الـفـصل الأول
علـم الإحصـاء تـعـريـفـه و أهميـتـه

أولا : تعريف علم الإحصاء .

ثانيا : أهمية علم الإحصاء .

ثالثا : تطور علم الإحصاء .

رابعا: علاقة علم الإحصاء بالعلوم الاجتماعية .

أولا : تعريف علم الإحصاء

هو فرع من فروع الرياضيات يشمل النظريات و الطرق الموجهة نحو جمع البيانات ووصف البيانات و الاستقراء و صنع القرارات .[1]

و عندما نتكلم عن علم الإحصاء لا نعنى بذلك البيانات الإحصائية وإنما نقصد حينئذ الطريقة الإحصائية . وهى الطريقة التى تمكننا من جميع الحقائق عن الظواهر المختلفة فى صورة قياسية رقمية وعرضها بيانيا ووضعها فى جداول تلخيصية بطريقة تسهل تحليلها بهدف معرفة اتجاهات هذه الظواهر وعلاقات بعضها ببعض [2] .

ولقد كان الهدف الرئيسى من علم الإحصاء قديما هو عد أو حصر الأشياء المراد توفير بيانات إحصائية عنها ، وكانت الجهة التى تقوم بإعداد الإحصاءات على مستوى الدولة تعرف بمصلحة التعداد ولذلك كان التعريف القديم لعلم الإحصاء أنه علم العد ، أي العلم الذي يشتمل على أساليب جمع البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة .

ولكن مع تطور المجتمعات وتشابه جوانب الحياة الاقتصادية والاجتماعية الحديثة بها ، لم يعد مجرد توفير البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة يفى بحاجات متخذى القرارات وصانعى السياسة العامة إلى تكوين صورة

متكاملة الجوانب عن مجتمعهم والمجتمعات المحيطة بـــه . فقـــام العلماء بتحديث نظريات علم الإحصاء وأساليبه وأدواته لكى يعين الباحثين وغيرهم على استخلاص استنتاجات معينة من البيانـــات الكمية التى أمكن لهم جمعها عن طريق العد .

من ذلك على سبيل المثال ، أن نظرية العينـــات ســـاعدت الباحثين على استخلاص استنتاجات عديدة من دراسة عدد صغير من الأفراد أو الأشياء – *العينة* – وتعميم تلك الاستنتاجات علـــى المجتمع الذى سحبت منه العينة بأسره ولذلك يعرف علم الإحصاء حديثاً بأنه : (علم متكامل يتضمن الأســـلوب العلمـــى الضـــرورى لتقصى حقائق الظواهر واستخلاص النتائج عنها ، كمـــا يتضـــمن أيضاً أيضا النظرية اللازمة للقياس واتخـــاذ القـــرار فـــى كافـــة الميادين الاقتصادية والاجتماعية والسياسية والعسكرية)[3]

ثـــانـــيـــا : أهميـه علـم الإحصـاء

لقد أصبح لعلم الإحصاء أهميه بالغه فى حياتنـــا الحديثـــة فصارت الإحصاءات مألوفة لدينا وتمثل جانبا مهما من المعلومات التى نطالعها كل يوم مثل جداول النقاط التى تحرزها أنديـــه كـــره القدم وتنشر فى الصحف والمجلات والتقديرات الخاصة بالتنبؤات الجوية ومؤشرات البورصة وانجازات الحكومة فى مجال الإسكان والتعمير والتغيرات التي تطرأ على أسعار العملات وأثمان السلع . وربما يتساءل المرء عن أهميه الإحصاء بالنسبة لـــدارس علـــم

الاجتماع أو علم النفس معتقداً أن الإحصاء موضوع يـدخل فـى صميم تخصص التجـاريين والاقتصـاديين والواقـع أن الباحـث الاجتماعي والمتخصص فى العلوم الاجتماعية بوجه عام يحتـاج فى كثير من الأحيان إلى استخدام الأرقام لكي يلخص ويعرض بها مجموعه من المشاهدات التى تتعلق بظاهرة يهتم بدراستها ، فقد يطلب منه أن يقدم تقريراً عن مدي التطور الذى حققـه برنـامج معين لمحو الأمية بين نزلاء المؤسسة التي يعمل بها ، وقد يكلف بدراسة الأسباب التى تجعل الذكور أكثر تقدما وحرصا على التعليم من الإناث فى المدرسة التى يشتغل فيها .

ففي كل مناسبة من هذه المناسبات سـيحتاج الباحـث أو الدارس إلى أداة من الأدوات الإحصـائية لكـي يسـتخدمها فـى تلخيص أفكاره والتعبير عنها بصورة محدده ومؤثرة ، فالعبـارة التى مؤداها " لقد نجحنا فى محو أميه 90%من العاملين الأميـين بالمصنع " أقوى وأشد من العبارة التى مفادها : " لقد نجحنا فـى محو أمية عدد كبير من العاملين الأميين بالمصـنع " : [4] يحتـل الإحصاء (أو الأساليب الإحصائية) أهميه خاصـة فـى الأبحـاث العلمية الحديثة ، إذ لا تخلو أى دراسة أو بحث من دراسة تحليليه إحصائية تتعرض لأصل الظاهرة أو الظاهرات المدروسة فتصـور واقعها فى قالب رقمى ، وتنتهي إلى ابرز اتجاهاتهـا وعلاقاتهـا بالظاهرات الأخرى . [5]

إن دراسة الإحصاء أمر له فوائد كثيرة بالنسبة لدارسي العلوم الاجتماعية وخاصة بعد أن تفتحت أمامهم مجالات عمل كثيرة فى تنظيمات الشرطة والعلاقات العامة بالشركات ومراكز البحوث وغير ذلك من مجالات العمل المختلفة . بل إن المعرفة بالإحصاء قد تفيد الإنسان على المستوى الشخصي فتكسبه مهارة التخطيط لحياته الاقتصادية الخاصة .

