نبذة تاريخية عن الاحصاء:

ارتبط مصطلح الاحصاء في اذهان الناس بالتعداد و الاحصاءات التي تجريها الدولة او الافراد لعد السكان، متطلبات التمويل و الادارة او حصر الثروة. فقد ظهر مصطلح الاحصاء قبل الفي سنة في الحضارة المصرية عند عد القبائل المنضوية تحت سيطرتها كذا عند الصينيين و البابليين في عد منتجاتهم الزراعية، اما في الحضارة الاسلامية فقد كان الخليفة الراشد عمر الفاروق يقدر عدد المتطوعين في الثغور بعدد ارغفة الخبز المستهلكة، وكان الخليفة عثمان بن عفان –رضي الله عنه– اول من استخدم التدوين لإحصاء المستفيدين من عطايا بيت المال.

سمي علم الاحصاء قديما علم الدولة فقد اعتبر بعض المؤرخين ان كلمة statistique مشتقة من الكلمة اللاتينية state بينما عزاها اخرون الى الكلمة الاغريقية statika وتعني الحساب و العد. وفي اوربا ترجع اول عملية تعداد مسجلة للأراضي سنة 1086 ببريطانيا، وانطلاقا من القرن 14 اصبح شائعا انشاء دواوين احصائيات السكان ، الضرائب و المنتجات الزراعية و الصناعية...اما في القرن 17 فقد برز Pascale ، Fermat و Bernoulli في دراسة هندسة العاب الحظ و يعتبر القرن 19 الفترة التي شهدت الانطلاقة العلمية للإحصاء الرياضي على يد Laplace وتابعه علماء اخرون في البحث و تطوير الطرق و الاساليب مثل Markov ، Pearson و Fisher.

وفي وقتنا الحاضر يرتبط تطور الاحصاء الرياضي كباقي فروع الرياضيات بمتطلبات التطبيقات المتعددة، فأصبحت طرق الاحصاء تستخدم بشكل واسع في كل الميادين ويلعب دورا هاما في البحوث التقنية، الاقتصادية، البيولوجية، الزراعية، الطبية والاجتماعية... مما يتطلب دراسة مستفيضة للإحصاء الرياضي ومختلف طرقه و اساليبه فهو يقيم بنية النموذج الاحصائي المعد لدراسة و تحليل التجارب الاحتمالية بملاحظة البيانات الاحصائية.

1– مفاهيم أساسية:

1–1 – الإحصاء: هو العلم الذي يهتم بالطرق الرياضية لجمع، وترتيب، وعرض، وتحليل البيانات ومن ثم استخدام المعطيات لاستخلاص نتائج و تعميمها على المجتمع –خاصة في حالات عدم اليقين – في مختلف العلوم من اجل اتخاذ القرارات المناسبة.

1–1–1– أقسام الإحصاء: ينقسم الاحصاء الى قسمين حسب التعريف السابق.

1-1-1-1- **الاحصاء الوصفي** :ويقصد به العلم الذي يختص بجمع البيانات و تبويبها وعرضها في شكل جداول احصائية او اشكال بيانية ثم تلخيصها ببعض المؤشرات الاحصائية الدالة على طبيعتها.

1-1-1-2- **الاحصاء الاستدلالي**: و هو يختص بتعميم النتائج على المجتمع من خلال فحص عينة ممثلة عنه ويشمل نظريتي التقدير و اختبار الفرضيات.

– **التوقع (التنبؤ)**: وفيه يتم استخدام نتائج الاستدلال الاحصائي لتوقع سلوك الظاهرة في المستقبل، ومن اهم الاساليب المتبعة في التنبؤ اسلوب الاتجاه العام.

1–1–2– الطريقة العلمية في البحوث الإحصائية: وتعتمد على النقاط التالية:

– تحديد اشكالية البحث والهدف و فرضيات الدراسة.

– تحديد المجتمع، الاطار، طريقة المعاينة و حجم العينة المناسبة رياضيا.

- جمع، تصنيف و عرض البيانات.
- حساب الوسائط – المؤشرات- كتقديرات لمعالم المجتمع.
- تحليل المعطيات.
- تفسير النتائج و اتخاذ القرار.

1-1-3- تصميم البحث الاحصائي:

- تحديد الغرض.
- تحديد امكانية التوقع.
- تحديد اطار البحث.

1-1-4- مصادر جمع البيانات: بعد تحديد الهدف من الدراسة لا بد من جمع بيانات عن الظاهرة من مصادرها:

1-1-4-1- **مصدر داخلي**: و تأخذ من السجلات و المستندات و الدراسات السابقة.

1-1-4-2- **مصدر خارجي**: ويتم تجميع البيانات من خلال مشاريع مشابهة او من مفردات المجتمع مباشرة عن طريق:

- المقابلة الشخصية وتسجيل الاجابات على الاستبيان او الاستمارة الخاصة.
- الاتصال الهاتفي، بالبريد العادي أو الالكتروني، وسائل التواصل الاجتماعي .
- المراقبة ، الملاحظة و المشاهدة أو التسجيل.
- التجربة أو العمل الميداني.

1-1-5- طرق جمع البيانات: يوجد أسلوبين لجمع المعلومات:

1-1-5-1- **المسح الشامل**: تعني دراسة شاملة لجميع افراد المجتمع

1-1-5-2- **الـعيـــنـــــة**: هي عبارة عن مجموعة من مفردات المجتمع يتم اختيارها لتكون ممثلة جيدا له.

المعاينة: وهي علم و فن التحكم و قياس دقة المعلومات الاحصائية باستخدام النظريات الرياضية ويفضل إستخدام أسلوب العينات لأسباب عديدة أهمها:

- تقليل الجهد و التكلفة و الوقت.
- صعوبة حصر أفراد المجتمع
- محدودية الامكانيات كنقص الخبرة و التخصص والإمكانيات اللوجستية.

1-1-6- أنواع العينات: يعتبر اختيار العينة من أهم خطوات البحث الاحصائي ويمكن أن تكون مستمرة –ثابتة- أو متغيرة، و يمكن تمييز نوعين من العينات:

1-1-6-1- **العينات العشوائية**: وتسمى ايضا العينات الاحتمالية، ويتم اختيار مفردات العينة بحيث يكون لكل مفردة نفس الحظ في الظهور . ومن أهم أنواعها :

- **العينة العشوائية البسيطة:** وتستخدم في حالة تماثل أو تجانس أفراد المجتمع، ويتم سحب المفردات باستخدام جداول الأرقام العشوائية أو بالقرعة بارجاع او دون إرجاع. فإذا أردنا اختيار n وحدة من N وحدة من المجتمع فإن عدد العينات التي يمكن اختيارها من المجتمع يساوي C_N^n.
- **العينة العشوائية المنتظمة:** و تستخدم في حالة صعوبة الحصول على إطار للمجتمع أو عدم جودته. وعادة يتم اختيار أول وحدة عشوائيا ، ثم نحدد باقي الوحدات المتتالية ذي بعد متساوي قدره k، أي $r,r+k,r+2k,\ldots,r+(n-1)k$ حيث $(1\leq r\leq k)$ و $k\leq N/n$.
- **العينة العشوائية الطبقية:** وهي اكثر الطرق استعمالا، وتقوم على تقسيم المجتمع إلى طبقات – مجموعات- متجانسة غير متداخلة، ويتم اختيار وحدات المعاينة عشوائيا من كل طبقة، ومن ميزاتها الحصول على تقدير اكثر دقة و أقل تباينا. أي $N=\sum_{i=1}^{L}N_i\,,n=\sum_{i=1}^{L}n_i\,,W_i=\frac{N_i}{N}$
- **العينة العشوائية العنقودية:** و تعتمد على تقسيم المجتمع إلى قطاعات – عناقيد- متشابهة غير متداخلة، ثم اختيار عدد من هذه العناقيد كعينة عشوائية بسيطة يتم دراسة جميع أفراد كل منها، يمكن ان تكون العينة العنقودية مكونة من مرحلة واحدة أو من عدة مراحل.

