جامعة محمد الصديق بن يحي - جيجل

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

الإحصـــاء I

دروس وأمثــلة تطبيقية

موجهة لطلبة السنة
أولى جذع مشترك

الدكتور: هشــام بورمة

السنة الجامعية 2018/2017

بسم الله الرحمن الرحيم

والصلاة والسلام على أشرف المرسلين، سيدنا محمد وعلى آله وصحبه أجمعين وبعد؛

يسرني أن أقدم هذه المطبوعة التي أتت كثمرة جهد سنوات عديدة في تدريس هذا المقياس لطلاب العلوم الاقتصادية والتجارية وعلوم التسيير، وكذا طلاب العلوم الاجتماعية والإنسانية، ولقد حرصت في تقديمها على البساطة في السرد والمنهجية في العرض، وعلى التوضيح عن طريق الأمثلة المباشرة في كل المواضيع، حرصا على الإيضاح والفهم السريع، والاستيعاب الجيد، من خلال سبعة فصول أساسية معروفة، مرتبة ومتسلسلة وفق مقرر وزارة التعليم العالي والبحث العلمي.

يعد علم الإحصاء واحدا من أكثر التخصصات الدراسية أهمية على الإطلاق ؛ لأنه أحد العناصر الرئيسية في أي عمل إداري أو مؤسسي مهما كان المجال الذي يتم تطبيقه فيه، فهو يعلم كيفية دراسة البيانات بطريقة صحيحة ، وتحليلها بشكل منهجي أكاديمي، وكيفية أخذ العينات والنماذج، .. ومن ثم يكون هدفها النهائي والأساسي هو اتخاذ القرارات السليمة والمناسبة.

وإذ اقدم هذا العمل لطلابنا؛ فإني أرجو أن أكون قد قدمت عملا مفيدا في تقريب وتبسيط مقياس الإحصاء 1 (الإحصاء الوصفي) من الطالب، وتوفير مرجع إضافي في مكتباتنا، كما أرجو أن يكون في مستوى الجهد المبذول في تأليفه وإخراجه على هذا النحو.

ونذّكر القارئ الكريم بأن هذه المطبوعة لا ندعي فيها أننا أتينا بشيء جديد أو مستحدث، وكل ما في الأمر أننا حاولنا أن نعرض هذا الموضوع (الإحصاء 1) بما يتماشى مع مقرر وزارة التعليم العالي والبحث العلمي، ويساعد الطلاب على الإلمام بالموضوع.

وحيث أن كل عمل قد تشوبه بعض النقائص، فإنني أرجو من القارئ الكريم التفضل بكل تنبيه أو ملاحظة على خطأ أو رأي، لنتمكن من تنقيحه وتحسينه في طبعة موالية إن شاء الله.

والله من وراء القصد

د/هشـــام بورمة

الفهـــرس

– الفصـــل التمهيدي: مدخل لعلم الإحصـــــاء

– الفصـــل الأول: عرض البيانات الإحصائية

– الفصـــل الثاني: مقاييس النزعة المركزية

– الفصـــل الثالث: مقاييس التشـــتت

– الفصـــل الرابع: مقاييس الشــــكل

– الفصـــل الخامس: الارتباط والانحدار البسيط

– الفصـــل السادس: مدخل للسلاسل الزمنية

– الفصـــل السابع: الأرقـــــام القياسية

الفصل التمهيدي

مدخــل لعلم الإحصــاء

الفصل التمهيدي

مدخـــل لعـلم الإحصـــاء

كلمة الإحصاء كلمة شائعة الاستخدام في مختلف حقول العلم والمعرفة، خاصة علم الاقتصاد والإدارة، إذ أن مختلف التقارير والقرارات تبنى على الدارسات والأرقام الإحصائية، أو تستخدم الإحصاء وطرقه في استدلالاتها.

1.تعريف الإحصاء

يعرف علم الإحصاء بأنه مجموعة من النظريات والطرق العلمية التي تهدف إلى جمع، وعرض، وتحليل البيانات المقاسة بشكل رقمي، ثم استخدام النتائج في التحليل واتخاذ القرارات.

كما يعرف أيضا الإحصاء بأنه ذلك الفرع من فروع الرياضيات الذي يشتمل على جمع المعلومات والبيانات لظاهرة ما، وتبويبها وعرضها وتنظيمها بواسطة جداول أو خطوط وتمثيلات بيانية، ثم تحليلها، وتفسير النتائج المستخلصة من أجل اتخاذ القرارات المناسبة.

ما يمكن استخلاصه مما سبق أن الإحصاء ليس هو البيانات الرقمية أو العددية العامة، وإنما هو فرع تطبيقي من فروع الرياضيات له نظرياته ورموزه ومصطلاحاته، وأساليب خاصة به. ولهذا فهو يخرج عن كونه مجرد جداول وأرقام أو تمثيلات وخطوط بيانية، وإنما يتعدى ذلك ليشمل ويبحث في كيفية الحصول على البيانات ثن تصنيفها فعرضها وتحليلها ثم استخلاص النتائج، وتفسيرها بشكل موضوعي وعلمي.

ويقسم علم الإحصاء إلى قسمين:

- **الإحصاء الوصفي**: وهو ذلك الفرع من الإحصاء الذي يتناول طرق تنظيم وتلخيص وعرض البيانات في صورة مبسطة (يدّرس الأن باسم الإحصاء 1).
- **الإحصاء الاستدلالي**: وهو ذلك الفرع من الإحصاء الذي يهتم بطرق الوصول إلى نتائج معينة أو توقعات ما عن المجتمع من خلال دراسة عينة من ذلك المجتمع (يدّرس الأن باسم الإحصاء 2 و 3).

2.خطوات العملية الإحصائية

حتى يسهل على الباحث قراءة وفهم البيانات لابد له من طريقة يتبعها، وخطواتها ينتهجها من أجل ذلك، هذه الخطوات تعرف بخطوات العملية الإحصائية وهي تنقسم إلى أربع كمايلي:

1.2. جمع البيانات

يقصد بجمع البيانات الحصول على بيانات رقمية أو وصفية تتصف بالصحة، الموضوعية والدقة عن الظاهرة المراد دراستها. والبيانات الإحصائية لا تجمع لذاتها ولكن لخدمة هدف معين، حيث يمهد لذلك الباحث كمايلي:

- تحديد البيانات المراد جمعها تحديدا دقيقا.
- التأكد من توافرها وأنه لا توجد مشكلة في الحصول عليها.
- معرفة الطرق السابقة تجنبا للتكرار ومعرفة للصعوبات التي واجهت من سبق.
- التأكد من حجم التكاليف المترتبة عن ذلك.
- مفاضلة هذه البيانات مع التكاليف المترتبة عنها، وأن هذه البيانات المجموعة ضرورية لدرجة تبرر التكاليف الناجمة عن عملية البحث.

1.1.2 أهمية الدقة في جمع البيانات: تعتبر الدقة في جمع البيانات في غاية الأهمية، كون أن الخطأ فيها يترتب عنه خطأ في الخطوات الموالية كالتحليل والتفسير، وبالتالي استنتاج خاطئ. فمهما بذل الباحث من جهد بعدها فإنه قد يصل لمتناقضات في أساليبه الإحصائية، كما قد تواجهه مشاكل إحصائية لن يجد لها تفسير ولا مخرج، لهذا وجب الحرص وتوخي الدقة عن جمع البيانات الإحصائية.

2.1.2 مصادر جمع البيانات: تصنف مصادر جمع البيانات في الغالب إلى قسمين هما:

أ. **المصادر المباشرة (الميدانية)**: تشمل البيانات التي يمكن الحصول عليها من مصادرها الأساسية عن طريق المواجهة أو المراسلة أو بأي طريقة أخرى وتسمى البيانات المجموعة بهذه الطريقة بالبيانات الأولية، وتتميز هذه البيانات بسهولة مراجعتها والتأكد من صحتها، لكن ما يعاب عليها احتياجها لوقت طويل وتكاليف كبيرة.

ومن الطرق التي تستخدم في هذا النوع: المقابلة الشخصية، المراسلة، الهاتف، الملاحظة... الخ. كما أنها تعتمد في جمعها على أسلوبين الحصر الشامل وهو أخذ المجتمع ككل للدراسة، وأسلوب المعاينة أو اختيار العينة من المجتمع الذي نلجأ إليه في حالة استحالة دراسة المجتمع ككل (سنتناول ذلك بشئ من التفصيل في النقاط الموالية).

ب.المصادر غير المباشرة (التاريخية): وتشمل الوثائق والمطبوعات والنشرات الإحصائية التي تصدرها الهيئات الحكومية في البلدان المختلفة وتسمى البيانات المجموعة بهذه الطريقة بالبيانات الثانوية مثل تعداد السكان والصادرات والواردات، لأنه قبل عملية جمع البيانات عن أي مشكلة لابد على الباحث من دراسة وافية ومستفيضة للمصادر التاريخية للموضوع محل الدراسة، وذلك توفيرا للوقت والجهد.

2.2. تبويب البيانات

بعد جمع البيانات يتحتم على الباحث مراجعتها وكتابتها وتلخيصها ثم عرضها في شكل جدوال أو تمثيلات وخطوط بيانية. وتقسم البيانات إلى نوعين:

1.2.2. البيانات الوصفية(النوعية): هذا النوع من البيانات – غالبا – لا يمكن تمثيلها بالأرقام، بل تمثل في العادة بنصوص باللغة الطببعية، وهي نوعين:

أ. **البيانات الإسمية**: بيانات تكون في صورة غير عددية ولا يمكن التفاضل بينها وتتكون من مجموعات متنافية مثل الأسئلة التي جوابها يكون بنعم أو لا، أو السؤال عن نوع الجنس: ذكر أو أنثى، أو السؤال عن الحالة الاجتماعية: متزوج ، أعزب، مطلق أو أرمل). وعادة ما نضع رموز أو أرقام لهذه الأجوبة كــ نعم بــ (1) ولا بــ (0)، ومن الخطأ أن نقوم بإجراء العمليات الحسابية على هذه الرموز أو على الأرقام الموضوعة المرافقة للصفات.

فتخيل بأنك تقوم بجمع معلومات عن الألوان المفضلة لدى مجموعة من الطلبة فلنقل مثلا: الأحمـــــر (1) الأخضر (2) الأسود (3) وهكذا...، هل يمكنك أن تجمع هذه البيانات مع بعضها البعض وتستخرج منها أي عملية حسابية ؟ بالطبع لا لأن النتائج الحسابية هنا ليس لها أي معنى وستلاحظ أنها بلا فائدة، ففي هذا المثال نلاحظ أن العملية الحسابية لا تدل على أي معلومة مفيدة.

ب. **البيانات الترتيبية**: هي أيضا تكون في صورة غير عددية ولا يمكن إجراء العمليات الحسابية عليها، والفرق بينها وبين البيانات الأسمية هو عملية المفاضلة والترتيب بين الطبقات، حيث نستطيع في البيانات الترتيبية ترتيب البيانات، مثل مستوى التعليم: ابتدائي، متوسط، ثانوي أو جامعي، أو عدد سنوات الخبرة: أقل من 5 سنوات، ما بين 5 و10 سنوات أو أكثر من 10 سنوات، هنا نلاحظ نوع من المفاضلة والطبقية بين الخيارات.

2.2.2. البيانات الكمية(الرقمية): هذا النوع من البيانات يمثل بالأرقام، وهي أنواع نذكر منها:

أ. **بيانات الفترة**: تكون بيانات الفترة في صورة عددية، توضح أو تمثل القيمة الفعلية للظاهرة، نستطيع أن نجري عليها العمليات الحسابية، لكن عملية الضرب والقسمة بين البيانات في هذا النوع لا معنى حقيقي لها. فمثلا لو أخذت بيانات عن طالبين في الامتحان؛ الأول ينهي أسئلته في نصف ساعة، والثاني ينهي أسئلته في ساعة كاملة، هنا نرى أن هناك فرقا مقداره نصف ساعة في حال الطرح بينهما، لكن الضرب والقسمة بينهما لا تفيد في شيء.

يمتاز هذا النوع أيضا بتساوي المسافات بين الرتب، حيث يستخدم هذا المقياس كثيرا في العلوم الإنسانية، فمن أمثلة هذا النوع من البيانات درجة الحرارة أو علامات الطلبة في الامتحانات؛ فهي تقاس

بمقدار بعدها عن الصفر، ولكن الصفر هنا لا يعني عدم وجود الظاهرة. فالطالب الذي يحصل على درجة صفر، لا يعني أنه لا يملك أي معلومة، وكذلك درجة الحرارة عندما تكون صفر لا تعني انعدام الحرارة، لهذا ينبغي علينا مراعاة قيمة الصفر في هذا النوع من البيانات.

ب. البيانات النسبية: بيانات تكون في مستوى أعلى من البيانات السابقة، بحيث تكون في صورة عددية ونستطيع أن نجري عليها جميع العمليات الحسابية، حيث تعطيك النتائج هنا معلومات ذات قيمة وفائدة. فمثلا، لو أخذنا سنوات الخبرة والعمل في مجال معين كشخص خبرته 5 سنوات والأخر 3 سنوات فمن خلال قسمتهم على بعض فإننا نستخرج مقدار النسبة والتناسب بين الشخصين.

يمتاز هذا النوع من البيانات في أن قيمة الصفر تعني انعدام الظاهرة، وهي أغلب البيانات المتعلقة بالعلوم الفيزيائية والهندسية.

كما أنه يمكن تقسم البيانات الكمية إلى:

– **بيانات كمية منفصلة (متقطعة)**: وهي البيانات التي لا تأخذ قيما رقمية عشرية، حيث تكون أرقامها طبيعية ، يحصل عليها من عملية العد، ولا يمكن تجزئة وحدة قياسها إلى وحدات أصغر، كعدد الأطفال داخل الأسرة، عدد الغرف... إلخ.

– **بيانات كمية متصلة (مستمرة)**: وهي بيانات نحصل عليها من عملية القياس، ولا تكون بالضرورة أرقامها طبيعية، ويمكن تجزئة وحدات قياسها إلى وحدات أصغر، كالطول، العمر ، الوزن... الخ، كلها تقاس بوحدات يمكن تجزئتها إلى وحدات أصغر. مثلا: الوزن بالكيلوغرام يمكن أيضا أن يقاس بالغرام وهكذا.

تعتبر البيانات النوعية بيانات غنية بالمعلومات أكثر من البيانات الرقمية، فهي تدرس في الغالب الأنماط والتوجهات بشكل عام، بينما غالبا ما ينظر إلى البيانات الكمية على أنها بيانات أكثر موضوعية للتحليل.

3.2. تحليل البيانات

بعد عملية جمع البيانات وتبويبها يتم معالجتها رياضيا باستخدام العلاقات، القوانين والعمليات الحسابية لاستخراج قيم عددية لها مؤشراتها ومدلولاتها وتفسيراتها، من بين العمليات والقوانين المستخدمة في عملية التحليل أو المعالجة: مقاييس النزعة المركزية، مقاييس التشتت، معاملات الارتباط وغيرها، بغية الوصول إلى النتائج.

4.2. تفسير نتائج تحليل البيانات

ويقصد بها توضيح دلالات النتائج الإحصائية التي تم الوصول والتوصل إليها، وبالتالي وضع الاستنتاجات وإصدار الأحكام على مجتمع الدراسة، بناء على معرفة الباحث للظروف المحيطة بها وكل نظرياتها المختلفة. هذه القراءات أو التفسيرات تختلف من باحث لأخر.

3.المجتمع والعينة

إن من المصطلحات الأكثر تداولا واستعمالا في الإحصاء الوصفي هي: المجتمع، العينة، والمفردة.

1.3.المجتـــــمع الإحصائي

إن مجتمع الدراسة أو المجتمع الاحصائي هو مفهوم إحصائي يقصد به جملة العناصر أو المفردات التي تستند اليها الدراسة، أي كل الوحدات المراد دراستها بغرض تعميم النتائج. وينبغي الإشارة إلى أن عملية تحديد المجتمع هي عملية نسبية ترتبط بالبحث وأهدافه ومشكلته.

ويمكن تصنيف المجتمع حسب متغيرين هما: المعلومية، والتجانس.

ماذا يعني مجتمع محدود؟

يعني أننا يمكننا معرفة جميع مفرداته وعناصره، فمثلا إذا كان المجتمع طلبة السنة أولى بجامعة جيجل، فيمكنا الرجوع إلى جميع الكليات في جامعة جيجل والوصول إلى أي طالب في أي قسم، فيكون هنا المجتمع محدود، حيث نعرف عددهم وتوزيعهم وبياناتهم.

متى يكون المجتمع غير محدودا؟

عندما لا يمكن حصر جميع مفردات المجتمع يكون المجتمع غير محدودا، فمثلاً إذا كان المجتمع هو مدمني المخدرات، والمدخنين، أو زبائن شركات الاتصالات، مشاهدي الفضائيات... الخ، فإنه لا يمكننا حصرهم عددا، ولا الوصول إلى أي منهم بشكل محدد ودقيق. ويكون المجتمع غير محدود وغير معروف.

ماذا يعني مجتمع متجانس؟

إذا كان لمفردات المجتمع نفس الصفات المرتبطة بالدراسة، مثلا: إذا كنا ندرس المشاكل التي تواجه الطلبة الذكور في السنة الأولى علوم اقتصادية وتجارية وعلوم التسير بجامعة جيجل، فإن مجتمع الدراسة هو جميع الطلبة الذكور الدارسين في السنة الأولى بجامعة جيجل، وهنا اشتراك في صفة الجنس (كلهم ذكور)، والتخصص (كلهم علوم اقتصادية وتجارية وعلوم التسير بجامعة جيجل) ، والمستوى (كلهم سنة أولى)، نجد أن مفردات المجتمع متجانسة.

ماذا يعني عدم تجانس مفردات المجتمع؟

إذا كان يوجد اختلافات بين أفراد المجتمع في إطار حيز الدارسة بمعنى: إذا كان مجتمع الدراسة هو طلبة السنة أولى بجامعة جيجل، وموضوع الدراسة هو التحصيل الدراسي، فإن طلبة السنة أولى ليسوا متجانسين، لأنهم ذكور وإناث، وليسوا بنفس التخصص، لأن فيهم العلمي والأدبي والرياضي.. فهنا لا يوجد تجانس بين مفردات المجتمع.

2.3.العينة الإحصائية

رأينا ان المجتمع الإحصائي أو مجتمع الدراسة هو جميع القيم والمفردات التي يمكن أن يشملها موضوع الدراسة، فهو قد يكون محدودا وبالتالي يسهل تحليل أو معالجة بياناته ومن تم تعميم النتائج، كما قد يكون غير محدودا بحيث يستحيل تحليل ومعالجة بياناته، وهنا يتحتم علينا اختيار مجموعة من المفردات أو العناصر بطريقة معينة، بحيث تمثل هذه العناصر أو المفردات المجتمع الأصلي أحسن تمثيل، حتى تكون نتائج التعميم صادقة وصحيحة وهذا ما يعرف بالعينة. إذن؛ العينة هي جزء من المجتمع اختير بطريقة ما تمثله أحسن تمثيل.

1.2.3. أسباب اللجوء إلى العينة: إن اجراء البحث على كامل مجتمع الدراسة الأصلي يكون مفضلا في معظم الحالات على اختيار عينة وإجراء الدراسة عليها، نظرا لما يعطيه دراسة كامل المجتمع من نتائج أقرب للواقع وأكثر قابلية للتعميم. إلا أن هناك أسباب عدة قد تدفع الباحث الى الاعتماد على العينة بدلا من دراسة مجتمع الدراسة بأكمله، ومن ضمن تلك الأسباب مايلي:

– ارتفاع التكلفة والوقت والجهد؛ ففي حالة كون مجتمع الدراسة الأصلي كبيرا ومتباعدا جغرافيا، فإن ذلك يتطلب تكلفة عالية وجهدا كبيرا ووقتا طويلا من الباحث.

– التجانس التام في خصائص مجتمع الدراسة الأصلي؛ فهناك بعض أنواع الظواهر التي تكون فيها عناصر المجتمع ومفرداته متجانسة بشكل كبير، وبالتالي فإن النتائج نفسها يتم الحصول عليها سواء أجريت الدراسة علة كامل المجتمع أو على أجزاء منه. ومن الأمثلة الواضحة في ذلك فحص الدم، فسواء تم اجراء الفحص على عينة من الـــدم أو كله، فإن النتائج ستكون واحدة، وبالتـــالي لا تكون هناك ضرورة لدراسة كامل المجتمع.

– عدم إمكانية اجراء الدراسة على كامل عناصر المجتمع الأصلي؛ ففي بعض أنواع الأطعمة المنتجة كالألبان والعصائر وبعض السلع الكهربائية كالتلفاز، تقوم معظم المصانع باختيار عينات من الإنتاج بشكل دوري ويتم فحص تلك العينات للتأكد من سلامتها ومطابقتها للمواصفات المحددة.

– عدم إمكـــانية حصر كامل عناصر مجتمع الدراسة الأصـــلي؛ كدراسة المدمنين على المخدرات، فقد لا تتوفر معلومات عن كامل المدمنين في الدولة، أو قد تكون المعلومات ســـرية ولا يمكن البوح بها عن هذه الفئة.

2.2.3. أنواع العينات: هناك عدة طرق يمكن استخدامها لاختيار العينة موضوع الدراسة، حيث يعتبر نوع العينة المختارة من الأمور الهامة التي يجب على الباحث أن يوليها اهتمام خاصا.

وبشكل عام لا توجد طريقة مثلى يمكن تفضيلها على غيرها من الطرق، فلكل طريقة من طرق اختيار العينات مزاياها كما لها عيوب، وما قد يفضل طريقة على غيرها هو طبيعة الظاهرة والبحث وظروف الباحث وطبيعة مجتمع الدراسة.

وبشكل عام تقسم العينات إلى مجموعتين رئيستين هما:

أ. **العينات الاحتمالية** (**العشوائية**): وفيها يتم اختيار العينة بطريقة عشوائية، بحيث يعطى لكل عنصر من عناصر مجتمع الدراسة فرصة الظهور في العينة مع ضرورة أن تكون فرصة الظهور لكل عنصر في العينة معروفة ومحددة مسبقا. ويمكن تقسيم العينات الاحتمالية (العشوائية) إلى أربعة أنواع هي:

– **العينة العشوائية البسيطة**: يتطلب استخدام هذه الطريقة حصر كامل العناصر التي يتكون منها مجتمع الدراسة الأصلي ومعرفتها ليتم لاحقا الاختيار من بين تلك العناصر. فمثلا اذا كانت الدراسة تتعلق بتخصصات الطلبة في الجامعة، فإن استعمال طريقة العينة العشوائية البسيطة تتطلب توفر لدى الباحث قائمة مفصلة ودقيقة بجميع تخصصات الطلبة الموجودة في الجامعة ليتم الاختيار من بينهم، حيث تعطى لكل عنصر من عناصر مجتمع الدراسة الأصلي نفس فرصة الظهور في العينة، والتي تكون معروفة ومحددة مسبقا.

وهناك وسائل عديدة يمكن استخدامها في تحديد مفردات العينة وفق هذه الطريقة منها:

– استخدام جداول الأرقام العشوائية، وهي جداول معدة خصيصا لذلك، حيث يتم من خلالها تحديد المفردات التي ستدخل في عينة الدراسة.

– استخدام الدواليب؛ كالتي تستخدم في السحب على جوائز اليناصيب.

– إعطاء أرقام متسلسلة لعناصر المجتمع الأصلي ووضع كل رقم في ورقة منفصلة في وعاء أو كيس، ثم سحب العدد المطلوب من الوعاء بشكل عشوائي.

تتميز العينة العشوائية البسيطة بسهولة استعمالها وتطبيقها، حيث تكون نتائجها قابلة للتعميم على المجتمع الأصلي للدراسة خصوصا إذا كان حجم العينة كبيرا نوعا ما.

لكن ما يعاب على هذه الطريقة؛ ارتفاع تكلفة استخدامها في بعض الأبحاث التي تكون عناصر مجتمع الظاهرة منتشرة في مناطق جغرافية متباعدة، كما يعاب عليها أيضا أن احتمالية عدم تمثيل العينة لبعض شرائح المجتمع، خصوصا إذا كان حجم العينة المختار صغيرا نسبيا.

– **العينة المنتظمة**: وفيها يتم حصر عناصر مجتمع الدراسة الأصلي، بحث يعطى لكل عنصر قم متسلسل، ثم نقسم عدد عناصر المجتمع الأصلي على عدد أفراد العينة المطلوبة، فينتج رقم معين هو الفاصل بين كل مفردة يتم اختيارها ضمن العينة والمفردة التي تليها. بعد ذلك يتم اختيار رقم عشوائي ضمن المجال

أو الرقم الذي تم حسابه في الخطوة السابقة، وتكون مفردات العينة هي أصحاب الأرقام المتسلسلة التي تفصل بين الرقم العشوائي المختار والترتيب الذي يليه.

مثال01: لنفرض ان باحثا يريد ان يختار عينة حجمها 50 طالبا من طلبة السنة الأولى علوم اقتصادية وتجارية وعلوم التسيير البالغ عددهم 2000 طالب، فوفق هذا الاسلوب يقسم حجم المجتمع 2000 على حجم العينة المراد اختياره 50، وذلك لتحديد المسافة او الفترة بين كل مفردة وأخرة ستختار ضمن العينة. هنا ينتج لنـــا (2000/50=40)، وبالتالي المجال أو الفترة التي سنخـــــــتار منها المفردة الأولى عشوائيا هي مـــا بين (1 – 40)، وليكن مثلا الرقم هو 6 فتكون العينة البالغ عددها 50 مفردة كمايلي:

6، 46، 86، 126، 166، ..... وهكذا.

ما يعاب على هذه الطريقة أنه إذا كان الرقم الأول اختير بطريقة متحيزة، فإن العينة كلها ستكوم متحيزة، مما سيؤثر على دقة النتائج وموضوعيتها.

- **العينة الطبقية:** عندما يكون المجتمع الأصلي للدراسة مقسما إلى طبقات وفئات متجانسة فيما بينها (ذكر، أنثى...الخ)، فإن أنسب طريقة لاختيار عينة من ذلك هي طريقة العينة الطبقية، حيث يتم تقسم مجتمع الدراسة الأصلي إلى طبقات أو فئات معينة وفق معيار معين، ثم يختار من كل فئة أو طبقة بشكل عشوائي يتناسب مع حجم تلك الفئة في مجتمع الدراسة الأصلي.

مثال02 : يتكون مجتمع من 2000 فرد منه 1100 ذكور، أردنا تشكيل عينة حجمها 15% من حجم المجتمع.

- ماهي الطريقة المناسبة لذلك؟ وضح كيف يتم ذلك؟

الحل: نلاحظ أن المجتمع مقسم الى طبقتين متجانستين داخليا، ومختلفتين فيما بينهما؛ وعليه فالطريقة المناسبة لاختيار حجم العينة هي طريقة العينة العشوائية الطبقية.

– نعـــلم أن العينة المختارة وفق طريقة العينة العشوائية الطبقية تحمل نفس نسب المجتمع الأصلي. لهـــذا سنحدد نسبة كل طبقة تكون المجتمع الأصلي أولا ومن تم نختـــار العينة التي ستكون بنفس هذه النسب، كمايلي:

- حجم المجتمع: 2000 فرد.
- حجم العينة: 15% من حجم المجتمع، اي: (15/100)x 2000= 300 فرد.
- الذكور 1100 فرد، أي بنسبة: 1100/2000= 55%.
- الاناث 2000–1100=900 فرد، أي بنسبة: 900/2000= 45%.

اذن: يكون حجم العينة التي ستختار كمايلي:

- الذكور: (55/100)x 300= 165 فرد.
- الاناث: (45/100)x 300= 135 فرد.

تتميز العينة الطبقية بأنها تضمن تمثيلا لجميع فئات مجتمع الدراسة الأصلي أو شرائحه، إلا أنها تتطلب أحيانا جهدا وتكلفة عالية من الباحث، كما تتطلب ضرورة معرفة وحصر عدد عناصر كل فئة وشريحة في مجتمع الدراسة الأصلي.

- **العينة العنقودية:** وتعرف أيضا بالعينة المتعددة المراحل، يلجأ إليها الباحث عندما يكون مجتمع الدراسة كبير جدا ومتناثرا على مساحات شاسعة تكلف الكثير من الوقت والجهد في التنقل بينها عند جمع البيانات، أيضا في حالة عدم وجود إطار يضم جميع مفردات المجتمع فيستحيل الاختيار العشوائي مباشر من المجتمع. لهذا يلجأ الباحث إلى أخذ العينة على مراحل متعددة متتالية. في المرحلة الأولى يتم تقسيم المجتمع إلى عدد محدد من وحدات المعاينة الكبيرة الحجم ومنها يختار بعضها عشوائيا ثم يتلو ذلك كمرحلة ثانية تقسيم الوحدات المختارة عشوائيا من المرحلة الأولى إلى وحدات أقل منها في الحجم ثم يختار بعضها عشوائيا، وهكذا تتابع مراحل التقسيم والاختيار العشوائي، وعدد هذه المراحل ليس ثابت بل يتوقف على طبيعة مجتمع الدراسة وإمكانيات الباحث. في المرحلة الأخيرة يصل الباحث إلى وحدات المعاينة التي سيجمع عنها بيانات البحث ويطلق عليها وحدات المعاينة الأولية.

ب. العينات غير **الاحتمالية** (غير **العشوائية**): وفيها يتم اختيار العينة بطريقة غير عشوائية، بحيث يتم مقدما استثناء بعض عناصر الدراسة من الظهور في العينة لأسباب معينة منها عدم توفر المعلومات المطلوبة للدراسة لدى تلك العناصر، أو لاستحالة الوصول لتلك العناصر، أو ارتفاع تكلفة الحصول على المعلومات المطلوبة فيما إذا تم اختيار العينة بشكل عشوائي. ويمكن تقسيم العينات غير الاحتمالية (غير العشوائية) إلى نوعين هما:

- **العينة العمدية أو المقصودة:** يلجأ الباحث إلى هذه الطريقة عندما يكون مجتمع الدراسة كبير جدا، وإمكانياته لا تسمح إلا بدراسة عينة حجمها صغير جدا بالنسبة لمجتمع الدراسة، في هذه الحالة يتعمد الباحث اختيار مفردات معينة كعينة لمجتمع الدراسة يرى بخبرته السابقة أن هذه العينة يمكن أن تعطي تمثيلا مقبولا لمجتمع الدراسة.

مثلا إذا أراد باحث دراسة الخصائص الاقتصادية أو الاجتماعية لقرية ما، وكانت إمكانياته المالية والإدارية لا تسمح له بعينة سوى سكان قرية واحدة، فإنه في هذه الحالة إذا تم اختيار القرية عشوائيا من بين آلاف القرى في تلك الدولة، فإن الصدفة قد تأتي بقرية بعيدة في خصائصها (من حيث الظاهرة موضوع الدراسة) عن خصائص معظم قرى تلك الدولة، كأن تأتي بالصدفة قرية ساحلية معظم سكانها من الصيادين أو قرية قريبة من مشروع صناعي ضخم يستوعب في قواه العاملة معظم سكانها، هذه القرية أو

تلك قد يأخذ النمط المعيشي لسكناها طابعا خاصا بعيدا عن النمط المعيشي المعتاد لبقية القرى، لذلك فأي منها لا يمكن أن يعطي تمثيلا مقبولا لريف تلك الدولة. لهذا فإن الباحث وعلى ضوء خبراته السابقة يتعمد اختيار قرية معينة يرى أنها يمكن أن تمثل الريف. وهذه الطريقة غير علميه وغالبا يتم اللجوء إليها في حالة البحوث التمهيدية.

