المحاضرة التاسعة

التباين والانحراف المعياري Variance and Standard Deviation

يعد كل من الانحراف المعياري والتباين كمقياس للتشتت من انسب المقاييس نظراً لتجاوزها المقاييس السابقة من ناحية واستخدامها على نطاق واسع في التحليل من ناحية ثانية ويعرف التباين بأنه معدل مجموع مربعات انحرافات القيم عن متوسطها .

أ- التباين في حالة البيانات غير المبوبة :-

تباين العينة $S^2 = \frac{\sum(yi - \bar{y})^2}{n-1}$ **الطريقة الاعتيادية**

$S^2 = \frac{\sum y^2 i - \frac{(\sum yi)^2}{n}}{n-1}$ **الطريقة السريعة**

حيث ان التباين $S^2 = \frac{SS}{d.f}$

$SS = \sum(yi - \bar{y})^2$

أو $SS = \sum y^2 i - \frac{(\sum yi)^2}{n}$

d.f = هي درجات الحرية او عدم السيطرة

d.f = n – 1

حيث n هي عدد القيم

$$\sigma^2 = \frac{\sum y^2 i - \frac{(\sum yi)^2}{N}}{N-1}$$

مثال// حسب التباين للقيم التالية التي تمثل درجات 6 طلاب من طلبة كلية الصيدلة في مادة كيمياء الادوية العملي ؟ yi = 9 , 4 , 6 , 8 , 10 , 5 , 7

Yi	$yi - \bar{y}$	$(yi - \bar{y})^2$
9	+2	4
4	- 3	9
6	- 1	1
8	+1	1
10	+ 3	9
7	0	0
5	- 2	4
49		28

$\bar{y} = \frac{\sum yi}{n} \quad \rightarrow \quad \bar{y} = \frac{49}{7}$

$S^2 = \frac{\sum (yi - \bar{y})^2}{n-1}$

$S^2 = \frac{28}{7-1} = \frac{28}{6} = 4.67$

الطريقة السريعة

Yi	yi^2
9	81
4	16
6	36
8	64
10	100
7	49
5	25
49	371

$S^2 = \frac{\sum y^2 i - \frac{(\sum yi)^2}{n}}{n-1}$

$S^2 = \frac{371 - \frac{(49)^2}{7}}{7-1}$

$$S^2 = \frac{371-343}{7-1} = \frac{28}{6} = 4.67$$

ب – في حالة البيانات المبوبة :

$$S^2 = \frac{\sum fi(yi-\overline{y})^2}{\sum fi-1}$$

$$S^2 = \frac{\sum fiy^2i - \frac{(\sum fiyi)^2}{\sum fi}}{\sum fi-1}$$

مثال// احسب التباين للبيانات التالية التي تمثل توزيع طلبة كلية الصيدلة حسب الوزن

الفئات	التكرار fi	مركز الفئات yi	$fiyi$	$(yi - \overline{y})$	$(yi - \overline{y})^2$	$f(yi - \overline{y})$
0 – 62	5	61	305	- 6.54	42.7716	213.858
3- 65	15	64	960	- 3.54	12.5316	187.974
6 – 68	45	67	3015	0.54	0.2916	13.122
9 – 71	27	70	1890	2.46	6.0516	163.3932
4 - 72	8	73	584	5.46	29.8226	238.4928
	100		6754			816.84

$$\overline{y} = \frac{\sum fiyi}{\sum fi} = \frac{6754}{100} = 67.54$$

$$S^2 = \frac{\sum fi(yi-\overline{y})^2}{\sum fi-1}$$

$$S^2 = \frac{816.84}{100-1} = \frac{816.84}{99} = 8.25$$

الفئات	التكرار *fi*	*yi*	*fiyi*	$y^2 i$	$fi\, y^2 i$
60 – 62	5	61	305	*3721*	18605
63- 65	15	64	960	4096	61440
66 – 68	45	67	3015	4489	202005
69 – 71	27	70	1890	4900	132300
74 - 72	8	73	584	5329	42632
	100		6754		456982

$$S^2=\frac{\sum fiy^2i-\frac{(\sum fiyi)^2}{\sum fi}}{\sum fi} = \frac{456982-\frac{(6754)^2}{100}}{100-1}$$

$$S^2=\frac{456982-456165.16}{99} = \frac{816.84}{99} = 8.25$$

Standard Deviation الانحراف المعياري

الجذر التربيعي لمجموع مربعات انحرافات القيم عند متوسطها ويستخدم على نطاق واسع كونه يتعامل مع نفس وحدات القياس للمشاهدات الاصلية ويعتبر الانحراف المعياري اهم مقاييس التشتت واكثرها استعمالاً في مجال التحليل الاحصائي .

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

مثال// اوجد الانحلااف المعياري اذا كان التباين (4.67) وكذلك اذا كان التباين (8.25)

S = $\sqrt{s^2}$

S = $\sqrt{4.67}$ **= 2.16**

S = $\sqrt{8.25}$ **= 2.87**

الخطأ القياسي Standard error :-

يسمى الانحراف المعياري لمتوسط العينة ويستخدم للدلالة على التشتت فكلما كان الخطأ القياسي قليلاً كلما كان هناك تقارب او تجانس اكثر بين القيم وكاما زاد الخطأ القياسي كلما قلت دقة القياس ودل ذلك على تشتت القيم

$$S\overline{y} = \frac{S}{\sqrt{n}}$$

مثال// اوجد الخطأ القياسي اذا كان التباين (4.67) وعدد القيم 6

$$S\overline{y} = \frac{S}{\sqrt{n}} = \frac{2.16}{\sqrt{6}} = \frac{2.16}{2.45} = +0.88$$

المحاضرة العاشرة

ب – مقاييس التشتت النسبى :-

من اهم مقاييس التشتت النسبي

1- معامل الاختلاف The coefficient variation

يستخدم للمقارنه بين المجموعات المختلفة او بين العينات فأننا لا نستطيع اجراء مقارنة بناء على الانحراف المعياري لكل مجموعة لاننا بحاجة الى توحيد القياس بالنسبة للمجموعتين لذلك يتم استخدام معامل الاختلاف

$$C.V = \frac{S}{\bar{y}} \times 100$$

فيما يلي درجات مجموعتين من الطلاب

المجموعة الاولى	المجموعة الثانية
2	2000
2	2000
4	4000
5	5000
12	12000

فعند حساب الوسط الحسابي والانحراف المعياري لكلتا المجموعتين

المجموعة	الوسط الحسابي	الانحراف المعياري
الاولى	5	4.123
الثانية	5000	4123.11

فهل نستطيع المقارنة بين المجموعتين بناءاً على الانحراف المعياري ، كما اشرنا سابقاً لا نستطيع حيث نحن بحاجة الى توحيد القياس لابد من استخدام معامل الاختلاف

معامل الاختلاف للمجموعة الاولى $C.V = \frac{4.123}{5} \times 100 = 82.46\ \%$

معامل الاختلاف للمجموعة الثانية $C.V = \frac{4123.11}{5000} \times 100 = 82.46\ \%$

بناءاً على هذه النتيجة فأن معامل الاختلاف هو واحد بالنسبة للمجموعتين او ان التباين متساوي .

يعرف معامل الاختلاف على انه النسبة المئوية التي يشكلها الانحراف المعياري

مثال// نتائج الامتحانات لدرس الاحصاء الحياتي والكيمياء السريرية لطلبة كلية الصيدلة كانت كما مبين ادناه اي من الدرجات بالنسبة للدرسين اكبر تشتتاً ؟

	الاحصاء	الكيمياء
الوسط الحسابي	78	73
الانحراف المعياري	8	76

معامل الاختلاف للاحصاء $C.V = \frac{8}{78} \times 100 = \%\ 10.25$

معامل الاختلاف للكيمياء $C.V = \frac{76}{73} \times 100 = \%\ 10.41$

ملاحظة // الحد الاعلى لمعامل الاختلاف للتجارب المختبرية يجب ان لا يتجاوز (10%) وللتجارب الحقيقية يجب ان لا يتجاوز (20%)

2- الدرجة المعيارية Standard Score

ان المقارنة بين الدرجات للفرد بناءاً على الدرجة الخام ليس له معنى وبالتالي لابد من تحويل هذه الدرجة الى درجة جديدة ، واحد هذه التحويلات تسمى بالدرجة المعيارية ومن خصائصها ان متوسطها (صفر) وانحرافها المعياري (1) وتستخدم الدرجة المعيارية لمقارنة اداء طالب معين في مواد مختلفة مثل :

مثال// لمقارنة اداء طالب من طلبة كلية الصيدلة في مواد دراسية مختلفة ؟

الدرجة المعيارية $Z = \frac{yi - \bar{y}}{s}$

	اللغة الانكليزية	الاحصاء الحياتي	كيمياء الادوية
الدرجة	75	70	60
الوسط الحسابي	70	55	50
الانحراف المعياري	10	15	5

عند النظر الى الدرجات نقول اداء الطالب في اللغة الانكليزية افضل ولكن عند التحويل الى الدرجة المعيارية

$$Z = \frac{y_i - \bar{y}}{S}$$

اللغة الانكليزية $Z = \frac{75-70}{10} = 0.5$

الاحصاء $Z = \frac{70-55}{15} = +1$

كيمياء الادوية $Z = \frac{60-50}{5} = +2$

اذن اداء الطالب افضل في مادة كيمياء الادوية .

المحاضرة الحادية عشرة

الاختبارات الاحصائية

الفرضية الاحصائية Statistical hypothesis :-

يفترض على الباحث ان يضع الفرضية الاحصائية لاختيارها قبل البدء بتنفيذ التجربة ، والفرضية الاحصائية (عبارة عن ادعاء او تصريح قد يكون صائباً او خطأ حول معلمة (صفة) او اكثر لمجتمع او مجموعة من المجتمعات والفرضية الاحصائية :

1- فرضية العدم Null hypothesis :- يرمز لها بالرمز Ho وهي التي تفترض عدم وجود فروق معنوية بين المتوسطات للمعاملات اي ان M1= M2
2- الفرضية البديلة Alemantive hypothesis :- ويرمز لها بالرمزH1 وهي التي تنص عن وجود فروقات معنوية بين متوسطات المعاملات اي ان M1≠M2

ولذلك فأن الباحث او الاحصائي دائماً يحاول ان يضع الفرضية بشكل يأمل ان يرفضها فمثلاً اذا اراد باحث ان يقارن بين عقار مصنع محلياً مع عقار مصنع خارج العراق في فعاليتهما في علاج مرض فأنه يضع فرضية فحواها بأنه لا توجد فروقات جوهرية او معنوية بين العقارين في فعاليتهما في علاج المرض وهكذا الفرضية التي يضعها الباحث على امل ان يرفضها تدعى فرضية العدم يقودنا الى قبول فرضية بديلة وعند رفض فرضية العدم وهي صحيحة تقع في خطأ من النوع الاول ويرمز له بالرمز (∝) اما اذا قبلنا فرضية العدم وهي خطأ نقع في خطأ من النوع الثاني (B) والذي يرمز له (β) وان خطأ القبول او الرفض للفرضيات الموضوعة يكون بدرجة احتمال او تسمى مستوى المعنوية والتي يرمز لها بالرمز (∝) وهي 1% و 5% ومستوى المعنوية (هي درجة الاحتمال التي ترفض فيها فرضية العدم عندما تكون صحيحة) ويكون اتخاذ القرار بدرجة احتمال 1% اقوى وبثقة اكبر وهذا يعني ان اعادة التجربة مئة مرة يكون احتمال ارتكاب الخطأ في اتخاذ القرار مرة واحدة اي اننا نرفض فرضية العدم وهي صحيحة واتخاذ القرار بمستوى 5% يحتمل ان تخطيء خمس مرات يرفضنا فرضية العدم وهي صحيحة .

