وزارة التعليم العالي والبحث العلمي
جامعة محمد بوضياف - مسيلة-
كلية العلوم الاقتصادية والتجارية وعلوم التسيير
قسم العلوم الاقتصادية
تخصص اقتصاد كمي

محاضرات موجهة لطلبة السنة الثالثة قسم العلوم الاقتصادية تخصص اقتصاد كمي في مقياس تطبيقات على الحاسوب

المحور الثالث:

اختبار t لعينة واحدة

إعداد الاستاذ: جحيرة عبد المنعم

2021-2020

اختبار t في حالة العينة الواحدة:

يفيد هذا الاختبار في الكشف عن ما إذا كان هناك فرق جوهري (دال إحصائيا) بين المتوسط الحسابي للمجتمع الذي سحبت منه العينة وقيمة ثابتة Constant ، كما يفيد ذات الاختبار في تقدير مجال الثقة لمتوسط المجتمع وهو اختبار يستخدم في حالة العينات الصغيرة $n<30$.

مبدأ الاختبار:

باستخدام معادلة الدرجات المعيارية يمكن اختبار متوسط عينة سحبت من مجتمع متوسطه الحسابي µ المعروف، وانحرافه المعياري σ المعروف أيضا، حيث تصبح معادلة الدرجات المعيارية كالتالي :

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

حيث:

$\bar{X}$: متوسط العينة.

μ : متوسط المجتمع.

$\sigma_{\bar{X}}$: الخطأ المعياري ويساوي $\frac{\sigma}{\sqrt{n}}$

مثال:

يتوزع افراد المجتمع الجزائري من حيث الطول وفق توزيع طبيعي بمتوسط حسابي قدره 170 سم وبانحراف معياري σ قدره 20 سم، تم سحب عينة تتكون من 100 فرد فوجد ان متوسط أطوالهم يساوي 175 سم. هل يمكن القول أن متوسط طول الفرد الجزائري قد تغير بشكل دال؟

الحل:

للاستدلال عن متوسط المجتمع فإننا نقارن بين المتوسط الحسابي للعينة $\bar{X}$ والمتوسط الحسابي للمجتمع μ باستخدام اختبار Z .

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{175 - 170}{20/\sqrt{100}} = 2.5$$

بالرجوع إلى جدول التوزيع الطبيعي نجد ان:

$$P(z > 2.5) = 0.0062$$

أي أن احتمال أن يزيد طول أفراد المجتمع عن 175 سم هو 0.0062.

إن العيب في استخدام الإحصاء z يكمن في كونه يتطلب معلومات لا يمكن الحصول عليها في الواقع فتطبيقه يتطلب معرفة الانحراف المعياري للمجتمع σ لحساب الخطأ المعياري، لكن في أغلب الأحيان يكون الانحراف المعياري للمجتمع غير معروف، في حين ان الغرض من استخدام الاستدلال الإحصائي هو الحصول على معلومات عن المجتمع، وهذه الحالة تخلق نوع من التناقض، فنحن نريد استخدام الدرجات المعيارية والتوزيع الطبيعي للحصول على معلومات عن المجتمع، لكن في نفس الوقت يجب ان تتوفر لدينا ذات المعلومات عن المجتمع لنتمكن من استخدام اختبار Z ؟ لحسن الحظ يوجد حل لهذه المعضلة ويتمثل في استخدام الانحراف المعياري للعينة s_d بدلا من الانحراف المعياري للمجتمع.

يمكن تعويض الخطأ المعياري للعينة في إحصائية اختبار Z ، غير ان الاختبار يصبح يحمل اسم اختبار t لعينة واحدة ويكتب كالتالي :

$$t = \frac{\bar{x} - \mu}{s_d/\sqrt{n}}$$

درجات الحرية في اختبار t لعينة واحدة تساوي : $df = n - 1.$

فكلما كانت درجات الحرية كبيرة، كان تباين العينة S^2 أكثر تمثيلا لتباين المجتمع σ^2 وزاد اقتراب اختبار t من اختبار Z ولهذا الاستنتاج تفسير مفاده أنه كلما كان حجم العينة كبيرا أصبحت العينة ممثلة للمجتمع.

