المحور الرابع: مقاييس النزعة المركزية

تعلمنا كيفية تنظيم المعطيات في جداول إحصائية وتمثيلها بيانياً، مما يعطينا فكرة عامة عن تطور الظاهرة المدروسة. لتحليل الظاهرة بدقة وسرعة، نستخدم مقاييس النزعة المركزية، والتي تساعد في فهم كيفية تمركز البيانات حول قيمة معينة. تشمل هذه المقاييس المتوسط الحسابي، المتوسط الهندسي، المتوسط التوافقي، المتوسط التربيعي، الوسيط، المنوال. هذه المقاييس توفر وصفاً مختصراً ومفيداً للبيانات، وتساعد في المقارنة بين مجموعات القيم أو الظواهر المختلفة.

1- الوسط الحسابي $\overline{x}$

يعتبر الوسط الحسابي أهم مقاييس النزعة المركزية وأكثرها استعمالا في التطبيقات العملية وخاصة في المقارنات الإحصائية وعادة ما يسمى المعدل. ويفضل الوسط الحسابي على جميع مقاييس النزعة المركزية لكونه يستعمل جميع البيانات ويستخدم الصيغ الرياضية.

1-1-الوسط الحسابي في حالة البيانات غير المبوبة.

وهو عبارة عن مجموع القيم مقسوما على عددها.

فإذا كانت لدينا القيم X_n ...., x_3 . x_2 . x_1 فإن متوسطها الحسابي يحسب بالعلاقة:

$$\overline{x} = \frac{\sum_{i=1}^{n} Xi}{n}$$

حيث : $\overline{x}$: المتوسط الحسابي.

X_i : قيم الظاهرة.

n : عدد البيانات. (المشاهدات)

1-2- المتوسط الحسابي في حالة البيانات المتكررة.

إذا كان لدينا المشاهدات $x_1 .x_2\ x_3 \ldots\ldots x_i$ ، وكانت تكرارات هذه المشاهدات $n_1.n_2 \ldots\ldots n_i$ على الترتيب ولحساب الوسط الحسابي في هذه الحالة فإننا نضرب المشاهدات بالتكرارات المناظرة ونقسم على مجموع هذه التكرارات. فيكون :

$$\bar{X}=\frac{\mathrm{n1x1} + \mathrm{n2x2} + \ldots + \mathrm{nixi}}{n1+n2+\cdots+ni} \qquad \Rightarrow \qquad \bar{X}=\frac{\sum_{\mathrm{i}=1}^{n}\mathrm{nixi}}{\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathrm{ni}}$$

3-1 المتوسط الحسابي في حالة توزيع تكراري

نستعمل نفس علاقة الوسط الحسابي في حالة البيانات المتكررة، غير الذي ينقصنا هي القيم النقطية للمتغير الإحصائي x_i ،لأن هذه القيم معطاة في حالة توزيع تكراري على شكل مجالات جزئية أو فئات، ولحل هذا الإشكال نستبدل هذه الفئات بمراكزها أي تستبدل x_i بمراكز الفئة c_i .

وعليه تصبح العلاقة :

$$\bar{X} = \frac{\sum \mathrm{ci\ ni}}{\sum \mathrm{ni}}$$

4-1 المتوسط الحسابي الموزون (المرجح)

في بعض الأحيان فإن القيم المراد حساب المتوسط الحسابي لها لا تكون لها نفس الأهمية بل أهميات نسبية مختلفة تختلف باختلاف عامل الترجيح الخاص بها، في مثل هذه الحالات فإن المتوسط الحسابي البسيط يمكن الاعتماد عليه في إيجاد المتوسط الصحيح والمنطقي، بل يتطلب الأمر استخدام صيغة أخرى تسمى بصيغة المتوسط الحسابي المرجح.

$$\bar{X} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathrm{nixi}}{\sum \mathrm{ni}}$$