- **العينة الحصصية:** تستخدم عادة في استطلاعات الرأي العام نظرا للسرعة التي تتم بها وقلة تكلفتها ، وفي هذه العينة يتم تقسيم المجتمع إلى طبقات طبقاً لخصائص معينة تمثل كل طبقة من طبقات العينة بنسبة وجودها في المجتمع ولكن يترك للباحث حرية اختيار مفردات كل طبقة وهذا يؤدي إلى عدم تمثيل المجتمع تمثيلا تاما.

على سبيل المثال في حالة دراسة صفاي الدخل لمنطقة ما، وطُلب أن يكون حجم العينة هو 100 فرد مثلا، كلف الباحث جامعوا البيانات بالحصول على البيانات من 20 موظفا، 45 من العمال الحرفيين، 35 من ذوي الأعمال الحرة، وتترك الحرية لجامعي البيانات في اختيار صفات الأفراد المطلوبة لكل طبقة من الطبقات المذكورة. واضح أنه رغما من أن هذه الطريقة في ظاهرها مماثلة للعينة الطبقية العشوائية، إلا أنه في حالة العينة الطبقية العشوائية يكون اختيار المفردات عشوائيا من داخل كل طبقة ولا يترك لجامع البيانات حرية اختيار المفردات من كل طبقة والذي قد يترتب عليه تميزا كبيرا.

4. تقريب الأرقام

التقريب جزء هام جدا في الرياضيات، ومعناه إزالة عدد كبير من الأرقام وتحويلها إلى عدد صحيح، أو عدد عشري منتهي. وهو أداة مفيدة جدا في الحياة اليومية، فبفضل التقريب نستطيع اختصار عدد هائل من الأعداد العشرية الضخمة إلى عدد صحيح يتكون من رقم إلى 5 أو 6 أرقام.

عند تقريب عدد عشري، فإننا نزيل جزءا كبيار جدا من الأعداد العشرية. فعلى سبيل المثال، العدد: 2.9848425 يعتبر هذا العدد كبيرا جدا حتى نقرأه، فبالتقريب يمكننا أن قراءته بسهولة ونزيل كل هذه الأعداد حتى يصبح الناتج هو 3، فلولا التقريب لعشنا في حياة مليئة بالأعداد العشرية التي يعجز الفرد عن قراءتها، ولكن الأمر يتطلب حفظ الأعداد الكريمة والبخيلة، فعند تقريب 909 سيكون 900 لأنه لا يحتوي على أعداد كريمة، أما عن 998 سيكون 1000 لأنه يحتوي على أعداد كريمة.

والأعداد البخيلة هي الأعداد التي لا يمكن الاستلاف منها، ويجب أن تكون أول خانة من القيمة التقريبية عدد منها، بعبارة أخرى هي الأعداد التي تكون أقل من الرقـــم 5 ، أي: 0، 1، 2، 3، 4، بينما الأعداد الكريمة هي الأعداد التي يمكن الاستلاف منها. ولا يجب أن تكون أول خانة من القيمة التقريبية عدد منها، أي الأعداد التي تكون أكبر من الرقـــم 5 وهي : 5، 6، 7، 8، 9.

مثال03:

1. قرّب الأعداد **53** و **346** إلى أقرب عشرة.

- **53** بما إن الآحاد هنا **3** أي أقل من **5** فحسب القاعدة يتم حذفها و استبدالها بـــ **0** ليصبح العدد **50.** أي: **53 ≈ 50**
- **346** ننظر إلى الرقم الذي يسبق الـــ **4** ونرى بأنه **6** وهو أكبر من 5، لذا نُقرّب العدد إلى الأعلى أي إلى **350**، إذن: **346 ≈ 350.**

2. قرّب العدد **855** إلى أقرب مئة.

- **855**.بما أن العشرات هنا 5 لذا وحسب القاعدة نقوم بإضافة واحد إلى **8** فيصبح **9** واحذف **55** و استبدلها بـــ **00** فيصبح العدد **900**.، أي: **855 ≈ 900**

وبنفس الطريقة و بموجب القاعدة يمكن التقريب إلى الألف والعشرة آلاف وإلى آخره.

3. قرّب الأعداد **0.455** و **3.4032** إلى رقمين بعد الفاصلة.

- **0.455** .بما أن الرقم العشري الموجود بعد الرقمين المراد التقريب لهما هو**5** فحسب القاعدة يتم التقريب إلى أعلى ليصبح العدد **0.46. أي: 0.455 ≈ 0.46**
- **3.4032** .بما أن الرقم العشري الموجود بعد الرقمين المراد التقريب لهما هو**3** وهو أصغر من **5** فحسب القاعدة يتم التقريب للأسفل أي إلى **3.40**، إذن: **3.4032 ≈ 3.40.**

وبنفس الطريقة و بموجب القاعدة يمكن التقريب إلى أي رقم بعد الفاصلة؛ وحتى إلى رقم صحيح.

5. المجالات وحدودها

توضع المجالات للدلالة على امتداد فترة أو لتلخيص بيانات معينة؛ تجنبا لطولها، والمجالات تكتب على ثلاثة أنواع كمايلي:

- **المجال النصف مفتوح:** وهو على نوعين:
  - النصف مفتوح من الأسفل: $]b, a]$ حيث: $a < b$ ؛ ويقرأ a وأقل من b. (b لا تحتسب في المجال)
  - النصف مفتوح من الأعلى: $[b, a[$ حيث: $a < b$ ؛ ويقرأ أكبر من a إلى b.(b تحتسب في المجال)
- **المجال المفتوح:** ويكون على الشكل: $]b, a[$ حيث: $a < b$ ؛ ويقرأ أكبر من a وأقل من b. (a و b لا تحتسبان في المجال)
- **المجال المغلق:** ويكون على الشكل: $[b, a]$ حيث: $a < b$ ؛ ويقرأ من a إلى b. (a و b تحتسبان في المجال).

الفصل الأول

عرض البيانات الاحصائية

الفصل الأول
عرض البيانات الإحصائية

1. العرض الجدولي للبيانات الإحصائية

حتى يسهل تفسر وترجمة البيانات بشكل أفضل، يلجأ الباحث عادة الى تلخيصها ووضعها في شكل جدول يسمى بالتوزيع التكراري.

1.1. التوزيع التكراري

التوزيع التكراري جدول يتكون في الغالب من عمودين؛ يمثل الأول الظاهرة المدروسة في حين يمثل الثاني عدد المفردات (العناصر) التابعة لكل مشاهدة أو ما تسمى بالتكرارات. وتنقسم التوزيعات التكرارية حسب تبويبها إلى قسمين:

1.1.1. التوزيع التكراري للبيانات غير المبوبة

وتتوزع فيه البيانات حسب نمط واحد، سواء كانت بيانات وصفية (نوعية) أو بيانات كمية (رقمية)؛ المتقطعة أو مستمرة.

مثال01-01 (بيانات وصفية)

تمثل البيانات التالية التوجيه النهائي بعد الطعن لـــ 32 طالب الى السنة الثانية حسب الشعبة في ميدان العلوم الاقتصادية والتجارية وعلوم التسيير.

علوم تجارية	علوم اقتصادية	علوم التسيير	علوم مالية	علوم تجارية	علوم تجارية	علوم اقتصادية	علوم اقتصادية
علوم تجارية	علوم التسيير	علوم اقتصادية	علوم مالية	علوم التسيير	علوم تجارية	علوم التسيير	علوم اقتصادية
علوم مالية	علوم تجارية	علوم تجارية	علوم التسيير	علوم التسيير	علوم مالية	علوم تجارية	علوم تجارية
علوم مالية	علوم التسيير	علوم تجارية	علوم تجارية	علوم تجارية	علوم مالية	علوم التسيير	علوم التسيير

الحل: نقوم بوضع جدول التوزيع التكراري كمايلي:

الشعبة	تفريغ البيانات	التكرارات
علوم تجارية	//// //// //	12
علوم اقتصادية	////	05
علوم التسيير	//// ////	09
علوم مالية	//// /	06
المجـــمــــوع		**32**

مثال01-02 (بيانات كمية)

تمثل البيانات التالية عدد الأرصدة المحصلة للطبة الموجهين الى السنة الثانية في **المثال01**، المطلوب كوّن الجدول التكراري .

58	30	60	60	30	33	48	30
30	60	39	40	39	39	30	41
59	56	40	37	37	40	48	41
59	56	41	41	41	30	30	39

الحل: نرتب أولا البيانات ترتيبا تنازليا أو تصاعديا، ثم نقوم بوضع جدول التوزيع التكراري كمايلي:

عدد الأرصدة	تفريغ البيانات	التكرارات (عدد الطلبة)
30	~~////~~ //	07
33	/	01
37	//	02
39	////	04
40	///	03
41	~~////~~	05
48	//	02
56	//	02
58	/	01
59	//	02
60	///	03
المجمـــوع		**32**

2.1.1. التوزيع التكراري للبيانات المبوبة

يتم اللجوء الى هذه الطريقة عندما يكون حجم البيانات كبير، لهذا يتم تجميع البيانات في مجموعات تسمى بالفئات(مجالات)، كل فئة يقابلها تكرار معين.

ومن أجل تشكيل جدول التوزيع التكراري في البيانات المبوبة نتبع مايلي:

– تحديد طول الفئة

يتم تحديد طول الفئة من خلال قسمة المدى على عدد الفئات (في الغالب لا يتجاوز عدد الفئات 15 فئة ولا يقل عن 5 فئات) كمايلي:

$$C=\frac{R}{K}=\frac{X_{Max}-X_{Min}}{1+3,322.\log N}$$

حيث: C : طول الفئة. X_{Max} : أكبر قيمة في البيانات. N: عدد القيم

K : عدد الفئات X_{Min} أصغر قيمة في البيانات Log: اللوغاريتم العشري

– كتابة حدود الفئات

يتم عادة تحديد وكتابة حدود الفئات في شكل مجالات مغلقة من الأدنى ومفتوحة من الأعلى(أي أن القيمة الدنيا للفئة يشملها المجال، بينما لا يشمل قيمتها العليا)، وهذا انطلاقا من الفئة الأولى والتي يكون حدها الأدنى أصغر قيمة في البيانات X_{Min} ، ثم نضيف له طول الفئة؛ وهكذا مع الفئات الموالية بإضافة في كل مرة طول الفئة C المحصل الى الحد الأدنى (الذي كان الحد الأعلى للفئة السابقة) حتى نغطي جميع البيانات.

– تحديد مراكز الفئات

نحصل على مركز الفئة من خلال جمع الحد الأعلى للفئة من الحد الأدنى لها (مدى الفئة) وقسمة الناتج على اثنين كما يلي:

$$x_i = \frac{x_{Max} + x_{Min}}{2}$$

حيث: x_i : مركز الفئة x_{Max} : الحد الأعلى للفئة. x_{Min} : الحد الأدنى للفئة

– وضع جدول التوزيع التكراري

يتم في الأخير وضع جدول التوزيع التكراري من خلال كتابة الفئات الناتجة وتكراراتها المقابلة n_i .

مثال01-03: تمثل البيانات أدناه علامات 64 طالبا من طلبة السنة الأولى علوم اقتصادية وتجارية وعلوم التسيير بجامعة جيجل في مقياس الإحصاء1. المطلوب تشكيل جدول التوزيع التكراري.

19	08	06	20	07	17	08	06
19	08	08	18	15	11	08	08
11	10	07	11	11	06	10	07
15	10	12	11	10	07	10	12
16	15	15	06	07	18	15	15
19	07	17	20	07	19	07	17
20	15	11	08	14	14	15	11
19	12	06	10	12	13	12	06

الحل:

1. تحديد طول الفئة

$$C = \frac{R}{K} = \frac{X_{Max} - X_{Min}}{1 + 3{,}322.\log N} = \frac{20 - 06}{7.00} = 2$$

2. كتابة حدود الفئات

نبدأ كتابة حدود الفئة الأولى في جدول التوزيع التكراري من أصغر قيمة في البيانات (06) وهي تمثل الحد الأدنى للفئة الأولى، ونضيف لها طول الفئة (C=2) لنتحصل على الحد الأعلى للفئة الأولى، وبالتالي تكون الفئة الأولى هي: $[6-8[$ أو يكتب اختصارا: 6 - 8 ، وهكذا دواليك مع الفئة الموالية حيث يصبح حدها الأدنى هو الحد الأعلى للفئة السابقة (8) وحدها الأعلى هو 8+2=10، وبإضافة كل مرة طول الفئة (C=2) نحصل على فئة جديدة الى أن نكمل عدد الفئات K ونغطي جميع البيانات الموجودة.

3.تحديد مراكز الفئات

نقوم بحساب مركز الفئة x_i لكل فئة موجودة في جدول التوزيع التكراري، فمثلا الفئة رقم واحد تكون x_1 (i=1) وتحسب كمايلي:

– الفئة الأولى هي $[6-8[$ إذن:

$$x_1 = \frac{8+\ 6}{2} = 7$$

وهكذا دواليك حتى نكمل آخر فئة في جدول التوزيع التكراري

4. وضع جدول التوزيع التكراري: بعد تفريغ البيانات يصبح الجدول كما يلي:

العلامة	مركز الفئة x_i	تفريغ البيانات	عدد الطلبة n_i
08 – 06	07	//// ~~////~~ ~~////~~	14
10 – 08	09	// ~~////~~	07
12 – 10	11	/// ~~////~~ ~~////~~	13
14 – 12	13	/ ~~////~~	06
16 – 14	15	~~////~~ ~~////~~	10
18 – 16	17	////	04
20 – 18	19	~~////~~ ~~////~~	10
المجـــمـــوع			**64**

2.1. التكرار المطلق والتكرار النسبي

التكرار المطلق هو تكرار كل فئة n_i، بينما يمثل التكرار النسبي تكرار كل فئة مقسوم على مجموع التكرارات كمايلي:

$$f_i = \frac{n_i}{\sum_{i=1}^{k} n_i}$$

حيث f_i : التكرار النسبي n_i : التكرار المطلق للفئة i $\sum_{i=1}^{k}$: مجموع التكرارات من الفئة i الى الفئة k

مثال04-01: بالرجوع الى المثال السابق نجد:

f_i	n_i	**العلامة**
0.22=64/14	14	08 — 06
0.11=64/07	07	10 — 08
0.20=64/13	13	12 — 10
0.09=64/06	06	14 — 12
0.16=64/10	10	16 — 14
0.06=64/04	04	18 — 16
0.16=64/10	10	20 — 18
1=64/64	**64**	**المجموع**

ونحصل على التكرار النسبي المئوي بضرب التكرار النسبي في مائة.

3.1. التكرارات التجميعية المطلقة والتكرارات التجميعية النسبية

1.3.1. التكرارات التجميعية المطلقة: هناك نوعين من التكرارات التجميعية المطلقة:

أ. **التكرارات التجميعية الصاعدة المطلقة:** يرمز لها بالرمز $N_i \uparrow$ وهي عبارة عن تراكم للتكرارات حسب الفئة المراد حسابها لها، وتحسب بالعلاقة التالية:

$$N_i \uparrow = \sum_{i=1}^{k} n_i$$
$$or$$
$$N_i \uparrow = N_{i-1} \uparrow + n_i$$

فمثلا:

– إذا ما أخذنا المثال السابق نجد أن التكرار التجميعي الصاعد المطلق للفئة الثانية $N_2 \uparrow$ هو:

$$N_2 \uparrow = \sum_{i=1}^{2} n_i = n_1 + n_2 = 14 + 07 = 21$$

–إذا ما أخذنا المثال السابق أيضا نجد أن التكرار التجميعي الصاعد المطلق للفئة الثالثة $N_3 \uparrow$ هو:

$$N_3 \uparrow = N_2 \uparrow + n_3 = 21 + 13 = 347$$

وهكذا دواليك حتى نصل الى التكرار التجميعي الصاعد للفئة الأخيرة والذي يكون مساويا لمجموع التكرارات.

ب. **التكرارات التجميعية النازلة المطلقة:** يرمز لها بالرمز $N_i \downarrow$ وهي عبارة عن تراكم للتكرارات من الأسفل إلى الأعلى (أي صعودا من الفئة الأخيرة إلى الفئة الأولى) حسب الفئة المراد حسابها لها، وتحسب بالعلاقة التالية:

$$N_i \downarrow = N_{i+1} \downarrow + n_i$$

فمثلا:

- إذا ما أخذنا المثال السابق نجد أن التكرار التجميعي النزال المطلق للفئة الأخيرة $N_7 \downarrow$ هو:

$$N_7 \downarrow = N_8 \downarrow + n_7 = 00 + 10 = 10$$

- إذا ما أخذنا المثال السابق نجد أن التكرار التجميعي الصاعد المطلق للفئة الثالثة $N_6 \downarrow$ هو:

$$N_6 \downarrow = N_7 \downarrow + n_6 = 10 + 04 = 14$$

وهكذا دواليك حتى نصل الى التكرار التجميعي النازل للفئة الأولى والذي يكون مساويا لمجموع التكرارات.

هناك طريقة أخرى لحساب التكرارات التجميعية النازلة المطلقة وهي بالبدء من الفئة الأولى والتي يكون تكرارها التجميعي النازل المطلق هو مجموعة التكرارات، ثم ننقص في كل مرة التكرار المطلق المقابل لنحصل على التكرار التجميعي النازل للفئة الموالية. ولحساب كل التكرارات التجميعية النازلة المطلقة نستعين بالعلاقة التالية:

$$N_1 \downarrow = \sum_{i=1}^{k} n_i$$
$$N_{i+1} \downarrow = N_i \downarrow - n_i$$
$$i \neq 0$$

2.3.1. التكرارات التجميعية النسبية: كما للتكرارات التجميعية المطلقة نوعين، فإن للتكرارت التجميعية النسبية أيضا نوعين:

أ. **التكرارات التجميعية الصاعدة النسبية:** يرمز لها بالرمز $F_i \uparrow$ وهي عبارة عن تراكم للتكرارات النسبية حسب الفئة المراد حسابها لها، وتحسب بالعلاقة التالية:

$$F_i \uparrow = \sum_{i=1}^{k} f_i$$
$$or$$
$$F_i \uparrow = F_{i-1} \uparrow + f_i$$

كما توجد طريقة أخرى لحساب التكرارات التجميعية الصاعدة النسبية وهي بقسمة التكرارات التجميعية المطلقة الصاعدة على مجموع التكرارات كمايلي:

$$F_i \uparrow = \frac{N_i \uparrow}{\sum_{i=1}^{k} n_i}$$

ب. **التكرارات التجميعية النازلة النسبية:** يرمز لها بالرمز $F_i \downarrow$ وهي عبارة عن تراكم للتكرارات النسبية من الأسفل إلى الأعلى (أي صعودا من الفئة الأخيرة إلى الفئة الأولى) حسب الفئة المراد حسابها لها، وتحسب بالعلاقة التالية:

$$F_i \downarrow = F_{i+1} \downarrow + f_i$$

أو بالبدء من الفئة الأولى والتي يكون تكرارها التجميعي النازل النسبي هو مجموعة التكرارات النسبية، ثم ننقص في كل مرة التكرار النسبي المقابل لنحصل على التكرار التجميعي النازل النسبي للفئة الموالية. ولحساب كل التكرارات التجميعية النازلة النسبية نستعين بالعلاقة التالية:

$$F_1 \downarrow = \sum_{i=1}^{k} f_i$$
$$F_{i+1} \downarrow = F_i \downarrow - f_i$$
$$i \neq 0$$

مثال01-05: وبالرجوع للمثال السابق الخاص بعلامات الطلبة في مقياس الإحصاء 1 نجد:

العلامة	n_i	f_i	$N_i \uparrow$	$N_i \downarrow$	$F_i \uparrow$	$F_i \downarrow$
08 – 06	14	0.22	14	64	0.22	1.00
10 – 08	07	0.11	21	50	0.33	0.78
12 – 10	13	0.20	34	43	0.53	0.67
14 – 12	06	0.09	40	30	0.62	0.47
16 – 14	10	0.16	50	24	0.78	0.38
18 – 16	04	0.06	54	14	0.84	0.22
20 – 18	10	0.16	64	10	1.00	0.16
المجموع	**64**	1.00	-	-	-	-

- **حالات من الجدول التكرارية المفتوحة:** هناك ثلاث حالات للجدول التكرارية المفتوحة وهي:

جدول مفتوح من الأسفل

العلامة	n_i
أقل من 08	14
10 – 08	07
12 – 10	13
14 – 12	06
16 – 14	10
18 – 16	04

جدول مفتوح من الأعلى

العلامة	n_i
08 – 06	14
10 – 08	07
12 – 10	13
14 – 12	06
16 – 14	10
16 فأكثر	04

جدول مفتوح من الطرفين

العلامة	n_i
أقل من 08	14
10 – 08	07
12 – 10	13
14 – 12	06
16 – 14	10
16 فأكثر	04

2. العرض البياني للبيانات الإحصائية

بالإضافة إلى العرض الجدولي للبيانات هناك طرق أخرى أسهل وأبسط لعرض البيانات الإحصائية تمكن من التحليل والتفسير السريع للظاهرة المدروسة، هذه الطرق تعتمد على التمثيلات والخطوط البيانية تصنف حسب نوع البيانات والمتغيرات المدروسة.

1.2. العرض البياني في حالة متغير كمي متقطع (منفصل)

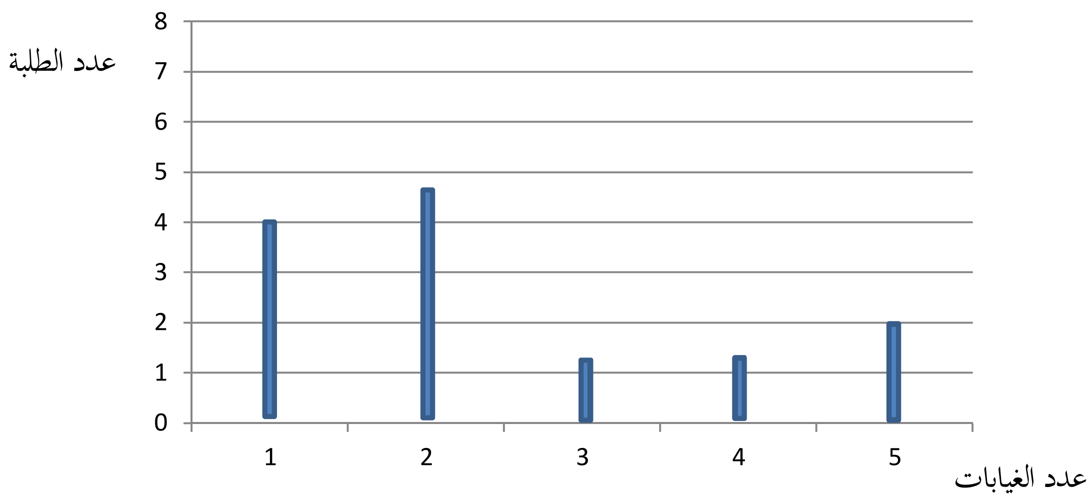
المتغير الكمي المتقطع هو المتغير الذي نحصل عليه من عملية العد، ويأخذ شكل أعداد طبيعية ولا يمكن تجزئة وحدة قياسه إلى وحدات أصغر، ولعرض هذا المتغير نميز مايلي:

1.1.2 العرض البياني بواسطة الأعمدة البسيطة: هي أعمدة غير متلاصقة؛ طولها يتناسب وقيمة التكـــرار المقابل للظاهرة المدروسة.

مثال01-06: يبين الجدول أدناه عدد غيابات 20 طالـــبا في مقياس الاقتصاد الجزئي1، موزعة كمايلي:

- المطلوب: تمثيل هذه البيانات بأعمدة بيانية.

الغيابات	عدد الطلبة
01	06
02	07
03	02
04	02
05	03



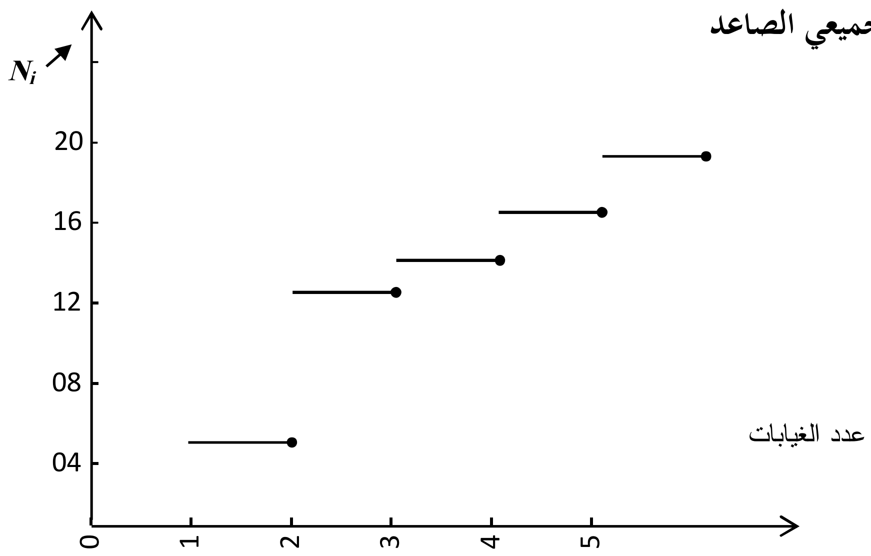
2.1.2 العرض البياني للتكرار التجميعي الصاعد والتكرار التجميعي النازل: هي قطع أفقيـــة مســـتقيمة متزايدة ومتصاعدة حسب التكرار التجميعي الصاعد، ومتناقصة ونازلة حسب التكرار التجميعي النازل.

مثال01-07: بالرجوع الى المثال السابق الخاص بعدد الغيابات في مقياس الاقتصاد الجزئي1، مثل التكرار التجميعي الصاعد والتكرار التجميعي النازل.

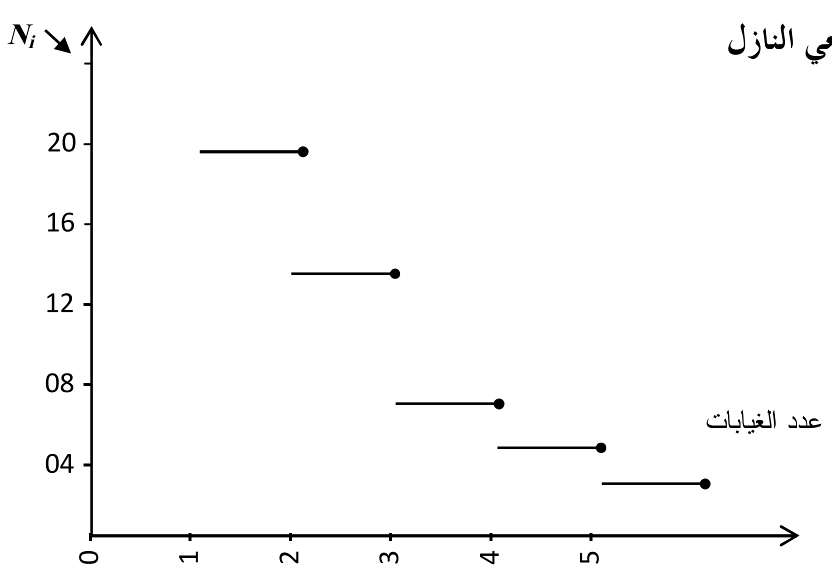
الحل: أولا نقوم بحساب كل من التكرارين التجمعيين الصاعد والنازل.

الغيابات	عدد الطلبة n_i	$N_i \uparrow$	$N_i \downarrow$
01	06	06	20
02	07	13	14
03	02	15	07
04	02	17	05
05	03	20	03

– تمثيل التكرار التجميعي الصاعد



– تمثيل التكرار التجميعي النازل



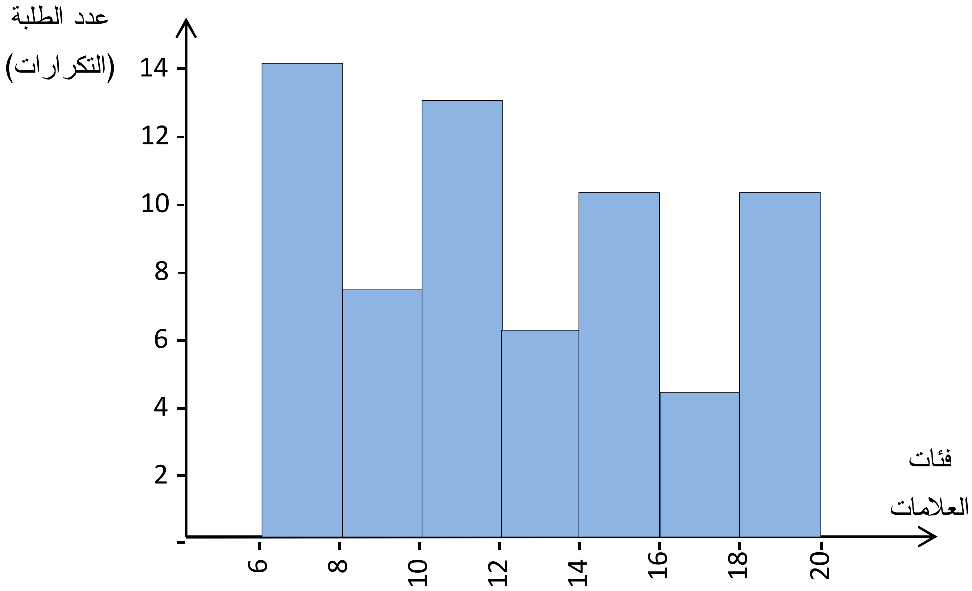
2.2. العرض البياني في حالة متغير كمي مستمر (متصل)

المتغير الكمي المستمر هو المتغير الذي نحصل عليه من عملية القياس، ولا يأخذ بالضرورة شكل أعداد طبيعية ويمكن تجزئة وحدة قياسه إلى وحدات أصغر، ولعرض هذا المتغير نميز مايلي:

1.2.2. العرض البياني بواسطة المدرج التكراري: المدرج التكراري هو عبارة مستطيلات متلاصقة جنبا الى جنبا، طولها هو تكرار الفئة المقابلة وعرضها هو طول الفئة.