الاختبارات الاحصائية :-

تستخدم عدة طرق احصائية لمعرفة الفروقات بين تأثير معاملة واخرى اضافة الى طرق التصميم المتبعة ولا تقل هذه الاختبارات الاحصائية في الاهمية في التحليل والاستنتاج عن طريق تصميم التجارب وهي الطرق الاحصائية ذات الاستخدام الواسع في مجال علوم الحياة والعلوم الاخرى.

اختبار -t :- كتب احد باحثي الاحصاء في بداية القرن العشرين المدعو *William Gossat* تحت اسم مستعار *Student* احد بحوثه الاحصائية عن هذه الطريقة استنبط فيها طريقة لفحص الاحصائية بأستخدام قياسات محسوبة (S^2 و $\bar{y}$) من العينات والمتغيرات وهذه الطريقة عبارة عن اختبار t ويقسم اختبار t الى

1- اختبار t يتعلق بمتوسط واحد

$$t = \frac{\bar{y}-M}{S\bar{y}}$$

t = t المحسوبة، M = متوسط المجتمع، $\bar{y}$ = متوسط العينة، $S\bar{y}$ = الخطأ القياسي للعينة.

$$S\bar{y} = \frac{S}{\sqrt{n}}$$

2- اختبار t يتعلق بمتوسطين

$$t = \frac{\bar{y}1-\bar{y}2}{S(\bar{y}i-\bar{y}i)}$$

حيث t= المحسوبة ، $\sqrt{\frac{2\ \text{mse}}{r}}$ = $S(\bar{y}i-\bar{y}i)$ الخطأ القياسي للفرق بين متوسطين ، r= عدد التكرارات

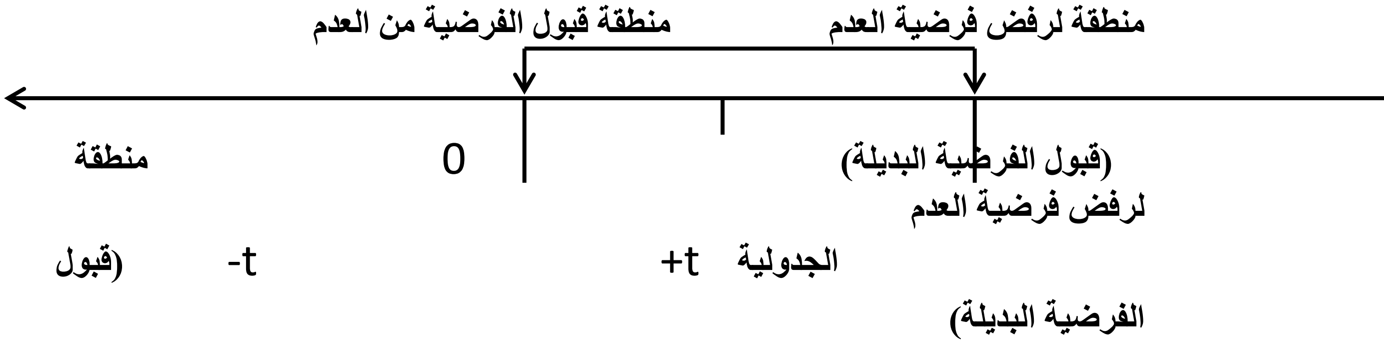
حيث تمثل t انحراف معدل العينة عن معدل المجتمع مقسوماً على الانحراف القياسي او المعياري للمعدلات ، ويستخدم للاستدلال فيما اذا كان انحراف معدل العينة عن معدل المجتمع طبيعياً او غير اعتيادي اذ من المفروض ان المشاهدات تتوزع توزيعاً طبيعياً حول المجتمع الذي اخذت منه ، كذلك فأن معدلات العينات لها توزيع طبيعي حول المجتمع .

تحسب قيمة *t* من العينة بصورة مباشرة *t- Calculate* وتقارن مع *t* الجدولية والتي على اساسها يتم قبول او رفض الفرضيات الموضوعه فأذا كانت قيمة *t* المحسوبة اكبر او تساوي قيمتها في الجدول (جدول *t*) لمستوى المعنوية المطلوب للاختبار عليه ودرجة الحرية (*n-1*) تعتبر في هذه الحالة العينة غير ممثلة للمجتمع ، كذلك يستخدم اختبار *t* لمقارنة معدلين او متوسطين من عينتين اذا كانت هاتان العينتان تعودان لنفس المجتمع ام لا سواء كانت هاتان العينتان متساويتان في عدد المشاهدات مزدوجة متساوية او غير متساوية (غير مزدوجة)وفي هذه الحالة نستخدم :-

$$t = \frac{\bar{y}1-\bar{y}2}{S(\bar{y}i-\bar{y}i)}$$

كذلك نستخدم t لاستخراج الفرق المعنوي الاصغر L.S.D   L.S.D =

$$t(\propto , v)\sqrt{\frac{2mse}{r}}$$



1- اختبار يتعلق بمتوسط واحد

مثال// اشار سجل مستشفى الولادة لمدينة كربلاء بأن معدل وزن الاطفال عند الولادة للسنين الماضية هو 5.5 كغم اخذت عينة عشوائية في الربع الاول من هذه السنة مؤلفة من 30 طفل وكان معدل وزنهم في تلك السنة 5.1 كغم وبأنحراف قياس قدره 0.9 كغم فهل هناك فرق معنوي في وزن الاطفال في هذه

السنة عما هو معروف في السنين الماضية اختبر ذلك تحت مستوى احتمال 0.01 علماً ان قيمة t الجدولية = 2.756 ؟

خطوات الاختبار

1- وضع الفرضيات $Ho : M1 = 5.5$

$H1 : M1 \neq 5.5$

2- اختبار الفرضية

$$t= \frac{\bar{y}-M}{\frac{s}{\sqrt{n}}}$$

$$t= \frac{5.1-5.5}{\frac{0.9}{\sqrt{30}}} = -2.434$$

3- استخراج قيمة t الجدولية لمستوى معنوية $\propto = 0.01$ ودرجة حرية =29

t= الجدولية = 2.756

4- الاستنتاج : - بما ان القيمة المطلقة (t المحسوبة = 2.434) اقل من t الجدولية لذا نقبل فرضية العدم *Ho* الى لا يوجد فرق معنوي بين اوزان الاطفال عند الولادة في هذه السنة عما هو في السنين الاخرى .

مثال//كان متوسط الزيادة وزن 12 فأرة بعد تغذيتها بغذاء يحتوي على 1% مضاد حيوي 145غم وبأنحراف قياسي للوسط الحسابي 2.3غم ففي مستوى احتمال 0.05 هل يمكن القول بأن الزيادة في الوزن نتيجة التغذية على هذا الغذاء لا تقل عن 150غم علماً ان قيمة t الجدولية 2.201

1- ضع الفرضيات

$Ho : M1 \geq 150$

$H1 : M1 < 150$

2- اختبار الفرضية

$$t= \frac{\bar{y}-M}{\frac{s}{\sqrt{n}}}$$

-2.174

$$t = \frac{145-150}{2-3} =$$

استخراج t الجدولية $\propto$ = 0.05 ودرجة حرية11 =2.201

4- الاستنتاج :- بما ان قيمة t المحسوبة (2.174) اقل من قيمتها في الجدوله 2.201 (ت الجدولية) نقبل فرضية العدم ونرفض الفرضية البديلة اي معدل الزيادة في الوزن لا تقل عن 150غم.

مثال3// ادعت احدى شركات انتاج السكاير بأن نسبة النيكوتين في انتاجها من السكاير لا يتجاوز 17.5ملغم ،اخذت عينة عشوائية مؤلفة من 9 سكاير وقيست نسبة النيكوتين فيها فكانت كالاتي :

yi= 18 , 18 ,16 , 20 ,19 , 19 ,18 , 18 , 17

فهل ادعاء الشركة صحيح تحت مستوى 0.05 علماً ان t الجدولية تحت مستوى 0.05 تساوي 2-30.6

1- ضع الفرضيات

$$Ho : M \leq 17.5$$

$$H1 \quad : \quad M \quad > 17.5$$

2- اختبار الفرضية

$$t = \frac{\bar{y}-M}{S\bar{y}}$$

$$\sum yi = \quad y^2 i = 2963$$

$$163$$

$$\bar{y} = \frac{\sum yi}{n} \qquad \bar{y} = \frac{163}{9}$$

$$= 18.1$$

$$SS= \sum y^2 i - \frac{(\sum yi)^2}{n}$$

$$SS= \ 2963 - \frac{(163)^2}{9} \ = \ 2963 - 2959.11 = 10.89$$

$$S^2 = \frac{SS}{d.f} = \frac{10.89}{9-1} = \frac{10.89}{8} = 1.36$$

$$S\sqrt{S^2} = \sqrt{1.36} = 1.17$$

$$t= \frac{\bar{y}-M}{S\bar{y}} = \frac{18.1-1.17}{\frac{1.17}{\sqrt{9}}} = 1.538$$

3- استخراج قيمة t الجدولية لمستوى معنوية $\propto$ = 0.05 ودرجة حرية11 =2.306

3- الاستنتاج :بما ان قيمة t المحسوبة 1.538 اقل من t الجدولية 2.306 لذا نقبل فرضية العدم ونرفض الفرضية البديلة اي ان ادعاء الشركة صحيح.

مستوى المعنوية ∝ / Rدرجة الحرية	0.1	0.025	0.05	0.01
1				
2				
3				
4				
5				
6				
7				
8			$\boxed{2.776}$	

المحاضرة الثانية عشرة

2- اختبار يتعلق بمتوسطين

أ- اختبار t للعينات المستقلة وهناك العدد من الافتراضات التي يقوم عليها اختبار t للعينات المستقلة:

1- ان العينتين تم اختيارها بشكل عشوائي من المجتمع الخاص لكل عينة

2- ان المجتمعان يتصفان بالسواء (التوزيع الطبيعي)

3- الملاحظات ، البيانات ، المشاهدات ، ضمن كل عينة مستقلة عن بعضها

4- العينات تم توزيعها بشكل عشوائي الى المجموعتين

5- لغرض تحديد العينتان متجانسة او غير متجانسة يجري استخدام فحص التجانس وعلى النحو التالي

$$F_{\text{المحسوبة}} = \frac{S^2\ larges}{S^2\ smallest}$$

$$F_{\text{المحسوبة}} = \frac{S^2\ L}{S^2\ s}$$

$$F_{\text{المحسوبة}} = \frac{\text{التباين الاكبر}}{\text{التباين الاصغر}}$$

تقارن F المحسوبة مع F الجدولية بدرجة حرية التباين الاكبر بالاتجاه الافقي وبدرجة حرية للتباين الاصغر بالاتجاه العمودي فأذا كانت F المحسوبة اصغر من F الجدولية فهناك تجانس العينتان واذا كانت F المحسوبة اكبر من F الجدولية فهناك عدم وجود تجانس العينتان .

