كلية التربية
الدراسات البيئية

المفردات الخاصة بمادة الإحصاء الجغرافي

الفصل الاول / مقدمة عامة عن الاحصاء

1-1 علم الاحصاء --- النشأة والتطور .
1-2 اهمية علم الاحصاء ومجالات تطبيقه .
1-3 مفهوم الاحصاء .
1-4 مراحل العملية الاحصائية .
1-5 زوايا النظر الى علم الاحصاء .
1-6 اقسام علم الاحصاء .
1-7 البيانات الاحصائية .

الفصل الثاني / الاحصاء وعلاقته بالجغرافية

2-1 مقدمة
2-2 الاسباب التي دعت الى دراسة الجغرافي الاحصاء .
2-3 بعض التعاريف المهمة في الاحصاء الجغرافي .
2-4 المتطلبات الاساسية لاجراء بحث .
2-4-1 المصادر التي يتم عن طريقها جمع المعلومات لاجراء البحث .
2-4-1 مصادر اولية
2-4-1-1 العينات
2-4-1-2 الاستبيان
2-4-1-3 المقابلات
2-4-2 مصادر ثانوية
2-5 تصنيف البيانات
2-5-1 اهمية التصنيف
2-5-2 انواع التصنيف
2-6 عمليات التصنيف

الفصل الثالث / التوزيعات التكرارية ومقاييس النزعة المركزية ومقاييس التشتت

3-1 التوزيع التكراري.

3-1-1 تعاريف موجزة عن التوزيع التكراري

3-1-2 مميزات التوزيع التكراري

3-2 مقاييس النزعة المركزية

3-2-1 الوسط الحسابي

3-2-2 المنوال

3-2-3 الوسيط

3-3 مقاييس التشتت

3-3-1 المدى

3-3-2 الانحراف المتوسط

3-3-3 التباين والانحراف المعياري

3-4 دقة تمثيل العينات للمجتمع

3-4-1 الخطأ المعياري

الفصل الرابع / الاحتمالية

4-1 الاحتمالية وتعريفها

4-2 قواعد الاحتمالات

4-3 تطبيقات جزافية حول نظرية الاحتمالات.

4-4 المعاينة

4-5 ايجاد حدود الثقة

الفصل الخامس / تحليل الأنماط المكانية والانماط الشبكية.

5-1 معدل التمركز المكاني

5-1-1 بيانات غير مبوبة

5-1-2 بيانات مبوبة

5-2 الوسيط المكاني

5-3 المسافة المعيارية

5-4 التمركز النسبي

5-5 تحليل الجار الأقرب
6-5 تحليل المربعات الصغرى
7-5 تحليل الانماط الشبكية (الانماط الخطية)
الفصل السادس / تحليل العلاقة والاشتراك المكاني.
6-1 مقدمة
6-2 مقاييس تحليل العلاقة
6-2-1 اختبار مربع كاي χ^2
6-2-2 نسبة التقاطع
6-2-3 معامل يول
6-2-4 معامل فاي
6-2-5 معامل كافا
6-2-6 معامل سبيرمان
6-2-7 معامل بيرسون
6-2-8 تحليل الانحدار
6-2-8-1 معادلة الانحدار (نموذج الانحدار)
6-2-8-2 تقدير معالم الانحدار
6-3-3 التنبؤ

سنة الولادة	٨١	٨٢	٨٣	٨٤	٨٥	٨٦	٨٧	المجموع
العدد	٢	٥	٢٢	٦٠	١٥	١٠	٢	١١٨

ب- التبويب الجغرافي :- وهو جمع البيانات وتصنيفها الى مجموعات ظاهرة بوحدة ـــــــ جغرافية معينة مثل : محافظة ، قضاء ، ... الخ

مثال (2) : توزيع طلبة الجامعة المستنصرية حسب المحافظات :

المحافظة	بغداد	بصرة	موصل	الانبار	ديالى	المجموع
العدد	٩٨	٢٥	١٥	١٠	٢	١٥١

ج- التبويب الكمي :- هو تصنيف البيانات في جداول كل مجموعة منها ظاهرة بوحدة ـــــــ كمية معينة كوحدات ، الوزن ، الطول ، ساعات العمل ، الاجور ...

مثال (3) :-

توزيع عدد العمال حسب اجورهم اليومية وكالاتي :-

الاجرة اليومية بالدينار العراقي	اقل من ١٠	١٠ - ١٥	١٥ - فاكثر	المجموع
عدد العمال	١٧٠	٩٠	٦٥	٢٢٥

د- التبويب على اساس صفة معينة :- وهو تصنيف البيانات وترتيبها في جداول على أساس ـــــــ كل مجموعة منها تشترك بصفة معينة كالحالة الاجتماعية ، الحالة الثقافية ... الخ .

مثال (4) :-

تصنيف طلبة قسم الجغرافية حسب الجنس :-

الجنس	ذكور	اناث	المجموع
العدد	١٢٠	٩٨	٢١٨

٤)- حساب المؤشرات الاحصائية لمقدرات العالم لمجتمع البحث او الدراسة.
٥)- تفسير النتائج بعد تحليل معطيات الدراسة.
٦)- عملية اتخاذ القرار بشأن فرضيات البحث او الدراسة.

1-5 زوايا النظر الى علم الاحصاء في البحث العلمي:-

مما سبق يمكن النظر الى علم الاحصاء من ثلاث زوايا متكاملة هي:-

١)- الطريقة الاحصائية Statistical Method:-

وهي مجموعة من الأساليب والمعادلات الرياضية والقوانين والاجراءات التي تفيدنا في بحث اي موضوع احصائي، وتمتاز هذه الطريقة بكونها تهيئ اسلوباً "علمياً" محايداً" دون ان يكون للباحث اي تدخل او اختيار تجاه نتائج البحث، هذه الميزة جعلت الطريقة الاحصائية تلاقي اهتمام الكثير من الباحثين في مجالات المعرفة المختلفة مثل تعطينا كيفية اختيار العينة Sample.

٢) النظرية الاحصائية Statistical Theory:-

وهي الدعامة العلمية التي تقوم عليها الطريقة الاحصائية من خلال استخدام النظريات التي تفسر القوانين والمعادلات والأساليب المستخدمة في الاحصاء والملائمة للبحث العلمي.

٣)- الاحصاء التطبيقي Applid statistics:-

هو تطبيق الاحصاء في البحث تحت الدراسة وحسب مجالات تطبيقه سواء كان ذلك في مجال العلوم الصرفة او العلوم الانسانية وغيرها.

1-6 اقسام علم الاحصاء:-

يمكن تقسيم علم الاحصاء بصورة عامة الى قسمين رئيسيين هما:-

أ)- الاحصاء الوصفي Descriptive Statistics:-

ويتضمن الطرق والأساليب المستخدمة في جمع البيانات والمعلومات عن ظاهرة معينة او مجموعة ظواهر وكيفية تنظيم وتصنيف وتبويب هذه البيانات مع امكانية عرضها في جداول ورسومات بيانية وحساب بعض المقاييس الاحصائية لها.

ب)- الاحصاء الاستدلالي Inferential Statistics:-

ويشمل الطرق الاحصائية التي تؤدي الى عمل استنتاجات او استدلالات حول المجتمع الذي جمعت منه البيانات ويضم هذا القسم فرعين رئيسيين هما التقدير Estimation واختبار الفرضيات testing hypotheses.

1-7 البيانات الاحصائية:-

هناك نوعين من البيانات الاحصائية هما:-

أ)- بيانات منفصلة Discrete Data:-

وهي المعالم التي يمكن الحصول عليها من عملية العد وتأخذ اي قيمة رقمية صحيحة (٩، ٥، ١٠٠، ...) كعدد الوحدات الانتاجية، عدد الاقسام، عدد الموظفين - الخ.

ب)- بيانات متصلة Continuous Data:-

وهي المعالم التي يتم الحصول عليها من عمليتي القياس والوزن والتي يمكن ان تأخذ أي قيمة في مدى معين (ارقام كسرية او عشرية غير صحيحة) مثل ٩٥،٥، $\frac{1}{2}$، ... الخ كأوزان الطلبة، اطوالهم، ... الخ ويمكن عدها بوحدات صغيرة قابلة للتجزئة.

"الفصل الثاني"
الاحصاء وعلاقته بالجغرافية

2-1 كما نعلم ان الجغرافية علم يختص بدراسة العلاقة بين الانسان والبيئة التي يعيش فيها، وعند شرح الخصائص المكانية والعلاقات المتداخلة بالظواهر الجغرافية يتم الاعتماد على البيانات الرقمية. وعلم الاحصاء هو علم من علوم يفسر هذه الظواهر ويقدرها وكذلك المكانية التنبؤ بها لتتم استغلالها.

2-2 الاسباب التي من اجلها يستخدم الجغرافي الطرق الرياضية والاحصائية في دراسته:-

---

1) الجغرافية علم تصانيف ويستفاد الجغرافيون من مختلف الطرق الاحصائية والرياضية لتقسيم العالم الى الاقاليم ومجموعات واصناف.

2) يلجأ الجغرافيون الى استخدام العينات في بحوثهم فتكون اكفأ الطرق الاحصائية لوصف البيانات التي تم جمعها والاستدلال من خلالها عن الكلية والاشمال.

3) يحتاج الجغرافيون الى اختبار الفرضيات الخاصة بملاحظاتهم من البيئة التي يعيشون فيها سواء كانت بيئة طبيعية او اجتماعية من خلال الاعتماد على طرق احصائية معينة وخصيصة لهذا الغرض او وضع نموذج رياضي خاص.

4) الجغرافيون يميلون الى استخدام الطرق الاحصائية لدراستهم الظواهر الجغرافية بسبب مرونة الواقع الجغرافي وعدم ثبوته لكونه قابل للتغير المستمر ويعتمد على الاحتمالية في التغير.

5) يدرس الجغرافي الظواهر وهو بذلك يهتم بدراسة العلاقات والعوامل التي شكلت هذه الظواهر او التي تؤثر عليها ومعرفة مقدار التباين بين الظواهر (أو بين) الاماكن. وتحدد الاقاليم الجغرافية في ضوء العلاقات او الظواهر او التباين وليس بالمكانية لقيام

يمثل هذه الدراسات دون استخدام الطرق الاحصائية .

٦) يحتاج الجزائي الى اعتماد الارقام التي تمثل ظاهرة معينة لوصف تلك الظاهرة اي انه يعتمد احصائية .

2-3 بعض التعاريف المهمة في الاحصاء الجزائي :-

المتغير Variable :-

هو اي ظاهرة تظهر اختلافات بين مفرداتها ويرمز له بأي رمز مثل X ، Y ، ... الخ .

عند جمع بيانات حول ظاهرة ما فاننا نرمز للظاهرة بالرمز (X) وكل مرة او مشاهدة مرفقا رمز لها بالرمز Xi فمثلا" عند دراسة اطوال الطلبة في احدى الجامعات فاننا نرمز لصفة الطول بالرمز (X) وطول اي طالب بالرمز (Xi) ويسمى بالمشاهدة او المفردة وان قيمة Xi تختلف من طالب الى آخر ولهذا نقول بأن X متغير .

والمتغيرات تنقسم الى :-

١) متغيرات وصفية او نوعية Qualitative variables :-

وهي تلك الظواهر او الصفات التي لا يمكن قياسها مباشرة بالارقام العددية مثل صفة لون العين (اسود ، بني ، ازرق) والحالة الاجتماعية (متزوج ، غير متزوج) والجنس (ذكر ، انثى) ... الخ .

٢) متغيرات كمية Quantitative variables :-

وهي تلك الظواهر او الصفات التي يمكن قياسها مباشرة بارقام عددية مثل :- صفة الوزن ، الطول ، العمر ، كمية المحصول ، ... الخ وتنقسم المتغيرات الكمية الى قسمين هما :-

أ) متغيرات متقطعة (او منفصلة) Continuous variables :-

ان المتغير المستمر هو المتغير الذي تأخذ المشاهدة او المفردة فيه اية قيمة رقمية في فترة معينة، فلو افترضنا بان اطوال طلبة جامعة ما تتراوح بين ١٢٠.٥ و ١٧٠.٥ سم فنقول بأن ($130.5 \leq X \leq 170.5$) اي ان المتغير X يمكن ان يأخذ اية قيمة

تقع بين ١٢٫٥ سم و ١٧٫٥ سم ، وكأمثلة اخرى على المتغيرات المستمرة هي درجة الحرارة ، الزمن ، ... الخ لانه يمكن قياسها باجزاء صغيرة جدا" وتأخذ اية قيمة تقع في حدود معينة .

وبصورة عامة فان كل البيانات التي تقاس تعتبر بيانات لمتغير مستمر .

ب)- متغيرات متقطعة (او منفصلة) Discrete Variables :-

المتغير المتقطع هو المتغير الذي تأخذ المشاهدة او المفردة منه قيما" متباعدة او متقطعة غير مستمرة ، فلو فرضنا ان عدد افراد الأسرة في اربع عوائل هي ٢ ، ٤ ، ٧ ، ١ متقول بأن ، X = 2, 4, 7, 1 .

وبصورة عامة فان كل البيانات التي تحقل عددا من العد تعتبر بيانات لمتغير منفصل .

المجتمع Population :-

هو عبارة عن جميع القيم او المفردات التي يمكن ان يأخذها المتغير . والمجتمع اما ان يكون :-

ا)- مجتمع" محدود Finite Population

اي ممكن حصر عدد مفرداته كما هو الحال في اطوال طلبة الجامعة المستنصرية او عدد الوحدات الانتاجية في مصنع ما في يوم معين ، ... الخ .

ب)- مجتمع غير محدود Infinite Population :

وهو المجتمع الذي من الصعب او المستحيل حصر عدد مفرداته مثل :- مجتمع نوع سمك معين في نهر دجلة ، عدد البكتريا في حقل ما ، ... الخ .

العينة Sample :-

هي جزء من المجتمع وهي عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع .

ان دراسة المجتمع ككل قد تكون صعبا" او تحتاج الى وقت وجهد ومال لذا فقد استعيض عن دراسة المجتمع بدراسة العينة وصفاتها وعليه نستطيع ان نستنتج خواص المجتمع الأصلي الذي اخذت منه هذه العينة .

البيانات Data :

هي عبارة عن مجموعة الأرقام التي يجمعها الباحث عند دراسته لمشكلة ما.

2-4 المتطلبات الأساسية لإجراء بحث او لاختباره احصائياً :

هناك اعتبارات كثيرة يتوقف عليها اجراء البحث حقيقة لكل اجراء يتوجب على الباحث الأخذ بنظر الاعتبار مسألة الحصول على البيانات والمعلومات بأقصر وقت واقل جهد وأوطأ كلفة وهذا يعني انه يجب مراعاة ما يلي :-

1) توفر المعلومات ، حيث لا يمكن القيام باجراء اي بحث مالم تتوفر معلومات عن الظاهرة قيد الدراسة فالمعلومات التي يجمعها الكرافي يجب ان تكون حقيقية وليست افتراضية .

2) بيان دقة هذه المعلومات : من الضروري جداً بيان دقة المعلومات وتحديد درجة الثقة بها كي تنعكس هذه الدقة على النتائج التي يتوصل اليها الباحث وذلك لأن البيانات الصحيحة والدقيقة تعطي نتائج منطقية وصحيحة .

3) اختيار العينة عند معالجة ما : بحيث تمثل مجتمع الدراسة افضل تمثيل للحصول على افضل دقة ممكنة بالإضافة الى ان حجم العينة يجب ان يكون ضمن الحدود المعقولة بحيث لا يكون كبير يؤدي الى صرف وقت وجهد وموارد مادية كبيرة ولا تكون صغيرة بحيث يكون ذلك على حساب دقة النتائج .

كما يجب ان تكون موزعة بصورة عادلة على المجتمع لتمثله بصورة صحيحة لكون العينة تعتبر وسيلة لجمع المعلومات .

٤)- تحديد الغرض من البحث، يجب ان يكون الهدف من البحث محدد بشكل واضح ودقيق بحيث يمكن التركيز على اوجه الاستفادة من نتائجه.

٥)- تحديد امكانية التنفيذ الفعلي للبحث :- من الضروري جدا "تحديد المتطلبات التي تستلزم عملية تنفيذ البحث وبشكل واضح ودقيق كالموارد المالية المطلوبة عند التنفيذ والامكانيات البشرية اللازمة المطلوبة في تحقيق بعض فقرات البحث.

٦)- تحديد اطار البحث، اي تحديد المجتمع الاحصائي على نحو واضح ودقيق.

المصادر التي يتم عن طريقها جمع المعلومات والبيانات :-

١)- مصادر اولية :-

_____ وتشمل :-

أ)- العينات.
ب)- الاستبيان.
ج)- اجراء المقابلات الشخصية.

٢)- العينات، (تم تعريفها سابقاً)

ب)- الاستبيان، هو عبارة عن استمارة يتم من خلالها جمع البيانات والمعلومات عن مفردات (او بعض المفردات) مجتمع الدراسة وذلك عن طريق مواجهة الباحث الشخصية للمفردة الاحصائية.

شروط تصميم الاستبيان الجيد :-

١)- اعداد مقدمة ايضاحية تكتب في بداية الاستمارة بأسلوب بسيط وواضح يوضح

خالياً الهدف من البحث وأغراضه وبالشكل الذي يجعل اجابة الفرد على فقرات الاستمارة بصورة صحيحة ودقيقة دون اي خلل.

٣- ان تكون فقرات الاستمارة (أسئلة الاستبيان) متسلسلة ومترابطة بحيث ان كل سؤال خالياً يخدم غرض معين.

٢- ان تكون الاسئلة واضحة من حيث اللفظ والغرض وليس خالياً اي غموض أو قصور في جانب معين.

٤- ان تكون الاسئلة متوسطة العدد بحيث ان الفرد المعني لا يمل عند الاجابة عليها وحتى امكانهم في الحصول على بيانات دقيقة ومتكاملة.

٥- ان تكون الاسئلة بالشكل الذي تكون اجابة الفرد عليها محددة وقصيرة بحيث لا تزيد عن دقيقة واحدة وتفضل الاسئلة التي يتم الاجابة عليها بـ (نعم) او (لا) أو التأشير بإشارة معينة كـ (✓) أو (X).

٥- اجراء المقابلات الشخصية: تعني بالمقابلات هي مجموعة لقاءات يجريها الباحث مع مجموعة معينة من الاشخاص المعنيين بموضوع البحث لغرض الحصول على معلومات تخدم البحث او الدراسة.

وشروط المقابلة الشخصية ان تكون جميع الاسئلة غرضها مباشرة وواضحة ومرتبطة بجميع افراد العينة التي سيتم مقابلتهم.

أسس اجراء المقابلات:-

هنالك مجموعة من الأسس العامة على الباحث اتباعها لغرض اجراء المقابلات وهي:-

١- كتابة تقرير عن الزيارة (المقابلة) لاكتشاف النواقص والاخطاء وتجاوزها او تكملتها وتشبيث الحقائق قبل نسيانها او اختلاطها مع غيرها.

٣)- ان تتم المقابلة بين الباحث والشخص المعني لكون معظم البحوث الجزائية بحثية بمعلومات حقيقية .

٢)- يفضل ان تكون المقابلة في مكان مناسب (المكتب ، المدرسة ، البيت ، ـــ)

٤)- تقديم معلومات وافية عن البحث والهدف منه والجهة المسؤولة عنه و يفضل كتابتها في رسالة وترسل الى الجهة قبل مدة كافية .

