جامعة محمد بوضياف بالمسيلة

كلية: العلوم الاقتصادية والتجارية وعلوم التسيير

قسم: العلوم الاقتصادية

مطبوعة دروس مقدمة إلى طلبة الجذع المشترك علوم اقتصادية، علوم التسيير والعلوم التجارية

العنوان

محاضرات في الإحصاء1

من إعداد:

د. بلقليل نورالدين

السنة الجامعية: 2025/2024

بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

الفهرس	الصفحة
مقدمة	أ
المحور الأول: نظرة عامة حول علم الاحصاء	
أولا: تطور علم الاحصاء	01
ثانيا: تعريف علم الاحصاء وأقسامه	02
1- تعريف علم الاحصاء	02
2- أقسامه	02
2-1 الاحصاء الوصفي	02
2-2 الاحصاء الاستدلالي	02
ثالثا: أهمية الاحصاء في الادارة والاقتصاد و مجالات استخدامه	03
1- أهمية الاحصاء في الادارة والاقتصاد	03
2- مجالات استخدامه	03
رابعا : مراحل أو خطوات المنهاج الاحصائي	03
خامسا : مصادر وأساليب جمع البيانات	04
1- مصادر جمع البيانات	04
1-1 المصادر المباشرة أو الأولية	04
1-2 المصادر غير المباشرة أو الثانوية	04
2- أساليب جمع البيانات	05
2-1 أسلوب الحصر الشامل	05
2-2 أسلوب العينات	05
سادسا : بعض المصطلحات والمفاهيم الاحصائية العامة	05
1- الوحدة الاحصائية	05
2- المجتمع الاحصائي	05
3- المتغير الاحصائي	06

3-1 متغيرات كيفية	06
3-1-1 متغيرات كيفية قابلة للترتيب	06
3-1-2 متغيرات كيفية غير قابلة للترتيب	06
3-2 متغيرات كمية	06
3-2-1 متغيرات كمية منفصلة أو متقطعة	06
3-2-2 متغيرات كمية متصلة أو مستمرة	06
4- العينة	06
المحور الثاني: العرض الجدولي للبيانات	
أولا : الجداول التكرارية	07
1- الجداول التكرارية البسيطة المطلقة	07
1-1 بالنسبة للمتغير الوصفي (النوعي)	08
1-2 بالنسبة للمتغير الكمي	09
1-2-1 حالة متغير كمي منفصل (متقطع)	09
1-2-2 حالة متغير كمي متصل (مستمر)	10
2-الجداول التكرارية البسيطة النسبية و النسبية المئوية	12
2-1 التكرار النسبي	12
2-2 التكرار النسبي المئوي	12
3- الجداول التكرارية المتجمعة	14
3-1 جدول التوزيع التكراري التجميعي الصاعد	14
3-2 جدول التوزيع التكراري التجميعي النازل	15
المحور الثالث: العرض البياني للبيانات	
أولا- العرض البياني في حالة متغير وصفي (نوعي)	17
1- العرض الدائري	17
2- الأعمدة المستطيلة	18
3- العمود المجزأ	19

ثانيا– العرض البياني في حالة متغير كمي منفصل (متقطع)	20
1– الأعمدة البسيطة	20
2– العرض البياني للتكرار التجميعي الصاعد والتكرار التجميعي النازل	21
ثالثا– العرض البياني في حالة متغير كمي متصل (مستمر)	22
1– المدرج التكراري	23
1-1 فئات الطول متساوية	23
2-1 فئات الطول غير متساوية	24
2 – المضلع التكراري	25
3 – المنحنى التكراري	26
4–المنحنى التجميعي الصاعد	27
5–المنحنى التجميعي النازل	27
المحور الرابع: مقاييس النزعة المركزية	
أولا–الوسط الحسابي	29
1– الوسط الحسابي للبيانات غير المبوبة	29
1-1 الطريقة المباشرة	29
2-1 طريقة الانحرافات البسيطة (الوسط الفرضي)	30
2– الوسط الحسابي للبيانات المبوبة	30
1-2 الوسط الحسابي في حالة متغير كمي منفصل	30
1-1-2 الطريقة المباشرة	30
2-1-2 طريقة الانحرافات البسيطة (الوسط الفرضي)	31
2-2 الوسط الحسابي في حالة متغير كمي متصل	32
1-2-2 الطريقة المباشرة	32
2-2-2 طريقة الانحرافات البسيطة (الوسط الفرضي)	33
03– خصائص الوسط الحسابي	34
ثانيا–الوسيط	35

1- الوسيط للبيانات غير المبوبة	35
2- الوسيط للبيانات المبوبة	36
2-1 الوسيط في حالة متغير كمي منفصل	36
2-2 الوسيط في حالة متغير كمي متصل	37
3- خصائص الوسيط	39
ثالثا-المنوال	39
1- المنوال للبيانات غير المبوبة	39
2- المنوال للبيانات المبوبة	39
2-1 المنوال في حالة متغير كمي منفصل	39
2-2 المنوال في حالة متغير كمي متصل	39
3- خصائص المنوال	41
رابعا-الوسط (المتوسط) الهندسي	41
1- الوسط الهندسي للبيانات غير المبوبة	41
2- حالة البيانات المبوبة	41
2-1 حالة متغير كمي منفصل	41
2-2 حالة متغير كمي متصل	42
خامسا-الوسط التوافقي	42
1- الوسط التوافقي للبيانات غير المبوبة	42
2- حالة البيانات المبوبة	42
2-1 حالة متغير كمي منفصل	43
2-2 حالة متغير كمي متصل	43
سادسا-الوسط التربيعي	43
1- الوسط التربيعي للبيانات غير المبوبة	43
2- حالة البيانات المبوبة	43
2-1 حالة متغير كمي منفصل	43

2-2 حالة متغير كمي متصل	43
سابعا-مشتقات الوسيط (الربيعات و العشيرات، والمئينات)	45
1- حالة البيانات غير المبوبة	46
2- حالة البيانات المبوبة	47
2-1 حالة متغير كمي منفصل	47
2-2 حالة متغير كمي متصل	48
المحور الخامس: مقاييس التشتت	
أولا-مقاييس التشتت المطلقة	50
1- المدى	50
1-1 المدى للبيانات غير المبوبة	51
1-2 المدى للبيانات المبوبة	51
1-2-1 حالة متغير كمي منفصل	51
1-2-2 حالة متغير كمي متصل	51
1-3 خواص المدى	51
2- المدى الربيعي	52
2-1 خواص المدى الربيعي	52
3- الانحراف المتوسط	52
3-1 الانحراف المتوسط للبيانات غير المبوبة	52
3-2 الانحراف المتوسط للبيانات المبوبة	53
3-2-1 حالة متغير كمي منفصل	53
3-2-2 حالة متغير كمي متصل	54
3-3 خواص الانحراف المتوسط	54
4-التباين والانحراف المعياري	55
4-1 التباين	55
4-1-1 التباين للبيانات غير المبوبة	55

4-1-2 التباين للبيانات المبوبة	55
4-1-2-1 حالة متغير كمي منفصل	55
4-1-2-2 حالة متغير كمي متصل	56
4-1-3 خواص التباين	57
4-2 الانحراف المعياري	57
4-2-1 الانحراف المعياري للبيانات غير المبوبة	57
4-2-2 الانحراف المعياري للبيانات المبوبة	57
4-2-2-1 حالة متغير كمي منفصل	57
4-2-2-2 حالة متغير كمي متصل	58
4-2-3 خواص الانحراف المعياري	58
ثانيا –مقاييس التشتت النسبية	59
1- المدى النسبي	59
2- المدى الربيعي النسبي	59
3- الانحراف المتوسط النسبي	59
4- الانحراف المعياري النسبي (معامل الاختلاف)	59
المحور السادس: مقاييس الشكل	
أولا – العزوم	62
1 – العزوم البسيطة (العزوم حول نقطة الصفر)	63
1-1 العزوم البسيطة للبيانات غير المبوبة	63
1-2 العزوم البسيطة للبيانات المبوبة	63
1-2-1 حالة متغير كمي منفصل	63
1-2-2 حالة متغير كمي متصل	63
2 – العزوم المركزية (العزوم حول المتوسط الحسابي)	63
2-1 العزوم المركزية للبيانات غير المبوبة	63

2-2 العزوم المركزية للبيانات المبوبة	63
2-2-1 حالة متغير كمي منفصل	63
2-2-2 حالة متغير كمي متصل	63
ثانيا – الالتواء	64
1 – مقارنة مقاييس النزعة المركزية	65
2 – قانون بيرسون للإلتواء	65
2-1 قانون بيرسون الأول	65
2-2 قانون بيرسون الثاني	65
3 – معامل فيشر للإلتواء	66
3-1 حالة البيانات غير المبوبة	66
3-2 حالة البيانات المبوبة	66
3-2-1 حالة متغير كمي منفصل	66
3-2-2 حالة متغير كمي متصل	66
ثالثا – التفرطح	68
1 – معامل بيرسون للتفرطح	68
1-1 حالة البيانات غير المبوبة	69
1-2 حالة البيانات المبوبة	69
1-2-1 حالة متغير كمي منفصل	69
1-2-2 حالة متغير كمي متصل	69
2 – معامل فيشر للتفرطح	69
2-1 حالة البيانات غير المبوبة	69
2-2 حالة البيانات المبوبة	69
2-2-1 حالة متغير كمي منفصل	69
2-2-2 حالة متغير كمي متصل	69
المحور السابع: مقاييس التمركز	

أولا – التحديد البياني للتمركز (مربع جيني)	71
ثانيا – التحديد الحسابي للتمركز (حساب معامل جيني للتمركز)	73
1 – الطريقة الأولى	73
2 – الطريقة الثانية	74
المحور الثامن : الأرقام القياسية	
أولا – الأرقام القياسية الفردية	81
1- لسعر سلعة	81
2- لكمية سلعة	81
ثانيا – الأرقام القياسية المركبة (التجميعية)	82
1- الارقام القياسية البسيطة (غير المرجحة) للأسعار	82
1-1 الطريقة النسبية	82
2-1 الطريقة التجميعية	82
2- الأرقام القياسية المرجحة للأسعار	83
1-2 رقم لاسبير المرجح ***laspeyres Index***	84
1-1-2 الطريقة النسبية	84
2-1-2 الطريقة التجميعية	84
2-2 رقم باش المرجح ***Paasche Index***	84
1-2-2 الطريقة النسبية	84
2-2-2 الطريقة التجميعية	84
3-2 رقم فيشر المرجح ***Fishers Index***	84
1-3-2 الطريقة النسبية	84
2-3-2 الطريقة التجميعية	84
4-2 رقم مارشال المرجح ***Marchal Index***	84
1-4-2 الطريقة النسبية	85

2-4-2 الطريقة التجميعية	85
المحور الثامن : الإرتباط والانحدار	
أولا: الارتباط	88
1-تحليل الارتباط	88
1-1 نوع العلاقة	88
1-2 قوة العلاقة	89
2-لوحة الانتشار وعلاقتها بالارتباط	89
3-معامل الارتباط	90
3-1 معامل الارتباط الخطى البسيط " لبيرسون " ***Pearson***	91
3-2 معامل الارتباط الرتبي (معامل الارتباط سبيرمان للرتب)	92
ثانيا: الانحدار	94
1-خط الانحدار البسيط	94
2- معادلة الانحدار	94
3- الخطأ المعياري	96
تمارين محلولة	99- 130
قائمة المراجع	

مقدمة:

تزداد أهمية علم الاحصاء يوما بعد يوم بفعل الاستخدامات العديدة له وعلى نطاق واسع وفي مختلف الميادين والانشطة فهو يساعد على الاستشراف واتخاذ القرارات المناسبة بفعل استخلاصه للنتائج التي بنيت على كم معين من المعلومات والبيانات التي تم تجميعها وتبويبها وتنظميها مسبقا . وقد ساير هذا الاستخدام المكثف له تطور الأساليب التقنية في التحليل بفعل استخدام برمجيات الحاسوب وغيرها وما توفره من معلومات هائلة وفي ظرف وجيز .

ويعتبرالإكتساب لمفاهيم مقياس الاحصاء 1 أو الاحصاء الوصفي أمر ضروري لمختلف طلبة كلية العلوم الاقتصادية والتجارية وعلوم التسييرلكونه قاعدة وركيزة مهمة قد يستند عليها في مختلف المواد والسنوات اللاحقة، وبهدف إلمام الطلبة بهذا المقياس وتسهيل فهمهم له، تم تصميم هذه المطبوعة والموجهة لطلبة السنة أولى(ل.م.د)جذع مشترك علوم اقتصادية، التسيير وعلوم تجارية والمتضمنة مجموعة من الدروس المدعمة بأمثلة وتمارين محلولة عن كل الحالات التي تهم الطالب و وفقا للمقرر الدراسي لوزارة التعليم العالي والبحث العلمي ، وقد تم تنظيم هذه الدروس في تسعة (09) محاور جاءت كالآتي:

- المحور الأول: نظرة عامة حول علم الاحصاء
- المحور الثاني : العرض الجدولي للبيانات
- المحور الثالث: العرض البياني للبيانات
- المحور الرابع : مقاييس النزعة المركزية
- المحور الخامس: مقاييس التشتت
- المحور السادس: مقاييس الشكل
- المحور السابع : مقاييس التمركز
- المحور الثامن: الارقام القياسية
- المحور التاسع : الارتباط والانحدار

المحور الأول: نظرة عامة حول علم الإحصاء

أولا: تطور علم الإحصاء

عـــرف علم الإحصاء قديما إذ وردت إشارة عن العد من قبل المؤرخ اليوناني هيرودوتس حين ذكر أنه في عام 480 ق.م إستعمل أحد قادة الجيوش طريقة بدائية بسيطة لمعرفة عدد جيشه.

و قديما أيضا من خلال دلالات بسيطة وبدائية استعمل فيها العد كأساس لحصر الأشياء المحيطة بالإنسان وغيرها، فقد شهدت مختلف الحضارات القديمة إستخدام أساليب حسابية وإحصائية عديدة أثبتتها انجازات وآثار تلك الحضارات كتلك التي أقيمت في مصر والصين .

إن ظهور الدولة وتطورها ساعد على زيادة الإهتمام بالبيانات والمعلومات واصبح من الضروري جمع هذه البيانات وتصنيفها وتبويبها من أجل خدمة أمور الدولة والمجتمع، كتحديد السكان الملزمين بدفع الضرائب، و إحصاء نفقات الدولة وإيراداتها ...إلخ، فهي كلها إستخدامات تشمل مختلف المجالات خاصة الاقتصادية والاجتماعية.

يعد القرن السادس عشر بداية تطورعلم الإحصاء في أوروبا حيث أصبح كمادة علمية تدرس في الجامعات لأول مرة في ألمانيا سنة 1748م، كما ورد في أول نشرة إحصائية صدرت في انجلترا عام 1971م ، ومنذ أواخر القرن التاسع عشر أصبح يدرس في معظم الكليات والمعاهد التعليمية مثل كليات الطب والصيدلة والزراعة والعلوم والآداب والسياحة الفندقية ...واستخدم الاحصاء في أمريكا لأول مرة في الصناعة عام 1920م وإنتشر إستعماله في هذا الميدان بعد الحرب العالمية الثانية[1]. وبداية من النصف الثاني للقرن العشرون أصبح علم الإحصاء أكثر ازدهارا وتقدما بفعل إنتشار الحاسبات الالكترونية ودقتها وصولا إلى إنتشار البرامج الإحصائية الحديثة في يومنا هذا والتي سهلت الوصول الى القراءات الإحصائية المناسبة بفعل التطبيقات الذكية التي تحوزها .

إن الإستخدامات العديدة لأساليب علم الإحصاء خاصة من خلال استخدام الأساليب الكمية اللازمة لذلك ساهم في فهم وتفسير عديد الظواهر وإعطاء إستنتاجات وتنبؤات لها في مختلف العلوم .

[1] – صلاح الدين حسين الهيتي ، الاساليب الاحصائية في العلوم الادارية: تطبيقات باستخدام spss، دار وائل للطباعة والنشر،عمان، الطبعة الثانية 2006، ص ص 20، 21.

ثانيا: تعريف علم الإحصاء وأقسامه

1- تعريف علم الإحصاء: يتكون مصطلح Statistics والذي بناؤه Stat-is-tics المشتق من كلمة State أي الدولة والذي يعني مجموعة الحقائق الخاصة بشؤون الدولة [1]، وهذه الكلمة ليست حديثة، فالمفهوم المحدود لعلم الإحصاء المرتبط باعتقادات عامة الناس هو مجرد عد أو حصر الأشياء والتعبير عنها بأرقام، ولكن الإحصاء كعلم يمكن تعريفه على أنه :" العلم الذي يهتم بطرق جمع البيانات، وتبويبها، وتلخيصها بشكل يمكن الإستفادة منها في وصف البيانات وتحليلها للوصول إلى قرارات سليمة في ظل ظروف عدم التأكد " [2]، ويخص هذا طبعا مجالات الإقتصاد والأعمال وغيرها من العلوم الاجتماعية والطبيعية .

في حين أن الإحصائيات هي البيانات العددية التي تخص موضوع معين، فهي تعتبر المادة الخام والأساسية التي يستند عليها علم الإحصاء. فوظيفة علم الإحصاء هي تقديم المعلومات الرقمية التي تساعد على فهم جوهر الحياة الاقتصادية والاجتماعية وبالتالي فإن الإحصاء هو أساس اتخاذ القرارات ورسم الخطط من أجل التنمية والتطور سواء كانت هذه الخطط على مستوى المؤسسة أو على مستوى القطاع أو الدولة [3].

2- أقسامه: ينقسم علم الإحصاء بصفة عامة إلى قسمين رئيسيين : الإحصاء الوصفي والإحصاء الاستدلالي

2-1 الإحصاء الوصفي: يهتم الإحصاء الوصفي بعمليات جمع وتنظيم وتلخيص البيانات العددية الرقمية بدلالة بعض المقاييس لأغراض الوصف والمقارنة، كما يتم عرض هذه البيانات على شكل جداول أو رسومات بيانية[4]. فالهدف من كل هذا هو إعطاء وصف أولي لهذه البيانات دون محاولة إستنتاج أي شيء .

2-2 الإحصاء الاستدلالي: هو عبارة عن ذلك الفرع من علم الإحصاء الذي يقدم ويستخدم الأدوات والوسائل الفنية للحصول على استنتاجات وتقديرات من خلال التحليل السليم للمعلومات الرقمية من العينات وتعميمها على المجتمع [5]. وعملية إختيار العينة يجب أن تتم بطريقة علمية صحيحة حتى يمكن تعميم نتائج الجزء على الكل .

---

[1] – نفس المرجع، ص 23.

[2] – شرف الدين خليل ، الاحصاء الوصفي، مكتبة شبكة الابحاث والدراسات الاقتصادية.

[3] – محمود محمد سليم صالح ، مقدمة في الاحصاء ، مكتبة المجتمع العربي للنشر والتوزيع،عمان، الطبعة الأولى 2008، ص 17.

[4] – سالم عيسى بدر، عماد غصاب عبابنة ، مبادئ الاحصاء الوصفي والاستدلالي ، دار المسيرة للنشر والتوزيع والطباعة،الاردن،الطبعة الاولى 2007، ص 13.

[5] – أحمد فوزي ملوخيه ، مبادئ علم الاحصاء ، مركز الاسكندرية للكتاب،مصر، 2009، ص 12.

ثالثا: أهمية الإحصاء في الادارة والاقتصاد و مجالات استخدامه

1- أهمية الإحصاء في الادارة والاقتصاد: يكتسب الإحصاء أهميته في مجالات الادارة والاقتصاد والمحاسبة حيث أن عملية جمع البيانات في هذه الميادين لم تقتصر على تسجيل و حصر الحوادث بل تعدتها الى تقييم البيانات المسجلة وتحليلها واتخاذ قرارات تتعلق بالمستقبل، ويظهر بوضوح مدى إرتباط الإحصاء بالادارة فالتخطيط السليم لا يمكن تحقيقه دون قدركاف من البيانات الإحصائية والقدرة على فهم البيانات الإحصائية وأساليب إستخدامها تعتبر ضرورة لكل من يشارك بعمليات التخطيط أو يستفيد منها حيث تشتمل عملية التنظيم على مراجعة الأداء و من ثم قياس كفاءة وفاعلية الأداء الذي لا يمكن قياسه دون استخدام البيانات والأساليب الإحصائية، ويستلزم (التوجيه، والتفويض والتحفيز، والتنسيق) إجراء دراسات ومقارنات عن نتائج إتباع سياسات معينة وهذا يتطلب إغناءها بالبيانات الإحصائية، كما أن الرقابة وتصميم التقارير وتطوير معايير وقياس النتائج والتقييم يحتاج لمعرفة الأساليب الإحصائية، وفي كل هذه العمليات تبرز أهمية اتخاذ القرار المناسب الذي هو من أهم مسؤوليات الادارة [1].

2- مجالات استخدامه: حقيقة ان استخدام الإحصاء لا يشمل مجال إدارة الأعمال والاقتصاد فقط باعتبار دوره في التخطيط والتنفيذ، ولكن الإحصاء يشمل عديد المجالات الحياتية ذات الإهتمامات المتنوعة، فهو يستخدم مثلا في مجالات البحوث العلمية المختلفة بكثرة، كالعلوم الطبية التي توظفه في حصر نتائج التجارب وتحليلها، والعلوم الفيزيائية والاجتماعية...الخ ، فهو يمكن القائمين على الابحاث و فروع العلم المختلفة من حل مشكلات واتخاذ قرارات مناسبة بناءا على مقارنات أوتقديرات .

رابعا : مراحل أو خطوات المنهاج الإحصائي:

تتطلب كل دراسة إحصائية خطوات مهمة نوجزها في مايلي [2]:

- **تحديد المشكلة و وضع الفروض:** فالعملية الإحصائية تبدأ بمشاهدة الظواهر التي نرغب في دراستها، ومن هنا يتولد الإحساس بالمشكلة و وضع فرض مبدئي لتفسير الظاهرة موضوع الدراسة .
- **جمع البيانات الاحصائية:** بغرض فحص الظاهرة المراد دراستها يتم تحديد طبيعة البيانات المطلوبة للدراسة وحجمها والإمكانيات المادية والبشرية اللازمة لإعداد هذه الدراسة وكذا الوسيلة المناسبة لجمع هذه البيانات .

[1] – صلاح الدين حسين الهيتي ، مرجع سابق، ص 22.

[2] – ابراهيم علي ابراهيم عبد ربه ، مقدمة في الاحصاء الوصفي ، الدار الجامعية، مصر، 2005، ص ص 8 – 9.

- **تجهيز وتبويب وعرض البيانات الإحصائية:** يتم من خلال فرز البيانات وتبويبها بطريقة تساعد على فهم مدلولها والإستفادة منها ويكون ذلك بعرضها إما في صورة جداول رقمية وتوضيحها في صورة بيانية أو أشكال هندسية مختلفة .
- **تحليل البيانات الإحصائية:** باستخدام بعض المقاييس الإحصائية كمقاييس النزعة المركزية ومقاييس التشتت والإلتواء والإرتباط والإنحدار ..الخ، يمكن إجراء عمليات التحليل المختلفة على الظاهرة المدروسة، كتحديد درجة تباين أو عدم تجانس توزيع بيانات الظاهرة ، بالإضافة إلى تحديد العلاقة وقوتها واتجاهها بين ظاهرتين أو أكثر..الخ.
- **إستخلاص وتفسير و استخدام النتائج الإحصائية:** بانتهاء مرحلة التحليل يصبح أمام الباحث الإحصائي نتائج رقمية محددة مقنعة ويتعين عليه بعد ذلك تفسير هذه النتائج بحكمة ومهارة وموضوعية مع التحليل الإحصائي الذي تم اجراؤه . ومن ثم استخلاص النتائج و التنبؤ للظاهرة موضوع البحث .

خامسا : مصادر وأساليب جمع البيانات

تعتبر عملية جمع البيانات مهمة جدا في أي عمل إحصائي، فكلما كانت البيانات الإحصائية المجمعة صحيحة ودقيقة كلما كان تشخيص المشكلة المدروسة سليم وبالتالي نتمكن من اتخاذ القرارات المناسبة ، وتختلف مصادر وأساليب جمع البيانات باختلاف الدراسة المراد القيام بها .

1- مصادر جمع البيانات: تنقسم مصادر جمع البيانات الى قسمين :

1-1 المصادر المباشرة أو الأولية: المصدر الأولي لجمع البيانات هو ذلك المصدر الذي يقوم بجمع البيانات بنفسه أو تحت إشرافه، فإذا ما قام باحث بدراسة عينة من المجتمع وجمع من أفراد العينة (بنفسه أو تحت إشرافه) البيانات المطلوبة لبحثه فهذا مثال على مصدر مباشر لجمع البيانات وتسمى البيانات التي يجمعها بيانات أولية. وتتعدد صيغ جمع البيانات الأولية فنجد منها المقابلات الشخصية المباشرة، الإستبانات، وغيرها[1] .

1-2 المصادر غير المباشرة أو الثانوية: في كثير من الأحيان يجد الباحث نفسه في وضع لايتمكن به من جمع البيانات بنفسه أو تحت إشرافه ولذلك يلجأ إلى المصادر غير المباشرة، أي البيانات التي جمعها غيره، وتكون هذه البيانات إما من مصادر منشورة كالتقارير والمنشورات الرسمية التي تصدرها الهيئات والدواوين

[1] – محمد صبحي أبو صالح ، مبادئ الاحصاء ، دار اليازوري العلمية للنشر والتوزيع، الأردن، 2010، ص ص 12- 13.

المختلفة ، أو الهيئات الدولية المختلفة ، أو من مصادر غير منشورة كتلك البيانات المدونة في سجلات المواليد في المستشفيات وغيرها .

2- أساليب جمع البيانات: تنقسم أساليب جمع البيانات الى :

2-1 أسلوب الحصر الشامل: وفيه يتم جمع البيانات عن الظاهرة موضوع الدراسة من جميع مفردات المجتمع الإحصائي المراد بحثه سواء أكان نطاقه أو مجاله واسعا أو محدودا وفي كلتا الحالتين يتطلب هذا الأسلوب توافر إمكانيات مادية وبشرية وعينية و زمنية أكبر نسبيا من أسلوب العينات. وعموما يتميز أسلوب الحصر الشامل بخلوه من الأخطاء العشوائية (الصدفة) كما أنه يعطي صورة مفصلة عن مفردات الظاهرة موضوع الدراسة بحكم اتساع نطاقه، وما يعاب عليه هو الزيادة الكبيرة في التكاليف المادية والعينية والبشرية والزمنية اللازمة لإجراء الدراسة [1].

2-2 أسلوب العينات: بمقتضى هذا الأسلوب يتم جمع البيانات عن جزء فقط من مفردات المجتمع الإحصائي، أي من عينة من هذا المجتمع، يتم سحبها بطريقة بطريقة علمية سليمة بما يساعد في تعميم نتائجها على مجتمع البحث، ومزايا هذا الأسلوب عديدة منها التوفير الكبير في التكاليف المادية والعينية والبشرية والزمنية اللازمة لاجراء الدراسة، أما سلبيات هذا الأسلوب فنتائجه تكون أقل دقة من نتائج أسلوب الحصر الشامل.

وتجدر الإشارة إلى أنه في بعض المرات نكون مجبرين على استخدام هذا الأسلوب في الحالات التي تتعرض فيها مفردات مجتمع البحث للتدمير أو الهلاك الجزئي أو الشامل عند فحصها .

سادسا : بعض المصطلحات والمفاهيم الإحصائية العامة

1- الوحدة الإحصائية: هي الوحدة الأساسية لتكوين المجتمع الإحصائي [2]، والتي تجرى عليها الدراسة الإحصائية ، ويمكن أن تكون هذه الوحدة إنسان أو حيوان أو شيء.

2- المجتمع الإحصائي: هو مجموع الوحدات الإحصائية المراد دراستها والمعرفة بشكل دقيق والتي تشترك في ما بينها في الصفة الأساسية محل إهتمام الباحث مثل: مجتمع من الطلبة ، مجتمع من الأسر ، مجتمع من المؤسسات [3].

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 14.

[2] – جلاطو جيلالي ، الاحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الجزائر، 2001، ص 5.

[3] – عبد الحميد قطوش ، محاضرات في الاحصاء 1، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة المسيلة، 2022-2023، ص 06.

3- المتغير الإحصائي: يعبر عن الظاهرة المدروسة، وهو الخاصية المدروسة في المجتمع الإحصائي، مثلا: طول القامة، السن، الوزن، العلامة المتحصل عليها في إمتحان معين، الإنتاج، الإدخار، الإستثمار، الإستهلاك ....إلخ.

وينقسم المتغير الإحصائي إلى قسمين:

3-1 متغيرات كيفية: هي تلك المتغيرات التي تصف الظاهرة المدروسة بشكل غير رقمي لإنها غير قابلة للقياس ، وتنقسم الى قسمين:

3-1-1 متغيرات كيفية قابلة للترتيب: مثل المستوى التعليمي (ابتدائي، متوسط، ثانوي، جامعي)..الخ.

3-1-2 متغيرات كيفية غير قابلة للترتيب: مثل نوع الشركات (عامة، خاصة، مساهمة)، أنواع المسافرين (قادمون، مغادرون)، الجنسية (جزائري، ليبي،..)، الحالة الاجتماعية (مطلق، أرمل، أعزب، متزوج)، جنس المولود (ذكر، أنثى)،...الخ.

3-2 متغيرات كمية: هي تلك المتغيرات التي تصف الظاهرة المدروسة بشكل رقمي أي "كمي"، وتنقسم إلى قسمين:

3-2-1 متغيرات كمية منفصلة أو متقطعة: وهي التي يتم التعبير عنها على شكل أعداد أو أرقام صحيحة ولا تحتوي على كسور مثل أعداد الطلاب أو أفراد الأسرة ...الخ[1].

3-2-2 متغيرات كمية متصلة أو مستمرة: وهي التي تأخذ كل القيم الممكنة لمجال الدراسة، كأن نقول أن عمر الطالب يتراوح بين(20-22) سنة أو أجرة العامل تتراوح بين (100 - 500) دينار وغيرها من الأمثلة، ونظرا للعدد اللامتناهي لهذه القيم نقسم مجال الدراسة الى مجالات جزئية تسمى الفئات .

4- العينة: هي جزء من بيانات المجتمع الإحصائي، ويجب أن تختار بحيث تمثل جميع خصائص وصفات المجتمع الإحصائي[2] .

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 32.

[2] – محمود محمد سليم صالح ، مرجع سابق، ص 19.

المحور الثاني: العرض الجدولي للبيانات

تمهيد:

بعد عملية الانتهاء من مرحلة جمع البيانات يجد الباحث أو الهيئة المشرفة على الدراسة كمًا هائلا من البيانات الخام ، وتوافر هذا الكم الهائل من البيانات الخام لا يعني أنه يفيد في إجراء الدراسات اللازمة وإظهار نتائج المشكلة موضوع الدراسة، فيجب هنا القيام بعمليات تجميع وتنسيق و ترتيب هذه البيانات، وبمعنى آخر تصنيفها وعرضها بما يسمح بسهولة إستيعابها من ناحية ، وامكانية وسهولة دراستها وتحليلها من ناحية أخرى. إن الهدف من عمليات التصنيف والتبويب هو تجميع وتلخيص البيانات التي تم جمعها في مجموعات متجانسة، تختلف باختلاف طبيعة هذه البيانات، وأيضا لكيفية استخدامها بعد إجراء التبويب لها، ومما لا شك فيه أن الجداول الاحصائية هي الوسيلة المثلى لإجراء عمليات التلخيص والتبويب المشار إليه [1].

أولا : الجداول التكرارية

إن أي ظاهرة مدروسة نلاحظ أنها تحتوي على مجموعة كبيرة من البيانات وعند عرضها بطريقة الجداول يصعب فهمها وأحيانا لا نستطيع المقارنة بين مفردات هذه البيانات، ولذلك لا بد من تلخيصها و عرضها بطريقة مبسطة تسهل فهمها. ومن هذه الطرق نجد التوزيعات التكرارية. فالتوزيعات التكرارية هي إحدى الطرق التي نتمكن من خلالها بتنظيم البيانات الكثيرة بحيث لا تخسر هذه البيانات من أهميتها إلا الشيء القليل أو ربما لا نخسر شيئا[2].

نسمي الجداول التي تعرض عن طريقة التوزيع التكراري بالجداول التكرارية ، ويتكون الجدول التكراري الذي هو أبسط وأهم طريقة لعرض البيانات الاحصائية من القيم المختلفة للمتغير المدروس بالإضافة الى التكرارات، وتتنوع هذه الجداول كما يلي :

1- الجداول التكرارية البسيطة المطلقة : يوضع هذا الجدول للبيانات المتعلقة بظاهرة واحدة سواء كانت وصفية أو كمية، و يحتوي الجدول التكراري البسيط المطلق على عمودين، العمود الأول يمثل صفات أو قيم المتغير الإحصائي المدروس أو فئات لقيم الظاهرة، وهذا حسب نوعية المتغير. أما العمود الثاني فيمثل التكرارات.

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 37.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص 47.

1-1 بالنسبة للمتغير الوصفي (النوعي) : يتم عرض بيانات هذا المتغير في جدول تكراري بسيط مطلق من خلال وضع صفات أو مستويات المتغير في العمود الأول وتكرارات كل صفة أو مستوى في العمود الثاني.

مثال 1: في مايلي التقديرات في مادة الإحصاء لـــ 30 طالبا في قسم العلوم الاقتصادية :

ممتاز	جيد	ضعيف	مقبول	مقبول
جيد	ضعيف	ضعيف	جيد جدا	ممتاز
مقبول	جيد جدا	ضعيف	جيد	مقبول
جيد جدا	جيد جدا	جيد	مقبول	ضعيف
ممتاز	جيد	مقبول	مقبول	جيد
جيد	جيد	جيد	جيد	جيد

والمطلوب: تبويب البيانات السابقة في صورة جدول تكراري

الحل: التقديرات هنا عبارة عن صفات، ويمكن تبويبها على أساس هذه الصفات كما يلي:

المرحلة الاولى: **جدول (2-1): التفريغ**

الصفة (التقدير)	عملية التفريغ
ممتاز	///
جيد جدا	////
جيد	//// //// /
مقبول	//// //
ضعيف	////

المرحلة الثانية: **جدول (2-2): التوزيع التكراري للطلبة حسب التقديرات**

الصفة (التقدير)	عدد التكرارات
ممتاز	3
جيد جدا	4
جيد	11
مقبول	7
ضعيف	5
المجموع	30

يطلق على الجدول التكراري السابق، بالجدول التكراري البسيط المطلق

2-1 بالنسبة للمتغير الكمي: يمكن كذلك عرض بيانات متغير كمي في جدول تكراري بسيط مطلق بنفس الطريقة، ونميز هنا بين نوعين من المتغير الكمي :

1-2-1 حالة متغير كمي منفصل (متقطع): يتم عرض بيانات هذا المتغير في جدول تكراري بسيط مطلق من خلال وضع قيم المتغير الإحصائي المدروس مرتبة ترتيبا تصاعديا في العمود الأول، وعدد التكرارات في العمود الثاني.

مثال 2: البيانات التالية تمثل عدد العاملين في 42 ورشة صغيرة للخياطة

1	4	2	3	2	5	3
7	1	3	5	7	2	4
6	7	6	2	3	4	6
4	1	7	3	7	4	7
2	4	2	7	5	1	3
7	6	1	4	5	2	2

والمطلوب: عمل جدول التوزيع التكراري لهذه البيانات . واشرح كلا من n_1 و n_7 .

الحل: عدد العمال متغير كمي منفصل(متقطع)، ويمكن عرض البيانات أعلاه في جدول تكراري بسيط مطلق كما يلي:

جدول (2-3): توزيع الورشات حسب عدد العمال

عدد العمال (x_i قيم المتغير)	التفريغ	عدد الورشات (n_i التكرار)
1	𝍸	5
2	𝍸 ///	8
3	𝍸 /	6
4	𝍸 //	7
5	////	4
6	////	4
7	𝍸 ///	8
المجموع		**42**

الشرح :

- $n_1 = 5$: هناك 5 ورشات صغيرة من بين 42 ورشة يعمل في كل منها عامل واحد .
- $n_7 = 8$: هناك 8 ورشات صغيرة من بين 42 ورشة يعمل في كل منها 7 عمال .

1-2-2 حالة متغير كمي متصل (مستمر): عندما يكون المتغير كميا متصلا فإنه يحتوي على قيم كثيرة نسبيا، وبالتالي فإنه من الصعب على الباحث أن يخصص خانة لكل قيمة من قيم المتغيركما فعلنا في الأمثلة السابقة. لذلك يتم جمع القيم المتقاربة في عدد مناسب من الفئات بحيث لا يكون عدد الفئات صغيرا جدا فتضيع معالم التوزيع، ولا يكون كبيرا جدا فيفقد التوزيع قيمته وبالتالي يصعب إستقراؤه . فلا توجد قاعدة ثابتة لاختيار العدد المناسب للفئات، ولكن عادة ما يتراوح عدد الفئات بين 5 إلى 12 فئة تبعا لحجم العينة محل الدراسة [1].

و لعرض بيانات متغير كمي مستمر في جدول توزيع تكراري، نتبع الخطوات التالية:

أ- تحديد نوع البيانات.

ب- حساب المدى وهو الفرق بين أعلى قيمة و أصغر قيمة للبيانات، ويحسب من العلاقة :

المدى= أكبر قيمة – أصغر قيمة : $E = Xmax - Xmin$

ت- تحديد عدد الفئات k ، وأهم طريقة لحساب عدد الفئات نجد طريقة ستورجس (*sturges*) التي تنص على ما يلي:

$k = 1 + 1.322\,\mathrm{In}(n)$ أو $k = 1 + 3.322\log(n)$

حيث :

k تمثل عدد الفئات و n تمثل حجم العينة أو المجتمع .

ث- نحسب طول الفئة C من خلال العلاقة التالية:

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} \quad : \quad C = \frac{E}{K}$$

ويتم التقريب إلى الأعلى (العدد الصحيح التالي) عندما يكون الناتج عددا كسريا[2] .

ج- نكون الجدول التفريغي، بحيث نضع الفئات في العمود الأول ، والتكرارات في العمود الثاني، وتجدر الإشارة إلى أن الفئة الأولى لابد أن تبدأ بـــ أو تشمل أصغر قيمة، والفئة الأخيرة لابد أن تنتهي بـــ أو تشمل أكبر قيمة .

[1] – محمود محمد سليم صالح ، مرجع سابق، ص 31.

[2] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 33.

ملاحظات مهمة:

- كل فئة تتميز بحدين، حد أدنى وحد أقصى، هذه الحدود وبالخصوص الحد الأقصى يمكن أن يكون حدا فعليا أو غير فعلي، كأن نقول الفئة من أ إلى ب أي: [أ- ب] ← ب يعتبر حدا فعليا أي ينتمي للفئة، أو من أ إلى أقل من ب أي: [أ- ب[← ب يعتبر حدا غير فعلي أي لا ينتمي للفئة [1].
- مركز (منتصف) الفئة = الحد الأدنى للفئة + الحد الأعلى للفئة\2 :

$$X_i' = \frac{lim_{sup} + lim_{inf}}{2}$$

- إذا كانت أطوال الفئات متساوية فإن الفروق بين المراكز تساوي أطوال الفئات .

مثال 2: البيانات التالية تمثل التوزيع الطولي لـــ 50 تلميذا بالسنتمتر بإحدى المدارس الأساسية للموسم الدراسي 2024/2023 .

142	134	156	142	134	151	142	138	130	123
128	153	135	147	138	152	126	150	140	139
132	136	141	153	136	141	131	135	141	134
148	138	146	129	146	145	137	145	144	137
127	143	147	131	140	144	145	144	133	140

والمطلوب: تبويب البيانات في جدول توزيع تكراري . واشرح n_3 .

الحل: المتغير الاحصائي يعبر عن الأطوال بالسنتمتر، وهو متغير كمي متصل(مستمر)، ويمكن عرض البيانات أعلاه في جدول تكراري بسيط مطلق حسب الخطوات التالية:

- حساب المدى: $E = Xmax - Xmin \quad = 156 - 123 \quad = 33$
- حساب عدد الفئات : $k = 1 + 3.322\log(n) \; = 1 + 3.322\log(50) \; = 6.64 \; \cong \; 7$
- حساب طول الفئة : $C = \frac{E}{K} \; = \frac{33}{7} \; = 4.71 \; \cong \; 5$

نقوم بتفريغ البيانات في الفئات المذكورة فنحصل على الجدول التالي:

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 18.

جدول (2-4): توزيع عدد التلاميذ حسب الأطوال

أطوال التلاميذ Xi	عدد التلاميذ (التكرار) n_i
]123- 128]	3
]128- 133]	6
]133- 138]	10
]138- 143]	13
]143- 148]	11
]148- 153]	4
]153- 158]	3
المجموع	$\sum n_i=50$

الشرح:

n_3= 10 : هناك 10 تلاميذ من بين 50 تلميذا أطوالهم تتراوح بين 133 و أقل من 138 سنتيمتر.

