وزارة التعليم العالي والبحث العلمي

المدرسة العليا للتجارة

مطبوعة الإحصاء الوصفي
الخاصة بطلبة الأقسام التحضيرية

من إعداد : الأستاذة نصيب حفيزة
أستاذ محاضر "أ"

السنة الجامعية

2019 - 2018

مفاهيم أساسية
Les concepts de base

تتطلب دراسة الإحصاء تناول مجموعة من المفاهيم الأساسية تشكل مفردات لغة علم الإحصاء، وتستعمل بصورة شائعة في كل الدراسات الإحصائية بمعاني دقيقة ومحددة ويجب أن تكون جلية في ذهن كل دارس للإحصاء ويمكن حصرها في الكلمات الدالة التالية:

- ❖ الإحصاء والإحصائيات – الإحصاء العام أو التحقيقات الإحصائية الشاملة – إحصاء العينة أو التحقيقات الإحصائية الجزئية – المجتمع الإحصائي – العينة – الوحدة الإحصائية أو الفرد الإحصائي – الصفة – المتغيرات الإحصائية – سلم قياس المتغيرات الإحصائية.

1- الإحصاء والإحصائيات La statistique et les statistiques

ويبحث **علم الإحصاء (**La statistique**)** طرق جمع المعطيات أو البيانات المتعلقة بمختلف الظواهر الطبيعية والإنسانية وتنظيمها وعرضها وتحليلها و استقراء النتائج، وذلك لاتخاذ القرار في ظل عدم اليقين.

يقصد بجمع المعطيات الإحصائية أو البيانات الإحصائية عملية الحصول على القياسات أو التعدادات أو قيم المشاهدات المعبرة عن الظاهرة موضوع الدراسة، وتسمى هذه القيم ب**الإحصائيات (**les statistiques)، والتي قد تكون كمية أو كيفية، وهي عملية بالغة الأهمية لأن الخطأ في جمع البيانات يقوض(dommages) أسس العملية الإحصائية برمتها.

وتشمل عملية تنظيم وعرض البيانات وضع ما تم جمعه من معطيات في جداول منسقة، وعرضها بطرق مناسبة و معبرة كالأشكال الهندسية والرسوم البيانية.

كما تشير عملية تحليل البيانات أو الإحصائيات إلى توظيف المعطيات قيد الدراسة في إيجاد قيم لمقاييس معينة تسمح بالتعبير الموجز أو الملخص عن الظاهرة المدروسة تسمى **المعلمات** (les paramètres).

أما استقراء النتائج واتخاذ القرار فهي من أهم أهداف علم الإحصاء، باعتبارها تشكل مقصدا لمعظم الدراسات والنظريات الإحصائية والتطبيقات العملية لها، وهي تتمحور حول الاستنتاجات التي يتوصل لها الباحث من تحليل البيانات وهي غالبا ما تكون على شكل تقديرات لقيم معبرة عن الظاهرة أو قياس لعلاقات بين مختلف عناصر الظاهرة المدروسة، أو تنبؤات بقيم لها، أو تعميمات تتخذ على ضوئها قرارات رفض أو قبول الفرضيات الإحصائية. ويقسم علم الإحصاء إلى قسمين:

- ❖ **الإحصاء الوصفي** La statistique descriptive

يختص بطرق وصف وجمع وتلخيص البيانات الإحصائية مع بعض من التحليل لها، ويمكن القول أن الإحصاء الوصفي هو قراءة بسيطة ولكن بتركيز واهتمام لاستيعاب أبعاده المختلفة، وباعتباره يقدم صورا ملموسة تشكل مرجعا لاستيعاب المفاهيم المجردة في الإحصاء الرياضي، فهو يمثل مدخلا تحضيريا لفهم الإحصاء الرياضي، وبذلك فهو طريقة وليس نظرية، وبإيجاز يمكن القول أن الإحصاء الوصفي هو قاعدة البداهة في علم الإحصاء.

- ❖ **الإحصاء الرياضي** La statistique mathématique

يرتبط الإحصاء بالاحتمالات ارتباطا وثيقا ولذا ينطلق من نظرية الاحتمالات ويتناول بعد ذلك استنتاج القوانين والنظريات الإحصائية وفق الأسس الرياضية، كذلك يبحث في التوزيعات الاحتمالية المختلفة، ويصيغ لها الدوال الرياضية بهدف التوصل إلى خصائصها ومعالمها أو مؤشراتها، كما تستخدم هذه الدوال في إعداد جداول التوزيعات الإحصائية.

- ❖ **الإحصاء الاستدلالي** أو **الاستقرائي** La statistique inférentielle

تبنئ الدراسات الإحصائية في غالب الأحيان على التحقيقات الإحصائية الجزئية بمعنى استعمال الجزء للاستدلال على الكل وبعبارة أخرى نستدل عن معلمات المجتمع الإحصائي بحسابها على مستوى العينة، وينشأ عدم اليقين من عدم معرفة القيم الأكيدة للمعلمات لكل المجتمع فنلجأ إلى الاستدلال عنها بواسطة قيم تقديرية غير أكيدة نحسبها من العينة ولذا يعتمد الاستدلال الإحصائي اعتمادا كليا على الاحتمالات وبإيجاز فإن الإحصاء الاستدلالي يتناول الاستدلال عن معالم المجتمع عبر الاستقراء أو الاستدلال بالجزء عن الكل.

❖ أهمية الإحصاء

تبرز الأهمية البالغة لعلم الإحصاء في وقتنا الحاضر من الدور الهام الذي يلعبه في مختلف القضايا التي تهم واقع الإنسان ومستقبله. فزيادة على أهميته في حياتنا اليومية سواء من جانب العد أو التصنيف أو إدراك طبيعة العلاقات بين مختلف مكونات الظواهر أو التنبؤ بقيمها المستقبلية ..إلخ، فإن رسم وتقييم السياسات و الخطط الاقتصادية والاجتماعية و التربوية ...إلخ يقوم في الأساس على توفر قواعد من المعطيات الإحصائية تغطي مختلف أوجه الحياة الإنسانية.

كما أن استخداماته في مختلف الفنون والعلوم بارزة، فلا يكاد يستغني علم من العلوم التجريبية أو الإنسانية عن الإحصاء. ففي العلوم الإنسانية عموما و العلوم التجارية خصوصا يلعب الإحصاء دورا بالغ الأهمية سواء من ناحية توصيف الواقع الاقتصادي أو عبر تطبيقاته الواسعة والضرورية التي تغطي مختلف أوجه النشاط الاقتصادي التي تسمح للباحثين بصياغة مختلف العلاقات المحاسبية والسلوكية في الاقتصاد.

❖ علاقة الإحصاء بالعلوم الأخرى

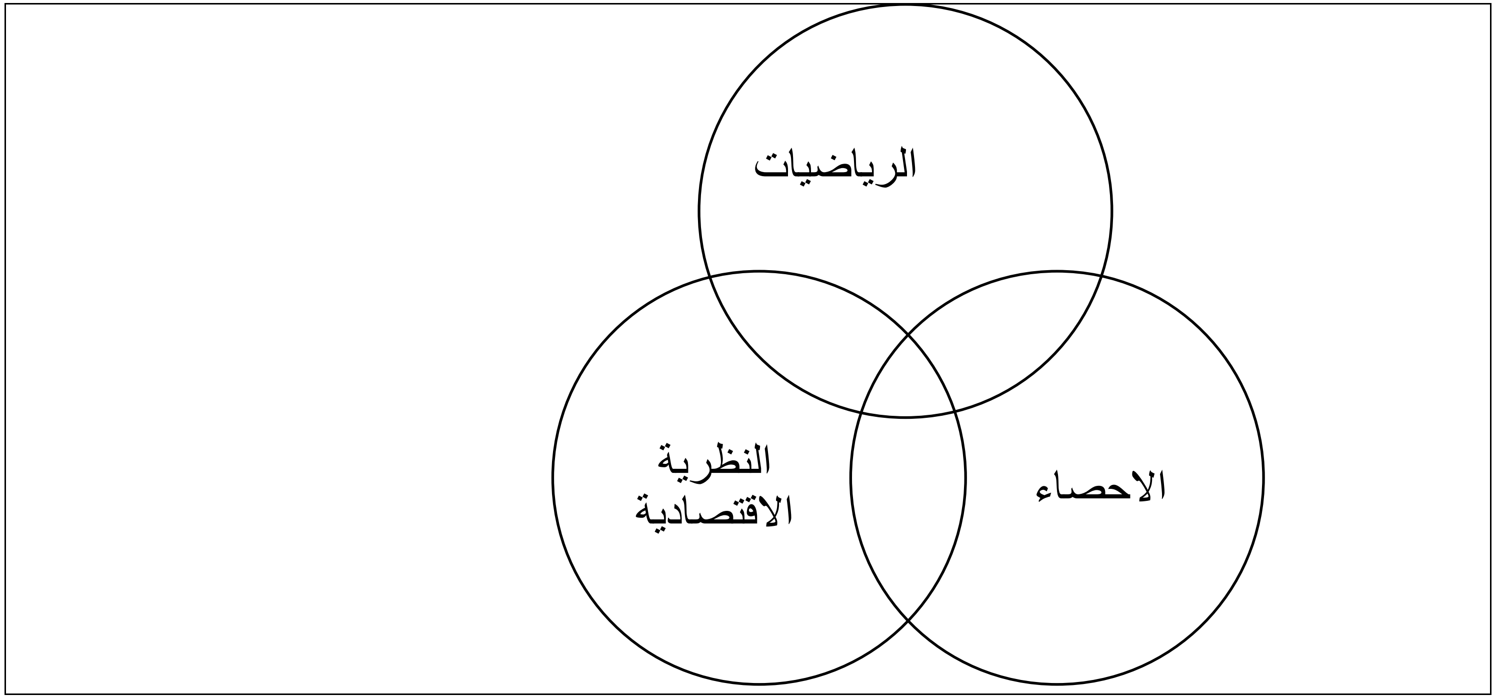
يمكن النظر إلى الإحصاء من زاوية نظرية ومن زاوية تطبيقية. فمن الناحية النظرية يعتبرا لإحصاء علم مستقل لكن تطبيقاته تغطي كل مجلات الحياة حيث يستخدم في الحياة اليومية وتستعمله العلوم التجريبية والعلوم الإنسانية على نطاق واسع.

ويغطي الإحصاء التطبيقي أو التجريبي كل مجالات حياة الإنسان. فالإحصاء الاقتصادي يشمل كل ما يتم إنتاجه من معطيات إحصائية تتعلق بالاقتصاد، فالمحاسبة الوطنية مثلا تمدنا بالإحصائيات عن الاقتصاد الوطني بالتعبير الكمي عن جميع العمليات الاقتصادية كالإنتاج والاستهلاك والتوزيع، ومعدل التضخم يعبر عن استقرار الأوضاع الاقتصادية، ومعدل البطالة يبرز الاختلال أو عدم التوازن الاقتصادي ومؤشر جيني هو قياس تقديري للعدالة الاجتماعية.

وبصورة عامة، فالإحصاء يمكننا من التعبير الكمي عن الظواهر الاقتصادية والاجتماعية عبر إعداد قياسات إحصائية في صورة مؤشرات ومعالم تترجم مستويات الظاهرة قيد الدراسة.

كما يستخدم علم الإحصاء لتفسير الظواهر الاقتصادية المختلفة كنظريات الطلب والعرض، والعلاقة بين مستويات الدخل والإنفاق الاستهلاكي مثلا، وقياس العلاقات الاقتصادية المختلفة وفي مراقبة الإنتاج في الشركات الصناعية من حيث الكمية والجودة و ومدى ملائمة كل ذلك لاحتياجات السوق وأذواق المستهلكين.

تستعمل الأبحاث الإحصائية الاحتمالات على نطاق واسع، فالاستدلال الإحصائي يعتمد اعتمادا كليا على الاحتمال كتعبير عن عدم اليقين، فمع غياب الحتمية وللابتعاد عن التخبط العشوائي يتم استعمال الاحتمال كنموذج للتفسير، وتبرز الصلة بين الإحصاء والرياضيات من خلال نظرية الاحتمال التي تمثل أحد فروع الرياضيات.



2- أساليب جمع المعطيات الإحصائية

تسمح العمليات الإحصـائية بالتعبير عن مختلف الظواهر الطبيعية و الاقتصـادية و الاجتماعية بمعطيات (إحصـائيات)، وترتبط أهداف جمع المعطيات الإحصائية بالهدف من الدراسة الإحصائية للظاهرة كقياسها وإبراز خصائصها وفرز مكوناتها وإيجاد العلاقات بين عناصـرها والتنبؤ بقيمها، كما أنها تتعلق بتطور الأوضـاع الاقتصـادية والاجتماعية للبلاد ويتم الحصـول على المعطيات الإحصائية بإجراء التحقيقات الإحصائية، ونميز بين نوعين من التحقيقات الإحصائية هي:

- **الإحصاء العام أو البحوث الشاملة** Recensement général ou Les enquêtes exhaustives

تشمل عملية جمع المعطيات الإحصـائية وفق هذه الطريقة كل أفراد المجتمع الإحصـائي، وأشـهر التحقيقات الإحصـائية الشـاملة في الجزائر هو الإحصـاء العام للسكان والسكن (RGPH) الذي ينجزه دوريا الديوان الوطني للإحصـائيات، ويسمح بجمع كم كبير من المعطيات الإحصائية عن كل شخص تم إحصاءه مثل الجنس، الحالة العائلية،

السـن، المسـتوى التعليمي، المهنة نوع السـكن ...الخ. كانت الجزائر تنظم هذا الإحصـاء العام كل عشـر (10) سـنوات وأصبحت تنظمه كل (5) خمس سنوات ويستغرق إنجازه 15 يوما من شـهر افريل، في سنة 2008 تم إجراء الإحصـاء العام الخامس للسكان والسكن. كما أنجز الديوان الوطني إحصاءات عامة أخرى حول الفلاحة والصناعة.

للإحصـــاء العام أهمية بالغة باعتباره قاعدة بيانات (base de données) فهو يمثل مصـــدرا للمعطيات التي تشـــكل مرجعا لا غنى عنه لرسم السياسات الاقتصادية والاجتماعية في البلاد، كما يشكل منطلقا لإجراء التحقيقات الإحصائية الجزئية كالتحقيق حول الإنفاق اسـتهلاكي للأسـر مثلا، إلا انه زيادة على ما يسـتغرقه الإحصـــاء العام من وقت في مختلف مراحل انجازه فانه يقتضـي رصـد إمكانيات بشرية ومادية كبيرة سواء في تحضيره أو في إنجازه أو في استغلال نتائجه مما يستدعي اللجوء في الكثير من الحالات إلى البحوث الإحصائية الجزئية.

- **التحري أو البحوث الجزئية** Le Sondage ou Les enquêtes partielles

تتعلق عملية جمع المعطيات الإحصـــائية في هذه الحالة بمجوعة جزئية من المجتمع الإحصـــائي هي العينة، التي يتم تحديد وحداتها الإحصائية باستغلال قاعدة البيانات التي يوفرها الإحصاء العام.

يتم اللجوء إلى التحقيقات الإحصائية الجزئية أو إحصائيات العينة تجنبا لتكلفة الإحصاء العام الكبيرة وربحا للوقت، كما انه لا يمكن عمليا في بعض الحالات إجراء الإحصـــاء العام، ففي حالة دراســـة مدى مطابقة المنتجات الغذائية المســـتوردة للمواصفات المطلوبة مثلا، يتم اخذ عينة منها ودراستها لأنه لا يمكن فحص كل الكمية المستوردة. و من بين أشهر التحقيقات الإحصـائية الجزئية تحقيقات اسـتهلاك الأسـر الذي يشـكل أسـاس إعداد الرقم القياسـي لأسـعار الاسـتهلاك وتحقيق اليد العاملة وسبر الآراء.الخ

3- طرق جمع المعطيات الإحصائية

لمباشـرة عملية جمع المعطيات الإحصـائية يلجأ الإحصـائي إلى الطرق التالية : المقابلة الشـخصـية أو المراسـلة أو الاتصـال عبر الهاتف أو البريد الالكتروني ويقوم جامع الإحصـائيات أو المسـتقصـي وفق الطريقة الأولى بمقابلة كل فرد من أفراد المجتمع الإحصـائي أو العينة وتوجيه إليه الأسئلة الموجودة في الاسـتبيان (Le questionnaire) أو الاسـتمارة وتدوين الإجابة في المكان المخصص أمام كل سؤال.

أما الطريقة الثانية فيقوم المسـتقصـي بإرسـال الاسـتبيان بالبريد إلى الأفراد المعنيين بالبحث الإحصـائي مرفقا أولا بتوضـيحات لأهداف البحث وأهميته وثانيا بتوجيهات للإجابة على أسـئلة الاسـتبيان، وعادة ما يرفق مع الاسـتبيان مظروف بعنوان الهيئة التي تقوم بالبحث وعليه طابع بريدي لإعادة الاسـتبيان بعد اسـتيفائه ويتم ذلك في ظل ما هو متاح من وسـائل اتصال منها استعمال وسائل الاتصال الالكترونية .

4- المجتمع الإحصائي La population statistique

هو مجموعة العناصر المشتركة في الصفة والمتعلقة بالظاهرة قيد الدراسة، والمحددة في الزمان والمكان، فمثلا عند دراسة حوادث السير على الطريق الوطني رقم 5 في ولاية سطيف في سنة 2006، فإن المجتمع الإحصائي هو جميع حوادث السير التي وقعت في إقليم ولاية سطيف وعلى الطريق الوطني رقم 5 وفي سنة 2005.

يتم تعريف المجتمع الإحصائي بحسب الظاهرة المدروسة مع تحديد الصفة المدروسة و الزمان و المكان بدقة تسمح للباحث بالجزم إن كان عنصر ما ينتمي للمجتمع الإحصائي أم لا، ويسمى هذا العنصر الوحدة الإحصائية أو الفرد الإحصـائي.

5 -العينة L’échantillon

العينة هي مجموعة جزئية من المجتمع الإحصائي، ولضمان تمثيلها للمجتمع الإحصائي تمثيلا صادقا يتم اختيارها وفق طرق علمية محددة، ونميز بين أنواع متعددة من العينات فهناك العينة غير العشوائية والعينة العشوائية التي قد تكون بسيطة أو منتظمة أو الطبقية...إلخ

ومصطلح العينة شائع في حياتنا اليومية، إذ عندما يمرض الشخص يطلب الطبيب إجراء تحاليل عن دمه ويتم بفحص عينة من دمه، كما أن السماح بتسويق المواد الاستهلاكية الغذائية وغير الغذائية تتم بعد إجراء اختبارات مطابقتها للمواصفات المطلوبة، فيتم ذلك عبر سحب عينة منها ودراستها في المخبر.

ويتم اللجوء إلى استخدام العينات في الدراسات الإحصائية كبديل عن البحوث الإحصائية الشاملة نظرا للتكلفة المرتفعة في غالب الأحيان لهذه الأخيرة و استحالة إجرائها في بعض الأحيان وربحا للوقت في أحيان أخرى.

6 - الفرد الاحصائي أو الوحدة الإحصائية L’individu statistique ou L’unité statistique

الوحدة الإحصائية أو الفرد الإحصائي هي عنصر من المجتمع الإحصائي نجمع عنه البيانات الإحصائية، والوحدات الإحصائية قد تكون وحدات طبيعية تتعلق بالجنس البشري كالموظف والأسرة والفرد والطالب..إلخ أو وحدات معنوية كالمؤسسة والحوادث ...إلخ، وينبغي عند تنفيذ البحوث الإحصائية تعريف الوحدة الإحصائية تعريفا واضحا يمنع التداخل مع الوحدات التي لا يشملها البحث.

نميز بين وحدة إحصائية بسيطة ووحدة إحصائية مركبة. فالوحدة الإحصائية البسيطة هي العنصر غير القابل للتجزئة مثلا إنسان، عامل، مصباح، كتاب...إلخ والوحدة الإحصائية المركبة هي العنصر الذي يمكن تجزئته إلى عناصر بسيطة أو عناصر مركبة من درجة اقل تركيبا، مثلا عند دراسة المؤسسات الصناعية التي تشغل أكثر من 10 عمال في ولاية تيبازة في الثلاثي الأخير من سنة 2005 يمكن أن نهتم بالمؤسسة ككل ونجمع المعطيات عن المتغيرات المدروسة مثل رقم الأعمال، عدد العمال...إلخ، ويمكن أن نهتم مثلا بالمصالح المالية أو الو رشات، وهذه الوحدة تتضمن وحدات أخرى التي تتألف بدورها من عناصر بسيطة هي العامل، الآلة ...إلخ

7-الصفة La modalité

وهي مختلف خصائص او أوضاع او حالات الوحدات الإحصائية، وتسمح بوصفها و تمييزها عن بعضها البعض داخل المجتمع الإحصائي.

فالمجتمع الإحصائي المتعلق بطلبة السنة الأولى في المدرسة العليا للتجارة في سنة **2019/2018** وحدته الإحصائية هي طالب مسجل في السنة الأولى بالمدرسة العليا للتجارة ولكل طالب مجموعة من الصفات تميزه عن غيره، وهذه الصفة قد تكون مثلا الجنس أو لون العينين أو الملاحظة في البكلوريا أو الوزن أو القامة...إلخ وللتعبير إحصائيا عن هذه الخصائص نستعمل مصطلح المتغيرات الإحصائية .

نسجل أن لكل صفة مجموعة من الحالات الممكنة لا تقل عن اثنين.

8 -المتغيرات الإحصائية Les variables statistiques

باعتبار أن كل صفة تأخذ مجموعة من الخصائص أو الحالات الممكنة وتختلف من فرد إلى آخر فإننا نستخدم تعبير متغير إحصائي للتعبير عن هذه القيم أو هذه الحالات الممكنة، مثلا إذا كان الجنس هو الصفة المدروسة، فإننا نقول عن الحالات الممكنة (ذكر ـــــ أنثى) و باعتبار أنها تتغير من فرد إحصائي إلى آخر أنها متغير إحصائي، أما إذا كانت الصفة المدروسة هي القامة فإننا نقول عن حالاته الممكنة ولنفس الاعتبارات أنها متغير من نوع ثان.

ولتسهيل استعمال هذه المتغيرات نوظف حرف من الحروف اللاتينية كرمز للمتغير مثل: x ، y، و إذا أردنا تحليل (P) متغير في آن واحد حيث (P) عدد صحيح موجب كأن نهتم بدراسة ثلاث صفات لدى طلبة السنة الاولى بالمدرسة العليا للتجارة في الموسم الدراسي 2016/2015 هي: الجنس و المعدل في البكلوريا والمستوى التعليمي للأب فإن في هذه الحالة (P=3) و(الجنس = x) و (المعدل في البكلوريا = y) و(المستوى التعليمي للأب = z). ونميز بين نوعين من المتغيرات الإحصائية المتغيرات الكيفية والمتغيرات الكمية.

1-8 المتغيرات الكيفية Les variables qualitatives

تتعلق بصفات لأفراد المجتمع الإحصائي غير قابلة للقياس الكمي فمثلا عند دراسة المجتمع الإحصائي لطلبة السنة الأولى في المدرسة العليا للتجارة في السنة الدراسية 2009/2008 فإن أفراد هذا المجتمع يمكن تصنيفهم حسب الجنس أو

حسب الإقامة أو حسب التقدير في البكلوريا أو حسب المستوى التعليمي للأب... إلخ. نلاحظ مما سبق أن هناك متغيرات كيفية مثل التقدير في البكلوريا حالاتها الممكنة لها أفضلية بينها الأمر الذي يسمح بترتيبها حسب هذه الأفضلية وهي بذلك على عكس الجنس ليس بين حالاته الممكنة أي أفضلية، لذا نميز بين نوعين من المتغيرات الكيفية

أ- **المتغيرات الكيفية الاسمية** Les variables qualitatives nominales

هي حالات لصفة من صفات الفرد الإحصائي ليس بينها أفضلية، بالتالي فهي لا تمكننا من إجراء أي ترتيب حسب الأفضلية لأفراد المجتمع الإحصائي أو العينة وفقها. إذا كان الجنس هو الصفة المدروسة و حالاته الممكنة كما هو معلوم هي: ذكر ــ أنثى فإن المتغير المدروس هو متغير كيفي وباعتبار غياب أي أفضلية بين الجنسين "كل ميسر لما خلق له" فيسمى هذا النوع من المتغيرات بالمتغيرات الكيفية الاسمية، ولتسهيل التعامل الآلي مع البيانات الإحصائية من هذا النوع نلجأ إلى ترقيمها بصفة اصطلاحية حيث نرقم الحالات الممكنة لصفة الجنس مثلا بالطريقة التالية: (ذكر= 1 أنثى= 2)

<u>مثال</u> :إذا كانت الصفة المدروسة هي الحالة العائلية وحالاتها الممكنة في المجتمع المسلم هي: أعزب ـ متزوج ـ مطلق ـ أرمل، فإن نوع المتغير في هذه الحالة هو متغير كيفي اسمي، ونرقمها اصطلاحا بالأرقام التالية على الترتيب:1، 2، 3، 4، دون أفضلية لحالة على أخرى.، لكن من الممكن أن نلجأ في بعض الاستبيانات إلى استعمال الطريقة الثنائية مثلا متزوج غير متزوج.

