الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي و البحث العلمي

جامعة د.مولاي الطاهر سعيدة

كلية العلوم الاقتصادية والعلوم التجارية و علوم التسيير

قسم : العلوم الاقتصادية

مطبوعة موجهة لطلبة السنة أولى جذع مشترك بعنوان :

# محاضرات في الإحصاء 1

من إعداد:

د.جلولي نسيمة

السنة الجامعية 2022–2023

| | |
|---|---|
| **مقدمة عامة** | 08 |
| **الفصل الأول**: مقدمة في علم الإحصاء | 09 |
| **I.** مفهوم علم الإحصاء | 10 |
| **1**. الإحصاء الوصفي | 12 |
| **2**. الإحصاء الإستدلالي | 13 |
| **II.** مفاهيم أساسية في علم الإحصاء | 14 |
| **1**. الطريقة الإحصائية | 14 |
| **2**. المجتمع الإحصائي | 15 |
| **3**. العينة الإحصائية | 16 |
| **4**. البيانات الإحصائية | 16 |
| **5**. المتغير الإحصائي | 19 |
| **6**. المعلمة | 21 |
| **7**. الإحصاءة | 21 |
| **III.** الرموز الإحصائية | 22 |
| **1**. المتغير الإحصائي | 22 |
| **2**. مجموع مربع المشاهدات | 23 |
| **3**.حاصل ضرب قيم متغيرين | 24 |

| | |
|---|---|
| 4. حاصل قسمة قيم متغيرين | 24 |
| 5. بعض القواعد تخص عملية الجمع و الطرح | 25 |
| **الفصل الثاني**: تبويب البيانات و العرض البياني | 27 |
| I. تبويب البيانات | 28 |
| 1.التوزيع التكراري البسيط | 28 |
| 1.1. البيانات الوصفية | 29 |
| 1. 2.البيانات الكمية | 30 |
| 2.التوزيع التكراري المتجمع | 35 |
| 2. 1. التوزيع التكراري المتجمع الصاعد | 35 |
| 2. 2. التوزيع التكراري المتجمع النازل | 36 |
| 3. التوزيع التكراري النسبي | 37 |
| 4. التوزيع التكراري النسبي المتجمع | 38 |
| II. العرض البياني | 39 |
| 1. الأعمدة البيانية | 39 |
| 1.1 . الأعمدة البيانية البسيطة | 39 |
| 1. 2. الأعمدة البيانية المزدوجة | 40 |
| 1. 3. الأعمدة البيانية المجزأة | 41 |
| 2.الشكل الدائري | 42 |

| | |
|---|---|
| **3**.المدرج التكراري | 44 |
| **4**.المضلع التكراري | 45 |
| **5**.المنحنى التكراري | 46 |
| **6**.المدرج التكراري الصاعد و النازل | 48 |
| **6. 1**. المدرج التكراري الصاعد | 48 |
| **6. 2**. المدرج التكراري النازل | 50 |
| **7**.منحنى التكرار المتجمع الصاعد و النازل | 50 |
| **7. 1**. منحنى التكرار المتجمع الصاعد | 50 |
| **7. 2**. منحنى التكرار المتجمع النازل | 51 |
| **الفصل الثالث**: مقاييس النزعة المركزية و غير المركزية | 52 |
| **I**. مقاييس النزعة المركزية | 53 |
| **1**. الوسط الحسابي | 53 |
| **1. 1**. الوسط الحسابي في حالة البيانات غير المبوبة | 53 |
| **1. 2**. الوسط الحسابي في حالة البيانات المبوبة | 54 |
| **2**. الوسيط | 57 |
| **2. 1**. الوسيط في حالة البيانات غير المبوبة | 57 |
| **2. 2**. الوسيط في حالة البيانات المبوبة | 58 |
| **2. 3**. تحديد الوسيط بيانيا | 61 |

| | |
|---|---|
| **3**. المنوال | 64 |
| **3**. **1**. المنوال في حالة البيانات غير المبوبة | 64 |
| **3**. **2**. المنوال في حالة البيانات المبوبة | 64 |
| **3.3** . تحديد المنوال بيانيا | 65 |
| **4**. العلاقة ما بين الوسط الحسابي، الوسيط و المنوال | 66 |
| **5**.الوسط الهندسي | 68 |
| **5**. **1**. الوسط الهندسي في حالة البيانات غير المبوبة | 69 |
| **5**. **2**. الوسط الهندسي في حالة البيانات المبوبة | 70 |
| **6**. الوسط التوافقي | 71 |
| **6**. **1**. الوسط التوافقي في حالة البيانات غير المبوبة | 71 |
| **6**. **2**. الوسط التوافقي في حالة البيانات المبوبة | 72 |
| **II**. المقاييس غير المركزية | 73 |
| **1**. الربيعيات | 73 |
| **1.1**. الربيعيات في حالة البيانات غير المبوبة | 73 |
| **1**. **2**. الربيعيات في حالة البيانات المبوبة | 75 |
| **2**.العشريات | 77 |
| **2**. **1**. العشريات في حالة البيانات غير المبوبة | 77 |
| **2.2**. العشريات في حالة البيانات المبوبة | 79 |

| | |
|---|---|
| **3**. المؤينات | 81 |
| **3. 1**. المؤينات في حالة البيانات غير المبوبة | 81 |
| **3. 2**. المؤينات في حالة البيانات المبوبة | 83 |
| **الفصل الرابع** : مقاييس التشتت | 86 |
| **1**. المدى | 88 |
| **2**. الإنحراف الربيعي | 89 |
| **3**. الإنحراف المتوسط | 90 |
| **3. 1**. الإنحراف المتوسط في حالة البيانات غير المبوبة | 90 |
| **3. 2**. الإنحراف المتوسط في حالة البيانات المبوبة | 91 |
| **4**. التباين | 92 |
| **4. 1**. التباين في حالة البيانات غير المبوبة | 93 |
| **4. 2**. التباين في حالة البيانات المبوبة | 95 |
| **5**. الإنحراف المعياري | 98 |
| **5. 1**. الإنحراف المعياري في حالة البيانات غير المبوبة | 98 |
| **5. 2**. الإنحراف المعياري في حالة البيانات المبوبة | 99 |
| **6**. معامل الإختلاف | 99 |
| **7**. الدرجة القياسية | 101 |
| **الفصل الخامس**: مقاييس الشكل | 103 |

| | |
|---|---|
| **1**.مقاييس الإلتواء | 107 |
| **1.1**. معامل بيرسون | 108 |
| **2.1**. معامل فيشر | 108 |
| **3.1**. معامل يول | 109 |
| **2**.مقاييس التفلطح | 112 |
| **1.2**. معامل بيرسون | 113 |
| **2.2**. معامل فيشر | 113 |
| **3.2**. معامل كيلي | 114 |
| **الفصل السادس**: تحليل الارتباط و الانحدار | 118 |
| **1**. تحليل الارتباط | 119 |
| **1.1**.معامل الارتباط البسيط | 119 |
| **2.1**.معامل الارتباط المتعدد | 121 |
| **3.1**. معامل الارتباط الجزئي | 124 |
| **2**. تحليل الإنحدار | 126 |
| **1.2**. التعريف بالنموذج الخطي البسيط | 126 |
| **2.2**. تقدير معالم النموذج باستخدام طريقة المربعات الصغرى | 128 |
| **3.2**. القدرة التفسرية للنموذج | 130 |
| **قائمة المراجع** | 137 |

**مقدمة عامة :**

مادة الإحصاء الوصفي تعد من بين المواد المهمة التي يكتسب من خلالها الطالب أساسيات التحليل الإحصائي ، حيث تساعد المتلقي في اتخاذ القرار السليم و المبني على منهجية علمية ، كما تتضمن هذه المادة مفاهيم أساسية حول البحث العلمي مثل المجتمع الإحصائي، العينة الإحصائية، المتغير الإحصائي و غيرها من المفاهيم الضرورية الواجب اكتسابها من طرف الطلبة. و تسمح هذه المادة كذلك للطالب تعلم حساب مختلف المقاييس الإحصائية التي تساعده في الكشف عن خصائص البيانات محل الدراسة مثل المتوسط الحسابي، التباين، و شكل البيانات، كما تساعده على الوصول إلى نتائج موضوعية عن الظاهرة المراد تحليلها.

و بهدف تسهيل اكتساب أساسيات الإحصاء الوصفي، حاولنا انجاز هذه المحاضرات بالشكل الذي يساعد الطالب على الفهم و التحصيل ، سعينا إلى إعداد سلسلة من المحاضرات مع أمثلة مبسطة كما هو مقرر تدريسه لطلبة السنة أولى جذع مشترك علوم اقتصادية (LMD) .

حيث قمنا بتقسيم هذه المطبوعة إلى ست فصول ، يتضمن الفصل الأول مقدمة في علم الإحصاء ، و الفصل الثاني تبويب البيانات و العرض البياني، و يحتوى الفصل الثالث على مقاييس النزعة المركزية و غير المركزية ، والفصل الرابع على مقاييس التشتت، كما يتضمن الفصل الخامس على مقاييس الالتواء و الفصل السادس تحليل الارتباط و الانحدار.

# الفصل الأول : مقدمة في علم الإحصاء

I. **مفهوم علم الإحصاء** : عرف علم الإحصاء ب:

- "هو العلم الذي يختص بالطرق العلمية لجمع البيانات و تنظيمها و تلخيصها، و ذلك عن طريق التعبير عنها أو عرضها بصورة علمية و تحليلها بغرض الوصول إلى استنتاج النتائج و القوانين التي تحكمها، واتخاذ القرارات الملائمة لذلك"، (أبو عمه، عبدالله، و هندي، 1990، صفحة 3).
- " أنه علم الابتكار و تطوير و تطبيق أساليب تمكننا من تقويم عدم التأكد و العشوائية للاستدلال الإحصائي ". (بري، هندي، و راضي، 1998، صفحة 4)
- "علم الإحصاء هو علم يدرس مختلف طرق ووسائل جمع البيانات الكمية عن مختلف الظواهر الاقتصادية والاجتماعية، و ترتيبها وتبويبها و تحليلها و تفسيرها، و تقديمها بأشكال و صور ملائمة بهدف تسهيل اتخاذ القرار على أساس سليم". (رتول، 2006، صفحة 7)
- " علم الإحصاء هو العلم الذي يبحث في جمع البيانات و عرضها و تبويبها و تحليلها و استخدام النتائج في التنبؤ أو التقرير أو التحقق". (رشيد، 2008، صفحة 15)
- " يشار لعلم الإحصاء من أنه مجموعة الطرق العلمية القياسية التي يمكن توظيفها لجمع المعطيات (البيانات والمعلومات) الإحصائية عن الظواهر، و تبويبها و تلخيصها و تقييمها و الخروج من خلالها باستنتاجات حول مجموع وحدات المجتمع اعتمادا على جزء صغير من هذا المجتمع ، و هذا الجزء يدعى العينة ". (البلداوي، 2009، صفحة 17)
- " علم الإحصاء هو مجموعة النظريات و الطرق العلمية التي تبحث في جمع البيانات و عرضها و تحليلها و استخدام النتائج في التنبؤ أو التقرير و اتخاذ القرارات بناءا على ذلك ". (فليفل و حمدان، 2013، صفحة 13).

- "مجموعة الطرق والوسائل و القواعد والقوانين المبنية على التحليل المنطقي التي تستخدم كأفضل وسيلة لقياس و تحليل الظواهر والحقائق لاستخلاص النتائج ووضعها بصورة مناسبة لتوضيح العلاقة القائمة بينها"، (البدري و نجم، 2014، صفحة 17).
- " الإحصاء هو العلم الذي يتعامل مع جمع، تصنيف و تبويب الحقائق الرقمية كأساس لتفسير ووصف ومقارنة الظواهر"، (السقاف، 2020، صفحة 11).

و بناءا على التعريفات السابقة ، نستخلص أن علم الإحصاء يهتم بوسائل جمع البيانات ، و ترتيب هذه البيانات بشكل علمي وعرضها على شكل جداول أو رسومات بيانية مناسبة، بالإضافة إلى تحليل هذه البيانات بالطرق القياسية العلمية ، و استدلال نتائج المجتمع انطلاقا من نتائج العينة ، حيث تستخدم نتائج العينة في تقدير نتائج المجتمع ، و كل ذلك من أجل اتخاذ القرار.

و يمكن تلخيص وظائف الإحصاء فيما يلي (محسن، 2011، الصفحات 7-8) :

- ✓ يساعد الباحث على إعطاء أوصاف على جانب كبير من الدقة العلمية.
- ✓ يساعد على تلخيص النتائج في شكل مفهوم ملائم، فالبيانات التي يجمعها الباحث لا تعطي صورة واضحة إلا إذا تم تلخيصها في شكل معامل أو رقم أو شكل توضيحي كالرسوم البيانية.
- ✓ يساعد على استخلاص النتائج العامة من النتائج الجزئية، فمثل هذه النتائج لا يمكن استخلاصها إلا تبعا لقواعد إحصائية، كما يستطيع الباحث أن يحدد درجة احتمال صحة التعميم الذي يصل إليه.
- ✓ يمكن الباحث من التنبؤ بالنتائج التي يحتمل أن يحصل عليها في ظروف خاصة عن الواقع.
- ✓ يساعد على تمييز تأثير العوامل المختلفة التي تدخل في ظاهرة معينة عن بعضها البعض.
- ✓ يوضح العلاقات بين جوانب الظاهرة ومدى ارتباط بين المتغيرات المختلفة.

و ينقسم علم الإحصاء وفق الباحثين إلى قسمين:

- الإحصاء الوصفي
- الإحصاء الاستدلالي

**I. 1. الإحصاء الوصفي** : وعرف الإحصاء الوصفي كما يلي:

- "و يشمل الطرق الخاصة بتنظيم البيانات و تلخيصها و عرضها في صورة جداول إحصائية أو رسوم بيانية ، أو أشكال هندسية ، أو تلخيصها، أو حساب مقاييس النزعة المركزية، ومقاييس التشتت وغيرها من المقاييس الأخرى"، (أبو عمه، عبدالله، و هندي، 1990، صفحة 3).
- " هو طرق تنظيم المعلومات و تلخيصها، و الغرض من التنظيم هو المساعدة على فهم المعلومات، والطرق الوصفية تحتوي على التوزيعات تكرارية (جداول تكرارية) ، ورسوم بيانية، وطرق مقاييس النزعة المركزية ( المتوسطات)، و مقاييس التشتت و مختلف القياسات الأخرى."، (بري، هندي، و عبدالله، مبادئ الإحصاء و الإحتمالات، 1997، صفحة 3).
- " يختص الإحصاء الوصفي بتلخيص البيانات المستخدمة للبحث الإحصائي في جداول إحصائية أو رسوم بيانية أو حساب مقاييس إحصائية مثل النزعة المركزية و التشتت و الالتواء و التفلطح و قياس قوة الارتباط بين ظاهرة و أخرى". (بري، هندي، و راضي، 1998، صفحة 5) .
- " وهو يهدف إلى وصف مجموعة من البيانات عندما تتوفر، ويركز هذا النوع من الإحصاء على وصف الظاهرة و ربما تصنيفها وذلك من خلال استخدام الرسوم والأشكال البيانية والتوزيعات التكرارية أو من خلال استخدام مقاييس النزعة المركزية والتشتت، أو من خلال استخدام معاملات الارتباط لدراسة العلاقة بين المتغيرات"، (غرايبة و المنيزل، 2007، صفحة 13).

❖ " وهو ما يتعلق بجمع و تحليل المعطيات ووصفها لتكون بصيغة ذات مدلول من دون التعامل مع تعميم النتائج."، (البلداوي، 2009، صفحة 17)

❖ "يحتوي الإحصاء الوصفي الأساليب (الطرق) المستخدمة لتلخيص ووصف البيانات و ذلك بغرض تسهيل تفسيرها، و هذه الأساليب ممكن أن تكون بيانيا أو التحليل الحسابي ووضعها في جداول مناسبة"، (السقاف، 2020، صفحة 10).

I. 2. **الإحصاء الاستدلالي**: بينما يعرف الإحصاء الاستدلالي ب:

❖ " هو عبارة عن مجموعة الطرق العلمية التي تعمل للاستدلال على المجتمع بناء على البيانات الإحصائية التي جمعت من عينة من هذا المجتمع وفق طرق إحصائية محددة، و تشمل على عدد من المفاهيم والنظريات ، مثل نظرية التقدير، واختبار الفرضيات، وفحوص جودة الإنتاج"، (أبو عمه، عبدالله، و هندي، 1990، صفحة 3).

❖ " هو الوسائل العلمية التي تجري لسبر معالم المجتمع بناء على المعلومات التي تم الحصول عليها من العينة المأخوذة منه وفق الطرق الإحصائية المعلومة"، (بري، هندي، و عبدالله، 1997، صفحة 3)

❖ " يختص الإحصاء الاستدلالي باستخلاص نتائج عامة من بعض البيانات حيث يتم ذلك باستخدام المعاينة الإحصائية" (بري، هندي، و راضي، 1998، صفحة 5).

❖ " يركز هذا النوع من الإحصاء على الوصول إلى استنتاجات حول خصائص المجتمع من خلال استخدام المعلومات المتوفرة عن العينة المسحوبة من هذا المجتمع، ولهذا فإن الإحصاء الاستدلالي يركز على اختبار الفرضيات المتعلقة بالفروق بين المتوسطات أو النسب المئوية المتعلقة بعينة واحدة أو العينتين أو أكثر أو الفروق بين معاملات الارتباط"، (غرايبة و المنيزل، 2007، صفحة 13)

- " يختص بطرق تحليل و تفسير و استخلاص الاستنتاجات بالاعتماد على جزء من المجتمع (عينة) للتوصل إلى قرارات تخص مجموع المجتمع الإحصائي، وعليه فإن الإحصاء الاستدلالي يتعامل مع التعميم والتنبؤ و التقدير، وتتسم الاستنتاجات في بعض الحالات بعدم التأكد عندها يتم قياسها باستخدام الاحتمالات." (البلداوي، 2009، صفحة 17).

- " يتم من خلاله اتخاذ القرارات حول المجتمع الإحصائي و ذلك من واقع العينة المسحوبة من هذا المجتمع الإحصائي، وهذه القرارات يتم اتخاذها تحت شروط احتمالية، و تسمى وصف العينة و خصائصها بإحصائية العينة ، بينما الخصائص التي تصف المجتمع تسمى معالم"، (السقاف، 2020، صفحة 10).

## II. مفاهيم أساسية في علم الإحصاء:

**II. 1. الطريقة الإحصائية**: حسب (رشيد، 2008، صفحة 16) ، تعني الطريقة الإحصائية مجموعة من النظريات و الطرق العلمية التي تستخدم في جمع البيانات و تبويبها و عرضها و استخلاص النتائج و تفسيرها ، لذلك فإن الطريقة الإحصائية تتكون من عناصر تعتبر وسائل وأدوات هامة في البحث العلمي و العناصر هي :

- **جمع البيانات**: وهي عملية الحصول على القياسات أو التعدادات أو قيم المشاهدات للتجارب التي يجريها الباحث و كلما كانت دقيقة كلما كانت النتائج أدق.
- **تبويب البيانات وعرضها**: يتم تبويب البيانات طبقا لأسلوب تصنيف محدد مسبقا و عرضها بطرق مناسبة كالجداول ، الأشكال البيانية و الهندسية .
- **وصف البيانات**: عن طريق إبراز الخصائص الأساسية لها والتي يمكن التعبير عنها بمقاييس معينة ، مثل مقاييس النزعة المركزية ، مقاييس التشتت ، مقاييس الشكل.

- **تحليل النتائج**: وهو إظهار الخصائص الأساسية على شكل أرقام والتي يهم الباحث الحصول عليها للوصول إلى نتائج محددة.
- **استقراء النتائج و اتخاذ القرارات**: وهي مجموعة الاستنتاجات التي يتوصل إليها الباحث من خلال تحليل النتائج وهي غالبا ما تكون على شكل تقديرات أو تنبؤات أو تعميمات أو قرارات رفض أو قبول للفرضيات الإحصائية.

II. 2. **المجتمع الإحصائي**: يمكن تعريف المجتمع الإحصائي كما يلي:

- " المجتمع الإحصائي هو مجموعة كل العناصر قيد الدراسة و يسمى أحيانا مجتمع الهدف" (فليفل و حمدان، 2013، صفحة 14)
- "يمثل المجتمع جميع المفردات محل الدراسة والتي لها خصائص مشتركة، فقد يكون المجتمع الإحصائي محدود أو قد يكون غير محدود" (بري، هندي، و راضي، 1998، صفحة 6)
- "هو عبارة عن مجموعة من الوحدات أو المفردات ذات الصفات المشتركة " (البدري و نجم، 2014، صفحة 78)
- "كلمة مجتمع تستخدم من قبل الإحصائيين للإشارة إلى كل العناصر التي تم اختيارها للدراسة، و خصائص المجتمع تسمى معلمة " (السقاف، 2020، صفحة 11).
- "مجموع الوحدات الإحصائية التي تقع عليها الدراسة الإحصائية" (رتول، 2006، صفحة 12)
- "المجتمع عبارة عن جميع القيم أو المفردات التي يمكن أن يأخذها المتغير" (الراوي، 1984، صفحة 13)

**II. 3. العينة الإحصائية** : في الكثير من الأحيان يصعب علينا الوصول إلى جميع أفراد المجتمع ، حيث يكلفنا ذلك مال وجهد ووقت ، لذلك نلجأ إلى اختيار جزء من المجتمع محل الدراسة ، و يسمى هذا الجزء بالعينة ، و يمكن تعريف العينة كما يلي:

- " جزء من المجتمع تختار بحيث تمثل جميع صفات المجتمع ، و ينفرد بها فرع خاص من علم الإحصاء يسمى نظرية العينات" (بري، هندي، و عبدالله، 1997، صفحة 04).
- "مجموعة جزئية من المجتمع" (النجار، 2015، صفحة 16).
- " هي مجموعة صغيرة من الأفراد تحمل جميع الخصائص المهمة المميزة للمجتمع الكلي بحيث تكون صورة مصغرة منه و تكون ممثلة لذلك المجتمع" (محسن، 2011، صفحة 09).
- " هي جزء من المجتمع، بحيث تكون ممثلة له تمثيلا جيدا، وخصائص العينة تسمى إحصائية " (السقاف، 2020، صفحة 11).
- " العينة عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع " (الراوي، 1984، صفحة 14).

**II. 4. البيانات الإحصائية**: و تعرف كما يلي :

- " هي قيم المشاهدات للتجارب التي يجريها الباحث و كلما كانت دقيقة كلما كانت القرارات المتخذة بصددها أكثر صدقا " (فليفل و حمدان، 2013، صفحة 13).
- " هي مجموعة من الحقائق و المشاهدات التي يتم جمعها من مجتمع إحصائي معين، و من أمثلة البيانات: الاسم و السن و المهنة و مستوى التعليم ، متوسط الدخل و الحالة الزوجية ..." (السقاف، 2020، صفحة 11) .

- وتنقسم البيانات الإحصائية إلى قسمين رئيسيين هما :

- **البيانات الوصفية ( النوعية)**: وهي ذلك النوع من البيانات الذي يعبر عن ظواهر لا يمكن قياسها رقميا و إنما يتم تصنيف المتغير فيها إلى مستويات ، و تنقسم إلى قسمين ، بيانات نوعية اسمية و بيانات نوعية ترتيبية . فأما البيانات النوعية الاسمية، فهي تلك البيانات غير قابلة للترتيب مثل : فصيلة الدم، جنس المواليد. و أما البيانات النوعية الترتيبية ، فهي تلك البيانات القابلة للترتيب ، مثل : تقديرات الطلبة ( ممتاز، جيد جدا، جيد ، متوسط) (صالحين، صفحة 15).

- **البيانات الكمية (العددية)** : وهي بيانات مقاسة بمقياس كمي ، و تنقسم البيانات الإحصائية العددية إلى نوعين، أحدهما منفصل و تكون قياساته ناتجة عن عملية عد أو تعداد، مثل عدد حوادث المرور اليومي خلال فترة زمنية محددة، أو العدد السنوي لحالات الولادة ، أو الزواج أو الوفاة، أو الطلاق في بلد معين، و تكون مثل هذه القياسات دائما أعداد صحيحة. و النوع الأخر هو النوع المتصل ( أو المستمر) ، و تكون قياساته ناتجة عن استخدام جهاز أو أداة للقياس ، مثل بيانات تتضمن قياسات الطول، أو وزن، أو درجة حرارة، أو مستوى التحصيل الدراسي (كنجو، 2000، صفحة 08).

- و يمكن جمع هذه البيانات بإتباع طريقتين :

  - **طريقة الحصر الشامل**: وهو الأسلوب الذي يدرس فيه الباحث جميع عناصر أو وحدات المجتمع المراد بحثه دون استثناء. و الشائع أن يتبع هذا الأسلوب عندما يراد الحصول على معلومات تفصيلية عن جميع مفردات المجتمع كما هو الحال في الأبحاث القومية الكبيرة التي تجرى لأغراض التخطيط القومي الشامل، سواء كان متوسط المدى أو بعيد المدى ، مثال على ذلك: التعداد السكاني القومي، تعداد المنشات الزراعية ،

والصناعية، و الاجتماعية . ولا يخفى أن مثل هذا النوع من الأبحاث يحتاج إلى فترة زمنية طويلة نسبيا، و إلى كادر إداري وأخر فني مدربين تدريبا عاليا على عمليات الإدارة و الحصر في الميدان، و إلى وسائل اتصال ومواصلات ذات كفاءة عالية علاوة على كبر حجمها، كما يحتاج إلى أموال طائلة (الصغير، 2011، صفحة 21) .

- **طريقة الحصر بالعينة ( المعاينة)** : العينة ببساطة هي جزء من مجتمع الدراسة ، ولكي يكون هذا الجزء ممثلا تمثيلا صادقا لمجتمع الدراسة ، و بالتالي يكون ممكنا تعميم النتائج المتحصل عليها من دراسة هذا الجزء فقط على جميع مفردات مجتمع الدراسة. حيث يجب أن تتصف هذه العينة بصفة ما نطلق عليها العينة الاحتمالية ، و سميت بهذا الاسم لأنها تعتمد على نظرية الاحتمالات ، و أهم ما يميز العينة الاحتمالية هو أنها عينة معروف احتمالات اختيار مفرداتها مسبقا، كما أن لكل فرد من أفراد المجتمع الذي سحبت منه العينة احتمال للاختيار ضمن أفرادها ، وهذا الاحتمال يمكن حسابه ولا يكون معدوما، و من أهم أنواع العينات الاحتمالية نجد : العينة العشوائية البسيطة ، العينة العشوائية المنتظمة، العينة العشوائية الطبقية، العينة العشوائية العنقودية. (الصغير، 2011، الصفحات 21-22).

- كما يمكن جمع البيانات بإتباع الأساليب (المصادر) التالية (فليفل و حمدان، 2013، صفحة 14):

  - الأسلوب المباشر : عن طريق الميدان مباشرة.
  - الأسلوب غير المباشر: عن طريق السجلات أو الوثائق التاريخية.

- أسلوب الاستبيان: وهي رزمة من الأوراق التي تحتوي على مجموعة من الأسئلة والاستفسارات تعبىء من قبل الشخص الخاضع للبحث .
- أسلوب المقابلات الشخصية : وهي السؤال المباشر من قبل الباحث.
- أسلوب الاختبارات الخاصة : أسلوب خاص يستخدم في حالات محددة فقط مثل اختبارات الذكاء.

- و صنف (طبيه، 2008، صفحة 13) مصادر جمع البيانات كما يلي :

- **المصدر الأول**: المصدر المباشر: النزول للميدان وجمع المعلومات مباشرة.
- **المصدر الثاني**: المصدر الغير مباشر: و يندرج تحت هذا المصدر كل ما يلي:
  - ✓ السجلات أو الوثائق التاريخية.
  - ✓ الإستبيان.
  - ✓ المقابلات الشخصية.
  - ✓ الإختبارات الخاصة.

## II. 5. المتغير الإحصائي:

❖ هو مقدار له خصائص رقمية (كمية) وغير رقمية (وصفية) تتغير قيمته من عنصر إلى أخر من عناصر المجتمع أو العينة. فمثلا إذا رغبنا في دراسة ظاهرة مثل الوزن أو الطول أو الذكاء أو الجنس أو لون البشرة أو لون الشعر أو لون العيون، فإن قراءة المفردات لمتغير الطول أو الوزن أو الذكاء تكون بيانات كمية أو رقمية ، و ظاهرة الجنس أو لون البشرة أو لون الشعر أو لون العيون تأخذ قيما وصفية أو غير رقمية (بري، هندي، و عبدالله، 1997، صفحة 05).

- ❖ يمثل المتغير قيما لظاهرة ما حيث تتغير قيم مفردات الظاهرة من مفردة إلى أخرى مثل ظاهرة الطول، الوزن و العمر، و قد يكون المتغير كمي ( متصل أو منفصل) مثل الطول و الوزن والعمر، أو قد يكون وصفي ( نوعي: اسمي أو ترتيبي) مثل لون العيون، لون البشرة، تقديرات الطلبة (بري، هندي، و راضي، 1998، صفحة 07).
- ❖ إن المعلومات التي تجمع من الأفراد و المتعلقة بسمة أو خاصية معينة تسمى بالبيانات وهي تمثل خصائص مجموعة من الأفراد قد تأخذ قيما مختلفة بالنسبة للأفراد المختلفين قيد الدراسة. مثل هذه الخصائص تسمى بالمتغير ، فالمتغير هي سمة أو خاصية تأخذ قيما متغيرة عند الأفراد المختلفين. فمثلا مجموعة من طلبة الجامعة قد يختلفون في الجنس أو السنة الدراسية، أو الذكاء أو التحصيل ، مثل هذه الخصائص تسمى متغيرات. (غرايبة و المنيزل، 2007، صفحة 13).

