المحور الرابع: اختبار الفروض Tests of Hypotheses

تمهيد:

المقصود بالفروض هنا الفروض الإحصائية statistical hypotheses بمعنى الفروض التي تتعلق بالمجتمع الإحصائي المسحوبة منه العينة، أو توزيع هذا المجتمع أو معالمه كالوسط الحسابي أو النسبة في المجتمع.

والفرض ما هو إلا تخمين أو استنتاج ذكي مبني على حيثيات معقولة أو منطقية ولكنه ليس مبنياً على حسابات دقيقة خاصة بالمجتمع لأننا نفترض أنه لا يمكن دراسة المجتمع بالكامل عن طريق الحصر الشامل بل نحاول استنتاج أو الاستدلال على مقاييس المجتمع باستخدام بيانات ونتائج العينة.

نهدف من خلال هذا المحور الى :

- التعرف على طريقة وضع الفرضيات؛
- التعرف على طريقة اثبات صحة الفرضيات من عدمها.

الفرض العدمي (أو الصفري) The Null Hypothesis

الفرض العدمي هو "الفرض الأساسي المراد اختباره"، ويرمز له عادة بالرمز: H_0

هذا الفرض يأخذ عادة شكل معادلة أو مساواة، فمثلاً إذا كان الفرض العدمي المراد اختباره هو أن متوسط دخل الفرد في إحدى المناطق هو 200 دولار شهرياً فإن هذا الفرض يكتب بالرموز كما يلي :

$$Ho : \mu = 200$$

ويقرأ بالشكل التالي الفرض العدمي هو أن متوسط دخل الفرد في المنطقة هو 200 دولاراً شهرياً.

الفرض البديل: The Alternative Hypothesis

في اختبارات الفروض يتحتم وضع فرض آخر غير الفرض العدمي المراد اختباره يسمى الفرض البديل. وهذا الفرض " هو الذي سيقبل في حالة رفض الفرض العدمي " أي لابد من تحديد فرض آخر بديل في الوقت الذي نحدد فيه الفرض العدمي، وبالتالي فإن الفرض البديل يعرف كما يلي :

"الفرض البديل هو الفرض الآخر الذي سيقبل في حالة رفض الفرض العدمي" ويرمز له عادة بالرمز: H_1

والفرض البديل له أهمية كبيرة وبالذات في قياس الظواهر الاجتماعية – كما سوف نرى – فهو الذي يحدد نوع الاختبار المستخدم لذلك فهو يأخذ أحد أشكال ثلاثة هي :

- أن يأخذ شكل " لا يساوي " وفي هذه الحالة نستخدم ما يسمى : اختبار الطرفين، فمثلاً : إذا كان الفرض العدمي هو أن متوسط الدخل الشهري لفئة معينة في المجتمع هو 200 دولار.

$$Ho: \mu = 200$$

فإن الفرض البديل في هذه الحالة يأخذ الشكل التالي :

$$H1: \mu \neq 200$$

بمعنى أن متوسط دخل هذه الفئة من المجتمع " لا يساوي " 200 دولار شهرياً.

- أن يأخذ شكل " أكبر من " وفي هذه الحالة نستخدم ما يسمى "اختبار الطرف الأيمن "، فمثلاً : قد يكون الفرض البديل كما يلي :

$$H1: \mu > 200$$

أي أن متوسط الدخل لهذه الفئة من المجتمع أكبر من 200 دولار شهرياً.

- قد يأخذ الفرض البديل شكل " أقل من " وفي هذه الحالة نستخدم ما يسمى " اختبار الطرف الأيسر"، فمثلاً: قد يكون الفرض البديل هو:

$$H1: \mu < 200$$

أي أن متوسط الدخل لهذه الفئة من المجتمع أقل من 200 دولاراً شهرياً.

لابد للباحث من تحديد الفرض البديل الذي لا يخرج عن أحد الأشكال الثلاثة السابقة، وهذا التحديد مهم جداً قبل الدخول في تفاصيل الاختبار الإحصائي وذلك لأنه هو الذي يحدد نوع الاختبار المستخدم كما سوف نرى.

الخطأ في اتخاذ القرار :

ففي حالة قبول الباحث لفرضه العدمي، فلا مجال للبحث في الفرض البديل، أما في حالة حدوث العكس بمعنى رفض الفرض العدمي فإنه يتحتم في هذه الحالة قبول الفرض البديل، على أنه من الجدير بالذكر أن الباحث هنا عرضة للوقوع في الخطأ عند اتخاذ قراره بقبول الفرض العدمي أو رفضه، فقد

يرفض فرضاً هو في الواقع صحيح، وقد يقبل فرضا هو في الواقع غير صحيح. لذلك فقد تم تصنيف هذه الأخطاء إلى نوعين هما:

1. **الخطأ من النوع الأول Type I error :** الخطأ من النوع الأول هو "رفض الفرض العدمي بينما هو صحيح ". أي أنه على الرغم من أن الفرض العدمي في الواقع صحيح وكان من الواجب قبوله فقد تم أخذ قرار خاطئ برفضه. وباختصار شديد فإن الخطأ من النوع الأول هو : " رفض فرض صحيح".
2. **الخطأ من النوع الثاني Type II error :** وفي المقابل فإن الخطأ من النوع الثاني يعني " قبول الفرض العدمي بينما هو خاطئ " أي أنه على الرغم من أن الفرض العدمي خاطئ وكان من الواجب رفضه فقد تم أخذ قرار خاطئ بقبوله وباختصار شديد فإن الخطأ من النوع الثاني هو " قبول فرض خاطئ ".

مستوى المعنوية: Level of Significance

يعتبر مصطلح " مستوى المعنوية " واحدا من أهم المصطلحات المستخدمة في دراسة نظرية اختبارات الفروض، والمقصود بمستوى المعنوية هو " احتمال حدوث الخطأ من النوع الأول ". أو نسبة حدوثه " أي احتمال رفض الفرض العدمي بينما هو صحيح ".

وعادة ما يرمز إلى مستوى المعنوية بالرمز اللاتيني ألفا α وأشهر قيمتين لمستوى المعنوية هما 5%، 1 %، ولكن ليس هناك ما يمنع من أن يأخذ قيما أخرى.