اولاً :في حالة التجانس

مثال// في تجربة لمقارنة نسبة المواد الفعالة التي تستخدم في صنع العقاقير في صنفين من نبات الكزبرة الصنف المحلي والصنف الباكستاني تم اختيار 12 نباتاً من كل صنف وقدرت نسبة المواد الفعالة فيهما وكانت النتائج كما يأتي فهل يختلف الصنفان تبعاً لنسبة المادة الفعالة تحت مستوى احتمال 0.05 علماً ان قيمة t الجدولية تساوي 2.075 تحت مستوى احتمال 0.05 ودرجة حرية 22

الصنف المحلي ملغم	الصنف الباكستاني ملغم
12.5	9.4
9.4	8.4
11.7	11.6
11.3	7.2
9.9	9.7
8.7	7.0
9.6	10.4
11.5	8.2
10.5	6.9
10.6	12.7
9.6	7.3
9.7	9.2
124.8 $\bar{y}$= 10.4	108 $\bar{y}$= 9

الحل :-

1- اجراء اختبار التجانس

$F = \frac{S^2\ L}{S^2\ s} = \frac{\text{التباين الاكبر}}{\text{التباين الاصغر}}$

$S^2 = \frac{SS}{d.f}$

$S^2 1 = \frac{\sum y^2 i - \frac{(\sum yi)^2}{n}}{n-1} = \frac{1312 - \frac{(124.8)^2}{12}}{12-1}$

$S^2 1 = \frac{1312-1297.92}{11} = \frac{14.08}{11} = 1.28$

$S^2 2 = \frac{1010.64-972}{11} = \frac{38.64}{11} = 3.51$

المحسوبة $F = \frac{3.51}{1.28} = \ = 2.74$

F المحسوبة = 2.82

بدرجة حرية بالاتجاه الافقي = 11

بدرجة حرية بالاتجاه العمودي = 11

بما ان F المحسوبة اصغر من F الجدولية : العينتان متجانستان

2- وضع الفرضيات

Ho : M1 – M2 = 0

H1 : M1 – M2 ≠ 0

اختبار الفرضيات

$t = \frac{\bar{y}1 - \bar{y}2}{S(\bar{y}i - \bar{y}i)}$

$\overline{y1}$ **= الوسط الحسابي للعينة الاولى**

$\bar{y}2$ **= الوسط الحسابي للعينة الثانية**

الخطأ القياسي للفرق = $S(\bar{y}i - \bar{y}i)$

متوسط معاملتين

$$S(\bar{y}i-\bar{y}i)=\sqrt{\frac{2\,mse}{r}}$$

التباين المشترك $mse = S^2\text{P}=$

$$S^2\text{P}=\frac{SS1+SS2}{n1+n2-2}$$

او
$$S^2\text{P}=\frac{(\text{n1}-1)S_1^2+(\text{n2}-1)S_2^2}{n1+n2-2}$$

حسب القانون الاول

$$S^2\text{P}=\frac{14.08+38.64}{12+12-2} = S^2\text{P}=\frac{52.72}{22} = 2.40$$

$$S(\bar{y}i-\bar{y}i)=\sqrt{\frac{2\,mse}{r}}=\sqrt{\frac{2\times 2.40}{12}}=0.63$$

$$t=\frac{\bar{y}1-\bar{y}2}{S(\bar{y}i-\bar{y}i)}=t=\frac{10.4-9.0}{0.63}=2.2$$

ويمكن استخدام القانون التالي لايجاد t المحسوبة

$$t=\frac{\bar{y}1-\bar{y}2}{SP\sqrt{\frac{1}{n1}+\frac{1}{n2}}}$$

SP هوالانحراف المعياري المشترك للفرق بين معاملتين

$$\text{SP}=\sqrt{S^2\text{P}}=\sqrt{mse}$$

$$\text{SP}=\sqrt{2.40}=1.55$$

$$t=\frac{10.4-9}{1.55\sqrt{\frac{1}{12}+\frac{1}{12}}}=2.2$$

3- استخراج t الجدولية لمستوى معنوية مطلوبة ورجة حرية $n1+n2-2$ = درجة الحرية ومستوى المعنوية المطلوب 0.05 حيث كانت t الدولية = 2.074

4- الاستنتاج : بما ان t المحسوبة اكبر من t الجدولية لذا نرفض فرضية العدم ونقبل فرضية البديلة وبما ان معدل المواد الفعالة في الصنف المحلي اعلى من الصنف الباكستاني لذا نوصي بأستخدام الصنف المحلي لأستخلاص المواد الفعالة .

ثانياً : في حالة عدم التجانس

مثال// اراد احد التدريسين في كلية الصيدلة ان يدرس اثر طريقتين من التدريس هما طريقة A و B في التعليم المختبري على التحصيل عند عينة من الطلبة اللذين يعانون من مشكلات تحصيلية في درس كيمياء الادوية فأختار عينة مؤلفة من 20 طالباً قام بتوزيعهم بشكل عشوائي الى مجموعتين 10 طلاب لكل مجموعة ثم عرض المجموعة الاولى للطريقة A وعرض المجموعة الثانية للطريقة B وبعد ذلك طبق عليهم امتحان تحصيلياً في التعليم المختبري وحصل على البيانات التالية علماً ان الدرجة القصوى للامتحان 30 درجة ، فهل هناك فرق معنوي بين الطريقتين على امتحان التحصيل للمادة العلمية اختبر ذلك تحت مستوى احتمال 0.01 ؟

المجموعة A	المجموعة B
06	12
05	13
04	28
04	05
07	10
04	18
05	23
05	06
07	04
06	30
53	149

$$\bar{y}A = \frac{53}{10} = 5.3$$

$$\bar{y}B = \frac{149}{10} = 14.9$$

$$S^2A = \frac{293 - 280.9}{9} = \frac{12.1}{9} = 1.34$$

$$S^2B = \frac{3027 - 2220.1}{9} = 89.7$$

المحسوبة F = $\frac{89.7}{1.34}$ = 66.94

الجدولية F=3.35

بما ان F المحسوبة اكبر من F الجدولية : العينتين غير متجانستين

1- وضع الفرضيات H0: M1 – M2 =0

H1 : M1 – M2 ≠0

2- اختبار الفرضية

$$t = \frac{\bar{y}1-\bar{y}2}{\sqrt{\frac{S_1^2}{n1}+\frac{S_2^2}{n2}}} = \frac{5.3-14.9}{\sqrt{\frac{1.34}{10}+\frac{89.3}{10}}} = \frac{-9.6}{3.02} = -3.18$$

نأخذ القيمة المطلقة

$$d.f = \frac{\left(\frac{S1^2}{n1}+\frac{S2^2}{n2}\right)^2}{\frac{\left(\frac{S1^2}{n1}\right)^2}{n1-1}+\frac{\left(\frac{S2^2}{n2}\right)^2}{n2-1}}$$

$$d.f = \frac{\left(\frac{1.34}{10}+\frac{89.7}{10}\right)^2}{\frac{\left(\frac{1.34}{10}\right)^2}{9}+\frac{\left(\frac{89.7}{10}\right)^2}{9}} = \frac{(9.104)^2}{0.001+8.940} = \frac{82.88}{8.94}$$

d.f = 9.27 ≃ 9

t الجدولية عند مستوى 0.01 ودرجة حرية 9= 3.25

بما ان t المحسوبة اصغر من t الجدولية نقبل فرضية العدم ونرفض البديلة اي لا يوجد فرق معنوي بين الطريقتين على امتحان التحصيل للمادة العلمية لمادة كيمياء الادوية .

المحاضرة الثالثة عشرة

ب - اختبار t للعينات المرتبطة

يستخدم اختبار t في هذه الحالة لاختبار الفروقات بين معاملتين مطبقة على نفس العينة او الوحدة التجريبية وتشكل البيانات بشكل ازواج تسجل قبل وبعد تطبيق المعاملة ، غالباً ما تستخدم في الدراسات الطبية وفي هذه الحالة نستخرج الفروقات بين ازواج المشاهدات وتعامل كعينة واحدة .

مثال // في تجربة لدراسة تأثير غذاء معين مع برنامج لإجراء بعض التمارين الرياضية لتقليل مستويات الكوليسترول بالدم ،طبق هذا البرنامج وكانت النتائج كالآتي ،اختبر ذلك تحت مستوى احتمال 0.01

ت	نسبة الكوليسترول قبل البرنامج y1	نسبة الكوليسترول قبل البرنامج y2	di
1	201	200	1
2	236	231	5
3	221	216	5
4	260	233	27
5	228	224	4
6	237	216	21
7	326	296	30
8	235	195	40
9	240	207	33
10	267	247	20
11	284	210	74
12	218	210	8
	2953	2685	268

الحل :

1- وضع الفرضيات

أ- ان الفرق بين مستوى الكولسترول قبل وبعد تطبيق البرنامج اكبر او يساوي صفر Ho: $M \geq 0$

ب- ان الفرق بين مستوى الكولسترول قبل وبعد تطبيق البرنامج اصغر Ho: $M < 0$

2- اختبار الفرضيات

t = $\frac{\acute{d}}{\acute{S}\, d}$

$\acute{d}$ = $\frac{\sum di}{n} = \frac{268}{12} = 22.33$

S^2 d= $\frac{\sum d^2 i - \frac{(\sum di)^2}{n}}{n-1}$ = $\frac{(1)^2+(5)^2+++(8)^2 - \frac{(268)^2}{12}}{11}$

= 434.61

S^2 d=$\frac{4780.67}{11}$

= 20.85

Sd= $\sqrt{434.61}$

$= \frac{20.85}{12}$ = 6.03

S $\acute{d}$=$\frac{Sd}{\sqrt{n}} =$

t = $\frac{\acute{d}}{\acute{S}\, d}$ = $\frac{22.33}{6.03}$ =3.70

3- استخراج t الجدولية لمستوى 0.01 ودرجة حرية 11 والتي تساوي 3.11

4- الاستنتاج : بما ان t المحسوبة 3.70 اكبر من t الجدولية 3.11 لذا نقبل الفرضية البديلة ونرفض فرضية العدم اي ان الفرق بين مستوى الكولسترول قبل وبعد تطبيق البرنامج اصغر من صفر لذا كان البرنامج فعالآ في تقليل مستوى الكولسترول بالدم .

