الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة محمد الصديق بن يحي - جيجل

كلية العلوم الإنسانية والاجتماعية

قسم علوم وتقنيات النشاطات البدنية والرياضية

محاضرات الإحصاء التطبيقي

السنة الاولى ماستر : تحضيربدني رياضي

إعداد الأستاذ/بولحليب مبروك

الموسم الجامعي 2023/2024

الفهرس

الفصل الأول ماهية الإحصاء

الفصل الثاني معاملات الارتباط

الفصل الثالث المقاييس الإحصائية

مقدمة

ليس علم الإحصاء علما حديثا، وإنما نشأ منذ القدم وتطور مع متطلبات البشرية حتى وصل إلى ما هو عليه الآن، وسيستمر هذا التطور مستقبلا، فقد ثبت لدارسي الآثار أن بناء الأهرامات في مصر قد سبقته عملية تعداد لسكان مصر وثرواتها وموسمية الأعمال بها، كما أنه كان يتم توزيع الأراضي على السكان عن طريق إجراء تعداد للسكان وكذلك الأراضي الصالحة للزراعة، كما كان خلفاء الدولة الإسلامية في العهد العباسي يقومون بتعداد السكان والثروات لتحديد الإمكانيات العسكرية والاقتصادية.

وحتى القرن السابع عشر اقتصر جمع البيانات على الطرق الوصفية، ولم تستخدم الأرقام للدلالة على ما يجمع من معلومات إلا في انجلترا خلال القرن السابع عشر. وقد أطلق على العلم الذي يبحث في طرق جمع البيانات الرقمية التي تهم الدولة تسمية "الحساب السياسي" والذي كان يشمل إحصاءات المواليد والوفيات وعدد السكان وكذلك مقدار الثروات المادية ودخول الأفراد وحصيلة الضرائب وغيرها من المنتجات، وبصفة عامة كل مصادر إيرادات الدولة ومجالات الإنفاق العامة.

وفي خلال القرن الثامن عشر تطورت الرياضيات مما دعا إلى تطور الإحصاء على أيدي الرياضيين والذي أدى إلى نشوء النظريات والطرق العلمية التي ما زالت تعتبر من أساسيات علم الإحصاء إلى يومنا هذا، ثم أصبح علم الإحصاء وخاصة في النصف الثاني من القرن التاسع عشر من العلوم الرئيسية التي تدرس بالجامعات.

أما في القرن العشرين، فقد تطور علم الإحصاء بشكل كبير، واتصل بالعلوم الإدارية، وأصبح مواكبا لتطورها، مما دعا بعض الكتاب إلى تعريف علم الإحصاء بأنه العلم الذي يبحث في طرق اتخاذ القرارات الإدارية في ظل عدم التأكد، وبذلك دخل علم الإحصاء مجال الصناعة والإنتاج، ثم بتحول عدة دول من النظام الرأسمالي إلى نظام التخطيط الاقتصادي العلمي ظهرت علوم جديدة هي مزج بين الإحصاء والعلوم الأخرى مثل القياس الاقتصادي (اقتصاد وإحصاء)، وبحوث العمليات (إدارة الأعمال والإحصاء)، وبذلك تظل هذه العلوم تطبيقا لنفس النظريات والطرق العلمية التي يتكون منها علم الإحصاء.

I. الفصل الأول

ماهية الإحصاء

1. مفهوم الإحصاء:

هو العلم الذي يهتمّ بجمع البيانات الرقمية، ومن ثم تنظيمها، وترتيبها، وتحليلها، بهدف الوصول إلى نتائج معينة لتوضيح ظاهرة أو حالة ما، أو بأنه العلم الذي يهتم بالطريقة التي يتم من خلالها جمع البيانات والمعلومات وتحويلها إلى صورة عددية، حيث تُجمَع البيانات من خلاله بشكل منتظم، وفيما يخص استخدامات علم الإحصاء فهي كثيرة؛ كاستخدامه في العلوم الطبية، وعلم الاجتماع، والاقتصاد، والصناعة، والكيمياء، والرياضة، والإدارة، وغيرها العديد من المجالات.

2. أنواع علم الإحصاء: وفيما يأتي أبرز أنواع علم الإحصاء :

أ. **الإحصاء الوصفي:** يتضمن علم الإحصاء كل ما يخص جمع وتحليل وتفسير المشاهدات، كما أنه يتضمن تمثيل البيانات؛ كحساب معدل الدخل الشهري والنفقات لعائلة ما، أو حساب نسب الطلاق والزواج في أحد الدول، أو عمل استبيان لتبيّن رأي المجتمع حول نقطة معينة، ولهذا يستخدم الإحصاء الوصفي لوصف البيانات وتحويلها إلى أرقام لعرضها بالصورة المناسبة سواء كان ذلك باستخدام الخرائط، أو الجداول الإحصائية، أو الرسومات والمنحنيات البيانية التي توضّح الظواهر بشكل أفضل من أي أسلوب آخر، كما يتضمن حساب بعض المؤشرات الإحصائية؛ كمقاييس النزعة التي تتضمن، المنوال، والوسط، والوسيط،...وغيرها، ومقاييس التشتت التي تتضمن الانحراف المعياري، والتباين، والمدى وغيرها ...

ب. **الإحصاء الاستدلالي:** يُطلق عليه أيضاً الإحصاء التحليلي، وهو يهتم بوضع القرارات المناسبة بناء على النتائج التي تم استنتاجها من البيانات التي تم جمعها، وتُستخدم لتحقيق ذلك عدة أساليب، وهي:

- **التقدير:** يعني تقدير معالم المجتمع المطلوب دراسته، عن طريق التقدير النقطي؛ كتقدير الوسط الحسابي للمجتمع، أو التقدير بفترة من خلال تقدير قيمة المجتمع ضمن فترة لها حدان: أدنى وأعلى.
- **اختبار الفرضيات:** يعني استخدام المشاهدات التي تم جمعها من المجتمع، والمؤشرات الإحصائية، بهدف الوصول إلى قرار نحو الفرضيات التي تم تنبؤها في بداية الدراسة، وبناءً عليه يتم قبول الفرضية أو رفضها.