1-1-6-2- **العينات غير العشوائية:** وهي التي يتم إختيار مفرداتها بطريقة غير عشوائية، بحيث تختار بالطريقة التي تحقق الهدف، ومن أهم أنواع العينات غير الإحتمالية : **العينة العمدية و العينة الحصصية.**

1-2- مصطلحات إحصائية :

1-2-1- المجتمع الإحصائي: وهو المجموعة التي تشكل أهمية للدراسة ويكون اما محدودا أو غير محدود أو مجموعة المشاهدات التي تخص ظاهرة وينقسم الى نوعين: المجتمع الهدف و مجتمع الدراسة، مثلا : مجموعة طلبة جامعة الوادي.

1-2-2- العينة: وهي كل مجموعة جزئية غير خالية من المجتمع الإحصائي، مثلا : 100 طالب من طلبة الجامعة.

1-2-3- الوحدة الإحصائية: وهو كل عنصر – فرد أو مفردة – من المجتمع الاحصائي، مثلا: طالب واحد من الجامعة.

1-2-4- الطبع الإحصائي: و تسمى الميزة الإحصائية أو الظاهرة الإحصائية، وهي الخاصية أو الصفة المدروسة أو الملاحظة على أفراد المجتمع، وهي نوعين:

1-2-4-1- **ميزة نوعية:** وهي الصفة غير القابلة للقياس وتنقسم الى نوعين: **ترتيبية** مثل مستوى الطالب و **غير ترتيبية** مثل الجنس ، اللون، الحالة الاجتماعية ... .

1-2-4-2- **ميزة كمية:** وهي الصفة القابلة للقياس مثل العمر، الوزن ... .

1-2-5- المتغير الإحصائي: نسمي قيم ظاهرة إحصائية متغيرا إحصائيا ويصنف إلى قسمين:

1-2-5-1- **المتغير الإحصائي المتقطع** :– المنفصل – وهو الذي يأخذ قيما معزولة مثل عدد الأخوة... .

1-2-5-2- **المتغير الإحصائي المستمر** : وهو يأخذ قيم في مجال $[a,b]$، يسمى كل مجال من الشكل $[a_k,a_{k+1}[$ فئة. نسمي العدد $L=a_{k+1}-a_k$ طول أو سعة –مدى- الفئة و العدد $c=(a_{k+1}+a_k)/2$ مركز الفئة.

1-2-6- السلسلة الإحصائية: نسمي نتائج ظاهرة إحصائية معينة بالسلسلة الاحصائية أي عبارة عن مجموعة الثنائيات (x_i,n_i) حيث x_i القيم و n_i عدد المرات – التكرارات، المشاهدات – التي تكرر فيها x_i .

1-2-7- التكرار النسبي: هو نسبة عدد التكرارات لقيم الطبع الإحصائي إلى العدد الإجمالي للتكرارات ونرمز له بالرمز f_i أي أن:

$$f_i = \frac{n_i}{N}; N = \sum_i n_i$$

ملاحظة: $f_i \in [0,1]$ و $\sum_i f_i = 1$

1-2-8- الجدول التكراري: وهي جداول تنظم و ترتب قيم المتغير الإحصائي، و تكون من الشكل:

القيمx_i	x_1	…					
التكرار المطلقn_i	n_1	…					
النسبة المئويةf_i 100	%						

و نميز بين الجداول التكرارية البسيطة ،جداول التوزيع التكرارية المزدوجة و الجداول التكرارية المفتوحة.

1-2-9- التكرار المجمع: وهو عدد المشاهدات التي تقل او تزيد عن قيمة معينة، ويمكن التفرقة بين نوعين:

1-2-9-1- **التكرار المتجمع الصاعد**: هو مجموع التكرارات للقيمة وللقيم الأصغر منها ونرمز له بـ n_i ↖

– هو مجموع عدد الوحدات التي تقل عن الحد الأعلى للفئة المقابلة.

$$\widetilde{n_k} = \sum_{i=1}^{k} n_i$$

1-2-9-2- **التكرار المتجمع النازل**: هو مجموع التكرارات للقيمة وللقيم الأكبر منها ونرمز له بـ n_i ↘

– هو مجموع عدد الوحدات التي تزيد عن الحد الأدنى للفئة المقابلة.

مثال: مصنع نسيج عند تجريب الة جديدة قام بإحصاء الاخطاء في 75 عينة قماش من 10 متر ، فكانت النتائج الاتية :

x_i	0	1	2	3	4	5
n_i	38	15	11	6	3	2

1- حدد المجتمع الاحصائي ، المتغير الاحصائي.

2- حدد التكرار النسبي ، التكرار المتجمع الصاعد.

الحل:

1- **المجتمع الإحصائي** قماش المصنع ذو 10 متر ، **الميزة الإحصائية** عدد الأخطاء في القطعة كمي منفصل.

2-

x_i	0	1	2	3	4	5
n_i	38	15	11	6	3	2
f_i%	50.6	20	14.7	8	4	2.7
↖ n_i	38	53	64	70	73	75

1-3- تبويب البيانات في جدول توزيع تكراري متصل: ان القاعدة العامة لإفراغ البيانات في جدول كمي مستمر هي ان يكون عدد الفئات N_s يتراوح بين 5 و **15** ، إلا ان هذا يعود للإحصائي في تحديد العدد و طول الفئة، رغم وجود طرق تحدد ذلك منها:

1-3-1- القاعدة البسيطة: $\boldsymbol{N_s = \sqrt{N}}$ و $\boldsymbol{L = \frac{x_{max}-x_{min}}{N_s}}$ حيث N عدد القيم و L طول الفئة.

1-3-2- قاعدة ستــورج: $\boldsymbol{N_s = 1 + 3.322log(N)}$

1-3-3- قاعدة يـــول: $N_s = 2.25\sqrt[4]{N}$

مثال: سجلت مدة نقل الطلبة من مقر سكناهم بالدقائق فكانت:

5-3-20-45-45-5-25-15-15-5-20-5-5-2-45-10-30-2-7-15-25-30-60-35-30-45-30-30-25-25-45-5-70-
60-35-35-18-5-5-65-10-50-25-30-2-28-30-25-20-45.

