الـجمهـوريـة الـجزائـريـة الـديـمقـراطيـة الـشعـبيـة

République Algérienne Démocratique et Populaire

وزارة الـتعلـيم الـعالـي و الـبحث الـعلمـي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جــامــعــة الـشهيـد حمه لـخضر - الـوادي -

Université Echahid Hamma Lakhdar-d'El Oued

Faculté des Sciences Exactes

كـليـة الـعلـوم الـدقـيقـة

قـسم الـريـاضيـات

مطبوعة موجهة لطلبة السنة اولى جذع مشترك

بـــعـنـو ا ن

محاضرات في الاحصاء الوصفي و مدخل الى الاحتمالات

من إعداد: د. تواتي إبراهيم محمد السعيد

الـسنة الـجامعية : 2020/2021

الفــــــــــــــهــــــــــــــرس

مقدمة

حظي **علم الإحصاء** باهتمام كبير من الناحية العلمية و العملية و الأكاديمية في السنوات الأخيرة، كونه وسيلة يستخدمها معظم الناس في حياتهم اليومية، و لما يشكله من أهمية في جميع التخصصات.

يعد **علم الإحصاء** أحد فروع الرياضيات التطبيقية، و هو عبارة عن مجموعة الأساليب العلمية القياسية التي يمكن توظيفها لجمع ووصف و تفسير البيانات و الاستقراء و صنع القرار خاصة في حالة عدم اليقين.

لكن يجب الإشارة أن دور الإحصاء يقتصر على توفير المؤشرات المبدئية التي تساعد الفرد على اتخاذ القرار.

هذه المطبوعة موجهة أساسا إلى طلبة سنة أولى جذع مشترك رياضيات وإعلام آلي حسب التحيين الجديد للبرنامج المقرر ، لذا قمنا بتقسيمها إلى سبعة فصول

تناول الفصل الأول المفاهيم الأساسية للإحصاء ، كما تضمن الفصل الثاني التمثيلات البيانية، فيما تم التطرق إلى مقاييس النزعة المركزية ومقاييس التشتت و مقاييس الشكل في الفصول اللاحقة الثالث و الرابع على الترتيب.

خصص الفصل الخامس للسلاسل الإحصائية ذات متغيرين، في حين يعرض الفصلين السادس و السابع أساسيات نظرية الاحتمالات.

تم عرض الدروس بطريقة مبسطة مع أمثلة بسيطة لترسيخ المفهوم و من ثم تمرين تطبيقي شامل مع الحل في نهاية الوحدة، و تستكمل بتمارين مقترحة.

أخيرا

إذا كان هدف الجامعة هو إعداد الطالب لمواجهة مشكلات الحياة الحقيقية فهذا لا يتأتى إلا بتنمية القدرات الفردية من خلال المثابرة و المصابرة و المبادرة، لذا فان أفضل طريقة لفهم المادة هو المراجعة المنتظمة و المستمرة.

أساسيات علم الإحصاء

نبذة تاريخية عن الإحصاء:

ارتبط مصطلح الإحصاء في أذهان الناس بالتعداد و الإحصاءات التي تجريها الدولة أو الأفراد لعد السكان، متطلبات التمويل و الإدارة أو حصر الثروة. فقد ظهر مصطلح الإحصاء قبل ألفي سنة في الحضارة المصرية عند عد القبائل المنضوية تحت سيطرتها كذا عند الصينيين و البابليين في عد منتجاتهم الزراعية، أما في الحضارة الإسلامية فقد كان الخليفة الراشد عمر الفاروق يقدر عدد المتطوعين في الثغور بعدد أرغفة الخبز المستهلكة، وكان الخليفة عثمان بن عفان –رضي الله عنه- أول من استخدم التدوين لإحصاء المستفيدين من عطايا بيت المال.

سمي علم الإحصاء قديما علم الدولة فقد اعتبر بعض المؤرخين أن كلمة statistique مشتقة من الكلمة اللاتينية state بينما عزاها آخرون إلى الكلمة الإغريقية statika وتعني الحساب و العد. وفي أوربا ترجع أول عملية تعداد مسجلة للأراضي سنة 1086 ببريطانيا، وانطلاقا من القرن 14 أصبح شائعا إنشاء دواوين إحصائيات السكان ، الضرائب و المنتجات الزراعية و الصناعية...اما في القرن 17 فقد برز Fermat ، Pascale وBernoulli في دراسة هندسة العاب الحظ و يعتبر القرن 19 الفترة التي شهدت الانطلاقة العلمية للإحصاء الرياضي على يد Laplace وتابعه علماء آخرون في البحث و تطوير الطرق و الأساليب مثل Pearson ، Markov و Fisher.

وفي وقتنا الحاضر يرتبط تطور الإحصاء الرياضي كباقي فروع الرياضيات بمتطلبات التطبيقات المتعددة، فأصبحت طرق الإحصاء تستخدم بشكل واسع في كل الميادين ويلعب دورا هاما في البحوث التقنية، الاقتصادية، البيولوجية، الزراعية، الطبية والاجتماعية... مما يتطلب دراسة مستفيضة للإحصاء الرياضي ومختلف طرقه و أساليبه فهو يقيم بنية النموذج الإحصائي المعد لدراسة و تحليل التجارب الاحتمالية بملاحظة البيانات الإحصائية.

1- مفاهيم أساسية:

1-1 – **الإحصاء**: هو العلم الذي يهتم بالطرق الرياضية لجمع، وترتيب، وعرض، وتحليل البيانات ومن ثم استخدام المعطيات لاستخلاص نتائج و تعميمها على المجتمع -خاصة في حالات عدم اليقين – في مختلف العلوم من اجل اتخاذ القرارات المناسبة.

1-1-1- **أقسام الإحصاء**: ينقسم الإحصاء إلى قسمين حسب التعريف السابق.

1-1-1-1- **الإحصاء الوصفي :**ويقصد به العلم الذي يختص بجمع البيانات و تبويبها وعرضها في شكل جداول إحصائية أو أشكال بيانية ثم تلخيصها ببعض المؤشرات الإحصائية الدالة على طبيعتها.

-1-1-1-2 **الإحصاء الاستدلالي:** و هو يختص بتعميم النتائج على المجتمع من خلال فحص عينة ممثلة عنه ويشمل نظريتي التقدير و اختبار الفرضيات.

- **التوقع (التنبؤ):** وفيه يتم استخدام نتائج الاستدلال الإحصائي لتوقع سلوك الظاهرة في المستقبل، ومن أهم الأساليب المتبعة في التنبؤ أسلوب الاتجاه العام.

-1-1-2 الطريقة العلمية في البحوث الإحصائية: وتعتمد على النقاط التالية:

- تحديد إشكالية البحث والهدف و فرضيات الدراسة.
- تحديد المجتمع، الإطار، طريقة المعاينة و حجم العينة المناسبة رياضيا.
- جمع، تصنيف و عرض البيانات.
- حساب الوسائط – المؤشرات- كتقديرات لمعالم المجتمع.
- تحليل المعطيات.
- تفسير النتائج و اتخاذ القرار.

-1-1-3 تصميم البحث الإحصائي:

- تحديد الغرض.
- تحديد إمكانية التوقع.
- تحديد إطار البحث.

-1-1-4 مصادر جمع البيانات: بعد تحديد الهدف من الدراسة لا بد من جمع بيانات عن الظاهرة من مصادرها:

-1-1-4-1 **مصدر داخلي:** و تأخذ من السجلات و المستندات و الدراسات السابقة.

-1-1-4-2 **مصدر خارجي:** ويتم تجميع البيانات من خلال مشاريع مشابهة أو من مفردات المجتمع مباشرة عن طريق:

- المقابلة الشخصية وتسجيل الإجابات على الاستبيان أو الاستمارة الخاصة.
- الاتصال الهاتفي، بالبريد العادي أو الالكتروني و وسائل التواصل الاجتماعي .
- المراقبة ، الملاحظة و المشاهدة أو التسجيل.
- التجربة أو العمل الميداني.

-1-1-5 طرق جمع البيانات: يوجد أسلوبين لجمع المعلومات:

-1-1-5-1 **المسح الشامل:** تعني دراسة شاملة لجميع أفراد المجتمع

-1-1-5-2 **العينة:** هي عبارة عن مجموعة من مفردات المجتمع يتم اختيارها لتكون ممثلة جيدا له.

المعاينة: وهي علم و فن التحكم و قياس دقة المعلومات الإحصائية باستخدام النظريات الرياضية ويفضل إستخدام أسلوب العينات لأسباب عديدة أهمها:

- تقليل الجهد و التكلفة و الوقت.
- صعوبة حصر أفراد المجتمع
- محدودية الإمكانيات كنقص الخبرة و التخصص والإمكانيات اللوجستية.

-1-1-6 أنواع العينات: يعتبر اختيار العينة من أهم خطوات البحث الإحصائي ويمكن أن تكون مستمرة – ثابتة- أو متغيرة، و يمكن تمييز نوعين من العينات:

-1-1-6-1 **العينات العشوائية:** وتسمى أيضا العينات الاحتمالية، ويتم اختيار مفردات العينة بحيث يكون لكل مفردة نفس الحظ في الظهور . ومن أهم أنواعها :

- **العينة العشوائية البسيطة:** وتستخدم في حالة تماثل أو تجانس أفراد المجتمع، ويتم سحب المفردات باستخدام جداول الأرقام العشوائية أو بالقرعة بإرجاع أو دون إرجاع. فإذا أردنا اختيار n وحدة من N وحدة من المجتمع فإن عدد العينات التي يمكن اختيارها من المجتمع يساوي C_N^n.
- **العينة العشوائية المنتظمة:** و تستخدم في حالة صعوبة الحصول على إطار للمجتمع أو عدم جودته. وعادة يتم اختيار أول وحدة عشوائيا ، ثم نحدد باقي الوحدات المتتالية ذي بعد متساوي قدره k، أي $r,r+k,r+2k,\dots,r+(n-1)k$ حيث $(r\leq k\geq 1)$ **و** $k\leq N/n$.
- **العينة العشوائية الطبقية:** وهي أكثر الطرق استعمالا، وتقوم على تقسيم المجتمع إلى طبقات – مجموعات- متجانسة غير متداخلة، ويتم اختيار وحدات المعاينة عشوائيا من كل طبقة، ومن ميزاتها الحصول على تقدير أكثر دقة و أقل تباينا. أي $N = \sum_{i=1}^{L} N_i \,, n = \sum_{i=1}^{L} n_i \,, W_i = \frac{N_i}{N}$
- **العينة العشوائية العنقودية:** و تعتمد على تقسيم المجتمع إلى قطاعات – عناقيد- متشابهة غير متداخلة، ثم اختيار عدد من هذه العناقيد كعينة عشوائية بسيطة يتم دراسة جميع أفراد كل منها، يمكن ان تكون العينة العنقودية مكونة من مرحلة واحدة أو من عدة مراحل.

2-6-1-1- **العينات غير العشوائية:** وهي التي يتم إختيار مفرداتها بطريقة غير عشوائية، بحيث تختار بالطريقة التي تحقق الهدف، ومن أهم أنواع العينات غير الإحتمالية : **العينة العمدية و العينة الحصصية.**

2-1- مصطلحات إحصائية :

1-2-1- المجتمع الإحصائي: وهو المجموعة التي تشكل أهمية للدراسة ويكون إما محدودا أو غير محدود أو مجموعة المشاهدات التي تخص ظاهرة وينقسم إلى نوعين: المجتمع الهدف و مجتمع الدراسة، مثلا : مجموعة طلبة جامعة الوادي.

1-2-2- العينة: وهي كل مجموعة جزئية غير خالية من المجتمع الإحصائي، مثلا : 100 طالب من طلبة الجامعة.

1-2-3- الوحدة الإحصائية: وهو كل عنصر – فرد أو مفردة – من المجتمع الإحصائي، مثلا: طالب واحد من الجامعة.

1-2-4- الطبع الإحصائي: و تسمى الميزة الإحصائية أو الظاهرة الإحصائية، وهي الخاصية أو الصفة المدروسة أو الملاحظة على أفراد المجتمع، وهي نوعين:

1-4-2-1- **ميزة نوعية:** وهي الصفة غير القابلة للقياس وتنقسم إلى نوعين: **ترتيبية** مثل مستوى الطالب و **غير ترتيبية** مثل الجنس ، اللون، الحالة الاجتماعية ... .

2-4-2-1- **ميزة كمية:** وهي الصفة القابلة للقياس مثل العمر، الوزن ... .

1-2-5- المتغير الإحصائي: نسمي قيم ظاهرة إحصائية متغيرا إحصائيا ويصنف إلى قسمين:

1-5-2-1- **المتغير الإحصائي المتقطع** :– المنفصل – وهو الذي يأخذ قيما معزولة مثل عدد الأخوة... .

2-5-2-1- **المتغير الإحصائي المستمر** : وهو يأخذ قيم في مجال$[a,b[$، يسمى كل مجال من الشكل$[a_k,a_{k+1}[$فئة.

نسمي العدد $L=a_{k+1} - a_k$ طول أو سعة –مدى- الفئة و العدد $c=(a_{k+1} + a_k)/ 2$ مركز الفئة.

1-2-6- السلسلة الإحصائية: نسمي نتائج ظاهرة إحصائية معينة بالسلسلة الإحصائية أي عبارة عن مجموعة الثنائيات (x_i,n_i) حيث x_i القيم و n_i عدد المرات – التكرارات، المشاهدات – التي تكرر فيها x_i .