3. مراحل العملية الإحصائية: تتضمن العملية الإحصائية مجموعة من المراحل، وهي:

أ. **جمع البيانات:** هي مرحلة جمع المعلومات العددية من مصادر موثوقة؛ كالمصادر الحكومية، أو يمكن الحصول على البيانات من خلال أخذ عينة من المشاهدات بدلاً من مسح الكل

ب. **تنظيم البيانات:** وهي مرحلة ترتيب وتنظيم المشاهدات ضمن جداول خاصة تُسمّى بالجداول الإحصائية، أو يمكن تنظيمها على شكل رسومات بيانية، وذلك بهدف تسهيل عرضها ومعالجتها بأسلوب رياضي .

ت. المعالجة الرياضية: وهي المرحلة التي يتم من خلالها الوصول إلى نتائج عددية، عن طريق معالجة المشاهدات والبيانات، وتتميز هذه النتائج بأن لها مؤشرات تدل على مدى تقاربها أو تشتتها عن بعضها البعض؛ كمقاييس النزعة المركزية، أو معاملات الارتباط.

ث. تحليل النتائج: وهي إحدى أهم المراحل التي تمر بها العملية الإحصائية، حيث إنها تعمل على تحويل البيانات الصماء إلى معلومات واضحة، فهذه العملية تتطلب الصدق والدقة، وعدم التحيّز، كما يجب أن يكون الباحث على معرفة جيدة واطلاع على موضوع البحث بشكل تام.

4. البيانات الإحصائية:

هي عبارة عن مجموعة من البيانات والمعلومات الخام التي تمثل علم الإحصاء، والمتعلقة بالظاهرة التي تتم دراستها، وتُصنّف البيانات إلى صنفين رئيسيين هما:

أ. **البيانات النوعية:** هي عبارة عن البيانات التي لا تقاس بأعداد؛ مثل الحالة الاجتماعية (غني، متوسط، فقير)، والجنس (ذكر، أنثى)، وهي تشمل البيانات الترتيبية التي يمكن ترتيبها تصاعديا وتنازلياً، والبيانات الاسمية التي لا يمكن ترتيبها.

ب. **البيانات الكمية:** هي البيانات التي تُقاس من خلال الأرقام مثل؛ أعداد العاملين، والطول، والوزن وغيرها...

5. مصادر البيانات الإحصائية: تُجلَب البيانات من مصادر عدة، منها ما يأتي:

أ. مصادر من الميدان: حيث يُحصل عليها بشكل مباشر عن طريق جمع المعلومات والتحري عن الحقائق حول دراسة معينة بنفسه؛ كالاستبيان مثلاً.

ب. مصادر رسمية: حيث تتولّى المؤسسات المختصة مسؤولية جمع البيانات الإحصائية عن الظواهر باختلاف أنواعها؛ مثل: الظواهر الصحية، والعلمية، والاقتصادية.

6. طرق جمع البيانات الإحصائية: يتم جمع البيانات الإحصائية من خلال عدة طرق، منها ما يأتي:

أ. **الطريقة المباشرة:** هي الطريقة التي يتم من خلالها جمع البيانات من موقع الحدث وأرض الواقع بشكل مباشر.

ب. **الطريقة غير المباشرة:** هي الطريقة التي يتم من خلالها جمع البيانات من خلال السجلّات والوثائق الرسمية والتاريخية.

ت. **طريقة الاستبيان:** هي عبارة عن حزمة من الأوراق التي يتم توزيعها على مجموعة من الأفراد بهدف الإجابة عن مجموعة من الأسئلة حول موضوع معين.

ث. **طريقة المقابلات الشخصية:** هي الطريقة التي يتم من خلالها سؤال الباحث لأفراد المجتمع المراد دراسته بشكل شخصي ومباشر.

ج. **طريقة الاختبارات الخاصة:** تستخدم هذه الطريقة في أوضاع خاصة؛ كامتحان مستوى الذكاء مثلا.

7. **أهمية الإحصاء في المجال الرياضي:** تعتمد جميع العلوم على الإحصاء في معالجة البيانات والتربية الرياضية لا تختلف عنهم في شيء ويمكننا إن ندرج أهمية علم الإحصاء أو مدى الاستفادة من الإحصاء في المجال الرياضي في النقاط الآتية:

ح. **اتخاذ القرار**: عندما يختلف رياضيان في صفة بدنية إي عندما يؤدي احدهما (12مرة) ثني ومد الذراعين ويؤدي الأخر (13مرة) فان الإحصاء هو الوحيد القادر على اعتبار فرق مرة واحدة بين الرياضيين وهو فرق يعود إلى الصدفة أم فرق حقيقي.

خ. **إيجاد العلاقة بين المتغيرات**: نستطيع باستخدام الإحصاء إن نقول بان الرياضي السريع هو رياضي قوي وان هذا الحديث لا يمكن التأكد منه إلا بإيجاد مقدار التغاير بين السرعة والقوة.

د. **التنبؤ:** عند توفر بيانات عن حالة لاعب من أسبوع إلى أخر أو شهر إلى أخر فان الإحصاء يستطيع إن يتنبأ كيف ستكون حالة اللاعب في الشهر القادم.

ذ. **الفرز:** للإحصاء القدرة على فرز العامل الأكثر أهمية من العوامل الأخرى التي تساهم في رفع المستوى الرياضي، فمثلا إي الصفات هي الأكثر أهمية في فعالية الوثب الطويل القوة أم السرعة أم المطاولة أم التوافق أم غير هذه الصفات ثم إي صفة تأتي بالدرجة الثانية.

