد العلامات التجارية، الزبائن اشتكوا من أن علب الحليب لا تحتوي على الكمية المشار اليها في الغلاف وهو 1 لتر (1 L) حيث معلوم أن الانحراف المعياري لكمية العلب في المجتمع $\sigma = 0,05$ L ، ومن أجل التحقق من هذه الادعاءات قامت مصالح مديرية التجارة باختيار عشوائيا 09 علب حليب من داخل المؤسسة المنتجة حيث وجد في هذه العينة أن المتوسط الحسابي والانحراف المعياري لكمية الحليب لهذه العلب هي على التوالي : $\bar{X} = 0,95\ L$, $S = 0,1$ L

المطلوب : اذا كان معلوما أن كمية علب الحليب تتبع التوزيع الطبيعي في مجتمع الدراسة، اختبر الفرضية الموالية عند مستوى معنوية 5 % مع توضيح منطقة القبول والرفض لهذه الفرضيات:

$$H_0 : \mu = 1\ L \quad , \quad H_1 : \mu \neq 1\ L$$ ؟

الحل :

بما أن حجم العينة أقل من 30 و **البيانات في المجتمع تتبع التوزيع الطبيعي** و **التباين في المجتمع معلوم** فان توزيع المعاينة للمتوسط الحسابي يتبع التوزيع الطبيعي بمعنى $\bar{X} \rightsquigarrow N(\mu, \sigma_{\bar{X}}^2)$

نعلم أن :

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 0,05/\sqrt{25} = 0,01$$

تعطى إحصائية التوزيع الطبيعي كما يلي :

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{0,95 - 1}{0,01} = -5$$

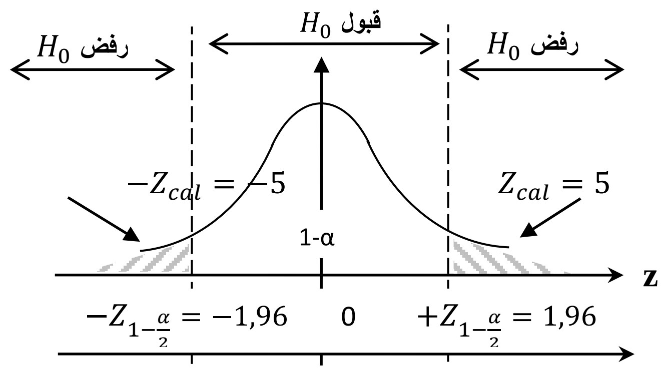
عند حد المعنوية $\alpha = 5\%$ نجد إحصائية التوزيع الطبيعي المجدولة (*الحرجة*)

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

بالمقارنة نجد :

$Z_{tab} = 1.96 < |Z_{cal}| = 5$ وبالتالي نرفض فرضية العدم و نقبل الفرضية البديلة

منطقة القبول والرفض في حالة الاختبار أحادي الجانب من اليمين



التمرين رقم 02

في احدى المؤسسات الانتاجية ، تزن كل قطعة منتجة من احدى المواد الغذائية 480 غ ، ان الزيادة او النقصان في وزن هذه السلعة يعتبر من المشاكل الكبرى التي تواجه هذه المؤسسة أن ذلك يؤدي الى تأثير سلبي على خط انتاج هذه المؤسسة ، من معطيات سابقة معلوم أن الانحراف المعياري لوزن هذا المنتوج في المصنع يساوي 5غ، لتجنب الآثار السلبية على خط الانتاج يقوم مراقب الجودة بمعاينة يومية بصورة عشوائية لـ 40 قطعة من هذا المنتوج ويقرر ما اذا يعلق عمل خط الانتاج من أجل تصحيح أم تواصل عملية الانتاج بصورة عادية.

المطلوب:

1. حدد الفرضيات التي سيعتمد عليها مراقب الجودة في اتخاذ القرار الاحصائي المناسب مع تحديد قاعدة القرار الاحصائي المتعلق بتعليق أو مواصلة خط الانتاج ؟
2. ما هي نصيحتك لمراقب الجودة في حالة ما اذا كان متوسط وزن القطع المنتجة في العينة المختارة يساوي 490غ ثم 479غ برر ذلك بطريقة احصائية ؟

الحل :

1. **الفرضيات التي سيعتمد عليها مراقب الجودة في اتخاذ القرار الاحصائي هي :**

$$H_0 : \mu = 480 \quad vs \quad , \quad H_1 : \mu \neq 480$$

ان هذا الاختبار هو اختبار ثنائي الاتجاه وبالتالي فان قاعدة اتخاذ القرار الاحصائي المناسب هي :

اذا كان: $Z_{1-\frac{\alpha}{2}} = 1.96 < |Z_{cal}|$ وبالتالي نرفض فرضية العدم و نقبل الفرضية البديلة

$$Z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \rightsquigarrow N(1;0)$$

2. النصيحة المقدمة لمراقب الجودة :

✓ في حالة ما اذا كان متوسط وزن القطع المنتجة في العينة المختارة يساوي 490غ

حتى نتخذ القرار المناسب فإننا نحسب احصائية التوزيع الطبيعي انطلاقات من المعطيات السابقة

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{40}} = 0.7905$$

$$Z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{490-480}{0.7905} = 12.6502$$

بما أن : $Z_{1-\frac{\alpha}{2}} = 1.96 < |Z_{cal}| = 12.6502$ وبالتالي فاننا نرفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 بأن متوسط وزن القطع المنتجة في المصنع تختلف عن 480غ وعليه ننصح مراقب الجودة بتعليق الانتاج واعادة تصحيحه.