- **فرضيات الاختبار:**

الفرضية الصفرية: المتوسط الحسابي للمجتمع الذي سحبت منه العينة يساوي القيمة النظرية b .

$$H_0: \mu = b$$

الفرضية البديلة : المتوسط الحسابي للمجتمع الذي سحبت منه العينة يختلف عن القيمة النظرية b .

$$H_1: \mu \neq b$$

- **إحصائية الاختبار:**

$$t = \frac{\bar{x} - \mu}{s_d/\sqrt{n}}$$

حيث:

$\bar{x}$: متوسط العينة الحسابي.

μ : متوسط المجتمع بموجب االفرضية الصفرية ويساوي القيمة النظرية b .

- **قاعدة اتخاذ القرار:**

نرفض الفرضية الصفرية إذا كانت القيمة المحسوبة أكبر أو تساوي القيمة المجدولة بمعنى ان الفرق الملاحظ بين متوسط العينة والدرجة الثابتة فرق دال إحصائيا، اما إذا كانت القيمة المحسوبة أصغر من القيمة المجدولة فذلك يعني أن الفرق الملاحظ بين متوسط العينة والدرجة الثابتة فرق دال إحصائيا بل هو فرق ظاهري يرجع إلى الصدفة.

- **تقدير مجال الثقة الذي يتراوح فيه متوسط المجتمع:**

يشكل التقدير الإحصائي للمتوسط الحسابي إحدى الأهداف الأساسية للإحصاء الاستدلالي، والمقصود بالتقدير هو إمكانية التعرف على معلمة معينة من المجتمع الإحصائي انطلاقا من الإحصائية المناسبة للعينة. عندما يختار الباحث عينة عشوائية ويتأكد من كونها حقيقة عشوائية وممثلة للمجتمع الذي أخذت منه، فغنه يمكنه القول بان متوسط المجتمع هو متوسط العينة، في هذه الحالة نقول أن الباحث استخدام التقدير النقطي، أي انه اعتمادا على متوسط معين قام بتقدير المتوسط الحسابي معتمدا على قيمة معينة في العينة وتعرف على المعلمة أو القيمة المناسبة لها. غير انه عادة ما يحدث أن يكون الباحث غير متأكد من العينة ممثلة للمجتمع الإحصائي ولو انها عشوائية، في هذه الحالة يلجأ إلى طريقة أخرى هي التقدير بمجال الثقة حيث يحدد مجال بقيمتين يتراوح فيه المتوسط الحسابي للمجتمع الإحصائي، ولتحديد هذا المجال نستخدم اختبار Z .

يتحدد المجال الذي يتراوح فيه المتوسط الحسابي لتوزيع المعاينة بين الحدين (H) و (L) بنسبة ثقة مناسبة مثل 95% أي ان احتمال الخطأ في تقدير المتوسط هو 0،05 ، تسمى القيم (H) و (L) بالقيم الحرجة وتحدد بالمعادلتين التاليتين :

$$L = \bar{X} - (z)s_x$$

$$H = \bar{X} + (z)s_x$$

أما إذا كان الإنحراف المعياري للمجتمع الإحصائي معرف σ_X فيجب استخدام نفس المعادلتين للحدين باستبدال s_x بـ σ_X.

