الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

كلية العلوم الاقتصادية، التجارية وعلوم التسيير

قسم علوم التسيير

مطبوعة في مقياس

الإحصاء 03
(الإحصاء الإستنتاجي أو الاستدلالي)

موجهة إلى طلبة السنة الثانية ليسانس نظام ل.م.د كل التخصصات

إعداد الدكتور: عابد علي

السنة الجامعية: 2018/2017

قائمة المحتويات

مقدمة

الاستنتاج الإحصائي أو(الاستدلال الإحصائي)، هو فرع من فروع علم الإحصاء يشمل كل الأساليب الإحصائية والنظريات القائمة عليها وتطبيقاتها العملية المستخدمة في تحليل البيانات (المعلومات) التي نحصل عليها من العينة، وذلك للاستنتاج أو الاستدلال على معالم المجتمع وخواص المجتمع الذي سحبت منه العينة، وتكون هذه الاستنتاجات على شكل تقديرات أو اختبارات فروض واتخاذ قرارات.

والاستدلال الإحصائي واحد من أكثر جوانب عملية اتخاذ القرارات أهمية وحيوية في الاقتصاد والأعمال والعلوم، ويتعلق الاستدلال الإحصائي بالتقدير واختبار الفروض، والتقدير الإحصائي هو عملية استنتاج أو تقدير معالم المجتمع (المتوسط الحسابي، أو التباين) من الإحصاء المناظر والخاص بعينة مسحوبة من المجتمع،ولكي يكون التقدير أو اختبار الفروض سليما، ينبغي أن يبنى على عينة ممثلة للمجتمع، ويمكن تحقيق ذالك بالمعاينة العشوائية حيث يكون لكل مفردة في المجتمع فرصة متكافئة للدخول في العينة.

وعليه فقد جاءت هذه المطبوعة في شكل محاضرات موجهة إلى طلبة السنة ثانية علوم التسيير والعلوم الاقتصادية وعلوم المالية والمحاسبة نظام ل.م.د وفق المقرر لهذا المقياس، وذلك من اجل إكمال المفاهيم المتعلقة بالإحصاء ألاستنتاجي أو الاستدلالي بعد دراستهم للإحصاء الوصفي والرياضي.

وقد تم تقسيمها إلى ثلاث محاور أساسية هي:

المحور الأول: توزيع المعاينة

المحور الثاني: تقدير معالم المجتمع

المحور الثالث:اختبار الفروض

المحور الأول:توزيعات المعاينة

تمهيد

قبل التطرق إلى توزيعات المعاينة سوف نوضح المصطلحات والمفاهيم الإحصائية التي لها علاقة مباشرة بموضوع الإحصاء ألاستنتاجي أو الاستدلالي وهذه المصطلحات هي: المجتمع، توزيع المجتمع، المعلمة، العينة، العينة العشوائية البسيطة، توزيع العينة، الإحصائية.

1-المجتمع(N): يعرف المجتمع الإحصائي بأنه مجموعة كل المفردات المقصودة بالدراسة، والمقصود هنا ضمنيا أن المجتمع يتكون من القيم التي نجمعها عن الظاهرة محل الدراسة من كل المفردات المقصودة بالدراسة،لأننا نحن بالطبع نتعامل مع القيم التي نحصل عليها من المفردات وليس مع المفردات نفسها فالمجتمع في أي دراسة إحصائية هو عبارة عن قيم عددية.

والمفردات في أي دراسة إحصائية قد تكون أشخاصا أو أشياء جامدة، وقد يكون المجتمع محدودا، أي نستطيع تحديد العد الكلي لقيمه أي العدد الكلي لمفرداته (عدد القيم هو نفسه عدد المفردات)، وقد يكون غير محدود (لانهائي) أي أن العدد الكلي لمفرداته كبيرا جدا لا يمكن تحديده أو حصره.

وبصفة عامة يطلق على عدد القيم التي تكون المجتمع أي العدد الكلي لمفردات المجتمع مصطلح حجم المجتمع ويرمز له بالرمز(N) وإذا رمزنا للمتغير العشوائي الذي يمثل الظاهرة محل الدراسة بالرمز X وكان المجتمع محدودا فسنرمز لقيم المتغير التي تكون هذا المجتمع كما يلي:

$$x_1, x_2, x_3 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots x_N$$

2-توزيع المجتمع: هو التوزيع التكراري أو التوزيع الاحتمالي لبيانات المجتمع، أو التوزيع الاحتمالي للقيم المجمعة عن كل مفردات المجتمع.

3-المعلمة: المعلمة هي أي مقياس إحصائي تحسب قيمته من بيانات المجتمع ككل دون استثناء، ونستخدمه لوصف المجتمع محل الدراسة وتحديد معالمه، وبالتالي يطلق عليه معلمة ، ومن المعالم أي المقاييس التي تصف لنا المجتمع،هي مقاييس النزعة المركزية(الوسط الحسابي، الوسيط ، المنوال،..) أو مقاييس التشتت (التباين، الانحراف المعياري،...) أو مقاييس الالتواء و لتفرطح، أو أكبر قيمة أو أصغر قيمة، أو أي مقاييس إحصائية أخرى تحسب من المجتمع، **والمعلمة عبارة عن قيمة ثابتة لا تتغير،** لأنها تحسب من المجتمع محل الدراسة، والمجتمع ثابت لا يتغير أثناء إجراء الدراسة ولذلك يطلق على المعالم أحيانا الثوابت لإحصائية، وعادة تستخدم الحروف اليونانية للتعبير عن المعالم، فيرمز للمتوسط الحسابي للمجتمع بالحرف μ (ميو)، ولتباين المجتمع بالرمز σ^2 (سيجما تربيع)، وللانحراف المعياري للمجتمع بالحرف σ ....الخ، وتعطى معادلة μ و σ^2 كالتالي:

$$\mu = \frac{\sum_{i=1}^{n} X_i}{N} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(01)$$

$$\sigma^2 = \frac{\sum_{i=1}^{n}\left(X_i - \mu\right)^2}{N} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (02)$$

4-العينة(n): في أي دراسة إحصائية يجب أن يكون الهدف هو دراسة المجتمع ككل وليس دراسة العينة، ولكن نستخدم العينة لأننا في اغلب الدراسات لا نستطيع أن نجمع بيانات عن كل مفردات المجتمع محل الدراسة وذلك للأسباب التالية:

- ✓إذا كان حجم المجتمع محل الدراسة كبيرا جدا وكانت إمكانات الباحث المادية محدودة، ولا تسمح له بجمع بيانات عن كل مفردة من مفردات المجتمع.
- ✓إذا كان حجم المجتمع لا نهائيا أي من المستحيل دراسته ككل.
- ✓إذا كانت دراسة المجتمع ككل تؤدي إلى تلف المجتمع بأكمله ومثال على ذلك الدراسة الخاصة بصلاحية طلبيه البيض، فالمجتمع في هذه الدراسة هو طلبيه البيض والمفردة هي البيضة، ودراسة المجتمع ككل تعني فحص كل بيضة أي كسر البيض جميعه، وهذا يؤدي للقضاء على الطلبية كلها، أي يتلف المجتمع بأكمله.
- ✓إذا كان المجتمع محل الدراسة متجانسا، أي أن جميع مفرداته تتمتع بنفس الخواص.

ويمكن تقسيم العينات إلى قسمين رئيسيين هما:

4-1عينات إحتمالية: وهي العينات التي يتم سحبها من المجتمع على أساس قوانين الاحتمالات، بحيث يمكن ربط سحب المفردات باحتمالات معينة، وهذا يعني أن تكون عملية المعاينة-أي سحب العينة- مرتبطة ارتباطا قويا بالنماذج الاحتمالية.

وفي هذه الطريقة نحصل على العينة بواسطة سحب المفردات بالتتابع، وكل منها لها احتمال معروف في الظهور في السحبة الأولى وفي أي سحبة تالية، ويكون ظهور أي مفردة من المفردات في السحبة التالية إما مستقلا عن ظهور مفردة السحبة الأولى أو غير مستقل عنه، فعمليات السحب المتتالية في العينات الاحتمالية، قد تكون بإرجاع المفردة المختارة قبل سحب المفردة التالية لها، وتسمى هذه الطريقة المعاينة مع الإرجاع، وفي هذه الحالة يكون ظهور المفردات مستقلا، وفد تكون مع عدم إرجاع المفردة المختارة قبل سحب التي تليها، وتسمى هذه الطريقة المعاينة مع عدم الإرجاع، وفي هذه الحالة يكون ظهور المفردات غير مستقل، ويجب أن تسحب العينات الاحتمالية بطريقة عشوائية لا يتدخل العامل البشري فيها، وذلك بإتباع طريقة السحب العشوائي، ومن أهم مميزات العينات الاحتمالية هو انه عند استخدامها نستطيع تحديد حجم الأخطاء الناجمة بسبب المعاينة أي حجم خطأ الصدفة، وكذلك قيمة خطأ التحيز في التقدير إن وجد، ومن أهم أنواع العينات الاحتمالية (العينة العشوائية البسيطة(عينة غير مقيدة)،أما باقي العينات وهي العينة المنتظمة،العينة الطبقية، العينة العنقودية، و العينة المضاعفة فهي عينات مقيدة.

4-2عينات غير احتمالية: وهي العينات التي لا يتم سحبها وفق قوانين الاحتمالات، وبالتالي لا نستطيع معرفة احتمال اختيار أي مفردة من مفرداتها، ولهذا لا نستطيع تطبيق نظرية الاحتمالات، وبالتالي هذه العينات لا تسمح بالتعميم من العينة إلى المجتمع، هذا وليس هناك أي طريقة إحصائية لتحديد حجم الخطأ وقياس دقة نتائج المعاينة غير الاحتمالية، ولهذا فالمعاينة غير الاحتمالية لا تعتبر طريقة من طرق المعاينة الجيدة.

5-العينة العشوائية البسيطة: العينة $x_1, x_2, x_3, \ldots\ldots\ldots\ldots x_n$ التي تحتوي على n من المفردات والمسحوبة من مجتمع ما، تكون عينة عشوائية بسيطة إذا كانت كل العينات ذات الحجم n الممكن سحبها من هذا المجتمع لها نفس فرصة الاختيار والعينة يمكن سحبها مع الإرجاع أو عدم الإرجاع.

5-1 السحب مع الإرجاع(معاينة غير نفاذية): عندما يكون السحب مع الإرجاع حيث يمكن أن تظهر المفردة أكثر من مرة في العينة تسمى هذه المعاينة غير نفاذية لان تكرار العملية لا يؤدي إلى تقليص عدد مفردات المجتمع، ويكون عدد العينات الممكن سحبها يساويN^n، وهي تحقق الشروط التالية:

-سحب جزء من الكل.
-الترتيب مهم.
-التكرار ممكن.
-السحب على التوالي.

ترتيبه تكرار (قائمة)
N^n

فمثلا إذا كان حجم المجتمع N=05 وحجم العينةn=03 وعملية السحب مع الإرجاع، فعندئذ عدد العينات الممكن سحبها من المجتمع في هذه الحالة يحدد كما يلي: $N^n = 05^{03} = 125$

5-2 السحب مع عدم الإرجاع (معاينة نفاذية): إذا كان سحب العينات من المجتمع تم مع عدم الإرجاع، أي عدم إرجاع المفردة المسحوبة قبل سحب المفردة التي تليها، في هذه الحالة توجد طريقتان لتحديد عدد العينات الممكن سحبها وهما:

✓ **الطريقة الأولى(ترتيبه بدون تكرار)**:تعطي أهمية لترتيب المفردات داخل العينة، والتي تحقق الشروط التالية:

-سحب جزء من الكل.
-الترتيب مهم.
-التكرار غير ممكن.
-السحب على التوالي.

ترتيبه بدون تكرار
P_N^n

وتحسب وفق العلاقة التالية: $P_N^n = \frac{N!}{(N-n)!} = N(N-1)(N-2)\ldots\ldots\ldots\ldots(N-(n-1))$

فمثلا عدد العينات ذات الحجم n=03 الممكن سحبها من مجتمع يحتوي على 05 مفردات مع عدم الإرجاع وعدم إهمال الترتيب هو: $P_N^n = P_5^3 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = \frac{5\times4\times3\times2!}{2!} = 60$

✓ **الطريقة الثانية (توفيقة بدون تكرار)**:لا تعطي أهمية لترتيب المفردات داخل العينة وهي تحقق الشروط التالية:

$$\left.\begin{array}{r}\text{-سحب جزء من الكل.}\\ \text{-الترتيب غير مهم.}\\ \text{-التكرار غير ممكن.}\\ \text{-السحب دفعة واحدة.}\end{array}\right\} \begin{array}{c}\text{توفيقة بدون تكرار}\\ C_N^n\end{array}$$

وتحسب وفق العلاقة التالية: $C_N^n = \frac{N!}{n!(N-n)!}$

فمثلا عدد العينات ذات الحجم n=03 الممكن سحبها من مجتمع يحتوي على 05 مفردات مع عدم الإرجاع و إهمال الترتيب هو $C_N^n = C_5^3 = \frac{5!}{3!(5-3)!} = \frac{5!}{3!\times2!} = \frac{5\times4\times3!}{3!\times2!} = 10$

6-توزيع العينة: نفرض أننا قمنا بسحب عينة عشوائية حجمها n إحصائي ما، فان توزيعها يشبه توزيع مجتمعها الإحصائي الذي أخذت منه لأنها قد أخذت منه بشكل عشوائي،أما المقاييس الإحصائية الوصفية المحسوبة بدلالتها كالوسط الحسابي والوسيط والمنوال والانحراف المعياري والنسبة المئوية لتكرار الصفة في العينة والتي يطلق عليها بالتوابع الإحصائية (أو المقدرات) والتي تتوافق تقريبا مع الثوابت الإحصائية المقابلة لها لان توزيع العينة يشبه توزيع المجتمع الإحصائي الذي سحبت منه هذه العينة، لكن قد تختلف التوابع الإحصائية عن الثوابت الإحصائية المقابلة لها أو عن توابع إحصائية أخرى، وهذا الاختلاف قد يكون:

-اختلافا ظاهريا ناشئا بسبب الأخطاء الاحتمالية أو أخطاء الحظ والمصادفات.

-إختلافا حقيقيا (جوهريا) ينشا عن عدم سحب العينة بشكل عشوائي، أو أنها قد سحبت من مجتمع إحصائي يختلف عن المجتمع الإحصائي المرغوب فيه.

7-الإحصائية: هي أي مقياس إحصائي تحسب قيمته من العينة المسحوبة من المجتمع محل الدراسة، فمثلا الوسط الحسابي للعينة عبارة عن إحصائية ويرمز له بالرمز $\overline{X}$، وتباين العينة عبارة عن إحصائية ويرمز له بالرمز S^2وهكذا... وحيث أن قيمة الإحصائية تعتمد على العينة المسحوبة وبما أننا نستطيع أن نسحب أكثر من عينة من نفس المجتمع فسنجد أن قيمة الإحصائية ستتغير من عينة إلى أخرى، وبالتالي فان الإحصائية **عبارة عن متغير** وهذا هو الفرق الجوهري بين المعلمة والإحصائية، فالمعلمة عبارة **عن قيمة ثابتة بينما الإحصائية عبارة عن متغير.**

ويتم حساب إحصاءة العينة لكل من $\overline{X}$ و S^2 كالتالي:

$$\overline{X}=\frac{\sum_{i=1}^{n}x_i}{n}\text{......................................(3)}$$

أما حساب التباين فهو متوقف على حجم العينة n ويتم تباين حسابه بالصيغتين التاليتين:

-إذا كان حجم العينة $n\geq30$

$$S^2=\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}\text{..............................(4)}$$

-إذا كان حجم العينة $n<30$

$$\hat{S}^2=\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}\text{..............................(5)}$$

وقد تكون إحصائية العينة مساوية لمعلمة المجتمع، أو قد تكون أصغر أو أكبر منها، ويسمى الفرق بين إحصائية العينة ومعلمة المجتمع بخطأ المعاينة sampling error أو الخطأ العشوائيrandom error، وهذا بافتراض أن العينة عشوائية وانه لا يوجد أخطاء غير عشوائية.

ويقصد بالأخطاء غير العشوائية الأخطاء في تجميع البيانات وتدوينها وتبويبها، وبما أن الفرق بين إحصائية العينة ومعلمة المجتمع قد يكون سالبا أو موجبا، لذلك يمكننا اخذ القيمة المطلقة لهذا الفرق لحساب خطأ المعاينة أي أن:

خطأ المعاينة= |إحصائية العينة-معلمة المجتمع|

والجدول التالي يوضح معالم المجتمع ومقدرات العينة(إحصاءة العينة)

الوصف	معلمة المجتمع	المقدر (إحصاءة العينة)
عدد المشاهدات	N	n
الوسط الحسابي	μ	$\overline{X}$
التباين	σ^2	S^2
الانحراف المعياري	σ	S
النسبة	P	$\hat{P}$
الفرق بين متوسطات مجتمعين	$\mu_1-\mu_2$	$\overline{X}_1-\overline{X}_2$
الفرق بين النسبة في مجتمعين	P_1-P_2	$\hat{P}_1-\hat{P}_2$

8-توزيع المعاينة: نفرض ان لدينا مجتمعا من المفردات يتبع توزيعا احتماليا معينا (سواء كان هذا المجتمع كبيرا او محدودا) واننا بصدد سحب عينة حجمها n من هذا المجتمع، بالطبع ليس معنى هذا ان هناك عينة واحدة يمكن سحبها ولكن يكون أمامنا عدد كبير من العينات التي يمكن سحبها من هذا المجتمع والتي حجم كل منها هو n من المفردات.

ولنفرض اننا سحبنا عينت عشوائية حجمها n من هذا المجتمع وحسبنا من هذه العينة مقياسا معينا (وليكن الوسط الحسابي) ثم سحبنا عينة ثانية لها نفس الحجم n وحسبنا منها نفس المقياس ثم سحبنا عينة ثالثة وحسبنا منها نفس المقياس وهكذا بالنسبة لجميع العينات التي يمكن سحبها من هذا المجتمع، سنجد أمامنا عددا كبيرا من القيم لنفس المقياس ولا نتوقع ان تكون جميع القيم التي حصلنا عليها من العينات لهذا المقياس متساوية وإ نما ستكون مختلفة عن بعضها وتكون مجتمعا آخر،عدد مفرداته أكبر بكثير من عدد مفردات المجتمع الاصلي، وعلى ذلك يمكن النظر الى هذا المقياس على انه متغير عشوائي يأخذ قيما مختلفة (هي التي حصلنا عليها من هذه العينات) ويتبع توزيعا معينا ربما يختلف او لا يختلف عن توزيع المجتمع الاصلي ويسمى بتوزيع المعاينة لهذا المقياس سواء كان هذا المقياس هو الوسط الحسابي اوالتباين او الانحراف المعياري او نسبة المفردات التي لها صفة معينة او غيره من المقايسس الاحصائية.

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

العينات المختلفة غالبا تكون قيم متوسطاتها الحسابية المختلفة، اي ان الوسط الحسابي للعينة $\overline{X}$ كأي إحصائية يتغير من عينة إلى أخرى، ولا نستطيع معرفة الوسط الحسابي لأي عينة عشوائية قبل سحبها، أي أن الوسط الحسابي للعينة هو متغير عشوائي وبالتالي له توزيع احتمالي، ويطلق على التوزيع الاحتمالي لأي إحصائية بتوزيع المعاينة وبالتالي يسمى التوزيع الاحتمالي للوسط الحسابي للعينة بتوزيع المعاينة للوسط الحسابي للعينة والذي يمكن تعريفه كما يلي:

1-1 تعريف توزيع المعاينة للوسط الحسابي للعينة $\overline{X}$: إذا كان لدينا مجتمع محدود حجمه N وسطه الحسابي μ وتباينه σ^2، وسحبنا منه كل العينات العشوائية البسيطة الممكنة والمتساوية في الحجم وليكن حجمها n، وحسبنا الوسط الحسابي $\overline{X}$ لكل عينة، فالتوزيع الاحتمالي للمتغير العشوائي يسمى توزيع المعاينة للوسط الحسابي للعينة.

2-1 خصائص التوقع والتباين: سيتم التطرق الى خصائص التوقع والتباين وذلك لان اي متغير عشوائي يتحدد معلماته من خلال التوقع الرياضي والتباين، وكذلك الى الحاجة اليهما في النقاط التي سوف تذكر من خلال هذه المطبوعة.

1-2-1 خصائص التوقع

1-إذا كان **X** متغير عشوائي متقطع فان $E(X)=\sum_{i=1}^{n} x_i P(X=x_i)$

2- إذا كان **X** متغير عشوائي مستمر فان $E(X)=\int_{-\infty}^{+\infty} xf(x)dx=\mu$

3-نفرض C قيمة ثابتة $E(C)=C$

4- $E(CX)=CE(X)$

5- $E(X+Y)=E(X)+E(Y)$

6- $E(X-Y)=E(X)-E(Y)$

7- $E(XY)=E(X)E(Y)$

8- $E\left(\frac{X}{C}\right)=\frac{1}{C}E(X)$

9- إذا كان **X** متغير عشوائي مستمر يخضع للتوزيع الطبيعي فان $E(X)=\mu$

1-2-1 خصائص التباين

1-العلاقة العامة للتباين هي $V(X)=E\left[X-E(X)\right]^2=E(X^2)-E^2(X)=\sigma^2$

2- إذا كان **X** متغير عشوائي متقطع فان $V(X)=\sum_{i=1}^{n}\left[X-E(X)\right]^2 P(X=x)$

3- إذا كان **X** متغير عشوائي مستمر فان

$$V(X)=\int_{-\infty}^{+\infty}\left[X-E(X)\right]^2 f(x)dx$$

$$=\int_{-\infty}^{+\infty}X^2 f(X)dX-\mu^2$$

4- نفرض C قيمة ثابتة إذن $V(C)=0$

5- $V(CX)=C^2V(X)$

6- $V\left(\frac{X}{C}\right)=\frac{1}{C^2}V(X)$

7- $V(XY)=V(X)V(Y)$

8-إذا كان X و y متغيرين عشوائيين مستقلين عن بعضهما البعض فان

$$V(X+Y)=V(X)+V(Y)$$

$$V(X-Y)=V(X)+V(Y)$$

نظرية1: إذا كان لدينا **X** متغير عشوائي يمثل مجتمع ما **(**لا محدود**)** وهو يخضع للتوزيع الطبيعي و $\overline{X}_i$ متغيرة عشوائية تمثل متوسط عينة مسحوبة من ذات المجتمع **بالارجاع او بدون ارجاع** فان القيمة المتوقعة لمتوسط العينة $E(\overline{X})$ تكتب كالتالي: $E(\overline{X})=\mu_{\overline{x}}=\mu$

البرهان: نعلم ان المتوسط الحسابي يعطى بالعلاقة التالية: $\overline{X}=\frac{\sum_{i=1}^{n}X_i}{n}$

بادخال التوقع الرياضي على طرفي المساواة نجد:

$$E\left(\overline{X}\right)=E\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=E\left(\frac{X_1+X_2+X_3+\dots\dots\dots\dots+X_n}{n}\right)$$

$$=\frac{1}{n}E\left(X_1+X_2+X_3+\dots\dots\dots\dots+X_n\right)$$

$$=\frac{1}{n}\left(E\left(X_1\right)+E\left(X_2\right)+E\left(X_3\right)+\dots\dots\dots\dots+E\left(X_n\right)\right)$$

$$=\frac{1}{n}\left(\mu_1+\mu_2+\mu_3+\dots\dots\dots\dots\dots\dots+\mu_n\right)=\frac{1}{\cancel{n}}\left(\cancel{n}\mu\right)=\mu$$

نظرية 2: إذا كان لدينا **X** متغير عشوائي يمثل مجتمع ما (محدود او لا محدود) وهو يخضع للتوزيع الطبيعي و $\overline{X}_i$ متغيرة عشوائية تمثل متوسط عينة مسحوبة من ذات المجتمع بالارجاع او بدون ارجاع فان تباين توزيع المعاينة للمتوسطات $V\left(\overline{X}\right)$ يكتب كالتالي: $V\left(\overline{X}\right)=\sigma_{\overline{x}}^2=\frac{\sigma^2}{n}$ .... $\frac{n}{N}\le 0.05$

$\frac{n}{N}$: يسمى كسر المعاينة ويتم حسابة فقط في حالة السحب مع عدم الارجاع من مجتمع محدود.

البرهان: نعلم ان المتوسط الحسابي يعطى بالعلاقة التالية: $\overline{X}=\frac{\sum_{i=1}^{n}X_i}{n}$

بادخال التباين على طرفي المساواة نجد:

$$V\left(\overline{X}\right)=V\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=V\left(\frac{X_1+X_2+X_3+\dots\dots\dots\dots+X_n}{n}\right)$$

$$=\frac{1}{n^2}V\left(X_1+X_2+X_3+\dots\dots\dots\dots+X_n\right)$$

$$=\frac{1}{n^2}\left(V\left(X_1\right)+V\left(X_2\right)+V\left(X_3\right)+\dots\dots\dots\dots+V\left(X_n\right)\right)$$

$$=\frac{1}{n^2}\left(\sigma_1^2+\sigma_2^2+\sigma_3^2+\dots\dots\dots\dots\dots\dots+\sigma_n^2\right)$$

$$=\frac{1}{n\times\cancel{n}}\left(\cancel{n}\sigma^2\right)=\frac{\sigma^2}{n}$$

نظرية 3: إذا كان لدينا X متغير عشوائي يمثل مجتمع ما (محدود) وهو يخضع للتوزيع الطبيعي و $\overline{X}_i$ متغيرة عشوائية تمثل متوسط عينة مسحوبة من ذات المجتمع مع عدم الارجاع فان تباين توزيع المعاينة للمتوسطات $V\left(\overline{X}\right)$ يكتب كالتالي:

$$\left.\begin{array}{l} V\left(\overline{X}\right)=\sigma_{\overline{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1} \\ V\left(\overline{X}\right)=\sigma_{\overline{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{n} \end{array}\right\}\dots\dots\dots\frac{n}{N}>0.05$$

بحيث:

$\frac{n}{N}$: يسمى كسر المعاينة ويتم حسابة فقط في حالة السحب مع عدم الارجاع وحجم المجتمع محدود.

$\frac{N-n}{N-1}$ أو $\frac{N-n}{N}$ يسمى معامل الارجاع او معامل التصحيح وهو يستخدم فقط في حالة السحب مع عدم الارجاع.

التمرين 01: إذا كان العدد الكلي للأساتذة في 05 تخصصات لأحد الجامعات في بلد ما يخضع للتوزيع الطبيعي، والبيانات التالية توضح قيم المتغير العشوائي **X** الذي يمثل عدد المدرسين في هذه التخصصات: $X_1=16$ ، $X_2=22$ ، $X_3=34$ ، $X_4=42$ ، $X_5=56$

المطلوب:

1-أحسب المتوسط الحسابي للمجتمع μ وتباينه σ^2.

2-أوجد كل العينات ذات الحجم n=02 الممكن سحبها من هذا المجتمع في حالة السحب مع الإرجاع.

3-أوجد خطأ المعاينة ثم توزيع المعاينة للوسط الحسابي للعينة، ثم احسب وسطه الحسابي $\mu_{\overline{X}}$ وتباينه $\sigma^2_{\overline{X}}$ في حالة السحب مع الإرجاع.

4-أوجد كل العينات ذات الحجم n=02 الممكن سحبها من هذا المجتمع في حالة السحب بدون الإرجاع.

5- أوجد توزيع المعاينة للوسط الحسابي للعينة، ثم احسب وسطه الحسابي $\mu_{\overline{X}}$ وتباينه $\sigma^2_{\overline{X}}$ في حالة السحب بدون الإرجاع.

6- قارن بين الإجابتين في السؤال رقم 03 و05 ثم اشتق العلاقة بين كل من التوزيع العيني والمجتمع لكل من الوسط الحسابي والتباين في حالتي السحب بالإرجاع والسحب بدون إرجاع.

الحل:

1-حساب كل من المتوسط الحسابي للمجتمع μ وتباينه σ^2.

$$\mu=\frac{\sum_{i=01}^{05} X_i}{N}=\frac{16+22+34+42+56}{05}=34$$

$$\sigma^2=\frac{\sum_{i=01}^{05}(X_i-\mu)^2}{N}=\frac{(16-34)^2+(22-34)^2+(34-34)^2+(42-34)^2+(56-34)^2}{05}=203.20$$

2- إيجاد كل العينات ذات الحجم n=02 الممكن سحبها من هذا المجتمع في حالة السحب مع الإرجاع

في حالة السحب مع الارجاع فان حجم المجتمع N لا يتغير عند سحب مفردة من مفردات المجتمع، وهذا يعني ان سحب اي مفردة لا يتاثر بسحب المفردات السابقة، ومن ثم تكون المشاهدات مستقلة عن بعضها البعض ويمكن سحب المفردة أكثر من مرة لتكوين العينة، فاذا أخذنا جميع العينات الممكنة والمكونة من مفردتين والتي يمكن سحبها من هذا المجتمع بالارجاع فانها ستكون ترتيبة بتكرار مع تحقق الشروط التالية:

$$\left.\begin{array}{l}\text{-سحب جزء من الكل.}\\ \text{-الترتيب مهم.}\\ \text{-التكرار ممكن.}\\ \text{-السحب على التوالي.}\end{array}\right\} \begin{array}{c}\text{ترتيبه بتكرار (قائمة)}\\ N^n\end{array}$$

عدد العينات الممكنة هو $N^n = 05^{02} = 25$

والجدول رقم 01 يوضح مختلف العينات ذات الحجم n=02 الممكن سحبها من هذا المجتمع بالارجاع.

العينات الممكنة				
(16.56)	(16.42)	(16.34)	(16.22)	(16.16)
(22.56)	(22.42)	(22.34)	(22.22)	(22.16)
(34.56)	(34.42)	(34.34)	(34.22)	(34.16)
(42.56)	(42.42)	(42.34)	(42.22)	(42.16)
(56.56)	(56.42)	(56.34)	(56.22)	(56.16)

3–إيجاد خطأ المعاينة ثم توزيع المعاينة للوسط الحسابي للعينة، ثم احسب وسطه الحسابي $\mu_{\overline{X}}$ وتباينه $\sigma^2_{\overline{X}}$ في حالة السحب مع الإرجاع.

لايجاد خطا المعاينة نحسب اولا متوسط العينات الممكنة وعلى سبيل المثال نحسب متوسط العينةالاولى كالتالي $\overline{X}_1 = \frac{16+16}{02} = 16$ وهكذا لجميع القيم المتبقية والجدول رقم 02 يوضح ذلك.

متوسط العينات الممكنة				
36	29	25	19	16
39	32	28	22	19
45	38	34	28	25
49	42	38	32	29
56	49	45	39	36

اما خطأ المعاينة فيسم حسابة وفق العلاقة التالية: خطأ المعاينة= |إحصائية العينة–معلمة المجتمع|

والجدول رقم 03 يوضح خطا المعاينة

خطأ المعاينة				
36–34=02	29–34=–05	25–34=–09	19–34=–15	16–34=–18
39–34=05	32–34=–02	28–34=–06	22–34=–12	19–34=–15
45–34=11	38–34=04	34–34=00	28–34=–06	25–34=–09
49–34=15	42–34=08	38–34=04	32–34=–02	29–34=–05
56–34=22	49–34=15	45–34=11	39–34=05	36–34=02

توزيع المعاينة للمتوسط الحسابي للعينة هو التوزيع الطبيعي لان المجتمع يتوزع توزيعا طبيعيا.

–حساب الوسط الحسابي والتباين للعينة في حالة السحب مع الارجاع: من اجل حساب الوسط الحسابي والتباين نشكل الجدول رقم 04 انطلاقا من الجدول رقم 02 (متوسط العينات) بحيث نوضح فيه التكرارات الممكنة لمتوسط العينات المختلفة كالتالي:

n_i	$\overline{X}_i$	n_i	$\overline{X}_i$
02	36	01	16
02	38	02	19
02	39	01	22
01	42	02	25
02	45	02	28
02	49	02	29
01	56	02	32
25	**المجموع**	01	34

$$\mu_{\overline{X}}=\overline{\overline{X}}=\frac{\sum_{i=01}^{15} n_i\overline{X}_i}{\sum_{i=01}^{15} n_i}=\frac{16\times01+19\times02+\dots\dots\dots\dots+56\times01}{25}=\frac{850}{25}=34$$

أو نحسب المتوسط لحسابي لهذه المتوسطات كالتالي:

$$\overline{\overline{X}}=\frac{\sum_{i=1}^{25}\overline{X}_i}{n}=\frac{850}{25}=34$$

–حساب التباين

$$V\left(\overline{X}\right)=\sigma^2_{\overline{X}}=\frac{\sum_{i=01}^{15}n_i\left(\overline{X}_i-\mu\right)^2}{\sum_{i=01}^{15}n_i}=\frac{01\times(16-34)^2+02\times(19-34)^2+\dots\dots\dots\dots+01\times(56-34)^2}{25}$$

$$=\frac{2540}{25}=101.60$$

4–إيجاد كل العينات ذات الحجم$n=02$ الممكن سحبها من هذا المجتمع في حالة السحب بدون الإرجاع

في حالة السحب بدون إرجاع فان اي مفردة في المجتمع لا يمكن ان تسحب الا مرة واحدة لتكوين العينة وفي هذه الحالة توجد لدينا طريقتين لايجاد عدد العينات الممكنة

الطريقة الاولى: ترتيبة بدون تكرار وتتحقق بتوفر الشروط التالية

–سحب جزء من الكل.

–الترتيب مهم.

–التكرار غير ممكن.

–السحب على التوالي.

ترتيبه بدون تكرار P_N^n

عدد العينات الممكنة هو $P_N^n=\frac{N!}{(N-n)!}\Rightarrow P_5^2=\frac{05!}{(05-02)!}=\frac{05\times04\times03!}{03!}=20$

والجدول رقم 05 يوضح مختلف العينات ذات الحجم $n=02$ الممكن سحبها من هذا المجتمع بدون إرجاع.

العينات الممكنة				
(16.56)	(16.42)	(16.34)	(16.22)	-------
(22.56)	(22.42)	(22.34)	-------	(22.16)
(34.56)	(34.42)	-------	(34.22)	(34.16)
(42.56)	-------	(42.34)	(42.22)	(42.16)
-------	(56.42)	(56.34)	(56.22)	(56.16)

5-حساب الوسط الحسابي والتباين في حالة السحب بدون إرجاع

من اجل حساب الوسط الحسابي والتباين نحسب اولا متوسط العينات الممكنة والموضحة في الجدول رقم 06.

متوسط العينات الممكنة				
36	29	25	19	-------
39	32	28	-------	19
45	38	-------	28	25
49	-------	38	32	29
-------	49	45	39	36

بعد حساب المتوسطات الممكنة نستخرج التكرارات المقابلة لهذه المتوسطات وندونها في الجدول رقم 07

$\overline{X}_i$	n_i	$\overline{X}_i$	n_i
19	02	38	02
25	02	39	02
28	02	45	02
29	02	49	02
32	02	------	---------
36	02	**المجموع**	**20**

$$\mu_{\overline{X}} = \overline{X} = \frac{\sum_{i=01}^{10} n_i \overline{X}_i}{\sum_{i=01}^{10} n_i} = \frac{19\times02+25\times02+\dots\dots\dots\dots+49\times02}{20} = \frac{680}{20} = 34$$

كما يمكن حسابه بطريقة اخرى كالتالي:

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{10} \overline{X}_i}{n} = \frac{680}{20} = 34$$

-حساب التباين:

$$V(\overline{X})=\sigma_{\overline{X}}^2=\frac{\sum_{i=01}^{10} n_i\left(\overline{X}_i-\mu\right)^2}{\sum_{i=01}^{15} n_i}=\frac{02\times(19-34)^2+02\times(25-34)^2+\dots\dots\dots+02\times(49-34)^2}{20}$$

$$=\frac{1524}{20}=76.20$$

الطريقة الثانية: توفيقة بدون تكرار وتتحقق بتوفر الشروط التالية

-سحب جزء من الكل.
-الترتيب غير مهم.
-التكرار غير ممكن.
-السحب دفعة واحدة.

توفيقة بدون تكرار C_N^n

عدد العينات الممكنة هو

$$C_N^n=\frac{N!}{n!(N-n)!}\Rightarrow C_{05}^{02}=\frac{05!}{02!(05-02)!}=\frac{05\times04\times03!}{02!\times03!}=10$$

والجدول رقم 08 يوضح مختلف العينات ذات الحجم n=02 الممكن سحبها من هذا المجتمع بدون إرجاع.

العينات الممكنة				
(16.56)	(16.42)	(16.34)	(16.22)	-------
(22.56)	(22.42)	(22.34)	-------	-------
(34.56)	(34.42)	-------	-------	-------
(42.56)	-------	-------	-------	-------
-------	-------	-------	-------	-------

-حساب الوسط الحسابي والتباين في حالة السحب بدون إرجاع

من اجل حساب الوسط الحسابي والتباين نحسب اولا متوسط العينات الممكنة والموضحة في الجدول رقم 09.

متوسط العينات الممكنة				
36	29	25	19	-------
39	32	28	-------	-------
45	38	-------	-------	-------
49	-------	-------	-------	-------
-------	-------	-------	-------	-------

بعد حساب المتوسطات الممكنة نستخرج التكرارات المقابلة لهذه المتوسطات وندونها في الجدول رقم 10

$\overline{X}_i$	n_i	$\overline{X}_i$	n_i
19	01	38	01
25	01	39	01
28	01	45	01
29	01	49	01
32	01	------	--------
36	01	**المجموع**	**10**

$$\mu_{\overline{X}}=\overline{\overline{X}}=\frac{\sum_{i=01}^{10} n_i\overline{X}_i}{\sum_{i=01}^{10} n_i}=\frac{19\times01+25\times01+\dots\dots\dots\dots+49\times01}{20}=\frac{340}{10}=34$$

كما يمكن حسابه بطريقة اخرى كالتالي:

$$\overline{\overline{X}}=\frac{\sum_{i=1}^{10}\overline{X}_i}{n}=\frac{340}{10}=34$$

-حساب التباين:

$$V\left(\overline{X}\right)=\sigma_{\overline{X}}^2=\frac{\sum_{i=01}^{10} n_i\left(\overline{X}_i-\mu\right)^2}{\sum_{i=01}^{15} n_i}=\frac{01\times(19-34)^2+01\times(25-34)^2+\dots\dots\dots\dots+01\times(49-34)^2}{10}$$

$$=\frac{762}{10}=76.20$$

6- إشتقاق العلاقة بين كل من التوزيع العيني والمجتمع لكل من الوسط الحسابي والتباين في حالتي السحب بالإرجاع والسحب بدون إرجاع.

6-1 في حالة السحب مع الإرجاع

أ-بالنسبة للوسط الحسابي: وجدنا ان الوسط الحسابي للعينة في حالة السحب مع الارجاع هو نفسه الوسط الحسابي للمجتمع او بعبارة أخرى ان الوسط الحسابي للاوساط الحسابية للعينات يساوي الوسط الحسابي للمجتمع اي انه:

$$\begin{cases}\overline{\overline{X}}=\mu_{\overline{X}}=\mu=34\\ E\left(\overline{X}\right)=\mu_{\overline{X}}=\mu=34\end{cases}$$

ب-بالنسبة للتباين: وجدنا مما سبق ان تيابن العينة لا يساوي تباين المجتمع في حالة السحب مع الارجاع ولكن توجد علاقة بينهما الا وهي: $V\left(\overline{X}\right)=\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}=\frac{203.20}{02}=101.60$

وقد حصلنا على هذه القيمة (101.60) عند حساب تباين العينة لتوزيع المعاينة ل $\overline{X}$ وتكون هذه

العلاقة بين تباين توزيع المعاينة وبين تباين توزيع المجتمع صحيحة اذا تم سحب العينات بالارجاع من مجتمع محدود الحجم، حيث لا يؤثر سحب مفردة من المجتمع على سحب مفردة أخرى ومن ثم يكون هذين الحدثين مستقلين احصائيا وتكون العلاقة الاخيرة صحيحة ايضا في حالة المجتمعات اللانهائية الحجم سواء كان السحب بالارجاع او بدون ارجاع حيث لا يتاثر حجم المجتمع بسحب مفردة منه، ومن ثم يكون حدث سحب مفردة مستقلا عن حدث سحب مفردة اخرى ويتحقق هذا الاستقلال الاحصائي عندما يكون حجم المجتمع كبيرا بالنسبة لحجم العينة اي عندما $\frac{n}{N} \leq 0.05$

2-6 في حالة السحب مع عدم الارجاع

أ-**بالنسبة للوسط الحسابي:** الوسط الحسابي لتوزيع العينة في حالة السحب مع عدم الارجاع هو ايضا يساوي الوسط الحسابي للمجتمع اي ان

$$\begin{cases} \overline{\overline{X}} = \mu_{\overline{X}} = \mu = 34 \\ E\left(\overline{X}\right) = \mu_{\overline{X}} = \mu = 34 \end{cases}$$

ب-**بالنسبة للتبيابن:** وجدنا فيما سبق ان تباين العينة لا يساوي تباين المجتمع في حالة السحب مع عدم الارجاع ولاكن توجد علاقة بينهما وهي كالتالي:

$$V\left(\overline{X}\right) = \sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$$
$$= \frac{203.20}{02} \times \frac{05-20}{05-01} = 76.20$$

عندما يتم السحب مع عدم الارجاع فان سحب مفردة من المجتمع لتكوين العينة يؤثر على سحب اي مفردة أخرى وهنا يكون الحدثين غير مستقلين ويتحقق عدم الاستقلال الاحصائي ايضا عندما يكون حجم المجتمع صغيرا نسبيا بالنسبة لحجم العينة اي عندما يتحقق الشرط $\frac{n}{N} > 0.05$.

الخلاصة:

في حالة السحب بالارجاع او عدم الارجاع فان متوسط العينة يساوي دائما متوسط المجتمع اي

$$E\left(\overline{X}\right) = \mu_{\bar{x}} = \mu$$

-في حالة السحب بالارجاع او عدم الارجاع و $\frac{n}{N} \leq 0.05$ فان $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}$

-في حالة السحب مع عدم الارجاع و $\frac{n}{N} > 0.05$ فان $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

2-توزيع المعاينة للوسط الحسابي $\overline{X}$ من مجتمع يتبع التوزيع الطبيعي

2-1 عندما يكون التبيابن σ^2 معلوم

نظرية04: إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتبياين σ^2 وسحبنا منه كل العينات العشوائية البسيطة ذات الحجم n، فتوزيع المعاينة للوسط الحسابي للعينة $\overline{X}$ ستوزع توزيعا طبيعيا بمتوسط حسابي قدره $\mu_{\overline{X}}=\mu$ وتباين قدره $\sigma^2_{\overline{X}}=\frac{\sigma^2}{n}$ اي $\overline{X}\sim N\left(\mu_{\overline{X}},\sigma^2_{\overline{X}}\right)$ وسيتوزع هذا المتغير توزيعا طبيعيا معياريا تبعا للاحصاءة Z كما يلي:

$$Z=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}=\frac{\overline{X}_i-\mu}{\frac{\sigma^2}{n}}\sim N(0,1)$$

هذه النظرية صحيحة بغض النظر عن حجم العينة سواء كان كبيرا ام صغيرا.

أما اذا كان حجم المجتمع N محدود والسحب مع عدم الارجاع فتوزيع المعاينة للوسط الحسابي للعينة $\overline{X}$ ستوزع توزيعا طبيعيا بمتوسط حسابي قدره $\mu_{\overline{X}}=\mu$ وتباين قدره $\sigma^2_{\overline{X}}=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}$

اي $\overline{X}\sim N\left(\mu_{\overline{X}},\sigma^2_{\overline{X}}\right)$ وسيتوزع هذا المتغير توزيعا طبيعيا معياريا تبعا للاحصاءة Z كما يلي:

$$Z=\frac{\overline{X}-\mu_{\overline{X}}}{\sigma_{\overline{X}}}=\frac{\overline{X}-\mu}{\sqrt{\frac{\sigma^2}{n}\times\frac{N-n}{N-1}}}\sim(0,1)$$

<u>تمرين02:</u> إذا كان عدد سائقي سيارات الاجرة في مدينة ما هو 1500 سائق، وعلمت ان اعمارهم تتوزيع توزيعا طبيعيا بمتوسط حسابي قدره 45 سنة، وانحراف معياري قدره 07 سنوات، فاذا سحبنا مع عدم الارجاع من هذا المجتمع عينة عشوائية بها 16 سائقا.

<u>المطلوب:</u>

1-أوجد توزيع المعاينة لمتوسط اعمار سائقي سيارات الاجرة؟

2-أحسب احتمال ان يكون متوسط العمر لهذه العينة أكبر من 48 سنة؟

<u>الحل:</u>

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 45 سنة وانحراف معياري 07 سنوات وهو يمثل اعمار سائق سيارت الاجرة $X\sim N\left(45,(07)^2\right)$

$n=16, N=1500, \mu=45\ ,\ \sigma=07$ والسحب مع عدم الارجاع

1-إيجاد توزيع المعاينة لمتوسط اعمار سائقي سيارات الاجرة

بما ان المجتمع يخضع للتوزيع الطبيعي فان توزيع المعاينة سوف يخضع هو الاخر للتوزيع الطبيعي حسب النظرية رقم 04 $\overline{X}\sim N\left(\mu_{\overline{X}},\sigma^2_{\overline{X}}\right)$

حساب التوقع الرياضي والتباين لتوزيع المعاينة

المتوسط الحسابي لتوزيع المعاينة $E(\overline{X})=\mu_{\overline{x}}=\mu=45$

حساب التباين: بما ان المجتمع محدود والسحب مع عدم الارجاع (الشرطان الكافيان) محققان ننتقل اذن الى الشرط اللازم وهو كسر المعاينة

نحسب اولا كسر المعاينة $\frac{n}{N}=\frac{16}{1500}=0.0106\leq 0.05$، إذن لا نستخدم معامل الارجاع .

$$\sigma_{\overline{X}}^{2}=\frac{\sigma^{2}}{n}=\frac{49}{16}=03.0625$$

إذن $\overline{X}\sim N(45,03.0625)$

2-حساب احتمال ان يكون متوسط العمر لهذه العينة أكبر من 48 سنة: بما ان توزيع المعاينة يخضع للتوزيع الطبيعي ولحساب الاحتمال السابق ننتقل الى القيمة المعيارية او الاحصاءة Z

$$Z=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\sim N(0,1)$$

$$P(\overline{X}>48)=P\left(Z>\frac{48-45}{\sqrt{03.0625}}\right)=P(Z>01.71)$$
$$=1-P(Z<01.71)=1-F(01.71)$$
$$=1-0.95637=0.04363$$

<u>تمرين03</u>:إذا علمت ان معدلات ان معدلات 420 طالب في امتحان الاحصاء الاستنتاجي تتوزع توزيعا طبيعيا بمتوسط حسابي قدره68 وتباين قدره 25، فاذا سحبنا من هؤلاء الطلبة مع عدم الارجاع عينة عشوائية تشمل 100 طالب.

<u>المطلوب</u>

1-أوجد توزيع المعاينة لمعدلات الطلبة؟

2-أحسب احتمال ان يكون المتوسط الحسابي لهذه العينة بين 67 و69؟

<u>الحل</u>

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي قدره 68 وتباين قدره 25 وهو يمثل معدلات الطلبة $X\sim N(68,25)$

السحب مع عدم الارجاع $N=420,\mu=68,\sigma^{2}=25,n=100$

1-إيجاد توزيع المعاينة لمعدلات الطلبة: بما ان المجتمع يتوزع توزيعا طبيعيا فان توزيع المعاينة لمعدلات الطلبة في العينة سيتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي وتباين كالتالي:

$$\overline{X}\sim N\left(\mu_{\overline{X}},\sigma_{\overline{X}}^{2}\right)$$

$$E(\overline{X})=\mu_{\overline{x}}=\mu=68$$

حساب التباين: بما ان المجتمع محدود والسحب مع عدم الارجاع (الشرطان الكافيان محققان) ننتقل الى الشرط اللازم وهو كسر المعاينة

نحسب كسر المعاينة $\frac{n}{N}=\frac{100}{420}=0.238>0.05$، إذن الشرط محقق وبالتالي نستخدم معامل الارجاع.

$$\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}=\frac{25}{100}\times\frac{420-100}{420-1}=0.1909$$

$$\overline{X}\sim N(68,0.1909)$$

2-حساب احتمال ان يكون المتوسط الحسابي لهذه العينة بين 67 و69

بما ان توزيع المعاينة يخضع للتوزيع الطبيعي ولحساب الاحتمال السابق ننتقل الى القيمة المعيارية او الاحصاءة Z

$$Z=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\sim N(0,1)$$

$$P(67<\overline{X}<69)=P\left(\frac{67-68}{\sqrt{0.1909}}<Z<\frac{69-68}{\sqrt{0.1909}}\right)=P(-02.29<Z<02.29)$$
$$=F(02.29)-F(-02.29)=F(02.29)-\left[1-F(02.29)\right]$$
$$=02F(02.29)-01=02\times0.98899-1=0.97798$$

<u>**تمرين04:**</u> مصنع للمصابيح الكهربائية يفترض ان اعمار المصابيح تتوزع توزيعا طبيعيا بمتوسط حسابي قدره 800ساعة وانحراف معياري قدره 40 ساعة، اخذت عينة عشوائية حجمها36 مصباح.

<u>**المطلوب:**</u>

1-أوجد توزيع المعاينة لمتوسط اعمار المصابيح؟

2-أحسب الاحتمالات التالية

2-1 ان يكون متوسط العمر اقل من 780 ساعة؟

2-2 ان يكون متوسط العمر اكبر من 810 ساعة؟

2-3 ان يكون متوسط العمر محصور بين 790 ساعة و810 ساعة؟

<u>**الحل**</u>

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 800 وانحراف معياري 40 وهو يمثل اعمار المصابيح الكهربائية.

$n=36$، $\sigma=40$، $\mu=800$، $N=\infty$(مجتمع لا محدود)

1-إيجاد توزيع المعاينة لمتوسط اعمار المصابيح: بما ان المجتمع يتوزع توزيعا طبيعيا فان توزيع المعاينة لمتوسط اعمار المصابيح سيتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي وتباين كالتالي:

$$\overline{X} \sim N\left(\mu_{\overline{X}}, \sigma^2_{\overline{X}}\right)$$

$$E\left(\overline{X}\right) = \mu_{\overline{x}} = \mu = 800$$

حساب التباين: من خلال معطيات التمرين نلاحظ ان المجتمع لا محدود وبالتالي لا نستخدم معامل الارجاع ويحسب التباين كالتالي: $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} = \frac{(40)^2}{36} = 44.44$

$$\overline{X} \sim N(800, 44.44)$$

2-حساب الاحتمالات

2- 1 $P\left(\overline{X} < 780\right) = P\left(Z < \frac{780-800}{\sqrt{44.44}}\right) = P(Z < -03) = F(-03) = 01 - F(03)$

$= 01 - 0.99865 = 0.00135$

2-2 $P\left(\overline{X} > 810\right) = P\left(Z > \frac{810-800}{\sqrt{44.44}}\right) = P(Z > 01.50) = 01 - P(Z < 01.50)$

$= 01 - F(01.50) = 01 - 0.93319 = 0.06681$

$P\left(790 < \overline{X} < 810\right) = P\left(\frac{790-800}{\sqrt{44.44}} < Z < \frac{810-800}{\sqrt{44.44}}\right) = P(-01.50 < Z < 01.50)$

2-3 $= F(01.50) - F(-01.50) = F(01.50) - \left[1 - F(01.50)\right]$

$= 02F(01.50) - 01 = 02 \times 0.93319 - 01 = 0.86638$

ملاحظة هامة:

-إذا كان السحب مع الارجاع فهذا يعني ان التوزيع هو توزيع ذي الحدين $X \sim BEN(np, npq)$

-إذا كان السحب مع عدم الارجاع فهذا يعني ان التوزيع هو توزيع مافوق الهندسي $X \sim HG\left(np, npq \times \frac{N-n}{N-1}\right)$

عند كبر حجم المجتمع وذهابه الى ما لانهاية، يقترب التوزيع مافوق الهندسي من التوزيع ذي الحدين ،اي يتساوى السحب مع عدم الارجاع مع السحب مع الارجاع وبالتالي يصبح معامل الارجاع مساويا الى الواحد الصحيح $\lim\limits_{N \to \infty} \frac{N-n}{N-1} = 1$

وبما ان حجم المجتمع في التمرين كبير جدا لم يذكر خاصية السحب (بالارجاع او عدم الارجاع) لانهما سوف يتساويان.

2-2 عندما يكون التباين σ^2 مجهول و$(n\geq30)$

في كثير من التطبيقات المتعلقة باستعمال الوسط الحسابي $\overline{X}$ لا يكون الانحراف المعياري للمجتمع σ معلوما ، ولذلك اذا كان حجم العينة n يساوي او اكثر من 30 فاننا نستعمل S(الانحراف المعياري للعينة) بدلا من الانحراف المعياري للمجتمع σ، ويكون المتغير العشوائي $\overline{X}$ خاضعا لتوزيع يمكننا تقريبه الى التوزيع الطبيعي المعياري(Z) مادامت$(n\geq30)$

نظرية05: إذا كانت لدينا $x_1,x_2,x_3......x_n$ عينة عشوائية بسيطة حجمها $(n\geq30)$ ومأخوذة من مجتمع يخضع للتوزيع الطبيعي $X\sim N\left(\mu,\sigma^2\right)$ وكانت σ^2 مجهولة، وكانت S^2 هي تباين العينة فان المتغير العشوائي $\overline{X}$سيكون له توزيع قريب من للتوزيع الطبيعي المعياري بمتوسط حسابي قدره $\mu_{\overline{X}}=\mu$ وتباين قدره $\sigma^2_{\overline{X}}=\frac{\sigma^2}{n}\simeq\frac{\hat{S}^2}{n}$ او بشكل مكافئ $\sigma^2_{\overline{X}}=\frac{\sigma^2}{n}\simeq\frac{S^2}{n-1}$ وتعطى علاقته كالتالي:

$$Z=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\simeq\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{\hat{S}}{\sqrt{n}}}=\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{S}{\sqrt{n-1}}}\sim N\left(0,1\right)$$

أما اذا كان المجتمع (N) محدود والسحب مع عدم الارجاع و $\frac{n}{N}>0.05$ فنستبدل العلاقات السابقة باضافة معامل الارجاع وتصبح العلاقات كالتالي:

$$\sigma^2_{\overline{X}}=\begin{cases}\frac{\sigma^2}{n}\simeq\frac{S^2}{n-1}\times\frac{N-n}{n-1}\\[2mm]\frac{\sigma^2}{n}\simeq\frac{\hat{S}^2}{n}\times\frac{N-n}{n-1}\end{cases}$$

اما المتغير العشوائي (Z) فتصبح علاقته كالتالي:

$$Z=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\simeq\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{\hat{S}}{\sqrt{n}}\times\sqrt{\frac{N-n}{n-1}}}=\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{S}{\sqrt{n-1}}\times\sqrt{\frac{N-n}{n-1}}}\sim N\left(0,1\right)$$

تمرين05: إذا كان X متغير عشوائي يمثل أطوال الطلاب لجامعة ابن خلدون، وكان هذا المتغير يخضع للتوزيع الطبيعي بوسط حسابي160سم وتباين مجهول، اخذت عينة عشوائية حجمها 37 طالب وكان تباين العينة هو60سم.

<u>المطلوب:</u>

1-أوجد توزيع المعاينة لمتوسط أطوال الطلبة؟

2-أحسب احتمال ان يكون طول الطلبة في هذه العينة اقل من 163سم؟

<u>الحل:</u>

X: متغير عشوائي يخضع للتوزيع الطبيعي بمتوسط حسابي 160سم وتباين مجهول، وهو يمثل اطوال

الطلبة لجامعة ابن خلدون بتيارت $X \sim N(160, \sigma^2)$

$N=\infty$(مجتمع لا محدود)، $\mu=160$ ،σ^2 ،$S^2=60$، $n=37$

1-إيجاد توزيع المعاينة لمتوسط اطوال الطلبة: بما ان المجتمع يخضع للتوزيع الطبيعي وتباينه مجهول وحجم العينة المسحوبة اكبر تماما من 30 فان توزيع المعاينة لمتوسط اطوال الطلبة سيكون توزيع قريب من التوزيع الطبيعي المعياري $\overline{X} \sim N\left(\mu_{\overline{X}}, \sigma^2_{\overline{X}}\right)$

$$E\left(\overline{X}\right) = \mu_{\overline{X}} = \mu = 160$$

حساب التباين: من خلال معطيات التمرين نلاحظ ان المجتمع لا محدود وبالتالي لا نستخدم معامل الارجاع ويحسب التباين كالتالي: $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} \simeq \frac{S^2}{n-1} = \frac{60}{37-1} = 01.67$

كما نعلم ان العلاقة بين تباين العينة المتحيز والعينة غير المتحيز تعطى بالعلاقة التالية:

$$nS^2 = (n-1)\hat{S}^2 \Rightarrow \hat{S}^2 = \frac{nS^2}{(n-1)} = \frac{37 \times 60}{(37-1)} = 61.67$$

إذن $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} \simeq \frac{\hat{S}^2}{n} = \frac{61.67}{37} = 01.67$

وبالتالي فان $\overline{X} \sim N(160, 01.67)$

2- حساب احتمال ان يكون طول الطلبة في هذه العينة اقل من 163سم

بتطبيق علاقة التوزيع الطبيعي المعياري $Z = \frac{\overline{X}_i - \mu_{\overline{x}}}{\sigma_{\overline{x}}} \simeq \frac{\overline{X}_i - \mu_{\overline{x}}}{\frac{\hat{S}}{\sqrt{n}}} = \frac{\overline{X}_i - \mu_{\overline{x}}}{\frac{S}{\sqrt{n-1}}} \sim N(0,1)$

$$P\left(\overline{X} < 163\right) = P\left(Z < \frac{163-160}{\sqrt{01.67}}\right) = P(Z < 02.32) = F(02.32) = 0.98983$$

3-2 عندما يكون التبيان σ^2 مجهول و(n<30)

في الحالات التي يكون فيها حجم العينة n اقل تماما من 30 فان الانحرافات المعيارية للعينة تكون ذات تغير كبير لدرجة ان S لا تكون تقديرا موثوقا للانحراف المعياري للمجتمع σ وبالتالي يكون المتغير العشوائي $\overline{X}$ خاضعا لتوزيع لا يمكننا تقريبه الى التوزيع الطبيعي المعياري (Z)، وانما يخضع لتوزيع آخر يطلق عليه توزيع ستودنت (او توزيع المعاينات الصغيرة) بدرجة حرية (V=n-1) وبالمعلمتين $X \sim T_V\left(0, \frac{V}{V-2}\right)$

نظرية 06: إذا كانت لدينا $x_1, x_2, x_3 \ldots\ldots x_n$ عينة عشوائية بسيطة حجمها (n<30) ومأخوذة من مجتمع يخضع للتوزيع الطبيعي $X \sim N(\mu, \sigma^2)$ وكانت σ^2 مجهولة، وكانت S^2 هي تباين العينة فان المتغير العشوائي $\overline{X}$ لن يخضع للتوزيع الطبيعي المعياري وانما لتوزيع آخر هو توزيع ستودنت(T) بدرجة

حرية $(V=n-1)$ ، اما معلمتي توزيع المتغير العشوائي ل $\overline{X}$ فهما $\mu_{\overline{X}}=\mu$ وتباين قدره $\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}\simeq\frac{\hat{S}^2}{n}$ او بشكل مكافئ $\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}\simeq\frac{S^2}{n-1}$ وتعطى علاقته كالتالي:

$$T=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\simeq\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{\hat{S}}{\sqrt{n}}}=\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{S}{\sqrt{n-1}}}$$

أما اذا كان المجتمع محدود والسحب مع عدم الارجاع و $\frac{n}{N}>0.05$ فنستبدل العلاقات السابقة باضافة معامل الارجاع وتصبح العلاقات كالتالي:

$$\sigma_{\overline{X}}^2=\begin{cases}\frac{\sigma^2}{n}\simeq\frac{S^2}{n-1}\times\frac{N-n}{n-1}\\[2mm]\frac{\sigma^2}{n}\simeq\frac{\hat{S}^2}{n}\times\frac{N-n}{n-1}\end{cases}$$

اما المتغير العشوائي (T) فتصبح علاقته كالتالي:

$$T=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\simeq\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{\hat{S}}{\sqrt{n}}\times\sqrt{\frac{N-n}{n-1}}}=\frac{\overline{X}_i-\mu_{\overline{X}}}{\frac{S}{\sqrt{n-1}}\times\sqrt{\frac{N-n}{n-1}}}$$

تمرين6: إذا كان X متغير عشوائي يمثل اوزان اكياس الطحين التي تنتجها احدى المؤسسات، وكان هذا المتغير يخضع للتوزيع الطبيعي بمتوسط حسابي 50 كغ وتباين مجهول، اخذت عينة عشوائية بسيطة حجمها 09 اكياس من انتاج هذه المؤسسة ووجد ان الانحراف المعياري لاوزان هذه الاكياس يساوي 01كغ.

المطلوب:

1-أوجد توزيع المعاينة لمتوسط اوزان الاكياس؟

2-أحسب احتمال ان يزيد الوسط الحسابي لهذه العينة عن 51كغ؟

الحل:

X: متغير عشوائي يخضع للتوزيع الطبيعي بمتوسط حسابي 50كغ وتباين مجهول، وهو يمثل اوزان اكياس الطحين التي تنتجها احدى المؤسسات.