– و تصنف المتغيرات كما يلي (غرايبة و المنيزل، 2007، صفحة 14):

- **المتغير المتصل** ( **المستمر**) : هي عبارة عن المتغيرات التي تأخذ أي قيمة على مقياس السمة ( الخاصية)، مثل الطول و الوزن ، وفي مثل هذا النوع من المتغيرات توجد قيم لا حصر لها بين أي قيمتين رقميتين .
- **المتغير المنفصل** ( **المتقطع**) : هي عبارة عن المتغيرات التي تأخذ قيما محددة بحيث لا توجد قيما كسرية أو عشرية ، مثال على ذلك ، عدد الطلاب في صف ما، أو عدد أفراد الأسرة أو عدد السيارات في كراج ما فلا يمكن أن تكون هذه الأعداد كسرية.

– كما يمكن تصنيف المتغيرات من حيث تأثير بعضهما في بعضها الأخر أو تأثر بعضها ببعضها الأخر إلى مجموعتين اثنتين ، إحداهما نطلق عليها متغيرات مستقلة ، و الأخرى نسميها متغيرات تابعة، ولا يتأتى هذا التصنيف إلا عند الزعم بوجود علاقة ما بين متغيرين على الأقل،

فمثلا نقول بأن هناك علاقة ما بين التفكك الأسري وانتشار الجريمة في المجتمع(معدل ارتكاب الجريمة )، ففي هذا المثال نجد أن التفكك الأسري وهو المتغير المستقل يؤثر في معدل ارتكاب الجريمة وهو المتغير التابع (الصغير، 2011، الصفحات 04-05) . و يمكن تعريف المتغير المستقل و المتغير التابع كما يلي (النجار، 2015، صفحة 20) :

- **المتغير المستقل** : وهو المتغير الذي يخضع لسيطرة الإحصائي أو الباحث.
- **المتغير التابع:** وهو المتغير الذي نتنبأ بقيمته من خلال معرفتنا لقيم المتغير المستقل.

### II. 6. المعلمة :

- شيء يميز المجتمع كله و ذلك مثل متوسط الدخل الشهري للأسر في دولة معينة ، أو متوسط الطول للذكر البالغ في دولة معينة ، أو نسبة الذين يدخنون بصفة دائمة في مجتمع معين، أو نسبة المعيب في الإنتاج لإحدى السلع و هكذا... (بري، هندي، و عبدالله، 1997، صفحة 05) .
- تمثل المعلمة مؤشر يصف المجتمع ، مثل المتوسط الحسابي للمجتمع، تباين المجتمع ،انحراف المعياري للمجتمع ، نسبة النجاح في المجتمع. (بري، هندي، و راضي، 1998، صفحة 06)

### II. 7.الإحصاءة ( الإحصائية) :

- شيء يميز العينة ، مثل متوسط الدخل الشهري لعينة مكونة من 100 أسرة في دولة ما أو متوسط الذكر البالغ لعينة مكونة من 50 ذكر و هكذا... (بري، هندي، و عبدالله، 1997، صفحة 05) .
- تمثل الإحصاءة مؤشرا يحسب من العينة و يستخدم تقديرا لمعلمة المجتمع محل الدراسة ، مثل المتوسط الحسابي للعينة ، تباين و الانحراف المعياري للعينة . (بري، هندي، و راضي، 1998، صفحة 06).

- و فيما يلي جدول يوضح بعض المعلمات و الإحصائيات:

| المعالم | | الإحصاءات | |
|---|---|---|---|
| $\mu$ | متوسط المجتمع | $\bar{X}$ | متوسط العينة |
| $\sigma_X$ | الإنحراف المعياري للمجتمع | $s_x$ | الإنحراف المعياري للعينة |
| $\sigma_X^2$ | تباين المجتمع | $s_x^2$ | تباين العينة |
| $N$ | حجم المجتمع | $n$ | حجم العينة |
| $\pi$ | نسبة المجتمع | $p$ | نسبة العينة |
| $X_i$ | مشاهدات المجتمع | $x_i$ | مشاهدات العينة |

## III. الرموز الإحصائية :

**III. 1. المتغير الإحصائي**: سنرمز للمتغير الإحصائي بالرمز $x$ ، و لكل قيمة له بالرمز $x_i$ ، فإذا كان أعمار خمس طلاب كالأتي: 16، 22، 24، 18، 20 ، نكتب :

$$x_i = 16, 22, 24, 18 ,20$$

حيث :

$$x_1 = 16\,, x_2 = 22, x_3 = 24, x_4 = 18, x_5 = 20$$

يعني : المشاهدة الأولى: 16 ، المشاهدة الثانية :22، المشاهدة الثالثة:24 ، المشاهدة الرابعة : 18، المشاهدة الخامسة : 20. وهنا عدد المشاهدات هو 5 ، أي $n = 5$، بينما مجموع المشاهدات ، تكتب بالشكل التالي:

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

من مثالنا ، نجد :

$$\sum_{i=1}^{n=5} x_i = x_1 + x_2 + x_3 + x_4 + x_5$$

$$\sum_{i=1}^{n=5} x_i = 16 + 22 + 24 + 18 + 20 = 100$$

و يرمز الرمز $\Sigma$ إلى المجموع و يطلق عليه سيغما ( Sigma) ، و يمكن حساب مجموع جزئي مثل $\sum_{i=1}^{3} x_i$، ويحسب بالشكل التالي:

$$\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3 = 16 + 22 + 24 = 62$$

**III. 2. مجموع مربع المشاهدات**: ويرمز لمربع المشاهدات ب $x_i^2$ ، أما مجموع مربع المشاهدات فيكتب بالشكل التالي:

$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

من مثالنا ، نجد :

$$2040 \sum_{i=1}^{n} x_i^2 = 16^2 + 22^2 + 24^2 + 18^2 + 20^2 =$$

ويرمز لمربع مجموع المشاهدات ب $(\sum_{i=1}^{n} x_i)^2$ و من مثالنا ، فإنه يساوي $100^2$، أي 10000، حيث :

$$(\sum_{i=1}^{n} x_i)^2 = (x_1 + x_2 + \cdots + x_n)^2$$

مع :

$$\sum_{i=1}^{n} x_i^2 \neq (\sum_{i=1}^{n} x_i)^2$$

**III. 3. حاصل ضرب قيم متغيرين $x$ و $y$**: ويرمز له بالرمز $\sum_{i=1}^{n} x_i y_i$ ، و يكتب بالشكل التالي:

$$\sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

ويرمز لحاصل ضرب مجموعين لقيم متغيرين بالرمز $\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)$، حيث :

$$\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) = (x_1 + x_2 + \cdots + x_n)(y_1 + y_2 + \cdots + y_n)$$

مع:

$$\sum_{i=1}^{n} x_i y_i \neq \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)$$

**III. 4. حاصل قسمة قيم متغيرين $x$ و $y$**: و يكتب بالشكل التالي :

$$\sum_{i=1}^{n} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \cdots + \frac{x_n}{y_n}$$

ويرمز لحاصل قسمة مجموع قيم متغيرين بالشكل التالي :

$$\frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} = \frac{x_1 + x_2 + \cdots + x_n}{y_1 + y_2 + \cdots + y_n}$$

مع :

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

**مثال** : نفرض أن قيم المتغير $x$ هي كالأتي: $x_i = 4,2,3,7$ ، و قيم المتغير $y$ هي كالأتي: $y_i =$ 3,9,6,2 ، أوجد قيم كل مما يأتي:

$$1.\sum_{i=1}^{n} x_i \quad 2.\sum_{i=1}^{n} y_i \quad 3.\sum_{i=1}^{n} x_i y_i \quad 4.\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)$$

الحل :

$$1.\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + x_4 = 4 + 2 + 3 + 7 = 16$$

$$2.\sum_{i=1}^{n} y_i = y_1 + y_2 + y_3 + y_4 = 3 + 9 + 6 + 2 = 20$$

$$3.\sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 = (4.3) + (2.9) + (3.6) + (7.2)$$

$$= 62$$

$$4.\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) = 16.20 = 320$$

### III. 5. بعض القواعد تخص عملية الجمع و الطرح :

1. إذا كان (c) عدد ثابت فإن:

$$\sum_{i=1}^{n} c = c_1 + c_2 + \cdots + c_n = nc$$

2. إذا كان (c) عدد ثابت فإن :

$$\sum_{i=1}^{n} cx_i = cx_1 + cx_2 + \cdots + cx_n$$

$$= c(x_1 + x_2 + \cdots + x_n) = c\sum_{i=1}^{n} x_i$$

3. إذا كان (c) عدد ثابت فإن :

$$\sum_{i=1}^{n}(x_i - c) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} c = \sum_{i=1}^{n} x_i - nc$$

مع :

$$\sum_{i=1}^{n}(x_i - c) \neq \sum_{i=1}^{n} x_i - c$$

4. جمع قيم متغيرين أو أكثر هو مجموع جمعهم ، أي:

$$\sum_{i=1}^{n}(x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

# الفصل الثاني: تبويب البيانات والعرض البياني

**تمهيد :**

بعد جمع البيانات من مختلف المصادر ، فإننا نحصل على بيانات أولية غير منتظمة عدديا و يصعب علينا دراستها أو استنتاج أي شيء منها ، و لذلك دعت الحاجة إلى تنظيم و تلخيص هذه البيانات بصورة يسهل فهمها واستنتاج بعض النتائج الأولية، و يتم تلخيص و تنظيم هذه البيانات في جدول يسمى بجدول التوزيع التكراري ، قد تكون هذه البيانات وصفية (نوعية) مثل تقديرات الطلاب : ممتاز ، جيد جدا، جيد ، ضعيف ، تخصص الطلبة : علوم اقتصادية، علم الاجتماع ، علوم سياسية ، الأدب ، أو بيانات كمية مثل أوزان الطلبة ، أطوال الطلبة ، درجات الطلبة ، عدد الحوادث الأسبوعية ، درجة الحرارة اليومية، و تنظم البيانات الإحصائية سواء كانت وصفية أم كمية بجدول التوزيع التكراري وهو عبارة عن جدول يلخص البيانات الأولية فيوزعها على فئات ، ويحدد عدد الأفراد الذين ينتمون إلى كل فئة و يسمى هذا العدد تكرار الفئة ، ويرمز له بالرمز $n_i$ ، و يطلق على عملية تلخيص البيانات في جدول تكراري بتبويب البيانات.

I. **تبويب البيانات** : تهدف عملية تبويب البيانات إلى وضع تلك البيانات في جدول مبسط بحيث يمكن للباحث دراستها و تحليلها و استخلاص النتائج بكل سهولة ، وقبل وضع البيانات في شكلها النهائي في جدول التوزيع التكراري، نقوم بعملية التعداد من خلال وضع تشطيبة عمودية صغيرة أمام المتغير الإحصائي و عندما نصل للتشطيبة الخامسة نضعها مقاطعة للأربعة الأولى ، و تسهل علينا هذه الخطوة حساب التكرارات المقابلة للمتغير الإحصائي .

I. 1. **التوزيع التكراري البسيط ( المطلق)** : و يقصد به توزيع قيم المتغير الإحصائي على فئات ، و فيما يلي كيفية تنظيم و تلخيص البيانات (تبويب البيانات) في جدول تكراري بسيط باختلاف شكلها وصفية كانت أو كمية :

1. **البيانات الوصفية** : أخذت عينة عشوائية من الطلبة تتكون من 25 طالبا ، ليتم استقصاءهم على شعب البكالوريا التي يحملونها ، من خلال ملء استمارات ، فكانت الإجابات كما يلي :

| | | | | |
|---|---|---|---|---|
| أداب | رياضيات | تسيير و اقتصاد | علوم | أداب |
| أداب | تسيير و اقتصاد | رياضيات | علوم | رياضيات |
| علوم | أداب | رياضيات | تسيير و اقتصاد | علوم |
| علوم | علوم | أداب | تسيير و اقتصاد | أداب |
| تسيير و اقتصاد | علوم | علوم | رياضيات | تسيير و اقتصاد |

يطلق على هذه البيانات بالبيانات غير مبوبة ، أي لا يتم عرض تكرارات المتغير الإحصائي فيها ، و لتبويب هذه البيانات نقوم أولا بخطوة التعداد ، كما يوضح الجدول الموالي :

| | |
|---|---|
| علوم | 𝍸 III |
| رياضيات | 𝍸 |
| تسيير و اقتصاد | 𝍸 I |
| أداب | 𝍸 I |

ثم نحصل على جدول التوزيع التكراري الذي يتكون من عمودين الأول يشير إلى المتغير الاحصائي الذي يمثل شعب البكالوريا ، و العمود الثاني يمثل التكرارات وكما هو موضح بالجدول التالي :

| المتغير الإحصائي $x_i$ | التكرار $n_i$ |
|---|---|
| علوم | 08 |
| رياضيات | 05 |
| تسيير و اقتصاد | 06 |

| | |
|---|---|
| أداب | 06 |
| المجموع | 25 |

## 2. البيانات الكمية :

### ❖ بالنسبة للمتغير الكمي المنفصل :

- **طول الفئة معدوم** :البيانات التالية تمثل إحصاءات حول عدد الحوادث المسجلة خلال شهر ( 30 يوما) ، و المطلوب وضع هذه البيانات في جدول تكراري مناسب :

| | | | | |
|---|---|---|---|---|
| 04 | 02 | 02 | 01 | 01 |
| 01 | 02 | 01 | 01 | 00 |
| 03 | 05 | 03 | 00 | 03 |
| 05 | 04 | 05 | 03 | 04 |
| 00 | 01 | 05 | 03 | 04 |
| 00 | 02 | 00 | 02 | 02 |

قبل عملية التبويب يتم أولا عملية التعداد كما يبين الجدول التالي :

| | |
|---|---|
| 00 | ~~\|\|\|\|~~ |
| 01 | ~~\|\|\|\|~~ \| |
| 02 | ~~\|\|\|\|~~ \| |
| 03 | ~~\|\|\|\|~~ |
| 04 | \|\|\|\| |
| 05 | \|\|\|\| |

و بذلك نحصل على جدول التوزيع التكراري الذي يتكون من عمودين ، الأول يمثل بيانات المتغير الإحصائي والمتمثل قي عدد الحوادث اليومية ، و العمود الثاني يمثل تكرار قيم المتغير الإحصائي ، كما هو موضح بالجدول التالي:

| عدد الحوادث $x_i$ | التكرار $n_i$ |
|---|---|
| 00 | 05 |
| 01 | 06 |
| 02 | 06 |
| 03 | 05 |
| 04 | 04 |
| 05 | 04 |
| المجموع | 30 |

- **طول الفئة غير معدوم** : البيانات التالية تمثل أعمار مجموعة من المستهلكين قاموا بشراء منتوج معين من إحدى المؤسسات، والمطلوب ترتيب هذه البيانات في جدول تكراري مناسب .

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 26 | 42 | 30 | 32 | 19 | 30 | 22 | 22 | 35 | 25 |
| 28 | 40 | 30 | 45 | 21 | 32 | 25 | 18 | 40 | 30 |
| 33 | 35 | 31 | 32 | 25 | 35 | 28 | 20 | 36 | 26 |
| 32 | 37 | 29 | 25 | 25 | 37 | 30 | 30 | 35 | 25 |
| 29 | 43 | 28 | 26 | 24 | 26 | 32 | 45 | 20 | 30 |

في مثل هذه الحالات يتم حساب طول الفئة ، و ترتيب هذه البيانات في جدول تكراري ذو فئات متساوية طولها L . و يتم حساب طول الفئة بتطبيق الصيغة التالية :

$$L = \frac{R}{C}$$

حيث $R$ تمثل المدى و هو الفرق ما بين أكبر قيمة بالبيانات و أدنى قيمة بها ، أي:

$$R = V_{max} - V_{min}$$

أما $C$ فتشير إلى عدد الفئات ، و يتم حسابها باستخدام واحدة من بين الصيغ الرياضية التالية:

- **معادلة ستيرجس (Staurges)** : و هي من الشكل التالي:

$$C = 1 + 3.322\log(n)$$

- **معادلة يول ( Yule)** : وهي من الشكل التالي:

$$C = 2.5\sqrt[4]{n}$$

و بذلك نجد طول الفئة بالشكل التالي :

$$C = 1 + 3.322\log(n) = 1 + 3.322\log(50) = 6.6439 \approx 7$$

أي أن عدد الفئات هو 7 فئات ، و بذلك نجد طول الفئة بالشكل التالي:

$$R = V_{max} - V_{min} = 45 - 18 = 27 \Rightarrow L = \frac{27}{7} = 3.85 \approx 4$$

قبل تبويب هذه البيانات ، سنقوم بعملية التعداد كأول مرحلة ، مثلما يوضح الجدول التالي:

<table>
<tr><td>|</td><td>43</td><td>-</td><td>38</td><td>|</td><td>33</td><td>|||</td><td>28</td><td>-</td><td>23</td><td>|</td><td>18</td></tr>
<tr><td>-</td><td>44</td><td>-</td><td>39</td><td>-</td><td>34</td><td>||</td><td>29</td><td>|</td><td>24</td><td>|</td><td>19</td></tr>
<tr><td>||</td><td>45</td><td>||</td><td>40</td><td>||||</td><td>35</td><td>|| 𝍸</td><td>30</td><td>| 𝍸</td><td>25</td><td>||</td><td>20</td></tr>
<tr><td></td><td></td><td>-</td><td>41</td><td>|</td><td>36</td><td>|</td><td>31</td><td>||||</td><td>26</td><td>|</td><td>21</td></tr>
</table>

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 22 | II | 27 | - | 32 | 𝍸 | 37 | II | 42 | I | | |

و بذلك يكون جدول التوزيع التكراري بالشكل التالي:

| **الفئات** | **التكرار المطلق** $n_i$ |
|---|---|
| [21-18 ] | 05 |
| [ 25-22 ] | 09 |
| [ 29-26 ] | 09 |
| [33-30 ] | 14 |
| [37-34 ] | 07 |
| [41-38 ] | 02 |
| [45-42 ] | 04 |
| المجموع | 50 |

❖ **بالنسبة للمتغير الكمي المتصل** : فيما يلي بيانات تمثل أجور العمال في إحدى المؤسسات الاقتصادية و الوحدة بـ $10^3$ ، و المطلوب تبويبها:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 41 | 40 | 44 | 39 | 38 | 35 | 36 | 36 | 37 | 35 |
| 42 | 41 | 43 | 40 | 44 | 41 | 40 | 43 | 44 | 42 |
| 42 | 47 | 48 | 49 | 45 | 41 | 43 | 41 | 42 | 40 |
| 48 | 52 | 55 | 46 | 50 | 45 | 48 | 41 | 45 | 48 |
| 47 | 41 | 58 | 54 | 59 | 50 | 55 | 51 | 49 | 43 |

تصنيف هذه البيانات في جدول تكراري يتطلب منا حساب طول الفئة على النحو التالي:

$$C = 1 + 3.322\log(n) = 1 + 3.322\log(50) = 6.6439 \approx 7$$

أي أن عدد الفئات هو 7 فئات ، و بذلك نجد طول الفئة بالشكل التالي:

$$R = V_{max} - V_{min} = 59 - 35 = 24$$
$$\Rightarrow L = \frac{24}{7} = 3.42 \approx 4$$

قبل تبويب هذه البيانات ، سنقوم بعملية التعداد كأول مرحلة ، مثلما يوضح الجدول التالي:

<table>
<tr><td>||</td><td>55</td><td>||</td><td>50</td><td>|||</td><td>45</td><td>||||</td><td>40</td><td>||</td><td>35</td></tr>
<tr><td>-</td><td>56</td><td>|</td><td>51</td><td>|</td><td>46</td><td>|||||||</td><td>41</td><td>||</td><td>36</td></tr>
<tr><td>-</td><td>57</td><td>|</td><td>52</td><td>||</td><td>47</td><td>||||</td><td>42</td><td>|</td><td>37</td></tr>
<tr><td>|</td><td>58</td><td>-</td><td>53</td><td>||||</td><td>48</td><td>||||</td><td>43</td><td>|</td><td>38</td></tr>
<tr><td>|</td><td>59</td><td>|</td><td>54</td><td>||</td><td>49</td><td>|||</td><td>44</td><td>|</td><td>39</td></tr>
</table>

و الجدول الموالي يبين الفئات و التكرارات المقابلة لكل فئة :

| التكرارات | الفئات |
|---|---|
| 06 | ]39-35] |
| 16 | ]43-39] |
| 11 | ]47-43] |
| 10 | ]51-47] |
| 03 | ]55-51] |
| 03 | ]59-55] |
| 01 | ]63-59] |
| 50 | المجموع |

**I.** **2. التوزيع التكراري المتجمع** : إن التوزيعات التكرارية تظهر لنا فقط عدد مرات تكرار الفئة ، ولمعرفة عدد التكرارات التي تقل، أو تزيد عن حد معين من حدود الفئات ، فإنه يتم اللجوء إلى جداول التوزيعات التكرارية المتجمعة ، وهي التوزيع التكراري المتجمع النازل، و التوزيع التكراري المتجمع الصاعد.

❖ **التوزيع التكراري المتجمع الصاعد:** يستخدم لغرض المعرفة السريعة لعدد التكرارات التي تقل عن حد معين من حدود الفئات، وفي حساب بعض مقاييس النزعة المركزية. في هذا التوزيع يكون عدد التكرارات التي تقل عن الحد الأعلى لأول فئة يساوي عدد تكرارات الفئة الأولى، و عدد التكرارات التي تقل عن الحد الأعلى للفئة الثانية تساوي إلى عدد تكرارات الفئة الأولى و الثانية، وهكذا يستمر التجميع حتى الوصول إلى التكرارات التي تقل عن الحد الأعلى للفئة الأخيرة ، حيث يساوي إلى مجموع التكرارات ، و يرمز له بالرمز $n_i^{\nearrow}$.

- **مثال(01)**: البيانات التالية توزيع مؤسسة ما حسب أجورهم الشهرية ( الوحدة : $10^3$ **دج)**

| الأجر | [ 20-10 [ | ]30-20 [ | ] 40-30 [ | ] 50-40 [ | ] 60-50 [ | ] 70-60 [ |
|---|---|---|---|---|---|---|
| عدد العمال | 18 | 30 | 25 | 17 | 12 | 8 |

**المطلوب**:

- أوجد التوزيع التكراري المتجمع الصاعد
- ما هو عدد العمال الذين تقل أجورهم عن 40000 دج
- ما هو عدد العمال الذين تقل أجورهم عن 60000 دج

**الحل**: نتائج عملية التجميع موضحة في الجدول الموالي:

| الأجر | عدد العمال $n_i$ | الحدود العليا | التكرار المتجمع الصاعد $n_i^{\nearrow}$ |
|---|---|---|---|
| [ 10-20 [ | 18 | أقل من 20 | 18 |
| [ 20-30 [ | 30 | أقل من 30 | 48 |
| [ 30-40 [ | 25 | أقل من 40 | 73 |
| [ 40-50 [ | 17 | أقل من 50 | 90 |
| [ 50-60 [ | 12 | أقل من 60 | 102 |
| [ 60-70 [ | 10 | أقل من 70 | 112 |
| المجموع | 112 | | |

عدد العمال الذين تقل أجورهم عن 40000 دج هو 73 عامل (18+30+25)

عدد العمال الذين تقل أجورهم عن 60000 دج هو 102 (18+30+25+17+12)

❖ **التوزيع التكراري المتجمع النازل**: يستخدم لغرض المعرفة السريعة لعدد التكرارات التي تساوي أو تزيد عن حد معين من حدود الفئات، في هذا الشكل من التوزيع ، يكون عدد التكرارات التي تساوي أو تزيد عن الحد الأدنى للفئة الأولى يساوي مجموع التكرارات. أما التكرارات التي تساوي أو تزيد عن الحد الأدنى للفئة الثانية يساوي مجموع التكرارات منقوص منه تكرار الفئة الأولى ، وهكذا حتى نصل إلى التكرارات التي تساوي أو تزيد عن الحد الأدنى للفئة الأخيرة حيث يساوي تكرار الفئة الأخيرة، و يرمز له بالرمز $n_i^{\searrow}$ .

**مثال (02)**: أوجد التكرار المتجمع النازل للمثال (01) ، و ما هو عدد العمال الذين تزيد أجورهم عن 40000 دج والذين تزيد أجورهم عن 60000 دج

**الحل:**

| التكرار المتجمع النازل $n_i^{\searrow}$ | الحدود الدنيا | عدد العمال $n_i$ | الأجر |
|---|---|---|---|
| 112 | 10 فأكثر | 18 | ] 20-10 ] |
| 94 | 20 فأكثر | 30 | ] 30-20 ] |
| 64 | 30 فأكثر | 25 | ] 40-30 ] |
| 39 | 40 فأكثر | 17 | ] 50-40 ] |
| 22 | 50 فأكثر | 12 | ] 60-50 ] |
| 10 | 60 فأكثر | 10 | ] 70-60 ] |
| | | 112 | المجموع |

عدد العمال الذين تزيد أجورهم عن 40000 دج هو 39 (112 – (18+30+25)) أو ( 17+12+10)

عدد العمال الذين تقل أجورهم عن 60000 دج هو 10 (112 – (18+30+25+17+12) )

**I. 3. التوزيع التكراري النسبي**: وهو حاصل قسمة التكرار المطلق على مجموع التكرارات أي : $f_i = \frac{n_i}{\sum_{i=1}^{k} n_i}$ ، حيث مجموع التكرار النسبي يساوي الواحد الصحيح، أي :

$$\sum_{i=1}^{k} f_i = \frac{n_1}{\sum_{i=1}^{k} n_i} + \frac{n_2}{\sum_{i=1}^{k} n_i} + \cdots + \frac{n_n}{\sum_{i=1}^{k} n_i} = 1$$

من بيانات مثال (01) ، يمكننا حساب التكرار النسبي بالشكل التالي :

| التكرار النسبي $f_i$ | عدد العمال $n_i$ | الأجر |
|---|---|---|
| 0,16 | 18 | $k=1$ ] 20-10 ] |
| 0,27 | 30 | $k=2$ ] 30-20 ] |

| | | |
|---|---|---|
| $k = 3$ ] 40-30 ] | 25 | 0,22 |
| $k = 4$ ] 50-40 ] | 17 | 0,15 |
| $k = 5$ ] 60-50 ] | 12 | 0,11 |
| $k = 6$ ] 70-60 ] | 10 | 0,09 |
| المجموع $\sum_{i=1}^{k} n_i$ | 112 | $\sum_{i=1}^{k} f_i = 1$ |

I. 4. **التوزيع التكراري النسبي المتجمع:** يتم فيه تجميع أو طرح التكرارات النسبية بدلا من التكرارات المطلقة ، و من بيانات مثال (01) ، نجد التكرار النسبي المتجمع الصاعد ويرمز له بالرمز $f_i^{\nearrow}$، و التكرار النسبي المتجمع النازل و يرمز له بالرمز $f_i^{\searrow}$ بالشكل التالي:

| الأجر | عدد العمال $n_i$ | التكرار النسبي $f_i$ | التكرار النسبي المتجمع الصاعد $f_i^{\nearrow}$ | التكرار النسبي المتجمع النازل $f_i^{\searrow}$ |
|---|---|---|---|---|
| ] 20-10 ] | 18 | 0,16 | 0,16 | 1 |
| ] 30-20 ] | 30 | 0,27 | 0,43 | 0,84 |
| ] 40-30 ] | 25 | 0,22 | 0,65 | 0,57 |
| ] 50-40 ] | 17 | 0,15 | 0,80 | 0,35 |
| ] 60-50 ] | 12 | 0,11 | 0,91 | 0,2 |
| ] 70-60 ] | 10 | 0,09 | 1 | 0,09 |
| المجموع | 112 | 1 | | |

## II. العرض البياني :

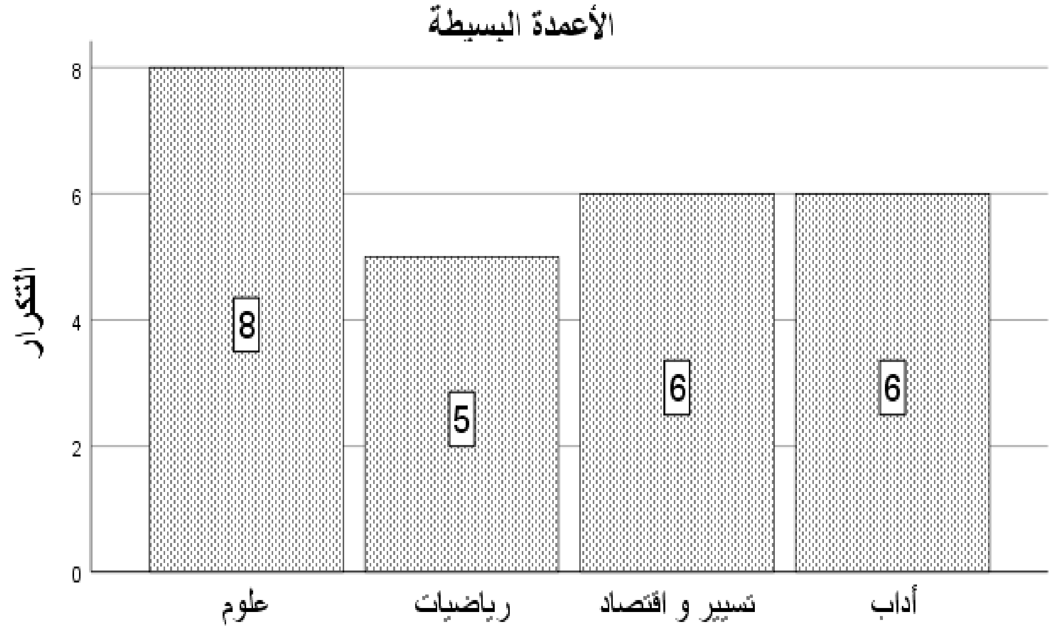
**II. 1. الأعمدة البيانية:** من أفضل الطرق البيانية وأوضحها ، وهي عبارة عن مستطيلات رأسية كل منها ذو سمك مناسب و متساو، و ارتفاعها يمثل قيم المشاهدات ( التكرارات) للظاهرة محل الدراسة ، و فيما يلي أنواع هذه الأعمدة البيانية :

- **الأعمدة البيانية البسيطة** : تستخدم الأعمدة البيانية البسيطة لتمثيل ظاهرة واحدة محل الدراسة.

