الجمهـــــورية الجـزائــــــــــرية الديـــــــــمقراطيـة الـــــــشعبيـة
وزارة التعليـــم العـــالـي والبحـــث العلمـــي
كلية العلوم الاجتماعية
قسم علم الاجتماع

محاضرات وتطبيقات في الإحصاء التربوي

موجهة لطلبة السنة الأولى ماستر علم اجتماع التربية

من إعداد الأستاذ المحاضرة :
شنافي فوزية

السنة الجامعية: 2022-2023

قسم علم الاجتماع

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
كلية العلوم الاجتماعية
قسم علم الاجتماع

محاضرات وتطبيقات في الإحصاء التربوي2

موجهة لطلبة السنة الأولى ماستر علم اجتماع التربية

من إعداد الأستاذ المحاضرة :
شنافي فوزية

السنة الجامعية: 2022-2023

الجمهـــــــورية الجـزائـــــــــــرية الديــــــــــمقراطيـة الـــــشعبيـة
وزارة التعليــــم العـــالـــي والبحـــث العلمـــي
كلية العلوم الاجتماعية
قسم علم الاجتماع

الفهرس

مقدمة

تهدف هذه المطبوعة إلى تعريف طلاب الماستر علم اجتماع التربية بعلم الإحصاء وبالأخص الإحصاء الوصفي والاستدلالي الذي بدوره يسهل العمل الميداني للطالب، وللباحث الاجتماعي. كما يساعده في تحديد مجتمع البحث، وكيفية تحديد حجم العينة وانتقائه الطريقة المناسبة لاستخراج العينة الممثلة لمجتمع البحث وجمع البياناتو تشفيرها وتفريغها وتنظيمها قبل جدولتها ووصفها (حساب مقاييس النزعة المركزية والتشتت). ودرجة ونوع العلاقات بين المتغيرين ومستوى قياسها ودلالتها واختباراتها البرامترية واللابرامترية.

بمعنى آخر يهدف محتوى البرنامج الى:

1. التعريف بمعنى كلمة الإحصاء وأنواعه وأهمية الإحصاء للباحث الاجتماعي.
2. العينات والمقصود بها وأنواعها المختلفة وطرق سحب العينات والطرق المختلفة لحساب حجم العينة من المجتمع المفتوح والمغلق.
3. اختبار الفروض باستخدام الاختبارين كا2 واختبار ت

اسم الوحدة: منهجية

اسم المادة: الاحصاء التربوي 2

الرصيد: 3

المعامل: 2

أهـــــــــــــــداف التعليـــــــــــــــــــم

-التحكم في التقنيات والأدوات الإحصائية المختلفة في مجال البحث.

- التحكم في استغلال المقاييس الإحصائية في مجال البحوث التخصصية.

- التحكم في تفريغ البيانات الإحصائية ومعالجتها وتفسيرها

المعـــــــــــارف المسبقـــــــــة المطلوبـــــــة

- التمكن من الإحصاء العام الاستدلالي والوصفي

- معرفة وظائف والشروط الواجب توفرها لاستعمال المقاييس الإحصائية

- معرفة ترجمة نتائج المقاييس الإحصائية المختلفة سوسيولوجيا

محتـــــــــــــــوى المــــــــــــــادة:

أولا: مقاييس النزعة المركزية و التشتت

1-مقاييس النزعة المركزية

- المنوال
- الوسط الحسابي
- الوسط الحسابي المرجح وخواصه
- الوسيط

2-مقاييس التشتت

- المدى
- الانحراف
- الانحراف المتوسط
- التباين
- الانحراف المعياري
- الخطأ المعياري للقياس

- معامل الاختلاف

ثانيا: معامل الارتباط والانحدار

1-تفسير معامل الارتباط

2-قياس الارتباط

3-معامل ارتباط بيرسون

4-معامل الارتباط النقطي

5-معامل ارتباط سبيرمان

6-معامل الاقتران (فاي)

7-مقياس التوافق

8-لامبادا

9-نسبة الارتباط (ايتا)

10-الارتباط والسببية

طريقة التقييم: متواصل+ امتحان

المحور الأول: مقاييس النزعة المركزية والتشتت

تعتبر مقاييس النزعة المركزية والتشتت من أهم المؤشرات الأولية التي يعتمد عليها الباحث في وصف الضاهرة والتي تسمح له بالحصول على فكرة سريعة وواضحة على طريقة تمركز البيانات وانتشارها. كما أنها الأكثر إستخداما في النواحي التطبيقية. ومن أهم هده المقاييس التي سنتناولها في هدا المحور يمكن ذكر المتوسط بأنواعه الحسابي. الهندسي والتوافقي. الوسيط، المنوال، المدى والتباين والانحراف المعياري.

المحاضرة رقم 1: المتوسط الحسابي

1. *المتوسط الحسابي*

يعرف المتوسط في الإحصاء بأنه القيمة التي تتجمع حولها مجموعة من القيم. يرمز له بالرمز $\bar{x}$ أما طريقة حسابه فإنها تختلف باختلاف البيانات المتوفرة:

1.1.المتوسط الحسابي حالة عدم تكرار المتغير x_i

يعتبر المتوسط الحسابي البسيط من أشهر أنواع المتوسطات الحسابية والأكثر أستخدام في حالة بيانات غير مبوبة والذي يعطى بالعلاقة التالية

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

مثال 1:

أوجد المتوسط الحسابي لنتائج اللغة العربية لعينة من تلاميذ متوسطة التالية: 11، 11، 7، 5، 3، 4، 12، 14

التطبيق

$$\bar{x} = \frac{1}{N}\sum xi = \frac{11+11+7+5+3+4+12+14}{8}$$

تمرين تطبيقي: إذا كانت علامات 9 تلاميذ في مادة الرياضيات كما يلي :

نتائج الرياضيات									
2	12	13	3	7	5	12	15	9	X_i

فما متوسط العلامات ؟

الحل:

$$\overline{x} = \frac{1}{N}\sum xi = \frac{9+15+12+5+7+3+13+12+2}{9}$$

2.1.المتوسط الحسابي حالة تكرار المتغير x_i: في هذه الحالة نسجل نوعين من التوزيع:

- **توزيع لمتغير كمي منفصل:** عند تجميع البيانات الرقمية التي لا تأخذ إلا قيمة صحيحة غير كسرية وتكرار الاجابة مثل عدد التلاميذ حسب المستوى، عدد المعلمين، عدد الاداريين في المؤسسة التعليمية. وعرضها في جداول بسيطة
- المتوسط الحسابي كما يلي:

فإذا كانت $x_1, x_2, \ldots\ldots\ldots\ldots, x_n$ تمثل مشاهدة عينة x. و كانت n_n ,

$n_1, n_2, \ldots\ldots\ldots\ldots,$ تمثل التكرارات المقابلة , فإن المتوسط الحسابي لهذه القيم

تعطى بالعلاقة التالية :

$$\bar{x} = \frac{1}{2}\sum nixi$$

مثال تطبيقي: ليكن لذينا توزيع البيانات الخاصة بعدد اخوة التلميذ:

ضرب

عدد الاخوة xi	التكرار ni	ni*xi
1	4	1x4=4
2	7	2x7=14
3	8	3x8=24
4	15	4x15=60
5	6	5x6=30
6	4	6x4=24
7	2	7x2=14
المجموع	46	170

متغير متصل: اذا كانت البيانات الرقمية المجمعة على شكل مجالات أو فئات. ويكون المتغير الكمي متصلا. فإن c_1,c_2 ,............ , c_n تمثل متوسطات الفئات و كانت n_n ,............ , n_2,n تمثل التكرارات المقابلة للمتوسط الحسابي لهده القيم تعطى بالعلاقة التالية: $\bar{X} = \frac{1}{N} \sum n_i c_i$

مثال تطبيقي:

معدلات مادة التاريخ xi	التكرار ni	مركز الفئة ci	ni*ci
0-5	5	(0+5)/2=2,5	5*2,5=12,5
5-10	12	(5+10)/2=7,5	12*7,5=90
10-15	6	(10+15)/2=12,5	6*12,5=75
15-20	15	(15+20)/2=17,5	15*17,5=262,5
المجموع	N=46		$\sum n_i c_i$=440

تطبيق عددي:

$\bar{X} = \frac{1}{46} * 440$

3.1.إيجابيات وسلبيات المتوسط الحسابي :

من ايجابياته سهل في الحساب والمقارنة مابين مجموعتين ولكن لا يمكن إجراء مقارنة شاملة بين مختلف البيانات. فقد تكون هناك مجموعتين من البيانات لهما نفس الوسط الحسابي ولأنها تختلف في خصائص أخرى. ولتوضيح دلك نفترض أن مجموعتين متكونة من 5طلبة في كل مجموعة تحصلوا على النتائج التالية في مادة اللإحصاء:

المجموعة الأولى:

18	19	2	6	5

المجموعة الثانية:

10	10	11	9	10

وبحساب الوسط الحسابي لمادة الإحصاء في المجموعة الأولى:

$$\overline{X}_A = \frac{18+19+2+6+5}{5}$$

$$\overline{X}_A = \frac{50}{5} = 10$$

وبحساب الوسط الحسابي لمادة الإحصاء في المجموعة الثانية:

$$\overline{X}_B = \frac{10+10+11+9+10}{5}$$

$$\overline{X}_B = \frac{50}{5} = 10$$

مقارنة بين المجموعتين يمكننا إستنتاج أن الغئتين لهما نفس المستوى في مادة الإحصاء. إلا أن الواقع غير ذلك، فلو تمعنا في النقاط لوجدنا 3 من 5 طلاب في المجموعة الأولى لم يتحصلوا على المعدل ومن ناحية أخرى لو أخذنا الفرق بين اعلى نقطة وأدناها في المجموعتين لوجدنا في المجموعة الأولى أن الفلرق كبير وهو 19−2=17 والمجموعة الثانية الفرق يساوي 11−9=2 بمعنى أن النقاط قريبة جدا من بعضها ومن وسطها الحسابي. وعليه يمكن إستنتاج أن المتوسط الحسابي (أو مقاييس النزعة المركزية عموما) ليست كافية لوصف البيانات أو تحليلها كميا ولابد أن تكون مصحوبة بمقاييس أخرى لقياس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها وتسمى هده المقاييس بمقاييس التشتت.

