

البرنامج التدريبي بعنوان

الأساليب الإحصائية ودورها في خدمة المجتمع

أهداف البرنامج التدريبي :

أهداف عامة تتمثل في معرفة الأساليب الإحصائية لجمع المتغيرات وإستخدام مقاييس لمعرفة مدي دلالة النتائج التي تم التوصل إليها وإتخاذ قرارات مناسبة تفيد المجتمع.

أهداف فرعية تتمثل في القدرة علي مناقشة علم الإحصاء من حيث مفهومه وأساليب جمع المتغيرات وتليخصها واتخاذ القرارات المناسبة. التميز بين أنواع العينات في المجتمع. التعرف علي الفرق بين مقاييس النزعة المركزية ومقاييس التشتت. ومن ثم التعبير عن الحقائق التي تم التوصل إليها في صورة رقمية دقيقة واضحة للباحث أو الدارس والمقارنة بين المجموعات المختلفة من المتغيرات والمعلومات ثم إيجاد علاقة بينها. التنبؤ لتسهيل عملية التخطيط والتنظيم والتطوير والتنمية على مستوى المجتمع وإستخلاص النتائج وإتخاذ القرارات الصحيحة التي تناسب المجتمع.

المحتوي

- الفصل الأول: المبادئ الأساسية لعلم الإحصاء وأساليب جمع البيانات
- الفصل الثاني: مقاييس النزعة المركزية
- الفصل الثالث: مقاييس التشتت
- الفصل الرابع: الارتباط والانحدار

الفصل الأول: المبادئ الأساسية لعلم الإحصاء وأساليب جمع البيانات

مقدمة

يعتبر علم الإحصاء في الوقت الحالي واحد من أهم العلوم الحديثة التي تلعب دوراً حيوياً في كثير من العلوم والدراسات المختلفة. حيث يعد من أقدم العلوم الذي ظهر مع حاجة الإنسان الأولى للتعامل مع القيم والأعداد لتسيير الحياة اليومية. ومع التطور الهائل في العلوم كافة في أواخر القرن العشرين تطور علم الإحصاء ليستفيد من تقنيات الحاسب الآلي بشكل يجعله العلم الأكثر تداخلاً مع العلوم الأخرى المختلفة، حيث أصبح يستخدم علم الإحصاء في العلوم التجارية وعلوم الطب والهندسة والأدب وجميع العلوم الأخرى دون استثناء. حيث يتم الاستفادة من علم الإحصاء في مجالات متنوعة تشمل ميادين عديدة كالصناعة والزراعة والطب والبحوث وغيرها من مجالات الإدارة والأعمال والعلم بشكل عام. ويتم تطبيق الأساليب الإحصائية في الجوانب المختلفة للمجتمع كالصناعة كمرقبة جودة المنتجات وتسويقها والتخزين وتشغيل خطوط الإنتاج. كما يتم استخدام علم الإحصاء في المجال الطبي لدراسة الأمراض المختلفة والبحث في مسبباتها وطرق علاجها وغيرها من الجوانب المختلفة.

مفهوم علم الإحصاء:

هو ذلك العلم الذي يعمل على استخدام الأسلوب العلمي في طرق جمع المتغيرات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول إلى إستنتاجات وقرارات مناسبة.

أقسام علم الإحصاء:

□ الإحصاء الوصفي : الطرق والأساليب المستخدمة في جمع وعرض وتنظيم المتغيرات في صورة مؤشرات رقمية مع امكانية عرضها في جداول ورسومات بيانية ، وحساب بعض المؤشرات الاحصائية منها.

□ الإحصاء الإستدلالي: وهو الطرق والأساليب المستخدمة في تعميم نتائج العينة على المجتمع واتخاذ القرارات المناسبة.

بعض المفاهيم الأساسية في الإحصاء

المجتمع: هو جميع القيم أو المفردات التي يمكن أن يأخذها المتغير.

مثال درجات الطلبة في كلية التجارة حيث تمثل درجات جميع الطلبة في تلك الكلية المجتمع محل الدراسة .