**مثال (03)** :الجدول الموالي يوضح شعب 25 طالب متحصل على بكالوريا تم استقصائهم أثناء عملية تسجيل سنة أولى جامعي، و المطلوب تمثيل هذه البيانات على شكل أعمدة بيانية.

| المتغير الإحصائي $x_i$ | التكرار $n_i$ |
|---|---|
| علوم | 08 |
| رياضيات | 05 |
| تسيير و اقتصاد | 06 |
| أداب | 06 |
| المجموع | 25 |

والشكل البياني التالي يوضح التمثيل البياني للبيانات أعلاه

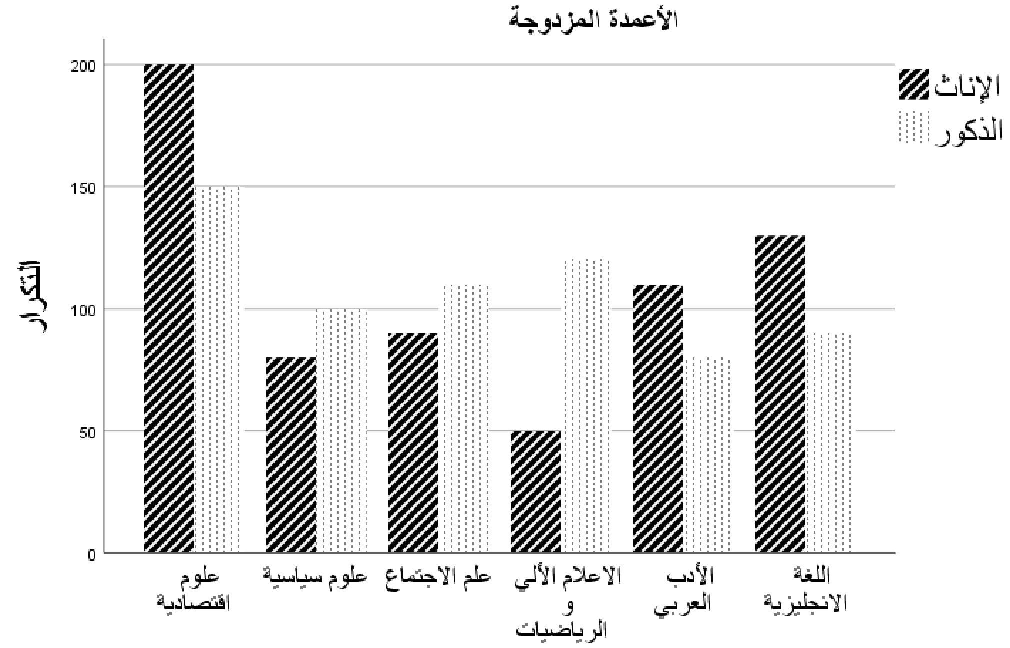



❖ **الأعمدة البيانية المزدوجة** : تستخدم الأعمدة البيانية المزدوجة في حالة ما إذا أردنا المقارنة بين ظاهرتين أو أكثر كالمقارنة ما بين طلاب جامعتين ، أو تحصيل الدراسي ما بين الذكور و الإناث، أو مقارنة الإنفاق و الدخل ، و تمثل كل ظاهرة بمستطيل يلاصق مستطيل الظاهرة الثانية .

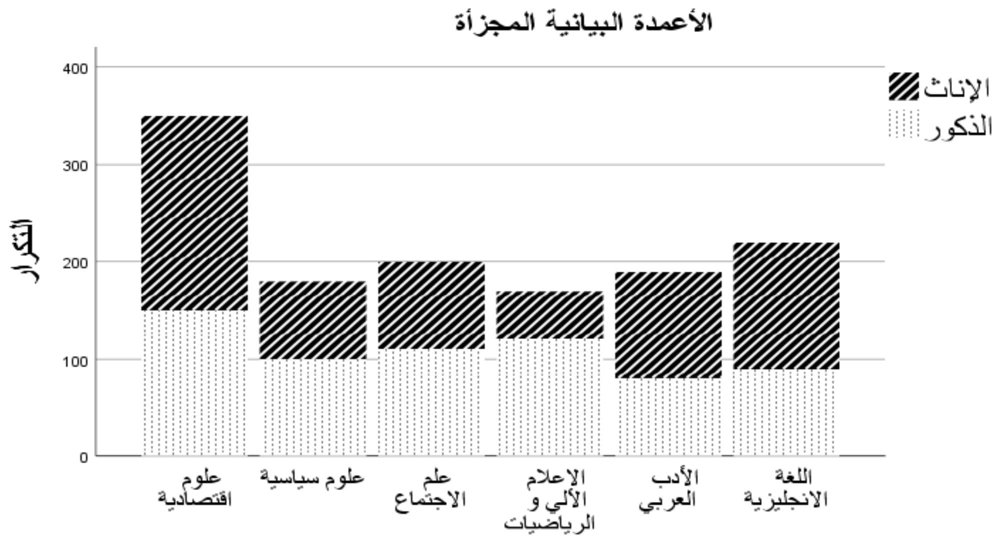
**مثال (04)** : البيانات التالية توضح عدد الطلبة إناث و ذكور المسجلين سنة أولى جذع مشترك في مجموعة من التخصصات ، و المطلوب تمثيل هذه البيانات على شكل أعمدة بيانية .

| التخصص | علوم اقتصادية | علوم سياسية | علم الاجتماع | الإعلام الآلي والرياضيات | الأدب العربي | اللغة الإنجليزية |
|---|---|---|---|---|---|---|
| الإناث | 200 | 80 | 90 | 50 | 110 | 130 |
| الذكور | 150 | 100 | 110 | 120 | 80 | 90 |

و الشكل التالي يوضح تمثيل البيانات أعلاه على شكل أعمدة مزدوجة




❖ **الأعمدة البيانية المجزأة** : تستخدم الأعمدة البيانية المجزأة في حالة وجود أكثر من ظاهرة كما هو الحال بالنسبة للأعمدة البيانية المزدوجة السابقة، ولكن في هذه الحالة يتم رسم عمود واحد لمجموع القيم لبيانات الظاهرتين المرغوب تمثيلها، ثم يقسم المستطيل بنسبة عدد البيانات لكل ظاهرة. وبالاعتماد على بيانات مثال (04) ، فإن التمثيل على شكل الأعمدة البيانية المجزأة يكون بالشكل التالي:




**II. 2.الشكل الدائري**: يستخدم الشكل الدائري في غالب الأحيان، لتقديم بيانات ظاهرة ما تتركب من مكونات جزئية خلال ظرف زماني أو مكاني محددين، كتوزيع كميات التساقط حسب الفصول في سنة ما، أو توزيع مساحة مزروعة حسب أنواع المزروعات، و تقديم البيانات عن طريق الشكل الدائري حساب الزاوية المقابلة لمكونات الظاهرة، حيث أن القيمة الكلية للظاهرة تقابل الزاوية الكلية للدائرة أي 360 درجة ، بينما القيم الجزئية للظاهرة يقابلها x الذي يشير الزوايا الجزئية للدائرة الكلية.

**مثال(05)** : البيانات التالية تظهر توزيع مساحة مزرعة ما حسب أنواع المزروعات خلال الموسم الفلاحي 2003- 2004 ، و المطلوب : تقديم هذه البيانات على شكل الدائرة النسبية .

| المزروعات | الفلاحة ( بالهكتار) |
|---|---|
| الحبوب | 350 |
| الخضر | 650 |
| الفواكه | 500 |
| م.أخرى | 500 |

لتمثيل البيانات أعلاه على شكل دائرة نسبية ، يتطلب منا إيجاد الزوايا و النسب المئوية المقابلة لكل نوع من المزروعات ، و الحسابات اللازمة موضحة بالشكل التالي :

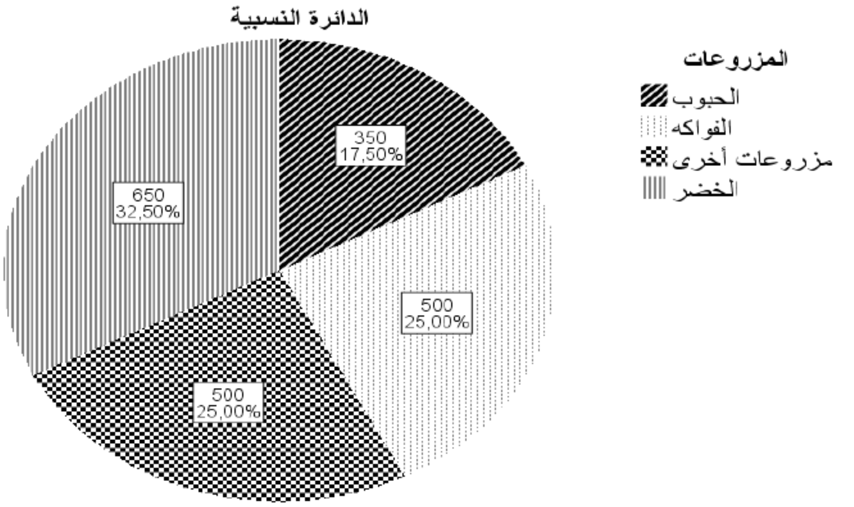
$$2000 \longrightarrow 360$$
$$350 \longrightarrow X$$

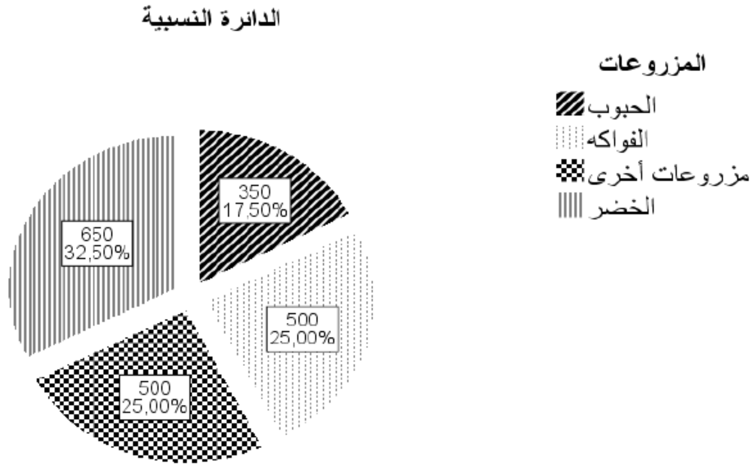
أما النسب المئوية يتم حسابها بالشكل التالي:

$$2000 \longrightarrow 100$$
$$350 \longrightarrow X$$

| المزروعات | الفلاحة ( بالهكتار) | الزاوية المقابلة بالدرجات | النسب المئوية المقابلة |
|---|---|---|---|
| الحبوب | 350 | 63 | %17.5 |
| الخضر | 650 | 117 | %32.5 |
| الفواكه | 500 | 90 | %25 |
| م.أخرى | 500 | 90 | %25 |

و الأشكال الموالية توضح الرسم البياني بالاعتماد على الدائرة النسبية على نوعين :





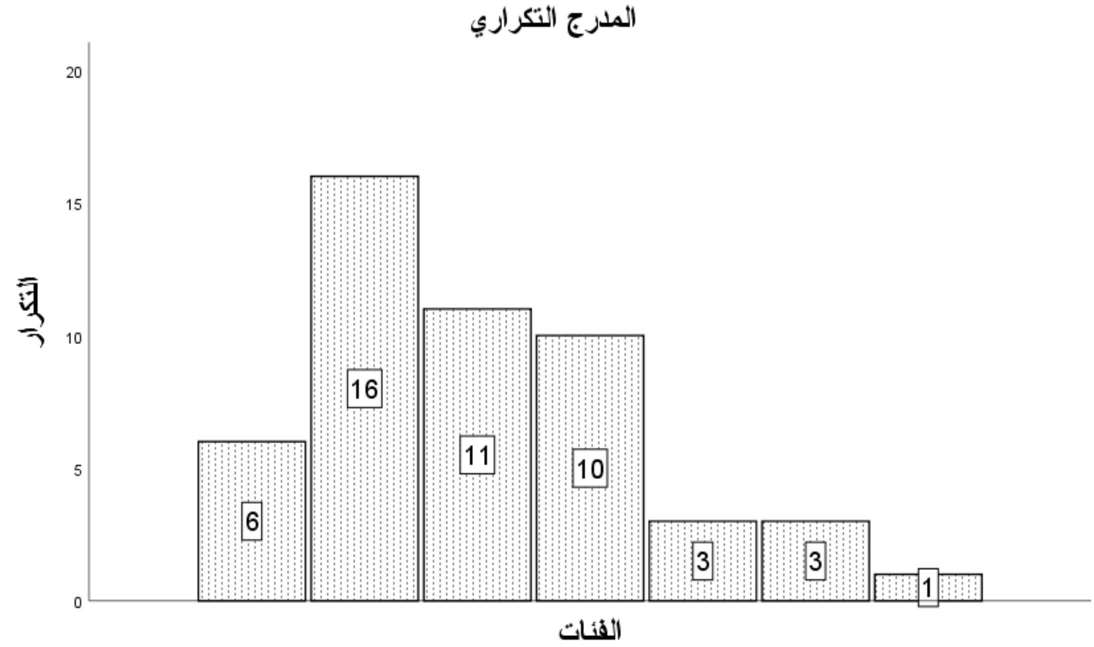

**II. 3. المدرج التكراري** : إن تمثيل البيانات في مدرج تكراري يعني تمثيل تكرار كل فئة من فئات التوزيع التكراري بمستطيل تكون حدود قاعدته هي حدود الفئات و ارتفاعه هو تكرار تلك الفئة ، لذلك نرسم محورين أحدهما المحور الأفقي و نضع عليه الحدود الفعلية للفئات و المحور العمودي و نضع عليه تكرار الفئات .

**مثال (06)** : البيانات التالية توزيع عمال مؤسسة ما حسب أجورهم الشهرية ( الوحدة : $10^3$ دج) ،والمطلوب تمثيل هذه البيانات في مدرج تكراري.

| الفئات | التكرارات |
|---|---|
| ]39-35] | 06 |
| ]43-39] | 16 |
| ]47-43] | 11 |
| ]51-47] | 10 |
| ]55-51] | 03 |
| ]59-55] | 03 |

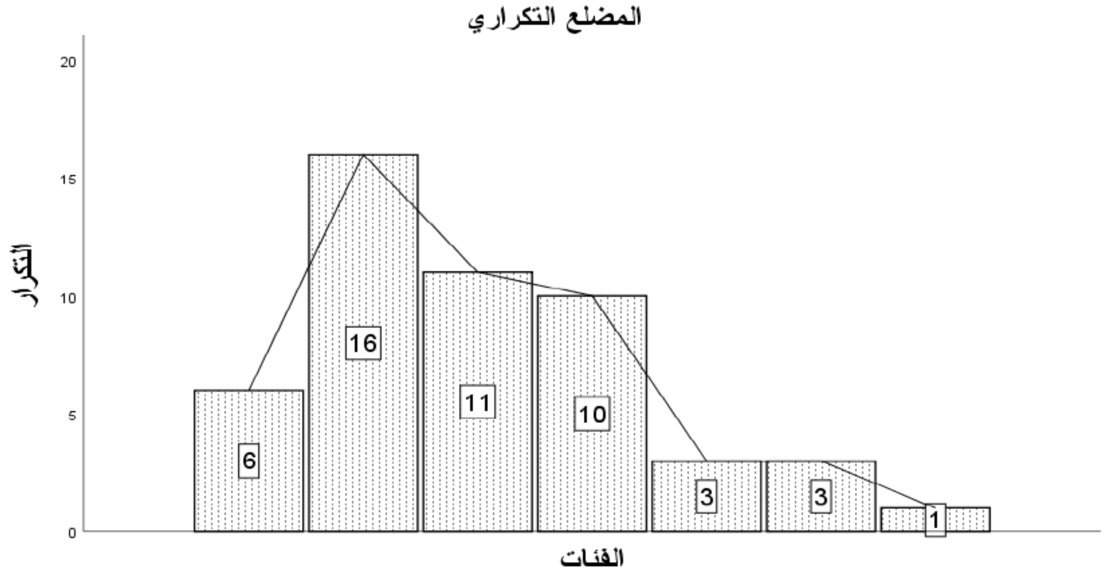
| | |
|---|---|
| ]63-59] | 01 |
| المجموع | 50 |

و الشكل التالي يوضح تمثيل البيانات أعلاه على شكل مدرج تكراري:

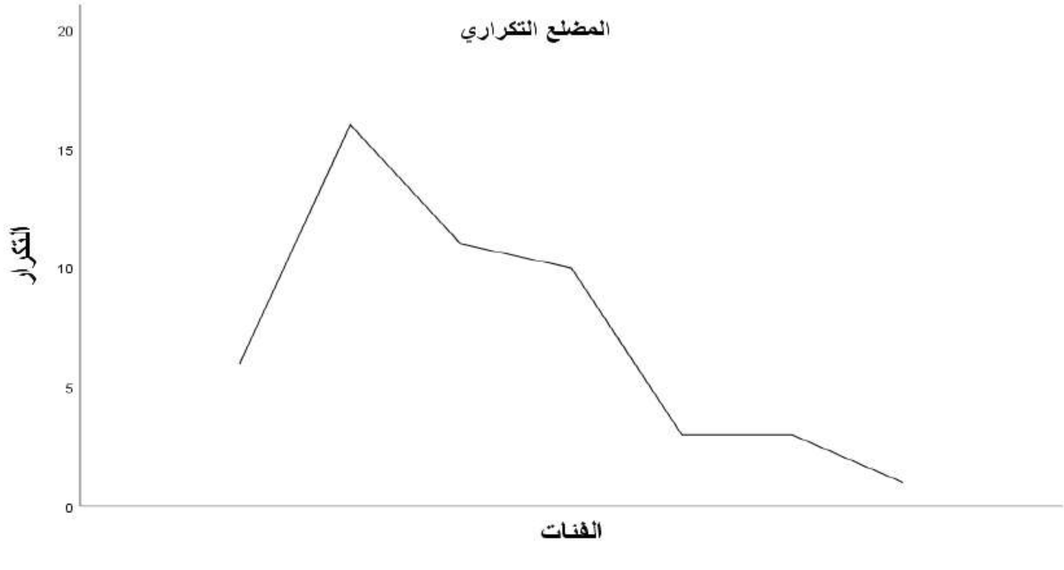



**II**. **4**. **المضلع التكراري** : يرسم المضلع التكراري على محورين الأفقي يمثل الفئات و العمودي يمثل التكرار ، مثلما ورد شرحه في المدرج التكراري ، و يمكن تعيين على المدرج التكراري نقاط إحداثياتها مراكز الفئات و التكرارات المقابلة لها، ثم نصل بينها بخط مستقيم ، و لكي نحصل على مضلع على شكل الخط المنكسر ، يتم إحداث مركز تصوري سابق لأول فئة ، و أخر لاحق لأخر فئة ، تكراراتهما معدومة ، ثم نصل طرف الخط المنكسر بهاتين النقطتين .

**مثال (07)** : باستخدام بيانات مثال (06) ، قدم هذه البيانات على شكل مضلع تكراري:

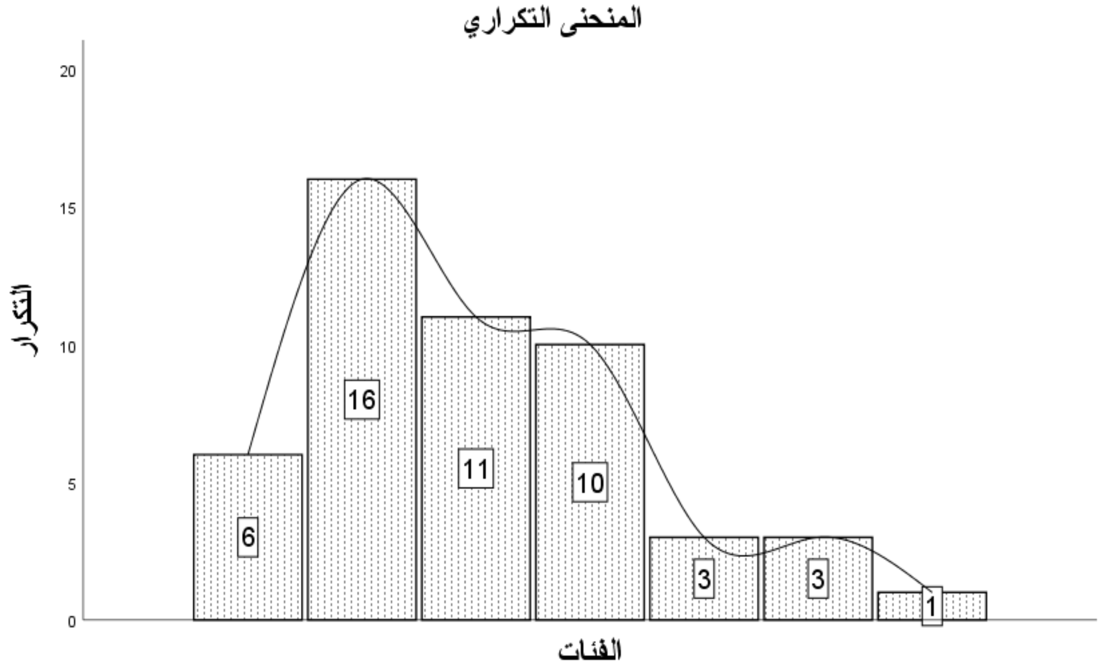



و الشكل النهائي للمضلع هو كالتالي:

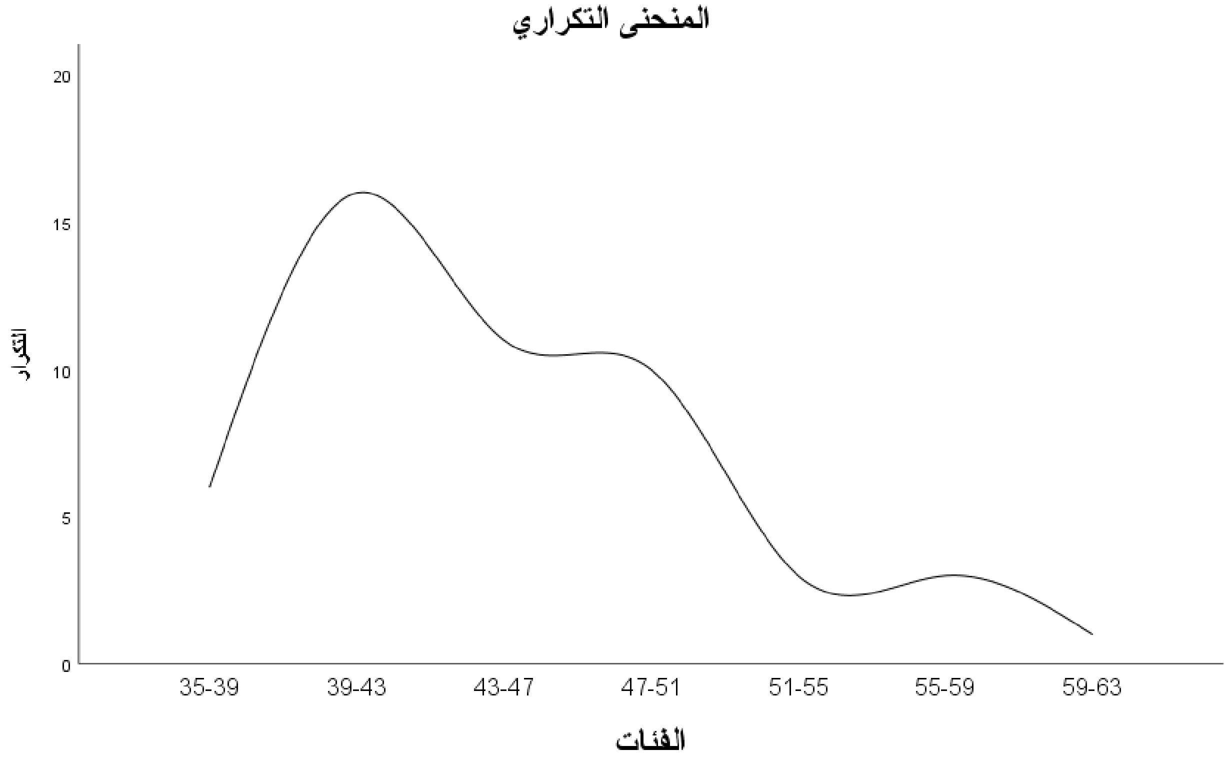



**II. 5. المنحنى التكراري**: يرسم المنحنى التكراري بنفس الكيفية التي رسم بها المضلع التكراري، لكن يتم الوصل ما بين النقاط بخط ممهد لنحصل على منحنى انسيابي .

**مثال (08)** : باستخدام بيانات مثال (06) ، قدم هذه البيانات على شكل منحنى تكراري:




لنحصل في الأخير على الشكل التالي:



## II. 6. المدرج التكراري الصاعد و النازل :

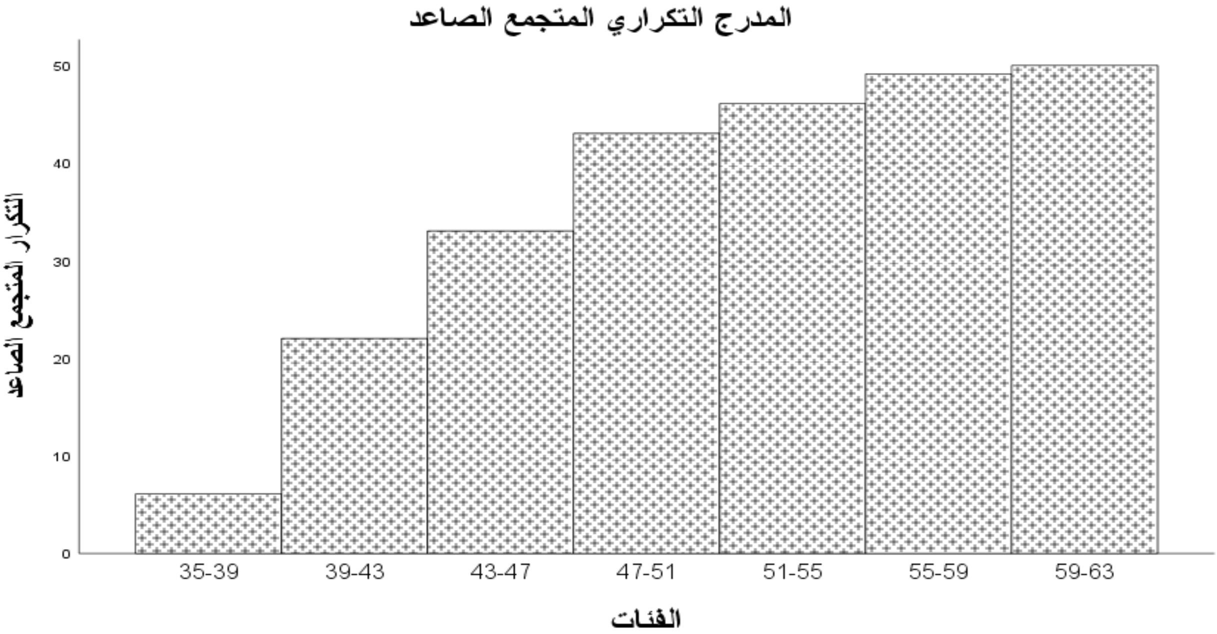
❖ **المدرج التكراري الصاعد** :يتم رسمه من خلال وضع الفئات على المحور الأفقي و التكرار المتجمع الصاعد على المحور العمودي ، و يرسم بالكيفية التي يرسم بها المدرج التكراري.

**مثال (09)** : باستخدام بيانات مثال (06) ، قدم هذه البيانات على شكل المدرج التكراري المتجمع الصاعد:

قبل رسم المدرج التكراري الصاعد، سيتم حساب التكرار المتجمع الصاعد ،كما هو موضح في الجدول التالي:

| الفئات | التكرارات | التكرار المتجمع الصاعد |
|---|---|---|
| [35-39[ | 06 | 6 |
| [39-43[ | 16 | 22 |
| [43-47[ | 11 | 33 |
| [47-51[ | 10 | 43 |
| [51-55[ | 03 | 46 |
| [55-59[ | 03 | 49 |
| [59-63[ | 01 | 50 |
| المجموع | 50 | |

و بذلك نحصل على المدرج التكراري المتجمع الصاعد بالشكل التالي:




❖ **المدرج التكراري النازل**: يتم رسمه من خلال وضع الفئات على المحور الأفقي و التكرار المتجمع النازل على المحور العمودي ، و يرسم بالكيفية التي يرسم بها المدرج التكراري.