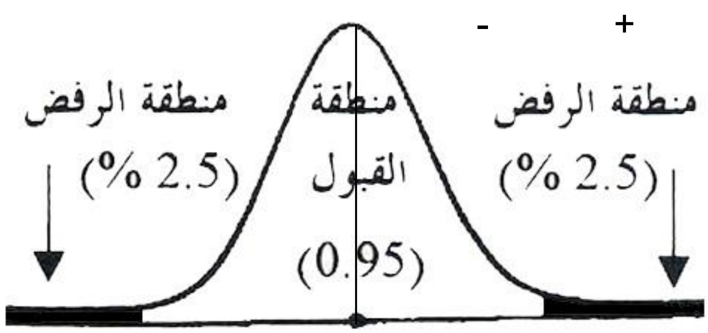
ومن الملاحظات المهمة هنا هو أن " مستوى المعنوية " والذي يسمى أحياناً " مستوى الدلالة " هو المكمل لدرجة الثقة " بمعنى أن مجموعهما يساوي 100% أو واحد صحيح. فإذا كانت درجة الثقة 95% فإن مستوى المعنوية يساوي 5%. والعكس صحيح فإذا كان مستوى المعنوية 5% فإن هذا يعني أن درجة الثقة 95 %. ولعل من أهم الملاحظات هنا هو استخدام تعبير "مستوى المعنوية" في حالات اختبارات الفروض، بينما يستخدم مصطلح "درجة أو مستوى الثقة" في حالات التقدير.

والفكرة الأساسية في اختبار الفرض هي تقسيم المساحة تحت المنحنى إلى منطقتين:

- أحداهما تسمى " **منطقة القبول** " أي منطقة قبول الفرض العدمي؛
- والأخرى تسمى " **منطقة الرفض**"، أي منطقة رفض الفرض العدمي والتي تسمى أحيانا " بالمنطقة الحرجة Critical region ".

والنقطة الجديرة بالملاحظة هنا هي أن منطقة القبول تمثل درجة الثقة، بينما تمثل منطقة الرفض مستوى المعنوية، وهناك ثلاث حالات مختلفة لمنطقتي القبول والرفض هي:

<u>الأولى</u> : إذا كان الفرض البديل يأخذ شكل " لا يساوي " كأن يكون الفرض في هذه الحالة هو أن متوسط دخل الفرد لا يساوي 200 دولاراً فإن منطقة الرفض تكون موزعة على طرفي المنحنى بالتساوي، ويسمى الاختبار في هذه الحالة " اختبار الطرفين "، والذي يأخذ الشكل التالي (بافتراض أن $\alpha = 5\%$) :

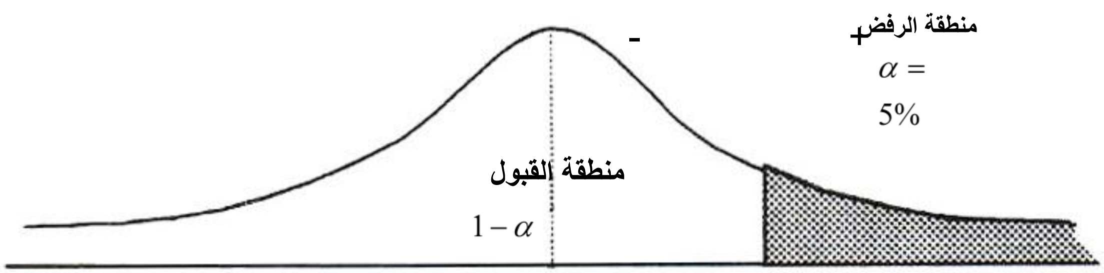


اختبار الطرفين

فالفرض العدمي هنا $Ho: \mu = 200$ يعني أن متوسط دخل الفرد يساوي 200 دولار شهريا، والفرض البديل في هذه الحالة هو $HI: \mu \neq 200$ بمعنى أن متوسط دخل الفرد لا يساوي 200 دولار شهرياً، حيث تمثل المنطقة البيضاء غير المظللة منطقة القبول والتي قد تساوي 95% وبالتالي فمنطقة الرفض مقسمة بالتساوي على طرفي المنحنى في هذه الحالة تكون قيمة كل منهما % 2.5.

والنتيجة هو أن القرار أيا كان نوعه سيكون بمستوى معنوية % 5 بمعنى أن احتمال أو نسبة الخطأ فيه من النوع الأول تساوي % 5.

<u>الثانية</u>: إذا كان الفرض البديل يأخذ شكل "أكبر من" فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيمن للمنحنى، ويسمى الاختبار في هذه الحالة اختبار الطرف الأيمن، والذي يأخذ الشكل التالي أدناه:



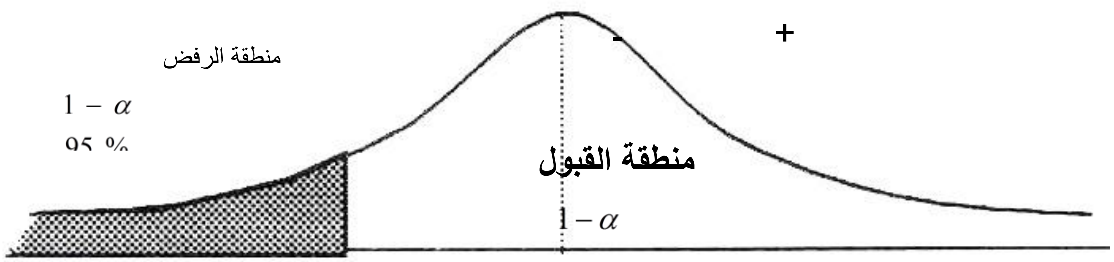
0

اختبار الطرف الأيمن

فالفرض العدمي هنا نفس فرض المثال السابق، بينما الفرض البديل هو $HI : \mu > 200$

بمعنى أن متوسط دخل الفرد أكبر من 200 دولاراً شهرياً، وبالتالي فإن مستوى المعنوية والذي يساوي مثلاً 5% مركز في الطرف الأيمن من المنحنى.