ولكن ينبغي أن نشير إلى أن النتائج التى تسفر عن تطبيق أداة إحصائية أو أكثر ليست نتائج قطعيه أو غير قابله للتمحيص والمراجعة . فإذا كانت الأدوات الإحصائية تستطيع أن تعين المرء على وصف البيانات وتصميم التجارب وعلى اختبار العلاقات بين الأشياء والوقائع التى يهتم بها إلا أن ذلك لا يلغى بصيرته السوسيولوجية وخبرته المهنية .

وبعبارة أخري ، يقتصر دور الأدوات الإحصائية على توفير المؤشرات المبدئية التى تساعد الباحث على رفض أو قبول الفروض التى يقوم بدراستها في حدود درجه معينه من الثقة . والإحصاء أيضا أداه لا تستخدم إلا في العثور على إجابات عن أسئلة تتصل ببيانات يمكن التعبير عنها بصيغ كميه . وهناك فى مجال العلوم الاجتماعية موضوعات لا حصر لها لا يمكن صياغة البيانات الخاصة بها فى صورة كميه على نحو دقيق ، ومن ثم لا يستطيع الباحث استخدام التحليل الإحصائي فى دراستها .

من ذلك على سبيل المثال ، دراسة التجربة الدينية بين جماعه المؤمنين بدين معين ، إذ أن عدد مرات تردد المرء على المسجد أو على الكنيسة في الشهر ليس دليلا فى حد ذاته على انه من الصالحين ، ولكنه مؤشر مبدئي على الصلاح .

ومما يعكس أهميه علم الإحصاء أنها يستخدم فى توجيه عمليه جمع البيانات وفى تفسير العلاقات التى تعكسها تلك البيانات . ومن ابرز المجالات التى تستخدم فيها المعالجات الإحصائية إجراء المقارنة بين عديد من الأشياء فى كثير من المناسبات . ويمكننا القول أن الحياة الإنسانية سلسله من المواقف التى يتخذ فيها الفرد قراره بناء على ما تسفر عنه المقارنة التى يجريها بين عديد من الاحتمالات وهذه المقارنة فى جوهرها عمليه إحصائية تقترن بالقياس والتقييم والتقدير . فنجاح الإنسان فى حياته يتحدد وفق مقياس معين فى ذهنه يقدر به هذا النجاح ، وحرية الفرد فى مجتمعه تقاس أيضا وفق معايير يتعارف عليها الأفراد فى مجتمعهم .

وبعبارة أخرى ، إن حياتنا تذخر بعمليات من القياس والتقدير الإحصائي فنحن علي سبيل المثال ،عندما ننزل إلى السوق لشراء سلعه معينه ، فى موسم التنزيلات ، نهتم وبطريقه لا شعورية بحساب ثمن هذه السلعة بالنسبة إلى إجمالي النقود التي فى حوزتنا ونقدر ما إذا كان الباقي من هذه النقود وسوف

يكفينا حتى نهاية الشهر أم لا وما إذ كانت نسبة التنزيلات على السلعة حقيقية أو مزيفة 00000الخ فى كل هذه العمليات الفكرية نحن نستعين بعمليات إحصائية ومقارنات مستمرة بين المواقف المختلفة . فضلا عن ذلك ، إن ما نطلق عليه ظاهرة اجتماعيه أو طبيعيه ما هو فى الواقع إلا سلسله متكررة من الواقع التى يمكن رصد حدوثها المستمر عبر فترة من الزمن وبنفس الوتيرة بطريقه إحصائية . (6)

ثالثا : تطور علم الإحصاء

تطور علم الإحصاء وتطبيقاته عبر سنوات طويلة ، وتم ذلك بجهود كثيرة من العلماء من دول مختلفة وكان . التطور بطيئا إلى أن جاء القرن العشرين ليشهد معدلا هائلا للتطور فى النظريات الإحصائية فى مجالات كثيرة .

ويرجع الاهتمام بالإحصاء إلى عصور قديمه ، وان تعداد السكان عند القدماء المصرين وفي الصين أمثلة توضح اهتمام الحكومات منذ القدم بالمعلومات الاجتماعية وذلك لأغراض التنظيم والتخطيط فى أحوال السلم والحرب .

ويبدو أن كلمه إحصاء (statistics) قد ظهرت لأول مره عام1749 وهى مشتقه من الكلمة اللاتينية (status) أو الايطالية (statista) وتعني كلاهما الدولة السياسية . ومن الطبيعي أن تكون الدولة أول من اهتم بجمع البيانات وذلك لإدارة شؤون البلاد

خاصة عن السكان لأغراض حربية وضريبية ، وامتدت بعد ذلك لتشمل إحصاءات حجم السكان والمواليد والوفيات والإنتاج والاستهلاك والثروة 00000الخ . وهكذا بداء العلم وتطوره باعتباره علم الدولة أو علم الملوك .[7]

ولقد تطور علم الإحصاء من مجرد فكره الحصر والعد إلى أن أصبح الآن علما له قواعده ونظرياته ويرجع الفضل فى ذلك إلى كثير من العلماء من أمثال عائله برونلي Bernoulli وفردريك جاوس F.gauss وكيتليه Quetlet وجولتونF.galton وأخيرا كارل بيرسون Karl.pearson وبولي A.bowley وبول U.yule فيشر L.fisher و.....الخ [8] .

وجاء التطور في علم الإحصاء بصفه عامه ملازما وموازيا للتطور في نظرية الاحتمالات . فقد نشأت نظريه الاحتمالات على أساس رياضى فى (1494) بواسطة باسيولي Lucapacidi . ومن الدراسات الفلكية لكل من كبلر (1517-1630) Keplr وجاليليو (1564-1642) Galilio قاما بتطوير نماذج الاحتمالات . غير أن التاريخ الحقيقي لنظريه الاحتمالات بدء فى القرن السابع عشر حيث وضعت أسسها في عام 1654 بواسطة كلا من العالمين : باسكان .Pascal,B (1623 1662) عالم الرياضيات والفيزياء

والفيلسوف الفرنسى – وكذا العالم فرمـــات Fermat (1608 – 1665) .