2-الجداول التكرارية البسيطة النسبية و النسبية المئوية :يمكن حساب التكرار النسبي والتكرار النسبي المئوي للمتغيرات الإحصائية السابقة، وتصبح الجداول تسمى جدول تكراري بسيط نسبي [1]، أو جدول تكراري بسيط نسبي مئوي حسب الحالة، ويحسب التكرار النسبي و التكرار النسبي المئوي كما يلي:

2-1 التكرار النسبي: هو حاصل قسمة التكرار المطلق لكل صفة أو قيمة من قيم المتغير الإحصائي أو كل فئة حسب نوعية المتغير على مجموع التكرارات، أي : $f_i = \frac{n_i}{\sum n_i}$

2-2 التكرار النسبي المئوي: هو عبارة عن التكرار النسبي مضروبا في مائة، أي :$f_i\% = \frac{n_i}{\sum n_i} \times 100$

مثال 4 : بالعودة الى الأمثلة 01 و 02 و 03 أحسب التكرارات النسبية و التكرارات النسبية المئوية :

الحل :

- **بالنسبة للمتغير الوصفي (النوعي):** يصبح الجدول كما يلي :

جدول (2-5): التوزيع التكراري النسبي والمئوي للطلبة حسب التقديرات

الصفة (التقدير)	التكرار n_i	f_i	$f_i\%$

[1] - ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 41.

ممتاز	3	0.1	10
جيد جدا	4	0.1333	13.33
جيد	11	0.3667	36.67
مقبول	7	0.2333	23.33
ضعيف	5	0.1667	16.67
المجموع	$\sum n_i$=30	**1**	**100**

- **بالنسبة للمتغير كمي منفصل (متقطع):** يصبح الجدول كما يلي :

جدول (2-6): التوزيع التكراري النسبي والمئوي للورشات حسب عدد العمال

عدد العمال (قيم المتغير) Xi	عدد الورشات (التكرار) n_i	f_i	f_i%
1	5	0.119	11.9
2	8	0.1905	19.05
3	6	0.1428	14.28
4	7	0.1667	16.67
5	4	0.0952	9.52
6	4	0.0952	9.52
7	8	0.1905	19.05
المجموع	$\sum n_i$=42	**1**	**100**

الشرح :

- $f_1\% = 11.9\ \%$: هناك 11.9 % من الورشات الصغيرة يعمل فيها عامل واحد .
- $f_7\% = 19.05\ \%$: هناك 19.05 % من الورشات الصغيرة يعمل فيها 08 عمال .

- **بالنسبة للمتغير كمي متصل (مستمر):** يصبح الجدول كما يلي :

جدول (2-7): التوزيع التكراري النسبي والمئوي للتلاميذ حسب الأطوال

أطوال التلاميذ **Xi**	عدد التلاميذ (التكرار) n_i	f_i	f_i%
]128 -123]	3	0.06	6

]133 -128]	6	0.12	12
]138 -133]	10	0.20	20
]143 -138]	13	0.26	26
]148 -143]	11	0.22	22
]153 -148]	4	0.08	8
]158 -153]	3	0.06	6
المجموع	$\sum n_i=50$	**1**	**100**

الشرح:

- $f_3\% = 20\ \%$: هناك 20% من التلاميذ أطوالهم تتراوح بين 133 و أقل من 138 سنتيمتر .

3- الجداول التكرارية المتجمعة : من الجداول التكرارية المطلقة أو النسبية في الامثلة السابقة (03) و (04) من السهل باستخدام هذه الجداول أن نجيب بسهولة و يسر على سؤال مثلا حول عدد التلاميذ الذين تتراوح أطوالهم بين 133 و أقل من 138 سنتمتر أو نسبتهم المئوية، لكن ليس سهلا باستخدام نفس الجداول تحديد عدد التلاميذ الذين تزيد (أو تقل) أطوالهم عن 133 سم أو نسبتهم المئوية ، و لكن باستعمال الجداول التكرارية المتجمعة الصاعدة أو النازلة يمكننا بمجرد النظر لمثل هذه الجداول الاجابة على السؤال السابق .

3-1 جدول التوزيع التكراري التجميعي الصاعد: يمكن القول أن التكرار المتجمع الصاعد لقيمة من قيم المتغير أو لفئة معينة هو تكرار هذه القيمة أو الفئة مضافا إليه مجموع تكرارات القيم أو الفئات السابقة، و لتكوين الجدول التكراري التجميعي الصاعد نقوم بتجميع التكرارات في الجدول البسيط من بداية الجدول إلى نهايته، كما يمكن تكوين جدول تكراري تجميعي صاعد مطلق أو نسبي أو نسبي مئوي، بحيث:

- نرمز للتوزيع التكراري التجميعي الصاعد المطلق بالرمز $N_i^{\uparrow}$.
- يحسب التوزيع التكراري التجميعي الصاعد النسبي $F_i^{\uparrow}$ من خلال الصيغة : $F_i^{\uparrow} = \frac{N_i^{\uparrow}}{\sum n_i}$
- كما يحسب التوزيع التكراري التجميعي الصاعد النسبي المئوي% $F_i^{\uparrow}$ من خلال الصيغة :

$F_i^{\uparrow}\% = F_i^{\uparrow} \times 100$

ملاحظة :

التكرار المتجمع الصاعد المطلق الأول يكون دوما مساوي للتكرار المطلق الأول، كما أن التكرار المتجمع الصاعد المطلق الأخير يكون دوما مساوي لمجموع التكرارات .

2-3 جدول التوزيع التكراري التجميعي النازل : يمكن القول أن التكرار المتجمع النازل لقيمة من قيم المتغير أو لفئة معينة هو عبارة عن مجموع التكرارات مطروحا منه تكرارات القيم أو الفئات اللاحقة و لتكوين الجدول التكراري التجميعي النازل نقوم بطرح التكرارات في الجدول البسيط من نهاية الجدول إلى بدايته، كما يمكن تكوين جدول تكراري تجميعي نازل مطلق أو نسبي أو نسبي مئوي ، بحيث:

- نرمز للتوزيع التكراري التجميعي النازل المطلق بالرمز $N_i^{\downarrow}$.
- يحسب التوزيع التكراري التجميعي النازل النسبي $F_i^{\downarrow}$ من خلال الصيغة: $F_i^{\downarrow} = \frac{N_i^{\downarrow}}{\sum n_i}$
- كما يحسب التوزيع التكراري التجميعي النازل النسبي المئوي% $F_i^{\downarrow}$ من خلال الصيغة :

$$F_i^{\downarrow}\% = F_i^{\downarrow} \times 100$$

ملاحظة :

التكرار المتجمع االنازل المطلق الأول يكون دوما يكون مساوي لمجموع التكرارات ، كما أن التكرار المتجمع النازل الأخير يكون دوما مساوي للتكرار المطلق الأخير .

مثال 4 : بالعودة الى المثال 03 أحسب كلا من $N_i^{\uparrow}$، $N_i^{\downarrow}$، $F_i^{\uparrow}$، $F_i^{\downarrow}$، $F_i^{\uparrow}\%$ ، $F_i^{\downarrow}\%$ ، ثم أشرح كلا من :

$N_3^{\uparrow}$، $N_5^{\downarrow}$، $F_3^{\uparrow}\%$، $F_5^{\downarrow}\%$

الحل:

جدول (2-8): التوزيع التكراري التجميعي الصاعد و النازل (المطلق ، النسبي والمئوي) للتلاميذ حسب الأطوال

أطوال التلاميذ **Xi**	عدد التلاميذ **n_i** (التكرار)	$N_i^{\uparrow}$	$N_i^{\downarrow}$	$F_i^{\uparrow}$	$F_i^{\downarrow}$	$F_i^{\uparrow}\%$	$F_i^{\downarrow}\%$
[123- 128[	3	3	50	0.06	1	6	100
[128- 133[	6	9	47	0.18	0.94	18	94
[133- 138[	10	19	41	0.38	0.82	38	82
[138- 143[	13	32	31	0.64	0.62	64	62
[143- 148[	11	43	18	0.86	0.36	86	36
[148- 153[	4	47	7	0.94	0.14	94	14
[153- 158[	3	50	3	1	0.06	100	6
المجموع	$\sum n_i=50$	**1**	**100**	/	/	/	/

الشرح:

- $N_3^{\uparrow} = 19$: هناك 19 تلميذ من بين 50 تلميذ أطوالهم أقل من 138 سنتيمتر .
- $N_5^{\downarrow} = 18$: هناك 18 تلميذ من بين 50 تلميذ أطوالهم أكبر أو تساوي 143 سنتيمتر .
- $F_3^{\uparrow}\ \% = 38\ \%$: هناك 38 % من التلاميذ أطوالهم أقل من 138 سنتيمتر .
- $F_5^{\downarrow}\% = 36\ \%$: هناك 36 % من التلاميذ أطوالهم أكبر أو تساوي 143 سنتيمتر.

المحور الثالث: العرض البياني للبيانات

تمهيد:

إن العرض البياني يعتبر كوسيلة أخرى لتلخيص وعرض البيانات الإحصائية، وهي أسلوب أسهل استيعابا للقارئ العادي على عكس أسلوب العرض الجدولي، لأن القارئ غير المتخصص لايستطيع قراءة الجداول التكرارية لذا فإن الهدف من التمثيل البياني تبسيط البيانات وعرضها بطريقة مرئية تعطي فكرة عامة عن الظاهرة المدروسة وتسمح حتى بإستعمال المقارنات بين الظواهر في مختلف الأمكنة والأزمنة .

إن طرق عرض البيانات بالرسم البياني كثيرة ومتعددة نحاول عرضها كما يلي:

أولا- العرض البياني في حالة متغير وصفي (نوعي) :

يمكن عرض البيانات الخاصة بمتغير وصفي في عدة أشكال هي:

1- العرض الدائري: إذا كانت البيانات المتوفرة لدينا عبارة عن مجموع مقسم إلى أجزاء فيمكن تمثيل هذه البيانات بمساحة دائرة[1]، فكل جزء يقابل زاوية مركزية تتناسب مع التكرارات المقابلة لكل خاصية من الخواص المعنية بالدراسة ، وعادة ما نضيف عمودا إلى جدول المعطيات به الزاوية المركزية لكل تكرار.

وتحسب الزاوية المركزية بالعلاقة : $\textbf{الزاوية المركزية} = \frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360°$ ،

مع ملاحظة أن مجموع الزوايا = 360°

مثال 1 : البيانات التالية تمثل مجموعة من الزوارق السياحية المتجهة في جولة سياحية واحدة إلى جزيرة معزولة في الساعة العاشرة صباحا ، حيث يرمز A للزورق الذي سعته 20 سائحا ، B للزورق الذي سعته 23 سائحا ، C للزورق الذي سعته 30 سائحا ، D للزورق الذي سعته 28 سائحا، E للزورق الذي سعته 25 سائحا وفقا للجدول التالي:

المجموعة	A	B	C	D	E	المجموع
التكرارات	20	23	30	28	25	**126**

المطلوب: عرض البيانات باستخدام القطع الدائرية .

الحل: يتم أولا حساب الزاوية المركزية كما يلي:

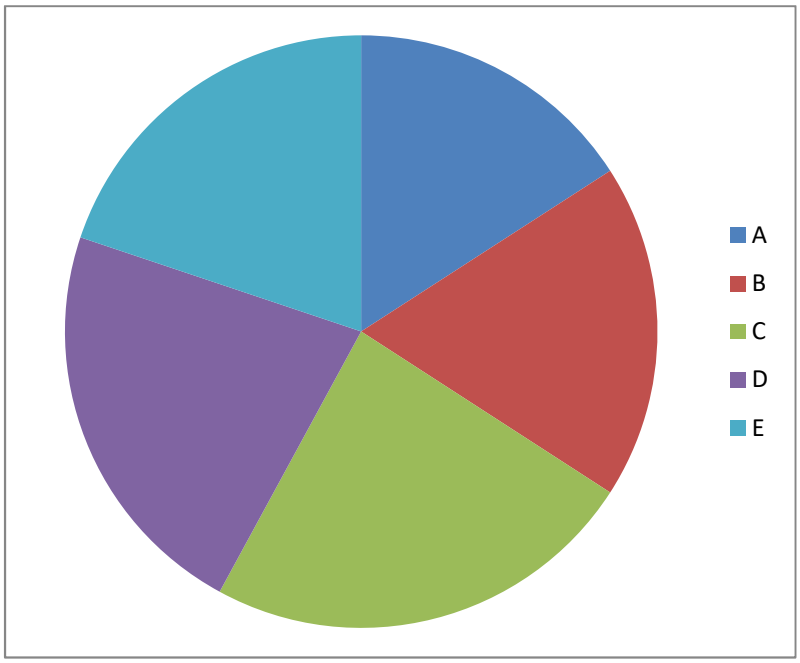
[1] – محمود محمد سليم صالح ، مرجع سابق، ص 43.

الجدول رقم (3-1) : يمثل توزيع عدد السياح حسب الزوارق السياحية

المجموعة	التكرارات	الزاوية المركزية
A	20	الزاوية المركزية $= \frac{20}{126} \times 360° = 57.14°$
B	23	الزاوية المركزية $= \frac{23}{126} \times 360° = 65.71°$
C	30	الزاوية المركزية $= \frac{30}{126} \times 360° = 85.71°$
D	28	الزاوية المركزية $= \frac{28}{126} \times 360° = 80°$
E	25	الزاوية المركزية $= \frac{25}{126} \times 360° = 71.43°$
المجموع	**126**	**360°**

إنطلاقا من الجدول السابق نرسم الدائرة كما يلي:

الشكل رقم (3-1): يمثل القطع الدائرية لتوزيع عدد السياح حسب الزوارق السياحية



2- الأعمدة المستطيلة: وهي عبارة عن مجموعة من الأعمدة التي تكون متجاورة (غير متلاصقة) و بنفس مسافة التباعد و بقواعد متساوية، كما أن إرتفاعها يتناسب مع تكرار الخاصية المقابل .

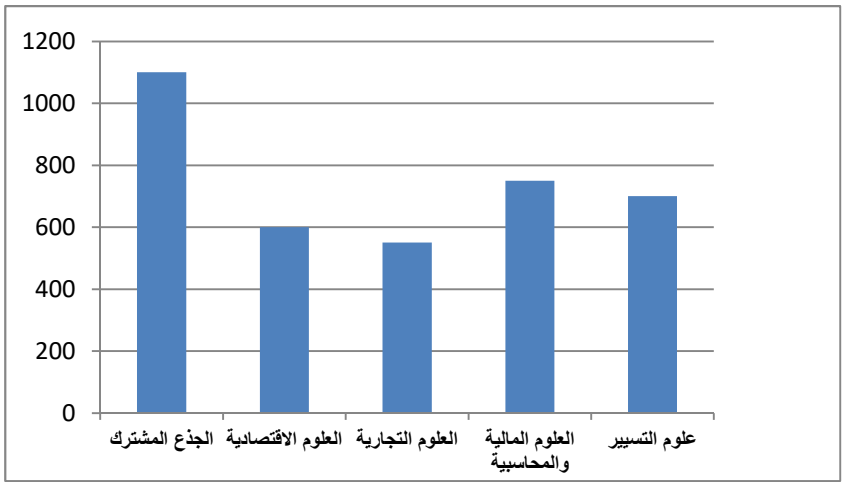
مثال 2 : الجدول التالي يمثل عدد الطلبة في أقسام كلية العلوم الاقتصادية وعلوم التسيير والعلوم التجارية بإحدى الجامعات

إسم القسم	الجذع المشترك	العلوم الاقتصادية	العلوم التجارية	العلوم المالية والمحاسبية	علوم التسيير	المجموع
عدد الطلبة	1100	600	550	750	700	**3700**

المطلوب: عرض البيانات باستخدام الأعمدة المستطيلة .

الحل: يتم تمثيل في شكل أعمدة مستطيلة، بوضع إسم أقسام الكلية على محور الفواصل، والتكرارات (عدد الطلبة) على محور التراتيب كما يلي:

الشكل رقم (3-2): يمثل أعمدة مستطيلة لتوزيع عدد طلبة الكلية حسب أقسام الكلية



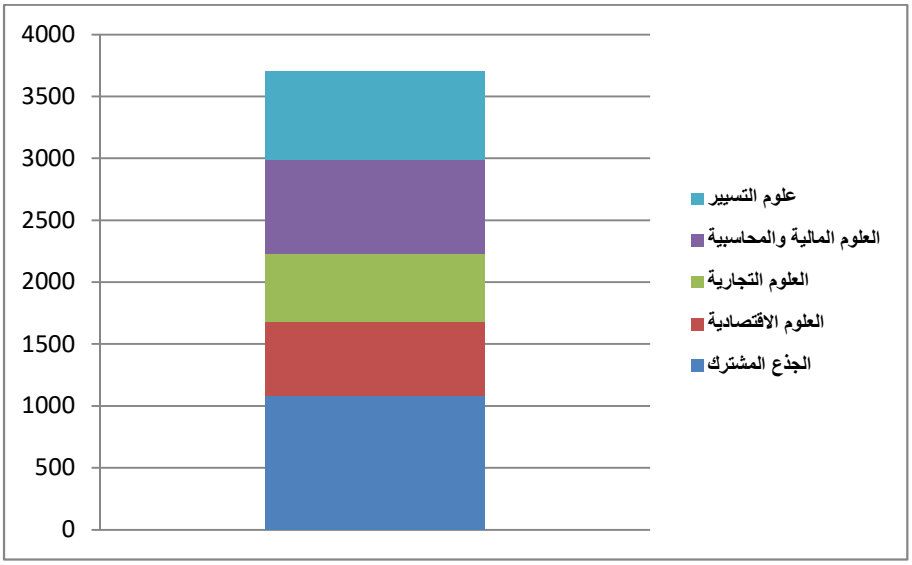
3- العمود المجزأ: وهو عبارة عن مستطيل يتكون من عدة أجزاء متراكمة فوق بعضها [1]، وكل جزء في العمود يتناسب مع تكراره.

مثال 3 : بعد أخذ نفس معطيات المثال رقم 02 السابق، قم بعرضها في عمود مجزأ

الحل: يتم التمثيل في شكل عمود مجزأ كما يلي :

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 40.

الشكل رقم (3-3) : يمثل عمود مجزأ لتوزيع عدد الطلبة حسب أقسام الكلية



ثانيا- العرض البياني في حالة متغير كمي منفصل (متقطع):

في هذا المتغير يمكن عرض الأعمدة البسيطة، بالإضافة إلى العرض البياني للتكرارات المتجمعة الصاعدة والنازلة.

1- الأعمدة البسيطة: يتم تمثيل المتغير الكمي المتقطع عن طريق الأعمدة (الأعمدة البسيطة) وذلك بتعيين النقاط المطلوبة على المحور الأفقي و إقامة الأعمدة من تلك النقط بما يتناسب أطوالها مع التكرارات المقابلة على المحور العمودي [1].

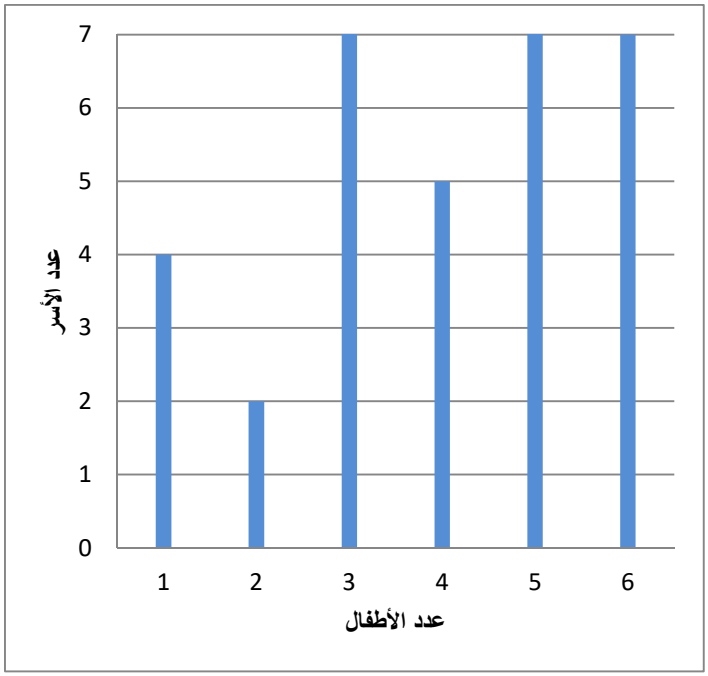
مثال 4 : يمثل الجدول التالي عدد الأطفال في عدد من الأسر في منطقة معينة

عدد الأطفال(قيم المتغير) X_i	عدد الأسر(التكرار) n_i
1	4
2	2
3	7
4	5
5	7
6	7
المجموع	$\sum n_i=32$

الحل: يتم التمثيل في شكل أعمدة بسيطة كما يلي :

[1] – نفس المرجع، ص 44.

الشكل رقم (3-4) : يمثل أعمدة بسيطة لتوزيع الأسر حسب عدد الأطفال



2- العرض البياني للتكرار التجميعي الصاعد والتكرار التجميعي النازل: العرض البياني هنا يكون على شكل منحنى سلمي ، فبالنسبة للتكرار التجميعي الصاعد يكون عبارة عن قطع مستقيمة متصاعدة تبعا لتصاعد التكرارات التجميعة الصاعدة المقابلة لكل قيمة من قيم المتغير الإحصائي المدروس، أما بالنسبة للتكرار التجميعي النازل فيكون عبارة عن قطع مستقيمة نازلة تبعا لتنازل التكرارات التجميعة النازلة المقابلة لكل قيمة من قيم المتغير الإحصائي المدروس .

مثال 5 : يمثل الجدول التالي توزيع عينة من الطلاب حسب عدد الغيابات خلال السداسي الثاني

عدد الغيابات (قيم المتغير) X_i	عدد الطلبة (التكرار) n_i
0	3
1	6
2	4
3	6
4	7
5	7
6	7
المجموع	$\sum n_i = 40$

المطلوب: تمثيل التكرار التجميعي الصاعد والنازل .

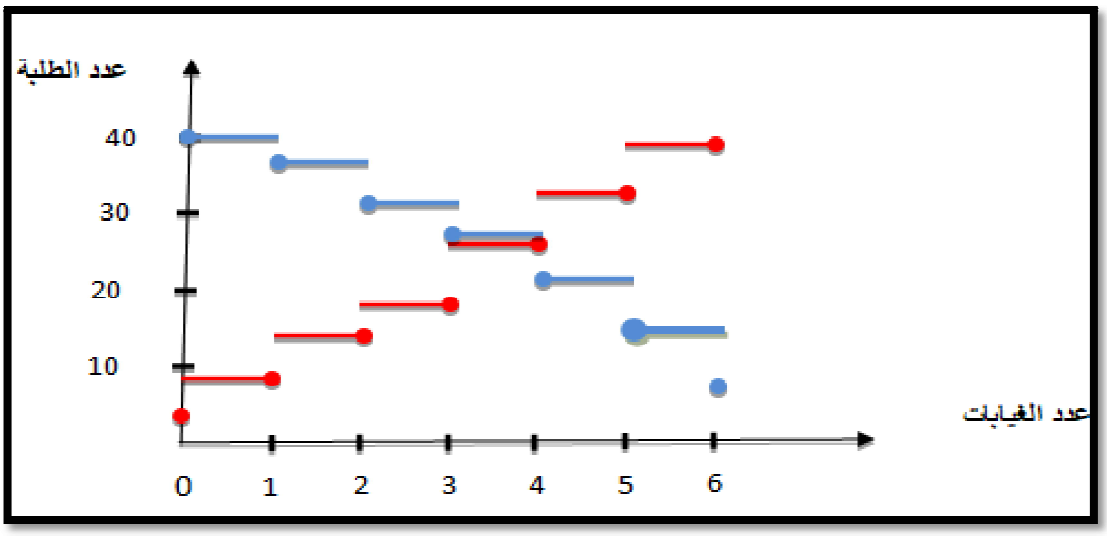
الحل: يتم أولا حساب التكرارات التجميعية الصاعدة والنازلة كما يلي :

الجدول رقم (3-2): التوزيع التكراري التجميعي الصاعد و النازل (المطلق) للبلديات حسب عدد المدارس

عدد المدارس (قيم المتغير) X_i	عدد البلديات (التكرار) n_i	التكرار التجميعي الصاعد $N_i^{\uparrow}$	التكرار التجميعي النازل $N_i^{\downarrow}$
0	3	3	40
1	6	9	37
2	4	13	31
3	6	19	27
4	7	26	21
5	7	33	14
6	7	40	7
المجموع	$\sum n_i = 40$	/	/

إنطلاقا من الجدول السابق نرسم التكرارات التجميعية الصاعدة والتكرارات التجميعية النازلة كما يلي:

الشكل رقم (3-5): يمثل عرض بياني على شكل منحنى سلمي للتكرارات التجميعية الصاعدة و النازلة لعينة من الطلاب حسب عدد الغيابات



ثالثا- العرض البياني في حالة متغير كمي متصل (مستمر):

في هذا المتغير يمكن عرض المدرج التكراري، المضلع التكراري، المنحنى التكراري بالإضافة إلى المنحنى التجميعي الصاعد والنازل .

1- المدرج التكراري: هو عبارة عن مجموعة من المستطيلات المتلاصقة قواعدها هي أطوال الفئات و إرتفاعها يمثل التكرارات المقابلة [1]. أي أننا نأخذ محورين متعامدين، ونرصد على المحور الأفقي الحدود الفعلية لكل فئة من فئات التوزيع التكراري ونقيم على كل فئة مستطيلا يتناسب طول إرتفاعه مع تكرار تلك الفئة [2]. وعند رسم المدرج التكراري نميز بين حالتين :

1-1 فئات الطول متساوية: في هذه الحالة فإننا لا نجري أي تعديلات حيث نمثل الفئة بمستطيل حدود قاعدته الحدود الفعلية لتلك الفئة وطول إرتفاعه يتناسب مع تكرارها الموجود في جدول المعطيات.

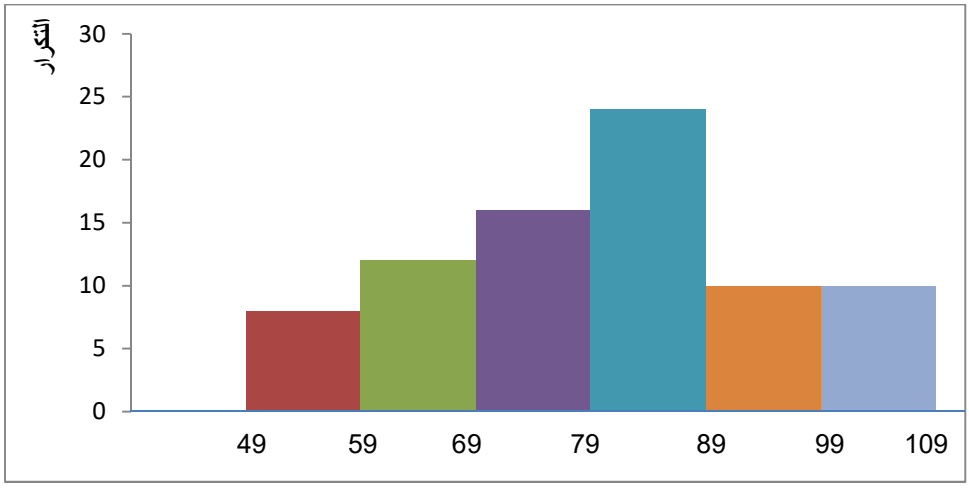
مثال 06 : الجدول التالي يبين توزيع عينة من 80 عامل في إحدى المؤسسات الاقتصادية حسب فئات الدخل الشهري (بالآلاف دينار جزائري):

فئات الأجور	[49- 59[	[59- 69[	[69- 79[	[79- 89[	[89- 99[	[99- 109[	المجموع
عدد العمال	8	12	16	24	10	10	**80**

المطلوب: رسم المدرج التكراري .

الحل: يتم تمثيل المدرج التكراري كمايلي :

الشكل رقم (3- 6): يمثل مدرج تكراري لتوزيع عدد العمال حسب فئات الدخل



[1] – نفس المرجع، ص 45.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص 65.

2-1 فئات الطول غير متساوية: في هذه الحالة فإن قاعدة المقارنة غير ثابتة فنقوم بتعديل التكرارات حتى يكون هناك تناسب بين طول الفئة والتكرار المقابل لها، وهو ما يجعل المدرج متزنا من حيث الشكل .

ونقوم بحساب التكرار المعدل من خلال الصيغة التالية: **التكرار المعدل** = $\frac{\text{تكرار الفئة}}{\text{طول الفئة}} \times$ **الطول الشائع** أي [1]:

$$n'_i = \frac{n_i}{a_i} \times a$$

مثال 07 : الجدول التالي يبين توزيع عينة من 80 عامل في إحدى المؤسسات حسب الأجر اليومي :

فئات الاجر	]110 -100]	]130 -110]	]135 -130]	]150 -135]	]200 -150]	]205 -200]	المجموع
عدد العمال	5	15	20	22	5	13	**80**

المطلوب: رسم المدرج التكراري .

الحل: نلاحظ أن فئات الطول غير متساوية و الطول الشائع هو : (a = 5) ، لذا يتم أولا تعديل التكرارات كما يلي:

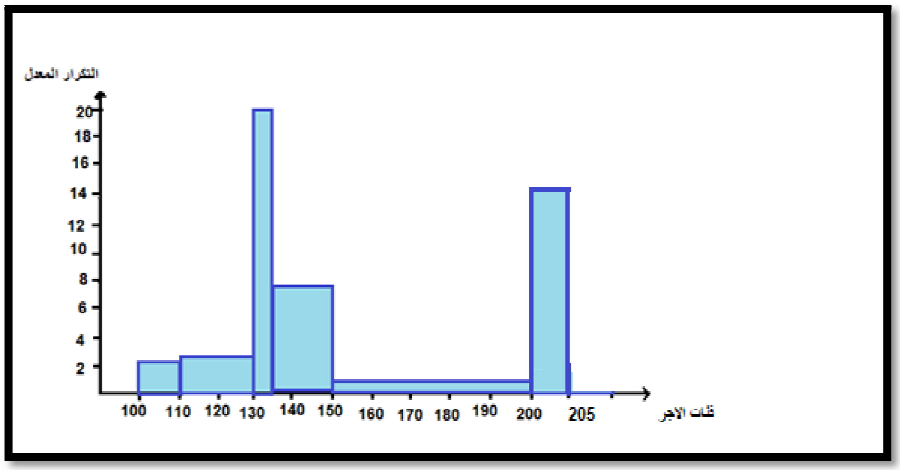
الجدول رقم (3-3) :التوزيع التكراري (المعدل) للعمال حسب فئات الأجر

فئات الاجر	عدد العمال (التكرار) n_i	طول الفئة a_i	$\frac{n_i}{a_i}$	$n'_i = \frac{n_i}{a_i} \times a$
]110 -100]	5	10	0.5	2.5
]130 -110]	15	20	0.75	3.75
]135 -130]	20	5	4	20
]150 -135]	22	15	1.47	7.35
]200 -150]	5	50	0.1	0.5
]205 -200]	13	5	2.6	13
المجموع	$\sum n_i = 80$	/	/	/

إنطلاقا من الجدول السابق نقوم برسم المدرج التكراري كما يلي:

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 23.

الشكل رقم (3-7): يمثل مدرج تكراري لتوزيع عدد العمال حسب فئات الأجر



ملاحظة: يتم اللجوء إلى تعديل التكرارات لحالة الفئات غير المتساوية في الحالتين التاليتين [1]:

- عند رسم المدرج التكراري .
- عند تحديد الفئة المنوالية وحساب المنوال .

2 – المضلع التكراري: هو مضلع مغلق نحصل عليه بتصنيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم نوصل هذه النقاط ببعضها البعض[2] ، ولكي نحافظ على نفس المساحة أسفل المدرج التكراري نقوم بغلق الخط المنكسر الذي حصلنا عليه بحيث نعتبر أن هناك فئتين إفتراضيتين واحدة إلى أقصى اليسار والثانية إلى أقصى اليمين وتكرار كل منهما يساوي صفر أي أن إرتفاعهما معدوم ، نأخذ مركز هاتين الفئتين الإفتراضيتين ونغلق المضلع.

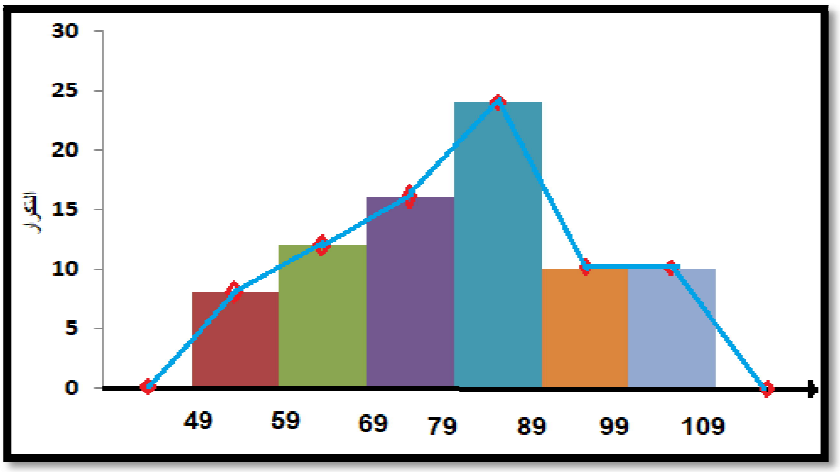
مثال 08 : مثل معطيات المثال 06 الخاص بتوزيع عينة من 80 عامل في إحدى المؤسسات الاقتصادية حسب فئات الدخل الشهري (بالآلاف دينار جزائري) في مضلع تكراري

الحل: يتم رسم المضلع التكراري في نفس الشكل الخاص بالمدرج التكراري كما يلي :

[1] – جلاطو جيلالي ، مرجع سابق، ص 24.

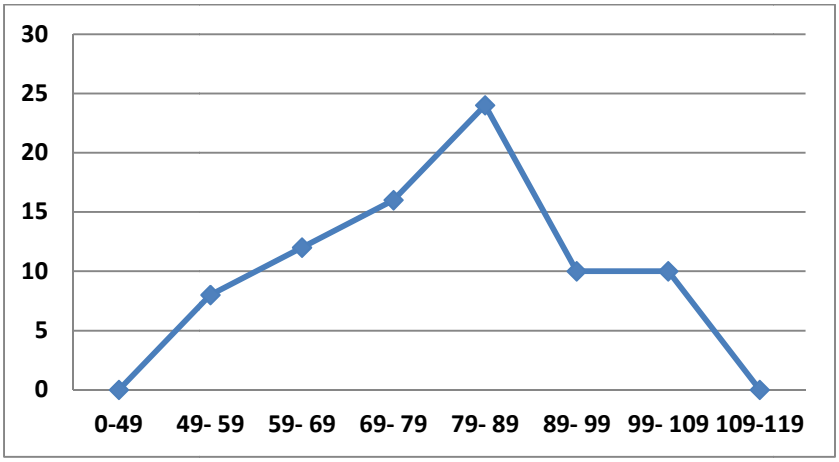
[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص 66.

الشكل رقم (3-8): يمثل مضلع تكراري و مدرج تكراري لتوزيع عدد العمال حسب فئات الدخل



وتجدر الإشارة إلى أنه يمكن رسم المضلع التكراري دون الإستعانة بالمدرج التكراري و ذلك بتعيين النقاط التي تمثل رؤوس المضلع ، بحيث تكون إحداثيات كل منها(مركز الفئة، التكرار المقابل) ثم نصل هذه النقاط بقطع مستقيمة و من ثم غلق المضلع التكراري لكي تكون المساحة أسفل المضلع التكراري مساوية لمساحات المستطيلات في المدرج التكراري و ذلك بإضافة فئة سابقة و فئة لاحقة و تكرار كل منها يساوي الصفر كما أشرنا سابقا. وهذا ما يوضحه الشكل التالي:

الشكل رقم (3-9): يمثل مضلع تكراري لتوزيع عدد العمال حسب فئات الدخل



3 – المنحنى التكراري: إذا مهدنا المضلع التكراري وجعلناه منحنى بدلا من خطوط منكسرة فإننا نحصل على المنحنى التكراري [1]، ويرسم المنحنى التكراري على محورين متعامدين ، حيث تكون مراكز الفئات في المحور الأفقي والتكرارات المقابلة تكون على المحور العمودي، و لكي تكون المساحة أسفل المدرج التكراري كما هو الشأن في المضلع التكراري مساوية لمساحات المستطيلات في المدرج التكراري نضيف كما أشرنا

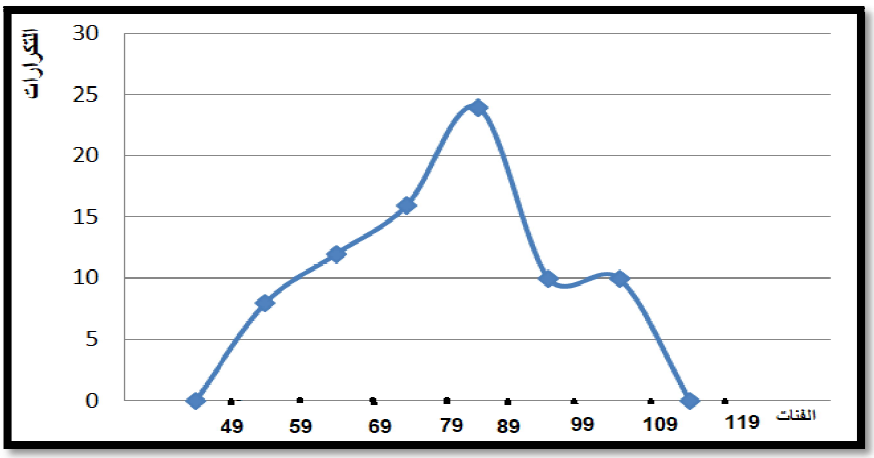
[1] – نفس المرجع ، ص 67.

سابقا فئة سابقة وفئة لاحقة و تكرار كل منها يساوي الصفر، ثم نقوم من خلال الرسم باليد المرور بالإحداثيات المعينة و أحيانا للحفاظ على شكل المنحنى الإنسيابي لا تلزمنا الإحداثيات كلها (ليس شرط المرور على جميع النقاط) .

مثال 09 : مثل معطيات المثال 06 الخاص بتوزيع عينة من 80 عامل في إحدى المؤسسات الاقتصادية حسب فئات الدخل الشهرى (بالآلاف دينار جزائري) في منحنى تكراري.

الحل: يتم رسم المدرج التكراري كما يلي :

الشكل رقم (3-10): يمثل منحنى تكراري لتوزيع عدد العمال حسب فئات الدخل



ملاحظة : يمكن رسم المدرج التكراري والمضلع التكراري والمنحنى التكراري على نفس الشكل .

4-المنحنى التجميعي الصاعد: يمكن تمثيل الجدول التكراري التجميعي الصاعد بواسطة منحنى التجميعي الصاعد، ويرسم المنحنى التجميعي الصاعد على محورين متعامدين ، حيث تكون إحداثيات كل نقطة الحدود العليا للفئات في المحور الأفقي والتكرارات التجميعية الصاعدة المقابلة لها في المحور العمودي ثم يتم إيصال النقاط ببعضها البعض .

5-المنحنى التجميعي النازل: يمكن تمثيل الجدول التكراري التجميعي النازل بواسطة منحنى التجميعي النازل، ويرسم المنحني التجميعي النازل على محورين متعامدين ، حيث تكون إحداثيات كل نقطة الحدود الدنيا للفئات في المحور الأفقي والتكرارات التجميعية النازلة المقابلة لها في المحور العمودي ثم يتم إيصال النقاط ببعضها البعض .

مثال 10 : بعد تكوين جدول التكرار التجميعي الصاعد والنازل مثل معطيات المثال 06 الخاص بتوزيع عينة من 80 عامل في إحدى المؤسسات الاقتصادية حسب فئات الدخل الشهرى (بالآلاف دينار جزائري) في منحنى التكرار التجميعي الصاعد و النازل .

