



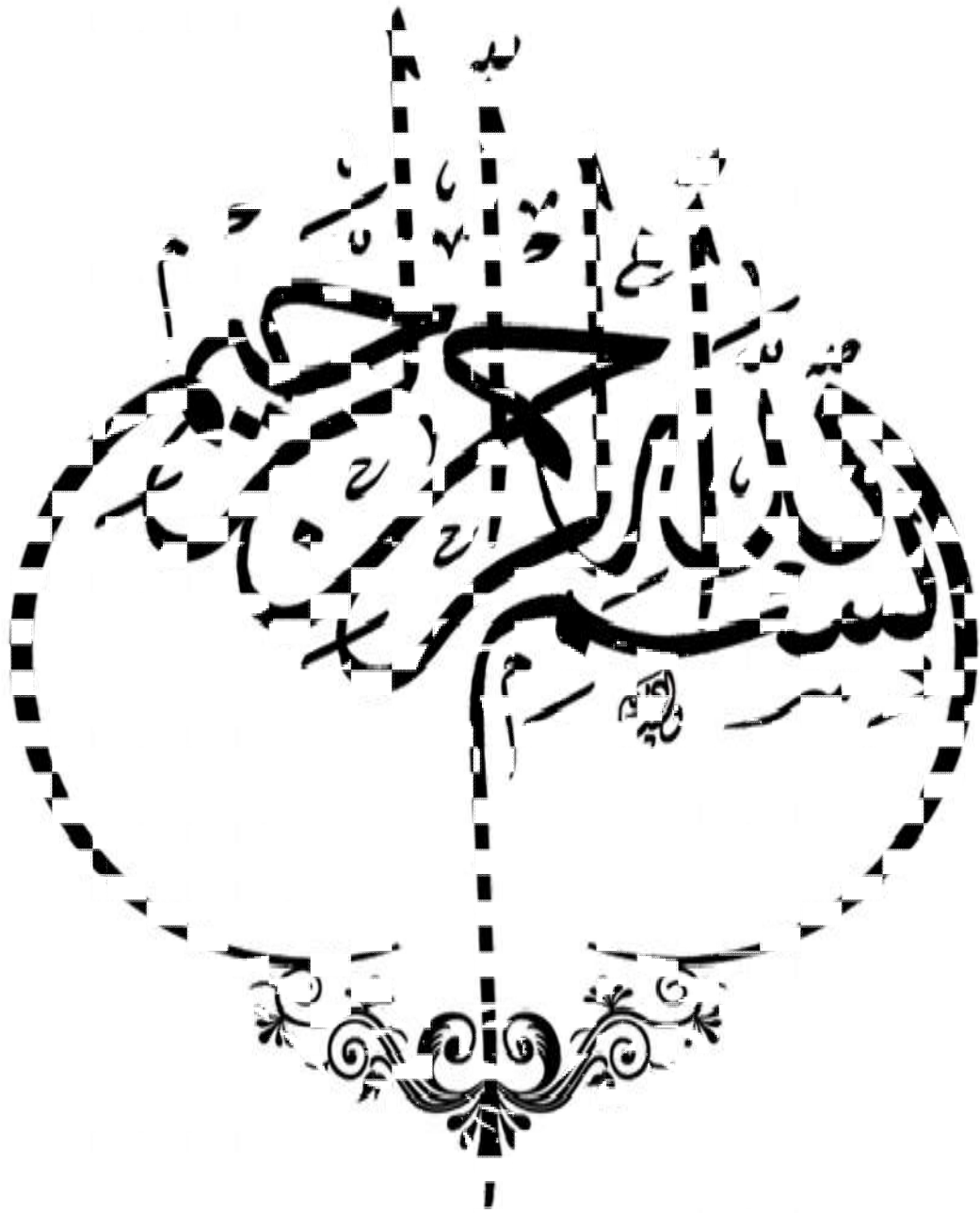
وزارة التعليم العالي والبحث العلمي
جامعة الشهيد حمه لخضر - الوادي
معهد العلوم الإسلامية
قسم الشريعة



محاضرات في الإحصاء الوصفي

مطبوعة موجهة لطلبة سنة أولى ماستر معاملات مالية معاصرة
إعداد الأستاذ: محمد زواري فرحات.

