المحور الرابع: اختبار الفرضيات الاحصائية:

المقصود بالفروض هنا الفروض الإحصائية 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.

فمثلاً: قد يفترض الباحث أن متوسط الدخل الشهري للفرد هو 20000 دينار جزائري (بناءً على ما يراه من مستوى المعيشة في البلد وأوضاعه الاقتصادية)، ويحتاج إلى اختبار علمي (إحصائي) لمعرفة مدى صحة هذا الفرض أو قد يفترض باحث آخر أن نسبة الناخبين في إحدى الدوائر الذين يؤيدون مرشحاً معيناً لا تقل عن 30 % وهكذا... والمطلوب هو اختبار مدى صحة هذه الفروض. أي أن يصل الباحث إلى قرار إما بقبول الفرض أو عدم قبوله (أي رفضه) وذلك باحتمال معين. وقبل تناول كيفية إجراء الاختبارات الإحصائية نستعرض أولاً بعض المفاهيم والتعريفات الأساسية اللازمة لهذا الموضوع حتى تكون الصورة أكثر وضوحاً..

4-1- مفاهيم أساسية:

4-1-1- الفرض العدم (أو الصفري)

فرضية العدم هي "الفرضية الأساسي المراد اختبارها". ويرمز لها عادة بالرمز H_0: . وهي فرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، فمثلاً إذا كان الفرض العدم المراد اختباره هو أن متوسط دخل الفرد هو 20000 دينار جزائري شهرياً فإن هذا الفرض يكتب بالرموز كما يلي :

$$H_0 : \mu = 20000$$

ويقرأ بالشكل التالي :

الفرض العدم هو : أن متوسط دخل الفرد هو 20000 دينار جزائري شهرياً.

وليس شرطاً أن يصاغ الفرض العدم بالرموز، فقد يتم التعبير عنه بدون رموز. فقد يريد الباحث أن يختبر ما إذا كانت هناك علاقة بين المؤهل العلمي ودرجة الوعي السياسي. فقد يصيغ الباحث الفرض العدم بالشكل التالي (على سبيل المثال) :

الأمية والاستعداد للانحراف مستقلان (أي لا توجد علاقة بينهما، أو أن العلاقة بينهما منعدمة).

4-1-2- الفرض البديل:

في اختبارات الفروض يتحتم وضع فرض آخر غير الفرض العدم المراد اختباره يسمى الفرض البديل. وهذا الفرض " هو الذي سيقبل في حالة رفض الفرض العدم " أي لابد من تحديد فرض آخر بديل في الوقت الذي نحدد فيه الفرض العدم، وبالتالي فإن الفرض البديل يعرف كما يلي: "الفرض البديل هو الفرض الآخر الذي سيقبل في حالة رفض الفرض العدم" ويرمز له عادة بالرمز : H_1

والفرض البديل له أهمية كبيرة وبالذات في قياس الظواهر الاقتصادية والاجتماعية، فهو الذي يحدد نوع الاختبار المستخدم لذلك فهو يأخذ أحد الأشكال الثلاثة هي:

أ- أن يأخذ شكل " لا يساوي ". وفي هذه الحالة نستخدم ما يسمى: اختبار ثنائي الاتجاه (اختبار الطرفين).

مثال1: إذا كان الفرض العدم هو أن متوسط الدخل الشهري لفئة معينة في المجتمع هو 20000 دينار جزائري.

$$H_0 : \mu = 20000$$

فإن الفرض البديل في هذه الحالة يأخذ الشكل التالي :

$$H_1 : \mu \neq 20000$$

بمعنى أن متوسط دخل هذه الفئة من المجتمع " لا يساوي " 20000 دينار جزائري. شهرياً.

(أ)ب- أو أن يأخذ شكل " أكبر من ". وفي هذه الحالة نستخدم ما يسمى الاختبار أحادي الاتجاه من اليمين (اختبار الطرف الأيمن).

مثال2: قد يكون الفرض البديل كما يلي :

$$H_1 : \mu > 20000$$

أي أن متوسط الدخل لهذه الفئة من المجتمع أكبر من 20000 دينار جزائري.شهرياً.

ج- وأخيراً قد يأخذ الفرض البديل شكل " أقل من ".وفي هذه الحالة نستخدم ما يسمى الاختبار أحادي الاتجاه من اليسار(اختبار الطرف الأيسر).

مثال3: قد يكون الفرض البديل هو :

$$H_1 : \mu < 20000$$

أي أن متوسط الدخل لهذه الفئة من المجتمع أقل من 20000 دينار جزائري.شهرياً.

و**الخلاصة:** هي 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.

