مدرس المادة :م.م. دعاء بشيرعباس

الاحصاء الوصفي

بعض المفاهيم الاساسية في علم الاحصاء

1- علم الاحصاء: هو العلم الذي يبحث في جمع البيانات وتنظيمها وعرضها وتحليلها بهدف الوصول الى النتائج واتخاذ القرارات السليمة .
2- المجتمع : هوعبارة عن جميع القيم او المفردات التي يمكن ان ياخذها المتغير ، والمجتمع ينقسم الى مجتمع محدود اي يمكن حصر عدد مفرداتة مثل اطوال طلبة كلية التربية للعلوم الانسانية.
اما المجتمع غير المحدود لايمكن حصر عدد مفرداتة مثلا عدد البكتريا في حقل ما.
3- العينة : عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع اي ان العينة هي جزء من المجتمع.
4- المتغير: هواي ظاهرة تظهر اختلافات بين مفرداتها ويرمز لة بالرمز X ،Y،....

المتغيرات نوعين هما

ا- المتغيرات نوعية (وصفية): هي تلك الظواهر او الصفات التي لايمكن قياسها مباشرة بالارقام العددية مثل صفة لون العيون(ازرق ، احضر، اسود) او الحالة الاجتماعية(متزوج ، اعزب،مطلق).
2- المتغيرات الكمية : هي الظواهر او الصفات التي يمكن قياسها مباشرة بالارقام كالطول او الوزن او كمية المحصول وتنقسم هذة المتغيرات الى نوعين هما:
1- متغير كمي متصل(مستمر): هوالمتغير الذي تاخذ المشاهدة او المفردة فية اي قيمة رقمية في مدى معين .

مثل: $0<x<1$

2- متغير كمي منفصل(غير مستمر): هوالمتغير الذي تاخذ المشاهدة او المفردة فية قيم متباعدة وغير مستمرة مثلاً:

$$X=1,2,3,4,....$$

مجالات علم الاحصاء

يعد الاحصاء احد الفروع ذات التطبيقات الواسعة لعلم الرياضيات وهو يبحث بشكل رسمي في الامور التالية:

1- جمع البيانات والمعلومات الخاصة بمختلف الظواهر وتسجيلها في صورة رقمية وتصنيفها في جداول منتظمة وتمثيلها بيانياً.
2- تحليل البيانات واستخلاص النتائج منها والخروج باستنتاجات حول المجتمع الاحصائي الذي يمثل جميع المفردات التي تتكون منها الظاهرة.
3- مقارنة الظواهر ببعضها ودراسة العلاقة بينهما واستخدامها في تفهم حقيقة الظواهر ومعرفة القوانين التي تسير تبعاً لها.

اهم فروع علم الاحصاء

1- الاحصاء الوصفي: وهو الاحصاء الذي يتناول تنظيم وعرض ووصف البيانات والمعلومات سواء كانت هذة المعلومات كمية كالمساحات والمسافات وكميات الانتاج اوكانت نوعية كانواع الترب او انواع المناخ او الحالة الاجتماعية للأفراد.
2- الاحصاء الاستدلالي : وهو يهتم بالعمليات التي تؤدي الى استخلاص النتائج العامة من النتائج الجزئية وفقا لقوانين احصائية معينة وتعميم النتائج التي يتم التوصل اليها من دراسة عينة مختارة من مجتمع البحث.

انواع البيانات الاحصائية

1- البيانات النوعية(الوصفيةاو الاسمية)Nominal Data:

نحصل على هذا النوع من البيانات عندما تكون السمة تحت الدراسة هي سمة نوعية والتي يمكن تصنيفها حسب الاصناف او الانواع وليس بقياسات عددية فمثلا تصنيف نوع فصيلة الدم وغيرها.

2- البيانات الترتيبية Ordinal Data

تمتاز البيانات الترتيبية عن البيانات الاسمية بان الفئات التي يتكون منها المجتمع الاحصائي تكون على هيئة رتب وليس مجرد فئات فمثلاً الرواسب النهرية تتصنف الى غليظة ومتوسطة ودقيقة وتتصنف المؤهلات العلمية الى دكتوراة وماجستير وبكلوريوس واعدادية ومتوسطة...الخ.

3- **البيانات الكمية Scale:**

هي البيانات التي يكون للأرقام فيها مدلول حسابي، اي انها تخضع للعمليات الحسابية العادية وهي افضل انواع البيانات واكثرها استخداماً في العملية الاحصائية مثلاً سنوات الخدمة والرواتب وغيرها.

تصنيف وتبويب البيانات وتمثيلها بيانياً

عند جمع البيانات الاولية الخاصة بدراسة ظاهرة او صفة ما فانة عادة لايمكن الاستفادة منها وهي بهذه الصورة لذا فغالباً ماتضع في جداول مبسطة او يعبر عنها في هيئة صور واشكال ورسوم بيانية لكي يسهل دراستها وتحليلها.

وهناك نوعين رئيسيين من الجداول الاحصائية هما:

1- **الجدول البسيط: وهو الجدول التي توزع فية البيانات حسب صفة واحدة ويتألف من عمودين احدهما بيمثل تقسيمات الظاهرة او الصفة الى مجموعات او فئات والعمود الاخر يمثل عدد المفردات او المشاهدات التابعة لكل مجموعة او فئة كما في المثال التالي:**

مثال// جدول توزيع اعضاء البعثات الموفدين الى الخارج حسب المواد الدراسية لسنة 1970-1971.

اسماء البعثات	عدد الطلبة
علوم اساسية	25
علوم زراعية	50
علوم هندسية	75
علوم طبية	50
علوم اجتماعية	30
علوم بيطرية	20
المجموع	250

2- الجدول المركب:هو الجدول الذي توزع فية البيانات حسب صفين او ظاهرتين في نفس الوقت ، فالجدول يتألف من صفوف واعمدة تمثل الصفوف المجاميع او الفئات احد الصفات وتمثل الاعمدة مجاميع او فئات الصفة الاخرى،اما المربعات التي تقابل الصفوف والاعمدة فتحتوي على عدد التكرارات او عدد المفردات المشتركة في فئات كلا الصفتين كما في المثال التالي:

مثال// الجدول التالي يبين مساحات وسكان بعض الدول العربية لعام 2015.

الدولة	المساحة كم2	عدد السكان
العراق	37072	36575000
الاردن	92300	9500000
الكويت	17820	4161000
سوريا	185180	23270000
مصر	1002000	90759000

الجداول التكرارية: وهو اسلوب احصائي مهم لختزال عدد كبير من البيانات الى عدد محدود من الفئات واحصاء عدد المشاهدات التي تقع ضمن حدود كل فئة ويتكون الجدول التكراري من عمودين:

العمود الاول: يقسم قيم المتغير او الظاهرة الى اقسام او مجموعات تدعى فئات Classes

العمود الثاني: يبين عدد المفردات التي تقع قيمها ضمن الفئة وتسمى تكراراتFrequence ويرمزلة بالرمز(f).

بعض المفاهيم الاساسية لموضوع الجداول التوزيع التكراري

1- البيانات غير المبوبة: وهي البيانات الاولية او الاصلية التي جمعت قبل تبويبها ووضعها في جدول توزيع تكراري.

2- البيانات المبوبة: وهي البيانات التي تم تبويبها ووضعها في جدول توزيع تكراري.

3- الفئات: وهي المجموعات التي تقسم قيم المتغير او الظاهرة مثل الفئة (10-15).

4- حدود الفئة : لكل فئة حدان حد ادنى وحد اعلى فالفئة (5-10) الحد الادنى لها 5 والحد الاعلى 10.

5- طول الفئة: هو المدى بين حدي الفئة ويرمز لة بالرمز L .

6- مركز الفئة: هو عبارة عن منتصف المدى بين حدي الفئة ويرمزلة بالرمز X، وقانونة هو:

$$X=\frac{\text{الحد الادنى الفئة+الحد الاعلى الفئة}}{2}$$

خطوات بناء جدول توزيع تكراري

1- استخراج المدى للبيانات وفق القانون التالي:

المدى = اكبر قيمة – اصغر قيمة

$$R= X_{max} - X_{min}$$

2- ايجاد عدد الفئات ويجب ان يكون عدد الفئات لايقل عن 5 فئات ولايزيد عت 15 يتم استخراجة حسب طريقة سترجس

$$n= 1+3.3\log X_i$$ **(تقريب العدد الكسري الى اقرب العدد الصحيح)**

حيث ان

X: المشاهدات او القيم

n: عدد الفئات

3- ايجاد طول الفئة

طول الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$ **(تقريب العدد الكسري الى اقرب عدد صحيح)**

$$L= \frac{R}{n}$$

4- ايجاد الحدالاعلى للفئة الاولى

$$X_{min}+ L -1$$

5- استخراج التكرارات لكل فئة

مثال// انشئ جدول توزيع تكراري من البيانات التالية:

65	72	73	73	77
77	82	82	85	87
88	91	93	94	75

Sol:-

1) $R = X_{max} - X_{min}$

$= 94 - 65$

$= 29$

2) $n = 1+3.3\log X_i$

$= 1+3.3 \log(15)$

$= 1+3.3(1.176)$

$= 4.880 \sim 5$

3) $L = \frac{R}{n}$

$= \frac{29}{5}$

$= 5.8 \sim 6$

4) $X_{min} + L - 1$

$65 + 6 - 1 = 70$

5)

Class	f_i
65 - 70	1
71 - 76	4
77 - 82	4
83 – 88	3
89- 94	3

مثال واجب/ كون جدول توزيع تكراري لعلامات (30) طالب في امتحان مادة الاحصاء والبيانات كانت كماياتي:

46	49	48	58	54	50
40	62	37	48	54	75
54	48	59	45	34	58
47	61	49	44	68	39
63	56	43	57	40	45

جدول التوزيع التكراري النسبي Relative Frequency

ان التكرار النسبي لكل فئة هو نسبة تكرار تلك الفئة الى مجموع التكرارات، فاذا كان مجموع التكرارات $\sum f_i$ وتكرار الفئة f فأن التكرار النسبي هو:

$$RF = \frac{f}{\sum f_i}$$

اما التكرارالمئوي هو ضرب التكرار النسبي * 100

اي ان

$$RF\% = RF * 100$$

مثال// من الجدول التوزيع التكراري التالي احسب التكرار النسبي والمئوي

Class	f_i
22 – 26	10
27 – 31	8
32 – 36	12
37 – 41	9
42 – 46	13
47 – 51	8

Sol:-

$$RF = \frac{f}{\sum f_i}$$

$$RF\% = RF * 100$$

Class	f_i	RF	RF%
22 - 26	10	0.166	16.6
27 - 31	8	0.133	13.3
32 - 36	12	0.2	20
37 – 41	9	0.15	15
42 - 46	13	0.216	21.6
47 - 51	8	0.133	13.3

مثال واجب/ من الجدول التوزيع التكراري التالي احسب التكرار النسبي والمئوي

Class	0 – 4	5 - 9	10 - 14	15 - 19	20 - 24	25 - 29
f_i	4	5	15	25	6	5

التوزيعات التكرارية المتجمعة

نحتاج في كثير من الاحيان معرفة عدد المشاهدات التي تساوي قيمة معينة او تكون اصغر منها فعلى سبيل المثال،اذا حصل طالب على درجة 80 في احد الامتحانات فانة يرغب في معرفة عدد الطلبة الحاصلين علة درجة 80 او اقل من ذلك ويوجد نوعان من هذه الجداول هما:

1- جدول التوزيع التكراري المتجمع الصاعد UCD

هو الجدول الذي يعطينا عدد المفردات التي تقل قيمتها عن الحد الادنى لفئة معينة.

ملاحظة/عندما نجمع جميع قيم جدول التوزيع التكرار الصاعد يجب ان يساوي مجموع التكرار الاصلي لجدول التكراري.

2- جدول التوزيع التكراري المتجمع النازل LCD

هو الجدول الذي يعطينا عدد المفردات التي تزيد قيمتها عن الحد الادنى لفئة معينة.

