العينات والاختبارات

SAMPLES AND TESTS

6-1- مقدمة: إن هدف الاختبار الإحصائي هو اختبار فرضية تتعلق إما بمؤشرات إحصائية أساسية (أو ما يسمى أحيانا وسيط) أو أنها تتعلق بالتوزيع المفترض لمجتمع ما.

فكثيرا من الأحيان يتطلب الأمر معرفة قانون توزيع مجتمع ما غير معروف فعلا وبالتالي لابد من الافتراض بان هذا المجتمع يتبع أحد التوزيعات المعروفة ومن ثم يتم اختبار هذه الفرضية لمعرفة مدى صحتها أي لرفضها أو قبولها.

وقد يكون العكس أحيانا حيث يتطلب الأمر معرفة المؤشرات الإحصائية ألأساسية لمجتمع ما مثل المتوسط الحسابي أو الانحراف المعياري وما شابه ذلك، وبالتالي لابد من وضع الفرضيات التي تفترض قيما افتراضية للمؤشرات الإحصائية ومن ثم يتم اختبار هذه الفرضيات لمعرفة مدى صحتها أي لرفضها أو قبولها.

6-2- الاختبار الإحصائي : يتضمن الاختبار الإحصائي العناصر الرئيسية التالية:

1. الفرضية الابتدائية ويرمز لها بـ :H0
2. منطقة الرفض والقبول
3. الفرضية البديلة ويرمز لها بـ :H1
4. احصاء الاختبار

إن العناصر الرئيسية الأربع المذكورة هي ضرورية لكل اختبار إحصائي وأي تغير في أحد من هذه العناصر الأربع يؤدي إلى اختبار جديد.

وفيما يلي سندرس بعض هذه العناصر الأربع بشيء من التفصيل :

1. الفرضية الابتدائية (:H0) : وهي الفرضية التي سيجري اختبار مدى صحتها أي لرفضها أو قبولها .وهذه الفرضية تنص تتضمن بشكل عام قيم افتراضية لوسيط ما للمجتمع أو العينة المدروسة؛مثل الافتراض بان هذا المجتمع يتبع أحد التوزيعات المعروفة أو تفترض قيما افتراضية للمؤشرات الإحصائية كان تفترض أن قيمة المتوسط μ يساوي قيمة افتراضية μ_0 وهكذا...

إن اتخاذ القرار للتأكد من صحة الفرضية الابتدائية أي لرفضها أو قبولها يعتمد أساسا على المعلومات التي تحويها العينة أو العينات المسحوبة بطريقة عشوائية من المجتمع المدروس.

2. منطقة الرفض والقبول :تقسم مجموعة كل القيم المأخوذة التي يمكن أن يأخذها احصاء الاختبار إلى مجموعتين أو منطقتين ندعو أحداهما منطقة الرفض والأخرى منطقة القبول. وان الرفض أو القبول يتوقف على أين تقع القيمة التي يأخذها احصاء الاختبار ، فإذا وقعت القيمة التي يأخذها احصاء الاختبار ،محسوبا من العينة التي بين أيدينا ،في منطقة الرفض ترفض الفرضية الابتدائية ونقبل الفرضية البديلة.أما وقعت القيمة التي يأخذها احصاء الاختبار ،محسوبا من العينة التي بين أيدينا ،في منطقة القبول نقبل الفرضية الابتدائية و ترفض الفرضية البديلة.

يخضع اتخاذ القرار للتأكد من صحة الفرضية الابتدائية ، أي لرفضها أو قبولها، إلى نوعين من الخطأ وهما التاليين:

1. الخطأ من النوع الأول ERROR TYPE I
2. الخطأ من النوع الثاني ERROR TYPE II

وينص الخطأ من النوع الأول ERROR TYPE I إلى أنه يمكن أن نرفض الفرضية الابتدائية بينما هي في الواقع صحيحة .أما الخطأ من النوع الثاني ERROR TYPE II فيحصل عندما نقبل الفرضية الابتدائية بينما هي في الواقع خاطئة.