4-1-3- الخطأ في اتخاذ القرار : ففي حالة قبول الباحث لفرضه العدم، فلا مجال للبحث في الفرض البديل، أما في حالة حدوث العكس بمعنى رفض الفرض العدم 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 العدم 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 :

4-1-3-1- الخطأ من النوع الأول:_الخطأ من النوع الأول هو "رفض الفرض العدم بينما هو صحيح ". أي أنه على الرغم من أن الفرض العدم في الواقع صحيح وكان من الواجب قبوله فقد تم أخذ قرار خاطئ برفضه. وباختصار شديد فإن الخطأ من النوع الأول هو : " رفض فرض صحيح". واحتماله α ، ويكتب : $p(RH_0 / H_0) = \alpha$

4-1-3-2- الخطأ من النوع الثاني:_وفي المقابل فإن الخطأ من النوع الثاني يعني " قبول الفرض العدم بينما هو خاطئ أي أنه على الرغم من أن الفرض العدم خاطئ وكان من الواجب رفضه فقد تم أخذ قرار خاطئ بقبوله وباختصار شديد فإن الخطأ من النوع الثاني هو " قبول فرض خاطئ ".

واحتماله 1- α ويكتب : $p(\overline{R}H_0 / H_1) = 1 - \alpha$

جدول: القرارات الخاطئة والصائبة في عملية اختبار الفرضيات

نتيجة الاختبار	الوضع الحقيقي	
	الادعاء غير صحيح H_oصحيحة	الادعاء صحيح رفض H_o وإقرار H_1
عدم رفض H_o	قرار صائب	قرار خاطئ خطاء من النوع الثاني
رفض H_o وقبول H_1	قرار خاطئ خطاء من النوع الأول	قرار صائب

وقد يتساءل البعض عند مدى إمكانية تصغير الخطأين معاً ولكن لسوء الحظ لا يمكن تصغيرهما معاً إلى أدنى حد ممكن، ويبدو أن الطريقة الوحيد ة المتاحة لذلك هي زيادة (أو تكبير) حجم العينة، الأمر الذي قد لا يكون ممكنا في كل الحالات. لذلك فإن الذي يحدث عادة هو تثبيت أحدهما كأن يكون نسبة أو احتمال حدوث الخطأ من النوع الأول ومحاولة تصغير الآخر، و يقيس احتمال رفض الفرضية الصفرية قوة الاختبار فيما يقيس احتمال قبولها فعالية الاختبار. ويتوقف كلا الاحتمالين على القيمة الحقيقية ل μ .

والفكرة الأساسية في اختبار الفرض هي تقسيم المساحة تحت المنحنى إلى منطقتين: أحداهما تسمى " منطقة القبول " أي منطقة قبول الفرض العدم. والأخرى تسمى " منطقة الرفض"، أي منطقة رفض الفرض العدم والتي تسمى أحيانا " بالمنطقة الحرجة، والنقطة الجديرة بالملاحظة هنا هي أن منطقة القبول تمثل درجة الثقة، بينما تمثل منطقة الرفض مستوى المعنوية.

4-1-3-3- خطوات الاختبار الإحصائي:

- تحديد الفرضيات (الصفرية والبديلة)

- تحديد قاعدة القرار
- حساب القيمة الجدولية للمتغيرة
- حساب القيمة الفعلية للمتغيرة.
- اتخاذ القرار.

تتحدد كيفية إتمام كل خطوة حسب طبيعة الاختبار (ثنائي أو أحادي الاتجاه)، حسب طبيعة المجتمع و طبيعة و حجم العينة، ... و تسخدم في ذلك نظريات توزيع المعاينة.

4-2- اختبار المتوسط

يتناول هذا الاختبار متوسط المجتمع (μ)، مثل متوسط الدخل، متوسط وزن منتج معين، ويؤكد اختبار المتوسط فرضية مساواته لقيمة ما μ_0. و للقيام بالاختبار نستخرج عينة عشوائية نحسب فيها المتوسط $\bar{x}$ ثم نستخدم التوزيع الاحتمالي ل $\bar{x}$ لقياس قرب أو بعد هذه القيمة من μ_0.

4-2-1- اختبار ثنائي الاتجاه للمتوسط.

(ب) نحتاج إلى الخطوات التالية: تحديد الفرضيات، تحديد قاعدة القرار، حساب القيمة الجدولية للمتغيرة، حساب القيمة الفعلية للمتغيرة، اتخاذ القرار.

(ج) **تحديد الفرضيات** (الصفرية والبديلة):

$$H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu \neq \mu_0$$

الفرضية H_0 الفرضية الصفرية أو فرضية العدم، ويؤدي الاختبار إما إلى رفضها ونكتب RH_0 وفي هذه الحالة نقبل الفرضية البديلة أو المعاكسة أو عدم رفضها ونكتب $\overline{RH_0}$.

μ_0 هي القيمة الافتراضية لـ μ عادة ما تكون μ_0 محددة بناءا على بيانات عينة عشوائية بسيطة، وفي هذه الحالة يمكن استخدام الخاصية $\bar{x} \to N\left(\mu, \frac{\delta^2}{n}\right)$، لإجراء الاختبار.