مثال/ اليك الجدول التكراري التالي بناءاً علية كون

1) جدول توزيع تكراري صاعد
2) جدول توزيع تكراري نازل

Class	4 - 6	7 - 9	10 - 12	13 - 15	16 - 18
f_i	7	5	10	8	10

Sol:-

Class	f_i	Ucd	LCD
4 - 6	7	7	40
7 - 9	5	12	33
10 - 12	10	22	28
13 – 15	8	30	18
16 - 18	10	40	10

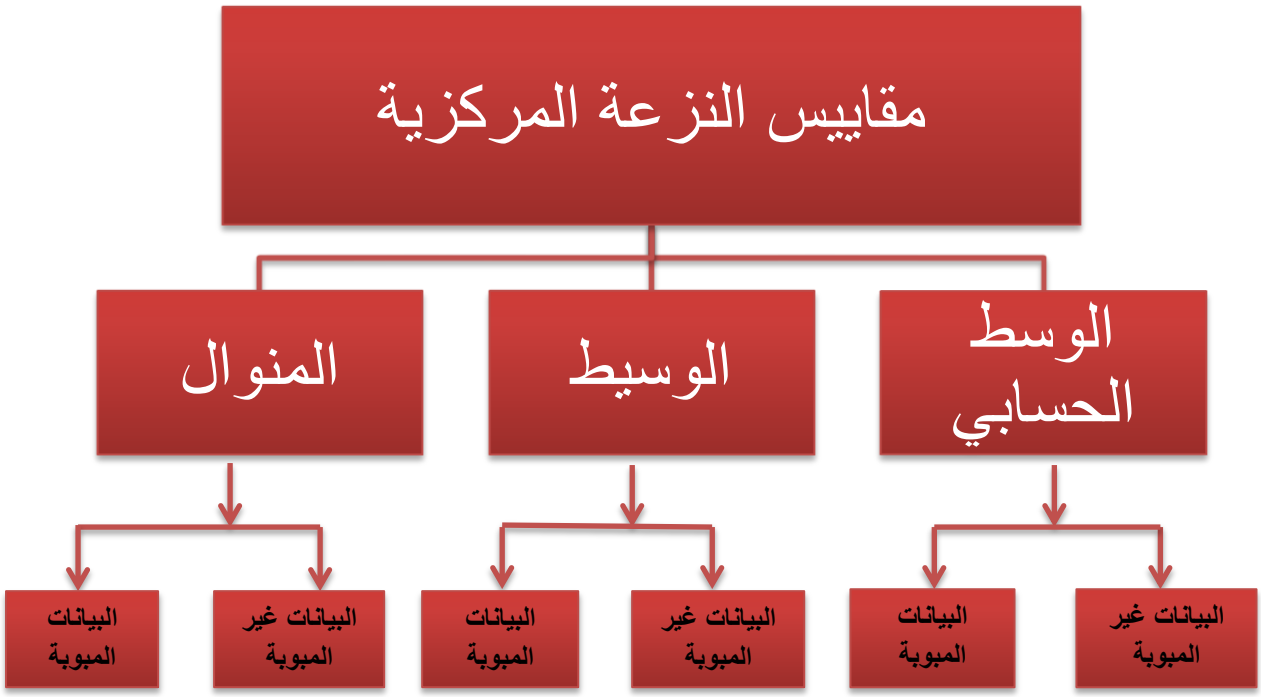
مثال واجب/ الجدول التالي يمثل علامات(30) طالب كون جدول تكرار متجمع صاعد ونازل

Class	34-39	40 - 45	46 - 51	52 - 57	58 - 63	64-69	70-75
f_i	3	6	8	5	6	1	1

مدرس المادة :م.م. دعاء بشير عباس

مقاييس النزعة المركزية

المقصود بالنزعة المركزية هو تجمع معظم المشاهدات حول قيمة متوسطة تمثل نقطة ارتكاز تلك المشاهدات او بؤرتها ومركز ثقلها، كالمتوسط الحسابي والوسيط والمنوال ، وكلما كانت تلك القيمة تمثل مركز ثقل فعلي تجذب اليها اكبر عدد ممكن من المشاهدات تكون اكثر قدرة على قياس النزعة المركزية للمشاهدات وتفقد تلك القيمة اهميتها اذا كانت اعداد كبيرة من المشاهدات تبتعد عنها، ومن اهم مقاييس النزعة المركزية هي:



اولاً: الوسط الحسابي Mean

يعد الوسط الحسابي اكثر مقاييس النزعة المركزية شيوعاً او يستخدمها الجغرافيون لوصف الكثير من الظواهرالجغرافية كمتوسط الزيادة السنوية للسكان ومتوسط حجم الاسرة ومتوسط دخل الفرد، واهم الخصائص التي يتميز بها الوسط الحسابي هي البساطة والوضوح، كما ان معظم اساليب التحليل الاحصائي تعتمد علية ولكي يتم حساب الوسط الحسابي في كلا الحالتين:

مدرس المادة :م.م. دعاء بشير عباس

1- حساب الوسط الحسابي في حالة البيانات لغير مبوبة

تعريف/ اذا كان للمتغير x قيماً عددها n ولتكن $x_1, x_2, \dots\dots, x_n$ فان:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

مثال1/ البيانات التالية تمثل اعداد افراد 10 اسر، احسب الوسط الحسابي لعدد افراد الاسر.

5,8,5,3,7,6,11,7,3,5

Sol:-

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{5+8+5+3+7+6+11+7+3+5}{10}$$

$$= \frac{60}{10} = 6$$

مثال 2/ من البيانات التالية استخرج المتوسط الحسابي

4,5,2,9

Sol:-

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{20}{4} = 5$$

مثال واجب 1/ اذا علمت قيمة $\overline{X} = 7$, n= 4 فماهي قيمة مجموع القيم.

مثال واجب 2/ اوجد قيمة $\sum_{i=1}^{4} 3i + 2$

2- حساب الوسط الحسابي في حالة البيانات المبوبة

تعريف/ اذا كان لدينا توزيع تكراري عدد فئاته n وكانت مراكز الفئات (او القيم) $X_1, X_2, \ldots\ldots, X_n$ والتكرارات المقابلة لها $f_1, f_2, \ldots\ldots, f_n$ فان الوسط الحسابي يكون:

$$\bar{X}=\frac{X_1f_1+X_2f_2+\cdots+X_nf_n}{f_1+f_2+\cdots+f_n}$$

وباستعمال رمز المجموع يكون:

$$\bar{X}=\frac{\sum_{i=1}^{n} X_i f_i}{\sum_{i=1}^{n} f_i}$$

حيث ان/ f_i : تكرار الفئة i

مثال// احسب الوسط الحسابي للجدول التوزيع التكراري التالي

Class	12 – 16	17 - 21	22 - 26	27 - 41
f_i	10	13	7	5

Sol:-

Class	f_i	X_i	$X_i f_i$
12 - 16	10	14	140
17 - 21	13	19	247
22 – 26	7	24	168
27 – 41	5	29	145
المجموع	35		700

$$\bar{X}=\frac{\sum_{i=1}^{n} X_i f_i}{\sum_{i=1}^{n} f_i}$$

$$\bar{X}=\frac{700}{35}$$

$$=20$$

مثال 2/ جد الوسط الحسابي للبيانات التالية:

Class	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49
f_i	8	11	21	28	17	15

Sol:-

Class	f_i	X_i	$X_i f_i$
20 - 24	8	22	176
25 - 29	11	27	297
30 – 34	21	32	672
35 - 39	28	37	1036
40 - 44	17	42	714
45 – 49	15	47	705
المجموع	100		3600

$$\overline{X}=\frac{\sum_{i=1}^{n} X_i f_i}{\sum_{i=1}^{n} f_i}$$

$$\overline{X}=\frac{\sum_{i=1}^{n} X_i f_i}{\sum_{i=1}^{n} f_i} = \frac{3600}{100}=36$$

مثال واجب/ من الجدول التكراري التالي احسب الوسط الحسابي

Class	10 -13	14 – 17	18 - 21	22 - 25	المجموع
f_i	5	11		6	30

2- الوسيط :Median

هو القيمة التي تقسم القيم الى نصفين متساويين، أي هو القيمة التي تقع في الوسط .ويستخرج بطريقتين :

أ- في حالة البيانات غير المبوبة

((الشرط ألساسي في هذه الطريقة ان ترتب القيم ترتيبا تصاعدياً اوتنازليا))

- في حالة اذا كانت القيم او الدرجات ذات عدد فردي فان الوسيط هو الدرجة التي تقع في الوسط بعد ترتيب البيانات، ويمكن ان نحدد موقعه في المعادلة التالية:

$$Me=\frac{n+1}{2}$$

حيث n تعني عدد الدرجات

مثال: استخرج الوسيط للبيانات (19 13، 23، 30، 15)

الحل/

اولاً: ينبغي ترتيب الارقام ترتيب تصاعدي او تنازلي: (30 23، 19، 15، 13)

ثانياً: نحدد موقع الوسيط بتطبيق القانون

$$Me=\frac{n+1}{2}$$

$$=\frac{5+1}{2}=\frac{6}{2}=3$$

اذن مرتبة الوسيط هي الثالث وهذا يعني انها الرقم (19) (30 23، 19، 15)

- اذا كانت القيم ذات عدد زوجي، فسيكون هنالك وسيطين، فالوسيط هو حاصل قسمة مجموع الدرجتين الوسطين بعد استخراج مرتبتيهما .

مثال: استخرج الوسيط للبيانات (7 6، 9، 5، 2، 3)

الحل/

اولاً: نرتب البيانات تصاعديا: (9 7، 6، 5، 3، 2)

ثانياً: نحدد موقع الوسيطين بتطبيق القانون التالي:

$$Me=\frac{\frac{n}{2}+\frac{n}{2}+1}{2}$$

$$=\frac{\frac{6}{2}+\frac{6}{2}+1}{2} \quad =\frac{3+4}{2} \quad =\frac{5+6}{2} \quad =5.5$$

ب- في حالة البيانات المبوبة (ذات الفئات)

نطبق القانون التالي:

$$Me=L+\frac{\frac{\Sigma f_i}{2}-f_1}{f_2}*W$$

حيث ان

L= الحد الادنى للفئة الوسيطية

$\frac{\Sigma f_i}{2}$ = ترتيب الوسيط

f_1= التكرار المتجمع الصاعد للفئة السابقة للفئة الوسيطية

f_2= التكرار الاصلي للفئة الوسيطية

W = طول الفئة

مثال/ جد الوسيط للتوزيع التكراري للمصروف الشهري لستين طالباً المعطى في الجدول التالي:

Classes	f_i
32 - 35	10
36 - 39	6
40 - 43	8
44- 47	4
48 - 51	8
52- 55	9
56 – 59	15
المجموع	60

الحل/

1- حساب التكرار المتجمع الصاعد

Classes	f_i	UCD
32 - 35	10	10
36 - 39	6	16
40 - 43	8	24
44- 47	4	28
48 - 51	8	36
52- 55	9	45
56 – 59	15	60
المجموع	60	

الفئة الوسيطية ← 48 - 51

→ $\frac{\sum f_i}{2}$

2- ايجاد ترتيب الوسيط

$$\frac{\sum f_i}{2} = \frac{60}{2} = 30$$

L= 48 ; f_1 = 28 ; f_2 = 8; W= 4

3- حساب الوسيط

$$Me = L + \frac{\frac{\sum f_i}{2} - f_1}{f_2} * W$$

$$Me = 48 + \frac{30-28}{8} * 4$$

$$= 48 + \frac{2}{8} * 4$$

$$= 48 + 1 = 49$$

مثال// جد الوسيط للجدول التكراري

Classes	10-14	15 - 19	20 - 24	25 - 29
f_i	4	8	9	3

الحل/

1- ايجاد التكرار المتجمع الصاعد

Classes	f_i	UCD
10 - 14	4	4
15 - 19	8	12
20 - 24	9	21
25 - 29	3	24
المجموع	24	

الفئة الوسيطية ← 15 - 19

$\frac{\sum f_i}{2}$ →

2- ايجاد ترتيب الوسيط

$$\frac{\sum f_i}{2} = \frac{24}{2} = 12$$

L= 15 ; f_1 = 4 ; f_2 = 8; W= 5

4- حساب الوسيط

$$Me = L + \frac{\frac{\sum f_i}{2} - f_1}{f_2} * W$$

$$= 15 + \frac{12-4}{8} * 5$$

$$= 15 + \frac{8}{8} * 5$$

$$= 15 + 5 = 20$$

3- المنوال Mode

أ / في حالة البيانات غير المبوبة

المنوال لمجموعة من القيم هو القيمة الاكثر تكراراً بين البيانات ،اي ان اذا كان لدينا n من البيانات او المشاهدات $X_1, X_2, \ldots\ldots, X_n$ فان المنوال لهذة البيانات هو القيمة الاكثر تكراراً ونرمز له بالرمز Mo.

مثال / جد المنوال للقيم التالية:

9,8,6,5,4,4,2

الحل/

المنوال هو(4) لانها القيمة التي تكررت اكثر من غيرها.

مثال 2/ جد المنوال للقيم التالية:

4,0,3,3,1,2

الحل/

المنوال هو(3) لانها القيمة التي تكررت اكثر من غيرها.

- اما اذا تكررت قيمتان او اكثر بالعدد نفسة بين مجموعة من القيم فان هذا يعني وجود اكثر من منوال للقيم.

مثال/ جد المنوال للقيم التالية:

9,2,4,8,2,6,6,5

الحل/

يوجد منوالين هما(2) و(6) لانهما تكررتا مرتين.

- اما اذا تكررت كل القيم بالعدد نفسة فهذا يعني لايوجد منوال

مثال/ جد المنوال للقيم التالية:

8,6,4,2

الحل/

لايوجد منوال لان ولا قيمة تكررت.

ب/ في حالة البيانات المبوبة

ان قيمة المنوال تكون تلك القيمة التي تقابل اكبر تكرار، اي ان اذا كان لدينا توزيع تكراري عدد فئاته n وكانت مراكز الفئات (او القيم) $X_1, X_2, \dots\dots, X_n$ والتكرارات المقابلة لها $f_1, f_2, \dots\dots, f_n$ وفي حالة وجود اكثر من مشاهدة يقابلها نفس التكرار فان التوزيع من هذا النوع يمتلك اكثر من قيمة واحدة اي التوزيع يمتلك اكثر من منوال.

