الفصل الثالث

اختبار الفرضيات البسيطة

3 − 1 مقدمة:

يتناول اختبار الفرضيات مسألة اتخاذ قرار احتمالي حول قبول أو رفض تلك الفرضيات،أما الفرضيات فقد تتناول أي مؤشر من مؤشرات المجتمع(كالمتوسط و الإجمالي والنسبة ...الخ)،حيث نفترض أنها تساوي قيمة معينة مسبقاً،ثم يتم إجراء الاختبار لاتخاذ قرار حول ذلك.

كما تتناول موضوعات أخرى مثل شكل أو نوع التوزيعات الاحتمالية فنفترض أن لها شكلاً منتظماً أو طبيعياً.

فمثلاً قد يدّعي أحدهم أن متوسط استهلاك الأسرة من الماء يساوي 60 ليتراً يومياً، أو متوسط الدخل الشهري للموظفين في بلدة يساوي 40000 ل.س ...الخ.

وعندما نسمع أو نقرأ هذه الإدعاءات لا نستطيع التسليم بها،لذلك نعتبر هذه الإدعاءات افتراضا ت حتى يتم اختبارها والتحقق من صحتها وعندها يتم قبولها أو رفضها.ويتم ذلك عن طريق سحب عينة من المجتمع والحصول منها على تقدير للثابت الإحصائي المفترض.وفي أغلب الأحيان سيكون هناك فرق بين القيمة المفروضة للثابت الإحصائي وبين قيمته المقدرة عن طريق العينة،فإذا كان الفرق بين القيمتين من الوجهة الإحصائية صغيراً، فإننا نعتبر أن نتائج العينة لا تخالف الفرضية،وبالتالي: نقرر صحة الفرضية وقبولها.أما إذا كان الفرق كبيراً فنقول بأن نتائج العينة لا تؤيد الفرضية وأن هناك عدم انسجام بينهما.

ولاختبار الفرضيات صلة وثيقة بمجالات الثقة التي تقابل احتمالات الثقة المطلوبة،لأننا لا نقبل الفرضية إلا إذا كانت واقعة ضمن مجال الثقة المفروض والمقابل لمعامل الثقة المعلوم.

3 − 2 اختبار الفرضيات:

مهما تكن عملية الاختبار التي نرغب القيام بها،فإن خطوات الاختبار تبقى ثابتة،وهي تهدف إلى تحديد – من خلال عملية عشوائية – قاعدة تسمح بتقرير إحدى فرضيتين:

الأولى تعطي قيمة محددة للثابت الإحصائي،والثانية تحدد مجالاً لهذا الثابت.

والفرضيتان هما :

– الفرضية الابتدائية أو فرضية العدم(الفرضية الصفرية) H_0 :وهي الفرضية التي نقوم باختبارها.وتتضمن قيمة افتراضية لكل أو بعض ثوابت المجتمع الإحصائي المدروس.

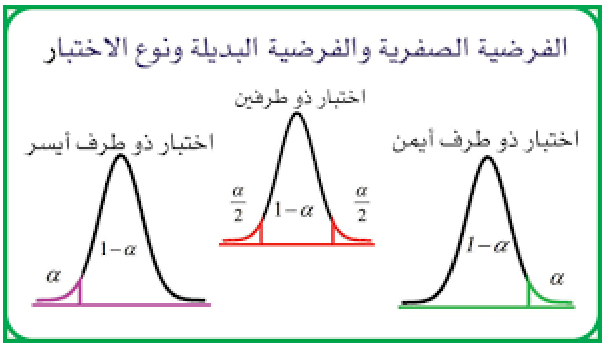
لنفترض أننا نريد اختبار المؤشر θ (ثيتا)وهو خاضع لقانون التوزيع الطبيعي حيث نفترض أن قيمة هذا المؤشر ه θ_0 ونكتب :

$$H_0 : \theta = \theta_0 \qquad (1-3)$$

– الفرضية البديلة H_1 :هي فرضية مخالفة للفرضية الأولى H_0 حيث نفترض أن المؤشر θ يختلف عن θ_0 . كما ويمكن أن تأخذ الفرضية H_1 أحد الأشكال الآتية :

$$\begin{aligned} H_1 &: \theta > \theta_0 \\ H_1 &: \theta < \theta_0 \\ H_1 &: \theta \neq \theta_0 \end{aligned} \qquad (2-3)$$

ففي الحالة الأولى نجد أن منطقة الرفض تقع بكاملها على يمين منطقة القبول.أما في الحالة الثانية فإنها تقع على يسار منطقة القبول.و في الحالة الثالثة فإنها تقع على طرفي منطقة القبول بشكل متناظر .حسب الشكل الآتي :



الشكل (3–1) : يبين الأشكال المختلفة للفرضية البديلة H_1

إن الاختبار في الحالتين الأولى والثانية ذو اتجاه واحد حيث يعطينا احتمال أن تكون θ أكبر من θ_0 في الحالة الأولى، واحتمال أن تكون θ أصغر من θ_0 في الحالة الثانية،أما الحالة الثالثة فتمثل اختباراً ذو اتجاهين حيث نحصل على احتمال أن تكون θ أصغر أو أكبر من θ_0 .

$3-2-1$ الأخطاء المتعلقة باختبار الفرضيات:

يعتمد الباحث أو الدارس في اختباره لفرض معين على دليل العينة وهو قيمة مشتقة أو محسوبة من بيانات عينة عشوائية مختارة من مجتمع الدراسة،حيث أن هذه العينة تشكل جزءاً صغيراً من مجتمع الدراسة ويستخدمها الباحث أو الدارس في التعميم على المجتمع بأكمله فإنه معرض للوقوع في بعض الأخطاء في اختبار الفرض موضوع البحث.

حيث اتفق على تقسيم هذه الأخطاء إلى نوعين:

- خطأ من النوع الأول Type 1 Error: وهو الخطأ المرتكب عند رفض فرضية العدم وهي صحيحة، ويرمز لاحتمال الوقوع في هذا الخطأ بـ α حيث يسمى هذا الاحتمال(مستوى المعنوية) Level of Significance والمقدار $1-\alpha$ يسمى(درجة الثقة) Degree of Confidence.

- خطأ من النوع الثاني Type 2 Error: وهو الخطأ المرتكب عند قبول فرضية العدم وهي خاطئة، ويرمز لاحتمال الوقوع في هذا الخطأ بـ β والمقدار $1-\beta$ يسمى(قوة الاختبار) Power o Test.