حيث أنه تحت H_0 فإن : $\bar{x} \to N\left(\mu_0, \frac{\delta^2}{n}\right)$

وبصفة عامة نكتب: $p\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu_0}{\delta_{\bar{x}}} \leq z_{1-\frac{\alpha}{2}}\right)=1-\alpha$

حيث:

- $\frac{\bar{x}-\mu_0}{\delta_{\bar{x}}}$: (متغيرة القرار) هي المتغيرة المعيارية ل $\bar{x}$ونرمز لها ب z_c.
- $\delta_{\bar{x}}$تحدد كما يلي: $\delta_{\bar{x}}=\frac{\delta}{\sqrt{n}}$ في حالة المعاينة بالإرجاع و $\delta_{\bar{x}}=\frac{\delta}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ في الحالة المعاينة بدون إرجاع.
- $1-\frac{\alpha}{2}$: المساحة على يسار z .
- n : حجم العينة.

يمكن إذا كان $\bar{x}$ خارج المجال $1-\alpha$، أن نرفض الفرضية الصفرية التي حدد على أساسها هذا المجال ونقبل بالتالي الفرضية البديلة.

تسمى هذه (الخطة) قاعدة القرار.

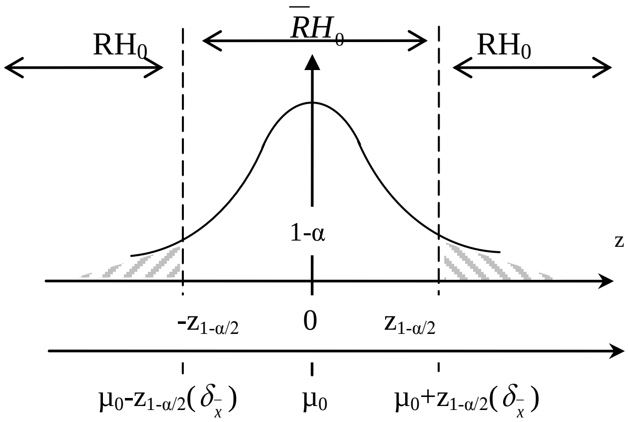
د) تحديد قاعدة القرار

تكتب قاعدة القرار، وهي قاعدة اختبار ثنائي الاتجاه، كما يلي:

$$\begin{cases} RH_0 \ si \ z_c=\dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}} \notin \left[-z_{1-\alpha/2};\ z_{1-\alpha/2}\right]. \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

أو

$$\begin{cases} RH_0 \ si \left|\dfrac{\bar{X}-\mu_0}{\sigma_{\bar{X}}}\right| > z_{1-\alpha/2}\ . \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$



الشكل: منطقتي القبول و الرفض في حالة قاعد القرار الثنائية

(هـ) حساب z الجدولية: ويرمز لها ب z_t حيث، وهي المشار إليها في قاعدة القرار.

(و) حساب z الفعلية: ويرمز لها ب z_c وهي المتغير المعيارية ل $\bar{x}$ (أنظر قاعدة القرار) :

$$z_c = \frac{\overline{X} - \mu_0}{\delta_{\bar{x}}}$$

(ز) القرار: نقرر قبول أو رفض H_0 حسب قاعدة القرار.

4-2-2- اختبار أحادي الاتجاه للمتوسط.

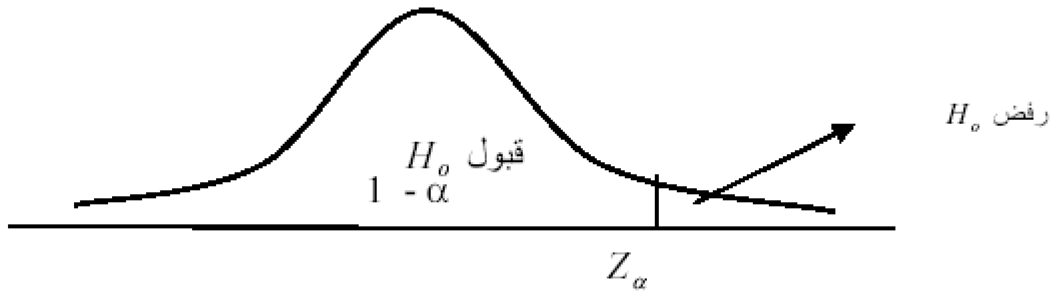
يتميز الاختبار الثنائي عن الأحادي في الفرضية البديلة التي هي عدم مساواة في الاختبار الثنائي وأكبر تماما أو أصغر تماما (حسب الحالة) في الاختبار الأحادي، وهذا يترتب عليه تغيير في قاعدة القرار.

4-2-2-1- اختبار أحادي الاتجاه من اليمين.

- **الفرضيات** : $H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu > \mu_0$

- **قاعدة القرار:**

$$\begin{cases} RH_0 \; si \dfrac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} > z_{1-\alpha} \; . \\ \overline{R}H_0 \; \text{sinon.} \end{cases}$$



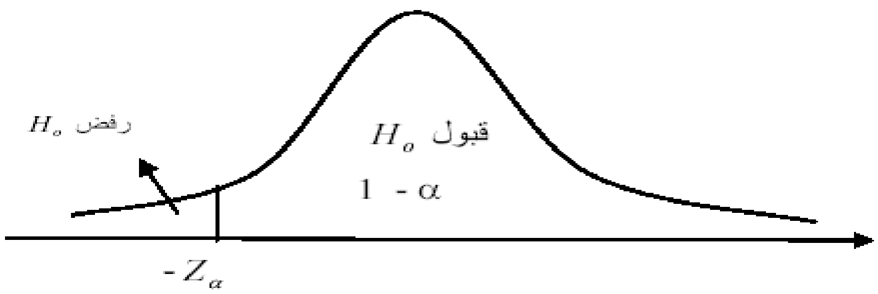
الشكل: منطقتي القبول و الرفض في حالة قاعد القرار اختبار أحادي الاتجاه من اليمين