والجدول التالي يبن الحالات السابقة عند اتخاذ القرار للتأكد من صحة الفرضية الابتدائية أو عدم صحتها :

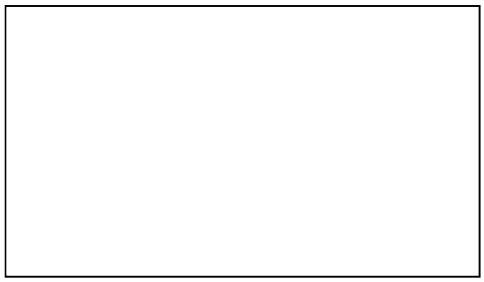
القرار	الفرضية الابتدائية	
	صحيحة	خاطئة.
الرفض	الخطأ من النوع الأول	قرار صحيح
القبول	قرار صحيح	الخطأ من النوع الثاني

وعادة نقيس جودة الاختبار الإحصائي باحتمالي الخطأ من النوع الأول و الخطأ من النوع الثاني ،ونرمز لهما بـ α و β على الترتيب.ويمكن التعبير عن الاحتمال α بأنه أحتما أن يقع احصاء الاختبار في منطقة الرفض علما بأن الفرضية الابتدائية صحيحة .ويدعى α أيضا بمستوى الأهمية أو المعنوية .ومن الواضح أن زيادة حجم منطقة الرفض سيزيد من

قيمة α وفي نفس الوقت فانه يؤدي إلى تناقص قيمة β وذلك من أجل حجم ثابت للعينة n .أما تخفيض منطقة الرفض فسيزيد من قيمة β وفي نفس الوقت فانه يؤدي إلى تناقص قيمة α.وعند زيادة حجم العينة n فانه من الطبيعي أن توفر لنا العينة الأكبر قدرا لأكبر من المعلومات نتخذ على ضوئها قرارنا وهذه ستؤدي إلى تناقص كل من α و β .

يمكن أن نقسم الفرضية الابتدائية إلى نوعين :

- الفرضية الابتدائية أحادية الجانب (البسيطة):نقول عن فرضية ابتدائية أنها بسيطة إذا كانت تختبر قيمة واحدة فقط بدون احتمالات أخرى .كمثال على ذلك نقول : H0: a > 50 ، ففي هذه الحالة لا يوجد سوى قيمة واحدة فقط بدون احتمالات أخرى و الفرضية البديلة تنص كما يلي : H1: a< 50 .
- الفرضية الابتدائية ثنائية الجانب(المركبة): نقول عن فرضية ابتدائية أنها مركبة إذا كانت تختبر قيمة واحدة فقط ولكن بعدة احتمالات.كمثال على ذلك نقول : H0: a = 50 وبالتالي فان الفرضية البديلة تنص كما يلي : H1: a ≠50 ، ففي هذه الحالة يوجد احتمالين وهما a>50 والاحتمال الآخر a< 50 .ويمكن أن تتوضع منطقة الرفض كما يلي :
  - إلى جانبي منطقة القبول عند ذلك نقول عن الاختبار الإحصائي بأنه ثنائي الجانب كما أشرنا قبل قليل.
  - إلى جانب واحد من منطقة القبول(إلى اليمين أو اليسار) عند ذلك نقول عن الاختبار الإحصائي بأنه أحادي الجانب كما أشرنا قبل قليل.والأشكال التالية توضح هذه الحالات.



3. الفرضية البديلة (H1:):توضع الفرضية البديلة كبديل في حال رفض الفرضية الابتدائية :H0 .
4. احصاء الاختبار: يتم استخدام قيم العينة المدروسة لحساب عدد واحد يأخذ دور صانع القرار و يدعى احصاء الاختبار.

3-6-الاختبارات المعنوية:تشمل الاختبارات المعنوية مجموعتين أساسيتين وذلك حسب مجال التطبيق وهما المجموعتان التاليتان:

- الاختبارات الخاصة بالعينات كبيرة الحجم :في حال العينات كبيرة الحجم يستخدم اختبار Z الذي يقوم على احصاء الاختبار الذي يتبع التوزيع الطبيعي .وهذا الاختبار يعتمد عليه بشكل خاص من أجل المقارنات التالية:
- مقارنة متوسطين لمجتمعين علما أن التشتت لهما معلوم.
- مقارنة متوسطين لمجتمعين علما أن التشتت لهما غير معلوم.

- مقارنة متوسط المجتمع مع متوسط العينة علما أن الانحراف المعياري للمجتمع معروف.

يكفي في هذا المجال أن نذكر أن العلاقة التي تعطينا اختبار Z-test التي سبق وأن شرحت في الفصل السابق كما تم شرح علاقة هذا الاختبار مع منحنى التوزيع الطبيعي .كذلك الأمر يمكن مراجعة الكتاب الثاني "الجزء العملي " للإطلاع على المزيد حول تطبيقات اختبار Z .