والمجتمع إما ان يكون:

□ مجتمع محدد أو منتهي أي يمكن حصر عدد أفراده مثل :عدد أفراد عائلة – عدد الطلبة الناجحين ... وغيرها

□ مجتمع غير محدد أو غير منتهي وهو المجتمع الذي من الصعب أو المستحيل حصر عدد أفرادهِ ، مثل : عدد الأسماك في النهر ، عدد الطيور ، عدد الميكروبات في الجو الخ.

العينة:

هي مجموعة جزئية من المجتمع اختيرت بطريقة ما من المجتمع .

مثال: عدد الطلاب الذين حصلوا علي جوائز تشجيعية من كلية التجارة.

ويجب أن تكون العينة ممثلة للمجتمع أي أنها تكون ممثلة لصفات المجتمع الأصلي ويمكن استنتاج خواصه منها.

أسباب اختيار العينة عن المجتمع:-

□ هو صعوبة دراسة المجتمع ككل إذ إنه قد يتطلب وقت وجهد وتكلفة كبيرين.

□ قد حيث دمار للموارد البشرية في حالة استخدام المجتمع ككل .

المتغير:

هو الصفة التي تكون محور الدراسة في المجتمع الإحصائي . بعض المتغيرات محل الدراسة لطلاب كلية التجارة لون الشعر للطلاب - مستوى الذكاء للطلاب - طول القامة للطلاب- النوع للطلاب (ذكر -انثى).

وتنقسم المتغيرات إلي نوعين:

هما متغيرات وصفية و متغيرات كمية. تنقسم المتغيرات الوصفية بدورها إلى قسمين رئيسيين هما متغيرات وصفية اسمية و متغيرات وصفية ترتيبية، حيث يتركز الفرق بين النوعين في كون المتغيرات الوصفية الاسمية تشير في الواقع إلى حقول مختلفة لا تمثل ترتيباً محدداً، بينما يكون للترتيب معناً محدداً في المتغيرات الوصفية الترتيبية. فالحالة الاجتماعية يمكن تصنيفها إلى أعزب و متزوج و مطلق و أرمل دون أن يكون هنالك حاجة إلى ترتيبها بشكل محدد، مما يشير إلى أن بيان الحالة الاجتماعية يمثل هنا بياناً وصفاً اسماً. في الجهة الأخرى نجد أن المتغيرات الوصفية النوعية تمثل حقول تشير إلى ترتيب محدد لا يمكن تجاهله. فالمستوى التعليمي يعكس ترتيباً محدداً حيث يتم ذكر مثل حقل " غير متعلم" يليه " اقل من ثانوي" ثم "اقل من جامعي" يليه "جامعي فأعلى."

يمثل النوع الآخر للمتغيرات الإحصائية متغيرات تأخذ أرقاماً و يطلق عليها متغيرات كمية. تمثل المتغيرات الكمية عصب الحياة للعملية الإحصائية. فمعظم العمليات الإحصائية تتم من خلال التعامل مع أرقاماً أكثر من التعامل مع كلمات أو صفات. وحتى في عملية التعامل مع الصفات فإن الهدف الرئيسي يكون غالباً في دراسة التكرارات أو عدد مرات ظهور الصفة المحددة والتي يمكن تسجيلها كمياً.

تختلف المتغيرات الكمية تبعاً لاختلاف القراءات للبيان، فالمتغيرات التي تأخذ أرقاماً محددة تمثل متغيرات كمية متقطعة بينما المتغيرات الكمية التي تأخذ قيماً محصورة في فترة محددة تمثل متغيرات كمية متصلة. فمثلاً، يمثل عدد الأطفال في البيت بياناً كميّاً متقطعاً بحكم أن عدد الأطفال لا بد وان يكون عدداً صحيحاً. كذلك يمثل عدد الحوادث المرورية بياناً كميّاً متقطعاً يأخذ قيماً بصيغ أعداد صحيحة لا تقبل أن تكون كسور، فعندما يقال أن عدد الحوادث أربعة فإنه يقصد فعلاً أنها أربعة بالتحديد دون تقريب حسابي. في الجهة الأخرى يمثل وزن الطفل أو التكاليف المادية للحوادث بياناً كميّاً متصلاً، غالباً ما يتم فيه التقريب لدرجة معينة للحصول على قيمة يمكن التعامل معها.

