وتسمى هذه الطريقة للاستدلال الإحصائي اختبار الفرضية،والغرض من هذا النوع من الاستدلال الإحصائي هو تحديد ما إذا كانت نتائج العينة تساند أو تفشل في مساندة اعتقاد معيّن أو فرضية محددة حول قيمة (معلمة) المجتمع التي يحددها الباحث ،وبذلك فان الفرضية الاحصائية هي عبارة عن تخمين أو تأكيد عن التوزيعات التي تحتوي على متغير عشوائي واحد أو أكثر. بعبارة أخرى يمكن القول بأن الفرضية الإحصائية هي تصريح أو ادعاء (قد يكون صائباً أو خاطئاً) يتعلق بالتوزيع الاحتمالي لمتغير عشوائي .

الهدف من اختبار الفرضية هو الاجابة عن هذا السؤال:

هل نتائج العينة متوافقة مع القيمة المفترضة لمعلمة المجتمع أو أن نتائج العينة تتعارض مع هذه القيمة؟

وبرفض إمكانية قبول القيمة المفترضة مبدئيا فاننا ننشئ بشكل غير مباشر إمكانية قبول القيمة المفترضة البديلة أو مدى من القيّم. وبهذه الطريقة يقوم الباحث بالاستدلال حول قيمة معلمة المجتمع.

فرضية العدم : (The Null Hypothesis)
وهي الفرضية الإحصائية التي يضعها الباحث على أمل أن يرفضها ويرمز لها (H_0) .

الفرضية البديلة : (The Alternative Hypothesis)
رفضنا لفرضية العدم يقودنا إلى قبول فرضية أخرى تسمى الفرضية البديلة ويرمز لها (H_1)وهي الفرضية التي ستقبل في حال رفضنا لفرضية العدم .

الخطأ من النوع الاول : هوالخطأ الحاصل بسبب رفض (H_0) عندما تكون صحيحة .

الخطأ من النوع الثاني : هوالخطأ الحاصل بسبب قبول(H_0) عندما تكون خاطئة .

القرار \ الحالة	H_0 صائبة	H_0 خاطئة
قبول H_0	قرار صائب	خطأ من النوع الثاني
رفض H_0	خطأ من النوع الاول	قرار صائب

مخطط يوضح الخطأ من النوع الاول والثاني واتخاذ القرار

مستوى المعنوية : Level of Significance

يعتبر هذا المصطلح واحدا من أهم المصطلحات المستخدمة في اختبار الفرضيات ، ويسمى احيانا (مستوى الدلالة) والمقصود به احتمال حدوث الخطأ من النوع الاول أو نسبة حدوثه أي احتمال رفض فرضية العدم عندما تكون صائبة ويرمز له (α) .

من الملاحظ ان تعبير مستوى المعنوية هو المكمل لمصطلح (درجة الثقة)، فاذا كانت درجة الثقة (95%)فان مستوى المعنوية (5%).

والفكرة الاساسية في اختبار الفرضية هي تقسيم المساحة تحت المنحنى الى منطقتين : احداهما تسمى " منطقة القبول " أي منطقة قبول فرضية العدم ،واخرى تسمى " منطقة الرفض"، أي رفض فرضية العدم والنقطة الجديرة بالملاحظة هنا هي أن منطقة القبول تمثل درجة الثقة بينما تسمى منطقة الرفض مستوى المعنوية .

خطوات أختبار الفروض الاحصائية:

1-صياغة فرضيتين احدهما فرضية العدم (H_0) والاخرى الفرضية البديلة (H_1) حول معلمة في مجتمع الدراسة .

2- يتم حساب احصاءة الاختبار من بيانات العينة .

3- نقارن بين احصاءة الاختبار والقيمة الجدولية من خلال منطقة القبول او الرفض.