✓ **في حالة ما اذا كان متوسط وزن القطع المنتجة في العينة المختارة يساوي 479غ**

حتى نتخذ القرار المناسب فاننا نحسب احصائية التوزيع الطبيعي انطلاقات من المعطيات السابقة

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{40}} = 0.7905$$

$$Z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{379-480}{0.7905} = -1.265$$

بما أن: $Z_{1-\frac{\alpha}{2}} = 1.96 > |Z_{cal}| = 1.26$ وبالتالي فاننا نقبل فرضية العدم H_0 ونرفض الفرضية البديلة H_1 وعليه ننصح مراقب الجودة بمواصلة الانتاج .

✓ في حالة ما اذا كان متوسط وزن القطع المنتجة في العينة المختارة يساوي 490غ

التمرين رقم 03

في احدى الجامعات تم اختيار عينة عشوائية من 64 طالب ، فوجد أن متوسط معدل الطلبة هو $\bar{X} = 10$، من المعلوم أن الانحراف المعياري لمعدل الطلبة في في المجتمع $\sigma = 1$، عند مستوى معنوية 0.05 وبافتراض أن البيانات في المجتمع تتبع التوزيع الطبيعي فاختبر الفرضية التالية :

$$\begin{cases} H_0 : \mu = 12 \\ H_1 : \mu \neq 12 \end{cases}$$

$\mu_0 = 12$, n=64 $\bar{x} = 10$

$$\sigma_{\bar{x}} = {}^{\sigma}/_{\sqrt{n}} = {}^{1}/_{\sqrt{64}} = 0.125$$

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{10 - 12}{0.125} = -16$$

بما أن الاختبار ثنائي فانه عند مستوى المعنوية $\alpha = 0.05$ تعطى إحصائية التوزيع الطبيعي المجدولة $Z_{tab} = Z_{1-\alpha/2} = 1.96$ ، و بمقارنة احصائية التوزيع الطبيعي المحسوبة مع المجدولة نجد أن :

$Z_{tab} = 1.96 < |Z_{cal}| = 16$ وبالتالي نرفض فرضية العدم ونقبل الفرضية البديلة .

التمرين رقم 04

في احدى المؤسسات يتم تدريب العمال بطريقتين مختلفتين طريقة حديثة وطريقة تقليدية وذلك لاجراء عملية تجميع منتوج معين بحيث تتطلب عملية التجميع فترة زمنية معينة (بالدقائق). تم سحب عينة عشوائية من مجموعة مدربة بطريقة حديثة A وعينة عشوائية ثانية من مجموعة مدربة بطريقة كلاسيكية B ، البيانات المتحصل عليها مبينة في الجدول الموالي.

المطلوب: هل توجد أدلة كافية للإشارة إلى اختلاف في متوسط أوقات التجميع الحقيقية (بالدقائق) لأولئك الذين تم تدريبهم باستخدام الطريقتين السابقتين. اختبر هذه الفرضية عند مستوى $\alpha = 0.05$ ؟

طريقة التدريب الحديثة A	طريقة التدريب الكلاسيكية B
$n_1 = 9$ $\bar{x}_1 = 35.22$ $\sum_{i=1}^{9}(x_{1i} - \bar{x}_1)^2 = 195.56$	$n_2 = 9$ $\bar{x}_1 = 31.56$ $\sum_{i=1}^{9}(x_{2i} - \bar{x}_2)^2 = 160.22$

الحل:

صياغة الفرضيات

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

الاختبار الاحصائي الذي يستخدم هو اختبار ستيودنت حيث تعطى صيغة هذه الاحصائية على الشكل:

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightsquigarrow T_{n_1+n_2-2}^{1-\alpha/2}$$

التباين يكون متساوي لأن العينتين مسحوبتين من نفس المجتمع الاحصائي $\sigma_1^2 = \sigma_2^2 = \hat{\boldsymbol{\sigma}}^2$

نقوم بحساب $\hat{\sigma}^2$ (التباين المرجح) كما يلي :

$$\hat{\boldsymbol{\sigma}}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{195.56+160.22}{9+9-2} = 22.2362$$

$$\hat{\boldsymbol{\sigma}}^2 = 22.2362$$

$$T = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\hat{S}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(35.22-31.56)-0}{\sqrt{22.2362\left(\frac{1}{9}+\frac{1}{9}\right)}} = \frac{3.66}{2.2229} = 1.6464$$

عند حد المعنوية $\alpha = 0.05$ و درجة الحرية $v = n_1+n_2-2 = 16$ تعطى إحصائية

ستيودنت الحرجة (المجدولة) $T_{n_1+n_2-2=16}^{\frac{\alpha}{2}=0.025} = 2.120$

بمقارنة احصائية توزيع ستيودنت المحسوبة مع المجدولة نجد أن:

$T_{tab} = 2.12 > |T_{cal}| = 1.6464$ وبالتالي نقبل فرضية العدم بعدم وجود اختلاف في متوسط أوقات التجميع الحقيقية لأولئك الذين تم تدريبهم باستخدام الطريقة الحديثة والطريقة التقليدية ونرفض الفرضية البديلة .