1-2-7- التكرار النسبي: هو نسبة عدد التكرارات لقيم الطبع الإحصائي إلى العدد الإجمالي للتكرارات ونرمز له بالرمز f_i أي أن:

$$f_i = \frac{n_i}{N}; N = \sum_i n_i$$

ملاحظة: $\sum_i f_i = 1$ □ $f_i \in [0,1]$

1-2-8- الجدول التكراري: وهي جداول تنظم و ترتب قيم المتغير الإحصائي، و تكون من الشكل:

القيم x_i	x_1	...					
التكرار المطلق n_i	n_1	...					
النسبة المئوية f_i 100	%						

و نميز بين الجداول التكرارية البسيطة ،جداول التوزيع التكرارية المزدوجة و الجداول التكرارية المفتوحة.

1-2-9- التكرار المجمع: وهو عدد المشاهدات التي تقل أو تزيد عن قيمة معينة، ويمكن التفرقة بين نوعين:

1-2-9-1- **التكرار المتجمع الصاعد:** هو مجموع التكرارات للقيمة وللقيم الأصغر منها ونرمز له بـ n_i

- هو مجموع عدد الوحدات التي تقل عن الحد الأعلى للفئة المقابلة.

$$n_k = \sum_{i=1}^{k} n_i$$

1-2-9-2- **التكرار المتجمع النازل:** هو مجموع التكرارات للقيمة وللقيم الأكبر منها ونرمز له بـ n_i

- هو مجموع عدد الوحدات التي تزيد عن الحد الأدنى للفئة المقابلة.

مثال: مصنع نسيج عند تجريب آلة جديدة قام بإحصاء الأخطاء في *75* عينة قماش من *10* متر ، فكانت النتائج الآتية :

x_i	*0*	*1*	*2*	*3*	*4*	*5*
n_i	*38*	*15*	*11*	*6*	*3*	*2*

1- حدد المجتمع الإحصائي ، المتغير الإحصائي.

2- حدد التكرار النسبي ، التكرار المتجمع الصاعد**.**

الحل:

1- **المجتمع الإحصائي** قماش المصنع ذو 10 متر ، **الميزة الإحصائية** عدد الأخطاء في القطعة كمي منفصل.

2-

x_i	*0*	*1*	*2*	*3*	*4*	*5*
n_i	*38*	*15*	*11*	*6*	*3*	*2*
$\%f_i$	*50.6*	*20*	*14.7*	*8*	*4*	*2.7*
$\nwarrow n_i$	*38*	*53*	*64*	*70*	*73*	*75*

1-3- تبويب البيانات في جدول توزيع تكراري متصل: إن القاعدة العامة لإفراغ البيانات في جدول كمي مستمر هي أن يكون عدد الفئات N_s يتراوح بين 5 **و 15** ، إلا أن هذا يعود للإحصائي في تحديد العدد و طول الفئة، رغم وجود طرق تحدد ذلك منها:

1-3-1- القاعدة البسيطة: $N_s = \sqrt{N}$ **و** $L = \frac{x_{max}-x_{min}}{N_s}$ حيث N عدد القيم و L طول الفئة.

2-3-1- قاعدة ستـــــــورج: $N_s = 1 + 3.322log(N)$

3-3-1- قاعدة يـــــــــول: $N_s = 2.25\sqrt[4]{N}$

مثال: سجلت مدة نقل الطلبة من مقر سكناهم بالدقائق فكانت:

5-3-20-45-45-5-25-15-15-5-20-5-5-2-45-10-30-2-7-15-25-30-60-35-30-45-30-30-25-25-45-5-70-60-35-35-18-5-5-65-10-50-25-30-2-28-30-25-20-45.

1- شكل القيم في جدول إحصائي بفئات-اعتمد طريقة ستورج-

الحل: نحسب: $N_s = 1 + 3.322log50 \simeq 7$ **وعليه** $L = \frac{70-2}{7} \cong 10$ **فتكون الفئات على الشكل:**

التمثيـــــــــــلات البيانيـــــــــــــــــــــــة

تمـــــــــــهيـــــــــــد: إن الهدف الأساسي لعرض المعلومات بواسطة المعطيات تسمح باستيعاب مجموعة العناصر المعروضة بشكل سريع و إظهار بعض الخصائص الهامة، رغم إهماله بالضرورة بعض الجزئيات.و الرسوم البيانية وسيلة لمراقبة التغيرات الشاذة و الانكسارات و التقطعات في مسار ظاهرة ما.

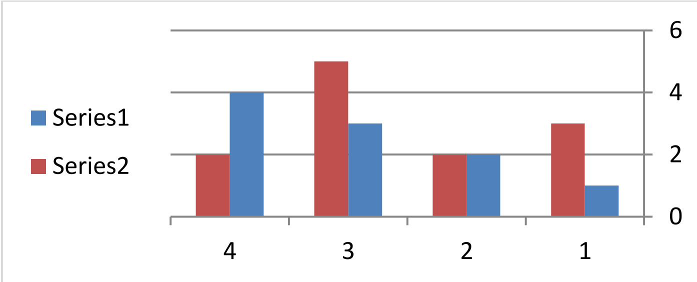
يقوم المبدأ الأساسي للتمثيل البياني على استخدام الإحداثيات الديكارتية، وهي بصورة عامة المستطيلات التي تسمح في علم الجبر بتمثيل التوابع على مستقيمين يتقاطعان في O و مرسومين على سطح مستو، و يختار لذلك اتجاه موجب و وحدة بيانية، وتمثل القيمة x_i للصفة على المحور الأفقي.

1-2- التمثيل البياني للتوزيعات التكرارية: يمكن تمثيل التوزيعات التكرارية بطرق مختلفة أهمها:

1-2- الأعمدة البيانية : ويسمى المخطط بالأعمدة، و نميز الأشكال التالية:

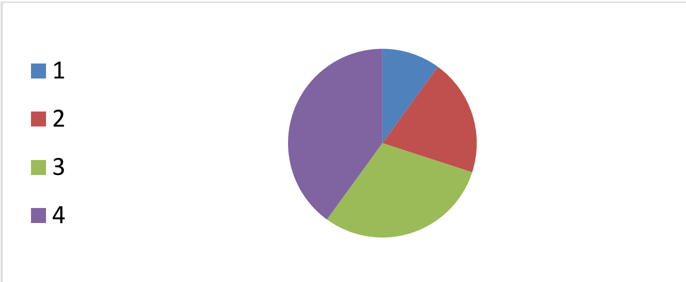
1-1-2- الأعمدة البسيطة: وهي أعمدة يتناسب طولها مع قيمة البيانات التي تمثلها.

2-1-2- الأعمدة المزدوجة: وتستخدم للمقارنة بين ظاهرتين أو أكثر

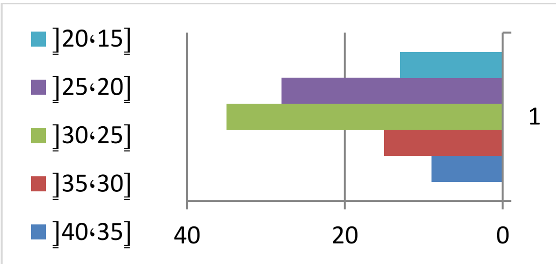


-2-2 الدائرة النسبية البيانية: و تستعمل خاصة في المتغيرات النوعية و يتم تخصيص لكل صفة قطاع، تحسب زاوية القطاع α كمايلي:

$$\alpha = f_i \times 2\pi$$



-2-3 المدرج التكراري: عندما يكون المتغير المدروس مستمرا فإن كل قيمة م مدى السلسلة يمكن أن يكون قيمة للصفة، لذا نرفق لكل فئة مجال على محور الفواصل ونمثل كل ثنائية (فئة،تكرار) بمستطيل مساحته تتناسب مع تكرار الفئة.



-2-3-1 تعترض مشكلة رسم المدرج في حالة أطوال الفئات غير متساوية ، وفي هذه الحالة نمثل الأطوال بالتكرار المعدل n'_i وفق القاعدة التالية:

$$n_\iota = \frac{\acute{l} \times n_\iota}{l_\iota}$$

حيث l أصغر طول فئة .

مثال : أجريت تجربة على *100* مصباح لمعرفة مدى صلاحيتها فكانت النتائج التالية:

الفئات			
n_i	15	45	40
l_i	100	300	200
$\acute{n}_\iota$	15	15	20

-2-3-2 أشكال التوزيعات و المدرجات التكرارية: نصادف في الحياة العملية عدة أشكال أساسية بسيطة من التوزيعات منها:

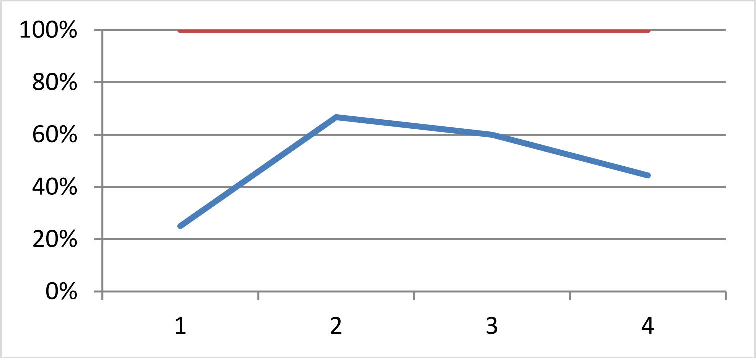
-2-3-2-1 التوزيعات المتناظرة: تتناقص فيها التكرارات نحو الصفر بشكل نظامي قبل و بعد القيمة المركزية الأعلى.

-2-3-2-2 التوزيعات غير المتناظرة

-2-3-2-3 التوزيعات على شكل حرف U, L ...

2-4- المنحنيات التكرارية: ويتم بتحديد نقطة لكل مشاهدة ومن ثم نصل هذه النقاط باليد أي تعليم النقاط (x_i,n_i)- x_i القيمة أو مركز الفئة -

إذا أوصلنا بين النقاط بالمسطرة تشكل لدينا المضلع التكراري.



2-4-1- المنحنيات التكرارية التجميعية: ويتم رسم المنحنى التكراري المتجمع الصاعد – النازل – لمتغير مستمر بإيصال مجموعة النقط

التي إحداثياتها الحدود العليا للفئات – الحدود الدنيا – والتكرارات المتجمعة الصاعدة المقابلة لها – التكرارات النازلة – على الترتيب .

2-4-2- الدرج التكراري: ويستخدم لرسم المنحنى التكراري المتجمع الصاعد – النازل – لمتغير منفصل .

مقاييس النزعة المركزية – مقاييس الموقع-

تــــمهيـــــــــد: من المهام الأساسية للإحصاء الوصفي اختصار القيم الكثيرة إلى بضع معلومات تعبر إلى أبعد حد عن السلسلة الإحصائية المدروسة، فالتمثيل البياني يزودنا بالملخص الأول، ونرغب الآن في تمييز السلسلة الإحصائية بعدد أو قيمة معيارية تمثل بدرجة كبيرة مجموعة عناصر السلسلة .

لذا يجب أن تحقق هذه القيمة بعض الشروط حددها ***يول*** في:

- أن تعين بشكل موضوعي و مستقل عن الرأي الشخصي للإحصائي.
- أن تتعلق بكل معطيات السلسلة. – أن تكون سهلة الحساب.
- أن تكون قليلة التأثر بتقلبات العينة - أن يكون لها معنى مادي.

مقاييس النزعة المركزية هي المؤشرات التي تعبر عن القيمة التي تتمركز حولها اغلب القياسات المعطاة منها:

3-1- المتوسط الحسابي: هو أكثر المقاييس إستخداما في الجانب التطبيقي، ويعد مركز التوازن للظاهرة.

3-1-1- حساب الوسط الحسابي :

3-1-1-1- الطريقة المباشرة:

أ- **في حالة البيانات غير المبوبة:** يعرف المتوسط الحسابي لسلسلة من n مفردة بالعدد الذي نرمز له بـ $\overline{X}$ الذي يحقق:

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

ب- **في حالة البيانات المبوبة:** في اغلب الحالات لا تكون للقيم نفس الأهمية بل لكل قيمة عامل ترجيح خاص، عندئذ يسمى المتوسط الحسابي بالمتوسط الحسابي المرجع و يحسب كمايلي:

-المتغير الإحصائي منفصل: إذا كانت x_1 ،x_2 ،...،x_n قيم الظاهرة X وكانت n_1 ،n_2 ،...،n_n تكرارتها على الترتيب فإن:

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

-المتغير الإحصائي مستمر: إذا كانت c_1 ،c_2 ،...،c_n مركز فئات الظاهرة X وكانت n_1 ،n_2 ،...،n_n تكرارتها على الترتيب فإن:

$$\overline{X} = \frac{\sum_{i=1}^{n} n_i c_i}{\sum_{i=1}^{n} n_i}$$

مثال: تحصل طالب على العلامات التالية 8-11-14-15-16 بمعاملات 3-1-1-5-2 على الترتيب:

$$\overline{X} = \frac{8 \times 3 + 11 \times 1 + 14 \times 1 + 15 \times 5 + 16 \times 2}{3 + 1 + 1 + 5 + 2} = 13$$

3-1-1-2- الطريقة غير المباشرة: وتسمى طريقة الوسط الفرضي، وتعتمد على استبدال المتغير الأساسي بمتغير جديد

يفضل اختيار الوسط*الفرضي* x_0 القيمة الوسيطية ويستحسن أن يقابلها أكبر تكرار.ففي حالة المتغير الإحصائي المتقطع:

$$\overline{X} = x_0 + \frac{\sum_{i=1}^{n} n_i(x_i - x_0)}{\sum_{i=1}^{n} n_i}$$

- إما في حالة المتغير المستمر: $\overline{X} = x_0 + a\frac{\sum_{i=1}^{n} n_i\frac{(c_i - x_0)}{a}}{\sum_{i=1}^{n} n_i}$ حيث: x_0 قيمة وسيطية و a طول الفئة

مثال:أجر 25 عاملا في مصنع تبعا للأجر في الساعة

الفئات	n_i	c_i	$n_i \times c_i$	$c_i - x_0$	z_i	$n_i \times z_i$
	5	3	15	-2	-1	-5
	8	5	40	0	0	0
	12	7	84	2	1	12

$$z_i = \frac{c_i - 5}{2}$$

ومنه نجد: $\overline{X} = \frac{139}{25}; \overline{Z} = \frac{7}{25}$ اذن: $\overline{X} = 5 + 2\frac{7}{25} = \frac{139}{25}$

3-1-2- مزايا وعيوب الوسط الحسابي:

- يتعلق بكل قيم السلسلة وسهل الحساب.
- يخضع للحسابات الجبرية.
- يتأثر بالقيم المتطرفة ولا يمكن تعيينه بيانيا.
- العمود المقام من نقطة الوسط يقسم المدرج التكراري إلى قسمين متساويين في المساحة.
- يصعب حسابه من الجداول التكرارية المفتوحة ومن البيانات الوصفية.