٥)- ان تكون لغة المقابلة سهلة وواضحة ومفهومة فكلما نال الباحث معلومات ادق وذات صلة مباشرة ببحثه تسهل عليه عملية تحليل الاجابات والربط بينها وبين الاستنتاج منها لذلك يفضل تفسير وتعريف بعض المفردات الصعبة الموجودة في اسئلة المقابلة .

٥)- المصادر الثانوية :- وتكون غيرها المعلومات قد جمعت من قبل مؤسسات او اشخاص غير الباحث ومن اهم المصادر الثانوية :-

١)- المصادر التاريخية :- قد يحتاج الجغرافي لمعرفة الخلفية التاريخية لظاهرة بحثية عند دراسته لها ليتمكن من رسم تطورها فالباحث الجغرافي يحتاج الى الاطلاع على السجلات القديمة والحديثة لبعض الدوائر الرسمية مثل البلدية والاحصاء ، الاحوال المدنية ، الانواء الجوية ، ـــ الخ .

٣)- الجهاز المركزي للاحصاء :- تتوفر لديه احصاءات عن السكان والنشاطات الاقتصادية المختلفة ، وعن المتطلبات لعملية التخطيط الحضري والاقليمي بالاضافة الى التعدادات العامة كل عشر سنوات .

٢)- مديرية المساحة العامة :- تتوفر لدى هذه المديرية خرائط لجميع ارجاء القطر وكذلك الصور الجوية .

٤)- مديرية الري :- يتم الحصول منها على سجلات القراءات عن منسوب المياه في الانهار

وتوعيه المياه ، وكذلك معلومات للمحيين بالزراعة والجيومورفولوجي ، اضافة الى المحيين بالمدينة والخدمات البلدية فيما .

٥) - محطات الرصد الجوي ومديرية الانواء الجوية : حيث يتم الحصول منها على معلومات عن المناخ والجو وعناصرهما اذ تصدر مديرية الانواء تقارير عن القراءات التي سجلتها اجهزتها وللسنوات المحددة .

٦) - مديرية المساحة العامة : تتوفر لدى هذه المديرية خرائط بجميع ارجاء القطر وكذلك الصور الجوية .

٧) - دوائر الزراعة : تتوفر لديها معلومات تفصيلية عن الانتاج الزراعي وانتاجية الارض .

٨) - الوزارات والمؤسسات العلمية : تحتفظ الوزارات بتقارير دوائرها المختلفة في المكتبات الخاصة بهذه الوزارات .

ملاحظة : اذا صعب الباحث نشاطه مع جهة واحدة للمعلومات فان بحثه سيكون ناقصاً اذ لابد من تنويع المصادر ، وان تكون ذات صلة بموضوع بحثه وان كان هناك تقارضاً .
اما اذا اعتمد الباحث على المصادر الاولية والثانوية في آن واحد فانه سيصل بالنتيجة على نتائج دقيقة وصحيحة .

2-5 تصنيف البيانات Classification of Data :-

تعتبر عملية التصنيف المعلومات مرحلة مهمة في تنظيم المعرفة و ترتيبها ووضع النظريات ، ولأهمية التصنيف في الجغرافية فالمختص يسمى بعلم التصانيف ، وهناك هدفان لعملية التصنيف هما :-

هدف عام :-

_____ وهو اعطاء مسميات للمفردات لتكون اعضاء في مجاميع مثل تصنيف المدن حسب أساسها الاقتصادي .

هدف خاص :-

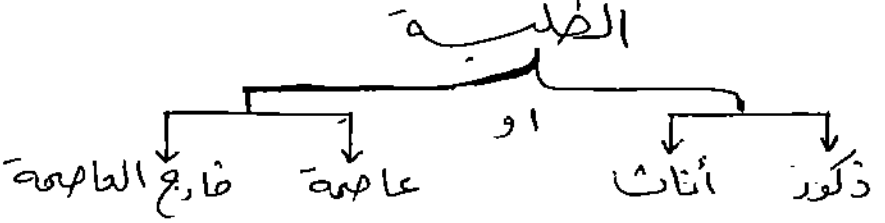
_____ الرجوع بمعلومات يُعتمد عليها في الاستدلال العلمي .

تعريف التصنيف :- هو تقسيم البيانات الى مجاميع بصورة نظامية وفق أسس معينة يحدد الخصائص والعلاقات المشتركة .

2-5-1 طرق التصنيف :-

_____ يتم تصنيف المعلومات بطريقتين :-

أ)- التصنيف على اساس التشابه :- وهو التصنيف المنطقي للمعلومات من خلال خطوات متسلسلة تعتمد على معايير واضحة ومثال ذلك تصنيف حركة من الطلبة الى اناث وذكور او الى سكنة العاصمة او خارج العاصمة :-

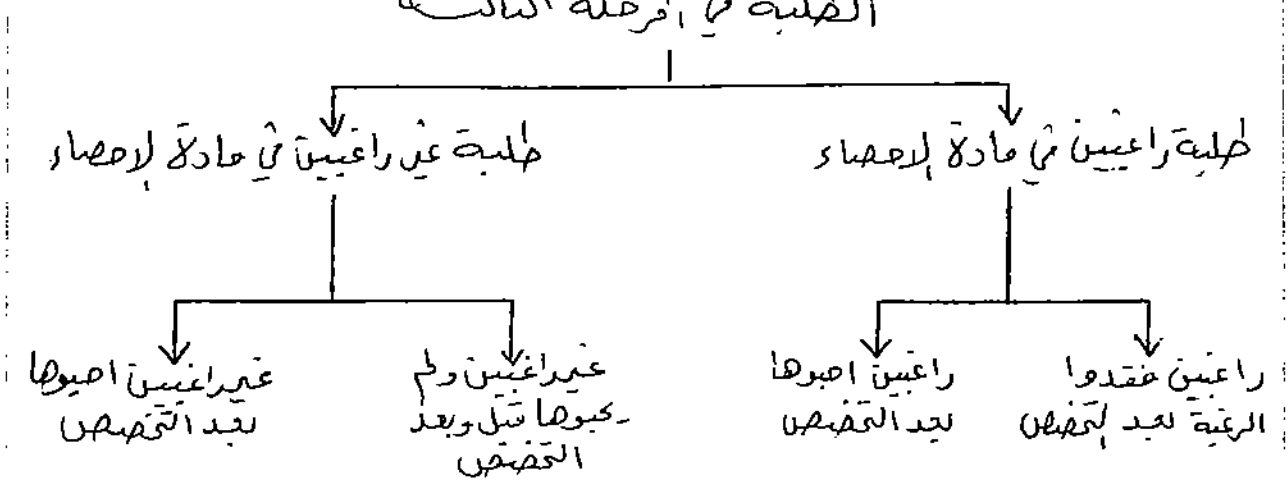


ويتميز تصنيف هذه العينة بأنه :-

١)- تصنيف شامل ، اي لا يوجد شخص خارج اطار التصنيف الجميع اما اناث او ذكور ، واما من العاصمة او من خارجها .

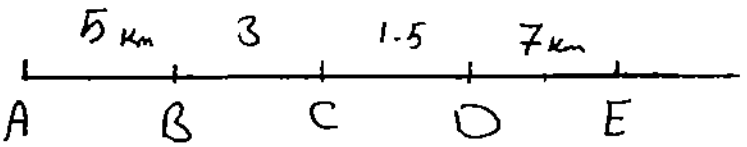
٢)- تعتبر الأصناف متميزة حيث لا يوجد من هو ضمن مجموعتين في آن واحد .

ج) تصنيف المجتمع الى مجاميع :- وهو عبارة عن تجميع المعلومات الى مجاميع منفصلة عن بعضها اي على أساس محدد فمثلاً يتم تصنيف الطلبة في هذه الطريقة الى طلبة راغبين في دراسة الجغرافية وطلبة غير راغبين في دراسته وهنا تم تصنيفهم حسب ظاهرة واحدة .
ويمكن تصنيفهم حسب ظاهرتين كما في الشكل ادناه :-



١. التصنيف الرقمي لمتغير واحد :-
وهو تصنيف المعلومات الى مجاميع بحيث ان اي قيمة في المجموعة يجب ان تكون اقرب ما يكون الى مجموعتها وان يكون التباين بينها وبين بقية القيم في المجاميع الاخرى واضحاً ، وبمعنى اخر ما يمكن بين افراد المجموعة الواحدة وكبير تباين بين المجاميع " .
عملية التصنيف هنا تعتمد على المسافة بين القيم اساساً في التصنيف .
مثال :- بالاعتماد على مثال قدمه العالم ABLER بهذا التصنيف نفترض وجود خمس قرى المسافة بينها كالاتي :-

بين A و B = 5 Km
بين B و C = 3 Km
بين C و D = 1.5 Km
بين D و E = 7 Km



اولاً تنظيم المسافات بين هذه القرى على شكل مصفوفة التي سماها ABLER مصفوفة التشابه .

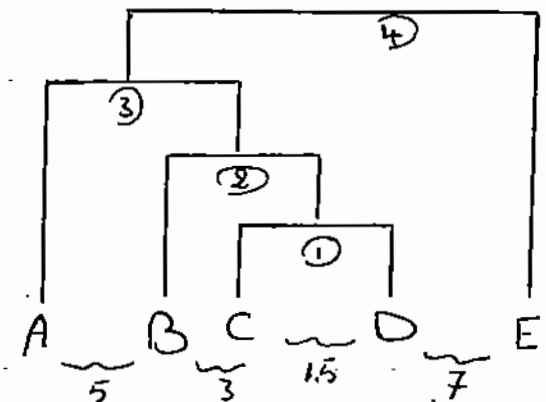
	A	B	C	D	E
A	Zero	5	8	9.5	16.5
B		Zero	3	4.5	11.5
C			Zero	1.5	8.5
D				Zero	7
E					Zero

لتطبيق البيانات في اعلاه وتحليلها الى مجاميع تعتمد على الطرق التالية :-

2-5-2 انواع التصنيف :-

1) - العلاقة المفردة :- في هذه الطريقة يتم اختيار اقصر المسافة بين اقرب نقطتين ——— (اقرب قريبين مثلاً) واعتبار النقطتين كمجموعة واحدة ونستطيع رسم هذه العلاقة على شكل يسمى بشجرة المسافات :-

- اقصر مسافة اولى هي بين (C - D) = 1.5
- " " ثانية " " (B - CD) = 3
- " " ثالثة هي بين (A - BCD) = 5
- " " رابعة " " (E - ABCD) = 7



2) - طريقة المعدل :- في هذه الطريقة يتم استخراج معدل المسافة بين النقاط ——— اي ايجاد المركز الوسط بينها .

تطبق هذه الطريقة على المثال السابق ثم نعمل على ايجاد المسافة بين النقاط على اساس مجاميع اي ايجاد النقطة الوسط لكل مجموعة وكما يلي :-

١- المجموعة الاولى (C - D) ⇐ النقطة الوسط = $\frac{1.5}{2}$ = 0.75

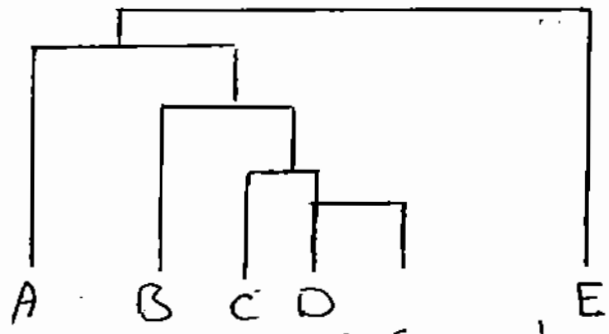
٢- المجموعة الثانية (C D - B) ⇐ النقطة الوسط = $\frac{3+1.5}{3}$ = 1.5

٣- المجموعة الثالثة (A - B C D) ⇐ ر ر = $\frac{5+3+1.5}{4}$ = 2.4

٤- المجموعة الرابعة (E - A B C D) ⇐ ر ر = $\frac{7+5+3+1.5}{5}$ = 3.24

٢- العلاقة الكاملة :-

تعتمد هذه الطريقة على وصفوفة التشابة فتحدد اولاً المجموعة التي تضم اقل مسافة وهي (C - D) البالغة بينهما (1.5) كم اي المجموعة التي لاتزيد عن (2) كم مثلاً والتي نستطيع ربطهما (C D - B) ونستطيع ربط A معهما ثم نربط المجاميع مع E .



٢-٥ سلبيات طرائق التصنيف :-

١- معظم طرائق التصنيف المعتمدة بسيطة ولاتؤدي الى تحليل رياضي للتتابع .

٢- تطورت هذه الطرق من علوم عدة لذا جعلت وفقا خصائص هذه العلوم وما يصلح لتخصص قد لايصلح لتخصص آخر .

٣- تؤدي طرائق التصنيف المختلفة الى نتائج مختلفة ولنفس البيانات .

٤- يفترض ان تؤدي طرائق التصنيف الى دراسة التركيب الداخلي للبيانات الا انها في الواقع تفرض تركيبا عليها .

٥- الفرق بين التحليل العاملي والطرائق الاخرى انه يحلل المتغيرات لتصنيف المناطق الجغرافية على اساسها بينما تصنف طرائق التحليل الاخرى المناطق الجغرافية مباشرة

H.W (1)

س/ عين نوع المتغير (مستمر ام متقطع) في كل من الحالات التالية :-

م- درجات الحرارة المقاسة كل نصف ساعة في محطة الأنواء الجوية في بغداد .

ب- عدد الكتب الموجودة على رفوف مكتبة كلية التربية .

ج- عدد السيارات المباعة يومياً من الشركة العامة للسيارات .

د- الدخل السنوي لموظف في احدى دوائر الدولة .

هـ- عدد ملمترات كمية الامطار النازلة في مدينة بغداد خلال اشهر السنة .

و- سرعة السيارة بالأميال في الساعة .

ز- عدد الطلبة المقبولين في الجامعة المستنصرية في عدة سنوات .

ح/ ما المقصود بالطريقة الاحصائية في البحث العلمي ؟ وما هي مراحل هذه الطريقة .

ك/ ما هي الخطوات الرئيسية التي يفترض اتباعها عند تصميم بحث معين ؟

ع/ ما المقصود بالمجتمع الاحصائي وما هي انواعه ؟

مع تحديد المفردة الاصلية التي سيتم جمع البيانات منها

س/ ما هي الوسيلة المفضلة في جمع البيانات لكل حالة من الحالات التالية ؟

م- دراسة تطور القبول في الجامعات العراقية .

ب- استطلاع رأي المواطنين بالخدمات الصحية المقدمة لهم .

ج- فحص فصيلة الدم لدى مجموعة من الافراد .

د- دراسة اوجه نفقات العائلة العراقية على السلع الاستهلاكية .

هـ- دراسة توقعات عدد سكان العراق حتى عام ٢٠١٠ .

و- اعداد دراسة حول المستوى العلمي لطلبة الجامعة المستنصرية من وجهة نظر الاساتذة .

س حدد نوع المجتمع (محدود ام غير محدود) مع تحديد المصدر الملائم لجمع البيانات والمعلومات لكل حالة من الحالات التالية :-

م- لتقدير كمية الأسماك الموجودة في بحيرة وكان .

ن- تحديد العوامل المسببة لمرض انفلونزا الطيور في المناطق الشمالية من الجزائر .

ع- لتقدير حجم الطلب على السيارات خلال عام ٢٠٠٨ .

ي- لتقدير حجم الطلب على الحليب للسنوات ١٩٩٨ - ٢٠٠٨ .

ه- حصر سكان الجزائر عام ٢٠٠٩ .

و- انتخاب اعضاء المجلس البلدي في الجزائر .

ز- التأكد من صلاحية اجهزة التلفزيون المنتجة في شركة الصناعات الالكترونية للتسويق .

ح- توقعات فوز منتخب كرة القدم قبل بدء المباراة .

ط- اعداد دراسة حول الاسباب الموءدية لانتشار ظاهرة التدخين في الجزائر .

ي- التعرف على المشاكل العلمية لطلبة الصف الثالث جغرافية .

ك- تحديد انتاجية الدونم الواحد من الحنطة في المحافظات الجنوبية من الجزائر .

ل- سحب قطرة دم من جسم الانسان لغرض تحليلها في مختبر .

م- تحديد العوامل المؤثرة في الاصابة بامراض الكلية في الجزائر .

ن- اختبار فاعلية دواء معين للشفاء من مرض معين .

ف- اختبار جودة اللبن المنتج في مصنع للألبان شهر تشرين الثاني .

ج ماهو الهدف من عملية تصنيف وتبويب البيانات ؟

ل اوجز ثلاثة امثلة تطبيقية عن كل من التبويب على الأساس :-

م- زمني ن- جغرافي ه- كمي ي- صفة معينة

ش طلب من باحث تصميم استمارة جمع البيانات لدراسة مفردات الخصبة الوراثية للمرض الجرامي ماهي نصيحتك لهذا الباحث لوضع انفضل نموذج لتصنيفات لهذه الاستمارة ؟

محاضرة الأسبوع الخامس

العينات

من خلال دراستنا للفصل الدراسي الأول كان اهم مواضيع الحصول على معلومات او بيانات الظاهرة التي نود دراستها هي المصادر وهنا ذكرنا انها المصادر نوعين

الأول : مصادر اولية (يقوم الباحث بنفسه الحصول عليها)

وهي : أ- الاستبانة الميداني ، ب- اجراء المقابلات الشخصية ، ج- العينات

الثاني : مصادر ثانوية (يقوم الباحث بالحصول على المعلومات من خلال تجميع من معلومات سابقة)

سنقوم الان بدراسة المصدر الاول للحصول على المعلومات وهي العينة الذي تم تأجيله في حينه .

ان دراسة مجتمع ما غالباً ما تكون مستحيلة كونها تعتبر عملية مكلفة جداً اقتصادية ، الجهد الكبير الكبير ، الوقت الكبير الممكن استغلاله ، لهذه الاسباب يلجأ الجغرافي لاستخدام العينة الممثلة للمجتمع لغرض اجراء دراسة عليها وهذه العينة تستخرج من المجتمع فمثلاً لو اراد جغرافي دراسة نسبة الامراض بمياه الامطار فلا يعقل انه سيجمع كل كميات مياه الامطار التي تمطرها المجتمع الذي يود ذلك الباحث اجراء دراسته عليه . وبالتالي سيختار عينة تمثل مياه الامطار .

تعريف العينة :- هي وحدة او مجموعة وحدات يتم اختيارها من المجتمع لتمثيل ذلك المجتمع بشرط ان تكون الاختيار بغرض متساوية لجميع الوحدات مع افتراض ان تكون العينة متجانسة مع بقية المجتمع اي ان العينة تكون مع بقية المجتمع بنفس الصفات . وهو ما يسمى بالاختيار العشوائي .