ولحساب المنوال للبيانات المبوبة نتبع الصيغة التالية:

$$\mathbf{Mo= L+\frac{f_1-f_0}{2f_1-f_0-f_2} *W}$$

حيث ان:

f_1= التكرار الذي يقابل الفئة المنوالية(اكبر تكرار بين التكرارات الاصلية).

f_0= التكرار السابق للفئة المنوالية.

f_2= التكرار اللاحق للفئة المنوالية.

مثال/ جد المنوال لجدول التكراري التالي:

Classes	f_i
25 - 29	5
30 - 34	7
35 – 39	8
40 - 44	15
45 – 49	9
50 - 54	12
55 - 59	4

الفئة المنوالية ← 40 - 44

15 → اكبر تكرار

الحل/

$$Mo= L+\frac{f_1-f_0}{2f_1-f_0-f_2} *W$$

- لايجاد قيمة المنوال يجب تحديد قيمة كل من:

$$L= 40 \ ; f_1 = 15 \ ; \ f_0 = 8 \ ; f_2 = 9 \ ; W = 5$$

$$Mo= 40 +\frac{15-8}{2(15)-8-9} *5$$

$$= 40 +\frac{7}{30-8-9} *5 = 40 + \frac{7}{13} * 5$$

$$=40 + \frac{35}{13} \quad = 40 + 2.692= 42.692$$

مثال / الجدول التالي يبين التوزيع التكراري لاعداد العمال الزراعيين في المزارع التعاونية لاحد الاقاليم الزراعة ، المطلوب ايجاد المنوال لهذا التوزيع.

Classes	f_i
6 - 10	8
11 - 15	12
16 - 20	20
21 - 25	10
26 - 30	96

الفئة المنوالية ← 16 - 20 ، 20 → اكبر تكرار

الحل/

$$Mo= L+\frac{f_1-f_0}{2f_1-f_0-f_2} *W$$

- لايجاد قيمة المنوال يجب تحديد قيمة كل من:

$$L= 16 \; ; f_1 = 20 \; ; \; f_0 = 12 \; ; f_2 = 10 \; ; W = 5$$

$$Mo= 16 +\frac{20-12}{2(20)-12-10} *5$$

$$= 16 +\frac{8}{40-12-10} *5 = 16 + \frac{8}{18} * 5$$

$$=16 + \frac{40}{18} \quad = 16 + 2.222= 18.222$$

مدرس المادة :م.م. دعاء بشير عباس

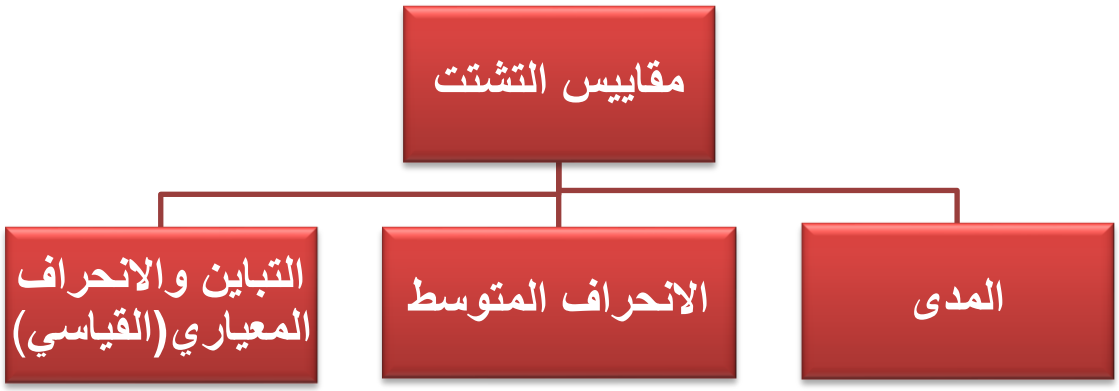
مقاييس التشتت Measures of Dispersion

التشتت يعني مدى انتشار او تباين قيم الظاهرة قيد الدراسة،والتشتت قد يكون قليلاً بمعنى ان الاعداد القريبة من بعضها تختلف عن بعضها ولكن في حدود ضيقة، وقد يكون التشتت كبيراً بمعنى ان الاعداد متفاوتة فيما بينها تفاوتاً كبيرا. ولمقارنة ظاهرتين او اكثر لانكتفي بمقاييس النزعة المركزية التي تم دراستها ، فقد يكون للظاهرتين المتوسط نفسه ولكنهما يختلفان في درجة التشتت فاذا اخذنا درجة حرارة خلال خمسة اشهر في مدينتين فكانتا على النحو التالي:

المدينة أ: 24,23,20,18,15 .

المدينة ب:36,26,20,10,8.

ان المدينتين السابقتين لهما نفس المتوسط هو (20) والوسيط نفسة هو (20)فاذا اكتفينا بمقارنة القيم الوسطى فاننا نستنتج ان درجات الحرارة متماثلة في المدينتين وهذا يخالف الواقع، لان درجات الحرارة في المدينة أ متقاربة مع بعضها وتتركز حول وسطها، والمدى هو (15-24=9)، اما المدينة ب فدرجات الحرارة الاشهر متباعدة ومتبعثرة والفرق بين اعلى قيمة وادنى قيمة هو (8-36= 28) وهو يزيد ثلاثة امثال المدى في المدينة الاولى. ان المتوسطات في الظاهرة المدروسة قد يخفي وراءة اختلافات كبيرة في شكل التوزيع ممايجعل الاعتماد على المتوسطات وحدها مضللاً في بعض الاحيان، لذا ينبغي البحث عن مقاييس تبين مدى تباعد الظواهر عن متوسطاتها وهذه المقاييس هي مقاييس التشتت ومن اهمها:



اولاً: المدى Rang

1- المدى للبيانات غير المبوبة: هو ابسط مقاييس التشتت وهو عبارة عن الفرق بين اكبر قيمة واقل قيمة في المجموعة او التوزيع اي ان:

$$Rang=X_{max}- X_{min}$$

مثال / احسب المدى للبيانات التالية:

5,7,20,13,10,15,18

الحل/

$$Rang=X_{max}- X_{min}$$

= 20 – 5 =15

2- المدى للبيانات المبوبة: يحسب حسب الصيغة التالية

المدى = الحد الاعلى الاكبر للفئات - الحد الادنى الاصغرللفئات

مثال/ جد المدى للفئات التالية:

1-10

11-20

21-30

31-40

الحل/

المدى = الحد الاعلى الاكبر للفئات - الحد الادنى الاصغرللفئات

1 -40 =

39 =

خصائص المدى

1- مقياس سهل وسريع الحساب.

2- لايمكن حسابة من توزيع تكراري مفتوح.

3- يتأثر بالقيم المتطرفة مما قد يؤدي الى اعطائة صورة غير صحيحة عن درجة تجانس توزيع الظاهرة.

ثانياً: الانحراف المتوسط Mean Deviation (MD)

أ- في حالة البيانات غير المبوبة

تعريف/ اذا كان لدينا n من المشاهدات $X_1, X_2, \ldots\ldots, X_n$ فان الانحراف المتوسط له هو الانحراف المطلق للقيم عن وسطها الحسابي اي ان:

$$\text{MD} = \frac{\sum|X_i - \bar{X}|}{n}$$

مثال/ اوجد الانحراف المتوسط للبيانات التالية:

7,6,10,8,4,5,7,9

الحل/

$$\text{MD} = \frac{\sum|X_i - \bar{X}|}{n}$$

- لحساب الانحراف المتوسط يجب ان نحسب الوسط الحسابي اولاً:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{7+6+10+8+4+5+7+9}{8}$$

$$= \frac{56}{8} = 7$$

X_i	$X_i - \bar{X}$	$\lvert X_i - \bar{X}\rvert$
7	0	0
6	-1	1
10	3	3
8	1	1
4	-3	3

5	-2	2
7	0	0
9	2	2
المجموع		12

$$MD = \frac{\sum|X_i - \bar{X}|}{n}$$

$$MD = \frac{12}{8}$$

$$= 1.5$$

مثال/ اوجد الانحراف المتوسط للبيانات التالية:

10,8,7,6,9,8

الحل/

$$MD = \frac{\sum|X_i - \bar{X}|}{n}$$

ـ لحساب الانحراف المتوسط يجب ان نحسب الوسط الحسابي اولاً:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{10+8+7+6+9+8}{6}$$

$$= \frac{48}{6} = 8$$

X_i	$X_i - \bar{X}$	$\lvert X_i - \bar{X}\rvert$
10	2	2
8	0	0
7	-1	1
6	-2	2
9	1	1
8	0	0
المجموع		6

$$MD = \frac{\sum|X_i - \bar{X}|}{n}$$

$$MD = \frac{6}{6}$$

$$= 1$$

ب- في حالة البيانات المبوبة

تعريف/ اذا كان لدينا $X_1, X_2, \ldots\ldots, X_n$ تمثل مراكز الفئات في جدول توزيع تكراري وكانت التكرارات المقابلة لها $f_1, f_2, \ldots\ldots, f_n$ على التوالي فان الانحراف المتوسط هو:

$$MD = \frac{\sum |X_i - \bar{X}| f_i}{\sum f_i}$$

مثال/ احسب الانحراف المتوسط للتوزيع التكراري في الجدول التالي:

Classes	f_i
4 - 10	1
11 - 17	2
18- 24	6
25 - 31	2
32 -38	1
المجموع	12

الحل/

$$MD = \frac{\sum |X_i - \bar{X}| f_i}{\sum f_i}$$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i f_i}{\sum_{i=1}^{n} f_i} = \frac{252}{12} = 21$$

Classes	f_i	X_i	$X_i f_i$	$X_i - \bar{X}$	$\lvert X_i - \bar{X} \rvert$	$\lvert X_i - \bar{X} \rvert f_i$
4 - 10	1	7	7	-14	14	14
11 - 17	2	14	28	-7	7	14
18- 24	6	21	126	0	0	0
25 - 31	2	28	56	7	7	14
32 -38	1	35	35	14	14	14
المجموع	12		252			56

$$MD = \frac{\sum|X_i - \bar{X}|f_i}{\sum f_i}$$

$$MD = \frac{56}{12} = 4.666$$

ثالثاً: التباين والانحراف المعياري (القياسي)

أ/ في حالة البيانات غير المبوبة

تعريف/ التباين هو مجموع مربعات انحرافات القيم عن وسطها الحسابي مقسوماً على n-1 اي ان اذا كان لدينا n من المشاهدات من $X_1, X_2, \ldots\ldots, X_n$ فان التباين لها هو:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

اما الانحراف المعياري(القياسي): هو الجذر التربيعي للتباين اي ان

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

مثال/ احسب التباين والانحراف القياسي للبيانات التالية:

4,8,6,7,5,8,4,9,7,2

الحل/

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

1- نحسب الوسط الحسابي

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\bar{X} = \frac{4+8+6+7+5+8+4+9+7+2}{10} = 6$$

2- نجد فرق البيانات عن وسطها الحسابي ثم نربع تلك الفروق

$$S^2 = \frac{(4-6)^2 + (8-6)^2 + (6-6)^2 + (7-6)^2 + (5-6)^2 + (8-6)^2 + (4-6)^2 + (9-6)^2 + (7-6)^2 + (2-6)^2}{9}$$

$$S^2 = \frac{4+4+0+1+1+4+4+9+1+1+16}{9} = \frac{45}{9} = 5$$

$$S = \sqrt{S^2} = \sqrt{5} = 2.236$$

مثال/ بلغت المعدلات السنوية للدخل الفردي في ستة اقاليم في بلد ما كما في الجدول الاتي، المطلوب ايجاد التباين والانحراف القياسي لمعدلات الدخول.

الاقاليم	معدل الدخل الفردي *$1000
A	10
B	8
C	7
D	6
E	9
F	8

الحل/

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

1- حساب الوسط الحسابي

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{48}{6} = 8$$

2- ايجاد انحرافات القيم عن وسطها الحسابي وتربيعها

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
10	2	4
8	0	0
7	-1	1
6	-2	4
9	1	1
8	0	0

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

$$= \frac{10}{5} = 2$$

$$S = \sqrt{2} = 1.414$$

ب/ في حالة البيانات المبوبة

تعريف/ اذا كان لدينا $X_1, X_2, \ldots\ldots, X_n$ تمثل مراكز الفئات في جدول توزيع تكراري وكانت التكرارات المقابلة لها $f_1, f_2, \ldots\ldots, f_n$ على التوالي فان التباين هو:

$$S^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i - 1}$$

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

مثال/ احسب التباين والانحراف القياسي للجدول التكراري التالي:

Classes	4 - 10	11-17	18 - 24	25 -31	32- 38
f_i	1	2	6	2	1

الحل/

$$S^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i - 1}$$

1- نحسب الوسط الحسابي

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i f_i}{\sum_{i=1}^{n} f_i} = \frac{252}{12} = 21$$

Class	f_i	x_i	$x_i f_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 f_i$
4- 10	1	7	7	-14	196	196
11 - 17	2	14	28	-7	49	98
18 - 24	6	21	126	0	0	0
25 - 31	2	28	56	7	49	98
32 - 38	1	35	35	14	196	196
المجموع	12		252			588

2- نحسب التباين والانحراف القياسي

$$S^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i - 1} = \frac{588}{11} = 53.454$$

$$S = \sqrt{53.454} = 7.311$$

معامل الاختلاف

يعرف بأنة نسبة حاصل قسمة الانحراف المعياري على الوسط الحسابي مضروباً * 100.