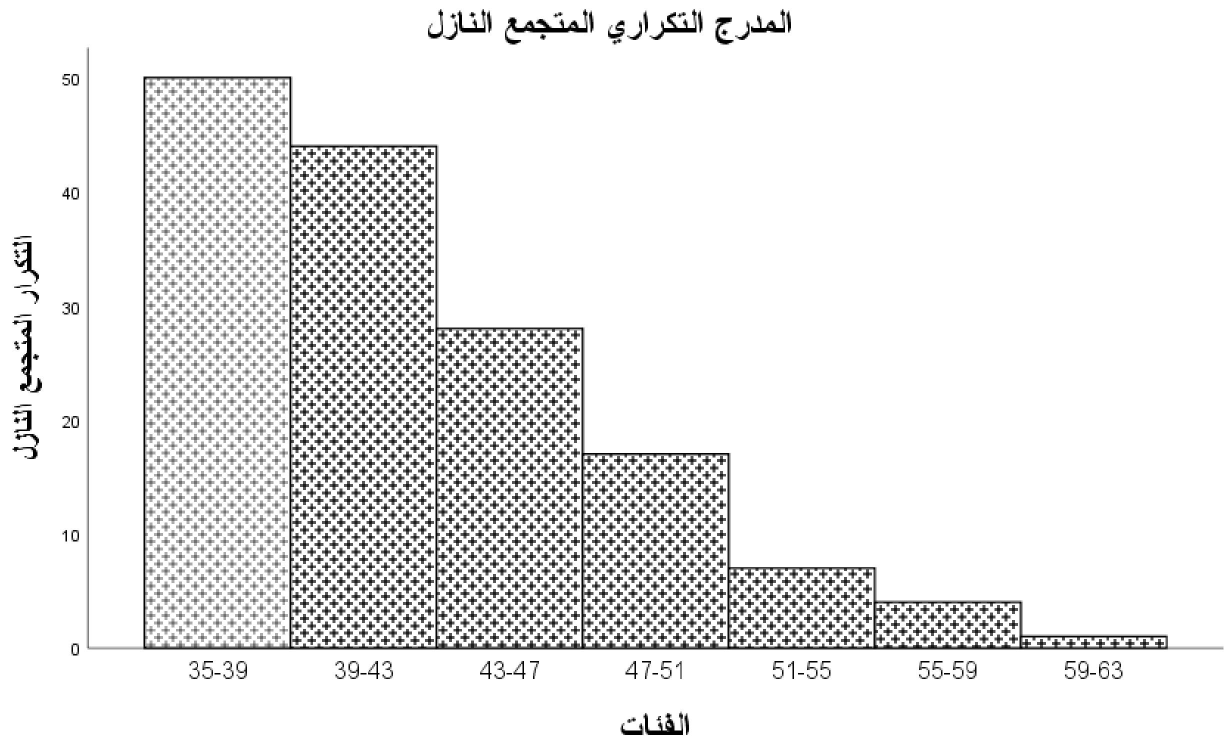
**مثال (10)** : باستخدام بيانات مثال (06) ، قدم هذه البيانات على شكل المدرج التكراري المتجمع النازل:

قبل رسم المدرج التكراري الصاعد، سيتم حساب التكرار المتجمع الصاعد ، كما هو موضح في الجدول التالي:

| الفئات | التكرارات | التكرار المتجمع النازل |
|---|---|---|
| [35-39[ | 06 | 50 |
| [39-43[ | 16 | 44 |
| [43-47[ | 11 | 28 |
| [47-51[ | 10 | 17 |
| [51-55[ | 03 | 7 |

| | | |
|---|---|---|
| [55-59[ | 03 | 4 |
| [59-63[ | 01 | 1 |
| المجموع | 50 | |

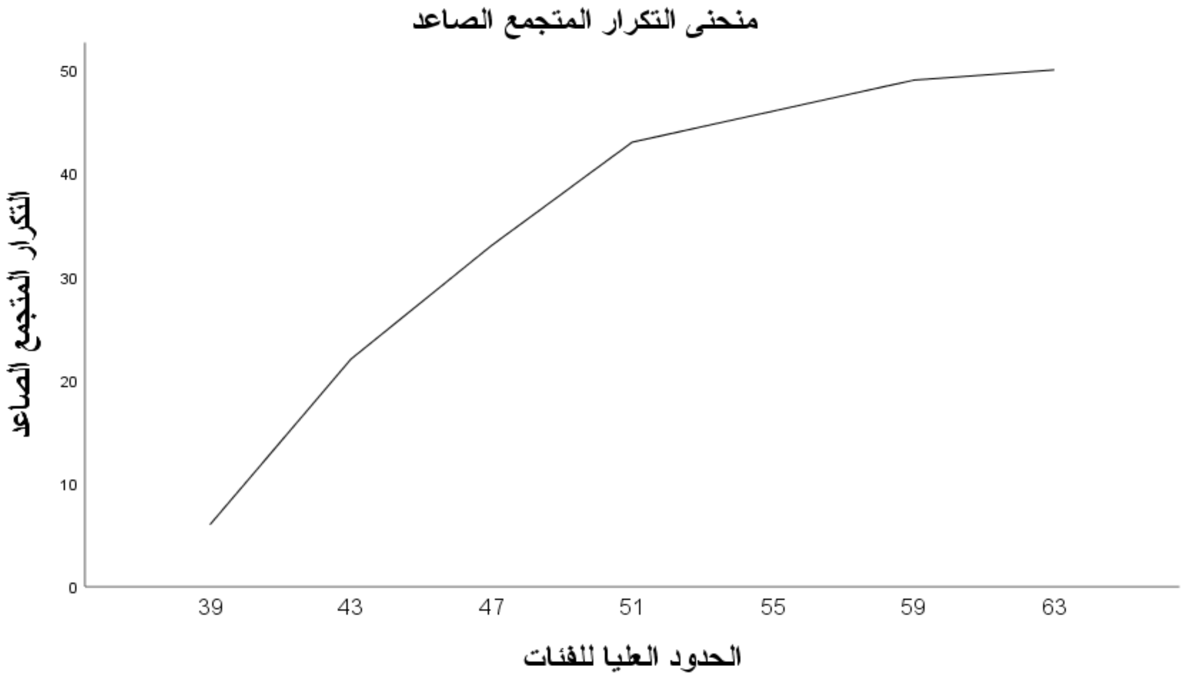
و بذلك نحصل على المدرج التكراري المتجمع الصاعد بالشكل التالي:




## II. 7. منحنى التكرار المتجمع الصاعد و النازل:

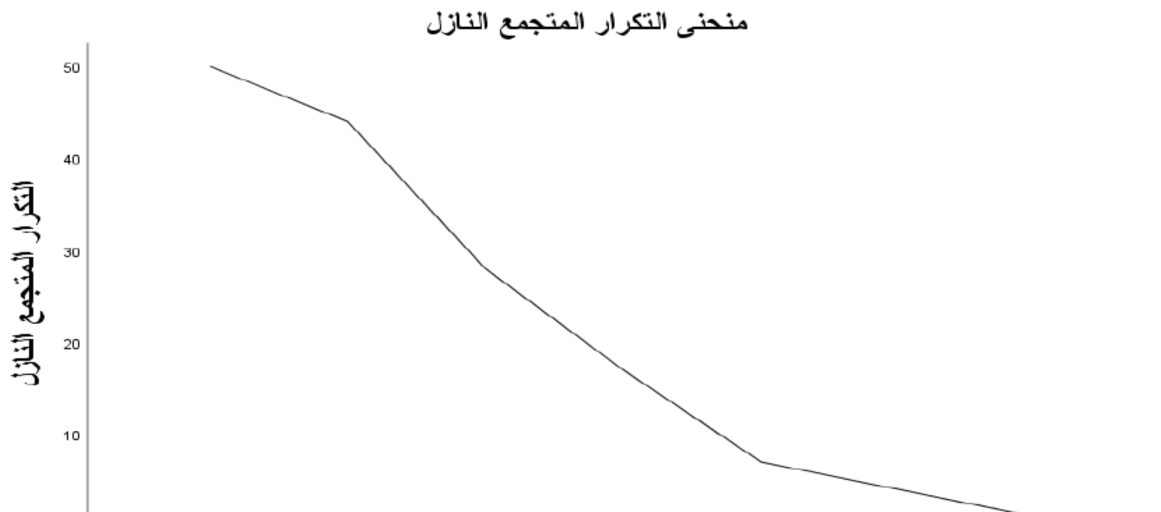
❖ **منحنى التكرار المتجمع الصاعد:** ويتم رسمه انطلاقا من المدرج التكراري المتجمع الصاعد ، و ذلك بوصل بين النقاط التي تمثل الحدود العليا للفئات و التكرارات المقابلة لها في المدرج التكراري المتجمع الصاعد.

**مثال (11)** : باستخدام بيانات مثال (06) ، قدم هذه البيانات على شكل المدرج التكراري المتجمع الصاعد:




❖ **منحنى التكرار المتجمع النازل:** ويتم رسمه انطلاقا من المدرج التكراري المتجمع النازل ، و ذلك بوصل بين النقاط التي تمثل الحدود العليا للفئات و التكرارات المقابلة لها في المدرج التكراري المتجمع النازل.

**مثال (12)** : باستخدام بيانات مثال (06) ، قدم هذه البيانات على شكل المدرج التكراري المتجمع النازل:



# مقاييس النزعة المركزية و غير المركزية

## أولا : مقاييس النزعة المركزية:

**I. الوسط الحسابي** : يعد من أهم مقاييس النزعة المركزية ، و أكثرها استخداما ، و سهل الحساب سواء في البيانات غير المبوبة أو البيانات المبوبة، و يرمز له بالرمز $\bar{x}$، ومن أهم خصائص المتوسط الحسابي نذكر (سيف النصر سيد، الديب، و رفيق جلال، 2018، صفحة 07):

- يأخذ بعين الاعتبار جميع القيم في حسابه، و يتأثر الوسط الحسابي بالقيم المتطرفة (الشاذة)، وبالتالي لا يفضل الاعتماد عليه في حالة وجود هذه القيم.
- مجموع انحرافات القيم عن المتوسط الحسابي يساوي الصفر.
- يصعب حساب الوسط الحسابي في حالة الفئات المفتوحة ، لأنه من الصعب تحديد مراكز الفئات .
- لا يمكن تحديده بيانيا

**1. حساب الوسط الحسابي في حالة البيانات غير المبوبة** : و يحسب بالشكل التالي :

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

حيث : $\sum_{i=1}^{n} x_i$ هو مجموع قيم وحدات العينة أو مجموع المشاهدات، و $n$ حجم العينة .

**مثال ( 13)** : في ما يلي درجات عشر طلاب في مادة الإحصاء ، و المطلوب حساب متوسط درجات الطلاب:

20، 18، 10، 08، 15، 06، 09، 12، 14، 05

**الحل**: في هذا المثال فإن حجم العينة هو 10 ، و متوسط درجات الطلاب يحسب بالشكل التالي :

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

$$= \frac{20 + 18 + 10 + 08 + 15 + 06 + 09 + 12 + 14 + 05}{10}$$

$$\bar{x} = 11.7$$

2. **حساب المتوسط الحسابي في حالة البيانات المبوبة** : و يحسب بالشكل التالي :

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

حيث $\sum_{i=1}^{n} x_i n_i$ تمثل مجموع ضرب المشاهدات مع التكرارات المقابلة لها ، و $\sum_{i=1}^{n} n_i$

تمثل مجموع التكرارات، في حالة الفئات فتصبح الصيغة من الشكل التالي:

$$\bar{x} = \frac{\sum_{i=1}^{n} c_i n_i}{\sum_{i=1}^{n} n_i}$$

حيث : $c_i$ تمثل مراكز الفئات.

**ملاحظة :**

يطلق على المتوسط الحسابي الخاص بالمجتمع بالتوقع الرياضي أو الأمل الرياضي ويرمز له بالرمز $\mu$ أو $E(X)$ و يحسب في حالة البيانات غير المبوبة بالشكل التالي :

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

حيث : $\sum_{i=1}^{n} X_i$ تمثل مجموع وحدات المجتمع و $N$ حجم المجتمع

أما في حالة البيانات المبوبة ، فإن الموسط الحسابي يحسب بالشكل التالي:

$$\mu = \frac{\sum_{i=1}^{N} X_i N_i}{\sum_{i=1}^{N} N_i}$$

حيث : $\sum_{i=1}^{N} X_i N_i$ تمثل حاصل ضرب المشاهدات مع التكرارات المقابلة لها ، و $\sum_{i=1}^{N} N_i$ تمثل مجموع التكرارات

**مثال (14)** : البيانات التالية تمثل عدد الحوادث في إحدى المدن خلال شهر، و المطلوب حساب متوسط الحوادث اليومية :

| عدد الحوادث $x_i$ | عدد الأيام $n_i$ |
|---|---|
| 0 | 05 |
| 1 | 06 |
| 2 | 04 |
| 3 | 06 |
| 4 | 05 |
| 5 | 04 |
| المجموع | 30 |

**الحل** :

| $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | $\Sigma$ |
|---|---|---|---|---|---|---|---|
| $n_i$ | 05 | 06 | 04 | 06 | 05 | 04 | 30 |
| $x_i.n_i$ | 0 | 06 | 08 | 18 | 20 | 20 | 72 |

متوسط الحوادث اليومية هي :

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} = \frac{72}{30} = 2.4$$

**مثال (15)** : لدينا جدول التوزيع التكراري الموالي والذي تتعلق بياناته بالخبرة المهنية بالسنوات لعينة عددها 74 عامل في الحقل الأكاديمي الجامعي ، والمطلوب حساب متوسط الخبرة المهنية للعاملين :

| الفئات ( الخبرة المهنية) $x_i$ | التكرارات ( العاملين) $n_i$ |
|---|---|
| [06-02 ] | 07 |
| [11-07] | 09 |

| | |
|---|---|
| [16-12] | 15 |
| [21-17] | 20 |
| [26-22] | 12 |
| [31-27] | 08 |
| [36-32] | 03 |
| المجموع | 74 |

**الحل:**

| $x_i$ | $n_i$ | $c_i$ | $c_i.n_i$ |
|---|---|---|---|
| [06-02] | 07 | 04 | 28 |
| [11-07] | 09 | 09 | 81 |
| [16-12] | 15 | 14 | 210 |
| [21-17] | 20 | 19 | 380 |
| [26-22] | 12 | 24 | 288 |
| [31-27] | 08 | 29 | 232 |
| [36-32] | 03 | 34 | 102 |
| المجموع | 74 | / | 1321 |

بتطبيق صيغة الوسط الحسابي نجد متوسط الخبرة المهنية للعاملين كما يلي:

$$\bar{x} = \frac{\sum_{i=1}^{n} c_i n_i}{\sum_{i=1}^{n} n_i} = \frac{1321}{74} = 17.85 \approx 18$$

**II. الوسيط** : يستخدم الوسيط كمقياس للنزعة المركزية بدلا من الوسط الحسابي عندما تكون بالبيانات قيم شاذة ، كما يفضل استخدامه في حالة الفئات المفتوحة ، لأنه يهتم بالقيم الواقعة

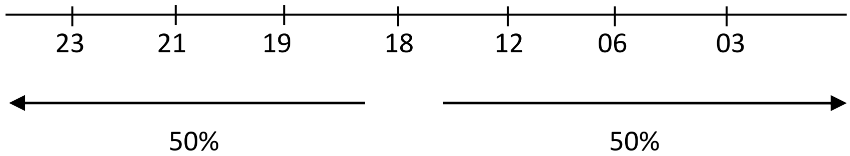
في المنتصف و يهمل الأطراف ، بالإضافة إلى أنه يمكن استخدامه في حالة المتغيرات الوصفية (الكيفية) التي لا نعبر عنها بالأرقام (غرايبة و المنيزل، 2007، صفحة 59)، كما يعرف الوسيط كذلك بالقيمة التي يسبقها 50% من المشاهدات بعد ترتيبها ، فإذا كان عدد المشاهدات فردي فالوسيط هو المشاهدة التي في المنتصف و إذا كان عدد المشاهدات زوجي فإن الوسيط هو متوسط المشاهدتين اللتين في المنتصف (بري، هندي، و راضي، أساسيات طرق التحليل الاحصائي ، 1998، صفحة 33).

**1. حساب الوسيط في حالة البيانات غير المبوبة** : لحساب الوسيط في حالة البيانات غير المبوبة ، يتم المرور على المراحل التالية :

- ترتيب البيانات ترتيبا تصاعديا أو تنازليا .
- تحديد رتبة الوسيط عن طريق :
  - إذا كان حجم العينة فرديا ، فيتم إيجاد الوسيط بتطبيق الصيغة التالية : $\frac{n+1}{2}$
  - إذا كان حجم العينة زوجيا ، فإن للوسيط ترتيبين ، يتم إيجاد الترتيب الأول بتطبيق الصيغة التالية : $\frac{n}{2}$، أما الترتيب الثاني بتطبيق الصيغة التالية : $\frac{n}{2}+1$
- استخراج القيمة التي تقابل ترتيب الوسيط.

**مثال (16)** : إذا كان لدينا البيانات التالية مرتبة تصاعديا: 03،06، 12، 18، 19، 21، 23، أوجد الوسيط .

**الحل:** بما أن حجم العينة فردي ، فإن ترتيب الوسيط يحسب بالشكل $\frac{n+1}{2} = \frac{7+1}{2} = 4$ ، أي ترتيب الوسيط هو الرتبة الرابعة ، و بالتالي فإن قيمة الوسيط هي 18 . حيث يقسم الوسيط البيانات إلى قسمين متساويين مثلما يبين الشكل التالي:




**مثال (17)** : إذا كان لدينا البيانات التالية مرتبة تصاعديا: 18، 23، 27، 28، 29، 40، 44، 46

**الحل:** بما أن حجم العينة زوجي، فإن للوسيط ترتيبين و هما : $\frac{8}{2} = \frac{n}{2} = 4$ و $\frac{n}{2} + 1 = \frac{8}{2} + 1 = 5$ ، إذا الوسيط يقع ما بين القيمة الرابعة والتي تساوي 28 و القيمة الخامسة والتي تساوي 29 و يحسب بالشكل التالي: $(28+29)/2 = 28.5$ .

### 2. حساب الوسيط في حالة البيانات المبوبة :

❖ **بالنسبة للمتغير الكمي المنفصل** : يتم حساب الوسيط باستخدام الخطوات التالية :

- حساب التكرار المتجمع الصاعد
- حساب ترتيب الوسيط باستخدام المعادلتين التاليتين :
  - إذا كان مجموع التكرارات زوجيا : $C = \frac{\sum n_i}{2}$
  - إذا كان مجموع التكرارات فرديا : $C = \frac{\sum n_i + 1}{2}$

**مثال (18)** : أوجد وسيط البيانات التالية :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| عدد العطل المرضية | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| عدد الموظفين | 15 | 10 | 09 | 07 | 03 | 04 | 02 |

**الحل:**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| عدد العطل المرضية | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| عدد الموظفين | 15 | 10 | 09 | 07 | 03 | 04 | 02 |
| التكرار المتجمع الصاعد | 15 | 25 | 34 | 41 | 44 | 48 | 50 |

إذا كان مجموع التكرارات زوجيا :

$$C = \frac{\sum n_i}{2} = \frac{50}{2} = 25$$

وعليه فإن الوسيط هو القيمة المقابلة لرتبة الوسيط وهي 02

**مثال (19)** : أوجد وسيط البيانات التالية :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| عدد العطل المرضية | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| عدد الموظفين | 15 | 10 | 09 | 07 | 03 | 04 | 03 |

**الحل:**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| عدد العطل المرضية | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| عدد الموظفين | 15 | 10 | 09 | 07 | 03 | 04 | 03 |
| التكرار المتجمع الصاعد | 15 | 25 | 34 | 41 | 44 | 48 | 51 |

إذا كان مجموع التكرارات فرديا :

$$C = \frac{\sum n_i + 1}{2} = \frac{51 + 1}{2} = 26$$

في هذه الحالة فإن قيمة الوسيط هي 3

❖ **بالنسبة للمتغير الكمي المتصل** : يتم حساب الوسيط باستخدام الخطوات التالية :

- إيجاد التكرار المتجمع الصاعد.
- إيجاد رتبة الوسيط بتطبيق الصيغة التالية : $\frac{\sum n_i}{2}$
- تطبيق الصيغة التالية :

$$\mathrm{M_e} = A + \frac{(\frac{\sum n_i}{2} - n_1^{\nearrow})}{n_2^{\nearrow} - n_1^{\nearrow}} \times L$$

$A$:الحد الأدنى للفئة الوسيطية

$\frac{\sum n_i}{2}$:ترتيب الوسيط

$n_1^{\nearrow}$:التكرار المتجمع الصاعد الذي يسبق ترتيب الوسيط

$n_2^{\nearrow}$:التكرار المتجمع الصاعد الذي يلي ترتيب الوسيط

**مثال (20)** : أحسب وسيط البيانات التالية :

| المتغير | التكرار المطلق $n_i$ |
|---|---|
| ] 16000-11000 ] | 13 |
| ] 21000-16000 ] | 16 |
| ] 26000-21000 ] | 12 |
| ] 31000-26000 ] | 3 |

| | |
|---|---|
| ] 36000-31000 ] | 4 |
| ] 41000-36000 ] | 1 |
| ] 46000-41000 ] | 1 |
| المجموع | 50 |

**الحل**: لحساب الوسيط ، يجب حساب التكرار المتجمع الصاعد في أول مرحلة :

| المتغير | التكرار المطلق $n_i$ | التكرار المطلق التجميعي الصاعد $n_i^{\nearrow}$ |
|---|---|---|
| ] 16000-11000 ] | 13 | $n_1^{\nearrow}$ 13 |
| ] 21000-16000 ] $A$ | 16 | $n_2^{\nearrow}$ 29 |
| ] 26000-21000 ] | 12 | 41 |
| ] 31000-26000 ] | 3 | 44 |
| ] 36000-31000 ] | 4 | 48 |
| ] 41000-36000 ] | 1 | 49 |
| ] 46000-41000 ] | 1 | 50 |
| المجموع | 50 | |

و بتطبيق صيغة الوسيط ، نجد :

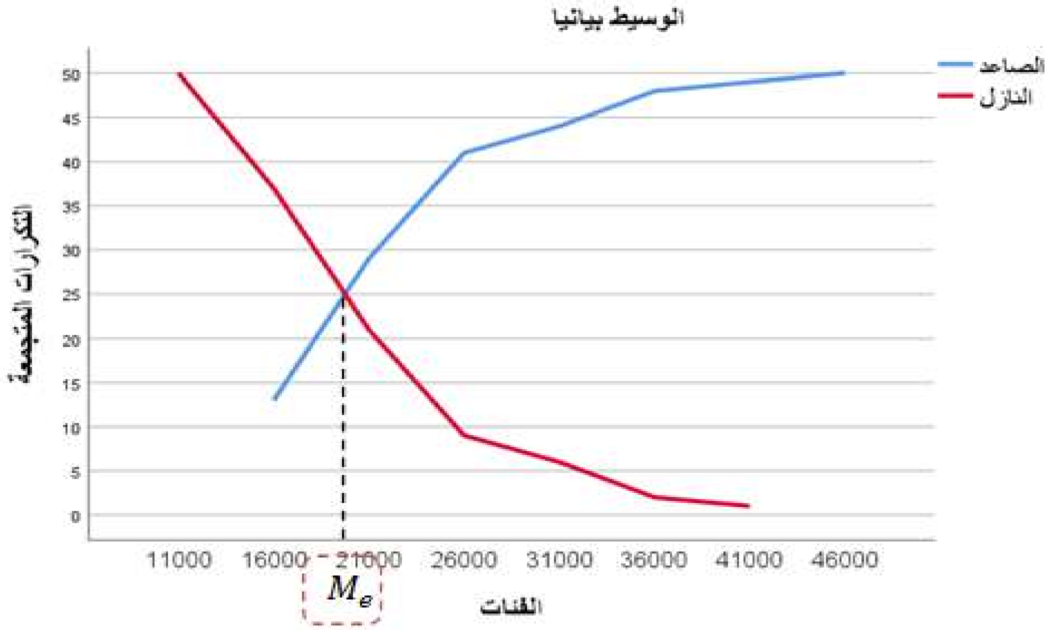
$$\mathrm{M_e} = 16000 + \frac{(25-13)}{29-13} \times 5000 = 19750$$

3. **تحديد الوسيط بيانيا** : نحصل على الوسيط بيانيا من خلال رسم المنحنيين المتجمعين الصاعد و النازل ، ومن ثم إنزال خط عمودي من نقطة التقاء المنحنين على المحور الأفقي ، و النقطة التي سيقع عليها الخط العمودي على المحور الأفقي ستمثل قيمة الوسيط.

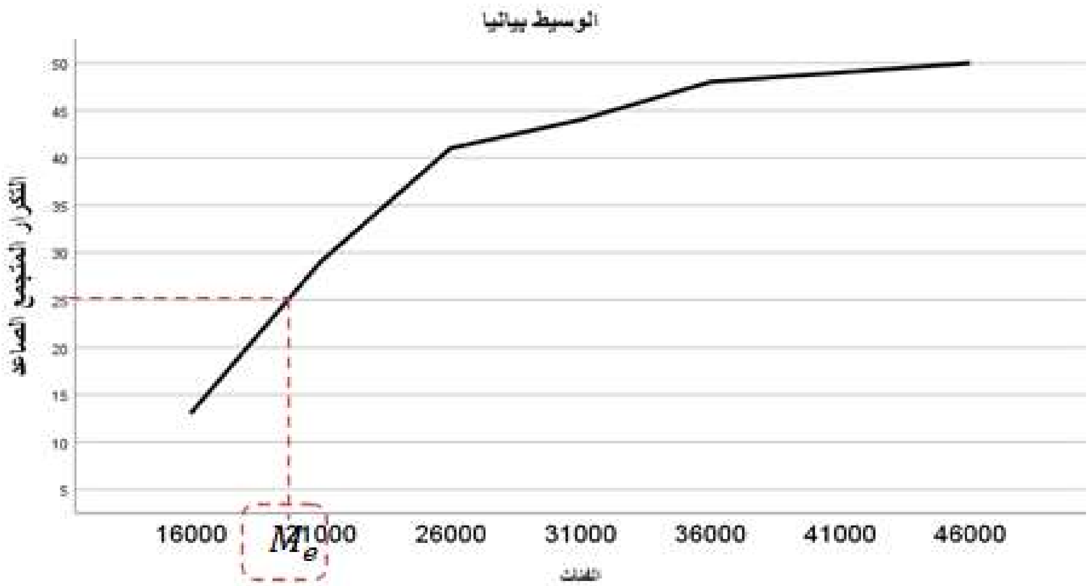
**مثال (21)** : باستخدام بيانات مثال رقم (20) ، حدد الوسيط بيانيا .

لتحديد الوسيط بيانيا ، يتم حساب التكرار المتجمع الصاعد و التكرار المتجمع النازل ، كما هو موضح في الجدول التالي:

| المتغير | التكرار المطلق $n_i$ | التكرار المطلق التجميعي الصاعد $n_i^{\nearrow}$ | التكرار المطلق التجميعي النازل $n_i^{\searrow}$ |
|---|---|---|---|
| ] 16000-11000 ] | 13 | 13 | 50 |
| ] 21000-16000 ] | 16 | 29 | 37 |
| ] 26000-21000 ] | 12 | 41 | 27 |
| ] 31000-26000 ] | 3 | 44 | 9 |
| ] 36000-31000 ] | 4 | 48 | 6 |
| ] 41000-36000 ] | 1 | 49 | 2 |
| ] 46000-41000 ] | 1 | 50 | 1 |
| المجموع | 50 | | |




كما يمكن الاكتفاء برسم أحد المنحنيين الصاعد و النازل ، و من ثم رسم خط أفقي من موقع الوسيط على المحور العمودي و إنزال خط عمودي من نقطة الالتقاء بالمنحنى إلى المحور الأفقي ليحدد قيمة الوسيط ،كما هو موضح في الشكل التالي:



**III. المنوال** : هو أبسط مؤشرات مقاييس النزعة المركزية ، و يعرف أنه القيمة الأكثر تكرارا (شيوعا) في التوزيع التكراري، و قد يكون للمشاهدات أكثر من منوال .

من مميزات المنوال (بري، هندي، و راضي، أساسيات طرق التحليل الاحصائي ، 1998، صفحة 39):

- ❖ سهل الحساب ولا يتأثر بالقيم المتطرفة .
- ❖ يمكن حسابه في حالة الجداول التكرارية المفتوحة من الطرفين
- ❖ يمكن حسابه في حالة البيانات الوصفية

و من عيوب المنوال (بري، هندي، و راضي، أساسيات طرق التحليل الاحصائي ، 1998، صفحة 39):

- ❖ مقياس غير دقيق لأنه لا يأخذ جميع القيم أثناء حسابه.
- ❖ قد يكون للبيانات أكثر من منوال
- ❖ قد لا يتواجد المنوال في مركز البيانات بل قد يكون في أحد الأطراف مما يمكنه أن لا يكون أحد مقاييس النزعة المركزية حسب بعض الإحصائيين .

**1. حساب المنوال في حالة البيانات غير المبوبة** : المنوال في البيانات غير المبوبة هي القيمة الأكثر تكرارا ، مثلا إذا كانت لدينا مجموعة من المشاهدات تمثل درجات الطلاب بالشكل التالي : 11، 04، 09، 15، 16، 11 ،12،11،06،09. في هذا المثال فإن المنوال هو الدرجة 11 لأنها تكررت 3 مرات أكثر من باقي الدرجات.

**2. حساب المنوال في حالة البيانات المبوبة** : قبل حساب المنوال في البيانات البوابة ، يجد تحديد الفئة المنوالية و التي تقابل أكبر تكرار ، ثم تطبيق الصيغة التالية :

$$M_d = A + \frac{d_1}{d_1 + d_2}.L$$

حيث :

$A$ : الحد الأدنى للفئة المنوالية .