الموسم الدراسي: 1442-1443 هـ / 2021-2022 م



المحور الأول

مفاهيم عامة حول الإحصاء الوصفي

تمهيد وتقسيم:

تعتمد البحوث في العلوم الإنسانية كثيرا على الجوانب المادية والإجرائية والتي هدفها دعم الدراسات النظرية، وذلك عن طريق جمع البيانات وتلخيصها وتمثيلها وعرضها، ثم تحليلها وتفسيرها وفق أساليب إحصائية معينة، للتوصل إلى اتخاذ القرار، وهو مرتبط الفرس والغاية التي وجد من أجلها علم الإحصاء، ما جعله من العلوم الضرورية والهامة التي يستخدمها جل الباحثين في شتى الميادين والمجالات والعلوم من أجل الوصول إلى نتائج تعتمد الموضوعية وتتسم بالمصداقية، تدعم فرضيات البحث التي أعدها الباحث أو ترفضها، من خلال تزويده بالأدوات التي تساعد على تحليل المعطيات بشكل علمي دقيق.

الموضوع الأول: النشأة التاريخية للعلوم الإحصائية وتطورها.

الموضوع الثاني: تعريف علم الإحصاء.

الموضوع الثالث: أقسام علم الإحصاء.

الموضوع الرابع: تعريف المصطلحات المستخدمة في علم الإحصاء.

الموضوع الأول

النشأة التاريخية للعلوم الإحصائية وتطورها

عبر السنين المتلاحقة تطور علم الإحصاء ككيان بدائي وسيادي تستخدمه الدولة حصراً، من خلال تعداد الجيوش وموظفيها والأراضي الإقطاعية التي تسيطر عليها والهدف من ذلك كان تحديد حجم الضرائب التي يجب أن يتم فرضها من أجل تغطية تكاليف ونفقات التي تنفقها الدولة لتمويل الجيش والقطاع الإداري التابع لها، لذلك كان يطلق عليه في العصور الوسطى باسم "علم الدولة" وهو ما جعل علم الإحصاء يشتق من الكلمة اللاتينية Status والتي تعني "الدولة" في اللغة اللاتينية.

استخدم علم الإحصاء بمفهومه البدائي أيضاً في عصر الدولة الإسلامية بداية بعهد الفاروق الذي اعتمد الدواوين والتي كان من بين مهامها إعداد بيانات وصفية دورية للمساكين والأرامل والأغنياء وعدد التجار وعدد موظفي الدولة في الدواوين، من أجل خلق موازنة في بيت المال.

في العصر الحديث كان الفضل يرجع لعلماء النفس والاجتماع في تطوير علم الإحصاء وتنظيمه، من أمثال عالم النفس الإنجليزي جالتون في تطبيق الانحدار والارتباط والذي طوره بعده كارل بيرسون، من رواده أيضاً عالم الفلك الاجتماعي كيتله الذي استخدم الإحصاء في تفسير الظواهر الاجتماعية.

توسع مجال استخدام العلوم الإحصائية في القرن الماضي إلى علم الأحياء، الطب، الزراعة والعلوم الاجتماعية والاقتصادية، كان نتيجتها زيادة تخصص تلك العلوم؛ كما هو الحال في الاقتصاد القياسي وبحوث العمليات في علم الاقتصاد، كميدان تخصصه وهدفه صنع القرارات الاقتصادية.

الموضوع الثاني

تعريف علم الإحصاء

يرتبط مفهوم الإحصاء عند الغالبية بأنه عد وحصر الأشياء والتعبير عنها بأرقام، وهو المفهوم اللغوي لمصطلح الإحصاء، بينما يتجاوز المعنى الحقيقي لعلم الإحصاء ذلك.

