المملكة العربية السعودية
وزارة التعليم العالي
الجامعة الإسلامية بالمدينة المنورة
كلية الأنظمة والاقتصاد

مقرر الإحصاء

جمع وإعداد

مدرس المقرر

د . فهد بن محمد بن بكر آل عابد

الأستاذ المشارك بقسم الاقتصاد – كلية الأنظمة والاقتصاد

بسم الله الرحمن الرحيم

مقدمات

الحمد لله رب العالمين، والصلاة والسلام على أشرف الأنبياء والمرسلين، نبينا محمد وعلى آله وصحبه أجمعين، وبعد:

- تستهدف هذه المذكرة تمكين الطلاب من تطبيق الإحصاء في العلوم الاجتماعية والسلوكية، ومساعدتهم على استخدام الإحصاء كأداة في خدمة أبحاثهم وكتابة تقاريرهم بصورة علمية، وتتطلع لتحقيق هذا الهدف واضعة في اعتبارها أن طالب العلوم الاجتماعية والسلوكية ليس لديه إلمام كبير بمادة الرياضيات، وبالتالي يستشعر الكثير منهم بالرهبة وعدم الثقة بنفسه في خوض هذا المجال.
- تتدرج الموضوعات، وفقا لتوصيف المقرر ووفقا لتصنيف مستويات البيانات، من الإحصاء الوصفي (تنظيم وتبويب البيانات، والأشكال، ومقاييس النزعة المركزية، والتشتت) إلى الإحصاء الاستدلالي (تقدير معالم المجتمع واختبار الفروض الإحصائية ومقاييس المعنوية ومقاييس الارتباط بين المتغيرات).
- مما لا شك فيه أن أفضل العلوم هو العلم الذي يقربك إلى الله عز وجل، وأما بالنسبة للحياة العملية فمن الطبيعي أن تسعى العلوم إلى أن ترتقي سلم التدرج الهرمي، هذا المدرج الهرمي يتربع على قمته العلوم التي تدرس الظواهر الطبيعية كالكيماء والفيزياء والأحياء، أما العلوم الإنسانية والاجتماعية (كالاقتصاد مثلا) فلا تزال تقبع في قاعدة الهرم، وتحاول جاهدة الصعود وإيجاد مكان منافس على قمة الهرم، ولكن حداثة عهد واستقلالية هذه العلوم كان له أثر كبير في عدم تبلور مناهج بحث وأدوات جمع بيانات ذات كفاءة عالية لفحص وتحليل الظواهر التي تتوافر على دراستها، إضافة إلى ذلك فإن هذه العلوم تتعامل مع ظواهر لا يستطيع الباحث التحكم فيها في ظروف معملية وإدخال المعالجات التجريبية عليها كما درج عليه علماء العلوم الطبيعية. ولكن رغم هذه الصعوبات لم يقف علماء هذه العلوم (أي العلوم الاجتماعية) مكتوفي الأيدي، بل دأبوا منذ قرون عديدة على إيجاد الكثير من الوسائل والمناهج لتصنيف وتنظيم المعلومات ومحاولة استقراء العلاقات بين الظواهر بطرق مباشرة وغير

مباشرة بغية الوصول إلى صياغة هذه العلاقات في تعميمها لتصبح فيما بعد قوانين تساعد على التنبؤ (التوقع) بطبيعة نشأة الظواهر والعوامل التي تؤثر فيها.

- لقد تطورت البرامج والتقنيات الإحصائية، وساعدت في إجراء العمليات الإحصائية، لكنها ليست بديلة عن معرفة كيفية إجراء المقاييس بالطرق التقليدية في حل المعادلات الإحصائية باستخدام الآلة الحاسبة. فأداء هذه العمليات بالطرق التقليدية يمكّن الطالب من فهم واستيعاب المفهومات الإحصائية ومواطن استخدامها بصورة واضحة، أما الاقتصار على إجراء العمليات الإحصائية باستخدام البرامج الحديثة دون معرفة قياسها بالطرق التقليدية ينطوي على الكثير من جوانب القصور، وبالتالي لن يتمكن الطالب من فهم واستيعاب المفاهيم الإحصائية ومن ثم لن يكون مؤهلا لاختيار المقاييس المناسبة التي تتناسب مع مقتضيات بحثه ونوعية بياناته.
- هذه المذكرة لا تقتصر فائدتها على طلاب الاقتصاد فقط، وإنما تشمل كافة العلوم الاجتماعية والسلوكية وغيرها كالعلوم الطبية والعلوم الزراعية؛ وذلك لأن المقاييس الإحصائية لا تختلف من علم إلى آخر.

المحاضرة الأولى

المفاهيم الأساسية في علم الإحصاء – أنواع البيانات

أولا: مراحل البحث العلمي:

إن الغرض من العلم هو البحث عن الحقيقة، والبحث العلمي هو الوسيلة للوصول إلى حقائق الأشياء والظواهر ومعرفة كل العلاقات التي تربط بعضها ببعض، سواء كانت هذه الظواهر اجتماعية أو اقتصادية أو طبيعية أو غير ذلك، لذلك، يستخدم البحث العلم لتحري غموض موضوع معين تحرياً منظماً دقيقاً بقصد اكتشاف حقائقه ومعرفة القواعد العامة التي تحكمه، كما إن الإحساس بوجود مشكلة يمثل شرطاً أساسياً للقيام ببحث علمي إلا أن الإحساس بوجود المشكلة لا يأتي إلا من خلال المشاهدة للظواهر المختلفة التي تثير الاهتمام وتتطلب الحاجة إلى بحث وتفسير لتلك المشكلة التي تكون محل الدراسة. وذلك الأمر يتطلب تحديد البيانات الواجب توافرها حتى يمكن إجراء البحث والوصول إلى نتائج دقيقة يمكن الاعتماد عليها في تفسير الظواهر المختلفة. ثم تأتي بعد ذلك مرحلة جمع البيانات من مصادرها المختلفة وتبويب تلك البيانات وعرضها في صور جدولية وبيانية.

ثم يتم استخدام تلك البيانات في حساب المقاييس المختلفة وإجراء عملية تحليل لها بما يساعد في تفسير النتائج المختلفة للبيانات واستخدامها في استنباط نظرية أو قاعدة أو قانون أو المساعدة في اتخاذ القرارات.

ويمكن توضيح مراحل البحث العلمي (في علم الاقتصاد على سبيل المثال) فيما يلي:

١- **المشاهدة:** المشاهدة تعني ملاحظة ظاهرة معينة قد تكون غير مفهومة أو غير مفسرة. تعتمد الملاحظة على متابعة الأسواق، الاقتصاد الكلي أو الجزئي، أو سلوك المستهلكين والشركات.

مثال: لاحظ أحد الباحثين الاقتصاديين ارتفاعًا ملحوظًا في أسعار العقارات في مدينة معينة مقارنةً بمناطق أخرى، مما أثار تساؤلات حول السبب وراء هذا الارتفاع.

٢- **الإحساس بمشكلة أو ظاهرة:** الإحساس بمشكلة اقتصادية يعني إدراك وجود تحدٍ اقتصادي أو ظاهرة تحتاج إلى تفسير، مثل التضخم، البطالة، أو التفاوت في الدخل.

مثال: شعر الباحث الاقتصادي بأن ارتفاع أسعار العقارات قد يكون مرتبطًا بتغيرات في العرض والطلب، أو بقرارات حكومية جديدة تخص الأراضي.

٣- **وضع الفرض العلمي المبدئي اللازم لتفسير الظاهرة:** في هذه المرحلة، يتم صياغة فرضية اقتصادية قابلة للاختبار تفسر الظاهرة بناءً على المشاهدة والتحليل الأولي.

مثال: الفرضية: ارتفاع أسعار العقارات سببه زيادة الطلب الناتج عن هجرة الأفراد من المناطق الريفية إلى المدينة بحثًا عن فرص عمل أفضل.

٤- **خطوات البحث الإحصائي:** تشمل هذه الخطوات تحديد المجتمع الإحصائي، اختيار العينة، وتصميم أدوات جمع البيانات، ثم استخدام أساليب إحصائية مناسبة لتحليلها.

مثال: اختيار عينة من ١٠ مناطق مختلفة داخل المدينة لتحليل بيانات الطلب على العقارات وأثر الهجرة عليها.

٥- **جمع البيانات والمعلومات:** يتم جمع البيانات الاقتصادية من مصادر مختلفة، مثل الإحصاءات الحكومية، بيانات السوق، أو الاستبيانات الميدانية.

مثال: جمع معلومات عن أسعار العقارات على مدى السنوات الخمس الماضية وعدد السكان المهاجرين إلى المدينة خلال نفس الفترة.

٦- **تحليل البيانات:** تحليل البيانات في الاقتصاد يتضمن استخدام برامج إحصائية أو أدوات مالية لتحليل العلاقات بين المتغيرات، مثل العرض والطلب أو تأثير السياسات الحكومية.

مثال: استخدام برنامج مثل Excel أو SPSS لتحليل العلاقة بين الزيادة السكانية وارتفاع أسعار العقارات.

٧- **تفسير البيانات:** في هذه المرحلة يتم تفسير النتائج الإحصائية وربطها بالفرضية الاقتصادية، وتحديد العوامل التي تؤثر على الظاهرة المدروسة.

مثال: تحليل البيانات أظهر أن ارتفاع أسعار العقارات ناتج عن زيادة الطلب بمعدل ١٥٪ خلال السنوات الثلاث الأخيرة، بسبب زيادة المهاجرين الباحثين عن السكن.

٨- **استنباط نظرية أو قاعدة أو قانون أو قرار أو توصية:** المرحلة النهائية التي يتم فيها تقديم توصيات اقتصادية أو صياغة نظرية توضح الظاهرة، مع اقتراح حلول للمشكلة الاقتصادية إذا لزم الأمر.

مثال: التوصية: لخفض أسعار العقارات، يجب على الحكومة زيادة العرض من خلال تقديم تسهيلات للبناء وتطوير ضواحي جديدة لتخفيف الضغط على المدينة.

ثانيا: تاريخ علم الإحصاء وتطوره:

لقد عرف العالم الإحصاء منذ زمن بعيد بشكل مبسط، إذ كان يقوم على أساس فكرة التعداد فقط، وقد بدأ استعماله عندما شعرت بعض الدول والإمارات بحاجتها إلى معرفة بُعْد البيانات العددية عن عدد سكانها وتكاثرهم وأحوالهم لتتمكن من معرفة إمكانياتها واحتياجاتها في حالتي السلم والحرب.

ولقد كان المسلمون من أوائل الشعوب التي استعانت بلغة الأرقام في تعريف أمورها، واستعملت التعداد لإحصاء مواردها، وحصر غنائمها وجندها وأعطياتهم وأسلحتهم ومعرفة الثروات لتحصيل الزكاة عنها.

وعلى الرغم من أن التعداد بمعناه المتقدم كان يتصف بالطابع البدائي البسيط، إلا أنه كان خطوة أولية نحو الإحصاء بمفهومه الحديث، حيث أرسى أولى قواعده، وأقام أولى أسسه، ومن ثم تطور الإحصاء

مع تطور الحضارة وازدهار المدنية مما أدى إلى انتشار استخدامه في جميع مناحي الحياة الاقتصادية والاجتماعية والعلمية، وبعد أن اكتسب الإحصاء صبغته العلمية، لم يعد استخدامه مقتصرا على أمور الدولة وتدبير شؤونها، بل تعداها إلى مختلف الميادين والمحالات وتحول بالتدريج إلى أداة بحث علمي لا غنى عنها في دراسة أغلب العلوم.

وجملة القول، فإن الإحصاء قد مر بثلاث مراحل للتطور ساير من خلالها حاجات الإنسان ورافق في تقدمه تقدم الحضارة وسد حاجاتها حتى أصبح اليوم يحتل مكانة رفيعة وهذه المراحل هي:

أ– مرحلة التعداد: في هذه المرحلة، كان الإحصاء يُستخدم لأغراض بسيطة مثل إحصاء عدد السكان والخيرات المتوافرة في البلاد بهدف التخطيط للإنتاج أو التجهيز للحروب.

مثال: في عهد الفراعنة، كان يتم إجراء تعداد للسكان والمواشي والأراضي الزراعية لتحديد الضرائب والموارد اللازمة لمشروعات البناء مثل الأهرامات.

ب– مرحلة الحساب السياسي: فقد تعدت هذه المرحلة عملية الوصف إلى عملية الوصول إلى القوانين التي تفسر مختلف الأحداث والعمليات الاجتماعية، ومن هذه المرحلة بدأ الإحصاء كعلم، وقد تبلورت هذه المرحلة مع مطلع القرن السادس عشر الميلادي.

مثال: في القرن السادس عشر الميلادي، استخدم الفيلسوف السياسي "توماس هوبز" الإحصاء لتحليل توزيع الثروات والنشاطات الاقتصادية في أوروبا لتقديم رؤى حول النظام السياسي والاجتماعي.

ج– مرحلة الإحصاء وحساب الاحتمالات: وفي هذه المرحلة تم استخدام الأساليب الإحصائية المتقدمة، والتعرف على التوزيعات الإحصائية بأنواعها، وقد بدأت هذه المرحلة تظهر خلال القرن الثامن عشر الميلادي.

مثال: خلال القرن الثامن عشر الميلادي، طور العالم "بيير سيمون لابلاس" نظرية الاحتمالات واستخدمها لتحليل الظواهر الطبيعية والاجتماعية، مثل دراسة أنماط الوفيات لتطوير جداول الحياة المستخدمة في حساب التأمينات.

ثالثا: مجالات استعمال علم الإحصاء في الحياة اليومية:

لقد شهد العالم في الآونة الأخيرة تزايدا سريعا لعجلة التنمية وتميز بخطى التقدم السريعة في المجال التكنولوجي، ومع مطلع القرن الواحد والعشرين بدا واضحا التغير الملحوظ في جميع النظم العالمية على جميع مستوياتها وقد جاء ظهور ذلك التغير نتيجة حتمية ومباشرة للتطور التقني الهائل والطفرة المتسارعة في عالم الاختراعات والاكتشافات، ولا شك بأن لهذا التغير الأثر الواضح على سلوك وحياة المجتمعات السكانية قاطبة وعلى طريقة تعاملها عريضة وتفاعلها، وكمطلب للتعايش ضمن هذا النظام العالمي فلا بد من توفر قاعدة الحقائق والركائز والمعلومات لمعرفة مضامين ذلك النظام من ناحية ومعرفة كيفية التعامل سعه واتخاذ القرارات المناسبة تجاهه من ناحية ثانية، وعلية فليس مستغرباً في وقتنا الحاضر ن يكثر الحديث عن البيانات والمعلومات والمؤشرات سواء كان ذلك الحديث عن ندرتها كيفية الحصول عليها أو توحيد مصادرها أو سبل إنتاجها، بحيث أدى ذلك إلى ظهور صناعة جديدة تسمى صناعة المعلومات.

وفيما يلي سنورد أمثلة لبعض المجالات التي يستعمل فيها الإحصاء والتي كان له دور بارز في حل كثير من مشاكلها وبالتالي تقدمها وتطورها:

- **يستخدم الإحصاء في تطوير التعليم وخططه.**

ومن الأمثلة على ذلك: تحليل نتائج الطلاب لتحديد نقاط الضعف والقوة. تقييم فعالية المناهج الدراسية. تخطيط احتياجات المدارس من المدرسين والمرافق بناءً على الإحصائيات السكانية.

- **يستخدم الإحصاء في مجال الدعاية والإعلانات التجارية.**

ومن الأمثلة على ذلك: تحليل سلوك المستهلكين لتحديد الفئة المستهدفة. قياس فعالية الحملات الإعلانية بناءً على المبيعات. دراسة اتجاهات السوق لتحديد توقيت إطلاق المنتجات.

- **يستخدم بشكل كبير في مجال التأمين.**

ومن الأمثلة على ذلك: استخدام الإحصائيات لتقدير احتمالات الحوادث أو الكوارث الطبيعية لتحديد قيمة الأقساط التأمينية.. تحليل البيانات الديموغرافية (العمر، الجنس، الحالة الصحية) لتحديد الأقساط المناسبة لكل فئة. دراسة نسب المطالبات السابقة للتنبؤ بالمطالبات المستقبلية وضمان استدامة الشركة ماليًا.

- **يستخدم في مجال الاقتصاد والصناعة.**

ومن الأمثلة على ذلك: تقدير معدلات التضخم والنمو الاقتصادي. تحليل تكاليف الإنتاج وأرباح الشركات. دراسة العرض والطلب لتحديد استراتيجيات الإنتاج والتسعير.

- **يستخدم الإحصاء في دراسة مختلف العلوم.**

رابعا: تعريف علم الإحصاء:

كلمة الإحصاء ترجمة عربية مشتقة من اللفظ اللاتيني ستاتوس أو ستاتو، وهذا لفظ يستعمل بمعنى الدولة، كما يستعمل ليشير إلى المعلومات المتصلة بنظام الدولة ومؤسساتها وأجهزتها المختلفة وأحوالها، لذلك أطلق على الإحصاء اسم ستاتستيك Statistic ليدل على مجموعة المعلومات الخاصة بالدولة في وقت من الأوقات، كذلك نلاحظ أن كلمة Statistic بالإنكليزية، وكلمة Statistique بالفرنسية، وكلمة Statistik بالألمانية، كلها مشتقة من كلمة State أي الدولة، الأمر الذي يدل على أن أصل تسمية الإحصاء بهذا الاسم يرتبط بتنظيم أمور الدولة.

ويمكن تعريف الإحصاء في اللغة بأنه: العدُّ الشامل، أما في الاصطلاح فيعرف الإحصاء بأنه: فرع من فروع الرياضيات يهدف إلى جمع وعرض وتنظيم ووصف وتحليل البيانات المقاسة رقميا مما يساعد على اتخاذ قرارات واستنتاجات وتوصيات مبنية على نظرية الاحتمالات.

والمتعمق في هذا التعريف يجد أن الإحصاء ليس هو البيانات الرقمية، وإنما هو فرع تطبيقي من فروع العلوم الرياضية له نظرياته ورموزه واصطلاحاته وأساليبه الخاصة به، لذا فعلم الإحصاء يعالج كيفية

الحصول على البيانات المختلفة وتصنيفها وعرضها واستخلاص النتائج منها، فعلم الإحصاء إذا هو: **العلم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وتلخيصها وعرضها بيانيًا ثم تحليلها وتفسيرها وإجراء المقارنات واستنتاج العلاقات؛ بهدف استخدامها في اتخاذ القرارات المناسبة.**

خامسا: أهداف علم الإحصاء:

من خلال التعريف السابق نخلص بأن علم الإحصاء يهدف إلى:

١. **جمع** البيانات عن الظواهر المختلفة التي تهم الباحث بطرق علمية محددة وبشكل مسبق.

٢. **تبويب** البيانات طبقا لأسلوب تصنيف محدد مسبقا.

٣. **عرض** البيانات باستخدام الجداول أو/و الأشكال البيانية أو/و الرسوم البيانية.

٤. **وصف** البيانات عن طريق إبراز الخصائص الأساسية لها والتي يمكن التعبير عنها بمقاييس معينة ومحددة، والخصائص الأساسية لأي مجموعة من البيانات تقاس بعدة مقاييس منها مقاييس النزعة المركزية، ومقاييس التشتت، وهذه المقاييس تبين مدى ميل المعلومات المجموعة إلى التمركز أو التشتت أو التماثل أو الاعتدال، كما إنها تزود الباحث بتقديرات عن سلوك المجتمع الإحصائي موضع البحث كمجموعة، وليس عن سلوك أي فرد من أفراده بشكل مستقل عن المجموعة.

٥. **تحليل** البيانات المبوبة عن طريق استعمال خصائصها الأساسية التي تم إبرازها للوصول إلى الأرقام ذات العلاقة بالمشكلة والتي يهم الباحث الحصول عليها للوصول إلى نتائج محددة.

٦. **استخدام** النتائج وتفسيرها تفسيراً منطقياً مناسباً لطبيعة المشكلة التي يبحثها حتى يتسنى للباحث الاستفادة منها وتطبيقها في الحياة الواقعية.

ويمكن تلخيص أهداف علم الإحصاء في الكلمات الست:

جمع – تبويب – عرض – وصف – تحليل – استخدام

سادسا: أهمية علم الإحصاء للباحث والبحوث العلمية:

يعتبر علم الإحصاء وسيلة لا غاية، فهو يساعد الباحث على التالي:

١- الوصف الدقيق للظواهر:

يساعد علم الإحصاء الباحثين على وصف الظواهر بدقة من خلال استخدام الأساليب الرياضية والإحصائية، مما يتيح لهم التعمق في البحوث والوصول إلى نتائج دقيقة. هذا الوصف الدقيق يسهل نقل الفهم إلى الآخرين.

مثال: تحليل بيانات الناتج المحلي الإجمالي لتقييم الأداء الاقتصادي لدولة معينة.

٢- التزام الدقة والتحديد:

يجبر استخدام الإحصاء الباحثين على الالتزام بأساليب دقيقة في القياس والتفكير، مما يؤدي إلى تحسين جودة النتائج. القياسات الدقيقة تحد من الأخطاء وتحسن مصداقية البحث.

مثال: قياس تأثير سعر الفائدة على الاستثمار باستخدام بيانات دقيقة ومقارنات مدروسة.

٣- تلخيص النتائج بشكل منظم:

يساهم الإحصاء في تنظيم البيانات الخام من خلال تصنيفها وتلخيصها في جداول أو رسوم بيانية ذات معنى، مما يسهّل استخراج النتائج الهادفة.

مثال: عرض نتائج استبيان عن معدلات الإنفاق الاستهلاكي في جداول تصنيفية مبسطة.

٤- استخلاص النتائج والتعميم:

الإحصاء الاستنتاجي يساعد على استخلاص النتائج من عينات صغيرة وتعميمها على المجتمع، مع تحديد درجة الثقة. يمكن بذلك التوصل إلى استنتاجات دقيقة حول ظواهر معقدة.

مثال: قياس رضا العملاء عن منتج جديد بناءً على عينة عشوائية.

٥- التنبؤ بالظواهر:

يستخدم الإحصاء للتنبؤ باحتمالية حدوث ظاهرة بناءً على بيانات معروفة، مع تحديد مستوى الثقة والخطأ في التقدير.