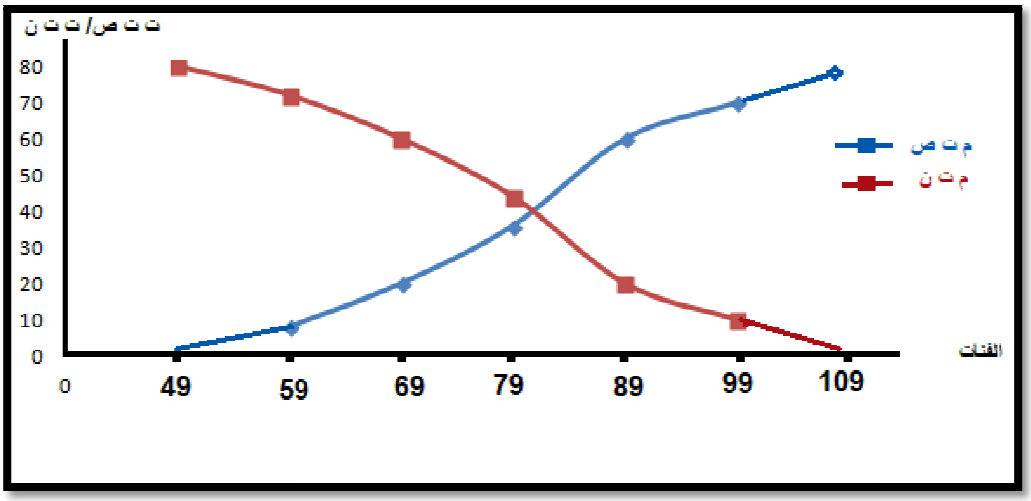
الحل: يتم أولا حساب التكرارات التجميعية الصاعدة والنازلة في الجدول التالي :

الجدول رقم (3-4) : التوزيع التكراري التجميعي الصاعد و النازل (المطلق) للعمال حسب فئات الأجر

فئات الأجور	عدد العمال (التكرار) n_i	التكرار التجميعي الصاعد $N_i^{\uparrow}$	التكرار التجميعي النازل $N_i^{\downarrow}$
]59 -49]	8	8	80
]69 -59]	12	20	72
]79 -69]	16	36	60
]89 -79]	24	60	44
]99 -89]	10	70	20
]109 -99]	10	80	10
المجموع	$\sum n_i = 80$	/	/

يتم رسم المنحنى التجميعي الصاعد والنازل في الشكل التالي:

الشكل رقم (3-9): يمثل منحنى التجميعي الصاعد والنازل لتوزيع عدد العمال حسب فئات الدخل



المحور الرابع: مقاييس النزعة المركزية

تمهيد:

رأينا في المحورين السابقين كيفية عرض البيانات الإحصائية جدوليا وبيانيا، حيث تساهم هذه العروض في إعطاء فكرة أولية و وصف عام عن الظاهرة المدروسة، و لكن هذا العرض الجدولي أو البياني غير كافي لوصف البيانات خاصة للباحث الإحصائي ففي كثير من المرات يكون في حاجة إلى حساب بعض المؤشرات العددية التي تمكنه من وصف الظاهرة المدروسة من حيث التركز للمشاهدات حول قيمة متوسطة أو معرفة مدى تجانس القيم من عدمها ...الخ، لذا وضعت مقاييس عددية وصفية تمكن الباحث عند استخدامها من القدرة على التحليل والتنبؤ واتخاذ القرارت، ونجد من بين هذه المقاييس ، مقاييس النزعة المركزية .

يشير مفهوم مقاييس النزعة المركزية إلى ميل البيانات للتمركز حول قيمة ممثلة أو نموذجية في التوزيع [1]. فعند التمعن في الظواهر المحيطة بنا نجد أن معظمها لديها قيم متقاربة حول بعضها و تتركز حول قيمة محددة متوسطة إذا ابتعدنا عنها فإن عدد المعلومات يبدأ في التناقص، فإذا أخذنا معدلات الطلبة مثلا أو ذكائهم أو أطوالهم ...الخ، نجد أن معظم القيم تتركز قرب بعضها وتنـــزع حول قيمة متوسطة ، لذا يطلق على هذه الظاهرة بالنزعة المركزية، والتي لها مجموعة من المقاييس هي :

أولا-الوسط الحسابي:

يعتبر من أهم مقاييس النزعة المركزية وأكثرها استخداما في النواحي التطبيقية ، ويمكن حسابه للبيانات المبوبة وغير المبوبة كما يلي:

1- الوسط الحسابي للبيانات غير المبوبة: يمكن إيجاد الوسط الحسابي من خلال الطريقتين التاليتين:

1-1 الطريقة المباشرة: تمكننا من حساب الوسط الحسابي بالطريقة البسيطة، حيث يعرف الوسط الحسابي هنا لمجموعة من القيم بأنه مجموع هذه القيم مقسوما على عددها، ويرمز له بالرمز $\bar{X}$.

فإذا كان لدينا n من القيم ولتكن : $x_1, x_2, x_3, \ldots, x_n$ فإن الوسط الحسابي لهذه القيم يحسب بالعلاقة التالية [2]:

$$\bar{X} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

حيث : X_i : تمثل قيم المتغير الإحصائي، و n : تمثل عدد القيم .

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 55.

[2] – شرف الدين خليل ، مرجع سابق، ص 31.

مثال 01: أوجد الوسط الحسابي للقيم التالية: 7، 15، 5، 0، 17، 22، 2، 12.

الحل:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{10} X_i}{n} = \frac{7+15+5+0+17+22+2+12}{8} = \frac{80}{8} = 10$$

1-2 طريقة الانحرافات البسيطة (الوسط الفرضي): تعتمد هذه الطريقة على مبدأ الإنحرافات (إنحراف القيم عن وسط فرضي) و الحكمة المتوخاة من وراء استخدام هذه الطريقة تسهيل عملية الحساب خصوصا عندما تكون قيم المشاهدات كبيرة [1]. كما تجنبنا احتمال الوقوع في الأخطاء لأنها تختصر علينا الحسابات، و بالتالي سهولة حساب الوسط الحسابي، حيث يتم حسابه في هذه الحالة باتباع الخطوات التالية:

- نختار عشوائيا وسطا فرضيا من بين القيم أو يكون القيمة الوسطى .
- نحسب الإنحرافات البسيطة بين القيم الموجودة والوسط الفرضي من خلال: $d_i = x_i - x_0$
- نحسب الوسط الحسابي من العلاقة : $\bar{X} = x_0 + \frac{\sum d_i}{n}$

حيث : x_0 : تمثل الوسط الفرضي، و $\sum d_i$: مجموع الإنحرافات ، و n : يمثل عدد القيم .

مثال 02 : أحسب الوسط الحسابي للقيم التالية : 10، 9، 21، 14، 5، 13، 7 عن طريق الوسط الفرضي .

الحل:

نفرض أن : $x_0 = 14$

وبالتالي :

$$\sum d_i = \sum(x_i - x_0) = (10-14) + (9-14) + (21-14) + (14-14) + (5-14) + (13-14) + (7-14) = -19$$

ومنه : $\bar{X} = 14 + \frac{-19}{7} = 11.28$

ونحصل على نفس النتيجة إذا طبقنا العلاقة : $\bar{X} = \frac{79}{7} = 11.28$

2- الوسط الحسابي للبيانات المبوبة: ونميز بين حالتين :

1-2 الوسط الحسابي في حالة متغير كمي منفصل: و يمكن إيجاده من خلال الطريقتين التاليتين:

1-1-2 الطريقة المباشرة: الوسط الحسابي هنا يكون عبارة عن مجموع ضرب قيم الظاهرة بالتكرارات المقابلة لها مقسوما على مجموع هذه التكرارات.

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 61.

فإذا كان لدينا مجموعة من القيم ولتكن : $x_1 , x_2 , x_3 , \dots . , x_k$ وتكرارها هي: $n_1 , n_2 , n_3 , \dots . , n_k$ ، فإن الوسط الحسابي لهذه القيم يحسب بالعلاقة التالية [1]:

$$\bar{X} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + \cdots + n_kx_k}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_iX_i}{\sum_{i=1}^{k} n_i}$$

حيث : X_i : تمثل قيم المتغير الاحصائي المتقطع، وn_i : تمثل التكرارات المطلقة المقابلة لها ، و $\sum_{i=1}^{k} n_i$: يمثل مجموع التكرارت.

مثال 03: البيانات التالية تمثل توزيع عينة من الطلاب حسب عدد الغيابات خلال السداسي الثاني

عدد الغيابات (قيم المتغير) X_i	0	1	2	3	4	5	المجموع
عدد الطلبة (التكرار) n_i	6	4	3	8	2	7	**30**

المطلوب: حساب الوسط الحسابي لهذه البيانات .

الحل:

جدول رقم (4-1): توزيع عينة من الطلاب حسب عدد الغيابات خلال السداسي الثاني

عدد الغيابات (قيم المتغير) X_i	عدد الطلبة (التكرار) n_i	$X_i \times n_i$
0	6	0
1	4	4
2	3	6
3	8	24
4	2	8
5	7	35
المجموع	**30**	**77**

$$\bar{X} = \frac{\sum_{i=1}^{k} n_iX_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_iX_i}{30} = \frac{77}{30} = 2.56$$

هذا يعني : متوسط عدد غيابات الطلبة خلال السداسي الثاني هو :2.56

2-1-2 طريقة الانحرافات البسيطة (الوسط الفرضي): يتم حساب الوسط الحسابي في هذه الحالة باتباع نفس الخطوات السابقة التي ذكرناها في الوسط الحسابي لحالة البيانات غير المبوبة ، ولكن باستخدام العلاقة :

$$\bar{X} = x_0 + \frac{\sum_{i=1}^{k} n_id_i}{\sum_{i=1}^{k} n_i}$$

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 30.

مثال 04 : من خلال بيانات الجدول التالية أحسب الوسط الحسابي عن طريق الوسط الفرضي، بافتراض : $x_0 = 3$

قيم المتغير X_i	1	2	3	4	5	المجموع
التكرارات n_i	5	7	13	11	4	**40**

الحل:

الجدول رقم (4-2) : توزيع التكرارات حسب قيم المتغير

قيم المتغير X_i	التكرارات n_i	$d_i = x_i - 3$	$d_i \times n_i$
1	5	-2	-10
2	7	-1	-7
3	13	0	0
4	11	1	11
5	4	2	8
المجموع	**40**	/	**02**

$$\bar{X} = x_0 + \frac{\sum_{i=1}^{k} n_i d_i}{\sum_{i=1}^{k} n_i} = 3 + \frac{\sum_{i=1}^{5} n_i d_i}{\sum_{i=1}^{5} n_i} = 3 + \frac{2}{40} = 3.05$$

2-2 الوسط الحسابي في حالة متغير كمي متصل: و يمكن إيجاده من خلال الطريقتين التاليتين أيضا:

2-2-1 الطريقة المباشرة: الوسط الحسابي هنا يحسب بنفس الصيغة السابقة للمتغير الكمي المنفصل. غير أنه هنا لابد من تحديد مراكز الفئات [1]، لأن قيم المتغير الإحصائي للظاهرة المدروسة معطاة على شكل فئات (مجالات) . وبالتالي يحسب بالعلاقة التالية :

$$\bar{X} = \frac{n_1 x'_1 + n_2 x'_2 + n_3 x'_3 + \cdots + n_k x'_k}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_i X'_i}{\sum_{i=1}^{k} n_i}$$

حيث : x'_i : تمثل مراكز الفئات، و n_i : تمثل التكرارات المطلقة المقابلة لها ، و $\sum_{i=1}^{k} n_i$: يمثل مجموع التكرارات.

مثال 05 : الجدول التالي يبين توزيع علامات (درجات) 50 طالب في مادة الإحصاء بقسم العلوم الاقتصادية

[1] – محمود محمد سليم صالح ، مرجع سابق، ص 56.

الفئات	]4 -0[	]8 -4[	]12 -8[	]16 -12[	]20 -16[	المجموع
التكرارات	5	10	15	16	4	**50**

المطلوب: حساب الوسط الحسابي لهذه البيانات .

الحل:

الجدول رقم (4-3) : توزيع علامات 50 طالب في مادة الاحصاء

العلامات(الفئات)	عدد الطلبة (التكرار) n_i	مراكز الفئات x'_i	$x'_i \times n_i$
]4 -0[	5	2	10
]8 -4[	10	6	60
]12 -8[	15	10	150
]16 -12[	16	14	224
]20 -16[	4	18	72
المجموع	**50**	/	**516**

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i x'_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i x'_i}{50} = \frac{516}{50} = 10.32$$

هذا يعني : متوسط علامات الطلبة في مادة الإحصاء هو :10.32

2-2-2 طريقة الانحرافات البسيطة (الوسط الفرضي): يتم حساب الوسط الحسابي في هذه الحالة بإتباع نفس الخطوات السابقة، التي ذكرناها في الوسط الحسابي لحالة البيانات غير المبوبة ، فقط أن الوسط الفرضي نختاره من مراكز الفئات و يستحسن أن يكون مركز الفئة الذي يقابل أكبر تكرار، و نستخدم العلاقة التالية :

$$\bar{X} = x_0 + \frac{\sum_{i=1}^{k} n_i d_i}{\sum_{i=1}^{k} n_i}$$

مع ملاحظة أن : $d_i = x'_i - x_0$

مثال 06 : أحسب الوسط الحسابي عن طريق الوسط الفرضي باستخدام معطيات المثال السابق (05).

الحل:

جدول رقم (4-4): توزيع علامات 50 طالب في مادة الاحصاء

العلامات(الفئات)	عدد االطلبة (التكرار) n_i	مراكز الفئات x'_i	$d_i = x'_i - 18$	$d_i \times n_i$
]4 -0]	5	2	-16	-80
]8 -4]	10	6	-12	-120
]12 -8]	15	10	-8	-120
]16 -12]	16	14	-4	-64
]20 -16]	4	18	0	0
المجموع	**50**	/	/	**-384**

$$\bar{X} = x_0 + \frac{\sum_{i=1}^{k} n_i d_i}{\sum_{i=1}^{k} n_i} = 18 + \frac{\sum_{i=1}^{5} n_i d_i}{\sum_{i=1}^{5} n_i} = 18 + \frac{-384}{50} = 10.32$$

03- خصائص الوسط الحسابي : للوسط الحسابي مجموعة من الخصائص هي [1]:

- مجموع إنحرافات مجموعة من القيم عن وسطها الحسابي يساوي صفر
  - ✓ حالة البيانات غير المبوبة: $\sum_{i=1}^{n}(X_i - \bar{X}) = 0$
  - ✓ حالة البيانات غير المبوبة: $\sum_{i=1}^{n} n_i(X_i - \bar{X}) = 0$
- إذا كان لدينا عدد من القيم فإن مجموع مربع إنحرافات القيم عن وسطها الحسابي هو أقل من مجموع مربع إنحرافاتها عن أي وسط فرضي آخر أي : $\sum_{i=1}^{n}(X_i - \bar{X})^2 < \sum_{i=1}^{n}(X_i - X_0)^2$
- الوسط الحسابي يتأثر بالقيم الشاذة أو المتطرفة .
- يتميز الوسط الحسابي بسهولة فهمه وحسابه و إدراك معناه مقارنة بغيره من مقاييس النزعة المركزية الأخرى.
- يستعمل الوسط الحسابي في حالة المتغيرات الكمية .
- لا يمكن حساب المتوسط الحسابي في حالة البيانات المفتوحة (أقل من ، أكبر من) [2].
- إذا كانت لدينا مجموعة من القيم وسطها الحسابي، وضربت كل قيمة من تلك القيم في عدد ثابت وليكن C فإن الوسط الحسابي للقيم الجديدة، يساوي إلى الوسط الحسابي للقيم الأصلية مضروبا في المقدار C ، أي أن : $\bar{X} = \bar{X}.C$ [1].

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 64.

[2] – عبد الحميد قطوش ، مرجع سابق ، ص 33.

ثانيا–الوسيط:

يعتبر الوسيط من مقاييس النزعة المركزية، فهو مقياس موضعي لفصله مجموعة من البيانات إلى قسمين متساويين ، ففي كثير من المرات يكون إهتمام الباحث معرفة القيمة التي يكون عدد الأفراد الحاصلين على قيمة أقل منها مساويا لعدد الحاصلين على قيمة أعلى منها ، وهو ما يجيبنا عليها الوسيط الذي تختلف عملية حسابه بين البيانات المبوبة وغير المبوبة كما يلي:

1– الوسيط للبيانات غير المبوبة: الوسيط لمجموعة من القيم المرتبة ترتيبا تصاعديا أو ترتيبا تنازليا هو العدد الأوسط منها إذ كان فرديا، والوسط الحسابي للعددين الأوسطين في المجموعة إذا كان عددها زوجيا [2].

أي أنه لحساب الوسيط هنا نتبع الخطوات التالية:

- نرتب البيانات ترتيبا تصاعديا أو تنازليا .
- إذا كان عدد البيانات n عددا فرديا فإن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$ ، أي أن :

$$M_e = X_{\left(\frac{n+1}{2}\right)}$$

- إذا كان عدد البيانات n عددا زوجيا فإن الوسيط هو المتوسط الحسابي للقيمة التي ترتيبها $\frac{n}{2}$ و القيمة التي ترتيبها $\frac{n}{2}+1$ ، أي أن : $M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}$

مثال 07: أوجد الوسيط لبيانات السلسلتين A و B التاليتين :

A:6، 2، 7، 8، 1، 5، 3

B:12، 10، 9، 13، 8، 11

الحل:

–السلسلة A:

نرتب البيانات ترتيبا تصاعديا : 1، 2، 3، 5، 6، 7، 8

بما أن عدد البيانات فردي فإن الوسيط هو : $M_e = X_{\left(\frac{n+1}{2}\right)} = X_{\left(\frac{7+1}{2}\right)} = X_{(4)} = 5$

إذن الوسيط هو : $M_e = 5$

–السلسلة B:

نرتب البيانات ترتيبا تصاعديا : 8، 9، 10، 11، 12، 13

[1] – عبد الحكيم بيصار، محاضرات في الاحصاء 1، مطبوعة دروس مقدمة الى طلبة الجذع المشترك علوم اقتصادية، التسيير والعلوم التجارية، جامعة المسيلة، 2019–2020 ، ص 26.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص 66.

بما أن عدد البيانات زوجي فإن الوسيط هو :

$$M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} = \frac{X_{\left(\frac{6}{2}\right)} + X_{\left(\frac{6}{2}+1\right)}}{2} = \frac{X_{(3)} + X_{(4)}}{2} = \frac{10+11}{2} = 10.5$$

إذن الوسيط هو : $M_e = 10.5$

2- الوسيط للبيانات المبوبة: ونميز بين حالتين :

2-1 الوسيط في حالة متغير كمي منفصل: و يمكن إيجاده من خلال إتباع نفس طريقة الوسيط للبيانات غير المبوبة، أو يمكن أن نستعمل طريقة أخرى باتباع الخطوات التالية [1]:

- نحسب التكرار المتجمع الصاعد للقيم.
- نحدد رتبة الوسيط: $\frac{n}{2}$ / (ملاحظة : n تمثل مجموع التكرارات)
- نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر من أو يساوي $\frac{n}{2}$ مباشرة وهي القيمة التي تمثل الوسيط .

مثال 08: البيانات الموجودة في الجدول التالي هي لعدد الأطفال في عدد من الأسر في منطقة ما

عدد الأطفال (قيم المتغير) X_i	0	1	2	3	4	5	المجموع
عدد الأسر (التكرار) n_i	3	8	10	3	4	2	**30**

المطلوب: حساب الوسيط لهذه البيانات .

الحل: نحسب أولا التكرار التجميعي الصاعد $N_i^{\uparrow}$

جدول رقم (4-5): توزيع الأسر حسب عدد الأطفال

عدد الأطفال (قيم المتغير) X_i	عدد الأسر (التكرار) n_i	التكرار التجميعي الصاعد $N_i^{\uparrow}$
0	3	3
1	8	11
2	10	21
3	3	24
4	4	28
5	2	30
المجموع	**30**	

- نحدد رتبة الوسيط كما يلي : $\frac{n}{2} = \frac{30}{2} = 15$

[1] – عبد الحكيم بيصار ، مرجع سابق، ص 28 .

- القيمة التي تكرارها المتجمع الصاعد يساوي أو أكبر مباشرة من $\frac{n}{2}$ وهي 2، ومنه: $M_e = 2$

2-2 الوسيط في حالة متغير كمي متصل: عندما تكون لدينا البيانات في شكل جدول تكراري ذو فئات فإننا نتبع الخطوات التالية :

- نحسب التكرار المتجمع الصاعد.
- تحديد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N^{\uparrow}_{M_e} \geq \frac{n}{2}$
- نحسب الوسيط بطريقة المد الداخلي (التوسيع الداخلي): $M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{M_e-1}}{n_{M_e}}\right] \times C_{M_e}$

حيث:

Lim_{M_e} : الحد الأدنى للفئة الوسيطية.

n : مجموع التكرارت .

$N^{\uparrow}_{M_e-1}$: التكرار المتجمع الصاعد السابق للفئة الوسيطية .

n_{M_e} : تكرار الفئة الوسيطية.

C_{M_e} : طول الفئة الوسيطية .

مثال 09: بالعودة على معطيات المثال 05 أحسب الوسيط .

الحل: نحسب أولا التكرار التجميعي الصاعد كما هو موضح في الجدول التالي :

الجدول رقم (4-6) : توزيع علامات 50 طالب في مادة الاحصاء

العلامات(الفئات)	عدد الطلبة (التكرار) n_i	التكرار التجميعي الصاعد $N^{\uparrow}_i$
]0 -4]	5	5
]4 -8]	10	15
]8 -12]	15	30
]12 -16]	16	46
]16 -20]	4	50
المجموع	**50**	/

- نحدد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N^{\uparrow}_{M_e} \geq \frac{n}{2} = \frac{50}{2} = 25$ ومنه الفئة الوسيطية هي :]8- 12].

- نحسب الوسيط من العلاقة :

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{M_e-1}}{n_{M_e}}\right] \times C_{M_e} = 8 + \left[\frac{25-15}{15}\right] \times 4 = 10.66$$

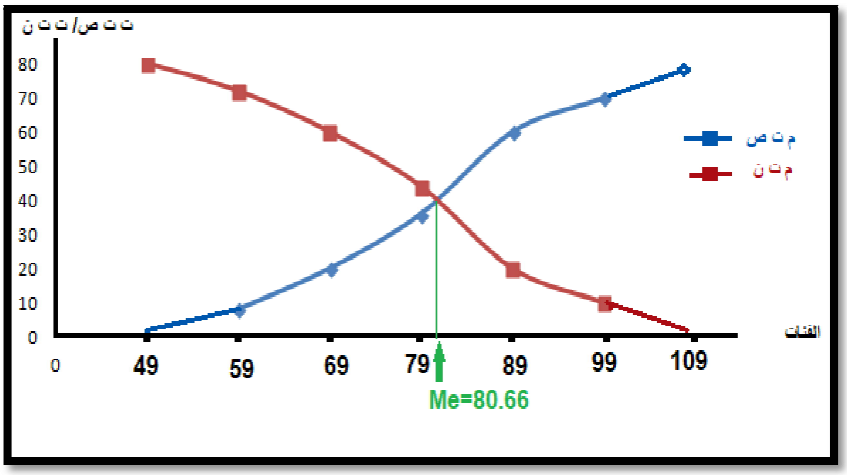
ملاحظة: يمكن إيجاد الوسيط في حالة المتغير الكمي المتصل بيانيا وذلك من خلال منحنى التوزيع التكراري التجميعي الصاعد والنازل، حيث يمثل نقطة التقاطع بين المنحنيين، ومن هذه النقطة نسقط عمودا رأسيا يلتقي مع محور الفئات في "نقطة" هي قيمة الوسيط [1].

مثال 10: حدد قيمة الوسيط بيانيا للبيانات المدرجة في الجدول التالي:

فئات الأجور	عدد العمال (التكرار) n_i	التكرار التجميعي الصاعد $N_i^{\uparrow}$	التكرار التجميعي النازل $N_i^{\downarrow}$
]59 -49]	8	8	80
]69 -59]	12	20	72
]79 -69]	16	36	60
]89 -79]	24	60	44
]99 -89]	10	70	20
]109 -99]	10	80	10
المجموع	$\sum n_i = 80$	/	/

الحل: يتم معرفة الوسيط بيانيا كما يلي:

الشكل رقم (4-1): إيجاد الوسيط بيانيا



[1] – محمود محمد سليم صالح ، مرجع سابق، ص 67. بتصرف

3- خصائص الوسيط : للوسيط مجموعة من الخصائص أهمها:

- لا يتأثر بالقيم الشاذة أو المتطرفة.
- أنه لا يأخذ عند حسابه كل القيم في الإعتبار، فهو يعتمد على قيمة أو قيمتين فقط[1].
- لا يمكن حسابه في حالات البيانات الوصفية.

ثالثا-المنوال:

المنوال لمجموعة من القيم هو القيمة التي تتكرر أكثر من غيرها أو القيمة الأكثر شيوعا ، وقد لا يكون للقيم منوال وقد يوجد أكثر من منوال [2]. ويرمز له بالرمز M_o ، ويحسب المنوال في حالة البيانات المبوبة وغير المبوبة كما يلي:

1- المنوال للبيانات غير المبوبة: يتم إيجاد المنوال هنا بالبحث مباشرة عن القيمة الأكثر تكرارا (شيوعا) .

مثال 11: أوجد المنوال للبيانات التالية : 18، 17، 15، 05، 12، 15، 11، 13، 15.

الحل: القيمة الأكثر تكرارا هي: 15 وبالتالي المنوال هو $M_o = 15$

2- المنوال للبيانات المبوبة: ونميز بين حالتين :

2-1 المنوال في حالة متغير كمي منفصل: المنوال هو القيمة التي تكرارها هو الأكبر في مجموعة القيم أو المستويات [3].

مثال 12: بالعودة إلى معطيات المثال 08، جد قيمة المنوال:

الحل: القيمة التي تكرارها هو الأكبر هي 2 لأن تكرارها هو 10، وبالتالي المنوال هو $M_o = 2$

2-2 المنوال في حالة متغير كمي متصل: لحساب المنوال في هذه الحالة نتبع الخطوات التالية:

- نحدد الفئة المنوالية وهي الفئة التي تقابل أكبر تكرار.
- نحسب المنوال من خلال العلاقة : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o}$

حيث:

Lim_{M_o} : الحد الأدنى للفئة المنوالية.

Δ_1 : هو الفرق بين أكبر تكرار والسابق له .

[1] – شرف الدين خليل ، مرجع سابق، ص 40.
[2] – محمود محمد سليم صالح ، مرجع سابق، ص 69.
[3] – عبد الحكيم بيصار ، مرجع سابق، ص 31 .

Δ_2 : هو الفرق بين أكبر تكرار واللاحق له .

C_{M_o} : طول الفئة المنوالية .

مثال 13: بالعودة إلى معطيات المثال 10 الموضحة في الجدول التالي :

فئات الأجور	[59 -49]	[69 -59]	[79 -69]	[89 -79]	[99 -89]	[109 -99]	المجموع
عدد العمال	8	12	16	24	10	10	**80**

المطلوب : أحسب قيمة المنوال:

الحل:

- نحدد الفئة المنوالية : لدينا أكبر تكرار هو 24 وبالتالي الفئة المنوالية هي [79- 89[

$$\Delta_1 = 24 - 16 = 8 \qquad \Delta_2 = 24 - 10 = 14$$

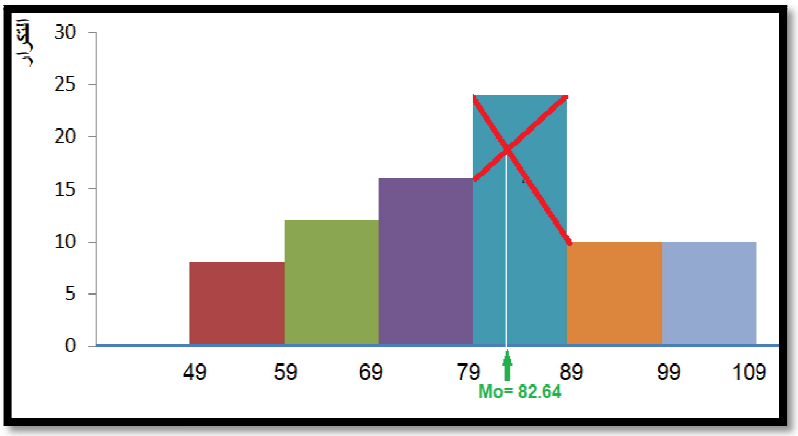
ومنه: قيمة المنوال هي : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o} = 79 + \left[\frac{8}{8+1}\right] \times 10 = 82.64$

ملاحظة: يمكن إيجاد المنوال في حالة المتغير الكمي المتصل بيانيا حيث بعد رسم المدرج التكراري نصل بين أركان الخلية الأكبر تكرارا وحواف الفئتين السابقة واللاحقة لها [1] .

مثال 14: بالعودة إلى معطيات المثال السابق أوجد قيمة المنوال بيانيا .

الحل : يتم إيجاد قيمة المنوال بيانيا من خلال الشكل التالي :

الشكل رقم (4-2): إيجاد المنوال بيانيا



[1] – محمود محمد سليم صالح ، مرجع سابق، ص 71.

3- خصائص المنوال : للمنوال مجموعة من الخصائص أهمها:

- سهل الحساب و لا يتأثر بالقيم الشاذة.
- يمكن حسابه من البيانات الوصفية.
- لا يأخذ جميع القيم في الحسبان .
- أحيانا يكون للقيم أكثر من منوال وبالتالي يصعب تحديد قيمة وحيدة له .

رابعا-الوسط (المتوسط) الهندسي: يعتبر الوسط الهندسي من مقاييس النزعة المركزية الثانوية، حيث يفضل استخدامه في حساب متوسط معدل النمو للمتغيرات المختلفة سواء أكانت لقيم متزايدة أو لقيم متناقصة، لذا فهو شائع الاستخدام عند حساب متوسط النسب –المنسوب- في الأرقام القياسية [1]، و يعرف الوسط الهندسي لمجموعة من القيم عددها (n) بأنه الجذر النوني لحاصل ضرب هذه القيم [2]، ويحسب وفقا لحالتين:

1- الوسط الهندسي للبيانات غير المبوبة: إذا كانت لدينا مجموعة من القيم عددها (n) قيمة هي : $X_1 , X_2 , X_3 \ldots\ldots , X_n$ فإن الوسط الهندسي $\bar{X}_G$ هو : $\bar{X}_G = \sqrt[n]{X_1.X_2.X_3 \ldots\ldots X_n}$

واختصارا هو : $\bar{X}_G = \sqrt[n]{\prod_{i=1}^{n} X_i}$

وفي كثير من الأحيان نستخدم اللوغاريتمات لحساب الوسط الهندسي لتسهيل الحسابات فتصبح العلاقة كالتالي: $\log_{\bar{X}_G} = \frac{1}{n}(\sum_{i=1}^{n} \log X_i)$

وبالتالي: $\bar{X}_G = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} \log(X_i)\right)}$

حيث :n: هو عدد القيم (البيانات) .

2- حالة البيانات المبوبة: ونميز بين حالتين :

1-2 حالة متغير كمي منفصل: إذا كانت لدينا مجموعة من القيم لمتغير إحصائي هي : $X_1 , X_2 , X_3 \ldots\ldots , X_k$ و تكراراتها على الترتيب هي: $n_1 , n_2 , n_3 \ldots\ldots , n_k$ فإن الوسط الهندسي $\bar{X}_G$ هو :

$$\bar{X}_G = \sqrt[n]{X_1^{n_1}.X_2^{n_2}.X_3^{n_3} \ldots\ldots X_k^{n_k}}$$

و إختصارا هو : $\bar{X}_G = \sqrt[n]{\prod_{i=1}^{k} X_i^{n_i}}$

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 160.

[2] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 97.

وفي كثير من الأحيان نستخدم اللوغاريتمات لحساب الوسط الهندسي لتسهيل الحسابات فتصبح العلاقة كالتالي: $\log_{\bar{X}_G} = \frac{1}{n}(\sum_{i=1}^{n} n_i \log X_i)$

وبالتالي: $\overline{X}_G = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i \log(X_i)\right)}$

ملاحظة: n: هو مجموع التكرارات $(\sum_{i=1}^{k} n_i)$في العلاقات السابقة.

2-2 حالة متغير كمي متصل: يحسب بنفس طريقة المتغير الكمي المنفصل فقط نستبدل قيم المتغير الإحصائي $X_1 , X_2 , X_3 \ldots\ldots, X_k$ بمراكز الفئات X_i' فإذا كانت لدينا: $X_1', X_2', X_3' \ldots\ldots, X_k'$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots, n_k$ فإن الوسط الهندسي $\bar{X}_G$ هو :

$$\bar{X}_G = \sqrt[n]{X_1'^{n_1}.X_2'^{n_2}.X_3'^{n_3} \ldots\ldots X_k'^{n_k}}$$

و إختصارا هو : $\bar{X}_G = \sqrt[n]{\prod_{i=1}^{k} X_i'^{n_i}}$

وفي كثير من الأحيان نستخدم اللوغاريتمات لحساب الوسط الهندسي لتسهيل الحسابات فتصبح العلاقة كالتالي: $\log_{\bar{X}_G} = \frac{1}{n}(\sum_{i=1}^{n} n_i \log X'_i)$

وبالتالي: $\overline{X}_G = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i \log(X_i')\right)}$

ملاحظة: n: هو مجموع التكرارات $(\sum_{i=1}^{k} n_i)$في العلاقات السابقة.

خامسا–الوسط التوافقي: هو مقياس آخر من مقاييس المتوسطات، يفضل استخدامه في حالات خاصة أي عندما يعبر عن المتغيرات في صورة معدلات زمنية ، كالمسافة التي تقطعها السيارة أو القطار أو الطائرة في وحدة الزمن، أو إنتاج ماكينة في الساعة مثلا، و أيضا متوسطات الأسعار إذا أعطيت بدلالة وحدة النقود...وهكذا...، ويعرف الوسط التوافقي لظاهرة أو متغير ما ، بأنه عبارة عن مقلوب الوسط الحسابي لمقلوبات قيم الظاهرة أو المتغير المشار إليهما [1]. ويحسب وفقا لحالتين:

1- الوسط التوافقي للبيانات غير المبوبة: إذا كانت لدينا مجموعة من القيم هي : $X_1 , X_2 , X_3 \ldots, X_n$ فإن الوسط التوافقي $\overline{X}_H$ هو : $\bar{X}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}}$

حيث :n: هو عدد القيم (البيانات) .

2- حالة البيانات المبوبة: ونميز بين حالتين :

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 168.

2-1 حالة متغير كمي منفصل: إذا كانت لدينا مجموعة من القيم لمتغير إحصائي هي : $X_1 , X_2 , X_3 \ldots\ldots , X_k$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots , n_k$ فإن الوسط التوافقي $\bar{X}_H$ هو [1]:

$$\bar{X}_H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{X_i}}$$

2-2 حالة متغير كمي متصل: يحسب بنفس طريقة المتغير الكمي المنفصل فقط نستبدل قيم المتغير الإحصائي $X_1 , X_2 , X_3 \ldots\ldots , X_k$ بمراكز الفئات X_i' فإذا كانت لدينا: $X_1' , X_2' , X_3' \ldots\ldots , X_k'$ و تكراراتها على الترتيب هي: $n_1 , n_2 , n_3 \ldots\ldots , n_k$ فإن الوسط التوافقي $\bar{X}_H$ هو :

$$\bar{X}_H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{X_i'}}$$

سادسا-الوسط التربيعي: يعرف الوسط التربيعي لمجموعة من القيم على أنه الجذر التربيعي لمتوسط مربعات القيم و يحسب على النحو التالي [2]:

1- الوسط التربيعي للبيانات غير المبوبة: إذا كانت لدينا مجموعة من القيم هي : $X_1 , X_2 , X_3 \ldots , X_n$ فإن الوسط التربيعي $\bar{X}_Q$ هو :

$$\bar{X}_Q = \sqrt{\frac{X_1^2 + X_2^2 + X_3^2 + \cdots + X_n^2}{n}}$$

و إختصارا هو:

$$\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n}}$$

حيث :n: هو عدد القيم (البيانات) .

2- حالة البيانات المبوبة: ونميز بين حالتين :

2-1 حالة متغير كمي منفصل: إذا كانت لدينا مجموعة من القيم لمتغير إحصائي هي : $X_1 , X_2 , X_3 \ldots\ldots , X_k$ و تكراراتها على الترتيب هي: $n_1 , n_2 , n_3 \ldots\ldots , n_k$ فإن الوسط التربيعي $\bar{X}_Q$ هو:

$$\bar{X}_Q = \sqrt{\frac{X_1^2 . n_1 + X_2^2 . n_2 + X_3^2 . n_3 + \cdots + X_n^2 . n_k .}{n_1 + n_2 + n_3 + \cdots + n_k}}$$

و إختصارا هو:

$$\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_i X_i^2}{\sum_{i=1}^{k} n_i}}$$

2-2 حالة متغير كمي متصل: يحسب بنفس طريقة المتغير الكمي المنفصل فقط نستبدل قيم المتغير الإحصائي $X_1 , X_2 , X_3 \ldots\ldots , X_k$ بمراكز الفئات X_i' فإذا كانت لدينا: $X_1' , X_2' , X_3' \ldots\ldots , X_k'$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots , n_k$ فإن الوسط التربيعي $\bar{X}_Q$ هو :

$$\bar{X}_Q = \sqrt{\frac{X_1'^2 . n_1 + X_2'^2 . n_2 + X_3'^2 . n_3 + \cdots + X_n'^2 . n_k .}{n_1 + n_2 + n_3 + \cdots + n_k}}$$

[1] – محمود محمد سليم صالح ، مرجع سابق، ص 62.

[2] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 49.

و إختصارا هو : $\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_i X_i'^2}{\sum_{i=1}^{k} n_i}}$

ملاحظة مهمة : دوما وفي جميع النتائج نجد :

الوسط التربيعي > الوسط الحسابي > الوسط الهندسي > الوسط التوافقي: $\bar{X}_H < \bar{X}_G < \bar{X} < \bar{X}_Q$

مثال 15: أحسب الوسط التربيعي، الوسط الحسابي، الوسط الهندسي، الوسط التوافقي للبيانات التالية:

5، 8 ، 4، 7

الحل:

-الوسط التربيعي: $\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{4} X_i^2}{4}} = \sqrt{\frac{5^2+8^2+4^2+7^2}{4}} = 6.20$

-الوسط الحسابي: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{4} X_i}{4} = \frac{5+8+4+7}{4} = 6$

-الوسط الهندسي: $\bar{X}_G = \sqrt[n]{\prod_{i=1}^{n} X_i} = \sqrt[4]{5 \times 8 \times 4 \times 7} = 5.78$

-الوسط التوافقي: $\bar{X}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}} = \frac{4}{\sum_{i=1}^{4} \frac{1}{X_i}} = \frac{4}{\frac{1}{5}+\frac{1}{8}+\frac{1}{4}+\frac{1}{7}} = 5.57$

نلاحظ أن: $\bar{X}_H < \bar{X}_G < \bar{X} < \bar{X}_Q$ محققة .

مثال 16: أحسب الوسط التربيعي، الوسط الحسابي، الوسط الهندسي، الوسط التوافقي لبيانات الجدول التكراري التالي:

الفئات	]8 -4]	]12 -8]	]16 -12]	]20 -16]	]24 -20]	المجموع
التكرارات	9	8	7	6	5	**35**

الحل:

الجدول رقم (4-7) : توزيع التكرارات حسب الفئات

الفئات	التكرارات n_i	مركز الفئة X_i'	$n_i X_i'^2$	$n_i X_i'$	$log(X_i')$	$n_i log(X_i')$	$\frac{n_i}{X_i}$
]8 -4]	9	6	324	54	0.778	7.002	1.5
]12 -8]	8	10	800	80	1	8	0.8
]16 -12]	7	14	1372	98	1.146	8.022	0.5
]20 -16]	6	18	1944	108	1.255	7.53	0.33
]24 -20]	5	22	2420	110	1.342	6.71	0.23
المجموع	$\sum n_i = 35$	/	**6860**	**450**	/	**37.264**	**3.36**

-الوسط التربيعي: $\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_i X_i'^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{\sum_{i=1}^{5} n_i X_i'^2}{35}} = \sqrt{\frac{6860}{35}} = 14$

-الوسط الحسابي: $\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i'}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i X_i'}{35} = \frac{450}{35} = 12.85$

-الوسط الهندسي:

$$\bar{X}_G = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i log(X_i')\right)} = 10^{\left(\frac{1}{35}\sum_{i=1}^{5} n_i log(X_i')\right)} = 10^{\left(\frac{1}{35} \; 37.264\right)} = 11.60$$

-الوسط التوافقي: $\bar{X}_H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{X_i'}} = \frac{35}{\sum_{i=1}^{5} \frac{n_i}{X_i'}} = \frac{35}{3.36} = 10.42$

نلاحظ أن: $\bar{X}_H < \bar{X}_G < \bar{X} < \bar{X}_Q$ محققة .

سابعا-مشتقات الوسيط (الربيعات و العشيرات، والمئينات):

إذا رتبت مجموعة من الأرقام حسب قيمتها تصاعديا أو تنازليا فان القيمة التي في المنتصف (أو الوسط الحسابي للقيمتين بالمنتصف) و التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد هي الوسيط.

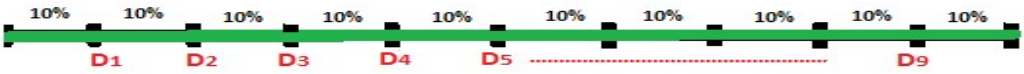
و بتعميم هذه الفكرة يمكن أن نفكر في القيم التي تقسم المجموعة إلى أربعة أجزاء متساوية هذه القيم يرمز لها بالرموز Q_3 ، Q_2 ، Q_1 تسمى الربيع الأول ،الربيع الثاني ،الربيع الثالث [1]، على الترتيب[2]:

- فالربيع الأول Q_1 هو القيمة التي يسبقها ربع البيانات ويليها ثلاثة أرباع البيانات.
- والربيع الثاني Q_2 وهو الوسيط، أي القيمة التي يسبقها نصف البيانات ويليها نصفها .
- والربيع الثالث Q_3 وهو القيمة التي يسبقها ثلاثة أرباع البيانات ويليها ربعها .



كذلك فإن القيم التي تقسم المجموعة إلى عشرة أجزاء متساوية يرمز لها بالرموز D_9 ..... ، D_2 ، D_1 تسمى العشير الأول، العشير الثاني، العشير الثالث، إلى غاية العشير التاسع على الترتيب :

- فالعشير الأول هو القيمة التي يسبقها $\frac{1}{10}$ من البيانات ويليها $\frac{9}{10}$ من البيانات.
- وهكذا تعرف العشيرات الأخرى بنفس الطريقة، مع ملاحظة أن العشير الخامس هو الوسيط.