ملاحظة عدم التحيز في اختيار العينة لأن التحيز يؤدي الى نتائج غير موضوعية وغير دقيقة

شرط الاختيار العشوائي :-

1. لابد الاختيار للعينة في ضوء النتائج التي يطمح الباحث الوصول اليها [يعني مثلاً يريد الباحث انتاج احد المزارع عليه ان لا يختار المزارع المجاورة بنتائجها العالية او العكس حيث كلهم لهم خصائص البحث]
2. يتم الاختيار على اساس خصائص معينة ، مثلاً عند اجراء بحث عن المقيم على الباحث عدم الاتجاه لاشخاص يتوقع منهم الاستجابة لغرض بحثه
3. كل احتمال العينة التي تم اختيارها عشوائياً يأخذ بعين الاعتبار

ملاحظة: هنا انواع العينات للترتيب بين عدم التركيز عليها او ان ناخذ وقت طويل لانه تعريف كل عينة هو كافي للفهم

تصنيف العينات الى :-

(أ) عينات احتمالية :- وتشمل على مجموعات من العينات كالآتي :

١- العينة العشوائية البسيطة : وهي العينة التي تؤخذ بصورة عشوائية من المجتمع بحيث تتيح لكل فردة من المجتمع الأصلي نفس من الظهور في العينة . وتستخدم طريقة سحب البطاقات (اليانصيب) او من اكبال للاعداد العشوائية لاختيار عينة مثلاً عندنا يكون لدينا مجتمع من حقول زراعية عددها 2000 واردنا عينة عددها (180) حقل زراعي مثل المجتمع السكاني 2000 مزرعة فنحصل على اعداد بطاقات باسماء الحقول الـ (2000) نرقمها من 1 الى 2000 ثم نختار (180) بطاقة بصورة عشوائية . حتى نحاسبه ذاتاً لكل واحد

ب) العينة العشوائية المنتظمة :- وهي العينة التي تؤخذ من المجتمع وفق نسب معينة تعتمد على تقسيم المجتمع الى اقسام حسب نسبة العينة ، هذه العينة تتميز بقلة التكاليف ، سهلة الاستخراج وتكلفة ما سوف اذا الاختيار ليكون عادلاً مثلاً عند دراسة منطقة تجارية فيها مجموعة من المحلات التجارية

مقارنة بين

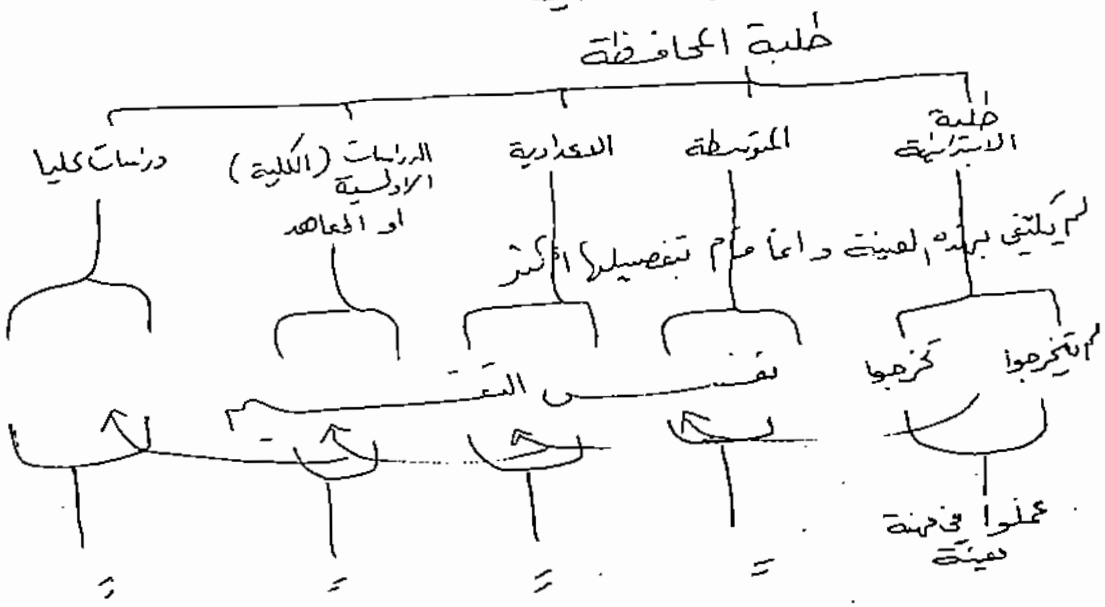
وبين منطقة سكنية فيها محلات تجارية ودراستنا لغرض المقارنة بين المحلات التجارية في المنطقتين فاننا عندما نأخذ عينة من المحلات التجارية ، طبعاً تقسيم المجتمع هنا سيتم وفق * موضوع العينة (محلات تجارية) وهذا يعني انه الاختيار لكل فردة من مفردات المجتمع ستكون لها ميزات خاصة تتم على ضوئها اختيار العينة .

ج- العينة الطبقية :- تؤخذ هذه العينة من مجتمع معروف للباحث معرفة جيدة تتيح له اختيار مفرداتها التي تمثل طبقات المجتمع الاحصائية بحيث تتناسب هذه المفردات مع نسبة الطبقة التي تمثلها من المجتمع ، مثلاً عند دراسة المستوى التعليمي لمحافظة معينة في القطر هنا تحدد الباحث بطبقة الطلبة لمعرفة وجهة نظرهم او يحدد بطبقة التدريسيين لمعرفة وجهة نظرهم ولذا الباحث لم يتجه الى فئة العمال او الموظفين . (يعني تقسيم مجتمع الدراسة الى طبقات واختيار الطبقة المناسبة لاجراء البحث) .

. عدد العاملين في احدى الشركات او المنشآت كالمدارس الازهر او الروابي (المنشآت الماشية) سوف يتم العاملين حسب كالة المعلمين

٤- العينة العنقودية: * تتميز هذه العينة بأن مفرداتها تتكون من مجموعات متقاربة من المفردات وتمثل هذه المجموعات بما يعرف بالعناقيد.

مثلاً لنأخذ عينة من طلبة محافظة معينة لغرض اجراء دراسة معينة عليهم قد يلجأ الباحث الى اسلوب العينة العنقودية مثل الاتي:



وايضاً هنا اراد تفاصيل اكثر

واخذ افراد من كل وحدة العشوائية

لاحظ الشكل اعلاه كأنه عنقود من العنب ← لذلك سميت بالعينة العنقودية

٥- العينة متعددة المراحل * عندما يكون حجم المفردات في العينة العشوائية البسيطة كبيراً لذا يتم اختيار عينة عشوائية اولية. واعلاه مثال وهكذا ذلك اختيار الاحياء الفقيرة داخل احدى المدن الكبرى كعينة عشوائية بسيطة ثم اختيار بعض العوائل من هذه الاحياء واتخاذها كعينة دراسية (هنا الاختلاف بين هذه العينة و العينة العنقودية انك الاخيرة تتفرع هذا الاساس اي حد من الوزع لكنه متعددة المراحل يتفرع الاساس لنوع واحد فقط وهكذا اي انه تخصص).

عينة بسيطة احياء فقيرة في مدينة ← بعض العوائل من هذه الاحياء ← مثلاً بعض العوائل في شارع من هذه الاحياء

١- عينات غير احتمالية :- وتشمل الآتي :-

١- العينة المقصودة (العمدية) : تؤخذ على اساس ان مفرداتها تمثل المتوسط او قريبة منه في المجتمع مثل اختيار عدة قرى من قرى محافظة معينة بشكل تكون قريبة من متوسط حياة سكان المحافظة ككل . تعتبر هذه العينة متحيزة للوسط وتستخدم بشكل قليل ونادراً

٢- العينة الحصصية : هي العينة التي تؤخذ مفرداتها حسب مؤشرات معينة وبذلك هي لا تمثل المجتمع تمثيلاً حقيقياً ولا تعطي نتائج يعتمد عليها

٣- العينة الترشيدية :

دقة تمثيل العينة للمجتمع :-

الغرض من استخدام العينة في الدراسات والبحوث لتقليل الكلفة والجهد في بعض الاحيان تكون مهام الصعب جداً دراسة مجتمع كامل نستعيض عن ذلك باختيار عينة بشرط ان تمثل المجتمع المسحوبة منه سواء بالحجم او بالصفات ، ولكي نقيس دقة تمثيل العينة للمجتمع نستخدم احد مقاييس الاحصائية الا وهو مقياس الخطأ المعياري Standard Error .

ملاحظة :-

- تختلف مقياس الخطأ المعياري عن المسافة المعيارية او الانحراف المعياري

& Standard error

الخطأ المعياري :: يقصد بالخطأ المعياري لتوزيع العينات بالانحراف المعياري لمتوسطات العينات وكلما كانت قيمته صغيرة دل ذلك على ان العينة التي تم اختيارها من المجتمع جيدة وتمثل المجتمع المسحوبة منه افضل تمثيل :

ولايجاد الخطأ المعياري نستخدم الصيغة

$$S.E = \frac{S.D}{\sqrt{n}}$$

S.E = الخطأ المعياري = Standard error

S.D = الانحراف المعياري (تم شرحه سابقاً في المسافة المعيارية) = Standard deviation

n = حجم العينة

مثال: عينة تتألف من (25) قرية زراعية تم اختيارها عشوائياً من (3) محافظات لدراسة ظاهرة الري بالتنقيط.
اذا علمت ان التباين (Var.) يساوي 10 هل ان العينة (25) قرية تمثل المجتمع المأخوذة منه؟

الحل: نرجع الى قانون الخطأ المعياري:

$$S.E = \frac{S.D}{\sqrt{n}}$$

التباين معلوم ونحن نعرف ان $S.D = \sqrt{Var.}$ ← الانحراف المعياري

$$\therefore S.D = \sqrt{10} = 3.16$$

$n = 25$ حجم العينة
قرية زراعية

$$S.E = \frac{3.16}{\sqrt{25}} = \frac{3.16}{5} = 0.63$$

هنا يعتبر الخطأ المعياري قليل نسبياً
∴ نستنتج من ذلك ان العينة المأخوذة من المجتمع تمثل ذلك المجتمع بدقة.

المعاينة: يمكن تعريفها بأسلوب مبسط فالمعاينة تعني الاسلوب او الاجراء المتبع لاختيار العينة وهذه العينة تحتوي على مفردات مأخوذة من مجتمع واحد المفردات تتبع توزيع احصائي وله مقاييس مثل الوسط الحسابي الوسيط، التباين الانحراف المعياري وغيرها... اي ان المعاينة هي التوزيع النظري لتكرارات متوسطات العينات الممكن سحبها
وان توزيع المعاينة للمتوسطات المسحوبة من المجتمع سيقترب من التوزيع الطبيعي Normal distribution كلما ازداد حجم العينة حتى لو كان توزيع المجتمع يبتعد عن التوزيع الطبيعي انه معظم الباحثين من خلال الممارسة او التحليل لاحدى الظواهر المدروسة تكون رغبتهم في ان تكون توزيع المفردات الخاصة بالعينة توزيع طبيعي وذلك للاهمية الكبيرة في حالة تطبيق الكثير من الاختبارات والطرق الاحصائية تشترط صفة الطبيعية في المجتمع قيد الدراسة.

تكملة للمثال ٥:

لمعرفة توزيع المعاينة لـ 10 عينات التي استخرجناها نستخرج الوسط الحسابي والانحراف المعياري

ولكن هنا الوسط الحسابي والانحراف المعياري هو للعينات وليس لعينة واحدة

مكونات طريقة استخراجهما كالآتي ←

العينات المسحوبة	$\bar{X}_i$ الوسط الحسابي لكل عينة	$(\bar{X}_i - \bar{X})$	$(\bar{X}_i - \bar{X})^2$
(3, 41)	3+41/2 = 22	22 - 18 = 4	16
(3, 30)	3+30/2 = 16.5	16.5 - 18 = -1.5	2.25
(3, 11)	3+11/2 = 7	7 - 18 = -11	121
(3, 7)	3+7/2 = 5	5 - 18 = -13	169
(41, 30)	41+30/2 = 35.5	35.5 - 18 = 17.5	306.25
(41, 11)	41+11/2 = 26	26 - 18 = 8	64
(41, 7)	41+7/2 = 24	24 - 18 = 6	36
(30, 11)	30+11/2 = 20.5	20.5 - 18 = 2.5	6.25
(30, 7)	30+7/2 = 18.5	18.5 - 18 = 0.5	0.25
(11, 7)	11+7/2 = 9	9 - 18 = -9	81
عدد العينات = 10	184		720.75

الطريقة التي تم استخراج $\bar{X}$ و S تم بالاعتماد على توزيع المعاينة وهو التوزيع النظري لتكرارات متوسطات العينات اذاً يعتمد على المتوسطات $(\bar{X})$ وليس على القيم نفسها:

$$\bar{X} = \frac{\sum \bar{X}_i}{n} = \frac{\text{مجموع متوسطات العينات}}{\text{عدد العينات}} = \frac{184}{10} = 18.4 \simeq \boxed{18}$$

مربنا العدد لتسهيل الحل

$$S_x = \text{الانحراف المعياري لمتوسطات العينات} = \sqrt{\frac{\sum(\bar{X}_i - \bar{X})^2}{n-1}} = \sqrt{\frac{720.75}{10-1}} = \boxed{8.48}$$

(720.75)

حيث أن:

$\bar{X}$ = الوسط الحسابي العام (للوسط الحسابية الخاصة بالعينات)

$\bar{X}_i$ = الوسط الحسابي للعينة i

n = عدد العينات

∴ المعاينة لها متوسط = 18 وانحراف معياري = 8.48 وهو قيمة كبيرة نوعاً ما

ملاحظة:

سبق ان اوضحت
م عملية الجمع
والطرح وعملية
لضرب في حالة
جود اعداد باشارات
سالبة وكيفية التعامل
ها بالسورة.

الجمع او الطرح

ذا الاشارات مختلفة
طرح وتضع اشارة
عدد الكبير،
ذا الاشارات متشابهة
ع وتضع الاشارة
ا موجبة او سالبة

عند الضرب اذا الاشارات مختلفة نضرب وتضع الاشارة السالبة
= متشابهة نضرب وتضع الاشارة الموجبة

ولد يجاد عدد العينات الممكن سحبها من المجتمع كما يأتي : Combination

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

بعني اختيار n من N

حيث N = حجم المجتمع
n = عدد المفردات في العينة .

وتحديد عدد العينات التي يمكن سحبها من المجتمع يمثل الخطوة الاولى في توزيع المعاينة
اما الخطوة الثانية فتتمثل باستخراج الوسط الحسابي (كمؤشر احصائي لكل عينة).
والانحراف المعياري (الذي يساوي جذر التباين).

مثال : في مجتمع احصائي مكون من (5) مفردات { باعتبار ان هذا المجتمع يمثل [اعمار عائلة مكونة من خمسة افراد] }

$$\left\{ \begin{array}{l} X_1 = 3 \\ X_5 = 7 \end{array} \right. , X_2 = 41 , X_3 = 30 , X_4 = 11 \Bigg\}$$

ماذا طلب منا معرفة توزيع المعاينة ، اذا علمت ان عدد المفردات في كل عينة مسحوبة هو (2) من المفردات

$$\therefore C_n^N = \frac{N!}{n!\,(N-n)!} \Longrightarrow C_2^5 = \frac{5!}{2!\,(5-2)!} = \frac{5.4.3.2.1!}{2.1\ (3)!}$$

$$= \frac{5.4.3}{3.2.1} = \boxed{10}$$

∴ نستطيع سحب 10 عينات من المجتمع المكون من 5 مفردات بشرط ان تكون كل عينة تحوي على مفردتين وكالآتي (للتوضيح)

و (41,7) , (41,11) , (41,30) , [illegible] و (3,41) , (3,30) , (3,11) , (3,7)

و (11,7) و (30,7) , (30,11) ═══ عددها 10 عينات .

يعني اخذنا $X_1 = 3$ مع البقية ، ثم اخذنا $X_2 = 41$ مع البقية بدون X_1 ، ثم $X_3 = 30$ مع البقية بدون X_1 و X_2 ثم X_4 مع X_5 فقط (يعني بدون تكرار) .

عملياً تكون لمؤشرات الاحصائية الكاملة للمجتمع (يعني الوسط الحسابي والتباين) غير معلومة لذا يتم تقديرها من بيانات العينة [اذ اننا نستخرج الوسط الحسابي $\bar{X}$ والانحراف المعياري S للعينة]

∴ $\bar{X}$ هو تقدير للوسط الحسابي للمجتمع M .

و Var او S^2 هو تقدير لتباين المجتمع V .

وهناك نوعين من التقديرات :-

١- تقدير بنقطة point estimates الذي يعني تقدير قيمة عددية واحدة لمعلمة المجتمع .

٢- تقدير بفترة Interval estimates الذي يعني تقدير فترة بين عددين تحددان الفترة التي نشعر انها تقع درجة الحرية تحتوي على المعلمة التي نقوم بتقديرها باحتمال Probability ويطلق على هذا الاحتمال (مستوى المعنوية او الثقة) وهاتان القيمتان تسميان بحدود الثقة والفترة التي تقع بينهما هي فترة الثقة بمعنى تقدير المعلمة يقع ضمن فترة الثقة .

ايجاد حدود الثقة :-

حدا الثقة هما حدا الوسط الحسابي للمجتمع M يعرفان الوسط الحسابي للمجتمع الاصلي نعرف الوسط الحسابي للعينة وقيمة Z القياسية وحجم وتباين العينة

علماً اننا سبق واشرنا انه يُعطى في السؤال مستوى المعنوية (α) المقابل لقيمة Z

حيث ان :

الحد الادنى $<$ M $<$ الحد الاعلى

الوسط الحسابي للمجتمع يقع بين الحدين

$$\text{الحد الادنى} = \bar{X} - Z_{\alpha} \cdot \frac{S}{\sqrt{n}}$$

$\bar{X}$ الوسط الحسابي — S الانحراف المعياري ($S = \sqrt{Var}$, $S^2 = Var$) — n حجم العينة

$$\text{الحد الاعلى} = \bar{X} + Z_{\alpha} \frac{S}{\sqrt{n}}$$

مثال 4: عينة عشوائية حجم (25) وسطها الحسابي = (2.7) وتباينها (9) اوجد حدي الثقة لمتوسط المجتمع (M) بمستوى معنوية 99% اذا علمت ان مستوى المعنوية $\alpha = 0.99$ ؟

$$\text{الحد الادنى} = \bar{X} - Z_{0.99}\frac{S}{\sqrt{n}} = 2.7 - (2.326)\frac{\sqrt{9}}{\sqrt{25}}$$

$$= \boxed{1.3}$$

$$\text{الحد الاعلى} = \bar{X} + Z_{0.99}\frac{S}{\sqrt{n}} = 2.7 - (2.326)\frac{\sqrt{9}}{\sqrt{25}} = \boxed{4.1}$$

$$\therefore 1.3 < M < 4.1$$

النهاية: ~~يمكن توزيع العينة جميع مفردات~~ العينة وهي الإجراء المتبع لاختيار العينة وهذه العينة تحتوي على مفردات أو الإجراء المتبع لاختيار العينة وهذه العينة تحتوي على مفردات مأخوذة من مجتمع وهذه المفردات تتبع توزيع احتمالي وله نهايته مثل الوسط الحسابي - الوسيط - التباين - الانحراف المعياري ... الخ

- اي أن المعاينة هي التوزيع النظري لتكرارات متوسطات العينات الممكن سحبها وأن توزيع المعاينة للمتوسطات المسحوبة من المجتمع سيقترب من التوزيع الطبيعي Normal distribution كلما ازداد حجم العينة حتى لو كان توزيع المجتمع ~~سيقترب~~ من يبتعد عن التوزيع الطبيعي.