1- شكل القيم في جدول احصائي بفئات-اعتمد طريقة ستورج-

الحل: نحسب: $\mathbf{N_s = 1 + 3.322log50 \simeq 7}$ وعليه $\boldsymbol{L = \frac{70-2}{7} \cong 10}$ فتكون الفئات على الشكل: $\boldsymbol{[2;12[, [12;22[...}$

التمثيـــلات البيانيـــــة

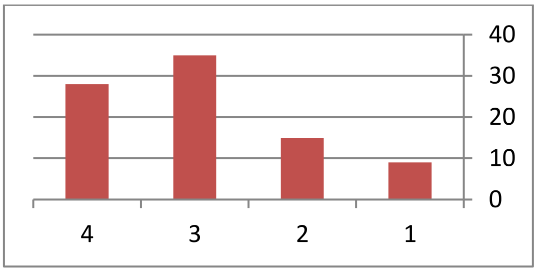
تمـــهيـــد: إن الهدف الأساسي لعرض المعلومات بواسطة المعطيات تسمح باستيعاب مجموعة العناصر المعروضة بشكل سريع و إظهار بعض الخصائص الهامة، رغم اهماله بالضرورة بعض الجزئيات.و الرسوم البيانية وسيلة لمراقبة التغيرات الشاذة و الانكسارات و التقطعات في مسار ظاهرة ما.

يقوم المبدأ الأساسي للتمثيل البياني على استخدام الاحداثيات الديكارتية، وهي بصورة عامة المستطيلات التي تسمح في علم الجبر بتمثيل التوابع على مستقيمين يتقاطعان في O و مرسومين على سطح مستو، و يختار لذلك اتجاه موجب و وحدة بيانية، وتمثل القيمة x_i للصفة على المحور الأفقي.

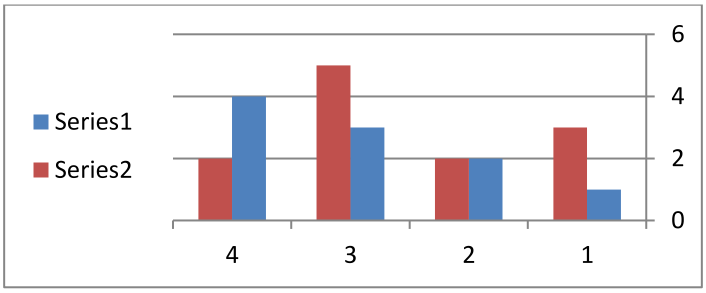
2-1- التمثيل البياني للتوزيعات التكرارية: يمكن تمثيل التوزيعات التكرارية بطرق مختلفة أهمها:

2-1- الأعمدة البيانية : ويسمى المخطط بالأعمدة، و نميز الاشكال التالية:

2-1-1- الأعمدة البسيطة: وهي أعمدة يتناسب طولها مع قيمة البيانات التي تمثلها.

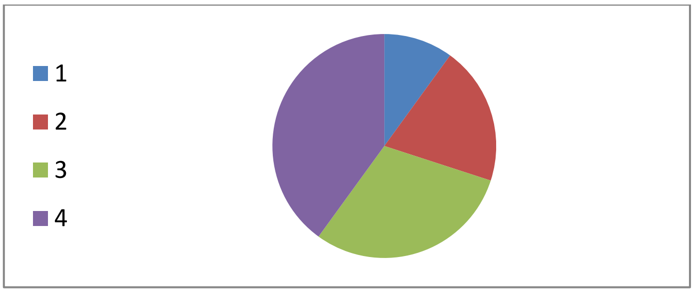


2-1-2- الأعمدة المزدوجة: وتستخدم للمقارنة بين ظاهرتين أو أكثر

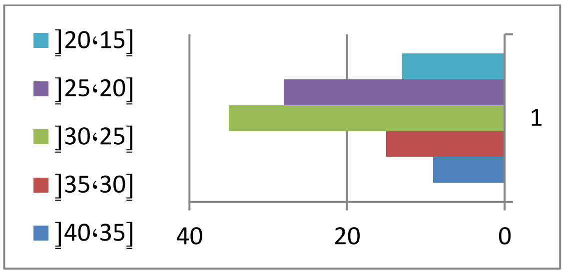


2-2- الدائرة النسبية البيانية: و تستعمل خاصة في المتغيرات النوعية و يتم تخصيص لكل صفة قطاع، تحسب زاوية القطاع α كمايلي:

$$\alpha = f_i \times 2\pi$$



2-3- المدرج التكراري: عندما يكون المتغير المدروس مستمرا فإن كل قيمة م مدى السلسلة يمكن ان يكون قيمة للصفة، لذا نرفق لكل فئة مجال على محور الفواصل ونمثل كل ثنائية (فئة،تكرار) بمستطيل مساحته تتناسب مع تكرار الفئة.



2-3-1- تعترض مشكلة رسم المدرج في حالة اطوال الفئات غير متساوية ، وفي هذه الحالة نمثل الاطوال بالتكرار المعدل n'_i وفق القاعدة التالية:

$$n_i' = \frac{l \times n_i}{l_i}$$

حيث l أصغر طول فئة .

مثال : أجريت تجربة على 100 مصباح لمعرفة مدى صلاحيتها فكانت النتائج التالية:

الفئات	[200,300[	[300,600[	[600,800[
n_i	15	45	40
l_i	100	300	200
n_i'	15	15	20

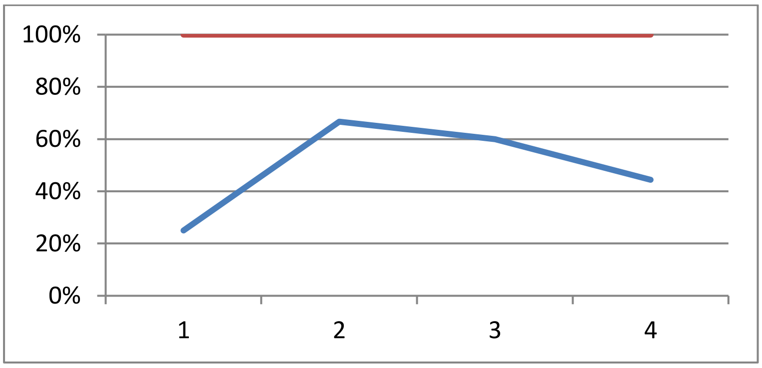
2-3-2- أشكال التوزيعات و المدرجات التكرارية: نصادف في الحياة العملية عدة أشكال أساسية بسيطة من التوزيعات منها:

2-3-2-1- التوزيعات المتناظرة: تتناقص فيها التكرارات نحو الصفر بشكل نظامي قبل و بعد القيمة المركزية الأعلى.

2-3-2-2- التوزيعات غير المتناظرة

2-3-2-3- التوزيعات على شكل حرف U, L ...

2-4- المنحنيات التكرارية: ويتم بتحديد نقطة لكل مشاهدة ومن ثم نصل هذه النقاط باليد أي تعليم النقاط (x_i,n_i)- x_i القيمة أو مركز الفئة - إذا اوصلنا بين النقاط بالمسطرة تشكل لدينا المضلع التكراري.



2-4-1- المنحنيات التكرارية التجميعية: ويتم رسم المنحنى التكراري المتجمع الصاعد – النازل – لمتغير مستمر بايصال مجموعة النقط التي احداثياتها الحدود العليا للفئات – الحدود الدنيا – والتكرارات المتجمعة الصاعدة المقابلة لها – التكرارات النازلة – على الترتيب .