التمرين رقم 05

من أجل معرفة مدى وجود فرق جوهري من عدمه بين أجور الأساتذة في الطور الثانوي **X1** و الطور المتوسط **X2** قام أحد الباحثين باختيار عينتين عشوائيتين من كل طور والنتائج المتحصل عليها ملخصة في الجدول التالي :

X1	20	21	18	25	15	23	21	25	21	20	19	14		
X2	21	16	15	15	18	22	16	15	17	19	13	16	14	20

المطلوب :

1. عند حد المعنوية $\alpha = 5\%$ بين هل يوجد فرق معنوي للأجور المتوسطة بين أساتذة طور التعليم الثانوي و طور التعليم المتوسط مع توضيح الفرضيات المراد اختبارها ، مع العلم أن التباين بين الطورين متساوي وفقا لدراسات سابقة ؟

الحل:

من خلال جدول المعطيات السابق نحسب المقاييس التالية :

طريقة التدريب التقليديةB	طريقة التدريب الحديثة A
$\bar{x}_2 = 16.93$ $S_2^2 = 7.3$	$\bar{x}_1 = 20.17$ $S_1^2 = 11.607$

$$\hat{\boldsymbol{\sigma}}^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = \frac{(12-1)11.607+(14-1)7.3}{12+14-2} = 9.27$$

$$T = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\hat{\boldsymbol{\sigma}}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(20.33-16.93)-0}{\sqrt{9.27\left(\frac{1}{12}+\frac{1}{14}\right)}} = \frac{3.66}{2.2229} = 2.706$$

عند حد المعنوية $\alpha = 0.05$ و درجة الحرية $v = n_1 + n_2 - 2 = 24$ تعطى إحصائية ستيودنت الحرجة (المجدولة) $T^{\frac{\alpha}{2}=0.025}_{n_1+n_2-2=24} = 2.064$

بمقارنة احصائية توزيع ستيودنت المحسوبة مع المجدولة نجد أن:

$T_{\text{tab}} = 2.064 > |T_{cal}| = 1.6464$ وبالتالي نقبل فرضية العدم بعدم وجود فروقات جوهرية بين متوسط أجور أساتذة التعليم المتوسط والتعليم الثانوي ونرفض الفرضية البديلة .

التمرين رقم 06

في دراسة حول المساواة في ادخار زبائن أحد البنوك التجارية ، تم اختيار عينة عشوائية من رواتب 50 زبونا ذكر و 50 زبونا أنثى ، الإحصاءات المتعلقة بالعينة موضحة في الجدول الموالي :

	متوسط ادخار زبائن البنك	الانحراف المعياري لادخار زبائن البنك
الزبائن الذكور	36400	360
الزبائن الاناث	34200	220

المطلوب: اختبر فرضية أن متوسط ادخار الزبائن الذكور أكبر من متوسط ادخار الزبائن الاناث عند مستوى معنوية $\alpha = 0.05$ مع العلم أن العينتين المختارتين مستقلتين ؟

الحل :

ليكن μ_1 و μ_2 متوسط ادخار الزبائن الذكور والاناث على التوالي في هذا البنك

صياغة الفرضيات

$$\begin{cases} H_0: \mu_1 - \mu_2 \leq 0 \\ H_1: \mu_1 - \mu_2 > 0 \end{cases}$$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} = \frac{(36400 - 34200) - 0}{\sqrt{\frac{(360)^2}{50} + \frac{(220)^2}{50}}} = 36.872$$

بما أن $n_1 \geq 30,\ n_2 \geq 30$ و التباين σ_1^2، σ_2^2 مجهولين فانه يتم تعويضهما بتباين العينة s_1^2، s_2^2

(من اجل مقارنة متوسطي مجتمعين مستقلتين في حالة عينتين صغيرتين $n_1 < 30,\ n_2 < 30$

و التباين σ_1^2 ، σ_2^2 مجهولين فانه يمكن افتراض أن: $\sigma_1^2 = \sigma_2^2$)

عند حد المعنوية $\alpha = 0.05$ تعطى إحصائية التوزيع الطبيعي $Z_\alpha = Z_{0.05} = 1.645$

هذا الاختبار هو اختبار أحادي من اليمين، نرفض فرضية العدم اذا كان : $Z_{0.05} = 1.645 < Z$

بمقارنة احصائية التوزيع الطبيعي المحسوبة مع المجدولة نجد أن:

$Z_{0.05} = 1.645 < Z = 36.872$: اذا نرفض فرضية العدم ونقبل الفرضية البديلة بأن متوسط ادخار الذكور في هذا البنك أكبر من متوسط ادخار الاناث .