<u>الثالثة:</u> إذا كان الفرض البديل يأخذ شكل " أقل من " فإن منطقة الرفض تكون مركزة بالكامل في الطرف الأيسر للمنحنى، ويسمى الاختبار في هذه الحالة اختبار الطرف الأيسر، والشكل التالي يوضح ذلك:



0

اختبار الطرف الأيسر

مع افتراض ثبات الفرض العدمي كما في المثال السابق، بينما الفرض البديل هو $HI : \mu < 200$

بمعنى أن متوسط دخل الفرد أقل من 200 دولار شهرياً، وبالتالي فإن مستوى المعنوية والذي يساوي مثلاً 5 % مركز في الطرف الأيسر من المنحنى.

خطوات الاختبار الإحصائي:

يمكن تلخيص خطوات الاختبار الإحصائي في <u>خمس</u> خطوات كما يلي :

1) **وضع الفرض العدمي H_o**، والذي يأخذ عادة شكل " يساوي " فمثلاً إذا كان المطلوب هو اختبار ما إذا كان متوسط عمر الناخب هو 20 سنة فإن هذا الفرض يصاغ كما يلي :

$$Ho : \mu = 20$$

2) **وضع الفرض البديل H1**، والذي يأخذ أحد أشكال ثلاثة إما:

" لا يساوي "

أو " أكبر من "

أو " أقل من "

وبالرموز فإن الفرض البديل قد يأخذ شكل أحد الصيغ التالية:

$$H1: \mu \neq 20$$

$$OR \mu > 20$$

$$OR \mu < 20$$

والذي يحدد شكل الفرض البديل هو مدى اقتناع الباحث بذلك أو مدى توفر المعلومات الأولية، فمثلاً إذا كانت وجهة نظر الباحث أن متوسط عمر الناخب لا يمكن أن يقل عن 20 سنة فإنه يختار الفرض البديل " أكبر من " والعكس صحيح إذا كان يعتقد أن متوسط عمر الناخب لا يزيد عن 20 سنة فإنه يختار الفرض البديل " أقل من " أما إذا لم يكن لديه أي تصور أو أي معلومات فإنه يختار الفرض البديل " لا يساوي ".

3) **إحصائية الاختبار :** وهي الإحصائية التي يتم حسابها من بيانات العينة بافتراض أن الفرض العدمي صحيح، ويتوقف شكل الإحصائية على العوامل التالية :

- توزيع المجتمع: وهل هو طبيعي أم لا، وهل تباينه معروف أم لا.
- حجم العينة: هل هو كبير أم صغير.
- والفرض العدمي المراد اختباره وهل هو عن الوسط أو النسبة أو التباين أو الارتباط... الخ.

الإحصائية في حالة اختبار الوسط الحسابي :

أ) بافتراض أن المجتمع الإحصائي المسحوبة منه العينة هو مجتمع طبيعي وانحرافه المعياري σ معروف، (أو) أن العينة كبيرة بدرجة كافية فإن إحصائية الاختبار والتي نرمز لها بالرمز $Z_{\bar{X}}$ تأخذ الشكل التالي :

الإحصائية في حالة اختبار الوسط للعينات الكبيرة

$$Z_{\bar{X}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

لاحظ أن البسط هو الفرق بين متوسطي المجتمع والعينة، والمقام هو الخطأ المعياري للوسط. ومن الناحية العملية فإن الانحراف المعياري للمجتمع عادة ما يكون غير معروف ولكن طالما أن العينة كبيرة بدرجة كافية فإنه يمكن استخدام الانحراف المعياري للعينة S بدلا من الانحراف المعياري للمجتمع σ.

ب) أما في حالة العينات الصغيرة وذلك عندما يكون المجتمع طبيعياً وانحرافه المعياري غير معروف فإن الإحصائية تأخذ الشكل التالي :

الإحصائية في حالة اختبار الوسط للعينات الصغيرة

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

والتي لها توزيع t بدرجات حرية n - 1

الإحصائية في حالة اختبار النسبة :

إذا كانت العينة كبيرة فإن إحصائية الاختبار تأخذ الشكل التالي:

الإحصائية في حالة اختبار النسبة

$$Z_{\hat{P}} = \frac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}}$$

والتي لها توزيع طبيعي معياري حيث $\hat{P}$ هي النسبة للعينة، P هي النسبة للمجتمع.

لاحظ أن البسط هو الفرق بين نسبتي المجتمع والعينة والمقام هو الخطأ المعياري للنسبة.

4) الخطوة الرابعة: في الاختبار هي تحديد منطقتي القبول والرفض وذلك بناءً على الجداول الإحصائية والتي تعتمد على:

أ- توزيع المعاينة (وهل هو طبيعي أو t أو...)

ب- والفرض البديل (وهل هو لا يساوي أو أكبر من أو أقل من أي هل يستخدم اختبار الطرفين أو الطرف الأيمن أو الأيسر).

ج- ومستوى المعنوية (وهل هو 1% أو 5% أو غير ذلك).

5) المقارنة والقرار: بمعنى أن نقارن قيمة الإحصائية (المحسوبة من الخطوة الثالثة) بحدود منطقتي القبول والرفض (والتي حددناها في الخطوة الرابعة)، فإذا وقعت قيمة الإحصائية داخل منطقة القبول فإن القرار هو: قبول الفرض العدمي، أما إذا وقعت قيمة الإحصائية في منطقة الرفض فإن القرار هو رفض الفرض العدمي، وفي هذه الحالة نقبل الفرض البديل، مع ملاحظة أن القرار مرتبط بمستوى المعنوية المحدد، بمعنى أن القرار قد يتغير إذا تغير مستوى المعنوية المستخدم (وفي بعض الحالات قد لا يتغير القرار، فهذا يتوقف على قيمة الإحصائية وما إذا كانت تقع في منطقة القبول أو منطقة الرفض).

مما سبق يمكن تلخيص خطوات الاختبار الإحصائي فيما يلي:

1- الفرض العدمي.

2- الفرض البديل.

3- الإحصائية.

4- حدود منطقتي القبول والرفض.

5- المقارنة والقرار.