(ح) 4-2-2-2- اختبار أحادي الاتجاه من اليسار

- **الفرضيات :** $H_0 : \mu = \mu_0 \leftrightarrow H_1 : \mu < \mu_0$

- **قاعدة القرار:**

$$\begin{cases} RH_0 \; si \; \dfrac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} < -z_{1-\alpha} \; . \\ \overline{R}H_0 \; \text{sinon.} \end{cases}$$



الشكل: منطقتي القبول و الرفض في حالة قاعد القرار اختبار أحادي الاتجاه من اليسار.

4-3- استخدام S كمقدر لـ σ والتوزيع t في اختبار المتوسط

4-3-1- استخدام S كمقدر لـ σ في اختبار المتوسط.

فيما سبق افترضنا أن σ معلوم وحجم العينة $n \geq 30$، في الواقع غالبا ما يكون الانحراف المعياري مجهولا ونحتاج بالتالي إلى استخدام الانحراف المعياري للعينة (S) عند حساب $\delta_{\bar{x}}$، حيث نعوض في عبارة متغيرة القرار أو قاعدة القرار:

$$\delta_{\bar{x}} = \frac{\widehat{S}}{\sqrt{n}} \quad \text{أو} \quad \delta_{\bar{x}} = \frac{S}{\sqrt{n-1}} \quad \text{،} \quad \delta_{\bar{x}} = \frac{\delta}{\sqrt{n}}$$

4-3-2- استخدام التوزيع t في اختبار المتوسط.

في حالة n< 30 و σ (الانحراف المعياري للمجتمع) مجهولا، لا يمكن استخدام التوزيع الطبيعي، ولكن لدينا :

$$\frac{\bar{X}-\mu}{S/\sqrt{n-1}} \sim t_{n-1} \quad \text{و تحت } H_0 \ (\mu = \mu_0) : \quad \frac{\bar{X}-\mu_0}{S/\sqrt{n-1}} \sim t_{n-1}$$

يمكن إذا استخدام التوزيع ستيودنت (بشرط أن يكون توزيع المجتمع طبيعيا) .

و تتغير قاعدة القرار تبعا لهذا التغيير فتكتب في حالة الاختبار الثنائي كما يلي:

$$\begin{cases} RH_0 \ si \left|\dfrac{\bar{X}-\mu_0}{S/\sqrt{n-1}}\right| > t_{n-1;\,1-\alpha/2} \ . \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

في حالة اختبار من اليمين:

$$\begin{cases} RH_0 \ si \dfrac{\bar{X}-\mu_0}{S/\sqrt{n-1}} > t_{n-1,\,1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

في حالة اختبار من اليسار:

$$\begin{cases} RH_0 \ si \dfrac{\bar{X}-\mu_0}{S/\sqrt{n-1}} < -t_{n-1;\,1-\alpha} \\ \overline{R}H_0 \ \text{sinon.} \end{cases}$$

مثال4: 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 الفرضية 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 للمجتمع يساوي 14 دولاراً.

<u>الحل</u> :

– الفرض الصفرية: هو أن متوسط المجتمع يساوي 72.

– الفرض البديل : هو أن المتوسط لا يساوي 72.

وبالرموز :
$$\begin{cases} H_0 : \mu = 72 \\ H_1 : \mu \neq 72 \end{cases}$$

– بما أن العينة كبيرة $n \geq 30$ و $\delta = 14$ و $\overline{X} = 75$ والاختبار هو اختبار ثنائي الاتجاه، فان متغيرة القرار تحسب كما يلي:

$$z_c = \frac{\overline{X} - \mu_0}{\delta_{\bar{x}}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}} = \frac{3}{2} = 1.5$$

أي أن قيمة الإحصائية تساوي 1.5

– حدود منطقتي القبول والرفض: نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% . وبالكشف في جدول التوزيع الطبيعي المعياري عن Z التي تقابل مستوى المعنوية 5% نجد أنها تساوي 1.96

– المقارنة والقرار : وبمقارنة قيمة الإحصائية المحسوبة والتي تساوي 1.5 $(z_c = 1.5)$ بـ z الفعلية $(z_t = 1.96)$ نلاحظ أن $z_c < z_t$، ومنه فإن القرار هو :

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

4-4- اختبار النسبة

يتعلق هذا الاختبار بنسبة مفردات المجتمع التي تتصف بخاصية ما (p)، حيث يؤكد الاختبار أو ينفي صحة فرضية معينة بخصوص قيمة p . يرمز للقيمة الافتراضية ب p_0 وتكتب الفرضية كما يلي:

$H_0 : p = p_0$

للقيام بالاختبار نستخدم خصائص p’ النسبة في العينة:

$$E(p') = \mu_{p'} = p \qquad ; \qquad \sigma^2_{p'} = \frac{pq}{n}$$

عند $n \geq 30$: $p' \rightarrow N\left(p, \delta^2_{p'}\right)$

استنادا إلى هذه الخصائص وتحت H0 :

$$p' \rightarrow N\left(p_0, \frac{p_0 q_0}{n}\right)$$

و من ثم يمكن تحديد قاعدة القرار بحسب طبيعة الاختبار كما يلي:

في حالة الاختبار الثنائي:

$$\begin{cases} RH_0 \; si \left|\dfrac{p'-p_0}{\sqrt{p_0 q_0 / n}}\right| > z_{1-\alpha/2} \; . \\ \overline{R}H_0 \; \text{sinon.} \end{cases}$$

في حالة اختبار من اليمين:

$$\begin{cases} RH_0 \; si \dfrac{p'-p_0}{\sqrt{p_0 q_0 / n}} > z_{1-\alpha} \\ \overline{R}H_0 \; \text{sinon.} \end{cases}$$

في حالة اختبار من اليسار:

$$\begin{cases} RH_0 \; si \dfrac{p'-p_0}{\sqrt{p_0 q_0 / n}} < -z_{1-\alpha} \\ \overline{R}H_0 \; \text{sinon.} \end{cases}$$

مثال5: تقدر الدوائر الرسمية نسبة المتخرجين الجامعيين الذين يحصلون على عمل في السنة الأولى التي تلي تخرجهم ب 70 % . وجدت دراسة أجريت على عينة من 900 طالب أن نسبة الحصول

على عمل 67 %. كيف يمكن اختبار ما إذا كانت النسبة الرسمية صحيحة أم مبالغ فيها، بمستوى معنوية 5 %.

$$H_0 : p = 0.70 \leftrightarrow H_1 : p < 0.70$$

$$\frac{p' - p_0}{\sqrt{p_0 q_0 / n}} = \frac{0.67 - 0.7}{\sqrt{0.7(0.3)/900}} \cong -196.34 < -z_{1-0.05} = -1.64$$

ومنه نرفض الفرضية H_0.

4-5- 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، وتكتب الفرضيات كما يلي:

- الفرضية الصفرية : أن متوسط المجتمع الأول يساوي متوسط المجتمع الثاني (أي لا يوجد فرق بين متوسطي المجتمعين).

وبالرموز : $H_0 : \mu_1 = \mu_2$

- الفرضية البديلة : أن المتوسطين غير متساويين وبالرموز :

$$H_1 : \mu_1 \neq \mu_2$$

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

لتحديد متغيرة القرار نعتمد في الاختبار على متغيرة القرار T أو T' بحسب الحالة، حيث نميز بين حالة كون تباينا المجتمعين معلومين وحالة كون تباينا المجتمعين مجهولين.

4-5-1- تباينا المجتمعين معلومين

- المجتمعين طبيعيين:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0.1)$$

حيث : يرمز بـ n1 إلى حجم العينة الأولى.

يرمز بـ n2 إلى حجم العينة الثانية.

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

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

يرمز بـ σ_1^2 إلى تباين المجتمع الأول.

يرمز σ_2^2 إلى تباين المجتمع الثاني.

- مجتمعين ما ($n_1 , n_2 \geq 30$):

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

مثال6: إذا كانت لدينا نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما حيث: $\overline{X}_1 = 40, \overline{X}_2 = 34, \delta_1^2 = 84, \delta_2^2 = 27, n_1 = 120, n_2 = 90$

- اختبر الفرضية أن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5% مقابل الفرض البديل أنهما غير متساويين؟

الحل:

- الفرض الصفرية أن المتوسطين متساويان.

– الفرضية البديلة أن المتوسطين غير متساويين.

ومنه:
$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

– الإحصائية: تأخذ الشكل التالي:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}}$$

وبالتعويض عن :

$$\overline{X}_1 = 40, \overline{X}_2 = 34, \delta_1^2 = 84, \delta_2^2 = 27, n_1 = 120, n_2 = 90$$

نحصل على:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}} = \frac{40 - 34}{\sqrt{\frac{84}{120} + \frac{27}{90}}}$$
$$= \frac{6}{\sqrt{0.7 + 0.3}} = 6$$

أي أن قيمة $T = z_c$ الإحصائية تساوي 6

– 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 +. أي سوف نقارن بالقيمة $z_t = 1.96$.

– المقارنة والقرار: بما أن قيمة الإحصائية T والتي تساوي 6 تقع في منطقة الرفض أي $T = z_c > 1.96$ فإن القرار هو رفض الفرض الصفرية وقبول الفرضية البديلة 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 %.