مثال// اراد احد الباحثين الاطباء ان يعرف فيما اذا كان متوسط ضغط الدم في الانسان يختلف في حالة قياسه والشخص معتدل القامة عنه في حالة استلقاء الشخص نفسه على ظهره فأخذ عينة عشوائية مؤلفة من 12 شخص والنتائج التالية تبين الفرق بين ضغط الدم وهو في حالة وقوفه وضغطه وهو في حالة استلقاء على ظهره فماذا كان قراره تحت مستوى 0.05 ؟

di= -4 , 1 , 1 , -5 , -6 , -3 , 2 , -9 , 1 , -4 , -7 , -7

الحل

H0 : M1 – M2 = do = 0

H1 : M1 – m2 $\neq d0 \neq 0$

2- اختبار الفرضية

$$t = \frac{\acute{d} - do}{S\acute{d}}$$

$$\acute{d} = \frac{\sum di}{n} = \frac{-\acute{4}0}{12} = -3.33$$

$$S^2\ d = \frac{\sum d^2 i - \frac{(\sum di)^2}{n}}{n-1} = 14.06$$

$$Sd = \sqrt{d\, S^2} = \sqrt{14.06} = 3.57$$

$$S\,\acute{d} = \frac{3.75}{12} = 1.08$$

$$t = \frac{3.33}{1.08} = -3.08$$

3- استخراج قيمة t الجدولية لمستوى 0.05 ودرجة حرية 11 t الجدولية 2.201

4- الاستنتاج : بما ان t المحسوبة اكبر من t الجدولية 3.11 لذا نرفض فرضية العدم ونقبل الفرضية البديلة اي ان معدل ضغط الفرد يكون اعلى والشخص المستلقي على ظهره من ضغطه وهو في حالة الاعتدال والوقوف .

المحاضرة الرابعة عشرة

اختبار مربع كاي (X^2) Chi- Square :-

اختبار مربع X^2 شائع الاستخدام مع البيانات العددية المتقطعة والتي تكون نوعية اكثر منها كمية ، كعدد الذكور والاناث في عينة عدد الاصحاء وعدد المرضى في مجتمع او عدد الاحياء والاموات او الاجابة بنعم او بدون نعم .

يعتبر توزيع X^2 من التوزيعات المستمرة ويعتمد على التوزيع الطبيعي في حين تكون التوزيعات التكرارية غير مستمرة لذا يكون اختبار التكرارات المشاهدة مع التكرارات النظرية (المتوقفة) ذات دقة تقريبية .

$$X^2 = \sum \frac{(0-e)^2}{e}$$

حيث 0 قيم المشاهدات المشاهدة او الواقفة

حيث e قيم المشاهدات المتوقفة

ملاحظة// تكون قيم X^2 صغيرة عندما تكون قيم المشاهدات المتوقفة قريبة جداً من قيم المشاهدات المشاهدة او الواقعة وكذلك لا تكون قيم X^2 سالبة .

وتقارن قيم كاي سكور الجدولية والتي تستخرج على اساس درجات الحرية ومستوى المعنوية المطلوبة فاذا كانت قيمة X^2 المحسوبة اكبر او تساوي X^2 الجدولية نرفض فرضية العدم ونقبل الفرضية البديلة اي هناك فروقات معنوية .

استخدامات اختبار X^2 :-

1- استخدام كاي سكور (X^2) لجودة المطابقة :

اي المطابقة بين القيم المشاهدة (او الملاحظة) والقيم المتوقفة وهذا الاختبار يقوم على اساس ان القيم المشاهدة (او الملاحظة) لها نفس توزيع القيم المتوقفة كما ويفيد بصورة خاصة لاختبار البيانات الوراثية لجودة تطابق انعزالات الجيل الثاني .

مثال// اذا كان عدد الذكور في مرحلة معينة من مراحل الدراسة في كلية الصيدلة 70 طالباً وعدد الاناث 90 طالبة هل ان عدد الذكور الى عدد الاناث متساوية ، اختبر ذلك تحت مستوى 0.05 علماً ان قيمة X^2 الجدولية = 3.84 ؟

الحل :

1- وضع الفرضيات : بتوزيع الطلاب حسب الجنس بالتساوي :Ho

لا يتوزع الطلاب حسب الجنس بالتساوي : H1

حساب التكرار المتوقع 70+90 = 160 حيث 160 = n (عدد افراد العينة)

التكرار المتوقع : $n \times \frac{1}{2}$

: $80 = 160 \times \frac{1}{2}$

القيمة المتوقعة للذكور = 80

القيمة المتوقعة للاناث = 80

2- ايجاد قيمة X^2 $X^2 = \sum \frac{(0-e)^2}{e}$

ملاحظة// بما ان التكرارات المشاهدات قيم متقطعة وبذلك تعطي قيم متقطعة (احصاءات متقطعة) لذلك يطرح ما يسمى بمعامل yeates والذي يساوي 0.5 حيث ان القيم المتقطعة لا تتطبق على توزيع X^2 الذي يكون مستمراً او قريباً منه

$$X^2 = \sum \frac{(0-e)^2}{e}$$
$$X^2 = \frac{[170-801-0.5]^2}{80} + \frac{[190-801-0.5]^2}{80}$$
$$X^2 = \frac{[905]^2}{80} + \frac{[905]^2}{80}$$
$$X^2 = 1.128 + 1.128 = 2.256$$

3-استخراج قيمة X^2 الجدولية لدرجة حرية n-1 هي ان n تساوي 2 ذكور واناث ودرجة الحرية 1 ومستوى معنوية 0.05 حيث قيمة X^2 الجدولية تساوي 3.84

القرار : لما كانت قيمة X^2 المحسوبة اقل من قيمتها في الجدول لذا يمكن الاستنتاج بأن القيم المتوقفة لا تختلف عن المشاهدة اي ان عدد الذكور مشابهة لعدد الاناث وبدون فرق معنوي وما موجود من فرق بينهم يرجع الى عامل الصدفة .

مثال// في تزاوج بين نباتين احدهما قصير والاخر طويل ظهرت نتائج الجيل الثاني 30 نبات طويل و20 نبات قصير ، بين الى اي نسبة تنتسب هذه النباتات ؟ اختبر تحت مستوى 0.05 علماً ان X^2 الجدولية 3.84 ؟

الحل :

نبات قصير نقي ← 11 × LL → نبات طويل نقي

L1 × L1 الجيل الاول

L1 × L1 نضرب الجيل الاول

$\frac{11}{1}$ $\frac{LL\ L1L1}{31}$ **الجيل الثاني**

في حالة سيادة صفة الطول

$\frac{11}{1}$ $\frac{LL}{1}$ **في حالة عدم سيادة صفة الطول**

وضع الفرضيات أ- تكون نسبة الانعزال 1:1 H0

ب- تكون نسبة الانعزال 3:1

أ- لا تكون نسبة الانعزال 1:1 H1

ب- لا تكون نسبة الانعزال 3:1

أ- افتراض ان نسبة الانعزال بنسبة 1:1 فتكون القيم المتوقعة

$\frac{1}{2} \times 50 = 25$

القيمة المتوقعة للنباتات الطويلة =25

القيمة المتوقعة للنباتات القصيرة = 25

	النباتات الطويلة	النباتات القصيرة	المجموع
القيم المشاهدة	30	20	50
القيم المتوقعة	25	25	50

انحراف القيم المشاهدة عن المتوقعة $|25 - 30|$ $|25 - 20|$

بطرح معامل yeats [151 – 0.5] [151 – 0.5]

$$X^2 \qquad \frac{(4.5)^2}{25} + \frac{(4.5)^2}{25}$$

$$X^2 \qquad \frac{\frac{20.25}{25}}{\text{النباتات الطويلة}} + \frac{\frac{20.25}{25}}{\text{النباتات القصيرة}}$$

X^2 **1.62=** $0.81 + 0.81$

2- X^2 الجدولية تحت مستوى 0.05 ودرجة حرية تساوي 1 هي (X^2 الجدولية تساوي (3.84)

لما كانت قيمة كاي سكور المحسوبة (1.62) اقل من قيمة X^2 الجدولية فهذا يدل على عدم وجود فروقات بين قيم المشاهدة والمتوقعة وما موجود من فرق يعود للصدفة اي النسبة المتوقعة لها هي 1:1

ب- اذا افترضنا ان الانعزال بنسبة 3:1 فأن قيمة X^2

القيمة المتوقعة للنباتات الطويلة 37.5 = 50 × $\frac{3}{4}$

القيمة المتوقعة للنباتات القصيرة 12.5 = 37.5 – 50

	النباتات الطويلة	النباتات القصيرة	المجموع
القيم المشاهدة	30	20	50
القيم المتوقعة	37.5	12.5	50

انحراف القيم المشاهدة عن المتوقعة │30 - 37.5│ │20 - 12.5│

بطرح معامل yeates [│7.5│ – 0.5] [│7.5│ – 0.5]

$$X^2 \qquad \frac{(7)^2}{12.7} + \frac{(7)^2}{37.5}$$

X^2 **5.23 =** $3.92 + 1.31$

X^2 الجدولية = 3.84

كون X^2 المحسوبة اكبر من X^2 الجدولية في هذه الحالة نرفض فرضية العدم التي وضعناها والتي تنص على الانعزال بنسبة 1:3 لوجود فرق معنوي بين النسبتين لدى مقارنتها معع قيمة X^2 الجدولية .

2- اختبار X^2 للاستقلال *Test intendance*

يستخدم X^2 لأختبار الفرضيات الموضوعة على اساس وجود معيارين من التصنيف لمكونات المجموعة لتحديد فيما اذا كان هناك ارتباط بين الصفتين او المعيارين ام انهما مستقلان .

حيث *r* هي عدد الصفوف تمثل مستويات مختلفة لاحد معايير التصنيف

حيث *c* هي تمثل الاعمدة وتمثل مستويات مختلفة للمعيار الاخر

درجات الحرية *d.f* = (*r-1*) (*c-1*)

وتستخدم المجاميع الحرية للفئات التي توزع عليها الصفات في تحديد التكرارات المتوقعة فأن هذه المجاميع يجب ان تعتبر ثوابت

فمثلاً // درجة الحرية لجدول التوافق $2X^2$

(*2-1*) (*2-1*)=1

كما يعتبر اختبار الاستقلال مقيداً لزوجين من العوامل صفات اختبار X^2 للاستقلال والتي تميزه عن الاختبارات الاخرى

1. **تسحب العينة من المجتمع موضوع الدراسة والتي تصنف فيه الفئات وفق المتغير التابع والمتغير المستقل على اساس اهمية المتغيرين**
2. **يعتمد حساب نسب التكرارات المتوقعة لكل فئة على قانون الاحتمال الذي [ينص على انه اذا وقع حدثان والحادث هنا معيار التصنيف بصورة مستقلة الواحد عن الاخر فأن احتمال حدوثهما معاً يكون مساوياً لحاصل ضرب احتمال كل منهما على انفراد .**

قانون ضرب الاحتمال للاحداث المستقلة *P(AB) = P(A)× P(B)*

$$\text{الاحتمال} = \frac{\text{عدد الحالات المؤاتية}}{\text{عدد الحالات الممكنة}}$$

3- توضع الفرضية على اساس المتغيرين مستقلين عن بعضهما .