$n=09$ ،$S^2=01$ ،σ^2 ،$\mu=50$ ،(مجتمع لا محدود)$N=\infty$

1-إيجاد توزيع المعاينة لمتوسط اوزان الاكياس: بما ان المجتمع يخضع للتوزيع الطبيعي وتباينه مجهول وحجم العينة العشوائية البسيطة المسحوبة اقل تماما من 30 فان توزيع المعاينة سوف يخضع لتوزيع ستودنت (T) بدرجة حرية $(V=09-01=08)$ أي $\overline{X}\sim T_V\left(\mu_{\overline{X}},\sigma_{\overline{X}}^2\right)$

ايجاد معلمتي التوزيع اي التوقع الرياضي والتباين

التوقع الرياضي يساوي $E\left(\overline{X}\right)=\mu_{\overline{x}}=\mu=50$

حساب التباين: من خلال معطيات التمرين نلاحظ ان المجتمع لا محدود وبالتالي لا نستخدم معامل الارجاع ويحسب التباين كالتالي: $\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}\simeq\frac{S^2}{n-1}=\frac{01}{09-1}=0.125$

وبالتالي فان $\overline{X}\sim T_{08}\left(50{,}0.125\right)$

2-حساب احتمال ان يزيد الوسط الحسابي لهذه العينة عن 51كغ

بتطبيق علاقة توزيع ستودنت $T=\frac{\overline{X}_i-\mu_{\overline{x}}}{\sigma_{\overline{x}}}\simeq\frac{\overline{X}_i-\mu_{\overline{x}}}{\frac{\hat{S}}{\sqrt{n}}\times\sqrt{\frac{N-n}{n-1}}}=\frac{\overline{X}_i-\mu_{\overline{x}}}{\frac{S}{\sqrt{n-1}}\times\sqrt{\frac{N-n}{n-1}}}$

$$P\left(\overline{X}<51\right)=P\left(T_{08}<\frac{51-50}{\sqrt{0.125}}\right)=P\left(T_{08}<02.828\right)=F\left(02.828\right)=0.988$$

بالعودة الى جدول ستودنت والبحث عن قيمة الاحتمال المقابل للقيمة 02.32 بدرجة حرية 08 لا نجد هذه القيمة وانما تكون محصورة بين قيمتين ولايجاد قيمة الاحتمال المقابل لها نستخدم نظرية الحصر بالنقصان او الزيادة كالتالي:

02.306	02.828	02.896
0.975	α	0.99

باستخدام نظرية الحصر بالنقصان:

$$(02.896-02.306)\longrightarrow(0.99-0.975)$$
$$(02.828-02.306)\longrightarrow(\alpha-0.975)$$

إذن

$$(00.59)\longrightarrow(0.015)$$
$$(0.522)\longrightarrow(\alpha-0.975)$$

إذن قيمة α هي:

$$\alpha=\frac{0.522\times0.015}{0.59}+0.975=0.988$$

3- توزيع المعاينة للوسط الحسابي $\overline{X}$ من مجتمع لا يتبع التوزيع الطبيعي

قد يحدث في كثير من الاحيان ان يكون المجتمع الذي تسحب منه العينات ليس معتدلا، فقد يكون ملتويا نحو اليمين او نحو اليسار في هذه الحالة نطبق نظرية النهاية المركزية.

إن اول من عرض هذه النظرية هو لابلاس عام 1812م وبرهنها ليبانوف 1901م،إن اهمية هذه النظرية تتاتى من كونها إعتمدت متغيرات عشوائية مختلفة، بصرف النظر عن طبيعة المجتمع

الاحصائي الذي جاءت منه (طبيعيا كان ام غير ذلك)إذ يكفي ان يكون لهذه المتغيرات العشوائية توقع رياضي وانحراف معياري ايا كان شكل توزيعها.

ونظرية النهاية المركزية لها أهمية كبيرة في الاستنتاج الاحصائي، فباستخدامها نعتبر توزيع المعاينة للوسط الحسابي للعينة توزيعا قريبا من التوزيع الطبيعي، وبالتالي نستطيع استخدام التوزيع الطبيعي في استنتاجات خاصة بالوسط الحسابي للمجتمع دون الحاجة الى معرفة التوزيع الاحتمالي لهذا المجتمع، بشرط ان يكون حجم العينة كبيرا **(n≥30)**، وكلما زاد حجم العينة كلما اقترب توزيع المعاينة للوسط الحسابي للعينة أكثر من التوزيع الطبيعي مهما كان توزيع المجتمع الذي سحبنا منه العينات.

نظرية 06 (نظرية النهاية المركزية):إذا كان لدينا مجتمع توزيعه لا يتبع التوزيع الطبيعي (مجهول) وسطه الحسابي μ وتبياين σ^2 وسحبنا منه كل العينات العشوائية البسيطة والممكنة وذات الحجم الكبير **(n≥30)** فتوزيع المعاينة للوسط الحسابي للعينة $\overline{X}$سيتوزع توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي قدره $\mu_{\overline{X}}$ وتباين قدره $\sigma^2_{\overline{X}}$ بعبارة أخرى فإن $Z=\frac{\overline{X}_i-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\sim(0,1)$ يقترب من التوزيع الطبيعي المعياري كلما كبرت n.

تمرين 07: في احد المصانع كان الوسط الحسابي لاجور العمال 300 ون وانحراف معياري 50 ون، وقد تم سحب عينة عشوائية من 40 عامل من عمال هذا المصنع.

المطلوب:

1-أوجد توزيع المعاينة للوسط الحسابي لاجور العمال؟

2-أحسب الاحتمالات التالية:

2-1إحتمال ان يكون الوسط الحسابي للعينة اقل من 275 ون؟

2-2 إحتمال ان يكون الوسط الحسابي للعينة أكبر من 320 ون؟

2-3 إحتمال ان يكون الوسط الحسابي للعينة بين 280 ون و315 ون؟

الحل:

X: متغير عشوائي مستمر يخضع لتوزيع مجهول بمتوسط حسابي 300 ون وانحراف معياري 50 ون، وهو يمثل متوسط اجور العمال لاحد المصانع $X\sim In(300,2500)$

n=40، $\sigma=50$، $\mu=300$، (مجتمع لا محدود)N=∞

1-إيجاد توزيع المعاينة: بما ان توزيع مجتمع اجور العمال مجهول، ونظرا لان حجم العينة كبير $(n=40\Rightarrow n\geq 30)$، إذا يمكننا تطبيق نظرية النهاية المركزية، ومن ثم فان توزيع المعاينة ل $\overline{X}$ سيكون معتدلا تقريبا اي $\overline{X}\sim N\left(\mu_{\overline{X}},\sigma^2_{\overline{X}}\right)$

$$E\left(\overline{X}\right)=\mu_{\overline{x}}=\mu=300$$

حساب التباين: من خلال معطيات التمرين نلاحظ ان المجتمع لا محدود وبالتالي لا نستخدم معامل الارجاع ويحسب التباين كالتالي:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} = \frac{(50)^2}{40} = 62.50$$

إذن

$$\bar{X} \sim N(300, 62.50)$$

2-حساب الاحتمالات

2- 1

$$P(\bar{X} < 275) = P\left(Z < \frac{275-300}{\sqrt{62.50}}\right) = P(Z < -03.16) = F(-03.16) = 01 - F(03.16)$$

$$= 01 - 0.99921 = 0.00079$$

2-2

$$P(\bar{X} > 320) = P\left(Z > \frac{320-300}{\sqrt{62.50}}\right) = P(Z > 02.53) = 01 - P(Z < 02.53)$$

$$= 01 - F(02.53) = 01 - 0.99430 = 0.00570$$

$$P(280 < \bar{X} < 315) = P\left(\frac{280-300}{\sqrt{62.50}} < Z < \frac{315-300}{\sqrt{62.50}}\right) = P(-02.53 < Z < 01.90)$$

2-3

$$= F(01.90) - F(-02.53) = F(01.90) - \left[1 - F(02.53)\right]$$

$$= F(01.90) + F(02.53) - 01$$

$$= 0.97128 + 0.99430 - 01 = 0.96558$$

04-توزيع المعاينة للفرق بين وسطين حسابيين لعينتين $(\bar{X}_1 - \bar{X}_2)$

تمهيد: في كثير من الدراسات يرغب الباحث في مقارنة مجتمعين لمعرفة الاختلاف بينهما، وذلك بمقارنة متوسطيهما اي بحساب الفرق بين وسطيهما الحسابيين $(\mu_1 - \mu_2)$ حيث μ_1 ترمز للوسط الحسابي للمجتمع و μ_2ترمز للوسط الحسابي للمجتمع الثاني، فمثلا يقارن الباحث بين أوزان الوحدات المنتجة بالمصنع "أ" والوسط الحسابي لاوزان الوحدات المنتجة بالمصنع "ب"، او يقارن مستوى الاسعار في سوقين مختلفين بحساب الفرق بين الوسط الحسابي لاسعار السوق الاول، والوسط الحسابي لاسعار السوق الثاني وهكذا....الخ،وحيث انه في أغلب الدراسات تكون بيانات المجتمعين غير متوافرة وبالتالي يكون μ_1 و μ_2 مجهولين، ولاجراء مثل هذه الدراسات نحاول استنتاج الفرق الحقيقي بين متوسطي المجتمعين $(\mu_1 - \mu_2)$، باستخدام الفرق بين متوسطي حسابيين لعينتين $(\bar{X}_1 - \bar{X}_2)$، حيث العينة الاولى مسحوبة من المجتمع الاول، والعينة الثانية مسحوبة من المجتمع الثاني، ويرمز للمجتمع الاول بالرقم 01 وللمجتمع الثاني بالرقم02، وبالتالي فان $\bar{X}_1$ يرمز للوسط الحسابي لعينة مسحوبة من المجتمع الاول، و $\bar{X}_2$يرمز للوسط الحسابي لعينة مسحوبة من المجتمع الثاني وبما ان الاحصائيتين $\bar{X}_1$و $\bar{X}_2$ متغيران عشوائيين، إذن الفرق بينهما $(\bar{X}_1 - \bar{X}_2)$هي الاخرى متغير عشوائي ولاجراء اي استنتاجات خاصة بالمعلمة $(\mu_1 - \mu_2)$ باستخدام الاحصائية $(\bar{X}_1 - \bar{X}_2)$ يجب معرفة ودراسة طبيعة توزيع المعاينة لهذه

الاحصائية، والذي يطلق عليه توزيع المعاينة للفرق بين متوسطين حسابيين.

4-1 توزيع المعاينة للفرق بين وسطين حسابيين لعينتين مستقلتين $\left(\overline{X}_1-\overline{X}_2\right)$

إذا تمت المعاينة من مجتمعين مستقلين، ولمعرفة توزيع المعاينة للفرق ، فان ذلك يتوقف على طبيعة توزيع المجتمعين(توزيع طبيعي او غير طبيعي) وكذلك معلومية او مجهولية تبايني المجتمعين وحجم العينتين المسحوبيتين اكبر أو يساوي 30 او اصغر تماما من 30 وسوف نتطرق الى مختلف هذه الحالات كالتالي:

أولا:توزيع المعاينة للفرق $\left(\overline{X}_1-\overline{X}_2\right)$ لمجتمعين يتوزعان توزيعا طبيعيا وبتباينين معلومين σ_1^2 و σ_2^2

نظرية08: إذا سحبنا كل العينات العشوائية الممكنة ذات الحجم n_1 من مجتمع لامحدود يتوزع توزيعا طبيعيا بوسط حسابي μ_1 وتباين σ_1^2، وسحبنا كل العينات العشوائية الممكنة ذات الحجم n_2 من مجتمع آخر لا محدود هو الاخر يتوزع توزيعا طبيعيا بوسط حسابي μ_2 وتباين σ_2^2، وكانت العينات المسحوبة من المجتمع الاول مستقلة عن العينات المسحوبة من المجتمع الثاني فان توزيع المعاينة للفرق بين المتوسطات $\left(\overline{X}_1-\overline{X}_2\right)$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي هو $\mu_{\overline{X}_1-\overline{X}_2}$ و تباين $\sigma^2_{\overline{X}_1-\overline{X}_2}$ اي $\overline{X}_1-\overline{X}_2 \sim N\left(\mu_{\overline{X}_1-\overline{X}_2},\sigma^2_{\overline{X}_1-\overline{X}_2}\right)$

ولايجاد معلمتي التوزيع نستعين بخصائص التوقع والتباين .

التوقع الرياضي:
$$E\left(\overline{X}_1-\overline{X}_2\right)=E\left(\overline{X}_1\right)-E\left(\overline{X}_2\right)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$

التباين :
$$V\left(\overline{X}_1-\overline{X}_2\right)=V\left(\overline{X}_1\right)+V\left(\overline{X}_2\right)=\sigma^2_{\overline{X}_1}+\sigma^2_{\overline{X}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$$

اما المتغير العشوائي (Z) فتصبح علاقته كالتالي:
$$Z=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim(0,1)$$

أما ذا كان احد المجتمعين او كليهما لا يتوزع توزيعا طبيعيا، فيكون توزيع المعاينة للفروق قريبا من التوزيع الطبيعي اذا كانت العينتان كبيرتين، اي ($n_1\geq30$،$n_2\geq30$) حسب نظرية النهاية المركزية.

تمرين 08: عينة عشوائية حجمها 20 سحبت من مجتمع يتوزع توزيعا طبيعيا بمعدل 50 وتباين 09، وعينة أخرى حجمها15 سحبت من مجتمع آخر يتوزع توزيعا طبيعيا بمعدل 40 وتباين 04، وكانت العينتان المسحوبتان مستقلتين عن بعضهما البعض.

المطلوب:

1-أوجد توزيع المعاينة للفرق بين متوسطي العينتين؟

2-أحسب احتمال ان يكون الفرق بين متوسطي العينتين اقل من 08.20؟

الحل:

1-إيجاد توزيع المعاينة للفرق بين متوسطي العينتين

المجتمع الاول

X_1:متغير عشوائي مستمر يخض للتوزيع الطبيعي بمتوسط حسابي 50 وتباين 09 اي $X_1 \sim N(50,09)$

ولدينا: $N_1=\infty$ ،$n_1=20$ ،$\mu_1=50, \sigma_1^2=09$

المجتمع الثاني

X_2: متغير عشوائي مستمر يخض للتوزيع الطبيعي بمتوسط حسابي 40 وتباين 04 اي $X_2 \sim N(40,04)$

ولدينا: $N_2=\infty$ ،$n_2=15$ ،$\mu_2=40, \sigma_2^2=04$

بما ان المجتمع الاول يخضع للتوزيع الطبيعي فان توزيع المعاينة له $\overline{X}_1$سوف يخضع هو الاخر للتوزيع الطبيعي، والمجتمع الثاني يخضع للتوزيع الطبيعي فان توزيع المعاينة له $\overline{X}_2$سوف يخضع هو الاخر للتوزيع الطبيعي، وبالتالي فان الفرق $(\overline{X}_1-\overline{X}_2)$هو متغير عشوائي جديد سوف يكون له توزيع قريب من التوزيع الطبيعي كالتالي: $\overline{X}_1-\overline{X}_2 \sim N\left(\mu_{\overline{X}_1-\overline{X}_2}, \sigma^2_{\overline{X}_1-\overline{X}_2}\right)$

التوقع الرياضي:

$$E(\overline{X}_1-\overline{X}_2)=E(\overline{X}_1)-E(\overline{X}_2)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$
$$=50-40=10$$

التباين:

$$V(\overline{X}_1-\overline{X}_2)=V(\overline{X}_1)+V(\overline{X}_2)=\sigma^2_{\overline{X}_1}+\sigma^2_{\overline{X}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$$
$$=\frac{09}{20}+\frac{04}{15}\simeq 0.72$$

إذن $\overline{X}_1-\overline{X}_2 \sim N(10,0.72)$

2-حساب إحتمال ان يكون الفرق بين متوسطي العينتين اقل من 08.20

بتطبيق الاحصاءة (Z) للتوزيع الطبيعي المعياري

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

$$P(\overline{X}_1-\overline{X}_2<08.20)=P\left(Z<\frac{08.20-10}{\sqrt{0.72}}\right)=P(Z<-02.12)=F(-02.12)=01-F(02.12)$$
$$=01-0.98300=0.0170$$

نظرية09: ليكن لدينا مجتمع أول حجمه N_1 يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم n_1 مع عدم الارجاع مع تحقق الشرط $\left(\frac{n_1}{N_1}>0.05\right)$وحسبنا كل المتوسطات الحسابية لها ولتكن $\overline{X}_1$ ،وليكن لدينا مجتمع ثاني محدود حجمه N_2 يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 وسحبنا منه كل العينات العشوائية

الممكنة ذات الحجم n_2 مع عدم الارجاع مع تحقق الشرط $\left(\frac{n_2}{N_2}>0.05\right)$ ، وحسبنا كل المتوسطات لها ولتكن $\overline{X}_2$، فان توزيع المعاينة للفرق بين المتوسطات $\left(\overline{X}_1-\overline{X}_2\right)$ سيتوزع توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي $\mu_{\overline{X}_1-\overline{X}_2}$ و تباين $\sigma^2_{\overline{X}_1-\overline{X}_2}$ أي $\overline{X}_1-\overline{X}_2\sim N\left(\mu_{\overline{X}_1-\overline{X}_2},\sigma^2_{\overline{X}_1-\overline{X}_2}\right)$

ولايجاد معلمتي التوزيع نستعين بخصائص التوقع والتباين .

التوقع الرياضي: $E\left(\overline{X}_1-\overline{X}_2\right)=E\left(\overline{X}_1\right)-E\left(\overline{X}_2\right)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$

التباين : $V\left(\overline{X}_1-\overline{X}_2\right)=V\left(\overline{X}_1\right)+V\left(\overline{X}_2\right)=\sigma^2_{\overline{X}_1}+\sigma^2_{\overline{X}_2}=\frac{\sigma_1^2}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{\sigma_2^2}{n_2}\times\frac{N_2-n_2}{N_2-1}$

اما المتغير العشوائي (Z) فتصبح علاقته كالتالي: $Z=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{\sigma_2^2}{n_2}\times\frac{N_2-n_2}{N_2-1}}}\sim(0,1)$

تمرين09: مجتمعان يتوزعان توزيعا طبيعيا الاول حجمه 1000 وسطه الحسابي 30 وتباينه 18، والثاني حجمه 1400 وسطه الحسابي 24 وتباينه 40، فإذا سحبنا من كل مجتمع عينة عشوائية مع عدم الارجاع وكانت العينتان مستقلتين، وكان حجم العينة الاولى 120 وحجم العينة الثانية 250.

المطلوب:

1- أوجد توزيع المعاينة للفرق بين متوسطي العينتين؟

2-احسب إحتمال ان يكون الفرق بين متوسطي اعينتين اقل من 05؟

الحل:

1-إيجاد توزيع المعاينة للفرق بين متوسطي العينتين

المجتمع الاول

X_1:متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 30 وتباين 18 اي

$X_1\sim N(50,09)$

ولدينا: $N_1=1000$، (السحب مع عدم الارجاع)$n_1=120$ ، $\mu_1=30,\sigma_1^2=18$

المجتمع الثاني

X_2: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 24 وتباين 40 اي

$X_2\sim N(24,40)$

ولدينا: $N_2=1400$، (السحب مع عدم الارجاع)$n_2=250$ ، $\mu_2=24,\sigma_2^2=40$

بما ان المجتمعين يتوزعان توزيعا طبيعيا، فان توزيع المعاينة لفروق المتوسطات سيكون توزيعا قريب من التوزيع الطبيعي بمتوسط حسابي وتباين على التوالي:

$$E\left(\overline{X}_1-\overline{X}_2\right)=E\left(\overline{X}_1\right)-E\left(\overline{X}_2\right)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$
$$=30-24=06$$

حساب التباين: بما ان المجتمعين محدودين والسحب مع عدم الارجاع (الشرطان الكافيان محققان) ننتقل الى الشرط اللازم وهو كسر المعاينة.

المجتمع الاول: $\frac{n_1}{N_1}=\frac{120}{1000}=0.12>0.05$ اذن نستخدم معامل الارجاع او التصحيح.

المجتمع الثاني: $\frac{n_2}{N_2}=\frac{250}{1400}=0.18>0.05$ اذن نستخدم معامل الارجاع او التصحيح.

$$V\left(\overline{X}_1-\overline{X}_2\right)=V\left(\overline{X}_1\right)+V\left(\overline{X}_2\right)=\sigma^2_{\overline{X}_1}+\sigma^2_{\overline{X}_2}=\frac{\sigma_1^2}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{\sigma_2^2}{n_2}\times\frac{N_2-n_2}{N_2-1}$$

$$=\frac{18}{120}\times\frac{1000-120}{1000-01}+\frac{40}{250}\times\frac{1400-250}{1400-01}\simeq 0.2636$$

إذن $\overline{X}_1-\overline{X}_2\sim N(06,0.2636)$

2-حساب إحتمال ان يكون الفرق بين متوسطي اعينتين اقل من 05

بتطبيق الاحصاءة (Z) للتوزيع الطبيعي المعياري $Z=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{\sigma_2^2}{n_2}\times\frac{N_2-n_2}{N_2-1}}}\sim(0,1)$

$$P\left(\overline{X}_1-\overline{X}_2<05\right)=P\left(Z<\frac{05-06}{\sqrt{0.2636}}\right)=P(Z<-01.95)=F(-01.95)=01-F(01.95)$$

$$=01-0.97441=0.02559$$

<u>تمرين10</u>: إذا كان الانفاق الاستهلاكي العائلي في مدينة "أ" بوسط حسابي 24ون وتباين 120 ون، والانفاق الاستهلاكي العائلي اليومي في مدينة "ب" بوسط حسابي 15 ون وتباين 72 ون، فإذا سحبنا من المدينة "أ" عينة عشوائية تشمل 240 عائلة، ومن المدينة "ب" عينة عشوائية تشمل 150 عائلة، وكانت العينتان مستقلتين عن بعضهما البعض.

<u>المطلوب:</u>

1- أوجد توزيع المعاينة للفرق بين متوسطي العينتين؟

2-أحسب احتمال ان يكون الفرق بين متوسطي العينتي اكبر من 10 ون؟

<u>الحل:</u>

1-إيجاد توزيع المعاينة للفرق بين متوسطي العينتين

المجتمع الاول

X_1:متغير عشوائي مستمر يخضع لتوزيع مجهول بمتوسط حسابي 24 وتباين 120 اي $X_1\sim In(24,120)$

ولدينا: $N_1=\infty$، $n_1=240$، $\mu_1=24,\sigma_1^2=120$

المجتمع الثاني

X_2: متغير عشوائي مستمر يخضع لتوزيع مجهول بمتوسط حسابي 15 وتباين 72 اي $X_2\sim In(15,72)$

ولدينا: $N_2=\infty$، $n_2=150$، $\mu_2=15,\sigma_2^2=72$

المتغير $\left(\overline{X}_1-\overline{X}_2\right)$ سيتوزع توزيعا قريب من التوزيع الطبيعي بالرغم من ان توزيع المجتمعين مجهولان، وذلك لان $n_1\geq 30$، و $n_2\geq 30$،(نظرية النهاية المركزية) وبما ان المجتمعين كبيران فلا نستخدم معامل الارجاع (او التصحيح) عند حساب تباين توزيع المعاينة للفروق اي

التوقع الرياضي:

$$E\left(\overline{X}_1-\overline{X}_2\right)=E\left(\overline{X}_1\right)-E\left(\overline{X}_2\right)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$
$$=24-15=09$$

التباين:

$$V\left(\overline{X}_1-\overline{X}_2\right)=V\left(\overline{X}_1\right)+V\left(\overline{X}_2\right)=\sigma^2_{\overline{X}_1}+\sigma^2_{\overline{X}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$$
$$=\frac{120}{240}+\frac{72}{150}\simeq 0.98$$

إذن $\overline{X}_1-\overline{X}_2 \sim N(09,0.98)$

2-حساب احتمال ان يكون الفرق بين متوسطي العينتي اكبر من 10 ون.

بتطبيق الاحصاءة (Z) للتوزيع الطبيعي المعياري

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

$$P\left(\overline{X}_1-\overline{X}_2>10\right)=P\left(Z>\frac{10-09}{\sqrt{0.98}}\right)=P(Z>01.01)=01-P(Z<01.01)$$
$$=01-F(01.01)=01-0.84375=0.15625$$

ثانيا:توزيع المعاينة للفرق $\left(\overline{X}_1-\overline{X}_2\right)$ لمجتمعين يتوزعان توزيعا طبيعيا وبتباينين مجهولين σ_1^2 و σ_2^2 وحجم العينتين كبير

نظرية11: إذا سحبنا كل العينات العشوائية الممكنة ذات الحجم $n_1 \geq 30$ من مجتمع لامحدود يتوزع توزيعا طبيعيا بوسط حسابي μ_1 وتباين σ_1^2 مجهول، وسحبنا كل العينات العشوائية الممكنة ذات الحجم $n_2 \geq 30$ من مجتمع آخر لا محدود هو الاخر يتوزع توزيعا طبيعيا بوسط حسابي μ_2 وتباين σ_2^2 مجهول، وكانت العينات المسحوبة من المجتمع الاول مستقلة عن العينات المسحوبة من المجتمع الثاني فان توزيع المعاينة للفرق بين المتوسطات $\left(\overline{X}_1-\overline{X}_2\right)$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي هو $\mu_{\overline{X}_1-\overline{X}_2}$ و تباين $\sigma^2_{\overline{X}_1-\overline{X}_2}$ اي $\overline{X}_1-\overline{X}_2\sim N\left(\mu_{\overline{X}_1-\overline{X}_2},\sigma^2_{\overline{X}_1-\overline{X}_2}\right)$

ولايجاد معلمتي التوزيع نستعين بخصائص التوقع والتباين .

التوقع الرياضي:

$$E\left(\overline{X}_1-\overline{X}_2\right)=E\left(\overline{X}_1\right)-E\left(\overline{X}_2\right)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$

التباين :

$$V\left(\overline{X}_1-\overline{X}_2\right)=V\left(\overline{X}_1\right)+V\left(\overline{X}_2\right)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$$
$$\simeq\frac{\hat{S}_1^2}{n_1}+\frac{\hat{S}_2^2}{n_2}\simeq\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}$$

اما المتغير العشوائي (Z) فتصبح علاقته كالتالي:

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

تمرين11: معمل ينتج 700 كغ من المكرونة كمعدل يومي، سحبت منه عينة عشوائية تمثل انتاج 40 يوما فبل وسطها الحسابي 740 كغ بانحراف معياري 40 كغ، معمل آخر ينتج 500 كغ كمعدل يومي، سحبت منه عينة عشوائية تمثل إنتاج 35 يوما فبل وسطها الحسابي 480 كغ بانحراف معياري 20 كغ.

المطلوب:

1-أوجد توزيع المعاينة للفرق بين متوسطي العينتين؟

2-أحسب احتمال ان يكون الفرق محصور بين 180كغ و210 كغ؟

الحل:

1-إيجاد توزيع المعاينة للفرق بين متوسطي العينتين

المجتمعان يخضعان لتوزيع مجهول وبما ان حجم العينتين المسحوبتين اكبر او تساوي 30 وبالاستفادة من نظرية النهاية المركزية فان توزيع المعاينة للفرق سوف يكون له توزيع قريب من التوزيع الطبيعي.

المجتمع الاول: $S_1=40$، $n_1=40$، $\mu_1=700$ ،$N_1=\infty$

المجتمع الثاني: $S_2=20$، $n_2=35$، $\mu_2=500$ ،$N_2=\infty$

حساب التوقع الرياضي:

$$E(\overline{X}_1-\overline{X}_2)=E(\overline{X}_1)-E(\overline{X}_2)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$
$$=700-500=200$$

حساب التباين:

$$V(\overline{X}_1-\overline{X}_2)=V(\overline{X}_1)+V(\overline{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}=\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}$$
$$=\frac{(40)^2}{40-1}+\frac{(20)^2}{35-1}=52.80$$

إذن $\overline{X}_1-\overline{X}_2\sim N(200,52.80)$

2-حساب احتمال ان يكون الفرق محصور بين 180كغ و210 كغ

بتطبيق الاحصاءة (Z) للتوزيع الطبيعي المعياري

$$P(180<\overline{X}_1-\overline{X}_2<210)=P\left(\frac{180-200}{\sqrt{52.80}}<Z<\frac{210-200}{\sqrt{52.80}}\right)=P(-02.75<Z<01.38)$$
$$=F(01.38)+F(02.75)-01$$
$$=0.91621+0.99702-01=0.91323$$

ثالثا:توزيع المعاينة للفرق$(\overline{X}_1-\overline{X}_2)$ لمجتمعين يتوزعان توزيعا طبيعيا وبتباينين مجهولين σ_1^2 و σ_2^2 وكان على الاقل حجم احد العينتين صغيير

تمهيد:نحتاج في بعض الدراسات الاحصائية لمقارنة متوسطين حسابيين لعينتين ومعرفة الفرق بينهما $(\overline{X}_1-\overline{X}_2)$فإذا كان لدينا مجتمع اول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2مجهول،

ومجتمع ثاني يتوزع هو الاخر توزيع طبيعي بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، وحجم العينة المسحوبة من المجتمع الاول $n_1<30$، وحجم العينة المسحوبة من المجتمع الثاني $n_2<30$، وكانت العينتان مستقلتين عن بعضهما البعض، ولدراسة توزيع المعاينة للفرق $(\overline{X}_1-\overline{X}_2)$ في هذه الحالة نكون أمام حالتين هما:

الحالة الاولى: عندما يكون التباينين مجهولين ومتساويين $\left(\sigma_1^2=\sigma_2^2=\sigma^2\right)$

نظرية12: إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 ، والمجتمع الثاني يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 ، وكان σ_1^2 و σ_2^2 مجهولين ومتساويين، وسحبنا من المجتمع الاول عينةعشوائية بسيطة حجمها $n_1<30$ ، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2<30$ مع فرضية تساوي تبايني المجتمعين، كما نعلم ان افضل مقدر لتباين المجتمع هو تباين العينة المسحوبة منه، فيعني ذلك ان افضل مقدر لتباين المجتمع الاول هو تباين العينة المسحوبة منه S_1^2، وان افضل مقدر لتباين المجتمع الثاني هو تباين العينة المسحوبة منه S_2^2، وبالطبع S_1^2 و S_2^2 مختلفان في القيمة وبما ان تباين المجتمع الاول يساوي تباين المجتمع الثاني فليس من المنطق ان نقدر معلمتين متساويتين بتقديرين مختلفين، لذلك نقدر التباينين بنفس المقدر، وهو عبارة عن الوسط المرجح لتباين العينة الاولى وتباين العينة الثانية، ويتم الترجيح بدرجة الحرية الخاصة بكل عينة، ويطلق على هذا المقدار التباين المشترك (the pooled variance) ، ويرمز له بالرمز S_p^2

وهو يحسب كمايلي:
$$S_p^2=\frac{(n_1-1)\hat{S}_1^2+(n_2-1)\hat{S}_2^2}{(n_1-1)+(n_2-1)}=\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}$$

ويخضع توزيع المعاينة للفرق بين متوسطي في حالة عدم معلومية تبايني المجتمعين وحجم العينين اقل تمام من 30 الى توزيع ستودنت بالعلاقة التالية:
$$\overline{X}_1-\overline{X}_2\sim T_{n_1+n_2-2}\left(\mu_{\overline{X}_1-\overline{X}_2},\sigma^2_{\overline{X}_1-\overline{X}_2}\right)$$

وتعطى معلمتي التوزيع كما يلي:

التوقع الرياضي:
$$E(\overline{X}_1-\overline{X}_2)=E(\overline{X}_1)-E(\overline{X}_2)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$

التباين:
$$V(\overline{X}_1-\overline{X}_2)=V(\overline{X}_1)+V(\overline{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\simeq\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}$$
$$=\frac{S_p^2}{n_1}+\frac{S_p^2}{n_2}=S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

وتعطى إحصاءة ستوندت (T) كالتالي:
$$T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

تمرين12: سحبت عينة عشوائية حجمها 16 بتباين 04 من مجتمع اول يخضع للتوزيع الطبيعي بمتوسط حسابي30 وتباينه مجهول، وسحبت عينة اخرى حجمها 14 بتباين 05 من مجتمع ثاني يخضع هو الاخر للتوزيع الطبيعي بمتوسط حسابي 28 وتباين مجهول.

المطلوب:

1-أوجد توزيع المعاينة للفرق بين متوسطي العينتين بافتراض تساوي التباينين؟

2-احسب احتمال ان يكون الفرق اقل من04؟

الحل:

1-إيجاد توزيع المعاينة للفرق بين متوسطي العينتين بافتراض تساوي التباينين

المجتمعان يخضعان للتوزيع الطبيعي وبتباينين مجهولين وبما ان حجم العينتين المسحوبتين أقل تماما من 30 ومع فرضية تساوي التباينين فان توزيع المعاينة للفرق سوف يخضع لتوزيع ستودنت بدرجة حرية n_1+n_2-2

المجتمع الاول: $S_1^2=04$ ،$n_1=16$، $\mu_1=30$

المجتمع الثاني: $S_2^2=05$ ،$n_2=14$، $\mu_2=28$

المعاينة تخضع لتوزيع ستودنت $\overline{X}_1-\overline{X}_2 \sim T_{n_1+n_2-2}\left(\mu_{\overline{X}_1-\overline{X}_2},\sigma^2_{\overline{X}_1-\overline{X}_2}\right)$

حساب التوقع الرياضي:

$$E\left(\overline{X}_1-\overline{X}_2\right)=E\left(\overline{X}_1\right)-E\left(\overline{X}_2\right)=\mu_{\overline{X}_1}-\mu_{\overline{X}_2}=\mu_1-\mu_2$$
$$=30-28=02$$

حساب التباين: نحسب اولا التباين المشترك

$$S_p^2=\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}=\frac{16\times04+14\times05}{16+14-2}=\frac{134}{28}=04.785$$

$$V\left(\overline{X}_1-\overline{X}_2\right)=V\left(\overline{X}_1\right)+V\left(\overline{X}_2\right)=S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$
$$=04.785\left(\frac{1}{16}+\frac{1}{14}\right)=0.64$$

إذن توزيع المعاينة هو $\overline{X}_1-\overline{X}_2 \sim T_{28}\left(02,0.64\right)$

2-حساب احتمال ان يكون الفرق اقل من04

لحساب هذا الاحتمال نستعين باحصاءة ستودنت (T) كالتالي:

$$T=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$P\left(\overline{X}_1-\overline{X}_2<04\right)=P\left(T_{28}<\frac{04-02}{\sqrt{0.64}}\right)=P\left(T_{28}<02.50\right)=F\left(02.50\right)=0.991$$

وقد تم ايجاد الاحتمال السابق باستخدام نظرية الحصر بالنقصان كما تم شرحها سابقا.

الحالة الثانية: عندما يكون التباينين مجهولين و غير متساويين $\left(\sigma_1^2 \neq \sigma_2^2\right)$

نظرية 13: إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 ، وكان σ_1^2 و σ_2^2 مجهولين وغير متساويين، وسحبنا من المجتمع الاول عينةعشوائية بسيطة حجمها $n_1<30$ ، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2<30$ مع فرضية عدم تساوي تبايني المجتمعين، كما نعلم ان افضل مقدر لتباين المجتمع هو تباين العينة المسحوبة منه S_1^2، وان افضل مقدر لتباين المجتمع الثاني هو S_2^2

وبتعويض S_1^2 و S_2^2 في العلاقة السابقة نتحصل على المتغير التالي:

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

والتوزيع الاحتمالي لهذا المتغير العشوائي معقد جدا، وقد درس هذه المشكلة كثير من الاحصائيين، ومن أهمهم فيشر وبهرين، ولذلك تسمى هذه المشكلة بمشكلة فيشر-بهرين(the fisher-Behren problem)، واقترحت عدة حلول لهذه المشكلة أكثرها استخداما هو اعتبار ان توزيع هذا المتغير هو توزيع قريب من توزيع ستودنت (T) بدرجة حرية V تحسب كالتالي:

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

وتعطى احصاءة ستونت (T) في هذه الحالة كالتالي :

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

تمرين13: سحبت عينة عشوائية حجمها 09 بتيابن 36 من مجتمع اول يخضع للتوزيع الطبيعي بمتوسط حسابي64 وتباينه مجهول، وسحبت عينة اخرى حجمها 16 بتباين 25 من مجتمع ثاني يخضع هو الاخر للتوزيع الطبيعي بمتوسط حسابي59 وتباين مجهول.

المطلوب:

1-أوجد توزيع المعاينة للفرق بين متوسطي العينتين بافتراض عدم تساوي التباينين؟

2-احسب احتمال ان يكون الفرق أكبر من08؟

الحل:

1-إيجاد توزيع المعاينة للفرق بين متوسطي العينتين بافتراض عدم تساوي التباينين؟

المجتمعان يخضعان للتوزيع الطبيعي وبتباينين مجهولين وبما ان حجم العينتين المسحوبتين أقل تماما

من 30 ومع فرضية عدم تساوي التباينين فان توزيع المعاينة للفرق سوف يخضع لتوزيع ستودنت بدرجة حرية V

المجتمع الاول: $S_1^2=36$ ،$n_1=09$، $\mu_1=64$

المجتمع الثاني: $S_2^2=25$ ،$n_2=16$، $\mu_2=59$

المعاينة تخضع لتوزيع ستودنت بدرجة حرية V $\quad \bar{X}_1-\bar{X}_2 \sim T_V\left(\mu_{\bar{X}_1-\bar{X}_2},\sigma^2_{\bar{X}_1-\bar{X}_2}\right)$

حساب التوقع الرياضي:

$$E(\bar{X}_1-\bar{X}_2)=E(\bar{X}_1)-E(\bar{X}_2)=\mu_{\bar{X}_1}-\mu_{\bar{X}_2}=\mu_1-\mu_2$$
$$=64-59=05$$

حساب التباين:

$$V(\bar{X}_1-\bar{X}_2)=V(\bar{X}_1)+V(\bar{X}_2)=\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)$$
$$=\left(\frac{36}{09}+\frac{25}{16}\right)=05.5625$$

بعد حساب التباين نستخرج درجة الحريةV كالتالي:

$$V=\frac{\left[\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right]^2}{\frac{\left[\frac{S_1^1}{n_1}\right]^2}{n_1-1}+\frac{\left[\frac{S_2^2}{n_2}\right]^2}{n_2-1}}=\frac{\left[\frac{36}{09}+\frac{25}{16}\right]^2}{\frac{\left[\frac{36}{09}\right]^2}{09}+\frac{\left[\frac{25}{16}\right]^2}{16}}\approx 16$$

إذن توزيع المعاينة هو $\quad \bar{X}_1-\bar{X}_2 \sim T_{16}(05,05.5625)$

2-حساب احتمال ان يكون الفرق أكبر من08

لحساب هذا الاحتمال نستعين باحصاءة ستودنت (T) كالتالي: $\quad T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

$$P(\bar{X}_1-\bar{X}_2>08)=P\left(T_{16}>\frac{08-05}{\sqrt{05.5625}}\right)=P(T_{16}>01.272)=01-P(T_{16}<01.272)$$
$$=01-F(01.272)=01-0.886=0.114$$

ملاحظة: ان الاحتمال المقابل للقيمة 01.272 عند درجة حرية 16 في جدول توزيع ستودنت لا يوجد لذا تم ايجاد الاحتمال المقابل لها (0.886) باستخدام نظرية الحصر بالنقصان كما تم شرحها سابقا في التمرين 07

2-4:توزيع المعاينة للفرق بين متوسطي عينتين$(\bar{X}_1-\bar{X}_2)$مرتبطتين

تمهيد: في كثير من الدراسات عند مقارنة متوسطي عينتين نجد ان العينتين ير مستقلتين، وتكون كل قيمة في العينة الاولى مقرونة بقيمة في العينة الثانية، أي تكون القيمة الاولى في العينة الاولى والقيمة الاولى في العينة الثانية تابعتين لنفس المفردة، والقيمة الثانية في العينة الالولى والقيمة الثانية في العينة

الثانية تابعتين لنفس المفردة وهكذا...،أي نجد ان القيم المشاهدة في العينتين تكون ازواج من القيم، ولذلك سميت العينتان في هذه الحالة بالعينتين المزدوجتين او المقرونتين او المرتبطتين.

نظرية14:إذا كان لدينا مجتمعين الاول يخضع للتوزيع الطبيعي بمتوسط حسابي μ_1 وتباين σ_1^2 ، ومجتمع ثاني هو الاخر يخضع للتوزيع الطبيعي بمتوسط حسابي μ_2 وتباين σ_2^2، وكان المجتمعان مرتبطان و سحبت من كل مجتمع عينة حجمها n $\{X_1,X_2,X_3,\ldots\ldots X_n\}$ و $\{Y_1,Y_2,Y_3,\ldots\ldots Y_n\}$ على الترتيب، بحيث تمثل X_i و Y_i القيمتين المتناظرتين في العينتين،إذا كانت الفروق بين قيم العينتين المتناظرتين هي $\{d_1,d_2,d_3,\ldots\ldots d_n\}$ حيث $d_i = X_i - Y_i$ لجميع قيم i=1,2,3,…………n، فان $\{d_1,d_2,d_3,\ldots\ldots d_n\}$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على أنها عينة عشوائية من مجتمع طبيعي وسطه الحسابي $\mu_d = \mu_1 - \mu_2$ وتباينه σ_d^2، ولنفرض ان وسط العينة هو $\bar{d}$ وتباينها هو S_d^2 ونميز الحالتين التاليتين:

الحالة الاولى: حجم العينة $n \geq 30$ و σ_d^2 مجهول

في هذه الحالة فان متوسط الفرق للعينتين سوف يخضع للتوزيع الطبيعي كالتالي: $\bar{d} \sim N\left(E(\bar{d}), V(\bar{d})\right)$

ويتم حساب معلمتي التوزيع كما يلي :

$$E(\bar{d}) = \mu_{\bar{d}} = \frac{\sum_{i=1}^{n} d_i}{n}$$

$$V(\bar{d}) = \sigma_{\bar{d}}^2 = \frac{\sigma^2}{n} \simeq \frac{S_d^2}{n-1} = \frac{\hat{S}_d^2}{n}$$

في حالةعدم معلومية التباين نستخدم الاحصاءة التالية:

$$Z = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n-1}}} = \frac{\bar{d} - \mu_d}{\frac{\hat{S}_d}{\sqrt{n}}} \sim N(0,1)$$

الحالة الثانية: حجم العينة $n < 30$ و σ_d^2 مجهول

في هذه الحالة فان متوسط الفرق للعينتين سوف يخضع للتوزيع ستودنت (T) بدرجة حرية V اي

$$\bar{d} \sim T_V\left(E(\bar{d}), V(\bar{d})\right)$$

ويتم حساب معلمتي التوزيع كما يلي :

$$E(\bar{d}) = \mu_{\bar{d}} = \frac{\sum_{i=1}^{n} |d_i|}{n} = \frac{\sum_{i=1}^{n} |X_i - Y_i|}{n}$$

$$V(\bar{d}) = \sigma_{\bar{d}}^2 = \frac{\sigma^2}{n} \simeq \frac{S_d^2}{n-1} = \frac{\hat{S}_d^2}{n}$$

في حالةعدم معلومية التباين وحجم العينة $n < 30$ نستخدم إحصاءة ستودنت:

$$T = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n-1}}} = \frac{\bar{d} - \mu_d}{\frac{\hat{S}_d}{\sqrt{n}}}$$

تمرين14: إذا كان أداء العمال يخضع للتوزيع الطبيعي، وفي في تجربة لبيان تحسن هذا الأداء تم سحب عينة عشوائية بحجم 16عاملا في المعمل فكانت قياس الكفاءة قبل وبعد دخولهم دورة تحسين الأداء كما هو مبين في الجدول التالي:

بعد (X_i	07	08	07	07	08	08	09	09	09	08	07	07	07	08	08	08
قبل(Y_i)	05	05	05	04	07	08	05	06	07	08	05	06	05	07	07	05

المطلوب:

1- أوجد توزيع المعاينة للفرق بين متوسطي العينتين المرتبطتين ؟

2-أحسب احتمال أن يكون الفرق في الأداء قبل وبعد الدورة لا يقل عن 02.50؟

الحل:

1-إيجاد توزيع المعاينة للفرق بين العينتين: بما ان آداء العمال متغير عشوائي يخضع للتوزيع الطبيعي وحجم العينة المسحوبة اقل تمام من 30 فالمعاينة سوف تخضع لتوزيع ستودنت.

إيجاد معلمتي التوزيع:

نحسب اولا الفرق $d_i = X_i - Y_i$ ونلخص النتائج في الجدول التالي:

بعد (X_i	07	08	07	07	08	08	09	09	09	08	07	07	07	08	08	08
قبل(Y_i)	05	05	05	04	07	08	05	06	07	08	05	06	05	07	07	05
$d_i = X_i - Y_i$	02	03	02	03	01	00	04	03	02	00	02	01	02	01	01	03

التوقع الرياضي:
$$E(\bar{d}) = \mu_{\bar{d}} = \frac{\sum_{i=1}^{n}|d_i|}{n} = \frac{\sum_{i=1}^{n}|X_i - Y_i|}{n} = \frac{30}{16} = 01.875$$

التباين:
$$S_d^2 = \frac{\sum_{i=1}^{n}(d_i - \bar{d})^2}{n} = \frac{(02-01.875)^2 + (03-01.875)^2 + \dots\dots\dots + (03-01.875)^2}{16}$$
$$= \frac{12.71875}{16} = 0.80$$
$$V(\bar{d}) = \frac{S_d^2}{n-1} = \frac{0.80}{15} = 0.053$$

إذن توزيع المعاينة هو $\bar{d} \sim T_{15}(01.875, 0.053)$

2-حساب احتمال أن يكون الفرق في الأداء قبل وبعد الدورة لا يقل عن 02.50

لحساب الاحتمال التالي نستعين بإحصاءة ستودنت
$$T = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n-1}}} = \frac{\bar{d} - \mu_d}{\frac{\hat{S}_d}{\sqrt{n}}}$$

$$P(\bar{d} > 02.50) = P\left(T_{15} > \frac{02.50 - 01.875}{\sqrt{0.053}}\right) = P(T_{15} > 02.715) = 01 - P(T_{15} < 02.715)$$
$$= 01 - F(02.715) = 01 - 0.9915 = 0.0085$$

ملاحظة: ان الاحتمال المقابل للقيمة 02.715 عند درجة حرية 15 لا يوجد في جدول توزيع ستودنت لذا تم ايجاد الاحتمال (0.9915) المقابل للقيمة (02.715) باستخدام نظرية الحصر بالنقصان كما تم شرحها سابقا في التمرين 07.

5-توزيع المعاينة لنسبة العينة $\hat{P}$

تمهيد:يحتاج الباحث في اغلب الدراسات لمعرفة نسبة ظاهرة معينة في المجتمع محل الدراسة، كنسبة المدخنين في مدينة ما، نسبة الذكور في جامعة ما، نسبة الوحدات التالفة في انتاج مصنع معين،....الخ،ففي كل هذه الحالات نجد ان المجتمع محل الدراسة منقسم الى قسمين، قسم تتوافر فيه الظاهرة محل الدراسة (الخاصية المدروسة) والقسم الثاني لا تتوافر فيه الظاهرة المدروسة، ومجتمعات من هذا النوع يكون فيها المتغير وصفيا أي نوعيا لا نستطيع قياسه قياسا كميا، وبالتالي تعاد صياغته وتحويله الى متغير عشوائي **X**، حيث ياخذ المتغير العشوائي القيمة (**X=01**) إذا توافرت الظاهرة محل الدراسة في المفردة المدروسة، وياخذ القيمة (**X=0**) إذا لم تتوافر الظاهرة في المفردة المدروسة، ويطلق على هذا النوع من التوزيع الذي ياخذ قيمتين فقط (0 او 01) بتوزيع ذي الحدين، ونتعامل في هذا النوع من المجتمعات مع نسبة الظاهرة محل الدراسة في المجتمع، ويرمز لها بالرمز P ويطلق عليها نسبة المجتمع وهي تحسب كالتالي:

$$P = \frac{\text{عدد مفردات المجتمع التي تتحقق فيها الظاهرة المدروسة}}{\text{العدد الكلي لمفردات المجتمع}} = \frac{X}{N}$$

وبالتالي فان P تمثل إحتمال ظهور هذه الظاهرة في المجتمع، ويرمز لاحتمال عدم ظهور هذه الظاهرة في المجتمع بالرمز q، حيث ان حدث ظهور الظاهرة وحدث عدم ظهورها هما حدثان مكملان لبعضهما البعض اي $p + q = 01$

وتعتبر النسبة P من اهم معالم المجتمع التي يرغب الباحث في معرفتها ليستطيع وصف المجتمع محل الدراسة وصفا جيدا، ولكن في الكثير من الاحيان لا نستطيع تحديد نسبة المجتمع لعدم توافر بيانات لدينا عن كل مفردات المجتمع، ولذلك نقوم بالاستدلال عليها، اي استنتاجها باستخدام نسبة الظاهرة محل الدراسة في العينة العشوائية المسحوبة من هذا المجتمع، ويرمز لنسبة العينة بالرمز $\hat{P}$ وتحسب كما يلي:

$$\hat{P} = \frac{\text{عدد مفردات العينة التي تتحقق فيها الظاهرة المدروسة}}{\text{العدد الكلي لمفردات العينة}} \quad \frac{X}{n}$$

نسبة العينة $\hat{P}$، كأي احصائية تتغير قيمتها من عينة الى اخرى وبالتالي فهي متغير عشوائي له توزيع إحتمالي يطلق عليه توزيع المعاينة لنسبة العينة.

نظرية15: ليكن **X** متغير عشوائي منفصل يخضع للتوزيع ثنائي الحدين بمتوسط حسابي $\mu = np$ وتباين $\sigma^2 = npq$ من مجتمع ما غير محدود، حيث P نسبة المفردات ذات صفة معينة، ولتكن $\hat{P}$متغير عشوائي تمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة n من ذات المجتمع بالارجاع او عدم الارجاع، و $\left(\frac{n}{N} \leq 0.05\right)$ نتحصل على توزيع المعاينة للاحصاءة $\hat{P}$ كالتالي:

$$\hat{P} \sim N\left(P, \frac{p \times q}{n}\right) \ldots\ldots\ldots\ldots\ldots\ldots \frac{n}{N} \leq 0.05$$

أما الاحصاءة **(Z)** للمتغير العشوائي $\hat{P}$ فتعطى بالعلاقة التالية: $Z = \frac{\hat{P} - \mu_{\hat{P}}}{\sigma_{\hat{P}}} = \frac{\hat{P} - P}{\sqrt{\frac{p \times q}{n}}} \sim N(0,1)$

البرهان:لاثبات معلمتي التوزيع نستخدم خاصية التقريب اي تقريب التوزيع ذي الحدين باستخدام التوزيع الطبيعي.

نعلم انه اذا كان لديناX متغير عشوائي مستمر يخضع للتوزيع الطبيعي فان معلمتي التوزيع هما $X \sim N(\mu, \sigma^2)$

كما نعلم ايضا اذا كان لدينا X متغير عشوائي منفصل يخضع لتوزيع ذي الحدين فان معلمتي التوزيع هما $X \sim ben(np, npq)$

باستخدام خاصية التقريب نجد:

$$E(X) = \mu = np$$
$$V(X) = \sigma^2 = npq$$

كما نعلم من جهة اخرى ان $\hat{P} = \frac{x}{n}$

بادخال التوقع الرياضي على طرفي المساوات وبالاستعانة بخصائص التوقع التي ذكرت فيما سبق نجد:

$$E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{\cancel{n}} \cancel{n}p = p$$

$$E(\hat{P}) = \mu_{\hat{P}} = p$$

بادخال التباين على طرفي العلاقة السابقة وبالاستعانة بخصائص التباين نجد:

$$V(\hat{P}) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{1}{n \times \cancel{n}} \cancel{n}pq = \frac{p \times q}{n}$$

$$V(\hat{P}) = \sigma^2_{\hat{P}} = \frac{p \times q}{n}$$

تمرين15:إذا كانت نسبة المدخنين في أحد المجتمعات تخضع لتوزيع ذي الحدين وتساوي 0.35 ، فاذا سحبت عينة عشوائية من 100 مفردة

المطلوب:

1-أوجد توزيع المعاينة للنسبة المدخنين؟

2-أحسب احتمال ان تكون نسبة المدخنين في هذه العينة أكبر من 0.4؟

3-ماهو إحتمال ان تكون نسبة المدخنين في العينة في حدود 0.05 من نسبة المجتمع؟

الحل:

X: متغير عشوائي منفصل يخضع لتوزيع ذي الحدين وهو يمثل نسبة المخنين في أحد المجتمعات.

1-إيجاد توزيع المعاينة للنسبة: لمعرفة طبيعة التوزيع نستخدم شروط التقريب من التوزيع ذي الحدين الى التوزيع الطبيعي:

1-الاحتمالP يقترب من 0.50

2-حجم العينة اكبر او يساوي 20

وبما ان الشرطين محققين فتوزيع المعاينة للنسية سيكون له توزيع قريب من التوزيع الطبيعي.

ولدينا من المعطيات: n=100،P=0.35

نستخرج معلمتي توزيع النسبة $\hat{P}$ بحساب التوقع الرياضي والتباين

$$E(\hat{P}) = \mu_{\hat{P}} = p = 0.35$$

$$V(\hat{P}) = \sigma^2_{\hat{P}} = \frac{p \times q}{n} = \frac{p \times (1-p)}{n}$$
$$= \frac{0.35 \times 0.65}{100} = 0.002275$$

إذن: $\hat{P} \sim N(0.35, 0.002275)$

2-حساب إحتمال ان تكون نسبة المدخنين في هذه العينة أكبر من 0.4.

لحساب الاحتمال نستعين بالاحصاءة (Z) كالتالي: $Z = \frac{\hat{P} - \mu_{\hat{P}}}{\sigma_{\hat{P}}} = \frac{\hat{P} - P}{\sqrt{\frac{p \times q}{n}}} \sim N(0,1)$

$$P(\hat{P} > 0.40) = P\left(Z > \frac{0.40 - 0.35}{\sqrt{0.002275}}\right) = P(Z > 01.05) = 01 - P(Z < 01.05)$$
$$= 01 - F(01.05) = 01 - 0.85314 = 0.14686$$

3-حساب إحتمال ان تكون نسبة المدخنين في العينة في حدود 0.05 من نسبة المجتمع.

نسبة العينة تكون مابين $P \pm 0.05 \Leftrightarrow 0.35 \pm 0.05 \Rightarrow (0.30, 0.40)$

ويحسب الاحتمال كالتالي:

$$P(0.30 < \hat{P} < 0.40) = P\left(\frac{0.30 - 0.35}{\sqrt{0.002275}} < Z < \frac{0.40 - 0.35}{\sqrt{0.002275}}\right) = P(-01.05 < Z < 01.05)$$
$$= 02F(01.05) - 01 = 02 \times 0.85314 - 01 = 0.70628$$

نظرية16: ليكن **X** متغير عشوائي منفصل يخضع للتوزيع ثنائي الحدين بمتوسط حسابي $\mu = np$ وتباين $\sigma^2 = npq$ من مجتمع ما محدود، حيث P نسبة المفردات ذات صفة معينة، ولتكن $\hat{P}$ متغير عشوائي تمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة n من ذات المجتمع مع عدم الارجاع، و $\left(\frac{n}{N} > 0.05\right)$ نتحصل على توزيع المعاينة للاحصاءة $\hat{P}$ كالتالي:

$$E\left(\hat{P}\right) = \mu_{\hat{P}} = P$$

$$V\left(\hat{P}\right) = \sigma^2_{\hat{P}} = \frac{p \times q}{n} \times \frac{N-n}{n-1} \ldots\ldots\ldots\ldots\ldots\ldots \frac{n}{N} > 0.05$$

إذن توزيع المعاينة هو

$$\hat{P} \sim N\left(P, \frac{p \times q}{n} \times \frac{N-n}{N-1}\right) \ldots\ldots\ldots\ldots\ldots\ldots \frac{n}{N} > 0.05$$

أما الاحصاءة **(Z)** للمتغير العشوائي $\hat{P}$ فتعطى بالعلاقة التالية:

$$Z = \frac{\hat{P} - \mu_{\hat{P}}}{\sigma_{\hat{P}}} = \frac{\hat{P} - P}{\sqrt{\frac{p \times q}{n} \times \frac{N-n}{N-1}}} \sim N(0,1)$$

تمرين16: إذا كان عدد الاسر في بلد ما هو 2500 أسرة وكانت نسبة الافراد فيها من الذكور تخضع للتوزيع ذي الحدين و هي تمثل 59%، سحبت عينة عشوائية من هذا المجتمع مع عدم الارجاع وكانت قيمة هذه العينة 200 اسرة.

المطلوب:

1-أوجد توزيع المعاينة لنسبة الذكور في الاسر؟

2-أحسب إحتمال ان تكون نسبة الذكور في هذا البلد تتراوح بين 51% و 61%

الحل:

X: متغير عشوائي منفصل يخضع لتوزيع ذي الحدين وهو يمثل نسبة الذكور في أحد المجتمعات.

1-إيجاد توزيع المعاينة لنسبة الذكور في الاسر: : لمعرفة طبيعة التوزيع نستخدم شروط التقريب من التوزيع ذي الحدين الى التوزيع الطبيعي:

1-الاحتمال P يقترب من 0.50

2-حجم العينة اكبر او يساوي 20

وبما ان الشرطين محققين فتوزيع المعاينة للنسبة سيكون له توزيع قريب من التوزيع الطبيعي.

ولدينا من المعطيات: P=0.59، n=200، N=2500

نستخرج معلمتي توزيع النسبة $\hat{P}$ بحساب التوقع الرياضي والتباين

حساب التوقع الرياضي:

$$E\left(\hat{P}\right) = \mu_{\hat{P}} = 0.59$$

حساب التباين: المجتمع محدود والسحب مع عدم الارجاع (الشرطان الكافيان محققان) ننتقل الى الشرط اللازم وهو كسر المعاينة $\frac{n}{N}=\frac{200}{2500}=0.08>0.05$

الشرط اللازم محقق اذن نستخدم معامل الارجاع او معامل التصحيح.

$$V\left(\hat{P}\right)=\sigma_{\hat{P}}^2=\frac{p\times q}{n}\times\frac{N-n}{n-1}$$
$$=\frac{0.59\times0.41}{200}\times\frac{2500-200}{2500-1}=0.0011$$

وعليه فتوزيع المعاينة هو $\hat{P}\sim N(0.59,0.0011)$

2-حساب إحتمال ان تكون نسبة الذكور في هذا البلد تتراوح بين 51% و 61%

من اجل حساب الاحتمال نستعين بالاحصاءة (Z)

$$P\left(0.51<\hat{P}<0.61\right)=P\left(\frac{0.51-0.59}{\sqrt{0.0011}}<Z<\frac{0.61-0.59}{\sqrt{0.0011}}\right)=P(-02.41<Z<0.60)$$
$$=F(0.60)+F(02.41)-1=0.72575+0.99202-1$$
$$=0.71777$$

نظرية 17 (نظرية النهاية المركزية): وفقا انظرية النهاية المركزية، توزيع المعاينة للنسبة $\hat{P}$ يقترب من التوزيع الطبيعي عندما يكون حجم العينة n كبيرا بما فيه الكفاية، ويتحقق ذلك عندما يكون كل من np و np أكبر او يساوي05 أي اذا تحقق:

$$\left.\begin{matrix}np\geq 05\\ nq\geq 05\end{matrix}\right\}\Rightarrow N$$

والمتغير العشوائي (z) والذي تعطى علاقته كالتالي $Z=\frac{\hat{P}-P}{\sqrt{\frac{p\times q}{n}}}\sim N(0,1)$ سيتبع توزيعا قريبا من التوزيع الطبيعي المعياري

وقد وضع العالم كوكرن (Cochran 1963) قواعد لحجم العينة اللازم من اجل استخدام التوزيع الطبيعي والجدول التالي يوضح ذلك:

يستخدم التوزيع الطبيعي إذا كان n على الاقل يساوي	إذا كانت P تساوي
30	0.50
50	0.40-0.60
80	0.30-0.70
200	0.20-0.80
600	0.10-0.90
1400	0.05-0.95

تمرين 17: بفرض ان مجتمعا ما افراده عدة آلاف يمثل مصنعا لانتاج بطاقات المعايدة، فإذا كان0.20

من البطاقات تالفة،اخذت عينة عشوائية بسيطة حجمها 300 بطاقة.

المطلوب:

1-أوجد توزيع المعاينة لنسبة البطاقات التالفة؟

2-أحسب احتمال ان تكون نسبة البطاقات التالفة اكبر من0.19؟

الحل:

X: متغير عشوائي منفصل يمثل بطاقات المعايدة وهو يخضع لتوزيع ذي الحدين بمتوسط جسابي 0.20

لدينا من معطيات السابقة: P=0.20، n=300،q=1-p=1-0.20=0.80

1-إيجاد اوزيع المعاينة لنسبة البطاقات التالفة: بما ان المجتمع يخضع لتوزيع ذي الحدين ولمعرفة توزيع المعاينة نستخدم نظرية النهاية المركزية

$$\left.\begin{aligned} np &= 0.20\times 300 = 60 \geq 05 \\ nq &= 0.80\times 300 = 240 \geq 05 \end{aligned}\right\} \Rightarrow N$$

إذن الشرطان محققان وبالتالي فتوزيع المعاينة سوف يكون له توزيع قريب من التوزيع الطبيعي.

$$\hat{P} \sim N\left(E\left(\hat{P}\right), V\left(\hat{P}\right)\right)$$

إيجاد معلمتي التوزيع

حساب التوقع الرياضي: $E\left(\hat{P}\right) = \mu_{\hat{P}} = 0.20$

حساب التباين: بما ان المجتمع كبير جدا لا نستخدم معامل الارجاع ويحسب التباين كالتالي

$$V\left(\hat{P}\right) = \sigma^2_{\hat{P}} = \frac{p\times q}{n} = \frac{0.20\times 0.80}{300} = 0.00053$$

إذن توزيع المعاينة هو كالتالي: $\hat{P} \sim N(0.20, 0.00053)$

2-حساب احتمال ان تكون نسبة البطاقات التالفة اكبر من0.19

لحساب الاحتمال نستعين بالمتغير العشوائي (Z)

$$P\left(\hat{P} > 0.19\right) = P\left(Z > \frac{0.19-0.20}{\sqrt{0.00053}}\right) = P(Z > -0.43) = 01 - P(Z < -0.43)$$

$$= 01 - F(-0.43) = F(0.43) = 0.66640$$

6-توزيع المعاينة للفرق بين نسبتي عينتين $\left(\hat{P}_1 - \hat{P}_2\right)$

تمهيد:إذا كانت الدراسة الخاصة بمقارنة ظاهرة معينة في مجتمعين مختلفين، أي محاولة معرفة الفرق بين النسبتين **(P_1-P_2)**، حيث $\hat{P}_1$ ترمز لنسبة الظاهرة في المجتمع الاول، و $\hat{P}_2$ ترمز انسبة الظاهرة في المجتمع الثاني، وعند عدم توافر بيانات عن كل مفردات المجتمع الاول و كل مفردات المجتمع الثاني،نقوم بالاستدلال على المعلمة **(P_1-P_2)**، أي استنتاجها باستخدام الفرق بين نسبتي العينتين

العشوائيتين المسحوبتين من هذين المجتمعين،أي باستخدام الاحصائية $\left(\hat{P}_1-\hat{P}_2\right)$، حيث $\hat{P}_1$ نسبة الظاهرة في العينة العشوائية المسحوبة من المجتمع الاول، و $\hat{P}_2$ نسبة الظاهرة في العينة العشوائية المسحوبة من المجتمع الثاني، ولذلك يجب دراسة توزيع المعاينة لهذه الاحصائية والذي يطلق عليه توزيع المعاينة للفرق بين نسبتي عينتين.

