وزارة التعليم العالي والبحث العلمي

جامعة العربي التبسي– تبسة –

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

ميدان التكوين: علوم اقتصادية تسيير وعلوم تجارية

محاضرات في الإحصاء 3

أمثلة وتمارين

المستوى: سنة ثانية ليسانس

د. عبد الحليم الحمزة

السنة الجامعية: 2023/2022

بسم الله الرحمن الرحيم

الفهرس

مقدمة

إن علم الإحصاء يعد أحد الأساليب العلمية الشائعة الاستخدام الذي يستعمل كوسيلة فعالة لتحليل المشكلات ومعالجتها في الحياة العملية بشكل موضوعي، ويعد أيضا أداة لخدمة متخذي القرار والعلماء في مختلف مجالات المعرفة عن طريق تزويدهم بالمؤشرات التحليلية التي تساعدهم على اتخاذ القرارات الرشيدة بشأن المشكلات قيد الدراسة.

وعلم الإحصاء كبقية العلوم الأخرى قد شهد تطورا سريعا خلال القرنين التاسع عشر والعشرين، مقترنا بتطور نظرية الاحتمالات، وتوزيعات المعاينة ونظريات التقدير واختبار الفرضيات، وبهدف إلمام الطلبة بمقياس الإحصاء 3 وتسهيل فهمهم له، تمت صياغة مطبوعة محاضرات في الإحصاء 3 بغرض تخفيف الصعوبات التي يواجهونها وفهم أفضل لهذا المقياس، وهذه المطبوعة هي عبارة عن سلسلة من المحاضرات في الإحصاء الاستدلالي موجهة لطلبة السنة الثانية في العلوم الاقتصادية والتجارية وعلوم التسيير نظام (LMD) نقدم من خلالها دروس مبسطة ومختصرة وسهلة الفهم مدعمة بالعديد من الأمثلة والتمارين.

سعينا إلى تقديم ما هو مقرر دراسته في مقياس الإحصاء 3 لطلبة العلوم الاقتصادية والتجارية وعلوم التسيير، لذلك قمنا بتقسيم محتوى هذه المطبوعة إلى ثلاثة فصول، يتضمن الفصل الأول نظرية توزيع المعاينة، ويتناول الفصل الثاني نظرية التقدير، أما الفصل الثالث تضمن نظرية اختبار الفرضيات.

الفصل الأول

نظرية توزيع المعاينة

1- مفاهيم في الإحصاء الاستدلالي

الاستدلال الإحصائي هو احد وظائف علم الإحصاء، وتظهر اهميته في كثير من النواحي التطبيقية التي يكون فيها معالم توزيع الظاهرة في المجتمع محل الدراسة غير معلومة، او النواحي التي يكون فيها التوزيع ذا معالم معروفة ويحتاج الباحث إلى اتخاذ قرارات حول صحة قيم هذه المعالم وذلك في الوقت الذي يصعب فيه إجراء حصر شامل لجميع مفردات المجتمع من اجل إمكانية حساب القيم الحقيقية لهذه المعالم. لذا يقوم الباحث بسحب عينة عشوائية من المجتمع محل الدراسة من اجل دراسة خصائص هذا المجتمع.

1-1- مفهوم المجتمع الاحصائي: المجتمع الإحصائي هو مجموعة من العناصر أو الأحداث المتشابهة التي تكون (بجميع عناصرها) موضوعا لدراسة علمية ما، قد يكون المجتمع الإحصائي مجموعة من العناصر المحسوسة مثل كل البشر الأحياء على كوكب الأرض أو كل النجوم في الكون، وقد تكون مجموعة من الأحداث الافتراضية مثل كل النتائج المحتملة لرمي 10 نرود معا، ترمي أغلب الدراسات الإحصائية إلى معرفة معلومات عن المجتمع الإحصائي المدروس.

1-2- مفهوم العينة الإحصائية: الجزء الذي يتم اختياره من المجتمع وهذه العينة تستخدم لتوفير الوقت والجهد، وتمثل المجتمع الأصلي وتحقق أغراض البحث وتغني الباحث عن مشقات دراسة المجتمع الأصلي. والعينة هي جزء ممثل لمجتمع البحث الأصلي.

1-3- أنواع العينات الإحصائية: تتمثل فيما يلي:

1-3-1- العينات (العشوائية) الاحتمالية: تقسم الى عدة أنواع تتمثل فيما يلي[1]:

- **العينة العشوائية البسيطة:** هي مجموعة جزئية من المجتمع الاحصائي لها نفس الفرصة لتختار كعينة من ذلك المجتمع، أي بمعنى أن جميع أفراد المجتمع لهم فرصة في أن يُختاروا، ويرجع ذلك إلى أن المجتمع متجانس إذا اختيرت منه عينة وبأي طريقة تستطيع تمثيله وتظهر فيها جميع خصائصه وسماته.

- **العينة العشوائية الطبقية:** المعاينة الطبقية يُلجأ إليها في حالة معرفة التركيب النسبي للمجتمع الأصلي عندما يكون هذا المجتمع مكونا من عدة طبقات بينها اختلاف واضح من حيث أحد أو مجموعة من الخصائص فتختار طريقة العينة الطبقية حرصا على أن تمثل جميع تلك الطبقات في العينة المختارة وعادة تكون العينة الطبقية متباينة فيما بينها ومتجانسة في داخلها، مثال ذلك، سوق ملابس به عدة أقسام: قسم الأطفال، قسم الصبيان، قسم الرجال، قسم النساء وغيرها، فهذه الأقسام هي عبارة عن طبقات.

[1] . نافذ محمد بركات، **توزيعات المعاينة**، مشورات الجامعة الإسلامية، غزة، 2020، ص 12.

– **العينة العشوائية العنقودية:** وهي تختلف عن المعاينة الطبقية في مبدأ العناقيد الذي يحدد أن تكون العناقيد متباينة في داخلها متجانسة فيما بينها أي عكس العينة الطبقية. لنأخذ نفس المثال في العينة الطبقية لكن هنا يكون شكل السوق بدون أقسام أي جميع الملابس توجد في محل واحد به الأطفال، الصبيان، الرجال النساء، وهذا ما نعني به متباينة في داخلها. أما متجانسة فيما بينها كأن تكون هنالك عدة أسواق بهذا الشكل. وبالتالي يمكنك أن تأخذ جميع أغراضك من محل واحد. وهذا ما يحدث في حالة العينة العنقودية عنقود واحد تجد فيه جميع أفراد المجتمع ولا تحتاج أن تذهب لكل العناقيد أي يمكنك الاستغناء عن البقية لأنها تحمل نفس الخصائص وهذا لا يحدث في العينة الطبقية حيث تقسم الطبقات على أساس خاصية واحدة محددة لا تتوفر في الطبقات الاخرى لذا لا بد عليك المرور على كل الطبقات (الأقسام) لتجد كل ما تحتاج إليه ولا تستطيع أن تستغني عن أي طبقة أو قسم، حسب المثال.

– **العينة العشوائية المنتظمة:** العينة المنتظمة يكون اختيار الوحدات منها على أساس تقسيم العدد الكلي للمجتمع على حجم العينة المطلوبة، ومن ثم توزيع وحدات المجتمع الأصلي وبشكل متساوٍ ومنتظم على الرقم الناتج من ذلك التقسيم مثلا: إذا كان العدد الكلي للمجتمع هو (3000) طالب وطالبة مثلاً وهو رقم يمثل عدد الطلبة في كلية ما، وكانت العينة المطلوبة هي (150) طالب وطالبة فقط فيكون توزيع الوحدات الكلية الأصلية للمجتمع على الشكل الآتي: 3000÷150=20 وعلى هذا الأساس يتحدد رقم العينة أي اسم الطالب الأول بحيث يكون أقل من الرقم (20) وليكن (3) مثلاً ويختار عشوائيا, ثم يبدأ الباحث بتوزيع العينة على بقية الأسماء بالشكل الآتي: أول رقم هو(3)، أما الرقم الثاني فهو(20+3=23) والثالث (43)، ثم (63) ثم (83) ثم (103) الى آخره، حتى نصل إلى (2983). وبهذا المنطق أعطينا فرصة لكل فرد من أفراد المجتمع المتمثل بما مجموعه (3000) طالب وطالبة أن يكونوا ضمن أفراد العينة وبشكل منتظم.

1–3–2– العينات غير الاحتمالية: تقسم الى عدة أنواع تتمثل فيما يلي[1]:

– **العينة العارضة:** يتجه الباحث في هذا النوع من العينات إلى اختيار الحالات التي تصادف الباحث مثل الاشخاص الذين يقابلهم الباحث قبل غيرهم في الطريق.

– **العينة العمدية (القصدية):** يختار الباحث في هذا النوع من العينات، حالات يعتقد أنها تمثل المجتمع في الجانب الذي يتناوله البحث، كأن يختار الباحث منطقة حسب اعتقاده هي الأكثر ملائمة للقيام بالبحث فيها. وتوفر هذه الطريقة على الباحث الكثير من الوقت والجهد الذي يبذله في اختيار العينة إلا أنها تستلزم

[1] . نافذ محمد بركات، مرجع سابق، ص 15.

معرفة المعالم الاحصائية بالنسبة للمجتمع الأصلي خاصة بالنسبة للوحدات التي يرغب الباحث في اختيارها وهو أمر لا يتيسر في كل الأحوال.

- **العينة الحصصية:** يكثر استخدام هذا النوع خاصة في استطلاعات الرأي العام لما يتميز به من سرعة حيث يقسم الباحث المجتمع إلى طبقات أو فئات حسب خصائص معينة ويعمل على تمثيل كل فئة من فئات العينة بالنسبة لوجودها في المجتمع، ويترك للباحث حرية اختيار مفردات الحصة بشرط أن يلتزم بالحدود العدية أو النوعية للعينة، إنما يخشى منه عدم تمثيل المجتمع الأصلي تمثيلا صحيحا.

1-4- مفهوم المعاينة: هو عمل إحصائي منظم مبني على أسس علمية ويقوم على مبدأ شمول جزء من المجتمع الإحصائي وتختار المفردات في الغالب باعتماد أحد أساليب المعاينة الاحتمالية، أو شمول جميع وحدات المجتمع وإخضاعها للمشاهدة من خال المسح الشامل، والمعاينة أيضا هي اختيار جزء من وحدات المجتمع بطرق معينة تتضمن تمثيل المجتمع الاصلي بجميع وحداته خير تمثيل، ويسمى الجزء المأخوذ بالعينة[1].

1-5- توزيعات المعاينة: توزيع المعاينة للإحصائية هو التوزيع الاحتمالي لهذه الإحصائية المحسوب لكل العينات الممكنة والمأخوذة من المجتمع الإحصائي المدروس أيا كان حجمه وأيا كانت طريقة السحب.

إذا أخذنا عينات متكررة من مجتمع ما وقمنا بقياس متوسط لكل عينة، فإننا نجد أن معظم هذه المتوسطات تختلف عن بعضها البعض، ويسمى التوزيع الاحتمالي لمتوسطات العينات بتوزيع المعاينة للمتوسط الحسابي، ولكن توزيع المعاينة للمتوسط الحسابي له أيضا متوسط حسابي يعبر عنة بالرمز $\mu_{\bar{x}}$ وانحراف معيارى (خطأ معياري) يعبر عنه بالرمز $\sigma_{\bar{x}}$، فتوزيع المعاينة للمتوسطات الحسابية هو عبارة عن التوزيع الاحتمالي للمتوسطات الحسابية لعدد كبير من العينات العشوائية المتساوية الحجم ومن مجتمع إحصائي واحد.

1-6- أنواع طرق المعاينة: تقسم المعاينة حسب طرقها الى قسمين هما[2]:

1-6-1- المعاينة بالإرجاع: هي المعاينة التي يتم فيها اختيار كل عنصر من المجتمع أكثر من مرة، ويمكن اعتبار هذا المجتمع الذي تتم فيه المعاينة بالإرجاع مجتمع غير منته أو مجتمع غير محدود طالما انه يمكن أن نسحب منه عينة أيا كان حجمها دون استنفاد المجتمع. وهناك قاعدة لمعرفة هل المعاينة تمت بالإرجاع او بدون الارجاع وهي:

المجتمع غير محدود أي أن المعاينة بالإرجاع اذا تحقق ما يلي: $n < 0,05N$

[1] . عودة، احمد سليمان، **الإحصاء للباحث في التربية والعلوم الإنسانية**، دار الأمل، الأردن، 2010، ص 18.

[2] . نفس المرجع السابق، ص ص: 28- 37.

N: حجم المجتمع، n: حجم العينة

إذا كان السحب بالإرجاع عدد العينات يساوي: N^n

مثال:

لنفترض ان لدينا مجتمع حجمه 3 سحبت كل العينات الممكنة والتي حجمها 2 فما هو عدد العينات الممكنة اذا كان السحب مع الارجاع.

عدد العينات يساوي: $N^n = 3^2 = 9$

إذا كان x_i متغير عشوائي يخضع للتوزيع الطبيعي متوسطه μ وتباينه σ^2 وكان $\bar{x}$ يمثل المتوسط الحسابي للعينات ذات الحجم n والمسحوبة من مجتمع غير محدود او غير منته (المعاينة بالإرجاع) فإن توزيع المعاينة للمتوسطات الحسابية له الخصائص التالية:

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

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

إذا لم يحدد حجم المجتمع فإننا نعتبر حجم المجتمع غير محدود.

مثال:

سحبت عينة عشوائية من مجتمع لا نهائي (السحب بالإرجاع) متوسطه 70 وتباينه 40

اذا كان حجم العينة 10 فأوجد:

المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية.

تباين توزيع المعاينة للمتوسطات الحسابية.

الانحراف المعياري لتوزيع المعاينة للمتوسطات الحسابية.

الحل:

حساب المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية

$$\mu_{\bar{x}} = \mu = 70$$

حساب تباين توزيع المعاينة للمتوسطات الحسابية

المجتمع غير محدود وعليه:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{40}{10} = 4$$

حساب الانحراف المعياري لتوزيع المعاينة للمتوسطات الحسابية

$$\sigma_{\bar{x}} = \sqrt{\sigma^2_{\bar{x}}} = \sqrt{4} = 2$$

1-6-2- المعاينة بدون إرجاع: هي المعاينة التي يتم فيها اختيار كل عنصر من المجتمع مرة واحدة فقط، ويُمكن اعتبار هذا المجتمع الذي تمت فيه المعاينة بدون إرجاع بالمجتمع المنته أو المجتمع المحدود.

المجتمع محدود أي أن المعاينة بدون إرجاع اذا تحقق ما يلي: $n \geq 0,05N$

نفترض أن لدينا مجتمع حجمه N مفردة هي:

$$X_1, X_2, X_3, X_4, \dots, X_N$$

نريد سحب عينة من هذا المجتمع حجمها n ومفرداتها:

$$x_1, x_2, x_3, \dots, x_n$$

عدد العينات ذات الحجم n والتي يمكن اختيارها من المجتمع ذات الحجم N هو:

إذا كان السحب بدون إرجاع عدد العينات يساوي: $C_n^N = \frac{N!}{n!(N-n)!}$

مثال:

إذا كان لدينا مجتمع حجمه 5 مفردات ونريد أخذ عينات بدون إرجاع حجمها 3 مفردات فإن عدد العينات الممكنة:

$$C_3^5 = \frac{5!}{3!\,(5-3)!} = 10$$

إذا كان x_i متغير عشوائي يخضع للتوزيع الطبيعي متوسطه μ وتباينه σ^2 وكان $\bar{x}$ يمثل المتوسط الحسابي للعينات ذات الحجم n والمسحوبة من مجتمع محدود او منته (المعاينة مع عدم الإرجاع) فإن توزيع المعاينة للمتوسطات الحسابية له الخصائص التالية:

$$\mu_{\bar{x}} = \mu$$

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

حيث أن $\frac{N-n}{N-1}$ يسمى بمعامل التصحيح ويستخدم إذا كان حجم المجتمع محدود.

مثال:

بافتراض أن لدينا مجتمع حجمه 4 مفردات هي 6، 3، 5، 2 وأردنا سحب عينة من هذا المجتمع حجمها مفردتين أوجد:

المتوسط الحسابي والتباين للمجتمع.

المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية.

التباين لتوزيع المعاينة للمتوسطات الحسابية.

الحل:

حساب المتوسط الحسابي للمجتمع

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = 4$$

حساب التباين للمجتمع

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} = 2.5$$

حساب المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية

$$\mu_{\bar{x}} = \mu = 4$$

حساب التباين لتوزيع المعاينة للمتوسطات الحسابية

لدينا: $\frac{n}{N} = \frac{2}{4} = 0,5 > 0,05$

اذن المجتمع محدود وعليه:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1} = \frac{2.5}{2} \times \frac{4-2}{4-1} = 0.833$$

وبالتالي فان الانحراف المعياري يساوي:

$$\sigma_{\bar{x}} = \sqrt{\sigma_{\bar{x}}^2} = \sqrt{0.833} = 0.913$$

1-7- معالم المجتمع: معلمة المجتمع هي عبارة عن خاصية أو مقياس يتم حسابها من المجتمع محل الدراسة. أي أن المعالم هي مقاييس تحدد خصائص المجتمع (التوزيع).

أهم معالم المجتمع هي:

المتوسط الحسابي للمجتمع يساوي: $\mu = \frac{\sum_{i=1}^{N} X_i}{N}$

التباين يساوي: $\sigma^2 = \frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}$

الانحراف المعياري (الجذر التربيعي للتباين) يساوي: $\sigma = \sqrt{\sigma^2}$

النسبة تساوي: $P = \frac{X}{N}$ حيث: X:عدد المفردات التي تتوافر فيهم الخاصية محل الدراسة.

1-8- إحصاءات العينة: إحصاءة العينة هي عبارة عن خاصية أو مقياس يتم حسابها من العينة المسحوبة من المجتمع محل الدراسة. أي أن الاحصاءة هي دالة في بيانات العينة.

أهم إحصاءات العينة هي:

المتوسط الحسابي للعينة يساوي: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

التباين يساوي: $S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$

الانحراف المعياري (الجذر التربيعي للتباين) يساوي: $S = \sqrt{S^2}$

النسبة تساوي: $\hat{P} = \frac{x}{n}$ حيث: x:عدد المفردات التي تتوافر فيهم الخاصية محل الدراسة.

2- توزيع المعاينة للمتوسطات الحسابية

هو عبارة عن توزيع احتمالي للمتوسط الحسابي للعينات، فالعينات المختلفة غالبا تكون قيم متوسطاتها الحسابية مختلفة، أي ان المتوسط الحسابي للعينة كأي إحصائية يتغير من عينة إلى أخرى، ولا نستطيع معرفة المتوسط الحسابي لأي عينة عشوائية قبل سحبها، أي أن المتوسط الحسابي للعينة هو متغير عشوائي وبالتالي له توزيع احتمالي، ويطلق على التوزيع الاحتمالي لأي إحصائية بتوزيع المعاينة وبالتالي يسمى التوزيع الاحتمالي للمتوسط الحسابي للعينة بتوزيع المعاينة للمتوسط الحسابي للعينة.

2-1- توزيع المعاينة للمتوسطات الحسابية في حالة مجتمع طبيعي

2-1-1- حالة تباين المجتمع σ^2 معلوم

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 وسحبنا منه كل العينات العشوائية البسيطة ذات الحجم n، فتوزيع المعاينة للمتوسط الحسابي للعينة $\bar{X}$ يتوزع توزيعا طبيعيا بمتوسط حسابي يساوي $\mu_{\bar{x}} = \mu$ وتباين يساوي $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$

أي: $\bar{X} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}^2)$

ويتوزع هذا المتغير توزيعا طبيعيا معياريا تبعا للاحصاءة Z كما يلي[1]:

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

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

[1] . المنيزل عبد الله فلاح، **الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الاحصائية (SPSS)**، دار وائل للنشر، عمان، 2000، ص 24.

مثال:

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.

أوجد توزيع المعاينة لمتوسط اعمار سائقي سيارات الأجرة.

أحسب احتمال ان يكون متوسط العمر لهذه العينة أكبر من 48 سنة.

الحل:

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 45 سنة وانحراف معياري 07 سنوات وهو يمثل اعمار سائقي سيارات الأجرة

$$X \sim N(45,49)$$

إيجاد توزيع المعاينة لمتوسط اعمار سائقي سيارات الاجرة

بما ان المجتمع يخضع للتوزيع الطبيعي فان توزيع المعاينة سوف يخضع هو الآخر للتوزيع الطبيعي

$$\bar{X} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}^2)$$

حساب المتوسط الحسابي والتباين لتوزيع المعاينة

المتوسط الحسابي لتوزيع المعاينة هو:

$$\mu_{\bar{x}} = \mu = 45$$

بما ان المجتمع محدود والسحب مع عدم الارجاع ننتقل الى حساب كسر المعاينة $\frac{n}{N}$

$$\frac{n}{N} = \frac{16}{1500} = 0{,}0106 < 0{,}05$$

حجم العينة صغير نسبيا بالنسبة لحجم المجتمع اذن لا نستعمل معامل الارجاع او التصحيح

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{49}{16} = 3.0625$$

اذن: $\bar{X} \sim N(45.03,3.0625)$

حساب احتمال ان يكون متوسط العمر لهذه العينة أكبر من 48 سنة

بما ان توزيع المعاينة يخضع للتوزيع الطبيعي لحساب الاحتمال نحسب القيمة المعيارية او الاحصاءة Z

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

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

2-1-2- حالة تباين المجتمع σ^2 مجهول وحجم العينة أكبر أو يساوي 30 $(n \geq 30)$

في كثير من الحالات يكون تباين للمجتمع σ^2 مجهولا، ولذلك اذا كان حجم العينة n اكبر او يساوي 30 فإننا نستعمل تباين العينة S^2 كتقدير لتباين المجتمع، ويكون توزيع المعاينة للمتوسط الحسابي $\overline{X}$ قريبا من التوزيع الطبيعي مادامت $n \geq 30$

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 مجهول وسحبنا منه كل العينات العشوائية البسيطة ذات الحجم n حيث $n \geq 30$، وكانت S^2 تباين العينة، فتوزيع المعاينة للمتوسط الحسابي للعينة $\overline{X}$ يتوزع توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي يساوي $\mu_{\overline{x}} = \mu$ وتباين يساوي $\sigma_{\overline{x}}^2 = \frac{S^2}{n}$

أي: $\bar{X} \sim N(\mu_{\overline{x}}, \sigma_{\overline{x}}^2)$

ويتوزع هذا المتغير توزيعا طبيعيا معياريا تبعا للاحصاءة Z كما يلي[1]:

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

مثال:

إذا كان X متغير عشوائي يمثل أطوال الطلبة، وكان هذا المتغير يخضع للتوزيع الطبيعي بمتوسط حسابي 160 سم وتباين مجهول، اخذت عينة عشوائية حجمها 37 طالب وكان تباين العينة هو 60 سم.

أوجد توزيع المعاينة لمتوسط أطوال الطلبة.

أحسب احتمال ان يكون طول الطلبة في هذه العينة اقل من 163سم.

الحل:

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 160 سم وتباين مجهول وهو يمثل أطوال الطلبة

$$X \sim N(160, \sigma^2)$$

[1] . المنيزل عبد الله فلاح، مرجع سابق، ص 34.

إيجاد توزيع المعاينة لمتوسط أطوال الطلبة

بما ان المجتمع يخضع للتوزيع الطبيعي وتباينه مجهول، وحجم العينة المسحوبة أكبر أو يساوي 30، فان توزيع المعاينة له توزيع قريب من التوزيع الطبيعي

$\bar{X} \sim N(\mu_{\bar{x}}, \sigma^2_{\bar{x}})$

حساب المتوسط الحسابي والتباين لتوزيع المعاينة

المتوسط الحسابي لتوزيع المعاينة هو:

$$\mu_{\bar{x}} = \mu = 160$$

بما ان المجتمع غير محدود اذن لا نستعمل معامل الارجاع او التصحيح

$$\sigma^2_{\bar{x}} = \frac{S^2}{n} = \frac{60}{37} = 1,62$$

اذن: $\bar{X} \sim N(160,1,62)$

حساب احتمال ان يكون أطوال الطلبة لهذه العينة أقل من 163 سم

بما ان توزيع المعاينة يخضع للتوزيع الطبيعي لحساب الاحتمال نحسب القيمة المعيارية او الاحصاءة Z

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

$$P(\overline{X} < 163) = P\left(Z < \frac{163 - 160}{\sqrt{1.62}}\right) = P(Z < 02.32) = F(02.32) = 0.98983$$

2-1-3- حالة تباين المجتمع σ^2 مجهول وحجم العينة أقل من 30 $(n < 30)$

في الحالات التي يكون فيها حجم العينة اقل تماما من 30 فان تباين العينات تكون ذات تغير كبير لدرجة ان S^2 لا تكون تقديرا موثوقا لتباين المجتمع وبالتالي يكون توزيع المعاينة للمتوسط الحسابي $\overline{X}$ خاضعا لتوزيع لا يمكننا تقريبه الى التوزيع الطبيعي وانما يخضع توزيع ستيودنت T (توزيع العينات الصغيرة) بدرجة حرية (V=n−1) أي:

$$\bar{X} \sim T_v\left(0, \frac{v}{v-2}\right)$$

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 مجهول وسحبنا منه كل العينات العشوائية البسيطة ذات الحجم n حيث $n < 30$، وكانت S^2 تباين العينة، فتوزيع المعاينة للمتوسط الحسابي للعينة $\overline{X}$ يتوزع وفق توزيع ستيودنت T بدرجة حرية (V=n−1)، وبمتوسط حسابي يساوي:

$\mu_{\bar{x}} = \mu$ وتباين يساوي $\sigma^2_{\bar{x}} = \frac{S^2}{n}$

ويتوزع هذا المتغير وفق توزيع ستيودنت تبعا للاحصاءة T كما يلي[1]:

$$T = \frac{\overline{X}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}} - \frac{\overline{X}_i - \mu}{\frac{S}{\sqrt{n}}}$$

مثال:

إذا كان X متغير عشوائي يمثل اوزان اكياس الدقيق التي تنتجها احدى المؤسسات، وكان هذا المتغير يخضع للتوزيع الطبيعي بمتوسط حسابي 50 كغ وتباين مجهول، اخذت عينة عشوائية بسيطة حجمها 9 اكياس من انتاج هذه المؤسسة، ووجد ان الانحراف المعياري لأوزان هذه الاكياس يساوي01 كغ.

أوجد توزيع المعاينة لمتوسط اوزان الاكياس.

أحسب احتمال ان يزيد المتوسط الحسابي لهذه العينة عن 51 كغ.

الحل:

X: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 50 كغ وتباين مجهول وهو يمثل أوزان الأكياس

$$X \sim N(50, \sigma^2)$$

إيجاد توزيع المعاينة لمتوسط أوزان الأكياس

بما ان المجتمع يخضع للتوزيع الطبيعي وتباينه مجهول، وحجم العينة المسحوبة أقل من 30، فان توزيع المعاينة يخضع لتوزيع ستيودنت T بدرجة حرية (V=8)

$$\bar{X} \sim T_v(\mu_{\bar{x}}, \sigma^2_{\bar{x}})$$

حساب المتوسط الحسابي والتباين لتوزيع المعاينة

المتوسط الحسابي لتوزيع المعاينة هو:

$$\mu_{\bar{x}} = \mu = 50$$

بما ان المجتمع غير محدود اذن لا نستعمل معامل الارجاع او التصحيح

$$\sigma^2_{\bar{x}} = \frac{S^2}{n} = \frac{1}{9} = 0{,}11$$

حساب احتمال ان يزيد المتوسط الحسابي لهذه العينة عن 51 كغ

بما ان توزيع المعاينة يخضع لتوزيع ستيودنت لحساب الاحتمال نحسب الاحصاءة T

$$T = \frac{\overline{X}_i - \mu\bar{x}}{\sigma_{\bar{x}}}$$

[1] . بدر سالم عيسى، **مبادئ الاحصاء الوصفي والاستدلالي**، دار المسيرة، عمان، 2007، ص 42.

$$P(\bar{X} < 51) = P\left(T_{08} < \frac{51-50}{\sqrt{0.11}}\right) = P(T_{08} < 02.828) = F(02.828) = 0.988$$

وذلك بعد العودة الى جداول توزيع ستيودنت

2-2- توزيع المعاينة للمتوسطات الحسابية في حالة مجتمع غير طبيعي أو غير معلوم

قد يحدث في كثير من الاحيان ان يكون المجتمع الذي تسحب منه العينات غير معتدل، فقد يكون ملتويا نحو اليمين او نحو اليسار، في هذه الحالة نطبق نظرية النهاية المركزية. تأتي اهمية هذه النظرية من كونها اعتمدت 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 أيا كان شكل توزيعها.