3-1-3- خواص الوسط الحسابي:

- الوسط الحسابي هو القيمة التي تعوض كل القيم ويبقى المجموع على حاله، أي: $\sum_{i=1}^{n} x_i = n\frac{\sum_{i=1}^{n} x_i}{n} = n\overline{X} = \sum_{i=1}^{n}\overline{X}$ يمكن كتابتها بالشكل $\sum_{i=1}^{n}(x_i - \overline{X}) = 0$
- الوسط الحسابي هو القيمة التي يكون مجموع مربعات الانحرافات عنها أقل ما يمكن،أي.

$$\sum_{i=1}^{n}(x_i - \overline{X})^2 = \sum_{i=1}^{n}(x_i - \overline{X} + x_0 - x_0)^2 = \sum_{i=1}^{n}(x_i - x_0)^2 - n(x_0 - \overline{X})$$

- إذا كان X متغيرا إحصائيا يأخذ القيم x_1،...،x_k وكان Yمتغيرا إحصائيا **حيث:** $y_i = ax_i + b$

فان: $\overline{Y} = a\overline{X} + b$ لأن: $\overline{Y} = \frac{\sum_1^n y_i}{n} = \frac{\sum_1^n(ax_i + b)}{n} = \frac{a\sum_1^n x_i}{n} + \frac{\sum_1^n b}{n} = a\overline{X} + b$

3-2- الوسط الهندسي: وهو مقياس شائع الاستعمال في الدراسات الاقتصادية

3-2-1- حساب المتوسط الهندسي:

أ- **في حالة البيانات غير المبوبة:** الوسط الهندسي لـ n قيمة موجبة والذي نرمز له بـ G العدد المعرف كمايأتي:

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

ب- في حالة البيانات المبوبة:1- الوسط الهندسي G للقيم x_n،...،x_1 مرفقة بالتكرارات n_n،...،n_1 على الترتيب:

$$G = \sqrt[\sum_1^n n_i]{x_1^{n_1} \times x_2^{n_2} \times \dots \times x_n^{n_n}}$$

ملاحظة : لحساب المتوسط الهندسي يمكن اللجوء إلى اللوغاريتم، ثم ندخل الأسية فنجد مثلا**:**

$$G = e^{\frac{\sum_{i=1}^n n_i ln x_i}{\sum_{i=1}^n n_i}}$$

2 - إذا كانت c_1 ،c_2 ،...،c_n مركز فئات الظاهرة X وكانت n_1 ،n_2 ،...،n_n تكرارتها على الترتيب فإن:

$$G = \sqrt[\sum_1^n n_i]{c_1^{n_1} \times c_2^{n_2} \times \dots \times c_n^{n_n}}$$

2-2-3- مزايا و عيوب المتوسط الهندسي: - يدخل في حسابه جميع القيم

- أقل تأثرا بالقيم المتطرفة من المتوسط الحسابي
- يفضل استخدامه في وصف الظواهر النسبية و حساب المعدلات.
- لا يمكن حسابه من الجداول المفتوحة.
- ليس له معنى إذا كانت احد القيم سالبة أو معدومة.

3-3- الوسط التوافقي: وهو مقياس خاص يستعمل في تحديد معدلات السرعة،...

1-2-3- حساب المتوسط التوافقي:

أ- في حالة البيانات غير المبوبة: الوسط التوافقي لـ n قيمة والذي نرمز له بـ H العدد المعرف كمايأتي:

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

ب- في حالة البيانات المبوبة:1- الوسط التوافقي H للقيم x_n،...،x_1 مرفقة بالتكرارات n_n،...،n_1 على الترتيب:

$$H = \frac{\sum_1^n n_i}{\sum_{i=1}^n \frac{n_i}{x_i}}$$

2 - إذا كانت c_1 ،c_2 ،...،c_n مركز فئات الظاهرة X وكانت n_1 ،n_2 ،...،n_n تكرارتها على الترتيب فإن:

$$H = \frac{\sum_1^n n_i}{\sum_{i=1}^n \frac{n_i}{c_i}}$$

2-3-3- مزايا و عيوب المتوسط التوافقي: - يدخل في حسابه جميع القيم

- أقل تأثرا بالقيم المتطرفة من المتوسط الحسابي
- يفضل استخدامه في حساب معدلات السرعة و الأسعار.
- لا يمكن حسابه من الجداول المفتوحة.
- لا يمكن حسابه إذا كانت إحدى القيم معدومة.

ملاحظة : $H < G < \overline{X}$

3-4- الوسط التربيعي: $Q = \sqrt{\frac{\sum_{i=1}^{n} n_i x_i^2}{\sum_{i=1}^{n} n_i}}$

3-5- الوسيط: وسيط سلسلة إحصائية ما والذي نرمز له بالرمز Med القيمة التي من أجلها يتساوى عدد القيم الأصغر منها مع القيم الأكبر منها والمرتبة تصاعديا أو تنازليا، أي التكرار النسبي للقيم الأصغر أو الأكبر من الوسيط يساوي على الأكثر 0.5 .

3-5-1- حساب الوسيط: يتطلب حساب الوسيط ترتيب حدود السلسلة ترتيبا تصاعديا – التكرار المتجمع الصاعد –

3-5-1-1- المتغير متقطع:

أ- إذا كانت عدد القيم N– التكرارات – فرديا فان: $Med = x_{\frac{N+1}{2}}$

ب- إذا كانت عدد القيم N– التكرارات – زوجيا فان: $Med = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}$

مثال 3. 5. 1: عين وسيط السلسلة التالية : 10-11-12-13-14-15

لدينا عدد القيم زوجي =6 و منه $Med = \frac{x_{\frac{6}{2}} + x_{\frac{6}{2}+1}}{2} = \frac{12+13}{2} = 12.5$

3-5-1-2- المتغير مستمر: يمكن باستعمال المضلع التكراري المتجمع الصاعد و بتطبيق نظرية طاليس ايجاد العلاقة التالية:

$$Med = a + \frac{\sum n_i}{2} - \frac{N_{i-1}^{+}}{n_e} l$$

حيث : a الحد الأدنى للفئة الوسيطية – تتحدد بناءا على رتبة الوسيط $\left(r = \frac{\sum n_i}{2}\right)$و التي يقابلها في التكرار المتجمع الصاعد –

N_{i-1}^{+} التكرار المتجمع الصاعد قبل الفئة الوسيطية . n_e تكرار الفئة الوسيطية . l طول فئة الوسيطية.

3-5-2- تحديد الوسيط بيانيا: الوسيط بيانيا هو فاصلة نقطة تقاطع المنحنى التكراري المتجمع الصاعد مع النازل، كما يمكن إيجاده من أحدهما فقط بتعيين فاصلة نقطة تقاطع المنحنى مع $y = \frac{N}{2}$.

مثال 3. 5. 2: انشئ المدرج يرغب صاحب مصنع معرفة رغبات زبائنه فأجرى سبرا للآراء فتحصل على معلومات موزعة حسب السن كماياتي:

العمر	[15□20[	[20□25[	[25□30[	[30□35[	[35□40[
f%	13	28	35	15	9

1- انشئ المنحنى للتكراري النسبي المجمع الصاعد ثم استنتج الوسيط.
2- عين الوسيط حسابيا.

الحل: 1-

2-الفئة الوسيطية ومنه: $Med = 25 + \frac{50-41}{35} \times 5 = 26.286$

2-5-3- خواص الوسيط: - لا يتأثر بالقيم الشاذة أو المتطرفة و لا يعتمد إلا على قيمة أو قيمتين.

- سهل الحساب إلا انه يصعب حسابه من البيانات الوصفية .
- مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيمة أخرى، أي:

$$\sum |x - med| \leq \sum |x - a|; a \neq med$$

2-6- المنوال: وهو مقياس يستعمل لمعرفة النمط الشائع، و يكثر استخدامه في البيانات النوعية، ونرمز له بالرمز Mod

2-6-1- حساب المنوال:

2-6-1-1- في حالة المتغير المتقطع: المنوال هو القيمة التي يقابلها أكبر تكرار.

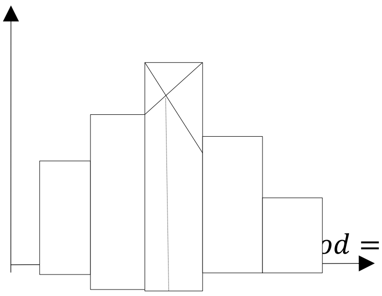
مثال 3. 6. 1: منوال السلسلة 7-8-8-8-9-9 هو 8

2-6-1-2- في حالة المتغير المستمر: باستعمال المدرج التكراري، وبتنفيذ طريقة الرافعة أو الفروق لبيرسون و بتطبيق نظرية طاليس ، نجد العلاقة التالية:

$$Mod = a + \frac{\Delta_1}{\Delta_1 + \Delta_2} l$$

حيث : a الحد الأدنى للفئة المنوالية – الفئة الأكبر تكرارا –، Δ_1 الفرق بين تكرار الفئة المنوالية و الفئة السابقة لها، Δ_2 الفرق بين تكرار الفئة المنوالية و الفئة الموالية لها ، l طول فئة المنوالية.

2-6-1-3- حساب المنوال بيانيا: و يتم برسم المدرج التكراري للتوزيع، ثم نصل بداية المستطيل للفئة المنوالية ببداية المستطيل للفئة اللاحقة، ونقوم بإيصال نهاية المستطيل للفئة المنوالية بنهاية المستطيل للفئة السابقة، المنوال هو فاصلة نقطة التقاطع.



مثال 3. 6. 2: أحسب منوال المثال 3. 5. 2 جبريا و بيانيا

الحل : الفئة المنوالية و منه $Mod = 25 + \frac{7}{7+20} \times 5 = 26.296$

ملاحظة: في التوزيعات غير المتناظرة بشكل معتدل ووحيدة المنوال لدينا: $Mod = 3(\bar{X} - Med)$

2-6-2- خواص المنوال: - أسهل المقاييس حسابا ولايتأثر بالقيم المتطرفة

- يمكن حسابه بيانيا.
- غير قابل للعمليات الجبرية.
- يمكن حسابه من الجداول المفتوحة.

3-7- الربيعيات: يمكن تقسيم البيانات المرتبة إلى أربعة أقسام متساوية – مثلا نحتاج 50 % من العناصر المركزية –

3-7-1- الربعي الأول: و نرمز له Q_1 هو أصغر قيمة للطبع الإحصائي بحيث يكون 25 % من الحدود المرتبة على الأقل لها قيمة أصغر أو

تساوي Q_1.

3-7-2- الربعي الثالث: و نرمز له Q_3 هو أصغر قيمة للطبع الإحصائي بحيث يكون 75 % من الحدود المرتبة على الأقل لها قيمة أصغر أو تساوي Q_3.

3-7-3- حساب الربيعيات: نرتب القيم تصاعديا – التكرار المتجمع الصاعد –

3-7-3-1- في حالة متغير متقطع: (i=1،3) Q_i هو القيمة التي رتبتها n حيث n أصغر عدد طبيعي يحقق: $n \geq \frac{}{4}$.

3-7-3-2- في حالة متغير مستمر: يمكن باستعمال المضلع التكراري المتجمع الصاعد و بتطبيق نظرية طاليس إيجاد العلاقة التالية:

$$Q_i = a + \frac{i\sum n_i}{4} - \frac{N_{i-1}^{+}}{n_Q} l$$

حيث : a الحد الأدنى للفئة الربيعية – تتحدد بناءا على رتبة الربيع $\left(r = i\frac{\sum n_i}{4}\right)$و ما يقابلها في التكرار المتجمع الصاعد –

N_{i-1}^{+} التكرار المتجمع الصاعد قبل الفئة الربيعية . n_Q تكرار الفئة الربيعية . l طول فئة الربيعية.

3-7-3-3- حساب الربيعيات بيانيا: (i=1،3) Q_i هي فاصلة النقطة من منحنى التواترات المجمعة الصاعدة التي ترتيبها$\frac{i}{4}$.

مثال3-7-1- عين الربيع الأول للسلسلة 3-3-4-4-4-4-5-7-7-8-9.

لدينا 11/4=2.75 ومنه Q_1=4 =x_3.

مثال 3-7-2- عين الربيع الثالث للمثال 3-5-2-

لدينا الفئة الربيعية ومنه: $Q_3 = 25 + \frac{75-41}{35} \times 5 = 29.857$.