مثال// درس مجموع من الباحثين العلاقة بين مجاميع الدم وشدة الاصابة بحالة مرضية معينة لمجتمع جمعت بيانات من (1500) شخصاً وكانت النتائج كما مبين ادناه ، هل الحالتين مرتبطتان ، اختبر ذلك تحت مستوى احتمال 0.05 علماً ان X^2 الجدولية = 12.592 ؟

الحالة المرضية					Total
	A	B	AB	0	
غير مصاب	543 (541.2)	211 (212.96)	90 (92.40)	476 (473.44)	1320
متوسط الاصابة	44 (43.05)	22 (16.94)	8 (7.33)	31 (73.66)	105
شديد الاصابة	28 (30.75)	9 (12.10)	7 (5.25)	31 (26.90)	75
Total	615	242	105	538	1500

الحل :

وضع الفرضيات الصفتان مستقلتان : *H0*

الصفتان مرتبطتان :*H1*

القيم المتوقعة

قانون ضرب الاحتمال للاحداث المستقلة × المجموع العام

$$e11 = \frac{1320}{1500} \times \frac{615}{\cancel{1500}} \times \cancel{1500}$$

المجموع العام قانون ضرب الاحتمال للاحداث المستقلة

$$e11 = \frac{1320}{1500} \times 615 = 451.20$$

$$e12 = \frac{1320}{1500} \times 242 = 212.96$$

وهكذا لبقية القيم

$$X^2 = \sum \frac{(0-e)^2}{e}$$

$$X^2 = \frac{[543-541.2]^2}{541.2} + + + + \frac{[31-26.90]^2}{26.90} = 5.12$$

استخرج قيمة X^2 الجدولية لدرجة حرية (r-1) (c-1)

(3-1) (4-1)=6

قيمة X^2المحسوبة اقل من قيمة X^2 الجدولية لذا نقبل فرضية العدم ونرفض الفرضية البديلة التي تنص على ان شدة الاصابات ومجموع الدم مستقلة عن بعضها .

المحاضرة الخامسة عشرة

3- استخدام اختبار(X^2) لاختبار التجانس :-

يتصف اختبار الاستقلال اننا نسحب العينة من المجتمع قبل تصنيف الفئات وفقاً لمعياري التصنيف وهذا يعني ان العدد المشاهد للفئات يحدد بقدر سحب العينة ولذلك فأن مجاميع الصفوف والاعمدة تعتبر مقادير محتملة وليس تحت سيطرة الباحث وان العينة المحسوبة تحت هذه الضروف هي عينة مفردة تسحب من مجتمع واحد .

اما في حالة اختبار التجانس فأن الباحث قد يحدد العينات المستقلة تسحب من مجتمعات عديدة وفي هذه الحالة تكون واحدة من المجاميع الحدية ثابتة بينما تكون المجموعة الاخرى وفق معيار التصنيف المستخدم غير ثابتة

واختبار X^2 للتجانس تكون فرضية العدم H0 ان العينات المسحوبة من المجتمعات متجانسة والفرضية البديلة ان المجتمعات غير متجانسة .

مثال // درس باحث مدى استخدام عقار معين بين طلبة كلية الصيدلة الذين اعلنو عن استخدام الادوية ، واختار من هذه المجموعة عينة مكونة من 150 طالباً من الصف الاول و 135 طالباً من الصف الثاني و 125 طالباً من الصف الثالث و 100 طالباً من الصف الرابع واجاب كل طالب الاستفتاء عن مدى استخدام العقار هل هذه البيانات مطابقة او موافقة للفرضية بأن المجتمعات الاربعة متجانسة فيما يخص تناول العقار وكانت النتائج كما في الجدول التالي :

المراحل الدراسية	استخدام العقار			Total
	اختيار احياناً	متقطع	متقطع كثيراً	
صف اول	57 (63.24)	50 (51.47)	43 (35.27)	150
صف ثاني	57 (56.91)	58 (46.33)	20 (31.76)	135
صف ثالث	56 (52.70)	45 (42.89)	24 (29.41)	125
صف رابع	45 (42.16)	22 (34.31)	33 (23.53)	100
Total	215	175	120	510

الحل:

1- وضع الفرضيات   المجتمعات متجانسة :H0

المجتمعات غير متجانسة :H1

2- استخراج قيمة X^2 المحسوبة

$$X^2 = \sum \frac{(0-e)^2}{e} \quad \text{-3}$$

$$X^2 = \frac{[57-63.24]^2}{63.24} + + + + \quad \frac{[33-23.53]^2}{23.53} = 19.4$$

3- استخراج X^2 الجدولية بدرجة حرية (r-1)(c-1)

(4-1) (3-1) =6

ومستوى معنوية 0.01

لما كانت قيمةX^2 المحسوبة اكبر من X^2 الجدولية نرفض فرضية العدم ونقبل الفرضية البديلة اي المجتمعات الاربعة غير متجانس .

امثلة :

1- في دراسة فيما اذا كان هناك ارتباط بين الاصابة بالملاريا وتضخم الطحال وجدت البيانات التالية فهل هناك علاقة ارتباط بين الاصابة بالملاريا وتضخم الطحال اختبر ذلك تحت مستوى احتمال 0.01 علماً ان قيمة X^2 الجدولية تساوي 5.41 ؟

الاصابة بالملاريا	تضخم الطحال		*Total*
	+	-	
+	740 (546.45)	743 (936.55)	1483
–	1287 (1480.55)	2731 (2537.45)	4018
Total	2027	3474	5501

الصفتان مستقلتان عن بعضهما :H0

الصفتان مرتبطتان : H1

$X^2 = \sum \frac{(0-e)^2}{e}$ -1

$e11 = \frac{2027}{5501} \times 1483$ =546.45

$e12 = \frac{3474}{5501} \times 1483$ =936.55

$e21 = \frac{2027}{5501} \times 4018$ =1480.55

$e22 = \frac{3474}{5501} \times 4018$ =2537.45

$$\frac{[740-546.45]^2}{546.45} + \frac{[1287-1480.55]^2}{1480.55} + \frac{[2731-2537.45]^2}{2537.45}$$

$$X^2 = \frac{[740-546.45]^2}{546.45}$$

$$X^2 = 68.55 + 40 + 25.30 + 14.76 = 148.6$$

بما ان قيمة X^2 المحسوبة اكبر من X^2 الجدولية نرفض فرضية العدم ونقبل الفرضية البديلة اي الصفتان مرتبطتان هناك علاقة بين الاصابة بالملاريا وتضخم الطحال .

2-على فرض ان احد الباحثين اراد يجد العلاقة بين الجنسين والاصابة بالسرطان فأختار عينة مؤلفة (15) فرداً (8 ذكور ، 7 اناث) وقد حصل على البيانات التالية ، اختبر ذلك تحت مستوى احتمال 0.05 علماً ان قيمة X^2 الجدولية تساوي 3.84

الجنس	الاصابة بالسرطان		Total
	غير مصاب	مصاب	
ذكور	6 (4.77)	2 (3.20)	8
اناث	3 (4.2)	4 (2.80)	4018
Total	9	6	5501

H0: الصفتان مستقلتان

H1: الصفتان مرتبطتان

$$X^2 = \sum \frac{(0-e)^2}{e}$$

$e11 = \frac{8}{15} \times 9$ =4.77

$$e12 = \frac{8}{15} \times 6 = 3.20$$

$$e21 = \frac{9}{15} \times 7 = 4.2$$

$$X^2 = \frac{[6-4.77]^2}{4.77} \; \frac{[2-3.20]^2}{3.20} + \frac{[3-4.2]^2}{4.2} + \frac{[4-2.80]^2}{2.80}$$

$$X^2 = 0.32 \;\; 0.45 + 0.34 + 0.51 = 1.62$$

بما ان X^2 المحسوبة اقل من X^2 الجدولية نقبل فرضية العدم ونرفض الفرضية البديلة اي الصفتان مستقلتان عن بعضهما .

المحاضرة السادسة عشرة

الارتباط والانحدار Correlation Regression

لقد كان في اهتمامنا في الاختبارات السابقة حول قضايا الاحصاء الاستنتاجي التي تعود الى متغير واحد اما الان سوف نتحول الى القضايا التي تخص توزيع ذو متغيرين وسنرمز لهذين المتغيرين بالرمز y , x

الارتباط Correlation : هو الاسلوب الذي يفسر درجة وقوة واتجاه العلاقة بين المتغيرين y , x دون النظر الى السببية بينهما فقط يرتبط هذين المتغيرين بعلاقة خطية او غير خطية وقد لا تكون بينهما اي علاقة على وجه الاطلاق فمثلاً لانتوقع ان تكون هناك علاقة بين طول الفرد (X) وعمر والده بينما نتوقع ان تكون هناك علاقة بين طول الفرد (x) ووزنه (y) وسوف نتناول الارتباط البسيط وان كلا المتغيرين (y , x) هما متغيرين مستقلين وان كلاهما يتبع التوزيع الطبيعي وتوضح احد الفرضيات التالية عندما تكون مشاهدات مزدوجة اي قيم y , x

1- عدم وجود علاقة بين المشاهدات وتحلل بصورة منفردة او منفصلة اي نقصد عدم وجود علاقة بين مشاهدات y , x

2- وجود علاقة بينهما وتحدد هذه العلاقة بأستخدام الارتباط

3- تقدير مقدار هذه العلاقة يستخدم تحليل الانحدار

ولمعرفة فيما هناك علاقة بين متغيرين ام يحسب بما يسمى بمعامل الارتباط وسنرمز له بالرمز r

ان معامل الارتباط يوضح العلاقة الخطية بين متغيرين وكمثال على وجود الارتباط الخطي البسيط بين متغيرين مستقلين هو عند دراسة العلاقة بين طول الاخ والاخت في عدة عوائل ففي هذه الحالة لا توجد علاقة وآليه بين المتغيرين لان التغير في طول الاخ لا يسبب تغير في طول الاخت لان كلا المتغيرين مستقلين لكون طول الاخ والاخت يتغيران سوية تبعاً لتغير طول الاباء هذا ويجب التأكيد ان يكون هناك تقريباً منطقياً لاختبار المتغيرين فلم نجد تغييرات الى الترابط بين التدخين والدرجة الامتحانية ، ان قيمة معامل الارتباط تتراوح بين (-1) و (+1)

$$-1 \leq r \leq 1$$

$$r = \frac{\sum xi\, yi - \frac{(\sum xi)(\sum yi)}{n}}{\sqrt{(\sum x^2 i - \frac{(\sum xi)^2}{n})(\sum y^2 i - \frac{(\sum yi)^2}{n})}}$$

وعندما يكون الارتباط الخطي ضعيفاً فان معامل الارتباط r يقترب من الصفر وعندما لا يوجد ارتباط تكون قيمة r= صفر وعندما يكون هناك ارتباط موجب فأن r تقترب من +1 وهذا يعني الزيادة والنقصان في احد المتغيرين يتبعها زيادة او نقصان في المتغير الاخر يعني ارتباط طردي .

وعندما تعتبر r من 1- تعني الزيادة او النقصان في احد المتغيرين يصاحبهما نقصان او زيادة في المتغير الاخر على التوالي علاقة الارتباط عكسي.