مثال: التنبؤ بحجم الطلب على منتج جديد بناءً على بيانات السوق السابقة.

٦- تحليل العوامل المعقدة:

يساعد الإحصاء على تحليل العوامل المتشابكة التي تؤثر على ظاهرة ما، من خلال استخدام اختبارات إحصائية لتحديد تأثير كل عامل.

مثال: دراسة تأثير التعليم والدخل على معدلات الادخار للأسر.

سابعا: أقسام علم الإحصاء:

بناء على تعريفنا السابق لعلم الإحصاء يمكن تقسيم هذا العلم إلى فرعين رئيسين هما: الإحصاء الوصفي Descriptive Statistics، والإحصاء الاستنتاجي أو التحليلي Inferential Statistics، ولكل منهما اختصاصاته، وفيما يلى نتناول كل قسم على حدة:

١- الإحصاء الوصفي Descriptive Statistics

يعرف الإحصاء الوصفي بأنه ذلك القسم من الإحصاء الذي يهتم بجمع بيانات المشكلة وتصنيفها وعرضها ثم إجراء الحسابات المختلفة للوصول إلى النتائج المختلفة التي تبرز خصائصها الأساسية.

ويختص الإحصاء الوصفي بتقديم وصف لمفردات مجموعات البيانات سواء أكانت ممثلة لمجتمع إحصائي أم ممثلة لعينة، ويهتم هذا الفرع بدراسة طرق تنظيم وتلخيص وعرض البيانات في صوره جداول إحصائية ورسوم بيانية، وكذلك مقاييس رقمية تجعل من السهولة بمكان الحصول على المعلومات الأساسية بأسلوب مناسب يفهمه العامة قبل الخاصة.

ويتركز دور الإحصاء الوصفي في تلخيص حجم كبير من البيانات الأولية المتاحة عديمة القيمة إلى مجموعه بسيطة من الصيغ الإحصائية مما يمهد الطريق لتحويل هذه البيانات إلى المعلومات سهله الفهم يستفيد منها الجميع، كما يتمثل هدف الإحصاء لوصفي في تحديد الخصائص الرئيسة للمتغير العشوائي والتعرف على اتجاهاته.

ويتحقق هدف "علم الإحصاء الوصفي" من خلال الآتي:

- تبويب وترتيب البيانات الأولية المتاحة في شكل جداول إحصائية.
- عرض البيانات المبوبة في صورة أشكال هندسية.

٢- الإحصاء الاستنتاجي أو الاستدلالي Inferential Statistics

أما الإحصاء الاستنتاجي أو الاستدلالي فيعرف بأنه العلم الذي يدرس الظروف والظواهر الاجتماعية والتربوية متعدياً العرض الوصفي للبيانات الإحصائية إلى تحليل هذه الحقائق والبيانات باستعمال عدد من الأساليب والطرق الإحصائية الاستنتاجية، وذلك باستنتاج معلومات جديدة، واتخاذ قرارات وتوصيات في ضوء تلك النتائج.

ويختص الإحصاء الاستنتاجي أو الاستدلالي بتعميم النتائج التي توصلنا إليها من خلال دراسة عينة عن اتجاهات وخصائص المتغير العشوائي في الواقع العملي والمتمثل في المجتمع الإحصائي المسحوبة منه هذه العينة والذي يهتم به المدير.

وتمثل نظريات الاحتمالات والتقدير حجر الزاوية الأساسي للإحصاء الاستنتاجي أو الاستدلالي في تحقيق هدفه وذلك باستخدام النتائج التي تم تعميمها على المجتمع الإحصائي من خلال دراسة بيانات العينة، كما أننا نذكر هنا أن الإحصاء الاستنتاجي أو الاستدلالي يقوم أيضاً بدراسة نسبة الخطأ في هذه النتائج عند تعميمها على المجتمع الإحصائي مع بيان كيفية التحكم في نسبة الخطأ إلى أقل ما يمكن.

ويعتمد الإحصاء الاستنتاجي أو التحليلي على وسائل لتحقيق هدفه نذكر منها ما يلي:

التقدير: حيث يتم التوصل لتقديرات ذات كفاءة لقيمة المتغير العشوائي في المجتمع من خلال دراسة بيانات العينة، وقد يكون التقدير بنقطة أو التقدير بفترة محددة بنسبة خطأ مسموح بها إحصائيا.

اختبارات الفروض الإحصائية: حيث يتم التوصل إلى قرار بقبول أو رفض أحد الفروض التي تم صياغتها في بداية الدراسة كتفسير مبدئي لسبب الظاهرة قيد البحث والتحليل.

ويلاحظ أن الإحصاء الاستنتاجي أو الاستدلالي يبدأ بالفعل حيث ينتهي الإحصاء الوصفي، فبعد إبراز الخصائص الأساسية للبيانات يبدأ الإحصاء الاستنتاجي، حيث يتم تحليل البيانات واستخدام نتائج التحليل في الاستنتاج ثم تفسير تلك النتائج منطقيا واتخاذ قرارات في ضوء ذلك.

والشكل التالي يوضح أقسام علم الإحصاء الرئيسة والهدف من كل قسم:

علم الإحصاء الوصفي		
جمع البيانات	تنظيم وعرض البيانات	تحليل البيانات وحساب بعض الخصائص الأساسية

علم الإحصاء الاستنتاجي أو الاستدلال الإحصائي
استنتاج معلومات جديدة وتوقع تنبوءات واتخاذ قرارات وتوصيات

مقدمة توضيحية لبعض المصطلحات الإحصائية

المفردة Entity: وهي تعني شخص أو مكان أو وقت أو تاريخ أو أي شيء محل الاهتمام في الدراسة، قابل للعدِّ أو القياس، وهي بمثابة العنصر أو الوحدة الواحدة.

وأمثلتها: شخص: موظف في شركة، طالب في مدرسة، زائر لمعرض. مكان: مدينة، مكتب، قاعة دراسية، مطار. وقت: يوم معين (١ يناير ٢٠٢٥)، فترة زمنية (الربع الأول من السنة)، ساعة معينة (الساعة ٣ مساءً). تاريخ: تاريخ ميلاد، تاريخ إصدار بطاقة، تاريخ بدء مشروع. شيء محدد: سيارة معينة (تويوتا ٢٠٢٠)، منتج معين (هاتف ذكي)، كتاب معين (كتاب تاريخ).

المتغير Variable: وهي أي خاصية كمية أو وصفية تأخذ مفرداتها قيمًا مختلفة عند قياسها أو مشاهدتها ولا يمكن التنبؤ بها مقدمًا. والمتغير العشوائي: عبارة عن جملة وصفية أو كلمة واحدة تتضمن مفردة أو أكثر. ومثالها: العمر، الدخل، النوع، وسيلة الإعلان، عدد الرحلات. فعمر الطالب مثلا يعتبر متغيرا عشوائيا يشتمل على **مفردة** واحدة محل الدراسة وهي **الطالب**.

المجتمع Population: هو عبارة عن دراسة جميع المفردات المكونة لمجموعة البيانات محل الاهتمام؛ بهدف الوصول إلى نتائج تستخدم في التطبيق العملي. مثاله: عند دراسة أحد البنوك، فإن المجتمع الإحصائي قد يكون عبارة عن كل الودائع أو المسحوبات خلال فترة محددة. ويختلف المجتمع الإحصائي في النوع (أشخاص، أماكن، سلع...) كما يتباين في الحجم أيضا. وينقسم إلى محدود وغير محدود:

المجتمع الإحصائي المحدود: هو المجتمع الذي يحتوي على عدد محدد ومعروف من المفردات، بحيث يمكن حصره أو عدّه بشكل دقيق. مثال: عدد الطلاب المسجلين في مدرسة معينة (١٠٠٠ طالب). عدد السيارات المسجلة في مدينة محددة خلال عام معين (٥٠٠٠ سيارة). عدد المنتجات في مخزن معين (٢٠٠ منتج).

المجتمع الإحصائي غير المحدود: هو المجتمع الذي لا يمكن تحديد عدد مفرداته بدقة أو يصعب حصرها لأنها تكون كبيرة جدًا أو متغيرة باستمرار. مثال: عدد النقاط الناتجة عن رمي قطعة نقدية مرات غير

محدودة. عدد الأسماك في المحيطات (لا يمكن عدها بدقة في وقت معين). عدد الزائرين لموقع إلكتروني معين بشكل مستمر وطويل الأمد.

العينة Sample: في معظم الحالات العملية لا يمكن إدخال كل مفردات المجتمع الإحصائي في الدراسة؛ لعدة أسباب منها: التكلفة أو الوقت. فنقوم باختيار جزء فقط من مفردات المجتمع الإحصائي حيث تُجرى عليه الدراسة والتحليل، وهذا الجزء هو ما يُسمى بالعينة. ويجب أن يتم اختيار العينة بطريقة علمية سليمة؛ لتمثل جميع ما يحتوي عليه المجتمع الإحصائي.

المَعْلَمة Parameter: هي قياس وصفي لأحد المتغيرات يتم باستخدام بيانات المجتمع. مثل: المتوسط والانحراف المعياري.

البيانات Data: هي القيمة الوصفية أو الرقمية التي نحتاج إليها لمساعدتنا في جعل القرارات التي نتخذها أكثر معلوماتية في موقف محدد.

أنواع البيانات

إن نوع البيانات التي يتم تجميعها يُحدد لنا طريقة التحليل الإحصائي التي يمكن استخدامها، نظرًا لوجود تحليل إحصائي يختص بكل نوع من البيانات المتوفرة، وغير مسموح بالخطأ في هذه الحالة؛ لأنه سيؤدي إلى نتائج لا يمكن الاعتماد عليها أو الاستفادة منها.

أولا: البيانات الوصفية Categorical Data

وهي ذلك النوع من البيانات التي لا يمكن قياسها رقميًا، ولا يمكن إجراء العمليات الحسابية عليها. مثل: نوع المولود [ذكر/ أنثى]، الحالة الاجتماعية [متزوج/ أعزب...].

وتنقسم البيانات الوصفية إلى قسمين:

أ– البيانات الوصفية الترتيبية Ordinal Data: وهي البيانات الوصفية التي يمكن ترتيبها تصاعديا أو تنازليا. مثل: درجة الموافقة [موافق بشدة/ موافق/ محايد...]، تقدير درجة الطالب [ممتاز/ جيد جدا..]

ب– البيانات الوصفية الاسمية Nominal Data: وهي البيانات الوصفية التي لا يمكن ترتيبها تصاعديا أو تنازليا. مثل: نوع المولود [ذكر/ أنثى]، الحالة الاجتماعية [متزوج/ أعزب...].

ثانيا: البيانات الكمية Quantitative Data

وهي ذلك النوع من البيانات التي يمكن قياسها رقميا ويمكن إجراء العمليات الحسابية عليها. مثل: العمر، المسافة، درجة الحرارة، الطول، الدخل الشهري.

وتنقسم البيانات الكمية إلى قسمين:

أ– البيانات الكمية المتقطعة أو المنفصلة Discrete Data

هي البيانات الرقمية التي تأخذ قيمًا قابلة للعد، لذلك دائما تعبر عن عدد صحيح من الوحدات. مثل: عدد أفراد الأسرة، عدد المقررات، عدد الحوادث.

ب– البيانات الكمية المستمرة أو المتصلة Continuous Data

هي البيانات الرقمية التي تأخذ قيمًا قابلة للقياس، ويمكن أن تشتمل على كسور. مثل: الطول، الوزن، العمر، معدل الذكاء، درجة الحرارة.

المحاضرة الثانية

مستويات قياس البيانات – تنظيم وعرض البيانات

مستويات قياس البيانات:

١- المقياس الاسمي. ٢- المقياس الرتبي. ٣- المقياس الفتري. ٤- المقياس النسبي.

أولا: المقياس الاسمي Nominal Scale:

هو أبسط مستويات القياس ويُستخدم لتصنيف البيانات إلى مجموعات أو فئات لا علاقة بينها من حيث الترتيب أو القيمة. وأمثلته: أسماء المناطق (شرق، غرب، شمال). الحالة الاجتماعية (متزوج، أعزب، مطلق). أنواع السيارات (تويوتا، هوندا، نيسان).

خصائص المقياس الاسمي:

أ- البيانات تُصنَّف فقط دون ترتيب أو مقارنة.

ب- لا تحتوي على قيمة عددية أو كمية.

ج- البيانات على هذا المقياس تُعامل كفئات متساوية (لا يمكن ترتيبها من الأعلى إلى الأدنى).

أبرز اختباراته:

- اختبار كاي مربع (Chi-Square Test) لتحليل التوزيع.

- اختبار فاي.

- اختبارات التكرارات والنسب المئوية.

١) مثال تطبيقي (عام) على الاختبار المستخدم للمقياس الاسمي:

البيانات: دراسة تفضيلات العملاء لنوع معين من القهوة (قهوة عربية، قهوة أمريكية، إسبريسو، لاتيه).

- الاختبار المستخدم: اختبار كاي مربع (Chi-Square Test):

الغرض: معرفة ما إذا كانت تفضيلات العملاء موزعة بالتساوي بين الأنواع أو يوجد فرق جوهري.

مثال عملي: العملاء الذين شملهم الاستطلاع: ١٠٠ عميل.

النتائج: قهوة عربية: ٣٠ ، قهوة أمريكية: ٢٥ ، إسبريسو: ٢٥ ، لاتيه: ٢٠

سؤال البحث: هل هناك اختلاف كبير في تفضيل العملاء لأنواع القهوة؟

٢) مثال تطبيقي (اقتصادي) على الاختبار المستخدم للمقياس الاسمي:

البيانات: تصنيف الدول حسب نظامها الاقتصادي (رأسمالي، اشتراكي، مختلط).

- الاختبار المستخدم: اختبار كاي مربع (Chi-Square Test):

الغرض: دراسة العلاقة بين النظام الاقتصادي ومستوى التنمية (دول نامية، دول متقدمة).

مثال عملي:

الدول الرأسمالية: ١٠ دول متقدمة، ٥ دول نامية.

الدول الاشتراكية: ٣ دول متقدمة، ١٢ دولة نامية.

الدول المختلطة: ٧ دول متقدمة، ٨ دول نامية.

سؤال البحث: هل يوجد ارتباط بين النظام الاقتصادي ومستوى التنمية؟

ثانيا: المقياس الرتبي Ordinal Scale:

يُستخدم لترتيب البيانات وفق مستوى معين من الأفضلية أو الأهمية، مع إمكانية المقارنة بين الفئات، ولكن بدون معرفة الفروقات الدقيقة بينها.

وأمثلته: مستويات الذكاء (مرتفع، متوسط، منخفض). تقييمات العملاء (ممتاز، جيد جدًا، جيد، ضعيف). الاتجاهات (أوافق بشدة، أوافق، لا أوافق).

خصائص المقياس الرتبي:

أ- البيانات مرتبة وفق تدرج معين.

ب- الفروق بين القيم قد تكون غير معروفة أو غير متساوية. (مثل ترتيب الطلاب في الفصل: الأول والثاني والثالث، يمكن ترتيب هذه القيم لكن أداء الطلاب غير معروف).

أبرز اختباراته:

- اختبار ويلكوكسون (Wilcoxon Signed-Rank Test) للمقارنة بين عينة واحدة.

- اختبار مان ويتني (Mann-Whitney U Test) للمقارنة بين مجموعتين.

- اختبار كروسكال واليس (Kruskal-Wallis Test) للمقارنة بين أكثر من مجموعتين.

- معامل سبيرمان للارتباط (Spearman's Rank Correlation).

١) مثال تطبيقي (عام) على الاختبار المستخدم للمقياس الرتبي:

البيانات: تقييم العملاء لخدمة مطعم (ممتاز، جيد جدًا، جيد، ضعيف).

- الاختبار المستخدم: اختبار مان ويتني (Mann-Whitney U Test):

الغرض: مقارنة تقييم العملاء بين فرعين مختلفين لنفس المطعم.

مثال عملي:

فرع ١: ٣٠ عميلًا قيّموا الخدمة (١٠ ممتاز، ١٠ جيد جدًا، ٥ جيد، ٥ ضعيف).

فرع ٢: ٣٠ عميلًا قيّموا الخدمة (٥ ممتاز، ١٥ جيد جدًا، ٥ جيد، ٥ ضعيف).

سؤال البحث: هل توجد فروق بين تقييم العملاء في الفرعين؟

- الاختبار المستخدم: اختبار سبيرمان للارتباط (Spearman's Correlation):

الغرض: قياس العلاقة بين ترتيب العملاء لخدمة الطعام وترتيبهم لجودة النظافة في المطعم.

مثال عملي:

جودة الطعام (ممتاز، جيد جدًا، جيد، ضعيف).

النظافة (ممتاز، جيد جدًا، جيد، ضعيف).

سؤال البحث: هل هناك ارتباط بين ترتيب تقييم جودة الطعام وترتيب تقييم النظافة؟

٢) مثال تطبيقي (اقتصادي) على الاختبار المستخدم للمقياس الرتبي:

البيانات: تصنيف الدول بناءً على مستوى الحرية الاقتصادية (مرتفع جدًا، مرتفع، متوسط، منخفض).

- الاختبار المستخدم: اختبار مان ويتني (Mann-Whitney U Test):

الغرض: مقارنة مستوى الحرية الاقتصادية بين مجموعتين من الدول (الدول ذات الدخل المرتفع والدول ذات الدخل المنخفض).

مثال عملي:

دول الدخل المرتفع: (١٥ مرتفع جدًا، ١٠ مرتفع، ٥ متوسط، ٠ منخفض).

دول الدخل المنخفض: (٥ مرتفع جدًا، ١٥ مرتفع، ١٠ متوسط، ١٠ منخفض).

سؤال البحث: هل هناك فروق جوهرية في مستوى الحرية الاقتصادية بين الدول ذات الدخل المرتفع والدخل المنخفض؟

- الاختبار المستخدم: اختبار سبيرمان للارتباط (Spearman's Correlation):

الغرض: قياس العلاقة بين ترتيب الدول في مؤشر الحرية الاقتصادية وترتيبها في مؤشر الفساد.

مثال عملي:

مؤشر الحرية الاقتصادية: (١، ٢، ٣، ٤، ...).

مؤشر الفساد: (١٠، ٨، ٥، ٣، ...).

سؤال البحث: هل هناك علاقة عكسية بين الحرية الاقتصادية ومستوى الفساد؟

ثالثا: المقياس الفتري Interval Scale:

مقياس يتميز بوجود فروق متساوية بين القيم، ولكنه لا يحتوي على نقطة صفر حقيقية (أي أن الصفر لا يعني انعدام الخاصية). وأمثلته: درجات الحرارة (بالدرجة المئوية أو الفهرنهايت). درجات الذكاء (IQ). التقويم الزمني (السنة الميلادية). ويتشابه في اختباراته مع المقياس النسبي إلى حد كبير.

خصائص المقياس الفتري:

أ- الفروق بين القيم متساوية. (يعني أن الفرق بين درجة الحرارة ٥ و ١٥ هو نفس الفرق بين درجة الحرارة ٢٠ و ٣٠ كلاهما الفرق يساوي ١٠ درجات)

ب- يمكن إجراء العمليات الحسابية (الجمع والطرح) فقط.

ج- لا يمكن حساب النسب؛ لأن الصفر غير حقيقي.

أبرز اختباراته:

- اختبار "t" للعينة الواحدة (One-Sample T-Test).
- اختبار "t" لمجموعتين مستقلتين (Independent T-Test)
- تحليل التباين (ANOVA).

– معامل ارتباط بيرسون (Pearson's Correlation).

– اختبار تحليل الانحدار (Regression Analysis).

١) مثال تطبيقي (عام) على الاختبار المستخدم للمقياس الفتري:

البيانات: درجات الطلاب في اختبار ذكاء (IQ).

- الاختبار المستخدم: اختبار "t" للعينة الواحدة (One-Sample T-Test):

الغرض: مقارنة متوسط درجات اختبار الذكاء مع المتوسط الوطني المعياري (١٠٠).

مثال عملي:

عينة: ٥٠ طالبًا حصلوا على متوسط IQ = 105.

المتوسط الوطني: ١٠٠.

سؤال البحث: هل متوسط IQ للطلاب أعلى من المتوسط الوطني؟

- الاختبار المستخدم: تحليل التباين (ANOVA):

الغرض: مقارنة درجات الذكاء بين ثلاثة فصول دراسية مختلفة.

مثال عملي:

الفصل ١: متوسط IQ = 110.

الفصل ٢: متوسط IQ = 100.

الفصل ٣: متوسط IQ = 95.

سؤال البحث: هل توجد فروق جوهرية في درجات الذكاء بين الفصول؟

٢) مثال تطبيقي (اقتصادي) على الاختبار المستخدم للمقياس الفتري:

البيانات: معدلات التضخم لعدد من الدول (بالنسبة المئوية).

- الاختبار المستخدم: اختبار "t" للعينة الواحدة (One-Sample T-Test):

الغرض: مقارنة متوسط معدل التضخم في مجموعة من الدول مع متوسط التضخم العالمي (٥٪).

مثال عملي: عينة: ١٠ دول، متوسط التضخم = ٦٪. المتوسط العالمي: ٥٪.

سؤال البحث: هل متوسط معدل التضخم في الدول المدروسة يختلف عن المتوسط العالمي؟

- الاختبار المستخدم: تحليل التباين (ANOVA):

الغرض: مقارنة معدلات التضخم بين ثلاث مجموعات من الدول (دول ذات دخل منخفض، دول ذات دخل متوسط، دول ذات دخل مرتفع).

مثال عملي:

الدول ذات الدخل المنخفض: متوسط التضخم = ١٠٪.

الدول ذات الدخل المتوسط: متوسط التضخم = ٦٪.

الدول ذات الدخل المرتفع: متوسط التضخم = ٣٪.

سؤال البحث: هل توجد فروق جوهرية في معدلات التضخم بين هذه المجموعات؟

رابعا: المقياس النسبي Ratio Scale:

مقياس يتميز بوجود فروق متساوية بين القيم، ويحتوي على نقطة صفر حقيقية، مما يسمح بإجراء جميع العمليات الحسابية (الجمع، الطرح، الضرب، القسمة). وأمثلته: العمر (بالسنوات). الطول (بالأمتار). الوزن (بالكيلوغرام). الدخل الشهري (بالعملة).