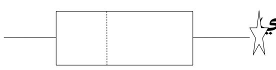
ملاحظة: نجد في بعض المراجع الطريقة التالية لتعيين الربيعيات لمتغير منفصل: نحدد رتبة الربيع $r = (N+1)\frac{i}{4}$ نميز حالتين:

اذا كان r عددا صحيحا فان $Q_i = x_{(r)}$ إما إذا كان r عددا كسريا فان: $x_{(l)} < Q_i < x_{(u)}$ و $Q_i = x_{(l)} + (r-l)\left(x_{(u)} - x_{(l)}\right)$

3-8-العشريات و الميئينيات: بنفس طريقة الربيعيات نعرف العشريات (i=1،...،9) D_i و المئينيات i=1،...،99) C_i (– تقسيم السلسلة إلى 10 أقسام ، 100 قسم متساو على الترتيب-.

3-9-المخطط بالعلبة: نرسم مخططا بالعلب حسب الطريقة التالية:

- نضع قيم الطبع على محور أفقي أو شاقولي.
- نعين على هذا المحور القيم $min; max; Q_1; Med; Q_3$
- نمثل القيم الشاذة بنجمة حيث: حدها الأدنى $Q_1 - \frac{Q_3-Q_1}{2}$ و حدها الأعلى $Q_3 + \frac{Q_3-Q_1}{2}$
- نكون مستطيلا بالتوازي مع المحور طوله Q_3-Q_1 وعرضه كيفي



مقاييس التشتت

تــــــمهيــــد:إن مقاييس النزعة المركزية لا تسمح لنا بالوصف الكامل لسلسلة إحصائية ما إذ لا تعطينا أي إرشاد حول توزيع الصفات داخل السلسلة.

فتشتت بيانات ظاهرة هي درجة تقارب أو تباعد البيانات عن بعضها البعض أو عن مقياس النزعة المركزية.

4-1- المقاييس الأولية للتشتت:

4-1-1- المدى: وهو طول سلسلة ونرمز له E و يساوي الفرق بين أكبر قيمة و أصغر قيمة للمتغير المدروس، أي:

$$E = x_{max} - x_{min}$$

أما في حالة البيانات المبوبة $E = c_F - c_p$. المدى سهل الحساب و يستعمل في مراقبة الجودة و المناخ و يتأثر بالقيم الشاذة.

4-1-2- المدى الربيعي:ونرمز له بـ W وهو $W = Q_3 - Q_1$، والأكثر استعمالا الإنحراف الربيعي L_Q ويساوي $\frac{W}{2}$

أما الانحراف الربيعي النسبي فقيمته $\frac{W}{Med}$.وهو مقياس لا يتأثر بالقيم المتطرفة و يمكن حسابه من الجداول المفتوحة،و يستخدم كمقياس للتشتت في التوزيعات شديدة الالتواء.

4-1-3- الإنحراف المتوسط: هو الوسط الحسابي لإنحرافات القيم عن وسطها الحسابي، ويتميز بصعوبة حسابه، ويعبر عن شدة تجمع البيانات حول بعضها و مدى تجانسها حول قيمة مركزية.

أ- **في حالة البيانات غير المبوبة:** متوسط الانحراف لـ n قيمة والذي نرمز له بـ e العدد المعرف كمايأتي:

$$e=\frac{\sum_{i=1}^{n}|x_i-\overline{X}|}{n}$$

ب- في حالة البيانات المبوبة:1- المتوسط المطلق للانحراف e للقيم x_n،...،x_1 مرفقة بالتكرارات n_n،...،n_1 على الترتيب هو:

$$e=\frac{\sum_{i=1}^{n}n_i|x_i-\overline{X}|}{\sum_{i=1}^{n}n_i}$$

2 - إذا كانت c_n،...،c_2 ،c_1 مركز فئات الظاهرة X وكانت n_n،...،n_2 ،n_1 تكرارتها على الترتيب فإن:

$$e=\frac{\sum_{i=1}^{n}n_i|c_i-\overline{X}|}{\sum_{i=1}^{n}n_i}$$

مثال: نعتبر السلسلة 12-16-11-14-17 متوسطها 14 وانحرافاتها المطلقة هي 2-2-3-0-3 ويكون e=2

ملاحظة : الانحراف المتوسط عن الوسيط أقل من الانحراف المتوسط عن الوسط الحسابي، أي **:** $e_{Med}<e_{\overline{X}}$

4-2- التبـايــن: وهو أكثر مقاييس التشتت تطبيقا، يرمز له σ^2 أو Var ويعبر عن متوسط مربعات انحراف المعطيات عن المتوسط الحسابي للسلسلة، أي أن: $\sigma^2=\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$

- **في حالة بيانات مبوبة:** $\sigma^2=\frac{\sum_{i=1}^{n}n_i(x_i-\mu)^2}{\sum_{i=1}^{n}n_i}$
- **في حالة العينات الصغيرة:** $s^2=\frac{\sum_{i=1}^{n}n_i(x_i-\overline{x})^2}{\sum_{i=1}^{n}n_i}$ وهو التقدير غير المتحيز لتباين المجتمع.

مثال: تم سحب عينة من طلبة سنة أولى و سجل عدد سنوات التمدرس النظامية فكانت النتائج التالية : 12-13-14-15-16

لحساب تباين السلسلة، نقوم بحساب الوسط الحسابي أولا فنجد: $\overline{x}=\frac{70}{5}=14$، ثم نحسب $(x_i-\overline{x})^2$ فنجد: *4-1-0-1-4*

و عليه: $s^2=\frac{4+1+0+1+4}{4}=2.5$

4-2-1- طرق أخرى لحساب التباين:

4-2-1-1- قاعدة كوينغ: - في حالة البيانات غير المبوبة: $\sigma^2=\frac{\sum_{i=1}^{n}x_i^2}{n}-\mu^2$

- في حالـة البيانـات المبـوبـة: $\sigma^2=\frac{\sum_{i=1}^{n}n_ix_i^2}{\sum_{i=1}^{n}n_i}-\mu^2$

- في حالة العينات : $s^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\overline{x}^2\right)$

برهان: $\sigma^2=\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}=\frac{\sum_{i=1}^{n}(x_i^2-2x_i\mu+\mu^2)}{n}=\frac{\sum_{i=1}^{n}x_i^2}{n}-2\frac{\sum_{i=1}^{n}x_i}{n}\mu+\mu^2=\frac{\sum_{i=1}^{n}x_i^2}{n}-\mu^2$

-4-2-1-2 $\sigma^2 = \sum_{i=1}^{n} x_i^2 - n\mu^2$

-4-2-1-3 الوسط الفرضي: $\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - x_0)^2}{n} - (\mu - x_0)^2$

-4-2-3 الانحراف المعياري: و هو الجذر التربيعي للتباين، و يرمز له ◷ أي$\sqrt{V}$◷

مثال: الانحراف المعياري للسلسلة 1-2-3-4-5 هو: $\sigma = \sqrt{\frac{1^2+2^2+3^2+4^2+5^2}{5} - 3^2}$

-4-2-3-1 مزايا و عيوب الانحراف المعياري: - قيمته صغيرة و يتعامل جميع القيم

- أكثر مقاييس التشتت استخداما وسهل الحساب.
- يتأثر بالقيم الشاذة، ولا يمكن حسابه من الجداول المفتوحة.
- يخضع للعمليات الجبرية.

ملاحظة 1: كلما صغرت قيمة الانحراف المعياري دل على أن القيم أقل تشتتا، وأن الوسط الحسابي يمثلها، و نعتبر القيم عير مشتتة اذا كان الانحراف المعياري يمثل أقل من 20% من الوسط الحسابي.

ملاحظة 2: $\sigma = \frac{5}{4}e = \frac{3}{2}LQ$ و $\frac{E}{6} < S < \frac{E}{4}$ *اذا وقع الانحراف المعياري خارج المدى دل ذلك على وجود قيم شاذة.*

-4-2-3-2 خـــــواص:

- الانحراف المعياري للمعطيات الثابتة يساوي 0 لأن $x_i - \mu = a - a = 0$.
- اذا كانت x_1،...،x_n قيم سلسلة إحصائية متوسطها $\bar{x}$ و انحرافها المعياري ◷$_x$ وكان
  حيث: $y_i = ax_i + b$ $b \in R, a \in R$ فإن: $\bar{y} = a\bar{x} + b$ و $\sigma_Y = |a|\sigma_X$

بـــــــــرهــــــــــــان: $\sigma_Y{}^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n} = \frac{\sum_{i=1}^{n}(ax_i + b - a\bar{x} - b)^2}{n} = \frac{a^2\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = a^2\sigma_X^2$

-4-2-3-3 إذا كان متوسط العينة♠1هو $\overline{x_1}$ و انحرافها المعياري s_1 حجمها n_1و كان متوسط العينة♠2 هو $\overline{x_2}$ و انحرافها المعياري s_2 حجمها n_2 فإن المتوسط و الإنحراف المعياري للعينتين هما على التوالي: $\overline{X} = \frac{n_1\overline{x_1} + n_2\overline{x_2}}{n_1 + n_2}$ و $S^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 1}$

-4-3 معامل الاختلاف : وهو مقياس دون وحدة يفضل استخدامه عند المقارنة بين مجموعتين خاصة مختلفتي الوحدة.

-4-3-1 معامل الاختلاف النسبي: وهو العدد الذي نرمز له بـ CV و يعرف بالعلاقة:

$CV = \frac{\sigma}{\mu} \times 100$ في حالة بيانات المجتمع، و $CV = \frac{S}{\bar{X}} \times 100$في حالة بيانات العينة.

-4-3-2 معامل الاختلاف الربيعي: وهو العدد الذي نرمز له بـ CVQ يستخدم في حالة الجداول التكرارية المفتوحة و يعرف بالعلاقة:

$CVQ = \frac{Q_3 - Q_1}{Med} \times 100$.

مثال: لتكن لدينا معطيات سلسلة إحصائية كالتالي: $Q_1 = 3, Q_3 = 8, \overline{x} = 6, Med = 5, s = 1$

$$CV = \frac{1}{6} \times 100, CVQ = \frac{8-3}{5} \times 100$$

***4-4-* معايرة سلسلة :** لتكن $A(x_i, n_i)$ سلسلة احصائية وسطها الحسابي $\overline{X}$ و انحرافها المعياري 🕘 ، نرمز بـ y_i للقيمة المعيرة للقيمة x_i حيث: $y_i = \frac{x_i - \overline{X}}{\sigma}$ ومنه: $\overline{y} = 0, \sigma_Y = 1$

بـــــرهـــــــــان: $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})}{\sigma \times n} = 0$ **و** $\sigma_Y^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{\sigma^2 \times n} = \frac{\sigma^2}{\sigma^2} = 1$

مثال: في سباق *100*م سرعة، حصل العداء ✌ على *10* ثواني في مجموعته التي زمنها المتوسط *9* ثا بانحراف*2* ثا، و تحصل العداء 👌 على *9* ثواني في مجموعته التي زمنها المتوسط *8* ثا بانحراف*1* ثا ، أيهما أسرع في مجموعته؟

المقارنة ممكنة بعد المعايرة، نجد: $y_B = 1, y_A = \frac{10-9}{2} = \frac{1}{2}$ اذن: ✌ أسرع في مجموعته.

مقاييس الشكل

تـــمـــهيـــد: يصادفنا عند رسم المنحنى التكراري للتوزيع عدة أشكال فقد يكون المنحنى متماثلا أي متناظرا بالنسبة لمستقيم عمودي، وقد يكون ملتو في اتجاه معين أو شديد الالتواء منخفضا أو مرتفعا ...

فمقاييس الموقع و التشتت تعطينا صورة عن تمركز القيم و مدى تجانسها، لكن لا توضح كيفية انتشار القيم، لذا يصبح ضرورة ملحة إيجاد مقاييس تعطي وصفا للبيانات من جهة و تسمح بالمقارنة بين التوزيعات من جهة أخرى.

***5-1-* الـــعــــــــزوم:** نميز بين نوعين من العزوم، قد يكون العزم حول نقطة معينة أو حول المتوسط الحسابي.

***5-1-1-* العزم البسيط – اللامركزي** – نسمي عزما بسيطا من الرتبة ❒ العدد m_r المعرف كالآتي:

أ- **في حالة البيانات غير المبوبة:** $m_r = \frac{\sum_{i=1}^{n} x_i^r}{n}$

ب- **في حالة البيانات المبوبة:** **1-** $m_r = \frac{\sum_{i=1}^{n} n_i x_i^r}{\sum_{i=1}^{n} n_i}$

2 - إذا كانت c_1 ،c_2 ،...،c_n مركز فئات المتغير X وكانت n_1 ،n_2 ،...،n_n تكراراتها على الترتيب فإن:

$$m_r = \frac{\sum_{i=1}^{n} n_i c_i^r}{\sum_{i=1}^{n} n_i}$$

***5-1-1*- العزم المركزي :** نسمي عزما مركزيا من الرتبة ❒ العدد M_r المعرف كالآتي:

أ- **في حالة البيانات غير المبوبة:** $M_r = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n}$

ب- **في حالة البيانات المبوبة:** **-1** $M_r = \frac{\sum_{i=1}^{n} n_i(x_i-\overline{x})^r}{\sum_{i=1}^{n} n_i}$ا

2 - إذا كانت c_1 ،c_2 ،...،c_n مركز فئات المتغير X وكانت n_1 ،n_2 ،...،n_n تكرارتها على الترتيب فإن:

$$M_r = \frac{\sum_{i=1}^{n} n_i(c_i-\overline{x})^r}{\sum_{i=1}^{n} n_i}$$

***5-2*- مقاييس الالتواء:** يقصد بالالتواء امتداد التوزيع و يعبر عن درجة توزع البيانات حول نقطة مركزية.

5-2-1- مقاييس بيرسون:

5-2-1-1- مقاييس بيرسون الأول: يعطى بالعلاقة التالية: $P_1 = \frac{\overline{x}-Mod}{\sigma}$

$P_1 = 0$ التوزيع متماثل ، $P_1 > 0$ التوزيع ملتوي نحو اليمين، $P_1 < 0$ التوزيع ملتوي نحو اليسار.