ر. **تصنيف وترتيب:** يمكننا باستخدام الإحصاء تصنيف اللاعبين إلى فئات وحسب المستويات كما باستطاعة الإحصاء وضع ترتيب اللاعبين باستخدام معايير معينة.

ز. **تقويم وبناء الاختبار:** تستطيع باستخدام الإحصاء تقويم مدى صلاحية اختبار معين لإيجاد الفروق بين اللاعبين، فمثلا هل باستطاعة اختبار ثني ومد الذراعين إيجاد الفروق الفردية بين مجموعة رياضيين في القدرة العضلية للذراعين. وإذا تم ذلك فيمكن بناء هذا الاختبار.

س. **التقييم:** تستطيع باستخدام الإحصاء وضع تقييم للاعب اي تثمين ما يمتلكه من صفة مهارية أو بدنية.

II. الفصل الثاني

معاملات الارتباط

1. **الارتباط Corrélation:** هو تعيين طبيعة و قوة العلاقة بين متغيرين أو عدمها

- معامل الارتباط Coefficient de Corrélation هو مؤشر هذه العلاقة
- أول خطوة في تحديد طبيعة العلاقة هي رسم شكل الانتشار

- إذا كان لدينا متغيران فقط. المتغير X و هو متغير يتم تحديده من قبل الباحث أو الشخص الذي يقوم بالدراسة و هو يسمى بالمتغير المستقل variable Independent

- يرافق المتغير X متغير آخر Y و يسمى بالمتغير التابع variable dépendent و هو متغير تابع لأن نتيجته غير محددة و تعتمد على قيم المتغير المستقل.

2. **أنواع الارتباط:**

أ. **الارتباط الموجب (الطردي) (corrélation positive):** بأنه علاقة بين متغيرين (x,y) بحيث إذا تغير أحد المتغيرين فإن الآخر يتبعه في نفس الاتجاه.

ب. **الارتباط السالب (العكسي) (corrélation négative):** بأنه علاقة بين متغيرين (x,y) بحيث إذا تغير أحد المتغيرين فإن الآخر يتبعه في الاتجاه المضاد.

3. **قياس الارتباط:**

- تستخدم معاملات الارتباط لقياس درجة الارتباط بين متغيرين (ظاهرتين) .
- **تعريف معامل الارتباط:**

يعرف معامل الارتباط و الذي يرمز له بالرمز **r** بأنه عبارة عن مقياس رقمي يقيس قوة الارتباط بين متغيرين، حيث تتراوح قيمته بين (+1) و (-1)، أي أن $-1 \leq r \leq +1$. و تدل إشارة المعامل الموجبة على العلاقة الطردية، بينما تدل إشارة المعامل السالبة على العلاقة العكسية.

- يمكن حساب العديد من معاملات الارتباط و يعتمد ذلك على مستوى القياس (اسمي-ترتيبي-فترة-نسبي) للمتغيرات التي تبدو مرتبطة.

و الجدول التالي يوضح أنواع الارتباط و اتجاه العلاقة و شكل الانتشار لكل نوع:

قيمة معامل الارتباط	المعنى
+1	ارتباط طردي تام
من 0.70 إلى 0.99	ارتباط طردي قوي
من 0.50 إلى 0.69	ارتباط طردي متوسط
من 0.01 إلى 0.49	ارتباط طردي ضعيف
0	لا يوجد ارتباط

و ما قيل عن الارتباط الطردي ينطبق على الارتباط العكسي (مع وضع إشارة سالبة)

أ. معامل بيرسون للارتباط الخطي Coefficient de corrélation linéaire de Pearson:

- معامل بيرسون للارتباط الخطي من أكثر معاملات الارتباط استخدامًا خاصة في العلوم الإنسانية و الاجتماعية.
- مستوى القياس المطلوب عند تطبيق معامل بيرسون للارتباط هو أن يكون كلا المتغيرين مقياس فترة أو نسبي أو بمعنى آخر أن تكون بيانات كلا المتغيرين(الظاهرتين) بيانات كمية.
- **حساب معامل بيرسون للارتباط الخطي:**

يمكن حساب معامل بيرسون بدلالة القراءات لبيانات المتغيرين x ,y ، باستخدام الصيغة التالية:

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

حيث: $\sum_{i=1}^{n} x_i y_i$: مجموع حاصل ضرب x في y

$\sum x$: مجموع قيم المتغير x

$\sum y$: مجموع قيم المتغير y

$\sum x^2$: مجموع مربعات قيم المتغير x

$\sum y^2$: مجموع مربعات قيم المتغير y

- **مثال:** سجلت ست قراءات تقريبية لحجم الإنتاج وحجم صادرات النفط الخام بالمملكة العربية السعودية (بالمليار برميل) خلال عدة سنوات كما يلي:

حجم الصادرات (y)	2	2	2	1	1	1
حجم الإنتاج (x)	3	4	2	2	2	2

ادرس وجود علاقة ارتباط خطية بين حجم الإنتاج وحجم صادرات النفط الخام.

- الحل:

	x	y	xy	x^2	y^2
	3	2	6	9	4
	4	2	8	16	4
	2	2	4	4	4
	2	1	2	4	1
	2	1	2	4	1
	2	1	2	4	1
Σ	15 $=\sum x$	9 $=\sum y$	24 $=\sum xy$	41 $=\sum x^2$	15 $=\sum y^2$

$$r_p = \frac{6(24)-(15)(9)}{\sqrt{\left((6\times 41)-15^2\right)\left((6\times 15)-9^2\right)}} =$$

$$\frac{144-135}{\sqrt{(246-225)(90-81)}} = \frac{9}{\sqrt{189}} = \frac{9}{13.75} = 0.65$$

من الملاحظ أن علاقة الارتباط الخطي بين حجم الإنتاج وحجم صادرات النفط الخام علاقة طردية متوسطة.