- الاختبارات الخاصة بالعينات صغيرة الحجم :في حال العينات صغيرة الحجم يستخدم الاختبارات التالية:

أ- اختبار t-test الذي يقوم على احصاء الاختبار الذي يتبع توزيع ستيودنت Student distribution .وهذا الاختبار يعتمد عليه بشكل خاص من أجل المقارنات التالية:

- مقارنة متوسط المجتمع مع متوسط العينة علما أن الانحراف المعياري للمجتمع غير معروف.
- اختبار الفرق بين نسبتين من مجتمع واحد.
- اختبار الفرق بين نسبتين من مجتمعين نفترض أن تباينهما متماثل
- اختبار الفرق بين متوسطين لعينتين نفترض أنهما متساويتين بعدد العناصر (العينات غير المستقلة)
- اختبار الفرق بين متوسطين لعينتين نفترض أنهما غير متساويتين بعدد العناصر (العينات المستقلة).
- اختبار القيم الشاذة وحذفها

وفيما يلي سنشرح بعضا من هذه الاختبارات والتي تلقى تطبيقا في مجال التجارب والاختبارات الزراعية والحيوية وهي التالية:

1. اختبارات النسبة المئوية بين نسبتين من مجتمع واحد: نستخدم مثل هذا الاختبار عندما تكون البيانات الإحصائية DATA عبارة عن نسبة مئوية،ومن أجل ذلك فإننا نستخدم احصاء الاختبار التالي :

$$T = p - p_0 / \sqrt{p_0 * q_0 / n}$$

مثـــــــال (1) :إذا كانت نسبة الأشجار المعتادة للإصابة بالمرض هي 15% ،تم رش الأشجار بأحد المبيدات فكانت نسبة الأشجار المريضة المتبقية هي 9% :فهل يعتبر المبيد المستخدم فعالا أم لا ؟

الحـــــل : نضع الفرضية الابتدائية وهي التالية :

$$H0: p = p_0 = 0.15$$

مقابل ذلك تكون الفرضية البديلة هي التالية :

$H1: p \neq p_0$

نطبق العلاقة السابقة الذكر لحساب قيمة احصاء الاختبار كما يلي :

$$T = 0.09 - 0.15 / \sqrt{0.15 * 0.85} / 100$$

$t = -1.6$

نخرج قيمة t_{tab} الجد ولية عند مستوى المعنوية 10% فنجد أن :

$$t_{tab} = 1.65$$

الآن نقارن قيمة مع قيمة t_{cct} المحسوبة من العلاقة السابقة أعلاه

نجد أن t_{tab} الجد ولية أكبر من قيمة t_{cct} المحسوبة وبالتالي نقبل الفرضية الابتدائية أي أن أثر المبيد غير معنوي و نستنتج بالنهاية أن المبيد غير فعال.

2. اختبارات النسبة المئوية بين نسبتين من مجتمعين لهما تباين واحد: نستخدم مثل هذا الاختبار عندما تكون البيانات الإحصائية DATA عبارة عن نسبة مئوية من مجتمعين تساوى فيها الانحراف المعياري،ومن أجل ذلك فإننا نستخدم احصاء الاختبار التالي :

$$T = p - p_0 / s * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

حيث أن :

$$s = \sqrt{(n_1 * p_1 * q_1 + n_2 * p_2 * q_2) / (n_1 + n_2 - 2)}$$

P_1 : النسبة الاولى

P_2 : النسبة الثانية

q_1 : متمم النسبة الاولى

q_2 : متمم النسبة الثانية

مثـــــــال (2) :تبين أن نسبة الأشجار المعتادة للإصابة بالمرض هي 20% ،تم رش الأشجار بأحد المبيدات فكانت نسبة الأشجار المريضة المتبقية هي 18 % :فهل تعتبر هاتان النسبتان مختلفتان علما أنه تم معاينة 20 شجرة في المرة الاولى و15 شجرة في المرة الثانية؟

الحـــــل : نضع الفرضية الابتدائية وهي التالية :

$H0: p_1 = p_2$

مقابل ذلك تكون الفرضية البديلة هي التالية :

$H1: p_1 \neq p_2$

$Df : n_1 + n_2 - 2 = 33$

نطبق العلاقة السابقة الذكر لحساب قيمة احصاء الاختبار كما يلي :

$$T = 0.20 - 0.18 / s * \sqrt{\frac{1}{20} + \frac{1}{18}}$$

حيث أن S تحسب من العلاقة:

$$s = \sqrt{(20 * 0.20 * 0.80 + 15 * 0.18 * 0.82) / (15 + 20 - 2)}$$

$T_{cct} = 0.14$

نخرج قيمة t_{tab} الجد ولية عند مستوى المعنوية 5% ودرجة حرية 33 فنجد أن :

$t_{tab} = 2.02$

الآن نقارن قيمة مع قيمة t_{cct} المحسوبة من العلاقة السابقة أعلاه

نجد أن t_{tab} الجد ولية أكبر من قيمة t_{cct} المحسوبة وبالتالي نقبل الفرضية الابتدائية أي أن تأثير المبيد غير معنوي و نستنتج أن المبيد لم يكن فعالا.