وبعد ذلك بـــثلاث ســـنوات قـــام هينجينـــز Huygens (1629 – 1695) بنشر كتيب صغير فـــي موضـــوع المعالجـــة الرياضية لفرص الفوز في مباريات ورق اللعب وزهرة النرد .

وفى نفس الوقت تقريبا قـــام جرونـــت grunt (1620 – 1674) بنشر ملاحظاته عن معالجة البيانات المتعلقة بالحكومـــة خاصة في النـــواحي الطبيعيـــة والسياســـية والتجاريـــة والنمـــو والوفيات والأمراض.

وقد كان العمل الذى قام به هيجيتر دافعا للكثيرين لدراســـة النظريات والمشاكل المتعلقة بمباريات الصدفة ومـــنهم برنـــوللى (1654 – 1705) ودي موافر De Moivre(1667 – 1754) واربوثنوت Arbuthnott ولابلاس laplace (1749 – 1827) وجاوس Gauss(1777 – 1855). [9] ويعد العالم البلجيكى كتيليه (1796 – 1874) أول من وضـــع قواعـــد محـــددة لعلـــم الإحصاء ، وكلمة إحصاء في الوقت الحاضر ذات معـــان متعـــددة فمنها يفهم جمع المعلومات التى تبين الحالة في الدولة مثل عـــدد المواليد والوفيات وبيانات عن المحاصـــيل والتجـــارة الخارجيـــة ..... الخ ويسمى نشر الأجهزة الحكومية لمثل هذه المعلومـــات في شكل كتب وتقارير " بالإحصاء الرسمي ".

وأخيرا يفهم بالإحصاء فرع من العلم له نظريته الخاصة . وعلم الإحصاء ، شأنه في ذلك شأن أى فرع آخر من فروع العلم له أسلوبه وموضوعات البحث الخاص به [10]

وكلمة إحصاء (Statistics) لها ثلاث معانى :

(1) الإحصاءات أو البيانات : مثال ذلك إحصاءات السكان والمواليد والوفيات والإنتاج – الصادرات – الاستهلاك .

(2) المؤشرات المحسوبة من عينة (العينة هى مجموعة جزئية من الوحدات محل الدراسة)

(3) علم الإحصاء : وهو فرع من فروع الرياضيات يشمل النظريات والطرق الموجهة نحو جميع البيانات ووصف البيانات والاستقراء وصنع القرارات . [11]

ولقد تطور علم الإحصاء وتنوعت طرائقه ، وأصبح له من القواعد ما يمكنه من القيام كعلم مستقل يمكن الاستعانة به في رسم وتحديد السياسات الاجتماعية التى ينتهجها المجتمع. كما برز دور الإحصاء – بما يقدمه من بيانات وإحصاءات – فى عمليات التخطيط والتنمية التي تمر بها مجتمعاتنا اليوم [12]

ويمكن القول أن الإحصاء تخدم الباحثين في جميع الميادين العلمية وصانعى القرارات في شتي المجالات العملية ، ولا يكاد يخلو ميدان من ميادين البحث العلمى إلا وطرقته الإحصاء وساهمت فيه مساهمة فعالة . وقد أثار روبرت

بارسوز فى مستهل كتابه " التحليل الإحصائي " أن كلمة إحصاء لها أكثر من استخدام إلا أن أكثر الاستخدامات شيوعاً هو ذلك الذى يرى أن كلمة إحصاء تشير إلى تلك الأساليب والإجراءات التحليلية المستخدمة في معالجة البيانات الرقمية.

بمعنى أنه للحصول علي معلومات ذات قيمة من تلك البيانات الرقمية فإنها يجب أن تخضع للتحليل الاحصائىStatistical Analysis بمساعدة تلك الأساليب والإجراءات والأدوات التي توفرها لنا الإحصاء.

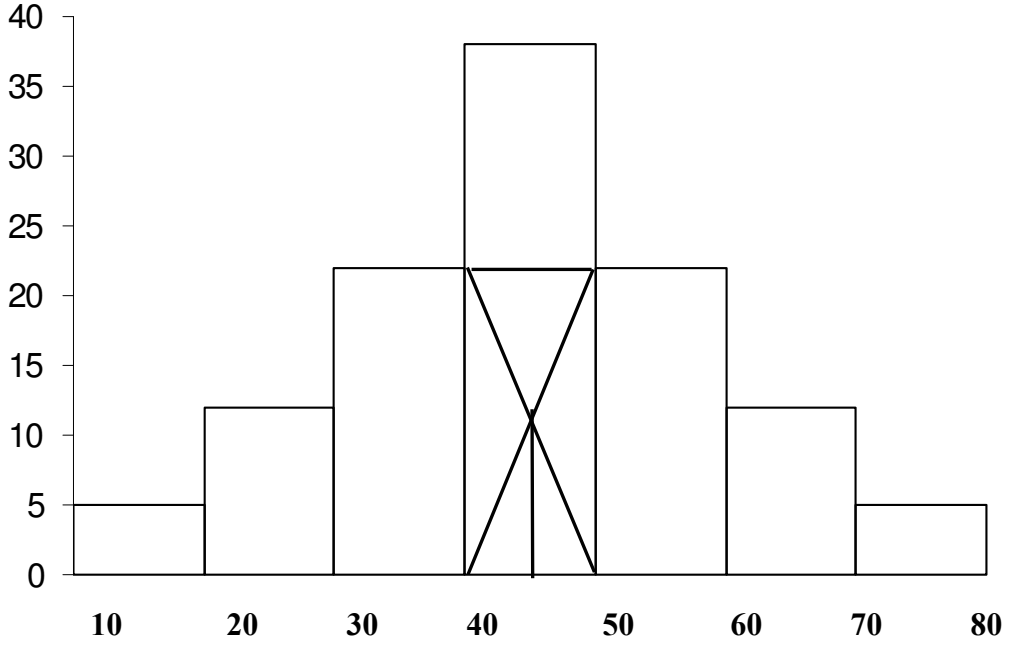
ويذهب كل من Whittaker, Startup إلى وجود ثلاثة استخدامات لكلمة إحصاء .