ويمكن توضيح ما سبق في الجدول الآتي:

البيان	H_0 صحيحة	H_0 خاطئة
رفض الفرضية H_0	خطأ من النوع الأول α	قرار صحيح واحتماله $\beta=1-\alpha$
قبول الفرضية H_0	قرار صحيح واحتماله $\beta=1-\alpha$	خطأ من النوع الثاني β

منطقتي القبول والرفض :

– المنطقة الأولى: وهي التي تقع ضمنها قيمة المؤشر المختبر المدروس وتسمى منطقة القبول، وتكون متمركزة حول القيمة الافتراضية θ_0 ويتحدد حجمها بحيث يكون الاحتمال الذي يقابلها مساوياً للمقدار $\beta = 1-\alpha$.

– المنطقة الثانية: وهي التي تقع ضمنها قيمة المؤشر المختبر المدروس وتسمى منطقة الرفض، ويكون الاحتمال الذي يقابلها مساوياً لمستوى الدلالة α .

ويمكن أن تكون هذه المنطقة موجودة على أحد جوانب منطقة القبول أو على كلا الجانبين حسبما يكون الاختبار أحادي أو ثنائي الاتجاه، كما في الشكل $(3-1)$ وعندما يكون ثنائي الجانب فإن مستوى الدلالة $\frac{\alpha}{2}$.

$3-2-2$ خطوات وعناصر الاختبار:

مهما كانت المسألة المراد اختبارها، فإن خطوات إجراء الاختبار واحدة وهي:

$1-$ تحديد الفرضيتين الابتدائية H_0 والبديلة H_1 وغالباً ما نفترض فرضية العدم التي تقول بأنه ليس هناك فرق جوهري بين θ و θ_0 .

تحديد مستوى الدلالة α ، وقد جرت العادة أن تكون قيمته $\alpha = 0.05$ أو $\alpha = 0.01$ ، وبالتالي: تحدد قيمة المتغير الطبيعي المعياري المقابل لهذا المستوى ويساوي عندما يكون الاختبار ذو اتجاهين $t = 1.96$ أو $t = 2.58$ على الترتيب. إن قيمة t التي نحصل عليها من الجدول تسمى بـ t النظرية ونرمز لها بـ t_α .

$2-$ حساب قيمة مؤشر الاختبار وهو يساوي إلى قيمة t الفعلية كالآتي:

$$t = \frac{\tilde{\theta} - \theta_0}{\tilde{\sigma}_{\tilde{\theta}}} \qquad (3-3)$$

حيث أن $\tilde{\theta}$ تقدير θ من العينة.

$\tilde{\sigma}_{\tilde{\theta}}$ الانحراف المعياري للتقدير.

إن المقدار t يخضع لتوزيع ستيودنت عندما يكون حجم العينة $n < 30$ مفردة وللتوزيع الطبيعي المعياري إذا كان حجم العينة $n \geq 30$ مفردة.

1 -المقارنة واتخاذ القرار الإحصائي السليم: نقارن قيمة t الفعلية مع t_{α} الحرجة،فإذا كانت $|t| > t_{\alpha}$ ذلك يعني أن t تقع في منطقة الرفض وبالتالي: نرفض الفرضية الصفرية.أما إذا كانت $|t| < t_{\alpha}$ ذلك يعني أن t تقع في منطقة القبول وبالتالي: نقبل الفرضية الصفرية.

3 – 2 – 3 اختبار حول متوسط المجتمع:

عندما يكون الاختبار حول متوسط المجتمع $\bar{y}$ ،فإن الفرضية الابتدائية تكون من الشكل:

$$H_0 : \bar{y} = y_0 \qquad (4-3)$$

وإن الفرضية البديلة يمكن أن تكون من الشكل:

$$H_1 : \bar{y} \neq y_0 \qquad (5-3)$$

وعندها يكون الاختبار ثنائي الجانب .

كما ويمكن للفرضية البديلة أن تكون من أحد الشكلين الآتيين:

$$H_1 : \bar{y} > y_0$$

أو

$$H_1 : \bar{y} < y_0$$

وعندها يكون الاختبار أحادي الجانب .

وإن قيمة مؤشر الاختبار تحسب من العلاقة الآتية :

$$t = \frac{\bar{x} - \bar{y}_0}{\tilde{\sigma}_{\bar{x}}} = \frac{\bar{x} - y_0}{\frac{s}{\sqrt{n}}} \qquad (6-3)$$

وهنا يجب أن نميز بين :

1 -عندما يكون حجم العينة $(n \geq 30)$:

إذا كانت الفرضية H_1 من الشكل $(H_1 : \bar{y} \neq y_0)$ أي الاختبار ثنائي الجانب،يكون القرار الإحصائي كالآتي :

- نقارن قيمة t مع القيمة $Z_{1-\frac{\alpha}{2}}$ والتي يمكن الحصول عليها من جدول التوزيع الطبيعي المعياري،والمقابلة للاحتمال $\left(1 - \frac{\alpha}{2}\right)$ ،فإذا كانت $|t| < z_{1-\frac{\alpha}{2}}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية H_1 ،أما إذا كانت $|t| \geq z_{1-\frac{\alpha}{2}}$ فإننا نرفض الفرضية H_0 ونقبل الفرضية البديل H_1 .

إذا كانت الفرضية H_1 من الشكل $(H_0 : \bar{y} > y_0)$ فإن الاختبار يكون أحادي الجانب يميني،يكون القرار الإحصائي كالآتي :

- نقارن قيمة t مع القيمة $Z_{(1-\alpha)}$ ،والمقابلة للاحتمال $(1-\alpha)$ ،فإذا كانت $|t| < z_{1-\alpha}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية H_1 ،أما إذا كانت $|t| \geq z_{1-\alpha}$ فإننا نرفض الفرضية H_0 ونقبل الفرضية H_1 .

إذا كانت الفرضية H_1 من الشكل $(H_1 : \bar{y} < y_0)$ فإن الاختبار يكون أحادي الجانب يساري،يكون القرار الإحصائي كالآتي :

- نقارن قيمة t مع القيمة $Z_{(1-\alpha)}$، والمقابلة للاحتمال $(1-\alpha)$، فإذا كانت $|t| > -z_{1-\alpha}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية H_1، أما إذا كانت $|t| < -z_{1-\alpha}$ فإننا نرفض الفرضية H_0 ونقبل الفرضية H_1.

2 -أما إذا كان حجم العينة $(n < 30)$:

إذا كانت الفرضية H_1 من الشكل $(H_1 : \bar{y} \neq y_0)$ فإن الاختبار يكون ثنائي الجانب، يكون القرار الإحصائي كالآتي:

- نقارن قيمة t مع القيمة $t_{1-\frac{\alpha}{2}}$ والتي يمكن الحصول عليها من جدول توزيع ستيودنت، والمقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ ولدرجة حرية $(k = n-1)$، فإذا كانت $|t| < t_{(1-\frac{\alpha}{2}, n-1)}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية H_1.

أما إذا كانت $|t| > t_{1-\alpha}$ فإننا نرفض الفرضية H_0 ونقبل الفرضية H_1.