ونظرية النهاية المركزية لها أهمية كبيرة في الاستنتاج الاحصائي، فباستخدامها نعتبر توزيع المعاينة للمتوسط الحسابي للعينة توزيعا قريبا من التوزيع الطبيعي، وبالتالي نستطيع استخدام التوزيع الطبيعي في استنتاجات خاصة بالمتوسط الحسابي للمجتمع دون الحاجة الى معرفة التوزيع الاحتمالي لهذا المجتمع، بشرط ان يكون حجم العينة كبيرا $n \geq 30$ وكلما زاد حجم العينة كلما اقترب توزيع المعاينة للمتوسط الحسابي للعينة أكثر من التوزيع الطبيعي مهما كان توزيع المجتمع الذي سحبنا منه العينات[1].

وتنص نظرية النهاية المركزية على أنه إذا كان لدينا مجتمع توزيعه لا يتبع التوزيع الطبيعي (مجهول متوسطه الحسابي μ وتباينه σ^2 وسحبنا منه كل العينات العشوائية البسيطة والممكنة وذات الحجم الكبير $n \geq 30$ فتوزيع المعاينة للمتوسط الحسابي للعينة $\bar{X}$ سيتوزع توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي قدره $\mu_{\bar{x}}$ وتباين قدره $\sigma^2_{\bar{x}}$

أي أن: $Z = \frac{\bar{X}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \sim N(0,1)$

يقترب من التوزيع الطبيعي المعياري كلما كبر حجم العينة n.

مثال:

في احد المصانع كان المتوسط الحسابي لأجور العمال 300 ون وانحراف معياري 50 ون، وقد تم سحب عينة عشوائية من 40 عامل من عمال هذا المصنع.

أوجد توزيع المعاينة للوسط الحسابي لأجور العمال.

---

[1] . بدر سالم عيسى، مرجع سابق، ص 49.

الحل:

X: متغير عشوائي مستمر يخضع لتوزيع مجهول بمتوسط حسابي 300 ون وانحراف معياري 50 ون وهو يمثل متوسط الأجور

إيجاد توزيع المعاينة لمتوسط الأجور

بما ان المجتمع يخضع لتوزيع مجهول ونظرا لأن حجم العينة المسحوبة أكبر من 30، إذا يمكننا تطبيق نظرية النهاية المركزية، وعليه فتوزيع المعاينة للمتوسط الحسابي قريب من التوزيع الطبيعي أي أن:

$\bar{X} \sim N(\mu_{\bar{x}}, \sigma^2_{\bar{x}})$

حساب المتوسط الحسابي والتباين لتوزيع المعاينة

المتوسط الحسابي لتوزيع المعاينة هو:

$$\mu_{\bar{x}} = \mu = 300$$

بما ان المجتمع غير محدود اذن لا نستعمل معامل الارجاع او التصحيح

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = \frac{50^2}{40} = 62{,}5$$

إذن: $\bar{X} \sim N(300{,}62.5)$

3– توزيع المعاينة للفروق بين المتوسطات الحسابية

في كثير من البحوث يرغب الباحث في مقارنة مجتمعين لمعرفة الاختلاف بينهما، وذلك بمقارنة متوسطيهما اي بحساب الفرق بين متوسطيهما الحسابيين (حيث μ_1 ترمز للمتوسط الحسابي للمجتمع الأول وترمز μ_2 للمتوسط الحسابي للمجتمع الثاني، وحيث انه في أغلب الدراسات تكون بيانات المجتمعين غير متوافرة وبالتالي يكون متوسطيهما الحسابيين مجهولين ولاجراء مثل هذه الدراسات نحاول استنتاج الفرق الحقيقي بين متوسطي المجتمعين، باستخدام الفرق بين متوسطين حسابيين لعينتين $\overline{X_1} - \overline{X_2}$ ، حيث العينة الاولى مسحوبة من المجتمع الاول، والعينة الثانية مسحوبة من المجتمع الثاني، ويرمز للمجتمع الأول بالرقم 01 وللمجتمع الثاني بالرقم 02، وبالتالي فان $\overline{X_1}$ يرمز للمتوسط الحسابي لعينة مسحوبة من المجتمع الاول، و $\overline{X_2}$ يرمز للمتوسط الحسابي لعينة مسحوبة من المجتمع الثاني، وبما ان الاحصائيتين $\overline{X_1}$ و $\overline{X_2}$ متغيران عشوائيين، إذن الفرق بينهما $\overline{X_1} - \overline{X_2}$ هو الآخر متغير عشوائي ولإجراء اي استنتاجات خاصة بالمعلمة $\mu_1 - \mu_2$ باستخدام الإحصائية $\overline{X_1} - \overline{X_2}$ يجب معرفة ودراسة طبيعة توزيع المعاينة لهذه الاحصائية، والذي يطلق عليه توزيع المعاينة للفرق بين متوسطين حسابيين.

3-1- توزيع المعاينة للفرق بين متوسطين حسابيين لعينتين مستقلتين

إذا تمت المعاينة من مجتمعين مستقلين، فلمعرفة توزيع المعاينة للفرق بين المتوسطين الحسابيين، فان ذلك يتوقف على طبيعة توزيع المجتمعين (طبيعي او غير طبيعي) وكذلك معلومية او مجهولية تبايني المجتمعين وكذا حجم العينتين المسحوبتين (أكبر أو يساوي 30 او أصغر تماما من 30) وسوف نتطرق الى مختلف هذه الحالات كالتالي:

3-1-1- توزيع المعاينة للفرق بين متوسطين حسابيين لمجتمعين طبيعيين بتباينين معلومين

إذا سحبنا كل العينات العشوائية الممكنة ذات الحجم n_1 من مجتمع لامحدود يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 ، وسحبنا كل العينات العشوائية الممكنة ذات الحجم n_2 من مجتمع آخر لا محدود يتوزع هو الاخر توزيعا طبيعيا بمتوسط 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 الحسابية $\overline{X_1} - \overline{X_2}$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي $\mu_{\overline{x_1}-\overline{x_2}}$ وتباين $\sigma^2_{\overline{x_1}-\overline{x_2}}$ أي[1]:

$$\overline{X}_1 - \overline{X}_2 \sim N\left(\mu_{\overline{x_1}-\overline{x_2}}, \sigma^2_{\overline{x_1}-\overline{x_2}}\right)$$

أما معلمتي التوزيع تساويان:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2$$

$$\sigma^2_{\overline{x_1}-\overline{x_2}} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

أما إذا كان احد المجتمعين او كليهما لا يتوزع توزيعا طبيعيا، فيكون توزيع المعاينة للفروق قريبا من التوزيع الطبيعي حسب نظرية النهاية المركزية اذا كانت العينتان كبيرتين $n_1, n_2 \geq 30$

مثال:

عينة عشوائية حجمها 20 سحبت من مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي 50 وتباين 09، وعينة أخرى حجمها 15 سحبت من مجتمع آخر يتوزع توزيعا طبيعيا بمتوسط 40 وتباين 04، وكانت العينتان المسحوبتان مستقلتين عن بعضهما البعض.

[1] . بدر سالم عيسى، مرجع سابق، ص 53.

أوجد توزيع المعاينة للفرق بين متوسطي العينتين.

أحسب احتمال ان يكون الفرق بين متوسطي العينتين اقل من 8,2.

الحل:

توزيع المعاينة للفرق بين متوسطي العينتين

المجتمع الأول

X_1: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 50 وتباين 9 أي:

$$\overline{X}_1 \sim N(50{,}9)$$

المجتمع الأول

X_2: متغير عشوائي مستمر يخضع للتوزيع الطبيعي بمتوسط حسابي 40 وتباين 4 أي:

$$\overline{X}_2 \sim N(40{,}4)$$

بما ان المجتمع الأول يخضع للتوزيع الطبيعي فان توزيع المعاينة $\overline{X}_1$ سوف يخضع هو الآخر للتوزيع الطبيعي، والمجتمع الثاني يخضع للتوزيع الطبيعي وبالتالي فان توزيع المعاينة $\overline{X}_2$ سوف يخضع هو الآخر للتوزيع الطبيعي، وعليه فان الفرق $\overline{X_1} - \overline{X_2}$ هو متغير عشوائي جديد سوف يكون له توزيع قريب من التوزيع الطبيعي.

معلمتي التوزيع تساويان:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2 = 50 - 40 = 10$$

$$\sigma^2{}_{\overline{x_1}-\overline{x_2}} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{9}{20} + \frac{4}{15} = 0.72$$

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

حساب احتمال ان يكون الفرق بين متوسطي العينتين اقل من 8.2

بتطبيق المتغير الطبيعي المعياري (Z) الذي علاقته كالتالي:

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

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

3-1-2- توزيع المعاينة للفرق بين متوسطين حسابيين لمجتمعين طبيعيين بتباينين مجهولين وحجم العينتين كبير

إذا سحبنا كل العينات العشوائية الممكنة ذات الحجم $n_1 \geq 30$ من مجتمع لامحدود يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، وسحبنا كل العينات العشوائية الممكنة ذات الحجم $n_2 \geq 30$ من مجتمع آخر لا محدود يتوزع هو الاخر توزيعا طبيعيا بمتوسط 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 الحسابية $\overline{X_1} - \overline{X_2}$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي $\mu_{\overline{x_1}-\overline{x_2}}$ وتباين $\sigma^2_{\overline{x_1}-\overline{x_2}}$ أي:[1]

$$\overline{X}_1 - \overline{X}_2 \sim N\left(\mu_{\overline{x_1}-\overline{x_2}}, \sigma^2_{\overline{x_1}-\overline{x_2}}\right)$$

أما معلمتي التوزيع تساويان:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2$$

$$\sigma^2_{\overline{x_1}-\overline{x_2}} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

مثال:

معمل ينتج 700 كغ من العجين الغذائي كمعدل يومي، سحبت منه عينة عشوائية تمثل انتاج 40 يوما متوسطها الحسابي 740 كغ بانحراف معياري 40 كغ، معمل آخر ينتج 500 كغ كمعدل يومي، سحبت منه عينة عشوائية تمثل إنتاج 35 يوما متوسطها الحسابي 480 كغ بانحراف معياري 20 كغ.

أوجد توزيع المعاينة للفرق بين متوسطي العينتين.

أحسب احتمال ان يكون الفرق محصور بين 180كغ و210 كغ.

---

[1] . أبو صالح محمد صبحي، **مقدمة في الإحصاء: مبادئ وتحليل باستخدام SPSS**، ط 6، دار المسيرة، عمان، 2012، ص 49.

الحل:

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.

معلمتي التوزيع تساويان:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2 = 700 - 500 = 200$$

$$\sigma^2{}_{\overline{x_1}-\overline{x_2}} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = \frac{40^2}{40} + \frac{20^2}{38} = 52{,}8$$

أي: $\overline{X}_1 - \overline{X}_2 \sim N(200{,}52{,}8)$

حساب احتمال ان يكون الفرق بين متوسطي العينتين محصور بين 180 و 210 كغ

بتطبيق المتغير الطبيعي المعياري (Z) الذي علاقته كالتالي:

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

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

3-1-3- توزيع المعاينة للفرق بين متوسطين حسابيين لمجتمعين طبيعيين بتباينين مجهولين وحجم أحد العينتين على الأقل صغير

إذا كان لدينا مجتمع اول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، ومجتمع ثان يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، العينة وحجم المسحوبة من المجتمع الأول $n_1 < 30$ ، وحجم العينة المسحوبة من المجتمع الثاني $n_2 < 30$ ، وكانت العينتان مستقلتين عن بعضهما البعض، فلدراسة توزيع المعاينة للفرق $\overline{X}_1 - \overline{X}_2$ في هذه الحالة نكون أمام حالتين هما:

- **الحالة الأولى: عندما يكون التباينين مجهولين ومتساويين** $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$

إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2، وكان σ_1^2 و σ_2^2 مجهولين ومتساويين، وسحبنا من المجتمع الأول عينة عشوائية بسيطة حجمها $n_1 < 30$، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2 < 30$ مع فرضية تساوي تبايني المجتمعين، كما نعلم ان افضل تقدير لتباين المجتمع هو

تباين العينة المسحوبة منه، فيعني ذلك ان افضل تقدير لتباين المجتمع الأول هو تباين العينة المسحوبة منه S_1^2 وان افضل تقدير لتباين المجتمع الثاني هو تباين العينة المسحوبة منه S_2^2 ، وبالطبع كلا من S_1^2 و S_2^2 مختلفان في القيمة وبما ان تباين المجتمع الأول يساوي تباين المجتمع الثاني فليس من المنطق ان نقدر معلمتين متساويتين بتقديرين مختلفين، لذلك نقدر التباينين بنفس التقدير، وهو عبارة عن المتوسط المرجح لتباين العينة الأولى وتباين العينة الثانية، ويتم الترجيح بدرجة الحرية الخاصة بكل عينة، ويطلق على هذا المقدار التباين المشترك، ويرمز له بالرمز S_p^2 .

وهو يحسب كما يلي:

$$S_p^2 = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2}$$

ويخضع توزيع المعاينة للفرق بين متوسطي في حالة عدم معلومية تبايني المجتمعين وحجم العينين اقل تماما من 30 الى توزيع ستيودنت 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[1]:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2$$

$$\sigma^2{}_{\overline{x_1}-\overline{x_2}} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$$

$$= \frac{S_p^2}{n_1} + \frac{S_p^2}{n_2} = S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

وتعطى إحصاءة ستيودنت T كما يلي:

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

مثال:

سحبت عينة عشوائية حجمها 16 بتباين 04 من مجتمع أول يخضع للتوزيع الطبيعي بمتوسط حسابي 30 وتباينه مجهول، وسحبت عينة اخرى حجمها 14 بتباين 05 من مجتمع ثاني يخضع هو الآخر للتوزيع الطبيعي بمتوسط حسابي 28 وتباين مجهول.

أوجد توزيع المعاينة للفرق بين متوسطي العينتين بافتراض تساوي التباينين.

---

[1] . أبو صالح محمد صبحي، مرجع سابق، ص 59.

احسب احتمال ان يكون الفرق اقل من 04.

الحل:

توزيع المعاينة للفرق بين متوسطي العينتين بافتراض تساوي التباينين

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

توزيع المعاينة يخضع لتوزيع ستيودنت 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:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2 = 30 - 28 = 2$$

قبل حساب تباين توزيع المعاينة نحسب أولا التباين المشترك

$$S_p^2 = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2} = 4.785$$

$$\sigma^2{}_{\overline{x_1}-\overline{x_2}} = \sigma_{\overline{X}_1}^2 + \sigma_{\overline{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = \frac{S_p^2}{n_1} + \frac{S_p^2}{n_2} = S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$
$$= 4.785\left(\frac{1}{16} + \frac{1}{14}\right) = 0.64$$

توزيع المعاينة كما يلي:

$$\overline{X}_1 - \overline{X}_2 \sim T_{28}(2,0.64)$$

حساب احتمال ان يكون الفرق اقل من 04

وتعطى إحصاءة ستيودنت T كما يلي:

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

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

– الحالة الثانية: عندما يكون التباينين مجهولين وغير متساويين $(\sigma_1^2 \neq \sigma_2^2)$

إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2، وكان σ_1^2 و σ_2^2 مجهولين وغير متساويين،

وسحبنا من المجتمع الأول عينة عشوائية 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 تقدير لتباين المجتمع الأول هو تباين العينة المسحوبة منه S_1^2 وان افضل تقدير لتباين المجتمع الثاني هو تباين العينة المسحوبة منه S_2^2.

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

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

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) واقترحت عدة حلول لهذه المشكلة أكثرها استخداما هو اعتبار ان توزيع هذا المتغير هو توزيع قريب من توزيع ستيودنت T بدرجة حرية V تحسب كالتالي[1]:

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

وتعطى إحصاءة ستيودنت T كما يلي:

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

مثال:

سحبت عينة عشوائية حجمها 09 بتباين 36 من مجتمع أول يخضع للتوزيع الطبيعي بمتوسط حسابي 64 وتباينه مجهول، وسحبت عينة اخرى حجمها 16 بتباين 25 من مجتمع ثاني يخضع هو الآخر للتوزيع الطبيعي بمتوسط حسابي 59 وتباين مجهول.

أوجد توزيع المعاينة للفرق بين متوسطي العينتين بافتراض عدم تساوي التباينين.

احسب احتمال ان يكون الفرق أكبر من 08.

[1] . أبو صالح محمد صبحي، مرجع سابق، ص 64.

الحل:

توزيع المعاينة للفرق بين متوسطي العينتين بافتراض عدم تساوي التباينين

المجتمعان يخضعان للتوزيع الطبيعي بتباينين مجهولين وبما ان حجم العينتين المسحوبتين أقل تماما من 30 ومع فرضية عدم تساوي التباينين فان توزيع المعاينة للفرق سوف يضع لتوزيع قريب من توزيع ستيودنت بدرجة حرية تساوي: V

توزيع المعاينة يخضع لتوزيع ستيودنت بالعلاقة التالية:

$$\overline{X}_1 - \overline{X}_2 \sim T_v\left(\mu_{\overline{x_1}-\overline{x_2}}, \sigma^2{}_{\overline{x_1}-\overline{x_2}}\right)$$

وتعطى معلمتي التوزيع كما يلي:

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2 = 64 - 59 = 5$$

$$\sigma^2{}_{\overline{x_1}-\overline{x_2}} = \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = \frac{36}{9} + \frac{25}{16} = 5{,}5625$$

درجة الحرية V تحسب كالتالي:

$$V = \frac{\left[\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right]^2}{\frac{\left[\frac{S_1^2}{n_1}\right]^2}{n_1 - 1} + \frac{\left[\frac{S_2^2}{n_2}\right]^2}{n_2 - 1}} = \frac{\left[\frac{36}{9} + \frac{25}{16}\right]^2}{\frac{\left[\frac{36}{9}\right]^2}{8} + \frac{\left[\frac{25}{16}\right]^2}{15}} = 16$$

توزيع المعاينة يعطى كما يلي:

$$\overline{X}_1 - \overline{X}_2 \sim T_{16}(5{,}5{,}5625)$$

حساب احتمال ان يكون الفرق أكبر من 08

تعطى إحصاءة ستيودنت T كما يلي:

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

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

3-2- توزيع المعاينة للفرق بين متوسطين حسابيين لعينتين مرتبطتين

في كثير من البحوث عند مقارنة متوسطي عينتين نجد ان العينتين غير مستقلتين، وتكون كل قيمة في العينة الاولى مقرونة بقيمة في العينة الثانية، أي تكون القيمة الأولى في العينة الأولى والقيمة الاولى في العينة الثانية تابعتين لنفس المفردة، والقيمة الثانية في العينة الاولى والقيمة الثانية في العينة الثانية تابعتين لنفس المفردة وهكذا، أي نجد ان القيم المشاهدة في العينتين تكون ازواجا من القيم، ولذلك تسمى العينتان في هذه الحالة بالعينتين المرتبطتين.

إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2، وكان المجتمعان مرتبطان وسحبت من كل مجتمع عينة حجمها n $(x_1,x_2,\ldots\ldots,x_n)$ و$(y_1,y_2,\ldots\ldots,y_n)$على الترتيب، بحيث تمثل x_i و y_i القيمتين المتناظرتين في العينتين، إذا كانت الفروق بين قيم العينتين المتناظرتين هي: $(d_1,d_2,\ldots\ldots.d_n)$، حيث: $(d_i=x_i-y_i)$ لجميع القيم $(i=1,2,3,\ldots\ldots,n)$ فان $(d_1,d_2,\ldots\ldots.d_n)$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على أنها عينة عشوائية من مجتمع طبيعي متوسطه الحسابي $\mu_d = \mu_1 - \mu_2$ وتباينهσ_d^2، ولنفرض ان متوسط العينة هو $\bar{d}$ وتباينها S_d^2 ونميز الحالتين التاليتين[1]:

3-2-1- الحالة الأولى: حجم العينة $n \geq 30$ و σ_d^2 مجهول

في هذه الحالة فان متوسط الفرق للعينتين سوف يخضع للتوزيع الطبيعي كالتالي:

$$\bar{d} \sim N\left(u_{\bar{d}}, \sigma_{\bar{d}}^2\right)$$

ويتم حساب معلمتي التوزيع كما يلي:

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

$$\sigma_{\bar{d}}^2 = \frac{S_d^2}{n}$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

3-2-2- الحالة الثانية: حجم العينة $n < 30$ و σ_d^2 مجهول

في هذه الحالة فان متوسط الفرق للعينتين سوف يخضع لتوزيع ستيودنت T بدرجة حرية V أي:

[1] . أبو صالح محمد صبحي، مرجع سابق، ص 70.

$$\bar{d} \sim T_v\left(u_{\bar{d}}, \sigma_{\bar{d}}^2\right)$$

ويتم حساب معلمتي التوزيع كما يلي:

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

$$\sigma_{\bar{d}}^2 = \frac{S_d^2}{n}$$

وتعطى إحصاءة ستيودنت T كما يلي:

$$T = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}}$$

مثال:

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

8	8	8	7	7	7	8	9	9	9	8	8	7	7	8	7	**X_i**
5	7	7	5	6	5	8	7	6	5	8	7	4	5	5	5	**Y_i**

X_i: قياس الكفاءة بعد دخول العمال دورة تحسين الأداء

Y_i: قياس الكفاءة قبل دخول العمال دورة تحسين الأداء

- أوجد توزيع المعاينة للفرق بين متوسطي العينتين المرتبطتين.
- أحسب احتمال أن يكون الفرق في الأداء قبل وبعد الدورة لا يقل عن 2.5.

الحل:

إيجاد توزيع المعاينة للفرق بين العينتين

بما ان أداء العمال متغير عشوائي يخضع للتوزيع الطبيعي وحجم العينة المسحوبة اقل تماما من 30 فتوزيع المعاينة سوف يخضع لتوزيع ستيودنت.

ايجاد معلمتي التوزيع

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

8	8	8	7	7	7	8	9	9	9	8	8	7	7	8	7	**X_i**
5	7	7	5	6	5	8	7	6	5	8	7	4	5	5	5	**Y_i**

d_i	2	3	2	3	1	0	4	3	2	0	2	1	2	1	1	3

$$\mu_{\bar{d}} = \frac{\sum_{i=1}^{n} |d_i|}{n} = \frac{\sum_{i=1}^{n} |X_i - Y_i|}{n} = \frac{30}{16} = 1,875$$

$$S_d^2 = \frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n} = \frac{12,71875}{16} = \frac{30}{16} = 0,80$$

$$\sigma_{\bar{d}}^2 = \frac{S_d^2}{n} = \frac{0,80}{16} = 0,053$$

متوسط الفرق للعينتين سوف يخضع لتوزيع ستيودنت T بدرجة حرية 15 أي:

$$\bar{d} \sim T_{15}(1,875,0,053)$$

حساب احتمال أن يكون الفرق في الأداء قبل وبعد الدورة لا يقل عن 2.5

نستعين بإحصاءة ستيودنت T كما يلي:

$$T = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}}$$

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

4- توزيع المعاينة للنسب

يحتاج الباحث في بعض الدراسات لمعرفة نسبة ظاهرة معينة في المجتمع محل الدراسة، كنسبة المدخنين في مدينة ما، نسبة الوحدات التالفة في انتاج مصنع وغيرها، ففي كل هذه الحالات نجد ان المجتمع محل الدراسة ينقسم الى قسمين، قسم تتوافر فيه الظاهرة محل الدراسة (الخاصية المدروسة) والقسم الثاني لا تتوافر فيه الظاهرة المدروسة، والمجتمعات من هذا النوع يكون فيها المتغير نوعيا لا نستطيع قياسه كميا، وبالتالي تعاد صياغته وتحويله الى متغير عشوائي X، حيث يأخذ المتغير العشوائي القيمة 1 إذا توافرت الظاهرة محل الدراسة في المفردة المدروسة، ويأخذ القيمة 0 إذا لم تتوافر الظاهرة في المفردة المدروسة، ويطلق على هذا النوع من التوزيعات الذي يأخذ قيمتين فقط بتوزيع ذي الحدين، ونتعامل في هذا النوع من المجتمعات مع نسبة الظاهرة محل الدراسة في المجتمع، ويرمز لها بالرمز P ويطلق عليها نسبة المجتمع تحسب كالتالي:

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

وبالتالي فان P تمثل احتمال ظهور هذه الظاهرة في المجتمع، ويرمز لاحتمال عدم ظهور هذه الظاهرة بالرمز Q، حيث ان حدث ظهور الظاهرة وحدث عدم ظهورها هما حدثان مكملان لبعضهما.

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 الظاهرة في العينة بالرمز $\hat{p}$ وتحسب كما يلي:

$$\hat{p} = \frac{x}{n} = \frac{\text{عدد مفردات العينة التى تتحقق فيها الظاهرة}}{\text{العدد الكلي لمفردات العينة}}$$

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

توزيع المعاينة للنسب يقترب من التوزيع الطبيعي وذلك إذا كان حجم العينة المسحوب كبيرا أي:

$$np \geq 5$$
$$n(1-p) \geq 5$$

4-1- حالة مجتمع غير محدود: ليكن X متغير عشوائي منفصل يخضع لتوزيع ثنائي الحدين بمتوسط حسابي μ وتباين σ^2 من مجتمع ما غير محدود، حيث p نسبة المفردات ذات صفة معينة في المجتمع، و $q = 1 - p$ ، ولتكن $\widehat{P}$ متغير عشوائي يمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة ذات الحجم n .

اذا كان المجتمع غير محدود (المعاينة بالإرجاع او عدم الارجاع) و $\frac{n}{N} \leq 0{,}05$

نتحصل على توزيع المعاينة للاحصاءة $\widehat{P}$ وسوف يخضع للتوزيع الطبيعي أي:

$$\widehat{P} \sim N\left(u_{\widehat{P}}, \sigma_{\widehat{P}}^2\right)$$

ويتم حساب معلمتي التوزيع كما يلي[1]:

4-1-1- حالة p للمجتمع معلومة:

$$u_{\widehat{P}} = p$$
$$\sigma_{\widehat{P}}^2 = \frac{pq}{n}$$

[1] . أبو صالح محمد صبحي، **الطرق الإحصائية**، دار اليازوري، عمان، 2007، ص 81.

ويعطى توزيع المعاينة للاحصاءة $\widehat{P}$ كما يلي:

$$\widehat{P} \sim N\left(p, \frac{pq}{n}\right)$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

4-1-2- حالة p للمجتمع مجهولة:

تستعمل نسبة العينة $\hat{p}$ كتقدير لنسبة المجتمع p المجهولة

$$u_{\widehat{P}} = \hat{p}$$
$$\sigma^2_{\widehat{P}} = \frac{\hat{p}(1-\hat{p})}{n}$$

ويعطى توزيع المعاينة للاحصاءة $\widehat{P}$ كما يلي:

$$\widehat{P} \sim N\left(\hat{p}, \frac{\hat{p}(1-\hat{p})}{n}\right)$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

مثال:

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

- أوجد توزيع المعاينة لنسبة المدخنين.
- أحسب احتمال ان تكون نسبة المدخنين في هذه العينة أكبر 0.4.

الحل:

إيجاد توزيع المعاينة للنسبة

لمعرفة طبيعة التوزيع نستخدم شروط التقريب من التوزيع ذي الحدين الى التوزيع الطبيعي:

$$np = 100 * 0.35 = 35 \geq 5$$

ويتم حساب معلمتي التوزيع (المجتمع غير محدود) كما يلي:

$$u_{\widehat{P}} = p = 0.35$$
$$\sigma^2_{\widehat{P}} = \frac{pq}{n} = \frac{0.35 * 0.65}{100} = 0.002275$$

ويعطى توزيع المعاينة للاحصاءة $\widehat{P}$ كما يلي:

$$\widehat{P} \sim N(0.35, 0.002275)$$

حساب احتمال ان تكون نسبة المدخنين في هذه العينة أكبر 0.4

المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

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

4-2- حالة مجتمع محدود: ليكن X متغير عشوائي منفصل يخضع لتوزيع ثنائي الحدين بمتوسط حسابي μ وتباين σ^2 من مجتمع ما محدود، حيث p نسبة المفردات ذات صفة معينة في المجتمع، و $q = 1 - p$ ، ولتكن $\widehat{P}$ متغير عشوائي يمثل نسبة المفردات ذات الصفة المذكورة في العينة المسحوبة ذات الحجم n .

اذا كان المجتمع محدود و $\frac{n}{N} > 0{,}05$

نتحصل على توزيع المعاينة للاحصاءة $\widehat{P}$ وسوف يخضع للتوزيع الطبيعي أي:

$$\widehat{P} \sim N\left(u_{\widehat{P}}, \sigma^2_{\widehat{P}}\right)$$

ويتم حساب معلمتي التوزيع كما يلي[1]:

4-2-1- حالة p للمجتمع معلومة:

$$u_{\widehat{P}} = p$$

$$\sigma^2_{\widehat{P}} = \frac{pq}{n}\left(\frac{N-n}{N-1}\right)$$

ويعطى توزيع المعاينة للاحصاءة $\widehat{P}$ كما يلي:

$$\widehat{P} \sim N\left(p, \frac{pq}{n}\left(\frac{N-n}{N-1}\right)\right)$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

$$Z = \frac{\widehat{P} - p}{\sqrt{\frac{pq}{n}\left(\frac{N-n}{N-1}\right)}} \sim N(0,1)$$

---

[1] . أبو صالح محمد صبحي، مرجع سابق، ص 92.