اي ان:

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري (القياسي)}}{\text{الوسط الحسابي}} * 100$$

$$C.V = \frac{S}{\bar{X}} * 100$$

ويستعمل معامل الاختلاف لاغراض المقارنة بين مجموعتين او اكثرفكلما كان معامل الاختلاف للمجموعة اصغر كانت قيم المجموعة اقل تشتتاً واكثر تجانساً وتشابهاً والعكس بالعكس.

مثال / من البيانات الجدول الاتي جد اي دولة هي اقل تشتتاً في توزيع الدخل الفردي السنوي ؟

الدولة	العملة النقدية	متوسط الدخل	الانحراف المعياري
A	يورو	1500	900
B	بأوند	2200	1100
C	دولار	1600	1040

الحل/

بتطبيق صيغة معامل الاختلاف نحصل على:

1- الدولة A

$$C.V(A) = \frac{900}{1500} * 100 = 60\%$$

2- الدولة B

$$C.V(A) = \frac{1100}{2200} * 100 = 50\%$$

3- الدولة C

$$C.V(A) = \frac{1040}{1600} * 100 = 65\%$$

نستنتج من ذلك ان التشتت في توزيع الدخل لدولة B اقل مما هو علية في الدولتين A,C والتشتت في الدولة A اقل من الدولة C، وهذا يعني ان الدولة B اكثر عدالة في توزيع الدخل الفردي السنوي لمجتمها.

الدرجة المعيارية

هي تعبير كمي عن انحراف المشاهدة عن الوسط الحسابي باستخدام الانحراف المعياري مقياساً، فهي تحدد موقع المشاهدة اتجاهاً او بعدا فالاتجاه تحددة الاشارة فاذاكانت موجبة فتكون المشاهدة اعلى من الوسط الحسابي وبالعكس اذا كانت سالبة فتكون اقل من الوسط الحسابي، اما البعد فيعني مقدار القيمة فكلما كبرت القيمة ابتعدت عن الوسط الحسابي وبالعكس.

وقانون الدرجة المعيارية هو:

$$\text{الدرجة المعيارية} = \frac{\text{القيمة – االوسط الحسابي}}{\text{الانحراف المعياري}}$$

$$Z=\frac{X-\bar{X}}{S}$$

مثال/ إذا كان متوسط درجات الطالب في مقرر الاحصاء هو (70) درجة وبانحراف معياري(10) درجات حصل أحد الطالب على (90) درجة.

المطلوب : أوجد الدرجة المعيارية التي حصل عليها الطالب.

الحل/

$$Z=\frac{X-\bar{X}}{S}$$

$$Z=\frac{90-70}{10}=2$$

مثال واجب /اذا كان عدد نزلاء فندق في الاسبوع الاول من شهر تموز هو(84) وكان معدل النزلاء في الشهر نفسه (67) والانحراف المعياري له (10) وكان عدد النزلاء في الاسبوع الاول من شهر اب (90) ومعدلهم لهذا الشهر (82) والانحراف المعياري (16) في أي شهرين كانت الايرادات اكثر حسب الدرجة المعيارية.

مدرس المادة :م.م. دعاء بشير عباس

الاحصاء الاستدلالي

الاحصاء الاستنتاجي او الاستدلالي :وهو الاحصاء الذي يهتم بتحليل البيانات واستخدام النتائج ثم تفسيرها واستعمالها لاتخاذ القرارات وعمل استنتاجات احصائية عن المجتمع الاحصائي الاصلي من العينات المحسوبة . ويضم هذا النوع نوعين رئيسيين هما:

1- التقدير: هو تقدير معالم المجتمع الاحصائي او (التوزيع الاحتمالي) والتي غالبا ماتكون مجهولة ويكون المطلوب ايجاد تقديرات لها من بيانات العينة فقد يكون المطلوب تقدير متوسط دخل الدولة او تباينها، وهناك نوعين من التقدير يسمى الاول تقدير بنقطة(تقدير قيمة واحدة اي نحصل على قيمة واحدة للعينة وتستخدم كتقدير لمعلمة المجتمع المجهولة)ويسمى الثاني تقدير بفترة او تقدير فترة الثقة(نحصل على مدى او فترة تتحدد بحد ادنى وحد اعلى).

2- اختبار الفرضيات:هو تفسير اولي يوضع لظاهرة المراد دراستها للوصول الى قرارات مناسبة بقبول الفرضية او رفضها.

انواع الاختبارات للاحصاء الاستدلالي

هناك نوعين من الاختبارات هما:

1- الاحصاء المعلمي . وهو مجموعة من الطرق التي تتطلب تحقيق افتراضات محددة حول المجتمع التي تسحب منه العينة . والافتراضات الواجب توفرها هي أ. التوزيع الطبيعي ب. تجانس التباين ج. الاستقلالية د. العشوائية في اختيار العينة .

2- الاحصاء اللامعلمي : وهو مجموعة من الطرق الاحصائية البديلة التي تستخدم في حالة عدم تحقيق الافتراضات حول المجتمع التي تسحب منه العينه او في حالة البيانات الاسمية او الرتيبة .

اخطاء في استخدام الاحصاء الاستدلالي :

عدم فحص متغيرات البحث كخطوه اولى وذلك للتعرف على نوع الاحصاء الملائم لطبيعة نتائج متغيرات البحث وهذا ما يجعلنا معرفة نوع الاحصاء الملائم ان كان معلمي او لا معلمي لان عند الفحص يجب التأكد على : تحقيق الشروط (الافتراضات) الواجب اتباعها للإحصاء المعلمي هي :

اولا: التوزيع الطبيعي : يمكن اعتبار التوزيع ان يكون طبيعيا بصورة تقريبية عندما يكون حجم العينة (30) فما فوق . فمثلا في اختبار(t) فان عدد العينة اذ كانت اقل من (30) فان شرط الاعتدالية اختل بذلك لابد من الانتقال الى الاحصاء البديل (اللامعلمي) .

ويمكن القول ان هذا الشرط لا يكون عائقا يسيرا امام استخدام الاساليب المعلمية اذ تجاوزناه بإحدى الطرق الإحصائية مثلا (معامل الاتواء ، مربع كأي ، عن طريق رسم البيانات على منحنى كأوس . وهدا يمكن استخدامه من خلال البرنامج الاحصائي .

ثانيا : تجانس التباين :ممكن ملاحظة هذا الشرط من خلال البرنامج الاحصائي (spss) مثلا في اختبار (t) للعينات المستقلة يوجد حقل خاص به هو (f) اي لمعرفة التباين بين المتغيرين. ويقصد بتجانس التباين للعينه هو الفرق بين تباين العينتين ويقاس بقسمه التباين الاكبر على التباين الاصغر .

ان تكون العينة كبيرة منتمية لمجتمع ذو تباين قليل والعينة صغيرة منتمية الى مجتمع ذو تباين كبير هنا تكمن المشكلة وتكون في حالتين :

1- اذا قبلت الفرضية الصفرية : يكون الباحث في امان اي انه لا يوجد خطا .

2- اذا قبلت الفرضية البديلة : اي ان الفرق معنوي وهنا نتوقف اذ يشير الاحصائيون ان شرط التباين هو من اهم الشروط الثلاثة التي تم ذكره .

وظيفة الإحصاء الاستدلالي

تكمن وظيفة الإحصاء الاستدلالي وأهميته في أنه يعطي الباحث القدرة على تحليل البيانات، ومن ثم يتيح له إمكانية التنبؤ بالنتائج أو استخلاص النتائج منها. لذلك يعد علم الإحصاء الاستدلالي هو أنسب أداة يمكن أن تستخدم في فهم وتفسير الظواهر المختلفة سواء في العلوم الاجتماعية أو التربوية أو الإنسانية، مع إمكانية التنبؤ بالأحداث المتعلقة بهذه الظواهر. وبالمثال يتضح المقال، فإذا أراد معلم ما الاستدلال على مدى فهم الطلاب واستيعابهم للمقرر الدراسي من خلال تحليل علاماتهم الدراسية في أحد الاختبارات. فمن خلال الإحصاء الاستدلالي يمكنه القيام بأخذ عينة عشوائية من الطلاب، وجمع البيانات اللازمة حول تقديراتهم وعلاماتهم الدراسية ليتمكن من تحليل هذه البيانات والخروج بنتائج يمكن تعميمها على باقي الطلاب.

أهمية الإحصاء الاستدلالي

تتلخص اهمية الاحصاء الاستدلالي بالنقاط التالية:

1. يمكن للإحصاء الاستدلالي التعامل مع كافة أنواع المتغيرات، سواء المتغيرات الكمية أو النوعية.
2. يمكن الباحثين من فهم خصائص العينة، ومن ثم فهو يتيح لهم فهم خصائص المجتمع ببساطة، لكن على نحو صحيح.
3. يساعد الباحثين في تعميم النتائج الخاصة بالدراسة، ويسهل عملية اتخاذ القرار.

بعض المصطلحات الاحصائية

1- المجتمع: هوعبارة عن جميع القيم او المفردات التي يمكن ان ياخذها المتغير ، والمجتمع ينقسم الى مجتمع محدود اي يمكن حصر عدد مفرداتة مثل اطوال طلبة كلية التربية للعلوم الانسانية. اما المجتمع غير المحدود لايمكن حصر عدد مفرداتة مثلا عدد البكتريا في حقل ما.

2- العينة: عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع اي ان العينة هي جزء من المجتمع.

3- اسلوب المعاينة الاحصائية :هو استخدام اسليب العينات في التوصل الى خصائص المجتمع الاصلي،وذلك من خلال تطبيق الطرق والنظريات الاحصائية.

4- الفرض الاحصائي: هو تعميم مبدئي عن ابعاد المشكلة موضوع الدراسة، وتظل صلاحيتة موضع اختبار حتى يتم الوصول الى النتائج بعد جمع البيانات وتحليلها وتفسيرها وليس من الظروري ان تاتي هذه النتائج متفقة تماماً مع الفروض التي سبق وضعها.

5- فضاء العينة : وهي مجموعة من النقاط التي تمثل جميع النتائج الممكنة لتجربة ما حيث كل نتيجة يمثل عنصر في فضاء العينة، فمثلا عند رمي حجر نرد مرة واحدة فان فضاء العينة هو:

$S=\{1,2,3,4,5,6\}$

6- الحدث: هونقطة في فضاء العينة ويرمزلة بالرمزE ، فمثلا عند رمي قطعة النقود مرة واحدة

$S=\{H,T\}$ فان $E1=\{H\}$ و$E2=\{T\}$

7- الحوادث المستقلة:وهي الحوادث التي اذا وقع احدها لايؤثر في وقوع الحوادث الاخرى.

8- الحوادث غير المستقلة: وهي الحوادث التي اذا وقع احدها يؤثر في وقوع الحوادث الاخرى.

9- الحالات الممكنة : وهي جميع النتائج المختلفة التي يمكن ان تظهر في تجربة ما.

10- الحوادث المتنافية او المستبعدة : وهي الحوادث التي يستحال حدوثهما معاً ،فمثلا عند رمي قطعة نقود فمن المستحيل الحصول على الصورة والكتابة في نفس الوقت.

11- الحالات المؤاتية: وهي الحالات التي تحقق ظهور الحدث المراد دراسة وتسمى ايظاً

حالات النجاح

12- الحالات المتماثلة : وهي الحالات المتساوية والمتكافئة في امكانية حدوثهما،فمثلا عند رمي قطعة النقود فان امكانية الحصول على الصورة اوالكتابة متساوية.

13- مضروب العدد n : يرمز للمضروب بالرمز ! فان مضروب العدد n يصبح n! ويحسب كالاتي:

$$n!=n \times (n-1) \times (n-2) \times ..... \times 2 \times 1$$

مثال/ جد مضروب الاعداد الاتية:

$$0!\ ،\quad 3!\ ،\ 1!\ ،\ 5!$$

الحل/

$$\mathbf{5!= 5\times 4\times 3 \times 2 \times 1 =120}$$

$$\mathbf{3!=3 \times 2 \times 1=6}$$

$$\mathbf{1!=1}$$

$$\mathbf{0!=1}$$

14- التوافيق

يقصد بالتوافيق بانة عدد طرق الاختيار غير مرتبة التي تكوينها من عد اشياء ياخذها كلها او بعضها ويرمز لة بالرمز C_r^n. ويحسب وفق القانون الاتي:

$$\boldsymbol{C_r^n = \frac{n!}{r!\,(n-r)!}}$$

مثال/ من بين (10) مدرسين بكم طريقة يمكن تشكيل لجنة تتكون من (3) مدرسين.