لا يمكن استخدامه في الجداول المفتوحة، متعددة المناويل أو شديدة الالتواء

. **5-2-1-2- مقاييس بيرسون الثاني:** يعطى بالعلاقة التالية: $P_2 = 3\frac{(\overline{x}-Med)}{\sigma}$

$P_2 = 0$ التوزيع متماثل ، $P_2 > 0$ التوزيع ملتوي نحو اليمين، $P_2 < 0$ التوزيع ملتوي نحو اليسار.

لا يمكن استخدامه في الجداول المفتوحة أو شديدة الالتواء

. **5-2-1-3- مقاييس بيرسون للالتواء العزمي:** يعطى بالعلاقة التالية: $\beta_1 = \frac{M_3^2}{M_2^3}$

لا يعطي فكرة جيدة عن الالتواء سوى تماثل التوزيع.

. **5-2-2- مقاييس فيشر:** يعطى بالعلاقة التالية: $F_1 = \frac{M_3}{\sigma^3}$

$F_1 = 0$ التوزيع متماثل ، $F_1 > 0$ التوزيع ملتوي نحو اليمين، $F_1 < 0$ التوزيع ملتوي نحو اليسار.

5-2-3- مقاييس يول- الالتواء الربيعي-: يعطى بالعلاقة التالية: $\gamma_1 = \frac{Q_3+Q_1-2Med}{Q_3-Q_1}$

$\gamma_1 = 0$ التوزيع متماثل ، $\gamma_1 > 0$ التوزيع ملتوي نحو اليمين، $\gamma_1 < 0$ التوزيع ملتوي نحو اليسار.

5-3- التفــرطــح: ويقصد به مقدار أو درجة علو- ارتفاع- أو انخفاض منحنى التوزيع عن التوزيع الطبيعي.

ويكون التوزيع مدببا – مرتفعا- إذا تركزت القيم في المنتصف و قلت في الطرفين، ومفرطحا عكس ذلك.

5-3-1- مقاييس بيرسون للتفرطح: يعطى بالعلاقة التالية: $\beta_2 = \frac{M_4}{M_2^2}$

$\beta_2 = 3$ التوزيع متماثل ، $\beta_2 > 3$ التوزيع مدبب ، $\beta_2 < 3$ التوزيع مفرطح.

5-3-2- مقاييس فيشر للتفرطح: يعطى بالعلاقة التالية: $F_2 = \beta_2 - 3$

5-3-3- مقاييس كيللي: يعطى بالعلاقة التالية: $K = \frac{Q_3 - Q_1}{D_9 - D_1}$

$K = 0.526$ التوزيع طبيعي ، $K > 0.526$ التوزيع مدبب ، $K < 0.526$ التوزيع مفرطح.

مثال : احسب مقياس للتفرطح و الالتواء للسلسلة التالية

x_i	2	3	4	5	6	المجموع
$(x_i - \overline{x})$	**-2**	**-1**	**0**	**1**	**2**	**0**
$(x_i - \overline{x})^2$	**4**	**1**	**0**	**1**	**4**	**10**
$(x_i - \overline{x})^3$	**-8**	**-1**	**0**	**1**	**8**	**0**
$(x_i - \overline{x})^4$	**16**	**1**	**0**	**1**	**16**	**34**

ومنه : $M_2 = 2; M_3 = 0; M_4 = \frac{34}{5}$ **و** $Med = 4; Q_1 = 3; Q_3 = 5$ **اذن:** $\beta_2 = \frac{34}{20}$ **و** $\gamma_1 = \frac{5+3-2\times4}{5-3}$

سلسة تمارين الوحدة الأولى

التمرين الأول:

ليكن توزيع عمال مصنع حسب المدة الزمنية بالأيام للعطل المرضية خلال سنة

2 3 6 4 5 3 3 2 4 5 6 7 8 4 3 2 10 3 4 5 6 8 3 2 4 5 6 7 8 9 12 11 10 15 3 2 3 2 4 5 6 7 8 9 10 11 12 13 14 10 3 6 4 2 6 4 2 3 7 4

1- عين الوحدة الإحصائية و الطبع الإحصائي
2- شكل القيم في جدول إحصائي مبرزا التكرار النسبي و التكرار النسبي المئوي المجمع الصاعد

3- انشئ المضلع التكراري للتوزيع

التمرين الثاني: مصنع نسيج عند تجريب آلة جديدة قام بإحصاء الأخطاء في 75 عينة من 10 متر ، فكانت النتائج الآتية :

x_i	0	1	2	3	4	5
n_i	38	15	11	6	3	2

3- حدد المجتمع الإحصائي ، الميزة الإحصائية.
4- حدد التكرار النسبي ، التكرار المتجمع الصاعد.
5- مثل السلسلة الإحصائية بأعمدة تكرارية.
6- احسب : الوسط الحسابي ، الوسيط ، المنوال ، الربعي الأول ، العشري التاسع و الانحراف المعياري للسلسلة الإحصائية.

التمرين الثالث:

سجلت مدة نقل الطلبة من مقر سكناهم بالدقائق فكانت:

5-3-20-45-45-5-25-15-15-5-20-5-5-2-45-10-30-2-7-15-25-30-60-35-30-45-30-30-25-25-45-5-70-60-35-35-18-5-5-65-10-50-25-30-2-28-30-25-20-45

2- شكل القيم في جدول إحصائي مبرزا التكرار النسبي و التكرار المجمع النازل.
3- احسب الوسيط و الوسط الحسابي.
4- شكل القيم في جدول إحصائي بفئات-اعتمد طريقة ستورج-
5- اعد حساب الوسيط و المنوال .

التمرين الرابع:

يرغب صاحب مصنع معرفة رغبات زبائنه فأجرى سبرا للآراء فتحصل على معلومات موزعة حسب السن كماياتي:

العمر	[20□15[	[25□20[	[30□25[	[35□30[	[40□35[
f%	13	28	35	15	9

3- انشئ المدرج التكراري لهذا التوزيع ثم استنتج المنوال.
4- انشئ المنحنى للتكراري النسبي المجمع الصاعد ثم استنتج الوسيط.
5- عين حسابيا المنوال و الوسيط
6- احسب الوسط ، الربيع الثالث و معامل الاختلاف.
7- احسب الوسط الهندسي .

التمرين الخامس: ليكن الجدول الإحصائي التالي و الذي يمثل مدة صلاحية دواء بالشهور:

30□25[[	25□15[[	[L،15[	[L□4 [	[0□4[	المدة
m	n	240	250	60	التكرار

[42□30[	[50□42[	المجموع

1000	*30*	*110*

1- عين n و m إذا علمت أن $Q_3 = 24.52$
2- عين L إذا علمت أن $\bar{x} = 17.16$
3- احسب العزوم المركزية من الدرجة الثالثة و الرابعة
4- احسب معاملي فيشر للالتواء و التفرطح
5- تحقق من صحة التواء التوزيع بحساب معامل يول وبيرسون

التمرين السادس:

1- لتكن الدالة المعرفة بـ: $f(a) = \frac{1}{N}\sum_{i=1}^{n} n_i(x_i - a)^2$

برهن إن الدالة f تقبل قيمة حدية صغرى عند قيمة يطلب تعيينها. حدد عندئذ f

2- ليكن المتغير الإحصائي التالي $y_i = \frac{x_i - \bar{X}}{\sigma_X}$ برهن ان :

$\bar{Y} = 0$ و $\sigma_Y = 1$

السلاسل الإحصائية ذات متغيرين

أثناء دراسة على مجتمع إحصائي ما يمكن أحيانا ملاحظة طبعين – كميين، نوعيين أو مختلط- . ومن ثم نتعامل مع متغيرين، نتناول في هذا الفصل دراسة العلاقة بين المتغيرين ، و نستخدم لذلك بعض طرق التحليل الإحصائي مثل : الارتباط الانحدار ... .

نعرف عندئذ متغيرين إحصائيينY ، X ، و تسمى الثنائية (X,Y) متغيرة إحصائية ذات بعدين.

مثلا: ملاحظة نقطة المراقبة المستمرة و نقطة الامتحان لطلبة سنة أولى.

2- السلسلة الاحصائية ذات بعدين:

ليكن Xو Y المتغيرتين المدروستين، p عدد قيم الطبع بالنسبة للمتغير X، q عدد قيم الطبع بالنسبة للمتغير Y و n العدد الاجمالي للمشاهدات.

مجموعة الثنائيات $(x_i, y_j)_{1\le i\le p, 1\le j\le q}$ تشكل سلسلة إحصائية ذات متغيرين.

ملاحظة: إذا كان المتغير الأول X هو الزمن تسمى سلسلة زمنية.

5-1 – **جدول التوزيع المشترك**: بفرض ان المتغيرين منفصلان و أن الطبعين كميان، الجدول التالي الذي يضم المتغيرتان يسمى جدول التوزيع المشترك، أو الجدول ذي المدخلين.

X Y	y_1	y_2	...	...	y_q	المجموع
x_1	n_{11}	n_{12}	...	...	n_{1q}	$n_{1.}$
x_2	n_{21}					
.						
x_p	n_{p1}				n_{pq}	$n_{p.}$
المجموع	$n_{.1}$				$n_{.q}$	n

في كل خانة من الجدول، نكتب التكرار n_{ij} عدد مشاهدة المعطيات حيث $X = x_i$ و $Y = y_j$.

نعرف التكرارات المطلقة التالية:

-5-1-1 التكرار الهامشي:

$n_{i.} = \sum_{j=1}^{q} n_{ij}$ أي مجموع مكونات السطر ♓

$n_{.j} = \sum_{i=1}^{p} n_{ij}$ أي مجموع مكونات العمود ℯɼ

إذن: التوزيع الهامشي $n_{i.}$ هو عدد الوحدات التي تتمتع بالطبع ♓ للمتغير X مهما يكن توزيع المتغير Y.

التوزيع الهامشي $n_{.j}$ هو عدد الوحدات التي تتمتع بالطبع ℯɼ للمتغير Y مهما يكن توزيع المتغير X.

مثال: كل الطلبة المتحصلين على العلامة 10 في الامتحان مهما كانت علامتهم في الرقابة المستمرة.

-5-1-2 **التكرار الشرطي:** ويعرف من أجل كل قيمة لـ ♓ و ℯɼ .

$$n_{j/i} = \frac{n_{ij}}{n_{i.}} \quad و \quad n_{i/j} = \frac{n_{ij}}{n_{.j}}$$

$n_{j/i}$ هي توزيع المتغير Yبتثبيت الطبع ♓للمتغير X.

مثال: توزيع علامات الامتحان للطلبة الذين لهم نفس نقطة الرقابة المستمرة.

-5-1-3 **التكرار النسبي**: و نعرف التكرار النسبي المشترك، التكرارات النسبية الهامشية f_{ij}، $f_{i.}$ و $f_{.j}$.

$$f_{.j} = \frac{n_{.j}}{n} \text{ و } f_{i.} = \frac{n_{i.}}{n} \quad و \quad f_{ij} = \frac{n_{ij}}{n}$$

قضية: التكرار النسبي يحقق: $\sum_{i=1}^{p}\sum_{j=1}^{q} f_{ij} = 1$ و $f_{ij} \geq 0; \forall i,j$ و $\sum_{j=1}^{q} f_{i.} = 1$ و $\sum_{i=1}^{p} f_{.j} = 1$

-5-1-4 **استقلال المتغيرين**: توزيعي ✠ و ✡ مستقلان اذا و فقط اذا كان من أجل كل قيمة للأدلة ♓ و ℯɼ:

$$f_{ij} = f_{i.} \times f_{.j}$$

-5-1-5 **الدالة التجميعية:** $F(x,y) = \sum_{i\leq x}\sum_{j\leq y} f_{ij}$

-5-2 **تمثيل البيانات:** يمكن تمثيل المعطيات بأحد الطرق:

-5-2-2 **سحابة النقط:** في معلم متعامد و مناسب، مجموعة النقط $M_i(x_i;y_j)$ تسمى سحابة النقط في R^2 لسلسلة ذات متغيرين ذوي طبع كمي.

ملاحظة: حتى يكون لدراسة سحابة النقط معنى ينبغي ان يكون عدد النقط كبيرا.

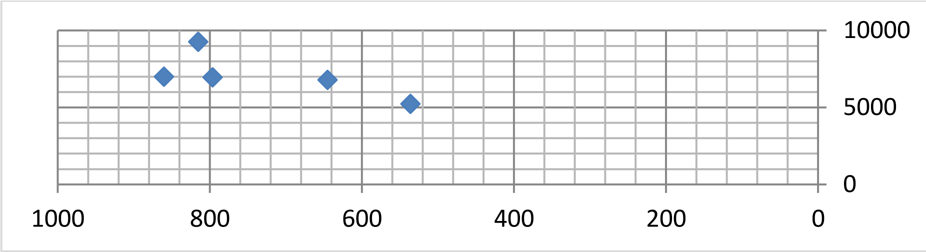
-5-2-1-1 **تغيير المبدأ:** نضع: $z_i = x_i - a$ و $t_i = y_i - b$.