4-2-2- حالة p للمجتمع مجهولة:

تستعمل نسبة العينة $\hat{p}$ كتقدير لنسبة المجتمع p المجهولة

$$u_{\hat{P}} = \hat{p}$$
$$\sigma_{\hat{P}}^2 = \frac{\hat{p}(1-\hat{p})}{n}\left(\frac{N-n}{N-1}\right)$$

ويعطى توزيع المعاينة للاحصاءة $\hat{P}$ كما يلي:

$$\hat{P} \sim N\left(\hat{p}, \frac{\hat{p}(1-\hat{p})}{n}\left(\frac{N-n}{N-1}\right)\right)$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

مثال:

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

أوجد توزيع المعاينة لنسبة الذكور في الاسر.

أحسب احتمال ان تكون نسبة الذكور في هذا البلد تتراوح بين 51 % و 61 %.

الحل:

إيجاد توزيع المعاينة للنسبة

لمعرفة طبيعة التوزيع نستخدم شروط التقريب من التوزيع ذي الحدين الى التوزيع الطبيعي:

$$np = 200 * 0.59 = 118 \geq 5$$

ويتم حساب معلمتي التوزيع كما يلي:

$$u_{\hat{P}} = p = 0.59$$

المجتمع محدود والسحب مع عدم الارجاع

كسر المعاينة $\frac{n}{N} = \frac{200}{2500} > 0,05$

$$\sigma_{\hat{P}}^2 = \frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right) = \frac{0,59 * 0,41}{200} * \frac{2500-200}{2500-1} = 0,0011$$

ويعطى توزيع المعاينة للاحصاءة $\widehat{P}$ كما يلي:

$$\widehat{P} \sim N(0{,}59{,}0.0{,}0011)$$

حساب احتمال ان تكون نسبة الذكور في هذا البلد تتراوح بين 51 % و 61 %

المتغير الطبيعي المعياري (Z) تصبح علاقته كالتالي:

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

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

5- توزيع المعاينة للفرق بين نسبتين

إذا كانت الدراسة الخاصة بمقارنة ظاهرة معينة في مجتمعين مختلفين، أي محاولة معرفة الفرق بين النسبتان $p_1 - p_2$ ، حيث p_1 ترمز لنسبة الظاهرة في المجتمع الاول، و p_2 ترمز لنسبة الظاهرة في المجتمع الثاني، وعند عدم توافر بيانات عن كل مفردات المجتمع الاول وكل 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 الإحصائية $\widehat{P}_1 - \widehat{P}_2$ حيث $\widehat{P}_1$ نسبة الظاهرة في العينة العشوائية المسحوبة من المجتمع الاول، و $\widehat{P}_2$ نسبة الظاهرة في العينة العشوائية المسحوبة من المجتمع الثاني، ولذلك يجب دراسة توزيع المعاينة لهذه الإحصائية والذي بطلق علبه توزيع المعاينة للفرق بين نسبتي عينتين.

إذا كانت $x_1,x_2,\dots\dots,x_{n1}$ تشكل عينة عشوائية من مجتمع إحصائي حجمه N_1، نسبة عناصرها التي لها الخاصية مدار البحث P_1 ونسبة عناصر العينة التي لها نفس الخاصية $\widehat{P}_1$ ، وكانت $x_1,x_2,\dots\dots,x_{n2}$ تشكل عينة عشوائية من مجتمع إحصائي حجمه N_2، نسبة عناصرها التي لها الخاصية مدار البحث P_2 ونسبة عناصر العينة التي لها نفس الخاصية $\widehat{P}_2$ وكانت العينتين مستقلتين عن بعضهما البعض، وقمنا بحساب الفرق بين نسب عينات المجتمع الاول ونسب عينات المجتمع الثاني، أي حسبنا كل قيم $\widehat{P}_1 - \widehat{P}_2$ فسنحصل على توزيع المعاينة للفرق بين نسبتي العينتين والذي يخضع للتوزيع الطبيعي تقريبا بمتوسط حسابي $u_{\widehat{P}_1-\widehat{P}_2}$ وتباين $\sigma^2{}_{\widehat{P}_1-\widehat{P}_2}$ ونميز الحالتين التاليتين[1]:

[1] . مراد صلاح أحمد، **الأساليب الإحصائية الاستدلالية في العلوم النفسية والتربوية والاجتماعية**، مكتبة الأتجلو، القاهرة، 2000، ص 77.

5-1- حالة مجتمعين غير محدودين: في حالة المجتمعين غير محدودين والسحب الارجاع او عدم الارجاع مع تحقق $\frac{n_1}{N_1} \leq 0,05$ و $\frac{n_2}{N_2} \leq 0,05$

يتم حساب معلمتي التوزيع كما يلي:

$$u_{\hat{P}_1-\hat{P}_2} = p_1 - p_2$$
$$\sigma^2{}_{\hat{P}_1-\hat{P}_2} = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \sim N(0,1)$$

5-2- حالة مجتمعين غير محدودين: في حالة المجتمعين غير محدودين والسحب مع عدم الارجاع مع تحقق $\frac{n_1}{N_1} > 0,05$ و $\frac{n_2}{N_2} > 0,05$

يتم حساب معلمتي التوزيع كما يلي:

$$u_{\hat{P}_1-\hat{P}_2} = p_1 - p_2$$
$$\sigma^2{}_{\hat{P}_1-\hat{P}_2} = \frac{p_1(1-p_1)}{n_1}\left(\frac{N_1-n_1}{N_1-1}\right) + \frac{p_2(1-p_2)}{n_2}\left(\frac{N_2-n_2}{N_2-1}\right)$$

اما المتغير الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}\left(\frac{N_1-n_1}{N_1-1}\right) + \frac{p_2(1-p_2)}{n_2}\left(\frac{N_2-n_2}{N_2-1}\right)}} \sim N(0,1)$$

بالنسبة لكسر المعاينة $\frac{n}{N}$ للمجتمع الاول والثاني ليس بالضرورة ان يتحققا معا، فقد يتحقق الأول ولا يتحقق الثاني او العكس.

مثال:

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 الأول والثانية بالنوع الثاني وتم اختيار عينتين من الحشرات حجم كل منها 200 ووضعت كل عينة في حجرة.

- اوجد توزيع المعاينة للفرق بين نسبتي العينتين.
- أحسب احتمال ان يكون الفرق ما بين نسبتي العينتين ما بين 15 و18%.

الحل:

إيجاد توزيع المعاينة للفرق بين نسبتي العينتين

بما ان المجتمعين يخضعان الى توزيع مجهول ولمعرفة توزيع المعاينة نستعين بنظرية النهاية المركزية.

المجتمع الأول $n_1p_1 = 0,9 * 200 = 180 \geq 5$ و $n_1q_1 = 0,1 * 200 = 20 \geq 5$

اذن توزيع المعاينة للنسبة $\widehat{P_1}$ يخضع للتوزيع الطبيعي

المجتمع الثاني $n_2p_2 = 0,8 * 200 = 160 \geq 5$ و $n_2q_2 = 0,2 * 200 = 40 \geq 5$

اذن توزيع المعاينة للنسبة $\widehat{P_2}$ يخضع للتوزيع الطبيعي

ما دام توزيع المعاينة للمجتمع الاول يخضع للتوزيع الطبيعي وتوزيع المعاينة للمجتمع الثاني يخضع للتوزيع الطبيعي فان توزيع المعاينة للفرق يعتبر متغير عشوائي جديد يخضع هو الاخر للتوزيع الطبيعي كالتالي:

$$\widehat{P_1} - \widehat{P_2} \sim N(u_{\widehat{P}_1-\widehat{P}_2}, \sigma^2{}_{\widehat{P}_1-\widehat{P}_2})$$

يتم حساب معلمتي التوزيع كما يلي:

$$u_{\widehat{P}_1-\widehat{P}_2} = p_1 - p_2 = 0.9 - 0.8 = 0.1$$

$$\sigma^2{}_{\widehat{P}_1-\widehat{P}_2} = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} = \frac{0.9 * 0.1}{200} + \frac{0.8 * 0.2}{200} = 0.00125$$

اذن توزيع المعاينة يخضع للتوزيع الطبيعي كالتالي:

$$\widehat{P_1} - \widehat{P_2} \sim N(0.1,0.00125)$$

حساب احتمال ان يكون الفرق ما بين نسبتي العينتين ما بين 15 و18%

نستعين بالمتغير الطبيعي المعياري

$$P(0.15 < \hat{P}_1 - \hat{P}_2 < 0.18) = P\left(\frac{0.15 - 0.10}{\sqrt{0.00125}} < Z < \frac{0.18 - 0.10}{\sqrt{0.00125}}\right)$$
$$= P(01.41 < Z < 02.26) = F(02.26) - F(01.41) = 0.98809 - 0.92073$$
$$= 0.06736$$

6- توزيع المعاينة للتباين

من المعلوم ان التباين S^2 احد مقاييس التشتت فهو يقيس مدى تشتت الظاهرة حول متوسطها الحسابي وكلما كان هذا التشتت ضعيفا كانت الظاهرة الإحصائية قريبة من المتوسط الحسابي لها وكلما زاد التشتت كانت الظاهرة الاحصائية بعيدة عن المتوسط الحسابي لها، وعليه فان تباين العينة S^2 يستخدم للاستدلال حول تباين المجتمع σ^2 .

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

اذا كان حجم العينة $n \geq 30$ فان: $S^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}$

اذا كان حجم العينة $n < 30$ فان: $\hat{S}^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$

إذا كانت $x_1, x_2, \ldots, x_{n1}$ تشكل عينة عشوائية حجمها n سحبت من مجتمع يتوزع توزيع طبيعي بمتوسط حسابي μ وتباين σ^2، فان تباين العينة S^2 او $\hat{S}^2$ سيتوزع توزيع كاي تربيع χ^2 بدرجة حرية n−1 وذلك كما يلي: $\frac{(n-1)\hat{S}^2}{\sigma^2} = \frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$

إذا كان X متغير عشوائي يخضع للتوزيع الطبيعي بمتوسط حسابي μ وتباين σ^2 من مجتمع ما، ولتكن S^2 متغيرة عشوائية تمثل تباين عينة حجمها n فان توزيع المعاينة للتباين نميز بين الحالات التالية[1]:

6−1− حالة مجتمع محدود او غير محدود والمعاينة بالإرجاع: تعطى معالم التوزيع كما يلي:

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

$$\mu_{\hat{S}^2} = \sigma^2$$

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

$$\sigma^2{}_{\hat{S}^2} = \frac{2}{n-1}\sigma^4$$

6−2− حالة مجتمع محدود او غير محدود والمعاينة مع عدم الإرجاع و $\frac{n}{N} \leq 0,05$: في هذه الحالة لا يتحقق كسر المعاينة وبالتالي تبقى العلاقات كما هي في حالة السحب مع الارجاع.

6−3− حالة مجتمع محدود والمعاينة مع عدم الإرجاع و $\frac{n}{N} > 0,05$: تعطى معالم التوزيع كما يلي:

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

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

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

---

[1] . مراد صلاح أحمد، مرجع سابق، ص 87.

$$\sigma^2{}_{\hat{S}^2} = \frac{2}{n-1}\sigma^4$$

إذا كان المتغير العشوائي X لا يخضع الى التوزيع الطبيعي فان تباين العينة يعطى بالعلاقة التالية:

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

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

مثال:

إذا كان لدينا مجتمع كبير جدا يخضع للتوزيع الطبيعي بمتوسط حسابي 25 وانحراف معياري 3.

– احسب المتوسط الحسابي والتباين لكل من توزيع المعاينة لتباين العينة وتباين العينة المعدل لجميع العينات الممكنة ذات الحجم 09 والتي يمكن سحبها من هذا المجتمع.

الحل:

تعطى معالم توزيع المعاينة لتباين العينة كما يلي:

$$\mu_{S^2} = \sigma^2\left(\frac{n-1}{n}\right) = 3^2\left(\frac{9-1}{9}\right) = 10.125$$

$$\sigma^2{}_{S^2} = \frac{2(n-1)}{n^2}\sigma^4 = \frac{2(9-1)}{9^2}3^4 = 16$$

تعطى معالم توزيع المعاينة لتباين العينة المعدل كما يلي:

$$\mu_{\hat{S}^2} = \sigma^2 = 3^2 = 9$$

$$\sigma^2{}_{\hat{S}^2} = \frac{2}{n-1}\sigma^4 = \frac{2}{9-1}3^4 = 20.25$$

7– توزيع المعاينة للنسبة بين تبايني عينتين

اذا كان لدينا تبايني عينتين عشوائيتين S_1^2 و S_2^2 بحجم n_1 وn_2 مسحوبتين من مجتمعين يتوزعان توزيعا طبيعيا بتباين σ_1^2 و σ_2^2 على التوالي فان النسبة $\frac{S_1^2}{S_2^2}$ لتباين العينتين تتوزع توزيع فيشر F بدرجتي حرية V_1 وV_2 حيث أن:

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

ومما تقدم يمكن الحصول على المتغير العشوائي F كما يلي[1]:

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

مثال:

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 $\sigma_1^2 = 9$ و $\sigma_2^2 = 16$ على التوالي.

– أحسب الاحتمال التالي: $P\left(\frac{S_1^2}{S_2^2} \geq 01,74\right)$

الحل:

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

نحسب الان الاحتمال بعد حساب إحصاءات فيشر كما يلي:

$$P\left(\frac{S_1^2}{S_2^2} > 01.74\right) = P\left[\frac{S_1^2\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{S_2^2\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}} > 01.74\frac{\left(\frac{n_1}{n_1-1}\right)\frac{1}{\sigma_1^2}}{\left(\frac{n_2}{n_2-1}\right)\frac{1}{\sigma_2^2}}\right]$$

$$= P\left(F_{15,20} > 01.74\frac{\left(\frac{16}{16-1}\right)\frac{1}{09}}{\left(\frac{21}{21-1}\right)\frac{1}{16}}\right) = P(F_{15,20} > 03.14)$$

$$= 01 - P(F_{15,20} < 03.14) = 01 - 0.99 = 0.01$$

[1] . مراد صلاح أحمد، مرجع سابق، ص 96.

تمارين محلولة

التمرين (01):

إذا كان لدينا مجتمع إحصائي مؤلف من العاملين في منشأة صغيرة حجمه N=4 مفردة، ومكون من القيم { 6 , 4 , 2 , 0 }. اوجد المتوسط الحسابي لتوزيع المعاينة للمتوسطات الحسابية.

الحل:

$$\mu = 3$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N} = 5$$

نفرض أننا سحبنا جميع العينات الممكنة مع الإعادة ذات الحجم $n = 2$ ثم حسبنا متوسطاتها

عدد العينات الممكن سحبها مع الإعادة يعطى بالعلاقة: $N^n = 4^2 = 16$

وإن متوسطات العينات العشوائية المسحوبة تتأرجح حسب الجدول التالي:

رقم العينة	العينة		المتوسط	رقم العينة	العينة		المتوسط
1	0	0	0	9	4	0	2
2	0	2	1	10	4	2	3
3	0	4	2	11	4	4	4
4	0	6	3	12	4	6	5
5	2	0	1	13	6	0	3
6	2	2	2	14	6	2	4
7	2	4	3	15	6	4	5
8	2	6	4	16	6	6	6

إن الجدول الاحتمالي لتوزيع المعاينة للمتوسطات الحسابية

	0	1	2	3	4	5	6
(P)	1/16	2/16	3/16	4/16	3/16	2/16	1/16

ولو رسمنا المدرج التكراري، سنلاحظ أن توزيع المعاينة للمتوسطات الحسابية للعينات يمكن أن يقترب وبشكل جيد من منحني التوزيع الطبيعي.

$$\mu_{\bar{x}} = 0\frac{1}{16} + 1\frac{2}{16} + 2\frac{3}{16} + 3\frac{4}{16} + 4\frac{3}{16} + 5\frac{2}{16} + 6\frac{1}{16} = \mu = 3$$

التمرين (02):

أخذت عينة عشوائية حجمها 9 من الأطفال حديثي الولادة في أحد المستشفيات، فإذا علم أن وزن الطفل حديث الولادة يخضع لتوزيع طبيعي بمتوسط 2900 غرام وانحرافه المعياري 600 غرام.

أوجد متوسط وتباين والانحراف المعياري لتوزيع المعاينة للمتوسطات الحسابية لأوزان الأطفال.

أوجد احتمال أن المتوسط الحسابي لأوزان الأطفال في العينة يزيد عن 3100 غرام.

أوجد احتمال أن المتوسط الحسابي لأوزان الأطفال في العينة يقع بين 2700 و 3200 غرام.

الحل:

ايجاد متوسط وتباين والانحراف المعياري لتوزيع المعاينة للمتوسطات الحسابية لأوزان الأطفال

$$\bar{X} \sim N(\mu_{\bar{x}}, \sigma^2_{\bar{x}})$$
$$\mu_{\bar{x}} = \mu = 2900$$
$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = \frac{600^2}{9} = 40000$$
$$\sigma = \sqrt{40000} = 200$$

ايجاد احتمال أن المتوسط الحسابي لأوزان الأطفال في العينة يزيد عن 3100 غرام.

$$\begin{aligned} P(\bar{X} > 3100) &= P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{3100-\mu}{\sigma/\sqrt{n}}\right) \\ &= P\left(Z > \frac{3100-2900}{600/\sqrt{9}}\right) \\ &= P\left(Z > \frac{200}{200}\right) \\ &= P(Z > 1) = 1 - P(Z < 1) = 1 - 0.8413 = 0.1587 \end{aligned}$$

إيجاد احتمال أن المتوسط الحسابي لأوزان الأطفال في العينة يقع بين 2700 و 3200 غرام.

$$\begin{aligned} P(2700 < \bar{X} < 3200) &= P\left(\frac{2700-\mu}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{3200-\mu}{\sigma/\sqrt{n}}\right) \\ &= P\left(\frac{2700-2900}{600/\sqrt{9}} < Z < \frac{3200-2900}{600/\sqrt{9}}\right) \\ &= P\left(\frac{-200}{200} < Z < \frac{300}{200}\right) \\ &= P(-1 < Z < 1.5) \\ &= P(Z < 1.5) + P(Z < 1) - 1 \\ &= 0.9332 + 0.8413 - 1 = 0.7745 \end{aligned}$$

التمرين (03):

إذا كانت درجات طلبة الإحصاء تأخذ شكل التوزيع الطبيعي بمتوسط حسابي مقداره 70 درجة.

أخذت عينة حجمها 9 طلبة، ووجد أن الانحراف المعياري لعلاماتهم 11 درجة.

احسب احتمال أن يزيد متوسط درجاتهم عن 75 درجة .

الحل:

توزيع المعاينة للمتوسط الحسابي يخضع لتوزيع ستيودنت T (توزيع العينات الصغيرة).

$$\begin{aligned} P(\bar{X} > 75) &= P\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} > \frac{75-\mu}{S/\sqrt{n}}\right) \\ &= P\left(T > \frac{75-70}{11/\sqrt{9}}\right) \qquad ,T\sim t_8 \\ &= P\left(T > \frac{5}{11/3}\right) \\ &= P\left(T > \frac{15}{11}\right) \\ &\approx P(T > 1.363) \\ &\approx P(T > 1.397) \\ &\approx 10\% \end{aligned}$$

التمرين (04):

اذا كان التوزيع لعمر جهاز كهربائي يقترب من التوزيع الطبيعي بمتوسط حسابي 800 ساعة وانحراف معياري 40 ساعة، سحبت عينة من 16 جهاز.

ما هو احتمال ألا يزيد المتوسط الحسابي عن 775 ساعة ؟

الحل:

توزيع المعاينة يخضع للتوزيع الطبيعي لحساب الاحتمال نحسب القيمة المعيارية او الاحصاءة Z

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

$$\mu_{\bar{x}} = \mu = 800$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{40^2}{16} = 100$$

$$\begin{aligned} P(\overline{X} < 775) \quad &= P\left(Z < \frac{775-800}{\sqrt{100}}\right) = P(Z < -2{,}5) \\ &= P(Z > 2{,}5) = 0.0062 \end{aligned}$$

التمرين (05):

اذا كان الانتاج لشركة الاولى A لأجهزة كهربائية يمتلك متوسط عمر 6.5 سنة بانحراف معياري 0.9 سنة اما الشركة الثانية B كان المتوسط 6 سنة بانحراف معياري 0.8 سنة.

اذا تم سحب عينة بحجم 36 من A و49 من B ما هو احتمال ان يكون على الاقل فرق المتوسط سنة واحدة ؟

الحل:

توزيع المعاينة يخضع للتوزيع الطبيعي المعياري (Z) فتصبح علاقته كالتالي:

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

$$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{X}_1} - \mu_{\overline{X}_2} = \mu_1 - \mu_2 = 6,5 - 6 = 0,5$$

$$\sigma^2{}_{\overline{x_1}-\overline{x_2}} = \sigma_{\overline{X}_1}^2 + \sigma_{\overline{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{0.9^2}{36} + \frac{0,8^2}{49} = 0,0225 + 0.0130 = 0.0355$$

$$P(\overline{X}_1 - \overline{X}_2 > 1) \quad = P\left(Z > \frac{1-0,5}{\sqrt{0.0355}}\right) = P(Z > 2,77) = 0.0028$$

تمارين غير محلولة

التمرين (01):

1 /العبارة الصحيحة من بين العبارات التالية هي:

أ) الاستدلال الإحصائي من أهم الأدوات لاتخاذ القرارات، ويشمل اختبار الفرضيات وجمع البيانات

ب) الاستدلال الإحصائي من أهم الأدوات لاتخاذ القرارات، ويشمل اختبار الفرضيات وعرض البيانات.

ج) الاستدلال الإحصائي من أهم الأدوات لاتخاذ القرارات، ويشمل اختبار الفرضيات والتقدير.

د) الاستدلال الإحصائي من أهم الأدوات لاتخاذ القرارات، ويشمل التقدير وحساب المتوسط لبعض البيانات.

2 /العبارة الصحيحة من بين العبارتين التاليتين:

أ) لا بد للحصول على تقدير سليم لمعالم مجتمع ما أن يتم اختيار عينة ممثلة لذلك المجتمع.

ب) ليس هناك حاجة لأن يتم اختيار عينة ممثلة لمجتمع ما للحصول على تقدير سليم لمعالم لك المجتمع.

3 /العبارة الصحيحة من بين العبارتين التاليتين:

أ) العينة العشوائية هي العينة التي لا يكون لكل مفردة من مفردات المجتمع نفس فرصة الاختيار في العينة.

ب) العينة العشوائية هي العينة التي يكون لكل مفردة من مفردات المجتمع نفس فرصة الاختيار في العينة.

4 /أي مجموعة من المفردات تشترك في صفة أو صفات وتكون موضوع دراسة أو بحث؛ فإن هذه المجموعة يطلق عليها إحصائيا:

أ) مجتمع الدراسة

ب) عينة الدراسة

5 /تصلح العبارة (تجميع البيانات عن كل مفردة من مفردات المجتمع، وهذا الأسلوب يتطلب وفرة في الوقت والمال والمجهود) لوصف:

أ) الحصر الشامل

ب) العينة العشوائية

ج) العينة المنتظمة

د) العينة العنقودية

6 /أي من الأسباب التالية يعد سببا في خطأ المعاينة العشوائية:

أ) الاختيار غير العشوائي للعينة

ب) التحيز المرصود

ج) استبدال وحدة بوحدة أخرى غيا مدرجة ضمن الإطار العام للدراسة

د) ليس من أي الأسباب أعلاه، وإنما هي الصدفة

7 /إذا كان المجتمع غير معروف، وكان متجانسا؛ فيمكن للباحث أن يستخدم طريقة:

أ) العينة الحصية

ب) العينة العمدية

8 /إذا كان المجتمع معروفا، وكان متجانسا؛ فيمكن للباحث أن يستخدم طريقة:

أ) العينة الطبقية

ب) العينة العنقودية

9 /إذا كان المجتمع معروفا، وكان غير متجانس؛ فيمكن للباحث أن يستخدم طريقة:

أ) العينة الطبقية

ب) العينة العنقودية

10 /يفترض أن يؤدي تدريب الباحثين بشكل جيد على جمع البيانات والتقيد بالتعليمات إلى:

أ) تقليل أخطاء البيانات الإحصائية الناتجة عن التحيز

ب) تقليل أخطاء البيانات الإحصائية الناتجة عن الصدفة

التمرين (02):

إذا كان لدينا مجتمع حجمه 5 مفردات ونريد أخذ عينة حجمها 3 مفردات.

احسب عدد العينات الممكنة.

التمرين (03):

بافتراض أن لدينا مجتمع حجمه 4 مفردات هي 6، 3، 5، 2 وأردنا سحب عينة من هذا المجتمع حجمها مفردتين أوجد:

الوسط الحسابي والتباين للمجتمع.

إيجاد جميع العينات الممكن سحبها من المجتمع.

تكوين توزيع المعاينة

حساب كل من المتوسط الحسابي لتوزيع المعاينة للمتوسطات ومقارنته بالمتوسط الحسابي للمجتمع.

أحسب التباين للمجتمع وكذلك الانحراف المعياري لتوزيع المعاينة للمتوسطات.

التمرين (04):

إذا كان لدينا مجتمع يتكون من 5 مفردات وأردنا سحب عينة عشوائية من هذا المجتمع حجمها 3 مفردات مع ملاحظة أن مفردات المجتمع هي 1، 2، 3، 4، 5.

أحسب المتوسط الحسابي والتباين لمفردات المجتمع.

كون توزيع معاينة للمتوسطات إذا كان السحب بدون إرجاع.

أحسب المتوسط الحسابي للعينات التي تم سحبها وقارن بينه وبين المتوسط الحسابي للمجتمع.

أحسب الخطأ المعياري لتوزيع المعاينة للمتوسطات.

التمرين (05):

سحبت عينة عشوائية من مجتمع لا نهائي معدله 70 وتباينه 40. إذا كان حجم العينة 10، فأوجد:

المتوسط الحسابي لتوزيع المعاينة للمتوسطات.

تباين توزيع المعاينة للمتوسطات.

التمرين (06):

تخضع علامات الطلاب في أحد المواد لتوزيع طبيعي متوسطه 65 وانحراف معياري 18. اخذت عينة عشوائية حجمها 36 طالب، احسب:

احتمال أن يزيد متوسط علامات العينة على 74.

التمرين (07):

إذا كانت أطوال الطلاب في أحد الصفوف المدرسية تتبع التوزيع الطبيعي بمتوسط يساوي 160سم، إذا سحبت عينة عشوائية من 4 طلاب فما احتمال أن يقل متوسطها الحسابي عن 166سم، إذا علمت أن الانحراف المعياري للعينة يساوي 10سم.

التمرين (08):

تخضع علامات الناجحين من امتحان الدراسة الثانوية في مدرسة ما لتوزيع طبيعي معدله 74 وانحرافه المعياري 12، وفي مدرسة أخرى تخضع لتوزيع الطبيعي معدله 70 وانحرافه المعياري 16، اخذت عينة عشوائية حجمها 16 طالب من المدرسة الأولى و9 طلاب من المدرسة الثانية أوجد:

احتمال أن يزيد الفرق بين المتوسطين عن 8.

الفصل الثاني

نظرية التقدير

تمهيد:

يقصد بالتقدير ان نقدر معالم المجتمع المجهولة عن طريق بيانات العينة المتاحة، ويقصد بمعالم المجتمع المجهولة المؤشرات أو الأدلة مثل: المتوسط الحسابي، كمتوسط عمر الفرد في دولة ما، متوسط دخل الاسرة في مدينة ما، هذه جميعها تسمى مؤشرات في مجتمع وهي مجهولة، نستطيع تقديرها عن طريق سحب عينة من المجتمع وحساب ما يقابل تلك المؤشرات بالعينة.

فإذا كنا نرغب في تقدير أحد معالم المجتمع وليكن θ عن طريق عينة من المشاهدات مسحوبة من المجتمع فإن القيمة التي يتم حسابها للمعلمة θ من واقع المشاهدات تسمى تقديرا، بينما الدالة أو الصيغة الرياضية الاحصائية التي تستخدم للوصول إلى هذا التقدير تسمى مقدرا، والمقدر هو دالة تعتمد على المشاهدات، 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.

والتقدير نوعان: التقدير بنقطة او التقدير وحيد القيمة، والتقدير بفترة ثقة.

1- التقدير بنقطة

1-1- مفهوم التقدير بنقطة: يقصد به 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، فمثلا نقدر المتوسط الحسابي للمجتمع μ بالمتوسط الحسابي للعينة $\bar{x}$ ، ونقدر تباين المجتمع σ^2 بتباين العينة S^2 ولذلك يسمى التقدير بقيمة وحيدة[1].