أيضا فإن القيم التي تقسم المجموعة إلى مائة جزء متساوية يرمز لها بالرموز P_{99} ..... ، P_2 ، P_1 تسمى المئين الأول، المئين الثاني، إلى غاية المئين التاسع والتسعون على الترتيب :

[1] – محمود محمد سليم صالح ، مرجع سابق ، ص 73.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص 109.

- فالمئين الأول هو القيمة التي يسبقها $\frac{1}{100}$ من البيانات ويليها $\frac{99}{100}$ من البيانات.
- وهكذا تعرف المئينات الأخرى بنفس الطريقة، مع ملاحظة أن المئين الخمسون هو الوسيط ،كما أن المئين الخامس و العشرون والمئين الخامس و السبعون يساويان الربيع الأول و الربيع الثالث على الترتيب.



ولحساب هذه المقاييس يجب أن نميز بين حالتي البيانات المبوبة وغير المبوبة وذلك كمايلي :

1- حالة البيانات غير المبوبة: ويتم تلخيص ذلك في الجدول التالي [1]:

الربيعات Q_i	العشيرات D_i	المئينات P_i
1- نقوم بترتيب البيانات ترتيبا تصاعديا 2- نحدد رتبة الربيع من : $\frac{i}{4}(n+1)$ فإذا كانت : - $\frac{i}{4}(n+1)$: دون فواصل نـــأخذ القيمة مباشرة. - $\frac{i}{4}(n+1)$: بالفواصل نـــأخذ متوسط القيمتين . و يحسب الربيع من العلاقة : $Q_i = X_{\frac{i}{4}(n+1)}$	1- نقوم بترتيب البيانات ترتيبا تصاعديا 2- نحدد رتبة العشير من : $\frac{i}{10}(n+1)$ فإذا كانت : - $\frac{i}{10}(n+1)$: دون فواصل نـــأخذ القيمة مباشرة. - $\frac{i}{10}(n+1)$: بالفواصل نـــأخذ متوسط القيمتين . و يحسب العشير من العلاقة : $D_i = X_{\frac{i}{10}(n+1)}$	1- نقوم بترتيب البيانات ترتيبا تصاعديا 2- نحدد رتبة المئين من : $\frac{i}{100}(n+1)$ فإذا كانت : - $\frac{i}{100}(n+1)$: دون فواصل نـــأخذ القيمة مباشرة. - $\frac{i}{100}(n+1)$: بالفواصل نـــأخذ متوسط القيمتين . و يحسب العشير من العلاقة : $P_i = X_{\frac{i}{100}(n+1)}$

مثال 17: احسب الربيع الثالث، والعشير السابع والمئين 37 للبيانات التالية :

18، 17، 05، 12، 11، 13، 15، 10، 9، 6.

الحل: حساب كلا من:

- الربيع الثالث:

نرتب البيانات ترتيبا تصاعديا : 5، 6، 9، 10، 11، 12، 13، 15، 17، 18

نحدد رتبة الربيع الثالث : $\frac{i}{4}(n+1) = \frac{3}{4}(10+1) = 8.25$

بما أن القيمة (8.25) بالفاصلة فإن قيمة الربيع الثالث هي متوسط القيمة التي رتبتها 8 والقيمة التي رتبتها 9.

أي : $Q_3 = X_{\frac{3}{4}(10+1)} = X_{(8.25)} = \frac{X_8+X_9}{2} = \frac{15+1}{2} = 16$

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 45.

– العشير السابع :

نرتب البيانات ترتيبا تصاعديا : 5، 6، 9، 10، 11، 12، 13، 15، 17، 18

نحدد رتبة العشير السابع : $\frac{i}{10}(n+1) = \frac{7}{10}(10+1) = 7.7$

بما أن القيمة (7.7) بالفاصلة فإن قيمة العشير السابع هي متوسط القيمة التي رتبتها 7 والقيمة التي رتبتها 8.

أي : $D_7 = X_{\frac{7}{10}(10+1)} = X_{(7.7)} = \frac{X_7+X_8}{2} = \frac{13+1}{2} = 14$

– المئين 37 :

نرتب البيانات ترتيبا تصاعديا : 5، 6، 9، 10، 11، 12، 13، 15، 17، 18

نحدد رتبة العشير السابع : $\frac{i}{100}(n+1) = \frac{37}{100}(10+1) = 4.07$

بما أن القيمة (4.07) بالفاصلة فإن قيمة المئين 37 هي متوسط القيمة التي رتبتها 4 والقيمة التي رتبتها 5.

أي : $P_{37} = X_{\frac{37}{100}(10+1)} = X_{(4.07)} = \frac{X_4+X_5}{2} = \frac{10+11}{2} = 10.5$

2- حالة البيانات المبوبة: ونميز بين حالتين :

2-1 حالة متغير كمي منفصل: ويتم ذلك كما يلي:

الربيعات Q_i	العشيرات D_i	المئينات P_i
1- نحسب التكرار المتجمع الصاعد للقيم. 2- نحدد رتبة الربيع: $\frac{in}{4}$ (ملاحظة : n تمثل مجموع التكرارات) 3- نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر من أو يساوي $\frac{in}{4}$ مباشرة وهي القيمة التي تمثل الربيع.	1- نحسب التكرار المتجمع الصاعد للقيم. 2- نحدد رتبة العشير: $\frac{in}{10}$ (ملاحظة : n تمثل مجموع التكرارات) 3- نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر من أو يساوي $\frac{in}{10}$ مباشرة وهي القيمة التي تمثل العشير.	1- نحسب التكرار المتجمع الصاعد للقيم. 2- نحدد رتبة المئين: $\frac{in}{100}$ (ملاحظة : n تمثل مجموع التكرارات) 3- نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر من أو يساوي $\frac{in}{100}$ مباشرة وهي القيمة التي تمثل المئين.

ملاحظة: يمكن حسابه بنفس الطريقة في حالة البيانات غير المبوبة .

مثال 18: من معطيات المثال 08 السابق والموجودة في الجدول التالي :

عدد الأطفال (قيم المتغير) X_i	عدد الأسر (التكرار) n_i	التكرار التجميعي الصاعد $N_i^{\uparrow}$
0	3	3
1	8	11
2	10	21
3	3	24
4	4	28
5	2	30
المجموع	**30**	

المطلوب: حساب الربيع الأول، والعشير الرابع، والمئين 45 .

الحل: حساب كلا من:

- الربيع الأول :
  - نحدد رتبة الربيع الأول كما يلي : $\frac{in}{4} = \frac{1(30)}{4} = 7.5$
  - القيمة التي تكرارها المتجمع الصاعد يساوي أو أكبر مباشرة من $\frac{in}{4}$ وهي 1، ومنه: $Q_1 = 1$
- العشير الرابع :
  - نحدد رتبة العشير الرابع كما يلي : $\frac{in}{10} = \frac{4(30)}{10} = 12$
  - القيمة التي تكرارها المتجمع الصاعد يساوي أو أكبر مباشرة من $\frac{in}{10}$ وهي 2، ومنه: $D_4 = 2$
- المئين 45 :
  - نحدد رتبة المئين 45 كما يلي : $\frac{in}{100} = \frac{45(30)}{100} = 13.5$
  - القيمة التي تكرارها المتجمع الصاعد يساوي أو أكبر مباشرة من $\frac{in}{100}$ وهي 2، ومنه: $P_{45} = 2$

2-2 حالة متغير كمي متصل: ويتم ذلك كما يلي:

الربيعات Q_i	العشيرات D_i	المئينات P_i
- نحسب التكرار المتجمع الصاعد. - تحديد الفئة الربيعية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـ $\frac{in}{4}$ أو يساويها ، أي: $N_{Q_i}^{\uparrow} \geq \frac{in}{4}$ - نحسب الربيع بالطريقة: $Q_i = Lim_{Q_i} + \left[\frac{\frac{in}{4} - N_{Q_i-1}^{\uparrow}}{n_{Q_i}}\right] \times C_{Q_i}$ حيث: Lim_{Q_i}: الحد الأدنى للفئة الربيعية. n : مجموع التكرارات . $N_{Q_i-1}^{\uparrow}$: التكرار المتجمع الصاعد السابق للفئة الربيعية . n_{Q_i} : تكرار الفئة الربيعية. C_{Q_i} : طول الفئة الربيعية .	- نحسب التكرار المتجمع الصاعد. - تحديد الفئة العشيرية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـ $\frac{in}{10}$ أو يساويها ، أي: $N_{D_i}^{\uparrow} \geq \frac{in}{10}$ - نحسب العشير بالطريقة: $D_i = Lim_{D_i} + \left[\frac{\frac{in}{10} - N_{D_i-1}^{\uparrow}}{n_{D_i}}\right] \times C_{D_i}$ حيث: Lim_{D_i}: الحد الأدنى للفئة العشيرية. n : مجموع التكرارات . $N_{D_i-1}^{\uparrow}$: التكرار المتجمع الصاعد السابق للفئة العشيرية . n_{D_i} : تكرار الفئة العشيرية. C_{D_i} : طول الفئة العشيرية .	- نحسب التكرار المتجمع الصاعد. - تحديد الفئة المئينية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـ $\frac{in}{100}$ أو يساويها ، أي: $N_{P_i}^{\uparrow} \geq \frac{in}{100}$ - نحسب المئين بالطريقة: $P_i = Lim_{P_i} + \left[\frac{\frac{in}{100} - N_{P_i-1}^{\uparrow}}{n_{P_i}}\right] \times C_{P_i}$ حيث: Lim_{P_i}: الحد الأدنى للفئة المئينية. n : مجموع التكرارات . $N_{P_i-1}^{\uparrow}$: التكرار المتجمع الصاعد السابق للفئة المئينية . n_{P_i} : تكرار الفئة المئينية. C_{P_i} : طول الفئة المئينية .

مثال19: من معطيات المثال 09 المبينة بالجدول التالي:

العلامات(الفئات)	عدد االطلبة (التكرار) n_i	التكرار التجميعي الصاعد $N_i^{\uparrow}$
]4 -0]	5	5
]8 -4]	10	15
]12 -8]	15	30
]16 -12]	16	46
]20 -16]	4	50
المجموع	**50**	/

المطلوب: أحسب الربيع الثالث، العشير الثامن، المئين 89.

الحل: حساب كلا من :

-الربيع الثالث:

- نحدد الفئة الربيعية الثالثة: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{in}{4}$ أو يساويها ، أي: $N_{Q_3}^{\uparrow} \geq \frac{in}{4} = \frac{3(50)}{4} = 37.5$ ومنه الفئة الربيعية الثالثة هي :]12- 16].
- نحسب الربيع الثالث من العلاقة :

$$Q_3 = Lim_{Q_3} + \left[\frac{\frac{in}{4} - N_{Q_3-1}^{\uparrow}}{n_{Q_3}}\right] \times C_{Q_3} = 12 + \left[\frac{37.5 - 30}{16}\right] \times 4 = 13.87$$

-العشير الثامن:

- نحدد الفئة العشيرية الثامنة: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{in}{10}$ أو يساويها ، أي: $N_{D_8}^{\uparrow} \geq \frac{in}{10} = \frac{8(50)}{10} = 40$ ومنه الفئة العشيرية الثامنة هي :]12- 16].
- نحسب العشير الثامن من العلاقة :

$$D_8 = Lim_{D_8} + \left[\frac{\frac{in}{10} - N_{D_8-1}^{\uparrow}}{n_{D_8}}\right] \times C_{D_8} = 12 + \left[\frac{40 - 30}{16}\right] \times 4 = 14.5$$

-المئين 89:

- نحدد الفئة المئينية 89: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{in}{100}$ أو يساويها ، أي: $N_{P_{89}}^{\uparrow} \geq \frac{in}{100} = \frac{89(50)}{100} = 44.5$ ومنه الفئة المئينية 89 هي :]12- 16].
- نحسب المئين 89 من العلاقة :

$$P_{89} = Lim_{P_{89}} + \left[\frac{\frac{in}{100} - N_{P_{89}-1}^{\uparrow}}{n_{p_{89}}}\right] \times C_{P_{89}} = 12 + \left[\frac{44.5 - 30}{16}\right] \times 4 = 15.62$$

المحور الخامس: مقاييس التشتت

تمهيد:

رأينا في المحاور السابقة طرق عرض البيانات الاحصائية جدوليا أو بيانيا ، ثم درسنا مقاييس النزعة المركزية التي تصف البيانات عدديا، غير أن هذه المقاييس غيركافية لوصف مجموعة من البيانات وصفا كاملا، فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من إختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات) .

فمثلا الوسط الحسابي للمجموعتين :

- A: 5، 5، 5، 5، 5
- B: 5، 3، 7، 4، 6

هو نفسه ويساوي 5، ولكن هناك اختلاف في البيانات ، فالوسط الحسابي يمثل مركز البيانات لكنه لا يبين مدى تقارب أو تشتت البيانات حول هذا الوسط . لذا لابد من إستعمال مقاييس أخرى توضح لنا مدى إختلاف البيانات فيما بينها، أي مدى تقاربها أو تباعدها عن بعضها البعض [1].

إذن من هنا تبرز الحاجة إلى إستعمال مقاييس التشتت التي تساعدنا في تقديم وصف دقيق حول مختلف البيانات والتوزيعات التكرارية، وفي المقارنة بين مجموعات لمختلف الظواهر، فمقاييس التشتت تحدد مدى تقارب أو تباعد أو تناثر(تشتت) أو إختلاف القيم عن بعضها البعض أو عن نقطة معينة تسمى الوسط الحسابي مثلا . فقد يكون صغيرا إذا كانت قيم المجموعة قريبة من بعضها وقد يكون كبيرا إذا كانت القيم بعيدة عن بعضها البعض .

ولحساب مقدار التشتت هناك عدة مقاييس منها مقاييس التشتت المطلقة كالمدى ، والمدى الربيعي ، الإنحراف المتوسط، التباين والانحراف المعياري ، ومنها مقاييس التشتت النسبية التي يعبر عنها على شكل نسب مئوية كمعامل الإختلاف [2]، وفي مايلي عرض لهذه المقاييس :

أولا–مقاييس التشتت المطلقة : وهي :

1– المدى: ويعتبر من أسهل وأبسط مقاييس التشتت، وإن كان ليس أدقها [3]، ويرمز له بالرمز *(R)* ويحسب وفقا لحالتين :

[1] – محمود محمد سليم صالح ، مرجع سابق ، ص 83.

[2] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 112.

[3] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 185.

1-1 المدى للبيانات غير المبوبة: يعرف المدى هنا بأنه الفرق بين أكبر قيمة و أصغر قيمة في البيانات [1]،

أي : المدى = أكبر قيمة – أقل قيمة : $R = X_{max} - X_{min}$

مثال 01: أوجد المدى للقيم التالية: 7، 15، 5، 2، 17، 22، 3، 12.

الحل: $R = X_{max} - X_{min} = 22 - 2 = 20$

1-2 المدى للبيانات المبوبة: ونميز بين حالتين :

1-2-1 حالة متغير كمي منفصل: ويحسب بنفس العلاقة السابقة : $R = X_{max} - X_{min}$

مثال 02: أحسب المدى لبيانات الجدول التكراري التالي :

قيم المتغير X_i	2	3	4	5	6	المجموع
التكرارات n_i	8	9	7	4	2	**30**

الحل: $R = X_{max} - X_{min} = 6 - 2 = 4$

1-2-2 حالة متغير كمي متصل : ويحسب بتطبيق نفس العلاقة السابقة، ولكن يعرف على أساس أن أكبر قيمة هي الحد الأعلى للفئة الأخيرة ، وأقل قيمة هي الحد الأدنى للفئة الأولى أي :

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى.

مثال 03: أحسب المدى لبيانات الجدول التكراري التالي :

الفئات	]10 -8]	]12 -10]	]14 -12]	]16 -14]	]18 -16]	المجموع
التكرارات	6	8	5	7	4	**30**

الحل: $R = X_{max} - X_{min} = 18 - 8 = 10$

1-3 خواص المدى: للمدى عدة خصائص هي:

- أنه بسيط وسهل الحساب.
- يعطي تصور عام عن طبيعة البيانات.
- يعتمد في حسابه على قيمتين فقط.

[1] – محمد صبحي أبو صالح ، مرجع سابق ، ص 118.

- مقياس تقريبي لا يعتمد عليه في وصف البيانات .

2- المدى الربيعي: هو الفرق بين الربيع الثالث والربيع الأول [1]، و يعتبر من مقاييس التشتت القليلة الاستخدام ، ويرمز له بالرمز (RQ) ، ويحسب بالعلاقة : $RQ = Q_3 - Q_1$

وهذه العلاقة تستخدم في حالتي البيانات المبوبة وغير المبوبة فقط يجب التفريق بين البيانات المبوبة و غير المبوبة عند حساب الربيع الثالث والربيع الأول كما رأينا في المحور السابق .

2-1 خواص المدى الربيعي : للمدى الربيعي عدة خصائص هي[2]:

- إستعمالاته محدودة نظرا لبساطته، غير أنه أحسن من المدى العام .
- يستعمل في المقارنة بين توزيعين إحصائيين أو أكثر.
- يتغير طوله مقارنة بالمدى العام حسب طبيعة التوزيع .
- يضم أو يعتمد على 50 % من البيانات .

3- الإنحراف المتوسط: إذا كان المدى و المدى الربيعي يقومان على قياس التشتت لمجموعة قيم أي ظاهرة عن بعضها البعض، فإن الإنحراف المتوسط يعتبر أحد مقاييس التشتت التي تقيس بُعد القيم عن الوسط الحسابي ولذلك فإن الإنحراف المتوسط يأخذ بالإعتبار جميع القيم في التوزيع ، ويعرف على أنه عبارة عن مجموع القيم المطلقة لإنحرافات القيم عن وسطها الحسابي مقسوما على عدد القيم أو مجموعها [3].

إن إدراج القيمة المطلقة يعني التخلص من مشكلة مجموع الإنحرافات $\sum(X_i - \bar{X})$ الذي يساوي دوما صفر وهذا يعني تجريد الانحرافات من إشاراتها السالبة ليصبح للانحراف المتوسط معنى[4]، و يرمز له بالرمز (MD) ويحسب وفقا لحالتين :

3-1 الإنحراف المتوسط للبيانات غير المبوبة: إذا كانت لدينا مجموعة من القيم هي :

$X_1 , X_2 , X_3 \ldots\ldots, X_n$ وكان وسطها الحسابي $\bar{X}$ فإن الإنحراف المتوسط MD هو :

$$MD = \frac{\sum_{i=1}^{n}|X_i - \bar{X}|}{n}$$

حيث n: هو عدد القيم (البيانات) .

[1] – نفس المرجع ، ص 118.
[2] – جلاطو جيلالي ، مرجع سابق، ص 71.
[3] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 117.
[4] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 193.

مثال 04: أحسب الإنحراف المتوسط للقيم التالية التي تمثل أوزان خمس طلاب بالكيلو غرام :

60، 70، 80، 75، 65.

الحل:

نحسب أولا الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{5} X_i}{5} = \frac{60+70+80+75+6}{5} = 70$$

نحسب الإنحراف المتوسط كما يلي :

$$MD = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n} = \frac{\sum_{i=1}^{5} |X_i - \bar{X}|}{5} = \frac{|60-70|+|70-70|+|80-70|+|75-70|+|65-7\ |}{5} = \frac{30}{5} = 6$$

3-2 الانحراف المتوسط للبيانات المبوبة: ونميز بين حالتين:

3-2-1 حالة متغير كمي منفصل: إذا كانت لدينا مجموعة من القيم لمتغير إحصائي هي :

$X_1 , X_2 , X_3 \ldots\ldots, X_k$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots, n_k$ فإن الانحراف المتوسط MD هو[1] : $MD = \frac{\sum_{i=1}^{k} n_i |X_i - \bar{X}|}{\sum_{i=1}^{k} n_i}$

مثال 05: أحسب الانحراف المتوسط لبيانات المثال (02) السابق.

الحل:

الجدول رقم (5-1) : توزيع التكرارات حسب قيم المتغير X_i

X_i (قيم المتغير)	n_i (التكرارات)	$n_i X_i$	$\lvert X_i - \bar{X} \rvert$	$n_i \lvert X_i - \bar{X} \rvert$
2	8	16	1.43	11.44
3	9	27	0.43	3.87
4	7	28	0.57	3.99
5	4	20	1.57	6.28
6	2	12	2.57	5.14
المجموع	**30**	**103**	/	**30.72**

نحسب أولا الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i X_i}{30} = \frac{103}{30} = 3.43$$

نحسب الانحراف المتوسط كما يلي :

[1] – محمود محمد سليم صالح ، مرجع سابق ، ص 91.

$$MD = \frac{\sum_{i=1}^{k} n_i|X_i - \bar{X}|}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i|X_i - \bar{X}|}{30} = \frac{30.72}{30} = 1.024$$

2-2-3 حالة متغير كمي متصل: يحسب بنفس طريقة المتغير الكمي المنفصل فقط نستبدل قيم المتغير الاحصائي X_k ,....., X_3 , X_2 , X_1 بمراكز الفئات X_i' فإذا كانت لدينا: X_k',....., X_3', X_2', X_1' و تكراراتها على الترتيب هي : n_k ,....., n_3 , n_2 , n_1 فإن الانحراف المتوسط MD هو: $MD = \frac{\sum_{i=1}^{k} n_i|X_i' - \bar{X}|}{\sum_{i=1}^{k} n_i}$

مثال 06: أحسب الإنحراف المتوسط بالإعتماد على نفس معطيات المثال (03) السابق.

الحل:

الجدول رقم (2-5) : توزيع التكرارات حسب الفئات

الفئات	التكرارات n_i	مراكز الفئات X_i'	$n_i X_i'$	$\lvert X_i' - \bar{X}\rvert$	$n_i\lvert X_i' - \bar{X}\rvert$
]10 -8]	6	9	54	3.66	21.96
]12 -10]	8	11	88	1.66	13.28
]14 -12]	5	13	65	0.34	1.7
]16 -14]	7	15	105	2.34	16.38
]18 -16]	4	17	68	4.34	17.36
المجموع	**30**	/	**380**		**70.68**

نحسب أولا الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i'}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i X_i'}{30} = \frac{380}{30} = 12.66$$

نحسب الإنحراف المتوسط كما يلي :

$$MD = \frac{\sum_{i=1}^{k} n_i|X_i' - \bar{X}|}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i|X_i' - \bar{X}|}{30} = \frac{70.68}{30} = 2.356$$

3-3خواص الانحراف المتوسط : للانحراف المتوسط عدة خصائص هي:

- بسبب وجود القيمة المطلقة فإنه لا يستعمل بشكل واسع.
- أحسن من المقاييس السابقة لأنه يأخذ في الإعتبار جميع القيم عند حسابه.
- يتأثر بالقيم المتطرفة، لأن إنحرافها عن المتوسط الحسابي يكون كبيرا.

4-التباين والانحراف المعياري: يعتبر مقياسا التبيان والإنحراف المعياري من مقاييس التشتت الأوسع إنتشارا ، والإنحراف المعياري هو الجذر التربيعي للتباين .

4-1 التباين: يعرف التباين بأنه متوسط مربعات إنحرافات القيم عن وسطها الحسابي [1]، وإذا كنا قد تخلصنا من مشكلة مجموع الإنحرافات $\sum(X_i - \bar{X})$ الذي يساوي دوما صفر في الإنحراف المتوسط من خلال القيمة المطلقة فإننا نعالج هذه المشكلة هنا بواسطة تربيع الإنحرافات، وعادة ما يرمز له بالرمز $V(X)$ ويحسب وفقا لحالتين :

4-1-1 التباين للبيانات غير المبوبة: إذا كانت لدينا مجموعة من القيم هي : $X_1 , X_2 , X_3 \ldots\ldots, X_n$ وكان وسطها الحسابي $\bar{X}$ فإن التباين $V(X)$ هو : $V(X) = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}$

حيث :n : هو عدد القيم (البيانات) .

مثال 07: أحسب التباين باستخدام معطيات المثال (04) السابق

الحل:

الوسط الحسابي $\bar{X}$ محسوب سابق و هو : $\bar{X} = 70$

نحسب التباين كما يلي :

$$V(X) = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n} = \frac{\sum_{i=1}^{5}(X_i-\bar{X})^2}{5} = \frac{(60-70)^2+(70-70)^2+(80-70)^2+(75-7\ )^2+(65-7\ )^2}{5} = 50$$

4-1-2 التباين للبيانات المبوبة: ونميز بين حالتين:

4-1-2-1 حالة متغير كمي منفصل: إذا كانت لدينا مجموعة من القيم لمتغير إحصائي هي :

$X_1 , X_2 , X_3 \ldots\ldots, X_k$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots, n_k$ فإن التباين $V(X)$ هو[2]:

$$V(X) = \frac{\sum_{i=1}^{k} n_i(X_i-\bar{X})^2}{\sum_{i=1}^{k} n_i}$$

مثال 08: أحسب التباين لبيانات المثال (02) السابق.

الحل:

[1] – نفس المرجع ، ص 94.

[2] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 120.

الجدول رقم (5-3) : توزيع التكرارات حسب قيم المتغير X_i

$n_i(X_i-\bar{X})^2$	$(X_i-\bar{X})^2$	n_iX_i	(التكرارات) n_i	(قيم المتغير) X_i
14.9292	2.0449	16	8	2
1.6641	0.1849	27	9	3
2.2743	0.3249	28	7	4
9.8596	2.4649	20	4	5
13.2098	6.6049	12	2	6
41.937	/	**103**	**30**	المجموع

الوسط الحسابي $\bar{X}$ محسوب سابقا وهو : $\bar{X} = 3.43$

نحسب التباين كما يلي :

$$V(X) = \frac{\sum_{i=1}^{k} n_i(X_i-\bar{X})^2}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i(X_i-\bar{X})^2}{30} = \frac{41.937}{30} = 1.4$$

4-1-2-2 حالة متغير كمي متصل: يحسب بنفس طريقة المتغير الكمي المنفصل فقط نستبدل قيم المتغير الإحصائي X_k ,..... X_1 , X_2 , X_3 بمراكز الفئات X_i' فإذا كانت لدينا: X_1' ,X_2' ,X_3' .....,X_k' و تكراراتها على الترتيب هي : n_k ,..... n_1 , n_2 , n_3 فإن التباين $V(X)$ هو: $V(X) = \frac{\sum_{i=1}^{k} n_i(X_i'-\bar{X})^2}{\sum_{i=1}^{k} n_i}$

مثال 09: أحسب التباين بالإعتماد على نفس معطيات المثال (03) السابق.

الحل:

الجدول رقم (5-4) : توزيع التكرارات حسب الفئات

$n_i(X_i'-\bar{X})^2$	$(X_i'-\bar{X})^2$	n_iX_i'	مراكز الفئات X_i'	التكرارات n_i	الفئات
80.3736	13.3956	54	9	6	]10 -8]
22.0448	2.7556	88	11	8	]12 -10]
0.578	0.1156	65	13	5	]14 -12]
38.3292	5.4756	105	15	7	]16 -14]
75.3424	18.8356	68	17	4	]18 -16]
216.668		**380**	/	**30**	المجموع

الوسط الحسابي $\bar{X}$ محسوب سابقا وهو : $\bar{X} = 12.66$

نحسب التباين كما يلي :

$$V(X) = \frac{\sum_{i=1}^{k} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i (X_i' - \bar{X})^2}{30} = \frac{216.668}{30} = 7.22$$

4-1-3 خواص التباين: للتباين عدة خصائص هي:

- قيمته تكون كبيرة ووحدات قياسه تكون مربعة، لأنه يأخذ مربعات القيم في الحساب، الشيء الذي لا يجعله يعطي نظرة في تمام الوضوح حول مدى تشتت القيم، لذلك يتم في غالب الأحيان استبداله بالإنحراف المعياري [1].
- يأخذ في الإعتبار جميع القيم عند حسابه.

4-2 الإنحراف المعياري : يعتبر الإنحراف المعياري أكثر مقاييس التشتت استخداما في النظرية الإحصائية وفي الإحصاء الإستدلالي وذلك لما يتميز به من خصائص تؤهله لذلك [2]، والإنحراف المعياري هو الجذر التربيعي للتباين [3]، و يرمز له بالرمز $\sigma(X)$ أو$s(X)$ ويحسب وفقا لحالتين :

4-2-1 الانحراف المعياري للبيانات غير المبوبة: إذا كانت لدينا مجموعة من القيم هي:

$X_1 , X_2 , X_3 \ldots\ldots, X_n$ وكان وسطها الحسابي $\bar{X}$ فإن الإنحراف المعياري $\sigma(X)$ هو :

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}}$$

حيث :n : هو عدد القيم (البيانات) .

مثال 10: أحسب الإنحراف المعياري باستخدام معطيات المثال (04) السابق

الحل:

التباين محسوب سابقا في المثال 07 و نحسب فقط الانحراف المعياري كما يلي :

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}} = \sqrt{50} = 7.07$$

4-2-2 الإنحراف المعياري للبيانات المبوبة: ونميز بين حالتين:

4-2-2-1 حالة متغير كمي منفصل: إذا كانت لدينا مجموعة من القيم لمتغير إحصائي هي :

[1] – عبد الحكيم بيصار ، مرجع سابق، ص 46 .
[2] – أحمد فوزي ملوخيه ، مرجع سابق، ص 218.
[3] – محمود محمد سليم صالح ، مرجع سابق ، ص 95.

$X_1 , X_2 , X_3 \ldots\ldots, X_k$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots, n_k$ فإن الإنحراف المعياري $\sigma(X)$ هو: $\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}}$

مثال 11: أحسب الانحراف المعياري لبيانات المثال (02) السابق.

الحل:

التباين محسوب سابقا في المثال 08 و نحسب فقط الإنحراف المعياري كما يلي :

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{1.4} = 1.18$$

4-2-2-2 حالة متغير كمي متصل: يحسب بنفس طريقة المتغير الكمي المنفصل فقط نستبدل قيم المتغير الإحصائي $X_1 , X_2 , X_3 \ldots\ldots, X_k$ بمراكز الفئات X_i' فإذا كانت لدينا: $X_1', X_2', X_3' \ldots\ldots, X_k'$ و تكراراتها على الترتيب هي : $n_1 , n_2 , n_3 \ldots\ldots, n_k$ فإن الانحراف المعياري $\sigma(X)$ هو:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i}}$$

مثال 12: أحسب الإنحراف المعياري بالإعتماد على نفس معطيات المثال (03) السابق.

الحل:

التباين محسوب سابقا في المثال 09 و نحسب فقط الإنحراف المعياري كما يلي :

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{7.22} = 2.69$$

4-2-3 خواص الانحراف المعياري : للانحراف المعياري عدة خصائص هي [1]:

- أكثر مقاييس التشتت استخداما، إتساقا وتفسيرا ومنطقا للتشتت.
- قيمته لا تتأثر مطلقا إذا طرحنا مقدارا ثابتا من جميع قيم الظاهرة .
- يتأثر بالقيم المتطرفة .
- عدم إمكانية حسابه من الجداول التكرارية المفتوحة .
- يأخذ في الإعتبار جميع القيم عند حسابه.

[1] – أحمد فوزي ملوخيه ، مرجع سابق، ص ص 228 – 229.

ثانيا –مقاييس التشتت النسبية :

اذا أردنا من خلال مقاييس التشتت المطلقة مقارنة تشتت مجموعتين أو أكثر من الظواهر، مختلفة في وحدات القياس فهو غير ممكن، حيث رأينا أن مقاييس التشتت المطلقة السابقة يعبر عنها بالوحدات الأصلية للمتغير أو للظاهرة التي نقيسها، وعليه لا يمكن مقارنة إنحراف معياري للأجور وليكن 6 دنانير بإنحراف معياري لوقت الإنتاج وليكن 6 دقائق، ذلك أن عملية المقارنة السابقة لتشتت الظاهرتين تكون مستحيلة لإختلاف وحدات القياس فيهما، فليس من المعقول مقارنة الدنانير بالدقائق [1].

ويفرض علينا المنطق العلمي أيضا عدم المقارنة بين ظاهرتين لهما نفس وحدة القياس و لكن تختلفان في متوسط كل منهما (إختلاف في طبيعة توزيع مفردات كل ظاهرة حول المتوسط) [2].

وعليه فإنه من الضروري عند مقارنة التشتت بين عدة ظواهر مراعاة:

- إختلاف وحدة القياس بين الظواهر.
- إختلاف المتوسط لكل ظاهرة وطبيعة توزيع المفردات حول المتوسط في كل ظاهرة.

وإجمالا يمكن القول أن استخدام مقاييس التشتت النسبية يكون لما نكون بصدد إجراء مقارنة بين الظواهر التي ليس لها نفس وحدة القياس أو ليس لها نفس المتوسط.

وتنتج مقاييس التشتت النسبية من حاصل قسمة مقياس التشتت المطلق على مقياس التوسط الذي حُسب حوله التشتت وتحويل هذا الناتج إلى نسبة مئوية كما يلي [3]: $\text{قيمة مقياس التشتت النسبي} = \frac{\text{قيمة مقياس التشتت المطلق}}{\text{مقياس التوسط المستخدم}} \times 100\%$

ومن أهم مقاييس التشتت النسبية نجد :

1– المدى النسبي: يرمز له بالرمز $(R\%)$ ويحسب بـــ : $R\% = \frac{R}{\bar{X}} \times 100\%$

2– المدى الربيعي النسبي: يرمز له بالرمز $(RQ\%)$ ويحسب بـــ [4] : $RQ\% = \frac{RQ}{M_e} \times 100\%$

3– الإنحراف المتوسط النسبي: يرمز له بالرمز $(MD\%)$ ويحسب بـــ : $MD\% = \frac{MD}{\bar{X}} \times 100\%$

4– الإنحراف المعياري النسبي (معامل الاختلاف) : يرمز له بالرمز (CV) ويحسب بـــ:

$$CV = \frac{\sigma(X)}{\bar{X}} \times 100\%$$

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 209.
[2] – أحمد فوزي ملوخيه ، مرجع سابق، ص 230.
[3] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 124.
[4] – أحمد فوزي ملوخيه ، مرجع سابق، ص 231.

مثال 13: يرغب أحد الباحثين في مقارنة تشتت ظاهرتي الوزن والطول لطلبة كلية العلوم الاقتصادية والتجارية وعلوم التسيير ، وقد أوضحت عينة عشوائية من 100 طالب أن :

- الوزن المتوسط هو : 75 كلغ.
- الانحراف المعياري للوزن هو : 7.12 كلغ

بينما كان التوزيع التكراري للطول بالسنتمتر كما يلي:

فئات الطول	]160 -150]	]170 -160]	]180 -170]	]190 -180]	]200 -190]	المجموع
عدد الطلبة	20	25	40	10	5	**100**

المطلوب: قارن بين تشتت الوزن وتشتت الطول.

الحل: كما أوضحنا سابقا فإنه لا نستطيع المقارنة بين الظاهرتين إلا من خلال مقياس تشتت نسبي بحكم إختلاف الظاهرتين في وحدة القياس (وزن، طول) وعليه من خلال معامل الاختلاف CV يمكن المقارنة بين تشتت الظاهرتين.

لدينا الوزن المتوسط والإنحراف المعياري لظاهرة الوزن ، ونحسب الطول المتوسط والإنحراف المعياري للطول كمايلي :

الجدول رقم (5-5) : توزيع الطلبة حسب الطول

فئات الطول	عدد الطلبة $\mathbf{n_i}$	مراكز الفئات X_i'	$n_i X_i'$	$(X_i' - \bar{X})^2$	$n_i(X_i' - \bar{X})^2$
]160 -150]	20	155	3100	240.25	4805
]170 -160]	25	165	4125	30.25	756.25
]180 -170]	40	175	7000	20.25	810
]190 -180]	10	185	1850	210.25	2102.5
]200 -190]	5	195	975	600.25	3001.25
المجموع	**100**	/	**17050**		**11475**

متوسط الطول هو : $\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i'}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i X_i'}{100} = \frac{17050}{100} = 170.5$

الإنحراف المعياري للطول:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{\sum_{i=1}^{5} n_i (X_i' - \bar{X})^2}{100}} = \sqrt{\frac{11475}{100}} = 10.71$$

إذن نحسب معامل الإختلاف للوزن :

$$CV_1 = \frac{\sigma(X)}{\bar{X}} \times 100\% = \frac{7.12}{75} \times 100\% = 9.49\%$$

و نحسب معامل الإختلاف للطول :

$$CV_2 = \frac{\sigma(X)}{\bar{X}} \times 100\% = \frac{10.71}{170.5} \times 100\% = 6.28\%$$

بما : $CV_2 < CV_1$: فإن تشتت الطلبة في الوزن أكبر من تشتت الطلبة في الطول ، أي أن الطلبة أكثر إختلافا في الأوزان منه في الأطوال .

المحور السادس: مقاييس الشكل

تمهيد:

أوضحنا سابقا أن تلخيص بيانات أي ظاهرة يحتاج إلى مقاييس عددية تصف البيانات وهي مقاييس النزعة المركزية، ثم أدركنا أهمية مقاييس التشتت التي تعطينا فكرة عن درجة التجانس أو الإختلاف بين قيم الظاهرة، حيث بتوافر كل من مقاييس النزعة المركزية ومقاييس التشتت أتاح لدينا وصفا مقبولا لمختلف التوزيعات، غير أن هذا الوصف تنقصه الدقة الكافية والمطلوبة للتعرف على خواص توزيع أي ظاهرة لأنه يمكن أن نجد بيانات لظاهرتين معينتين يتساويان في المتوسط وفي الإنحراف المعياري مثلا ولكن يختلفان في الشكل، وهو ما يتطلب الإستعانة بمقاييس أخرى لوصف البيانات بشكل دقيق، وهذه المقاييس تسمى بمقاييس الشكل، حيث يكون إهتمامنا باختلاف الشكل من ناحيتين هما : الإلتواء و التفرطح .

وقبل التطرق إلى كل ما يخص الإلتواء والتفرطح نعرج في البداية على ما يعرف بـــ "العزوم" التي تساعدنا في استخدام مقاييس الشكل .

أولا – العزوم :

إن العزم لأي قوة هو مقدار العمل الذي تحدثه هذه القوة، ويتوقف ذلك العمل على عنصرين هما، القوة نفسها، والمسافة بين هذه القوة والنقطة التي عندها تُحدث أثرها، ذلك أن قوة مقدارها (08 كيلو غرام) على بعد (متر واحد) تعادل في مفعولها قوة أخرى مقدارها (02 كيلو غرام) على بعد (04 متر)، فطبقا لقانون الرافعة يتحقق التعادل (أو التوازن) عند تساوي طرفي هذا القانون[1].

فالعزوم إذن هي مصطلح فيزيائي يستعمل في علم الميكانيك و قد إستعمله الإحصائيون بمعنى قريب من المعنى الفيزيائي، فالعزم يقيس تأثير القيم أو مراكز الفئات أو قواها موزونة بتكراراتها النسبية في الإنسحاب عن نقطة الأصل أو مركز معين [2].

و وفقا للأساس السابق فإنه بالنسبة للتوزيعات التكرارية، فإن تكرارات أي توزيع تكون هي القوة المؤثرة عليه، وبالتالي عزوم أي تكرارات تُقاس بحاصل ضرب هذه التكرارات في إنحرافاتها عن نقطة معينة، و قد

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 241.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص 135.

تكون هذه النقطة هي نقطة الوسط الحسابي، أو نقطة الصفر، أو عند قيمة ثابتة أخرى، مقسوما على تلك التكرارات [1]، وسنتطرق إلى نوعين من العزوم:

1 – العزوم البسيطة (العزوم حول نقطة الصفر): ونميز بين حالتيين :

1-1 العزوم البسيطة للبيانات غير المبوبة: العزم k حول نقطة الصفر(الأصل) للقيم : $X_1 , X_2 , X_3 \ldots\ldots , X_n$ هو [2]: $m_k = \frac{\sum_{i=1}^{n} X_i^k}{n}$

حيث :n : هو عدد القيم (البيانات) .

1-2 العزوم البسيطة للبيانات المبوبة: ونميز بين حالتين:

1-2-1 حالة متغير كمي منفصل: العزم k حول نقطة الصفر للقيم: X_k ,$\ldots\ldots$,X_3 , X_2 , X_1 و تكراراتها المقابلة على الترتيب : n_k ,$\ldots\ldots$,n_3 , n_2 , n_1 هو: $m_k = \frac{\sum_{i=1}^{k} n_i X_i^k}{\sum_{i=1}^{k} n_i}$

1-2-2 حالة متغير كمي متصل: العزم k حول نقطة الصفر لمراكز الفئات : X'_k,$\ldots\ldots$,X'_3 ,X'_2 ,X'_1 و تكراراتها المقابلة على الترتيب : n_k ,$\ldots\ldots$,n_3 , n_2 , n_1 هو: $m_k = \frac{\sum_{i=1}^{k} n_i X_i'^k}{\sum_{i=1}^{k} n_i}$

2 – العزوم المركزية (العزوم حول المتوسط الحسابي): ونميز بين حالتيين :

2-1 العزوم المركزية للبيانات غير المبوبة: العزم k للقيم : X_n ,$\ldots\ldots$,X_3 , X_2 , X_1 حول وسطها الحسابي هو [3]: $\mu_k = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^k}{n}$

حيث :n : هو عدد القيم (البيانات) .