مثال تطبيقي :

أجريت دراسة عام 2009 حول سن ربات البيوت في الجزائر وقد بلغ متوسط سن هؤلاء 45،20 سنة. نريد معرفة هل تغير متوسط سن ربات البيوت في الجزائر اليوم، لهذا الغرض جمعت بيانات حول سن 35 امراة متزوجة فكانت كالتالي :

الرقم	1	2	3	4	5	6	7	8	9	10
السن	25	60	45	35	32	30	47	51	52	24
الرقم	11	12	13	14	15	16	17	18	19	20
السن	19	33	46	52	50	57	43	36	30	31
الرقم	21	22	23	24	25	26	27	28	29	30
السن	50	24	29	30	67	32	27	37	32	44
الرقم	31	32	33	34	35					
السن	48	36	55	43	32					

1. هل يختلف متوسط سن ربات البيوت في الجزائر اليوم عن ما كان عليه في سن 2009 ؟
2. حدد مجال يتراوح فيه متوسط سن ربات البيوت في الجزائر بنسبة خطأ $\alpha = 0.05$ ؟

الحل :

- **كتابة الفرضيات الإحصائية:**

الفرضية الصفرية : المتوسط الحسابي للمجتمع الذي سحبت منه العينة يساوي القيمة النظرية (45.20).

$$H_0 : \mu = 45.20$$

الفرضية البديلة : المتوسط الحسابي للمجتمع الذي سحبت منه العينة يختلف عن القيمة النظرية (45.20)

$$H_0 : \mu \neq 45.20$$

- **إحصائية اختبار t:**

تكون إحصائية اختبار t المستخدمة في الاختبار كالتالي :

$$t = \frac{\bar{x} - \mu}{s_d/\sqrt{n}}$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1384}{35} = 39.54$$

$$S^2 = \frac{\sum x_i^2 - (\sum x_i)^2 / n}{n - 1} = \frac{59400 - (\frac{1384^2}{35})}{34} = 137.54$$

$$s = \sqrt{S^2} = \sqrt{137.54} = 11.72$$

دالة اختبار t لعينة واحدة هي :

$$t = \frac{\bar{x} - \mu}{s_d/\sqrt{n}} = \frac{39.54 - 45.20}{11.72/\sqrt{35}} = -\frac{5.66}{1.98} = -2.85$$

بعد حساب قيمة t نبحث في الجدول الخاص بتوزيع t عن القيمة المجدولة بدرجات حرية $df = n - 1 = 34$ وعند مستوى الخطأ $\alpha = 0.05$ فنجدها تساوي : $t_{0.05}^{34} = 2.021$.

- **القرار:**

نرفض الفرضية الصفرية لأن قيمة t المحسوبة أكبر من القيمة الحرجة 2.021 وعليه فإن الفرق الملاحظ بين متوسط العينة والقيمة الثابتة هو فرق جوهري، بمعنى ان متوسط سن ربات البيوت الجزائريات قد تراجع فعلا من 45.20 في سنة 2009 إلى مايقارب 39.54 في سنة 2010 . نقول هذا لأنه يتبقى تحديد مجال الثقة الذي يتراوح فيه المتوسط الجديد، وعليه يجب حساب مجال ثقة المتوسط بنسبة خطا قدرها 0.05.

- **تقدير مجال متوسط سن ربات البيوت خلال سنة 2010**

لقد تم حساب الخطأ المعياري S_x حيث :

$$s_x = \frac{S}{\sqrt{n}} = \frac{11.72}{\sqrt{35}} = 1.98$$

تحديد المجال :

الحد الأعلى :

$$H = \bar{X} + (z)s_x$$

الحد الأدنى :