إذا كانت المعلمة المجهولة هي المتوسط الحسابي للمجتمع μ فيقدر بالمتوسط الحسابي للعينة، أي ان:

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

واذا كانت المعلمة المجهولة هي تباين المجتمع σ^2 فيقدر بتباين العينة S^2 أي ان:

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

[1] . علام صلاح الدين، **الأساليب الإحصائية الاستدلالية البارامترية واللابارامترية**، دار الفكر العربي، القاهرة، 2005، ص 43.

1-2- خصائص المقدر الجيد: لكي يكون المقدر جيد يجب أن يحقق بعض المعايير، وترجع كلمة المقدر الجيد إلى العالم فيشر، ويعتبر المقدر مقدرا جيدا طبقا لمعيار فيشر إذا توافرت فيه الخصائص التالية[1]:

1-2-1- خاصية عدم التحيز: 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 مسحوبة من المجتمع فان المقدر $\hat{\theta}$ غير متحيز اذا كان: $E(\hat{\theta}) = \theta$.

كما أنه يفضل التقدير الذي له اقل متوسط مربع خطأ.

1-2-2- خاصية الاتساق: يقال بأن $\hat{\theta}$ مقدرا متسقا لمعلمة المجتمع θ إذا كانت $\hat{\theta}$ تؤول إلى θ (أي تقترب منها) كلما زاد حجم العينة، ويكون المقدر متسقا اذا كان:

$$\begin{cases} \lim_{n\to\infty} E(\hat{\theta}) = \theta \\ \lim_{n\to\infty} V(\hat{\theta}) = 0 \end{cases}$$

1-2-3- خاصية الكفاءة: 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 فان المقدر $\hat{\theta}_1$ اكثر كفاءة من المقدر $\hat{\theta}_2$ اذا تحقق ما يلي:

$$e = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)} < 1$$

1-2-4- خاصية الكفاية: يقال أن المقدر $\hat{\theta}$ مقدرا كافيا للمعلمة θ إذا كان التقدير $\hat{\theta}$ 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 .

1-2-5- خاصية التقارب: يقال أن المقدر $\hat{\theta}_1$ أكثر تقاربا من $\hat{\theta}_2$ في تقديره للمعلمة θ اذا كان احتمال أو درجة الثقة للأول أكبر من الثاني أي:

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

2- التقدير بفترة

إن المقدر المستخدم عند تقدير معلمة المجتمع تختلف نتائجه التقديرية من عينة إلى أخرى، 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

[1] . علام صلاح الدين، مرجع سابق، ص 48.

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 مستوى الثقة.

التقدير بفترة هو التقدير الذي يتألف من قيمتين عدديتين تحددان مجالا من القيم الذي نتوقع أن يتضمن المعلمة المطلوب تقديرها بمستوى ثقة معلوم.

وفقا لهذا التعريف لإيجاد فترة الثقة للمعلمة المجهولة ولتكن مثلا θ يجب إيجاد إحصاءتين مثلا U و L بحيث يكون: $P(L \le \theta \le U) = 1 - \alpha$

ويطلق على $(1-\alpha)\%$ فترة الثقة لتقدير المعلمة المجهولة للمجتمع θ.[1]

3- التقدير بفترة للمتوسط الحسابي

3-1- التقدير بفترة للمتوسط الحسابي عندما يكون تباين المجتمع معلوم

إذا كان لدينا مجتمع غير محدود بمتوسط حسابي μ وتباين σ^2 ليس بالضرورة يتوزع توزيعا طبيعيا وسحبنا منه كل العينات العشوائية ذات الحجم n، بحيث n كبيرة $n \ge 30$ فتوزيع المعاينة للمتوسط الحسابي للعينة $\bar{X}$ سيكون قريبا من التوزيع الطبيعي، حسب نظرية النهاية المركزية، بمتوسط حسابي يساوي $\mu_{\bar{x}} = \mu$ وتباين يساوي $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ أي: $\bar{X} \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}^2)$ ، أما اذا كان المجتمع يتوزع توزيعا طبيعيا فتوزيع المعاينة للمتوسط الحسابي للعينة $\bar{X}$ سيتوزع توزيعا طبيعيا، وذلك سواء كان حجم العينة n صغيرا أو كبيرا، وفي هذه الحالة يمكن تحويل المتغير $\bar{X}$ الى المتغير الطبيعي المعياري Z كالتالي:[2]

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

ويعطى مجال الثقة للمتوسط الحسابي للمجتمع عندما يكون تباين المجتمع معلوما او عند مستوى ثقة $(1-\alpha)\%$ كالتالي: $P\left(\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$

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

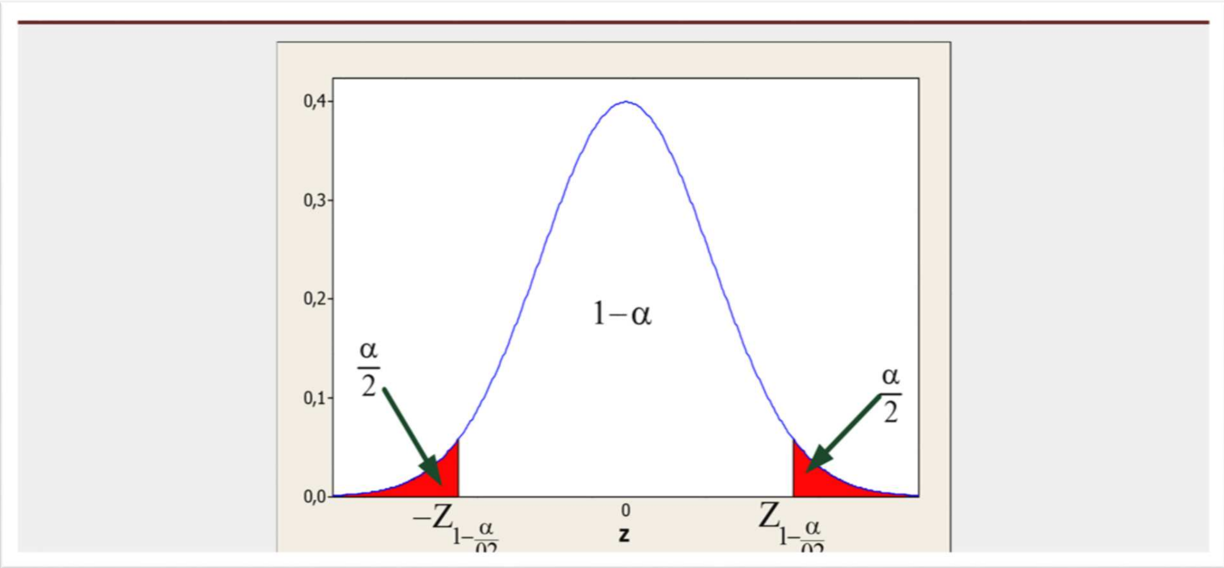
يسمى المقدار $\bar{X} + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ بالحد الاعلى لفترة الثقة.

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

والشكل الموالي يوضح المعطيات السابقة بيانيا.

[1] . علام صلاح الدين، مرجع سابق، ص 54.

[2] . المنيزل عبد الله فلاح، مرجع سابق، ص 74.



أما اذا كان المجتمع محدود والسحب مع عدم الارجاع مع كسر المعاينة $\frac{n}{N} > 0.05$ يعطى مجال الثقة للمتوسط الحسابي للمجتمع كالتالي:

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

والجدول التالي يوضح مختلف القيم والدرجات المعيارية التي يأخذها مجال الثقة كالتالي:

$\mathbf{Z_{1-\frac{\alpha}{2}}}$	$\boldsymbol{\alpha}$	$\mathbf{(1-\alpha)}$	$\mathbf{(1-\alpha)\%}$
1.645	0.1	0.9	%90
1.96	0.05	0.95	%95
2.575	0.01	0.99	%99

مثال:

سحبت عينة عشوائية من 100 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.

– اوجد تقدير نقطة لمتوسط عمر المصباح.

– اوجد تقدير فترة الثقة 95 % لمتوسط عمر المصباح في هذا المصنع، وماذا تستنتج ؟

الحل:

إيجاد تقدير نقطة لمتوسط عمر المصباح

$\hat{\mu} = \bar{X} = 1000$

إيجاد تقدير نقطة لمتوسط عمر المصباح

بما ان حجم العينة أكبر من 30 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 لمتوسط مجتمع معلوم التباين بالعلاقة التالية:

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

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} = 1000 - \left(1.96 * \frac{150}{\sqrt{100}}\right) = 970.6$$
$$\bar{X} + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} = 1000 + \left(1.96 * \frac{150}{\sqrt{100}}\right) = 1029.4$$

بالتعويض في مجال التقدير نجد:

$$P(970.6 \le \mu \le 1029.4) = 0.95$$

الاستنتاج: نحن متأكدون بدرجة ثقة 95 % أن عمر المصباح سوف لن يقل عن 970.60 ساعة ولن يزيد عن 1029.40 ساعة.

3-2- التقدير بفترة للمتوسط الحسابي عندما يكون تباين المجتمع مجهول

بالرغم من انه في بعض الأحيان يكون المتوسط الحسابي غير معلوم وتباين المجتمع معلوم، إلا انه في الحقيقة وفي كثير من الحالات يكون تباين المجتمع غير معلوم أيضا، وإذا أردنا تكوين فترة ثقة حول المتوسط الحسابي عندما يكون تباين المجتمع غير معلوم، فان المدخل المناسب هو تقدير تباين المجتمع باستخدام تباين العينة، وهنا نميز حالتين حسب حجم العينة.

3-2-1- حالة حجم العينة أكبر أو يساوي 30: إذا كان لدينا مجتمع غير محدود يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 مجهول، وسحبنا عينة عشوائية ذات الحجم n حيث $n \ge 30$، في هذه الحالة يؤخذ تباين العينة S^2 كتقدير غير متحيز لتباين المجتمع، يعطى مجال الثقة للمتوسط الحسابي للمجتمع عندما يكون تباين المجتمع مجهولا عند مستوى ثقة $(1-\alpha)\%$ كالتالي[1]:

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

أو بشكل مكافئ كما يلي:

$$P\left(\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{\hat{S}}{\sqrt{n}} \le \mu \le \bar{X} + Z_{1-\frac{\alpha}{2}}\frac{\hat{S}}{\sqrt{n}}\right) = 1-\alpha$$

[1] . بدر سالم عيسى، مرجع سابق، ص 89.

أما ذا كانت المعاينة من مجتمع محدود والسحب مع عدم الإرجاع مع تحقق العلاقة $\frac{n}{N} > 0.05$ يعطى مجال الثقة للمتوسط الحسابي للمجتمع عندما يكون تباين المجتمع مجهولا عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

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

أو بشكل مكافئ كما يلي:

$$P\left(\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{\hat{S}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{X} + Z_{1-\frac{\alpha}{2}}\frac{\hat{S}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right) = 1-\alpha$$

مثال:

نظرا للشكاوى التي تتلقاها إدارة شركة منتجة لسائل كيميائي بخصوص العمر الزمني الذي يعمره السائل، قامت باختبار عينة عشوائية تتكون من 50 قارورة سائل من بين عدد كبير جدا من القارورات فوجدت أن متوسط العمر الزمني لهذه العينة يساوي 02.67 سنة وانحراف معياري يساوي 01.94 سنة.

– أوجد فترة ثقة حول متوسط العمر الزمني الذي تعمره هذه القارورات عند مستوى 95 %.

الحل:

بما أن تباين المجتمع مجهول وحجم العينة كبير (اكبر من 30)، إذن يمكن استخدام التوزيع الطبيعي المعياري، بتطبيق العلاقة التالية:

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

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}} = 2.67 - \left(1.96 * \frac{1.94}{\sqrt{50-1}}\right) = 2.1268$$

$$\bar{X} + Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}} = 2.67 + \left(1.96 * \frac{1.94}{\sqrt{50-1}}\right) = 3.2132$$

بالتعويض في مجال التقدير نجد:

$$P(2.1268 \leq \mu \leq 3.2132) = 0.95$$

3-2-2- حالة حجم العينة أقل من 30: إذا كان لدينا مجتمع غير محدود يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 مجهول، وسحبنا عينة عشوائية ذات الحجم n حيث $n < 30$، في هذه الحالة يؤخذ

تباين العينة S^2 كتقدير غير متحيز لتباين المجتمع، يعطى مجال الثقة للمتوسط الحسابي للمجتمع عندما يكون تباين المجتمع مجهولا عند مستوى ثقة $(1-\alpha)\%$ كالتالي[1]:

$$P\left(\bar{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\leq\mu\leq\bar{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\right)=1-\alpha$$

أو بشكل مكافئ كما يلي:

$$P\left(\bar{X}-T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}}{\sqrt{n}}\leq\mu\leq\bar{X}+T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}}{\sqrt{n}}\right)=1-\alpha$$

أما ذا كانت المعاينة من مجتمع محدود والسحب مع عدم الإرجاع مع تحقق العلاقة $\frac{n}{N}>0.05$ يعطى مجال الثقة للمتوسط الحسابي للمجتمع عندما يكون تباين المجتمع مجهولا عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

$$P\left(\bar{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\leq\mu\leq\bar{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\sqrt{\frac{N-n}{N-1}}\right)=1-\alpha$$

أو بشكل مكافئ كما يلي:

$$P\left(\bar{X}-T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\leq\mu\leq\bar{X}+T_{1-\frac{\alpha}{2},n-1}\frac{\hat{S}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right)=1-\alpha$$

مثال:

إذا علمت أن متوسط أوزان عينة تتكون من 16 طفلا أعمارهم 10 سنوات يساوي 32.40 كغ، وبانحراف معياري يساوي 05.40 كغ.

أوجد 99 % فترة ثقة لتقدير المتوسط الحسابي للأوزان على افتراض أن مجتمع المعاينة مجتمع طبيعي.

الحل:

بما أن المجتمع يخضع للتوزيع الطبيعي بتباين مجهول وحجم العينة المأخوذة اقل من 30 فان عملية التقدير يتم إجراؤها باستخدام توزيع ستيودنت (T) بالعلاقة التالية:

$$P\left(\bar{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\leq\mu\leq\bar{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\right)=1-\alpha$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\bar{X}-T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}=32.4-\left(2.947*\frac{5.4}{\sqrt{16-1}}\right)=28.291$$

[1] . بدر سالم عيسى، مرجع سابق، ص 94.

$$\bar{X}+T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}=32.4+\left(2.947*\frac{5.4}{\sqrt{16-1}}\right)=36.509$$

بالتعويض في مجال التقدير نجد:

$$P(28.291\leq\mu\leq 36.509)=0.99$$

4- التقدير بفترة للفرق بين متوسطي مجتمعين

نميز هنا بين حالتين هما:

4-1- التقدير بفترة للفرق بين متوسطي مجتمعين مستقلين

لإيجاد فترة الثقة حول الفرق بين متوسطي مجتمعين مستقلين فإننا نميز الحالات التالية[1]:

4-1-1- التقدير بفترة للفرق بين متوسطي مجتمعين معلومي التباين: إذا افترضنا عينة عشوائية ذات حجمها n_1 مسحوبة من مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 معلوم، وعينة عشوائية حجمها n_2 من مجتمع يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 معلوم، وكانت العينتين مستقلتين عن بعضهما البعض، او كان المجتمعين غير طبيعيين بشرط حجم كل عينة اكبر او تساوي 30، فان توزيع المعاينة للفرق بين المتوسطات الحسابية $\overline{X_1}-\overline{X_2}$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي $\mu_{\overline{x_1}-\overline{x_2}}=\mu_1-\mu_2$ وتباين $\sigma^2_{\overline{x_1}-\overline{x_2}}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ أي أن:

المتغير الطبيعي المعياري (Z) تعطى علاقته كالتالي:

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

ويعطى مجال الثقة في هذه الحالة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

$$P\left((\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\leq(\mu_1-\mu_2)\leq(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)=1-\alpha$$

أما إذا كان المجتمعان محدودين مع تحقق العلاقة $\frac{n}{N}>0.05$ للمجتمعين، يعطى مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

[1] . بدر سالم عيسى، مرجع سابق، ص ص: 101-110.

$$P\left((\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}*\frac{N_1-n_1}{N_1-1}+\frac{\sigma_1^2}{n_1}*\frac{N_2-n_2}{N_2-1}}\leq(\mu_1-\mu_2)\right.$$
$$\left.\leq(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}*\frac{N_1-n_1}{N_1-1}+\frac{\sigma_1^2}{n_1}*\frac{N_2-n_2}{N_2-1}}\right)=1-\alpha$$

مثال:

في دراسة خاصة بمقارنة متوسط الدخل الشهري للأسر القاطنة بمدينة تبسة بمتوسط الدخل الشهري للأسر القاطنة بمدينة خنشلة، فإذا كان تباين الدخول في مدينة تبسة هو 6400 ون، وتباين الدخول في مدينة خنشلة هو 3600 ون، فإذا اخترنا من مدينة تبسة عينة عشوائية تحتوي على 400 أسرة ووجدنا أن متوسط الدخل الشهري لهذه الأسر يساوي 250 ون، واخترنا من مدينة خنشلة عينة عشوائية مستقلة عن العينة السابقة تحتوي على 300 أسرة ووجدنا أن الدخل الشهري لهذه الأسر يساوي 210 ون.

– أحسب مجال الثقة للفرق بين متوسطي الدخول في المدينتين عند مستوى ثقة 90%.

الحل:

باستخدام العلاقة التالية:

$$P\left((\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\leq(\mu_1-\mu_2)\leq(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$
$$=1-\alpha$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$(\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=(250-210)-\left(1.645*\sqrt{\frac{6400}{400}+\frac{3600}{300}}\right)$$
$$=31.3$$
$$(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=(250-210)+\left(1.645*\sqrt{\frac{6400}{400}+\frac{3600}{300}}\right)$$
$$=48.7$$

بالتعويض في مجال التقدير نجد:

$$P(31.3\leq\mu_1-\mu_2\leq48.7)=0.90$$

4-1-2- التقدير بفترة للفرق بين متوسطي مجتمعين مجهولي التباين وحجم العينتين كبير: إذا افترضنا عينة عشوائية ذات حجمها n_1 مسحوبة من مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مهول، وعينة عشوائية حجمها n_2 من مجتمع يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، وكانت العينتين مستقلتين عن بعضهما البعض، او كان المجتمعين غير طبيعيين بشرط حجم كل عينة اكبر او تساوي 30، فان توزيع المعاينة للفرق بين المتوسطات الحسابية $\overline{X_1}-\overline{X_2}$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي $\mu_{\overline{x_1}-\overline{x_2}}=\mu_1-\mu_2$ وتباين

$$\sigma^2{}_{(\bar{x}_1-\bar{x}_2)}=\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}=\frac{\hat{S}_1^2}{n_1}+\frac{\hat{S}_2^2}{n_2}$$

أي أن: المتغير الطبيعي المعياري (Z) تعطى علاقته كالتالي:

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

أو بشكل مكافئ كما يلي:

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

ويعطى مجال الثقة في هذه الحالة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

$$P\left((\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}\leq(\mu_1-\mu_2)\leq(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}\right)=1-\alpha$$

مثال:

أجري امتحان في مادة الإحصاء لمجموعتين مستقلتين من الطلبة، الأولى تشمل 75 طالبا، والثانية تشمل 50 طالبا وكانت نتائج هذا الامتحان للمجموعتين كما يلي:

المجموعة الأولى: $\bar{X}_1=80$ ، $S_1^2=49$

المجموعة الثانية: $\bar{X}_2=70$ ، $S_2^2=36$

- أوجد مجال الثقة للفرق بين المتوسطين عند مستوى ثقة 95 %.

الحل:

ويعطى مجال الثقة كالتالي:

$$P\left((\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}\leq(\mu_1-\mu_2)\right.$$

$$\left.\leq(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}\right)=1-\alpha$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$(\bar{X}_1-\bar{X}_2)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}$$

$$=(80-70)-\left(1.96*\sqrt{\frac{49}{75-1}+\frac{36}{50-1}}\right)=7.68$$

$$(\bar{X}_1-\bar{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}$$

$$=(80-70)+\left(1.96*\sqrt{\frac{49}{75-1}+\frac{36}{50-1}}\right)=12.32$$

بالتعويض في مجال التقدير نجد:

$$P(7.68\leq\mu_1-\mu_2\leq12.32)=0.90$$

4-1-3- التقدير بفترة للفرق بين متوسطي مجتمعين مجهولي التباين وحجم العينتين صغير: إذا كان لدينا مجتمع اول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، ومجتمع ثان يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، العينة وحجم المسحوبة من المجتمع الأول $n_1 < 30$ ، وحجم العينة المسحوبة من المجتمع الثاني $n_2 < 30$ ، وكانت العينتان مستقلتين عن بعضهما البعض، فعند التقدير في هذه الحالة نكون أمام حالتين هما:

4-1-3-1- الحالة الأولى: عندما يكون التباينين مجهولين ومتساويين $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$

إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، وكان σ_1^2 و σ_2^2 متساويين، وسحبنا من المجتمع الأول عينة عشوائية بسيطة حجمها $n_1 < 30$، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2 < 30$ مع فرضية تساوي تبايني المجتمعين والعينتين مستقلتين، فان المقدر بقيمة واحدة للمعلمة σ^2 ويطلق عليه التباين أو المقدر المشترك، ويرمز له بالرمز S_p^2 .

وهو يحسب كما يلي:

$$S_p^2 = \frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2 - 2}$$

ويخضع توزيع الاحصاءة الى توزيع ستيودنت بالعلاقة التالية:

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

ويعطى مجال الثقة في هذه الحالة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

$$P\left(-t_{1-\frac{\alpha}{2},n_1+n_2-2} \le T \le t_{1-\frac{\alpha}{2},n_1+n_2-2}\right) = 1 - \alpha$$

ويمكن إعادة كتابتها كالتالي:

$$P\left((\overline{X}_1 - \overline{X}_2) - t_{1-\frac{\alpha}{2},n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le (\mu_1 - \mu_2) \le (\overline{X}_1 - \overline{X}_2) + t_{1-\frac{\alpha}{2},n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right) = 1 - \alpha$$

مثال:

لمعرفة الفارق الحقيقي بين معدل الإجازات السنوية للعاملين، ومعدل الإجازات السنوية للعاملات في إحدى الشركات، تم سحب عينتين عشوائيتين مستقلتين من سجلات العاملين والعاملات في هذه الشركة لسنة معينة، فوجد من سجلات 12 عاملا أن متوسط أيام الإجازات السنوية هو 81 يوما بانحراف معياري 05 أيام، ومن سجلات 10 عاملات وجد متوسط أيام الإجازات السنوية يساوي 85 يوما بانحراف معياري 04 أيام، وبفرض أن التباينين في المجتمعين المدروسين متساويان.

– أوجد تقدير الفارق بين معدل الإجازات السنوية للعمال والعاملات في هذه الشركة عند مستوى ثقة 90 %، مع فرضية أن المجتمعين يتوزعان بشكل طبيعي أو قريب من الطبيعي.

الحل:

نحسب التباين المشترك كالتالي:

$$S_p^2 = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2 - 2} = \frac{(12 * 25) + (10 * 16)}{12 + 10 - 1} = 23$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$(\overline{X}_1 - \overline{X}_2) - t_{1-\frac{\alpha}{2}, n_1+n_2-2}\, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= |81 - 85| - \left(1.725 * \sqrt{23} * \sqrt{\frac{1}{12} + \frac{1}{10}}\right) = 0.45$$

$$(\overline{X}_1 - \overline{X}_2) + t_{1-\frac{\alpha}{2}, n_1+n_2-2}\, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= |81 - 85| + \left(1.725 * \sqrt{23} * \sqrt{\frac{1}{12} + \frac{1}{10}}\right) = 7.54$$

بالتعويض في مجال التقدير نجد:

$$P(0.45 \leq \mu_1 - \mu_2 \leq 7.54) = 0.90$$

4–1–3–2– الحالة الثانية: عندما يكون التباينين مجهولين وغير متساويين ($\sigma_1^2 \neq \sigma_2^2$): إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2، وكان σ_1^2 و σ_2^2 مجهولين وغير متساويين، وسحبنا من المجتمع الأول عينة عشوائية بسيطة حجمها $n_1 < 30$، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2 < 30$ مع فرضية عدم تساوي تبايني المجتمعين والعينتين مستقلتين، كما نعلم ان افضل تقدير لتباين المجتمع الأول هو تباين العينة المسحوبة منه S_1^2 وان افضل تقدير لتباين المجتمع الثاني هو تباين العينة المسحوبة منه S_2^2.

ويعطى مجال الثقة في هذه الحالة عند مستوى ثقة $(1 - \alpha)\%$ كالتالي:

$$P\left((\bar{X}_1-\bar{X}_2)-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\leq(\mu_1-\mu_2)\right.$$

$$\left.\leq(\bar{X}_1-\bar{X}_2)+t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)=1-\alpha$$

وتحسب درجة حرية V لتوزيع ستيودنت T كالتالي:

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

مثال:

تم قياس مستوى السكر في الدم لمجموعتين من المرضى بداء السكري، المجموعة الأولى عددها 10 مرضى، يعانون من السكري المعتمد على الأنسولين، والمجموعة الثانية عددها 20 مريضا يعانون من السكري غير المعتمد على الأنسولين، وجد من المجموعة الأولى أن متوسط السكر لديهم يساوي 310 ملغ بانحراف معياري 165 ملغ، ومن المجموعة الثانية وجد أن متوسط السكر لديهم 235 ملغ بانحراف 100 ملغ، وبفرض أن التجانس بين المجتمعين غير متساويان والمجتمعين يخضعان إلى التوزيع الطبيعي.

– أحسب تقدير الفرق في متوسط السكر للنوعين (المعتمد على الأنسولين وغير المعتمد على الأنسولين) وذلك عند مستوى ثقة 99%.

الحل:

بما أن المجتمعان يخضعان إلى التوزيع الطبيعي وحجم العينتين اقل من 30 فان فترة الثقة للفرق بين متوسطي المجتمعين سوف يخضع لتوزيع ستيودنت بدرجة حرية V.

نحسب قيمة درجة الحرية لتوزيع ستيودنت كما يلي:

$$V=\frac{\left[\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right]^2}{\frac{\left[\frac{S_1^2}{n_1}\right]^2}{n_1-1}+\frac{\left[\frac{S_2^2}{n_2}\right]^2}{n_2-1}}=\frac{\left[\frac{165^2}{10}+\frac{100^2}{20}\right]^2}{\frac{\left[\frac{165^2}{10}\right]^2}{10-1}+\frac{\left[\frac{100^2}{20}\right]^2}{20-1}}=12.41\approx 12$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$(\bar{X}_1-\bar{X}_2)-(\bar{X}_1-\bar{X}_2)-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

$$=(310-235)-\left(3.055*\sqrt{\frac{165^2}{10}+\frac{100^2}{20}}\right)=-98.42$$

$$(\bar{X}_1-\bar{X}_2)+(\bar{X}_1-\bar{X}_2)-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

$$=(310-235)+\left(3.055*\sqrt{\frac{165^2}{10}+\frac{100^2}{20}}\right)=248.42$$

بالتعويض في مجال التقدير نجد:

$$P(-98.42\le \mu_1-\mu_2\le 248.42)=0.99$$

4-2- التقدير بفترة للفرق بين متوسطي مجتمعين مرتبطين

في كثير من البحوث تكون القيمة الأولى في العينة الأولى والقيمة الاولى في العينة الثانية تابعتين لنفس المفردة، والقيمة الثانية في العينة الاولى والقيمة الثانية في العينة الثانية تابعتين لنفس المفردة وهكذا، أي نجد ان القيم المشاهدة في العينتين تكون ازواجا من القيم، ولذلك تسمى العينتان بالعينتين المرتبطتين.

إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2، وكان المجتمعان مرتبطان وسحبت من كل مجتمع عينة حجمها n $(x_1,x_2,......,x_n)$ و$(y_1,y_2,......,y_n)$على الترتيب، بحيث تمثل x_i و y_i القيمتين المتناظرتين في العينتين، إذا كانت الفروق بين قيم العينتين المتناظرتين هي: $(d_1,d_2,........d_n)$، حيث: $(d_i=x_i-y_i)$ لجميع القيم $(i=1,2,3,.......,n)$ فان $(d_1,d_2,........d_n)$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على أنها عينة عشوائية من مجتمع طبيعي متوسطه الحسابي $\mu_d=\mu_1-\mu_2$ وأفضل مقدر للمعلمة μ_d هي المتوسط الحسابي لعينة الفروق $\bar{d}$ وتباينهσ_d^2، وبما أن تباين مجتمع الفروق مجهول اذن قدره بتباين عينة الفروق S_d^2 .

في هذه الحالة يعطى مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي[1]:

$$P\left(\bar{d}-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_d^2}{n-1}}\le(\mu_1-\mu_2)\le\bar{d}+t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_d^2}{n-1}}\right)=1-\alpha$$

[1] . أبو صالح محمد صبحي، مرجع سابق، ص 115.

