المحاضرة النظرية السادسة
في الإحصاء

إعداد:

د. حيدر الحسن

15/11/2020

تعديل السلاسل الزمنية TIME SERIES

نظراً لأنَّ السلاسل الزمنية تتعرّض لعدد من المؤثّرات الخارجية لذا يلزم إجراء تعديل للسلاسل الزمنية. من هذه التعديلات نورد ما يلي:

1. **التعديل البسيط:** في هذا النوع من التعديل يتمّ تقسيم السلسلة إلى عدّة أقسام أو فترات متساوية ويحسب المتوسط الحسابي لكل فترة، ومن ثمَّ يتمّ دراسة تطوّر السلسلة وتحديد الاتجاه العام لها.

مثـــال: لدينا السلسلة الزمنية التالية:

العام	الصادرات	المجموع لثلاث سنوات	المتوسط لثلاث سنوات
1985	40		
1986	45	126	42
1987	41		
1988	50		
1989	45	150	50
1990	55		
1991	51		
1992	60	179	58.6
1993	65		
1994	70		
1995	75	219	73
1996	74		

نلاحظ من هذا الجدول أنَّ المتوسط الحسابي يتزايد مع الزمن وهذا يدلّ على أنَّ الاتجاه العام تصاعدي.

طريقة المتوسط المنزلق: تُستخدم هذه الطريقة عندما يكون تقسيم السلسلة إلى عدة أقسام أو فترات متساوية غير ممكن، وبالتالي نلجأ إلى طريقة المتوسط المنزلق وهي تتمّ من خلال حساب المتوسط الحسابي لفترة ما ولتكن (3-6) سنوات، ثمَّ نحذف العنصر الأول من القسم الأول أو الفترة الاولى ونضيف عنصر جديد من السلسلة وهكذا ..أخيرا يتم دراسة تطوّر السلسلة وتحديد الاتجاه العام لها.

- **مثـــال: لدينا السلسلة الزمنية التالية:**

نلاحظ من الجدول أعلاه أنَّ المتوسط الحسابي يتقلب بين الزيادة والنقصان مع الزمن، وهذا يدلّ على أنَّ الاتجاه العام يتخذ شكلاً منحنياً ومتذبذباً.

العام	الصادرات	المجموع لثلاث سنوات	المتوسط لثلاث سنوات
1988	3500		
1989	4000	11600	3866.6
1990	4100	12300	4100
1991	4200	12200	4066.6
1992	3900	12400	4133.3
1993	4300	12200	4066.6
1994	4000	12800	4266.6
1995	4500		

التحليل الاقتصادي للسلاسل الزمنية TIME SERIES :

يعتمد التحليل الاقتصادي للسلاسل الزمنية على **معرفة القوى المؤثّرة فيها**، فالسلاسل الزمنية تخضع لتغيرات دورية شبه منتظمة تعكس تأثير عوامل مختلفة. **وتصنف القوى المؤثرة في السلاسل الزمنية بشكل عام ضمن الفئات التالية:**

1- قوى الاتجاه العام

2- التغيرات الموسمية

3- التغيرات الدورية

4- القوى العشوائية أو العرضية.

1. - **الاتجاه العام:** إنَّ استخراج معادلة الاتجاه العام يماثل حساب معاملات معادلة الانحدار البسيط من الشكل التالي:

$$Y = a + b*X$$

حيث أنَّ : Y القيمة النظرية للصفة المختارة للظاهرة المدروسة

a ثابت المعادلة الممثّلة للسلسلة الزمنية

b عامل المعادلة الممثّلة للسلسلة الزمنية

X العامل المؤثّر (أي المتغيّر)