المحاضرة رقم 2: المتوسط الحسابي الهندسي، التوافقي والتربيعي

2.المتوسط الهندسي :

2-1-المتوسط الهندسي للبيانات الخامة:

يستخدم الوسط الهندسي عندما تكون في البيانات قيم متطرفة أو عندما يكون توزيع البيانات ملتو نحو اليمين أو موجبة الإلتواء. ودلك بهدف تقليل تأتير القيم المتطرفة فإدا كانت x2,x ,.... تمتل مشاهدات عينة X فإن المتوسط الهندسي لهده القيم معطى بالعلاقة التالية :

$$G = \sqrt[N]{x_1.x_2..............x_n}$$

و لتبسيط العلاقة . يفضل استخدام لوغاريتم لتصبح العلاقة :

$$G = \sqrt[n]{x_1.x_2..............x_n} = (x_1.x_2..............x_n)^{1/N}$$

$$\ln G = \ln(x_1.x_2..............x_n)^{1/N} = \frac{1}{N}(\ln x_1 + \ln x_2 + ...\ln x_n)$$

$$\ln G = \frac{1}{N}\sum_{i=1}^{n} \ln x_i \rightarrow G = e^{\frac{1}{n}\sum_{i=1}^{n}\ln x_i}$$

2-2-المتوسط الحسابي الهندسي للبيانات المبوبة:

إدا كانت x2,x ...., xn تمثل مشاهدات عينة X و كانت , nn

n2,n .... تمثل التكرارات المقابلة فإن المتوسط الحسابي لهده القيم معطى

بالعلاقة التالية:

$$\boxed{G=\sqrt[N]{x_1^{n_1}.x_2^{n_2}\ldots\ldots\ldots\ldots..x_n^{n_n}}}$$

و لتبسيط العلاقة يفضل إستخدام اللوغريتم لتصبح العلاقة :

$$\boxed{G=\left(x_1^{n_1}.x_2^{n_2}\ldots\ldots\ldots\ldots..x_n\right)^{1/N}}$$

في حالة بيانات مبوبة في جدول توزيع تكراري متصل فإنxi تعوض بمركز الفئات ci

3.2الوسط التوافقي:

هو أحد مقاييس النزعة المركزية وهو مقلوب الوسسط الحسابي لمقلوبات هده القيم

ويفضل أستخدامه على باقي المتوسطات في حالة أيجاد معدل السرعات ومعدل التغيير.

و لا يمكن إستخدامه في حالة إدا كانت إحدى هده القيم مساوية للصفر.

إدا كانت 1x ,x2 ...., xn تمثل مشاهدات عينة x فإن الوسط التوافقي H

لهده القيم. يعطى بالعلاقة التالية:

$$H = \frac{N}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

حساب الوسط التوافقي من بيانات مبوبة منفصلة أو متصلة:

إذا كانت x1, x2, ..., xn تمثل مشاهدات عينة x و كانت , nn

n,n2 ...., تمثل التكرارات المقابلة فإن المتوسط التوافقي لهذه القيم يعطى

بالعلاقة التالية :

$$H = \frac{N}{\sum_{i=1}^{n} \frac{1}{x_i n_i}}$$

3.1.المتوسط التربيعي:

المتوسط التربيعي للأي مجموعة قيم هو الجدر التربيعي للمتوسط الحسابي لمربعات تلك القيم. ويتم حساب المتوسط التربيعي للبيانات الأولية كما يلي:

$$Q = \frac{\sqrt{\sum_{i=1}^{n} x_i^{\,2}}}{\sqrt{N}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^{\,2}}{N}}$$

يحسب المتوسط التربيعي للبيانات المبوبة كما يلي :

$$Q = \frac{\sqrt{\sum_{i=1}^{n} x_i^{\,2}.n_i}}{\sqrt{N}} =$$

مثال 1:

أحسب كل من المتوسط الحسابي الهندسي التوافقي والتربيعي في المثال التالي:

البيانات التالية تمثل أوزان 10 أطفال بالكيلوغرام 20-25-15-20-19-18-19-23-17-19

- الوسط الحسابي:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{N} = \frac{18+19+20+15+25+20+23+17+18+19}{10}$$
$$= \frac{194}{10} = 19.4$$

- الوسط الهندسي:

$$G = \sqrt[N]{x_1.x_2..............x_n} = (18.19.20.15.25.20.23.17.18.19)^{1/10}$$
$$= 19.21$$

- الوسط التوافقي:

$$H = \frac{N}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{10}{0.525} = 19.03$$

- الوسط التربيعي:

$$Q = \frac{\sqrt{\sum_{i=1}^{n} x_i^{\,2}}}{\sqrt{N}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^{\,2}}{N}} == \sqrt{\frac{3838}{10}} = 19.59$$

عند حساب المتوسطات السابقة يجب التأكد من المعادلة التالية:

$$\boxed{H < G < \bar{X} < Q}$$

مثال2:

الجدول التالي يمثل توزيع 38 تلميذ على حسب عدد الأطفال:

5	4	3	2	1	Xi
2	4	15	10	7	Ni

الوسط الحسابي:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i n_i}{N} = \frac{98}{38} = 2.57$$

- الوسط الهندسي:

$$G = \sqrt[N]{x_1^{n1} . x_2^{n2} \ldots\ldots\ldots\ldots x_n^{nn}} = (1^7 . 2^{10} \ldots\ldots\ldots\ldots\ldots 5^2)^{1/10}$$
$$= (94036996915200)^{1/10} = 2.33$$

- الوسط التوافقي:

$$H = \frac{N}{\sum_{i=1}^{n} \frac{n_i}{x_i}} = \frac{38}{18.4} = 2.06$$

- الوسط التربيعي:

$$Q = \frac{\sqrt{\sum_{i=1}^{n} x_i^2 n_i}}{\sqrt{N}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 n_i}{N}} = \sqrt{\frac{296}{38}} = 2.79$$

مثال3:

الجدول التالي يمثل توزيع العمال على حسب الأجور في مؤسسة تعليمية(x 10)

ni	Xi
3	30 – 20
9	40 – 30
6	50 – 40

5	60 – 50
2	70 – 60

- الوسط الحسابي:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i c_i}{N} = \frac{1065}{25} = 42.6$$

- الوسط الهندسي:

$$G = \sqrt[N]{c_1^{n1} . c_2^{n2} .... c_n^{nn}} = (25^3 . 35^9 .... 65^2)^{1/25}$$
$$= (217444120180937 * 10^{26})^{1/25} = 41.07$$

- الوسط التوافقي:

$$H = \frac{N}{\sum_{i=1}^{n} \frac{n_i}{c_i}} = \frac{25}{0.632} = 39.54$$

- الوسط التربيعي:

$$Q = \frac{\sqrt{\sum_{i=1}^{n} c_i^{\,2} n_i}}{\sqrt{N}} = \sqrt{\frac{\sum_{i=1}^{n} c_i^{\,2} n_i}{N}} = \sqrt{\frac{48625}{25}} = 44.10$$

المحاضـرة رقم 3: المنـوال

3.المنوال

تعريف: يعرف المنوال Mo بأنه القيمة الأكثر تكرارا أو الأكثر شيوعا. نرمز له بالرمز Mo يمكن إستخراجه من البيانات مختلفة الأنواع كما يجب الإشارة أن المجتمع الإحصائي ليس بالضرورة له منوال واحد. يسمى التوزيع أحادي المنوال إدا وجد منوال واحد وتنائي المنوال إذا وجد منوالان، وفي حالة وجود عدة منوالات فقد يسمى متعدد المنوال.

متلا لو سألنا 10 تلاميذ الابتدائية عن المواد الأساسية المفظلة المفضلة لديهم وكانت الإجابة كالتالي: رياضيات–عربية – عربية – فرنسية – عربية – عربية – رياضيات – رياضيات – فرنسية – عربية.

فالمنوال هو الإجابة الغالبة وبالتالي هي مادة اللغة العربية.

بنفس المعنى يمكن استخراج المنوال في حالة القيم المنفردة والقيم المبوبة المنقطعة. وهو الذي يقابل أكبر تكرار.

متال 1:

البيانات التالية تمثل أوزان 10 تلاميذ السنة الأولى ابتدائي بالكيلوغرام 20–25–15–20–19–18–20–19–17–23 –20–18–20–19–18–20–20.

Mo=20 لانه مقارنة ببقية الأوزان .فإنه تكرر 7 مرات. أي أن الوزن الغالب في العينة الإحصائية هو 20 كلغ.

مثال2:

الجدول التالي يمثل توزيع 43 أسرة على حسب عدد الأطفال

xi	0	1	2	3	4	5	Σ
ni	5	7	10	15	4	2	43

إن عدد الأطفال الغالب في العينة الإحصائية هو 3.

أما إدا كانت البيانات مبوبة في جدول توزيع تكراري دات فئات. فالمنوال يحسب وفق الخطوات التالية

- نعين الفئة المنوالية و هي الفئة المقابلة لأكبر تكرار.
- حساب المنوال وفق القاعدة التالية:

$$\boxed{Mo = A_{Mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2}.L}$$

حيث:

Δ1: الفرق بين تكرار الفئة منوالية وتكرار الفئة التي تقابلها.

Δ2 :الفرق بين تكرار الفئة منوالية وتكرار الفئة التي بعدها

L : طول الفئة المنوالية

A : الحد الأدنى للفئة المنوالية.

مثال2 : من خلال البيانات التالية أوجد المنوال

ni	xi
1	50- 40
5	60- 50
4	70- 60
8	80- 70
12	90- 80
2	100 -90
5	110- 100

- الفئة المنوالية التي تقابل أكبر تكرار 80 -90

تطبيق عددي:

$$Mo = A_{Mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2}.L = 80 + \frac{(12-8)}{(12-8)+(12-2)}x10 = 82.85$$

المحاضرة رقم4: الوسيط

4.الوسيط

تعريف:

يعرف الوسيط Me بقيمة المشاهدة التي تتوسط البيانات بعد ترتيبها تصاعديا.

وعليه تكون عدد القيم (المفردات) قبله مساوي لعدد القيم بعده.