<u>مثال:</u> عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 دولاراً، كيف يمكن اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72 دولاراً مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية 5 % إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 دولاراً.

<u>الحل</u> :

1- الفرض العدمي : هو أن متوسط المجتمع يساوي 72 وبالرموز :

$$Ho : \mu = 72$$

2- الفرض البديل : هو أن المتوسط لا يساوي 72 وبالرموز :

$$HI : \mu \neq 72$$

3- الإحصائية : بما أن العينة كبيرة فإن الإحصائية في حالة اختبار الوسط تأخذ الشكل التالي :

$$Z_{\overline{X}} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

حيث $n = 49, \sigma = 14, \overline{X} = 75, \mu = 72$

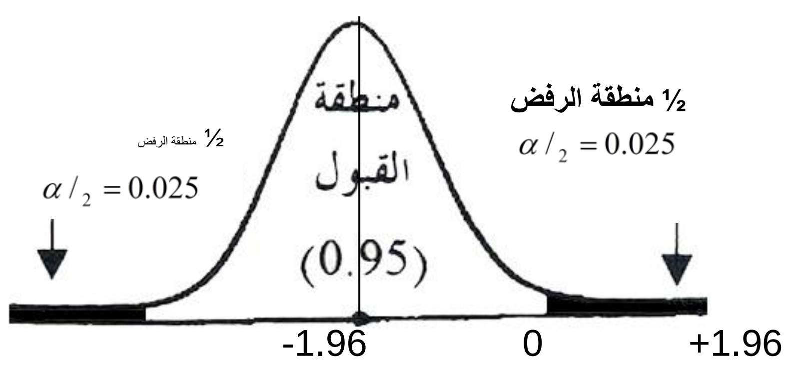
وبالتعويض نحصل على :

$$Z_{\overline{X}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}}$$

$$Z_{\overline{X}} = \frac{3}{\frac{4}{7}} = \frac{3}{2} = 1.5$$

أي أن قيمة الإحصائية تساوي 1.5

4- حدود منطقتي القبول والرفض : نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% وبما أن الفرض البديل هو :" لا يساوي " فإن ما يستخدم في هذه الحالة هو اختبار الطرفين كما في الشكل التالي :



حدود منطقتي القبول والرفض وذلك بقسمة درجة الثقة (المكملة لمستوى المعنوية) والتي تساوي 0.95 على 2 فنحصل على 0.4750 وبالكشف في جدول التوزيع الطبيعي المعياري عن Z التي تقابل المساحة 0.4750 نجد أنها تساوي 1.96 وحيث أنها موزعة على طرفي المنحنى بالتساوي فنضع إشارة

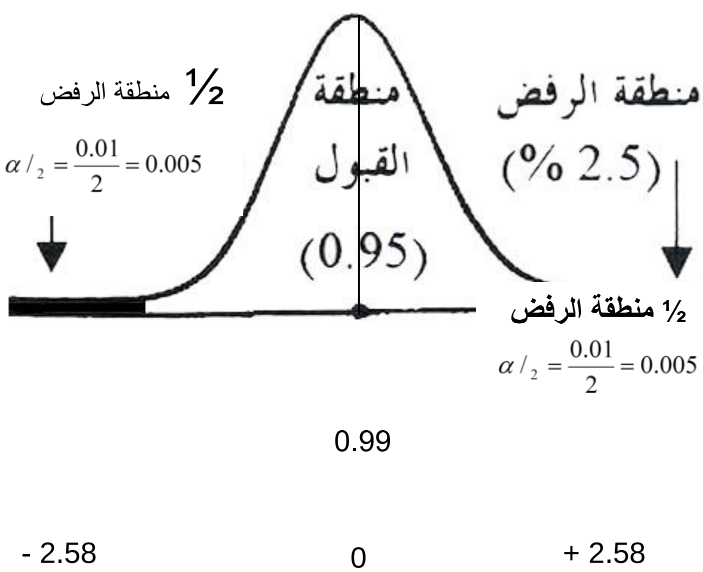
موجبة في النصف الأيمن، وإشارة سالبة في النصف الأيسر، أي أن منطقة القبول تبدأ من القيمة 1.96- وتستمر حتى القيمة 1.96 + (أي أن أي قيمة محصورة بين هاتين القيمتين تكون في منطقة القبول، وأي قيمة خارج هذه الحدود تكون في منطقة الرفض).

5- المقارنة والقرار: وبمقارنة قيمة الإحصائية المحسوبة من الخطوة رقم 3 (والتي تساوي 1.5) بحدود منطقتي القبول والرفض (من الخطوة رقم 4) نجد أنها تقع في منطقة القبول لذلك فإن القرار هو :

قبول الفرض العدمي بأن متوسط دخول الأفراد الأسبوعية في هذه الدولة يساوي 72 دولاراً وذلك بمستوى معنوية % 5.

ملاحظة :

لو استخدمنا مستوى معنوية 1% بدلاً من 5% كما في المثال أعلاه فإن حدود منطقتي القبول والرفض تصبح كما يلي :



بمقارنة قيمة الإحصائية 1.5 بحدود منطقتي القبول والرفض نجد أنها تقع في منطقة القبول أي أن القرار هو نفسه قبول الفرض العدمي ولن يتغير بل يتأكد باستخدام مستوى معنوية 1%.

مثال: يدّعِي أحد المرشحين في الانتخابات أنه سيحصل على نسبة 70% من أصوات الناخبين عندما تجري الانتخابات، ولاختبار هذا الادعاء تم اختيار عينة عشوائية من الناخبين حجمها 100 ناخب، ووجد أن نسبة من يؤيدون المرشح في العينة هي % 60 اختبر مدى صحة ادعاء المرشح بأن النسبة في المجتمع هي % 70 مقابل الفرض البديل أن النسبة أقل من 70% وذلك بمستوى معنوية % 5.