إذا كانت الفرضية H_1 من الشكل $(H_1 : \bar{y} > y_0)$ فإن الاختبار يكون أحادي الجانب يميني، يكون القرار الإحصائي كالآتي:

- نقارن قيمة t مع القيمة $t_{1-\alpha}$، والمقابلة للاحتمال $(1-\alpha)$ ولـ $(n-1)$ درجة حرية في توزيع ستيودنت، فإذا كانت $|t| < t_{(1-\alpha)}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية H_1. أما إذا كانت $|t| \geq t_{1-\alpha}$ فإننا نرفض الفرضية H_0 ونقبل الفرضية H_1.

مثال $(1-3)$:

تعاقدت إحدى الشركات على شراء كمية من الأنابيب الحديدية، بلغ متوسط قطرها 2 سم وانحرافه المعياري 0.01 سم. سحبت الشركة من الشحنة عينة عشوائية مؤلفة من 100 أنبوب بلغ متوسط القطر 2.02 سم، فهل تقبل الشركة الشحنة أم ترفضها إذا علمت أن مستوى الدلالة $\alpha = 0.05$.

الحل:

الفرضية الابتدائية H_0 : لا توجد فرق جوهري بين الوسط الحسابي للعينة والمجتمع

$$H_0 : \mu = \mu_0$$

الفرضية البديلة H_1 : توجد فرق جوهري بين الوسط الحسابي للعينة والمجتمع

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

نحسب مؤشر الاختبار:

$$t = \frac{\bar{x} - \mu}{\tilde{\sigma}_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

لدينا: $\bar{x} = 2.02$ و$\tilde{\sigma}_{\bar{x}} = 0.01$ و $\mu = 2$

$$t = \frac{\bar{x} - \mu}{\tilde{\sigma}_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

نعوض في t فنجد:

$$t = \frac{2.02 - 2}{0.01} = 2$$

نجد أن القيمة الحرجة لـ t والمقابلة لمستوى دلالة $\alpha = 0.05$ تساوي(1.96)، نقارن قيمة t الفعلية(2)مع t الحرجة (1.96)، نجد أن القيمة الفعلية هي أكبر من الحرجة ،وبالتالي: نعتبر الفرق بين الوسط الحسابي للعينة والمجتمع جوهرياً، أي نرفض فرضية العدم .

مثال $(3-2)$:

ادعى أحدهم أن متوسط دخل الأسرة في بلده يساوي 15000 وحدة في الشهر وللتحقق من ذلك سُحبت عينة عشوائية مؤلفة من 35 أسرة، فكان متوسط دخولهم 12000 وحدة في الشهر، و الانحراف المعياري للعينة 6000 وحدة. اختبر صحة الادعاء بمستوى دلالة $\alpha = 0.05$.

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : \bar{y} = 15000$$
$$H_1 : \bar{y} \neq 15000$$

نلاحظ أن حجم العينة كبير نسبياً، وأن $\alpha = 0.05$.

نحسب قيمة مؤشر الاختبار:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{12000 - 15000}{\frac{6000}{\sqrt{35}}} = \frac{-3000}{101419} = -2.96$$

الاختبار ثنائي الجانب، نبحث عن قيمة $Z_{1-\frac{\alpha}{2}}$ المقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ في جدول التوزيع الطبيعي المعياري، فنجد أن: $Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$

بالمقارنة نجد أن $|t| > Z_{1-\frac{\alpha}{2}} \Rightarrow |-2.96| < 1.96$ ، وبالتالي: نرفض الفرضية H_0 والتي تقول بأن متوسط الدخل يساوي 15000 وحدة .

مثال $(3-3)$:

نفترض أن متوسط سعر مادة ما خلال العام كان 12ل.س وعندما سحبنا عينة بحجم15 قياساً لسعرها كانت كالآتي:

$X: 12, 13, 14, 15, 11, 10, 9, 8, 10, 13, 14, 12, 15, 14, 12$

اختبر صحة ذلك القول باحتمال قدره $\beta = 0.99$ (بمستوى دلالة $\alpha = 0.01$) .

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : \bar{y} = 12$$
$$H_1 : \bar{y} > 12$$

نلاحظ أن حجم العينة صغير نسبياً، وأن $\alpha = 0.01$.

قبل أن نحسب قيمة مؤشر الاختبار، لابد من حساب كلاً من $\bar{x}$ و s^2 كالآتي:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{15} x_i \Rightarrow \bar{x} = 12.13$$
$$s^2 = \frac{1}{n-1}\sum^{n}(x_i - \bar{x})^2 \Rightarrow s^2 = 3.85 \Rightarrow s = 1.96$$

نعوض في العلاقة الآتية:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{12.13 - 12}{\frac{1.96}{\sqrt{15}}} = \frac{0.13}{0.5060} = 0.25$$

الاختبار أحادي الجانب يميني، نبحث عن قيمة $t_{1-\alpha}$ المقابلة للاحتمال $(1-\alpha)$ و$(n-1)$ لدرجة حرية في جدول توزيع ستيودنت، فنجد أن: $t_{(0.99,14)} = 2.60$

بالمقارنة نجد أن $|0.25| < 2.60 \Rightarrow |t| > t_{1-\alpha}$، وبالتالي: نقبل الفرضية H_0 والتي تقول بأن متوسط السعر يساوي 12ل.س.

مثال $(4-3)$:

سحبنا عينة بحجم $n = 100$ عنصراً من مجتمع طبيعي، فوجدنا أن متوسطها $\bar{x} = 2.7$ وأن $s^2 = 23$

اختبر الفرضية $H_0 : \bar{y} = 3$ بمستوى دلالة $\alpha = 0.10$.

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : \bar{y} = 3$$
$$H_1 : \bar{y} \neq 3$$

نلاحظ أن حجم العينة كبير نسبياً، وأن $\alpha = 0.10$.

نحسب قيمة مؤشر الاختبار :

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{2.7-3}{\frac{4.80}{\sqrt{100}}} = \frac{-0.3}{0.48} = -0.625$$

الاختبار ثنائي الجانب، نبحث عن قيمة $Z_{1-\frac{\alpha}{2}}$ المقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ في جدول التوزيع الطبيعي المعياري، فنجد أن: $Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.10}{2}} = Z_{0.95} = 1.64$.

بالمقارنة نجد أن: $|-0.626| < 1.64 \Rightarrow |t| < Z_{1-\frac{\alpha}{2}}$، وبالتالي: نقبل الفرضية H_0 .