2-2 العزوم المركزية للبيانات المبوبة: ونميز بين حالتين:

2-2-1 حالة متغير كمي منفصل: العزم k للقيم : X_k ,$\ldots\ldots$,X_3 , X_2 , X_1 و تكراراتها المقابلة على الترتيب : n_k ,$\ldots\ldots$,n_3 , n_2 , n_1 حول وسطها الحسابي [4]: $\mu_k = \frac{\sum_{i=1}^{n} n_i (X_i - \bar{X})^k}{\sum_{i=1}^{k} n_i}$

2-2-2 حالة متغير كمي متصل: العزم k لمراكز الفئات : X'_k,$\ldots\ldots$,X'_3 ,X'_2 ,X'_1 و تكراراتها المقابلة على الترتيب : n_k ,$\ldots\ldots$,n_3 , n_2 , n_1 حول وسطها الحسابي: $\mu_k = \frac{\sum_{i=1}^{n} n_i (X'_i - \bar{X})^k}{\sum_{i=1}^{k} n_i}$

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 241.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص ، 135.

[3] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 139.

[4] – جلاطو جيلالي ، مرجع سابق، ص 78.

ملاحظة: من خلال المعادلات السابقة نجد أن :

- العزم الأول حول المتوسط الحسابي يساوي دوما للصفر: $\mu_1 = 0$
- العزم الثاني حول المتوسط الحسابي يساوي التباين : $\mu_2 = V(X)$

ثانيا – الإلتواء :

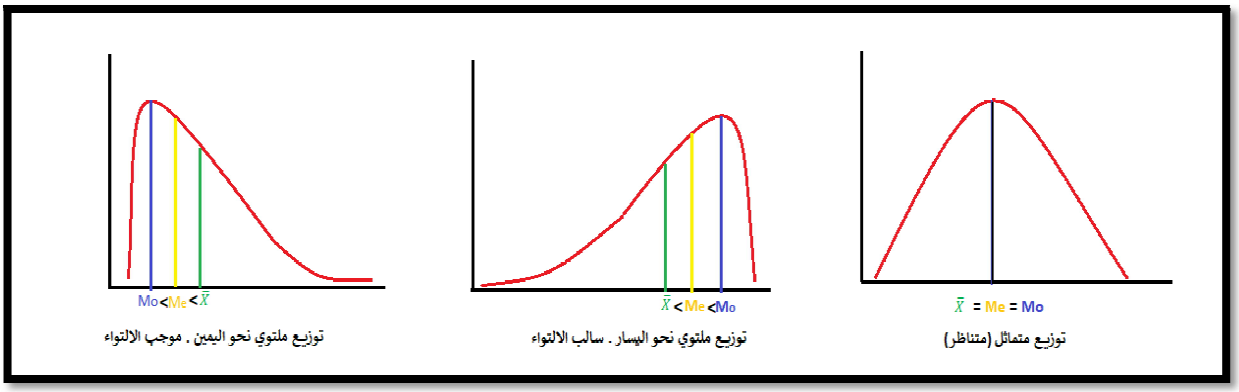
إلتواء توزيع ما يعني مدى بعده عن التماثل، وتكون المنحنيات التكرارية إما متماثلة، والمنحنيات غير المتماثلة تسمى (ملتوية) والإلتواء كما أشرنا هو بعد المنحى عن التماثل، وهو بمثابة قيمة تعطي فكرة عن تمركز قيم المتغير[1].

- فيكون التوزيع التكراري متماثلا إذا كانت التكرارات موزعة توزيعا متماثلا حول الوسط الحسابي، بمعنى أن تكون لقيم المتغير المتساوية البعد عن الوسط الحسابي نفس التكرارات [2]، فإذا أسقطنا عمودا من قمة المنحنى لوجدنا أن المنحنى قد إنشطر إلى نصفين متماثلين ومن هنا يطلق على المنحنى بأنه منحنى متماثل. ويعتبر منحنى التوزيع الطبيعي أشهر المنحنيات المتماثلة بل إن البعض قد يقرن صفة الطبيعي مع صفة التماثل[3]. و هنا يكون (الوسط الحسابي = الوسيط = المنوال).
- وعنما يكون التوزيع ملتويا نحو اليمين فإن القيم المتطرفة الكبيرة نحو اليمين تؤثر على الوسط الحسابي وتسحبه نحو اليمين [4]، وبذلك يكون (الوسط الحسابي > الوسيط > المنوال). وفي هذه الحالة نقول أن التوزيع ملتوي نحو اليمين موجب الالتواء.
- أما إذا كان التوزيع ملتويا نحو اليسار فإن القيم المتطرفة الصغيرة نحو اليسار تؤثر على الوسط الحسابي وتسحبه نحو اليسار وبذلك يكون (الوسط الحسابي < الوسيط < المنوال). وفي هذه الحالة نقول أن التوزيع ملتوي نحو اليسار سالب الالتواء.

والشكل الموالي يميز بين الأشكال الثلاثة :

[1] – صلاح الدين حسين الهيتي ، مرجع سابق، ص 128.
[2] – محمود محمد سليم صالح ، مرجع سابق ، ص 103.
[3] – أحمد فوزي ملوخيه ، مرجع سابق، ص 248.
[4] – محمد صبحي أبو صالح ، مرجع سابق ، ص 139.

الشكل رقم (6-1): يوضح الحالات الثلاثة لناحية الالتواء



توجد عدة طرق لقياس الإلتواء نذكر منها:

1 – مقارنة مقاييس النزعة المركزية: يمكن إستخدام مقاييس النزعة المركزية في معرفة شكل التوزيع من ناحية الإلتواء، حيث بعد حساب مقاييس النزعة المركزية الثلاثة تتم المقارنة بينها كما يلي :

- إذا كان : الوسط = الوسيط = المنوال، فمنحنى التوزيع متماثل (متناظر).
- إذا كان : الوسط > الوسيط > المنوال، فمنحنى التوزيع ملتوي جهة اليمين ، موجب الالتواء.
- إذا كان : الوسط < الوسيط < المنوال، فمنحى التوزيع ملتوي جهة اليسار ، سالب الإلتواء.

2 – قانون بيرسون للإلتواء: يستعمل هذا المقياس عادة إذا كان التوزيع قريب من التماثل و ليس شديد الإلتواء، فإذا كانت ($\bar{X}$) تمثل الوسط الحسابي، وكانت (M_e) تمثل الوسيط و (M_o) تمثل المنوال و $\sigma(X)$ تمثل الإنحراف المعياري فإن [1]:

1-2 قانون بيرسون الأول: يكون معطى في الصيغة التالية : $SK = \frac{\bar{X}-M_o}{\sigma(X)}$

2-2 قانون بيرسون الثاني: يكون معطى في الصيغة التالية : $SK = \frac{3(\bar{X}-M_e)}{\sigma(X)}$

ويحدد نوع الإلتواء هنا من خلال الحالات الثلاثة كما يلي :

- $SK = 0$ ، فمنحنى التوزيع متماثل (متناظر).
- $SK > 0$ ، فمنحنى التوزيع ملتوي جهة اليمين ، موجب الإلتواء.
- $SK < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الإلتواء.

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص ص 144- 145.

3 – معامل فيشر للإلتواء: يقيس هذا المعامل التواء شكل التوزيع الإحصائي، ويعتمد في ذلك على قيمة العزم المركزي من الدرجة الثالثة، ولإستبعاد وحدة القياس نقسمه على الإنحراف المعياري من نفس المرتبة [1]، ويرمز له بالرمز F_1 ، ويحسب بالعلاقة :

3-1 حالة البيانات غير المبوبة[2]: $F_1 = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^3}{n}}{[\delta(X)]^3}$

حيث n: هوعدد القيم

3-2 حالة البيانات المبوبة: ونميز فيها بين :

3-2-1 حالة متغير كمي منفصل: $F_1 = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{\sum_{i=1}^{k} n_i(X_i-\bar{X})^3}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^3}$

3-2-2 حالة متغير كمي متصل: $F_1 = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{\sum_{i=1}^{k} n_i(X'_i-\bar{X})^3}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^3}$ حيث : X'_i: تمثل مراكز الفئات

ويحدد نوع الإلتواء هنا من خلال الحالات الثلاثة كما يلي :

- $F_1 = 0$ ، فمنحنى التوزيع متماثل (متناظر).
- $F_1 > 0$ ، فمنحنى التوزيع ملتوي جهة اليمين ، موجب الالتواء.
- $F_1 < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الالتواء.

مثال 01: الجدول التالي يمثل الدخل (بالالف دينار) وساعات العمل الأسبوعية في أحد المصانع

فئات الدخل	]20 -10[	]30 -20[	]40 -30[	]50 -40[	]60 -50[	المجموع
عدد الساعات (التكرارات)	200	250	300	350	400	**1500**

المطلوب:

1-أحسب مقاييس النزعة المركزية الثلاث ثم قارن بينها لمعرفة شكل التوزيع من ناحية الإلتواء.

2-من خلال قانون بيرسون الأول والثاني وكذا معامل فيشر بين نوع الإلتواء حسب كل قانون.

[1] – جلاطو جيلالي ، مرجع سابق، ص 78.

[2] – عبد الحميد قطوش ، مرجع سابق ، ص 57.

الحل:

نكون الجدول التالي :

الجدول رقم (6-1) : توزيع ساعات العمل حسب الدخل

فئات الدخل	التكرارات n_i	مراكز الفئات X'_i	$n_i X'_i$	$N_i^{\uparrow}$	$(X'_i - \bar{X})^2$	$n_i(X'_i - \bar{X})^2$	$(X'_i - \bar{X})^3$	$n_i(X'_i - \bar{X})^3$
]20 -10]	200	15	3000	200	544.2889	108834.45	-12698.26004	-2539652.008
]30 -20]	250	25	6250	450	177.6889	44408.895	-2368.593037	-592148.2593
]40 -30]	300	35	10500	750	11.0889	3323.34	-36.926037	-11077.8111
]50 -40]	350	45	15750	1100	44.4889	15577.785	296.740963	103859.3371
]60 -50]	400	55	22000	1500	277.8889	111172.23	4632.407963	1852963.185
المجموع	**1500**	/	**57500**	/	/	**283316.7**	/	**-1186055.556**

1-حساب مقاييس النزعة المركزية :

- الوسط الحسابي: $\overline{X} = \frac{\sum_{i=1}^{k} n_i X'_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{5} n_i X'_i}{1500} = \frac{57500}{1500} = 38.33$

- الوسيط :

نحدد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N_{M_e}^{\uparrow} \geq \frac{n}{2} = \frac{1500}{2} = 750$ ومنه الفئة الوسيطية هي :]30- 40].

نحسب الوسيط من العلاقة :

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times C_{M_e} = 30 + \left[\frac{750 - 450}{300}\right] \times 10 = 40$$

- المنوال :

نحدد الفئة المنوالية : لدينا أكبر تكرار هو 400 وبالتالي الفئة المنوالية هي]50- 60]

$$\Delta_1 = 400 - 350 = 50$$
$$\Delta_2 = 400 - 0 = 400$$

ومنه: قيمة المنوال هي : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times C_{M_o} = 50 + \left[\frac{50}{50+400}\right] \times 10 = 51.11$

بما أن : $\bar{X} = 38.33 < M_e = 40 < M_o = 51.11$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الإلتواء.

2-حساب قانون بيرسون الأول والثاني وكذا معامل فيشر وتبيين نوع الإلتواء حسب كل قانون :

- قانون بيرسون الأول:

حساب الإنحراف المعياري : $\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{283316.7}{1500}} = \sqrt{188.8778} = 13.74$

ويصبح: $SK = \frac{\bar{X} - M_o}{\sigma(X)} = \frac{38.33 - 51.11}{13.74} = -0.93$

بما أن : $SK < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الإلتواء.

– قانون بيرسون الثاني: $SK = \frac{3(\bar{X} - M_e)}{\sigma(X)} = \frac{3(38.33 - 40)}{13.74} = -0.36$

بما أن : $SK < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الإلتواء.

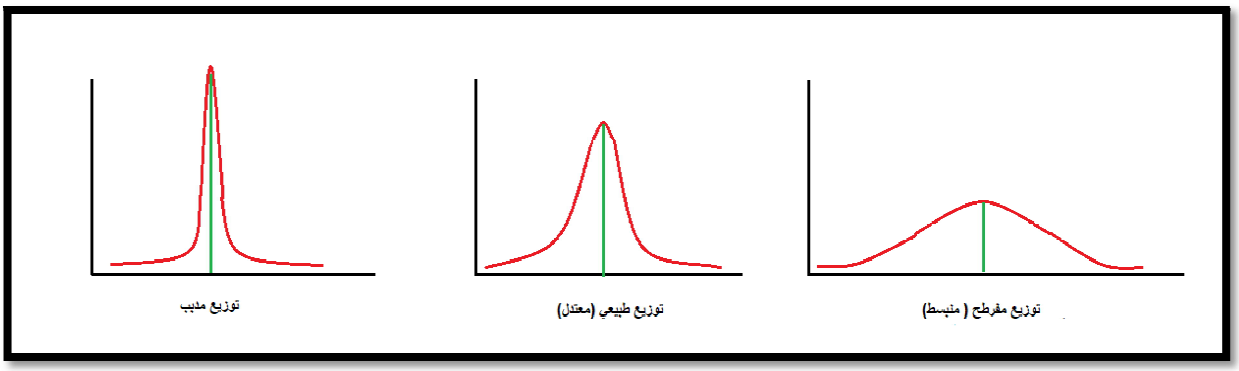
–معامل فيشر للإلتواء:

$$F_1 = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{\sum_{i=1}^{k} n_i (X_i' - \bar{X})^3}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^3} = \frac{\frac{-1186055.556}{1500}}{[13.74]^3} = -0.30$$

بما أن : $F_1 < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الإلتواء.

ثالثا – التفرطح : المقصود بالتفرطح هو مقدار أو درجة تدبب قمة منحنى التوزيع و الذي عادة ما يقارن بالتوزيع الطبيعي، أو هو قياس درجة علو أو إنخفاض قمة التوزيع بالنسبة للتوزيع الطبيعي و تقسم المنحنيات التكرارية من حيث تفرطح قمتها الى ثلاثة أنواع [1]، كما يوضحه الشكل التالي :

الشكل رقم (6–2): يوضح الحالات الثلاثة للتفرطح



تجدر الاشارة إلى أنه من خلال الشكل السابق يكون التشتت كبيرا جدا عندما يكون منحنى التوزيع مفرطح ، ويكون تشتت ضعيف جدا عندما يكون منحنى التوزيع مدبب، ويمكن قياس التفرطح باستخدام عدد من الطرق، وأهمها :

1 – معامل بيرسون للتفرطح: وهو عبارة عن النسبة بين العزم المركزي من المرتبة الرابعة والإنحراف المعياري مرفوعا إلى القوة أربعة، ومعامل التفرطح يساوي 3 في حالة التوزيع الطبيعي أي التوزيع على الشكل الجرسي، ويرمز لهذا المعامل بالرمز K ، ويحسب بالعلاقة [1]:

[1] – محمود محمد سليم صالح ، مرجع سابق ، ص 107.

1-1 حالة البيانات غير المبوبة: $K = \frac{\mu_4}{[\delta(X)]^4} = \frac{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^4}{n}}{[\delta(X)]^4}$

حيث n: هوعدد القيم

2-1 حالة البيانات المبوبة: ونميز فيها بين :

1-2-1 حالة متغير كمي منفصل: $K = \frac{\mu_4}{[\delta(X)]^4} = \frac{\frac{\sum_{i=1}^{k} n_i(X_i-\bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4}$

2-2-1 حالة متغير كمي متصل: $K = \frac{\mu_4}{[\delta(X)]^4} = \frac{\frac{\sum_{i=1}^{k} n_i(X'_i-\bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4}$ حيث : X'_i: تمثل مراكز الفئات .

ويحدد نوع التفرطح هنا من خلال الحالات الثلاثة كما يلي :

- $K = 3$ ، فمنحنى التوزيع طبيعي (معتدل).
- $K > 3$ ، فمنحنى التوزيع مدببا (متطاولا).
- $K < 3$ ، فمنحنى التوزيع مفرطح (منبسط).

2 – معامل فيشر للتفرطح: وهو عبارة عن معامل بيرسون مطروحا منه 3 ، ويرمز له بالرمز F_2 [2] ، و تكون العلاقات السابقة كالتالي:

1-2 حالة البيانات غير المبوبة: $F_2 = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^4}{n}}{[\delta(X)]^4} - 3$

حيث n: هوعدد القيم

2-2 حالة البيانات المبوبة: ونميز فيها بين :

1-2-2 حالة متغير كمي منفصل: $F_2 = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{\frac{\sum_{i=1}^{k} n_i(X_i-\bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4} - 3$

2-2-2 حالة متغير كمي متصل: $F_2 = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{\frac{\sum_{i=1}^{k} n_i(X'_i-\bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4} - 3$ حيث : X'_i: تمثل مراكز الفئات .

ويحدد نوع التفرطح هنا من خلال الحالات الثلاثة كما يلي :

- $F_2 = 0$ ، فمنحنى التوزيع طبيعي (معتدل).

[1] – عبد الحكيم بيصار ، مرجع سابق، ص 54 .

[2] – نفس المرجع، ص 55 .

- $F_2 > 0$ ، فمنحنى التوزيع مدببا (متطاولا).
- $F_2 < 0$ ، فمنحنى التوزيع مفرطح (منبسط).

مثال 02: باستخدام معطيات المثال 02 حدد شكل التوزيع من ناحية التفرطح باستخدام معامل بيرسون للتفرطح ومعامل فيشر للتفرطح

الحل:

نكون الجدول التالي :

الجدول رقم (6-2) : توزيع ساعات العمل حسب الدخل

فئات الدخل	التكرارات n_i	مراكز الفئات X'_i	$(X'_i - \bar{X})^4$	$n_i(X'_i - \bar{X})^4$
]20 -10]	200	15	296250.4067	59250081.33
]30 -20]	250	25	31573.34518	7893336.296
]40 -30]	300	35	122.9637032	36889.11096
]50 -40]	350	45	1979.262223	692741.7781
]60 -50]	400	55	77222.24074	30888896.3
المجموع	**1500**	/		**98761944.82**

-معامل بيرسون للتفرطح: $K = \frac{\mu_4}{[\delta(X)]^4} = \frac{\frac{\sum_{i=1}^{k} n_i (X'_i - \bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4} = \frac{\frac{98761944.82}{1500}}{[13.74]^4} = 1.85$

- بما أن : $K < 3$: فمنحنى التوزيع مفرطح (منبسط).

-معامل فيشر للتفرطح: $F_2 = k - 3 = 1.85 - 3 = -1.15$

- بما أن : $F_2 < 0$: فمنحنى التوزيع مفرطح (منبسط).

المحور السابع: مقاييس التمركز

تمهيد:

مقاييس التمركز هي مقاييس إحصائية تطبق على التوزيعات التكرارية (الجداول) التي يكون موضوعها توزيع ثروة معينة (دخل، أجور، أراضي، ...الخ) ، على فئات متفاوتة (فئات من الأسر، فئات من العمال، فئات من المزارعين...الخ) ، لكي نعرف هل هناك تمركز في توزيع هذه الثروة أي هل هناك توزيع عادل أو غير عادل لهذه الثروة على مختلف هذه الفئات [1]. ونعني بالعدالة أو المساواة فيما سبق من حيث الدخل أو الملكية للأراضي الزراعية هو أن الدخل ككل أو الأراضي الزراعية ككل موزعة على كافة السكان ككل أيضا أي بنسب متناظرة ، ونقصد بذلك أن نسبة معينة من الدخل ولتكن 25% مثلا يحصل عليها 25% من السكان، وهكذا 75% من الدخل يحصل عليها 75% من السكان أو 90% من الدخل يحصل عليها 90% من السكان وهكذا الأمر بالنسبة لتوزيع ملكية الأراضي الزراعية كظاهرة ثانية أو لأي ظاهرة أخرى يراد معرفة وضعها التوزيعي [2].

إن مؤشر التمركز هو مقياس للتشتت النسبي لوصف بعض التوزيعات الإحصائية ، حيث لا يطبق إلا على المتغيرات الاحصائية المستمرة ذات القيم الموجبة، وتتم دراسة قضية التمركز بطريقتين [3]:

- طريقة بيانية بواسطة مربع جيني للتمركز.
- طريقة حسابية بواسطة مؤشر جيني للتمركز.

أولا – التحديد البياني للتمركز (مربع جيني):

إن منحنى لورونز "***Lorenz Curve***" الذي يُنسب لعالم الإحصاء الأمريكي كونراد لورانز "***conrad lorenz***" الذي بين سنة 1905 كيفية توضيح العلاقة بين السكان والدخل والمساحة ...إلخ، أي أنه يحاول التعرف على درجة البعد وتركز السكان وهل تركز السكان مثالي [4]. يعدّ أهم طريقة لتمثيل التوزيع بيانيا، حيث يستخدم منحنى لورنز لغرض تمثيل التفاوت في توزيع الدخل أو الإنفاق أو متغيرات أخرى بشكل بياني، ولدراسة التمركز بيانيا وتمثيل منحنى لورونز نقوم بمجموعة من الخطوات هي:

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 63.

[2] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 220.

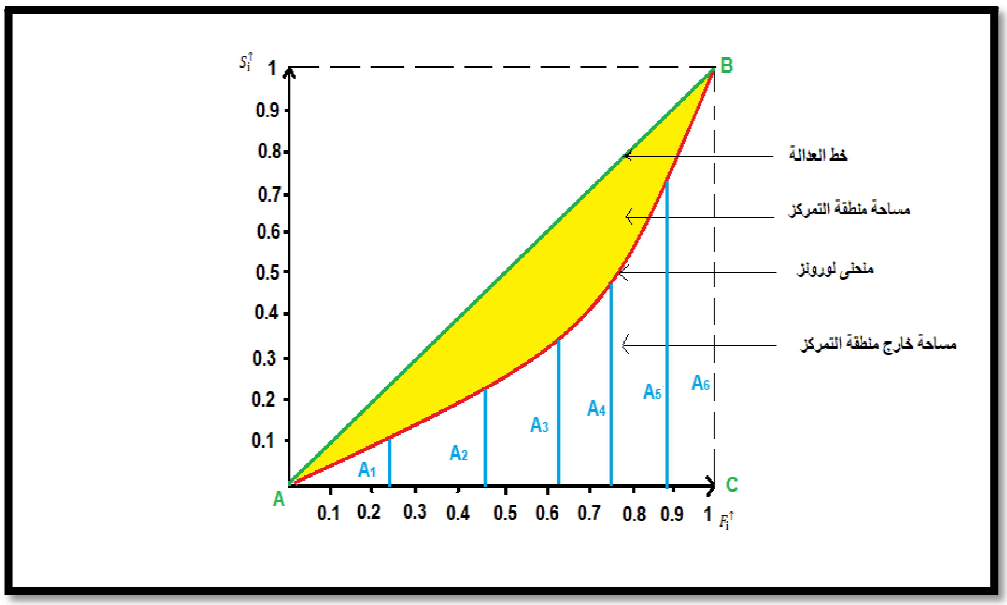
[3] – سهام بشكيط ، محاضرات في مقياس الاحصاء 1، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة جيجل، السنة الجامعية 2021-2022 ، ص 133.

[4] – ميدون سيساني ، مطبوعة دروس بعنوان "التحليل الديمغرافي" محاضرات في مقياس الاحصاء 1، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة تيارت، السنة الجامعية 2019-2020 ، ص 79.

- حساب التكرارات النسبية المتجمعة الصاعدة $F_i^{\uparrow}$.
- حساب النسب المتجمعة الصاعدة $S_i^{\uparrow}$ بعد حساب النسب s_i من العلاقة : $s_i = \frac{n_i X_i'}{\sum n_i X_i'}$
- نرسم مربع طول ضلعه 100 ملم موجه على معلم متعامد ومتجانس ، وهذا المربع طول ضلعه يمثل 100 % حيث إن الإحداثي الأفقي له يمثل التكرارات النسبية المتجمعة الصاعدة $F_i^{\uparrow}$ و الإحداثي العمودي يمثل النسب المتجمعة الصاعدة $S_i^{\uparrow}$ [1].
- نصل القطر الرئيسي للشكل أي النقطتين [(0 ، 0) ، (100 ، 100)] وهما نقطة الأصل والنهاية في الشكل بخط مستقيم (AB) يطلق عليه خط العدالة وهو الخط الأمثل للتوزيع . كما نعين المثلث (ABC).
- نرسم منحنى لإحداثيات النقاط($F_i^{\uparrow}$،$S_i^{\uparrow}$) الخاصة بالمنحنى النسبي المتجمع الصاعد للمتغيرين [2].

ويتم إعطاء صورة تقريبية على الرسم في الشكل التالي :

الشكل رقم (7-1): مربع جيني



[1] – طارق الحاج ، مشكلات التنمية في الدول النامية، المحور الرابع ، متاح على الرابط :
https://faculty.ksu.edu.sa/sites/default/files/ pdf : تاريخ الاطلاع : 2024/10/20 على الساعة 14.00

[2] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 221.

إن المساحة المحصورة بين المنحنى والخط الأمثل للتوزيع تعبر عن عدم العدالة أو عدم المساواة وتسمى "بمنطقة التمركز" فكلما صغرت هذه المساحة دل ذلك على الإقتراب من العدالة في التوزيع ، والعكس صحيح ، وعليه يمكن أن يأخذ منحنى لورونز الحالات التالية [1]:

- إذا إنطبق منحنى لورونز على خط العدالة (AB) تنعدم مساحة التمركز وبالتالي نقول أن توزيع الثروة عادل تماما (إنعدام التمركز).
- إذا إنطبق منحنى لورونز على المثلث (ABC) تكون مساحة التمركز تساوي مساحة المثلث وبالتالي نقول أن توزيع الثروة جائر تماما (تمركز كلي) .
- اذا إقترب منحنى لورونز من خط العدالة (AB) تقل مساحة التمركز مقارنة بمساحة المثلث (ABC) وبالتالي نقول أن التوزيع أكثر عدالة .
- إذا إبتعد منحنى لورونز عن خط العدالة (AB) تكبر مساحة التمركز مقارنة بمساحة المثلث (ABC) وبالتالي نقول أن التوزيع أقل عدالة .

ثانيا – التحديد الحسابي للتمركز (حساب معامل جيني للتمركز):

معامل جيني في علم الاقتصاد الذي يُطلق عليه أحيانا مؤشر جيني أو نسبة جيني، هو مقياس للتشتت الإحصائي الذي يهدف إلى تمثيل عدم المساواة في الدخل أو عدم المساواة في الثروة داخل دولة أو أي مجموعة أخرى من الناس. تم تطويره بواسطة الإحصائي وعالم الاجتماع الإيطالي كواردو جيني ونشر في مقالته التي صدرت عام 1912 [2].

ويمكن حساب معامل جيني بعدة طرق منها:

1 – الطريقة الأولى : جنبا إلى جنب مع منحنى لورنز، يمكننا حساب مؤشر جيني ببساطة ، حيث نقوم بتقسيم المساحة بين منحنى لورونز وخط العدالة على إجمالي المساحة التي تظل تحت المنحنى، بهذه الطريقة نحصل على المعامل [3]، إذا كانت البيانات معبر عنها بقيم نسبية:

$$I_{Gini} = \frac{\text{مساحة التمركز S}}{\text{مساحة المثلث } ABC} = \frac{S}{1/2} = 2S \quad \text{.....(01)}$$

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 63.

[2] – عبد الحليم الحمزة ، كتاب بيداغوجي في الاحصاء 1 ، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة تبسة، الموسم الجامعي 2022-2023، ص 92.

[3] – نفس المرجع، نفس الصفحة.

أما إذا كانت البيانات معبر عنها بالنسب المئوية، فإن معامل جيني يكون : $I_{Gini} = \frac{\text{مساحة التمركز S}}{\text{مساحة المثلث ABC}} = \frac{S}{5000}$

2 – الطريقة الثانية : (طريقة شبه المنحرف) وذلك كما يلي [1]:

$$S = \frac{1}{2} - (A_1 + A_2 + A_3 + \cdots + A_k)$$

حيث : $A_1, A_2, A_3, \ldots, A_k$ تمثل مساحات أشباه المنحرف تحسب كل واحدة منها بالعلاقة التالية :

$$A_i = \frac{\text{القاعدة الصغرى+القاعدة الكبرى}}{2} \times \text{الارتفاع}$$

$$S = \frac{1}{2} - \left(\frac{s_1^{\uparrow} \times f_1}{2} + \frac{(s_2^{\uparrow} + s_1^{\uparrow}) f_2}{2} + \frac{(s_3^{\uparrow} + s_2^{\uparrow}) f_3}{2} + \cdots + \frac{(s_k^{\uparrow} + s_{k-1}^{\uparrow}) f_k}{2}\right)$$

$$S = \frac{1}{2} - \left(\frac{(s_1^{\uparrow} + s_0^{\uparrow}) f_1}{2} + \frac{(s_2^{\uparrow} + s_1^{\uparrow}) f_2}{2} + \frac{(s_3^{\uparrow} + s_2^{\uparrow}) f_3}{2} + \cdots + \frac{(s_k^{\uparrow} + s_{k-1}^{\uparrow}) f_k}{2}\right)$$

حيث: $s_0^{\uparrow} = 0$

$$S = \frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k} f_i (s_i^{\uparrow} + s_{i-1}^{\uparrow})$$

بالتعويض في العلاقة (01) نجد: $I_{Gini} = 2\left(\frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k} f_i (s_i^{\uparrow} + s_{i-1}^{\uparrow})\right)$

وبالتالي معامل جيني اذا كانت البيانات معبر عنها بقيم نسبية يكون: $I_{Gini} = 1 - \sum_{i=1}^{k} f_i (s_i^{\uparrow} + s_{i-1}^{\uparrow})$

حيث : f_i :تمثل التكرار النسبي. $s_i^{\uparrow}$: تمثل النسبة المتجمعة الصاعدة و: $s_0^{\uparrow} = 0$

و يستخلص مما سبق أن قيمة معامل جيني المعبر عنها تقع بين صفر و1 ، وكلما كانت قيمة هذا المعامل أعلى كان توزيع الدخل أكثر تفاوتا، حيث يأخذ معامل جيني الحالات التالية:

- $I_{Gini} = 0$ هذا يعني أن منحنى لورونز ينطبق على خط العدالة (AB) وبالتالي فالتوزيع عادل تماما.
- $I_{Gini} = 1$ هذا يعني أن منحنى لورونز ينطبق على المثلث (ABC) وبالتالي فالتوزيع جائر تماما.
- $I_{Gini} \leq 0.3$ هذا يعني أن منحنى لورونز يقترب من خط العدالة (AB) وبالتالي فالتوزيع أكثر عدالة .
- $I_{Gini} > 0.3$ هذا يعني أن منحنى لورونز يبتعد من خط العدالة (AB) وبالتالي فالتوزيع أقل عدالة .

–ولحسابه إذا كانت البيانات معبر عنها بالنسب المئوية تكون العلاقة:

$$I_{Gini} = 1 - \frac{1}{10000}\sum_{i=1}^{k} f_i\%(s_i^{\uparrow}\% + s_{i-1}^{\uparrow}\%)$$

[1] – عبد الحميد قطوش ، مرجع سابق ، ص 67.

مثال 01: الجدول التالي يمثل توزيع العمال حسب أجورهم الشهرية بالدينار الجزائري في إحدى المؤسسات الجزائرية لسنة 1990

الاجور الشهرية	[3000 -2000]	[4000 -3000]	[5000 -4000]	[6000 -5000]	[7000 -6000]	[8000 -7000]	المجموع
عدد العمال	28	60	40	40	30	02	**200**

المطلوب:

1-أدرس قضية تمركز الأجور بيانيا وحسابيا .

2-في دراسة مماثلة على مؤسسة أخرى، قدرت مساحة خارج منطقة التمركز بـــ 1800 ملم^2 ، ما مدى عدالة هذا التوزيع .

الحل:

1-دراسة قضية التمركز بيانيا وحسابيا :

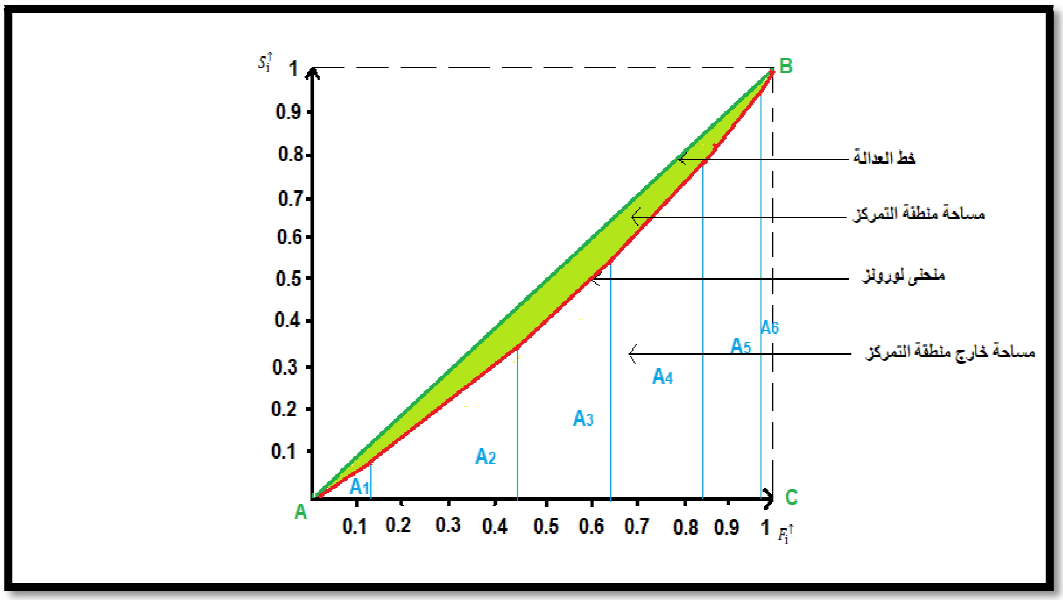
1-1 بيانيا : ويتم ذلك بواسطة منحنى لورونز حيث نقوم أولا بحساب $F_i^{\uparrow}$ و $S_i^{\uparrow}$ ، ثم نقوم برسم منحنى لورونز عن طريق وصل النقاط التي احداثياتها ($F_i^{\uparrow}$ ، $S_i^{\uparrow}$) ببعضها ضمن مربع جيني .

الجدول رقم (7-1) : جدول تحضير الحسابات لتوزيع العمال حسب الأجور الشهرية لرسم منحنى لورونز

الاجور الشهرية (دج)	عدد العمال n_i	مراكز الفئات X_i'	$n_i X_i'$	نسبة العمال f_i	نسب الاجر الشهري s_i	$F_i^{\uparrow}$	$S_i^{\uparrow}$
] 3000 -2000]	28	2500	70000	0.14	0.079	0,14	0.079
] 4000 -3000]	60	3500	210000	0.3	0.236	0,44	0.315
] 5000 -4000]	40	4500	180000	0.2	0.202	0,64	0.517
] 6000 -5000]	40	5500	220000	0.2	0.247	0,84	0.764
] 7000 -6000]	30	6500	195000	0.15	0.219	0,99	0.983
] 8000 -7000]	02	7500	15000	0.01	0.017	1	1
المجموع	**200**	/	**890000**	**1**	/	/	/

ويكون منحنى لورونز في الشكل التالي :

الشكل رقم (7-2): مربع جيني لتوزيع العمال حسب الأجور الشهرية



نلاحظ من خلال الشكل أن منحنى لورونز يقترب نوعا ما من خط العدالة ومنه فإن توزيع العمال حسب الأجر الشهري أكثر عدالة .

2-1 حسابيا: ويتم ذلك بواسطة حساب معامل جيني ، وذلك كمايلي:

الجدول رقم (7-2) : جدول تحضير الحسابات لتوزيع العمال حسب الأجور الشهرية لحساب معامل جيني

$f_i(s_i^{\uparrow} + s_{i-1}^{\uparrow})$	$s_i^{\uparrow} + s_{i-1}^{\uparrow}$	$S_i^{\uparrow}$	f_i
0.01106	0.079	0.079	0,14
0.1182	0.394	0.315	0,3
0.1664	0.832	0.517	0,2
0.2562	1.281	0.764	0,2
0.26205	1.747	0.983	0,15
0.01983	1.983	1	0.01
0.8337	/	/	**1**

$I_{Gini} = 1 - \sum_{i=1}^{6} f_i(s_i^{\uparrow} + s_{i-1}^{\uparrow})$

$I_{Gini} = 1 - 0.8337$

$I_{Gini} = 0.1663 = 16.63\%$

نلاحظ أن $I_{Gini} = 16{,}63\%$ وهي أقل من 30% وعليه يمكن القول أن هذا التوزيع أكثر عدالة، حيث توجد هناك فعلا فوارق في توزيع الأجر الشهري على مختلف فئات العمال ، لكن هذه الفوارق يمكن تفسيرها بأسباب موضوعية.

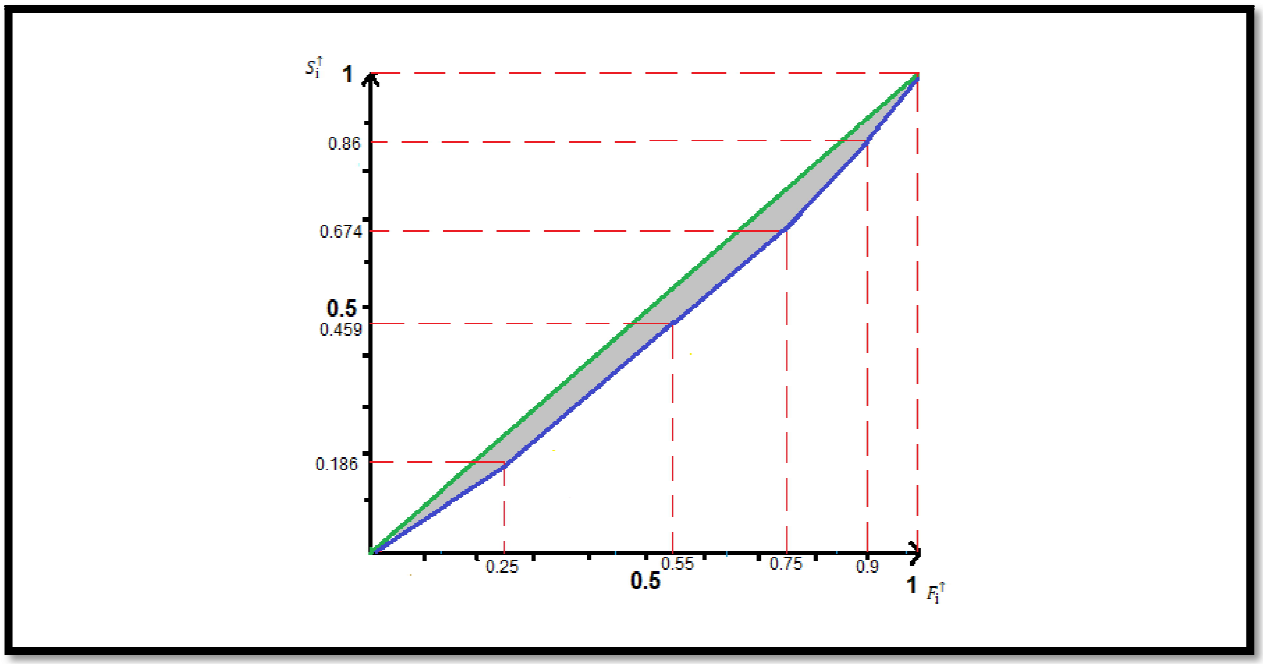
2-مدى عدالة التوزيع في الدراسة الثانية :

- مساحة منطقة التمركز(S) = 5000 – مساحة خارج منطقة التمركز
- مساحة منطقة التمركز(S) = 5000 – 1800 = 3200 ملم^2.

ومنه: $I_{\text{Gini}} = \frac{\text{S}}{\text{ABC}} = \frac{3200}{5000} = 0{,}54 = 54\%$

-بما أن $I_{Gini} > 0{,}3$ فإن توزيع الأجور الشهرية في الدراسة الثانية أقل عدالة.

مثال 02: إليك الشكل التالي :



حيث $F_i^{\uparrow}$: يمثل التكرار التجميعي الصاعد النسبي للعمال للمؤسسة A، $S_i^{\uparrow}$: تمثل النسبة المتجمعة الصاعدة للأجور اليومية للعمال للمؤسسة A.

إليك الجدول التالي :

الأجر اليومي	]50 -40]	]60 -50]	]70 -60]	]80 -70]	]90 -80]
$F_i^{\uparrow}$	؟	؟	؟	؟	؟
$S_i^{\uparrow}$	؟	؟	؟	؟	؟
f_i	؟	؟	؟	؟	؟

المطلوب:

1- ماذا يمثل الشكل البياني السابق؟ ماذا نستنتج بالنسبة لمدى عدالة التوزيع؟

2- أكمل بيانات الجدول أعلاه، ثم أكد على إستنتاجك الذي أعطيته في السؤال الأول حسابيا.

3- إفترض انه في دراسة مماثلة تم دراسة التمركز على توزيع مؤسسة أخرى B، حيث قدرت المساحة خارج منطقة التمركز بـــ 3800 ملم^2 ، حدد أي التوزيعين أكثر عدالة.

الحل:

1-الشكل البياني السابق يمثل : منحنى لورونز لتوزيع العمال حسب أجرهم اليومي.