التمرين رقم 07

لأغراض تجارية يريد مدير التسويق لأحد العلامات التجارية المشهورة الخاصة بصناعة الهواتف النقالة معرفة نسبة الزبائن المعجبين بالجيل الجديد من الهواتف الذي تم تداوله في السوق، لتحقيق هذا الغرض تم اختيار عينة عشوائية من 500 زبون لهذه العلامة التجارية حيث وجد أن من بينهم يوجد 200 زبون معجب بخصائص الجيل الجديد من الهاتف الذي تم طرحه في السوق لهذه العلامة التجارية.

المطلوب: اذا كان الجيل الجديد من الهواتف يعتبر ناجحا يعتبر ناجح في حالة ما كانت نسبة الاعجاب والرضى به أكثر 30% من الزبائن، انطلاقات من المعطيات السابقة عند مستوى معنوية 10% في رأيك هل يعتبر طرح هذا الهاتف الجديد ناجحا مع توضيح الفرضيات الاحصائية المراد اختبارها ؟

الحل :

الفرضيات المراد اختبارها :

$$H_1 : p > 0.3 \quad vs \quad H_0 : p \leq 0.3$$

هذا الاختبار هو اختبار أحادي الاتجاه نحو اليمين بحيث نرفض فرضية العدم اذا كان :

$$Z_\alpha < Z_{cal} = \frac{\hat{p} - p_0}{\sqrt{\frac{(\hat{p})(1-\hat{p})}{n}}}$$

$$\hat{p} = \frac{200}{500} = 0.4$$

$$Z_{cal} = \frac{\hat{p} - p_0}{\sqrt{\frac{(\hat{p})(1-\hat{p})}{n}}} = \frac{0.4 - 0.3}{\sqrt{\frac{(0.3)(0.7)}{500}}} = 4.8795$$

بما أن : $Z_\alpha = 1.28 < Z = 4.8795$

فاننا نرفض فرضية العدم H_0 ونقبل الفرضية البديلة $H_1 : p > 0.3$ وبالتالي نعتبر هذا الجيل الجديد من الهواتف الخاصة بهذه العلامة التجارية ناجح .

التمرين رقم 08

أقرت احدى الحكومات قوانين خاصة بمعايير المنافسة في السوق موجه للمصانع الانتاجية، وبعد مدة كافية من طرح هذا القانون وبدأ تطبيقه، قام أحد ملاك المصانع بتقديم شكوى الى السلطات المختصة يدعي فيها أن نسبة المصانع التي تخضع لمعايير المنافسة في المنطقة A أقل منه في المنطقة B،و من أجل التحقق من هذا الادعاء قام فريق من الخبراء التابعين للحكومة باجراء دراسة احصائية للتحقق من هذا الادعاء تم اختيار عينة عشوائية من 60 مصنع من المنطقة A حيث وجد ضمنها 50% المصانع تخضع لمعايير المنافسة التي أقرتها الحكومة (تخص منتوج معين)، بينما في المنطقة B تم اختيار عينة عشوائية من 40 مصنع حيث وجد ضمنها نسبة 40% تخضع لمعايير المنافسة.

المطلوب: في رأيك ماهي النتيجة التي سيتوصل اليها خبراء التابعين للحكومة، برر اجابتك احصائيا عند مستوى ثقة 95% مبينا الفرضيات الاحصائية المراد اختبارها ؟

الحل:

الفرضيات المراد اختبارها هي :

$$H_1 : p_1 > p_2 \quad vs \quad H_0 : p_1 \leq p_2$$

أو بعبارة اخرى : $H_1 : p_1 - p_2 < 0\ , \quad vs \quad H_0 : p_1 - p_2 \geq 0$

$\boldsymbol{p_2, p_1}$: نسبة المصانع التي تخضع لمعايير المنافسة التي أقرتها الحكومة في المنطقة A والمنطقة B على التوالي

ان هذا الاختبار الاحصائي ذو اتجاه واحد(أحادي الاتجاه) نحو اليمين اذا $Z_\alpha = 1.645$

يقترب توزيع للمتغيرة العشوائية $\hat{p}_1 - \hat{p}_2$ من التوزيع الطبيعي وفقا لنظرية النهاية المركزية اذا كان حجم العينة كبير بتحقق الشروط التالية :

$n_1 p_1 = n_1 \hat{p}_1 = (60)(0.5) = 30 > 5 \quad n_1(1-\hat{p}_1) = (60)(0.5) = 30 > 5$

$n_2 p_2 = n_2 \hat{p}_2 = (40)(0.4) = 16 > 5 \quad , n_2(1-\hat{p}_2) = (40)(0.6) = 24 > 5$

اذا: $(\hat{p}_1 - \hat{p}_2) \rightsquigarrow N(p_1 - p_2; \sigma^2_{(\hat{p}_1-\hat{p}_2)})$

وتعطى صيغة المتغيرة المعيارية **Z** التي تتبع التوزيع الطبيعي المعياري كما يلي :

$$Z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sigma_{\hat{p}_1-\hat{p}_2}} \rightsquigarrow N(1;0)$$

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

في حالة اختبار أحادي من اليمين فانه يتم رفض فرضية العدم اذا كان : $Z_\alpha < Z_{cal}$

$$Z = \frac{(0.5-0.4)-0}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

$$Z = \frac{(0.5-0.4)-0}{\sqrt{\frac{0.5(0.5)}{60} + \frac{0.4(0.6)}{40}}} = \frac{0.1}{0.1008} = 0.992$$

بما أن : $\boldsymbol{Z_\alpha = 1.64 > Z_{cal} = 0.992}$

وبالتالي نقبل فرضية العدم H_0 ونرفض الفرضية البديلة H_1 ومنه نستنتج أن الادعاءات التي تضمنتها شكوى صاحب المصنع غير صحيحة وننصح خبراء الحكومة برفض الشكوى.

