**المحور الخامس: تحليل البيانات الإحصائية**

1. **تمهيد**
2. **مقاييس النزعة المركزية**
	1. **– المتوسط الحسابي**
	2. **الوسيط**
	3. **المنوال**

**مقدمة:**

 بعد تلخيص البيانات في جداول تكرارية وعرضها في أشكال هندسية بيانية ، يكون الباحث في حاجة إلى أن ينتقل إلى التحليل الاحصائي أو تحليل البيانات لهذا التوزيع والتي تعني إيجاد مقاييس محددة من تلك البيانات نريد اعتمادها في دراسة الفرضية أو المسألة المطروحة من خلال حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة وذلك من خلال ثلاثة مباحث هي:

1. البحث عن القيمة التي تتمركز حولها القيم التي تتخذها الظاهرة في تغيرها من وحدة إلى أخرى وهو ما يعرف بالمتوسطات أو مقاييس النزعة المركزية
2. البحث عن مدى تشتت القيم التي يأخذها المتغير، بالنسبة للقيمة المتوسطة وإذا ما إذا كان هناك قيم شاذة أم لا وهو ما يعرف بمقاييس التشتت
3. البحث عن تماثل التوزيع حول القيمة المتوسطة وهو ما يعرف بالالتواء

مقاييس النزعة المركزية " المتوسطات " Measures of Central Tendency

 الفكرة الأساسية التي تعتمد عليها مقاييس النزعة المركزية هي تمثيل مجموعة كبيرة من البيانات بقيمة واحدة ، وهذه القيمة عادة تكون في وسط البيانات أي في مركزها وهي القيمة التي تميل إليها بقية القيم وتتجمع حولها

**أولا: الوسط الحسابي (المتوسط الحسابي)**

ويعرف الوسط الحسابي لمجموعة من القيم بأنه القيمة التي لو حلت محل جميع القيم لا يتغير مجموعها. كما يعرف الوسط الحسابي بأنه القيمة التي تمثل مركز ثقل البيانات أي نقطة الارتكاز التي يحصل عندها التوازن

**طرق حساب الوسط الحسابي :**

1. **الوسط الحسابي لبيانات غير مبوبة:** ونقصد بذلك أن جميع المفردات التي نريد حساب متوسطها يمكن حسابها مباشرة دون إعداد جدول توزيع تكراري لها وبهذه الطريقة فإن الوسط الحسابي لمجموعة من القيم هو :

 **الوسط الحسابي =** $\frac{القيم مجموع}{القيم عدد}$

$$\overbar{x}= \frac{\sum\_{}^{}x}{N}= \frac{x\_{1}+x\_{2}+….x\_{N}}{N} $$

1. **الوسط الحسابي لبيانات مبوبة:**

أ - إذا كانت البيانات مبوبة في جدول توزيع تكراري بحيث تمثل كل فئة من فئات الجدول قيمة واحدة فقط ، ففي هذه الحالة قيمة الوسط الحسابي تعطى بالعلاقة التالية :

حيث:

$ x\_{i} $مركز الفئة

 $f\_{i}$ التكرار المطلق للفئةi

 $∑f=N$ مجموع التكرارات

**الوسط الحسابي المرجح :** إذا كانت القيم المشاهدة ليس لها نفس الأهمية أو الوزن ، فعندئذ لحساب الوسط الحسابي يجب ألا نعامل جميع القيم نفس المعاملة بل نكرر كل قيمة عدداً من المرات حسب أهميتها أو وزنها أي نرجح كل قيمة بوزنها، وعليه يطلق على الوسط الحسابي في هذه الحالة الوسط

الحسابي المرجح ويرمز له بالرمز  **x**w حيث:

 = $\frac{\sum\_{i=1}^{n}xiwi}{\sum\_{i=1}^{n}wi}$ **x**w



حيث w هي الأوزان المختلفة المقابلة لكل قيمة .

 $x1$هي قيمة المتغير الأول $w1$هي وزن أوو معامل القيمة الأولى

**خصائص الوسط الحسابي:**

* سهولة حسابه والتعامل معه جبريا
* يدخل في حسابه جميع القيم
* لا يحتاج في حسابه ترتيب البيانات
* يعتبر أكثر مقاييس النزعة المركزية استخداما في عمليات الإحصاء الاستدلالي.

عيوبه :

* يستخدم في البيانات الكمية فقط
* يتأثر بالقيم الشاذة أو المتطرفة
* لا يمكن استخدامه في حالة الفئات المفتوحة لأن حسابه يعتمد على معرفة مراكز الفئات

**– الوسيط ( The Median)** :

 هو المفردة التي تقسم مفردات العينة بعد ترتيبها تصاعديا أو تنازليا إلى قسمين متساويين

الوسيط لبيانات غير مبوبة

الوسيط يحسب كالتالي:

* نرتب البيانات تصاعدياً أو تنازلياً.
* نوجد موقع الوسيط (𝒏+𝟏)/𝟐 .