أ- للإشارة إلى الحقائق الرقمية التى جمعت بطريقة منتظمة من الواقع الاجتماعي.

ب- تشير إلى الأساليب المستخدمة فى جمع ، وتصنيف وتحليل البيانات الرقمية.

ج- للإشارة إلى صفة أو خاصية للعينة تحت الدراسة.

والقاموس الحديث لعلم الاجتماع الذى وضعه كل من George and Achilles Theocorson يقدم رؤية لا تختلف عما سبق فيما يتعلق بكلمة إحصاء سواء من حيث المعنى أو الاستخدام فهى تعني مجموعة من الأساليب التي



العلاقة بين الوسط والوسيط والمنوال

المنوال = 3 × الوسيط – 2 × الوسط

مثال :

إذا علمت أن قيمة الوسط = 5 وقيمة الوسيط = 10 احسب قيمة المنوال .

الحل :

المنوال = 3 × الوسيط – 2 × الوسط

المنوال = 3 × 10 – 2 × 5

المنوال = 30 – 10 = 20

تحديد التواء التوزيع مباشرة من مقاييس النزعة المركزية :

1- المنحنى معتدل التوزيع :

عندما يكون :

الوسط = الوسيط = المنوال

2- المنحنى ملتوى التواء موجب :

عندما يكون :

الوسط > الوسيط > المنوال

3- المنحنى ملتوى التواء سالب :

عندما يكون :

الوسط < الوسيط < المنوال

مـثـال

إذا علمت أن قيمة الوسط = 5 وقيمة الوسيط = 10 احسب قيمة المنوال ، ثم حدد نوع التواء التوزيع .

الحل :

المنوال = 3 × الوسيط – 2 × الوسط

المنوال = 3 × 10 – 2 × 5

المنوال = 30 – 10 = 20

نلاحظ أن

الوسط < الوسيط < المنوال

التوزيع ملتوي التواء سالب .

تمـارينـن

1- احسب الوسط الحسابى والوسيط للدرجات الخام التالية :

5 ــ 3 ــ 2 ــ 8 ــ 17 ــ 4 ــ 10

من قيمة الوسط والوسيط احسب قيمة المنوال ثــم حــدد التــواء التوزيع .

2- أوجد الوسط الحسابى والوسيط فى كل حالة مــن الحــالات التالية ومنها أوجد قيمة المنوال ثم حدد التواء التوزيع .

- **8 - 11 - 9 - 12 - 7**
- **111 - 102 – 103 - 104 - 107 - 105**
- **24 - 20 - 9 - 18 - 35 - 3 - 39 - 3623 - 22**

3- احسب الوسيط والمنوال لكل حالة من الحالات التالية:

- **6 ــ 10 ــ 9 ــ 4 ــ 2 ــ 8 ــ 5**
- **8 ــ 5 ــ 4 ــ 10 ــ 7 ــ 9 ــ 6**
- **20 – 15 – 15 – 12 – 15 – 10 – 12 – 10**
- **70 – 60 – 40 – 20 – 30 – 25 – 20**
- **18 – 15 – 18 – 10 – 12 – 18 – 15 – 13**

4- الجدول التالى يمثل فئات الأجر الأسبوعى لعمال مصنع .

الأجر الأسبوعى	2-	4-	6-	8-	10-12
عدد العمال	10	40	70	50	30

<u>والمطلوب :</u>

- **احسب الوسط الحسابي بطريقة مراكز الفئات**
- **احسب الوسط الحسابي بطريقة الانحرافات**
- **احسب الوسط الحسابي بطريقة الانحرافات المختصرة**
- **احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد**
- **احسب الوسيط باستخدام جدول التكرار المتجمع الهابط**
- **احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد**
- **احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط**
- **احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد والهابط**
- **المنوال بطريقة بيرسون**
- **المنوال بيانياً بطريقة بيرسون**
- **المنوال بطريقة الرافعة كينج**
- **المنوال بيانياً بطريقة الرافعة كينج**

5- من واقع بيانات الجدول التالي :-

ف	20-	30-	40-	50-	60-70
ك	20	40	100	30	10

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون
- المنوال بطريقة كينج .

كلية الاداب / قسم علم الاجتماع
المادة / الاحصاء التطبيقي
مدرس المادة/ أ.د.معاذ احمد حسن

الفصل السادس
مقاييس التشتت

أولاً : المدى .

ثانياً : التباين والانحراف المعيارى .

ثالثاً : الانحراف المتوسط .

رابعاً : الالتواء وتحديد اعتدالية التوزيع .

مقاييس التشتت

لا تعتبر مقاييس التمركز كافية لوصف مجموعة من البيانات وصفاً كاملاً فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات). فالعينات التالية ذات وسط حسابي واحد (8) ولكنها بلا شك تختلف عن بعضها.

عينة 1	8	8	8	8	8
عينة 2	4	3	6	16	11

فالوسط الحسابي يمثل مركز البيانات لكنه لا يبين مدى التفاف أو بعثرة البيانات حول هذا الوسط ، ولهذا لا بد من وجود مقياس آخر مع المقاييس المركزية لقياس درجة التجانس أو التشتت في داخل هذه البيانات.

إن الدرجة التي تتجه بها البيانات الرقمية للانتشار حول قيمة وسطى تسمى تشتت أو توزيع البيانات .