- **أمَّا كيفية حساب كل من a ثابت المعادلة الممثّلة للسلسلة الزمنية وb عامل المعادلة الممثَّلة للسلسلة الزمنية فهي على الشكل التالي:**
- حساب a ثابت المعادلة الممثّلة للسلسلة الزمنية باستخدام العلاقة الرياضية التالية:

$$a = \bar{y} - b * \bar{x}$$

حساب b عامل المعادلة الممثّلة للسلسلة الزمنية باستخدام العلاقة الرياضية التالية:

$$b = \sum (x - \bar{x}) * ((y - \bar{y}) / \sum (x - \bar{x}))^2$$

كما ويمكن حساب b عامل لمعادلة الممثلة للسلسلة الزمنية باستخدام العلاقة الرياضية التالية:

$$b = (\sum x * y - \sum x * \sum x / n) / \sum x^2 - (\sum x)^2 / n$$

العينات Samples

نلجأ إلى بحث العينات هذا والتي تعدّ أسلوباً من الأساليب الرئيسة في تحليل معطيات المراقبة الإحصائية، وفي السنوات الأخيرة تقدّمت البحوث الإحصائية تقدّماً كبيراً نتيجة تطبيق نظرية العينات على هذه البحوث

فعند دراسة المجتمع يتحتّم علينا أن نكتفي بفحص جزء منه، فالذي يريد أن يأخذ كمّية من الدواء لا يتفحّص كل قطرة منه، ولكنه يفحص كمّية صغيرة من هذا الدواء

واختبار جزء من المجتمع للدراسة يُعرف **بالمعاينة** والجزء المختار يعرف **بالعينة**

وتعدّ طريقة المعاينة العشوائية أبسطّ وأدقّ طريقة للمعاينة

حيث أنَّ اختيار أفراد العينة يُبنى على محض الصدفة، ويشترط في العينة المختارة من المجتمع أن تتّصف بصفتين:

1- أن تكون متّسقة: ويعني ذلك أن الوسط الحسابي للعينة ($\overline{X}$) يقترب من الوسط الحسابي للمجتمع (M) ويتحقق ذلك كلّما كبر حجم العينة.

2- أن تكون غير متميّزة: فإذا أخذنا من المجتمع عدّة عيّنات فإن الوسط الحسابي لأوساط هذه العينات يكون مساوياً للوسط الحسابي للمجتمع (M).

اختبارات المعنوية:

- لكي نتمكّن من تعميم نتائج البحث يجب أن نتأكّد **من صحّتها**، ويتم ذلك بوضع النظرية الفرضية لاختبار مدى الاعتماد على الإحصاءات المقدّرة من التجربة أو العيّنة ثم تختبر إحصائياً وبنتيجة هذه الاختبارات تقبل النظرية الفرضية أو ترفض.
- وللحكم على صحّة النظرية الفرضية من عدمها يجري اختبار المعنوية لمحاولة معرفة هل الفروق بين العيّنات أو المعاملات هي فروق عشوائية بين عيّنات المجتمع الواحد، أم أنَّ الفروق بين العيّنات أو المعاملات هي في الحقيقة فروق أساسية ناتجة عن تأثير المعاملات المختلفة، وأنَّ تلك العينات مأخوذة من مجتمعين مختلفين، **وفي الحالة الأخيرة ترفض النظرية الفرضية.**
- فعند فحص النظرية الفرضية (H_1)، **نقول أنَّ هذه الفرضية خطأ إذا كانت قيمة فحص الاحتمال المحسوبة أقل من مستوى الدلالة المفروض، وإن النتيجة ذات دلالة إحصائية، أمَّا إذا كانت قيمة الاحتمال المحسوبة أكبر من مستوى الدلالة المفروض فإننا نقول بصحّة النظرية الفرضية، وإنَّ النتيجة ليست بذات دلالة.**
- والمتّبع في الإحصاء غالباً مستويان للدلالة **5% و 1% .**

اختبار مربّع كاي Chi Square (X^2):

- تقسم البيانات التي يحصل عليها الباحث من التجارب إلى قسمين:

أ- القياسات Measurement Data:

تعبّر القياسات عن البيانات التي يحصل عليها الباحث عن طريق قياس أفراد المتغيّر العشوائي لصفة ما كالطول أو الوزن أو كمّية المحصول، ولقد تكلّمنا عن طريقة جمع البيانات وكيفية تبويبها ثمَّ تحليلها في الفصول السابقة.