2-4-2- الدرج التكراري: ويستخدم لرسم المنحنى التكراري المتجمع الصاعد – النازل – لمتغير منفصل .

مقاييس النزعة المركزية – مقاييس الموقع–

تـمهيـــد: من المهام الأساسية للإحصاء الوصفي اختصار القيم الكثيرة إلى بضع معلومات تعبر إلى أبعد حد عن السلسلة الاحصائية المدروسة، فالتمثيل البياني يزودنا بالملخص الأول، ونرغب الآن في تمييز السلسلة الاحصائية بعدد او قيمة معيارية تمثل بدرجة كبيرة مجموعة عناصر السلسلة .

لذا يجب أن تحقق هذه القيمة بعض الشروط حددها ***يول*** في:

- أن تعين بشكل موضوعي و مستقل عن الرأي الشخصي للإحصائي.
- أن تتعلق بكل معطيات السلسلة. – أن تكون سهلة الحساب.
- أن تكون قليلة التأثر بتقلبات العينة – أن يكون لها معنى مادي.

مقاييس النزعة المركزية هي المؤشرات التي تعبر عن القيمة التي تتمركز حولها اغلب القياسات المعطاة منها:

3–1– المتوسط الحسابي: هو أكثر المقاييس إستخداما في الجانب التطبيقي، ويعد مركز التوازن للظاهرة.

3–1–1– حساب الوسط الحسابي :

3–1–1–1– الطريقة المباشرة:

أ– **في حالة البيانات غير المبوبة:** يعرف المتوسط الحسابي لسلسلة من n مفردة بالعدد الذي نرمز له ب $\overline{X}$ الذي يحقق:

$$\overline{\mathbf{X}} = \frac{\sum_{i=1}^{n} x_i}{n}$$

ب–**في حالة البيانات المبوبة:** في اغلب الحالات لا تكون للقيم نفس الأهمية بل لكل قيمة عامل ترجيح خاص، عندئذ يسمى المتوسط الحسابي بالمتوسط الحسابي المرجع و يحسب كمايلي:

–**المتغير الاحصائي منفصل:** إذا كانت x_1، x_2،...، x_n قيم الظاهرة X وكانت n_1، n_2، ...، n_n تكرارتها على الترتيب فإن:

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

–**المتغير الاحصائي مستمر:** إذا كانت c_1، c_2، ...، c_n مركز فئات الظاهرة X وكانت n_1، n_2، ...، n_n تكرارتها على الترتيب فإن:

$$\overline{\mathbf{X}} = \frac{\sum_{i=1}^{n} n_i c_i}{\sum_{i=1}^{n} n_i}$$

مثال: تحصل طالب على العلامات التالية 8–11–14–15–16 بمعاملات 3–1–1–5–2 على الترتيب:

$$\overline{X} = \frac{8 \times 3 + 11 \times 1 + 14 \times 1 + 15 \times 5 + 16 \times 2}{3 + 1 + 1 + 5 + 2} = 13$$

3-1-2- الطريقة غير المباشرة: وتسمى طريقة الوسط الفرضي، وتعتمد على استبدال المتغير الاساسي بمتغير جديد

يفضل اختيار الوسط *الفرضي* x_0 القيمة الوسيطية ويستحسن ان يقابلها أكبر تكرار.ففي حالة المتغير الاحصائي المتقطع:

$$\overline{X} = x_0 + \frac{\sum_{i=1}^{n} n_i(x_i - x_0)}{\sum_{i=1}^{n} n_i}$$

- اما في حالة المتغير المستمر: $\overline{X} = x_0 + a\frac{\sum_{i=1}^{n} n_i\frac{(c_i - x_0)}{a}}{\sum_{i=1}^{n} n_i}$ حيث: x_0 قيمة وسيطية و a طول الفئة

مثال:أجر 25 عاملا في مصنع تبعا للأجر في الساعة

الفئات	n_i	c_i	$n_i \times c_i$	$c_i - x_0$	z_i	$n_i \times z_i$
$[2,4[$	5	3	15	-2	-1	-5
$[4,6[$	8	5	40	0	0	0
$[6,8[$	12	7	84	2	1	12

حيث $z_i = \frac{c_i - 5}{2}$

ومنه نجد: $\overline{Z} = \frac{7}{25}; \bar{X} = \frac{139}{25}$ اذن: $\bar{X} = 5 + 2\frac{7}{25} = \frac{139}{25}$

3-1-2- مزايا وعيوب الوسط الحسابي:

- يتعلق بكل قيم السلسلة وسهل الحساب.
- يخضع للحسابات الجبرية.
- يتأثر بالقيم المتطرفة ولا يمكن تعيينه بيانيا.
- العمود المقام من نقطة الوسط يقسم المدرج التكراري الى قسمين متساويين في المساحة.
- يصعب حسابه من الجداول التكرارية المفتوحة ومن البيانات الوصفية.

3-1-3- خواص الوسط الحسابي:

- الوسط الحسابي هو القيمة التي تعوض كل القيم ويبقى المجموع على حاله، أي:

$\sum_{i=1}^{n} x_i = n\frac{\sum_{i=1}^{n} x_i}{n} = n\bar{X} = \sum_{i=1}^{n} \bar{X}$ يمكن كتابتها بالشكل $\sum_{i=1}^{n}(x_i - \bar{X}) = 0$

- الوسط الحسابي هو القيمة التي يكون مجموع مربعات الانحرافات عنها أقل ما يمكن،أي:

$$\sum_{i=1}^{n}(x_i-\bar{X})^2=\sum_{i=1}^{n}(x_i-\bar{X}+x_0-x_0)^2=\sum_{i=1}^{n}(x_i-x_0)^2-n(x_0-\bar{X})$$

- إذا كان X متغيرا إحصائيا يأخذ القيم x_1،...،x_k وكان Yمتغيرا إحصائيا حيث: $y_i=ax_i+b$

فان: $\bar{Y}=a\bar{X}+b$ لأن: $\bar{Y}=\frac{\sum_1^n y_i}{n}=\frac{\sum_1^n(ax_i+b)}{n}=\frac{a\sum_1^n x_i}{n}+\frac{\sum_1^n b}{n}=a\bar{X}+b$:

3-2- الوسط الهندسي: وهو مقياس شائع الاستعمال في الدراسات الاقتصادية

3-2-1- حساب المتوسط الهندسي:

أ- **في حالة البيانات غير المبوبة**: الوسط الهندسي لـ n قيمة موجبة والذي نرمز له بـ G العدد المعرف كمايأتي:

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

ب-**في حالة البيانات المبوبة**:**1-** الوسط الهندسي G للقيم x_n،...،x_1 مرفقة بالتكرارات n_n،...،n_1 على الترتيب:

$$G=\sqrt[\sum_1^n n_i]{x_1^{n_1}\times x_2^{n_2}\times\ldots\times x_n^{n_n}}$$

ملاحظة : لحساب المتوسط الهندسي يمكن اللجوء الى اللوغاريتم، ثم ندخل الاسية فنجد مثلا:

$$G=e^{\frac{\sum_{i=1}^n n_i ln x_i}{\sum_{i=1}^n n_i}}$$

2 - إذا كانت c_1، c_2، ...،c_n مركز فئات الظاهرة X وكانت n_1، n_2، ...،n_n تكرارتها على الترتيب فإن:

$$G=\sqrt[\sum_1^n n_i]{c_1^{n_1}\times c_2^{n_2}\times\ldots\times c_n^{n_n}}$$

3-2-2- مزايا و عيوب المتوسط الهندسي: - يدخل في حسابه جميع القبم

- أقل تأثرا بالقيم المتطرفة من المتوسط الحسابي
- يفضل استخدامه في وصف الظواهر النسبية و حساب المعدلات.
- لا يمكن حسابه من الجداول المفتوحة.
- ليس له معنى إذا كانت احد القيم سالبة او معدومة.