يدعى مربع r معامل التحديد r^2 او ما يسمى بالقدرة التنبؤية وهو نسبة مربعات الانحدار SSR الى مجموع المربعات الكلية

$$r^2 = \frac{SSR}{SST}$$

مثال// فيما يلي اوزان واطوال عشرة اشخاص هل توجد علاقة بين اوزانهم واطوالهم اختبر ذلك تحت مستوى 0.05 علماً r الجدولية تساوي 0.63

xiyi	xi^2	yi^2	الطول سم xi	الوزن كغم yi
8448	16384	4356	128	66
9588	19881	4624	141	68
7552	13924	4096	118	64
10710	23409	4900	153	70
9522	19044	4761	138	69
12410	28900	5329	170	73
9180	18225	4624	135	68
8710	16900	4489	130	67
8125	15625	4225	125	65
12024	27889	5184	167	72
96269	200181	46588	1405	682

1- وضع الفرضيات H0: r= 0

$$H1 : r \neq 0$$

2- حساب معامل الارتباط

$$r = \frac{\sum xi\ yi - \frac{(\sum xi)(\sum yi)}{n}}{\sqrt{(\sum x^2 i - \frac{(\sum xi)^2}{n})(\sum y^2 i - \frac{(\sum yi)^2}{n})}}$$

$$r = \frac{96269 - \frac{(1405)(682)}{10}}{\sqrt{(200181 - \frac{(1405)^2}{10})(46588 - \frac{(682)^2}{10})}}$$

$$r = \frac{96269 - 95821}{\sqrt{(200181 - 1974025)(46588 - 46512.4)}}$$

$$r = \frac{448}{\sqrt{27785 \times 75.6}}$$

$$r = \frac{448}{\sqrt{2100546}} = \frac{448}{1449.33} = 0.31$$

3-الاستنتاج

لما كانت r المحسوبة اقل من r الجدولية لذا نقبل فرضية العدم ونرفض الفرضية البديلة اي توجد علاقة ارتباط بين وزن الطلاب واطوالهم ولكن هذه العلاقة لم تصل الى المستوى المعنوي وهي علاقة طردية ضعيفة .

المحاضرة السابعة عشرة

مثال// الجدول التالي يحوي على بيانات الكميات الناتجة في سلسلة من التفاعلات اجريت في درجة حرارة مختلفة المطلوب ايجاد معامل الارتباط بين درجات الحرارة وكمية الناتج من التفاعل بغية ا

لتعرف على مدى ارتباط الكميات الناتجة بدرجة الحرارة اختبر ذلك تحت مستوى احتمال 0.01 علما r الجدولية تحت مستوى المعنوية 0.01 ودرجة حرية 5= تساوي 0.87 ؟

xiyi	xi^2	yi^2	الكمية الناتجة yi من التفاعل	xi درجة الحرارة
615	1681	225	41	15
800	1600	400	40	20
950	1444	625	38	25
1050	1225	900	35	30
1120	1024	1225	32	35
1200	900	1600	30	40
1000	400	3500	20	50
6735	8274	7475	236	215

1- وضع الفرضيات H0: r= 0

H1 : r $\neq$ 0

2- حساب معامل الارتباط

$$r = \frac{\sum xi\, yi - \frac{(\sum xi)(\sum yi)}{n}}{\sqrt{(\sum x^2 i - \frac{(\sum xi)^2}{n})(\sum y^2 i - \frac{(\sum yi)^2}{n})}}$$

$$r = \frac{6735 - \frac{(215)(236)}{7}}{\sqrt{(7475 - \frac{(215)^2}{7})(8274 - \frac{(236)^2}{7})}}$$

$$r = \frac{6735 - 7248.57}{\sqrt{(7475 - 6603.57)(8274 - 7956.57)}}$$

$$r = \frac{-513.57}{\sqrt{871.43 \times 317.43}}$$

$$r = \frac{-513.57}{525.94} = 0.98$$

3-الاستنتاج

بما ان القيمة المطلقة r المحسوبة 0.98 اكبر من r الجدولية هناك ارتباط معنوي قدره 0.98 بين درجة الحرارة والكمية الناتجة من التفاعل وهذه العلاقة عكسية اي كلما زادت درجة الحرارة قلت كمية المواد الناتجة من التفاعل .

الانحدار Regression:

يعرف الانحدار مقدار التغير في المتغير y والذي يسمى المتغير التابع نتيجة تغير وحدة واحدة في المتغير x الذي يسمى المتغير المستقل

ففي حالة اجراء بحث نعين قيم المتغير المستقل مسبقاً فمثلاً تأثير درجة الامتحان الفصلي للطالب في مادة الاحصاء على درجة الامتحان النهائي ، تأثير سرعة نبضات القلب على القلق عند الكبار وتأثير عدد اطفال العائلة على معيار ذكاء الاطفال وتأثير الوزن على مستوى الكلوكوز بالدم وتأثير انقباض ضغط الدم على كمية الدم المفقودة خلال العمليات الجراحية وتأثير مكونات النظام الغذائي على مقياس الشحوم في البلازما وتأثير العمر على ضغط الدم وتأثير عقار معين على الانخفاض في النبض (ضربة/دقيقة) والانحدار يفرق عن الارتباط اننا نعرف ان هذا المتغير سوف يؤثر في المتغير الاخر بينما الارتباط ندرس فيما اذا كانت علاقة بين المتغيرين.

ان قيمة معامل الانحدار يعبر عنها بنفس الوحدات المستخدمة للصفة وتاخذ قيم سالبة او موجبة وقيمة الانحدار تأخذ قيمة غير محدودة وعندما تكون :

أ- قيمة الانحدار موجبة يعني كل زيادة في قيم x يتبعها زيادة في قيم y او كل نقصان في قيم x يتبعها نقصان في قيم y

ب- قيمة الانحدار سالبة فأن كل زيادة في x يتبعها انخفاض في قيمة y

ومن خلال معادلة الخط المستقيم يمكن ان نتنبأ بالمتغير التابع y بدلالة المتغير المستقل x مثلاً التنبأ بضغط الدم المتغير التابع من خلال العمر المتغير المستقل .

$$\hat{y} = a + bx$$

حيث ان $\hat{y}$ هي القيمة المتوقعة للمتغير التابع

a هي نقطة تقاطع خط الانحدار مع المحور الصادي

b معامل الانحدار

X المتغير المستقل

$$a = \bar{y} - \overline{bx}$$

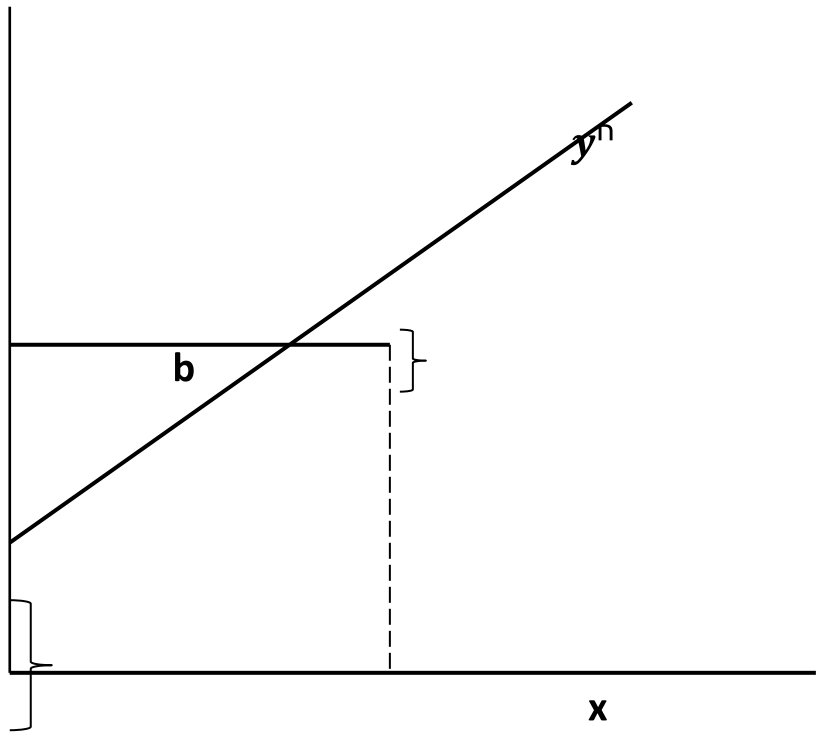
$$b = \frac{\sum xi\, yi - \frac{(\sum xi)(\sum yi)}{n}}{\sum x^2 i - \frac{(\sum xi)^2}{n}}$$ معامل الانحدار b

y

$$\hat{y} = a + bx$$

2

1 a



1 2

مثال// من البيانات التالية التي تمثل الدرجة الفصلية والدرجة النهائية لأثنى عشر طالباً في مادة الاحصاء من طلبة كلية الصيدلة

1-اوجد معامل الارتباط واختبر معنويته علماً ان r الجدولية = 0.71

2- اوجد معامل الانحدار واختبر معنويته على مستوى احتمال 0.01 وكذلك معامل الارتباط عن نفس مستوى المعنوية

3- تنبأ بدرجة الطالب النهائية في مادة الاحصاء علماً ان درجة الفصلية a=65 b= 50 c= 70

الدرجة الفصلية xi	الدرجة النهائية yi	xi^2	yi^2	xiyi	المقدرة y $y^{\cap}$
65	85	4225	7225	5525	88.4
50	74	2500	5476	3700	74.9
55	76	3025	5776	4180	79.4
65	90	4225	8100	5850	88.4
55	85	3025	7225	4675	79.4
70	87	4900	7569	6090	92.8
65	94	4225	8836	6110	88.4
70	98	4900	9604	6860	92.8
55	81	3025	6561	4455	79.4
70	91	4900	8281	6370	92.8
50	76	2500	5776	3800	74.90
55	74	3025	5476	4070	79.4
725	1011	44475	85905	61685	

1- معامل الارتباط

$$r = \frac{\sum xi\, yi - \frac{(\sum xi)(\sum yi)}{n}}{\sqrt{(\sum x^2 i - \frac{(\sum xi)^2}{n})(\sum y^2 i - \frac{(\sum yi)^2}{n})}}$$

$$r = \frac{61685 - \frac{(725)(1011)}{12}}{\sqrt{(44475 - \frac{(725)^2}{12})(85905 - \frac{(1011)^2}{12})}}$$

$$r = \frac{61685 - 61081.25}{\sqrt{(44475 - 43802.08)(85905 - 85176.75)}}$$

$$r = \frac{603.75}{\sqrt{672.92 \times 723.25}}$$

$$r = \frac{603.75}{700.4} = 0.86$$

الاستنتاج

بما ان r المحسوبة اكبر من r الجدولية هناك ارتباط معنوي وطردي .

