المحاضـــــره السابعه (الجزء الاول)

المقاييس الإحصائية للبيانات غير المبوبة

أولا: مقاييس النزعة المركزية

سبق و أن أستعرضنا مراحل البحث العلمى و أتضح لنا أن البحث الإحصائى له نفس المراحل فبعد جمع البيانات و المعلومات Data Collection لا بد من عرض هذة البيانات فى شكل جدولى او فى شكل الرسومات بيانية Data Presentation and Tabulation مما يسهل من فهم و أستيعاب مضمونها . و تأتى بعد ذلك المرحلة التالية وهى تحليل البيانات Data Analysis و التى فيها يتم أستخدام الأدوات الإحصائية المختلفة لوصف البيانات من خلال حساب المقاييس الإحصائية المختلفة التى سوف نستعرضها في هذه المحاضرة بمشيئة الله.

المقاييس الإحصائية:

تتمثل أهمية عملية وصف البيانات كمياً من خلال محاولة الوصول إلى فهم ورؤية أوضح للمعلومة المحتواه فى القيم الكمية للمتغيرات محل الدراسة ومحاولة التعبير عن تلك البيانات الكمية بقيم تصف طبيعة وشكل المتغيرات محل الدراسة بالطريقة التي تمكننا من التعامل معها بشكل أدق وأفضل ويطلق على تلك القيم المقاييس الإحصائية.

المقاييس الإحصائية لم توجد من تلقاء نفسها وأنما دعت الحاجة إلى وجودها حيث تساعدنا فى وصف المتغيرات المختلفة عن طريق معرفة القيم التى تتركز حولها البيانات ومدى التفاوت بين قيم المفردات محل الدراسة وتلك القيم، كما تساعدنا فى المقارنة بين المتغيرات المختلفة من حيث مدى نزعتها نحو مراكز معينة وتحديد مدى تجانس البيانات بعضها مع بعض.

أقسام المقاييس الإحصائية :

تنقسم المقاييس الإحصائية إلى نوعيين رئيسيين هما:

- مقاييس النزعة المركزية Central Tendency Measures
- مقاييس التشتت أو الأنتشار _Dispersion Measures

فى هذه المحاضرة سنتعرض لكيفية حساب مقاييس النزعة المركزية ومقاييس التشتت فى حالة أستخدام البيانات الخام غير المبوبة، أي تلك التى لم يتم تصنيفها فى صورة جداول تكرارية، وذلك هو الأصل فى التحليل الإحصائي للبيانات، لأنة يعطى الصورة الحقيقية للنتائج بدون أي تدخل شخصى فيها، إلا ان ذلك لا يقلل إيضا من أهمية الحاجة لدراسة كيفية حساب المقاييس الإحصائية المختلفة من البيانات المبوبة والتى سنتعرض لها فى المحاضرات التالية إن شاء الله.

اولا- مقاييس النزعة المركزية Central Tendency Measures

نقصد بمقاييس النزعة المركزية تلك القيم الوسطى التى توضح القيمة التي تجمع أكبر عدد من القيم الخاصة بمجموعة معينة عندها . أو هي قيمة تلك الدرجة التي يمكن أن تعتبر ممثلة لكافة الدرجات الموجودة في تلك المجموعة . ولتحديد القيمة المتوسطة للتوزيع يوجد هناك عدة مقاييس أهمها :

- المتوسط الحسابي
- الوسيط
- المنوال (الشائع)

كما يوجد عدة مقاييس أخرى أقل شيوعا مثل:

- الوسط الهندسى
- الوسط التوافقى
- العشير
- المئين

أهمية حساب مقاييس النزعة المركزية :

حساب مقاييس النزعة المركزية يساعد على التالي:

- ايجاد ذلك الرقم المتوسط الذي يدل على خصائص أرقام مجموعة من المجموعات فيكفي أن ننظر الى ذلك الرقم المتوسط لنعرف الكثير عن خصائص هذه المجموعة من الأرقام
- المقارنة بين عدة مجموعات في وقت واحد ، فنقول أن هذه المجموعة أقوى من تلك ، وذلك اعتمادا على مقارنة هذه المتوسطات بعضها ببعض