ومن أهم مقاييس التشتت المدى والتباين والانحراف المعيارى والانحراف المتوسط .

أولاً : المدى

المدى هو الفرق بين أكبر قيمة وأصغر قيمة .

حساب المدى من البيانات الغير مبوبة

المدى = أكبر قيمة – أصغر قيمة

مثال :

احسب المدى للبيانات التالية :

95 – 200 – 250 – 300 – 110 – 90 – 150 – 80 – 350 – 100

الحل :

نرتب القيم أولاً : (80–90–95–100–110–150–200–250–300–350)

المدى = 350 – 80 = 270

حساب المدى من البيانات المبوبة

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

مثال:

احسب المدى للجدول التالى :

الفئات	16–	20–	24–	28–	32–36
عدد المبحوثين	10	15	40	20	15

الحل :

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

المدى = 36 – 16 = 20

ثانياً : التباين والانحراف المعيارى

يرمز للتباين بالرمز ع^2

بينما يرمز للانحراف المعيارى بالرمز ع

أي أنه إذا تم حساب أحدهما فيمكن حساب الآخر لأن الانحراف المعيارى هو جذر التباين .

التباين من البيانات الغير مبوبة

هناك طريقتان لحساب التباين من البيانات الغير مبوبة:

الأولى : باستخدام القانون العام من الدرجات الخام كالتالى

$$ع^2 = \frac{مجـ س^2}{ن} - \left[\frac{مجـ س}{ن}\right]^2$$

مثال :

احسب التباين والانحراف المعيارى للقيم التالية ومنه احسب الانحراف المعيارى لكل من المتغيرين س ، ص على حده .

س	23	21	19	19	18
ص	19	19	18	14	15

الحل :

نكون الجدول التالى :

س	$س^2$	ص	$ص^2$
23	529	19	361

361	19	441	21
324	18	361	19
196	14	361	19
225	15	314	18
1467	85	2016	100

ثم نعوض فى القانون العام لحساب التباين :

<u>بالنسبة للمتغير (س)</u>

$$ع^2_{س} = \frac{مجـ\ س^2}{ن} - \left[\frac{مجـ\ س}{ن}\right]^2$$

$$ع^2_{س} = \frac{2016}{5} - \left[\frac{100}{5}\right]^2 = 3.2$$

وبالتالي فان قيمة تباين المتغير س = $ع^2$= 3.2

ومنها فان قيمة الانحراف المعيارى = جذر التباين

ع = $\sqrt{3.2}$ = 1.78

<u>بالنسبة للمتغير (ص)</u>

$$ع^2_{ص} = \frac{مجـ\ ص^2}{ن} - \left[\frac{مجـ\ ص}{ن}\right]^2$$

$$ع^2_{ص} = \frac{1467}{5} - \left[\frac{85}{5}\right]^2 = 4.4$$

وبالتالي فان قيمة تباين المتغير ص = $ع^2$= 4.4

ومنها فان قيمة الانحراف المعيارى = جذر التباين

$$ع = \sqrt{4.4} = 2.1$$

الثانية : باستخدام الطريقة المختصرة "طريقة الانحرافات"

$$ع^2 = \frac{مجـ ح^2}{ن}$$

حيث ح هو الانحراف = س – $م_{س}$

مثال :

احسب الانحراف المعيارى للقيم التالية :

20	19	13	48	19	32	22	17	35

الحل :

نكون الجدول التالى :

س	حس	$ح^2_{س}$
35	10	100
17	8-	64
22	3-	9
32	7	49
19	6-	36
48	23	529
13	12-	144
19	6-	36
20	5-	25
225	-	992

حساب المتوسط :

$$م_{س} = \frac{مجـ\ س}{ن_1} = \frac{225}{9} = 25$$

بعد حساب $م_{س}$ نحسب عمود ح ومنه نحسب $ح^2$ ثم نعـوض فـى القانون :

$$ع^2 = \frac{مجـ\ ح^2}{ن}$$

$$ع^2 = \frac{992}{9} = 110.22$$

ومنه نحسب ع = $\sqrt{110.22}$ = 10.5 .

التباين والانحراف المعيارى من البيانات المبوبة :

يحسب التباين من البيانات المبوبة من العلاقة التالية :

$$ع^2 = ل^2 \times \left\{ \frac{مجـ (ح^2 \times ك)}{مجـ ك} - \left\{ \frac{مجـ (ح \times ك)}{مجـ ك} \right\}^2 \right\}$$

حيث :

$ع^2$ = التباين

ل = طول الفئة

ح = الانحراف ويحسب عن طريق وضع صفر فى عمـوده أمـام الفئة التى يقابلها أكبر تكرار ثم نضع مـن أسـفل (1 ، 2 ، 3 ، وهكذا) ومن أعلى نضع (−1 ، −2 ، −3 ، وهكذا)

مثال :

احسب الانحراف المعيارى من الجدول التالى :

فئات الدخل	10−	20−	30−	40−	50−	60−	80−70
عدد العمال	5	12	22	38	22	12	5

الحل :

نكون الجدول التالى :

الانحراف المتوسط من البيانات الغير مبوبة (المفردة)

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} \mid \text{س} - \text{س}/ \mid}{\text{ن}}$$

حيث :

س = القيمة

س/ = متوسط القيم

ن = عدد القيم

مثال :

لمجموعة البيانات التالية احسب الانحراف المتوسط:-

2 – 3 – 5 – 6 – 7 – 8 – 8 – 9

الحل :

نحسب س/ = (2+3+5+6+7+8+8+9) / 8 = 48/8 = 6

نكون الجدول التالى :

س	| س – س/ |
2	4
3	3
5	1
6	0
7	1
8	2
8	2
9	3

مج	16

$$\text{الانحراف المتوسط} = \frac{16}{8} = 2$$

الانحراف المتوسط من البيانات المبوبة

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} (| \text{س} - \text{س}_{/} | \times \text{ك})}{\text{مجـ ك}}$$