من جانب آخر قد نصادف أن بعض أفراد المجتمع الإحصائي أو أفراد العينة لا يجدون الحالة المناسبة لوضعيتهم في الاستبيان، لأن وضعيتهم غير مذكورة في الاستبيان أو أنهم لا يريدون الكشف عنها لحساسيتها، الأمر الذي يقتضي عند وضع الاستبيان إضافة حالات ممكنة خاصة توقعا لمثل هذه الوضعيات مثل: إجابة أخرى أو رفض الإجابة أو غير ذلك (أسئلة مفتوحة)

ب- **المتغيرات الكيفية الترتيبية** Les variables qualitatives ordinales

هي حالات لصفة من صفات الفرد الإحصائي بينها أفضلية الأمر الذي يسمح على ضوء ذلك بإجراء ترتيب (تصاعدي عادة) حسب الأفضلية لأفراد المجتمع الإحصائي أو العينة، فالمستوى التعليمي كصفة مدروسة لعينة من العمال حالاته

الممكنة هي: ابتدائي ـــ متوسط ـــ ثانوي ـــ جامعي، تسمح بتصنيف أفراد هذه العينة حسب المستوى التعليمي من الأدنى إلى الأعلى.

8 -2- المتغيرات الكمية Les variables quantitatives

نقول عن متغير إحصـــائي انه متغير كمي عندما تكون جميع الحالات الممكنة للصـــفة المدروســـة قابلة للقياس الكمي وبالتالي فإن جميع الصـــفات التي تقاس بأرقام ســـواء كانت طبيعية أم حقيقية مثل عدد الطلبة في الفوج أو الوقت المســـتغرق للإجابة على سؤال ما ،هي متغيرات كمية، وهي على نوعين.

أ- المتغيرات الكمية المنفصلة (المنقطعة) Les variables quantitatives discontinues

وهي المتغيرات الكمية التي تأخذ أرقاما معزولة وبالتعبير الرياضـــي هي المتغيرات التي تأخذ قيمها في مجموعة الأعداد الطبيعية $\{0,1,2,3.....\} = \mathbb{N}$ وواضـح أن عدد الأطفال في الأسرة أو عدد الغرف في المنزل أو عدد العمال في المؤسسة هي أعداد تنتمي إلى مجموعة الأعداد الطبيعية

ب- المتغيرات الكمية المتصلة Les variables quantitatives continues

وهي المتغيرات الكمية التي تأخذ، بالتعبير الرياضـــي، قيمها في مجموعة الأعداد الحقيقية، وبالتالي فإن الأرقام المعبرة عن الحقائق الاقتصـــادية والفيزيائية مثل الأســـعار والأجور والتكاليف والصـــادرات والكتلة النقدية...إلخ والأعمار و الأوزان والأطوال والأحجام ...إلخ هي متغيرات كمية متصـــلة باعتبار أنها تأخذ قيم أجزاء وحدة القياس فطول الفرد يمكن التعبير عنه بالمتر وأجزاء الوحدة حسب دقة القياس المعتمدة.

9 - سلم قياس المتغيرات الإحصائية L’ Echelle de mesure

يسمح سلم القياس بتعريف طريقة قياس متغير ما، أي كيف نرفق بكل فرد إحصائي قيمة من قيم المتغير الإحصائي x ونميز عادة بين أربع أنواع من سلم قياس المتغيرات الإحصائية .

1-9 السلم الاسمي L’échelle nominale

يسمح هذا السلم بتصنيف أفراد المجتمع الإحصائي حسب فئات اسمية دون أفضلية لإحداها على الأخرى ويتعلق بالمتغيرات الكيفية الاسمية مثلا متغير الجنس يسمح بتوزيع وحدات المجتمع الإحصائي أو العينة على فئتين: ذكور إناث وبالطريقة الموجزة التالية.

وواضح أن هذا السلم يوافق المتغيرات الكيفية الاسمية

الجنس x_i	العدد n_i
ذكر	n_1
أنثى	n_2
المجموع	N

2-9 السلم الترتيبي L’échelle ordinale

يسمح هذا السلم بترتيب الوحدات الإحصائية تنازليا أو تصاعديا لكن دون تحديد بدقة للفروق بين القيم الكيفية الممكنة للصفة المدروسة وبالتالي فإن هذا السلم يوافق المتغيرات الكيفية الترتيبية.

<u>مثال</u> : قدم كل تلميذ من مدرسة ما عملا أدبيا وقيمه المعلم بإحدى الملاحظات التالية: ضعيف - متوسط - جيد تسمح لنا هذه الملاحظات وبتوظيف هذا السلم بعرض نتائج التلاميذ بالطريقة الموجزة التالية:

عدد التلاميذ n_i	الملاحظة x_i
n_1	ضعيف
n_2	متوسط
n_3	جيد
N	المجموع

9- 3 السلم ألمجالي L’échelle d’intervalle

يسمح هذا السلم بقياس الخواص الكمية و تحديد الفروق بين قيمها بدقة، ويتميز بوجود الصفر الذي لا يعني عدم توفر الصفة المدروسة مثلا في سلم قياس درجة الحرارة الصفر يعبر عن خيار مرتبط بحقيقة فيزيائية هي المرور من الحالة الصلبة إلى الحالة السائلة ولا يعبر عن غياب الظاهرة كما يناسب هذا السلم مع اختبارات الذكاء، ويجب الإقرار بأن استعمال هذا المجال خارج الأمثلة المذكورة سابقا محدود جدا .

9-4 السلم النسبي L’échelle de rapport

يسمح بقياس المتغيرات الكمية و يماثل السلم السابق لكن الصفر يمثل غياب الظاهرة المدروسة مثل المتغيرات الزمنية والمسافة و الأسعار والأوزان ...الخ، ويمكن تلخيص ما سبق في الجدول التالي

سلم القياس	**نوع المتغيرات**
اسمي أو ترتيبي	المتغيرات الكيفية
مجالي أو نسبي	المتغيرات الكمية

10- مصادر المعطيات الإحصائية

نميز بين نوعين من المعطيات الإحصائية من حيث مصادرها، المعطيات الإحصائية العمومية والمعطيات الإحصائية الخاصة :

الأولى يتم إعدادها من طرف مختلف الهيئات العمومية كالديوان الوطني للإحصائيات ومختلف الوزارات والهيئات العمومية كالبنك المركزي والجمارك...الخ والمنظمات الدولية كالأمم المتحدة وصندوق النقد الدولي والبنك الدولي.

أما الإحصائيات الخاصة فمصدرها هو المعاهد المتخصصة والجمعيات المهنية والجمعيات الخيرية وأعمال البحث كالدراسات والبحوث والمذكرات...الخ.

يشكل الديوان الوطني للإحصائيات (office national des statistiques) محور نظام المعلومات الإحصائية في الجزائر وهو مكلف بجمع الإحصائيات الإدارية وانجاز الإحصاء العام للسكان والسكن وإجراء التحقيقات الإحصائية حول استهلاك الأسر واليد العاملة ومتابعة تطور الإنتاج الصناعي بحساب الرقم القياسي للإنتاج الصناعي والرقم القياسي لأسعار الإنتاج الصناعي ...الخ.

<u>ملاحظة:</u> عند توظيف المعطيات الإحصائية في موضوع ما كالبحث أو المذكرة أو الدراسة يجب ذكر مصدرها وذلك بغرض السماح بتأكد منها أو تفصيلها.

<u>مثال</u> : أجريت دراسة عن التسرب الدراسي في ولاية تيبازة في شهر جويلية 2017 وتم جمع معطيات تتعلق بجوانب شخصية وتربوية واجتماعية واقتصادية وثقافية مختلفة للتلاميذ المتسربين من المدرسة هي : سن التلميذ ألمتسرب جنس التلميذ المتسرب، معدله الدراسي، مهنة الأب، المستوى التعليمي للأب، المستوى التعليمي للأم ، نوع السكن (جماعي، فردي) منطقة السكن (حضرية، ريفية) ، عدد أفراد الأسرة، عدد الغرف في البيت، الدخل الشهري للأسرة، الارتباط بشبكة المياه الصالحة للشرب، الارتباط بشبكة الصرف الصحي، الارتباط بشبكة الكهرباء، بعد المنزل عن المدرسة .

1. ما هو المجتمع الإحصائي المدروس؟
2. ما هي الصفات المدروسة؟
3. ما هي أنواع المتغيرات المدروسة ؟
4. ما هو سلم القياس المناسب؟

<u>الإجابة</u>:

المجتمع الإحصائي المدروس هو التلاميذ المتسربين من المدرسة في شهر جويلية من سنة 2012 في ولاية تيبازة.

الصفات المدروسة	نوع المتغير	سلم القياس المناسب
منطقة السكن	كيفي	اسمي
مهنة الأب	كيفي	ترتيبي
المستوى التعليمي للأب	كيفي	ترتيبي
عدد أفراد الأسرة	كمي منفصل	نسبي
الدخل الشهري للأسرة	كمي متصل	نسبي
الارتباطبشبكةالمياهالصالحةللشرب	كيفي	اسمي

الفصل الثاني:
تجميع وتنظيم وعرض البيانات الإحصائية

الباحث يقوم بالتقاط المعلومات المتحصل عليها، ويضعها في جداول أولية، تسمى بالسلسلة الخامة، وهي غير مرتّبة، وذلك بعد فرز الاستمارات ثم يجهزها لاستعمالات نهائية بعد تلخيصها في جداول نهائية، تسمى جداول المعطيات، وهي جداول مرتّبة.

السلسلة الخامة:

هي سلسلة تحتوي على معطيات أخذت كما لوحظت.

إذا كانت تحتوي القيم $x_i = \{x_1 ; x_2 ;\cdots; x_n\}$ المرقمة من 1 إلى n والخاصة بـ n فرد من المجتمع الإحصائي، فمتتالية القيم $x_1 ; x_2 ;\cdots; x_n$ تسمّى بالسلسلة الخامة.

1. عرض البيانات الإحصائية أو توزيع المشاهدات:

هناك مجموعة من الطرق لعرض البيانات الإحصائية، والتي تتلخص بطريقتين هامتين:

- طريقة العرض الجدولي.
- طريقة العرض البياني.

1.1. طريقة العرض الجدولي:

وهي أبسط طريقة لعرض المشاهدات الإحصائية، وذلك حسب الصفة المدروسة.

أ – إذا كانت الصفة المدروسة كيفية:

باستخدام العرض الجدولي، نميّز بين مختلف مسميات الصفة (Les modalités du caractères)، وإذا كان لدينا مجتمع إحصائي مكوّن من n، وحدة إحصائية بالصفة المميزة، والتي تتكون من k مسمية (k modalités)، وهي مختلف الحالات التي يمكن للصفة أن تأخذها في الفرد أو المجتمع، والمعبّر عنها بـ x_i، حيث:

$$\overline{i = 1,k} \,/\, x_i = \{x_1 ; x_2 ;\cdots; x_k\}$$

نحدد عدد الأفراد الذي لهم المواصفات المتشابهة، فنسمي هذا العدد بالتكرارات أو التكرارات المطلقة ونرمز لها بـ:

$$\sum n_i = N \qquad n_i = \{n_1 ; n_2 ;\cdots; n_k\}$$

ويمكن الحصول على التكرارات النسبية لكل صفة بمواصفاتها أو مسمياتها.

أ.1 – التكرارات المطلقة (n_i Fréquences absolues):

n_i : عبارة عن عدد أفراد المجتمع الذين لهم الصفة x_i.

k: هي عدد القيم أو المسمّيات التي يمكن للصفة x_i أن تأخذها.

N : عدد أفراد المجتمع المدروس، حيث: $N = n_1 + n_2 + \cdots + n_k = \sum n_i$

فيظهر جدول المعطيات كالآتي:

x_i	n_i	f_i
x_1	n_1	$\frac{n_1}{N}$
x_2	n_2	$\frac{n_2}{N}$
$\vdots$	$\vdots$	$\vdots$
x_k	n_k	$\frac{n_k}{N}$
$\sum_{i=1}^{k} n_i$	N	1

أ.2 – التكرارات النسبية (f_i Fréquences relatives):

عبارة عن نسبة الأفراد الذين لديهم x_i ، بالنسبة للمجتمع المدروس وتحسب كالآتي:

$$f_i = \frac{n_i}{N} \quad \text{و} \quad \sum_{i=1}^{k} f_i = 1$$

$$= \sum_{i=1}^{k} \frac{n_i}{1} = \frac{n_1}{N} + \frac{n_2}{N} + \cdots + \frac{n_k}{N}$$

ويمكن تقدير التكرار النسبي بالمائة حيث المجموع يعطينا 100.

مثال: توزيع 50 شخصا حسب الجنس كالآتي:

الجدول (1): توزيع 50 شخصا حسب الجنس.

التكرارات النسبية f_i **(%)**	**التكرارات المطلقة** n_i	**الجنس** x_i
60	30	رجل
40	20	امرأة
100	50	المجموع

المصدر: غير معروف

ب – الصفة المدروسة كمية:

نميز بين الصفة المدروسة الكمية المتقطعة (المنفصلة أو المنقطعة)، والصفة المدروسة الكمية المستمرة أو المتصلة.

ب.1 – الكمية المتقطعة أو التوزيعات التكرارية المطلقة:

وهي عبارة عن جداول، نأخذ بعين الاعتبار كل متغير بتكراره.

مثال: ليكن توزيع 100 عامل في مؤسسة ما حسب مناصب الشغل كما هو مبين في الجدول الآتي:

الجدول (2): توزيع 100 عامل في مؤسسة حسب مناصب الشغل.

التكرار النسبي f_i **(%)**	**التكرار المطلق** n_i	**مناصب الشغل** x_i

مهندس	5	$f_1 = 5\% = 0,05$
تقني سامي	10	$f_2 = 10\% = 0,1$
عامل	15	$f_3 = 15\% = 0,15$
عامل بسيط	70	$f_4 = 70\% = 0,7$
المجموع	100	$100\% = 1$

المصدر: من ملفات المؤسسة 'م'

ملاحظة:

التكرارات أو التوزيعات التكرارية خاصة بالمتغيرات الكيفية أو الكمية المنقطعة أو المنفصلة.

مثال: نفرض أن مجموعة من 60 طالبا أجرت فحصا لقياس حدة البصر، وكانت النتائج كالآتي:

8	4	3	6	8	5	9	7	6	5	8	7

ترتّب هذه المعلومات وتوضع في جدول يوضح التوزيع التكراري لعلامات الطلبة في فحص النظر.

الجدول (3): التوزيع التكراري لعلامات 60 طالبا في فحص النظر.

x_i	n_i	f_i (%)
2	1	1,67
3	2	3,33
4	4	8,33
5	8	13,33
6	12	20
7	15	25
8	9	15

9	6	10
10	2	10
$\sum$	60	100%

المصدر: ملفات الجامعة 'ع'

ب.2 – الكمية مستمرة أو حالة التوزيعات المبوّبة في فئات:

إذا كان مدى المشاهدات صغيرا، فيمكن إعطاء التوزيع، كما في المثال أعلاه، أي الجدول التكراري المطلق.
أما إذا كان المدى كبيرا أو عدد المعطيات كبير، فنقسّم قيم المعطيات إلى فئات يتراوح عددها بين 5 و 15 حسب المعطيات.
قد تكون هذه الفئات غير متساوية الطول، حيث يراعى فيها تراكم الأفراد قصد تعديلها بواسطة تعديل التكرارات، وقد تكون هذه الفئات متساوية الطول، وبالتالي يجب معرفة كيف يتم تشكيل المعطيات.

- **تشكيل المعطيات في جداول ذات فئات متساوية الطول:**

- ***تحديد طول الفئة h***:

تحدّد عدد الفئات حسب حجم العيّنة وحسب توزيع الوحدات الإحصائية قيد الدراسة.
قد يختار هذا العدد أو يحدّد من طرف الباحث أو من خلال الطريقة الرياضية المعروفة بطريقة ***"STURGES"*** .

إذا كان طول الفئة يحدّد بـ $h > \frac{E}{J}$ ، حيث:

h: طول الفئة
E: مدى السلسلة
J: عدد الفئات

إذن: عدد الفئات J يحسب كالآتي في حالة ما إذا كان:

- **غير معروف**:

$$J = 1 + 3,32 \log(N)$$

أو

$$J = 1 + 1,32 \ln(N)$$ حيث:

Ln: لوغاريتم نيبري
Log: لوغاريتم عشري (أكثر استعمالا)
نشكل فئات ذات مجالات نصف مفتوحة (مثال $]3,1[$)
نحدد E
حيث: $E = x_{\max} - x_{\min}$
E : المدى العام للسلسلة
$x_{\max}$: القيمة العظمى من السلسلة
$x_{\min}$: القيمة الدنيا من السلسلة

- **معلومًا**:
فمراحل تكوين الفئات يكون كالآتي:

1 – حساب h طول الفئة
2 – مقارنة عدد القيم الأصلية $E + d$ وعدد القيم المجمّعة $h \times J$.

ونحسب الفرق Δ حيث:

$$\Delta = (h * j) - (E+d)$$

d : دقة القياس

3 – نشكل الفئات النظرية من الشكل $\left[x^{-} ; x^{+}\right]$.
4 – نشكّل الفئات الحقيقية ويستلزم تحديد الحد الأعلى للفئة الأخيرة، والحد الأدنى للفئة الأولى، وهذا مرتبط بقيمة Δ:

●● إذا كان Δ عدد زوجي:
- الحد الأدنى: $x_{\min} = x^{-} - \dfrac{\Delta}{2}$
- الحد الأعلى: $x_{\max} = x^{+} + \dfrac{\Delta}{2}$
●● إذا كان Δ عدد فردي:

- الحد الأدنى: $x_{\min} = x^{-} - \frac{\Delta + 1}{2}$

- الحد الأعلى: $x_{\max} = x^{+} + \frac{\Delta - 1}{2}$

5 – حساب التكرارات بحيث نراعي فيها دقّة القياس، حيث نحوّل الفئات المغلقة إلى فئات نصف مفتوحة بحدّيها x^{-} و x^{+}.

حيث x^{-} يكتب: $x^{-} = x^{-} - \frac{d}{2}$ (d : دقة القياس)

و x^{+} يكتب: $x^{+} = x^{+} + \frac{d}{2}$

يمكن تحديد d بـ :

- إذا كانت أعداد مطلقة فـ $d = 1$.
- إذا كانت أعداد نسبية عشرية فـ $d = 0,1$.
- إذا كانت أعداد نسبية مئوية فـ $d = 0,01$.

7 – حساب مراكز الفئات ويحسب كالآتي:

$$x_i = \frac{x_i^{-} + x_i^{+}}{2}$$

i : الفئة ، $x_i^{+} + x_i^{-}$ مجموع الحدّين ، x_i: مركز الفئة

فتصبح مراكز الفئات هي تسمية أو مسميات المتغير (Les modalités du caractère).

❖ **دالة التوزيع التراكمي:**

لكل توزيع، يمكن حساب دالة التوزيع التراكمي الصاعدة ودالة التوزيع التراكمي النازلة.

❖ **دالة التوزيع التراكمي الصاعدة:**

ونرمز لها بـ $N(x)$ وتأخذ الشكل :

$$N(x = x_p) = \sum_{i=1}^{p} x_i$$

ويمكن حسابها من الجدول من خلال التكرارات حيث:

$$N(x)\begin{cases} 0 \longrightarrow x > x_1 \\ n_1 \longrightarrow x \leq x_1 \\ n_1 + n_2 \longrightarrow x_1 \leq x \leq x_2 \\ n_1 + n_2 + n_3 \longrightarrow x_2 \leq x \leq x_3 \\ n_1 + n_2 + \cdots + n_p \longrightarrow x > x_p \end{cases}$$

هذه الدالة، يمكن وضعها في جدول التوزيع " $N^{\nearrow}$ " (N صاعد)، ويعني التكرارات المجمعة الصاعدة، بحيث:
$= \left(n_1 + n_2 + \cdots + n_p\right)$ هو مجموعة الأفراد الذين لديهم على الأكثر القيمة x_i.

لنعتبر الجدول أعلاه فنحصل على ما يلي:

$N^{\searrow}$	$N^{\nearrow}$	f_i	n_i	x_i
60	$n_1 = 1$		1	2
$60 - 1 = 59$	$n_2 = 1 + 2 = 3$		2	3
57	$n_3 = 3 + 5 = 8$		5	4
52	$n_4 = 8 + 8 = 16$		8	5
44	$n_5 = 16 + 12 = 28$		12	6

7	15		$n_6 = 43$	32
8	9		$n_7 = 52$	17
9	6		$n_8 = 58$	8
10	2		$n_9 = \boxed{60}$	2
Σ	$\boxed{60}$	1		

. طريقة العرض البياني:

هي الطريقة المستعملة لتسهيل قراءة الجدول ولإعطاء التفسير اللازم، كما يوضح لنا المتغير المدروس مهما كان نوعه.

أ – المتغير المدروس كيفي:

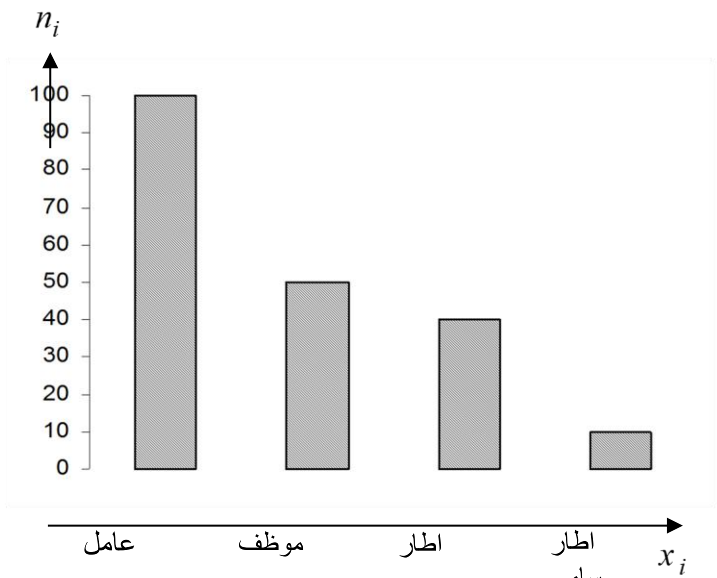
هناك طريقتين لتمثيل المتغير الكيفي:

أ.1 – طريقة المستطيلات: (Les graphiques en tuyaux d’orgues)

عندما نذهب لتمثيل السلاسل التكرارية، فيجب أن نمثل قيمة x_i للصفة المدروسة على محور الفواصل والتكرارات n_i على محور التراتيب.

يشكل هذا النوع من الرسم البياني من عدة مستطيلات مصفوفة، بحيث تكون المسافة التي تفصلها عن بعضها البعض ثابتة، ولها نفس العرض وتوضع قواعدها على نفس المحور الأفقي، أما أطوال هذه المستطيلات فتتناسب طردا مع عدد التكرارات.

مثال: إذا كان لدينا مؤسسة تشغل 100 عامل و 50 مستخدم، 40 إطار، 10 إطارات سامية، فيكون التمثيل البياني كالآتي:



ويمكن أن نمثل البيان في مستطيل واحد، ويمكن أن يقسم إلى أجزاء، ويسمى **القضبان البياني**.

وتستعمل هذه الطريقة عندما يراد إظهار القيم النسبية لمختلف أقسام هذه الظاهرة مثلا تجزئة عدد العمال إلى ذكور وإناث، يمثل المستطيل العدد الكلي للعمال.