نظرية18: إذا كانت $x_1,x_2,x_3......x_{n1}$ تشكل عينة عشوائية من مجتمع إحصائي حجمه N_1 نسبة عناصرها التي لها الخاصية مدار البحث تساوي P_1 ونسبة عناصر العينة التي لها نفس الخاصية هي $\hat{P}_1$، وكانت لدينا $x_1,x_2,x_3......x_{n2}$ والتي تشكل عينة عشوائية من مجتمع إحصائي ثاني حجمه N_2 نسبة عناصرها التي لها الخاصية مدار البحث تساوي P_2 ونسبة عناصر العينة التي لها نفس الخاصية $\hat{P}_2$ وكانت العينتين مستقلتين عن بعضهما البعض، وقمنا بحساب الفرق بين بين نسب عينات المجتمع الاول ونسب عينات المجتمع الثاني، أي حسبنا كل قيم $\left(\hat{P}_1-\hat{P}_2\right)$ فسنحصل على توزيع المعاينة للفرق بين نسبتي عينتين $\left(\hat{P}_1-\hat{P}_2\right)$ والذي يخضع للتوزيع الطبيعي تقريبا بمتوسط حسابي $\mu_{\hat{P}_1-\hat{P}_2}$ وتباين $\sigma_{\hat{P}_1-\hat{P}_2}$ ، ونميز الحالتي التاليتين:

- **في حالة السحب بالارجاع او عدم الارجاع و** $\left(\frac{n_1}{N_1}\le 0.05\right)$ **و** $\left(\frac{n_2}{N_2}\le 0.05\right)$

نستنتج معلمتي التوزيع انطلاقا من خصائص التوقع الرياضي والتباين

$$\hat{P}_1-\hat{P}_2 \sim N\left(E\left(\hat{P}_1-\hat{P}_2\right),V\left(\hat{P}_1-\hat{P}_2\right)\right)$$

التوقع الرياضي:

$$E\left(\hat{P}_1-\hat{P}_2\right)=E\left(\hat{P}_1\right)-E\left(\hat{P}_2\right)=P_1-P_2$$

التباين:

$$V\left(\hat{P}_1-\hat{P}_2\right)=V\left(\hat{P}_1\right)+V\left(\hat{P}_2\right)=\frac{P_1\left(1-P_1\right)}{n_1}+\frac{P_2\left(1-P_2\right)}{n_2}$$

اما المتغير العشوائي (Z) فيعطى بالعلاقة التالية:

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

- **في حالة السحب مع عدم الارجاع و** $\left(\frac{n_1}{N_1}> 0.05\right)$ **و** $\left(\frac{n_2}{N_2}> 0.05\right)$

نستنتج معلمتي التوزيع انطلاقا من خصائص التوقع الرياضي والتباين

$$\hat{P}_1-\hat{P}_2 \sim N\left(E\left(\hat{P}_1-\hat{P}_2\right),V\left(\hat{P}_1-\hat{P}_2\right)\right)$$

التوقع الرياضي:

$$E\left(\hat{P}_1-\hat{P}_2\right)=E\left(\hat{P}_1\right)-E\left(\hat{P}_2\right)=P_1-P_2$$

التباين:

$$V\left(\hat{P}_1-\hat{P}_2\right)=V\left(\hat{P}_1\right)+V\left(\hat{P}_2\right)=\frac{P_1\left(1-P_1\right)}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{P_2\left(1-P_2\right)}{n_2}\times\frac{N_2-n_2}{N_2-1}$$

اما المتغير العشوائي (Z) فيعطى بالعلاقة التالية:

$$Z=\frac{\left(\hat{P}_1-\hat{P}_2\right)-\left(P_1-P_2\right)}{\sqrt{\frac{P_1\left(1-P_1\right)}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{P_2\left(1-P_2\right)}{n_2}\times\frac{N_2-n_2}{N_2-1}}}\sim N(0,1)$$

ملاحظة: بالنسبة لكسر المعاينة للمجتمع الاول والثاني ليس بالضرورة ان يتحققا معا، فقد يتحقق الاول ولا يتحقق الثاني او العكس.

تمرين18: إذا علمت انه يوجد نوعين من المبيدات الحشرية بالسوق، وتدعي الشركة المصنعة للنوع الاول انه يقضي على 90% من الحشرات عند استعماله وتدعي الشركة المصنعة للنوع الثاني انه يقضي على 80% من الحشرات عند استعماله فإذا تم رش حجرتين لهما نفس الحجم بالنوع الاول والثانية بالنوع الثاني وتم اختيار عينتين من الحشرات حجم كل منها 200 ووضعت كل عينة في حجرة.

المطلوب:

1-إيجاد توزيع المعاينة للفرق بين نسبتي العينتين؟

2-أحسب احتمال ان يكون الفرق مابين نسبتي العينتي مابين 15% و 18%؟

الحل:

لدينا المعطيات التالية: $P_1=0.90$، $P_2=0.80$، $n_1=n_2=200$

1-إيجاد توزيع المعاينة للفرق بين نسبتي العينتين: بما ان المجتمعين يخضعان الى توزيع مجهول ولمعرفة توزيع المعاينة نستعين بنظرية النهاية المركزية

المجتمع الاول: بتطبيق نظرية النهاية المركزية

$$\left.\begin{aligned} n_1p_1&=0.90\times200=180\geq05\\ n_1q_1&=0.10\times200=20\geq05\end{aligned}\right\}\Rightarrow N$$

إذن توزيع المعاينة للنسبة ل $\hat{P}_1$ يخضع للتوزيع الطبيعي.

المجتمع الثاني: بتطبيق نظرية النهاية المركزية

$$\left.\begin{aligned} n_2p_2&=0.80\times200=160\geq05\\ n_2q_2&=0.20\times200=40\geq05\end{aligned}\right\}\Rightarrow N$$

إذن توزيع المعاينة للنسبة ل $\hat{P}_2$ يخضع للتوزيع الطبيعي.

مادام توزيع المعينة للمجتمع الاول يخضع للتوزيع الطبيعي وتوزيع المعاينة للمجتمع الثاني يخضع للتوزيع الطبيعي فان الفرق يعتبر متغير عشوائي جديد يخضع هو الاخر للتوزيع الطبيعي كالتالي:

$$\hat{P}_1-\hat{P}_2\sim N\left(E\left(\hat{P}_1-\hat{P}_2\right),V\left(\hat{P}_1-\hat{P}_2\right)\right)$$

ايجاد معلمتي التوزيع للمتغير العشوائي $\left(\hat{P}_1-\hat{P}_2\right)$

التوقع الرياضي: $E\left(\hat{P}_1-\hat{P}_2\right)=E\left(\hat{P}_1\right)-E\left(\hat{P}_2\right)=P_1-P_2=0.90-0.80=0.10$

التباين: $V\left(\hat{P}_1-\hat{P}_2\right)=V\left(\hat{P}_1\right)+V\left(\hat{P}_2\right)=\frac{P_1\left(1-P_1\right)}{n_1}+\frac{P_2\left(1-P_2\right)}{n_2}$

$$=\frac{0.90(1-0.90)}{200}+\frac{0.80(1-0.80)}{200}=0.00125$$

إذن توزيع المعاينة هو $\hat{P}_1-\hat{P}_2 \sim N(0.10,0.00125)$

2-حساب إحتمال ان يكون الفرق مابين نسبتي العينتي مابين 15%و 18%

لحساب الاحتمال نستعين بالمتغير العشوائي (Z)

$$P\left(0.15<\hat{P}_1-\hat{P}_2<0.18\right)=P\left(\frac{0.15-0.10}{\sqrt{0.00125}}<Z<\frac{0.18-0.10}{\sqrt{0.00125}}\right)=P(01.41<Z<02.26)$$

$$=F(02.26)-F(01.41)=0.98809-0.92073=0.06736$$

تمرين19:إذا كان لدينا مجتمعين يخضعان لتوزيع ذي الحدين، الاول حجمه 1400 نسبة النجاح فيه في امتحان الاحصاء الاستنتاجي هي 70%، والمجتمع الثاني حجمه 1200 نسبة النجاح فيه في نفس الامتحان هي 65%،أخذت عينة عشوائية بسيطة من المجتمع الاول حجمها 350 مع عدم الارجاع ،واخذت عينة عشوائية بسيطة من المجتمع الثاني حجمها 150 مع عدم الارجاع وكانت العينتين مستقلتان عن بعضهما البعض.

المطلوب:

1-أوجد توزيع المعاينة للفرق بين نسبتي العينتين؟

2-أحسب احتمال ان يكون الفرق بين نسبتي العينتين اقل من 20%؟

الحل:

لدينا المعطيات التالية: $P_1=0.70$،$P_2=0.65$، $n_1=350$، $n_2=150$، $N_1=1400$، $N_2=1000$

1-إيجاد توزيع المعاينة للفرق بين نسبتي العينتين: بما ان المجتمعين يخضعان لتوزيع ذي الحدين وحجم العينتين كبير جدا فان توزيع المعاينة سوف يكون له توزيع قريب من التوزيع الطبيعي.

$$\hat{P}_1-\hat{P}_2 \sim N\left(E\left(\hat{P}_1-\hat{P}_2\right),V\left(\hat{P}_1-\hat{P}_2\right)\right)$$

ايجاد معلمتي التوزيع للمتغير العشوائي $\left(\hat{P}_1-\hat{P}_2\right)$

التوقع الرياضي: $E\left(\hat{P}_1-\hat{P}_2\right)=E\left(\hat{P}_1\right)-E\left(\hat{P}_2\right)=P_1-P_2=0.70-0.65=0.05$

التباين: بما ان المجتمعين محدودين والسحب مع عدم الارجاع (الشرطان الكافيان محققان) ننتقل الى الشرط اللازم (كسر المعاينة)

كسر المعاينة للمجتمع الاول: $\frac{n_1}{N_1}=\frac{350}{1400}=0.25>0.05$ إذن الشرط محقق نستخدم معامل الارجاع.

كسر المعاينة للمجتمع الثاني: $\frac{n_2}{N_2}=\frac{150}{1000}=0.15>0.05$ إذن الشرط محقق نستخدم معامل الارجاع.

$$V\left(\hat{P}_1-\hat{P}_2\right)=V\left(\hat{P}_1\right)+V\left(\hat{P}_2\right)=\frac{P_1(1-P_1)}{n_1}\times\frac{N_1-n_1}{N_1-1}+\frac{P_2(1-P_2)}{n_2}\times\frac{N_2-n_2}{N_2-1}$$

$$=\frac{0.70(1-0.70)}{350}\times\frac{1400-350}{1400-1}+\frac{0.65(1-0.65)}{150}\times\frac{1000-150}{1000-1}$$

$$=0.001741$$

إذن توزيع المعاينة هو $\hat{P}_1-\hat{P}_2 \sim N(0.05,0.001741)$

2-حساب احتمال ان يكون الفرق بين نسبتي العينتين اقل من 20%

لحساب الاحتمال نستعين بالمتغير العشوائي (Z)

$$P\left(\hat{P}_1-\hat{P}_2<0.20\right)=P\left(Z<\frac{0.20-0.05}{\sqrt{0.001741}}\right)=P(Z<03.60)=F(03.60)=0.99984$$

7- توزيع المعاينة للتباين S^2 والانحراف المعياري S

من المعلوم والشائع لدى الاحصائيين ان التباين S^2 هو احد مقاييس التشتت فهو يقيس مدى تشتت الظاهرة حول متوسطها الحسابي وكلما كان هذا التشتت ضعيف كانت الظاهرة الاحصائية قريبة من المتوسط الحسابي لها وكلما زاد التشتت كانت الظاهرة الاحصائية بعيدة عن المتوسط الحسابي لها وعليه فان تباين العينة S^2 يستخدم للاستدلال عن σ^2 **و** لها نفس اهمية $\overline{X}$ عند الاستدلال عن μ ، وسف يتم التفصيل في هذا النوع من المعاينة كالتالي:

7-1 توزيع المعاينة لتباين العينة S^2

إذا أخذنا كل العينات العشوائية الممكنة والتي حجمها n والمسحوبة من مجتمع ما ثم حسبنا بعد ذلك التباينات (التغايرات) لكل عينة، فاننا نحصل بذلك على توزيع المعاينة للتباين S^2، وقبل الخوض في هذا النوع من التوزيع لابد ان نفرق بين تباين العينة S^2 وتباين العينة المعدل $\hat{S}^2$والذين تعطى معادلتهما كالتالي:

-إذا كان حجم العينة n≥30
$$S^2=\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n}$$

-إذا كان حجم العينة n<30
$$\hat{S}^2=\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}$$

نظرية19: لنفرض أن $x_1,x_2,x_3......x_n$ عينة عشوائية حجمها n سحبت من مجتمع يتوزع توزيع طبيعي بمعدل μ و σ^2 فإن تباينة العينة S^2 **او** $\hat{S}^2$ سيتوزع توزيع مربع كاي (χ^2) بدرجة حرية n-1 كالتالي:
$$\frac{(n-1)\hat{S}^2}{\sigma^2}=\frac{nS^2}{\sigma^2}\sim\chi^2_{n-1}$$

مقدمة عن توزيع مربع كاي(χ^2): إن الاساس النظري لاشتقاق توزيع مربع كاي هو التوزيع الطبيعي، بمعنى آخر ان توزيع مربع كاي مشتق من الدرجة المعيارية (Z)، وعلى افتراض كان لدينا المتغير

العشوائي (X) يخضع للتوزيع الطبيعي بمتوسط حسابي μ وتباين σ^2 أي ان $X \sim N(\mu, \sigma^2)$ ، عندئذ تكون الدرجة المعيارية (Z) كالتالي: $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ ، وبتربيع قيمة الدرجة المعيارية سنتحصل على المتغير (Z^2)، وعند البحث عن التوزيع الاحتمالي لهذا المتغير سيتضح بانه يخضع لتوزيع مربع كاي (χ^2) بدرجة حرية واحدة اي $Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi_1^2$

خصائص توزيع مربع كاي(χ^2): يتصف مربع كاي بالخصائص التالية

1- المساحة الكلية تحت منحنى التوزيع تساوي الواحد الصحيح.

2-إن قيمة مربع كاي موجبة دوما اي $(\chi^2 > 0)$

3-الوسط الحسابي μ للتوزيع هو عدد درجات الحرية (n-1)

4-إن التباين σ^2 للتوزيع هو ضعف درجات الحرية (2n-2)

5-منحنى توزيع مربع كاي(χ^2) ملتو نحو اليمين (إلتواء موجب)

6-عندما يكون عدد درجات الحرية (n>30) فان توزيع مربع كاي (χ^2) يقترب من التوزيع الطبيعي المعياري (Z).

نظرية20: إذا كان X متغير عشوائي يخضع للتوزيع الطبيعي بمتوسط حسابي μ و تباين σ^2 من مجتمع ما (محدود او لا محدود)، ولتكن S^2 متغيرة عشوائية تمثل تباين عينة حجمها n مسحوبة بالارجاع او بدون ارجاع فان التوقع الرياضي والتباين يعطى كالتالي:

أ-حالة المعاينة بالارجاع:

1-التوقع الرياضي

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

البرهان:

نعلم ان

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

بإدخال التوقع الرياضي على طرفي المساواة وبالاستعانة بخصائص التوقع نجد:

$$E(S^2) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2) - E(\overline{X}^2) \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

كما نعلم من جهة اخرى ان:

$$V(X_i) = E(X_i^2) - E^2(X_i) \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$V(\overline{X}) = E(\overline{X}^2) - E^2(\overline{X}) \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

بتعويض (2) و(3) في (1) نجد:

$$E\left(S^2\right)=\frac{1}{n}\sum_{i=1}^{n}\left(V\left(X_i\right)+\overset{2}{E}\left(X_i\right)\right)-\left(V\left(\overline{X}\right)+\overset{2}{E}\left(\overline{X}\right)\right)$$
$$=\frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2+\mu^2\right)-\left(\sigma_{\overline{X}}^2+\mu^2\right)$$

وبما ان السحب مع الارجاع فهذا يعني $V\left(\overline{X}\right)=\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}$

بالتعويض نجد

$$E\left(S^2\right)=\frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2+\mu^2\right)-\left(\frac{\sigma^2}{n}+\mu^2\right)$$
$$=\sigma^2+\cancel{\mu^2}-\frac{\sigma^2}{n}-\cancel{\mu^2}=\sigma^2\left(1-\frac{1}{n}\right)$$
$$=\sigma^2\left(\frac{n-1}{n}\right)$$

في التطبيقات العملية نعتبر n كبيرة عندما $(n\geq 30)$ وبشكل عام يعرف متغير إحصائي آخر يسمى متغير تباين العينة المعدل $\hat{S}^2$

نعلم انه توجد علاقة بين تباين العينة S^2 وتباين العينة المعدل $\hat{S}^2$ كالتالي:

$$\hat{S}^2=\frac{n}{n-1}S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$$

بادخال التوقع على طرفي المساوات نجد:

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

أي ان $E\left(\hat{S}^2\right)=\sigma^2$ هو مقدر أجود من S^2 كما هو واضح الان، وكما سنجده لاحقا اثناء دراسة محور التقدير.

2-التباين: نعلم ان تباين العينة يتوزع وفق توزيع مربع كاي بدرجة حرية $E\left(X\right)=n-1$ وتباين $V\left(X\right)=2\left(n-1\right)$

ويعطى التباين في حالة السحب مع الارجاع كالتالي: $V\left(\frac{nS^2}{\sigma^2}\right)=\sigma^2\left(\frac{nS^2}{\sigma^2}\right)$

بمساوات تباين العينة بتباين مربع كاي و بالاستعانة بخصائص التباين التي سبق وان تطرقنا لها نجد:

$$\sigma^2\left(\frac{nS^2}{\sigma^2}\right)=2\left(n-1\right)\Rightarrow\frac{n^2}{\sigma^4}\sigma^2\left(S^2\right)=2\left(n-1\right)$$

وبالتالي نجد: $\sigma^2\left(S^2\right)=\frac{2\left(n-1\right)}{n^2}\sigma^4$

أما تباين العينة المعدل $\hat{S}^2$ فيعطى بالعلاقة التالية:

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

$$=\frac{\cancel{n^2}}{\cancel{(n-1)}(n-1)}\times\frac{2\cancel{(n-1)}}{\cancel{n^2}}\sigma^4=\frac{2}{n-1}\sigma^4$$

ب-في حالة المعاينة مع عدم الارجاع و $\frac{n}{N}\leq 0.05$

في هذه الحالة لا يتحقق كسر المعاينة وبالتالي تبقى العلاقات كما هي في حالة السحب مع الارجاع.

تمرين20: إذا كان لدينا مجتمع كبير جدا يخضع للتوزيع الطبيعي بمتوسط حسابي 25 وإنحراف معياري 03

المطلوب: احسب التوقع الرياضي والتباين لكل من تباين العينة وتباين العينة المعدل لجميع العينات الممكنة ذات الحجم 09 والتي يمكن سحبها من هذا المجتمع.

الحل:

لدينا المعطيات التالية: $N=\infty, \mu=25, \sigma=03$

1-حساب التوقع الرياضي والتباين لتباين العينة S^2

التوقع الرياضي: نعلم ان $E\left(S^2\right)=\sigma^2\left(\frac{n}{n-1}\right)\Rightarrow E\left(S^2\right)=(03)^2\left(\frac{9}{9-1}\right)=10.125$

التباين: نعلم مما سبق ان $\sigma^2\left(S^2\right)=\frac{2(n-1)}{n^2}\sigma^4\Rightarrow\sigma^2\left(S^2\right)=\frac{2(9-1)}{(9)^2}(03)^4=16$

2-حساب التوقع الرياضي والتباين لتباين العينة المعدل $\hat{S}^2$

التوقع الرياضي: نعلم ان $E\left(\hat{S}^2\right)=\sigma^2\Rightarrow E\left(\hat{S}^2\right)=(03)^2=09$

التباين: نعلم مما سبق ان $V\left(\hat{S}^2\right)=\frac{2}{n-1}\sigma^4\Rightarrow V\left(\hat{S}^2\right)=\frac{2}{9-1}(03)^4=20.25$

نظرية21: إذا كان X متغير عشوائي يخضع للتوزيع الطبيعي بمتوسط حسابي μ و تباين σ^2 من مجتمع ما (محدود)، ولتكن S^2 متغيرة عشوائية تمثل تباين عينة حجمها n مسحوبة مع عدم الارجاع وكان $\frac{n}{N}>0.05$ فان التوقع الرياضي والتباين يعطى كالتالي:

حالة المعاينة مع عدم الارجاع: نستخرج التوقع الرياضي والتباين لتباين العينة وتباين العينة المعدل كالتالي:

1-التوقع الرياضي: $E\left(S^2\right)=\mu_{S^2}=\left(\frac{N}{N-1}\right)\left(\frac{n-1}{n}\right)\sigma^2$

البرهان: وجدنا مما سبق

$$E\left(S^2\right)=\frac{1}{n}\sum_{i=1}^{n}\left(V\left(X_i\right)+\overset{2}{E}\left(X_i\right)\right)-\left(V\left(\overline{X}\right)+\overset{2}{E}\left(\overline{X}\right)\right)$$

$$=\frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2+\mu^2\right)-\left(\sigma_{\overline{X}}^2+\mu^2\right)$$

وبما ان السحب مع عدم الارجاع فهذا يعني $V\left(\overline{X}\right)=\sigma_{\overline{X}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}$

بالتعويض وباختصار العلاقات الرياضية نجد $E\left(S^2\right)=\left(\frac{N}{N-1}\right)\left(\frac{n-1}{n}\right)\sigma^2$

أما تباين العينة المعدل فان توقعه يعطى كالتالي:

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

2-التباين:

تباين العينة ل S^2 هو $\sigma^2\left(S^2\right)=\frac{2(n-1)}{n^2}\sigma^4$

أما تباين العينة المعدل $\hat{S}^2$ فهو $\sigma^2\left(\hat{S}^2\right)=\frac{2}{n-1}\sigma^4$

ملاحظة: إذا كان المتغير العشوائي **X** لا يخضع الى التوزيع الطبيعي فان تباين العينة يعطى بالعلاقة الرياضية التالية: $\sigma_{S^2}=\sqrt{\frac{\mu_4-\sigma^4}{n}}$

ومن اجل $(n\geq 100)$ فان توزيع S^2 يقترب كثيرا من التوزيع الطبيعي.

<u>تمرين21</u>: إذا كان لدينا مجتمع حجمه 1000 يخضع للتوزيع الطبيعي بمتوسط حسابي 80 وإنحراف معياري 06، فإذا تم سحب عينة عشوائية حجمها 200 مع عدم الارجاع

<u>المطلوب:</u> احسب التوقع الرياضي والتباين لكل من تباين العينة وتباين العينة

<u>الحل:</u>

1-حساب التوقع الرياضي والتباين لتباين العينة S^2

التوقع الرياضي: نعلم ان

$$E\left(S^2\right)=\left(\frac{N}{N-1}\right)\left(\frac{n-1}{n}\right)\sigma^2\Rightarrow E\left(S^2\right)=\left(\frac{1000}{1000-1}\right)\left(\frac{200-1}{200}\right)(06)^2=35.85$$

التباين: $$\sigma^2\left(S^2\right)=\frac{2(n-1)}{n^2}\sigma^4\Rightarrow\sigma^2\left(S^2\right)=\frac{2(200-1)}{(200)^2}(06)^4=12.8952$$

2-حساب التوقع الرياضي والتباين لتباين العينة المعدل $\hat{S}^2$

التوقع الرياضي: نعلم ان $E\left(\hat{S}^2\right)=\left(\frac{N}{N-1}\right)\sigma^2 \Rightarrow \left(\frac{1000}{1000-1}\right)(06)^2=36.03$

التباين: نعلم ان $\sigma^2\left(\hat{S}^2\right)=\left(\frac{2}{n-1}\right)\sigma^4 \Rightarrow \sigma^2\left(\hat{S}^2\right)=\left(\frac{2}{200-1}\right)(06)^4=13.025$

تمرين22: سحبت عينة عشوائية بحجم 20 مشاهدة من مجتمع يتوزع توزيع طبيعي بمتوسط حسابي μ وتباين 09.

المطلوب:

1-أحسب احتمال ان يزيد تباين العينة عن 15؟

الحل:

لدينا من المعطيات: $n=20, X\sim N(\mu,09)$

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي μ وتباين 09

لحساب الاحتمال التالي، يجب استخدام احصاءة مربع كاي والتي تعطى كالتالي:

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

نحسب درجة الحرية $V=n-1 \Rightarrow V=20-1=19$

$$P\left(S^2>15\right)=P\left(\frac{nS^2}{\sigma^2}>\frac{20\times15}{09}\right)=P\left(\chi^2_{19}>33.34\right)$$

بالبحث عن هذه القيمة بجدول مربع كاي عند درجة حرية19 لا نجدها وانما تكون محصورة بين قيمتين ولايجاد الاحتمال المقابل لها نستخدم نظرية الحصر بالنقصان كالتالي:

36.191	33.34	32.852
0.99	α	0.975

$$(36.191-32.852)\longrightarrow(0.99-0.975)$$

$$(33.34-32.852)\longrightarrow(\alpha-0.975)$$

بالطرح نجد:

$$(03.339)\longrightarrow(0.015)$$

$$(0.488)\longrightarrow(\alpha-0.975)$$

إذن قيمة α هي:

$$\alpha=\frac{0.488\times0.015}{03.339}+0.975=0.977$$

إذن $P\left(\chi^2_{19}>33.34\right)=01-P\left(\chi^2_{19}<33.34\right)=01-0.977=0.023$

2-7 توزيع المعاينة للانحراف المعياري العينة S

عندما نسحب عددا كبيرا من العينات العشوائية متساوية الحجم من مجتمع إحصائي ما، نحصل من كل منها على قيمة للانحراف المعياري، فتشكل هذه القيم من الانحرافات المعيارية لتلك العينات فيما بينها توزيعا تكراريا يطلق عليه بتوزيع معاينة الانحرافات المعيارية، وهذا التوزيع يقترب من التوزيع الطبيعي مع تزايد حجم العينة.

ويطلق على الانحراف المعياري للتوزيع التكراري لعدد كبير من جدا من الانحرافات المعيارية لعينات عشوائية من حجم واحد، ومن مجتمع إحصائي واحد بالخطأ المعياري للانحرافات المعيارية ويرمز له σ_s والذي يحسب من العلاقات التالية:

-إذا كان الانحراف المعياري للمجتمع معلوما: $\sigma_s = \frac{\sigma}{\sqrt{2n}}$

-أما إذا كان الانحراف المعياري للمجتمع غير معلوم، يستعاض عنه بدلالة الانحراف المعياري للعينة كالتالي: $\sigma_s = \frac{S}{\sqrt{2n}}$

أما إذا كان المجتمع لا يخضع للتوزيع الطبيعي فان الانحراف المعياري يعطى بالعلاقة التالية:

$$\sigma_s = \sqrt{\frac{\mu_4 - \sigma^4}{4n\sigma^2}}$$

تمرين23: إذا كان لدينا مجتمع يمثل أعمار خمسة أطفال، وكانت أعمارهم على التوالي 03،02،06،08،11، وقمنا بسحب جميع العينات العشوائية البسيطة من حجم طفلين من هذا المجتمع.

1-أوجد المتوسط الحسابي للمجتمع وتباين المجتمع؟

2-أوجد الانحراف المعياري لتوزيع معاينة الانحرافات المعيارية؟

الحل:

1-حساب المتوسط الحسابي والتباين للمجتمع

نعلم ان المتوسط الحسابي يعطى بالصيغة التالية:

$$\mu = \frac{\sum_{i=1}^{5} X_i}{N} = \frac{02+03+06+08+11}{05} = 06$$

أما التباين فيحسب بالصيغة التالية:

$$\sigma^2 = \frac{\sum_{i=1}^{05}(X_i-\mu)^2}{N} = \frac{(02-06)^2+(03-06)^2+(06-06)^2+(08-06)^2+(11-06)^2}{05} = 10.80$$

2-حساب الانحراف المعياري لتوزيع معاينة الانحرافات المعيارية

نحسب اولا الانحراف المعياري للمجتمع كالتالي: $\sigma^2 = 10.80 \Rightarrow \sigma = \sqrt{10.80} = 03.29$

بتطبيق قانون الانحراف المعياري لتوزيع المعاينة للانحرافات المعيارية نجد:

$$\sigma_s = \frac{\sigma}{\sqrt{2n}} \Rightarrow \sigma_s = \frac{03.29}{\sqrt{2\times 2}} = 01.643$$

8- توزيع المعاينة للنسبة بين تبايني عينتين S_1^2 و S_2^2

نظرية22: إذا كان لدينا S_1^2 و S_2^2 تباين عينتين عشوائيتين بحجم n_1 و n_2 مسحوبة من مجتمعين يتوزعان توزيع طبيعي بتباين σ_1^2 و σ_2^2 على التوالي فان النسبة $\left[\frac{S_1^2}{S_2^2}\right]$ لتباين العينتين تتوزع توزيع فيشر (F) بدرجة حرية V_1 و V_2 بحيث ان:

$$F=\frac{\left[\frac{S_1^2 n_1}{n_1-1}\right]\frac{1}{\sigma_1^2}}{\left[\frac{S_2^2 n_2}{n_2-1}\right]\frac{1}{\sigma_2^2}}=\frac{\frac{\hat{S}_1^2}{\sigma_1^2}}{\frac{\hat{S}_2^2}{\sigma_2^2}}\sim F_{n_1-1,n_2-1}$$

مقدمة عن توزيع فيشر(F): إن الاساس النظري لاشتقاق توزيع فيشر (F) هو في الحقيقة توزيع مشتق من نسبة توزيعين مستقلين كل منهما عبارة عن توزيع مربع كاي (χ^2) مقسوما على درجة حريته.

وعلى إفتراض كان لدينا المتغير العشوائي (χ_1^2) كمتغير عشوائي يخضع الى توزيع مربع كاي بدرجة حرية (n_1-1) حيث ان:

$$\chi_1^2=\frac{n_1 S_1^2}{\sigma_1^2}=\frac{(n_1-1)\hat{S}_1^2}{\sigma_1^2}\sim\chi^2_{n_1-1}$$

وكان لدينا ايضا (χ_2^2) كمتغير عشوائي يخضع الى توزيع مربع كاي بدرجة حرية (n_2-1) حيث ان:

$$\chi_2^2=\frac{n_2 S_2^2}{\sigma_2^2}=\frac{(n_2-1)\hat{S}_2^2}{\sigma_2^2}\sim\chi^2_{n_2-1}$$

ومما تقدم يمكن الحصول على المتغير العشوائي (F) من خلال النسبة بين التوزيعين السابقين مقسوما كل منهما على درجة حريته، حيث سيعتمد المتغير الجديد (F) على نوعين من درجات الحرية تتمثل في بدرجات الحرية للبسط (n_1-1) ودرجات الحرية للمقام (n_2-1) أي ان

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

خصائص توزيع فيشر (F): يتصف توزيع فيشر بالخصائص التالية

1-قيم المتغير العشوائي لفيشر لا تكون سالبة (F>0)، فيبدا منحنى التوزيع من الصفر على المحور الافقي، ويمتد الى اليمين وعند القيم الكبيرة يقترب من المحور الافقي ولكن لا يلتقيان.

2-منحنى توزيع فيشر يشبه توزيع مربع كاي فهو ملتو ناحية اليمين.

3-المساحة الكلية تحت منحنى التوزيع تساوي الواحد الصحيح.

4-يقترب توزيع فيشر من التوزيع الطبيعي كلما زادت درجتا الحرية.

5-يعتمد توزيع فيشر على معلمتين وهما درجة البسط ($V_1=n_1-1$) ودرجة حرية المقام ($V_2=n_2-1$)

تمرين24: إذا كانت S_1^2 و S_2^2 تمثل التباينات لعينتين عشوائيتين مستقلتين عن بعضهما البعض بحجم $n_1=16$ و $n_2=21$ سحبت من مجتمعين يتوزعان توزيعا طبيعيا بتباين قدره $\sigma_1^2=09$، $\sigma_2^2=16$ على التوالي

المطلوب:

1-أحسب الاحتمال التالي $P(\frac{S_1^2}{S_2^2}\geq 01,74)$

الحل:

بما ان المطلوب هو نسبة تباين عينتين إذن سوف نستخدم توزيع فيشر.

نحسب اولا درجة الحرية للبسط والمقام كالتالي:

$$V_1=n_1-1\Rightarrow V_1=16-01=15$$
$$V_2=n_2-1\Rightarrow V_2=21-01=20$$

نحسب الان الإحتمال

$$P\left(\frac{S_1^2}{S_2^2}>01.74\right)=P\left[\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}>01.74\frac{\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}\right]=P\left(F_{15,20}>01.74\frac{\left(\frac{16}{16-1}\right)\frac{1}{09}}{\left(\frac{21}{21-1}\right)\frac{1}{16}}\right)$$
$$=P\left(F_{15,20}>03.14\right)=01-P\left(F_{15,20}<03.14\right)=01-0.99=0.01$$

تمارين المحور الاول

التمرين الأول: إذا كان عدد سائقي سيارات الأجرة في مدينة ما هو 2000 سائق، وعلمت أن أعمارهم تتوزع توزيعا طبيعيا بمتوسط حسابي قدره 50 سنة، وتباين قدره 49 سنة، فإذا سحبنا مع عدم الإرجاع من هذا المجتمع عينة بها 25 سائقا.

المطلوب:

1-أوجد توزيع المعاينة لمتوسط أعمار سائقي سيارات الأجرة في العينة؟

2-أحسب احتمال أن يكون متوسط العمر لهذه العينة محصورا بين 54 و48.

3-نفرض أن حجم العينة المسحوبة مع عدم الإرجاع تغير وأصبح 144 سائقا.

3-1 أوجد توزيع المعاينة للمتوسط أعمار سائقي سيارات الأجرة في العينة؟ واحسب الاحتمال السابق؟

التمرين الثاني: تخضع أوزان علب سائل غسيل معين لتوزيع وسطه الحسابي 1000غ وانحراف معياري قدره 77غ، إذا أخذت عينة عشوائية حجمها 35 علبة.

المطلوب:

1- أوجد توزيع المعاينة لأوزان العلب في للعينة.

2-احسب احتمال أن يزيد المتوسط الحسابي لأوزان العلب في العينة عن 980غ؟

التمرين الثالث: مصنع لإنتاج نوع معين من المصابيح الكهربائية، إذا علم أن عمر المصباح من إنتاج هذا المصنع يخضع للتوزيع الطبيعي بمتوسط حسابي قدره 800 ساعة وانحراف معياري قدره 60 ساعة، سحبت عينة عشوائية حجمها 64 من إنتاج هذا المصنع.

المطلوب:

1- أوجد توزيع المعاينة لمتوسط عمر المصابيح في العينة ؟

2-أحسب احتمال أن يكون متوسط العمر في المصابيح محصور بين 790 ساعة و815 ساعة.

3- إذا فرضنا أن الانحراف المعياري لعمر المصابيح مجهولا، وأخذت عينة عشوائية حجمها 26 مصباح وكان الانحراف المعياري للعينة المسحوبة هو 45 ساعة.

3-1 أوجد توزيع المعاينة لمتوسط عمر المصابيح في العينة ؟ واحسب الاحتمال السابق؟

4-إذا فرضنا أن الانحراف المعياري لعمر المصابيح مجهولا، وأخذت عينة عشوائية حجمها 32 مصباح وكان الانحراف المعياري للعينة المسحوبة هو 40 ساعة.

4-1 أوجد توزيع المعاينة لمتوسط عمر المصابيح في العينة ؟ واحسب الاحتمال السابق؟

5-إذا فرضنا أن الانحراف المعياري لعمر المصابيح هو30 ساعة، وأخذت عينة عشوائية حجمها 25 مصباح.

5-1 أوجد توزيع المعاينة لمتوسط عمر المصابيح في العينة ؟ واحسب الاحتمال السابق؟

التمرين الرابع: مجتمعان يتوزعان توزيعا طبيعيا، الأول حجمه 1000 وسطه الحسابي 30 وتباينه 18، والثاني حجمه 1400 وسطه الحسابي 24 وتباينه 40، فإذا سحبنا من كل مجتمع عينة عشوائية مع عدم الإرجاع وكانت العينتين مستقلتين عن بعضهما البعض، وكان حجم العينة الأولى 120 وحجم العينة الثانية 250.

المطلوب:

1- أوجد توزيع المعاينة للفرق بين متوسطي العينتين ؟

2-أحسب احتمال أن يكون الفرق بين متوسطي العينتين اقل من 05؟

3-نفرض أن حجم العينتين المسحوبتين مع عدم الإرجاع تغير وأصبح في العينة الأولى 50 وفي العينة الثانية 70.

3-1أوجد توزيع المعاينة للفرق بين متوسطي العينتين ؟ واحسب الاحتمال السابق؟

التمرين الخامس: يقوم مصنع بإنتاج نوعين من المصابيح الكهربائية A وB المصابيح من النوع A تخضع في إنتاجها إلى التوزيع الطبيعي بمتوسط عمر 700 ساعة وتباين σ_1^2، فإذا تم سحب عينة عشوائية من النوع الأول تقدر ب 18 مصباحا بانحراف معياري 30، أما المصابيح من النوع B فهي الأخرى تخضع للتوزيع الطبيعي بمتوسط عمر 500 ساعة وتباين، σ_2^2فإذا تم سحب عينة عشوائية من النوع الثاني تقدر ب 12 مصباحا بانحراف معياري 25 ساعة.

المطلوب:

1- أوجد توزيع المعاينة للفرق بين متوسطي العينتين ، بافتراض تجانس تبايني المجتمعين؟

2-أحسب احتمال أن يكون الفرق بين متوسطي عمر المصابيح اكبر من 215؟

3-أوجد توزيع المعاينة للفرق بين متوسطي العينتين ، بافتراض عدم تجانس تبايني المجتمعين؟

4- أحسب احتمال أن يكون الفرق بين متوسطي عمر المصابيح اكبر من 215؟

التمرين السادس:إذا كان أداء العمال يخضع للتوزيع الطبيعي، وفي تجربة لبيان تحسن هذا الأداء تم سحب عينة عشوائية بحجم 13عاملا في المعمل فكانت قياس الكفاءة قبل وبعد دخولهم دورة تحسين الأداء كما هو مبين في الجدول التالي:

بعد (X_i)	07	08	07	07	08	08	09	09	09	08	07	07	08
قبل(Y_i)	05	05	05	04	07	08	05	06	07	08	05	06	05

المطلوب:

1- أوجد توزيع المعاينة للفرق بين متوسطي العينتين المرتبطتين ؟

2-أحسب احتمال أن يكون الفرق في الأداء قبل وبعد الدورة لا يقل عن 02.50؟

التمرين السابع: في دراسة قام بها مركز متخصص لمعالجة آثار الإدمان على التدخين في مجتمع

إحصائي متكون من 3000 شخص، فإذا علمت أن نسبة المدخنين تخضع لتوزيع ذي الحدين وتمثل 58.27%، و سحبنا عينة عشوائية من هؤلاء المدخنين مع عدم الإرجاع تشمل 100 مدخنا.

المطلوب:

1- أوجد توزيع المعاينة لنسبة المدخنين في هذه المدينة؟

2-احسب احتمال أن تكون نسبة المدخنين في هذه العينة تتراوح بين 55% و70%؟

3-نفرض أن حجم العينة المسحوبة مع عدم الإرجاع تغير وأصبح 180 مدخنا.

3-1 أوجد توزيع المعاينة لنسبة المدخنين في هذه المدينة؟ واحسب الاحتمال السابق؟

التمرين الثامن: إذا كانت نسبة الأشخاص الذين يحملون صنف الدم من نوع B في مجتمع هو 0.30 فإذا تم اختيار عينة عشوائية حجمها 500 شخص

المطلوب:

1-أوجد توزيع المعاينة لنسبة الأشخاص الذين يحملون صنف الدم من نوع B؟

2-أحسب احتمال أن يكون الأشخاص الذين يحملون صنف الدم من نوع B في الحالتين التاليتين:

2-1 أكبر من 0.35

3-1 محصور بين 0.25 و 0.35

التمرين التاسع: إذا علمت أن شركة متخصصة في صناعة الالكترونيات تستورد قطع الغيار اللازمة لها من مصدرين مختلفين وان القطع الواردة إليها من المصدر الأول ترفض بمعدل 08% بسبب عيوب بها، بينما القطع الواردة إليها من المصدر الثاني ترفض بمعدل 05% لنفس السبب، فإذا كانت خطوط الإنتاج تستخدم 150 قطعة من المصدر الأول و300 قطعة من المصدر الثاني يوميا.

المطلوب:

1- أوجد توزيع المعاينة للفرق بين نسبتي العينتين؟

2-أحسب احتمال أن يكون الفرق ما بين نسبتي العينتين اقل من أو يساوي 07%؟

التمرين العاشر: إذا كان لدينا مجتمع يخضع للتوزيع الطبيعي و مكون من القيم التالية (01، 03، 05، 07، 09)، وسحبنا منه العينات العشوائية الممكنة والمتساوية في الحجم وليكم حجمها (n=02)

المطلوب:

1-أحسب المتوسط الحسابي للمجتمع وتباينه؟

2-إيجاد كل العينات الممكنة ذات الحجم (n=02) الممكن سحبها من هذا المجتمع مع الإرجاع؟

3-أحسب القيمة المتوقعة لتباين العينة S^2 في حالة السحب مع الإرجاع ومقارنتها بتباين المجتمع؟

4- إيجاد كل العينات الممكنة ذات الحجم (n=02) الممكن سحبها من هذا المجتمع مع عدم الإرجاع؟

5- أحسب القيمة المتوقعة لتباين العينة S^2 في حالة السحب مع عدم الإرجاع ومقارنتها بتباين المجتمع؟

6-أحسب القيمة المتوقعة لتباين العينة $\hat{S}^2$ في حالة السحب بالإرجاع وعدم الإرجاع؟ ومقارنته بتباين المجتمع؟

7-إيجاد توزيع المعاينة $\frac{nS^2}{\sigma^2}$ في حالة السحب بالإرجاع وعدم الإرجاع؟

8- إيجاد توزيع المعاينة $\frac{n\hat{S}^2}{\sigma^2}$ في حالة السحب بالإرجاع وعدم الإرجاع؟

المحور الثاني: نظرية التقدير

تمهيد

إن أحد أهداف النظرية الإحصائية هو إيجاد الطرق التي تمكننا من الاستدلال على قيمة معالم المجتمع وذلك عن طريق النتائج التي تحصل عليها من عينة عشوائية نختارها من بين مفردات المجتمع.
فإذا كنا نرغب في تقدير أحد معالم المجتمع وليكن θ عن طريق عينة من المشاهدات $x_1, x_2, x_3 \ldots\ldots x_n$ مسحوبة من المجتمع فإن القيمة التي يتم حسابها للمعلمة θ من واقع هذه المشاهدات تسمى تقديرا (Estimate) بينما الدالة أو الصيغة الرياضية الإحصائية التي تستخدم للوصول إلى هذا التقدير تسمى مقدرا (Estimator)، والمقدر هو دالة تعتمد على المشاهدات ولتكن $\hat{\theta}=f\left(X_1,X_2,X_3,\ldots\ldots\ldots\ldots,X_n\right)$ بينما التقدير هو قيمة هذه الدالة عند التعويض بقيم المشاهدات فيها، ولهذا فإن التقدير يختلف من عينة إلى أخرى رغم استخدام نفس المقدر وهذا أمر طبيعي حيث أن هناك اختلافا بين قيم المشاهدات من عينة إلى أخرى رغم أن المقدر له نفس الصيغة التي يتم التعويض فيها.
يمكن تقدير معلمة المجتمع بقيمة واحدة وهي قيمة المقدر التي نحصل عليها من بيانات لعينة، فالتقدير الذي نحصل عليه عن طريق إحصاء العينة والذي أطلقنا عليه تعبير المقدر يمثل قيمة تقديرية لمعلمة المجتمع مقدرة بقيمة واحدة، وبصورة أخرى فانه يمكن إيجاد قيمة تقديرية واحدة لمعلمة المجتمع ويسمى هذا التقدير بتقدير النقطة.
والتقدير في علم الإحصاء لا تختلف أسسه عن هذا المنطق حيث يكون لدينا توزيعا احتماليا معلوم الصيغة ولكن مجهول المعلمات وباختيار عينة عشوائية من هذا التوزيع (المجتمع) والحصول على دالة لقيم العينة يستخدمها كمقدر لمعلمة المجتمع المجهولة، هذه الدالة سميت بالمقدر لمعلمة المجتمع والقيمة التي يتم الحصول عليها من هذه الدالة تستخدم كتقدير لقيمة المعلمة.
فالوسط الحسابي $\overline{X}$ للعينة العشوائية هو تقدير لمتوسط المجتمع μ، وتباين العينة S^2 هو تقدير لتباين المجتمع σ^2، ونسبة صفة معينة في العينة $\hat{P}$ هي تقدير لنسبة نفس الصفة في المجتمعP.
وتتلخص مشكلة التقدير في استخدام بيانات عينة عشوائية مسحوبة من المجتمع في إيجاد تقدير معالم المجتمع المجهولة، وهناك نوعان من التقدير هما:

أ–التقدير النقطي

ب–التقدير ضمن فترة.

وسوف نستعرض كلا التقديرين بنوع من التفصيل كالتالي:

01–التقدير النقطي (التقدير بقيمة واحدة)

المقصود بهذا النوع من التقدير، هو تقدير معلمة المجتمع المجهولة بإحصائية تحسب قيمتها من بيانات عينة عشوائية مسحوبة من هذا المجتمع، أي نقوم بتقدير المعلمة بقيمة واحدة فقط تحسب من العينة، ولذلك يسمى التقدير بقيمة، فمثلا نقدر الوسط الحسابي للمجتمع μ بالوسط الحسابي للعينة $\overline{X}$، ونقدر

تباين المجتمع σ^2 بتباين العينة S^2 وهكذا......الخ.

وبصفة عامة إذا كان لدينا مجتمع معلمته المجهولة θ ونستخدم لتقديرها دالة تعتمد على بيانات العينة العشوائية $x_1,x_2,x_3......x_n$ يرمز لها بالرمز $\hat{\theta}$ بحيث الإشارة (^) تعني في علم الإحصاء القيمة المقدرة وبالتالي: $\hat{\theta}=f\left(X_1,X_2,X_3,..............,X_n\right)$

فهذه الدالة يطلق عليها مقدر أي المقدر هو عبارة عن إحصائية، وعند التعويض في هذه الدالة ببيانات العينة المسحوبة، فالقيمة العددية التي نحصل عليها تسمى تقديرا، أي أن المقدر متغير عشوائي تتغير قيمته من عينة إلى أخرى بينما التقدير هو إحدى قيم هذا المتغير.

مثال:إذا كانت المعلمة المجهولة هي الوسط الحسابي للمجتمع μ، فيقدر بالوسط الحسابي للعينة $\overline{X}$ أي أن: $\hat{\mu}=\overline{X}=\dfrac{\sum_{i=1}^{n}X_i}{n}$

فإذا سحبنا عينة عشوائية من هذا المجتمع، وحسبنا وسطها الحسابي ووجدنا أن قيمته 27 مثلا فان:

μ: هو المعلمة المجهولة θ

$\overline{X}$: هي مقدر بقيمة $\hat{\theta}$

$\overline{X}=27$: هي تقدير بقيمة.

وإذا كانت المعلمة المجهولة هي تباين المجتمع σ^2، فيقدر بتباين العينة S^2 أي أن $\hat{\sigma}^2=S^2=\dfrac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)}{n-1}$

فإذا سحبنا عينة عشوائية من هذا المجتمع، وحسبنا تباينها ووجدنا أن قيمته هي 06 فإن:

σ^2: هي المعلمة المجهولة θ.

S^2: هي مقدر بقيمة $\hat{\theta}$.

$S^2=06$: هي تقدير بقيمة.

02-خصائص المقدر الجيد(معيار فيشر للمقدر الجيد)

إذا أخذنا مقدرا معينا فان هذا المقدر لا يعطي قيما قريبة جدا من قيمة المعلمة لجميع العينات الممكن اختيارها وعلى ذلك فلا يمكن الحكم على المقدر في ضوء تقدير محسوب من عينة واحدة ذا قد تكون هذه القيمة قريبة جدا من القيمة الحقيقية وقد تكون بعيدة عنها.

ولذلك يعتبر المقدر جيدا إذا كان في المتوسط (كعدد كبير من العينات) يعطي قيما قريبة من القيمة الحقيقية أو بعبارة أخرى إذا كان يعطي في غالب الأحيان قيما تكون قريبة جدا من القيمة ونادرا ما يعطي قيما بعيدة جدا عن القيمة الحقيقية.

ولكي يكون المقدر جيد يجب أن يحقق بعض المعايير، وترجع كلمة المقدر الجيد إلى العالم فيشر (Fisher)،ويعتبر المقدر مقدرا جيدا طبقا لمعيار فيشر إذا توافرت فيه الخصائص التالية:

1-خاصية عدم التحيز(Unbiasness): يقال إن المقدر غير متحيز إذا كان توقعه الرياضي يساوي قيمة المعلمة الحقيقية للمجتمع، فإذا كانت معلمة المجتمع هي θ وكان $\hat{\theta}$ هو المقدر المحسوب من عينة $X_1,X_2,X_3,......,X_n$ مسحوبة من المجتمع فان المقدر $\hat{\theta}$ غير متحيز إذا كان: $E\left(\hat{\theta}\right)=\theta$

أما إذا كان $E\left(\hat{\theta}\right)=\theta+b$ فان المقدار $b=E\left(\hat{\theta}\right)-\theta$ يسمى مقدار التحيز.

كما انه من النادر أن نجد تقديرا يكون متوسط مربع خطأه اقل ما يمكن لجميع قيم θ فإذا كان $\hat{\theta}$ تقديرا للمعلمة θ فان متوسط مربع الخطأ هو $MSE\left(\hat{\theta}\right)=E\left(\hat{\theta}-\theta\right)^2$.

1-1تعريف متوسط مربع الخطأ(Mean Squared error): إذا كانت $X=\left(X_1,X_2,X_3,......,X_n\right)$ عينة عشوائية مسحوبة من مجتمع X توزيعه الاحتمالي $f\left(X,\theta\right)$ يعتمد على معلمة θ، وذا أخذنا $\hat{\theta}=f(X)$ كتقدير للمعلمة θ فان الخطأ في هذا التقدير أو الخسارة الناجمة من اتخاذ $\hat{\theta}$ كتقدير للمعلمة θ يمكن أن يقاس بالمقدار $\left(\hat{\theta}-\theta\right)$، ولكن هذا المقدار يمكن أن يكون موجبا أو سالبا حسب قيمة $\hat{\theta}$ المحسوبة من العينة، وحتى نتخلص من الإشارة فإننا نأخذ مربع هذا المقدار أي $\left(\hat{\theta}-\theta\right)^2$ كمقياس لقرب التقدير $\hat{\theta}$ من المعلمة θ ولكن هذا المقدار يتغير من عينة لأخرى لذلك فإننا نأخذ القيمة المتوقعة له أي نأخذ القيمة المتوقعة له $E\left(\hat{\theta}-\theta\right)^2$، والتي تسمى متوسط مربع خطا التقدير ل $\hat{\theta}$ ويرمز له بالرمز $MSE\left(\hat{\theta}\right)$ أي أن $MSE\left(\hat{\theta}\right)=E\left(\hat{\theta}-\theta\right)^2$ كمقياس لجودة المقدر $\hat{\theta}$ للمعلمة θ.

ويمكن إعادة كتابته على الصورة التالية

$$MSE\left(\hat{\theta}\right)=E\left[\left(\hat{\theta}-E\left(\hat{\theta}\right)\right)+\left(E\left(\hat{\theta}\right)-\theta\right)\right]^2$$

$$=E\left(\hat{\theta}-E\left(\hat{\theta}\right)\right)^2+E\left(E\left(\hat{\theta}\right)-\theta\right)^2+2\left\{E\left(\hat{\theta}\right)-\theta\right\}E\left\{\hat{\theta}-E\left(\hat{\theta}\right)\right\}$$

$$=V\left(\hat{\theta}\right)+\left(E\left(\hat{\theta}\right)-\theta\right)^2+2\left\{E\left(\hat{\theta}\right)-\theta\right\}E\left\{\hat{\theta}-E\left(\hat{\theta}\right)\right\}$$

الحد الأخير في الطرف الأيمن يساوي صفرا لان $E\left\{\hat{\theta}-E\left(\hat{\theta}\right)\right\}=E\left(\hat{\theta}\right)-E\left(\hat{\theta}\right)=0$

وعلى ذلك فان $MSE\left(\hat{\theta}\right)=V\left(\hat{\theta}\right)+b^2(\theta)$

حيث $b^2(\theta)=\left(E\left(\hat{\theta}\right)-\theta\right)^2$

أي أن متوسط مربع الخطأ لأي تقدير يساوي تباين التقدير مضافا إليه مقدار آخر هو مربع $b^2(\theta)$

يسمى المقدار $b(\theta)=E\left(\hat{\theta}\right)-\theta$ بمقدار التحيز في الإحصاء $\hat{\theta}$ كتقدير للمعلمة θ وهو يبين مقدار الاختلاف بين متوسط توزيع الإحصاء $\hat{\theta}$ وبين المعلمة θ، ويكون متوسط مربع الخطأ أقل ما يمكن إذا كان كل من التباين ومربع التحيز أقل ما يمكن، وهذا يعني أن التقدير ل $\hat{\theta}$ يكون قريبا من المعلمة θإذا كان كل من التباين ومربع التحيز اقل ما يمكن.

وعلى ذلك فإننا نفضل التقدير الذي له اقل متوسط مربع خطا، فإذا كان $\hat{\theta}_1$ و $\hat{\theta}_2$ مقدران للمعلمة θ فإننا نفضل $\hat{\theta}_1$ عن $\hat{\theta}_2$ إذا كان $MSE\left(\hat{\theta}_1\right)<MSE\left(\hat{\theta}_2\right)$

2-الاتساق (Consistency): يقال بأن $\hat{\theta}$ مقدرا متسقا لمعلمة المجتمع θ إذا كانت $\hat{\theta}$ تؤول إلى θ (أي تقترب منها) كلما زاد حجم العينة أي $\lim\limits_{n\to\infty} P\left(\left|\hat{\theta}_n-\theta\right|>\varepsilon\right)=0$ أو بشكل مكافئ

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n-\theta\right|\leq\varepsilon\right)=1$$

ويكون $\hat{\theta}$ مقدرا متسقا إذا كانت $\hat{\theta}$ تقديرا غير متحيزا للمعلمة θ ويقترب تباينه من الصفر كلما زاد حجم العينة.

معنى أن تؤول $\hat{\theta}$ إلى θ هو أن التوزيعات العينية للمقدر $\hat{\theta}$ لأحجام العينات المتتابعة في الكبر سوف تولد لنا متسلسلة من متوسطات $\hat{\theta}$ ومتسلسلة من تباينات $\hat{\theta}$ بحيث أن متسلسلة المتوسطات تؤول إلى θ ونهاية متسلسلة التباينات تؤول إلى الصفر مع زيادة n زيادة لا نهائية.

هذا ويمكن ملاحظة أن المقدر المتسق يكون غير متحيزا في النهاية بينما المقدر الغير متحيز قد يكون أو لا يكون مقدرا متسقا وعموما يكون أي مقدر متسقا إذا كان:

$$\begin{cases} \lim\limits_{n\to\infty} E\left(\hat{\theta}\right)=\theta \\ \lim\limits_{n\to\infty} V\left(\hat{\theta}\right)=0 \end{cases}$$

3-خاصية الكفاءة (Efficiency): قد يوجد للمعلمة الواحدة أكثر من مقدر ويمكننا المقارنة بين هذه المقدرات من خلال المقارنة بين تبايناتهم حيث نعتبر أن المقدر الأقل تباينا هو المقدر الأكثر كفاءة، فإذا كان لدينا المقدران $\hat{\theta}_1$ و $\hat{\theta}_2$ للمعلمة θ وكان تباين المقدر الأول هو $V\left(\hat{\theta}_1\right)$ وتباين المقدر الثاني هو $V\left(\hat{\theta}_2\right)$ فان المقدر $\hat{\theta}_1$ يكون أكثر كفاءة من المقدر $\hat{\theta}_2$ ، وبفرض أن $\hat{\theta}_1$ و $\hat{\theta}_2$ مقدرات غير متحيزة أي $V\left(\hat{\theta}_1\right)<V\left(\hat{\theta}_2\right)$ ،وأن الكفاءة النسبية(Relative efficiency) للمقدران $\hat{\theta}_1$ و $\hat{\theta}_2$ هي

كالتالي: $e=\frac{V\left(\hat{\theta}_1\right)}{V\left(\hat{\theta}_2\right)}$ ، فإذا كانت قيمتها اقل من الواحد الصحيح فان $\hat{\theta}_1$ يكون مقدرا ذي تباين اقل من $\hat{\theta}_2$ ويكون هو الأكفأ، وإذا كانت النسبة اكبر من الواحد فان $\hat{\theta}_2$ يكون هو الأكفأ.

4-خاصية الكفاية (Sufficienc): يقال أن المقدر $\hat{\theta}$ مقدرا كافيا للمعلمة θ إذا كان تقدير $\hat{\theta}$ يحتوي على كل المعلومات الموجودة في العينة عن المعلمة θ، وهذا يعني انه بعد معرفة $\hat{\theta}$ فان المعلومات المتبقية في العينة لا تفيد في معرفة θ.

فإذا كانت $\left(X_1,X_2,X_3,......,X_n\right)$ عينة عشوائية مسحوبة من مجتمع يتبع التوزيع الاحتمالي $f\left(X,\theta\right)$ فان احتمال اختيار العينة هو: $f\left(x_1,x_2,x_3,......,x_n,\theta\right)=\prod_{i=1}^{n}f\left(x_i,\theta\right)$

وإذا أمكن تحليل هذا الاحتمال إلى حاصل ضرب دالتين تعتمد الأولى على المعلمة والتقدير والثانية لا تعتمد على المعلمة أي إذا كان $f\left(x_1,x_2,........,x_n,\theta\right)=f_1\left(\hat{\theta},\theta\right)\times f_2\left(x_1,x_2,.....,x_n\right)$

فإن هذا يعني أن كل المعلومات عن θ والمتوفرة في العينة قد لخصها وامتصها واحتوى عليها التقدير $\hat{\theta}$ بحيث أن أي تقدير آخر لا يحتوي على هذا القدر من المعلومات التي يحتويها التقدير الكافي، وعلى ذلك يكون التقدير الكافي هو أفضل التقديرات وخاصة إذا كان غير متحيز.

5-خاصية التراصف والتركيز(Closeness): يقال عن المقدر $\hat{\theta}_1$ لتقدير الوسيط θ بأنه أكثر تراصفا أو تركيزا من المقدر $\hat{\theta}_2$ في تقديره إذا كان احتمال أو درجة الثقة بالأول أكبر من الثاني أي:

$$P\left(\theta-\lambda<\hat{\theta}_1\le\theta+\lambda\right)>P\left(\theta-\lambda<\hat{\theta}_2\le\theta+\lambda\right)$$

ولكفافة قيم $\lambda>0$

ومن جانب آخر يمكن استخدام تعريف بتمان (Bitman) في هذا الصدد بحيث:

$$P\left(\left|\hat{\theta}_2-\theta\right|<\left|\hat{\theta}_1-\theta\right|\right)\ge\frac{1}{2}$$

<u>تمرين1:</u> ليكن X_1 و X_2 عينة عشوائية مأخوذة من مجتمع إحصائي له التوزيع الطبيعي $X\sim N(\mu,1)$ ولتكن T_1،T_2،T_3 ثلاث مقدرات ل μ ، معرفة بالعلاقات التالية:

$$T_3=\frac{1}{4}X_1+\frac{3}{4}X_2,T_2=\frac{1}{3}X_1+\frac{2}{3}X_2,T_1=\frac{1}{2}X_1+\frac{1}{2}X_2$$

<u>المطلوب:</u>

1-أي المقدرات الثلاثة غير متحيزة ل μ؟

2-أي المقدرات الثلاثة أكفأ؟

الحل:

1–تحديد أي المقدرات الثلاثة غير متحيز ل μ، نوجد التوقع الرياضي للمقدرات الثلاثة كالتالي:

$$E(T_1)=E\left(\frac{1}{2}X_1+\frac{1}{2}X_2\right)=E\left(\frac{1}{2}X_1\right)+E\left(\frac{1}{2}X_2\right)$$

$$=\frac{1}{2}E(X_1)+\frac{1}{2}E(X_2)=\frac{1}{2}\mu+\frac{1}{2}\mu=\mu$$

$$E(T_2)=E\left(\frac{1}{3}X_1+\frac{2}{3}X_2\right)=E\left(\frac{1}{3}X_1\right)+E\left(\frac{2}{3}X_2\right)$$

$$=\frac{1}{3}E(X_1)+\frac{2}{3}E(X_2)=\frac{1}{3}\mu+\frac{2}{3}\mu=\mu$$

$$E(T_3)=E\left(\frac{1}{4}X_1+\frac{3}{4}X_2\right)=E\left(\frac{1}{4}X_1\right)+E\left(\frac{3}{4}X_2\right)$$

$$=\frac{1}{4}E(X_1)+\frac{3}{4}E(X_2)=\frac{1}{4}\mu+\frac{3}{4}\mu=\mu$$

أي أن المقدرات الثلاثة T_3،T_2،T_1 هي مقدرات غير متحيزة ل μ.