أساليب جمع البيانات:

الحصص الشامل: عملية جمع البيانات جميع المفردات من المجتمع محل الدراسة أسلوب العينة : عملية اختيار جزء من مفردات المجتمع(عينة) بطريقة مدروسة تجعله يمثل المجتمع بغرض تعميم النتائج علي المجتمع كله.

أنواع العينات العشوائية

العينة العشوائية البسيطة:

إذا كان المجتمع يتضمن عدد ن من المفردات المتجانسة وتم اختيار عينة منها وكان لكل مفردة لها نفس الفرصة في الظهور أو الاختيار فإن العينة تسمى بالعينة العشوائية البسيطة

العينة العشوائية الطبقيّة :

يستخدم هذا النوع من العينات في المجتمعات الغير متجانسة والتي تتباين مفرداتها وفقاً لخواص معينة، مثل المستوى التعليمي لمفردات مجتمع الدراسة، الجنس، نوع التخصص. ويمكن تقسيم مجتمع الدراسة إلى طبقات وفقاً لهذه الخواص وعادة تتجانس مفردات الطبقة الواحدة فيما بينها وتختلف الطبقات عن بعضها البعض. ويعتبر هذا النوع من العينات الأنسب للمجتمعات المتباينة حيث تكون العينة ممثلة لكافة فئات مجتمع الدراسة. ويتم اختيار العينة العشوائية الطبقيّة عبر الخطوات التالية :

1- تقسيم المجتمع إلى فئات أو مجموعات متجانسة وفقاً لخاصية معينة .

2- تحديد عدد مفردات العينة الكلية

3- تحديد نسبة كل طبقة في العينة المختارة إلى إجمالي حجم المجتمع الأصلي

4- تحديد عدد الأفراد لكل طبقة في العينة المختارة.

العينة العشوائية المنتظمة :

يستخدم هذا النوع من العينات عند دراسة المجتمعات المتجانسة والتي لا تتباين مفرداتها كثيراً وسميت بالعينة المنتظمة لانتظام المسافات بين المفردات المختارة من مجتمع الدراسة. ويتم عادة اختيار العينة المنتظمة من خلال حصر مفردات مجتمع الدراسة الأصلي ثم يعطى كل فرد رقماً متسلسل. بعدها يتم قسمة عدد مفردات مجتمع البحث على حجم العينة المطلوبة فينتج الرقم الذي

سيفصل بين كل مفردة يتم اختيارها في عينة الدراسة والمفردة التي تليها. وعادة يتم اختيار المفردة الأولى عشوائياً.

الفصل الثاني : مقاييس النزعة المركزية مقدمة :

تعتبر المقاييس الكمية على اختلاف أنواعها مؤشرات مهمة جدا حيث تلعب دورا حيويا في معظم مراحل العملية الإحصائية الاستدلالية. وتتنوع تلك المقاييس مع توحيد الهدف من وجودها، لتعطي الباحث إمكانية التعامل مع جميع أنواع البيانات الكمية ومن ثم الوصول إلى وصف كمي دقيق لها ومن ثم فإن مقاييس النزعة المركزية (أو المتوسطات) من أهم المقاييس الإحصائية التي يفكر الباحث في حسابها، بل هي أول مقاييس إحصائية يفكر فيها الباحث السياسي عموماً. فمقياس النزعة المركزية لظاهرة سياسية ما تعني التعرف على القيمة التي تقع عادة عند مركز التوزيع العددي للقيم المبحوثة.

إن متوسط أي ظاهرة يعبر عن المستوى العام لهذه الظاهرة. فمتوسط مجموعة من القيم هو القيمة التي تعبر عن جميع القيم، أو هو القيمة التي تدور (أو تتركز) حولها باقي القيم. فمتوسط الدخل لأي بلد يعبر عن المستوى العام للدخل في هذا البلد كما أن متوسط دخل عمال أحد المصانع هو الدخل الذي تتركز حوله دخول العمال بهذا المصنع. ولذا تسمى المتوسطات بمقاييس النزعة المركزية. وذلك لأن قيم أي ظاهرة - عادة - تميل أو تنزع للتركز حول قيمة معينة هي متوسط هذه الظاهرة أو مقياس نزعتها المركزية. فأطوال البالغين تتركز حول رقم معين هو متوسط الطول، وكذلك أوزانهم ومعدلات ذكائهم، وأي ظاهرة أخرى. فالمتوسط - بصفة عامة - هو الذي يعبر عن المستوى العام للظاهرة أي هو الذي يعبر عن جميع قيمها، بمعنى أنه القيمة التي تتركز حولها باقي القيم.