$d_1$ : الفرق ما بين تكرار الفئة المنوالية و تكرار الفئة السابقة لها.

$d_2$: الفرق ما بين تكرار الفئة المنوالية و تكرار الفئة اللاحقة لها .

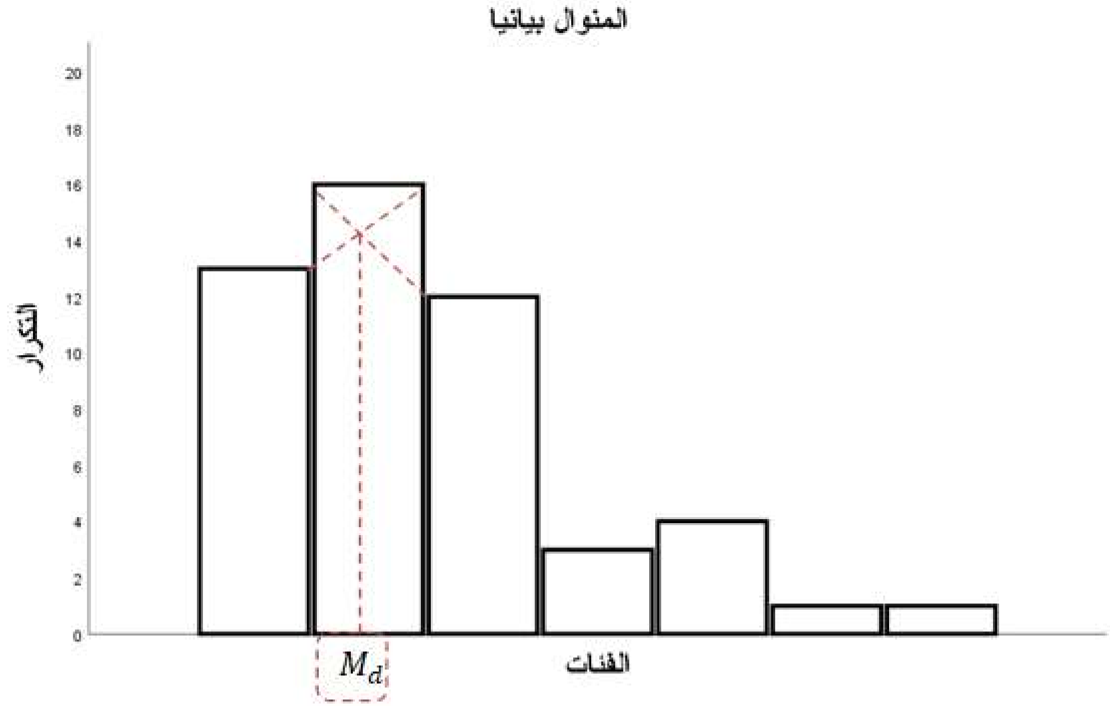
$L$ : طول الفئة .

**مثال(22)** : باستخدام بيانات مثال رقم (20) ، أوجد المنوال .

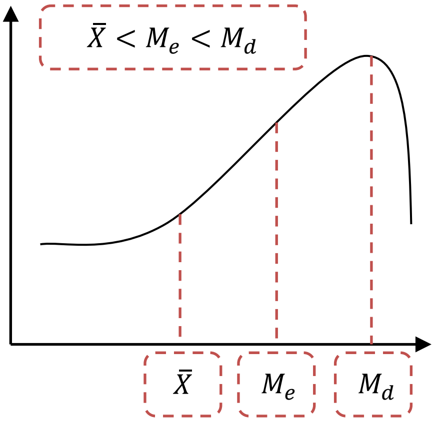
بتطبيق الصيغة الرياضية للمنوال في حالة البيانات المبوبة ، نجد :

$$M_d = 16000 + \frac{3}{3+4} \times 5000 = 18142{,}85$$

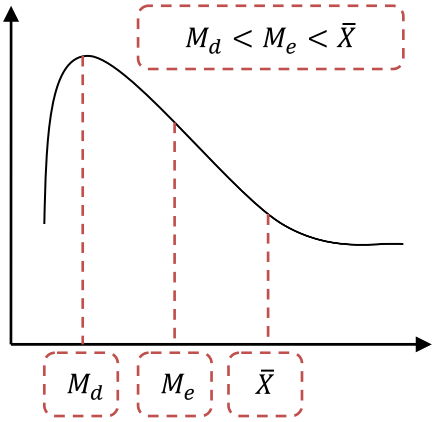
**3. تحديد المنوال بيانيا** : يمكن تحديد المنوال باستخدام المدرج التكراري ذو الفئات متساوية الطول ، و ذلك بربط زوايا أعلى عمود مع زوايا الأعمدة المجاورة له ، و من ثم إنزال خط عمودي من نقطة التقاء الخطوط على المحور الأفقي للحصول على قيمة المنوال ، و باستخدام بيانات مثال رقم (20) ، نجد المنوال بيانيا بالشكل التالي:



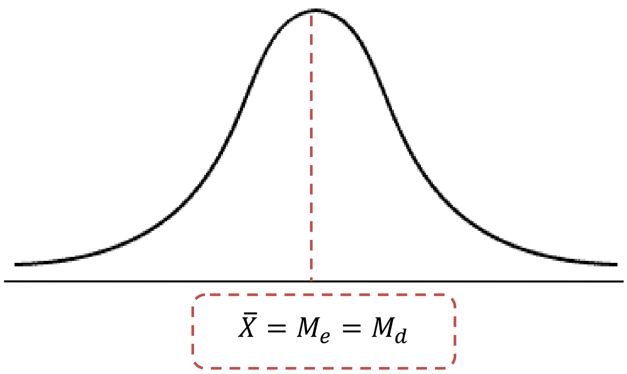

**IV. العلاقة ما بين الوسط الحسابي ، الوسيط و المنوال** : من خلال مقارنة الوسط الحسابي مع الوسيط مع المنوال ، نجد أمامنا ثلاث حالات ، أولها عندما تكون قيمهم متساوية ، في هذه الحالة نقول أن التوزيع التكراري هو توزيع متماثل أي عديم الالتواء ، لكن عندما يكون بالتوزيع التكراري التواء سواء إلى اليمين أو إلى اليسار ، فإن قيم المقاييس الثلاث تختلف عن بعضها البعض. وجود الالتواء بالمنحنى يعني أن التكرارات تتجمع في إحدى جهتي المنحنى أكثر مما تتجمع في الجهة الأخرى، على عكس المنحنى الاعتدالي (المتماثل) الذي يتساوى فيه توزيع التكرارات على جانبي المنحنى، ففي حالة الالتواء نحو اليمين تكون قيمة الوسط الحسابي أكبر المقاييس ثم قيمة الوسيط و المنوال يكون أصغر المقاييس، أما في حالة الالتواء نحو اليسار تكون قيمة الوسط الحسابي أصغر المقاييس ، ثم الوسيط ، و تكون قيمة المنوال أكبر المقاييس ، وفيما يلي أشكال توضح بشكل أفضل العلاقة ما بين المقاييس الثلاث :




الإلتواء نحو اليسار (سالب)




الإلتواء نحو اليمين (موجب)




توزيع متماثل

يلاحظ أن المتوسط الحسابي يستغل في حسابه جميع القيم ، و لذا هو أدق المتوسطات الثلاثة التي ذكرت ، كما أنه أكثر ثباتا أي أنه لا يختلف اختلافا كبيرا باختلاف العينات المختارة ، إلا أنه كثيرا ما يحدث أن تشمل المجموعة على قيم متطرفة لا تمثل المجموعة ، فالمتوسط الحسابي في هذه الحالة لابد و أن تتأثر قيمته بالحالات المتطرفة ، كما أنه في حالات الجداول التكرارية المفتوحة يتعذر حساب المتوسط الحسابي ، حيث لا يكون من الممكن معرفة مركز الفئة المفتوحة ، في مثل هذه الحالات نضطر إلى الاستعانة إما بالوسيط أو المنوال فكلاهما لا

يتأثران بالقيم المتطرفة، ذلك لأن حسابهما ينحصر في القيم و التكرارات المتوسطة، كما أنه يمكن إيجادهما إذا ما كان الجدول مفتوحا من أحد طرفيه أو كليهما. و في حالة التوزيع المتماثل نجد أن قيم هذه المتوسطات الثلاثة متطابقة، و تختلف فيما عدا ذلك، فالمتوسط الحسابي في التوزيعات الملتوية يتجه عادة نحو الطرف الملتوي، فهو يمثل مركز الثقل بالنسبة للمجموعة ، ذلك أن مجموع القيم يكون متعادلا على جانبيه، أما الوسيط فإنه يقع في منتصف مساحة التوزيع، أي أن مجموع التكرارات يكون متساويا على جانبيه، و أما المنوال فهو يمثل أعلى نقطة في منحنى التوزيع (خيثري، 1975، الصفحات 59- 60) .

**V.** **الوسط الهندسي** : بما أن الوسط الحسابي يتأثر بالقيم المتطرفة ، أي القيم الكبيرة جدا ، أو القيم الصغيرة جدا مقارنة بباقي القيم ، و لذا دعت الحاجة إلى إيجاد مقاييس تكون أقل تأثرا بالقيم الشاذة، و خاصة القيم الكبرى و من هذه المقاييس الوسط الهندسي الذي يعطي قيما أدق من الوسط الحسابي (أبو عمه، عبدالله، و هندي، 1990، صفحة 79). و الوسط الهندسي لأية مجموعة من القيم ، هو الجذر النوني لجداءات هذه القيم (رتول، 2006، صفحة 119) ، ويستخدم الوسط الهندسي في البيانات التي تشكل أو تكاد تشكل متوالية هندسية، كبيانات تطور عدد السكان، أو بيانات المبالغ المستثمرة وفق فائدة مركبة ، أما غير ذلك فهو نادر الاستخدام لصعوبة حسابه (رتول، 2006، صفحة 122). وعادة ما يستخدم الوسط الهندسي في الأرقام القياسية للأسعار أو إيجاد متوسط عدد من النسب أو في إيجاد معدلات التغير في المبيعات أو السكان، كما أنه لا يمكن إيجاد الوسط الهندسي إلا إذا كانت مجموع القيم موجبة (الراوي، 1984، صفحة 76).

**1. حساب الوسط الهندسي في حالة البيانات غير المبوبة** : إذا كانت لدينا القيم :

$x_1, x_2 \ldots, x_3$ ، فإن وسطها الهندسي هو :

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

و اختصارا يكتب على بالصيغة التالية :

$$G = \left[\prod_{i=1}^{n} x_i\right]^{\frac{1}{n}}$$

و يمكن حسابه بإدخال اللوغاريتم العشري ، و تصبح الصيغة بالشكل التالي:

$$logG = \frac{1}{n}\log(x_1 \times x_2 \times \ldots \times x_n) \Rightarrow$$

$$logG = \frac{1}{n}(logx_1 + logx_2 + \cdots + logx_n) \Rightarrow$$

$$logG = \frac{1}{n}\sum_{i=1}^{n} logx_i \Rightarrow G = 10^{\frac{1}{n}(\sum_{i=1}^{n} logx_i)}$$

حيث القيمة 10 هي أساس اللوغاريتم العشري ، و يتم استبدالها بالقيم e=2.718 في حالة استخدام اللوغاريتم النيبيري.

**مثال (23)** : أوجد الوسط الهندسي للبيانات التالية :

| | | | | | |
|---|---|---|---|---|---|
| 20 | 10 | 7 | 3 | 5 | 2 |

**الحل** :

$$G = \sqrt[6]{20 \times 10 \times 7 \times 3 \times 5 \times 2} \Rightarrow logG = \frac{1}{n}\log(20 \times 10 \times 7 \times 3 \times 5 \times 2) \Rightarrow$$

$$logG = \frac{1}{n}(log20 + log10 + log7 + log3 + log5 + log2) = 0.77 \Rightarrow$$

$$G = 10^{0.77} = 5.89$$

**2. حساب الوسط الهندسي في حالة البيانات المبوبة** : يحسب الوسط الهندسي في حالة البيانات المبوبة بالشكل التالي:

$$G = \sqrt[\sum_{i=1}^{k} n_i]{x_1^{n_1} \times x_2^{n_2} \times \ldots \times x_k^{n_k}} = \left[x_1^{n_1} \times x_2^{n_2} \times \ldots \times x_k^{n_k}\right]^{\frac{1}{\sum_{i=1}^{k} n_i}}$$

و بادخال اللوغاريتم على طرفي المعادلة نجد:

$$logG = \frac{1}{\sum n_i}\sum_{i=1}^{k} n_i logx_i \Rightarrow G = 10^{\frac{1}{\sum_{i=1}^{k} n_i}\sum_{i=1}^{k} n_i logx_i}$$

حيث القيمة 10 هي أساس اللوغاريتم العشري ، و يتم استبدالها بالقيم e=2.718 في حالة استخدام اللوغاريتم النيبيري.

**مثال (24)** : أوجد الوسط الهندسي للبيانات التالية :

| $x_i$ | 5 | 8 | 10 | 15 | $\Sigma$ |
|---|---|---|---|---|---|
| $n_i$ | 20 | 25 | 15 | 10 | 70 |

**الحل:**

| $x_i$ | $n_i$ | $logx_i$ | $n_i logx_i$ |
|---|---|---|---|
| 5 | 20 | 0.6989700043 | 13.98 |
| 8 | 25 | 0.903089987 | 22.58 |
| 10 | 15 | 1 | 15 |
| 15 | 10 | 1.1760912591 | 11.77 |
| / | 70 | / | 63.33 |

$$logG = \frac{1}{70} \times 63.33 = 0.9047 \Rightarrow G = 10^{0.9047} = 8.03$$

**ملاحظة :**

يتم استبدال القيم$x_i$ في حالة البيانات التي مدى فئاتها تفوق الصفر

**VI. الوسط التوافقي :** الوسط التوافقي لمجموعة من القيم ، هو مقلوب الوسط الحسابي لمقلوب تلك القيم .

**1. حساب الوسط التوافقي في حالة البيانات غير المبوبة** : إذا كانت لدينا القيم : $x_1,x_2...,x_3$ ، فإن وسطها التوافقي هو من الشكل التالي :

$$\overline{H} = \frac{n}{\sum_{i=1}^{k}\left[\frac{1}{x_i}\right]}$$

**مثال (25)** : أوجد الوسط التوافقي لبيانات المثال رقم(23) .

**الحل :**

| $x_i$ | 2 | 5 | 3 | 7 | 10 | 20 | Σ |
|---|---|---|---|---|---|---|---|
| $\frac{1}{x_i}$ | $\frac{1}{2}$ | $\frac{1}{5}$ | $\frac{1}{3}$ | $\frac{1}{7}$ | $\frac{1}{10}$ | $\frac{1}{20}$ | 1.326 |

$$\overline{H} = \frac{6}{1.326} = 4.52$$

**2. حساب الوسط التوافقي في حالة البيانات المبوبة** : إذا كانت لدينا القيم : $x_1, x_2 \ldots, x_3$ ، فإن وسطها التوافقي هو من الشكل التالي :

$$\overline{H} = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \left[\frac{1}{x_i}\right] n_i}$$

**مثال (26)** : أوجد الوسط التوافقي لبيانات المثال رقم (24) .

| $x_i$ | $n_i$ | $\frac{1}{x_i}$ | $\left[\frac{1}{x_i}\right] . n_i$ |
|---|---|---|---|
| 5 | 20 | $\frac{1}{5}$ | 4 |
| 8 | 25 | $\frac{1}{8}$ | 3.125 |
| 10 | 15 | $\frac{1}{10}$ | 1.5 |
| 15 | 10 | $\frac{1}{15}$ | 0.67 |
| / | 70 | / | 9.295 |

$$\overline{H} = \frac{70}{9.295} = 7.53$$

### ثانيا: المقاييس غير المركزية :

إن المقاييس المركزية تستهدف تحديد المركز الذي تتمحور حوله البيانات، و باستثناء المنوال فإن جميعها تتمثل بقيمة مفردة واحدة تكون ممثلة للبيانات محل الدراسة، أما في حالة المقاييس التي تحدد لنا مواقع غير مركزية ، فتظهر إليها الحاجة خاصة عند حساب مقاييس التشتت ( سيتم التطرق إليها في الفصل اللاحق) ، وذلك بهدف التخلص من القيم المتطرفة التي تكون في غالب الأحيان عند بداية أو نهاية القيم بعد ترتيبها تصاعديا (البلداوي، 2009، صفحة 88). و إذا رتبت عينة من البيانات حسب قيمها تصاعديا أو تنازليا فإن القراءة التي تكون في المنتصف و التي تقسم العينة إلى مجموعتين متساويتين في العدد هي الوسيط، و بتعميم الفكرة وتقسيم البيانات بعد ترتيبها إلى أربعة أجزاء متساوية فإن نقاط التقسيم يطلق عليها الربيعيات ، أما إذا قسمت البيانات إلى عشرة أجزاء ، فإن نقاط التقسيم يطلق عليها العشريات، و إذا قسمت إلى مائة فيطلق عليها المؤينات.

**I. الربيعيات Quartiles**: و يرمز لها بالرمز $Q_i$ حيث $i = 1,2,3$ ، و يشير $Q_1$ إلى الربيع الأول أو الربيع الأدنى ، و $Q_2$ يشير إلى الربيع الثاني أو الربيع الأوسط وهو نفسه الوسيط ، و $Q_3$ يشير إلى الربيع الثالث أو الربيع الأعلى.

الربيع الأدنى للبيانات هو القيمة التي يكون قبلها 25 % على الأكثر من البيانات و بعدها 75% على الأكثر من البيانات، و يحسب كما يلي:

**1. حالة البيانات غير المبوبة** : قبل معرفة قيمة الربيعيات ، يجب تحديد رتبها أولا (موقع كل ربيع) ، و فيما يلي كيفية حساب الرتبة في حالة البيانات غير المبوبة:

$$r_i = \frac{(i)(n+1)}{4}$$

- موقع الربيع الأول $Q_1$ هو $r_1 = \frac{(1)(n+1)}{4}$

- موقع الربيع الثاني $Q_2$ هو $r_2 = \frac{(2)(n+1)}{4}$
- موقع الربيع الثالث $Q_3$ هو $r_3 = \frac{(3)(n+1)}{4}$

و يتم حساب قيم الربيعات بتطبيق القاعدة التالية :

$$Q_i = x_l + (r_i - l)(x_u - x_l)$$

حيث $x_l$ تمثل القيمة التي تسبق قيمة الربيع ، $r$ رتبة الربيع ، $l$ الرتبة التي تسبق رتبة الربيع ، $u$ الرتبة التي تلي رتبة الربيع ، $x_u$ القيمة التي تلي قيمة الربيع .

**مثال (27)** : البيانات التالية مرتبة ترتيبا تصاعديا ، و المطلوب إيجاد الربيع الأول، الثاني و الثالث

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| الرتبة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| القيمة | 2 | 3 | 5 | 6 | 7 | 9 | 10 | 11 | 14 | 18 |
| الرتبة | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| القيمة | 21 | 22 | 26 | 28 | 29 | 30 | 34 | 35 | 38 | 40 |

$Q_1$ $Q_2$ $Q_3$

- حساب الربيع الأول $Q_1$:

رتبة الربيع الأول هي :

$$r_1 = \frac{(1)(20+1)}{4} = 5.25$$

قيمة الربيع الأول هي:

$$Q_1 = x_5 + (5.25 - 5)(x_6 - x_5) = 7 + 0.25(9 - 7) = 7.5$$

- حساب الربيع الثاني $Q_2$:

$$r_2 = \frac{(2)(20+1)}{4} = 10.5$$

قيمة الربيع الثاني هي :

$$Q_2 = x_{10} + (10.5 - 10)(x_{11} - x_{10}) = 18 + 0.5(21 - 18) = 19.5$$

❖ حساب الربيع الثالث $Q_3$:

$$r_3 = \frac{(3)(20+1)}{4} = 15.75$$

قيمة الربيع الثالث هي :

$$Q_3 = x_{15} + (15.75 - 15)(x_{16} - x_{15}) = 29 + (0.75)(30 - 29) = 29.75$$

2. **حالة البيانات المبوبة** : في حالة البيانات المبوبة يتم تحديد رتبة الربيع بالصيغة التالية :

$$r_i = \frac{(i)\sum_{i=1}^{k} n_i}{4}$$

و يتم حساب قيم الربيعيات بتطبيق القاعدة التالية :

$$Q_i = A + \left(\frac{r_i - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L$$

حيث $A$ تمثل الحد الأدنى لفئة الربيع ، $r$ رتبة الربيع ، $n_1^{\nearrow}$ قيمة التكرار المتجمع الصاعد الذي يسبق رتبة الربيع ، $n_2^{\nearrow}$ قيمة التكرار المتجمع الصاعد الذي يلي رتبة الربيع ، $L$ طول الفئة.

**مثال(28)** : الجدول الإحصائي التالي يبين توزيع عمال مؤسسة ما حسب أجورهم الشهرية ( الوحدة : $10^3$ دج) :

| الأجر | [10-20[ | [20-30[ | [30-40[ | [40-50[ | [50-60[ | [60-70[ |
|---|---|---|---|---|---|---|
| عدد العمال | 18 | 30 | 25 | 17 | 12 | 8 |

**المطلوب:** حساب الربيع الأول ، الربيع الثاني و الربيع الثالث.

**الحل:**

- حساب الربيع الأول $Q_1$:

رتبة الربيع الأول هي :

$$r_1 = \frac{(1)(110)}{4} = 27.5$$

حساب التكرار المتجمع الصاعد :

| الفئات | التكرار | التكرار المتجمع الصاعد |
|---|---|---|
| ]20-10] | 18 | 18 |
| ]30-20] | 30 | 48 |
| ]40-30] | 25 | 73 |
| ]50-40] | 17 | 90 |
| ]60-50] | 12 | 102 |
| ]70-60] | 8 | 110 |
| المجموع | 110 | / |

قيمة الربيع الأول هي:

$$Q_1 = 20 + \left(\frac{27.5 - 18}{48 - 18}\right).10 = 23.17 \times 10^3$$

- حساب الربيع الثاني $Q_2$:

- رتبة الربيع الثاني هي :

$$r_2 = \frac{(2)(110)}{4} = 55$$

قيمة الربيع الثاني هي:

$$Q_1 = 30 + \left(\frac{55-48}{73-48}\right).10 = 32.8 \times 10^3$$

❖ حساب الربيع الثالث $Q_2$:

❖ رتبة الربيع الثالث هي :

$$r_3 = \frac{(3)(110)}{4} = 82.5$$

قيمة الربيع الثالث هي:

$$Q_1 = 40 + \left(\frac{82.5-73}{90-73}\right).10 = 45.59 \times 10^3$$

## II. العشريات Deciles

: و يرمز لها بالرمز $D_i$ حيث $i = 1,2,\dots,9$ ، و يشير $D_1$ إلى العشير الأول ، و $D_5$ يشير إلى العشير الخامس وهو نفسه الوسيط ، و $D_9$ يشير إلى العشير التاسع . والعشير الأول هو القيمة التي يكون قبلها عشر البيانات و بعدها تسعة أعشارها، أما العشير التاسع فهو القيمة التي يكون قبلها تسعة أعشار البيانات و بعدها عشر البيانات. تحسب العشريات كما يلي:

1. **حالة البيانات غير المبوبة** : يتم تحديد رتبة العشير قبل حساب قيمة العشير ، بالشكل التالي:

$$r_i = \frac{(i)(n+1)}{10}$$

و يتم حساب قيم العشريات بتطبيق القاعدة التالية :

$$D_i = x_l + (r_i - l)(x_u - x_l)$$

حيث $x_l$ تمثل القيمة التي تسبق قيمة العشير ، $r$ رتبة العشير ، $l$ الرتبة التي تسبق رتبة العشير ، $u$ الرتبة التي تلي رتبة العشير ، $x_u$ القيمة التي تلي قيمة العشير .

**مثال (29)** : باستخدام بيانات مثال (27) ، أوجد قيمة العشير الأول، الخامس و التاسع

**الحل:**

| الرتبة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| القيمة | 2 | 3 | 5 | 6 | 7 | 9 | 10 | 11 | 14 | 18 |
| الرتبة | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| القيمة | 21 | 22 | 26 | 28 | 29 | 30 | 34 | 35 | 38 | 40 |

❖ حساب العشير الأول $D_1$:

رتبة العشير الأول $D_1$هي:

$$r_1 = \frac{(1)(20+1)}{10} = 2.1$$

قيمة العشير $D_1$ هي:

$$D_1 = x_2 + (2.1-2)(x_3-x_2) = 3 + (0.1)(5-3) = 3.2$$

❖ حساب العشير الخامس $D_5$:

رتبة العشير الخامس $D_5$هي:

$$r_5 = \frac{(5)(20+1)}{10} = 10.5$$

قيمة العشير الخامس $D_5$ هي:

$$D_5 = x_{10} + (10.5 - 10)(x_{11} - x_{10}) = 18 + (0.5)(21 - 18) = 19.5$$

❖ حساب العشير التاسع $D_9$:

رتبة العشير التاسع $D_9$هي:

$$r_9 = \frac{(9)(20+1)}{10} = 18.9$$

قيمة العشير التاسع $D_9$ هي:

$$D_9 = x_{18} + (18.9 - 18)(x_{19} - x_{18}) = 35 + (0.9)(38 - 35) = 37.7$$

**2. حالة البيانات المبوبة** : في حالة البيانات المبوبة يتم تحديد رتبة العشير بالصيغة التالية :

$$r_i = \frac{(i)\sum_{i=1}^{k} n_i}{10}$$

و يتم حساب قيم العشريات بتطبيق القاعدة التالية :

$$D_i = A + \left(\frac{r_i - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L$$

حيث $A$ تمثل الحد الأدنى لفئة العشير ، $r$ رتبة العشير ، $n_1^{\nearrow}$ قيمة التكرار المتجمع الصاعد الذي يسبق رتبة العشير ، $n_2^{\nearrow}$ قيمة التكرار المتجمع الصاعد الذي يلي رتبة العشير ، $L$ طول الفئة.

**مثال (30)** : باستخدام بيانات مثال (28) ، أوجد قيمة العشير الأول، الخامس و التاسع

**الحل:**

❖ حساب العشير الأول $D_1$:

رتبة العشير الأول هي :

$$r_1 = \frac{(1)(110)}{10} = 11$$

من جدول التكرار المتجمع الصاعد :

| الفئات | التكرار | التكرار المتجمع الصاعد |
|---|---|---|
| ]20-10] | 18 | 18 |
| ]30-20] | 30 | 48 |
| ]40-30] | 25 | 73 |
| ]50-40] | 17 | 90 |
| ]60-50] | 12 | 102 |
| ]70-60] | 8 | 110 |
| المجموع | 110 | / |

قيمة العشير الأول هي:

$$Q_1 = 10 + \left(\frac{11-0}{18-0}\right).10 = 16.11 \times 10^3$$

❖ حساب العشير الخامس $D_5$:

❖ رتبة العشير الخامس $D_5$ هي :

$$r_5 = \frac{(5)(110)}{10} = 55$$

قيمة العشير الخامس $D_5$هي:

$$D_5 = 30 + \left(\frac{55-48}{73-48}\right).10 = 32.8 \times 10^3$$

❖ حساب العشير التاسع $D_9$:

❖ رتبة العشير التاسع $D_9$هي :

$$r_9 = \frac{(9)(110)}{10} = 99$$

قيمة العشير التاسع $D_9$ هي:

$$D_9 = 50 + \left(\frac{99-90}{102-90}\right).10 = 57.5 \times 10^3$$

**III. المؤينات:** مؤين مجموعة من البيانات هو القيمة التي تفصل بين أقسام هذه المجموعة بعد تجزئتها إلى مائة جزء متساو، وعلى هذا فالمؤين الأول هو القيمة التي يكون قبلها $\frac{1}{100}$ من البيانات و بعدها $\frac{99}{100}$ من تلك البيانات، والمؤين العاشر هو القيمة التي يكون قبلها $\frac{10}{100}$ من البيانات و $\frac{90}{100}$ من تلك البيانات ، و يحسب المؤين بالشكل التالي:

**3. 1. حالة البيانات غير المبوبة** : يتم تحديد رتبة المؤين قبل حساب قيمته بالشكل التالي:

$$r_i = \frac{(i)(n+1)}{100}$$

و يتم حساب قيم المؤينات بتطبيق القاعدة التالية :

$$P_i = x_l + (r_i - l)(x_u - x_l)$$

حيث $x_l$ تمثل القيمة التي تسبق قيمة المؤين ، $r$ رتبة المؤين ، $l$ الرتبة التي تسبق رتبة المؤين ، $u$ الرتبة التي تلي رتبة المؤين ، $x_u$ القيمة التي تلي قيمة المؤين .