خصائص المقياس النسبي:

أ- يحتوي على صفر حقيقي يمثل انعدام الخاصية.

ب- يمكن إجراء جميع العمليات الرياضية عليه.

ج- يسمح بحساب النسب بين القيم.

أبرز اختباراته:

- جميع الاختبارات المستخدمة للمقياس الفتري.

- اختبارات تعتمد على النسب (مثل اختبار F-Statistic).

١) مثال تطبيقي (عام) على الاختبار المستخدم للمقياس النسبي:

البيانات: نصيب الفرد من الناتج المحلي الإجمالي (GDP Per Capita) بالدولار الأمريكي.

- الاختبار المستخدم: اختبار تحليل الانحدار (Regression Analysis):

الغرض: دراسة تأثير التعليم والاستثمار الأجنبي المباشر (FDI) على نصيب الفرد من الناتج المحلي الإجمالي.

مثال عملي:

التعليم: عدد سنوات الدراسة.

الاستثمار الأجنبي المباشر: النسبة المئوية من الناتج المحلي.

نصيب الفرد من الناتج المحلي: بالدولار.

سؤال البحث: ما مدى تأثير التعليم والاستثمار الأجنبي على الناتج المحلي للفرد؟

- الاختبار المستخدم: اختبار "t" لمجموعتين مستقلتين (Independent T-Test):

الغرض: مقارنة نصيب الفرد من الناتج المحلي بين دول ذات نظام ضريبي منخفض وأخرى ذات نظام ضريبي مرتفع.

مثال عملي:

الدول ذات الضرائب المنخفضة: متوسط نصيب الفرد = ٥٠,٠٠٠$.

الدول ذات الضرائب المرتفعة: متوسط نصيب الفرد = ٤٠,٠٠٠$.

سؤال البحث: هل هناك فرق جوهري في نصيب الفرد من الناتج المحلي بين الدول ذات الضرائب المنخفضة والعالية؟

- الاختبار المستخدم: تحليل التباين (ANOVA):

الغرض: مقارنة نصيب الفرد من الناتج المحلي بين ثلاث مناطق جغرافية (آسيا، أوروبا، أفريقيا).

مثال عملي:

آسيا: متوسط نصيب الفرد = ١٥,٠٠٠$.

أوروبا: متوسط نصيب الفرد = ٣٠,٠٠٠$.

أفريقيا: متوسط نصيب الفرد = ٥,٠٠٠$.

سؤال البحث: هل توجد فروق جوهرية في نصيب الفرد من الناتج المحلي بين المناطق الثلاث؟

٢) مثال تطبيقي (اقتصادي) على الاختبار المستخدم للمقياس النسبي:

البيانات: دراسة دخل الموظفين في شركة.

الاختبار المستخدم: اختبار تحليل الانحدار: (Regression Analysis)

الغرض: دراسة تأثير مستوى التعليم وعدد سنوات الخبرة على الدخل الشهري.

مثال: التعليم (عدد السنوات الدراسية) .الخبرة (عدد السنوات في العمل). الدخل الشهري (بالعملة).

سؤال البحث :هل يؤثر التعليم والخبرة على مستوى الدخل الشهري؟

- الاختبار المستخدم: اختبار "t" لمجموعتين مستقلتين (Independent T-Test):

الغرض: مقارنة متوسط الدخل بين الرجال والنساء في الشركة.

مثال عملي: الرجال: متوسط دخل = ٥٠٠٠$. النساء: متوسط دخل = ٤٥٠٠$.

سؤال البحث: هل يوجد فرق جوهري بين متوسط دخل الرجال والنساء؟

- الاختبار المستخدم: تحليل التباين (ANOVA):

الغرض: مقارنة دخل الموظفين في ثلاث إدارات مختلفة (الإدارة المالية، إدارة التسويق، إدارة المبيعات).

مثال عملي: الإدارة المالية: متوسط دخل = ٦٠٠٠$. إدارة التسويق: متوسط دخل = ٥٠٠٠$.

إدارة المبيعات: متوسط دخل = ٤٥٠٠$.

سؤال البحث: هل توجد فروق جوهرية في متوسط الدخل بين الإدارات؟

أهمية معرفة مستويات القياس:

- اختيار التحليل الإحصائي المناسب: يساعد في تحديد الطريقة الإحصائية الأنسب لتحليل البيانات (اختبارات تباين، ارتباط، إلخ).
- تصميم الأدوات البحثية: يُستخدم لتحديد نوع الأسئلة المناسبة في الاستبيانات أو الدراسات الميدانية.
- تفسير البيانات بدقة: يُسهم في التمييز بين البيانات الكيفية (مثل المقياس الاسمي والرتبي) والبيانات الكمية (مثل المقياس الفتري والنسبي).
- اتخاذ قرارات فعّالة: فهم نوع البيانات يساعد الباحثين على استخلاص استنتاجات دقيقة تدعم اتخاذ القرارات.

جدول يوضح المقارنة بين مستويات القياس وأبرز الاختبارات المستخدمة لها

الفرق أراد باحث معرفة الفرق أو إجراء مقارنة بين ...	مستوى البيانات	العلاقة أراد باحث معرفة العلاقة أو الارتباط بين ...
– اختبار "t" للعينة الواحدة: مثال: مقارنة متوسط درجات اختبار الذكاء مع المتوسط الوطني المعياري (١٠٠). مثال٢: مقارنة متوسط معدل التضخم في مجموعة من الدول مع متوسط التضخم العالمي (٥٪). **– اختبار "t" لمجموعتين مستقلتين:** مثال١: مقارنة نصيب الفرد من الناتج المحلي بين دول ذات نظام ضريبي منخفض وأخرى ذات نظام ضريبي مرتفع. مثال٢: مقارنة متوسط الدخل بين الرجال والنساء في الشركة. **– تحليل التباين (ANOVA):** مثال: مقارنة درجات الذكاء بين ثلاثة فصول دراسية مختلفة. مثال٢: مقارنة معدلات التضخم بين ثلاث مجموعات من الدول (مجموعة أ ، مجموعة ب ، مجموعة ج).	**البيانات الكمية** **المستوى النسبي** **المستوى الفتري**	**– معامل بيرسون للارتباط:** مثال١: معرفة الارتباط بين ساعات المذاكرة وبين درجات الطلاب. مثال٢: معرفة العلاقة بين الصادرات والنمو الاقتصادي. **– اختبار تحليل الانحدار:** مثال١: دراسة تأثير التعليم والاستثمار الأجنبي المباشر (FDI) على نصيب الفرد من الناتج المحلي الإجمالي. مثال٢: دراسة تأثير مستوى التعليم وعدد سنوات الخبرة على الدخل الشهري.
– اختبار ويلكوكسون للعينة الواحدة: مثال١: الفرق بين مستوى رضا العملاء الحالي عن الخدمة وبين مستوى الرضا المستهدف الذي حددته الإدارة: "جيد جدًا". **– اختبار مان ويتني لمجموعتين:** مثال١: الفروق بين تقييمات العملاء في فرعين لنفس المطعم؟ مثال٢: هل هناك فروق جوهرية في مستوى ترتيب الحرية الاقتصادية بين مجموعتين من الدول (مجموعة أ ، مجموعة ب)؟ **– اختبار كروسكال واليس لأكثر من مجموعة:** مثال: مقارنة مستوى رضا المواطنين عن خدمات الرعاية الصحية في ثلاث مدن مختلفة: بمقياس ترتيب يتضمن: "ممتاز"، "جيد جدًا"، "جيد"، "متوسط"، "ضعيف".	**المستوى الرتبي**	**– معامل سبيرمان للارتباط:** مثال١: هل هناك ارتباط (علاقة) بين ترتيب تقييم جودة الطعام وترتيب تقييم النظافة؟ مثال٢: قياس العلاقة بين ترتيب الدول في مؤشر الحرية الاقتصادية وترتيبها في مؤشر الفساد.
– اختبار كاي مربع: مثال: هل هناك اختلاف كبير في تفضيل العملاء لأنواع القهوة (قهوة عربية، قهوة أمريكية، إسبريسو، لاتيه)؟	**المستوى الاسمي**	**– اختبار كاي مربع:** مثال: هل يوجد ارتباط بين النظم الاقتصادية (رأسمالي، اشتراكي، مختلط) وبين مستوى التنمية (دول نامية، دول متقدمة)؟

تنظيم وعرض البيانات

مما سبق تبين لنا أن البيانات تنقسم إلى: بيانات وصفية (اسمية، ترتيبية) وبيانات كمية (متقطعة، مستمرة)، وفيما يلي تنظيمها وعرضها في جداول وأشكال بيانية على النحو التالي:

أولا: العرض الجدولي للبيانات الوصفية الاسمية:

فيما يلي بيان بألوان السيارات التي تم بيعها، والمطلوب عرضها في شكل جدول، ثم احسب التكرار النسبي والتكرار المئوي لجميع الألوان، والبيانات كما يلي:

أحمر	أصفر	أسود	أحمر	أبيض	أبيض	أبيض	أحمر	أسود	أسود	أزرق
أبيض	أزرق	أصفر	أبيض	أبيض	أحمر	أسود	أسود	أبيض	أبيض	أبيض
أسود	أبيض	أحمر	أزرق	أحمر	أسود	أسود	أبيض	أسود	أبيض	أبيض

الحل:

أولا: نقوم بحساب مجموع السيارات، وتدوين النتيجة في خانة: المجموع.

ثانيا: نقوم بحساب التكرارات لكل لون بشكل مستقل، وتدوينه في خانة التكرار.

ثالثا: نقوم بحساب التكرار النسبي: بقسمة التكرار على مجموع التكرارات. $\left\{\frac{\text{التكرار}}{\text{مجموع التكرارات}}\right\}$

رابعا: نقوم بحساب التكرار المئوي: بضرب التكرار النسبي في ١٠٠. $\left\{\frac{\text{التكرار}}{\text{مجموع التكرارات}}\right\} \times 100$

والجدول التالي يوضح طريقة الحل:

اللون	العلامات	التكرار	التكرار النسبي	التكرار المئوي %
أبيض		١٣	(١٣÷٣٣) = ٠,٣٩	(١٣÷٣٣) × ١٠٠=٣٩,٣٩
أسود		٩		
أحمر				
أزرق	/ / /	٣		(٣ ÷٣٣) × ١٠٠ =٩,٠٩
أصفر				
المجموع		**٣٣**	**١**	**١٠٠٪**

* يمكن حساب المعدل التراكمي للطالب بحساب تكرار النقاط لكل مقرر ثم الضرب في ٥ بدلا من ١٠٠

ثانيا: العرض الجدولي للبيانات الوصفية الترتيبية:

فيما يلي بيان مكان المنشأ الأصلي لعينة من الطلاب مكونة من ٢٥ طالبا:

مدينة كبيرة	مدينة صغيرة	قرية	مدينة كبيرة	مدينة كبيرة
قرية	مدينة كبيرة	مدينة متوسطة	قرية	مدينة صغيرة
مدينة صغيرة	مدينة كبيرة	مدينة كبيرة	مدينة متوسطة	قرية
مدينة كبيرة	مدينة صغيرة	مدينة صغيرة	مدينة صغيرة	مدينة متوسطة
مدينة متوسطة	قرية	مدينة كبيرة	مدينة متوسطة	مدينة صغيرة

المطلوب تنظيم البيانات النوعية (الوصفية) الترتيبية أعلاه في جدول توزيع تكراري مبوب:

الحل:

مكان المنشأ	العلامات	التكرار	التكرار النسبي	التكرار المئوي
قرية				
مدينة صغيرة				
مدينة متوسطة				
مدينة كبيرة				
المجموع				

ثالثا: العرض الجدولي للبيانات الكمية المنفصلة:

أجري بحث شمل عينة من ٢٠ أسرة، وتم تسجيل لكل أسرة عدد الأبناء بها، وحصلنا على النتائج التالية. المطلوب عرضها في صورة جدول، والبيانات كما يلي:

٢	٤	٣	١	٠	٢	٤	٢	٤	٢
٣	١	٢	٤	٣	٤	٣	٤	٠	١

الحل:

عدد الأبناء	العلامات	التكرار	التكرار النسبي	التكرار المئوي
المجموع				

رابعا: العرض الجدولي للبيانات الكمية المتصلة:

البيانات التالية تمثل أوزان عينة من الطلاب:

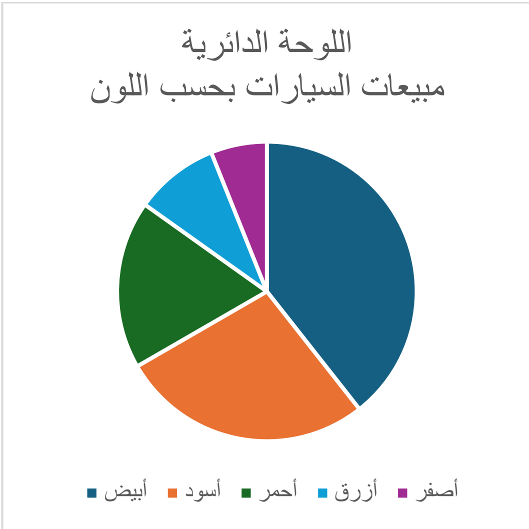
٦٩	٥٨,٢	٧٥	٥٢	٦٠	٨٣	٧٧	٧١,٧	٦٩	٥٥
٨٩	٧٥	٦٥	٦٢	٩٠	٧٣	٥٧	٩١	٨٩	٦٥
٥٥	٥٩	٨٥	٥٢	٨٠	٦٣	٧٧	٨١	٧٩	٨٥,٤
٥٢	٦٦,٥	٥٥	٨٢,١	٧٠	٩٣,٦	٦٧	٩٧	٦٩,٥	٨٥
٦١	٥٧	٧٥	٦٢	٦٠	٧٣	٦٧	٥١	٧٩	٥٥

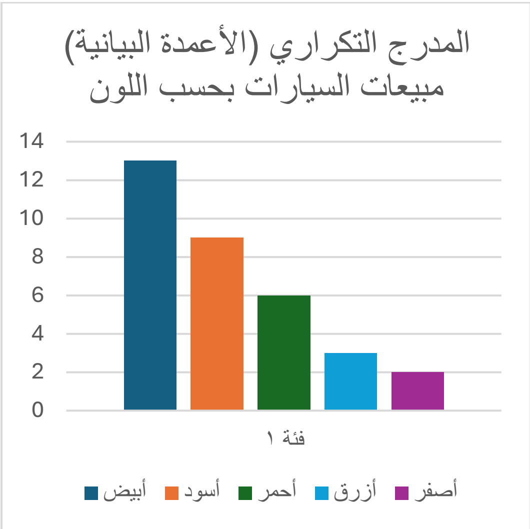
المطلوب عرض البيانات السابقة في صورة جدول تكراري معتبرا أول فئة (٥١ – ٦٠) وطول الفئة = ١٠ أما آخر فئة (٩١ – ١٠٠)، ثم احسب التكرار النسبي والتكرار المئوي لكل تكرار.

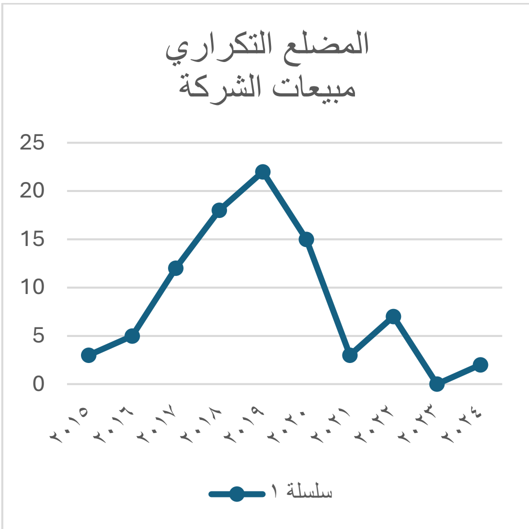
الفئات	العلامات	التكرار	التكرار النسبي	التكرار المئوي
٥١ – ٦٠				
٩١ – ١٠٠				
المجموع				

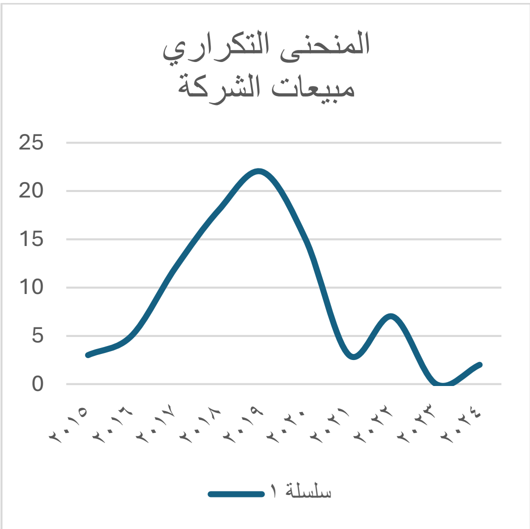
تمثيل التوزيعات الكمية بيانيا: تنقسم الرسوم البيانية إلى عدة أشكال، أهمها ما يلي:

العرض البياني باللوحة الدائرية: $\left\{\frac{\text{التكرار}}{\text{مجموع التكرارات}}\right\} \times 360$ درجة، الأعمدة البيانية (المدرج التكراري)، المضلع التكراري، المنحنى التكراري.









مميزات الأشكال البيانية:

- إثارة انتباه المشاهد، خاصة إذا كانت جيدة التصميم.
- توفير وقت المشاهدة.
- إمكانية معرفة الاتجاهات العامة للظواهر.
- سهولة فهم وتذكر العلاقات بين الظواهر محل الدراسة.

عيوب الأشكال البيانية:

- التضحية بدقة البيانات.
- أحيانا تكون الرسوم البيانية معقدة، خاصة إذا اشتملت على مجموعات من البيانات المتباينة.
- كثرة التكاليف، خاصة إذا كانت البيانات متعددة ومتنوعة وبحاجة إلى رسومات كثيرة.

المحاضرة الثالثة

مقاييس النزعة المركزية

نقصد بمقاييس النزعة المركزية (Measures of Central Tendency) تلك القيم الوسطى التي توضح القيمة التي تجمع أكبر عدد من القيم الخاصة بمجموعة معينة عندها، أو هي تلك الدرجة التي يمكن أن تعتبر ممثلة لكافة الدرجات الموجودة في تلك المجموعة. إذا هي قيم نموذجية يمكن أن تمثل أو تصف مجموعة من البيانات.

ولتحديد القيمة المتوسطة للتوزيع يوجد هناك عدة مقاييس، أهمها: الوسط الحسابي، الوسيط، المنوال، الوسط الهندسي، الوسط التوافقي.

ولمقاييس النزعة المركزية أهمية كبيرة عند معرفتها، منها:

١- النظر إلى الرقم المتوسط يكفي لمعرفة الكثير عن خصائص هذه المجموعة من الأرقام.

٢- إمكانية المقارنة بين عدة مجموعات في وقت واحد.

الوسط الحسابي Mean

الوسط الحسابي أو المتوسط لمجموعة من القيم: هو مجموع هذه القيم مقسوما على عددها.

ويُرمز للوسط الحسابي بـ $\bar{x}$ ويُرمز لعدد القيم بـ n ويُرمز للقيم بـ $x_1\ x_2\ x_3$

ويأخذ شكل المعادلة التالية: $\bar{x} = \frac{\sum x}{n}$

مثال١: أعمار عينة من الطلاب هي:

٢١	٢٤	١٩	١٨	٢٠	٢٣

أوجد قيمة الوسط الحسابي.

الحل:

مجموع القيم (٢١ + ٢٤ + ١٩ + ١٨ + ٢٠ + ٢٣) ÷ عدد القيم ٦ = الوسط الحسابي

مجموع القيم ١٢٥ ÷ عدد القيم ٦ = الوسط الحسابي

٢٠,٨٣ = الوسط الحسابي

مثال٢: إذا كانت درجات ٥ طلاب في مقرر ما، لاختبار محدد من ٢٠ درجة:

١٠	١٢	٧	٢	٩

فما هو الوسط الحسابي لدرجات هؤلاء الطلاب؟

الحل: $\bar{x} = \frac{\sum x}{n} = \frac{9+2+7+12+10}{5} = \frac{\quad}{\quad} =$

إذا الوسط الحسابي لهذه الدرجات =

مثال٣: البيانات التالية تعبر عن المبيعات الشهرية لأحد المحلات التجارية خلال عام ١٤٤٤هـ بالألف ريال، والمطلوب حساب الوسط الحسابي للمبيعات:

الشهر	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

الحل: $\bar{x} = \frac{\sum x}{n} = \frac{\qquad\qquad\qquad\qquad}{\qquad\qquad\qquad\qquad} = \frac{\quad}{\quad} =$

تعقيب: من خلال الأمثلة السابقة يمكن ملاحظة الآتي:

- يمكن تحديد قيمة الوسط الحسابي بالضبط.
- طريقة تحديده سهلة.
- يأخذ في الاعتبار جميع البيانات.
- لا يتأثر بترتيب البيانات.
- لا يُشترط أن يكون الوسط الحسابي أو القيم عددا صحيحا.
- يتأثر بالقيم المتطرفة، مثل: $\frac{10+15+12+13+900}{5} = 190$ ، $\frac{10+15+12+13+9}{5} = 11.8$
- حاصل ضرب قيمة الوسط الحسابي في عدد البيانات = مجموع قيم البيانات

$$\bar{x} = \frac{\sum x}{n} \qquad\qquad n\,\bar{x} = \sum x$$

- إذا أضفنا عددا ثابتا c لكل قيمة من قيم البيانات، فإن:

الوسط الحسابي الجديد = الوسط الحسابي القديم + العدد الثابت c (وأيضا في الضرب)

- المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفرا.

مزايا الوسط الحسابي:

- إذا كان توزيع البيانات توزيعا طبيعيا، فإن المتوسط الحسابي أفضل المقاييس للنزعة المركزية.
- يدخل في حسابه جميع القيم، دون إهمال أي قيمة منها.
- أكثر المقاييس استخداما وأسهلها فهما؛ وذلك نتيجة لسهولة حسابه.