لرسم المدرج التكراري هناك حالتين هما:

أ. حالة تساوي طول الفئة: بالرجوع للمثال01-03 الخاص بعلامات الطلبة في مقياس الاحصاء 1 يكون المدرج التكراري كمايلي:



ب. حالة عدم تساوي طول الفئة: في هذه الحالة قبل رسم المدرج التكراري نقوم أولا بتعديل التكرارات حتى تتناسب وأطوال الفئات المقابلة، وذلك استعانة بالعلاقة التالية:

$$n_i^{/} = \frac{n_i}{c_i}.c$$

حيث: $n_i^{/}$: تكرار الفئة المعدل الجديد — C_i : طول الفئة القديم

n_i : تكرار الفئة القديم — C : طول الفئة المقترح

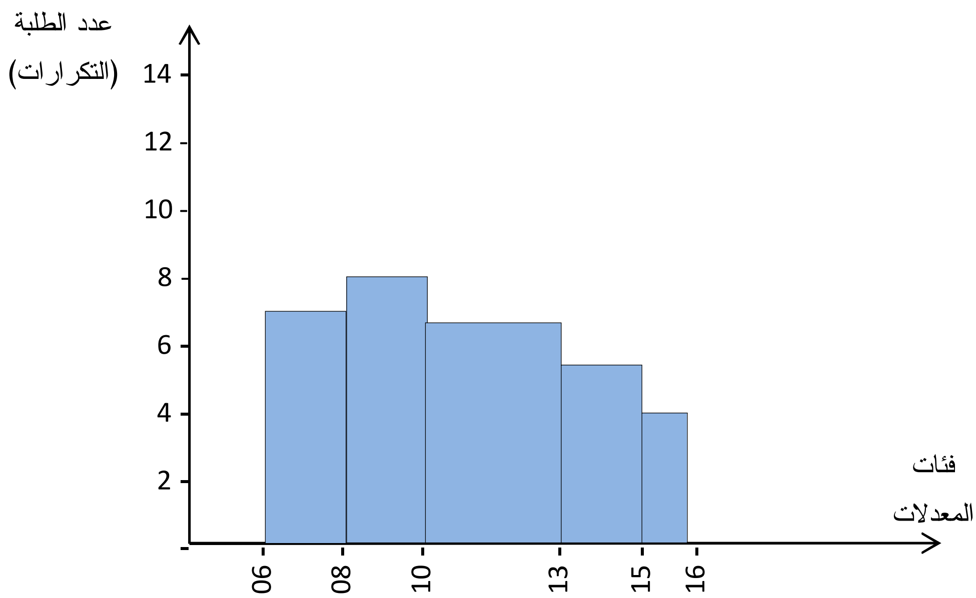
- نشير هنا أن طول الفئة المقترح، يعين من بين الأطوال الأكثر تكرارا في جدول التوزيع التكراري.

مثال08-01: يتكون فوج من أفواج طلبة السنة الأولى؛ ميدان العلوم الاقتصادية والتجارية وعلوم التسيير من 32 طالبا، كانت معدلاتهم خلال السداسي الأول ملخصة في الجدول أدناه. المطلوب رسم المدرج التكراري.

الحل: نلاحظ من الجدول أن طول الفئة غير متساوي، لهذا قبل رسم المدرج التكراري يجب تعديل التكرارات بما يتناسب وطول الفئة كمايلي: (أنظر الجدول)

معدل السداسي الأول	التكرار n_i	طول الفئة C_i	الطول المقترح C	التكرار المعدل (الجديد) $n_i^{/}$
08 – 06	07	02	02	07
10 – 08	08	02		08
13 – 10	10	03		6.67
15 – 13	05	02		05
16 – 15	02	01		04
المجموع	**32**	-	-	-

– رسم المدرج التكراري: (الفئات الأصلية بالتكرارات المعدلة)

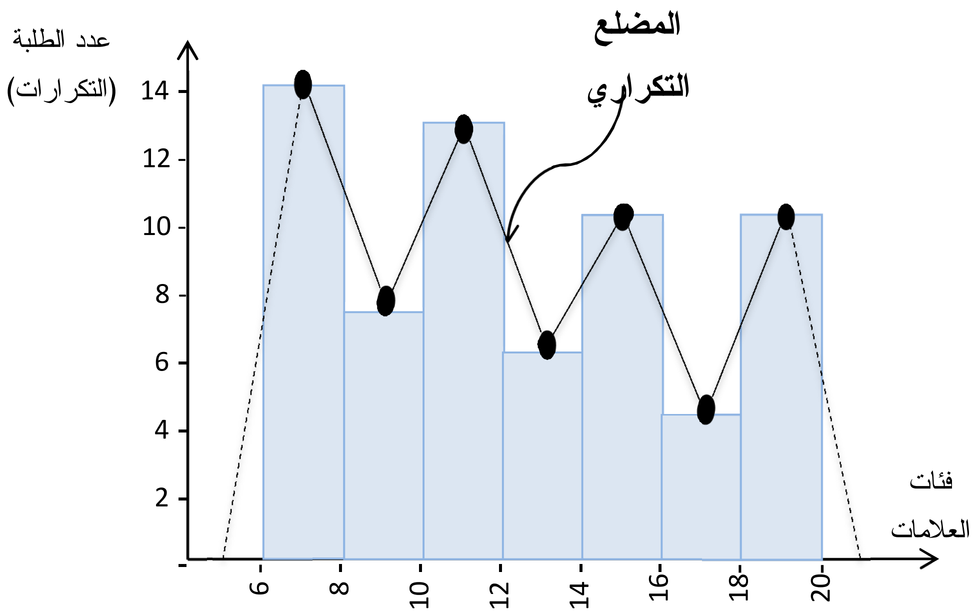


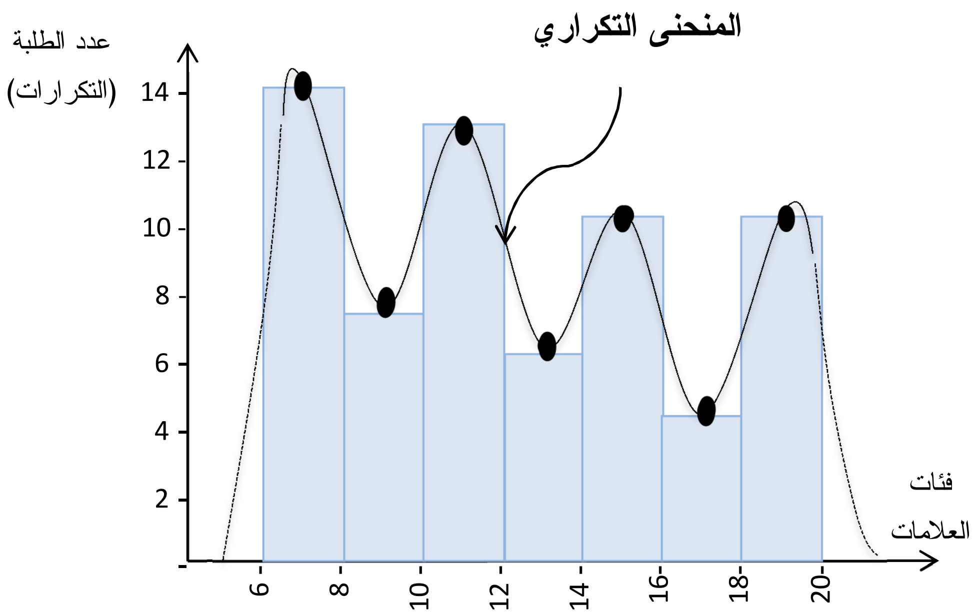
ملاحظة: نقوم بتعديل التكرارات في حالة الفئات غير المتساوية في الحالتين التاليتين:

- عند رسم المدرج التكراري؛
- عند تحديد الفئة المنوالية وحساب المنوال.

2.2.2. العرض البياني بواسطة المضلع التكراري والمنحنى التكراري: خطوط متصلة مع بعض، احداثياتها هي مراكز الفئات والتكرارات المقابلة، هذه الخطوط تكون مستقيمة في حالة المضلع التكراري ومنحنية في حالة المنحنى التكراري.

بالرجوع للمثال01-03 الخاص بعلامات الطلبة في مقياس الاحصاء 1 يكون المضلع والمنحنى التكرارين كمايلي:





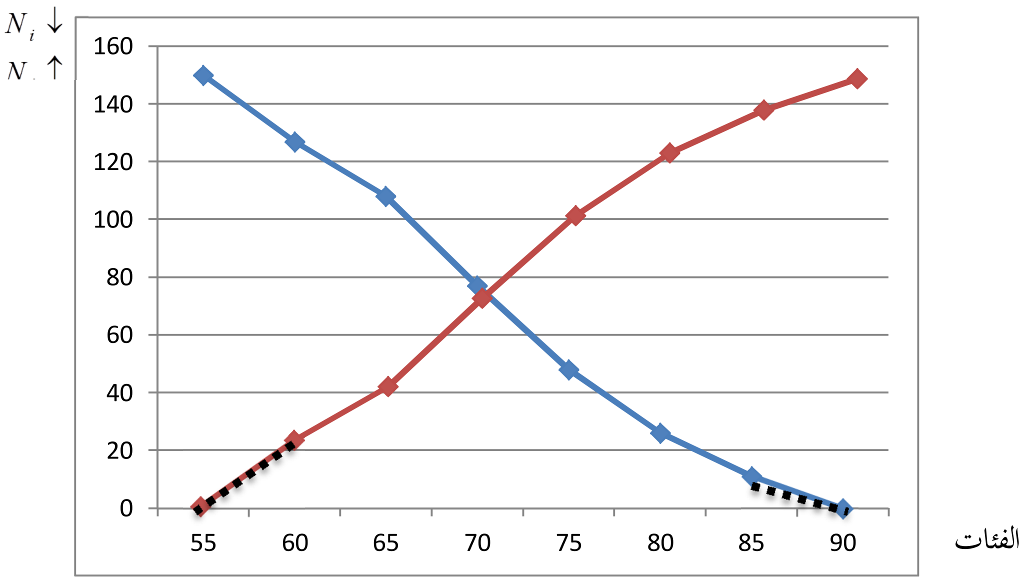
إن المساحة التي تقع تحت المضلع التكراري هي نفسها المساحة الواقعة تحت المدرج التكراري، وحتى نحافظ على المساحة التي تقع أسفل هذا المضلع، نفترض أن لهذا التوزيع فئتين افتراضيتين إحداهما في بدايته والأخرى في نهايته، تكرار كل منهما يساوي صفر، بحيث ننطلق في رسم المضلع من مركز الفئة الافتراضية الأولى (الفئة ما قبل الأولى)، وننتهي عند مركز الفئة الافتراضية الأخيرة.

3.2.2. العرض البياني للتكرار التجميعي الصاعد والتكرار التجميعي النازل: يرسم منحنى التكرار التجميعي الصاعد في المتغير الكمي المستمر اعتمادا على نقاط إحداثياتها الحدود العليا للفئات والتكرارات التجميعية الصاعدة المقابلة، بينما يرسم منحنى التكرار التجميعي النازل بنفس الطريقة، لكن اعتمادا على نقاط احداثياتها الحدود الدنيا للفئات والتكرارات التجميعية النازلة المقابلة.

يبين كل من منحنى التكرار التجميعي الصاعد ومنحنى التكرار التجميعي النازل شدة أو ضعف تطور الظاهرة المدروسة عن مستوى معين من مجال الدراسة، وتسمى نقطة تقاطع المنحنيين بالوسيط.

مثال09-01: تمثل البيانات أدناه أوزان مجموعة طلبة من طلبة السنة الأولى؛ ميدان العلوم الاقتصادية والتجارية وعلوم التسيير. المطلوب رسم منحنيي التكرار التجميعي الصاعد والتكرار التجميعي النازل على نفس المعلم.

$N_i \downarrow$	$N_i \uparrow$	n_i	**الوزن**
150	23	23	60 — 55
127	42	19	65 — 60
108	73	31	70 — 65
77	102	29	75 — 70
48	124	22	80 — 75
26	139	15	85 — 80
11	150	11	90 — 85
- -		**150**	**المجموع**



3.2. العرض البياني في حالة متغير نوعي (وصفي)

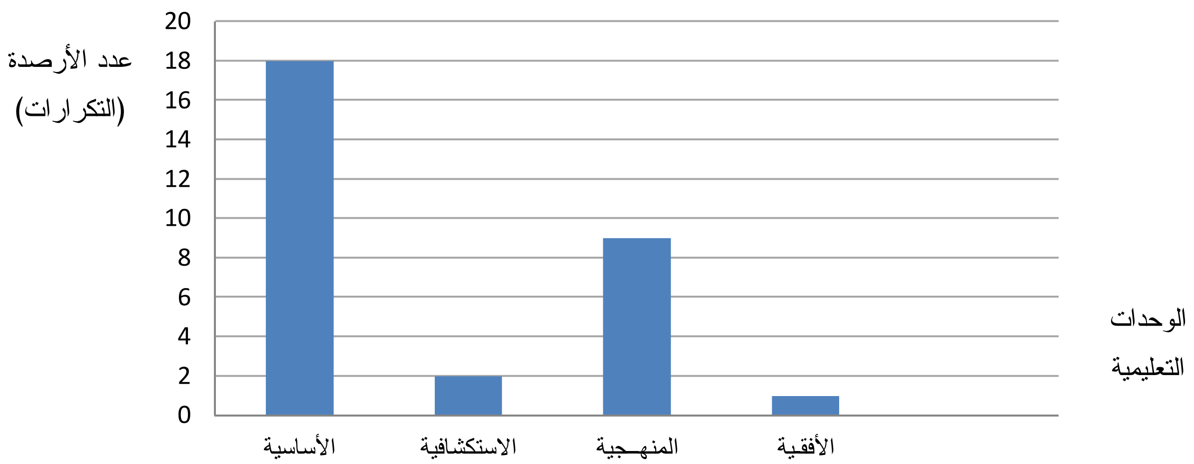
أكثر العروض والتمثيلات البيانية انتشارا هي الاعمدة البيانية والدوائر النسبية.

1.3.2. العرض البياني بواسطة الأعمدة البيانية: وهي مستطيلات غير متلاصقة؛ تفصلها مسافات متساوية، قاعدتها هي الصفة أو النوع المدروس في الظاهرة وأطوالها هي التكرارات المقابلة.

مثال10-01: يضم السداسي الأول من السنة الأولى ميدان العلوم الاقتصادية والتجارية وعلوم التسيير أربع وحدات تعليمية، كل وحدة لها رصيد معين كمايلي:

الوحدات التعليمية	عدد الأرصدة
الأساسية	18
الاستكشافية	02
المنهجية	09
الأفقية	01
المجموع	30

- **المطلوب:** مثل البيانات بأعمدة بيانية.



2.3.2. العرض البياني بواسطة الدائرة النسبية: وهي دائرة مقسمة إلى عدة أجزاء كل جزء يقابل زاوية مركزية تتناسب مع التكرارات المقابلة لكل خاصية أو صفة من الصفات المدروسة.

لحساب زاوية كل صفة أو جزء نستخدم العلاقة التالية:

$$\frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360°$$

مثال01-11: عند الانتقال من السنة الأولى الى السنة الثانية ميدان العلوم الاقتصادية والتجارية وعلوم التسيير ، هناك أربع شعب، كل شعبة يمكنها استعاب عدد من الطلبة نلخصه في الجدول التالي:

الشعبة	عدد الطلبة
علوم تجارية	250
علوم التسيير	340
علوم المالية والمحاسبة	240
علوم اقتصادية	250
المجموع	1080

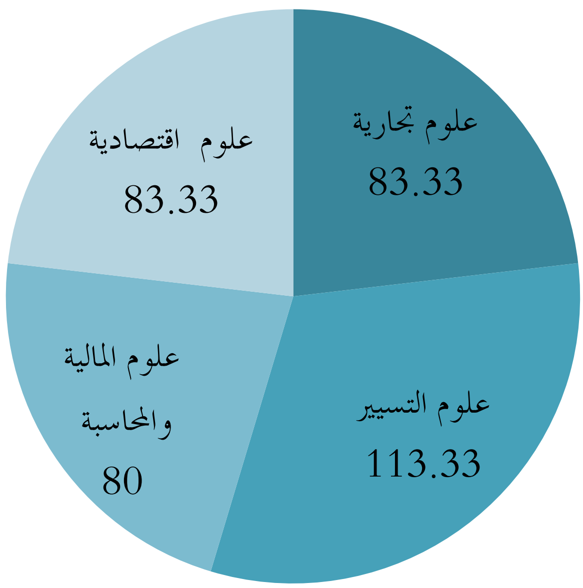
– **المطلوب**: مثل البيانات بدائرة نسبية.

الحل: لتمثيل البيانات بواسطة الدائرة النسبية يجب أولا حساب الزوايا المركزية لكل شعبة باستخدام العلاقة:

$$\frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360°$$

الشعبة	عدد الطلبة	الزوايا المركزية
علوم تجارية	250	83.33^0
علوم التسيير	340	113.33^0
علوم المالية والمحاسبة	240	80^0
علوم اقتصادية	250	83.33^0
المجموع	1080	360^0

رسم الدائرة:



الفصل الثاني

مقاييس النزعة المركزية

الفصل الثاني
مقاييس النزعة المركزية

عند دراسة الظواهر والقيم نلاحظ أن غالبيتها تتموقع وتتموضع قريبة من بعضها البعض، بل نجدها تتجمع وتتركز حول قيم معينة، هذه القيم هي ما يسمى بمقاييس النزعة المركزية أو المتوسطات.

وتستخدم هذه المقاييس لتلخيص البيانات عدديا، إذ أنها تعتبر قيم نموذجية أو مثالية للبيانات، كما أنها تستخدم لوصف مجموعات البيانات المختلفة ومقارنتها.

من أهم هذه المقاييس أو المتوسطات: المتوسط الحسابي، المتوسط الهندسي، المتوسط التوافقي، الوسيط، المنوال والتي سنتطرق لها في هذا الفصل.

1. المتوسط الحسابي

المتوسط الحسابي $\overline{X}$ من أهم مقاييس النزعة المركزية وأكثرها استخداما في النواحي التطبيقية، يمكن حسابه للبيانات المبوبة وغير المبوبة كمايلي:

1.1.المتوسط الحسابي في البيانات غير المبوبة

ونميز طريقتين لحسابه هما:

1.1.1. الطريقة المباشرة: فإذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \ldots, X_n$ فإن متوسطها الحسابي هو:

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots\ldots + X_n}{n} = \frac{\sum_{i=1}^{n} Xi}{n}$$

مثال01-02:إذا كانت العلامات التي تحصل عليها طالب في خمس مواد هي: 10، 08، 13، 16، 09. أحسب متوسط علامات هذا الطالب.

الحل:

$$\overline{X} = \frac{\sum_{i=1}^{5} X_1}{5} = \frac{10 + 08 + 13 + 16 + 9}{5}$$

$$\overline{X} = 11.20$$

اذن: متوسط علامات هذا الطالب هو 11.20، أي كأنه أخذ في كل مقياس علامة تقدر بـــــ 11.20.

2.1.1. طريقة الانحرافات عن وسط فرضي: تقوم هذه الطريقة على فرض قيمة من البيانات أو من خارجها؛ موجبة أو سالبة A، ثم نضيفها لمتوسط الفروق حولها. فإذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \dots, X_n$ وكانت القيمة A هي الوسط الفرضي المختار فإن المتوسط الحسابي يعطى بالعلاقة التالية:

$$\overline{X} = A + \overline{d}$$

حيث: $$d_i = X_i - A \quad \text{و} \quad \overline{d} = \frac{\sum_{i=1}^{n} d_i}{n}$$

مثال02-02: بالرجوع للمثال02-01؛ أحسب متوسط علامات الطالب باستخدام الانحرافات عن وسط فرضي.

الحل: علامات المواد هي: 10، 08، 13، 16، 09.

$$\overline{X} = A + \frac{\sum_{i=1}^{n} d_i}{n}$$

ليكن: A=10

$$d_1 = 10 - 10 = 0 \quad d_2 = 08 - 10 = -2 \quad d_3 = 13 - 10 = 3 \quad d_4 = 16 - 10 = 6 \quad d_5 = 09 - 10 = -1$$

ومنه:

$$\overline{X} = A + \frac{\sum_{i=1}^{n} d_i}{n} = 10 + \frac{0 + (-2) + 3 + 6 + (-1)}{5} = 10 + \frac{6}{5} = 11.20$$

اذن: متوسط علامات هذا الطالب هو 11.20.

- **المتوسط الحسابي المرجح:** يستخدم في حالة البيانات التي تكون لها أوزان أو معاملات (تكرارات)، فإذا كانت لدينا القيم: $X_1, X_2, X_3, \dots, X_n$، ولها تكرارات(معاملات): $W_1, W_2, W_3, \dots, W_n$، فإن المتوسط الحساب لها يعطي بالعلاقة:

$$\overline{X} = \frac{W_1X_1 + W_2X_2 + W_3X_3 + \dots\dots\dots\dots n_n x_n}{W_1 + W_2 + W_3 + W_4 + \dots\dots\dots\dots + W_n} = \frac{\sum_{i=1}^{n} W_i X_i}{\sum W_i}$$

مثال02-03: في امتحان الأعمال الموجهة لمقياس المحاسبة 1 تحصل فوج من طلبة السنة الأولى على العلامات التالية (الامتحان على 10 نقاط): -أحسب متوسط علامات الطلبة.

العلامات	عدد الطلبة
05.75	06
07.25	07
08.50	06
09.00	04
09.25	03
10.00	05

الحل:

$$\overline{X} = \frac{\sum_{i=1}^{n} W_i X_i}{\sum W_i} = \frac{05.75 * 06 + 07.25 * 07 + 08.50 * 06 + 09.00 * 04 + 09.25 * 03 + 10.00 * 05}{06 + 07 + 06 + 04 + 03 + 05}$$

$$= \frac{250}{31} = 8.06$$

2.1.المتوسط الحسابي في البيانات المبوبة

تعتمد طريقة حساب المتوسط الحسابي لبيانات مبوبة (بيانات مقسمة إلى فئات) على مراكز الفئات.

1.2.1. الطريقة المباشرة: ويحسب فيها المتوسط الحسابي بالعلاقة التالية:

$$\overline{X} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + ...n_nx_n}{n_1 + n_2 + n_3 + n_4 + ...n_n} = \frac{\sum_{i=1}^{k} n_ix_i}{\sum_{i=1}^{k} n_i}$$

حيث: x_i : مركز الفئة n_i التكرار المقابل

2.2.1. طريقة الانحرافات عن وسط فرضي: ويحسب فيها المتوسط الحسابي بنفس الطريقة التي تناولناها في البيانات غير المبوبة تقريبا.

$$\overline{X} = A + \overline{d}$$

حيث: $d_i = x_i - A$ و $\overline{d} = \frac{\sum_{i=1}^{k} n_id_i}{\sum_{i=1}^{k} n_i}$

مثال02-04 : بالرجوع للمثال السابق (المثال01-03) الخاص بعلامات الطلبة في مقياس الإحصاء 1يكون المتوسط الحسابي كمايلي:

- **الطريقة المباشرة:** $\overline{X} = \frac{\sum_{i=1}^{k} n_ix_i}{\sum_{i=1}^{k} n_i}$

-**طريقة الانحرافات عن وسط فرضي:** $\overline{X} = A + \frac{\sum_{i=1}^{k} n_id_i}{\sum_{i=1}^{k} n_i}$ **حيث:** $A=13$

نستخدم الجدول في الحساب أفضل وأسهل:

العلامة	n_i	x_i	n_ix_i	A	d_i	n_id_i
08—06	14	07	98		6-	84-
10—08	07	09	63		4-	28-
12—10	13	11	143		2-	26-
14—12	06	13	78	13	0	0
16—14	10	15	150		2	20
18—16	04	17	68		4	16
20—18	10	19	190		6	60
المجموع	**64**	-	**790**	-	-	**42-**

ومنه:

$$\overline{X}=\frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}=\frac{790}{64}=12.34$$

$$\overline{X}=A+\frac{\sum_{i=1}^{k} n_i d_i}{\sum_{i=1}^{k} n_i}=13+\frac{-42}{64}=13+(-0.66)=12.34$$

أي أن متوسط ما يحصل عليه كل طالب في مقياس الإحصاء 1 كعلامة هو 12.34.

إضافة للطريقة المباشرة وطريقة الانحرافات عن الوسط الفرضي، هناك طرق أخرى يعتمد فيها لحساب المتوسط الحسابي على التكرار النسبي كمايلي:

$$\overline{X}=\sum_{i=1}^{k} f_i x_i$$

أو:

$$\overline{X}=A+\sum_{i=1}^{k} f_i d_i$$

- **خواص المتوسط الحسابي**

- يعتبر المتوسط الحسابي أبسط مقاييس المركزية حسابا وأكثرها استخداما؛
- يأخذ المتوسط الحسابي بعين الاعتبار جميع قيم الظاهرة المدروسة؛
- يتأثر المتوسط الحسابي بالقيم المتطرفة والشاذة؛
- لا يمكن حسابه في الجداول التكرارية المفتوحة؛
- مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفر $\sum_{i=1}^{n}(x_i-\overline{X})=0$

2.المتوسط الهندسي

المتوسط الحسابي مقياس لا يعتمد عليه بشكل دائم في وصف جميع الظواهر وصفا سليما، بل هناك ظواهر يفضل فيها استخدام متوسطات أخرى لها من القدرة على وصف الظاهرة بشكل صحيح؛ ففي الحالات التي تكون فيها قيم الظاهرة المدروسة عبارة عن نسب أو معدلات، أي في الحالات التي نرغب فيها بدراسة معدل تغيرات ظاهرة ما فإن أفضل متوسط لذلك هو المتوسط الهندسي.

1.2المتوسط الهندسي في البيانات غير المبوبة

إذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \ldots, X_n$فإن متوسطها الهندسي هو الجذر النوني لحاصل ضرب هذه القيم في بعضها البعض.

$$G = \sqrt[n]{x_1.x_2x_3......x_n}$$

أو عن طريق ادخال اللوغاريتم العشري كمايلي:

$$LogG = \frac{1}{N}\left[\log X_1 + \log X_2 + \log X_3 + ... + \log X_n\right]$$

$$LogG = \frac{1}{n}\sum_{i=1}^{n} LogX_i$$

$$G = 10^{\log G}$$

مثال05-02: لو أخذنا نفس المثال الذي استخدمناه في حساب المتوسط الحسابي لبيانات غير مبوبة، فإن المتوسط الهندسي يحسب كمايلي:

- العلامات هي: 10، 08، 13، 16، 09 ومنه:

$$G = \sqrt[n]{x_1.x_2x_3......x_n} = \sqrt[5]{10 * 08 * 13 * 16 * 09} = \sqrt[5]{149760} = 10.84$$

أو:

$$LogG = \frac{1}{N}\left[\log X_1 + \log X_2 + \log X_3 + ...+ \log X_n\right]$$

$$LogG = \frac{1}{5}\left[\log 10 + \log 08 + \log 13 + \log 16 + \log 09\right] = \frac{5.1754}{5} = 1.0351$$

$$G = 10^{1.0351} = 10.84$$

أمافي حالة البيانات الإحصائية التي تكون لها أوزان أو معاملات (تكرارات)، فإن المتوسط الهندسي في هذه الحالة يعرف بالمتوسط الهندسي المرجح ويحسب بالعلاقة:

$$G = \sqrt[N]{x_1^{n_1}.x_2^{n_2}.x_3^{n_3}......x_3^{n_4}}$$

$$LogG = \frac{1}{\sum_{i=1}^{n} w_i}\left[W_1 \log X_1 + W_2 \log X_2 + W_3 \log X_3 + ... + W_n \log X_n\right]$$

أو:

$$LogG = \frac{\sum_{i=1}^{n} W_i \log X_i}{\sum_{i=1}^{n} W_i}$$

$$G = 10^{\log G}$$

مثال02-06: بالتطبيق على نفس المثال الخاص بعلامات امتحان الأعمال الموجهة لمقيـــــاس المحاسبة 1(مثال استخدمناه للمتوسط الحسابي المرجح)نحصل على المتوسط الهندسي المرجح كمايلي:

العلامات	عدد الطلبة
05.75	06
07.25	07
08.50	06
09.00	04
09.25	03
10.00	05

$$LogG = \frac{1}{\sum_{i=1}^{n} w_i}\left[W_1 \log X_1 + W_2 \log X_2 + W_3 \log X_3 + \dots\dots\dots + W_n \log X_n\right]$$

$$LogG = \frac{1}{31}\left[6\log 5.75 + 7\log 7.25 + 6\log 8.50 + 4\log 9 + 3\log 9.25 + 5\log 10\right]$$

$$LogG = \frac{27.8723}{31} = 0.8991$$

$$G = 10^{0.8991} = 7.93$$

2.2.المتوسط الهندسي في البيانات المبوبة

يحسب المتوسط الهندسي في بيانات مبوبة وفق العلاقة التالية:

$$LogG = \frac{1}{\sum_{i=1}^{k} n_i}\left[n_1 \log x_1 + n_2 \log x_2 + n_3 \log x_3 + \dots\dots\dots + n_k \log x_4\right]$$

$$LogG = \frac{\sum_{i=1}^{k} n_i Logx_i}{\sum_{i=1}^{k} n_i}$$

$$G = 10^{\log G}$$

مثال02-07:اليك جدول التوزيع التكرار الموالي، والمطلوب حساب المتوسط الهندسي.

$n_i \log x_i$	$\log x_i$	x_i	n_i	**العلامة**
11.8314	0.8451	07	14	08—06
6.6797	0.9542	09	07	10—08
13.5381	1.0414	11	13	12—10
6.6837	1.1139	13	06	14—12
11.7609	1.1761	15	10	16—14
4.9218	1.2304	17	04	18—16
12.7875	1.2788	19	10	20—18
68.2031	-	-	**64**	**المجموع**

$$LogG = \frac{\sum_{i=1}^{k} n_i Logx_i}{\sum_{i=1}^{k} n_i} = \frac{68.2031}{64} = 1.0657$$

$$G = 10^{1.0657} = 11.63$$

- **خواص المتوسط الهندسي**

- يدخل في حساب جميع القيم ولكنه أقل تأثر بالقيم المتطرفة من المتوسط الحسابي؛
- لا يمكن حسابه من الجداول التكرارية المفتوحة ،ولا في حالة وجود قيمة سالبة أو معدومة؛
- يستخدم بأكثر واقعية عند وصف الظواهر النسبية (حساب نسب، معدلات...)؛
- قيمة المتوسط الهندسي لأي ظاهرة أصغر دائما من قيمة المتوسط الحسابي $G < \overline{X}$.

3.المتوسط التوافقي

يتميز المتوسط الهندسي عن المتوسط الحسابي في أنه يفضل استخدامه في الحالات التي تكون فيها قيم الظاهرة المدروسة عبارة عن نسب أو معدلات، بينما يفضل في الظواهر التي تقيس وتدرس معدلات السرعة ومتوسط الأسعار، ومتوسط الكثافة السكانية، متوسط آخر غيرهما هو المتوسط التوافقي.

1.3.المتوسط التوافقي في البيانات غير المبوبة

إذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \ldots, X_n$ فإن متوسطها التوافقي هو مقلـــوب المتوســـط الحسابي لمقاليب هذه القيم، ويكتب بالعلاقة التالية:

$$H = \frac{n}{\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + \ldots\ldots\ldots \frac{1}{X_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}$$

مثال08-02: بالتطبيق على المثال الذي استخدمناه في حساب المتوسط الحسابي والمتوسط الهندسي لبيانات غير مبوبة، فإن المتوسط الهندسي يحسب كمايلي:

- العلامات هي: 10، 08، 13، 16، 09. ومنه

$$H = \frac{5}{\frac{1}{10} + \frac{1}{8} + \frac{1}{13} + \frac{1}{16} + \frac{1}{9}} = \frac{5}{0.4755} = 10.52$$

أما في حالة البيانات الإحصائية التي تكون لها أوزان أو معاملات W_i (تكرارات)، فإن المتوسط التوافقي في هذه الحالة يعرف بالمتوسط التوافقي المرجح ويحسب بالعلاقة:

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

مثال02-09: بالتطبيق على نفس المثال الخاص بعلامات امتحان الأعمال الموجهة لمقياس المحاسبة 1(مثال استخدمناه للمتوسط الحسابي المرجح)نحصل على المتوسط التوافقي المرجح كمايلي:

العلامات	عدد الطلبة
05.75	06
07.25	07
08.50	06
09.00	04
09.25	03
10.00	05

$$H=\frac{31}{\frac{6}{5.75}+\frac{7}{7.25}+\frac{6}{8.50}+\frac{4}{9}+\frac{3}{9.25}+\frac{5}{10}+}=\frac{31}{3.9836}=7.78$$

2.3.المتوسط التوافقي في البيانات المبوبة

يحسب المتوسط التوافقي في بيانات مبوبة اعتمادا على مراكز الفئات x_i والتكرارات المقابلة n_i

وفق العلاقة التالية:

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

مثال02-10: اليك جدول التوزيع التكرار الموالي،

والمطلوب حساب المتوسط التوافقي.