2–تحديد أي المقدرات الثلاثة هو مقدر كفؤ: من أجل تحديد أي من المقدرات الثلاثة هو المقدر الكفء نحسب التباين لكل مقدر كالتالي

المقدر الأول T_1

$$V(T_1)=V\left(\frac{1}{2}X_1+\frac{1}{2}X_2\right)=V\left(\frac{1}{2}X_1\right)+V\left(\frac{1}{2}X_2\right)$$

وذلك لان المتغيرين العشوائيين مستقلين، وبالاستفادة من خواص التباين نجد:

$$V(T_1)=\frac{1}{4}V(X_1)+\frac{1}{4}V(X_2)=\frac{1}{4}(1)+\frac{1}{4}(1)=\frac{1}{2}$$

المقدر الثاني T_2

$$V(T_2)=V\left(\frac{1}{3}X_1+\frac{2}{3}X_2\right)=V\left(\frac{1}{3}X_1\right)+V\left(\frac{2}{3}X_2\right)$$

وذلك لان المتغيرين العشوائيين مستقلين، وبالاستفادة من خواص التباين نجد

$$V(T_2)=\frac{1}{9}V(X_1)+\frac{4}{9}V(X_2)=\frac{1}{9}(1)+\frac{4}{9}(1)=\frac{5}{9}$$

المقدر الثالث T_3

$$V(T_3)=V\left(\frac{1}{4}X_1+\frac{3}{4}X_2\right)=V\left(\frac{1}{4}X_1\right)+V\left(\frac{3}{4}X_2\right)$$

وذلك لان المتغيرين العشوائيين مستقلين، وبالاستفادة من خواص التباين نجد

$$V(T_3)=\frac{1}{16}V(X_1)+\frac{9}{16}V(X_2)=\frac{1}{16}(1)+\frac{9}{16}(1)=\frac{10}{16}$$

نلاحظ أن $V(T_1)=\frac{1}{2}<V(T_2)=\frac{5}{9}<V(T_3)=\frac{10}{16}$

أي أن المقدر T_1 أفضل من المقدر T_2 وكلاهما أفضل من المقدر T_3.

تمرين02: إذا كان $X=(X_1,X_2,X_3,.......,X_n)$ عينة عشوائية مسحوبة من مجتمع يخضع للتوزيع الطبيعي بمتوسط حسابي μ و تباين σ^2

المطلوب:

1-أثبت أن $\overline{X}=\frac{\sum_{i=1}^{n}X_i}{n}$ تقديرا غير متحيز ومتسقا وكافيا للمعلمة μ.

الحل:

إثبات أن $\overline{X}$ هو تقدير غير متحيز

التوقع الرياضي $$E(\overline{X})=E\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}\mu=\frac{1}{\cancel{n}}\times\cancel{n}\mu=\mu$$

التباين $$V(\overline{X})=V\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=\frac{1}{n^2}V\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}V(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=\frac{1}{n\times\cancel{n}}\times\cancel{n}\sigma^2=\frac{\sigma^2}{n}$$

وبالتالي يكون $\overline{X}$ تقديرا غير متحيز للمعلمة μ وان تباينه هو $\frac{\sigma^2}{n}$

-إثبات أن $\overline{X}$ هو تقدير متسق

$$\lim_{n\to\infty}E(\overline{X})=\mu$$

$$\lim_{n\to\infty}V(\overline{X})=\lim_{n\to\infty}\frac{\sigma^2}{n}=0$$

إذن $\overline{X}$ هو تقدير متسق للمعلمة μ

- إثبات أن $\overline{X}$ هو تقدير كافيا

$$f(x_1,x_2,..........,x_n)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\Pi\sigma}}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$=\left(\frac{1}{\sqrt{2\Pi\sigma}}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

وحيث أن $$\sum_{i=1}^{n}(X_i-\mu)^2=\sum_{i=1}^{n}(X_i-\overline{X})^2+n(\overline{X}-\mu)^2$$

إذن
$$=\left(e^{-\frac{n(\overline{X}-\mu)}{2\sigma^2}}\right)\left(\frac{1}{\sqrt{2\Pi\sigma}}\right)^n e^{-\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{2\sigma^2}}$$

الحد الأول على اليسار يعتمد فقط على μ، $\overline{X}$، بينما الحد الثاني خال من المعلمة μ وعلى ذلك يكون $\overline{X}$ تقديرا كافيا للمعلمة μ.

03-التقدير بفترة(التقدير المجالي)

مقدمة: إن المقدر الذي نستخدمه عند تقدير معلمة المجتمع تختلف نتائجه التقديرية من عينة إلى أخرى (رغم استخدام نفس الصيغة الإحصائية في التقدير لكل عينة يتم اختيارها)، وكما نعلم فإننا نختار عينة واحدة نستخدمها في تقدير معلمة المجتمع ونأمل أن يكون التقدير المحسوب منها أقرب ما يكون للمعلمة الحقيقية التي لا نعلم قيمتها بالتحديد، ولكن ليس هناك ما يدعو إلى الاعتقاد بان معلمة المجتمع سوف تساوي تماما قيمة التقدير بنقطة الذي نحصل عليه من الصيغة، لذلك يكون من الأفضل وضع حد أعلى وحد أدنى للتقدير بحيث يمكن القول أن معلمة المجتمع التي لا نعلم قيمتها بدقة تقع بين هذين الحدين وبصورة أخرى يمكن تكوين فترة ثقة تحدد بعددين بحيث نتوقع أن تقع معلمة المجتمع داخل هذه الفترة باحتمال معين يطلق عليه معامل الثقة.

3-1 تعريف التقدير ضمن فترة (التقدير بمجال): هو التقدير الذي يتألف من قيمتين عدديتين تعرفان مجال من القيم الذي نتوقع أن يتضمن المعلمة المطلوب تقديرها بدرجة ثقة محدودة.

إذن بصفة عامة، وفقا لهذا التعريف لإيجاد فترة الثقة للمعلمة المجهولة ولتكن مثلا θ يجب إيجاد إحصائيتين ولتكونا L و U بحيث يكون $P(L \leq \theta \leq U)=1-\alpha$

وهاتين الإحصائيتين تكونا متغيرين عشوائيين وذلك لأنها دوال في بيانات العينة، وان الفترة الناتجة من ذلك ستكون $L \leq \theta \leq U$ ، ويطلق عليها تسمية $(1-\alpha)100\%$ فترة الثقة لتقدير المعلمة المجهولة θ أو مستوى الثقة ، ويطلق على "L " تسمية الحد الأدنى لفترة الثقة بينما يطلق على "U" الحد الأعلى لفترة الثقة، ويطلق على تسمية معامل الثقة $1-\alpha$ ، كما وتسمى α أو مستوى الدلالة أو مستوى المعنوية.

ولإيجاد فترة الثقة المطلوبة نتبع الخطوات التالية:

1-نفتش عن كمية تكون تابعة في متغيرات العينة وتحوي الوسيط المجهول ويكون توزيعها هو أحد التوزيعات الشهيرة، تسمى هذه الكمية بالكمية المحورية.

2-نحصر هذه الكمية المحورية بين قيمتين معياريتين مأخوذتين من توزيع العينة بحيث يكون احتمال أن تقع هذه الكمية المحورية المحصورة بين هاتين القيمتين المعياريتين يساوي $1-\alpha$.

3-نعزل الوسيط المجهول ونوجد فترة الثقة المطلوبة.

04- فترة الثقة حول المتوسط (μ)

4-1 فترة الثقة حول المتوسط (μ) عندما يكون التباين σ^2 معلوما

نظرية01: إذا كان لدينا مجتمع حجمه لا محدود بمتوسط حسابي μ و تباين σ^2 (ليس من الضروري ان يكون توزيعه توزيعا طبيعيا) وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم n بحيث تكون كبيرة بما فيه الكفاية $(n\geq 30)$، فإن توزيع المعاينة للوسط الحسابي للعينة $\overline{X}$ سيكون قريبا من من التوزيع الطبيعي(نظرية النهاية المركزية) بوسط حسابي وتباين قدرهما على التوالي $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

أما إذا كان المجتمع يتوزع توزيعا طبيعيا، فان توزيع المعاينة للمتوسط الحسابي للعينة $\overline{X}$ سيكون توزيعا طبيعيا، وذلك سواء كان حجم العينة n صغيرا او كبيرا، وفي هذه الحالة نستطيع تحويل $\overline{X}$ الى المتغير المعياري Z كالتالي

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

ونلاحظ هنا أن Z هي دالة في المقدر $\overline{X}$، والمعلمة المجهولة هي μ والتوزيع الاحتمالي للمتغير Z هو التوزيع الطبيعي المعياري الذي لا يعتمد على μ إذن Z هي كمية محورية نستطيع استعمالها في إيجاد فترة الثقة للمتوسط الحسابي للمجتمع μ.

ويعطى مجال الثقة للمتوسط الحسابي للمجتمع μ عندما يكون تباين المجتمع معلوما او عند مستوى ثقة $(1-\alpha)100\%$ كالتالي

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

بحيث:

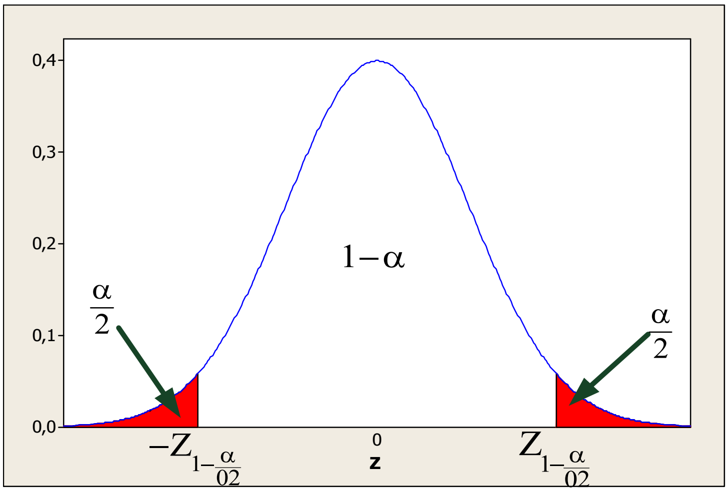
يسمى المقدار $\overline{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ بالحد الادنة لفترة الثقة.

يسمى المقدار $\overline{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ بالحد الاعلى لفترة الثقة.

يسمى المقدار $Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ بخطأ التقدير.

البرهان:

لدينا الشكل رقم01 الذي يوضح مجال الثقة



ومن خلال الشكل يتضح أن

$$P\left(-Z_{1-\frac{\alpha}{2}} \le Z \le Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

وبتعويض بقيمة المتغير Z في العلاقة السابقة نجد

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

وبإجراء العمليات الجبرية على الحدود الثلاثة للمتباينة الموجودة بين قوسين، بحيث نجعل الحد الأوسط للمتباينة يحتوي على المعلمة المجهولة فقط، ويتم ذلك بضرب الأطراف الثلاثة للمتباينة في $\frac{\sigma}{\sqrt{n}}$ فنحصل على ما يلي:

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

ثم بطرح $\overline{X}$ من الأطراف الثلاثة للمتباينة نجد:

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

وبضرب الأطراف الثلاثة للمتباينة في (-1) وإعادة ترتيب أطراف المتباينة نتحصل على ما يلي:

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

ويمكن كتابة العلاقة الأخيرة بالكتابة المختصرة كالتالي:

$$CI(\mu)_{(1-\alpha)} = \overline{X} \pm Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

–أما ذا كانت المعاينة من مجتمع (محدود) والسحب مع عدم الإرجاع مع تحقق العلاقة $\frac{n}{N} > 0.05$ يتم استبدال العلاقة السابقة بإضافة معامل الإرجاع أو التصحيح كالتالي:

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

والجدول التالي يوضح مختلف القيم والدرجات المعيارية التي يأخذها مجال الثقة كالتالي:

درجة الثقة $(1-\alpha)100\%$	$(1-\alpha)$	α	$Z_{1-\frac{\alpha}{2}}$
%90	0,90	0,10	01,645
%95	0,95	0,05	01,960
%99	0,99	0,01	02,575

تمرين03: سحبت عينة عشوائية من100 مصباح كهربائي من إنتاج أحد المصانع فوجد أن المتوسط الحسابي لعمر المصباح هو 1000 ساعة، فإذا علمت أن الانحراف المعياري لعمر المصباح في المجتمع هو 150 ساعة.

المطلوب:

1-إيجاد تقدير نقطة لمتوسط عمر المصباح ؟

2-إيجاد تقدير فترة الثقة 95% لمتوسط عمر المصباح في هذا المصنع؟وماذا تستنتج؟

3- إيجاد تقدير فترة الثقة 99% لمتوسط عم المصباح في هذا المصنع؟ وماذا تستنتج؟

4-بمقارنة النتائج المتحصل عليها في السؤالين 1و2، استنتج العلاقة بين مستوى الثقة وطول فترة الثقة؟

5- إيجاد تقدير فترة الثقة 95% لمتوسط عمر المصباح إذا كان حجم العينة 200 مصباح؟

الحل:

لدينا من معطيات التمرين ما يلي: $n=100, \overline{X}=1000, \sigma=150$

1-إيجاد تقدير النقطة لمتوسط عمر المصباح: $\hat{\mu}=\overline{X}=1000$

2- إيجاد تقدير فترة الثقة 95% لمتوسط عمر المصباح في هذا المصنع.

بما ان حجم العينة أكبر من30 فيمكننا تطبيق نظرية النهاية المركزية، ومن ثم فان توزيع المعاينة يتبع التوزيع الطبيعي بصرف النظر عن التوزيع الأصلي للمجتمع، ومن ثم يمكننا حساب فترة الثقة $(1-\alpha)100\%$ لمتوسط مجتمع معلوم التباين بالعلاقة التالية:

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

بما أن $\%95=\%100(1-\alpha)$

إذن $1-\alpha=0.95 \Leftarrow \alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\Rightarrow L=1000-(1.96)\frac{150}{\sqrt{100}}=970.60$$

$$U=\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\Rightarrow U=1000+(1.96)\frac{150}{\sqrt{100}}=1029.40$$

بالتعويض في العلاقة الأساسية لمجال التقدير نجد

$$P(970.60 \le \mu \le 1029.40) = 0.95$$

الاستنتاج: نحن متأكدون بدرجة ثقة 95% أن عمر المصباح سوف لن يقل عن 970.60 ساعة ولا يزيد عن 1029.40 ساعة.

3- إيجاد تقدير فترة الثقة 99% لمتوسط عمر المصباح في هذا المصنع: بتطبيق العلاقة الأساسية في التقدير

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

بما أن $\%99 = \%100(1-\alpha)$

لدينا $1-\alpha = 0.99 \Leftarrow \alpha = 0.01$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,01}{2}} = Z_{0,995} = 2,575$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L = \overline{X} - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \Rightarrow L = 1000 - (2.575)\frac{150}{\sqrt{100}} = 931.375$$

$$U = \overline{X} + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \Rightarrow U = 1000 + (2.575)\frac{150}{\sqrt{100}} = 1038.625$$

بالتعويض في العلاقة الأساسية لمجال التقدير نجد

$$P(961.375 \le \mu \le 1038.625) = 0.99$$

الاستنتاج: نحن متأكدون بدرجة ثقة 99% أن عمر المصباح سوف لن يقل عن 931.375 ساعة ولا يزيد عن 1038.625 ساعة.

4-استنتاج العلاقة بين مستوى الثقة وطول فترة الثقة: من النتائج المتحصل عليها في السؤالين 1و2 نجد:

طول فترة الثقة عند مستوى ثقة 95%=1029.40−970.60=58.80

طول فترة الثقة عند مستوى ثقة 99%=1038.625−961.375=77.25

نستنتج أن العلاقة بين مستوى الثقة وطول فترة الثقة هي علاقة طردية.

4-2 فترة الثقة حول المتوسط (μ) عندما يكون التباين مجهولا

بالرغم من انه في بعض الأحيان تكون μ غير معلومة و σ معلومة، إلا انه في الحقيقة وفي كثير من الحالات تكون σ غير معلومة أيضا،وإذا أردنا تكوين فترة ثقة حول μ عندما تكون σ غير معلومة فان المدخل المناسب هو تقدير σ باستخدام S (الانحراف المعياري للعينة) وفي هذه الحالة فان فترة الثقة حول المتوسط μ تتوقف على حجم العينة ، وبالتالي نميز حالتين هما إما أن تكون$(n\geq30)$ وفي هذه الحالة يتم التقدير باستخدام التوزيع الطبيعي المعياري أما إذا كان $(n<30)$ يتم التقدير باستخدام توزيع ستودنت بدرجة حرية $(V=n-1)$.

نظرية02: إذا كان لدينا مجتمع حجمه (لا محدود) يخضع للتوزيع الطبيعي بمتوسط حسابي μ مجهولة وتباين σ^2 مجهول أيضا، أخذت عينة عشوائية $X=(X_1,X_2,X_3,.......,X_n)$ حجمها $(n\geq 30)$، في هذه الحالة يؤخذ S^2 تباين العينة كتقدير غير متحيز لتباين المجتمع،و يعطى مجال الثقة للمتوسط الحسابي للمجتمع μ عندما يكون تباين المجتمع σ^2 مجهولا أيضا أو عند مستوى ثقة $(1-\alpha)100\%$ بالعلاقة التالية:

$$P\left(\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\leq\mu\leq\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\right)=1-\alpha$$

او بشكل مكافئ قدره

$$P\left(\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{\hat{S}^2}{\sqrt{n}}\leq\mu\leq\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{\hat{S}^2}{\sqrt{n}}\right)=1-\alpha$$

- أما ذا كانت المعاينة من مجتمع (محدود) والسحب مع عدم الإرجاع مع تحقق العلاقة $\frac{n}{N}>0.05$ يتم استبدال العلاقة السابقة بإضافة معامل الإرجاع أو التصحيح كالتالي:

$$P\left(\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\leq\mu\leq\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\right)=1-\alpha$$

او بشكل مكافئ قدره

$$P\left(\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{\hat{S}^2}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\leq\mu\leq\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{\hat{S}^2}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right)=1-\alpha$$

تمرين04: نظرا للشكاوى التي تتلقاها إدارة الشركة المنتجة للنضائد السائلة بخصوص العمر الزمني الذي تعمره النضائد المنتجة لها قامت باختيار عينة عشوائية تتكون من 50 نضيدة من بين عدد كبير جدا من هذه النضائد فوجدت أن متوسط العمر الزمني لهذه العينة يساوي 02.67 سنة وانحراف معياري يساوي 01.94.

المطلوب:

1-أوجد 95% فترة ثقة حول متوسط العمر الزمني الذي تعمره هذه النضائد بصفة عامة؟

الحل:

لدينا من معطيات التمرين ما يلي: $n=50,\overline{X}=02.67,S=01.94$

1-إيجاد تقدير فترة الثقة 95 % لمتوسط العمر الزمني الذي تعمره هذه النضائد

بما أن تباين المجتمع مجهول وحجم العينة كبير(n=50) اكبر من 30، إذن يمكن استخدام التوزيع الطبيعي المعياري، وبتطبيق العلاقة الأساسية في التقدير

$$P\left(\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\leq\mu\leq\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\right)=1-\alpha$$

بما أن $\%100(1-\alpha)=\%95$

لدينا $1-\alpha=0.95 \Rightarrow \alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=\overline{X}-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}} \Rightarrow L=2.67-(1.96)\frac{1.94}{\sqrt{50-1}}=02.1268$$

$$U=\overline{X}+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}} \Rightarrow U=2.67+(1.96)\frac{1.94}{\sqrt{50-1}}=03.2132$$

بالتعويض في العلاقة الأساسية لمجال التقدير نجد

$$P(02.1268 \leq \mu \leq 03.2132)=0.95$$

نظرية03: إذا كان لدينا مجتمع حجمه (لا محدود) يخضع للتوزيع الطبيعي بمتوسط حسابي μ مجهولة وتباين σ^2 مجهول أيضا، أخذت عينة عشوائية $X=(X_1,X_2,X_3,......,X_n)$ حجمها (n<30)، في هذه الحالة يؤخذ S^2 تباين العينة كتقدير غير متحيز لتباين المجتمع،و يعطى مجال الثقة للمتوسط الحسابي للمجتمع μ عندما يكون تباين المجتمع σ^2 مجهولا أيضا أو عند مستوى ثقة $(1-\alpha)100\%$ بالعلاقة التالية:

$$P\left(\overline{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}} \leq \mu \leq \overline{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\right)=1-\alpha$$

او بشكل مكافئ قدره

$$P\left(\overline{X}-T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}^2}{\sqrt{n}} \leq \mu \leq \overline{X}+T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}^2}{\sqrt{n}}\right)=1-\alpha$$

- أما ذا كانت المعاينة من مجتمع (محدود) والسحب مع عدم الإرجاع مع تحقق العلاقة $\frac{n}{N}>0.05$ يتم استبدال العلاقة السابقة بإضافة معامل الإرجاع أو التصحيح كالتالي:

$$P\left(\overline{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \overline{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\right)=1-\alpha$$

او بشكل مكافئ قدره

$$P\left(\overline{X}-T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}^2}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \overline{X}+T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}^2}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right)=1-\alpha$$

تمرين05: إذا علمت أن متوسط أوزان عينة تتكون من 16 طفلا أعمارهم 10 سنوات يساوي 32.40 كغ، وبانحراف معياري يساوي 05.40 كغ.

المطلوب:

1- أوجد 99% فترة ثقة لتقدير المتوسط الحسابي (μ) لأوزان مجتمع العينة على افتراض أن مجتمع المعاينة مجتمع طبيعي؟

الحل:

لدينا من معطيات التمرين ما يلي: $n=16, \overline{X}=32.40, S=05.40$

1-إيجاد 99% فترة ثقة لتقدير المتوسط الحسابي (μ) لأوزان مجتمع العينة

بما أن المجتمع يخضع للتوزيع الطبيعي بتباين مجهول وحجم العينة المأخوذة اقل من30 فان عملية التقدير سيتم إجرائها باستخدام توزيع ستودنت (T) وبالعلاقة التالية:

$$P\left(\overline{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\leq\mu\leq\overline{X}+T_{1-\frac{\alpha}{2},\alpha}\frac{S}{\sqrt{n-1}}\right)=1-\alpha$$

بما أن $\%99=\%100(1-\alpha)$

لدينا $1-\alpha=0.99\Leftarrow\alpha=0.01$ ودرجة الحرية (V=n-1=16-1=15) وبالتالي فان قيمة ستودنت تساوي $T_{1-\frac{\alpha}{2},n-1}=T_{1-\frac{0.01}{2},15}=T_{0.995,15}=02.947$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=\overline{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\Rightarrow L=32.40-(02.947)\frac{05.40}{\sqrt{16-1}}=28.291$$

$$U=\overline{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\Rightarrow L=32.40+(02.947)\frac{05.40}{\sqrt{16-1}}=36.509$$

بالتعويض في العلاقة الأساسية لمجال التقدير نجد

$$P(28.291\leq\mu\leq36.509)=0.99$$

3-4 تحديد حجم العينة للنوعيات لتقدير النسبة في المجتمع: العينة عبارة عن جزء من المجتمع، وبالتالي فان مقدر القيمة المحسوب من العينة لا نتوقع انه يساوي القيمة الحقيقية للمعلمة المجهولة التي يقدرها، والفرق بين تقدير القيمة لأي معلمة وبين القيمة الحقيقية للمعلمة يطلق عليه خطأ المعاينة لان هذا الخطأ ينتج بسبب استخدامنا للعينة بدلا من المجتمع ككل، وعند تقدير الوسط الحسابي للمجتمع μ بمقدر القيمة $\overline{X}$ أي بالوسط الحسابي للعينة ، فسيكون خطأ المعاينة في هذه الحالة مساويا إلى $(\overline{X}-\mu)$ ، ونعلم مما سبق أن فترة الثقة للمتوسط الحسابي للمجتمع تعطى بالعلاقة التالية:

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

فإذا كانت μ فعلا تقع في مركز الفترة فان $\overline{X}$ تقدر μ وبدون خطأ، ولكن في معظم الأحيان $\overline{X}$ لا تساوي μ بالضبط وبالتالي سيكون هناك خطأ في التقدير بقيمة واحدة، وحجم هذا الخطأ يساوي الفرق مابين μ و $\overline{X}$ وبالتالي نحن على ثقة قدرها $\%100(1-\alpha)$ بأن هذا الفرق سوف لن يكون أكبر من $Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ والذي سوف يرمز له بالرمز E إذن $E=Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$

ويتحدد حجم العينة وفق الحالات التالية:

-**المجتمع لا محدود ويخضع للتوزيع الطبيعي:** في هذه الحالة حجم العينة هو $n=\frac{Z_{1-\frac{\alpha}{2}}^{2}\sigma^{2}}{E^{2}}$

-**المجتمع محدود والسحب مع عدم الإرجاع و يخضع للتوزيع الطبيعي:** في هذه الحالة حجم العينة هو

$$n=\frac{NZ_{1-\frac{\alpha}{2}}^{2}\sigma^{2}}{(N-1)E^{2}+Z_{1-\frac{\alpha}{2}}^{2}\sigma^{2}}$$

-**المجتمع لا محدود ويخضع لتوزيع ستودنت:** في هذه الحالة حجم العينة هو $n=\frac{T_{1-\frac{\alpha}{2},n-1}^{2}\sigma^{2}}{E^{2}}$

- **المجتمع محدود والسحب مع عدم الإرجاع و يخضع لتوزيع ستودنت:** في هذه الحالة حجم العينة هو

$$n=\frac{NT_{1-\frac{\alpha}{2},n-1}^{2}\sigma^{2}}{(N-1)E^{2}+T_{1-\frac{\alpha}{2},n-1}^{2}\sigma^{2}}$$

-أما ذا كانت σ^2 مجهولة فيمكن تقديرها باستخدام تباين العينة S^2 المحسوب من عينة مبدئية ذات حجم صغير متفق عليه، أو يمكن تقديرها من عينة سبق الحصول عليها من دراسات سابقة أو دراسات مماثلة، وإذا كان توزيع المجتمع معتدلا تقريبا فان:

المدى=06 σ، كما أن المدى =أكبر قيمة-أصغر قيمة إذن

الانحراف المعياري(σ)=المدى/06 $\sigma=\frac{Rang}{06}=\frac{Max(X_i)-Min(X_i)}{06}$

<u>**تمرين06:**</u> كم يجب أن يكون حجم العينة العشوائية n التي يجب أن تؤخذ من مجتمع إحصائي للمحركات والذي يخضع للتوزيع الطبيعي لتحديد المتوسط الزمني الذي تستغرقه آلة معينة في تحريك عجلة سيارة على أساس 95% من الثقة، بان أعلى خطأ في تقدير المعلمة μ هو 0.50 دقيقة، وإذا علمت من تجارب سابقة ميدانية بان تباين المجتمع الإحصائي معروف ويساوي 02.56 دقيقة.

<u>**الحل:**</u>

لدينا من معطيات التمرين $1-\alpha=95\%, E=0.50, \sigma^2=02.56, N=\infty$

لدينا $1-\alpha=0.95 \Leftarrow \alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

بتطبيق العلاقة الأساسية لحساب العينة في حالة مجتمع لا محدود ويخضع للتوزيع الطبيعي نجد

$$n=\frac{Z_{1-\frac{\alpha}{2}}^{2}\sigma^{2}}{E^{2}}\Rightarrow n=\frac{(01.96)^{2}(02.56)}{(0.50)^{2}}\simeq 40$$

أي أننا يجب أن نفحص 40 محركا من مجموع كافة المحركات على الأقل كي يكون أكبر خطأ في التقدير هو 0.50

تمرين07: ما حجم العينة العشوائية الواجب سحبها من مجتمع الأغنام لتقدير وسطي وزن الخروف، مع العلم أن حجم المجتمع هو 2000 وتباين المجتمع هو 600، وان أعلى خطا في تقدير المعلمة μ هو 05، وان قيمة الدرجة المعيارية هو 03.

الحل:

لدينا من معطيات التمرين $Z=03, E=05, \sigma^2=600, N=2000$

بتطبيق العلاقة الأساسية لحساب العينة في حالة مجتمع محدود والسحب مع عدم الإرجاع ويخضع للتوزيع الطبيعي نجد

$$n=\frac{NZ^2_{1-\frac{\alpha}{2}}\sigma^2}{(N-1)E^2+Z^2_{1-\frac{\alpha}{2}}\sigma^2} \Rightarrow n=\frac{(2000)(03)^2(600)}{(2000-1)(05)^2+(03)^2(600)}=195$$

تمرين08: يرغب أحد خبراء الكفاءة باستخدام المتوسط الحسابي لعينة عشوائية ذات حجم 150 لتقدير متوسط الكفاءة الميكانيكية لعدد معين من العمال في إحدى الصناعات الكبيرة للسيارات، فإذا افترضنا وبناءا على الخبرة بان الانحراف المعياري للمجتمع هو 06.20.

المطلوب:

1-كم يكون أعلى خطأ في تقدير المتوسط الحسابي μ للمجتمع العمالي على أساس 99% من الثقة؟

2-إذا افترضنا بان المتوسط الحسابي للعينة السابقة هو 69.50، فكم هو تقدير المتوسط الحسابي μ لكفاءة العمال كافة في تلك الصناعة وذلك بنفس الدرجة 99% من الثقة؟

الحل:

لدينا من معطيات التمرين: $n=150, \sigma=06.20, 1-\alpha=0.99, \overline{X}=69.50$

1-تحديد أعلى خطأ في تقدير المتوسط الحسابي μ للمجتمع العمالي على أساس 99% من الثقة.

نعلم أن أعلى خطأ في تقدير المتوسط الحسابي ل μ يعطى بالعلاقة التالية: $E=Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$

لدينا $1-\alpha=0.99 \Leftarrow \alpha=0.01$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,01}{2}}=Z_{0,995}=2,575$

إذن

$$E=Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \Rightarrow E=\frac{(02.575)(06.20)}{\sqrt{150}}=01.30$$

2-تقدير المتوسط الحسابي μ لكفاءة العمال كافة في تلك الصناعة وذلك بنفس الدرجة 99% من الثقة

نعلم أن $CI(\mu)_{(1-\alpha)}=\overline{X}\pm Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ وبالتالي فان التقدير هو

$$CI(\mu)_{(1-\alpha)}=\overline{X}\pm Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \Rightarrow \begin{cases}\mu_1=\overline{X}+E=69.50-1.30=68.20\\ \mu_2=\overline{X}-E=69.50+1.30=70.80\end{cases}$$

تمرين 09: ترغب وزارة الصناعة في معرفة متوسط عمر العمال في صناعة ما، وكان لديها معلومات تفيد بان العمال يبدؤن العمل في الثامنة عشرة من العمر، ويحالون على التقاعد في الستين منه.

المطلوب:

1- ما هو حجم العينة المراد سحبها لتقدير هذا المتوسط، إذا ربنا في التأكد من كل الحالات العملية على أن لا يزيد الخطأ في التقدير عن 01.50 سنة؟

الحل:

من معطيات التمرين لدينا: $E=01.50, Max(X_i)=60, Min(X_i)=18, Z=03$

بما أن تباين المجتمع مجهول، يمكننا تقديره ب 6/1 المدى العددي الذي يمكن أن يقع به متوسط الظاهرة المدروسة.

نحسب قيمة المدى

$$Rang = Max(X_i) - Min(X_i) \Rightarrow Rang = 60 - 18 = 42$$

إذن قيمة الانحراف المعياري سوف تساوي $\sigma = \frac{Rang}{06} = \frac{42}{06} = 07$

وبتطبيق العلاقة الأساسية في حساب حجم العينة نجد $n = \frac{Z^2_{1-\frac{\alpha}{2}}\sigma^2}{E^2} \Rightarrow n = \frac{(03)^2(07)^2}{(01.50)^2} \simeq 196$

05- فترة الثقة بإستخدام نظرية تشيبيشيف (Techebecheff)

تستخدم نظرية تشيبيشيف إذا كان توزيع المجتمع مجهول أو لا يخضع للتوزيع الطبيعي وحجم العينة المسحوبة من المجتمع اقل تماما من 30 (n<30) وتعطى معادلة التقدير بالعلاقة التالية

$$P\left(\left|\overline{X} - \mu\right| \le K\sigma\right) \ge 1 - \frac{1}{K^2}$$

بحيث $1 - \alpha = 1 - \frac{1}{K^2}$

ويتم التقدير بالعلاقة التالية $CI(\mu)_{(1-\alpha)} = \overline{X} \pm K\frac{\sigma}{\sqrt{n}}$

تمرين10: ليكن لدينا دخول أفراد مؤسسة ما بمتوسط حسابي μ وانحراف معياري 30 ون، أخذت عينة عشوائية بسيطة حجمها 25، فكان متوسطها الحسابي 80 ون.

المطلوب:

1- أوجد فترة الثقة لمتوسط أجور العمال عند مستوى ثقة 95%؟

الحل:

بما أن المجتمع لا يخضع للتوزيع الطبيعي وحجم العينة صغير(أقل تماما من30) إذن لا يمكننا استخدام توزيع ستودنت، وإنما يتم استخدام احد الاختبارات اللامعلمية ألا وهو متراجحة تشبيشيف.

لدينا من معطيات التمرين $\sigma = 30, n = 25, \overline{X} = 80, 1 - \alpha = 0.95$

لتطبيق علاقة التقدير السابقة باستخدام متراجحة تشيبيشيف نستخرج أولا قيمة K كالتالي:

$$1-\alpha=1-\frac{1}{K^2}\Rightarrow 0.95=1-\frac{1}{K^2}\Rightarrow K=\sqrt{20}=02\sqrt{05}$$

بعد ذلك نستخرج فترة التقدير

$$P\left(\overline{X}-K\frac{\sigma}{\sqrt{n}}\leq\mu\leq\overline{X}+K\frac{\sigma}{\sqrt{n}}\right)=1-\alpha\Rightarrow$$

$$P\left(80-\left(02\sqrt{05}\right)\frac{30}{\sqrt{25}}\leq\mu\leq 80+\left(02\sqrt{05}\right)\frac{30}{\sqrt{25}}\right)=0.95\Rightarrow$$

$$P\left(53.1672\leq\mu\leq 106.8328\right)=0.95$$

06- فترة الثقة حول الفرق بين متوسطي مجتمعين مستقلين $\left(\mu_1-\mu_2\right)$

لإيجاد فترة الثقة حول الفرق بين متوسطي مجتمعين مستقلين فإننا نميز الحالات التالية:

6-1 فترة الثقة حول الفرق بين متوسطي مجتمعين معلومي التباين

نظرية04:إذا افترضنا أن $\left(X_1,X_2,X_3,......,X_{n_1}\right)$ عينة عشوائية من مجتمع يخضع للتوزيع الطبيعي بمتوسط حسابي μ_1 غير معلوم وتباين σ_1^2 معلوم وكانت $\left(X_1,X_2,X_3,......,X_{n_2}\right)$ عينة عشوائية من مجتمع إحصائي ثاني متوسطه الحسابي μ_2 غير معلوم وتباين σ_2^2 معلوم، وكانت العينتين مستقلتين عن بعضهما البعض وكان $\overline{X}_1$ يمثل متوسط العينة الأولى و $\overline{X}_2$ تمثل متوسط العينة الثانية، وكان مجتمعي العينتين طبيعيين (أو كان حجم كل عينة أكبر من 30، إذا كان المجتمعين غير طبيعيين)، فإن توزيع المعاينة للفرق ما بين متوسطي العينتين $\left(\overline{X}_1-\overline{X}_2\right)$ سوف يكون له توزيع قريب من التوزيع الطبيعي بمتوسط حسابي $\mu_{\left(\overline{X}_1-\overline{X}_2\right)}=\mu_1-\mu_2$ وتباين $\sigma^2_{\left(\overline{X}_1-\overline{X}_2\right)}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ وبالتالي فان المتغير العشوائي (Z) يعطى كالتالي:

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

ونلاحظ أن Z هي دالة في المقدر $\left(\overline{X}_1-\overline{X}_2\right)$ والمعلمة المجهولة $\left(\mu_1-\mu_2\right)$ ، والتوزيع الاحتمالي للمتغير Z هو التوزيع الطبيعي المعياري الذي لا يعتمد على المعلمة المجهولة $\left(\mu_1-\mu_2\right)$، إذن Z هي كمية محورية نستطيع استعمالها في إيجاد فترة الثقة للفرق بين متوسطي مجتمعين $\left(\mu_1-\mu_2\right)$، وباستخدام خواص التوزيع الطبيعي المعياري نجد أن

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

وبالتعويض في العلاقة السابقة بقيمة Z نجد:

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

وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة الموجودة بين قوسين، بحيث نجعل الأوساط للمتباينة تحتوي فقط على المعلمة المجهولة $\left(\mu_1-\mu_2\right)$ وهي الفرق بين متوسطي المجتمعين، فنحصل على علاقة التقدير التالية:

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

وبالتالي فان فترة الثقة للفرق بين متوسطي مجتمعين $\left(\mu_1-\mu_2\right)$ عندما يتوزع المجتمعان توزيعا طبيعيا بتباينين معلومين وعند مستوى ثقة $(1-\alpha)100\%$ هي

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

-أما إذا كان المجتمعان محدودان وكسر المعاينة للمجتمع الأول هو $\frac{n_1}{N_1}>0.05$ وكسر المعاينة للمجتمع الثاني هو $\frac{n_2}{N_2}>0.05$ فان مجال التقدير يعطى بالعلاقة التالية

$$CI\left(\mu_1-\mu_2\right)_{(1-\alpha)}=\left(\overline{X}_1-\overline{X}_2\right)\pm Z_{1-\frac{\alpha}{2}}\left(\sqrt{\frac{\sigma_1^2}{n_1}\times\frac{N_1-n_1}{N_1-1}}+\sqrt{\frac{\sigma_2^2}{n_2}\times\frac{N_2-n_2}{N_2-1}}\right)$$

تمرين11: في دراسة خاصة بمقارنة متوسط الدخل الشهري للأسر القاطنة بمدينة تيارت بمتوسط الدخل الشهري للأسر القاطنة بمدينة مستغانم، فإذا كان تباين الدخول في مدينة تيارت هو 6400 ون، وتباين الدخول في مدينة مستغانم هو 3600ون،فإذا اخترنا من مدينة تيارت عينة عشوائية تحتوي على 400 أسرة ووجدنا أن متوسط الدخل الشهري لهذه الأسر يساوي250ون، واخترنا من مدينة مستغانم عينة عشوائية مستقلة عن العينة السابقة تحتوي على 300 أسرة ووجدنا أن الدخل الشهري لهذه الأسر يساوي 210 ون

المطلوب:

1-أحسب فترة الثقة للفرق بين متوسطي الدخول في المدينتين عند درجة ثقة 90%؟

الحل:

لدينا من معطيات التمرين

المجتمع الأول مدينة تيارت: $\sigma_1^2=6400, n_1=400, \overline{X}_1=250, 1-\alpha=0.90$

المجتمع الثاني مدينة مستغانم: $\sigma_2^2=3600, n_2=300, \overline{X}_2=210, 1-\alpha=0.90$

إيجاد فترة الثقة عند مستوى معنوية 90%

بتطبيق العلاقة الأساسية في التقدير نجد

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

نستخرج أولا مقدار الدرجة المعيارية كالتالي

لدينا $1-\alpha=0.90\Leftarrow\alpha=0.10$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,10}{2}}=Z_{0,950}=1,645$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=\left(\overline{X}_1-\overline{X}_2\right)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\Rightarrow L=(250-210)-(1,645)\left(\sqrt{\frac{6400}{400}+\frac{3600}{300}}\right)=31,30$$

$$U=\left(\overline{X}_1-\overline{X}_2\right)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\Rightarrow U=(250-210)+(1,645)\left(\sqrt{\frac{6400}{400}+\frac{3600}{300}}\right)=48.70$$

إذن فترة التقدير هي $P\left(31,30\le\left(\mu_1-\mu_2\right)\le48,70\right)=0,90$

6-2 فترة الثقة حول الفرق بين متوسطي مجتمعين مجهولي التباين وحجم العينتين كبير

نظرية05: إذا افترضنا أن $\left(X_1,X_2,X_3,\dots\dots,X_{n_1}\right)$ عينة عشوائية من مجتمع يخضع للتوزيع الطبيعي بمتوسط حسابي μ_1 غير معلوم وتباين σ_1^2 معلوم وكانت $\left(X_1,X_2,X_3,\dots\dots,X_{n_2}\right)$ عينة عشوائية من مجتمع إحصائي ثاني متوسطه الحسابي μ_2 غير معلوم وتباين σ_2^2 معلوم، وكانت العينتين مستقلتين عن بعضهما البعض وكان $\overline{X}_1$ يمثل متوسط العينة الأولى و S_1^2 يمثل تباين العينة الأولى الغير متحيز لتباين المجتمع و $\overline{X}_2$ تمثل متوسط العينة الثانية، و S_2^2 يمثل تباين العينة الثانية الغير متحيز لتباين المجتمع وكان مجتمعي العينتين طبيعيين (أو كان حجم كل عينة أكبر من 30، إذا كان المجتمعين غير طبيعيين)، فإن توزيع المعاينة للفرق ما بين متوسطي العينتين $\left(\overline{X}_1-\overline{X}_2\right)$ سوف يكون له توزيع قريب من التوزيع الطبيعي بمتوسط حسابي $\mu_{\left(\overline{X}_1-\overline{X}_2\right)}=\mu_1-\mu_2$ وتباين $\sigma^2_{\left(\overline{X}_1-\overline{X}_2\right)}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\simeq\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}\simeq\frac{\hat{S}_1^2}{n_1}+\frac{\hat{S}_2^2}{n_2}$ وبالتالي فان المتغير العشوائي (Z) يعطى كالتالي: $Z=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}}\sim N(0,1)$ أو بشكل مكافئ $Z=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\hat{S}_1^2}{n_1}+\frac{\hat{S}_2^2}{n_2}}}\sim N(0,1)$

ويعطى مجال الثقة بالعلاقة التالية

$$P\left(\left(\overline{X}_1-\overline{X}_2\right)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}\le\left(\mu_1-\mu_2\right)\le\left(\overline{X}_1-\overline{X}_2\right)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}\right)=1-\alpha$$

تمرين12:أجري امتحان في مادة الإحصاء لمجموعتين مستقلتين من الطلبة، الأولى تشمل 75طالبا، والثانية تشمل50 طالبا وكانت نتائج هذا الامتحان للمجموعتين كما يلي:

المجموعة الأولى: $\overline{X}_1 = 80, S_1^2 = 49$ ، المجموعة الثانية: $\overline{X}_2 = 70, S_2^2 = 36$

المطلوب:

1-أوجد فترة الثقة ل $(\mu_1 - \mu_2)$ عند درجة ثقة 95%؟

الحل:

لدينا من معطيات التمرين

المجموعة الأولى: $\overline{X}_1 = 80, S_1^2 = 49, n_1 = 75$ ، المجموعة الثانية: $\overline{X}_2 = 70, S_2^2 = 36, n_2 = 50$

بتطبيق العلاقة الأساسية في التقدير نجد

$$P\left(\left(\overline{X}_1 - \overline{X}_2\right) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}} \le (\mu_1 - \mu_2) \le \left(\overline{X}_1 - \overline{X}_2\right) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}\right) = 1 - \alpha$$

نستخرج أولا مقدار الدرجة المعيارية كالتالي

لدينا $1 - \alpha = 0.95 \Leftarrow \alpha = 0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}} = Z_{0,975} = 1,960$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L = \left(\overline{X}_1 - \overline{X}_2\right) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}} \Rightarrow L = (80 - 70) - (1,960)\sqrt{\frac{49}{75 - 1} + \frac{36}{50 - 1}} = 07,68$$

$$U = \left(\overline{X}_1 - \overline{X}_2\right) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}} \Rightarrow U = (80 - 70) + (1,960)\sqrt{\frac{49}{75 - 1} + \frac{36}{50 - 1}} = 12.32$$

إذن فترة التقدير هي $P\left(07.68 \le (\mu_1 - \mu_2) \le 12.32\right) = 0,95$

3-6 فترة الثقة حول الفرق بين متوسطي مجتمعين مجهولي التباين وحجم العينتين صغير

عندما يكون لدينا مجتمعين إحصائيين غير معلومي التباين ولكن توزيعهما التكراري من النوع الطبيعي أو المعتدل وعندما يكون حجم كل من العينتين المأخوذتين من هذين المجتمعين صغيرا ($n_1<30$،$n_2<30$) فإننا نستخدم عادة متباينة الثقة الخاصة بتوزيع ستودنت (T) بدلا من توزيع (Z) ونميز حالتين هما

الحالة الأولى:عندما يكون التباينان مجهولين ومتساويين $\left(\sigma_1^2 = \sigma_2^2 = \sigma^2\right)$

نظرية06: إذا كانت $\left(X_1, X_2, X_3, \ldots\ldots, X_{n_1}\right)$عينة عشوائية من توزيع طبيعي بمتوسط حسابي μ_1 مجهول وتباين σ_1^2 مجهول،وكانت $\left(X_1, X_2, X_3, \ldots\ldots, X_{n_2}\right)$عينة عشوائية من توزيع طبيعي بمتوسط حسابي μ_2 مجهول وتباين σ_2^2وكانت $\left(\sigma_1^2 = \sigma_2^2 = \sigma^2\right)$ والعينتين مستقلتين عن بعضهما البعض، فان المقدر بقيمة واحدة للمعلمة σ^2 (التباين المشترك) يطلق عليه تسمية المقدر المشترك ويرمز له بالرمز

S_p^2 حيث : $S_p^2=\frac{(n_1-1)\hat{S}_1^2+(n_2-1)\hat{S}_2^2}{(n_1-1)+(n_2-1)}=\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}$

وقد اثبت رياضيا بأن التوزيع الاحتمالي للإحصاءة هو $T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

يتبع توزيع ستودنت بدرجة حرية n_1+n_2-02 وعليه:

$$P\left(-t_{1-\frac{\alpha}{2},n_1+n_2-2}\le T\le t_{1-\frac{\alpha}{2},n_1+n_2-2}\right)=1-\alpha$$

وبالتعويض عن قيمة T نجد: $P\left(-t_{1-\frac{\alpha}{2},n_1+n_2-2}\le T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\le t_{1-\frac{\alpha}{2},n_1+n_2-2}\right)=1-\alpha$

والتي يمكن إعادة كتابتها بعد إجراء مختلف العمليات الجبرية كالتالي:

$$P\left((\overline{X}_1-\overline{X}_2)-t_{1-\frac{\alpha}{2},n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le(\mu_1-\mu_2)\le(\overline{X}_1-\overline{X}_2)+t_{1-\frac{\alpha}{2},n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

تمرين13: لمعرفة الفارق الحقيقي بين معدل الإجازات السنوية للعاملين، ومعدل الإجازات السنوية للعاملات في إحدى الشركات الضخمة، تم سحب عينتين عشوائيتين مستقلتين من سجلات العاملين والعاملات في هذه الشركة لعام (2010) فوجد من سجلات 12 عاملا أن متوسط أيام الإجازات السنوية هو 81 يوما بانحراف معياري 05أيام، ومن سجلات 10 عاملات وجد متوسط أيام الإجازات السنوية يساوي 85 يوما بانحراف معياري 04 أيام، وبفرض أن التباينين في المجتمعين المدروسين متساويان.

المطلوب:

1-أوجد قيمة الفارق الحقيقي باحتمال 90% بين معدل الإجازات السنوية للعمال والعاملات في هذه الشركة، مع فرضية أن المجتمعين يتوزعان بشكل طبيعي أو قريب من الطبيعي).

الحل:

لدينا من معطيات التمرين

المجتمع الأول الإجازات السنوية للعاملين: $\overline{X}_1=81, S_1=05, n_1=12$

المجتمع الثاني الإجازات السنوية للعاملات: $\overline{X}_2=85, S_2=04, n_2=10$

نستخرج أولا مقدار درجة الحرية ومقدار T كالتالي

درجة الحرية تساوي $n_1+n_2-2\Rightarrow 12+10-2=20$

ومن جدول ستودنت من جهة اليسار (القيم السفلى) وعند درجة حرية20 ومستو المعنوية 0.01

نجد:

لدينا $1-\alpha = 0.90 \Rightarrow \alpha = 0.10$ ودرجة الحرية (20) وبالتالي فان قيمة ستودنت تساوي

$$T_{1-\frac{\alpha}{2},n_1+n_2-2} \Rightarrow T_{1-\frac{0.10}{2},20} \Rightarrow T_{0.95,20} = 01.725$$

نحسب ألان التباين المشترك كالتالي:

$$S_p^2 = \frac{n_1S_1^2 + n_2S_2^2}{n_1+n_2-2} \Rightarrow S_p^2 = \frac{12\times 25 + 10\times 16}{12+10-2} = 23$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L = \left(\overline{X}_1 - \overline{X}_2\right) - t_{1-\frac{\alpha}{2},n_1+n_2-2} S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \Rightarrow L = |81-85| - (1.725)\left(\sqrt{23}\right)\left(\sqrt{\frac{1}{12}+\frac{1}{10}}\right) = 0,45$$

$$U = \left(\overline{X}_1 - \overline{X}_2\right) + t_{1-\frac{\alpha}{2},n_1+n_2-2} S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \Rightarrow U = |81-85| + (1.725)\left(\sqrt{23}\right)\left(\sqrt{\frac{1}{12}+\frac{1}{10}}\right) = 07,54$$

إذن فترة التقدير هي $P\left(0,45 \le (\mu_1-\mu_2) \le 07,54\right) = 0,90$

الحالة الثانية:عندما يكون التباينان مجهولين وغير متساويين $\left(\sigma_1^2 \neq \sigma_2^2\right)$

نظرية07: إذا كانت $\left(X_1,X_2,X_3,.......,X_{n_1}\right)$عينة عشوائية من توزيع طبيعي بمتوسط حسابي μ_1 مجهول وتباين σ_1^2 مجهول، وكانت $\left(X_1,X_2,X_3,.......,X_{n_2}\right)$عينة عشوائية من توزيع طبيعي بمتوسط حسابي μ_2 مجهول وتباين σ_2^2وكانت $\left(\sigma_1^2 \neq \sigma_2^2\right)$ والعينتين مستقلتين عن بعضهما البعض، فإن أفضل مقدر لتباين المجتمع الأول σ_1^2 هو تباين العينة المسحوبة منه أي S_1^2، وان أفضل مقدر لتباين المجتمع الثاني σ_2^2 هو تباين العينة المسحوبة منه أي S_2^2 ، ومن ثم استخدام توزيع ستودنت وبدرجة حرية معينة (V)، وعليه وفقا لذلك فان **%**$100(1-\alpha)$ فترة ثقة لتقدير الفرق مابين متوسطي المجتمعين $(\mu_1-\mu_2)$ عندما يكون تباين المجتمعين مجهولا ولكنهما غير متساويين تكون كما يلي:

$$P\left(\left(\overline{X}_1-\overline{X}_2\right) - t_{1-\frac{\alpha}{2},V}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \le (\mu_1-\mu_2) \le \left(\overline{X}_1-\overline{X}_2\right) + t_{1-\frac{\alpha}{2},V}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right) = 1-\alpha$$

وقد تم اقتراح العديد من الطرق في كيفية اختيار درجة الحرية من بينها الطريقة التالية، والتي هي الأكثر قبولا

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

تمرين14:تم قياس مستوى السكر في الدم لمجموعتين مريضتين بداء السكري، المجموعة الأولى عددها10 مرضى، يعانون من السكري المعتمد على الأنسولين، والمجموعة الثانية عددها20 مريضا يعانون من السكري غير المعتمد على الأنسولين، وجد من المجموعة الأولى أن متوسط السكر لديهم يساوي 310 ملغ بانحراف معياري165ملغ، ومن المجموعة الثانية وجد أن متوسط السكر لديهم

235 ملغ بانحراف100ملغ، وبفرض أن التجانس بين المجتمعين غير متساو وان المجتمعين يخضعان إلى التوزيع الطبيعي.

المطلوب:

أحسب فترة الثقة للفارق الحقيقي في متوسط السكر للنوعين (المعتمد على الأنسولين وغير المعتمد على الأنسولين) وذلك عند مستوى ثقة 99%؟

الحل:

لدينا من المعطيات

المجموعة الأولى: $n_1 = 10, \overline{X}_1 = 310, S_1 = 165$

المجموعة الثانية: $n_2 = 20, \overline{X}_2 = 235, S_2 = 100$

بما أن المجتمعان يخضعان إلى التوزيع الطبيعي وحجم العينتين اقل من 30 فان فترة الثقة للمجتمعين سوف يخضع لتوزيع ستودنت بدرجة حرية v

نستخرج قيمة درجة الحرية V ومن ثم نحسب قيمة الدرجة المعيارية لتوزيع ستودنت.

درجة الحرية

$$V = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} \Rightarrow V = \frac{\left(\frac{(165)^2}{10} + \frac{(100)^2}{20}\right)^2}{\frac{\left(\frac{(165)^2}{10}\right)^2}{10 - 1} + \frac{\left(\frac{(100)^2}{20}\right)^2}{20 - 1}} = 12{,}41 \simeq 12$$

لدينا

لدينا $1 - \alpha = 0.99 \Leftarrow \alpha = 0.01$ ودرجة الحرية (12) وبالتالي فان قيمة ستودنت تساوي

$$T_{1-\frac{\alpha}{2},V} \Rightarrow T_{1-\frac{0.01}{2},12} \Rightarrow T_{0.995,12} = 03.055$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L = \left(\overline{X}_1 - \overline{X}_2\right) - t_{1-\frac{\alpha}{2},V}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \Rightarrow L = (310 - 235) - (03.055)\sqrt{\frac{(165)^2}{10} + \frac{(100)^2}{20}} = -98.42$$

$$U = \left(\overline{X}_1 - \overline{X}_2\right) + t_{1-\frac{\alpha}{2},V}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \Rightarrow U = (310 - 235) + (03.055)\sqrt{\frac{(165)^2}{10} + \frac{(100)^2}{20}} = 248.42$$

إذن فترة التقدير هي

$$P\left(-98.42 \le (\mu_1 - \mu_2) \le 248.42\right) = 0{,}99$$

07-فترة الثقة حول الفرق بين متوسطي مجتمعين غير مستقلين $(\mu_1 - \mu_2)$

تمهيد: إن العينات غير المستقلة (المفردات المزدوجة) تكون في حالة عدم تغيير مفردات العينة، وذلك بعد إعطاء المفردات معالجة من شكل جديد، من شان هذه المعالجة إما أن تغير أو تقلل أو تزيد من القياسات التي تؤخذ لهذه المفردات.

فطريقة إدارية جديدة يمكن أن تغير من إنتاجية الشركات، وطريقة تربوية جديدة قد تؤدي إلى رفع مستوى الطلاب، ودواء جديد من المحتمل أن يقلل من ضغط الدم.....الخ.

وبالتالي سوف يكون هناك عينتين الأولى تمثل النتائج قبل التجربة والثانية تمثل النتائج بعد التجربة وفي هذه الحالة تكون العينتين غير مستقلتين، إذن يمكن القول بأنه إذا ظهرت مفردات العينتين في هيئة أزواج بحيث إن كل مفردتين لهما علاقة ببعضهما البعض (أي أن القياسات أو القراءات على نفس الوحدة التجريبية) فان العينتين غير مستقلتين، وإذا كانت العينتين غير مستقلتين فيجب أن يكون عددهما متساوي.

نظرية08: لتكن لدينا $\{(X_1,Y_1),(X_2,Y_2),..............(X_n,Y_n)\}$ ترمز لعينة عشوائية من المفردات المزدوجة، أي أن (X_i,Y_i) ترمز لقياسين تم أخذهما على نفس وحدة المعاينة حيث المجتمع (X) له توزيع طبيعي بمتوسط حسابي μ_1 وتباين σ_1^2 ، والمجتمع (Y) له توزيع طبيعي بمتوسط حسابي μ_2 وتباين σ_2^2 فإذا رمزنا للفرق مابين كل زوج من المفردات بالرمز $d_i=X_i-Y_i$ و i=1،2،3،......،n فإن d_i لها توزيع طبيعي بمتوسط يساوي μ_d وتباين σ_d^2، حيث:

$$\mu_d=E(d)=(X_1-Y_1)=E(X_1)-E(Y_1)=\mu_1-\mu_2$$

وأفضل مقدر للوسط الحسابي لمجتمع الفروق $(\mu_d=\mu_1-\mu_2)$ هو الوسط الحسابي لعينة الفروق $\bar{d}$ ، وبما أن تباين مجتمع الفروق مجول،إذن سنقدره بتباين عينة الفروق S_d^2 ، ونجد المتغير العشوائي التالي:

$\dfrac{\bar{d}-\mu_d}{\sqrt{\dfrac{S_d^2}{n}}}$ يتوزع توزيع ستودنت (T) بدرجة حرية (V=n−1).

وتعطى فترة الثقة للفرق بين متوسطي مجتمعين في حالة عينتين مزدوجتين عند مستوى ثقة $(1-\alpha)100\%$ كالتالي:

$$P\left(\bar{d}-t_{1-\frac{\alpha}{2},V}\sqrt{\frac{S_d^2}{n-1}}\leq(\mu_1-\mu_2)\leq\bar{d}+t_{1-\frac{\alpha}{2},V}\sqrt{\frac{S_d^2}{n-1}}\right)=1-\alpha$$

أو بشكل مكافئ

$$P\left(\bar{d}-t_{1-\frac{\alpha}{2},V}\sqrt{\frac{\hat{S}_d^2}{n}}\leq(\mu_1-\mu_2)\leq\bar{d}+t_{1-\frac{\alpha}{2},V}\sqrt{\frac{\hat{S}_d^2}{n}}\right)=1-\alpha$$

بحيث يتم حساب كل من $\bar{d}$ و S_d^2 و $\hat{S}_d^2$ كما يلي:

$$\bar{d}=\frac{\sum_{i=1}^{n}d_i}{n}=\frac{\sum_{i=1}^{n}(X_i-Y_i)}{n},S_d^2=\frac{\sum_{i=1}^{n}(d_i-\bar{d})^2}{n},\hat{S}_d^2=\frac{\sum_{i=1}^{n}(d_i-\bar{d})^2}{n-1}$$

تمرين15: هناك إدعاء من إحدى الشركات المنتجة للأدوية أن هناك نوع جديد من العقاقير التي يمكن استخدامها لتخفيف الوزن بمتوسط قدره 04.50 كغ، خلال شهر من تناولها، فإذا تناول هذه العقاقير سبعة أشخاص وسجلت أوزانهم قبل بداية البرنامج وبعد شهر من تناولها فكانت النتائج كما يلي:

الأشخاص	01	02	03	04	05	06	07
الوزن قبل تناول العقاقير (X)	58.50	60.30	61.70	69.00	64.00	62.60	56.70
الوزن بعد تناول العقاقير (Y)	60.00	54.90	58.10	62.10	58.50	59.90	54.40

المطلوب:

أوجد فترة الثقة لمتوسط الفرق في متوسط الوزن(قبل وبعد تناول العقار) وذلك عند مستوى ثقة 95%، بافتراض أن المجتمعين يتوزعان توزيعا طبيعيا؟

الحل:

نجد أولا عينة الفروق ثم نحسب الوسط الحسابي $\bar{d}$ و S_d^2 و $\hat{S}_d^2$ كما يلي:

الأشخاص	X_i	Y_i	$d_i = X_i - Y_i$	$\left(d_i - \bar{d}\right)^2$
01	58.50	60.00	−01.50	25.573
02	60.30	54.90	+05.40	03.397
03	61.70	58.10	+03.60	0.002
04	69.00	62.10	+06.90	11.176
05	64.00	58.50	+05.50	03.775
06	62.60	59.90	+02.70	0.734
07	56.70	54.40	+02.30	01.580
المجموع	**432.80**	**407.90**	**24.90**	**46.237**

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} \Rightarrow \bar{d} = \frac{24.90}{07} = 03.557$$

$$S_d^2 = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n} \Rightarrow S_d^2 = \frac{46.237}{07} = 06.605$$

$$\hat{S}_d^2 = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1} \Rightarrow \hat{S}_d^2 = \frac{46.237}{06} = 07.706$$

وحيث أن $1-\alpha = 0.95 \Leftarrow \alpha = 0.05$ ودرجة الحرية (06) وبالتالي فان قيمة ستودنت تساوي

$$T_{1-\frac{\alpha}{2},V} \Rightarrow T_{1-\frac{0.05}{2},06} \Rightarrow T_{0.975,06} = 02.447$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=\bar{d}-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_d^2}{n-1}} \Rightarrow L=03.557-(02.447)\sqrt{\frac{06.605}{07-01}}=0.99$$

$$U=\bar{d}+t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_d^2}{n-1}} \Rightarrow U=03.557-(02.447)\sqrt{\frac{06.605}{07-01}}=06.124$$

أو بشكل مكافئ باستخدام تباين العينة المعدل

$$L=\bar{d}-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{\hat{S}_d^2}{n}} \Rightarrow L=03.557-(02.447)\sqrt{\frac{07.706}{07}}=0.99$$

$$U=\bar{d}+t_{1-\frac{\alpha}{2},v}\sqrt{\frac{\hat{S}_d^2}{n}} \Rightarrow U=03.557+(02.447)\sqrt{\frac{07.706}{07}}=06.124$$

إذن فترة التقدير هي $P\left(0.99\le(\mu_1-\mu_2)\le 06.124\right)=0,95$

08-فترة الثقة لنسبة المجتمع P

إن التقدير بقيمة واحدة لنسبة مفردات المجتمع الإحصائي التي تحمل الصفة مدار البحث يرمز له بالرمز $\hat{P}$ حيث $\hat{P}=\frac{x}{n}$ ويطلق عليها تسمية نسبة العينة،أما نسبة مفردات العينة التي لا تحمل الصفة مدار البحث يرمز لها بالرمز $\hat{q}=1-\hat{P}$، أن التقدير بقيمة واحدة عرضة للأخطاء وبالتالي يفضل استخدام التقدير ضمن فترة، وكما نعلم أن توزيع المعاينة لنسبة العينة $\hat{P}$ عندما يكون حجم العينة كبيرا وتكون قيمة p ليست بقريبة من الصفر أو الواحد يتبع التوزيع الطبيعي بمتوسط حسابي $\mu_{\hat{p}}=p$ وتباين قدره $\sigma_{\hat{p}}^2=\frac{p(1-p)}{n}=\frac{p\times q}{n}$ وعليه فان $Z=\frac{\hat{P}-P}{\sqrt{\frac{P(1-P)}{n}}}\sim N(0,1)$

وبما أن نسبة المجتمع P مجهولة، وهي التي نرغب في تقديرها بإيجاد فترة الثقة، فلا نستطيع حساب تباين المعاينة لنسبة العينة ($\sigma_{\hat{p}}^2=\frac{p(1-p)}{n}=\frac{p\times q}{n}$) ولكننا سنقدره باستخدام أفضل مقدر بالقيمة لنسبة المجتمع المجهولة P وهو نسبة العينة $\hat{P}$ وعند استخدام مقدر التباين وهو: $\sigma_{\hat{P}}^2=\frac{\hat{P}\left(1-\hat{P}\right)}{n}$ حيث $\hat{q}=1-\hat{P}$

وبالتعويض بقيمة التباين الجديد في صيغة المتغير العشوائي (Z) السابقة وبالتالي سيتوزع المتغير الجديد توزيعا قريبا من التوزيع الطبيعي المعياري، عندما يكون حجم العينة n كبيرا إذن:

$$P\left(-Z_{1-\frac{\alpha}{2}}\le Z\le Z_{1-\frac{\alpha}{2}}\right)=1-\alpha \Rightarrow P\left(-Z_{1-\frac{\alpha}{2}}\le\frac{\hat{P}-P}{\sqrt{\frac{\hat{P}\hat{q}}{n}}}\le Z_{1-\frac{\alpha}{2}}\right)=1-\alpha$$

وبإجراء العمليات الجبرية على الحدود الثلاثة للمتباينات سنتحصل على فترة التقدير التالية:

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

نظرية09: إذا كانت $\hat{P}=\frac{x}{n}$ نسبة النجاح في عينة عشوائية حجمها n، وكان n كبيرا فان فترة الثقة التقريبية $(1-\alpha)100\%$ لنسبة النجاح P هي

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

أما إذا كان المجتمع محدود والسحب مع عدم الإرجاع مع تحقق الشرط $\frac{n}{N}>0.05$ فإننا نستخدم معامل الإرجاع في العلاقة السابقة ويصبح مجال التقدير كالتالي:

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

في الحقيقة أن فترة الثقة أعلاه فترة تقريبية وذلك لان التقريب قد يستخدم في مرحلتين أو ثلاثة مراحل مختلفة، فمن الممكن استخدام توزيع ذي الحدين لتقريب التوزيع فوق الهندسي، ثم استخدام التوزيع الطبيعي لتقريب توزيع ذي الحدين واستخدام النسبة $\hat{P}$ لتقريب النسبة الحقيقية ل P، ومع ذلك فان فترة ثقة هذه ستكون جيدة طالما حجم العينة n كبير بشكل كاف.