الوسط الحسابى (المتوسط) Mean

يُعرف المتوسط الحسابي بأنه قيمة التى اذا أعطيت لكل مفرد من مفردات الظاهرة لكان مجموع القيم الجديدة مساويا للمجموع الفعلي للقيم الأصلية للظاهرة، أي أن الوسط الحسابي يساوي مجموع القيم مقسوما على عددها، ويتم حساب المتوسط الحسابي للبيانات غير المبوبة من خلال المعادلة التالية:

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

مثال:

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلى:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعدة	ذي الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

حساب المتوسط الحسابى للمبيعات الشهرية.

ويجب ملاحظة عدة أمور في الوسط الحسابي وهي:

- انه لا يشترط أن يكون المتوسط الحسابي عددا صحيحا.
- ان المتوسط الحسابي دائما محصور بين أقل القيم وأعلاها، ولكن هذا لا يعني أنه يقع في الوسط تماما بين هذين الحدين.
- إن المجموع الجبري لانحراف القيم عن المتوسط يكون دائما صفر.
- ومن أهم خصائص الوسط الحسابى هو تأثره بجميع العمليات الجبرية تجرى على البيانات من إضافة قيمة لجميع البيانات أو طرحها أو ضربها أو قسمتها.

مثال:

بسؤال خمسة أشخاص عن أجرهم الشهرى فكانت إجاباتهم كما يلى بالألف ريال:

3 , 5 , 2, 7,3

المطلوب:

- أحسب متوسط الأجر الشهرى
- وإذا قررت إدارة الشركة زيادة أجورهم أحسب متوسط الأجر الجديد فى الحالتين التاليتين

1- زيادة اجور العاملين بمقدار 2000 ريال

2- زيادة أجور العاملين بنسبة 5 %

مزايا وعيوب المتوسط الحسابي:

المزايا:

- يعد المتوسط الحسابي من اكثر مقاييس النزعة المركزية استخداما، واسهلها فهما وذلك نتيجة لسهولة حسابه
- يدخل في حسابه كل القيم دون اهمال أي قيمة منها.

العيوب:

- يتأثر بالقيم المتطرفة الشاذة قلة أو كثرة، فقد يرتفع لمجرد وجود قيمة مرتفعة، وقد يقل كثيرا لمجرد وجود قيمة واحدة صغيرة وهذا بالتالي يؤدي الى عدم تمثيل المتوسط لواقع المعلومات.
- لايمكن ايجاده من خلال الرسم

الوسيط Median

هو الدرجة التي تتوسط مجموعة من الدرجات المرتبة ترتيبا تصاعديا أو تنازليا، أى هو القيمة التى يصغرها عدد من القيم يساوى العدد الذى يكبر هذة القيمة ويمكن حساب الوسيط باتباع الخطوات التالية:

- ترتيب الدرجات تصاعديا أو تنازليا
- إيجاد ترتيب الوسيط و يقصد بة إيجاد مكان الوسيط، ويختلف ترتيب الوسيط إذ كان عدد المشاهدات فردى أو زوجي كما يلي:

عدد المشاهدات n	ترتيب الوسيط
فردى	$(n+1)/2$
زوجى	يوجد ترتيبين هما $(n/2)+1$, $n/2$

- - إيجاد قيمة الوسيط.

مثال:

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بلألف ريال كما يلى:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعدة	ذي الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

إيجاد قيمة الوسيط للبيانات السابقة.

مزايا وعيوب الوسيط:

المزايا:

- لا يتأثر بالقيم الشاذة.
- يمكن استخدام الوسيط في البيانات الناقصة.
- يمكن الحصول على الوسيط وحسابه من خلال الرسم.
- يمكن استخدام الوسيط في البيانات التي يعرف ترتيبها ولا تعرف قيمتها.

العيوب:

- لا يعتمد على جميع القيم، حيث أنه لا يدخل في حسابه سوى قراءة واحدة أو قراءتين من البيانات كلها.

المنوال Mode

هو القيمة التي تعتبر اكثر القيم شيوعا، وعلى ذلك فتحديده يتوقف على تكرار القيم في المجموعة.