عيوب الوسط الحسابي:

- لا يمكن إيجاده بالرسم.
- يتأثر بالقيم المتطرفة الشاذة قلة أو كثرة. وبالتالي لا يمثل المتوسط الواقعي للبيانات.
- قد لا يساوي عددا صحيحا، أي يمكن أن نجد متوسط عدد الأطفال = ٣,٥ (مجازيا).

الوسيط **Median** ويرمز له بالرمز M

الوسيط: هو القيمة المفردة التي تقع في منتصف البيانات وذلك بعد ترتيبها ترتيبا تصاعديا أو تنازليا.

أوجد القيمة الوسيطية للبيانات التالية:

(أ) ١٢ ، ١٧ ، ١٤ ، ١٩ ، ٢١ ، ١٦ ، ٢١ (عدد فردي من القيم)

(ب) ٦٧ ، ٥٦ ، ٦٢ ، ٧١ ، ٦٦ ، ٥٨ ، ٥١ ، ٤٨ (عدد زوجي من القيم)

الحل: في البداية لابد من ترتيب البيانات تصاعديا أو تنازليا.

(أ) ترتيب البيانات تصاعديا: ١٢ ، ١٤ ، ١٦ ، **١٧** ، ١٩ ، ٢١ ، ٢١

الوسيط = القيمة المفردة التي تقع في منتصف الترتيب التصاعدي أو التنازلي السابق = ١٧

(ب) ترتيب البيانات تصاعديا: ٤٨ ، ٥١ ، ٥٦ ، **٥٨** ، **٦٢** ، ٦٦ ، ٦٧ ، ٧١

الوسيط = القيمة المفردة التي تقع في منتصف الترتيب التصاعدي السابق أي بين القيمتين: ٥٨ ، ٦٢

لذلك نحتاج استخراج الوسط الحسابي لهاتين القيمتين. أي أن الوسيط = (٥٨ + ٦٢) ÷ ٢ = ٦٠

أمثلة: أوجد الوسط الحسابي والوسيط للبيانات التالية:

(أ)	٩	٦	٥	٣	٢	٧	٦	٤	٣	الوسط الحسابي	الوسيط
(ب)	١٨	١٢	٩	٥	١٥	١١	٧	٥		الوسط الحسابي	الوسيط

المنوال Mode

يعرف المنوال لمجموعة من القيم أنه: القيمة التي تتكرر أكثر من غيرها، أو القيمة الأكثر شيوعًا. ولذا يسمى "بالشائع".

فمثلا مجموعة القيم التالية:

١٨	١٢	١١	١٠	١٠	**٩**	**٩**	**٩**	٧	٥	٢	لها منوال: ٩

ومجموعة القيم:

١٦	١٥	١٢	١٠	٨	٥	٣	٩	١	٧	٦	....

أما مجموعة القيم:

٩	**٧**	**٧**	**٧**	٥	٥	**٤**	**٤**	**٤**	٣	٢	....

أي أن مجموعة القيم قد تكون وحيدة المنوال (لها منوال واحد) أو تكون عديدة المنوال، أو عديمة المنوال.

وبالنسبة لمجموعة القيم:

٧	٧	٦	٦	٥	٥	٤	٤

فقد تتسرع وتقول بأنه رباعية المنوال ٧، ٦، ٥، ٤، لكن حيث إن جميع القيم لها نفس التكرار، فنقول بأنها: عديمة المنوال.

والمنوال به العديد من العيوب، منها:

- أنه لا يأخذ في الاعتبار جميع القيم، ولكنه يهتم فقط بالقيم الأكثر تكرارا.
- أنه قد لا يتواجد أو قد يكون هناك أكثر من منوال للبيانات.

إلا أنه يتميز ببعض المزايا، منها:

- أنه أسرع في تحديده من الوسط الحسابي والوسيط.

تمرين: أوجد الوسط الحسابي $\bar{x}$، الوسيط M، والمنوال للقيم التالية:

٦	٨	٢	٥	٩	٥	٦	٢	٥	٣

$\bar{x} =$

$M =$

المنوال =

تمرينات على الوسط الحسابي، الوسيط، المنوال

٩	٢	٧	٣	٨	٤	٥

الوسط الحسابي: الوسيط: المنوال:

٢٠,٠	١٩,٧	١٨,٨	٢٠,٢	٢٢,٤	١٩,٣	٢٠,٦	١٨,٣

الوسط الحسابي: الوسيط: المنوال:

٩٤	٨٢	٩٣	٧٦	٨٥

الوسط الحسابي: الوسيط: المنوال:

س/ الأجر (بالريال) في الساعة لخمسة موظفين في مكتب هو: ٣٢ ، ٣٩ ، ٢٥ ، ٣٧ ، ٩٢

احسب الوسط الحسابي للأجور، ووسيط هذه الأجور. ثم بين أيهما تفضل كمقياس لمتوسط أجر الساعة؟ ولماذا؟

س/ أجور خمسة أشخاص (بالألف ريال): ٣ ، ٧ ، ٢ ، ٥ ، ٣

أ- احسب الوسط الحسابي للأجور الشهرية.

ب- إذا قررت الإدارة زيادة أجورهم، فأيهما يحسن الوسط الحسابي للأجور بصورة أفضل: زيادة أجر كل شخص بمقدار ٢ ألف ريال، أم زيادة الأجور بنسبة ٥٪ ؟

اختر الإجابة الصحيحة:

١) إذا كان الوسط الحسابي لمجموعة من القيم هو ٢٠، **وأضفنا** لكل قيمة من القيم العدد ٢، فإن الوسط الحسابي للقيم الجديدة يكون:

أ- ٢٠ ب- ٢٢ ج- ٤٠ د- ١٨

٢) إذا كان الوسط الحسابي لمجموعة من القيم هو ٢٠، **وضربنا** كل قيمة من القيم في العدد ٢، فإن الوسط الحسابي للقيم الجديدة يكون:

أ- ٢٠ ب- ٢٢ ج- ٤٠ د- ١٨

٣) إذا كان الوسط الحسابي لمجموعة من القيم هو ٢٠، **وضربنا** كل قيمة من القيم في العدد **–٢**، فإن الوسط الحسابي للقيم الجديدة يكون:

أ) ٢٠ ب) ٢٢ ج) ٤٠ د) –٤٠

خاص بالأسئلة من ٤ – ٦: لمجموعة القيم ٤ ، ٥ ، ٨ ، ٩ ، ٤

٤) الوسط الحسابي يساوي: أ- ٨ ب- ٥ ج- ٤ د- ٦

٥) الوسيط يساوي: أ- ٨ ب- ٥ ج- ٤ د- ٦

٦) المنوال يساوي: أ- ٨ ب- ٥ ج- ٤ د- ٦

٧) بفرض توفر البيانات التالية: الوسط الحسابي ومجموع القيم، احسب عدد القيم:

$\bar{x} = 30$ $\sum x = 210$ $n =$

$\bar{x} = 24$ $\sum x = 216$ $n =$

$\bar{x} = 84$ $\sum x = 924$ $n =$

الوسط الهندسي Geometric Mean

نتيجة أن الوسط الحسابي يتأثر بالقيم الشاذة؛ دعت الحاجة إلى وجود مقاييس لا تتأثر بقدر الإمكان بالقيم الشاذة والمتطرفة. ومن تلك المقاييس: الوسط الهندسي، والذي يفيد في بعض التطبيقات الاقتصادية ودراسات نمو الظواهر الديموغرافية وكذلك في حساب الأرقام القياسية[١].

تعريف الوسط الهندسي: هو الجذر النوني لحاصل ضرب مجموعة من القيم الموجبة. ونرمز له: GM.

$$GM = \sqrt[n]{x_1\, x_2 \dots x_n}$$

فمثلا: الوسط الهندسي للقيم: ٢ ، ٤ ، ٨ هو:

$$GM = \sqrt[3]{2 \times 4 \times 8} = \sqrt[3]{64} = 4$$

ويمكن استخدام اللوغاريتمات لتسهيل حساب الوسط الهندسي، كالآتي:

$$\log GM = \frac{\sum \log x}{n} = \frac{\log(x_1 \times x_2 \times \dots x_n)}{n}$$

ومن خلال المعادلة السابقة يمكن تفسيرها بأن: لوغاريتم الوسط الهندسي لمجموعة من القيم: هو الوسط الحسابي للوغاريتمات هذه القيم.

خصائص الوسط الهندسي:

١- يعطي نتائج أكثر اعتدالا من الوسط الحسابي.

٢- لا بد من استخدام جميع القيم دون إهمال لأحدها.

٣- أقل تأثرا بالقيم المتطرفة عن الوسط الحسابي.

مزايا الوسط الهندسي:

١- أكثر تمثيلا للقيم عن الوسط الحسابي، باعتبار عدم تأثره بالقيم المتطرفة مقارنة بالوسط الحسابي.

---

(١) يتم استخدام الوسط الهندسي لحساب الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر) أو (الرقم الأمثل)، كما سيأتي.

٢- يُعد من أفضل المقاييس لحساب متوسطات النسب ومعدلات النمو.

٣- يُعد من أكثر مقاييس النزعة المركزية ملاءمة لحساب الأرقام القياسية.

عيوب الوسط الهندسي:

١- لا يمكن حسابه إذا كانت إحدى القيم تساوي صفرا.

٢- لا يمكن استخدامه إذا كانت إحدى القيم سالبة.

٣- صعوبة حسابه يدويا، وإنما يمكن ذلك بالآلات الحاسبة.

مثال: أوجد كلا من الوسط الحسابي والوسط الهندسي لمجموعة القيم: ٣، ٥، ٦، ٦، ٧، ١٠، ١٢

الحل:

$$\bar{x} = \frac{3+5+6+6+7+10+12}{7} = \frac{49}{7} = 7$$

$$GM = \sqrt[7]{3\times5\times6\times6\times7\times10\times12} = \sqrt[7]{453600} \cong 6.43$$

ولتسهيل الحساب، يمكن استخدام اللوغاريتمات على النحو الآتي:

$$\log GM = \frac{log3+log5+log6+log6+log7+log10+log12}{7} \cong 0.8081$$

$$GM \cong 10^{0.8081} \cong 6.43$$

الوسط التوافقي Harmonic Mean:

يُعد الوسط التوافقي من المقاييس التي تحدُّ من تأثير القيم المتطرفة، وخاصة حالة التطرف نحو الكبر.

تعريف الوسط التوافقي HM: مقلوب الوسط الحسابي لمقلوبات القيم. ويأخذ شكل المعادلة:

$$HM = \frac{n}{\sum\frac{1}{x}}$$

مثال الوسط التوافقي للقيم: ٢ ، ٤ ، ٨

$$HM = \frac{n}{\sum\frac{1}{x}} = \frac{3}{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}} = \frac{3}{0.5+0.25+0.13} = \frac{3}{0.88} = 3.43$$

تمرينات على الوسط الهندسي والوسط التوافقي

أوجد الوسط الهندسي والتوافقي للقيم التالية:

أ- ١٢ ، ١٠ ، ٧ ، ٦ ، ٦ ، ٥ ، ٣

ب- ٥ , ٤ , ٨ , ٣ , ٧ , ٢ , ٩

ج- ٨٥ , ٧٦ , ٩٣ , ٨٢ , ٩٤

د- ١٥ , ٥ , ٩ , ١٢ , ١٨

المحاضرة الرابعة

مقاييس التشتت

كما تميل القيم إلى التمركز، فإنها تميل إلى التشتت أو الانتشار، وبالتالي فإن أي توزيع من القيم له صفة التمركز وصفة التشتت، فمقاييس التشتت (Measures of Dispersion) هي تلك المقاييس التي تعبر عن مدى تباعد القيم أو تقاربها في المجموعات التي يشملها البحث.

فالدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة (أحد مقاييس النزعة المركزية) تُسمى تشتت أو تغير البيانات.

فمثلا إذا كان لدينا ٣ مجموعات من الطلاب، وكل مجموعة مكونة من خمسة طلاب، وكانت لها الدرجات التالية (من ١٠ درجات) في أحد المقررات:

المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
5 , 5 , 5 , 5 , 5	3 , 4 , 5 , 6 , 7	1 , 2 , 5 , 8 , 9
وسطها الحسابي 5	**وسطها الحسابي 5**	**وسطها الحسابي 5**

المجموعات الثلاثة لها وسط حسابي 5، لكن في المجموعة الأولى: جميع القيم متساوية وتساوي 5، في حين **تنتشر** البيانات في المجموعة الثانية حول هذا الوسط بقدر ما، وفي المجموعة الثالثة **تنتشر** البيانات حول الوسط بقدر آخر، أي أن الوسط الحسابي وحده [وهو ممثل لمقياس نزعة مركزية، أي قيمة نموذجية ممثلة للبيانات] ليس كافياً وحده لوصف البيانات، بتعبير آخر نقول أنه كما تميل القيم إلى التمركز فإنها تميل أيضا إلى التشتت أو الانتشار حول قيمة مركزية، وبالتالي فإن أي توزيع من القيم له صفة التمركز وصفة التشتت. ومقاييس التشتت هي تلك المقاييس التي تعبر عن مدى تباعد القيم أو تقاربها في المجموعات التي يشملها البحث.

وهناك العديد من المقاييس التي يمكن استخدامها لقياس هذا التشتت، من أهمها:

المدى، الانحراف المتوسط، الانحراف المعياري، التباين، معامل الاختلاف، القيم المعيارية.

المدى Range

يعرف المدى لمجموعة من البيانات الكمية على أنه: **الفرق بين أكبر قيمة في البيانات وأقل قيمة فيها**. ويُرمز له بالرمز: R.

فمثلا لمجموعة القيم التالية:

15	13	3	5	**18**	12	6	7	**3**	15
R = 18 – 3 = 15					يكون المدى:				

ولمجموعة القيم التالية:

16	14	13	17	**18**	17	15	14	**3**	16
R = 18 – 3 = 15					يكون المدى:				

أي أن المدى نفسه لكلا المجموعتين، في حين تبدو للعين المجردة أن هناك تشتت للبيانات أكبر في المجموعة الأولى عنه في المجموعة الثانية، مما يعني أن المدى هنا لا يُظهر هذا الفارق، ولذا يُعد المدى مقياساً للتشتت لكنه **غير جيد** في كثير من الأحيان. ويصلح كلما كانت البيانات معتدلة التوزيع.

وبالرغم من سهولة تحديده إلا أن به بعض العيوب مثل: تأثره بالقيم المتطرفة كما اتضح من المثال السابق عند حسابه للمجموعة الثانية، حيث تأثرت بالقيمة المتطرفة 3، وإذا تم استبعاد هذه القيمة يكون المدى للمجموعة الثانية مساويا لـ: $R = 18 - 13 = 5$.

الانحراف المتوسط [أو متوسط الانحرافات المطلقة] Absolute Mean Deviation

الانحراف المتوسط [أو متوسط الانحرافات المطلقة] ويرمز له بـ M.D: هو ذلك المقياس الذي يقيس تباعد كافة القيم عن المتوسط الحسابي، ويُعرف الانحراف المتوسط أو متوسط الانحرافات على أنه: متوسط القيم المطلقة للانحرافات عن القيمة المتوسطة للبيانات [عادة تكون الوسط الحسابي]، فإذا اعتبرنا أن القيمة المتوسطة للبيانات هي **الوسط الحسابي**، فإن الانحراف المتوسط لمجموعة من البيانات يأخذ شكل المعادلة التالية: $M.D = \frac{\sum|d|}{n} = \frac{\sum x-\bar{x}}{n}$: حيث $d = x - \bar{x}$ هي انحراف القيمة x عن الوسط الحسابي $\bar{x}$ لمجموعة القيم، أما $|d|$ فهي القيمة المطلقة للانحراف d.

ملحوظة: القيمة المطلقة لأي عدد d مثلا: هي القيمة العددية له دون إشارة $(-\ +)$، ونرمز له بنفس الرمز ولكن بين خطين رأسيين $|d|$، مثلا: $|-3| = 3$ و $|-2.5| = 2.5$ و $|11| = 11$

إذا لحساب الانحراف المتوسط: يلزم حساب الوسط الحسابي أولا، ثم نحسب انحرافات كل قيمة من هذه القيم عن الوسط الحسابي، ثم القيم المطلقة لهذه الانحرافات، ثم متوسط هذه القيم المطلقة على النحو الآتي:

احسب الوسط الحسابي ← احسب الانحرافات عن الوسط ← احسب القيم المطلقة للانحرافات ← احسب متوسط القيم المطلقة

مثال رقم (١) على الانحراف المتوسط:

15	13	3	5	18	12	6	7	3	15

$$\bar{x} = \frac{15 + 13 + 3 + 5 + 18 + 12 + 6 + 7 + 3 + 15}{10} = 9.7$$

وبطرح هذا الوسط من كل قيمة من القيم السابقة: نحصل على **الانحرافات عن هذا الوسط**:

5.3	3.3	**-6.7**	**-4.7**	8.3	2.3	**-3.7**	**-2.7**	**-6.7**	3.5

لاحظ أن مجموع الانحرافات يجب أن = صفر

وبالتالي تكون | القيم المطلقة | لهذه الانحرافات هي:

5.3	3.3	6.7	4.7	8.3	2.3	3.7	2.7	6.7	3.5

إذا الانحراف المتوسط هو متوسط القيم المطلقة للانحرافات، أي:

$$M.D = \frac{5.3 + 3.3 + 6.7 + 4.7 + 8.3 + 2.3 + 3.7 + 2.7 + 6.7 + 3.5}{10} = 4.9$$

ويمكن أن يتم حل المثال السابق بتنظيم الخطوات من خلال جدول على النحو الآتي:

x	$\bar{x}$	$d = x - \bar{x}$	$\lvert d\rvert$	$M.D$
15	9.7	15 – 9.7 = 5.3	5.3	$M.D = \frac{\sum\lvert d\rvert}{n}$ $M.D = \frac{49}{10} = 4.9$
13	9.7	13 – 9.7 = 3.3	3.3	
3	9.7	3 – 9.7 = -6.7	6.7	
5	9.7	-4.7	4.7	
18	9.7	8.3	8.3	
12	9.7	2.3	2.3	
6	9.7	-3.7	3.7	
7	9.7	-2.6	2.6	
3	9.7	-6.7	6.7	
15	9.7	5.3	5.3	
$\sum =$ **97**	**97**	**0**	**49**	

مثال رقم (٢) لمجموعة القيم التالية أوجد الانحراف المتوسط:

x	$\bar{x}$	$d = x - \bar{x}$	$\lvert d \rvert$	$M.D$
16				$M.D = \frac{\sum \lvert d \rvert}{n}$ $M.D = \frac{}{} =$
14				
13				
17				
18				
17				
15				
14				
3				
16				
$\sum =$				

من خلال المثالين السابقين يتضح أن الانحراف المتوسط للمجموعة الثانية أقل من الانحراف المتوسط للمجموعة الأولى، مما يعني أن المجموعة الثانية: أقل تشتتاً من المجموعة الأولى، وهذا ما لم يمكن ملاحظته عند استخدام المدى كمقياس للتشتت.

- من تعريف الانحراف المتوسط يتضح لنا أن الانحراف المتوسط يعتمد تماماً في حسابه على الوسط الحسابي، وبالتالي يكون له نفس مزايا وعيوب الوسط الحسابي: **المزايا**: من السهل حسابه، يأخذ في الاعتبار جميع البيانات، لا يحتاج لترتيب معين للبيانات. **العيوب**: يتأثر بالقيم المتطرفة، لا يمكن إيجاده بالرسم بيانيا.

التباين والانحراف المعياري Variance and Standard Deviation

يُعرف متوسط مربعات الانحراف عن الوسط الحسابي على أنه **تباين** مجموعة البيانات، ويُرمز له بـ S^2، ويُعرف الجذر التربيعي للتباين على أنه **الانحراف المعياري** للبيانات، ويُرمز له بـ S.

ويأخذ التباين شكل المعادلة التالية: $S^2 = \frac{\sum d^2}{n}$

ويأخذ الانحراف المعياري شكل المعادلة التالية: $S = \sqrt{S^2}$

أمثلة على التباين والانحراف المعياري: من خلال قيم المجموعتين التالية، أوجد التباين والانحراف المعياري:

15	13	3	5	18	12	6	7	3	15

16	14	13	17	18	17	15	14	3	16

الحل: نقوم أولا بحساب الوسط الحسابي للبيانات، ثم حساب انحرافات البيانات عن هذا الوسط، ثم مربعات هذه الانحرافات، ثم متوسط هذه المربعات، فنكون حصلنا على تباين البيانات، ويكون الانحراف المعياري هو الجذر التربيعي لهذا التباين. هذه الخطوات يمكن تلخيصها في الجدول التالي:

x	$\overline{x}$	$d = x - \overline{x}$	$\lvert d \rvert$	d^2	الانحراف المتوسط
15					$M.D = \frac{\sum\lvert d\rvert}{n}$
13					
3					$M.D = \frac{\quad}{\quad} =$
5					التباين
18					$s^2 = \frac{\sum d^2}{n}$
12					
6					$s^2 = \frac{\quad}{\quad} =$
7					الانحراف المعياري
3					$s = \sqrt{s^2}$
15					
$\sum =$					$s =$

x	$\overline{x}$	$d = x - \overline{x}$	$\lvert d \rvert$	d^2	الانحراف المتوسط
16					
14					$M.D = \frac{\quad}{\quad} =$
13					
17					التباين
18					
17					$s^2 = \frac{\quad}{\quad} =$
15					
14					الانحراف المعياري
3					
16					
$\sum =$					$s =$

❖ من خلال ما سبق يتضح لنا أن الانحراف المتوسط والانحراف المعياري يعتمد في حسابهما على الوسط الحسابي، وبالتالي يكون لهما نفس مزايا وعيوب الوسط الحسابي.

❖ خاصيتان هامتان للانحراف المتوسط والانحراف المعياري:

- إضافة عدد ثابت c لكل قيمة من قيم البيانات لا يؤثر على قيمة الانحرافين المتوسط والمعياري.

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم

- ضرب كل قيمة من قيم البيانات في عدد ثابت يجعل:

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم × القيمة المطلقة للثابت c

مثال: درجات 5 طلاب بأحد الفصول في مقرر العلوم، كانت كالتالي: 10 , 12 , 7 , 2 , 9

أ) احسب الوسط الحسابي والانحراف المتوسط والانحراف المعياري للدرجات.