- أن معظم الباحثين من خلال المقارنة او التحليل لاحد الظواهر المدروسة تكون رغبتهم في ان يكون توزيع المفردات الخاصة بالعينة توزيع طبيعي وذلك للاهمية الكبيرة في حالة توفره لكثير من الاختبارات والطرق الاحصائية تشترط صفة الطبيعية في المجتمع قيد الدراسة.

** أيجاد عدد العينات :-

لايجاد عدد العينات الممكن سحبها من المجتمع تستخدم صيغة القانون التالي :-

Combination :- التوافيق

$$C_n^N = \frac{N!}{n!(N-n)!}$$

حيث أن :-
N : حجم المجتمع
n : عدد مفردات في العينة

والناتج يمثل عدد العينات المحتملة من N من المفردات بحيث كل عينة تحتوي كل n مفردة وبدون تكرار

- وتعتبر عملية تحديد عدد العينات التي يمكن سحبها من المجتمع الخطوة الاولى في توزيع المعاينة اما الخطوة الثانية فتتمثل باستخراج الوسط الحسابي (كمؤشر احصائي لعينة) والانحراف المعياري (الذي يسمى خطأ التباين).

مثال: في مجتمع إحصائي يمثل (5) مفردات هي:

$(X_1 = 3 \ , X_2 = 41 \ , X_3 = 30 \ , X_4 = 11 \ , X_5 = 7)$

بافتراض أن هذا المجتمع يمثل [أعمار عائلة مكونة من خمس أفراد].

ماذا طلب منا معرفته توزيع المعاينة إذا علمت أن عدد المفردات في كل عينة مسحوبة هو (2) ؟ (حجم العينة)

الحل

① الخطوة الأولى هو إيجاد عدد العينات:

$$C_n^N = \frac{N!}{n!(N-n)!}$$

عدد العينات الممكنة

$$\Rightarrow C_2^5 = \frac{5!}{2!(5-2)!} = \frac{5*4*3*2*1}{2*1*(3)!} = \frac{5*4*3}{3*2*1} = \boxed{10}$$

أي مثلاً ( )

∴ تستطيع سحب (10) عينات من المجتمع، لكون ق 5 مفردات بشرط أن تكون كل عينة تحتوي على مفردتين وتختلف (بالترتيب) (ودون تكرار) :

(3, 41), (3, 30), (3, 11), (3, 7), (41, 30), (41, 11), (41, 7), (30, 11), (30, 7), (11, 7).

وعددها 10 عينات

② الخطوة الثانية هو استخراج الوسط الحسابي والانحراف المعياري:

العينات الممكنة	الوسط الحسابي لكل عينة $\bar{X}_i$	$(\bar{X}_i - \bar{\bar{X}})$	$(\bar{X}_i - \bar{\bar{X}})^2$
(3,41)	3+41/2 = 22	22 − 18 = 4	16
(3,30)	3+30/2 = 16.5	16.5 − 18 = −1.5	2.25
(3,11)	3+11/2 = 7	7 − 18 = −11	121
(3,7)	3+7/2 = 5	5 − 18 = −13	169
(41,30)	41+30/2 = 35.5	35.5 − 18 = 17.5	306.25
(41,11)	41+11/2 = 26	26 − 18 = 8	64
(41,7)	41+7/2 = 24	24 − 18 = 6	36
(30,11)	30+11/2 = 20.5	20.5 − 18 = 2.5	6.25
(30,7)	30+7/2 = 18.5	18.5 − 18 = 0.5	0.25
(11,7)	11+7/2 = 9	9 − 18 = −9	81
	$\Sigma \bar{X}_i = 184$		$\Sigma(\bar{X}_i - \bar{\bar{X}})^2 = 720.75$

$$\bar{\bar{X}} = \frac{184}{10} = 18.4$$

الطريقة التي تم استخراج $\bar{X}$ و S_x تم الاعتماد على تعريف العينة وهو التوزيع الذي لمتغيرات متوسطات العينات اذاً يعتمد على المتوسطات ($\bar{X}$) وليس على القيم نفسها:

$$\Rightarrow ① \quad \bar{X} = \frac{\sum \bar{X}_i}{n} = \frac{\text{مجموع متوسطات العينات}}{\text{عدد العينات}} = \frac{184}{10} = 18.4 \simeq 18$$

$$② \quad S_{\bar{X}} = \frac{\text{الانحراف المعياري}}{\text{لمتوسطات العينات}} = \sqrt{\frac{\sum(\bar{X}_i - \bar{X})^2}{n}} = \sqrt{\frac{730.75}{10}} = 8.48$$

حيث أن:

$\bar{X}$ = الوسط الحسابي العام (للأوساط الحسابية للعينات)

$\bar{X}_i$ = الوسط الحسابي للعينة i.

n = عدد المفردات.

∴ العينة لها متوسط = 18 وانحراف معياري = 8.48 وهي قيمة كبيرة نوعاً ما وذلك لأن حجم العينة (يتناسب عكسياً).

- عملياً تكون المؤشرات الإحصائية الخاصة بالمجتمع (يعني الوسط الحسابي والتباين) غير معلومة لذا يتم تقديرها من بيانات العينة [أي أننا نستخرج الوسط الحسابي $\bar{X}$ والانحراف المعياري S للعينة].

∴ $\bar{X}$ هو تقدير للوسط الحسابي للمجتمع μ.

∴ Var أو S^2 هو تقدير لتباين المجتمع σ^2.

- هناك نوعين من التقديرات:

① تقدير بنقطة point estimation والذي يعني تقدير قيمة عددية واحدة لمعلمة المجتمع.

② تقدير بفترة Interval estimation والذي يعني تقدير قيمتين عدديتين تحددان الفترة التي تنتشر احتمالية درجة الحرية فتحتوي على المعلمة التي نقوم بتقديرها باحتمال probability وتطلق على هذا الاحتمال (مستوى الثقة أو الثقة) وهنا بالإحصاء نسبيان عددي يحدد الثقة والفترة التي تقع بينهما هي فترة (الثقة) تعني تقدير المعلمة ضمن فترة الثقة تعني تقدير المعلمة يقع ضمن فترة الثقة.

حدود الثقة:

- حدا الثقة هما حدا الوسط الحسابي للمجتمع (M) يعرفان الوسط الحسابي للمجتمع الأعلى من معرفة الوسط الحسابي للعينة وقيمة (Z) القياسية وحجم وتباين العينة.
- وحيث أننا وجدنا في موضوع العينة كيفية حساب الوسط الحسابي والانحراف المعياري فأننا سوف نستفاد منهما في حساب حدود الثقة بالإضافة الى أننا سوف نحتاج الى ما يسمى بمستوى المعنوية والذي نرمز له بالرمز (α) والذي يقابل قيمة Z وغالبا ما يعطى في السؤال.
- عليه فأن حدود الثقة للوسط الحسابي للمجتمع M تكون:

$$\text{الحد الأعلى} > M > \text{الحد الأدنى}$$

(الوسط الحسابي للمجتمع يقع بين الحدين)

حيث أن:

$$\text{الحد الأدنى} = \bar{X} - Z_{\alpha} * \frac{S'}{\sqrt{n}}$$

$$\text{الحد الأعلى} = \bar{X} + Z_{\alpha} * \frac{S'}{\sqrt{n}}$$

مثال: عينة عشوائية حجم (25) وسطها الحسابي = 3.7 وتباينها (9)، أوجد حدود الثقة للوسط الحسابي للمجتمع (M) بمستوى معنوية ($\alpha = 0.99$)، علماً أن قيمة Z عند مستوى المعنوية 0.99 تساوي (2.326)

الحل:

$S' = \sqrt{Var} \Rightarrow S' = \sqrt{9}$

$$\text{الحد الأدنى} = \bar{X} - Z_{\alpha} * \frac{S'}{\sqrt{n}} = 3.7 - (2.326) * \frac{\sqrt{9}}{\sqrt{25}}$$

$$= 1.3$$

$$\text{الحد الأعلى} = \bar{X} + Z_{\alpha} * \frac{S'}{\sqrt{n}} = 3.7 - (2.326) * \frac{\sqrt{9}}{\sqrt{25}}$$

$$= 4.1$$

∴ حدود الثقة للوسط للمجتمع (M) تكون:

$$1.3 < M < 4.1$$

يقصد بالخطأ المعياري لتوزيع المعاينة بالانحراف المعياري لمتوسطات العينات وكلما كانت قيمته صغيرة دل ذلك على أن العينة التي تم اختيارها من المجتمع جيدة وتمثل المجتمع المسحوبة منه أفضل تمثيل.

- يتم إيجاد الخطأ المعياري [Standard Error] باستخدام الصيغة التالية :-

$$S.E = \frac{S.D}{\sqrt{n}}$$

حيث أن :-

Standard Error = S.E = الخطأ المعياري

Standard deviation = S.D = الانحراف المعياري

numbers = n = حجم العينة

- إذا كان لدينا مجتمع مكون من (25) قرية وتم اختيار عينة منها من (5) قرى لدراسة ظاهرة ما وتم استخراج أن التباين (Var.) يساوي (10). هل أن العينة (5) مؤشر تمثل المجتمع المأخوذة منه ؟

الحل/

نرجع لايجاد الخطأ المعياري

$$S.E = \frac{S.D}{\sqrt{n}}$$

حيث الانحراف المعياري هو جذر التباين $S.D = \sqrt{Var.}$

$S.D = \sqrt{10} = 3.16$

$n = 25$

$$S.E = \frac{S.D}{\sqrt{n}} = \frac{3.16}{\sqrt{25}} = \frac{3.16}{5} = 0.63$$

هنا يجب الحذر فالعينة (5) ليست مؤشر تمثل المجتمع

الفصل الثالث

عند جمع البيانات الأولية الخاصة بدراسة ظاهرة ما فإنه عادة لا يمكن الاستفادة منها وهي بهذه الصورة، لذلك فغالبا ما توضع في جداول مبسطة أو يعبر عنها في صورة أشكال ورسوم بيانية لكي يسهل دراستها وتحليلها.

3-1 جدول التوزيع التكراري Frequency distribution Table :-

وهو عبارة عن جدول بسيط لتلخيص وترتيب البيانات للمتغير العشوائي التي سبق ان جمعت وصنفت الى عدد من المجاميع كل منها تسمى الفئة class وهذه الفئات قد تكون مرتبة ترتيبا "تصاعديا" او تنازليا".
ويسمى توزيع القيم (X) حسب الفئات (بالتوزيع التكراري). وقد تكون فئات التوزيع متساوية الطول فيسمى توزيعا "منتظما" او قد تكون غير متساوية الطول وذلك يعتمد على طبيعة الدراسة ومتطلباتها.

3-1-1 تعاريف مهمة :-

1- المدى الكلي Total Range (T.R) :- وهو الفرق بين اكبر واصغر قيمة في المجموعة مضافا" له العدد واحد ونرمز له بالرمز T.R :-

$$T.R = X_L - X_S + 1$$

حيث ان :-
X_L : هي اكبر قيمة
X_S : : اصغر :

2- عدد الفئات Number of Classes :- ويمثل عدد المجاميع التي تحوي عليها التوزيع وكل مجموعة تسمى

بالفئة ، وهناك صيغ يمكن من خلالها تحديد عدد الفئات التوزيع الامثل :

صيغة سترجس Sturges :-

وهي :-

$$m = 1 + 3.322 \log_{10} n$$

حيث أن :-

m : يمثل عدد الفئات

$\log_{10}$: اللوغاريتم العشري للأساس عشرة

n : حجم العينة (عدد المفردات)

وعند التطبيق يتم التقريب الناتج لأقرب عدد صحيح .

٢- طول الفئة Length of Classes :-

وهو مقدار المدى بين حدي الفئة [الاعلى والادنى] ، وان طول الفئة يتناسب عكسياً مع عدد الفئات فكلما كبر طول الفئة قل عدد الفئات والعكس صحيح ويستحسن ان تكون اطوال الفئات متساوية لتسهيل استخراج المؤشرات الاحصائية (كالوسط الحسابي والانحراف المعياري ، ... الخ) ويتوزع طول الفئة كالاتي :-

$$\text{طول الفئة } L = \frac{T - R}{m}$$

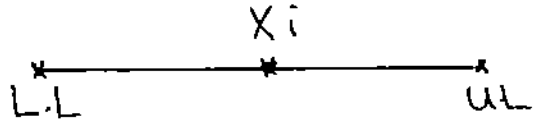
٤) الحد الأدنى والحد الأعلى للفئة :-

الحد الادنى هو بداية الفئة والحد الاعلى هو نهاية الفئة .

٥- مركز الفئة Class mid-point :-

وهو عبارة عن منتصف المدى بين حدي الفئة ولكل فئة مركز ويرمز له

بالرمز Xi .



حيث ان : L.L الحد الادنى للفئة
U.L الحد الاعلى للفئة

~~مركز الفئة هو عبارة عن~~

$$Xi = \frac{L.L + U.L}{2}$$

٦- تكرار الفئة Class frequency :

يمثل تكرار الفئة جزء من مفردات العينة التي تتصف بكونها تقع من حيث القيمة العددية ما بين حدي الفئة بحيث ان مجموع التكرارات يساوي عدد مفردات العينة كلها وهو (n) . ونرمز لتكرار الفئة بالرمز f_i وان تكرار الفئة الاولى يرمز لها بالرمز f_1 والثانية بـ f_2 وهكذا بحيث ان :-

$$f_1 + f_2 + f_3 + \cdots + f_m = n$$

الخطوات الحاقة في انشاء جداول التوزيع التكراري :-

١) - استخراج مدى المتغير Range .
٢) - اختيار وتحديد عدد الفئات .
٣) - ايجاد طول مدى الفئة .
٤) - كتابة حدود الفئات
٥) - استخراج عدد التكرارات لكل فئة .

ملاحظة (١) ليس من الضروري عند التقسيم ان يكون الحد الادنى للفئة الاولى مساوٍ تماماً لاصغر قيمة في المجموعة بل قد يكون اقل منها لاعتبارات تتعلق بتسهيل العمليات الحسابية اللاحقة .

ملاحظة (2):
قد يكون التوزيع التكراري توزيع مغلق او توزيع مفتوح وذلك بالاعتماد على طبيعة الدراسة. ونقصد بالتوزيع المغلق بانه ذلك التوزيع الذي يمتلك حدا" ادنى للفئة الاولى وحدا" اعلى للفئة الاخيرة.
اما التوزيع المفتوح هو ذلك التوزيع الذي لايمتلك حدا" ادنى للفئة الاولى او حدا" اعلى للفئة الاخيرة او كليهما معا".

3-1-2 مميزات التوزيع التكراري:

1- ان تكون فئات التوزيع متساوية الطول.
2- ان يكون التوزيع التكراري توزيع مغلق.
3- ان تبدأ حدود الفئات التوزيع وتنتهي باعداد صحيحة.
4- ان لايقل عدد فئات التوزيع عن خمسة ولايزيد عن خمسة عشر فئة.
فاذا قل عدد فئات عن خمسة فان عملية التبويب تؤدي الى عدم كشف الصفات الاساسية للمجتمع واعطاء صورة واضحة عنه، اما اذا زاد عددها عن 15 فئة فان الباحث سيلاقي بعض الصعوبة في اجراء العمليات الحسابية.
5- ان تكون حدود الفئات محددة بشكل واضح بحيث ان كل قيمة من قيم المفردات يجب ان تضم في فئة واحدة فقط من فئات التوزيع.

مثال: القيم التالية تمثل اطوال 20 نباتا من نباتات القطن والمطلوب انشاء جدول توزيع تكراري لاطوال هذه النباتات علما" ان طول الفئة قد استخرج مسبقا" وسيأتي (5):

80, 87, 98, 81, 74, 84, 79, 80
78, 82, 93, 91, 70, 90, 80, 84
71, 70, 92, 88.

$\frac{98-25}{19-40} = 10$ 5

$$T.R = X_L - X_S + 1$$

$$= 98 - 70 + 1 = 29$$

$$L = \frac{T.R}{m}$$

$$= \frac{29}{5} = 5.8 \simeq 6$$

التكرار رقماً	التكرار علامة	حدود الفئات بشكل نهائي	الحد الاعلى للفئة	الحد الادنى للفئة
4	1111	70 - 75	70+6 -1	70 -
6	1 ⧸⧸⧸⧸	76 - 81	70+12-1	70+6
4	1111	82 - 87	70+18-1	70+12
5	⧸⧸⧸⧸	88 - 93	70+24-1	70+18
1	1	94 - 99	70+30-1	70+24
n=20				

مثال (2) :- كون جدول توزيع تكراري للبيانات التالية وافرض ان عدد الفئات يساوي 5 .

10 , 23 , 30 , 35 , 41 , 55 , 12 , 25 , 45 , 59 , 24 , 37
42 , 39 , 38 .

الحل :-

$$T.R = X_L - X_S + 1$$

$$= 59 - 10 + 1 = 50$$

$$L = \frac{T-R}{m}$$

$$= \frac{50}{5} = 10 \qquad \text{طول الفئة}$$

ان جدول التوزيع التكراري سيكون بالشكل الآتي :-

الفئات class	تكرار الفئة fi	مركز الفئة Xi
10 - 19	2	14.5
20 - 29	3	24.5
30 - 39	5	34.5
40 - 49	3	44.5
50 - 59	2	54.5
	15	

ملاحظة تختلف الفئات حسب نوع المتغير العشوائي كما موضح ادناه :-

الفئات	الفئات	الفئات	الفئات
10 - 19	10 -	10 - 20	less than 20
20 - 29	20 -	20 - 30	20 -
30 - 39	30 -	30 - 40	30 -
40 - 49	40 -	40 - 50	40 -
60 - 59	50 -	50 - 60	50 and more
متغير عشوائي متقطع	متغير عشوائي مستمر	متغير عشوائي مستمر	متغير عشوائي مستمر

"مقاييس التمركز او التوسط"
Measures of Central Tendency

ان معظم القيم لمختلف الظواهر الطبيعية تتمركز عادة في الوسط او قريبة منه ومقاييس التمركز او التوسط لأي مجموعة من البيانات التابعة لظاهرة ما هي تلك المقاييس التي تبحث في تقدير قيمة تتمركز حولها اغلبية هذه البيانات وان هذه القيمة المتوسطة هي رقم وحيد يعبر ويمثل جميع بيانات تلك المجموعة الهدف منها اعطاء صورة سريعة عن ماهية تلك المجموعة.

واهم مقاييس التوسط هي:-

1- الوسط الحسابي Arithmetic Mean
2- المنوال Mode
3- الوسيط Median

3-2-1 الوسط الحسابي The Arithmetic Mean

الوسط الحسابي او المتوسط لقيم متغير ما هو القيمة الناتجة من قسمة مجموع تلك القيم على عددها ويرمز له بالرمز $\bar{X}$.

طرق حساب الوسط الحسابي:-

1- في حالة البيانات الغير مبوبة:-
اذا كان لدينا n من القيم او المشاهدات فان الوسط الحسابي لها هو مجموع هذه القيم مقسوما" على عددها اي:-

$$X_1, X_2, \dots, X_n$$

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

مثال ٢- البيانات التالية تمثل كمية المطر الساقطة سنوياً (بالملليمترات) على مدينة بغداد خلال فترة من سنوات 450, 350, 520
400, 380.

$$\bar{X} = \frac{520 + 350 + 450 + 380 + 400}{5} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{2100}{5} = 420 \text{ mm.}$$

اي ان معدل سقوط الامطار خلال تلك الفترة هو 420 ملم.