الحل/

$$\boldsymbol{C_3^{10} = \frac{10!}{3!\,(10-3)!}}$$

$$\boldsymbol{= \frac{10 \times 9 \times 8 \times 7!}{3\times 2\times 1 \quad (7)!} = 120}$$

مثال واجب/ بكم طريقة يمكن تكوين لجنة مكونة من ثمانية اعضاء من بين ثمان مدرسات وعشر طالبات؟

مدرس المادة :م.م. دعاء بشيرعباس

التوقع الرياضي Expectation

تعريف / اذا كانت المتغير العشوائي x يأخذ القيم $x_1, x_2, x_3, \dots, x_n$ باحتمال $p(x_1), p(x_2), p(x_3), \dots, p(x_n)$ على التوالي فان القيمة المتوقعة لـ(x) والتي يرمز لها بالرمز E(x)

$$E(x) = x_1 p(x_1) + x_2 p(x_2) + x_3 p(x_3) + \cdots . + x_n p(x_n)$$

$$E(x) = \sum_{i=1}^{n} x_i p(x_i)$$

ملاحظة /

- القيم المتوقعة تمثل الوسط للمجتمع اي ان $E(x) = \mu$
- التباين يحسب من خلال القانون الاتي:

$$\sigma^2 = E(x^2) - (E(x))^2$$

خواص التوقع الرياضي

- $E(a) = a$
- $E(a\,x) = a\,E(x)$
- $E(ax + b) = a\,E(x) + b$

مثال/ نلقي قطعة نقدية معدنية مرتين حيث x يمثل ظهور الصورة وكانت النتائج الاتية

x	0	1	2
P(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

احسب/

$E(x)$ ، $E(x+2x^2)$ ، σ^2

الحل/

$$E(x) = x_1\,p(x_1) + x_2\,p(x_2) + x_3\,p(x_3)$$

$$= 0(\frac{1}{4}) + 1(\frac{1}{2}) + 2(\frac{1}{4})$$

$$= 0 + \frac{1}{2} + \frac{1}{2} = 1$$

$$E(x + 2x^2) = E(x) + 2E(x^2)$$

$$E(x^2) = x_1^2\,p(x_1) + x_2^2\,p(x_2) + x_3^2\,p(x_3)$$

$$= (0)^2\ \frac{1}{4} + (1)^2\ \frac{1}{2} + (2)^2\ \frac{1}{4}$$

$$= 0 + \frac{1}{2} + 1 = \frac{3}{2}$$

$$E(x + 2x^2) = E(x) + 2E(x^2) = 1 + 2\left(\frac{3}{2}\right) = \frac{8}{2} = 4$$

$$\sigma^2 = E(x^2) - (E(x))^2$$

$$\sigma^2 = \frac{3}{2} - (1)^2 = \frac{1}{2}$$

مثال واجب/ افرض لديك احتمال نجاح راديو صالح للعمل هو 0.7 وتكلفة الراديو هو 50 دينار، وتكلفة انتاج الراديو غير صالح للعمل هو 20 دينار، ماهي القيمة المتوقعة لتكلفة راديو واحد.

مدرس المادة :م.م. دعاء بشيرعباس

التوزيع الاحتمالي Probability distribution

هناك نوعين من التوزيعات الاحتمالية تبعاً للمتغير العشوائي سواء كان متقطع او مستمر.

1- التوزيعات الاحتمالية المنفصلة (المتقطعة)

سميت التوزيعات الاحتمالية المتقطعة بهذا الاسم لان قيم المتغير العشوائي تحتوي على قفزات اي غير مستمرة فمثلا x= 1,2,4,5,6 ، ومن شروط التوزيع الاحتمالي المتقطع هي:

$$\mathbf{P(x) \geq 0} \quad , \quad \sum \boldsymbol{p(x) = 1}$$

ومن اهم التوزيعات الاحتمالية المتقطعة التي سوف نتطرق لها هو توزيع ثنائي الحدين Binomial distribution وتوزيع بواسن Possion bistribution.

1- توزيع ثنائي الحدين Binomial distribution

يستخدم هذا التوزيع لايجاد الاحتمالات في الحالات التي يكون للتجربة العشوائية نتيجتان فقط النجاح أوالفشل ومن الامثلة على التوزيع ثنائي الحدين نذكر منها:

- عند رمي قطعة نقدية ، لها نتيجتان (ظهور الوجة الذي يحمل الصورة ، او الوجة الذي يحمل الكتابة).
- نتيجة الطالب في الاختبار (نجاح ، رسوب).
- نتيجة سباق الجري (الفوز ، الخسارة).

وصيغة توزيع ثنائي الحدين تكون بالشكل الاتي:

$$\boldsymbol{p(x) = C_x^n p^x q^{n-x}} \qquad \boldsymbol{x = 0, 1, 2, 3 \ldots . n}$$

ولغرض الترميز لتوزيع بواسن ياخذ الشكل الاتي:

$$\mathbf{X \sim b(n, p)}$$

حيث ان

P: يمثل احتمال النجاح

q: يمثل احتمال الفشل ويحسب من خلال **q=1-p** ، **p+q=1**

n: عدد المحاولات

- التوقع الرياضي (الوسط الحسابي) والتباين والانحراف المعياري لتوزيع ثنائي الحدين

$$E(x) = np$$
$$V(x) = npq$$
$$S(x)=\sqrt{npq}$$

مثال/ اذا كانت نسبة نجاح الطلبة الذين يحضرون المحاضرة في مادة الاحصاء هي 0.6 ، واذا كان عدد الطلبة الذين يحضرون المحاضره بانتظام 10 طلبة ، فاذا افترضنا المتغير العشوائي x يمثل عدد الطلبة الناجحين الذين يحضرون المحاضرة.

م/ 1- حدد شكل دالة الكتلة الاحتمالية لهذا المتغير

2- احسب الاحتمالات الاتية:

أ- ماهو احتمال نجاح 3 طلاب من الذين يحضرون المحاضرة؟

ب- ماهو احتمال نجاح طالب واحد على الاقل؟

ج- احسب الوسط الحسابي والانحراف المعياري لهذا التوزيع.

الحل/ 1-

P= 0.6 **q=1-p = 0.4** **n=10**

$$p(x) = C_x^{10}(0.6)^x(0.4)^{10-x} \qquad x = 0, 1, 2, 3 \ldots\ldots 10$$

-2

$$P(x=3) = C_3^{10}(0.6)^3(0.4)^{10-3}$$

$$= \frac{10!}{3!\,7!}\ (0.216)\ (0.4)^7$$

$= 120\ (0.216)\ (0.0016) = 0.041$

-3

$P(x \geq 1) = 1 - p(x < 1) = 1 - p(x=0)$

$= 1 - C_0^{10}(0.6)^0(0.4)^{10-0}$

$= 1 - \frac{10!}{0!\ 10!}\ (1)\ (0.4)^{10}$

$= 1 - [(1)\ (1)\ (0.0001)] = 0.999$

-4

$E(x) = np = 10(0.6) = 6$

$V(x) = npq = 10(0.6)(0.4) = 2.4$

$S(x) = \sqrt{npq} = \sqrt{2.4} = 1.549$

2-- توزيع بواسن Possion distribution

يعد توزيع بواسن من التوزيعات المتقطعة المهمة جداً في الكثير من التطبيقات الاحصائية ويطلق عليها في بعض الاحيان توزيع الحوادث نادرة الحصول ، مثال على ذلك عدد الوحدات المعيبة في انتاج مصنع معين وكذلك عدد النداءات الهاتفية المستلمة من قبل بدالة هاتف خلال فترة زمنية محددة ، وسمي هذا التوزيع باسم العالم الرياضي الفرنسي Simon D.Possion(1840-1781)

- يقال ان المتغير العشوائي (x) يتوزع على وفق دالة توزيع بواسن وتاخذ دالة الكتلة الاحتمالية لهذا المتغير الشكل الاتي:

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, \ldots\ldots, \lambda > 0 \quad , e = 2.718$$

ولغرض الترميز لتوزيع بواسن ياخذ الشكل الاتي:

$X \sim p_0(\lambda)$

ولغرض الترميز لتوزيع بواسن ياخذ الشكل الاتي:

(λ): معلمة التوزيع

- التوقع الرياضي (الوسط الحسابي) والتباين والانحراف المعياري لتوزيع ثنائي الحدين

$$\mathbf{E(x) = \lambda}$$
$$\mathbf{V(x) = \lambda}$$
$$\mathbf{S(x)=\sqrt{\lambda}}$$

مثال/ من خلال مراقبة احدى البدلات في احد المستشفيات تبين انها تستقبل (6) نداءات هاتفية بالمتوسط خلال فترة ساعة واحدة، ماهو احتمال ان هذه البدلالة خلال ساعة واحدة سوف تستقبل:

1- على الاكثر 3 نداءات

2- على الاقل ندائيين

3- اربعة نداءات

4- احسب E(x) ، V(x) ، S(x)

الحل/

1- $\boldsymbol{\lambda = 6}$

$$\boldsymbol{p(x) = \frac{\lambda^x e^{-\lambda}}{x!}}$$

$$\boldsymbol{p(x \leq 3) = p(x = 0) + p(x = 1) + p(x = 2) + p(x = 3)}$$

$$\boldsymbol{= \frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!} + \frac{6^2 e^{-6}}{2!} + \frac{6^3 e^{-6}}{3!}}$$

$$\boldsymbol{= e^{-6} + 6\, e^{-6} + 18 e^{-6} + 36 e^{-6} = 0.151}$$

-2

$$\boldsymbol{p(x \geq 2) = 1 - p(x < 2) = 1 - [p(x = 0) + p(x = 1)]}$$

$$= 1 - \left[\frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!}\right] = 1 - e^{-6} + 6\,e^{-6} = 0.983$$

-3

$$p(x = 4) = \frac{6^4 e^{-6}}{4!} = 0.134$$

-4

$$E(x) = \lambda = 6$$
$$V(x) = \lambda = 6$$
$$S(x) = \sqrt{6} = 2.449$$

التوزيعات الاحتمالية المستمرة

من اهم التوزيعات الاحتمالية المستمرة هو (التوزيع الطبيعي ، T ، F، مربع كاي)

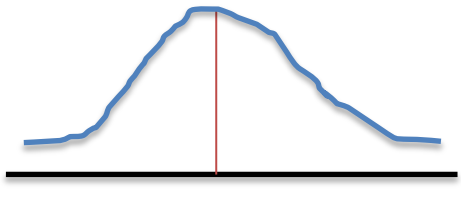
1- التوزيع الطبيعي Normal distrbution

هو التوزيع الاهم في مجال الاحصاء وان صيغة هذا التوزيع نشرت لاول مرة بواسطة العالم ابرهام في عام 1935 وقد ساهم الكثير من الاحصائيين في تاريخ التوزيع الطبيعي ومنهم Gauss والذي غالبا مايدعى التوزيع باسمة ..Gauss dis وان دالة التوزيع الطبيعي تكتب بالصيغة الاتية:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\frac{-(x-m)^2}{2\sigma^2}} \qquad -\infty < x < \infty\ ;\ -\infty < \mu < \infty\ ;\ \sigma^2 > 0$$

$$\pi = 3.14 \quad , \quad e = 2.178$$

وياخذ شكل التوزيع شكل الناقوسي او الجرس في الرسم



μ

خواص التوزيع الطبيعي:

1- شكلة يشبة الجرس وعلية فهو متماثل حول العمود المقام على الوسط الحسابي μ.

2- التوزيع الطبيعي لة قيمة واحدة وهذا يعطينا استنتاج بان لة منوال واحد يساوي الوسيط ويساوي μ. اي ان:

3- يتقارب طرفا منحنى هذا التوزيع من الصفر عندما $x \sim +\infty$ و $x \sim -\infty$

4- المساحة الكلية تحت المنحنى تساوي واحد.

2- التوزيع الطبيعي القياسي Standard Normal Distrbution

هو التوزيع الطبيعي الذي وسطة الحسابي يساوي صفر وتباينة واحد $(\boldsymbol{\mu = 0; \sigma^2 = 1})$ فلو سمينا المتغير Z على انه متغير عشوائي يخضع للتوزيع الطبيعي القياسي(المعياري) فهذا يدل على انة توزيع المتغير Z هو التوزيع الطبيعي بوسط صفر وتباين واحد اي ان دالة التوزيع الطبيعي القياسي هي:

$$\boldsymbol{f(z) = \frac{1}{\sqrt{2\pi}}\, e^{\frac{-1}{2}z^2}} \qquad -\infty < z < \infty$$

ويمكن الاعتماد على النظرية التي تنص اذا كان التوزيع الاحتمالي للمتغيرالعشوائي X هو التوزيع الطبيعي بوسط $\boldsymbol{\mu}$ وتباين $\boldsymbol{\sigma^2}$ فان توزيع المتغير العشوائي $Z=\frac{x-\mu}{\sigma}$ هو التوزيع الطبيعي القياسي.

ونلاحظ هنا ان كل قيمة من x تقابلها قيمة من قيم z حسب ماتم تعرفة اعلاة وتسمى قيم z القيم المعيارية او القياسية المقابلة لقيم x.