أو بشكل مكافئ كما يلي:

$$P\left(\bar{d}-t_{1-\frac{\alpha}{2},v}\sqrt{\frac{\hat{S}_d^2}{n}}\leq(\mu_1-\mu_2)\leq\bar{d}+t_{1-\frac{\alpha}{2},v}\sqrt{\frac{\hat{S}_d^2}{n}}\right)=1-\alpha$$

بحيث تحسب كل من: $\bar{d}$ و S_d^2 و $\hat{S}_d^2$ كما يلي:

$$\bar{d}=\frac{\sum_{i=1}^{n}d_i}{n}=\frac{\sum_{i=1}^{n}(X_i-Y_i)}{n},S_d^2=\frac{\sum_{i=1}^{n}\left(d_i-\bar{d}\right)^2}{n},\hat{S}_d^2=\frac{\sum_{i=1}^{n}\left(d_i-\bar{d}\right)^2}{n-1}$$

مثال:

هناك ادعاء 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:

الاشخاص	1	2	3	4	5	6	7
الوزن قبل تناول العقاقير (X)	58,5	60,3	61,7	69	64	62,6	56,7
الوزن قبل تناول العقاقير (Y)	60	54,9	58,1	62,1	58,5	59,9	54,4

- أوجد فترة الثقة لمتوسط الفرق في متوسط الوزن (قبل وبعد تناول العقار) وذلك عند مستوى ثقة 95% بافتراض أن المجتمعين يتوزعان توزيعا طبيعيا.

الحل:

الاشخاص	X_i	Y_i	d_i	$\left(d_i-\bar{d}\right)^2$
1	58,5	60	1.5−	25.573
2	60,3	54,9	5.4	3.397
3	61,7	58,1	3.6	0.002
4	69	62,1	6.9	11.176
5	64	58,5	5.5	3.775
6	62,6	59,9	2.7	0.734

1.580	2.3	54,4	56,7	**7**
46.237	24.9	407.9	432.8	**المجموع**

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} = \frac{\sum_{i=1}^{n}(X_i - Y_i)}{n} = \frac{24.9}{7} = 3.557$$

$$S_d^2 = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n} = \frac{46.237}{7} = 6.605$$

$$\hat{S}_d^2 = \frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1} = \frac{46.237}{6} = 7.706$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\bar{d} - t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_d^2}{n-1}} = 3.557 - \left(2.447 * \sqrt{\frac{6.605}{7-1}}\right) = 0.99$$

$$\bar{d} + t_{1-\frac{\alpha}{2},v}\sqrt{\frac{S_d^2}{n-1}} = 3.557 + \left(2.447 * \sqrt{\frac{6.605}{7-1}}\right) = 6.124$$

بالتعويض في مجال التقدير نجد:

$$P(0.99 \leq \mu_1 - \mu_2 \leq 6.124) = 0.95$$

5- التقدير بفترة لنسبة المجتمع

إن التقدير بقيمة واحدة لنسبة مفردات المجتمع 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 $\hat{q}$ حيث: $\hat{q} = 1 - \hat{p}$ ، وحيث أن التقدير بقيمة واحدة عرضة للأخطاء وبالتالي يفضل استخدام التقدير ضمن فترة، وكما نعلم ان توزيع المعاينة لنسبة العينة $\hat{P}$ عندما يكون حجم العينة كبيرا وتكون قيمة p ليست قريبة من الصفر او الواحد، يتبع التوزيع الطبيعي بمتوسط حسابي $u_{\hat{P}} = p$ وتباين يساوي القيمة $\sigma_{\hat{P}}^2 = \frac{pq}{n}$

وعليه فان: $Z = \frac{\hat{P}-p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$

وبما أن نسبة المجتمع p مجهولة، وهي التي نرغب في تقديرها بإيجاد فترة الثقة، فلا نستطيع حساب تباين المعاينة لنسبة العينة $\sigma_{\hat{P}}^2$ ولكن سنقدره باستخدام أفضل مقدر بالقيمة وهو نسبة العينة $\hat{p}$ وعند استخدام مقدر التباين $\sigma_{\hat{P}}^2 = \frac{\hat{p}(1-\hat{p})}{n}$

في هذه الحالة يعطى مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي[1]:

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

إذا كانت $\widehat{P} = \frac{x}{n}$ نسبة النجاح في عينة عشوائية حجمها n، وكان n كبيرا، فان مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

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

أما إذا كان المجتمع محدود والسحب مع عدم الإرجاع مع تحقق الشرط $\frac{n}{N} > 0{,}05$ فإننا نستخدم معامل الإرجاع ويصبح مجال التقدير كالتالي:

$$P\left(\hat{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}\sqrt{\frac{N-n}{N-1}} \le p \le \hat{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}\sqrt{\frac{N-n}{N-1}}\right) = 1-\alpha$$

مثال:

تريد إحدى 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 وأهدت لهم عبوة مجانية، وبعد استعمالها وجدت الشركة أن 140 منهم فضلوا هذا المسحوق.

- اوجد التقدير النقطي لنسبة المستهلكين الذين يفضلون هذا المسحوق.
- اوجد التقدير بفترة ثقة لنسبة المستهلكين الذين يفضلون هذا المسحوق عند مستوى ثقة 95 %.

الحل:

ايجاد التقدير النقطي لنسبة المستهلكين الذين يفضلون هذا المسحوق

$\hat{p}$ هي تقدير نقطي لنسبة المجتمع p

$$\hat{p} = \frac{x}{n} = \frac{140}{200} = 0.7$$

ايجاد التقدير بفترة ثقة لنسبة المستهلكين الذين يفضلون هذا المسحوق

بتطبيق نظرية النهاية المركزية نجد:

$$n\hat{p} = 200 * 0.7 = 140 \ge 5$$
$$n(1-\hat{p}) = 200(0.3) = 60 \ge 5$$

وبالتالي توزيع المعاينة للنسبة سوف يخضع للتوزيع الطبيعي المعياري

[1] . أبو صالح محمد صبحي، مرجع سابق، ص 125.

يعطى مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

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

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\hat{p}-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}=0.7-\left(1.96*\sqrt{\frac{0.7*0.3}{200}}\right)=0.64$$

$$\hat{p}+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}=0.7+\left(1.96*\sqrt{\frac{0.7*0.3}{200}}\right)=0.76$$

بالتعويض في مجال التقدير نجد:

$$P(0.64\leq p\leq 0.76)=0.95$$

6- التقدير بفترة للفرق بين نسبتي مجتمعين

إذا كانت $x_1,x_2,\ldots,x_{n1}$ تشكل عينة عشوائية من مجتمع أول يخضع لتوزيع ذو الحدين، فان التقدير النقطي لنسبة المجتمع $\hat{p}_1=\frac{x_1}{n_1}$ ، واذا كانت n_1 كبيرة جدا فحسب نظرية النهاية المركزية فان توزيع المعاينة للنسبة يؤول للتوزيع الطبيعي كما يلي: $\hat{P}_1\sim N\left(P_1,\frac{P_1(1-P_1)}{n_1}\right)$.

وكذلك إذا كانت $x_1,x_2,\ldots,x_{n2}$ تشكل عينة عشوائية من مجتمع ثان يخضع لتوزيع ذو الحدين، فان التقدير النقطي لنسبة المجتمع $\hat{p}_2=\frac{x_2}{n_2}$ ، واذا كانت n_2 كبيرة جدا فحسب نظرية النهاية المركزية فان توزيع المعاينة للنسبة يؤول للتوزيع الطبيعي كما يلي: $\hat{P}_2\sim N\left(P_2,\frac{P_2(1-P_2)}{n_2}\right)$.

اذا كانت العينتين مستقلتين، فان توزيع المعاينة للفرق بين نسبتي العينتين $\widehat{P}_1-\widehat{P}_2$ يخضع للتوزيع الطبيعي تقريبا بمتوسط حسابي $u_{\hat{P}_1-\hat{P}_2}=p_1-p_2$ 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 توزيع المعاينة للفرق بين نسبتي العينتين يؤول للتوزيع الطبيعي، والمتغير المعياري يعطى كما يلي[1]:

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

وبما أن نسبة المجتمعين مجهولة ولكن سنقدرهما باستخدام أفضل مقدر بالقيمة وهو نسبة العينتين.

في هذه الحالة يعطى مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

[1] . مراد صلاح أحمد، مرجع سابق، ص 97.

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

$$\left.\leq\left(\hat{P}_1-\hat{P}_2\right)+Z_{1-\frac{a}{2}}\sqrt{\frac{\hat{P}_1\left(1-\hat{P}_1\right)}{n_1}+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}}\right)=1-\alpha$$

أما إذا كان المجتمعان محدودان والسحب مع عدم الإرجاع مع تحقق الشرط $0,05 < \frac{n}{N}$ فإننا نستخدم معامل الإرجاع ويصبح مجال التقدير كالتالي:

$$P\left(\left(\hat{P}_1-\hat{P}_2\right)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\left(1-\hat{P}_1\right)}{n_1}\left(\frac{N_1-n_1}{N_1-1}\right)+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}\left(\frac{N_2-n_2}{N_2-1}\right)}\leq\left(P_1-P_2\right)\right.$$

$$\left.\leq\left(\hat{P}_1-\hat{P}_2\right)+Z_{1-\frac{a}{2}}\sqrt{\frac{\hat{P}_1\left(1-\hat{P}_1\right)}{n_1}\left(\frac{N_1-n_1}{N_1-1}\right)+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}\left(\frac{N_2-n_2}{N_2-1}\right)}\right)$$

$$=1-\alpha$$

مثال:

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.

– 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:

$$\hat{p}_1=\frac{x_1}{n_1}=\frac{6}{120}=0.05$$

$$\hat{p}_2=\frac{x_2}{n_2}=\frac{9}{200}=0.045$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\left(\hat{P}_1-\hat{P}_2\right)-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1\left(1-\hat{P}_1\right)}{n_1}+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}}$$

$$=(0.05-0.045)-1.96\sqrt{\frac{0.05(1-0.05)}{120}+\frac{0.\hat{0}45(1-0.045)}{200}}$$

$$=-0.0434$$

$$(\hat{P}_1-\hat{P}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}$$

$$=(0.05-0.045)+1.96\sqrt{\frac{0.05(1-0.05)}{120}+\frac{0.045(1-0.045)}{200}}$$

$$=0.0534$$

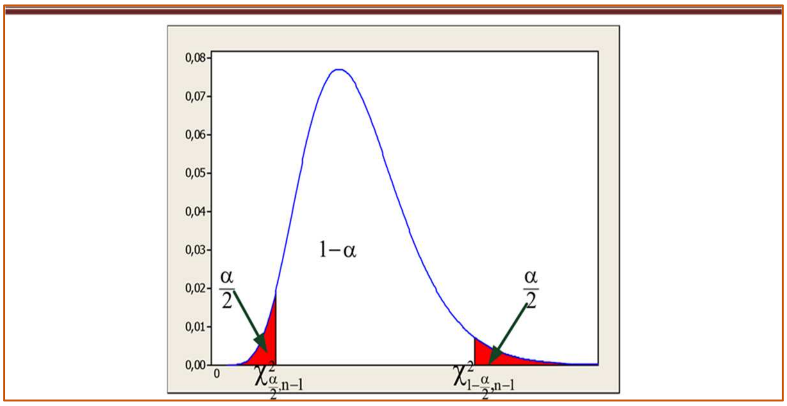
بالتعويض في مجال التقدير نجد:

$$P(-0.0434\leq P_1-P_2\leq 0.0534)=0.95$$

7- التقدير بفترة لتباين المجتمع

في بعض الدراسات الاحصائية نحتاج الى معرفة تباين المجتمع σ^2 وكثيرا ما يكون هذا التباين مجهولا، لذلك نستخدم في هذه الحالة تباين العينة S^2 كتقدير لتباين المجتمعσ^2 ، وهذا التقدير يسمى التقدير بنقطة.

إذا كانت $x_1,x_2,\ldots\ldots,x_{n1}$ تشكل عينة عشوائية حجمها n سحبت من مجتمع يتوزع توزيع طبيعي بمتوسط حسابي μ وتباين σ^2، فان تباين العينة S^2 او $\hat{S}^2$ سيتوزع توزيع كاي تربيع χ^2 بدرجة حرية $n-1$ وذلك كما يلي: $\frac{(n-1)\hat{S}^2}{\sigma^2}=\frac{nS^2}{\sigma^2}\sim\chi^2_{n-1}$ والشكل التالي يوضح ذلك:



ومن الشكل السابق يمكننا أن نستنتج مجال الثقة كالتالي:

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

ولإيجاد فترة ثقة بمستوى ثقة $1-\alpha$ لتباين هذا المجتمع، يكون هناك حالتين[1]:

7-1- اذا كان متوسط المجتمع معلوم: اذا كان متوسط المجتمع μ معلوم فان $S^2=\frac{\sum_{i=1}^{n}(X-\mu)^2}{n}$ تمثل تباين العينة، وان الكمية المحورية المطلوبة في هذه الحالة $\frac{nS^2}{\sigma^2}$ ونعلم أن توزيعها هو توزيع مربع كاي بدرجة حرية n، وهو توزيع مستقل عن σ^2 وبالتالي تعطى فترة الثقة $1-\alpha$ % بالعلاقة التالية:

[1] . مراد صلاح أحمد، مرجع سابق، ص 103.

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

7-2- اذا كان متوسط المجتمع غير معلوم: اذا كان متوسط المجتمع μ غير معلوم 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 $\frac{(n-1)\hat{S}^2}{\sigma^2}$ ونعلم أن توزيعها هو توزيع مربع كاي بدرجة حرية $(n-1)$، وهو توزيع مستقل عن σ^2 وبالتالي تعطى فترة الثقة $1-\alpha$ % بالعلاقة التالية:

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

مثال:

إذا علمت أن تباين عينة عشوائية ذات حجم 25 مسحوبة من مجتمع له التوزيع الطبيعي وكان:

$$X \sim N(10,\sigma^2)$$

$$S^2 = \frac{\sum_{i=1}^{25}(X_i-10)^2}{25} = 09$$

اوجد فترة الثقة عند مستوى ثقة 95 % لتباين هذا المجتمع.

الحل:

تعطى فترة الثقة $1-\alpha$ % بالعلاقة التالية:

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

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\frac{nS^2}{\chi^2_{1-\frac{\alpha}{2},n}} = \frac{25*9}{40.646} = 5.5356$$

$$\frac{nS^2}{\chi^2_{\frac{\alpha}{2},n}} = \frac{25*9}{13.12} = 17.1493$$

مجال التقدير هو:

$$P(5.5356 \leq \sigma^2 \leq 17.1493) = 0.95$$

8- التقدير بفترة للنسبة بين تبايني مجتمعين

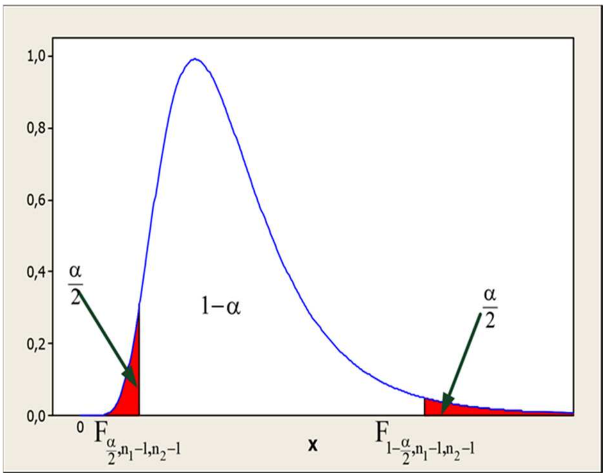
اذا كان لدينا تبايني عينتين عشوائيتين مستقلتين S_1^2 و S_2^2 بحجم n_1 و n_2 مسحوبتين من مجتمعين يتوزعان توزيعا طبيعيا بمتوسط حسابي μ_1 و μ_2 على التوالي، وبتباين σ_1^2 و σ_2^2 على التوالي فان النسبة $\frac{S_1^2}{S_2^2}$ لتباين العينتين تتوزع توزيع فيشر F بدرجتي حرية V_1 و V_2 حيث أن[1]:

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

ومما تقدم يمكن الحصول على المتغير العشوائي F كما يلي:

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

الشكل أدناه يوضح توزيع فيشر بيانيا.



تعطى فترة الثقة $1 - \alpha$ % للنسبة بين تبايني المجتمعين بالعلاقة 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:

---

[1] . مراد صلاح أحمد، مرجع سابق، ص 108.

$$P\left(\frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{1-\frac{\alpha}{2},v_1,v_2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\left(\frac{n_1}{n_1-1}\right)\left(\frac{n_2-1}{n_2}\right)\frac{S_1^2}{S_2^2}}{F_{\frac{\alpha}{2},v_1,v_2}}\right) = 1-\alpha$$

مثال:

تم امتحان مجموعة من الطلبة في مقياس الإحصاء الرياضي وعددهم 16 ولمجموعة أخرى عددها 10، فكانت نتائجهم بانحراف معياري للمجموعة الأولى والثانية على التوالي هو 05 و02.

أوجد فترة الثقة لنسبة التباين عند مستوى ثقة 90%.

الحل:

تعطى فترة الثقة % $1-\alpha$ للنسبة بين تبايني المجتمعين بالعلاقة التالية:

$$P\left(\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{1-\frac{\alpha}{2},v_1,v_2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\hat{S}_1^2/\hat{S}_2^2}{F_{\frac{\alpha}{2},v_1,v_2}}\right) = 1-\alpha$$

نحسب أولا تباين العينة المعدل للمجموعة الأولى والثانية كالتالي:

$$\hat{S}_1^2 = \left(\frac{n_1}{n_1-1}\right)S_1^2 \Rightarrow \hat{S}_1^2 = \left(\frac{16}{16-1}\right)(05)^2 = 26.67$$

$$\hat{S}_2^2 = \left(\frac{n_2}{n_2-1}\right)S_2^2 \Rightarrow \hat{S}_2^2 = \left(\frac{10}{10-1}\right)(02)^2 = 04.44$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{1-\frac{\alpha}{2},v_1,v_2}} = \frac{\frac{26.67}{04.44}}{2.85} = 2.108$$

$$\frac{\hat{S}_1^2/\hat{S}_2^2}{F_{\frac{\alpha}{2},v_1,v_2}} = \frac{\frac{26.67}{04.44}}{0.3937} = 15.257$$

مجال التقدير هو:

$$P\left(02.108 \le \frac{\sigma_1^2}{\sigma_2^2} \le 15.257\right) = 0.90$$

تمارين محلولة

التمرين (01):

تريد إدارة البحوث لإحدى الشركات المنتجة للسيارات تقدير متوسط سعر السيارة المنتجة من قبل هذه الشركة والمباعة في سوق السيارات المستعلمة فقامت باختيار عينة من 225 سيارة من سجلات السيارات للسوق المستعمل في العام الماضي فوجد أن متوسط سعر السيارة 42000 دولار بانحراف معياري 8000 دولار.

اوجد التقدير بنقطة لمتوسط سعر بيع السيارة في سوق المستعمل.

قم بإنشاء فترة ثقة 95 % لمتوسط سعر بيع السيارة ثم فسر النتيجة التي توصلت إليها.

هل تعتقد أنه من المعقول أن يكون متوسط سعر السيارة في السوق 45000 دولار؟ ولماذا ؟

الحل:

ايجاد التقدير بنقطة لمتوسط سعر بيع السيارة في سوق المستعمل

$$\hat{\mu} = \bar{X} = 42000$$

إنشاء فترة ثقة 95 % لمتوسط سعر بيع السيارة

بما أن تباين المجتمع مجهول وحجم العينة كبير (اكبر من 30)، إذن يمكن استخدام التوزيع الطبيعي المعياري، بتطبيق العلاقة التالية:

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

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}} = 42000 - \left(1.96 * \frac{8000}{\sqrt{225-1}}\right) = 42000 - 1045{,}33$$
$$= 40954{,}33$$

$$\bar{X} - Z_{1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}} = 42000 + \left(1.96 * \frac{8000}{\sqrt{225-1}}\right) = 42000 + 1045{,}33$$
$$= 43045{,}33$$

التفسير هذا يعني أن قيمة μ من المحتمل أنها ستقع بين 43045,33 دولار و 40954,33 دولار بنسبة 95 % وهذا يمكن كتابته كالتالي:

$$P(40954{,}33 \leq \mu \leq 43045{,}33) = 0.95$$

لا يمكن القول أن متوسط سعر السيارة في السوق 45000 دولار نظرا لأن هذه القيمة تقع خارج حدود فترة الثقة في المطلوب السابق.

التمرين (02):

أخذت عينة عشوائية حجمها 16 عامل من عمال أحد المصانع فوجد أن متوسط الأجر اليومي للعامل هو 100 دولار بانحراف معياري 10 دولار .

ما هو تقديرك للأجر اليومي للعامل في المصنع كله عند مستوى ثقة 90 % ؟

الحل:

بما أن المجتمع بتباين مجهول وحجم العينة المأخوذة اقل من 30 فان عملية التقدير يتم إجراؤها باستخدام توزيع ستيودنت (T) بالعلاقة التالية:

$$P\left(\bar{X} - T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}} \leq \mu \leq \bar{X} + T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}}\right) = 1-\alpha$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\bar{X} - T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}} = 100 - \left(1,753 * \frac{10}{\sqrt{16-1}}\right) = 95,62$$

$$\bar{X} + T_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n-1}} = 100 + \left(1,753 * \frac{10}{\sqrt{16-1}}\right) = 104,38$$

بالتعويض في مجال التقدير نجد:

$$P(95,62 \leq \mu \leq 104,38) = 0.90$$

التمرين (03):

في استطلاع للرأي من خلال شبكة الانترنت على عينة من 2000 زائر لموقع معين على هذه الشبكة وجد أن 1600 شخص من هذه العينة يفضلون استخدام هذا الموقع.

أوجد التقدير بنقطة لمن يفضلون هذا الموقع في المجتمع ؟

أنشئ فترة ثقة 99 % لنسبة الذين يفضلون هذا الموقع ؟

الحل:

ايجاد التقدير بنقطة

$$\hat{p} = \frac{x}{n} = \frac{1600}{2000} = 0.8$$

إنشاء فترة الثقة

يعطى مجال الثقة عند مستوى ثقة $(1-\alpha)\%$ كالتالي:

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

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$\hat{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0,8 - \left(2,58 * \sqrt{\frac{0.8 * 0.2}{2000}}\right) = 0,777$$

$$\hat{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0,8 + \left(2,58 * \sqrt{\frac{0.8 * 0.2}{2000}}\right) = 0,823$$

بالتعويض في مجال التقدير نجد:

$$P(0,777 \leq p \leq 0,823) = 0.99$$

التمرين (04):

النتائج التالية تمثل نتائج عينتين مستقلتين تم سحبهما من مجتمعين وكانت البيانات في الجدول التالي:

	العينة الأولى	العينة الثانية
حجم العينة	100	200
المتوسط الحسابي	90	80
التباين	25	64

انشئ فترة الثقة للفرق بين متوسطي المجتمعين بدرجة 95 %.

الحل:

يعطى مجال الثقة كالتالي:

$$P\left((\bar{X}_1 - \bar{X}_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}} \leq (\mu_1 - \mu_2) \leq (\bar{X}_1 - \bar{X}_2) + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}\right) = 1 - \alpha$$

بالتعويض في الحد الأدنى والأعلى لفترة الثقة نجد:

$$(\bar{X}_1 - \bar{X}_2) - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}} = (90 - 80) - \left(1.96 * \sqrt{\frac{25}{100 - 1} + \frac{64}{200 - 1}}\right) = 8,52$$

$$(\overline{X}_1-\overline{X}_2)+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1-1}+\frac{S_2^2}{n_2-1}}$$

$$=(90-80)+\left(1.96*\sqrt{\frac{25}{100-1}+\frac{64}{200-1}}\right)=11{,}48$$

بالتعويض في مجال التقدير نجد:

$$P(8{,}52\leq\mu_1-\mu_2\leq 11{,}48)=0.95$$

تمارين غير محلولة

التمرين (01):

1/عند تقدير الوسط الحسابي لمجتمع يتبع توزيع طبيعي، ما هي العبارة الخاطئة فيما يلي:

أ) يتم استخدام التوزيع الطبيعي المعياري إذا كان الانحراف المعياري للمجتمع معلوما.

ب) يتم استخدام التوزيع الطبيعي المعياري إذا كان حجم العينة كبيرا.

ج) يتم استخدام توزيع t إذا كان الانحراف المعياري للمجتمع مجهولا.

د) يتم استخدام توزيع t إذا كان الانحراف المعياري للمجتمع معلوماً .

2 /العبارة الصحيحة من بين العبارات التالية هي:

أ) درجات الحرية بأنها عدد المشاهدات المستقلة في العينة والتي تساوي حجم العينة.

ب) درجات الحرية بأنها عدد المشاهدات المستقلة في العينة والتي تساوي حجم العينة مطروحاً منه عدد القيود أو معالم المجتمع التي يتم تقديرها من بيانات العينة.

ج) درجات الحرية بأنها عدد المشاهدات غير المستقلة في العينة والتي تساوي حجم العينة

د) درجات الحرية بأنها عدد المشاهدات غير المستقلة في العينة والتي تساوي حجم العينة مطروحاً منه عدد القيود أو معالم المجتمع التي يتم تقديرها من بيانات العينة.

التمرين (02):

عينة عشوائية من 64 مفردة وسطها 50 وانحرافها المعياري 20 أخذت من مجتمع عدد مفرداته 800 أوجد تقدير بفترة لمتوسط المجتمع نكون معه واثقين 95 % أن الفترة تتضمن وسط المجتمع.

التمرين (03):

عينة عشوائية من 25 مفردة لمتوسط 80 أخذت من مجتمع عدد مفرداته 1000 توزيعه طبيعي بانحراف معياري 30، أوجد فترات الثقة الآتية لمتوسط المجتمع غير المعلوم عند مستويات ثقة 90% ، 95% ، 99%، وعلام تدل الفروق في النتائج السابقة ؟

التمرين (04):

أخذت عينة عشوائية من 36 طالبا من بين 500 طالب بمدرسة ثانوية، متقدمين لامتحان القبول بالجامعة، ووجد أن متوسط درجات العينة هو 380، والانحراف المعياري للمجتمع كله المكون من 500 طالب هو 40، أوجد فترة الثقة 95% للمتوسط غير المعلوم للدرجات في المجتمع كله.

التمرين (05):

يرغب باحث في تقدير متوسط الأجر الأسبوعي لعدة آلاف من العاملين بأحد المصانع في حدود زائد وناقص 20 دولار وبدرجة ثقة 99%، ويعرف الباحث من خبرته الماضية أن توزيع الأجر الأسبوعي للعاملين يتبع التوزيع الطبيعي بانحراف معياري قدره 40 دولار، ما هو الحد الأدنى للعينة المطلوب ؟

الفصل الثالث

اختبار الفرضيات

تمهيد:

تم التطرق سابقا إلى وسائل دراسة معالم المجتمع المجهولة وذلك من خلال إنشاء فترات ثقة لهذه المعالم واستخدامها كمعلومة مساندة لاتخاذ القرارات، حيث يتم استخدام بيانات عينة عشوائية مسحوبة من المجتمع المراد تقدير معالمه لإنشاء فترة الثقة المطلوبة عند مستوى ثقة.

يلاحظ أن فترة الثقة يتم إنشاؤها بالاعتماد على بيانات عينة عشوائية، ليتم استخدام تلك الفترة في عمليات الاستدلال الإحصائي حول القيمة الحقيقية لمعلمة المجتمع، ولكن في الواقع العملي غالبا ما يكون هنالك ادعاء مسبق حول قيمة المعلمة المجهولة، وليس بالضرورة أن يكون الادعاء مرتبط بقيمة محددة حيث يمكن أن يكون الادعاء ذا صيغة رياضية، كأن ينص مثلا على أن قيمة المعلمة لا تزيد عن قيمة محددة أو أن تكون أكبر من قيمة محددة، في هذه الحالة يكون الهدف من الاستدلال الإحصائي أكثر تحديدا منه في عملية إنشاء فترة ثقة، حيث يكون منصبا حول البحث في مصداقية الادعاء المطروح وبالتالي الوصول إلى قرار بقبول أو رفض الادعاء.

يطلق على عملية التعامل مع الافتراضات والحكم على مصداقيتها بعملية اختبار الفرضيات، وتوجد علاقة بين كل من إنشاء فترة ثقة واختبار الفرضيات، حيث يمكن القول بان اختبار الفرضيات يعطي معلومة أكثر استخداما في اتخاذ القرارات من المعلومة المحصلة من إنشاء فترات الثقة، بيد انه يمكن الاعتماد على فترات الثقة في بعض الحالات للوصول إلى نتائج حول صحة فرضية من عدمها.