حالة قيم غير مبوبة:

يمكن إيجاد الوسيط للقيم المفردة من خلال:

- عدد المشاهدات فرديا

- ترتيب هده القيم تصاعديا
- رتبة الوسيط هو RMe= N+1/2
- الوسيط هي المشاهدات التي تقابل هده الفئة

مثال1:

جد الوسيط من خلال المعطيات التالية: 7-10-9-8-10-8-12-12-6-1-3

- ترتيب هده القيم تصاعديا: 1-3-6-7-8-8-9-10-10-12-12
- رتبة الوسيط هو RMe= N+1/2=11+1/2=6
- الوسيط هي المشاهدة التي توافق هده الرتبة Me=8

عدد المشاهدات زوجيا

يمكن ايجاد الوسيط للقيم المفردة من خلال:

- ترتيب هده القيم تصعديا
- رتبة الوسيط هو $RMe= (N/2,n/2+1)$
- الوسيط هي المشاهدة التي تتوسط المشاهدتين $(N/2 , N+2+1)$

مثال2:

جد الوسيط من خلال المعطيات التالية: 7-10-9-8-10-8-12-12-6-1-3-12

- ترتيب هده القيم تصاعديا: 1-3-6-7-8-8-9-10-10-12-12-12
- رتبة الوسيط هو $RMe= (N/2 , N/2+1)= (5.6)$
- الوسيط هي المشاهدة التي توافق متوسط الرتبتين $Me = 8+8/2$

حساب الوسيط من البيانات المصنفة في جدول تكراري

متغير منقطع:

لحساب الوسيط من جدول تكراري دات متغير منفصل. نتبع الخطوات التالية:

- حساب التكرار المجتمع الصاعد
- رتبة الوسيط هو $RMe= (N/2)$
- الوسيط وهي قيمة المشاهدة المقابلة للتكرار المجتمع الدي يساوي أو يزيد عن رتبة الوسيط

مثال3:

لو أخذنا نفس المثال السابق وحاولنا حساب الوسيط

- حساب التكرار المتجمع الصاعد

الجدول التالي يمثل توزيع 43 تلميذ حسب عدد الاخوة

xi	0	1	2	3	4	5	Σ
ni	5	7	10	15	4	2	43
N↗	5	12	22	37	41	43	

- رتبة الوسيط هو RMe = (42/2) = 21.5
- الوسيط وهي قيمة المشاهدة المقابلة للتكرار الذي يساوي أو يزيد عن رتبة الوسيط

Me = 3. (37)

- رتبة الوسيط هو RMe= (N/2)
- الفئة الوسيطة وهي قيمة المشاهدة المقابلة للتكرار الدي يساوي أو يزيد عن رتبة الوسيبط
- يطبق القانون التالي لحساب قيمة الوسيط:

$$Me = A_{ME} + \frac{\frac{N}{2} - n\nearrow_{iMe-1}}{n_{Me}} . L$$

حيث:

A_{me} : الحد الأدنى للفئة الوسيطة

$N\nearrow$: التكرارات الصاعدة التي تسبق الفئة الوسيطية

L : طول الفئة الوسيطية

مثال4:

لو أخذنا نفس المتال السابق ونحاول حساب الوسيط:

- حساب التكرار المتجمع الصاعد

n	ni	xi
1	1	50 -40
6	5	60 -50
10	4	70 -60
18	8	80 -70
30	12	90 -80
32	2	100 -90
37	5	110 -100

- رتبة الوسيط هي RMe = (37/2) = 18.5
- الفئة الوسيطة 80 – 90

$$Me = A_{ME} + \frac{\frac{N}{2} - n_{iMe-1}}{n_{Me}}.L = 80 + \frac{18.5 - 18}{12}.10 = 80.41$$

مقاييــس التشتــت

المحاضرة رقم 5: المـــدى الربيعي

5.المدى الربيعي:

تعريف: يعرف بأنه الفرق بين أكبر قيمة وأصغر قيمة لمجموعة من البيانات. في حالة البيانات الغير المبوبة أو بيانات تكرارية منفصلة فلا مشكلة في حساب المدى. بينما في حالة البيانات التكرارية دات فئات. فإن المدى يساوي الفرق بين مركزي الفئتين الأخيرة والأولى. ويرمز له ب E وفق العلاقة التالية:

$$E= X_{MAX} - X_{MIN}$$

يعد المدى من أبسط مقاييس التشتت وأقلها دقة من حيت إتخاده.

مثال:

لو فرضنا أطوال 6 تلاميد في مؤسسة تربوية كما يلي:

130	128	125	133	111	135

فالمدى هو : $E= XMAX - XMIN$

$$E= 24= 111-135$$

بالإضافة ألى الخصائص السابقة فهو شديد التأتر بالقيم المتطرفة.

من أهم عيوب المدى أنه يتأتر بالقيم المتطرفة وبالتالي فهو يعطي أحيانا نتائج مضلة.

لدلك دعت الحاجة إلى إيجاد مقياس أخر يتم من خلاله التخلص من تأتير هده القيم وهو ما يسمى بالمدى الربيعي

ويعرف كما يلي:

إذا كانت لدينا مجموعة من البيانات عددها n قيمة فيتم ترتيب القيم تصاعديا وتقسم إلى أربعة أقسام متساوية. تسمى القيمة التي يسبقها ربع البيانات بالربيع الأول ويرمز له بالرمز Q1و رتبته n/4. و تسمى القيمة التي يسبقها ثلاثة أرباع البيانات بالربيع التالت و يرمز له بالرمز وQ3 و رتبته 3n/4.كما يسمى المقدار الناتج من الفرق بين Q1 و ، Qبالمدى الربيعي Q و يعطى من خلال العلاقة :

$$Q = Q3 - Q1$$

بالنسبة للبيانات الغير مبوبة:

أوجد المدى لمجموعة تابيانات المعطاة في الجدول التالي:

6	8	5	4	6	1	9	2	3	7	9	2

1-ترتيب الأعداد تصاعديا

1	2	2	3	4	5	6	6	7	8	9	9

2-حساب:

R(Q1)=12/4 =3 القيمة المقابلة له هي **2:Q1**

R(Q3) = 3*12/4= 36/4= 9 **القيمة المقابلة له هي 7: Q3**

$$Q = 7-2 = 5$$

مثال:

بالنسبة للبيانات مبوبة المفصلة :

أوجد المدى لمجموعة تابيانات المعطاة في الجدول التالي:

X_i	1	2	3	4	5	6	7	8
n_i	12	14	11	13	14	12	19	17

حساب التكرار التصاعدي

X_i	n_i	$N\nearrow$
1	12	12
2	14	26
3	11	37
4	13	50
5	14	64
6	12	76
7	19	95
8	17	112

1-حساب

$R(Q1) = 112/4 = 28$ القيمة المقابلة له هي Q1 :3

$R(Q3) = 3*112 = 84$ القيمة المقابلة له هي Q3: 7

$Q=7-3=4$

في حالة ما كانت رتبة الربيعي محصورة بين قيمتين. فالقيمة العليا هي التي تأخذ بعين الإعتبار.

متال :

بالنسبة للبيانات المبوبة دات فئات

البيانات التالية تمتل توزيع أجور عمال مؤسسة تعليمية الوحدة: 10^3 دج

X_i	28 −18	38 − 28	48 − 38	58 − 48	68 − 58
N_i	18	22	17	15	4

أوجد المدى

1−حساب التكرار التصاعدي

Xi	Ni	ni
28 −18	18	18
38 − 28	22	40
48 − 38	17	57
58 − 48	15	72
68 − 58	4	76

2-حساب الفئة الربيعية الأولى والتالتة

R(Q1) = 76/4 = 19 الئة المقابلة له هيQ1 : 38 − 28

R(Q3)=3x76/4=57 القيمت المقابلت لهQ3 :48 – 38

حساب Q1 و Q3

$$Q_1 = A_{Q_1} + \left[\frac{\frac{n}{4} - N^{\uparrow}_{Q_1-1}}{n_{Q_1}}\right] \times l_{Q_1} = 28 + \left[\frac{19-18}{22}\right] \times 10 = 28.45$$

$$Q_3 = A_{Q_3} + \left[\frac{\frac{3n}{4} - N^{\uparrow}_{Q_3-1}}{n_{Q_3}}\right] \times l_{Q_{31}} = 38 + \left[\frac{57-40}{17}\right] \times 10 = 48$$

IQ = 48 – 28.45 = 20.45

المحاضرة رقم 6: التباين والانحراف المعياري

6.التباين و التشتت

1.6.التباين

يعتبر التباين من بين مقاييس التشتت المهمة. وهو عبارة عن المتوسط الحسابي امربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي. يتأثر التباين بالقيم المتباعدة أو المتطرفة ولكنه لا يتأثر كثيرا بالتغيرات التي تطرأ على العينة. يرمز له بالرمز **V(X).**

حساب التباين من البيانات الغير مبوبة:

إذا كانت القيم $X_1, X_2, \ldots\ldots\ldots\ldots, X_n$ تمثل بيانات مجتمع أو عينة . فإن التباين يمكن صياغته وفق العلاقة التالية:

- **تباين المجتمع:**

$$\boxed{\sigma^2 = \frac{\sum_{i=1}^{n} (Xi - \mu)^2}{n} =}$$

او بالصيغة المختصرة :

$$\sigma^2 = \frac{\sum x_i^2}{n} - \mu^2$$

- **تباين العينة:**

$$\boxed{S^2 = \frac{\sum_{i=1}^{n} (Xi - \overline{X})^2}{n-1}}$$

او بالصيغة المختصرة :

$$\boxed{S^2 = \frac{1}{n-1} \left(\sum X_i^2 - n\overline{X}^2\right)}$$

يطرح العدد واحد من المقام n−1(أو من مجموع التكرارات و دلك عند تقدير مقياس التباين و الإنحراف المعياري للعينات الصغيرة و أن العدد واحد يمتلدرجات الحرية) تعتبر العينة كبيرة إدا زاد حجمها عن 30 مشاهدة.

مثال:

أحسب تباين المعطيات التالية: 13 − 17 − 14 − 19 − 15 − 16 − 11 .

الحل:

1−حساب الوسط الحسابي:

$$\overline{X} = \frac{11+16+15+19+14+17+13}{7} = \frac{105}{7} = 15$$

2−حساب إنحرافات القيم عن الوسط الحسابي:

−4 1 0 4 −1 2 −2

3−تربيع الإنحرافات:

16 1 0 16 1 4 4

4−التباين يساوي متوسط هده المربعات , أي يساوي :

$$\frac{16+1+0+16+1+4+4}{6}=\frac{42}{6}=7$$

حساب التباين من البيانات المبوبة:

1-متغير منفصل:

إدا كانت القيم X,X2..............., XN تمثل بيانات مجتمع أو عينة . و كانت n,n2............nn تمثل التكرارات المقابلة لها. فإن التباين يعطى بالصيغة التالية

❖ تباين المجتمع: $\sigma^2=\frac{\sum_{i=1}^{n}(Xi-\mu)^2.n_i}{n}=$

او بالصيغة المختصرة : $\sigma^2=\frac{\sum x_i^2.n_i}{n}-\mu^2$

❖ تباين العينة: $S^2=\frac{\sum_{i=1}^{n}(Xi-\overline{X})^2.n_i}{n-1}$

او بالصيغة المختصرة : $S^2=\frac{1}{n-1}\left(\sum X_i^2 n_i - n\overline{X}^2\right)$

مثال:

المطلوب وبإستعمال المعطيات التالية التالي في الجدول حساب التباين:

18	16	15	12	11	10	Xi
2	6	7	10	9	6	ni

الحل: يمكن تنضيم الحل في الجدول التالي:

$Xi^2.n_i$	Xi^2	$X_i.n_i$	n_i	X_i
600	100	60	6	10
1089	121	99	9	11
1440	144	120	10	12
1575	225	105	7	15
1536	256	96	6	16
648	324	36	2	18
6888		516	40	Σ
172.2		12.9		

و بالتالي و باستعمال الصيغة المختصرة : $\sigma^2 = \frac{\sum x_i^2 . n_i}{n} - \mu^2$

$\sigma^2 = \frac{6888}{40} - (12.9)^2$ $\qquad$ $\sigma^2 = 5.79$

إدا كانت القيم x2,x.............., xN تمثل مراكز الفئات لبيانات مجتمع أو عينة و كانت

n2,n1...........nn تمثل التكرارات المقابلة لها .فإن التباين يعطى بالصيغة التالية:

❖ تباين المجتمع: $\boxed{\sigma^2 = \frac{\sum_{i=1}^{n} (xi - \mu)^2 . n_i}{n} =}$

