علم الإحصاء:-

هو ذلك العلم الذي يعمل على استخدام الأسلوب العلمي في طرق جمع البيانات وعرضها وتحليلها للوصول منها الى استنتاجات وقرارات مناسبة.

لقد كان علم الإحصاء في بدايته يهتم فقط بعملية العد والحصر للأشياء ، ومن هنا جاءت تسميته العربية " إحصاء" فهي مشتقة من كلمة أحصى ، وتعني استخدام الحصى أو الحجارة الصغيرة كوسيلة بدائية لعد الأشياء الكثيرة ، فقد كان الإنسان قديماً يستعين بالحصى في عملية العد .

وكان علم الإحصاء مقصوراً على تعدادات السكان وثرواتهم ، وعدد المواليد والوفيات لمعرفة القوى البشرية المتوفرة في الدولة ، وذلك للأهتداء بها في تصريف أمور الدولة ورسم سياستها ، وحيث إن الإحصاء كان مقصوراً على الحقائق الخاصة بالدولة فمن هنا جاءت التسمية باللغة الأجنبية

" **Statistics**" فهي مشتقة من كلمة " **State** " أي الدولة. وهناك تعريفات عدة للإحصاء تتراوح بين ما كان مألوفاً وشائعاً في الماضي إلى ما هو حديث وجامع .

فقديماً عرف الإحصاء على أنه مجرد جمع المعلومات وترتيبها في جداول أو إبرازها في رسوم بيانية . وقد أخذ هذا المعنى للإحصاء يتلاشى من الأذهان فاسح المجال للمعنى الحديث الذي يعرف علم الإحصاء بأنه العلم الذي يبحث في جمع البيانات ، وتنظيمها ، وعرضها ، وتحليلها ، واستقراء النتائج ، واتخاذ القرارات بناء عليها.

ونعرض فيما يلي علاقة لإحصاء ببعض العلوم :

*ففي علم الاقتصاد ، استخدم علم الإحصاء لتفسير الظواهر الاقتصادية المختلفة كنظريات الطلب والعرض ، والعلاقة بين مستويات الدخل والإنفاق الاستهلاكي ، ونوع العلاقات الاقتصادية المختلفة وكيفية قياسها .. الخ .

وفي مراقبة الإنتاج في الشركات الصناعية من حيث كمية ودرجة وجودة ومدى ملاءمة كل ذلك لاحتياجات السوق وأذواق المستهلكين . وغيرها من الدراسات الاقتصادية. ولايمكن للدراسات الاقتصادية أن تتطور دون استخدام النظريات والطرق الإحصائية . فالإحصاء يعتبر أداة مهمة في تلك الدراسات . ويمكن القول بأن المجال الاقتصادي هو الإطار الذي نشأ وتطور في كنفه علم الإحصاء .

*وفي علم النفس استخدمت الطرق الإحصائية في قياس درجة ذكاء الأشخاص،وفي دراسة العلاقة بين ذكاء الأشخاص ومهارتهم ..الخ.

*وفي علم الفلك استخدمت الطرق الإحصائية في دراسات خاصة بتحديد مدارات الكواكب والنجوم وغيرها من الأبراج السماوية ..الخ .

*وفي علم الجغرافيا (بشقيها الطبيعي والبشري) استخدمت الأساليب الإحصائية في دراسة أشكال سطح الأرض ، والجغرافية المناخية وجغرافية البحار والمحيطات . فضلا عن تطبيق الطرق الإحصائية في جغرافية المدن وعلم الخرائط .. الخ.

*وفي العلوم الطبية ، يطبق علم الإحصاء في أغلبية الدراسات الطبية لمقارنة الأمراض المختلفة وسبل علاجها وتحديد العلاقة بين بعض الأمراض ومسبباتها، ولقياس كفاءة الأدوية .. الخ.

وفي الواقع لا يخلو علم من العلوم في عصرنا هذا من استخدام علم الإحصاء .

أنواع علم الإحصاء

1- الإحصاء الوصفي:-

الإحصاء الوصفي هو عبارة عن وصف مختصر لمجموعة كبيرة من البيانات أو مجموعة طرق تستخدم لتسهيل وصف الخصائص الرئيسة للبيانات كميًا وذلك باستخدام الجداول والمخططات البيانية ليسهّل فهمها لدى مستخدمها، فعلى سبيل المثال، يقدّم المتوسّط الحسابي معلومات عن الأداء للطالب ويعكس محتوى البيانات الكبيرة والتي تحتوي على العلامات الخاصّة بكل مادة لتصبح مفهومة لدى قارئها دون الحاجة إلى معرفة مجموعة البيانات الكبيرة.
ينقسم الإحصاء الوصفي إلى قسمين رئيسيين، قسم يهتم بمقاييس النزعة المركزيّة والمتمثّلة في قيم المتوسّط الحسابي، الوسيط والمنوال، وقسمُ آخر يعنى بمقاييس التشتّت مثل الانحراف المعياري، التباين والقيم القصوى، والتي تدلّ على كيفية توزيع البيانات عن طريق تحليلها ورسم منحنى يضم جميع هذه البيانات

الإحصاء الوصفي هو تلخيص لمجموعة البيانات الكبيرة بحيث يسهّل على المستخدم فهم البيانات دون الحاجة إلى الاطّلاع عليها بالتفصيل.

2- الإحصاء الاستنتاجي أو الاستدلالي:-

يعرف الإحصاء الاستدلالي أو الاستنتاجي Inferential analysis بأنه النوع الثاني من أنواع علم الإحصاء، والذي يعنى بدراسة العلاقات بين المتغيرات المتعلقة بالدراسة الإحصائية، حيث يركز على الاستنتاجات الناتجة من الحسابات الرياضية الناتجة عن الإحصاء الوصفي (النوع الأول) ويعمل على تحليلها بقصد إطلاق التنبؤات والتعميمات وفقًا لما جاء فيها، كما يتم استخدامه كوسيلة للحكم على بعض البيانات غير المرئية، حيث يتم تحليلها واستخلاص النتائج منها

تساهم عمليات الإحصاء الاستدلالي في شرح وتفسير الظواهر المختلفة والتنبؤ بالأحداث المتعلقة بها، مما يسهل علينا فهم الواقع الحالي، ومن الأمثلة على المجالات التي يدخل فيها الإحصاء الاستدلالي؛ الهندسة، والتعليم، والخدمات المصرفية، وتطبيقات الذكاء الصناعي، فضلًا عن أهميته في مجال الأعمال، فعلى سبيل المثال يمكن الاستفادة من الإحصاء الاستدلالي باختيار أسواق جديدة والتنبؤ بالمبيعات المستقبلية عن طريق دراسةعينة عشوائية صغيرة

طبيعة البيانات الإحصائية: -

عند جمع بيانات حول ظاهرة ما فإننا نرمز للظاهرة بالرمز(x) وكل مفردة أو مشاهدة منها نرمز لها بالرمز (xi).

المتغير: -

هو أي ظاهرة تظهر اختلافات بين مفرداتها. ويرمز له بالرمز x أو أي رمز أخر مثل y او z-----الخ.

وينقسم الى: -

1- متغيرات وصفية او نوعية: -

وهي تلك الظواهر او الصفات التي لايمكن قياسها بالأرقام العددية. مثل صفة لون العيون(اسود،ازرق، بني)----الخ.

2- متغيرات كمية:-

وهي تلك الظواهر او الصفات التي يمكن قياسها بأرقام عددية. مثل صفة العمر، او الوزن،-----الخ.

المجتمع:-

المجتمع الإحصائي هو عبارة عن جميع المفردات موضع الدراسة والتي نرغب في معرفة حقائق عنها سواء كانت على شكل إنسان أو حيوان أو جماد أو درجات امتحان أو منازل أو مزارع أو سفن ... الخ. وقد يتكون المجتمع من عدد محدود Finite من المفردات مثل عدد أفراد مدينة ما أو عدد المنازل بهذه المدينة ... الخ، أو يتكون المجتمع من عدد غير محدود Infinite مثل الأسماك في الخليج العربي أو عدد النجوم أو عدد حبات القمح في مزارعة عينة وإذا تم جمع البيانات لجميع مفردات المجتمع فتسمى هذه العملية بالحصر الشامل. وفى بعض الحالات لا نتمكن من حصر كل المفردات للمجتمع مثل مجتمعات الأسماك أو النباتات، أو أن تؤدي عملية الحصول على البيانات لمفردات المجتمع إلى إتلافها أو هلاكها، مثال ذلك فحص دم المريض كله يؤدي إلى وفاة الشخص، وكذلك فحص جميع أعواد الثقاب يؤدي إلى إتلاف هذا المنتج بالكامل... الخ وبالتالي لا يمكن جمع البيانات من كل المفردات، أو قد تحتاج عملية جمع البيانات إلى وقت طويل أو جهد كبير أو تكاليف باهظة. وفى مثل هذا الحالات السابقة يتم جمع من البيانات عن جزء فقط من مفردات المجتمع يسمى العينة . sample

العينة:-

عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع . وعلية فالعينة جزء من المجتمع.

الرموز الإحصائية:-

يرمز عادة لمجموع قيم المتغير بالرمز

$$\sum_{i=1}^{n} xi$$

فالرمز $\sum$ يسمى (Summation of) أي المجموع.