- نستنتج مايلي :
- لقد تكونت منطقة تمركز على الشكل وبالتالي فإنه يوجد هناك فوارق بين مختلف الفئات.
- و بما أن منحنى لورونز يقترب نوعا ما من خط العدالة، وإذا ما قارنا مساحة منطقة التمركز بمساحة المثلث A B C بالعين المجردة نجد أنها صغيرة ، وبالتالي يمكن القول أن الفوارق في توزيع الدخل مقبولة أي أن التوزيع أكثر عدالة.

2-إكمال بيانات الجدول، ثم التأكيد على الإستنتاج المعطى في السؤال الأول حسابيا:

الأجر اليومي	[40- 50[	[50- 60[	[60- 70[	[70- 80[	[80- 90[
$F_i^{\uparrow}$	0.25	0.55	0.75	0.90	1.00
$S_i^{\uparrow}$	0.186	0.459	0.674	0.86	1.00
f_i	0.25	0.30	0.20	0.15	0.10

- التأكيد حسابيا : يكون من خلال حساب معامل جيني :

الجدول رقم (7-3) : جدول تحضير الحسابات لتوزيع العمال حسب الأجور اليومية لحساب معامل جيني

f_i	$S_i^{\uparrow}$	$S_i^{\uparrow} + S_{i-1}^{\uparrow}$	$f_i(S_i^{\uparrow} + S_{i-1}^{\uparrow})$
0.25	0.186	0.186	0.0465
0.30	0.459	0.645	0.1935
0.20	0.674	1.133	0.2266
0.15	0.86	1.534	0.2301
0.1	1	1.86	0.1860
1	/	/	**0.8827**

$I_{Gini} = 1 - \sum_{i=1}^{5} f_i(S_i^{\uparrow} + S_{i-1}^{\uparrow})$

$I_{Gini} = 1 - (0{,}8827)$

$I_{Gini} = 0{,}1173 = 11.73\%$

نلاحظ أن $I_{Gini} = 11.73\%$ وهي أقل من 30% وعليه يمكن القول أن هذا التوزيع أكثر عدالة، حيث توجد هناك فعلا فوارق في توزيع الأجر الإجمالي على مختلف الفئات ، ولكن يمكن أن نعتبر هذا التمركز ضعيفا، أي أن هذه الفوارق يمكن تفسيرها بأسباب موضوعية.

3- بفرض أنه تم دراسة التمركز على توزيع مؤسسة أخرى B، حيث قدرت المساحة خارج منطقة التمركز بـــ 3800 ملم^2،

- تحديد أي التوزيعين أكثر عدالة:

المؤسسة A: $I_{Gini} = 0,1173 = 11.73\%$

المؤسسة B: مساحة منطقة التمركز(S) = 5000 – مساحة خارج منطقة التمركز

مساحة منطقة التمركز(S) = 5000 – 3800 = 1200 ملم^2.

ومنه: $I_{Gini} = \frac{S}{ABC} = \frac{1200}{5000} = 0,24 = 24\%$

بما أن معامل جيني للمؤسسة A أقل منه في المؤسسة B فإن توزيع المؤسسة الأولى أكثر عدالة.

المحور الثامن : الأرقام القياسية

تمهيد :

غالبا ما نصدر أحكاما في حياتنا اليومية، تتضمن تلخيص الفروق التي حدثت في متغيرات إقتصادية خلال فترة زمنية أو الفروق التي تحصل بين مكانين أو أكثر. فمثلا قد يزداد دخل أسرة بمقدار 20% خلال فترة 05 سنوات فإذا كانت هذه الفترة هي فترة تضخم أو إرتفاع عام في الأسعار فهل أن الدخل الحقيقي للعائلة قد إزداد ؟

هل تستطيع العائلة شراء بضائع وخدمات بدخلها أكثر مما كانت تشتري قبل 05 سنوات ؟ فإذا كان المستوى العام لسعر هذه المواد قد ازداد أكثر من 20% فإنها لن تتمكن أن تشتري ما كانت تفعله قبل هذه الفترة .

ومن جهة أخرى فإذا كانت الزيادة في الأسعار أقل من 20% فإن الدخل الحقيقي للعائلة قد ازداد . وكمثال آخر، إذا رغبت إحدى الشركات الكبرى نقل أحد موظفيها من دولة عربية إلى دولة أوروبية مثلا فما الحد الأدنى للزيادة في راتب هذا الموظف كي يستطيع مواجهة تكاليف المعيشة في المدن الأوروبية ؟

تتطلب هاتان الحالتان مقاييس للمستوى العام لهذه الأسعار. فمما لا شك فيه أن أسعار المواد المختلفة التي تشتريها العائلة قد إزدادت بمعدلات مختلفة و ربما القليل من هذه الأسعار قد نقص . وبالمثل فإن أسعار بعض المواد في المدن الأوروبية أعلى بكثير من أسعارها في الدولة العربية ، بينما البعض منها قد يكون أقل . فكيف نستطيع أن نلخص برقم مركب واحد معدل الفروق التي تحدث بين فترتين زمنيتين أو مكانين ؟ إن الأرقام القياسية تفيدنا في الإجابة عن مثل هذه الأسئلة [1].

فالأرقام القياسية هي طريق مختصر لوصف المتغيرات الاقتصادية ، وهي كثيرا تستعمل لوصف التغير في الأسعار أو الكميات أو القيمة مع الزمن، فالرقم القياسي للأسعار هو عدد واحد يرينا كيف تغيرت مجموعة كاملة من الأسعار، والرقم القياسي لتكاليف المعيشة يعطينا نسبة التغير في تكلفة الحياة بين سنة وأخرى مقاسة بتغير سلة من المواد من وقت لآخر . فمثلا لو قلنا أن الرقم القياسي لتكاليف المعيشة في بلد ما بلغ 180 في عام 1995 مقارنة لعام 1990 فهذا يعني أن ما تحتاج إلى إنفاقه عائلة في عام 1995 لشراء عدد من المواد (عادة تحدد هذه المواد عند تعريف الرقم القياسي) يزيد 80% عما كانت تحتاجه هذه العائلة عام 1990 .أي أن تكاليف الحياة ارتفعت 80% في عام 1995 عما كانت عليه في عام 1990 [2].

[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 205.

[2] – محمد صبحي أبو صالح ، مرجع سابق ، ص ، 239.

وعموما فالرقم القياسي في أبسط صورة هو عبارة عن نسبة مئوية تستخدم لقياس التغير الذي يطرأ على ظاهرة ما بالنسبة إلى فترتين زمنيتين مختلفتين أو مكانين مختلفين [1]. (فترتي : أساس ، مقارنة و مكاني : أساس ، مقارنة).

وتوجد عدة أنواع من الأرقام القياسية سندرسها كما يلي:

أولا – الأرقام القياسية الفردية: ويمكن أن نحسبه لسعر سلعة أو كمية سلعة :

1- لسعر سلعة : و يقصد بها إظهار سعر سلعة واحدة معينة في فترة المقارنة منسوبا إلى نفس السلعة في فترة الأساس [2]، ونعبر عن ذلك كما يلي: $I_{(P)} = \frac{P_1}{P_0} \times 100$

حيث : P_1 : سعر السلعة في فترة المقارنة ، P_0 : سعر السلعة في فترة الأساس

ملاحظة: بعد حساب الرقم القياسي و لتحديد نسبة التغير في السعر نحسب Δ حيث : $\Delta = I_{(P)} - 100$

- فإذا كان : $\Delta > 0$ نقول هناك زيادة في السعر.
- و إذا كان : $\Delta < 0$ نقول هناك إنخفاض في السعر.
- و إذا كان : $\Delta = 0$ نقول هناك ثبات في السعر.

مثال01: إذا كان سعر برميل النفط الخام في سنة 2017 هو 65 دولار و أصبح في سنة 2024، 80 دولار، ما هو الرقم القياسي لسعر برميل نفط خام على إعتبار سنة 2017 هي سنة الأساس ؟

الحل:

$$I_{(P)} = \frac{P_1}{P_0} \times 100 = \frac{80}{65} \times 100 = 130.07\%$$

ومنه نسبة التغير في السعر : $\Delta = I_{(P)} - 100 = 130.07 - 100 = 30.07\%$

أي هناك زيادة في الأسعار بـــ : 30.07%

2- لكمية سلعة : و يقصد بها إظهاركمية سلعة واحدة معينة في فترة المقارنة منسوبا إلى نفس السلعة في فترة الأساس ، ونعبر عن ذلك كما يلي: $I_{(Q)} = \frac{Q_1}{Q_0} \times 100$

حيث : Q_1 : سعر السلعة في فترة المقارنة ، Q_0 : سعر السلعة في فترة الأساس .

[1] – أحمد فوزي ملوخيه ، مرجع سابق، ص 341.

[2] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 349.

مثال02: إذا كان إنتاج سلعة معينة (A) في سنة 2017 هو 20 طنا و أصبح إنتاجها في سنة 2024، 80 طنا ، ما هو الرقم القياسي لإنتاج هذه السلعة على إعتبار سنة 2017 هي سنة الأساس ؟

الحل:

$$I_{(Q)} = \frac{Q_1}{Q_0} \times 100 = \frac{80}{20} \times 100 = 400\%$$

ومنه نسبة التغير في الكمية : $\Delta = I_{(Q)} - 100 = 400 - 100 = 300\%$

أي هناك زيادة في انتاج الكمية بـــ : 300%

ثانيا – الأرقام القياسية المركبة (التجميعية) : وتعتمد هذه الأرقام في تركيبها على قيمة ظاهرة أو قيم ظواهر (متغيرات) أزمنة أو أماكن مختلفة [1]، ويكون التصنيف للأرقام القياسية بدلالة الأرقام القياسية للسعر، أو الكميات ، أو القيمة ، حيث سيتم الإكتفاء و التركيز على الأرقام القياسية للأسعار، ويتم ذلك كما يلي :

1- الأرقام القياسية البسيطة (غير المرجحة) للأسعار: لتركيب هذا الرقم لا تتم المفاضلة أو الترجيح لسلعة على سلعة أخرى، حيث نميز بين طريقتين :

1-1 الطريقة النسبية:

$$I_{(S.R.P)} = \frac{\frac{P_{1,1}}{P_{0,1}} + \frac{P_{1,2}}{P_{0,2}} + \cdots + \frac{P_{1,n}}{P_{0,n}}}{m} \times 100 = \frac{\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}}}{m} \times 100$$

حيث m : هو عدد السلع

1-2 الطريقة التجميعية :

$$I_{(S.A.P)} = \frac{P_{1,1} + P_{1,2} + \cdots + P_{1,n}}{P_{0,1} + P_{0,2} + \cdots + P_{0,n}} \times 100 = \frac{\sum_{i=1}^{n} P_{1,i}}{\sum_{i=1}^{n} P_{0,i}} \times 100$$

مثال03: كانت الأسعار لبعض المواد الاستهلاكية كما في الجدول :

السلعة	الكميات	السعر عام **1992** (دج)	السعر عام **1998** (دج)
السكر	1 كلغ	20	40
الأرز	1 كلغ	25	35
الشاي	1 علبة	10	15
القهوة	1 كلغ	200	300

المطلوب : أحسب الرقم القياسي البسيط للأسعار بالطريقيتين النسبية والتجميعية باعتبار عام 1992 سنة الأساس.

[1] – محمد صبحي أبو صالح ، مرجع سابق ، ص 240.

الحل : حساب الرقم القياسي البسيط للأسعار بالطريقتين النسبية والتجميعية باعتبار عام 1992 سنة الأساس:

-الطريقة النسبية: $I_{(S.R.P)} = \frac{\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}}}{m} \times 100 = \frac{\frac{40}{20}+\frac{35}{25}+\frac{15}{10}}{3} \times 100 = 163.33\%$

نسبة التغير في السعر:

$$\Delta = I_{(P)} - 100 = 163.33 - 100 = 63.33\%$$

أي هناك زيادة في الأسعار بـــ :63.33 %

-الطريقة التجميعية : $I_{(S.A.P)} = \frac{\sum_{i=1}^{n} P_{1,i}}{\sum_{i=1}^{n} P_{0,i}} \times 100 = \frac{40+35+15}{20+25+10} \times 100 = 163.63\%$

نسبة التغير في السعر:

$$\Delta = I_{(P)} - 100 = 163.33 - 100 = 63.63\%$$

أي هناك زيادة في الأسعار بـــ:63.63 %

ملاحظة: من سلبيات الأرقام القياسية البسيطة أنها تعطي لجميع السلع نفس الأهمية، أي أنها لا تعطي أهمية أو وزنا للتغير في سعر السلعة الأكثر إستعمالا أكبر مما تعطيه للسلعة قليلة الإستعمال . فعلى سبيل المثال، الأرقام القياسية البسيطة للأسعار تعامل التغير في سعر التلفزيون أو الآلة الحاسبة اليدوية مثلما تعامل التغير في سعر السكر والأرز والشاي مع أن شراء السلع الأولى يكاد يكون مرة واحدة في الفترة الزمنية الطويلة (سنة أو أكثر) بينما يكون استعمال السلع الثلاث الأخيرة متكررا في اليوم الواحد أو الأسبوع الواحد [1].

وهذا العيب يعتبر واضحا في هذا الرقم ، لأنه ليس من المعقول مثلا أن نساوي بين الأهمية النسبية للبترول أو الصلب في الاقتصاد الوطني مع الأهمية النسبية للمربى أو الملح ونضعها جميعا على قدم المساواة .

2- الأرقام القياسية المرجحة للأسعار: للتغلب على العيب الخطير في الأرقام القياسية البسيطة، فإن الأمر يستلزم ترجيح السلع بحسب درجة أهميتها ، فتعطى السلع ذات الأهمية الكبيرة وزنا كبيرا أما السلع ذات الأهمية القليلة فتعطي وزنا صغيرا ، لذا يتطلب الأمر هنا البحث عن معيار معقول يعطي وزنا حقيقيا للأهمية النسبية لكل سلعة تدخل في تركيب الرقم القياسي البسيط، وبعد الدراسة والتحليل وجد أن أهم و أفضل الأوزان الترجيحية المناسبة لكل سلعة هو الكميات سواء أكانت كمياته المستهلكة أو كمياته المشتراة أو كمياته المنتجة على حسب الغرض من تركيب الرقم القياسي [2]. وهذه الصيغ هي :

[1] – نفس المرجع ، ص 243.

[2] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 354.

2-1- رقم لاسبير المرجح *laspeyres Index*: و هنا يتم ترجيح أسعار كل سلعة بكميات سنة الأساس ، ويتم بطريقتين:

2-1-1 الطريقة النسبية: $I_{(\boldsymbol{L.R.P})} = \left(\frac{P_{1,1}}{P_{0,1}} \times Q_{0,1}\% + \frac{P_{1,2}}{P_{0,2}} \times Q_{0,2}\% + \cdots + \frac{P_{1,n}}{P_{0,n}} \times Q_{0,n}\%\right) \times 100$

$$I_{(\boldsymbol{L.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times Q_{0,i}\,\%\right) \times 100$$

2-1-2 الطريقة التجميعية : $I_{(\boldsymbol{L.A.P})} = \frac{P_{1,1}Q_{0,1}+P_{1,2}Q_{0,2}+\cdots+P_{1,n}Q_{0,n}}{P_{0,1}Q_{0,1}+P_{0,2}Q_{0,2}+\cdots+P_{0,n}Q_{0,n}} \times 100$

$$I_{(\boldsymbol{L.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{0,i}}{\sum_{i=1}^{n} P_{0,i}Q_{0,i}} \times 100$$

2-2- رقم باش المرجح *Paasche Index*: و هنا يتم ترجيح أسعار كل سلعة بكميات سنة المقارنة ، ويتم بطريقتين:

2-2-1 الطريقة النسبية: $I_{(\boldsymbol{P.R.P})} = \left(\frac{P_{1,1}}{P_{0,1}} \times Q_{1,1}\% + \frac{P_{1,2}}{P_{0,2}} \times Q_{1,2}\% + \frac{P_{1,n}}{P_{0,n}} \times Q_{1,n}\%\right) \times 100$

$$I_{(\boldsymbol{P.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times Q_{1,i}\,\%\right) \times 100$$

2-2-2 الطريقة التجميعية : $I_{(\boldsymbol{P.A.P})} = \frac{P_{1,1}Q_{1,1}+P_{1,2}Q_{1,2}+\cdots+P_{1,n}Q_{1,n}}{P_{0,1}Q_{1,1}+P_{0,2}Q_{1,2}+\cdots+P_{0,n}Q_{1,n}} \times 100$

$$I_{(\boldsymbol{P.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{1,i}}{\sum_{i=1}^{n} P_{0,i}Q_{1,i}} \times 100$$

2-3- رقم فيشر المرجح *Fishers Index*: و يعتمد هذا الرقم في تركيبه على كل من رقمي لاسبير وباش للاسعار السابقين ، و بحكم أن لكل من الرقمين السابقين عيوب و مميزات، فقد استنتج فيشر رقم قياسي جديد والذي هو عبارة عن الوسط الهندسي لرقمي لاسبير وباش التجميعيين ، ويتم ذلك بطريقتين :

2-3-1 الطريقة النسبية: $I_{(\boldsymbol{F.R.P})} = \sqrt{I_{(L.R.P)} \times I_{(P.R.P)}}$

2-3-2 الطريقة التجميعية : $I_{(\boldsymbol{F.A.P})} = \sqrt{I_{(L.A.P)} \times I_{(P.A.P)}}$

2-4- رقم مارشال المرجح *Marchal Index*: و هنا يتم ترجيح الأسعار بكميات سنتي الأساس و المقارنة ، ويتم بطريقتين:

2-4-1 الطريقة النسبية :

$$I_{(M.R.P)} = \left(\left(\frac{P_{1,1}}{P_{0,1}} \times \frac{Q_{0,1\%}+Q_{1,1\%}}{2}\right) + \left(\frac{P_{1,2}}{P_{0,2}} \times \frac{Q_{0,2\%}+Q_{1,2\%}}{2}\right) + .. + \left(\frac{P_{1,n}}{P_{0,n}} \times \frac{Q_{0,n\%}+Q_{1,n\%}}{2}\right)\right) \times 100$$

$$I_{(M.R.P)} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times \frac{Q_{0,i\%}+Q_{1,i\%}}{2}\right) \times 100$$

2-4-2 الطريقة التجميعية :

$$I_{(M.A.P)} = \frac{\left(P_{1,1} \times \frac{Q_{1,1}+Q_{0,1}}{2}\right) + \left(P_{1,2} \times \frac{Q_{1,2}+Q_{0,2}}{2}\right) + \cdots + \left(P_{1,n} \times \frac{Q_{1,n}+Q_{0,n}}{2}\right)}{\left(P_{0,1} \times \frac{Q_{1,1}+Q_{0,1}}{2}\right) + \left(P_{0,2} \times \frac{Q_{1,2}+Q_{0,2}}{2}\right) + \cdots + \left(P_{0,n} \times \frac{Q_{1,n}+Q_{0,n}}{2}\right)} \times 100$$

$$I_{(M.A.P)} = \frac{\sum_{i=1}^{n} P_{1,i} \frac{Q_{0,i}+Q_{1,i}}{2}}{\sum_{i=1}^{n} P_{0,i} \frac{Q_{0,i}+Q_{1,i}}{2}} \times 100$$

مثال04: الجدول التالي يبين أسعار وكميات أربع مواد استهلاكية لسنتي 1998 و 2002 كما يلي:

السلعة	P0 1998	Q0 1998	Q0%	P1-2002	Q1-2002	Q1%
الاولى	0.5	46	0.144	2	300	0.236
الثانية	1.6	161	0.505	4	500	0.394
الثالثة	0.9	86	0.270	4	400	0.315
الرابعة	0.3	26	0.081	6	70	0.055

المطلوب: باعتبار سنة 1998 سنة الاساس ، أحسب:

-الرقم القياسي المرجح : باش- لاسبير-فيشر-مارشال.

الحل:

الجدول رقم (8-1) : مجاميع كميات المواد استهلاكية

السلعة	P0 1998	Q0 1998	Q0%	P1-2002	Q1-2002	Q1%	$\frac{Q_{0,i}+Q_{1,i}}{2}$	$\frac{Q_{0,i}\%+Q_{1,i}\%}{2}$
الاولى	0.5	46	0.144	2	300	0.236	173	0.19
الثانية	1.6	161	0.505	4	500	0.394	330.5	0.4495
الثالثة	0.9	86	0.270	4	400	0.315	243	0.2925
الرابعة	0.3	26	0.081	6	70	0.055	48	0.068

حساب الرقم القياسي المرجح للأسعار: لاسبير، باش، فيشر، ومارشال، على إعتبار أن 1998 سنة الأساس:

1- حساب الرقم القياسي المرجح للأسعار للاسبير:

- الطريقة النسبية:

$$I_{(\boldsymbol{L.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times Q_{0,i}\,\%\right) \times 100$$

$$I_{(\boldsymbol{L.R.P})} = \left(\frac{2}{0,5} \times 0,144 + \frac{4}{1,6} \times 0,505 + \frac{4}{0,9} \times 0,270 + \frac{6}{0,3} \times 0,081\right) \times 100$$

$$I_{(\boldsymbol{L.R.P})} = 4,657 \times 100 = 465,70\%$$

نسبة التغير في السعر: $\Delta = 465,7 - 100 = 365,7\%$ ،أي أن هناك زيادة في الأسعار بـــــ 365,7%.

- الطريقة التجميعية:

$$I_{(\boldsymbol{L.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{0,i}}{\sum_{i=1}^{n} P_{0,i}Q_{0,i}} \times 100 = \left(\frac{2\times46+4\times161+4\times86+6\times26}{0,5\times46+1,6\times161+0,9\times86+0,3\times26}\right) \times 100$$

$$I_{(\boldsymbol{L.A.P})} = \frac{1236}{365.8} \times 100 = 3,379 \times 100 = 337,9\%$$

نسبة التغير في السعر: $\Delta = 337,9 - 100 = 237,9\%$ ، أي أن هناك زيادة في الأسعار بـــــ 237,9%.

2- حساب الرقم القياسي المرجح للأسعار لباش:

- الطريقة النسبية:

$$I_{(\boldsymbol{P.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times Q_{1,i}\,\%\right) \times 100$$

$$I_{(\boldsymbol{P.R.P})} = \left(\frac{2}{0,5} \times 0,236 + \frac{4}{1,6} \times 0,394 + \frac{4}{0,9} \times 0,315 + \frac{6}{0,3} \times 0,055\right) \times 100$$

$$I_{(\boldsymbol{P.R.P})} = 4,428 \times 100 = 442,8\%$$

نسبة التغير في السعر: $\Delta = 442,8 - 100 = 342,8\%$ ، أي أن هناك زيادة في الأسعار بـــــ 342,8%.

- الطريقة التجميعية:

$$I_{(\boldsymbol{P.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{1,i}}{\sum_{i=1}^{n} P_{0,i}Q_{1,i}} \times 100 = \left(\frac{2\times300 \quad \times500 \quad \times400 \quad \times70}{0,5\times300 \quad ,6\times500 \quad ,9\times400 \quad ,3\times70}\right) \times 100$$

$$I_{(\boldsymbol{P.A.P})} = \frac{4620}{1331} \times 100 = 3,471 \times 100 = 347,10\%$$

نسبة التغير في السعر: $\Delta = 347,10 - 100 = 247,10\%$ ، أي أن هناك زيادة في الأسعار بـــــ 247,10%.

3- حساب الرقم القياسي المرجح للأسعار لفيشر:

- الطريقة النسبية: $I_{(\boldsymbol{F.R.P})} = \sqrt{I_{(L.R.P)} \times I_{(P.R.P)}} = \sqrt{465,70 \times 442,8} = 454,10$

نسبة التغير في السعر: $\Delta = 454,10 - 100 = 354,10\%$ ،أي أن هناك زيادة في الأسعار بـــــ 354,10%.

- الطريقة التجميعية: $I_{(\boldsymbol{F.A.P})} = \sqrt{I_{(L.A.P)} \times I_{(P.A.P)}} = \sqrt{337,9 \times 347,10} = 342,47$

نسبة التغير في السعر: $\Delta = 342,47 - 100 = 242,47\%$ ، أي أن هناك زيادة في الأسعار بـ $242,47\%$.

4- حساب الرقم القياسي المرجح للأسعار لمارشال:

- الطريقة النسبية:

$$I_{(\boldsymbol{M.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times \frac{Q_{0,i}\% + Q_{1,i}\%}{2}\right) \times 100$$

$$I_{(\boldsymbol{M.R.P})} = \left(\frac{2}{0,5} \times 0,19 + \frac{4}{1,6} \times 0,4495 + \frac{4}{0,9} \times 0,2925 + \frac{6}{0,3} \times 0,068\right) \times 100$$

$$I_{(\boldsymbol{M.R.P})} = 4,542 \times 100 = 454,20\%$$

نسبة التغير في السعر: $\Delta = 454,20 - 100 = 354,20\%$ ، أي أن هناك زيادة في الأسعار بـــــ $354,20\%$.

- الطريقة التجميعية:

$$I_{(\boldsymbol{M.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i} \frac{Q_{0,i}+Q_{1,i}}{2}}{\sum_{i=1}^{n} P_{0,i} \frac{Q_{0,i}+Q_{1,i}}{2}} \times 100 = \left(\frac{2\times173+4\times330.5+4\times243+6\times48}{0,5\times173+1,6\times330+0,9\times243+0,3\times48}\right) \times 100$$

$$I_{(\boldsymbol{M.A.P})} = \frac{2928}{847.6} \times 100 = 3,454 \times 100 = 345,40\%$$

نسبة التغير في السعر: $\Delta = 345,40 - 100 = 245,40\%$ ، أي أن هناك زيادة في الأسعار بـــــ $245,4\%$.

المحور التاسع : الإرتباط والإنحدار

تمهيد:

إنحصرت دراستنا من خلال الفصول السابقة في موضوع ظاهرة واحدة ، أي أن البيانات التي كانت لدينا كانت لقيم متغير واحد، فمثلا كانت إهتماماتنا حول علامات مجموعة من الطلبة في أحد الامتحانات، أو أطوال الطلبة لقسم معين، أو أوزان الاطفال حديثي الولادة...الخ، حيث كانت دراستنا تنصب على وضع هذه البيانات في توزيع تكراري ، أو حساب مقاييس النزعة لها أو مقاييس التشتت وغيرها.

وفي هذا المحور سندرس متغيرين عن كل عنصر من عناصر الدراسة، كأن نهتم بطول و وزن الطلبة في أحد الكليات ومن ثم ندرس العلاقة بين هذين المتغيرين ، حيث يمكن الحكم على نوع وقوة هذه العلاقة من خلال تحليل الإرتباط، ثم إذا كانت هناك علاقة بين المتغيرين سنعرف كيف نترجم ذلك في معادلة رياضية تساعدنا في التنبؤ بقيمة أحدهما إذا عُلمت قيمة الآخر وهذا من خلال أسلوب الإنحدار.

إن الأمثلة على هذه الازواج (x ، y) من المتغيرات عديدة فقد نجد :

- الإنفاق، والدخل العائلي.
- معدل الجرائم والبطالة .
- عدد الساعات الدراسية للطالب وعلامته في الامتحان .
- وزن الجسم، وضغط الدم.

وسنحاول التطرق بداية إلى الإرتباط ثم الانحدار.

أولا: الارتباط :

1-تحليل الارتباط : يقيس الارتباط مدى العلاقة بين الظواهر المختلفة أو المتغيرات ، فيساعدنا في معرفة ما إذا كان تغير أحدهما مرتبطا بتغير الآخر [1]، فتحليل الإرتباط يعني دراسة هذه العلاقة من خلال معرفة نوع وقوة العلاقة بين هذين المتغيرين، وذلك كمايلي[2]:

1-1 نوع العلاقة : وتأخذ ثلاث أنواع حسب إشارة معامل الارتباط كما يلي:

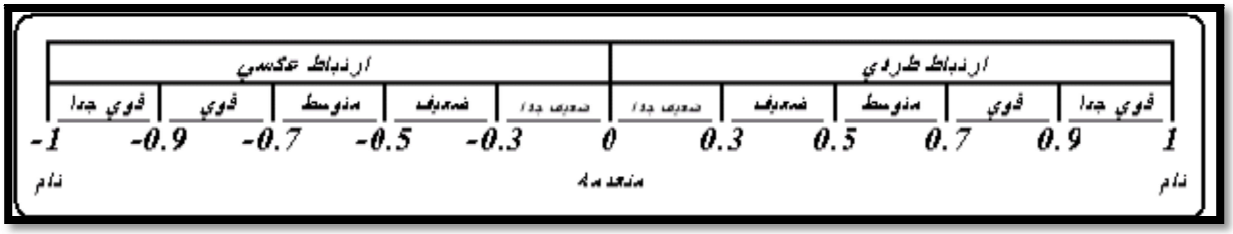
[1] – سالم عيسى بدر، عماد غصاب عبابنة ، مرجع سابق، ص 151.

[2] – شرف الدين خليل ، مرجع سابق، ص 81.

- إذا كانت إشارة معامل الإرتباط سالبة ($r < 0$) توجد علاقة عكسية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه إنخفاض في المتغير الثاني، والعكس.
- إذا كانت إشارة معامل الإرتباط موجبة ($r > 0$) توجد علاقة طردية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني، والعكس.
- إذا كان معامل الإرتباط قيمته صفرا ($r = 0$) دل ذلك على إنعدام العلاقة بين المتغيرين.

1-2 قوة العلاقة : تختلف العلاقة بين المتغيرات من حيث قوتها، فقد نجد أن تغير أحد المتغيرات يعتمد كليا على تغير الآخر، وقد نجد خلاف ذلك ، بحيث يمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن (∓ 1) حيث أن قيمة معامل الارتباط تقع في المدى ($-1 < r < +1$) وقد صنف بعض الإحصائيين درجات لقوة العلاقة يمكن تمثيلها على الشكل التالي:

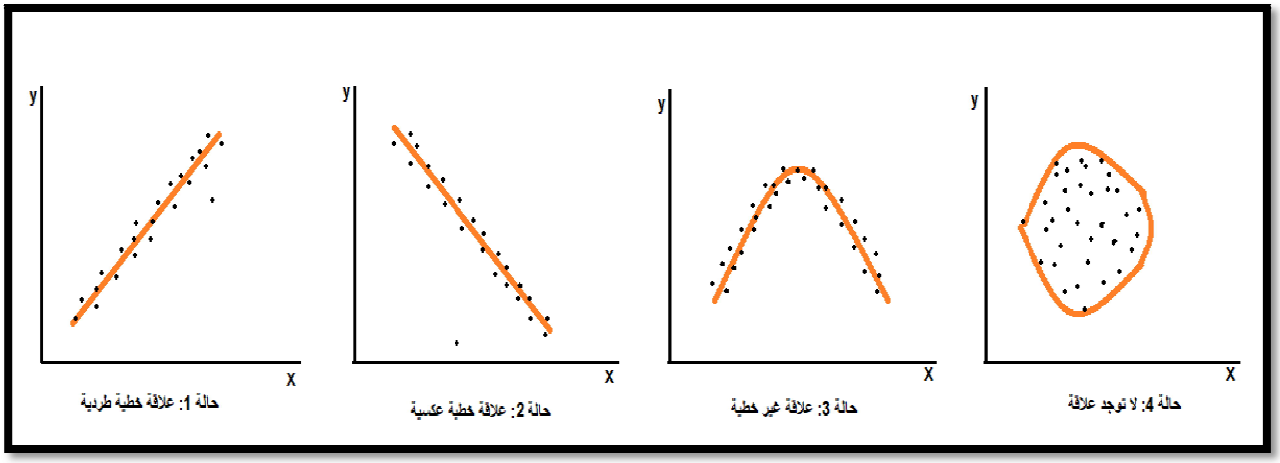
الشكل رقم (9-1): درجات قوة معامل الارتباط



2-لوحة الانتشار وعلاقتها بالارتباط: إن إحدى الطرق التي تساعدنا على معرفة وجود علاقة أو علاقة غير خطية بين متغيرين x ، y هي لوحة الانتشار ، ففي تحليل العلاقة بين متغيرين نجد أنه من المفضل عادة عرض أزواج القيم المتناظرة لهذين المتغيرين في شكل بياني ، حيث تمثل قيم المتغير x على المحور الأفقي وقيم المتغير y على المحور العمودي ، وبالتالي فإن كل زوج من أزواج القيم المتناظرة لهذين المتغيرين ولتكن (x_1 ، y_1) ، (x_2 ، y_2) ، ..... ، (x_n ، y_n) يمكن تمثيلها بيانيا بنقطة ، وتُكون جميع النقط ما يعرف بشكل الإنتشار وهذا الشكل يعطينا صورة مرئية عن طبيعة العلاقة بين هذين المتغيرين ، وشكل الإنتشار قد يأخذ إحدى الحالات التالية [1]:

[1] – أحمد فوزي ملوخيه ، مرجع سابق، ص 268.

الشكل رقم (9-2): حالات شكل الإنتشار



من خلال الشكل السابق يتضح أن شكل الانتشار يعبر عن عدة حالات :

-الحالة (01) : نلاحظ أن شكل نقط الإنتشار تنتشر حول خط مستقيم يتجه من أدنى اليسار إلى أعلى اليمين وهذا يوضح أن المتغيرين x ، y مرتبطان بعلاقة طردية أو موجبة . وهذا يعني أن كل زيادة في المتغير x يصاحبها زيادة في المتغير الآخر y ، والعكس كل نقص في المتغير x يصاحبه نقص في المتغير y . وبعبارة أخرى فإن التغير في كل من هذين المتغيرين سواء بالزيادة أو بالنقص يسيران في نفس الإتجاه .

-الحالة (02) : نلاحظ أن شكل نقط الإنتشار تنتشر حول خط مستقيم ولكن بصورة مخالفة للحالة السابقة، وهذا يوضح أن المتغيرين x ، y مرتبطان بعلاقة عكسية أو سالبة . وهذا يعني أن كل زيادة في المتغير x يصاحبها نقصان في المتغير الآخر y ، والعكس كل نقص في المتغير x يصاحبه زيادة في المتغير y . وبعبارة أخرى فإن التغير في كل من هذين المتغيرين يسيران في اتجاهين متضادين .

-الحالة (03) : نلاحظ أن شكل نقط الإنتشار تنتشر حول خط منحنى وهذا يعني أن المتغيرين x ، y مرتبطان بعلاقة غير خطية .

-الحالة (04) : نلاحظ أن شكل نقط الإنتشار لا تنتشر حول خط مستقيم أو منحنى بل نجدها مبعثرة في جميع أنحاء الشكل البياني بشكل غير منتظم، وهو ما يعني عدم وجود علاقة بين المتغيرين ، أي أن هذين المتغيرين يعتبران مستقلان .

3-معامل الإرتباط: يساعدنا شكل الإنتشار كما لاحظنا سابقا في الحصول على فكرة عامة حول العلاقة بين متغيرين ، إلا أننا لا نستطيع من خلال رسم شكل الإنتشار أن نعرف بالضبط قوة واتجاه العلاقة بين المتغيرين ، لذلك نلجأ إلى حساب معامل الارتباط "***Coefficient Correlation***"

ونحسب معامل الإرتباط للحالات التالية :

1-3 معامل الإرتباط الخطي البسيط " لبيرسون " *Pearson*: في حالة جمع بيانات عن متغيرين كميين يمكن قياس الإرتباط بينهما، باستخدام طريقة " بيرسون " *Pearson* [1] ، حيث يصلح مقياس بيرسون لقياس الإرتباط بين متغيرين كميين لمعرفة نوع وقوة العلاقة ، ويرمز له بالرمز (r) ، و يحسب بالعلاقة :

$$r = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{\sigma_x . \sigma_y}$$

حيث: $\bar{x} = \frac{\sum x}{n}$ و $\bar{y} = \frac{\sum y}{n}$

$\sigma_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$ و $\sigma_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2}$ و n: هو عدد أزواج البيانات.

ويمكن حسابه بطريقة الإنحرافات وفقا للعلاقة : $r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2}\sqrt{\sum (y-\bar{y})^2}}$

مثال (01) : أحسب معامل الارتباط بين المتغيرين x ، y وحدد نوع وقوة العلاقة بينهما ، حيث تكون قيمة كل من x و y كما في الجدول التالي:

4	7	7	10	4	6	1	8	X
2	4	5	6	5	7	3	4	Y

الحل :

لحساب معامل الإرتباط نحتاج الى تكوين الجدول التالي :

الجدول رقم (9-1) : مجاميع قيم المتغيرين x و y

	X	Y	XY	X^2	Y^2
	8	4	32	64	16
	1	3	3	1	9
	6	7	42	36	49
	4	5	20	16	25
	10	6	60	100	36
	7	5	35	49	25

[1] – شرف الدين خليل ، مرجع سابق، ص 81.

	7	4	28	49	16
	4	2	8	16	4
	47	**36**	**228**	**331**	**180**

$$\bar{x} = \frac{\sum x}{n} = \frac{47}{8} = 5.87$$

$$\bar{y} = \frac{\sum y}{n} = \frac{36}{8} = 4.5$$

$$\sigma_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{331}{8} - 5.87^2} = 2.63$$

$$\sigma_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2} = \sqrt{\frac{180}{8} - 4.5^2} = 1.5$$

$$r = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{\sigma_x . \sigma_y} = \frac{\frac{228}{8} - (5.87)(4.5)}{(2.63)(4.5)} = 0.18$$

يوجد هناك إرتباط طردي (لأنه موجب) وضعيف جدا (لأنه أقل من 0.3) فالعلاقة بين المتغيرين طردية ولكن ضعيفة .

2-3 معامل الإرتباط الرتبي (معامل الارتباط سبيرمان للرتب) : يستخدم هذا المعامل لتحديد الارتباط بين البيانات الوصفية التي لا يمكن تطبيق معادلة بيرسون للارتباط الخطي عليها، كما يمكن استخدامه أيضا في حالة البيانات الكمية لتسهيل الحسابات ، وعادة ما يلاحظ أن معامل إرتباط الرتب يكون قريب من معامل بيرسون ويمتاز عنه بالسهولة و الدقة و خاصة عندما يكون عدد أزواج البيانات أقل من 30 (n<30)، و واضح من هذه العلاقة أنه إذا أمكن تحديد الرتب (أو البيانات التي يمكن ترتيبها) نستطيع تعيين r_s. والمقصود بالرتب هو تحديد رتب البيانات دون تحويلها من وضعها، أي ترتيب البيانات ترتيبا تصاعديا أو تنازليا ثم إعطائها ترتيبا يوضع أمام كل قراءة، ففي حالة الترتيب التصاعدي تأخذ أصغر قراءة الواحد الصحيح ثم القيمة التالية في الكبر تأخذ الرقم 2 وهكذا... وعند تساوي قيمتين تأخذ كل منهما الوسط الحسابي لرتبتيهما [1].

ويحسب هذا المعامل من العلاقة:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

حيث: d : هو الفرق بين رتب x و رتب y و n: يمثل عدد أزواج البيانات .

[1] – محمود محمد سليم صالح ، مرجع سابق ، ص 258.

مثال (02) : البيانات التالية تعطي تقدير خمسة طلاب في مادتي الكيمياء و الفيزياء :

مادة الكيمياء X	جيد	راسب	مقبول	ممتاز	مقبول
مادة الفيزياء Y	ممتاز	جيد	جيد	ممتاز	مقبول

المطلوب: إيجاد معامل سبيرمان لارتباط الرتب بين هذه التقديرات .

الحل:

نقوم بترتيب كل متغير على حدى ثم نحدد رتب كل متغير وفقا للجدول التالي :

مادة الكيمياء X	الرتب
راسب	1
مقبول	2.5
مقبول	2.5
جيد	4
ممتاز	5

مادة الفيزياء Y	الرتب
مقبول	1
جيد	2.5
جيد	2.5
ممتاز	4.5
ممتاز	4.5

ثم نقوم بحساب المجاميع كالتالي :

الجدول رقم (9-2) : مجاميع قيم المتغيرين x و y

	مادة الكيمياء X	مادة الفيزياء Y	رتب X	رتب Y	d_i	d_i^2
	جيد	ممتاز	4	4.5	-0.5	0.25
	راسب	جيد	1	2.5	-1.5	2.25
	مقبول	جيد	2.5	2.5	0	0
	ممتاز	ممتاز	5	4.5	0.5	0.25
	مقبول	مقبول	2.5	1	1.5	2.25
Σ	/	/	/	/	/	**5**

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} = 1 - \frac{6(5)}{5(5^2-1)} = 0.75$$

بما أن $r_s = 0.75$ فإن هناك علاقة طردية قوية بين المتغيرين .

ثانيا: الإنحدار :

ذكرنا سابقا بأن الإرتباط يحدد نوع وقوة العلاقة بين متغيرين ، أما في الإنحدار فإننا نحاول توضيح شكل العلاقة التي تربط بين المتغيرين من خلال معادلة رياضية تسمح لنا بالتنبؤ في أحد المتغيرات على أساس المتغير الآخر .

و الإنحدار ظاهرة طبيعية ترجمت إلى مفهوم إحصائي، ويهدف تحليل الإنحدار إلى تقدير معالم (مجاهيل) المعادلة الرياضية التي تعبر عن العلاقة السببية القائمة بين المتغيرات تمهيدا للوصول إلى أفضل تقدير أو (التنبؤ) للمتغير التابع y أي تقدير بيانات غير معروفة مبنية على بيانات معروفة وذات صلة بالظاهرة المدروسة [1].