ملاحظة هامة

إذا كانت قيمة معامل بيرسون تساوي صفرا لا يعني بالضرورة عدم وجود ارتباط بين المتغيرين و لكن قد يوجد ارتباط غير خطي. على سبيل المثال عند دراسة العلاقة بين القلق (Anxiété) مع الانجاز (Performance) نجدها علاقة غير خطية. فمع زيادة قلق الشخص يزداد معه الاهتمام بانجاز العمل و حتى مرحلة معينة نجد أن زيادة القلق تصاحبها قلة الانجاز. في الواقع يكون القلق مصاحبا بقلة الانجاز، لذلك في هذه الحالة فإن معامل بيرسون للارتباط ليس مناسبا لوصف علاقة منحنية. و إذا استخدم معامل بيرسون لوصف مثل هذه البيانات فإن النتيجة هي بخس للعلاقة الحقيقية بين المتغيرين. و بذلك يتضح عمليا فائدة شكل الانتشار قبل الشروع في حساب معامل بيرسون للارتباط.

ب. معامل سبيرمان لارتباط الرتب Spearman Rank Corrélation Coefficient:

- نستخدم معامل سبيرمان لارتباط الرتب إذا كان قياس المتغيرين كليهما مقياس ترتيبي.
- **طريقة حساب معامل سبيرمان لارتباط الرتب:**

إذا فرضنا أن المتغير X له الرتب R_x و أن المتغير Y له الرتب R_y. و بفرض أن d ترمز لفرق الرتبتين، بمعنى $d=R_x-R_y$ فإن معامل سبيرمان لارتباط الرتب يعطى بالصيغة التالية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث n هي عدد الأزواج المرتبة.

- **مثال:** لدراسة علاقة ارتباط تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات، اخترنا خمس طلاب وكانت تقديراتهم كما يلي:

تقديرات الإحصاء (x)	F	A	C	D	B
تقديرات الرياضيات (y)	D	C	B	F	A

هل توجد علاقة ارتباط؟ ما نوعها ومدى قوتها؟

x	y	رتب x	رتبy	d	d^2
F	D	1	2	-1	1
A	C	5	3	2	4
C	B	3	4	-1	1
D	F	2	1	1	1
B	A	4	5	-1	1
			Σ	0	8
				$\sum d$	$\sum d^2$

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{(6)(8)}{5(25-1)} = 1 - \frac{48}{120} = 1 - 0.4 = 0.6$$

نلاحظ وجود علاقة ارتباط طردية متوسطة بين تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات.

III. الفصل الثالث

المقاييس الإحصائية

i. مقاييس النزعة المركزية:

هي مقاييس عددية تحدد موقع التوزيع للبيانات، ويمكن تعريف المتوسطات بأنها القيمة النموذجية الممثلة لمجموعة من البيانات، والتي تميل إلى الوقوع في المركز، لذلك تسمى المتوسطات بمقاييس النزعة المركزية. وهي مهمة في حالة المقارنة بين التوزيعات المختلفة للبيانات وتكون فائدتها أكثر في حالة التوزيعات المتشابهة في طبيعتها وأشكالها ولكنها مختلفة في مواقعها، فمثلا: عند دراسة الإنفاق لعينة من الأسر في الريف، وأخرى في المدينة، فإنه يمكننا المقارنة بينهما من خال هذه المقاييس. وسوف نستعرض أهم مقاييس النزعة المركزية أدناه.

1. المتوسط الحسابي (يرمز له بالرمز $\bar{X}$):

يعرف المتوسط الحسابي لمجموعة من البيانات على انه نسبة إجمالي البيانات إلى عددها .و يختلف حسابه باختلاف طبيعة المتغيرة المدروسة، و مبوبة آو غير مبوبة.

- **المتوسط الحسابي لبيانات غير مبوبة:**

لتكن $x_1.x_2..........x_n$ سلسلة بيانات من n مشاهدة، فإن متوسطها الحسابي هو: $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$

مثال: احسب المتوسط الحسابي للبيانات التالية: 2. 3. 5. 8. 10. 12. 14. 16. 20

المتوسط الحسابي هو: $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{2+3+5+8+10+12+14+16+20}{9} = 10$

- **المتوسط الحسابي لبيانات مبوبة متغيرة كمية متقطعة:** لتكن المتغيرة X. $x_1.x_2..........x_k$ القيم التي تأخذها المتغيرة و $n_1.n_2........n_k$ تكرارات هذه القيم لسلسلة بيانات من n مشاهدة، متوسطها الحسابي هو: $\bar{X} = \frac{\sum_{i=1}^{k} n_i x_i}{n}$

مثال: الجدول الموالي يمثل توزيع الأسر حسب عدد الأولاد لحي يتكون من 50 أسرة

عدد الأولاد (x_i)	0	1	2	3	4	5	6	7	*المجموع*
عدد الأسر(n_i)	1	4	6	9	10	8	7	5	50
$x_i\ n_i$	0	4	12	27	40	40	42	35	172

متوسطها الحسابي هو: $\bar{X} = \frac{\sum_{i=1}^{k} n_i x_i}{n} = \frac{200}{50} = 4$

- **المتوسط الحسابي لبيانات مبوبة متغيرة كمية مستمرة:** لتكن المتغيرة X موزعة في جدول توزيع تكراري يحتوي l فئة و لتكن $x_1.x_2..........x_l$ مراكز هذه الفئات و $n_1.n_2.........n_l$ تكرارات هذه الفئات ولتكن n تمثل إجمالي التكرارات أي أن: $\left(\sum_{i=1}^{l} n_i = n\right)$ متوسطها الحسابي هو: $\bar{X} = \frac{\sum_{j=1}^{l} n_j x_j}{n}$

مثال: الجدول التالي يمثل توزيع الإنتاج السنوي للتمور (بالقنطار) لمجموعة من البساتين الفلاحية (واحد هكتار لكل بستان) في محيطات حاسي بن عبد الله.