أ.2 – طريقة الدوائر أو القطاعات (Méthode circulaire ou par secteur):

غالبا ما تستعمل هذه الطريقة لتمثيل أجزاء من ظاهرة كلية، وتعرف هذه الطريقة على شكل دائرة مقسّمة إلى قطاعات متناسبة طردا مع مختلف القيم التي تتميز بها. حيث تسمح بتقدير النسب بين أحجام القطاعات بشكل دقيق.

<u>مثال:</u> تأخذ المثال السابق: ننجز جدولاً:

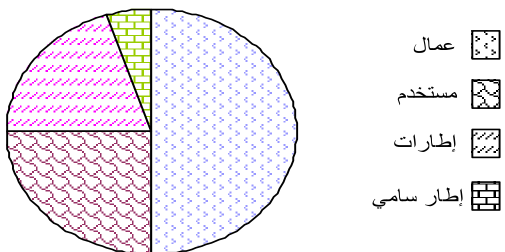
المشاهدة، المتغير x_i	التكرار n_i	$f_i = \frac{n_i}{N}$	$\alpha°$
عامل	100	0,50	180
مستخدم	50	0,25	90
إطارات	40	0,20	72
إطار سامي	10	0,05	18
N	200	1	360°

تحوّل قيم التكرارات إلى درجات، وذلك بتطبيق القاعدة على زاوية مركزية.

بحيث: $\begin{cases} 360° \to 200 \\ \alpha° \to 100 \end{cases}$

مثال: $\alpha = 180°$

التمثيل البياني بواسطة الدائرة أو القطاعات لتوزيع 200 عامل حسب الوظيفة



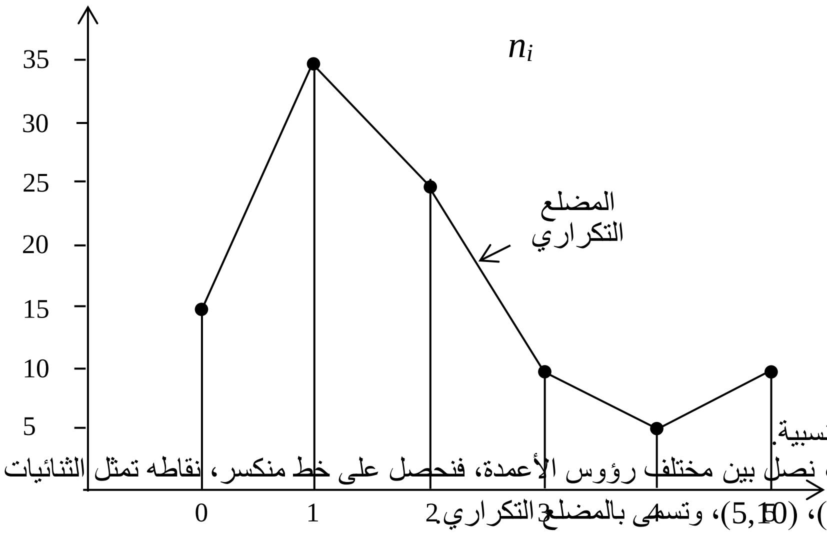
ب – المتغير المدروس كمّي:

ب.1 – المتغير الكمي المنفصل أو المتقطع:

● **التمثيل البياني بواسطة الأعمدة:** هناك طريقة بسيطة لتمثيل هذا النوع من التوزيعات، وهي أن نعين نقاط متباعدة بشكل منتظم على محور الفواصل (ox)، وتكون المسافة بين نقطتين متتاليتين هي وحدة القياس. نكتب هذه النقاط على (ox) والتي تمثل الصفة المدروسة x_i حسب ترتيبها الطبيعي.

ونختار وحدة قياس لمحور التراتيب، نضع فيه التكرارات، وبما أن لكل x_i تكرار n_i فنحصل على جملة من النقاط، ونسقط كل نقطة على محور الفواصل.

التمثيل البياني بواسطة الأعمدة لـ 100 أسرة موزعة حسب عدد أطفالها.



لنحصل على خطوط تسمى الأعمدة البيانية.

يمكن الحصول على نفس التمثيل البياني بالنسبة للقيم النسبية.

● **التمثيل البياني بواسطة المنحنى:** للحصول على المنحنى، نصل بين مختلف رؤوس الأعمدة، فنحصل على خط منكسر، نقاطه تمثل الثنائيات (x_i, n_i)، أي: (0,15)، (4,5)، (3,10)، (2,25)، (1,35)، (5,10)، وتسمى بالمضلع التكراري.

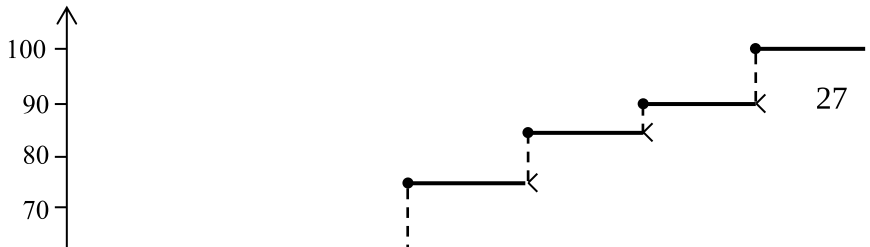
بالنسبة للمثال أعلاه، نُعطي دالة التوزيع الصاعدة:

$$N(x)\begin{cases} n_0 = 0 \ ; \ x \leq 0 \\ n_1 = 15 \ ; \ 0 \leq x < 1 \\ n_1 + n_2 = 50 \ ; \ 1 \leq x < 2 \\ n_1 + n_2 + n_3 = 75 \ ; \ 2 \leq x < 3 \\ n_1 + n_2 + n_3 + n_4 = 85 \ ; \ 3 \leq x < 4 \\ n_1 + n_2 + n_3 + n_4 + n_5 = 90 \ ; \ 4 \leq x < 5 \\ n_1 + n_2 + n_3 + n_4 + n_5 + n_6 = 100 \ ; \ 5 \leq x < 6 \end{cases}$$

<u>مثال:</u> ليكن لدينا سلسلة إحصائية بصفة منفصلة تمثل مجموعة 100 أسرة، حسب عدد أطفالها، كما هو مبين في الجدول التالي:

عدد الأطفال x_i	عدد الأسر n_i	التكرار النسبي $f_i = \frac{n_i}{N}$	N □	F □
0	15	0,15	15	0,15
1	35	0,35	50	0,50
2	25	0,25	75	0,75
3	10	0,1	85	0,85
4	5	0,05	90	0,9
5	10	0,1	100	0,1
Σ	100	1	/	/

التمثيل البياني لدالة التوزيع التراكمي للدالة الصاعدة (100 أسرة حسب عدد أطفالها).



X_i

يعطي هذا التمثيل البياني لدالة التوزيع التراكمي، بالنسبة للكمية المنفصلة والكيفية.

ب.2 – المتغير الكمي المستمر أو المتصل:

ويتم التمثيل البياني عن طريق المدرج التكراري أو المضلع التكراري أو المنحنى البياني.

- **المدرج التكراري (L'histogramme):** هناك حالتين:

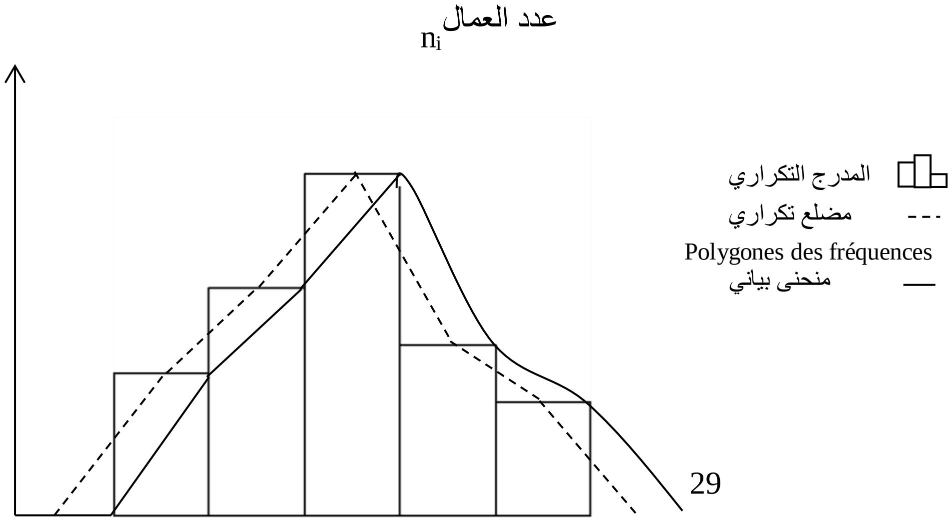
- حالة فئات متساوية الأطوال.
- حالة فئات غير متساوية الأطوال.

❖ في حالة فئات غير متساوية الأطوال، نعدّل الفئات بواسطة التعديل في التكرارات للحصول على فئات متساوية الأطوال، هذه الفئات تكون على شكل مجالات نصف مفتوحة $\left[x^{-};x^{+}\right[$ ، ونهاية الفئة هي بداية الفئة التي تليها. ويتم إنجاز المدرج التكراري بواسطة رسم مستطيلات متلاصقة، حيث قاعدة كل مستطيل تمثل طول الفئة والارتفاع، يتناسب طردا مع التكرارات.

مثال: ليكن لدينا توزيع مجموعة من 35 عامل حسب أجرهم الشهري.

الأجر الشهري 10^4 DA	عدد العمال	مركز الفئة	N ↗
[2;3[	5	2,5	5
[3;4[	8	3,5	13
[4;5[	12	4,5	25
[5;6[	6	5,5	31
[6;7[	4	6,5	35

في هذا المثال، الأطوال متساوية ويمكن إعطاء التمثيل البياني لهذا التوزيع كما يلي:



0 2 3 4 5 6 7 الأجر

x_i

- **المضلعات التكرارية (Polygones des fréquences):**

لنعد إلى المثال أعلاه، ونعيّن منتصف العرض الأعلى لكل مستطيل من المستطيلات المتلاصقة، فنحدّد بذلك منتصفي أو مراكز الفئات مع الأخذ بعين الاعتبار منتصفي الفئتين التي تسبق والتي تلي السلسلة، أي مراكز الفئة الأولى التي حدودها 1 و2، والأخيرة 7 و8. لأن تكرار هاتين هو 0.

نحصل على ثنائيات $(x_i ; n_i)$، ونصل بين مختلف هذه النقاط، ونحصل على خطوط منكسرة تعطي ما يسمّى بالمضلّع التكراري للسلسلة الإحصائية.

ويمكن الحصول على مراكز الفئات x_i، حيث: $x_i = \frac{\text{الاعلى الحد} + \text{الحدالادنى}}{2}$

ووضعها في جدول.

- **المنحنيات التكرارية والبيانية (Les courbes des fréquences):**

وهو التمثيل الأكثر شيوعا في الإحصاء الرياضي، ويتم رسمه بوصل النهايات العليا للفئات، مع الأخذ بعين الاعتبار الفئتين ما بعد الأخيرة، وما قبل الأولى.

ونحصل على منحنيات على شكل ∪ أو / أو مقلوب ∪ (∩) ومقلوب / (\)، وهذا يتوقف على طبيعة المعطيات.

- **التمثيل البياني لدالة التوزيع التراكمي:**

نعطي صيغة الدالة التراكمية

$$N(x)\begin{cases} n=0 \ ; \ x<2 \\ n_1=5 \ ; \ 2\le x<3 \\ n_1+n_2=13 \ ; \ 3\le x<4 \\ n_1+n_2+n_3=25 \ ; \ 4\le x<5 \\ n_1+n_2+n_3+n_4=31 \ ; \ 5\le x<6 \\ n_1+n_2+n_3+n_4+n_5=35 \ ; \ 6\le x<7 \ \wedge \ x\ge 7 \end{cases}$$

التمثيل البياني لدالة التوزيع التراكمي الصاعدة، يكون من خلال إيصال النهايات العليا للفئات، بينما التمثيل البياني لدالة التوزيع التراكمي النازلة، يكون من خلال إيصال النهايات الدنيا للفئات، مع أخذ بعين الاعتبار الفئة قبل الأولى وبعد الأخيرة.

الفصل الثالث:

معالم المجتمع وتلخيص المعلومات

مقدمة:

بعد جمعنا للمعلومات وترتيبها ووضعها في جداول، وتمثيلها بيانيا بالنسبة لمختلف المتغيرات الكيفية والكمية بأنواعها، لا يمكن إعطاء التحليل الكامل الذي نريده للدراسة الإحصائية للظاهرة، ففي كثير من الأحيان، وعند وجود عدد كبير من الملاحظات والمشاهدات، يجد الباحث نفسه مضطرا لتلخيصها في قيم، سميت مقاييس، وهي أنواع، نجد منها:

- **مقاييس النزعة المركزية.**
- **مقاييس التشتت.**
- **مقاييس الشكل والتمركز**.

1. مقاييس النزعة المركزية:

هذه المقاييس تخضع إلى بعض الشروط، منها:

- تبنى على جميع الملاحظات أو المشاهدات.
- تكون سهلة التفسير والفهم.
- تكون سهلة الحساب حيث تخضع للعمليات الجبرية بسهولة.
- لا تتأثر القيم الحدية للقيمة.

وهذه المقاييس أنواع، نجد المتوسطات بأنواعها: المنوال والوسيط.
وسندرس من المتوسطات ما يلي:

1.1. المنوال (Le mode):

وهي القيمة الأكثر تكرارا في السلسلة الإحصائية، ونرمز لها بـ M_θ ، ويتم تحديده رياضيا وبيانيا.

أ – تحديد M_θ رياضيا:

ويختلف حسابه باختلاف نوع السلسلة.

- **حالة السلسلة الخامة:**

$$x_i = \{x_1, x_2, \dots, x_n\}$$

نرتب هذه السلسلة ترتيبا تصاعديا ثم نعطي منوالها.

مثال: $x_i = \{1, 2, 2, 2, 3, 3, 5\}$.

قيمة M_θ هي الأكثر تكرارا = 2.

- **حالة التوزيع التكراري المنفصل:**

M_θ هي قيمة x التي يقابلها أكبر تكرار:

$$M_\theta = x_i \longrightarrow \max(n_i)$$

$$M_\theta = x_2 / n_2 = 12$$

$$M_\theta = 10$$

مثال:

n_i	x_i
5	9
12	10
6	11
3	12
26	Σ

- **حالة التوزيعات التكرارية المجمعة في فئات:**

في حالة التوزيعات التكرارية ذات الفئات، يمكن أن نعطي التعاريف التالية:

- الفئة التي يقابلها أكبر تكرار تسمّى **الفئة المنوالية.**

ومركز الفئة المنوالية يسمّى **المنوال التقريبي.**

M_θ هي قيمة تقديرية للقيم الحقيقية، وتحسب من خلال العبارة التالية:

$$M_\theta = x_{M_\theta}^{-} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$$

Δ_1: تكرار الفئة المنوالية – تكرار الفئة ما قبلها.

Δ_2: تكرار الفئة المنوالية – تكرار الفئة ما بعدها.

h: طول الفئة المنوالية.

$x_{M_\theta}^{-}$: الحد الأدنى للفئة المنوالية.

M_θ هي قيمة تقديرية للقيم الحقيقية، وتحسب من خلال العبارة التالية:

$$M_\theta = x_{M_\theta}^{-} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$$

Δ_1: تكرار الفئة المنوالية – تكرار الفئة ما قبلها.

Δ_2: تكرار الفئة المنوالية – تكرار الفئة ما بعدها.

h: طول الفئة المنوالية.

$x_{M_\theta}^{-}$: الحد الأدنى للفئة المنوالية.

مثال: ليكن لدينا التوزيع التكراري الآتي:

الفئات	$[37,40[$	$[40,43[$	$[43,46[$	$[46,49[$	$\sum$
التكرارات	2	7	4	2	15

$$x^{-}_{M_\theta} = 40 \qquad \Delta_1 = 7 - 2 = 5$$

$$h = 3 \qquad \Delta_2 = 7 - 4 = 3$$

$$M_\theta = 40 + 3 \cdot \frac{5}{5+3}$$

$$\boxed{M_\theta = 41,875}$$

$$M_\theta = [40\ ,\ 43[$$

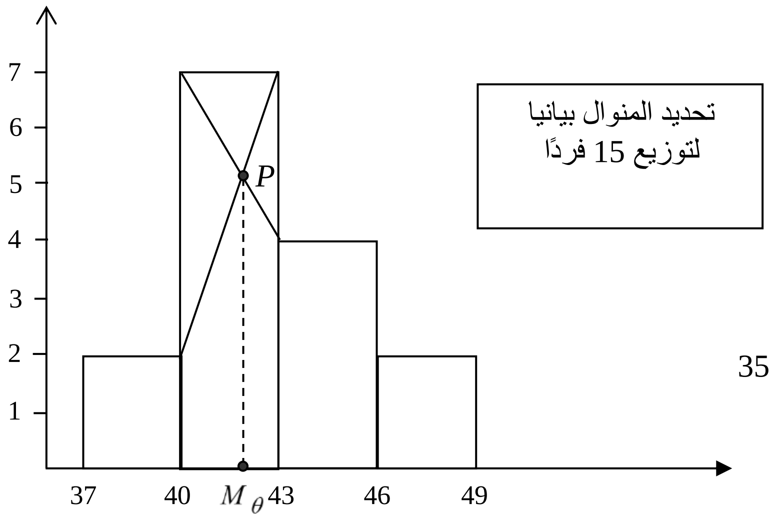
أ – تحديد M_θ بيانيا:

يتم تحديده على المدرج التكراري للتوزيع وهو عبارة عن نقطة التقاطع بين قطعتين مستقيمتين حيث:

القطعة الاولى تبدأ من النهاية العليا للفئة ما قبل الفئة المنوالية، وتنتهي عند الحد الأعلى للفئة المنوالية

والقطعة الثانية تبدأ من الحد الأدنى للفئة المنوالية وتنتهي عند الحد الأدنى للفئة ما بعد الفئة المنوالية.

باعتبار المدرّج التكراري للتوزيع أعلاه الخاص بالدراسة نحصل على ما يلي.



حيث نصل الحد الأعلى للفئة ما قبل الفئة المنوالية $]37,40]$ بالحد الأعلى للفئة المنوالية $]40,43]$.

ونصل بين الحد الأدنى للفئة المنوالية، بالحد الأدنى للفئـة مـا بعـد الفئة المنوالية $]43,46]$، نحصل على نقطة تقاطع P.

نسقط P على محور الفواصل فتعطينا M_θ .

<u>ملاحظة:</u>

1) يمكن أن يكون السلسلة أكثر من منوال، فقد تكون ثنائية المنوال أو متعددة المنوال.

2) يمكن أن تكون السلسلة ثنائية المنوال، ولكن لكل منوال تكرار مختلف، شرط أن يكونـان متقاربين.

3) لا يمكن حساب المنوال بالطريقة الجبرية أو البيانية، إلاّ بعد التأكد من أن كل الفئات متساوية الطول.

- **حالة عدم تساوي أطوال الفئات:**

في هذه الحالة يجب تعديل الفئات بتعديل التكرارات.

التكرارات المعدّلة نرمز لها بـ: $n_i^{'}$.

$$n_i^{'} = n_i \times \frac{h'}{h_i}$$

n_i : تكرار الفئة.

h' : أصغر طول فئة.

h_i: طول الفئة.

مثال: أحسب منوال هذا التوزيع:

الفئات	n_i	h_i	h'	$n_i' = n_i \times \frac{h'}{h_i}$
$[37,40[$	2	3		$n_1' = n_1 \times \frac{h'}{h_1} = 2 \times \frac{3}{3} = 2$
$[40,43[$	7	3	3	$n_2' = n_2 \times \frac{h'}{h_2} = 7 \times \frac{3}{3} = 7$
$[43,49[$	4	6		$n_3' = n_3 \times \frac{h'}{h_3} = 4 \times \frac{3}{6} = 2$
$[49,57[$	2	8		$n_4' = n_4 \times \frac{h'}{h_4} = 2 \times \frac{3}{8} = 0,75$
المجموع	15	/	/	/

الفئة المنوالية هي: $[40,43[$

$\Delta_1 = 7 - 2 = 5$ $\qquad x_{\bar{M}_\theta} = 40$

$\Delta_2 = 7 - 2 = 5$ $\qquad h = 3$

لحساب المنوال في هذه الحالة، نأخذ بعين الاعتبار التكرار المعدّل n_i'.

$$M_\theta = x_{\bar{M}_\theta} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$$

$$M_\theta = 40 + 3 \cdot \frac{5}{5+5} = 41,5$$

$$M_\theta = 41,5 \in [40,43[$$

ملاحظة هامة:

يتم تعديل الفئات في حالتين:

- في إعطاء الرسم البياني عندما تكون الأطوال غير متساوية.

- لحساب المنوال رياضيا وبيانيا.

. الوسيط (La médiane)**:**

يعطي الوسيط قيمة خاصة للمتغير x، حيث هذه القيمة تقسّم السلسلة من حيث العدد إلى قسمين متساويين، ويرمز له بـ M_e.

ويمكن القول أنه: يوجد $\frac{N}{2}$ من المعطيات $\leq M_e$.

ويوجد $\frac{N}{2}$ من المعطيات $> M_e$.

ويمكن تحديد رتبة M_e في الدالة التراكمية، أو من خلال المجمّع الصاعد، أو النازل حيث الرتبة هي:

$N(M_e) = \frac{N}{2}$ للقيم المطلقة في الدالة التراكمية.

و

$F(M_e) = \frac{1}{2} = 0{,}5 = 50\%$ للدالة التراكمية النسبية،

ويمكن حساب M_e جبريا وبيانيًا:

- **تحديد M_e جبريا:**

ـ حالة السلسلة الخامة:

ترتب السلسلة ترتيبا تصاعديا ثم تحدد قيمة M_e.
ويراعى فيها حالة n (فردي أو زوجي).

- **<u>لما يكون n فردي</u>:**

$$M_e = x_{\left(\frac{n+1}{2}\right)}$$

ما هو M_e في السلسلة التالية:

$$x_i = \{14, 7, 10, 5, 11, 8, 9\}$$

نرتب هذه السلسلة ترتيبا تصاعديا:

$$x_i = \{5, 7, 8, 9, 10, 11, 14\}$$

ثم نحسب M_e:

$$M_e = x_{\left(\frac{n+1}{2}\right)} = x_{\left(\frac{7+1}{2}\right)} = x_4$$

$$x_4 = 9$$

$$\boxed{M_e = 9}$$

توجد 50% من المعطيات ≤ 9 وتوجد 50% من المعطيات $> M_e$.

<u>لما يكون n زوجي</u>:

$$M_e = \frac{x_{\frac{N}{2}} + x_{\left(\frac{N}{2}+1\right)}}{2}$$

مثال:

$$x_i = \{6, 9, 3, 4, 9, 4\}$$

نرتبها:

$$x_i = \{3, 4, 4, 6, 9, 9\}$$

نحسب M_e:

$$M_e = \frac{x_3 + x_4}{2} = \frac{4+6}{2} = 5$$

$$M_e = 5$$

ونقول أن الوسيط يحتل المرتبة الثالثة والرابعة، وتحسب هذه الطريقة عندما تكون المعطيات صغيرة الحجم وبدون تكرارات.

- <u>حالة التوزيعات التكرارية:</u>

في هذه الحالة، نحسب N لتحديد رتبة الوسيط، حيث يرمز لهذه الرتبة بـ $N(M_e)$ و M_e هو أول قيمة لـ x يقابلها $\boxed{\frac{N}{2} \le N\nearrow}$.

مثال: ليكن توزيع 200 أسرة حسب عدد أطفالها.

- جدول يبين توزيع 200 أسرة حسب عدد أطفالها.

عدد الأطفال	n_i عدد الأسر	$N\nearrow$	$\frac{n_i}{N} = f_i$	$F\nearrow$
0	20	20	0,1	0,1
1	65	85	0,325	0,425
2	70	155	0,35	0,775

0,925	0,15	185	30	3
0,975	0,05	195	10	4
1	0,025	200	5	5
/	1	/	200	المجموع

نحسب الرتبة $N(M_e)$:

$$N(M_e) = \frac{N}{2} = \frac{200}{2} = 100$$

نبحث عن 100 في N↗ .

$$85 \leq 100 \leq 155$$

ونسقط القيمة 155 على x.