4-5-2- تباينا المجتمعين مجهولين

- المجتمعان طبيعيان:

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:

$$T' = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \to t_{n_1+n_2-2}$$

حيث :

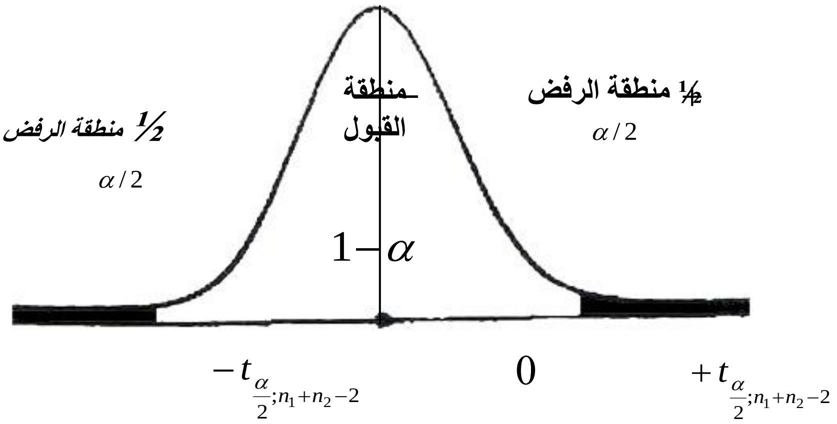
$$S^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n1+n_2-2}$$

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

هي :

- الفرضية الصفرية $H_0:\mu_1=\mu_2$
- الفرضية البديلة $H_1:\mu_1\neq\mu_2$

- الإحصائية هي المكتوبة أعلاه (وهي في هذه الحالة t وليست Z)
- حدود منطقتي القبول والرفض ونحصل عليها في هذه الحالة من جدول t عند درجات حرية تساوي n_1+n_2-2 وعند مستوى معنوية يساوي $\frac{\alpha}{2}$ كما في الشكل التالي:



– المقارنة والقرار : كما سبق.

– **مجتمعين ما $(n_1\ ,\ n_2 \geq 30)$:** أن تباين المجتمعين غير متساويين:

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

مثـال7: البيانـات التاليـة تمثل نتـائج عينتين عشوائيتين مستقلتين مسحوبتين من مدينتين عن أعمـار 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 الفرضية الصفرية: $Ho:\mu_1=\mu_2$ مقابل الفرضية البديلة $HI:\mu_1\neq\mu_2$

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

<u>الحل</u> :

– الفرضية الصفرية: $Ho:\mu_1=\mu_2$

أي متوسط أعمار الناخبين في المدينتين متساوٍ

– الفرضية البديلة : $HI:\mu_1\neq\mu_2$

أي متوسط أعمار الناخبين في المدينتين غير متساوٍ

– الإحصائية لاحظ (أن العينات صغيرة، وأن تباين المجتمعين هو نفسه، وأن المجتمعين طبيعيان).
فإن الإحصائية المناسبة في هذه الحالة هي T' :

$$T' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

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 :

$$T' = \frac{28-26}{\sqrt{40\left(\frac{1}{10}+\frac{1}{10}\right)}}$$

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

أي أن قيمة الإحصائية تساوي 0.7

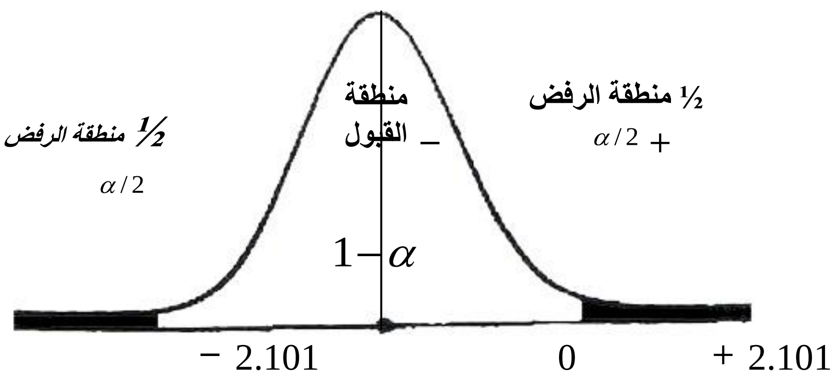
– حدود منطقتي القبول والرفض:

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 $\frac{0.05}{2} = 0.025$ أي أن $t_{0.025,18} = 2.101$

كما في الشكل التالي :



أي أن منطقة القبول تبدأ من -2.101 وحتى 2.101 +

- المقارنة والقرار :

حيث أن قيمة الإحصائية تساوي 0.7 أقل من قيمة ستيودنت الجدولية والتي تساوي 2.101 ، وهي تقع في منطقة القبول وبالتالي فإن القرار هو قبول الفرضية الصفرية 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).