فئات الإنتاج بالقنطار	-5	-10	-15	-20	-25	-30	40-35	المجموع
عدد البساتين n_j	2	8	12	30	26	12	10	100
مركز الفئة x_j	7.5	12.5	17.5	22.5	27.5	32.5	37.5	
$n_j x_j$	15	100	210	675	715	390	375	2480

المتوسط الحسابي: $\bar{X} = \frac{\sum_{j=1}^{l} n_j x_j}{n} = \frac{2480}{100} = 24.8$

خواص المتوسط الحسابي:

- سهل الفهم و الاستيعاب و الحساب و يستخدم في حسابه جميع القيم.
- تأثر بالقيم المتطرفة.
- يتأثر بالبيانات المبوبة لمتغيرة مستمرة التي لها الفئة الأولى المفتوحة من الأسفل والفئة الأخيرة مفتوحة منن الأعلى.

2. المنوال (يرمز له بالرمز M_o):

و يحسب المنوال حسب حالة البيانات و طبيعة المتغير الإحصائية.

أ. بيانات غير مبوبة:

المنوال: هو القيمة الأكثر تكرارا أو الأكثر شيوعا لمجموعة من البيانات. يمكن توضيح ذلك بالأمثلة التالية:

مثال 1: لتكن البيانات التالية: 1. 2. 4. 5. 4. 2. 7. 4. 6. 5. 4. 8. 4. 9. 4. 7. 4 المنوال هو: $M_o=4$

مثال 2: إذا كان لدينا القيم التالية: 5. 2. 4. 5. 5. 4. 7. 4. 6. 5. 4. 5. 8. 4 يوجد منوالان: $M_{o1}=4$ و $M_{o2}=5$

مثال 3: لتكن البيانات التالية: 6. 5. 4. 1. 8. 3. 9. 7 لا يوجد منوال.

ب. بيانات مبوبة متغيرة متقطعة:

المنوال: هو القيمة المقابلة لأكبر تكرار.

مثال: اوجد المنوال لتوزيع الأسر حسب عدد الأولاد ل 50 أسرة.

التكرار (عدد الأسر)	المتغير (عدد الأولاد)
2	0
5	1
9	2
11	3
15	4
8	5
50	المجموع

القيمة المقابلة لأكبر تكرار (15) هي: 4

المنوال يساوي: $M_o=4$

ت. **بيانات مبوبة متغيرة مستمرة:**

للمنوال في هذه الحالة عدة تعاريف.

- يمكن قبول (تقدير) مركز الفئة المنوالية كمنوال، خاصة إذا كانت الفئة المنوالية هي الفئة الأولى أو الأخيرة. والفئة المنوالية هي الفئة المقابلة لأكبر تكرار.

عيوب ومزايا المنوال:

أولا: عيوب المنوال

1. نجد بعض التوزيعات التي لا تحتوي على منوال أو لها أكثر من منوال.
2. يتأثر بأخطاء المعاينة.
3. لا يعتمد في إيجاده على كافة البيانات.
4. لا يخضع للعمليات الجبرية.

ثانيا: مزايا المنوال

1. سهل الفهم و الحساب.
2. لا يتأثر بالقيم المتطرفة (الشاذة).
3. يستخدم في أيجاد النزعة المركزية للبيانات الكيفية.
4. لا يتأثر بالبيانات المفتوحة في طرفي التوزيع.

3. المتوسط التربيعي (يرمز له بالرمز Q):

يعرف المتوسط التربيعي لمجموعة من البيانات على انه الجذر التربيعي لمتوسط مربعات البيانات و يختلف حسابه باختلاف طبيعة المتغيرة المدروسة، و مبوبة آو غير مبوبة.

أ. **المتوسط التربيعي لبيانات غير مبوبة :** لتكن $x_1.x_2..........x_n$ سلسلة بيانات من n مشاهدة، فإن متوسطها التربيعي هو:

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

مثال: احسب المتوسط التربيعي للبيانات التالية: 2. 3. 5. 8. 10. 12. 14. 16. 20

المتوسط التربيعي هو:

$$Q = \sqrt{\frac{2^2+3^2+5^2+8^2+10^2+12^2+14^2+16^2+20^2}{9}} = 11.53$$

ب. **المتوسط التربيعي لبيانات مبوبة متغيرة كمية متقطعة:** لتكن المتغيرة X. $x_1.x_2..........x_k$ القيم التي تأخذها المتغيرة و $n_1.n_2........n_k$ تكرارات هذه القيم لسلسلة بيانات من n مشاهدة، متوسطها التربيعي هو:

$$\boldsymbol{Q = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{n}} = \sqrt{\frac{n_1 x_1^2 + n_2 x_2^2 + \cdots + n_{n_k} x_{n_k}^2}{n}}}$$

مثال: حساب المتوسط التربيعي للبيانات التالية:

$\boldsymbol{x_i}$	1	3	5	7	*المجموع*
$\boldsymbol{n_i}$	15	20	15	10	60
$\boldsymbol{n_i x_i{}^2}$	15	180	375	49	1060

متوسطها التربيعي هو: $Q = \sqrt{\frac{1060}{60}} = 4.2$

ت. **المتوسط التربيعي لبيانات مبوبة متغيرة كمية مستمرة:** لتكن المتغيرة X موزعة في جدول توزيع تكراري يحتوي k فئة و لتكن $c_1.c_2........c_k$ مراكز هذه الفئات $n_1.n_2........n_k$ تكرارات هذه الفئات ولتكن n تمثل إجمالي التكرارات أي أن: $(\sum_{i=1}^{l} n_i = n)$ متوسطها التربيعي هو:

$$\boldsymbol{Q = \sqrt{\frac{\sum_{i=1}^{k} n_i c_i^2}{n}} = \sqrt{\frac{n_1 c_1^2 + n_2 c_2^2 + \cdots + n_{n_k} c_{n_k}^2}{n}}}$$

مثال: الجدول التالي يمثل توزيع الإنتاج السنوي للتمور (بالقنطار) لمجموعة من البساتين الفلاحية (واحد هكتار لكل بستان) في محيطات حاسي بن عبد الله.