1-خط الإنحدار البسيط : إن المعادلة الرياضية المراد تحديدها تمثل أقرب خط مستقيم يعبر عن بيانات المتغيرين في صورة خطية وتسمى بمعادلة خط الإنحدار البسيط [2]، لأنه يصف العلاقة الخطية المتوسطة بين متغيرين إثنين فقط أحداهما المتغير التابع و الآخر المتغير المستقل، أما معادلة خط الإنحدار في حالة متغير مستقل وعدة متغيرات تابعة فتسمى بمعادلة خط الإنحدار المتعدد . وقد تكون هذه العلاقة غير خطية أصلا أي تختلف عن شكل الخط المستقيم ، لكن ما يهمنا هنا هو ايجاد معادلة الإنحدار الخطي البسيط .

2- معادلة الإنحدار : إن الغرض من استخدام أسلوب تحليل الإنحدار الخطي البسيط، هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر، حيث يعتبر نموذج الإنحدار الخطي البسيط من أكثر نماذج الإنحدار بساطة وسهولة واستخداما في التقدير والتحليل الإحصائي للعلاقة بين متغيرين إثنين.

و إذا كان لدينا عدد n من الأزواج المرتبة للمتغيرين X و Y أي : $\{(X_1,Y_1),(X_2,Y_2)\ldots\ldots(X_n,Y_n)\}$ فإن معادلة خط الإنحدار للمتغيرين يمكن تقديرها (معادلة إنحدار Y على X) بالشكل التالي:

$$\hat{Y}_i=\hat{\alpha}+\hat{\beta}X_i \qquad i=1,2\ldots.n$$

حيث: $\hat{Y}_i$: هي القيم المقدرة للمتغير التابع Y (قيمة Y المتنبأ بها وقيمة X المختارة).

X_i : المتغير المستقل .

$\hat{\alpha}$: معلمة التقاطع.

$\hat{\beta}$: معلمة الميل .

[1] – ابراهيم علي ابراهيم عبد ربه ، مرجع سابق، ص 271.

[2] – محمود محمد سليم صالح ، مرجع سابق ، ص 262.

ويتم تقدير معاملات الإنحدار $\hat{\alpha}$ و $\hat{\beta}$ باستخدام طريقة المربعات الصغرى ، ويحسب هذا التقدير بالمعادلة التالية [1]:

$$\hat{\beta} = \frac{\Sigma xy - n\bar{x}\,\bar{y}}{\Sigma x^2 - n\bar{x}^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

ملاحظة: يمكن إيجاد معادلة انحدار X على Y وذلك من خلال إجراء تعديلات على تقديرات معاملات الإنحدار بحيث تصبح على النحو التالي:

$$\hat{\beta}' = \frac{\Sigma x \ - n\bar{x}\,\bar{y}}{\Sigma y^2 - n\bar{y}^2}$$

$$\hat{\alpha}' = \bar{X} - \hat{\beta}'\bar{Y}$$

وتكون معادلة إنحدار X على Y كالتالي: $\hat{X}_i = \hat{\alpha}' + \hat{\beta}' Y_i$

مثال 03: أوجد معادلتي الإنحدار (Y على X و X على Y) للبيانات الموجودة بالجدول التالي :

5	8	12	9	10	4	X
4	5	11	8	6	2	Y

الحل: لكتابة معادلتي الانحدار نستعين بالجدول التالي:

الجدول رقم (9-3) : مجاميع قيم المتغيرين x و y

	X	Y	XY	X^2	Y^2	$\hat{Y}_i$	$Y_i - \hat{Y}_i$	$(Y_i - \hat{Y}_i)^2$
	4	2	8	16	4	2.16	−0.16	0.0256
	10	6	60	100	36	7.92	−1.92	3.6864
	9	8	72	81	64	6.96	1.04	1.0816
	12	11	132	144	121	9.84	1.16	1.3456
	8	5	40	64	25	6	−1	1
	5	4	20	25	16	3.12	0.88	0.7744
	48	**36**	**332**	**430**	**263**	/	/	**7.9136**

[1] – محمد صبحي أبو صالح ، مرجع سابق ، ص 230.

1-معادلة إنحدار Y على X :

$\bar{x} = \frac{\Sigma x}{n} = \frac{48}{6} = 8$

$\bar{y} = \frac{\Sigma y}{n} = \frac{36}{6} = 6$

$\widehat{\beta} = \frac{\Sigma xy - n\bar{x}\,\bar{y}}{\Sigma x^2 - n\bar{x}^2} = \frac{332-6(8)(6)}{430-6(8)^2} = 0.96$

$\widehat{\alpha} = 6 - 0.96(8) = -1.68$

معادلة الانحدار التقديرية هي: $\hat{Y}_i = -1.68 + 0.96X_i$

2-معادلة إنحدار **X** على Y:

$\widehat{\beta}' = \frac{\Sigma xy - n\bar{x}\,\bar{y}}{\Sigma y^2 - n\bar{y}^2} = \frac{332-6(8)(6)}{263-6(6)^2} = 0.94$

$\widehat{\alpha}' = 8 - 0.94(6) = 2.36$

معادلة الانحدار التقديرية هي: $\hat{X}_i = 2.36 + 0.94Y_i$

3- الخطأ المعياري : هو معيار مهم يقيس كفاءة نموذج الإنحدار المقدر أي مقدار دقة تمثيل النموذج المقدر للعلاقة بين المتغيرين على ضوء البيانات المتاحة لهما وكلما كانت قيمة الخطأ المعياري صغيرة دل ذلك على كفاءة النموذج المقدر والعكس صحيح [1]، ويتم حسابه كالآتي:

$$\sigma_{e_i} = \sqrt{\frac{\Sigma(Y_i - \hat{Y}_i)^2}{n}}$$

مثال 04: يمكن حساب الخطأ المعياري لمعادلة الإنحدار Y على X من خلال معطيات المثال السابق كالتالي:

$\sigma_{e_i} = \sqrt{\frac{\Sigma(Y_i - \hat{Y}_i)^2}{n}} = \sqrt{\frac{7.9136}{6}} = 1.49$ لدينا خطأ في تقديرالنموذج قدره: 1.49

مثال 05: البيانات المدرجة بالجدول التالي تعطي علامات 12 طالبا في الإمتحان الأول X و الإمتحان الثاني Y

X	18	14	10	15	7	12	13	8	9	12	17	15	12
Y	20	11	14	16	10	10	17	11	12	20	18	12	

المطلوب:

1-أحسب معامل الإرتباط بين المتغيرين X ، Y .

2-أوجد معادلة الإنحدار Y على X .

[1] – طه حسين الزبيدي ، مبادئ الاحصاء ، دارعيداء للنشر والتوزيع، عمان– الاردن، 2012 ، ص 203.

3-أخذ الطالب في الإمتحان الأول العلامة 15 وغاب عن الإختبار الثاني، ما هي العلامة التقديرية التي يحصل عليها الطالب في الإمتحان الثاني ؟

4- مثل بيانيا خط الانحدار .

5-ما هو الخطأ في تقديرك للعلامة في الاختبار الثاني إذا حصل الطالب على 10 في الاختبار الأول ؟

الحل:

1-لحساب معامل الإرتباط نحتاج إلى تكوين الجدول التالي :

الجدول رقم (9-4) : مجاميع قيم المتغيرين x و y

	X	Y	XY		
	18	20	360	324	400
	14	11	154	196	121
	10	14	140	100	196
	15	16	240	225	256
	7	10	70	49	100
	12	10	120	144	100
	13	17	221	169	289
	8	11	88	64	121
	9	12	108	81	144
	17	20	340	289	400
	15	18	270	225	324
	12	12	144	144	144
Σ	**150**	**171**	**2255**	**2010**	**2595**

$$\bar{x} = \frac{\sum x}{n} = \frac{150}{12} = 12.5$$

$$\bar{y} = \frac{\sum y}{n} = \frac{171}{12} = 14.25$$

$$\sigma_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{2010}{12} - 12.5^2} = 3.35$$

$$\sigma_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2} = \sqrt{\frac{2595}{12} - 14.25^2} = 3.63$$

$$r = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{\sigma_x . \sigma_y} = \frac{\frac{2255}{12} - (12.5)(14.25)}{(3.35)(3.63)} = 0.8$$

يوجد هناك إرتباط طردي (لأنه موجب) وقوي (لأنه قريب من الواحد) ، يمكن القول بوجود علاقة قوية بين بين المتغيرين X ، Y و هي علاقة قوية وطردية في نفس الإتجاه.

2-إيجاد معادلة الإنحدار Y على X :

$$\widehat{\beta} = \frac{\Sigma xy - n\bar{x}\,\bar{y}}{\Sigma x^2 - n\bar{x}^2} = \frac{2255 - (12.5)(14.25)}{2010 - (12.5)^2} = 0.87$$

$$\widehat{\alpha} = \bar{Y} - \widehat{\beta}\bar{X}$$

$$\widehat{\alpha} = 14.25 - 0.87(12.5) = 3.375$$

معادلة الانحدار هي: $\hat{Y}_i = 3.375 + 0.87X_i$

3-إذا أخذ الطالب في الإمتحان الأول العلامة 15 وغاب عن الإختبار الثاني، فالعلامة التقديرية التي يحصل عليها الطالب في الإمتحان الثاني تكون :

$$\hat{Y}_i = 3.375 + 0.87X_i$$

$$\hat{Y}_i = 3.375 + 0.87(15)$$

$$\hat{Y}_i \cong 16$$

إذن العلامة التقديرية التي يحصل عليها الطالب في الإمتحان الثاني هي : 16

4- التمثيل البياني لخط الانحدار:

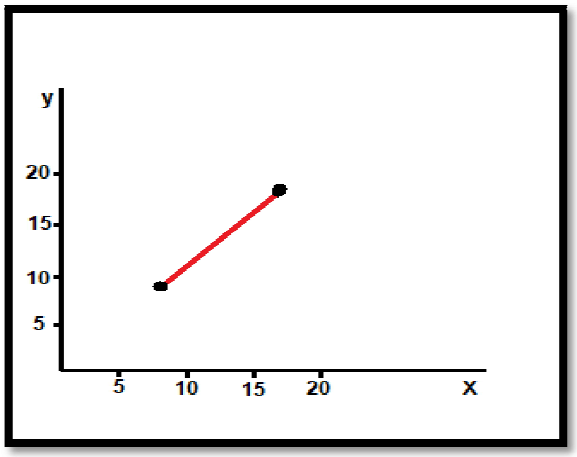
نضع الجدول التالي : والذي يساعدنا في تشكيل خط الانحدار

الجدول رقم (9-5) : نقاط رئيسية لرسم خط الانحدار

X	7	18
$\hat{Y}_i$	9	19

نصل بين النقطتين بخط مستقيم كما هو موضح في الشكل التالي :

الشكل رقم (9-3): خط الانحدار



5-الخطأ في تقدير العلامة في الإختبار الثاني إذا حصل الطالب على 10 في الإختبار الأول، يكون:

$\hat{Y}_i = 3.375 + 0.87(10)$

$\hat{Y}_i \cong 12$

$\hat{e}_i = (Y_i - \hat{Y}_i) = 14 - 12 = 2$

تمارين محلولة :

تمرين 01:

الجدول التالي يعرض الحالة الإجتماعية لـــ 42 عاملا بإحدى المؤسسات :

مطلق	أرمل	أعزب	متزوج	متزوج
أعزب	متزوج	مطلق	متزوج	أرمل
مطلق	متزوج	أرمل	مطلق	متزوج
متزوج	أعزب	متزوج	مطلق	أعزب
متزوج	مطلق	متزوج	أرمل	متزوج
متزوج	متزوج	أعزب	متزوج	متزوج
أرمل	أعزب	أعزب	مطلق	أعزب
متزوج	أعزب	متزوج	أعزب	متزوج
متزوج	متزوج	/	/	/

المطلوب:

1- ما هو نوع المتغير الإحصائي .

2- تلخيص البيانات في جدول توزيع تكراري حسب الحالة الإجتماعية للعمال .

الحل:

1- نوع المتغير الإحصائي:

الحالة الإجتماعية (متزوج، أرمل، مطلق، أعزب) متغيركيفي غير قابل للترتيب.

2- تلخيص البيانات في جدول توزيع تكراري :

يمكن عرض هذه البيانات في جدول تكراري بسيط كما يلي:

الجدول رقم (10-1) : التوزيع التكراري للعمال حسب الحالة الإجتماعية

الحالة الإجتماعية	التفريغ	التكرارات
متزوج	///// ///// ///// /////	20
أرمل	/////	05
مطلق	// /////	07
أعزب	///// /////	10
المجموع		**42**

تمرين 02:

يبين الجدول التالي عدد العقود المبرمة بإحدى الأسواق المالية لدولة العربية لسنتي 2023 و 2024 :

القطاع	عدد العقود المبرمة	
	2023	**2024**
البنوك	55	100
الخدمات	25	30
التأمين	80	70
الصناعة	110	95
المجموع	**270**	**295**

المطلوب:

1- عرض هذه البيانات ضمن عمود مجزأ .

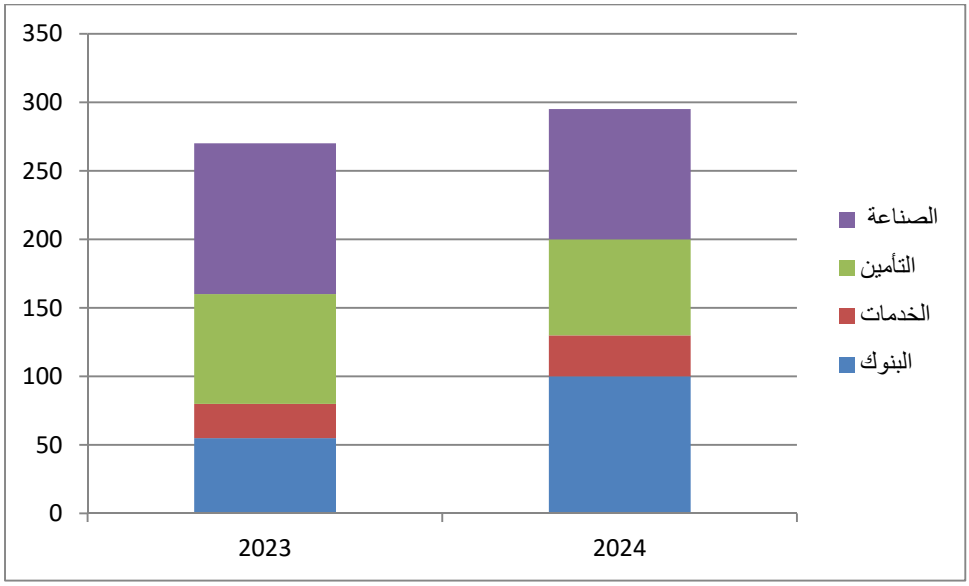
2- عرض بيانات سنة 2023 باستخدام الأعمدة المستطيلة .

3- عرض بيانات سنة 2024 باستخدام القطع الدائرية.

الحل:

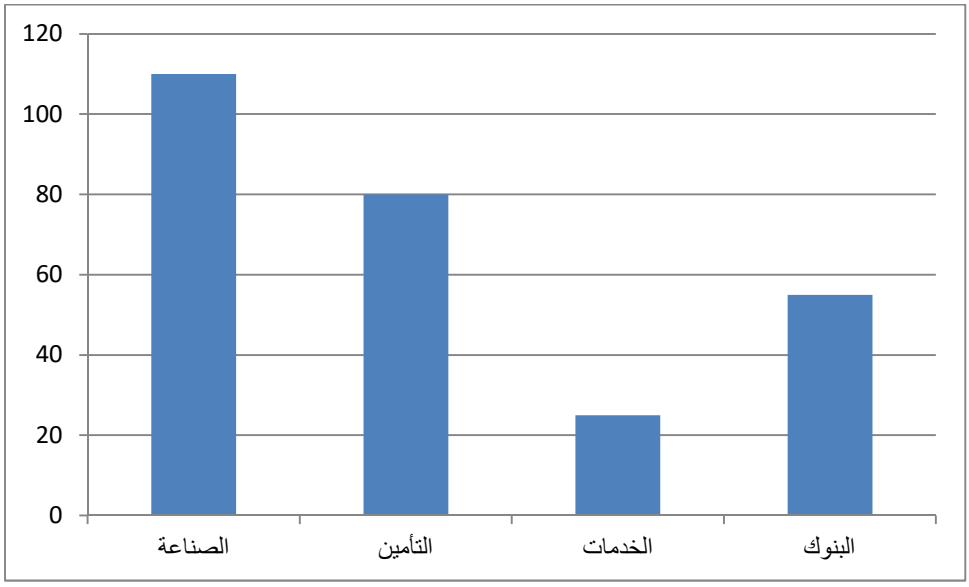
1- عرض البيانات ضمن عمود مجزأ ، كما يلي:

الشكل رقم (10-1) : يمثل عمود مجزأ لتوزيع العقود المبرمة حسب القطاعات لسنتي 2023 و 2024



2- عرض بيانات سنة 2023 باستخدام الأعمدة المستطيلة، كما يلي:

الشكل رقم (10-2): يمثل أعمدة مستطيلة لتوزيع عدد العقود المبرمة حسب القطاعات لسنة 2023



3- عرض بيانات سنة 2024 باستخدام القطع الدائرية:

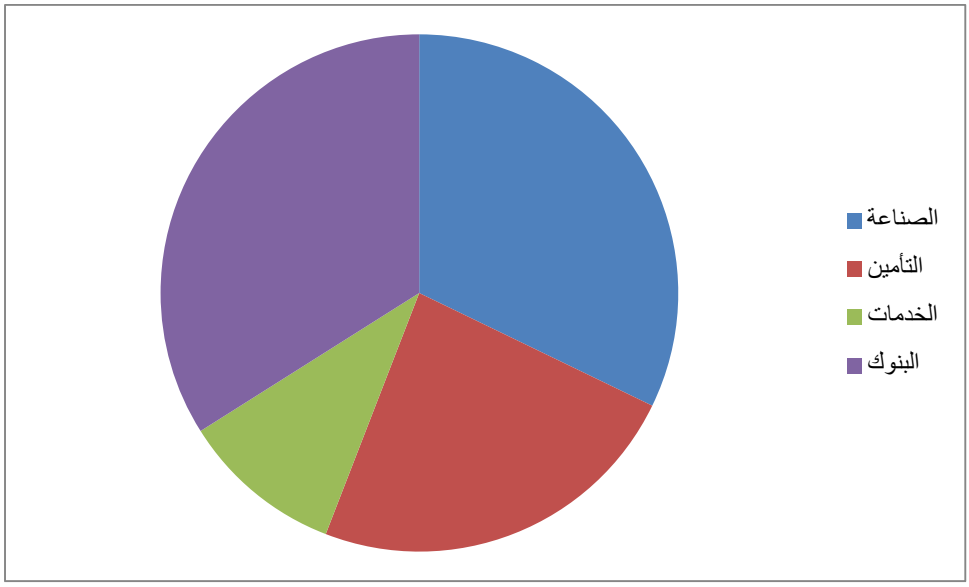
يتم أولا حساب الزاوية المركزية كما يلي:

الجدول رقم (10-2) : يمثل توزيع عدد العقود المبرمة حسب القطاعات لسنة 2024

القطاع	العقود	الزاوية المركزية
البنوك	100	الزاوية المركزية $= \frac{100}{295} \times 360° = 122.03°$
الخدمات	30	الزاوية المركزية $= \frac{30}{295} \times 360° = 36.61°$
التأمين	70	الزاوية المركزية $= \frac{70}{295} \times 360° = 85.42°$
الصناعة	95	الزاوية المركزية $= \frac{95}{295} \times 360° = 115.93°$
المجموع	**295**	**360°**

إنطلاقا من الجدول السابق نرسم الدائرة كما يلي:

الشكل رقم (10-3): يمثل القطع الدائرية لتوزيع عدد العقود المبرمة حسب القطاعات لسنة 2024



تمرين 03:

إذا كانت أعمار ثمانية من أصدقائك هي : 23 ، 21 ، 24 ، 26 ، 28 ، 20 ، 26 ، 27

المطلوب:

1- ماهي قيمة كل من الوسط الحسابي، الوسيط، المنوال للعمر .

الحل:

1- إيجاد قيمة كل من الوسط الحسابي، الوسيط، المنوال للعمر :

-الوسط الحسابي: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{8} X_i}{8} = \frac{23+21+24+26+28+20+26+27}{8} = \frac{195}{8} = 24.375$

-الوسيط:

نرتب البيانات ترتيبا تصاعديا : 20، 21، 23، 24، 26، 26، 27، 28.

بما أن عدد البيانات زوجي فإن الوسيط هو :

$$M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} = \frac{X_{\left(\frac{8}{2}\right)} + X_{\left(\frac{8}{2}+1\right)}}{2} = \frac{X_{(4)} + X_{(5)}}{2} = \frac{24+26}{2} = 25$$

إذن الوسيط هو : $M_e = 25$

-المنوال:

القيمة الأكثر تكرارا هي : 26 وبالتالي المنوال هو $M_o = 26$

تمرين 04:

فيما يلي القيم التي تمثل عدد الحوادث المرورية المسجلة خلال شهر واحد في 40 مفترق طرق في إحدى الولايات الجزائرية:

3	4	2	0	1	5	6	4	3	2
1	0	5	6	3	2	4	5	1	0
6	5	3	2	1	4	5	6	0	2
5	3	5	6	1	0	3	2	6	5

المطلوب:

1-حدد كلا من :المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي، نوع المتغير.

2-أعرض هذه البيانات في جدول توزيع تكراري و مثلها بيانياً.

3-أحسب كلا من f_i، $f_i\%$، $N_i^{\uparrow}$، $N_i^{\downarrow}$، $F_i^{\uparrow}$، $F_i^{\downarrow}$، $F_i^{\uparrow}\%$ ، $F_i^{\downarrow}\%$ ، ثم أشرح كلا من : n_2 ، $N_2^{\uparrow}$، $N_5^{\downarrow}$، $F_2^{\uparrow}\%$ ، $F_5^{\downarrow}\%$

4-أحسب الوسط الحسابي ، الوسط التربيعي للبيانات.

الحل:

1- تحديد كلا من :المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي، نوع المتغير:

المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوع المتغير
جميع مفترقات الطرق في الولاية	مفترق الطرق الواحد	عدد الحوادث في مفترق الطرق خلال الشهر	كمي متقطع

2-عرض البيانات في جدول توزيع تكراري و تمثيلها بيانياً.

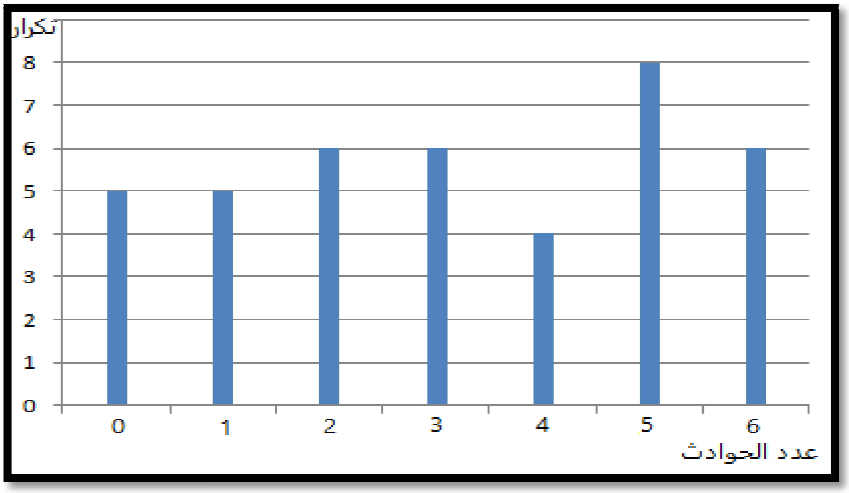
-العرض الجدولي:

الجدول رقم (10-3): توزيع مفترقات الطرق حسب عدد الحوادث

عدد الحوادث Xi	مفترقات الطرق (التكرار) n_i
0	5
1	5
2	6
3	6
4	4
5	8
6	6
المجموع	**40**

-التمثيل البياني : يتم التمثيل في شكل أعمدة بسيطة كما يلي :

الشكل رقم (10-4) : يمثل أعمدة بسيطة لتوزيع مفترقات الطرق حسب عدد الحوادث



3-حساب كلا من f_i، $f_i\%$، $N_i^{\uparrow}$، $N_i^{\downarrow}$، $F_i^{\uparrow}$، $F_i^{\downarrow}$، $F_i^{\uparrow}\%$ ، $F_i^{\downarrow}\%$:

نضع الجدول التالي:

الجدول رقم (10-4): التوزيع التكراري التجميعي الصاعد و النازل (المطلق ، النسبي والمئوي)

$F_i^{\downarrow}\%$	$F_i^{\downarrow}$	$F_i^{\uparrow}\%$	$F_i^{\uparrow}$	$N_i^{\downarrow}$	$N_i^{\uparrow}$	$f_i\%$	f_i	التكرار n_i	قيم المتغير Xi
100	1	12.5	0.125	40	5	12.5	0.125	5	0
87.5	0.875	25	0.25	35	10	12.5	0.125	5	1
75	0.750	40	0.4	30	16	15	0.15	6	2
60	0.6	55	0.55	24	22	15	0.15	6	3
45	0.45	65	0.65	18	26	10	0.10	4	4
35	0.35	85	0.85	14	34	20	0.20	8	5
15	0.150	100	1	6	40	15	0.15	6	6
/	/	/	/	/	/	**100**	**1**	$\sum n_i$=40	المجموع

-شرح كلا من : n_2 ، $N_2^{\uparrow}$، $N_5^{\downarrow}$، $F_2^{\uparrow}\%$ ، $F_5^{\downarrow}\%$:

- n_2: هناك 6 مفترقات طرق من بين 40 مفترق عدد الحوادث فيها يساوي 2 خلال شهر
- $N_2^{\uparrow}$: هناك 16 مفترق طريق من بين 40 مفترق عدد الحوادث فيها على الأكثر خلال شهر هو 2
- $N_5^{\downarrow}$: هناك 14 مفترق طريق من بين 40 مفترق عدد الحوادث فيها على الأقل خلال شهر هو 5
- $F_2^{\uparrow}\%$: هناك 40 % من مفترقات الطرق عدد الحوادث فيها على الأكثر خلال شهر هو 2
- $F_5^{\downarrow}\%$: هناك 35 % من مفترقات الطرق عدد الحوادث فيها على الأقل خلال شهر هو 5

4-حساب الوسط التربيعي ، الوسط الحسابي للبيانات.

نضع الجدول التالي الذي يساعد في الحسابات:

الجدول رقم (10-5): توزيع مفترقات الطرق حسب عدد الحوادث

عدد الحوادث (X_i) (قيم المتغير)	مفترقات الطرق (n_i) (التكرار)	$X_i \times n_i$	$n_i X_i^2$
0	5	0	0
1	5	5	5
2	6	12	24
3	6	18	54
4	4	16	64
5	8	40	200
6	6	36	216
المجموع	**40**	**127**	**563**

-الوسط التربيعي: $\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_i X_i^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{\sum_{i=1}^{7} n_i X_i^2}{40}} = \sqrt{\frac{563}{40}} = 3.751$

-الوسط الحسابي: $\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{7} n_i X_i}{40} = \frac{127}{40} = 3.175$

تمرين 05:

يبين الجدول التالي التوزيع التكراري للدخل السنوي لمجموعة من العمال باحدى المؤسسات: (الوحدة:بالألف)

فئات الدخل	[200 -240[	[240 -280[	[280 -320[	[320 -360[	[360 -400[	[400 -440[	المجموع
عدد العمال	6	15	21	21	5	12	**80**

المطلوب:

إيجاد كل من:

1- الوسط الحسابي للدخل .

2- الوسيط بالحساب وبالرسم .

3- المنوال بالحساب وبالرسم .

الحل:

1- إيجاد الوسط الحسابي للدخل .

نضع الجدول التالي الذي يساعد في الحسابات:

الجدول رقم (6-10) : توزيع العمال حسب فئات الدخل السنوي

فئات الدخل	عدد العمال n_i	مراكز الفئات X'_i	$n_i X'_i$	$N_i^{\uparrow}$	$N_i^{\downarrow}$
]240 -200]	06	220	1320	06	80
]280 -240]	15	260	3900	21	74
]320 -280]	21	300	6300	42	59
]360 -320]	21	340	7140	63	38
]400 -360]	05	380	1900	68	17
]440 -400]	12	420	5040	80	12
المجموع	**80**	/	**25600**	/	/

-الوسط الحسابي: $\bar{X} = \frac{\sum_{i=1}^{k} n_i X'_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{6} n_i X'_i}{80} = \frac{25600}{80} = 320$

2- إيجاد الوسيط الحسابي بالحساب وبالرسم :

- حسابيا:

نحدد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N_{Me}^{\uparrow} \geq \frac{n}{2} = \frac{80}{2} = 40$ ومنه الفئة الوسيطية هي :]320 -280].

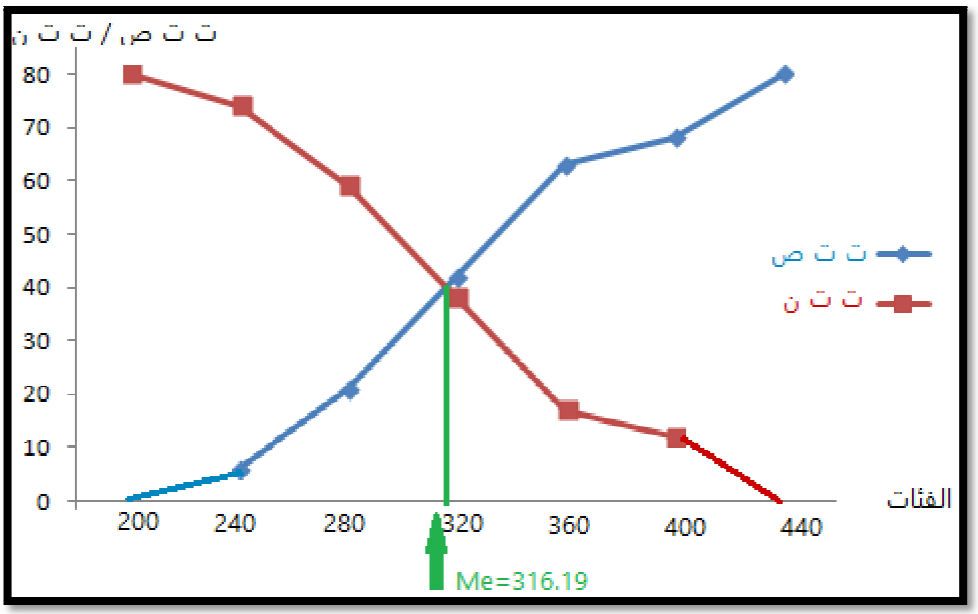
نحسب الوسيط من العلاقة :

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times C_{M_e} = 280 + \left[\frac{40-21}{21}\right] \times 40 = 316.19$$

- بيانيا:

يتم ذلك في الشكل التالي :

الشكل رقم (10-5): إيجاد الوسيط بيانيا



3- إيجاد المنوال بالحساب وبالرسم :

- حسابيا:

نلاحظ أنه يوجد منوالين إثنين :

- المنوال الأول:

نحدد الفئة المنوالية : لدينا أكبر تكرار هو 21 وبالتالي الفئة المنوالية هي [280- 320[

$$\Delta_1 = 21 - 15 = 6 \qquad \Delta_2 = 21 - 21 = 0$$

ومنه: قيمة المنوال الاول هي : $M_{O\,1} = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o} = 280 + \left[\frac{6}{6+0}\right] \times 40 = 320$

- المنوال الثاني:

نحدد الفئة المنوالية : لدينا أكبر تكرار هو 21 وبالتالي الفئة المنوالية هي [320- 360[

$$\Delta_1 = 21 - 21 = 0 \qquad \Delta_2 = 21 - 5 = 16$$

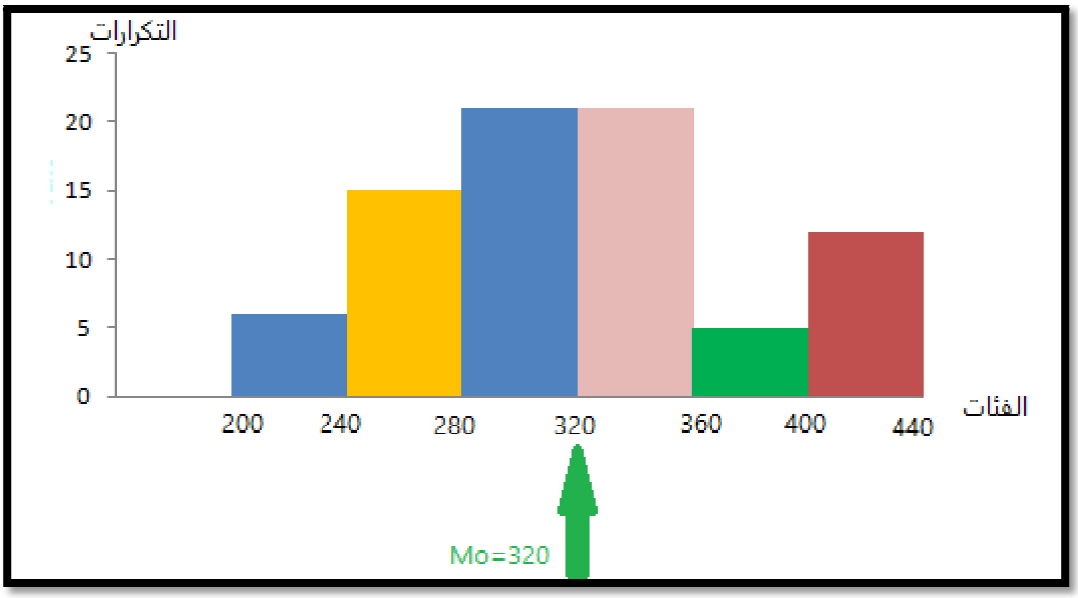
ومنه: قيمة المنوال الثاني هي : $M_{O\,2} = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o} = 320 + \left[\frac{0}{0+16}\right] \times 40 = 320$

نلاحظ أن : $M_{O\,1} = M_{O\,2}$ وبالتالي لدينا منوال واحد هو : $M_O = 320$

- بيانيا:

يتم من خلال المدرج التكراري، حيث بعد رسم هذا الأخير نصل بين أركان الخلية الأكبر تكرارا وحواف الفئتين السابقة واللاحقة لها، لكن في هذه الحالة نلاحظ وجود فئتين منواليتين متجاورتين لهما نفس التكرار الأكبر، وبالتالي يقع المنوال البياني عند الحد المشترك بين الفئتين، أي: $M_O = 320$

الشكل رقم (10-6): إيجاد المنوال بيانيا



تمرين 06:

من جدول المعطيات التالي :

فئات	]10 -05]	]15 -10]	]20 -15]	]25 -20]	]30 -25]	]35 -30]	المجموع
تكرارات	5	3	12	4	6	10	**40**

المطلوب:

1- أوجد الوسط الحسابي ، الوسط التربيعي و الوسط الهندسي والوسط التوافقي للبيانات.

الحل:

1- إيجاد الوسط الحسابي ، الوسط التربيعي و الوسط الهندسي والوسط التوافقي للبيانات:

نضع الجدول التالي الذي يساعد في الحسابات:

الجدول رقم (10-7) : توزيع التكرارات حسب الفئات

فئات	التكرارات n_i	مراكز الفئات X'_i	$n_i X'_i$	$n_i X'^2_i$	$log(X'_i)$	$n_i log(X'_i)$	$\frac{n_i}{X_i}$
]10 -5]	05	7.5	37.5	7031.25	0.875	4.375	0.67
]15 -10]	03	12.5	37.5	4218.75	1.079	3.237	0.24

0.68	14.916	1.243	3675	210	17.5	12	]20 -15]
0.18	5.408	1.352	2025	90	22.5	04	]25 -20]
0.22	8.634	1.439	4537.5	165	27.5	06	]30 -25]
0.31	15.12	1.512	10562.5	325	32.5	10	]35 -30]
2.3	**51.69**	/	**32050**	**865**	/	**40**	المجموع

- الوسط الحسابي : $\overline{X} = \frac{\sum_{i=1}^{k} n_i X_i'}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{6} n_i X_i'}{40} = \frac{865}{40} = 21.62$

- الوسط التربيعي: $\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_i X_i'^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{\sum_{i=1}^{6} n_i X_i'^2}{40}} = \sqrt{\frac{32050}{40}} = 28.30$

- الوسط الهندسي:

$$\bar{X}_G = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i log(X_i')\right)} = 10^{\left(\frac{1}{40}\sum_{i=1}^{6} n_i log(X_i')\right)} = 10^{\left(\frac{1}{40} 51.69\right)} = 19.60$$

- الوسط التوافقي: $\bar{X}_H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} \frac{n_i}{X_i'}} = \frac{40}{\sum_{i=1}^{6} \frac{n_i}{X_i'}} = \frac{40}{2.3} = 17.39$

نلاحظ أن: $\bar{X}_H < \bar{X}_G < \bar{X} < \bar{X}_Q$ محققة .

تمرين 07:

أخذت عينة عشوائية من سبعة طلاب من كل من المجموعتين الأولى والثانية بأحد الكليات، وتم إجراء إختبار للذكاء لهم فكانت نتائج العينات كمايلي:

- نتائج العينة من المجموعة الأولى: 61 ، 79 ، 83 ، 61 ، 76 ، 69 ، 73 .
- نتائج العينة من المجموعة الثانية: 94 ، 98 ، 91 ، 86 ، 96 ، 93 ، 99 .

المطلوب:

1-إيجاد المدى والانحراف المعياري لكلا المجموعتين.

2-قارن بين تشتت المجموعتين .

الحل:

1- إيجاد المدى والإنحراف المعياري:

أ- المجموعة الأولى :

- المدى: $R = X_{max} - X_{min} = 83 - 61 = 22$

- الإنحراف المعياري:

نحسب المتوسط الحسابي أولا: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{7} X_i}{7} = \frac{61+79+83+61+76+69+73}{7} = \frac{502}{7} = 71.71$

ويكون الانحراف المعياري:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}} = \sqrt{\frac{\sum_{i=1}^{7}(X_i-\bar{X})^2}{7}}$$

$$= \sqrt{\frac{(61-71.71)^2+(79-71.71)^2+(83-71.71)^2+(61-71.71)^2+(76-71.71)^2+(69-71.71)^2+(73-71.71)^2}{7}} = \sqrt{\frac{437.43}{7}} = 7.9$$

ب- المجموعة الثانية :

- المدى: $R = X_{max} - X_{min} = 99 - 86 = 13$

- الانحراف المعياري:

نحسب المتوسط الحسابي أولا: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{7} X_i}{7} = \frac{94+98+91+86+96+93+99}{7} = \frac{657}{7} = 93.86$

ويكون الانحراف المعياري:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}} = \sqrt{\frac{\sum_{i=1}^{7}(X_i-\bar{X})^2}{7}}$$

$$= \sqrt{\frac{(94-93.86)^2+(98-93.86)^2+(91-93.86)^2+(86-93.86)^2+(96-93.86)^2+(93-93.86)^2+(99-93.86)^2}{7}} = \sqrt{\frac{118.86}{7}} = 4.12$$

2-المقارنة بين تشتت المجموعتين :

بما أن المتوسطين الحسابيين مختلفان فإننا نستعمل معامل الاختلاف CV للمقارنة بين التشتت النسبي للمجموعتين:

أ- المجموعة الأولى : $CV_1 = \frac{\sigma(X)}{\bar{X}} \times 100\% = \frac{7.9}{71.71} \times 100\% = 11.02\%$

ب- المجموعة الثانية : $CV_2 = \frac{\sigma(X)}{\bar{X}} \times 100\% = \frac{4.12}{93.86} \times 100\% = 4.39\%$

بما أن: $CV_1 > CV_2$ فإننا نقول أن تشتت القيم في المجموعة الثانية أقل من تشتت القيم في المجموعة الأولى وبالتالي فإن نتائج المجموعة الثانية أكثر تجانسا.

تمرين 08:

إذا كانت القيم التالية تمثل معدل (متوسط) الدخل للعامل في دولتين A و B

- الدولة A : $\bar{X} = 1500$ دينار ، $\sigma(X) = 900$

- الدولة B : $\bar{X} = 3000$ دينار ، $\sigma(X) = 1200$

المطلوب:

أي الدولتين يكون تشتت الدخول فيها أقل .

الحل:

– المقارنة بين تشتت الدخول في الدولتين :

بما أن المتوسطين الحسابيين مختلفان فإننا نستعمل معامل الإختلاف CV للمقارنة بين التشتت النسبي للدخول في الدولتين:

– الدولة A: $CV_1 = \frac{\sigma(X)}{\bar{X}} \times 100\% = \frac{900}{1500} \times 100\% = 60\%$

– الدولة B: $CV_2 = \frac{\sigma(X)}{\bar{X}} \times 100\% = \frac{1200}{3000} \times 100\% = 40\%$

بما أن: $CV_1 > CV_2$ فإننا نقول أن تشتت الدخول في الدولة B أقل من تشتت الدخول في الدولة A وبالتالي فإن الدخول في الدولة B أكثر تجانسا.