3 – 2 – 4 اختبارات حول النسبة في المجتمع:

إن إجراء اختبار حول نسبة ما في المجتمع R لا يختلف كثيراً عن إجرائه حول المتوسط حيث نضع الفرضية الابتدائية على الشكل الآتي:

$$H_0 : R = R_0 \qquad (7-3)$$

و نضع الفرضية البديلة على الشكل الآتي:

$$H_1 : R \neq R_0 \qquad (8-3)$$

ثم نسحب عينة عشوائية ونفترض أن نسبة الخاصة المدروسة في المجتمع كانت r .

أما مؤشر الاختبار فيحسب من العلاقة الآتية:

$$t = \frac{r - R_0}{\tilde{\sigma}_r} = \frac{r - R_0}{\sqrt{\frac{r.q}{n}}}$$

ومن ثم نقوم بمقارنة قيمة t المحسوبة مع معامل الاحتمال الطبيعي Z_0 والمقابل لنصف مستوى الدلالة المحدد. ونقبل الفرضية H_0 إذا كانت القيمة المطلقة لـ t أصغر من Z_0 .وفي حال العينات الصغيرة نقارن t مع t_0 لتوزيع ستيودنت والمقابلة لنصف مستوى الدلالة و $(n-1)$ درجة حرية .

مثال $(3-5)$:

أدعى أحد المرشحين أن 80% من السكان سينتخبونه ،ولكن عينة من $n = 500$ شخص بينت أن 500 شخص يؤيدون ذلك المرشح.اختبر ادعاء المرشح باحتمال قدره 0.95 .

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : R = R_0 = 0.80$$
$$H_1 : R > 0.80$$

نلاحظ أن حجم العينة كبير نسبياً، وأن $\alpha = 0.05$.

نحسب قيمة مؤشر الاختبار:

$$t = \frac{r - R_0}{\sqrt{\frac{r.q}{n}}} = \frac{0.60 - 0.80}{\sqrt{\frac{(0.60).(0.40)}{500}}} = -9.13$$

الاختبار أحادي الجانب، نبحث عن قيمة $Z_{1-\alpha}$ المقابلة للاحتمال $(1-\alpha)$ في جدول التوزيع الطبيعي المعياري، فنجد أن: $Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.64$

بالمقارنة نجد أن $|t| > Z_{1-\alpha} \Rightarrow |-9.13| > 1.64$ ، وبالتالي: نرفض الفرضية H_0. والتي تقول بأن نسبة المؤيدين للمرشح كانت 0.80 .

مثال $(6-3)$:

لمعرفة نسبة الأسر التي تملك سيارة سياحية. سحبنا عينة من $n = 625$ أسرة فوجدنا أن $r = 0.48$ فإذا كان مستوى الدلالة $\alpha = 0.05$. **المطلوب**: اختبار فيما إذا كانت نسبة الأسر التي تملك سيارة سياحية لم تتغير بشكل جوهري.

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : R = R_0 = 0.50$$
$$H_1 : R < 0.50$$

نلاحظ أن حجم العينة كبير نسبياً، وأن $\alpha = 0.05$.

نحسب قيمة مؤشر الاختبار:

$$t = \frac{r - R_0}{\sqrt{\frac{r.q}{n}}} = \frac{0.48 - 0.50}{\sqrt{\frac{(0.48).(0.52)}{625}}} = -1.0008$$

الاختبار أحادي الجانب، نبحث عن قيمة $Z_{1-\alpha}$ المقابلة للاحتمال $(1-\alpha)$ في جدول التوزيع الطبيعي المعياري، فنجد أن: $Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.64$

بالمقارنة نجد أن $|-1.0008| < 1.64 \Rightarrow |t| > Z_{1-\alpha}$، وبالتالي: نقبل الفرضية H_0. والتي تقول بأن نسبة الأسر التي تملك سيارة سياحية لم تتغير بشكل جوهري.

3 – 2 – 5 اختبارات حول الفرق بين متوسطي مجتمعين طبيعيين:

عندما يكون لدينا مجتمعان متوسطهما $\bar{y}_1$ و $\bar{y}_2$ ، فإننا نحتاج إلى مقارنة هذين المتوسطين عن طريق اختبار الفرق بينهما. لذلك نقوم بسحب عينتين مستقلتين n_1 و n_2 على الترتيب. فإذا كان متوسطا الخاصة المدروسة في هاتين العينتين يساويان $\bar{x}_1$ و $\bar{x}_2$ ، وكان تباينهما s_1^2 و s_2^2 على الترتيب. فإن تقدير الفرق $(\bar{y}_1 - \bar{y}_2)$ يعطى بواسطة الفرق $(\bar{x}_1 - \bar{x}_2)$ ، أي أن:

$$(\bar{y}_1 - \bar{y}_2) = (\bar{x}_1 - \bar{x}_2) \qquad (9-3)$$

ونعلم أن تقدير تباين الفرق $(\bar{x}_1 - \bar{x}_2)$ يعطى بالعلاقة:

$$\tilde{\sigma}^2_{(\bar{x}_1 - \bar{x}_2)} = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \qquad (10-3)$$

لذلك فإننا نضع الفرضية الابتدائية الآتية:

$$H_0 : \bar{y}_1 - \bar{y}_2 = \left(\bar{y}_1 - \bar{y}_2\right)_0 \qquad (11-3)$$

أما الفرضية البديلة تأخذ الشكل الآتي:

$$H_1 : \bar{y}_1 - \bar{y}_2 \neq \left(\bar{y}_1 - \bar{y}_2\right)_0 \qquad (12-3)$$

إذا كان الاختبار ثنائي الجانب .

أما إذا كان الاختبار أحادي الجانب (يميني)،تأخذ الفرضية البديلة الشكل الآتي:

$$H_1 : \bar{y}_1 - \bar{y}_2 > \left(\bar{y}_1 - \bar{y}_2\right)_0 \qquad (13-3)$$

أما إذا كان الاختبار أحادي الجانب (يساري)،تأخذ الفرضية البديلة الشكل الآتي:

$$H_1 : \bar{y}_1 - \bar{y}_2 < \left(\bar{y}_1 - \bar{y}_2\right)_0 \qquad (14-3)$$

ويحسب مؤشر الاختبار من العلاقة الآتية:

$$t = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\bar{y}_1 - \bar{y}_2\right)_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (15-3)$$

نقارن قيمة t مع $Z_{1-\frac{\alpha}{2}}$ (إذا كان الاختبار ثنائي الجانب)،فإذا كانت $|t| < Z_{1-\frac{\alpha}{2}}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية البديلة H_1 .والعكس صحيح.

أما إذا كان الاختبار أحادياً يمينياً،فإننا نقبل الفرضية H_0 إذا كانت $t < Z_{1-\frac{\alpha}{2}}$ ونرفض الفرضية البديلة H_1 .والعكس صحيح .

وإذا الاختبار أحادياً يسارياً، فإننا نقبل الفرضية H_0 إذا كانت $t > -Z_{1-\frac{\alpha}{2}}$ ونرفض الفرضية البديلة H_1 .