ونجد:

$$M_e = x_3 = 2$$

أو نحسب الرتبة $F(M_e)$:

$$F(M_e) = \frac{F}{2} = 0,5$$

$$0,425 \leq 0,5 \leq 0,775$$

نختار 0,775 نسقطها على x.

ونجد:

$$M_e = x_3 = 2$$

- <u>حالة التوزيعات المجمعة في فئات:</u>

نحسب قيمة الوسيط من خلال العلاقة التقديرية لـ M_e ، حيث:

$$M_e = x_{\bar{M}_e} + h_{M_e} \cdot \frac{\frac{N}{2} - N^{\uparrow}_{-1}}{n_{M_e}}$$

$x_{\bar{M}_e}$: الحد الأدنى للفئة الوسيطية.

h_{M_e} : طول الفئة الوسيطية.

$\frac{N}{2}$: رتبة الوسيط.

$N^{\uparrow}_{-1}$: التكرار المجمع الصاعد ما قبل الفئة الوسيطية.

n_{M_e} : تكرار الفئة الوسيطية.

وتحدد الفئة الوسيطية من $N \uparrow$ فنحدد الرتبة:

$$N_{M_e} = \frac{N}{2}$$

أو من خلال $F \uparrow$ فتحدد الرتبة:

$$F_{M_e} = \frac{F}{2}$$

مثال: ليكن لدينا توزيع 140 عامل حسب الأجور.

- جدول يبين توزيع 140 عامل حسب الأجور.

الأجور	**عدد العمال**	$N \nearrow$
$[30,40[$	26	26
$[40,50[$	33	59
$[50,60[$	64	123
$[60,70[$	7	130
$[70,80[$	10	140
المجموع	140	/

- تحديد رتبة الوسيط:

$$N\,(M_e) = \frac{N}{2} = \frac{140}{2} = 70$$

- نبحث على 70 في: $N\nearrow$:

$$59 \leq 70 \leq 123$$

- نختار 123 (لأنها القيمة الأكبر).
- نسقط قيمة 123 على x_i.
- نحدّد الفئة الوسيطية $[50,60[$.
- نطبق العلاقة:

$$M_e = 50 + 10 \cdot \frac{70-59}{64} = 50 + \frac{110}{65} = 51,71875$$

$$M_e \simeq 51,72$$

وتقرأ هذه النتيجة:

توجد 50% من العمال أجورهم $\geq 51,72$ وحدة نقدية.
توجد 50% من العمال أجورهم $\leq 51,72$ وحدة نقدية.

- **تحديد M_e بيانيًا:**

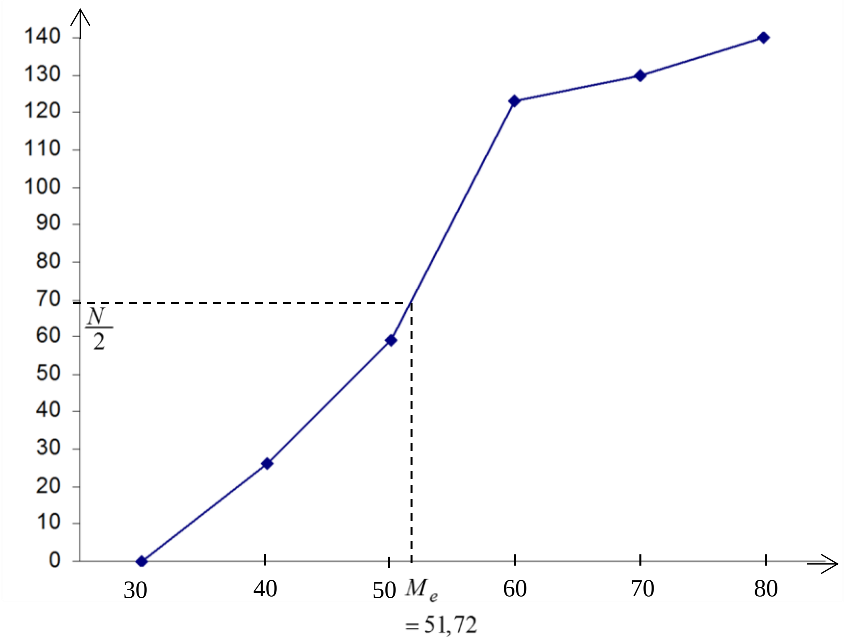
يحدّد الوسيط بيانيا من خلال $N\nearrow$ و $N\searrow$.
تقاطع المنحنيين وإسقاط نقطة التقاطع على محور الفواصل، يعطينا قيمة M_e التقريبية والتقديرية.

أو:

مباشرة من خلال منحنى N، حيث تؤخذ بعين الاعتبار قيمة $\frac{N}{2}$.

تحديد الوسيط بيانيا بالنسبة للمثال السابق.

تمثيل بياني يبين توزيع 140 عامل حسب الأجور:



● **تعميم فكرة الوسيط** (الرُّبيعيات، العشيريات، المؤينات)**:**

هي مقاييس تسمح بتقسيم السلسلة إلى أجزاء متساوية.

- الربيعيات (*Quartiles*):
تسمح بتقسيم السلسلة إلى 4 أجزاء متساوية، و هي الرُّبيع الأول، الثاني والثالث.

- **الرُّبيع الأول Q_1 :**

يرمز له بـ Q_1.

ويعطى من خلال العلاقة:

$$Q_1 = x_{Q_1}^{-} + h_{Q_1} \times \frac{\frac{1N}{4} - N_{-1}^{\square}}{n_{Q_1}}$$

وتقرأ هذه النتيجة:

توجد 25% من المعطيات $\geq Q_1$ وتوجد 75% من المعطيات $\leq Q_1$.

- **الرُّبيع الثاني Q_2:**

ما هو إلا علاقة الوسيط:

$$Q_2 = x_{Q_2}^{-} + h_{Q_2} \times \frac{\frac{2N}{4} - N_{-1}^{\uparrow}}{n_{Q_2}} = M_e$$

- **الرُّبيع الثالث $Q3$:**

ويعطى من خلال العلاقة:

$$Q_3 = x_{Q_3}^{-} + h_{Q_3} \times \frac{\frac{3N}{4} - N_{-1}^{\uparrow}}{n_{Q_3}}$$

وتقرأ هذه النتيجة:

توجد 25% من المعطيات $\leq Q_3$.

توجد 75% من المعطيات $\geq Q_3$.

ويمكن حساب هذه القيم بعد تحديد الرتبة $N(Q_1)$، $N(Q_2)$، $N(Q_3)$.

$$N(Q_1) = \frac{1}{4}N \quad , \quad N(Q_2) = \frac{2}{4}N \quad , \quad N(Q_3) = \frac{3}{4}N$$

وذلك لتحديد كل من $x_i^{-}\ ,\ h_i\ ,\ N_{-1}^{\uparrow}\ ,\ n_i$ ، من خلال تحديد الفئات المناسبة.

- **العشيريات (*Déciles*):**

تسمح بتقسيم السلسلة إلى 10 أجزاء متساوية، يحتوي كل جزء على 10% من المعطيات. ويمكن حساب "9 قيم" نسميها عشيريات، ونرمز لها بـ D_i $(i: 1 \rightarrow 9)$، حيث:

$$D_i = x_{d_i} + h_{d_i} \times \frac{\frac{i}{10} N - N_{-1}}{n_{d_i}}$$

وتحدد العشريات من خلال تحديد الرتبة، حيث:

$$N(D_i) = \frac{i}{10} N$$

- ***المؤينات (Percentiles)***:

و تكمن في تقسيم السلسلة إلى 100 جزء. كل جزء يحتوي على $\frac{1}{100}$ من المعطيات. ويمكن حساب 99 جزء نسميها المؤينات، ونرمز لها بـ ***P***

وحساب P_j يعطى من خلال العلاقة:

$$P_j = x_{\bar{P}_j} + h_{P_j} \times \frac{\frac{j}{100} \cdot N - N_{-1}}{n_{P_j}}$$

j= 1 ,100

ويمكن حساب المؤين بعد تحديد الرتبة:

$$N(P_j) = \frac{j}{100} \cdot N$$

ويمكن تحديد كل هذه القيم من خلال تطبيق المثال السابق:

- **الرُّبيع الأول Q_1:**

$$N(Q_1) = \frac{1}{4} \cdot N = 140 \cdot \frac{1}{4}$$

$$N\,(Q_1) = 35$$

نبحث عن 35 في N لتحديد الفئة:

$$26 < 35 < 59$$

نأخذ 59 تقابله الفئة $[40,50[$.

$$Q_1 = x_{Q_1}^{-} + h_{Q_1} \times \frac{\frac{N}{4} - N_{-1}}{n_{Q_1}}$$

$$Q_1 = 40 + 10 \times \frac{35-26}{33}$$

$$Q_1 = 42,72 \in [40,50[$$

توجد 25% من المعطيات $x \geq 42,72$ و توجد 75% من المعطيات $x \leq 42,72$.

- **الرُّبيع الثاني Q_2:**

$$N\,(Q_2) = \frac{N}{2} = \frac{140}{2}$$

$$N\,(Q_2) = 70$$

نبحث عن 70 في N لتحديد الفئة.

$$59 < 70 < 123$$

نأخذ 123 تقابله الفئة $[50,60[$.

$$Q_2 = x_{\bar{Q}_2} + h_{Q_2} \times \frac{\frac{N}{2} - N_{-1}^{\square}}{n_{Q_1}}$$

$$Q_2 = 50 + 10 \times \frac{70 - 59}{64}$$

$$Q_2 = 51,72 = M_e$$

- **الرُّبيع الثالث Q_3:**

$$N\,(Q_3) = \frac{3N}{4} = \frac{3 \times 140}{4}$$

$$N\,(Q_3) = 105$$

$$59 < 105 < 123$$

123 تقابله الفئة $[50,60[$.

$$Q_3 = x_{\bar{Q}_3} + h_{Q_3} \times \frac{\frac{3N}{4} - N_{-1}^{\square}}{n_{Q_3}}$$

$$Q_3 = 50 + 10 \times \frac{105 - 59}{64}$$

$$Q_3 = 57,19 \in [50,60[$$

ونقول: توجد 75% من القيم ≥ 57,19 وتوجد 25% من القيم ≤ 57,19.

- **العشيريات:**

مثلا حساب D_7:

$$N(D_7)=\frac{7}{10}\times 140=\frac{7}{10}\cdot N$$
$$N(D_7)=98$$

$$59<98<123$$

123 تقابله الفئة $[50,60[$.

$$D_7=x_{\bar{D}_7}+h_{D_7}\times\frac{\frac{7}{10}N-N_{-1}}{n_{D_7}}$$

$$D_7=50+10\times\frac{98-59}{64}$$

$$D_7=56,09\in[50,60[$$

توجد 70% من القيم $\geq 56,09$ وتوجد 30% من القيم $\leq 56,09$.

- **المؤينات:**

مثلا حساب P_{20}:

$$N(P_{20})=\frac{20}{100}\times N=\frac{20}{100}\times 140$$
$$N(P_{20})=28$$

$$26<28<59$$

59 تقابله الفئة $[40,50[$ و بتطبيق العلاقة

$$P_{20} = x_{\bar{P}_{20}} + h_{P_{20}} \times \frac{\frac{20}{100} \times N - N_{-1}}{n_{D_7}}$$

نحصل على :

$$P_{20} = 40 + 10 \times \frac{28 - 26}{33}$$

$$P_{20} = 40,61 \in [40,50[$$

توجد 20% من القيم $\leq 40,61$ وتوجد 80% من القيم $\leq 40,61$.

ملاحظة:

* $\%50 = P_{50} = D_5 = Q_2$

* $P_{25} = Q_1$ * $P_{75} = Q_3$

3.1. المتوسط الحسابي (*Moyenne arithmétique*):

أ – حساب المتوسط الحسابي باستعمال العلاقة المناسبة:

● حالة السلسلة الخامة:

نقول في هذا المجال:

المتوسط الحسابي في السلسلة الإحصائية x_i، $x_i = \{1, 2, \dots, n\}$ هو المقياس الأكثر شهرة لتعريف القيمة المركزية.

وهو عبارة عن جمع لكل المعطيات المقسوم على العدد n للسلسلة الإحصائية، ويرمز لها بـ $\bar{x}$. ويعطى من خلال العلاقة:

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

مثال: احسب الوسط الحسابي للعلامات الآتية:

$$\{18 , 15 , 08 , 09 , 14 , 16 , 12 , 13 , 07\}$$

بتطبيق العلاقة: $N = 9 , \sum x_i = 112$

$$\bar{x} = \frac{1}{9}\cdot\sum_{1}^{9} x_i = \frac{1}{9}\times 112$$

$$\bar{x} = 12,44$$

حالة التوزيع التكراري:

إذا كان لدينا توزيع تكراري $(x_i , n_i) \,/\, i = 1,\ldots,n$.

x_i: مشاهدات n_i: عدد التكرارات

نحصل على $\bar{x}$ من خلال هذه المعادلة:

$$N = \sum n_i \qquad \bar{x} = \frac{1}{N}\sum_{i=1}^{n} x_i\, n_i$$

يمكن أيضا حساب $\bar{x}$ بالقيم النسبية، حيث:

$$\bar{x} = \sum_{i=1}^{n} \frac{n_i \cdot x_i}{N} = \sum_{i=1}^{n} \frac{n_i}{N} \cdot x_i$$

$$\bar{x} = \sum_{i=1}^{n} f_i \cdot x_i$$

مثال: اوجد المتوسط الحسابي للقيم التالية:

7(3) , 13(2) , 12(2) , 16(1) , 14(1) , 9(1) , 8(3) , 15(1) , 18(1)

استعمال القانون لحساب $\bar{x}$:

$$\bar{x} = \frac{1}{N} \sum x_i n_i$$

$$\bar{x} = \frac{167}{15} \qquad \bar{x} = 11{,}13$$

$n_i x_i$	n_i	x_i
21	3	07
24	3	08
9	1	09
24	2	12
26	2	13
14	1	14
15	1	15
16	1	16
18	1	18
167	15	المجموع

حالة التوزيع ذو فئات:

لتطبيق العلاقة السابقة، ألا وهي: $\boxed{\overline{x} = \frac{1}{N}\sum x_i n_i}$.

نحسب مركز الفئة x_i.

مثال: ضع القيم السابقة في توزيع تكراري، ذو فئات متساوية عددها أربعة (4).

$$E = x_{\max} - x_{\min}$$

$$\begin{cases} E = 18 - 07 \Rightarrow E = 11 \\ j = 4 \end{cases}$$

$$h > \frac{E}{j}$$

$$h > \frac{11}{4} = 2,75 \qquad h = 3$$

$n_i \cdot x_i$	x_i	n_i	الفئات الحقيقية	الفئات النظرية
56	8	7	$[6,5\ ,\ 9,5[$	$[07,09]$
22	11	2	$[9,5\ ,\ 12,5[$	$[10,12]$
48	14	4	$[12,5\ ,\ 15,5[$	$[13,15]$
34	17	2	$[15,5\ ,\ 18,5[$	$[16,18]$
160	50	15	المجموع	

$d = 1$

$x^- = x^- - \frac{d}{2}$

$x^+ = x^+ + \frac{d}{2}$

خصائص المتوسط الحسابي:

إذا تعلق متغير بمتغير آخر، فبإمكان إيجاد المتوسط الحسابي لمتغير بدلالة متغير آخر.

مثال: إذا كان: $x_i = k \cdot V_i$ $V_i = \frac{x_i}{k}$

$\overline{x} = k \cdot \overline{V}$ $x_i = bU_i + a$ $U_i = \frac{x_i - a}{b}$

$$\overline{x} = b\overline{U} + a$$

ملاحظة: المتوسط الحسابي للقيمة الثابتة هو نفسه، حيث:

$$\overline{b} = b \ , \ \overline{a} = a \ , \ \overline{k} = k$$

ب – حساب المتوسط الحسابي باستعمال تغيير المتغير:

يمكن حساب المتوسط الحسابي $\overline{x}$ باستعمال تغيير المتغير، وذلك في حالة القيم الكبيرة جدا لتسهيل الحسابات. فنستعمل متغير جديد x_i'.

ويمكن حساب x_i' في حالتين:

- **حالة التوزيع التكراري المنفصل:**

نحسب x_i' من خلال العلاقة:

$$x_i' = x_i - x_0$$

حيث x_0 هو المتوسط الفرضي القريب من $\overline{x}$، أو M_e، أو M_θ.

حيث:

$$\overline{x'} = \frac{\sum x_i'}{N}$$

وفي حالة وجود تكرارات:

$$\overline{x'} = \frac{\sum n_i x_i'}{N}$$

بما أنّ $x_i' = x_i - x_0$، نعوض في العلاقة:

$$\overline{x'} = \frac{\sum n_i (x_i - x_0)}{N}$$

$$\overline{x'} = \frac{\sum n_i x_i - x_0 \sum n_i}{N}$$

$$\overline{x'} = \frac{\sum n_i x_i}{N} - \cancel{\frac{x_0 \sum n_i}{N}}$$

$$\overline{x'} = \overline{x} - x_0$$

$$\overline{x} = \overline{x'} + x_0$$

- **حالة التوزيع التكراري المبوّب في فئات:**

يستعمل x_i' و $x_i' = \frac{x_i - x_0}{h}$.

$(h \vee k \vee a$ تدل على طول الفئة). ومنه يكون المتوسط الحسابي للمتغير الجديد x_i':

$$\overline{x'} = \frac{\sum n_i x_i'}{N}$$

$$\overline{x'} = \frac{\sum n_i \left(\frac{x_i - x_0}{h}\right)}{N} = \frac{\sum n_i x_i - \sum n_i x_0}{Nh}$$

$$\overline{x'} = \frac{\sum n_i x_i}{Nh} - \frac{x_0 \sum n_i}{Nh}$$

$$\overline{x'} = \frac{1}{h}\overline{x} - \frac{1}{h}x_0$$

$$\overline{x'} = \frac{1}{h}(\overline{x} - x_0)$$

$$\overline{x} = h\overline{x'} + x_0$$

من المثال السابق

مركز الفئة	$x'_i = \frac{x_i - x_0}{h}$	n_i	$n_i x'_i$
8	-1	7	-7
11	0	2	0
14	1	4	4
17	2	2	4
المجموع		15	1

$\bar{x} = 11,2$

$x_0 = 11$

$h = 3$

$$\overline{x'} = \frac{\sum n_i x'_i}{N} = \frac{1}{15} = 0,0666$$

$$\boxed{\overline{x'} = 0,0666}$$

$$\bar{x} = h \cdot \overline{x'} + x_0$$

$$\bar{x} = 3(0,0666) + 11$$

$$\boxed{\bar{x} = 11,2}$$

نلاحظ أن $\bar{x}$ قيمتها لا تتغير.

ج ـ تعميم فكرة المتوسط:

يستعمل المتوسط الحسابي عادة للقيم المباشرة المقاسة على أفراد، ولكنه ليس بالمقياس الوحي. بل يوجد عدد كبير من المتوسطات تستعمل في بعض الحالات، نجد منها المتوسط الهندسي (***Moyenne géométrique***)، المتوسط التوافقي (***Moyenne harmonique***)، المتوسط التربيعي (**Moyenne quadratique**).

- **<u>المتوسط الهندسي</u>** (***La moyenne géométrique***)**:**

يستعمل المتوسط الهندسي في حالة المعطيات الزمنية لظاهرة ما، بغية معرفة وقياس طول هذه الظاهرة.

- **<u>حالة سلسلة عادية:</u>**

المتوسط الهندسي لـ n قيم موجبة هو الجذر النوني***Racine n^iemm^e***) لجدائها.
ليكن لدينا n قيمة موجبة $(x_1, x_2, \ldots, x_n)$، فمتوسطها الهندسي يحسب من خلال القانون التالي:

$$G = \sqrt[n]{x_1 \times x_2 \times \ldots \times x_n}$$
$$G = (x_1 \times x_2 \times \ldots \times x_n)$$

ويمكن أن يكتب أيضا:

$$G = \left[\prod_{i=1}^{n} x_i\right]^{\frac{1}{n}}$$

وتقرأ: جداء x_i من أجل i يتغير من $(n \leftarrow 1)$.
مثال: ليكن لدينا سلسلة مؤلفة من 6 أعداد:

$$\{1, 2, 5, 7, 10, 13\}$$

متوسطها الهندسي: $n = 6$

$$G = \sqrt[6]{1 \times 2 \times 5 \times 7 \times 10 \times 13} \approx 4,57$$

<u>ملاحظة:</u>

يتطلب حساب المتوسط الهندسي اللجوء إلى اللوغاريتمات العشرية.
ويظهر عندئذ أن $\log G$ هو المتوسط الحسابي للوغاريتمات العناصر التي تؤلف السلسلة.

وتكتب:

$$\boxed{\log_{10} G = \frac{\sum \log x_i}{N}}$$

تطبق على المثال:

$$\log G = \frac{\log 1 + \log 2 + \log 5 + \log 7 + \log 10 + \log 13}{6}$$

$$\log G = 0,6598$$

$$G = 4,569 \approx 4,57$$

- **<u>حالة التوزيع التكراري:</u>**

إلى جانب المتوسط الهندسي البسيط المحسوب أعلاه، نجد أيضا المتوسط الهندسي المرجح (بتكرارات) (Moyenne géométrique pondéré).

مثلا: ليكن لدينا مكتبا يستخدم 22 عاملا تتوزع أجورهم الشهرية كالآتي:

الأجر الشهري	الترجيح (عدد العمال)
100	10
120	05
125	04
140	03
المجموع	22

فإذا أخذنا المتوسط الهندسي المرجّح للتعبير عن الأجر المتوسط، فإن:

$$G_P = \sqrt[22]{100^{10} \cdot 120^5 \cdot 125^4 \cdot 140^3}$$
$$G_P = 113,64$$
$$G_P = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \cdot \cdots \cdot x_n^{n_k}}$$

وباستخدام $\log_{10}$ ، نحصل على:

$$\log_{10} G_p = \frac{10\log 100 + 5\log 120 + 4\log 125 + 3\log 140}{22}$$

$$\log_{10} G_P = \frac{45,22193}{22} = 2,0555$$

ومنه:

$$G_P = 113,64$$

القانون الذي يسمح بالحصول على G_P يكتب:

$$\log G_P = \frac{\sum n_i \log x_i}{N}$$

إذا كان المراد حساب نسبة التطور لسلسلة ما نستعمل في:

- **<u>حالة التوزيع البسيط</u> (بدون تكرار):**

$$G = \sqrt[n]{\prod_{i=1}^{n}(1+r_i) - 1}$$

- **حالة التوزيع التكراري:**

$$G = \sqrt[n]{\prod_{i=1}^{n}(1+r_i)^k} - 1$$

مثال:

1) معدل التطور للناتج المحلي الخام الذي نرمز له PIB في الجزائر، في المخطط الرباعي الأول هو كالتالي :

السنة	70	71	72	73
المعدّل	7,2	6,3	7,0	4,3

ما هو معدل النمو السنوي لـ PIB في الفترة، $n = 4$.

$$G = \sqrt[4]{(1+r_1)(1+r_2)(1+r_3)(1+r_4)} - 1$$
$$G = \sqrt[4]{(1+0,072)(1+0,063)(1+0,07)(1+0,043)} - 1$$
$$G = \sqrt[4]{1,072 \times 1,063 \times 1,07 \times 1,043} - 1$$
$$G = 1,0619 - 1 = 0,0619$$

$$G = 6,2\%$$

2) إذا ارتفعت أرباح المؤسسة في 3 سنوات الأولى بـ 5,8% في السنة الرابعة: 4,6%، في السنة الخامسة والسادسة: 11,2%.

ما هو معدل النمو السنوي للأرباح المحققة في هذه الفترات الثلاثة.
(الفترة الاولى فيها 3 سنوات، الفترة الثانية سنة واحدة والفترة الثالثة سنتين)

n=6

$$G=\sqrt[n]{(1+r_1)^{k_1}(1+r_2)^{k_2}(1+r_3)^{k_3}}$$
$$G=\sqrt[6]{(1,058)^3(1,046)^1(1,112)^2}-1$$
$$G=1,0736-1=0,0736$$

$$G=7,36\%$$

<u>مفهوم معدل النمو أو التطور:</u>

هو المعدل الذي يقيس الفرق في القيم بين الفترات، ويرمز له بـ i أو v أو r ، ويحسب من خلال العلاقة:

$$r=\frac{G_t-G_0}{G_0}=\frac{G_t}{G_0}-1$$
$$r=\frac{G_{t+1}-G_t}{G_t}=\frac{G_{t+1}}{G_t}-1$$

ملاحظة:

$G_t>G_0 \Leftarrow r>0 \Leftarrow$ زيادة.