العلامة	n_i	x_i	$\frac{n_i}{x_i}$
08—06	14	07	2.0000
10—08	07	09	0.7778
12—10	13	11	1.1818
14—12	06	13	0.4615
16—14	10	15	0.6667
18—16	04	17	0.2353
20—18	10	19	0.5263
المجموع	**64**	-	**5.8494**

$$H=\frac{\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n}\frac{n_i}{x_i}}=\frac{64}{5.8494}=10.94$$

– **خواص المتوسط التوافقي**

- يأخذ بعين الاعتبار جميع القيم وتأثره بالقيم الشاذة أقل من تأثر المتوسط الحسابي؛
- لا يمكن حسابه في حالة وجود بيانات معدومة؛
- يعطي نتائج أكثر دقة في حالة الظواهر التي تهتم بمتوسطات الأسعار والسرعة؛
- قيمته دائما أقل من قيمة المتوسط الهندسي والمتوسط الحسابي $H<G<\overline{X}$.

4.الوسيط

تبين لنا عند دراستنا للمتوسط الحسابي والهندسي والتوافقي أنها تتأثر بالقيم المتطرفة والشادة، مما ينعكس على واقعية القيم والظواهر، لهذا وجد متوسط آخر أكثر واقعية ودلالة للحصول على فكرة عامة حول البيانات التي بها قيم متطرفة، هذا المتوسط يسمى بالوسيط Me.

والوسيط هو القيمة المشاهدة التي تتوسط البيانات بعد ترتيبها تصاعدياً أو تنازلياً، فعدد القيم (المفردات) قبله مساوي لعدد القيم بعده.

1.4.الوسيط في البيانات غير المبوبة

لتحديد قيمة الوسيط في بيانات غير مبوبة نتبع الخطوات التالية:

- يجب أولا أن نرتب البيانات ترتيبا تصاعديا أو تنازليا.
- نبحث عن الوسيط وذلك بحساب رتبته: $(n+1)/2$، سواء كانت السلسلة زوجية أو فردية.
- اذا كانت الرتبة هي n فإننا نميز حالتين :

أ- لما n عدد طبيعي: هنا الوسيط هو مباشرة ترتيب تلك القيمة.

ب- لما n عدد عشري : فإن الوسيط يحسب كمايلي:

الوسيط = ما يقابل الجزء الصحيح في الرتبة n+(ما يقابل الجزء الصحيح+1 في الرتبة n - ما يقابل الجزء الصحيح في الرتبة n) x الجزء العشري المتبقي من الرتبة n

مثال02-11: لتكن لدينا القيم التالية ؛التي تمثل درجات الحرارة في عطلة الشتاء بولاية جيجل:

12، 15، 22، 22، 16، 18، 25، 18، 24، 25 ، 9، 10، 17، 15، 12

- **المطلوب:** تحديد الوسيط.

الحل: لتحديد الوسيط في هذه البيانات نتبع مايلي: **لدينا:** $n=15$

- **يجب أولا أن نرتب البيانات ترتيبا تصاعديا أو تنازليا:** نرتب البيانات تصاعديا.

9، 10، 12، 12، 15، 15، 16، 17، 18، 18، 22، 22، 24، 25، 25

- **نبحث عن الوسيط وذلك بحساب رتبته:** $\frac{(n+1)}{2}=\frac{15+1}{2}=8$

- n **عدد طبيعي:** $n=8$، إذن الوسيط هو العدد ذي الرتبة **8** في السلسلة وهو:**17**

نقول: 50% من درجات الحرارة أكبر من 17 درجة.

50% من درجات الحرارة أكبر من 17 درجة.

مثال02-12: لو أخذنا المثال السابق وأنقصنا منه ثلاث قيم، فإن الوسيط سيكون بعد ترتيب البيانات تصاعديا هو: 9، 10، 12، 12، 15، 15، 17، 22، 22، 24، 25، 25

- **لدينا:** $n=15$
- **نبحث عن الوسيط وذلك بحساب رتبته:** $\frac{(n+1)}{2}=\frac{12+1}{2}=6.5$
- n **عدد عشري** : إذن الوسيط يحسب كمايلي:

الوسيط = ما يقابل الجزء الصحيح في الرتبة n +(ما يقابل الجزء الصحيح+1 في الرتبة n - ما يقابل الجزء الصحيح في الرتبة n) x الجزء العشري المتبقي من الرتبة n

$$Me=15+(17-15)*0.5=16$$

نقول: 50% من درجات الحرارة أكبر من 16 درجة.

50% من درجات الحرارة أقل من 16 درجة.

2.4.الوسيط في البيانات المبوبة

هناك طريقتين حسابيتين لإيجاد الوسيط في بيانات مقسمة إلى فئات هما:

1.2.4. الوسيط بالاعتماد على التكرار التجميعي الصاعد $N_i\uparrow$: ونتبع مايلي:

- نحسب التكرار التجميعي الصاعد؛
- نحدد ترتيب الوسيط وهو عبارة عن نصف مجموع التكرارات $n/2$؛
- نحدد الفئة الوسيطية؛ وهي الفئة التي تقابل التكرار التجميعي الصاعد الذي يساوي ترتيب الوسيط أو أكبر منه مباشرة؛
- نحدد ونحسب الوسيط بتطبيق العلاقة التالية:

$$Me = L_1 + \left(\frac{\frac{n}{2}-\sum N_1}{n_{me}}\right).c$$

حيث: C : طول الفئة. L_1 : الحد الأدنى للفئة الوسيطية

$\sum N_1$: التكرار التجميعي الصاعد السابق للفئة الوسيطية n_{me} تكرار الفئة الوسيطية $n/2$: رتبة الوسيط.

2.2.4. الوسيط بالاعتماد على التكرار التجميعي النازل $N_i\downarrow$: ونتبع مايلي:

- نحسب التكرار التجميعي النازل؛
- نحدد ترتيب الوسيط وهو عبارة عن نصف مجموع التكرارات $n/2$؛
- نحدد الفئة الوسيطية ، وهي الفئة التي تقابل التكرار التجميعي النازل الذي يساوي ترتيب الوسيط أو أكبر منه مباشرة.
- نحدد ونحسب الوسيط بتطبيق العلاقة التالية:

$$Me = L_1 + \left(\frac{N_{me}\downarrow - \frac{n}{2}}{n_{me}}\right).c$$

حيث: C : طول الفئة. L_1 : الحد الأدنى للفئة الوسيطية
$N_{me}\downarrow$: التكرار التجميعي النازل للفئة الوسيطية n_{me} تكرار الفئة الوسيطية $n/2$: رتبة الوسيط.

مثال02-13: بالرجوع للمثال السابق (مثال01-03)الخاص بعلامات الطلبة في مقياس الإحصاء1 يكون الوسيط كمايلي:

العلامة	n_i	$N_i\uparrow$	$N_i\downarrow$
08–06	14	14	64
10–08	07	21	50
12–10	13	34	43
14–12	06	40	30
16–14	10	50	24
18–16	04	54	14
20–18	10	64	10
المجموع	**64**	-	-

- **بالاعتماد على التكرار التجميعي الصاعد:**
  - نحسب التكرار التجميعي الصاعد؛
  - ترتيب الوسيط: $\frac{n}{2}=32$؛
  - الفئة الوسيطية: $12 - 10$؛

ومنه:

$$Me = L_1 + \left(\frac{\frac{n}{2} - \sum N_1}{n_{me}}\right).c$$

$$Me = 10 + \left(\frac{32-21}{13}\right)2 = 11.69$$

نقول: 50% من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أكبر من 11.69.

50% من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أصغر من 11.69.

- **بالاعتماد على التكرار التجميعي النازل:**

- نحسب التكرار التجميعي النازل؛
- ترتيب الوسيط: $\frac{n}{2}=32$؛
- الفئة الوسيطية: $10-12$؛

ومنه:

$$Me=L_1+\left(\frac{N_{me}\downarrow-\frac{n}{2}}{n_{me}}\right).c$$

$$Me=10+\left(\frac{43-32}{13}\right).2=11.69$$

نقول: 50% من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أكبر من 11.69.

50% من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أصغر من 11.69.

- **خواص الوسيط**

- لا يتأثر بالقيم المتطرفة وبالتالي فإنه يعتبر أصلح المقاييس عند وجود مثل هذه القيم؛
- يمكن إيجاده الوسيط بيانيا.
- يمكن حسابه من الجداول التكرارية المفتوحة.

5.المقاييس الشبيهة بالوسيط

بما أنه يمكن تقسيم بيانات أي ظاهرة إلى عدة أقسام متساوية وليس إلى قسمين فقط كما يفعل الوسيط، فإنه يمكن التعامل مع هذه القيم بنفس طريقة التعامل مع الوسيط، حيث نميز:

- الربيعات Q: هي القيم التي تقسم مجموع البيانات إلى أربعة أقسام متساوية، كل قسم يأخذ 25% .
- الخميسات S: هي القيم التي تقسم مجموع البيانات إلى خمسة أقسام متساوية، كل قسم يأخذ 20% .
- العشيرات D: هي القيم التي تقسم مجموع البيانات إلى عشرة أقسام متساوية، كل قسم يأخذ 10% .
- الميئينات P: هي القيم التي تقسم مجموع البيانات إلى مئة قسم متساوي، كل قسم يأخذ 1% .

1.5.المقاييس الشبيهة بالوسيط في البيانات غير المبوبة

- يجب أولا أن نرتب البيانات ترتيبا تصاعديا (وليس تنازليا)؛
- نحسب التكرار التجميعي الصاعد؛

– نبحث عن المقياس الشبيه بالوسيط المراد حسابه لهذه البيانات وذلك بحساب رتبته، والرتب هي كمايلي:

– **الربيعات Q:** $NQ_M = \left(\frac{M(n+1)}{4}\right)$ – **العشيرات D:** $ND_M = \left(\frac{M(n+1)}{10}\right)$

– **الخميسات S:** $NS_M = \left(\frac{M(n+1)}{5}\right)$ – **الميئينات P:** $NP_M = \left(\frac{M(n+1)}{100}\right)$

– بعد تحديد الرتبة نقوم بنفس خطوات تحديد الوسيط.

مثال14-02: بالعودة للمثال السابق الخاص بدرجات الحرارة في عطلة الشتاء بولاية جيجل، ونحسب الربيع الأول Q_1، الخميس الثالث S_3، العشير الثانيD_2 ، والميئين السايع والثلاثونP_{37} نجد:

17، 10، 12، 12، 15، 22، 24، 25، 9، 25، 15، 22

– يجب أولا أن نرتب البيانات ترتيبا تصاعديا:

9، 10، 12، 12، 15، 15، 17، 22، 22، 24، 25، 25

– نبحث عن المقياس الشبيه بالوسيط المراد حسابه لهذه البيانات وذلك **بحساب رتبته**، والرتب هي كمايلي:

$$NQ_1 = \left(\frac{1(n+1)}{4}\right) = \frac{13}{4} = 3.25 \qquad ND_2 = \left(\frac{2(n+1)}{10}\right) = \frac{26}{10} = 2.6$$

$$NS_3 = \left(\frac{3(n+1)}{5}\right) = \frac{39}{5} = 7.8 \qquad NP_{37} = \left(\frac{37(n+1)}{100}\right) = \frac{481}{100} = 4.81$$

ومنه:

– **الربيع الأول:**

$$Q_1 = 12 + (12 - 12) * 0.25 = 12$$

نقول: 75% من درجات الحرارة أكبر من 12 درجة.

25% من درجات الحرارة أقل من 12 درجة.

– **الخميس الثالث:**

$$S_3 = 17 + (22 - 17) * 0.8 = 21$$

نقول: 40% من درجات الحرارة أكبر من 21 درجة.

60 % من درجات الحرارة أقل من 21 درجة.

- **العشير الثاني:**

$$D_2 = 10 + (12 - 10) * 0.3 = 10.6$$

نقول: 60% من درجات الحرارة أكبر من 10.6 درجة.

40 % من درجات الحرارة أقل من 10.6 درجة.

- **الميئين السابع والثلاثون:**

$$P_{37} = 12 + (15 - 12) * 0.81 = 14.43$$

نقول: 63% من درجات الحرارة أكبر من 14.43 درجة.

37 % من درجات الحرارة أقل من 14.43 درجة.

2.5.المقاييس الشبيهة بالوسيط في البيانات المبوبة

- في البيانات المبوبة نعتمد فقط على التكرار التجميعي الصاعد $N_i \uparrow$ ؛
- نحسب التكرار التجميعي الصاعد؛
- نحدد ترتيب المقياس الشبيه بالوسيط المراد حسابه كمايلي:

  - الربيعات **Q:** $MQ_M = \left(\frac{M(n)}{4}\right)$
  - الخميسات **S:** $MS_M = \left(\frac{M(n)}{5}\right)$
  - العشيرات **D:** $MD_M = \left(\frac{M(n)}{10}\right)$
  - الميئينات **P:**...... $MP_M = \left(\frac{M(n)}{100}\right)$

- نحدد فئة المقياس الشبيه بالوسيط المراد حسابه، وهي التي تقابل التكرار المتجمع الصاعد الذي يساوي ترتيب المقياس الشبيه بالوسيط المراد حسابه أو أكبر منه مباشرة؛
- نحدد ونحسب المقياس الشبيه بالوسيط المراد حسابه بتطبيق العلاقة السابقة الخاصة بالوسيط كمايلي:

$$B = L_1 + \left(\frac{B^* - \sum N_1}{n_B}\right).c$$

حيث: B : المقياس الشبيه بالوسيط المراد حسابه .

B^*: رتبة المقياس الشبيه بالوسيط المراد حسابه.

مثال02-15: بالرجوع للمثال السابق (الفصل الأول؛ الجدول في الصفحة 05)الخاص بعلامات الطلبة في مقياس الإحصاء1 يكون الربيع الثالث Q_3، الخميس الثالث S_2 كمايلي:

العلامة	n_i	$N_i \uparrow$
08—06	14	14
10—08	07	21
12—10	13	34
14—12	06	40
16—14	10	50
18—16	04	54
20—18	10	64
المجموع	**64**	-

الحل:

- نحسب التكرار التجميعي الصاعد؛
- نحدد ترتيب المقياس الشبيه بالوسيط المراد حسابه كمايلي:

- الربيع الثالث Q_3 **:**

$$MQ_3 = \left(\frac{3(n)}{4}\right) = \frac{3*64}{4} = 48$$

- الفئة الربيعية الثالثة: 14 – 16؛

ومنه: $Q_3 = 14 + (\frac{48-40}{10}).2 = 15.6$

نقول: 25% من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أكبر من 15.6.

75 % من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أقل من 15.6.

- الخميس الثالث S_2 **:**

$$S_2 = \left(\frac{2*64}{5}\right) = \frac{128}{5} = 25.6$$

- الفئة الخميسية الثانية: 10 – 12؛

ومنه: $S_2 = 12 + (\frac{25.6-21}{13}).2 = 12.71$

نقول: 60 % من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أكبر من 12.71.

40 % من الطلبة علاماتهم في مقياس الإحصاء 1 الحرارة أقل من 12.71.

❖ **الوسيط و المقاييس الشبيه به بيانيا**

- الوسيط هو تقاطع منحنى التكرار التجميعي الصاعد مع منحنى التكرار التجميعي النازل.
- الوسيط في منحنى التكرار التجميعي الصاعد أو التكرار التجميعي النازل هو تمثيل لرتبته n/2 واسقاطها على المنحنى.
- الوسيط في المدرج التكراري يحدد انطلاقا من حساب المساحة الكلية للمدرج وقسمتها على 2 (مساحة المدرج هي عبارة عن مجموع مساحة المستطيلات التي يتكون منها)، ثم نقوم بجمع مساحة كل مستطيل تلوى الأخر حتى تنطبق مع نصف مساحة المدرج، النقطة التي تنطبق عليها مجاميع مساحة المستطيلات ونصف مساحة المدرج الكلية هي الوسيط.
- المقاييس الشبيه بالوسيط ترسم على منحنى التكرار التجميعي الصاعد فقط، وذلك من خلال تمثيل الرتبة واسقاطها على المنحنى.

6.المنوال

المنوال Mo هو القيمة الأكثر تكراراً من غيرها في مجموعة من البيانات.

1.6.المنوال في البيانات غير المبوبة

المنوال Mo هو القيمة الأكثر شيوعا أو تكرار في مجموعة القيم (سواء كانت البيانات كيفية أو كمية)، والمنوال قد يكون وحيد القيمة كما قد يكون هناك أكثر من منوال لنفس التوزيع.

مثال16-02: لتكن لديك سلسلة البيانات التالية، والمطلوب تحديد المنوال في كل منها.

- السلسلة أ: 12، 10، 12، 9، 3، 4، 5
- المنوال هو: $Mo= 12$
- السلسلة ب: 12، 10، 12، 9، 3، 4، 5 ،9
- المنوال هو: $Mo_1= 12$ و $Mo_2= 9$
- السلسلة ج: 12، 10، 12، 9، 3، 4، 5 ،9 ، 9
- المنوال هو: $Mo_1=9$
- السلسلة د: 12، 10، 13، 9، 3، 4، 5 ،19

المنوال هو: لا يوجد

- السلسلة هـــ: 12، 10، 12، 9، 4، 4 ،9 ، 10

المنوال هو: لا يوجد لان كل الأرقام تكررت مرتين ولا يوجد من بينها رقم أكثر تكرار.

2.6.المنوال في البيانات المبوبة

يمكن إيجاد المنوال من الرسم، وذلك من خلال رسم المدرج التكراري للفئة المنوالية (الفئة التي يقابلها أكبر تكرار) وللفئتين التي قبلها والتي بعدها، ثم نقوم بعد ذلك بإيصال نهاية المستطيل للفئة المنوالية من الناحية اليسرى بنهاية المستطيل للفئة التي بعدها من الناحية اليسرى، كذلك نهاية المستطيل للفئة المنوالية من الجهة اليمنى بنهاية المستطيل للفئة التي قبلها من الجهة اليمنى، ومن نقطة تقاطع المستقيمين نترل عمودا على المحور الأفقي فتكون نقطة تقاطع هذا العمود مع المحور الأفقي هي قيمة المنوال.

كما يمكن حساب المنوال بالاعتماد على الفرق بين تكرار الفئة المنوالية والفئتين التي قبلها والتي بعدها وفق طريقة بيرسون (الفروق)كمايلي:

$$Mo = M_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2}.c$$

حيث

C : طول الفئة. Δ_1 الفرق بين تكرار الفئة المنوالية والفئة التي قبلها

M_1 : الحد الأدنى للفئة المنوالية. Δ_2 الفرق بين تكرار الفئة المنوالية والفئة التي بعدها

n_i	الفئات
07	10—08
13	12—10
06	14—12
10	16—14
04	18—16
10	20—18
50	المجموع

مثال02-17: اليك البيانات التالية: أحسب المنوال وحدده بيانيا.

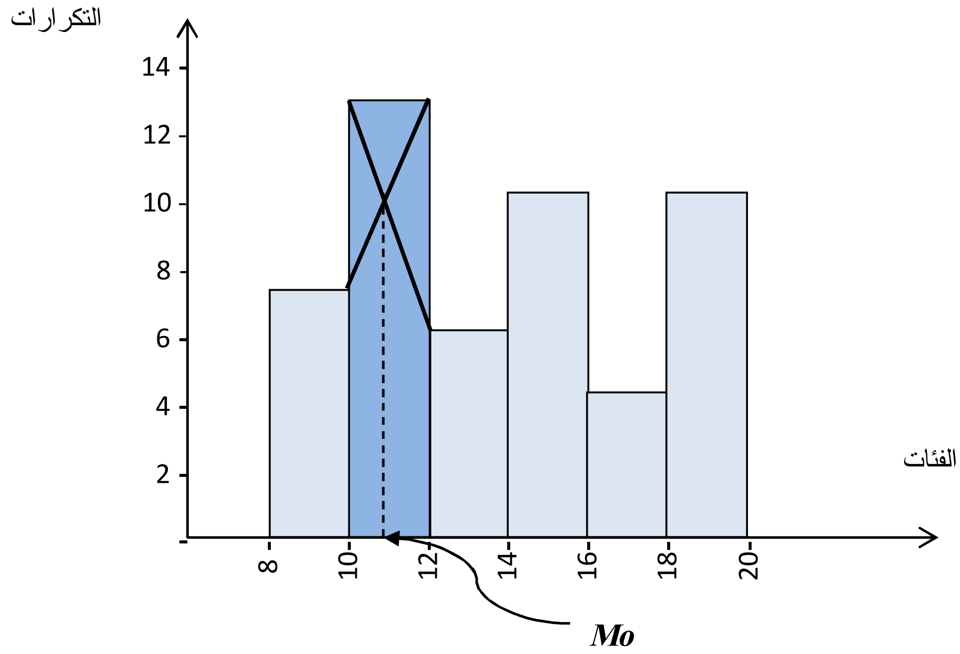
الحل:

- **المنوال حسابيا:** $$Mo = M_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2}.c$$
- الفئة المنوالية: 10 – 12

ومنه: $$Mo \; = 10 + \frac{13-7}{13-7+13-6}.2 = 10.62$$

اذن المنوال في هذه البيانات هو 10.62

- **المنوال بيانيا:** لتحديد المنوال بيانيا يكفي رسم مدرج تكراري للفئة المنوالية والفئة التي قبلها والفئة التي بعدها، ثم نقوم بمايلي(الشكل):



خواص المنوال

- لا يأخذ بعين الاعتبار جميع البيانات المعطاة، وبالتالي فهو لا يتأثر بالقيم المتطرفة؛
- يمكن حسابه في حالة الجداول التكرارية ذات الفئات المفتوحة؛
- يمكن تحديده بيانيا.
- يمكن أن يوجد أكثر من منوال لتوزيع واحد.

الفصل الثالث

مقــاييس التشتت

الفصل الثالث
مقـــاييس التشتت

رأينا في الفصل السابق أن مقاييس النزعة المركزية تسمح لنا بتحديد القيم المتوسطة للبيانات أو أماكن تجمعها، غير أن هذه المقاييس لا تعطي فكرة واضحة ووافية عن اختلاف قيم مفردات مجموعة البيانات، ولا تحقق كل الأغراض التي يريد الباحث الوصول إليها من دراسته وتحليلاته للبيانات. فالمتوسطات لا تبين طبيعة المجموعة ولا كيفية توزيع مفرداتها حول القيمة المركزية، ولتوضيح ما سبق نورد المثال التالي الخاص ببيانات تساقط الأمطار خلال الأسبوع الأول من شهر ديسمبر (المجموعة الأولى)، ثم خلال الأسبوع الأول من شهر جانفي (المجموعة الثانية) كمايلي:

- المجموعة الأولى: 10، 20، 30، 40، 50، 60، 70.
- المجموعة الثانية: 10، 30، 30، 40، 50، 50، 70.

نلاحظ أن المتوسط الحسابي لكل واحدة من هاتين المجموعتين هو 40، وكذلك قيمة وسيطها تساوي 40، لكن الفرق بين قيم مختلف وحدات المجموعتين المرتبة يختلف من المجوعة الأولى عنه في المجموعة الثانية، إذ يساوي 10 بين مختلف قيم عناصر المجموعة الأولى، ويتراوح بين الصفر و 20 بين قيم المجموعة الثانية. وبذلك فإن الاختلافات بين مفردات المجموعة الأولى أقل منه بين مفردات المجموعة الثانية، ويقال اصطلاحا أن المجموعة الأولى أقل تشتتا من المجموعة الثانية. إن هذه الحقيقية تبين أن مقاييس النزعة المركزية فعلا لا تعطي فكرة وافية عن اختلاف قيم الظواهر، لذلك فإن مقاييس النزعة المركزية لا بد أن تكون مصحوبة بمقاييس أخرى لقياس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها، تسمى هذه المقاييس بمقاييس التشتت.

ما معنى التشتت؟

تشتت بيانات ظاهرة ما يقصد به درجة أو مقدار التفاوت أو الاختلاف بين مفردات هذه الظاهرة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمتها قريبة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة. أما إذا كانت بيانات الظاهرة متباعدة وغير متجانسة فنقول أن مفردات الظاهرة مشتتة وغير مركزة. ويقاس تشتت البيانات بعدة مقاييس منها:

1.المدى

1.1.المدى في البيانات غير المبوبة

المدى R هو الفرق بين أكبر قيمة في البيانات وأصغر قيمة فيها، ويعطى بالعلاقة:

$$R = X_{Max} - X_{Min}$$

مثال01-03: اليك درجات الحرارة المسجلة في مدينة خلال يوم ما: 8، 9، 10، 9، 6، 12، 15، 11، 20، 24،3، -2، المطلوب تحديد المدى.

الحل: نرتب البيانات ترتيب تصاعديا أو تنازليا لتسهيل التعامل معها:

-2،3،6،، 8، 9، 9، 10، 11، 12، 15، 20، 24

إذن المدى هو: $R = X_{Max} - X_{Min} = 24-(-2) = 26$

2.1.المدى في البيانات المبوبة

يحسب المدى في البيانات المبوبة بعدة طرق؛ فمنها من يعتبره الفرق بين مركز الفئة الأخيرة ومركز الفئة الأولى، ومنها من يعتبره الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى.

من خلال المثال01-03 نلاحظ أن أغلب درجات الحرارة المسجلة متقاربة، باستثناء القيمة -2، وهي قيمة متطرفة أثرت على حساب المدى. بسبب هذا العيب فإن المدى كمقياس للتشتت لا يستخدم بشكل واسع إلا عندما نريد الحصول على نتائج سريعة لا يهمنا فيها الدقة بشكل كبير.

أما إذا أردنا أن نقلل من أثر القيم المتطرفة فإننا نقوم باستبعادها باستخدام الطرق التالية:

- المدى الربيعي = الربيع الثالث – الربيع الأول.
- المدى العشري = العشر التاسع – العشر الأول.
- المدى المئيني = المئين التاسع والتسعون – المئين الأول.

- **خواص المدى**

- يتصف المدى بسهولة حسابه.
- يعتمد في حساب على قيمتين فقط هما القيمة الكبرى والقيمة الصغرى،
- شديد التأثر بالقيم المتطرفة؛
- لا يمكن استعماله في حالة البيانات المبوبة التي تتضمن فئات مفتوحة.

2.المدى الربيعي

يحسب بالفرق بين الربيع الثالث و الربيع الأول، ويعطينا فكرة عن المجال الذي تنتشر فيه نصف عدد البيانات متوسطة القيمة، ويحسب بالعلاقة:

$$IQ = Q_3 - Q_1$$

– خصائص المدى الربيعي

- يمتاز بسهولة حسابه؛
- لا يتأثر بالقيم المتطرفة لأنه يستبعد نصف البيانات (ربع من كل جهة)؛
- يضم نصف مفردات المجتمع؛
- يستعمل في المقارنة بين توزيعين احتماليين أو أكثر.

3.التباين والانحراف المعياري

1.3.التباين

التباين V_X أو δ^2 هو أحد مقاييس التشتت وأكثرها استخداما في النواحي التطبيقية ، ويعبر عن متوسط مربعات انحرافات القيم عن وسطها الحسابي.

1.1.3.التباين في البيانات غير المبوبة

إذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \ldots, X_n$ فإن التباين يعطي بالعلاقة (الصيغة العادية):

$$V_X = \frac{(X_1-\overline{X})^2+(X_2-\overline{X})^2+(X_3-\overline{X})^2+\ldots\ldots+(X_n-\overline{X})^2}{n}$$

$$V_X = \frac{\sum_{i=1}^{n}(X_1-\overline{X})^2}{n}$$

كما يمكن حسابه بطريقة مختصرة كمايلي:

$$V_x = \frac{\sum_{i=1}^{n} X_i^2}{n} - \overline{X}^2$$

مثال03-02: يعمل بقسم التعليم الأساسي **15** أستاذ دائم، عدد سنوات الخبرة لديهم كما يلي :

10 12 11 6 14 13 10 8 6 9 12 14 7 13 5

- **المطلوب:** حساب التباين بطريقتين.

الحل:

- نحسب أولا المتوسط الحسابي:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{10+12+11+...+13+5}{15} = \frac{150}{15} = 10$$

ومنه:

$$V_X = \frac{(X_1-\overline{X})^2 + (X_2-\overline{X})^2 + (X_3-\overline{X})^2 + ... + (X_n-\overline{X})^2}{n}$$

$$V_X = \frac{(10-10)^2 + (12-10)^2 + (11-11)^2 + ... + (5-10)^2}{15} = \frac{130}{15} = 8.67$$

أما باستخدام الطريقة المختصرة فيكون التباين كمايلي:

$$V_x = \frac{\sum_{i=1}^{n} X_i^{\,2}}{n} - \overline{X}^2$$

$$V_x = \frac{10^2 + 12^2 + 11^2 + 6^2 + 14^2 + ... + 5^2}{15} - 10^2$$

$$V_x = 108.67 - 100 = 8.67$$

اذن: مقدار التشتت في سنوات خبرة هؤلاء الأساتذة هو **8.67**.

2.1.3.التباين في البيانات المبوبة

يحسب التباين في البيانات المبوبة بالعلاقة التالية (الصيغة العادية):

$$V_X = \frac{\sum_{i=1}^{k} n_i (x_i - \overline{X})^2}{\sum_{i=1}^{k} n_i}$$

أما باستخدام الصيغة المختصرة؛ فيحسب كمايلي:

$$V_x = \frac{\sum_{i=1}^{k} n_i x_i^{\,2}}{\sum_{i=1}^{k} n_i} - \overline{X}^2$$

مثال03-03: أوجد التباين للبيانات التالية:

الفئة	8-4	13-9	18-14	23-19	28-24	المجموع
التكرار	3	4	6	2	4	**19**

الحل:

- نحسب أولا المتوسط الحسابي:

$$\overline{X}=\frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}$$

$n_i x_i^2$	x_i^2	$n_i(x_i-\overline{X})^2$	$(x_i-\overline{X})^2$	$x_i-\overline{X}$	$n_i x_i$	x_i	n_i	**العلامة**
108	36	300	100	-10	18	06	03	08 - 04
484	121	100	25	-5	44	11	04	13 - 09
1536	256	0	0	0	96	16	06	18 - 14
882	441	50	25	5	42	21	02	23 - 19
2704	676	400	100	10	104	26	04	28 – 24
5714	-	**850**	-	-	**304**	-	**19**	**المجموع**

ومنه:

$$\overline{X}=\frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}=\frac{304}{19}=16$$

-حساب التباين بالصيغة العادية (نستعين بالجدول):

$$V_X=\frac{\sum_{i=1}^{k} n_i (x_i-\overline{X})^2}{\sum_{i=1}^{k} n_i}=\frac{850}{19}=44.74$$

-حساب التباين بالصيغة المختصرة (نستعين بالجدول):

$$V_x=\frac{\sum_{i=1}^{k} n_i x_i^2}{\sum_{i=1}^{k} n_i}-\overline{X}^2$$

$$V_x=\frac{5714}{19}-16^2=44.74$$

اذن: مقدار التشتت هذه البيانات هو **44.74**.