و(n,1) هما حدا المجموع

وعليه فالرمز $\sum_{i=1}^{n} xi$ يقرا مجموع القيم x مبتدأ من المشاهدة الأولى وحتى الأخيرة إي أن

$$\sum_{i=1}^{n} xi = x_1 + x_2 + --- + x_n$$

*هناك مجموع جزئي مثل

$$\sum_{i=3}^{5} xi = x_3 + x_4 + x_5$$

تصنيف وتبويب البيانات _

ان البيانات المستحصل عليها بخصوص الظاهرة المعينة تسمى بالبيانات الخام او الاولية وتلك البيانات تكون غير منتظمة ويصعب على الباحث تكوين فكرة عن الظاهرة المدروسة ،كذلك يصعب الاعتماد عليها لاغراض التحليل الاحصائي لذا فان الخطوات الهامة التي تلي عملية جمع البيانات هي عملية مراجعة وتصنيف البيانات وتبويبها.

مراجعة البيانات

بعد اتمام عملية جمع البيانات يجب مراجعة تلك البيانات وترقيمها لغرض التاكد من مطابقتها وتكاملها لمتطلبات الدراسة.

تبويب البيانات

بعد اتمام تصنيف البيانات نبدا بعملية تبويب البيانات ويقصد بتبويب البيانات ،تفريغ البيانات المصنفة في جداول خاصة بحيث ان كل جزء من البيانات المصنفة عن الظاهرة المعينة يعود الى مستوى معين لتلك الظاهرة ،والهدف من عملية التبويب هو ابراز البيانات وتوضيحها في اضيق حيز ممكن كي نتمكن من تكوين فكرة عنها .وهناك عددا من اشكال التبويب اهمها الاتي :

1- التبويب الزمني

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها يعود الى وحدة زمنية معينة كاليوم او الاسبوع او الشهر او السنة.

2- التبويب الجغرافي.

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة جغرافية معينة او تقسيم اداري معين.

3- التبويب الكمي

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة كمية معينة لوحدات الوزن ،الطول ،العرض ،المسافة.

4- التبويب على اساس صفة معينة.

عبارة عن عملية تجميع البينات وترتيبها في جداول خاصة على اساس ان كل جمع منها يشترك بصفة معينة كالجنس ،الحالة الاجتماعية ،عنوان الوظيفة ،القومية.

الأخطاء الشائعة في جمع البيانات

هناك بعض الأخطاء التي تحدث عندما يقوم الباحث بجمع البيانات والمعلومات التي تخص بحثه :

1- خطأ التحيز :هنا هذا الخطأ يرتكبه المصدر أو المفردة الإحصائية التي تزود الباحث بالمعلومات سواء بقصد أو بغير قصد أو يحدث هذا الخطأ أحيانا عندما يستقي الباحث معلومات بحثه ليست من مصادرها الأصلية بل من مصادرها الغير مباشرة

2- خطأ الصدفة: هذا الخطأ يرتكبه الباحث بنفسه سواء بتعمد أو بصورة غير متعمدة حيث يستقي معلومات بحثه بالاعتماد على ذاكرته بسبب بعد المفردة الإحصائية عنه أو لأي سبب شخصي أخر .هذا سيؤدي إلى الحصول على نتائج واستنتاجات غير دقيقة وبعيدة عن الواقع **.**

مقاييس النزعة المركزية :-

مقاييس النزعة المركزية والتي يطلق عليها غالبًا مصطلح المتوسطات بأنها مجموعة القيم المركزية أو النموذجية المتخصصة في توزيع الاحتمالات، ويطلق عليها في بعض الأحيان مراكز التوزيع، ومن أهم مقاييس النزعة المركزية الأكثر شيوعًا المقاييس الوسط الحسابي والمتوسط، والمنوال وسوف يتم شرح كل واحدة منها .

مقاييس النزعة المركزية: Central Tendency

في كثير من النواحي التطبيقية يكون الباحث في حاجة الى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم ، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير، وايضاً ما اذا كان هناك قيم شاذة او لا .
والاعتماد على العرض البياني وحده لا يكفي ، لذا يتناول هذا الفصل والذي يليه عرض بعض المقاييس الاحصائية والتي يمكن من خلالها التعرف على خصائص الظاهرة محل البحث وكذلك امكانية مقارنة ظاهرتين او اكثر، ومن اهم هذه المقاييس مقاييس النزعة المركزية والتشتت .
تسمى مقاييس النزعة المركزية بمقاييس الموضع او المتوسطات ، وهي القيم التي تتركز القيم حولها ، ومن هذه المقاييس ؛ الوسط الحسابي ، المنوال ، الوسيط ، الوسط الهندسي ، والوسط التوافقي ، وفيما يلي عرض لأهم هذه المقاييس .

*** الوسط الحسابي Arithmetic mean :** من أهم مقاييس النزعة المركزية ، وأكثرها استخداما في النواحي التطبيقية ، ويمكن حسابه للبيانات المبوبة وغير المبوبة ، كما يلي :

أولا: الوسط الحسابي للبيانات غير المبوبة: يعرف الوسط الحسابي بشكل عام على أنه مجموع القيم مقسوما على عددها. فإذا كان لدينا n من القيم ، ويرمز لها بالرمز $x_1, x_2, \ldots, x_n$ فإن الوسط الحسابي لهذه القيم ، ونرمز له بالرمز $\bar{x}$ يحسب بالمعادلة التالية :

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

حيث يدل الرمز Σ على المجموع .
مثال(3-1) فيما يلي درجات 8 طلاب في مادة الإحصاء 40، 36، 40، 35، 37، 42، 32، 34 .
والمطلوب إيجاد الوسط الحسابي لدرجة الطالب في الامتحان .
الحل : لإيجاد الوسط الحسابي للدرجات تطبق المعادلة السابقة كما يلي:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$= \frac{34+32+42+37+35+40+36+40}{8} = \frac{296}{8} = 37$$

أي أن الوسط الحسابي لدرجة الطالب في اختبار الإحصاء يساوي 37 درجة.
ثانيا: الوسط الحسابي للبيانات المبوبة: من المعلوم أن القيم الأصلية، لا يمكن معرفتها من جدول التوزيع التكراري، حيث أن هذه القيم موضوعة في شكل فئات، ولذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز هذه الفئة، ومن ثم يؤخذ في الاعتبار أن مركز الفئة هو القيمة التقديرية لكل مفردة تقع في هذه الفئة.

فإذا كانت k هي عدد الفئات ، وكانت $x_1, x_2, \ldots, x_k$ هي مراكز هذه الفئات، $f_1, f_2, \ldots, f_k$ هي التكرارات ، فإن الوسط الحسابي يحسب بالمعادلة التالية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_k f_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i}$$

مثال الجدول التالي يعرض توزيع 40 تلميذ حسب أوزانهم.

فئات الوزن	32-34	34-36	36-38	38-40	40-42	42-44
عدد التلاميذ	4	7	13	10	5	1

والمطلوب إيجاد الوسط الحسابي.
الحل: لحساب الوسط الحسابي باستخدام المعادلة السابقة يتم إتباع الخطوات التالية :
1- إيجاد مجموع التكرارات $\sum f$
2- حساب مراكز الفئات x
3- ضرب مركز الفئة في التكرار المناظر له $\sum f$ وحساب المجموع $\sum xf$
4- حساب الوسط الحسابي بتطبيق المعادلة.

فئات الوزن (c)	التكرارات f	مراكز الفئات x	x f
32 - 34	4	33 =2÷(34+33)	123=4×33
34 – 36	7	35 =2÷(34+35)	245=7×35
36 – 38	13	37	481=13×37
38 - 40	10	39	390=10×39
40 – 42	5	41	205=5×41
42 – 44	1	43	43=1×43
المجموع	40		1496

إذا الوسط الحسابي لوزن التلميذ هو :

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i f_i}{\sum_{i=1}^{6} f_i} = \frac{1496}{40} = 37.4 \ k.g$$

أي أن متوسط وزن التلميذ يساوي kg37.4

خصائص الوسط الحسابي: يتصف الوسط الحسابي بعدد من الخصائص ، ومن هذه الخصائص ما يلي :

1- الوسط الحسابي للمقدار الثابت يساوى الثابت نفسه ، أي أنه إذا كانت قيم x هي

$x : a , a ,...., a$ ، فإن الوسط الحسابي هو:

$$\boxed{\bar{x} = \frac{a + a + ... + a}{n} = \frac{na}{n} = a} \quad (٣-٣)$$

ومثال على ذلك ، لو اخترنا مجموعة من 5 طلاب ، ووجدنا أن كل طالب وزنه 63 كيلوجرام فإن متوسط وزن الطالب في هذه المجموعة هو :

$$\bar{x} = \frac{63+63+63+63+63}{5} = \frac{315}{5} = 63 \ k.g$$

2-مجموع انحرافات القيم عن وسطها الحسابي يساوى صفرا ، ويعبر عن هذه الخاصية بالمعادلة:

ويمكن التحقق من هذه الخاصية باستخدام بيانات المثال السابق : نجد أن درجات الطلاب هي :

34, 32, 42, 37, 35, 40, 36, 40، والوسط الحسابي للدرجة هو $\bar{x} = 37$ ، إذا :

x	34	32	42	37	35	40	36	40	296
$(x-\bar{x})$	34-37	32-37	42-37	37-37	35-37	40-37	36-37	40-37	0
$(x-37)$	-3	-5	5	0	-2	3	-1	3	

أي أن : $\sum (x-37) = 0$

مزايا وعيوب الوسط الحسابي:

يتميز الوسط الحسابي بالمزايا التالية:

- أنه سهل الحساب. .
- يأخذ في الاعتبار كل القيم . .
- أنه أكثر المقاييس استخداما وفهما.

ومن عيوبه:

- أنه يتأثر بالقيم الشاذة والمتطرفة.
- يصعب حسابه في حالة البيانات الوصفية .
- يصعب حسابه في حالة الجداول التكرارية المفتوحة.

* الوسيط median:

الوسيط هو قيمة المفردة التي يسبقها عدد من المفردات يساوي عدد المفردات التي تعقبها، بعد ترتيبها تصاعديا أو تنازليا.