X_i : القيم

n_i : المعاملات

6-1 خصائص المتوسط الحسابي

- ✓ يعتبر المتوسط الحسابي أبسط مقاييس النزعة المركزية حسابا وأكثرها استخداما.
- ✓ يأخذ المتوسط الحسابي بعين الاعتبار جميع قيم الظاهرة المدروسة.
- ✓ يتعذر حساب المتوسط الحسابي في الجداول المفتوحة.
- ✓ لا يمكن تحديد المتوسط الحسابي بيانيا.
- ✓ يستعمل الوسط الحسابي في حالة المتغيرات الكمية أي القابلة للقياس.
- ✓ أساس حساب المتوسط الحسابي هو الحساب التجميعي.
- ✓ لا يمكن أن يكون لأي توزيع تكراري أكثر من وسط حسابي.
- ✓ يتأثر الوسط الحسابي بالقيم المتطرفة أو الشاذة (القيم التي تقع في طرفي مجال الدراسة).

2- المتوسط الهندسي G

في الحالات التي تكون فيها قيم الظاهرة المدروسة عبارة عن نسب أو معدلات، أي في الحالات التي نرغب فيها بدراسة معدل تغيرات ظاهرة ما نستعمل المتوسط الهندسي و هذا ما يجعله واسع الاستخدام في الحياة الاقتصادية، لأن التركيز يكون غالبا منصبا على إيجاد متوسط نسب التغير لبعض الظواهر مثل : تطور الدخول، زيادة الأجور، والنمو السكاني..... إلخ.

1.2. تعريف :

إن المتوسط الهندسي ل n قيمة من المتغير الإحصائي يعطى بالعلاقة :

- **البيانات غير مبوبة**

$$G = \sqrt[n]{x_1 x_2 x_3 \ldots\ldots x_i}$$

ولتسهيل العمليات الحسابية نقوم بإدخال اللوغاريتم على الصيغة السابقة لتصبح من الشكل:

$$\text{Log } G = \frac{1}{n} \log (x_1 \,.\, x_2 \ldots . x_n)$$

$$= \frac{1}{n} [\log x_1 + \log x_2 + \ldots . \log x_n]$$

$$\text{Log } G = \frac{1}{n} \sum_{i=1}^{n} logxi$$

- **البيانات المتكررة**

$$G = \sqrt[n]{x_1^{n1} x_2^{n2} x_3^{n3} \ldots\ldots x_n^{ni}}$$

حيث : $\sum n_i = N$ **(مجموع التكرارات)**

وبإدخال اللوغاريتم $\log G = \frac{1}{n} \sum_{i=1}^{n} ni \log Xi$

- **حالة التوزيع التكراري**

- $G = \sqrt[n]{C_1^{n1} C_2^{n2} C_3^{n3} \ldots . C_n^{ni}}$

وبإدخال اللوغاريتم تصبح

$$\log G = \frac{1}{N} \sum_{i=1}^{n} ni \log Ci$$

2.2. خصائص المتوسط الهندسي

✓ يدخل في حساب جميع القيم ولكنه أقل تأثر بالقيم المتطرفة من المتوسط الحسابي.

- ✓ لا يمكن حسابه من الجداول التكرارية المفتوحة من البداية إلى النهاية.
- ✓ لا يمكن حسابه في حالة وجود قيمة سالبة أو معدومة.
- ✓ يستخدم بأكثر واقعية عند وصف الظواهر النسبية.
- ✓ قيمة المتوسط الهندسي لأي ظاهرة أصغر دائما من قيمة المتوسط الحسابي $G < \bar{X}$

3. المتوسط التوافقي H

1.3. تعريف الوسط التوافقي H هو مقلوب الوسط الحسابي لمقلوب هذه القيم، فإذا كان لدينا المتغير X_i يأخذ القيم X_n ....., X_1 والتي عددها N يحسب الوسط التوافقي على الشكل التالي:

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

أما إذا كانت البيانات متكررة أو مبوبة في جداول توزيع تكراري فإن :

$$\frac{\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{ni}}{\sum_{\mathrm{i=1}}^{\mathrm{n}} \frac{ni}{\mathrm{Xi}}}$$

حيث :

N_i : تمثل التكرار

X_i: تمثل القيم أو مراكز الفئات

يستعمل حساب المتوسط التوافقي في حالة وجود العلاقة العكسية بين الظاهرتين.