فى نفس المثال السابق للمبيعات الشهرية . أحسب المنوال؟

نجد أن المبيعات الأكثر تكراراً هنا هى 3 ألف ريال لذلك

فان المنوال هنا = 3

وقد يكون في التوزيع منوالين أو أكثر وذلك كالمثال الأتي:

٦ ، 5 ، 5 ، 5 ، 4 ، 4 ، 4

فالمنوال هنا = 4 ، 5 أي أنه يوجد منوالين .

وقد لا يكون في التوزيع منوال وذلك كالمثال الآتي:

2 ، 5 ، 7 ، 9 ، 11

مزايا وعيوب المنوال:

المزايا:

- سهل الحساب سواء بالرسم أو بالحساب
- لا يتأثر كثيرا بالقيم الشاذة
- لا يتأثر كثيرا لو تغيرت قيم بعض مفردات البيانات

العيوب:

- أقل مقاييس النزعة المركزية استعمالا
- عديم الفائدة في البيانات القليلة العدد

الوسط الهندسى Geometric Mean

نتيجة أن الوسط الحسابى يتأثر بالقيم الشاذة دعت الحاجة إلى وجود مقاييس لا تتأثر بقدر الإمكان بالقيم الشاذة والمتطرفة ومن تلك المقاييس الوسط الهندسى GM والذى يكون مفيد فى بعض التطبيقات الاقتصادية ودراسات نمو الظواهر الديموجرافية وكذلك فى حساب الأرقام القياسية، فالوسط الهندسى هو الجذر النونى لحاصل ضرب القيم محل الدراسة ويمكن حسابة من خلال المعادلة التالية:

$$GM = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n}$$

ويحسب الجذر النوني من خلال استخدام الآلة الحاسبة العلمية بكتابة حاصل ضرب القيم محل الدراسة ثم الضغط على الجذر النوني ثم إدخال قيمة n ثم الضغط على يساوي فتظهر بالتالي قيمة الوسط الهندسي.

مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام1427 هـ بلألف ريال كما يلى:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعدة	ذي الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

إيجاد قيمة الوسط الهندسي للبيانات السابقة.

الحل:

يمكن تطبيق المعادلة السابقة على البيانات الموجودة بالمثال ولكن قد يكون الأمر صعب فى حالة ما تكون المشاهدات محل الدراسة (n) كبيرة الحجم . لذا يمكن حساب الوسط الهندسي كما يلى:

$$GM = \sqrt[12]{3\times5\times8\times3\times6\times4\times12\times5\times4\times3\times7\times9} =$$

$$= \sqrt[12]{391910400} = 5.2014$$

خواص الوسط الهندسى:

١. يعطى نتائج أكثر اعتدالا من المتوسط الحسابى
٢. تتوقف قيمتة على سائر القيم دون استثناء أو استبعاد، شأنه شأن الوسط الحسابى
٣. أقل تأثرا بالقيم المتطرفة عن الوسط الحسابى

مزايا وعيوب الوسط الهندسي:

المزايا :

١. أكثر تمثيلا للقيم عن الوسط الحسابى بأعتبار أنة لا يتأثر بالقيم المتطرفة بنفس درجة الوسط الحسابى

٢. يعتبر من أنسب المقاييس لحساب متوسطات النسب ومعدلات النمو

٣. يعتبر من أكثر مقاييس النزعة المركزية ملائمة لحساب الأرقام القياسية للمناسيب

العيوب :

١. لا يمكن حسابه اذا كانت احدى القيم صفر

٢. لا يمكن استخدامه اذا كان ناتج حاصل ضرب قيم المشاهدات محل الدراسة سالب

٣. صعوبة حسابه يدويا وإنما يمكن ذلك بأستخدام الحاسب الألى (الآلة الحاسبة).

تابع المحاضره السابعه (الجزء الثاني)

تابع ... المقاييس الإحصائية للبيانات غير المبوبة

ثانيا: مقاييس التشتت أو الأنتشار Dispersion Measures

كما تميل القيم الى التمركز فانها تميل أيضا إلى التشتت أو الانتشار، فبالتالي فان أي توزيع من القيم له صفة التمركز، وصفة التشتت.