ب- التعدادات Enumeration data:

ويعبّر التعداد عن البيانات التي يحصل عليها الباحث عن طريق تسجيل عدد الأفراد أو عدد القياسات أو التكرارات التي تقع في قسم أو فئة معيّنة

ذلك كما في حالة عمل جداول التوزيع التكراري، حيث يقوم الباحث بتقسيم الصفة المدروسة إلى فئات أو أقسام أو مجموعات يقع كل منها داخل مدى معيّن، ثم حصر عدد الأفراد أو تكرار الأفراد الذي يقع في كل فئة أو قسم منها،

- **سنتناول في هذا القسم طريقة اختبارات البيانات العددية**
- يعدّ الإحصاء المسمّى اختبار **مربع كاي لحسن المطابقة أو اختبار التطابق النسبي** من أهم الطرق التي تستعمل في مقارنة مجموعة من النتائج المشاهدة أو المتحصّل عليها من تجربة حقيقية بمجموعة أخرى فرضية وضعت على أساس النظرية الفرضية التي يُراد اختبارها
- تعتمد هذه الطريقة على افتراض وجود عيّنة عشوائية بها عدد N من الأفراد، قسّمت إلى عدد من الفئات المتشابهة بحيث يقع كل فرد في العينة في إحدى هذه الفئات
- ثمَّ مقارنة التكرارات المشاهدة بالتكرارات النظرية بقصد معرفة مدى انطباق التكرارات المشاهدة على تلك النظرية، وذلك باستعمال **اختبار مربع كاي**.

اختبار المعنوية بوساطة مربع كاي:

يستخدم اختبار مربع كاي أحياناً لاختبار حسن المطابقة والتوفيق بين البيانات التي تحصلنا عليها وبين منحنى أو معادلة معيّنة

فإذا كان لدينا بيانات والمطلوب اختبارها لمعرفة فيما إذا كانت هذه البيانات تتّفق مع معادلة وراثية معروفة فإنَّه يمكن حساب قيمة **مربع كاي** من المعادلة:

$$v^2 = \sum \frac{(O-E)^2}{E}$$

حيث: O: التكرار المشاهد أو القيم النظرية.

E: التكرار المتوقّع.

وتدلّ هذه المعادلة على أنَّ **مربع كاي عبارة عن مجموع مربعات الفروق بين التكرارات المشاهدة مقسوماً على التكرارات المتوقعة.**

<u>شروط إجراء اختبار مربع كاي</u>

1-عدد مشاهدات العينة أكبر من 50

2- التكرار المتوقع المناظر لكل فئة لا يقل عن5

أمثلة:

- **أ- مربع كاي لدرجة حريَّة واحدة:**

تعتمد المقارنة هنا على مجموعتين أو فئتين فقط، ولذلك تكون درجات الحرية في هذا النوع من المقارنة عبارة عن درجة حرية واحدة فقط.

- **مثال:**

قام أحد الأشخاص باختبار قدرة أحد المبيدات الحشرية على قتل نوع معيَّن من الحشرات، وكانت التعليمات المرفقة بهذا المبيد تفيد بأن المبيد يمكنه أن يقتل 80% من عدد الحشرات بعد رشّها به

ولاختبار ما إذا كانت هذه التعليمات تنطبق على الظروف العملية أو المشاهدة بالحقل، أجريت تجربة حقيقية رشّ فيها المبيد على النباتات التي تحوي عدداً معروفاً من الحشرات، ثم حصرت عدد الحشرات الحيّة والميتة

فكانت النتائج كالتالي: حشرات حيّة عددها 100، حشرات ميتة وعددها 300، فهل تُشير هذه النتائج إلى صحة التعليمات المرفقة بالمبيد ؟

- الحل:

لحل هذه المسألة نضع النظرية الفرضية التي تقول بصحة التعليمات المرفقة بالمبيد، أي أنه يقتل 80% من الحشرات ثم تقوم بحساب التكرارات النظرية والمتوقعة كالتالي:

- عدد الحشرات الحيّة بعد استعمال المبيد $400.\frac{20}{100}=80$

- عدد الحشرات الميتة بعد استعمال المبيد $400.\frac{80}{100}=320$

المشاهدات	التكرارات المتوقعة O	التكرارات النظرية E	$\frac{(O-E)^2}{E}$
الحشرات الحية	100	80	5
الحشرات الميتة	300	320	1.25
المجموع	400	400	6.25

بالنظر إلى جدول K^2 عند احتمال 5% ودرجة حرية 1 نجد أنّ تساوي كاي الجدولية 3.84، **وبما أنَّ قيمة معامل كاي مربع أكبر من الجدولية، لذا ترفض النظرية الفرضية،** أي أن النسبة المشاهدة في الحقل لا تنطبق على النظرية الموجودة في التعليمات المرفقة مع المبيد (80%).

ب- إختبار مربع كاي في حال درجات الحرية المتعدّدة:

مثال2:

حصل باحث في مجال وراثة النباتات على البيانات التالية: 32 – 108 – 101 – 315 وافترض أنَّها تتبع نسبة الأشكال المظهرية: 1 :3 :3 :9، فهل تتّفق هذه البيانات مع النظرية المذكورة ؟

نوجد القيم المتوقعة بضرب المجموع الكلي في النسبة الخاصة بكل نوع على الترتيب.

المشاهدة O	المتوقع E	O - E	$\frac{(O-E)^2}{E}$
315	312.8	2.2	0.02
101	104.3	-3.3	0.10
108	104.3	3.7	0.13
32	34	-2.6	0.20
556	556	0	0.45

إذاً $K^2 = 0.45$ ، درجة الحرية $3 = 1 - 4$.

من الجدول نجد أنَّ كاي الجدولية عند 5% = 7.81

نجد أنَّ K^2 المحسوبة أقل من K^2 الجدولية،

ولذلك فهي غير معنوية، أي أنَّ قيمة K^2 التي حصلنا عليها من هذه البيانات لا تسمح برفض النظرية الفرضية، ونستنتج أن البيانات تتوافق مع النسبة 1 :3 :3 :9 .

اختبار العلاقة بين صفتين:

- إذا قمنا بفحص كل فرد في المجتمع لخاصتين معينتين، وكل خاصَّة صنّفت إلى مجموعات فإنّنا قد نرغب في معرفة إذا كانت هذه الخصائص مستقلّة بعضها عن بعضها الآخر
- فمثلاً إذا كان لدينا عدد من النباتات في خط طويل، وتمّت دراستها من حيث لون الأوراق ومعدّل النمو،
- يمكننا وضع النظرية الفرضية الدالّة على أنَّ لون الأوراق ومعدّل النمو مستقلان عن بعضهما الآخر ومن ثم نُخضع هذه النظرية للاختبار
- وبالمثل إذا كان لدينا مجموعة من الطلاب وقسّمناهم حسب لون الشعر ولون العيون، ونريد أن نعرف إن كانت هناك علاقة ما بين لوني الشعر والعيون، ففي هذه الحالات جميعها يمكننا الاستفادة من اختبار مربع كاي لإجراء الاختبار.

مثال:

- الجدول الآتي يبيّن توزيع 250 بادرة من بادرات الفول البلدي حسب معدّل النمو ولون الأوراق، باستخدام هذه البيانات اختبر إمكانية وجود علاقة بين الصفتين.