تمرين16: تريد أحد الشركات القيام بتسويق نوع جديد من مسحوق الغسيل، وقبل القيام بذلك أرادت الشركة القيام بأبحاث للسوق لمعرفة مدى تفضيل الناس لهذا المسحوق، فسحبت عينة عشوائية من 200 مستهلك و أهدت لهم عبوة مجانية، وبعد استعمالها وجدت الشركة أن 140 منهم فضلوا هذا المسحوق.

المطلوب:

1-تقدير نقطة لنسبة المستهلكين الذين يفضلون هذا المسحوق؟

2-تقدير فترة ثقة 95% لنسبة المستهلكين الذين يفضلون هذا المسحوق؟

الحل:

لدينا من المعطيات $n=200, X=140$

1-تقدير نقطة لنسبة المستهلكين الذين يفضلون هذا المسحوق.

$\hat{P}=\frac{x}{n}\Rightarrow \hat{P}=\frac{140}{200}=0.70$ إذن $\hat{P}$ هي مقدر لنسبة المجتمع P.

2- تقدير فترة ثقة 95% لنسبة المستهلكين الذين يفضلون هذا المسحوق

وبتطبيق نظرية النهاية المركزية نجد

$$\begin{cases} n\hat{P} \geq 05 \\ n\hat{q} \geq 05 \end{cases} \Rightarrow \begin{cases} 200 \times 0.70 = 140 \geq 05 \\ 200 \times 0.30 = 60 \geq 05 \end{cases}$$

إذن الشرطان محققان وبالتالي النسبة سوف تخضع للتوزيع الطبيعي المعياري

لدينا $1-\alpha = 0.95 \Leftarrow \alpha = 0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}} = Z_{0,975} = 1,960$

وبتطبيق علاقة التقدير للنسبة $P\left(\hat{P} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}} \leq P \leq \hat{P} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}}\right) = 1-\alpha$

نستخرج الحد الأعلى والأدنى كما يلي:

$$L = \hat{P} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}} \Rightarrow L = 0.70 - (01.960)\sqrt{\frac{0.70 \times 0.30}{200}} = 0.64$$

$$U = \hat{P} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}} \Rightarrow U = 0.70 + (01.960)\sqrt{\frac{0.70 \times 0.30}{200}} = 0.76$$

إذن فترة التقدير هي $P(0.64 \leq P \leq 0.76) = 0.95$

8-1 تحديد حجم العينة للنوعيات المناسب لتقدير النسبة في المجتمع: إذا كانت P تقع في مركز الفترة فان $\hat{P}$ تقدر P وبدون خطأ ولكن في اغلب الأحيان سوف لن تكون بالضبط $\hat{P}$ مساوية للنسبة P وبالتالي هناك خطأ في التقدير، وهذا الخطأ يرمز له بالرمز $E = \left|\hat{P} - P\right|$، ونحن على ثقة قدرها تقريبا $(1-\alpha)100\%$ إن هذا الخطأ سوف لن يكون أكبر من $Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\hat{q}}{n}}$ ، وبالتالي في الحالات التي يمكن فيها اختيار حجم العينة n فانه يمكن اختياره بحيث إذا كانت $\hat{P}$ سوف تستخدم كتقدير للنسبة P، نحن على ثقة $(1-\alpha)100\%$ بان هذا الخطأ سوف لن يكون أكبر من E، وان حجم العينة المناسب إذا كان السحب مع الإرجاع هو

$$n = \frac{Z^2_{1-\frac{\alpha}{2}}(1-P)(P)}{E^2}$$

أما إذا كان السحب من مجتمع محدود مع عدم الإرجاع فان حجم العينة يعطى كالتالي:

$$n = \frac{N(1-P)(P)Z^2_{1-\frac{\alpha}{2}}}{(N-1)E^2 + (1-P)(P)Z^2_{1-\frac{\alpha}{2}}}$$

-وبما أن P غير معلومة ولا نستطيع تقديرها بالقيمة $\hat{P}$ لأننا لم نأخذ العينة بعد، ولذلك نضع بدلا من (P)(1−P) أكبر قيمة لها وهي 4/1 لان الاقتران (P)(1−P) على الفترة (1,0) له قيمة عظمى عند p=1/2 والقيمة العظمى تكون 4/1 إذن حجم العينة هو:

$$n=\frac{Z^2_{1-\frac{\alpha}{2}}}{E^2}\times\frac{1}{4}$$

-أما إذا كان السحب من مجتمع محدود مع عدم الإرجاع فان حجم العينة يعطى كالتالي:

$$n=\frac{NZ^2_{1-\frac{\alpha}{2}}}{04(N-1)E^2+Z^2_{1-\frac{\alpha}{2}}}$$

-وفي بعض الأحيان تتوافر لدينا معلومات عن قيمة تحقق الظاهرة المدروسة في المجتمع (P)، إلا أنها تتراوح بين قيمتين كبرى وصغرى، فإننا نأخذ بالقيمة الأقرب إلى 50%، وهنا نميز حالتين هما:

- $(05\leq P\leq 10)$ فتكون قيمة P=10
- $(60\leq P\leq 80)$ فتكون قيمة P=60

تمرين17: ترغب إحدى الشركات السياحية في معرفة نسبة الطلبة بإحدى الكليات الذين يسافرون إلى لندن في فترة العطلة الصيفية وحيث أن مقابلة وسؤال جميع طلبة الكلية البالغ عددهم 2000 طالب يعتبر أمرا بالغ الصعوبة فقد قررت الشركة اختيار عينة عشوائية لتقدير النسبة المطلوبة وذلك في حدود خطأ يساوي 0.04

المطلوب:

تحديد حجم العينة بدرجة ثقة 95%؟

الحل:

لدينا من معطيات التمرين: $N=2000, E=0.04, 1-\alpha=0.95$

حيث لا توجد لدى الشركة أية معلومات أو تقديرات عن نسبة الطلبة الذين يسافرون إلى لندن خلال العطلة الصيفية فإن:

لدينا $1-\alpha=0.95 \Leftarrow \alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

$$n=\frac{NZ^2_{1-\frac{\alpha}{2}}}{04(N-1)E^2+Z^2_{1-\frac{\alpha}{2}}}\Rightarrow n=\frac{(2000)(01.96)^2}{04(2000-01)(0.04)^2+(01.96)^2}\simeq 462$$

تمرين18: في أحد الأحياء من مدينة ما تقطن 4000 أسرة، وكان المطلوب تقدير نسبة الأسر المستأجرة بخطأ مرتكب لا يزيد عن 05%، وكانت النسبة الحقيقية للمستأجرين مساوية ل 45%.

المطلوب:

1-تحديد حجم العينة الواجب سحبها مع الإرجاع لتقدير النسبة الحقيقية للمستأجرين بدرجة ثقة 95%؟

2- تحديد حجم العينة الواجب سحبها مع عدم الإرجاع لتقدير النسبة الحقيقية للمستأجرين بدرجة ثقة 95%؟

الحل:

من معطيات التمرين لدينا: $N=4000, E=0.05, P=0.45, 1-\alpha=0.95$

1-تحديد حجم العينة في حالة السحب مع الإرجاع عند درجة ثقة 95%

لدينا $1-\alpha=0.95 \Leftarrow \alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

إذن

$$n=\frac{Z^2_{1-\frac{\alpha}{2}}(1-P)(P)}{E^2} \Rightarrow n=\frac{(01.96)^2(0.45)(01-0.45)}{(0.05)^2}=380$$

2- تحديد حجم العينة في حالة السحب مع عدم الإرجاع عند درجة ثقة 95%

$$n=\frac{N(1-P)(P)Z^2_{1-\frac{\alpha}{2}}}{(N-1)E^2+(1-P)(P)Z^2_{1-\frac{\alpha}{2}}} \Rightarrow n=\frac{(4000)(0.45)(0.55)(01.96)^2}{(4000-1)(0.05)^2+(0.45)(0.55)(01.96)^2}=347$$

تمرين19: لقد كان من المرغوب في تحديد نسبة العمال حملة الإجازة الجامعية في الصناعات الكيميائية، وقد قدرت هذه النسبة بين 02% و22%.

المطلوب:

1-ما هو حجم العينة الواجب سحبها عشوائيا بغية تقدير النسبة الحقيقية لحملة الإجازة الجامعية في هذه الصناعة على أن لا يزيد الخطأ المرتكب في التقدير عن05%، وان لا يقل احتمال الدقة عن 95%؟

الحل:

لدينا من معطيات التمرين: $(0.12 \le P \le 0.22), E=0.05, 1-\alpha=0.95$

وبما أن نسبة الاحتمال محصورة بين قيمتين نأخذ القيمة الأقرب إلى 50% وبالتالي فقيمة P=0.22

لدينا $1-\alpha=0.95 \Leftarrow \alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

إذن حجم العينة هو $n=\frac{Z^2_{1-\frac{\alpha}{2}}(1-P)(P)}{E^2} \Rightarrow n=\frac{(01.96)^2(0.22)(01-0.22)}{(0.05)^2}=264$

تمرين20: رغب أحد الباحثين في تقدير نسبة الأسر التي لا يزيد عدد أفرادها عن ثلاثة أشخاص، وكان الاعتقاد السائد بان نسبة هذه الأسر تترواح بين 60% و70%.

المطلوب:

1-تحديد حجم العينة المراد سحبها عشوائيا لتقدير النسبة الحقيقية للأسر التي لا يزيد عدد أفرادها عن ثلاثة أشخاص، على أن لا يزيد الخطأ المسموح ارتكابه في التقدير عن 04%، وان لا يقل إحتمال الدقة عن 95%؟

الحل:

من معطيات التمرين لدينا: $(0.60 \le P \le 0.70), E = 0.04, 1-\alpha = 0.95$

وبما أن نسبة الاحتمال محصورة بين قيمتين نأخذ القيمة الأقرب إلى 50% وبالتالي فقيمة P=0.60

لدينا $1-\alpha = 0.95 \Leftarrow \alpha = 0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}} = Z_{0,975} = 1,960$

إذن حجم العينة هو $$n = \frac{Z^2_{1-\frac{\alpha}{2}}(1-P)(P)}{E^2} \Rightarrow n = \frac{(01.96)^2(0.60)(01-0.60)}{(0.04)^2} = 576$$

09- فترة الثقة للفرق بين نسبتي مجتمعين $(P_1 - P_2)$

نظرية10: إذا كانت $(X_1, X_2, X_3, \dots\dots, X_{n_1})$ تمثل عينة عشوائية تخضع لتوزيع ذي الحدين فان التقدير بقيمة واحدة يعطى لهذه العينة $\hat{P}_1 = \frac{X}{n}$، وإذا كانت n_1 كبيرة جدا (نظرية النهاية المركزية) فإن $\hat{P}_1 \sim N\left(P_1, \frac{P_1(1-P_1)}{n_1}\right)$، وبالمثل إذا كانت $(X_1, X_2, X_3, \dots\dots, X_{n2})$ تمثل عينة عشوائية تخضع لتوزيع ذي الحدين فان التقدير بقيمة واحدة يعطى لهذه العينة $\hat{P}_2 = \frac{X}{n_2}$ وإذا كانت n_2 كبيرة جدا (نظرية النهاية المركزية) فان $\hat{P}_2 \sim N\left(P_2, \frac{P_2(1-P_2)}{n_2}\right)$، وإذا كانت العينتين مستقلتين عن بعضهما البعض فان $\hat{P}_1 - \hat{P}_2$ تشكل صيغة خطية لمتغير عشوائي يتبع التوزيع الطبيعي تقريبا بمتوسط حسابي $\mu_{\hat{P}_1-\hat{P}_2} = P_1 - P_2$ وتباين $\sigma^2_{\hat{P}_1-\hat{P}_2} = \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}$ وعليه فان:

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

وبالتالي يمكننا القول بان
$$P\left(-Z_{1-\frac{\alpha}{2}} \le Z \le Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

أي أن
$$P\left(-Z_{1-\frac{\alpha}{2}} \le \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} \le Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

وبما أن نسبة المجتمع الأول P_1 مجهولة، ونسبة المجتمع الثاني P_2 مجهولة، فلا نستطيع حساب الخطأ المعياري لتوزيع المعاينة للفرق بين نسبتي العينتين $\hat{P}_1 - \hat{P}_2$ ولكننا سنقدره باستخدام أفضل مقدر بالقيمة لنسبة المجتمع الأول المجهولة P_1 وهو نسبة العينة المسحوبة منه $\hat{P}_1$، وأفضل مقدر بالقيمة لنسبة المجتمع الثاني المجهولة P_2 وهو نسبة العينة المسحوبة منه $\hat{P}_2$، وعليه فان $(1-\alpha)100\%$ فترة ثقة لتقدير الفرق $P_1 - P_2$ انطلاقا من المعادلة الأخيرة وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة سنتحصل على الصيغة التالية للتقدير:

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

إذا كان المجتمعان محدودان والسحب مع عدم الإرجاع مع تحقق الشرط اللازم (كسر المعاينة) نضيف للعلاقة السابقة معامل الإرجاع أو معامل التصحيح.

تمرين21: سحبت عينتان عشوائيتان مستقلتان عن بعضهما البعض مع الإرجاع، الأولى تحتوي على 120 وحدة منتجة بالآلة (أ) ووجدنا بها 06 وحدات بها عيوب، والثانية تحتوي على 200 وحدة منتجة بالآلة (ب) ووجدنا بها 09 وحدات بها عيوب.

المطلوب:

1-قدر الفرق بين نسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (أ) ونسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (ب) وذلك باستخدام مستوى ثقة 95%؟

الحل:

لدينا من معطيات التمرين: $X_1=06, n_1=120, X_2=09, n_2=200, 1-\alpha=0.95$

1-تقدير الفرق بين نسبتي الوحدات التالفة

نحسب أولا نسبة الوحدات التي بها عيوب في إنتاج الآلتين كما يلي:

-نسبة الوحدات التي بها عيوب لعينة إنتاج الآلة (أ) هي: $\hat{P}_1=\frac{X_1}{n_1}\Rightarrow\hat{P}_1=\frac{06}{120}=0.05$

- نسبة الوحدات التي بها عيوب لعينة إنتاج الآلة (ب) هي: $\hat{P}_2=\frac{X_2}{n_2}=\frac{09}{200}=0.045$

وبما أن $95\%=(1-\alpha)100\%$

إذن $1-\alpha=0.95\Leftarrow\alpha=0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=\left(\hat{P}_1-\hat{P}_2\right)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\left(1-\hat{P}_1\right)}{n_1}+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}}\Rightarrow$$

$$L=(0.05-0.045)-(01.96)\sqrt{\frac{0.05(1-0.05)}{120}+\frac{0.045(1-0.045)}{200}}=-0.0434$$

$$U=\left(\hat{P}_1-\hat{P}_2\right)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\left(1-\hat{P}_1\right)}{n_1}+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}}\Rightarrow$$

$$U=(0.05-0.045)+(01.96)\sqrt{\frac{0.05(1-0.05)}{120}+\frac{0.045(1-0.045)}{200}}=0.0534$$

إذن فترة التقدير هي $P\left(-0.0434 \leq (P_1 - P_2) \leq 0.0534\right) = 0,95$

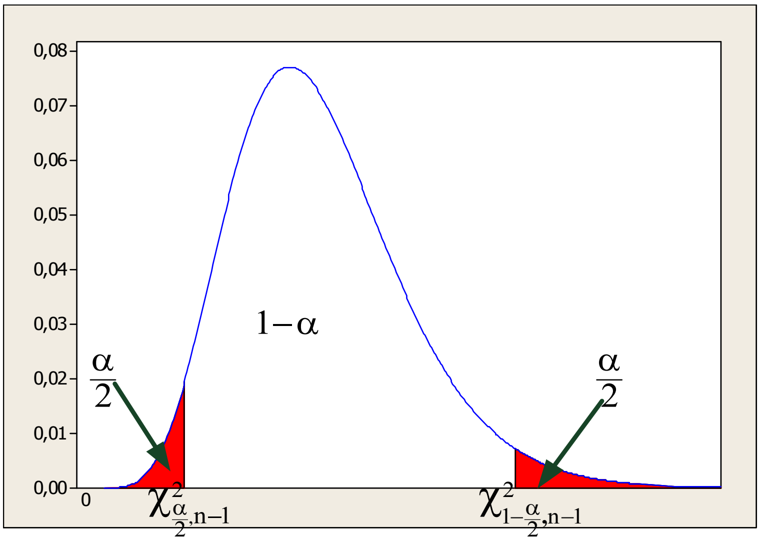
أي بثقة قدرها 95% نستطيع القول بان الفرق بين نسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (أ) ونسبة الوحدات التي بها عيوب في الإنتاج الكلي للآلة (ب) يقع بين 04.34%- و05.34%.

10- فترة الثقة لتباين المجتمع σ^2

تمهيد: في بعض الدراسات الإحصائية نجد أنفسنا بحاجة إلى معرفة تباين المجتمع σ^2 وكثيرا ما يكون هذا التباين مجهولا وغير معروف لنا، لهذا نلجأ إلى بيانات العينة لتقدير تباين المجتمع، فمثلا عند تقدير متوسط المجتمع باستخدام متوسط العينة استطعنا إيجاد فترات ثقة مناسبة ولكن كانت فترات الثقة تعتمد على الانحراف المعياري للمجتمع σ، وأحيانا تكون σ مجهولة لذلك نستخدم في هذه الحالة الانحراف المعياري للعينة S كتقدير للانحراف المعياري للمجتمع σ، أي أننا نستخدم تباين العينة S^2 كتقدير لتباين المجتمع σ^2، وهذا التقدير يسمى التقدير بنقطة، أي أننا من بيانات العينة يمكننا الحصول على قيمة عددية وحيدة لتباين العينة S^2 واعتبرت تقديرا لتباين المجتمع، أي أن التقدير بنقطة هو التقدير بقيمة عددية وحيدة وهو تقدير مقبول وجيد ويستخدم في كثير من الدراسات الإحصائية، ولكننا الآن نرغب في إيجاد فترة ثقة لتباين المجتمع.

نظرية11: إذا كانت لدينا $(X_1, X_2, X_3, ..., X_n)$ تشكل عينة عشوائية من التوزيع الطبيعي بمتوسط حسابي يساوي μ وتباين σ^2 ، فان توزيع المعاينة للإحصاءة $\chi^2 = \frac{(n-1)\hat{S}^2}{\sigma^2} = \frac{nS^2}{\sigma^2}$

يتبع توزيع يطلق عليه توزيع مربع –كاي وبدرجة حرية $n-1$، ومن الجداول الإحصائية نحصل على الجدول الإحصائي الخاص بهذا التوزيع بحيث نتحصل على قيمتين للمتغير χ^2 عند درجة حرية معينة، بحيث تكون المساحة على يمين القيمة الكبرى تساوي المساحة على يسار القيمة الصغرى وتساوي $\frac{\alpha}{2}$ ، وإذا رمزنا لكل قيمة باستعمال المساحة التي على يمينها سنرمز للقيمة الكبرى بالرمز $\chi^2_{\frac{\alpha}{2}}$ وللقيمة الصغرى بالرمز $\chi^2_{1-\frac{\alpha}{2}}$ وحيث أن المساحة الكلية تحت منحنى توزيع χ^2 تساوي الواحد الصحيح،إذن المساحة بين هاتين القيمتين تساوي $1-\alpha$ و الشكل التالي يوضح ذلك:



ومن الشكل السابق يمكننا أن نستنتج مجال الثقة كالتالي:

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

نظرية12: إذا كانت لدينا $(X_1, X_2, X_3, \ldots\ldots\ldots, X_n)$ عينة عشوائية مسحوبة من مجتمع له التوزيع الطبيعي $X \sim N(\mu, \sigma^2)$ وبفرض أننا نريد إيجاد فترة ثقة بمعامل ثقة $(1-\alpha)$ لتباين هذا المجتمع، في هذه الحالة يكون هناك حالتين:

1-إذا كان متوسط المجتمع μ معلوم: إذا كانت μ معلومة فإن $S^2 = \frac{\sum_{i=1}^{n}(X-\mu)^2}{n}$ تمثل تباين العينة على أساس متوسط المجتمع μ، وان الكمية المحورية المطلوبة في هذه الحالة هي $\frac{nS^2}{\sigma^2}$ ، ونعلم أن توزيعها هو توزيع مربع كاي بدرجة حرية n، وهو توزيع مستقل عن σ^2 وبالتالي نحصر هذه الكمية المحورية بين القيمتين المعياريتين $\chi^2_{\frac{\alpha}{2},n}, \chi^2_{1-\frac{\alpha}{2},n}$ ثم نعزل الوسيط المجهول σ^2 ، فنحصل على فترة الثقة $(1-\alpha)100\%$ المطلوبة والتي تعطى بالعلاقة التالية:

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

2-إذا كان متوسط المجتمع μ غير معلوم: في حالة كون متوسط المجتمع μ مجهولا فإن أفضل تقدير لتباين المجتمع هو $\hat{S}^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$ ، في هذه الحالة فان الكمية المحورية هي $\frac{(n-1)\hat{S}^2}{\sigma^2}$ وتوزيعها هو توزيع مربع كاي بدرجة حرية n-1، وهو توزيع مستقل عن σ^2 وبالتالي نحصر هذه الكمية المحورية بين القيمتين المعياريتين $\chi^2_{\frac{\alpha}{2},n-1}, \chi^2_{1-\frac{\alpha}{2},n-1}$ ثم نعزل الوسيط المجهول σ^2 ، فنحصل على فترة الثقة $(1-\alpha)100\%$ المطلوبة والتي تعطى بالعلاقة التالية:

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

تمرين 22: إذا علمت أن تباين عينة عشوائية ذات حجم 25 مسحوبة من مجتمع له التوزيع الطبيعي

$X \sim N(10,\sigma^2)$ وكان $S^2=\frac{\sum_{i=1}^{25}(X_i-10)^2}{25}=09$

المطلوب:

1-تحديد فترة الثقة عند درجة ثقة 95% لتباين هذا المجتمع؟

الحل:

لدينا من معطيات التمرين: $\mu=10, n=25, 1-\alpha=0.95, S^2=09$

بما أن $100(1-\alpha)\%=95\%$ لدينا $1-\alpha=0.95 \Rightarrow \alpha=0.05$

ودرجة الحرية n=25 وباستخدام جدول مربع كاي نستخرج الدرجتين المعياريتين الصغرى والكبرى

كالتالي

$$\begin{cases}\chi^2_{\frac{\alpha}{2},n} \Rightarrow \chi^2_{0.025,25}=13.120\\ \chi^2_{1-\frac{\alpha}{2},n} \Rightarrow \chi^2_{0.975,25}=40.646\end{cases}$$

وبالتعويض في حدي الثقة نستخرج الحد الأعلى والأدنى كما يلي:

$$L=\frac{nS^2}{\chi^2_{1-\frac{\alpha}{2},n}} \Rightarrow L=\frac{25\times 09}{40.646}=05.5356$$

$$U=\frac{nS^2}{\chi^2_{\frac{\alpha}{2},n}} \Rightarrow U=\frac{25\times 09}{13.120}=17.1493$$

إذن مجال التقدير هو $P(05.5356 \leq \sigma^2 \leq 17.1493)=0.95$

تمرين23: أخذت عينة عشوائية حجمها 10 من مجتمع يخضع للتوزيع الطبيعي $X \sim N(\mu,\sigma^2)$ ،

ووجد أن تباينها هو $\hat{S}^2=117.12$

المطلوب:

1- تحديد فترة الثقة عند درجة ثقة 95% لتباين هذا المجتمع؟

الحل:

لدينا من معطيات التمرين: $n=10, 1-\alpha=0.95, \hat{S}^2=117.12$

بما أن $100(1-\alpha)\%=95\%$ لدينا $1-\alpha=0.95 \Rightarrow \alpha=0.05$

ودرجة الحرية n−1=10−1=09 وباستخدام جدول مربع كاي نستخرج الدرجتين المعياريتين الصغرى

والكبرى كالتالي
$$\begin{cases}\chi^2_{\frac{\alpha}{2},n-1} \Rightarrow \chi^2_{0.025,09} = 02.700 \\ \chi^2_{1-\frac{\alpha}{2},n-1} \Rightarrow \chi^2_{0.975,09} = 19.023\end{cases}$$

وبالتعويض في حدي الثقة نستخرج الحد الأعلى والأدنى كما يلي:

$$L = \frac{(n-1)\hat{S}^2}{\chi^2_{1-\frac{\alpha}{2},n}} \Rightarrow L = \frac{09 \times 117.12}{19.023} = 55.411$$

$$U = \frac{(n-1)\hat{S}^2}{\chi^2_{\frac{\alpha}{2},n}} \Rightarrow U = \frac{09 \times 117.12}{02.700} = 390.40$$

إذن مجال التقدير هو
$$P\left(55.411 \le \sigma^2 \le 390.40\right) = 0.95$$

11- فترة الثقة للانحراف المعياري للمجتمع σ

نظرية13: إذا كانت لدينا $(X_1, X_2, X_3, ..., X_n)$ عينة عشوائية مسحوبة من مجتمع له التوزيع الطبيعي $X \sim N(\mu, \sigma^2)$ وبفرض أننا نريد إيجاد فترة ثقة بمعامل ثقة $(1-\alpha)$ للانحراف المعياري لهذا المجتمع، نكون أمام حالتين:

1-إذا كان حجم العينة $(n < 30)$: في هذه الحالة مجال الثقة للانحراف المعياري هو الجذر ألتربيعي للمجال الثقة للتباين، وبأخذ الجذر ألتربيعي على كل حد من حدود المتباينة **(فترة الثقة للتباين)** فإن $100\%(1-\alpha)$ فترة ثقة حول الانحراف المعياري σ للمجتمع الإحصائي مدار البحث ستكون كالتالي:

1-1 إذا كان متوسط المجتمع μ **معلوم**: يعطى مجال الثقة للانحراف المعياري كالتالي

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

1-2 إذا كان متوسط المجتمع μ **غير معلوم**: يعطى مجال الثقة للانحراف المعياري كالتالي

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

2- **إذا كان حجم العينة** $(n \ge 30)$: عندما يكون حجم العينة كبيرا $(n \ge 30)$ فإنه من الممكن اعتبار توزيع العينات في المجتمع الإحصائي توزيعا يقترب كثيرا من التوزيع الطبيعي ذي المتوسط الحسابي المساوي للانحراف المعياري للمجتمع $(\mu = \sigma)$والانحراف المعياري الذي يساوي $\frac{\sigma}{\sqrt{n}}$ وعلى هذا الأساس تحسب الدرجة المعيارية Z والتي تمثل متغيرا عشوائيا بالعلاقة التالية: $Z = \frac{S-\sigma}{\frac{\sigma}{\sqrt{n}}}$

و الذي يقع ضمن المتباينة التالية: $P\left(-Z_{1-\frac{\alpha}{2}} \leq \frac{S-\sigma}{\frac{S}{\sqrt{2n}}} \leq Z_{1-\frac{\alpha}{2}}\right)=1-\alpha$

وبإعادة ترتيب المتراجعة السابقة نحصل على فترة الثقة لتقدير الانحراف المعياري σ

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

ومن أجل $(n \geq 30)$ فإنه يمكننا أن نستخدم علاقة التقريب التالية $\left(\sqrt{2\chi^2}-\sqrt{2v-1}\right)$ وهو قريب جدا من التوزيع الطبيعي بمتوسط حسابي صفر وتباين واحد، وعليه إذا كان Z_p هو المئينة ل $(100)_p$ للتوزيع الطبيعي المعياري فإننا يمكن أن نكتب بدرجة كبيرة من التقريب

$$\sqrt{2\chi_p^2}-\sqrt{2v-1}=Z_p \quad \text{أو} \quad \sqrt{2\chi_p^2}=Z_p+\sqrt{2v-1}$$

ومنه فان $\chi_p^2=\frac{1}{2}\left(Z_p+\sqrt{2v-1}\right)^2$

وتعطى حدود الثقة للانحراف المعياري σ ب $CI(\sigma)_{1-\alpha}=S \pm Z_c\frac{\sigma}{\sqrt{2n}}$

بحيث Z_c تشير إلى مستوى الثقة، وسوف نستعمل الانحراف المعياري للعينة لتقدير σ

تمرين24: يستورد أحد المصانع عدد من الألواح الزجاجية لصناعة العدسات ومن خبرة المصنع السابقة تم تحديد تباين معامل الانكسار لهذا النوع من الألواح الزجاجية بالقيمة $\sigma^2=1.26\times10^{-4}$ ولما كان من الأهمية بمكان أن تكون كافة الألواح الزجاجية من نفس معامل الانكسار فان المصنع يرفض الشحنة المستوردة إذا كان تباين أي عينة عشوائية ذات حجم 20 لوحة زجاجية هو $\hat{S}^2=2.00\times10^{-4}$ ، و لو افترضنا بأن مجتمع الألواح الزجاجية هو مجتمع طبيعي التوزيع.

المطلوب:

1-ما هو احتمال رفض الشحنة المستوردة حتى ولو كان $\sigma^2=1.26\times10^{-4}$؟

2-إذا افترضنا بان تباين معامل الانكسار للألواح الزجاجية في عينة عشوائية ذات حجم 20 هو $\hat{S}^2=1.20\times10^{-4}$، فاحسب فترة التقدير للثقة للانحراف المعياري للمجتمع الإحصائي وبدرجة ثقة 95%؟

الحل:

لدينا من معطيات التمرين : $n=20, \hat{S}^2=1.20\times10^{-4}, \sigma^2=1.26\times10^{-4}$

1-إيجاد احتمال رفض الشحنة المستوردة حتى ولو كان $\sigma^2=1.26\times10^{-4}$

نلاحظ بان الكمية المستوردة هي كمية محدودة أي أن المجتمع الإحصائي من النوع المحدود والمتناهي وان حجم العينة ليس كبيرا لذا نستخدم قاعدة المتغير العشوائي لمربع كاي والمبين حسب العلاقة التالية:

$$\chi^2=\frac{(n-1)\hat{S}^2}{\sigma^2}=\frac{(20-1)(2.00\times10^{-4})}{1.26\times10^{-4}}=30.16$$

وباستخدام جدول مربع –كاي من جهة اليسار وبدرجة حرية $V=n-1 \Rightarrow V=20-1=19$ نجد الاحتمال المقابل للعدد المذكور $\chi^2_{1-\alpha}=30.16$ هو $1-\alpha=0.95$

2–حساب فترة التقدير للثقة للانحراف المعياري للمجتمع الإحصائي وبدرجة ثقة 95%

بما أن $(1-\alpha)100\%=95\%$ لدينا $1-\alpha=0.95 \Leftarrow \alpha=0.05$

ودرجة الحرية $V=n-1 \Rightarrow V=20-1=19$ وباستخدام جدول مربع كاي نستخرج الدرجتين المعياريتين الصغرى والكبرى كالتالي

$$\begin{cases}\chi^2_{\frac{\alpha}{2},n-1} \Rightarrow \chi^2_{0.025,19}=08.907\\ \chi^2_{1-\frac{\alpha}{2},n-1} \Rightarrow \chi^2_{0.975,09}=32.852\end{cases}$$

وبالتعويض في حدي الثقة نستخرج الحد الأعلى والأدنى كما يلي:

$$L=\frac{(n-1)\hat{S}^2}{\chi^2_{1-\frac{\alpha}{2},n}} \Rightarrow L=\frac{(20-1)(1.20\times10^{-4})}{32.852}=0.69\times10^{-4}$$

$$U=\frac{(n-1)\hat{S}^2}{\chi^2_{\frac{\alpha}{2},n}} \Rightarrow U=\frac{(20-1)(1.20\times10^{-4})}{08.907}=02.56\times10^{-4}$$

وعليه يمكن القول بان $100\%(1-\alpha)$ فترة ثقة حول σ^2 ستكون كالتالي

$$0.69\times10^{-4} \le \sigma^2 \le 02.56\times10^{-4}$$

ويلاحظ من متراجحة التباين أن الحد الأعلى هو $U=02.56\times10^{-4}$ ومن البديهي أن نستنتج قبول كل الشحنات لو كان الحد الأعلى لتباين المجتمع الإحصائي هو اقل من شروط القبول السابق كأن يكون مثلا $U=01.99\times10^{-4}$

أما فترة الثقة حول الانحراف المعياري للمجتمع الإحصائي سيكون كالتالي:

$$0.0083 \le \sigma \le 0.016$$

تمرين25: تم حساب الانحراف المعياري لعمر عينة تتألف من200 مصباح كهربائي فوجد بأنه 100 ساعة.

المطلوب:

1–أوجد فترة الثقة للانحراف المعياري لكل المصابيح الكهربائية عند درجة الثقة 95% و99% في الحالتين التاليتين:

أ–باستخدام نظرية المعاينات الكبيرة.

ب–باستخدام نظرية المعاينة الصغيرة أو الصحيحة؟

<u>الحل:</u>

لدينا من معطيات التمرين: $n=200,S=100$

أ–إيجاد فترة الثقة للانحراف المعياري باستخدام نظرية المعاينات الكبيرة عند درجة الثقة 95% و 99%

1–عند درجة ثقة 95%

بتطبيق علاقة حدود الثقة في حالة العينات الكبيرة $P\left(S-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{2n}}\le\sigma\le S+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{2n}}\right)=1-\alpha$

بما أن $(1-\alpha)100\%=95\%$ لدينا $1-\alpha=0.95\Leftarrow\alpha=0.05$

وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=1,960$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=S-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{2n}}\Rightarrow L=100-(01.960)\frac{100}{\sqrt{2\times200}}=90.20$$

$$U=S+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{2n}}\Rightarrow L=100+(01.960)\frac{100}{\sqrt{2\times200}}=109.80$$

وبالتالي فنحن واثقون بدرجة ثقة 95% أن الانحراف المعياري للمجتمع سوف يقع بين 90.20 و 109.80.

2– عند درجة ثقة 99%

بما أن $(1-\alpha)100\%=99\%$ لدينا $1-\alpha=0.99\Leftarrow\alpha=0.01$

وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,01}{2}}=Z_{0,995}=2.575$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$L=S-Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{2n}}\Rightarrow L=100-(02.575)\frac{100}{\sqrt{2\times200}}=87.125$$

$$U=S+Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{2n}}\Rightarrow L=100+(02.575)\frac{100}{\sqrt{2\times200}}=112.875$$

وبالتالي فنحن واثقون بدرجة ثقة 99% أن الانحراف المعياري للمجتمع سوف يقع بين 87.125 و 112.875.

ب–إيجاد فترة الثقة للانحراف المعياري باستخدام نظرية المعاينات الصغيرة عند درجة الثقة95% و 99%

1–عند درجة ثقة 95%

بتطبيق علاقة حدود الثقة في حالة العينات الصغيرة $P\left(\sqrt{\frac{(n-1)\hat{S}^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}}\le\sigma\le\sqrt{\frac{(n-1)\hat{S}^2}{\chi^2_{\frac{\alpha}{2},n-1}}}\right)=1-\alpha$

نحسب أولا تباين العينة المعدل كالتالي

$$\hat{S}=\sqrt{\left(\frac{n}{n-1}\right)\hat{S}^2}\Rightarrow \hat{S}=\sqrt{\left(\frac{200}{200-1}\right)\times(100)^2}=100.25$$

بما أن $100\%(1-\alpha)=95\%$ لدينا $0.01=\alpha \Leftarrow 0.99=\alpha-1$

وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}}=Z_{0,975}=01.960$

وبما أن حجم العينة n=200 نستخدم علاقة التقريب التالية: $\left(\sqrt{2\chi^2}-\sqrt{2v-1}\right)$ فنجد

$$\chi^2_p=\frac{1}{2}\left(Z_p+\sqrt{2v-1}\right)^2\Rightarrow \chi^2_{0.975}=\frac{1}{2}\left(Z_{0.975}+\sqrt{2v-1}\right)^2$$

$$\Rightarrow \chi^2_{0.975}=\frac{1}{2}\left(01.960+\sqrt{2(199)-1}\right)^2=239$$

$$\chi^2_p=\frac{1}{2}\left(Z_p+\sqrt{2v-1}\right)^2\Rightarrow \chi^2_{0.025}=\frac{1}{2}\left(Z_{0.025}+\sqrt{2v-1}\right)^2$$

$$\Rightarrow \chi^2_{0.025}=\frac{1}{2}\left(-01.960+\sqrt{2(199)-1}\right)^2=161$$

نجد ألان حدود الثقة كالتالي

$$L=\sqrt{\frac{(n-1)\hat{S}^2}{\chi^2_{1-\frac{\alpha}{2}}}}\Rightarrow L=\sqrt{\frac{(200-1)(100.25)^2}{239}}=91.47$$

$$U=\sqrt{\frac{(n-1)\hat{S}^2}{\chi^2_{\frac{\alpha}{2},}}}\Rightarrow U=\sqrt{\frac{(200-1)(100.25)^2}{161}}=111.45$$

2- عند درجة ثقة 99%

بما أن $100\%(1-\alpha)=99\%$ لدينا $0.01=\alpha \Leftarrow 0.99=\alpha-1$

وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,01}{2}}=Z_{0,995}=2.575$

وبما أن حجم العينة n=200 نستخدم علاقة التقريب التالية: $\left(\sqrt{2\chi^2}-\sqrt{2v-1}\right)$ فنجد

$$\chi^2_p=\frac{1}{2}\left(Z_p+\sqrt{2v-1}\right)^2\Rightarrow \chi^2_{0.995}=\frac{1}{2}\left(Z_{0.995}+\sqrt{2v-1}\right)^2$$

$$\Rightarrow \chi^2_{0.995}=\frac{1}{2}\left(02.575+\sqrt{2(199)-1}\right)^2=253$$

$$\chi^2_p=\frac{1}{2}\left(Z_p+\sqrt{2v-1}\right)^2\Rightarrow \chi^2_{0.005}=\frac{1}{2}\left(Z_{0.005}+\sqrt{2v-1}\right)^2$$

$$\Rightarrow \chi^2_{0.005}=\frac{1}{2}\left(-02.575+\sqrt{2(199)-1}\right)^2=150$$

نجد ألان حدود الثقة كالتالي

$$L=\sqrt{\frac{(n-1)\hat{S}^2}{\chi^2_{1-\frac{\alpha}{2}}}}\Rightarrow L=\sqrt{\frac{(200-1)(100.25)^2}{253}}=88.91$$

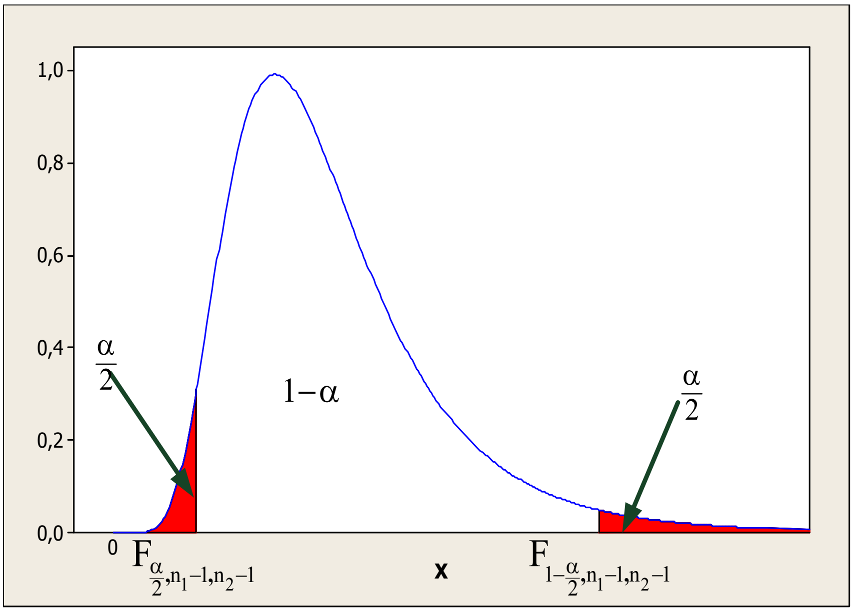
$$U=\sqrt{\frac{(n-1)\hat{S}^2}{\chi^2_{\frac{\alpha}{2},}}}\Rightarrow U=\sqrt{\frac{(200-1)(100.25)^2}{150}}=115.47$$

12- فترة الثقة حول التناسب مابين تباينين $\frac{\sigma_1^2}{\sigma_2^2}$

نظرية14:إذا كانت $(X_1,X_2,X_3,............,X_{n1})$تمثل عينة عشوائية من مجتمع طبيعي بمتوسط حسابي μ_1وتباين σ_1^2 وكلاهما غير معلوم وكانت S_1^2 تمثل تباين هذه العينة، و كانت $(X_1,X_2,X_3,............,X_{n2})$ تمثل عينة عشوائية من مجتمع طبيعي بمتوسط حسابي μ_2 وتباين σ_2^2 حيث كلاهما غير معلوم أيضا، وكانت S_2^2 تمثل تباين هذه العينة وكانت العينتين مستقلتين عن بعضهما البعض فإن توزيع المعاينة للأحصاءة

$$F=\frac{\left[\left(\frac{n_1}{n_1-1}\right)(S_1^2)\right]\frac{1}{\sigma_1^2}}{\left[\left(\frac{n_2}{n_2-1}\right)(S_2^2)\right]\frac{1}{\sigma_2^2}}=\frac{\frac{\hat{S}_1^2}{\sigma_1^2}}{\frac{\hat{S}_2^2}{\sigma_2^2}}\sim F_{n_1-1,n_2-1}$$

وبدرجات حرية V_1 وV_2 ويأخذ الشكل التالي



من الشكل التالي نستطيع الحصول على قيمتين للمتغير F عند درجتي معينتين للحرية V_1 وV_2 حيث V_1 درجة حرية البسط وV_2 درجة حرية المقام، بحيث تكون المساحة على يمين القيمة الكبرى تساوي المساحة على يسار القيمة الصغرى وتساوي $\frac{\alpha}{2}$ ، وإذا رمزنا لكل قيمة باستعمال المساحة على يمينها سنرمز للقيمة الكبرى بالرمز $F_{1-\frac{\alpha}{2}}$ ، وللقيمة الصغرى بالرمز $F_{\frac{\alpha}{2}}$ ، وحيث أن المساحة الكلية تحت

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

$$P\left(F_{\frac{\alpha}{2},V_1,V_2} \leq F \leq F_{1-\frac{\alpha}{2},V_1,V_2}\right)=1-\alpha$$

وبالتعويض عن إحصاءة فيشر F في العلاقة السابقة نجد:

$$P\left(F_{\frac{\alpha}{2},V_1,V_2} \leq \frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2} \leq F_{1-\frac{\alpha}{2},V_1,V_2}\right)=1-\alpha$$

وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة سنتحصل على فترة الثقة التالية:

$$P\left(\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{1-\frac{\alpha}{2},V_1,V_2}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{\hat{S}_1^2/\hat{S}_2^2}{F_{\frac{\alpha}{2},V_1,V_2}}\right)=1-\alpha$$

كما يمكن إعطائه بشكل مكافئ كالتالي:

$$P\left(\frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{1-\frac{\alpha}{2},V_1,V_2}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{\frac{\alpha}{2},V_1,V_2}}\right)=1-\alpha$$

<u>تمرين26</u>: تم امتحان مجموعة من الطلبة في مقياس الإحصاء الرياضي وعددهم 16 ولمجموعة أخرى عددها 10، فكانت نتائجهم بانحراف معياري للمجموعة الأولى والثانية على التوالي هي 05 و02.

<u>المطلوب:</u>

1-أوجد فترة الثقة لنسبة التباين عند درجة ثقة 90%؟

2- أوجد فترة الثقة لنسبة التباين عند درجة ثقة 98%؟

<u>الحل:</u>

لدينا من معطيات التمرين: $n_1=16, S_1=05, n_2=10, S_2=02$

1-إيجاد فترة الثقة لنسبة التباين عند درجة ثقة 95%؟

بتطبيق العلاقة الأساسية في التقدير نجد:

$$P\left(\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{1-\frac{\alpha}{2},V_1,V_2}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{\hat{S}_1^2/\hat{S}_2^2}{F_{\frac{\alpha}{2},V_1,V_2}}\right)=1-\alpha$$

نحسب أولا تباين العينة المعدل للمجموعة الأولى والثانية كالتالي:

$$\hat{S}_1^2=\left(\frac{n_1}{n_1-1}\right)S_1^2\Rightarrow\hat{S}_1^2=\left(\frac{16}{16-1}\right)(05)^2=26.67$$

$$\hat{S}_2^2=\left(\frac{n_2}{n_2-1}\right)S_2^2\Rightarrow\hat{S}_2^2=\left(\frac{10}{10-1}\right)(02)^2=04.44$$

بما أن $(1-\alpha)100\%=90\%$ لدينا $1-\alpha=0.90\Leftarrow\alpha=0.10$ ودرجة الحرية للبسط والمقام هي $V_1=n_1-1\Rightarrow V_1=16-1=15, V_2=n_2-1\Rightarrow V_2=10-1=09$ نجد القيمة الصغرى والكبرى لتوزيع فيشر.

$$F_{1-\frac{\alpha}{2},V_1,V_2}\Rightarrow F_{0.95,15,10}=02.85$$

$$F_{\frac{\alpha}{2},V_1,V_2}=\frac{01}{F_{1-\frac{\alpha}{2},V_2,V_1}}\Rightarrow F_{\frac{\alpha}{2},V_1,V_2}=\frac{01}{F_{0.95,10,15}}=\frac{01}{02.54}=0.3937$$

نجد ألآن الحد الأعلى والحد الأدنى لمجال الثقة:

$$L=\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{1-\frac{\alpha}{2},V_1,V_2}}\Rightarrow L=\frac{26.67/04.44}{02.85}=02.108$$

$$U=\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{\frac{\alpha}{2},V_1,V_2}}\Rightarrow U=\frac{26.67/04.44}{0.3937}=15.257$$

إذن مجال التقدير هو $P\left(02.108\le\frac{\sigma_1^2}{\sigma_2^2}\le15.257\right)=0.90$

كما يمكن الحصول على نفس النتيجة بدون استخدام تبايني العينتين المعدل كما يلي:

$$P\left(\frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{1-\frac{\alpha}{2},V_1,V_2}}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{\frac{\alpha}{2},V_1,V_2}}\right)=1-\alpha\Rightarrow$$

$$P\left(\frac{\left(\frac{16}{16-1}\right)\left(\frac{10-1}{10}\right)\frac{25}{04}}{02.85}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{\left(\frac{16}{16-1}\right)\left(\frac{10-1}{10}\right)\frac{25}{04}}{0.3937}\right)\Rightarrow$$

$$P\left(02.10\le\frac{\sigma_1^2}{\sigma_2^2}\le15.240\right)=0.90$$

2-أيجاد فترة الثقة لنسبة التباين عند درجة ثقة 98%

بما أن $(1-\alpha)100\%=98\%$ لدينا $1-\alpha=0.98\Leftarrow\alpha=0.02$ ودرجة الحرية للبسط والمقام هي $V_1=n_1-1\Rightarrow V_1=16-1=15, V_2=n_2-1\Rightarrow V_2=10-1=09$ ، نجد القيمة الصغرى والكبرى لتوزيع فيشر.

$$F_{1-\frac{\alpha}{2},V_1,V_2} \Rightarrow F_{0.99,15,10} = 04.56$$

$$F_{\frac{\alpha}{2},V_1,V_2} = \frac{01}{F_{1-\frac{\alpha}{2},V_2,V_1}} \Rightarrow F_{\frac{\alpha}{2},V_1,V_2} = \frac{01}{F_{0.99,10,15}} = \frac{01}{03.80} = 0.2632$$

نجد الآن الحد الأعلى والحد الأدنى لمجال الثقة:

$$L = \frac{\hat{S}_1^2/\hat{S}_2^2}{F_{1-\frac{\alpha}{2},V_1,V_2}} \Rightarrow L = \frac{26.67/04.44}{04.56} = 01.317$$

$$U = \frac{\hat{S}_1^2/\hat{S}_2^2}{F_{\frac{\alpha}{2},V_1,V_2}} \Rightarrow U = \frac{26.67/04.44}{0.2632} = 22.822$$

إذن مجال التقدير هو

$$P\left(01.317 \le \frac{\sigma_1^2}{\sigma_2^2} \le 22.822\right) = 0.98$$

كما يمكن الحصول على نفس النتيجة بدون استخدام تبايني العينتين المعدل كما يلي:

$$P\left(\frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{1-\frac{\alpha}{2},V_1,V_2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{\frac{\alpha}{2},V_1,V_2}}\right) = 1-\alpha \Rightarrow$$

$$P\left(\frac{\left(\frac{16}{16-1}\right)\left(\frac{10-1}{10}\right)\frac{25}{04}}{04.56} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\left(\frac{16}{16-1}\right)\left(\frac{10-1}{10}\right)\frac{25}{04}}{0.2632}\right) \Rightarrow$$

$$P\left(01.315 \le \frac{\sigma_1^2}{\sigma_2^2} \le 22.796\right) = 0.98$$

13- فترة الثقة للفرق بين الانحرافات المعيارية $(\sigma_1 - \sigma_2)$

نظرية15: إذا كان لدينا مجتمعان يتوزع كل منهما توزيعا طبيعيا، وسحبنا من المجتمع الأول الذي تباينه σ_1^2 كل العينات العشوائية ذات الحجم n_1، وحسبنا منها قيم المتغير S_1^2 ، وسحبنا من المجتمع الثاني الذي تباينه σ_2^2 كل العينات العشوائية الممكنة ذات الحجم n_2، وحسبنا منها قيم المتغير S_2^2 ، وكانت العينات المسحوبة من المجتمع الأول مستقلة عن العينات المسحوبة من المجتمع الثاني، فان فترة الثقة للانحرافات المعيارية عند درجة ثقة $(1-\alpha)100\%$ تعطى كالتالي:

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

تمرين27: أخذت عينة عشوائية من 150 مصباح من الصنف "أ"، فكان متوسط عمرها الإنتاجي هو 1400 ساعة وانحرافها المعياري 120 ساعة، كما أخذت عينة عشوائية أخرى من 100 مصباح من الصنف "ب"، فوجد متوسط عمرها الإنتاجي 1200 ساعة وانحرافها المعياري 80 ساعة.

المطلوب:

1-أوجد حدود الثقة للفرق بين الانحرافات المعيارية لمجتمع الصنفين "أ" و"ب" عند درجة ثقة 95%؟

2-أوجد حدود الثقة للفرق بين الانحرافات المعيارية لمجتمع الصنفين "أ" و"ب" عند درجة ثقة 99%؟

الحل:

لدينا من معطيات التمرين:

المجتمع الأول "أ": $n_1 = 150, \overline{X}_1 = 1400, S_1 = 120$

المجتمع الثاني "ب": $n_2 = 100, \overline{X}_2 = 1200, S_2 = 80$

1-إيجاد حدود الثقة للفرق بين الانحرافات المعيارية لمجتمع الصنفين "أ" و"ب" عند درجة ثقة 95%

بما أن $100\%(1-\alpha) = 95\%$

لدينا $1-\alpha = 0.95 \Leftarrow \alpha = 0.05$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,05}{2}} = Z_{0,975} = 01.960$

نجد الآن الحد الأدنى و الحد الأعلى لمجال الثقة:

$$L = (S_1 - S_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}} \Rightarrow L = (120-80) - (01.960)\sqrt{\frac{(120)^2}{2\times150} + \frac{(80)^2}{2\times100}} = 22.47$$

$$U = (S_1 - S_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}} \Rightarrow U = (120-80) + (01.960)\sqrt{\frac{(120)^2}{2\times150} + \frac{(80)^2}{2\times100}} = 57.53$$

إذن مجال التقدير هو $P\left(22.47 \le (\sigma_1 - \sigma_2) \le 57.53\right) = 0.95$

2- إيجاد حدود الثقة للفرق بين الانحرافات المعيارية لمجتمع الصنفين "أ" و"ب" عند درجة ثقة 99%

بما أن $100\%(1-\alpha) = 99\%$

لدينا $1-\alpha = 0.99 \Leftarrow \alpha = 0.01$ وبالتالي فان الدرجة المعيارية تساوي $Z_{1-\frac{0,01}{2}} = Z_{0,995} = 02.575$

نجد الآن الحد الأدنى و الحد الأعلى لمجال الثقة:

$$L = (S_1 - S_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}} \Rightarrow L = (120-80) - (02.575)\sqrt{\frac{(120)^2}{2\times150} + \frac{(80)^2}{2\times100}} = 16.97$$

$$U = (S_1 - S_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{2n_1} + \frac{S_2^2}{2n_2}} \Rightarrow U = (120-80) + (02.575)\sqrt{\frac{(120)^2}{2\times150} + \frac{(80)^2}{2\times100}} = 63.03$$

إذن مجال التقدير هو $P\left(16.97 \le (\sigma_1 - \sigma_2) \le 63.03\right) = 0.99$

تمارين المحور الثاني

التمرين الأول: إذا كان عمر المصابيح الكهربائية في مصنع ما، يخضع للتوزيع الطبيعي بانحراف معياري 40 ساعة، اختيرت عينة عشوائية من 25 مصباحا ووجد أن متوسط العمر هو 780 ساعة،

المطلوب:

1-أوجد فترة الثقة لمتوسط المجتمع لكل المصابيح المنتجة في هذا المصنع وذلك عند درجة ثقة 99%؟.

التمرين الثاني: إذا علمت أن الزمن الذي يقضيه الزبائن في احد المصارف يتوزع وفق التوزيع الطبيعي بانحراف معياري يساوي 04 دقائق، فإذا تم إختيار عينة عشوائية من 16 زبون ووجد أن متوسط الزمن الذي أمضوه بالمصرف كان 28.50 دقيقة.

المطلوب:

1-أوجد فترة الثقة حول متوسط الزمن الذي يمضيه الزبائن بصفة عامة بهذا المصرف وذلك عند درجة ثقة 99%؟

التمرين الثالث: أختيرت عينتين عشوائيتين ومستقلتين من 25 طالب و16 طالبة من جامعة ما، وتم قياس أوزانهم فكان متوسط الوزن لدى الطلبة هو 82.50كغ و58.14كغ لدى الطالبات، أما الانحراف المعياري للطلبة والطالبات فكان على التوالي 12.25كغ و08.75 كغ.

المطلوب:

1-أوجد فترة الثقة لمتوسط أوزان الطلبة عند درجة ثقة 99%؟

2- أوجد فترة الثقة لمتوسط أوزان الطالبات عند درجة ثقة 99%؟

3-أوجد فترة الثقة للفرق بين متوسطي أوزان الطلبة والطالبات وذلك عند درجة ثقة 99%؟

التمرين الرابع: في عينة عشوائية تتكون من 100 حذاء من إنتاج احد المصانع وجد بينهم 10 بها عيوب، وفي عينة عشوائية أخرى تتكون من 150 حذاء من إنتاج مصنع آخر ولنفس النوعية وجد من بينهم 12 بها عيوب.

المطلوب:

1-أوجد 90% فترة ثقة لتقدير الفرق بين نسبتي الأحذية المنتجة من قبل المصنعين وبها عيوب؟

2- أوجد 98% فترة ثقة لتقدير الفرق بين نسبتي الأحذية المنتجة من قبل المصنعين و بها عيوب؟

التمرين الخامس: عينة عشوائية مسحوبة من مجتمع يتوزع توزيعا طبيعيا بتباين مجهول، فإذا كان حجم العينة 13، وكان $\sum_{i=1}^{13}\left(X_i-\overline{X}\right)^2=128.41$.

المطلوب:

1-قدر تباين المجتمع والانحراف المعياري للمجتمع باستخدام فترة الثقة عند مستوى ثقة 95%؟

التمرين السادس: مجتمعان الأول توزيعه $X_1 \sim N\left(\mu_1,\sigma_1^2\right)$ ، والثاني توزيعه $X_2 \sim N\left(\mu_2,\sigma_2^2\right)$ ، سحبنا من المجتمع الأول عينة عشوائية حجمها05، ومن المجتمع الثاني عينة عشوائية مستقلة عن العينة الأولى حجمها 07، وحصلنا على النتائج التالية:

عينة المجتمع الأول: $\overline{X}_1 = 525.30, S_1^2 = 2273$

عينة المجتمع الثاني: $\overline{X}_2 = 510.80, S_2^2 = 1759$

المطلوب:

1-أوجد فترة الثقة لنسبة تباين المجتمع الأول إلى تباين المجتمع الثاني وذلك باستخدام مستوى ثقة 95%؟

2- أوجد فترة الثقة لنسبة الانحراف المعياري للمجتمع الأول إلى الانحراف المعياري للمجتمع الثاني وذلك باستخدام مستوى ثقة 95%؟

3-أوجد فترة الثقة للفرق بين الانحرافات المعيارية وذلك عند درجة ثقة 99%؟

التمرين السابع: يدعي احد المصانع انه أنتج نوع معين من الغذاء يحدث تخفيفا في الوزن بمقدار 04.50 كغ في المتوسط خلال فترة زمنية معينة مدتها أسبوعين من تناوله، ولتأكيد ادعائه أجريت دراسة على عينة من سبعة أشخاص فكانت أوزانهم قبل وبعد استخدام هذا النوع من الغذاء لمدة أسبوعين كما يلي:

الوزن	01	02	03	04	05	06	07
قبل	58.50	60.30	61.70	69.00	64.00	62.60	56.70
بعد	60.00	54.90	58.10	62.10	58.50	59.90	54.40

المطلوب:

1-أوجد 90% فترة ثقة لتقدير الفرق في متوسط الوزن على افتراض أن توزيع الأوزان يخضع للتوزيع الطبيعي؟

التمرين الثامن: إذا كان P هو تناسب النجاحات المشاهد في عينة حجمها n

المطلوب:

1- اثبت بان حدود ثقة تقدير تناسبات نجاح المجتمع P عند مستوى الثقة المحدد ب Z_c تعطى بالعلاقة التالية:

$$P = \frac{P + \frac{Z_c^2}{2n} \pm Z_c\sqrt{\frac{P(1-P)}{n} + \frac{Z_c^2}{4n^2}}}{1 + \frac{Z_c^2}{n}}$$

2-بافتراض أن حجم العينة هو 100 وان النسبة P=0.55،أوجد حدي الثقة عند مستوى ثقة 99.73%؟

التمرين التاسع: من خلال واقع السجلات بمصلحة الرصد الجوي في السنوات الخمس عشر الماضية تبين أن متوسط الأمطار التي سقطت على مدينة تيارت خلال شهر جانفي كانت 05.33 سم بانحراف معياري يساوي 01.20 سم، ومتوسط كمية الأمطار التي سقطت على مدينة مستغانم خلال نفس الشهر ولمدة عشر سنوات كان 03.64 سم بانحراف معياري يساوي 0.99 سم.