-5-2-1-2 **تغيير الوحدة:** نضع $u_i = k y_i$

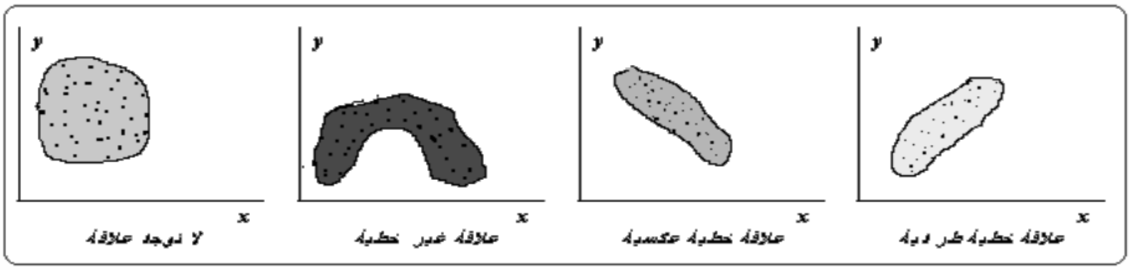
مثال: الجدول التالي يبين عدد خريجي جامعة لمدة 5 سنوات بدلالة عدد الطلاب

x_i	645	796	536	860	815
y_i	6780	6951	5219	6990	9253

يمكن إختيار $\dot{O}(500;5000)$ كمبدأ و 2سم لكل 60 خريجا و 10 سم لـ 800 طالب



5-2-1-3- أشكال الإنتشار:



5-2-2- **البيان المتعدد:** ويتم استعماله في الطبع الوصفي.

5-3- النقطة المتوسطة: ونرمز لها بالرمز ☝ حيث $G(\overline{x};\overline{y})$ مع $\overline{x}=\frac{1}{n}\sum_i n_{i.}x_i$ و $\overline{y}=\frac{1}{n}\sum_i n_{.j}y_j$.

5-3-1- **المتوسطات المتحركة:** هي واحدة من اقدم تقنيات التحليل الاحصائي وتستخدم بشكل رئيسي كأداة لمعرفة الإتجاه.

تعريف 1-3-1: لتكن السلسلة الزمنية الاحصائية $x_1,x_2,\dots,x_n$ في الأزمنة $t_1,t_2,\dots,t_n$.

*- الوسط المتحرك من الرتبة□ حيث: $p=2k+1$ من أجل t_i $(2\le i\le n-1)$ هو وسط القيم $t_{i-k},\dots,t_i,\dots,t_{i+k}$.

أي: $\frac{t_{i-k}+\cdots+t_{i-1}+t_i+t_{i+1}+\cdots+t_{i+k}}{p}$

*- الوسط المتحرك من الرتبة□ حيث: $p=2k$ من أجل t_i $(2\le i\le n-1)$ هو العدد:

$$\frac{\frac{t_{i-k}}{2}+\cdots+t_{i-1}+t_i+t_{i+1}+\cdots+\frac{t_{i+k}}{2}}{p}$$

5-3-2- **التمليس بالأوساط المتحركة:** تمليس منحنى سلسلة زمنية بالأوساط المتحركة من الرتبة □ هو إنجاز منحنى سلسلة أخرى قيمها الأوساط المتحركة من الرتبة □ لقيم السلسلة الاصلية.

ملاحظة: عند تمليس سلسلة زمنية بالأوساط المتحركة نهمل & قيمة عند الأطراف.

مثال:

t_i	1974	1975	1976	1977	1978	1979	1980
x_i	5	7	6	4	8	6	5
الوسط المتحرك من الرتبة 3		$\frac{5+7+6}{3}=6$	5.67	6	6	6.33	

الوسط المتحرك من الرتبة 4			$\frac{\frac{5}{2}+7+6+4+}{4} = 5.875$	6.125	5.875		

5-4- مقاييس الارتباط: يستخدم تحليل الارتباط لتحديد نوع و قوة العلاقة بين المتغيرين، وتقاس طبيعة، قوة و أثر العلاقة بين المتغيرين بعدة مقاييس منها:

5-4-1- التغاير: و يسمى أيضا التباين المشترك وهو عدد حقيقي يقيس ارتباط و كيفية انتشار النقاط حول النقطة المتوسطة.

$$cov(X,Y) = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}(x_i - \overline{x})(y_j - \overline{y})$$

باستعمال قاعدة **كوينغ** نجد: $cov(X,Y) = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}x_i y_j - \overline{xy}$.

5-4-2- مقياس الارتباط الخطي: ويعتبر مقياسا للارتباط الخطي بين متغيرين كميين وهو نسبة دون تمييز – بدون وحدة-.

ونرمز له بالرمز ρ في حالة المجتمع و ❐ في حالة العينات حيث:

$$\rho = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

5-4-2-1- خواص:

أ- $\rho(X,Y) = \rho(Y,X)$ ب- $-1 \le \rho \le 1$.

ت- $\rho(X,X) = 1$ **لان** $cov(X,X) = V(X)$:   د- $\rho(X,-X) = -1$.

ه- $\rho(aX+b, cY+d) = \rho(X,Y); a,c \neq 0$.

برهــان: باستعمال خاصيتي التبديل و التجميع للضرب و الجمع، وخواص التباين نجد المطلوب.

ملاحظة: توجد أزواج من المتغيرات لها نفس التوزيعات الهامشية لكنها تختلف في التغاير و الارتباط لهذا فإن التغاير و معامل الارتباط الخطي يقيسان التداخل بين المتغيرين.

مثــال: ليكن X و Y متغيران إحصائيان لهما التوزيع المشترك التالي:

Y \ X	4	10	
1	5	2	7
3	1	0	1
	6	2	8

التواتر الهامشي للمتغير Xهي:

X	1	3
$f_{i.}$	$\frac{7}{8}$	$\frac{1}{8}$

نجد: $\overline{x} = 1 \times \frac{7}{8} + 3 \times \frac{1}{8}$ و $cov(X,Y) = \frac{1}{8}(5\times 1\times 4 + 1\times 3\times 4 + 2\times 1\times 10 + 0\times 3\times 10) - \frac{10}{8}\times\frac{44}{8}$

ملاحظة: ونستخدم في حالة العينات: $r_{X,Y} = \frac{s_{X,Y}}{s_X s_Y}$

5-4-3- معامل الارتباط للرتب: و تستخدم في حالة متغيرات وصفية ترتيبية- مثلا: تقديرات المواد(جيد، حسن،...)...-

-5-4-3-1 معامل سبيرمان: $r_S = 1 - \frac{6\sum_i d_i}{n(n^2-1)}$

حيث : $d_i = R_{i,X} - R_{i,Y}$.أي الفرق بين رتب مستويات المتغير الأول و رتب مستويات المتغير الثاني.

-5-4-3-2 معامل كندال: $\tau = \frac{2S}{n(n-1)}$ حيث S مجموع التوافيق لترتيب الثنائيات.

ملاحظة: في حالة استعمال هذه المعاملات من أجل ثنائية المتغيرات الطبيعية، وبمعرفة معامل الارتباط الخطي. فمن أجل n كبير بالقدر الكافي لدينا: $r_S = \frac{6}{\pi} arcsin \frac{\rho}{2}$

مثال : ليكن تقدير تفضيل 5 سلع حسب مستوى الدخل .

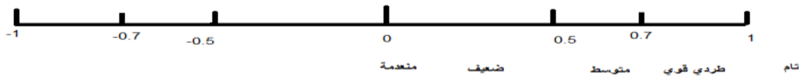
تفضيل السلعة	أ	ب	ب	جـ	ب
مستوى الدخل	جيد	حسن	متوسط	متوسط	حسن

الرتب	1	2	3	4	5
تقدير 1	أ	ب	ب	ب	جـ
رتب X	1	$(2+3+4/3)=3$			5
تقدير2	جيد	حسن	حسن	متوسط	متوسط
رتبY	1	$(2+3/2)=2.5$		4.5	

X	Y	رتبX	رتبY	♎	♎²
أ	جيد	1	1	0	0
ب	حسن	3	2.5	0.5	0.25
ب	متوسط	3	4.5	1.5-	2.25
جـ	متوسط	5	4.5	0.5	0.25
ب	حسن	3	2.5	0.5	0.25

ومنه $r_s = 1 - \frac{6\times3}{5(25-1)} = 0.85$ و يدل على وجود ارتباط طردي قوي بين تفضيل سلعة معينة و مستوى الدخل.

ملاحظة: يمكن استخدام المعاملات السابقة لحساب الارتباط بين متغيرين كميين.



-5-4-4 درجة قوة معامل الارتباط:

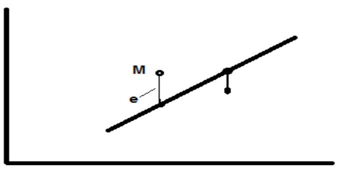
-5-5 التعديل – الانحدار- الخطي : القيام بتسوية خطية أو بتعديل خطي لسحابة النقط يعني ايجاد دالة خطية تعبر بطريقة تقريبية عن Y بدلالة X أي تعيين عددين حقيقيين ♌، ♋ بحيث:$y = ax + b$.

ويهدف الى دراسة و تحليل أثر أحد المتغيرين على الآخر،كذا عملية التنبؤ.

-5-5-1 **الانحدار الخطي البسيط:** عندما يكون لسحابة النقط $M_i(x_i, y_i)$ المرفقة بسلسلة إحصائية لمتغيرين عدديين (X, Y) طبيعيين شكل متطاول فإنه يمكن إنشاء مستقيم تقع حوله نقط السحابة.

معادلة مستقيم الإنحدار هي: $Y = aX + b + \varepsilon$ حيث :

Yالمتغير التابع – المفسر- ، X التابع المستقل –المؤثر- ✧ الخطأ العشوائي.



-5-5-1-1 **تقدير نموذج الانحدار الخطي البسيط:** باستعمال طريقة المربعات الدنيا يمكن تعيين معاملات المستقيم
و هو أفضل تقدير و الذي يجعل $\sum_{i=1}^{n}\varepsilon_i^2$ أصغريا.

مبرهنة: مستقيم الانحدار بالمربعات الدنيا هو المستقيم الذي يشمل النقطة المتوسطة لسحابة النقط و معادلته المختصرة:

$$y = ax + b \quad \text{حيث:} \quad a = \frac{cov(X,Y)}{V(X)} \;\text{و}\; b = \bar{y} - a\bar{x}$$

البرهان: المتغير Y عشوائي بينما X مقاس دون خطأ، ليكن $y = ax + b$ المعادلة المختصرة لمستقيم الانحدار بالمربعات الدنيا، وليكن $y_i = ax_i + b$ القيمة المحسوبة . اذن $S(a,b) = \sum_{i=1}^{n}\varepsilon_i^2 = \sum_{i=1}^{n}(y - y_i)^2 = \sum_{i=1}^{n}(\acute{y}_i - ax_i - b)^2$ يكون أصغريا من أجل:

$$\frac{dS}{da} = 0 \Rightarrow \sum_{i=1}^{n} x_i\,(\acute{y}_i - ax_i - b) = 0 \Longrightarrow a = \frac{cov(X,Y)}{V(X)} \;\text{و}\; \frac{dS}{db} = 0 \Rightarrow \sum_{i=1}^{n}(\acute{y}_i - ax_i - b) = 0 \Longrightarrow b = \bar{y} - a\bar{x}$$

-5-5-2 **مستقيم مايير:** ويتم بكتابة معادلة المستقيم (G_1G_2) حيث G_1 النقطة المتوسطة لنصف النقط الاولى و G_2 النقطة المتوسطة للنقط المتبقية.

مثال : ندرس نسبة تكاليف – ماء، غاز، كهرباء – من دخل عائلي .

$x_i = a_i - 1970$	8	14	22	24	30	34
$y_i\%$	4.4	5.2	4.3	3.2	3.3	2.8

أ- مستقيم الانحدار: $a = \frac{(8\times4.4+\cdots+34\times2.8)-6\times22\times3,86}{(8^2+\cdots+34^2)-6\times22^2} = -0.08$ و $b = 3.86 - (-0.08)\times 22 = 5.58$

ب- مستقيم مايير: $G_1(14.67,4.63)$ و $G_2(29.33,3.1)$ ومنه $a = -0.06$ و $b = 4.89$

-5-6 **التعديل الدالي :** يرتكز على البحث عن دوال مناسبة، و بتغيير في المتغير نعيد المعادلة للبحث عن مستقيم الانحدار.أي
نضع: $t_i = g(x_i)$ و $z_i = f(y_i)$ ونعين معادلة الانحدار بالمربعات الدنيا للمتغيرين الجديدين.
أشهر التعديلات الدالية: التعديل الأسي، اللوغاريتمي، التعديل بقطع مكافئ

مثال: الجدول التالي يمثل درجة التلوث في مدينة ما على مدار سبع سنوات.

x_i	*1*	*2*	*3*	*4*	*5*	*6*	*7*
y_i	*9.8*	*9.6*	*9.7*	*10.5*	*11.5*	*12.7*	*14.6*

برسم سحابة النقط فيوحي شكل الانتشار بوجود قطع مكافئ ذروته $x = 2$.

اذن نبحث عن تعديل من الشكل $y = a(x-2)^2 + b$ فنضع: $t_i = (x_i - 2)^2$ ونكتب معادلة الانحدار $y = at + b$ ثم نستنتج.

تمارين مقترحة للوحدة الثانية

التمرين الأول:

الجدول أدناه نشر في مجلة اقتصادية وهو يبين الوقت الجزئي للطبقة العاملة

السنة x	1980	1985	1990	1995	1997
الوقت y %	8.3	11	12	15.2	16.8

1- مثل في معلم متعامد سحابة النقط
2- عين إحداثيات النقطة المتوسطة
3- اوجد معامل الارتباط الخطي . هل التعديل الخطي مبرر؟
4- حدد المعادلة المختصرة لمستقيم التعديل باستعمال طريقة المربعات الدنيا
5- قدر الوقت الجزئي سنة 2019.