$$L = \bar{X} - (z)s_x$$

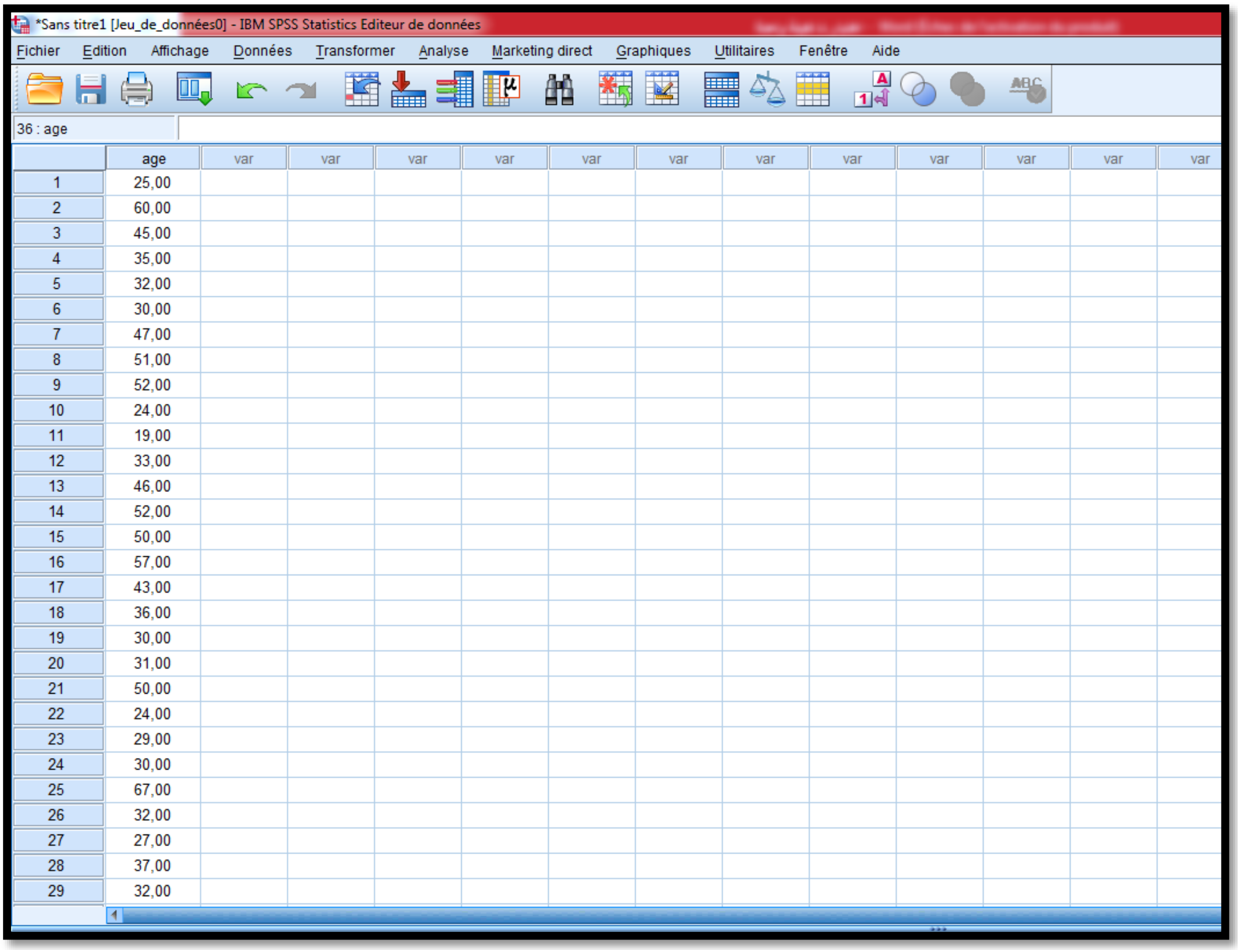
لتحديد قيمة Z نعرف أن متوسط الخطأ $\alpha = 0.05$ تقابله القيمة 1.96 .