3-3- الوسط التوافقي: وهو مقياس خاص يستعمل في تحديد معدلات السرعة،...

3-2-1- حساب المتوسط التوافقي:

أ- **في حالة البيانات غير المبوبة**: الوسط التوافقي لـ n قيمة والذي نرمز له بـ H العدد المعرف كمايأتي:

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

ب- في حالة البيانات المبوبة:1- الوسط التوافقي H للقيم $x_n،...،x_1$ مرفقة بالتكرارات $n_1،...،n_n$ على الترتيب:

$$H = \frac{\sum_1^n n_i}{\sum_{i=1}^{n} \frac{n_i}{x_i}}$$

2 - إذا كانت $c_1، c_2، ...، c_n$ مركز فئات الظاهرة X وكانت $n_1، n_2، ...، n_n$ تكرارتها على الترتيب فإن:

$$H = \frac{\sum_1^n n_i}{\sum_{i=1}^{n} \frac{n_i}{c_i}}$$

3-3-2- مزايا و عيوب المتوسط التوافقي: - يدخل في حسابه جميع القيم

- أقل تأثرا بالقيم المتطرفة من المتوسط الحسابي
- يفضل استخدامه في حساب معدلات السرعة و الأسعار.
- لا يمكن حسابه من الجداول المفتوحة.
- لا يمكن حسابه إذا كانت احدى القيم معدومة.

ملاحظة : $H < G < \bar{X}$

3-4- الوسط التربيعي: $Q = \sqrt{\frac{\sum_{i=1}^{n} n_i x_i^2}{\sum_{i=1}^{n} n_i}}$

3-5- الوسيط: وسيط سلسلة احصائية ما والذي نرمز له بالرمز Med القيمة التي من أجلها يتساوى عدد القيم الاصغر منها مع القيم الاكبر منها والمرتبة تصاعديا او تنازليا، أي التكرار النسبي للقيم الأصغر أو الأكبر من الوسيط يساوي على الأكثر 5.0 .

3-5-1- حساب الوسيط: يتطلب حساب الوسيط ترتيب حدود السلسلة ترتيبا تصاعديا – التكرار المتجمع الصاعد –

2-5-1-1- المتغير متقطع:

أ- إذا كانت عدد القيم N– التكرارات – فرديا فان: $Med = x_{\frac{N+1}{2}}$

ب-إذا كانت عدد القيم N– التكرارات – زوجيا فان: $Med = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}$

مثال 3. 5. 1: عين وسيط السلسلة التالية : 10-11-12-13-14-15

لدينا عدد القيم زوجي =6 و منه $Med = \frac{x_{\frac{6}{2}} + x_{\frac{6}{2}+1}}{2} = \frac{12+13}{2} = 12.5$

2-5-1-2- المتغير مستمر: يمكن باستعمال المضلع التكراري المتجمع الصاعد و بتطبيق نظرية طاليس ايجاد العلاقة التالية:

$$Med = a + \frac{\frac{\sum n_i}{2} - N_{i-1}^{+}}{n_e} l$$

حيث : a الحد الادنى للفئة الوسيطية – تتحدد بناءا على رتبة الوسيط $\left(r = \frac{\sum n_i}{2}\right)$و التي يقابلها في التكرار المتجمع الصاعد –

N_{i-1}^{+} التكرار المتجمع الصاعد قبل الفئة الوسيطية . n_e تكرار الفئة الوسيطية . l طول فئة الوسيطية.

2–5–2– تحديد الوسيط بيانيا: الوسيط بيانيا هو فاصلة نقطة تقاطع المنحنى التكراري المتجمع الصاعد مع النازل، كما يمكن ايجاده من أحدهما فقط بتعيين فاصلة نقطة تقاطع المنحنى مع $y = \frac{N}{2}$.

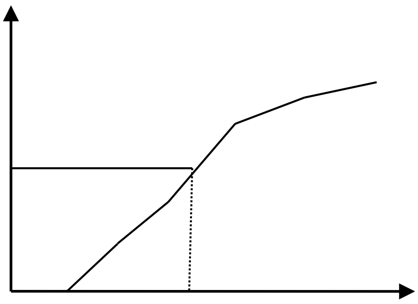
مثال 3. 5. 2: انشئ المدرج يرغب صاحب مصنع معرفة رغبات زبائنه فأجرى سبرا للآراء فتحصل على معلومات موزعة حسب السن كماياتي:

العمر	[20□15[	[25□20[	[30□25[	[35□30[	[40□35[
f%	13	28	35	15	9

1– انشئ المنحنى للتكراري النسبي المجمع الصاعد ثم استنتج الوسيط.

2– عين الوسيط حسابيا.

الحل: 1–



2–الفئة الوسيطية [25,30[ومنه: $Med = 25 + \frac{50-41}{35} \times 5 = 26.286$

2–5–3– خواص الوسيط: – لا يتأثر بالقيم الشاذة أو المتطرفة و لا يعتمد إلا على قيمة أو قيمتين.

– سهل الحساب إلا انه يصعب حسابه من البيانات الوصفية .

– مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيمة أخرى، أي:

$$\sum |x - med| \leq \sum |x - a|; a \neq med$$

2–6– المنوال: وهو مقياس يستعمل لمعرفة النمط الشائع، و يكثر استخدامه في البيانات النوعية، ونرمز له بالرمز Mod

2–6–1– حساب المنوال:

2–6–1–1– في حالة المتغير المتقطع: المنوال هو القيمة التي يقابلها أكبر تكرار.

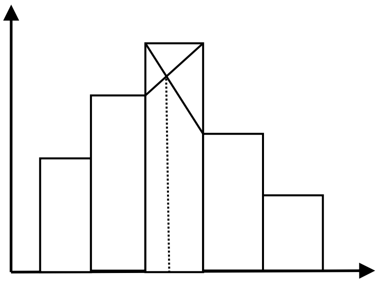
مثال 3. 6. 1: منوال السلسلة 7–8–8–8–9–9–9 هو 8

2–6–1–2– في حالة المتغير المستمر: باستعمال المدرج التكراري، وبتنفيذ طريقة الرافعة أو الفروق لبيرسون و بتطبيق نظرية طاليس ، نجد العلاقة التالية:

$$Mod = a + \frac{\Delta_1}{\Delta_1 + \Delta_2} l$$

حيث : a الحد الادنى للفئة المنوالية – الفئة الأكبر تكرارا –، Δ_1 الفرق بين تكرار الفئة المنوالية و الفئة السابقة لها، Δ_2 الفرق بين تكرار الفئة المنوالية و الفئة الموالية لها ، l طول فئة المنوالية.