من تعريفات علم الإحصاء:

(علم الإحصاء فرع من فروع علم الرياضيات يشمل النظريات والطرق الموجهة نحو جمع البيانات، وصف البيانات، الاستقراء، صنع القرارات، ويتميز باستخدام الأرقام والرموز والدوال الرياضية والمقاييس والجداول والرسومات البيانية)

ويعرف أيضا على أنه:

(الإحصاء هو العلم الذي يبحث طرق ووسائل جمع البيانات الكمية عن مختلف الظواهر الاقتصادية والاجتماعية وغيرها، وترتيب هذه البيانات وتبويبها وتحليلها وتفسيرها وتقديمها بأشكال وصور ملائمة بهدف تسهيل اتخاذ القرار على أساس سليم)

وتجدر الإشارة في هذا الشأن بأن مصطلح الإحصائيات يقصد به البيانات العددية المتعلقة بموضوع ما والمنظمة في جداول أو رسومات بيانية حول نشاط أو قطاع معين في الدولة، مثال ذلك إحصائيات السكان للتعبير عن مجموعة البيانات الخاصة بالسكان في بلد ما من العدد الإجمالي للسكان، توزيع السكان حسب العمر والجنس والتوزيع الجغرافي للسكان حسب الولايات أو المقاطعات.

الموضوع الثالث

أقسام علم الإحصاء

يقسم علماء الإحصاء طرق تحليل البيانات الإحصائية إلى قسمين:

القسم الأول: التحليل الوصفي للبيانات الإحصائية.

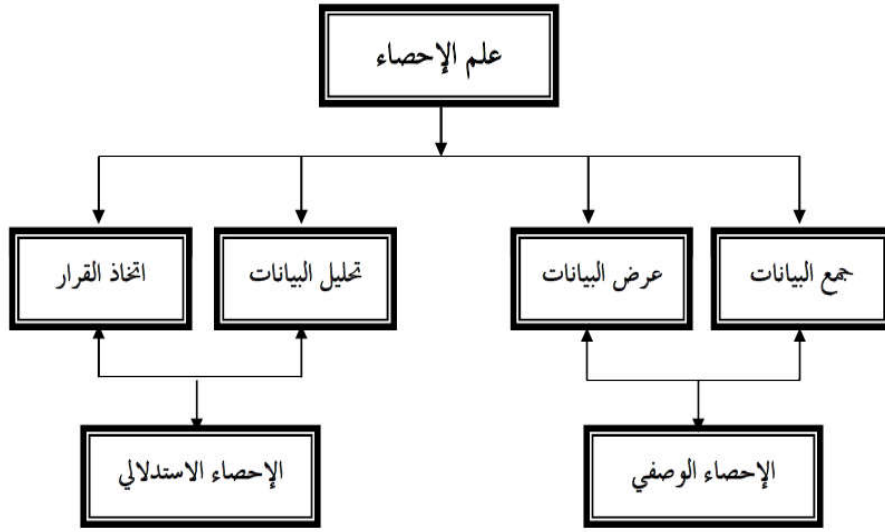
يهدف هذا القسم إلى وصف الظاهرة المدروسة عن طريق جمع وتنظيم وتصنيف وعرض البيانات المتوفرة، واستكشاف خصائصها الأساسية وتلخيصها في صورة مؤشرات رقمية، كمية كانت أو نوعية بطرق مناسبة كأشكال هندسية ورسوم بيانية وتوزيعات تكرارية، أو من خلال استخدام مقاييس النزعة المركزية ومقاييس التشتت أو مقاييس العلاقات الارتباطية بين المتغيرات.

القسم الثاني: التحليل الاستدلالي للبيانات الإحصائية.

يركز هذا النوع من التحليل الإحصائي للوصول إلى استنتاجات ودلائل حول خصائص المجتمع المدروس عن طريق البيانات المتوفرة من العينة المسحوبة من المجتمع المراد تقدير معالمه أو اختبار فرضيات بشأنه، أي أنه يهدف إلى التعميم من العينة إلى المجتمع.