المطلوب:

1-أوجد 95% فترة الثقة لتقدير الفرق مابين متوسطي كمية الأمطار التي سقطت على المدينتين وذلك على افتراض أن البيانات من مجتمعين طبيعيين تباينهما متساوي؟

2- أوجد 95% فترة الثقة لتقدير الفرق مابين متوسطي كمية الأمطار التي سقطت على المدينتين وذلك على افتراض أن البيانات من مجتمعين طبيعيين تباينهما غير متساوي؟

التمرين العاشر: لتكن $X_4،X_3،X_2،X_1$ عينة عشوائية مأخوذة من مجتمع إحصائي له توزيع طبيعي $X \sim N(\mu,\sigma^2)$ ولتكن $T_4،T_3،T_2،T_1$ أربعة مقدرات معرفة بالعلاقات التالية:

$$T_1 = X_1, T_2 = \frac{(X_1 + X_3)}{02}, T_3 = \frac{(02X_1 + 03X_2 - 02X_3)}{03}, T_4 = \overline{X}$$

المطلوب:

1-أي المقدرات الأربعة هي مقدرات غير متحيزة لمتوسط الحسابي μ ؟

2-أي المقدرات الأربعة أفضل مقدر؟

3-أحسب الكفاءة النسبية لهذه التقديرات بالنسبة للتقدير T_4 ؟

التمرين الحادي عشر: ليكن لديك مجتمع يخضع للتوزيع الطبيعي بمتوسط حسابي μ وتباين σ^2، سحبت منه عينة عشوائية حجمها n بمتوسط حسابي قدره $\overline{X}$ وانحراف معياري قدره 01.80، فإذا علمت أن مجال الثقة للمتوسط الحسابي للمجتمع يعطى بالعلاقة التالية:

$$P(21.412 \le \mu \le 22.588) = 0.95$$

المطلوب:

1-أوجد قيمة المتوسط الحسابي للعينة $\overline{X}$؟

2- أحسب قيمة الخطأ المعياري للعينة؟

3-أوجد قيمة العينة المستخدمة في التقدير؟

المحور الثالث: إختبار الفروض
(إختبار معالم المجتمع)

تمهيد

ينقسم الإحصاء ألاستنتاجي أو الاستدلالي إلى نوعين وهما التقدير واختبارات الفروض (اختبار معالم المجتمع) وقد تطرقنا في المحور السابق إلى نظرية التقدير وكيفية استنتاج معالم المجتمع المجهولة وذلك بتقديرها إما بقيمة واحدة تحسب من عينة عشوائية مسحوبة من المجتمع، أو تقديرها بفترة حيث نكون واثقين بمقدار ثقة معين أن المعلمة المجهولة تقع داخل هذه الفترة، أما النوع الثاني والمتعلق باختبارات الفروض فسوف يتم التطرق فيه إلى كيفية استنتاج معالم المجتمع المجهولة عن طريق إجراء اختبارات لفروض إحصائية، وهي الطريقة الأكثر أهمية في ميدان الاستنتاج الإحصائي، ويعود الفضل الأول في اقترح هذا الإجراء أو الاختبار لكل من نيمان وبيرسون (J.Neman and E.Pearson) عام 1930م.

وقبل الخوض في هذا المحور سوف نتطرق إلى المفاهيم الأساسية لاختبار الفروض و التي نذكرها في النقاط التالية:

01- مفاهيم أساسية لاختبار الفروض

1-1 الفرضية الإحصائية والرفض والقبول: تعرف الفرضية الإحصائية على أنها عبارة أو تصور يتعلق بكامل المجتمع الإحصائي الأم مأخوذة عن عينة عشوائية منه،أما الرفض والقبول للفرضية الإحصائية و الصح والخطأ في تعابير تستخدم في هذا المجال ولابد من الإشارة هنا إلى أن رفض الفرضية لا يعني خطأها أما قبول الفرضية الإحصائية فلا يعني بالضرورة صحتها ولكنه يعني عدم وجود حجة والدليل القاطع لرفضها، والهدف من الاختبار الإحصائي هو اختبار فرضية حول معلمة أو أكثر من معالم المجتمع الإحصائي، وبالتالي فان الاختبار الإحصائي يتكون من العناصر التالية:

أ-الفرضية الصفرية او فرضية العدم (Null Hypothèses): إذا صيغت فرضية حول توزيع أو خصائص متغير عشوائي (أو عدة متغيرات عشوائية) بهدف اختبار صحتها فتدعى عادة بفرضية العدم ويرمز لها بالرمز (H_0)، وتعني انعدام الاختلاف والتباين أي التساوي(=)،كما تدعى ببعض الأدبيات الإحصائية بالفرضية الصفرية أو الأساسية، وعادة يقوم الباحثون في ميادين المعرفة المختلفة بصياغة فرضية العدم ومهمة الإحصائي تكمن في مساعدتهم في اختيار فرضية العدم المناسبة واتخاذ القرار بقبول أو رفض تلك الفرضية، إذن فرضية العدم هي الفرضية التي يتم اختبارها وعادة إلى جانب فرضية العدم تصاغ فرضية أخرى تدعى بالفرضية البديلة.

ب-الفرضية البديلة (Alternative Hypothèses): والتي يرمز لها بالرمز(H_1)عادة وتعني وجود الاختلاف والتباين وعدم انعدامها، كما تسمى أحيانا بفرضية القبول أي القبول بوجود الاختلاف والتباين ($\neq$),($>$,$<$) ، وتأسيسا على ما تقدم، يمكن صياغة فرضية العدم (H_0) والفرضية البديلة (H_1)، وفقا لأحد الأشكال الثلاثة التالية:

1-اختبار ذو جانبين أو ذيلين: يتحقق هذا الاختبار وفق العلاقة التالية

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

2-اختبار من الجانب الأيمن أو الجانب العلوي : يتحقق هذا الاختبار وفق العلاقة التالية

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

3- إختبار من الجانب الأيسر أو الجانب السفلي: يتحقق هذا الاختبار وفق العلاقة التالية

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

2-1 إحصاءة الاختبار: تعرف إحصاءة الاختبار بأنها متغير عشوائي لها توزيع احتمالي معروف، وتستخدم لوصف العلاقة بين القيم النظرية للمجتمع والقيم المحسوبة من العينة، وتعتمد إحصاءة الاختبار على ما سيتم اختباره من معلمات المجتمع المدروس، لذلك فهي تختلف باختلاف الحالة المدروسة للمعلمة، وتسمى عادة باحصاءة الاختبار المحسوبة، وتكون إحصاءات الاختبار على أنواع عدة نذكر منها(إحصاءة الاختبار(Z)، احصاءة الاختبار(T)، احصاءة الاختبار $\left(\chi^2\right)$ ، واحصاءة الاختبار (F)).

ويتم تقسيم كل النتائج الممكن الحصول عليها، أي كل القيم التي يمكن أن تأخذها إحصائية الاختبار لمجموعتين غير متداخلتين، إحداهما تشمل النتائج التي إذا ظهرت نقبل فرض العدم وتسمى منطقة القبول، والأخرى تشمل النتائج التي إذا ظهرت نرفض فرض العدم وتسمى منطقة الرفض، وبالتالي يقسم توزيع المعاينة لإحصائية الاختبار إلى منطقتين يمكن تعريفهما كما يلي:

2-1منطقة القبول(acceptance region):هي المنطقة التي تحتوي على قيم إحصائية الاختبار التي تؤدي إلى عدم رفض العدم (H_0)، أي قبول فرض العدم (H_0).

2-2 منطقة الرفض (rejection region): وهي المنطقة التي تحتوي على قيم إحصائية الاختبار التي تؤدي إلى رفض فرض العدم (H_0) ، وتسمى كذلك بالمنطقة الحرجة (the critical region).

والقيمة أو القيم التي تفصل بين هاتين المنطقتين تسمى بالقيمة أو القيم الحرجة، وتعرف كالتالي:

2-3 القيمة الحرجة (the critical region): هي القيمة التي تفصل بين منطقة الرفض ومنطقة القبول.

3-1 الخطأين من النوع الأول والثاني: إن اتخاذ القرار يعتمد على قيمة المشاهدة لإحصائية الاختبار، أي على القيمة المحسوبة من العينة العشوائية المسحوبة من المجتمع، وقد تكون هذه العينة لا تمثل المجتمع الذي سحبت منه تمثيلا صحيحا، مما يؤدي إلى وقوع متخذ القرار في خطأ من اثنين هما:

أ-الخطأ من النوع الأول: يحدث هذا الخطأ إذا كان فرض العدم في الحقيقة صحيحا، ولكن بيانات العينة تظهر انه غير صحيح، أي أن نتائج العينة تؤدي الر رفض فرض العدم مع انه في الواقع صحيح، ويرمز لاحتمال وقوع خطأ من النوع الأول أي لإحتمال رفض فرض العدم مع انه في الواقع صحيح بالرمز α ويطلق عليه اسم مستوى المعنوية أي أن:

$\alpha = P$ (ارتكاب خطأ من النوع الأول)

$=P(H_0$ فرض العدم H_0 صحيح/رفض فرض العدم)

ب- الخطأ من النوع الثاني:يحدث هذا الخطأ نتيجة لقبول فرض العدم مع أنه في الواقع غير صحيح، أي أن بيانات العينة تؤيد فرض العدم، مع أن فرض العدم في الحقيقة غير صحيح، ويرمز إلى احتمال وقوع الخطأ من النوع الثاني، أي احتمال قبول فرض العدم مع أن فرض العدم في الواقع خطأ بالرمزβ، أي أن:

$\beta=P($ ارتكاب خطأ من النوع الثاني)

$=P(H_0$ فرض العدم H_0 خطأ/قبول فرض العدم)

$=P(H_0$ الفرض البديلH_1 هو الصحيح /قبول فرض العدم)

ويمكن تلخيص الحالات السابقة من خلال الجدول التالي:

القرار	H_0	
	H_0 صحيحة	H_0 غير صحيحة
قبول H_0	قرار سليم	قرار خاطئ (خطأ من النوع الثاني)
رفض H_0	قرار خاطئ (خطأ من النوع الأول)	قرار سليم

إن من الملائم أن اختبار الفروض يقاس بالاحتمالات الناتجة عن الخطأ الأول والثاني و التي رمزنا لها بالرموزα وβ، وحيث أن α هو الاحتمال الناتج عن رفض H_0 عندما تكون صحيحة فان منطقة الرفض تزداد بازدياد α وفي نفس الوقت ينقص β عند ثبات حجم العينة، وعند إنقاص منطقة الرفض فان قيمة β تزداد، أما إذا زدنا حجم العينة فان معلومات أكثر ستوفر لاتخاذ القرار على أساسه وعندئذ ستنقص قيمتي α وβ.

إن قيمة β وهو الاحتمال الناتج عن الخطأ من النوع الثاني يتغير بالاعتماد أساسا على القيم الحقيقية لمعلمة المجتمع.

تمرين01:إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بوسط حسابي μ مجهول، وتباينه $\sigma^2=625$، وأردنا إجراء الاختبار الإحصائي التالي بخصوص الوسط الحسابي للمجتمع μ التالي:

$$H_0: \mu = 100$$
$$H_1: \mu = 80$$

إذا كانت القاعدة التي وضعها متخذ القرار اعتمادا على خبرته السابقة كما يلي:

يرفض H_0 إذا كان الوسط الحسابي للعينة $\overline{X}$ أقل من 67، ويقبل H_0 إذا كان $\overline{X}$أكبر من أو يساوي67.

المطلوب:

-أحسب احتمال وقوع الخطأ من النوع الأول، واحتمال وقوع الخطأ من النوع الثاني، إذا كان قراراه سيعتمد على عينة عشوائية مسحوبة من هذا المجتمع حجمها 05.؟

الحل:

لدينا من معطيات التمرين : $\overline{X} \sim (\mu, 625), \overline{X} = 67, n = 05$

بناءا على القاعدة التي وضعها متخذ القرار، نجد أن القيمة الحرجة تساوي67، وبما أن احتمال الوقوع في الخطأ من النوع الأول هو (فرض العدم H_0 صحيح/رفض فرض العدم α=P(H_0

إذن $\alpha = P(\overline{X} < 67 / \mu = 100)$

وبما أن توزيع المجتمع هو توزيع طبيعي، إذن سيكون توزيع المعاينة للإحصائية $\overline{X}$ توزيعا طبيعيا بوسط حسابي وتباين قدرهما على التوالي

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{625}{5} = 125 \quad ، \quad \mu_{\overline{X}} = \mu$$

فعندما يكون فرض العدم هو الفرض الصحيح، يكون $\mu_{\overline{X}} = \mu = 100$

وعندما يكون الفرض البديل هو الفرض الصحيح يكون $\mu_{\overline{X}} = \mu = 80$

لحساب α نحسب القيمة المعيارية (Z) المقابلة للقيمة الحرجة $\overline{X} = 67$ كالتالي:

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} = \frac{67 - 100}{\sqrt{125}} = -02.95$$

إذن $\alpha = P(\overline{X} < 67) = P(Z < -02.95) = F(-02.95) = 01 - F(02.95) = 01 - 0.99841 = 0.00159$

وبما أن احتمال الوقوع في الخطأ من النوع الثاني هو:

(الفرض البديلH_1 هو الصحيح /قبول فرض العدم β=P(H_0

إذن $\beta = P(\overline{X} \geq 67 / \mu = 80)$

ولحساب قيمة β نحسب القيمة المعيارية المقابلة للقيمة الحرجة $\overline{X} = 67$ كما يلي:

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} = \frac{67 - 80}{\sqrt{125}} = -01.16$$

إذن $\beta = P(\overline{X} \geq 67) = P(Z \geq -01.16) = F(01.16) = 0.87698$

1-4 مستوى المعنوية(Level of Singificance) (α): يعرف مستوى المعنوية بأنه (إحتمال رفض الفرضية العدمية H_0 عندما تكون هذه الفرضية صحيحة)، بمعنى آخر هي (إحتمال الوقوع في الخطأ من النوع الأول)، ويرمز إلى مستوى المعنوية عادة بالرمز (α) حيث أن:

(فرض العدم H_0 صحيح/رفض فرض العدم α=P(H_0

ويسمى مستوى المعنوية (α) أحيانا بحجم منطقة الرفض في حين يسمى المقدار ($1-\alpha$) بدرجة الثقة.

ويقوم الباحث قبل البدء بعملية الاختبار بتحديد مستوى المعنوية (α) عند تصميم التجربة منذ البداية، حيث يتوقف تحديد مستوى المعنوية (α) على طبيعة البحث أو الدراسة وغالبا ما يتم تحديد مستوى المعنوية (α) حجمها 1% او5% او10%.

فعلى سبيل المثال، لو تم اختيار مستوى المعنوية (α) مساو إلى (0.05) فهذا يعني انه من المحتمل الوقوع بالخطأ من النوع الأول، أي احتمال رفض الفرضية العدمية H_0 وهي صحيحة بمقدار خمس مرات في كل (100) مرة، في حين يكون الاستنتاج صحيحا بدرجة ثقة ($1-\alpha=0.95$).

5-1 قوة الاختبار (Power of the Test)($1-\beta$): تعرف قوة الاختبار بأنها احتمال رفض الفرضية العدمية H_0 عندما تكون هذه الفرضية خاطئة أي أن :

P.O.T=P(فرض العدم H_0 هو خاطئ/رفض فرض العدم H_0)

=1- P(فرض العدم H_0 خطأ/قبول فرض العدم H_0)

$= 1-\beta$

يتضح من خلال المعادلة الأخيرة، بان (β) تمثل احتمال الوقوع في الخطأ من النوع الثاني، أي أن

β=P (فرض العدم H_0 خطأ/قبول فرض العدم H_0)

وبالتالي يتضح بأنه كلما كان احتمال الوقوع في الخطأ من النوع الثاني (β) قليل كلما أدى ذلك إلى زيادة قوة الاختبار ($1-\beta$)، وهذا يعني زيادة شدة رفض الفرضية العدمية (H_0) عندما تكون هذه الفرضية خاطئة.

وتأسيسا على ما تقدم، يمكن توضيح العلاقة بين الخطأين من النوع الأول (α) والثاني (β) على النحو التالي:

1-ن انخفاض احد الخطأين يؤدي إلى زيادة الخطأ الآخر.

2-إن زيادة حجم العينة (n)، يقلل من إحتمال الوقوع في كلا الخطأين (α) و(β)، وبالتالي إلى زيادة درجة الثقة.

3-يحسب احتمال الوقوع في الخطأ من النوع الأول (α) على أساس الفرضية العدمية (H_0)، في حين يحسب إحتمال الوقوع في الخطأ من النوع الثاني (β) على أساس الفرضية البديلة (H_1).

تمرين02: إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا، وسطح الحسابي μ مجهول، وتباينه $\sigma^2=81$ ، وأردنا إجراء الاختبار الإحصائي التالي بخصوص الوسط الحسابي للمجتمع μ:

$$H_0:\mu=100$$
$$H_1:\mu=112$$

إذا كانت القاعدة التي وضعها متخذ القرارات اعتمادا على خبرته السابقة كما يلي:

يرفض (H_0) إذا كان الوسط الحسابي للعينة $\overline{X}$ أكبر من 107، ويقبل (H_0) إذا كان الوسط الحسابي للعينة $\overline{X}$ اقل أو يساوي107

المطلوب:

1-أحسب احتمال وقوع خطأ من النوع الأول، واحتمال وقوع خطأ من النوع الثاني، إذا كان قراره سيعتمد على عينة عشوائية مسحوبة من هذا المجتمع حجمها 09.؟

2- إذا قام الباحث بتثبيت القيمة الحرجة والمساوية ل107، مع تغيير حجم العينة إلى 12، فما هي قيمة كل من الخطأ الأول والثاني وماذا تستنتج؟

الحل:

لدينا من معطيات التمرين : $\overline{X} \sim (\mu,81), \overline{X}=107, n=09$

بناءا على القاعدة التي وضعها متخذ القرار، نجد أن القيمة الحرجة تساوي107، وبما أن احتمال الوقوع في الخطأ من النوع الأول هو α=P(رفض فرض العدم H_0/فرض العدم H_0 صحيح)

إذن $\alpha = P\left(\overline{X} > 107 / \mu = 100\right)$

وبما أن احتمال الوقوع في الخطأ من النوع الثاني هو:

β=P(قبول فرض العدم H_0/ الفرض البديل H_1 هو الصحيح)

إذن $\beta = P\left(\overline{X} \le 107 / \mu = 112\right)$

وبما أن توزيع المجتمع هو توزيع طبيعي، إذن سيكون توزيع المعاينة للإحصائية $\overline{X}$ توزيعا طبيعيا بوسط حسابي وتباين قدرهما على التوالي $\mu_{\overline{X}} = \mu$ ، $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{81}{09} = 09$

فسنجد أن قيمة (α) و(β) في هذه الحالة كما يلي:

$$\alpha = P\left(\overline{X} > 107\right) = P\left(Z > \frac{107-100}{\sqrt{09}}\right) = P(Z > 02.33) = 01 - P(Z < 02.33) = 01 - F(02.33)$$

$$= 01 - 0.99010 = 0.0099$$

$$\beta = P\left(\overline{X} \le 107\right) = P\left(Z \le \frac{107-112}{\sqrt{09}}\right) = p(Z \le -01.67) = F(-01.67) = 01 - F(01.67)$$

$$= 01 - 0.95254 = 0.04746$$

2-حساب قيمة (α) و(β) في حالة تثبيت القيمة الحرجة مع تغيير حجم العينة إلى 12

نحسب التباين الجديد كالتالي: $\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{81}{12} = 06.75$

$$\alpha = P\left(\overline{X} > 107\right) = P\left(Z > \frac{107-100}{\sqrt{06.75}}\right) = P(Z > 02.69) = 01 - P(Z < 02.69) = 01 - F(02.69)$$

$$= 01 - 0.99643 = 0.00357$$

$$\beta = P\left(\overline{X} \le 107\right) = P\left(Z \le \frac{107-112}{\sqrt{06.75}}\right) = p(Z \le -01.92) = F(-01.92) = 01 - F(01.92)$$

$$= 01 - 0.97257 = 0.002743$$

الاستنتاج: نلاحظ انه عندما يزداد حجم العينة، يقل إحتمال ظهور الخطأ من النوع الأول (α) ويقل احتمال ظهور الخطأ من النوع الثاني (β)، وهذا أمر منطقي لان هذه الأخطاء سببها دراسة العينة بدلا

من دراسة المجتمع ككل، وبالتالي كلما زاد حجم العينة كلما قل هذان الخطآن إلى أن يساويا الصفر عندما تشمل العينة كل مفردات المجتمع محل الدراسة.

تمرين03: بفرض انه تم اختيار عينة عشوائية حجمها 09 من مجتمع طبيعي بمتوسط μ وتباين $\sigma^2 = 16$ ، وان الفرضية كانت كالتالي

$$H_0 : \mu = 20$$
$$H_1 : \mu = 24$$

فإذا علمت أن قيمة (α) الخطأ من النوع الأول هو 0.05 وأننا نرفض H_0 إذا كانت $\overline{X} \geq a$

المطلوب:

1-أوجد قيمة a (القيمة الحرجة)؟

2-أوجد قوة الاختبار؟

الحل:

لدينا من المعطيات $\overline{X} \sim (\mu, 16), \overline{X} = a, n = 09, \alpha = 0.05$

بما أن العينة من مجتمع طبيعي إذن $\mu_{\overline{X}} = \mu$ ، $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n} = \frac{16}{09}$

وحيث انه من تعريف α (فرض العدم H_0 صحيح/رفض فرض العدم H_0)α=P

إذن $= P\left(\overline{X} \geq a / \mu = 20\right) = P\left(\frac{\overline{X} - 20}{\sigma_{\overline{X}}} \geq \frac{a - 20}{\sqrt{16/09}}\right) = P\left(Z \geq \frac{03a - 60}{04}\right) = P\left(Z \geq Z_0\right)$

ومن جدول التوزيع الطبيعي المعياري نجد أن $P\left(Z \geq 01.645\right) = 0.05$

إذن $\frac{03a - 60}{04} = 01.645 \Rightarrow a = 22.193$

إذن سوف نرفض H_0 إذا كانت $\overline{X} \geq 22.193$

2-قوة الاختبار:إن متوسط العينة $\overline{X}$ يطلق عليه في هذه الحالة تسمية إحصاءة الاختبار، وان قوة الاختبار تكون كالتالي: (H_0 خاطئة /رفض H_0)1−β=P

$$1 - \beta = P\left(\overline{X} \geq 22.193 / \mu = 24\right)$$
$$= P\left(\frac{\overline{X} - 24}{04/03} \geq \frac{22.193 - 24}{04/03}\right) = P\left(Z \geq -01.36\right) = F\left(01.36\right) = 0.91309$$

1-6 مستوى المعنوية المشاهد (P–Value): هو إحتمال الحصول على قيمة لإحصاءة الاختبار كبيرة مثلما القيمة المحسوبة فعلا لهذه الإحصاءة أو أكبر منها عندما يكون فرض العدم صحيحا وفي نفس اتجاه فرض البديل.

وبعبارة أخرى إن مستوى المعنوية المشاهد لأي اختبار هو أصغر قيمة لمستوى المعنوية (α) التي يمكن عندها رفض فرض العدم، وبصفة عامة سوف نرفض فرض العدم إذا كان مستوى المعنوية المشاهد اصغر من أو يساوي (α) ولا نرفضه خلاف ذلك.

1-7 المنطقة الحرجة والقيم الحرجة: تعرف المنطقة الحرجة بأنها المنطقة التي عندها يتم رفض الفرضية العدمية (H_0) والتي تقع فيها إحصاءة الاختبار المحسوبة، بمعنى آخر تعرف المنطقة بأنها جزء من المساحة تحت منحنى دالة التوزيع الاحتمالي لإحصاءة الاختبار، حيث إن هذه المساحة تمثل احتمال رفض الفرضية العدمية (H_0) عندما تكون هذه الفرضية صحيحة.

مما تقدم يتضح بان مساحة المنطقة تحت منحنى دالة التوزيع الاحتمالي لإحصاءة الاختبار يمثل (α) في حالة الاختبار من جانب واحد، أو تمثل $\left(\frac{\alpha}{2}\right)$ في حالة الاختبار من جانبين.

في حين تعرف القيم الحرجة بأنها قيم جدولية يتم استخراجها من قيم التوزيع الاحتمالي لإحصاءة الاختبار، والتي تتحدد بموجبها مناطق رفض الفرضية العدمية (H_0) ومناطق قبولها، وتعتمد القيم الحرجة على ما يلي:

1-مستوى المعنوية (α).

2-الفرضية البديلة (H_1)، كأن تكون ذات جانب واحد أو جانبين.

3-التوزيع الاحتمالي لإحصاءة الاختبار، كأن يكون احد التوزيعات الآتية $\left[Z,\chi^2,T,F\right]$

4-عدد درجات الحرية، فيما إذا كان التوزيع الاحتمالي أحد توزيعات المعاينة.

ولتوضيح مفهوم المنطقة الحرجة ومفهوم القيم الحرجة نفترض لدينا التوزيع الاحتمالي لإحصاءة الاختبار هو التوزيع الطبيعي المعياري (Z)، ويراد إختبار الفرضية الإحصائية الآتية: $H_0:\mu=\mu_0$

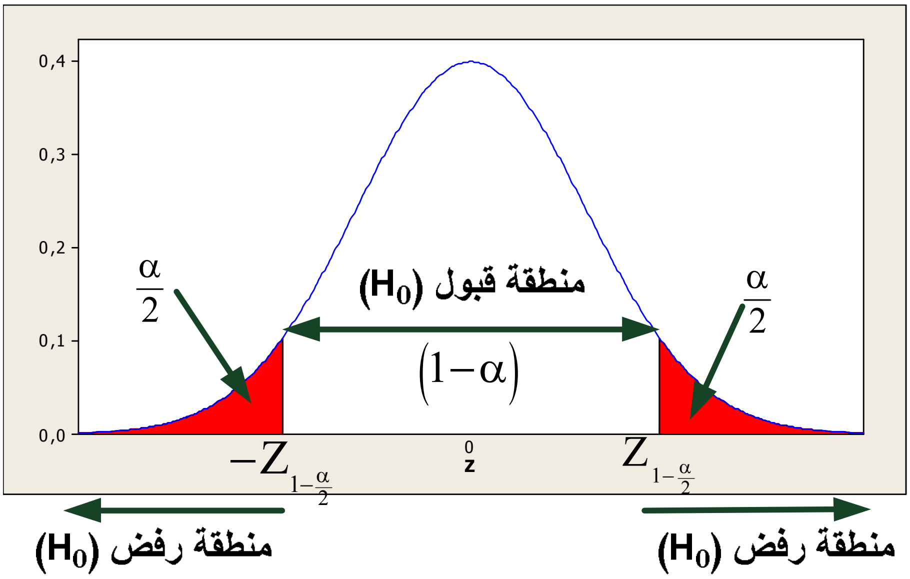
ضد الفرضية البديلة (H_1) التي تأخذ أحد الأشكال التالية، عند مستوى معنوية (α).

$$1-H_1:\mu\neq\mu_0$$
$$2-H_1:\mu>\mu_0$$
$$3-H_1:\mu<\mu_0$$

الحالة الأولى:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

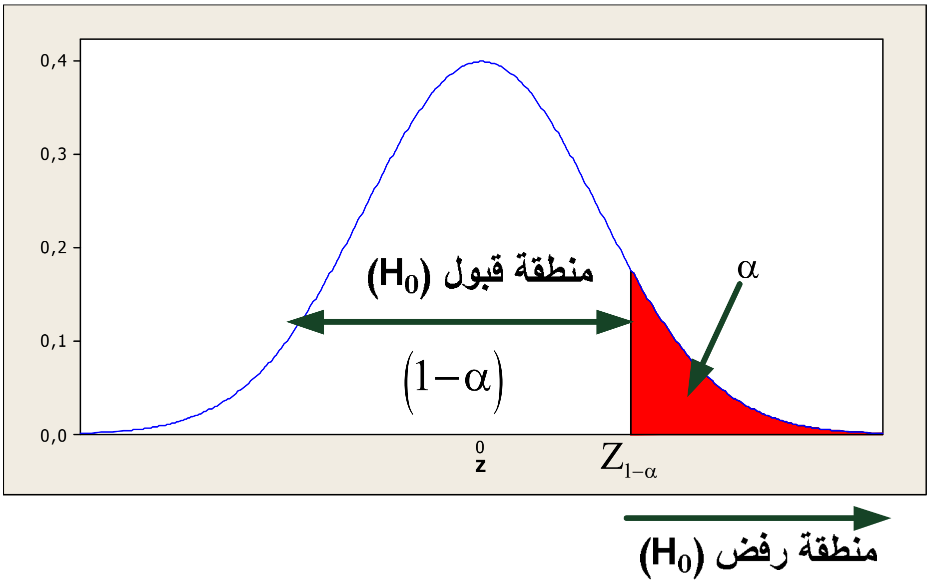
تأخذ الشكل التالي



الحالة الثانية:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

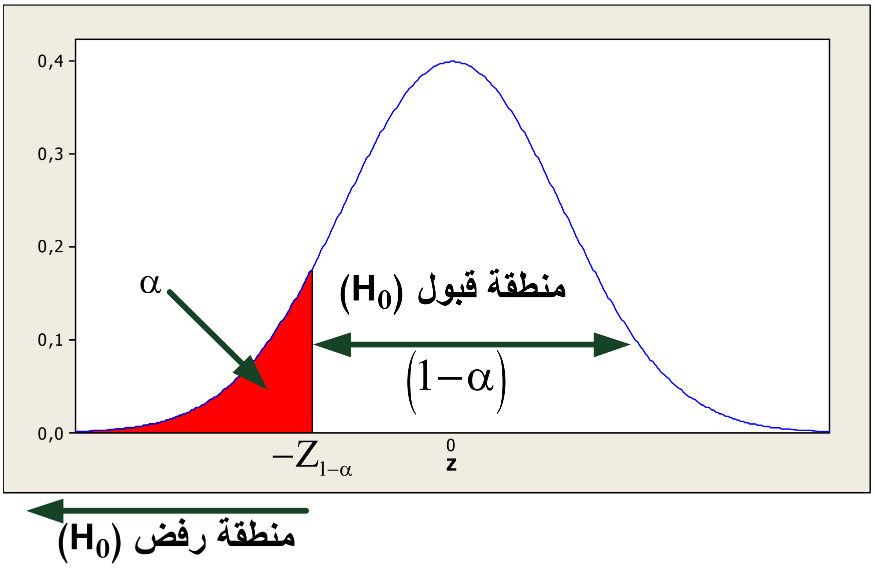
تأخذ الشكل التالي



الحالة الثالثة:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

تأخذ الشكل التالي



1-8 الخطوات المتبعة في اختبار الفرضيات: فيما يلي أهم الخطوات المتبعة في اختبار الفرضيات الإحصائية، نوجزها بما يلي

1-تحديد فرضية العدم(H_0) واختيار الفرضية البديلة المناسبة (H_1) من بين الحالات الثلاث:

طرف واحد $\theta > \theta_0$ أو $\theta < \theta_0$، أو ذات طرفين $\theta \neq \theta_0$

2-تحديد مستوى المعنوية (α) (احتمال الرفض) كان يكون $(0.01$ أو 0.05 أو$0.10)$ و القبول $P=(1-\alpha)100\%$

3-تحديد نوعية التوزيع واختيار الإحصائية المناسبة $\left[Z,\chi^2,T,F\right]$ وإيجاد الدرجة أو القيمة الحرجة لها من الجداول الخاصة بها ونرمز لها بالرمز $\left[Z_{tab},\chi_{tab}{}^2,T_{tab},F_{tab}\right]$

4-حساب القيمة الإحصائية θ_{cal} المحددة في الفقرة (03) من بيانات العينة بحيث $\theta_{cal} = \left[Z_{cal},\chi_{cal}{}^2,T_{cal},F_{cal}\right]$

5-تحديد القيمة الحرجة حسب نوعية التوزيع المستخدم، واعتمادا على مستوى المعنوية (α)، ونوع الفرضية البديلة (H_1)، ودرجات الحرية (ddl) في حالة توزيعات المعاينة.

6-التحليل ومقارنة قيمة الإحصائية المحسوبة θ_{cal} من العينة بالدرجة الحرجة θ_{tab}

7-إتخاذ القرار بشان رفض أو قبول الفرضية العدمية (H_0)، بعد مقارنة قيمة إحصاءة الاختبار المحسوبة θ_{cal} مع القيم الحرجة θ_{tab} حسب نوع الفرضية البديلة (H_1)، مستوى المعنوية (α) فإذا كانت

7-1 قيمة إحصاءة الاختبار المحسوبة θ_{cal} تقع في منطقة رفض الفرضية العدمية (H_0)، فإن ذلك يدل على رفض الفرضية العدمية (H_0)، وقبول الفرضية البديلة (H_1).

2-7 أما إذا وقعت قيمة إحصاءة الاختبار المحسوبة θ_{cal} في منطقة قبول الفرضية العدمية (H_0)، فإن ذلك يدل على قبول الفرضية العدمية (H_0)، ورفض الفرضية البديلة (H_1).

2-إختبارت الفروض حول المتوسط الحسابي (μ)

بافتراض لدينا $\left(X_1, X_2, X_3, \ldots\ldots\ldots\ldots\ldots, X_n\right)$ تمثل مشاهدات عينة عشوائية قوامها (n) مشاهدة، تم إختيارها من مجتمع طبيعي بوسط حسابي (μ) وتباين σ^2، وان $\left(\overline{X}, S^2\right)$ يمثلان الوسط الحسابي والتباين لمشاهدات هذه العينة على التوالي، يتم تقديرهما على أساس مشاهدات العينة، علما أن هذا الاختبار يكون من جانب واحد أو جانبين.

وعلى افتراض أننا نرغب في إختبار الفرضية العدمية $\left(H_0 : \mu = \mu_0\right)$ ضد أي فرضية بديلة أخرى، عندئذ هنالك حالتين هما:

2-1عندما يكون تباين المجتمع $\left(\sigma^2\right)$معلوما، وحجم العينة (n≥30)

<u>نظرية01:</u>إن من ابسط وأكثر الفرضيات اختبارا إذا توافرت بيانات كافية عن المجتمع الإحصائي مدار البحث هي الفرضيات الخاصة بالمتوسط، فإذا كانت $\left(X_1, X_2, X_3, \ldots\ldots\ldots\ldots\ldots, X_n\right)$ تمثل عينة عشوائية من مجتمع توزيعه طبيعي (أو غير طبيعي ولكن شروط تطبيق نظرية النهاية المركزية تبقى صحيحة)، متوسطه الحسابي μ مجهول وتباينه $\left(\sigma^2\right)$ معلوم، وأردنا اختبار ما ذا كانت μ تساوي قيمة معينة ولتكن μ_0 مثلا أم لا، ومن خلال المحور السابق (التقدير بنقطة) علمنا أن $\overline{X}$ مقدر جيد للمتوسط μ، وعليه فان أي قرار يتعلق بالمتوسط μ سيكون مبني على أساس $\overline{X}$، كما يمكن ربط هذا الاختبار بفترات الثقة حول المتوسط μ والتي سبق وان تعرضنا إليها في (المحور الثاني: التقدير) وذلك على أساس انه إذا كانت فترة الثقة الناتجة أي $\overline{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ تحتوي μ_0 فإنه لا يمكن رفض الفرضية أما إذا كانت μ_0 تقع خارج حدي هذه الفترة فنه يجب رفض الفرضية وفي هذه الحالة إما أن تكون:

$$\mu_0 < \overline{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad \text{أو} \quad \mu_0 > \overline{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

وأنه يمكن إعادة كتابة المتباينتين على التوالي كما يلي:

$$\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > Z_{1-\frac{\alpha}{2}} \quad \text{أو} \quad \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -Z_{1-\frac{\alpha}{2}}$$

وبتعريف $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ حيث $Z \sim N(0,1)$، فإنه وفقا لما سبق نرفض الفرضية $\mu = \mu_0$ إذا كانت $Z < -Z_{1-\frac{\alpha}{2}}$ أو $Z > Z_{1-\frac{\alpha}{2}}$ ، ولكن الفرضية الإحصائية قد تكون من طرف واحد أو من طرفين وذلك حسب الفرضية البديلة (H_1)، وعليه سيكون هناك فرق بين منطقتي الرفض وذلك كما يتضح من خلال الجدول التالي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	نرفض (H_0) إذا كانت $Z_{cal} \geq Z_{tab}$ أو $Z_{cal} \leq -Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$
ب-إختبار من طرف واحد: $H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	نرفض (H_0) إذا كانت $Z_{cal} \geq Z_{tab}$ بحيث $Z_{tab} = Z_{1-\alpha}$ أو إذا كانت $Pvalue < \alpha$
ج- إختبار من طرف واحد: $H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	نرفض (H_0) إذا كانت $Z_{cal} \leq -Z_{tab}$ بحيث $Z_{tab} = -Z_{1-\alpha}$ أو إذا كانت $Pvalue < \alpha$

ولغرض إتخاذ القرار الإحصائي حول رفض أو عدم رفض الفرضية العدمية(H_0) بعد المقارنة اعتمادا على مستوى المعنوية ونوع الفرضية البديلة (H_1)، والجدول التالي يوضح بعض القيم الجدولية الشائعة الاستخدام ل (Z) التي تخضع للتوزيع الطبيعي المعياري .

(H_1)	$H_0: \mu \neq \mu_0$	$H_1: \mu > \mu_0$ أو $H_1: \mu < \mu_0$
α	$Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$
0.01	02.575	02.326
0.05	01.960	01.645
0.10	01.645	01.285

<u>تمرين04:</u> في إختبار القدرات الذكائية والذي يتبع التوزيع الطبيعي بمتوسط حسابي يساوي 110 وتباين يساوي 100، فإذا تقدم لهذا الاختبار عينة تتكون من 25 طالبا بمدرسة معينة وكان متوسط درجاتهم بهذا الاختبار يساوي115.

<u>المطلوب:</u>

1-أكتب فرضية الاختبار؟

2-هل يمكن القول بان متوسط درجات الطلبة بصفة عامة في هذه المدرسة يختلف عن المتوسط العام وذلك عند مستوى معنوية 05%؟

الحل:

1-كتابة فرضية الاختبار

بفرض أن μ تمثل متوسط درجات الطلبة بصفة عامة بهذه المدرسة وان $\sigma^2 = 100$ تمثل تباين درجات الاختبار بصفة عامة، و $\overline{X} = 115$ تمثل متوسط درجات الاختبار لعينة الطلبة بهذه المدرسة، إذن خطوات الاختبار الإحصائي تكون كالتالي:

$$H_0 : \mu = 110$$
$$H_1 : \mu \neq 110$$

-2

-**إحصاءة الاختبار :** لدينا من المعطيات $\overline{X} = 115, \sigma^2 = 100, n = 25, \mu_0 = 110$ وبالتالي فان إحصاءة الاختبار تكون كالتالي

$$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{115 - 110}{10/\sqrt{25}} = 02.50$$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية $Z_{1-\frac{\alpha}{2}} = Z_{0.975} = Z_{tab} = 01.960$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $(Z_{cal} = 02.50)$ هي أكبر من القيمة الجدولية $(Z_{tab} = 01.960)$، وهذا يعني رفض الفرضية العدمية (H_0)، أي أن الاختبار معنوي ونقبل الفرضية البديلة (H_1).

وباستخدام مفهوم مجال الثقة يتضح لنا أننا نرفض (H_0) إذا كانت:

$$\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > 01.960 \quad \text{أو} \quad \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -01.960$$

أي إذا كانت $\overline{X} > \mu_0 + 01.960\frac{\sigma}{\sqrt{n}}$ أو $\overline{X} < \mu_0 - 01.960\frac{\sigma}{\sqrt{n}}$

أي انه إذا كانت $\overline{X} > 113.92$ أو $\overline{X} < 106.08$

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$\begin{aligned} P - value &= P\left(Z \leq -02.50 \vee Z \geq 02.50\right) \\ &= F\left(-02.50\right) + 01 - F\left(02.50\right) = 01 - F\left(02.50\right) + 01 - F\left(02.50\right) \\ &= 02 - 02F\left(02.50\right) = 02 - 02 \times 0.99379 = 0.01242 \end{aligned}$$

نلاحظ أن $0.01242 < 0.05$ $\Rightarrow$ $Pvalue < \alpha$، وبالتالي نرفض (H_0).

تمرين05: اختيرت عينة عشوائية قوامها 64 مشاهدة من مجتمع يخضع للتوزيع الطبيعي بوسط حسابي μ مجهول وتباين $\sigma^2 = 25$ ، وقد لوحظ بان متوسط مشاهدات العينة بلغ 20.

المطلوب:

1-أكتب فرضية الاختبار؟

2-هل تعتقد أن متوسط المجتمع يقل عن 23، وذلك عند مستوى معنوية $(\alpha=0.01)$؟

الحل:

1-كتابة فرضية الاختبار

$$H_0 : \mu = 23$$
$$H_1 : \mu < 23$$

-2

إحصاءة الاختبار: لدينا من المعطيات $\overline{X} = 20, \sigma^2 = 25, n = 64, \mu_0 = 23$ وبالتالي فان إحصاءة الاختبار تكون كالتالي

$$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{20-23}{05/\sqrt{64}} = -04.80$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد وفي الجانب الأيسر (القيم السفلى) و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.01 نستخرج قيمة Z الجدولية

$$-Z_{1-\alpha} = -Z_{1-0.01} = -Z_{0.99} = -Z_{tab} = -02.326$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal} = -04.80)$،هي أقل من القيمة الجدولية $(Z_{tab} = -02.326)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يدل على أن متوسط المجتمع يقل عن 23، أي أن $(\mu < 23)$ وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.01.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$P - value = P(Z \leq -04.80) = F(-04.80) = 01 - F(04.80) = 01 - 01 = 0$$

نلاحظ أن $Pvalue < \alpha \Rightarrow 0 < 0.01$ ، وبالتالي نرفض (H_0).

تمرين06: عينة عشوائية تتكون من16 حالة طوارئ تم اختيارها من واقع سجلات خدمات سيارات الإسعاف بأحد المستشفيات، فوجد أن متوسط الزمن المطلوب لوصول سيارة الإسعاف إلى المكان المطلوب يساوي 13 دقيقة، فإذا فرضنا أن مجتمع الزمن المطلوب لوصول سيارة الإسعاف إلى المكان المطلوب يتوزع وفق التوزيع الطبيعي بانحراف معياري يساوي 03 دقائق.

المطلوب:

1- أكتب فرضية الاختبار؟

2- وعلى ضوء هذه البيانات هل يمكن القول بان متوسط مجتمع الزمن المطلوب لوصول سيارة الإسعاف إلى المكان المطلوب أكبر من 10 دقائق وذلك عند مستوى معنوية (α=0.05)؟

الحل:

1- كتابة فرضية الاختبار

$$H_0 : \mu = 10$$
$$H_1 : \mu > 10$$

-2

إحصاءة الاختبار: لدينا من المعطيات $\overline{X} = 13, \sigma = 03, n = 16, \mu_0 = 10, 1-\alpha = 0.05$ وبالتالي فان إحصاءة الاختبار تكون كالتالي

$$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{13-10}{03/\sqrt{16}} = 04$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد وفي الجانب الأيمن (القيم العليا) و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية

$$Z_{1-\alpha}=Z_{1-0.05}=Z_{0.95}=Z_{tab}=01.645$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $\left(Z_{cal}=04\right)$،هي أكبر من القيمة الجدولية $\left(Z_{tab}=01.645\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يدل على أن متوسط المجتمع أكبر من 10، أي أن $\left(\mu>10\right)$ وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$P-value=P\left(Z\geq 04\right)=01-P\left(Z<04\right)=01-F\left(04\right)=01-01=0$$

نلاحظ أن $Pvalue<\alpha\Rightarrow 0<0.01$ ، وبالتالي نرفض (H_0).

2-2عندما يكون تباين المجتمع $\left(\sigma^2\right)$غيير معلوم، وحجم العينة ($n\geq 30$)

نظريـــة02: إذا أخـذت عينـة عشـوائية حجمهـا ($n\geq 30$)مـن مجتمـع يخضـع للتوزيـع الطبيعـي $X\sim N\left(\mu,\sigma^2\right)$ وكانــت σ^2 غيــر معلومــة فــان إحصــاءة الاختبــار للفرضــية $H_0:\mu=\mu_0$ هـو $Z_{cal}=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $Z_{cal}=\dfrac{\overline{X}-\mu_0}{\hat{S}/\sqrt{n}}$ وهو يخضع للتوزيع الطبيعي المعياري حيث $\overline{X}$ هو الوسط الحسابي للعينة.

تمرين07: إذا أعطت عينة إحصائية حجمها42 وسطا حسابيا قدره 11.50 بانحراف معياري 03.30

المطلوب:

1- أكتب فرضية الاختبار؟

2-أختبر الفرضية القائلة بان $\left(\mu=10\right)$ مقابل الفرضية القائلة بان $\left(\mu\neq 10\right)$ ، علما بان مستوى المعنوية (α=0.05)؟

الحل:

1-كتابة فرضية الاختبار

$$H_0:\mu=15$$
$$H_1:\mu\neq 15$$

2-بما أن المجتمع لا يخضع للتوزيع الطبيعي وحيث أن حجم العينة كبير جدا ($n>30$) وحسب نظرية النهاية المركزية فان توزيع المعاينة سوف يكون له توزيع قريب من التوزيع الطبيعي المعياري.

إحصاءة الاختبار: لدينا من المعطيات $\overline{X}=11.50,S=03.30,n=64,\mu_0=10,1-\alpha=0.05$

وبالتالي فان إحصاءة الاختبار تكون كالتالي $Z_{cal}=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n-1}}=\dfrac{11.50-10}{03.30/\sqrt{42-1}}=02.91$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى

معنوية 0.05 نستخرج قيمة Z الجدولية $Z_{1-\frac{\alpha}{2}}=Z_{0.975}=Z_{tab}=01.960$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة$\left(Z_{cal}=02.91\right)$،هي أكبر من القيمة الجدولية $\left(Z_{tab}=01.960\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يدل على أن متوسط المجتمع أكبر من 10، أي أن $(\mu>10)$ وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p–value) كالتالي:

$$\begin{aligned}P-value&=P\left(Z\leq -02.91\vee Z\geq 02.91\right)\\&=F(-02.91)+01-F(02.91)=01-F(02.91)+01-F(02.91)\\&=02-02F(02.91)=02-02\times 0.99819=0.00362\end{aligned}$$

نلاحظ أن $Pvalue<\alpha\Rightarrow 0.00362<0.05$ ، وبالتالي نرفض (H_0).

2-3 عندما يكون تباين المجتمع $\left(\sigma^2\right)$ غير معلوم وحجم العينة (n<30)

نظرية03: إذا كانت لدينا $\left(X_1,X_2,X_3,.................,X_n\right)$ تمثل عينة عشوائية بحيث (n<30) من مجتمع يخضع للتوزيع الطبيعي بمتوسط حسابي μ وتباين σ^2 مجهول، فان توزيع المعاينة للمتغير العشوائي (أو الإحصاءة) هو $T=\dfrac{\overline{X}-\mu}{S/\sqrt{n-1}}=\dfrac{\overline{X}-\mu}{\hat{S}/\sqrt{n}}$

يتبع توزيع ستونت (t) بدرجة حرية (V=n–1) وبناءا على ذلك إذا كانت مفردات العينة العشوائية تم الحصول عليها من توزيع طبيعي والتباين مجهول وحجم العينة (n<30)، فإن إحصاءة الاختبار المناسبة لإختبار الفرضيات الإحصائية المتعلقة بالمتوسط الحسابي للمجتمع (μ) ستكون مبنية على (T) المعرفة سابقا، وان قيم المنطقة الحرجة لهذا الاختبار يتم الحصول عليها من جدول ستونت (t) بدرجات حرية تساوي(V=n–1)، ومستوى معنوية α.

وباستخدام هذه الإحصاءة يمكن اختبار الفرضيات التالية:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ–إختبار من طرفين: $H_0:\mu=\mu_0$ $H_1:\mu\neq\mu_0$	$T_{cal}=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal}=\dfrac{\overline{X}-\mu_0}{\hat{S}/\sqrt{n}}$	نرفض (H_0) إذا كانت $T_{cal}\geq T_{tab}$ أو $T_{cal}\leq -T_{tab}$ بحيث $T_{tab}=T_{1-\frac{\alpha}{2},n-1}$ أو إذا كانت $Pvalue<\alpha$

ب-إختبار من طرف واحد: $H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$T_{cal} = \frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal} = \frac{\overline{X}-\mu_0}{\hat{S}/\sqrt{n}}$	نرفض (H_0) إذا كانت $T_{cal} \geq T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2},n-1}$ أو إذا كانت $Pvalue < \alpha$
ج- إختبار من طرف واحد: $H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$T_{cal} = \frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal} = \frac{\overline{X}-\mu_0}{\hat{S}/\sqrt{n}}$	نرفض (H_0) إذا كانت $T_{cal} \leq -T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2},n-1}$ أو إذا كانت $Pvalue < \alpha$

تمرين08: تمت صناعة آلة لتعطي 230 غرام من الزيت، وذلك عند وضع قطعة النقود المناسبة، ولمعرفة ما ذا كانت هذه الآلة تعمل بحب المواصفة السابقة، سحبت عينة عشوائية من 06 عبوات لقياس كمية الزيت في كل عبوة، فوجدت النتائج التالية:

رقم العبوة	01	02	03	04	05	06
مقدار الزيت بالغرام	190	180	220	200	160	250

المطلوب:

1-أكتب فرضية الاختبار؟

2-هل يمكن القول بأن متوسط عمل هذه الآلة يختلف عن 230 غرام وذلك عند مستوى معنوية $(\alpha=0.10)$؟

الحل:

1-كتابة فرضية الاختبار

$$H_0: \mu = 230$$
$$H_1: \mu \neq 230$$

-2

بما أن المجتمع يخضع للتوزيع الطبيعي بتباين σ^2 مجهول، وحجم العينة المختارة $(n<30)$، إذن سنستخدم توزيع ستونت بدلا من التوزيع الطبيعي المعياري.

إحصاءة الاختبار:

لدينا من المعطيات $n = 06, \mu_0 = 230, 1-\alpha = 0.10$ وبالتالي فان إحصاءة الاختبار تكون كالتالي

$$T_{cal} = \frac{\overline{X}-\mu_0}{S/\sqrt{n-1}} \quad \text{أو بشكل مكافئ} \quad T_{cal} = \frac{\overline{X}-\mu_0}{\hat{S}/\sqrt{n}}$$

نحسب أولا المتوسط الحسابي للعينة والتباين أو التباين المعدل

$$\overline{X}=\frac{\sum_{i=1}^{n=6}X_i}{n}=\frac{190+180+220+200+160+250}{06}=200$$

$$S^2=\frac{\sum_{i=1}^{n=6}\left(X_1-\overline{X}\right)^2}{n}=\frac{\left(190-200\right)^2+\left(180-200\right)^2+\text{........}+\left(250-200\right)^2}{06}=833.34$$

نحسب الآن الإحصاءة (T) $T_{cal}=\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}=\frac{200-230}{28.87/\sqrt{6-1}}=-02.32$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول توزيع ستودنت وعند مستوى معنوية 0.10

نستخرج قيمة T الجدولية $T_{1-\frac{\alpha}{2},n-1}=T_{0.950,05}=T_{tab}=02.015$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة$\left(\left|T_{cal}\right|=02.32\right)$،هي أكبر من القيمة الجدولية $\left(T_{tab}=02.015\right)$، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، ونؤكد احتمال كبير أن هذه الآلة لا تعمل بحسب المواصفات التي صنعت لأجلها (230غرام)، بل ما تعطيه هذه الآلة يختلف اختلافا حقيقيا (ذا دلالة إحصائية) عن $\left(\mu=230\right)$،وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.10.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$\begin{aligned}P-\text{value}&=P\left(T\le-02.32\vee T\ge02.32\right)\\&=F\left(-02.32\right)+01-F\left(02.32\right)=01-F\left(02.32\right)+01-F\left(02.32\right)\\&=02-02F\left(02.32\right)=02-02\times0.963=0.074\end{aligned}$$

الاحتمال 0.963 المقابل للقيمة F(02.32) لا يوجد في توزيع ستودنت وإنما تم الحصول عليه باستخدام نظرية الحصر التي تم التطرق لها في المحورين الأول والثاني.

نلاحظ أن $0.10 > 0.074 \Rightarrow \text{Pvalue}<\alpha$ ، وبالتالي نرفض (H_0).

تمرين09: تنتج إحدى شركات الأدوية مستحضرا دوائيا جديدا، حيث تحتوي كل مضغوطة منه على 10 ملغ بالمتوسط من المادة الفعالة، قررت وزارة الصحة أنها ستعفي الشركة من الرقابة الصحية (وذلك لتشجيعها على الإنتاج الجديد)، إذا كان هناك ما يؤكد أن متوسط المادة الفعالة في المضغوطة الواحدة تزيد عن 10 ملغ وبشكل جوهري، أما ذا كان متوسط المادة الفعالة أقل أو تساوي 10 ملغ فلن تعفى الشركة من هذه الرقابة.

تم سحب عينة عشوائية بسيطة مؤلفة من 15 مضغوطة من إنتاج المستحضر الجديد فوجد أن متوسط المادة الفعالة يساوي 10.42 ملغ، بانحراف معياري يساوي 01.53ملغ،

المطلوب:

1- أكتب فرضية الاختبار؟

2-هل سيتم تشجيع الشركة على الإنتاج الجديد وإعفائها من الرقابة الصحية، وذلك عند مستوى معنوية $(\alpha=0.05)$؟

الحل:

1-كتابة فرضية الاختبار

$$H_0: \mu = 10$$
$$H_1: \mu > 10$$

-2

إحصاءة الاختبار: لدينا من المعطيات $\overline{X}=10.42, S=01.53, n=15, \mu_0=10, 1-\alpha=0.05$ وبالتالي فان إحصاءة الاختبار تكون كالتالي $T_{cal}=\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal}=\frac{\overline{X}-\mu_0}{\hat{S}/\sqrt{n}}$

$$T_{cal}=\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}=\frac{10.42-10}{01.53/\sqrt{15-1}}=01.027$$

-**القيمة الحرجة**: بما أن الاختبار من جانب واحد فقط ومن اليمين (القيم العليا) و من جدول توزيع ستودنت وعند مستوى معنوية 0.05 نستخرج قيمة T الجدولية

$$T_{1-\alpha,n-1}=T_{0.950,14}=T_{tab}=01.761$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $(T_{cal}=01.027)$،هي أكبر من القيمة الجدولية $(T_{tab}=01.761)$، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1)، وان الفارق بين متوسط العينة ومتوسط المجتمع ما هو إلا فارق ظاهري ليس له دلالة إحصائية ولن تعفى الشركة من الرقابة الصحية وذلك عند مستوى معنوية$(\alpha=0.05)$.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$P-value=P(T\geq 01.027)=01-P(Z<01.027)=01-F(01.027)=01-0.833=0.167$$

نلاحظ أن $Pvalue>\alpha \Rightarrow 0.167>0.05$، وبالتالي نقبل ($H_0$).

تمرين10:إذا كان متوسط الوقت اللازم لإنجاز الطلبة لاختبار معين هو 50 دقيقة، تم تطبيق طريقة جديدة للاختبار باستخدام الحاسبات الحديثة من قبل 12 طالب وكان متوسط وقت الانجاز للعينة هو 42 دقيقة، بانحراف معياري 11.90.

المطلوب:

1- أكتب فرضية الاختبار؟

2-أختبر فيما إذا كان متوسط الانجاز للطريقة الجديدة اقل من 50 دقيقة وذلك عند مستوى معنوية $(\alpha=0.05)$؟

الحل:

1-كتابة فرضية الاختبار

$$H_0 : \mu = 50$$
$$H_1 : \mu < 50$$

-2

إحصاءة الاختبار: لدينا من المعطيات $\overline{X} = 42, S = 11.90, n = 12, \mu_0 = 50, 1-\alpha = 0.05$ وبالتالي فان

إحصاءة الاختبار تكون كالتالي $T_{cal} = \frac{\overline{X} - \mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal} = \frac{\overline{X} - \mu_0}{\hat{S}/\sqrt{n}}$

$$T_{cal} = \frac{\overline{X} - \mu_0}{S/\sqrt{n-1}} = \frac{42-50}{11.90/\sqrt{12-1}} = -02.229$$

-**القيمة الحرجة**: بما أن الاختبار من جانب واحد فقط ومن اليسار (القيم السفلى) و من جدول توزيع ستودنت وعند مستوى معنوية 0.05 نستخرج قيمة T الجدولية

$$-T_{1-\alpha, n-1} = -T_{0.950, 11} = T_{tab} = -01.796$$

-**القرار**: بما أن قيمة إحصاءة الاختبار المحسوبة $(T_{cal} = -02.229)$،هي أصغر من القيمة الجدولية $(T_{tab} = -01.796)$، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1).

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$P-\text{value} = P(T \le -02.23) = F(-02.23) = 01 - F(02.23) = 01 - 0.976 = 0.024$$

نلاحظ أن $\text{Pvalue} < \alpha \Rightarrow 0.024 < 0.05$ ، وبالتالي نرفض(H_0).

03-إختبار الفروض للفرق بين وسطين حسابيين لمجتمعين $(\mu_1 - \mu_2)$ مستقلين

إذا كان لدينا مجتمعان، الأول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 ، والثاني يتوزع توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 ، وأردنا وضع فروض إحصائية حول الفرق بين متوسطي هذين المجتمعين $(\mu_1 - \mu_2)$ ، فان توزيع الفرق سوف يخضع للتوزيع الطبيعي بمتوسط حسابي $\mu_{(\overline{X}_1 - \overline{X}_2)} = \mu_1 - \mu_2$ وتباين $\sigma^2_{(\overline{X}_1 - \overline{X}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ وبالتالي فان المتغير العشوائي (Z) يعطى كالتالي:

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

وعلى افتراض أننا نرغب في إختبار الفرضية العدمية $(H_0 : \mu_1 = \mu_2)$ ضد أي فرضية بديلة أخرى، عندئذ هنالك ثلاث حالات هي :

3-1: عندما يكون التباينين σ_1^2 و σ_2^2 معلومين وحجم العينتين كبير

نظرية04: إذا أخذت عينتان عشوائيتان مستقلتان عن بعضهما البعض من مجتمعين، الأول يخضع للتوزيع الطبيعي بمتوسط حسابي μ_1 مجهول وتباين σ_1^2 معلوم، والمجتمع الثاني يخضع للتوزيع

الطبيعي بمتوسط حسابي μ_2 مجهول وتباين σ_2^2 معلوم، وكان حجم العينة الأولى n_1 كبيرا وحجم العينة الثانية n_2 كبيرا أيضا، فان إحصاءة الاختبار للفرضية $\left(H_0:\mu_1=\mu_2\right)$ هو

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

وبالتالي يمكن استخدام الإحصاءة (Z) كأساس لاختبار الفرضيات التالية:

القرار	إحصاءة الاختبار	الفرضية الإحصائية
نرفض (H_0) إذا كانت $Z_{cal}\geq Z_{tab}$ أو $Z_{cal}\leq -Z_{tab}$ بحيث $Z_{tab}=Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue<\alpha$	$Z_{cal}=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	أ-إختبار من طرفين: $H_0:\mu=\mu_0$ $H_1:\mu\neq\mu_0$
نرفض (H_0) إذا كانت $Z_{cal}\geq Z_{tab}$ بحيث $Z_{tab}=Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue<\alpha$	$Z_{cal}=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	ب-إختبار من طرف واحد: $H_0:\mu=\mu_0$ $H_1:\mu>\mu_0$
نرفض (H_0) إذا كانت $Z_{cal}\leq -Z_{tab}$ بحيث $Z_{tab}=Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue<\alpha$	$Z_{cal}=\frac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	ج- إختبار من طرف واحد: $H_0:\mu=\mu_0$ $H_1:\mu<\mu_0$

تمرين11:أخذت عينة عشوائية بسيطة حجمها 22 من مجتمع أول يخضع للتوزيع الطبيعي بمتوسط حسابي μ_1 و تباين 110، فأعطت متوسط حسابي $\overline{X}_1=83$ ، وأخذت عينة عشوائية ثانية مستقلة عن الأولى وحجمها 27 من مجتمع ثاني يخضع للتوزيع الطبيعي بمتوسط حسابي μ_2 وتباين81 فأعطت متوسط حسابي $\overline{X}_2=69$

المطلوب:

1-أختبر الفرضية $\left(H_0:\mu_1=\mu_2\right)$ مقابل الفرضية $\left(H_1:\mu_1>\mu_2\right)$ وذلك عند مستوى معنوية $(\alpha=0.05)$؟

2-إختبر الفرضية $\left(H_0:\mu_1-\mu_2=10\right)$ مقابل الفرضية $\left(H_1:\mu_1-\mu_2>10\right)$ وذلك عند مستوى معنوية $(\alpha=0.05)$؟

الحل:

1-كتابة الفرضية الإحصائية

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 > \mu_2$$

إحصاءة الاختبار:

لدينا من المعطيات $\overline{X}_1 = 83, \sigma_1^2 = 110, n_1 = 22, \overline{X}_2 = 69, \sigma_2^2 = 81, n_2 = 27, 1-\alpha = 0.05$ وبالتالي فان

إحصاءة الاختبار تكون كالتالي

$$Z_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(83-69)-0}{\sqrt{\frac{110}{22} + \frac{81}{27}}} = 04.95$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد وفي الجانب الأيمن (القيم العليا) و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية

$$Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = Z_{tab} = 01.645$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $(Z_{cal} = 04.95)$،هي أكبر من القيمة الجدولية $(Z_{tab} = 01.645)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، وذلك عند مستوى معنوية 0.05.