فمقاييس التشتت هي تلك المقاييس التي تعبر عن مدى تباعد القيم أو تقاربها في المجموعات التي يشملها البحث

مثال

مجموعة (أ) : 8 ، 8 ، 8 ، 8 ، 8

مجموعة (ب) : 1 ، 2 ، 3 ، 5 ، 6

نلاحظ أن المجموعة الأولى (أ) لا يوجد بها تشتت، فهذه المجموعة متجانسة.

في حين نلاحظ ان المجموعة الثانية (ب) يوجد بها تشتت

يمكن ان يقاس تشتت البيانات عن طريق مقاييس التشتت المختلفة، وأهم هذه المقاييس:

- المدى
- المدى الربيعي
- الإنحراف عن المتوسط
- التباين
- الإنحراف المعياري

- لماذا نستخدم مقاييس التشتت؟

نستخدم هذه المقاييس اذا كان عندنا مجموعتين ونريد ان نقارن بينهما، وكان المتوسط فيما يبينهما متساوي ، كما في المثال التالي:

مجموعة (أ): (45 ، 50 ، 55) المتوسط هنا = 50

مجموعة (ب): (30 ، 50 ، 70) المتوسط هنا = 50

فلذا لا نستطيع ان نقول هنا ان المجموعتين متساويتين لأننا إذا رجعنا الى المجموعتين وجدنا انهما مختلفتين في الدرجات رغم تساوي المتوسطين حيث أن المتوسط الحسابي في المجموعتين يساوي (50) .

لكن اذا استخدمنا احد مقاييس التشتت مثل المدى والذي يحسب من خلال العلاقة التالية: المدى = أعلى درجة – أقل درجة

وعلى ذلك فإن:

مدى مجموعة (أ) = 55 - 45 = 10

مدى مجموعة (ب) = 70 – 30 = 40

نرى ان درجة التشتت في المجموعة (أ) أقل منها في المجموعة (ب)، أي ان المجموعة (أ) تكون أكثر تجانسا من المجموعة (ب)

المدى Range

المدى هو الفرق بين أعلى درجة وأقل درجة في التوزيع.

ويعتبر المدى الوسيلة المباشرة لمعرفة مدى تقارب القيم أو تباعدها في أي توزيع، وهو وسيلة سهلة، إلا أنها أقل الوسائل دقة وذلك لأن حسابه يتوقف على قيمتين فقط من قيم المجموعة، ولا يهتم مطلقا بما بينهما من قيم أخرى.

فالمدى لا يصلح الا اذا اراد الباحث أن يأخذ فكرة سريعة عن مدى تشتت بينات التوزيع موضع الدراسة، الا أن استخدامه والاعتماد عليه قد يؤديان الى نتائج خادعة، وخاصة اذا كان هناك انفصال بين الدرجات المتطرفة وباقي الدرجات موضع البحث.

مثال:

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعده	ذي الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

حساب المدى للمبيعات الشهرية.

الحل:

نلاحظ أن أكبر قيمة هى 12 وأقل قيمة للمبيعات الشهرية هى 3 لذلك يكون المدى 9

Range=12-3=9

عيوب المدى:

نجد أن من أهم عيوب المدى أنة يتم حسابة بناءا على أكبر و أصغر قيميتين وبالتالى فى حالة كونهما أو أحدهما متطرفتين أو قيم شاذة فأن المدى يعطى نتائج مضللة.

- متوسط الانحرافات المطلقة Average Absolute Deviation

متوسط الانحرافات المطلقة AAD هو ذلك المقياس الذي يقيس تباعد كافة القيم عن المتوسط الحسابي .

وعلى الرغم من أن حساب نصف المدى الربيعي يقضي على أثر القيم المتطرفة، والتي تؤثر على حساب المدى المطلق، إلا أنها جميعا (المدى، ونصف المدى الربيعي) يتناولان التباعد بين قيمتين فقط (أعلى قيمة وأدنى قيمة) في المدى، (وقيمة الربيع الأدنى وقيمة الربيع الأعلى) في نصف المدى الربيعي، وذلك من بين القيم موضع الدراسة، أما بقية القيم تبقى مهملة .