أي أن الوسيط هو النقطة التي تقسم التوزيع إلى قسمين بحيث يكون عدد الدرجات التي أعلى هذه النقطة يساوي عدد الدرجات التي تقع أسفل النقطة.

أ- حساب الوسيط فى حالة البيانات الغير مبوبة :

ويعتمد حساب الوسيط على ما إذا كان عدد البيانات فرديا أم زوجيا.

وفيما يلي طريقة حساب الوسيط:

. إذا كان عدد البيانات فردياً :

فهنا يكون الوسيط هو الدرجة الوسطى

مثال: 2 ، 5 ، 6 ، 7 ، 10

تعتبر الدرجة 6 تقسم التوزيع الى نصفين، نظرا لأن الدرجتين 3 ، 5 أقل من 6 ، والدرجتين 7 ، 10 أكبر من 6
اما إذا كان عدد الدرجات زوجياً.
فهنا يكون الوسيط مساويا لمتوسط الدرجتين اللتين تقعان في وسط التوزيع. فإذا كان لدينا الدرجات (3 ، 5 ، 6 ، 7 ، 10 ، 11) فإن الدرجة التي تقسم هذا التوزيع الى نصفين تقع بين 6 ، 7 وهنا يكون الوسيط مساويا لـ 6,5

ب - حساب الوسيط في حالة البيانات المبوبة :

قانون الوسيط بالنسبة للبيانات المبوبة هو :

$$\mathrm{Me} = A + \{\frac{\frac{n}{2} - N1}{n1}\} * \mathrm{L}$$

n مجموع التكرارات
N1 التكرار المتجمع للفئة التي تسبق الفئة الوسيطية
n1 تكرار الفئة الوسيطية
A الحد الادنى للفئة الوسيطية
L طول الفئة

يتم استخراج رتبة الفئة الوسيطية من قسمة مجموع التكرارات على 2

مثال : احسب الوسيط واشرح النتيجة في حالة كان توزيع الاجور اليومية للعمال على النحو التالي :

الاجور اليومية (الفئات)	التكرارات ni	التكرار المتجمع الصاعد Ni
550 - 400	17	17
700 – 550	13	30
850 – 700	10	40
1000 – 850	5	45
1150 – 1000	2	47
1300 – 1150	2	49
1450 – 1300	1	50

الحل :

تحديد الفئة الوسيطية : وهي اول فئة تكرار متجمع الصاعد اكبر او يساوي n/2 وبما ان n مجموع التكرارات

اذا $\frac{n}{2} = \frac{50}{2} = 25$

$$\mathrm{Me} = A + \{\frac{\frac{n}{2} - N1}{n1}\} * \mathrm{L}$$

$$\mathrm{Me} = 550 + \left\{\frac{25 - 17}{13}\right\} * 150 = 642{,}30$$

الشرح : هناك 50 % من العمال اجورهم اليومية اقل من 642,30 د ج و 50 % من العمال اجورهم اليومية اكبر من 642,30 د ج .
الوسيط بيانيا هو نقطة التقاطع بين المنحني المتجمع الصاعد والمنحني المتجمع النازل .

- خصائص الوسيط :

1- سهولة معناه رغم عدم شيوعه
2- قيمة الوسيط محددة حيث يسبقه 50% من قيم درجات الأفراد ويتلوه 50% من قيم درجات الأفراد ولذلك فهو متوسطاً مكانياً وليس متوسطاً حسابياً
3- قيمة الوسيط لا تتأثر بالقيم المتطرفة مثل المتوسط الحسابي ولكنه يتأثر بالقيم القريبة منه ويتأثر أيضا بعدد الأفراد

فمثلاً: الدرجات 2 4 5 8 9 13 15
متوسطها الحسابي = 8، ووسيطها= 8
بينما نفس الدرجات عندما نستبدل الدرجة 15 ب 50 نجد أن متوسطها الحسابي أصبح 13 بينما وسيطها ظل كما هو= 8

* المنوال Mode

حساب المنوال

أ-حساب المنوال في حالة البيانات الغير مبوبة :

حدد المنوال للقيم التالية: 1، 2، 3، 4، 5 لا يوجد منوال

ب-حساب المنوال في حالة توزيع تكراري

لا يستدعي تحديد المنوال في هذه الحالة أي عمليات حسابية، بحيث يتم تحديد المفردة أو العنصر أو القيمة التي حصلت أكثر تكرار

مثال: حدد المنوال للبيانات التالية:
ذكر، أنثى، أنثى، أنثى، ذكر
المنوال في هذه الحالة هو: أنثى، لأنها تكررت ثلاث مرات في حين تكررت ذكر مرتين فقط.
جد المنوال للقيم التالية : 1، 3، 7، 11، 5، 1، 3، 1 المنوال 3 فقط لأنه اكثر تكرار

<u>حساب المنوال في حالة البيانات المبوبة :</u>

قانون المنوال بالنسبة للبيانات المبوبة هو :

$$\mathrm{Mo} = A + \{\frac{d1}{d1+d2}\} * \mathrm{L}$$

A هو الحد الادنى للفئة المنوالية
d1 تكرار الفئة المنوالية – تكرار الفئة السابقة
d2 تكرار الفئة المنوالية – تكرار الفئة اللاحقة
L طول الفئة المنوالية

<u>مثال :</u>

الفئات	التكرارات
2 -	4
5 -	7
8 -	10
11 -	5
14 - 17	4

فئة المنوال

$A = 8$

$d_1 = 10 - 7 = 3$

أكبر تكرار

$d_2 = 10 - 5 = 5$

- حساب الفروق d ، حيث أن :

$$d_1 = (10-7) = 3 \quad d_2 = (10-5) = 5$$

- تحديد الحد الأدنى للفئة المنوالية $(A = 8)$ ، وكذلك طول الفئة $(L = 3)$
- وبتطبيق المعادلة الخاصة بحساب المنوال في حالة البيانات المبوبة . نجد أن :

$$Mod = A + \frac{d_1}{d_1 + d_2} \times L$$

$$= 8 + \frac{3}{3+5} \times 3 = 8 + 1.125 = 9.125$$

<u>خصائص المنوال</u>

إن أهمية المنوال تتمثل فيما إذا كان الهدف معرفة القيمة التي يتفق فيها أغلب أفراد المجموعة، إن هذا المقياس المركزي يمكن الحصول عليه في أقصر وقت ممكن، إلا أنه لا يهتم كثيرا بالدقة.

ج-حساب المنوال في حالة بيانات مبوبة في فئات من خلال القانون التالي:

البيانات غير المبوبة

معامل الارتباط البسيط (معامل بيرسون) :

ويرمز له بالرمز rp وهو معامل ارتباط خطي بسيط يقيس قوة واتجاه العلاقة بين متغيرين فقط ، وهذان المتغيران هما متغيران كميان أي يعبر عنهما بالأرقام ، ويحسب المعامل وفق القانون الآتي:

$$rp = \frac{n\Sigma xY - \Sigma x\, \Sigma Y}{\sqrt{[n\Sigma x2 - (\Sigma x)2]\, [n\Sigma Y2 - (\Sigma Y)2]}}$$

مثال: جد قوة واتجاه العلاقة بين سعر السلعة x والكمية المطلوبة Y باستخدام مقابل ارتباط بيرسون :

X	Y	XY	X^2	Y^2
1	6	6	1	36
2	6	12	4	36
3	4	12	9	16
4	2	8	16	4
5	1	5	25	1
Σ=15	19	43	55	93

$$rp = \frac{n\Sigma xY - \Sigma x\, \Sigma Y}{\sqrt{[n\Sigma x2 - (\Sigma x)2]\, [n\Sigma Y2 - (\Sigma Y)2]}}$$

$$rp = \frac{5\,(43) - 15\,(19)}{\sqrt{[5\,(55) - (15)2][5\,(93) - (19)2]}}$$

$$rp = \frac{215-285}{\sqrt{[275-225][465-361]}} = \frac{-70}{\sqrt{5200}} = \frac{-70}{72} = -0{,}97$$

من قيمة معامل الارتباط يتضح ما يلي :

1 - هناك ارتباط قوي بين سعر السلعة والكمية المطلوبة كون أن قيمة معامل الارتباط يساوي 0,97 - وهي قريبة جداً من 1-

2 - العلاقة بين السعر والكمية المطلوبة هي علاقة عكسية بدليل الإشارة السالبة لمعامل الارتباط.

3 - هذه العلاقة تتفق مع المنطق (كلما ارتفع سعر السلعة انخفض الطلب عليها)

ملاحظات :

1 - يكون الارتباط محصور بين الواحد والسالب واحد $[-1 < (rp) < 1]$.

2 - تختلف درجة قوة معامل الارتباط باختلاف العلوم ، فعلى سبيل المثال في العلوم الطبية لا يمكن اعتماد معامل ارتباط أقل من 90 %، أما في العلوم الزراعية فيمكن اعتماد 60 % إلا أنه بصورة عامة إذا زادت قيمة r عن 70 % يعتبر هناك ارتباط قوي بين المتغيرين.

معامل ارتباط الرتب (سبيرمان) :

وهو معامل ارتباط ثنائي يصلح في المتغيرات الكمية والنوعية وهو أقل من معامل ارتباط بيرسون ، ويمكن إيجاد معامل ارتباط سبيرمان حسب الخطوات الآتية:

1- إعطاء رموز رقمية للبيانات النوعية (الرتب) لكل من X - وY

2 - نستخرج الفرق بين رتب X ورتب Y بعمود جديد هوd

3 - نقوم بتربيع الفرق بعمود آخر d^2 -

4 - تطبيق قانون سبيرمان لإيجاد معامل الارتباط.

$$rs = 1 - \frac{6\Sigma d2}{n(n2-1)}$$

مثال :

كانت تقديرات مجموعة من الطلبة في مادتي التحليلات X والإحصاء Y كما يأتي :

المطلوب : إيجاد معامل الارتباط بينهما والتعليق عليه.