3. اختبار الفرق بين متوسطين لعينتين نفترض أنهما متساويتين بعدد العناصر (العينات غير المستقلة):في مثل هذه الحالة نستخدم مثل هذا الاختبار عندما تكون البيانات الإحصائية DATA عبارة عن عينتين متساويتين بعدد العناصر أي $n_1 = n_2$،ومن أجل ذلك فإننا نستخدم احصاء الاختبار التالي :

$$t = \bar{x}_1 - \bar{x}_2 / s_{\bar{d}}$$

$$s_{\bar{d}} = s_d / \sqrt{n}$$

$$s_d = \sqrt{\sum (d - \bar{d})^2 / (n-1)}$$

حيث أن : s_d الانحراف المعياري للفروق بين عناصر العينتين

$x`_1$ متوسط العينة الاولى

$x`_2$ متوسط العينة الثانية

مثـــــال (3): لدينا العينتين التاليتين :

والمطلوب اختبار الفرق بين متوسطين العينتين حيث أنهما متساويتين بعدد العناصر:

العينة الاولى	العينة الثانية
Y	X
5	10
13	20
16	30
7.73	40
5.51	50
10	60

نطبق العلاقة السابقة الذكر لحساب قيمة احصاء الاختبار t – test فنحصل على النتائج كما يلي :

t-Test: Two-Sample Assuming Equal Variances

النتائج	*X*	*Y*
Mean	35	9.54
Variance	350	18.80868
Observations	6	6
Pooled Variance	184.40434	
Hypothesized Mean Difference	0	
df	10	
t Stat	3.247382587	
P(T<=t) one-tail	0.004379611	
t Critical one-tail	1.812461505	
P(T<=t) two-tail	0.008759222	
t Critical two-tail	2.228139238	

الآن نضع الفرضية الابتدائية و هي التالية :

$H0: x`_1 = x`_2$

مقابل ذلك تكون الفرضية البديلة هي التالية :

$H1: x`_1 \neq x`_2$

نخرج قيمة t_{tab} الجد ولية عند مستوى المعنوية 5% ودرجة حرية 10 كما هو وارد في الجدول السابق فنجد أن :

$t_{tab} = 2.22$

الآن نقارن قيمة مع قيمة t_{cct} المحسوبة من العلاقة السابقة كما هو وارد في الجدول السابق أعلاه حيث أن $t_{clt} = 3.2473$

نجد أن t_{clt} المحسوبة أكبر من قيمة t_{tab} الجد ولية وبالتالي نرفض الفرضية الابتدائية أي أن متوسطي العينتين مختلفين والفرق بينهما معنوي مؤكد إحصائيا وليس عائدا للصدفة..

4. اختبار الفرق بين متوسطين لعينتين نفترض أنهما ليستا متساويتين بعدد العناصر (العينات المستقلة):في مثل هذه الحالة نستخدم مثل هذا الاختبار عندما تكون البيانات الإحصائية DATA عبارة عن عينتين غير متساويتين بعدد العناصر أي $n_1 \neq n_2$،ومن أجل ذلك فإننا نستخدم احصاء الاختبار التالي :

$$t = \bar{x}_1 - \bar{x}_2 / s_{\bar{d}}$$

$$s_{\bar{d}} = \sqrt{\frac{sp^2}{n_1} + \frac{sp^2}{n_2}}$$

$$sp^2 = s_1 + s_2 /(n_1 + n_2 - 2)$$

مثـــال (4) : لدينا العينتين التاليتين :

والمطلوب اختبار الفرق بين متوسطين العينتين حيث أنهما غير متساويتين بعدد العناصر:

العينة الاولى	العينة الثانية
Y	X

10	5
20	13
30	16
40	**10**
50	
60	

نطبق العلاقة السابقة الذكر لحساب قيمة احصاء الاختبار t – test فنحصل على النتائج كما يلي :

t-Test: Two-Sample Assuming Unequal Variances

	X	*Y*
Mean	35	11
Variance	350	22
Observations	6	4
Hypothesized Mean Difference	0	
df	6	
t Stat	3.003913896	
P(T<=t) one-tail	0.01194363	
t Critical one-tail	1.943180905	
P(T<=t) two-tail	0.023887261	
t Critical two-tail	2.446913641	

الآن نضع الفرضية الابتدائية وهي التالية :

$$H0: \ x`_1 = x`_2$$

مقابل ذلك تكون الفرضية البديلة هي التالية :

$$H1: \ x`_1 \neq x`_2$$

نخرج قيمة t_{tab} الجد ولية عند مستوى المعنوية 5% ودرجة حرية 6 كما هو وارد في الجدول السابق فنجد أن :

$$t_{tab} = 2.446913$$

الآن نقارن قيمة مع قيمة t_{cct} المحسوبة من العلاقة السابقة كما هو وارد في الجدول السابق أعلاه حيث أن $t_{clt} = 3.00391$

نجد أن t_{clt} المحسوبة أكبر من قيمة t_{tab} الجد ولية وبالتالي نرفض الفرضية الابتدائية أي أن متوسطي العينتين مختلفين والفرق بينهما معنوي مؤكد إحصائيا وليس عائدا للصدفة مثل النتيجة السابق.