**مثال (31)** : باستخدام بيانات مثال (27) ، أوجد قيمة المؤين العاشر، الخمسون و التسعون

**الحل:**

| الرتبة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| القيمة | 2 | 3 | 5 | 6 | 7 | 9 | 10 | 11 | 14 | 18 |
| الرتبة | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| القيمة | 21 | 22 | 26 | 28 | 29 | 30 | 34 | 35 | 38 | 40 |

❖ حساب المؤين العاشر $P_{10}$:

رتبة المؤين العاشر $P_{10}$هي:

$$r_{10} = \frac{(10)(20+1)}{100} = 2.1$$

قيمة المؤين العاشر $P_{10}$ هي:

$$P_{10} = x_2 + (2.1-2)(x_3 - x_2) = 3 + (0.1)(5-3) = 3.2$$

❖ حساب المؤين الخمسون $P_{50}$:

رتبة المؤين الخمسون $P_{50}$هي:

$$r_{50} = \frac{(50)(20+1)}{100} = 10.5$$

قيمة المؤين الخمسون $P_{50}$ هي:

$$P_{50} = x_{10} + (10.5 - 10)(x_{11} - x_{10}) = 18 + (0.5)(21 - 18) = 19.5$$

❖ حساب المؤين التسعون $P_{90}$:

رتبة المؤين التسعون $P_{90}$هي:

$$r_{90} = \frac{(90)(20+1)}{100} = 18.9$$

قيمة المؤين التسعون $P_{90}$ هي:

$$P_{90} = x_{18} + (18.9 - 18)(x_{19} - x_{18}) = 35 + (0.9)(38 - 35) = 37.7$$

**2. حالة البيانات المبوبة** : في حالة البيانات المبوبة يتم تحديد رتبة المؤين بالصيغة التالية :

$$r_i = \frac{(i)\sum_{i=1}^{k} n_i}{100}$$

و يتم حساب قيم المؤينات بتطبيق القاعدة التالية :

$$P_i = A + \left(\frac{r_i - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L$$

حيث $A$ تمثل الحد الأدنى لفئة المؤين ، $r$ رتبة المؤين ، $n_1^{\nearrow}$ قيمة التكرار المتجمع الصاعد الذي يسبق رتبة المؤين ، $n_2^{\nearrow}$ قيمة التكرار المتجمع الصاعد الذي يلي رتبة المؤين ، $L$ طول الفئة.

**مثال (32)** : باستخدام بيانات مثال (28) ، أوجد قيمة المؤين العشرون ، الخمسون و السبعون

**الحل:**

❖ حساب المؤين العشرون $P_{20}$:

رتبة المؤين العشرون $P_{20}$ هي :

$$r_{20} = \frac{(20)(110)}{100} = 22$$

من جدول التكرار المتجمع الصاعد :

| الفئات | التكرار | التكرار المتجمع الصاعد |
|---|---|---|
| ]20-10] | 18 | 18 |
| ]30-20] | 30 | 48 |
| ]40-30] | 25 | 73 |
| ]50-40] | 17 | 90 |
| ]60-50] | 12 | 102 |
| ]70-60] | 8 | 110 |
| المجموع | 110 | / |

قيمة المؤين العشرون $P_{20}$هي:

$$Q_1 = 20 + \left(\frac{22-18}{48-18}\right).10 = 21.33 \times 10^3$$

❖ حساب المؤين الخمسون $P_{50}$:

❖ رتبة المؤين الخمسون $P_{50}$هي :

$$r_{50} = \frac{(50)(110)}{100} = 55$$

قيمة المؤين الخمسون $P_{50}$ هي:

$$P_{50} = 30 + \left(\frac{55-48}{73-48}\right).10 = 32.8 \times 10^{3}$$

- ❖ حساب المؤين السبعون $P_{70}$:
- ❖ رتبة المؤين السبعون $P_{70}$هي :

$$r_{70} = \frac{(70)(110)}{100} = 77$$

قيمة المؤين السبعون $P_{70}$ هي:

$$P_{70} = 40 + \left(\frac{77-73}{90-73}\right).10 = 42.35 \times 10^{3}$$

# الفصل الرابع: مقاييس التشتت

**تمهيد :**

**تعريف التشتت [Dispersion]**: يقصد بالتشتت للبيانات ، دراسة مدى التباعد أو التقارب لهذه البيانات عن وسطها الحسابي ، فكلما كانت البيانات قريبة من وسطها الحسابي كانت هذه البيانات غير مشتتة ، وكلما كانت هذه البيانات بعيدة عن وسطها الحسابي ، كانت هذه البيانات مشتتة (السقاف، 2020، صفحة 52). وعرف (رتول، 2006، صفحة 138) التشتت هو مدى تباعد مجموعة القيم عن بعضها البعض أو عن القيمة التي تمثل مركز تلك المجموعة. و لتوضيح مفهوم التشتت ، نأخذ المثال التالي.

**مثال(33)** : كانت نتائج عينة تتكون من خمس طلبة في مادة الإحصاء لقسمان دراسيان كالتالي:

| | | | | | |
|---|---|---|---|---|---|
| القسم الأول | 10 | 12 | 13 | 10 | 10 |
| القسم الثاني | 3 | 20 | 20 | 5 | 7 |

لو أخذنا الوسط الحسابي لنقاط المجموعتين لوجدناهما متساويين ، حيث :

$$\bar{x}_1 = \bar{x}_2 = \frac{\sum_{i=1}^{5} x_i}{n} = \frac{55}{5} = 11$$

ويعني ذلك ، أن حسب مقياس الوسط الحسابي، فإن المستوى البيداغوجي للقسمين متساو ، غير أنه لو لاحظنا نقاط القسم الأول ، لوجدنا أن طلبة القسم الأول كلهم ناجحون، بينما لم ينجح من القسم الثاني سوى طالبين، لذلك مستوى القسمين غير متساو. ولو لاحظنا الفرق ما بين أكبر علامة و أصغر علامة في القسم الأول لوجدناها : 13−10=3 ، بينما الفرق ما بين أكبر علامة و أصغر علامة في القسم الثاني لوجدناها : 20−3=17، أما الفرق ما بين أكبر علامة و متوسط علامات القسم الأول تساوي 13−11=2، بينما الفرق ما بين أكبر علامة و متوسط علامات القسم الثاني تساوي 20−11=9 . و هذا يشير إلى أن القسم الثاني أكثر

تشتتا من القسم الأول ، و الخلاصة أن مقاييس النزعة المركزية وحدها غير كافية لإعطاء صورة دقيقة عن البيانات، و ذلك لأن الكثير منها يتأثر بالقيم المتطرفة، لذلك لابد من دراسة تشتت البيانات لاستخلاص نتائج أكثر دقة (رتول، 2006، صفحة 138).

**I. المدى [Rang]**: وهو أبسط مقاييس التشتت، و يحسب المدى في حالة البيانات غير المبوبة بالشكل التالي:

$$R = V_{max} - V_{min}$$

حيث $V_{max}$ تمثل أكبر قيمة في البيانات ، أما $V_{min}$ فتمثل أقل قيمة بالبيانات .

**مثال (34)**: أوجد مدى البيانات التالية :

| | | | | |
|---|---|---|---|---|
| 05 | 12 | 10 | 20 | 15 |

**الحل:**

$$R = V_{max} - V_{min} = 20 - 5 = 15$$

أما المدى في حالة البيانات المبوبة ، فيحسب بالشكل التالي:

$$R = V_{max} - V_{min}$$

حيث $V_{max}$ تمثل مركز الفئة الأخيرة ، أما $V_{min}$ فتمثل مركز الفئة الأولى .

**مثال (35)**: أوجد مدى البيانات التالية :

| الفئات | 20-15 | 25-20 | 30-25 | 35-30 | 40-35 | 45-40 |
|---|---|---|---|---|---|---|
| التكرارات | 3 | 9 | 15 | 18 | 12 | 3 |

**الحل** :

المدى= مركز الفئة الأخيرة – مركز الفئة الأولى

مركز الفئة الأخيرة هي (40+45)/2 = 42.5 ، مركز الفئة الأولى : (15+20)/2 = 17.5 ، إذا :

$$R = 42.5 - 17.5 = 25$$

من مميزات المدى أنه بسيط و سهل الحساب، يستخدم كثيرا في مجال المناخ ( درجة الحرارة، الرطوبة، الضغط الجوي)، و من عيوبه أنه يعتمد على قيمتين فقط ولا يأخذ في اعتباره بقية القيم ، لذلك يتأثر بالقيم الشاذة (خليل، دون ذكر السنة، الصفحات 53-54) .

**ملاحظة** : يمكن حساب المدى في حالة البيانات المبوبة بالشكل التالي:

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

**II. الإنحراف الربيعي [Quartile Deviation]** : يعرف الانحراف الربيعي بأنه نصف المدى ما بين الربيعين الأول و الثالث ، و نلجأ إليه للتخلص من تأثير القيم الشاذة ( المتطرفة) أو في حالة التوزيعات التكرارية المفتوحة، و تكتب صيغته الرياضية بالشكل التالي:

$$Q_d = \frac{Q_3 - Q_1}{2}$$

**III. الإنحراف المتوسط [Average Deviation]**: هو أحد مقاييس التشتت ، و يعتمد على معدل الانحرافات المطلقة لكافة القيم $x_i$ عن و سطها الحسابي $\bar{x}$ ، إلا أن هذا المقياس نادر الاستخدام بسبب كونه يعتمد على القيم المطلقة و يهمل الإشارة ، و يحسب بالشكل التالي:

**1. في حالة البيانات غير المبوبة** : و تعطى صيغته الرياضية بالشكل التالي :

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

**مثال (36)** : البيانات التالية خاصة بالأجور الأسبوعية لعمال مؤسسة ما ، الوحدة : ألف دينار ، والمطلوب إيجاد الإنحراف المتوسط:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 70 | 80 | 50 | 60 | 50 | 70 | 50 | 70 | 90 |
| 70 | 100 | 50 | 80 | 60 | 70 | 50 | 60 | 60 | 80 |

**الحل**: لإيجاد الانحراف المتوسط ، يتطلب أولا إيجاد الوسط الحسابي ، و قيمته تساوي:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i}{n} = \frac{1320}{20} = 66.10^3$$

و الجدول الموالي يوضح الفروقات ( الانحرافات) عن الوسط الحسابي :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| -16 | 4 | 14 | -16 | -6 | -16 | 4 | -16 | 4 | 24 |
| 4 | 34 | -16 | 14 | -6 | 4 | -16 | -6 | -6 | 14 |

ومعلوم أن مجموع هذه القيم معدوم، حسب خواص الوسط الحسابي ، لذلك يتم أخذ القيم المطلقة لإيجاد الإنحراف المتوسط ، و كما يلي :

$$MD = \frac{\sum_{i=1}^{k}|x_i - \bar{x}|}{n} = \frac{240}{20} = 12.10^3$$

**2. في حالة البيانات المبوبة** : و تعطى صيغته الرياضية بالشكل التالي :

$$MD = \frac{\sum_{i=1}^{k} n_i |x_i - \bar{x}|}{\sum_{i=1}^{k} n_i}$$

**مثال (37)** : البيانات التالية تبين أجور لـ60 عامل بالدينار الجزائري في إحدى المؤسسات الاقتصادية الوحدة ( $10^3$) ، و المطلوب حساب الانحراف المتوسط :

| الأجور | عدد العمال |
|---|---|
| ] 15-10 ] | 4 |
| ] 20-15 ] | 12 |
| ] 25-20 ] | 10 |
| ] 30-25 ] | 16 |
| ] 35-30 ] | 8 |
| ] 40-35 ] | 4 |
| ] 45-40 ] | 6 |
| المجموع | 60 |

**الحل**: لحساب الانحراف المتوسط ، يجب حساب المتوسط الحسابي أولا ، ثم انحرافات القيم عن المتوسط الحسابي ، ثم ضرب كل انحراف بالتكرار المقابل له ، و حاصل مجموع القيم تقسم على مجموع التكرارات ، و الجدول الموالي يبين مختلف الحسابات:

| $n_i\lvert x_i-\bar{X}\rvert$ | $\lvert x_i-\bar{X}\rvert$ | $n_i x_i$ | $x_i$ | $n_i$ | الأجور |
|---|---|---|---|---|---|
| 56 | 14 | 50 | 12,5 | 4 | ] 15-10 ] |
| 108 | 9 | 210 | 17,5 | 12 | ] 20-15 ] |
| 40 | 4 | 225 | 22,5 | 10 | ] 25-20 ] |
| 16 | 1 | 440 | 27,5 | 16 | ] 30-25 ] |
| 48 | 6 | 260 | 32,5 | 8 | ] 35-30 ] |
| 44 | 11 | 150 | 37,5 | 4 | ] 40-35 ] |
| 96 | 16 | 255 | 42,5 | 6 | ] 45-40 ] |
| 408 | | 1590 | | 60 | المجموع |

حساب الوسط الحسابي:

$$\bar{X}=\frac{\sum x_i n_i}{N}=\frac{1590}{60}=26,5$$

حساب الانحراف المتوسط:

$$MD=\frac{\sum n_i\lvert x_i-\bar{X}\rvert}{N}=\frac{408}{60}=6,8$$

إن الخاصية الإيجابية للانحراف المتوسط هي أنه يأخذ في الحسبان جميع القيم ، لذلك فدرجة تأثره بالقيم الشاذة ضعيفة، على عكس المدى ، كما أنه لا يخضع للعمليات الجبرية ، إذ أن مجموع انحرافات القيم عن وسطها الحسابي يكون معدوما ، لذلك تم استخدام القيمة المطلقة (رتول، 2006، صفحة 141).

**IV. التباين [Variance]**: لكي يمكننا التغلب على مشكلة الإشارات عند جمع الانحرافات والتي تؤدي دائما لأن يكون مجموع انحرافات أي عينة عن وسطها الحسابي يساوي صفرا ، وبدلا من أخذ القيم المطلقة للانحرافات ، يمكننا تربيع قيم الانحرافات و بذلك تصبح جميعها موجبة، أي نحصل على مجموع مربعات الانحرافات والتي يرمز لها بالرمز SS وعلى ذلك فإن مجموع مربع الإنحرافات في حالة البيانات غير المبوبة هي من الشكل التالي :

$$SS = \sum_{i=1}^{k} (x_i - \bar{x})^2$$

أما مجموع مربع الانحرافات في حالة البيانات المبوبة فهي من الشكل التالي:

$$SS = \sum_{i=1}^{k} n_i (x_i - \bar{x})^2$$

ولكي نأخذ في الاعتبار حجم العينة حتى يمكن مقارنة العينات المختلفة الأحجام فإننا نقسم مجموع مربعات الانحرافات على درجات الحرية (n-1) ، و بذلك نتحصل على التباين $S^2$

**1. التباين في حالة البيانات غير المبوبة** : و تعطى صيغته الرياضية بالشكل التالي :

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \Rightarrow S^2 = \frac{1}{n-1} \left[ \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} \right]$$

**مثال (38)** : احسب تباين البيانات التالية :

| | | | | |
|---|---|---|---|---|
| 10 | 12 | 13 | 10 | 10 |

**الحل**: 1.حساب الوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i}{n} = \frac{55}{5} = 11$$

2. حساب مجموع مربع الانحرافات :

$$SS = \sum_{i=1}^{k}(x_i - \bar{x})^2$$

فيما يلي جدول يوضح انحراف كل قيمة عن وسطها الحسابي مع مربع هذه الانحرافات :

| | | | | | |
|---|---|---|---|---|---|
| $x_i - \bar{x}$ | −1 | −1 | 2 | 1 | −1 |
| $(x_i - \bar{x})^2$ | 1 | 1 | 4 | 1 | 1 |

إذا مجموع مربع الانحرافات هو :

$$SS = 8$$

3. حساب التباين :

$$S^2 = \frac{\sum_{i=1}^{k}(x_i - \bar{x})^2}{n-1} = \frac{8}{4} = 2$$

**عن طريق تطبيق القانون الثاني نجد** :

$$S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}\right]$$

كل الحسابات اللازمة ، موضحة في الجدول التالي :

| | | | | | |
|---|---|---|---|---|---|
| $x_i$ | 10 | 10 | 13 | 12 | 10 |
| $x_i^2$ | 100 | 100 | 169 | 144 | 100 |

$$S^2 = \frac{1}{5-1}\left[613 - \frac{(55)^2}{5}\right] = \frac{8}{4} = 2$$

إثبات أن : $S^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \Rightarrow S^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2-\frac{(\sum_{i=1}^{n}x_i)^2}{n}\right]$

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

$$\Rightarrow S^2=\frac{1}{n-1}\sum(x_i^2-2x_i\bar{x}+\bar{x}^2) \Rightarrow S^2=\frac{1}{n-1}(\sum x_i^2-2\bar{x}\sum x_i+\sum \bar{x}^2)$$

$$\Rightarrow S^2=\frac{1}{n-1}(\sum x_i^2-2\bar{x}\sum x_i+n\bar{x}^2) \Rightarrow S^2 =\frac{1}{n-1}(\sum x_i^2-2\bar{x}\sum x_i+n\frac{\sum x_i}{n}\bar{x}$$

$$\Rightarrow S^2=\frac{1}{n-1}(\sum x_i^2-2\bar{x}\sum x_i+\bar{x}\sum x_i) \Rightarrow S^2=\frac{1}{n-1}(\sum x_i^2-\bar{x}\sum x_i)$$

$$\Rightarrow S^2=\frac{1}{n-1}(\sum x_i^2-\frac{\sum x_i}{n}\sum x_i) \Rightarrow S^2=\frac{1}{n-1}(\sum x_i^2-\frac{(\sum x_i)^2}{n})$$

**2. التباين في حالة البيانات المبوبة** : و تعطى صيغته الرياضية بالشكل التالي :

$$S^2=\frac{\sum_{i=1}^{k}n_i(x_i-\bar{x})^2}{\sum_{i=1}^{k}n_i-1} \Rightarrow S^2=\frac{1}{\sum_{i=1}^{k}n_i-1}\left[\sum_{i=1}^{k}n_ix_i^2-\frac{(\sum_{i=1}^{k}n_ix_i)^2}{\sum_{i=1}^{k}n_i}\right]$$

**مثال (39)** : باستخدام بيانات مثل رقم (37) ، احسب التباين :

**الحل** : الجدول التالي يبين الحسابات اللازمة :

| الأجور | $n_i$ | $x_i$ | $n_ix_i$ | $(x_i-\bar{X})^2$ | $n_i(x_i-\bar{X})^2$ |
|---|---|---|---|---|---|
| ] 15-10 ] | 4 | 12,5 | 50 | 196 | 784 |
| ] 20-15 ] | 12 | 17,5 | 210 | 81 | 972 |
| ] 25-20 ] | 10 | 22,5 | 225 | 16 | 160 |
| ] 30-25 ] | 16 | 27,5 | 440 | 1 | 16 |

| | | | | | |
|---|---|---|---|---|---|
| 288 | 36 | 260 | 32,5 | 8 | ] 35-30 ] |
| 484 | 121 | 150 | 37,5 | 4 | ] 40-35 ] |
| 1536 | 256 | 255 | 42,5 | 6 | ] 45-40 ] |
| 4240 | | 1590 | | 60 | المجموع |

حساب الوسط الحسابي:

$$\bar{X} = \frac{\sum x_i n_i}{N} = \frac{1590}{60} = 26,5$$

حساب التباين :

$$S^2 = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}{\sum_{i=1}^{k} n_i - 1} = \frac{4240}{59} = 71.86$$

و باستخدام الصيغة الرياضية الثانية للتباين و المعرفة كالتالي :

$$S^2 = \frac{1}{\sum_{i=1}^{k} n_i - 1}\left[\sum_{i=1}^{k} n_i x_i^2 - \frac{(\sum_{i=1}^{k} n_i x_i)^2}{\sum_{i=1}^{k} n_i}\right]$$

نجد التباين بعد إجراء الحسابات اللازمة و الموضحة في الشكل التالي:

| $n_i x_i^2$ | $x_i^2$ | $n_i x_i$ | $x_i$ | $n_i$ | الأجور |
|---|---|---|---|---|---|
| 625 | 156.25 | 50 | 12,5 | 4 | ] 15-10 ] |
| 3675 | 306.25 | 210 | 17,5 | 12 | ] 20-15 ] |
| 5062.5 | 506.25 | 225 | 22,5 | 10 | ] 25-20 ] |
| 12100 | 756.25 | 440 | 27,5 | 16 | ] 30-25 ] |
| 8450 | 1056.25 | 260 | 32,5 | 8 | ] 35-30 ] |
| 5625 | 1406.25 | 150 | 37,5 | 4 | ] 40-35 ] |

| | | | | | |
|---|---|---|---|---|---|
| 10837.5 | 1806.25 | 255 | 42,5 | 6 | ] 45-40 ] |
| 46375 | | 1590 | | 60 | المجموع |

و بتطبيق الصيغة الرياضية الثانية للتباين ، نجد :

$$S^2 = \frac{1}{60-1}\left[46375 - \frac{1590^2}{60}\right] = 71.86$$

إثبات أن :

$$S^2 = \frac{\sum_{i=1}^{k} n_i(x_i-\bar{x})^2}{\sum_{i=1}^{k} n_i - 1} \Rightarrow S^2 = \frac{1}{\sum_{i=1}^{k} n_i - 1}\left[\sum_{i=1}^{k} n_i x_i^2 - \frac{(\sum_{i=1}^{k} n_i x_i)^2}{\sum_{i=1}^{k} n_i}\right]$$

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

$$\Rightarrow S^2 = \frac{1}{n-1}(\sum n_i x_i^2 - 2\bar{x}\sum n_i x_i + \bar{x}^2 \sum n_i)$$

$$\Rightarrow S^2 = \frac{1}{n-1}\left(\sum n_i x_i^2 - 2\bar{x}\sum n_i x_i + \bar{x}\frac{\sum n_i x_i}{\sum n_i}\sum n_i\right)$$

$$\Rightarrow S^2 = \frac{1}{n-1}(\sum n_i x_i^2 - 2\bar{x}\sum n_i x_i + \bar{x}\sum n_i x_i)$$

$$\Rightarrow S^2 = \frac{1}{n-1}\left(\sum n_i x_i^2 - \bar{x}\sum n_i x_i\right)$$

$$\Rightarrow S^2 = \frac{1}{n-1}(\sum n_i x_i^2 - \frac{\sum n_i x_i}{\sum n_i}\sum n_i x_i)$$

$$\Rightarrow S^2 = \frac{1}{\sum n_i - 1}(\sum n_i x_i^2 - \frac{(\sum n_i x_i)^2}{\sum n_i})$$

**ملاحظة** : في حالة حساب التباين من بيانات المجتمع ، يتم قسمة مجموع مربع الإنحرافات على حجم المجتمع و يرمز له بالرمز $\sigma^2$، حيث يحسب التباين بالصيغ الرياضية التالية :

**1. في حالة البيانات غير المبوبة :**

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} \Rightarrow \sigma^2 = \frac{\sum_{i=1}^{N} X_i^2}{N} - \mu^2$$

**2. في حالة البيانات المبوبة :**

$$\sigma^2 = \frac{\sum_{i=1}^{k} N_i(C_i - \mu)^2}{\sum_{i=1}^{k} N_i} \Rightarrow \sigma^2 = \frac{\sum_{i=1}^{k} N_i C_i^2}{\sum_{i=1}^{k} N_i} - \mu^2$$

**V. الإنحراف المعياري[Standard Deviation]**: الجذر التربيعي للتباين ينتج عنه مقياس من أهم و أدق مقاييس التشتت وهو الانحراف المعياري و يرمز له بالرمز $S$ أي أن : $S = \sqrt{S^2}$. و يحسب بالشكل التالي:

**1. في حالة البيانات غير المبوبة :**

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

**مثال (40)** : باعتماد بيانات مثال (38) ، أوجد الانحراف المعياري

**الحل:**

$$S = \sqrt{2} = 1.414$$

**2. في حالة البيانات المبوبة :**

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

**مثال (41)** : باعتماد بيانات مثال (37) ، أوجد الانحراف المعياري

**الحل:**

$$S=\sqrt{71.86}=8.477$$

**VI. معامل الإختلاف[Coefficient of Variation]**: من المعلوم أن الانحراف المعياري لمجموعة من البيانات يأخذ وحدات البيانات نفسها ، فإذا كانت البيانات تمثل الأطوال مقاسة بالسنتيمتر فإن الانحراف المعياري يكون بالسنتميتر و إذا كانت البيانات تمثل الأوزان فإنها تكون مقاسة بالكيلوغرام ويكون الانحراف المعياري بالكيلوغرام ، فإذا أردنا مقارنة تجانس مجموعة من الأوزان أو تشتتها بمجموعة من الأطوال فلا يمكن استخدام الانحراف المعياري للمقارنة لأنه لا يمكن مقارنة السنتيمتر بالكيلوغرام ، لذلك دعت الحاجة إلى إيجاد مقياس لا يعتمد على الوحدات وهذا المقياس هو ما يسمى بمعامل الاختلاف أو التشتت النسبي (بري، هندي، و عبدالله، مبادئ الإحصاء و الإحتمالات، 1997، صفحة 114) . و مقياس التشتت النسبي لها أهميتها عند مقارنة تشتت مجموعتين أو أكثر تختلف في وحدات القياس ، لأنه يكون خالي من وحدات القياس و تعطى صيغته بالشكل التالي :

إذا كان $S$ و $\bar{X}$ هما الانحراف المعياري و الوسط الحسابي لمجموعة من القيم على التوالي فإن معامل الاختلاف لها و يرمز له بالرمز $CV$ يكتب بالشكل التالي :

$$CV = \frac{S}{\overline{\overline{X}}} \times 100$$

**مثال (42):** نتائج الامتحانات النهائية لدرسي الإحصاء و الرياضيات لسنة أولى جذع مشترك علوم اقتصادية كالأتي :

| | الإحصاء | الرياضيات |
|---|---|---|
| الوسط الحسابي | 15 | 10 |
| الانحراف المعياري | 6 | 5 |

ففي أي الموضوعين كان تشتت الدرجات أكثر ؟

**الحل** : بتطبيق الصيغة التالية :

$$CV = \frac{S}{\overline{\overline{X}}} \times 100$$

نجد :

1. بالنسبة للإحصاء:

$$CV = \frac{6}{15} \times 100 = 40\%$$

2. بالنسبة للرياضيات:

$$CV = \frac{5}{10} \times 100 = 50\%$$

أي أن تشتت الرياضيات أكبر من تشتت الإحصاء. لو اعتمادنا على مقياس الانحراف المعياري لمقارنة التشتت ما بين المجموعتين لوجدنا أن التشتت في مقياس الإحصاء أكبر من التشتت في مقياس الرياضيات.

وهناك مقاييس أخرى للتشتت النسبي ، و نذكر منها :

في حالة استخدام المدى الربيعي و يسمى بمعامل الاختلاف الربيعي ، فإن معامل الاختلاف يكتب بالصيغة التالية :

$$CV = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

أما في حالة استخدام الانحراف المتوسط ، فإن صيغة معامل الاختلاف تعطى بالشكل التالي :

$$CV = \frac{MD}{\bar{X}} \times 100$$

**VII.** **الدرجة القياسية [Standardized Scores]**: في كثير من الأحيان نحتاج إلى مقارنة مفردتين من مجموعتين مختلفتين ، و في هذه الحالة يجب تحويل وحدات كل مفردة إلى وحدات قياسية حتى تكون المقارنة ذات معنى و ذلك باستخدام الوسط الحسابي و الانحراف المعياري لكل مجموعة ، و يرمز لهذه الدرجة القياسية بالرمز $z_i$ ، و تعطى صيغتها الرياضية بالشكل التالي :

$$z_i = \frac{x_i - \bar{X}}{S}$$

ومن هنا يتضح بأن الدرجات القياسية خالية من الوحدات المستخدمة في القياس ، و إذا حولنا جميع قيم عينة ما إلى درجات قياسية فإن الوسط الحسابي لهذه الدرجات القياسية يساوي 0 و التباين 1.

**مثال (43)** : حصل طالب من طلبة العلوم الاقتصادية على درجة 16 في مادة الرياضيات ، و حصل على درجة 18 في مادة الإحصاء ، إذا علمت أن الوسط الحسابي لجميع الطلبة في مادة الرياضيات هو 12 بانحراف معياري 5 ، و كان الوسط الحسابي لجميع الطلبة في مادة الإحصاء هو 15 بانحراف معياري 6 ، أي المادتين كان تحصيل هذا الطالب أعلى ؟

بالاعتماد على مقياس الدرجة القياسية و التي تعطى صيغتها الرياضية بالشكل التالي:

$$z_i = \frac{x_i - \bar{X}}{S}$$

نجد الدرجة القياسية بالنسبة للمادتين و الخاصة بالطالب بالشكل التالي:

1. بالنسبة للرياضيات:

$$z_i = \frac{16 - 12}{5} = 0.8$$

2. بالنسبة للإحصاء:

$$z_i = \frac{18 - 15}{6} = 0.5$$

عند مقارنة درجة الامتحانين نجد أن تحصيل الطالب في مادة الإحصاء أعلى منه في مادة الرياضيات ، لكن بالاعتماد على الدرجة القياسية ، يتضح لنا أن تحصيل الطالب في مادة الرياضيات أعلى منه في مادة الإحصاء.