-2

$$b = \frac{\sum xi\ yi - \frac{(\sum xi)(\sum yi)}{n}}{\sum x^2 i - \frac{(\sum xi)^2}{n}}$$

$$b = \frac{61685 - \frac{(725)(1011)}{12}}{44475 - \frac{(725)^2}{12}}$$

$$b = \frac{61685 - 61081.25}{44475 - 43802.08}$$

$$b = \frac{603.75}{672.92} = 0.897$$

وبالتعويض في معادلة خط الانحدار نحصل على قيمة $y^{\cap}$

$$y^{\cap} = a + bx$$

$\bar{x}$ **=60.417** $\quad \bar{y} = 84.220$

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

a= 84.220 – (0.897)(60.417)

a= 30.056

اوجد قيمة y عندما تكون x=65

50 =X

70 =X

$$y^{\cap}65 = 30.05 + 0.897 \times 65 = 88.40$$

$$y^{\cap}50 = 30.05 + 0.897 \times 50 = 74.9$$

$$y^{\cap}70 = 30.05 + 0.897 \times 70 = 92.8$$

وعند رسم خط الانحدار بتحديد نقطتين ولتكن احدهما $(\bar{x}, \bar{y})$ ونحدد نقطة اخرى على ان تكون قيمة x محددة ونستخرج قيمة y حسب المعادلة $y^{\cap} = a + bx$

ولاختبار معنوية معامل الانحدار يستخدم تحليل البيانات

H0: b =0

H1 : b $\neq 0$

المحاضرة الثامنة عشرة

A Nova Table

S.o .v	درجات الحرية df	SS مجموع المربعات	التباين MS	F cal	F table
egression	n-1 ونقصد n هنا عدد المتغير وهي y ,x	SSR= $b^2 \times (\sum x^2 i - \frac{(\sum xi)^2}{n})$	$\frac{SSR}{d.f}$	$\frac{MSR}{Mse}$	من جدول F بدرجة حرية R بالاتجاه الافقي ودرجة حرية الخطأ بالاتجاه العمودي
Residual الخطأ التجري error	n-2 ونقصد n هنا عدد المشاهدات لكل متغير	Sse = SST - SSR	$\frac{sse}{d.f}$		
Total	n-1 n هنا هي عدد المشاهدات	SST =$\sum y^2 i - \frac{(\sum yi)^2}{n}$			

SSR = $b^2 SSx = (0.897)^2 \times (672.92) = 541.44$

SSX =$\sum x^2 i - \frac{(\sum xi)^2}{n}$

SSX =$44475 - \frac{(725)^2}{12}$ = 672.92

SST = SSy =$\sum y^2 i - \frac{(\sum yi)^2}{n}$

SST = 85905 $- \frac{(1011)^2}{12}$ =728.25

SSe= SST = SSR

Sse = 728.25 – 541.44 = 186.81

A Nova Table

S.o .v	df	SS	MS	Fcal	F table
egression الانحدار	1	541.44	541.44	28.98	10.04 الجدوليةF عند مستوى 0.01 ودرجة حرية 10 للخطأ و 11 للانحدار
Residual الخطأ التجريبي error	10	186.81	18.681		
Total	11	728.25			

بما ان F المحسوبة اكبر من F الجدولية لذا نرفض فرضية العدم فأن معامل الانحدار معنوي يختلف عن الصفر وهنا تجدر الاشارة ان نسبة مجموع المربعات للانحدار الى مجموع المربعات الكلية التي يمكن تفسيرها بالعلاقة الخطية الموجودة بين المتغيرين y ,x وعادة يرمز لها بالرمز r^2 يسمى معامل التحديد او القدرة التنبؤية اما نسبة مجموع المربعات التي لا تفسر العلاقة الخطية بين المتغيرين y,x والتي تعزى الى الخطأ العشوائي ويطلق عليها بمعامل عدم الارتباط ويرمز لها K

K=1 - r^2

$r^2 = \frac{SSR}{SST} = \frac{541.44}{728.25}$ **وفي المثال السابق**

$r^2 = 0.74$

اي ان 75% من الاختلافات الكلية في قيم y تعود الى وجود العلاقة الخطية بين المتغيرين y , x وان 0.74 – 1 =26% من الاختلافات الكلية تعود الى الخطأ العشوائي

$r^2 = (r)^2$

$$\boldsymbol{r^2 = (0.86)^2 = 0.74}$$

المحاضرة التاسعة عشرة

بعض المفاهيم الاساسية في تصميم التجارب

1- التجربة Experiment :-هي الخطط التي ترسم مقدماً لتشكل اساس جيد ومأمون لكي يتم الحصول على معلومات جديدة لرفض او تأكيد فرضيات سابقة او استنتاج قواعد او قوانين جديدة او هي الاسلوب العلمي لاختيار الفرضيات واستكشاف العلاقات بين المتغيرات ويمكن تلخيص التجربة بالنقاط التالية :

أ- تحديد المشكلة :

ب- تحديد المتغير المتأثر او مايسمى بمتغير الاستجابة Response او المتغير التابع وهي القيمة او المستوى الذي يتم التوصل للمتغير او المتغيرات المدروسة لكل وحدة تجريبية على انفراد وعليه كمية المحصول للوحدة التجريبية هو استجابة للمعاملة المعنية على سبيل المثال عند دراسة تأثير مستويات عقار معين على معايير الدم في الانسان فمعايير الدم هي بمتغير الاستجابة .

جـ- تحديد العامل او العوامل التي سيجري تغيرها ويطلق عليها العوامل المؤثرة حيث يدرس تأثير عامل بعدة صيغ او مستويات مثال العامل المدروس او المؤثر الاصناف ، مواعيد الزراعة ، مستوى العقار حيث يدرس تأثير عامل واحد فقط مع ثبوت العوامل الاخرى تسمى التجارب البسيطة وحيث يدرس تأثير اكثر من عامل وتسمى بالتجارب العاملية.

2- التصميم Design

تصميم التجربة ببساطة يعني تخطيطها ، بحيث يصبح بالامكان جمع المعلومات المتعلقة بالمشكلة المراد دراستها لكي تضمن امكانية الحصول على البيانات المناسبة التي تسمح بتحليلها تحليلاً سليماً وموضوعياً للوصول الى استنتاجات صحيحة فيما يتعلق بالمشكلة .

وعند اجراء تصميم التجربة يجب تحديد ما يأتي :-

أ- عدد المشاهدات المطلوب تسجيلها وحجم العينة يجب ان يكون ملائم اي يجب ان يكون حجم العينة اكبر كي يؤدي الى قلة الخطأ.

ب- الاسلوب التجريبي الذي ستجري عليه التجربة

يجب ان يكون الاسلوب العشوائي (بدون تدخل شخص) من حيث ان العشوائية تميل لموازنة العوامل وبالاسلوب العشوائي الذي يقلل الخطأ العشوائي .

3- النموذج الرياضي :

يمكن للباحث بعد ان يتخذ الاسلوب التجريبي العشوائي من وضع نموذج يصف التجربة بحيث يظهر هذا النموذج المتغير المتأثر كدالة لكل العوامل التي سنذكر وهي فروض او قيود فرضية على التجربة كنتيجة لتطبيق الاسلوب العشوائي

4- التحليل الاحصائي Analysis

يعتبر التحليل المرحلة الاخيرة وفي هذه المرحلة يتم جمع البيانات ثم جدولتها ثم اجراء الاختبارات الاحصائية والرياضية ثم الوصول الى قرارات مفيدة لاختبار فرضيات متعلقة بالنموذج الرياضي يصف التجربة .

ويمكن تلخيص التحليل بثلاث نقاط :

1- جمع البيانات وجدولتها

2- اجراء الاختبارات الاحصائية

3- مناقشة النتائج وتفسيرها

الوحدة التجريبية Experimental Unit

هي اصغر وحدة اساسية او اصغر جزء في مواد التجربة تطبق عليها المعاملات او المعالجات وتستخدم في قياس المتغيرات تحت الدراسة وقد تكون الوحدة التجريبية حيواناً او نباتاً .

المعالجات او المعاملات Treatment

تمثل مجموعة الظروف التجريبية المتغيرة التي تخضع تحت سيطرة الباحث والذي يقوم بتوزيعها على الوحدات التجريبية حسب التصميم وقد تكون المعالجات عدة مستويات لعامل واحد تسمى بالتجارب البسيطة او تكون عدة مستويات لاكثر من عامل كما هو الحال في التجارب العاملية مثلاً في النبات نوع السماد (سماد عضوي ، سماد اليوريا ، سماد فوسفاتي) وكذلك كمية السماد ، عمق البذار ، ميعاد الزراعة وفي الحيوان تكون المعاملة نوع الحيوان ، جنس الحيوان ، سلالة الحيوان ، نوع الغذاء .

معاملة المقارنة Control treatment

عادة ما تشمل التجربة على معاملة او اكثر بهدف اتخاذها اساساً في جدوى اضافة عامل (مبيد ، تسميد ......الخ) فقد تشمل تجربة مخصصة لدراسة تأثير مبيد معين بتراكيز (2% ، 4% ، 6% ، 8%) اضافة الى معاملة اضافية تمثل عدم استخدام المبيد كي تكون اساساً لمعرفة جدوى اضافة المبيد من حيث المبدأ ، وكذلك عندما لا تعطي بعض الوحدات التجريبية الى كمية من السماد (اي صفر سماد) بينما تعطي مجموعات اخرى من الوحدات التجريبية مستويات محدودة من السماد قيد الدرس .

<u>التكرارات Replication</u>

تظهر المعاملة اكثر من مرة في التجربة ويطلق على ذلك التكرارات ويجب تكرار المعاملة الواحدة اكثر من مرة في التجربة كي يمكن تقدير قيمة الخطأ التجريبي وبالتالي فصله عن تأثير المعاملة حيث تمثل كل معاملة بوحدة تجريبية واحدة لا يعطي فكرة صحيحة عن تأثير المعاملة .

كما ان زيادة تكرار المعاملات يؤدي الى زيادة دقة كفاءة التجربة وذلك كنتيجة مباشرة لتقليل قيمة الخطأ لمتوسط المعاملة وان تكرار المعاملة يفيد في حالة فقد او تلف احدى القطع التجريبية .

المحاضرة العشرون

تحليل التباين Analysis of variance

ان مصطلح تحليل التباين يطلق على مدى واسع من الاساليب الاحصائية الفنية ويكاد اغلب الاحصائيين في المواضيع السابقة مثل اختبار t للتعرف على معنوية الاختلافات بين الوسطين الحسابيين لمجموعتين تجريبيتين وبالتالي التباين هو اسلوب احصائي يتم بواسطته مقارنة الاختلافات بين اكثر من متوسطين حسابيين تعود لاكثر من مجموعتين وبالتالي فأن تحليل يتم بتجزئة التباين بين مجموعة من المشاهدات يتم بواسطته تجزئة الاختلافات الكلية الى مكوناتها المختلفة وحسب مصادرها المعروفة وغير المعروفة وبالتالي الكشف عن وجود او عدم وجود فروق معنوية احصائية بين عدد من المتوسطات الحسابية ذلك من خلال اختبار المعنوية والتي اطلق Snedecor اسم اختبار t نسبة الى العالم Fisher ، وفي تحليل التباين يتم اختبار عدة متوسطات لعوامل او مجموعة عينات دفعة واحدة مما يسهل على الباحثين في ميادين البحوث التجريبية كالتجارب الزراعية والبايلوجية والصناعية مما يسهل العمل كثيراً .