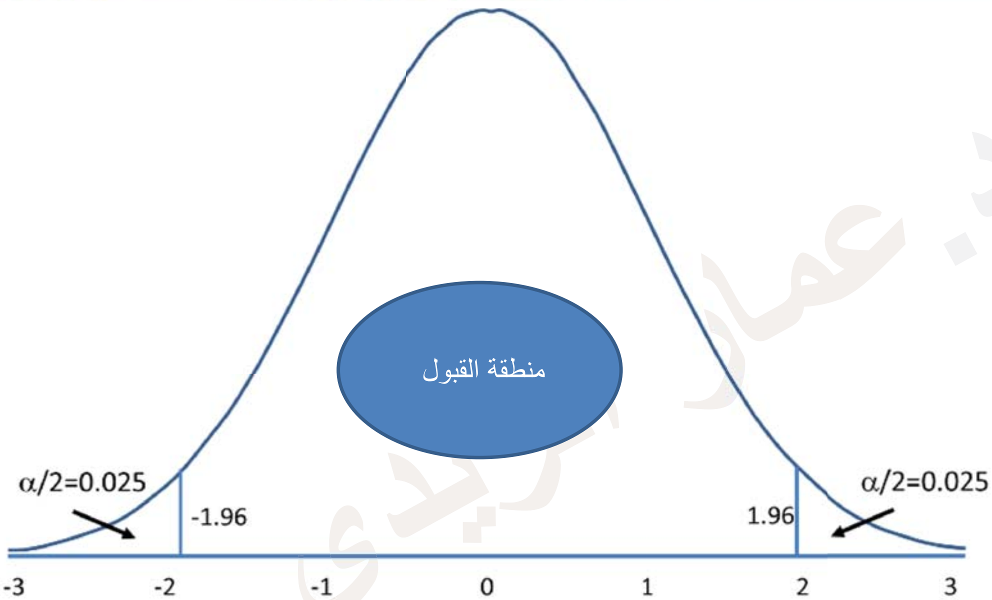
4-نحدد القرار برفض او قبول الفرضية (H_0) .

صياغة فرضية العدم والفرضية البديلة لمتوسط المجتمع (μ_0):

1- الاختبار ذو الاتجاهين(الطرفين) :

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

وياخذ الشكل التالي بافتراض ان (α=5%)

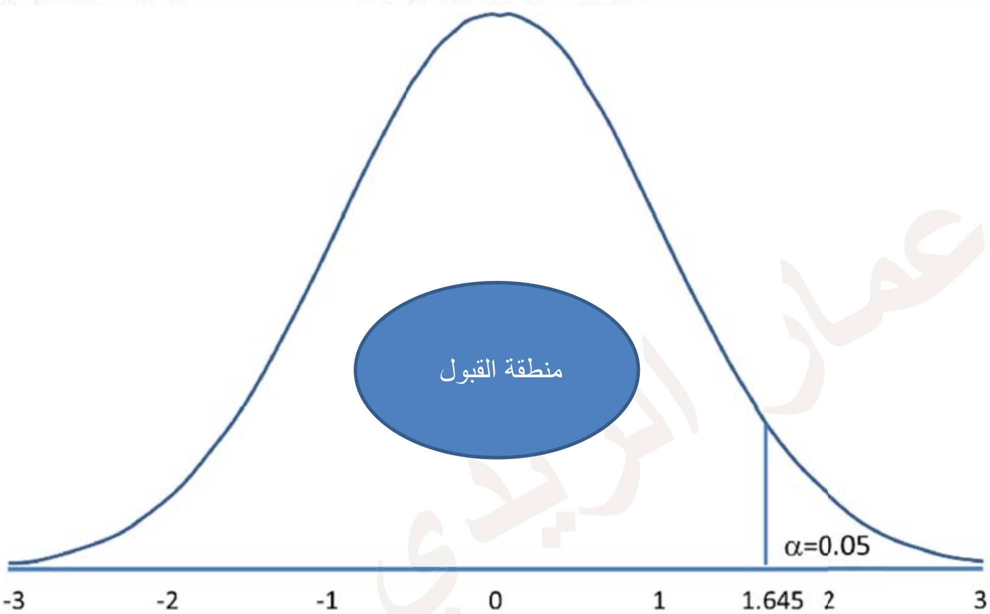


Rejection Region for Two-Tailed Z Test (H_1: $\mu \neq \mu_0$) with $\alpha = 0.05$
The decision rule is: Reject H_0 if $Z \leq -1.960$ or if $Z \geq 1.960$.