وهذا ما أدى الى تطبيق متوسط الانحرافات المطلقة AAD الذي يقيس تباعد كافة القيم عن متوسطها الحسابي.

ويمكن حساب متوسط الأنحرافات المطلقة من خلال المعادلة التالية :

$$AAD = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بلألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعده	ذي الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

أحسب متوسط الأنحرافات المطلقة للمبيعات الشهرية.

- التباين والانحراف المعياري:

التباين Variance هو متوسط مربعات انحرافات القيم عن وسطها الحسابى. ويرمز له بالرمز (تقراء سيجما تربيع) وذلك إذا كان محسوب لبيانات المجتمع أما فى حالة حسابة لبيانات عينة من المجتمع فيرمز له بالرمز S^2 .

الانحراف المعياري Standard Deviation وهو الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابى أى هو جذر التباين لذلك يرمز له بالرمز (تقراء سيجما) و ذلك إذا كان محسوب لبيانات المجتمع أما فى حالة حسابة لبيانات عينة من المجتمع فيرمز له بالرمز S .

ويعتبر الانحراف المعياري والتباين من أهم مقاييس التشتت جميعا أو اكثرها استعمالا، وهما قريبين في خطوات ايجادهما من الانحراف عن المتوسط.

فالتباين واللإنحراف المعياري يختلف عن الإنحراف عن المتوسط في طريقة التخلص من اشارات الفروق بين القيم والمتوسط الحسابي، فبينما نتخلص من هذه الاشارات في طريقة الانحراف عن المتوسط بإهمال الاشارات كلية، نحتال على ذلك في طريقة التباين والانحراف المعياري بتربيع هذه الفروق (أي نضربها في نفسها) فتصبح بالتالي جميع الاشارات موجبة.

حساب التباين والانحراف المعياري :

- يمكن حساب التباين من خلال المعادلة التالية:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

- بالتالى يكون حساب الإنحراف المعياري كما يلي:

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

مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلى:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذو القعده	ذو الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

أحسب قيمة التباين وقيمة الإنحراف المعياري للمبيعات الشهرية.

ملاحظة هامة:

يعتبر من أهم خصائص الانحراف المعيارى هو عدم تأثره بعمليات الجمع والطرح وإنما يتأثر فقط بعمليات الضرب والقسمة.

فنلاحظ عدم تغير قيمة الانحراف المعيارى في حالة الجمع أو الطرح وإنما تظل قيمتة كما هى بالرغم من طرح مقدار ثابت من جميع قيم التوزيع.

أما في حالة الضرب أو القسمة فنلاحظ تغير قيمة الانحراف المعيارى وهى نفس قيمة الانحراف المعيارى القديمة مضروبة فى القيمة التي ضرب فيها أو قسم عليها.

مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلى:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذو القعده	ذو الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

فإذا تم طرح 2 من جميع بيانات المبيعات الشهرية أى تم تخفيض المبيعات الشهرية بمقدار 2 أحسب قيمة الانحراف المعيارى الجديد؟

نلاحظ عدم تغير قيمة الانحراف المعيارى وإنما ظلت قيمتة كما هى بالرغم من طرح مقدار ثابت 2 من جميع قيم المبيعات الشهرية.

مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلى:

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذو القعده	ذو الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

أحسب قيمة الانحراف المعيارى للمبيعات الشهرية إذا تم زيادة المبيعات الشهرية إلى ثلاث أمثال الموجود حالياً ؟

نلاحظ تغير قيمة الانحراف المعيارى وهى نفس قيمة الانحراف المعيارى القديمة مضروبة فى 3

وبالتالى يمكن أن نكون حصلنا ع لى كافة المقاييس الإحصائية الوصفية التى تصف المبيعات الشهرية فكانت كما يلى:

المتوسط	الوسيط	المنوال	الوسط الهندسى
5.75	5	3	5.20114

المدى	متوسط الانحرافات المطلقة	التباين	الانحراف المعيارى
9	2.20833	7.840909	2.80016