2.3. خواص المتوسط التوافقي

- ✓ يأخذ بعين الاعتبار جميع القيم وتأثره بالقيم الشاذة أقل من تأثر المتوسط الحسابي.
- ✓ لا يمكن حسابه في حالة وجود بيانات معدومة.
- ✓ يعطي نتائج أكثر واقعية في حالة حساب متوسطات الأسعار والسرعة.
- ✓ قيمته دائما أقل من المتوسط الهندسي :

$$\bar{X} > G > H$$

4. المتوسط التربيعيQ

يعرف المتوسط التربيعي لقيم المتغيرات الإحصائية على أنه الجذر التربيعي لمجموع مربعات هذه القيم على عدد المفردات ويعطى بالعلاقة :

$$Q = \sqrt{\frac{\sum X_i^2}{N}}$$

أما في حالة ما إذا كانت البيانات مبوبة وحالة التوزيع التكراري فإن

$$Q = \sqrt{\frac{\sum (nixi)^2}{\sum ni}}$$

حيث

N_i : التكرارات

X_i: القيم أو مركز الفئات.

ملاحظة :

يستعمل المتوسط التربيعي عندما يكون لدينا قيمة سالبة لأن القيم السالبة تصبح موجبة بالتربيع فيصعب في هذه الحالة استعمال المتوسطات الأخرى.

العلاقة بين المتوسطات الحسابية والهندسية والتوافقية والتربيعية

-عندما تكون السلسلة متناظرة تكون جميع المتوسطات متساوية.

$$H = G = X = M_0 = M_e$$

في ما عدا ذلك يكون لدينا

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

5-المنوال M_0

المنوال هو أحد مقاييس النزعة المركزية ويعرف بأنه القيمة الأكثر تكرارا أو شيوعا أو الصفة الأكثر شيوعا أي القيمة المسيطرة.

5-1 المنوال في حالة البيانات المبوبة

في هذه الحالة للبيانات المبوبة فإن الفئة التي يقابلها أكبر تكرار تدعى الفئة المنوالية (الفئة المسيطرة) ويمكن استخراج المنوال من هذه الفئة بطريقتين :

أ-الطريقة الحسابية

بعد تحديد الفئة المسيطرة نستطيع حساب المنوال بالعلاقة التالية :

$$M_0 = L_1 + \frac{\Delta_1}{\Delta_1+\Delta_2} .a$$

حيث

L_1: الحد الأدنى للفئة المنوالية.

Δ_1: الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي سبقتها.

Δ_2 : الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي تليها.

a : طول الفئة المنوالية.