$G_t<G_0 \Leftarrow r<0 \Leftarrow$ انخفاض.

$G_t=G_0 \Leftarrow r=0 \Leftarrow$ ثبات.

ومنه، نستنتج أنه في حالة:

- الزيادة $\Leftarrow$ $\frac{G_t}{G_0} = 1 + r$.
- الانخفاض $\Leftarrow$ $\frac{G_t}{G_0} = 1 - r$.
- الثبات $\Leftarrow$ $\frac{G_t}{G_0} = 0$.

ونسمي المضاعف $\Leftarrow$ $\frac{G_t}{G_0} = 1 - r$

إذا كانت القيمة G_0 الأصلية، تطورت بمعدل سنوي ثابت r خلال n سنة، فتصبح هذه القيمة في السنة G_0^n.

ونحسب:

$$G_t = G_0 (1 + r)^n$$

المناقشة:

إذا كانت القيمة لمعدلات غير ثابتة (متفاوتة خلال الفترة n). حيث: $n = \sum n_i$ ، فنتحصل على:

$$G_t = G_0 (1 + r_1)^{n_1} (1 + r_2)^{n_2} (1 + r_3)^{n_3} \cdots (1 + r_n)^{n_k}$$

مثال:

1) إذا كانت قيمة مبلغ قدره 4000 دج، قد وظف في البنك بمعدل سنوي ثابت r، حيث $r = 8\%$.
احسب هذه القيمة عند نهاية السنة السادسة.

$$G_t = G_0 (1 + r)^n$$
$$G_t = 4000 (1 + 0,08)^6$$
$$G_t = 4000 (1,08)^6$$

$$\boxed{G_t = 6347,49}$$

2) إذا كان هذا المبلغ 4000 دج وظّف لفترات، حيث في كل فترة معدل النمو يتغير بين 2%: سنتين ، 4%: 3 سنوات ، 6%: 5 سنوات.

احسب هذا المبلغ عند نهاية السنة 10.

$$G_t = G_0 (1 + r_1)^{n_1} (1 + r_2)^{n_2} \cdots (1 + r_k)^{n_k}$$
$$G_t = 4000 (1,02)^2 (1,04)^3 (1,06)^5$$
$$G_t = 6264,54 \, \text{DA}$$

3) إذا كانت القيم موجودة والمراد حساب نسبة التطور، فنحسب مثلا:
معدل التطور السنوي:

$$\frac{G_t}{G_0} = (1+r)^n$$

$$\Rightarrow \quad (1+r) = \sqrt[n]{\frac{G_t}{G_0}}$$

ومنه:

$$r = \sqrt[n]{\frac{G_t}{G_0}} - 1$$

إذا كانت المبيعات تقدم في سنة t بـ 20 و.ن ، بعد 7 سنوات أصبحت تقدر بـ 35 و. ن.
ما هو معدل النمو السنوي r؟

$$r = \sqrt[7]{\frac{35}{20}} - 1$$

$$r = 8,32\,\%$$

هذه العلاقة تمكّننا من إيجاد n, r, G_t, G_0.

$$\frac{G_t}{G_0} = (1+r)^n$$

مجالات تطبيق المتوسط الهندسي:

مقارنة بالمتوسط الحسابي، فمجالات تطبيق المتوسط الهندسي قليلة ومن أهمها:

- ❖ حساب الأرقام القياسية حيث يعد من أحسن المتوسطات.
- ❖ حساب المعدلات كمعدل الفائدة، معدل السكان،،،،،،،،،،

➢ **<u>المتوسط التوافقي</u> (*La moyenne harmonique*)**

المتوسط التوافقي المرموز له بـ h لـ n قيمة x_i، حيث: $x_i = \{x_1, x_2, \ldots, x_n\}$

هو العدد الذي يكون مقلوبه هو المتوسط الحسابي لمقلوبات هذه القيم.

إذا كان لدينا n قيمة، فمقلوبات هذه القيم هي:

$$\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}$$

ويستعمل هذا المتوسط في حالة المتغيرات الناتجة عن نسبة متغيرتين، مثل المردودية:

. السرعة = $\frac{\text{المسافة}}{\text{الزمن}}$. المردودية = $\frac{\text{الإنتاج}}{\text{عدد العمال}}$ نسبة التأطير = $\frac{\text{عدد الأطباء}}{\text{عدد السكان}}$

الأسعار/الوزن

- **حالة سلسلة خامة:**

نحسب المتوسط التوافقي البسيط

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

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

➢ حالة التوزيعات التكرارية:

نستعمل نفس الطريقة مع إدخال التكرارات المعتبرة، فيصبح H هو:

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

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

مثال: يقطع دراج مسافة بسرعة 60 كم في الساعة ذهابا و 30 كم إيابا، احسب السرعة المتوسطة.

ويكون حساب السرعة المتوسطة بحساب H البسيط.

$$x_1 = 60\,\text{km}\ ,\ x_2 = 30\,\text{km}\ ,\ N = 2\text{h}$$

$$H = \frac{2}{\frac{1}{30}+\frac{1}{60}} = \frac{2}{\frac{3}{60}} = \frac{120}{3} = 40$$

السرعة المتوسطة = 40 كم/سا

أما إذا كانت المسافات متغيرة، فنحسب باستخدام $H=\frac{N}{\sum\frac{n_i}{x_i}}$، حيث $N = \sum n_i$

في حالة مسافتين مختلفتين نحسبه بالتوزيع التكراري.

بصورة عامة:

المتوسط التوافقي يتأثر بالقيم الصغيرة في السلسلة، بينما يتأثر المتوسط الحسابي أكثر بالقيم الكبيرة.
المتوسط التوافقي يستعمل في حالة وجود علاقة بين ظاهرتين، أي في حالة وجود قيم متناسبة عكسا.

ملاحظة:

●يمكن استخراج العلاقة بين المتوسطات: $H\ ,\ G\ ,\ \overline{x}$، حيث تكون العلاقة كالآتي:

$$H < G < \overline{x}$$

● محاولات تطبيق المتوسط الهندسي G:
مقارنة بالمتوسط الحسابي، فمجالات تطبيق G قليلة، ومن أهمها:

- حساب الأرقام القياسية، حيث يعدّ من أحسن المتوسطات.
- حساب المعدلات، كمعدل الفائدة ومعدل النمو السكاني.

المتوسط التربيعي (Q) (*La moyenne quadratique*):

يعرّف المتوسط التربيعي بواسطة مربّعه، إذ أن مربعه يساوي المتوسط الحسابي لمربعات السلسلة، ويرمز له بـ Q.

● حالة السلسلة الخامة:

نحسب Q البسيط.

ليكن لدينا n قيمة لـ x_n / $x_i = \{x_1,\dots,x_n\}$، فيكون لدينا:

$$Q^2 = \frac{1}{N}\left(x_1^2 + x_2^2 + \cdots + x_n^2\right)$$

$$Q = \sqrt{\frac{1}{N}\sum_{i=1}^{n} x_i^2} = \left(\frac{1}{N}\sum x_i^2\right)^{\frac{1}{2}}$$

● **حالة التوزيع التكراري:**

نحسب Q المرجح في حالة التوزيعات التكرارية، فنحصل على:

$$Q = \left(\frac{1}{N}\sum_{i=1}^{n} n_i\, x_i^2\right)^{\frac{1}{2}}$$

ملاحظة:

إذا أردنا الحصول على العلاقة بين كل المتوسطات:

$$H < G < \bar{x} < Q$$

ويستعمل Q في:

- حساب الفروقات، ولتفادي الفروقات، نربّع القيم.

مثال: ليكن التوزيع التالي: المطلوب حساب كل أنواع المتوسطات

x_i	1	2	3	4	5	6	$\sum$

n_i	20	30	15	10	5	2	82
$x_i n_i$	20	60	45	40	25	12	202

$$\overline{x} = \frac{\sum n_i x_i}{N} =$$

$$\overline{x} = 2,46$$

$$H = \frac{82}{\frac{20}{1} + \frac{30}{2} + \frac{15}{3} + \frac{10}{4} + \frac{5}{5} + \frac{2}{6}} = \frac{N}{\sum \frac{n_i}{x_i}}$$

$$H = 1,87$$

$$G = \sqrt[n]{x_1^{n_1}, x_2^{n_2}, \ldots, x_2^{n_2}}$$

$$G = \sqrt[82]{1^{20} \cdot 2^{30} \cdot 3^{15} \cdot 4^{15} \cdot 5^{5}} \cdot 6^{2}$$

$$\log G = \frac{1}{82}\left(20\log 1 + 30\log 2 + 15\log 3 + 10\log 4 + 5\log 5 + 2\log 6\right)$$

$$G = 2,15$$

$$Q^2 = 7,71$$
$$Q = 2,78$$

$$H = 1,87 < G = 2,15 < \bar{x} = 2,46 < Q = 2,78$$

2. مقاييس التشتت (Caractéristique de dispersion):

لإجراء المقارنة بين السلاسل الإحصائية، قد نجد سلسلتين متشابهتين ولهما متوسطين حسابيين مختلفينوقد نجد سلسلتين مختلفتين، ولكن لهما نفس مقاييس النزعة المركزية لذلك، أصبحت هذه الأخيرة غير كافية لإعطاء معلومات حول توزيع القيم.

ولإجراء المقارنة، يجب قياس شدة أو قوة التجميع، قوة التركيز أو قوة التشتت بالنسبة لمقاييس النزعة المركزية.
إذ تسمح مقاييس التشتت هذه، بتقييم التوزيع العددي للسلسلة الإحصائية بالنسبة للقيمة المركزية.

1.2. المقاييس الأولية لقياس التشتت:

وهي أنواع، نجد منها:

- المدى العام للسلسلة.
- المدى الرُّبيعي والشعيري.
- الانحراف المطلق والمتوسط.
- التباين.
- الانحراف المعياري.

- معامل التغير والعزوم.

أ – المدى العام أو مجال التغير:

وهو المقياس الأكثر بساطة للدراسة، وهو الفرق بين القيمتين المتطرفتين في السلسلة للمتغير المدروس ويرمز له بـ E، نسبة إلى Etendue.

حيث:

$$E = x_{\max} - x_{\min}$$

ويستعمل في بعض الحالات مثل: -مراقبة الإنتاج، وبما أنه غير كاف لتمييز التشتت بصفة مقبولة، فهو قليل الاستعمال، لأنه يستعمل فقط قيمتين في السلسلة، ويتأثر بالقيم الشاذّة (كبيرة).

مثال:

$$x_i = \{2, 4, 6, 8, 10, 14, 16\}$$

$$E = 16 - 2 \qquad E = 14$$

ب – الانحراف المتوسط أو الفرق الحسابي:

الفرق الحسابي هو المتوسط الحسابي للفروق بالقيمة المطلقة، لجميع عناصر السلسلة والانحراف المتوسط يكتب حسب نوع السلسلة، ونرمز لـه بـ $\overline{E}$ نسبة إلى Ecart moyen.

- **حالة السلسلة الخامة:**

يكتب كالآتي:

$$\overline{E} = EM = \frac{1}{N} \sum \left| x_i - \overline{x} \right|$$

- **حالة السلسلة ذات التوزيع التكراري:**

ويكتب كالآتي:

$$\overline{E} = \frac{1}{N}\sum n_i \left| x_i - \overline{x} \right|$$

ملاحظة:

يمكن حساب الانحراف المتوسط بالنسبة لمقاييس النزعة المركزية الأخرى (M_θ, M_e)، بدلا من $\overline{x}$، حيث:

$$\overline{E} = \frac{1}{N}\sum n_i \left| x_i - M_\theta \right|$$

أو:

$$\overline{E} = \frac{1}{N}\sum n_i \left| x_i - M_e \right|$$

وبما أنه يأخذ القيم المطلقة، فهو لا يُستعمل، بالرغم من كونه المقياس الذي يعطي أكثر تمثيلا للتوزيع، ولذا نفضل استعمال الانحراف المعياري، بدلا من الانحراف المتوسط.

ج – المدى الرُّبيعي والعشيري:

ج.1. المدى الرّبيعي:

المدى أو المجال الرّبيعي هو: $Q_3 - Q_1$،ويعطي لنا مجال تغير 50% من المعطيات.

ج.2. المدى العشيري:

المدى أو المجال العشيري هو $D_9 - D_1$ ، ويعطي لنا مجال تغير 80% من المعطيات، ويعتبر هذين غير دقيقين، لأنهما يستعملان القيم القصوى.

د – التباين، الانحراف المعياري ومعامل التغير:

د.1. التباين (La variance):

ويرمز لها بـ: $V(x) = \sigma_x^2 = S_x^2$.

S_x^2 في حالة العيّنة.

σ_x^2 في حالة دراسة المجتمع.

ويعطى التباين بالعلاقة التالية:

- **حالة السلسلة الخامة:**

يكتب كالآتي:

$$V(x) = \frac{1}{N}\sum\left(x_i - \bar{x}\right)^2$$

- **حالة التوزيع التكراري:**

يكتب كالآتي:

$$V(x) = \sum f_i\left(x_i - \bar{x}\right)^2$$

$$V(x) = \frac{1}{N}\sum n_i\left(x_i - \bar{x}\right)^2$$

يمكن اختصار هذه العلاقة:

$$V(x) = \frac{1}{N}\sum n_i x_i^2 - \bar{x}^2$$

$$V(x) = \sum f_i\, x_i^2 - \bar{x}^2$$

د.2. الانحراف المعياري (L'écart - type):

ويرمز لها بـ: $\sqrt{V(x)} = \sigma_x = S_x$.

ويعطى بالكتابة الآتية:

●حالة السلسلة الخامة:

$$\sigma_x = \sqrt{V(x)} = \sqrt{\frac{1}{N}\sum\left(x_i - \overline{x}\right)^2}$$

●حالة التوزيع التكراري:

$$\sigma_x = \sqrt{\frac{1}{N}\sum n_i\left(x_i - \overline{x}\right)^2}$$

مثال: احسب التباين والانحراف المعياري للسلسلة التالية:

الفئـات	n_i	x_i	$n_i x_i$	x_i^2	$n_i x_i^2$
$[50,60[$	20	55	1100	3025	60500
$[60,70[$	40	65	2600	4225	169000
$[70,80[$	25	75	1875	5625	140625
$[80,90[$	5	85	425	7225	36125
$[90,100[$	10	95	950	9025	90250
Σ	100	/	6950	/	496500

$$\overline{x} = \frac{1}{N}\sum n_i x_i = \frac{1}{100}\times 6950$$

$$\bar{x} = 69,50$$

$$V(x) = \frac{1}{N}\sum n_i x_i^2 - \bar{x}^2$$

$$\bar{x}^2 = 4830,25$$

$$V(x) = \frac{1}{100}(496500) - 4830,25$$
$$V(x) = 4965 - 4830,25$$

$$V(x) = 134,75$$

$$\sigma_x = \sqrt{V(x)} = \sqrt{134,75}$$
$$\sigma_x = 11,61$$

د.3. خصائص التباين والانحراف المعياري:

●تباين القيمة الثابتة a: $V(a)$

$$V(a) = 0$$

إذا كان: $x_i = ax_i' - x_0$

$$V(x) = a^2 V(x_i')$$

إذا كان: $y = az_i$

$$V(y) = a^2 V(z_i)$$

●حالة وجود ثابتة ومتغير:

$$a+x \Rightarrow V(a+x)=V(x)$$
$$V(a+x)=V(x)=V(-x)$$

●حالة وجود متغيرين x وy:

$$(x+y) \Rightarrow V(x+y)=V(x)+V(y)+2\operatorname{cov}(x,y)$$

$\operatorname{cov}(x,y)$: التباين المشترك أو التغاير (**La covariance**).

$$\operatorname{cov}(x,y)=\frac{1}{N}\sum(x_i-\overline{x})(y_i-\overline{y})$$

ملاحظة:

يعتبر الانحراف المعياري من أهم مقاييس التشتت، وذلك لاستعماله لحساب معامل الارتباط وتحديد أشكال التوزيعات الإحصائية، بحيث يستعمل في تحديد نسب عدد الوحدات بالنسبة للتوزيع القريب من المتناظر، حسب الحالات التالية:

●إذا كان المجال يحتوي على:

* 66% من المعطيات، فالمجال يصبح $\left[\overline{x}\pm\sigma_x\right]$.

* 50% من المعطيات، فالمجال يصبح $\left[\overline{x}\pm 0,67\sigma_x\right]$.

* 95% من المعطيات، فالمجال يصبح $\left[\overline{x}\pm 2\sigma_x\right]$.

* 99% من المعطيات، فالمجال يصبح $\left[\overline{x}\pm 3\sigma_x\right]$.

- يتأثر σ_x بالقيم الشاذة، مثل المتوسط الحسابي $\overline{x}$.

- كون مقاييس التشتت المدروسة متعلقة بوحدة القياس، فوضعت أخرى مثل:

المجال الرُّبيعي النسبي:

$$I_Q = \frac{Q_3 - Q_1}{Q_2} = \frac{Q_3 - Q_1}{M_e}$$

د.4. معامل التغير (معامل الاختلاف):

ونرمز له بـ: $C.V.$

ويعطى من خلال العلاقة بين $\bar{x}$ و σ_x، حيث:

$$C.V. = \frac{\sigma_x}{\bar{x}} \times 100\%$$

$\bar{x}$ و σ_x من نفس الوحدة، ولإجراء المقارنة بين سلسلتين بالنسبة للتشتت، حيث توزيعها يختلف من حيث الوحدة، نستعمل هذا المعامل.

وهو يسمح بإجراء المقارنة بين مختلف السلاسل، وهو قيمة نسبية(%) ولا تعتمد على الوحدة.

مثال:

إذا كان لدينا توزيعين a وb، حيث:

$$\bar{x}_b = 11,32 \quad , \quad \bar{x}_a = 6 \text{ unités}$$

$$\sigma_b = 4,22 \quad , \quad \sigma_a = 2,27 \text{ unités}$$

فما هي السلسلة أو التوزيع الأكثر تشتُّتا؟

نحسب معامل التغير: $C.V.$ بالنسبة لـ aو**b**:

$$C.V_{\cdot a} = \frac{\sigma_a}{x_a} = \frac{2,27}{6} = 0,3783$$

$$C.V_{\cdot a} = 37,83\%$$

$$C.V_{\cdot b} = \frac{\sigma_b}{x_b} = \frac{4,22}{11,32} = 0,3728$$

$$C.V_{\cdot b} = 37,28\%$$

نستنتج أن نسبة تشتتهما **متقاربة**.

أما إذا قارنّا تشتت السلسلة في المثال.2حيث:

$$C.V_{\cdot c} = \frac{\sigma_c}{\overline{x}_c} \quad \Leftarrow \quad \begin{cases} \overline{x}_c = 69,5 \\ \sigma_c = 11,61 \end{cases}$$

$$= \frac{11,61}{69,5} = 0,1671$$

$$C.V_{\cdot c} = 16,71\%$$

ونقول أن التغير في المجموعة a أكبر منه في المجموعة c.

أوالتغير في المجموعة c أقل تشتّتًا بالنسبة لـ a و b.

هـ – العـزوم (Les moments):

هي مؤشرات نستعملها في تحديد شكل التوزيع، وتعتمد في حسابها على $\overline{x}$ و σ_x^2، ونجد منها:

- العزوم البسيطة m.

- العزوم المركزية µ.

هـ.1. العزوم البسيطة (Les moments simples)**:**

العزم m_r للبيانات x_i، حيث $x_i = \{x_1, \dots, x_n\}$، هو عبارة عن المتوسط الحسابي لقيم المتغيرة مرفوع إلى القوة r ويرمز له بـ m_r.

ويقرأ العزم البسيط من الدرجة r (Moment simple de degré r)، ويكتب حسب نوع التوزيع.

- **حالة السلسلة الخامة:**

$$m_r = \frac{1}{N} \sum x_i^r$$

- **حالة التوزيع التكراري:**

$$m_r = \frac{1}{N} \sum n_i x_i^r$$

ومراتب العزوم البسيطة تتراوح بين 0 و rولدينا:

$r=0$: $$m_0 = 1$$

$r=1$: $$m_1 = \overline{x}$$

$r=2$: $$m_2 = \frac{1}{N} \sum n_i x_i^2$$

هـ.2. العزوم المركزية (Moments centrés)**:**

نرمز له بـ μ_r (Moment centré d’ordre *r*)

- **حالة السلسلة الخامة:**

$$\mu_r = \frac{\sum\left(x_i - \overline{x}\right)^r}{N}$$

- **حالة التوزيع التكراري:**

$$\mu_r = \frac{\sum n_i \left(x_i - \overline{x}\right)^r}{N}$$

$r = 0$: $\mu_0 = 1$

$r = 1$: $\mu_1 = 0$

$r = 2$: $\mu_2 = V(x)$

العزوم البسيطة والعزوم المركزية مرتبطة في حالة حساب التباين.

$$V(x) = \frac{1}{N}\sum x_i^2 - \overline{x}^2$$

$$V(x) = m_2 - m_1^2 = \mu_2$$

هـ.3. العلاقة بين العزوم المركزية والعزوم البسيطة:

نعمّم μ_r حسب علاقة نيوتن (Newton)، وهي تستخدم معاملات مستخرجة من مثلث باسكال.

$r=3$:

$$\begin{aligned}(x_i-\bar{x})^3 &= (x_i-\bar{x})^2(x_i-\bar{x})\\ &= (x_i-\bar{x})\left(x_i^2-2x_i\bar{x}+\bar{x}^2\right)\\ &= x_i^3-3\bar{x}x_i^2+3x_i\bar{x}^2-\bar{x}^3\end{aligned}$$

$$\frac{1}{N}\sum n_i(x_i-\bar{x})^3=\frac{1}{2}\sum n_i\left(x_i^3-3\bar{x}x_i^2+3x_i\bar{x}^2-\bar{x}^3\right)$$

$$\mu_3=\frac{1}{N}\left(\sum n_i x_i^3-3\bar{x}\sum x_i n_i+3\bar{x}^2\sum x_i-\bar{x}^3\sum n_i\right)$$

$$\mu_3=\frac{1}{N}\sum_n n_i x_i^3-\frac{3\bar{x}\sum x_i n_i}{N}+3\bar{x}^2\frac{\sum x_i}{N}-\bar{x}^3\cancel{\frac{\sum n_i}{N}}$$

$$\mu_3=m_3-3m_1\cdot m_2+3m_1^3-m_1^3$$

$$\mu_3=m_3-3m_1m_2-2m_1^3$$

احسب أيضا بالنسبة لـ μ_4 وغيرها.

ملاحظة:

إذا كانت لدينا عيّنة أولى بحجم n_1، ومتوسطها الحسابي $\overline{x}_1$، وعينة ثانية بحجم n_2، ومتوسطها الحسابي $\overline{x}_2$، وتباين الأولى هو $V(x_1)$، وتباين الثانية هو $V(x_2)$.

فمتوسط العينتين معا هو متوسط المتوسطات:

$$\overline{X} = \frac{1}{n}\left(n_1\overline{x}_1 + n_2\overline{x}_2\right)$$

وتباين العينتين معا هو:

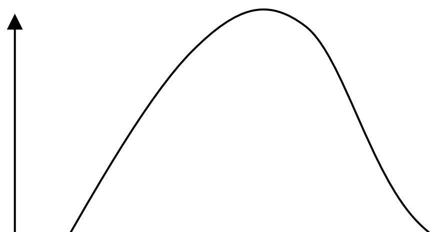
$$V(x) = \underbrace{\frac{1}{n}\left(n_1 \cdot V(x_1) + n_2 \cdot V(x_2)\right)}_{\text{متوسط التباينات}} + \underbrace{\frac{1}{n}\left(n_1\left(\overline{x}_1 - \overline{X}\right)^2 + n_2\left(\overline{x}_2 - \overline{X}\right)^2\right)}_{\text{تباين المتوسطات}}$$

3.مقاييس الشكل :

هذه المقاييس تستعمل لتحديد شكل التوزيع الإحصائي، ويطلق على نوعين من المعاملات، وذلك بحساب معامل الالتواء لدراسة تناظر التوزيع ومعامل التفلطح والتطاول، لدراسة التوزيع الطبيعي.