فئات الإنتاج بالقنطار	-5	-10	-15	-20	-25	-30	40-35	المجموع
عدد البساتين n_i	2	8	12	30	26	12	10	100
مركز الفئة c_i	7.5	12.5	17.5	22.5	27.5	32.5	37.5	
$n_i c_i^2$	112.5	1250	3675	15187.5	19662.5	12675	14062.5	66625

المتوسط التربيعي : $Q = \sqrt{\frac{66625}{100}} = 25.81$

خواص المتوسط التربيعي:

- صعب الفهم.
- يستخدم في حسابه جميع القيم (موجبة، سالبة، معدومة).
- يستخدم أكثر في مجال الفيزياء.

4. **المتوسط الهندسي (يرمز له بالرمز G):**

يعرف المتوسط الهندسي لمجموعة من البيانات بأنه الجدر النوني لجداء المشاهدات ولوغاريتم المتوسط الهندسي هو المتوسط الحسابي للوغاريتمات المشاهدات. و يختلف حسابه باختلاف طبيعة المتغيرة المدروسة، و مبوبة آو غير مبوبة.

- **المتوسط الهندسي لبيانات غير مبوبة:** لتكن $x_1.x_2..........x_n$ سلسلة بيانات من n مشاهدة، فإن متوسطها الهندسي هو:

$$G = \sqrt[n]{\prod_{i=1}^{n} x_i} = [x_1 . x_2 \dots \dots x_n]^{\frac{1}{n}} \rightarrow \log G = \frac{1}{n} \sum_{i=1}^{n} \log x_i$$

مثال: احسب المتوسط الهندسي للبيانات التالية: 2. 4. 8

المتوسط الهندسي هو: $G = \sqrt[3]{2 \times 4 \times 8} = 4$

- **المتوسط الهندسي لبيانات مبوبة متغيرة كمية متقطعة:** لتكن المتغيرة X. $x_1.x_2..........x_k$ القيم التي تأخذها المتغيرة و $n_1.n_2.........n_k$ تكرارات هذه القيم لسلسلة بيانات من n مشاهدة، متوسطها الهندسي هو:

$$G = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}} = \left[x_1^{n_1} . x_2^{n_2} \dots \dots x_k^{n_k}\right]^{\frac{1}{n}} \rightarrow \log G = \frac{1}{n} \sum_{i=1}^{k} n_i \log x_i$$

مثال: حساب المتوسط الهندسي للبيانات التالية:

x_i	1	3	5	7	*المجموع*
n_i	15	20	15	10	60
$n_i \log x_i$	0	9.54	10.48	8.45	28.48

متوسطها الهندسي هو: $\log G = \frac{1}{n}\sum_{i=1}^{k} n_i log x_i = \frac{28.48}{60} = 0.47 \rightarrow G = 2.98$

- **المتوسط الهندسي لبيانات مبوبة متغيرة كمية مستمرة:** لتكن المتغيرة X موزعة في جدول توزيع تكراري يحتوي k فئة و لتكن $c_1.c_2.........c_k$ مراكز هذه الفئات و $n_1.n_2.........n_k$ تكرارات هذه الفئات ولتكن n تمثل إجمالي التكرارات أي أن: $(\sum_{i=1}^{k} n_i = n)$ متوسطها الهندسي هو:

$$G = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}} = [c_1^{n_1}.c_2^{n_2}\ldots\ldots c_k^{n_k}]^{\frac{1}{n}} \rightarrow \log G = \frac{1}{n}\sum_{i=1}^{k} n_i log c_i$$

مثال: الجدول التالي يمثل توزيع الإنتاج السنوي للتمور (بالقنطار) لمجموعة من البساتين الفلاحية (واحد هكتار لكل بستان) في محيطات حاسي بن عبد الله.

فئات الإنتاج بالقنطار	-5	-10	-15	-20	-25	-30	40-35	المجموع
عدد البساتين n_i	2	8	12	30	26	12	10	100
مركز الفئة c_i	7.5	12.5	17.5	22.5	27.5	32.5	37.5	
$n_i \log x_i$	1.75	8.77	14.91	40.56	37.42	18.14	15.74	137.31

المتوسط الهندسي: $\log G = \frac{1}{100}\sum_{i=}^{7} n_i \log c_i = \frac{137.31}{100} = 3.93$

خوص المتوسط الهندسي

- صعب الفهم و الحساب.
- تستخدم في حسابه جميع القيم المعطاة.
- يستخدم في حساب نسب الزيادة في الظواهر كالمبيعات و الأسعار و الأرقام القياسية.
- مجموع انحرافات لوغاريتمات قيم المتغيرة المدروسة عن لوغاريتم متوسطها الهندسي يساوي صفر.
- لا يجب أن تأخذ قيمة من قيم المتغيرة الإحصائية الصفر.

5. المتوسط التوافقي (يرمز له بالرمز H):

يعرف المتوسط التوافقي لمجموعة من البيانات على انه مقلوب المتوسط الحسابي لمقلوبات المشاهدات. أي أن مقلوب المتوسط التوافقي هو المتوسط الحسابي لمقلوبات البيانات. ويحسب رياضيا حسب الحالات التالية:

أ. **بيانات غير مبوبة:** لتكن $x_1.x_2..........x_n$ سلسلة بيانات من n مشاهدة، فإن متوسطها التوافقي هو:

$$\boldsymbol{H}=\frac{\mathbf{n}}{\sum_{\mathbf{i=1}}^{\mathbf{n}}\frac{\mathbf{1}}{\mathbf{x}_{\boldsymbol{i}}}}=\frac{\mathbf{n}}{\frac{\mathbf{1}}{\mathbf{x_1}}+\frac{\mathbf{1}}{\mathbf{x_2}}+\cdots+\frac{\mathbf{1}}{\mathbf{x_n}}}$$