وهكذا تصنف طرق معالجة وتحليل البيانات الإحصائية إلى طرق وصفية متمثلة في إيجاد قيم المقاييس والارتباطات التي تتحدد قيمتها من البيانات التي حصلنا عليها من العينة المدروسة كالمتوسط الحسابي للبيانات أو مدى تشتت وتجانس تلك البيانات، أو بعض المقاييس التي تظهر مدى اقتران أو تباعد تلك البيانات بعضها من بعض، وطرق استدلالية والتي تؤدي إلى الاستنتاجات التي يتوصل إليها الباحث من خلال تطبيق بعض الاختبارات الاستدلالية كاختبار χ^2 وغالبا ما تكون على شكل تقديرات وتنبؤات أو تعميمات أو قرارات رفض أو قبول الفرضيات الإحصائية.

أقسام علم الإحصاء



الموضوع الرابع

تعريف المصطلحات المستخدمة في علم الإحصاء

من المصطلحات المستخدمة في علم الإحصاء:

المصطلح الأول: الوحدة الإحصائية.

تسمى أيضا بالعنصر أو المفردة التي تجري عليها الدراسة الإحصائية أو المعاينة والتي نتحصل منها على المعلومات والبيانات، وهي عنصر فعال في عملية التحليل، فيشترط في الوحدة الإحصائية أن تكون خاضعة لتعريف دقيق وواضح، فهي قد تكون شيئا حيويا مثل: شخص، طالب، موظف وقد تكون شيئا ماديا مثل: مؤسسة، دولة وقد تكون شيئا معنويا مثل: فكرة، الدافعية.

المصطلح الثاني: المجتمع الإحصائي.

هو عبارة عن مجموع الوحدات الإحصائية المراد دراستها، والتي تشترك فيما بينها في الصفة الأساسية المراد تحليلها، ويشترط في المجتمع الإحصائي أن يكون معرفا تعريفا جيد.

المصطلح الثالث: العينة الإحصائية.

هي جزء من مجتمع الظاهرة قيد الدراسة تؤخذ بطريقة معينة بحيث تكون ممثلة تمثيلا صحيحا للمجتمع بقصد التعرف على خصائص هذا المجتمع.

المصطلح الرابع: المتغير الإحصائي.

هو الخاصية التي يرغب الباحث في دراستها أو هو القاسم المشترك بين عناصر المجتمع، وتكون قابلة لتغيير من فرد إلى آخر ومن مشاهدة إلى أخرى، فهي تسمح بالتفريق بين وحدات المجتمع.

ويمكن تقسيم المتغيرات الإحصائية إلى قسمين:

أولاً: متغيرات كمية.

وتسمى أيضا المتغيرات النوعية أو المتغيرات الغير كمية، وهي تلك المتغيرات التي لا يمكن قياسها كما، أي غيرا قابلة للقياس بل يقاس تكرارها فقط وهي عبارة عن صفات، وتنقسم بدورها إلى قسمين:

1- متغيرات كمية قابلة للترتيب: وهي تلك المتغيرات التي يمكن ترتيبها حسب رتبة ما، إما

تصاعديا أو تنازليا، مثل مستوى التأهيل العلمي.

2- متغيرات كمية غير قابلة للترتيب: وهي تلك المتغيرات الوصفية التي لا يمكن ترتيبها مثل

الجنسية، الجنس، الحالة العائلية.

ثانيا: متغيرات كمية.

هي عبارة عن متغيرات تأخذ طابع عددي أي يكون معبر عنها في شكل أرقام، فهي تلك المتغيرات التي يمكن قياسها، وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام مثل الانتاج، الاستهلاك.

والمتغيرات الكمية تنقسم بدورها إلى قسمين:

- 1- متغيرات كمية منقطعة: هي تلك المتغيرات التي يتم التعبير عنها على شكل أرقام صحيحة لا يمكن تجزئتها، مثل عدد الأطفال في الأسرة الواحدة.
- 2- متغيرات كمية منقطعة: هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد الغير متناهي لهذه القيم تقسم مجال الدراسة إلى مجالات جزئية تسمى فئات، مثل الطول، الوزن.