ملاحظة : هناك جداول خاصة لاستخراج قيم z

ملاحظة: الصيغ المعتمدة في ايجاد قيمة الاحتمال

1-P(z<a)

2- $P(z \geq a)= 1-P(z< a)$

3- P(z< -a) = 1-P(z < a)

4-P(z>-a) =1-P(z< -a)=1-[1-P(z<a)]

مثال/ اذا كان x متغير عشوائي (يمثل اوزان الطلبة) له توزيع طبيعي بمعدل $\boldsymbol{\mu = 62}$ وتباين $\boldsymbol{\sigma^2 = 36}$

م/ 1- حول المتغير x الى متغير طبيعي قياسي 2- اوجد القيم المعيارية المقابلة للقيم

$$\boldsymbol{x_1 = 64 \quad ; \quad x_2 = 54 \quad ; \quad x_3 = 70}$$

الحل/ 1-

$$\boldsymbol{Z = \frac{x - \mu}{\sigma} = \frac{x - 62}{6}}$$

-2

$$\boldsymbol{Z_1 = \frac{x - \mu}{\sigma} = \frac{64 - 62}{6} = 0.333}$$

$$\boldsymbol{Z_2 = \frac{x - \mu}{\sigma} = \frac{54 - 62}{6} = -1.333}$$

$$\boldsymbol{Z_3 = \frac{x - \mu}{\sigma} = \frac{70 - 62}{6} = 1.333}$$

مثال/ اذا علمت ان المتغير Z له التوزيع الاتي: Z ~ N(0,1) جد

P(z<1.34) ، P(z> 1.22)، P(1.3<z<2.4)

الحل/1-

نلاحظ الاحتمال المطلوب يتعلق بالخاصية الاولى من خواص التوزيع الطبيعي

P(x< 1.34)= 0.90900

2- نجد من الجدول لخاص $p(z\leq 1.22)= 0.8888$

$$P(z>1.22)=1-P(z<1.22)=1- 0.8888=0.1112$$

3- نجد من الجداول الخاصة $P(z<1.3)=0.9032$ ، $P(z<2.4)=0.9918$

$$P(1.3<z<2.4)=P(z<2.4)-P(z<1.3)$$

$$= 0.9918 - 0.9032 = 0.0886$$

الاختبار (التائي)

T- Test

هو احد اهم الاختبارات الاحصائية واكثرها استخداما في الابحاث والدراسات التي تهدف عن كشف دلالة الفروق الاحصائية بين متوسطي عينتين .

امثلة:

1- الفرق بين متوسطي الذكور والاناث في الاختبار التحصيلي لمادة العلوم.

2- الفرق بين طريقتين من طرق التدريس (بأستخدام الحاسوب او الطريقة التقليدية)

شروط عامة في الاختبارات المعلمية (مثل اختبار ت وتحليل التباين)

هناك مجموعة من الافتراضات او الشروط العامة لاستخدام الاختبارات المعلمية مثل الاختبارات او تحليل التباين:

1- مستوى قياس المتغير التابع كمي (نسبي أو فئوي).

2- المعاينة العشوائية: استخدام الاسلوب العشوائي في اختيار العينات.

3- استقلالية القياس أو المشاهدات.

4- التوزيع الاعتدالي للمتغير التابع.

5- تجانس التباين: تماثل تشتت درجات المجموعات.

انواع الاختبارات التائية

تقوم فكرة اختبار التائي على حساب نسبة انحراف فرق أي متوسطين من متوسطات التوزيع الاحصائي الى الخطأ المعياري المصاحب.

- اختبار التائي لعينة واحدة
- اختبار التائي لعينتين مستقلتين
- اختبار التائي لعينتين مترابطتين

اختبار التائي لعينة واحدة

يستخدم هذا الاختبار في مقارنة متوسط عينة بقيمة مفترضة للمجتمع

- فروض الاختبار (الفرض الصفري والفرض البديل)

$H_0 = \mu = a$

$H_a = \mu \neq a$

- مستوى الدلالة $(\alpha = 0.05)$

- قانون الاختبار التائي لعينة واحدة:

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

مثـــال / اراد باحـــث قيـــاس مســـتوى الطمـــوح لـــدى عينـــة مكونـــة مـن(x=14,5,6,8,10,11,6,12) مـن خـلال مقارنـة درجـاتهم علـى مقيـاس مسـتوى الطموح مع متوسط المجتمع والبالغ (10)، علما ان القيمة التائية الجدولية هي(2.306) عند مستوى دلاله 0.05.

الحل/

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

1- حساب الوسط الحسابي

$$\bar{X} = \frac{\sum X_i}{n}$$

$$= \frac{72}{8} = 9$$

x	X^2
14	196
5	25
6	36
8	64
10	100
11	121
6	36
12	144
$\sum = 72$	722

2- حساب الانحراف المعياري

$$S = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}}$$

$$= \sqrt{\frac{722 - \frac{(72)^2}{8}}{7}}$$

$$= \sqrt{\frac{722-648}{7}}$$

$$= \sqrt{10.571} = 3.251$$

$$t = \frac{9-10}{\frac{3.251}{\sqrt{8}}}$$

$$= \frac{-1}{\frac{3.251}{2.828}}$$

$$= \frac{-1}{1.149}$$

= -0.870

نقارن t الحسابية بالجدولية(2.306)

بمان القيمة التائية الحسابية اصغر من الجدولية ،هذا يعني لاتوجد فروق معنوية

اي ان

نقبل الفرضية الصفرية $H_0 = \mu_1 = \mu_2$

ونرفض الفرضية البديلة $H_1 = \mu_1 \neq \mu_2$

مثال2 / عينة من الطلبة مقدارها(10) حصلت على متوسط حسابي مقدارة (3.222) في اختبار الصحة النفسية وبانحراف معياري مقدارة (0.999) وبوسط فرضي (140) اختبر الدلالة الاحصائية علماً ان القيمة التائية الجدولية مقدارها (1.96).

الحل/

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{140 - 3.222}{\frac{0.999}{\sqrt{10}}} = \frac{136.778}{\frac{0.999}{3.162}} = \frac{136.778}{0.315} = 434.215$$

t =1.96 الجدولية

t الحسابية اكبر من t الجدولية وهذا يدل على وجود فروق احصائية

اي ان نرفض الفرضية الصفرية التي تنص على عدم وجود فروق ذات دلالة بين الطلبة في اختبار الصحة النفسية ونقبل الفرضية البديلة التي تنص على وجود الفروق بين الطلبة في اختبار الصحة النفسية.

- تقدير فترة الثقة لعينة واحد

ان تقدير فترة الثقة100%(1- α) للوسط الحسابي للمجتمع μ عندما يكون تباين المجتمع غير معلوم σ^2 وان (n≤ 30) هي:

$$\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}$$

حيث ان

$\bar{X}$: الوسط الحسابي.

S: الانحراف المعياري (القياسي).

$t_{\frac{\alpha}{2},n-1}$: قيمة توزيع t بمساحة $\frac{\alpha}{2}$ الى اليمين وبدرجة حرية n-1.

مثال/ اذا كانت درجات 7 طلاب في امتحان مادة الاحصاء هي:

$$X_i = 14,12,18,9,17,16,15$$

اوجد فترة الثقة عند 90% للوسط الحسابي للمجتمع التي تعود اليه درجات الطلاب.

الحل/

$$\bar{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}$$

X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
14	-0.428	0.183
12	-2.428	5.895
18	3.572	12.759
9	-5.428	29.463
17	2.572	6.615
16	1.572	2.471
15	0.572	0.327

$$\bar{X}=\frac{\sum X_i}{n}=\frac{101}{7}=14.428$$

$$S=\sqrt{\frac{\sum(X_i-\bar{X})^2}{n-1}}=\sqrt{\frac{57.713}{6}}=3.101$$

n-1= 7-1=6

1- α = 0.90 $\Longrightarrow$ α = 1-0.90 $\Longrightarrow$ α = 0.10

$$t_{\frac{\alpha}{2},n-1}=t_{\frac{0.10}{2},6}=t_{0.05,6}=1.943$$

$$\sqrt{n}=\sqrt{7}=2.645$$

$$\bar{X}-t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\leq \mu \leq \bar{X}+t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}$$

$$14.428-(1.943)\frac{3.101}{2.645}\leq \mu \leq 14.428+(1.943)\frac{3.101}{2.645}$$

$$14.428-(1.943)(1.172)\leq \mu \leq 14.428+(1.943)(1.172)$$

$$14.428-2.277\leq \mu \leq 14.428+2.277$$

$$12.151\leq \mu \leq 16.705$$

اذن فان الوسط الحسابي لفترة الثقة 90% التي تعود لها درجات الطلاب تقع بين (12.151)و (16.705)

2- الاختبار التائي لعينتين مستقلتين

يستخدم هذا النوع من الاختبارات للحكم على معنوية الفروق بين متوسطين لعينتين مستقلتين ، على فرض ان تباين المجتمع σ_2^2, σ_1^2 غير معلوميين ولكنهما متجانسين وان حجم العينة $n_1, n_2 \leq 30$ ولحساب هذا الاختبار نتبع القانون الاتي:

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

مثال/ اراد باحث ان يعرف اثر استخدام نظم مساندة على كفاءة القرارات الذي تتخذها الادارة بمساعدة تلك النظم فوزع (50) مديراً على منشات صناعية عشوائية على مجموعتين ثم عين احد المجموعتين بطريقة عشوائية لتكون مجموعة تجريبية والاخرى ضابطة وفي نهاية التجربة وزع على المجموعتين استقصاء يقيس درجة فعالية القرار وكفاءة عندما يتم اتخاذ القرار باستخدام النظم المساندة بدلاً من الطريقة التقليدية وكانت النتائج كالاتي:

المجموعة التجريبية	المجموعة الضابطة
$n_1 = 25$	$n_2 = 25$
$\bar{x}_1 = 7.60$	$\bar{x}_2 = 6$
$s_2 = 2.27$	$s_1 = 1.78$

فهل تدل هذة البيانات على ان اداء المجموعة التجريبية كان افضل من المجموعة الضابطة عند مستوى معنوية 0.05 وان القيمة التائية الجدولية هي (1.68).

الحل/

$$H_0: \mu_1 = \mu_2 = 0$$

$$H_1: \mu_1 \neq \mu_2 = 0$$

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

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

$$= \frac{(25-1)(2.27)^2 + (25-1)(1.78)^2}{25+25-2} = \frac{(24)(5.1529) + (24)(3.1684)}{48}$$

$$= \frac{123.6696 + 76.0416}{48} = \frac{199.7112}{48} = 4.160$$

$$s_p = \sqrt{s_p^2} = \sqrt{4.160} = 2.039$$

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

$$= \frac{(7.60-6)+0}{2.039\sqrt{\frac{1}{25}+\frac{1}{25}}} = \frac{1.6}{2.039\sqrt{(0.04+0.04)}} = \frac{1.6}{2.039\sqrt{0.08}} = 2.774$$

الجدولية $t_{(48,0.05)} = 1.68$

اذن t الحسابية اكبر من t الجدولية فعلية نقبل الفرضية البديلة ونرفض الفرضية الصفرية ، اي ان هذا يدل على وجود فروق معنوية بين المجموعتين.

تقدير فترة الثقة للفرق بين متوسطين حسابين(العينتين مستقلتين) للمجتمع

- اذا كان σ_1^2, σ_2^2 غير معلومتان ولكنهما متساويتان وان العينتين العشوائيتين مستقلتان وان حجم العينة $n_1, n_2 \leq 30$، ان فترة الثقة 100(1-α) الى $(\mu_1 - \mu_2)$ عندما تكون σ_1^2, σ_2^2 غير معلومتان ولكنهما متساويتان. نتبع القانون التالي لحساب فترة الثقة لعينتين مستقلتين:

$$\boldsymbol{(\bar{y}_1 - \bar{y}_2) - t_{(\frac{\alpha}{2}, n_1+n_2-2)} s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < (\mu_1 - \mu_2) < (\bar{y}_1 - \bar{y}_2) + t_{(\frac{\alpha}{2}, n_1+n_2-2)} s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

اذ ان

$(\bar{y}_1 - \bar{y}_2)$: المتوسطين الحسابيين للعينتين المستقلتين

s_p: انحراف العينتين المستقلتين ويحسب وفق القانون الاتي:

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

$$s_p = \sqrt{s_p^2}$$

مثال/ اعطيت دروس في الاحصاء بطريقتي(الحضوري والالكتروني) لشعبتين من نفس المستوى وبعد الانتهاء من الدروس اجري نفس الامتحان لكلا الشعبتين وكانت النتائج هي:

الشعبة الاولى بالطريقة التقليدية(الحضوري)	الشعبة الثانية بالطريقة الالكترونية
$n_1 = 12$	$n_2 = 10$
$\bar{y}_1 = 85$	$\bar{y}_1 = 81$
$s_1 = 4$	$s_1 = 5$

م/ اوجد فترة الثقة 90% للفرق بين متوسطي المجتمع على فرض ان المجتمع يقترب من التوزيع الطبيعي وان تباينهما غير معلوميين ولكنهما متساويين.