2- كتابة الفرضية الإحصائية

$$H_0: \mu_1 - \mu_2 = 10$$
$$H_1: \mu_1 - \mu_2 > 10$$

إحصاءة الاختبار: إحصاءة الاختبار تكون كالتالي

$$Z_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(83-69)-(10)}{\sqrt{\frac{110}{22} + \frac{81}{27}}} = 01.41$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد وفي الجانب الأيمن (القيم العليا) و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية

$$Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = Z_{tab} = 01.645$$

القرار:بما أن قيمة إحصاءة الاختبار المحسوبة $(Z_{cal} = 01.41)$،هي أصغر من القيمة الجدولية $(Z_{tab} = 01.645)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1)، أي انه لا يوجد دلالة بان الفرق بين الوسطين يزيد عن 10 وذلك عند مستوى معنوية 0.05.

تمرين12: للمقارنة بين الأجور اليومية للعمال في شركتين (A وB) في مدينة ما، والتي تخضع الأجور فيها للتوزيع الطبيعي بحيث $X_1 \sim N(\mu_1, 160)$و $X_2 \sim N(\mu_2, 90)$على الترتيب، اختيرت عينة عشوائية من عمال الشركة (A) بحجم (n_1=64)عامل، وتبين أن متوسط أجورهم بلغ 25 ون، واختيرت عينة عشوائية من عمال الشركة (B) بحجم (n_2=36) عامل وتبين أن متوسط أجورهم بلغ 15 ون.

المطلوب:

1-هل يمكن الاستنتاج بأن متوسط الأجور اليومية للعمال في الشركة(A) يختلف عن متوسط الأجور اليومية للعمال في الشركة(B) وذلك عند مستوى معنوية$(\alpha=0.05)$؟

الحل:

الفرضية الإحصائية المراد اختبارها هي
$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

إحصاءة الاختبار:

لدينا من المعطيات $\overline{X}_1 = 25, \sigma_1^2 = 160, n_1 = 64, \overline{X}_2 = 15, \sigma_2^2 = 90, n_2 = 36, 1-\alpha = 0.05$ وبالتالي فان إحصاءة الاختبار تكون كالتالي
$$Z_{cal} = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(25-15)-0}{\sqrt{\frac{160}{64} + \frac{90}{36}}} = 04.47$$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية $Z_{1-\frac{\alpha}{2}} = Z_{0.975} = Z_{tab} = 01.960$

القرار:بما أن قيمة إحصاءة الاختبار المحسوبة$\left(Z_{cal} = 04.47\right)$،هي أكبر من القيمة الجدولية $\left(Z_{tab} = 01.960\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يدل على أن متوسط الأجور في الشركة (A) يختلف عن متوسط الأجور اليومية لعمال الشركة (B)وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:
$$\begin{aligned} P-value &= P\left(Z \le -04.47 \vee Z \ge 04.47\right) \\ &= F(-04.47) + 01 - F(04.47) = 01 - F(04.47) + 01 - F(04.47) \\ &= 02 - 02F(04.47) = 02 - 02 \times 01 = 0 \end{aligned}$$

نلاحظ أن $Pvalue < \alpha \Rightarrow 0 < 0.05$ ، وبالتالي نرفض (H_0).

3-2:عندما يكون التباينين σ_1^2 و σ_2^2 مجهولين وحجم العينتين كبير

نظرية05: بافتراض أن لدينا مجتمعين يتوزعان توزيعا طبيعيا الأول هو $X_1 \sim N\left(\mu_1, \sigma_1^2\right)$ وكلا منهما مجهول، والثاني $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$ وكان أيضا مجهولين، وسحبنا جميع أزواج العينات المستقلة عن بعضهما البعض على التوالي (n_1) و (n_2)، وكان حجم العينتين كبير جدا فان التوزيع الاحتمالي (توزيع المعاينة) للاحصاءة يعطى كالتالي:
$$Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}} \sim N(0,1)$$
أو بشكل مكافئ
$$Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\hat{S}_1}{n_1} + \frac{\hat{S}_2}{n_2}}} \sim N(0,1)$$

وعليه يمكن استخدام هذه الإحصاءة لاختبار الفرضيات التالية:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}}$ أو بشكل مكافئ $Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{S}_1}{n_1} + \frac{\hat{S}_2}{n_2}}}$	نرفض (H_0) إذا كانت $Z_{cal} \geq Z_{tab}$ أو $Z_{cal} \leq -Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$
ب-إختبار من طرف واحد: $H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}}$ أو بشكل مكافئ $Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{S}_1}{n_1} + \frac{\hat{S}_2}{n_2}}}$	نرفض (H_0) إذا كانت $Z_{cal} \geq Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$
ج- إختبار من طرف واحد: $H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}}$ أو بشكل مكافئ $Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{S}_1}{n_1} + \frac{\hat{S}_2}{n_2}}}$	نرفض (H_0) إذا كانت $Z_{cal} \leq -Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$

تمرين13: للمقارنة بين رواتب أعضاء هيئة التدريس في جامعة أ و ب والتي تخضع الرواتب فيها إلى التوزيع الطبيعي $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$ على الترتيب، اختيرت عينة عشوائية من أساتذة جامعة أ بحجم (n_1=50)، وتبين أن متوسط رواتبهم بلغ 100 ون، وتباين قدره 250 ون، واختيرت عينة عشوائية من أساتذة جامعة "ب" بحجم (n_2=60)، وتبين أن متوسط رواتبهم بلغ 120 ون وتباين قدره 360 ون.

المطلوب:

1-أختبر الفرضيات التالية:

1- متوسط رواتب أعضاء هيئة التدريس في جامعة "أ" يختلف عن متوسط رواتب نظرائهم في جامعة "ب" وذلك عند مستوى معنوية $(\alpha=0.05)$؟

2- متوسط رواتب أعضاء هيئة التدريس في جامعة "أ" يقل عن متوسط رواتب نظرائهم في جامعة "ب" وذلك عند مستوى معنوية $(\alpha=0.05)$؟

3- متوسط رواتب أعضاء هيئة التدريس في جامعة "أ" يزيد عن متوسط رواتب نظرائهم في جامعة "ب" وذلك عند مستوى معنوية $(\alpha=0.05)$؟

الحل:

1-الفرضية الإحصائية المطلوب اختيارها هي:
$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

إحصاءة الاختبار:

لدينا من المعطيات $\overline{X}_1 = 100, S_1^2 = 250, n_1 = 50, \overline{X}_2 = 120, S_2^2 = 360, n_2 = 60, 1-\alpha = 0.05$ وبالتالي

فان إحصاءة الاختبار تكون كالتالي
$$Z_{cal} = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}} = \frac{(100-120)-(0)}{\sqrt{\frac{250}{50-1} + \frac{360}{60-1}}} = -05.97$$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية
$$Z_{1-\frac{\alpha}{2}} = Z_{0.975} = Z_{tab} = 01.960$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $\left(Z_{cal} = -05.97\right)$،هي أصغر من القيمة الجدولية $\left(Z_{tab} = -01.960\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يدل على أن متوسط الرواتب في الجامعة "أ"يختلف عن متوسط الرواتب في الجامعة "ب"وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

2- الفرضية الإحصائية المطلوب اختيارها هي:
$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

إحصاءة الاختبار: مما سبق نعلم أن الإحصاءة هي $Z_{cal} = -05.97$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد فقط ومن جهة اليسار (القيم السفلى)، ومن جدول التوزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية
$$-Z_{1-\alpha} = -Z_{0.950} = -Z_{tab} = -01.645$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $\left(Z_{cal} = -05.97\right)$،هي أصغر من القيمة الجدولية $\left(Z_{tab} = -01.645\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يدل

على أن متوسط الرواتب في الجامعة "أ " يقل عن متوسط الرواتب في الجامعة "ب"وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

3- الفرضية الإحصائية المطلوب اختيارها هي:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 > \mu_2$$

<u>إحصاءة الاختبار</u>: مما سبق نعلم أن الإحصاءة هي $Z_{cal} = -05.97$

<u>-القيمة الحرجة:</u> بما أن الاختبار من جانب واحد فقط ومن جهة اليمين (القيم العليا)، ومن جدول التوزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج قيمة Z الجدولية

$$Z_{1-\alpha} = Z_{0.950} = Z_{tab} = 01.645$$

<u>القرار</u>:بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal} = -05.97)$،هي أصغر من القيمة الجدولية $(Z_{tab} = 01.645)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1)، مما يدل على أن متوسط الرواتب في الجامعة "أ " يقل عن متوسط الرواتب في الجامعة "ب"وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

3-3 عندما يكون التباينين σ_1^2 و σ_2^2 مجهولين وحجم العينتين صغير

عندما يكون σ_1^2 و σ_2^2 مجهولين وحجم العينتين صغير فننا نميز حالتين:

الحالة الأولى: عندما يكون التباينان مجهولين ومتساويين $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$

نظرية06: بافتراض أن لدينا مجتمعين يتوزعان توزيعا طبيعيا الأول هو $X_1 \sim N(\mu_1, \sigma_1^2)$ وكلا منهما مجهول، والثاني $X_2 \sim N(\mu_2, \sigma_2^2)$ وكان أيضا مجهولين، وإذا كان معلوم مسبقا بأن $(\sigma_1^2 = \sigma_2^2)$ وسحبنا جميع أزواج العينات المستقلة عن بعضهما البعض على التوالي (n_1) و (n_2)، وكان حجم العينتين صغير فان التوزيع الاحتمالي (توزيع المعاينة) للاحصاءة يعطى كالتالي:

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

والذي يتبع توزيع يطلق عليه توزيع ستودنت (t) وبدرجات حرية تساوي ($V=n_1+n_2-2$) حيث

$$S_p^2 = \frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$$

يمثل التباين المشترك للعينتين، وعليه يمكن استخدام هذه الإحصاءة لاختبار الفرضيات التالية:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	نرفض (H_0) إذا كانت $T_{cal} \geq T_{tab}$ أو $T_{cal} \leq -T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2}, n_1+n_2-02}$ أو إذا كانت $Pvalue < \alpha$
ب-إختبار من طرف واحد: $H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	نرفض (H_0) إذا كانت $T_{cal} \geq T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2}, n_1+n_2-02}$ أو إذا كانت $Pvalue < \alpha$
ج- إختبار من طرف واحد: $H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	نرفض (H_0) إذا كانت $T_{cal} \leq -T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2}, n_1+n_2-02}$ أو إذا كانت $Pvalue < \alpha$

<u>تمرين14:</u> عينتان عشوائيتان من توزيعين مستقلين ويخضعان للتوزيع الطبيعي بحيث $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$ و كانت $(\sigma_1^2 = \sigma_2^2)$ و مجهولتان ووجد أن حجم العينة الأولى هو $(n_1=64)$، وحجم العينة الثانية هو $(n_2=49)$، وكان المتوسط الحسابي للعينة الأولى هو 16 بتباين 09 والمتوسط الحسابي للعينة الثانية هو 20 بتباين 16

<u>المطلوب:</u>

1-إختبر الفرضية $H_0: \mu_1 - \mu_2 = 0$ مقابل الفرضيات التالية وذلك عند مستوى معنوية$(\alpha=0.05)$؟

أ- $H_1: \mu_1 \neq \mu_2$

ب- $H_1: \mu_1 > \mu_2$

ج- $H_1: \mu_1 < \mu_2$

<u>الحل:</u>

أ-الفرضية الإحصائية المطلوب اختبارها هي:

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

إحصاءة الاختبار:

لدينا من المعطيات $\overline{X}_1=16, S_1^2=09, n_1=15, \overline{X}_2=12, S_2^2=16, n_2=10, 1-\alpha=0.05$ وبالتالي فان

إحصاءة الاختبار تكون كالتالي $T_{cal}=\dfrac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$

نحسب أولا تباين العينة المشترك كالتالي: $S_p^2=\dfrac{n_1S_1^2+n_2S_2^2}{n_1+n_2-02}=\dfrac{(15)(09)+(10)(16)}{15+10-02}=12.82$

إذن $T_{cal}=\dfrac{\left(\overline{X}_1-\overline{X}_2\right)-\left(\mu_1-\mu_2\right)}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}=\dfrac{(16-20)-(0)}{(03.58)\sqrt{\dfrac{1}{15}+\dfrac{1}{10}}}=-02.74$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول توزيع ستودنت وعند درجة حرية (23) نجد

$$T_{1-\frac{\alpha}{2},n_1+n_2-02}=T_{1-\frac{0.05}{2},15+10-02}=T_{0.975,23}=02.069$$

القرار:بما أن قيمة إحصاءة الاختبار المحسوبة$\left(T_{cal}=-02.74\right)$،هي أصغر من القيمة الجدولية $\left(T_{tab}=-02.069\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1).

ب- الفرضية الإحصائية المطلوب اختيارها هي:

$$H_0:\mu_1=\mu_2$$
$$H_1:\mu_1>\mu_2$$

إحصاءة الاختبار: لدينا مما سبق $T_{cal}=-02.74$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد فقط ومن جهة اليمين ومن من جدول توزيع ستودنت وعند درجة حرية (23) نجد $T_{1-\alpha,n_1+n_2-02}=T_{1-0.05,15+10-02}=T_{0.950,23}=01.714$

-**القرار**:بما أن قيمة إحصاءة الاختبار المحسوبة$\left(T_{cal}=-02.74\right)$،هي أصغر من القيمة الجدولية $\left(T_{tab}=01.714\right)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1).

ج- الفرضية الإحصائية المطلوب اختيارها هي:

$$H_0:\mu_1=\mu_2$$
$$H_1:\mu_1<\mu_2$$

إحصاءة الاختبار: لدينا مما سبق $T_{cal}=-02.74$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد فقط ومن جهة اليسار ومن من جدول توزيع ستودنت وعند درجة حرية (23) نجد $-T_{1-\alpha,n_1+n_2-02}=-T_{1-0.05,15+10-02}=-T_{0.950,23}=-01.714$

-**القرار**:بما أن قيمة إحصاءة الاختبار المحسوبة$\left(T_{cal}=-02.74\right)$،هي أصغر من القيمة الجدولية $\left(T_{tab}=-02.069\right)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1).

الحالة الثانية: عندما يكون التباينان مجهولين و غير متساويين $\left(\sigma_1^2 \neq \sigma_2^2\right)$

نظرية07: بافتراض أن لدينا مجتمعين يتوزعان توزيعا طبيعيا الأول هو $X_1 \sim N\left(\mu_1, \sigma_1^2\right)$ وكلا منهما مجهول، والثاني $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$ وكان أيضا مجهولين، ففي هذه نقدر تباين كل مجتمع بتباين العينة المسحوبة منه، أي نقدر تباين المجتمع الأول σ_1^2 بتباين العينة المسحوبة منه S_1^2 ، ونقدر تباين المجتمع الثاني σ_2^2 بتباين العينة المسحوبة منه S_2^2 ، نتحصل على إحصاءة الاختبار التالية:

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

وهذا التوزيع قريب من التوزيع الاحتمالي لستودنت بدرجة حرية (V) تعطى كالتالي:

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

وسيكون القرار لكل فرضية من الفرضيات السابقة كما سبق الإشارة إليها في الجدول السابق مع مراعاة درجة الحرية فقط.

تمرين15: بفرض أن البيانات التالية تم الحصول عليها من عينتين عشوائيتين مستقلتين

المجتمع الأول: $\bar{X}_1 = 24.20, S_1^2 = 10, n_1 = 15$ **المجتمع الثاني:** $\bar{X}_2 = 23.90, S_2^2 = 20, n_2 = 10$

وعلى افتراض أن $\left(\sigma_1^2 \neq \sigma_2^2\right)$ والمجتمعين طبيعيين

المطلوب:

1-أختبر الفرضية التالية $H_0: \mu_1 = \mu_2$ ، $H_1: \mu_1 \neq \mu_2$ وذلك عند مستوى معنوية $(\alpha=0.01)$؟

الحل:

إختبار الفرضية التالية عند مستوى معنوية 0.01

- **إحصاءة الاختبار:** حيث أن تباين المجتمعين غير متساويين، وعليه فإن إحصاءة الاختبار تكون كالآتي

$$T_{cal} = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(24.20 - 23.90) - (0)}{\sqrt{\frac{10}{15} + \frac{20}{10}}} = 0.1873$$

- **القيمة الحرجة:** بما أن الاختبار من جانبين و يخضع لتوزيع ستودنت، نستخرج أولا درجة الحرية

كالتالي
$$V=\frac{\left[\left(S_1^2/n_1\right)+\left(S_2^2/n_2\right)\right]^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1-1}+\frac{\left(S_2^2/n_2\right)^2}{n_2-1}}=\frac{\left[\left(10/15\right)+\left(20/10\right)\right]^2}{\frac{\left(10/15\right)^2}{15-1}+\frac{\left(20/10\right)^2}{10-1}}\simeq 15$$

ويتم استخراج قيمة ستودنت الجدولية عند درجة حرية 15 كما يلي:
$$T_{1-\frac{\alpha}{2},V}=T_{1-\frac{0.01}{2},15}=T_{0.995,15}=T_{tab}=02.947$$

-القرار:بما أن قيمة إحصاءة الاختبار المحسوبة $(T_{cal}=0.1873)$،هي أصغر من القيمة الجدولية $(T_{tab}=02.947)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1).

04-إختبار الفروض للفرق بين وسطين حسابيين لمجتمعين $(\mu_1-\mu_2)$ مرتبطين

نظرية08: نفرض أن $\overline{X}\sim N(\mu_1,\sigma_1^2)$ و $\overline{Y}\sim N(\mu_2,\sigma_2^2)$ وكان المجتمعان مرتبطان وسحبت من كل مجتمع عينة حجمها (n) بحيث عينة المجتمع الأول هي $\{X_1,X_2,X_3.................X_n\}$ وعينة المجتمع الثاني هي $\{Y_1,Y_2,Y_3.....................Y_n\}$، بحيث بحيث تمثل X_i و Y_i القيمتين المتناظرتين في العينتين،إذا كانت الفروق بين قيم العينتين المتناظرتين هي $\{d_1,d_2,d_3................d_n\}$ حيث $d_i=X_i-Y_i$ لجميع قيم n............i=1,2,3، فان $\{d_1,d_2,d_3................d_n\}$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على أنها عينة عشوائية من مجتمع طبيعي وسطه الحسابي $\mu_d=\mu_1-\mu_2$ وتباينه σ_d^2، ولنفرض ان وسط العينة هو $\bar{d}$ وتباينها هو S_d^2 ، وإذا رغبنا في اختبار تساوي وسطي المجتمعين امام الفرضية التالية $H_0:\mu_1=\mu_2$ كما يمكن صياغتها بطريقة أخرى هي $H_0:\mu_d=0$ وتاخذ الفرضية البديلة احدى الصور الثلاث التالية:

$$\begin{cases}H_1:\mu_1\neq\mu_2\\H_1:\mu_1>\mu_2\\H_1:\mu_1<\mu_2\end{cases}\Rightarrow\begin{cases}H_1:\mu_d\neq 0\\H_1:\mu_d>0\\H_1:\mu_d<0\end{cases}$$

ونميز الحالتين التاليتين:

الحالة الاولى: حجم العينة $n\geq 30$ و σ_d^2 مجهول

في هذه الحالة فان إحصاءة الاختبار هي
$$Z_{cal}=\frac{\bar{d}-\mu_d}{\sigma_{\bar{d}}}=\frac{\bar{d}-\mu_d}{\frac{S_d}{\sqrt{n-1}}}=\frac{\bar{d}-\mu_d}{\frac{\hat{S}_d}{\sqrt{n}}}\sim N(0,1)$$

بحيث
$$E(\bar{d})=\mu_{\bar{d}}=\frac{\sum_{i=1}^{n}d_i}{n}$$
$$V(\bar{d})=\sigma_{\bar{d}}^2=\frac{\sigma^2}{n}\simeq\frac{S_{\bar{d}}^2}{n-1}=\frac{\hat{S}_{\bar{d}}^2}{n}$$

$$S_d^2=\frac{\sum_{i=1}^{n}\left(d_i-\bar{d}\right)^2}{n}=\frac{\sum_{i=1}^{n}d_i^2}{n}-\left(\bar{d}\right)^2$$

$$\hat{S}_d^2=\frac{\sum_{i=1}^{n}\left(d_i-\bar{d}\right)^2}{n-1}=\frac{1}{n-1}\left(\sum_{i=1}^{n}d_i^2-\frac{\left(\sum_{i=1}^{n}d_i\right)^2}{n}\right)$$

وسيكون القرار في حالة إختبار من طرفين او من طرف واحد كما تم التطرق له في حالة اختبار المتوسط الحسابي للمجتمع (μ) في الفقرة 2

الحالة الثانية: حجم العينة $n<30$ و σ_d^2 مجهول

في هذه الحالة فان إحصاءة الاختبار هي

$$T_{cal}=\frac{\bar{d}-\mu_d}{\frac{S_d}{\sqrt{n-1}}}=\frac{\bar{d}-\mu_d}{\frac{\hat{S}_d}{\sqrt{n}}}$$

وسيكون القرار في حالة الاختبار من طرفين او من طرف واحد كما تم التطرق له في حالة إختبار المتوسط الحسابي للمجتمع (μ) في الفقرة 2-3

<u>تمرين16:</u> إذا كان $\bar{X}\sim N\left(\mu_1,\sigma_1^2\right)$ و $\bar{Y}\sim N\left(\mu_2,\sigma_2^2\right)$ وكان المجتمعان مرتبطان وسحبت من كل مجتمع عينة حجمها 30 وكان مجموع الفروق بين القيم المتناظرة للعينتين هو 180 والانحراف المعياري للفروق هو 09

<u>المطلوب:</u>

1-هل يمكن القول أن المجتمعين لهما نفس المتوسط عند مستوى معنوية(α=0.05)؟

<u>الحل:</u>

1- الفرضية الإحصائية المطلوب اختيارها هي:

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

- **<u>إحصاءة الاختبار</u>**: بما أن حجم العينة هو 30 نستخدم التوزيع الطبيعي المعياري، وبما أن الاختبار من جانبين و حسب المعطيات لدينا $n_1=n_2=n=30,\sum_{i=1}^{30}d_i=180,S_d=09,1-\alpha=0.05$

إذن إحصاءة الاختبار هي

$$Z_{cal}=\frac{\bar{d}-\mu_d}{\sigma_{\bar{d}}}=\frac{\bar{d}-\mu_d}{\frac{S_{\bar{d}}}{\sqrt{n-1}}}=\frac{\bar{d}-\mu_d}{\frac{\hat{S}_{\bar{d}}}{\sqrt{n}}}\sim N(0,1)$$

نستخرج أولا قيمة $\bar{d}$ كالتالي

$$\bar{d}=\frac{\sum_{i=1}^{n}d_i}{n}\Rightarrow\bar{d}=\frac{180}{30}=06$$

إذن قيمة Z المحسوبة هي $Z_{cal}=\dfrac{\bar{d}-\mu_d}{\dfrac{S_d}{\sqrt{n-1}}}=\dfrac{06-0}{\dfrac{09}{\sqrt{30-1}}}=03.59$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول توزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج القيمة الحرجة كالتالي $Z_{tab}=Z_{1-\frac{\alpha}{2}}=Z_{1-\frac{0.05}{2}}=Z_{0.975}=01.960$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal}=03.59)$،هي أكبر من القيمة الجدولية $(Z_{tab}=01.960)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، وبالتالي نعبر انه يوجد فرق معنوي بين متوسطي المجتمعين وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$\begin{aligned}P-\text{value}&=P\left(Z\le-03.59\vee Z\ge03.59\right)\\&=F(-03.59)+01-F(03.59)=01-F(03.59)+01-F(03.59)\\&=02-02F(03.59)=02-02\times0.99983=0.00034\end{aligned}$$

نلاحظ أن $Pvalue<\alpha\Rightarrow0.00034<0.05$ ، وبالتالي نرفض (H_0).

تمرين17: عشر أشخاص يربون في تخفيض أوزانهم، اشتركوا في نظام غذائي للمساعدة في تخفيض الوزن فكانت أوزانهم قبل وبعد الاشتراك في النظام كالآتي:

رقم الشخص	01	02	03	04	05	06	07	08	09	10
قبل الاشتراك	62	53	39	37	92	53	68	67	47	82
بعد الاشتراك	60	50	40	35	90	55	70	65	45	80

المطلوب:

1-هل أدى الاشتراك في النظام الغذائي إلى تخفيض متوسط الوزن؟ وذلك عند مستوى معنوية$(\alpha=0.05)$؟

الحل:

1- الفرضية الإحصائية المطلوب اختيارها هي: $\begin{cases}H_0:\mu_1=\mu_2\\H_1:\mu_1<\mu_2\end{cases}\Rightarrow\begin{cases}H_0:\mu_d=0\\H_1:\mu_d<0\end{cases}$

- **إحصاءة الاختبار**: بما أن حجم العينة هو اقل من 30 وتباين المجتمع مجهول نستخدم توزيع ستودنت بدرجة حرية (V=n-1) وتعطى إحصاءة الاختبار كالتالي $T_{cal}=\dfrac{\bar{d}-\mu_d}{\dfrac{S_d}{\sqrt{n-1}}}=\dfrac{\bar{d}-\mu_d}{\dfrac{\hat{S}_d}{\sqrt{n}}}$

لحساب احصاءة الاختبار نحسب أولا $\bar{d}$ و S_d^2

الأشخاص	X_i	Y_i	$d_i = X_i - Y_i$	d_i^2	$\left(d_i - \bar{d}\right)^2$
01	62	60	+02	04	01
02	53	50	+03	09	04
03	39	40	−01	01	04
04	37	35	+02	04	01
05	92	90	+02	04	01
06	53	55	−02	04	09
07	68	70	−02	04	09
08	67	65	+02	04	01
09	47	45	+02	04	01
10	82	80	+02	04	01
المجموع	------	------	**+10**	**42**	**32**

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} \Rightarrow \bar{d} = \frac{10}{10} = 01$$

$$S_d^2 = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n} = \frac{\sum_{i=1}^{n} d_i^2}{n} - \left(\bar{d}\right)^2 \Rightarrow S_d^2 = \frac{42}{10} - (01)^2 = 03.20$$

$\vee$

$$\hat{S}_d^2 = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1} = \frac{1}{n-1}\left(\sum_{i=1}^{n} d_i^2 - \frac{\left(\sum_{i=1}^{n} d_i\right)^2}{n}\right) \Rightarrow \hat{S}_d^2 = \frac{1}{10-1}\left(42 - \frac{(10)^2}{10}\right) = 03.55$$

إذن قيمة ستونت المحسوبة هي

$$T_{cal} = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n-1}}} = \frac{\bar{d} - \mu_d}{\frac{\hat{S}_d}{\sqrt{n}}} \Rightarrow T_{cal} = \frac{01 - 0}{\frac{01.78}{\sqrt{10-01}}} = \frac{01 - 0}{\frac{01.88}{\sqrt{10}}} = 01.68$$

–<u>القيمة الحرجة:</u> بما أن الاختبار من جانب واحد فقط ومن جهة اليسار ومن من جدول توزيع ستودنت وعند درجة حرية (09) نجد $-T_{1-\alpha, n-1} = -T_{1-0.05, 10-1} = -T_{0.950, 09} = -01.833$

-القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $(T_{cal}=01.68)$، هي أصغر من القيمة الجدولية $(T_{tab}=-01.833)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1)، وبالتالي فان النظام الغذائي لم يؤدي إلى تخفيض متوسط الوزن وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

5-إختبار الفروض حول النسبة (P)

نظرية09: إذا كان لدينا (X) متغير عشوائي متقطع يمثل عدد حالات النجاح في (n) من المحاولات المستقلة باحتمال نجاح قدره(P)، أي أن المتغير العشوائي يخضع لتوزيع ذي الحدين بالمعلمتين $X\sim Ben(np,npq)$، وفي حالات كون عدد المحاولات (n) كبيرا جدا، وان احتمال نجاح المحاولة(P) أو فشلها (q) ليس صغيرين جدا، ففي هذه الحالة إن توزيع ثنائي الحدين يقترب من التوزيع الطبيعي المعياري (Z)عندما تكون (n) كبيرة، وفقا إلى نظرية النهاية المركزية$(np\geq 05, nq\geq 05)$ فإن توزيع المعاينة للإحصاءة (Z) تعطى كما يلي:

$$Z=\frac{X-np}{\sqrt{np(1-p)}}=\frac{\frac{X}{n}-P}{\sqrt{P(1-P)/n}}=\frac{\hat{P}-P}{\sqrt{P(1-P)/n}}\sim N(0,1)$$

يتبع التوزيع الطبيعي المعياري ، وبالتالي يمكن استخدام هذه الاحصاءة لإختبار الفرضيات الخاصة بنسبة عناصر المجتمع الإحصائي التي تحمل صفة معينة وذلك كما يلي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0: P=P_0$ $H_1: P\neq P_0$	$Z_{cal}=\frac{\hat{P}-P_0}{\sqrt{P_0(1-P_0)/n}}$	نرفض (H_0) إذا كانت $Z_{cal}\geq Z_{tab}$ أو $Z_{cal}\leq -Z_{tab}$ بحيث $Z_{tab}=Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue<\alpha$
ب-إختبار من طرف واحد: $H_0: P=P_0$ $H_1: P>P_0$	$Z_{cal}=\frac{\hat{P}-P_0}{\sqrt{P_0(1-P_0)/n}}$	نرفض (H_0) إذا كانت $Z_{cal}\geq Z_{tab}$ بحيث $Z_{tab}=Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue<\alpha$
ج- إختبار من طرف واحد: $H_0: P=P_0$ $H_1: P<P_0$	$Z_{cal}=\frac{\hat{P}-P_0}{\sqrt{P_0(1-P_0)/n}}$	نرفض (H_0) إذا كانت $Z_{cal}\leq -Z_{tab}$ بحيث $Z_{tab}=Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue<\alpha$

تمرين18: أدعت إحدى شركات إنتاج الأجهزة الالكترونية أن نسبة الأجهزة التالفة في إنتاجها بلغت 04%، قام احد الوكلاء المعتمدين لهذه الشركة باختيار عينة من الأجهزة الالكترونية المجهزة من قبل الشركة قوامها 500 جهاز وتم فحصها، فتبين للوكيل أن عدد الأجهزة التالفة كان 45 جهازا.

المطلوب:

1-هل يمكن الاستنتاج بان الشركة صادقة في إدعائها مع الوكيل، وذلك عند مستوى معنوية ($\alpha=0.05$)؟

الحل:

1- الفرضية الإحصائية المطلوب اختبارها هي:

$$H_0 : P = 0.04$$
$$H_1 : P > 0.04$$

- **إحصاءة الاختبار**: بما أن حجم العينة هو أكبر من 30 سوف نستخدم إحصاءة التوزيع الطبيعي المعياري التالية:

$$Z_{cal} = \frac{\hat{P} - P_0}{\sqrt{P_0(1-P_0)/n}}$$

لدينا من المعطيات $P_0 = 0.04, n = 500, X = 45, 1-\alpha = 0.05$

لحساب الإحصاءة نحسب أولا نسبة عدد الأجهزة التالفة في العينة $\hat{P}$

$$\hat{P} = \frac{X}{n} = \frac{45}{500} = 0.09$$

إذن

$$Z_{cal} = \frac{\hat{P} - P_0}{\sqrt{P_0(1-P_0)/n}} = \frac{0.09 - 0.04}{\sqrt{0.04(1-0.04)/500}} = 05.70$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد فقط ومن جهة اليمين و من جدول توزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج القيمة الحرجة كالتالي

$$Z_{tab} = Z_{1-\alpha} = Z_{1-0.05} = Z_{0.950} = 01.645$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $(Z_{cal} = 05.70)$،هي أكبر من القيمة الجدولية $(Z_{tab} = 01.645)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، مما يعني أن الشركة غير صادقة في إدعائها مع الوكيل وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

تمرين19: إذا علمت أن 15% من الأجهزة المرئية التي تنتجها شركة الإلكترونيات يتم إرجاعها بسبب وجود خلل بها، وفي محاولة للتغلب على هذه المشكلة قامت الشركة بتعديل في بعض قطع الغيار اللازمة للإنتاج ثم قامت بتسويق عينة عشوائية من 250 جهاز حيث تم إرجاع 30 منها للشركة وذلك بسبب خلل بها.

المطلوب:

1-هل يمكن القول بأن هناك تحسن في إنتاج الشركة عند مستوى معنوية ($\alpha=0.01$)؟

الحل:

1- الفرضية الإحصائية المطلوب اختيارها هي:

$$H_0: P=0.15$$
$$H_1: P<0.15$$

- **إحصاءة الاختبار**: بما أن حجم العينة هو أكبر من 30 سوف نستخدم إحصاءة التوزيع الطبيعي المعياري التالية:

$$Z_{cal}=\frac{\hat{P}-P_0}{\sqrt{P_0(1-P_0)/n}}$$

لدينا من المعطيات $P_0=0.15, n=250, X=30, 1-\alpha=0.01$

لحساب الإحصاءة نحسب أولا نسبة عدد الأجهزة التالفة في العينة $\hat{P}$

$$\hat{P}=\frac{X}{n}=\frac{30}{250}=0.12$$

إذن

$$Z_{cal}=\frac{\hat{P}-P_0}{\sqrt{P_0(1-P_0)/n}}=\frac{0.12-0.15}{\sqrt{0.15(1-0.15)/250}}=-01.33$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد فقط ومن جهة اليسار و من جدول توزيع الطبيعي المعياري وعند مستوى معنوية 0.01 نستخرج القيمة الحرجة كالتالي

$$-Z_{tab}=-Z_{1-\alpha}=-Z_{1-0.01}=-Z_{0.990}=-02.326$$

القرار:بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal}=-01.33)$،هي أصغر من القيمة الجدولية $(Z_{tab}=-02.326)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1)، مما يعني أنه لم يطرأ تحسن على إنتاج الشركة وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

6-إختبار الفرق بين نسبتي مجتمعين (P_1-P_2)

نظرية10: إذا كانت لدينا $(X_1,X_2,X_3 \ldots\ldots X_{n_1})$ تمثل عينة عشوائية من مجتمع ذي الحدين بالمعلمتين $X_1 \sim Ben(n_1P_1, n_1p_1q_1)$ حيث P_1 تمثل نسبة عناصر المجتمع التي تحمل الصفة قيد الدراسة، وكانت X_1 تمثل عدد هذه العناصر، وكانت لدينا $(X_1,X_2,X_3 \ldots\ldots X_{n_2})$عينة عشوائية من مجتمع ذي الحدين بالمعلمتين$X_2 \sim Ben(n_2P_2, n_2P_2q_2)$ حيث P_2 تمثل نسبة عناصر هذا المجتمع التي تحمل الصفة قيد الدراسة، وكانت X_2 تمثل عدد هذه العناصر، وكان حجم العينتين كبيرا بحيث يمكن استخدام التقريب الطبيعي لتوزيع ذي الحدين وكانت العينتين مستقلتين عن بعضهما البعض، فإنه من نظرية النهاية المركزية فان توزيع المعاينة للاحصاءة يعطى بالعلاقة التالية:

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

حيث $\hat{P}_1=\frac{X_1}{n_1}$ و $\hat{P}_2=\frac{X_2}{n_2}$ ، وعليه يمكن استخدام هذه الإحصاءة كأساس لاختبار الفرضية

الإحصائية $H_0: P_1 = P_2$ مقابل جميع البدائل الممكنة، ولكن قبل أن نوضح كيفية استخدام هذه الإحصاءة في الاختبار تجدر الإشارة هنا إلى انه إذا كانت $P_1 = P_2 = P$ فان $\frac{X_1}{n_1}$ و $\frac{X_2}{n_2}$ سيكون لكلاهما تقدير بقيمة واحدة لنفس المعلمة، وبالتالي يمكننا تجميع هذه التقديرات من خلال النظر إلى مجموع حالات النجاح وقسمته على إجمالي عدد المحاولات،أي أننا نستخدم التقدير $\frac{X_1 + X_2}{n_1 + n_2}$ عوضا عن $\frac{X_1}{n_1}$ و $\frac{X_2}{n_2}$ وسوف نطلق على هذا التقدير تسمية التقدير المشترك ونرمز له بالرمز $\hat{P}$ فإذا كان $P_1 = P_2 = P$ فإن:

$$\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}} = \sqrt{\frac{P(1-P)}{n_1} + \frac{P(1-P)}{n_2}}$$

والتقدير بقيمة واحدة لهذا المقدر هو

$$\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n_1} + \frac{\hat{P}\left(1-\hat{P}\right)}{n_2}}$$

وعليه يمكن استبدال المقام في الإحصاءة السابقة بهذه القيمة لتصبح كما يلي:

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

وستكون الفرضيات وإحصاءة الاختبار لكل منها كما يلي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0: P_1 = P_2$ ، $H_1: P_1 \neq P_2$	$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n_1} + \frac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}$	نرفض (H_0) إذا كانت $Z_{cal} \le -Z_{tab}$ أو $Z_{cal} \ge Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$
ب-إختبار من طرف واحد: $H_0: P_1 = P_2$ ، $H_1: P_1 > P_2$	$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n_1} + \frac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}$	نرفض (H_0) إذا كانت $Z_{cal} \ge Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$

ج- إختبار من طرف واحد: $H_0: P_1 = P_2$ $H_1: P_1 < P_2$	$Z_{cal} = \dfrac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\dfrac{\hat{P}\left(1-\hat{P}\right)}{n_1} + \dfrac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}$	نرفض (H_0) إذا كانت $Z_{cal} \le -Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$

تمرين20: في دراسة للمقارنة مابين نسبتي السائقين الذين يستخدمون حزام الأمان عند القيادة في سيارات الركوب الخاصة والركوب العامة اختيرت عينة عشوائية تتكون من 400 سائق سيارة خاصة فوجد من بينهم 240 يستخدمون حزام الأمان عند القيادة، ومن عينة عشوائية تتكون من 200 سائق سيارة ركوب عامة وجد من بينهم 135 يستخدمون حزام الأمان عند القيادة.

المطلوب:

1- هل يمكن القول بان هناك اختلاف بين نسبة السائقين الذين يستخدمون حزام الأمان عند القيادة في نوعي السيارات عند مستوى معنوية (α=0.05)؟

الحل:

1- الفرضية الإحصائية المطلوب اختيارها هي:

$$H_0: P_1 = P_2$$
$$H_1: P_1 \neq P_2$$

- **إحصاءة الاختبار**: بما أن حجم العينة هو أكبر من 30 سوف نستخدم إحصاءة التوزيع الطبيعي المعياري التالية:

$$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\dfrac{\hat{P}\left(1-\hat{P}\right)}{n_1} + \dfrac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}$$

لحساب الإحصاءة نحسب القيم التالية أولا:

$$\hat{P}_1 = \frac{X_1}{n_2} = \frac{240}{400} = 0.40$$

$$\hat{P}_2 = \frac{X_2}{n_2} = \frac{135}{200} = 0.675$$

$$\hat{P} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{240+135}{400+200} = 0.625$$

إذن

$$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\dfrac{\hat{P}\left(1-\hat{P}\right)}{n_1} + \dfrac{\hat{P}\left(1-\hat{P}\right)}{n_2}}} = \frac{(0.600 - 0.675) - 0}{\sqrt{\dfrac{0.625(1-0.625)}{400} + \dfrac{0.625(1-0.625)}{200}}} = -01.79$$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول توزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج القيمة الحرجة كالتالي $Z_{tab}=Z_{1-\frac{\alpha}{2}}=Z_{1-\frac{0.05}{2}}=Z_{0.975}=01.960$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal}=-01.79)$،هي أكبر من القيمة الجدولية $(Z_{tab}=01.960)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، أي انه توجد فروق معنوية ما بين نسبتي السائقين الذين يستخدمون حزام الأمان عند القيادة وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

تمرين21: اختيرت عينة عشوائية قوامها 700 طالب، بواقع 360 طالبا و340 طالبة، وقد لوحظ أن عدد الطلاب المدمنين على التدخين كان 50 طالبا، وعدد الطالبات المدمنات على التدخين كان 30 طالبة.

المطلوب:

1-هل يمكن الاستنتاج أن نسبة الطلاب المدمنين على التدخين هي أعلى من نسبة الطالبات المدمنات على التدخين، وذلك عند مستوى معنوية (α=0.05)؟

الحل:

1- الفرضية الإحصائية المطلوب اختيارها هي:

$$H_0: P_1=P_2$$
$$H_1: P_1>P_2$$

حيث أن:

P_1: نسبة الطلاب المدمنين على التدخين.

P_2: نسبة الطالبات المدمنات على التدخين.

- **إحصاءة الاختبار**: بما أن حجم العينة هو أكبر من 30 سوف نستخدم إحصاءة التوزيع الطبيعي المعياري التالية:

$$Z_{cal}=\frac{\left(\hat{P}_1-\hat{P}_2\right)-\left(P_1-P_2\right)}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n_1}+\frac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}$$

لحساب الإحصاءة نحسب القيم التالية أولا:

$$\hat{P}_1=\frac{X_1}{n_2}=\frac{50}{360}\simeq 0.14$$
$$\hat{P}_2=\frac{X_2}{n_2}=\frac{30}{340}\simeq 0.090$$
$$\hat{P}=\frac{X_1+X_2}{n_1+n_2}=\frac{50+30}{360+340}=0.114$$

إذن
$$Z_{cal}=\frac{\left(\hat{P}_1-\hat{P}_2\right)-\left(P_1-P_2\right)}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n_1}+\frac{\hat{P}\left(1-\hat{P}\right)}{n_2}}}=\frac{(0.14-0.09)-0}{\sqrt{\frac{0.114(1-0.114)}{360}+\frac{0.114(1-0.114)}{340}}}=02.080$$

-**القيمة الحرجة:** بما أن الاختبار من جانب واحد(من اليمين) و من جدول توزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج القيمة الحرجة كالتالي

$$Z_{tab}=Z_{1-\alpha}=Z_{1-0.05}=Z_{0.950}=01.645$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal}=02.080)$،هي أكبر من القيمة الجدولية $(Z_{tab}=01.645)$ ، وهذا يعني رفض الفرضية العدمية(H_0) وقبول الفرضية البديلة (H_1)، أي انه توجد فروق معنوية ما بين نسبتي المدخنين وبالتالي نستنتج أن نسبة الطلاب المدمنين على التدخين أكبر من نسبة الطالبات المدمنات على التدخين وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

07-اختبار الفروض حول التباين σ^2

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا تباينه σ^2 مجهول، وأردنا إجراء اختبارات حول المعلمة المجهولة σ^2، فإن إحصائية الاختبار المناسبة في هذه الحالة هي الإحصائية التي تعتمد على أفضل مقدر بالقيمة للمعلمة المجهولة، وهو تباين العينة المعدل $\hat{S}^2$ أو تباين العينة S^2 ، وهذه الإحصائية هي المتغير العشوائي لتوزيع كاي-المربع (χ^2) .

نظرية 11: إذا كان لدينا $(X_1,X_2,X_3,\ldots\ldots X_n)$ تمثل مشاهدات عينة عشوائية حجمها (n) مشاهدة تم اختيارها من مجتمع يخضع للتوزيع الطبيعي بتباين σ^2 غير معلوم، ولأغراض التطبيق يتم تقدير تباين العينة S^2 على أساس مشاهدات العينة، فان توزيع المعاينة للاحصاءة $\chi^2=\frac{(n-1)\hat{S}^2}{\sigma_0^2}$ أو بشكل مكافئ $\chi^2=\frac{(n)S^2}{\sigma_0^2}$ يتبع توزيع مربع-كآي بدرجة حرية (n-1)، والجدول التالي يوضح الفرضيات وإحصاءة الاختبار لكل منها كما يلي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0:\sigma^2=\sigma_0^2$ $H_1:\sigma^2\neq\sigma_0^2$	$\chi_{cal}^2=\frac{(n-1)\hat{S}^2}{\sigma_0^2}$ أو بشكل مكافئ $\chi_{cal}^2=\frac{(n)S^2}{\sigma_0^2}$	نرفض (H_0) إذا كانت $\chi_{cal}^2>\chi_{1-\frac{\alpha}{2},n-1}^2$ و $\chi_{cal}^2<\chi_{\frac{\alpha}{2},n-1}^2$ بحيث $\chi_{tab}^2=\chi_{\frac{\alpha}{2},n-1}^2$ و $\chi_{tab}^2=\chi_{1-\frac{\alpha}{2},n-1}^2$ أو إذا كانت $Pvalue<\alpha$

ب-إختبار من طرف واحد: $H_0:\sigma^2=\sigma_0^2$ $H_1:\sigma^2>\sigma_0^2$	$\chi^2_{cal}=\frac{(n-1)\hat{S}^2}{\sigma_0^2}$ أو بشكل مكافئ $\chi^2_{cal}=\frac{(n)S^2}{\sigma_0^2}$	نرفض (H_0) إذا كانت $\chi^2_{cal}>\chi^2_{1-\alpha,n-1}$ بحيث $\chi^2_{tab}=\chi^2_{1-\alpha,n-1}$ أو إذا كانت $Pvalue<\alpha$
ج- إختبار من طرف واحد: $H_0:\sigma^2=\sigma_0^2$ $H_1:\sigma^2<\sigma_0^2$	$\chi^2_{cal}=\frac{(n-1)\hat{S}^2}{\sigma_0^2}$ أو بشكل مكافئ $\chi^2_{cal}=\frac{(n)S^2}{\sigma_0^2}$	نرفض (H_0) إذا كانت $\chi^2_{cal}<\chi^2_{\alpha,n-1}$ بحيث $\chi^2_{tab}=\chi^2_{\alpha,n-1}$ أو إذا كانت $Pvalue<\alpha$

تمرين22: لوحظ أن الانحراف المعياري لعمر نوع من المصابيح المنتجة (تخضع للتوزيع الطبيعي) في أحد المصانع بلغ 30 ساعة، وبعد فترة من الزمن طلب مدير الإنتاج التأكد من أن تشتت عمر المصابيح المنتجة لم يتأثر بالرغم من تقادم أعمار المكائن المستخدمة، ولغرض تنفيذ طلب مدير الإنتاج اختيرت عينة عشوائية من الإنتاج قوامها 31 مصباح، وبعد إجراء فحص المصابيح تبين أن الانحراف المعياري بلغ 35 ساعة.

المطلوب:

1-هل يمكن الاستنتاج بان إنتاج المصنع سيبقى بنفس مستوى الجودة، عند مستوى معنوية(α=0.05)؟

الحل:

1-الفرضية الإحصائية المطلوب اختيارها هي

$$H_0:\sigma^2=900$$
$$H_1:\sigma^2\neq 900$$

- **إحصاءة الاختبار**: بما أن الاختبار متعلق بالتباين سوف نستخدم توزيع مربع-كاي (χ^2)

لدينا من معطيات التمرين: $\sigma_0^2=30, n=31, S=35, \alpha=0.05$

نحسب أولا تباين العينة المعدل كما يلي: $\hat{S}^2=\frac{nS^2}{n-1}=\frac{(31)(35)^2}{31-1}=1265.83$

نحسب الآن قيمة كاي-المربع المحسوبة كما يلي $\chi^2_{cal}=\frac{(n-1)\hat{S}^2}{\sigma_0^2}=\frac{(31-1)(1265.83)}{(30)^2}=42.19$

أو بشكل مكافئ كما يلي $\chi^2_{cal}=\frac{(n)S^2}{\sigma_0^2}=\frac{(31)(35)^2}{(30)^2}=42.19$

-**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول توزيع مربع-كاي وعند مستوى معنوية 0.05 نستخرج القيمة الحرجة العليا والدنيا كالتالي

$$\begin{cases}\chi^2_{tab}=\chi^2_{1-\frac{\alpha}{2},n-1}\\ \chi^2_{tab}=\chi^2_{\frac{\alpha}{2},n-1}\end{cases}\Rightarrow\begin{cases}\chi^2_{tab}=\chi^2_{0.975,30}\\ \chi^2_{tab}=\chi^2_{0.025,30}\end{cases}\Rightarrow\begin{cases}\chi^2_{tab}=46.979\\ \chi^2_{tab}=16.791\end{cases}$$

القرار: بما أن قيمة مربع كاي المحسوبة والتي تساوي$\left(\chi^2_{cal}=42.19\right)$تقع بين القيمتين الحرجتين لمربع كاي $\left(\chi^2_{tab}=16.791\right)$،$\left(\chi^2_{tab}=46.979\right)$ ، مما يدل ذلك على قبول فرضية العدم H_0 ورفض الفرضية البديلة H_1، وعليه نستنتج بان إنتاج المصنع سيبقى بنفس مستوى الجودة بعد مضي فترة من الزمن بمعنى آخر أن تباين العمر الإنتاجي للمصابيح في العينة لا يختلف عن تباين المجتمع وقد يساوي 900ساعة وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$\begin{aligned}P-\text{value}&=P\left(\chi^2<16.791\vee\chi^2>46.979\right)\\&=F(46.979)-F(16.791)=0.975-0.025=0.95\end{aligned}$$

نلاحظ أن $P-\text{value}>\alpha\Rightarrow0.95>0.05$ ، وبالتالي نقبل (H_0).

تمرين23: أختيرت عينة عشوائية قوامها 40 مشاهدة وتم حساب الانحراف المعياري للعينة وقد بلغ15

المطلوب:

1-هل يمكن الاستنتاج بان العينة مسحوبة من مجتمع طبيعي تباينه أكبر من144، وذلك عند مستوى معنوية (α=0.05)؟

الحل:

1-الفرضية الإحصائية المطلوب اختيارها هي

$$\begin{aligned}H_0&:\sigma^2=144\\H_1&:\sigma^2>144\end{aligned}$$

- **إحصاءة الاختبار**: بما أن الاختبار متعلق بالتباين سوف نستخدم توزيع مربع-كاي $\left(\chi^2\right)$

لدينا من معطيات التمرين: $n=40,S=15,\sigma_0^2=144,\alpha=0.05$

نحسب أولا تباين العينة المعدل كما يلي: $\hat{S}^2=\frac{nS^2}{n-1}=\frac{(40)(15)^2}{40-1}=230.77$

نحسب الآن قيمة كاي-المربع المحسوبة كما يلي $\chi^2_{cal}=\frac{(n-1)\hat{S}^2}{\sigma_0^2}=\frac{(40-1)(230.77)}{144}=62.50$

أو بشكل مكافئ كما يلي $\chi^2_{cal}=\frac{nS^2}{\sigma_0^2}=\frac{(40)(15)^2}{144}=62.50$

-**القيمة الحرجة**: بما أن الاختبار من جانب واحد (من جهة اليمين) ودرجة الحرية 39 أكبر من 30 وهي غير متوفرة في جدول مربع-كاي إذن يمكن إيجاد قيمة مربع-كاي الحرجة $\left(\chi^2_{\alpha}\right)$ على النحو

التالي: $\chi^2_{1-\alpha}=n\left[1-\frac{2}{9n}+Z_{1-\alpha}\left\{\sqrt{\frac{2}{9n}}\right\}\right]^3$

نبحث أولا عن قيمة $Z_{1-\alpha}$ كالتالي: $\alpha=0.05 \Rightarrow Z_{1-\alpha}=Z_{1-0.05}=Z_{0.950}=01.645$

إذن

$$\chi^2_{1-\alpha}=n\left[1-\frac{2}{9n}+Z_{1-\alpha}\left\{\sqrt{\frac{2}{9n}}\right\}\right]^3 \Rightarrow \chi^2_{0.995}=40\left[1-\frac{2}{9\times40}+01.645\left\{\sqrt{\frac{2}{9\times40}}\right\}\right]^3=55.75$$

كما يمكن إيجادها من خلال علاقة التقريب التالية:

$$\chi^2_p=\frac{1}{2}\left(Z_p+\sqrt{2v-1}\right)^2 \Rightarrow \chi^2_{0.995}=\frac{1}{2}\left(Z_{0.995}+\sqrt{2v-1}\right)^2$$

$$\Rightarrow \chi^2_{0.995}=\frac{1}{2}\left(01.645+\sqrt{2(40)-1}\right)^2=55.47$$

<u>القرار</u>: بما أن قيمة مربع كاي المحسوبة والتي تساوي $\left(\chi^2_{cal}=62.50\right)$ هي أكبر من القيمة الحرجة لمربع-كآي $\left(\chi^2_{1-\alpha}=\chi^2_{tab}=55.75\right)$، مما يدل ذلك على رفض فرضية العدم H_0 وقبول الفرضية البديلة H_1، وعليه نستنتج بان العينة محسوبة من مجتمع طبيعي تباينه اكبر من 144وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

08-اختبار تجانس عدة تقديرات مستقلة لتباين المجتمع(اختبار بارتليت)

يعد اختبار تجانس عدة تقديرات مستقلة لتباين المجتمع σ^2من الاختبارات الإحصائية التي تستخدم في اختبار أن التقديرات المستقلة S_i^2 لتباين المجتمع σ^2 هي تقديرات متجانسة محسوبة على أساس K من العينات ويسمى هذا الاختبار باختبار بارتليت (Bartlett Test)

نظرية12: إذا أخذنا مجموعة r من العينات من مجتمعات طبيعية مستقلة عن بعضها البعض، وكانت تباينات هذه العينات هي $S_1^2,S_2^2,S_3^2,.........,S_r^2$ وأحجامها $n_1,n_2,n_3,.........n_r$ على الترتيب، وذا أردنا اختبار الفرضية $H_0=\sigma_1^2=\sigma_2^2=.........=\sigma_r^2$ حيث σ_i^2هو تباين المجتمع i ضد الفرضية H_1 القائلة بأن تباينين على الأقل غير متساويين، فننا نرفض الفرضية الصفرية H_0 عند مستوى معنوية α إذا كان $\chi^2_{cal}>\chi^2_{1-\alpha}(r-1)$

حيث يتم حساب χ^2_{cal} حسب الصيغة التالية:

$$\chi^2_{cal}=\frac{(n-r)\ln\left[\frac{1}{n-r}\sum_{i=1}^{r}(n_i-1)S_i^2\right]-\sum_{i=1}^{r}(n_i-1)\ln S_i^2}{1+\frac{1}{3(r-1)\left[\sum_{i=1}^{r}\frac{1}{n_i-1}-\frac{1}{n-r}\right]}}$$

حيث أن:

n_i: تمثل حجم العينة ذات الرقم (i)

r : تمثل عدد العينات.

Ln: تمثل اللوغاريتم الطبيعي

كما يمكن حساب χ^2_{cal} بالصيغة التالية:

$$\chi^2_{cal} = \frac{\sum_{i=1}^{r}(n_i-1)\mathbf{ln}\left(S^2/S_i^2\right)}{1+\dfrac{1}{3(r-1)\left[\sum_{i=1}^{r}\left(\dfrac{1}{n_i-1}\right)-\dfrac{1}{\sum_{i=1}^{r}(n_i-1)}\right]}}$$

حيث أن

$$S^2 = \frac{\sum_{i=1}^{r}(n_i-1)S_i^2}{\sum_{i=1}^{r}(n_i-1)}$$

<u>تمرين24:</u> أخذت أربع عينات من أربعة مجتمعات طبيعية مستقلة عن بعضها البعض ورتبت أحجام هذه العينات وتبايناتها في الجدول التالي:

حجم العينة	تباين العينة	تباين المجتمع
07	0.47	σ_1^2
07	0.4614	σ_2^2
07	0.5714	σ_3^2
07	0.1995	σ_4^2

<u>المطلوب:</u>

1-أستخدم اختبار بارتليت في اختبار الفرضية الصفرية: تساوي المجتمعات الأربعة مقابل الفرضية البديلة: اثنان على الأقل مختلفان عند مستوى معنوية $(\alpha=0.01)$؟

<u>الحل:</u>

1-فرضية الاختبار $H_0 = \sigma_1^2 = \sigma_2^2 = ... = \sigma_r^2$

H_1:على الأقل أحد التباينات مختلف

التوزيع والقيم الحرجة

التوزيع المستخدم هو توزيع مربع-كآي وحيث أن $(\alpha=0.01)$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة $\chi^2_{1-\alpha,r-1} \Rightarrow \chi^2_{0.99,03} = 11.345$

نقوم الآن بترتيب الحسابات في الجدول التالي وذلك من اجل سهولة إيجاد χ^2_{cal} ومقارنتها مع القيمة الحرجة.

المجتمع	n_i-1	S_i^2	$(n_i-1)S_i^2$	$\ln(S_i^2)$	$(n_i-1)\ln S_i^2$
01	06	0.47	02.82	−0.755	−04.5301
02	06	0.4614	02.7684	−0.7734	−04.6409
03	06	0.5714	03.4282	−0.5596	−03.3579
04	06	0.1995	01.197	−01.6119	−09.6716
المجموع	**24**		**10.2138**		**−22.2007**

$$\chi^2_{cal}=\frac{(n-r)\ln\left[\frac{1}{n-r}\sum_{i=1}^{r}(n_i-1)S_i^2\right]-\sum_{i=1}^{r}(n_i-1)\ln S_i^2}{1+\frac{1}{3(r-1)\left[\sum_{i=1}^{r}\frac{1}{n_i-1}-\frac{1}{n-r}\right]}}\Rightarrow$$

$$\chi^2_{cal}=\frac{(24)\ln\left[\frac{10.2138}{24}\right]+22.2007}{1+\frac{1}{3\times3}\left[\frac{4}{6}-\frac{1}{24}\right]}=\frac{01.6986}{01.0694}=01.588$$

-<u>القرار</u>: نلاحظ أن قيمة كآي-المربع الجد ولية $\left(\chi^2_{tab}=11.345\right)$ أكبر من قيمة كآي-المربع المحسوبة $\left(\chi^2_{cal}=01.588\right)$أي أننا نقبل الفرضية الصفرية ونرفض الفرضية البديلة، ومنه نستنتج أن تباينات المجتمعات الأربعة متجانسة.

<u>تمرين25:</u>أجريت ثلاث دراسات مسحية مستقلة للأسرة للتعرف إلى طبيعة العلاقة بين دخل الأسرة الشهري (بالدينار) ومقدار الإنفاق الشهري للأسرة على بعض السلع الغذائية والخدمات الأساسية وقد تم الحصول على المعلومات التالية:

الدراسة	حجوم العينات (n_i)	تباين الإنفاق (S_i^2)
01	30	190
02	25	210
03	35	240

<u>المطلوب:</u>

1-هل يمكن الاستدلال بان عينات الأسر للدراسات الثلاث أعلاه مسحوبة من مجتمع طبيعي تباينه $\left(\sigma^2\right)$ ، وذلك عند مستوى معنوية $(\alpha=0.01)$؟

الحل:

1- فرضية الاختبار

H_0 : إن عينات الأسر للدراسات الثلاث مسحوبة من مجتمع طبيعي تباينه$\left(\sigma^2\right)$.

H_1 : إن عينات الأسر للدراسات الثلاث مسحوبة من مجتمعات مختلفة التباين.

2-**إحصاءة الاختبار**: لحساب احصاءة الاختبار $\left(\chi^2_{cal}\right)$، ينبغي عمل الجدول التالي:

الدراسة	n_i	S_i^2	n_i-1	$\frac{1}{n_i-1}$	$(n_i-1)S_i^2$	$\mathbf{ln}\left(S^2/S_i^2\right)$	$(n_i-1)\mathbf{ln}\left(S^2/S_i^2\right)$
01	30	190	29	0.034	5510	0.124	03.596
02	25	210	24	0.042	5040	0.024	0.576
03	35	240	34	0.029	8160	-0.110	-03.740
المجموع			**87**	**0.105**	**18710**		**0.432**

حيث أن

$$S^2=\frac{\sum_{i=1}^{r}(n_i-1)S_i^2}{\sum_{i=1}^{r}(n_i-1)}=\frac{18710}{87}=215.057$$

$$\chi^2_{cal}=\frac{\sum_{i=1}^{r}(n_i-1)\mathbf{ln}\left(S^2/S_i^2\right)}{1+\frac{1}{3(r-1)\left[\sum_{i=1}^{r}\left(\frac{1}{n_i-1}\right)-\frac{1}{\sum_{i=1}^{r}(n_i-1)}\right]}}\Rightarrow\chi^2_{cal}=\frac{0.432}{1+\frac{1}{3(3-1)}\left[0.105-\frac{1}{87}\right]}=0.425$$

-**التوزيع والقيم الحرجة**

التوزيع المستخدم هو توزيع مربع-كآي وحيث أن $(\alpha=0.01)$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة $\chi^2_{1-\alpha,r-1}\Rightarrow\chi^2_{0.99,02}=09.210$

-**القرار**: بما أن قيمة مربع -كآي المحسوبة$\left(\chi^2_{cal}=0.425\right)$ هي أقل من القيمة الحرجة لمربع كآي $\left(\chi^2_{tab}=09.210\right)$، مما يعني قبول الفرضية الصفرية ورفض الفرضية العدمية، وبالتالي نستنتج بان عينات الأسر للدراسات الثلاث مسحوبة من مجتمع طبيعي تباينه$\left(\sigma^2\right)$، بمعنى آخر إن تباينات الإنفاق $\left(S_i^2\right)$ الثلاثة هي تقديرات متجانسة وذلك عند مستوى معنوية $(\alpha=0.01)$.