او بالصيغة المختصرة : $\sigma^2 = \frac{\sum x_i^2 . n_i}{n} - \mu^2$

❖ تباين العينة: $\boxed{S^2 = \frac{\sum_{i=1}^{n} (xi - \bar{X})^2 . n_i}{n-1}}$

او بالصيغة المختصرة : $\boxed{S^2 = \frac{1}{n-1} \left(\sum x_i^2 n_i - n\bar{X}^2 \right)}$

مثال :

البيانات التالية تمثل توزيع أجور عمال مؤسسة وطنية الوحدة: 10^3 دج

68 − 58	58 − 48	48 − 38	38 − 28	28 −18	Xi
4	15	17	22	18	Ni

أوجد التباين:

يمكن تتضيم الحل في الجدول التالي:

Xi^{2} ni	Xi^2	xi.ni	Xi	Ni	Xi
9522	529	414	23	18	28 −18
23958	1089	726	33	22	38 − 28
31433	1849	731	43	17	48 − 38
42135	2809	795	53	15	58 − 48
15876	3969	252	63	4	68 − 58
122924		2918		76	Σ
161742105		383947368			

و بالتالي و باستعمال الصيغة المختصرة : $\sigma^2 = \frac{\sum x_i^2 . n_i}{n} - \mu^2$

$$\sigma^2 = \frac{122924}{76} - (38.39)^2 \qquad \sigma^2 = 143.26$$

6.2. الإنحراف المعياري:

الإنحراف المعياري هو الجدع التربيعي للتباين. يعتبر القيمة الأكثر إستخداما في النضريات والقوانين الإحصائية من بين مقاييس التشتت الإحصائي لقياس مدى التبعثر الإحصائي. أي أنه يدل على مدى إمتداد مجالات القيم ضمن مجموعة البيانات الإحصائية. ويرمز له بالرمز

في حالة المجتمع و S في حالة بيانات العينة فمثلا لو أردنا حساب الإنحراف المعياري في البيانات المبوبة فإنه يعطى بالصيغة التالية:

❖ الانحراف المعياري للمجتمع:

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(Xi-\mu)^2.n_i}{N}}=$$

او بالصيغة المختصرة :

$$\sigma=\sqrt{\frac{\sum x_i^2.n_i}{n}-\mu^2}$$

❖ الانحراف المعياري للعينة:

$$S=\sqrt{\frac{\sum_{i=1}^{n}(Xi-\bar{X})^2.n_i}{n-1}}$$

او بالصيغة المختصرة :

$$S=\sqrt{\frac{1}{n-1}\left(\sum X_i^2 n_i-n\bar{X}^2\right)}$$

بعض خصائص الإنحراف المعياري:

1-الإنحراف المعياري للمقدار الثابت يساوي الصفر

2-قيمة الإنحراف المعياري دائما موجبة

3-كلما كان التشتت كبيرا حول الوسط كلما كان الإنحراف المعياري كبيرا و العكس صحيح

4-إدا أضيف مقدار ثابت إلى كل قيمة من قيم البيانات. فغن الانحراف المعياري للقيم الجديدة تساوي الانحراف المعياري للبيانات الأصلية.

3.6 .الإنحراف المتوسط

يعرف الانحراف المتوسط بأنه المتوسط الحسابي للقيم المطلقة للإنحرافات القيم عن متوسطها الحسابي ونرمز له بالرمز:

$$E_{\bar{X}} = \frac{\sum_{i=1}^{k} |X_i - \bar{X}|}{n}$$

المحاضرة 7: معامل الاحتلاف

1.7معامل الإختلاف CV (Coefficient de Variation)

يعرف معامل الإختلاف بأنه نسبة مقياس التشتت إلى المتوسط الحسابي. يستعمل غالبا عند مقارنة التشتت بين مجموعتين حيث لا يكفي مقارنة القيم المطلقة للإنحرافات المعيارية مع بعضها لأن نتائج هده المقارنة ستعطي أحكاما خاطئة نضرا لتأثر كل مجموعة بحجمها.فكلما كانت قيمة هدا المعامل كبيرة دل دلك على وجود تشتت كبير بين مفردات التوزيع و العكس صحيح.

فإدا رمزنا لمعامل الإختلاف بالرمز CV فإن:

معامل الإختلاف لبيانات العينة:

$$\boxed{C.V = \frac{S}{\overline{\overline{X}}} \times 100}$$

معامل الإختلاف لبيانات المجتمع:

$$\boxed{C.V = \frac{\sigma}{\overline{\overline{X}}} \times 100}$$

مثال:

إدا علمت أن الوسط الحسابي والإنحراف المعياري لمساهمة مجموعتين من رجال الأعمال في الجمعيات الخيرية مدونة في الجدول التالي:

	المجموعة الأولى	المجموعة الثانية
X	60	45
S	32	22

أوجد معامل الإختلاف بين المجموعتين.

الحل:

معامل الإختلاف المجموعة الأولى:

$$c.v_1 = \frac{s_1}{\overline{x}_1} \times 100 = \frac{32}{60} \times 100 = 53.33\%$$

معامل الإختلاف المجموعة الثانية:

$$c.v_2 = \frac{s_2}{\overline{x}_2} \times 100 = \frac{22}{45} \times 100 = 48.88\%$$

بما أن معامل اختلاف المجموعة الأولى أكبر من معامل اختلاف المجموعة الثانية. فإن تشتت المجموعة A أكبر من تشتت المجموعة B التي تعتبر أكثر تجانسا وأكثر تجانسا وأقل ابتعادا عن قيمتها المتوسطة.

ملاحظة: يمكن استعمال في حالة الجداول معامل اخر وهو معامل التغير الربيعي وفقا للقاعدة:

$$\boxed{CV_Q = \frac{Q_3 - Q_1}{Q_3 + Q_1}}$$

المحور الثاني: معامل الارتباط

محاضرة رقم8 :دراسة العلاقة بين المتغيرين

مقدمة:

تناولنا في الفصـــول الســـابقة دراســـة متغير واحد أو ظاهرة واحدة من حيث قياس وحساب متوسط هذه الظاهرة وكذلك حساب مقياس لتشتتها. ولكن في الحياة العملية كثيراً ما يحتاج الباحث لدرا سة العلاقة بين ظاهرتين (أو متغيرين) لمعرفة مدى الارتباط بينهما ونوع هذا الارتباط. فقد يريد الباحث معرفة ما إذا كان هناك علاقة بين التدخين والإصابة بمرض في الرئة، أو بين درجة تعليم الشـــخص ومســـتوى دخله. أو بين الحالة التعليمية والحالة الاجتماعية للناخب. وكما نرى فإنه يمكن أن نذكر الكثير بين الأمثلة في مختلف المجالات بل قد يرغب الباحث في دراســـة العلاقة بين أكثر من متغيرين في وقت واحد. فمثلاً قد يريد الباحث معرفة تأثير درجة التعليم ومستوى الدخل وحجم الأسرة على درجة الوعي الـ سيا سي للـ شخص. في هذا المثال يريد الباحث معرفة تأثير ثلاثة متغيرات على متغير رابع. وفي هذا الكتاب سوف نركز على دراسة العلاقة بين **متغيرين اثنين فقط** وهو ما يعرف **بالارتباط " البسيط "** Simple Correlation. بينما الحالات التي تتناول الدراسة فيها **أكثر من متغيرين** تعرف **بالارتباط المتعدد** Multiple Correlation وهي – كما ذكرنا – خارج نطاق هذا الكتاب.

1.8. معامل الارتباط

أنواع العلاقة بين المتغيرين:

إذا كان المتغيران يتغيران معاً في الاتجاه نفسه بمعنى أنه إذا زاد أو نقص أحدهما، زاد أو نقص الأخر، فإن العلاقة بينهما تكون طردية والارتباط بينهما يكون موجباً. مثال ذلك العلاقة بين زيادة حجم الطبقة الوسطى في المجتمع وزيادة الاستقرار السياسي.

وإذا كان المتغيران يتغيران معاً ولكن في عكس الاتجاه بمعنى أنه إذا زاد أحدهما نقص الأخر، أو إذا نقص أحدهما زاد الأخر، فإن العلاقة بينهما تكون عكسية والارتباط بينهما يكون سالباً. مثال العلاقة بين تدني مستوى الفرد التعليمي ودرجة الوعي الاجتماعي.

وتختلف العلاقات بين الظواهر من حيث القوة. فقد تكون العلاقة قوية جداً (أو حتى تامة)، وقد تكون متوسطة، أو ضعيفة، أو منعدمة تماماً. وفي هذا الفصل سوف نتناول بالتفصيل كيفية حساب الارتباط بين متغيرين سواء كان المتغيران كميين أو وصفيين (ترتيبيين أو اسميين)، أو أحدهما كمياً والأخر وصفياً.

تفسير بيانيا علاقة الارتباط بين المتغيرين المستقل والتابع:

يعتبر التمثيل البياني لتوزيع احداثية القيم (x, y) طريقة من طرق التي تسمح للباحث على تحديد نوع الارتباط بين المتغيرين وما إذا كان الارتباط قوياً وضعيفاً أو منعدماً، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة.

والمقـ صود بـ شكل التوزيع الاحداثية هو تمثيل قيم الظاهرتين بيانياً على المحورين، المتغير الأول X على المحور الأفقي، والمتغير الثاني Y على المحور الرأ سي، حيث يتم تمثيل كل احداثية من القيم بنقطة، فنحـ صل على شكل يمثل كيفية توزيع القيم وتنتـ شارها القيم على المـ ستوى، وهو الذي يـ سمى ايـ ضا بانتـ شار القييم وتـ شتتها وتبعثرها. وطريقة انتشـــار القيم تدل على وجود أو عدم وجود علاقة بين المتغيرين ومدى قوتها ونوعها. ويمكن تفسيرها كما يلي:

- إذا كانت تتوزع بشكل منتظم دل ذلك على وجود علاقة (يمكن استنتاجها)،
- إذا وقعت جميع النقاط على خط مستقيم، دل ذلك على أن العلاقة بينهما خطية وأنها ثابتة أو تامة. وهذه تمثل أقوى أنواع الارتباط بين المتغيرين " ارتباط تام
- إذا كانت النقط مبعثرة ولا تنتشر حسب نظام معين دل ذلك على عدم وجود علاقة بين المتغيرين أو أن العلاقة بينهما ضعيفة.

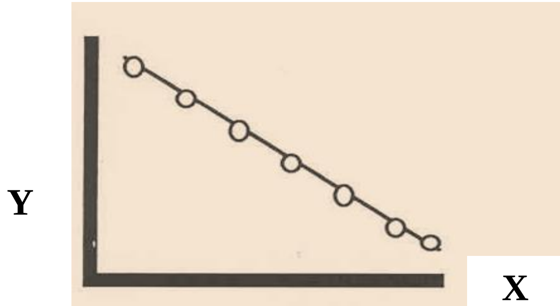
والأشكال التالية تظهر بعض أشكال الارتباط المعروفة:

". فإذا كانت العلاقة طردية فإن " الارتباط طردي تام " كما في الشـــكل الأول (أ).