مثال: اوجد المتوسط التوافقي للباتات التالية: 2 .5.3. 8. 10. 12. 14. 16. 20.

$$H=\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}=\frac{9}{\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{8}+\frac{1}{10}+\frac{1}{12}+\frac{1}{14}+\frac{1}{16}+\frac{1}{20}}=5.89$$

ب. **بيانات مبوبة متغيرة متقطعة:** لتكن المتغيرة X. $x_1.x_2..........x_k$ القيم التي تأخذها المتغيرة

و $n_1.n_2.........n_k$ تكرارات هذه القيم لسلسلة بيانات من n مشاهدة، متوسطها التوافقي هو:

$$\boldsymbol{H}=\frac{\mathbf{n}}{\sum_{\mathbf{i=1}}^{\boldsymbol{k}}\frac{\boldsymbol{n_i}}{\mathbf{x}_{\boldsymbol{i}}}}=\frac{\mathbf{n}}{\frac{\boldsymbol{n_1}}{\mathbf{x_1}}+\frac{\boldsymbol{n_2}}{\mathbf{x_2}}+\cdots+\frac{\boldsymbol{n_k}}{\mathbf{x}_{\boldsymbol{k}}}}$$

مثال: حساب المتوسط التوافقي للبيانات التالية:

المجموع	7	5	3	1	x_i
60	10	15	20	15	n_i
26.1	10 /7	15/5	20/3	15	n_i/x_i

$$H=\frac{60}{26.1}=2.3$$

ت. **بيانات مبوبة متغيرة مستمرة:** لتكن المتغيرة X موزعة في جدول توزيع تكراري يحتوي k فئة و

لتكن $c_1.c_2..........c_k$ مراكز هذه الفئات و $n_1..n_2.........n_k$ تكرارات هذه الفئات ولتكن n تمثل إجمالي

التكرارات أي أن: $(\sum_{i=1}^{k} n_i = n)$ متوسطها التوافقي هو

$$\boldsymbol{H}=\frac{\mathbf{n}}{\sum_{\mathbf{i=1}}^{\boldsymbol{K}}\frac{\boldsymbol{n_i}}{\mathbf{c}_{\boldsymbol{i}}}}=\frac{\mathbf{n}}{\frac{\boldsymbol{n_1}}{\boldsymbol{c_1}}+\frac{\boldsymbol{n_2}}{\mathbf{c_2}}+\cdots+\frac{\boldsymbol{n_k}}{\mathbf{c_K}}}$$

مثال: الجدول التالي يمثل توزيع الإنتاج السنوي للتمور (بالقنطار) لمجموعة من البساتين الفلاحية (واحد هكتار لكل بستان) في محيطات حاسي بن عبد الله.

100	10	12	26	30	12	8	2	عدد البساتين n_i
المجموع	40-35	-30	-25	-20	-15	-10	-5	فئات الإنتاج بالقنطار
	37.5	32.5	27.5	22.5	17.5	12.5	7.5	مركز الفئة c_i
4.51	10/37.5	12/32.5	26/27.5	30/22.5	12/17.5	8/12.5	2/7.5	n_i/c_i

$$H=\frac{100}{4.51}=22.18$$

خواص المتوسط التوافقي:

- تستخدم في حسابه جميع البيانات المتاحة.
- صعب الحساب و الفهم و الاستيعاب.
- يتأثر بالقيم السالبة و المعدومة.

6. **الوسيط (يرمز له بالرمز Me):**

يرمز له بالرمز Me ويعرف على انه القيمة الوسطى لقيم مرتبة ترتيبا تصاعديا أو تنازليا، أي عدد القيم الأقل منها يساوي عدد القيم الأكبر منها، أو انه القيمة التي اقل منها 50% من القيم.

a) **الوسيط لبيانات غير مبوبة عدد القيم n فردي**: يعرف الوسيط على انه القيمة ذات الرتبة:

$$\frac{n+1}{2}$$

مثال: إذا كانت لدينا البيانات التالية: 7. 11. 2. 15. 5. 13. 9. إننا نرتبها ترتيبا تصاعديا:

2. 5. 7. 9. 11. 13. 15 و الوسيط هو القيمة التي ترتيبها: $\frac{7+1}{2} = 4$ وبالتالي فإن الوسيط هو:

.Me=9

b) **الوسيط لبيانات غير مبوبة عدد القيم n زوجي:** يعرف الوسيط على انه القيمة التي تتوسط القيمة ذات الرتبة: $\frac{n}{2}+1$ والقيمة ذات الرتبة: $\frac{n}{2}$.

مثال: إذا كانت لدينا البيانات التالية: 7. 11. 2. 15. 5. 13. 9. 17 إننا نرتبها ترتيبا تصاعديا

2. 5. 7. 9. 11. 13. 15. 17. و الوسيط هوا لقيمة التي توسط القيم (9 و11) و بالتالي فإن الوسيط:

.Me = 10

c) **الوسيط لبيانات مبوبة متغيرة متقطعة:** إذا كانت $x_1.x_2..........x_k$ قيم المتغير المتقطع X و $n_1.n_2..........n_k$ تكراراتها لإيجاد الوسيط نقوم بالخطوات التالية:

1) إيجاد التكرار التجمع الصاعد.

2) تحديد رتبة الوسيط $\frac{n}{2}$.

3) استخراج الوسيط مباشرة من الجدول باعتباره القيمة المقابلة للتكرار المتجمع الصاعد الذي يحوي رتبة الوسيط .

مثال: حساب الوسيط للبيانات التالية

x_i	n_i	F_i
1	15	15
3	20	35
5	15	50
7	10	60
المجموع	60	

رتبة الوسيط هي: 30

فئة الوسيط هي: الفئة الثانية

قيمة الوسيط هي: 3

d) الوسيط لبيانات مبوبة متغيرة مستمرة: لتكن المتغيرة X موزعة في جدول توزيع تكراري يحتوي k فئة و لتكن $c_1.c_2..........c_k$ مراكز هذه الفئات و $n_1..n_2.........n_k$ تكرارات هذه الفئات ولتكن n تمثل إجمالي التكرارات لإيجاد الوسيط نتبع الخطوات التالية:

1) إيجاد التكرار المتجمع الصاعد.