والعكس صحيح .

حالة خاصة1 :

إذا كان تباينا المجتمعين متساويين،أي كان $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ،فإنه يمكننا كتابة تباين الفرق $(\bar{x}_1 - \bar{x}_2)$ على الشكل الآتي:

$$\tilde{\sigma}^2_{(\bar{x}_1-\bar{x}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \qquad (16-3)$$

كما ويعطى تقدير القيمة المشتركة للتباين σ^2 بواسطة الوسط الحسابي المثقل لتباين العينتين s_1^2 و s_2^2 من خلال العلاقة الآتية :

$$\tilde{\sigma}^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = s^2 \qquad (17-3)$$

وبذلك نجد أن تقدير $\tilde{\sigma}^2_{(\bar{x}_1-\bar{x}_2)}$ في هذه الحالة يعطى بالعلاقة:

$$\tilde{\sigma}^2_{(\bar{x}_1-\bar{x}_2)} = s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right) \qquad (18-3)$$

ويأخذ مؤشر الاختبار في هذه الحالة الشكل الآتي:

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\bar{y}_1-\bar{y}_2)_0}{s.\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \qquad (19-3)$$

وهو مقدار خاضع لتوزيع ستيودنت ذي (n_1+n_2-2) درجة حرية. لذلك فإننا نقارن قيمة t مع قيمة $t_{1-\frac{\alpha}{2}}$ أو $t_{1-\alpha}$ المقابلة للاحتمال $1-\frac{\alpha}{2}$ أو $1-\alpha$ ولـ(n_1+n_2-2) درجة حرية.

فإذا كانت $|t| \le t_{1-\frac{\alpha}{2}}$، وكان الاختبار ثنائي الجانب فإننا نقبل H_0 ونرفض H_1. أما إذا كان الاختبار أحادي الجانب (يميني) فإننا نقبل H_0 إذا كان $t < t_{1-\alpha}$. وعندما يكون أحادي الجانب(يساري) فإننا نقبل H_0 إذا كان $t > -t_{1-\alpha}$.

حالة خاصة 2: (اختبار حول متوسطي عينتين مستقلتين من مجتمع واحد) وتباينهما متساويان $\sigma_1^2=\sigma_2^2$

إذا سحبنا عينتين بحجمين n_1 وn_2 من مجتمع واحد، فإنه يكون لدينا حكماً في هذه الحالة ما يلي:

$\bar{y}_1 = \bar{y}_2 = \bar{y}$ ، $\sigma_1^2 = \sigma_2^2 = \sigma^2$

ويأخذ مؤشر الاختبار العلاقة الآتية:

$$t = \frac{(\bar{y}_1 - \bar{y}_2) - 0}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (20-3)$$

ويتحول اختبار الفرضية H_0 إلى اختبار جوهرية الفرق بين متوسطي العينتين $\bar{x}_1$ و $\bar{x}_2$ وليس بين $(\bar{y}_1 - \bar{y}_2)$.

ولاتخاذ قرار حول جوهرية الفرق بين $\bar{x}_1$ و $\bar{x}_2$ ، فإننا نقارن قيمة t مع $Z_{1-\frac{\alpha}{2}}$ ، عند الاحتمال $\left(1-\frac{\alpha}{2}\right)$ فإذا كانت $|t| \leq Z_{1-\frac{\alpha}{2}}$ فإننا نعتبر الفرق $(\bar{x}_1 - \bar{x}_2)$ غير جوهري وعائد لأسباب عشوائية. أما إذا كانت $|t| \geq Z_{1-\frac{\alpha}{2}}$ فإننا نعتبر الفرق $(\bar{x}_1 - \bar{x}_2)$ جوهري وعائد لأسباب غير عشوائية.

مثال $(7-3)$:

نفترض أننا نريد أن ندرس أداء الطالب في مقرر الرياضيات في مدرستين مختلفتين، سحبنا عينة من طلاب المدرسة الأولى بحجم $n_1 = 50$ وعينة من طلاب المدرسة الثانية بحجم $n_2 = 60$ فكان متوسط الدرجات فيها $\bar{x}_1 = 70$ و $\bar{x}_2 = 65$ وتباينهما $s_1^2 = 150$ و $s_2^2 = 300$. **والمطلوب** اختبار الفرضية التي تقول أن الفرق بين متوسطي الدرجات في المدرستين غير جوهري وذلك بمستوى دلالة قدره $\alpha = 0.05$.

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : \bar{y}_1 - \bar{y}_2 = y_0 = 0$$
$$H_1 : \bar{y}_1 - \bar{y}_2 \neq 0$$

نحسب قيمة مؤشر الاختبار :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - y_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(70-65)}{\sqrt{\frac{150}{50} + \frac{300}{60}}} = 1.77$$

الاختبار ثنائي الجانب،نبحث عن قيمة $Z_{1-\frac{\alpha}{2}}$ المقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ في جدول التوزيع الطبيعي المعياري، فنجد أن: $Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$

بالمقارنة نجد أن $1.96 > |1.77| \Rightarrow |t| < Z_{1-\frac{\alpha}{2}}$، وبالتالي: نقبل الفرضية H_0 .والتي تقول بعدم وجود فرق جوهري بين متوسطي درجات الطلاب.

مثال $(8-3)$:

سحبنا عينتين بحجمين $n_1 = n_2 = 20$ عنصراً من مجتمعين طبيعيين متوسطهما $\bar{y}_1$ و $\bar{y}_2$ ولهما تباين موحد، فوجدنا أن $\bar{x}_1 = 60$ و $\bar{x}_2 = 65$ وتباينهما $s_1^2 = 30$ و $s_2^2 = 45$.فهل تقدم هذه المعلومات دلالة كافية باحتمال قدره 0.95 على وجود فرق بين المتوسطين $\bar{y}_1$ و $\bar{y}_2$.

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : \bar{y}_1 - \bar{y}_2 = y_0 = 0$$
$$H_1 : \bar{y}_1 - \bar{y}_2 \neq 0$$

نحسب قيمة مؤشر الاختبار :

$$t=\frac{(\bar{x}_1-\bar{x}_2)-y_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}=\frac{(60-65)-0}{\sqrt{\frac{30}{20}+\frac{45}{20}}}=-2.58$$

الاختبار ثنائي الجانب،نبحث عن قيمة $Z_{1-\frac{\alpha}{2}}$ المقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ في جدول التوزيع الطبيعي المعياري،فنجد أن: $Z_{1-\frac{\alpha}{2}}=Z_{1-\frac{0.05}{2}}=Z_{0.975}=1.96$.

بالمقارنة نجد أ $|t|>Z_{1-\frac{\alpha}{2}} \Rightarrow |-2.56|>1.96$ ،وبالتالي: نرفض الفرضية H_0 .والتي تقول بعدم وجود فرق جوهري بين متوسطي درجات الطلاب.