التمرين الثاني:

تحصلنا على الجدول أسفله في دراسة متغيرين حول مجتمع متكون 25 فردا

Y X	$[0;5[$	$[5;10[$	$[10;15[$	$[15;20[$
0	2	0	3	0
1	0	1	2	3
2	0	0	1	1
3	4	3	0	0
4	1	0	4	0

1- اعط التوزيع الهامشي لكل من X وY
2- اعط التوزيع الشرطي لـY علما أن $X=2$
3- هل X وY مستقلان.
4- اكتب معادلة الانحدار لـY على X.

التمرين الثالث:

شرعنا في تربية نوع من البكتيريا داخل وسط مناسب، و قمنا بحساب العدد الناتج عن التكاثر في كل يوم خلال *14* يوم، فتحصلنا:

اليوم	*1*	*2*	*3*	*4*	*5*	*6*	*7*
العدد	*16*	*30*	*58*	*124*	*236*	*485*	*960*
اليوم	*8*	*9*	*10*	*11*	*12*	*13*	*14*
العدد	*965*	*969*	*981*	*984*	*992*	*996*	*1000*

1- ما هي الوسيلة التطبيقية التي نستعملها لاختيار طبيعة دالة التسوية.
2- لا تتطور بعض الظواهر خلال عمرها بنفس الطريقة، هل أننا في حالة انفصام في النموذج.
3- عبر عن تطور عدد البكتيريا بواسطة دالتين موافقتين لمرحلتي الظاهرة.

التمرين الرابع:

الجدول التالي يمثل تطور ميزانية الإشهار بالمليون دينار لمؤسسة جزائرية من 2009 الى 2016 .

رتبة السنة	1	2	3	4	5	6	7	8
الميزانية	0.4	0.45	0.5	0.56	0.63	0.68	0.75	0.83

1- اوجد معادلة مستقيم الانحدار بالمربعات الدنيا – تعطى النتائج مدورة

الى 10^{-2} -.

2- قدر الميزانية المتوقعة سنة 2020 –بفرض التطور يتم بنفس الطريقة-
3- عين احداثيي G_1 النقطة المتوسطة لسحابة النقط الأربعة الأولى و G_2 النقطة المتوسطة لسحابة النقط الأربعة الأخرى
4- اكتب معادلة المستقيم(G_1 G_2) –يسمى مستقيم مايير-

العناصر الأساسية لنظرية الاحتمالات

نبذة تاريخية : إن السمة البارزة للقرن العشرين هي التقدم الواضح في العلوم التجريبية عامة و في الفيزياء خاصة مما أنتج معارف كثيرة و علاقات فيزيائية جديدة فلقد تم اكتشاف النظرية النسبية،نظرية الكم وعلاقة المادة بالإشعاع ...وتم إدراك أن القوانين الكلاسيكية لا تعد صحيحة دائما – الأجسام الصغيرة ذات السرعات الكبيرة- لذا توجب استخدام نظرية الاحتمالات في الفيزياء و في مختلف العلوم . فلقد استخدم مصطلح الاحتمالات منذ نشوئها في القرن 17 لدى محاولة دراسة العاب الحظ بأشكال مختلفة أهمها : الاحتمالات الفيزيائية و الذي يرتبط بالنظم الفيزيائية العشوائية مثل : الذرات المشعة، عجلة الروليت... وفي مثل هذه النظم يميل حدث ما إلى الوقوع بمعدل ثابت أو بتكرار نسبي على المدى البعيد. وتبلورت في هذا القرن بشكل قوانين مبرهن عليها رياضيا، وتم استخدامها في نطاق واسع في كل العلوم الكونية، الطبية، الاجتماعية، الاقتصادية، ...

ارتبط مفهوم الاحتمال في أذهان الكثيرين بالإمكانية – وهي حدث – بينما هي في العلم عدد يقيس حظوظ وقوع حدث أو إمكانية.

6-1-أساسيات نظرية الاحتمالات:

6-1-1- مصطلحات:

6-1-1-1- **التجربة العشوائية:** وهي كل تجربة لا يمكن توقع نتيجتها رغم معرفة مجموعة النتائج الممكنة لها.

6-1-1-2- **المجموعة الشاملة:** وهي مجموعة النتائج الممكنة في تجربة عشوائية و نرمز لها بΩ- فضاء العينة-.

6-1-1-3- **الحادث:** نسمي كل عنصر من Ω حدثا أوليا – بسيطا-،إذا كان $A \in \Omega$ نقول ان A حادث.

6-1-1-4- **عشيرة الأحداث:** نقول عن جماعة Λ مؤلفة من مجموعات جزئية منΩ أنها عشيرة أحداث إذا تحققت الشروط التالية:

- ✓ $\Omega \in \Lambda$
- ✓ $A \in \Lambda \Rightarrow \overline{A} \in \Lambda$
- ✓ $\forall A, B \in \Lambda: A \cup B \in \Lambda$

مثال :عند رمي زهرة نرد متوازنة نجد :

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

$\Lambda = \{\phi, \Omega, \{1,3,5\}, \{2,4,6\}\}$ تمثل عشيرة أحداث من Ω.

*عين جماعة من Ω لاتمثل عشيرة أحداث؟

6-1-1-5- **الفضاء القابل للاحتمال** نسمي الثنائية (Ω, Λ) فضاءا قابلا للاحتمال. مثلا $(\Omega, P(\Omega))$

*- إذا كان $A \in \Lambda; r \in A / r \in \Omega$ نقول أن الحدث A قد أنجز.

6-2- العمليات على الأحداث:

- **الحادث الأكيد و الحادث المستحيل:** Ω هي الحادث الأكيد – لابد أن يتحقق احد نواتج Ω-، بينما ϕ تسمى الحادث المستحيل (ظهور 3 صور عند رمي قطعة نقدية متزنة مرة فقط)
- **الحادث العكسي:** $\overline{A}$ هي الحادثة العكسية للحادث A وتحوي كل عناصر Ω ماعدا عناصر A (تقع عندما لا تقع A).

- **الحادثان غير المتلائمين** : كان $A \cap B = \emptyset$ نقول ان الحادثتين غير متلائمتين – لا يقعا معا -.

ملاحظة: يمكن إجراء مقارنة بين نظرية الحوادث و نظرية المجموعات.

6-3- طرق العد: يهتم التحليل التوفيقي بإعطاء عدد الطرق الممكنة لحساب عدد الحالات الملائمة للمجموعات ضمن شروط معينة من خلال القواعد الرياضية.

6-3-1- المبدأ الأساسي للعد:

6-3-1-1- قاعدة الضرب: إذا أمكن القيام بعملية ما ب n طريقة ممكنة ، وقمنا بعمل آخر ب m طريقة ممكنة من اجل كل طريقة من الطرق السابقة فانه يمكن القيام بالعمليتين معا ب n x m طريقة ممكنة.

6-3-1-2- قاعدة الجمع: إن قاعدة الضرب تطبق في حالة الحوادث غير المتنافية بالتبادل ، إما إذا كان الحدثان متنافيين فان عدد مرات وقوع احدهما أو الآخر هو $n+m$ طريقة ممكنة.

مثال:

1- عدد الخيارات المتاحة لاختيار طالب حسب الجنس و القسم المنتمي له هو 2×4=8.

2- عدد الإمكانيات المتوفرة لاختيار ممثلين من جنسين مختلفين لفوج متكون من 4 ذكور و 6 إناث هو 6×4+4×6=48

6-3-2- قوائم عناصر مجموعة منتهية

لتكن E مجموعة منتهية ذات n عنصرا و m عدد طبيعي (nوm اكبر أو يساوي من 1).

نسمي قائمة ذات m عنصرا من E كل متتالية مرتبة من m عنصرا من عناصر E، أ ي كل عنصر $(a_1, \dots, a_m)$ من E^m.

*- عدد قوائم E ذات m عنصرا يساوي n^m

مثال:إن كل عدد مكون من 3 أرقام مشكلة من الأرقام 1.2.3.4.5 هو قائمة ذات 3 عناصر إذن عدد الأعداد هو $5^3 = 125$

6-3-3- تراتيب عناصر مجموعة منتهية

نسمي القائمة التي عناصرها متمايزة مثنى مثنى ترتيبة ويرمز لعدد الترتيبات ذات m عنصرا من بين n عنصرا بالرمز A_n^m ونكتب:

$$A_n^m = n(n-1)(n-2)\dots(n-m+1) = \frac{n!}{(n-m)!}$$

6-3-3-1- التبديلة: ترتيبة ذات n عنصرا من مجموعة ذات n عنصرا تسمى تبديلة ذات n عنصرا ، عدد التبديلات

إذن هو $n!$ و نكتب: $P_n = A_n^n = n!$

مثال: في المثال كل عدد ذي 3 أرقام متمايزة مثنى مثنى هو ترتيبة ذات 3 عناصر ، عدد الأعداد هو $3 \times 4 \times 5 = 60$.

6-3-3-2- حالة خاصة: إذا تباديل العناصر من وضعية دائرية فان عدد الطرق هو$P_{\overline{n}} = (n-1)!$.

مثال: عدد الطرق التي يمكن لـ 3 أخوة الجلوس حول طاولة مستديرة هو $P_{\overline{3}} = 2!$.

6-3-3-3- التباديل مع التكرار: يراد أحيانا معرفة عدد تباديل مجموعة من العناصر بعضها متماثلا - متشابه أو مكرر- .

عدد تبديلات ■ و التي يكون منها ■1 عنصرا متماثلا، ...، □■ عنصرا متماثلا هو: $P_n^{n_1,\dots,n_r} = \frac{n!}{\prod_{i=1}^{r} n_i!}$.

مثال: عدد الأعداد من ستة أرقام و التي يمكن تكوينها بتكرار الرقم 1 مرتين و 3 مرات الرقم 2 ومرة واحدة الرقم 3 هو $P_6^{2,3,1} = \frac{6!}{2!\times 3!\times 1!} = 60$ مثلا: 121223،112223....

6-3-4- التوفيقات

لتكن E مجموعة منتهية ذات n عنصرا و m عدد طبيعي حيث $0 \le m \le n$، نسمي توفيقة ذات m عنصرا من E كل جزء منE ذي m عنصرا. نرمز لعدد التوفيقات ذات m عنصرا من E بالرمز C_n^m

ملاحظة: $C_n^n = 1, C_n^1 = n, C_n^0 = 1$

مبرهنة: من اجل كل عددين طبيعيين n وm حيث $0 \le m \le n$ نجد:

$$C_n^m = \frac{A_n^m}{m!} = \frac{n!}{m!\,(n-m)!}$$

6-3-4-1- خواص : من اجل كل عددين طبيعيين n وm حيث $0 \le m \le n$ لدينا:

1- $C_n^m = C_n^{n-m}$

2- $C_n^m = C_{n-1}^m + C_{n-1}^{m-1}$

برهـــان: 1- $C_n^m = \frac{n!}{(n-m)!(n-n+m)!} = C_n^{n-m}$

2-بتوحيد المقامات و التبسيط نحصل على الطرف الثاني انطلاقا من الأول.

6-3-4-2- التوافيق مع الإعادة: عدد التوفيقات ذات ❒ عنصرا مع إمكانية تكرار العنصر نفسه من ■ مختلف هو:

$$K_n^r = C_{n+k-1}^r$$

مثال: عدد العينات المكونة من 3 طلاب و التي يمكن سحبها مع الإعادة من مجموعة ذات 6 طلبة هو = C_{6+3-1}^3

56

6-3-5- دستور ثنائي الحد

ليكن a و bعددان طبيعيان ، n عدد طبيعي اكبر أو يساوي1 لدينا:

$$(a+b)^n = \sum_{m=0}^{n} C_n^m a^m b^{n-m}$$

برهـــــــان: نستعمل البرهان بالتراجع

مثال: 1- يحتوي صندوق على 5 كرات سوداء و4 كرات بيضاء متشابهة لانفرق بينها في اللمس ، نسحب في آن واحد 3كرات.