وسوف نتناول في هذا الفصل أهم المتوسطات وهي : الوسط الحسابي والوسيط والمنوال. وفيما يلي عرض لهذه المتوسطات نبين فيه تعريف كل مقياس وكيفية حسابه ومزاياه وعيوبه.

الوسط الحسابي :

يعتبر الوسط الحسابي أكثر المتوسطات شهرة وأكثرها استخداماً، بل لعله من أهم المقاييس الإحصائية على الإطلاق، وذلك لما يتمتع به من مزايا وخواص، ولدخوله في حساب الكثير من المقاييس الإحصائية الأخرى كما سيتضح فيما بعد.

والفكرة الأساسية في حساب الوسط الحسابي لمجموعة من القيم أنه يساوي خارج قسمة مجموع القيم على عددها.

$$\text{الوسط الحسابي لمجموعة قيم} = \frac{\text{مجموع هذه القيم}}{\text{عددها}}$$

(ويعرف الوسط الحسابي لمجموعة من القيم بأنه القيمة التي لو حلت محل جميع القيم لا يتغير مجموعها).

يتضح مما يلي :

أولاً : أنه لحساب الوسط الحسابي يجب أن تكون لدينا بيانات كمية. أي لا يصلح الوسط الحسابي إذا كانت البيانات وصفية أسمية أو ترتيبية، إذ لا معنى له في هذه الحالات.

ثانياً : أن جميع القيم – بلا إستثناء – تدخل في حساب الوسط الحسابي. أي أنه يعبر عن جميع القيم فعلاً. ولذا فإنه إذا كان من بين القيم قيمة شاذة أو متطرفة (بمعنى أنها كبيرة جداً أو صغيرة جداً بالنسبة لباقي القيم) فإنها سوف تؤثر في قيمة الوسط الحسابي. أي أنه يتأثر بالقيم الشاذة أو المتطرفة.

الوسيط :

يعرف الوسيط بأنه القيمة التي تقع في منتصف القيم بعد ترتيبها (تصاعدياً أو تنازلياً). فالوسيط هو القيمة التي تتوسط القيم بعد ترتيبها. فإذا كان عدد القيم فردياً فإنه توجد قيمة واحدة في المنتصف (بعد الترتيب) تكون هي الوسيط. أما إذا كان عدد القيم زوجياً فإنه توجد قيمتان في المنتصف نجمعهما ونقسم على 2 فنحصل على قيمة الوسيط. وبديهي أننا سنحصل على النتيجة نفسها لو كان الترتيب تصاعدياً أو تنازلياً.

بعض خصائص الوسيط :

1 – لا يتأثر الوسيط بالقيم الشاذة أو المتطرفة. وهذا منطقي لأنه يقع في منتصف القيم، والقيم الشاذة إما أن تكون في أول القيم أو آخرها (بعد ترتيب القيم تصاعدياً أو تنازلياً).

المنوال :

المنوال وهو ثالث المتوسطات ويعرف بأنه القيمة الأكثر تكراراً أو شيوعاً بين القيم، فهو القيمة التي تتكرر أكثر من غيرها. وأحياناً يسمى المنوال "القيمة الشائعة" أي الأكثر شيوعاً بين القيم. والمنوال من أكثر المتوسطات استخداماً في الحياة التجارية. حيث تعتمد – على سبيل المثال – مصانع الملابس الجاهزة على المقاييس الشائعة بين الناس لتحديد المقاييس المختلفة لهذه الملابس.

ويتميز المنوال بالسهولة والبساطة سواء في فكرته أو في إيجاد قيمته. حيث أنه لا يعتمد علي ترتيب البيانات ولا تجمع ولا أي شيء من هذا القبيل. فقط نبحث عن القيمة التي تتكرر أكثر من غيرها لتكون منوال القيم.