فعلى سبيل المثال لو اردنا اختبار 4 متوسطات عينات دفعة واحدة بأستخدام اختبار t فأن ذلك يتطب كل زوج من متوسطات العينات على حدة بصورة عامة

(طرق الاختيار غير المرتب) n من r توافيق $C_r^n = \frac{n!}{r!(n-r)!}$

$$C_2^4 = \frac{4!}{2!\,(4-2)!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = 6$$

اي 6 اختبارات *t* للتعرف على الاختلافات بين العينات ولكن بأستخدام تحليل التباين يعني استخدام اختبار واحد لجميع البيانات او المتوسطات في نفس الوقت للحصول على استنتاج عام بوجود او عدم وجود اختلاف معنوي بين العينات وهذا يعتبر ادق وافضل من الاسلوب السابق ، اي ان التحليل يعني تجزئة الى مركباته وتحليل التباين عبارة عن عملية رياضية يقسم فيها التباين الكلي الى مكوناته المحتملة ويوضح ذك في جدول تحليل التباين *Analysis of variance* ويسمى *A Nova table*

خطوات تحليل التباين

y_{ij} = قيمة اي مشاهدة *y* في المعاملة *i* في التكرار

اي المشاهدة *y* في المعاملة *i* في التكرار *j*

حيث (من العاملات) *i= 1 ,2 , 3 , ... t*

(من التكرارات) *j= 1 ,2 , 3....... r*

<u>خطوات تحليل التباين :</u>

1- نستخرج قيمة معامل التصحيح *Correction Factor CF*

وذلك من خلال **C.F =**

$$\frac{y_{..}}{tr} = \frac{\text{مربع المجموع الكلي}}{\text{عدد المشاهدات الكلية}}$$

2- استخراج *SST* **SST=**

$$\sum y^2 ij - \frac{(y_{..})^2}{tr}$$

3- استخراج *SS treatment*

$$SSt = \frac{\sum y^2 i}{r} - \frac{(y_{..})^2}{tr}$$

4- استخراج مجموع المربعات للخطأ التجريبي

$$SSe = SST - SSt$$

5- عمل جدول تحليل التباين

6- استخراج قيمة *F* المحسوبة *F-calculated*

7- استخراج قيمة *F* الجدولية من جدول *F* اعتماداً على درجات الحرية للمعاملات بالاتجاه الافقي ودرجات حرية الخطأ التجريبي بالاتجاه العمودي ومستوى المعنوية 0.01 او اكثر 0.05

8- اذا ساوت *F* المحسوبة او اكبر من *F* الجدولية نستخدم احد طرق اختبار المعنوية للمتوسطات للتعرف على المعاملة ذات التأثير الاكبر لان هناك تأثير معنوي للمعاملة في الصفة المدروسة .

<u>الصيغة العامة لجدول تحليل التباين</u>

Analysis of Variance
A Nova table

مصادر الاختلاف Source of variation	*درجات الحرية* Degree of freedom df	*مجموع المربعات* Sum of Squares SS	*متوسطات المربعات التباين* Mean Squares MS	*قيمة* F*المحسوبة* Calculator	F table
يشمل جميع مسببات الاختلافات بين مواد التجربة للتصميم المستعمل في التجربة	وهي عدد المقارنات المستقلة يمكن اجراءها في كل مصدر من مصادر الاختلاف	مجموع مربعات الانحرافات المسؤول عنها كل مصدر في مصادر الاختلاف	هو التباين لكل مصدر من مصادر الاختلاف ويحسب بقسمة مجموع المربعات على درجات الحرية لكل مصدر	تحسب بقسمة تباين كل مصدر على تباين الخطأ التجريبي اي انها نسبة التباين	تستخرج من جدول *F* لكل مصدر اختلاف اعتماداً على مستوى المعنوية ودرجة حرية البسط بالاتجاه الافقي ودرجات حرية الخطأ بالاتجاه العمودي

اختيار التصميم التجريبي المناسب يعتمد على عدة امور هي:

أ- طبيعة تجانس الوحدات التجريبية في المجال المخصص لتنفيذ التجربة
ب- عدد المعاملات المشمولة بالتجربة

جـ- عدد العوامل المراد دراستها والاهمية النسبية لكل تصميم

2- التصميم العشوائي الكامل مع تسجيل اكثر من مشاهدة

Analysis of Variance with sub sample

في بعض تجارب التصميم العشوائي الكامل قد نسجل عدة مشاهدات في كل وحدة تجريبية طبقت عليها المعاملة وعلى ذلك فان المعاملة تحتوي على عدد من الوحدات التجريبية وقد ياخذ من كل وحدة تجريبية عدد من العينات العشوائية ونتيجة لذلك نلاحظ اضافة الى تباين المعاملات وتباين الخطأ التجريبي error

سوف يظهر تباين العينات حيث ان عدم حساب تباين العينات سوف يؤدي الى خطا في التحليل الاحصائي حيث ان تباين المعاملات سوف يزداد وتباين الخطأ كذلك وان هذا النوع من التجارب يسمى المتعشية ومن امثلة هذه التجارب الحقلية فقد لا يتوفر الوقت الكافي للباحث مثلا كي يحصد كلا من القطع التجريبية بأكمالها وعلى ذلك فقد يلجأ الباحث الى اختيار قطعة عشوائية من كل قطعة تجريبية ثم يحصر بيانات كل من الاجزاء المختارة وكذلك في تجارب تغذية الحيوانات قد تكون الوحدات التجريبية عدد من حظائر الحيوانات ويوجد في كل حظيرة عدد من الحيوانات التي نطلق عليها العينات يوزع على كل منها احد العلائق (المعاملات) حيث ان الحيوانات تمثل مشاهدات داخل الوحدة التجريبية التي تمثل الحظيرة .

اذا افترضنا t من المعاملات وكررت على عدد r من الوحدات التجريبية ثم سجلت المشاهدات على عدد S من العينات داخل كل وحدة تجريبية

i = 1 , 2 , ...t
j = 1, 2 , ...r
k= 1 , 2 , ...S

معادلة النموذج الرياضي yijk=M + ti + eij +Sijk

حيث yijk قيمة المشاهدة y في المعاملة i في الوحدة التجريبية t في العينة k

M المتوسط العام للتجربة

ti تأثير المعاملة ti = $\bar{y}..-\bar{y}$ .... ..

eij تأثير الخطأ العشوائي للوحدة التجريبية j في المعاملة i

eij =$\bar{y}ij-\bar{y}i$ ...

Sijk تأثير الخطأ العشوائي للعينة S في الوحدة التجريبية j في المعاملة i

Sijk =$\bar{y}ijk-\ \bar{y}\bar{ı}\bar{ȷ}$.

A Nova table

S.o.v	d.f	SS	MS	Fcal	F table
Treat	t - 1	$\frac{\sum y^2 i..}{rs}-\frac{(y..)^2}{trs}$	$\frac{SSt}{d.f}$	$\frac{MSt}{MSe}$	

Ex:p.uni error	T(r-1)	$\frac{\sum y^2 ij.}{s} - \frac{\sum y^2 i..}{rs}$	$\frac{SSe}{d.f}$	$\frac{MSe}{MSS}$	
Sampling error	tr(s-1)	SSs = SST- SSt - SSe	$\frac{SSs}{d.f}$		
		$SST = \sum y^2 ijk - \frac{(y..)^2}{trs}$			

مثال // في احدى تجارب تغذية الحيوانات وضعت 4 حيوانات اختيرت عشوائياً في كل 18 قفص وكانت المعاملات هي 3 انواع من العلائق تختلف في ما تحتويه من الذرة وقد وزعت هذه المعاملات الثلاث بطريقة عشوائية على الاقفاص (الوحدات التجريبية) حيث خصصت لكل عليقة ست حظائر وبعد فترة محددة سجلت الزيادة في اوزان الحيوانات وكانت موضحة كما في الجدول التالي ، هل هناك تأثير معنوي للعلائق في الزيادة الوزنية للحيوانات اختبر ذلك تحن مستوى احتمال 0.05 ؟

Treat المعاملات	الوحدات التجريبية الاقفاص		مجموع الوحدات التجريبية yij	مجموع المعاملات yi..
T1 8% corn	1	2.9 3.1 3.0 3.0	12	77.5
	2	3.2 3.3 3.2 3.1	12.8	
	3	3.2 3.3 3.4 3.3	13.2	
	4	3.2 2.8 2.9 3.0	11.9	
	5	3.3 3.5 3.4 3.4	13.6	
	6	3.4 3.5 3.5 3.6	14.0	
T2 8% corn	1	3.1 3.4 3.4 3.3	13.2	78.7
	2	2.9 2.9 3.0 3.1	11.9	
	3	3.3 3.4 3.2 3.2	13.1	
	4	3.7 3.5 3.6 3.5	14.3	
	5	2.9 3.1 2.8 3.0	11.8	
	6	3.7 3.6 3.5 3.6	14.4	
T3 0% corn	1	2.7 2.7 2.8 2.9	11.1	68.6
	2	3.1 2.9 3.0 3.1	12.1	
	3	2.7 2.7 2.6 2.7	10.7	
	4	2.8 2.9 2.9 3.0	11.6	
	5	3.2 3.0 3.0 3.2	12.4	
	6	2.8 2.7 2.6 2.6	10.7	

y...=(224.8)

$$SST = \sum y^2 ijk - \frac{(y..)^2}{trs}$$

$$SST = (2.9)^2 + (3.1)^2 + + + +(2.9)^2 - \frac{(224.8)^2}{3 \times 6 \times 4}$$

SST = 6.205

$$SSt = \frac{\sum y^2 i..}{rs} - \frac{(y..)^2}{trs}$$
$$= \frac{(77.5)^2 + (78.7)^2 + (68.6)^2}{6 \times 4} - \frac{(224.8)^2}{3 \times 6 \times 4}$$

SSt = 2.557

$$SSe = \frac{\sum y^2 ij.}{s} - \frac{\sum y^2 i..}{rs}$$
$$= \frac{(12.0)^2 + +(10.7)^2}{4}$$
$$- \frac{(77.5)^2 + + + (68.6)^2}{6 \times 4}$$

Exp.unet

SSe = 3.112

SSs = SST – SSt –Sse

SSs = 6.205 – 2.557 – 3.112 = 0.556

S.o.v	d.f	SS	MS	Fcal	F table
Treat	2	2.557	1.268	6.13	(0.05$\frac{2}{15}$)=3.17
Ex:p.uni error	15	3.112	0.207	20.70	(0.05$\frac{2}{45}$)=3.17
Sampling error	54	0.556	0.010		

	71	6.205			

الاستنتاج : هناك تأثير حقيقي بين المعاملات العلائق وهذه الاختلافات ليست بسيطة وانما اختلافات جوهرية

لتحقيق النموذج الرياضي **yijk=M + ti + eij +Sijk**

y111=2.9

M= 3.12

علماً ان متوسط المعاملة=3.22

ti = $\bar{y}\,..-\overline{y\,...}$

ti =$3.22\,..-3.12=0.10$

تأثير الخطأ العشوائي للوحدة التجريبية j في المعاملة i

علماً ان متوسط الوحدة التجريبية =3

eij =$\bar{y}ij-\bar{y}i$ ...

eij =$3-3.22=0.22$

علماً ان قيمة الوحدة التجريبية =2.9

Sijk =$\bar{y}ijk-\ \bar{y}\bar{\imath}j$.

Sijk =$2.9-3.0=0.1$

Sijk =$3.12+0.10-0.22-0.1$

Sijk =$3.22+0.32=2.9$