4-6- اختبار الفرق بين نسبتين :

ربما نرغب 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 :

- الفرضية الصفرية: أن النسبة في المجتمعين متساوية أي: $H_0 : p_1 = p_2$
- الفرضية البديلة: أن النسبتين في المجتمعين غير متساوية: $H_0 : p_1 \neq p_2$

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

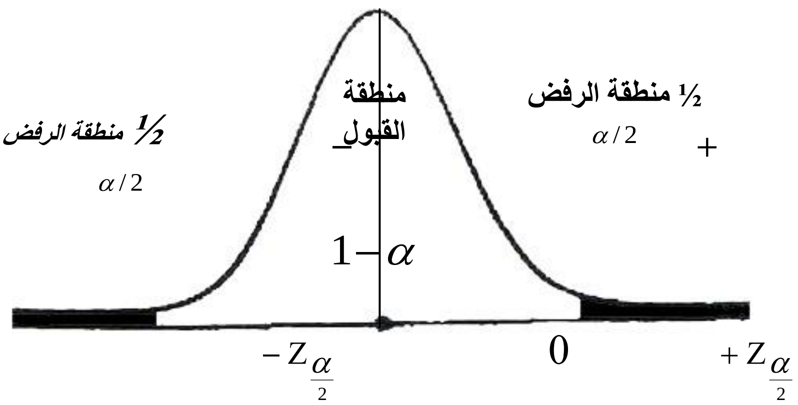
- الإحصائية (متغيرة القرار): بافتراض أن العينتين كبيرتان بدرجة كافية تكون الإحصائية كما يلي:

$$Z_{p'_1-p'_2}=\frac{p'_1-p'_2}{\sqrt{p'q'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\rightarrow N(0,1)$$

حيث:
$$p'=\frac{n_1p'_1+n_2p'_2}{n_1+n_2}$$
$$q'=1-p'$$

أي يتم أولاً حساب p' (والتي تمثل متوسط مرجح من نسبتي العينتين) قبل التعويض في الإحصائية والتي لها توزيع طبيعي معياري.

– 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:



– المقارنة والقرار: كما سبق

مثال8: لاختبار ما إذا كانت نسبة المؤيدين لبرنامج اقتصادي معين في المدينة A يساوي نسبة المؤيدين لهذا البرنامج في المدينة B تم اختيار عينتين عشوائيتين مستقلتين من المدينتين حيث: حجم العينة الأولى يساوي حجم العينة الثانية يساوي 100 وكانت نسبة المؤيدين للبرنامج في عينة المدينة A هي: $p'_1=0.70$ ونسبة المؤيدين للبرنامج في عينة المدينة B هي: $p'_2=0.50$.

– اختبر الفرضية الصفرية أن النسبة في المدينتين متساوية مقابل الفرضية البديلة أنها غير متساوية وذلك بمستوى معنوية 1 %؟.

الحل:

- الفرضية الصفرية: النسبة في المدينة A تساوي النسبة في المدينة B.
- الفرضية البديلة : النسبة في المدينتين غير متساوية.

ومنه تكتب الفرضيتين كما يلي: $\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}$

- متغيرة القرار:

$$Z_{p_1^/ - p_2^/} = \frac{p_1^/ - p_2^/}{\sqrt{p^/ q^/ \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

حيث: $p^/ = \dfrac{n_1 p_1^/ + n_2 p_2^/}{n_1 + n_2}$

$q^/ = 1 - p^/$

وبالتعويض عن : $p_1^/ = 0.7; p_2^/ = 0.5; n_1 = 100; n_2 = 100$ نجد:

$$p^/ = \frac{(100)(0.7) + (100)(0.5)}{100 + 100}$$

$$= \frac{70 + 50}{200} = \frac{120}{200}$$

$$p^/ = 0.60$$

$$q^/ = 1 - 0.6 = 0.40$$

وبالتعويض في متغيرة القرار نجد:

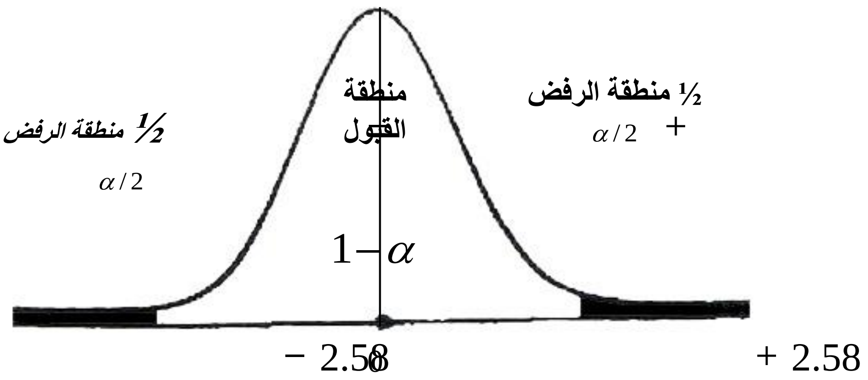
$$Z_{p_1^/ - p_2^/} = \frac{p_1^/ - p_2^/}{\sqrt{p^/ q^/ \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.70 - 0.50}{\sqrt{(0.6)(0.4)\left(\frac{1}{100} + \frac{1}{100}\right)}}$$

$$= \frac{0.20}{0.069} = 2.899$$

أي أن قيمة الإحصائية تساوي 2.899

- 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 + أي أن $z_t = 2.58$

- المقارنة والقرار:

بما أن قيمة الإحصائية تساوي 2.899 فهي تقع في منطقة الرفض أي $z_c > z_t$ وبالتالي فإن القرار هو: رفض الفرضية الصفرية وقبول الفرضية البديلة أي رفض الفرض القائل بأن نسبة المؤيدين للبرنامج الاقتصادي في المدينة A تساوي نسبة المؤيدين له في المدينة B وذلك بمستوى معنوية 1%، وقبول الفرض البديل بأن النسبتين غير متساويتين.