في عمليات اختبار الفرضيات يكون هنالك ادعاء أو افتراض يراد اختباره، ويتم في البداية افتراض عدم صحة الادعاء ومن ثم استخدام بيانات الدراسة لإثبات العكس، أي إثبات صحة الادعاء، وتلك الآلية تعطي اختبار الفرضيات قوة نابعة من تلافي التحيز وعدم الدقة، حيث أن الضعف في أداء الدراسة وجمع البيانات يصب في مصلحة عكس الادعاء ومن ثم لا يمكن قبول ادعاء إلا إذا كان هنالك مؤشر إحصائي قوي على ذلك، وتحاكي تلك السياسة في التعامل مع الفرضيات آلية التحقيق في القضايا الجنائية، حيث تقوم على قاعدة أساسية فحواها أن المتهم بريئا حتى تثبت إدانته، وعليه فان الادعاء بان المتهم مذنب يوضع جانبا ويتم تبني العكس، وتتمثل قوة تلك الآلية في انه لا يمكن قبول الفرض بان المتهم مذنب إلا في حال كان هنالك أدلة قوية تشير إلى ذلك، أما في حال كون الأدلة ضعيفة أو أن يكون المتهم بريئا فانه لا يتم قبول الادعاء أو بالأحرى لا يتم رفض الافتراض بان المتهم بريء في الأصل، وتحتوى تلك الآلية أيضا على هدف جوهري يتمثل في تفضيل عدم رفض افتراض بان المتهم بريء وهو مذنب على أن يتم رفض الافتراض بأنه بريء ومن ثم إدانته وهو بريء في الأصل.

في عملية اختبار الفرضيات تمثل العينة العشوائية والبيانات المستخلصة منها دور الأدلة المستخدمة لإثبات إدانة المتهم في القضايا الجنائية، لذا فانه يمكن استخدام بيانات عينة عشوائية وبياناتها لإثبات صحة ادعاء من عدمه، وبالطبع يتم قبول الادعاء في حال كون الأدلة المتمثلة في العينة العشوائية وبياناتها تشير بقوة إلى صحة الادعاء، أما إذا كانت البيانات لا تدعم الادعاء بقوة أو أن يكون الادعاء غير صحيح في الأصل فإننا لا نرفض الافتراض بان الادعاء غير صحيح.

1- مفاهيم نظرية في اختبار الفرضيات

هناك بعض المفاهيم المتعلقة باختبارات الفرضيات لابد من معرفتها:

1-1- الفرضية الإحصائية: هو عبارة عن ادعاء او تخمين معين حول معلمة من معالم المجتمع ويكون المطلوب اختبار صحة هذا الادعاء أو التخمين.

والهدف من الاختبار الإحصائي هو اختبار فرضية حول معلمة أو أكثر من معالم المجتمع الإحصائي، وبالتالي فان الاختبار الإحصائي يتكون من العناصر التالية:

- **فرضية العدم او الفرضية الصفرية:** في عمليات الاستدلال الإحصائي يتم وضع رموز تمثل الادعاء وعكس الادعاء، فبالنسبة للفرض الذي ينص على عدم وجود ظاهرة ومن ثم عدم صحة الادعاء يتم استخدام الرمز H_0 ويطلق عليها فرضية العدم دلالة على عدم وجود أدلة قوية تساند الادعاء المطروح.

- **الفرضية البديلة:** يتم في مقابل فرضية العدم استخدام الرمز H_1 للدلالة على الفرض المغاير للفرض العدمي. وبالطبع عند إجراء اختبار لفرضية باستخدام الطرق المطلوبة فان الادعاء يقع دوما في الفرضية البديلة، وبمعنى آخر يمكن القول بان الادعاء الجيد والقابل للاختبار إحصائيا يجب أن يكون في الفرضية البديلة لا في فرضية العدم، وذلك تجنبا للتحيز وتجنبا لنتائج قد يكون سببها غياب المعلومة وضعف العينة المسحوبة لإجراء الاختبار.

تتطلب عملية اختبار الفرضيات أن تكون الفرضية البديلة عبارة عن جملة كاملة تحتمل الصواب والخطأ، كذلك يفترض أن تكون الفرضية البديلة متعلقة بقيمة معلمة محددة كمتوسط أو نسبة حدوث حدث معين، وبهدف الحصول على فرضية بديلة يتطلب الأمر أن يكون الادعاء متعلق بإحدى حالات ثلاث هي:

أن تكون قيمة معلمة المجتمع اقل من قيمة محددة.

أن تكون قيمة معلمة المجتمع أكثر من قيمة محددة.

أن تكون قيمة معلمة المجتمع مختلفة عن قيمة محددة.

وبالطبع يتم صياغة فرضية العدم، حيث لا تخرج عن إحدى ثلاث صياغات تقابل الصياغات السابقة على التوالي، كما يلي:

قيمة معلمة المجتمع لا تقل من قيمة محددة.

قيمة معلمة المجتمع لا تزيد عن قيمة محددة.

قيمة معلمة المجتمع تساوي قيمة محددة.

مثال:

نفترض اننا نرغب في إجراء اختبارات لثلاث معالم مجهولة القيم هي γ و ρ و φ، وان تلك الاختبارات مرتبطة بفرضيات مستقلة هي:

الفرضية الأولى: قيمة معلمة المجتمع المجهولة γ اقل من 65

الفرضية الثانية: قيمة معلمة المجتمع المجهولة ρ أكثر من 88

الفرضية الثالثة: قيمة معلمة المجتمع المجهولة φ لا تساوى 0.25

حيث يمكن صياغة الفرضيات الثلاث السابقة رياضيا كالتالي:

الفرضية الأول: $\gamma < 65$

الفرضية الثانية: $\rho > 88$

الفرضية الثالثة: $\varphi \neq 0,25$

وبما أن الفرضيات الثلاثة السابقة لا تحتوي في مضمونها على افتراض قيم مساوية لقيمة المعلمة فانه يمكن اعتبارها فرضيات جيدة. وعليه فان عكس تلك الفرضيات يمثل فرضيات العدم للاختبارات لها، وبحكم أسلوب اختبار الفرضيات يتم كتابة فرضية العدم أولا، ثم يليها كتابة الفرضية البديلة، وبتطبيق تلك السياسة فان فرضيات العدم والفرضية البديلة يتم كتابتها بالشكل الرياضي للادعاءات الثالث كالتالي:

- **اختبار ذو جانبين أو ذيلين:** يتحقق هذا الاختبار وفق العلاقة التالية:

$$\begin{Bmatrix} H_0: \varphi = 0,25 \\ H_1: \varphi \neq 0,25 \end{Bmatrix}$$

- **اختبار من الجانب الأيمن:** يتحقق هذا الاختبار وفق العلاقة التالية:

$$\begin{Bmatrix} H_0: \rho = 65 \\ H_1: \rho > 65 \end{Bmatrix}$$

- **اختبار من الجانب الأيسر:** يتحقق هذا الاختبار وفق العلاقة التالية:

$$\begin{Bmatrix} H_0: \gamma = 65 \\ H_1: \gamma < 65 \end{Bmatrix}$$

1-2- 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: احصاءة اختبار (Z)، احصاءة اختبار (T)، احصاءة اختبار (χ^2)، احصاءة اختبار (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 فرضية العدم وتسمى منطقة القبول، والأخرى تشمل النتائج التي إذا ظهرت نرفض فرضية العدم وتسمى منطقة الرفض، وبالتالي يقسم توزيع المعاينة لإحصائية الاختبار إلى منطقتين يمكن تعريفهما كما يلي:

– **منطقة القبول:** هي المنطقة التي تحتوي على قيم إحصائية الاختبار التي تؤدي إلى قبول فرضية العدم.

– منطقة الرفض: وهي المنطقة التي تحتوي على قيم إحصائية الاختبار التي تؤدي إلى رفض فرضية العدم، وتسمى كذلك بالمنطقة الحرجة.

– **القيمة الحرجة:** القيمة أو القيم التي تفصل بين منطقة الرفض ومنطقة القبول.

1-3- أخطاء اختبار الفرضيات وأنواعها: يرتبط الاستدلال الإحصائي بالتعامل مع المجاهيل، ومن ثم لا يمكن الجزم أبدا بان النتائج المحصلة صحيحة تماما، عندما توجد فرضية قائمة على ادعاء فان القرار النهائي يكون إما قبول الفرضية أو رفضها، وبحكم احتواء عملية اختبار الفرضيات على فرضيتان شاملتان ومتضادتان (فرضية العدم والفرضية البديلة) لذا فان القرار المتعلق بإحداها يمثل القرار المعاكس للفرضية الأخرى، فقبول فرضية العدم يعني رفض الفرضية البديلة، والعكس صحيح.

يوجد نوعان من الأخطاء التي يحتمل حدوثها في عملية اختبار الفرضيات، لتوضيح هذين النوعين نفترض أننا على علم بالوضع الحقيقي أو القيمة الحقيقية لمعلمة المجتمع الواقع عليها الاختبار، وأننا نرغب في إجراء الاختبار دون استخدام تلك المعلومة، أي على افتراض أننا نجهل قيمة المعلمة الحقيقية، في هذه الحالة نصبح أمام إحدى خيارين إما أن نفشل في رفض فرضية العدم (لكون اختبار الفرضيات قائم في الأصل على تبني فرضية العدم ومحاولة إثبات عدم صحتها) أو أن ننجح في رفضها ومن ثم إثبات صحة الادعاء الموجود في الفرضية البديلة، فعندما نفشل في رفض فرضية العدم وتكون في الأصل صحيحة فان هذا القرار يعتبر قرار صائب وليس خاطئ، كذلك الوضع عند رفض فرضية عدم خاطئة.

– **الخطأ من النوع الأول:** عندما نقبل فرضية عدم خاطئة أو نرفض فرضية عدم صحيحة فان القرار هنا يصبح خاطئ، يطلق على خطأ رفض فرضية عدم صحيحة بالخطأ من النوع الأول ويرمز لاحتمال حدوثه بالرمز α ويطلق عليه أيضا مستوى المعنوية.

– **الخطأ من النوع الثاني:** يطلق على خطأ قبول فرضية عدم خاطئة بالخطأ من النوع الثاني ويرمز لاحتمال حدوثه بالرمز β .

ويمكن تلخيص القرارات الخاطئة والصائبة في عملية اختبار الفرضيات في الجدول الموالي.

القرار	الوضع الحقيقي	
	H_0 صحيحة	H_0 غير صحيحة
قبول H_0	قرار صائب	قرار خاطئ خطأ من النوع الثاني
رفض H_0	قرار خاطئ خطأ من النوع الأول	قرار صائب

وتبعا لسياسة اختبار الفرضيات يعتبر الخطأ من النوع الأول أكثر خطورة وضررا من الخطأ من النوع الثاني، فإدانة بريء أكثر ضررا وخطورتها من تبرئة مذنب، كما أن الادعاء محل الاهتمام في عملية اختبار الفرضيات يقع دوما في الفرضية البديلة، وبالتالي يفضل في المقام الأول تقليل احتمال قبول ادعاء خاطئ، وعليه فان قيمة احتمال الخطأ من النوع الأول تصبح محدودة بسقف يضعه متخذ القرار إشارة إلى أن احتمال رفض فرضية عدم صحيحة ومن ثم قبول ادعاء خاطئ يجب أن لا يتجاوز حد معين، يطلق إحصائيا على ذلك الحد بمستوى المعنوية (α) للاختبار.

تتحدد قيمة (β) احتمال الوقوع في الخطأ من النوع الثاني من خلال عدة عوامل من أهمها مستوى المعنوية للاختبار (α) وحجم العينة (n) والقيمة الحقيقية لمعلمة المجتمع، وبتقدير احتمال الوقوع في الخطأ من النوع الثاني يتم الحصول على قوة الاختبار حيث تمثل المكمل لاحتمال الوقوع في الخطأ من النوع الثاني. تمثل قوة الاختبار احتمال قبول الفرضية البديلة عندما تكون صحيحة فعلا، وبالطبع لا يمكن حسابها إلا تحت افتراض أن القيمة الحقيقية لمعلمة المجتمع تحت الاختبار معلومة، كاختبار أن متوسط مجتمع يساوي قيمة مختلفة عن القيمة الموجودة في فرضية العدم ويمكن تلخيص ما سبق فيما يلي:

- الخطأ من النوع الأول هو رفض فرضية عدم صحيحة ويرمز لاحتمال وقوعه بالرمز (α) ويطلق علية مصطلح مستوى المعنوية، ويقوم الباحث قبل البدء بعملية الاختبار بتحديد مستوى المعنوية عند تصميم التجربة منذ البداية، حيث يتوقف تحديد مستوى المعنوية على طبيعة البحث أو الدراسة وغالبا ما يتم تحديد مستوى معنوية يساوي 1% او 5% او 10%.
- الخطأ من النوع الثاني هو قبول فرضية عدم خاطئة ويرمز لاحتمال وقوعه بالرمز (β).
- مستوى الثقة $(1-\alpha)$ هو احتمال قبول فرضية عدم صحيحة.
- قوة الاختبار $(1-\beta)$ هو احتمال رفض فرضية عدم خاطئة، وبالتالي يتضح بأنه كلما كان احتمال الوقوع في الخطأ من النوع الثاني قليل كلما أدى ذلك الى زيادة قوة الاختبار، إن العلاقة بين الخطأين من

النوع الأول والثاني، حيث أن انخفاض احد الخطأين يؤدي إلى زيادة الخطأ الآخر، كما أن زيادة حجم العينة يقلل من احتمال الوقع في كلا الخطأين وبالتالي زيادة درجة الثقة.

1-4- المنطقة الحرجة والقيم الحرجة: تعرف المنطقة الحرجة بأنها المنطقة التي عندها يتم رفض الفرضية العدمية، 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 عندما تكون هذه الفرضية صحيحة، حيث أن مساحة المنطقة تحت منحنى دالة التوزيع الاحتمالي لإحصاءة الاختبار يمثل α في حالة الاختبار من جانب واحد، أو تمثل $\frac{\alpha}{2}$ في حالة الاختبار من جانبين.

أما القيم الحرجة فهي قيم جدولية يتم استخراجها من قيم التوزيع الاحتمالي لإحصاءة الاختبار، والتي تتحدد بموجبها مناطق رفض الفرضية العدمية ومناطق قبولها، وتعتمد القيم الحرجة على مستوى المعنوية، وعلى الفرضية البديلة كأن تكون ذات جانب واحد أو جانبين، وعلى التوزيع الاحتمالي لإحصاءة الاختبار، وعلى عدد درجات الحرية، فيما إذا كان التوزيع الاحتمالي أحد توزيعات المعاينة.

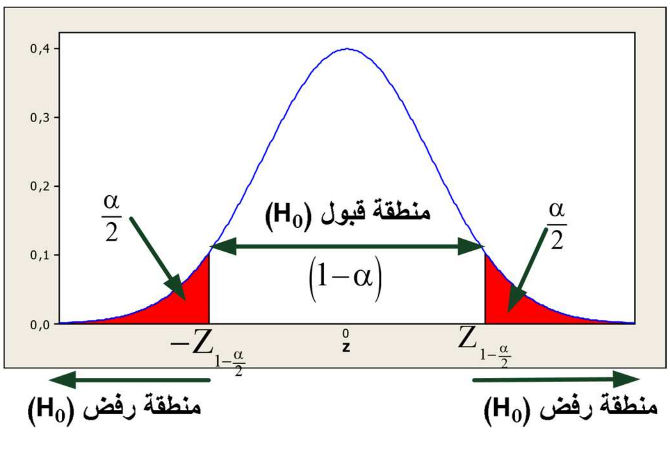
ولتوضيح مفهوم المنطقة الحرجة والقيم الحرجة نفترض لدينا التوزيع الاحتمالي لإحصاءة الاختبار، وهو التوزيع الطبيعي المعياري (Z)، ونريد اختبار الفرضية الاحصائية الآتية: $H_0: \mu = \mu_0$ ضد الفرضية البديلة H_1 التي تأخذ أحد الأشكال التالية عند مستوى معنوية α كما يلي:

$$H_1: \mu \neq \mu_0$$
$$H_1: \mu > \mu_0$$
$$H_1: \mu < \mu_0$$

نستعرض الحالات المختلفة لاختبار الفرضيات كما يلي:

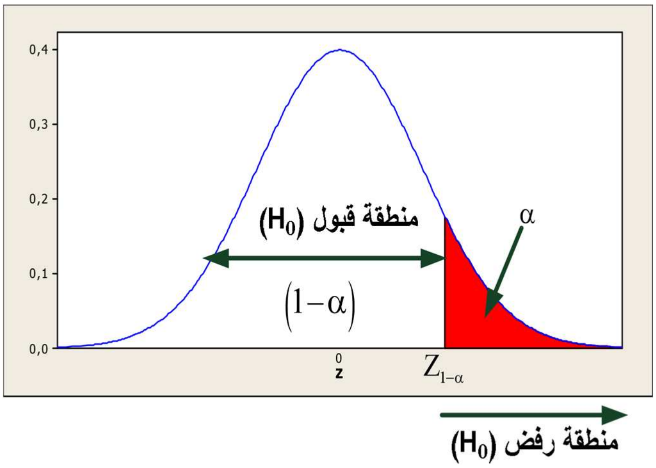
- **الحالة الأولى:** تأخذ الشكل التالي:

$$\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{Bmatrix}$$



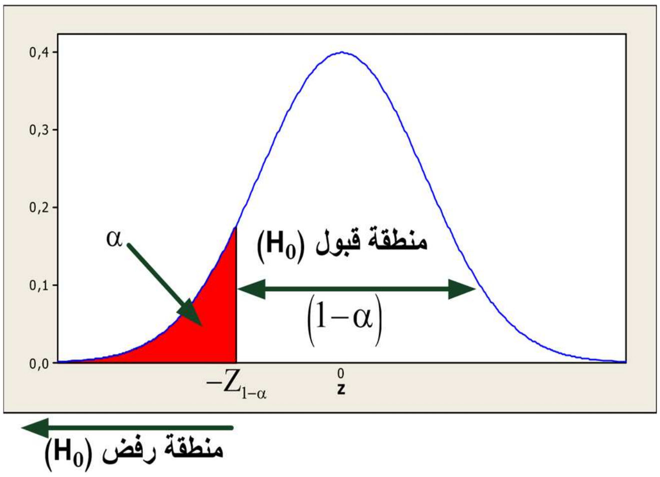
- **الحالة الثانية:** تأخذ الشكل التالي:

$$\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{Bmatrix}$$



- **الحالة الثالثة:** تأخذ الشكل التالي:

$$\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{Bmatrix}$$



1-5- مراحل اختبار الفرضيات: يمر اختبار الفرضيات بالمراحل التالية:

- يجب في البداية تحديد المعلمة المجهولة القيمة والمطلوب إجراء الاختبار عليها، حيث يمكن أن تكون متوسط مجتمع أو الفرق بين متوسطين أو نسبة حدوث حدث في مجتمع أو الفرق بين نسبتين أو تباين مجتمع أو نسبة تباينين.

- يتم تحديد القيمة المقابلة للمعلمة المجهولة والمتعلقة بالادعاء المطلوب اختباره.

- يتم تحديد اتجاه العلاقة بين المعلمة والقيمة المقابلة، والتي يلزم أن تكون في إحدى ثلاث صيغ هي: ($<$ أو $>$ أو $\neq$) ، والتي ستكون الفرضية البديلة الممثلة للادعاء اما من طرفين او طرف واحد.

– يتم في هذه الخطوة صياغة فرضية العدم، حيث تضم مكونات الفرضية البديلة مع تبديل العلاقة الرياضية بين المعلمة والقيمة المقابلة مع تغيير إشارة المتباينة لتعكس الحالة المقابلة للفرضية البديلة، وبالتالي لتمثل عكس الادعاء.

– تحديد مستوى المعنوية α .

– تحديد نوعية التوزيع واختيار الإحصائية المناسبة $[Z, \chi^2, T, F]$ وإيجاد الدرجة أو القيمة الحرجة لها من الجداول الخاصة بها ونرمز لها بالرمز $[Z_{tab}, \chi_{tab}{}^2, T_{tab}, F_{tab}]$

– حساب القيمة الإحصائية θ_{cal} المحددة في النقطة السابقة من بيانات العينة بحيث:

$$\theta_{cal} = [Z_{cal}, \chi^2_{cal}, T_{cal}, F_{cal}]$$

– تحديد القيمة الحرجة حسب مستوى المعنوية، والفرضية البديلة كأن تكون ذات جانب واحد أو جانبين، والتوزيع الاحتمالي لإحصاءة الاختبار، وعدد درجات الحرية، فيما إذا كان التوزيع الاحتمالي أحد توزيعات المعاينة.

– التحليل ومقارنة قيمة الإحصائية المحسوبة من العينة θ_{cal} بالدرجة الحرجة θ_{tab} .

– اتخاذ القرار بشأن رفض أو قبول الفرضية العدمية بعد مقارنة قيمة إحصاءة الاختبار المحسوبة θ_{cal} مع القيمة الحرجة θ_{tab} حسب نوع الفرضية البديلة ومستوى المعنوية فاذا كانت:

- قيمة إحصاءة الاختبار المحسوبة θ_{cal} تقع في منطقة رفض الفرضية العدمية فإن ذلك يدل على رفض الفرضية العدمية وقبول الفرضية البديلة.
- قيمة إحصاءة الاختبار المحسوبة θ_{cal} تقع في منطقة قبول الفرضية العدمية فإن ذلك يدل على قبول الفرضية العدمية ورفض الفرضية البديلة.

2– اختبار الفرضيات حول المتوسط الحسابي

إذا كان لدينا مجتمع طبيعي بمتوسط حسابي μ وتباين σ^2 وسحبنا منه عينة عشوائية ذات الحجم n، متوسطها الحسابي $\overline{X}$ وتباينها S^2، وعلى افتراض أننا نرغب في اختبار الفرضية العدمية $(H_0: \mu = \mu_0)$ ضد أي فرضية بديلة، علما أن هذا الاختبار سواء من جانب واحد او جانبين، فاننا أمام حالتين هما:

2–1– عندما يكون تباين المجتمع معلوم وحجم العينة أكبر أو يساوي 30

إذا كانت لدينا عينة عشوائية من مجتمع توزيعه طبيعي (أو غير طبيعي ولكن شروط تطبيق نظرية النهاية المركزية تبقى متوفرة) متوسطه الحسابي μ مجهول وتباينه σ^2 معلوم، وأردنا اختبار $(H_0: \mu = \mu_0)$، علما أن $\overline{X}$ مقدر جيد للمتوسط μ، حيث يمكن ربط هذا الاختبار بفترات الثقة حول المتوسط μ وذلك مبني على أساس انه إذا كانت فترة الثقة الناتجة أي: $\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ تحتوي μ_0

فانه لا يمكن رفض الفرضية العدمية، اما اذا كانت μ_0 تقع خارج حدي هذه الفترة فنه يجب رفض الفرضية، وفي هذه الحالة إما أن تكون: $\mu_0 < \bar{X} - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ أو $\mu_0 > \bar{X} + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$

ولكن الفرضية الإحصائية قد تكون من طرف واحد أو من طرفين وذلك حسب الفرضية البديلة، وعليه سيكون هناك فرق بين منطقتي الرفض وذلك كما يتضح من خلال الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{Bmatrix}$	$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	رفض H_0 اذا كانت $Z_{cal} \le -Z_{tab}$ او $Z_{cal} \ge Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{Bmatrix}$	$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	رفض H_0 اذا كانت $Z_{cal} \ge Z_{tab}$ بحيث: $Z_{tab} = Z_{1-\alpha}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{Bmatrix}$	$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	رفض H_0 اذا كانت $Z_{cal} \le -Z_{tab}$ بحيث: $Z_{tab} = -Z_{1-\alpha}$

ولغرض اتخاذ القرار الإحصائي حول رفض أو عدم رفض الفرضية العدمية بعد المقارنة اعتمادا على مستوى المعنوية ونوع الفرضية البديلة والجدول التالي يوضح بعض القيم الجدولية الشائعة الاستخدام (Z) التي تخضع للتوزيع الطبيعي المعياري.

	$\mathbf{H_1: \mu \neq \mu_0}$	$\mathbf{H_1: \mu > \mu_0}$ او $\mathbf{H_1: \mu < \mu_0}$
α	$Z_{1-\frac{\alpha}{2}}$	$Z_{1-\alpha}$
0.01	2.575	2.326
0.05	1.96	1.645
0.1	1.645	1.285

مثال:

في اختبار 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 يساوي 115.

– أكتب فرضية الاختبار.

– هل يمكن القول بان متوسط درجات الطلبة بصفة عامة في هذه المدرسة يختلف عن المتوسط العام وذلك عند مستوى معنوية 5% ؟

الحل:

خطوات الاختبار الاحصائي كالتالي:

$$H_0: \mu = 110$$
$$H_1: \mu \neq 110$$

احصاءة الاختبار تكون كالتالي:

$$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{115 - 110}{10/\sqrt{25}} = 2.5$$

بما أن الاختبار من جانبين ومن جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 تكون قيمة Z الجدولية كما يلي: $Z_{1-\frac{\alpha}{2}} = Z_{0.975} = Z_{tab} = 1.96$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $Z_{cal} = 2.5$ وهي أكبر من القيمة الجدولية $Z_{tab} = 1.96$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة

2-2- عندما يكون تباين المجتمع غير معلوم وحجم العينة أكبر أو يساوي 30

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 غير معلوم، وسحبنا عينة عشوائية ذات الحجم n حيث $n \geq 30$، فان احصاءة الاختبار للفرضية $(H_0: \mu = \mu_0)$ تتمثل فيما يلي:

$Z_{cal} = \frac{\bar{X}-\mu_0}{S/\sqrt{n-1}}$ او بشكل مكافئ: $Z_{cal} = \frac{\bar{X}-\mu_0}{\hat{S}/\sqrt{n}}$ وهو يخضع للتوزيع الطبيعي المعياري، حيث $\bar{X}$ هو المتوسط الحسابي للعينة.

مثال:

إذا أعطت عينة إحصائية حجمها 42 متوسطا حسابيا قدره 11.50 بانحراف معياري 03.30.

– أكتب فرضية الاختبار.

– أختبر الفرضية القائلة بان $(H_0: \mu = 10)$ مقابل الفرضية البديلة $(H_1: \mu \neq 10)$ وذلك عند مستوى معنوية $\alpha = 0.05$

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: \mu = 15$$
$$H_1: \mu \neq 15$$

بما أن المجتمع لا يخضع للتوزيع الطبيعي وحيث أن حجم العينة كبير جدا $n \geq 30$ وحسب نظرية النهاية المركزية فان توزيع المعاينة سوف يكون له توزيع قريب من التوزيع الطبيعي المعياري.

احصاءة الاختبار تكون كالتالي:

$$Z_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} = \frac{11.5 - 10}{3.3/\sqrt{42-1}} = 2.91$$

بما أن الاختبار من جانبين ومن جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0.05 تكون قيمة Z الجدولية كما يلي: $Z_{1-\frac{\alpha}{2}} = Z_{0.975} = Z_{tab} = 1.96$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $Z_{cal} = 2.91$ وهي أكبر من القيمة الجدولية $Z_{tab} = 1.96$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة

2-3- عندما يكون تباين المجتمع غير معلوم وحجم العينة أقل من 30

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ وتباين σ^2 مجهول، وسحبنا عينة عشوائية حجمها n حيث $n < 30$، فان توزيع المعاينة للمتغير العشوائي (الإحصاءة) هو $T = \frac{\bar{X}-\mu}{S/\sqrt{n-1}} = \frac{\bar{X}-\mu}{\hat{S}/\sqrt{n}}$ يتبع توزيع ستيونت (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 تتوزع وفق توزيع ستيودنت (T)، وأن القيم الحرجة لهذا الاختبار يتم الحصول عليها من جدول ستيونت (T) بدرجات حرية $V = n - 1$ ومستوى معنوية α .

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{Bmatrix}$	$T_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal} = \frac{\bar{X} - \mu_0}{\hat{S}/\sqrt{n}}$	رفض H_0 اذا كانت $T_{cal} \leq -T_{tab}$ او $T_{cal} \geq T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2},n-1}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{Bmatrix}$	$T_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal} = \frac{\bar{X} - \mu_0}{\hat{S}/\sqrt{n}}$	رفض H_0 اذا كانت $T_{cal} \geq T_{tab}$ بحيث: $T_{tab} = T_{1-\alpha,n-1}$

اختبار من طرف واحد		
$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$	$T_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}}$ أو بشكل مكافئ $T_{cal} = \frac{\bar{X} - \mu_0}{\hat{S}/\sqrt{n}}$	رفض H_0 اذا كانت $T_{cal} \leq -T_{tab}$ بحيث: $T_{tab} = -T_{1-\alpha,n-1}$

مثال:

تمت صناعة آلة لتعطي 230 غرام من الزيت، وذلك عند وضع قطعة النقود المناسبة، ولمعرفة ما ذا كانت هذه الآلة تعمل بحسب المواصفة السابقة، سحبت عينة عشوائية من 06 عبوات لقياس كمية الزيت في كل عبوة، فوجدت النتائج التالية:

رقم العبوة	1	2	3	4	5	6
كمية الزيت بالغرام	190	180	220	200	160	250

- أكتب فرضية الاختبار.
- هل يمكن القول بأن متوسط عمل هذه الآلة يختلف عن 230 غرام وذلك عند مستوى معنوية 0.1 ؟

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: \mu = 230$$
$$H_1: \mu \neq 230$$

بما أن المجتمع لا يخضع للتوزيع الطبيعي بتباين مجهول وحيث أن حجم العينة صغير $n < 30$ ان توزيع المعاينة سوف يكون له توزيع ستيودنت.