2- الاختبار ذو الطرف الايمن :

$$H_0 : \quad \mu \leq \mu_0$$
$$H_1 : \quad \mu > \mu_0$$

وياخذ الشكل التالي بافتراض ان (α=5%)



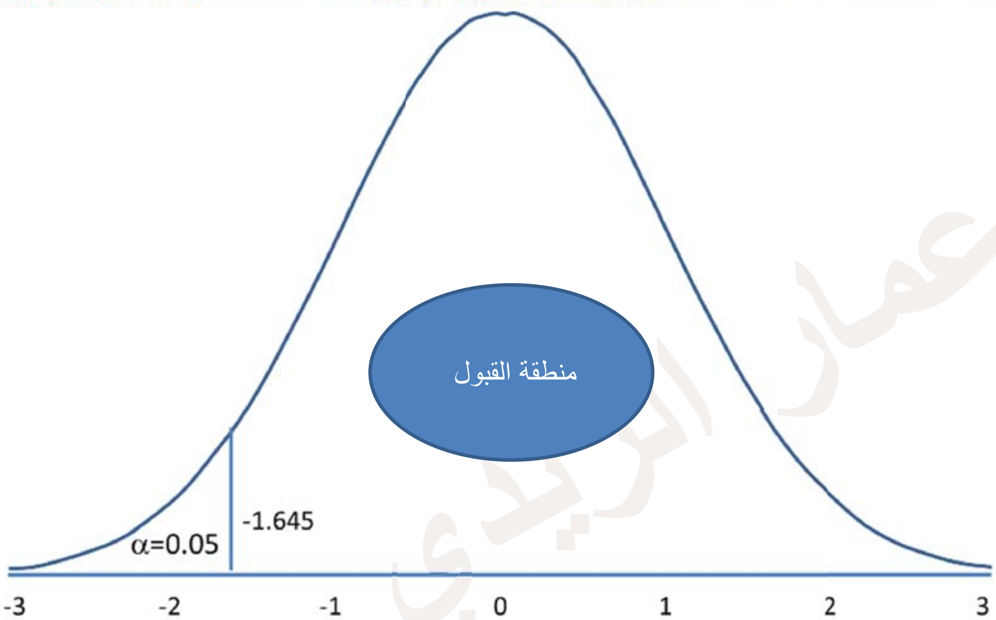
Rejection Region for Upper-Tailed Z Test (H_1: $\mu > \mu_0$) with α=0.05

The decision rule is: Reject H_0 if $Z \geq 1.645$.

3- الاختبار ذو الطرف الايسر :

$$H_0 : \quad \mu \geq \mu_0$$
$$H_1 : \quad \mu < \mu_0$$

وياخذ الشكل التالي بافتراض ان (α=5%)



Rejection Region for Lower-Tailed Z Test (H_1: $\mu < \mu_0$) with α =0.05
The decision rule is: Reject H_0 if $Z \le 1.645$.

أحصاءة الاختبار:

الحالة الاولى : عندما تكون قيمة σ^2 معلومة وحجم العينة كبير .

$$Z=\frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$$

الحالة الثانية : عندما تكون قيمة σ^2 غير معلومة ويتم تقديرها s^2 وحجم العينة صغير .

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

مثال : عينة عشوائية بحجم 49 شخصا تم اختيارها من مجتمع معين فاذا كان الوسط الحسابي لدخول الافراد الاسبوعية في العينة 75 دولار والانحراف المعياري للمجتمع لدخول الافراد 14 دولار ،هل يمكن الاستنتاج ان متوسط الدخل الاسبوعي هو 72 دولار عند مستوى دلالة 5 % علما ان $z_{0.025} = 1.96$

الحل : **n=49, σ= 14, $\bar{X}$=75, μ=72**

فرضية الاختبار :