من المهم أن نميز تناظر وعدم تناظر (الالتواء نحو اليمين أو اليسار)، للتوزيع المدروس وكذلك تفلطح وتطاول السلسلة وذلك بواسطة معاملات الالتواء لدراسة التناظر ومعامل التفلطح لدراسة تطاول السلسلة.

1.3. معامل الالتواء (Coefficient d’asymétrie):



أ – تعريف الالتواء:

هناك 3 أنواع من التوزيعات:

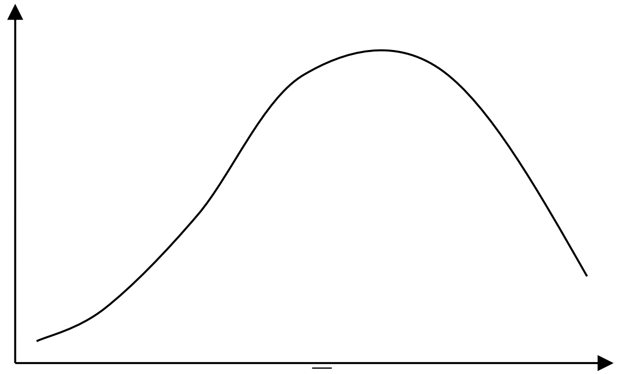
التوزيع الملتوي نحو اليسار:

يكون شكله كالآتي:

$$\bar{x} > M_e > M_\theta$$

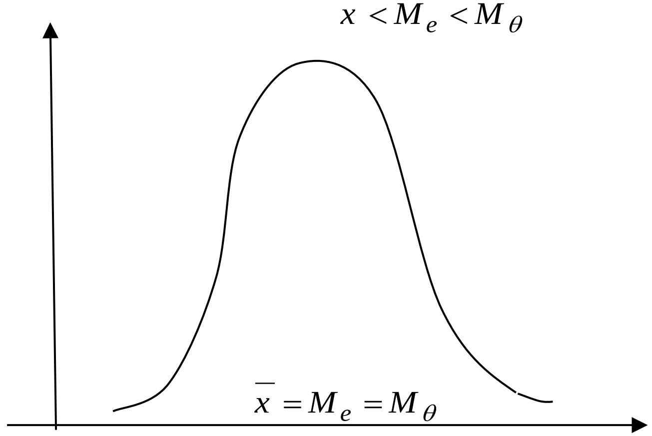
التوزيع الملتوي نحو اليمين:

يكون شكله كالآتي:



$$\bar{x} < M_e < M_\theta$$

التوزيع المتناظر:



$$\bar{x} = M_e = M_\theta$$

إذن، لدراسة التناظر، نستعمل القيم الثلاثة للنزعة المركزية M_θ، $\bar{x}$، M_e.

ب – معامل الالتواء:
لدينا 3 معاملات للالتواء:

- معامل يول (Kendall et Yull).
- معامل بيرسون (Pearson).
- معامل فيشر (Fisher)

ب.1. Coefficient Yull et Kendall (y_k):

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

$$y_k = \frac{Q_1 + Q_3 - 2M_e}{Q_3 - Q_1}$$

$$\boxed{-1 < y_k < +1}$$

$y_k = 0 \Leftarrow$ التوزيع متناظر.

$y_k < 0 \Leftarrow$ التوزيع ملتوي نحو اليمين.

$y_k > 0 \Leftarrow$ التوزيع ملتوي نحو اليسار.

ب.2. Coefficient Pearson:
لـ Pearson معاملين لتحديد شكل التوزيع.

- المعامل الأول:
يستخدم فيه $\overline{x}$ و M_θ و σ_x، ويستخدم في حساب معامل الالتواء بالنسبة للتشتت، ورمزه β_1.

ويكتب:

$$\beta_1 = \frac{\overline{x} - M_\theta}{\sigma_x}$$

$\beta_1 = 0 \Leftarrow$ التوزيع متناظر.

$\beta_1 < 0 \Leftarrow$ التوزيع ملتوي نحو اليمين.

$\beta_1 > 0 \Leftarrow$ التوزيع ملتوي نحو اليسار.

ـ المعامل الثاني:

يحسب من خلال العلاقة:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$\beta_1 = 0 \Leftarrow$ التوزيع متناظر.

$\beta_1 > 0 \Leftarrow$ التوزيع ملتوي نحو اليسار.

$\beta_1 < 0 \Leftarrow$ التوزيع ملتوي نحو اليمين.

الاتجاه عادة ما يكون معطى من خلال إشارة μ_2.

ب.3. Coefficient Fisher (γ_1):

يكتب:

$$\gamma_1 = \frac{\mu_3}{\sigma_x^3}$$

$\gamma_1 = 0 \Leftarrow$ تناظر.

$\gamma_1 > 0 \Leftarrow$ ملتوي نحو اليسار.

$\gamma_1 < 0 \Leftarrow$ ملتوي نحو اليمين.

2.3. معامل التفلطح (Coefficient d’aplatissement):

أ – مقياس Pearson (β_2):

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma_x^4}$$

$\beta_2 = 3 \leftarrow$ متناظر

$\beta_2 > 3 \leftarrow$ متطاول

β_2 قريب من $1 \leftarrow$ التوزيع مفلطح

ب – مقياس Fisher (γ_2):

$$\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\sigma_x^4} - 3$$

$\gamma_2 = 0 \Leftarrow$ تناظر.

$\gamma_2 > 0 \Leftarrow$ متطاول.

$\gamma_2 < 0 \Leftarrow$ مفلطح.

4. مقاييس التمركز (Les caractéristiques de concentration):

مفهوم التمركز مرتبط بمفهوم التشتت، وقياس التمركز هو نتيجة قياس التشتت، حيث أنه في حالة قياس التشتت، درسنا تقارب أو تباعد قيم x_i بالنسبة لـ $\overline{x}$ و r.

ولكن لم نراعي في هذه الدراسة توزيع x_i على مجموع الأفراد N.

أما في حالة قياس التمركز، فإننا نهتم بالعدالة أو التساوي في توزيع x_i على الأفراد N.

وهذا القياس مهم جدا في الاقتصاد، ويخص المتغيرات الكمية المستمرة الموجبة، حيث نهتم بقياس تمركز الأجور، أرقام الأعمال، عدد العمال، كمية الإنتاج...

وهي لا تخص الوحدات الإحصائية الخاصة، كالأوزان والأطوال.

وعمومًا، هناك طريقتين لتحديد قيمة التمركز: الطريقة الجبرية والطريقة البيانية

1.4. الطريقة الجبرية:

تكون هذه الطريقة بالخطوات التالية:

1 – حساب وسيط السلسلة M_e (Médiane).

2 – حساب الوسطى M_L (Médiale).

3 – حساب الفرق Δ_M حيث $\Delta_M = M_L - M_e$.

4 – مقارنة هذا الفرق Δ_M بـ E (المدى العام).

ويكون التعليق على النتائج كالآتي:

إذا كان: $\Delta_m >>> E$: نقول هناك: تمركز قوي جدا.

عدم التساوي في التوزيع.

لا عدالة في التوزيع.

$\Delta_M < E$: نقول هناك: تمركز ضعيف لوجود عدالة أو توازن في التوزيع.

$M_L = M_e \Leftarrow \Delta_M = 0$: هذا يعني أننا في وضعية تساوي في التوزيع.

وفي حالة توزيع الأجور، نقول أن العمال يتقاضون نفس الأجر.

●كيفية تحديد الوسطى (Médiale):

تعريف: الوسطى عبارة عن وسيط يحسب من خلال الجداء $n_i x_i$ ، وهي عبارة عن قيمة x_i ، وتقسم السلسلة $n_i x_i$ إلى قسمين متساويين و M_L تعتبر من مقاييس النزعة المركزية، وتحسب من خلال القانون التالي:

$$M_L = x^- + h \frac{\frac{F}{2} - F_{-1}^{\square}\left(\frac{n_i x_i}{\sum n_i x_i}\right)}{F^{\square}\left(\frac{n_i x_i}{\sum n_i x_i}\right) - F_{-1}^{\square}\left(\frac{n_i x_i}{\sum n_i x_i}\right)}$$

مثال: يبين الجدول التالي توزيع 44 عاملا حسب الأجور. والمطلوب حساب تمركز السلسلة.

Lf_i	$F^{\nearrow}y_i + F^{\nearrow}y_i - 1$	f_i	$F^{\nearrow}(y_i - 1)$	$F^{\nearrow}\left(\frac{n_i x_i}{\sum n_i x_i}\right)$	$\frac{n_i x_i}{\sum n_i x_i}$	$n_i x_i$	x_i	$N^{\nearrow}$	n_i	فئات	$F^{\nearrow}$
0,068	0,0601	0,1136	0	0,0601	0,0601	25	5	5	5	[4,6[	0,1136
0,0463	0,2548	0,1818	0,0601	0,1947	0,1346	56	7	13	8	[6,8[	0,2954
0,1770	0,649	0,2727	0,1947	0,4543	0,2596	108	9	25	12	[8,10[	0,5681
0,2667	1,1732	0,2273	0,4543	0,7187	0,2644	110	11	35	10	[10,12[	0,7961
0,3517	1,7187	0,2046	0,7187	1	0,2813	117	13	44	09	[12,14[	0,9997
0,8484	/	/	/	/	1	416	/	/	44	المجموع	

لحساب M_e: $N(M_e) = \frac{N}{2} = \frac{44}{2} = 22$

$13 < 22 < 25$

الفئة الوسيطية $[8,10[$

$$M_e = x_{M_e}^{-} + h \cdot \frac{\frac{N}{2} - N_{-1}^{\nearrow}}{nM_e} = 8 + 2\,\frac{22-13}{12}$$

$M_e = 9,5 \times 10^3$

لحساب M_L:

$$FM_L = \frac{F}{2} = 0,5$$

$$0,4543 < 0,5 < 0,7187$$

الفئة الوسطى $[10,12[$

$$M_L = x^- + h \frac{\frac{1}{2} - F_{-1}^{\square}\left(\frac{n_i x_i}{\sum n_i x_i}\right)}{F^{\square}\left(\frac{n_i x_i}{\sum n_i x_i}\right) - F_{-1}^{\square}\left(\frac{n_i x_i}{\sum n_i x_i}\right)}$$

$$M_L = 10 + 2\frac{0,5 - 0,7543}{0,7187 - 0,4543}$$

$$M_L = 10 + 2\frac{0,0457}{0,2644}$$

$$M_L = 10,3457 \times 10^3$$

تحديد الفرق Δ_M:

$$\Delta_M = M_L - M_e$$

$$\Delta_M = 10,3457 \times 10^3 - 9,5 \times 10^3$$

$$\Delta_M = 0,84 \times 10^3$$

نحسب E:

$$E = x_{\max} - x_{\min}$$
$$E = 14 - 4$$
$$E = 10 \times 10^3$$

$\Delta_M < E$: التمركز ضعيف، وهذا يعني تقارب في توزيع الأجور وليس مساواة.

ويمكن أيضا حساب مؤشر **التمركز** ***IC*** (Indice de concentration).

$$IC = \frac{\Delta_M}{E}$$

حيث:

$$IC = \frac{0,84}{10} = 0,084$$

$$IC = 8,4\%$$

لا وجود لمساواة في توزيع الأجور

$$\left.\begin{array}{l} IC \neq 0 \\ \Delta_M \neq 0 \end{array}\right\} \Rightarrow$$

خصائص *IC*:

$$0 < IC < 1$$

$IC = 0$: مساواة مطلقة (***égalité parfaite***).

$IC = 1$: تفاوت مطلق (***inégalité parfaite***)عدم مساواة مطلقة.

2.4. الطريقة البيانية:

تعود هذه الطريقة إلى الإيطالي Cossado Gini، خلال أبحاثه حول الاختلاف في توزيع الأجور، والذي كان يشكل النزاع الطبيعي، وبالتالي، أدّت أبحاثه إلى ما يسمى بمنحنى التمركز (La courbe de concentration)، وتحديد النسبة المسماة **بمؤشر Gini** أو **معامل Gini**.

أ – الرسم البياني للتمركز:

يمثَّل على معلم متعامد ومتجانس بالطريقة التالية:

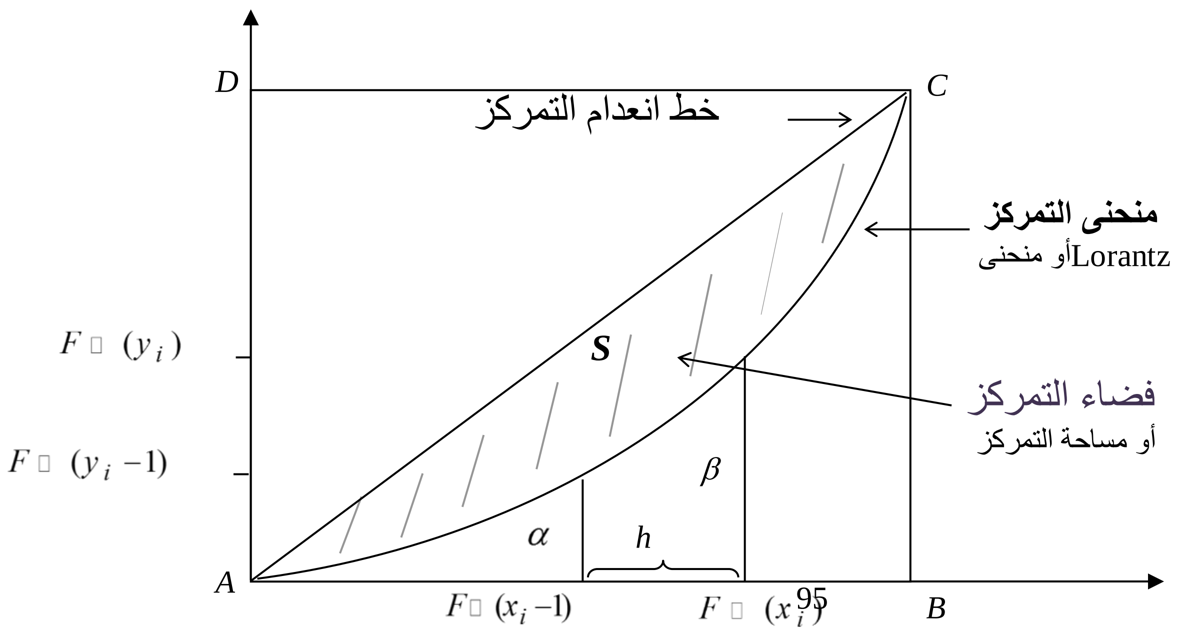
على (xx') نضع $F\uparrow(x_i)$.

و على (yy') نضع $F\uparrow\left(\dfrac{n_i x_i}{\sum n_i x_i}\right)$.

وهي عبارة عن الكتلة المجمّعة الصاعدة،

ونحصل على محور فواصل قياسه (1) لأن $F\uparrow(x_i)=1$.

ونحصل على محور تراتيب قياسه (1) لأن $F\uparrow(y_i)=1$.



$[AC]$ تمثل منصّف المربع $ABCD$

مساحة المثلث ABC = مساحة المربع / 2.

مساحة المربع = 1.

مساحة المثلث $= \frac{1}{2}$.

ب – حساب معامل Gini:

عبارة عن نسبة مساحة التمركز إلى مساحة المثلث ABC.

$$IG = S/(\frac{1}{2})$$

$$IG = 2S$$

$$0 < IG < 1$$

كلما ابتعد منحنى التمركز من المستقيم (AC)، كلما كانت المساواة منعدمة، ويعتبر التمركز قوي.

كلما اقترب منحنى التمركز من المستقيم (AC)، كلما كانت المساواة مطلقة، ويعتبر التمركز ضعيف.

$$IG = 2S$$

$$S = \sum_{1}^{k} S^*$$

S^*: مساحة شبه المنحرف.

وليكن لدينا k شبه منحرف.

$$S^{*}=\frac{(\alpha+\beta)}{2}\cdot h$$

$$h=F(x_i)-F(x_i-1)$$

$$h=f_i$$

$$\alpha=F(y_i-1)$$

$$\beta=F(y_i)$$

$$S^{*}=\frac{(\alpha+\beta)h}{2}=\frac{1}{2}\left[F(y_i-1)+F(y_i)\right]f_i$$

مساحة المربع $=1\times1=1$

مساحة المثلث $=\frac{1}{2}$

$$S=\frac{1}{2}-\sum_{1}^{k}S^{*}$$

$$S=\frac{1}{2}-\frac{1}{2}\sum_{1}^{k}\left[F(y_i-1)+F(y_i)\right]f_i$$

من خلال الجدول، يمكن لنا إعطاء نتائج هذه الحسابات:

$$IG=2\cdot S$$

$$IG=\cancel{2}\times\frac{1}{\cancel{2}}\left[1-\sum_{1}^{k}\left[F\square\left(y_i-1\right)+F\square\left(y_i\right)\right]f_i\right]$$

$$IG=1-\sum_{1}^{k}\left[F\square\left(y_i-1\right)+F\square\left(y_i\right)\right]f_i$$

حساب IG للمثال السابق:

$$IG=1-\sum_{1}^{k}\left[F\square\left(y_i-1\right)+F\square\left(y_i\right)\right]f_i$$

$$IG=1-0,8484=0,1516$$

$$IG=15,16\%$$

الفصل الرابع:

الأرقام القياسية (Les indices)

مفهوم الأرقام القياسية:

في مجال العلوم الاقتصادية والاجتماعية، تتغير القيم الهامة، كالأسعار، الإنتاج، معدل البطالة،...، الخ، عبر الزمن أو من مكان لآخر، أو عبر الزمن والمكان معا.

فعادة، الإحصائيات المعطاة على شكل جداول، تكون صعبة المقارنة، ولسهولة التفسير، نلجأ إلى استخدام مفهوم الأرقام القياسية.

والرقم القياسي هو مؤشر يسمح لنا بقياس تطور القيم على فترات، قد تكون الشهر، السداسي، الثلاثي أو السنة. وعمومًا، تستخدم هذه الأرقام القياسية لمعرفة تطور القدرة الشرائية وتكلفة المعيشة ومعدل التضخم. وقد يستعمل هذا المؤشر لحساب قيم مجموعة من السلع في فترة محدّدة بقيمة العملة لسنة مختارة، تسمّى سنة الأساس يشترط أن تكون هذه السنة، سنة عادية وأكثر استقرارًا من الناحية الاقتصادية.

ويمكن أن نميز نوعين من الأرقام القياسية، منها البسيطة والتجميعية.

1. الرقم القياسي البسيط: (L' indice élémentaire)

وهو عبارة عن نسبة القيم من نفس المقادير على فترتين، ويعبّر عنه ب $I_{\frac{1}{0}}$

الفترة الحالية يرمز لها بـ $t_1 \leftarrow 1$

الفترة الأساس يرمز لها بـ $t_0 \leftarrow 0$

ونحسب $I_{\frac{1}{0}}$.

$$I_{\frac{1}{0}} = I_{\frac{t_1}{t_0}} = \frac{v_1}{v_0} \times 100$$

والرقم القياسي هو نسبة بدون بُعد (وحدة) و يضرب في 100.

حيث:

v_1: القيمة في الفترة 1.

v_0: القيمة في الفترة 0.

ونميّز 3 حالات للرقم القياسي البسيط:

زيــادة $I_{\frac{1}{0}} > 1 \longrightarrow v_1 > v_0 \longrightarrow$

انخفاض $I_{\frac{1}{0}} < 1 \longrightarrow v_1 < v_0 \longrightarrow$

ثبــات $I_{\frac{1}{0}} = 1 \longrightarrow v_1 = v_0 \longrightarrow$

ويمكن استنتاج معدّل التطور وذلك بحساب:

$$\Delta V = \frac{v_1 - v_0}{v_0}$$

$$\Delta V = \frac{v_1}{v_0} - 1$$

$$\Rightarrow \boxed{\Delta V = I_{\frac{1}{0}} - 1}$$

$$\boxed{I_{\frac{1}{0}} = \Delta V + 1}$$

أو، إذا كان الرقم القياسي محسوب بالنسبة:

$$\boxed{\Delta V = I_{\frac{1}{0}} - 100}$$

مثال1:

إذا كان سعر السكر في 2014 هو 80 دج، وأصبح في عام 2015 يقدر بـ 95 دج، احسب الرقم القياسي البسيط لسعر السكر بين عامي 2014 و 2015.

$$I_{\frac{2014}{2015}} = \frac{V2015}{V2014} \times 100 = \frac{95}{80} \times 100 = 118,75\ \%$$

ونقول أن:

الرقم القياسي لسعر السكر في 2015، نسبة إلى سنة الأساس 2014 هو % 118,75.

ومنه يمكن حساب ΔV:

$$\Delta V = 118,75\% - 100\% = 18,75\ \%$$

أي نسبة الزيادة لسعر السكر هي % 18,75

.

.2 خواص الأرقام القياسية:

- **خاصية الثبات (Propriété de constance):**

$$I_{\frac{1}{1}} = \frac{v_1}{v_1} \times \frac{v_0}{v_0} \times 100 = 1 \times 100$$

- **خاصية الانعكاس (Propriété de réversibilité):**

$$I_{1/0} = \frac{V_1}{V_0} = \frac{1}{\frac{V_0}{V_1}} = \frac{1}{I_{0/1}} \times 100$$

$$I_{1/0} \times I_{0/1} = 1^2 \ (100^2)$$

- **خاصية الدوران أو الدورانية :(Propriété de la circularité ou de la transférabilité)**

(Indice chaîne حسب)

$$I_{2/0} = I_{2/1} \times I_{1/0}$$

$$I_{2/0} = \frac{V_2}{V_0} \times 100$$

$$I_{2/0} = \frac{V_2}{V_0} \times \frac{V_1}{V_1} \times 100$$

$$I_{2/0} = \frac{V_2}{V_1} \times \frac{V_1}{V_0} \times 100$$

و منه يمكن استنتاج $I_{2/1}$

$$I_{2/1} = \frac{I_{2/0}}{I_{1/0}} \times 100$$

يمكن حساب الرقم القياسي $I_{2/0}$ بالنسبة إلى سنة الأساس 0، باستخدام **الأرقام القياسية الوسيطية كالاتي :**

$$I_{\frac{2}{0}} = \frac{I_{\frac{2}{1}} \times I_{\frac{1}{0}}}{100^{2-1}}$$

$$I_{\frac{2}{0}} \times 100 = I_{\frac{2}{1}} \times I_{\frac{1}{0}} \times 100$$

ملاحظة:

➢ إذا كان الرقم القياسي لا يخضع لخاصية الدورانية، نستطيع إنشاء هذه الخاصية إذا علمنا $I_{\frac{1}{0}}$ و $I_{\frac{2}{1}}$.

فنكتب:

$$I_{\frac{2}{0}} = I_{\frac{2}{1}} \times I_{\frac{1}{0}}$$

ويمكن تعميمها إلى ما يسمى بـ **الرقم القياسي المتسلسل (*Indice chaîne*):**

إذا كانت السنة الحالية t وسنة الأساس 0، فنكتب:

$$I_{t/0} = I_t/t_{-1} \times I_{t_{-1}}/t_{-2} \times \dots\dots I_{1/0} \times 100$$

➢ وعموما:

إذا أردنا مقارنة قيمتين بسيطتين لفترتين، نحسب نسبة أرقامهما القياسية.

$$I_{t/t'} = \frac{I_{t/0}}{I_{t'/0}} \times 100$$

وتمكننا هذه الطريقة من تغيير نسبة الأساس بتعويض t' محل 0.

مثال 2:

الرقم القياسي	الأسعار	الفترة
100	150	2000 (0)
140	210	2005 (t')
153,3	250	2008 (t)

$$I_{\frac{t'}{0}} = \frac{153,3}{140} \times 100 = 109,5$$

$$\Delta = 109,5 - 100 = +9,5\%$$

مما يدل على زيادة(**+**) الأسعار بين 2008 و 2005 تقدر ب % 9,5

$$I_{\frac{t'}{0}} = \frac{140}{100} \times 100 = 140$$

$$\Delta = 140 - 100 = +40\%$$

مما يدل على زيادة الأسعار بين سنتي 2005 و2000.