بعض الملاحظات على المنوال :

- 1 – أنه يمكن إيجاد المنوال لكل انواع البيانات (كمية أو ترتيبية أو اسمية).
- 2 – حسب تعريف المنوال قد لا تتكرر قيمة أكثر من غيرها، وبالتالي قد لا يوجد منوال لبعض البيانات.

الفصل الثالث: مقاييس التشتت مقدمة

تمثل مقاييس التشتت الجانب الآخر من المقاييس الإحصائية الأساسية بجانب مقاييس النزعة المركزية، حيث تستخدم تلك المقاييس في وصف البيانات والتعرف على خصائصها. كما تعمل مقاييس التشتت كجزئية مكملة ومهمة جدا بجانب مقاييس النزعة المركزية في عمليات الاستدلال الإحصائي المبنية على عملية التعامل مع البيانات. وينصب الاهتمام عند التعامل مع مقاييس التشتت حول قياس درجة الاختلاف بين القيم المختلفة للمتغير الكمي المدروس، ويتم ذلك من خلال عدة مقاييس مختلفة يهتم كل واحد منها بقياس درجة الاختلاف من زاوية مختلفة. يمثل التباين والانحراف المعياري والانحراف المتوسط والانحراف الربيعي بالإضافة إلى المدى مقاييس مختلفة لقياس تشتت المتغيرات الكمية.

يتم الحصول على تصور دقيق عن خصائص المتغير الكمي في حال توفر كل من مقياس النزعة المركزية ومقياس التشتت، حيث تعطي مقاييس النزعة المركزية تصور عن تمركز القيم بينما تعطي مقاييس التشتت تصور عن درجة اختلاف تلك القيم عن بعضها البعض. لذا يمكن القول بأن الاعتماد على مقياس واحد قد لا يغني عن الآخر في عملية الاستدلال الإحصائي، حيث ينتج عنه دوما قصور في المعلومة المعتمد عليها ومن ثم عدم القدرة على قراءة البيانات إحصائيا بشكل سليم. وسوف نتناول في هذا الفصل أهم مقاييس التشتت المطلقة وهي: المدى والانحراف المعياري والتباين وأيضا مقاييس التشتت النسبية: معامل الاختلاف.

المدى:

يستخدم المدى لتمثيل الفرق بين أكبر قيمة وأصغر قيمة كمقياس بسيط وسطحي عن درجة تشتت قيم المتغير الكمي. ولكن لا يجب الأخذ بهذا المقياس والاعتماد عليه في العمليات الاستدلالية الإحصائية حيث انه يتأثر بشدة بالقيم المتطرفة بالإضافة إلى عدم استخدامه لباقي قيم المتغير الكمي. وكمقياس أدق يعمل في حال وجود قيم متطرفة، يمكن استخدام الانحراف الربيعي والذي يعتمد على ترتيب القيم كما هو معمول به في عملية حساب الوسيط. في المقابل عندما لا تكون مشكلة القيم المتطرفة حاضرة فانه يمكن استخدام الانحراف المعياري كمقياس للتشتت، حيث يتم استخدام كافة القيم في عملية حساب المقاييس السابقة.

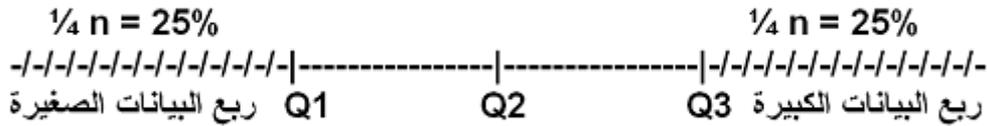
يتم في الواقع حساب مقاييس التشتت للبيانات الكمية بصيغتها الخام، ولكن الطريقة المتبعة في عمليات الحساب تختلف باختلاف طبيعة البيانات المدروسة. وبالطبع، كما تم الإشارة إليه سابقا، تمثل التقديرات المحصلة من البيانات الخام معلومة أدق وأكثر صحة من المعلومة المحصلة من البيانات المبوبة، لذا فانه يجب الاعتماد على البيانات الخام في حال توفرها في عملية حساب مقاييس التشتت، كما يجب قصر الاعتماد على البيانات المبوبة ليتم فقط في حالة عدم توفر الصيغة الخام للبيانات حيث أنها تعطي قيم تقريبية لا ترقى إلى دقة التقديرات المحصلة من خلال استخدام البيانات الغير مبوبة.