5. اختبار القيم الشاذة وحذفها: كثيرا ما ينتج لدينا في تجربة ما بين قيم العينة قيم شاذة ومتطرفة تؤثر سلبا 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 أ ننا نستخدم احصاء الاختبار التالي :

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

حيث أن : x_0 القيمة المراد اختبارها

s الانحراف المعياري للعينة

x` المتوسط الحسابي للعينة

وتكون الفرضيات الإحصائية كما يلي:

الفرضية الابتدائية :

$$H0: \quad x_0 = x`$$

الفرضية البديلة :

$$H1: x_0 \neq x`$$

مثـــــــال (5) :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 :

$$x` = 122.93$$

$$t = 550 – 383.75 / 122.93$$

$$t = 1.35$$

نخرج قيمة t_{tab} الجد ولية عند مستوى المعنوية 10% ودرجة حرية 7 فنجد أن :

$$t_{tab} = 1.89$$

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} الجد ولية 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.

ب- اختبار مربع كاي χ^2 : يستخدم هذا الاختبار عند وجود نوعين من المؤشرات أو البيانات وهي التالية :

- مؤشرات أو بيانات حقيقية أو تجريبية observed يحصل عليها الباحث من خلال تجربة معينة.

- مؤشرات أو بيانات نظرية : يحصل عليها الباحث من خلال تطبيق معادلة أو نموذج رياضي arithmetical model أو معلومات مرجعية من خبرات سابقة في أماكن أخرى وظروف مختلفة.

في مثل هذه الحالة باستخدم اختبار مربع كاي χ^2 لقياس الفرق بين المؤشرات أو البيانات الحقيقية أو التجريبية وبين المؤشرات أو البيانات النظرية.

ولهذه الغاية يمكننا أن نستخدم احصاء الاختبار التالي :

$$\chi^2 = \sum((O_i - E_i)^2 / E_i)$$

حيث أن : O_i القيمة الحقيقية أو التجريبية

E_i القيمة النظرية.

بعد الآن نقارن قيمة χ^2 الجد ولية مع قيمة χ^2 المحسوبة عند درجة حرية $df = k-1$ حيث k عدد قيم أو المشاهدات للعينة أو المجموعات إذا كانت العينة مقسمة إلى مجموعات وعند مستوى المعنوية 5% لمعرفة هل الفروق بين البيانات الحقيقية (التجريبية) وبين المؤشرات أو البيانات النظرية هل هي معنوية حقيقية مؤكدة إحصائيا أم غير ذلك.

مثـــال (6) : لدينا العينتين التاليتين الاولى عينة من البيانات الحقيقية أو التجريبية وأخرى من البيانات النظرية:

البيانات النظرية	البيانات الحقيقية
y	x
12	10
22	20
25	30
44	40
66	50
48	60

والمطلوب اختبار هل الفروق بين البيانات الحقيقية أو التجريبية وبين المؤشرات أو البيانات النظرية هل هي معنوية حقيقية مؤكدة إحصائيا أم لا:

نطبق العلاقة التي تساعدنا في حساب قيمة مربع كاي المحسوبة وهي التالية :

$$\chi^2 = \sum((O_i - E_i)^2 / E_i)$$

بعد تبديل المتغيرات في العلاقة السابق نحصل على الجدول التالي :

	البيانات الحقيقية	البيانات النظرية	الجد χ^2 = 11.07 ولية	
	x	y	(x-y)^2	(x-y)^2/y
	10	12	4	0.333333
	20	22	4	0.181818
	30	25	25	1
	40	44	16	0.363636
	50	66	256	3.878788
	60	48	144	3
Total Σ	210	217	449	8.757576

الآن نقارن قيمة χ^2 الجد ولية 11.07 وهي تساوي مع قيمة χ^2 المحسوبة عند درجة حرية df = k-1= 5 (حيث k عدد قيم أو المشاهدات للعينة أو المجموعات إذا كانت العينة مقسمة إلى مجموعات) وعند مستوى المعنوية 5% فنجد أن χ^2 الجد ولية أكبر من قيمة χ^2المحسوبة والتي تساوي 8.757576 كما هو وارد في الجدول السابق وبالتالي نستنتج أن الفروق بين البيانات الحقيقية (التجريبية) وبين المؤشرات أو البيانات النظرية هي ليست معنوية وليست حقيقية وغير مؤكدة إحصائيا وتعود للصدفة أو لعوامل خارجية مجهولة.