و باستخدام الأرقام القياسية الوسيطية نجد:

$$I_{\frac{t'}{0}} = \frac{I_{\frac{t}{t'}} \times I_{\frac{t'}{0}}}{100^{2-1}} = \frac{109,5 \times 140}{100^{2-1}}$$

$$I_{\frac{t'}{0}} = 153,3$$

$$\Delta = 153,3 - 100 = +53,3\%$$

مما يدل على زيادة الأسعار بين 2000 و 2008 بنسبة 53,3%

مثال 3:

إذا كان سعر الزّبدة في سنة 2000 مقدّر بـ 65 دج وأصبح يساوي في 2007 إلى 85 دج، وفي 2009 أصبح يقدّر بـ 110 دج.

احسب الرقم القياسي لسعر الزّبدة عام:

2007 نسبة إلى 2000

2009 نسبة إلى 2007

ثم للفترة 2000 - 2009 باعتبار سنة 2000 كسنة أساس :

$$I_{\frac{2007}{2000}} = \frac{V2007}{V2000} \times 100 = \frac{85}{65} \times 100 = 130,76\ \%$$

$$\Delta = +30,76\ \%$$ (↗ السعر)

$$I_{\frac{2009}{2007}} = \frac{\boldsymbol{V2009}}{\boldsymbol{V2007}} \times \mathbf{100} = \frac{\mathbf{110}}{\mathbf{85}} \times \mathbf{100} = \mathbf{129,41}\%$$

(السعر ↗) $\Delta = +\mathbf{29,41}\ \%$

❖ وإذا أردنا حساب الرقم القياسي للفترة بين 2000- 2009، نحسب باستخدام خاصية الدورانية حيث:

$$I_{2009}/_{2000} = \frac{I_{2009}/_{2007} \times I_{2007}/_{2000}}{100^{2-1}}$$

$$I_{2009}/_{2000} = \frac{\mathbf{129,41}\% \times \mathbf{130,76}\ \%}{100^{2-1}}$$

$$I_{2009}/_{2000} = 169{,}21\ \%$$

$$\Delta = +\mathbf{69,21}\ \%$$

مما يدل على زيادة (+) أسعار الزبدة خلال الفترة بين 2000 و 2009 بنسبة 69,21 %.

مثال 4:

تحصّلنا على زيادة قدرها 62,5% بين 2004 و 2005، علما أن سنة الأساس هي 2004،

و اذا كان لدينا زيادة قدرها 45% في سنة 2004 بالنسبة لسنة 2003 فما هو الرقم القياسي ل 2005 بالنسبة ل 2003

أي نقارن أسعار 2004 بـ 2005.

$$\boldsymbol{I_{2005}/_{2003}} = \frac{I_{2005/2004}}{I_{2004/2003}}$$

$$I_{2005/2003} = \frac{62,5\%}{45\%} \times 100 = 138,88\%$$
$$\Delta = 138,88\% - 100\% = 38,88\%$$

زيادة الأسعار بـ 38,88% خلال 2003 و 2005.

2. الأرقام القياسية التجميعية (*Les indices synthétiques*):

عبارة عن نسبة بين أسعار وكميات مجموعة من المواد في السنة الحالية، ومجموع أسعار وكميات هذه المواد في السنة الأساس.

أ – الأرقام القياسية غير المرجّحة (*Indices synthétiques simples ou non pondérés*):

و نعني بذلك أن مجموع السلع التي نقارنها، لها نفس الوزن أو الترجيح، ويتم حسابها بطريقتين:

- **الطريقة الاولى طريقة الأرقام القياسية المتوسطة:**

و هي عبارة عن نسبة متوسط قيم مجموع السلع في فترة المقارنة (أو الفترة الحالية (1))، ومجموع قيم السلع لفترة الأساس، أي إذا كان لدينا k سلعة وفترتين:

- فترة الأساس: 0
- فترة المقارنة: 1

فقيم هذه السلع تكون في فترة المقارنة (1) v_1^i، حيث نحصل على $v_1^1, v_1^2, \dots, v_1^k$.

وتكون قيم هذه السلع في فترة الأساس (0) v_0^i، حيث نحصل على $v_0^1, v_0^2, \dots, v_0^k$.

فالرقم القياسي المتوسط لـ k سلعة هو:

$$I_{\frac{1}{0}} = \frac{\frac{1}{k}\sum_{1}^{k} v_1^i}{\frac{1}{k}\sum_{1}^{k} v_0^i}$$

$$I_{\frac{1}{0}} = \frac{\sum_{1}^{k} v_1^i}{\sum_{1}^{k} v_0^i} \times 100$$

- **الطريقة الثانية طريقة متوسط الأرقام القياسية:**

وهو المتوسط الحسابي للأرقام القياسية البسيطة لكل سلعة، بحيث الرقم القياسي البسيط للسلعة (1) يكون $I_{1\frac{1}{0}}$.

$$I_{1\frac{1}{0}} = \frac{v_1^1}{v_0^1} \times 100$$

للسلعة j يكون $I_{j\frac{1}{0}}$.

$$I_{j\frac{1}{0}} = \frac{v_1^j}{v_0^j} \times 100$$

إذا كان لدينا kسلعة، فمتوسط الأرقام القياسية في فترة (1) وفترة (0) يكتب كالآتي:

$$\overline{I_{\frac{1}{0}}} = \frac{1}{k}\sum_{j=1}^{k} I_{j\,\frac{1}{0}} \times 100$$

ملاحظة:

- إن متوسط الأرقام القياسية هو الأكثر استعمالا، لأنه لا يتعلق بالوحدة، بالرغم من أنه لا يحقق خاصية التحويل ولا خاصية الدوران.

- في الاقتصاد نهتم بتغيرات الأسعار أو الكميات، أو تغيرات جدائهما، والذي نسميه القيمة الإجمالية (***Valeur globale***)، وهي قيمة نقدية تحسب كالآتي:

$$V_G = \text{السعر} \times \text{الكمية}$$

القيمة الإجمالية = السعر × الكمية

ب – الأرقام القياسية التجميعية المرجّحة:

هي أرقام قياسية قريبة من الواقع ونعطي التطور الإجمالي للقيم، آخذين بعين الاعتبار قيمة السلعة ***i*** ووزنها w_i.

مثلا: السلع أو المواد التي تدخل في تكوين ميزانية المستهلك المخصصة للإنفاق، تختلف فيما بينها من حيث الأهمية.

ومن أهم الأرقام القياسية المرجحة المستعملة، هي الأرقام القياسيـة لـ *Laspeyrs* وPaashes.

- **الرقم القياسي للاسبير (*Laspeyrs*):**

وهوعبارة عن متوسط حسابي مرجّح للأرقام القياسية البسيطة لكل سلعة، وتسمى هذه الترجيحات بأوزان (weight)، ونرمز لها بـ W_i، حيث الترجيح يكون بالنسبة لسنة الأساس ويكتب :

$$W_i = P_0^i \cdot Q_0^i$$

حيث:

P_0^i : سعر السلعة ***i*** في سنة الأساس.

و Q_0^i: كمية السلعة ***i*** في سنة الأساس.

فيكون الرقم القياسي لـ Laspeyrs على الشكل:

$$I_L = \frac{\sum W_i G_i}{\sum W_i}$$

حيث:

$$G_i = \frac{v_1^i}{v_0^i}$$

ويمكن التمييز بين ***3 أرقام قياسية***:

- **رقم قياسي للأسعار.**
- **رقم قياسي للكميات.**
- **رقم قياسي التجميعي.**

وبالتالي: G_i يكتب:

إذا كان خاص بالأسعار: $G_i = \frac{P_1^i}{P_0^i}$.

إذا كان خاص بالكميات: $G_i = \frac{Q_1^i}{Q_0^i}$.

1* - الرقم القياسي للأسعار:

يرمز له بـ $I_{LP\frac{1}{0}}$:

$$I_{LP\frac{1}{0}} = \frac{\sum W_i G_i}{\sum W_i} = \frac{\sum \cancel{P_0} \cdot Q_0 \times \frac{P_1}{\cancel{P_0}}}{\sum P_0 \cdot Q_0}$$

$$\boxed{I_{LP\frac{1}{0}} = \frac{\sum P_1 \cdot Q_0}{\sum P_0 Q_0} \times 100}$$

الترجيح يكون بالنسبة لسنة الأساس أي ترجح الأسعار بكميات سنة الأساس.

2* - الرقم القياسي للكميات:

يرمز له بـ $I_{LQ\frac{1}{0}}$:

$$I_{LQ\frac{1}{0}} = \frac{\sum W_i G_i}{\sum W_i} = \frac{\sum P_0 \cdot \cancel{Q_0} \times \frac{Q_1}{\cancel{Q_0}}}{\sum P_0 \cdot Q_0}$$

$$\boxed{I_{LQ\frac{1}{0}} = \frac{\sum P_0 \cdot Q_1}{\sum P_0 Q_0} \times 100}$$

وترجّح الكميات بأسعار سنة الأساس.

3* - الرقم القياسي للقيم الإجمالية:

يرمز له بـ $I_{L_{V.G}}$:

$$\boxed{I_{L_{V.G}} = \frac{\sum P_1 \cdot Q_1}{\sum P_0 Q_0} \times 100}$$

مثال5: من خلال الجدول أدناه، وإذا كانت سنة الأساس هي 2000، والسنة الحالية هي 2008.

احسب الأرقام القياسية للأسعار I_{LP}، وللكميات I_{LQ}، وللقيم الإجمالية $I_{L_{V.G}}$.

السنة / البيان / المواد	2000 (0)		2008 (1)					
	P_0	Q_0	P_1	Q_1	P_0Q_0	P_1Q_0	P_0Q_1	P_1Q_1
المادة 1	15	50	20	80	750	1200	1000	1600
المادة 2	25	70	35	100	1750	2500	2450	3500
المادة 3	45	300	30	350	13500	15750	9000	10500
Σ					16000	19450	12450	15600

$$I_{LP_{\frac{08}{00}}} = \frac{\sum P_1 Q_0 \times 100}{\sum P_0 Q_0} = \frac{12450}{16000} \times 100 = 0,7781$$

$$I_{LP_{\frac{08}{00}}} = 77,81\%$$

$$\Delta = I_{\frac{08}{00}} - 100 = 77,81 - 100$$

$\Delta = -22,19\% \Rightarrow$ **انخفاض سعر السلعة بين 2008 و 2000**

$$I_{LQ_{\frac{2008}{2000}}} = \frac{\sum P_0 Q_1}{\sum P_0 Q_0} \times 100 = \frac{19450 \times 100}{16000} 1,2156 \times 100$$

$$I_{LQ_{\frac{08}{00}}} = 121,56\%$$

$$\Delta = I_{\frac{08}{00}} - 100 = 121,56 - 100$$

$\Delta = +21,56\% \Rightarrow$ **زيادة كميات السلع بين 2008 و 2000**

$$I_{L_{VG_{\frac{08}{00}}}} = \frac{\sum P_1 Q_1}{\sum P_0 Q_0} \times 100 = \frac{15600}{16000} \times 100 = 0,975 \times 100$$

$$I_{L_{VG_{\frac{08}{00}}}} = 97,5\%$$

$$\Delta = I_{\frac{08}{00}} - 100 = 100 - 97,5$$

$\Delta = +2,5\% \Rightarrow$ **انخفاض في القيمة الإجمالية للسلع بـ 2,5% بين 2008 و 2000**

4. بعض الأرقام القياسية الأخرى:

1.4.الرقم القياسي لباش (*Paashes*):

ونرمز له بـ I_P.

وهو عبارة عن المتوسط التوافقي المرجّح للأرقام القياسية البسيطة.

ويرمز له بـ I_P.

ويكتب:

$$I_P = \frac{1}{\frac{1}{\sum W_i} \cdot \sum W_i \cdot \frac{1}{G_i}}$$

$$\boxed{I_P = \frac{\sum W_i}{\sum W_i \cdot \frac{1}{G_i}}}$$

$$I_{P_{P\frac{1}{0}}} = \frac{\sum P_1 Q_1}{\sum P_1 Q_1 \cdot \frac{1}{\frac{P_1}{P_0}}} \times 100$$

$$I_{P_{P\frac{1}{0}}} = \frac{\sum P_1 Q_1}{\sum \cancel{P_1} Q_1 \cdot \frac{P_0}{\cancel{P_1}}} \times 100$$

$$\boxed{I_{P_{P\frac{1}{0}}} = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100}$$

ترجّح الأسعار بكميات سنة المقارنة (1).

2* - الرقم القياسي للكميات:

يرمز له بـ $I_{P_{Q\frac{1}{0}}}$:

$$I_{P_{Q\frac{1}{0}}} = \frac{\sum P_1Q_1}{\sum P_1Q_1 \cdot \frac{1}{\frac{Q_1}{Q_0}}} \times 100$$

$$I_{P_{Q\frac{1}{0}}} = \frac{\sum P_1Q_1}{\sum P_1 \cancel{Q_1} \cdot \frac{Q_0}{\cancel{Q_1}}} \times 100$$

$$\boxed{I_{P_{Q\frac{1}{0}}} = \frac{\sum P_1Q_1}{\sum P_1Q_0} \times 100}$$

ترجّح الكميات بأسعار سنة المقارنة أو السنة الحالية.

3* - الرقم القياسي للقيم الإجمالية:

يرمز له بـ $I_{P_{VG\frac{1}{0}}}$:

$$= \boxed{I_{P_{VG\frac{1}{0}}} = \frac{\sum P_1Q_1}{\sum P_0Q_0} \times 100}$$

$$\boxed{I_{P_{VG\frac{1}{0}}} = I_{L_{VG\frac{1}{0}}}}$$

2.4. الرقم القياسي لـ فيشر (***Fisher***):

ونرمز له بـ I_F.
وهو عبارة عن المتوسط الهندسي للرقمين القياسيين لـ Laspeyrs و Paashe.

$$I_F = \sqrt{I_P \times I_L}$$

وهناك 3 أرقام قياسية لـ Fisher (VG, Q, P).

***1 - الرقم القياسي للأسعار:**

يرمز له بـ I_{FP}:

$$I_{FP\frac{1}{0}} = \sqrt{I_{PP\frac{1}{0}} \times I_{LP\frac{1}{0}}}$$

***2 - الرقم القياسي للكميات:**

يرمز له بـ I_{FQ}:

$$I_{FQ\frac{1}{0}} = \sqrt{I_{PQ\frac{1}{0}} \times I_{LQ\frac{1}{0}}}$$

*** - الرقم القياسي للقيم الإجمالية:**

يرمز له بـ I_{FVG}:

$$I_{FVG\frac{1}{0}} = \sqrt{I_{PVG\frac{1}{0}} \times I_{LVG\frac{1}{0}}}$$

ملاحظة:

الأرقام القياسية لـ ***Laspers*** و***Paashes*** لا تحقق خاصية الانعكاس، أما ***Fisher*** فيحقق ذلك.

الأرقام القياسية الثلاثة (F, P, L) لا تحقق خاصية الدورانية.

3.4. الرقم القياسي لـ مارشال (*Marshall*):

ونرمز له بـ I_M.

***1 - الرقم القياسي للأسعار:**

يرمز له بـ $I_{MP\frac{1}{0}}$:

$$I_{MP\frac{1}{0}} = \frac{\sum P_1 (q_0 + q_1)}{\sum P_0 (q_0 + q_1)}$$

***2 - الرقم القياسي للكميات:**

يرمز له بـ $I_{MQ\frac{1}{0}}$:

$$I_{MQ\frac{1}{0}} = \frac{\sum Q_1 (p_0 + p_1)}{\sum Q_0 (p_0 + p_1)}$$

4.4. بعض الأرقام القياسية الأخرى:

***1 - الرقم القياسي للأجور (*Indices des salaires*):**

ويرمز له بـ I_S.

وهو مؤشر يقيس تطور الأجور بين فترتين زمنيتين مختلفتين، لفئة مهنية اجتماعية معينة.
ويمكن صياغته على طريقتين:

ط1: الترجيح بواسطة عدد الأفراد المنتمون إلى فئة اجتماعية مهنية معينة.
ويكتب:

$$I_{S\frac{1}{0}} = \frac{\sum n_0^i \times \frac{S_1^i}{S_0^i}}{\sum n_0^i} \times 100$$

حيث:

n_0^i: عدد أفراد الفئة المهنية الاجتماعية i في سنة الأساس.

S_0^i: أجور الفئة i في سنة الأساس.

S_1^i: أجور الفئة i في السنة الحالية.

ط2: تعتمد أساسا على علاقة Laspeyrs، وتكتب:

$$I_{S\frac{1}{0}} = \frac{\sum n_0^i S_0^i \times \frac{S_1^i}{S_0^i}}{\sum S_0^i \times n_0^i} \times 100$$

*** - الرقم القياسي للبورصة I_B (*Les indices boursiers*)**

وهي مؤشرات مالية في العالم، ومن أشهرها

- Dow Jones معمول بها في بورصة New York،
- Financial Times معمول به في بورصة London،
- Cac 40 معمول به في بورصة Paris.

مثلا:

مؤشر Dow Jones يقيس تطور أسعار الأسهم في بعض الشركات العامة، مثل General Motors، وقد يحسب حسب ترجيح Laspeyrs أو Paashes، أو استخدام الرّسملة(capitalisation) في السنة الحالية وفي سنة الأساس.

*3 - الرقم القياسي للقدرة الشرائية (*Indice de pouvoir d'achat*):

القدرة الشرائية للعملة وتحديد الأجر الاسمي والأجر الحقيقي

$$\text{الاجر الحقيقي} = \frac{\text{الأجر الاسمي}}{\text{الرقم القياسي لأسعار الاستهلاك}}$$

$$S_R = \frac{S_N}{I_{PC}}$$

$$\text{Salaire réel} = \frac{\text{salaire nominal}}{\text{Indice des prix de consommation}}$$

- ➢ لدراسة المستوى المعيشي للفرد:
- ✓ نحسب نسبة الأجور بين الفترتين المعنية بالدراسة.

$$\frac{S_1}{S_0} \times 100$$

- ✓ نقارن بين نسبة الزيادة في الأجر ونسبة الزيادة في أسعار الاستهلاك، أي نحسب Δ_S % و Δ_P %.
- ✓ وهذا الفرق بين النسبتين يبين من أجل الحفاظ على نفس القدرة الشرائية، حيث يجب أن يصبح أجر العامل الحقيقي مساويا لـ $I_P \times S_0$ الحالي.

$$S_R = S_0 \times I_P$$

ومنه:

$$\Delta_S = (S - S_1) \times 100$$

- ✓ ومنه، نسبة التغير في القدرة الشرائية $\frac{\Delta}{S}\%$ ، وهي نتيجة لارتفاع أسعار مواد الاستهلاك.
- ✓ ويمكن تقديرها بحساب الرقم القياسي للقدرة الشرائية:

$$I_{PA\frac{1}{0}} = \frac{I_{S\frac{1}{0}}}{I_{P\frac{1}{0}}} \times 100 = \frac{\frac{S_1}{S_0} \times 100}{I_{P\frac{1}{0}}}$$

والنسبة المطلوبة هي: $\Delta = 100 - I_{PA}$.

مثال6:

بلغ الرقم القياسي لأسعار الاستهلاك في شهر جوان 2007: 198,4%، علما أن سنة الأساس هي 2000.
نفرض أن الأجر الاسمي السنوي لعامل ما مساوي إلى 200000 دج سنة 2007.

حدّد الأجر الحقيقي لهذا العامل.

لنفرض ان أجر نفس العامل كان يساوي 130000 دج سنة 2000.
هل تحسّن مستوى معيشة هذا العامل؟

الحل:

- الأجر الحقيقي للعامل S_R لسنة 2000/2007:

$$S_R = \frac{S_N}{I_{PC}} = \frac{200000}{198,4} \times 100$$

$$\boxed{S_R = 100806,45\,\mathrm{DA}}$$

نظرا لارتفاع الأسعار وانخفاض مستوى المعيشة، فإن 200000 دج في سنة 2007، مقارنة بسنة 2000 يساوي 100806,5 دج.

- لدراسة المستوى المعيشي لهذا العامل:

نحدّد نسبة زيادة الأجرين في 2007 و 2000:

$$\frac{200000}{130000} \times 100 = 153,846\%$$

نلاحظ أن نسبة الزيادة في الأجر بين الفترتين بلغت 53,85+%.
أما الأسعار فقد ارتفعت بـ 98,40+%.

والفرق بين النسبتين:

$$98,4\% - 53,85\% = 44,55\%$$

هذا لا يعني أن القدرة الشرائية انخفضت في سنة 2007، وإنما هذه الأخيرة.

نبيّن أنه من أجل المحافظة على نفس القدرة الشرائية لسنة الأساس، يجب أن يصبح أجر العامل لسنة 2007 = $S_{2000} \times I_P$.

$$S_{2007} = S_{2000} \times I_P$$
$$= 130000 \times 198,4\%$$

$$\boxed{S_{2007} = 257920\,\text{DA}}$$

بناء على هذا، يفقد العامل بين 2000 و 2007 مقدار:

257920 – 200000 = 57920 DA

نتيجة لارتفاع أسعار مواد الاستهلاك، وهذا بنسبة تقدّر بـ:

$$\frac{57920}{257920} \times 100 = 22,46\%$$

ويمكن أن نصل إلى نفس النتيجة باستعمال I_{PA} ، وهو عبارة عن $\frac{I_S}{}$.

$$I_{PA} = \frac{I_S}{I_P} = \frac{153,85\%}{198,4\%}$$

$$I_{PA} = 77,54\%$$

$$\Delta = 100 - 77,54 = 22,46\%$$

ونقول لم يتحسن مستوى معيشة هذا العامل، بل ***انخفض***.

تمرين :

1) تحديد الرقم القياسي للأجور بترجيح ***Laspeyrs***:

$$I_{LS\frac{01}{00}} = \frac{\sum \frac{S^i_{2001}}{S^i_{2000}} \times S^i_{2000} \times S^i_{2000}}{\sum S^i_{2000} \times n^i_0} \times 100$$

$\frac{S^i_1}{S^i_0} \times n^i_0$	$S^i_0 n^i_0 \times \frac{S^i_1}{S^i_0}$	$\frac{S^i_1}{S^i_0}$	$S^i_0 n^i_0$	S^i_1	S^i_0	n^i_0	الفئة
159,539	1041471,89	1,0496	992256	6852	6528	152	عمّال
72,991	634074,55	1,0734	590716	9325	8687	68	إطارات متوسطة
18,3906	230397,44	1,0217	225504	12800	12528	18	إطارات عليا
250,9206	1905943,88	3,14477	1808476	28977	27743	238	Σ

$$I_{LS} = \frac{1905943,98}{1808476}$$

$$I_{LS} = \frac{1905943,98}{1808476}$$

$$I_{LS} = 105,38\%$$

$$\Delta = 105,38 - 100$$

$$\Delta = +5,38\%$$

⇐ ارتفعت الأجور بنسبة ***%5,38***.

2) تحديد الرقم القياسي للأجور بالترجيح العددي :

$$I_{S\frac{1}{0}} = \frac{\sum \frac{S_1}{S_0} \times n_0^i}{\sum n_0^i}$$

$$I_{S\frac{1}{0}} = \frac{250,9206}{238}$$

$$\Delta = +5,43\%$$

إذن، ارتفعت الأجور بـ %5,43.

الفصل الخامس

السلاسل الإحصائية ثنائية البعد

والعلاقة بين المتغيرات الإحصائية

مقدمة:

في الفصول الأخرى، تعرّفنا على التوزيعات ذات متغير واحد، مع توظيف خصائص هذا المتغير. وفي الجداول المتحصل عليها، نعتبر المتغير x_i بتكراره n_i.

في هذا الفصل، نتعرّض إلى دراسة المجتمعات، من خلال جداول تأخذ بعين الاعتبار صفتين كمّيتين، حيث تقدّر كل واحدة بمميزات مختلفة، نرمز لها بـ i و j معًا؛ فنحصل على جداول ثنائية البعد بـ x_i و y_i متغيرة ، حيث يمكن دراسة الخواص، كالمتوسطات والتباينات، ودراسة ما إذا كانت هناك علاقة فيما بين المتغيرات، مما يعرّف بالارتباط ومعامل الارتباط، وتحديد معادلة خط الانحدار.

I. الجداول الإحصائية ذات البعدين أو الجداول المتقاطعة:

نعتبر توزيع إحصائي المتغيرين x و y،

حيث: x بمميزات i فيصبح x_i ؛

y بمميزات j فيصبح y_j.