المحور الثاني

عرض البيانات الإحصائية

تمهيد وتقسيم:

بعد الانتهاء من جمع البيانات والمعلومات يقوم الباحث بعملية تفرغ هذه البيانات التي نجد الكثير منها في صورة غير معبرة وغير منظمة مما يصعب استيعابها والمقارنة بين مفرداتها واستنتاج المعلومات منها، لذا يجب تنظيمها وترتيبها وعرضها بطرق مناسبة تسهل دراستها والاستفادة منها، ويتم عرض هذه البيانات وفق عدة طرق، حسب احتياجات البحث ومقصد الباحث من الدراسة.

الموضوع الأول: التكرار المطلق.

الموضوع الثاني: التكرار المتجمع الصاعد.

الموضوع الثالث: التكرار المتجمع النازل.

الموضوع الأول

التكرار المطلق

هو عبارة عن صورة تنقل المعلومات دون الانقاص منها من حالتها الأولى إلى حالة جديدة تتسم بالتنظيم والترتيب والسهولة والوضوح، فهو جدول يضم قيم المتغير والتكرارات المقابلة له، ويستخدم هذا التوزيع لوصف البيانات التي تتعلق بظاهرة واحدة فقط، ويحتوي هذا الجدول في صورته البسيطة على عمودين:

أولاً: قيم المتغير الإحصائي.

تظهر هذه القيم في العمود الأول من الجدول وتتمثل في مختلف القيم التي يأخذها المتغير الإحصائي

وتكون مرتبة ترتيباً تصاعدياً ويرمز لها بالرمز X_i حيث i يمثل عدد الأسطر في الجدول، ومثال ذلك:

ثانياً: التكرار.

ويتمثل في عدد المرات التي يتكرر فيها كل قيمة للمتغير ويرمز له بالرمز n_i أو بالرمز f_i في حالة عرض

البيانات في شكل فئات، مثال ذلك:

التكرار	المتغير
F_1	X_1
F_2	X_2
F_3	X_3
F_4	X_4
F_5	X_5

الموضوع الثاني

التكرار المتجمع الصاعد

يمثل التكرار المتجمع الصاعد مجموع القيم التي تقل قيمهم الإحصائية عن القيمة المقابلة يرمز له بالرمز f_i^{\uparrow} أو بالرمز Ni^{\uparrow} ففي حساب التكرار المتجمع الصاعد نبدأ من أعلى الجدول إلى أسفله ونقوم بجمع التكرارات، وفق القاعدة الآتية:

$$f_i^{\uparrow} = f_1 + f_2 + \dots + f_i$$

مثال ذلك:

التكرار المتجمع الصاعد	التكرار	المتغير
F_1	F_1	X_1
$F_1 + F_2$	F_2	X_2
$F_1 + F_2 + F_3$	F_3	X_3
$F_1 + F_2 + F_3 + F_4$	F_4	X_4
$F_1 + F_2 + F_3 + F_4 + F_5$	F_5	X_5

الموضوع الثالث

التكرار المتجمع النازل

يمثل التكرار المتجمع النازل مجموع القيم التي تفوق قيمهم الإحصائية عن القيمة المقابلة يرمز له بالرمز $f_i \downarrow$ أو بالرمز $N_i \downarrow$ ففي حساب التكرار المتجمع النازل نبدأ من أسفل الجدول إلى إعلاه ونقوم بجمع التكرارات، وفق القاعدة الآتية:

$$f_i \downarrow = f_i + f_{i-1} + \dots + f_1$$

مثال ذلك:

التكرار المتجمع الصاعد	التكرار	المتغير
$F_5 + F_4 + F_3 + F_2 + F_1$	F_1	X_1
$F_5 + F_4 + F_3 + F_2$	F_2	X_2
$F_5 + F_4 + F_3$	F_3	X_3
$F_5 + F_4$	F_4	X_4
F_5	F_5	X_5

المحور الثالث

مقاييس النزعة المركزية

تمهيد وتقسيم:

يتم تلخيص البيانات عن طريق مقاييس تساهم في معالجتها إحصائياً، من تلكم المقاييس؛ مقاييس النزعة المركزية وهي القيم التي تقترب منها جمع القيم في التوزيع بهدف تلخيص البيانات في قيمة واحدة، وفيما يلي عرض لأهم هذه المقاييس:

الموضوع الأول: الوسط الحسابي.