2.3. الانحراف المعياري

يعتمد التباين علي مجموع مربعات الانحرافات، وهو ما لا يتماشى مع وحدات قياس المتغير محل الدراسة، فإذا رجعنا إلى المثال03-02، نجد أن تباين سنوات الخبرة 8.67؛ فليس من المنطق عند تفسير هذه النتيجة أن نقول، تباين سنوات الخبرة هو8.67 سنة مربع، لأن وحدات قياس المتغير هو عدد السنوات، من أجل ذلك لجأ الإحصائيين إلى مقياس منطقي يأخذ في الاعتبار الجذر التربيعي للتباين يناسب وحدات قياس المتغير؛ هذا المقياس هو الانحراف المعياري δ_X. إذا الانحراف المعياري هو الجذر التربيعي الموجب للتباين، أي أن: $\delta_x = \sqrt{V_x}$

1.2.3. الانحراف المعياري في البيانات غير المبوبة

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{(X_1-\overline{X})^2+(X_2-\overline{X})^2+(X_3-\overline{X})^2+........+(X_n-\overline{X})^2}{n}}$$

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{\sum_{i=1}^{n}(X_1-\overline{X})^2}{n}}$$

الطريقة المختصرة:

$$\delta_X = \sqrt{V_x} = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n} - \overline{X}^2}$$

2.2.3. الانحراف المعياري في البيانات المبوبة

$$\delta_x = \sqrt{V_X} = \sqrt{\frac{\sum_{i=1}^{k} n_i(x_i-\overline{X})^2}{\sum_{i=1}^{k} n_i}}$$

الطريقة المختصرة:

$$\delta_X = \sqrt{V_x} = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{\sum_{i=1}^{k} n_i} - \overline{X}^2}$$

مثال03-04: من خلال المثال مثال03-03 وجدنا أن : $V_x = 44.74$ ومنه الانحراف المعياري يكون:

$$\delta_X = \sqrt{V_x} = \sqrt{44.74} = 6.69$$

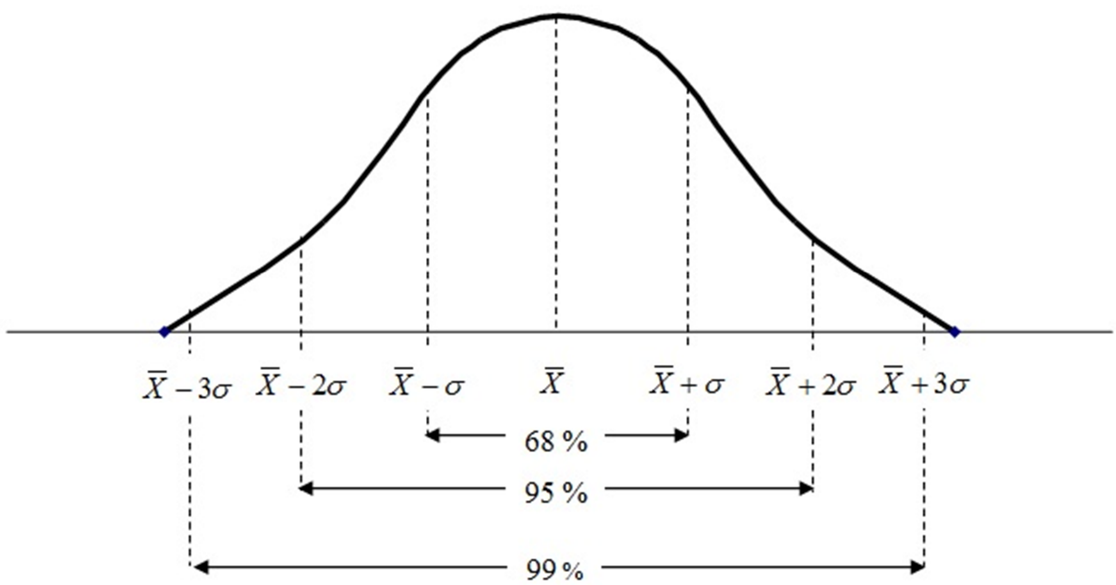
– خصائص الانحراف المعياري

- سهل التطبيق والحساب؛
- يستعمل في المقارنة بين الظواهر، حيث يعتمد عليه كمقياس لقياس تشتت الظواهر في حالة تساوي المتوسطات الحسابية لأنه يتأثر بها؛
- يأخذ الانحراف المعياري نفس وحدة القياس للمتغير الأصلي (كلغ، متر، لتر ...الخ) لذلك لا يمكن استخدامها كأساس للمقارنة بين تشتت توزيعين لهما وحدات قياس مختلفة؛
- لا يمكن إيجاده بالنسبة للتوزيعات التكرارية المفتوحة؛ سواء كانت من البداية أو النهاية أو من الطرفين؛
- الانحراف المعياري لقيمة ثابتة معدوم : $\delta_a = 0$
- $\delta_{aX} = a\delta_X$
- $\delta_{a+X} = \delta_a + \delta_X = \delta_X$

في حالة التوزيع الطبيعي (التوزيع المتناظر الذي يأخذ شكل الجرس) تكون:

- 68 % من القيم تقع على بعد انحراف معياري واحد من المتوسط الحسابي $\overline{X} = \pm\delta$ ؛
- 95 % من القيم تقع على بعد انحرافين معياريين من المتوسط الحسابي $\overline{X} = \pm 2\delta$ ؛
- 99 % من القيم تقع على بعد ثلاثة انحرافات معيارية من المتوسط الحسابي $\overline{X} = \pm 3\delta$.

والشكل التالي يلخص ذلك:



4.معامل الاختلاف

رأينا في خصائص الانحراف المعياري أنه إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها مختلفة، فإن المقارنة اعتمادا عليه ستكون غير منطقية وغير صحيحة، لهذا وجدت مقاييس أخرى تعرف بمقاييس التشتت النسبي تعتمد على فكرة مقارنة البيانات والظواهر في شكل نسبة مئوية اعتمادا على متوسطاتها، من أهم هذه المقاييس معامل الاختلاف CV.

يحسب معامل الاختلاف بقسمة الانحراف المعياري للبيانات على المتوسط الذي حسبت حوله، ويعطى بالعلاقة التالية:

$$CV = \frac{\delta_X}{\overline{X}} \times 100$$

مثال03-05: إذا كانت لديك البيانات التالية حول ظاهرتين من نفس النوع:

- الظاهرة أ: $\delta_x = 1.8$ $\overline{X} = 14$
- الظاهرة ب: $\delta_x = 1.3$ $\overline{X} = 10$

أي الظاهرتين أكثر تشتت؟

الحل:

إذا ما استخدمنا الانحراف المعياري في المقارنة، فإننا نلاحظ أن الظاهرة ب أكثر تجانسا من الظاهرة أ، أو أن الظاهرة أ أكثر تشتتا من الظاهرة ب، لكن بما أن المتوسط الحسابي الذي حسب من حوله الانحراف المعياري مختلف، فإن المقارنة هنا لا تصح، بل يجب الاحتكام الى مقياس يوحد نمط المقارنة وهو معامل الاختلاف CV كمايلي:

$$CV_{أ} = \frac{\delta_X}{\overline{X}} \times 100 \qquad CV_{ب} = \frac{\delta_X}{\overline{X}} \times 100$$

$$CV_{أ} = \frac{1.8}{14} \times 100 = 12.86\% \qquad CV_{ب} = \frac{1.3}{10} \times 100 = 13\%$$

إذن بعد استخدام معامل الاختلاف تبين أن الظاهرة ب أكثر تشتتا من الظاهرة أ لأن: $CV_{أ} < CV_{ب}$.

- **خصائص معامل الاختلاف**

ما يميز معامل الاختلاف عن بقية مقاييس التشتت أنه غير مرتبط بوحدة القياس المستعملة، و هذا ما يسمح بقياس ومقارنة تشتت ظاهرتين مختلفتين من حيث الوحدات المستعملة.

الفصل الرابع

مقــاييس الشكل

الفصل الرابع
مقاييس الشــكل

تسمح مقاييس النزعة المركزية ومقاييس التشتت بإعطاء فكرة سريعة عن خصائص توزيع البيانات ودرجة تجانسها أو تشتتها، إلا أن ما يعاب عليها هو نقص الدقة في التعرف على خصائص التوزيع من حيث انتشار البيانات على المنحى البياني وشكله مقارنة بالتوزيع أو الحالة الطبيعية (منحنى تتوزع فيه البيانات بشكل متناظر ويأخذ شكل الجرس)، الأمر الذي دعى إلى وجود مقاييس أخرى أكثر دقة وتفصيل تبين لنا شكل التوزيع وكيفية انتشار بياناته؛ هذه المقاييس تسمى بمقاييس الشكل أو ما مقاييس الالتواء والتفرطح.

قبل أن نتناول مقاييس الالتواء والتفرطح سنتطرق لما يعرف بـــــ "العزوم" التي نعتمد عليها كثيراً في تحديد مقاييس الشكل.

1.العـــزوم

تكون العزوم حول أي نقطة معينة؛ تتحد رتبتها بدرجة القوة (الأس). وتستخدم في إيجاد المعامل العزمي للالتواء والتفرطح.

1.1.العـــزوم حول نقطة الأصل

ونميز بين البيانات غير المبوبة والبيانات المبوبة كمايلي:

1.1.1.العـــزوم حول نقطة الأصل في البيانات غير المبوبة

إذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \ldots, X_n$ فإن العزم من الدرجة n يعطى بالعلاقة التالية:

$$m_n = \frac{\sum_{i=1}^{n} X_i^{\,n}}{n}$$

حيث أن: m_n : العزم من الدرجة n

X_i : القيمة i

n : درجة العزم

مثال04-01: إذا كانت لديك البيانات التالية: 5، 6، 4، 8، 2، 9، المطلوب حساب العزم الأول والثالث والرابع حول نقطة الأصل.

الحل:

لدينا:
$$m_n = \frac{\sum_{i=1}^{n} X_i^{\,n}}{n}$$

ومنه:

- العزم الأول:
$$m_1 = \frac{\sum_{i=1}^{n} X_i^{\,1}}{n} = \frac{5+6+4+8+2+9}{6} = \frac{34}{6} = 5.67$$

- العزم الثالث:
$$m_3 = \frac{\sum_{i=1}^{n} X_i^{\,3}}{n} = \frac{5^3+6^3+4^3+8^3+2^3+9^3}{6} = \frac{1654}{6} = 275.67$$

- العزم الرابع:
$$m_4 = \frac{\sum_{i=1}^{n} X_i^{\,4}}{n} = \frac{5^4+6^4+4^4+8^4+2^4+9^4}{6} = \frac{12850}{6} = 2141.67$$

2.1.1.العـــزوم حول نقطة الأصل في البيانات المبوبة

وتعطى بالعلاقة التالية:

$$m_n = \frac{\sum_{i=1}^{k} n_i x_i^{\,n}}{\sum_{i=1}^{k} n_i}$$

حيث أن: m_n : العزم من الدرجة n

x_i : مركز الفئة i

n : درجة العزم

مثال02-04: بالرجوع الى **المثال 03–03** أحسب: العزم الأول، العزم الخامس والعزم السابع.

الحل:

لدينا **:**
$$m_n = \frac{\sum_{i=1}^{k} n_i x_i^{\,n}}{\sum_{i=1}^{k} n_i}$$

ومنه يكون العزم الأول، العزم الخامس والعزم السابع كمايلي: (نستعين بالجدول لتسهيل الحساب).

$n_i x_i^7$	x_i^7	$n_i x_i^5$	x_i^5	$n_i x_i$	x_i	n_i	العلامة
839808	279936	23328	7776	18	06	03	08 - 04
77948684	19487171	644204	161051	44	11	04	13 - 09
1610612736	268435456	6291456	1048576	96	16	06	18 - 14
3602177082	1801088541	8168202	4084101	42	21	02	23 - 19
32127240704	8031810176	47525504	11881376	104	26	04	28 – 24
37418819014	**-**	**62652694**	-	**304**	-	**19**	المجموع

ومنه:

- العزم الأول:

$$m_1 = \frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i} = \frac{304}{19} = 16$$

- العزم الخامس:

$$m_5 = \frac{\sum_{i=1}^{k} n_i x_i^5}{\sum_{i=1}^{k} n_i} = \frac{62652694}{19} = 3297510.21$$

- العزم السابع:

$$m_4 = \frac{\sum_{i=1}^{k} n_i x_i^7}{\sum_{i=1}^{k} n_i} = \frac{37418819014}{19} = 1969411527.06$$

2.1.العـــزوم حول المتوسط الحسابي

يرمز لها بالرمز μ_n وتكون المتغيرات متمركزة حول المتوسط الحسابي.

1.2.1.العـــزوم حول المتوسط الحسابي في البيانات غير المبوبة

إذا كانت لدينا القيم التالية: $X_1, X_2, X_3, \ldots, X_n$ فإن العزم من الدرجة n حول المتوسط الحسابي يمثل متوسط انحرافات قيم التوزيع عن المتوسط الحسابي، ويعطى بالعلاقة التالية:

$$\mu_n = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^n}{\sum n}$$

حيث أن: μ_n : العزم حول المتوسط الحسابي من الدرجة n

X_i : القيمة i

$\overline{X}$: المتوسط الحسابي للبيانات

n : درجة العزم

مثال04-03: بالرجوع إلى **المثال04-01**؛ أحسب العزم الأول والثالث.

الحل: لدينا سلسلة البيانات 5، 6، 4، 8، 2، 9

ولدينا :

$$\mu_n = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^n}{\sum n}$$

قبل حساب العزوم المطلوبة، نقوم بحساب المتوسط الحسابي كمايلي:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{5+6+4+8+2+9}{6} = \frac{34}{6} = 5.67$$

ومنه:

- العزم الأول:

$$\mu_1 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^1}{\sum n} = \frac{(5-5.67)+(6-5.67)+(4-5.67)+(8-5.67)+(2-5.67)+(9-5.67)}{6}$$

$$\mu_1 = -0.003$$

– العزم الثالث:

$$\mu_3 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^3}{\sum n} = \frac{(5-5.67)^3+(6-5.67)^3+(4-5.67)^3+(8-5.67)^3+(2-5.67)^3+(9-5.67)^3}{6}$$

$$\mu_3 = -0.80$$

2.2.1.العـــزوم حول المتوسط الحسابي في البيانات المبوبة

وتعطى بالعلاقة التالية:

$$\mu_n = \frac{\sum_{i=1}^{k} n_i\left(x_i - \overline{X}\right)^n}{\sum_{i=1}^{k} n_i}$$

حيث أن: μ_n : العزم حول المتوسط الحسابي من الدرجة n

x_i : مركز الفئة i

n : درجة العزم

مثال04-04: من خلال معطيات المثال 03–03 أحسب: العزم الأول، العزم الخامس.

الحل:

لدينا :
$$\mu_n = \frac{\sum_{i=1}^{k} n_i \left(x_i - \overline{X}\right)^n}{\sum_{i=1}^{k} n_i}$$

قبل أن نحسب العزوم المطلوبة، نقوم بحساب المتوسط الحسابي كمايلي:

$$\overline{X} = \frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i} = \frac{304}{19} = 16$$

ومنه يكون العزم الأول والعزم الخامس كمايلي: (نستعين بالجدول لتسهيل الحساب).

العلامة	n_i	x_i	$n_i x_i$	$\left(x_i - \overline{X}\right)$	$n_i\left(x_i - \overline{X}\right)$	$\left(x_i - \overline{X}\right)^5$	$n_i\left(x_i - \overline{X}\right)^5$
08 - 04	03	06	18	-10	-30	-100000	-300000
13 - 09	04	11	44	-5	-20	-3125	-12500
18 - 14	06	16	96	0	0	0	0
23 - 19	02	21	42	5	10	3125	6250
28 – 24	04	26	104	10	40	100000	400000
المجموع	**19**	-	**304**	**-**	**0**	**-**	**93750**

ومنه:

- العزم الأول:

$$\mu_1 = \frac{\sum_{i=1}^{k} n_i \left(x_i - \overline{X}\right)}{\sum_{i=1}^{k} n_i} = \frac{0}{19} = 0$$

- العزم الخامس:

$$\mu_5 = \frac{\sum_{i=1}^{k} n_i \left(x_i - \overline{X}\right)^5}{\sum_{i=1}^{k} n_i} = \frac{93750}{19} = 4934.21$$

2. تحديد شكل التوزيع

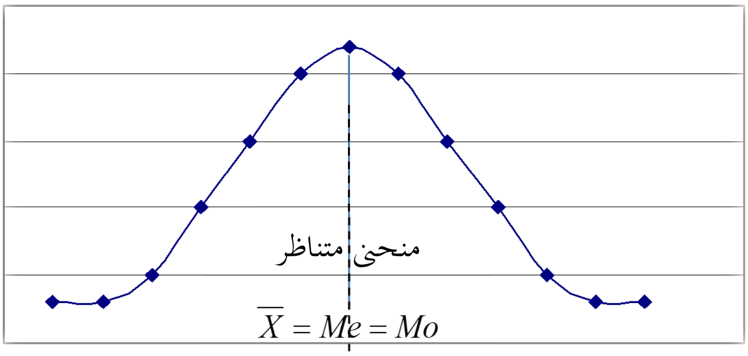
يتحدد شكل التوزيع اعتمادا على مقاييس الالتواء ومعامل التفرطح.

1.2. الالتواء

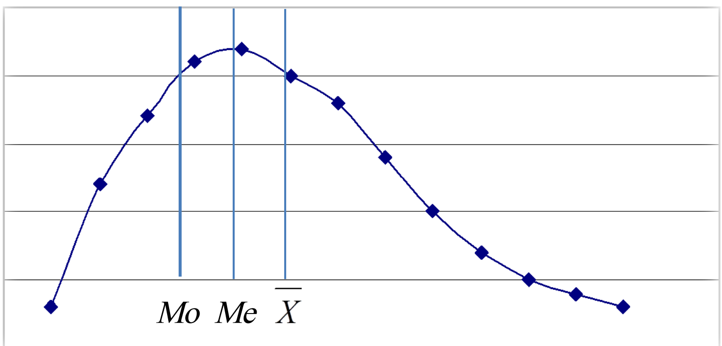
عند تساوى المتوسطات الثلاث $\left(\overline{X}=Me=Mo\right)$ يكون شكل التوزيع معتدلا (البيانات تتوزع طبيعيا أي تنتشر بشكل متناظر حول قيمة وسطية)، ويعتبر التوزيع الطبيعي توزيع هام في الدراسات الإحصائية، ويعد منحناه ونادر الوقوع، لأن أغلب منحنيات الظواهر التي نحصل عليها عادة ما تكون ملتوية ومائلة نحو جهة معينة أو قريبة من التماثل والاعتدال.

1.1.2. أشكال الالتواء: تقيس مقاييس الالتواء مدى التواء وميلان التوزيع نحو جهة معينة، حيث نميز ثلاث أشكال أساسية للتوزيعات نلخصها فيمايلي:

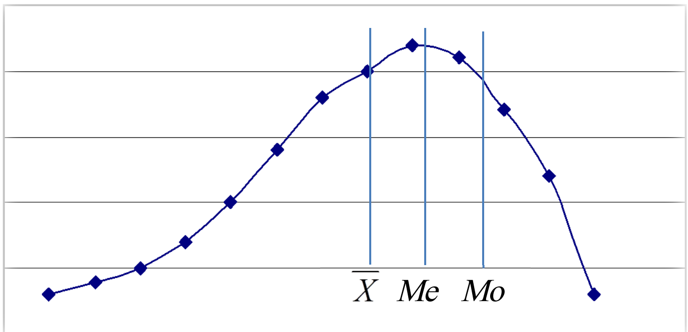
أ. **التوزيع المتماثل:** وتتوزع فيه البيانات بشكل متناظر بالنسبة الى نقطة هي الوسيط، حيث تأخذ المتوسطات الثلاث العلاقة: $\left(\overline{X}=Me=Mo\right)$



ب.**التوزيع موجب الالتواء(مائل نحو اليمين):** وتتركز فيه البيانات في الفئات الصغرى من التوزيع، حيث تأخذ المتوسطات الثلاث العلاقة: $\left(\overline{X}>Me>Mo\right)$



ج. التوزيع سالب الالتواء(مائل نحو اليسار): وتتركز فيه البيانات في الفئات الكبرى من التوزيع، حيث تأخذ المتوسطات الثلاث العلاقة: $\left(\overline{X} < Me < Mo\right)$



2.1.2. مقاييس الالتواء: تتعدد مقاييس الالتواء نذكر أهمها فيمايلي:

أ. معامل بيرسون الأول للالتواء: ويحسب بالعلاقة التالية:

$$SK_1 = \frac{\overline{X} - Mo}{\sigma}$$

حيث أن: SK_1 : معامل بيرسون الأول للالتواء δ الانحراف المعياري في للتوزيع.

$\overline{X}$: المتوسط الحسابي للتوزيع Mo المنـــــوال في التوزيع

ب. معامل بيرسون الثاني للالتواء: ويحسب بالعلاقة التالية:

$$SK_2 = \frac{3\left(\overline{X} - Me\right)}{\sigma}$$

حيث أن: SK_2 : معامل بيرسون الثاني للالتواء δ الانحراف المعياري في للتوزيع.

$\overline{X}$: المتوسط الحسابي للتوزيع Me الوسيـــــط في للتوزيع

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع أو شكل الالتواء، كمايلي:

– إذا كان $SK = 0$ ، معنى هذا أن: (المتوسط الحسابي = المنوال و المتوسط الحسابي = الوسيط)، وهذا يدل على أن منحنى التوزيع متماثل.

– إذا كان $SK > 0$ ، معنى هذا أن: (المتوسط الحسابي > المنوال و المتوسط الحسابي > الوسيط)، وهذا يدل على أن منحنى التوزيع ملتوي جهة اليمين.

– إذا كان $SK < 0$ ، معنى هذا أن: (المتوسط الحسابي < المنوال و المتوسط الحسابي < الوسيط)، وهذا يدل على أن منحنى التوزيع ملتوي جهة اليسار.

ج. معامل فيشر: ويحسب بالعلاقة التالية:

$$F_1 = \frac{\mu_3}{\sigma^3}$$

حيث أن: F_1 : معامل فيشر للالتواء.

μ_3 : العزم من الدرجة الثالثة.

δ^3 : مكعب الانحراف المعياري.

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع أو شكل الالتواء، كمايلي:

– إذا كان $F_1 = 0$ ، معنى هذا أن منحنى التوزيع متماثل.

– إذا كان $F_1 > 0$ ، معنى هذا أن أن منحنى التوزيع ملتوي جهة اليمين.

– إذا كان $F_1 < 0$ ، معنى هذا أن منحنى التوزيع ملتوي جهة اليسار.

د. معامل يول وكندال: ويحسب بالعلاقة التالية:

$$C_Y = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

$$C_Y = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$$

حيث أن: C_Y : معامل يول وكندال للالتواء

Q_1 : الربيع الأول Q_2 : الربيع الثاني Q_3 : الربيع الثالث

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع أو شكل الالتواء، كمايلي:

– إذا كان $C_Y = 0$ ، معنى هذا أن منحنى التوزيع متماثل.

– إذا كان $C_Y > 0$ ، معنى هذا أن أن منحنى التوزيع ملتوي جهة اليمين.

– إذا كان $C_Y < 0$ ، معنى هذا أن منحنى التوزيع ملتوي جهة اليسار.

هـــ. معامل الالتواء الميئيني: يفضل استخدام هذا المقياس في حالة الجداول التكرارية المفتوحة، أو عند وجود قيم شادة، ويحسب بالعلاقة التالية:

$$CP = \frac{\left(P_{100-x} - P_{50}\right) - \left(P_{50} - P_x\right)}{\left(P_{100-x} - P_x\right)}$$

حيث أن: $P_x < P_{50} < P_{100-X}$.

CP : معامل الالتواء الميئيني.

P_{50} : الميئين الخمسون أو الوسيط.

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع أو شكل الالتواء، كمايلي:

- اذا كان بعد P_{100-X} عن P_{50} هو نفسه بعد P_X عن P_{50}، نقول أن التوزيع متماثل.
- اذا كان بعد P_{100-X} عن P_{50} أكبر من بعد P_X عن P_{50}، نقول أن التوزيع موجب الالتواء.
- اذا كان بعد P_{100-X} عن P_{50} أصغر من بعد P_X عن P_{50}، نقول أن التوزيع سالب الالتواء.

مثال04-05: تمثل البيانات التالية معدلات مجموعتين من طلبة السنة أولى علوم اقتصادية وتجارية وعلوم التسيير بجامعة محمد الصديق بن يحي جيجل بعد مداولات الدورة الأولى.

المجموعة الأولى	
المعدل	n_i
06.90 – 06.30	03
07.50 – 06.90	07
08.10 – 07.50	08
08.70 – 08.10	12
09.30 – 08.70	15
09.90 – 09.30	30
10.50 – 09.90	15
11.10 – 10.50	12
11.70 – 11.10	08
12.30 – 11.70	07
12.90 – 12.30	03
المجموع	**112**

المجموعة الثانية	
المعدل	n_i
06.90 – 06.30	1
07.50 – 06.90	4
08.10 – 07.50	8
08.70 – 08.10	11
09.30 – 08.70	13
09.90 – 09.30	15
10.50 – 09.90	16
11.10 – 10.50	16
11.70 – 11.10	15
12.30 – 11.70	12
12.90 – 12.30	1
المجموع	**112**

- **المطلوب:** معرفة شكل توزيع معدلات الطلبة في كل مجموعة.

الحل: المجموعة الأولى.

نقوم أولا بحساب مقاييس النزعة المركزية الثلاث: المتوسط الحسابي، الوسيط والمنوال ثم نقارن بينها.

- المتوسط الحسابي: $\overline{X}=\frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}$ الوسيط: $Me = L_1 + (\frac{\frac{n}{2}-\sum N_1}{n_{me}}).c$

- المنـــــوال: $Mo = M_1 + \frac{\Delta_1}{\Delta_1+\Delta_2}.c$

إذن نستعين بالجدول من أجل تسهيل عملية الحساب:

$N_i \uparrow$	$n_i.x_i$	x_i	n_i	**المعدل**
3	19.8	6.6	3	06.90 – 06.30
10	50.4	7.2	7	07.50 – 06.90
20	78	7.8	10	08.10 – 07.50
33	109.2	8.4	13	08.70 – 08.10
48	135	9	15	09.30 – 08.70
64	153.6	9.6	16	09.90 – 09.30
79	153	10.2	15	10.50 – 09.90
92	140.4	10.8	13	11.10 – 10.50
102	114	11.4	10	11.70 – 11.10
109	84	12	7	12.30 – 11.70
112	37.8	12.6	3	12.90 – 12.30
-	**1075.2**	-	**112**	**المجموع**

- **المتوسط الحسابي:**

$$\overline{X}=\frac{\sum_{i=1}^{k} n_i x_i}{\sum_{i=1}^{k} n_i}=\frac{1075.2}{112}=9.60$$

متوسط معدلات هؤلاء الطلبة هو 9.60.

- **الوسيط:**

- ترتيب الوسيط: $\frac{n}{2}=\frac{112}{2}=56$؛
- الفئة الوسيطية: 09.90 – 09.30؛

ومنه:

$$Me = L_1 + \left(\frac{\frac{n}{2} - \sum N_1}{n_{me}}\right).c$$

$$Me = 9.30 + \left(\frac{56-48}{16}\right).0.6 = 9.60$$

نقول: 50% من الطلبة معدلاتهم في الدورة الأولى الحرارة أكبر من 9.60.

50% من الطلبة معدلاتهم في الدورة الأولى الحرارة أصغر من 9.60.

- **المنوال حسابيا:**
- الفئة المنوالية: 09.30 – 09.90

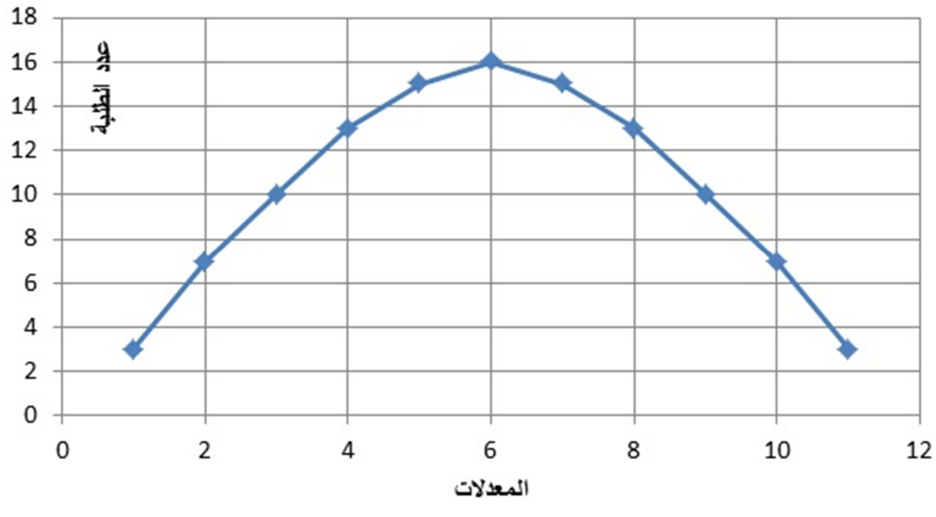
ومنه:

$$Mo \ = 09.30 + \frac{16-15}{16-15+16-15}.0.6 = 9.60$$

اذن: المعدل الأكثر تكرارا في نتائج مداولات الدورة الأولى هو 9.60.

بعد حساب المتوسطات النزعة المركزية الثلاث: المتوسط الحسابي، الوسيط والمنوال وجدناها متساوية؛ وهذا ما يدل على أن شكل التوزيع سيكون متناظر أو متماثل.

- **رسم شكل توزيع المعدلات:**



أيضا لو قمنا بحساب كل معاملات الالتواء التي عرفناها لوجدناها معدومة.

الحل: المجموعة الثانية.

نقوم أولا بحساب مقاييس النزعة المركزية الثلاث: المتوسط الحسابي، الوسيط والمنوال ثم نقارن بينها.