تحليلاتX	Y	رتبX	رتبY	d	d^2
امتياز	جيد جدا	1	2	1-	1
جيد جدا	امتياز	2	1	1	1
جيد	متوسط	3,5	4,5	1-	1
جيد	جيد	3,5	3	0,5	0,25
متوسط	متوسط	5	4	0,5	0,25
					3,5

$$rs = 1 - \frac{6\Sigma d2}{n(n2-1\,)} \quad = 1 - \frac{6(3,5)}{5(25-1\,)}$$

$$= 1 - \frac{21}{5(24)} \quad = 1 - \frac{21}{120} \quad = 1 - 0,175$$

$$rs \;=\; 0,82$$

1- ان قيمة rs هي 0,82 وهذا يعني أن هناك علاقة ارتباط قوية بين تحصيل الطالب في التحليلات وتحصيله في الاحصاء .

2 - أن هذه العلاقة هي علاقة طردية أي أن الطالب المتفوق في التحليلات ويكون متفوق في الإحصاء.

3 - تتفق هذه العلاقة من النظريات التربوية.

مقاييس التشتت للبيانات Measures of variation

تعتبر مقاييس التشتت للبيانات ذات أهمية بالغة في وصف البيانات حيث أن مقاييس النزعة المركزية المتضمنة الوسط الحسابي ، والوسيط ، والمنوال لا تعطينا الصورة الكاملة والحقيقية في توزيع البيانات ، فقد يكون لدينا مجموعتين من البيانات لديها نفس قيمة الوسط الحسابي ولكنها مختلفة تماما من حيث التشتت والانتشار أو مدى التقارب والتباعد للبيانات من مقاييس النزعة المركزية الخاصة بها

أولا / المدى للبيانات (غير المبوبة (التي ليس لها جدول توزيع تكراري)

يعتبر من أسهل مقاييس التشتت للبيانات ويمكن تعريفه على أنه الفرق بين أعلى وأصغر قيمة في البيانات ، فإذا كان المدى الخاص بالبيانات صغير يدل على تجانسها وتقاربها من بعضها لبعض وإذا كان عكس ذلك فهو يدل على تشتتها وتباعدها عن بعضها لبعض

R رمز المدى وهو = أعلى قيمة ــ أدنى قيمة

مثال :

لدينا مجموعتين من الطلاب ، وكانت درجات المجموعة الأولى والثانية على النحو التالي

الجموعة الأولى	40	45	55	70	75	80	85	88	90	100
المجموعة الثانية	65	68	70	72	73	74	75	76	77	78

المدى للمجموعة الأولى = 100 ــ 40 = 60

المدى للمجموعة الثانية = 78 ــ 65 = 13

نلاحظ في هذا المثال بأن المدى في المجموعة الثانية أكثر تجانسا وتقاربا حيث يساوي 13 بينما في المجموعة الأولى أكثر تشتتا وتباعدا حيث يساوي 60 وهو قرابة خمسة أضعاف مدى المجموعة الثانية

ثانيا / المدى للبيانات المبوبة التي لها جدول توزيع تكراري

المدى = مركز الفئة الأخيرة ــ مركز الفئة الأولى

مثال / أوجد المدى للبيانات التالية

الدرجات	عدد الطلاب
10 - 0	2
20 - 10	12
30 - 20	22
40 - 30	8
50 - 40	6

مركز الفئة الأخيرة = 50 + 40 ÷ 2 = 90 ÷ 2 = 45

مركز الفئة الأولى = 10 + 0 ÷ 2 = 10 ÷ 2 = 5

المدى = 45 ــ 5 = 40

ملاحظة : هنا تم استخراج مركز الفئة للفئة الاولى والفئة الاخيرة ثم تم الطرد (استخراج مركز الفئة يتم من جمع الحد الأدنى للفئة + الحد الاعلى لنفس الفئة مقسوما على 2)

المميزات والعيوب التي يتصف بها المدى

قليل الإستخدام في مقاييس التشتت

سهولة استخدامه حسابيا

يعتمد على قيمتين فقط في عملية حسابه

يتأثر بالقيم الشاذة والمتطرفة

لا نعطيه اهتمام بالغ في قياس التشتت للبيانات

ثالثا / التباين والانحراف المعياري :

يعتبر الإنحراف المعياري أكثر استخداما لقياس تشتت البيانات ، قيمة الإنحراف المعياري تخبرنا عن مدى تشتت وانتشار البيانات حول الوسط الحسابي ، فكلما كانت قيمة الإنحراف المعياري متدنية دلت على أن قيم البيانات متقاربة في مداها حول الوسط الحسابي بينما لو كانت قيمة الإنحراف المعياري عالية دلت على أن قيم البيانات متباعدة في مداها حول الوسط الحسابي . الإنحراف المعياري يمكن الحصول عليه بعد ايجاد قيمة التباين فهو يساوي الجذر التربيعي لقيمة التباين

وقانون التباين بالصيغة التالية :

يرمز لتباين العينة بالرمز S^2 ويرمز لتباين المجتمع بالرمز σ^2 ويرمز للانحراف المعياري بالرمز M.D

صيغة التباين للمجتمع $$\sigma^2 = \frac{\sum (X-\mu)^2}{N}$$

صيغة التباين للعينة $$s^2 = \frac{\sum (x_i - \overline{X})^2}{n-1}$$

صيغة الانحراف المعياري للعينة $s = \sqrt{\frac{\sum(x-\overline{x})^2}{N-1}}$

ملاحظة: S = M.D الانحراف المعياري يساوي جذر التباين

أولاً: بيانات غير مبوبة :

مثال:

احسب كلاً من التباين والانحراف المعياري للقيم 12 ، 15 ، 11 ، 17 ، 18 ، 20 ، 19

الحل: الحل باستخدام SPSS

نكون جدول المعلومات التالي:

$(X_i - \overline{X})^2$	$(X_i - \overline{X})$	X_i
16	12 – 16 = –4	12
1	15 – 16 = – 1	15
25	11 – 16 = – 5	11
1	17 – 16 = 1	17
4	18 – 16 = 2	18
16	20 – 16 = 4	20
9	19 – 16 = 3	19
$\sum (X_i - \overline{X})^2 = 72$	$\overline{X} = 112/7 = 16$	$\sum X_i = 112$

نحسب التباين من القانون أعلاه:

$$S^2 = [\sum (x_i - \overline{X})^2] / (n - 1)$$
$$= 72 / 6$$
$$= 12$$

الانحراف المعياري يساوي الجذر ألتربيعي للتباين أي:

$$S.D = 3.46 = S$$

يمكن استخراج التباين دون الوسط الحسابي القانون كما يلي :

$$S^2 = [\sum (X_i^2) - (\sum X_i)^2 / n] / (n - 1) \ldots (1)$$

حساب الانحراف المعياري من البيانات المبوبة:

$$S^2 = \frac{\Sigma(\text{fixi})2 - (\Sigma(fixi))2}{\text{n}-1}$$

$$S^2 = [\Sigma(f_i x_i)^2 - (\Sigma f_i x_i)^2] / n-1$$

مثال:

احسب التشتت باستخدام الانحراف الانحراف من جدول التوزيع التكراري الآتي والذي يبين درجات 30 طالب في امتحان ما.

Total	24 – 26	21 – 23	18 – 20	15 – 17	12 – 14	الفئات
30	2	7	10	8	3	التكرار

الحل:

الحل باستخدام SPSS

نكون الجدول الشامل للبيانات المطلوبة للصيغ الرياضية الخاصة بالانحراف المعياري :

نجد أن:

$f_i X_i^2$	X_i^2	$f_i X_i$	X_i	f_i	الفئات
507	169	39	13	3	12 – 14
2048	256	128	16	8	15 – 17
3610	361	190	19	10	18 – 20
3388	484	154	22	7	21 – 23
1250	625	50	25	2	24 – 26
10803	1895	561		30	Total

$$\sigma = \sqrt{\frac{\sum f_i X_i^2 - \frac{\sum (f_i X_i)^2}{n}}{n-1}}$$

$$\sigma = \sqrt{\frac{10803 - \frac{(561)^2}{30}}{30-1}}$$

$$\sigma = \sqrt{\frac{10803 - 10490.7}{29}}$$

$$\sigma = \sqrt{\frac{312.3}{29}}$$

$$\sigma = \sqrt{10.769}$$

$$\sigma = 3.28$$

الانحراف المعياري = 3.28

يمكن استخدام الصيغة أعلاه لحساب الانحراف المعياري كما يلي

نكون الجدول الشامل للبيانات المطلوبة للصيغ الرياضية الخاصة بالانحراف المعياري و باستخدام الصيغة (2) أعلاه نجد أن:

$$\bar{X} = 561 / 30 = 18.7$$

$f_i (X_i - \bar{X})^2$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})$	$f_i X_i$	X_i	f_i	الفئات
97.47	32.49	– 5.7	39	13	3	12 – 14
58.32	7.29	– 2.7	128	16	8	15 – 17
0.90	0.09	0.3	190	19	10	18 – 20
76.23	10.89	3.3	154	22	7	21 – 23
79.38	39.69	6.3	50	25	2	24 – 26
312.3	90.45		561		30	Total

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

$$\sigma = \sqrt{\frac{312.3}{30}}$$

$$\sigma = 3.23$$

معامل الاختلاف:

هو معامل نسبي يستخدم للمقارنة بين تشتت بيانات ظاهرتين او اكثر مختلفتين في وحدات القياس او متفقتين ، او مختلفتين في القيمة المتوسطة لهما الظاهرة التي معامل اختلافها أصغر تكون اقل تشتتا والظاهرة التي معامل اختلافها أكبر تكون أكثر تشتتا

ويمكن حساب معامل الاختلاف بالمعادلة : $\%100 * \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}}$ = معامل الاختلاف

ويرمز لة بالرمز $V.C = \frac{S}{\bar{X}} * \%100$

مثال : قارن بي تشتت درجات كل من الطلاب والطالبات من واقع البيا نات التالية :

الحل : للمقارنة بين تشتت درجات كل من الطلاب والطالبات يتم حساب معامل الاختلاف لدرجات الطلاب ومعامل الاختلاف لدرجات الطالبات والمقارنة بينهما كما يلي :

الطالبات	الطلاب	النوع \ المقاييس
80	70	الوسط الحسابي
6	7	الانحراف المعياري

اولا : معامل الاختلاف لدرجات الطلاب $V.C = \frac{S}{\bar{X}} * \%100 = \frac{70}{7} * \%100 = \%\ 10$

معامل الاختلاف لدرجات الطلاب $V.C = \frac{S}{\bar{X}} * \%100 = \frac{80}{6} * \%100 = \%\ 7{,}5$

من الملاحظ : ان معامل الاختلاف لدرجات الطلاب اكبر من معامل الاختلاف لدرجات الطالبات وبالتالي درجات الطلاب اكثر تشتتا من درجات الطالبات

الدرجة المعيارية : هو احد مقاييس التشتت النسبية ويرمز له بالرمز Z في توزيع متوسطه الحسابي يساوي صفر والانحراف المعياري له يساوي واحد في كثير من الاحيان نحتاجها الى مقرنة قيمتين مختلفتين تنتمي كل منهما الى مجموعة معينة ولا تتم المقارنة بشكل مطلق وانما تحويل القيم المراد مقارنتها الى وحدات معيارية خالية من تأثير الوسط الحسابي والانحراف المعياري ومن وحدات القياس للمتغير الاصلي X .

التحويل الى الدرجة المعيارية

فاذا كانت (x_1 ، x_2 ، x_3 ، x_4 ، x_5 ،........ ، x_n) تمثل قياسات عينة من المفردات قوامها n وان x يمثل الوسط الحسابي للمتغير x وان s يمثل الانحراف المعياري

عندئذ لتحويل المتغير (x) الى الدرجة المعيارية $Z = \frac{x - \bar{X}}{s}$

وبذلك اصبح الوسط الحسابي = 0 والانحراف المعياري =1

مثال : اذا كانت تقديرات احد العاملين في مصنع ما على مقياسين اولهما لمستوى الانتاج وثانيهما لمستوى الرضا عن العمل . علما بان المقياسين طبقا على مجموعة من العاملين في ذلك المصنع ، كما موضح في الجدول ادناه : المطلوب في أي حالة كان وضع هذا العامل افضل ؟

مقياس الرضا عن العمل	مقياس الانتاج	
80	42	تقديرات العامل
60	35	متوسط التقديرات
12	5	الانحراف المعياري للتقديرات

تحويل تقدير العامل لمقياس الانتاج الى الدرجة المعيارية

$$Z_1 = \frac{x - \bar{X}}{S} = \frac{42-35}{5} = 1,2$$

تحويل تقدير العامل لمقياس الرضا عن العمل الى الدرجة المعيارية

$$Z_2 = \frac{x - \bar{X}}{S} = \frac{80-60}{12} = 1,67$$

ومن مقارنة الدرجات المعيارية للمقياسين نجد ان Z_2 اكبر من Z_1

وهكذا ان وضع تقدير العامل لمقياس الرضا عن العمل أفضل حالا من وضعه لمقياس الانتاج بالنسبة لجميع العاملين

معامل ارتباط الرتب (سبيرمان)

مثال :

لدراسة علاقة ارتباط تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات، اخترنا ثمان 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؟

الحل :

X	Y	رتبة X	رتبة Y	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
					8

$$r_s = 1 - \frac{6(8)}{5(25-1)} = 1 - \frac{48}{120}$$

$$r_s = 1 - 0,4 = 0,6$$

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

ملاحظات هامة:

1- أن معامل ارتباط الرتب يمكن حسابة سواء اكانت البيانات كمية او وصفية بينما معامل الارتباط (بيرسون) لا يمكن حسابه إلا على المتغيرات الكمية .

2 - يتميز معامل سبيرمان لارتباط الرتب بسهولة حسابه حتى لو كانت البيانات غير مبوبة .

3 - يعاب علية اهماله لفروق الاعداد عند حساب الرتب وبالتالي فهو اقل دقة .

4 - يصعب حساب المتغيرات العادية اذا كانت كبيرة العدد ولذلك يفضل استخدامه لتحديد درجة ارتباط بيانات بيانات كمية اقل من 30 .

الانحدار

1 - التنبؤ : هو تقدير القيمة المستقبلية لمتغير واحد بناء على معرفة قيم متغير آخر و يفيدنا في

تحديد شكل العلاقة بين المتغيرين رياضيا وبيانيا (خط الانحدار) .

توضيح اتجاه العلاقة بين المتغيرين.

التنبؤ بقيمة أحد المتغيرين بدلالة المتغير الآخر.

والانحدار هو أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلة الانحدار ، وله أنواع:

الانحدار الخطي البسيط : فكلمة " بسيط " تعني أن المتغير التابع Y يعتمد على متغير مستقل واحد وهو X وكلمة " خطي " تعني أن العلاقة بين المتغيرين (X , Y) علاقة خطية.

الانحدار المتعدد: إذا كان المتغير Y يعتمد على أكثر من متغير مستقل .

الانحدار غير الخطي: إذا كانت العلاقة بين المتغير Y والمتغيرات المستقلة غير خطية كأن تكون من الدرجة الثانية أو آسية .

الانحدار الخطي البسيط

بعد تمثيل الازواج المرتبة بالمستوي نحصل على شكل الانتشار فأذا اظهر الشكل الانتشاري للبيانات ان هناك علاقة خطية بين المتغيرين نقوم بتقدير خط الانحدار Y على X بواسطة العلاقة التالية :

$$\hat{y} = a + bx$$

حيث ان a : ثابت الانحدار أو الجزء المقطوع من محور y

B :ميل الخط المستقيم أو معامل انحدار XعلىY أو (Y/X)

$$a = \frac{\Sigma Y - b\,\Sigma x}{n} \qquad b = \frac{n\Sigma xY - \Sigma x\,\Sigma Y}{n\Sigma x2 - (\Sigma x)2}$$

لإيجاد قيمة مقدرة جديدة $\hat{y}_h$ نعوض بقيمة معلومة للمتغير المستقل ولتكن x_h في معادلة تقدير خط الانحدار Y/X

$$\hat{y} = a + bx$$

ملاحظات هامة:

* ميل الخط يمثل كمية التغير في Y المناظرة للتغير في X بمقدار وحدة واحدة.

* إشارة معامل الانحدار تدل على نوع الارتباط (طردي او عكسي) .

* توجد علاقة بين معامل الانحدار ومعامل الارتباط الخطي .

مثال :

لدراسة علاقة الاستهلاك المحلي (y) بالإنتاج (x) لمادة الإسفلت (بالمليون برميل) خلال عدة سنوات، أخذنا عشر قراءات تقريبية كما يلي :

y	6	8	9	8	7	6	5	6	5	5
x	10	13	15	14	9	7	6	6	5	5

أوجد معادلة الانحدار الخطي البسيط وتوقع قيمة الاستهلاك عندما يصل إنتاج 16,000,000 برميل .

الحل :

X^2	XY	Y	X
100	60	6	10
169	104	8	13
225	135	9	15
196	112	8	14
81	63	7	9
49	42	6	7
36	30	5	6
36	36	6	6
25	25	5	5
25	25	5	5
Σ =942	Σ =632	Σ =65	Σ =90

$$b = \frac{n\Sigma xY - \Sigma x\, \Sigma Y}{n\Sigma x2 - (\Sigma x)2} = \frac{6320 - (90)(65)}{9420 - (90)2}$$

$$b = \frac{6320 - 5850}{9420 - 8100} = \frac{470}{1320} = 0{,}36$$

$$a = \frac{\Sigma Y - b\, \Sigma x}{n} = \frac{(65) - (0{,}36 * 90)}{10} = 3{,}26$$

.:. معادلة خط الانحدار البسيط في هذه الحالة :

$$\hat{y} = a + bx$$

$$\hat{y} = 3{,}26 + 0{,}36\, X$$

ولتوقع قيمة الاستهلاك المحلي عندما يصل الإنتاج 16000000 برميل نحول وحدة هذه القيمة من برميل إلى مليون برميل بالقسمة على مليون أي أن القيمة المستخدمة في توقع الاستهلاك هي $x_h = 16$ وبالتعويض في المعادلة السابقة نجد أن :

$$\hat{y} = a + bx$$

$$\hat{y} = 3{,}26 + 0{,}36\,(16) = 9.02$$

أي أن الاستهلاك قد يصل إلى 9.02 مليون برميل، أي ما يعادل 9020000 برميل خلال السنة .