2–6–1–3– حساب المنوال بيانيا: و يتم برسم المدرج التكراري للتوزيع، ثم نصل بداية المستطيل للفئة المنوالية ببداية المستطيل للفئة اللاحقة، ونقوم بايصال نهاية المستطيل للفئة المنوالية بنهاية المستطيل للفئةالسابقة، المنوال هو فاصلة نقطة التقاطع.



مثال 3. 6. 2: أحسب منوال المثال 3. 5. 2 جبريا و بيانيا

الحل : الفئة المنوالية $[25,30[$ و منه $Mod = 25 + \frac{7}{7+20} \times 5 = 26.296$

ملاحظة: في التوزيعات غير المتناظرة بشكل معتدل ووحيدة المنوال لدينا: $\overline{X} - \text{Mod} = 3(\overline{\text{X}} - \text{Med})$

2–6–2– خواص المنوال: - أسهل المقاييس حسابا ولايتأثر بالقيم المتطرفة

- يمكن حسابه بيانيا.
- غير قابل للعمليات الجبرية.
- يمكن حسابه من الجداول المفتوحة.

3–7– الربيعيات: يمكن تقسيم البيانات المرتبة الى أربعة اقسام متساوية – مثلا نحتاج 50 % من العناصر المركزية –

3–7–1– الربعي الأول: و نرمز له Q_1 هو أصغر قيمة للطبع الاحصائي بحيث يكون 25 % من الحدود المرتبة على الأقل لها قيمة أصغر أو تساوي Q_1.

3–7–2– الربعي الثالث: و نرمز له Q_3 هو أصغر قيمة للطبع الاحصائي بحيث يكون 75 % من الحدود المرتبة على الأقل لها قيمة أصغر أو تساوي Q_3.

3–7–3– حساب الربيعيات: نرتب القيم تصاعديا – التكرار المتجمع الصاعد –

3–7–3–1– في حالة متغير متقطع: (i=1،3) Q_i هو القيمة التي رتبتها n حيث n أصغر عدد طبيعي يحقق: $n \geq \frac{iN}{4}$.

3–7–3–2– في حالة متغير مستمر: يمكن باستعمال المضلع التكراري المتجمع الصاعد و بتطبيق نظرية طاليس ايجاد العلاقة التالية:

$$Q_i = a + \frac{\frac{i\sum n_i}{4} - N_{i-1}^{+}}{n_Q} l$$

حيث : a الحد الادنى للفئة الربيعية – تتحدد بناءا على رتبة الربيع $\left(r = i\frac{\sum n_i}{4}\right)$ و ما يقابلها في التكرار المتجمع الصاعد –

N_{i-1}^{+} التكرار المتجمع الصاعد قبل الفئة الربيعية . n_Q تكرار الفئة الربيعية . l طول فئة الربيعية.

3-7-3-3- حساب الربيعات بيانيا: Q_i (i=1،3) هي فاصلة النقطة من منحنى التواترات المجمعة الصاعدة التي ترتيبها $\frac{i}{4}$.

مثال3-7-1- عين الربيع الأول للسلسلة 3-3-4-4-4-4-5-7-7-8-9.

لدينا 11/4=2.75 ومنه $Q_1=x_3=4$.

مثال 3-7-2- عين الربيع الثالث للمثال 3-5-2-

لدينا الفئة الربيعية [25,30[ومنه: $Q_3 = 25 + \frac{75-41}{35} \times 5 = 29.857$.

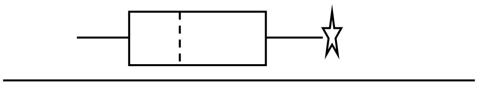
ملاحظة: نجد في بعض المراجع الطريقة التالية لتعيين الربيعات لمتغير منفصل: نحدد رتبة الربيع $r = (N+1)\frac{i}{4}$ نميز حالتين:

اذا كان r عددا صحيحا فان $Q_i = x_{(r)}$ اما اذا كان r عددا كسريا فان: $x_{(l)} < Q_i < x_{(u)}$ و $Q_i = x_{(l)} + (r-l)\left(x_{(u)} - x_{(l)}\right)$

3-8-العشريات و المئينيات: بنفس طريقة الربيعات نعرف العشريات D_i (i=1،...،9) و المئينيات C_i (i=1،...،99) – تقسيم السلسلة الى 10 أقسام ، 100 قسم متساو على الترتيب-.

3-9-المخطط بالعلبة: نرسم مخططا بالعلب حسب الطريقة التالية:

- نضع قيم الطبع على محور أفقي او شاقولي.
- نعين على هذا المحور القيم $min; max; Q_1; Med; Q_3$
- نمثل القيم الشاذة بنجمة حيث: حدها الأدنى $Q_1 - \frac{Q_3-Q_1}{2}$ و حدها الأعلى $Q_3 + \frac{Q_3-Q_1}{2}$
- نكون مستطيلا بالتوازي مع المحور طوله Q_3-Q_1 وعرضه كيفي



مقاييس التشتت

تـــمهيــد: إن مقاييس النزعة المركزية لا تسمح لنا بالوصف الكامل لسلسلة احصائية ما إذ لا تعطينا أي إرشاد حول توزيع الصفات داخل السلسلة.

فتشتت بيانات ظاهرة هي درجة تقارب أو تباعد البيانات عن بعضها البعض أو عن مقياس النزعة المركزية.

4–1– المقاييس الأولية للتشتت:

4–1–1– المدى: وهو طول سلسلة ونرمز له E و يساوي الفرق بين أكبر قيمة و أصغر قيمة للمتغير المدروس، أي: $E = x_{max} - x_{min}$

أما في حالة البيانات المبوبة $E = c_F - c_p$. المدى سهل الحساب و يستعمل في مراقبة الجودة و المناخ و يتأثر بالقيم الشاذة.

4–1–2– المدى الربيعي: ونرمز له بـ W وهو $W = Q_3 - Q_1$، والأكثر استعمالا الإنحراف الربيعي L_Q ويساوي $\frac{W}{2}$

أما الانحراف الربيعي النسبي فقيمته $\frac{W}{Med}$. وهو مقياس لا يتأثر بالقيم المتطرفة و يمكن حسابه من الجداول المفتوحة،و يستخدم كمقياس للتشتت في التوزيعات شديدة الالتواء.

4–1–3– الإنحراف المتوسط: هو الوسط الحسابي لإنحرافات القيم عن وسطها الحسابي، ويتميز بصعوبة حسابه، ويعبر عن شدة تجمع البيانات حول بعضها و مدى تجانسها حول قيمة مركزية.