- المتوسط الحسابي: $\overline{X}=\frac{\sum_{i=1}^{k}n_i x_i}{\sum_{i=1}^{k}n_i}$ الوسيط: $Me=L_1+(\frac{N_{me}\downarrow-\frac{n}{2}}{n_{me}}).c$

- المنـــــوال: $Mo=M_1+\frac{\Delta_1}{\Delta_1+\Delta_2}.c$

إذن نستعين بالجدول من أجل تسهيل عملية الحساب:

$N_i\uparrow$	$n_i(x_i-\overline{X})^3$	$(x_i-\overline{X})^3$	$n_i.x_i^2$	$N_i\downarrow$	$n_i.x_i$	x_i	n_i	المعدل
2	73.85-	36.93-	87.12	112	13.2	6.6	2	06.90 – 06.30
6	81.39-	20.35-	207.36	110	28.8	7.2	4	07.50 – 06.90
14	77.31-	9.66-	486.72	106	62.4	7.8	8	08.10 – 07.50
25	39.40-	3.58-	776.16	98	92.4	8.4	11	08.70 – 08.10
38	10.46-	0.80-	1053	87	117	9	13	09.30 – 08.70
53	0.54-	0.04-	1382.4	74	144	9.6	15	09.90 – 09.30
69	0.31	0.02	1664.64	59	163.2	10.2	16	10.50 – 09.90
84	9.88	0.66	1749.6	43	162	10.8	15	11.10 – 10.50
98	44.47	3.18	1819.44	28	159.6	11.4	14	11.70 – 11.10
110	106.44	8.87	1728	14	144	12	12	12.30 – 11.70
112	38.07	19.03	317.52	2	25.2	12.6	2	12.90 – 12.30
-	**83.77-**		**11271.96**	-	**1111.8**	-	**112**	المجموع

- **المتوسط الحسابي:**

$$\overline{X}=\frac{\sum_{i=1}^{k}n_i x_i}{\sum_{i=1}^{k}n_i}=\frac{1111.8}{112}=9.93$$

متوسط معدلات هؤلاء الطلبة هو 9.93.

- **الوسيط:**

- ترتيب الوسيط: $\frac{n}{2}=\frac{112}{2}=56$ ؛

- الفئة الوسيطية: 09.90 – 10.50؛

ومنه:

$$Me = L_1 + (\frac{N_{me} \downarrow - \frac{n}{2}}{n_{me}}).c$$

$$Me = 9.90 + \frac{59 - 56}{16}.0.6 = 10.01$$

نقول: 50% من الطلبة معدلاتهم في الدورة الأولى الحرارة أكبر من 10.01.

50% من الطلبة معدلاتهم في الدورة الأولى الحرارة أصغر من 10.01.

- **المنوال حسابيا:**

- الفئة المنوالية: 09.90 – 10.50.

ومنه:

$$Mo = 09.90 + \frac{16-15}{16-15+16-15}.0.6 = 10.20$$

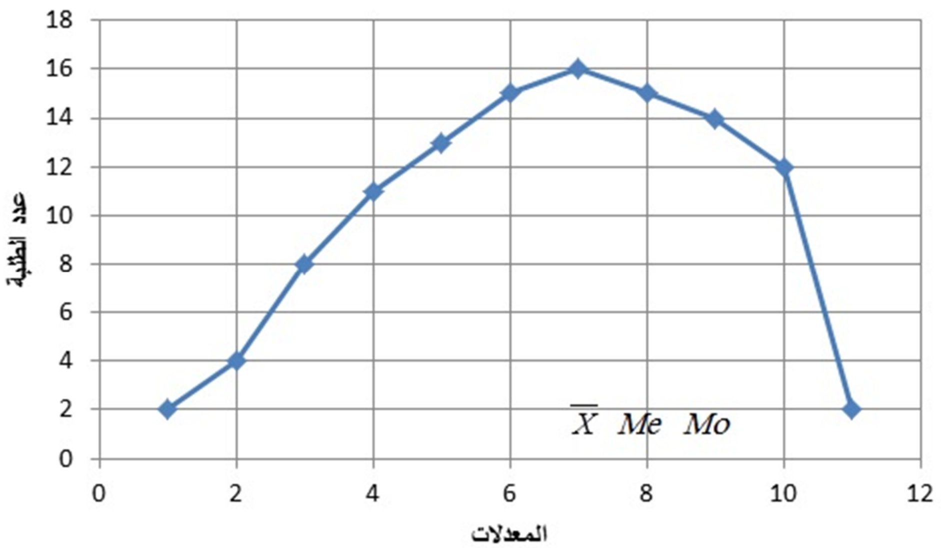
اذن: المعدل الأكثر تكرارا في نتائج مداولات الدورة الأولى هو 10.20.

بعد حساب المتوسطات النزعة المركزية الثلاث: المتوسط الحسابي، الوسيط والمنوال وجدنا أن:

$$(\overline{X} < Me < Mo)$$

وهذا ما يدل على أن شكل التوزيع مائل نحو اليسار أو سالب الالتواء.

- **رسم شكل توزيع المعدلات:**



وإذا أردنا التأكد من ذلك، نقوم بحساب معاملات الالتواء لهذا التوزيع كما يلي:

- **معامل بيرسون الأول للالتواء:**

$$SK_1 = \frac{\overline{X} - Mo}{\sigma}$$

لدينا: $\overline{X} = 9.93$ $\quad$ $Mo = 10.20$ $\quad$ $Me == 10.01$

تبقى حساب الانحراف المعياري بالطريقة المختصرة كمايلي: (نستعين بالجدول لتسهيل الحساب)

$$\delta_X = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{\sum_{i=1}^{k} n_i} - \overline{X}^2} = \sqrt{\frac{11271.96}{112} - 9.93^2} = 1.43$$

ومنه:

$$SK_1 = \frac{\overline{X} - Mo}{\sigma} = \frac{9.93 - 10.20}{1.43} = -\ 0.19$$

- $SK_1 < 0$ ، معنى هذا أن منحنى التوزيع ملتوي جهة اليسار(سالب الالتواء).

- **معامل بيرسون الثاني للالتواء:**

$$SK_2 = \frac{3(\overline{X} - Me)}{\sigma} = \frac{3(9.93 - 10.01)}{1.43} = -0.17$$

- $SK_2 < 0$ ، معنى هذا أن منحنى التوزيع ملتوي جهة اليسار(سالب الالتواء).

- **معامل فيشر:**

$$F_1 = \frac{\mu_3}{\sigma^3}$$

نقوم بحساب العزم من الدرجة الثالثة كمايلي: (نستعين بالجدول لتسهيل الحساب)

$$\mu_3 = \frac{\sum_{i=1}^{k} n_i \left(x_i - \overline{X}\right)^3}{\sum_{i=1}^{k} n_i} = \frac{-83.77}{112} = -0.75$$

ومنه:

$$F_1 = \frac{\mu_3}{\sigma^3} = \frac{-0.75}{1.43^3} = -0.26$$

– $F_1 < 0$ ، معنى هذا أن منحنى التوزيع ملتوي جهة اليسار(سالب الالتواء).

- **معامل يول وكندال:**

$$C_Y = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

$$C_Y = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$$

نقوم بحساب الربيع الأول والثالث، أم الربيع الثاني فهو نفسه الوسيط.

- الربيع الأول Q_1:

$$MQ_1 = \left(\frac{1(n)}{4}\right) = \frac{112}{4} = 28$$

- الفئة الربيعية الأولى: 8.70 – 9.30؛

ومنه: $$Q_1 = 8.70 + (\frac{28-25}{13}).0.6 = 8.84$$

- الربيع الثالث Q_3:

$$MQ_3 = \left(\frac{3(n)}{4}\right) = \frac{3*112}{4} = 84$$

- الفئة الربيعية الثالثة: 10.50 – 11.10؛

ومنه: $$Q_3 = 10.50 + (\frac{84-69}{15}).0.6 = 11.10$$

إذن معامل يول وكندال:

$$C_Y = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1} = \frac{8.84 + 11.10 - (2*10.01)}{11.10 - 8.84} = -0.035$$

– $C_Y < 0$ ، معنى هذا أن منحنى التوزيع ملتوي جهة اليسار (سالب الالتواء).

- **معامل الالتواء الميئيني:**

$$CP=\frac{(P_{100-x}-P_{50})-(P_{50}-P_x)}{(P_{100-x}-P_x)}$$

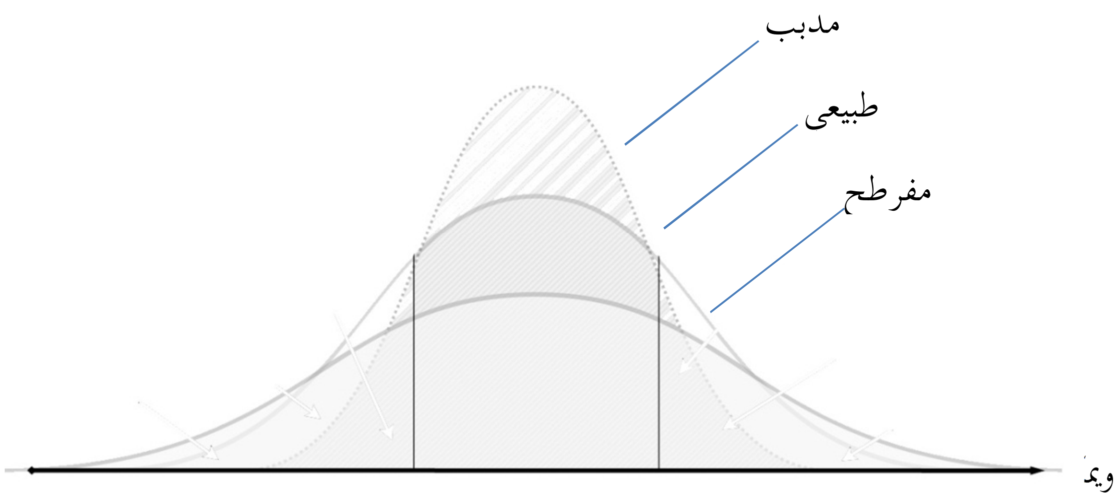
إذا كان: X=25 فإننا نكون بصدد معامل يول وكندال وبالتالي:

$$CP=C_Y=\frac{(P_{75}-P_{50})-(P_{50}-P_{25})}{(P_{75}-P_{25})}=-0.035$$

ومنحنى التوزيع ملتوي جهة اليسار (سالب الالتواء).

2.2. التفرطح

يأخذ منحنى التوزيع عدة أشكال؛ فإذا تركزت القيم بشكل كبير في المنتصف يكون المنحنى مدبب، أما إذا تركزت القيم في الأطراف كان المنحنى مفرطحا أو منبسطا. الأشكال الموالية تبين تدبب وتفرطح المنحنى مقارنة بمنحنى لتوزيع معتدل.



1.2.2. معامل بيرسون: ويحسب بالعلاقة التالية:

$$K=\frac{\mu_4}{\sigma^4}$$

حيث: K : معامل التفرطح.

μ_4 : العزم من الدرجة الرابعة.

δ^4 : الانحراف المعياري مرفوع للقوة أربعة.

- في حالة البيانات غير المبوبة $\mu_4 = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^4}{n}$

- في حالة البيانات المبوبة $\mu_4 = \dfrac{\sum_{i=1}^{k} n_i\left(x_i - \overline{X}\right)^4}{\sum_{i=1}^{k} n_i}$

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع: مفرطح، معتدل أو مدبب، كمايلي:

- إذا كان $K = 3$ ، معنى هذا أن منحنى التوزيع متماثل.
- إذا كان $K > 3$ ، معنى هذا أن منحنى التوزيع مدبب.
- إذا كان $K < 3$ ، معنى هذا أن منحنى التوزيع مفرطح.

2.2.2. معامل فيشر: ويحسب بالعلاقة التالية:

$$K = \frac{\mu_4}{\sigma^4} - 3$$

ويمكن من خلال الإشارة التي يأخذها المعامل معرفة شكل التوزيع: مفرطح، معتدل أو مدبب، كمايلي:

- إذا كان $K = 0$ ، معنى هذا أن منحنى التوزيع متماثل.
- إذا كان $K > 0$ ، معنى هذا أن منحنى التوزيع مدبب.
- إذا كان $K < 0$ ، معنى هذا أن منحنى التوزيع مفرطح.

3.2.2. معامل كيلي: ويحسب بالعلاقة التالية:

$$K = \frac{Q_3 - Q_1}{2\left(P_{90} - P_{10}\right)}$$

حيث:

- اذا كان $K = 0.263$، نقول أن التوزيع طبيعي أي معتدل أومتماثل.
- اذا كان $K > 0.263$، نقول أن التوزيع مدبب.
- اذا كان $K < 0.263$، نقول أن التوزيع مفرطح.

مثال06-04: بالرجوع للمثال04-05 المجموعة الأولى، أحسب معامل التفرطح بمختلف الطرق.

$n_i x_i^2$	$n_i(x_i-\overline{X})^4$	$(x_i-\overline{X})^4$	$N_i\uparrow$	x_i	n_i	المعدل
130.68	243	81	3	6.6	3	06.90 – 06.30
362.88	232.2432	33.1776	10	7.2	7	07.50 – 06.90
608.4	104.976	10.4976	20	7.8	10	08.10 – 07.50
917.28	26.9568	2.0736	33	8.4	13	08.70 – 08.10
1215	1.944	0.1296	48	9	15	09.30 – 08.70
1474.56	0	0	64	9.6	16	09.90 – 09.30
1560.6	1.944	0.1296	79	10.2	15	10.50 – 09.90
1516.32	26.9568	2.0736	92	10.8	13	11.10 – 10.50
1299.6	104.976	10.4976	102	11.4	10	11.70 – 11.10
1008	232.2432	33.1776	109	12	7	12.30 – 11.70
476.28	243	81	112	12.6	3	12.90 – 12.30
10569.6	**1218.24**	-	-	-	**112**	**المجموع**

- **معامل بيرسون:**

$$K=\frac{\mu_4}{\sigma^4}$$

لدينا: $\overline{X}=Me=Mo=9.60$

- نحسب العزم من الدرجة أربعة، والانحراف المعياري (نستعين بالجدول لتسهيل الحساب)

- العزم من الدرجة الرابعة:

$$\mu_4=\frac{\sum_{i=1}^{k}n_i\left(x_i-\overline{X}\right)^4}{\sum_{i=1}^{k}n_i}=\frac{1218.24}{112}=10.88$$

$$\delta_X=\sqrt{\frac{\sum_{i=1}^{k}n_i x_i^2}{\sum_{i=1}^{k}n_i}-\overline{X}^2}=\sqrt{\frac{10569.6}{112}-9.60^2}=1.49$$

- الانحراف المعياري:

ومنه:

$$K=\frac{\mu_4}{\sigma^4}=\frac{10.88}{1.49}=7.30$$

- $K>3$ ، معنى هذا أن منحنى التوزيع مدبب .

- معامل فيشر:

$$K = \frac{\mu_4}{\sigma^4} - 3$$

$$K = \frac{10.88}{1.49} - 3 = 4.30$$

- $K > 0$ ، معنى هذا أن منحنى التوزيع مدبب .

- معامل كيلي:

$$K = \frac{Q_3 - Q_1}{2\left(P_{90} - P_{10}\right)}$$

- الربيع الأول Q_1 **:**

ومنه: $$Q_1 = 8.10 + (\frac{28-20}{13}).0.6 = 8.47$$

- الربيع الثالث Q_3 **:**

ومنه: $$Q_3 = 10.50 + (\frac{84-79}{13}).0.6 = 10.73$$

- الميئين العاشر P_{10} **:**

ومنه: $$P_{10} = 7.5 + (\frac{11.2-10}{10}).0.6 = 7.52$$

- الميئين التسعون P_{90} **:**

ومنه: $$P_{90} = 11.10 + (\frac{100.8-92}{10}).0.6 = 11.63$$

ومنه:

$$K = \frac{Q_3 - Q_1}{2\left(P_{90} - P_{10}\right)} = \frac{10.73 - 8.47}{2\left(11.63 - 7.52\right)} = 0.550$$

$K > 0.263$ ، نقول أن التوزيع مدبب.

الفصل الخامس

الارتباط والانحدار الخطي البسيط

الفصل الخامس
الارتباط والانحدار الخطي البسيط

بعدما عرضنا في الفصول السابقة بعض المقاييس الوصفية مثل مقاييس النزعة المركزية، والتشتت ومقاييس الالتواء وغيرها، التي يمكن أن تصف لنا شكل توزيع البيانات المجموعة من متغير واحد، وبحكم تأثر الظواهر ببعضها، فإن دراسة أي ظاهرة سواء كانت اقتصادية أو غيرها؛ لا يمكن أن تتم بصورة علمية إذا لم تدرس علاقتها بالظواهر المحيطة بها. وتتم هذه الدراسة بواسطة طريقتين؛ الأولى في قوة العلاقة بين المتغيرين، هل هي علاقة قوية، متوسطة أو ضعيفة، علاقة طردية، عكسية أم غير موجودة أصلا. بينما تبحث الثانية تبحث في شكل العلاقة ونوعيتها: هل هي خطية أم لا؟ وبأسلوب رياضي يفرق بين المتغير المؤثر (المستقل) والمتغير التابع (المتأثر).

1. الارتبـــــاط البسيط

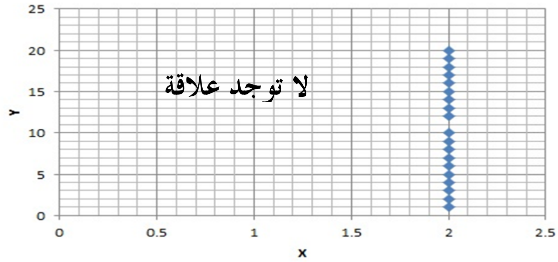
كما سبق الإشارة فإن الارتباط يبحث في قوة العلاقة بين متغيرين أو أكثر، وأبسط أنواع الارتباط هو الارتباط الذي يبحث في قوة العلاقة بين متغيرين(موضوع دراستنا)، فوجود ارتباط بين متغيرين يعني أنه إذا زادت قيم المتغير الأول وصاحبها زيادة في قيم المتغير الثاني تكون العلاقة طردية، أما إذا صاحبها انخفاض في قيم المتغير الثاني نكون أمام علاقة عكسية.

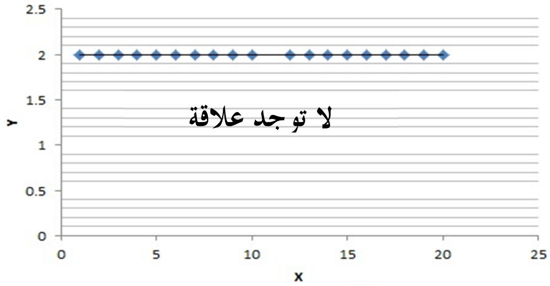
وتوجد أكثر من طريقة لمعرفة طبيعة العلاقة بين متغيرين أو ظاهرتين أهمها:

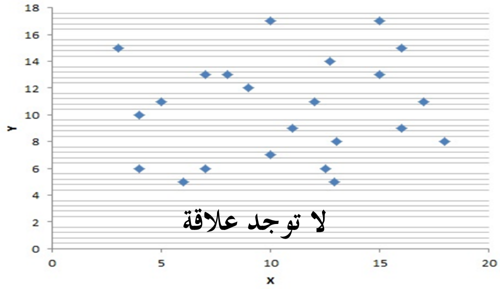
1.1.شكل الانتشار

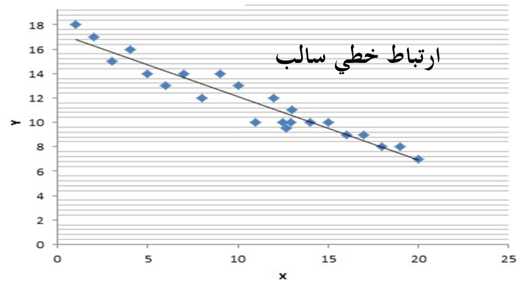
شكل الانتشار؛ شكل بياني يعطي فكرة مبدئية عن اتجاه وقوة العلاقة بين متغيرين، وذلك من خلال تمثيل احداثيات قيم المتغيرين في معلم متعامد.

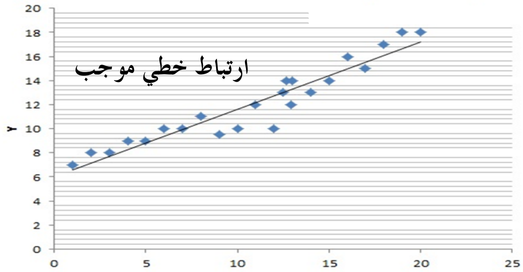
والأشكال التالية تلخص ما سبق:

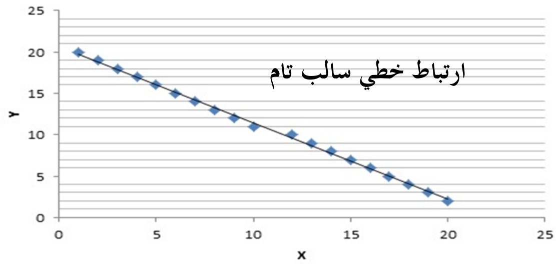


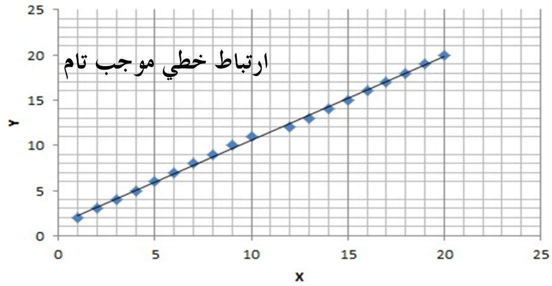












2.1.معامل الارتباط

مقياس يقيس قوة العلاقة الارتباطية واتجاهها بين ظاهرتين أو متغيرين في قيمة يطلق عليها معامل الارتباط **r**. ومعامل الارتباط يأخذ قيمة محصورة بين $-1 \le r \le +1$ ، حيث يدل الرقم على قوة الارتباط أما الإشارة فتدل على اتجاه العلاقة (طردية في حال الموجبة وعكسية في حال الشارة السالبة)، حيث يصنف في المجالات التالية:

- إذا كان: $r < 0.3$ فإن العلاقة بين المتغيرين ضعيفة جدا.
- إذا كان: $0.3 \le r < 0.5$ فإن العلاقة بين المتغيرين ضعيفة.
- إذا كان: $0.5 \le r < 0.7$ فإن العلاقة بين المتغيرين متوسطة.
- إذا كان: $0.7 \le r < 0.9$ فإن العلاقة بين المتغيرين قوية.
- إذا كان: $0.9 \le r < 1$ فإن العلاقة بين المتغيرين قوية جدا.
- إذا كان: $r = 1$ فإن العلاقة بين المتغيرين علاقة ارتباط تام.

هناك عدة طرق لحساب معامل الارتباط، لعل أهمها:

1.2.1. معامل الارتباط البسيط في بيانات كمية: ونميز عدة أنواع؛ نذكر منها معامل الارتباط بيرســـون ويسمى كذلك بمعامل حاصل ضرب العزوم للارتباط، هذا المعامل من أكثر المقاييس استخداما في مختلف الدراسات، ويمكن حساب معامل بيرسون بدلالة القراءات لبيانات المتغيرين X,Y باستخدام الصيغة التالية:

$$r=\frac{n\sum XY-\sum X\sum Y}{\sqrt{\left[n\sum X^2-(\sum X)^2\right]}\sqrt{\left[n\sum Y^2-(\sum Y)^2\right]}}$$

و يمكن حساب معامل الارتباط البسيط بالعلاقة التالية أيضاً :

$$r=\frac{\sum XY-n\overline{X}.\overline{Y}}{\sqrt{\left[\sum X^2-n\overline{X}^2\right]}\sqrt{\left[\sum Y^2-n\overline{Y}^2\right]}}$$

كما يمكن أيضا حساب معامل الارتباط بطريقة مختصرة كمايلي:

$$r=\frac{\sum x\,y}{\sqrt{\sum x^2.\sum y^2}}\qquad / x=X-\overline{X}\ ,\quad y=Y-\overline{Y}$$

كما يمكن حساب معامل الارتباط اعتمادا على التباين المشترك كمايلي:

$$r=\frac{Cov(x,y)}{\sigma_x.\sigma_y}$$

حيث أن: $Cov(x,y)$: هو التباين المشترك لِــ: X و Y و يحسب كما يلي:

$$Cov(x,y)=\frac{\sum(x_i-\overline{X})(y_i-\overline{Y})}{n}$$

σ_x : الانحراف المعياري لِــ X.

σ_y : الانحراف المعياري لِــ Y.

مثال05-01: أدرس العلاقة بين علامات مقيــــاسي الاقتصاد الجزئي1 والاقتصاد الكلي1 لبعض طلاب السنة الثانية علوم تجارية.

الاقتصاد الجزئي1 (X)	10	11	8	16	7	13	6	9
الاقتصاد الكلي1 (Y)	3	7	2	12	1	9	2	4

الحل: نستعين بالجدول لتسهيل الحساب

لدينا:

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - (\sum X)^2\right]}\sqrt{\left[n\sum Y^2 - (\sum Y)^2\right]}}$$

ومنه:

$x.y$	y^2	$y^2 = (Y_i - \overline{Y})^2$	x^2	$x = (X_i - \overline{X})$	$X.Y$	Y^2	X^2	Y	X
0	4	-2	0	0	30	9	100	3	10
2	4	2	1	1	77	49	121	7	11
6	9	-3	4	-2	16	4	64	2	8
42	49	7	36	6	192	144	256	12	16
12	16	-4	9	-3	7	1	49	1	7
12	16	4	9	3	117	81	169	9	13
12	9	-3	16	-4	12	4	36	2	6
1	1	-1	1	-1	36	16	81	4	9
87	**108**	**0**	**76**	**0**	**487**	**308**	**876**	**40**	**80**

$$r = \frac{8\times 487 - 80\times 40}{\sqrt{\left[8\times 876 - 80^2\right]}\sqrt{\left[8\times 308 - 40^2\right]}} = 0.96$$

ومنه توجد علاقة طردية قوية جدا بين علامات مقياس الاقتصاد الجزئي1 وعلامات مقياس الاقتصاد الكلي1.

أما معامل الارتباط البسيط بالعلاقة التالية:

$$r = \frac{\sum XY - n\overline{X}.\overline{Y}}{\sqrt{\left[\sum X^2 - n\overline{X}^2\right]}\sqrt{\left[\sum Y^2 - n\overline{Y}^2\right]}}$$

- نحسب أولا المتوسط الحسابي لكل من المتغير X والمتغير Y كمايلي:

$$\overline{X} = \frac{\sum_{i=1}^{n} Xi}{n} = \frac{80}{8} = 10$$

$$\overline{Y} = \frac{\sum_{i=1}^{n} Yi}{n} = \frac{40}{8} = 5$$

ومنه:

$$r=\frac{487-8\times10\times5}{\sqrt{\left[876-8\times10^2\right]}\sqrt{\left[308-8\times5^2\right]}}=0.96$$

ومنه توجد علاقة طردية قوية جدا بين علامات مقياس الاقتصاد الجزئي1 وعلامات مقياس الاقتصاد الكلي1.

أما بالطريقة المختصرة فيكون معامل الارتباط كمايلي:

$$r=\frac{\sum x\ y}{\sqrt{\sum x^2.\sum y^2}}\qquad /\,x=X-\overline{X}\ \ ,\ \ y=Y-\overline{Y}$$

- نحسب ومنه:

$$r=\frac{\sum x\ y}{\sqrt{\sum x^2.\sum y^2}}\quad=\frac{87}{\sqrt{76\times108}}=0.96$$

ومنه توجد علاقة طردية قوية جدا بين علامات مقياس الاقتصاد الجزئي1 وعلامات مقياس الاقتصاد الكلي1.

أما اعتمادا على التباين المشترك كمايلي:

$$r=\frac{Cov(x,y)}{\sigma_x.\sigma_y}$$

- نحسب أولا: $Cov(x,y)$

لدينا:

$$Cov(x,y)=\frac{\sum\left(x_i-\overline{X}\right)\left(y_i-\overline{Y}\right)}{n}$$

$$=\frac{87}{8}=10.875$$

- ثم نحسب σ_x و σ_y:

$$\delta_X=\sqrt{\frac{\sum_{i=1}^{n}X_i^2}{n}-\overline{X}^2}$$

$$\delta_X=\sqrt{\frac{876}{8}-10^2}=3.08$$

$$\delta_Y = \sqrt{\frac{\sum_{i=1}^{n} Y_i^2}{n} - \overline{Y}^2}$$

$$\delta_Y = \sqrt{\frac{308}{8} - 5^2} = 3.67$$

ومنه:

$$r = \frac{10.875}{3.08 \times 3.67} = 0.96$$

ومنه توجد علاقة طردية قوية جدا بين علامات مقياس الاقتصاد الجزئي1 وعلامات مقياس الاقتصاد الكلي1.

1.2.1. معامل الارتباط البسيط في بيانات نوعية: هناك عدة أنواع هناك عدة مقاييس لقياس الارتباط بين المتغيرات النوعية من أهمها معامل الارتباط الرتبي أو ما يعرف بمعامل الارتباط سبيرمان. ولحسابه فإنه يتم ترتيب المتغير الأول من الأقل إلى الأكبر (أو العكس) ثم ترتب المتغير الثاني بنفس الطريقة، ثم تطبيق العلاقة.

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_1^2}{n(n^2-1)}$$

حيث: d_i إلى الفرق بين ترتيب المتغير الأول والمتغير الثاني.

n عدد أزواج المتغيرين.

ونأخذ متوسط الترتيب لهذه القيم، عندما تكون هناك أكثر من قيمة لها نفس الرتبة.

مثال05-02: تبين البيانات الموالية؛ تقديرات نتائج 8 تلاميذ مدرسة تحصلوا عليها في مادتي الكتابة X والإملاء Y ، المطلوب أحسب معامل الارتباط سبيرمان .

لدينا :

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_1^2}{n(n^2-1)}$$

نستعين بالجدول كمايلي:

ومنه:

الكتابة X	الإملاء Y	ترتيب X	ترتيب Y	d_i	d_i^2
حسن جدا	متوسط	5	4	1	1
جيد	حسن جدا	6	6	0	0
متوسط	ضعيف	3	2.5	0.5	0.25
ممتاز	جيد جدا	8	8	0	0
ضعيف	ضعيف جدا	2	1	1	1
جيد جدا	جيد	7	7	0	0
ضعيف جدا	ضعيف	1	2.5	-1.5	2.25
حسن	حسن	4	5	-1	1
-	-	-	-	**0**	**5.50**

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_1^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6 \times 5.50}{8(8^2-1)} = 1 - 0.066 = 0.93$$

ومنه توجد علاقة طردية قوية جدا بين نتائج مادة الكتابة ونتائج مادة الإملاء.

2. الانحدار الخطي البسيط

الانحدار هو الطريقة التي تهتم ببناء علاقة تفسيرية بين المتغير التابع والمتغير (أو عدة متغيرات) المستقل، هذه العلاقة قد تكون من الدرجة الأولى أو من الدرجة الثانية أو علاقة أسية ...أو غيرها، ولو أننا في دراستنا هذه سوف نهتم فقط بالعلاقة الخطية بين متغيرين.

و يمكن كتابة معادلة الانحدار الخطي البسيط بالصيغة التالية:

$$Y_i = aX_i + b + e_i$$

حيث : Y_i : المتغير التابع.

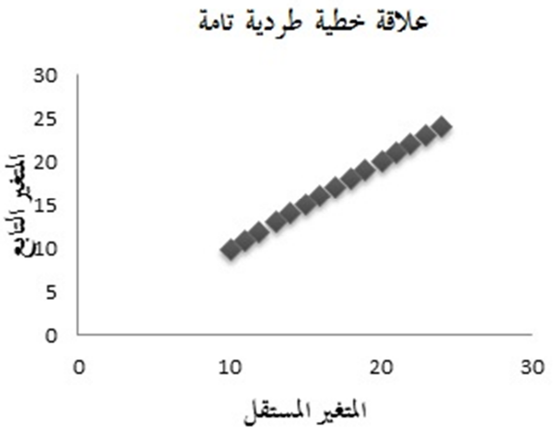
X_i : المتغير المستقل.