ومثاله العلاقة بين الكمية المشتراة من سلعة والمبلغ المدفوع لشراء هذه الكمية.

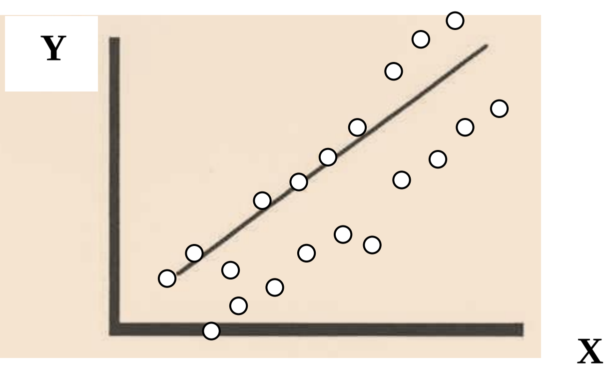
أما إذا كانت العلاقة عكســـية (وجميع النقاط تقع على خط مســـتقيم واحد) فإن " الارتباط عكسي تام " كما في الشكل الأول (ب). ومثال على ذلك العلاقة بين السرعة والزمن.

الشكل الأول (أ) ارتباط طردي تام(موجب)

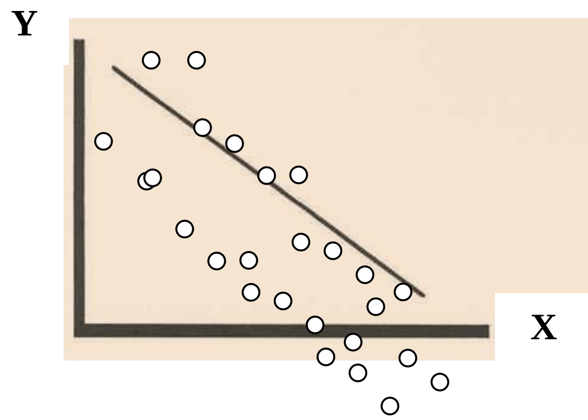


الشكل الأول (ب) ارتباط عكسي تام (سالب)

أما إذا كانت النقاط تأخذ شكل خط مستقيم ولكن لا تقع جميعها على الخط قيل أن العلاقة خطية (موجبة أو سالبة) كما في الشكل الثاني أ، ب.

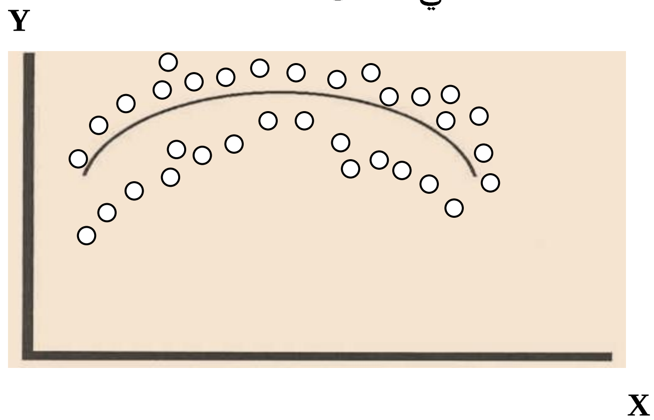


الشكل الثاني (أ) ارتباط موجب قوي (ارتباط خطي طردي)



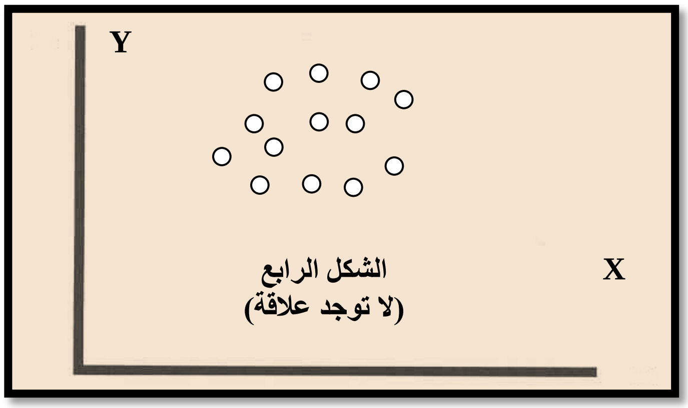
الشكل الثاني (ب) ارتباط سالب قوي (ارتباط خطي عكسي)

وإذا كانت العلاقة تأخذ شكل منحنى فإن الارتباط لا يكون خطياً "ارتباط غير خطي " Non Linear Correlation كما في الشكل الثالث :



الشكل الثالث
(ارتباط غير خطي)

أما إذا كانت النقاط تتبعثر بدون نظام معين فإن ذلك يدل على عدم وجود علاقة بين المتغيرين (أو أن العلاقة بينهما ضعيفة جداً) كالعلاقة مثلاً بين دخل الشخص وطوله كما في الشكل الرابع:



الشكل الرابع
(لا توجد علاقة)

محاضرة رقم 9 :قياس معامل الارتباط

1.9معامل ارتباط بيرسون

تمهيد:

يقاس الارتباط بين متغرين بمقياس إحصائي يسمى " معامل الارتباط " ويعكس هذا المقياس درجة أو قوة العلاقة بين المتغيرين واتجاه هذه العلاقة. وتنحصـــــر قيمة معامل الارتباط بين + 1، − 1. ويتم تفسير طبيعة الارتباط كمايلي كمايلي:

- إذا كانت قيمة معامل الارتباط تساوي + 1 فمعنى ذلك أن الارتباط بين المتغيرين طردي تام، وهو أقوى أنواع الارتباط الطردي بين متغيرين.
- إذا كانت قيمة معامل الارتباط تساوي − 1 فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام، وهو أقوى أنواع الارتباط العكسي بين متغيرين.
- إذا كانت قيمة معامل الارتباط تساوي صفر، فمعنى ذلك أنه لا يوجد ارتباط بين المتغيرين.
- كلما اقتربت قيمة معامل الارتباط من + 1 أو − 1 كلما كان الارتباط قوياً،
- كلما اقترب من الصفر كلما كان الارتباط ضعيفاً.
- كلما كانت العلاقة قوية بين المتغيرين كلما اقترب معامل الارتباط من + 1 أو − 1
- فإذا وصلت قيمة المعامل إلى + 1 أو − 1 كان الارتباط تاماً بين المتغيرين.

- وأنه كلما كانت العلاقة ضـــــعيفة بين المتغيرين كلما اقترب معامل الارتباط من الصفر،
- فإذا وصلت قيمة المعامل إلى الصفر كان الارتباط منعدماً بين المتغيرين. ومعنى ذلك أيضـــاً أنه لا يوجد ارتباط بين متغيرين تكون قيمة المعامل فيه أكبر من + 1 ولا أصغر من – 1.

يتم قياس العلاقة بين متغيرين كميين، أومتغيرين ترتيبين، وأخيراً متغيرين أسميين.

معامل بيرسون للارتباط الخطي البسيط:

يفترض بيرسون Pearson أن المتغيرين كميان، وأن العلاقة بينهما خطية (أي تأخذ شكل خط مستقيم، أنظر الشكل الثاني من أشكال الانتشار). بمعنى آخر، فإن أفضل مقياس للارتباط بين مثل هذين المتغيرين – حســـب رأي بيرســـون – هو عن طريق حســـاب انحرافات كل من المتغيرين عن وسطه الحسابي وقسمة هذه الانحرافات على الانحراف المعياري لكل منهما، فنحصـــل على ما يســـمى بالوحدات المعيارية لكل متغير. ويكون معامل ارتباط بير سون هو " متوسط حاصل ضرب هذه الوحدات المعيارية ". ومعامل الارتباط يكون بدون تمييز.

وبالرموز، إذا فرضنا أن المتغيرين هما X , Y وأن لدينا عدد n من أزواج القيم هي :

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

وأن الوسط الحسابي للمتغير X هو $\overline{X}$ وللمتغير Y هو $\overline{Y}$ وأن الانحراف المعياري للمتغير X هو S_x وللمتغير Y هو S_y فإن معامل بيرســـون للارتباط الخطي والذي يرمز له بالرمز r هو :

القانون:

$$r = \frac{n\,\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{\left[n\,\Sigma x^2 - (\Sigma x)^2\right]\left[n\,\Sigma y^2 - (\Sigma y)^2\right]}}$$

مثال (1) :

البيانات التالية تمثل أعمار ثمانية من الناخبين ودخولهم اليومية بالدولار، والمطلوب حساب معامل بيرسون للارتباط الخطي بين الأعمار والدخول.

الأعمار x :	25	32	29	43	38	51	47	35
الدخول y :	10	18	15	35	40	62	100	50

الحل:

لحساب معامل بيرسون للارتباط الخطي يلزم حساب المجاميع:

$\sum x , \sum y , \sum xy , \sum x^2 , \sum y^2$ لذلك يتم تنظيم حساب هذه المجاميع كما في الجدول التالي:

x الأعمار	y الدخول	xy	x^2	y^2
25	10	250	625	100
32	18	576	1024	324
29	15	435	841	225
43	35	1505	1849	1225
38	40	1520	1444	1600
51	62	3162	2601	3844
47	100	4700	2209	10000
35	50	1750	1225	2500
300	330	13898	11818	19818

ثم نطبق في الصيغة المختصرة رقم (2) لمعامل الارتباط حيث $n = 8$:

أي أن معامل بيرسون للارتباط الخطي بين أعمار الناخبين ودخولهم اليومية يساوي 0.81 وهو ارتباط طردي (لأن إشارته موجبة) وقوى (لأنه قريب من الواحد الصحيح). بمعنى أخر، إن هناك علاقة طردية قوية بين عمر الناخب ودخله مقدارها 81 %. فمع زيادة عمر الناخب يزيد دخله، والعكس صحيح.

ملاحظة:

من خواص معامل بيرســـون للارتباط الخطي أنه لا يتأثر بالعمليات الحســـابية التي تجري على المتغيرين x , y. بمعنى أنه لا يتأثر بالطرح (أو الجمع)، ولا بالقســـمة (أو الضرب). أي إذا طرحنا (أو جمعنا) قيمة معينة من كل قيم x وقيمة أخرى من كل قيم y، أو قسمنا (أو ضربنا) قيم x على قيمة معينة وكل قيم y على قيمة أخرى فإن قيمة معامل الارتباط لا تتغير أي نحصل على القيمة نفسها.