الحل :

1- الفرض العدمي هو أن النسبة في المجتمع (نسبة من يؤيدون المرشح في المجتمع) هي 0.70 أي أن الفرض العدمي هو أن الادعاء صحيح وأن المرشح سيحصل على النسبة التي ادعاها وهي 70 % بالرموز HO : P = 0.70

2- الفرض البديل والمنطقي: في هذه الحالة هو أن النسبة في المجتمع أقل من هذا الادعاء وبالرموز :

$$HI : P < 0.70$$

3- الإحصائية : وتأخذ الإحصائية في حالة اختبار النسبة الشكل التالي:

$$Z_{\hat{P}} = \frac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}}$$

حيث $n = 100, \hat{P} = 0.60, P = 0.70، 1 - p = 1 - 0.70 = 0.30$

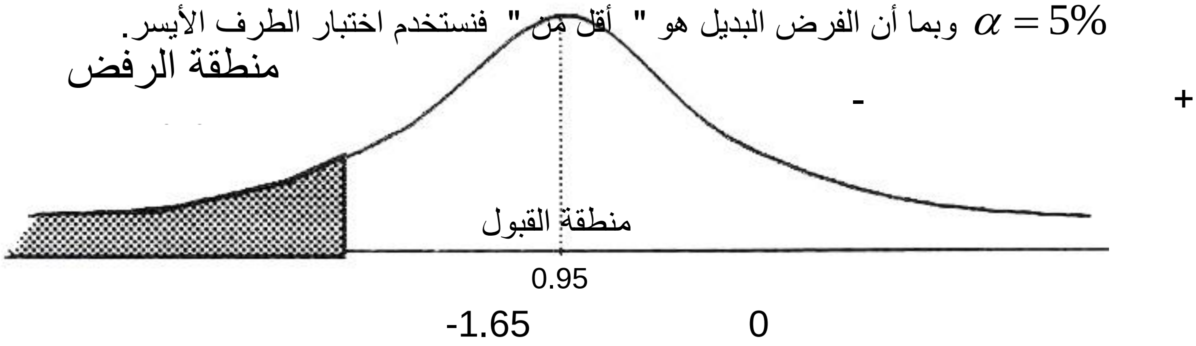
$$Z_{\hat{P}} = \frac{0.60 - 0.70}{\sqrt{\frac{0.70 \times 0.30}{100}}}$$

$$= \frac{-0.10}{0.046}$$

$$Z_{\hat{P}} = -2.17$$

أي أن قيمة الإحصائية تساوي 2.17 -

4- حدود منطقتي القبول والرفض نحصل عليها من التوزيع الطبيعي المعياري، حيث مستوى المعنوية $\alpha = 5\%$ وبما أن الفرض البديل هو " أقل من " فنستخدم اختبار الطرف الأيسر.



أي أن منطقة القبول تشمل النصف الموجب (اليمين) من المنحنى وحتى القيمة السالبة 1.65- وبالتالي فإن منطقة الرفض تشمل القيم التي أقل من 1.65- وقد حصلنا على هذا الرقم من جدول Z حيث تتركز منطقة الرفض والتي تساوي 0.05 في الطرف الأيسر للمنحنى. فنقوم بطرح هذه المنطقة (أو المساحة) من (نصف مساحة المنحنى) فنحصل على ما يلي :

$$0.5 - 0.05 = 0.4500$$

ونكشف في جدول التوزيع الطبيعي عن Z التي تقابل المساحة 0.4500 مع ملاحظة مهمة جداً وهي أن منطقة الرفض تقع في الطرف الأيسر أي السالب للمنحنى، لذلك لابد من وضع إشارة سالبة لقيمة Z التي نحصل عليها.

المقارنة والقرار : وبمقارنة قيمة الإحصائية التي حصلنا عليها في الخطوة رقم (3) التي تساوي 2.17 - بحدود منطقتي القبول والرفض (من الخطوة رقم 4) نجد أن قيمة الإحصائية تقع في منطقة الرفض لأن - 2.17 أصغر من 1.65 - فإن القرار هو :

رفض الفرض العدمي بادعاء المرشح بأن نسبة مؤيديه في المجتمع هي % 70 وقبول الفرض البديل بأن النسبة أقل من % 70 وذلك بمستوى معنوية % 5 (أي أن احتمال الخطأ في هذا القرار لا يتعدى % 5).

اختبار الفرق بين وسطين حسابيين :

قد يرغب الباحث في إجراء اختبار عما إذا كان متوسط الدخل في إحدى الدول يساوي متوسط الدخل في دولة أخرى، أو إجراء اختبار عما إذا كان متوسط عمر الناخب في إحدى المناطق يساوي متوسط عمر الناخب في منطقة أخرى... وهكذا بمعنى آخر قد يرغب الباحث في إجراء اختبار عما إذا كان متوسط المجتمع الأول يساوي متوسط المجتمع الثاني.

في مثل هذه الحالات يسمى الاختبار اختبار الفرق بين وسطين حسابيين، وتكون خطوات هذا الاختبار في حالة العينات الكبيرة كما يلي :

1- **الفرض العدمي:** أن متوسط المجتمع الأول يساوي متوسط المجتمع الثاني (أي لا يوجد فرق بين متوسطي المجتمعين).

وبالرموز : $Ho: \mu_1 = \mu_2$

2- الفرض البديل : أن المتوسطين غير متساويين وبالرموز :

$$HI : \mu_1 \neq \mu_2$$

ويمكن للباحث استخدام أكبر من أو أقل من بدلاً من لا يساوي إذا كان لديه معلومات تشير إلى ضرورة ذلك.

3- الإحصائية: وبافتراض أن المجتمعين طبيعيان وأن العينتين مستقلتان وكبيرتان فإن إحصائية الاختبار في هذه الحالة تأخذ الشكل التالي:

$$Z_{\bar{X}1-\bar{X}2} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

حيث : يرمز بـ n1 إلى حجم العينة الأولى.

يرمز بـ n2 إلى حجم العينة الثانية.

يرمز بـ $\bar{X}1$ إلى الوسط الحسابي للعينة الأولى.

يرمز بـ $\bar{X}2$ إلى الوسط الحسابي للعينة الثانية.

يرمز بـ σ_1^2 إلى تباين المجتمع الأول.