4-7- اختبار التباين وتساوي تبايني مجتمعين.

4-7-1- اختبار التباين.

لاختبار صدقية فرضية بخصوص قيمة تباين مجتمع ما،

$$H_0 : \sigma^2 = \sigma_0^2 \quad \leftrightarrow H_1 : \sigma^2 = \sigma_0^2$$

نستعمل المقدر غير المنحاز $\hat{S}^2 = \dfrac{\sum_i (X_i - \bar{X})^2}{n-1}$. حيث في حالة العينة الكبيرة ($n \geq 50$ في أحسن الأحوال)، وتحت H_0 فإن $\mu_4 = E(X-\mu)^4$

$$T = \frac{\hat{S}^2 - \sigma_0^2}{\sqrt{(\mu_4 - \sigma_0^4)/n}} \approx N(0.1).$$

حيث μ_4 هو العزم المركزي من الدرجة الرابعة. وبهذا الشكل تكتب قاعدة القرار للاختبار الثنائي كما يلي:

$$\begin{cases} RH_0 \;\; si \left| \dfrac{\hat{S}^2 - \sigma_0^2}{\sqrt{(\mu_4 - \sigma_0^4)/n}} \right| > z_{1-\alpha/2} \\ \overline{R}H_0 \;\; \text{sinon.} \end{cases}$$

وفي حالة μ_4 مجهول يمكن استخدام كمقدر:

$$\bar{x}_4 = E\left(x_i - \bar{x}\right)^4$$

وإذا كان المجتمع طبيعيا، حيث $\mu_4 = 3\sigma^4$ ، فإن متغيرة القرار يمكن أن تكتب كما يلي:

$$T = \frac{\hat{S}^2 - \sigma_0^2}{\sigma_0^2 \sqrt{2/n}} \approx N(0.1).$$

4-7-2- اختبار تساوي تبايني مجتمعين

الغرض من الاختبار هو تأكيد أو نفي تساوي تباينا مجتمعين من خلال عينتين عشوائيتين بسيطتين مستقلتين.

تكتب الفرضيات (في حالة الاختبار الثنائي)كما يلي:

$$H_0 : \sigma^2_1 = \sigma^2_2 \leftrightarrow H_1 : \sigma^2_1 \neq \sigma^2_2$$

نعتمد في الاختبار على متغيرة القرار T أو T’ بحسب الحالة، حيث نميز بين حالة كون المجتمعين طبيعيين أم غير ذلك.

4-7-2-1- مجتمعين طبيعيين.

- الحالة العامة:

$$T' = \frac{\hat{S}_1^2}{\hat{S}_2^2} \sim F_{n_1-1;\ n_2-1}$$

- في حالة $n_1 , n_2 \geq 30$

$$T = \frac{1}{2}\ln\left(\frac{\hat{S}_1^2}{\hat{S}_2^2}\right) / \sqrt{\frac{1}{2}\left(\frac{1}{n_1-1} + \frac{1}{n_2-1}\right)} \approx N(0;1)$$

4-7-2-2- مجتمعين ما $(n_1, n_2 \geq 50)$

- $\mu_4^{(1)}$; $\mu_4^{(2)}$ معروفين :

$$T = (\hat{S}_1^2 - \hat{S}_2^2) / \sqrt{\frac{\mu_4^{(1)} - \hat{S}_1^4}{n_1} - \frac{\mu_4^{(2)} - \hat{S}_2^4}{n_2}} \approx N(0;1)$$

– في حالة $\mu_4^{(2)}$; $\mu_4^{(1)}$ غير معروفين : نعوض μ_4 بـ $\bar{x}_4$.

مثال9: نسحب من مجتمعين طبيعيين عينتين حجم الأولى 18 وحجم الثانية 21. وجدنا النتائج التالية:

$S_1^2 = 9; S_2^2 = 8;$

– كيف يمكن إجراء اختبار تأكيد أو نفي تساوي تباينا مجتمعين بمستوى معنوية 5 % ؟

الحل:

$$H_0 : \sigma^2_1 = \sigma^2_2 \leftrightarrow H_1 : \sigma^2_1 \neq \sigma^2_2$$

– متغيرة القرار في هذه الحالة هي:

$$T' = \frac{\hat{S}_1^2}{\hat{S}_2^2} \sim F_{n_1-1;\, n_2-1}$$

لدينا:

$$\widehat{S}_1^2 = S_1^2 \frac{n_1}{(n_1-1)} = 9\frac{18}{17} \approx 9.53$$

$$\widehat{S}_2^{\prime} = S_2^2 \frac{n_2}{n_2-1} = 8\frac{21}{20} \approx 8.4$$

$$T^{\prime} = \frac{\widehat{S}_1^2}{\widehat{S}_2^{\prime}} = \frac{9.53}{8.4} \approx 1.135$$

$$F_{0.05;17;20} \approx 2.17$$

$$T^{\prime} = 1.135 < F_{0.05;17;20} \approx 2.17$$

$$\Rightarrow \overline{R}H_0$$