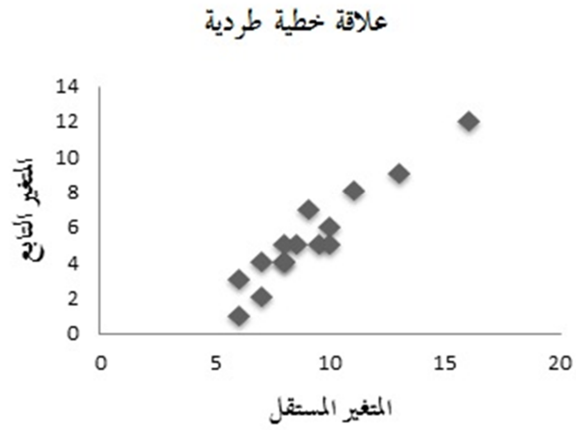
a, b : معلمات.

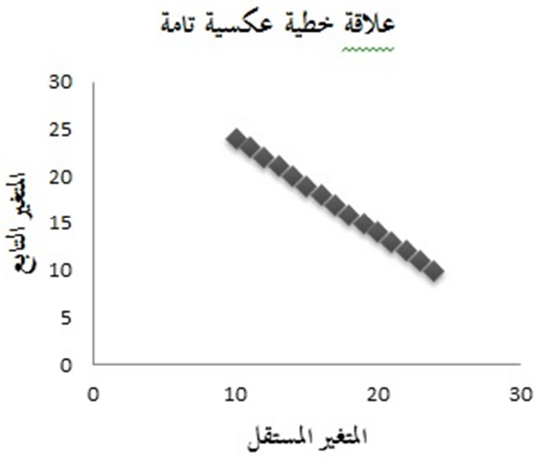
e_i : الخطأ العشوائي

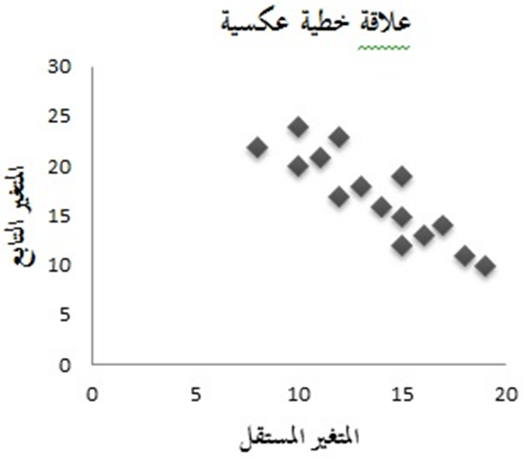
1.2.شكل الانتشار

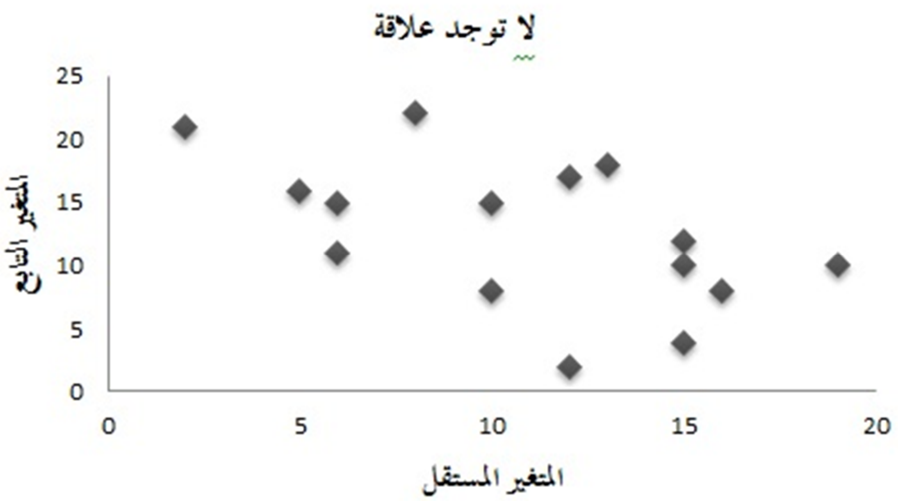
و نميز عدة أشكال للعلاقة الخطية البسيطة بين المتغير التابع والمتغير المستقل:











2.2. تقدير نموذج الانحدار الخطي البسيط

تحدد العلاقة بين المتغير التابع و المتغير المستقل بعدة طرق أشهرها طريقة المربعات الصغرى، حيـــث تعتمد على مبدأ تقدير المعلمات a ، b بشرط أن يكون الخطأ العشوائي عند حده الأدنى(مجموع مربعـــات الفروق بين القيم المشاهدة والقيم المقدرة أصغر ما يمكن).

ويمكن تقدير المعلمتين (a,b) باستخدام طريقة المربعات الصغرى كمايلي:

– **عن طريق حل جملة المعادلة:**

$$\begin{cases} \sum Y = a\sum X + nb \\ \sum YX = b\sum X + a\sum X^2 \end{cases}$$

– **عن طريق القانون العادي:**

$$a = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$

$$b = \overline{Y} - a\overline{X}$$

– **عن طريق القانون المختصر:**

$$a = \frac{\sum xy}{\sum x^2} \qquad / \, x = X - \overline{X} \; , \; y = Y - \overline{Y}$$

$$b = \overline{Y} - a\overline{X}$$

مثال03-05: يبين الجدول الآتي إنتاج محصول الذرة Yمن المساحة المزروعة به X

المطلوب: 1. أرسم شكل الانتشار.

2. أوجد معادلة انحدار Y على X .

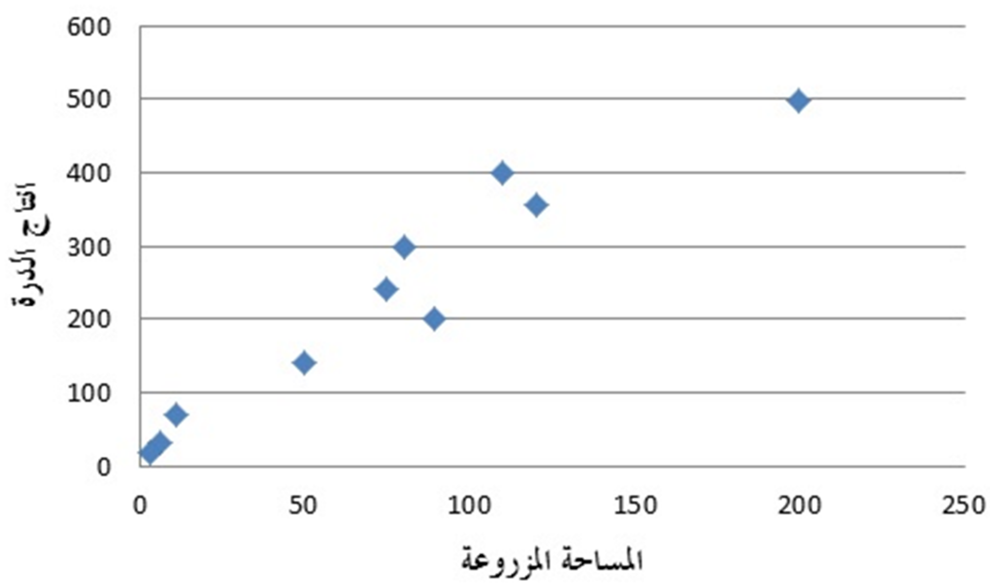
3. ما هو مستوى الانتاج المتوقع عندما تكون المساحة المزروعة هي 300 هكتار؟

المنطقة	X المساحة المزروعة بالهكتار	Y إنتاج الذرة بآلاف الكيلوغرام
1	50	140
2	200	500
3	110	400
4	80	300
5	120	356
6	74.5	240.5
7	88.9	200.6
8	5.7	33.5
9	11	69.8
10	3.2	18.7

الحل:(نستعين بالجدول من أجل سهولة الحساب)

المنطقة	X المساحة المزروعة	Y إنتاج الذرة	X^2	$X.Y$	x	y	$x.y$	x^2
1	50	140	2500	7000	24.33-	85.91-	2090.19	591.95
2	200	500	40000	100000	125.67	274.09	34444.89	15792.95
3	110	400	12100	44000	35.67	174.09	6209.79	1272.35
4	80	300	6400	24000	5.67	74.09	420.09	32.15
5	120	356	14400	42720	45.67	130.09	5941.21	2085.75
6	74.5	240.5	5550.25	17917.25	0.17	14.59	2.48	0.03
7	88.9	200.6	7903.21	17833.34	14.57	25.31-	368.77-	212.28
8	5.7	33.5	32.49	190.95	68.63-	192.41-	13205.10	4710.08
9	11	69.8	121	767.8	63.33-	156.11-	9886.45	4010.69
10	3.2	18.7	10.24	59.84	71.13-	207.21-	14738.85	5059.48
$\sum$	**743.3**	**2259.1**	**89017.19**	**254489.18**	**-**	**-**	**86570.28**	**33767.70**

1.رسم شكل الانتشار:



من خلال شكل الانتشار نلاحظ أن العلاقة بين المساحة المزروعة وإنتاج الذرة علاقة خطية طردية.

2. إيجاد معادلة انحدار Y على X

وجدنا أن العلاقة بين المتغيرين علاقة خطية من الشكل : $Y_i = aX_i + b + e_i$، ولتقدير المعلمتين (a,b) نستخدم مايلي:

- **عن طريق حل جملة المعادلة**

$$\begin{cases}\sum Y = a\sum X + nb \\ \sum YX = b\sum X + a\sum X^2\end{cases} \Rightarrow \begin{cases}2259.1 = 743.3a + 10b \\ 254489.18 = 743.3b + 89017.19a\end{cases}$$

$$\Rightarrow \begin{cases}b = \dfrac{2259.1 - 743.3a}{10} \\ 254489.18 = 743.3b + 89017.19a\end{cases}$$

$$\Rightarrow \begin{cases}b = 225.91 - 74.33a \\ 254489.18 = 743.3b + 89017.19a\end{cases}$$

بالتعويض بقيمة b في المعادلة الثانية نجد:

$$\begin{cases}b = 225.91 - 74.33a \\ 254489.18 = 743.3b + 89017.19a\end{cases} \Rightarrow \begin{cases}b = 225.91 - 74.33a \\ 254489.18 = 743.3(225.91 - 74.33a) + 89017.19a\end{cases}$$

$$\Rightarrow \begin{cases}b = 225.91 - 74.33a \\ 86570.277 = 33767.701a\end{cases}$$

$$\Rightarrow \begin{cases}b = 225.91 - 74.33a \\ \mathrm{a} = \dfrac{86570.277}{33767.701} = 2.564\end{cases}$$

$$\Rightarrow \begin{cases}b = 35.35 \\ \mathrm{a} = 2.564\end{cases}$$

ومنه معادلة الانحدار تأخذ الشكل: $Y_i = 2.564X_i + 35.35$

- **عن طريق القانون العادي**

$$a = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$

$$b = \overline{Y} - a\overline{X}$$

- نحسب أولا المتوسط الحسابي لكل من المتغير X والمتغير Y كمايلي:

$$\overline{X} = \frac{\sum_{i=1}^{n} Xi}{n} = \frac{743.3}{10} = 74.33$$

$$\overline{Y}=\frac{\sum_{i=1}^{n}Yi}{n}=\frac{2259.1}{10}=225.91$$

ومنه:

$$a=\frac{\sum XY-n\overline{X}\overline{Y}}{\sum X^2-n\overline{X}^2}=\frac{254489.18-10\times 74.33\times 225.91}{89017.19-10\times(74.33)^2}$$

$$=\frac{86570.277}{33767.701}=2.564$$

$$a=2.564$$

بالتعويض بقيمة a نجد:

$$a=2.564$$

$$b=\overline{Y}-a\overline{X}=225.91-2.564\times(74.33)=35.35$$

$$a=2.564$$
$$b=35.35$$

ومنه معادلة الانحدار تأخذ الشكل: $Y_i=2.564X_i+35.35$

- **عن طريق القانون المختصر**

$$a=\frac{\sum xy}{\sum x^2} \qquad / x=X-\overline{X}\ ,\ y=Y-\overline{Y}$$

$$b=\overline{Y}-a\overline{X}$$

$$a=\frac{\sum xy}{\sum x^2}=\frac{86570.28}{33767.70}=2.564$$

$$a=2.564$$

$$a=2.564$$

ومنه:

$$b=\overline{Y}-a\overline{X}=225.91-2.564\times(74.33)=35.35$$

$$a=2.564$$
$$b=35.35$$

3. العلاقة بين معامل الارتباط و معامل الانحدار الخطي البسيط

إذا كان لدينا:

- a معامل انحدار Y على X أي: $\left(Y_{(x)} = a_1X \ + b_1\right)$
- a معامل انحدار X على Y أي: $\left(Y_{(x)} = a_2X \ + b_2\right)$

فإن معامل الارتباط البسيط بين X و Y يمكن حسابه كما يلي:

$$r = \sqrt{a_1 . a_2}$$

الفصل السادس

مدخل للسلاسـل الزمنية

الفصل السادس
مدخل للسلاسل الزمنية

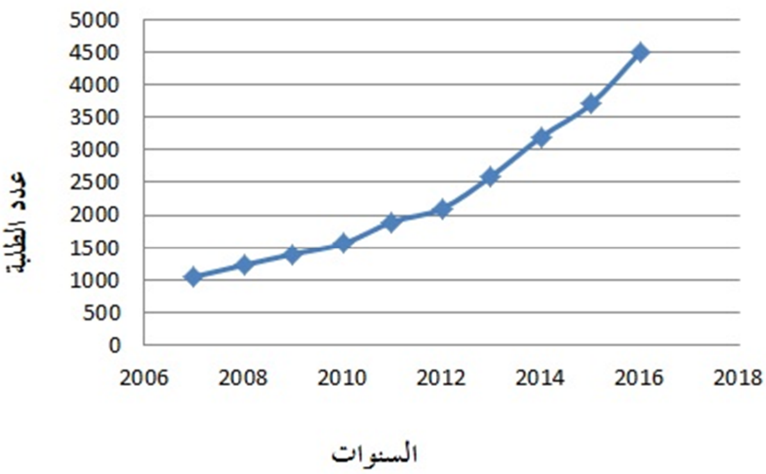
لا يمكن لأي ظاهرة أن تتغير بمعزل عن الزمن، لذا تعتبر دراسة تغيرات الظواهر المختلفة بالنسبة للزمن (السلسلة الزمنية) من أهم مجالات الإحصاء الوصفي. والسلسلة الزمنية هي مجموعة القياسات المسجلة لمتغير واحد أو أكثر مرتبة حسب زمن وقوعها (سنة، فصل، شهر، أسبوع، يوم)، ونعبر عنها رياضيا حيث نقول أن متغير الزمن المستقل t والقيم المناظرة له المتغير التابع Y ، وإن كل قيمة في الزمن t يقابلها قيم للمتغير التابع Y فإن Y دالة في الزمن t أي:

$$Y = f(t)$$

من الأمور الطبيعية والواجبة للحكومات والمؤسسات التخطيط لمستقبلها بغية تحقيق الأهداف الخاصة والعامة، وتقديم كافة الخدمات للوصول إلى حالة الاستقرار والعمل على اتخاذ قرارات التنبؤ بوقوع الأحداث قبل وقوعها في كافة أوجه النشاط التي تخص المجتمع، وتعتبر السلاسل الزمنية من أهم أساليب التنبؤ للمستقبل انطلاقا من الماضي والحاضر.

إن من أهم السلاسل الزمنية تلك التي تهتم بالمؤشرات الاقتصادية والمبيعات السنوية للشركات بكافة أوجه نشاطاتها والتعليم وحجم السكان وما شابه ذلك .والتغير الذي يحدث في قيم متغير السلسلة الزمنية أو قيم متغيراتها يعتبر دالة في الزمن يمكن تمثيلها بيانيا، كما هو مبين بالشكل الآتي لجدول البيانات الآتي والدال على تطور عدد الطلبة في كلية العلوم الاقتصادية والتجارية وعلوم التسيير جامعة جيجل خلال الفترة 2007-2016.

السنة	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
عدد الطلبة	1050	1230	1400	1560	1900	2100	2600	3200	3700	4500



1. مكونات (مركبات) السلسلة الزمنية

تتكون السلسلة الزمنية في الغالب من أربع مركبات يمكن تلخيصها فيمايلي:

1.1. الاتجاه العـــام "T"

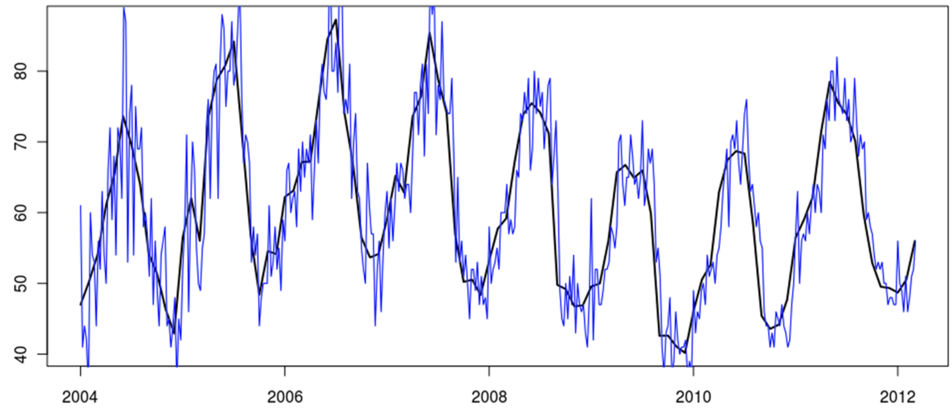
الاتجاه العام وتغيراته من أهم عناصر السلسلة الزمنية، وهو يعكس اتجاه تغير الظاهرة المدروسة عبر الزمن؛ هل هي متزايدة أم متناقصة. الشكلين المواليين يوضحين ذلك:

2.1 التغيرات الدورية "C"

التغيرات الدورية تغيرات تنتج عن تأثير القوى التي تظهر من حين لآخر ويكون تأثيرها على قيم السلسلة الزمنية على شكل تزايد لهذه القيم أو تناقصها حتى تبلغ دورة عظمى (صغرى) ثم تعود لتتناقص (تتزايد) حتى تبلغ دورة صغرى (كبرى) وهكذا.

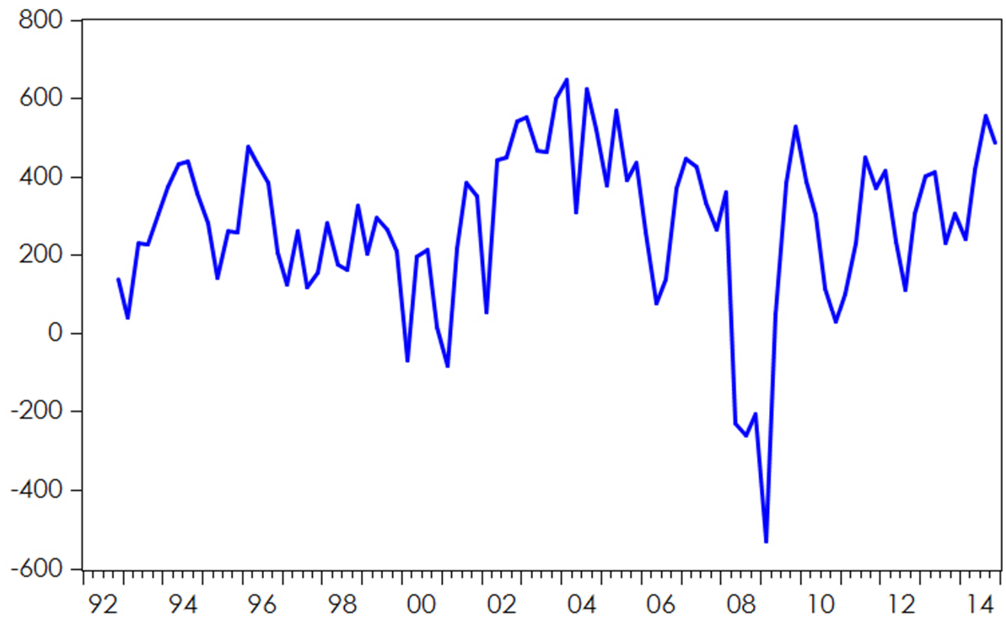
وقد تتكرر هذه التغيرات في فترات زمنية أكثر من سنة ولا تتبع نفس النظام والنسق من حيث الفترات الزمنية. وترجع التغيرات الدورية لعدة أسباب أهمها التغير في العلاقات الدولية والسياسية

الحكومية، وكذا التغير في عرض السلع والخدمات والطلب عليها. الشكل الموالي يوضح شكل التغيرات الدورية:



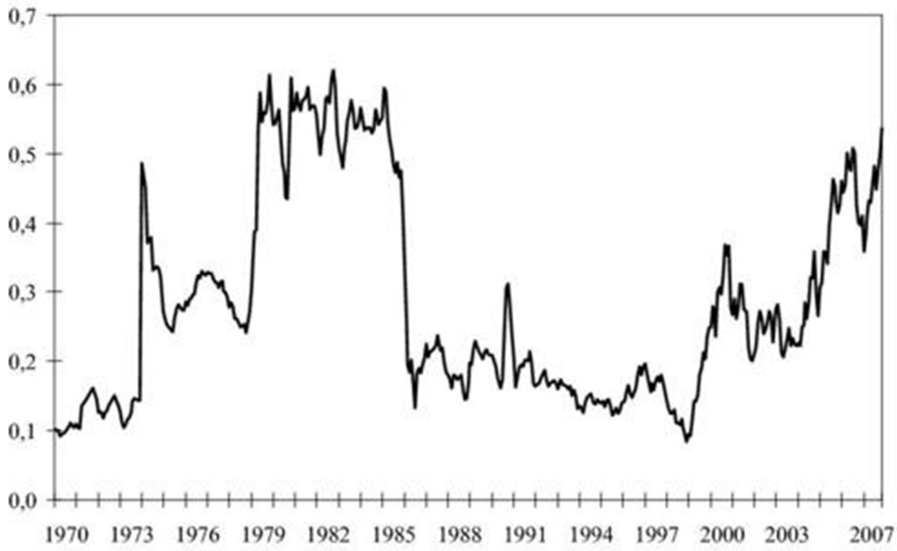
3.1 التغيرات الموسمية "*S*"

التغيرات الموسمية هي التغيرات الناتجة عن تأثير فصول السنة أو تأثير الأشهر خلال الفصول أو تأثير الأيام خلال الأشهر، وقد يرجع السبب في هذه التغيرات الموسمية إلى العادات الاجتماعية أو الطقس ...الخ. والشكل الموالي يوضح ذلك:



4.1. التغيرات العشوائية "*R*"

وهي تغيرات طارئة تحدث نتيجة حوادث فجائية غالبا لا تكون في الحسبان، وبالتالي لا تحدث هذه التغيرات مفعولها طبقا لقاعدة ثابتة أو تأثيرا ثابت على قيم السلسلة الزمنية، فقد يكون التأثير تارة بالزيادة وتارة بالنقصان وعلى فترات قصيرة. وفجائية عوامل حدوثها تجعل من الصعوبة بمكان التنبؤ بها وتقديرها من حيث حجمها واتجاهها. من أهم عوامل حدوث التغيرات العرضية الحروب والإضرابات والزلازل والفيضانات ...الخ.



2. تحليل السلاسل الزمنية

الهدف من تحليل السلسلة الزمنية هو التعرف على مكوناتها (الاتجاه العام – التغيرات الموسمية – التغيرات الدورية – التغيرات العشوائية) كلاً على حدة حيث يستخدم نموذجين، الأول يعرف بنموذج الجمع وأما الآخر فيعرف بنموذج الضرب للسلسلة الزمنية بقصد تجزئة السلسلة الزمنية وذلك بتحديد علاقة السلسلة بمكوناتها، وسنرمز بالرموز الآتية:

t : الزمن.

Yt : قيمة الظاهرة عند اللحظة *t*.

T : القيمة الاتجاهية للظاهرة عند اللحظة الزمنية *t*.

C : التغير الموسمي للظاهرة عند اللحظة الزمنية *t*.

S : التغير الموسمي للظاهرة عند اللحظة *t*.

R : التغير العشوائي (التغير العرضي) للظاهرة عند اللحظة الزمنية t.

وهناك نموذجان شائعا الاستخدام هما:

- نموذج الجمع ويأخذ الشكل: $Yt = T+C+S+R$
- نموذج الضرب ويأخذ الشكل: $Yt= T.C.S.R$

سنكتفي بنموذج الضرب؛ كون السلاسل الزمنية التي سندرسها قصيرة وبالتالي لا يوجد أثر دوري، وعلى هذا فإن النموذج الذي سنركز عليه هو نموذج الجمع المختصر $Yt = T+ S+R$

3.تحديد مكونات السلسلة الزمنية

يتلخص الاتجاه العام للسلسلة الزمنية في الخط الذي يبين حركة السلسلة خلال فترة من الزمن، وهو من أهم عناصر السلسلة الزمنية، حيث يمكن تحديده بواسطة طريقة المتوسطات المتحركة أو طريقة المربعات الصغرى.

مثال06-01: فيما يلي عدد الأسهم الفصلية المباعة لشركة ما خلال الفترة 2014-2017.

المطلوب: تحديد الاتجاهين للظاهرة باستخدام طريقة المربعات الصغرى، ثم استخدام نموذج الجمع للسلسلة الزمنية للتنبؤ بمبيعات الفصول الأربعة لسنة 2018؟

السنة	2014				2015				2016				2017			
الفصل	1	2	3	4	1	2	3	4	1	2	3	4	1	3	3	4
الإنتاج	20	12	10	18	23	14	13	22	28	17	17	26	30	21	20	28

الحل:

1.تحديد معادلة الاتجاه العام:

لتحديد معادلة الاتجاه العام نتبع نفس الخطوات التي درسناها في تحديد معادلة الانحدار، حيث نجد أن الإنتاج هنا دالة في الزمن، أي: $Y = f(t)$. وتأخذ معادلة الاتجاه العام الشكل التالي:

$$Y_i = at_i + b$$

نستعين بالجدول من أجل تسهيل الحسابات وذلك بفرض ان السلسلة تبدأ بصفر من الفصل الرابع لسنة 2015 كمايلي:

السنة	الفصل	Y	t	t^2	$t.Y$	T	$Y-T=S+R$
2014	1	20	7-	49	-140	14.32	5.68
	2	12	6-	36	-72	15.07	-3.07
	3	10	5-	25	-50	15.82	-5.82
	4	18	4-	16	-72	16.57	1.43
2015	1	23	3-	9	-69	17.32	5.68
	2	14	2-	4	-28	18.07	-4.07
	3	13	1-	1	-13	18.82	-5.82
	4	22	0	0	0	19.57	2.43
2016	1	28	1	1	28	20.32	7.68
	2	17	2	4	34	21.07	-4.07
	3	17	3	9	51	21.82	-4.82
	4	26	4	16	104	22.57	3.43
2017	1	30	5	25	150	23.32	6.68
	2	21	6	36	126	24.07	-3.07
	3	20	7	49	140	24.82	-4.82
	4	28	8	64	224	25.57	2.43
	$\sum$	**319**	**8**	**344**	**413**		

باستخدام طريقة المربعات الصغرى وبواسطة القانون العادي:

$$a = \frac{\sum tY - n\bar{t}\bar{Y}}{\sum t^2 - n\bar{t}^2}$$

$$b = \bar{Y} - a\bar{t}$$

- نحسب أولا المتوسط الحسابي لكل من المتغير t والمتغير Y كمايلي:

$$\bar{t} = \frac{\sum_{i=1}^{n} t_i}{n} = \frac{8}{16} = 0.5$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{319}{16} = 19.94$$

ومنه:

$$a = \frac{\sum tY - n\bar{t}\bar{Y}}{\sum t^2 - n\bar{t}^2} = \frac{413 - 16 \times 0.5 \times 19.94}{344 - 16 \times 0.5^2}$$

$$a = 0.75$$

$$a = 0.75$$

$$b = 19.94 - 0.75 \times 0.5 = 19.57$$

ومنه معادلة الاتجاه العام تأخذ الشكل: $Y_i = 0.75t_i + 19.57$

2.حساب القيم الاتجاهية:

انطلاقا من معادلة الاتجاه نحسب القيم الاتجاهية لكل فصل. (الحسابات في الجدول)

3.حساب الفرق بين القيم الأصلية والقيم الاتجاهية: $Y - T = S + R$ (الحسابات في الجدول)

4.استبعاد التغير العشوائي ***R*** : وذلك من خلال حساب المتوسط الحسابي أو الوسيط لنفس الفترة الزمنية من كل سنة، بهدف اظهار التغير الموسمي فقط.

السنة	الفصل	2014	2015	2016	2017	المتوسط الحسابي
الفصل	1	5.68	5.68	7.68	6.68	6.43
	2	-3.07	-4.07	-4.07	-3.07	-3.57
	3	-5.82	-5.82	-4.82	-4.82	-5.32
	4	1.43	2.43	3.43	2.43	2.43
	$\sum$	-	-	-	-	**0.03-**

إذا كان متوسط المتوسطات (متوسط كل الفروق) لا يساوي صفر فإننا نقوم بإجراء تصحيح لهذه المتوسطات من خلال طرح متوسط كل الفروق من متوسط كل فصل، وه ما وجدناه في مثالنا.

السنة	الفصل	المتوسط الحسابي	تصحيح	تصحيح	تصحيح
الفصل	1	6.43	6.4	6.25	5.5
	2	-3.57	-3.6	-3.75	-4.5
	3	-5.32	-5.35	-5.5	-6.25
	4	2.43	2.4	2.25	1.5
	$\sum$	**0.03-**	**-0.15**	**-0.75**	**0**

5.استخدام نموذج الجمع للسلسلة الزمنية للتنبؤ بمبيعات الفصول الأربعة لسنة 2018

نقوم بحساب القيمة الاتجاهية لكل فصل من فصول السنة ثم نضيف إليها التغير الموسمي لكل فصل.

- بالنسبة للفصل الأول فإن t=9 وبالتالي:

$$T_9 = 0.75 \times 9 + 19.57$$
$$T_9 = 26.32$$

وبما أن القيمة الموسمية لهذا الفصل هي S=5.5 فإن القيمة المتوقعة لهذا الفصل هي:

$$P_9 = 26.32 + 5.5 = 31.82$$

- بالنسبة للفصل الثاني فإن t=10 وبالتالي:

$$T_{10} = 0.75 \times 10 + 19.57$$
$$T_{10} = 27.07$$

وبما أن القيمة الموسمية لهذا الفصل هي 4.5 -=S فإن القيمة المتوقعة لهذا الفصل هي:

$$P_{10} = 27.07 - 4.5 = 22.57$$

- بالنسبة للفصل الثالث فإن t=11 وبالتالي:

$$T_{11} = 0.75 \times 11 + 19.57$$
$$T_{11} = 27.82$$

وبما أن القيمة الموسمية لهذا الفصل هي 6.25 -=S فإن القيمة المتوقعة لهذا الفصل هي:

$$P_{11} = 27.07 - 6.25 = 21.57$$

- بالنسبة للفصل الثالث فإن t=12 وبالتالي:

$$T_{12} = 0.75 \times 12 + 19.57$$
$$T_{12} = 28.57$$

وبما أن القيمة الموسمية لهذا الفصل هي 6.25 -=S فإن القيمة المتوقعة لهذا الفصل هي:

$$P_{12} = 28.57 + 1.5 = 30.07$$

مثال06-02:

تمثل البيانات التالية مبيعات أحد المتاجر خلال شهر(من يوم السبت إلى يوم الأربعاء).