ج- اختبار f-test : يعد اختبار f-test من الاختبارات المعنوية الهامة جدا وهو يتبع توزيع فيشر ، Fisher distribution كما أن هذا الاختبار يتعلق بتحليل التباين بشكل كامل لذلك سيتم شرحه بالتفصيل في الفصل القادم.

4-6- مجال أو حدود الثقة Confidence interval :إن المقصود من مجال أو حدود الثقة هو تحديد المجال (الحد الأدنى والحد الأعلى) لأحد المؤشرات الإحصائية Parameter مع حساب الاحتمال وقوع هذا المؤشر ضمن المجال المحدد.وعادة يتم استخراج مجال أو حدود الثقة من العلاقة التي تستخدم لحساب احصاء اختبار المعنوية لهذا المؤشر.

يقوم الاحصائيون عادة بحساب مجال أو حدود الثقة للحالات التالية :

1. حساب مجال أو حدود الثقة لمتوسط مجتمع طبيعي تباينه غير معروف
2. حساب مجال أو حدود الثقة لمتوسط مجتمع طبيعي تباينه معروف
3. حساب مجال أو حدود الثقة لمتوسط عينة كبيرة الحجم .
4. حساب مجال أو حدود الثقة لمعاملات الانحدار
5. حساب مجال أو حدود الثقة للقيم التابعية لمعادلة الانحدار .

تعتبر طريقة حساب مجال أو حدود الثقة للحالات الخمسة السابقة الذكر أمرا صعبا وخاصة بالنسبة لغير المتخصصين ،إلا أن توفر برامج التحليل الإحصائي في برمجيات الجداول الالكترونية مثل برنامج Excel وبرنامج SPSS وبرنامج LOTUS 123 وغيرها جعل إنجاز مثل هذه الحسابات أمرا سهلا جدا لذلك سنكتفي بشرح مثال واحد عن تطبيق هذه البرمجيات لحساب مجال أو حدود الثقة لمعاملات الانحدار نظرا لأهميتها وكفايتها كما في المثال التالي:

مثــال (7) : عند دراسة العلاقة ما بين كمية السماد المضاف كغ/ها وصفة الإنتاج لأحد المحاصيل وجاء كان لدينا التحليل الإحصائي للبيانات الإحصائية DATA كان لدينا ما يلي :

الإنتاج	السماد
Y	**X**
5	**1**

2	**13**
3	**16**
4	**23**
5	**33**
6	**38**
7	**40**

أجري التحليل الإحصائي للانحدار وحصلنا على النتائج التالية:

معاملات الانحدار	Coefficients	الحد الأدنى Lower 95%	الحد الأعلى Upper 95%
a	0.57142857	-5.274208	4.131351
b	6.14285714	5.0912837	7.194431

نلاحظ من ألجدول السابق أن عامل الانحدار b = 6.14285714 يقع ضمن مجال الثقة الوارد في الجدول السابق مابين 5.0912837 و 7.194431 باحتمال قدره : (1-α = 5%) .

*****

مسائل وحلول متقدمة

اختبار مربع كايCHITEST

إرجاع اختبار الاستقلالية. تقوم CHITEST بإرجاع القيمة الناتجة من التوزيع كاي تربيع (γ2) لإحصاء البيانات ودرجات الحرية المناسبة. يمكنك استخدام اختبارات γ2 في تحديد ما إذا كانت النتائج الفرضية تحققها تجربة ما.

بناء الجملة لاستخدامها في البرنامج في الحاسب:

CHITEST(actual_range,expected_range)

Actual range (النطاق الفعلي) نطاق البيانات الذي يحتوي على الملاحظة المراد اختبارها مقارنةً بالقيم المتوقعة.

Expected range (النطاق المتوقع) نطاق البيانات الذي يحتوي على نسبة حاصل ضرب مجاميع الصفوف ومجاميع الأعمدة إلى المجموع الكلي.

تنويهات

- إذا كان actual range و expected range لهما عدد مختلف من نقاط البيانات، تقوم CHITEST بإرجاع قيمة الخطأ #N/A.
- يحسب اختبار γ2 أولاً إحصائية γ2 ثم يلخص فرق القيم الفعلية عن القيم المتوقعة.