مثال :

من بيانات الجدول التالى احسب الانحراف المتوسط :-

الفئات	-16	-20	-24	-28	36-32
عدد المبحوثين	10	15	40	20	15

الحل :

نكون الجدول التالى :

ف	ك	س	ح/	ح/ × ك	| س - س/ |	| س - س/ | × ك
16-	10	18	2-	20-	8.6	86
20-	15	22	1-	15-	4.6	69
24-	40	26	0	0	0.6	24
28-	20	30	1	20	3.4	68
36-32	15	34	2	30	7.4	111

مج	100	مج		15	مج	358

$$س_{/} = 26 + \frac{15}{100} \times 4 = 26 + 0.6 = 26.6$$

$$\text{الانحراف المتوسط} = \frac{358}{100} = 3.58$$

الالتواء وتحديد اعتدالية التوزيع

$$\text{الالتواء} = \frac{3\ (م - و)}{ع}$$

حيث :

م : المتوسط

و : الوسيط

ع : الانحراف المعيارى .

مثال :

حدد نوع التوزيع التالى :

10 – 40 – 60 – 50 – 20

الحل :

حساب المتوسط :

مجـ س 10+40+60+50+20

$$م = \frac{\text{مج س}}{\text{ن}} = \frac{180}{5} = 36$$

حساب الوسيط :

نرتب القيم تصاعدياً :

10	20	40	50	60

الوسيط = 40

حساب الانحراف المعيارى :

$$ع = \sqrt{\frac{\text{مج ح}^2}{\text{ن}}}$$

نكون الجدول التالى :

س	ح	$ح^2$
20	16-	256
50	14	196
60	24	576
40	4	16
10	26-	676
مج	0	1720

$$ع = \sqrt{\frac{مجـ ح^2}{ن}} = \sqrt{\frac{1720}{5}} = 18.54$$

$$\text{الالتواء} = \frac{3\ (م - و)}{ع} = \frac{3\ (36 - 40)}{18.54} = -0.64$$

الالتواء قيمته سالبة فيكون التواء التوزيع سالب .

تمـارين

1- فيما يلى مجموعة بيانات هى :

95 – 200 – 250 – 300 – 110 – 90 – 150 – 100 – 350 – 80

المطلوب حساب :

- **المدى**
- **التباين**
- **الانحراف المعيارى**
- **المتوسط**
- **الوسيط**
- **المنوال**
- **الانحراف المتوسط**
- **حدد نوع الالتواء**

2- لمجموعة البيانات التالية احسب الانحراف المتوسط :-

2 – 3 – 5 – 6 – 7 – 8 – 8 – 9

كلية الاداب / قسم علم الاجتماع
المادة / الاحصاء التطبيقي
مدرس المادة/ أ.د.معاذ احمد حسن

الفصل السابع
تحليل التباين

مقدمه :

أولاً : طريقة حساب نسبة ف

ثانياً : تحديد مدى دلالة نسبة ف من عدمه .

مقدمه :

دلت الأبحاث الإحصائية التى قام بها فيشر على أهمية تحليل التباين فى الميادين المختلفة لعلوم الحياة وخاصة فى الكشف عن مدى تجانس العينات ومدى انتسابها إلى أصل واحد أو أصول متعددة .

وبالطبع هناك تساؤل لماذا نستخدم تحليل التباين "النسبة الفائية" للحكم على دلالة الإحصائية للعلاقة بين متغيرين وقد استخدمنا من قبل اختبار"ت" لنفس الغرض .

الإجابة بمنتهى السهولة هو أن اختبار "ت" يستخدم لدراسة العلاقة بين متغيرين فقط لا غير أما إذا زاد عدد المتغيرات عن اثنين فلا يمكن استخدام اختبار "ت" بل نستخدم "نسبة ف" .

وبالتالي فان "نسبة ف" تصلح فى حالة متغيرين أو أكثر .

ويعتمد تحليل التباين فى صورته النهائية على قياس مدى اقتراب التباين الداخلي من التباين الخارجي أو مدى ابتعاده عنه وتقاس هذه الناحية بالنسبة التباينية أو النسبة الفائية من خلال العلاقة :

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

حيث أن التباين الكبير هو الأكبر فى القيمة والتباين الصغير هو الأصغر فى القيمة .

طريقة حساب نسبة ف

- **حساب التباين بين المجموعات**

$$\text{التباين بين المجموعات} = \frac{\text{مجموع المربعات بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}}$$

مجموع المربعات بين المجموعات = $ن_1 ق^2_1 + ن_2 ق^2_2 + ن_3 ق^2_3 +$ 00000000000

حيث :

$ن_1$ ، $ن_2$ ، $ن_3$ ، .... هى عدد أفراد المجموعات

$ق^2_1$ ، $ق^2_2$ ، $ق^2_3$ ، .... هى مربع انحراف متوسط كل مجموعة عن المتوسط الكلى للمجموعات ويحسب من العلاقة :

$$ق^2_1 = (م - م_1)^2$$

حيث "م" هو المتوسط الوزنى أو الكلى لكافة المجموعات .

درجة حرية التباين بين المجموعات = عدد المجموعات – 1

- **حساب التباين داخل المجموعات**

$$\text{التباين داخل المجموعات} = \frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}}$$

مجموع المربعات داخل المجموعات = $ن_1 ع^2_1 + ن_2 ع^2_2 + ن_3 ع^2_3$ +

00000000000

حيث :

$ن_1$ ، $ن_2$ ، $ن_3$ ، .... هى عدد أفراد المجموعات

$ع^2_1$، $ع^2_2$ ، $ع^2_3$ ، .... هو التباين لكل مجموعة ويحسب من العلاقة :

$$ع^2_1 = \frac{مجـ س^2}{ن} - \left[\frac{مجـ س}{ن}\right]^2$$

درجة حرية التباين داخل المجموعات = مجموع أفراد جميع المجموعات – عدد المجموعات

تحديد مدى دلالة "نسبة ف" من عدمه

على أى حال نحصل من قانون "نسبة ف" على "ف" المحسوبة نقوم بمقارنتها بـ "ف" الجدوليه ونتبع الآتى :

إذا كانت قيمة "ف" المحسوبة > قيمة "ف" الجدولية فان "نسبة ف" تكون دالة إحصائية .