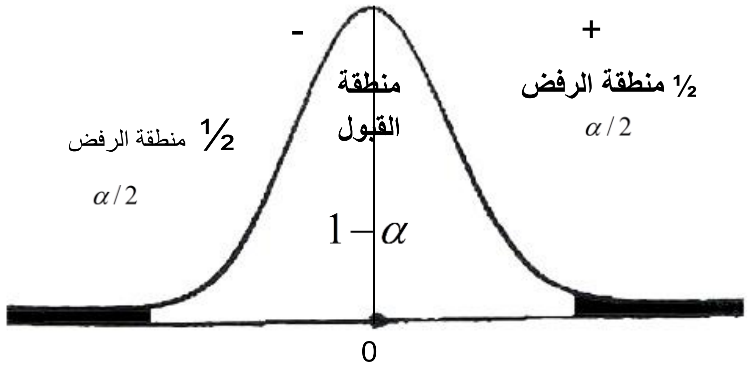
يرمز σ_2^2 إلى تباين المجتمع الثاني.

4- حدود منطقتي القبول والرفض ويمثلها الشكل التالي مع ملاحظة أن :

أ- التوزيع طبيعي (نحصل على القيم من توزيع Z).

ب- اختبار الطرفين (لأن الفرض البديل لا يساوي).

جـ- مستوى المعنوية يساوي α



5- المقارنة والقرار نقارن قيمة الإحصائية بحدود منطقتي القبول والرفض، فإذا وقعت في منطقة القبول نقبل الفرض العدمي، وإذا وقعت في منطقة الرفض نرفض الفرض العدمي، ونقبل الفرض البديل.

<u>مثال</u>: البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما:

حيث $\bar{X}_1 = 35, \bar{X}_2 = 29, n_2 = 80, n_1 = 100$

اختبر الفرض العدمي : أن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5% مقابل الفرض البديل أنهما غير متساويين إذا علمت أن :

$$\sigma_1^2 = 60, \sigma_2^2 = 32$$

الحل :

1- الفرض العدمي أن المتوسطين متساويان وبالرموز :

$$Ho: \mu_1 = \mu_2$$

2- الفرض البديل أن المتوسطين غير متساويين وبالرموز :

$$HI : \mu_1 \neq \mu_2$$

3- الإحصائية : تأخذ الشكل التالي :

$$Z_{\bar{X}1-\bar{X}2} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

وبالتعويض عن:

$$n_1 = 100, n_2 = 80, \bar{X}_1 = 35, \bar{X}_2 = 29, \sigma_1^2 = 60, \sigma_2^2 = 32$$

نحصل على :

$$Z_{\bar{X}1-\bar{X}2} = \frac{35-29}{\sqrt{\frac{60}{100} + \frac{32}{80}}}$$

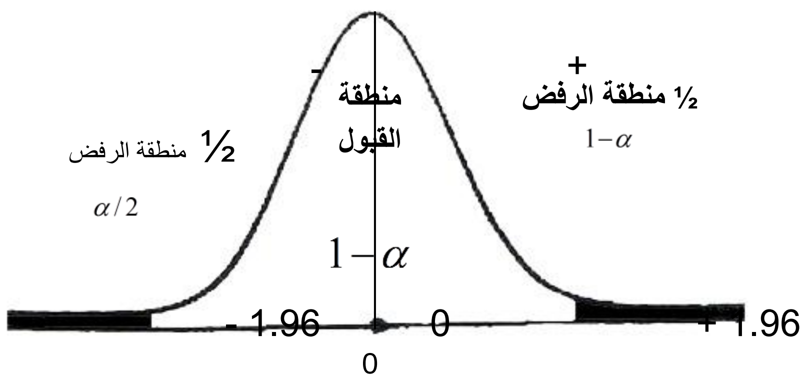
$$= \frac{60}{\sqrt{0.60+0.40}}$$

$$= \frac{6}{\sqrt{1}} = 6$$

أي أن قيمة الإحصائية تساوي 6

4- حدود منطقتي القبول والرفض التي نحصل عليها من جدول التوزيع الطبيعي Z لأن العينات كبيرة، والاختبار هو اختبار الطرفين (لأن الفرض البديل لا يساوي)

ومستوى المعنوية المطلوب هو % 5.



أي أن منطقة القبول تبدأ من 1.96- إلى 1.96+ ومنطقة الرفض هي القيم التي أصغر من 1.96 - والتي أكبر من 1.96 +.

5- المقارنة والقرار ولما كانت قيمة الإحصائية (والتي تساوي) 6 تقع في منطقة الرفض فإن القرار هو **رفض الفرض العدمي وقبول الفرض البديل** بمستوى معنوية 5% أي أننا نرفض الفرض القائل بأن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية وذلك بمستوى معنوية 5 %.

اختبار الفرق بين وسطين في حالة العينات الصغيرة :

إذا كانت العينات صغيرة (مجموع العينتين أقل من 30 مفردة أو حتى 31 مفردة) فإن الإحصائية في هذه الحالات بافتراض أن المجتمعين طبيعيان، وأن تباين المجتمع الأول يساوي تباين المجتمع الثاني ولكنه مجهول (بمعنى أن $\sigma_1^2 = \sigma_2^2$ ولكن قيمة هذا التباين غير معروفة) وأن العينتين مستقلتان فإن إحصائية الاختبار تأخذ الشكل التالي والتي لها توزيع t بدرجات حرية تساوي $n_1 + n_2 - 2$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}}}$$

حيث :

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n1 + n_2 - 2}$$

أي يتم حساب S^2 أولاً قبل التعويض في الإحصائية وتكون خطوات الاختبار هي :

هي :

1- الفرض العدمي (هو نفسه)