احصاءة الاختبار تكون كالتالي:

$$T_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}}$$

نحسب أولا المتوسط الحسابي للعينة والتباين أو التباين المعدل

$$\bar{X} = \frac{\sum_{i=1}^{n=6} X_i}{n} = \frac{190 + 180 + 220 + 200 + 160 + 250}{06} = 200$$

$$S^2 = \frac{\sum_{i=1}^{n=6} (X_1 - \bar{X})^2}{n}$$
$$= \frac{(190-200)^2 + (180-200)^2 + \cdots \dots \dots + (250-200)^2}{06}$$
$$= 833.34$$

$$T_{cal} = \frac{\bar{X}-\mu_0}{S/\sqrt{n-1}} = \frac{200-230}{28.87/\sqrt{6-1}} = -2.32$$

بما أن الاختبار من جانبين ومن جدول توزيع ستيودنت (T) وعند مستوى معنوية 0,1 تكون قيمة T الجدولية كما يلي: $T_{tab} = T_{1-\frac{\alpha}{2},n-1} = T_{0,95,5} = -2,015$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $T_{cal} = -2.32$ وهي أصغر من القيمة الجدولية $T_{tab} = -2,015$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة

3- اختبار الفرضيات حول الفرق بين متوسطي مجتمعين

نميز هنا بين حالتين هما:

3-1- اختبار الفرضيات حول الفرق بين متوسطي مجتمعين مستقلين

إذا كان لدينا مجتمع أول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 ، ومجتمع ثان يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 ، وكانت العينات مستقلة عن بعضهما البعض، وأردنا وضع فرضيات حول الفرق بين متوسطي المجتمعين، فان توزيع المعاينة للفرق بين المتوسطات الحسابية $\overline{X_1}-\overline{X_2}$ سيكون توزيعا طبيعيا بمتوسط حسابي $\mu_{\overline{x_1}-\overline{x_2}} = \mu_1-\mu_2$ وتباين $\sigma^2_{\overline{x_1}-\overline{x_2}} = \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$

أي أن:

المتغير الطبيعي المعياري (Z) تعطى علاقته كالتالي:

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

وعلى افتراض أننا نرغب في اختبار الفرضية العدمية $(H_0\colon \mu_1=\mu_2)$ ضد أي فرضية بديلة، علما أن هذا الاختبار سواء من جانب واحد او جانبين، فإننا أمام ثلاث حالات هي:

- **الحالة الأولى: عندما يكون تبايني المجتمعين معلومين وحجم العينتين كبير**: إذا افترضنا عينة عشوائية ذات حجمها n_1 مسحوبة من مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 معلوم، وعينة عشوائية حجمها n_2 من مجتمع يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 معلوم، وكانت العينتين مستقلتين عن بعضهما البعض، بشرط حجم كل عينة اكبر او تساوي 30، فان احصاءة الاختبار للفرضية $(H_0\colon \mu_1=\mu_2)$ تتمثل فيما يلي:

$Z_{cal} = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ وهو يخضع للتوزيع الطبيعي المعياري.

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{Bmatrix}$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \leq -Z_{tab}$ او $Z_{cal} \geq Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{Bmatrix}$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \geq Z_{tab}$ بحيث: $Z_{tab} = Z_{1-\alpha}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{Bmatrix}$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \leq -Z_{tab}$ بحيث: $Z_{tab} = -Z_{1-\alpha}$

مثال:

أخذت عينة عشوائية بسيطة حجمها 22 من مجتمع أول يخضع للتوزيع الطبيعي بمتوسط حسابي μ_1 وتباين 110، فأعطت متوسط حسابي $\bar{X}_1 = 83$، وأخذت عينة عشوائية ثانية مستقلة عن الأولى، وحجمها 27 من مجتمع ثاني يخضع للتوزيع الطبيعي بمتوسط حسابي μ_2 وتباين 81 فأعطت متوسط حسابي يساوي $\bar{X}_1 = 69$.

– أختبر الفرضية $H_0: \mu_1 = \mu_2$ مقابل الفرضية $H_1: \mu_1 > \mu_2$ وذلك عند مستوى معنوية 0.05.

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 > \mu_2$$

احصاءة الاختبار تكون كالتالي:

$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(83 - 69) - 0}{\sqrt{\frac{110}{22} + \frac{81}{27}}} = 4.95$$

بما أن الاختبار من جانب واحد ومن الجانب الأيمن ومن جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0,05 تكون قيمة Z الجدولية كما يلي: $Z_{tab} = Z_{1-\alpha} = Z_{0,95} = 1,645$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $Z_{cal} = 4.95$ وهي أكبر من القيمة الجدولية $Z_{tab} = 1,645$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة

– **الحالة الثانية: عندما يكون تبايني المجتمعين مجهولين وحجم العينتين كبير**: إذا افترضنا عينة عشوائية ذات حجمها n_1 مسحوبة من مجتمع يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، وعينة عشوائية حجمها n_2 من مجتمع يتوزع هو الاخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، وكانت العينتين مستقلتين عن بعضهما البعض، بشرط حجم كل عينة اكبر او تساوي 30، فان توزيع المعاينة للفرق بين المتوسطات الحسابية $\overline{X_1} - \overline{X_2}$ سيكون توزيعا قريبا من التوزيع الطبيعي بمتوسط حسابي $\mu_{\overline{X_1}-\overline{X_2}} = \mu_1 - \mu_2$ وتباين $\sigma^2{}_{(\bar{x}_1-\bar{x}_2)} = \frac{S_1^2}{n_1-1} + \frac{S_2^2}{n_2-1} = \frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}$

أي أن: المتغير الطبيعي المعياري (Z) تعطى علاقته كالتالي:

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

أو بشكل مكافئ كما يلي:

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

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	**احصاءة الاختبار**	**القرار**
اختبار من طرفين $\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{Bmatrix}$	$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}}$ أو بشكل مكافئ $Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \leq -Z_{tab}$ او $Z_{cal} \geq Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$

اختبار من طرف واحد $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$Z_{cal} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1 - 1} + \dfrac{S_2^2}{n_2 - 1}}}$ أو بشكل مكافئ $Z_{cal} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\hat{S}_1^2}{n_1} + \dfrac{\hat{S}_2^2}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \geq Z_{tab}$ بحيث: $Z_{tab} = Z_{1-\alpha}$
اختبار من طرف واحد $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$	$Z_{cal} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1 - 1} + \dfrac{S_2^2}{n_2 - 1}}}$ أو بشكل مكافئ $Z_{cal} = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\hat{S}_1^2}{n_1} + \dfrac{\hat{S}_2^2}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \leq -Z_{tab}$ بحيث: $Z_{tab} = -Z_{1-\alpha}$

مثال:

للمقارنة بين رواتب أعضاء هيئة التدريس في جامعة 1 و2 والتي تخضع الرواتب فيها إلى التوزيع الطبيعي على الترتيب، اختيرت عينة عشوائية من أساتذة جامعة 1 بحجم $n_1 = 50$ وتبين أن متوسط رواتبهم بلغ 100 ون، وتباين قدره 250 ون، واختيرت عينة عشوائية من أساتذة جامعة 2 بحجم $n_2 = 60$ وتبين أن متوسط رواتبهم بلغ 120 ون وتباين قدره 360 ون.

– أختبر الفرضية التالية:

متوسط رواتب أعضاء هيئة التدريس في جامعة 1 يختلف عن متوسط رواتب أعضاء هيئة التدريس في جامعة 2 وذلك عند مستوى معنوية 0.05.

الحل:

فرضية الاختبار الاحصائي كالتالي:

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

احصاءة الاختبار تكون كالتالي:

$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}} = \frac{(100 - 120) - 0}{\sqrt{\frac{250}{50 - 1} + \frac{360}{60 - 1}}} = -5.97$$

بما أن الاختبار من الجانبين ومن جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0,05 تكون قيمة Z الجدولية كما يلي: $Z_{tab} = Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $Z_{cal} = -5.97$ وهي أصغر من القيمة الجدولية $Z_{tab} = -1.96$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة.

- **الحالة الثالثة: عندما يكون تبايني المجتمعين مجهولين وحجم العينتين صغير:** هنا نميز حالتين:

- **عندما يكون التباينان متساويين:** إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2 مجهول، وكان σ_1^2 و σ_2^2 متساويين، وسحبنا من المجتمع الأول عينة عشوائية بسيطة حجمها $n_1 < 30$، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2 < 30$ مع فرضية تساوي تبايني المجتمعين والعينتين مستقلتين، وبالتالي فان التوزيع الاحتمالي للاحصاءة يعطى كما يلي: $T = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

والذي يتبع توزيع ستيودنت (T) بدرجات حرية $V = n_1 + n_2 - 2$

حيث أن التباين المشترك، ويرمز له بالرمز S_p^2 وهو يحسب كما يلي:

$$S_p^2 = \frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2 - 2}$$

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{Bmatrix}$	$T_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	رفض H_0 اذا كانت $T_{cal} \geq T_{tab}$ او $T_{cal} \leq -T_{tab}$ بحيث $T_{tab} = T_{1-\frac{\alpha}{2}, n_1+n_2-2}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{Bmatrix}$	$T_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	رفض H_0 اذا كانت $T_{cal} \geq T_{tab}$ بحيث: $T_{tab} = T_{1-\alpha, n_1+n_2-2}$

اختبار من طرف واحد $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$	$T_{cal} = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$	رفض H_0 اذا كانت بحيث: $T_{cal} \le -T_{tab}$ $T_{tab} = -T_{1-\alpha, n_1+n_2-2}$

مثال:

عينتان عشوائيتان من توزيعين مستقلين ويخضعان للتوزيع الطبيعي بحيث وكان تباينا المجتمعين مجهولان ومتساويان، ووجد أن حجم العينة الأولى هو 24، وحجم العينة الثانية هو49، وكان المتوسط الحسابي للعينة الأولى هو 16 بتباين 09 والمتوسط الحسابي للعينة الثانية هو 20 بتباين 16.

– اختبر الفرضية $H_0: \mu_1 - \mu_2 = 0$ مقابل الفرضية $H_1: \mu_1 - \mu_2 \neq 0$ وذلك عند مستوى معنوية 0.05.

الحل:

فرضية الاختبار الاحصائي كالتالي:

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

احصاءة الاختبار تكون كالتالي:

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

حيث أن التباين المشترك يحسب كما يلي:

$$S_p^2 = \frac{n_1\,S_1^2 + n_2\,S_2^2}{n_1 + n_2 - 2} = \frac{15 * 9 + 10 * 16}{15 + 10 - 2} = 12.82$$

$$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(16 - 20) - 0}{3.58\sqrt{\frac{1}{15} + \frac{1}{10}}} = -2.74$$

بما أن الاختبار من الجانبين ومن جدول توزيع ستيودنت (T) وعند مستوى معنوية 0,05 تكون قيمة T الجدولية كما يلي: $T_{tab} = T_{1-\frac{\alpha}{2}, n_1+n_2-2} = T_{0,975,23} = -2{,}069$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $T_{cal} = -2.74$ وهي أصغر من القيمة الجدولية $T_{tab} = -2{,}069$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة.

- **عندما يكون التباينان غير متساويين:** إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2 مجهول، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي

μ_2 وتباين σ_2^2 مجهول، وكان σ_1^2 و σ_2^2 غير متساويين، وسحبنا من المجتمع الأول عينة عشوائية بسيطة حجمها $n_1 < 30$، ومن المجتمع الثاني عينة عشوائية بسيطة حجمها $n_2 < 30$ مع العينتين مستقلتين، ونعلم ان افضل تقدير لتباين المجتمع الأول هو تباين العينة المسحوبة منه S_1^2 وان افضل تقدير لتباين المجتمع الثاني هو تباين العينة المسحوبة منه S_2^2 وبالتالي فان التوزيع الاحتمالي للاحصاءة يعطى كما يلي: $T = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

والذي يتبع توزيع قريب من توزيع ستيودنت (T) بدرجة حرية V تعطى كما يلي:

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

وسيكون القرار لكل فرضية من الفرضيات السابقة كما سبق الإشارة إليها في الجدول السابق مع مراعاة طريقة حساب درجة الحرية فقط.

3-2- اختبار الفرضيات حول الفرق بين متوسطي مجتمعين مرتبطين

إذا كان لدينا مجتمعين الاول يتوزع توزيعا طبيعيا بمتوسط حسابي μ_1 وتباين σ_1^2، والمجتمع الثاني يتوزع هو الآخر توزيعا طبيعيا بمتوسط حسابي μ_2 وتباين σ_2^2، وكان المجتمعان مرتبطان وسحبت من كل مجتمع عينة حجمها n $(x_1,x_2,\dots\dots,x_n)$ و$(y_1,y_2,\dots\dots,y_n)$على الترتيب، بحيث تمثل x_i وy_i القيمتين المتناظرتين في العينتين، إذا كانت الفروق بين قيم العينتين المتناظرتين هي: $(d_1,d_2,\dots\dots\dots d_n)$، حيث: $(d_i=x_i-y_i)$ لجميع القيم $(i=1,2,3,\dots\dots\dots,n)$ فان $(d_1,d_2,\dots\dots\dots d_n)$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على أنها عينة عشوائية من مجتمع طبيعي متوسطه الحسابي $\mu_d = \mu_1 - \mu_2$ وأفضل مقدر للمعلمة μ_d هي المتوسط الحسابي لعينة الفروق $\bar{d}$ وتباينهσ_d^2، وبما أن تباين مجتمع الفروق مجهول اذن نقدره بتباين عينة الفروق S_d^2 .

إذا أردنا اختبار فرضية تساوي وسطي المجتمعين $H_0: \mu_1 = \mu_2$ أو $H_0: \mu_d = 0$ نميز الحالتين:

- **الحالة الأولى: حجم العينة $n \geq 30$ و σ_d^2 مجهول:** في هذه الحالة تعطى احصاءة الاختبار كما يلي:

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

بحيث:

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

$$\sigma_{\bar{d}}^2 = \frac{\sigma^2}{n} \simeq \frac{S_{\bar{d}}^2}{n-1} = \frac{\hat{S}_{\bar{d}}^2}{n}$$

$$S_{\bar{d}}^2 = \frac{\sum_{i=1}^{n} \left(d_i - \bar{d}\right)^2}{n}$$

$$\hat{S}_{\bar{d}}^2 = \frac{\sum_{i=1}^{n} \left(d_i - \bar{d}\right)^2}{n-1}$$

وسيكون القرار في حالة إختبار الفرضيات كما تم التطرق له في حالة اختبار المتوسط الحسابي للمجتمع.

– **الحالة الثانية: حجم العينة $n < 30$ و $\sigma_{\bar{d}}^2$ مجهول:** في هذه الحالة تعطى احصاءة الاختبار كما يلي:

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

وسيكون القرار في حالة إختبار الفرضيات كما تم التطرق له في حالة اختبار المتوسط الحسابي للمجتمع.

4– اختبار الفرضيات حول نسبة المجتمع

إذا كان لدينا متغير عشوائي متقطع (X) يمثل عدد حالات النجاح في (n) من المحاولات المستقلة باحتمال نجاح قدره (p)، أي أن المتغير العشوائي يخضع لتوزيع ذي الحدين، وفي حالات كون عدد المحاولات (n) كبير جدا، وان احتمال النجاح (p) واحتمال الفشل (q) ليسا صغيرين جدا، ففي هذه الحالة فإن توزيع ثنائي الحدين يقترب من التوزيع الطبيعي وفقا لنظرية النهاية المركزية، وبالتالي فان توزيع المعاينة للاحصاءة (Z) يعطى كما يلي: $Z = \frac{\hat{P}-p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: p = p_0 \\ H_1: p \neq p_0 \end{Bmatrix}$	$Z_{cal} = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	رفض H_0 اذا كانت $Z_{cal} \leq -Z_{tab}$ او $Z_{cal} \geq Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: p = p_0 \\ H_1: p > p_0 \end{Bmatrix}$	$Z_{cal} = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	رفض H_0 اذا كانت $Z_{cal} \geq Z_{tab}$ بحيث: $Z_{tab} = Z_{1-\alpha}$

اختبار من طرف واحد $\begin{cases} H_0: p = p_0 \\ H_1: p < p_0 \end{cases}$	$Z_{cal} = \dfrac{\widehat{P} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	رفض H_0 اذا كانت $Z_{cal} \leq -Z_{tab}$ بحيث: $Z_{tab} = -Z_{1-\alpha}$

مثال:

ادعت إحدى شركات إنتاج الأجهزة الإلكترونية أن نسبة الأجهزة التالفة في إنتاجها بلغت 4%، قام احد الوكلاء المعتمدين لهذه الشركة باختبار عينة من الأجهزة الإلكترونية المجهزة من قبل الشركة قوامها 500 جهاز وتم فحصها، فتبين للوكيل أن عدد الأجهزة التالفة كان 45 جهازا.

– هل يمكن الاستنتاج بان الشركة صادقة في ادعائها مع الوكيل، وذلك عند مستوى معنوية 0.05 ؟

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: p = 0.04$$
$$H_1: p > 0.04$$

احصاءة الاختبار تكون كالتالي:

$$Z_{cal} = \frac{\widehat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.09 - 0.04}{\sqrt{\frac{0.04(1-0.04)}{500}}} = 5.7$$

بما أن الاختبار من جانب واحد وهو الجانب الايمن ومن جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0,05 تكون قيمة Z الجدولية كما يلي: $Z_{tab} = Z_{1-\alpha} = Z_{0,95} = 1,645$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $Z_{cal} = 5.7$ وهي أكبر من القيمة الجدولية $Z_{tab} = 1,645$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة.

5– اختبار الفرضيات حول الفرق بين نسبتي مجتمعين

إذا كانت $x_1, x_2, \ldots, x_{n1}$ تشكل عينة عشوائية من مجتمع أول يخضع لتوزيع ذو الحدين، واذا كانت n_1 كبيرة جدا، وكذلك إذا كانت $x_1, x_2, \ldots, x_{n2}$ تشكل عينة عشوائية من مجتمع ثان يخضع لتوزيع ذو الحدين، واذا كانت n_2 كبيرة جدا فحسب نظرية النهاية المركزية اذا كانت العينتين مستقلتين، فان توزيع المعاينة للاحصاءة يعطى كما يلي:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \sim N(0,1)$$

وعلى افتراض أننا نرغب في اختبار الفرضية العدمية $(H_0: p_1 = p_2)$ ضد أي فرضية بديلة، علما أن هذا الاختبار سواء من جانب واحد او جانبين.

ولكن قبل أن نوضح كيفية استخدام الإحصاءة في الاختبار تجدر الإشارة هنا إلى انه إذا كان ما يلي: $p_1 = p_2 = p$ فان: $\frac{x_1}{n_1}$ و $\frac{x_2}{n_2}$ سيكون لكلاهما تقدير بقيمة واحدة لنفس المعلمة أي أننا نستخدم التقدير التالي: $\frac{x_2+x_1}{n_2+n_1}$ ويطلق عليه التقدير المشترك ونرمز له بالرمز $\widehat{P}$

وبالتالي تعطى الاحصاءة السابقة كما يلي:

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (p_1 - p_2)}{\sqrt{\frac{\widehat{P}\left(1-\widehat{P}\right)}{n_1} + \frac{\widehat{P}\left(1-\widehat{P}\right)}{n_2}}} \sim N(0,1)$$

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: p_1 = p_2 \\ H_1: p_1 \neq p_2 \end{Bmatrix}$	$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (p_1 - p_2)}{\sqrt{\frac{\widehat{P}\left(1-\widehat{P}\right)}{n_1} + \frac{\widehat{P}\left(1-\widehat{P}\right)}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \le -Z_{tab}$ او $Z_{cal} \ge Z_{tab}$ بحيث $Z_{tab} = Z_{1-\frac{\alpha}{2}}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: p_1 = p_2 \\ H_1: p_1 > p_2 \end{Bmatrix}$	$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (p_1 - p_2)}{\sqrt{\frac{\widehat{P}\left(1-\widehat{P}\right)}{n_1} + \frac{\widehat{P}\left(1-\widehat{P}\right)}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \ge Z_{tab}$ بحيث: $Z_{tab} = Z_{1-\alpha}$
اختبار من طرف واحد $\begin{Bmatrix} H_0: p_1 = p_2 \\ H_1: p_1 < p_2 \end{Bmatrix}$	$Z_{cal} = \frac{\left(\hat{P}_1 - \hat{P}_2\right) - (p_1 - p_2)}{\sqrt{\frac{\widehat{P}\left(1-\widehat{P}\right)}{n_1} + \frac{\widehat{P}\left(1-\widehat{P}\right)}{n_2}}}$	رفض H_0 اذا كانت $Z_{cal} \le -Z_{tab}$ بحيث: $Z_{tab} = -Z_{1-\alpha}$

مثال:

في دراسة للمقارنة بين 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.

– هل يمكن القول بان هناك اختلاف بين نسبة السائقين الذين يستخدمون حزام الأمان عند القيادة بين نوعي السيارات عند مستوى معنوية 0.05 ؟

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: p_1 = p_2$$
$$H_1: p_1 \neq p_2$$

احصاءة الاختبار تكون كالتالي:

$$Z_{cal} = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n_1} + \frac{\hat{P}(1-\hat{P})}{n_2}}} = \frac{(0.6 - 0.675) - 0}{\sqrt{\frac{0.625(1-0.625)}{400} + \frac{0.625(1-0.625)}{200}}}$$
$$= -1.79$$

بما أن الاختبار من جانبين ومن جدول التوزيع الطبيعي المعياري (Z) وعند مستوى معنوية 0,05 تكون قيمة Z الجدولية كما يلي: $Z_{tab} = Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $Z_{cal} = -1.79$ وهي أصغر من القيمة الجدولية $Z_{tab} = -1.96$ وهذا يعني رفض الفرضية العدمية أي أن الاختبار معنوي ونقبل الفرضية البديلة.

6- اختبار الفرضيات حول تباين المجتمع

إذا كانت $x_1, x_2, \dots, x_{n1}$ تشكل عينة عشوائية حجمها n سحبت من مجتمع يتوزع توزيع طبيعي بمتوسط حسابي μ وتباين σ^2 مجهول، وبالتالي فان توزيع المعاينة للاحصاءة $\chi^2 = \frac{(n-1)\hat{S}^2}{\sigma_0^2}$ او بشكل مكافئ $\chi^2 = \frac{nS^2}{\sigma_0^2}$ سيتوزع توزيع كاي تربيع χ^2 بدرجة حرية n-1

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{Bmatrix} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{Bmatrix}$	$\chi_{cal}^2 = \frac{(n-1)\hat{S}^2}{\sigma_0^2}$ او بشكل مكافئ $\chi_{cal}^2 = \frac{nS^2}{\sigma_0^2}$	رفض H_0 اذا كانت $\chi_{cal}^2 < \chi_{\frac{\alpha}{2},n-1}^2$ و $\chi_{cal}^2 > \chi_{1-\frac{\alpha}{2},n-1}^2$ بحيث $\chi_{tab}^2 = \chi_{\frac{\alpha}{2},n-1}^2$ و $\chi_{tab}^2 = \chi_{1-\frac{\alpha}{2},n-1}^2$
اختبار من طرف واحد $\begin{Bmatrix} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 > \sigma_0^2 \end{Bmatrix}$	$\chi_{cal}^2 = \frac{(n-1)\hat{S}^2}{\sigma_0^2}$ او بشكل مكافئ	رفض H_0 اذا كانت $\chi_{cal}^2 > \chi_{1-\alpha,n-1}^2$ بحيث: $\chi_{tab}^2 = \chi_{1-\alpha,n-1}^2$

	$\chi^2_{cal} = \frac{nS^2}{\sigma_0^2}$	
اختبار من طرف واحد $\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 < \sigma_0^2 \end{cases}$	$\chi^2_{cal} = \frac{(n-1)\hat{S}^2}{\sigma_0^2}$ او بشكل مكافئ $\chi^2_{cal} = \frac{nS^2}{\sigma_0^2}$	رفض H_0 اذا كانت $\chi^2_{cal} < \chi^2_{\alpha,n-1}$ بحيث: $\chi^2_{tab} = \chi^2_{\alpha,n-1}$

مثال:

لوحظ أن الانحراف المعياري لعمر نوع من المصابيح المنتجة (تخضع للتوزيع الطبيعي) في أحد المصانع يبلغ 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.

– هل يمكن الاستنتاج بان إنتاج المصنع سيبقى بنفس مستوى الجودة، عند مستوى معنوية 0.05 ؟

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: \sigma^2 = 900$$
$$H_1: \sigma^2 \neq 900$$

احصاءة الاختبار تكون كالتالي:

نحسب أولا تباين العينة المعدل كما يلي:

$$\hat{S}^2 = \frac{nS^2}{n-1} = \frac{31(35)^2}{31-1} = 1265{,}83$$
$$\chi^2_{cal} = \frac{(n-1)\hat{S}^2}{\sigma_0^2} = \frac{(31-1)1265{,}83}{(30)^2} = 42{,}19$$

بما أن الاختبار من جانبين ومن جدول توزيع كاي تربيع وعند مستوى معنوية 0,05 نستخرج القيم الحرجة العليا والدنيا كما يلي:

$$\begin{cases} \chi^2_{tab} = \chi^2_{1-\frac{\alpha}{2},n-1} = \chi^2_{0.975,30} = 46.979 \\ \chi^2_{tab} = \chi^2_{\frac{\alpha}{2},n-1} = \chi^2_{0.025,30} = 16.791 \end{cases}$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $\chi^2_{cal} = 42{,}19$ تقع بين القيمتين الحرجتين لكاي تربيع $\chi^2_{tab} = 16.791$, $\chi^2_{tab} = 46.979$ وهذا يعني قبول الفرضية العدمية ورفض الفرضية البديلة.

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 900 ساعة وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.05.

7– اختبار الفرضيات حول النسبة بين تبايني مجتمعين

اذا كان لدينا تبايني عينتين عشوائيتين مستقلتين S_1^2 و S_2^2 بحجم n_1 و n_2 مسحوبتين من مجتمعين يتوزعان توزيعا طبيعيا بمتوسط حسابي μ_1 و μ_2 على التوالي، وبتباين σ_1^2 و σ_2^2 على التوالي فان توزيع المعاينة للنسبة بين تبايني العينتين يتوزع وفق توزيع فيشر F بدرجتي حرية V_1 و V_2 حيث أن:

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

يمكن اختبار الفرضيات حسب الجدول التالي:

الفرضية الاحصائية	احصاءة الاختبار	القرار
اختبار من طرفين $\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$	$F_{cal} = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2}$ او بشكل مكافئ $F_{cal} = \frac{\left[\frac{S_1^2 n_1}{n_1 - 1}\right]\frac{1}{\sigma_1^2}}{\left[\frac{S_2^2 n_2}{n_2 - 1}\right]\frac{1}{\sigma_2^2}}$	رفض H_0 اذا 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 $F_{tab} = F_{1-\frac{\alpha}{2},V_1,V_2}$
اختبار من طرف واحد $\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 > \sigma_2^2 \end{cases}$	$F_{cal} = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2}$ او بشكل مكافئ $F_{cal} = \frac{\left[\frac{S_1^2 n_1}{n_1 - 1}\right]\frac{1}{\sigma_1^2}}{\left[\frac{S_2^2 n_2}{n_2 - 1}\right]\frac{1}{\sigma_2^2}}$	رفض H_0 اذا كانت $F_{cal} > F_{1-\alpha,V_1,V_2}$ بحيث: $F_{tab} = F_{1-\alpha,V_1,V_2}$
اختبار من طرف واحد $\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 < \sigma_2^2 \end{cases}$	$F_{cal} = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2}$ او بشكل مكافئ	رفض H_0 اذا كانت $F_{cal} < F_{\alpha,V_1,V_2}$ بحيث: $F_{tab} = F_{\alpha,V_1,V_2}$

	$F_{cal}=\dfrac{\left[\dfrac{S_1^2 n_1}{n_1-1}\right]\dfrac{1}{\sigma_1^2}}{\left[\dfrac{S_2^2 n_2}{n_2-1}\right]\dfrac{1}{\sigma_2^2}}$	

مع العلم أن: $F_{\alpha,V_1,V_2}=\frac{1}{F_{1-\alpha,V_2,V_1}}$

مثال:

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.

– أختبر ما إذا كان هناك فرق بين تباين مجتمع أطوال الطالبات وتباين مجتمع أطوال الطلبة وذلك عند مستوى معنوية 0.1.