ب) إذا أضفنا 5 درجات لكل طالب، ما قيمة الوسط الحسابي والانحرافين المتوسط والمعياري الجديدة؟

ج) إذا أردنا تحسين الدرجات بزيادة كل درجة 50% من قيمتها الأصلية، ما قيمة الوسط الحسابي والانحراف المتوسط والمعياري للدرجات الجديدة؟

x	$\bar{x}$	$d = x - \bar{x}$	$\lvert d \rvert$	d^2	
					$M.D =$
					$s^2 =$
					$s =$

معامل الاختلاف $c.v$

لا تتوقف عملية وصف البيانات للمتغيرات محل الدراسة على حساب مقاييس النزعة المركزية، ولا على مقاييس التشتت فقط، وإنما هناك مقاييس أخرى لا بد من دراستها؛ لأنها تساعد الباحث في الحكم على البيانات محل التحليل والدراسة من حيث درجة التشتت والمقارنة فيما بينها، ومن ذلك ما يسمى بمقاييس التشتت النسبي، ومنها: معامل الاختلاف والقيم المعيارية.

إن تشتت البيانات بمقدار 10 درجات عن قيمة متوسطة 50 درجة مثلا، يختلف عن تشتت قدره 10 درجات عن قيمة متوسطة 200، لذا من المناسب تعريف ما يسمى بالتشتت النسبي.

لمقارنة تشتت مجموعتين من البيانات لا يصح مقارنة التباين أو الانحراف المعياري لكل من المجموعتين، حيث يكون لهما وحدات قياس تختلف على حسب طبيعة الظاهرة محل الدراسة، ولكن مع مقاييس التشتت النسبي يمكن ذلك، ومن أكثر مقاييسه استخداما ما يسمى بمعامل الاختلاف، ويأخذ شكل المعادلة التالية: معامل الاختلاف = الانحراف المعياري ÷ الوسط الحسابي × 100

$$c.v = \frac{s}{\bar{x}} \times 100$$

حيث يتم الاعتماد على كل من الوسط الحسابي والانحراف المعياري، ويُعبر عنه كنسبة مئوية.

مثال: في بحث ما لتحديد مستوى أداء الطلاب في المرحلة الثانوية في المدارس الحكومية والخاصة من خلال درجاتهم في الاختبار النهائي، كانت النتائج على النحو الآتي:

نوع المدارس	الوسط الحسابي	الانحراف المعياري	معامل الاختلاف
المدارس الحكومية	77	16	$c.v = \frac{16}{77} \times 100 = 20.8$
المدارس الخاصة	72	9	$c.v = \frac{9}{72} \times 100 = 12.5$

التعليق: من خلال ما سبق يتضح لنا بأن درجات طلاب المدارس الخاصة أكثر تجانسا من درجات طلاب المدارس الحكومية، مما يشير إلى تفاوت (تشتت) ملحوظ بين مستويات طلاب المدارس الحكومية مقارنة بطلاب المدارس الخاصة.

القيم المعيارية Standard Values

الدرجة المعيارية هي قيمة تقيس مدى انحراف قيمة مفردة من مفردات الدراسة عن الوسط الحسابي لها، وذلك بوحدات من الانحراف المعياري، وبالتالي يمكن الاعتماد على الدرجة المعيارية في المقارنة بين القيم المطلقة للظواهر المختلفة، وتُسمى القيمة المعيارية للقيمة x ويرمز لها عادة بالرمز: z، ومعادلتها:

$$z = \frac{x - \bar{x}}{s}$$

مثال١: في الاختبار النهائي لمقرر الإحصاء حصل طالب على 84 درجة [حيث كان متوسط درجات الطلاب 76 بانحراف معياري قدره 10]، وحصل في مقرر الاقتصاد على 90 درجة [حيث كان متوسط درجات الطلاب 82 بانحراف معياري قدره 16].

<u>هل يمكن القول:</u> بأن درجة استيعاب الطالب لمقرر الاقتصاد كانت أعلى من درجة استيعابه لمقرر الإحصاء؟

الجواب: بالاعتماد على درجة الطالب في المقررين: [الإحصاء: 84، الاقتصاد: 90] تجعل الإجابة: <u>نعم</u> درجة استيعاب الطالب لمقرر الاقتصاد أعلى من درجة استيعابه لمقرر الإحصاء، ولكن <u>الإجابة الصحيحة</u> تعتمد على الدرجة المعيارية للطالب في كل من المقررين وليس على درجة الطالب الحقيقية:

الدرجة المعيارية في مقرر الإحصاء	الدرجة المعيارية في مقرر الاقتصاد الكلي
$z = \frac{x - \bar{x}}{s} = \frac{84 - 76}{10} = \frac{8}{10} = \mathbf{0.8}$	$z = \frac{x - \bar{x}}{s} = \frac{90 - 82}{16} = \frac{8}{16} = \mathbf{0.5}$

أي أن درجة استيعاب الطالب لمقرر الإحصاء أعلى من درجة استيعابه لمقرر الاقتصاد.

(عند المقارنة في القيم المعيارية: لا ننظر إلى القيم المطلقة لها، بل يجب أن نأخذ في الاعتبار الإشارة أيضا؛ لفهم الاتجاه النسبي للقيمة بالنسبة للمتوسط، فلو كانت –٢ فهذا يعني أقل من المتوسط بانحرافين معياريين).

مثال٢: إذا كان مستوى أداء أحد العمال في المؤسسة (أ) = 38 درجة [$\bar{x} = 40 \quad s = 10$]، ومستوى أداء أحد العمال في المؤسسة (ب) = 32 درجة [$\bar{x} = 30 \quad s = 5$]. فأي العاملين درجته أفضل؟

مثال٣: إذا كان تقييم موظف في مصرف الراجحي (أ) = 4 من 5 [$\bar{x} = 3.9 \quad s = 0.5$]، وتقييم موظف آخر في مصرف الإنماء = 3.7 من 5 [$\bar{x} = 3.5 \quad s = 0.2$]. حدد أي الموظفَين أفضل؟

المحاضرة الخامسة

تحليل الارتباط

في دراستنا للفصول السابقة كنا نتعامل مع بيانات ذات متغير واحد [كنا نرمز له بالرمز X]، وفي هذه الفصول السابقة قمنا بـ: جمع البيانات وتنظيمها وعرضها، واستخراج مقاييس خاصة بها. وكل ذلك من خلال القسم الأول من علم الإحصاء، وهو الإحصاء الوصفي.

وفي هذا الفصل سنتناول الأساليب الإحصائية لتقييم **العلاقات** بين عدة متغيرات من خلال ما يسمى بتحليل الارتباط؛ لمعرفة ما إذا كان هناك علاقة (ارتباط) بين ظاهرتين أو أكثر، وأيضا **التنبؤ** بأداء الظاهرة في المستقبل، وذلك من خلال ما يُسمى بـ تحليل الانحدار.

في هذا الفصل سنتعامل مع بيانات يمثلها متغير x وبيانات أخرى يمثلها متغير آخر y، ونبحث في:

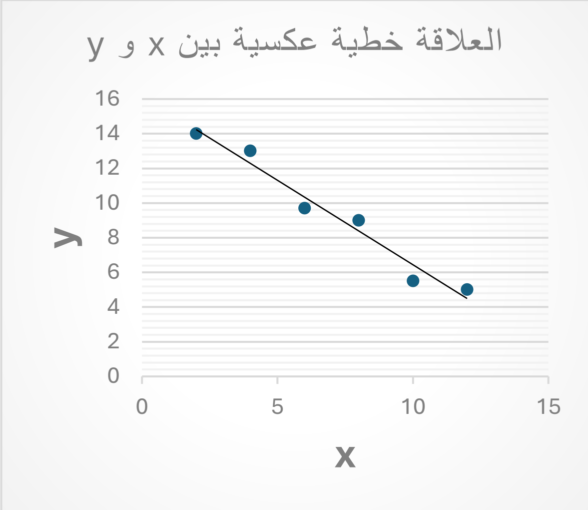
١) هل هناك علاقة بين هاتين المجموعتين من البيانات أم لا: فإذا كانت هناك علاقة نقول إن المتغيرين X ، Y مرتبطان، وإلا فهما غير مرتبطين.

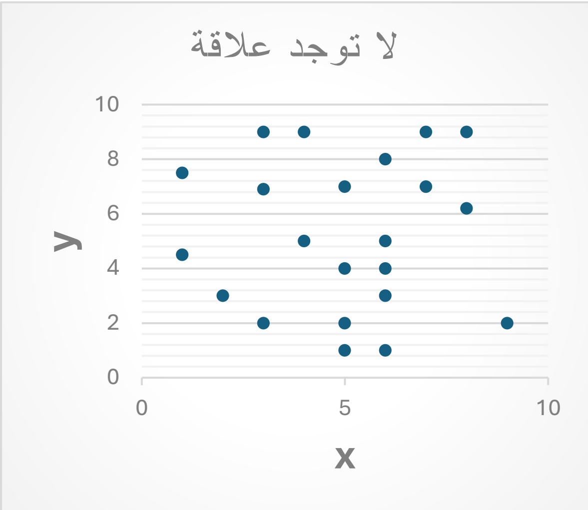
٢) مدى قوة هذه العلاقة إن وُجدت، وهل هي قوية جدا أم قوية أم متوسطة أم ضعيفة أم ضعيفة...

٣) نوع هذه العلاقة [إن وُجدت]، وهل هي طردية أم عكسية.

الطردية: كلما زادت قيمة x زادت قيمة y. والعكسية: كلما زادت قيمة x نقصت قيمة y.

x	2	4	6	8	10	12
y	5	5.5	9	9.7	13	14





قوة الارتباط: إذا أمكن رسم خط مستقيم يمر بجميع نقاط شكل الانتشار سُمي الارتباط "ارتباط تاما". وإذا أمكن رسم خط مستقيم بحيث تكون انحرافات النقاط عنه ضعيفة جدا سُمي الارتباط "ارتباط قوي". وإذا كانت بعيدة بشكل معقول سمي بـ "ارتباط متوسط"، وإذا كانت الانحرافات بعيدة جدا سُمي الارتباط "ارتباط ضعيف".

معامل الارتباط البسيط: يُستخدم في تحديد ما إذا كان هناك علاقة بين المتغيرين، وكذلك تحديد نوع وقوة العلاقة إن وُجدت، أما في حالة دراسة مدى وجود علاقة ارتباطية بين أكثر من متغيرين فإنه يتم الاعتماد على معامل الارتباط المتعدد و/أو معامل الارتباط الجزئي. حيث يتم دراسة تأثير أحد المتغيرات مع تثبيت المتغيرات الأخرى [كما في حالة كثير من المشكلات الاقتصادية حيث يتم دراسة تأثير السعر على الكمية المطلوبة بفرض ثبات الجودة ومستوى الذوق].

ويتم استخدام معامل الارتباط في الحكم على نوع العلاقة بين المتغيرين [حيث تكون علاقة طردية أو عكسية]، وكذلك بالنسبة لقوة العلاقة [فقد تكون قوية أو متوسطة أو ضعيفة].

وعادة ما يتم تقسيم المتغيرات محل الدراسة إلى:

متغيرات مستقلة x: وهي التي بتغير قيمتها تؤثر في تغيير قيمة متغير أو متغيرات أخرى.

متغيرات تابعة y: وهي التي تتغير قيمتها بتغير المتغيرات المستقلة أو إحداها.

وسيتم قياس الارتباط البسيط من خلال كل من:

- معامل الارتباط الخطي البسيط لبيرسون[١] Person's Correlation Coefficient
- معامل ارتباط الرتب لسبيرمان[٢] Spearman Rank Correlation Coefficient

معامل الارتباط الخطي البسيط لبيرسون:

يُعتبر معامل الارتباط الخطي البسيط لبيرسون والذي يُرمز له بـ r_p من أكثر الأدوات الإحصائية استخداماً في تحديد قوة العلاقة بين متغيرين، كما يُستعمل لتحديد مدى وجود علاقة خطية بين متغيرين، وهناك أكثر من صيغة يمكن الاعتماد عليها في حساب هذا المعامل على النحو التالي:

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\ \sqrt{n\sum y^2 - (\sum y)^2}}$$

وتتراوح قيمة معامل الارتباط بين 1+ و 1- بمعنى: $[-1 \leq r_p \leq 1]$، وطبقا لتلك القيمة تكون قوة العلاقة على النحو التالي:

قوة معامل ارتباط بيرسون:

القيمة	0	0.3 – 0.1	0.5 – 0.3	0.7 – 0.5	0.9 – 0.7	0.99 – 0.9	1

(٢) كارل بيرسون (١٩٣٦م): هو رياضياتي وزميل للجمعية الملكية، وزمالة الجمعية الملكية: هي جائزة وزمالة تُعطى من الجمعية الملكية في لندن للأفراد الذي ساهموا بمساهمات كبيرة في تحسين المعرفة الطبيعية كالرياضيات والهندسة والعلوم والطب. وضع أسس الإحصاء الرياضي. وفي عام ١٩١١م أسس أول قسم للإحصاء في العالم في كلية لندن الجامعية.

(٨) تشارلز إدوارد سبيرمان (١٩٤٥م): هو عالم نفس إنجليزي له أيضا مساهمات في علم الإحصاء حيث كان من رواد التحليل العاملي، وقد ابتكر معامل ارتباط الرتب، الذي يُعرف باسمه حتى الآن. كما كانت له أبحاث رائدة في مجال الذكاء البشري.

القوة	منعدمة	ضعيفة جدا	ضعيفة	متوسطة	قوية	قوية جدا	تامة

- إذا كانت قيمة معامل الارتباط أكبر من ١، فهذا لا يعني إلا: **وجود خطأ في الحسابات**.
- إذا كانت إشارة معامل الارتباط **موجبة**، فهذا يعني أن العلاقة **طردية**.
- وإذا كانت إشارة معامل الارتباط **سالبة**، فهذا يعني أن العلاقة **عكسية**.

معامل ارتباط بيرسون	المعنى
0.9	ارتباط طردي قوي جدا
-0.87	ارتباط عكسي قوي
-0.21	ارتباط عكسي ضعيف
0.43	ارتباط طردي متوسط
1	ارتباط طردي تام
-0.51	ارتباط عكسي متوسط

مميزات معامل بيرسون:

- سهولة الحساب: يمكن حساب معامل بيرسون بسهولة نسبيا.
- التفسير: يمكن تفسير قيمة معامل بيرسون بسهولة.
- التعميم: يمكن استخدام معامل بيرسون مع مجموعة متنوعة من البيانات.

عيوب معامل بيرسون:

- افتراض التوزيع الطبيعي للبيانات: فيجب أن تكون البيانات موزعة بشكل طبيعي.
- تأثيره بالقيم المتطرفة، وبالعينة الصغيرة: يمكن أن تؤثر القيم المتطرفة أو العينة الصغيرة بشكل كبير على قيمة معامل بيرسون، فتظهر القيمة المحسوبة أكبر من القيمة الجدولية.
- عدم ملاءمته للبيانات الترتيبية.

مثال: فيما يلي بيان المبالغ التي تم إنفاقها على الإعلان x [بالمليون ريال]، والمبيعات y [بالمليون ريال]

x	2	3	2	7	6	5	10	15	4	11	9	8
y	10	12	9	22	18	19	26	33	18	22	15	17

احسب معامل بيرسون للارتباط الخطي البسيط.

الحل:

x	y	xy	x^2	y^2
2	10	20	4	100
3	12	36	9	144
2	9	18	4	81
7	22	154	49	484
6	18	108	36	324
5	19	95	25	361
10	26	260	100	676
15	33	495	225	1089
4	18	72	16	324
11	22	242	121	484
9	15	135	81	225
8	17	136	64	289
82	221	1771	734	4581

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\ \sqrt{n\sum y^2 - (\sum y)^2}}$$

$$r_p = \frac{12(1771) - (82)(221)}{\sqrt{12(734) - (82)^2}\ \sqrt{12(4581) - (221)^2}} = 0.8756$$

معامل التحديد:

معامل التحديد هو مربع معامل الارتباط، ويرمز له بالرمز r^2، ويُشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغيُّر في المتغير التابع، فمثلا نجد أن $r^2 = 0.8756^2 = 0.76675$ مما يعني أن المنفق على الإعلان يفسر نسبة 0.76675% من التغير في المبيعات، بينما 23.32% من التغير في المبيعات ترجع إلى متغيرات أو عوامل أخرى منها الخطأ العشوائي.

الدلالة الإحصائية:

لا نكتفي فقط بحساب معامل الارتباط لبيرسون، بل لابد من مقارنة معامل الارتباط (r_p) المحسوبة مع معامل الارتباط (r_p) الجدولية = 0.576. حيث درجة الحرية لمعامل ارتباط بيرسون = n-2

فنقول: بما أن (r_p) المحسوبة أكبر من (r_p) الجدولية وفق مستوى الدلالة 0.05 فإننا نرفض الفرض الصفري ونقبل الفرض البديل.

وبالتالي: **توجد علاقة طردية قوية بين الإنفاق على الإعلان والمبيعات، فكلما زاد الإنفاق زادت المبيعات.**

مثال٢: على افتراض مثال آخر، كان عدد القيم 12. ووفق المعطيات التالية احسب معامل بيرسون للارتباط الخطي البسيط **واحسب نسبة التأثير** مع بيان اتخاذ القرار والنتيجة. الحل:

x	y	xy	x^2	y^2
2	10	20	4	100
3	12	36	9	144
…	…	…	…	…
142 $\sum x$	442 $\sum y$	5722 $\sum xy$	1854 $\sum x^2$	18324 $\sum y^2$

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\ \sqrt{n\sum y^2 - (\sum y)^2}}$$

$$r_p = \frac{\qquad}{\qquad} =$$

نسبة التأثير= معامل التحديد = $r^2 =$

حساب r الجدولية:

الدلالة الإحصائية:

ملاحظات متعلقة بالدلالة الإحصائية:

١– عند مقارنة قيمة R المحسوبة مع قيمة R الجدولية فإننا ننظر إلى القيم المطلقة لها.

٢– أحيانا تكون قيمة R المحسوبة أصغر من الجدولية في جدول القيم الحرجة ذات الطرفين وتكون أكبر في القيم ذات الطرف الواحد. لماذا يمكن أن تختلف النتيجة بين الاختبارين؟

- في الاختبار ثنائي الطرف: توزيع الدلالة الإحصائية موزع على طرفين، مما يعني أن القيمة الحرجة تكون أعلى (أكثر صرامة) لأنه يجب أن تأخذ في الاعتبار احتمالية الخطأ في كلا الاتجاهين (الإيجابي والسالب). لذلك، من الصعب الوصول إلى مستوى دلالة إحصائية في اختبار ثنائي الطرف مقارنة بأحادي الطرف.
- في الاختبار أحادي الطرف: نظرًا لأنك تفحص فقط احتمال التأثير في اتجاه واحد (مثلاً في الاتجاه الإيجابي فقط)، يكون من الأسهل تجاوز القيمة الحرجة لأن نسبة الخطأ موزعة في اتجاه واحد فقط. وبالتالي، قد تظهر العلاقة كذات دلالة إحصائية في هذا الاتجاه.

الخلاصة:

يمكن أن تظهر النتيجة "لا توجد علاقة" في الاختبار ثنائي الطرف لأنه أكثر صرامة ويأخذ في الاعتبار كلا الاتجاهين، بينما يظهر في الاختبار أحادي الطرف وجود علاقة ذات دلالة إحصائية في اتجاه واحد.

معامل ارتباط الرتب لسبيرمان r_s

نلاحظ مما سبق أن معامل الارتباط لبيرسون لا يمكن استخدامه في حساب قوة العلاقة بين متغيرين إلا إذا كانت البيانات المتوفرة عنها في صورة كمية فقط، أما إذا كانت في صورة وصفية فلا يمكن تطبيقه في حساب الارتباط بين المتغيرين محل الدراسة، لذا ظهر معامل ارتباط الرتب لسبيرمان r_s والذي يمكن استخدامه في قياس الارتباط خاصة في حالة البيانات الوصفية الترتيبية مثل تقديرات الطلاب: ممتاز/جيدجدا/جيد/مقبول/ضعيف، وكذلك قوة المركز المالي: جيد/متوسط/ضعيف، ودرجة الموافقة على الرأي في أسئلة الاستبانة: موافق تماما/موافق/محايد/غير موافق/غير موافق تماما.

ويُعتبر معامل ارتباط الرتب لسبيرمان من الاختبارات اللابارامترية، ويمكن استخدامه في إيجاد العلاقة بين المتغيرات الكمية ولكن يتوجب تحويل قيمها إلى رتب، أي أن معامل سبيرمان لا يتعامل مع البيانات الخام للمتغيرين، وإنما على رتب قيمهما. ولا يشترط أن تكون العلاقة بين المتغيرين خطية، ولا البيانات أن تتبع التوزيع الاعتدالي (الطبيعي) للمتغيرين.

ويتم حساب معامل ارتباط الرتب لسبيرمان r_s باستخدام المعادلة التالية:

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

حيث d هو الفرق (في الرتب) بين المتغيرين، n عدد المشاهدات.

قوة معامل ارتباط الرتب لسبيرمان:

القيمة	0	0.3 – 0.1	0.5 – 0.3	0.7 – 0.5	0.9 – 0.7	0.99 – 0.9	1
القوة	منعدمة	ضعيفة جدا	ضعيفة	متوسطة	قوية	قوية جدا	تامة

شروط استخدامه:

١- عندما تكون البيانات محل الدراسة رتبية.

٢- عندما يمكن تحويل البيانات الكمية للمتغيرين إلى بيانات رتبية.

ملاحظات يجب مراعاتها عند ترتيب المتغيرات:

- يتم ترتيب قيم مشاهدات المتغير x [وتسمى القيم الترتيبية للمتغير x]، وكذلك للمتغير y، ويكون الترتيب لكلا المتغيرين تصاعديا أو تنازليا بنفس الترتيب.
- في حالة الترتيب التصاعدي مثلا يتم إعطاء أقل قيمة الرتبة 1 والقيمة التي هي أكبر منها الرتبة 2، ... وهكذا.
- في حالة تكرار أو تساوي بعض القيم لأي متغير تُعطى كل منهم رتبة كما لو كانت القيم غير متساوية، ثم نحسب الوسط الحسابي (مجموع الرتب / عددها) لتلك الرتب ويُعطى الوسط الحسابي كرتبة لتلك القيم المتساوية.