التمرين رقم 09

يدعي طبيب أن التباين في مستويات الكوليسترول لدى الرجال البالغين في مختبر معين هو 100 على الأقل.أنتجت عينة عشوائية من 25 ذكرًا بالغًا من هذا المختبر انحراف معياري لمستويات الكوليسترول في الدم بمستوى 12.

المطلوب : اختبر صحة ادعاء الطبيب عند مستوى معنوية بنسبة 5٪ ؟

الحل:

الفرضيات هي:

$$H_0 : \sigma^2 \geq 144 \quad vs \quad H_1 : \sigma^2 < 144$$

هذه الاختبار هو أحادي من اليسار بحيث نرفض فرضية العدم اذا كان : $\frac{(n-1)s^2}{\sigma_0^2} < {\chi^2}_{1-\alpha}^{(n-1)}$

عند مستوى معنوية 5% نجد احصائية مربع كاي المجدولة:

$$\frac{(n-1)s^2}{\sigma_0^2} = \frac{(25-1)12^2}{100} = 34.56$$

$${\chi^2}_{1-\alpha}^{(n-1)} = {\chi^2}_{1-0.05}^{(25-1)} = {\chi^2}_{0.95}^{(24)} = 13.85$$

اذا نجد أن:

$${\chi^2}_{1-\alpha}^{(n-1)} = 13.85 < \frac{(n-1)s^2}{\sigma_0^2} = 34.56$$

وبالتالي نقبل فرضية العدم $H_0 : \sigma^2 \geq 144$ و نرفض الفرضية البديلة $H_1 : \sigma^2 < 144$

التمرين رقم 10

في مصنع لانتاج أحد المواد الغذائية يريد مراقب الجودة اختبار دقة وزن المنتوج حيث اذا كان الانحراف المعياري $\sigma = 1.2$ فهذا موافق لمعايير الانتاج في المؤسسة. تم اختيار عينة عشوائية n=10 قطعة من انتاج المصنع بانحراف معياري يساوي 1.5 . مع العلم أن متغير وزن المنتوج في المصنع يتبع التوزيع الطبيعي.

المطلوب :

- عند مستوى معنوية $\alpha = 0.1$ اختبر الفرضية التالية :

$$H_0 : \sigma^2 \neq 1.44 \quad vs \quad H_1 : \sigma^2 \neq 1.2$$

الحل :

$$\sigma^2 = (1.2)^2 = 1.44$$

نعلم أن :

$$\frac{(n-1)s^2}{\sigma^2} \rightsquigarrow {\chi^2}_{(n-1)}$$

$$Pr\left({\chi^2}_{1-\alpha/2}^{(n-1)} \leq \frac{(n-1)s^2}{\sigma^2} \leq {\chi^2}_{\alpha/2}^{(n-1)}\right) = 1-\alpha$$

في حالة اختبار ثنائي الاتجاه فاننا نرفض فرضية العدم H_0 اذا كانت احصائية الاختبار $\frac{(n-1)s^2}{\sigma_0^2}$ لا تقع في مجال القبول التالي : $\left[{\chi^2}_{1-\alpha/2}^{(n-1)}, {\chi^2}_{\alpha/2}^{(n-1)}\right]$

عند مستوى معنوية $\alpha = 0.1$ نجد احصائية مربع كاي المجدولة (حدي مجال الثقة) بحيث:

$${\chi^2}_{1-\alpha/2}^{(n-1)} = {\chi^2}_{0.95}^{(9)} = 3.33$$
$${\chi^2}_{\alpha/2}^{(n-1)} = {\chi^2}_{0.05}^{(9)} = 16.92$$

مجال قبول فرضية العدم $[3.33,16.92]$

$$\frac{(n-1)s^2}{\sigma^2} = \frac{(10-1)1.5^2}{1.2^2} = 14.0625$$

نلاحظ أن قيمة الاختبار لا تنتمي الى مجال القبول وبالتالي نرفض فرضية العدم ونقبل الفرضية البديلة $H_1 : \sigma^2 \neq 1.2$

التمرين رقم 11

في مصنع لانتاج أنابيب النحاس يسمح مراقب الانتاج بتسويق الانابيب اذا كان الانحراف المعياري في سمك الأنبوب في حدود 0.003، يقوم مراقب الانتاج بمراقبة روتينية للانتاج خلال كل بداية شهر حيث قام باختيار عينة عشوائية من 20 أنبوب نحاس حيث وجد أن الانحراف المعياري في سمك الأنابيب يساوي 0.004 .