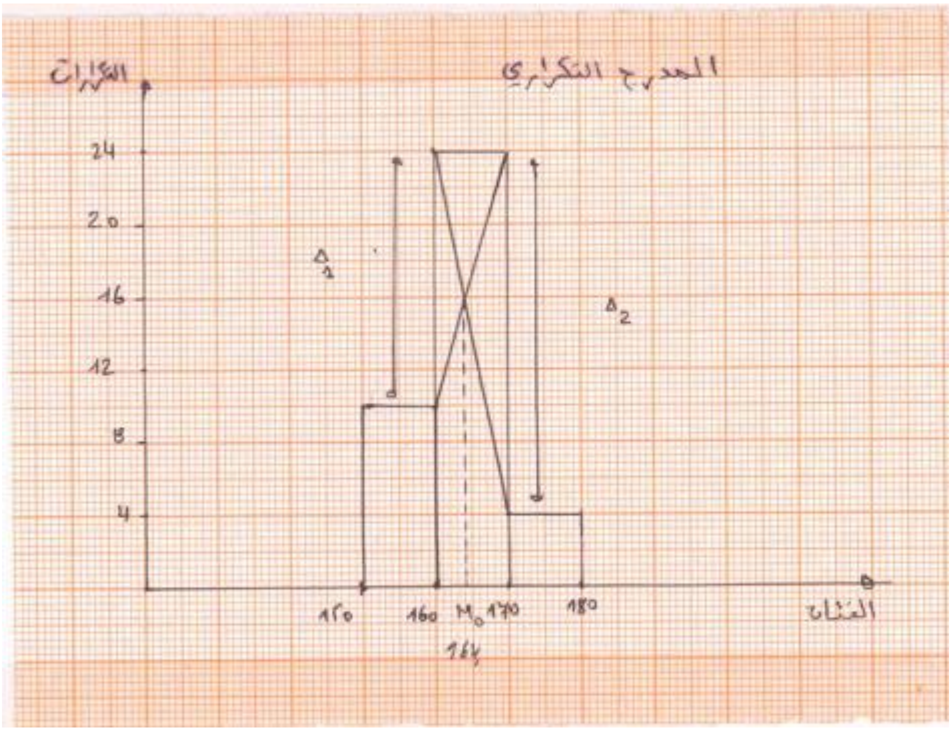
<u>ب- الطريقة البيانية</u>

1- نرسم المدرج التكراري للتوزيع المدروس (يمكن رسم الفئات الثلاثة فقط).

نصل الحد الأعلى للفئة قبل المنوالية بالحد الأعلى للفئة المنوالية بخط مستقيم

نصل الحد الأدنى للفئة المنوالية بالحد الأدنى للفئة بعد المنوالية بخط مستقيم.

نقطة تلاقي العود النازل من نقطة تقاطع الخطين المستقيمين على المحور الأفقي تمثل قيمة المنوال.



5-2- خواص المنوال :

- ✓ لا يأخذ بعين الإعتبار جميع البيانات وبالتالي فهو لا يتأثر بالقيم المتطرفة.
- ✓ لا يمكن حسابه بيانيا.
- ✓ يمكن أن يوجد أكثر من منوال لتوزيع واحد.
- ✓ يمكن الحساب من الجداول الإحصائية المفتوحة.

- ✓ يعتبر أفضل المتوسطات لوصف الظاهرة النوعية.

6-الوسيط M_e

يعرف الوسيط بأنه القيمة التي تقع تماما في منتصف مجموع البيانات بعد ترتيبها تصاعديا أو تنازليا، بحيث يكون مجموع التكرارات التي تسبقها مساويا لمجموع التكرارات التي تليها أو يكون عدد القيم التي تسبقها يساوي عدد القيم التي تليها بعد ترتيبها.

كما يوجد مقاييس تشبه في مفهومها مفهوم الوسيط مثل الربيعيات، والعشيريات والمثينات كما سنرى لاحقا.

6-1- الوسيط في حالة البيانات المفردة

وهنا أيضا نميز بين حالتين

أ- إذا كان عدد البيانات فردي

في هذه الحالة تكون قيمة الوسيط هي القيمة التي ترتيبها $\frac{n+1}{2}$

مثال : أحسب الوسيط للبيانات التالية 9، 6، 4، 8، 1

الحل : نرتب القيم تصاعديا على النحو : 9، 8، 6، 4، 1

بما أن $n = 5$ (فردي)

فإن الوسيط يساوي المشاهدة التي ترتيبها $\frac{n+1}{2}$ أي $\frac{5+1}{2}=3$

ومنه نجد أن $M_e = 6$

يمكن الحصول على نفس النتيجة بترتيب القيم تنازليا.

ب- إذا كان عدد البيانات زوجي

إذا كان عدد المشاهدات زوجي فإن الوسيط هو متوسط القيمتين اللتين ترتيبهما $\frac{n}{2}$ و $\frac{n}{2}+1$

مثال : أوجد وسيط مجموعة البيانات التالية (2، 4، 7، 12، 9، 5)

الحل :

ترتب القيم ترتيبا تصاعديا على النحو التالي : 12، 9، 7، 5، 4، 2

بما أن $n = 6$ زوجي.

فإن الوسيط هو معدل المشاهدتين اللتين ترتيبهما $\frac{n}{2}$ و $\frac{n}{2}+1$

حيث : $\frac{n}{2} = 3$ ، $\frac{n}{2}+1 = 4$ المشاهدتان هما 5 و 7 على التوالي

وبالتالي الوسيط $M_e = \frac{7+5}{2} \Rightarrow M_e = 6$

6-2 الوسيط في حالة البيانات المتكررة

إذا كانت البيانات المراد حساب الوسيط لها متكررة فإن الوسيط نقوم بحسابه باتباع الخطوات التالية:

- نحسب التكرار المتجمع الصاعد لقيم الظاهرة.
- نحدد ترتيب الوسيط $\frac{n}{2}$ (N = مجموع التكرارات)
- نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر أو أكبر أو تساوي من $\frac{N}{2}$ مباشرة، وهي القيمة التي تمثل الوسيط.