مثال: على بيانات المثال السابق، احسب معامل الارتباط لسبيرمان بين المنفق على الإعلان والمبيعات.

الحل:

يتم أولا ترتيب قيم كل من x و y كما يتضح من الجدول التالي:

المنفق x	المبيعات y	رتب x		رتب y		d	d^2
2	10	~~1~~	1.5	~~2~~	2	-0.5	0.25
3	12	~~3~~	3	~~3~~	3	0	0
2	9	~~2~~	1.5	~~1~~	1	0.5	0.25
7	22	~~6~~	6	~~10~~	9.5	-3.5	12.25
9	18	~~8~~	8.5	~~7~~	6.5	2	4
5	19	~~5~~	5	~~8~~	8	-3	9
10	26	~~10~~	10	~~11~~	11	-1	1
15	33	~~12~~	12	~~12~~	12	0	0
4	18	~~4~~	4	~~6~~	6.5	-2.5	6.25
11	22	~~11~~	11	~~9~~	9.5	1.5	2.25
9	15	~~9~~	8.5	~~4~~	4	4.5	20.25
8	17	~~7~~	7	~~5~~	5	2	4
الأرقام المشطوبة تمثل الترتيب المبدئي وليس النهائي (والذي يمثل الرتب)						0	59.5

نلاحظ من الجدول أنه:

- تم ترتيب المتغيرين تصاعديا عند إدخال الرتب.
- عند ترتيب قيم المتغير x نجد أن القيمة 2 تكررت مرتين لتأخذ الرتب 1 و 2 لذلك نحسب المتوسط لهما وهو 1.5 ونضعه أما القيمة 2، وكذلك بالنسبة للقيمة 9 فإنها تأخذ الرتبة 8 و 9 لذلك وضعنا أمام القيمة 9 الرتبة 8.5. وهكذا بنفس الطريقة لقيم المتغير y.
- عند حساب الفرق بين رتب المتغيرين، نلاحظ أن مجموع الفروق يجب أن يكون صفراً، وألا يكون هناك خطأ في الترتيب لأحد المتغيرين أو كليهما، ولابد من مراجعة الترتيب مرة أخرى.

وبما أن عدد المشاهدات $n=12$ فإنه يمكن حساب معامل ارتباط الرتب لسبيرمان كما يلي:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(59.5)}{12(144-1)} = 0.7919$$

أي أن معامل ارتباط الرتب لسبيرمان 0.7919 مما يدل على وجود ارتباط طردي قوي بين المنفق على الإعلان وبين المبيعات، وهي قيمة قريبة من التي تم حسابها باستخدام معامل الارتباط لبيرسون التي بلغت 0.8756

مثال٢: البيانات التالية تمثل التقديرات التي حصل عليها عشر طلاب في مقرري المحاسبة x والقانون y:

x	A	B	C	D	F	C	D	B	C	D
y	C	C	D	C	B	B	A	D	C	B

حيث A تعني "ممتاز"، B "جيد جدا"، C "جيد"، D "مقبول"، F "راسب".

المطلوب حساب معامل الارتباط المناسب. مع بيان الدلالة الإحصائية.

الحل: هنا لا يمكن حساب معامل بيرسون للارتباط الخطي؛ لأن البيانات وصفية وليست كمية، لذا يكون معامل الارتباط الذي يمكن حسابه هو معامل ارتباط الرتب "معامل سبيرمان"، لذا يتم ترتيب المشاهدات وحساب الفروق بين الرتب ومربعاتها كما يتضح من الجدول التالي:

المحاسبة x	القانون y	رتب x		رتب y		d	d^2
A	C	~~10~~	10	~~3~~	4.5	5.5	0.25
B	C	~~8~~	8.5	~~4~~	4.5	4	0
C	D	~~5~~	6	~~1~~	1.5	4.5	0.25
D	C	~~2~~	3	~~5~~	4.5	-1.5	12.25
F	B	~~1~~	1	~~7~~	8	-7	4
C	B	~~6~~	6	~~8~~	8	-2	9
D	A	~~3~~	3	~~10~~	10	-7	1
B	D	~~9~~	8.5	~~2~~	1.5	7	0
C	C	~~7~~	6	~~6~~	4.5	1.5	6.25
D	B	~~4~~	3	~~9~~	8	-5	2.25
الأرقام المشطوبة تمثل الترتيب المبدئي وليس النهائي (والذي يمثل الرتب)						0	247

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(247)}{10(100-1)} = -0.4969$$

الدلالة الإحصائية

لا نكتفي فقط بحساب معامل الارتباط لسبيرمان، بل لابد من مقارنة معامل الارتباط (r_s) المحسوبة مع معامل الارتباط (r_s) الجدولية = 0.648. حيث لا نحتاج إلى درجات الحرية، وإنما إلى حجم العينة لمعرفة القيمة الجدولية. **وعند المقارنة ننظر إلى القيم المطلقة r_s**

فنقول: بما أن (r_s) المحسوبة أصغر من (r_s) الجدولية وفق مستوى الدلالة 0.05 فإننا نقبل الفرض الصفري ونرفض الفرض البديل.

وبالتالي: لا توجد علاقة ذات دلالة إحصائية بين تقديرات مقرر المحاسبة وبين تقديرات مقرر القانون.

المحاضرة السادسة

تحليل الانحدار

تمهيد: في الفصل السابق (تحليل الارتباط) تمت دراسة درجة أو قوة الارتباط بين ظاهرتين (متغيرين)، دون بتحديد درجة تأثير أحدهما في الآخر.

بينما في هذا الفصل (تحليل الانحدار) تتم دراسة وتحليل أثر متغير كمي (يُوصف بالمتغير المستقل) على متغير كمي آخر (يُوصف بالمتغير التابع)، ومن الأمثلة على ذلك:

- دراسة أثر الإنتاج على التكلفة.
- دراسة أثر الدخل على الاستهلاك.
- دراسة أثر الإنفاق الإعلاني على المبيعات.
- دراسة أثر الصادرات على الناتج المحلي الإجمالي

ويهتم تحليل الانحدار بصياغة العلاقة بين المتغيرين على شكل معادلة رياضية يمكن الاستفادة منها بالتنبؤ بقيمة أحد المغيرين. وإذا كان هناك متغير مستقل واحد فقط، فيُوصف بالانحدار الخطي البسيط.

مفهوم تحليل الانحدار:

هو أسلوب رياضي قدمه (جالتون) بهدف الاستفادة من الارتباط في التنبؤ أو تقدير قيمة متغير ما. فقط وجد من خلال بحثه حول (وراثة طول القامة) أنَّ الأطفال الذين يأتون من آباء طوال القامة يميلون لأن يكونوا أقصر قامة من آبائهم، والذين يأتون من آباء قصار القامة يميلون لأن يكونوا أطول قامة من آبائهم، أي أن طول الأبناء يميل إلى التراجع أو الانحدار نحو المتوسط العام، لذا أطلق على هذه القاعدة اسم قاعدة الانحدار، وأطلق على الخط الذي يُوضح هذه القاعدة اسم خط الانحدار.

فإذا كان لدينا درجات اختبارين للطلاب وقمنا بقياس العلاقة الارتباطية بينهما، فإنه من خلال معادلة الانحدار يمكننا أن نتنبأ بدرجة كل فرد في الاختبار الثاني بناءً على درجاته في الاختبار الأول، وقد سُمي هذا الأسلوب الإحصائي بالانحدار؛ لأنه ينحدر في تقديره للدرجات المختلفة نحو المتوسط.

أهداف تحليل الانحدار:

١) تقدير العلاقة بين المتغيرين والتي من خلالها يمكن معرفة التغيّر في أحد المتغيرين على أساس تأثره بالمتغير الآخر.

٢) قياس مدى الارتباط بين المتغير المستقل والمتغير التابع.

٣) تقدير نسبة تفسير كل متغير مستقل بناءً على تغيّر المتغير التابع.

٤) معرفة اتجاه تأثير المتغير المستقل على المتغير التابع، هل هو تأثير إيجابي أم سلبي؟

أنواع تحليل الانحدار:

١) الانحدار البسيط: يُستخدم لدراسة العلاقة بين متغيرين فقط، متغير مستقل ومتغير تابع.

٢) الانحدار الجزئي: يدرس العلاقة بين المتغير التابع وواحد فقط من المتغيرات المستقلة، بفرض أن العوامل الأخرى ثابتة.

٣) الانحدار المتعدد: يدرس العلاقة بين المتغير التابع والمتغيرات المستقلة كلها.

والنوع الأول هو محل دراستنا في هذا الفصل.

شروط استخدام الانحدار الخطي البسيط:

١) العشوائية في اختيار العينة.

٢) أن يكون المتغيرين (المستقل والتابع) مصنفين ضمن المقاييس الكمية.

٣) التوزيع الاعتدالي (الطبيعي) لدرجات المتغيرين (المستقل والتابع).

٤) وجود علاقة خطية بين المتغيرين (المستقل والتابع).

٥) أن يكون تباين المتغير المستقل أكبر من الصفر؛ والغرض منه أن يسهم المتغير المستقل في تفسير التباين في درجات المتغير التابع.

٦) أن يكون متوسط البواقي أو الأخطاء العشوائية يساوي صفرا.

٧) أن تكون الأخطاء العشوائية (البواقي) موزعة توزيعيا طبيعيا (اعتداليا).

معادلة الانحدار الخطي البسيط:

$$\hat{y} = \alpha + \beta x$$

حيث: $\hat{y}$ هي قيمة y **التقديرية** والمناظرة لقيمة x .

α هي ثابت الانحدار، وتمثل طول الجزء المقطوع من المحور الرأسي، وهي تمثل أيضا قيمة y عندما تكون $x = 0$. (وقد تكون موجبة أو سالبة).

β هي (معامل الانحدار) أو (مقدار الميل) أو معدل التغير في y . (وتكون موجبة أو سالبة)

x تمثل قيمة المتغير المستقل الذي يؤثر في المتغير التابع y .

ولتقدير المعلمتين نستخدم طريقة رياضية تسمى (طريقة المربعات الصغرى OLS) والتي تجعل مربعات انحرافات الأخطاء العشوائية المقدرة أقل ما يمكن. وتأخذ شكل المعادلتين التالية:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} \qquad\qquad \alpha = \bar{y} - \beta \bar{x}$$

علاقة معامل الانحدار β بمعامل الارتباط r:

القانون التالي يوضح العلاقة بين معامل الانحدار ومعامل الارتباط، حيث يمكن إيجاد معامل الانحدار إذا تم حساب معامل الارتباط من خلال القانون:

$$\beta = r\frac{s_y}{s_x}$$

حيث: r هو معامل الارتباط بين المتغيرين y،x ويمكن إيجاده من خلال معامل ارتباط بيرسون.

s_x هو الانحراف المعياري لقيم x.

s_y هو الانحراف المعياري لقيم y.

مثال١: البيانات التالية تمثل رأس المال المستثمر (بالمليون ريال) وكمية الإنتاج (بالألف وحدة شهريا):

رأس المال x	5	7	9	10	12	15
كمية الإنتاج y	12	20	22	25	21	19

المطلوب:

(أ) إيجاد معادلة الانحدار مع تفسير العلاقة.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي ١١.

(ج) أوجد الدلالة الإحصائية (اختبار المعنوية).

الحل: من خلال المعادلات السابقة المتعلقة بتحليل الانحدار، نقوم أولا بإنشاء الجدول التالي وتعبئته:

رأس المال x	كمية الإنتاج y	xy	x^2
5	12	60	25
7	20	140	49
9	22	198	81
10	25	250	100
12	21	252	144
15	19	285	225
$\sum x =$	$\sum y =$	$\sum xy =$	$\sum x^2 =$
$\bar{x} =$	$\bar{y} =$		

ثانيا: نقوم بتعويض معادلات تحليل الانحدار على النحو التالي:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} = \frac{6(1185) - (58)(119)}{6(624) - (58)^2} = \frac{7110 - 6902}{3744 - 3364} = \frac{208}{380} \cong 0.547$$

$$\alpha = \bar{y} - \beta\bar{x} \;=\; 19.83 - 0.547(9.67) \;=\; 19.83 - 5.29 \;\cong\; 14.54$$

(أ) إذا معادلة الانحدار هي: $\hat{y} = 14.54 + 0.547\,x \quad = \quad \hat{y} = \alpha + \beta x$

تفسير العلاقة: بما أن معامل الانحدار يساوي (0.547) والذي يمثل التغير في كمية الإنتاج عندما يتغير رأس المال بوحدة واحدة، وبما أن إشارة المعامل موجبة فهذا يعني أن زيادة وحدة واحدة (بمقدار مليون ريال) في رأس المال سيؤدي إلى زيادة كمية الإنتاج بمقدار (0.547) ألف ريال.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي ١١:

$$\hat{y} = 14.54 + 0.547(11) = 20.557$$

(ج) الدلالة الإحصائية (اختبار المعنوية) (تحليل التباين) سيتم بيانها في الدرس القادم.

إعادة مثال رقم ١

البيانات التالية تمثل رأس المال المستثمر (بالمليون ريال) وكمية الإنتاج (بالألف وحدة شهريا):

رأس المال x	5	7	9	10	12	15
كمية الإنتاج y	12	20	22	25	21	19

المطلوب:

(أ) إيجاد معادلة الانحدار مع تفسير العلاقة.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي ١١.

(ج) أوجد الدلالة الإحصائية (اختبار المعنوية).

الحل: من خلال المعادلات السابقة المتعلقة بتحليل الانحدار، نقوم أولا بإنشاء الجدول التالي وتعبئته:

رأس المال x	**كمية الإنتاج y**	xy	x^2
5	12		
7	20		
9	22		
10	25		
12	21		
15	19		
$\sum x =$	$\sum y =$	$\sum xy =$	$\sum x^2 =$
$\bar{x} =$	$\bar{y} =$		

ثانيا: نقوم بتعويض معادلات تحليل الانحدار على النحو التالي:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} =$$

$$\alpha = \bar{y} - \beta\bar{x} \quad =$$

(أ) إذا معادلة الانحدار هي: $\hat{y} = \alpha + \beta x \quad = \quad \hat{y} = \quad + \quad x$

تفسير العلاقة: بما أن .... يساوي (   ) والذي يمثل التغير في كمية الإنتاج عندما يتغير رأس المال ....، وبما أن إشارة المعامل موجبة فهذا يعني أن زيادة وحدة واحدة (بمقدار مليون ريال) في رأس المال سيؤدي إلى .... كمية الإنتاج بمقدار (   ) ألف ريال.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي ١١:

$$\hat{y} = 14.54 + 0.547(\quad) =$$

مثال٢: البيانات التالية تمثل رأس المال المستثمر (بالمليون ريال) وكمية الإنتاج (بالألف وحدة شهريا):

رأس المال x	4	8	4	12	13	15
كمية الإنتاج y	5	12	6	15	18	22

المطلوب:

(أ) إيجاد معادلة الانحدار مع تفسير العلاقة.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي 14.

(ج) أوجد الدلالة الإحصائية (اختبار المعنوية).

الحل: من خلال المعادلات السابقة المتعلقة بتحليل الانحدار، نقوم أولا بإنشاء الجدول التالي وتعبئته:

رأس المال x	**كمية الإنتاج y**	xy	x^2
4	5		
8	12		
4	6		
12	15		
13	18		
15	22		
$\sum x =$	$\sum y =$	$\sum xy =$	$\sum x^2 =$
$\bar{x} =$	$\bar{y} =$		

ثانيا: نقوم بتعويض معادلات تحليل الانحدار على النحو التالي:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} =$$

$$\alpha = \bar{y} - \beta\bar{x} \quad =$$

(أ) إذا معادلة الانحدار هي:

تفسير العلاقة: بما أن .... يساوي ( ) والذي يمثل التغير في كمية الإنتاج عندما يتغير رأس المال ....، وبما أن إشارة المعامل موجبة فهذا يعني أن زيادة وحدة واحدة (بمقدار مليون ريال) في رأس المال سيؤدي إلى .... كمية الإنتاج بمقدار ( ) ألف ريال.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي 14:

مثال٣: البيانات التالية تمثل رأس المال المستثمر (بالمليون ريال) وكمية الإنتاج (بالألف وحدة شهريا):

رأس المال x	6	12	20	22	25	29
كمية الإنتاج y	10	15	21	28	33	30

المطلوب:

(أ) إيجاد معادلة الانحدار مع تفسير العلاقة.

(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي 21.

(ج) أوجد الدلالة الإحصائية (اختبار المعنوية).

الحل: من خلال المعادلات السابقة المتعلقة بتحليل الانحدار، نقوم أولا بإنشاء الجدول التالي وتعبئته:

رأس المال x	**كمية الإنتاج y**	xy	x^2
6	10		
12	15		
20	21		
22	28		
25	33		
29	30		
$\sum x =$	$\sum y =$	$\sum xy =$	$\sum x^2 =$
$\bar{x} =$	$\bar{y} =$		

ثانيا: نقوم بتعويض معادلات تحليل الانحدار على النحو التالي:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} =$$

$$\alpha = \bar{y} - \beta\bar{x} \ =$$

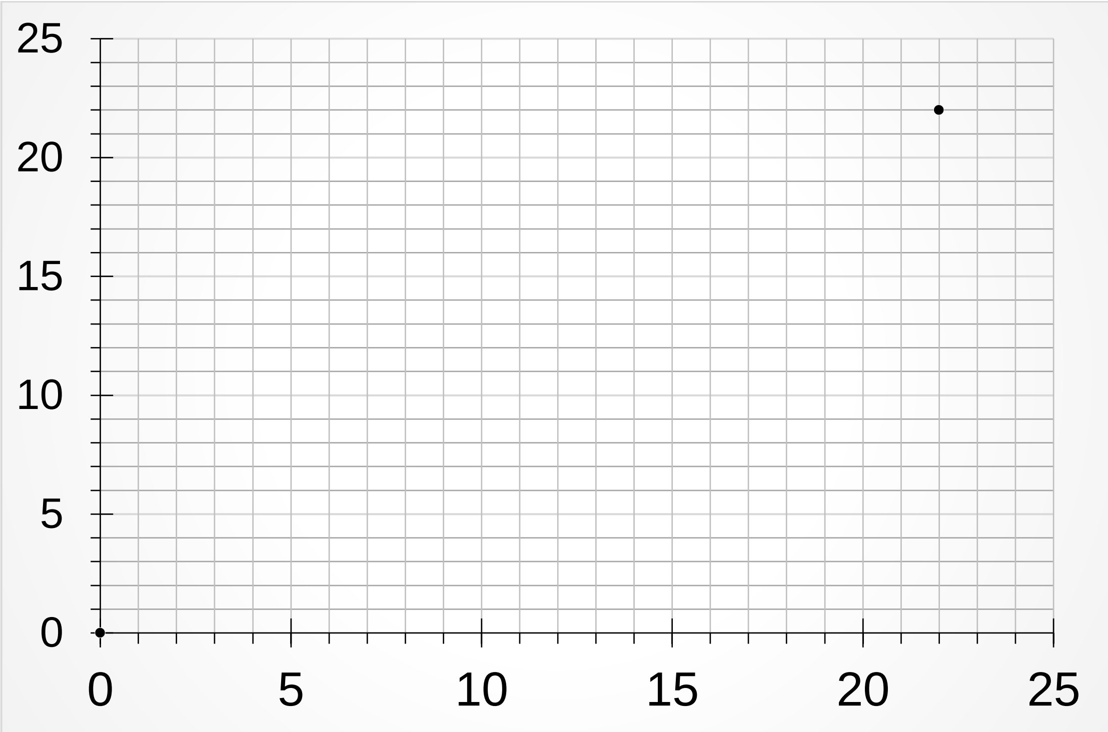
(أ) إذا معادلة الانحدار هي:

تفسير العلاقة: بما أن .... يساوي (   ) والذي يمثل التغير في كمية الإنتاج عندما يتغير رأس المال ....، وبما أن إشارة المعامل موجبة فهذا يعني أن زيادة وحدة واحدة (بمقدار مليون ريال) في رأس المال سيؤدي إلى .... كمية الإنتاج بمقدار (   ) ألف ريال.