الموضوع الثاني: الوسيط.

الموضوع الثالث: المنوال.

الموضوع الأول

الوسط الحسابي

$$\bar{X}$$

أولاً: تعريف الوسط الحسابي.

الوسط الحسابي أو المتوسط الحسابي هو من أهم مقاييس النزعة المركزية التي يعتمد عليها كثيرا لوصف البيانات الكمية ويعرف المتوسط الحسابي بأنه حاصل قسمة مجموع القيم على عددها وفي حالة البيانات المبوبة يتم جمع مراكز الفئات مضروبة في تكرارها وتقسم على مجموع التكرارات.

$$\bar{X} = \frac{\sum f_i x_i}{\sum f_i}$$

وفق القاعدة الآتية:

مثال ذلك:

المتغير	التكرار
X ₁	F ₁
X ₂	F ₂
X ₃	F ₃
X ₄	F ₄
X ₅	F ₅

$$\bar{x} = \frac{x_1 \times f_1 + x_2 \times f_2 + \dots + x_i \times f_i}{f_1 + f_2 + \dots + f_i}$$

مثال:

المتغير	التكرار
5	4
3	3
2	2
4	4
7	4

$$\bar{x} = \frac{5 \times 4 + 3 \times 3 + 2 \times 2 + 4 \times 4 + 7 \times 4}{4 + 3 + 2 + 4 + 4} = \frac{20 + 9 + 4 + 16 + 28}{17} = 4.52$$

ثانيا: خصائص الوسط الحسابي.

- ✓ يأخذ في الاعتبار جميع القيم في حسابه.
- ✓ يمكن حسابه بسهولة وسرعة.
- ✓ يتناسب مع البيانات الكمية.
- ✓ أكثر المقاييس استخداما لوصف البيانات والاختبارات الإحصائية.
- ✓ من السهل فهمه وتفسيره.
- ✓ يتأثر بالقيم الشاذة والمتطرفة.

✓ يصعب حسابه في حالة الفئات التكرارية المبوبة المفتوحة، لأنه من الصعب تحديد مراكز الفئات، وفي هذه الحالة يمكن تحديد مراكز الفئات بصورة تقريبية أو البحث عن الوسيط كبديل عنه.

الموضوع الثاني

الوسيط

 M_e

أولاً: تعريف الوسيط الحسابي.

يعرف الوسيط بالقيمة التي تقسم مجموع القيم إلى نصفين بعد ترتيبها ترتيباً تصاعدياً، ويعبر عنه بالقيمة العددية التي يكون 50% من البيانات أقل منها و 50% من البيانات أكبر منها.

1- في حالة القيم الغير مبوبة يتم حساب الوسيط وفق الخطوات الآتية:

- نرتب القيم ترتيباً تصاعدياً.

- الوسيط هو القيمة $M_e = X_{(n+1)/2}$ اذا كان عدد القيم فردياً.

- الوسيط هو القيمة $M_e = (X_{(n+1)} + X_{(n/2+1)})/2$ اذا كان عدد القيم زوجياً.

مثال:

27-15-11-8-9-6-5-4-2

الوسيط هو القيمة الخامسة التي تساوي 9 لأن عدد القيم فردي تسعة قيم ولأن عدد القيم الأقل منها يساوي عدد القيم الأصغر منها وهي أربعة قيم.

مثال 02:

22-21-20-17-12-11-10-09-7-3

الوسيط هو مجموع القيمتين الخامسة والسادسة مقسوما على اثنين $(12+11)/2 = 11.5$ لأن عدد القيم زوجيا عشرة قيم.

-2 في حالة البيانات المبوبة يتم حساب الوسيط بالقاعدة التالية:

$$Me = L_{me} + \left[\frac{(\sum f)/2 - f_{me-1}^{\uparrow}}{f_{me}} \right] \times C_{me}$$

حيث:

L_{me} : الحد الأدنى للفئة الوسيطة.

F_{me-1}^{\uparrow} : التكرار المتجمع الصاعد للفئة قبل الوسيطة.

f_{me} : تكرار الفئة الوسيطة.