H_0 : μ = **72**

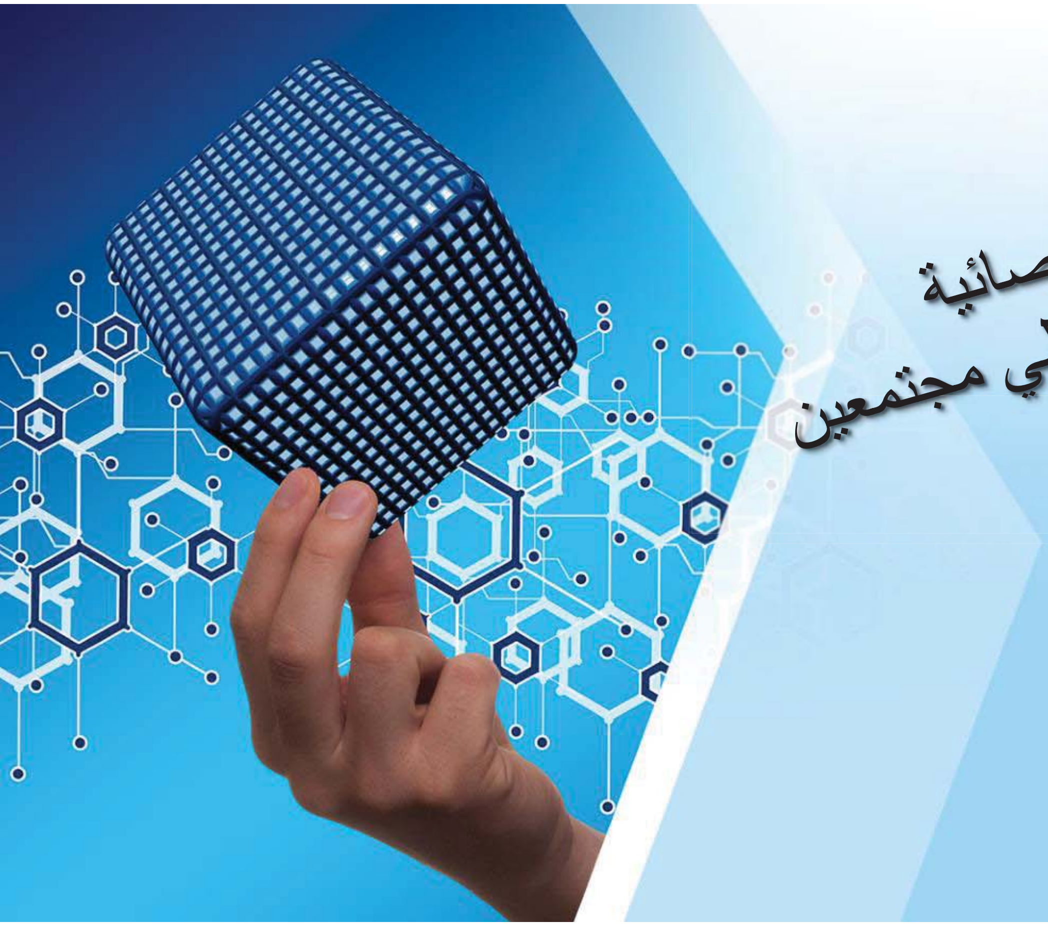
H_1: $\mu \neq$ **72**

احصاءة الاختبار ستكون :

$$Z= \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{75-72}{\frac{14}{\sqrt{49}}} = \frac{3}{2} = 1.5$$

قيمة جدولية تعطى بالسؤال $z_{table} = z_{\left(\frac{\alpha}{2}\right)} = z_{\left(\frac{0.05}{2}\right)} = z_{0.025} = 1.96$

بما انه قيمة Z الحسابية تساوي (1.5) اصغر من Z الجدولية والتي تساوي (1.96) بذلك نقبل فرضية العدم H_0 اي ان متوسط دخل الفرد الاسبوعي 72



محاضرة 7
الفرضيات الاحصائية
المتعلقة بمتوسط و متوسطي مجتمعين

أ.م.د.عمار الزيدي

مثال :اختار باحث عينة عشوائية بحجم (9) من طلبة المرحلة الثالثة واجرى لهم اختبار في مادة الاحصاء وتم الحصول على الدرجات الاتية ، هل يمكن اعتبار ان المتوسط العام يختلف عن (5.7) اختبر ذلك عند مستوى دلالة 5%