مثـــــــال:

1. قم بإنشاء مصنف أو ورقة عمل فارغة.
2. حدد المثال في موضوع التعليمات. لا تحدد رؤوس الصفوف أو الأعمدة.
3. اضغط CTRL+C
4. في ورقة العمل، حدد خلية A1، واضغط على CTRL+V.
5. للتبديل بين عرض النتائج وعرض الصيغ التي تقوم بإرجاعها، اضغط CTRL+`، أو في القائمة **أدوات**، أشر إلى **تدقيق الصيغة**، ثم انقر فوق **وضع تدقيق الصيغة**.

	A	B	C
	رجال (فعلي)	نساء (فعلي)	الوصف
1	58	35	موافقة
2	11	25	محايد
3	10	23	عدم موافقة
4	رجال (متوقع)	نساء (متوقع)	الوصف
5	45.35	47.65	موافقة
6	17.56	18.44	محايد
7	16.09	16.91	عدم موافقة
8	الصيغة	وصف (الناتج(	
	=CHITEST(A2:B4,A6:B8)	إحصائية γ2 للبيانات أعلاه هي 16.16957 مع درجتين من درجات الحرية.	

مجال الثقة لوسط مجموعة بياناتCONFIDENCE

إرجاع فترة الثقة لوسط مجموعة بيانات. فترة الثقة هي النطاق الواقع على أي من جانبي وسط مجموعة البيانات. فعلي سبيل المثال، إذا طلبت أحد المنتجات عبر البريد، يمكنك تحديد، بمستوى معين من الثقة أقرب آخر موعد لوصول المنتج.

بناء الجملة

CONFIDENCE(alpha,standard_dev,size)

Alpha (ألفا) مستوى الأهمية المُستخدم في حساب مستوى الثقة. يساوي مستوى الثقة 100*(1 - ألفا)%؛ بمعنى أن ألفا 0.05 تشير إلى مستوى ثقة قدره 95 بالمائة.

Standard_dev (الانحراف المعياري) الانحراف المعياري لمحتوى نطاق البيانات ومفترض أنه معطى.

Size (الحجم) حجم العينة.

تنويهات

- إذا كانت أية وسيطة غير رقمية، تُرجع SERIESSUM القيمة الخطأ !VALUE#.
- إذا كانت alpha ≥ 0 أو إذا كانت alpha ≤ 1، تقوم CONFIDENCE بإرجاع قيمة الخطأ !NUM#.
- إذا كانت standard_dev ≥ 0، تقوم CONFIDENCE بإرجاع قيمة الخطأ !NUM#.
- إذا لم يكن size عدداً صحيحاً، يتم اقتصاصه.
- إذا كان size < 1، تقوم CONFIDENCE بإرجاع قيمة الخطأ !NUM#.
- بفرض أن ألفا تساوي 0.05، نحتاج إلى حساب الناحية الواقعة تحت المنحنى المعياري الطبيعي الذي يساوي (1 - ألفا)، أو 95 بالمائة. تساوي هذه القيمة ± 1.96. ولهذا تكون فترة الثقة:

مثـــــــــال:

بفرض أننا قد لاحظنا في عينة تتكون من 50 راكباً أن متوسط طول الرحلة إلى العمل هو 30 دقيقة مع انحراف معياري للمحتوى 2.5. فيمكن الثقة بنسبة 95 بالمائة أن وسط المحتوى يقع في الفاصل الزمني:

1. قم بإنشاء مصنف أو ورقة عمل فارغة.
2. حدد المثال في موضوع التعليمات. لا تحدد رؤوس الصفوف أو الأعمدة.
3. اضغط CTRL+C
4. في ورقة العمل، حدد خلية A1، واضغط على CTRL+V.
5. للتبديل بين عرض النتائج وعرض الصيغ التي تقوم بإرجاعها، اضغط CTRL+`، أو في القائمة **أدوات**، أشر إلى **تدقيق الصيغة**، ثم انقر فوق **وضع تدقيق الصيغة**.

	A	B
1	**بيانات**	**الوصف**
2	0.05	مستوى الأهمية

3	2.5	الانحراف المعياري للمحتوى
4	50	حجم العينة

الصيغة	**وصف (الناتج)**
=CONFIDENCE(A2,A3,A4)	فترة الثقة لوسط مجموعة البيانات .بمعنى آخر، أن متوسط طول الرحلة إلى العمل يساوي 30 ± 0.692951 دقيقة، أو من 29.3 إلى 30.7 دقيقة (0.692951).

توزيع الاحتمال F : **FDIST**

لإرجاع توزيع الاحتمال F. يمكنك استخدام هذه الدالة لتحديد ما إذا كان هناك درجات للاختلاف بين مجموعتي البيانات. على سبيل المثال، يمكنك فحص نقاط الاختبار للفتيان والفتيات المتقدمين لمدرسة ثانوية جديدة وتحديد إذا كانت التباين بين الفتيات مختلف عن ذلك الموجود بين الفتيان .