موضوع : السلسلة الزمنية

السلسلة الزمنية بكل بساطة هي مجموعة القياسات المسجلة لمتغير واحد أو أكثر مرتبة حسب زمن وقوعها .
رياضياً : نقول إن متغير الزمن المستقل (t) والقيم المناظرة له للمتغير التابع (y) وإن كل قيمة في الزمن t يقابلها قيم للمتغير التابع y فإن y دالة في الزمن t

أي: $$y = F(t)$$

من الأمور الطبيعية والواجبة للحكومات والمؤسسات والشركات التجارية منها والصناعية والتعليمية وغيرها بالتخطيط لمستقبلها لتحقيق الأهداف الخاصة والعامة وتقديم كافة الخدمات والوصول لحالة العدل والاستقرار للمجتمع والعمل على أتحاذ قرارات التنبؤ بوقوع الأحداث قبل وقوعها في كافة أوجه النشاط التي تخص المجتمع، وتعتبر السلاسل الزمنية من أهم أسليب التنبؤ حول المستقبل من خلال وقائع الأمس واليوم.
من أهم السلاسل الزمنية تلك الخاصة بالمؤشرات الاقتصادية والمبيعات السنوية للشركات بكافة أوجه نشاطاتها والتعليم وحجم السكان وما شابه ذلك.
والتغير الذي يحدث في قيم متغير السلسلة الزمنية أو قيم متغيراتها يعتبر دالة في الزمن يمكن تمثيلها بيانياً باتخاذ المحور الأفقي للزمن والرأسي لقيم المتغير

وسنتعرض هنا لمكونات هذه السلسلة الزمنية وكيفية قياس المتغيرات التي تخص السلسلة خلال فترة زمنية (سنوية - نصف سنوية - شهرية - ....) ونخرج منها بالتنبؤ بافتراض أن التطبيقات الاقتصادية تفترض تمتع السلسلة الزمنية بخاصية السكون والاستقرار، ويرجع عدم الاستقرار لمكونات السلسلة الزمنية الأربع وهي :

*** الاتجاه العام(Secular Trend)**
*** التغيرات الموسمية(Seasonal Variations)**
*** التغيرات الدورية(Cyclical Variations)**
*** التغيرات العشوائية أو العرضية(Irregular Variations)**

إن هذه المكونات (العناصر) الأربع الخاصة بالسلسلة الزمنية والتي تتأثر بالعوامل الاقتصادية والبيئية والاجتماعية والسياسية وما إلى ذلك وسنتعرض لكل من هذه العناصر بصورة تفي بالغرض المطلوب .

* الاتجاه العام : -

اتجاه السلسلة الذي تأخذه السلسلة الزمنية للظاهرة محل الدراسة من خلال فترة زمنية سواء في اطراد متزايد (اتجاه موجب) أو متناقص (اتجاه سالب) أو الأمرين معاً كالنمو السكاني في حالة التزايد والأمية بالتناقص وكمبيعات مادة ما تتطور بشكل واضح كجهاز التلفزيون الأسود والأبيض والملون أو عدد العمال للشركات التي تستخدم التكنولوجيات وفي كل الحالات يكون التغيير فيها ليس مفاجئاً بل بالتدريج وهو ميزة للاتجاه العام الذي يعتبر من أهم عناصر السلسلة الزمنية .

الاتجاه العام يبين الحركة المنتظمة لحالات التزايد (النمو) والتناقص (الركود) لفترات زمنية طويلة .
الفترة الزمنية تشمل دورتين اقتصاديتين على الأقل بقصد الحصول على نتائج وافية .
الاتجاه العام يقيس متوسط التغير لكل فترة زمنية واحدة

* التغيرات الموسمية :-

فترات خاصة كالأعياد أو بداية العام الدراسي مثلاً حيث يكثر بيع سلعة معينة وتعد هذه الفترات مجالاً جيداً للدراسة وقد يلعب الطقس والتقاليد والاحتفالات الوطنية والدينية كالحج بالتأثير على التغير الموسمي الذي لا يزيد طول فترته عن السنة او قد يكون أسبوعياً أو يومياً .

* التغيرات الدورية :-

التغيرات التي تطرأ على الدورات الاقتصادية من ارتفاع وهبوط بمدة تتجاوز السنة وبيانها كبيان دالة الجيب أو الجيب تمام مع وجود اختلاف في الطول والسعة وتضم عدة خمسة مراحل في الدورة الكاملة هي الارتفاع الأولي - التراجع - الركود - الانتعاش - الارتفاع النهائي وقد تمتد طول الفترة (الدورة الكاملة) من ثماني سنوات إلى عشر سنوات وترجع لعوامل كثيرة مثل سياسة الحكومة والعلاقات الدولية وغيرها ويقاس طول الدورة (التجارية) بطول الفترة الزمنية بين مرحلتي ازدهار متتاليتين أو ركود متتاليتين .

* التغيرات العشوائية :-

تشير هذه التغيرات وهي غير منتظمة لتحركات السلسلة الزمنية لأعلى ولأسفل بعد استبعاد التغيرات الأخرى والاتجاه العام وتنشأ هذه التغيرات لعوامل لا يمكن التحكم بها كالزلازل والبراكين والفيضانات والحروب وإفلاس بنك وما شابه ذلك، ومن الواضح بأنه لا يمكن التنبؤ بها لعدم انتظامها من جهة وللفترة الزمنية الصغيرة التي تحدث فيها ويسهل تأثيرها عند دراسة العناصر الأخرى للسلسلة الزمنية وغالباً يشار إليها بالتغيرات المتبقية Residual Variations لكونها تضم ما تبقى من العوامل التي لم يشار إليها في عناصر السلسلة الثلاثة السابق ذكرها وبالطبع هذا العنصر عشوائي لأنه يقع فجأة أو للصدفة .

الهدف من تحليل السلسلة الزمنية هو التعرف على مكوناتها (الاتجاه العام - التغيرات الموسمية - التغيرات الدورية - التغيرات العشوائية) كلاً على حدة حيث يستخدم نموذج يعرف بنموذج الجمع (Additive model) وآخر بنموذج الضرب (Multiplicative) للسلسلة الزمنية بقصد تجزئة السلسلة الزمنية وذلك بتحديد علاقة السلسلة بمكوناتها والنموذجين الجمع والضرب هما تقدير جيد للعلاقة الحقيقية التي تظهرها البيانات هذه وسنرمز بالرموز الآتية:

T قيمة الاتجاه العام(Secular Trend)

S قيمة التغيرات الموسمية(Seasonal Variations)

C قيمة التغيرات الدورية(Cyclical Variations)

I قيمة التغيرات العشوائية أو العرضية(Irregular Variations)

Y قيمة الظاهرة عند زمن معين (مشاهدات السلسلة الزمنية)

نموذج الضرب هو $Y = T . S . C . I$:مع التأكد بأن T قيمة عددية ، S , C , I نسب مئوية.

نموذج الجمع هو $Y = T + S + C + I$: يعبر عن كل منها بقيمة عددية.

طريقة المربعات الصغرى :Least Square Method

الطريقة الأكثر استخدام (سبق ذكرها في الانحدار) بها يتم التقليل من مجموع مربعات الفروق بين القيم الفعلية والقيم المحسوبة حيث القيم الفعلية هي الزمن والقيم المحسوبة قيم المتغير المطلوب له إيجاد اتجاهه العام وسنرمز بالرمز X للقيم الفعلية وبالرمز $\hat{Y}$ لقيم الاتجاه المحتسبة .

نقاط خط الانحدار تمثل المتوسط الشرطي للمتغير التابع Y لقيمة المتغير المستقل X والفرق (الانحراف) بين قيم المتغير Y عن المتوسطات الشرطية هي الأخطاء العشوائية وتمثل الانحرافات لقيم السلسلة عن خط الاتجاه العام للبيانات باستثناء المتغيرات الموسمية وعند توفيق خط الاتجاه العام بهذه الطريقة سيكون $\hat{Y}$ ممثلة للقيم الاتجاهية و X تمثل الزمن وسنعتمد الصيغ الرياضية الاتية :

(1) $$\hat{Y} = a + b X + e$$

(2) $$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

(3) $$a = \bar{Y} - b\bar{X}$$

مثال :

لإيجاد معادلة الاتجاه العام للبيانات المبينة في الجدول التالي :

X^2	X Y	Y	X	السنة X
1	40	40	1	1985
4	66	33	2	1986
9	87	29	3	1987
16	100	25	4	1988
25	105	21	5	1989
36	192	32	6	1990
49	280	40	7	1991
64	360	45	8	1992
51	369	41	9	1993
100	400	40	10	1994
$\sum X^2 = 385$	$\sum XY = 1999$	$\bar{Y} = 346 / 10 = 346$	$\bar{X} = 55/10 = 55$	

باستخدام الصيغة السابقة :

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b = \frac{10(1999) - 55 \times 346}{10(385) - (55)^2}$$

$$b = \frac{19990 - 19030}{3850 - 3025}$$

$$b = \frac{960}{825} = 1.16$$

$$a = \bar{Y} - b\ \bar{X}$$

$$a = 34.6 - 1.16(5.5) \quad = \quad 34.6 - 6.38 \ = 28.22$$

$$\hat{Y} = a + b\,X \qquad\qquad \hat{Y} = 28.22 + 1.16\,X$$

معادلة خط الاتجاه العام يمكن استخدامها للتنبؤ بالأجور مثلاً حسب مثالنا هنا لمعرفة المتوقع للأجور سنة 1998 كما يلي:
سنة 1998 تأخذ الترتيب 14 حسب الجدول أعلاه الذي يبدأ سنة 1985 حيث نستبدل X بالقيمة 14 كما يلي:

$$\hat{Y} = 28.22 + 1.16\,X$$

$$\hat{Y} = 28.22 + 1.16(14)$$

$$\hat{Y} = 28.22 + 16.24$$

$$\hat{Y} = 44.6$$

وهذه معادلة الاتجاه العام لقيم الأجور للفترة المبينة بالجدول .