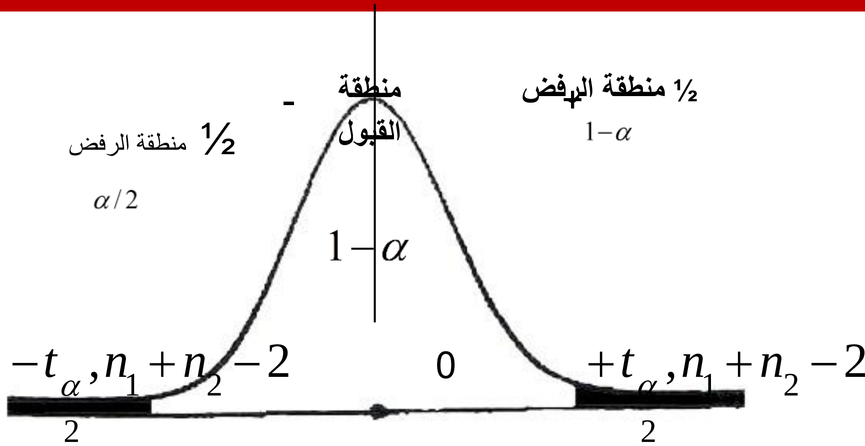
$$HI : \mu_1 = \mu_2$$

2- الفرض البديل : (هو نفسه) :

$$Ho : \mu_1 \neq \mu_2$$

3- الإحصائية هي المكتوبة أعلاه (وهي في هذه الحالة t وليست Z)

4- حدود منطقتي القبول والرفض ونحصل عليها في هذه الحالة من جدول t عند درجات حرية تساوي $n1 + n_2 - 2$ وعند مستوى معنوية يساوي $\frac{\sigma}{2}$ كما في الشكل التالي:



5- المقارنة والقرار : كما سبق :

أما إذا فرضنا أن تباين المجتمعين غير متساويين، فإن الإحصائية في هذه الحالة تأخذ الشكل التالي :

$$t=\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

مثال: البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من مدينتين عن أعمار الناخبين بهما (بافتراض أن تباينهما هو نفسه) :

$$n1=10, n_2=10, \bar{X}1=28\bar{X}2=26, S_1^2=50, S_2^2=30$$

اختبر الفرض العدمي : $Ho: \mu_1=\mu_2$ مقابل الفرض البديل $HI: \mu_1\neq\mu_2$

وذلك بمستوى معنوية 5 % بافتراض أن الأعمار في المدينتين لهما توزيع طبيعي.

<u>**الحل**</u> :

1- الفرض العدمي : $Ho: \mu_1=\mu_2$

أي متوسط أعمار الناخبين في المدينتين متساوٍ

2- الفرض البديل : $HI: \mu_1\neq\mu_2$

أي متوسط أعمار الناخبين في المدينتين غير متساوٍ

3- الإحصائية لاحظ (أن العينات صغيرة، وأن تباين المجتمعين هو نفسه، وأن المجتمعين طبيعيان). فإن الإحصائية المناسبة في هذه الحالة هي t :

$$t=\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S^2}{n_1}+\frac{S^2}{n_2}}}$$

حيث

$$S^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n1+n_2-2}$$

نحسب أولاً S^2 كما يلي :

$$S^2=\frac{(10-1)\times 50+(10-1)\times 30}{10+10-2}$$

$$=\frac{9\times 50+9\times 30}{18}$$

$$=\frac{450+270}{18}$$

$$=\frac{720}{18}$$

$$S^2 = 40$$

وبالتعويض في الإحصائية عن :

$$\bar{X}_1=28,\ \bar{X}_2=26,\ S^2=40,\ n1=10,\ n_2=10$$

نحصل على :

$$t=\frac{28-26}{\sqrt{\frac{40}{10}+\frac{40}{10}}}$$

$$= \frac{2}{\sqrt{8}} = \frac{2}{2.828} = 0.7$$

أي أن قيمة الإحصائية تساوي 0.7

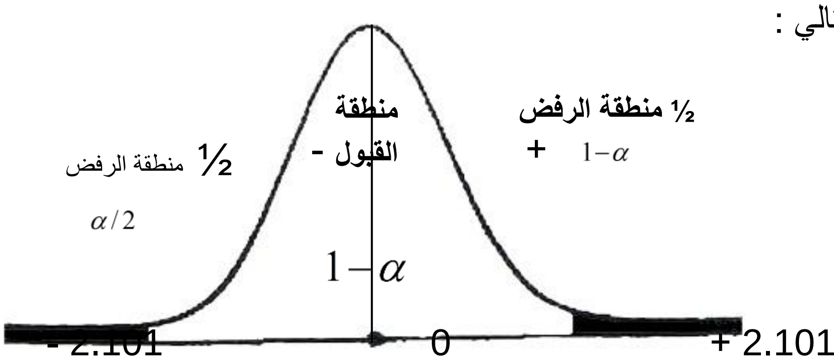
4- حدود منطقتي القبول والرفض :

ونحصل عليها في هذه الحالة من جدول t عند درجات حرية تساوي $n_1 + n_2 - 2$

أي تساوي 2 - 10 + 10 والتي تساوي 18 وذلك عند مستوى معنوية يساوي

$\alpha = 0.05$ أي أن نصف مستوى المعنوية $0.025 = \frac{0.05}{2}$ أي أن $t_{0.025,18} = 2.101$

كما في الشكل التالي :



أي أن منطقة القبول تبدأ من 2.101- وحتى 2.101 +

قيمة الإحصائية تساوي 0.7 فإنها تقع في منطقة القبول وبالتالي فإن القرار هو **قبول الفرض العدمي** بأن متوسط أعمار الناخبين في المدينة الأولى يساوي متوسط أعمار الناخبين في المدينة الثانية وذلك بمستوى معنوي 5.% (حل المثال السابق بافتراض أن تباين المجتمعين غير متساويين).

اختبار الفرق بين نسبتين:

كذلك قد يرغب الباحث في اختبار ما إذا كانت نسبة المؤيدين لمرشح ما في الانتخابات التشريعية تساوي نسبة المؤيدين لمرشح آخر في الانتخابات نفسها، في مثل هذه الحالات فإن المطلوب هو اختبار ما إذا كانت النسبة في المجتمع الأول تساوي النسبة في المجتمع الثاني، ويسمى الاختبار : اختبار الفرق بين نسبتين

وتكون خطوات هذا الاختبار ما يلي :

1- الفرض العدمي : هو أن النسبة في المجتمعين متساوية وبالرموز :

$$Ho: \mathrm{P}_1 = \mathrm{P}_2$$

2- الفرض البديل : هو أن النسبتين في المجتمعين غير متساوية وبالرموز

$$H1: p_1 \neq p_2$$

(ويمكن اختيار شكل آخر للفرض البديل مثل: أكبر من أو أقل إذا دعت الحاجة لذلك).