2) تحديد رتبة الوسيط $\frac{}{2}$.

3) تحديد فئة الوسيط وهي الفئة التي تكرارها المتجمع الصاعد يحتوي على رتبة الوسيط.

4) حساب الوسيط وفق العلاقة

$$\mathbf{Me} = \mathbf{Min} + \frac{\frac{\boldsymbol{n}}{\mathbf{2}} - \boldsymbol{F}_{-\mathbf{1}}}{\boldsymbol{F}_{\boldsymbol{Me}} - \boldsymbol{F}_{-\mathbf{1}}} . \boldsymbol{L}$$

Min: الحد الأدنى لفئة الوسيط.

F_{-1}: التكرار المتجمع الصاعد للفئة السابقة.

ملاحظة: لهذه المتوسطات الأربع صفة مشتركة تتمثل في استخدام جميع البيانات ووجد أن:

$$\boldsymbol{Q} \geq \boldsymbol{X} \geq \boldsymbol{H} \geq \boldsymbol{G}$$

علما أن المساواة تحقق في حالة واحدة فقط إذا كانت جميع القيم متساوية.

ii. مقاييس التشتت:

مقاييس التشتت هي مقاييس عددية تستخدم لقياس درجة تجانس (تقارب) أو تشتت(تباعد)مفردات البيانات عن بعضها البعض. ومقاييس التشتت تستخدم لوصف مجموعة البيانات، وكذلك لمقارنة مجموعات البيانات المختلفة، إذ أن مقاييس النزعة المركزية لا تكفي وحدها لوصف مجموعة البيانات أو مقارنة مجموعات البيانات المختلفة.

ومن أشهر مقاييس التشتت نذكر:

1. المدى:

يعتبر المدى من أسهل مقاييس التشتت تعريفا وحسابا، حيث أنه يعطينا فكرة سريعة عن مدى تفرق البيانات ويرمز له بالرمز (R) . ويعرف المدى لمجموعة من البيانات بالصيغ التالية:

أ. **حالة البيانات غير المبوبة:**

• المدى (R) = أكبر قيمة – أصغر قيمة

مثال: المدى للبيانات التالية : 54، 89 ، 65 ، 70 ،95 ، 47

الحل: المدى: 95-47=84

ب. **حالة البيانات المبوبة:**

المدى يعرف بأكثر من طريقة، نذكر منها الطريقتين الآتيتين:

• المدى (R)= مركز الفئة العليا - مركز الفئة الدنيا.

• المدى (R)= الحد الأعلى للفئة العليا - الحد الأدنى للفئة الدنيا.

2. الانحراف المعياري:

يعتبر الانحراف المعياري أهم مقاييس التشتت و أكثرها استخداما، يعرف على انه الجذر التربيعي لمتوسط مربعات انحرافات القيم عن متوسطها الحسابي و يحسب حسب حالة البيانات المدروسة.

أ. **بيانات غير مبوبة:**

يحسب الانحراف المعياري ل n مشاهدة $x_1.x_2..........x_n$ بالعلاقة:

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

ب. **بيانات مبوبة متغيرة متقطعة:**

إذا كانت $x_1.x_2..........x_k$ قيم المتغير المتقطع X و $n_1.n_2..........n_k$ تكراراتها يعرف الانحراف المعياري بالعلاقة:

$$\sigma = \sqrt{\frac{\sum n_i (x_i - \bar{x})^2}{n}}$$

ت. **بيانات مبوبة متغيرة مستمرة :**

إذا كانت $x_1.x_2..........x_k$ قيم المتغير المتقطع X و $n_1.n_2..........n_k$ تكراراتها يعرف الانحراف المعياري بالعلاقة:

$$\sigma = \sqrt{\frac{\sum_{j=1}^{k} n_j (x_j - \bar{x})^2}{n}}$$

حيث Xj مركز الفئة j .

3. التباين:

يعرف على انه متوسط مربعات انحرافات المشاهدات عن وسطها الحسابي فهو مربع الانحراف المعياري أي أن:

$$V = \sigma^2$$

4. انحراف المتوسط (يرمز له بالرمز e):

يعرف الانحراف المتوسط على انه متوسط انحرافات القيم عن متوسطها الحسابي مأخوذة بالقيم المطلقة ويحسب حسب طبيعة البيانات

ث. **بيانات غبر مبوبة:** يحسب الانحراف المتوسط ل n مشاهدة $x_1.x_2..........x_n$ بالعلاقة:

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

ج. **بيانات مبوبة متغيرة متقطعة:** إذا كانت $x_1.x_2..........x_k$ قيم المتغير المتقطع X و $n_1.n_2..........n_k$ تكراراتها يعرف الانحراف المتوسط بالعلاقة:

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

ح. **بيانات مبوبة متغيرة مستمرة:** لتكن المتغيرة X موزعة في جدول توزيع تكراري يحتوي k فئة و لتكن $c_1.c_2..........c_k$ مراكز هذه الفئات و $n_1.n_2........n_k$ تكرارات هذه الفئات ولتكن n تمثل إجمالي التكرارات أي أن: $(\sum_{i=1}^{k} n_i = n)$ يعرف الانحراف المتوسط بالعلاقة:

$$\mathbf{e} = \frac{\sum_{i=1}^{k} n_i |c_i - \bar{x}|}{n}$$

حيث $c\,j$ مركز الفئة j .

خواص الانحراف المتوسط:

- سهل الفهم والحساب والتطبيق.
- يقيس مدى تباعد المشاهدات عن وسطها الحسابي.
- يعتمد في حسابه على جميع المشاهدات