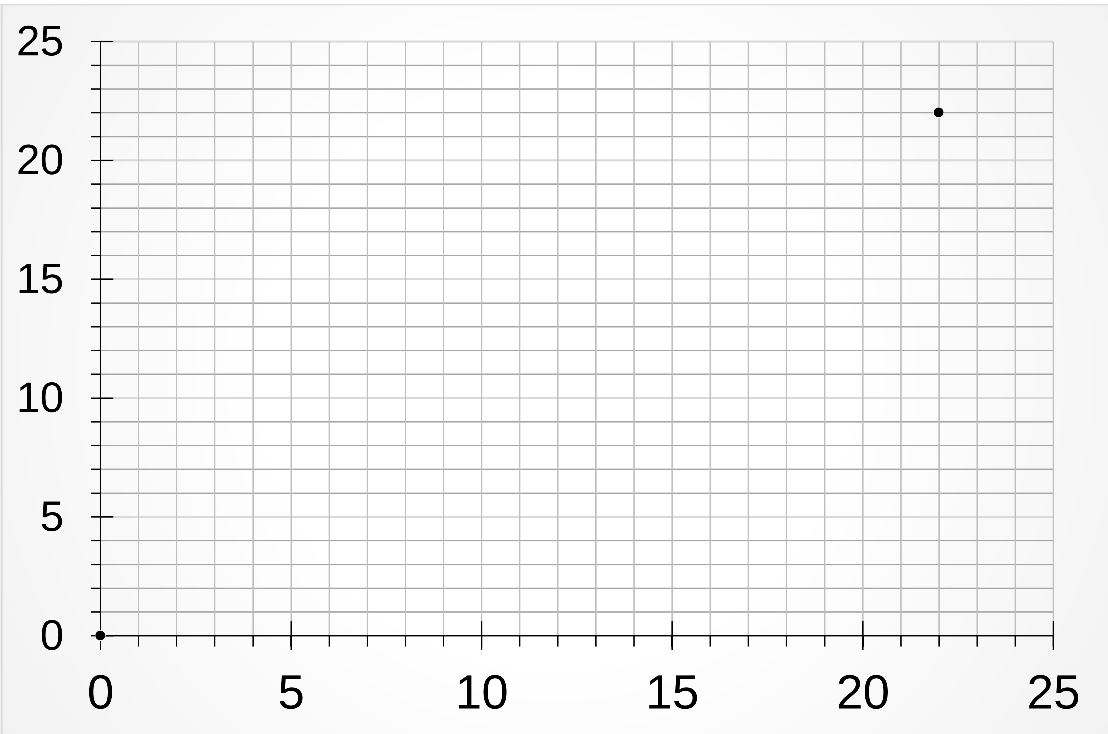
(ب) تقدير كمية الإنتاج في حال كان رأس المال يساوي 21:

تمرينات على خط الانحدار

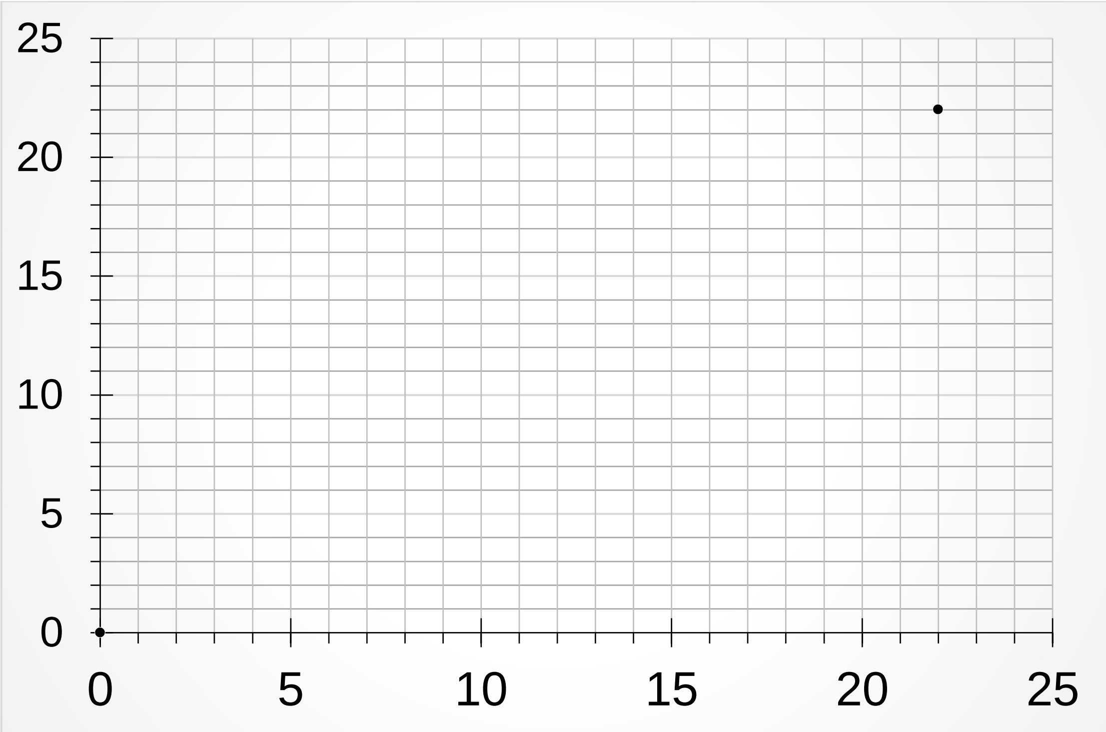
ارسم خط الانحدار: β = ١ ، α = صفر



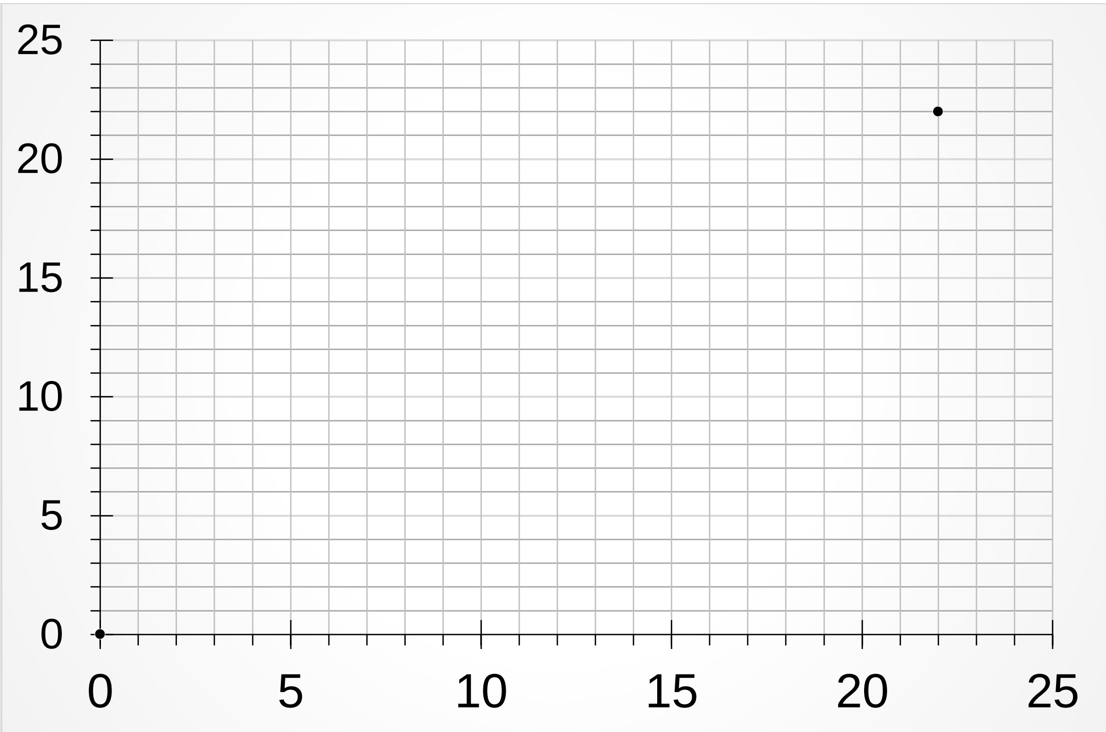
ارسم خط الانحدار: β = ٢ ، α = ٥



ارسم خط الانحدار: β = ٢ ، α = ...



ارسم خط الانحدار: β = ... ، α = ٧



تمرينات على خط الانحدار من خلال جدول الأرقام

المحاضرة السابعة

تحليل التباين

تحليل التباين الأحادي الاتجاه (ANOVA)

تمهيد: يُنسب ظهور تحليل التباين إلى العالم (رونالد فيشر) الذي وضع منطق وأسس تحليل التباين وذلك في عام ١٩١٨م وتبعه (Sedecor) عام ١٩٣٤م الذي اقترح إحصاءة اختبار تحليل التباين، وأطلق عليها اسم اختبار (F) نسبة لفيشر، وتحليل التباين هو البحث عن مكونات الاختلاف أو التباين، وهذا يعني دراسة مكونات الاختلاف بين مجموعة من الأفراد في ظاهرة معينة. ويعتبر تحليل التباين الأحادي الاتجاه أو تحليل التباين في اتجاه واحد نفس المصطلح امتدادا لاختبار الفرق بين المتوسطين، لأن تحليل التباين في اتجاه واحد هو اختبار الفروق بين متوسطات ثلاثة مجتمعات فأكثر.

ويهدف التحليل إلى بيان ما إذا كانت متوسطات المجتمعات متساوية تقريباً وأن أية اختلافات بينها تعزى للصدفة ويمكن توقعها، أم أن المتوسطات مختلفة (غير متساوية) والفروق بينها جوهرية أو ذات دلالة.

وسمي بتحليل التباين الأحادي الاتجاه لأنه يعتمد في تحليله على تأثير متغير مستقل واحد (له عدة مستويات أو تصنيفات) على متغير تابع واحد.

وأسلوب تحليل التباين هو أسلوب إحصائي الهدف منه تقسيم مجموع مربعات الانحرافات الكلي إلى مكوناته الأساسية ومن ثم إرجاع كل من هذه المكونات إلى سببه.

والتباين الذي نسعى إلى تحليله والكشف عنه يكون في درجات المتغير التابع فقط، وهذا طبقا لاختلافات مجموعات المتغير المستقل، فمثلا لدينا متغيرين المعدل والتخصص العلمي، حيث المعدل هو المتغير التابع والتخصص هو المتغير المستقل، وبالتالي سنسعى إلى تحليل معدلات الطلبة بالنسبة لمستويات المتغير المستقل والمتمثل في التخصصات العلمية (علم الاجتماع – علم النفس – علم التربية..)

مكونات تحليل التباين

أبسط مكونين لتحليل التباين هما **التباين بين المجموعات** (الاختلاف بين الأفراد) و**التباين داخل المجموعات** (الاختلاف داخل الأفراد)، فمثلا إذا كان لدينا تخصصين لطلبة هما التخصص الأدبي والتخصص العلمي) فإن الاختلاف بين درجات التخصصين الأدبي والعلمي يعتبر تباين بين المجموعات، أما الاختلاف بين درجات الطلبة فيما بينهم في التخصص الأدبي أو التخصص العلمي، فهذا يعتبر تباين داخل المجموعات.

ويمكن التوسع في شرح تحليل التباين وأنواعه وأهميته في فصل مستقل لاحقا، لكن هنا سنقوم باستخدام تحليل التباين على الدلالة الإحصائية لتحليل الانحدار الخطي البسيط الذي تناولنا في الفصل السابق.

اختبار معنوية الانحدار الخطي البسيط:

لتحديد مدى صحة ودقة تقديرات العلاقة بين المتغيرين المستقل والتابع، والتي يتم تقديرها من خلال معادلة الانحدار، والتي تفترض أن العلاقة بين المتغيرين خطية فلابد من إجراء اختبار لهذه التقديرات؛ حتى يتبين لنا هل فعلاً العلاقة بين المتغيرين خطية؟ أم أن معادلة الانحدار المتحصل عليها لا تعبر عن التعبير الصحيح عن العلاقة بين المتغيرين، ويجب التحقق من الشكل الدقيق للعلاقة بينهما.

ويُستخدم تحليل التباين Analysis of Variance ويتم اختصاره بـ [ANOVA] في إجراء اختبار معنوية معادلة الانحدار، ويعتمد تحليل التباين هذا على تحليل مجموع مربعات انحرافات القيم للمتغير التابع y عن وسطها الحسابي، كالآتي:

<table>
<tr><th colspan="5">ANOVA</th></tr>
<tr><th>المصدر</th><th>درجات الحرية
$(d.f)$</th><th>مجموع المربعات
Sum of Squares</th><th>متوسط المربعات
Mean of Squares</th><th>إحصائية الاختبار
$\hat{F}$</th></tr>
<tr><td>معادلة الانحدار
Regression</td><td>$(d.f)_R$
[عدد المتغيرات المستقلة]</td><td>SSR</td><td>$\text{MSR} = \frac{\text{SSR}}{(d.f)_R}$</td><td>$\hat{F} = \frac{\text{MSR}}{\text{MSE}}$</td></tr>
<tr><td>الخطأ العشوائي (البواقي)
Error</td><td>$(d.f)_E$
[تساوي n-2]</td><td>SSE</td><td>$\text{MSE} = \frac{\text{SSE}}{(d.f)_E}$</td><td></td></tr>
<tr><td>المجموع
Total</td><td>$(d.f)_t$</td><td>SST
[SSR + SSE]</td><td></td><td></td></tr>
</table>

وللاختصار فإن عمود مجموع المربعات يشتمل على:

- مجموع مربعات الانحدار SSR، وهو يشير إلى مجموع مربعات انحرافات القيم المقدرة $\hat{y}$ عن الوسط الحسابي، ويأخذ شكل المعادلة: $\text{SSR} = \sum(\hat{y} - \bar{y})^2$.
- مجموع مربعات الأخطاء العشوائية (البواقي) ، وهو يشير إلى مجموع مربعات انحرافات القيم عن خط الانحدار (مجموع مربعات الأخطاء العشوائية أو البواقي)، ويأخذ شكل المعادلة: $\text{SSE} = \sum(y - \hat{y})^2$. ويمكن التعبير عن هذه المعادلة بمعادلة أخرى تسهل عملية الحساب: $\text{SSE} = \sum y^2 - \beta \sum xy - \alpha \sum y$

- مجموع المربعات الكلي ، وهو يشير إلى مجموع مربعات انحرافات القيم عن الوسط الحسابي، ويأخذ شكل المعادلة: $SST = \sum(y-\bar{y})^2$ ويمكن التعبير عن هذه المعادلة بمعادلة أخرى تسهل عملية الحساب: $SST = \sum y^2 - n\bar{y}^2$، وأيضا يمكن إثبات أن: $SST = SSR + SSE$.

ويتم اختبار معنوية معادلة الانحدار كما يلي:

الفرض الصفري (فرض العدم): [عدم وجود علاقة انحدار خطي]	$H_0\colon\ \beta = 0$
الفرض البديل: [وجود علاقة انحدار خطي]	$H_1\colon\ \beta \neq 0$

ثم نقوم بمقارنة $\hat{F}$ المحسوبة في جدول تحليل التباين ANOVA بـ $\hat{F}$ الجدولية، فإذا كانت المحسوبة أكبر من الجدولية فإننا نرفض الفرض الصفري ونقبل البديل، وإلا العكس.

بناء على مثال رقم ١ في درس (تحليل الانحدار)، حصلنا على المعطيات التالية:

$\sum x^2 = 624$	$\sum xy = 1185$	$\sum y = 119$	$\sum x = 58$
$\beta = 0.547$	$\alpha = 14.54$	$\bar{y} = 19.83$	$\bar{x} = 9.67$

وبناء على المعادلات السابقة، نقوم بحساب التالي:

$$SST = \sum y^2 - n\bar{y}^2 = 2455 - 6(19.83) \cong 94.833$$
$$SSE = \sum y^2 - \beta\sum xy - \alpha\sum y = 2455 - 0.547(1185) - 14.542(119) \cong 76.307$$
$$SSR = SST - SSE = 18.526$$

مما سبق يمكن تكوين جدول تحليل التباين كالتالي:

ANOVA				
المصدر	درجات الحرية $(d.f)$	مجموع المربعات Sum of Squares	متوسط المربعات Mean of Squares	إحصائية الاختبار $\hat{F}$
معادلة الانحدار Regression	1	18.526	$MSR = \frac{18.526}{1}$	$\hat{F} = \frac{18.526}{19.077} = 0.971$
الخطأ العشوائي (البواقي) Error	4	76.307	$MSE = \frac{76.307}{4}$	
المجموع Total	5	94.833		

الدلالة الإحصائية: بما أن $\hat{F}$ المحسوبة (0.971) أقل من $\hat{F}$ الجدولية التي تساوي (7.71)، إذا نقبل الفرض الصفري ونرفض البديل، وبالتالي لا توجد علاقة انحدار خطي معنوية بين المتغيرين.

المحاضرة الثامنة

السلاسل الزمنية

تمهيد: السلسلة الزمنية هي مجموعة القياسات المسجلة لمتغير واحد أو أكثر، مرتبة حسب زمن وقوعها. ورياضيا نقول إن متغير الزمن (t) هو المتغير المستقل، والقيم المناظرة له هي المتغير التابع (y)، وإن كل قيمة في الزمن يقابلها قيم للمتغير التابع. وبالتالي فإن دالة في الزمن [أي متأثرة بالزمن].

t	٢٠١٩	٢٠٢٠	٢٠٢١	٢٠٢٢	٢٠٢٣	٢٠٢٤
Y	23	20	26	28	33	34

وتعتبر السلاسل الزمنية من الأمور الطبيعية والواجبة التي ينبغي للحكومات والمؤسسات والشركات التجارية منها والتعليمية وغيرها لاستخدامها؛ من أجل التخطيط لمستقبلها ولتحقيق أهدافها العامة والخاصة. وتعتبر السلاسل الزمنية من أهم أساليب التنبؤ حول المستقبل من خلال وقائع الأمس واليوم.

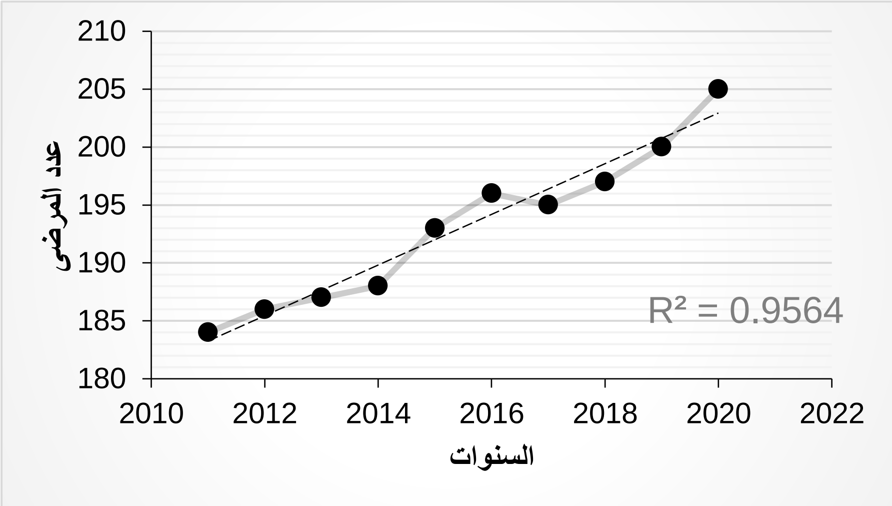
تعريف السلسلة الزمنية:

هي عبارة عن مجموعة من المشاهدات الإحصائية التي تصف الظاهرة مع مرور الزمن. أو هي: البيانات الإحصائية التي تُجمع أو تُشاهد أو تُسجل لفترات متتالية من الزمن.

وقد تكون السلسة الزمنية بالأرقام المطلقة، أو قد تكون بالقيم النسبية مثل تلك الجداول التي تبين معدلات الزيادة الطبيعية للسكان، أو قد تكون بالمتوسطات. ومن أمثلة هذه السلاسل الزمنية:

مرضى عيادات الأسنان المترددين شهريا.	عدد العاطلين سنويا عن العمل.
معدلات الإنجاب السنوية.	درجات حرارة المريض كل ساعة لمدة يوم واحد.
المبيعات اليومية في مركز لبيع الكتب لمدة شهر.	الإنتاج الشهري لمدة سنة في إحدى الشركات.

الرسم البياني للسلاسل الزمنية:



عدد المرضى	السنة
١٨٤	٢٠١١
١٨٦	٢٠١٢
١٨٧	٢٠١٣
١٨٨	٢٠١٤
١٩٣	٢٠١٥
١٩٦	٢٠١٦
١٩٥	٢٠١٧
١٩٧	٢٠١٨
٢٠٠	٢٠١٩
٢٠٥	٢٠٢٠

تحليل السلسة الزمنية:

لغرض فهم السلسلة الزمنية لابد من تحليلها إلى عناصرها ومركباتها الأساسية بطريقة تمكننا من معرفة تطور الظاهرة مع الزمن، والتنبؤ بمعالمها خلال الفترات المقبلة؛ لتُتخذ أساسا للتخطيط الاقتصادي أو الإداري الطويل الأجل. وتتألف السلسلة الزمنية من **أربعة عناصر أساسية**، هي:

١) الاتجاه العام [التحركات طويلة المدى]: وهي تشير إلى الاتجاه العام الذي يظهر به الشكل البياني للسلسلة الزمنية على فترة طويلة من الزمن. وفي الشكل البياني السابق: تمثل النقاط المتصلة الشكل البياني للسلسلة الزمنية، في حين يمثل الخط المتقطع خط الاتجاه العام. وقد سبق دراسة مثل في تحليل الانحدار، وأن أفضل منحنى له هو ما يكون بطريقة المربعات الصغرى.

٢) التغيرات الموسمية: وهي تشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتقابلة خلال السنوات المتتالية، مثل الزيادة المفاجئة في مبيعات المحلات في الفترة السابقة للأعياد.

٣) التغيرات الدورية: وهي تشير إلى الذبذبات طويلة المدى حول خط (أو منحنى) الاتجاه العام، وهذه الدورات قد تكون وقد لا تكون على فترات، بمعنى أنها قد تتبع وقد لا تتبع نفس النمط بعد كل فترة زمنية متساوية، مثل دورات الأعمال التي تمثل فترات الرخاء، الركود الكساد.

٤) التغيرات العشوائية أو الفجائية: مثل الفيضانات، الاضطرابات، الانتخابات، ونحوها.

وفي هذا الدرس سنوضح الاتجاه العام فقط على النحو التالي:

الاتجاه العام:

تغيرات الاتجاه العام تعني الزيادة أو الانخفاض طويل الأجل في البيانات عبر الزمن، ويتم التعرف على ذلك من خلال تمثيل السلسلة الزمنية بيانيا، فنحصل بالتالي على خط بياني واتجاه خط السلسلة الزمنية صعودا أو هبوطا، ويدل ذلك على تزايد الظاهرة أو تناقصها مع مرور الزمن، ويمكن تقدير الاتجاه العام بعدة طرق منها:

طريقة الانتشار (التمهيد باليد).

طريقة المتوسطات المتحركة.

طريقة متوسط نصف السلسة.

طريقة المربعات الصغرى (وهي أكثر الطرق دقة، وهي ما سنوضحه في هذا الدرس).

طريقة المربعات الصغرى:

تعتبر طريقة المربعات الصغرى أكثر الطرق دقة لحساب خط الاتجاه العام، وذلك من خلال استخدام أسلوب الانحدار الخطي البسيط المعتمد على طريقة المربعات الصغرى التي تجعل مجموع مربعات انحرافات القيم المقدرة عن القيم الفعلية أقل ما يمكن، وذلك من خلال المعادلة التالية:

$$\hat{y} = \alpha + \beta t$$

حيث: $\hat{y}$ هي قيمة y **التقديرية** والمناظرة لقيمة الفترة الزمنية t .

α هي ثابت الانحدار، وتمثل طول الجزء المقطوع من المحور الرأسي، وهي تمثل أيضا قيمة y عندما تكون $t = 0$.