مثال $(3-9)$:

لدى استطلاع للرأي حول الإنفاق الشهري على المواد الغذائية في محافظة ما،فقد تم سحب عينة مؤلفة من 327 أسرة عمالية وأخرى مؤلفة من 286 أسرة من ذوي الدخل المحدود وقد تم دراسة العينتين فتبين ما يلي:

معلومات عن الأسر العمالية: $n_1=327 \quad \bar{x}_1=612 \quad s_1=104$

معلومات عن الأسر ذوي الدخل المحدود: $n_2=286 \quad \bar{x}_2=642 \quad s_2=118$

فهل يمكننا استنتاج أن أسر ذوي الدخل المحدود تنفق أكثر من الأسر العمالية على المواد الغذائية ضمن مستوى دلالة.

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0 : \bar{y}_1 - \bar{y}_2 = y_0 = 0$$
$$H_1 : \bar{y}_1 - \bar{y}_2 \neq 0$$

نحسب قيمة مؤشر الاختبار :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - y_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(612 - 642) - 0}{\sqrt{\frac{10816}{327} + \frac{13924}{286}}} = -3.32$$

نجد أن $(n_1 + n_2 - 2)$ كبير وبما أن $\alpha = 0.05$ والاختبار ثنائي الجانب، نبحث عن قيمة $Z_{1-\frac{\alpha}{2}}$ المقابلة

للاحتمال $\left(1 - \frac{\alpha}{2}\right)$ في جدول التوزيع الطبيعي المعياري، فنجد أن: $Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.975} = 1.96$

بالمقارنة نجد أن $1.96 < |-3.32| \Rightarrow |t| > Z_{1-\frac{\alpha}{2}}$ ، وبالتالي: نرفض الفرضية H_0 .والتي تقول بعدم وجود

فرق جوهري بين النفقات الغذائية للأسرة ذوي الدخل المحدود وبين النفقات للأسر العمالية.

مثال $(3-10)$:

أراد أحد أساتذة كلية الطب أن يدرس فترة حياة مرض ما. علماً بأن هذا المرض قد يُصاحب بارتفاع درجة الحرارة وقد لا يُصاحب. ونتيجة للدراسة كان متوسط حياة هذا المرض هو 11 يوماً لخمسة لم يصابوا بارتفاع درجة الحرارة و 17 يوماً لأربعة مرضى أصيبوا بارتفاع درجة الحرارة، وقد تم تقدير تباين مدة المرض فكان $s^2 = 20$ ، وأن مستوى الدلالة قد حدد بـ $\alpha = 0.01$.

ماذا نستخلص من ذلك ؟ وما هي الفرضية التي يمكن تطبيقها.

الحل:

تم سحب العينتين $n_1 = 4$ و $n_2 = 5$ من مجتمع واحد. لذلك تتحول العملية من اختبار فرضيات إلى حساب مؤشر t فقط، بحيث يعطى المؤشر كالآتي:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s.\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

حيث أن $\bar{x}_1$ متوسط مدة المرض للمرضى الذين يصابون أيضاً بارتفاع درجة الحرارة أثناء المرض.

$\bar{x}_2$ متوسط مدة المرض للمرضى الذين لا يصابون أيضاً بارتفاع درجة الحرارة أثناء المرض.

إن المتغير $(\bar{x}_1 - \bar{x}_2)$ يتبع قانون توزيع احتمالي t ستيودنت عند $(n_1 + n_2 - 2)$ درجة حرية، وبالتالي: فإن قيمة مؤشر الاختبار هي:

$$t = \frac{(17-11)}{4.47.\sqrt{\frac{1}{4} + \frac{1}{5}}} = 2.0095$$

إن قيمة t التي تخضع لتوزيع ستيودنت التي نحصل عليها من خلال مستوى الدلالة $\alpha = 0.05$ وعدد درجات حرية $(\nu = n_1 + n_2 - 2 = 7)$ نجدها في جدول توزيع ستودنت كالآتي: $t_{(\alpha,\nu)} = t_{(0.01,7)} = 3.499$.

بالمقارنة نجد أن: $t < t_{(\alpha,\nu)} \Rightarrow 2 < 3.499$، لذلك لا نستطيع أن نقول بوجود فرق جوهري بين متوسطي العينتين وهذا يعني أن الفرق الحاصل بينهما ناتج عن الجوانب العشوائية .

3 – 2 – 6 اختبارات حول الفرق بين نسبتين في مجتمعين طبيعيين:

عندما نريد اختبار الفرق بين النسبتين $(R_1 - R_2)$ في مجتمعين مختلفين، نستخدم الفرق بين هاتين النسبتين في العينة وهو $(r_1 - r_2)$.

تكون الفرضية الابتدائية كالآتي:

$$H_0 : R_1 - R_2 = (R_1 - R_2)_0 \qquad (21-3)$$

أما الفرضية البديلة تأخذ الشكل الآتي:

$$H_1 : R_1 - R_2 \neq \left(R_1 - R_2\right)_0 \qquad (22-3)$$

إذا كان الاختبار ثنائي الجانب .

أما إذا كان الاختبار أحادي الجانب(يميني)،تأخذ الفرضية البديلة الشكل الآتي:

$$H_1 : R_1 - R_2 > \left(R_1 - R_2\right)_0 \qquad (23-3)$$

أما إذا كان الاختبار أحادي الجانب (يساري)،تأخذ الفرضية البديلة الشكل الآتي:

$$H_1 : R_1 - R_2 < \left(R_1 - R_2\right)_0 \qquad (24-3)$$

ويحسب مؤشر الاختبار من العلاقة الآتية:

$$t = \frac{\left(r_1 - r_2\right) - \left(R_1 - R_2\right)_0}{\sqrt{\frac{r_1 q_1}{n_1} + \frac{r_2 q_2}{n_2}}} \qquad (25-3)$$

ولاتخاذ القرار الإحصائي السليم، نقارن قيمة t الفعلية مع $Z_{1-\frac{\alpha}{2}}$ أو $Z_{1-\alpha}$ الحرجة حسب اتجاه الاختبار (ثنائي أم آحادي)،فإذا كانت $|t| < Z_{1-\frac{\alpha}{2}}$ فإننا نقبل الفرضية H_0 ونرفض الفرضية البديلة H_1 .والعكس صحيح.