مثال ٢- البيانات التالية تمثل اوزان عينة من الطلبة حوامها 15 طالب وطالبة
اوجد متوسط وزن الطالب لهذه العينة.

50.2, 60.9, 68.3, 59.2, 58.1, 62.3, 52.9, 61.5, 63.2, 59.1
69.3, 64.2, 65.2, 56.6, 65.3

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{50.2 + 60.9 + \text{----} + 56.6 + 65.9}{15}$$

$$= \frac{916.3}{15} = 61.087$$

ملاحظة: في حالة البيانات المتقطع يتم تقريب الناتج الى اقرب عدد صحيح.

٢- في حالة البيانات المبوبة ٢-

اذا كانت $X_1, X_2, \ldots, X_k$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها $f_1, f_2, \ldots, f_k$ على التوالي فالوسط الحسابي هو ٢-

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

خطوات إيجاد الوسط الحسابي في حالة البيانات المبوبة :-

١- تعيين مراكز الفئات X_i .

٢- ضرب مركز كل فئة بمقدار تكرارها ($f_i X_i$) .

٣- قسمة مجموع (حاصل ضرب مركز كل فئة × تكرارها) ÷ مجموع التكرارات .

مثال :- الآتي توزيع تكراري لدرجات الحرارة في مدينة معينة والمسجلة لمدة 95 يوماً متتالياً، يطلب حساب متوسط درجة الحرارة في هذه المدينة خلال هذه الفترة .

درجات الحرارة :	0-	1-	2-	3-	4-	5-	6-	7-8
عدد الأيام :	4	8	12	16	20	25	6	4

الحل :-

الفئات	التكرارات f_i	مركز الفئات X_i	$f_i X_i$
0-	4	0.5	2
1-	8	1.5	12
2-	12	2.5	30
3-	16	3.5	56
4-	20	4.5	90
5-	25	5.5	137.5
6-	6	6.5	39
7-	4	7.5	30
المجموع	95	—	396.5

متوسط الحرارة في هذه المدينة

$$\bar{X} = \frac{\sum_{i=1}^{8} f_i X_i}{\sum_{i=1}^{8} f_i} = \frac{396.5}{95} = 4.174$$

مثال (2) :- في ادناه توزيع تكراري لعينة من الأسر عددها 75 أسرة حسب عدد افراد الأسرة (بضمنها الوالدين)، يطلب حساب متوسط عدد افراد الأسرة في هذه العينة.

عدد الافراد :	2-4	5-7	8-10	11-13	14-16	17-19	20-22
عدد الأسر :	8	12	20	13	10	8	4

الحل :-

الفئات Class	التكرارات f_i	مراكز الفئات X_i	$f_i X_i$
2-4	8	3	24
5-7	12	6	72
8-10	20	9	180
11-13	13	12	156
14-16	10	15	150
17-19	8	18	144
20-22	4	21	84
المجموع	75	—	810

$$\bar{X} = \frac{\sum_{i=1} f_i X_i}{\sum_{i=1} f_i} = \frac{810}{75}$$

$$= \frac{810}{75} = 10.8 \simeq 11 \text{ فرد}$$

وحيث ان عدد افراد الأسرة متغير من النوع المتقطع فانه لايوجد مقياس يجزء من الفرد عليه يتم تقريب الناتج لاقرب عدد صحيح.

3-2-2 المنوال The Mode :-

يعرف المنوال بانه تلك القيمة التي تتكرر اكثر من غيرها من بين مجموعة من القيم او ازاء القيمة الشائعة من بين مجموعة من القيم ويرمز له بالرمز ($\overline{Mo}$) .

وقد يكون هناك اكثر من منوال او قد لايوجد .

مثال (3) :-

اوجد المنوال للبيانات التالية :-

5 , 6 , 2 , 3 , 6 , 8 , 2 , 5 , 9 , 5

الحل :- العدد 5 هي اكثر المفردات تكرارا" فهي المنوال

$\therefore \overline{Mo} = 5$

مثال (4) :-

جد المنوال للبيانات التالية :-

2 , 4 , 3 , 6 , 8 , 7 , 10 , 12

الحل :-

واضح من هذه المجموعة انه لايوجد عدد متكرر اكثر من غيره عليه فانه لايوجد منوال لهذه المجموعة .

ايجاد المنوال في حالة البيانات المبوبة :-

هنا يحتسب حساب المنوال على نوع المتغير وكما يلي :-

م- حساب المنوال اذا كان المتغير متقطع :-

ان المنوال هنا يمثل قيمة مركز الفئة التي تقابل اكبر تكرار في الجدول.

ملاحظة : في حالة وجود فئتين او اكثر تقابل نفس التكرار فان التوزيع في هذا النوع سوف يمتلك اكثر من قيمة واحدة للمنوال كل منها تمثل مركز الفئة التي تقابل ذلك التكرار.

مثال (5): اوجد المنوال للجدول التوزيع التكراري التالي :-

Class	f_i
50 - 74	2
75 - 89	6
90 - 104	14
105 - 119	8
120 - 134	10

الحل:- نلاحظ ان الفئة (90-104) لها اكبر تكرار

$$X_i = \frac{90 + 104}{2} = 97$$

$\therefore \bar{M}o = 97$ ⇐

ب- اذا كان المتغير مستمر :-

يحسب حساب المنوال حسب الصيغة التالية :-

$$\bar{M}o = L_1 + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2}\right) \times C$$

حيث أن :-

L_1 : الحد الأدنى الحقيقي للفئة المنوالية.

$\Delta 1$: تكرار الفئة المنوالية - التكرار السابق لها.

$\Delta 2$: تكرار الفئة المنوالية - التكرار اللاحق لها.

C : طول الفئة المنوالية.

مثال (6) :- الجدول التالي يمثل التوزيع التكراري لأطوال عينة من الأشخاص عددها 50 شخصاً، يطلب حساب القيمة الشائعة لطول الشخص في هذه العينة .

فئات الطول (سم) :-	150-	160-	170-	180-	190-200
عدد الأشخاص :-	8	12	15	9	6

الحل :- من الواضح ان اكبر تكرار في التوزيع هو 15 عليه فأن المنوال فئته هي 170 - 180 اي الفئة الثالثة .

$$L_1 = 170 - 0.5 = 169.5$$

$$\Delta_1 = 15 - 12 = 3$$

$$\Delta_2 = 15 - 9 = 6$$

$$C = 10$$

$$\therefore \bar{M}o = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) \times C$$

$$= 169.5 + \left(\frac{3}{3+6}\right)(10)$$

$$= 172.83 \text{ c.m}$$

3-2-3 الوسيط The Median

وهو تلك القيمة من قيم المتغير العشوائي التي تقسم مجموعة البيانات (القيم) إلى قسمين متساويين ، اي انها القيمة التي تجعل عدد القيم قبلها مساوٍ لعدد القيم بعدها. ويرمز له بالرمز $\bar{M}e$.

1- ايجاد الوسيط لبيانات غير مبوبة :

اذا كان لدينا n من القيم $x_n, \dots, x_2, x_1$ ورتبت هذه القيم ترتيباً "تصاعدياً" او تنازلياً

فانه :-

١- اذا كانت n عدد فردي فأن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$.

٢- اذا كانت n عدد زوجي فأن الوسيط هو الوسط الحسابي للقيمتين اللتين تسلسلها $\frac{n}{2}$ ، $\frac{n}{2}+1$ اي

$$\overline{M}e = \frac{X_{n/2} + X_{n/2+1}}{2}$$

مثال :- اوجد الوسيط لدرجات طالب في 5 امتحانات يدرس لاحصاء اذا كانت الدرجات هي :- 84 ، 87 ، 76 ، 82 ، 80 .

الحل :-

نرتب القيم تصاعديا" ⟶ 76 ، 80 ، 82 ، 84 ، 87

بما ان عدد الارقام فردي $n=5$

اذن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2} = \frac{5+1}{2} = 3$

$X_3 \Rightarrow 82$

$\therefore \overline{M}e = 83$

مثال (7) :- الآتي اعمار عينة من الافراد عددها عشرة افراد ، جد الوسيط لعمر الفرد في هذه المجموعة .

20 ، 22 ، 19.5 ، 26 ، 24.5 ، 27 ، 28 ، 18 ، 20 ، 23

الحل :- نرتب هذه القيم ترتيب تنازلي .

28 ، 27 ، 26 ، 24.5 ، 23 ، 22 ، 20 ، 20 ، 19.5 ، 18 ⟶

X_1 ، X_2 ، X_3 ، X_4 ، X_5 ، X_6 ، X_7 ، X_8 ، X_9 ، X_{10}

بما ان n عدد زوجي اذن :-

القيمة التي ترتيبها السادس $\frac{n}{2}+1 = 6 \Rightarrow X_6 = 22$

القيمة التي ترتيبها الخامس $\frac{n}{2} = \frac{10}{2} = 5 \Rightarrow X_5 = 23$

وبذلك فان الوسيط لهذه المجموعة يمثل الوسط الحسابي لهاتين القيمتين اي :

$$\frac{22+23}{2} = 22-5 \text{ سنة}$$

ايجاد الوسيط لبيانات غير مبوبة :-

اذا كانت $X_1, X_2, \ldots X_K$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها $f_1, f_2, \ldots f_K$ فقيمة الوسيط لهذه البيانات بعد ترتيبها ترتيباً تصاعدياً او تنازلياً هو

$$Me = L_1 + \left[\frac{\frac{\Sigma f_i}{2} - F_i}{f_i}\right] w$$

حيث أن :-

L_1 = الحد الأدنى الحقيقي لفئة الوسيط

Σf_i = مجموع التكرارات

F_i = التكرار المتجمع عند بداية فئة الوسيط

f_i = تكرار فئة الوسيط = [التكرار المتجمع عند نهاية فئة الوسيط - التكرار المتجمع عند بداية فئة الوسيط]

w = طول فئة الوسيط

خطوات ايجاد الوسيط :-

١- عمل جدول توزيع تكراري تجميعي تصاعدي او تنازلي .

٢- ايجاد ترتيب الوسيط وهو $\frac{\Sigma f_i}{2}$.

٣- ايجاد فئة الوسيط

أ- ايجاد حدودها الحقيقية .

ب- كتابة التكرار التجميعي التصاعدي او التنازلي امام كل منها .

بمعنى تحديد فئة الوسيط وهي الفئة التي تقع قيمة الوسيط بين حديها وذلك عن طريق ايجاد قيمتين متتاليتين في التكرار التجميعي تقع بينهما ترتيب الوسيط.
تقابل هاتين القيمتين حدا الفئة الوسيطية الادنى والاعلى الحقيقي.

٤- تطبيق القانون.

ملاحظة: الحد الادنى الحقيقي لأي فئة = الحد الادنى لتلك الفئة - $\frac{1}{2}$

الحد الاعلى الحقيقي لأي فئة = الحد الاعلى لتلك الفئة + $\frac{1}{2}$

مثال (8): اوجد الوسيط للتوزيع التكراري التالي:-

فئات الطول	التكرارات f_i	الحدود العليا للفئات	التكرار المتجمع الصاعد F_i
60-62	5	less than 62	5
63-65	18	less than 65	23 ← فئة الوسيط
66-68	42	less than 68	65 ←
69-71	27	less than 71	92
72-74	8	less than 74	100
	100	—	—

ترتيب الوسيط $= \frac{\sum_{i=1} f_i}{2} = \frac{100}{2} = 50$

اي ان قيمة الوسيط هو طول الشخص الذي ترتيبه 50 بعد ترتيب القيم تصاعديا"

او(تنازلياً)، وفي جدول التوزيع التكراري التجميعي الصاعدي نرى بان 50 تقع واقعة بين الرقمين 23 و 65

	التكرار التجميعي	الحدود الحقيقية
فئة الوسيط	23	65.5
	65	68.5

$\therefore$ $L_1 = 65.5$ الحد الأدنى الحقيقي لفئة الوسيط

$F_i = 23$ التكرار المتجمع عند بداية فئة الوسيط

$f_i = 65 - 23 = 42$ تكرار فئة الوسيط

$w = 68.5 - 65.5 = 3$ طول فئة الوسيط

$$\therefore \overline{M}e = L_1 + \left[\frac{\left(\frac{\Sigma f_i}{2}\right) - F_i}{f_i}\right] w$$

$$= 65.5 + \left[\frac{50 - 23}{42}\right](3)$$

$$= 67.43 \text{ c.m}$$

ملاحظة

التكرار المتجمع الصاعد less Than Cumulative frequency ويرمز له بالرمز F_i وتختصر less Than C.F.

التكرار المتجمع النازل more Than Cumulative frequency يرمز له بالرمز F_i وتختصر more Than C.F.

مقاييس التشتت او الاختلاف
Measures of Dispersion or Variation

التشتت : Dispersion

التشتت او الاختلاف هو التباعد او التقارب الموجود بين المشاهدات لمتابعة لمتغيرها . ومقاييس التشتت تستخدم لقياس تشتت القيم عن وسطها الحسابي وكلما كان مقياس التشتت كبيراً دل ذلك على عدم التجانس بين المفردات وكلما كان مقياس التشتت صغيراً دل ذلك على التجانس بين المفردات (اي ان الاختلاف بين القيم قليلة) .

ان لمقاييس التشتت اهمية كبيرة في وصف التوزيعات ومقارنتها مع بعضها حيث ان مقاييس الوسط وحدها لا تكفي لهذا الغرض فقد يتساوى الوسط الحسابي لمجموعتين من القيم بينما يختلف مدى انتشار القيم في المجموعة الاولى عن المجموعة الثانية .

مثال :- بيانات المجموعة الاولى هي :- 15 ، 12 ، 9 ، 6 ، 3

بيانات المجموعة الثانية هي :- 17 ، 1 ، 13 ، 9 ، 5

ان الوسط الحسابي لكلا المجموعتين هو (9) ولكن المجموعة الاولى اكثر تجانساً من المجموعة الثانية .

وهناك عدة مقاييس للتشتت اهمها :-

1- المدى The Range

2- الانحراف المتوسط The Mean Deviation

3- التباين والانحراف القياسي The variance and standard Deviation

1- المدى The Range :-

وهو ابسط انواع مقاييس التشتت المطلقة

وهو الفرق بين اكبر قيمة واقل قيمة في المجموعة في حالة البيانات غير المبوبة :-

$$R = X_L - X_S$$

اما في الجداول المبوبة فأن المدى هو الفرق بين الحد الأدنى للفئة الأولى والحد الاعلى للفئة الاخيرة .

مثال :- اوجد المدى للبيانات الآتية :-

X_i = 12 , 6 , 7 , 3 , 15 , 10 , 18 , 5

الحل :-

$$R = X_L - X_S$$
$$= 18 - 3 = 15$$

مثال :- اوجد المدى لجدول التوزيع التكراري التالي :-

classes	التكرار
3 - 5	7
6 - 8	8
9 - 11	2
12 - 15	4

الحل :-

المدى = الحد الاعلى للفئة الاخيرة - الحد الادنى للفئة الاولى

$$R = 15 - 3 = 12$$

2- الانحراف المتوسط The Mean Deviation :-

وهو احد مقاييس التشتت المطلق ويمكن ايجاد الانحراف المتوسط لحالة البيانات المبوبة وغير المبوبة ويرمز له بالرمز M.D .

أ- حالة البيانات غير المبوبة :-

يقيس تشتت المفردات حول وسطها الحسابي .

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

مثال :- اوجد الانحراف المتوسط للبيانات التالية :-

9 , 8 , 6 , 5 , 7 :

الحل :-

X_i	$X_i - \bar{X}$	$\lvert X_i - \bar{X} \rvert$
9	2	2
8	1	1
6	-1	1
5	-2	2
7	0	0
35		6

$$\bar{X} = \frac{\sum X_i}{n} = \frac{35}{5} = 7$$

$$M.D = \frac{\sum_{i=1}^{5} |X_i - \bar{X}|}{n}$$

$$= \frac{6}{5} = 1.2$$

ملاحظة ان السبب في أخذ الانحرافات المطلقة هو ان بقاء الإشارات الموجبة والسالبة يجعل مجموع الانحرافات ساوية الى الصفر حيث ان $\sum(X_i - \bar{X}) = 0$ ولذلك نأخذ القيم المطلقة لإفقاد الإشارات السالبة وجعلها موجبة .

ب- حالة البيانات المبوبة :-

$$M.D = \frac{\sum_{i=1}^{n} f_i |X_i - \bar{X}|}{\sum_{i=1}^{n} f_i}$$

حيث أن :-

X_i = مركز الفئة

f_i = التكرارات ، مجموع التكرارات $\sum f_i = n$

$\bar{X}$ = الوسط الحسابي

مثال :- اوجد الانحراف المتوسط لجدول التوزيع التكراري التالي :-

الفئات	التكرار f_i	X_i	$X_i f_i$	$\lvert X_i - \bar{X} \rvert$	$f_i \lvert X_i - \bar{X} \rvert$
60-62	5	61	305	+6.45	32.25
63-65	18	64	1152	+3.45	62.10
66-68	42	67	2814	+0.45	18.90
69-71	27	70	1890	+2.55	68.85
72-74	8	73	584	+5.55	44.40
	100		6745		226.50

$$\bar{X} = \frac{\sum X_i f_i}{\sum f_i} = \frac{6745}{100} = 67.45$$

$$M.D = \frac{\sum f_i |X_i - \bar{X}|}{\sum f_i} = \frac{226.50}{100} = 2.265$$

3-3 الانحراف المعياري والتباين Standard Deviation and Variance:

يعتبر الانحراف المعياري من اهم مقاييس التشتت واكثرها استخداماً وانتشاراً ويسمى هذا المقياس في بعض الاحيان بالانحراف القياسي.

ويعرف بانه الجذر التربيعي لمتوسط مجموع انحراف اي عينة عن وسطها الحسابي ولكني نأخذ في نظر الاعتبار حجم العينة حتى يمكن مقارنة العينات المختلفة الاحجام فاننا نقسم مجموع مربعات الانحرافات على $n-1$ ويرمز للانحراف المعياري بالرمز S اما التباين فهو مربع الانحراف المعياري اي S^2.

عند حساب التباين (S^2) ستكون قيمته مقاسه بمربع الوحدات المستخدمة في قياس المشاهدات فاذا كانت المشاهدات مقاسة بالمتر فان التباين تكون مقاساً بالمتر المربع وللتوحيد مثلاً ولكن اذا كانت المشاهدات عبارة عن (اوزان او مبالغ بالدينار او عمال او عدد افراد الاسرة ـ ـ ـ) فان التباين ستكون مقاساً بالدينار المربع او العامل المربع ـ ـ ـ وهذه كلها ليس لها معنى ولحل لترجع وحدات القياس الى اصلها يأخذ الجذر التربيعي للتباين (اي S).

ب)- حساب الانحراف المعياري لبيانات غير مبوبة:

لتكن $X_1, X_2, \ldots, X_n$ تمثل قياسات مفردات عينة قوامها n و ليكن $\bar{X}$ يمثل الوسط الحسابي لهذه القياسات فان الانحراف المعياري S لهذه المجموعة هو:

الانحراف المعياري للعينة

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

التباين للعينة

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

مثال: البيانات التالية تمثل كمية محصول القطن بالطن في عشر مزارع اوجد التباين والانحراف المعياري لها؟

7 ، 5 ، 6 ، 8 ، 9 .