أ-إذا كان n عدد فردي فإن الناتج يكون عدد صحيح و بالتالي الوسيط هو القيمة التي يكون ترتيبها (𝒏+𝟏)/𝟐 .

أو أن الوسيط يساوي: **x** ($\frac{n+1}{2})$ **m** =

ب-إذا كان n عدد زوجي فإن الناتج يكون عدد غير صحيح و بالتالي الوسيط هو الوسط الحسابي للقيمتين اللتين يكون ترتيبهما 𝒏/𝟐 +𝟏)) و 𝒏/𝟐.

أو $\frac{ x \left(1+\frac{n}{2}\right)+ x (\frac{n}{2}) }{2}$ m =

**حساب الوسيط بيانيا:**

يرسم منحني التكرار المتجمع الصاعد على المحور العمودي، والحدود العليا للفئات ترسم على المحور الأفقي.

يحسب ترتيب الوسيط $\frac{\sum\_{}^{}f\_{i}}{2}$، ويثبت على المحور العمودي، ومنه يرسم خط افقي مواز للمحور الافقي سيتقاطع مع منحنى التكرار المتجمع الصاعد. من نقطة التقاطع ينزل عمود على المحور الأفقي وسيتقاطع معه في نقطة

هي الوسيط.



ويمكن ايجاد الوسيط ايضا

• يرسم منحني التكرار المتجمع الصاعد على المحور العمودي، وحدود الفئات ترسم على المحور الافقي.

يرسم منحنى التكرار المتجمع النازل على المحور العمودي، وحدود الفئات ترسم على المحور الأفقي.

من نقطة تقاطع الرسمين ينزل عمود على المحور الأفقي وسيتقاطع معه في نقطة هي الوسيط.

****

**بعض خصائص الوسيط :**

1 – لا يتأثر الوسيط بالقيم الشاذة أو المتطرفة. وهذا منطقي لأنه يقع في منتصف القيم، والقيم الشاذة إما أن تكون في أول القيم أو أخرها (بعد ترتيب القيم تصاعدياً أو تنازلياً). ففي المثال التالي لدينا عدة مجموعات من القيم المرتبة.

2-يمكن حساب ومعرفة الوسيط في حالة البيانات الوصفية الترتيبية

- 3يمكن استخدامه في الجداول ذات الفئات المفتوحة . ..

4 -مجموع الانحرافات المطلقة للقيم عن وسيطها أقل ما يمكن

**ثالثا: المنوال**:

 يعبر المنوال عن القيمة الأكثر تكرارا أو شيوعا من بين قيم المشاهدات

**حساب المنوال لبيانات غير مبوبة**: حسب التعريف هي القيمة الأكثر تكرارا

مثال: (1)

البيانات التالية تمثل أعمار مجموعة من الناخبين :

25 , 29 , 34 , 29 , 36 , 42 , 29 , 50 , 29 , 36

فما هو منوال هذه الأعمار ؟

الحل :

بما أن العمر 29 سنة هو العمر الذي تكرر أكثر من غيره من الأعمار (تكرر 4 مرات) فإن : منوال العمر = 29 سنة

**استخراج المنوال للقيم المبوبة :**

**البيانات الوصفية أو الكمية المنفصلة** : يحسب المنوال من التعريف مباشرة أي القيمة التي يقابلها أكبر تكرار

**حساب المنوال في حالة البيانات الكمية المتصلة:** توجد عدة طرق تقريبية لحساب المنوال منها:

**طريقة بيرسون** :

وتسمى طريقة الفروق لبيرسون نسبة ألي كارل بيرســـون الذي توصل لها. وهو يرى أن الذي يحدد موقع المنوال داخل الفئة المنوالية هو الفرق بين تكرار الفئة المنواليـة وتكراري الفئتين المجاورتين أي السابقة واللاحقة لها .

وتعطى بالعلاقة:

$$D=h\_{k}+ \left(\frac{∆\_{1}}{∆\_{1}+∆\_{2}}\right)×l\_{k} $$

حيث:

 نحدد الفئة المنوالية و هي الفئة التي لها أكبر تكرار.

$∆\_{1}$**: الفرق بين تكرار الفئة المنوالية والتي قبلها**

$∆\_{2}$**: الفرق بين تكرار الفئة المنوالية والتي بعدها**

**:**$ h\_{k}$ **الحد الأدنى للفئة المنوالية**

$l\_{k}$ **: طول الفئة المنوالية**

**خصائص المنوال**:

سهولة حسابه و إيجاده

لا يتأثر بالقيم الشاذة.

يمكن حسابه في الجداول التكرارية المفتوحة.

يمكن إيجاده لجميع أنواع البيانات الوصفية و الكمية.

يمكن إيجاده من خلال الرسم

يعتبر المقياس الوحيد الذي يمكن استخدامه للبيانات الاسمية

**عيوبه:**

لا تدخل في حسابه جميع البيانات.

قد لا يقع في مركز البيانات بل في طرفها.

تتغير قيمته باختلاف طريقة اختيار الفئات.

يصعب التعامل معه في الإحصاء الاستدلالي لأنه قد تكون له أكثر من قيمة.