حالة خاصة1:

إذا كان للمجتمعين المدروسين تباينان متساويان،أي كان $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ،فإن مؤشر الاختبار لهذا الفرق يعطى بالعلاقة:

$$t = \frac{(r_1 - r_2) - (R_1 - R_2)_0}{s.\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (26-3)$$

حيث أن:

$$s = \sqrt{\frac{n_1 r_1 q_1 + n_2 r_2 q_2}{n_1 + n_2 - 2}} \qquad (27-3)$$

وأن t خاضع لتوزيع(ستودينت) ذي$(n_1 + n_2 - 2)$ درجة حرية. ولتحديد نتيجة الاختبار ثنائي الجانب نقارن قيمة t مع قيمة $t_{\left(1-\frac{\alpha}{2}, n_1+n_2-2\right)}$ الجدولية (أو مع $t_{(1-\alpha, n_1+n_2-2)}$ إذا كان الاختبار آحادي الجانب).

فإذا كانت $|t| < t_{\left(1-\frac{\alpha}{2}, n_1+n_2-2\right)}$ نقبل H_0 ونرفض H_1 .والعكس صحيح .

حالة خاصة 2 :

إذا سحبنا عينتين بحجمين n_1 و n_2 من مجتمع واحد،فإنه يكون لدينا حكماً في هذه الحالة كالآتي:

$R_1 - R_2 = 0$ ، $\sigma_1^2 = \sigma_2^2 = \sigma^2$

ويأخذ مؤشر الاختبار العلاقة الآتية:

$$t=\frac{(r_1-r_2)-0}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \qquad (28-3)$$

حيث أن s تعطى بالعلاقة الآتية :

$$s=\sqrt{\frac{n_1r_1q_1+n_2r_2q_2}{n_1+n_2-2}} \qquad (29-3)$$

ويتحول اختبار الفرضية H_0 إلى اختبار جوهرية الفرق بين النسبتين (r_1-r_2) في العينتين.

ولاتخاذ قرار حول جوهرية الفرق بين $\bar{x}_1$ و $\bar{x}_2$ ،فإننا نقارن قيمة t مع $Z_{1-\frac{\alpha}{2}}$ ،عند الاحتمال $\left(1-\frac{\alpha}{2}\right)$ فإذا كانت $|t| \le Z_{1-\frac{\alpha}{2}}$ فإننا نعتبر الفرق (r_1-r_2) غير جوهري وعائد لأسباب عشوائية.أما إذا كانت $|t| \ge Z_{1-\frac{\alpha}{2}}$ فإننا نعتبر الفرق (r_1-r_2) جوهري وعائد لأسباب غير عشوائية.

مثال $(11-3)$:

لدراسة نسبة المدخنين في مجتمعين فيهما $\sigma_1^2=\sigma_2^2=\sigma^2$ ،سحبنا عينتين بحجمين $n_1=10$ و $n_2=15$ شخصاً،فوجدنا أن نسبة المدخنين فيهما تساويان $r_1=0.20$ و $r_2=0.25$ على الترتيب.فهل هنالك دلالة كافية لاعتبار أن نسبتي المدخنين R_1 و R_2 مختلفتان وذلك بمستوى دلالة قدره $\alpha=0.05$

الحل:

لإجراء الاختبار نضع الفرضيتين الآتيتين:

$$H_0: R_1-R_2=0$$
$$H_1: R_1-R_2\neq 0$$

نحسب قيمة مؤشر الاختبار :

$$t=\frac{(r_1-r_2)-R_0}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{(0.20-0.25)}{\sqrt{\frac{10(0.20)(0.80)+15(0.25)(0.75)}{10+15-2}}\sqrt{\frac{1}{10}+\frac{1}{15}}}=-0.2796$$

الاختبار ثنائي الجانب،نبحث عن قيمة $t_{1-\frac{\alpha}{2}}$ المقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ ولـ (n_1+n_2-2) درجة حرية في جدول توزيع ستودينت،فنجد أن: $t_{1-\frac{\alpha}{2}}=t_{1-\frac{0.05}{2}}=t_{(0.97523)}=2.07$

بالمقارنة نجد أن $|t|<t_{1-\frac{\alpha}{2}}\Rightarrow|-0.2796|<2.97$ وبالتالي: نقبل الفرضية H_0. والتي تقول بعدم وجود فرق جوهري بين النسبتين المذكورتين في المجتمعين .

مثال $(3-12)$:

في دراسة لنسبة المصابين بأمراض قلبية في مجتمع ما،سحبنا عينتين بحجمين $n_1=12$ و $n_2=15$، فوجدنا أن نسبتي المصابين بمرض قلبي فيهما تساويان $r_1=0.10$ و $r_2=0.15$. فهل يعتبر الفرق بين هاتين النسبتين فرقاً جوهرياً إذا كان مستوى الدلالة $\alpha=0.05$.

الحل:

بما أن العينتين سحبتا من مجتمع واحد، فإن مؤشر الاختبار يحسب من العلاقة الآتية:

$$t=\frac{(r_1-r_2)-R_0}{\sqrt{\frac{n_1r_1q_1+n_2r_2q_2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$
$$=\frac{(0.10-0.15)}{\sqrt{\frac{12(0.10)(0.90)+15(0.15)(0.85)}{12+15-2}}\sqrt{\frac{1}{12}+\frac{1}{15}}}=-0.0289$$

الاختبار ثنائي الجانب،نبحث عن قيمة $t_{1-\frac{\alpha}{2}}$ المقابلة للاحتمال $\left(1-\frac{\alpha}{2}\right)$ و لـ (n_1+n_2-2) درجة حرية في جدول توزيع ستودينت،فنجد أن: $t_{1-\frac{\alpha}{2}} = t_{1-\frac{0.05}{2}} = t_{(0.97523)} = 2.07$.

بالمقارنة نجد $2.07 > |-0.12| \Rightarrow |t| < t_{1-\frac{\alpha}{2}}$ ،وبالتالي: نقول بعدم وجود فروق جوهرية بين النسبتين المذكورتين .

3 – 2 – 7 اختبارات حول القياسات الشاذة لحذفها:

القياسات الشاذة هي بعض القياسات المتطرفة والتي تختلف كثيراً عن باقي القيم الأخرى،ويعود ذلك لأسباب مختلفة منها خطأ النقل أو النسخ أو التسجيل ...الخ.

خطأ القياس ينجم عن سوء استخدام الأجهزة أو تعطلها.وحتى تبقى هذه القياسات تعطي نتائج واضحة وجيدة وبعيدة عن التشويه،نقوم بدراسة صحة انتمائها إلى مجموعة القياسات المفروضة وذلك باستخدام مبادئ اختبار الفرضيات. فإذا كان الاختبار يؤيد انتمائها نتركها ضمن المجموعة،وإذا كان العكس نقوم بحذفها من المجموعة. ولنفترض أنه لدينا القياسات الآتية:

$x_0,\ x_1,\ x_2,\ x_3,\ ....x_n$

حيث نعتبر أن x_0 هو قياس مشكوك بعدم انتمائه للمجموعة بسبب اختلاف قيمته عن الآخرين.