لون الأوراق	معدَّل النمو			المجموع
	جيد	مقبول	ضعيف	
أخضر	55	79	4	138
أخضـــــر مصفر	11	60	15	86
أصفر	1	6	19	26
المجموع	67	145	38	250

والاتحادات المتوقعة من هذه البيانات إذا كانت الصفتان مستقلتين هي:

أخضر جيد : 37 = 250 / (138 . 167)

أخضر مقبول: 80 = 250 / (138 . 145)

أخضر /ضعيف: 21 = 250 / (138 . 38)

أخضر مصفر/جيد: 23 = 250/ (86 . 67)

أخضر مصفر/مقبول: .49.9 = 250/ (86 . 145)

أخضر مصفر/ضعيف: 13.1 = 250 / (86 . 38) .

أصفر جيد: 7 =250 / (26 . 67)

أصفر مقبول: 15.1 = 250/ (26 . 145)

أصفر ضعيف: 3.9 = 250 / (26 . 38) .

نضع البيانات الآتية في جدول:

لون الأوراق	معدل النمو			المجموع
	جيد	مقبول	ضعيف	
أخضر	37	80	21	138
أخضر - مصفر	23	49.9	13.1	86
أصفر	7	15.1	3.9	26
المجموع	67	145	38	250

عدد الصفوف في الجدول (r) = 3

وعدد الأعمدة في الجدول C = 3

ودرجات الحرية = $(r-1)\ (C-1)$

$2.\ 2 = 4$

مربع كاي المحسوب:

$$K^2=\frac{(55-37)^2}{37}+\frac{(79-80)^2}{80}+\frac{(4-21)^2}{21}+\frac{(11-23)^2}{23}+\frac{(60-49.9)^2}{49.9}+$$

$$\frac{(15-13.1)^2}{13.1}+\frac{(1-7)^2}{7}+\frac{(6-15.1)^2}{15.1}+\frac{(19-3.9)^2}{3.9}=\frac{(18)^2}{37}+\frac{(1)^2}{80}+\frac{(17)^2}{21}+$$

$$\frac{(12)^2}{23}+\frac{(10.1)^2}{49.9}+\frac{(1.9)^2}{13.1}+\frac{(6)^2}{7}+\frac{(9.1)^2}{15.1}+\frac{(15.1)^2}{3.9}=100.2$$

ثم نستخرج قيمة K^2 من الجدول عند درجة حرية 4 ومعنوية 5%، وهي 9.48،

نجد أنَّ K^2 المحسوبة أكبر من K^2 الجدولية

فمعنى ذلك نرفض النظرية الفرضية القائلة باستقلال العينتين بمستوى معنوية 5%

أي نستنتج أنَّ هناك علاقة بين الصفتين.

- **الخطوات الأساسية لبحث العينة:**
- عند إجراء البحوث الإحصائية بوساطة العيّنات، على الباحث أن يكون ملمّاً بأهمية البيانات المطلوبة، وكيفية استخدامها، والخطوات المتّخذة عند تصميم بحث العينة:

1- تعريف المشكلة، وضع الأسئلة التي يجب أن نجد لها أجوبةً من خلال البحث.

2- تحديد المجتمع الإحصائي المراد معاينته والمفردات الداخلة فيه.

3- نتائج الأبحاث السابقة لهذه المشكلة.

4- تحديد البيانات المطلوبة، هل هي عن الحبوب، أو الأبقار، أو الأغنام.

5- تحديد طريقة جمع البيانات: الاتّصال المباشر أو غير المباشر، أو كليهما أو الطريقة الميدانية، أو السجلاّت الرسمية.

6- تحديد وحدة القياس: هل هي الهكتار، أو الفدان، أو الدونم، والكغ... إلخ.

7- تكوين إطار العينة: مثلاً تحديد المساحة المزروعة بالقمح في القطر العربي السوري يكون إطار العينة هو أراضي القطر العربي السوري المزروعة بالقمح كافة.

8- اختيار وحدة المعاينة: الإنسان، أو الأسرة، .... 9- اختيار نوع العيّنة: سنتكلّم عنها لاحقاً.

10- تهيئة العناصر البشرية والمادية، مثل تدريب العدادين والباحثين، وغير ذلك.