الدرجات :

X_i	10	9	7	8	9	3	5	8	5

الحل : كتابة الفرضية

$H_0 : \mu = 5.7$
$H_1 : \mu \neq 5.7$

الوسط الحسابي للعينة ($\bar{X}$) :

$$\bar{X} = \frac{\sum x}{n} = \frac{64}{9} = 7.1$$

S= 2.32

$$t = \frac{7.1 - 5.7}{\frac{2.32}{\sqrt{9}}} = \frac{1.4}{\frac{2.32}{\sqrt{9}}} = 1.82$$

ان القيمة الجدولية عند مستوى دلالة ($\frac{0.05}{2}$) ودرجة حرية 8 تكون (2.306)

بما انه قيمة t الحسابية والبالغة (1.82) اقل من قيمة t الجدولية والبالغة (2.306) بذلك نقبل فرضية العدم اي انه لايمكن اعتبار المتوسط العام يختلف عن (5.7)

مثال : يدعي احد الاطباء ان من الاثار الجانبية لاستعمال دواء معين هو انخفاض ضغط الدم بمتوسط (75) ، تم سحب عينة بحجم (49) مريض ،وتم قياس ضغط الدم بعد تعاطي هذا الدواء فوجد ان متوسط ضغط الدم في العينة (65.5) بانحراف معياري مقداره (6.4) هل يمكن الاستنتاج ان ادعاء الطبيب صائب عند مستوى دلالة 5% ؟

الحل : من السؤال يتوفر لدينا

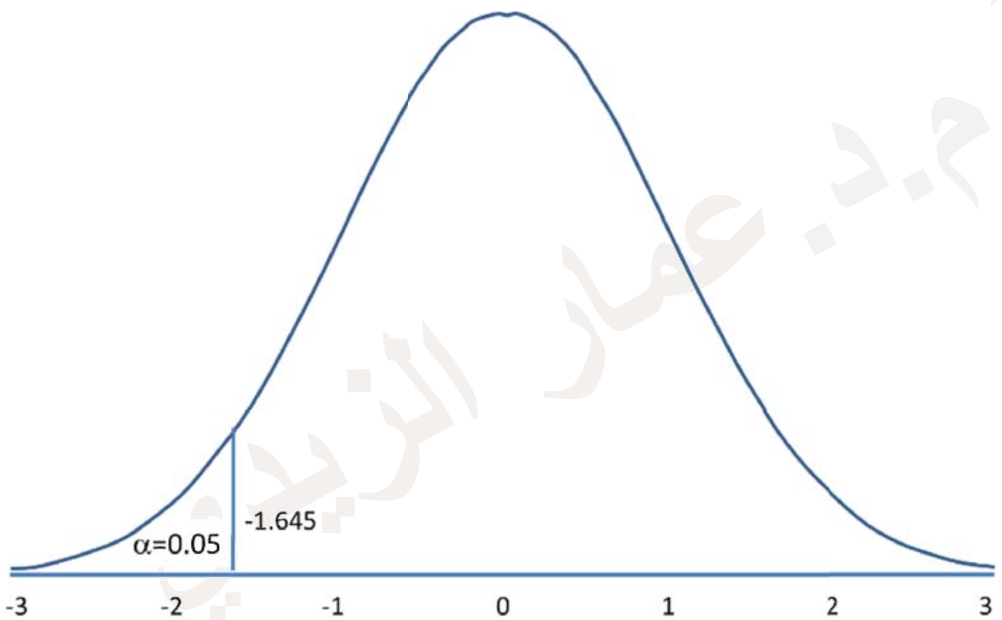
$n = 49 \quad , \quad s = 6.4 \quad , \bar{X} = 65.5 \quad , \quad \mu = 75$

$$H_0 : \ \mu \geq 75$$
$$H_1 : \ \mu < 75$$

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{65.5-75}{\frac{6.4}{\sqrt{49}}} = \frac{-9.5}{0.91} = -10.44$$

$z_{table} = z_{\alpha} = z_{0.05} = 1.64$ قيمة جدولية

بما انه قيمة Z الحسابية تساوي (-10.44) اصغر من Z الجدولية والتي تساوي (-1.64) بذلك نرفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 اي ان ادعاء الطبيب كان صائب .