β هي (معامل الانحدار) أو (مقدار الميل) أو معدل التغير في y . (وتكون موجبة أو سالبة)

t تمثل الفترة الزمنية (المتغير المستقل) الذي يؤثر في المتغير التابع y .

تنبيه: نقوم بتحديد السنوات المشاهدة على النحو التالي: $t_1 = 0,\ t_2 = 1, \ldots$

ولتسهيل المعادلة يمكن استبدال رمز الفترة الزمنية (t) بـ (x)، لتشابه هذه المعادلة مع معادلة الانحدار الخطي البسيط، لتصبح المعادلة كالتالي: $\hat{y} = \alpha + \beta x$

ولتقدير المعلمتين نستخدم طريقة رياضية تسمى (طريقة المربعات الصغرى OLS) والتي تجعل مربعات انحرافات الأخطاء العشوائية المقدرة أقل ما يمكن. وتأخذ شكل المعادلتين التالية:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} \qquad \alpha = \bar{y} - \beta\bar{x}$$

مثال توضيحي للفرق بين الانحدار الخطي وبين السلسلة الزمنية:

الانحدار الخطي	رأس المال x	6	12	20	22	25	29
	كمية الإنتاج y	10	15	21	28	33	30

السلسلة الزمنية	السنوات t	2019	2020	2021	2022	2023	2024
	كمية الإنتاج y	10	15	21	28	33	30

مثال١: البيانات التالية تمثل الدخل السنوي بملايين الريالات لإحدى الشركات الكبرى:

السنوات	2020	2021	2022	2023	2024
الدخل	3	4	5	7	6

المطلوب:

(أ) إيجاد معادلة الاتجاه العام مع تفسير العلاقة.

(ب) قدر قيمة الدخل لعام 2025.

(ج) أوجد الدلالة الإحصائية (اختبار المعنوية).

الحل: من خلال المعادلات السابقة المتعلقة بتحليل الانحدار، نقوم أولا بإنشاء الجدول التالي وتعبئته:

السنوات	الزمن	الدخل	xy	x^2
2020	0	3		
2021	1	4		
2022	2	5		
2023	3	7		
2024	4	6		
	$\sum x =$	$\sum y =$	$\sum xy =$	$\sum x^2 =$
	$\bar{x} =$	$\bar{y} =$		

ثانيا: نقوم بتعويض معادلات تحليل الانحدار على النحو التالي:

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \left(\sum x\right)^2} =$$

$$\alpha = \bar{y} - \beta x \quad =$$

(أ) إذا معادلة الاتجاه العام هي:

تفسير العلاقة: بما أن معامل الانحدار β يساوي (   ) والذي يمثل التغير في الدخل عندما تتغير الفترة الزمنية بفترة واحدة، وبما أن إشارة المعامل β موجبة فهذا يعني أن زيادة فترة واحدة (بمقدار سنة واحدة) سيؤدي إلى .... قيمة الدخل بمقدار (   ) مليون ريال.

(ب) تقدير قيمة الدخل في عام 2025:

المحاضرة التاسعة

الأرقام القياسية

تمهيد: يعود استخدام الأرقام القياسية إلى أكثر من قرنين من الزمن، حيث استخدمها الإحصائي الإيطالي كارلي عام ١٧٦٤م لمقارنة الأسعار في إيطاليا لسنة ١٧٥٠م بالأسعار في سنة ١٥٠٠م. ثم شاع استخدامها بصورة أوسع منذ ذلك الوقت.

وتستخدم الأرقام القياسية في التطبيقات الإحصائية في مجال الدراسات الاقتصادية حيث يمكن من خلالها التعرف على الأحوال الاقتصادية للدول المختلفة من خلال دراسة التغيرات الاقتصادية في البلد أو البلدان قيد الدراسة، للمساعدة في التنبؤ بما يمكن أن يحدث للمتغيرات المختلفة في المستقبل. كما تستخدم لقياس ظواهر متعددة مثل مقارنة أسعار السلع الغذائية في سنة محددة بسنة أخرى سابقة أو مقارنة إنتاج قطاع اقتصادي معين في دولة ما بنظيره في دولة أخرى؛ للوقوف على التطور الذي طرأ على إنتاج هذا القطاع عبر الزمن.

ولم تعد تطبيقات الأرقام القياسية حكرا على الاقتصاديين فقط بل أصبحت وسيلة في أيدي المهتمين في العلوم الاجتماعية والإدارية والزراعية لعمل المقارنات وقياس التغيرات، وهناك أرقام قياسية في ميادين مختلفة مثل الرقم القياسي لأسعار الجملة والرقم القياسي للصادرات والرقم القياسي للاستيرادات، كما تؤخذ أرقام قياسية للإنتاج الزراعي والإنتاج الصناعي وتكاليف المعيشة.

تعريف الرقم القياسي:

الرقم القياسي هو **مؤشر إحصائي (رقم نسبي %) يُستخدم في قياس التغير النسبي الذي يطرأ على ظاهرة من الظواهر الاقتصادية أو الاجتماعية،** فهو يُستخدم لقياس التغير في أسعار السلع أو في حجم إنتاجها أو في كميات المبيعات منها أو في حجم السكان أو أجور العمال، وذلك وفقاً لأساس معين سواء كان هذا الأساس فترة زمنية معينة أو مكاناً معيناً. وتسمى أحياناً المجموعة من الأرقام القياسية لسنوات أو أماكن مختلفة، وما إلى ذلك بالسلسلة القياسية.

فترة الأساس:

الأساس هو فترة زمنية معينة أو مكان معين أو أي خاصية أخرى مثل الدخل، الوظيفة، وغير ذلك، وعادة تكون فترة الأساس فترة سابقة للفترة التي نريد مقارنتها، وفي حالات نادرة جداً قد تكون فترة الأساس فترة لاحقة لفترة المقارنة.

ويجب أن تمتاز فترة الأساس بما يلي: الاستقرار الاقتصادي، وخلوها من العوامل المؤثرة على الأسعار (مثل الحروب)، وأن تكون بعيدة عن سنوات المقارنة.

أما عند اختيار مكان الأساس، فلابد أن يكون لهذا المكان أهمية خاصة وأن يكون مركزاً أساسياً لإنتاج السلعة المراد استخراج الرقم القياسي لها.

الأرقام القياسية للأسعار Price Index Numbers:

تعتبر الأرقام القياسية للأسعار من أهم أنواع الأرقام القياسية وأكثرها شيوعاً، فهي (أي الأرقام القياسية للأسعار) تُساهم في قياس التغير في المستوى العام للأسعار أو التغير في تكاليف المعيشة في فترة زمنية معينة مقارنة بفترة زمنية أخرى.

ومن أشهر الأرقام القياسية للأسعار ما يلي:

- مؤشر أسعار المستهلكين Consumer Price Index ويرمز له (CPI).
- مخفض الناتج القومي الإجمالي Gross National Product Deflator.
- مؤشر أسعار المنتجين Producer Price Index ويرمز له (PPI).
- مخفض الناتج المحلي الإجمالي Gross Domestic Product Deflator.
- مؤشر أسعار الأسهم.

ويختلف مؤشر أسعار المستهلكين عن مخفض الناتج القومي في أن مؤشر أسعار المستهلكين يتم حسابه باستخدام السلع والخدمات التي لها علاقة بالمستهلك العادي كالمواد الغذائية والملابس والسكن والمواصلات والاتصالات والرعاية الصحية والتعليم، وما إلى ذلك، بينما يتم حساب مخفض الناتج القومي الإجمالي باستخدام جميع السلع والخدمات التي يستهلكها النظام الاقتصادي في الدولة من مستهلكين عاديين وقطاع خاص وقطاع حكومي.

ويهتم النظام الاقتصادي السعودي بنشر الأرقام القياسية للأسعار وتكاليف المعيشة على شكل تقارير شهرية، ومن هذه الأرقام ما يلي:

الرقم القياسي لتكاليف المعيشة لجميع السكان: ويشمل المواد الغذائية، السكن، وتوابعه، الأقمشة والملابس الأثاث المنزلي، الرعاية الطبية، النقل والاتصالات، التعليم والترفيه، النفقات والخدمات الأخرى.

الرقم القياسي لأسعار الجملة: ويشمل المواد الغذائية المشروبات، المواد الخام، الوقود المعدني وزيوت التشحيم، الدهون والزيوت الحيوانية والنباتية، الكيماويات والمواد ذات الصلة، السلع المصنعة مصنفة حسب المادة الآلات ومعدات النقل والاتصالات، التعليم والترفيه النفقات والخدمات الأخرى.

دور الأرقام القياسية في حساب معدلات التضخم:

المقصود بالتضخم هو الارتفاع المستمر في المستوى العام للأسعار. والذي على ضوئه تنخفض القيمة الشرائية للوحدة النقدية (الريال مثلا)، وتقوم الجهات الاقتصادية في الدول باستخدام الأرقام القياسية للأسعار لإيجاد معدلات التضخم السنوية، وفي معظم الأحيان يستخدم مؤشر أسعار المستهلكين (CPI) لسنتين متتاليتين لحساب معدل التضخم السنوي في السنة الأخيرة وذلك من خلال العلاقة التالية:

$$i_r = \frac{CPI_r - CPI_{r-1}}{CPI_{r-1}} \times 100$$

حيث i_r هو معدل التضخم للسنة r [وهو نسبة مئوية].

CPI_r هو مؤشر أسعار المستهلكين للسنة r.

CPI_{r-1} هو مؤشر أسعار المستهلكين للسنة r-1 [أي السنة السابقة].

فمثلا إذا افترضنا أن مؤشر أسعار المستهلكين في المملكة لسنة 2020 هو 120 ومؤشر أسعار المستهلكين لسنة 2021 هو 123، فإن معدل التضخم في سنة 2021 يكون:

$$i_r = \frac{CPI_r - CPI_{r-1}}{CPI_{r-1}} \times 100 \quad = \quad i_r = \frac{123 - 120}{120} \times 100 = 2.5\%$$

أي أن معدل التضخم في سنة 2021 يساوي: 2.5%

فوائد الأرقام القياسية واستعمالاتها:

تستخدم الأرقام القياسية عادة لقياس التغير الذي يطرأ على الحياة بمجملها بشكل عام والجوانب الاقتصادية بشكل خاص، كما تساعد الأرقام القياسية على تحليل العوامل التي تساهم في تغير الظاهرة فتبين مدى مساهمة كل من هذه العوامل في إحداث التغير الكلي، وتستخدم كذلك في الرقابة على تنفيذ الخطط.

منسوب السعر لسلعة واحدة (ظاهرة بسيطة) [الرقم القياسي لسلعة واحدة]:

يسمى الرقم القياسي للسعر بـ منسوب السعر (حيث يُمثل هذا الرقم القياسي التغير في سعر السلعة أو الخدمة في سنة معينة مقارنة بسنة الأساس)، ويُرمز له بالرمز: P_r ويأخذ شكل المعادلة التالية:

$$P_r = \frac{P_1}{P_0} \times 100$$

حيث: P_r هو منسوب السعر للسلعة [وهو نسبة مئوية].

P_0 هو السعر خلال فترة الأساس.

P_1 هو السعر خلال فترة المقارنة.

مثال: البيانات التالية تمثل سعر سلعة معينة من الفترة ٢٠٢٠ إلى ٢٠٢٤م. المطلوب: إيجاد منسوب السعر لهذه السلعة، باعتبار سنة **٢٠٢٠م** هي **سنة الأساس**، مع تفسير النتائج المتحصلة.

السنة	**2020**	2021	2022	2023	2024
السعر (بالريال)	25	30	24	32	36

الحل: يتم مباشرة حساب قيمة منسوب السعر لهذه السلع من خلال المعادلة السابقة كما يلي:

السنة	**سعر السلعة**	**الطريقة**	**منسوب السعر (الرقم القياسي)**	**تفسير النتائج**
2020	25	$\frac{25}{25} \times 100$	100%	سنة الأساس، ودائما تكون 100
2021	30	$\frac{30}{25} \times 100$	120%	زيادة في السعر بنسبة 20%
2022	24	$\frac{24}{25} \times 100$	96%	انخفاض في السعر بنسبة 4%
2023	32	$\frac{32}{25} \times 100$	128%	زيادة في السعر بنسبة 28%
2024	36	$\frac{36}{25} \times 100$	144%	زيادة في السعر بنسبة 44%

منسوب السعر لمجموعة من السلع (ظاهرة معقدة):

الرقم القياسي السابق يُوضح منسوب السعر لسلعة واحدة، إلا أن كثيرا من الحالات تكون أكثر تعقيدا، فقد يكون لدينا عدة سلع متغيرة ونرغب في حساب منسوب السعر (أو الرقم القياسي) لها، ففي حالة استخراج الرقم القياسي لمثل هذا الوضع فإنه يدخل في الحساب جميع قيم السلع التي تتألف منها الظاهرة. ويتم ذلك من خلال استخدام الطرق التالية:

- الوسط الحسابي لمناسيب الأسعار.
- الرقم القياسي التجميعي البسيط للأسعار.
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير).
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش).
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر).

الوسط الحسابي لمناسيب الأسعار:

عند توفر عدة سلع، فإن الوسط الحسابي لمناسيب الأسعار هو بإيجاد منسوب السعر (الرقم القياسي) لكل سلعة بشكل مستقل، ثم يتم حساب الوسط الحسابي لهذه المناسيب. ويأخذ شكل المعادلة:

$$\frac{\sum P_r}{n}$$

حيث: $\sum P_r$ مجموع مناسيب الأسعار.

n عدد السلع.

الرقم القياسي التجميعي البسيط للأسعار:

يُرمز لهذا الرقم بـ I_s ويتم حسابه من خلال المعادلة التالية:

$$I_s = \frac{\sum P_1}{\sum P_0} \times 100$$

حيث: $\sum P_1$ مجموع أسعار السلع والخدمات خلال فترة المقارنة.

$\sum P_0$ مجموع أسعار السلع والخدمات خلال فترة الأساس.

وتكمن مشكلة الرقم القياسي التجميعي البسيط للأسعار في أنه لا يعطي للكميات المستهلكة من السلع والخدمات أوزاناً، وبالتالي قد لا يكون دقيقاً عندما يكون هناك تباينا في الكميات المستهلكة.

الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير):

ويُرمز له بالرمز I_r، وهذا الرقم يُعبر عن أثر التغير في السعر كما لو بقيت الكميات المشتراة في سنة الأساس هي نفسها في سنة المقارنة، ويتم حسابه بنفس الطريقة السابقة مع ترجيح وزن كل سعر بكميته Q_0 خلال فترة الأساس، ويتم ذلك من خلال المعادلة التالية:

$$I_r = \frac{\sum P_1 Q_0}{\sum P_0\, Q_0} \times 100$$

حيث: $\sum P_1 Q_0$ مجموع أسعار السلع والخدمات خلال فترة المقارنة مرجحة بكيمات فترة الأساس.

$\sum P_0 Q_0$ مجموع أسعار السلع والخدمات خلال فترة الأساس مرجحة بكميات فترة الأساس.

ويُفضل استخدام هذه الطريقة عند حساب مؤشر أسعار المستهلكين (CPI)؛ وذلك لتوفير الجهد والوقت والمال؛ لأن كمية فترة الأساس ثابتة عند إيجاد رقم لاسبير لأي فترة لاحقة لفترة الأساس.

الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش):

ويُرمز له بالرمز I_p، وهذا الرقم يُعبر عن أثر التغير في السعر كما لو أن الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الأساس. وتختلف طريقة حساب هذا الرقم من حيث إنه يُرجح كل سعر بكميته Q_1 خلال فترة المقارنة، ويتم ذلك من خلال المعادلة التالية:

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0\, Q_1} \times 100$$

حيث: $\sum P_1 Q_1$ مجموع أسعار السلع والخدمات خلال فترة المقارنة مرجحة بكميات فترة المقارنة.

$\sum P_0 Q_1$ مجموع أسعار السلع والخدمات خلال فترة الأساس مرجحة بكميات فترة المقارنة.

والمشكلة الرئيسة في هذه الطريقة هي الحاجة لتحديد الكميات المستهلكة من كل سلعة إذا تم حساب هذا الرقم سنويا. بينما يتم الاكتفاء بحساب كميات السلع في سنة واحدة (سنة الأساس) عند حساب (رقم لاسبير).

الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر) أو (الرقم الأمثل):

ويُرمز له بالرمز I_f، وهو عبارة عن الوسط الهندسي لكل من رقمي لاسبير وباش، [أي الجذر التربيعي لحاصل ضرب رقم لاسبير في رقم باش]، وهذا الرقم يهتم بالناحية الرياضية، ولكنه لا معنى له اقتصاديا، وهذا أبرز عيوبه. ويأخذ شكل المعادلة التالية:

$$I_f = \sqrt{I_r I_p} = \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1}} \times 100$$

مثال: يبين الجدول أسعار وكميات ثلاث منتجات استهلاكية للسنتين 2021 ، 2024م:

	سنة 2021		سنة 2024	
المنتج	**الكمية**	**السعر**	**الكمية**	**السعر**
السلعة الأولى	5000	9	8500	12
السلعة الثانية	8000	25	15000	31
السلعة الثالثة	9000	14	19000	17

وعلى اعتبار أن سنة 2021 هي سنة الأساس، المطلوب حساب ما يلي مع تفسير النتائج:

(أ) الوسط الحسابي لمناسيب الأسعار.

(ب) الرقم القياسي التجميعي البسيط للأسعار.

(ج) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير).

(د) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش).

(هـ) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (فيشر).

الحل: يمكن إعادة كتابة بيانات الجدول المعطى بعد إضافة الرموز المختلفة وأعمدة مختلفة توضح الكميات المطلوبة لحساب الأرقام القياسية المطلوبة كالتالي [وذلك باعتبار سنة 2021 هي سنة الأساس وسنة 2024 هي سنة المقارنة]:

	سنة الأساس 2021		سنة المقارنة 2024						
السلعة	كمية Q_0	سعر P_0	كمية Q_1	سعر P_1	منسوب السعر P_r	P_0Q_0	P_0Q_1	P_1Q_0	P_1Q_1
الأولى	5000	9	8500	12	133.33%	45000	76500	60000	102000
الثانية	8000	25	15000	31	124%	200000	375000	248000	465000
الثالثة	9000	14	19000	17	121.43%	126000	266000	153000	323000
		48		**60**	**378.78**	**371000**	**717500**	**461000**	**890000**
		$\sum P_0$		$\sum P_1$	$\sum P_r$	$\sum P_0Q_0$	$\sum P_0Q_1$	$\sum P_1Q_0$	$\sum P_1Q_1$

(أ) الوسط الحسابي لمناسيب الأسعار:

$$\frac{\sum P_r}{n} = \frac{378.78}{3} = 126.26$$

(ب) الرقم القياسي التجميعي البسيط:

$$I_s = \frac{\sum P_1}{\sum P_0} \times 100 = I_s = \frac{60}{48} \times 100 = 125\%$$

وهذا يعني أن المستوى العام لأسعار المنتجات الثلاثة قد ارتفع في سنة 2024 بمعدل 25% وذلك مقارنة بسنة 2021م.

(ج) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير):

$$I_r = \frac{\sum P_1 Q_0}{\sum P_0\, Q_0} \times 100 = I_r = \frac{461000}{371000} \times 100 = 124.26$$

وهذا يعني أن المستوى العام لأسعار المنتجات الثلاثة قد ارتفع في سنة 2024 بمعدل 24.26% وذلك مقارنة بسنة 2021م.

(د) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش):

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0\, Q_1} \times 100 = I_p = \frac{890000}{717500} \times 100 = 124.04$$

وهذا يعني أن المستوى العام لأسعار المنتجات الثلاثة قد ارتفع في سنة 2024 بمعدل 24.04% وذلك مقارنة بسنة 2021م.

(هـ) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر: الرقم القياسي الأمثل):

$$I_{f\,=}\sqrt{I_r I_p} = \sqrt{124.26\ \times\ 124.04} \cong 124.15$$

وهذا يعني أن المستوى العام لأسعار المنتجات الثلاثة قد ارتفع في سنة 2024 بمعدل 24.15% وذلك مقارنة بسنة 2021م.

ملاحظات عامة على الأرقام القياسية:

- الرقم القياسي للظاهرة في سنة الأساس يساوي 100.
- إذا كان الرقم القياسي للظاهرة في سنة المقارنة أكبر من 100 فهذا يعني أن هناك ارتفاعاً في المستوى العام للظاهرة مقارنة بسنة الأساس.
- إذا كان الرقم القياسي للظاهرة في سنة المقارنة أصغر من 100 فهذا يعني أن هناك انخفاضاً في المستوى العام للظاهرة مقارنة بسنة الأساس.

تمرين على الأرقام القياسية

السلعة	سنة الأساس 2021		سنة المقارنة 2024						
	كمية Q_0	سعر P_0	كمية Q_1	سعر P_1	منسوب السعر P_r	P_0Q_0	P_0Q_1	P_1Q_0	P_1Q_1
الأولى	5000	9	8500	12					
الثانية	8000	25	15000	31					
الثالثة	9000	14	19000	17					
		48		**60**					
		$\sum P_0$		$\sum P_1$	$\sum P_r$	$\sum P_0Q_0$	$\sum P_0Q_1$	$\sum P_1Q_0$	$\sum P_1Q_1$

(أ) الوسط الحسابي لمناسيب الأسعار:

$$\frac{\sum P_r}{n} = \text{———} =$$

(ب) الرقم القياسي التجميعي البسيط:

$$I_s = \frac{\sum P_1}{\sum P_0} \times 100 = I_s = \text{——} \times 100 = \quad \%$$

(ج) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير):

$$I_r = \frac{\sum P_1 Q_0}{\sum P_0\, Q_0} \times 100 = I_r = \text{————————} \times 100 =$$

(د) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش):

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0\, Q_1} \times 100 = I_p = \text{————————} \times 100 =$$

(هـ) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر):

$$I_{f} = \sqrt{I_r I_p} = \sqrt{\qquad\qquad \times \qquad\qquad} \cong$$

المحاضرة العاشرة

الأساليب البارامترية واللابارامترية – الفروض الإحصائية

(تم شرح هذه المحاضرة عند ذكر مستويات مقاييس البيانات واختبارات الدلالة الإحصائية)