رابعاً: دراسة العينات الكبيرة Large Samples:

ذكرنا سابقاً أنّه يشترط في العيّنة المختارة من المجتمع أن تتّصف بصفتين: متّسقة، وغير متحيّزة.

فإذا كان عدد المتغيّرات المأخوذة لعيّنة من مجتمع يساوي أو أكبر من 30 متغيّر فإن الانحراف المعياري للعيّنة $s_{\bar{x}}$ يساوي الانحراف المعياري للمجتمع σ

وتسمَّى هذه بنظرية العيّنات الكبيرة، أمَّا إذا كان عد المتغيّرات في العينة المدروسة أقل من 30 متغيّراً فعند ذلك نتّبع نظرية العيّنات الصغيرة وتكون قيمة $s_{\bar{x}}$

توزيع الأوساط الحسابية للعيّنات:

- نفرض أننا نأخذ من مجتمع ما موزّعاً طبيعياً عدداً من العيّنات، عدد المتغيّرات كل منها Nمتغيّراً (30 < N)، فإنَّ الأوساط الحسابية لهذه العيّنات هي:

$$(\bar{x}_n, \ldots\ldots, \bar{x}_3, \bar{x}_2, \bar{x}_1)$$

وانحرافاتها المعيارية هي:

$$(S\bar{x}_{12}, \ldots\ldots, S\bar{x}_2, S\bar{x}_1)$$

- وهي متساوية وتساوي **S** .
- هذه الأوساط الحسابيةً للعيّنات تختلف بعضها عن بعضها الآخر، ولذلك تتوزّع إحصائياً وفق نظام معيّن يُطلق عليه اسم **توزيع الأوساط الحسابية**، ولهذا التوزيع الإحصائي وسط حسابي $\bar{x}_n$ وانحراف معياري $S\bar{x}$
- ولقد وجد أنَّ قيمة الوسط الحسابي لهذه الأوساط الحسابية يساوي **الوسط الحسابي للمجتمع**، أي أنَّ:

$$\bar{x}_m = M$$

- وأن الانحراف المعياري لها : $S\bar{x} = \dfrac{\sigma}{\sqrt{n}}$

حيث: $S\bar{x}$ الانحراف المعياري للعينة .

σ : الانحراف المعياري للمجتمع.

n : عدد أفراد كل عينة من العينات

- وبذلك يمكن تطبيق خصائص المنحني الطبيعي، حيث تمثّل الأوساط الحسابية للعيّنات على المحور الأفقي وتكراراتها على المحور العمودي، وبالتالي يمكن قياس المساحة **Z** وحساب الاحتمال، حيث أن:

$$Z = \frac{\bar{x} - M}{S\bar{x}}$$

حيث: $\bar{x}$ هو الوسط الحسابي للعينة .

M الوسط الحسابي للمجتمع .

$S\bar{x}$ الانحراف المعياري للعينة.

مثال:

- وجد أنَّ أطوال جنود أحد المعسكرات البالغ عددهم 1000 جندي موزعة توزيعاً طبيعياً بوسط حسابي قدره 168.2 سم بانحراف معياري قدره 6.07 سم،
- احسب الاحتمال لأن يكون الوسط الحسابي لأطوال 100 جندي منهم أكثر من 169.9سم.
- الحل:

$$\bar{x}_m = M = 168.2$$

$$S\bar{x} = \frac{\sigma}{\sqrt{N}} = \frac{6.07}{\sqrt{100}} = 0.607$$

- والاحتمال المطلوب هو المساحة تحت منحني الاحتمال الطبيعي الواقعة إلى يمين Z :

$$Z = \frac{\bar{x} - M}{S\bar{x}} \Rightarrow \frac{169.9 - 168.2}{0.607} = 2.81$$

ومن جدول Z نجد أن المساحة المقابلة لـ $\mathbf{Z}$ = 0.4975 .

إذاً الاحتمال المطلوب: 0.5 – 0.4975 = 0.0025 .