عند التعامل مع بيانات خام فانه يمكن أن يتم تغطية مجتمع الدراسة إذا كان المجتمع محدود الحجم وكانت الدراسة موجهة للتعامل مع المجتمع كاملا مع ما يصاحب تلك العملية من جهد ووقت وتكاليف مادية عالية. وإذا كان مجتمع الدراسة غير محدود أو كانت الدراسة لا تستطيع

تحمل الوقت أو الجهد أو التكلفة العالية المطلوبة لتغطية مجتمع الدراسة، فإنه يمكن الاعتماد على عينة عشوائية مسحوبة من المجتمع وممثلة له للحصول على تقديرات لمقاييس التشتت المطلوبة. تجدر الإشارة إلى أن جميع مقاييس التشتت هي قيم موجبة، وذلك شرط أساسي يجب توفره في جميع مقاييس التشتت. لذا فإن مقاييس التشتت لا تأخذ قيم سالبة أبداً بل تكون قيمها موجبة دوماً أو مساوية للصفر فقط وذلك إذا كانت جميع قيم المتغير الكمي محل الدراسة متساوية، أي أنه لا يوجد تباين أو تشتت أصلاً. مزايا المدى أنه سهل الحساب و الفهم . أما عيوبه أنه يأخذ في الاعتبار فقط أصغر وأكبر قيمة للتوزيع ، ومن ثم فإنه يتأثر كثيراً بالقيم المتطرفة ، هذه العيوب للمدى تجعل فائدته محدودة.

نصف المدى الربيعي:

رأينا أن المدى يتأثر كثيراً بالقيم الشاذة أو المتطرفة. ولذلك دعت الحاجة لإيجاد مقاييس أخرى للتشتت لا تتأثر بالقيم المتطرفة. وأحد هذه المقاييس هو نصف المدى الربيعي. وحيث أن القيم المتطرفة هي تلك القيم الصغيرة جداً أو الكبيرة جداً فإنه عند حساب نصف المدى الربيعي لا يؤخذ في الاعتبار ربع البيانات الصغيرة (25%) ولا ربع البيانات الكبيرة (25%). الشكل التالي يوضح موضع القيم الشاذة والمتطرفة.



يرمز لنصف المدى الربيعي بالرمز Q ويعرف بالصيغة التالية:

$$Q = \frac{Q_3 - Q_1}{2}$$

حيث أن Q1 هو الربع الأول و Q3 هو الربع الثالث.

بعض مميزات وعيوب نصف المدى الربيعي:

1. من المميزات أنه لا يتأثر بالقيم الشاذة أو المتطرفة.
2. من العيوب أنه لا يأخذ في الاعتبار جميع البيانات.

الانحراف المعياري:

يعتبر الانحراف المعياري والتباين من أهم مقاييس التشتت الإحصائية. ويرتبط المقياسين بعلاقة رياضية قوية، حيث يمكن دوماً الحصول على المقياس الآخر في حال معرفة قيمة أحدهما. يرمز للتباين بالرمز σ^2 في حال الحصول على قيمته من خلال تغطية مجتمع الدراسة، بينما يتم استخدام الرمز s^2 للدلالة على مقدر التباين المحصل من خلال بيانات عينة عشوائية مسحوبة

من مجتمع الدراسة. وبأخذ الجذر التربيعي للتباين يتم الحصول على قيمة الانحراف المعياري وذلك في الحالتين، حالة المجتمع وحالة العينة،

$$\sigma = \sqrt{\sigma^2} \quad \text{or} \quad S = \sqrt{S^2}$$

وبحكم العلاقة الرياضية القوية بين كل من التباين والانحراف المعياري فإنه يمكن اعتبارهما وجهين لعملة واحدة لهما نفس الأهمية.