$$H = 39.54 + (1.96)(1.98) = 43.42$$

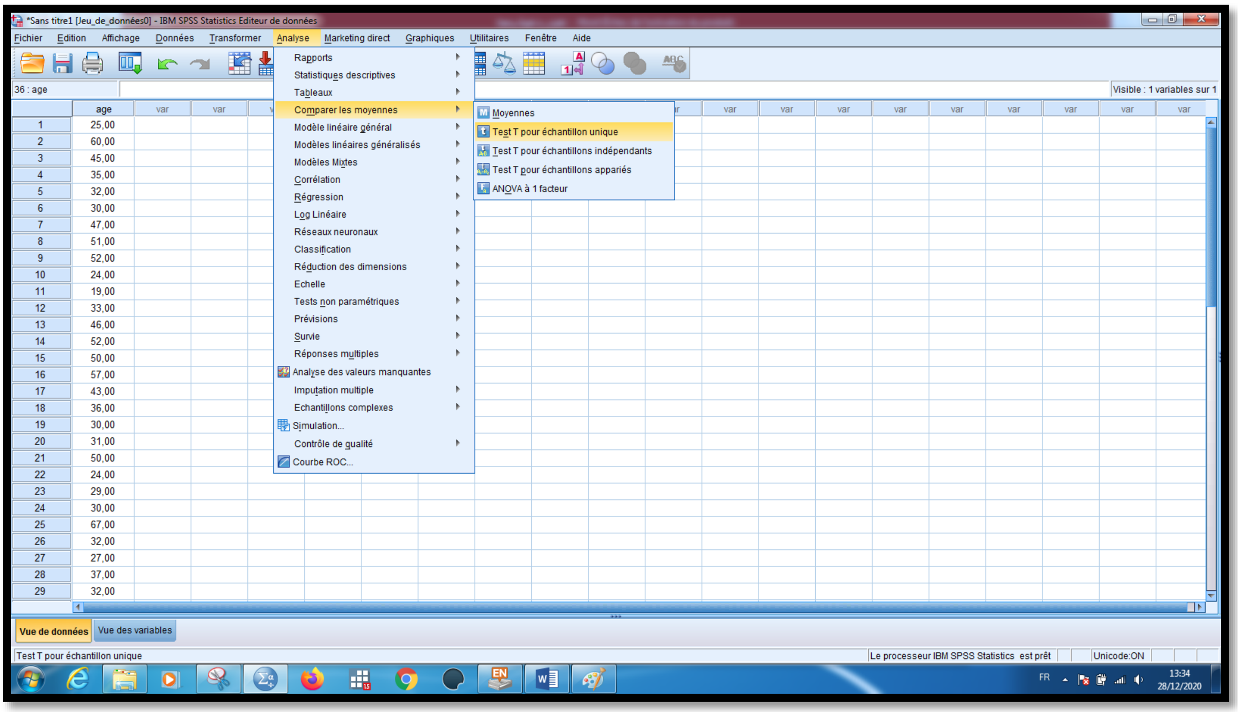
$$L = 39.54 - [(1.96)(1.98)] = 35.66$$

- حساب اختبار t لعينة واحدة ومجال الثقة باستخدام برنامج SPSS :

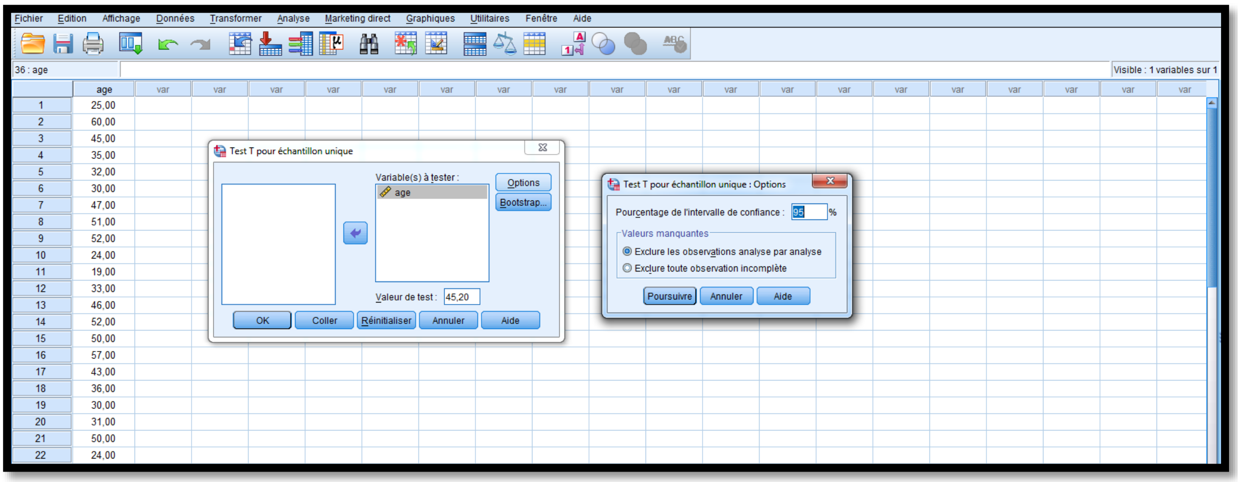
- يتم فتحج برنامج SPSS وإدخال البيانات أو فتح ملف بيانات سبق إدخاله قبل ذلك.