أما إذا كانت قيمة "ف" المحسوبة < قيمة "ف" الجدولية فان "نسبة ف" ليست دالة إحصائية .

مثال :

درجات الذكور	23	21	19	19	18
درجات الإناث	19	19	18	14	15

الجدول السابق يوضح درجات خمس ذكور وخمس إناث فى اختبار ما والمطلوب حساب النسبة الفائية وبيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 وكذلك عند مستوى دلالة 0.01 ؟

الحل :

نفترض أن درجات الذكور هى "س" ودرجات الإناث هى "ص" ثم نكون الجدول التالى :

س	$س^2$	ص	$ص^2$
23	529	19	361
21	441	19	361
19	361	18	324
19	361	14	196
18	314	15	225
100	2016	85	1467

حساب المتوسطات :

$$م_{س} = \frac{مجـ\ س}{ن_1} = \frac{100}{5} = 20$$

$$م_{ص} = \frac{مجـ ص}{ن_2} = \frac{85}{5} = 17$$

حساب المتوسط الكلى :

$$م = \frac{م_{س} + م_{ص}}{2} = \frac{17 + 20}{2} = 18.5$$

حساب مربع انحراف كل متوسط عن المتوسط الكلى :

$$ق^2_{س} = (م - م_{س})^2 = (20 - 18.5)^2 = (-1.5)^2 = 2.25$$

$$ق^2_{ص} = (م - م_{ص})^2 = (17 - 18.5)^2 = (1.5)^2 = 2.25$$

حساب التباين :

$$ع^2_{س} = \frac{مجـ س^2}{ن} - \left[\frac{مجـ س}{ن}\right]^2$$

$$ع^2_{س} = \frac{2016}{5} - \left[\frac{100}{5}\right]^2 = 3.2$$

$$\frac{مجـ ص^2}{} - \left[\frac{مجـ ص}{}\right]^2$$

$$ع^2_{ص} = \frac{\text{مج ص}^2}{ن} - \frac{\text{مج ص}}{ن}$$

$$ع^2_{ص} = \frac{1467}{5} - \left[\frac{85}{5}\right]^2 = 4.4$$

حساب مجموع المربعات بين المجموعات :

مجموع المربعات بين المجموعات = $ن_1 ق^2_{س} + ن_2 ق^2_{ص}$

مجموع المربعات بين المجموعات = $5 \times 2.25 + 5 \times 2.25$

مجموع المربعات بين المجموعات = 22.5

حساب مجموع المربعات داخل المجموعات :

مجموع المربعات داخل المجموعات = $ن_1 ع^2_{س} + ن_2 ع^2_{ص}$

مجموع المربعات داخل المجموعات = $5 \times 3.2 + 5 \times 4.4$

مجموع المربعات داخل المجموعات = 38

حساب درجات الحرية :

درجة حرية التباين بين المجموعات = عدد المجموعات – 1

درجة حرية التباين بين المجموعات = 2 – 1 = 1

درجة حرية التباين داخل المجموعات = عدد أفراد جميع المجموعات – عدد المجموعات

درجة حرية التباين داخل المجموعات = 5 + 5 – 2 = 8

حساب التباين بين المجموعات :

$$\text{التباين بين المجموعات} = \frac{\text{مجموع المربعات بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}}$$

$$\text{التباين بين المجموعات} = \frac{22.5}{1} = 22.5 \quad \text{(الأكبر)}$$

حساب التباين داخل المجموعات :

$$\text{التباين داخل المجموعات} = \frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}}$$

$$\text{التباين داخل المجموعات} = \frac{38}{8} = 4.75 \quad \text{(الأصغر)}$$

حساب نسبة ف :

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

$$\text{نسبة ف} = \frac{22.5}{4.75} = 4.7368$$

حساب "ف" الجدولية :

لحساب "ف" الجدولية نستخدم درجة حرية التباين الكبير = 1 ودرجة حرية التباين الصغير = 8 ونبحث في جداول النسبة الفائية بدرجتى الحرية السابقتين فنحصل على القيمتين :

"ف" الجد ولية = 5.32 عند مستوى دلالة 0.05

"ف" الجد ولية = 11.26 عند مستوى دلالة 0.01

تحديد مدى دلالة "نسبة ف"

- **"نسبة ف" المحسوبة = 4.7 < "ف" الجدولية عند مستوى دلالة 0.05 = 5.32 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.05 .**
- **"نسبة ف" المحسوبة = 4.7 < "ف" الجدولية عند مستوى دلالة 0.01 = 11.26 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.01 .**

التعليق :

يمكن القول بأن جميع الفروق التي حصل عليها الباحث ليس لها دلالة إحصائية ولا توجد فروق معنوية بين المجموعتين وهذه الفروق ليست إلا مجرد صدفة .