إن كل سحبة هي توفيقة ذات 3 عناصر من9 عناصر و بالتالي عدد السحبات هو:

$$C_8^3 = 84$$

$\sum_{k=0}^{n} C_n^k \frac{3^{n-k}}{4^k} = \left(3+\frac{1}{4}\right)^n = \left(\frac{13}{4}\right)^n$-2

2-3-5-1- تقريب ستيرلينغ: من أجل ■ كبيرة، يمكن أن نكتب : $n! \approx \sqrt{2\pi n} n^n e^{-n}$

2-3-5-2- مثلث باسكال –الكرخي-: باستعمال الخاصية 2 نستطيع حساب عدد التوفيقة باستعمال الجدول التالي

$n \quad m \quad 1 \quad 1 \quad 1$

$1 \quad 2 \quad 1$
$1 \quad 3 \quad 31 \quad 4$
$1 \quad 4 \quad 6$

6-4- مسلمات كلوموغوروف

لتكن Ω فضاء عينة ، Λ عائلة من الاحداث و ليكن P تابع حقيقي معرف على Λ .نقول عن P انه دالة احتمال و ان $P(A)$ احتمال الحادث A إذا تحققت مسلمات كالموغوروف التالية:

$$\forall A \in \Lambda: 0 \le P(A) \le 1$$
$$P(\Omega) = 1$$
$$\forall i,j \in N, i \neq j: A_i \cap A_j = \emptyset \Longrightarrow P(iA_i) = \sum_i P(A_i)$$

6-4-1 الفضاء الاحتمالي

عندما يكون عدد مخارج التجربة العشوائية منتهيا ، نعرف على مجموعة المخارج متتالية اعداد $\{p_1, p_2 ... p_k\}$تحقق $\sum_{i=1}^{k} p_i = 1$ □ $p_i \ge 0$ نسمي هذه الاعداد احتمال

نسمي الثلاثية (Ω, Λ, p)فضاء احتمالي

مبرهنة : في حالة تساوي احتمال على Ω ، من اجل كل حادثة A لدينا:

$$p(A) = \frac{card(A)}{card(\Omega)} = \frac{A□□□□□□□□□□□□□□□□□□□□□□□□□□□□□}{□□□□□□□□□□□□□□□□□□□□□□□□□□□□□}$$

مثال:نفرض ان احتمال ميلاد ذكر او انثى متساويان، نختار عشوائيا عائلة ذات 5 أطفال ، احتمال ان يكون عدد الاناث اكبر من عدد الذكور هو:

$$P(A) = P(A_1) + P(A_2) + P(A_3) = \frac{C_5^5 + C_5^4 + C_5^3}{2^5}$$

6-4-2- خواص و نتائج: من اجل كل حادثتين A و B من Λ

1- $P(\emptyset) = 0$ $\qquad$ **2** -$P(\overline{A}) = 1 - P(A)$

3- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4- $P(A - B) = P(A) - P(A \cap B)$
5- $A \subseteq B \Rightarrow P(A) \leq P(B)$

برهـــان: 1- لدينا $\Omega \cup \emptyset = \Omega$ و $\Omega \cap \emptyset = \emptyset$ ومنه: $P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) = P(\Omega)$

5- $A \subseteq B \Rightarrow B = A \cup B \setminus A$ **و منه** $P(B) = P(A) + P(B \setminus A)$

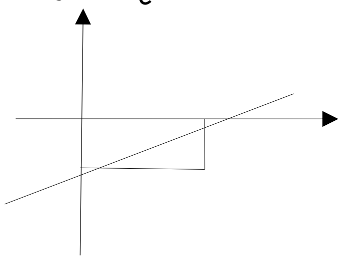
ملاحظة:

- الصيغة السابقة لحساب الاحتمال تستخدم فقط في حالة فضاء عينة منتظم ويعبر عنه "بطريقة عشوائية" للدلالة عليه.
- في فضاء العينة غير المنته و غير القابل للعد، نعتبر تلك التي لها قياس محدود و نجد $P(A) = \frac{\square\square\square\square A}{\square\square\square\square \Omega}$.

مثال:
نختار نقطتان ♋ و ♌ عشوائيا حيث: $3 \leq a \leq 0$ و $0 \leq b \leq 2-$. *ما احتمال أن تكون المسافة بين ♋ و ♌ أكبر من* 3.

$\Omega = \{(a,b), 0 \leq a \leq 3, -2 \leq b \leq 0\}$ $A = \{(a,b) \in \Omega / a - b > 3\}$

$P(A) = \frac{\square\square\square\square\square A}{\square\square\square\square\square \Omega} = \frac{1}{3}$



الاحتمالات الشرطية

7-1- تعريف: لتكن 👌 حادثة من فضاء العينة ✸ بحيث: $P(B) > 0$.

يعرف إحتمال وقوع الحدث ✌ بفرض أن 👌 قد وقع، بعبارة أخرى الإحتمال الشرطي للحدث ✌ اذا وقع 👌 كمايلي:

$$P(A/B) = P_B(A) = \frac{P(A \cap B)}{P(B)}$$

7-1-1- حالة خاصة: إذا كان ✸ فضاءا منتظما منتهيا فإن: $P(A/B) = \frac{card(A \cap B)}{card(B)}$

مثال: ألقينا زهرتي نرد منتظمتين، إذا كان مجموع الرقمين الظاهرين مساويا 6، أوجد احتمال أن يكون الرقم الظاهر في إحدى الزهرتين هو 2 .

لدينا: $B = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$ و $A = \{(1,2), (3,2), \dots, (2,6)\}$

$$P(A/B) = \frac{card(A \cap B)}{card(B)} = \frac{2}{5}$$

7-2- خواص:

أ- $P_B(A) + P_B(\overline{A}) = 1$

ب- $P(A \cap B) = P(B)P_B(A) = P(A)P_A(B)$

ت- $P_B(iA_i) = \sum_i P_B(A_i); A_i \cap A_j = \emptyset$

مثال: يحتوي صندوق على 5 كرات حمراء و 3 خضراء، لا نفرق بينها في اللمس.

نسحب كرتين على التوالي دون إرجاع، نضع ✌ الكرة المسحوبة الأولى حمراء و 👌 الكرة المسحوبة الثانية خضراء.

$$P(A)=\frac{5}{8}, P_A(B)=\frac{3}{7}, P(A\cap B)=\frac{5}{8}\times\frac{3}{7}$$

7-3- دستور الإحتمالات الكلية:

لتكن الاحداث $A_1, A_2, \dots, A_n$ تجزئة لفضاء العينة Ω أي: غير خالية و متنافية مثنى مثنى و $i=1nA_i=\Omega$.

لتكن 👌 حدث كيفي آخر، لدينا: $P(B)=P(A_1\cap B)+\dots+P(A_n\cap B)=P(A_1)P_{A_1}(B)+\dots+P(A_n)P_{A_n}(B)$

اذن: $P(B)=\sum_{i=1}^{n}P(A_i)\times P_{A_i}(B)$

***7-4*- نظرية بابيز:** بنفس شروط دستور الاحتمالات الكلية، من أجل كل $\mathcal{H}$لدينا:

$$P_B(A_i)=\frac{P(A_i\cap B)}{P(B)}=\frac{P(A_i)\times P_{A_i}(B)}{\sum_{i=1}^{n}P(A_i)\times P_{A_i}(B)}$$

ملاحظة: يمكن اعتبار دستور بابيز كإعادة تقييم.

مثال: في مصنع به *3* آلات تنتج *50*% ،*30*% و *20*% من الانتاج الكلي للمصنع، نسبة التالف هي *3*% ،*4*% و*5*% من انتاج الآلات على الترتيب.

أختيرت قطعة عشوائيا من إنتاج المصنع ، إحتمال ان تكون معيبة$\times 0.2$ $P(B)=0.5\times 0.03+0.3\times 0.04+0.2$ $0.05=0.037$

إذا وجد ان القطعة المختارة تالفة فإحتمال ان تكون من انتاج الآلة الاولى هو $P_B(A_1)=\frac{0.5\times 0.03}{0.037}$

7-5- الحوادث المستقلة: ليكن (Ω, A, P) فضاء إحتمالي ، نقول عن حادثتين ✌ و👌 أنهما مستقلتان اذا وفقط اذا كان:

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

مثال: احتمال أن يصيب ✌ الهدف هو $\frac{5}{7}$ و إحتمال أن يصيب 👌 الهدف هو $\frac{7}{8}$ ، صوب كل منهما نحو الهدف مرة واحدة.

احتمال ان يصاب الهدف هو $P(A\cup B)=P(A)+P(B)-P(A\cap B)=P(A)+P(B)-P(A)P(B)=\frac{5}{7}+\frac{7}{8}-\frac{5}{7}\times\frac{7}{8}$

7-5-1- خواص: أ- $A\amalg B\Rightarrow\begin{cases}\overline{A}\amalg B\\ A\amalg\overline{B}\\ \overline{A}\amalg\overline{B}\end{cases}$

ب- $A\amalg B_1\wedge A\amalg B_2\Rightarrow A\amalg(B_1\cup B_2)$

ت- $A_i \coprod B \Rightarrow iA \coprod B; A_i \cap A_j = \emptyset$

برهـــان: $P(\overline{A} \cap B) = P(B) - P(A \cap B) = P(B) - P(A)P(B) = P(B)(1 - P(A)) = P(\overline{A})P(B)$

6-7- الاشجار البيانية: من بين الطرق المناسبة لوصف العمليات العشوائية المتتالية ما يسمى بشجرة الاحتمالات، و نستخدم حاصل الضرب لاحتمال ظهور اية نتيجة ممثلة في مسار معطى في هذه الشجرة.

تمارين الوحدة الثالثة

التمرين الاول: A، B و C حوادث من عشيرة الاحداث،

أ- عبر بالترميز المجموعي عن الحوادث التالية:

1-الحادثA فقط أنجز 2- أحد الأحداث على الاقل قد تحقق

ب- أحسب : $Card(A \cap B \cap C)$

التمرين الثاني: أ- بسط مايلي:$\frac{11!}{9!2!}$

ب- احسب: $C_n^0 + C_n^1 + \cdots + C_n^n$

ج- عين قيم العدد الطبيعي n حيث:$A_n^3=56$

التمرين الثالث: *1*- ماهو عدد الاعداد ذات *3* ارقام مختلفة مثنى مثنى يمكن تشكيلها باستعمال الارقام *2□3□5□6□7□9*

من بين هذه الاعداد ، ما عدد الاعد اد:

أ-الزوجية ب-مضاعفات العدد *5* ج- الاصغر من *500*

التمرين الرابع: نريد وضع *4* كتب رياضيات مختلفة ،*6* كتب دينية و *4* كتب تاريخ في رف. ماهو عدد الطرق لترتيب الكتب :

أ- حسب المواد ب- كتب الرياضيات فقط مرتبة معا

* **التمرين السادس:** A وB حادثتان حيث

$$P(\overline{A \cap B}) = \frac{3}{4}, P(A) = \frac{3}{8}; P(B) = \frac{1}{3}$$

أحسب احتمالات الاحداث التالية:

$\overline{A}$□$\overline{A} \cap \overline{B}$□$\overline{A} \cap B; A/B$

التمرين السابع: يحتوي صندوق على 4 كرات خضراء و 4كرات بيضاء و كرتين حمراوين، نسحب في آن واحد 3كرات عشوائيا. احسب احتمال :

1- الحصول على 3كرات مختلفة اللون
2- الحصول على 3كرات من نفس اللون.

التمرين الثامن:

ثلاث حوادث مستقلة حيث A،B وC:

$$P(A) = 0.3; P(B) = 0.4; P(C) = 0.5$$

احسب قيمة مايلي:

1- احتمال وقوعها معا
2- احتمال وقوع واحدة على الاقل
3- احتمال عدم وقوع اي منها

التمرين التاسع: شركة تأمينات تقسم زبائنها الى 3 اصناف ، اشخاص ذوي الخطر الضعيف ، وذوي الخطر المتوسط و اصحاب الخطر العالي

فحسب احصائيات سابقة فان احتمال وقوع حادث لأحدهم خلال سنة هي على الترتيب 0.05،0.15 و0.3.

وتشير التقديرات الى ان نسب هؤلاء الاشخاص هي على الترتيب 20%، 50% و 30%.

1- عين نسبة الاشخاص الذين يقع لهم حادث على الاقل سنويا.
2- اذا اختير شخص وعلم انه تعرض لحادث فما احتمال انه من المجموعة الاولى.

***التمرين العاشر:** احتمال ان تقوم مصلحة البريد بسحب الرسائل من صندوق ما في اليوم ذي الرتبة n من السنة هو ½ في حالة ما اذا كانت قد سحبت الرسائل بالامس وهو 1 في الحالة الاخرى. ليكن p_n احتمال ان تقوم مصلحة البريد بسحب الرسائل من الصندوق في اليوم ذي الرتبة n.

احسب p_n اذا علمت ان $p_1 = 1$

الملخص:

الإحصاء الوصفي هو مجموعة طرق لوصف الخصائص الرئيسية لمجموعة بيانات كمية أو نوعية باستخدام الجداول والمخططات البيانية،يُعدّ الإحصاء الوصفي أحد أهم الأدوات التحليلية في البحث العلمي، حيث يهدف إلى وصف وتلخيص البيانات المتاحة بطرق كمية وموضوعية. يعتبر الإحصاء الوصفي جزءًا أساسيًا من عملية التحليل البياني والاستنتاج العلمي، حيث يساعد الباحث على فهم الظواهر وتفاصيلها بشكل دقيق ومفصل.

تهدف هذه المطبوعة إلى استكشاف الإحصاء الوصفي وأهميته في البحث العلمي. سيتم تناول تعريف الإحصاء الوصفي وشرح أهم الأساليب المستخدمة فيه، . سنتناول أيضًا الأدوات الإحصائية المستخدمة في تحليل البيانات الوصفية، مثل المتوسط والانحراف المعياري والترددات وحساب الاحتمال.

الكلمات المفتاحية: الإحصاء الوصفي- مقاييس النزعة المركزية- مقاييس التشتت- مقاييس الشكل- التحليل التوفيقي- الاحتمال الشرطي

Résumé

Les statistiques descriptives sont un ensemble de méthodes permettant de décrire les principales caractéristiques d'un ensemble de données quantitatives ou qualitatives à l'aide de tableaux et de graphiques. Les statistiques descriptives sont l'un des outils analytiques les plus importants dans la recherche scientifique, car elles visent à décrire et à résumer les données disponibles de manière quantitative. et des moyens objectifs. Les statistiques descriptives constituent une partie essentielle du processus d'analyse graphique et de conclusion scientifique, car elles aident le chercheur à comprendre les phénomènes et leurs détails de manière précise et détaillée

Cette publication vise à explorer les statistiques descriptives et leur importance dans la recherche scientifique. La définition des statistiques descriptives sera discutée et les méthodes les plus importantes utilisées seront expliquées. Nous couvrirons également les outils statistiques utilisés dans l'analyse des métadonnées, tels que la moyenne, l'écart type, les fréquences et les distributions de probabilité.

Mots clés : Statistiques descriptives - mesures de tendance - mesures de dispersion - mesures de forme - analyse combinatoire - probabilité conditionnelle.