X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
9	$9-7=2$	4
8	1	1
6	-1	1
5	-2	4
7	0	0
35	0	10

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{35}{5} = 7$$

التباين $S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} = \frac{10}{4} = 2.5$

الانحراف المعياري $S = \sqrt{S^2} = \sqrt{2.5} = 1.58$

ب)- حساب الانحراف المعياري لبيانات مبوبة :-

لتكن $X_1, X_2, \dots, X_K$ تمثل مراكز فئات توزيع تكراري عدد فئاته K وان $f_1, f_2, \dots, f_K$ تمثل التكرارات المقابلة لهذه الفئات فان الانحراف المعياري يكون :-

الانحراف المعياري للعينة

$$S = \sqrt{\frac{\sum_{i=1}^{K} f_i (X_i - \bar{X})^2}{\sum_{i=1}^{n} f_i - 1}}$$

$$\bar{X} = \frac{\sum_{i=1}^{K} f_i X_i}{\sum_{i=1}^{K} f_i} \quad , \sum f_i = n$$

التباين للعينة

$$S^2 = \frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i - 1}$$

مثال : اوجد الانحراف المعياري والتباين لجدول التوزيع التكراري التالي:

الفئات	f_i	x_i	$f_i x_i$	$(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^2$
60 - 62	5	61	305	41.60	208.01
63 - 65	18	64	1152	11.90	214.24
66 - 68	42	67	2814	0.202	8.505
69 - 71	27	70	1890	6.50	175.56
72 - 74	8	73	584	30.80	246.42
	100		6745		852.75

$$\bar{X} = \frac{\sum_{i=1}^{5} f_i x_i}{\sum_{i=1}^{5} f_i}$$

$$= \frac{6745}{100} = 67.45$$

① $$S^2 = \frac{\sum_{i=1}^{5} f_i (x_i - \bar{x})^2}{\sum_{i=1}^{5} f_i - 1} = \frac{852.75}{100-1} = 8.41$$

② $$S = \sqrt{S^2} = \sqrt{8.41} = 2.9$$

الفصل الرابع

"الاحتمالية"

الاحتمالية The probability : تعني بالاحتمالية او الاحتمال هو دراسة التجارب العشوائية وهو كلمة تستخدم للتعبير عن احتمال وقوع شيء لم يحدث بعد. ولقد نشأت من الناحية التاريخية نظرية الاحتمالات بدراسة بعض العاب الحظ والورق. P

بعض التعاريف المهمة :-

التجربة العشوائية :- هي التجربة التي لا يمكن معرفة نتيجتها لخضوعها لقوانين الاحتمال.

فضاء العينة Sample Space :- تعني الحالات الكلية للتجربة العشوائية وهي مجموعة النقاط التي تمثل جميع النتائج او الحالات الممكنة التي لها نفس فرصة الظهور لتجربة ما.

حيث ان كل نتيجة تمثل بنقطة او عنصر في فضاء العينة، ويرمز لفضاء العينة بالرمز S.

فعند رمي قطعة نقود مرة واحدة، فان فضاء العينة (الحالات الكلية) لهذه التجربة هي :-

$$S = \{ H , T \}$$

اي يتكون من نتيجتين ممكنتين.

اما اذا رميتا قطعتين من النقود فأن فضاء العينة يتكون من 4 نتائج هي :-

$$S = \{ TT , HT , TH , HH \}$$

الحادثة (او الحدث) The event :-

هو نقطة او عدة نقاط في فضاء العينة اي هو مجموعة جزئية من فضاء العينة S ويرمز له بالرمز (E) .
فالحصول على صورة (H) من رمي قطعة النقود مرة واحدة تسمى حادثاً وهو يتكون من نقطة واحدة هي { H } من مجموعة نقاط (الحالات الكلية) :-

$$S = \{ T , H \}$$

والحادث قد يكون بسيطاً اذا كان فضاء العينة من نقطة واحدة [نتيجة تجربة هي حالة واحدة] ، او يكون الحادث مركباً اذا شمل فضاء العينة على حالتين فأكثر

ان احتمال ظهور الحادثة E يرمز له بالرمز $P(E)$ ويساوي :-

$$P(E) = \frac{\text{عدد حالات ظهور الحادث}}{\text{عدد الحالات الكلية}} = \frac{n(E)}{N}$$

مجموع كل الاحتمالات للتجربة العشوائية الواحدة = 1 عدد صحيح

قواعد الاحتمالات :

1) - قيمة الاحتمال دائماً موجبة وهي تتراوح مابين الصفر والواحد اي $0 \leq P(E) \leq 1$ وتمثل قيمة الاحتمال تساوي صفر يعني فشل المحاولة وتمثل قيمة الاحتمال تساوي 1 يعني نجاح المحاولة وتأكدة الوقوع (المستحيلة) .

مثال (1) : صندوق يحتوي على 20 كرة بيضاء فاذا سحبت منه كرة واحدة عشوائية فما هو احتمال :-

أ - ان تكون بيضاء .
ب - ان تكون حمراء .

الحل :

أ - نفرض ان ظهور كرة بيضاء هو ω

فأن احتمال ان تكون الكرة بيضاء هو :-

$$P(W) = \frac{n(W)}{N} = \frac{20}{20} = 1$$

ب- نفرض ان ظهور كرة حمراء هو R ،

احتمال ان تكون الكرة حمراء هو :- $P(R) = \frac{n(R)}{N} = \frac{0}{20} = 0$

٢)- ان احتمال عدم ظهور الحادث (اي فشله) ويرمز له بالرمز $P(\bar{E})$ وحسابه

$$P(\bar{E}) = 1 - P(E)$$

لأن :-

٢)- احتمال النجاح + احتمال الفشل = 1 اي أن :-

$$P(E) + P(\bar{E}) = 1$$

٤) $\sum P_i = 1$ مجموع الاحتمالات للتجربة العشوائية = 1 .

قواعد عامة في نظرية الاحتمالات General rules in probability Theroy

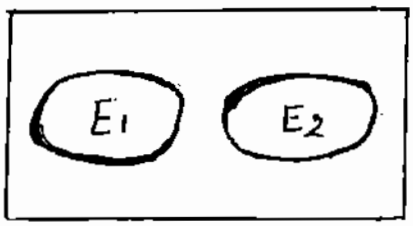
بعد ان عرفنا مفهوم الاحتمال وكيفية قياسه ، سنعرض الآن بعض القواعد العامة في نظرية الاحتمالات من حيث عمليات الجمع والضرب :-

اولاً : قاعدة الجمع The addition rule :-

أ- عندما تكون الحوادث متنافية :

اذا كانت E_1 و E_2 حادثان متنافيان فأنه من غير الممكن حدوثهما معاً بعبارة أخرى اي انهما غير متقاطعين ، $E_1 \cap E_2 = \phi$

اي ليس هناك علاقة بين الحادثين



فإن احتمال حدوث أياً منهما اي E_1 و E_2 هو حاصل جمع كل منهما :-

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

مثال (2) :

ألقي حجر النرد (الزار) مرة واحدة ما هو احتمال ظهور عدد زوجي وظهور عدد فردي في نفس الوقت ؟

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

نفرض ان A هي حادثة ظهور عدد زوجي .

$$A = \{2, 4, 6\}$$

و و B و و و فردي .

$$B = \{1, 3, 5\}$$

$$A \cap B = \phi$$

$$P(A) = \frac{n(A)}{N} = \frac{3}{6} = \frac{1}{2}$$

$$P(B) = \frac{n(B)}{N} = \frac{3}{6} = \frac{1}{2}$$

$$\therefore P(A \underset{or}{\cup} B) = P(A) + P(B)$$

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

ومعناه ان ظهور احد هذه الارقام هو مؤكد

مثال (3) :- صندوق يحتوي على 4 كرات سوداء و5 بيضاء و3 حمراء فاذا سحبت كرة واحدة منها هو احتمال ان تكون أما سوداء او بيضاء ؟

4 B
5 W
3 R

الحل :-

نفرض ان الحادثة B هي ظهور كرة سوداء

و و W و و و بيضاء

$$P(B \cup W) = P(B + W) = P(B) + P(W)$$

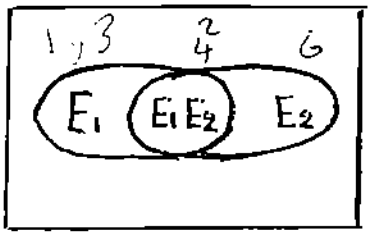
$$P(B) = \frac{n(B)}{N} = \frac{4}{12} \quad \text{و} \quad P(W) = \frac{n(W)}{N} = \frac{5}{12}$$

فإن احتمال ان تكون الكرة البيضاء او السوداء $P(B+W) = \frac{4}{12} + \frac{5}{12}$

$= \frac{9}{12} = \frac{3}{4}$

ب - عندما تكون الحوادث غير متنافية :- لتكن E_1 و E_2 حادثتين ممكنتين الحدوث معاً في تجربة معينة عندئذ فإن احتمال وقوع E_1 أو E_2 هو حاصل جمع احتمال وقوع E_1 مع احتمال وقوع E_2 مطروحاً منه احتمال وقوعهما معاً :

أي ان :-



$$P(E_1+E_2) = P(E_1) + P(E_2) - P(E_1E_2)$$

$E_{12}\{2, 4, 6\}$

$E_2 = \{1, 2, 3, 4\}$

مثال : اذا كان اللاعب A يصيب هدفاً ما باحتمال $\frac{1}{4}$ وان اللاعب B يصيب نفس الهدف باحتمال $\frac{2}{5}$ فما هو احتمال اصابة الهدف اذا صوب كل من A و B نحو الهدف ؟

الحل :- المقصود هنا ما هو احتمال A أو B او كلاهما يصيب الهدف

$P(A) = \frac{1}{4}$ ، $P(B) = \frac{2}{5}$

$\therefore P(A+B) = P(A) + P(B) - P(AB)$

$= \frac{1}{4} + \frac{2}{5} - (\frac{1}{4})(\frac{2}{5})$

$= \frac{1}{4} + \frac{2}{5} - \frac{2}{20}$

$= \frac{11}{20}$

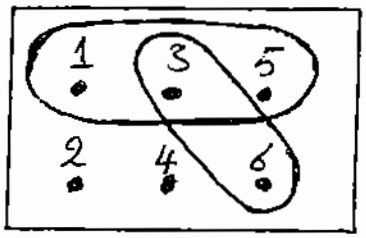
مثال (2) :- اذا القيَ زهر النرد مرة واحدة فما هو احتمال ظهور عدد يكون فرديا" او يقبل القسمة على 3 ؟

الحل :- ان فضاء العينة لهذه التجربة هو :-

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

نفرض ان A تمثل ظهور عدد فردي

نفرض ان B يمثل حادثة ظهور عدد يقبل القسمة على 3 .



$$P(A) = \frac{n(A)}{N} = \frac{3}{6} = \frac{1}{2}$$

$$P(B) = \frac{n(B)}{N} = \frac{2}{6} = \frac{1}{3}$$

$$\therefore P(A+B) = P(A) + P(B) - P(AB)$$

$$= \frac{1}{2} + \frac{1}{3} - \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) = \frac{4}{6}$$

ثانياً :- قاعدة الضرب The multiplication rule :-

P- عندما تكون الحوادث مستقلة :- وهي الحوادث التي اذا وقع احدها لا يمنع او يؤثر على وقوع الاحداث الاخرى .

مثلاً" عند رمي قطعتي نقود فالحصول على صورة في القطعة الأولى لا يؤثر في نتيجة القطعة الثانية .

اذا كان لدينا E_1 و E_2 حادثتين مستقلتين عندئذ فان احتمال حدوث E_1 و E_2 معاً" هو :-

$$P(E_1 E_2) = P(E_1) \cdot P(E_2)$$

مثال: في تجربة رمي قطعتي نقود، افرض ان E_1 تمثل ظهور الوجه الذي يحمل صورة للقطعة الاولى وان E_2 تمثل ظهور الوجه الذي يحمل كتابة للقطعة الثانية فما هو احتمال ظهور صورة في القطعة الاولى وظهور كتابة في القطعة الثانية ؟

$S_1 = \{T, H\}$

$S_2 = \{T, H\}$

الحل:-

نفرض ان احتمال ظهور الوجه هو E_1 وهو يساوي $P(E_1) = \frac{1}{2}$

وان E_2 هو احتمال ظهور الكتابة وهو يساوي $P(E_2) = \frac{1}{2}$

$$P(E_1 \text{ and } E_2) = P(E_1 E_2) = P(E_1) \cdot P(E_2)$$

$$= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

واجب: صندوقان الاول يحوي على 4 كرات بيضاء و3 سوداء والثاني يحوي على 3 بيضاء و5 سوداء فاذا سحبت كرة من كل منهما فما هو احتمال ان تكونا سوداوين ؟

4 W 2 B	3 W 5 B
الصندوق الاول	الصندوق الثاني

ب- عندما تكون الحوادث غير مستقلة:-

وهي الحوادث التي اذا وقع احدها يؤثر في وقوع الاحداث الأخرى.

أي ان :-

$$P(E_1 E_2) \neq P(E_1) \cdot P(E_2)$$

مثال:-

اذا القينا قطعة نقود 3 مرات فأننا سنحصل على فضاء العينة التالي:

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

نفرض ان A هو حادثة ظهور صورة في الرمية الاولى.

" B " " " " " " الثانية.

" C " " " " " مرتان متتاليتان.

الحل :-

من الواضح ان A و B مستقلان ولكن العلاقة مابين A و C أو B و C ليست واضحة.

$$P(A) = P(\{HHH, HHT, HTH, HTT\}) = \frac{4}{8} = \frac{1}{2}$$

$$P(B) = P(\{HHH, HHT, THH, THT\}) = \frac{4}{8} = \frac{1}{2}$$

$$P(C) = P(\{HHT, THH\}) = \frac{2}{8} = \frac{1}{4}$$

$$P(AB) = P(A \cap B) = P(\{HHH, HHT\}) = \frac{2}{8} = \frac{1}{4}$$

$$P(AC) = P(A \cap C) = P(\{HHT\}) = \frac{1}{8}$$

$$P(BC) = P(B \cap C) = P(\{HHT, THH\}) = \frac{2}{8} = \frac{1}{4}.$$

الآن نبرهن هل ان الاحداث مستقلة ام غير مستقلة.

$$P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = P(A \cap B)$$

اذاً A و B مستقلان

$$P(A) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} = P(A \cap C)$$

اذاً A و C مستقلان

$$P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \neq P(B \cap C)$$

اذاً B و C غير مستقلين.

$$ABC = \{HHT\} \Rightarrow P(ABC) = \frac{n(ABC)}{N} = \frac{1}{8}$$

$$P(A) \cdot P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{16} \neq \frac{1}{8}$$

واجب (3-H-2)

س1/ نفرض ان الحدث A هو «في العائلة اطفال من النوعين ضبيط وبنت 9) والحدث B هو «في العائلة ولدين واحد على الاكثر» ؛ اثبت أن :-

أ- الحدثين A و B مستقلان اذا كان في العائلة 3 اطفال .

ii- الحدثين A و B غير مستقلين اذا كان في العائلة طفلان .

س2/ ألقيت قطعتي نقود وكان A يمثل حادثة ظهور صورة في القطعة الاولى. و B يمثل حادثة ظهور الصورة في القطعة الثانية ، اما هو يمثل حادثة ظهور الصورة مرة واحدة فقط ، اطلوب ايجاد الحوادث المستقلة والحوادث غير المستقلة .

تطبيقات جغرافية حول نظرية الاحتمالات :-

لتطبيق نظرية الاحتمالات في الدراسات الجغرافية لابد من تحديد حجم العينة وحجم المجتمع الذي سحبت منه العينة ولابد ايضاً من معرفة مجال العينة فمثلاً لدراسة حالة الجو يتم الاعتماد على المعرفة السابقة لحالة الجو .

فمثلاً لمعرفة احتمال سقوط المطر في ما خلال شهر آذار فعلى الباحث معرفة حالات المطر السجلة في منطقة دراسته ولفترة طويلة .

والعينة هنا هو عدد سنطقة الدراسة وعدم اخذ جميع المناطق التي يسقط فيها المطر وهذا مستقبل ومجال العينة هنا هو شهر آذار لسنوات مختلفة لمعرفة كمية سقوط المطر فيها لعدة سنوات المقارنة .

مثال :- حللت احدى محطات الرصد الجوي لمدينة بغداد الأيام التي تسقط فيها الامطار لشهر آذار ولمدة عشرة سنوات وكان مجموعها (296) يوماً من مجموع ايام شهر آذار لعشر سنوات والذي يساوي (310) يوماً والمطلوب ايجاد :

أ- احتمال المطر لأي يوم من ايام شهر اذار ؟

ii- احتمال ان يكون احد ايام شهر آذار جاف (غير ممطر) ؟

iii) - احتمال ان يكون 3 أيام ممطرة ؟

iv) - احتمال ان يكون احد الايام ممطرة ؟

v) - احتمال ان يكون اليوم الاول جاف او اليوم الثاني جاف او يوم الثالث جاف ؟

الحل :-

شهر أذار = 31 يوماً

2- عدد ايام أشهر اذار = (310) يوماً عبر مدى عشر سنوات

i) - احتمال ان تمطر لأي يوم من ايام شهر آذار

لتكن A حادثة سقوط المطر لأي يوم

$$P(A) = \frac{n(A)}{N} = \frac{296}{310}$$

$$= 0.63$$

ii) - احتمال ان يكون احد ايام شهر آذار جاف (غير ممطر)

كما نعلم أن احتمال النجاح + احتمال الفشل = 1

احتمال النجاح هو سقوط المطر وهو $P(A)$

احتمال الفشل هو عدم سقوط المطر وليكن $P(\bar{A})$

$$\therefore P(A) + P(\bar{A}) = 1$$

$$0.63 + P(\bar{A}) = 1 \quad \Rightarrow P(\bar{A}) = 1 - 0.63 = 0.37$$

iii) احتمال ان تكون 3 أيام ممطرة = احتمال اليوم الاول X احتمال اليوم الثاني X احتمال اليوم الثالث

$$P(A) = (0.63)(0.63)(0.63)$$

$$= 0.25$$

iv) احتمال ان يكون احد الايام الثلاثة جافاً = احتمال يوم ممطر X احتمال يوم جاف X احتمال يوم ممطر

$$P(A) = (0.63) \times (0.37) \times (0.63) = 0.15$$

v) احتمال ان يكون اليوم الاول والثاني والثالث جافاً = احتمال اليوم الاول جاف X احتمال اليوم الثاني جاف X احتمال اليوم الثالث جاف

$$P(\bar{A}) = (0.37) \times (0.37) \times (0.37) = 0.45$$

$${}_nP_r = \frac{n!}{(n-r)!}$$

التباديل Permutations

هو عدد طرق ترتيب r من الأشياء المختلفة حيث n من الأشياء مختلفة (هناك إشارة إلى الترتيب) وقانونه هو:

كم كلمة مختلفة يمكن أن تتكون من الحروف الآتية a, b, C, U, R ؟

الحل: عدد الكلمات الممكنة هو عدد طرق ترتيب الأحرف بجانب بعضها في صف واحد

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$

إن إذا كان الترتيب غير مهم مهما يعني أن ab والاختيار a b هما اختيار واحد

التوافيق Combinations

هو عدد طرق اختيار r من الأشياء من بين n من الأشياء المختلفة عندما لا يكون الترتيب مهماً حيث n توافيق r ويرمز له بالرمز ${}_nC_r$ أو $\binom{n}{r}$ وقانونه هو:

$$C^n_r = \frac{n!}{r!(n-r)!}$$

Ex: يوجد 12 طالب و 8 طالبات، بكم طريقة يمكن تكوين لجنة مكونة من ثلاثة طلاب وطالبتين؟

Sol: نلاحظ هنا أن الترتيب غير مهم حيث نريد لجنة من غير أن نحدد وظائف معينة لأعضائها وبما أنه يريد اختيار 3 من الطلاب

$$C^{12}_3 = \frac{\quad}{\quad} = 220$$

$$C^8_2 = 28 \qquad C \cdot C = 220 \times 28 = 6160$$

Ex: | 3 R 7 G | 5 B | كم طريقة مختلفة نستطيع وضع هذه الكرات بأصف؟

Sol: لو كانت جميع الكرات مختلفة عن بعضها البعض فإن عدد طرق ترتيبها بأصف هو 15! ولكن يوجد تشابه بين بعض الكرات 3! 5! 7! إذن عدد الترتيب المختلفة هو $\frac{5!}{3!\,5!\,7!}$

Ex: إذا كان على الطالب أن يجيب في أحد الامتحانات عن 5 أسئلة من بين 8 أسئلة فأوجد التالي:

1- بكم طريقة يمكن أن يجيب الطالب عن خمسة أسئلة

2- " " " " " " " " على أن يكون السؤال الأول من المسائل الـ 3 الأولى.