يعتمد الانحراف المعياري والتباين على فكرة تربيع الفروق بين قيم المتغير الكمي X ووسطها الحسابي، وفي حال التعامل مع مجتمع بحجم N يتم الحصول على الانحراف المعياري من خلال الدالة التالية،

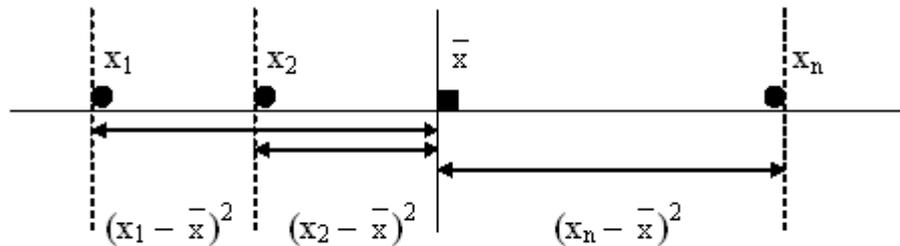
$$\sigma_x = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu_x)^2}{N}}$$

أما في حالة التعامل مع عينة عشوائية بحجم n مسحوبة من مجتمع الدراسة وممثلة له، فإن صيغة حساب الانحراف المعياري هي،

$$S_x = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$$

التباين :

فكرة التباين تعتمد على تشتت أو تباعد البيانات عن متوسطها. فالتباين يكون كبيراً إذا كانت البيانات متباعدة عن متوسطها والعكس بالعكس. ويعرف التباين بأنه متوسط مربع انحرافات القيم عن الوسط الحسابي ويرمز له بالرمز S^2 . الشكل التالي يبين مربعات انحرافات القيم عن الوسط الحسابي.



معامل الاختلاف (التغير):

ذكرنا سابقاً أن التباين والانحراف المعياري من المقاييس المفيدة لقياس التشتت لتوزيع متغير ما. ولكن في كثير من الأحيان نكون مهتمين بمقارنة التشتت والاختلاف لتوزيعين متغيرين مختلفين. وبما أن التباين والانحراف المعياري مقياسان يعتمدان على وحدة البيانات فإنه يصعب استخدامهما لمقارنة تجانس المجموعات المختلفة من البيانات وذلك لاختلاف الوحدة المستخدمة. وبشكل عام فإن مقاييس التشتت التي ذكرناها آنفاً تكون غير مناسبة لمقارنة تجانس مجموعات البيانات المختلفة في الحالتين التاليتين:

(1) إذا كانت وحدتا المتغيرين مختلفتين حيث لا نستطيع مقارنة الوحدات المختلفة ببعض.
(2) إذا كان متوسط المتغيرين مختلفين وذلك لأن تباين توزيع المتغير ذي المتوسط الصغير ينزع لأن يكون صغيراً والعكس بالعكس.

لذلك دعت الحاجة إلى مقياس لا يعتمد على وحدة المتغير ويقاس ما يسمى بالتشتت النسبي. وأحد هذه المقاييس هو ما يسمى بمعامل الاختلاف أو معامل التغير. فمعامل الاختلاف هو أحد مقاييس التشتت النسبي وهو مقياس عديم الوحدة ويستخدم لمقارنة التشتت النسبي أو التجانس لمجموعات البيانات المختلفة. فمجموعة البيانات ذات معامل الاختلاف الأكبر يكون تشتتها النسبي أكبر أي أنها تكون أقل تجانساً والعكس بالعكس. ويعرف معامل الاختلاف للعينة التي متوسطها \bar{x} وانحرافها المعياري s بالصيغة التالية:

$$c. v. = \frac{s}{\bar{x}}$$

الفصل الرابع : الإرتباط والإنحدار الخطي البسيط

مقدمة

عادة ما يهتم الباحثين بدراسة ووصف متغير واحد عن طريق حساب احدي المقاييس الاحصائية كالوسط الحسابي والانحراف المعياري وغيرها ؛ ولكن عادة ما يكون الاهتمام بدراسة عدة متغيرات والعلاقة بينهما . ففي علم التسويق علي سبيل المثال يكون من الضروري دراسة العلاقة بين السعر والمبيعات لمعرفة هل سياسة تخفيض الاسعار تزيد من المبيعات ام لا ؟ ويتم تحديد مدي قوة العلاقة بين متغيرين م خلال معامل الإرتباط .