نضع الفرضيتين الآتيتين:

x_0 ينتمي إلى المجموعة لأنه قريب من $\bar{x}$

$$H_0 : x_0 = \bar{x}$$

x_0 لا ينتمي إلى المجموعة لأنه أكبر من $\bar{x}$ بكثير

$$H_0 : x_0 > \bar{x}$$

نقوم بعزل القياس x_0 عن المجموعة، ثم نحسب متوسط العناصر المتبقية $\bar{x}$ وانحرافها المعياري s ونحسب قيمة مؤشر الاختبار كالآتي:

$$t = \frac{x_0 - \bar{x}}{s} \qquad (30-3)$$

مؤشر الاختبار السابق هو متحول خاضع لتوزيع ستودنيت ذي $(n-1)$ درجة حرية.

ولاتخاذ قرار حول x_0 نقارن قيمة t مع قيمة متحول(ستودينت) $t_{(1-\alpha,n-1)}$ المقابلة للاحتمال $(1-\alpha)$ ولـ $(n-1)$ درجة حرية.

فإذا كان الاختبار أحاديا(يمينياً) وكانت $t < t_{(1-\alpha,n-1)}$، فإننا نقبل الفرضية H_0 ونعتبر أن x_0 ينتمي إلى المجموعة. وإذا كان العكس، فإننا نعتبر أن x_0 غير منتم للمجموعة وهو عبارة عن قيمة شاذة، لذلك نقوم بحذفه من المجموعة بشكل نهائي حتى لا يشوه لنا نتائج التحليل الإحصائي.

أما إذا كان الاختبار أحاديا(يسارياً) وكانت $|t| \geq -t_{(1-\alpha,n-1)}$، فإننا نرفض الفرضية H_0 ونعتبر أن x_0 لا ينتمي إلى المجموعة وهو عبارة عن قيمة شاذة، لذلك نقوم بحذفه من المجموعة بشكل نهائي حتى لا يشوه لنا نتائج التحليل الإحصائي. وإذا كان العكس، فإننا نعتبر أن x_0 منتم للمجموعة.

مثال $(13-3)$:

لتكن لدينا القياسات الآتية عن وزن سبعة صناديق تفاح:

25.5, 24.1, 23.5, 23.4, .23.3, 23.2, 23.8

فهل القياس $x_0 = 25.5$ يعد من قياسات هذه المجموعة. اختبر بمستوى دلالة $\alpha = 0.05$.

الحل:

نضع الفرضيتين الآتيتين:

$$H_0 : x_0 = \bar{x}$$
$$H_1 : x_0 > \bar{x}$$

نعزل x_0 ثم نحسب $\bar{x}$ والانحراف المعياري s للقياسات الأخرى كالآتي:

$$\bar{x} = \frac{1}{n}\sum x_i = \frac{1}{6}(141.3) = 23.55$$

$$s^2 = \frac{1}{n-1}\sum \left(x_i - \bar{x}\right)^2 = 0.115 \Rightarrow s = 0.34$$

نجد مؤشر الاختبار كالآتي:

$$t = \frac{25.5 - 23.5}{0.34} = 5.74$$

ولاتخاذ قرار حول x_0 نقارن قيمة t مع قيمة متحول(ستودينت $t_{(0.95,6)} = 2.02$ المقابلة للاحتمال 0.95 ولـ 6 درجات حرية .

بالمقارنة نجد $5.43 < 2.02 \Rightarrow t < t_{1-\alpha}$ ، وبالتالي: نعتبر القياس $x_0 = 25.5$ قياساً شاذاً و نحذفه من المجموعة.

تمارين عامة

1 – سحبنا عينة بحجم $n=10$ عنصراً من مجتمع طبيعي، فوجدنا أن متوسطها $\bar{x}=2.7$ ، وأن $s^2=23$.
والمطلوب: اختبر الفرضية $H_0: \bar{y}=3$ بمستوى دلالة $\alpha=0.10$.

2 – يدعي تاجر أن متوسط سعر بيع الوحدة من سلعة معينة في يوم معين لا يزيد عن 150 ل.س، فإذا وجد من عينة عشوائية حجمها 25 من الوحدات التي بيعت ذلك اليوم، وأن متوسط سعر البيع لهذه الوحدات $\bar{x}=130$ ،
والمطلوب: اختير بمستوى دلالة $\alpha=0.10$ صحة ادعاء التاجر، إذا كان من المعلوم أن سعر البيع له توزيع طبيعي بتوقع μ وتباين معلوم σ^2 ويساوي 80 ل.س .

3 – يدعي أحدهم أن متوسط الدخل الشهري في بلده يساوي 7500 ل.س، سحبت عينة عشوائية تضم 9 أفراد من مجتمع طبيعي تباينه معلوم $\sigma^2=81000$ ، فكانت النتيجة أن متوسط الدخل في العينة $\bar{x}=7000$ ، **والمطلوب** :
اختير بمستوى دلالة $\alpha=0.5$ صحة الإدعاء السابق .

4 – مصنعان ينتجان مصابيح كهربائية صغيرة بأسلوبين مختلفين، تم سحب عينة من كل مصنع بهدف دراسة متوسط عمر المصابيح فحصلنا على المعلومات الآتية :

$n_1=25$ و $n_2=20$ و $\bar{x}_1=1261$ و $\bar{x}_2=1832$ و $\sigma_1^2=497$ و $\sigma_2^2=501$ ، بالإضافة إلى ذلك نفترض أن اختلاف أسلوب الإنتاج لا يؤثر على تباين حياة المصابيح، أي أن التباين متشابه في المصنعين σ^2. **فهل** نستطيع أن نعتبر أن الفرق بين متوسط حياة المصابيح جوهري مع مستوى دلالة $\alpha=0.5$.

5 – لتقدير نسبة المتعطلين عن العمل في مدينتين A و B سحبنا منهما عينتين بحجمين $n_1=35$ و $n_2=45$ شخصاً، فوجدنا أن نسبة المتعطلين عن العمل فيهما $R_1=0.12$ و $R_2=0.15$ ، فهل يعطينا هذا دلالة كافية لاعتبار هاتين النسبتين متساويتين وذلك عند مستوى دلالة $\alpha=0.10$.

6 – لمعرفة أراء العاملين في شركة كبيرة بشأن خطة جديدة للحوافز، اختيرت عينة عشوائية من الفئة الدنيا حجمها 100 مستخدم، ووجد أن عدد المؤيدين للخطة في هذه العينة يساوي 40 ، واختيرت عينة عشوائية من الفئة المتوسطة حجمها 200 مستخدم، وتبين أن عدد المؤيدين للخطة في هذه العينة يساوي 100. **فهل** يمكن القول بمستوى معنوية $\alpha=0.05$ أن نسبة المؤيدين للخطة الجديدة للحوافز متساوية في الفئتين.