تمرين 09:

تم جمع نقاط 150 طالب في مقياس الرياضيات للسنة أولى جذع مشترك، فتحصلنا على المعلومات التالية :

النقاط	[0 - 5[	[5 - 10[	[10 - 15[	[15 - 20[	المجموع
عدد الطلبة	25	45	56	24	**150**

المطلوب:

1–حدد المتغير الإحصائي وطبيعته؟

2–أحسب كل من المتوسط الحسابي والوسيط والمنوال؟

3–أحسب الرُبيع الأول (Q_1) ثم استنتج العُشير الخامس (D_5) والمئوي 25 (P_{25})؟

4–أدرس شكل التوزيع من ناحية الالتواء بطريقتين؟

5–حدد شكل التوزيع من ناحية التفرطح ؟

الحل:

1– تحديد المتغير الاحصائي وطبيعته :

المتغير الاحصائي: نقاط طلبة في مقياس الاحصاء .

طبيعته : كمي متصل .

2- حساب كل من المتوسط الحسابي و الوسيط والمنوال :

نضع الجدول التالي الذي يساعد في الحسابات:

الجدول رقم (8-10) : توزيع علامات 150 طالب في مادة الرياضيات

فئات العلامات	عدد الطلبة n_i	مراكز الفئات X'_i	$N_i^{\uparrow}$	$(X'_i-\bar{X})^2$	$n_i(X'_i-\bar{X})^2$	$(X'_i-\bar{X})^4$	$n_i(X'_i-\bar{X})^4$
]0- 5]	25	2.5	25	58.22	1455.5	3389.21	84730.25
]5- 10]	45	7.5	70	6.92	311.4	47.84	2152.80
]10- 15]	56	12.5	126	5.62	314.72	31.55	1766.80
]15- 20]	24	17.5	150	54.32	1303.68	2950.32	70807.68
المجموع	**150**	/	/	/	**3385.3**	/	**159457.53**

-الوسط الحسابي: $\bar{X} = \frac{\sum_{i=1}^{k} n_i X'_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{4} n_i X'_i}{150} = \frac{1520}{150} = 10.13$

-الوسيط:

نحدد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N_{M_e}^{\uparrow} \geq \frac{n}{2} = \frac{150}{2} = 75$ ومنه الفئة الوسيطية هي :]10- 15].

نحسب الوسيط من العلاقة :

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times C_{M_e} = 10 + \left[\frac{75-70}{56}\right] \times 5 = 10.96$$

-المنوال:

نحدد الفئة المنوالية : لدينا أكبر تكرار هو 56 وبالتالي الفئة المنوالية هي]10- 15]

$$\Delta_1 = 56 - 45 = 11 \qquad \Delta_2 = 56 - 24 = 32$$

ومنه: قيمة المنوال هي : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o} = 10 + \left[\frac{11}{11+32}\right] \times 5 = 11.28$

3- حساب الرُبيع الأول (Q_1) ثم استنتاج العُشير الخامس (D_5) والمئوي 25 (P_{25}):

-الربيع الأول (Q_1):

نحدد الفئة الربيعية الأولى: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{in}{4}$ أو يساويها ، أي: $N_{Q_1}^{\uparrow} \geq \frac{in}{4} = \frac{1(150)}{4} = 37.5$ ومنه الفئة الربيعية الأولى هي :]05- 10].

نحسب الربيع الأول من العلاقة :

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{in}{4} - N^{\uparrow}_{Q_1-1}}{n_{Q_1}}\right] \times C_{Q_1} = 5 + \left[\frac{37.5 - 25}{45}\right] \times 5 = 6.39$$

- إستنتاج العشير الخامس (D_5) :

$$D_5 = M_e = 10.96$$

- إستنتاج المئوي 25 (P_{25}) :

$$P_{25} = Q_1 = 6.39$$

4- إيجاد شكل التوزيع من ناحية الإلتواء بطريقتين :

- طريقة 1: مقارنة مقاييس النزعة المركزية :

بما أن : $\bar{X} = 10.13 < M_e = 10.96 < M_o = 11.28$

فإن منحى التوزيع مائل لليسار سالب الإلتواء.

- طريقة 2: طريقة معامل بيرسون للإلتواء :

من خلال قانون بيرسون الأول: $SK = \frac{\bar{X} - M_o}{\sigma(X)}$

حساب الانحراف المعياري : $\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X'_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{3385.3}{150}} = \sqrt{22.569} = 4.75$

ويصبح: $SK = \frac{\bar{X} - M_o}{\sigma(X)} = \frac{10.96 - 1\ .28}{4.75} = -0.067$

بما أن : $SK < 0$ ، فمنحنى التوزيع مائل لليسار ، سالب الالتواء.

5- تحديد شكل التوزيع من ناحية التفرطح :

-معامل بيرسون للتفرطح: $K = \frac{\mu_4}{[\delta(X)]^4} = \frac{\frac{\sum_{i=1}^{k} n_i (X'_i - \bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4} = \frac{\frac{159457.53}{150}}{[4.75]^4} = 2.09$

بما أن : $K < 3$: فمنحنى التوزيع مفرطح (منبسط).

تمرين 10:

البيانات التالية تمثل أطوالا بالسنتيمتر لقطع حديدية قبل تحويلها إلى قطع منتظمة الشكل :

326	296	286	266	235	346	221	381	401	376	257	345	246	342	235
356	346	231	396	291	266	351	251	311	336	391	226	361	291	391

المطلوب:

1- تفريغ البيانات في جدول تكراري ذو فئات.

2- ما هو معدل طول القطع، الطول الوسيط للقطع، والطول الشائع للقطع .

3- من خلال مقارنة مقاييس النزعة المركزية حدد شكل التوزيع من ناحية الالتواء واشارته .

4- باستخدام معامل "بيرسون" حدد شكل التوزيع من ناحية الالتواء واشارته .

الحل:

1- تفريغ البيانات في جدول تكراري ذو فئات:

نقوم بالخطوات التالية :

- حساب المدى: $E = Xmax - Xmin \quad = 401 - 221 \quad = 180$
- حساب عدد الفئات : $k = 1 + 3.322\log(n) = 1 + 3.322\log(30) \quad = 5.90 \cong 6$
- حساب طول الفئة : $C = \frac{E}{K} \quad = \frac{180}{6} \quad = 30$
- نكون الجدول التفريغي:

الجدول رقم (9-10) : توزيع القطع الحديدية حسب الطول

فئات الطول	عدد القطع $\mathbf{n_i}$	مراكز الفئات X'_i	$N_i^{\uparrow}$	$n_i X'_i$	$(X'_i - \overline{X})^2$	$n_i(X'_i - \overline{X})^2$
]221- 251]	6	236	6	1416	6241	37446
]251- 281]	4	266	10	1064	2401	9604
]281- 311]	4	296	14	1184	361	1444
]311- 341]	3	326	17	978	121	363
]341- 371]	7	356	24	2492	1681	11767
]371- 401]	6	386	30	2316	5041	30246
المجموع	**30**	/	/	**9450**	/	**90870**

2-حساب معدل طول القطع، الطول الوسيط للقطع، والطول الشائع للقطع .

- معدل طول القطع (الوسط الحسابي) : $\overline{X} = \frac{\sum_{i=1}^{k} n_i X'_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{6} n_i X'_i}{30} = \frac{9450}{30} = 315$
- الوسيط :

نحدد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N_{Me}^{\uparrow} \geq \frac{n}{2} = \frac{30}{2} = 15$ ومنه الفئة الوسيطية هي :]311- 341].

نحسب الوسيط من العلاقة :

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{M_e-1}}{n_{M_e}}\right] \times C_{M_e} = 311 + \left[\frac{15-14}{3}\right] \times 30 = 321$$

- المنوال :

نحدد الفئة المنوالية : لدينا أكبر تكرار هو 7 وبالتالي الفئة المنوالية هي [341- 371[

$$\Delta_1 = 7 - 3 = 4$$
$$\Delta_2 = 7 - 6 = 1$$

ومنه: قيمة المنوال هي : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o} = 341 + \left[\frac{4}{4+1}\right] \times 30 = 365$

3-تحديد شكل التواء منحنى التوزيع من خلال مقارنة مقاييس النزعة المركزية :

بما أن : $\bar{X} = 315 < M_e = 321 < M_o = 365$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الالتواء.

4-تحديد شكل التواء منحنى التوزيع من خلال معامل "بيرسون":

- قانون بيرسون الأول:

حساب الانحراف المعياري : $\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{90870}{30}} = \sqrt{3029} \approx 55.04$

ويصبح: $SK = \frac{\bar{X} - M_o}{\sigma(X)} = \frac{315-365}{55.04} \approx -0.91$

بما أن : $SK < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الالتواء.

- قانون بيرسون الثاني: $SK = \frac{3(\bar{X} - M_e)}{\sigma(X)} = \frac{3(315-321)}{55.04} \approx -0.33$

بما أن : $SK < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الالتواء.

تمرين 11:

إذا كان لدينا الجدول التكراري التالي:

فئات	[03- 09[	[09- 15[	[15- 21[	[21- 27[	المجموع
تكرارات	8	13	24	25	**70**

المطلوب:

1- أحسب مقدار و نوع الالتواء باستخدام العزوم.

2- أحسب مقدار و نوع التفرطح باستخدام العزوم.

الحل:

نضع الجدول التالي الذي يساعد في الحسابات:

الجدول رقم (10-10) : توزيع التكرارات حسب الفئات

فئات	التكرارات n_i	مراكز الفئات X'_i	$n_i X'_i$	$n_i(X'_i-\bar{X})^2$	$n_i(X'_i-\bar{X})^3$	$n_i(X'_i-\bar{X})^4$
]9 -3]	08	06	48	1075.9848	- 12681.9384	147871.4014
]15 -9]	13	12	156	410.8028	- 2357.1794	13341.6357
]21 -15]	24	18	432	3.1144	0.9433	0.3207
]27 -21]	25	24	600	1011.23	6371.0026	40392.1565
المجموع	**70**	/	**1236**	**2501.132**	**- 8667.1719**	**201605.5143**

1- حساب مقدار و نوع الالتواء باستخدام العزوم:

-نحسب الوسط الحسابي والانحراف المعياري أولا:

$$\overline{X} = \frac{\sum_{i=1}^{k} n_i X'_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{4} n_i X'_i}{70} = \frac{1236}{70} = 17.66$$

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X'_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{2501.132}{70}} = \sqrt{35.73} \approx 5.98$$

-معامل فيشر للالتواء:

$$F_1 = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{\sum_{i=1}^{k} n_i (X'_i - \bar{X})^3}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^3} = \frac{\frac{-8667.1719}{70}}{[5.98]^3} = -0.58$$

بما أن : $F_1 < 0$ ، فمنحنى التوزيع ملتوي جهة اليسار ، سالب الالتواء.

2- حساب مقدار و نوع التفرطح باستخدام العزوم:

-معامل فيشر للتفرطح:

$$F_2 = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{\frac{\sum_{i=1}^{k} n_i (X'_i - \bar{X})^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4} - 3 = \frac{\frac{201605.5143}{70}}{[5.98]^4} - 3 = -0.75$$

بما أن : $F_2 < 0$: فمنحنى التوزيع مفرطح (منبسط).

تمرين 12:

إذا أعطيت البيانات التالية :

فئات	]10 -06]	]14 -10]	]18 -14]	]22 -18]	]26 -22]	]30 -26]	المجموع
تكرارات	7	5	13	6	7	12	**50**

المطلوب:

حساب مايلي:

1- المدى الربيعي .

2- الانحراف المتوسط والانحراف المعياري .

3- مقياس الالتواء .

4- مقياس التفرطح.

الحل:

نضع الجدول التالي الذي يساعد في الحسابات:

الجدول رقم (10-11) : توزيع التكرارات حسب الفئات

فئات	التكرارات n_i	مراكز الفئات X'_i	$n_i X'_i$	$N_i^{\uparrow}$	$n_i\|X'_i - \bar{X}\|$	$n_i(X'_i - \bar{X})^2$	$n_i(X'_i - \bar{X})^3$	$n_i(X'_i - \bar{X})^4$
]10 -6]	07	8	56	07	76.72	840.8512	9215.7291	101004.3915
]14 -10]	05	12	60	12	34.8	242.208	1685.7677	11732.9430
]18 -14]	13	16	208	25	38.48	113.9008	337.1464	997.9532
]22 -18]	06	20	120	31	6.24	6.4896	6.7492	7.0191
]26 -22]	07	24	168	38	35.28	177.8112	896.1684	4516.6890
]30 -26]	12	28	336	50	108.48	980.6592	8865.1592	80141.0389
المجموع	**50**	/	**948**	/	**300**	**2361.92**	**21006.72**	**198400.03**

1- حساب المدى الربيعي:

نحسب الربيع الأول والربيع الثالث:

-الربيع الأول (Q_1):

نحدد الفئة الربيعية الأولى: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{in}{4}$ أو يساويها ، أي:12.5 $= \frac{1(50)}{4} = \frac{in}{4} \geq N_{Q_3}^{\uparrow}$ ومنه الفئة الربيعية الأولى هي : [14- 18[.

نحسب الربيع الأول من العلاقة :

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{in}{4} - N_{Q_1-1}^{\uparrow}}{n_{Q_1}}\right] \times C_{Q_1} = 14 + \left[\frac{12.5 - 12}{13}\right] \times 4 = 14.15$$

-الربيع الثالث (Q_3):

نحدد الفئة الربيعية الثالثة: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{in}{4}$ أو يساويها ، أي:37.5 $= \frac{3(50)}{4} = \frac{in}{4} \geq N_{Q_3}^{\uparrow}$ ومنه الفئة الربيعية الثالثة هي : [22- 26[.

نحسب الربيع الثالث من العلاقة :

$$Q_1 = Lim_{Q_3} + \left[\frac{\frac{in}{4} - N_{Q_3-1}^{\uparrow}}{n_{Q_3}}\right] \times C_{Q_3} = 22 + \left[\frac{37.5 - 31}{07}\right] \times 4 = 25.71$$

ومنه: $RQ = Q_3 - Q_1 = 25.71 - 14.15 = 11.56$

2- حساب الانحراف المتوسط والانحراف المعياري :

-الإنحراف المتوسط :

نحسب أولا الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i'}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{6} n_i X_i'}{50} = \frac{948}{50} = 18.96$$

نحسب الإنحراف المتوسط كما يلي :

$$MD = \frac{\sum_{i=1}^{k} n_i |X_i' - \bar{X}|}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{6} n_i |X_i' - \bar{X}|}{50} = \frac{300}{50} = 6$$

-الإنحراف المعياري :

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n} n_i (X_i' - \bar{X})^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{\sum_{i=1}^{6} n_i (X_i' - \bar{X})^2}{50}} = \sqrt{\frac{2361.92}{50}} = 6.87$$

3- حساب مقياس الإلتواء :

يمكن تبيين نوع الإلتواء بحساب أحد المقاييس التالية:

- قانون بيرسون الأول:

نحسب المنوال أولا:

نحدد الفئة المنوالية : لدينا أكبر تكرار هو 13 وبالتالي الفئة المنوالية هي [14- 18[

$$\Delta_2 = 13 - 6 = 7$$

ومنه: قيمة المنوال هي : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times C_{M_o} = 14 + \left[\frac{8}{8+7}\right] \times 4 = 16.13$

ويصبح: $SK = \frac{\bar{X}-M_o}{\sigma(X)} = \frac{18.96-16.13}{6.87} = 0.41$

بما أن : $SK > 0$ ، فمنحنى التوزيع ملتوي جهة اليمين ، موجب الالتواء.

– قانون بيرسون الثاني:

نحسب الوسيط أولا:

نحدد الفئة الوسيطية: وهي الفئة التي يزيد تكرارها المتجمع الصاعد عن القيمة المساوية لـــ $\frac{n}{2}$ أو يساويها ، أي: $N_{M_e}^{\uparrow} \geq \frac{n}{2} = \frac{50}{2} = 25$ ومنه الفئة الوسيطية هي : [14- 18[.

نحسب الوسيط من العلاقة :

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times C_{M_e} = 14 + \left[\frac{25-12}{13}\right] \times 4 = 18$$

ويصبح: $SK = \frac{3(\bar{X}-M_e)}{\sigma(X)} = \frac{3(18.96-18)}{6.87} = 0.42$

بما أن : $SK > 0$ ، فمنحنى التوزيع ملتوي جهة اليمين ، موجب الالتواء.

–معامل فيشر للإلتواء:

$$F_1 = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{\sum_{i=1}^{k} n_i\left(X_i'-\bar{X}\right)^3}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^3} = \frac{\frac{21006.72}{50}}{[6.87]^3} = 1.29$$

بما أن : $F_1 > 0$ ، فمنحنى التوزيع ملتوي جهة اليمين ، موجب الالتواء.

4– حساب مقياس التفرطح :

يمكن تبيين نوع التفرطح بحساب أحد المقاييس التالية:

– معامل بيرسون للتفرطح:

$$K = \frac{\mu_4}{[\delta(X)]^4} = \frac{\frac{\sum_{i=1}^{k} n_i\left(X_i'-\bar{X}\right)^4}{\sum_{i=1}^{k} n_i}}{[\delta(X)]^4} = \frac{\frac{198400.03}{50}}{[6.87]^4} = 1.87$$

بما أن : $K < 3$: فمنحنى التوزيع مفرطح (منبسط).

–معامل فيشر للتفرطح:

$$F_2 = k - 3 = 1.87 - 3 = -1.13$$

بما أن : $F_2 < 0$: فمنحنى التوزيع مفرطح (منبسط).

تمرين 13:

الجدول التالي يمثل توزيع الأراضي الزراعية في إحدى الدول العربية لفترة ما حسب فئات المساحة (فدان) للفرد الواحد

فئات المساحة (فدان)	]10 -05[	] 15 -10[	] 20 -15[	] 25 -20[	] 30 -25[	المجموع
عدد الملاك	200	500	100	400	300	1500

المطلوب:

-إظهار مدى التفاوت في توزيع ملكية الأراضي الزراعية بين الملاك بيانيا وحسابيا .

الحل:

دراسة قضية التمركز بيانيا وحسابيا :

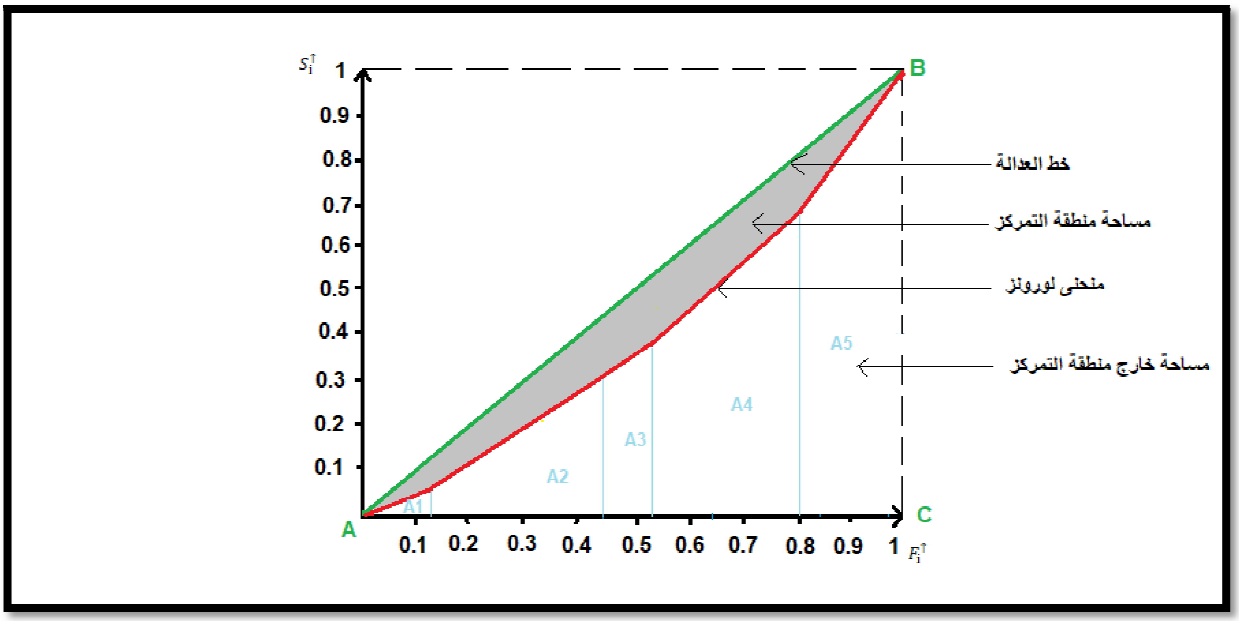
1- بيانيا : ويتم ذلك بواسطة منحنى لورونز حيث نقوم أولا بحساب $F_i^{\uparrow}$ و $S_i^{\uparrow}$ ، ثم نقوم برسم منحنى لورونز عن طريق وصل النقاط التي إحداثياتها ($F_i^{\uparrow}$ ، $S_i^{\uparrow}$) ببعضها ضمن مربع جيني .

الجدول رقم (10-12) : جدول تحضير الحسابات لتوزيع عدد الملاك حسب فئات المساحة لرسم منحنى لورونز

فئات المساحة (فدان)	عدد الملاك n_i	مراكز الفئات X_i'	$n_i X_i'$	نسبة الملاك f_i	نسب المساحة s_i	$F_i^{\uparrow}$	$S_i^{\uparrow}$
] 10 -05]	200	7.5	1500	0.13	0.050	0.13	0.050
] 15 -10]	500	12.5	6250	0.33	0.210	0.46	0.260
] 20 -15]	100	17.5	1750	0.07	0.059	0.53	0.319
] 25 -20]	400	22.5	9000	0.27	0.303	0.8	0.622
] 30 -25]	300	37.5	11250	0.20	0.378	1	1
المجموع	**1500**	/	**29750**	**1**	/	/	/

ويكون منحنى لورونز في الشكل التالي :

الشكل رقم (10-7): مربع جيني لتوزيع عدد الملاك حسب فئات المساحة



نلاحظ من خلال الشكل أن منحنى لورونز يقترب نوعا ما من خط العدالة ومنه فإن توزيع الملاك حسب المساحة أكثر عدالة .

1-1 **حسابيا**: ويتم ذلك بواسطة حساب معامل جيني ، وذلك كمايلي:

الجدول رقم (10-13):جدول تحضير الحسابات لتوزيع الملاك حسب المساحة لرسم منحنى لورونز لحساب معامل جيني

$f_i(s_i^{\uparrow} + s_{i-1}^{\uparrow})$	$s_i^{\uparrow} + s_{i-1}^{\uparrow}$	$S_i^{\uparrow}$	f_i
0.0065	0.050	0.050	0.13
0.1023	0.310	0.260	0.33
0.04053	0.579	0.319	0.07
0.25407	0.941	0.622	0.27
0.3244	1.622	1	20
0.7278	/	/	**1**

$I_{Gini} = 1 - \sum_{i=1}^{5} f_i(s_i^{\uparrow} + s_{i-1}^{\uparrow})$

$I_{Gini} = 1 - 0.7278$

$I_{Gini} = 0.2722 = 27.22\%$

نلاحظ أن $I_{Gini} = 27,22\%$ وهي أقل من 30%وعليه يمكن القول أن هذا التوزيع أكثر عدالة، حيث توجد هناك فعلا فوارق في توزيع مساحات الاراضي على مختلف الملاك ، لكن هذه الفوارق يمكن تفسيرها بأسباب موضوعية.

تمرين 14: كانت الأسعار لبعض المواد الإستهلاكية كما في الجدول :

السلعة	الكميات	السعر عام **1993** (دج)	السعر عام **1997** (دج)
A	1 كلغ	18	38
B	1 كلغ	20	30
C	1 علبة	08	15
D	1 كلغ	100	200
E	1 كلغ	30	50

المطلوب : أحسب الرقم القياسي البسيط للأسعار بالطريقيتين النسبية والتجميعية بإعتبار عام 1993 سنة الأساس.

الحل : حساب الرقم القياسي البسيط للأسعار بالطريقيتين النسبية والتجميعية بإعتبار عام 1993 سنة الأساس:

-الطريقة النسبية: $I_{(\boldsymbol{S.R.P})} = \frac{\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}}}{m} \times 100 = \frac{\frac{38}{18}+\frac{30}{20}+\frac{15}{8}+\frac{200}{100}+\frac{50}{30}}{5} \times 100 = 183.05\%$

نسبة التغير في السعر:

$$\Delta = I_{(\boldsymbol{P})} - 100 = 183.05 - 100 = 83.05\%$$

أي هناك زيادة في الأسعار بـــ :83.05 %

-الطريقة التجميعية : $I_{(\boldsymbol{S.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}}{\sum_{i=1}^{n} P_{0,i}} \times 100 = \frac{38+30+15+200+5}{18+20+8+100} \times 100 = 189.20\%$

نسبة التغير في السعر:

$$\Delta = I_{(\boldsymbol{P})} - 100 = 189.20 - 100 = 89.20\%$$

أي هناك زيادة في الأسعار بـــ:89.20 %

تمرين 15:

الجدول التالي يبين أسعار وكميات أربع مواد استهلاكية في الجزائر لسنتي 1983 و 1984 كما يلي:

السلعة	1983		1984	
	السعر P0	الكمية Q0	السعر P1	الكمية Q1
أ	2.5	3	3.5	4
ب	0.4	2	0.7	3
ج	1.2	4	1.4	5
د	0.8	1	1.1	3

المطلوب: باعتبار سنة 1983 سنة الأساس ، أحسب:

-الرقم القياسي المرجح : باش-لاسبير-فيشر-مارشال.

الحل:

الجدول رقم (10-14) : مجاميع كميات المواد استهلاكية

السلعة	P0 1983	Q0 1983	Q0%	P1-1984	Q1-1984	Q1%	$\frac{Q_{0,i}+Q_{1,i}}{2}$	$\frac{Q_{0,i}\%+Q_{1,i}\%}{2}$
أ	2.5	3	0.3	3.5	4	0.267	3.5	0.2835
ب	0.4	2	0.2	0.7	3	0.2	2.5	0.2
ج	1.2	4	0.4	1.4	5	0.333	4.5	0.3665
د	0.8	1	0.1	1.1	3	0.2	2	0.15

حساب الرقم القياسي المرجح للأسعار: لاسبير، باش، فيشر، ومارشال، على إعتبار أن 1983سنة الأساس

1- حساب الرقم القياسي المرجح للأسعار للاسبير:

- الطريقة النسبية:

$$I_{(L.R.P)} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times Q_{0,i}\,\%\right) \times 100$$

$$I_{(L.R.P)} = \left(\frac{3.5}{2.5} \times 0{,}3 + \frac{0.7}{0.4} \times 0{,}2 + \frac{1.4}{1.2} \times 0{,}4 + \frac{1.1}{0{,}8} \times 0{,}1\right) \times 100$$

$$I_{(L.R.P)} = 1.367 \times 100 = 136{,}7\%$$

نسبة التغير في السعر: $\Delta = 136.7 - 100 = 36{,}7\%$ ،أي أن هناك زيادة في الأسعار بـــــ 36,7%.

– الطريقة التجميعية:

$$I_{(\boldsymbol{L.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{0,i}}{\sum_{i=1}^{n} P_{0,i}Q_{0,i}} \times 100 = \left(\frac{3.5\times3+0.7\times2+1.4\times4+1.1\times1}{2.5\times3+0.4\times2+1.2\times4+0.8\times1}\right) \times 100$$

$$I_{(\boldsymbol{L.A.P})} = \frac{18.6}{13.9} \times 100 = 1{,}338 \times 100 = 133{,}8\%$$

نسبة التغير في السعر: $\Delta = 133.8 - 100 = 33{,}8\%$ ، أي أن هناك زيادة في الأسعار بـــــ 33,8%.

2– حساب الرقم القياسي المرجح للأسعار لباش:

– الطريقة النسبية:

$$I_{(\boldsymbol{P.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times Q_{1,i}\,\%\right) \times 100$$

$$I_{(\boldsymbol{P.R.P})} = \left(\frac{3.5}{2.5} \times 0.267 + \frac{0.7}{0.4} \times 0.2 + \frac{1.4}{1.2} \times 0.333 + \frac{1.1}{0.8} \times 0.2\right) \times 100$$

$$I_{(\boldsymbol{P.R.P})} = 1{,}387 \times 100 = 138{,}7\%$$

نسبة التغير في السعر: $\Delta = 138{,}7 - 100 = 38{,}7\%$ ، أي أن هناك زيادة في الأسعار بـــــ 38,7%.

– الطريقة التجميعية:

$$I_{(\boldsymbol{P.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{1,i}}{\sum_{i=1}^{n} P_{0,i}Q_{1,i}} \times 100 = \left(\frac{3.5\times0.267+0.7\times0.2+1.4\times0.333+1.1\times0.2}{2.5\times0.267+0.4\times0.2+1.2\times0.333+0.8\times0.2}\right) \times 100$$

$$I_{(\boldsymbol{P.A.P})} = \frac{1.7607}{1.3071} \times 100 = 1{,}347 \times 100 = 134{,}7\%$$

نسبة التغير في السعر: $\Delta = 134{,}7 - 100 = 34{,}7\%$ ، أي أن هناك زيادة في الأسعار بـــــ 34,7%.

3– حساب الرقم القياسي المرجح للأسعار لفيشر:

– الطريقة النسبية: $I_{(\boldsymbol{F.R.P})} = \sqrt{I_{(L.R.P)} \times I_{(P.R.P)}} = \sqrt{136{,}7 \times 138{,}7} = 137{,}70$

نسبة التغير في السعر: $\Delta = 137{,}70 - 100 = 37{,}70\%$ ،أي أن هناك زيادة في الأسعار بـــــ 37,70%.

– الطريقة التجميعية: $I_{(\boldsymbol{F.A.P})} = \sqrt{I_{(L.A.P)} \times I_{(P.A.P)}} = \sqrt{133{,}8 \times 134{,}7} = 134{,}25$

نسبة التغير في السعر: $\Delta = 134{,}25 - 100 = 34{,}25\%$ ، أي أن هناك زيادة في الأسعار بـ 34,25%.

4– حساب الرقم القياسي المرجح للأسعار لمارشال:

– الطريقة النسبية:

$$I_{(\boldsymbol{M.R.P})} = \left(\sum_{i=1}^{n} \frac{P_{1,i}}{P_{0,i}} \times \frac{Q_{0,i}\% + Q_{1,i}\%}{2}\right) \times 100$$

$$I_{(\boldsymbol{M.R.P})} = \left(\frac{3.5}{2.5} \times 0.2835 + \frac{0.7}{0.4} \times 0.2 + \frac{1.4}{1.2} \times 0.3665 + \frac{1.1}{0.8} \times 0.15\right) \times 100$$

$$I_{(\boldsymbol{M.R.P})} = 1{,}380 \times 100 = 138\%$$

نسبة التغير في السعر: $\Delta = 138 - 100 = 38\%$ ، أي أن هناك زيادة في الأسعار بـــــ 38%.

– الطريقة التجميعية:

$$I_{(M.A.P)} = \frac{\sum_{i=1}^{n} P_{1,i}\frac{Q_{0,i}+Q_{1,i}}{2}}{\sum_{i=1}^{n} P_{0,i}\frac{Q_{0,i}+Q_{1,i}}{2}} \times 100 = \left(\frac{3.5\times3.5+0.7\times2.5+1.4\times4.5+1.1\times2}{2.5\times3.5+0.4\times2.5+1.2\times4.5+0.8\times2}\right) \times 100$$

$$I_{(M.A.P)} = \frac{22.5}{16.75} \times 100 = 1{,}343 \times 100 = 134{,}33\%$$

نسبة التغير في السعر: $\Delta = 134{,}33 - 100 = 34{,}33\%$ ، أي أن هناك زيادة في الأسعار بـــــ 34,33%.

تمرين 16:

البيانات التالية تبين أسعار مجموعة من السلع بالدينار الجزائري في سنتي 2000 و 2024 كما يلي:

البيان / السلعة	وحدة القياس	السعر في سنة **2000**	الكمية المستهلكة في سنة **2000**	السعر في سنة **2024**	الكمية المستهلكة في سنة **2024**
حليب	**لتر**	50	110	130	160
سميد	**قنطار**	3200	08	4000	12
سكر	**كلغ**	50	30	95	35
موز	**كلغ**	140	40	400	05

المطلوب:

1- أحسب الرقم القياسي البسيط للأسعار (بالطريقتين النسبية والتجميعية)، مع تحديد نسبة التغير Δ؟

2- أحسب الرقم القياسي للأسعار مرجحا بالكميات للاسبير و باش (بالطريقة التجميعية)، مع تحديد نسبة التغير Δ؟

الحل:

1-حساب الرقم القياسي البسيط للأسعار:

- الطريقة النسبية:

$$I_{(S.R.P)} = \frac{\sum_{i=1}^{n}\frac{P_{1,i}}{P_{0,i}}}{M} \times 100 = \frac{\frac{130}{50}+\frac{4000}{3200}+\frac{95}{50}+\frac{400}{140}}{4} \times 100 = 215.25\ \%$$

$$\Delta = I_{(S.R.P)} - 100\ \% = 215.25 - 100\ \% = 115.25\ \%$$

أي أن هناك زيادة في الأسعار بـــــ 115,25%.

- الطريقة التجميعية:

$$I_{(S.A.P)} = \frac{\sum_{i=1}^{n} P_{1,i}}{\sum_{i=1}^{n} P_{0,i}} \times 100 = \frac{130+4000+95+400}{50+3200+50+1} \times 100 = 134.45\ \%$$

$$\Delta = I_{(S.A.P)} - 100\ \% = 134.45 - 100\ \% = 34.45\ \%$$

أي أن هناك زيادة في الأسعار بـــــ 34,45%.

2- الرقم القياسي للأسعار مرجحا بالكميات للاسبير و باش(بالطريقة التجميعية) :

- الرقم القياسي المرجح للأسعار للاسبير:

$$I_{(\boldsymbol{L.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{0,i}}{\sum_{i=1}^{n} P_{0,i}Q_{0,i}} \times 100 = \left(\frac{130\times110+4000\times8+95\times30+400\times40}{50\times110+3200\times8+50\times30+1 \quad \times40}\right) \times 100 = \frac{65150}{38200} \times 100 = 170.55\%$$

$$\Delta = I_{(\boldsymbol{L.A.P})} - 100\,\% = 170.55 - 100\,\% = 70.55\,\%$$

أي أن هناك زيادة في الأسعار بـــــ 70.55%.

- الرقم القياسي المرجح للأسعار لباش:

$$I_{(\boldsymbol{P.A.P})} = \frac{\sum_{i=1}^{n} P_{1,i}Q_{1,i}}{\sum_{i=1}^{n} P_{0,i}Q_{1,i}} \times 100 = \left(\frac{130\times160+4000\times12+9 \times35+4 \quad \times05}{50\times160+3200\times12+5 \times35+1 \quad \times05}\right) \times 100 = \frac{74125}{48850} \times 100 = 151.74\%$$

$$\Delta = I_{(\boldsymbol{P.A.P})} - 100\,\% = 151.74 - 100\,\% = 51.74\,\%$$

أي أن هناك زيادة في الأسعار بـــــ 51.74%.

تمرين 17:

الجدول التالي يعطي الضغط العالي و الضغط السفلي لدم عينة مكونة من 6 أشخاص :

ضغط الدم العالي X	130	120	130	150	110	140
ضغط الدم السفلي Y	90	80	70	80	80	90

المطلوب: إيجاد معامل سبيرمان للارتباط .

الحل: نقوم بترتيب كل متغير على حدى ثم نحدد رتب كل متغير وفقا لما يلي :

ضغط الدم السفلي Y	الرتب
70	1
80	3
80	3
80	3
90	5.5
90	5.5

ضغط الدم العالي X	الرتب
110	1
120	2
130	3.5
130	3.5
140	5
150	6

ثم نقوم بحساب المجاميع كالتالي :

الجدول رقم (10-15) : مجاميع قيم المتغيرين x و y

	ضغط الدم العالي X	ضغط الدم السفلي Y	رتب X	رتب Y	d_i	d_i^2
	130	90	3.5	5.5	-2	4
	120	80	2	3	-1	1
	130	70	3.5	1	2	4
	150	80	6	3	3	9
	110	80	1	3	-2	4
	140	90	5	5.5	-0.5	0.25
Σ	/	/	/	/	/	**22.25**

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} = 1 - \frac{6(22.25)}{6(6^2-1)} = 0.36$$

بما أن $r_s = 0.36$ فإن هناك علاقة طردية ضعيفة بين المتغيرين .

تمرين 18:

البيانات المدرجة بالجدول التالي الجدول التالي تعطي دخل 06 عائلات و مقدار انفاقهم (بآلاف الدينارات – جزائري) :

الدخل X	300	180	200	150	230	250
الانفاق Y	250	160	170	130	180	220

المطلوب:

1- أحسب معامل الارتباط بين المتغيرين X ، Y .

2- أوجد معادلة الانحدار Y على X .

3- قدر انفاق الأسرة التي يبلغ دخلها 280 الف دج.

الحل:

1-لحساب معامل الإرتباط نحتاج الى تكوين الجدول التالي :

الجدول رقم (10-16) : مجاميع قيم المتغيرين x و y

	X	Y	XY	X^2	Y^2
	300	250	75000	90000	62500
	180	160	28800	32400	25600
	200	170	34000	40000	28900
	150	130	19500	22500	16900
	230	180	41400	52900	32400
	250	220	55000	62500	48400
Σ	**1310**	**1110**	**253700**	**300300**	**214700**

$$\bar{x} = \frac{\sum x}{n} = \frac{1310}{6} = 218.33$$

$$\bar{y} = \frac{\sum y}{n} = \frac{1110}{6} = 185$$

$$\sigma_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{300300}{6} - 218.33^2} = 48.80$$

$$\sigma_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2} = \sqrt{\frac{214700}{6} - 185^2} = 39.47$$

$$r = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{\sigma_x . \sigma_y} = \frac{\frac{253700}{6} - (218.33)(185)}{(48.80)(39.47)} = 0.98$$

يوجد هناك إرتباط طردي (لأنه موجب) وقوي (لأنه قريب من الواحد) ، يمكن القول بوجود علاقة قوية بين المتغيرين X ، Y (الدخل و الانفاق) و هي علاقة قوية وطردية في نفس الاتجاه.

2-إيجاد معادلة الانحدار Y على X :

$$\widehat{\beta} = \frac{\Sigma xy - n\bar{y}}{\Sigma x^2 - n\bar{x}^2} = \frac{253700 - 6(218.33)(185)}{300300 - 6(218.33)^2} = 0.79$$

$$\widehat{\alpha} = \bar{Y} - \widehat{\beta}\bar{X}$$

$$\widehat{\alpha} = 185 - 0.79(218.33) = 12.52$$

معادلة الانحدار هي: $\hat{Y}_i = 12.52 + 0.79X_i$

3- تقدير انفاق الأسرة التي يبلغ دخلها 280 ألف دج:

لدينا معادلة الانحدار هي :

$$\hat{Y}_i = 12.52 + 0.79X_i$$
$$\hat{Y}_i = 12.52 + 0.79(280) = 233.72$$

إذا كان دخل الأسرة : 280 الف دج ، فيقدر إنفاقها بمبلغ 233.72 ألف دج.

قائمة المراجع

1- ابراهيم علي ابراهيم عبد ربه ، مقدمة في الاحصاء الوصفي ، الدار الجامعية، مصر، 2005.

2- أحمد فوزي ملوخيه ، مبادئ علم الاحصاء ، مركز الاسكندرية للكتاب،مصر، 2009.

3- جلاطو جيلالي ، الاحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الجزائر، 2001.

4- شرف الدين خليل ، الاحصاء الوصفي، مكتبة شبكة الابحاث والدراسات الاقتصادية.

5- طه حسين الزبيدي ، مبادئ الاحصاء ، دارعيداء للنشر والتوزيع، عمان- الاردن، 2012.

6- محمد صبحي أبو صالح ، مبادئ الاحصاء ، دار اليازوري العلمية للنشر والتوزيع، الأردن، 2010.

7- محمود محمد سليم صالح ، مقدمة في الاحصاء ، مكتبة المجتمع العربي للنشر والتوزيع،عمان، الطبعة الأولى 2008.

8- ميدون سيساني ، مطبوعة دروس بعنوان "التحليل الديمغرافي" محاضرات في مقياس الاحصاء 1، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة تيارت، السنة الجامعية 2019-2020.

9- صلاح الدين حسين الهيتي ، الاساليب الاحصائية في العلوم الادارية: تطبيقات باستخدام spss، دار وائل للطباعة والنشر،عمان، الطبعة الثانية 2006.

10- عبد الحكيم بيصار، محاضرات في الاحصاء 1، مطبوعة دروس مقدمة الى طلبة الجذع المشترك علوم اقتصادية، التسيير والعلوم التجارية، جامعة المسيلة، 2019-2020.

11- عبد الحليم الحمزة ، كتاب بيداغوجي في الاحصاء 1 ، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة تبسة، الموسم الجامعي 2022-2023.

12- عبد الحميد قطوش ، محاضرات في الاحصاء 1، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة المسيلة، 2022-2023.

13- سالم عيسى بدر، عماد غصاب عبابنة ، مبادئ الاحصاء الوصفي والاستدلالي ، دار المسيرة للنشر والتوزيع والطباعة،الاردن،الطبعة الاولى 2007.

14- سهام بشكيط ، محاضرات في مقياس الاحصاء 1، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة جيجل، السنة الجامعية 2021-2022.