# الفصل الخامس: مقاييس الشكل

**تمهيد:**

عند تمثيل بيانات الجداول التكرارية على شكل محنى ، فإن هذا المنحنى يأخذ أشكالا مختلفة ، فقد نحصل على منحنى متماثل أو منحنى غير متماثل ، و يعتبر المنحنى متماثلا إذا كان بالإمكان إقامة عمود على المحور الأفقي ويقسم هذا العمود المنحنى إلى قسمين متساويين، حيث في حالة المنحنى ( التوزيع) المتماثل تتساوى قيم كل من الوسط الحسابي ، الوسيط و المنوال . لكن في الكثير من الأحيان تكون هناك قيم شاذة بالبيانات تجذب إليها الوسط الحسابي ، فإذا كان الوسط الحسابي يقع في يمين المنحنى فهذا يعني وجود التواء جهة اليمين ( التواء موجب)، و إذا كان الوسط الحسابي يقع في يسار المنحنى فهذا يعني وجود التواء جهة اليسار ( التواء سالب). كما يمكننا من خلال تمثيل التوزيع التكراري للبيانات معرفة ما إذا كان شكل المنحنى مدبب أو مفرطح ، و في حالة ما إذا كان التوزيع ملتوي يمينا أو يسارا، مدبب أو مفرطح فهذا يعني أن التوزيع غير متماثل ، ويمكننا تحديد شكل الالتواء و التفرطح حسابيا من خلال مقاييس الالتواء و مقاييس التفرطح. وقبل شرح مقاييس الإلتواء والتفرطح ، سيتم التطرق إلى ما يسمى بالعزوم لأنها تدخل في تقدير هذه المقاييس :

## 1. العزم الرائي حول الصفر:

❖ **للبيانات غير مبوبة** : إذا كان لدينا n من المشاهدات التابعة للمتغير x ، فإن العزم الرائي حول الصفر هو:

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

فالعزم الأول حول الصفر هو الوسط الحسابي أي:

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

أما العزم الثاني حول الصفر فهو:

$$\overline{x^2} = \frac{\sum_{i=1}^{n} x_i^2}{n}$$

❖ **للبيانات المبوبة**: إذا كان لدينا $x_1, x_2, \dots, x_k$ تمثل مراكز الفئات في جدول التوزيع التكراري مع التكرارات المقابلة لها $n_1, n_2, \dots, n_k$ على التوالي ، فإن العزم الرائي حول الصفر هو :

$$\overline{m^r} = \frac{\sum_{i=1}^{k} n_i x_i^r}{\sum_{i=1}^{k} n_i}$$

**2. العزم الرائي حول الوسط الحسابي**:

❖ **للبيانات غير المبوبة**: العزم الرائي حول الوسط الحسابي في حالة البيانات غير المبوبة هو:

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

إذا كان $r = 1$ ، فإن $m_1 = 0$

❖ **البيانات المبوبة** : العزم الرائي حول الوسط الحسابي في حالة البيانات المبوبة هو:

$$m_r = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^r}{\sum_{i=1}^{k} n_i}$$

**مثال (44)** : أوجد العزم الأول ، الثاني و الثالث حول الصفر و حول الوسط الحسابي للبيانات التالية :

$$x_i = 4,7,5,9,8,3,6$$

❖ العزم الأول حول الصفر هو:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{4+7+5+9+8+3+6}{7} = \frac{42}{7} = 6$$

❖ العزم الثاني حول الصفر هو :

$$\overline{x^2}=\frac{\sum_{i=1}^{n}x_i^2}{n}=\frac{4^2+7^2+5^2+9^2+8^2+3^2+6^2}{7}=\frac{280}{7}=40$$

❖ العزم الثالث حول الصفر:

$$\overline{x^3}=\frac{\sum_{i=1}^{n}x_i^3}{n}=\frac{4^3+7^3+5^3+9^3+8^3+3^3+6^3}{7}=\frac{2016}{7}=288$$

❖ العزم الأول حول الوسط الحسابي:

$$m_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^1}{n}=0$$

❖ العزم الثاني حول الوسط الحسابي:

$$m_2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

$$=\frac{(4-6)^2+(7-6)^2+(5-6)^2+(9-6)^2+(8-6)^2+(3-6)^2+(6-6)^2}{7}$$

$$m_2=4$$

❖ العزم الثالث حول الوسط الحسابي:

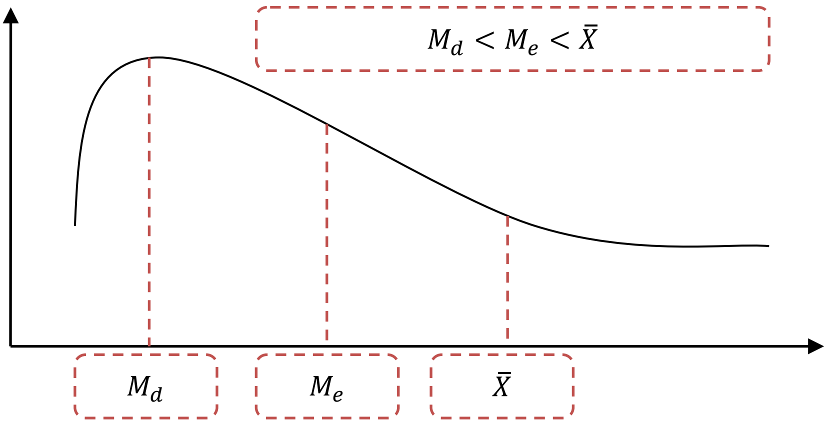
$$m_3=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{n}$$

$$=\frac{(4-6)^3+(7-6)^3+(5-6)^3+(9-6)^3+(8-6)^3+(3-6)^3+(6-6)^3}{7}$$

$$m_3=0$$

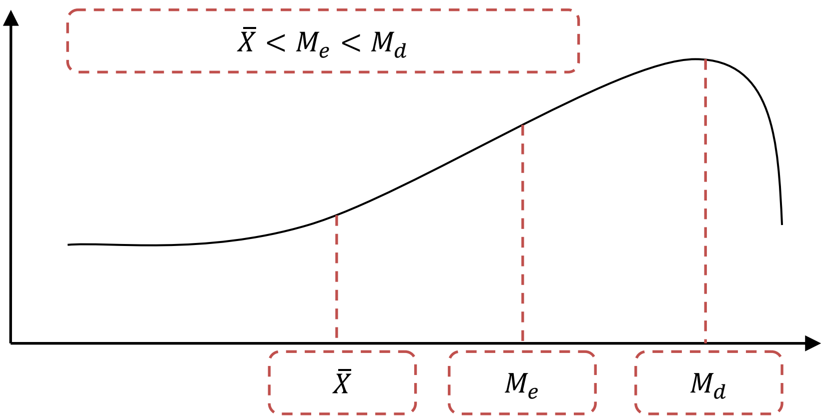
*

**I. مقاييس التماثل و الالتواء** : الإلتواء هو انحراف منحنى التوزيع التكراري عن التماثل وقد يكون الإلتواء موجب (أي الإلتواء نحو اليمين) أو سالب ( الإلتواء نحو اليسار)، ومنحى التوزيع ذو الإلتواء الموجب تكون مفرداته متمركزة في الجهة اليسرى ( عند الفئات الدنيا) و طرفه يمتد إلى اليمين كما في الشكل التالي:




الإلتواء نحو اليمين (موجب)

أما منحنى التوزيع ذو الإلتواء السالب فإن مفرداته تتمركز في الجهة اليمنى (عند الفئات العليا) بينما يمتد طرفه إلى اليسار ، كما هو موضح في الشكل التالي:




الإلتواء نحو اليسار (سالب)

و يقاس الإلتواء باستخدام المقاييس التالية :

1. **معامل بيرسون** : و يحسب بإحدى الصيغتين :

$$P = \frac{\bar{X} - M_0}{S}$$

$$P = \frac{3(\bar{X} - M_e)}{S}$$

و يتم تحديد شكل الإلتواء كما يلي :

- إذا كان : $P = 0$ فإن منحنى التوزيع يكون متماثل .
- إذا كان : $P > 0$ فإن منحنى التوزيع يكون مائل لليمين .
- إذا كان : $P < 0$ فإن منحنى التوزيع يكون مائل لليسار .

2. **معامل الإلتواء بدلالة العزوم ( معامل فيشر)**: و يحسب بالشكل التالي:

$$F = \frac{m_3}{S^3}$$

يعتمد على العزم المركزي من الدرجة الثالثة بشكل خاص في حساب معامل فيشر للالتواء لأن قيمته في حال توزيع متناظر تساوي الصفر $m_3 = 0$ و يتم تحديد شكل الالتواء كما يلي :

- إذا كان : $F = 0$ فإن منحنى التوزيع يكون متماثل .
- إذا كان : $F > 0$ فإن منحنى التوزيع يكون مائل لليمين .
- إذا كان : $F < 0$ فإن منحنى التوزيع يكون مائل لليسار .

**3. معامل يول للالتواء $C_y$**: يستخدم هذا المعامل في حالة التوزيعات التكرارية المفتوحة ، و يسمى أيضا معامل الالتواء الربيعي ، و هو معطى بالعلاقة التالية :

$$C_y = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}, C_y = \frac{(Q_3 - Q_2) + (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

و يتم تحديد شكل الالتواء كما يلي :

- ❖ إذا كان : $C_y = 0$ فإن منحنى التوزيع يكون متماثل .
- ❖ إذا كان : $C_y > 0$ فإن منحنى التوزيع يكون مائل لليمين .
- ❖ إذا كان : $C_y < 0$ فإن منحنى التوزيع يكون مائل لليسار .

**مثال (45)** : باستخدام بيانات التالية ، حدد شكل التواء البيانات :

| الأجور | $n_i$ |
|---|---|
| ] 15-10 ] | 10 |
| ] 20-15 ] | 6 |
| ] 25-20 ] | 18 |
| ] 30-25 ] | 16 |
| ] 35-30 ] | 8 |
| ] 40-35 ] | 12 |
| ] 45-40 ] | 10 |
| المجموع | 80 |

**الحل**: الجدول التالي يبين الحسابات اللازمة :

| $n_i(x_i-\bar{X})^2$ | $(x_i-\bar{X})^2$ | $n_i x_i$ | $x_i$ | $F_i^{\nearrow}$ | $n_i$ | الأجور |
|---|---|---|---|---|---|---|
| 2287.65 | 228.765 | 125 | 12,5 | 10 | 10 | ] 15-10 ] |
| 615.09375 | 102.515 | 105 | 17,5 | 16 | 6 | ] 20-15 ] |
| 472.78125 | 26.265 | 405 | 22,5 | 34 | 18 | ] 25-20 ] |
| 0.25 | 0.015 | 440 | 27,5 | 50 | 16 | ] 30-25 ] |
| 190.125 | 23.765 | 260 | 32,5 | 58 | 8 | ] 35-30 ] |
| 1170.1875 | 97.515 | 450 | 37,5 | 70 | 12 | ] 40-35 ] |
| 2212.65 | 221.265 | 425 | 42,5 | 80 | 10 | ] 45-40 ] |
| 6948.7375 | | 2210 | | | 80 | المجموع |

بعد الحسابات اللازمة ، تحصلنا على الحسابات التالية :

$$\bar{X}=27.625, M_e=26.875, M_o=24.285, S^2=87.958$$

**1. معامل بيرسون :**

$$P=\frac{\bar{X}-M_0}{S}=\frac{27.625-24.285}{\sqrt{87.958}}=0.356$$

$$P=\frac{3(\bar{X}-M_e)}{S}=\frac{3(27.625-26.875)}{\sqrt{87.958}}=0.239$$

بما أن $P>0$ فإن منحنى التوزيع يكون مائل لليمين.

**2.معامل الإلتواء بدلالة العزوم ( معامل فيشر)**: و يحسب بالشكل التالي:

$$F=\frac{m_3}{S^3}$$

$$m_3 = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^3}{\sum_{i=1}^{k} n_i}$$

| $n_i(x_i - \bar{X})^3$ | $(x_i - \bar{X})^3$ | $x_\mathrm{i}$ | $n_\mathrm{i}$ | الأجور |
|---|---|---|---|---|
| -34600.8 | -3460.08 | 12,5 | 10 | ] 15-10 ] |
| -6227.824 | -1037.97 | 17,5 | 6 | ] 20-15 ] |
| -2423.003 | -134.611 | 22,5 | 18 | ] 25-20 ] |
| -0.03125 | -0.0019 | 27,5 | 16 | ] 30-25 ] |
| 925.859 | 115.857 | 32,5 | 8 | ] 35-30 ] |
| 11555.601 | 962.966 | 37,5 | 12 | ] 40-35 ] |
| 32913.261 | 3291.326 | 42,5 | 10 | ] 45-40 ] |
| 2143.06275 | | | 80 | المجموع |

$$m_3 = \frac{2143.06275}{80} = 26.788$$

$$F = \frac{m_3}{S^3} = \frac{26.788}{\sqrt{87.958}^{\mathbf{3}}} = 0.032$$

بما أن $F > 0$ فإن منحنى التوزيع يكون مائل لليمين .

**3.معامل يول للالتواء $C_y$**:

$$C_y = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

$$Q_i = A + \left(\frac{r_i - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L$$

$$r_i = \frac{(i)\sum_{i=1}^{k} n_i}{4} \Rightarrow$$

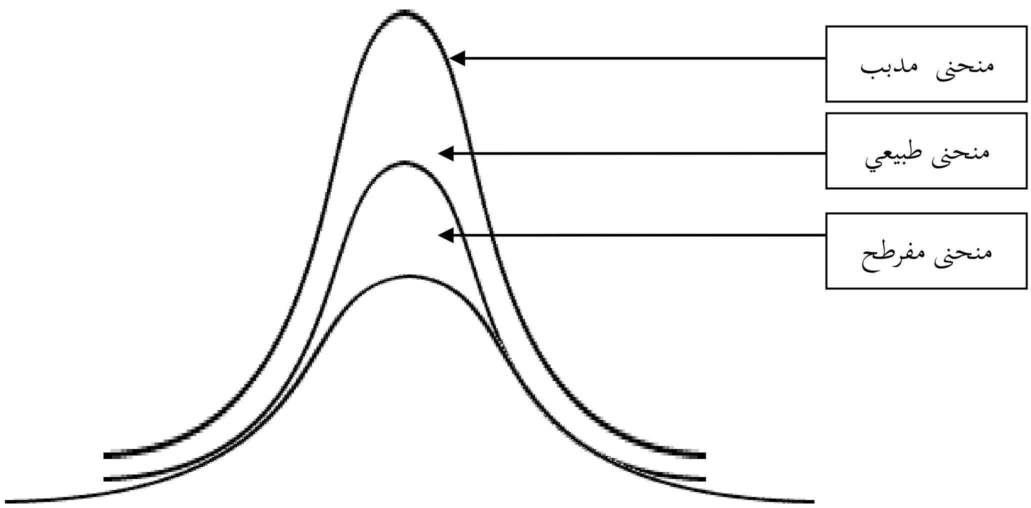
$$r_1 = \frac{(1)\sum_{i=1}^{k} n_i}{4} = \frac{80}{4} = 20 \Rightarrow Q_1 = 20 + \left(\frac{20-16}{34-16}\right).5 = 21.11$$

$$r_3 = \frac{(3)\sum_{i=1}^{k} n_i}{4} = \frac{3 \times 80}{4} = 60 \Rightarrow Q_3 = 35 + \left(\frac{60-58}{70-58}\right).5$$
$$= 35.83$$

$$C_y = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{35.83 - (2)(26.875) + 21.11}{35.83 - 21.11} = 0.216$$

بما أن $C_y > 0$ ،فإن منحنى التوزيع يكون مائل لليمين.

**II.** **مقاييس التفرطح (التفلطح)** : التفرطح يعني قياس درجة علو قمة التوزيع بالنسبة للتوزيع الطبيعي، فكلما كان الشكل أكثر ارتفاعا من الشكل الطبيعي نقول أن شكل المنحنى غير مفرطح أي أن الشكل مدبب ، أما إذا كان الشكل أقل ارتفاعا من شكل المنحنى الطبيعي ، نقول أن الشكل مفرطح ، و التمثيل البياني التالي يوضح ذلك :



و يمكن قياس التفرطح باستخدام المعاملات التالية:

**1. معامل بيرسون :**

$$\beta_2 = \frac{\mathrm{M}_4}{\mathrm{M}_2^2} = \frac{\mathrm{M}_4}{\sigma_\mathrm{x}^2}$$

حيث العزم المركزي من الدرجة الثانية $\mathrm{M}_2$ و العزم المركزي من الدرجة الرابعة $\mathrm{M}_4$، و يتم تحديد شكل التوزيع كما يلي:

- إذا كان $\beta_2 = 3$ ، فإن منحنى التوزيع طبيعي .
- إذا كان $\beta_2 > 3$ ، فإن منحنى التوزيع مدبب.
- إذا كان $\beta_2 < 3$ ، فإن منحنى التوزيع مفرطح(متفلطح).

**2.معامل فيشر** : و يعتبر من أكثر المقاييس استعمالا ، و يعتمد في حسابه على العزم المركزي من الدرجة الرابعة ، ونرمز له بالرمز F ، و صيغته من الشكل التالي:

$$F = \frac{\mathrm{M}_4}{\sigma_x^4} - 3 = \beta_2 - 3$$

و يتم تحديد شكل التوزيع كما يلي:

- إذا كان $F = 0$ ، فإن منحنى التوزيع طبيعي .
- إذا كان $F > 0$ ، فإن منحنى التوزيع مدبب.
- إذا كان $F < 0$ ، فإن منحنى التوزيع مفرطح(متفلطح).

**3.معامل كيلي Kelly** : و صيغته الرياضية هي من الشكل التالي:

$$C_k = \frac{0.5(Q_3 - Q_1)}{D_9 - D_1}$$

و يتم تحديد شكل التوزيع كما يلي:

- ❖ إذا كان $0.15 < C_k < 0.25$، فإن منحنى التوزيع طبيعي .
- ❖ إذا كان $C_k > 0.25$، فإن منحنى التوزيع مدبب.
- ❖ إذا كان $0 < C_k < 0.15$، فإن منحنى التوزيع مفرطح(متفلطح).

**مثال (46)** : باستخدام بيانات التالية، أدرس تفرطح هذا التوزيع

| الأجور | $n_{\mathrm{i}}$ |
|---|---|
| ] 15-10 ] | 10 |
| ] 20-15 ] | 12 |
| ] 25-20 ] | 10 |
| ] 30-25 ] | 16 |
| ] 35-30 ] | 6 |
| ] 40-35 ] | 5 |
| ] 45-40 ] | 1 |
| المجموع | 60 |

الحل:

**1.باستخدام معامل بيرسون:**

$$\beta_2 = \frac{\mathrm{M}_4}{\mathrm{M}_2^2} = \frac{\mathrm{M}_4}{\sigma_{\mathrm{x}}^4}$$

**الحل:**

بعد الحسابات اللازمة ، وجدنا:

$$\bar{X} = 23.75, S^2 = 62.1875$$

| $(x_i - \bar{X})^4$ | $x_i - \bar{X}$ | $x_{\mathrm{i}}$ | $n_{\mathrm{i}}$ | الأجور |
|---|---|---|---|---|
| 16018 ,0664 | -11 ,25 | 12,5 | 10 | ] 15-10 ] |
| 1525,8789 | -6 ,25 | 17,5 | 12 | ] 20-15 ] |
| 2,4414 | -1,25 | 22,5 | 10 | ] 25-20 ] |
| 197,753 | 3 ,75 | 27,5 | 16 | ] 30-25 ] |
| 5861,8164 | 8,75 | 32,5 | 6 | ] 35-30 ] |
| 35744,628 | 13,75 | 37,5 | 5 | ] 40-35 ] |
| 123596,191 | 18,75 | 42,5 | 1 | ] 45-40 ] |
| 182946,7751 | | | 60 | المجموع |

$$m_4 = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^4}{\sum_{i=1}^{k} n_i} = \frac{182946,7751}{60} = 3049,1129$$

$$\beta_2 = \frac{\mathrm{M}_4}{\sigma_{\mathrm{x}}^4} = \frac{3049,1129}{(\sqrt{62.1875})^4} = 0,7884$$

بما أن $\beta_2 = 0,7884$ أقل من 3 ، فإن منحنى التوزيع متفلطح

**2- باستخدام معامل فيشر :**

$$F = \frac{\mathrm{M}_4}{\sigma_x^4} - 3 = \beta_2 - 3 = 0,7884 - 3 = -2,2115$$

بما أن $F = -2,2115$ أقل من 0 ، فإن منحنى التوزيع متفلطح

**3- باستخدام معامل كيلي Kelly :**

$$C_k = \frac{0.5(Q_3 - Q_1)}{D_9 - D_1}$$

| الأجور | $n_{\mathrm{i}}$ | |
|---|---|---|
| ] 15-10 ] | 10 | 10 |
| ] 20-15 ] | 12 | 22 |
| ] 25-20 ] | 10 | 32 |
| ] 30-25 ] | 16 | 48 |
| ] 35-30 ] | 6 | 54 |
| ] 40-35 ] | 5 | 59 |
| ] 45-40 ] | 1 | 60 |
| المجموع | 60 | |

لدينا :

$$Q_i = A + \left(\frac{r_i - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L\ , r_i = \frac{(i)\sum_{i=1}^{k} n_i}{4}$$

$$Q_1 = A + \left(\frac{r_1 - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L\,, r_1 = \frac{(1)60}{4}$$

$$r_1 = 15 \;\Rightarrow Q_1 = 15 + \left(\frac{15-10}{22-10}\right).5 \Rightarrow Q_1 = 17{,}08$$

$$Q_3 = A + \left(\frac{r_3 - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L\,, r_3 = \frac{(3)60}{4}$$

$$r_3 = 45 \;\Rightarrow Q_3 = 25 + \left(\frac{45-32}{48-32}\right).5 \Rightarrow Q_3 = 29{,}0625$$

و لدينا :

$$D_i = A + \left(\frac{r_i - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L, r_i = \frac{(i)\sum_{i=1}^{k} n_i}{10} \Rightarrow$$

$$D_9 = A + \left(\frac{r_9 - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L, r_9 = \frac{(9)60}{10}$$

$$r_9 = \frac{(9)60}{10} = 54 \Rightarrow D_9 = 30 + \left(\frac{54-48}{54-48}\right).5 = 35$$

$$D_1 = A + \left(\frac{r_1 - n_1^{\nearrow}}{n_2^{\nearrow} - n_1^{\nearrow}}\right).L, r_1 = \frac{(1)60}{10}$$

$$r_1 = 6 \Rightarrow D_1 = 10 + \left(\frac{6-0}{10-0}\right).5 = 13$$

$$\Rightarrow C_k = \frac{0.5(29{,}0625 - 17{,}08)}{35-13} = 0{,}272$$

حسب معامل كيلي ، $C_k > 0.25$، فإن منحنى التوزيع مدبب

# الفصل السادس: تحليل الارتباط والانحدار

**1.تحليل الارتباط** :

يهدف تحليل الارتباط الخطي إلى معرفة ما إذا كانت هناك علاقة خطية ما بين متغيرين أو مجموعة من المتغيرات ، فإذا كان لدينا متغيرين فقط فيطلق عليه ارتباط بسيط ويرمز له بالرمز $r$، أما إذا كان لدينا مجموعة من المتغيرات (ثلاث فأكثر) ، فيطلق عليه ارتباط متعدد ويرمز له بالرمز $R$، كما توجد العديد من معاملات الارتباط تختلف باختلاف نوع المعطيات، فهناك متغيرات كمية وأخرى نوعية، ولكلاهما معامل خاص بها، وبهذا الفصل سنقتصر على دراسة معامل الارتباط الخاص بالمتغيرات الكمية.

وفيما يلي شرح مفصل لمعامل الارتباط البسيط، المتعدد، والجزئي (البلداوي، 2008، الصفحات 172-183):

**1-** **معامل الارتباط البسيط**: أو كما يطلق عليه معامل بيرسون ، من خواص معامل الارتباط أن قيمته محصورة ما بين (+1) و (-1) ، حيث إذا كانت قيمته أكبر من الصفر نقول أن نوع الارتباط موجب ، وتزداد قوة هذه العلاقة الموجبة كلما اقتربت من (+1)، أما إذا كانت قيمته أصغر من الصفر، نقول أن نوع الارتباط سالب، وتزداد قوة هذه العلاقة السالبة كلما اقتربت من (-1)، أما إذا كانت قيمة معامل الارتباط معدومة، فنقول أنه لا توجد علاقة أو ارتباط خطي بين المتغيرين، و كما توضح الأشكال الموالية:

وهو معرف بالعلاقة التالية:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} \dots..(01)$$

حيث :

$\bar{x}$ : الوسط الحسابي للمتغير $x_i$.

$\bar{y}$: الوسط الحسابي للمتغير $y_i$.

$S_x$ : الانحراف المعياري للمتغير $x_i$ $\quad S_x = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$.

$S_y$ : الانحراف المعياري للمتغير $y_i$ $\quad S_y = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}$ ، $n$ : حجم العينة .

**مثال (47)**: يبين الجدول الموالي تطور كل من الاستهلاك $y_i$ و الدخل $x_i$ خلال 15 سنة:

| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 500 | 400 | 350 | 450 | 300 | 350 | 420 | 510 | 360 | 270 | 320 | 350 | 430 | 520 | 550 |
| $y_i$ | 230 | 210 | 200 | 210 | 200 | 210 | 220 | 240 | 215 | 180 | 200 | 210 | 240 | 250 | 300 |

أحسب معامل الارتباط بين $x_i$ و $y_i$.

**الحل:** بتطبيق العلاقة رقم (01) نجد :

| $x_i$ | $y_i$ | $(x_i-\bar{x})$ | $(y_i-\bar{y})$ | $(x_i-\bar{x})\,(y_i-\bar{y})$ |
|---|---|---|---|---|
| 500 | 230 | 94,6666667 | 9 | 852 |
| 400 | 210 | -5,33333333 | -11 | 58,6666667 |
| 350 | 200 | -55,3333333 | -21 | 1162 |
| 450 | 210 | 44,6666667 | -11 | -491,333333 |
| 300 | 200 | -105,333333 | -21 | 2212 |
| 350 | 210 | -55,3333333 | -11 | 608,666667 |
| 420 | 220 | 14,6666667 | -1 | -14,6666667 |

| | | | | |
|---|---|---|---|---|
| 510 | 240 | 104,666667 | 19 | 1988,66667 |
| 360 | 215 | -45,3333333 | -6 | 272 |
| 270 | 180 | -135,333333 | -41 | 5548,66667 |
| 320 | 200 | -85,3333333 | -21 | 1792 |
| 350 | 210 | -55,3333333 | -11 | 608,666667 |
| 430 | 240 | 24,6666667 | 19 | 468,666667 |
| 520 | 250 | 114,666667 | 29 | 3325,33333 |
| 550 | 300 | 144,666667 | 79 | 11428,6667 |
| 86,5090141 | 28,5482048 | / | / | $\Sigma = 29820$ |

حيث :

86,5090141 هو الانحراف المعياري للمتغير $x_i$، 28,5482048 هو الانحراف المعياري للمتغير $y_i$ ، إذا :

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} \Rightarrow r = \frac{29820}{(15-1)(86,509)(28,548)} = 0,86246$$

**2- معامل الارتباط المتعدد:** قيمة هذا المعامل تنحصر ما بين (1) و (0)، أي بمعنى أن هذا المعامل لا يكشف عن نوع الاتجاه ما إذا كان موجب أم سالب، لأن هذا الاتجاه لا يكون موحدا لجميع المتغيرات فقد يكون الارتباط بين المتغير الأول والثاني موجب، لكن ما بين المتغير الأول والثالث سالب، نفترض أن لدينا ثلاث متغيرات ($y_i$ ، $x_{i1}$ ، $x_{i2}$ ) ، ففي هذه الحالة يتم حساب معامل الارتباط المتعدد بالعلاقة التالية :

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}} \dots (02)$$

حيث: $r_{yx_1}$، $r_{yx_2}$، $r_{x_1x_2}$ ، هي ثلاث معاملات ارتباط بين $x_1$ و $y$ ، وبين $x_2$ و $y$ ، وبين

$x_1$ و $x_2$ على التوالي و تحسب بالعلاقات التالية :

$$r_{yx_1} = \frac{\left(n\sum_{i=1}^{n} y_i\, x_{i1}\right) - \left(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_{i1}\right)}{\sqrt{\left[\left(n\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2\right]\left[\left(n\sum_{i=1}^{n} x_{i1}^2\right) - \left(\sum_{i=1}^{n} x_{i1}\right)^2\right]}} \ldots (03)$$

$$r_{yx_2} = \frac{\left(n\sum_{i=1}^{n} y_i\, x_{i2}\right) - \left(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_{i2}\right)}{\sqrt{\left[\left(n\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2\right]\left[\left(n\sum_{i=1}^{n} x_{i2}^2\right) - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right]}} \ldots (04)$$

$$r_{x_1x_2} = \frac{\left(n\sum_{i=1}^{n} x_{i1}x_{i2}\right) - \left(\sum_{i=1}^{n} x_{i1} \sum_{i=1}^{n} x_{i2}\right)}{\sqrt{\left[\left(n\sum_{i=1}^{n} x_{i1}^2\right) - \left(\sum_{i=1}^{n} x_{i1}\right)^2\right]\left[\left(n\sum_{i=1}^{n} x_{i2}^2\right) - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right]}} \ldots (05)$$

**ملاحظة:**إذا زاد عدد المتغيرات عن ثلاث متغيرات، يتم الاستعانة بالمصفوفات لحساب معدل الارتباط المتعدد ، وهذا سيتم التطرق له في فصل الانحدار المتعدد، بعد إلقاء نظرة عن العمليات على المصفوفات.