مثال :

حساب "ف" الجدولية :

لحساب "ف" الجدولية نستخدم درجة حرية التباين الكبير = 12 ودرجة حرية التباين الصغير = 2 ونبحث فى جداول النسبة الفائية بدرجتى الحرية السابقتين فنحصل على القيمتين :

"ف" الجد ولية = 19.41 عند مستوى دلالة 0.05

"ف" الجد ولية = 99.42 عند مستوى دلالة 0.01

تحديد مدى دلالة "نسبة ف"

- **"نسبة ف" المحسوبة = 1.13 < "ف" الجدولية عند مستوى دلالة 0.05 = 19.41 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.05**
- **"نسبة ف" المحسوبة = 1.13 < "ف" الجدولية عند مستوى دلالة 0.01 = 99.42 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.01**

تمـارين

1- الجدول التالى يوضح درجات مجموعتين مـن الطـلاب فـى اختبار فى مادة الإحصاء الاجتماعي :

س	19	20	14	17	25	مجـ = 95
ص	14	11	12	13	20	مجـ = 70

والمطلوب حساب نسبة "ف" مع بيان عما إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

2- الجدول التالى يوضح درجات مجموعتين مـن الطـلاب فـى اختبار فى مادة الحاسب الآلى :

س	18	19	13	16	24	مجـ = 90
ص	15	12	13	14	21	مجـ = 75

والمطلوب حساب نسبة "ف" مع بيان عما إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

3- الجدول التالى يوضح درجات مجموعتين مـــن الطـــلاب فـــى اختبار فى مادة اللغة الفرنسية :

س	7	15	15	11	12
ص	7	2	8	7	6

احسب الدلالة للفروق القائمة بين تلك الدرجات بطريقـــة تحليـــل التباين عند مستوى دلالة 0.05

4- الجدول التالى يوضح درجات 4 مجموعات من الطـــلاب فـــى اختبار فى مادة اللغة العربية :

س	49	59	61	60	61
ص	68	55	60	67	60
ع	64	63	54	52	62
هـــ	67	55	65	64	59

احسب الدلالة الإحصائية للفروق القائمة بين تلك الدرجات بطريقة تحليل التباين عند مستوى دلالة 0.05 وبين مدى تجـــانس هـــذه المجموعات بالنسبة لأصل واحد أو لأصول متعددة .

كلية الاداب / قسم علم الاجتماع
المادة / الاحصاء التطبيقي
مدرس المادة/ أ.د.معاذ احمد حسن

الفصل الثامن
اختبار "ت"

مقدمه :

أولاً : شروط استخدام اختبار "ت"

ثانياً : الحالات المختلفة لحساب "ت"

مقدمة :

يعد اختبار "ت" من أكثر اختبارات الدلالة شيوعاً فى الأبحاث النفسية والاجتماعية والتربوية ، وترجع نشأته الأولى إلى أبحاث العالم "ستودنت" ولهذا سمى الاختبار بأكثر الحروف تكراراً فى اسمه وهو حرف التاء .

ومن أهم المجالات التى يستخدم فيها هذا الاختبار الكشف عن الفروق بين تحصيل الذكور والإناث فى مادة دراسية ما وذلك عن طريق حساب دلالة فرق متوسط تحصيل الذكور عن متوسط تحصيل الإناث .

ويمكن القول أن اختبار "ت" يستخدم لقياس دلالة فروق المتوسطات غير المرتبطة والمرتبطة للعينات المتساوية والغير متساوية .

شروط استخدام اختبار "ت" لدلالة فروق المتوسطات

لا يحق للباحث أن يستخدم اختبار "ت" قبل أن يدرس خصائص متغيرات البحث من النواحى التالية :-

1- حجم كل عينة .

2- الفرق بين حجم عينتي البحث .

3- مدى تجانس العينة .

4- مدى اعتدالية التوزيع التكرارى لكل من عينتى البحث .

1- حجم كل عينة

يجب أن يزيد حجم كل من العينتين عن "5" ويفضل أن يزيد عن "30" أما إذا قل حجم أى من العينتين عن "5" فلا يمكن استخدام اختبار "ت" .

2- الفرق بين حجم عينتى البحث : شرط التقارب

يجب أن يكون حجم عينتى البحث متقارباً فلا يكون مثلاً حجم أحد العينتين "500" وحجم الأخرى "30" لأن للحجم أثره على مستوى دلالة "ت" .

3- مدى تجانس العينتين

يقصد بتجانس العينات مدى انتسابها إلى أصل واحد أو أصول متعددة . فإذا انتسبت العينات إلى أصل واحد فهى متجانسة وإذا لم تنتسب العينات إلى أصل واحد فهى غير متجانسة .

وبالطبع يصعب بالنسبة للباحث تحديد أصول العينات لتحديد تجانسها لذا يمكنه استخدام النسبة الفائية لتحديد التجانس .

يحدد تجانس العينتين من خلال حساب قيمة النسبة الفائية حيث تحسب من العلاقة :

$$ف = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}}$$

حيث أن التباين الأكبر هو التباين الأكبر فى القيمة دون التحيز لأحد العينتين ، والتباين الأصغر هو الأصغر فى القيمة دون التحيز لأحد العينتين .

بالطبع نحصل من القانون السابق على قيمة لـ "ف" تسمى بقيمة ف المحسوبة ولتحديد التجانس نحسب قيمة أخرى تسمى ف الجدولية ونحصل عليها من جداول "ف" الإحصائية عند درجة حرية التباين الأكبر ودرجة حرية التباين الأصغر ومستوى الدلالة الذى قيمته إما "0.05" أو "0.01" حيث نحسب درجات الحرية من القانون التالى :

درجة حرية التباين الأصغر = ن – 1

حيث "ن" هى عدد أفراد العينة التى تبيانها هو الأكبر .

درجة حرية التباين الأصغر = ن – 1

حيث "ن" هى عدد أفراد العينة التى تبيانها هو الأصغر .

تحديد التجانس

- إذا كانت قيمة "ف" المحسوبة > قيمة "ف" الجدولية فلا يوجد هناك تجانس .
- أما إذا كانت قيمة "ف" المحسوبة < قيمة "ف" الجدولية فيوجد هناك تجانس .

4- مدى اعتدالية التوزيع التكرارى لكل من العينتين

يكون التوزيع التكرارى معتدلاً عندما تكون قيمة الالتواء الخاص به محصورة بين القيمتين [–3 ، +3] أى واقعة فى الفترة المغلقة –3 و +3 .

ويحسب الالتواء من القانون التالى :-