09-اختبار الفروض حول الانحراف المعياري σ

نظرية13: بافتراض لدينا مشاهدات عينة عشوائية حجمها (n_1) اختيرت من مجتمع طبيعي تباينه σ_1^2 غير معلوم، وان S_1 يمثل الانحراف المعياري تم تقديره على أساس مشاهدان العينة، وبافتراض لدينا

مشاهدات عينة عشوائية ثانية حجمها(n_2) اختيرت من مجتمع طبيعي آخر مستقل عن المجتمع الأول تباينه σ_2^2 غير معلوم أيضا، وان S_2 يمثل الانحراف المعياري تم تقديره على أساس مشاهدات العينة الثانية، ولاختبار الفرضية العدمية مقابل الفرضيات الأخرى نلخصها في الجدول التالي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0: \sigma_1 = \sigma_2$ $H_1: \sigma_1 \neq \sigma_2$	$Z_{cal} = \frac{(S_1 - S_2) - (\sigma_1 - \sigma_2)}{\sqrt{S_1^2/2n_1 + S_2^2/2n_2}}$	نرفض (H_0) إذا كانت $Z_{cal} \geq Z_{tab}$ أو $Z_{cal} \leq -Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$
ب-إختبار من طرف واحد: $H_0: \sigma_1 = \sigma_2$ $H_1: \sigma_1 \geq \sigma_2$	$Z_{cal} = \frac{(S_1 - S_2) - (\sigma_1 - \sigma_2)}{\sqrt{S_1^2/2n_1 + S_2^2/2n_2}}$	نرفض (H_0) إذا كانت $Z_{cal} \geq Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$
ج- إختبار من طرف واحد: $H_0: \sigma_1 = \sigma_2$ $H_1: \sigma_1 \leq \sigma_2$	$Z_{cal} = \frac{(S_1 - S_2) - (\sigma_1 - \sigma_2)}{\sqrt{S_1^2/2n_1 + S_2^2/2n_2}}$	نرفض (H_0) إذا كانت $Z_{cal} \leq -Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$ أو إذا كانت $Pvalue < \alpha$

تمرين26:في دراسة حول متوسط الدخل الشهري لإحدى الصناعات، أخذت عينة عشوائية من 600 عامل من معمل تابع لتلك الصناعة، فوجد أن متوسط الدخل الشهري للعامل مساو ل5950 ون، وبانحراف معياري قدره 750 ون، كما أخذت عينة عشوائية من 400 عامل من معمل آخر تابع للصناعة نفسها، فوجد أن متوسط الدخل الشهري للعامل فيها مساو ل5250 ون، وبانحراف معياري قدره 700 ون.

المطلوب:

1-هل يمكن الاستنتاج بان هناك تماثلا في الدخل الشهري للمعمل الأول وللمعمل الثاني وذلك عند مستوى معنوية (α=0.05)؟

الحل:

1-الفرضية الإحصائية المطلوب اختبارها هي $\begin{cases} H_0: \sigma_1 = \sigma_2 \\ H_1: \sigma_1 \neq \sigma_2 \end{cases} \Rightarrow \begin{cases} H_0: \sigma_1 - \sigma_2 = 0 \\ H_1: \sigma_1 \neq \sigma_2 \end{cases}$

<u>إحصاءة الاختبار</u>: بما أن الاختبار متعلق بالفرق بين إنحرافيين معياريين سوف نستخدم التوزيع الطبيعي المعياري (Z)

لدينا من معطيات التمرين: $n_1 = 600, \overline{X}_1 = 5950, S_1 = 750, n_2 = 400, \overline{X}_2 = 5250, S_2^2 = 700$

نحسب أولا قيمة (Z) المحسوبة كالتالي:

$$Z_{cal} = \frac{(S_1 - S_2) - (\sigma_1 - \sigma_2)}{\sqrt{S_1^2/2n_1 + S_2^2/2n_2}} \Rightarrow Z_{cal} = \frac{(750-700)-0}{\sqrt{(750)^2/2\times600 + (700)^2/2\times400}} = 01.520$$

<u>-القيمة الحرجة</u>: بما أن الاختبار من جانبين و من جدول توزيع الطبيعي المعياري وعند مستوى معنوية 0.05 نستخرج القيمة الحرجة كالتالي $Z_{tab} = Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 01.960$

<u>القرار</u>: بما أن قيمة إحصاءة الاختبار المحسوبة$(Z_{cal} = 01.520)$،هي أقل من القيمة الجدولية $(Z_{tab} = 01.960)$ ، وهذا يعني قبول الفرضية العدمية(H_0) ورفض الفرضية البديلة (H_1)، أي انه لا توجد فروق معنوية بين المصنعين، وبالتالي نستنتج أن هناك تماثلا في توزيع الدخل الشهري للمعمل الأول والمعمل الثاني وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

10-إختبار الفروض حول نسبة التباين $\frac{\sigma_1^2}{\sigma_2^2}$

نظرية14:إذا كانت $(X_1, X_2, X_3 \ldots\ldots\ldots X_{n_1})$تمثل عينة عشوائية من توزيع طبيعي بحيث $X \sim N(\mu_1, \sigma_1^2)$ و$(X_1, X_2, X_3 \ldots\ldots\ldots X_{n_2})$عينة عشوائية من توزيع طبيعي آخر $X \sim N(\mu_1, \sigma_1^2)$ ، وكانت العينتين مستقلتين و (μ_1, σ_1^2) و (μ_2, σ_2^2) معلمات مجهولة، وكانت S_1^2 تمثل تباين العينة الأولى و S_2^2 تمثل تباين العينة الثانية فان توزيع المعاينة للنسبة بين التباينين يعطى بالاحصاءة التالية

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2} \quad \text{أو بشكل مكافئ} \quad F = \frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}$$

بدرجة حرية ($V_1 = n_1 - 1$) و($V_2 = n_2 - 1$)، وعليه يمكن استخدام مئويات توزيع فيشر كقيم حرجة لاختبار الفرضيات التالية والموضحة في الجدول الموالي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0:\sigma_1^2=\sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2}=1$ $H_1:\sigma_1^2\neq\sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2}\neq 1$	$F_{cal}=\frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}$ أو بشكل مكافئ $F_{cal}=\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}$	نرفض (H_0) إذا كانت $F_{cal}<F_{\frac{\alpha}{2},V_1,V_2}$ و $F_{cal}>F_{1-\frac{\alpha}{2},V_1,V_2}$ بحيث $F_{tab}=F_{\frac{\alpha}{2},V_1,V_2}$ و $F_{tab}=F_{1-\frac{\alpha}{2},V_1,V_2}$ أو إذا كانت $Pvalue<\alpha$
ب-إختبار من طرف واحد: $H_0:\sigma_1^2=\sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2}=1$ $H_1:\sigma_1^2>\sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2}>1$	$F_{cal}=\frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}$ أو بشكل مكافئ $F_{cal}=\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}$	نرفض (H_0) إذا كانت $F_{cal}>F_{tab}$ بحيث $F_{cal}>F_{1-\alpha,V_1,V_2}$ بحيث $F_{tab}=F_{1-\alpha,V_1,V_2}$ أو إذا كانت $Pvalue<\alpha$
ج- إختبار من طرف واحد: $H_0:\sigma_1^2=\sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2}=1$ $H_1:\sigma_1^2<\sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2}<1$	$F_{cal}=\frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}$ أو بشكل مكافئ $F_{cal}=\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}$	نرفض (H_0) إذا كانت $F_{cal}<F_{tab}$ بحيث $F_{cal}<F_{\alpha,V_1,V_2}$ بحيث $F_{tab}=F_{\alpha,V_1,V_2}$ أو إذا كانت $Pvalue<\alpha$

مع العلم أن $F_{\alpha,v_1,v_2}=\frac{1}{F_{1-\alpha,V_2,V_1}}$

تمرين27: إذا علمت أن مجتمع أطوال الطالبات، ومجتمع أطوال الطلبة في جامعة ما يتبع التوزيع الطبيعي، وسحبنا من الطالبات عينة عشوائية حجمها 31 طالبة، ومن الطلبة عينة عشوائية حجمها 21 طالبا وكانت العينتان مستقلتين، ووجدنا ان تباين أطوال عينة الطالبات يساوي 64، وتباين عينة أطوال الطلبة يساوي 36.

المطلوب:

1–أختبر ما إذا كان هناك فرق بين تباين مجتمع أطوال الطالبات وتباين مجتمع أطوال الطلبة وذلك عند مستوى معنوية $(\alpha=0.10)$؟

الحل:

– الفرضية الإحصائية المطلوب اختبارها هي

$$H_0 : \sigma_1^2 = \sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2 \Leftrightarrow \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

– **إحصاءة الاختبار**: بما أن المجتمعين يتوزعان توزيعا طبيعيا، والعينتين مستقلتين إذن الاختبار المناسب هو اختبار فيشر **(F)** $F_{cal} = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2}$ أو بشكل مكافئ

$$F_{cal} = \frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}$$

لدينا من معطيات التمرين $n_1 = 31, n_2 = 21, S_1^2 = 64, S_2^2 = 36$

نحسب أولا تباين العينتين المعدل كما يلي:

$$\begin{cases} \hat{S}_1^2 = \frac{n_1 \times S_1^2}{n_1 - 1} \\ \hat{S}_2^2 = \frac{n_2 \times S_2^2}{n_2 - 1} \end{cases} \Rightarrow \begin{cases} \hat{S}_1^2 = \frac{31 \times 64}{31-1} = 66.13 \\ \hat{S}_2^2 = \frac{21 \times 36}{21-1} = 37.80 \end{cases}$$

نحسب الآن احصاءة فيشر المحسوبة كما يلي:

$$F_{cal} = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2} \Rightarrow F_{cal} = \frac{66.13}{37.80} \times 1 = 01.75$$

أو بشكل مكافئ كما يلي

$$F_{cal} = \frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}} \Rightarrow F_{cal} = \frac{64\left(\frac{31}{31-1}\right)}{36\left(\frac{21}{21-1}\right)} = 01.75$$

–**القيمة الحرجة:** بما أن الاختبار من جانبين و من جدول توزيع فيشر وعند مستوى معنوية 0.10 نستخرج القيمة الحرجة العليا والدنيا كالتالي

$$F_{tab} = \begin{cases} F_{1-\frac{\alpha}{2},V_1,V_2} = F_{1-\frac{0.10}{2},30,20} = F_{0.95,30,20} = 02.04 \\ F_{\frac{\alpha}{2},V_1,V_2} = \frac{1}{F_{1-\frac{\alpha}{2},V_2,V_1}} = \frac{1}{F_{0.95,20.30}} = \frac{1}{01.93} = 0.518 \end{cases}$$

القرار: بما أن قيمة فيشر المجدولة تقع بين القيمتين 0.518 و 02.04 $(0.518 \leq F_{tab} \leq 02.04)$ أي في منطقة القبول، أما منطقة الرفض فتقع على يسار القيمة 0.518 وعلى يمين القيمة 02.04، و بما

أن قيمة فيشر المحسوبة تساوي 01.75 أي تقع في منطقة القبول، إذن نقبل فرضية العدم H_0 و نرفض الفرضية البديلةH_1، أي أن الاختبار ليس ذا معنوية، أي أن تباين مجتمع الطالبات يساوي تباين مجتمع الطلبة ولا يوجد فرق حقيقي بينهما وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.10

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$P-\text{value}=P\left(F<0.518\vee F>02.04\right)$$
$$=F\left(02.04\right)-F\left(0.518\right)=0.95-0.05=0.90$$

نلاحظ أن $P-\text{value}>\alpha\Rightarrow 0.90>0.10$ ، وبالتالي نقبل (H_0).

تمرين28: نظرا لعدم التزام بعض الشركات المصنعة لإطارات السيارات بالمواصفات القياسية المطلوبة للتصنيع قررت الشركة العامة للإطارات استيراد نوع واحد من الإطارات الذي يعمر أطول مسافة واقل اختلافا، ولهذا قامت الشركة بتجريب نوعين من الإطارات هما A و b حيث استخدمت 16 إطار من كل نوع واستعملت هذه الإطارات حتى استهلكت بالكامل فكانت النتائج كما يلي:

	المتوسط	الانحراف المعياري
النوع A	36800	5100
النوع B	38600	5900

المطلوب:

1-من هذه البيانات هل يمكن القول بان النوع A اقل اختلافا من النوع B وذلك عند مستوى معنوية (α=0.05)، وعلى افتراض أن المجتمعين طبيعيين؟

الحل:

- الفرضية الإحصائية المطلوب اختبارها هي

$$H_0:\sigma_1^2=\sigma_2^2\Leftrightarrow\frac{\sigma_1^2}{\sigma_2^2}=1$$
$$H_1:\sigma_1^2<\sigma_2^2\Leftrightarrow\frac{\sigma_1^2}{\sigma_2^2}<1$$

- **إحصاءة الاختبار**: بما أن المجتمعين يتوزعان توزيعا طبيعيا، والعينتين مستقلتين إذن الاختبار المناسب هو اختبار فيشر (F) $F_{cal}=\frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}$ أو بشكل مكافئ

$$F_{cal}=\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}$$

لدينا من معطيات التمرين $n_1=16, n_2=16, S_1^2=\left(5100\right)^2, S_2^2=\left(5900\right)^2$

نحسب الآن احصاءة فيشر المحسوبة كما يلي:
$$F_{cal}=\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}\Rightarrow F_{cal}=\frac{(5100)^2}{(5900)^2}=0.48$$

-**<u>القيمة الحرجة:</u>** بما أن الاختبار من جانب واحد فقط ومن جهة اليسار و من جدول توزيع فيشر وعند مستوى معنوية 0.02 نستخرج القيمة الحرجة الدنيا كالتالي

$$F_{tab}=F_{\alpha,V_1,V_2}=\frac{1}{F_{1-\alpha,V_2,V_1}}=\frac{1}{F_{0.95,15.15}}=\frac{1}{02.40}=0.416$$

<u>القرار:</u> بما أن قيمة فيشر المجدولة تساوي 0.28 $(F_{tab}=0.416)$ اقل من قيمة فيشر المحسوبة $(F_{cal}=0.48)$، إذن نقبل فرضية العدم H_0 و نرفض الفرضية البديلةH_1، أي أن الاختبار ليس ذا معنوية، أي أن تباين النوع A يساوي تباين النوع B ولا يوجد فرق حقيقي بينهما وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05

كما يمكن إثبات صحة القرار باستخدام مستوى المعنوية المشاهدة (p-value) كالتالي:

$$P-value=P(F<0.416)=F(0.416)=0.05$$

نلاحظ أن $P-value>\alpha\Rightarrow0.90>0.10$ ، وبالتالي نقبل (H_0).

11-تحليل الانحدار وتقييم المعالم

تمهيد: يعتبر تحليل الانحدار أسلوب إحصائي يختص بدراسة العلاقة بين متغيرين أو أكثر وكيفية إيجاد معادلة جبرية تمثل هذه العلاقة رياضيا أفضل تمثيل، وذلك لاستخدامها في التقدير أو التنبؤ بمتوسط احد المتغيرات، ويسمى المتغير التابع (The dépendent variable) بدلالة بقية المتغيرات التي تؤثر فيه وتسمى المتغيرات المستقلة (The Independent variable).

وأول من استخدم تعبير الانحدار العالم "**فرانسيس جالتون**" (**Francis Galton**) عام 1886م، عند دراسته للعلاقة بين طول الأباء وأبنائهم، فقد وجد انه عندما يكون الأباء طوال القامة جدا فان أبنائهم يكونون طوال القامة، ولكن في المتوسط اقصر قامة من الاباء، وإذا كان الاباء قصار القامة جدا فان أبنائهم يكونون قصار القامة، ولكن في المتوسط يكونون أطول قامة من الاباء، فقد استنتج أن أطوال الأبناء تتجه أو ترتد نحو متوسط الطول لجميع أفراد المجتمع، وقد أطلق **جالتون** على هذا الاتجاه نحو المتوسط لفظ الانحدار.

11-1 تعريف الانحدار الخطي البسيط: يكون الانحدار بسيطا، إذا كان يدرس العلاقة بين متغيرين فقط،أحدهما متغير تابع ويرمز له عادة بالحرف (Y)، والثاني متغير أي مستقل ويرمز له عادة بالحرف (X)، ونعبر رياضيا على أن المتغير (Y) يتبع المتغير (X) بالقول أن المتغير Yدالة للمتغير X ويرمز لذلك كما يلي: Y=f(X)

وبالتالي فان نموذج الانحدار الخطي البسيط هو معادلة رياضية تصف العلاقة بين المتغيرين X،Y في المجتمع، والصورة العامة لنموذج الانحدار الخطي البسيط هي كالتالي: $Y = B_0 + B_1X + \mu$

حيث:

Y: قيمة المتغير التابع للمفردة (المشاهدة).

X: قيمة المتغير المستقل للمفردة.

B_1،B_0:معالم المجتمع المجهولة وهي عبارة عن قيم ثابتة.

μ:قيمة الحد الذي يمثل الخطأ العشوائي للمفردة.

وبما أن بيانات المجتمع يصعب الحصول عليها لذلك تستخدم بيانات العينة لتقدير معادلة الانحدار حيث يمكن التعبير عن قيم y_i وبالشكل التالي: $y_i = b_0 + b_1x_i + \varepsilon_i$

وتعطى معادلة الانحدار التقديرية كما يلي: $\hat{y}_i = b_0 + b_1x_i$

حيث:

b_1،b_0 هي التقديرات لمعالم نموذج الانحدار B_0، B_1

y_i هي قيمة Y الحقيقية.

$\hat{y}_i$ هي قيمة y التقديرية.

أما قيمة $\varepsilon_i = \left(y_i - \hat{y}_i\right)$ يساوي الفرق بين القيمة الحقيقية ل Y والقيمة التقديرية $\hat{y}_i$ ويكون الفرق إما موجبا أو سالبا أو معدوما.

وعادة مجموع الفروق الكلية بين القيم الحقيقية والقيم التقديرية يساوي صفرا، أي أن

$$\sum_{i=1}^{n}\varepsilon_i = \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right) = 0$$

2-11 تقدير معادلة الانحدار الخطي: لنفرض انه لدينا عينة من أزواج قيم X،Y بحيث $\left(X_1,Y_1\right),\left(X_2,Y_2\right),........,\left(X_n,Y_n\right)$ عددها n ونرغب بتقدير معادلة الانحدار الخطية $Y = B_0 + B_1X + \mu$ أي إيجاد التقديرات لمعالم خط الانحدار B_1،B_0، حيث توجد عدة طرق لتقدير هذه المعادلة أهمها طريقة المربعات الصغرى.

-**طريقة المربعات الصغرى:** الهدف من تحليل الانحدار هو تحديد أفضل خط يمثل العلاقة بين المتغيرين X و y أو الخط الذي يطابق نقاط شكل الانتشار للبيانات أفضل ما يمكن حسب طريقة المربعات الصغرى فان أفضل خط يطابق نقاط شكل الانتشار هو ذلك الخط الذي تكون أبعاد نقاط شكل الانتشار عنه أصغر ما يمكن، وبما أن أبعاد نقاط شكل الانتشار عن خط الانحدار موجبة وسالبة تؤخذ مربعات هذه الفروق (الأبعاد) أي جعل مجموع مربعات الفروق

$\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$ أصغر ما يمكن ، وللتوصل إلى القيمة الصغرى لمجموع مربعات الفروق $\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$ نقوم بالاشتقاق الجزئي نسبة لكل من b_1،b_0 وتبسيط المعادلات نحصل على تقديرات المربعات الصغرى كما يلي:

$$\frac{\partial\left(\sum\varepsilon_i^2\right)^2}{\partial b_0}=-02\sum_{i=1}^{n}\left(y_i-b_0-b_1x_i\right)=0\text{..................}(1)$$

$$\frac{\partial\left(\sum\varepsilon_i^2\right)^2}{\partial b_1}=-02\sum_{i=1}^{n}x_i\left(y_i-b_0-b_1x_i\right)=0\text{..................}(2)$$

ومن المعادلتين السابقتين نستنتج المعادلتين التاليتين:

$$\sum y_i=nb_0+b_1\sum X_i\text{............................}(3)$$

$$\sum X_iY_i=b_0\sum X_i+b_1\sum X_i^2\text{.....................}(4)$$

وتسمى هاتان المعادلتان بالمعادلتين الطبيعيتين لخط المربعات الصغرى

ومن بيانات العينة نستطيع الحصول على $\sum X_iY_i,\sum X_i,\sum Y_i$ ، وبالتعويض في المعادلتين (03) و(04) وحلهما نستطيع الحصول على b_1،b_0 اللتين تحققان مبدأ طريقة المربعات الصغرى وهما:

$$b_1=\frac{n\sum X_iY_i-\sum X_i\sum Y_i}{n\sum X_i^2-\left(\sum X_i\right)^2}\text{..............................}(05)$$

وبعد معرفة قيمة b_1 نعوض في إحدى المعادلتين السابقتين ولتكن المعادلة (03) فنحصل على قيمة b_0

$$b_0=\frac{1}{n}\left(\sum Y_i-b_1\sum X_i\right)=\overline{Y}-b_1\overline{X}\text{............................}(06)$$

حيث $\overline{X}=\frac{\sum X_i}{n}$ ، $\overline{Y}=\frac{\sum Y_i}{n}$ وهما الوسط الحسابي لقيم X والوسط الحسابي لقيم Y على التوالي، والخط الذي نحصل عليه يسمى بخط الانحدار المقدر أو بخط المربعات الصغرى.

تمرين29: يبين الجدول تطور العلاقة بين استهلاك الكهرباء (X) ومردودية العامل (Y) في 10 مصانع كما يلي:

رقم المصنع	01	02	03	04	05	06	07	08	09	10
Y_i	02	03	06	04	02	03	04	05	06	08
X_i	06	6.10	6.80	7.20	7.40	7.90	8.20	8.50	8.90	9.10

المطلوب:

1-ما هو نوع العلاقة بين مردودية العامل واستهلاك الكهرباء؟

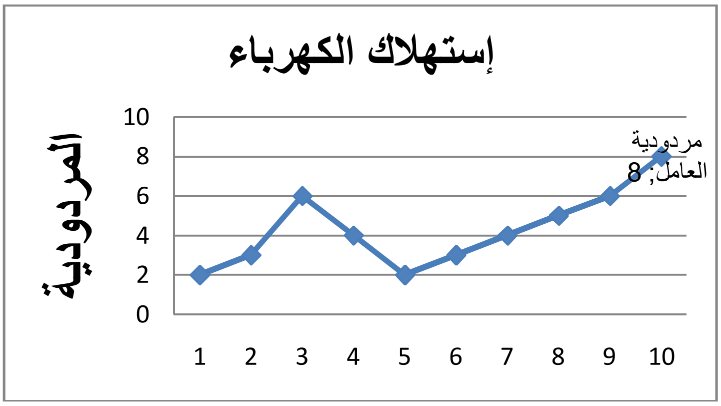
2-قدر معلمات النموذج الخطي البسيط؟

3-ما هي قيمة مردودية العامل عند مستوى X=12؟

4-كيف نتأكد من أن المعلمات المقدرة هي معلمات طريقة المربعات الصغرى؟

الحل:

1-نوع العلاقة بين مردودية العامل واستهلاك الكهرباء



نلاحظ أن العلاقة بين المردودية واستهلاك الكهرباء هي علاقة خطية تقريبا.

2-تقدير معلمات النموذج الخطي البسيط بطريقة المربعات الصغرى العادية

يوضح الجدول التالي مختلف العمليات الحسابية:

X_i^2	$X_i Y_i$	X_i	Y_i	رقم المصنع
36	12	06	02	01
37.20	18.30	06.10	03	02
46.20	40.80	06.80	06	03
51.80	28.80	07.20	04	04
54.70	14.80	07.40	02	05
62.40	23.70	07.90	03	06
67.20	32.80	08.20	04	07
72.20	42.50	08.50	05	08
79.20	53.40	08.90	06	09
82.80	72.80	09.10	08	10
589.97	**339.90**	**76.10**	**43**	**المجموع**

أولا: حساب معامل الانحدار كالتالي:

$$b_1=\frac{n\sum X_iY_i-\sum X_i\sum Y_i}{n\sum X_i^2-\left(\sum X_i\right)^2}=\frac{(10)(339.90)-(76.10)(43)}{(10)(589.97)-(76.10)^2}=01.167$$

ثانيا: حساب b_0

نحسب أولا المتوسط الحسابي لكل من $\overline{X}=\frac{\sum X_i}{n}$ ، $\overline{Y}=\frac{\sum Y_i}{n}$

$$\begin{cases}\overline{X}=\frac{\sum X_i}{n}=\frac{76.10}{10}=07.610\\ \overline{Y}=\frac{\sum Y_i}{n}=\frac{43}{10}=04.30\end{cases}$$

إذن $b_0=\overline{Y}-b_1\overline{X}=04.30-(01.167)(07.610)=-04.580$

وتصبح معادلة الانحدار التقديرية كما يلي: $\hat{y}_i=-04.580+01.167X_i$

3-التنبؤ بالمردودية عند مستوى X=12

$$\hat{y}_i=-04.580+01.167(12)=09.424$$

4-التأكد من أن المعلمات المقدرة هي معلمات طريقة المربعات الصغرى: للتأكد من النتيجة المتحصل عليها من أن b_1،b_0 هي معلمات طريقة المربعات الصغرى العادية نقوم بحساب محدد مصفوفة المشتقات الجزئية من الدرجة الثانية بالنسبة للمعلمات المقدرة

$$\begin{bmatrix}\frac{\partial^2\sum\varepsilon_i^2}{\partial b_0^2},\frac{\partial^2\sum\varepsilon_i^2}{\partial b_0\partial b_1}\\ \frac{\partial^2\sum\varepsilon_i^2}{\partial b_0\partial b_1},\frac{\partial^2\sum\varepsilon_i^2}{\partial b_1^2}\end{bmatrix}=\begin{bmatrix}02n,02\sum X_i\\ 02\sum X_i,02\sum X_i^2\end{bmatrix}$$

إذا كان محدد هذه المصفوفة موجب فان المعلمتان تجعلان مربع البواقي اصغر ما يمكن، وبما أن محدد المصفوفة هو

$$\mathbf{det}=04n\sum X_i^2-04\left(\sum X_i\right)^2\Rightarrow\mathbf{det}=04\times10\times589.97-04(76.10)^2=433.96$$

إن هذا المقدار لا يمكن أن يكون سالبا وبالتالي فان b_1،b_0 معلمتان لطريقة المربعات الصغرى العادية.

3-11 الاستدلال الإحصائي للانحدار الخطي البسيط: حيث أننا افترضنا أن نموذج الانحدار الخطي هو نموذج خطي بسيط وبالتالي نود تقييم جودة هذا النموذج، أي مدى توفيق معادلة الانحدار للبيانات، وعادة ما يتم تقييم نموذج الانحدار الخطي البسيط من خلال اختبار الفرضيات الإحصائية ذات العلاقة بمعلمات النموذج وتكوين فترات الثقة حول هذه المعلمات، ولكن للقيام بمثل هذا الإجراء يتطلب الأمر معرفة خواص مقدرات المربعات لمعلمات نموذج الانحدار وتوزيعاتها الاحتمالية وذلك لأنها المدخل الرئيسي لهذا التقييم.

حيث انه عند دراستنا لأزواج النقاط (X_i,Y_i) اعتبرنا أن y_i تمثل القيمة المشاهدة للمتغير العشوائي Y، ومن خلال النظر للصيغ الرياضية الخاصة بكل من b_1،b_0 نلاحظ أن هاتين الصيغتين تعتمدان على قيم y_i، وعليه فإنهما يمثلان قيم مشاهدات لمتغيرات عشوائية وبالتالي يمكن إيجاد القيمة المتوقعة والتباين لهذه الصيغ، وان لهذه المقدرات الخواص التالية:

1-b_0 ،b_1 مقدرين غير متحيزين أي أن $E(b_1)=B_1,E(b_0)=B_0$

2-نفترض أن التوزيع الاحتمالي للخطأ يتوزع توزيعا طبيعيا بمتوسط حسابي يساوي الصفر وتباين

يساوي σ^2 أي $\varepsilon_i \sim N\left(0,\sigma^2\right)$ ($i=1,2,........,n$)، وعليه فان توزيع المعاينة (التوزيع الاحتمالي) لكل من b_1،b_0 يتبع التوزيع الطبيعي كما يلي:

$$b_0 \sim N\left(B_0,\frac{\sigma^2_{Y/X}\sum_{i=1}^{n}X_i^2}{n\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right)$$

$$b_1 \sim N\left(B_1,\frac{\sigma^2_{Y/X}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right)$$

3-حيث أن غالبا ما تكون $\sigma^2_{Y/X}$ مجهولة وبالتالي يتم استبدالها بالقيمة التقديرية لها وهي $\hat{\sigma}^2_{Y/X}$ ، وبما أن توزيع المعاينة لمعلمتي نموذج الانحدار توزيع طبيعي وبالتالي يمكن أيضا الإثبات بان:

$$\frac{b_0-B_0}{\hat{\sigma}_{b_0}} \sim t_{(n-2)}$$

$$\frac{b_1-B_1}{\hat{\sigma}_{b_1}} \sim t_{(n-2)}$$

إذن مما سبق يمكننا القيام باختبار الفرضيات الإحصائية الخاصة بمعلمات نموذج الانحدار الخطي البسيط وذلك كما يلي:

اختبار الفرضيات الإحصائية: إذا تحقق الشرط الخاص بالخطأ العشوائي $\varepsilon_i \sim N\left(0,\sigma^2\right)$، والذي يمكن التحقق منه باستخدام ما يطلق عليه تسمية تحليل البواقي، فانه يمكن اختبار الفرضيات التالية:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
أ-إختبار من طرفين: $H_0: B_1 = b$ $H_1: B_1 \neq b$	$t_{cal} = \dfrac{b_1 - B_1}{\hat{\sigma}_{b_1}}$ حيث $\hat{\sigma}_{b_1} = \sqrt{\dfrac{\hat{\sigma}^2_{Y/X}}{SS_{XX}}}$	نرفض (H_0) عند مستوى المعنوية α إذا كانت $T_{cal} \geq T_{tab}$ أو $T_{cal} \leq -T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2},n-2}$ أو إذا كانت $Pvalue < \alpha$
ب-إختبار من طرف واحد: $H_0: B_1 = b$ $H_1: B_1 > b$	$t_{cal} = \dfrac{b_1 - B_1}{\hat{\sigma}_{b_1}}$ حيث $\hat{\sigma}_{b_1} = \sqrt{\dfrac{\hat{\sigma}^2_{Y/X}}{SS_{XX}}}$	نرفض (H_0) عند مستوى المعنوية α إذا كانت $T_{cal} \geq T_{tab}$ بحيث $T_{tab} = T_{1-\alpha,n-2}$ أو إذا كانت $Pvalue < \alpha$
ج- إختبار من طرف واحد: $H_0: B_1 = b$ $H_1: B_1 < b$	$t_{cal} = \dfrac{b_1 - B_1}{\hat{\sigma}_{b_1}}$ حيث $\hat{\sigma}_{b_1} = \sqrt{\dfrac{\hat{\sigma}^2_{Y/X}}{SS_{XX}}}$	نرفض (H_0) عند مستوى المعنوية α إذا كانت $T_{cal} \leq -T_{tab}$ بحيث $T_{tab} = T_{1-\alpha,n-2}$ أو إذا كانت $Pvalue < \alpha$

وبالمثل عند الاختبار بالنسبة للمعلمة B_0

تمرين30: إذا افترضنا انه تم اختيار عينة عشوائية من 12 شخص وتم قياس طول ووزن كل منهم وكانت النتائج كما يلي:

الشخص	01	02	03	04	05	06	07	08	09	10	11	12
الطول	150	150	150	155	155	155	155	160	160	175	175	175
الوزن	50	61	54	54	63	59	61	68	65	77	83	72

المطلوب:

1-أوجد معادلة الانحدار الخطي البسيط؟

2-إختبر الفرضية التالية عند مستوى معنوية 05%

$$H_0: B_1 = 0$$
$$H_1: B_1 \neq 0$$

الحل:

1-أيجاد معادلة الانحدار الخطي البسيط

يوضح الجدول التالي مختلف العمليات الحسابية:

X_i^2	$X_i Y_i$	X_i	Y_i	الشخص
22500	7500	150	50	01
22500	9150	150	61	02
22500	8100	150	54	03
24050	8370	155	54	04
24050	8765	155	63	05
24050	9145	155	59	06
24050	9455	155	61	07
25600	10880	160	68	08
25600	10400	160	65	09
30625	13475	175	77	10
30625	14525	175	83	11
30625	12600	175	72	12
306675	**123365**	**1915**	**767**	**المجموع**

أولا: حساب معامل الانحدار كالتالي:

$$b_1 = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - \left(\sum X_i\right)^2} = \frac{(12)(123365)-(1915)(767)}{(12)(306675)-(1915)^2} = 0.899$$

ثانيا: حساب b_0

نحسب أولا المتوسط الحسابي لكل من $\overline{X} = \frac{\sum X_i}{n}$ ، $\overline{Y} = \frac{\sum Y_i}{n}$

$$\begin{cases} \overline{X} = \frac{\sum X_i}{n} = \frac{1915}{12} = 159.583 \\ \overline{Y} = \frac{\sum Y_i}{n} = \frac{767}{12} = 63.916 \end{cases}$$

إذن

$$b_0 = \overline{Y} - b_1\overline{X} = 63.916 - (0.899)(159.583) = -79.60$$

وتصبح معادلة الانحدار التقديرية كما يلي:

$$\hat{y}_i = -79.60 + 0.899X_i$$

2-إختبار الفرضية التالية

$$H_0 : B_1 = 0$$
$$H_1 : B_1 \neq 0$$

قبل إجراء الاختبار نحسب أولا الانحراف المعياري للأخطاء العشوائية كما يلي:

$$\hat{\sigma}^2_{Y/X}=\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{n-2}=\frac{1}{n-2}\left[\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2-b_1SS_{xy}\right]$$

بحيث $$\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2=SS_{yy}=\sum_{i=1}^{n}y_i^2-\frac{\left(\sum_{i=1}^{n}y_i\right)^2}{n}$$

$$SS_{yy}=\sum_{i=1}^{n}y_i^2-\frac{\left(\sum_{i=1}^{n}y_i\right)^2}{n}=50075-\frac{(767)^2}{12}=1050.917$$

$$SS_{xy}=\sum_{i=1}^{n}X_iY_i-\frac{\left(\sum_{i=1}^{n}X_i\right)\left(\sum_{i=1}^{n}Y_i\right)}{n}=123365-\frac{(1915)(767)}{12}=964.583$$

إذن $$\hat{\sigma}^2_{Y/X}=\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{n-2}=\frac{1}{n-2}\left[\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2-b_1SS_{xy}\right]\Rightarrow$$

$$=\frac{1}{12-2}\left[1050.917-(0.899)(964.583)\right]=18.3757$$

نحسب الآن قيمة ستونت المحسوبة كما يلي: $t_{cal}=\frac{b_1-B_1}{\hat{\sigma}_{b_1}}$ بحيث $\hat{\sigma}_{b_1}=\sqrt{\frac{\hat{\sigma}^2_{Y/X}}{SS_{XX}}}$

نحسب أولا قيمة SS_{xx} كالتالي : $$SS_{xx}=\sum_{i=1}^{n}X_i^2-\frac{\left(\sum_{i=1}^{n}X_i\right)^2}{n}=306675-\frac{(1915)^2}{12}=1072.917$$

$$\hat{\sigma}_{b_1}=\sqrt{\frac{\hat{\sigma}^2_{Y/X}}{SS_{XX}}}=\sqrt{\frac{18.3757}{1072.916}}=0.13086$$

إذن $$t_{cal}=\frac{b_1-B_1}{\hat{\sigma}_{b_1}}\Rightarrow t_{cal}=\frac{0.899-0}{0.13086}=06.896$$

ومن جدول ستونت وبدرجة حرية تساوي 10 وعند مستوى معنوية 05% نجد أن

$$T_{tab}=T_{1-\frac{\alpha}{2},n-2}\Rightarrow T_{tab}=T_{1-\frac{0.05}{2},12-2}=T_{0.975,10}=0.228$$

وحيث أن قيمة ستودنت المحسوبة وهي 06.868 أكبر من قيمة ستونت المجدولة وهي 02.228 وبالتالي نرفض الفرضية العدمية ونقبل الفرضية البديلة.

تمارين المحور الثالث

التمرين الأول:يدعي احد موردي آلات التعبئة أن المتوسط الحسابي للعبوة هو 0.50 كغ، وللتأكد من صحة إدعاء المورد تم سحب عينة من العبوات حجمها 36 عبوة فبلغ متوسط العبوة 0.495كغ فإذا علم أن الأوزان المعبأة تتبع التوزيع الطبيعي بانحراف معياري 24 كغ.

المطلوب:

1-هل يمكن للشركة اتخاذ قرار حول شراء مثل هذه الآلات وذلك عند مستوى معنوية (α=0.05)؟

التمرين الثاني: تعتقد الشركة العامة للبريد أن دخلها الشهري من المكالمات الهاتفية الدولية قد تزايد بشكل واضح، ولتأكيد هذا الاعتقاد اختيرت عينة عشوائية تتكون من 100 سجل من سجلات زبائنها فوجد أن متوسط القسط الشهري المدفوع عن المكالمات الدولية يساوي 22.10 ون، فإذا كانت سجلات الشركة تشير إلى أن متوسط القسط الشهري الذي يدفعه الزبائن عن المكالمات الدولية هو 17.10ون وبتباين يساوي 450 ون.

المطلوب:

1-هل تؤيد اعتقاد الشركة عند مستوى معنوية (α=0.01)، مع فرضية أن الدخل الشهري يتبع التوزيع الطبيعي؟

التمرين الثالث: يدعي مدير مصنع للروائح العطرية، أن المتوسط الحسابي لكمية العطر الذي تحتويه زجاجة الكولونيا على الأقل 50 سم^3، بتباين يساوي 16، ولاختبار هذا الادعاء سحبت من الإنتاج الكلي للمصنع 144 زجاجة، وكان المتوسط الحسابي لعبوة زجاجات العينة 49.20 سم^3.

المطلوب:

1-اختبر صحة ادعاء المدير باستخدام مستوى معنوية (α=0.05)؟

التمرين الرابع: يعتقد باحث في مركز البحوث الزراعية أن هناك نوع جديد من الطماطم تم تهجينه ينمو بسرعة أكثر من النوع المتداول بين المزارعين حاليا تحت نفس الظروف البيئية، حيث يكون متوسط طول شتلة الطماطم من النوع الجديد 07.60 سم بعد فترة زمنية معينة، ولتأكيد الاعتقاد قام بزراعة 10 بذور من النوع الجديد وبعد نفس الفترة الزمنية تم قياس الأطوال فكانت كما يلي:06.60،80.10،08.50،08.10، 07.80، 08.50، 09.10، 07.90 ، 07.90، 07.50

المطلوب:

بناء على هذه البيانات هل اعتقاد هذا الباحث صحيحا عند مستوى معنوية (α=0.05)،إذا فرضنا أن أطوال هذه النباتات تتبع التوزيع الطبيعي تقريبا؟

التمرين الخامس: إذا علمت أن متوسط درجات الحرارة خلال 10 أيام من شهر جويلية في ولاية تيارت بلغت 35 درجة بانحراف معياري قدره 01.30 درجة، وان متوسط درجات الحرارة لنفس الفترة في ولاية

وهران بلغت 30 درجة، بانحراف معياري قدره 01.50، علما أن تبايني المجتمعين متساويين.

المطلوب:

هل تعتقد أن متوسط درجات الحرارة في ولاية تيارت يختلف عن متوسط درجات الحرارة في ولاية وهران، وذلك عند مستوى معنوية (α=0.05)؟

التمرين السادس: للمقارنة بين مستوى كفاءة الذكور والإناث في استخدام برنامج كتابة النصوص (Word)، حيث طلب من كل فرد في عينة عشوائية من عشرة ذكور وعينة عشوائية من ثماني إناث، ممن لهم تقريبا نفس مستوى التدريب،كتابة مقال معين بإستخدام هذا البرنامج وفي نهاية التجربة تم قياس الزمن الذي استغرقه كل منهم في الكتابة فكان متوسط الزمن المستغرق في عينة الذكور 06.40 دقيقة بانحراف معياري 60 ثانية، وكان متوسط الزمن في عينة الإناث 05.20 دقيقة بانحراف معياري 48 ثانية.

وعلى فرض أن تباين زمن الكتابة متساوي في مجتمعي الدراسة وان كلا منهما يتبع توزيعا طبيعيا تقريبا.

المطلوب:

1-هل تؤيد هذه البيانات القول بان الذكور أفضل من الإناث في استخدام هذا البرنامج وذلك عند مستوى معنوية(α=0.01)؟

التمرين السابع: ترغب إدارة أحد المطاعم في تحديد ما إذا كان للدعاية أثر على زيادة متوسط دخلها اليومي، وبالتالي قامت بتسجيل الدخل اليومي ولمدة 50 يوم قبل الدعاية فكان المتوسط يساوي 1255 ون، ثم قامت بدعاية عن طريق الصحف والإذاعة المرئية، وبعد 20 يوما من الدعاية قامت بتسجيل الدخل اليومي ولمدة 30 يوما فكان المتوسط 1330 ون، فإذا افترضنا أن الانحراف المعياري للدخل اليومي قبل الدعاية يساوي 215 ون بينما الانحراف المعياري للدخل اليومي بعد الدعاية يساوي 238 ون.

المطلوب:

1-من هذه البيانات هل يمكن القول بان متوسط الدخل اليومي لهذا المطعم قد زاد بعد الدعاية، وذلك عند مستوى معنوية(α=0.10)؟

التمرين الثامن: فيما يلي العلامات التي حصل عليها 08 منتجين في اختبار لكفاءة الأداء قبل وبعد انتهاء دورة تدريبية على مقياس من 1-5 درجات

رقم الشخص	01	02	03	04	05	06	07	08	09	10
قبل الدورة	08	05	04	09	06	08	05	07	08	05
بعد الدورة	10	07	05	06	11	07	08	09	10	07

المطلوب:

1-أختبر فيما ذا كانت الدورة تزيد من كفاءة الأداء للمنتجين، وذلك عند مستوى معنوية (α=0.01)؟

التمرين التاسع: وكالة مبيعات ترب بتقدير الفرق بين نسبة اللذين يستخدمون نوعين من معجوني الأسنان، في عينة حجمها 500 من الذين يستخدمون النوع الأول "أ"، 20 % منهم لا يرغبون التغيير لنوع آخر، وفي عينة حجمها 400 من مستخدمي النوع الثاني "ب"، 13% لا يرغبون في التغيير لنوع آخر.

المطلوب:

1-أختبر فيما إذا كانت النسبة للمجتمع الأول أكبر من النسبة للمجتمع الثاني، وذلك عند مستوى معنوية (α=0.01)؟

التمرين العاشر: تم سؤال عينة عشوائية من 200 طالب وعينة عشوائية من 150 طالبة عن مدى التزامهم بحضور جميع المحاضرات أثناء تواجدهم داخل الحرم الجامعي، فوجد أن 124 طالبا و105 طالبة من طلبة العينة يلتزمون بحضور جميع المحاضرات.

المطلوب:

1-عند مستوى معنوية 05%، هل تعطي هذه المعلومات دلالة كافية على أن الطالبات أكثر التزاما من الطلاب بحضور المحاضرات؟

التمرين الحادي عشر: يدعي مدير مصنع لصناعة نوع معين من الأنابيب، أن طول الأنابيب المنتجة في مصنعه تتوزع توزيعا طبيعيا بتباين يساوي0.09 أو اقل، ولكن يشك المسؤولون على رقابة الإنتاج في ذلك، ويعتقدون أن تباين أطوال الأنابيب المنتجة في هذا المصنع اكبر من 0.09، لذلك سحبوا عينة عشوائية تحتوي على 25 أنبوبة ووجدوا أن تباين أطوال أنابيب العينة يساوي 0.10.

المطلوب:

1-إختبر صحة ادعاء المدير عند مستوى معنوية (α=0.05)؟

التمرين الثاني عشر: تستخدم الشركة العامة للمشروبات الغازية آلة لتعبئة قوارير سعتها 200 مل، ولكن هناك نوع من العشوائية في كمية المشروب المعبئة بهذه القوارير، هذه الكمية تتوزع توزيع طبيعي بانحراف معياري يساوي 03 مل، وبسبب هذه العشوائية في الكمية المعبئة في القوارير، قررت دارة الشركة شراء آلة من نوع آخر تدعي الشركة المصنعة لها أنها أفضل من النوع المستخدم حاليا، ولكن هذه الإدارة سوف لن تقدم على الشراء إلا إذا كانت الكمية المعبئة باستخدام هذه الآلة اقل عشوائية من الآلة السابقة، ولذلك قامت بتجربة الآلة الجديدة واختارت عينة عشوائية تتكون من 71 قارورة فكان الانحراف المعياري لهذه العينة يساوي02.50 مل.

المطلوب:

1-من هذه البيانات هل هناك مبرر لإدارة هذه الشركة بشراء هذه الآلة وذلك عند مستوى معنوية (α=0.05)؟

التمرين الثالث عشر: قام مكتب دراسات اقتصادية بدراسة العلاقة بين الدخل (دخل العائلات) واستهلاك الحليب فكانت النتائج كالتالي:

t	01	02	03	04	05	06	07	08	09	10	11
الاستهلاك	31.6	28.9	30.2	25.8	25	23.6	24.3	22.4	21.3	20	20.4
الدخل	10	12	11	15	16	18	17	20	22	25	24

المطلوب:

1-ما هو شكل العلاقة بين الاستهلاك والدخل؟

2-قدر معلمات النموذج الخطي البسيط بطريقة المربعات الصغرى العادية؟

3-أدرس صلاحية النموذج وذلك عند مستوى معنوية 05%؟

التمرين الرابع عشر: ليكن لديك الجدول التالي والذي يوضح الطبيعة بين متغيرين

X	01	02	03	04	05	06
Y	06	04	03	05	04	02

المطلوب:

1-ارسم الشكل الانتشاري؟

2-أوجد معادلة انحدار Y على X التقديرية؟

3-إختبر الفرضية التالية عند مستوى معنوية 0.05

$$H_0 : B_1 = 0$$
$$H_1 : B_1 \neq 0$$

4-إختبر الفرضية التالية عند مستوى معنوية 0.01

$$H_0 : B_0 = 0$$
$$H_1 : B_1 \neq 0$$

قائمة المراجع

1-قائمة المراجع باللغة العربية

1. إمتثال محمد حسن، عادل محمود حلاوة، لبيبة حسب النبي العطار، مقدمة في أساليب الإستدلال الإحصائي والتنبؤ، الطبعة الأولى، الناشر مكتبة الوفاء القانونية، مصر -الإسكندرية،2012

2. إياد محمد الهوبي، الإحصاء التطبيقي، الطبعة الأولى، الناشر الكلية الجامعية للعلوم والتكنولوجيا -خان يونس-، فلسطين،2014

3. برنار غريه، طرق الإحصاء، ترجمة هيثم لمع، الطبعة الأولى، المؤسسة الجامعية للدراسات والنشر والتوزيع، بيروت -لبنان،1989

4. ثروت محمد عبد المنعم، مدخل حديث للإحصاء و الاحتمالات، الطبعة الثانية، الناشر العبيكان للنشر، الرياض-المملكة العربية السعودية،2007

5. جبار عبد مضحي،مقدمة في نظرية الاحتمالات، الطبعة الأولى، دار المسيرة للنشر والتوزيع، عمان-الأردن،2011

6. جلاطو جيلالي، الإحصاء التطبيقي مع تمارين ومسائل محلولة، الطبعة الاولى، دار الخلدونية للنشر والتوزيع، الجزائر،2007.

7. جلال الصياد، عبد الحميد محمد ربيع، مبادئ الطرق الإحصائية، الطبعة الأولى، جدة-المملكة العربية السعودية، 1983

8. جلال مصطفى الصياد، مصطفى جلال مصطفى، مقدمة في طرق المعاينة الإحصائية، الطبعة الأولى، الناشر مكتبة الصباح، المملكة العربية السعودية، 1990

9. حسن ياسين طعمة، الاختبارات الإحصائية أسس وتطبيقات، الطبعة الثانية ، دار صفاء للنشر والتوزيع،عمان-الأردن،2015

10. حسن ياسين طعمة،إيمان حسين حنوش، الإحصاء الاستدلالي، الطبعة الثانية، دار صفاء للنشر والتوزيع،عمان-الأردن،2015

11. الحليم عشماوي،صلاح جلال،محمد حسين صادق، الإحصاء الحيوي وتصميم التجارب، المكتبة الأكاديمية، مصر-القاهرة،2008

12. سلطان محمد الصلخدي، الإحصاء الرياضي، الناشر جامعة دمشق-كلية العلوم، 2013-2014

13. سهير فهمي حجازي، محمود الدريني، الإحصاء التطبيقي، الناشر الشركة المصرية لإعادة التامين، مصر-القاهرة،2004/2003

14.عبد الحفيظ محمد فوزي مصطفى، نظرية إختبار الفرضيات (الإستدلال الاحصائي02)، الطبعة الأولى، مجموعة النيل العربية، القاهرة،2002

15.عبد الحفيظ محمد فوزي مصطفى، نظرية التقدير (الإستدلال الاحصائي01)، ، مجموعة النيل العربية، القاهرة،2002

16.عبد الحميد عبد المجيد البلداوي، الإحصاء للعلوم الإدارية والتطبيقية، الطبعة الأولى، دار الشروق للنشر والتوزيع،عمان–الأردن،1997

17.عبد العزيز فهمي هيكل، مبادئ الأساليب الإحصائية، الطبعة الأولى، بيروت–لبنان، 1966

18.عبد اللطيف حسن شومان،مقدمة في الاحصاء التطبيقي، الطبعة الاولى، دار الجنان للنشر والتوزيع، عمان–الاردن،2015

19.عبد اللطيف حسن شومان،مقدمة في الإحصاء والاستنساخ الإحصائي، الطبعة الأولى، دار الجنان، عمان–الأردن،2008

20.عدنان عباس حميدان،مطانيوس مخول،فريد جاعوني،عمار ناصر آغا،الإحصاء التطبيقي، منشورات جامعة دمشق،2015–2016

21.عزام صبري، الإحصاء الرياضي، الطبعة الثانية، دار صفاء للنشر والتوزيع، عمان–الأردن، 2014

22.علاء الدين القبانجي،حسام حمامة كمرجي، الإحتمال والإحصاء، منشورات جامعة دمشق، سوريا– دمشق،2012.

23.علي عبد السلام العماري،علي حسين العجيلي،الإحصاء والاحتمالات النظرية والتطبيق، منشورات ELGA، فاليتا–مالطا،2000

24.علي نصر السيد الوكيل، دروس في الإحصاء المتقدم،دار مصطفى للطباعة، بنها–القاهرة،2011

25.فتحي أحمد عاروري، المعاينة الإحصائية طرقها واستخدامها، الطبعة الأولى، الأكاديميون للنشر والتوزيع، عمان–الأردن،2013

26.لحسن عبد الله، مقدمة في الاحتمالات، الطبعة الأولى، دار الوراق للنشر والتوزيع، عمان–الأردن، 2014

27.ليونارد ج.كازمير، الإحصاء التجاري (ملخصات إيزي شوم)، ترجمة مصطفى جلال مصطفى، الطبعة الأولى،الدار الدولية للإستثمارات الثقافية، مصر –القاهرة،2004

28.مبارك أسبر ديب، مبادئ في الإحتمالات والإحصاء، الناشر جامعة تشرين، الجمهورية العربية السورية،2008/2009

29.محمد أبو يوسف، الإحصاء في بحوث العمليات، الناشر المكتبة الأكاديمية، مصر –القاهرة،1989

30.محمد حسين محمد رشيد (القادري)، منى عطا الله الشويلات،مبادئ الإحصاء والاحتمالات ومعالجتها باستخدام برنامج SPSS،الطبعة الثانية، دار صفاء للنشر والتوزيع، عمان–الأردن،2014

31.محمد صبحي أبو صالح،عدنان محمد عوض، مقدمة في الإحصاء (مبادئ وتحليل باستخدام SPSS)، الطبعة الثانية، دار المسيرة للنشر والتوزيع والطباعة، عمان–الأردن،2008.

32.محمد صبحي أبو صالح،عدنان محمد عوض، مقدمة في الإحصاء،مركز الكتب الأردني، إربد–عمان،1982

33.مديحة السيد محمد موسى،أساسيات الإحصاء الرياضي وتطبيقاتها، دار الكتاب الحديث، مصر –القاهرة،2016

34.نبيل جمعه صالح النجار، الإحصاء التحليلي مع تطبيقات برمجية SPSS، الطبعة الأولى، دار الحامد للنشر والتوزيع، عملن–الأردن، 2014

2–قائمة المراجع باللغة الاجنبية

1-Anderson-Sweeney-Williams,Traduction de la 5e edition americaine par Claire Borsenberger, Statistiques pour l'economie et la gestion,3eedition,Distribution Nouveaux Horizons-A.R.S-Paris2010.

2-Douglas C.Montgomery, George C.Runger, Applied Statistics and Probability for Engineers, Third Edition, Printed in the United States of America,2003.

3-E.E Bassett,JM Bremner,BJT Morgan.IT Jolliffe.B Jones,PM North, STATISTICS Problems and solutions, Second Edition, British Library Cataloguing-in-Publication Data,2000.

4-Eva Cantoni, Philippe Huber, Elvezio Ronchetti, Maitriser L'aleatoire, Exercices résolus de probabilités et statistique, Springer-Verlag France, Paris, 2006.

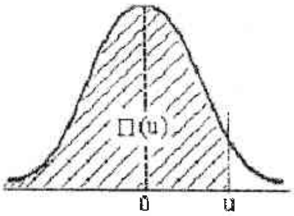
5-Jean-Pierre Boulay,Statistique mathématique « Applications Commentees »,Ellipses Edition Marketing.S.A.2010,paris.

6-Renée Veysseyre, Aide-memoire statistique et probabilités,2 édition, DUNOD,2006,PARIS.

7-Wackerly-Mendenhall-Scheaffer,Mathematical Statistics With Applications,7th Edition, 2008 Thomson Learning, Inc. All Rights Reserved.

الملاحق

Table de Loi Normale
Fonction de répartition F de la loi normale centrée réduite.
Probabilité de trouver une valeur inférieure à z .
F (-z) = 1 - F (z)



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
4.0	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

B.3.3 Loi du Khi-deux à ν ddl $X \sim \chi^2_\nu$

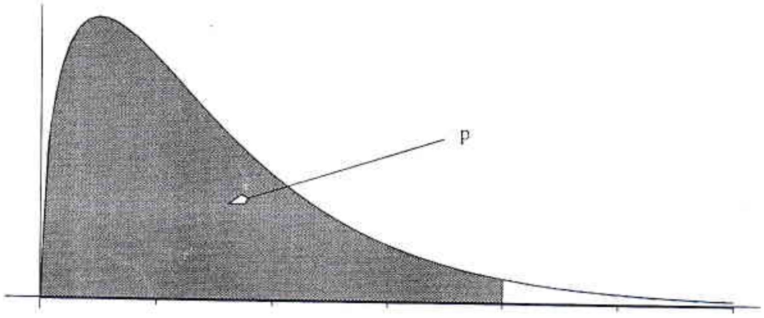


Table des fractiles $c_\nu(p)$ pour une loi du χ^2_ν : $p = \Pr\{X \leq c_\nu(p)\}$

ν \ p	0.001	0.005	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99	0.995	0.999
1	0.000	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879	10.827
2	0.002	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597	13.815
3	0.024	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838	16.266
4	0.091	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860	18.466
5	0.210	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.086	16.750	20.515
6	0.381	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548	22.457
7	0.599	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278	24.321
8	0.857	1.344	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955	26.124
9	1.152	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589	27.877
10	1.479	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188	29.588
11	1.834	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757	31.264
12	2.214	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300	32.909
13	2.617	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	29.819	34.527
14	3.041	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319	36.124
15	3.483	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801	37.698
16	3.942	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267	39.252
17	4.416	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718	40.791
18	4.905	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156	42.312
19	5.407	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582	43.819
20	5.921	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997	45.314
21	6.447	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401	46.796
22	6.983	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796	48.268
23	7.529	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181	49.728
24	8.085	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558	51.179
25	8.649	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928	52.619
26	9.222	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290	54.051
27	9.803	11.808	12.878	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645	55.475
28	10.391	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994	56.892
29	10.986	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.335	58.301
30	11.588	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672	59.702
40	17.917	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766	73.403
50	24.674	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490	86.660
60	31.738	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952	99.608
70	39.036	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.42	104.21	112.32
80	46.520	51.172	53.540	57.153	60.391	64.278	96.578	101.88	106.63	112.33	116.32	124.84
90	54.156	59.196	61.754	65.647	69.126	73.291	107.56	113.14	118.14	124.12	128.30	137.21
100	61.918	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169	149.449

C.3 Loi de STUDENT

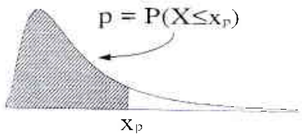
Fractiles de la loi de STUDENT

p	0.5	0.6	0.7	0.8	0.9	0.95	0.975	0.99	0.995	0.999	0.9995
1	0.	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	0.	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.	0.254	0.526	0.846	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.	0.254	0.526	0.845	1.290	1.660	1.984	2.364	2.626	3.174	3.390
200	0.	0.254	0.525	0.843	1.286	1.653	1.972	2.345	2.601	3.131	3.340
∞	0.	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Le fractile $t_{\nu,p}$ est tel que $\mathbb{P}(T_\nu < t_{\nu,p}) = p$.

1 Table des quantiles de la v.a. de Fisher

Fournit les quantiles x_p tels que
$P(X \leq x_p) = p$
pour $X \sim F_{n1;n2}$



p=0.95

n1 / n2	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	inf
1	161	200	216	225	230	234	237	239	241	242	246	248	250	252	253	254
2	18,5	19,0	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,43	19,5	19,5	19,5	19,5	19,5
3	10,1	9,6	9,3	9,1	9,0	8,9	8,9	8,8	8,8	8,8	8,70	8,7	8,6	8,6	8,5	8,5
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80	5,75	5,70	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56	4,50	4,44	4,41	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87	3,81	3,75	3,71	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,45	3,38	3,32	3,28	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15	3,08	3,02	2,98	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94	2,86	2,80	2,76	2,71
10	4,97	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77	2,70	2,64	2,59	2,54
11	4,84	3,98	3,59	3,36	3,20	3,10	3,01	2,95	2,90	2,85	2,72	2,65	2,57	2,51	2,46	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54	2,47	2,40	2,35	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46	2,38	2,31	2,26	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39	2,31	2,24	2,19	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33	2,25	2,18	2,12	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28	2,19	2,12	2,07	2,01
17	4,45	3,59	3,20	2,97	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23	2,15	2,08	2,02	1,96
18	4,41	3,56	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19	2,11	2,04	1,98	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16	2,07	2,00	1,94	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12	2,04	1,97	1,91	1,84
25	4,24	3,39	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,09	2,01	1,92	1,84	1,78	1,71
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,17	2,02	1,93	1,84	1,76	1,70	1,62
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11	1,96	1,88	1,79	1,70	1,64	1,56
40	4,09	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	1,92	1,84	1,74	1,66	1,59	1,51
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	1,89	1,81	1,71	1,63	1,55	1,47
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,87	1,78	1,69	1,60	1,53	1,44
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,84	1,75	1,65	1,56	1,48	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,81	1,72	1,62	1,53	1,45	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,79	1,70	1,60	1,51	1,43	1,32
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,78	1,69	1,59	1,49	1,41	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,98	1,93	1,77	1,68	1,57	1,48	1,39	1,28
150	3,90	3,06	2,67	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,73	1,64	1,54	1,44	1,35	1,22
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,99	1,93	1,88	1,72	1,62	1,51	1,41	1,32	1,19
inf	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,67	1,57	1,46	1,35	1,24	1,00

p=0.99

n1 / n2	1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	inf
1	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6157	6209	6261	6303	6334	6366
2	99	99	99	99	99	99	99	99	99	99	99,4	99,4	99,5	99,5	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	26,9	26,7	26,5	26,4	26,2	26,1
4	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,20	14,02	13,84	13,69	13,58	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,72	9,55	9,38	9,24	9,13	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,56	7,40	7,23	7,09	6,99	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,31	6,16	5,99	5,86	5,75	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,52	5,36	5,20	5,07	4,96	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,96	4,81	4,65	4,52	4,41	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,56	4,41	4,25	4,12	4,01	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,25	4,10	3,94	3,81	3,71	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,01	3,86	3,70	3,57	3,47	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,82	3,66	3,51	3,38	3,27	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,66	3,51	3,35	3,22	3,11	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,52	3,37	3,21	3,08	2,98	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,41	3,26	3,10	2,97	2,86	2,75
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,31	3,16	3,00	2,87	2,76	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,23	3,08	2,92	2,78	2,68	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,15	3,00	2,84	2,71	2,60	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,09	2,94	2,78	2,64	2,54	2,42
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,85	2,70	2,54	2,40	2,29	2,17
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,70	2,55	2,39	2,25	2,13	2,01
35	7,42	5,27	4,40	3,91	3,59	3,37	3,20	3,07	2,96	2,88	2,60	2,44	2,28	2,14	2,02	1,89
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,52	2,37	2,20	2,06	1,94	1,81
45	7,23	5,11	4,25	3,77	3,45	3,23	3,07	2,94	2,83	2,74	2,46	2,31	2,14	2,00	1,88	1,74
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,78	2,70	2,42	2,27	2,10	1,95	1,82	1,68
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,35	2,20	2,03	1,88	1,75	1,60
70	7,01	4,92	4,07	3,60	3,29	3,07	2,91	2,78	2,67	2,59	2,31	2,15	1,98	1,83	1,70	1,54
80	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64	2,55	2,27	2,12	1,94	1,79	1,65	1,49
90	6,93	4,85	4,01	3,53	3,23	3,01	2,84	2,72	2,61	2,52	2,24	2,09	1,92	1,76	1,62	1,46
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	2,22	2,07	1,89	1,74	1,60	1,43
150	6,81	4,75	3,91	3,45	3,14	2,92	2,76	2,63	2,53	2,44	2,16	2,00	1,83	1,66	1,52	1,33
200	6,76	4,71	3,88	3,41	3,11	2,89	2,73	2,60	2,50	2,41	2,13	1,97	1,79	1,63	1,48	1,28
inf	6,64	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,04	1,88	1,70	1,52	1,36	1,00