6-3- الوسيط في حالة التوزيع التكراري (البيانات المستمرة)

لتحديد الوسيط لبيانات مبوبة في جداول توزيع تكراري فإننا نتبع الخطوات التالية:

- نحسب التكرار المتجمع الصاعد أو النازل.

- نحدد ترتيب الوسيط وهو عبارة عن نصف مجموع التكرارات $\frac{N}{2}$
- نحدد الفئة الوسيطية أي الفئة التي يقع فيها الوسيط.
- نحدد ونحسب الوسيط بتطبيق إحدى العلاقتين:

* **الوسيط عن طريق التكرار المتجمع الصاعد**

$$M_e = L_1 + \frac{\frac{\sum ni}{2} - F'_{n-1}}{n_{me}} .a$$

حيث

L_1: الحد الأدنى للفئة الوسيطية

$\frac{N}{2} = \frac{\sum \text{ni}}{2}$ رتبة الوسيط

F'_{n-1} : التكرار المتجمع الصاعد الذي يسبق الفئة الوسيطية.

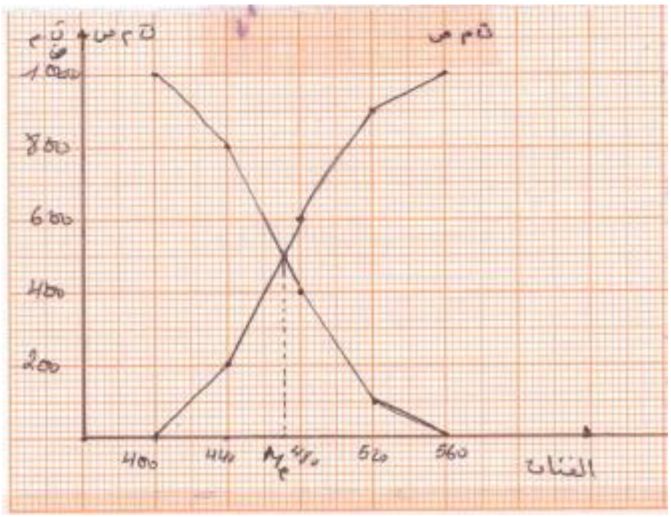
a: طول الفئة الوسيطية.

n_{m0}: التكرار المقابل للفئة الوسيطية.

4-6 حساب الوسيط بيانيا

يمكن استخراج M_e من منحنى التكرار المتجمع الصاعد والتكرار المتجمع النازل وذلك بإسقاط نقطة تقاطعها على محور الفواصل وبذلك نحصل على قيمة الوسيط.

مثال:



6- خواص الوسيط

يتصف الوسيط بعدة خصائص أهمها :

* يتغير كلما غيرنا أطوال الفئات بالنسبة لنفس التوزيع التكراري إذن يتميز بعدم الثبات.
* لا يتأثر الوسيط بالقيم المتطرفة أو الشاذة.
* يقسم المدرج التكراري إلى مساحتين متساويتين.
* يستعمل الوسيط في عدة مجالات منها على وجه الخصوص، دراسة الأجور، الأسعار، في نظرية أخطاء القياس، في دراسة الوفيات، في دراسة المدة المتوسطة للحياة.....الخ
* يمكن تحديده بيانيا.
* يمكن استخدامه في البيانات النوعية.
* يمكن حسابه من الجداول التكرارية المفتوحة.

7- أشباه الوسيط

الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين متساويين، بحيث نصف عدد البيانات أقل منه ونصف البيانات أكبر منه، ومادام يمكن تقسيم بيانات أي ظاهرة إلى

عدة أقسام متساوية وليس إلى قسمين فقط فإنه يمكن التعامل مع القيم التي تقسم هذه البيانات بنفس طريقة التعامل مع الوسيط.