المطلوب: عند مستوى معنوية $\alpha = 0.05$ بما ذا تنصح مراقب الانتاج فيما يتعلق بتسويق هذا المنتوج ؟

الحل :

الفرضيات هي :

$$H_0 : \sigma^2 = 0.003^2 \quad vs \quad H_1 : \sigma^2 \neq 0.003^2$$

في حالة اختبار ثنائي الاتجاه فاننا نرفض فرضية العدم H_0 اذا كانت احصائية الاختبار $\frac{(n-1)s^2}{\sigma_0^2}$ لا تقع في مجال القبول التالي : $\left[{\chi^2}_{1-\alpha/2}^{(n-1)}, {\chi^2}_{\alpha/2}^{(n-1)}\right]$

عند مستوى معنوية $\alpha = 0.1$ نجد احصائية مربع كاي المجدولة:

$${\chi^2}_{1-\alpha/2}^{(n-1)} = {\chi^2}_{0.975}^{(19)} = 8.91$$
$${\chi^2}_{\alpha/2}^{(n-1)} = {\chi^2}_{0.025}^{(9)} = 32.85$$
$$[8.91,32.85]$$
$$\frac{(n-1)s^2}{\sigma_0^2} = \frac{(20-1)0.004^2}{0.003} = 33.7777$$

نلاحظ أن قيمة الاختبار الاحصائي تقع خارج مجال الثقة وبالتالي نرفض فرضية العدم ونقبل الفرضية البديلة $H_1: \sigma^2 \neq 0.003^2$

التمرين رقم 12

خبير يرى أن التباين في جهاز معين أنه ينتج انحراف معياري قدره 02 ، تم اختيار عشوائيا 16 جهازا وتم تجريبه أسفر عن انحراف معياري في القياس يساوي 6.1 .

المطلوب : في رأيك هل أنت مع ادعاء هذا الخبير عند حد المعنوية 0.05 ؟

الحل:

صياغة الفرضيات :

$$H_0 : \sigma^2 = 4 \quad vs \quad H_1: \sigma^2 \neq 4$$

$$\sigma^2 = 2^2 = 4$$

هذا الاختبار هو ثنائي الاتجاه و تعطى قيمة اختبار مربع كاي بالصيغة التالية :

$$\frac{(n-1)s^2}{\sigma^2} = \frac{(16-1)6.1^2}{4} = 22.875$$

من خلال الجدول الاحصائي لتوزيع مربع كاي عند حد المعنوية 0.05 نجد احصائية مربع كاي المجدولة :

$$\chi^{2\,(n-1)}_{\alpha/2} = \chi^{2\,(15)}_{0.025} = 27.49$$

$$\chi^{2\,(n-1)}_{\alpha/2} = \chi^{2\,(15)}_{0.025} = 6.26$$

نلاحظ أن قيمة اختبار مربع كاي المحسوبة تقع داخل مجال القبول وبالتالي نقل فرضية العدم $H_0 : \sigma^2 = 4$ وعليه أتفق مع ادعاء هذا الخبير

كان الهدف من خلال هذه المطبوعة هو تقديم دروس وتمارين في مقياس احصاء 03 وبالخصوص الاحصاء الاستدلالي بطريقة سهلة ومبسطة انطلاقا من توزيع المعاينة ثم نظرية التقدير وأخيرا اختبار الفرضيات، من خلال الخبرة المكتسبة في تدريس هذا المقياس يبقى عنصر الفهم ضعيفا نظرا لوجود تسلل منطقي بين المبادئ الأساسية التي من المفترض أن يتم اكتسابها في الاحصاء 01 و الاحصاء 02 بحيث يعتبر الاحصاء 03 تكملة لما سبق ويعتمد بدرجة كبيرة على المعارف السابقة المتحصل عليها، الا أن الملاحظ من خلال التجربة هو وجود تقريبا قطيعة معرفية مع ما تم دراسته مسبقا لدى أغلبية الطلبة وأحيانا بالنسبة حتى للبديهيات، وهو ما طرح اشكالا حقيقيا في درجة استيعاب وفهم هذا المقياس وحتى أي مقياس آخر يعتمد تدريسه على المعارف السابقة.

ولهذا وجب اعادة النظر في برنامج مقياس الاحصاء بكامل فروعة (01 و 02 و 03) والتركيز على التعليم النوعي بدلا من الكمي وخفض البرنامج لمضاعفة درجة الاستيعاب، والعمل على تدريس مقياس الاحصاء بفروعه المختلفة في كل سداسي ابتداءا من السنة الأولى ليسانس الى غاية آخر سنة في الماستر مع ضرورة حث الطلبة على حضور المحاضرات لأنها تعتبر الركيزة الأساسية لفهم حصص الأعمال الموجهة.

قائمـــــة المراجـــع

❖ المراجع **باللغة العربية**

1) عبد الحميد عبد المجيد البلداوي (2007)، أساليب البحث العلمي والتحليل الاحصائي، التخطيط للبحث وجمع وتحليل البيانات يدويا وباستخدام برنامج SPSS، الطبعة العربية الأولى، الاصدار الثالث، دار الشروق للنشر والتوزيع عمان الأردن.