المحاضرة رقم 10: معامل ارتباط الرتب

10. معامل الترتيب

قد يرغب الباحث في حســـاب معامل الارتباط بين رتب المتغيرين وليس بين القيم ذاتها، فقد يكون المتغيران **وصـــفيين ترتيبين** Ordinal **أو أن يكون أحد المتغيرين كمياً بينما الأخر و صفياً ترتيبياً**، أو أن يكون **المتغيران كميين**، ويكون اهتمام الباحث منصباً على الرتب أكثر من القيم

فإذا كانت رتب المتغيرين تسير في الاتجاه نفسه: بمعنى أن الرتب الأعلى للمتغير الأول تناظرها رتب أعلى للمتغير الثاني كانت العلاقة طردية بينهما. وإذا كانت الرتب الأعلى للمتغير الأول تناظرها رتب أدنى للمتغير الثاني كانت العلاقة بينهما عكسية. ففي مثالنا السابق عن العلاقة بين دخل الناخب وعمره، كان الناخب الأكبر عمراً (بصفة عامة) هو الأعلى دخلاً، فمن الواضح أن العلاقة بينهما طردية، أما إذا كان الناخب الأكبر عمراً (بصفة عامة) هو الأقل مشاركة في العمل السياسي فإننا في هذه الحالة نكون أمام علاقة عكسية.

ولحساب معامل ارتباط الرتب هناك طرق مختلفة أهمها معاملي سبيرمان وكيندال.

أ.معامل سبيرمان لارتباط الرتب:

لحســـاب معامل ســـبيرمان لارتباط الرتب يقوم بترتيب كل من المتغيرين ترتيباً تـ صاعدياً أو تنازلياً (أما تـ صاعدياً لكلا المتغيرين أو تنازلياً لكليهما). وفي حالة الترتيب التصاعدي تأخذ أقل قيمة من قيم المتغير الرتبة رقم 1، والقيمة الأعلى منها مباشرة الرتبة رقم 2 وهكذا (بالنســـبة لكل من المتغيرين). أما في حالة الترتيب التنازلي تأخذ أكبر قيمة من قيم المتغير الرتبة رقم 1، والقيمة الأقل منها مباشرة الرتبة رقم 2 وهكذا (بالنسبة لكل من المتغيرين). وعند تساوي قيمتين (أو أكثر) من قيم المتغير نعطي كل قيمة رتبة مختلفة (كما لو كانت القيم غير متساوية) ثم نحسب متوسط هذه الرتب، ويعطى هذا المتوسط لكل من هذه القيم المتساوية.

وبعد ترتيب المتغيرين نحســـب الفروق بين رتب كل من المتغيرين (ونرمز للفروق بالرمز d) ثم نقوم بتربيع هذه الفروق ونحصل على مجموعها أي نحصل على $\sum d^2$ ثم نعوض في معامل سبيرمان لارتباط الرتب والذي يأخذ الشكل التالي:

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

حيث: $\sum d^2$ هو مجموع مربعات الفروق بين رتب المتغيرين، n هي عدد أزواج القيم.

مثال (4):

البيانات التالية تمثل إجابات عينة من ســـبعة أشـــخاص حول برامج الضـــمان الاجتماعي، ومدى ملاءمتها لحاجات الناس.

السؤال الأول	جيدة	مقبولة	ممتازة	جيدة	جيدة جداً	مقبولة	جيدة
السؤال الثاني	جيدة جداً	مقبولة	جيدة جداً	جيدة	جيدة	جيدة	ممتازة

والمطلوب حساب معامل سبيرمان لارتباط الرتب بين هذين السؤالين؟

<u>الحل :</u>

تنظم الحل في الجدول التالي مع ملاحظة ما يلي:

1 – بالنسبة للسؤال الأول، فإن التقدير الأعلى سيحصل على الرتبة رقم 1 والأقل منه مباشرة سيحصل على الرتبة رقم 2 وهكذا.. أي أن الترتيب تنازلي. ونكرر العمل نفسه مع السؤال الثاني.

2 – عند حصول إجابتين أو أكثر على التقدير نفسه نعطي لكل إجابة مبدئياً رتبة **<u>كما لو كانوا مختلفين</u>** ثم نحسب متوسط هذه الرتب، وهذا المتوسط هو الذي يعطى لكل إجابة.

3 – ثم نحســـب الفروق بين رتب الســـؤالين ونرمز لها بالرمز d ثم نربع هذه الفروق

فنحصل على d^2 ونعوض في القانون عن $\sum d^2$ مع ملاحظة أن n = 7.

السؤال الأول X	السؤال الثاني Y	رتب X	رتب Y	d الفرق بين الرتب	d^2 مربعات الفرق
جيدة	جيدة جداً	4	2.5	1.5	2.25
مقبولة	مقبولة	6.5	7	- 0.05	0.25
ممتازة	جيدة جداً	1	2.5	- 1.5	2.25
جيدة	جيدة	4	5	- 1.0	1.00
جيدة جداً	جيدة	2	5	- 3.0	9.00
مقبولة	جيدة	6.5	5	1.5	2.25
جيدة	ممتازة	4	1	3.0	9.00
المجموع					**26.0**

$$r_s = 1 - \frac{6(Sd^2)}{n(n^2-1)} = 1 - \frac{6\left(\sum d^2\right)}{n(n^2-1)} = 1 - \frac{6(26)}{7(49-1)}$$

$$= 1 - \frac{156}{336} = 1 - 0.46$$

$$r_s = 0.54$$

وهذا يعني أن الارتباط بين إجابات المبحوثين بالنسبة للسؤالين هو ارتباط طردي متوسط. وبالتالي فليس بالضرورة أن يكون رأي المجيبين في برامج الضمان الاجتماعي تعني ملاءمتها لحاجات الناس.

مثال (5):

البيانات التالية تمثل أعداد الساعات التي ذاكرها عشرة طلاب والدرجات التي حصلوا عليها في امتحان أحد المقررات:

عدد الساعات X	10	6	12	14	11	6	19	16	3	9
الدرجات y	12	5	14	15	13	9	16	13	2	11

أحسب معامل سبيرمان لارتباط الرتب.

الحل :

كما في المثال السابق ننظم الحل في الجدول التالي مع ملاحظة أن $n = 10$

عدد الساعات X	الدرجات Y	رتب X	رتب Y	الفروق d	d^2

0	0	6	6	12	10
2,25	1.5	9	8.5	5	6
1	1	3	4	14	12
1	1	2	3	15	14
0,25	0.5	4.5	5	13	11
0,25	0.5	8	8.5	9	6
0	0	1	1	16	19
2,25	−1.5	4.5	2	13	16
25	5	5	10	2	3
1	−1	8	7	11	9
33					المجموع

وبالتعويض في القانون حيث $\sum d^2 =$ ، n = 10 نحصل على :

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

$$r_s = 0.973$$

$$r_s = 1 - \frac{6(33)}{10(10^2-1)} =$$

$$r_s = 0.80$$

مما يعني أننا أمام علاقة طردية قوية بين المتغيرين. فكلما زادت عدد الساعات التي يدرســـها الطالب في هذا المثال، كلما زادت درجاته في الامتحانات قوة وذلك بما نســـبته .82 %

ب.معامل كيندال لارتباط الرتب Kendall.s Tau :

من المعاملات والاختبارات اللابرامترية المهمة لقياس ارتباط الرتب معامل كيندال (Kendall's. Tau)

وهو نســـبة الفرق بين عدد الأزواج المتجانســـة (المتطابقة) وعدد الازاج غير المتجانسة (المتنافرة). يتطلب حساب (مقلوبات) الترتيب بدلا من معالجة الرتب نفسها كما هو متبع في ســـبيرمان، فمعامل كندال تاو يعتمد فقط على عدد المقلوبات لكل زوج من المشاهدات أو أفراد في فئات الترتيب.

يســـتخدم هذا النوع من الارتباط للتأكيد من أزواج رتب المتغيرات تحتفظ بنفس المستوى ومعالجة التطابق أو عدم التطابق في حال حدوثه بين هذه الأزواج في المتغيرات التي نخضعها للدراسة.

بالرمز kt يأخذ الشكل التالي :

$$\frac{C-D}{C+D}$$

حيث :

C: عدد قيم **الأكبر** من قيمة **رتبة نتيجة الفصـــــل الثاني بناءا على ترتيب نتائج الفصل الأول**

D: عدد قيم **الأصـــــغر** من قيمة **رتبة نتيجة الفصـــــل الثاني بناءا على ترتيب نتائج الفصل الأول**

مثال(7) : البيانات التالية تمثل درجات 12 تلميذا في في مادة الفيزياء:

رقم الطالب	1	2	3	4	5	6	7	8	9	10	11	12
الفصل الأول X	15	7	3	18	2	6	10	11	13	14	12	9
الفصل الثاني y	14	10	12	17	6	4	8	5	13	9	7	11

أحسب معامل كيندال لارتباط الرتب بين درجات الطلاب في المقررين.

الحل : نعيد ترتيب البيانات تصـــاعدياً حســـب قيم X ونحســـب أعداد الأزواج المتوافقة والأزواج المختلفة كما يلي :

D	C	رتبة نتيجة الفصل الثاني بناءا على ترتيب نتائج الفصل الأول	ترتيب التلاميذ حسب نتائج امتحان الأول تصاعديا	ترتيب التلميذ في الفصل الثاني	ترتيب التلميذ في الفصل الأول	الفصل الثاني	الفصل الأول	ترتيب التلاميذ
0	11	1	1	2	2	14	15	1

2	7	10	9	6	2	2	10	0
3	3	12	11	4	3	7	5	4
4	18	17	1	1	4	3	8	0
5	2	6	12	10	5	9	3	4
6	6	4	10	12	6	11	1	5
7	10	8	7	8	7	8	2	3
8	11	5	6	11	8	5	3	1
9	13	13	4	3	9	6	2	1
10	14	9	3	7	10	12	0	2
11	12	7	5	9	11	4	1	0
12	9	11	8	5	12	10	0	0
	$N_c =$ 49	$N_d =$ 12					46	20

وبالتعويض في معامل كيندال لارتباط الرتب نحصل على :

$$\frac{46-20}{46+20} = 0{,}39$$

أي أن هناك ارتباط رتب موجب متوسط بين درجات الطلاب في المقررين.

مقارنة معامل سبيرمان ومعامل ارتباط كاندال تاو:

له خصائص أفضل من سبيرمان

يقدم حل وتفسير لأزواج المتغيرات التطابقة والمتنافرة في الرتب

قيمة الدلالة الإحصائية لمعامل ارتباط كندل أكثر دقة في العينات صغيرة الحجم

المحاضرة رقم 11: معامل الاقتران (فاي) ومقياس التوافق

11.2معامل الاقتران ومقاييس التوافق: سبق لنا وأن درسنا العالقة بين المتغيرين(x,y)وإيجاد معامل االرتباط بينهما ، وذلك بقياس قوة االرتباط وإيجاد العالقة بينهما (طردية أو عكسية)، كما هي في معامل بيرسون وسبيرمان. أي معامل بيرسون الذي يعطي لنا قوة االرتباط في حالة البيانات الكمية، ومعامل ارتباط سبيرمان الذي يستخدم إليجاد قوة ارتباط للرتب في حالة البيانات الكمية الوصفية التي لها صفة الترتيب، لكن قد تكون هناك بيانات وصفية لكنها غير مرتبة مثل: رأي اجابي أوسلبي أو مكان الاقامة،......، ولكي نقيس قوة االرتباط لهذه البيانات لزم البحث عن مقياس مناسب لقياس االرتباط بين هذه الصفات، ومن بين هذه المقاييس معامل الإقتران الذي يستخدم عندما تكون لكل من الظاهرتين صفتين فقط، أما معامل التوافق فهي تستخدم في حالة لكل من الظاهرتين أو إحداهما على الأقل لها أكثر من صفتين.