3- بكم طريقة يمكنه أن يجيب عن خمسة أسئلة من ضمن السؤال 8.

الحادي عشر : براءة الاختراع المسجلة خارج أو داخل العراق

عنوان البراءة	رقم البراءة	تاريخ منح البراءة	جهة منح البراءة

**ملاحظة :
تمنح سبع نقاط عن كل براءة إختراع خارج العراق .
تمنح خمس نقاط لكل براءة إختراع مسجلة داخل العراق .

"الفصل الخامس"

تحليل الانماط المكانية والانماط الشبكية

تحليل الانماط المكانية :-

منذ القدم تم الاعتماد على النقاط لتأشير المواقع على الخرائط حيث تعتبر النقاط وسيلة من وسائل عرض التوزيع الجغرافي وان استثمار استخدام هذه الطريقة هي لسهولة اسقاطها وسهولة فهم التوزيعات الجغرافية التي تمثلها .

١- تحليل الانماط النقطية :- هي عبارة عن اشتقاق الظاهرة المدروسة من العالم الحقيقي وتحويلها الى هيكل يودع الظاهرة مكانياً بواسطة النقاط نستطيع ان نؤشر على خارطة معينة مواقع لمصانع أو قرى أو مستشفيات او مراكز تجارية او مدارس اوغيرها وهذه الخرائط تمثل مرحلة اولى في التحليل الجغرافي تأتي بعدها مرحلة تحديد فيما اذا كان توزيع الظاهرة المدروسة يشكل نمط معين ام لا .

فكل نقطة تؤشر على الخارطة يتم بها تأشير موقع ما اي انها عبارة عن عمليات تفاعلت في الزمان والمكان لتكون هذا الموقع . وهذه الحقيقة جعلت الجغرافيون لايهتمون في انماط التوزيع الجغرافي فقط بل بالعمليات التي تؤدي الى تطور هذا النمط لأن النمط المكاني يمثل الزمان الذي تكون فيه (هو عند ثابت) فالزمن لايتوقف بالعمليات التي تشكل الانماط لكنها مستمرة في عملها وبدون توقف وفي دراستنا للانماط وتحليلها لايشترط ان تكون وحدة القياس لرسم شبكة مربعة فقط تكون مثلثه او دائره اومستطيل او أي شكل هندسي اخر ولكن للسهولة يؤخذ المربع كمقياس .

١- معدل المركز الكاني :-

وهو مركز التوزيع الكاني وهو مشابه للمعدل في مجموع الارقام غير الكانية والجرافيون معتادون على استخدام خطوط الطول والعرض ومعتادون ايضا على استخدام اكرائط ذات التوزيع الشبكي لهذا فان عملية ايجاد مركز المعدل امرا ليس بالصعب او الغريب على الجرافيون . وكذلك معدل المركز الكاني (مركز المعدل) وحسب نوع البيانات تتبع الخطوات التالية :-

1-5 معدل التمركز الكاني (مركز المعدل) في حالة البيانات غير المبوبة :

١- رسم شبكه من المربعات على الخارطة بشرط ان تكون المقاييس فيها متساوية افقيا وعموديا.

٢- يتم تحديد مواقع (النقاط) التي تمثل ظاهرة معينة (متغيرة) برسم محورين الافقي (X) والعامودي (Y) وتكون مواقع النقاط تابعة لكل من (X) و (Y) .

٣- نقوم بتكوين جدول من ثلاثة اعمدة ، العامود الاول هو مواقع الظاهرة ، والعامود الثاني يمثل المسافة لهذه المواقع او بُعدها عن (X) ، اما العامود الثالث فيمثل المسافة مابين نقاط المواقع (Y) .

٤- نستخرج الوسط الحسابي لبعد المواقع عن (X) وليكن ($\bar{X}$) ، والوسط الحسابي لبعد المواقع عن (Y) وليكن ($\bar{Y}$) .

٥- يتم تعيين نقطة التقاطع بيانيا على شبكه المربعات مابين قيمة ($\bar{X}$) وقيمة ($\bar{Y}$) حيث تمثل هذه النقطة معدل المركز الكاني (مركز المعدل) .

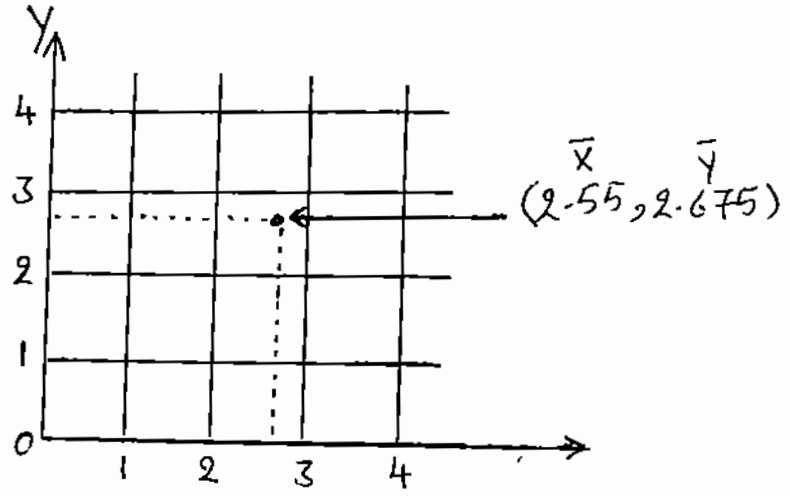
مثال (١) ، اوجد معدل المركز الكاني (مركز المعدل) للبيانات المتوفرة الآتية والتي تمثل توزيعا جغرافيا لبعض المدن في منطقة ما :-

الحل :-

المدن	A	B	C	D	E	F	G	H	المجموع
From X	4.0	3.2	3.0	3.2	2.2	2.1	1.7	1.0	20.4
From Y	1.2	1.8	2.7	3.7	2.3	3.6	3.2	2.9	21.4

① - $\bar{X} = \frac{\sum_{i=1}^{8} X_i}{n} = \frac{20.4}{8} = 2.55$

② - $\bar{Y} = \frac{\sum_{i=1}^{8} Y_i}{n} = \frac{21.4}{8} = 2.675$



مركز المعدل الوزني :-

يستخدم هذا المقياس في حالة حساب مركز المعدل (معدل المركز المكاني) لظاهرة معينة تختلف قيمتها في الحجم او في العدد فمثلاً عند دراسة مواقع المحلات التجارية فان هذه المحلات تختلف في حجمها وتختلف في مقدار المردود المالي لها عندها يجب استخدام مركز المعدل الوزني او وزن قيم الظاهرة المراد دراستها.

مثال :- لدراسة اهمية المواقع محلات بيع الكتب للمدن التالية (A, B, C, D, E) اوجد مركز المعدل الوزني، اذا علمت ان اهمية وزن كل مكتبة من المكتبات وموقعها كما في ادناه

	Li	Wi	Si	Wi	LiXwi	SiXwi
A	5	8	2	8	40	16
B	4	5	1	5	20	5
C	2	10	3	10	20	30
D	1	42	4	42	42	168
E	3	20	5	20	60	100
	مكتبات عامة	أوزان	محلات بيع الكتب	85	182	319

الحل :-

نستخرج الوسط الحسابي للمكتبات العامة ($\bar{L}$) :-

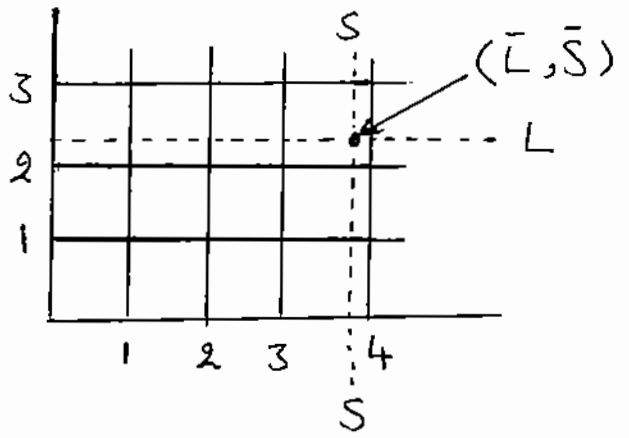
$$\textcircled{1}\ \bar{L} = \frac{\sum Liwi}{\sum Wi} = \frac{182}{85} = 2.14$$

نستخرج الوسط الحسابي لمحلات بيع الكتب ($\bar{S}$)

$$\textcircled{2}\ \bar{S} = \frac{\sum Siwi}{\sum Wi} = \frac{319}{85}$$

$$= 3.75$$

النقطة الاخيرة هي اسقاط قيم النواتج الاخيرة $(\bar{S}, \bar{L})$ وهي (3.75 و 2.14)



③- الوسيط المكاني :-

يتناول الوسيط المكاني مفهومين :

م- الوسيط الذي يحقق اقل مسافة بين النقاط والمواقع ويعتمد على هذا المفهوم في دراسات النقل والدراسات الاقتصادية ولايجاد الوسيط يتم تكرار عملية اختيار نقطة معينة وحساب المسافة بينها وبين النقاط الباقية إلا ان نصل الى نقطة تكون المسافة بينها وبين البقية اقل من المسافات التي تم الحصول عليها بالتكرار وتعتبر هذه النقطة الاخيرة هي الوسيط .

ب- المفهوم التقليدي للوسيط اي هو القيمة التي تفصل البيانات اي بين مجموعتين متساويتين اي ان القيم فوق الوسيط تساوي القيم التي تقع اسفل الوسيط (في العدد) ، يمكن تحديد الوسيط المكاني من خلال حساب عددي ونقسم شبكة المربعات الى جزئين متساويتين في عدد القيم اي اننا لانتباع اي عمليات حسابية واحصائية.

مثال :

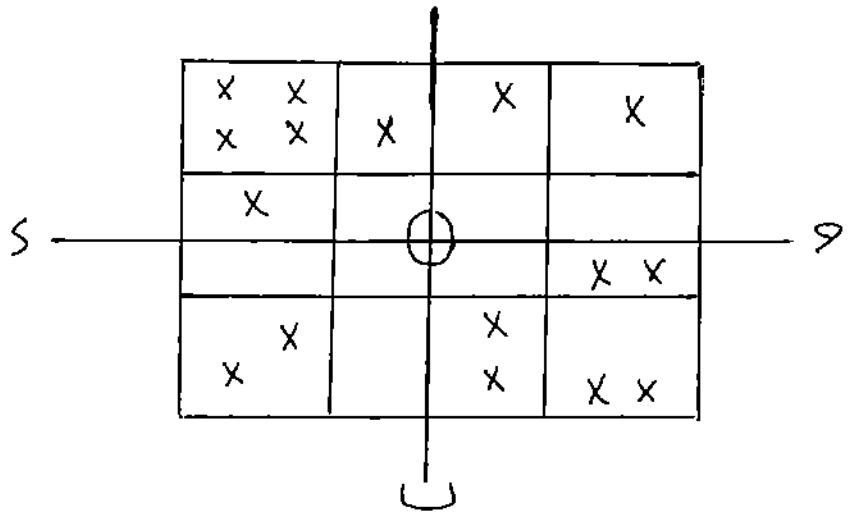
شبكة المربعات ادناه تم توزيع المناطق التي تشهد إنتاج محصول (التي عليها بشكل نقاط) حدد مكان الوسيط على هذه الشبكة ؟

الحل :-

نرسم خطين متعامدين مثل (م - ب) و (ح - ك) .

* لرسم الخط (م - ب) بصورة عمودية بحيث ان عدد المواقع على يمينه يساوي عدد المواقع على يساره .

* لرسم الخط (ص - س) بصورة افقية بحيث يقسم الواقع الى قسمين متساويين بحيث اعلى الخط يساوي اسفل الخط .
* نقطة التقاطع تسمى بالوسيط الحسابي م



- المسافة المعيارية :-

وهي من مقاييس التشتت وهي تمثل الانحراف المعياري في الاحصاء الوصفي للمتغيرات حيث يوفر هذا المقياس وصفاً دقيقاً لانتشار النقاط حول مركز معلوم (معدل المركز الحسابي) وهناك عدة طرق لقياس المسافة المعيارية وتتبع الطريقة المباشرة والتي تكون خطواتها كالاتي :-

① - ايجاد الانحراف المعياري (لمواقع الظاهرة المراد دراستها من المتغير (X) وكما يلي :-

$$\text{الانحراف المعياري لـ (X)} \quad = \quad S_X = \sqrt{\frac{\Sigma X_i^2}{n} - \left(\frac{\Sigma X_i}{n}\right)^2}$$

② - ايجاد الانحراف المعياري (لمواقع الظاهرة المراد دراستها من متغير (Y)) وكما يلي :-

$$\text{الانحراف المعياري لـ (Y)} \quad = \quad S_Y = \sqrt{\frac{\Sigma Y_i^2}{n} - \left(\frac{\Sigma Y_i}{n}\right)^2}$$

③ - ايجاد المسافة المعيارية الكلية وذلك بجمع الانحرافين المعياريين وكالاتي :-

$$\text{المسافة المعيارية} \quad S = S_X + S_Y$$

بعد ايجاد وتحديد المسافة المعيارية يمكن رسم دائرة حول مركز المعدل حيث ان نصف القطر لهذه الدائرة يمثل المسافة المعيارية الذي يوضح مدى انتشار المواقع

حول مركز المعدل وباعتماد المسافة المعيارية الذي يوضح مدى انتشار المواقع حول مركز المعدل وباعتماد المسافة المعيارية يمكن دراسة توزيع السكان وتوزيع النشاطات التجارية وغيرها من الظواهر المراد دراستها.

يكون مركز المعدل لمعظم الظواهر متقارباً وفي بعض الاحيان متساوياً اما المسافة المعيارية فمعظم الاحيان لابد ان تكون متباينة اي مختلفة فيما بينها.

مثال :-

استخرج المسافة المعيارية لثمانية مدن كانت بياناتها كالاتي :-

City مدن	fromX	from Y	X^2	Y^2
A	4	1.2	16	1.44
B	3.2	1.8	10.24	3.24
C	3	2.7	9	7.29
D	3.2	3.7	10.24	13.69
E	2.2	2.3	4.84	5.29
F	2.1	3.6	4.41	12.96
G	1.7	3.2	2.89	10.29
H	1	2.9	1	8.41
	20.4	21.4	58.62	62.56

$$S_X = \sqrt{\frac{\Sigma X_i^2}{n} - \left(\frac{\Sigma X_i}{n}\right)^2}$$

$$S_X = \sqrt{\frac{58.62}{8} - \left(\frac{20.4}{8}\right)^2} = \sqrt{7.33 - 6.50} = 0.91$$

انتشار المواقع البعد

$$S_Y = \sqrt{\frac{\Sigma Y_i^2}{n} - \left(\frac{\Sigma Y_i}{n}\right)^2}$$

$$= \sqrt{\frac{62.56}{8} - \left(\frac{21.4}{8}\right)^2} = \sqrt{7.82 - 7.15} = 0.82$$

انتشار المواقع اقرب

المسافة المعيارية $S = S_x + S_y = 0.91 + 0.82 = 1.73$

5-5 التبعثر النسبي :-

كجل المقارنة موضوعية وذات معيار موحد نوجد مقياس التبعثر النسبي والذي يمثل قيمة المسافة المعيارية مقسوماً على نصف القطر لمسافة الظاهرة المدروسة أي انه :-

$$\text{التبعثر النسبي} = \frac{\text{المسافة المعيارية}}{\text{نصف قطر المسافة للظاهرة المدروسة}}$$

مثال :-

لاربعة مدن اوجد مقياس التبعثر النسبي لمدى تمركز النشاط التجاري فيها اذا توفرت لديك المعلومات في الجدول ادناه :-

المدن	المسافة المعيارية	نصف مسافة الظاهرة المدروسة	التبعثر النسبي
A	2.64	3.40	$\frac{2.64}{3.40} = 0.77$
B	15.98	1.5	10.6
C	9.11	3	3.03
D	5.6	5.8	0.96

نلاحظ ان المدينة A تبعثرها النسبي اقل قيمة من قيم التبعثر النسبي للمدن الاخرى وهذا يعني مدى تمركز النشاط التجاري في تلك المدن كما نلاحظ ان المسافة المعيارية لتلك المدينة اقل من قيم المسافات المعيارية لبقية المدن والذي يوضح ان تشتت وتبعثر النشاطات التجارية لتلك المدينة اقل من المدن الأخرى.

تحليل الجار الأقرب :-

وهو احد تقنيات تحليل الانماط المكانية والواسعة الانتشار والاستخدام والذي أخذ بنظر الاعتبار المسافات بين النقاط المتجاورة وذلك يعني :-

١- عندما تكون المسافة بين النقاط قصيرة بمعنى ان النقاط متكتلة اي المسافة = صفر .

٢- عندما تكون المسافة بين النقاط كبيرة بمعنى ان النقاط موزعة بشكل غير عشوائي

٣- " " " " " مساوية للواحد فأن النقاط تتوزع بشكل عشوائي.

ان هذا التحليل يعتمد على مسافات كل نقطة مع النقاط الاقرب اليها من هنا جاءت تسمية هذا التحليل بإجار الأقرب ، ويتم حساب مسافات التجاور كالاتي :-

أولاً :- يتم تحديد النقاط المطلوب تحليل المسافة بينها ويجب ان تكون الظاهرة المدروسة التي يطلب تحليل المسافة لمواقعها ذات احجام متساوية ومتقاربة او تكون وظائفها متشابهة .

ثانياً :- ايجاد معدل المسافة الحقيقية لقيم المقارنة الجغرافية في درجة تقارب او تباعد النقاط عن بعضها من خلال الصيغة التالية :-

$$\text{Real Average} = A = \frac{\sum_{i=1}^{n} X_i}{n}$$

حيث أن :-

A = Real Average ، هو معدل المسافة الحقيقية .