- فإذا تم تقسيم البيانات إلى أربعة أقسام فإن المقياس يسمى بالربيع.
- إذا تم تقسيم البيانات إلى عشرة أقسام فإن المقياس يسمى بالعشير.
- إذا تم تقسيم البيانات إلى مئة قسم فإن المقياس يسمى بالمثين.

7-1- الربيعيات Q

نستعمل نفس الطريقة المتبعة لإيجاد الوسيط، غير أن الذي يتغير هو الترتيب وما يترتب عنه.

تنقسم الربيعيات إلى ثلاثة أقسام : الربيع الأول، الربيع الثاني (الوسيط)، الربيع الثالث.

أ-<u>الربيع الأول Q_1</u> : هي قيمة المتغير الإحصائي التي تقسم المجتمع الإحصائي إلى قسمين، يحتوي القسم الأول على 25% من الوحدات الإحصائية، أما القسم الثاني فيحتوي على 75% من هذه الوحدات. تقع قيمة الربيع الأول في نهاية الربع الأول من التوزيع الإحصائي، وفي هذا الموضع تكون مرتبة تساوي 25% أو $\frac{\sum ni}{4}$ حسب الترتيب التصاعدي لقيم المتغير الإحصائي ويرمز له بالرمز Q_1 ويعطى بالعلاقة :

$$Q_1 = L_1 + \frac{\frac{\sum ni}{4} - F_{n-1}^{\nearrow}}{n_Q} * a$$

ب-<u>الربيع الثالث</u> : هو عبارة عن قيمة المتغير الإحصائي التي تقسم المجتمع إلى قسمين، يحتوي القسم الأول على 75% من المجتمع المدروس، أما القسم الثاني فيحتوي على 25% المتبقية منه حسب الترتيب التصاعدي للمتغير المدروس ويرمز لها بالرمز Q_3 ومرتبته 75% أو $\frac{3\sum ni}{4}$ ويعطى بالعلاقة :

$$Q_3 = L_1 + \frac{\frac{3\sum ni}{4} - F_{n-1}^{\nearrow}}{n_a} * a$$

7-2- العشيريات D

يوجد على طول مجال الدراسة 9 عشيريات هي : $D_1, D_2, \ldots\ldots, D_9, D_{10}$

مرتبة كل منها على التوالي :

%100 ، %90، ....%20، %10 أو $\sum ni, \frac{9\sum ni}{10}, \ldots\ldots, \frac{2\sum ni}{10}, \frac{\sum ni}{10}$

نستعمل نفس فكرة الوسيط لتحديد هذه العلاقة وهي كالتالي:

$$D_i = L_1 + \frac{\frac{i\sum ni}{10} - F_{n-1}^{\nearrow}}{n_{Di}} * a$$

7-3 المثينات

يمكن تعميم طريقة حساب الوسيط والعشيريات إلى بقية مراتب قيم المتغير الإحصائي، حيث يمكن تحديد وحساب قيمة المتغير الإحصائي إذا عرفت مرتبته، ومن بين هذه القيم المثينات ومراتبها على التوالي :

%100 ، %99، ....%2، %1

أما العلاقة العامة لها فهي :

$$P_i = L_1 + \frac{\frac{i\sum ni}{100} - F_{n-1}^{\nearrow}}{n_{pi}} * a$$

حيث i مرتبة المتغير الإحصائي.

8-العلاقة بين الوسط الحسابي والوسيط والمنوال

بعد دراسة مقاييس النزعة المركزية لابد من الدخول في معرفة وتحديد العلاقة بينها. فإذا كان التوزيع متماثل فإن الوسط الحسابي والوسيط والمنوال متساوية أي

$$\bar{X} = M_e = M_0$$

وهذه التوزيعات المتماثلة توزيعات نظرية مثل التوزيع الطبيعي، بمعنى أن في الحياة العملية قد لا يكون هذا التماثل كاملا وإنما يقال بأن التوزيع قريبا من التماثل وفي هذه الحالة تصبح العلاقة بين مقاييس النزعة المركزية تربطهم العلاقة التالية:

$$\bar{X} - M_0 = 3(\bar{X} - M_e)$$