أ. **معامل الاقتران:** يستخدم للعلاقة بين متغيرين اسميين كل منهما ثنائي التقسيم.بمعنى آخر رقياس قوة االرتباط بين ظاهرتين كل منهما ذات صفتين فقط، ويرمز لها بالرمز (CC) مثل: دراسة الإرتباط بين xو y في الجدول التالي:

	X1	X2	Sum
Y1	a	b	a+b
Y2	c	d	C+d
المجموع	a+c	b+d	a+b+c+d

$$r_{cc} = \frac{\boldsymbol{a \times d \ - b \times C}}{\sqrt{(\boldsymbol{a+b})(\boldsymbol{c+d})(\boldsymbol{a+c})(\boldsymbol{b+d})}}$$

مثال تطبيقي: أوجدي قيمة معامل الاقتران بين النوع (ذكر/ أنثى) وبين رأي التلاميد حول التوجيه (موافق/ غير موافق) للبيانات التالية:

الجنس	غير موافق	موافق	المجموع
ذكر	12	7	19
أنثى	10	5	15
المجموع	22	12	36

$$r_{cc} = \frac{\boldsymbol{12 \times 5 \ - 7 \times 10}}{\sqrt{\boldsymbol{22 \times 12 \times 19 \times 15}}} =$$

$$= \frac{\boldsymbol{60 - 70}}{\sqrt{\boldsymbol{75240}}} = \frac{\boldsymbol{-10}}{\boldsymbol{274.299}} = \boldsymbol{-0.037}$$

ب.**معامل التوفيق:**

يـ ستخدم في درا سة العلاقة مابين البيانات الو صفية التي لكل منهما أو إحداهما على الأقل مق سمة إلى أكثر من صفتين، هنا لا يمكن ا ستخدام مقياس االقتران، بل نلجأ إلى ا ستخدام مقياس آخر وهو معامل التوافق، ونرمز له بالرمز C ، ولحـ ساب معامل التوافق نفترض أن لدينا ظاهرة x لها n من الصفات، والظاهرة y والتي لها y من الصفات، كما هي مبينة في الجدول التالي:

رغبة الآباء رغبة الأبناء	علمي	رياضي	أدبي
علمي	15	7	22
رياضي	25	11	18
أدبي	11	13	16
المجموع			

المحاضرة رقم 12: لامبادا

12.2مفهوم معامل لامبادا

يعد معامل لامبادا أو جثمان من المقاييس الاسميه لقياس قوه الاقتران بين المتغيرات الاحصائيه او ما يسمى بمقياس قوه الترجيح، اذ يقوم على فكره التقليل النسبي للخطا. او قيمه الخطا الذي تم تخفيضه عند استخدام قيمه المتغير(س) للتنبؤ بقيمة ص. أي يعتمد هذا الأخير، على كيفية التنبؤ بقيمة المتغير التابع عندما نعرف قيمة المتغير المستقل ان هذا النوع من المقياس انه يقارن الاخطاء في موقعين مختلفين:

الخطأ الاول (E1): عندما لا نستخدم المتغير المستقل لغايات التنبؤ

الخطأ الثاني (E2ُ) عندما نستخدم المتغير المستقل لغايه التنبؤ.

نقدم في الجدول التالي حساب الخطأين:

E 1= N – المجموع الأكبر لصف من الصفوف ↓ لا نستخدم المتغير المستقل لغايات التنبؤ
E2= (مجموع العمود الأول – التكرارات الأعلى الموجودة في خلية من الخلايا التي تنتمي إلى العمود الأول) +.....+ (مجموع العمود الأخير – التكرارات الأعلى الموجودة في خلية من الخلايا التي تنتمي إلى العمود الأخير) ↓

نستخدم المتغير المستقل لغايه التنبؤ

يحسب معامل لامبادا باستعمال القانون التالي:

قيمة لامبادا ƛ= $\frac{E1-E2}{E1}$

مثال تطبيقي:

أراد أحد الباحثين في علم الإجتماع أن يدرس العلاقة بيالتوجيه الى تخصصات علم الاجتماع والرضا عن هذا الأخير، فأختار عينة تتكون من 305 فردا من مستويات إقتصادية – إجتماعية عليا، متوسطة ودنيا، فتحصل على النتائج التالية:

المستوى الجامعي / الرضا عن التوجيه	السنة الأولى	السنة الثانية	السنة الثالثة	المجموع
راضي بشكل عالي	60	25	30	115
راضي	10	50	25	85

غير راضي	10	20	75	105
المجموع	80	95	130	305

المطلوب:

-أحسب قيمة لامبادا λ.

الحل:

-حساب قيمة لامبادا λ:

قيمة لامبادا λ= $\frac{E1-E2}{E1}$

-حساب الخطأ الأول:

خ1= ن – المجموع الأكبر لصف من الصفوف

خ1= 305 -115

خ1=190.

-حساب الخطأ الثاني:

خ2= (مجموع العمود الأول – التكرارات الأعلى الموجودة في خلية من الخلايا التي تنتمي إلى العمود الأول) +.....+ (مجموع العمود الأخير –التكرارات الأعلى الموجودة في خلية من الخلايا التي تنتمي إلى العمود الأخير).

خ2= (80–40) + (95 –50) + (130 –75) = 40+ 45+ 55

خ2= 140 .

–حساب قيمة لامبادا λ:

قيمة لامبادا λ= $\frac{E1-E2}{E1}$

$$\text{قيمة لامبادا } λ = \frac{190 - 140}{190} = 0.26$$

المحاضرة رقم 13: نسبة الارتباط (ايتا)

2.13نسبة الإرتباط (إيتا)

نحاول من خلال هذه المحاضرة التطرق إلى معامل من معاملات الإرتباط ألا وهو نسبة الإرتباط (إيتا)، حيث أنه إذا كنا بصدد تقدير العلاقة بين متغيرين كميين فإن أول خطوة يجب القيام بها هي التأكد من طبيعة العلاقة، هل هي علاقة خطية أم غير خطية، فإذا كانت خطية نستخدم معامل إرتباط بيرسون، أما إذا كانت غير خطية فنستخدم نسبة الإرتباط إيتا".

1- مفهوم نسبة الإرتباط (إيتا):

إن نسبة الإرتباط (إيتا) هو معامل، يستخدم لتحديد العلاقة بين متغيرين كميين منفصلين، على مستوى القياس الفتري (المسافات المتساوية) النسبي، حيث أن قيمة معامل إيتا موجبة دوما، أي محصورة بيت 0 و+1.

مجال استخدامه

فهذا المعامل يستخدم عندما تكون العلاقة غير خطية، وتكون البيانات بالنسبة للمتغير التابع مصنفة ضمن مقياس مسافات على الأقل، والبيانات بالنسبة للمتغير المستقل مصنفة ضمن مقياس إسمي أو تراتيبي. وفي هذا الصدد فإن مربع إيتا يشير إلى نسبة التباين الكلي في المتغير التابع والتي تعزى إلى المتغير المستقل.

2-كيفية حساب نسبة الإرتباط (إيتا):

لحساب نسبة الإرتباط (إيتا) نتبع عدة خطوات وعمليات حسابية إستنادا على القانون الخاص بهذا المعامل، والذي هو كما يلي:

القانون الخاص بنسبة الإرتباط (إيتا) هو :

$$\boldsymbol{n}_{\text{إيتا}} = \sqrt{\mathbf{1} - \frac{\sum(\boldsymbol{y} - \overline{\boldsymbol{y}\mathbf{1}})^2}{\sum(\boldsymbol{y} - \overline{\boldsymbol{y}})^2}}$$

خطوات حساب نسبة الإرتباط (إيتا):

يتم حساب نسبة الإرتباط (إيتا) كما يلي:

- في حالة غياب ترتيب قيم المتغير (x) أي عدم تقديم قيم x مرتبة في الجدول فإننا نرتب قيم المتغير x (المتغير المستقل) ترتيبا تصاعديا ، ثم نرتب قيم المتغير y (المتغير التابع) تبعا لترتيب قيم المتغير x .
- حساب $\bar{y}1$ وهو متوسط قيم المتغير y التابع المقابلة لقيم x المكررة. ويحسب بجمع مجموع قيم عدد القيم لتلك الفئة.
- **حساب** $\bar{y}$ وهو المتوسط الحسابي الكلي لقيم المتغير التابع. والذي يحسب كما يلي:

- ❖ إيجاد إنحراف القيم y عن متوسط قيم المتغير y التابع المقابلة لقيم x المكررة، وبعبارة أخرى (قيمة y – متوسط قيم المتغير y التابع المقابلة لقيم x المكررة) أي $(y - \bar{y}1)$
- ❖ –إيجاد مربع انحراف القيم y عن متوسط قيم المتغير y التابع المقابلة لقيم x المكررة ، أي مربع (قيمة y – متوسط قيم المتغير y التابع المقابلة لقيم x المكررة) والذي هو $(y - \bar{y}1)^2$.
- ❖ –حساب إنحراف قيم y عن المتوسط الحسابي الكلي لقيم المتغير y، أي (قيمة y – المتوسط الحسابي الكلي لقيم المتغير y) والذي هو $(y - \bar{y}1)$.
- ❖ –إيجاد مربع إنحراف قيم y عن المتوسط الحسابي الكلي لقيم المتغير y، أي مربع (قيمة y – المتوسط الحسابي الكلي لقيم المتغير y) والذي هو $(y - \bar{y}1)^2$
- ❖ –إيجاد قيمة إيتا.

ويمكن إختصار العمليات الحسابية السابقة في جدول كما يلي:

ملاحظة: المتغير x هو المتغير المستقل والمتغير y هو المتغير التابع.

مثال:

دراسة العلاقة بين متغير درجة رضى طلاب تجاه توجههم الى تخصص علم اجتماع التربية x ونتائج الامتحان y لعينة مكونة من 30 فرد وفق النتائج الموضحة في الجدول التالي:

نتائج الامتحان	درجة الرضى	ترتيب الطلبة
7	2	1
6	2	2
3	2	3
8	4	4
10	4	5
9	4	6
11	4	8
4	5	9
7	5	10
12	6	12
13	6	13
11	6	14
10	7	15
14	7	16
15	8	17
12	8	18
13	8	19
15	9	20
180		المجموع

المطلوب:

-أوجد قيمة إيتا.