- نضغط على كلمة Analyse في شريط الاختيارات الأفقي فتظهر قائمة فرعية بها عدد من الاختبارات، نختار منها Comparer les moyenness فتظهر قائمة فرعية أخرى بها اختبارات t المختلفة فنختار منها اختبار t لعينة واحدة .



- بالضغط على عبارة Test T pour échantillon unique يظهر صندوق حوار اختبار t لعينة واحدة.
- ننقل المتغيرات التي نريد إجراء التحليل عليها من قائمة المتغيرات الموجودة في المستطيل على يسار صندوق الحوار إلى المستطيل الذي يحمل عنوان Variables à tester ، وقد تم إدخال متغير السن.
- نكتب في المستطيل الصغير الذي يحمل عنوان أسفل صندوق الحوار القيمة المقارن بها، وهذه القيمة في هذا المثال تساوي 45.20.
- نضغط على زر options لتحديد مجال الثقة، يمكن اختيار 95% أو 99% مع العلم ان النسبة الأولى هي النسبة الافتراضية للبرنامج.



- نضغط على الزر poursuivre ثم ok فتظهر نتائج اختبار العينة واحدة في جدولين .

النتائج:

```
T-TEST
  /TESTVAL=45.20
  /MISSING=ANALYSIS
  /VARIABLES=age
  /CRITERIA=CI(.95).
```

Test T

[Jeu_de_données0]

Statistiques sur échantillon uniques

	N	Moyenne	Ecart type	Moyenne erreur standard
age	35	39,5429	11,72314	1,98157

Test sur échantillon unique

	Valeur de test = 45.20					
					Intervalle de confiance de la différence à 95 %	
	t	ddl	Sig. (bilatéral)	Différence moyenne	Inférieur	Supérieur
age	-2,855	34	,007	-5,65714	-9,6842	-1,6301

يوضح الجدول الأول متغير السن « age» في العمود الأول، وحجم العينة في العمود الثاني (عدد الحالات) وهي 35، ويحتوي العمود الثالث على المتوسطات الحسابية، حيث يظهر أن المتوسط الحسابي لمتغير السن قد بلغ 39.5429، أم العمود الرابع فيظهر الإنحراف المعياري ، وقد بلغ في التطبيق الحالي 11.7231 ، وأما العمود الأخير فيظهر الخطأ المعياري وقد بلغ 1.9816.

أم الجدول الثاني فيحتوي على متغير السن والقيمة الإحصائية t=-2.855 ودرجات الحرية ddl التي تساوي 34 ، ثم الدلالة الغخصائية بمخرجين للإحصائية t وفي أعلى الجدول تظهر القيمة المقارن بها .(b) ، يظهر من نتائج الجدول بأن الفرق بين المتوسط الملاحظ والقيمة الثابتة 45.20 هو فرق دال إحصائيا كون الاحتمال $p<0.01$ ، ويظهر في العمود الخامس قيمة الفرق بين المتوسط الملاحظ والدرجة الثابتة وقد بلغ -5.6571 ، أما العمود الأخير الذي ينقسم إلى عمودين فإنه يحتوي على درجات الثقة للفروق بين المتوسط والدرجة الثابتة عند مستوى 95% حيث يشمل الحدين : الأعلى والأدنى، وحيث يقع الفرق بنسبة 95% بين هذين الحدين:

$H= a-1.9816$ و $L=a-9.6842$