2) خالد زهدي خواجه (1-2001/8/2)، تطبيقات أساليب المعاينة في الأقطار العربية، اصدار المعهد العربي للتدريب والبحوث الإحصائية (وقائع ندوة تطبيقات أساليب المعاينة في الأقطار العربية دمشق. سوريا)

3) اياد محمد الهوبي (2014)، الاحصاء التطبيقي، الطبعة الأولى الكلية الجامعية للعلوم والتكنولوجيا

4) بوعبد الله صالح، محاضرات الإحصاء الرياضي 2005-2006

5) معتوق أمحمد، الاحصاء الرياضي والنماذج الاحصائية، ديوان المطبوعات الجامعية (بدون سنة نشر)

6) علي عبد السلام العماري، علي حسين العجيلي، الاحصاء والاحتمالات النظرية والتطبيق ، دار الحكمة (بدون سنة نشر)

7) نبيل جمعة النجار(2015)، الإحصاء التحليلي مع تطبيقات برمجية SPSS ، الطبعة الأولى، دار ومكتبة الحامد للنشر والتوزيع عمان.

8) سمير فهمي حجازي (2004-2003) ، محمود الدريني، الاحصاء التطبيقي، الشركة المصرية لإعادة التأمين. مصر.

9) عزام عبد الرحمن صبري(2015) ، الإحصاء التطبيقي بنظام SPSS، الدار المنهجية للنشر والتوزيع، عمان الأردن.

10) عدنان عباس حميدان، مطانيوس مخول، فريد جاعوني، عمار ناصر آغا (2005.2006) ، الاحصاء التطبيقي، منشورات جامعة دمشق .

11) دومينيك سالفاتور، ترجمة سعدية حافظ منتصر(1982)، ملخصات شوم نظريات ومسائل في الاحصاء والاقتصاد القياسي، الدار الدولية للنشر والتوزيع

12) عماد نشوان (2005)، الدليل العملي لمقرر الإحصاء التطبيقي، جامعة القدس المفتوحة

13) البشير عبد الكريم (2006)، إحصاء02 (الإحتمالات و الإحصاء) دروس مع تمارين محلولة، السنة الثانية علوم تسيير، علوم اقتصادية، علوم تجارية، دار الكتاب العربي، الجزائر.

❖ المراجع باللغة الأجنبية

14) Weiss, N. A. , Neil A. (2012) Introductory statistics; biographies by Carol A. Weiss. – 9th ed. Pearson Education

15) KandethodyM.Ramachandran (2009), Chris P.Tsokos, Mathematical Statistics with Applications, Elsevier Academic Press .

16) Renée Veysseyre (2006), Statistique et Probabilités pour l’ingénieur, 2 emme Edition Dunod ,PARIS .

17) Kandethody M. Ramachandran, Chris P.Tsokos (2009), Mathematical Statistics with Applications, Elsevier Academic Press .

18) Michel Lejeune (2010), Statistique La théorie et ses applications, Deuxième édition
avec exercices corrigés, Sprin-ger-Verlag Paris France .

19) CHRISTOPHE HURLIN, VALÉRIE MIGNON (2015), Statistique et Probabilités en Economié-Gestion, Dunod, Paris France.

20) Dennis D. Wackerly , William Mendenhall III, Richard L. Scheaffer (2008) Mathematical Statistics with Applications, Thomson Learning

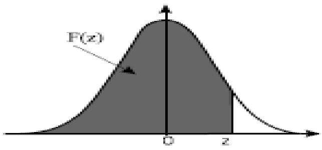
21) Dominick Salvatore et Georges Loudière (1986), Économétrie et statistique appliquées : cours et problèmes , McGraw-Hill, Auckland

قائمـــــة الملاحق

Standard Normal Distribution Table الملحق رقم 01: جدول التوزيع الطبيعي المعياري

Probability of finding a value less than : z

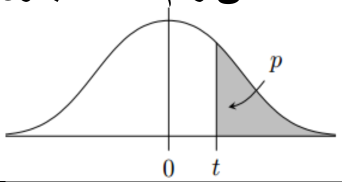
$f(-z) = 1 - f(z)$ $z = (x - \mu) / \sigma$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