-1.645
α=0.05
-3
-2
-1
0
1
2
3

اختبار الفرق بين متوسطين حسابيين :

صياغة فرضية العدم والفرضية البديلة للفرق بين متوسطين حسابيين :

1- الاختبار ذو الاتجاهين(الطرفين) :

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

2- الاختبار ذو الطرف الايمن :

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

3- الاختبار ذو الطرف الايسر :

$$H_0:\ \mu_1 \ge \mu_2$$
$$H_1:\ \mu_1 < \mu_2$$

أحصاءة الاختبار:

الحالة الاولى : بافتراض ان المجتمعين طبيعيان وان العينتين مستقلتان وكبيرتان وتباين المجتمع للعينتان معلوم $\sigma^2{}_2,\ \sigma^2{}_1$

$$Z= \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}}$$

الحالة الثانية : بافتراض ان العينات صغيرة (مجموع العينتين أقل من 30) وبافتراض ان المجتمعين طبيعيان وان العينتين مستقلتان وتباين المجتمع الاول يساوي تباين المجتمع الثاني ولكنه مجهول (بمعنى ان $\sigma^2{}_1 = \sigma^2{}_2$) ولكن قيمة هذا التباين غير معروفة سيكون لها توزيع t بدرجة حرية $(n_1 + n_2 - 2)$

فان احصاءة الاختبار ستكون :

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{(s)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1{}^2 + (n_2 - 1)s_2{}^2}{n_1 + n_2 - 2}}$$

مثال :البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما اذ ان ($n_1 = 100$, $n_2 = 80$, $\bar{X}_1 = 35$, $\bar{X}_2 = 29$, $\sigma^2{}_1 = 60$, $\sigma^2{}_2 = 32$)

هل يمكن الاستنتاج ان متوسط عمر الناخب في المنطقة الاولى يساوي متوسط عمر الناخب في المنطقة الثانيةعند مستوى دلالة 5% ؟

الحل : كتابة الفرضية

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

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}} = \frac{(35 - 29) - (0)}{\sqrt{\frac{60}{100} + \frac{32}{80}}}$$

$$= \frac{6}{\sqrt{0.6 + 0.4}} = 6$$

قيمة جدولية $z_{table} = z_{(\frac{\alpha}{2})} = z_{(\frac{0.05}{2})} = z_{0.025} = 1.96$

بما انه قيمة Z الحسابية تساوي (6) اكبر من Z الجدولية والتي تساوي (1.96)
بذلك نرفض فرضية العدم H_0 اي ان متوسط عمر الناخب في المنطقتين غير متساوي

مثال :البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما اذ ان ($s^2{}_2$= 30 , $s^2{}_1$= 50 , $\bar{X}_2 = 26$, $\bar{X}_1 = 28$, $n_2 = 10$, $n_1 = 10$)

هل يمكن الاستنتاج ان متوسط عمر الناخب في المنطقة الاولى يساوي متوسط عمر الناخب في المنطقة الثانيةعند مستوى دلالة 5% ؟

الحل : كتابة الفرضية

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{(s)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1{}^2 + (n_2 - 1)s_2{}^2}{n_1 + n_2 - 2}}$$

$$s = \sqrt{\frac{(10-1)50+(10-1)30}{10+10-2}} = 6.32$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{(s)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(28 - 26) - (0)}{(6.32)\sqrt{\frac{1}{10}+\frac{1}{10}}} = 0.71$$

$t_{table} = t_{(\frac{\alpha}{2}, n_1 + n_2 - 2)} = t_{(\frac{0.05}{2},18)} = t_{(0.025,18)} = 2.101$ قيمة جدولية

بما انه قيمة t الحسابية تساوي (0.71) اصغر من t الجدولية والتي تساوي (2.101)
بذلك نقبل فرضية العدم H_0 اي ان متوسط عمر الناخب في المنطقتين متساوي .