معامل الارتباط

ان الهدف من دراسة الارتباط هو الكشف عن قوة أو درجة العلاقة بين متغيرين أو أكثر، وتتراوح درجة العلاقة بين أي متغيرين والتي يعبر عنها باصطلاح معامل الارتباط بين + 1 ، -1. فكلما كانت درجة الارتباط قريبة من 1 فإن ذلك يعني ان الارتباط قوياً بين المتغيرين، وكلما قلت درجة الارتباط كلما ضعفت العلاقة بين المتغيرين.

وقد تتخذ العلاقة الارتباطية بين المتغيرين أحد شكلين:-

- علاقة طردية : زيادة قيمة أحد المتغيرات تؤدي إلي زيادة قيمة المتغير الاخر وكذلك نقصان قيمة أحد المتغيرين تؤدي الى نقصان قيمة المتغير الآخر كالعلاقة بين المصروف علي الإعلان والمبيعات.
- علاقة عكسية :زيادة قيمة أحد المتغيرين تؤدي الي نقصان قيمة المتغير الاخر مثل العلاقة بين معدل دوران العمل والإنتاجية .يمكن أن تكون العلاقة بالعكس فنقصان قيمة أحد المتغيرين قد يؤدي إلي زيادة قيمة المتغير الاخر.

بشكل عام فإنه يمكن اعتبار ان العلاقة ضعيفة اذا كانت قيمة معامل الارتباط أقل من 0.30 , ويمكن اعتبارها متوسطة اذا تراوحت قيمة معامل الارتباط بين 0.30 الى 0.70 أما اذا كانت قيمة معامل الارتباط أكثر من 0.70 فتعتبر العلاقة قوية بين المتغيرين.ومن الجدير بالذكر ان الارتباط يدل على وجود علاقة ما بين متغير وآخر

أ) معامل ارتباط بيرسون Pearson

يستخدم معامل ارتباط بيرسون لقياس قوة العلاقة بين قيم متغيرين كالعلاقة بين مصروف الاعلان وحجم المبيعات أو العلاقة بين التدريب وانتاجية العاملين. ويمكن استخراج معامل الارتباط من خلال تطبيق المعادلة التالية:

$$R = \frac{N(\sum x \times y) - \sum x \times \sum y}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

الانحدار الخطي البسيط :

الهدف الاساسي من تحليل الانحدار هو تقدير الصورة الرياضية للعلاقة بين متغير مستقل ومتغير تابع. ويستخدم تحليل الانحدار لدراسة مدى تأثير متغير مستقل واحد أو أكثر على متغير تابع محدد بحيث نستطيع التنبؤ بقيم المتغير التابع اذا علمنا قيم المتغير المستقل أو المتغيرات المستقلة. ويجب ان تتوفر شروط أساسية لاجراء تحليل الانحدار حتى تكون النتائج دقيقة ويمكن الوثوق بها, حيث ينبغي ان يكون توزيع المتغيرين المستقل والتابع توزيعاً طبيعياً, كما ينبغي ان تكون العينة مختارة بشكل عشوائي.

وهناك نوعين من الانحدار الخطي:

- الانحدار الخطي البسيط : يبحث في تأثير متغير مستقل واحد في متغير تابع واحد.
- الانحدار الخطي المتعد : يبحث في تأثير أكثر من متغير مستقل في متغير تابع واحد

معادلة الانحدار الخطي البسيط:

يعد الانحدار الخطي البسيط من أكثر الموضوعات استخداماً في العمليات الإحصائية . ويقول سمير كامل عاشور وسامية أبو الفتوح سالم (2005 ، ص 16) بأن عملية الانحدار الخطي في أبسط صورها تبدأ بوجود متغير واحد مستقل ومتغير آخر تابع ، فإذا توفرت بيانات للمتغيرين يكون المطلوب الحصول على أحسن خط يمثل العلاقة بين المتغيرين باستخدام هذه البيانات.

ويمكن تمثيل العلاقة بين المتغير المستقل والتابع على شكل معادلة كما يلي:

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

حيث أن : $Y =$ المتغير التابع

$x =$ المتغير المستقل

a : الجزء الثابت من محور الصادات

b : ميل معادلة الانحدار.

يمكنك التنبؤ بقيم المتغير التابع اعتماداً على قيم المتغير المستقل من خلال استخدام معادلة الانحدار الخطي البسيط .