الحل:

-قيمة إيتا هي:

-حساب المتوسط الحسابي الكلي لقيم المتغير y:

$$\boldsymbol{n}_{\text{إيتا}} = \sqrt{\mathbf{1} - \frac{\sum(\boldsymbol{y} - \overline{\boldsymbol{y}\mathbf{1}})^2}{\sum(\boldsymbol{y} - \overline{\boldsymbol{y}})^2}}$$

قيمة إيتا في الجدول التالي:

ترتيب الطلبة	درجة الرضى	نتائج الامتحان	متوسط نفس درجات الرضى	نتائج - درجة	مربع متوسط الفرق	متوسط النتائج - النتائج	مربع الفارق
1	2	7	5,33	1,67	2,78	-2	4
2	2	6	5,33	0,67	0,44	-3	9
3	2	3	5,33	-2,33	5,44	-6	36
4	4	8	9,50	-1,50	2,25	-1	1
5	4	10	9,50	0,50	0,25	1	1
6	4	9	9,50	-0,50	0,25	0	0
8	4	11	9,50	1,50	2,25	2	4
9	5	4	5,50	-1,50	2,25	-5	25
10	5	7	5,50	1,50	2,25	-2	4
12	6	12	12,00	0,00	0,00	3	9
13	6	13	12,00	1,00	1,00	4	16
14	6	11	12,00	-1,00	1,00	2	4
15	7	10	12,00	-2,00	4,00	1	1
16	7	14	12,00	2,00	4,00	5	25

17	8	15	13,33	1,67	2,78	6	36
18	8	12	13,33	−1,33	1,78	3	9
19	8	13	13,33	−0,33	0,11	4	16
20	9	15	15,00	0,00	0,00	6	36
المجموع		180			32,83		236

بعد حساب المتوسط الحسابي الكلي لقيم المتغير y ونختصر العمليات الحسابية من أجل

حساب

$$\boldsymbol{n}\text{إيتا} = \sqrt{\mathbf{1} - \frac{\mathbf{32,83}}{\mathbf{236}}} \;= \mathbf{0,86}$$

من النتيجة السابقة يمكن القول بأن هناك إرتباط قوي بين متغير درجة رضى الطلاب

تجاه توجههم الى تخصص علم اجتماع التربية و النتائج المتحصل عليها.

المحاضرة رقم 14: الإرتباط والسببية

2.14 .الارتباط والسببية

نتطرق في هذه المحاضرة إلى الإرتباط والسببية، فهما مصطلحين إحصائيين، فالإرتباط ليس هو السببية، حيث أن الإرتباط يركز على ما مدى وجود علاقة بين متغيرين، واحد مستقل والآخر تابع. فهو يركز على القوة والدرجة الموجودة بين المتغير المستقل والمتغير التابع. أما السببية فتركز على مدى تسبب المتغير المستقل في حدوث المتغير التابع على العموم.

1-مفهوم الإرتباط:

مصطلح إحصائي يُستخدم للتعبير على مدى قوة العلاقة الموجودة بين متغيرين عشوائيين أو أكثر، ويعني **الارتباط** التلازم في التغير، أي أن التغير في أحد المتغيرين يتبعه تغيّر في المتغير الثاني، فعلى سبيل المثال، حين يزداد حجم ساعات المراجعة ، يرتفع مستوى التحصيل الدراسي ، والعكس صحيح2

-مفهوم السببية :

السببية هي مصطلح مرتبطة بالسبب والنتيجة، وتعني ما مدى تسبب متغير في حدوث أو ضهور متغير آخر، كيف يسبب المتغير المستقل في حدوث المتغير التابع، أي أن وقوع ظاهرة معينة مرتبط بسبب معين.

3-الإرتباط والسببية:[1]

إن الإرتباط والسببية هما كلمتين مختلفتين، حيث أن معامل الإرتباط البسيط هو قياس الإقتران أو الإرتباط بين متغيرين (متغير مستقل ومتغير تابع). إذ أنه في غياب أية معلومات إضافية، فإن هذا المعامل لا يقدم لنا أية معلومات عن العلاقة السببية بين المتغيرين.

---

[1]- عبد الله فلاح المنيزل، عايش موسى غرايبة، الإحصاء التربوي، تطبيقات بإستخدام الرزم الإحصائية للعلوم الإجتماعية، دار المسيرة، بدون سنة. ص170، 171. بصيغة PDF.

إن الباحث يقوم بمعالجة متغير مستقل واحد أو عدة متغيرات مستقلة، وهذا من أجل التوصل إلى إستنتاج حول تأثير المتغير المستقل على المتغير التابع، ففي هذه الحالة فإن التجربة كفيلة، حيث تساعد الباحث على التأكيد على أن المتغير المستقل سبب التغيرات التي حدثت بالنسبة للمتغير التابع أي أن المتغير المستقل أدى إلى تغير المتغير التابع. هذا في حالة القيام بالتجربة، حيث أن الباحث يقوم بإختيار عشوائي لمجموعة أو أكثر من المجموعات التجريبية ولمجموعة أو أكثر من المجموعات الضابطة.

أما في الدراسات الإرتباطية فإن الباحث يختار متغيرين ويحاول معرفة إلى أي مدى يتغير أحدهما بتغير الآخر، فهو لا يقوم بمعالجة منظمة للمتغير المستقل لملاحظة أثره على المتغير التابع، حيث أن الدراسات الإرتباطية لا تتضمن مجموعة ضابطة. إذ أن إجراءات الدراسة الإرتباطية تجعل من غير الممكن التحقق من أن معامل الإرتباط قياس للسببية.

من خلال ما سبق ذكره، فإنه توجد ثلاثة فرضيات أساسية حول العلاقة بين المتغيرين المترابطين (x و y) ، والتي تكون كما يلي:

أولا: x يسبب y .

ثانيا: y يسبب x .

ثالثا: العلاقة بين x و y تسبب بواسطة متغير ثالث F.

عندما يكون هناك ثلاثة متغيرات في الدراسة، فإنه يوجد العديد من النماذج السببية:

x

–النموذج الأول: y، في هذا النموذج فإن x يسبب y و s يسبب y.

F

ولكن لا توجد علاقة سببية بين x و F.

y

–النموذج الثاني: x ، في هذا النموذج فإن x يسبب y و x يسبب s.

F

ولكن لا توجد علاقة سببية بين y و s.

–النموذج الثالثة: x ← F ← y ، في هذا النموذج فإن العلاقة السببية بين x و y يتوسطها متغير ثالث، وبالتالي فإن المتغير x له تأثير غير مباشر على المتغير y.

يمكن القول أن فحص الإرتباطات السببية في حالة الدراسات الإرتباطية يعتمد على النظرية، أي يتم الربط بين النظرية والبيانات. فالبيانات المتوفرة لدى الباحث عند إجراء أية دراسة تحدد النموذخ الدقيق للعلاقة السببية. فالإرتباط الموجود بين متغيرين لا يعني بالضرورة أن أحد المتغيرين يسبب المتغير الآخر.

إختبر كفاءاتك المكتسبة

الأسئلة:

3-لماذا يلجأ الباحث الى حساب التشتت

4-ماهو الهذف من حساب المنوال

5-هل معامل الارتباط يؤكد صحة الفرضيات الاحصائية

6-ماهي المعاملات الرتبية، أذكر اجابياتها.

7- ماهو الفرق بين المتغيرين المستقل و التابع.

8-هل يجب تقديم نتائج بدون متغير مستقل

التمرين: في دراسة ميدانية حول الدروس أثرالدروس الخصوصية على التحصيل الدراسي، قام الباحث طرح العديد من الأسئلة ومن أهمها

السؤال الأول:

هو المعدل المتحصل عليه قبل وبعد اللجوء الى الدروس الخصوصية. بعد تفريغ السؤال تحصل على النتائج التالية:

ترتيب التلميد	نتائج الرياضيات قبل الجوء الى الدروس الخصوصية	نتائج الرياضيات قبل الجوء الى الدروس الخصوصية
1	4	12
2	8	10
3	14	14
4	10	11
5	8	4
6	8	7
7	12	13
8	13	14
9	3	6
10	2	8
11	7	3
12	16	18
13	2	10
14	12	14
15	13	13
16	4	5
17	5	9

1-حسب رأيك ماهي مقاييس النزعة المركزية والتشتت التي يمكن حسابها
2-إذا أراد الباحث دراسة العلاقة مابين التحصيل الدراسي والدروس الخصوصية، ماهو معامل الارتباط المناسب لهذه الدراسة.

السؤال الثاني:

حسب رأيك، هل توافق لجوء التلاميذ الى الدروس الخصوصي.

الجنس	ذكور	اناث
موافق تماما	15	34
موافق	20	54
حيادي	30	54
غير موافق	14	23
غير موافق تماما	45	12

1-ادرس العلاقة مابين المتغيرين باستعمال معامل ارتباط التوفيق ومعامل لامبادا، ماذا تستنتج.
2-هل يمكن أن يكون متغير آخر يؤثر عل المتغير التابع، فسر ذالك باستخدام معامل الارتباط والسببية

المراجع:

المراجع

المراجع باللغة العربية

1. عبد الله فلاح المنيزل وعايش موسى غرايبية"الاحصاء التربوي وتطبيقات باستخدام الرزم الاحصائية للعلوم الاجتماعية.
2. أمل محمد سلامة عبارى "طرائق الاحصاء الاجتماعي: التطبيقات العلمية في العلوم الاجتماعية، الاسكندرية، دار الوفاء لدنيا الطباعة والنشر، الطبعة الأولى 2003.
3. أنيس كنجو" الاحصاء وطرق تطبيقها في البحث العلمي"، بيروت، الطبعة الأولى، 1977.
4. باسمة المنلا "قياس العلاقات الاجتماعية السوسيومتري وتطبيقاته في الميدان التربوي وفي جماعات العمل"، بيروت، الطبعة الأولى، 1990.
5. جلاطو جيلالي " الاحصاء""، الجزائر، ديوان المطبوعات الجامعية، 1999.
6. عبد الكريم بوحفص"الاحصاء المطبق في العلوم الاجتماعية والانسانية"، الجزائر، ديوان المطبوعات الجامعية، 2009.
7. فضيل دليو، أسس البحث وتقنياته في العلوم الاجتماعية. قسنطينة: ديوان المطبوعات الجامعية،1997.
8. كلاس محمد' محاضرات في الاحصاء التطبيقي"، الجزائر، ديوان المطبوعات الجامعية، 1993.
9. نداء محمد الصوص"مبادئ الاحصاء"الرياض المملكة السعودية، الطبعة الأولى 2007.

المراجع باللغة الفرنسية

1. BBAERNARD PY ‘’ Statistique descriptive’’ Paris, Economica, 4emme édition, 2000.

2. WILLIAM FOX ‘’ Statistiques sociales’’ Bruxelles, de Boeck Université,3emme édition, 1998.
3. DAVID C HOWELL ‘’Méthodes Statistiques en sciences humaines, de Boeck Université,3emme édition, 2008.
4. GERARD COLOT ‘’ cours de statistique descriptive’, Paris, dunod décision
5. Pascal Bressoux La modelisation statistique appliquée aux sciences sociales, Bruxelles, de Boeck Université,2emme édition, 2011.