3- الإحصائية : بافتراض أن العينتين كبيرتان بدرجة كافية تكون الإحصائية كما يلي

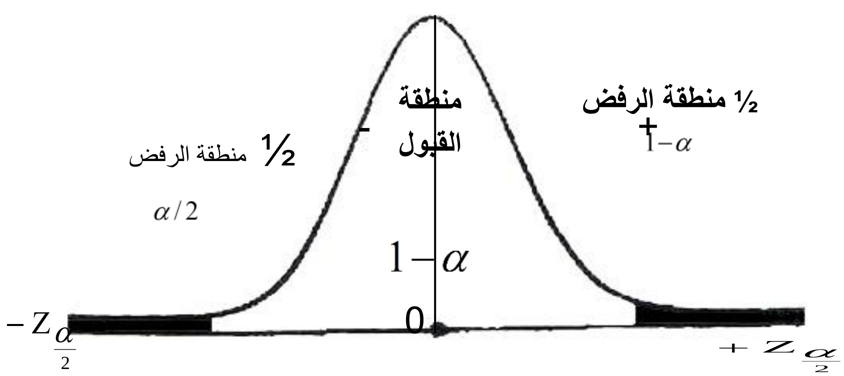
$$Z_{\hat{P}_1-\hat{P}_2} = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n1} + \frac{\hat{P}(1-\hat{P})}{n2}}}$$

حيث :

$$\hat{P} = \frac{n_1\hat{P}_1 + n_2\hat{P}_2}{n_1 + n_2}$$

أي يتم أولاً حساب $\hat{P}$ (والتي تمثل متوسط مرجح من نسبتي العينتين) قبل التعويض في الإحصائية والتي لها توزيع طبيعي معياري.

4- حدود منطقتي القبول والرفض ونحصل عليها من جدول التوزيع الطبيعي، والاختبار هنا هو اختبار الطرفين (لأن الفرض البديل لا يساوي) وتحدد المنطقتين بناءً على مستوى المعنوية المطلوب، وذلك كما في الشكل التالي :



مثال: لاختبار ما إذا كانت نسبة المؤيدين لبرنامج اقتصادي معين في المدينة (أ) يساوي نسبة المؤيدين لهذا البرنامج في المدينة (ب) تم اختيار عينتين عشوائيتين مستقلتين من المدينتين حيث : حجم العينة الأولى يساوي حجم العينة الثانية يساوي 100 وكانت نسبة المؤيدين للبرنامج في عينة المدينة (أ) هي : $\hat{P}=0.70$ ونسبة المؤيدين للبرنامج في عينة المدينة ب هي $\hat{P}_2=0.50$.

اختبر الفرض العدمي أن النسبة في المدينتين متساوية مقابل الفرض البديل أنها غير متساوية وذلك بمستوى معنوية 1 %.

الحل :

1- الفرض العدمي :

النسبة في المدينة أ تساوي النسبة في المدينة ب وبالرموز :

$$Ho:\mathrm{P}_1=\mathrm{P}_2$$

2- الفرض البديل : النسبة في المدينتين غير متساوية وبالرموز

$$H1:\mathrm{P}_1\neq\mathrm{P}_2$$

3- الإحصائية :

$$Z_{\hat{P}_1-\hat{P}_2}=\frac{\hat{P}_1-\hat{P}_2}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n1}+\frac{\hat{P}(-\hat{P})}{n2}}}$$

حيث

$$\hat{P}=\frac{n_1\hat{P}_1+n_2\hat{P}_2}{n_1+n_2}$$

وبالتعويض عن : $n_1=100, n_2=100, \hat{P}_1=0.70, \hat{P}_2=0.50$

نحصل على :

$$\hat{p}=\frac{100\times0.70+100\times0.50}{100+100}$$

$$=\frac{70+50}{200}$$

$$=\frac{120}{200}$$

$$\hat{p}=0.60$$

وبالتعويض في الإحصائية نحصل على:

$$z=\frac{0.70-0.50}{\sqrt{\frac{0.60\times0.40}{100}+\frac{0.60\times0.40}{100}}}$$

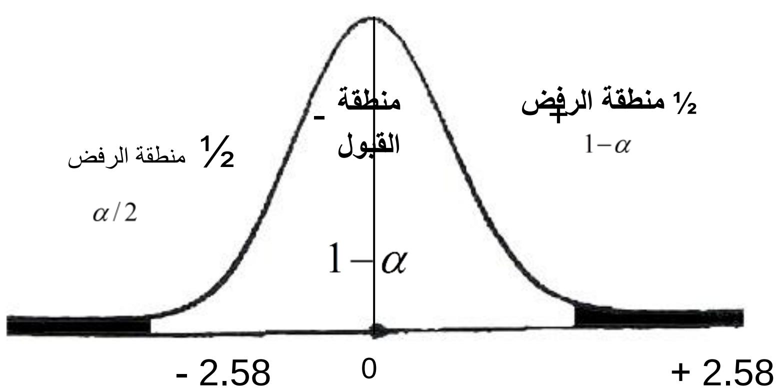
$$=\frac{0.20}{0.069}$$

$$=2.899$$

أي أن قيمة الإحصائية تساوي 2.899

4- حدود منطقتي القبول والرفض

نحصل عليها من التوزيع الطبيعي، واختبار الطرفين بمستوى معنوية 1 % كما في الشكل التالي :



أي أن منطقة القبول تبدأ من 2.58- وحتى 2.58 +

5- المقارنة والقرار :

وحيث أن قيمة الإحصائية تساوي 2.899 فهي تقع في منطقة الرفض وبالتالي فإن القرار هو :

رفض الفرض العدمي وقبول الفرض البديل أي رفض الفرض القائل بأن نسبة المؤيدين للبرنامج الاقتصادي في المدينة (أ) تساوي نسبة المؤيدين له في المدينة (ب) وذلك بمستوى معنوية 1% (بمعنى أن احتمال الخطأ في هذا القرار لا تتعدى 1%). وقبول الفرض البديل بأن النسبتين غير متساويتين.