الحل:

فرضية الاختبار الاحصائي كالتالي:

$$H_0: \sigma_1^2=\sigma_2^2$$
$$H_1: \sigma_1^2\neq\sigma_2^2$$

احصاءة الاختبار تكون كالتالي:

$$F_{cal}=\frac{\hat{S}_1^2}{\hat{S}_2^2}\times\frac{\sigma_2^2}{\sigma_1^2}$$

أو بشكل مكافئ:

$$F_{cal}=\frac{\left[\dfrac{S_1^2 n_1}{n_1-1}\right]\dfrac{1}{\sigma_1^2}}{\left[\dfrac{S_2^2 n_2}{n_2-1}\right]\dfrac{1}{\sigma_2^2}}=\frac{\dfrac{64*31}{31-1}}{\dfrac{36*21}{21-1}}=1.75$$

بما أن الاختبار من جانبين ومن جدول توزيع فيشر وعند مستوى معنوية 0.1 نستخرج القيم الحرجة العليا والدنيا كما يلي:

$$\left\{\begin{array}{c} F_{tab}=F_{1-\frac{\alpha}{2},V_1,V_2}=F_{095,30,20}=2.04 \\ F_{tab}=F_{\frac{\alpha}{2},V_1,V_2}=\dfrac{1}{F_{1-\frac{\alpha}{2},V_2,V_1}}=\dfrac{1}{1.93}=0.518 \end{array}\right\}$$

القرار: بما أن قيمة إحصاءة الاختبار المحسوبة $F_{cal} = 1.75$ تقع بين القيمتين الحرجتين لفيشر $F_{tab} = 0.518\,, F_{tab} = 2.04$ وهذا يعني قبول الفرضية العدمية ورفض الفرضية البديلة.

أي أن تباين مجتمع الطالبات يساوي تباين مجتمع الطلبة ولا يوجد فرق حقيقي بينهما وفقا لمعطيات العينة وذلك عند مستوى معنوية 0.1.

تمارين محلولة

التمرين (01):

نفرض أن باحثا اجتماعيا ادعى ان متوسط اعمار طلبة الجامعة لا يختلف عن متوسط أعمار الطالبات.

قم بصياغة كل من فرضية العدم والفرضية البديلة.

الحل:

فرضية العدم تصاغ في صورة عدم وجود فرق أو عدم وجود علاقة أو عدم وجود تغير.

H_0: عدم وجود اختلاف بين متوسطي اعمار الطلبة والطالبات.

H_1: يوجد اختلاف حقيقي وليس ظاهري بين متوسط اعمار الطلة والطالبات .

التمرين (02):

شركة متخصصة في صناعة لعب الأطفال تعاقدت لشراء نوع جديد من الخيوط الصناعية، يدعي صانع هذه الخيوط أن متوسط قوة تحمل الخيط 15 كغ بانحراف معياري نصف كغ .

ولاختبار صحة ادعاء الصانع أخذت عينة عشوائية من 50 خيطا وتم اختبارها فوجد أن متوسط قوة التحمل في العينة 14.8 كغ. فهل يمكننا تأييد ادعاء المدير مع استخدام مستوى معنوية 5 % ؟

الحل:

$n=50 \qquad \mu_0=15\ kg$

$\overline{X}=14.8\ kg \quad \sigma=0.5\ kg$

صياغة الفرضيات الإحصائية

$$\begin{Bmatrix} H_0: \mu = 15 \\ H_1: \mu \neq 15 \end{Bmatrix}$$

إيجاد قيمة إحصاء الاختبار

$$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{14,8 - 15}{0,5/\sqrt{50}} = -2,83$$

تحديد القيمة الجدولية

$$Z_{tab} = Z_{1-\frac{\alpha}{2}} = \mp 1.96$$

اتخاذ القرار

بما أن القيمة المحسوبة وقعت في منطقة الرفض، فإن القرار هو رفض فرضية العدم أي أن الادعاء غير صحيح وأن هناك فرق معنوي بين المتوسط الحقيقي والمتوسط المدعى.

التمرين (03):

إذا كان من المعروف أن جسم الإنسان البالغ يحتاج يوميا في المتوسط 800 ميللغرام من الكالسيوم لكي يقوم بوظائفه، ويعتقد احد علماء التغذية أن الأفراد ذوي الدخل المنخفض لا يستطيعون تحقيق هذا المتوسط، ولاختبار ذلك تم اختيار عينة من 50 شخصا بالغا من بين ذوي الدخل المنخفض فكان متوسط ما يتناولونه من كالسيوم يوميا هو 755.3 ميللغرام والانحراف المعياري هو 239.3 ميللجرام.

فهل تدل هذه النتائج على أن متوسط ما يتناوله الأشخاص البالغون من ذوي الدخل المنخفض من كالسيوم يقل عن 800 ميللغرام عند مستوى معنوية 0.05 ؟

الحل:

صياغة الفرضيات الإحصائية

$$\begin{Bmatrix} H_0: \mu = 800 \\ H_1: \mu < 800 \end{Bmatrix}$$

إيجاد قيمة إحصاء الاختبار

لان تباين المجتمع مجهول والعينة كبيرة فانه يمكن استخدام تباين العينة بديلا لتباين المجتمع.

وبالتعويض نجد ان قيمة احصاءة الاختبار

$$Z_{\text{cal}} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} = \frac{755{,}3 - 800}{239{,}3/\sqrt{50-1}} = -1{,}32$$

تحديد القيمة الجدولية

نلاحظ هنا أن الاختبار ذو جانب (طرف) واحد هو الجانب (الطرف) الأيسر وحيث أن مستوى المعنوية 0,05 فان القيمة الحرجة

$$Z_{\text{tab}} = -Z_{1-\alpha} = -Z_{0.95} = -1.64$$

اتخاذ القرار

بما أن قيمة الاحصاءة اكبر من القيمة الحرجة وهي تقع في منطقة القبول وبالتالي فإننا لا نرفض فرضية العدم وهي أن متوسط ما يتناوله الإنسان البالغ ذو الدخل المنخفض من الكالسيوم يساوي 800 ميللغرام.

التمرين (04):

في عينة عشوائية مكونة من تسجيل 100 حالة وفاة في قرية معينة تبين أن متوسط العمر في العينة 67.5 عاما والانحراف المعياري 8 أعوام.

فهل هذا يوضح أن متوسط العمر في هذه القرية اكبر من 65 عاما عند مستوى معنوية 5 % ؟

الحل:

صياغة الفرضيات الإحصائية

$$\begin{cases} H_0: \mu = 65 \\ H_1: \mu > 65 \end{cases}$$

إيجاد قيمة إحصاء الاختبار

$$Z_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} = \frac{67{,}5 - 65}{8/\sqrt{100-1}} = 3{,}125$$

تحديد القيمة الجدولية

نلاحظ هنا أن الاختبار ذو جانب (طرف) واحد هو الجانب (الطرف) الأيمن وحيث أن مستوى المعنوية 0,05 فان القيمة الحرجة

$$Z_{tab} = Z_{1-\alpha} = Z_{0.95} = 1.64$$

اتخاذ القرار

نجد أن القيمة المحسوبة اكبر من القيمة الجدولية لهذا فان القرار هو رفض فرضية العدم، ونستنتج من ذلك أن متوسط العمر في هذه القرية اكبر من 65.

التمرين (05):

أجرى بحث لمعرفة مستوى الدخل في المناطق الريفية، وقد أخذت عينة من 25 أسرة وبقياس مستوى الدخل وجد أن متوسط الدخل للأسرة في العينة 8000 دينار في العام وبإنحراف معيارى قدره 666 دينار علما بأن متوسط الدخل حسب ما كشفته بيانات الإحصاء العام في المناطق الريفية يبلغ 6000 دينار.

هل يمكن القول أن تقديرات العينة للدخل تختلف إختلافا حقيقيا عن مستوى الدخل العام عند مستوى معنوية 5 %.

الحل:

صياغة الفرضيات الإحصائية

$$\begin{Bmatrix} H_0: \mu = 6000 \\ H_1: \mu \neq 6000 \end{Bmatrix}$$

إيجاد قيمة إحصاء الاختبار

الإختبار من طرفين نستخدم توزيع ستيودنت عند درجة حرية 24

$$T_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} = \frac{8000 - 6000}{666/\sqrt{25-1}} = 14.68$$

تحديد القيمة الجدولية

$$T_{tab} = T_{1-\frac{\alpha}{2},n-1} = T_{0,975,24} = 2,064$$

إتخاذ القرار

نرفض فرضية العدم ونقبل الفرضية البديلة أي أن مستوى الدخل حسب تقديرات العينة تختلف عن تقديرات الإحصاء العام.

التمرين (06):

إذا كانت أعمار بطاريات السيارات المنتجة بواسطة أحد المصانع تتبع توزيعا طبيعيا، ويدعي صاحب المصنع أن متوسط أعمار هذه البطاريات هو 36 شهرا . ولاختبار صحة هذا الإدعاء اختيرت عينة عشوائية حجمها عشرة بطاريات وقيست أعمارها بالشهور فكان متوسط أعمارها هو 30.33 شهر بانحراف معياري 4.01 شهرا.

فهل تدل هذه البيانات على أن متوسط أعمار البطاريات أقل من 36 شهرا عند مستوى معنوية 5 % ؟

الحل:

صياغة الفرضيات الإحصائية

$$\begin{Bmatrix} H_0: \mu = 36 \\ H_1: \mu < 36 \end{Bmatrix}$$

إيجاد قيمة إحصاء الاختبار

الإختبار من الطرف الأيسر نستخدم توزيع ستيودنت عند درجة حرية 9

$$T_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n-1}} = \frac{30,33 - 36}{4,01/\sqrt{10-1}} = -0.47$$

تحديد القيمة الجدولية

$$T_{tab} = -T_{1-\alpha,n-1} = -T_{0,95,9} = -1,83$$

اتخاذ القرار

نرفض فرضية العدم ونقبل الفرضية البديلة أي أن مستوى الدخل حسب تقديرات العينة تختلف عن تقديرات الإحصاء العام.

التمرين (07):

لمقارنة اتجاهات الذكور والإناث فيما يتعلق بإتجاهاتهم نحو الإنفاق على السلع الكمالية صمم استبيان يضم أسئلة وأعطيت درجات معينة بحيث كانت أعلى درجات تشير إلى الرغبة في إقتناء الأشياء الكمالية وأدنى الدرجات تشير إلى عدم الرغبة في شرائها.

أختيرت عينة عشوائية من 10 رجلا و 15 امرأة وبعد إختيارهم كان متوسط درجات الذكور 115 درجة بإنحراف معياري قدره 14، بينما متوسط درجات الإناث 125 بإنحراف معياري قدره 9 .

هل الإناث أكثر ميلا من الذكور في الإنفاق على الكماليات عند مستوى معنوية 5 % ؟

الحل:

صياغة الفرضيات الإحصائية

μ_1: متوسط مجتمع الذكور

μ_2: متوسط مجتمع الإناث

$$\begin{Bmatrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{Bmatrix}$$

إيجاد قيمة إحصاء الاختبار

الإختبار من الطرف الأيسر وبفرض أن تبايني المجتمعين متساويين نستخدم توزيع ستيودنت عند درجة حرية 23

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

$$S_p^2 = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2 - 2} = \frac{10 \times (14)^2 + 15 \times (9)^2}{23} = 138{,}04$$

$$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(115 - 125) - (0)}{138{,}04\sqrt{\frac{1}{10} + \frac{1}{15}}} = -0{,}18$$

تحديد القيمة الجدولية

$$T_{tab} = -T_{1-\alpha, n_1+n_2-2} = -T_{0{,}95,23} = -1{,}71$$

اتخاذ القرار

نرفض فرضية العدم ونقبل الفرضية البديلة أي أن الإناث أكثر ميلا من الذكور في الإنفاق على الكماليات.

التمرين (08):

إذا كانت أوزان الأفراد الذكور تتبع توزيعا طبيعيا وللمقارنة بين أوزان الأفراد الذكور في فئتي العمر (25–34) و(45–54) اختيرت عينة من أفراد كل فئة، حيث اختيرت عينة حجمها 10 أشخاص من الفئة (25–34) فكان متوسط أوزانهم يساوي 70.19 كلغ بتباين قدرة 8 كلغ، واختيرت عينة حجمها 15 شخصا من الفئة (45–54) فكان متوسط أوزانهم يساوي 68.58 كلغ بتباين قدره 12 كلغ.

فهل تدل هذه البيانات على أن الأفراد الذكور من فئة العمر (25–34) أثقل وزنا من الأفراد في فئة العمر (45–54) عند مستوى معنوية 5 %.

الحل:

صياغة الفرضيات الإحصائية

μ_1: متوسط أوزان مجتمع الذكور من فئة العمر (25–34)

μ_2: متوسط أوزان مجتمع الذكور من فئة العمر (45–54)

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$$

إيجاد قيمة إحصاء الاختبار

الاختبار من الطرف الأيمن وبفرض أن تبايني المجتمعين متساويين نستخدم توزيع ستيودنت عند درجة حرية 23

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

$$S_p^2 = \frac{n_1\, S_1^2 + n_2\, S_2^2}{n_1 + n_2 - 2} = \frac{10 \times (8)^2 + 15 \times (12)^2}{23} = 121{,}73$$

$$T_{cal} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(70{,}19 - 68{,}58) - (0)}{121{,}73\sqrt{\frac{1}{10} + \frac{1}{15}}} = 0.03$$

تحديد القيمة الجدولية

$$T_{tab} = T_{1-\alpha, n_1+n_2-2} = T_{0{,}95{,}23} = 1{,}71$$

اتخاذ القرار

نقبل فرضية العدم أي أن متوسط أوزان الذكور في فئة العمر (25–34) لا يختلف عن متوسط أوزان الذكور في فئة العمر (45–54).

تمارين غير محلولة

التمرين (01):

بين فيما يلي معطيات الفرضيات (المعلمة، القيمة المقابلة والعلاقة الرياضية) وقم بصياغة كل من فرضية العدم والفرضية البديلة.

– ادعاء مدير إدارة الصيانة في إحدى الشركات بان الوقت المستغرق في المتوسط لصيانة أي آلة اقل من 12 ساعة.

– يدعي أحد المصانع الوطنية للبطاريات الكهربائية بان متوسط عمر البطارية المنتجة بواسطة المصنع أكثر من 1.5 سنة.

– يدعي أحد الباحثين بان نسبة الطلاب الحاصلين على إنذارات أكاديمية في جامعة ما اقل من 0.30 من إجمالي عدد الطلاب في الجامعة.

– يدعي أحد المستثمرين بان نسبة الربح في المتوسط من الاستثمار في الأسهم لا تساوي (تختلف عن) 0.10.

التمرين (02):

إذا كانت لديك البيانات التالية التي تمثل عينة أطوال قطع معينة من القماش الذي ينتجه مصنع، اختبر فرضية العدم القائلة أن متوسط أطوال الأقمشة تساوي 150 عند مستوى دلالة 5%.

140	120	150	140	125	140	130	125	140

التمرين (03):

ادعى استاذ الفيزياء أن متوسط درجات طلابه هو (68) فسحب عينة من درجات الطلاب وكانت كما مبينة في الجدول أدناه، فهل ادعاء الاستاذ صحيح أم خاطئ عند مستوى الدلالة 0.05 ؟

80	45	25	30	33	44	52	55	65	75	85	80

التمرين (04):

شركة متخصصة في صناعة الدرجات الهوائية تعاقدت لشراء نوع جديد من العجلات، يدعي صانع هذه العجلات أن متوسط قوة تحمل العجلة 15 كغ بانحراف معياري نصف كجم.

ولاختبار صحة ادعاء الصانع أخذت عينة عشوائية من 50 عجلة وتم اختبارها فوجد أن متوسط قوة التحمل في العينة 14,2 كغ، فهل يمكننا تأييد ادعاء المدير عند مستوى معنوية 1 % ؟

التمرين (05):

إذا كان من المعروف أن البقرة تحتاج يوميا في المتوسط 50 كغ من العلف لكي تنمو نموا سريعا، ويعتقد احد علماء الحيوان أن المربون ذوي الإمكانيات المنخفضة لا يستطيعون تحقيق هذا المتوسط، ولاختبار ذلك تم اختيار عينة من 50 مربي من بين ذوي الإمكانيات المنخفضة فكان متوسط ما يقدمونه من علف يوميا هو 38 كلغ والانحراف المعياري هو 18 كلغ فهل تدل هذه النتائج على أن متوسط ما يقدمه الأشخاص ذوي الإمكانيات المنخفضة من علف يقل عن 50 كلغ عند مستوى معنوية 0.05 ؟

التمرين (06):

في عينة عشوائية مكونة من تسجيل 100 مولود جديد في مدينة معينة تبين أن نسبة المواليد الميتة في العينة 5 % والانحراف المعياري 1%، فهل هذا يوضح أن نسبة المواليد الميتة في هذه المدينة اكبر من 5 % عند مستوى معنوية 5 %؟

التمرين (07):

تدعى منظمة لمكافحة الجوع أن أكثر من 80 % من السكان في منطقة معينة يعانون من سوء التغذية ولكن الحكومة لا تصدق هذا الادعاء. أخذت عينة عشوائية من البيانات المنشورة عن مكافحة الفقر في في المنطقة وتجد أن منها 70 % من السكان يعانون من سوء التغذية.

هل تؤيد بيانات العينة ادعاء الحكومة عند مستوى معنوية 5 % ؟

التمرين (08):

يرغب مشتر كبير للإسمنت أن يقرر عند مستوى معنوية 5%، أي صنف يشتري من بين صنفين لهما نفس السعر، لعمل هذا أخذ عينة عشوائية من 100 كيس من كل صنف فيجد أن الصنف الأول يعيش في المتوسط 15 يوما، مع انحراف معياري قدره 3 يوم، وبالنسبة للصنف الثاني يعيش في المتوسط 22 يوم، مع انحراف معياري قدره 5 يوم، أي الصنفين يجب شراءه إذا كان المشتري يرغب في أن يصل إلى قرار عند مستوى معنوية 5 % ؟

خاتمة

في نهاية هذه المطبوعة البيداغوجية يجدر التوضيح الى اهمية دراسة علم الإحصاء الاستدلالي، لكون موضوعاته موجودة في حياتنا اليومية، فالإحصاء الاستدلالي هو عبارة عن النوع الثاني من علم الإحصاء، ومن خلال هذا النوع يتم دراسة جميع العلاقات بين المتغيرات الخاصة بالدراسة، ويقوم بالتركيز على كافة الاستنتاجات التي تنتج من جميع العلاقات والحسابات الرياضية التي تتم داخل الإحصاء الوصفي.

من خلال الإحصاء الاستدلالي ستتمكن من تحليل كافة البيانات؛ وذلك بغرض التنبؤ، ويعتبر هذا النوع من الإحصاءات وسيلة تستخدم للحكم على كافة البيانات الغير مرئية، ويتم تحليل جميع هذه البيانات واستخلاص كافة النتائج منها.

بالنسبة للعمليات الخاصة بالإحصاء الاستدلالي فإنها الوسيلة الملائمة لشرح وتفسير كافة الظواهر أو المشكلات، كما تستخدم في التنبؤ بكافة الأحداث التي تتعلق بهذه الظواهر، وبالتالي جعل عملية فهم الواقع الحالي أكثر سهولة.

كان ولا يزال علم الإحصاء الاستدلالي يدرس في مختلف المعاهد والجامعات، وبسبب أهميته والحاجة الماسة إليه من قبل طلبتنا الأعزاء، لذا قمنا بوضع هذه المطبوعة بين أيدي طلبتنا، في محاولة منا لإثراء هذا الموضوع من بعض المصادر المهمة، حيث تناولت أمثلة كثيرة في كل فصل لتعزيز وتوضيح ما ورد من تعاريف أو قوانين أو نظريات. كما وضعت في نهاية كل فصل عدد من التمارين وذلك للمساعدة في التدريب على المواضيع ولتحسين قدرة القارئ على فهم هذه المادة.

قائمة المراجع

قائمة المراجع

1- أبو صالح محمد صبحي، **الطرق الإحصائية**، دار اليازوري، عمان، 2007.

2- أبو صالح محمد صبحي، **مقدمة في الإحصاء: مبادئ وتحليل باستخدام SPSS**، ط 6، دار المسيرة، عمان، 2012.

3- المنيزل عبد الله فلاح، **الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الاحصائية (SPSS)**، دار وائل للنشر، عمان، 2000.

4- بدر سالم عيسى، **مبادئ الاحصاء الوصفي والاستدلالي**، دار المسيرة، عمان، 2007.

5- مراد صلاح أحمد، **الأساليب الإحصائية الاستدلالية في العلوم النفسية والتربوية والاجتماعية**، مكتبة الأنجلو، القاهرة، 2000.

6- نافذ محمد بركات، **توزيعات المعاينة**، مشورات الجامعة الإسلامية، غزة، 2020.

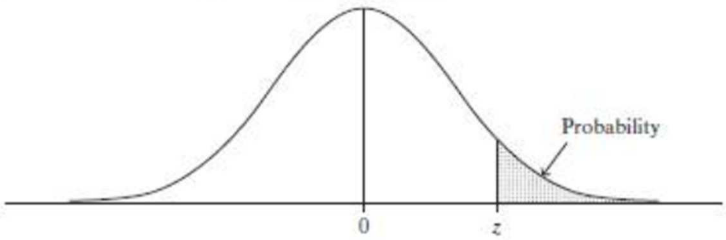
7- علام صلاح الدين، **الأساليب الإحصائية الاستدلالية البارامترية واللابارامترية**، دار الفكر العربي، القاهرة، 2005.

8- عودة، احمد سليمان، **الإحصاء للباحث في التربية والعلوم الإنسانية**، دار الأمل، الأردن، 2010.

قائمة الملاحق

الملحق رقم (01): جدول التوزيع الطبيعي المعياري

TABLE A: Normal curve tail probabilities. Standard normal probability in right-hand tail (for negative values of *z*, probabilities are found by symmetry).

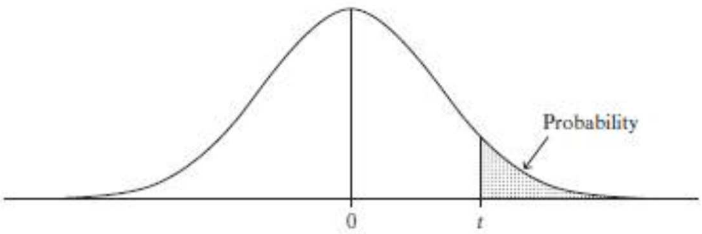


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
	Second Decimal Place of *z*									
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000233									
4.0	.0000317									
4.5	.00000340									
5.0	.000000287									

Source: R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).

الملحق رقم (02): جدول توزيع ستيودنت (t)

TABLE B: *t* Distribution Critical Values

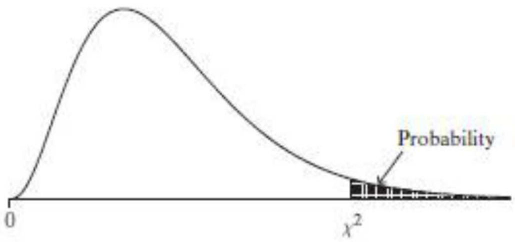


	Confidence Level					
	80%	90%	95%	98%	99%	99.8%
	Right-Tail Probability					
df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	$t_{.001}$
1	3.078	6.314	12.706	31.821	63.656	318.289
2	1.886	2.920	4.303	6.965	9.925	22.328
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.894
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.611
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.091

Source: "Table of Percentage Points of the *t*-Distribution." Computed by Maxine Merrington, Biometrika, 32 (1941): 300. Reproduced by permission of the Biometrika trustees.

الملحق رقم (03): جدول توزيع كاي تربيع (χ^2_n)

TABLE C: Chi-Squared Distribution Values for Various Right-Tail Probabilities

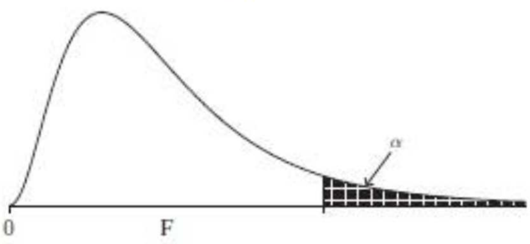


df	Right-Tail Probability						
	0.250	0.100	0.050	0.025	0.010	0.005	0.001
1	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	6.63	9.24	11.07	12.83	15.09	16.75	20.52
6	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	10.22	13.36	15.51	17.53	20.09	21.96	26.12
9	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	13.70	17.28	19.68	21.92	24.72	26.76	31.26
12	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	23.83	28.41	31.41	34.17	37.57	40.00	45.32
25	29.34	34.38	37.65	40.65	44.31	46.93	52.62
30	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	45.62	51.80	55.76	59.34	63.69	66.77	73.40
50	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	77.58	85.53	90.53	95.02	100.4	104.2	112.3
80	88.13	96.58	101.8	106.6	112.3	116.3	124.8
90	98.65	107.6	113.1	118.1	124.1	128.3	137.2
100	109.1	118.5	124.3	129.6	135.8	140.2	149.5

Source: Calculated using *StaTable*, software from Cytel Software, Cambridge, MA.

الملحق رقم (04): جدول توزيع فيشر (F)

TABLE D: *F* Distribution



	$\alpha = .05$									
	df_1									
df_2	1	2	3	4	5	6	8	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9	249.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25
∞	3.84	2.99	2.60	2.37	2.21	2.09	1.94	1.75	1.52	1.00

Source: From Table V of R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, published by Longman Group Ltd., London, 1974. (Previously published by Oliver & Boyd, Edinburgh.) Reprinted by permission of the authors and publishers.

TABLE D: (continued)

	$\alpha = .01$									
	df_1									
df_2	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98.49	99.01	99.17	99.25	99.30	99.33	99.36	99.42	99.46	99.50
3	34.12	30.81	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.27	9.89	9.47	9.02
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91
11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.59	3.16
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.17	2.80	2.36
22	7.94	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00

TABLE D: (continued)

$\alpha = .001$

df_2 \ df_1	1	2	3	4	5	6	8	12	24	∞
1	405284	500000	540379	562500	576405	585937	598144	610667	623497	636619
2	998.5	999.0	999.2	999.2	999.3	999.3	999.4	999.4	999.5	999.5
3	167.5	148.5	141.1	137.1	134.6	132.8	130.6	128.3	125.9	123.5
4	74.14	61.25	56.18	53.44	51.71	50.53	49.00	47.41	45.77	44.05
5	47.04	36.61	33.20	31.09	29.75	28.84	27.64	26.42	25.14	23.78
6	35.51	27.00	23.70	21.90	20.81	20.03	19.03	17.99	16.89	15.75
7	29.22	21.69	18.77	17.19	16.21	15.52	14.63	13.71	12.73	11.69
8	25.42	18.49	15.83	14.39	13.49	12.86	12.04	11.19	10.30	9.34
9	22.86	16.39	13.90	12.56	11.71	11.13	10.37	9.57	8.72	7.81
10	21.04	14.91	12.55	11.28	10.48	9.92	9.20	8.45	7.64	6.76
11	19.69	13.81	11.56	10.35	9.58	9.05	8.35	7.63	6.85	6.00
12	18.64	12.97	10.80	9.63	8.89	8.38	7.71	7.00	6.25	5.42
13	17.81	12.31	10.21	9.07	8.35	7.86	7.21	6.52	5.78	4.97
14	17.14	11.78	9.73	8.62	7.92	7.43	6.80	6.13	5.41	4.60
15	16.59	11.34	9.34	8.25	7.57	7.09	6.47	5.81	5.10	4.31
16	16.12	10.97	9.00	7.94	7.27	6.81	6.19	5.55	4.85	4.06
17	15.72	10.66	8.73	7.68	7.02	6.56	5.96	5.32	4.63	3.85
18	15.38	10.39	8.49	7.46	6.81	6.35	5.76	5.13	4.45	3.67
19	15.08	10.16	8.28	7.26	6.61	6.18	5.59	4.97	4.29	3.52
20	14.82	9.95	8.10	7.10	6.46	6.02	5.44	4.82	4.15	3.38
21	14.59	9.77	7.94	6.95	6.32	5.88	5.31	4.70	4.03	3.26
22	14.38	9.61	7.80	6.81	6.19	5.76	5.19	4.58	3.92	3.15
23	14.19	9.47	7.67	6.69	6.08	5.65	5.09	4.48	3.82	3.05
24	14.03	9.34	7.55	6.59	5.98	5.55	4.99	4.39	3.74	2.97
25	13.88	9.22	7.45	6.49	5.88	5.46	4.91	4.31	3.66	2.89
26	13.74	9.12	7.36	6.41	5.80	5.38	4.83	4.24	3.59	2.82
27	13.61	9.02	7.27	6.33	5.73	5.31	4.76	4.17	3.52	2.75
28	13.50	8.93	7.19	6.25	5.66	5.24	4.69	4.11	3.46	2.70
29	13.39	8.85	7.12	6.19	5.59	5.18	4.64	4.05	3.41	2.64
30	13.29	8.77	7.05	6.12	5.53	5.12	4.58	4.00	3.36	2.59
40	12.61	8.25	6.60	5.70	5.13	4.73	4.21	3.64	3.01	2.23
60	11.97	7.76	6.17	5.31	4.76	4.37	3.87	3.31	2.69	1.90
120	11.38	7.31	5.79	4.95	4.42	4.04	3.55	3.02	2.40	1.56
∞	10.83	6.91	5.42	4.62	4.10	3.74	3.27	2.74	2.13	1.00