فنحصل على جدول إحصائي لبعدين $x_i \ , i=\overline{1,p}$ و $y_j \ , j=\overline{1,k}$ ،

Y \ X	y_1	y_2	....	y_k	$n_{\bullet i}$
x_1	n_{11}	n_{12}	....	n_1k	$n_{1\bullet}$
x_2	n_{21}	n_{22}	....	n_2k	$n_{2\bullet}$
⋮	⋮	⋮	....	⋮	⋮
x_p	n_{p1}	n_{p2}	....	n_{pk}	$n_{\bullet p}$
$n_{\bullet j}$	$n_{\bullet 1}$	$n_{\bullet 2}$	....	$n_{\bullet k}$	$n_{\bullet\bullet}$

ملاحظات:

1) مميزات المتغير y تظهر على السطر.
مميزات المتغير x تظهر على العمود.

2) تظهر تكرارات n_{ij} داخل الجدول، فنحصل على مصفوفة ذات البعد (p, k) لأن: $i = \overline{1, p}$ و $j = \overline{1, k}$.

3) يصبح التوزيع: (x_i, y_i, n_{ij}).

وللتعبير عن التكرارات الهامشية (المجمّعة) في الهوامش، نعوض المؤشر أو الرمز i أو j الذي يتغير بنقطة، فنحصل على المجموع $\sum n_{ij}$.

بحيث: i متغير $\sum_{i=1}^{p} n_{ij} = n_{\bullet j} \longrightarrow$

$$\sum_{j=1}^{k} n_{ij} = n_{i\cdot} \longrightarrow \text{متغير } j$$

والتكرارات الإجمالية تكتب $n_{\cdot\cdot}$.

وهو عدد أفراد المجتمع المدروس.

حيث:

$$\sum n_{i\cdot} = \sum n_{\cdot j} = n_{\cdot\cdot}$$

II. دراسة العلاقة بين المتغيرات الإحصائية (الانحدار والارتباط):

عند الحصول على جدول التقاطع ذو قطر مساوي إلى 1 وباقي التكرارات مساوية لـ 0.
فهذا لا يعني وجود علاقة تامة بين x و y، وإنما يعني إمكانية وجود علاقة نسبية بين x و y.
ويمكن استبدال الجدول التقاطعي بجدول آخر من الشكل:

x_i	y_i	n_i
x_1	y_1	n_1
x_2	y_2	n_2
$\vdots$	$\vdots$	$\vdots$
x_k	y_k	n_k

وهذا الجدول هو الأكثر استعمالا في دراسة المتغيرات لـ n فرد أو لـ t فترة من الزمن.
ومن أجل التحقق من وجود علاقة بين x و y ونسبة هذه العلاقة، نقوم بتحديد معادلة خط الانحدار وحساب معامل الارتباط.

وإن وجدت هذه العلاقة، نحدّد معاملات خط الانحدار الذي هو خط مستقيم يربط بين x و y، وذلك باستخدام طريقة المربعات الصغرى ***(MCO) Moindres carrés ordinaires***.

والتي تكمن في رسم سحابة النقاط وتحديد اتجاه نقاط هذه السحابة برسم خط مستقيم.

مبدأ MCO هو إيجاد أفضل خط مستقيم يربط بين x و y، بحيث مجموع مربعات البواقي (الأخطاء) يكون أصغر ما يمكن، والبواقي عبارة عن الفرق بين قيم y التقديرية، أي المحصورة عن طريق أفضل خط مستقيم ؛ وذلك تحت اعتبار الفرضية الأساسية التي تشمل النقاط التالية:

- متوسط الأخطاء u_i هو: $\bar{u} = 0$.
- تباين u_i ثابت ومعلوم: $V(u_i) = \sigma^2$.
- u_i له توزيع طبيعي (متناظر).

$$y_i = ax_i + b + u_i$$

ويمكن تحديد ثلاث علاقات بين x و y:

ـ علاقة كلية (دالية):
تحديد x يعطي مباشرة تحديد y.

ـ علاقة جزئية:
حالة عامة، بحيث x يحدّد بـ y أو y بـ x، ولكن بدرجة معيّن:

مثال (الدخل والاستهلاك):

ـ علاقة منعدمة بين x و y:

أي لا وجود لعلاقة بين x و y، ويعطي استقلالية بين x و y.

إيجاد قيم a و b لمعادلة خط الانحدار، والتي تجعل من $\sum\left(y_i - y\right)^2$ أصغر ما يمكن حيث $y = ax_i + b$.

(a، b المعاملات المقدرة لـ a و b)

لإيجاد هذه القيمة الصغرى، نقوم بالاشتقاق بالنسبة لـ a و b، ومع شرط المشتق الثاني يكون سالبا.

$$\sum\left(y_i - y\right)^2 = \sum\left(y_i - ax_i - b\right)^2$$

نقوم بالاشتقاق بالنسبة لـ a:

$$a : \frac{\partial \sum\left(y_i - ax_i - b\right)^2}{\partial a} = 0$$

$$\Rightarrow \sum 2\left(-x_i\right)\left(y_i - ax_i - b\right) = 0$$

$$\Rightarrow -2\sum x_i\left(y_i - ax_i - b\right) = 0 \cdots\cdots\cdots ①$$

نقوم بالاشتقاق بالنسبة لـ b:

$$b : \frac{\partial \sum\left(y_i - ax_i - b\right)^2}{\partial b} = 0$$

$$\Rightarrow \sum 2\left(-1\right)\left(y_i - bx_i - b\right) = 0$$

$$\Rightarrow -2\sum\left(y_i - bx_i - b\right) = 0 \cdots\cdots\cdots ②$$

$$\begin{cases} \cancel{-2}\sum x_i \left(y_i - \widehat{ax_i} - \widehat{b}\right) = 0 \\ \cancel{-2}\sum \left(y_i - \widehat{ax_i} - \widehat{b}\right) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \sum \left(x_i y_i - a x_i^2 - b x_i\right) = 0 \\ \sum \left(y_i - a x_i - b\right) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \sum x_i y_i - a \sum x_i^2 - b \sum x_i = 0 \cdots\cdots ① \\ \sum y_i - a \sum x_i - nb = 0 \cdots\cdots\cdots ② \end{cases}$$

من ②:

$$nb = \sum y_i - a \sum x_i$$

$$b = \frac{\sum y_i - a \sum x_i}{n}$$

$$\boxed{b = \overline{y} - a\overline{x}} \cdots\cdots\cdots ③$$

نعوض ③ في ①:

$$\sum x_i y_i - a\sum x_i^2 - \left(\overline{y} - a\overline{x}\right)\sum x_i = 0$$

$$\sum x_i y_i - a\sum x_i^2 - \overline{y}\sum x_i + a\overline{x}\sum x_i = 0$$

$$a\left(\overline{x}\sum x_i - \sum x_i^2\right) = \overline{y}\sum x_i - \sum x_i y_i$$

$$a = \frac{\overline{y}\sum x_i - \sum x_i y_i}{\overline{x}\sum x_i - \sum x_i^2}$$

$$a = \frac{\sum x_i y_i - \overline{y}\sum x_i}{\sum x_i^2 - \overline{x}\sum x_i}$$

$$a = \frac{\frac{1}{N}\left(\sum x_i y_i - \overline{y}\sum x_i\right)}{\frac{1}{N}\left(\sum x_i^2 - \overline{x}\sum x_i\right)}$$

$$a = \frac{\frac{1}{N}\sum x_i y_i - \frac{1}{N}\overline{y}\sum x_i}{\frac{1}{N}\sum x_i^2 - \frac{1}{N}\overline{x}\sum x_i}$$

$$a = \frac{\frac{1}{N}\sum x_i y_i - \overline{x}\,\overline{y}}{\frac{1}{N}\sum x_i^2 - \overline{x}^2}$$

$$\boxed{a = \frac{\operatorname{cov}(x, y)}{v(x)}}$$

. معامل الارتباط (*Coefficient de régression*):

ويسمى أيضا معامل الارتباط الخطي، ويقيس العلاقة الخطية أو القريبة من الخطية، ونرمز له بـ r أو $r_{(x,y)}$. وقيمته محصورة بين 1− و 1+.أي $-1 \le r \le +1$

عند:

$r = 0 \Leftarrow$ **استقلالية** بين x و y.

$r < 0 \Leftarrow$ توجد **علاقة خطية سالبة** أو عكسية بين x و y.

$r > 0 \Leftarrow$ توجد **علاقة طردية** أو موجبة بين x و y.

- كلما اقتربت قيمة r من **(1–) أو (2+)**، يدل على **وجود علاقة قوية عكسية أو طردية** بين x و y.
- كلما اقترب من **0**، يدل على **وجود علاقة ضعيفة عكسية أو طردية** بين x و y.
- دون ذلك العلاقة عكسية أو طردية نسبية بين x و y.

ويعطى r من خلال العلاقة:

$$\boxed{r = \frac{\operatorname{cov}(x, y)}{\sigma_x, \sigma_y}}$$

يمكن حساب نسبة التأثير $\frac{x}{y}$ بحساب ما يسمى معامل التحديد R.

حيث: $R = r^2$

حيث: $0 \le R \le 1$

تطبيق:

إذا كان $\overline{x} = 3,5 \quad , \quad v(x) = 2,917$

$\overline{y} = 5 \quad , \quad v(y) = 2,78$

ولدينا

$\operatorname{cov}(x, y) = 2,83$

حدّد معادلة خط الانحدار y بدلالة x ومعامل الارتباط، وفسّر درجة التأثير.

معادلة الانحدار y مقدرة من خلال y:

$$y = ax + b$$
$$b = \overline{y} - a\overline{x}$$
$$\overline{b} = 5 - 3,5a = 5 - 3,5(0,97)$$
$$a = \frac{\cos(x, y)}{v(x)} = \frac{2,83}{2,78}$$
$$\boxed{a = 0,97}$$

ومنه:

$$\boxed{b = 1,605}$$

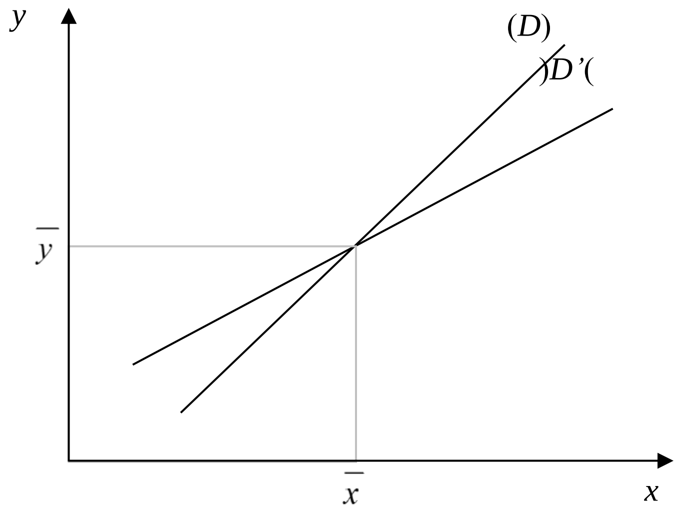
$$y = 0,97x - 1,605$$
$$r = \frac{\operatorname{cov}(x, y)}{\sigma_x \cdot \sigma_x} = \frac{2,83}{1,71 \times 1,67}$$
$$r = 0,99$$

هناك علاقة قوية طردية بين x و y.

$$R = r^2 = (0,99)^2$$
$$R = 0,98$$

y تفسر بـ x بدرجة 98%.

تفسير x لِـ y.



y: المتغير التابع.

x: المتغير المستقل.

Y يفسر بـ x بنسبة ما

.

$$y = f(x) = \begin{cases} y_i = ax_i + b' \\ D: y_i = ax_i + b \end{cases}$$

$$x = f(y) = \begin{cases} x_i = a'y_i + b' \\ D': x_i = a'y_i + b' \end{cases}$$

عند تقرير خط الانحدار (D')، نقوم بتقدير المعادلة $x_i = a'y_i + b'$.

وتقدم من خلال: $x_i = a'y_i + b'$.

حيث:

$$b = \overline{y} - a\overline{x}$$
$$b' = \overline{x} - a'\overline{y}$$

$$a = \frac{\mathrm{cov}(x, y)}{v(x)}$$ و

$$a' = \frac{\mathrm{cov}(x, y)}{v(y)}$$

a و a' هما ميلا D و D'.

ويمكن قياس الزاوية المكونة، والتي نحصل عليها من جراء تقاطع D و D'.

- كلما كانت الزاوية كبيرة أو منفرجة، كانت العلاقة بين x و y ضعيفة، كلما كانت الزاوية صغيرة أو حادة كانت العلاقة بين x و y قوية.

- إذا كانت x و y مستقلتان، فإن $D \perp D'$ ومنه $a \times a' = 0$.
- إذا كانت x و y مرتبطتان بشدة، حيث x يفسر بالضرورة y والعكس صحيح، فإن D و D' متطابقين و $a \times a' = a \times \frac{1}{a} = 1$

.

ومنه يمكن تعميم العلاقة إلى:

$$0 \leq a \times a' \leq 1$$
$$0 \leq R = r^2 \leq 1$$
$$\Downarrow$$
$$r^2 = a \times a'$$
$$\Downarrow$$
$$r = \pm\sqrt{a \times a'}$$
$$-1 \leq r \leq 1$$

مثال: تطبيق المثال السابق على x:

$$x_i = a' y_i + b'$$
$$b' = \overline{x} - a'\overline{y} = 3,5 - \overline{a}' \cdot 5$$
$$a' = \frac{\operatorname{cov}(x, y)}{v(y)} = \frac{2,83}{2,78} = 1,018$$
$$\Longrightarrow \ b' = 3,5 - (1,018) \times 5 = -1,59$$
$$x_i = 1,018 y_i - 1,59$$
$$r^2 = a \cdot a' = 0,97 \times 1,018$$
$$R = r^2 = 0,98$$

$$0 \le R \le 1$$

$$r = \sqrt{0,98}$$
$$r = 0,99$$

$$-1 \le r \le 1$$

لإعطاء الرسم البياني لـ D و D' على نفس المعلم:

يجب كتابة y حيث:

$$y_i = \frac{1}{a'} x_i - \frac{b'}{a'}$$
$$y_i = \frac{1}{1,018} x_i - \left(\frac{-1,59}{1,018} \right)$$
$$y_i = 0,98 x_i + 1,56$$

ملاحظة حول الجدول التقاطعي:

- عناصر الجدول:

التكرارات الجزئية:

تقرأ داخل الجدول وترمز لها بـ n_{ij} ، وهي تعبّر عن عدد الأفراد الذين لديهم في نفس الوقت البعد x_i والبعد y_j .

التكرارات النسبية الإجمالية:

أ – نسبة التكرارات الجزئية على التكرارات الإجمالية وتكتب:

$$f_{ij} = \frac{n_{ij}}{n_{\square\square}}$$

ب – التكرارات النسبية الهامشية: نسبة التكرارات الهامشية على التكرارات الإجمالية.

إذا اعتبرنا j يتغير: $$f_{\square j} = \frac{n_{\square j}}{n_{\square\square}}$$

j يتغير: $$f_{i\,\square} = \frac{n_{i\,\square}}{n_{\square\square}}$$

التكرارات النسبية الشرطية:

ونرمز لها بـ $f_{i/j}$ أو $f_{j/i}$ وتقرأ (f_i علما j).

$$f_{j/i} = \frac{n_{ij}}{n_{i\cdot}} \qquad\qquad f_{i/j} = \frac{n_{ij}}{n_{\cdot j}}$$

التكرارات الهامشية:

هي التكرارات الموجودة على هامشي الجدول، إما المطلقة أو النسبية.

- بعض الخاصيات:

التوزيعات الهامشية:

$$\bullet \sum_{j=1}^{k} f_{\cdot j} = \sum_{j=1}^{k} \frac{n_{\cdot j}}{n_{\cdot\cdot}} = \frac{1}{n_{\cdot\cdot}} \sum_{j=1}^{k} n_{\cdot j} = \frac{1}{\cancel{n_{\cdot\cdot}}} \times \cancel{n_{\cdot\cdot}} = 1$$

$$\bullet \sum_{i=1}^{p} f_{i\cdot} = \sum_{i=1}^{p} \frac{n_{i\cdot}}{n_{\cdot\cdot}} = \frac{1}{n_{\cdot\cdot}} \sum_{i=1}^{p} n_{i\cdot} = \frac{1}{n_{\cdot\cdot}} \times n_{\cdot\cdot} = 1$$

$$\Rightarrow \sum_{i=1}^{p} f_{i\cdot} = \sum_{j=1}^{k} f_{\cdot j} = 1$$

التوزيعات الشرطية:

$$\bullet \sum_{i=1}^{p} f_{i/j} = \sum_{i=1}^{p} \frac{n_{ij}}{n_{\cdot j}} = \frac{1}{n_{\cdot j}} \sum_{i=1}^{p} n_{ij} = \frac{1}{n_{\cdot j}} \times n_{\cdot j} = 1$$

$$\bullet \sum_{j=1}^{k} f_{j/i} = \sum_{j=1}^{k} \frac{n_{ij}}{n_{i\cdot}} = \frac{1}{n_{i\cdot}} \sum_{j=1}^{k} n_{ij} = \frac{1}{n_{i\cdot}} \times n_{i\cdot} = 1$$

$$\Rightarrow \sum_{1}^{p} f_{i/j} = \sum_{1}^{k} f_{j/i} = 1$$

- **العلاقة بين التكرارات النسبية:**

$$* \; f_{i\cdot} \times f_{j/i} = \frac{\cancel{n_{i\cdot}}}{n_{\cdot\cdot}} \times \frac{n_{ij}}{\cancel{n_{i\cdot}}} = \frac{n_{ij}}{n_{\cdot\cdot}} = f_{ij}$$

$$* \; f_{\cdot j} \times f_{i/j} = \frac{\cancel{n_{\cdot j}}}{n_{\cdot\cdot}} \times \frac{n_{ij}}{\cancel{n_{\cdot j}}} = \frac{n_{ij}}{n_{\cdot\cdot}} = f_{ij}$$

$$\Longrightarrow \; f_{i\cdot} \times f_{j/i} = f_{\cdot j} \times f_{i/j} = f_{ij}$$

- **العلاقات غير الخطية:**

يمكن تطبيق طريقة المربعات الصغرى **MCO** على النماذج غير الخطية إلى العلاقة غير الخطية إلى العلاقة الخطية بين المتغيرتين، وذلك بالاستعانة بالنماذج اللوغاريتمية أو النماذج النصف لوغاريتمية.

النماذج نصف لوغاريتمية:

يتنقل من الشكل الأسي إلى الشكل نصف لوغاريتمي، إذا كان لدينا علاقة من الشكل: $y = \alpha\beta^{x}$.لتقدير مَعْلمات (Paramètres) α و β ، ننتقل أولا إلى الشكل الخطي لاستخدام صيغة اللوغاريتم على هذا النموذج، فنحصل على:

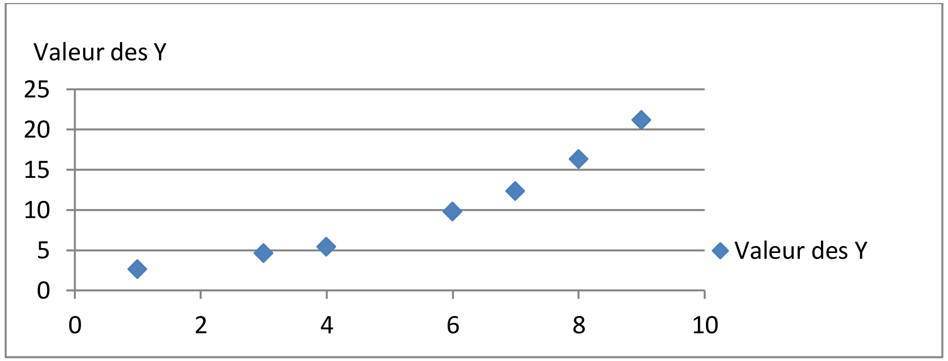
$$\log y = \log \alpha + x \log \beta$$

والقيم الأصلية تصبح قيم لوغاريتمية.
وبعد تقدير α و β ، نرجع إلى النموذج الأصلي ونضع قيم لهذه المَعْلمات.

مثال: ليكن لدينا الجدول التالي:

x	1	3	4	6	7	8	9
y	2,6	4,6	5,4	9,8	12,3	16,3	21,2

1) تحديد العلاقة بين x و y.



2) تقدير مَعْلَمَات النموذج α و β .

نلاحظ من خلال الشكل أن العلاقة بين x و y هي من الشكل $y = \alpha\beta^{x}$.
و هو نموذج أسّي يمكن إرجاعه إلى نموذج نصف لوغاريتمي، فنحصل على:

$$\log y = \log \alpha + x \log \beta$$

x	1	3	4	6	7	8	9
y	2,6	4,6	5,4	9,8	12,3	16,3	21,2

Log *y*	0,414	0,662	0,732	0,991	1,089	1,212	1,326

إذا وضعنا:

$$\log y = z$$
$$\log \alpha = b$$
$$\log \beta = a$$

$$z = ax + b$$

نقدر $\begin{cases} \hat{a} \\ \hat{b} \end{cases}$ بـ a و بـ b حيث: $a = \dfrac{\text{cov}(x, z)}{v(x)}$

$$b = \overline{z} - a\overline{x}$$

عندما نحصل على a و b، نرجع إلى $\log\alpha$ و $\log\beta$ ونستخرج قيم α و β .يمكن استخدام هذا النموذج للتنبؤ في المستقبل إما:

أن يكون *x* معلوما فنتنبأ بقيمة *y*.
أو أن تكون *y* معلوما فنتنبأ بقيمة *x*.

مثلا: $x = 11$ ما هي قيمة *y*.

النماذج اللوغاريتمية:

في هذه الحالة، نأخذ اللوغاريتمات للمتغير التابع *y* والمتغير المستقل *x*، ويمكن أن نحصل على علاقة من الشكل:

$$y = \frac{\alpha}{x^{\beta}} = \alpha x^{\beta}$$

وننتقل من الصيغة غير الخطية إلى الخطية، بإدخال log على العلاقة فنحصل على:

$$\log y = \log \alpha + \log x^{-\beta}$$
$$\log y = \log \alpha + (-\beta) \log x$$

نضع:

$$-\beta = a$$
$$\log x = T$$
$$\log \alpha = b$$
$$\log y = z$$

ونحصل على: $z = b + aT$

نقدّر كل من a بـ a و b بـ b.

ثم نرجع إلى الصيغة log، فنحصل على قيمة α و β، والقيم المتحصل عليها نضعها في النموذج الأصلي، ويمكن لنا أن نتنبأ

بالقيم المستقبلية.

مثال: إذا كان لدينا:

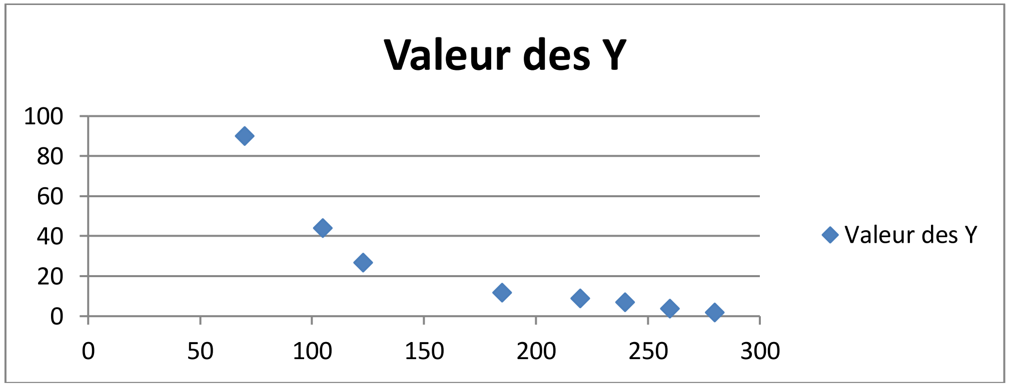
x	70	105	123	185	220	240	260	280
y	90	44	27	12	9	7	4	2

إذا كان

x: الكميات المباعة.

y: أسعار هذه الكميات

1) تحديد شكل العلاقة.



تقدير معلمات النموذج.

قدّر y عندما $x = 60$.

العلاقة:

$$y = \frac{\alpha}{x^{\beta}} = \alpha x^{\beta}$$