C_{me} : طول الفئة الوسيطة.

حيث يتم حساب طول الفئة بفارق الحد الأعلى من الأدنى، مثال ذلك:

$$C = 40 - 10 = 30 \text{ هو: طول الفئة } [10-40]$$

ويتم تحديد الفئة الوسيطة بطريقة قسمة جميع التكرارات بالنصف ثم تحديد التكرار المتجمع الصاعد الذي يكون أكبر مباشرة منه.

ثانيا: خصائص الوسيط الحسابي.

✓ لا يأخذ في عين الاعتبار جميع قيم التوزيع في حسابه، فهو يعتمد على قيمة واحدة أو قيمتين.

✓ يمكن حسابه بسرعة وسهولة.

✓ لا يتأثر بالقيم الشاذة أو المتطرفة.

الموضوع الثالث

المنوال

M_0

أولاً: تعريف المنوال.

يعرف المنوال بأنه القيمة الأكثر تكراراً أو انتشاراً من بين قيم التوزيع

1- في حالة القيم الغير مبوبة يتم حساب المنوال باختيار الأكبر قيمة:

مثال:

9 -27-15-11-8-9-6-5-4-2

المنوال يساوي 9 لأنه القيمة الأكثر تكراراً.

مثال 02:

10-11-22-21-20-17-12-11-10-09-7-3

المنوال يساوي قيمتين وهما: 10-11 لأنهما الأكثر تكراراً.

مثال 03:

30-13-11-09-08-05-02-01

لا يوجد منوال، لأنه جميع التكرارات متساوية.

2- في حالة البيانات المبوبة يتم حساب المنوال بالقاعدة التالية:

$$M_0 = L_{mo} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2} \right] \times C_{mo}$$

حيث:

L_{mo} : الحد الأدنى للفئة المنوالية.

1 : الفارق بين تكرار الفئة المنوالية والفئة التي تسبقها .

2 : الفارق بين تكرار الفئة المنوالية والفئة التي تليها.

C_{mo} : طول الفئة المنوالية.

و تكون الفئة المنوالية هي الفئة التي تقابل أكبر تكرار.

المحور الرابع

مقاييس التشتت

تمهيد وتقسيم:

سبق وأن تطرقنا إلى مقاييس النزعة المركزية التي تصف لنا مجموعة من قيم التوزيع في قيمة واحدة، ولكن هذا الوصف غير كافي لإعطاء فكرة دقيقة عن مجموعة من قيم من حيث طبيعة توزيعها، ومدى تقاربها وتباعدها عن بعضها البعض، أو عند المقارنة بين المجموعات، فقد يتساوى متوسط مجموعتين أو أكثر ولكن نجد قيم إحدى المجموعتين متجمعة قريبا من متوسطها، بينما قيم مجموعة أخرى متباعدة عن متوسطها.

ومن هنا جاء مفهوم التشتت لمجموعة من القيم الذي يقصد به درجات التفاوت أو الاختلاف بين قيم هذه المجموعة، فإذا كانت قيم المجموعة متقاربة من بعضها البعض يكون التشتت صغيرا، وإذا كانت القيم متباعدة عن بعضها البعض يكون التشتت كبيرا.

الموضوع الأول: التباين.

الموضوع الثاني: الانحراف المعياري.

الموضوع الأول

التباين

أولاً: التباين.

هو احد مقاييس التشتت وأكثرها استخداما، وهو مجموع مربعات انحرافات القيم عن متوسطها الحسابي مقسوما على عددها، ويرمز له بالرمز δ^2

ويحسب التباين وفق القاعدة الآتية:

$$\delta^2 = \frac{\sum (X_i - \bar{X})^2 \times f_i}{\sum f_i}$$

الموضوع الثاني

الانحراف المعياري

يعرف الانحراف المعياري بأنه الجذر التربيعي لمعدل مجموع مربعات انحرافات القيم عن متوسطها، أي أن الانحراف المعياري هو الجذر التربيعي للتباين.

$$\delta = \sqrt{\delta^2}$$