أ– **في حالة البيانات غير المبوبة**: متوسط الانحراف لـ n قيمة والذي نرمز له بـ e العدد المعرف كمايأتي:

$$e = \frac{\sum_{i=1}^{n} |x_i - \bar{X}|}{n}$$

ب–**في حالة البيانات المبوبة**:**1**– المتوسط المطلق للانحراف e للقيم $x_1، \dots، x_n$ مرفقة بالتكرارات $n_1، \dots، n_n$ على الترتيب هو:

$$e = \frac{\sum_{i=1}^{n} n_i |x_i - \bar{X}|}{\sum_{i=1}^{n} n_i}$$

2 – إذا كانت $c_1، c_2، \dots، c_n$ مركز فئات الظاهرة X وكانت $n_1، n_2، \dots، n_n$ تكرارتها على الترتيب فإن:

$$e = \frac{\sum_{i=1}^{n} n_i |c_i - \bar{X}|}{\sum_{i=1}^{n} n_i}$$

مثال: نعتبر السلسلة 12–16–11–14–17 متوسطها 14 وانحرافاتها المطلقة هي 2–2–3–0–3 ويكون e=2

ملاحظة : الانحراف المتوسط عن الوسيط أقل من الانحراف المتوسط عن الوسط الحسابي، أي: $e_{Med} < e_{\bar{X}}$

4–2– التبايـن:

4–2– التبايـن: وهو أكثر مقاييس التشتت تطبيقا، يرمز له σ^2 أو Var ويعبر عن متوسط مربعات انحراف المعطيات عن المتوسط الحسابي للسلسلة، أي أن:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}$$

- **في حالة بيانات مبوبة**: $\sigma^2 = \frac{\sum_{i=1}^{n} n_i (x_i-\mu)^2}{\sum_{i=1}^{n} n_i}$
- **في حالة العينات الصغيرة**: $s^2 = \frac{\sum_{i=1}^{n} n_i (x_i-\bar{x})^2}{\sum_{i=1}^{n} n_i}$ وهو التقدير غير المتحيز لتباين المجتمع.

مثال: تم سحب عينة من طلبة سنة اولى و سجل عدد سنوات التمدرس النظامية فكانت النتائج التالية : 12-13-14-15-16

لحساب تباين السلسلة، نقوم بحساب الوسط الحسابي اولا فنجد: $\bar{x} = \frac{70}{5} = 14$، ثم نحسب $(x_i - \bar{x})^2$ فنجد: 4-1-0-1-4

و عليه: $s^2 = \frac{4+1+0+1+4}{4} = 2.5$

4-2-1- طرق أخرى لحساب التباين:

4-2-1-1- قاعدة كوينغ: - **في حالة البيانات غير المبوبة**: $\boldsymbol{\sigma^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \mu^2}$

- **في حالة البيانات المبوبة**: $\boldsymbol{\sigma^2 = \frac{\sum_{i=1}^{n} n_i x_i^2}{\sum_{i=1}^{n} n_i} - \mu^2}$
- **في حالة العينات** : $\boldsymbol{s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)}$

برهان: $\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} = \frac{\sum_{i=1}^{n}(x_i^2-2x_i\mu+\mu^2)}{n} = \frac{\sum_{i=1}^{n} x_i^2}{n} - 2\frac{\sum_{i=1}^{n} x_i}{n}\mu + \mu^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \mu^2$

4-2-1-2- $\boldsymbol{\sigma^2 = \sum_{i=1}^{n} x_i^2 - n\mu^2}$

4-2-1-3- الوسط الفرضي: $\boldsymbol{\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-x_0)^2}{n} - (\mu - x_0)^2}$

4-2-3- الانحراف المعياري: و هو الجذر التربيعي للتباين، و يرمز له أي $\sqrt{V}$ =

مثال: الانحراف المعياري للسلسلة 1-2-3-4-5 هو: $\sigma = \sqrt{\frac{1^2+2^2+3^2+4^2+5^2}{5} - 3^2}$

4-2-3-1- مزايا و عيوب الانحراف المعياري: - قيمته صغيرة و يتعامل جميع القيم

- أكثر مقاييس التشتت استخداما وسهل الحساب.
- يتأثر بالقيم الشاذة، ولا يمكن حسابه من الجداول المفتوحة.
- يخضع للعمليات الجبرية.

ملاحظة 1: كلما صغرت قيمة الانحراف المعياري دل على ان القيم أقل تشتتا، وأن الوسط الحسابي يمثلها، و نعتبر القيم عير مشتتة اذا كان الانحراف المعياري يمثل أقل من 20% من الوسط الحسابي.

ملاحظة 2: $\sigma = \frac{5}{4}e = \frac{3}{2}LQ$ و $\frac{E}{6} < \hat{S} < \frac{E}{4}$ *اذا وقع الانحراف المعياري خارج المدى دل ذلك على وجود قيم شاذة.*

4-2-3-2- خــواص:

- الانحراف المعياري للمعطيات الثابتة يساوي 0 لأن $x_i - \mu = a - a = 0$.
- اذا كانت $x_1، ...، x_n$ قيم سلسلة احصائية متوسطها $\bar{x}$ و انحرافها المعياري σ_X وكان

حيث: $y_i = ax_i + b$ $b \in \mathbb{R}, a \in \mathbb{R}^*$ فإن: $\bar{y} = a\bar{x} + b$ و $\sigma_Y = |a|\sigma_X$

بــــرهـــــــان: $${\sigma_Y}^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n} = \frac{\sum_{i=1}^{n}(ax_i+b-a\bar{x}-b)^2}{n} = \frac{a^2\sum_{i=1}^{n}(x_i-\bar{x})^2}{n} = a^2\sigma_X^2$$

4-2-3-3- اذا كان متوسط العينة1 هو $\overline{x_1}$ و انحرافها المعياري s_1 حجمها n_1 وكان متوسط العينة2 هو $\overline{x_2}$ و انحرافها المعياري s_2 حجمها n_2 فإن المتوسط و الإنحراف المعياري للعينتين هما على التوالي: $\bar{X} = \frac{n_1\overline{x_1}+n_2\overline{x_2}}{n_1+n_2}$ و $S^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-1}$

4-3- معامل الاختلاف : وهو مقياس دون وحدة يفضل استخدامه عند المقارنة بين مجموعتين خاصة مختلفتي الوحدة.

4-3-1- معامل الاختلاف النسبي: وهو العدد الذي نرمز له بـ CV و يعرف بالعلاقة:

$CV = \frac{\sigma}{\mu} \times 100$ في حالة بيانات المجتمع، و $CV = \frac{S}{\bar{X}} \times 100$ في حالة بيانات العينة.

4-3-2- معامل الاختلاف الربيعي: وهو العدد الذي نرمز له بـ CVQ يستخدم في حالة الجداول التكرارية المفتوحة و يعرف بالعلاقة:

$$CVQ = \frac{Q_3-Q_1}{Med} \times 100.$$

مثال: لتكن لدينا معطيات سلسلة احصائية كالتالي: $Q_1 = 3, Q_3 = 8, \bar{x} = 6, Med = 5, s = 1$

$$CV = \frac{1}{6} \times 100, CVQ = \frac{8-3}{5} \times 100$$

4-4- معايرة سلسلة : لتكن $A(x_i, n_i)$ سلسلة احصائية وسطها الحسابي $\bar{X}$ و انحرافها المعياري σ ، نرمز بـ y_i للقيمة المعيرة للقيمة x_i

حيث: $y_i = \frac{x_i-\bar{X}}{\sigma}$ ومنه: $\bar{y} = 0, \sigma_Y = 1$

بــرهـــان: $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})}{\sigma \times n} = 0$ و $\boldsymbol{\sigma_Y^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2 \times n} = \frac{\sigma^2}{\sigma^2} = 1$

مثال: في سباق 100م سرعة، حصل العداء A على 10 ثواني في مجموعته التي زمنها المتوسط 9 ثا بانحراف2 ثا، و تحصل العداء B على 9 ثواني في مجموعته التي زمنها المتوسط 8 ثا بانحراف1 ثا ، أيهما أسرع في مجموعته؟

المقارنة ممكنة بعد المعايرة، نجد: $y_A = \frac{10-9}{2} = \frac{1}{2}, y_B = 1$ اذن: A أسرع في مجموعته.