**مثال تطبيقي(48)**: يبين الجدول الموالي 10 مشاهدات لثلاث متغيرات ($y_i$ ، $x_{i1}$ ، $x_{i2}$ ) :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_i$ | 120 | 80 | 75 | 100 | 110 | 128 | 140 | 142 | 145 | 150 |
| $x_{i1}$ | 35 | 60 | 75 | 70 | 72 | 80 | 85 | 90 | 92 | 100 |
| $x_{i2}$ | 120 | 105 | 90 | 87 | 81 | 70 | 72 | 68 | 62 | 60 |

أحسب معامل الارتباط المتعدد

**الحل:**

| $y_i$ | $x_{i1}$ | $x_{i2}$ | $y_i^2$ | $x_{i1}^2$ | $x_{i2}^2$ | $y_i x_{i1}$ | $y_i x_{i2}$ | $x_{i1}x_{i2}$ |
|---|---|---|---|---|---|---|---|---|
| 120 | 35 | 120 | 14400 | 1225 | 14400 | 4200 | 14400 | 4200 |
| 80 | 60 | 105 | 6400 | 3600 | 11025 | 4800 | 8400 | 6300 |
| 75 | 75 | 90 | 5625 | 5625 | 8100 | 5625 | 6750 | 6750 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 100 | 70 | 87 | 10000 | 4900 | 7569 | 7000 | 8700 | 6090 |
| 110 | 72 | 81 | 12100 | 5184 | 6561 | 7920 | 8910 | 5832 |
| 128 | 80 | 70 | 16384 | 6400 | 4900 | 10240 | 8960 | 5600 |
| 140 | 85 | 72 | 19600 | 7225 | 5184 | 11900 | 10080 | 6120 |
| 142 | 90 | 68 | 20164 | 8100 | 4624 | 12780 | 9656 | 6120 |
| 145 | 92 | 62 | 21025 | 8464 | 3844 | 13340 | 8990 | 5704 |
| 150 | 100 | 60 | 22500 | 10000 | 3600 | 15000 | 9000 | 6000 |
| 1190 | 759 | 815 | 148198 | 60723 | 69807 | 92805 | 93846 | 58716 |

بالتعويض في العلاقة (03) ،(04)، و(05) نجد :

$$r_{yx_1} = \frac{(10)(92805) - (1190)(759)}{\sqrt{[(10)(148198) - (1190)^2][(10)(60723) - (759)^2 \ ]}} = 0{,}548343$$

$$r_{yx_2} = \frac{(10)(93846) - (1190)(815)}{\sqrt{[\ (10)(148198) - \ (1190)^2 \ ][(10)(69807) - (815)^2 \ \ ]}}$$
$$= -0{,}6647636$$

$$r_{x_1x_2} = \frac{(10)(58716) - (759)(815)}{\sqrt{[\ (10)(60723) - \ (759)^2 \ ][(10)(69807) - (815)^2 \ \ ]}}$$
$$= -0{,}9678453$$

وبالتعويض في العلاقة (02)، نجد:

$$R = \sqrt{\frac{(0{,}54834 \ \ )^2 + (-0{,}66476)^2 - 2(0{,}54834)(-0{,}66476)(-0{,}96784)}{1 - (-0{,}96784)^2}}$$

$$\Rightarrow R = 0{,}76464$$

**3- معامل الارتباط الجزئي**: هو مقياس لارتباط أي زوج من المتغيرات عندما تبقى باقي المتغيرات ثابتة، فمثلا عندما يكون لدينا خمس متغيرات $x_1$ ، $x_2$ ، $x_3$ ، $x_4$ ، $x_5$ ، فإيجاد الارتباط الجزئي بين المتغيرين $x_1$، و $x_3$، يتم بإبقاء باقي المتغيرات ثابتة ، وهي $x_2$ ،$x_4$ ، $x_5$ وعندها يرمز لمعامل الارتباط في مثل هذه الحالة ب $r_{13.245}$ ، ويطلق على مربع معامل الارتباط الجزئي بمعامل التحديد الجزئي، والفرق بين معامل الارتباط البسيط ومعامل الارتباط الجزئي ، هو أن الأول يقيس العلاقة بين المتغيرين ضمن تأثيرات المتغيرات الأخرى، في حين يقيس الثاني العلاقة بين متغيرين بعد مراقبة المتغير أو المتغيرات الأخرى ، و بافتراض أنه لدينا ثلاث متغيرات $x_1, x_2, y$ ، ونريد حساب معامل الارتباط الجزئي بين المتغيرين $x_2, y$ مع ثبات المتغير $x_1$ فيحسب معامل الارتباط الجزئي بالعلاقة التالية:

$$r_{yx_2.x_1} = \frac{r_{yx_2} - (r_{yx_1})(r_{x_2.x_1})}{\sqrt{(1 - r^2_{yx_1})(1 - r^2_{yx_2})}} \dots (06)$$

حيث تحسب المعاملات، $r_{yx_1}$، $r_{yx_2}$، $r_{x_2.x_1}$ بتطبيق العلاقة (03)، (04)، (05) على التوالي.

**مثال (49)**: باعتماد معطيات المثال التطبيقي رقم (48)، أحسب معاملات الارتباط الجزئية التالية:

1/ بين $y$ و $x_2$ مع مراقبة $x_1$ ، ($r_{yx_2.x_1}$ ).

2/ بين $y$ و $x_1$ مع مراقبة $x_2$ ،($r_{yx_1.x_2}$).

3/ بين $x_1$ و$x_2$ مع مراقبة$y$ ،($r_{x_1x_2.y}$).

**الحل**: لدينا :

$$r_{yx_1} = \frac{(10)(92805) - (1190)(759)}{\sqrt{[(10)(148198) - (1190)^2][(10)(60723) - (759)^2\ \ ]}} = 0{,}548343$$

$$r_{yx_2} = \frac{(10)(93846) - (1190)(815)}{\sqrt{[\ (10)(148198) -\ (1190)^2\ ][(10)(69807) - (815)^2\ \ ]}}$$
$$= -0{,}6647636$$

$$r_{x_1x_2} = \frac{(10)(58716) - (759)(815)}{\sqrt{[\ (10)(60723) -\ (759)^2\ ][(10)(69807) - (815)^2\ \ ]}}$$
$$= -0{,}9678453$$

1/حساب معامل الارتباط الجزئي بين $y$ و $x_2$ مع مراقبة $x_1$:

$$r_{yx_2.x_1} = \frac{r_{yx_2} - \left(r_{yx_1}\right)\left(r_{x_2.x_1}\right)}{\sqrt{\left(1 - r^2_{yx_1}\right)\left(1 - r^2_{x_2.x_1}\right)}}$$

$$\Rightarrow r_{yx_2.x_1} = \frac{(-0{,}66476) - (0{,}548343)(-0{,}96784)}{\sqrt{(1 - (0{,}548343)^2)(1 - (-0{,}96784)^2)}} = -0.637$$

2/ حساب معامل الارتباط الجزئي بين $y$ و $x_1$ مع مراقبة $x_2$ :

$$r_{yx_1.x_2} = \frac{r_{yx_1} - \left(r_{yx_2}\right)\left(r_{x_1.x_2}\right)}{\sqrt{\left(1 - r^2_{yx_2}\right)\left(1 - r^2_{x_1.x_2}\right)}}$$

$$\Rightarrow r_{yx_2.x_1} = \frac{(0{,}5483) - (-0{,}6647)(-0{,}96784)}{\sqrt{(1 - (-0{,}6647)^2)(1 - (-0{,}96784)^2)}} = -0.5058$$

3/حساب معامل الارتباط الجزئي بين $x_1$ و$x_2$ مع مراقبة$y$ :

$$r_{x_2.x_1.y} = \frac{r_{x_2.x_1} - (r_{x_2 y})(r_{x_1.y})}{\sqrt{(1 - r_{x_2 y}^2)(1 - r_{x_1.y}^2)}}$$

$$\Rightarrow r_{x_2.x_1.y} = \frac{(-0,9678453) - (-0,6647636)(0,548343)}{\sqrt{(1 - (-0,6647636)^2)(1 - (0,548343)^2)}} = -0.9657$$

## 2. تحليل الانحدار:

### 1- التعريف بالنموذج الخطي البسيط:

يهدف تحليل الانحدار الخطي البسيط إلى دراسة تأثير متغير مستقل واحد و ليكن X بالمتغير المفسر أو المتغير الخارجي، والإشارة إلى المتغير Y بالمتغير المفسر، وتعددت العلاقات الاقتصادية التي تم قياسها بالأسلوب الخطي ، وتكتب معادلة النموذج الخطي البسيط أو دالة انحدار المجتمع PRF (*) (النموذج النظري) بالشكل التالي:

$$\mathrm{y_i} = \alpha + \beta \mathrm{x_i} + \varepsilon_\mathrm{i} \ldots (06)$$

حيث :

$y_i$: المتغير التابع ، $x_i$: المتغير المستقل ، $\alpha$: الثابت أو القاطع، $\beta$: معامل الانحدار أو الميل ويعبر عن التغير في y الناتج عن التغير في x ، $\varepsilon_i$: الخطأ العشوائي.

ويرجع وجود حد الخطأ إلى عدة أسباب منها (شيخي، 2012، صفحة 19) :

- إهمال بعض المتغيرات المستقلة التي يمكن أن تؤثر على المتغير التابع في النموذج.
- الصياغة الرياضية غير السليمة للنموذج.

(*) Population Regression Function

- حدوث خطأ في كل من تجميع البيانات وقياس المتغيرات الاقتصادية.

وتعتمد صحة هذا النموذج على صحة الفرضيات الأساسية التالية (فروخي، 1993، صفحة 01):

1. الأمل الرياضي للأخطاء معدوم، وتعني هذه الفرضية أن الأخطاء $e_i$ لا تدخل في تفسير Y ، إذ تعبر عن حدود عشوائية لا يمكن قياسها أو تحديدها بدقة ويمكن التعبير عن هذه الفرضية رياضيا بالشكل التالي:

$$E(\varepsilon_i) = 0 , \forall i, i = \overline{1, n}$$

2. تتعلق بافتراض ثبات تباين الأخطاء أو تشتتها، وهو ما يعني أن تبعثرها حول المتوسط ثابت، ونعبر رياضيا بالكتابة التالية:

$$E\left(\varepsilon_i^2\right) = \sigma_\varepsilon^2, \forall i, i = \overline{1, n}$$

3. لا يوجد ارتباط بين الأخطاء المرتكبة على مشاهدات مختلف مكونات العينة، ونعبر عن هذه الفرضية رياضيا كما يلي:

$$\operatorname{cov}\left(\varepsilon_i, \varepsilon_j\right) = 0, \text{أو}, E(\left(\varepsilon_i, \varepsilon_j\right) = 0, \forall\ i \neq j$$

4. هذه الفرضية تتعلق بـX، وتفترض هذه الفرضية أن $(x_i)$ يمكن السيطرة عليها و ليس بمتغير عشوائي تماما.

وأضاف (بني هني، 2014، صفحة 22) بخصوص الفرضية الرابعة، أن المتغير $(x_i)$ متغير غير عشوائي، أي ثابت القيم من عينة إلى أخرى وتباينه محدود، وأكبر من صفر دائما ، والمقصود بثبات قيم المتغير $(x_i)$ هو أنه في حالة سحب عينات عشوائية عن x لمرات عديدة ، فإن قيم $(x_i)$ لا تتغير لكن $(y_i)$ تتغير مقابل كل قيمة

مثبتة لـ($X_i$)، كما أضاف (شيخي، 2012، صفحة 21)أن هذه الفرضية تشير إلى استقلالية الأخطاء عن المتغير x، ويمكن التعبير عنها رياضيا بالشكل التالي:

$$\mathrm{cov}(X_i, \varepsilon_j) = 0, \forall i, i = \overline{1, n}$$

هذه الفرضيات تشكل حجر الأساس الذي يرتكز عليه نظرية الانحدار، وفي حالية انتهاك واحدة أو أكثر من هذه الفرضيات ، فإن المعلمات المقدرة لن تتميز بالصفات التي تجعلها مرغوبة (بني هني، 2014، صفحة 03).

## 2- تقدير معالم النموذج باستخدام طريقة المربعات الصغرى:

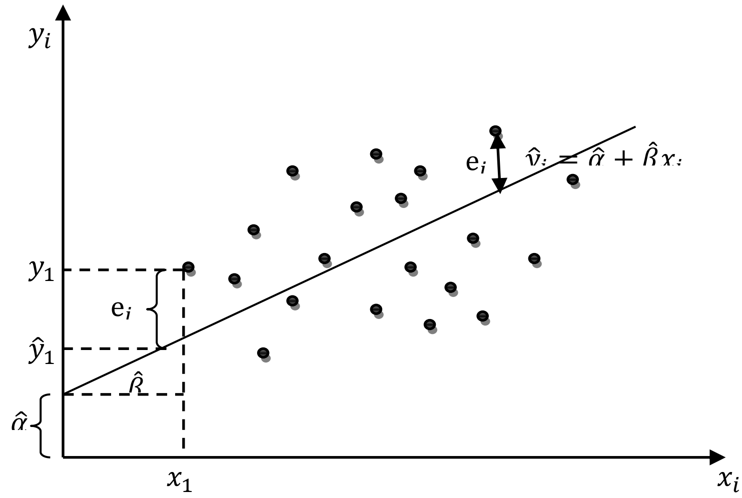
نظرا لعدم إمكانية الحصول على البيانات الخاصة بكافة أفراد المجتمع، يتم اختيار عينة عشوائية بهدف تقدير معالم المجتمع، وتسمى بدالة انحدار العينة SRF (*)أو معادلة الانحدار المقدرة (النموذج النظري)، وهي معرفة بالشكل التالي:

$$y_i = \hat{\alpha} + \hat{\beta}x_i + e_i \ldots..(07)$$

وعند سحب عينة مكونة من n ثنائية : $(x_1, y_1)$ ، .....، $(x_n, y_n)$ ، يكون التساؤل عن الخط الذي يعبر بكيفية جيدة عن العلاقة (07) و يكمن التساؤل حول الخط الأمثل الذي يعبر بكيفية جيدة عن هذه العلاقة الخطية بحيث يأخذ كل المشاهدات بعين الاعتبار، لذلك تم اقتراح طريقة المربعات الصغرى في تقدير معالم العلاقة (07) المتمثلة في الثابت (القاطع) $\hat{\alpha}$ و الميل $\hat{\beta}$ ، حيث تهدف هذه الطريقة إلى تدنية المسافة بين القيمة الفعلية للمتغير التابع $(y_i)$ وقيمتها المقدرة $\hat{y}_i$ حيث : $\hat{y}_i = \hat{\alpha} + \hat{\beta}x_i$ وهذه المسافة تشير إلى الأخطاء العشوائية$\varepsilon_i$، حيث: $\varepsilon_i = y_i - \hat{y}_i$ ، كما هو مبين في الشكل التالي:

(*)Sample Regression Function

الأخطاء العشوائية




المصدر: (شيخي، 2012، صفحة 02)

هذه الطريقة تحاول إيجاد أحسن تصحيح خطي من خلال تدنئة مربعات الانحرافات بين المشاهدات الفعلية والمشاهدات المقدرة $\sum_{i=1}^{n} \varepsilon_i^2$ ، ويمكن كتابة ذلك رياضيا كما يلي:

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - E(y_i))^2 = \sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta}x_i)^2 \dots (08)$$

حيث تشير العلاقة (08) إلى مجموع مربع انحرافات قيم $( y_i )$ عن وسطها $\alpha + \beta \mathrm{X_i}$، حيث تهدف تقديرات طريقة المربعات الصغرى إلى أن يكون الفرق بين كل نقطة من النقاط المبعثرة فوق وتحت الخط المستقيم ، والنقطة المقابلة لها على الخط المستقيم أقل ما يمكن ،ولإيجاد تقديرات كل من $\hat{\alpha}$ و $\hat{\beta}$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{y}\bar{x}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \Rightarrow \hat{\beta} = \frac{s_{xy}^2}{s_x^2}, \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

### 3– القدرة التفسيرية للنموذج (شيخي، 2012، 38–41):

تساعد البواقي $(\varepsilon_i)$ على قياس مدى تمثيل المعادلة المفروضة في النموذج لمشاهدات العينة، حيث أن القيمة الكبيرة للبواقي تعني بأن التمثيل يكون غير جيد و القيمة الصغيرة لها تعني تمثيلا جيدا للنموذج، إن المشكلة في استعمال البواقي كمقياس لجودة التوفيق هو أن قيمة البواقي تعتمد على المتغير التابع $y_i$، الذي نعرفه حول متوسطه انطلاقا من الشكل (02–02) كما يلي:

$$y_i = \hat{y}_i + \varepsilon_i$$

$$y_i - \bar{y} = \hat{y}_i - \bar{y} + \varepsilon_i$$

وبتربيع طرفي المعادلة أعلاه وجمعها بالنسبة لكل $i$ نجد:

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}{\varepsilon_i}^2$$

وتستعمل هذه المعادلة في قياس القدرة التفسيرية للنموذج ، و تعني حدودها ما يلي:

$\sum_{i=1}^{n}(y_i - \bar{y})^2$: هو مجموع مربعات الانحرافات الكلية للمتغير $y_i$ (TSS) (*).

$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$: هو مجموع مربعات الانحرافات المشروحة (ESS) (**).

$\sum_{i=1}^{n}{\varepsilon_i}^2$: هو مجموع مربعات البواقي (RSS) (***).

وعند إعادة صياغة المعادلة السابقة تصبح:

(*)Total Sum of Squares

(**)Sum of Squares

(***) Residual Sum of Squares

$$TSS = ESS + RSS$$

و بتقسيم كل الأطراف على الانحرافات الكلية TSS ، نجد:

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

وعليه نجد معامل التحديد $R^2$ ، كما يلي:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{\mathrm{n}} {\varepsilon_i}^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \dots (09)$$

حيث

$$0 \leq R^2 \leq 1$$

و يقيس معامل التحديد ويشرح الانحرافات الكلية أو التغيرات التي تحدث في المتغير التابع $y_i$، و المشروحة بواسطة تغيرات المتغير المستقل $x_i$، فهي نسبة تأثير المتغير المستقل على المتغير التابع ، وبالتالي فإنه يقيس القدرة التفسيرية للنموذج ، ويعتبر $R^2$ من أهم المعاملات التي تقيس علاقة الارتباط بين متغيرين ، ووجود مثل هذه العلاقة يعني ضمنيا أن أحد هذين المتغيرين يعتمد في تغيره أو في حدوثه على المتغير الأخر.

لما يأخذ $R^2$ أكبر قيمة وهي 1، فنقول أن القدرة التفسيرية للنموذج عالية جدا، أي هناك جودة في التوفيق والارتباط بين المتغير التابع و المستقل، ولما يأخذ أصغر قيمة له وهي الصفر، فنقول أن ليس للنموذج قوة تفسيرية ، ويعود ذلك لسببين إما أن العلاقة الموجودة بين المتغيرين هي غير خطية أو غياب السببية بينهما.

ويكمن الفرق بين معامل التحديد، ومعامل الارتباط في السببية ، حيث يقيس معامل الارتباط العلاقة بين متغيرين بغض النظر عن الدور الذي يلعبه كل متغير، بينما معامل التحديد فيقيس أيضا الارتباط ولكن يأخذ في الاعتبار السببية حيث أن متغير $x_i$ هو الذي يشرح الظاهرة $y_i$.

ويمكن حساب معامل التحديد المعدل بالعلاقة التالية:

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-2}$$

**مثال (50)**: نفس معطيات المثال رقم(47):

1- تقدير $\hat{\alpha}$ و $\hat{\beta}$: نعلم أن $\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$ ، ولحسابها نستخدم الجدول التالي:

| $x_i$ | $y_i$ | $x_i - \bar{x}$ | $y_i - \bar{y}$ | $(x_i - \bar{x})(y_i - \bar{y})$ | $(x_i - \bar{x})^2$ |
|---|---|---|---|---|---|
| 500 | 230 | 94,6666667 | 9 | 852 | 8961,77778 |
| 400 | 210 | -5,33333333 | -11 | 58,6666667 | 28,4444444 |
| 350 | 200 | -55,3333333 | -21 | 1162 | 3061,77778 |
| 450 | 210 | 44,6666667 | -11 | -491,333333 | 1995,11111 |
| 300 | 200 | -105,333333 | -21 | 2212 | 11095,1111 |
| 350 | 210 | -55,3333333 | -11 | 608,666667 | 3061,77778 |
| 420 | 220 | 14,6666667 | -1 | -14,6666667 | 215,111111 |
| 510 | 240 | 104,666667 | 19 | 1988,66667 | 10955,1111 |
| 360 | 215 | -45,3333333 | -6 | 272 | 2055,11111 |
| 270 | 180 | -135,333333 | -41 | 5548,66667 | 18315,1111 |
| 320 | 200 | -85,3333333 | -21 | 1792 | 7281,77778 |

| | | | | | |
|---|---|---|---|---|---|
| 350 | 210 | -55,3333333 | -11 | 608,666667 | 3061,77778 |
| 430 | 240 | 24,6666667 | 19 | 468,666667 | 608,444444 |
| 520 | 250 | 114,666667 | 29 | 3325,33333 | 13148,4444 |
| 550 | 300 | 144,666667 | 79 | 11428,6667 | 20928,4444 |
| 405,33333 | 221 | | | 29820 | 104773,333 |

حيث:

$$\bar{x} = 405{,}333333;\ \bar{y} = 221$$

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 29820$$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = 104773{,}333$$

$$\Rightarrow \hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{29820}{104773{,}333} = 0{,}28461$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}\ \Rightarrow \hat{\alpha} = 221 - (0{,}28461)(405{,}333333)$$

$$\Rightarrow \hat{\alpha} = 105{,}63629$$

$$\Rightarrow y_i = 105{,}63629 + 0{,}28461x_i + e_i$$

$$\Rightarrow \hat{y}_i = 105{,}63629 + 0{,}28461x_i$$

$$\Rightarrow e_i = y_i - \hat{y}_i$$

والجدول الموالي يبين تقديرات $\hat{y}_i$ و البواقي $e_i$:

| $y_i$ | $x_i$ | $\hat{y}_i$ | $e_i$ | $e_i^2$ |
|---|---|---|---|---|
| 230 | 500 | 247,943497 | −17,9434971 | 321,969087 |
| 210 | 400 | 219,482057 | −9,4820565 | 89,9093955 |
| 200 | 350 | 205,251336 | −5,25133622 | 27,5765321 |
| 210 | 450 | 233,712777 | −23,7127768 | 562,295783 |
| 200 | 300 | 191,020616 | 8,97938407 | 80,6293382 |
| 210 | 350 | 205,251336 | 4,74866378 | 22,5498077 |
| 220 | 420 | 225,174345 | −5,17434462 | 26,7738422 |
| 240 | 510 | 250,789641 | −10,7896411 | 116,416356 |
| 215 | 360 | 208,09748 | 6,90251973 | 47,6447786 |
| 180 | 270 | 182,482184 | −2,48218376 | 6,16123623 |
| 200 | 320 | 196,712904 | 3,28709595 | 10,8049998 |
| 210 | 350 | 205,251336 | 4,74866378 | 22,5498077 |
| 240 | 430 | 228,020489 | 11,9795113 | 143,508692 |
| 250 | 520 | 253,635785 | −3,63578519 | 13,2189339 |
| 300 | 550 | 262,174217 | 37,8257826 | 1430,78983 |
| 221 | 405,333333 | | | 2922,79842 |

**2- القدرة التفسيرية للنموذج (معامل التحديد):**

لدينا :

$$TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2 \,;\; \mathrm{ESS} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \,; \mathrm{RSS} = \sum_{i=1}^{\mathrm{n}} \mathrm{e}_i^2$$

وبالتعويض نجد :

$$TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2 = 11410$$

$$\mathrm{ESS} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = 8487{,}2016$$

$$\mathrm{RSS} = \sum_{i=1}^{\mathrm{n}} \mathrm{e}_i^2 = 2922{,}8$$

والجدول الموالي يوضح طريقة الحساب:

| $y_i$ | $\hat{y}_i$ | $e_i$ | $e_i^2$ | $\hat{y}_i - \bar{y}$ | $(\hat{y}_i - \bar{y})^2$ | $y_i - \bar{y}$ | $(y_i - \bar{y})^2$ |
|---|---|---|---|---|---|---|---|
| 230 | 247,9435 | -17,943 | 321,969 | 26,9435 | 725,95203 | 9 | 81 |
| 210 | 219,4821 | -9,4821 | 89,9094 | -1,51794 | 2,3041525 | -11 | 121 |
| 200 | 205,2513 | -5,2513 | 27,5765 | -15,7487 | 248,02041 | -21 | 441 |
| 210 | 233,7128 | -23,713 | 562,296 | 12,71278 | 161,61469 | -11 | 121 |
| 200 | 191,0206 | 8,97938 | 80,6293 | -29,9794 | 898,76347 | -21 | 441 |
| 210 | 205,2513 | 4,74866 | 22,5498 | -15,7487 | 248,02041 | -11 | 121 |
| 220 | 225,1743 | -5,1743 | 26,7738 | 4,174345 | 17,425153 | -1 | 1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 240 | 250,7896 | -10,79 | 116,416 | 29,78964 | 887,42272 | 19 | 361 |
| 215 | 208,0975 | 6,90252 | 47,6448 | -12,9025 | 166,47502 | -6 | 36 |
| 180 | 182,4822 | -2,4822 | 6,16124 | -38,5178 | 1483,6222 | -41 | 1681 |
| 200 | 196,7129 | 3,2871 | 10,805 | -24,2871 | 589,86303 | -21 | 441 |
| 210 | 205,2513 | 4,74866 | 22,5498 | -15,7487 | 248,02041 | -11 | 121 |
| 240 | 228,0205 | 11,9795 | 143,509 | 7,020489 | 49,287261 | 19 | 361 |
| 250 | 253,6358 | -3,6358 | 13,2189 | 32,63579 | 1065,0945 | 29 | 841 |
| 300 | 262,1742 | 37,8258 | 1430,79 | 41,17422 | 1695,3162 | 79 | 6241 |
| 221 | | | 2922,8 | | 8487,2016 | | 11410 |

$$R^2 = \frac{ESS}{TSS} = \frac{8487,2016}{11410} = 0.74$$

# قائمة المراجع

**قائمة المراجع:**

1. أحمد عبدالسميع طبيه. (2008). *مبادئ الإحصاء*. عمان ، الأردن: دار البداية .

2. أنيس إسماعيل كنجو. (2000). *الإحصاء و الإحتمال*. الرياض، السعودية: مكتبة العبيكان.

3. جمال فروخي. (1993). *نظرية الاقتصاد القياسي*. الجزائر : ديوان المطبوعات الجامعية.

4. خاشع محمود الراوي. (1984). *المدخل إلى الإحصاء*. الموصل، العراق: جامعة الموصل.

5. شرف الدين خليل. (دون ذكر السنة). *الإحصاء الوصفي*. شبكة الأبحاث و الدراسات الإقتصادية.

6. شوقي سيف النصر سيد، علي سعيد الديب، و مروة رفيق جلال. (2018). *الإحصاء في مجال الأعمال* . القاهرة ، مصر: كلية التجارة، جامعة القاهرة.

7. صلاح العيادي صالحين. *مسائل و حلول في الإحصاء و الاحتمالات مع حلول نماذج امتحانات* .

8. طارق البدري، و سهيلة نجم. (2014). *الإحصاء في المناهج البحثية التربوية و النفسية*. الأردن: دار الثقافة للنشر و التوزيع.

9. عايش موسى غرايبة، و عبدالله فلاح المنيزل. (2007). *الإحصاء التربوي، تطبيقات باستخدام الرزم الإحصائية للعلوم الإجتماعية*. الأردن: الطبعة الثانية ، دار المسيرة للنشر و التوزيع و الطباعة.

10. عبد الحميد عبدالمجيد البلداوي. (2009). *أساليب علم الإحصاء للعلوم الإقتصادية و إدارة الأعمال مع استخدام برنامج SPSS*. عمان ، الأردن : الطبعة الأولى ،دار وائل للنشر.

11. عبد الحميد عبدالمجيد البلداوي. (2008). *الأساليب الإحصائية التطبيقية* . عمان، الأردن: دار الشروق للنشر و التوزيع.

12. عبدالرحمن بن محمد سليمان أبو عمه، أنور أحمد محمد عبدالله، و محمود محمد ابراهيم هندي. (1990). *الإحصاء التطبيقي*. المملكة العربية السعودية: مطابع جامعة الملك سعود.

13. عبدالرزاق بني هني. (2014). *الاقتصاد القياسي (نظرية الانحدار البسيط و المتعدد)* . عمان ، الأردن: دار وائل للنشر و التوزيع ، الطبعة الأولى.

14. عدنان بن ماجد عبدالرحمن بري، محمود محمد إبراهيم هندي، و الحسيني عبدالبر ابراهيم راضي. (1998). *أساسيات طرق التحليل الاحصائي* . المملكة العربية السعودية: النشر العلمي و المطابع .

15. عدنان بن ماجد عبدالرحمن بري، محمود محمد إبراهيم هندي، و أنور أحمد محمد عبدالله. (1997). *مبادئ الإحصاء و الإحتمالات*. المملكة العربية السعودية: مطابع جامعة الملك سعود.

16. علي أحمد السقاف. (2020). *الإحصاء الوصفي و الإستدلالي* . برلين، ألمانيا : المركز الديمقراطي العربي .

17. كامل فليفل، و فتحي حمدان. (2013). *الإحصاء*. عمان، الأردن: دار المناهج للنشر و التوزيع.

18. لطفي أحمد محسن. (2011). *مقدمة في الإحصاء الإجتماعي*. المملكة العربية السعودية: النشر العلمي و المطابع ، جامعة الملك سعود.

19. محمد بن صالح الصغير. (2011). *مقدمة في الإحصاء الإجتماعي* . الرياض، المملكة العربية السعودية: النشر العلمي و المطابع، جامعة الملك سعود.

20. محمد حسين محمد رشيد. (2008). *الإحصاء الوصفي و التطبيقي و الحيوي* . عمان ، الأردن: الطبعة الأولى ، دار للنشر و التوزيع .

21. محمد خيثري. (1975). *الإحصاء النفسي التربوي*. الرياض ، السعودية: مطبوعات جامعة الرياض .

22. محمد رتول. (2006). *الإحصاء الوصفي* . الجزائر: الطبعة الثانية، ديوان المطبوعات الجامعية .

23. محمد شيخي. (2012). *طرق الاقتصاد القياسي (محاضرات و تطبيقات)*. دار الحامد للنشر و التوزيع ، الطبعة الأولى.

24. نبيل جمعة صالح النجار. (2015). *الإحصاء التحليلي مع تطبيقات برمجية SPSS*. عمان، الأردن: دار حامد للنشر و التوزيع.