أي أن احتمال الوسط الحسابي لعينة مؤلفة من 100 جندي أن يتجاوز 169.9سم هو 0.0025 أي 0.25%

حدود الثقة Confidence Limits :

- إنّه لمجال الشك أو لفترة الثقة حدين أعلى وأدنى، وفي التالي جدول يبين العلاقة بين مجال الثقة وقيمة Z المقابلة:
- %50 ، %68 ، %80 ، %95 ، %95.45 ، %96 ، %98 ، %99 ، %99.73: **مجال الثقة**
- Z : 3 – 2.58 – 2.33 – 2.05 – 2 – 1.96 – 1.645 – 1.28 – 1 – 0.65
- الوسط الحسابي لعينة مؤلفة من 64 متغيّراً هو 60 بانحراف حيادي قدره /3/ أوجد حدود الثقة %98 للوسط الحسابي للمجتمع .

$$S\bar{x} = \frac{\sigma}{\sqrt{N}} = \frac{3}{\sqrt{64}} = 0.375$$

- ومن الجدول إن قيمة Z المناظرة للمساحة 0.49 = 0.98/2 على جانبي خط التناظر هي =Z 2.33 ، وبالتعويض بالمعادلة:

$$Z = \frac{|x - M|}{S\bar{x}}$$

$$2.33 = \frac{|60 - M|}{0.375}$$

ومنه : $\boldsymbol{m = 6 \pm 0.87}$

أي أنه باحتمال 0.98 يقع الوسط الحسابي لهذا المجتمع بين (60.87 – 59.13).

نظرية العينات الصغيرة (توزيع ستودنت):
Small Sampling Theory, Student's Distribution

- درس ستودنت هذا الموضوع وتوصّل إلى معرفة أنَّه لا يمكن تطبيق الجدول المبني على المنحني الطبيعي Z لتحليل البيانات التي نحصل عليها من العينات الصغيرة (أقل من 30)
- لأنه كلما قلّ عدد الأفراد في العينة، كلّما كبر الفرق بين متوسط العينة ومتوسط المجموع الحقيقي، مما ينتج عنه زيادة فرصة الوقوع في استنتاج خاطئ من النتائج نتيجة استعمال جداول الاحتمالات.
- وعند عدم معرفة الانحراف المعياري للمجتمع (σ) يمكن استبداله بالانحراف المعياري للعينة $S\bar{x}$ وذلك عندما يكون عدد المتغيّرات في العينة أكثر من 30، أما عندما يكون عدد المتغيّرات N في العينة أقل من 30 فيلجأ إلى تقدير الانحراف المعياري للمجتمع من الانحراف المعياري للعينة من المعادلة:

$$\sigma = S\bar{x}.\sqrt{\frac{N}{N-1}}$$

- حيث : σ: أحسن تقدير للانحراف المعياري للمجتمع .
- $S\bar{x}$: الانحراف المعياري للعينة.
- N: عدد المتغيّرات في العينة.

أولاً: مقارنة متوسطة عينة بمتوسط المجتمع الذي أخذت منه في هذه تكون:

$$t = \frac{\bar{x} - M}{Sd'}$$

$$Sd' = \frac{\sigma}{\sqrt{N-1}}$$

- حيث: $\bar{x}$: المتوسط الحسابي للعينة .
- M : الوسط الحسابي للمجتمع .
- Sd' : الخطأ المعياري .
- تستخرج قيمة t من جدول المقابلة لدرجة الحرية N-1 الموجودة في التجربة عند احتمال 0.05 و 0.01 .