مقاييس الشكل

تـمـهيـد: يصادفنا عند رسم المنحنى التكراري للتوزيع عدة أشكال فقد يكون المنحنى متماثلا أي متناظرا بالنسبة لمستقيم عمودي، وقد يكون ملتو في اتجاه معين او شديد الالتواء منخفضا أو مرتفعا ...

فمقاييس الموقع و التشتت تعطينا صورة عن تمركز القيم و مدى تجانسها، لكن لا توضح كيفية انتشار القيم، لذا يصبح ضرورة ملحة ايجاد مقاييس تعطي وصفا للبيانات من جهة و تسمح بالمقارنة بين التوزيعات من جهة أخرى.

5-1- الـعــزوم: نميز بين نوعين من العزوم، قد يكون العزم حول نقطة معينة أو حول المتوسط الحسابي.

5-1-1- العزم البسيط – اللامركزي – نسمي عزما بسيطا من الرتبة □ العدد m_r المعرف كالآتي:

أ- **في حالة البيانات غير المبوبة:** $m_r = \frac{\sum_{i=1}^{n} x_i^r}{n}$

ب- **في حالة البيانات المبوبة:** 1- $m_r = \frac{\sum_{i=1}^{n} n_i x_i^r}{\sum_{i=1}^{n} n_i}$

2 – إذا كانت c_1، c_2،...، c_n مركز فئات المتغير X وكانت n_1، n_2، ...، n_n تكرارتها على الترتيب فإن:

$$m_r = \frac{\sum_{i=1}^{n} n_i c_i^r}{\sum_{i=1}^{n} n_i}$$

5-1-1- العزم المركزي : نسمي عزما مركزيا من الرتبة □ العدد M_r المعرف كالآتي:

أ- **في حالة البيانات غير المبوبة:** $M_r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$

ب- **في حالة البيانات المبوبة:** 1- $M_r = \frac{\sum_{i=1}^{n} n_i (x_i-\bar{x})^r}{\sum_{i=1}^{n} n_i}$

2 – إذا كانت c_1، c_2،...، c_n مركز فئات المتغير X وكانت n_1، n_2، ...، n_n تكرارتها على الترتيب فإن:

$$M_r = \frac{\sum_{i=1}^{n} n_i (c_i-\bar{x})^r}{\sum_{i=1}^{n} n_i}$$

5-2- مقاييس الالتواء: يقصد بالالتواء امتداد التوزيع و يعبر عن درجة توزع البيانات حول نقطة مركزية.

5-2-1- مقاييس بيرسون:

5-2-1-1- مقاييس بيرسون الاول: يعطى بالعلاقة التالية: $P_1 = \frac{\bar{x}-Mod}{\sigma}$

$P_1 = 0$ التوزيع متماثل ، $P_1 > 0$ التوزيع ملتوي نحو اليمين، $P_1 < 0$ التوزيع ملتوي نحو اليسار.

لا يمكن استخدامه في الجداول المفتوحة، متعددة المناويل أو شديدة الالتواء

. **5-2-1-2- مقاييس بيرسون الثاني**: يعطى بالعلاقة التالية: $P_2 = 3\frac{(\bar{x}-Med)}{\sigma}$

$P_2 = 0$ التوزيع متماثل ، $P_2 > 0$ التوزيع ملتوي نحو اليمين، $P_2 < 0$ التوزيع ملتوي نحو اليسار.

لا يمكن استخدامه في الجداول المفتوحة أو شديدة الالتواء

. **5-2-1-3- مقاييس بيرسون للالتواء العزمي**: يعطى بالعلاقة التالية: $\beta_1 = \frac{M_3^2}{M_2^3}$

لايعطي فكرة جيدة عن الالتواء سوى تماثل التوزيع.

. **5-2-2- مقاييس فيشر**: يعطى بالعلاقة التالية: $F_1 = \frac{M_3}{\sigma^3}$

$F_1 = 0$ التوزيع متماثل ، $F_1 > 0$ التوزيع ملتوي نحو اليمين، $F_1 < 0$ التوزيع ملتوي نحو اليسار.

5-2-3- مقاييس يول- الالتواء الربيعي-: يعطى بالعلاقة التالية: $\gamma_1 = \frac{Q_3+Q_1-2Med}{Q_3-Q_1}$

$\gamma_1 = 0$ التوزيع متماثل ، $\gamma_1 > 0$ التوزيع ملتوي نحو اليمين، $\gamma_1 < 0$ التوزيع ملتوي نحو اليسار.

5-3- التفرطح: ويقصد به مقدار أو درجة علو- ارتفاع- او انخفاض منحنى التوزيع عن التوزيع الطبيعي.

ويكون التوزيع مدببا – مرتفعا- اذا تركزت القيم في المنتصف و قلت في الطرفين، ومفرطحا عكس ذلك.

5-3-1- مقاييس بيرسون للتفرطح: يعطى بالعلاقة التالية: $\beta_2 = \frac{M_4}{M_2^2}$

$\beta_2 = 3$ التوزيع متماثل ، $\beta_2 > 3$ التوزيع مدبب ، $\beta_2 < 3$ التوزيع مفرطح.

5-3-2- مقاييس فيشر للتفرطح: يعطى بالعلاقة التالية: $F_2 = \beta_2 - 3$

5-3-3- مقاييس كيللي: يعطى بالعلاقة التالية: $K = \frac{Q_3-Q_1}{D_9-D_1}$

$K = 0.526$ التوزيع طبيعي ، $K > 0.526$ التوزيع مدبب ، $K < 0.526$ التوزيع مفرطح.

مثال : احسب مقياس للتفرطح و الالتواء للسلسلة التالية

x_i	2	3	4	5	6	المجموع
$(x_i - \bar{x})$	-2	-1	0	1	2	0

$(x_i - \bar{x})^2$	**4**	**1**	**0**	**1**	**4**	**10**
$(x_i - \bar{x})^3$	**−8**	**−1**	**0**	**1**	**8**	**0**
$(x_i - \bar{x})^4$	**16**	**1**	**0**	**1**	**16**	**34**

ومنه : $M_4 = \frac{34}{5};\ M_3 = 0; M_2 = 2$ و $Med = 4; Q_1 = 3; Q_3 = 5$ اذن: $\boldsymbol{\beta_2} = \frac{\mathbf{34}}{\mathbf{20}}$ و $\boldsymbol{\gamma_1} = \frac{\mathbf{5+3-2\times4}}{\mathbf{5-3}}$