الحل/

1- نكتب القانون الرياضي المطلوب وهو

$$(\bar{y}_1 - \bar{y}_2) - t_{\left(\frac{\alpha}{2}, n_1+n_2-2\right)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < (\mu_1 - \mu_2) < (\bar{y}_1 - \bar{y}_2) + t_{\left(\frac{\alpha}{2}, n_1+n_2-2\right)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

2- نحسب قيمة s_p باستخدام القانون الاتي:

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

$$= \frac{(12-1)(4)^2 + (10-1)(5)^2}{12+10-2} = \frac{(11)(16)+(9)(25)}{20}$$

$$= \frac{176+225}{20} = \frac{401}{20} = 20.05$$

$$s_p = \sqrt{s_p^2} = \sqrt{20.05} = 4.478$$

3- ايجاد قيمة α

$$(1-\alpha) = 0.90 \rightarrow \alpha = 1 - 0.90 = 0.1$$

$$t_{\left(\frac{\alpha}{2}, n_1+n_2-2\right)} = t_{\left(\frac{0.1}{2}, 12+10-2\right)} = t_{(0.05, 20)} = 1.725$$

4- نطبق قانون فترة الثقة

$$(\bar{y}_1 - \bar{y}_2) - t_{(\frac{\alpha}{2}, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < (\mu_1 - \mu_2) < (\bar{y}_1 - \bar{y}_2) + t_{(\frac{\alpha}{2}, n_1+n_2-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$(85 - 81) - (1.725)(4.478)\sqrt{\frac{1}{12} + \frac{1}{10}} < (\mu_1 - \mu_2 < (85-81) + (1.725)(4.478)\sqrt{\frac{1}{12} + \frac{1}{10}}$$

$$4 - (1.725)(4.478)(0.428) < (\mu_1 - \mu_2 < 4 + (1.725)(4.478)(0.428)$$

$$4 - 3.306 < \mu_1 - \mu_2 < 4 + 3.306$$

$$0.694 < \mu_1 - \mu_2 < 7.306$$

اي ان احتمال 90% يقع الفرق الحقيقي لمتوسطي درجات طريقتي التعلم بين 0.694 و 7.306،وبما ان حدي الثقة قيم موجبة فهذا يعني ان طريقة التدريس الحضوري اكثر كفاءة من الطريقة الالكترونية.

3- الاختبار التائي لعينتين مترابطتين

يستخدم هذا الاختبار في مقارنة متوسط عينتين مترابطتين (نفس افراد العينة قبل وبعد الاختبار).

شروط اختبار عينتين مترابطتين

1- ان يكون المتغير المستقل متغيراً تصنيفياً ذات مستتويين اثنين مثل (ذكر ، انثى ، متعلم ،غير متعلم).

2- ان يتبع توزيع الفروق التوزيع الطبيعي.

3- ان يكون المتغير التابع مقاساً.

4- العينة المختارة عشوائية.

القانون الرياضي لاختبار التائي لعينتين مترابطتين هو:

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

$\bar{d}$: الوسط الحسابي للفروق، s_d : الانحراف القياسي للفروق

مثال/ مقارنة متوسط تحميل الطلاب في اختبار مادة الحاسوب الالي قبل وبعد اعطائهم دورة تدريبية وكانت النتائج الاتية:

ت	1	2	3	4	5	6	7	8	9	10
الدرجات قبل	87	75	73	94	69	96	86	85	92	88
الدرجات بعد	72	80	64	90	69	81	77	83	87	76

م/ استخدم الاختبار المناسب للمقارنة بين الفروق

الحل/ بما ان نفس العينة اجرى عليها الاختبار قبل وبعد، فان الاختبار المناسب هو الاختبار التائي لعينتين مترابطتين.

خطوات الحل:

1- نضع الفرضيات

$H_0 = \mu_1 - \mu_2 = 0$

$H_1 = \mu_1 - \mu_2 = 0$

2- نحسب الفرق بين الدرجات قبل وبعد الدورة التدريبية ثم ناخد مربع الفرق وكما يلي:

ت	1	2	3	4	5	6	7	8	9	10	
الدرجات قبل	87	75	73	94	69	96	86	85	92	88	
الدرجات بعد	72	80	64	90	69	81	77	83	87	76	
d (الفرق)	15	-5	9	4	0	15	9	2	5	12	
d^2	225	25	81	16	0	225	81	4	25	144	$\sum$ 826

3- نطبق قانون الاختبار التائي لعينتين مترابطتين

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

- اولاً: نحسب متوسط الفرق باستخدام القانون الاتي:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{66}{10} = 6.6$$

-ثانياً: نجد قيمة s_d باستخدام القانون الاتي:

$$s_d^2 = \frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1} = \frac{826 - \frac{(66)^2}{10}}{9} = \frac{826-435.6}{9} = 43.378$$

$$s_d = \sqrt{s_d^2} = \sqrt{43.378} = 6.586$$

$$\sqrt{n} = \sqrt{10} = 3.162$$

$$t = \frac{6.6 - 0}{6.586/3.162} = \frac{6.6}{2.083} = 3.169$$

الجدولية t =2.228

نلاحظ t المحسوبة اكبر من t الجدولية وهذا يدل على رفض الفرضية الصفرية التي تنص على عدم وجود فروق معنوية بين الدرجات قبل وبعد الدورة التدريبية وقبول الفرضية البديلة التي تنص على وجود فروق معنوية بين الدرجات.

ـ تقدير فترة الثقة للعينتين مترابطتين

تكون العينتان مترابطتين(غير مستقلة) عندما تكون المشاهدات مرتبطة مع بعضها البعض، فمثلا تجربة لدراسة نظام غذائي صحي معين لـ20 شخص يعانون من السمنة فعادة يتم قياس اوزانهم قبل وبعد النظام الغذائي الصحي او دراسة تاثير دواء ما على سكرالدم لمجموعة من المرضى وقبل تطبيق لدواء في العادة يتم قياس السكر لهولاء المرضى قبل وبعد العلاج.

والتجارب من هذا النوع يكون حجم العينتان متساوي اي ان:

$$n_1 = n_2 = n$$

وفي كل حالة نحسب الفرق بين المشاهدات المزدوجة ونرمز لة بالرمز di اي ان

$$d_1, d_2, \ldots\ldots d_n$$

هي الفرق بين n من المشاهدات فان تقدير نقطة لـ

$$\mu_1 - \mu_2 = \mu_D$$

لتقدير فترة الثقة % 100(1-α) الى $(\mu_1 - \mu_2) = \mu_D$ لازواج المشاهدات نتبع القانون الاتي:

$$\bar{d} - t_{\frac{\alpha}{2},n-1}\frac{S_d}{\sqrt{n}} \leq \mu_D \leq \bar{d} + t_{\frac{\alpha}{2},n-1}\frac{S_d}{\sqrt{n}}$$

مثال/ تم تقسيم 20 طالباً الى عشرة ازواج حسب درجة ذكائهم بحيث ان كل زوج من الطلاب لهم نفس درجة الذكاء تقريباً، وتم اختيار طالب من كل زوج عشوائياً لتشكيل شعبة لاعطائهم دروساً في الاحصاء من قبل استاذ المادة مباشرةً بينما بقية الطلاب (ايظاً يمثلون طلاب من كل زوج) فقد تم اعطائهم دروساً في الاحصاء بالطريقة الالكترونية.

بعد الانتهاء من الفصل الدراسي تم اعطائهم نفس الامتحان فكانت النتائج كالاتي:

ت	1	2	3	4	5	6	7	8	9	10
الالكتروني y_{1i}	76	60	85	58	91	75	82	64	79	88
الحضوري y_{2i}	81	52	85	70	86	77	90	63	85	83

م/ اوجد فترة الثقة 98% للفرق الحقيقي بين الطريقتين.

الحل/

$$\bar{d} - t_{\frac{\alpha}{2},n-1}\frac{S_d}{\sqrt{n}} \leq \mu_D \leq \bar{d} + t_{\frac{\alpha}{2},n-1}\frac{S_d}{\sqrt{n}}$$

خطوات الحل:

1- نحسب الفروق بين الطريقتين d كمايلي:

ت	1	2	3	4	5	6	7	8	9	10
الالكتروني y_{1i}	76	60	85	58	91	75	82	64	79	88
الحضوري y_{2i}	81	52	87	70	86	77	90	63	85	83
$d_i = y_{1i} - y_{2i}$	-5	8	-2	-12	5	-2	-8	1	-6	5
$\mathbf{d_i^2}$	25	64	4	144	25	4	64	1	36	25

2- نحسب الوسط الحسابي للفروق $\bar{d}$ كالاتي:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{-16}{10} = -1.6$$

3- حساب الانحراف القياسي للفروق كالاتي:

$$s_d^2 = \frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1} = \frac{392 - \frac{(-16)^2}{10}}{9} = \frac{392 - 25.6}{9} = \frac{366.4}{9} = 40.711$$

$$s_d = \sqrt{s_d^2} = \sqrt{40.711} = 6.381$$

$$\sqrt{n} = \sqrt{10} = 3.162$$

$$(1-\alpha) = 0.98 \rightarrow \alpha = 1 - 0.98 = 0.02$$

$$t_{(\frac{\alpha}{2},n-1)} = t_{(\frac{0.02}{2},9)} = t_{(0.01,9)} = 2.821$$

$$\bar{d} - t_{\frac{\alpha}{2},n-1}\frac{S_d}{\sqrt{n}} \leq \mu_D \leq \bar{d} + t_{\frac{\alpha}{2},n-1}\frac{S_d}{\sqrt{n}}$$

$$-1.6 - (2.821)\frac{6.381}{3.162} \leq \mu_D \leq -1.6 + (2.821)\frac{6.381}{3.162}$$

$$-1.6-(2.821)(2.018) \leq \mu_D \leq -1.6+(2.821)(2.018)$$

$$-1.6-5.693 \leq \mu_D \leq -1.6+5.693$$

$$-7.293 \leq \mu_D \leq 4.093$$

فنحن على ثقة بنسبة 98% ان الفرق الحقيقي بين متوسطي الطريقتين يقع بين

(-7.293) و(4.093)

اختبار مربع كاي للاستقلالية χ^2

يستخدم هذا الاختبار اثنين من المتغيرات سواء كانا وصفيان ام احداهما وصفي او كمي ويكشف عن العلاقة بينهما مستقلة ام لا وتكون بيانات المتغيرين من النوع الثنائي ويطلق على جدول البيانات احياناً بجدول التوافق ويرمز لة بالرمز χ^2.

شروط اختبار مربع كاي χ^2

- يجب الحصول على تكرارات لعينة عشوائية للمجتمع.
- يجب ان يكون كل تكرار اكبر اويساوي 5.
- مجموع تكرارات المشاهدات الفعلية = مجموع التكرارات المتوقعة

$$\sum_{i}^{n}\sum_{j}^{m} O_{ij} = \sum_{i}^{n}\sum_{j}^{m} E_{ij}$$

جدول التوافق:

اذا كانت مستويات الصفة الاولى $x_1, x_2, x_3, \dots, x_{n;\ i=1,2,\dots n}$ (عدد الصفوف)، واذا كانت مستويات الصفة الثانية $y_1, y_2, y_3, \dots, y_{m;j=1,2,\dots m}$ (عدد الاعمدة) فان جدول التوافق يكون كالاتي:

y / x	$y_{.1}$	$y_{.2}$	.....	$y_{.m}$	$\sum x_{i.}$
$x_{1.}$	O_{11}	O_{12}	.....	O_{1m}	$O_{1.}$
$x_{2.}$	O_{21}	O_{22}	.....	O_{2m}	$O_{2.}$
⋮	⋮	⋮		⋮	⋮
$x_{n.}$	O_{n1}	O_{n2}	.....	O_{nm}	$O_{n.}$
$\sum y_{.j}$	$O_{.1}$	$O_{.2}$	.....	$O_{.m}$	$O_{..}$

خطوات اختبار مربع كاي $\aleph^2$

1- تحديد الفرضية

H_0 : لاتوجد علاقة بين تصنيفات المتغيرين X و Y

H_1 : توجد علاقة بين تصنيفات المتغيرين X و Y

2- تحديد مستوى المعنوية

3- حساب احصاءة الاختبار $\aleph^2$ وفق القانون الاتي:

$$\aleph^2 = \sum_i^n \sum_j^m \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

اذ ان

O_{ij} : تكرار المشاهدة الفعلية في الصف i وعمود j

E_{ij} : تكرار المتوقع في الصف i وعمود j

4- حساب التكرارات المتوقعة E_{ij} وفق القانون الاتي:

$$E_{ij} = \frac{O_{i.} - O_{.j}}{O_{..}}$$

5- درجات الحرية لاختبار الاستقلالية

df=(n-1)(m-1)

حيث ان

n: عدد الصفوف

m: عدد الاعمدة

6- القرار

اذا كانت قيمة χ^2 المحسوبة > (اكبر من) χ^2 الجدولية

نرفض فرضية العدم ونقبل الفرضية البديلة ، اي ان توجد علاقة بين المتغيرين.

مثال/ لدراسة العلاقة بين التعليم والتدخين سُحبت عينة من 400 شخص فكانت النتائج الموضحة في الجدول ادناة.المطلوب هل توجد علاقة بين التعليم والتدخين عند مستوى معنوية 0.05 ، علماً ان قيمة χ^2 الجدولية هي (5.99).