Student's Law Table الملحق رقم 02 : : جدول توزيع ستيودنت

***t** as a function of p such that p = P[T > t]*
.for T according to a Student law

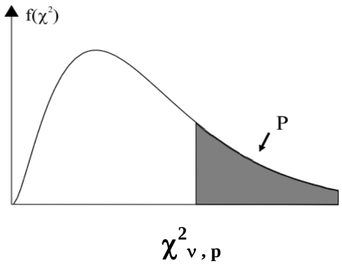


	p (α)										
v	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

الملحق رقم 03: جدول توزيع مربع كاي Chi–Square Distribution Table

Percentage Points of the χ^2 Distribution; $\chi^2_{v,p}$

$$P(\chi^2 > \chi^2_{v,p}) = p$$



	p (α)												
V	**0.995**	**0.990**	**0.975**	**0.950**	**0.900**	**0.750**	**0.500**	**0.250**	**0.100**	**0.050**	**0.025**	**0.010**	**0.005**
1	0.00004	0.00016	0.00098	0.00393	0.01579	0.10153	0.45494	1.32330	2.70554	3.84146	5.02389	6.63490	7.87944
2	0.01003	0.02010	0.05064	0.10259	0.21072	0.57536	1.38629	2.77259	4.60517	5.99146	7.37776	9.21034	10.59663
3	0.07172	0.11483	0.21580	0.35185	0.58437	1.21253	2.36597	4.10834	6.25139	7.81473	9.34840	11.34487	12.83816
4	0.20699	0.29711	0.48442	0.71072	1.06362	1.92256	3.35669	5.38527	7.77944	9.48773	11.14329	13.27670	14.86026
5	0.41174	0.55430	0.83121	1.14548	1.61031	2.67460	4.35146	6.62568	9.23636	11.07050	12.83250	15.08627	16.74960
6	0.67573	0.87209	1.23734	1.63538	2.20413	3.45460	5.34812	7.84080	10.64464	12.59159	14.44938	16.81189	18.54758
7	0.98926	1.23904	1.68987	2.16735	2.83311	4.25485	6.34581	9.03715	12.01704	14.06714	16.01276	18.47531	20.27774
8	1.34441	1.64650	2.17973	2.73264	3.48954	5.07064	7.34412	10.21885	13.36157	15.50731	17.53455	20.09024	21.95495
9	1.73493	2.08790	2.70039	3.32511	4.16816	5.89883	8.34283	11.38875	14.68366	16.91898	19.02277	21.66599	23.58935
10	2.15586	2.55821	3.24697	3.94030	4.86518	6.73720	9.34182	12.54886	15.98718	18.30704	20.48318	23.20925	25.18818
11	2.60322	3.05348	3.81575	4.57481	5.57778	7.58414	10.34100	13.70069	17.27501	19.67514	21.92005	24.72497	26.75685
12	3.07382	3.57057	4.40379	5.22603	6.30380	8.43842	11.34032	14.84540	18.54935	21.02607	23.33666	26.21697	28.29952
13	3.56503	4.10692	5.00875	5.89186	7.04150	9.29907	12.33976	15.98391	19.81193	22.36203	24.73560	27.68825	29.81947
14	4.07467	4.66043	5.62873	6.57063	7.78953	10.16531	13.33927	17.11693	21.06414	23.68479	26.11895	29.14124	31.31935
15	4.60092	5.22935	6.26214	7.26094	8.54676	11.03654	14.33886	18.24509	22.30713	24.99579	27.48839	30.57791	32.80132
16	5.14221	5.81221	6.90766	7.96165	9.31224	11.91222	15.33850	19.36886	23.54183	26.29623	28.84535	31.99993	34.26719
17	5.69722	6.40776	7.56419	8.67176	10.08519	12.79193	16.33818	20.48868	24.76904	27.58711	30.19101	33.40866	35.71847
18	6.26480	7.01491	8.23075	9.39046	10.86494	13.67529	17.33790	21.60489	25.98942	28.86930	31.52638	34.80531	37.15645
19	6.84397	7.63273	8.90652	10.11701	11.65091	14.56200	18.33765	22.71781	27.20357	30.14353	32.85233	36.19087	38.58226
20	7.43384	8.26040	9.59078	10.85081	12.44261	15.45177	19.33743	23.82769	28.41198	31.41043	34.16961	37.56623	39.99685
21	8.03365	8.89720	10.28290	11.59131	13.23960	16.34438	20.33723	24.93478	29.61509	32.67057	35.47888	38.93217	41.40106
22	8.64272	9.54249	10.98232	12.33801	14.04149	17.23962	21.33704	26.03927	30.81328	33.92444	36.78071	40.28936	42.79565
23	9.26042	10.19572	11.68855	13.09051	14.84796	18.13730	22.33688	27.14134	32.00690	35.17246	38.07563	41.63840	44.18128
24	9.88623	10.85636	12.40115	13.84843	15.65868	19.03725	23.33673	28.24115	33.19624	36.41503	39.36408	42.97982	45.55851
25	10.51965	11.52398	13.11972	14.61141	16.47341	19.93934	24.33659	29.33885	34.38159	37.65248	40.64647	44.31410	46.92789
26	11.16024	12.19815	13.84390	15.37916	17.29188	20.84343	25.33646	30.43457	35.56317	38.88514	41.92317	45.64168	48.28988
27	11.80759	12.87850	14.57338	16.15140	18.11390	21.74940	26.33634	31.52841	36.74122	40.11327	43.19451	46.96294	49.64492
28	12.46134	13.56471	15.30786	16.92788	18.93924	22.65716	27.33623	32.62049	37.91592	41.33714	44.46079	48.27824	50.99338
29	13.12115	14.25645	16.04707	17.70837	19.76774	23.56659	28.33613	33.71091	39.08747	42.55697	45.72229	49.58788	52.33562
30	13.78672	14.95346	16.79077	18.49266	20.59923	24.47761	29.33603	34.79974	40.25602	43.77297	46.97924	50.89218	53.67196

V : is the number of degrees of freedom. For V> 30, we can assume that the quantity

$\sqrt{2\chi^2} - \sqrt{2V-1}$ follows **the Standard Normal Distribution**