بناء الجملة

FDIST(x,degrees_freedom1,degrees_freedom2)

X القيمة التي يتم تقييم الدالة عندها.

Degrees_freedom1 (درجات الحرية 1) هي بسط درجات الحرية.

Degrees_freedom2 (درجات الحرية 2) هي مقام درجات الحرية.

تنويهات

- إذا كانت أي وسيطة منها غير رقمية، تقوم FDIST بإرجاع قيمة الخطأ !VALUE#.
- إذا كانت x سالبة، تقوم FDIST بإرجاع قيمة الخطأ !NUM#.
- إذا لم تكن (درجات_الحرية 1) degrees_freedom1 أو (درجات_الحرية 2) degrees_freedom2 أعداداً صحيحة، يتم اقتصاصها.
- إذا كانتdegrees_freedom1 < 1 أو degrees_freedom1 ≥ 10^10، تقوم FDIST بإرجاع قيمة الخطأ !NUM#.

- إذا كانتdegrees_freedom2 < 1 أو degrees_freedom2 ≤ 10^10، تقوم FDIST بإرجاع قيمة الخطأ !NUM#.
- يتم حساب FDIST كـ (F)P=FDIST<x، حيث F متغير عشوائي له التوزيع F.

مثـــــــــــال:

1. قم بإنشاء مصنف أو ورقة عمل فارغة.
2. حدد المثال في موضوع التعليمات. لا تحدد رؤوس الصفوف أو الأعمدة.
3. اضغط CTRL+C
4. في ورقة العمل، حدد خلية A1، واضغط على CTRL+V.
5. للتبديل بين عرض النتائج وعرض الصيغ التي تقوم بإرجاعها، اضغط `+CTRL، أو في القائمة **أدوات**، أشر إلى **تدقيق الصيغة**، ثم انقر فوق **وضع تدقيق الصيغة**.

	A	B
	بيانات	**الوصف**
1	15.20675	القيمة التي يتم تقييم الدالة عندها
2	6	بسط درجات الحرية
3	4	مقام درجات الحرية
4	**الصيغة**	**وصف (الناتج)**
	=FDIST(A2,A3,A4)	Fاحتمالية التوزيع حسب الشروط الموضحة أعلاه(0.01)

اختبار_F **FTEST**

لإرجاع ناتج اختبار_F. يقوم اختبار F بإرجاع الاحتمال وحيد الطرف الذي تتباين فيه Array1 و Array2 بشكل غير مختلف اختلافاً كبيراً. استخدم هذه الدالة لتحديد ما إذا كانت عينتان بهما تباينات مختلفة. على سبيل المثال، عند معرفة درجات اختبارات من المدارس الحكومية والمدارس الخاصة، يمكنك اختبار ما إذا كانت هذه المدارس بها مستويات مختلفة من تباين درجات الاختبارات.

بناء الجملة

FTEST (array2،array1)

Array1 (مصفوفة 1) مصفوفة أو نطاق البيانات الأول.

Array2 (مصفوفة 2) مصفوفة أو نطاق البيانات الثاني.

تنويهات

- يجب أن تكون الوسائط إما أرقاماً أو أسماء، أو مصفوفات، أو مراجع تحتوي على أرقام.
- إذا احتوت وسيطة مصفوفة أو مرجع على نص، أو قيم منطقية، أو خلايا فارغة، يتم تجاهل تلك القيم؛ رغم ذلك يتم تضمين الخلايا التي تحتوي على قيمة الصفر (0).
- إذا كان عدد نقاط البيانات في array1 أو array2 أقل من 2، أو إذا كان تباين array1 أو array2 صفراً، تقوم FTEST بإرجاع قيمة الخطأ !DIV/0 #.

مثـــــــال :

1. قم بإنشاء مصنف أو ورقة عمل فارغة.
2. حدد المثال في موضوع التعليمات. لا تحدد رؤوس الصفوف أو الأعمدة.
3. اضغط CTRL+C
4. في ورقة العمل، حدد خلية A1، واضغط على CTRL+V.
5. للتبديل بين عرض النتائج وعرض الصيغ التي تقوم بإرجاعها، اضغط CTRL+`، أو في القائمة **أدوات**، أشر إلى **تدقيق الصيغة**، ثم انقر فوق **وضع تدقيق الصيغة**.

	A	B
1	مجموعة بيانات 1	مجموعة بيانات 2
2	6	20
3	7	28
4	9	31
5	15	38
6	21	40