- إذا كانت قيمة t المحسوبة أكبر من قيمة t الجدولية عند احتمال 5% و1% دلّ أن الفرق **مؤكد جداً.**
- أما إذا كانت t المحسوبة أكبر من t عند مستوى 5% وأصغر من t عند 1% دلّ على أن **الفرق مؤكد وليس راجعاً للصدفة،**
- **وفي كلتا الحالتين ترفض النظرية الفرضية القائلة:** أنَّه لا يوجد فرق بين الأفراد والمجموعات.
- أمَّا إذا كانت قيمة t المحسوبة أصغر من t الجدولية عند المستويين 1% و 5% فإنَّ ذلك يدل **على أن الفرق غير مؤكّد (معنوي)، وتقبل النظرية الفرضية.**
- من المعادلة يمكن أن نحدد الحدين الأعلى والأدنى:

$$\bar{x} - M = \pm t S'd$$

$$M = \bar{x} \pm t S'd$$

- نستخرج t عند المستوى 5% وعند المستوى 1% .

- **مثال:**

أخذ من مجتمع طبيعي ذي وسط حسابي 30 عينة مؤلّفة من 16متغيّراً، وجد أن انحرافها المعياري هو 3 وكان انحراف وسطها عن الوسط الحسابي للمجتمع هو 2 فهل يعد هذا الانحراف ذا دلالة:

- **الحل:** توضع النظرية الفرضية وهي أنَّ انحراف الوسط الحسابي للعينة عن الوسط الحسابي للمجتمع ليس بذات دلالة.

$$T = \frac{\bar{x} - M}{S'd}$$

$$S'd = \frac{\sigma}{\sqrt{N-1}} = \frac{3}{\sqrt{16-1}} = 0.78$$

$$T = \frac{2}{0.78} = 2.56$$

من الجدول t بدرجة حرية N-1 =15 والمستوى 0.05 هي 2.131 نجد أن t المحسوبة أكبر من t الجدولية عند مستوي 5% **فالنتيجة ذات دلالة**، أمّا قيمة t الجدولية عند مستوى 1% هي 2.947، نجد أنَّ t المحسوبة أصغر من t الجدولية عند مستوى 1% **فالنتيجة ليست بذات دلالة بل يرجع للصدفة.**

اختبار القيم الشاذة وحذفها:

- كثيرا ما ينتج لدينا في تجربة ما بين قيم العينة قيم شاذة ومتطرفة تؤثّر سلباً على التحليل الإحصائي ولابد من إزالتها.
- **والمهم هنا هو كيفية التعرف على القيم الشاذة والمتطرفة.** هذا ما سنشرحه فيما يلي :
- **يستخدم اختبار t-test لهذه الغاية حيث أننا نستخدم احصاء الاختبار التالي:**

$$t = x_0 - \bar{x} / s$$

حيث أنَّ: x_0 القيمة المراد اختبارها

s الانحراف المعياري للعينة

`x المتوسط الحسابي للعينة

- وتكون الفرضيات الإحصائية كما يلي:
- الفرضية الابتدائية: $H0: x_0 = x`$
- الفرضية البديلة: $H1: x_0 \neq x`$

- مثـــــــال: لدينا العينة التالية:
- $X = \{200, 300,250,400,400,500,550\}$
- هل يمكن أن نعتبر القيمة 550 قيمة شاذة أم لا؟
- نطبق العلاقة السابقة الذكر لحساب قيمة احصاء الاختبار t – test فنحصل على النتائج كما يلي:
- $Sx` = 122.93$
- $t = 550 – 383.75 / 122.93$
- $t = 1.35$

$$t = x_0 - \bar{x} / s$$

- نخرج قيمة t_{tab} الجدولية عند مستوى المعنوية 10% ودرجة حرية 7 فنجد أنَّ:
- $t_{tab} = 1.89$
- الآن نقارن قيمة مع قيمة t_{cct} المحسوبة من العلاقة السابقة كما هو وارد في الجدول السابق أعلاه حيث أن $t_{clt} = 1.35$
- نجد أن t_{clt} المحسوبة أصغر من قيمة t_{tab} الجدولية وبالتالي نقبل الفرضية الابتدائية أي أن القيمة المعنية 550 ليست قيمة شاذة لأنها لا تختلف عن قيمة المتوسط بشكل معنوي والفرق غير حقيقي وغير مؤكد إحصائيا.

إلى اللقاء في المحاضرة القادمة