1. أحسب متوسط متحرك لخمسة أيام؟
2. أحسب المبيعات بعد تصحيحها موسميا؟
3. أحسب التغير العشوائي؟
4. أرسم منحنى المبيعات الفعلية والمبيعات المصححة موسميا على رسم بياني واحد؟

	الاسبوع1	الأسبوع 2	الأسبوع 3	الأسبوع 4
السبت	390	380	350	360
الأحـــد	450	440	430	400
الاثنين	500	490	490	480
الثلاثاء	600	590	580	600
الاربعاء	690	690	680	660

الحل:

1. من أجل حساب المتوسط المتحرك نقوم بحساب عدد من المتوسطات المتتابعة من القيم الأصلية للظاهرة، بهدف القضاء على التغيرات الموسمية والعشوائية، حيث تمثل قيم المتوسطات المتحركة قيم اتجاهية تقريبية.

ولحساب المتوسط المتحرك لخمسة أيام والذي لا يتضمن أي أثر موسمي نقوم بمايلي:

- نحسب مجموع خمسة قيم متتالية، ونضعها في عمود نسميه "مجموع خمسة أيام" وحيث أن العدد هنا فردي فإن المجموع يوضع أمام القيمة الوسطى* (أمام مبيعات يوم الاثنين).
- نحسب المتوسط المتحرك (القيمة الاتجاهية) وهو مجموع القيم على عددها، ونضعه في عمود نسميه (T).

الأسبوع	اليوم	Y	مجموع 05 أيام	T	$Y-T=S+R$	$Y-S$	R
الأسبوع 1	السبت	390	-	-	-	501	-
	الأحد	450	-	-	-	474	-
	الاثنين	500	2500	500	-20	504	4
	الثلاثاء	600	2490	498	102	522	24
	الأربعاء	690	2520	504	156	499	-5
الأسبوع 2	السبت	380	2500	500	-150	491	-9
	الأحد	440	2480	496	-66	504	8
	الاثنين	490	2500	500	-10	514	14
	الثلاثاء	590	2530	506	74	502	-4
	الأربعاء	690	2540	508	172	519	11
الأسبوع 3	السبت	350	2540	508	-128	521	13
	الأحد	430	2550	510	-70	514	4
	الاثنين	490	2560	512	-22	514	2
	الثلاثاء	580	2570	514	76	512	-2
	الأربعاء	680	2580	516	174	529	13
الأسبوع 4	السبت	360	2590	518	-128	531	13
	الأحد	400	2600	520	-70	524	4
	الاثنين	480	2600	520	-20	524	4
	الثلاثاء	600	-	-	-	522	-
	الأربعاء	660	-	-	-	529	-

2. لإجراء التصحيح الموسمي نقوم بالتالي:

- نحسب التغير الموسمي والتغير العشوائي (S+R) وذلك بطرح (T) من (Y)، ولإيجاد متوسطات قيم (S+R) ننقل هذه القيم إلى الجدول الثاني الذي يمثل الأعمدة فيه الأسابيع والأسطر الأيام، وبالتالي فإن المتوسطات المطلوبة تعطى في السطور، وقد أخذنا في مثالنا هذا المتوسط للقيم الواردة.

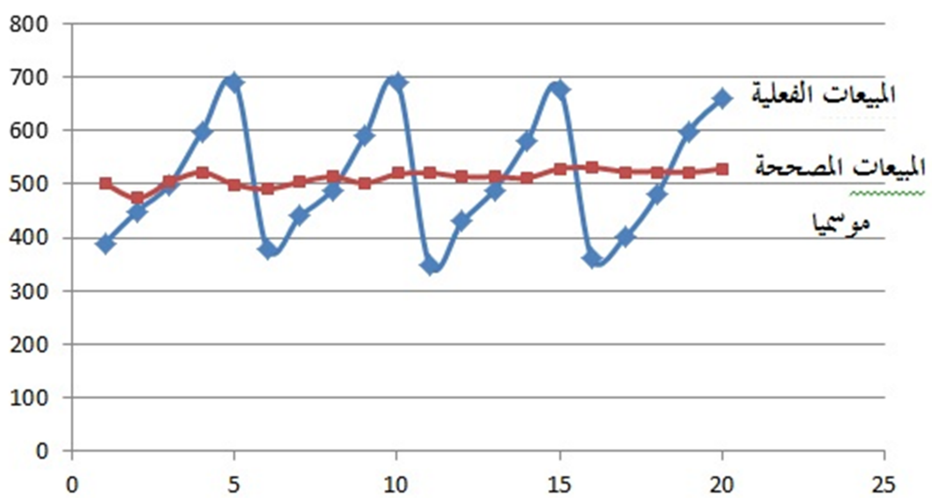
	الاسبوع1	الأسبوع 2	الأسبوع 3	الأسبوع 4	المتوسط	متوسط الفروق	التصحيح	التقريب
السبت	-	-150	-128	-128	-135.33		-141.196	-141
الأحــــد	-	-66	-70	-70	-68.67		-74.536	-75
الاثنين	-20	-10	-22	-20	-18	5.866	-23.866	-24
الثلاثاء	102	74	76	-	84		78.134	78
الاربعاء	156	172	174	-	167.33		161.464	161
					29.33	-	-	-

– بما أن متوسط المتوسطات (متوسط الفروق) يختلف عن الصفر ($\frac{29.33}{5} = 5.866$)، فإننا نقوم بإجراء تصحيح لهذه المتوسطات (لأن التغيرات الموسمية حسب التعريف المقدم في بداية هذا الفصل هي الاختلاف من أحد فصول السنة إلى فصل آخر (أو من يوم إلى آخر في مثالنا) فإنه لا يوجد تأثير موسمي على مدى أيام الأسبوع مجمعة) من خلال طرح متوسط كل الفروق (5.866) من متوسط كل يوم ونضعه في عمود نسميه (Sالمصححة)، أما العمود (S المقربة) فهو تقريب لقيم S المصححة.

– نرجع للجدول الأول ونطرح القيم المقربة لـــ (S) من قيم (Y) فنحصل على عمود (Y-S) أي العمود الذي يبين المبيعات المصححة موسميا.

3. لحساب التغيرات العشوائية (R) فإننا نقوم بطرح القيم المقربة لـــ (S) من العمود (Y-T) فنحصل على العمود (R) والذي يساوي (Y-T-S).

4. المبيعات الفعلية والمصححة موسميا يمكن تمثيلها بالمنحنيات التالية:



الفصل السابع

الأرقـــام القياسية

الفصل السابع
الأرقـــــام القياسية

الرقم القياسي هو عبارة عن مؤشر إحصائي يقيس التغير النسبي الذي طرأ على ظاهرة معينة، سعرا، كمية، قيمة أو أجرًا، بالنسبة لأساس معين قد يكون فترة زمنية معينة او مكانا جغرافيا معينا، حيث تؤخذ قيمة الظاهرة كأساس لحساب الرقم القياسي. ويسمى الوقت أو المكان الذي تنسب اليه الظاهرة بفترة أو مكان الاساس، كما يسمى الوقت او المكان الذي ننسبه الى فترة او مكان، المقارنة.

يرجع استخدام الأرقام القياسية إلى أكثر من قرنين من الزمن، حيث استخدمها الإحصائي الإيطالي كارلي عام 1764 لمقارنة الأسعار في إيطاليا لسنة 1750 بالأسعار في سنة 1500، ثم شاع استخدامها بصورة أوسع منذ ذلك الحين، حيث اهتمت الحكومات بتركيب وحساب بعض الأرقام القياسية.

من الأمور الهامة عند تركيب الرقم القياسي اختيار فترة الأساس أو مكان الأساس التي تعتمد لتركيب الرقم، وعادة ما تكون فترة الأساس سابقة لفترة المقارنة. كما يجب اختيار فترة أو مكان الأساس بحيث تكون متميزة بالاستقرار الاقتصادي وخالية من الاضطرابات العنيفة التي قد تتعرض لها الظاهرة كالحروب والأزمات الاقتصادية، كما يفضل أن لا تكون بعيدة جدًا عن سنوات المقارنة.

واستخدام الأرقام القياسية كأساس للمقارنة لا يقتصر فقط على مقارنة التغير في ظاهرة ما زمانيا أو مكانيا، بل يمكن استخدامها للمقارنة بين ظاهرتين مختلفتين أو أكثر، فعلى سبيل المثال يمكن المقارنة بين التغيرات في أسعار سلعة ما والتغيرات في الكميات المستهلكة منها، أو المقارنة بين التغيرات في مستوى المعيشة للعمال ومستويات أجورهم.

وتجدر الإشارة إلى أن تطبيقات الأرقام القياسية لم تعد مقتصرة على الاقتصاديين فقط، بل أصبحت وسيلة في أيدي المهتمين في العلوم الاجتماعية والإدارية والزراعية ... وذلك لعمل المقارنات وقياس التغيرات.

يمكن تمييز صيغتين أساسيتين للأرقام القياسية هما:

1.الصيغ البسيطة للأرقام القياسية

وتشمل مايلي:

1.1.منسوب السعر: يعتبر منسوب السعر من أبسط الأمثلة للرقم القياسي، وهو نسبة قيمة المتغير في فترة المقارنة إلى قيمة نفس المتغير في فترة الأساس. ويعطى منسوب السعر بالعلاقة:

$$P_{1/0} = \frac{p_1}{p_0}.100$$

حيث:

$P_{1/0}$: منسوب السعر

p_1 : السعر خلال السنة n.

p_o : السعر خلال سنة الأساس

مثال07-01: إذا كان سعر مادة الدقيق في سنة 2016 هو 870 دينار، وفي سنة 2017 بلغ 990 دينار، باتخاذ سنة 2016 كسنة أساس فإن منسوب السعر:

$$P_{2017/2016} = \frac{990}{870}.100$$

$$P_{2017/2016} = 113.79\%$$

وهذا يعني أن سعر مادة الدقيق زاد بنسبة 13.79% في سنة 2017 عما كان عليه سنة 2016

2.1.منسوب الكمية: في حالة مقارنة كميات السلع بدلا من الأسعار، كما هو الحال بالنسبة لحجم الإنتاج والاستهلاك والتصدير مثلا، فإننا نتكلم عن منسوب الكمية كما في حالة الأسعار. ويعطى منسوب الكمية بالعلاقة:

$$q_{1/0} = \frac{q_1}{q_0}.100$$

حيث: $q_{1/0}$: منسوب الكمية

q_1 : كمية السلعة خلال السنة n.

q_o : كمية السلعة خلال سنة الأساس

3.1.منسوب القيمة: نعرف أن القيمة الاجمالية للسلعة هي عبارة عن كمية هذه السلعة مضروبة في سعرها ($V=P.q$)، فإذا كانت q_0,P_0 تعبر عن سعر السلعة والكمية المنتجة منها في سنة الأساس و q_1,P_1 تعبر عن سعر السلعة والكمية المنتجة منها في سنة المقارنة، فإن القيمة الإجمالية خلال سنة الأساس هي V_0 وخلال سنة المقارنة هي V_n وعليه فإن:

$$V_{1/0} = \frac{V_1}{V_0} \times 100 = \frac{P_1 q_1}{P_0 q_0} \times 100$$

مثال07-02: بلغت مبيعات إحدى معاصر الزيتون بولاية جيجل 4500 لتر من الزيت سنة 2016 بسعر 800 دج/ل، في حين كانت مبيعاتها في سنة 2017 من نفس المادة 2300 لتر بسعر 1200 دج/ل، أوجد منسوب القيمة إذا اعتبرنا سنة 2016 هي سنة الأساس.

الحل:

$$V_{1/0} = \frac{V_1}{V_0} \times 100 = \frac{P_1 q_1}{P_0 q_0} \times 100$$

$$V_{2017/2016} = \frac{4500 \times 800}{2300 \times 1200} \times 100 = 130.43\%$$

وهذا يعني أن قيمة مبيعات المعاصرة زادت بنسبة 43. 30% في سنة 2017 عما كانت عليه في سنة 2016.

2.الصيغ المجمعة للأرقام القياسية

نرغب في بعض الأحيان دراسة تطور بعض الظواهر المعقدة مثل تطور المستوى العام للأسعار، تطور حجم الصادرات، وتطور حجم الواردات ... الخ، هذا التطور لا يمكن التعبير عنه بالأرقام القياسية البسيطة لأنها تبين تطور ظاهرة واحدة، بل يستخدم في ذلك ما يسمى بالأرقام القياسية التجميعية التي تنقسم إلى:

1.2.الأرقام القياسية التجميعية البسيطة: وهي تمثل النسبة بين أسعار (أو كميات) مجموعة من المواد بين فترتين زمنيتين (مكانين) مختلفتين، و يعطى بالصيغة التالية:

1.1.2.الرقم القياسي التجميعي البسيط للأسعار:

$$Ip_{1/0} = \frac{\sum_{j=1}^{n} P_1^j}{\sum_{j=1}^{n} P_0^j} \times 100$$

حيث: n هو عدد المواد

P_0^j : سعر المادة J في سنة الأساس.

P_1^j : سعر المادة J في سنة المقارنة.

2.1.2. الرقم القياسي التجميعي البسيط للكميات:

$$Iq_{1/0}=\frac{\sum_{j=1}^{n}q_1^j}{\sum_{j=1}^{n}q_0^j}\times 100$$

حيث: n هو عدد المواد

q_0^j : كمية المادة J في سنة الأساس.

q_1^j : كمية المادة J في سنة المقارنة.

مثال07-03: البيانات التالية تبين أسعار وكميات خمسة أنواع من زيوت المحركات، اشتريت من طـــرف مؤسسة عمومية ما في سنتي 2014 ، 2017.

	السنة			
	2014		2017	
نوع الزيت	السعر (دج/لتر)	الكمية (لتر)	السعر (دج/لتر)	الكمية (لتر)
النوع الأول	340	45	340	70
النوع الثاني	370	65	400	100
النوع الثالث	450	210	470	300
النوع الرابع	350	80	300	90
النوع الخامس	200	70	190	90

- الرقم القياسي التجميعي البسيط للأسعار يحسب كما يلي:

$$Ip_{2017/2016}=\frac{\sum_{j=1}^{5}P_{2017}^j}{\sum_{j=1}^{5}P_{2016}^j}\times 100=\frac{340+400+470+300+190}{340+370+450+350+200}\times 100=\frac{1700}{1710}\times 100=99.42\%$$

أي أن الأسعار التي اشترت بها المؤسسة العمومية قد انخفضت سنة 2017 بـــ:

$$99.42\%-100\%=-0.58\%$$

- الرقم القياسي التجميعي البسيط للكميات يحسب كما يلي:

$$Iq_{2017/2016}=\frac{\sum_{j=1}^{5}q_{2017}^{j}}{\sum_{j=1}^{5}q_{2016}^{j}}\times100=\frac{70+100+300+90+90}{45+65+210+80+70}\times100=\frac{650}{470}\times100=138.30\%$$

أي أن الكمية التي اشترتها المؤسسة العمومية قد زادت سنة 2017 بـــ:

$$138.30\%-100\%=38.30\%$$

2.2.الأرقام القياسية التجميعية المرجحة: تعتمد الأرقام القياسية التجميعية المرجحة على طريقة التـــرجيح بواسطة معامل معين يبين الأهمية النسبية.

وهناك ثلاث صيغ للأرقام القياسية المرجحة تعتمد على ما إذا كنا نستخدم كميات أو أسعار سنة الأساس أو المقارنة أو السنة النموذجية كمايلي:

1.2.2.الرقم القياسي لاسبير *laspeyres* : اعتمد "لاسبير" في حسابه للأرقام القياسية على أن التـــرجيح يكون بواسطة سنة الأساس. و هنا نميز الأرقام القياسية التالية:

أ. **الرقم القياسي لاسبير للأسعار**: يفترض في هذه الصيغة ثبات أذواق المستهلكين واستمرارهم في استهلاك نفس كميات السلع حتى ان تغيرت أسعارها، لكن في الواقع القاعدة معروفة في الحالات العادية: وهي العلاقة العكسية بين السعر والكميات المطلوبة. ويعطي الرقم القياسي لاسبير للأسعار بالصيغة:

$$L_{(P)\,1/0}=\frac{\sum_{j=1}^{n}P_1^j\cdot q_0^j}{\sum_{j=1}^{n}P_0^j\cdot q_0^j}\times100$$

ب.**الرقم القياسي لاسبير للكميات**: يفترض في هذه الصيغة ثبات الأسعار في سنتي الأساس والمقارنة بغض النظر عن تغير الكميات المستهلكة في السنتين. ويعطي الرقم القياسي لاسبير للكميات بالصيغة:

$$L_{(q)\,1/0}=\frac{\sum_{j=1}^{n}P_0^j\cdot q_1^j}{\sum_{j=1}^{n}P_0^j\cdot q_0^j}\times100$$

ج.**الرقم القياسي لاسبير للقيم الكلية**: ويعطي الرقم القياسي لاسبير للقيم الكلية بالصيغة:

$$L_{(VG)\,1/0}=\frac{\sum_{j=1}^{n}P_1^j\cdot q_1^j}{\sum_{j=1}^{n}P_0^j\cdot q_0^j}\times100$$

مثال07-04: البيانات التـــالية تبين أسعـــار وكميات خمس مواد استهلاكية، اشتريت من طرف عائلة عبد الرحمن خلال سنتي 2015 و2017.

المادة	السنة			
	2015		2017	
	السعر P_0	الكمية q_0	السعر P_1	الكمية q_1
الـــدقيق دج/كغ	39	45	50	70
الـــزيت دج/لتر	52	65	59	100
الكهرباء دج/كيلواط	1.2	210	1.7	300
الغاز الطبيعي دج/م^3	0.23	80	0.36	90
القمـــاش دج/متر	170	70	240	90

المطلوب:

1. أحسب الرقم القياسي لاسبير للأسعار.
2. أحسب الرقم القياسي لاسبير للكميات.
3. أحسب الرقم القياسي لاسبير للقيم الكلية.

الحل:

1. **حساب الرقم القياسي لاسبير للأسعار:**

$$L_{P\,1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_0^j}{\sum_{j=1}^{n} P_0^j \cdot q_0^j} \times 100$$

$$= \frac{(50\times45)\times(59\times65)+(1.7\times210)+(0.36\times80)+(240\times70)}{(39\times45)\times(52\times65)+(1.2\times210)+(0.23\times80)+(170\times70} \times 100$$

$$= \frac{23270.8}{17305.4} \times 100 = 134.47\%$$

أي أن الأسعار التي اشترت بها عائلة عبد الرحمن قد زادت سنة 2017 بـــ:

$$134.47\% - 100\% = 34.47\%$$

2. حساب الرقم القياسي لاسبير للكميات:

$$Lq_{1/0}=\frac{\sum_{j=1}^{n}P_0^j\cdot q_1^j}{\sum_{j=1}^{n}P_0^j\cdot q_0^j}\times 100$$

$$=\frac{(70\times 39)\times(100\times 52)+(300\times 1.2)+(90\times 0.23)+(90\times 170)}{(39\times 45)\times(52\times 65)+(1.2\times 210)+(0.23\times 80)+(170\times 70}\times 100$$

$$=\frac{23610.7}{17305.4}\times 100=136.44\%$$

أي أن كمية المواد التي اشترتها عائلة عبد الرحمن قد زادت سنة 2017 بـــ:

$$136.44\%-100\%=36.44\%$$

3. حساب الرقم القياسي لاسبير للقيم الكلية:

$$L_{(VG)1/0}=\frac{\sum_{j=1}^{n}P_1^j\cdot q_1^j}{\sum_{j=1}^{n}P_0^j\cdot q_0^j}\times 100$$

$$=\frac{(70\times 50)\times(100\times 59)+(300\times 1.7)+(90\times 0.36)+(90\times 240)}{(39\times 45)\times(52\times 65)+(1.2\times 210)+(0.23\times 80)+(170\times 70}\times 100$$

$$=\frac{31542.4}{17305.4}\times 100=182.27\%$$

أي أن هناك زيادة في القيم الكلية سنة 2017 بـــ:

$$182.27\%-100\%=82.27\%$$

2.2.2 .الرقم القياسي باش *Pache*: اعتمد "باش" في حسابه للأرقام القياسية علـــى أن التـــرجيح يكون بواسطة سنة المقارنة. و هنا نميز الأرقام القياسية التالية:

أ. **الرقم القياسي باش للأسعار**: يقيس التغير في النفقات للحصول على كميات السلع في سنة المقارنة مرجعة بأسعار سنة المقارنة وأسعار سنة الأساس. وبذلك يفترض هذا الرقم أن نفس الكميات سنة المقارنة كانت قد استهلكت في سنة الأساس. ويعطي الرقم القياسي باش للأسعار بالصيغة:

$$P_{(P)1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_1^j}{\sum_{j=1}^{n} P_0^j \cdot q_1^j} \times 100$$

ب.**الرقم القياسي باش للكميات**: يفترض أن المستهلك يقيم ما يستهلكه في كل من سنتي الأساس والمقارنة بنفس سنة المقارنة. ويعطي الرقم القياسي باش للكميات بالصيغة:

$$P_{(q)1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_1^j}{\sum_{j=1}^{n} P_1^j \cdot q_0^j} \times 100$$

ج.**الرقم القياسي باش للقيم الكلية**: ويعطي الرقم القياسي باش للقيم الكلية بالصيغة:

$$P_{(VG)1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_1^j}{\sum_{j=1}^{n} P_0^j \cdot q_0^j} \times 100$$

مثال07-05: بالرجوع لبيانات **المثال07-04** يكون الرقم القياسي لباش كمايلي:

1. **حساب الرقم القياسي باش للأسعار:**

$$P_{(P)1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_1^j}{\sum_{j=1}^{n} P_0^j \cdot q_1^j} \times 100$$

$$= \frac{(50\times70)\times(59\times100)+(1.7\times300)+(0.36\times90)+(240\times90)}{(39\times70)\times(52\times100)+(1.2\times300)+(0.23\times90)+(170\times90)} \times 100$$

$$= \frac{31542.4}{23610.7} \times 100 = 133.59\%$$

أي أن الأسعار التي اشترت بها عائلة عبد الرحمن قد زادت سنة 2017 بـــ:

$$133.59\% - 100\% = 34.47\%$$

حساب الرقم القياسي باش للكميات:

$$P_{(q)1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_1^j}{\sum_{j=1}^{n} P_1^j \cdot q_0^j} \times 100$$

$$= \frac{(50\times 70)\times(59\times 100)+(1.7\times 300)+(0.36\times 90)+(240\times 90)}{(50\times 45)\times(59\times 65)+(1.7\times 210)+(0.36\times 80)+(240\times 70}\times 100$$

$$= \frac{31542.4}{23270.8}\times 100 = 135.55\%$$

أي أن كمية المواد التي اشترتها عائلة عبد الرحمن قد زادت سنة 2017 بـــ:

$$135.55\% - 100\% = 35.55\%$$

2. **حساب الرقم القياسي باش للقيم الكلية:**

$$P_{(VG)1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot q_1^j}{\sum_{j=1}^{n} P_0^j \cdot q_0^j} \times 100$$

$$= \frac{(70\times 50)\times(100\times 59)+(300\times 1.7)+(90\times 0.36)+(90\times 240)}{(39\times 45)\times(52\times 65)+(1.2\times 210)+(0.23\times 80)+(170\times 70}\times 100$$

$$= \frac{31542.4}{17305.4}\times 100 = 182.27\%$$

أي أن هناك زيادة في القيم الكلية سنة 2017 بـــ:

$$182.27\% - 100\% = 82.27\%$$

3.2.2.الرقم القياسي فيشر *Fisher*: ما يعاب على الرقم القياسي لاسبير أنه متحيز إلى أعلى بالنظر إلى أنه مبني على الترجيح بأوزان نسبة الأساس، في حين أنه ما يعاب على الرقم القياسي باش أنه متحيز إلى أسفل لأنه يستند على الترجيح بأوزان سنة المقارنة، وعليه فقد اقترحت عدة صيغ لمعالجة الفرق بين الترجحين فكانت صيغة فيشر.

إن الرقم القياسي لــ "فيشر" هو المتوسط الهندسي البسيط للرقمين القياسيين لــ "لاسبير" و"باش" ويحسب كما يلي:

$$F_{1/0} = \sqrt{L_{1/0} \times P_{1/0}}$$

و هنا نميز الأرقام القياسية التالية:

أ. **الرقم القياسي فيشر للأسعار:**

$$F_{(P)1/0} = \sqrt{L_{(p)\,1/0} . P_{(P)\,1/0}}$$

ب. الرقم القياسي فيشر للكميات:

$$F_{(q)1/0} = \sqrt{L_{(q)\,1/0} . P_{(q)\,1/0}}$$

ج. **الرقم القياسي فيشر للقيم الكلية:**

$$F_{(VG)1/0} = \sqrt{L_{(VG)\,1/0} . P_{(VG)\,1/0}}$$

مثال07-06: بالرجوع لبيانات **المثال07-04** يكون الرقم القياسي لفيشر كمايلي:

- **الرقم القياسي فيشر للأسعار:**

$$F_{(P)1/0} = \sqrt{L_{(p)\,1/0} . P_{(P)\,1/0}}$$
$$= \sqrt{134.47 \times 133.59} = 134.03$$

- **الرقم القياسي فيشر للكميات:**

$$F_{(q)1/0} = \sqrt{L_{(q)\,1/0} . P_{(q)\,1/0}}$$
$$= \sqrt{136.44 \times 135.55} = 135.99$$

- **الرقم القياسي فيشر للقيم الكلية:**

$$F_{(VG)1/0} = \sqrt{L_{(VG)\,1/0} . P_{(VG)\,1/0}}$$
$$= \sqrt{182.27 \times 182.27} = 182.27$$

4.2.2. الرقم القياسي مارشال *Marshall*: اعتمد "مارشال" في حسابه للأرقام القياسية على أن الترجيح يكون بواسطة سنة المقارنة وسنة الأساس معا ، و هنا نميز بين أيضا بين نوعين من الأرقام القياسية:

أ. **الرقم القياسي مارشال للأسعار**: يعطى بالصيغة التالية:

$$Mp_{1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot \left(q_0^j + q_1^j\right)}{\sum_{j=1}^{n} P_0^j \cdot \left(q_0^j + q_1^j\right)} \times 100$$

ب. الرقم القياسي مارشال للكميات: يعطى بالصيغة التالية:

$$Mq_{1/0} = \frac{\sum_{j=1}^{n} \left(P_0^j + P_1^j\right) \cdot q_1^j}{\sum_{j=1}^{n} \left(P_0^j + P_1^j\right) \cdot q_0^j} \times 100$$

مثال07-07: بالرجوع لبيانات **المثال07-04** يكون الرقم القياسي مارشال كمايلي:

- الرقم القياسي مارشال للأسعار:

$$Mp_{1/0} = \frac{\sum_{j=1}^{n} P_1^j \cdot \left(q_0^j + q_1^j\right)}{\sum_{j=1}^{n} P_0^j \cdot \left(q_0^j + q_1^j\right)} \times 100$$

$$= \frac{50\times(45+70)+59\times(65+100)+1.7\times(210+300)+0.36\times(80+90)+240\times(70+90)}{39\times(45+70)+52\times(65+100)+1.2\times(210+300)+0.23\times(80+90)+170\times(70+90)} \times 100$$

$$Mp_{1/0} = \frac{54813.2}{40916.1} \times 100 = 133.97$$

- الرقم القياسي مارشال للكميات:

$$Mq_{1/0} = \frac{\sum_{j=1}^{n} \left(P_0^j + P_1^j\right) \cdot q_1^j}{\sum_{j=1}^{n} \left(P_0^j + P_1^j\right) \cdot q_0^j} \times 100$$

$$= \frac{70\times(50+39)+100\times(59+52)+300\times(1.7+1.2)+90\times(0.36+0.23)+90\times(240+170)}{45\times(50+39)+65\times(59+52)+210\times(1.7+1.2)+80\times(0.36+0.23)+70\times(240+170)} \times 100$$

$$Mp_{1/0} = \frac{55153.1}{40576.2} \times 100 = 135.92\%$$

3.الأرقام القياسية ذات الأساس المتحركة: تعتمد الأرقام القياسية ذات الأساس المتحرك على فكرة تحريك سنة الأساس دوريا كل سنة، لأنه إذا كان هناك فاصل طويل بين سنة المقارنة وسنة الأساس، فإن الظروف المحيطة تتغير وبشكل كبير في بعض الأحيان.

مثال07-08: تمثل البيانات التالية تطور أسعار الأسهم بالدينار لشركة ما خلال الفترة 2014–2017

السنة	2014	2015	2016	2017
السعر	230	275	280	277

الحل: نستعين بالجدول لحساب الأرقام القياسية للأسعار كمايلي:

	الأرقام القياسية للأسعار		
	$P_{2015/2014}$	$P_{2016/2015}$	$P_{2017/2016}$
	275/230	280/275	277/280
	119.57%	101.82%	98.93%
التفسير	%19.57+	+1.82%	-1.07%

المراجـــــع

- جلاطو جيلالي، الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الطبعة الثامنة، الجزائر، 2010.
- حليمي عبد القادر، مدخل إلى الإحصاء، ديوان المطبوعات الجامعية، الطبعة الخامسة، الجزائر، 2004.
- حيدوشي عاشور، محاضرات في الإحصاء الوصفي، مطبوعة موجهة لطلبة السنة الأولى جذع مشترك، جامعة أكلي محند أولحاج البويرة، الجزائر، 2016.
- خزار محمد بونوارة، مبـــادئ الإحصاء، منشورات جامعة باتنة، الجزائر، 1996.
- خليل شرف الدين، الإحصاء الوصفي، شبكة الأبحاث والدراسات الاقتصادية.
- راتول محمد، الإحصـــاء الوصفي، ديوان المطبوعات الجامعية، الطبعة الثالثة، الجزائر، 2009.
- عبد الناصر موسى، دروس في الإحصاء الوصفي، مطبوعة موجهة لطلبة السنة الأولى جذع مشترك، جامعة محمد خيضر بسكرة، الجزائر، 2007.
- عبيدات محمد وآخرون، منهجية البحث العلمي، القواعد والمراحل والتطبيقات، دار وائل للطباعة والنشر، الطبعة الثانية، عمان، 2006.
- عوض مراد كمال، أســاسيات الإحصاء، دار البداية ناشرون وموزعون، الطبعة الأولى، عمان، 2008.
- موساوي عبد النور، بركان يوسف، الإحصاء 1، دروس-تمارين محلولة-تطبيقات، دار العلوم للنشر والتوزيع، عنابة، الجزائر، 2009.

- *Dominick SALVATORE & Derrick REAGLE, Theory and Problems of Statistics and Econometrics, Second Edition, Schaum's Outline Series, McGRAW-HILL, companies, New York,2002.*
- *Fabrice MAZEROLLE, Statistique Descriptive,Notes de cours, 2009.*