التدخين / التعلم	مدخن	غير مدخن
غيرمتعلم	170	50
تعليم متوسط	50	70
تعليم عالي	20	40

الحل/

1-

$\mathbf{H_0}$: لاتوجد علاقة بين تصنيفات المتغيرين التدخين و التعليم

$\mathbf{H_1}$: توجد علاقة بين تصنيفات المتغيرين التدخين و التعليم

2- نجد مجموع كل صف وكل عمود

التدخين / التعلم	مدخن	غير مدخن	المجموع
غيرمتعلم	170	50	220
تعليم متوسط	50	70	120
تعليم عالي	20	40	60
المجموع	240	160	400

3- وضع جدول القيم المتوقعة لكل مشاهدة باستخدام القانون الاتي:

$$E = \frac{\text{مجموع العمود} * \text{مجموع الصف}}{\text{مجموع الكلي للمشاهدات}} \quad \text{(القيم المتوقعة)}$$

4- نحسب قيمة $\aleph^2$ المحسوبة باستخدام القانون الاتي:

$$\aleph^2 = \sum_{i}^{n}\sum_{j}^{m}\frac{(\boldsymbol{O_{ij}} - \boldsymbol{E_{ij}})^2}{\boldsymbol{E_{ij}}}$$

- نجد القيم المتوقعة

$$E_{11} or\ (170) = \frac{220 * 240}{400} = 132$$

$$E_{12} or\ (50) = \frac{220 * 160}{400} = 88$$

$$E_{21} or\ (50) = \frac{120*240}{400} = 72$$

$$E_{22} or\ (70) = \frac{120 * 160}{400} = 48$$

$$E_{31} or\ (20) = \frac{60 * 240}{400} = 36$$

$$E_{32} or\ (40) = \frac{60*160}{400} = 24$$

5- درجة الحرية

df =(n-1)(m-1)

= (3-1)(2-1)=2*1=2

نضع القيم المتوقعة في الجدول التالي لاستخراج قيم χ^2 المحسوبة

المشاهدات الفعلية O	المشاهدات المتوقعة E	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
170	132	38	1444	10.939
50	88	-38	1444	16.409
50	72	-22	484	6.722
70	48	22	484	10.083
20	36	-16	256	7.111
40	24	16	256	10.667
المجموع				61.931

اذن قيمة χ^2 المحسوبة هي:

$$\chi^2 = \sum_i^n \sum_j^m \frac{(O_{ij}-E_{ij})^2}{E_{ij}} = 61.931$$

$\chi^2 = 5.99$ الجدولية

نلاحظ قيمة χ^2 المحسوبة > من χ^2 الجدولية

اذن نرفض فرضية العدم (الصفرية) ونقبل الفرضية البديلة ، اي ان توجد علاقة بين التدخين ومستوى التعليم (انهما غير مستقلين).

مثال واجب/ سحبت عينة عشوائية من 100 شخص تم تصنيفها حسب ظاهرتين لاجراء اختبار الاستقلالية بين الدخل ومستوى التعليمي عند مستوى معنوية 0.05 ، علماً بان قيمة χ^2 الجدولية هي 9.488

الاولى / الثانية	حالة أ	حالة ب	حالة ج
منخفض	28	6	6
متوسط	22	8	10
عالي	10	6	4

مدرس المادة :م.م. دعاء بشيرعباس

تحليل الانحدارRegression Analysis

تحليل الانحدار هو اداة احصائية تقوم ببناء نموذج احصائي وذلك لتقدير العلاقة بين متغيركمي واحد اوعدة متغيرات كمية (متغيرات مستقلة) ومتغير تابع، بحيث تنتج معادلة احصائية توضح العلاقة بين المتغيرات.ويمكن استخدام هذة المعادلة في معرفة نوع العلاقة بين المتغيرات وتقدير المتغير التابع باستخدام المتغيرات الاخرى. وعندما تكون العلاقة في النموذج الاحصائي بين متغير تابع واحد مع متغيرمستقل واحد فان هذا النموذج هو ابسط نماذج الانحدار ويسمى نموذج الانحدار الخطي البسيط Simple Leaner Regression ، وعندما تكون اكثر من متغير مستقل مع متغير تابع واحد يسمى نموذج الانحدار المتعدد multiple Regression.

نموذج الانحدارالخطي البسيط Simple Leaner Regression

نموذج الانحدار البسيط هو نموذج احصائي يقوم بتقدير العلاقة التي تربط بين متغير تابع واحد مع متغير مستقل واحد، وينتج من هذا النموذج معادلة احصائية خطية يمكن استخدامها لتفسير العلاقة بين المتغيرين او لتقدير قيمة المتغير التابع عند معرفة قيمة المتغير المستقل، ويمكن صياغة العلاقة الاحصائية بالنموذج الاتي:

$$\mathbf{y} = \widehat{\boldsymbol{B_0}} + \widehat{\boldsymbol{B_1}}\boldsymbol{x} + \boldsymbol{u}$$

حيث ان

y: المتغير التابع ، x : المتغير المستقل ،u: الخطأ العشوائي ، $\boldsymbol{B_0}$ هي قيمة ثابتة تعبر عن قيمة y عندما تكون قيمة x تساوي صفراً، $\boldsymbol{B_1}$ تمثل ميل خط المستقيم الذي يوضح العلاقة بين المتغيرات، ويمكن تقدير معالم النموذج وفق طريقة المربعات الصغرى الاعتيادية .

لتقدير معالم نموذج الانحدار الخطي البسيط $\boldsymbol{B_0}, \boldsymbol{B_1}$سنستخدم الصيغ الاتية:

$$\widehat{\boldsymbol{B_0}} = \bar{\boldsymbol{y}} - \widehat{\boldsymbol{B_1}}\bar{\boldsymbol{x}}$$

$$\widehat{\boldsymbol{B_1}} = \frac{\sum \boldsymbol{xy} - \boldsymbol{n}\bar{\boldsymbol{x}}\bar{\boldsymbol{y}}}{\sum \boldsymbol{x^2} - \boldsymbol{n}\bar{\boldsymbol{x}}^{\boldsymbol{2}}}$$

مثـال/ مـن الجـدول الاتـي يمثـل العلامـات النهائيـة لــ8 طـلاب فـي مـادتي الاحصـاء والرياضيات

الرياضـــيات (x)	الاحصاء (y)
85	77
65	67
60	62
45	52
80	72
65	57
85	82
95	87

م/ جد معادلة خط الانحدار

الحل/

لايجـاد معادلـة خـط الانحـدار يجـب تقـدير معلمـات النمـوذج وهـي $\boldsymbol{B_0, B_1}$ باسـتخدام القوانين الاتية:

$$\widehat{\boldsymbol{B_0}} = \bar{y} - \widehat{\boldsymbol{B_1}}\bar{x}$$

$$\widehat{\boldsymbol{B_1}} = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

الاحصـــــاء (y)	الرياضـيات (x)	xy	$\boldsymbol{x^2}$
77	85	6545	7225
67	65	4355	4225
62	60	3720	3600
52	45	2340	2025
72	80	5760	6400
57	65	3705	4225
82	85	6970	7225
87	95	8265	9025
مج =556	مج=580	41660	43950

$$\bar{y} = \frac{\sum y_i}{n} = \frac{556}{8} = 69.5$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{580}{8} = 72.5$$

$$\widehat{\boldsymbol{B_1}} = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{41660 - 8 * 72.5 * 69.5}{43950 - 8 * (72.5)^2}$$

$$= \frac{41660-40310}{43950-42050} = \frac{1350}{1900} = 0.711$$

$$\widehat{\boldsymbol{B_0}} = \bar{y} - \widehat{\boldsymbol{B_1}}\bar{x}$$

$$= 69.5 - (0.711)(72.5) = 17.952$$

اذن معادلة خط الانحدار هي:

$$y=17.952+0.711x$$

مقاييس العلاقة

يواجة الباحث في العلوم التربوية والنفسية حالات عديدة في التعرف على العلاقة بين متغيرين او اكثر .

اما كيفية الكشف عن درجة الارتباط فهناك مؤشران رئيسيان هما معامل الارتباط وشكل الانتشار، ويعد معامل الارتباط اكثر الاكثر اهمية لانة مؤشر كمي لقوة واتجاة العلاقة ، وقوة العلاقة تقاس بمدى اقترابها من القيمة (1,-1) وعندما تقترب هذة القيمة من الصفر هذا يدل على ضعف العلاقة وينبغي التنبية الى عندما تكون قيمة معامل الارتباط صفراً فان هذا لايعني عدم وجود علاقة بل قدتكون هناك علاقة بين المتغيرين

منحنية ومن هنا لابد من رسم مخطط الانتشار لغرض ملاحظة توزيع نقاط الاستدلال على وجود او عدم وجود علاقة بين المتغيرين.

ومن جهة اخرى فان الحكم على قوة وضعف العلاقة الارتباطية ينبغي الايكون على اساس معامل الارتباط من دون ان الاخذ بعين الاعتبار العوامل عديدة مثل نوع المتغيرات ، حجم العينة وهدف البحث وغيرها ولكن بصورة عامة اذا اردنا ان نستدل على قوة وضعف العلاقة هناك عدة طرق للارتباط الخطي البسيط Simple Leanear Corrlation منها(معامل ارتباط بيرسون، ومعامل ارتباط سبيرمان ، معامل ارتباط فاي ، معامل ارتباط كاندال)وسنتطرق التوضيح هذه المعاملات لاحقاً.

اما اتجاة العلاقة فنستدل علية من خلال الاشارة فيما اذا كانت موجبة او سالبة حيث اذا كانت الاشارة سالبة تدل على ان العلاقة بين المتغيرين عكسية بمعنى ان الزيادة في احد المتغيران يصاحبة نقصان في المتغيرالاخر، وبالعكس فيما اذا كانت الاشارة موجبة تشير الى العلاقة الطردية بين المتغيرين بمعنى الزيادة في احد المتغيرين يصاحبة زيادة في المتغير الاخر وبالعكس في حالة النقصان احداهما يصاحبة نقصان في الاخر.

1- معامل الارتباط بيرسون

ذكرنا فيما سبق ان معامل الارتباط يعطي تقديراً كمياً للعلاقة بين المتغيرين اذ يكون لكل فرد قيمتان احداهما على المتغير x وتقابلها قيمة اخرى على المتغير y
اي ان معامل الارتباط لعينة عشوائية حجمها n

$$(x_1, y_1), (x_2, y_2), \dots\dots\dots, (x_n, y_n)$$

والذي يرمز لة بالرمز r هو تقدير غير متحيز لمعامل ارتباط المجتمع والقانون الرياضي لهذا المعامل ياخذ الشكل الاتي:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

او

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \sqrt{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}}$$

مثال/ احسب معامل الارتباط بيرسون بين المتغيرين x,y حيث تكون قيم كل منهما كما في الجدول الاتي:

x	y
8	6
1	4
6	3
4	4
7	5
7	4
4	2

الحل/

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\sum {x_i}^2 - \frac{(\sum x_i)^2}{n}} \sqrt{\sum {y_i}^2 - \frac{(\sum y_i)^2}{n}}}$$

x_i	y_i	$x_i y_i$	${x_i}^2$	${y_i}^2$
8	6	48	64	36
1	4	4	1	16
6	3	18	36	9
4	4	16	16	16
7	5	35	49	25
7	4	28	49	16
4	2	8	16	4
$\sum = 37$	$\sum = 28$	$\sum = 157$	$\sum = 231$	$\sum = 122$

$$r = \frac{157 - \frac{(37)(28)}{7}}{\sqrt{231 - \frac{(37)^2}{7}} \sqrt{122 - \frac{(28)^2}{7}}}$$

$$r = \frac{157 - 148}{\sqrt{231 - 195.571}\sqrt{122 - 112}} = \frac{9}{\sqrt{35.429}\sqrt{10}}$$

$$= \frac{9}{(5.952)(3.162} = \frac{9}{18.820} = 0.478$$

اذن توجد علاقة طردية متوسطة بين المتغيرين x,y

2- معامل ارتباط سبيرمان (لارتباط الرتب)

هوقياس لمستوى لارتباط الاحصائي بين متغيرين انطلاقاً من ترتيب البيانات من خلال اعطاء ارقام تسلسلية ويحسب معامل ارتباط الرتب سبيرمان باستخدام القانون الرياضي الاتي:

$$\boldsymbol{r = 1 - \frac{6\sum d^2}{n(n^2 - 1)}}$$

حيث ان

d: الفروق بين رتب المتغيرين

n: حجم العينة

مثال/احسب معامل سبيرمان لارتباط الرتب بين للمعدلات 10 طلاب في الشهادة الثانوية والفصل الجامعي الاول كما في الجدول الاتي:

Y (المعدل الثانوي)	X (الفصل الجامعي الاول)
61	77.5
72.8	85.5
65.5	79.5
71	88.8
80	93.5
74.5	85
82	94.1
81	90.5
76.1	87.2
78	89.8

الحل/

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

X (الفصل الجامعي الاول)	Y (المعدل الثانوي)	رتب x	رتب y	d	d^2
77.5	61	10	10	0	0
85.5	72.8	7	7	0	0
79.5	65.5	9	9	0	0
88.8	71	5	8	-3	9
93.5	80	2	3	-1	1
85	74.5	8	6	2	4
94.1	82	1	1	0	0
90.5	81	3	2	1	1
87.2	76.1	6	5	1	1
89.8	78	4	4	0	0
مج					16

$$r = 1 - \frac{6(16)}{10(100-1)} = 1 - \frac{96}{990}$$

$$= 1 - 0.097 = 0.903$$

اذن هناك علاقة ارتباط طردية وقوية جداً بين المتغيرين x,y

مثال واجب/ جد معامل ارتباط سبيرمان لتقديرات الطلبة في مادتي الفلسفة والاحصاء

الفلسفة	امتياز	امتياز	جيد	متوسط	ضعيف	ضعيف
الاحصاء	امتياز	جيد	جيد	متوسط	متوسط	ضعيف