ب – حساب الوسيط في البيانات المبوبة

قانون الوسيط بالنسبة للبيانات المبوبة هو : $Me = A + [\frac{\frac{n}{2} - N1}{n1}] * L$

مجموع التكرارات

N1 التكرار المتجمع للفئة التي تسبق الفئة الوسيطية

n1 تكرار الفئة الوسيطية

A الحد الادنى للفئة الوسيطية

L طول الفئة

يتم استخراج رتبة الفئة الوسيطية من قسمة مجموع التكرارات على 2

مثال : احسب الوسيط واشرح النتيجة في حالة كان توزيع الاجور اليومية للعمال على النحو التالي :

الاجور اليومية (الفئات)	التكرارات ni	Ni التكرار المتجمع الصاعد
550 - 400	17	17
700 – 550	13	30
850 – 700	10	40
1000 – 850	5	45
1150 – 1000	2	47
1300 – 1150	2	49
1450 – 1300	1	50

الحل :

تحديد الفئة الوسيطية : وهي اول فئة تكرار متجمع الصاعد اكبر او يساوي n/2 وبما ان n مجموع التكرارات

اذا كان $\frac{n}{2} = \frac{50}{2} = 25$

$$Me = A + [\frac{\frac{n}{2} - N1}{n1}] * L = 550 + [\frac{25 - 17}{13}] * 150 = 642 . 30$$

الشرح : هناك 50 % من العمال اجورهم اليومية اقل من 642,30 د و 50 % من العمال اجورهم اليومية اكبر من 642,30 د .

الوسيط بيانيا هو نقطة التقاطع بين المنحني المتجمع الصاعد والمنحني المتجمع النازل .

الاحتمالات

النظرية الاحتمالية (Probability theory) : هي النظرية التي تدرس احتمال الحوادث العشوائية، بالنسبة للرياضيين، **الاحتمال :** هو قياس إمكانية وقوع حدث ما . يُقاس الاحتمال بأنه رقم بين الصفر والواحد حيث يشير الصفر إلى الاستحالة ويشير الواحد إلى التأكيد . كلما زاد احتمال الحدث، كلما زادت إمكانية وقوع هذا الحدث. أحد الأمثلة البسيطة هي رمي العملة (غير المنحاز). لأن العملة غير منحازة، فإن الناتجين (كتابة وصورة) متساويان في الاحتمال تماما أي أن احتمالية ظهور الكتابة تساوي احتمالية ظهور الصورة ، ولأنه لا يوجد احتمالات أخرى فإن إمكانية ظهور "الكتابه" أو "الصورة" هي ½ (والتي يمكن كتابتها 0.5 أو 50%)

ملاحظة : **كلمة الاحتمال تستخدم للتعبير عن قياس فرصة حدوث حادثة معينة غير مؤكدة الحدوث .**
إن المقياس الكمي الذي يقيس فرصة حدوث حادثة معينة يسمى بمقياس الاحتمال وقيمة هذا المقياس تتراوح بين الصفر والواحد .
كلما زادت فرصة وقوع الحادثة كلما اقتربت قيمة هذا المقياس من الواحد . وكلما قلت فرصة وقوع الحادثة كلما اقتربت قيمة هذا المقياس من الصفر .

فضاء العينة (Sample Space) : ♣**فضاء العينة** للتجربة العشوائية هي المجموعة المكونة من جميع النتائج الممكنة للتجربة**.** ونرمز لفضاء العينة بالرمز S ويرمز لعدد عناصر فضاء العينة بالرمز n(s)
نقطة العينة هي أي نتيجة من نتائج التجربة العشوائية أي أنها أي عنصر من عناصر فضاء العينة

مثال : إذا ألقيت قطعة نقود مرة واحدة، فإن فضاء العينة $S = \{ T , H \} \Leftarrow n(S) = 2$
مثال : إذا ألقيت قطعة نقود مرتين متتاليتين، فأن فضاء العينة $S = \{(H,H) , (H,T) , (T,H) , (T,T)\}$
$\Leftarrow n(S) = 4$
مثال : إذا ألقي حجر النرد مرة واحدة، فإن فضاء العينة: $S = \{ 1 , 2 , 3 , 4 , 5 , 6 \} \Leftarrow n(S) = 6$
مثال : إذا ألقي حجر النرد مرتين متتاليتين ، فإن فضاء العينة :
$S = \{ (1,1) , (1,2) , (1,3) , (1,4) , (1,5) , (1,6)\ (2,1) , (2,2) , (2,3) , (2,4) , (2,5) , (2,6)$
$(3,1) , (3,2) , (3,3) , (3,4) , (3,5) , (3,6)\ (4,1) , (4,2) , (4,3) , (4,4) , (4,5) , (4,6)\ (5,1) ,$
$(5,2) , (5,3) , (5,4) , (5,5) , (5,6)\ (6,1) , (6,2) , (6,3) , (6,4) , (6,5) , (6,6) \}$
$\Leftarrow n(S) = 36 \Leftarrow$ **لايجاد فضاء العينة :** باستخدام حاصل الضرب الديكارتي او الكارتيزي :

$$S = \{ 1 , 2 , 3 , 4 , 5 , 6 \} \times \{ 1 , 2 , 3 , 4 , 5 , 6 \}$$

التجربة العشوائية : (Random Experiment) التجربة العشوائية هي تجربة أو عملية تحقق الشروط التالية: ♣ جميع النتائج الممكنة للتجربة تكون معلومة مسبقا قبل إجرائها .
♣ لا يمكن التنبؤ بنتيجة التجربة بشكل قطعي ومؤكد قبل إجرائها .
♣يمكن معرفة أو قياس فرصة ظهور كل نتيجة من نتائج التجربة قبل إجراء التجربة .

الحادثة أو الحدث : Event هي مجموعة جزئية من فضاء العينة (S) **.**
A حادثة إذا وفقط إذا كانت $A \subseteq S$.

♣ يقال بأن الحادثة A وقعت إذا كانت نتيجة التجربة هي أحد عناصر الحادثة A

♣ الحادثة المستحيلة $\phi \subseteq S$ ، حيث أن ϕ هي المجموعة الخالية .

♣ الحادثة المؤكدة $S \subseteq S$

♣ يرمز لعدد عناصر الحادثة A بالرمز n(A) ونقول بأن الحادثة وقعت إذا كانت نتيجة التجربة هي أحد عناصر الحادثة A

مثال : احسب الحوادث التالية وعدد عناصرها وذلك في تجربة قذف قطعة النقود مرتين متتاليتين:

A = { الحصول على صورة في الرمية الأولى }

B = { الحصول على كتابة في الرمية الأولى }

C = { الحصول على صورة واحدة على الأقل }

الحل :

S= { (H,H) , (H,T) , (T,H) ,(T,T) } ; n(S) = 4

A= { (H,H) , (H,T) } ; n(A) = 2

B= { (T,H) , (T,T) } ; n(B) = 2

C= { (H,H) , (H,T) , (T,H) } ; n(C) = 3

مثال : في مثال تجربة رمي حجر النرد مرة واحدة وتسجيل الرقم الظاهر على الوجه العلوي فإن المجموعات التالية تشكل حوادث لأنها مجموعات جزئية من فضاء العينة S .

الحل :

عدد العناصر	الحادثة Event	
n(A)=3	$A \subseteq S$	A = (ظهور عدد زوجي) = (2 ، 4 ، 6);
n(B)=3	$B \subseteq S$	B = (ظهور عدد فردي) = (1 ، 3 ، 5);
n(C)=5	$C \subseteq S$	C = (ظهور عدد اقل من ستة) = (1 ، 2 ، 3، 4، 5);
n(D)=1	$D \subseteq S$	D = (ظهور العدد ستة) = (6);
n(ϕ)= 0	$\phi \subseteq S$	ϕ = (ظهور عدد سالب) = ( ) ;
n(S)=6	$S \subseteq S$	S = (ظهور عدد موجب) = (1، 2 ، 3، 4 ، 5 ، 6);

الاحتمال : Probability احتمال الحادثة Probability of An Event

• احتمال الحادثة A هو مقياس عددي يرمز له بالرمز P (A) ويقيس فرصة وقوع الحادثة A عند إجراء التجربة. وتتراوح قيمة هذا المقياس بين الواحد الصحيح والصفر .

• إذا كان لدينا تجربة عشوائية جميع نتائجها متساوية الفرصة وعدد عناصر فضاء العينة لها محدود ويساوي n(S) فإن احتمال الحادثة A يعرف بالصيغة التالية :

$$P(A) = \frac{n(A)}{n(S)} = \frac{\text{عدد عناصر الحادثة } (A)}{\text{عدد عناصر فضاء العينة } (S)}$$

مثال : أوجد احتمال الحوادث في مثال تجربة رمي حجر النرد المتزن مرة واحدة

الحل : بما أن نتائج تجربة رمي حجر النرد المتزن متساوية الفرصة وعدد عناصر فضاء العينة محدود n(S) = 6 ، فإن احتمالات الحوادث هي :

عدد العناصر	الحادثة Event
n(A)=3	A ⊆ S A = (ظهور عدد زوجي) = (2 ، 4 ، 6);
n(B)=3	B ⊆ S B = (ظهور عدد فردي) = (1 ، 3 ، 5);
n(C)=5	C ⊆ S C = (ظهور عدد اقل من ستة) = (1 ، 2 ، 3، 4، 5);
n(D)=1	D ⊆ S D = (ظهور العدد ستة) = (6);
n(ϕ)= 0	ϕ ⊆ S ϕ = (ظهور عدد سالب) = ( ) ;
n(S)=6	S ⊆ S S = (ظهور عدد موجب) = (1، 2 ، 3، 4 ، 5 ، 6);

هذا يؤدي الى

الاحتمال	عدد العناصر	الحادثة Event
P(A) = n(A)/n(S)= 3/6 = 0.5	n(A) = 3	A = (2 ، 4 ، 6);
P(B) = n(B)/n(S)= 3/6 = 0.5	n(B) = 3	B = (1 ، 3 ، 5);
P(C) = n(C)/n(S)= 5/6 = 0.833	n(C) = 5	C = (1 ، 2 ، 3، 4، 5);
P(D) = n(D)/n(S)= 1/6 = 0.1667	n(D) = 1	D = (6);
P(ϕ) = n(ϕ)/n(S)= 0/6 = 0.0	n(ϕ) = 0	ϕ = ( ) ;
P(S) = n(S)/n(S)= 6/6 = 1	n(S) = 6 S	= (1، 2 ، 3، 4 ، 5 ،6)

مثال :احسب احتمالات الحوادث التالية لتجربة قذف قطعة النقود المتزنة مرتين متتاليتين :

A = { الحصول على صورة في الرمية الأولى }
B = { الحصول على كتابة في الرمية الأولى }
C = { الحصول على صورة واحدة على الأقل }

الحل :بما أن نتائج تجربة قذف قطعة النقود المتزنة متساوية الفرصة وعدد عناصر فضاء العينة محدود فإن احتمالات الحوادث هي:

الاحتمال	عدد العناصر	الحادثة Event
P(A) = n(A)/n(S)= 2/4 = 0.5	; n(A) = 2	A= { (H,H) , (H,T) }
P(B) = n(B)/n(S)= 2/4 = 0.5	; n(B) = 2	B= { (T,H) , (T,T) }
P(C) = n(C)/n(S)= 3/4 = 0.75	; n(C) = 3	C= { (H,H) , (H,T) , (T,H) }

مفهوم الاحتمال: -هو إمكانية وقوع أمر ما لسنا على ثقة تامة بحدوثه ، ويلعب الاحتمال دوراً أساسياً في الحياة اليومية بالتنبؤ بإمكانية وقوع حدث ما وهو النظرية التي يستخدمها الإحصائي لتساعده في معرفة مدى تمثيل العينة العشوائية محل الدراسة للمجتمع المأخوذ منه العينة . والاحتمال يبحث في ثلاثة مسائل هامة معتمدة على القواعد الخاصة بالاحتمال التي سنذكرها في حينها والمسائل الثلاثة هي :

1 - حساب الاحتمال المتمثل بالتكرار النسبي .
2 - حساب الاحتمال بدلالة احتمالات أخرى معلومة من خلال عمليات مثل الاتحاد والتقاطع والفرق و ....
3 - طرق إجراء التقدير كالتوزيعات الاحتمالية .

انواع الحدث في نظرية الاحتمالات :

الحدث البسيط :(Simple event) وهو الحدث المكون من عنصر واحد مثل {1} في تجربة إلقاء حجر النرد .

الحدث المركب :(Compound event)الحدث المكون من أكثر من عنصر مثل {2، 4، 6} حدث العدد زوجي في تجربة إلقاء حجر النرد .

الحدث المستحيل:الحدث الذي لا يحوي أي عنصر كحدث ظهور العدد 7 في تجربة إلقاء حجر النرد .

الحدث المؤكد : الحدث الذي يضم كافة عناصر الفضاء كحدث ظهور عدد أقل من 7 في تجربة إلقاء حجر النرد .

الحدثان المتنافيان :(Mutually Exclusive events)الحدثان اللذان لا يشتركا في أي عنصر وتقاطعهم المجموعة الخالية أي $A \cap B = \phi$ مثل {2} ، {3} ، وتعرف بالأحداث غير المتصلة.

لأحداث المنتظمة : (dependent events) المتساوية في احتمالاتها. ففي تجربة إلقاء حجر النرد مرة واحدة يكون :

$P(1) = P(2) = P(3) =P(4) = P(5) = P(6) = 1:6$

الأحداث الشاملة :(Exhaustive events) إذا كان S فضاء عينة ما فإن الأحداث A, B, C شاملة إذا تحقق الشروط الثلاثة الآتية:

1 - متنافية فيما بينها أي $A \cap B = \phi$: و $A \cap C = \phi$ و $C \cap B = \phi$
2 - أياً منها ليست خالية أي $A \neq \phi$ و$B \neq \phi$ و$C \neq \phi$
3 - اتحادها يساوي S أي $A \cup B \cup C = S$

الأحداث المكملة:(Complementary events) الحدثان اللذان اتحادهم يساوي فضاء العينة بمعنى A حدث فإن $\bar{A}$ الحدث المكمل حيث $\bar{A} \cup A = S$ بمعنى $p(\bar{A}) = 1 - p(A)$

الحدثان المستقلان :(Independent events) اللذان لا يتأثر أي منهم بالآخر (وقع أحدهم لا يؤثر أو يتأثر بوقوع أو عدم وقوع الآخر) .

قاعدة الضرب للاحتمالات للإحداث المستقلة $P(A \cap B) = P(B) \times P(A)$

يمكن تعميم هذه القاعدة لأكثر من حدثين

$P(A \cap B \cap C \cap \ldots \cap Z) = P(A) \times P(B) \times P(C) \times \ldots \times P(Z)$

(الأحداث الغير مستقلة)المشروطة : Conditional Probability

حدثان وقوع أحدهما يؤثر في وقوع الآخر مثل سحب ورقة من أوراق اللعب دون إرجاع مما يؤدي لتأثير سحب ورقة جديدة لنقص الفرصة بنقص عدد الأوراق (من 52 إلى 51) فالحدثان A , B نكتب حدث وقوع A بشرط وقوع B بالصورة A / B ويكون

$$P(A / B) = \frac{P(A \cap B)}{P(B)} , \quad P(B) \neq 0$$

OR $P(A \cap B) = P(B) \times P(A / B)$

لاحظ أن العلامة / ليست علامة القسمة بل علامة شرط وقوع ما يليها من أحداث

P(A / B) وهو احتمال وقوع الحدث A بشرط وقوع الحدث B، قد ترد عبارة أخرى تفيد الشرط كالقول علماً بأن ... ,

وفي حالة الحدثان مستقلان أي لا يؤثر وقوع أحدهما على الآخر يصبح القانون : $P(A \cap B) = P(B) \times P(A)$

ملاحظة : في بعض الأحيان يتم استبدال كلمة "و" بـ "∩" ، وهي رمز من مجموعة النظرية التي تشير إلى تقاطع مجموعتين.

قانون جمع الاحتمالات: اذا كان B و A حدثان معرفان على فراغ العينة S فان احتمال وقوع الحدث A او احتمال وقوع الحدث B او احتمال وقوعهما معا او بمعنى اخر وقوع احد الحدثين على الاقل ويرمز له(P(A ∪ B هذا يعني

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

مثال : طائرة بها محركين B و A فاذا احتمال تعطل المحرك A = 0.005 واحتمال تعطل المحرك B = 0.004 واحتمال تعطل المحركين معا 0.001 فما احتمال تعطل احد محركي الطائرة على الاقل .

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B) = 0.005 + 0.004 - 0.001 = 0.009 - 0.001 = 0.008$$

مثال صف يحتوي على 60 طالب، 12/7 من الطلاب يرتدي قميص لونه ابيض، و 3/1 الطلاب يرتدي قميص لونه ازرق، أما باقي الطلاب فيرتدون قمصاناً رصاصي اللون، فإذا تم اختيار طالب بشكل عشوائي من الصف فما هو احتمال أن يكون قميصه رصاصي اللون ؟

الحل: لمعرفة احتمال أن يكون لون قميص الطالب رصاصي يجب أولاً معرفة عدد الطلاب اللذين يرتدون قمصاناً رصاصي اللون، ويمكن إيجادها كما يلي:

عدد الطلاب اللذين يرتدون قمصاناً ابيض = 12/7 × 60 = 35 طالب.

عدد الطلاب اللذين يرتدون قمصاناً ازرق اللون = 3/1 × 60 = 20 طالب.

عدد الطلاب اللذين يرتدون قمصاناً رصاصيا = 60-35-20 = 5 طلاب.

وبالتالي فإن احتمال أن يكون لون قميص الطالب رصاصي = 60/5 = 12/1.

قواعد الاحتمال :

1 - إذا كان A حدث من S أي أنَّ A مجموعة جزئية من S فإن : الرمز P(A) يعبر عن احتمال وقوع الحدث A

$$P(A) = \frac{\text{عدد حالات وقوع الحدث } (A) \text{ بالفعل}}{\text{كل الحالات التي يمكن وقوعها}}$$

حيث ان $0 < P(A) < 1 \;,\; P(S) = 1 \;,\; P(\phi) = 0$

2 - الحدثان المتكاملان (المتتامان) A ∪ $\bar{A}$ = S يكون : $P(A) + P(\bar{A}) = 1$

ويمكن استنتاج $P(\bar{A}) = 1 - P(A)$ **: أو** $P(A) = 1 - P(\bar{A})$

أيضاً نقول أن الحدث $\bar{A}$ هو حدث عدم وقوع A .

3 - مجموع احتمالات الأحداث الشاملة يساوي الواحد الصحيح لأن اتحادها يساوي S

4 - الحدثان المتنافيان A, B أي تقاطعهم ϕ فإن : $P(A \cap B) = 0 \;,\; P(A \cup B) = P(A) + P(B)$

ويمكن تعميم ذلك على أكثر من حدثين متنافيين .

5 - إذا كان A, B حدثان غير متنافيين (متصلين) أو احتمال وقوع أحدهم على الأقل فإن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

عملية الطرح هنا للاحتمال $P(A \cap B)$ لتكراره مرتين عند حساب الاحتمال للجزء المشترك بين A, B حيث يحسب مرة مع A وأخرى مع B

يمكن تعميم القاعدة السابقة لأكثر من حدثين متصلين كالتالي :

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C)$$