$\sum X_i$ ، مجموع المسافات بين النقاط

n ، عدد النقاط

ثالثاً :- ايجاد كثافة التوزيع والتي يرمز لها بالرمز (K) من خلال الصيغة التالية :-

$$K = \frac{n}{Volume}$$

حيث أن :-

K = كثافة التوزيع .

n = عدد النقاط .

Volume = المساحة [ونقصد بها في المثال] .

رابعاً :- ايجاد معدل المسافة المتوقع والذي يرمز اليه بـ (E.A) وذلك من خلال استخدام الصيغة التالية :-

$$Expected\ Average = E.A = \frac{1}{2\sqrt{K}}$$

حيث أن :-

E.A = معدل المسافة المتوقع .

$\sqrt{K}$ = جذر قيمة كثافة التوزيع المستخرج سابقاً .

خامساً :- حساب دليل المجاورة والذي يرمز له (N.D) من خلال الصيغة التالية :-

حيث أن :-

N.D = دليل المجاورة

A = المعدل الحقيقي للمسافة .

E.A = المعدل المتوقع للمسافة .

سادساً :- اختبار الفرضيات بعد طرحها وتحديدها والفرضيات تتوعين هما :-

Ho : The distribution of points is random

H1 : The distribution of points is non-random

ويقصد ذلك ان الفرضيات (Hypothesis) نوعين الاول تسمى بفرضية العدم (الصفرية) ويرمز لها بالرمز (H_0) وغالبا تثبت الفرضية التي نريدها فمثلا ان التوزيع عشوائي.

اما الفرضيات الثانية هي الفرضية البديلة ويرمز لها بالرمز (H_1) وهي عكس او تنفي الفرضية الاولى دائما اي أن :-

H_0 : التوزيع عشوائي

H_1 : التوزيع غير عشوائي

سابعاً :- ايجاد قيمة (Z) المحتسبة وهي قيمة نستخدمها لغرض المقارنة لبيان قبول او رفض الفرضيات وحسب كما في الصيغة التالية :-

$$Z_c = \frac{A - E.A}{Sx/\sqrt{n}}$$

حيث أن :-

Z_c = قيمة Z_c المحتسبة

A = معدل المسافة الحقيقي

$E.A$ = معدل المسافة المتوقع

n = عدد النقاط

Sx = الانحراف المعياري والذي يحسب كما في السابق $Sx = \sqrt{\frac{\Sigma Xi^2}{n} - \left(\frac{\Sigma Xi}{n}\right)^2}$

ثامناً :- ايجاد قيمة Z_t وتعني الجدولية عن طريق الجداول الاحصائية (التوزيع الطبيعي) والقيم التالية تمثل قيم Z_t لكل مستوى المعنوية α (اي مدى الثقة) [تحفظ].

مستوى المعنوية α	قيمة Z_t
0.99	2.326
0.975	1.96
0.95	1.645
0.90	1.282

مثال :- اذا علمت ان المسافات بين مواقع (6) مكاتب للخدمات الكاسوية في منطقة مساحتها ($5\ Km^2$) موضحة في ادناه ، بين فيما اذا كان توزيع هذه المواقع عشوائياً ام لا بأستخدام تحليل الجار الأقرب اذا علمت ان مستوى المعنوية هو $\alpha = 0.95$ ؟

الحل :

	نسبة المسافات لأقرب جار	X_i	X_i^2
1	0.90	0.90	0.81
2	0.62	0.69	0.38
3	0.75	0.75	0.56
4	0.65	0.65	0.42
5	0.78	0.78	0.61
6	0.79	0.79	0.62
		4.49	3.4

① $A = \frac{\Sigma X_i}{n} = \frac{4.49}{6} = 0.75$

② $K = \frac{n}{Volum} = \frac{6}{5} = 1.2$

③ $E.A = \frac{1}{2\sqrt{K}} = \frac{1}{2\sqrt{1.2}} = 0.45$

④ $D.N = \frac{A}{E.A} = \frac{0.75}{0.45} = 1.66$

⑤ $Z_c = \frac{A - E.A}{Sx / \sqrt{n}}$

$Sx = \sqrt{\frac{\Sigma X_i^2}{n} - \left(\frac{\Sigma X_i}{n}\right)^2} = \sqrt{\frac{3.4}{6} - \left(\frac{4.49}{6}\right)^2} = \sqrt{0.007} = 0.08$

$Z_c = \frac{0.75 - 0.45}{0.08 / \sqrt{6}} = \frac{0.3}{0.03} = 10$

$\because \alpha = 0.95 \Longrightarrow Z_t = 1.64$

$1.64 < 10 \Rightarrow Z_t < Z_c$

بما ان Z_c المحتسبة > اكبر من Z_t الجدولية اذن نرفض الفرضية الصفرية اي اننا نرفض ان التوزيع عشوائي و نقبل الفرضية البديلة اي ان توزيع هذا الواقع غير عشوائي.

تحليل المربعات القياسية :- يستخدم هذا التحليل في تحليل كثافة التوزيع وعناصر تمركز النقاط وكذلك مناطق تبعثرها اي تعبير رقمي عن نمط توزيع تلك النقاط وذلك من خلال تسجيل تكرار النقاط في مواقع مختلفة من الخارطة وحسب الخطوات التالية :-

١) - تقسيم الخارطة الى مجموعة من المربعات الصغيرة المتساوية في القياس.

٢) - حساب تكرار وجود النقاط في المربعات وتحويلها الى جدول تكراري.

٣) - الجدول التكراري يمثل التكرارات الملاحظة او المشاهدة ويتم مقارنتها بالتكرارات المتوقعة.

أي أننا نستخدم طريقة احصائية لغرض المقارنة بين التكرارات المشاهدة والمتوقعة وهذا (اختبار مربع كاي χ^2 واختبار t واختبار كولمكروف - سميرنوف للعينة الواحدة).

٤) - يمكن اختصار خطوات الحل كما في المخطط ومن اليسار الى اليمين وكما موضح ادناه :-

← حساب تكرار متجمع صاعد ③ ← حساب التكرار المشاهد ② ← حساب عدد النقاط ①

← حساب التكرار المتوقع = مجموع التكرارات / عددها ⑤ ← حساب نسبة التراكم المشاهد = التكرار المتجمع / مجموع التكرارات ④

← نسبة التكرار المتوقع = المتجمع المتوقع / مجموع التكرارات ⑦ ← حساب التكرارات المتوقعة المتجمعة ⑥

اختبار كولمكروف ⑨ ← خطوة ⑤ - خطوة ④ ⑧

مثال :- باستخدام تحليل المربعات اللاتينية بين فيما اذا كان هناك نمط معين ——— (فرقه) لانتشار ظاهرة السكن العمودي لشبكة المربعات الآتية عند مستوى المعنوية $\alpha = 0.05$ ؟

على فرض ان هذه فاطرة وتم تأشير بعض المواقع

2		6
5	4	3
	1	3

الحل :-

أولاً" وضع الفرضيات والتابتر وهي :-

H_0 : لايوجد فرقين لتوزيع المشاهد والتوزيع المتوقع

H_1 : يوجد فرق ر ر ر ر ر

عدد النقاط	التكرارات المشاهدة	تكرار متجمع صاعد	تكرار متجمع / مجموع التكرارات = نسبة التراكم المشاهد
عدم وجود نقطة	2	2	2/9 = 0.22
وجود نقطة واحدة	1	3	3/9 = 0.33
وجود نقطتين	1	4	4/9 = 0.44
ر ثلاث نقاط	2	6	6/9 = 0.66
ر اربع نقاط	1	7	7/9 = 0.77
ر خمسة نقاط	1	8	8/9 = 0.88
ر ستة نقاط	1	9	9/9 = 1
	9		

التكرار المتوقع = مجموع التكرارات / عددها	التكرارات المتوقعة المجمعة	نسبة التراكم المتوقع = المتجمع المتوقع / مجموع التكرارات	عامود 7 - عامود 4
9/7 = 1.28	1.28	1.28/9 = 0.14	0.22 - 0.14 = 0.08
9/7 = 1.28	2.56	2.56/9 = 0.28	0.33 - 0.28 = 0.05
=	3.84	3.84/9 = 0.43	0.44 - 0.43 = 0.01
=	5.12	5.12/9 = 0.57	0.66 - 0.57 = 0.09
=	6.4	6.4/9 = 0.71	0.77 - 0.71 = 0.06
=	7.68	7.68/9 = 0.85	0.88 - 0.85 = 0.03
=	8.96	8.96/9 = 0.99	1 - 0.99 = 0.01

من العامود الاخير نستخرج اكبر قيمة (اكبر صفر) وهو 0.09 ويمثل قيمة χ^2 المحسوبة والتي تقارن مع القيمة المجدولة لاختبار مربع كاي χ^2 لدرجة حرية = مجموع التكرارات = 9 . وبما ان القيمة المحسوبة < القيمة الجدولية $\Leftarrow$ $0.09 < 4.17$

∴ نقبل الفرضية الصفرية اي ليس هناك فرق بين التوزيع الشاهد والتوزيع العامودي للسكن .

ملاحظة :- لاستخراج قيم χ^2 المجدولة يمكن اعتماد الجدول التالي لدرجة حرية ومستوى المعنوية (α) وكما يلي :-

درجة الحرية d.f	مستوى المعنوية (α) 0.90	0.95	0.99
1	⋮	⋮	⋮
2	⋮	⋮	⋮
⋮			
8	3.49	2.73	1.65
9	4.17	3.33	2.09
10	4.87	3.93	2.56

تحليل الأنماط الشبكية :- وتسمى ايضاً تحليل الانماط الخطية ويمكن تعريف الانماط الشبكية انه عند اسقاط مجموعة من الخطوط على الخارطة تمثل العلاقة بين مجموعة من النقاط او المناطق مثل حركة المرور ، هجرة السكان ، هجرة الطيور ، خطوط السكك الحديدية الحدود الدولية بين دولتين ، --- الخ وتتباين اشكال الخطوط تبعاً للحالة التي تمثل يعني ذلك اذا كانت الخطوط ممثلة للحركة الحقيقية (حركة المرور مثلاً) فيكون عرض الخط مناسباً بحجم حركة المرور بين المواقع المدروسة (اذا كانت بحركة كثيفة فان الخط يكون عريضاً ، واذا كانت الحركة خفيفة فيكون الخط رفيعاً وهكذا ...).

اما اذا كانت الخطوط رمزية فانها تعبر عن وجود الصلة بين الموقعين دون الالتزام بالمسار الحقيقي (الطرق وتبدلاته او تعرجاته) مثل خطوط النقل البحري او الاتصالات بين مناطق المدينة الواحدة - - الخ .

اذ تستخدم في هذه الحالة انواع مختلفة من الخطوط المتصلة او المتقطعة أو

المتعرجة وان مثل هذه الخطوط على الخارطة تقدم معلومات عن منطقة معينة في زمن معين ، اي انها تعطي معلومات ثابتة و ببعدين (الطول والعرض) ويمكن وصف الحالة وفق وحدة زمنية محددة مثل (عالدن / يوم ، او سيارة / ساعة ، أو لدادات / دقيقة) ومن انواع الانماط الشبكية هي شبكة المواصلات.

خصائص شبكة المواصلات :-

١) - لكل شبكة عدد محدد من النقاط (الأماكن) .

٢) - كل طريق يوصل بين نقطتين مختلفتين فقط .

٣) - كل نقطتين مرتبطتين بأكثر من طريق واحد .

٤) - جميع الطرق مزدوجة الاتجاه (ذهاباً وأياباً) مالم تتم الاشارة بغير ذلك

ايجاد درجة الاتصال لشبكة الاتصالات :-

لحساب درجة الاتصال لأي شبكة من شبكات الاتصالات يجب الاعتماد على ماقدمه العالم (ابلر Abler) في هذا المجال حيث وجد ثلاث معادلات لحساب ولتوضيح واقع شبكة الاتصالات وهذه المعادلات هي :-

١) - الاتصال الكلي Real Connectivity :-

والذي يمكن حساب قيمته من خلال الصيغة التالية :-

$$\text{Real con.} = \frac{w}{(n^2-n)/2}$$

حيث أن :-

Real con. = الاتصال الكلي

w = (Ways) = عدد الطرق

n = عدد النقاط

٦- اعلى اتصال Maximum Connectivity (Max. con.) :-
والذي يمكن حسابه رياضيا من خلال الصيغة التالية :-

$$\therefore \text{Max. Con.} = \frac{(n^2-n)/2}{(n^2-n)/2} = 1$$

اي ان اعلى اتصال = 1 دائما .

٧- ادنى اتصال Minimum Connectivity (Min. Con.) :-
وهو معادلة ادنى اتصال والذي يمكن حسابها باستخدام الصيغة التالية :-

$$\text{Min. Con} = \frac{(n-1)}{(n^2-n)/2}$$

حساب درجة تطور شبكة الاتصالات :-

هناك عدة طرق لقياس أو حساب درجة تطور شبكة الاتصالات وهنا (دليل بيتا ، ودليل كاما و دليل ألفا) ، ممكن السهولة حساب كل من دليل بيتا ودليل كاما ودليل الفا وكالتالي :-

① - Beta Index $= \frac{w}{n}$ دليل بيتا

② - Gamma Index $= \frac{w}{3/(n-2)}$ دليل كاما

حيث أن :-
w = عدد الطرق
n = عدد النقاط

اما دليل ألفا فيعبر عنه في حالة الظروف الاقتصادية لأي دولة بين نقاط معينة حيث يعتمد في حسابه على مجموع اطوال الطرق المستخدمة في شبكة الاتصالات .

ملاحظات مهمة :-

١)- اذا كانت قيمة دليل بيتا = صفر يدل ذلك على عدم وجود اي اتصال.

٢)- اذا كانت قيمة دليل بيتا اكبر من الواحد يدل ذلك على وجود اكثر من طريق يصل بين نقطتين او اكثر .

٣)- اذا كانت قيمة دليل بيتا اقل من الواحد يدل ذلك على عدم وجود اي طريق بين النقاط بصورة مباشرة .

٤)- اذا كانت قيمة دليل گاما = 1 فان ذلك يدل على وجود اتصال كامل بين النقاط .

مثال :- تمكن باحث من تحديد وجود (7) مراكز لمدن مرتبطة مع بعضها ـــــ بـ (27) طريق سياحي، جد درجة الاتصال بين هذه المراكز ثم أحسب كل من معامل بيتا ومعامل گاما.

الحل :- حساب درجة الاتصال نقوم بايجاد كل من :-

① Real Con. $= \frac{w}{(n^2-n)/2}$

$= \frac{27}{(7^2-7)/2} = \frac{27}{(49-7)/2} = \frac{27}{21} = 1.2$ الاتصال الحالي

② Max. Con. $= \frac{(n^2-n)/2}{(n^2-n)/2} = \frac{(7^2-7)/2}{(7^2-7)/2} = \underline{1}$

③ Min. Con $= \frac{(n-1)}{(n^2-n)/2} = \frac{7-1}{(7^2-7)/2} = \frac{6}{21} = 0.21$

من خلال نتائج المعادلات فان شبكة الاتصالات جيدة حيث ان واقعها الحقيقي اكثر بقليل من اعلى اتصال ممكن ان تصل اليه الشبكة وحساب معامل (دليل) بيتا و معامل (دليل) گاما :-

① Beta Index $= \frac{w}{n} = \frac{27}{7} = 3.8$

② Gamma Index $= \frac{w}{3/(n-2)} = \frac{27}{3/(7-2)} = 45$

"الفصل السادس"

تحليل العلاقة والاشتراك المكاني

عندما يهتم الباحث الجغرافي بسلوك ظاهرة او متغير واحد فيستخدم احد مقاييس النزعة المركزية او احد مقاييس التشتت واذا كانت بياناته تخص ظاهرتين او متغيرين او اكثر ففي هذه الحالة يستخدم مقاييس تحليل الارتباط (تحليل العلاقة).

خصائص الارتباط :-

١) - معامل الارتباط يتراوح قيمته مابين (١-) و(١+) ويرمز اليه بالرمز (R) أي ان : $-1 \leq R \leq +1$ فلا يتجاوز هاتين القيمتين.

٢) - الارتباط يشير الى الاقتران بين ظاهرتين او اكثر (اي متغيرين او اكثر)، ويقصد بذلك ان تغير قيم احد المتغيرين في اتجاه معين قد تغير قيم المتغير الآخر.

٣) - عندما يتغير قيم احد المتغيرين في اتجاه وقيم المتغير الآخر في الاتجاه نفسه للأول فأننا بذلك نحصل على مايسمى بالارتباط ذو اتجاه طردي (اي موجب +) مثلاً" عندما ترتفع قيم احد المتغيرين ترتفع قيم الآخر او بالعكس عندما تقل قيم احد المتغيرين تقل قيم احد المتغيرين تقل قيم الآخر.

٤) - عندما يتغير قيم احد المتغيرين في اتجاه وقيم المتغير الآخر في اتجاه عكسي للأول فأن الارتباط الناتج تكون ذو اتجاه عكسي (اي سالب -) مثلاً" عند ارتفاع قيم احد المتغيرين تقل قيم المتغير الآخر او بالعكس عند انخفاض قيم احد المتغيرين ترتفع قيم الآخر.

٥) - ان وجود الارتباط بين متغيرين قد لاينشأ عن علاقة سببية مباشرة بينهما او قد يحصل نتيجة لعوامل اخرى اي عند تحليل الارتباط مابين متغيرين لابد من الأخذ بنظر الاعتبار وجود علاقة منطقية بين المتغيرين اي المعرفة بطبيعة المشكلة الجغرافية موضوع الدراسة ويتم ذلك بتحديد اي المتغيرين متغير مستقل واياهما المتغير المعتمد (اي الذي يعتمد في قيمه على المتغير المستقل).

ملاحظة مهمة :-

١)- عندما $R=+1$ ، $R=0.9$ ، $R=0.8$ ، $R=0.7$. معناه أن الارتباط قوي وذو اتجاه طردي (موجب) .

٢)- عندما $R=-1$ ، $R=-0.9$ ، $R=-0.8$ ، $R=-0.7$ معناه أن الارتباط قوي وذو اتجاه عكسي (سالب) .

٣)- عندما $R=0.6$ ، $R=0.5$ معناه ان الارتباط متوسط القوة وذو اتجاه طردي.

٤)- عندما $R=-0.6$ ، $R=-0.5$ معناه ان الارتباط متوسط القوة وذو اتجاه عكسي.

٥)- $R=0.1, 0.2, 0.3, 0.4$ معناه ان الارتباط ضعيف وذو اتجاه طردي .
وعندما $R=0.1, 0.2, 0.3, 0.4$ معناه ان الارتباط ضعيف وذو اتجاه عكسي.

٦)- عندما يكون صفر $=R$ معناه عدم وجود اي ارتباط بين المتغيرين (عدم وجود اي علاقة)

ومن اهم مقاييس تحليل العلاقة ماياتي :-

اولاً : اختبار مربع كاي χ^2 :-

يستخدم هذا الاختبار في اجراء الموازنة مابين التكرارات المشاهدة والمتوقعة. وايضاً يستخدم عندما تكون حجم العينة (5) فاكثر ووجود عدد كبير من التكرارات ولاستخراج قيمة χ^2 نتبع الخطوات الآتية الموضحة في المثال :-

مثال :- باستخدام اختبار مربع كاي χ^2 ، وضح فيما اذا كانت هناك علاقة مابين توزيع القرى والتكوين الجيولوجي في منطقة معينة من السهول في ادناه (5 تكوينات) .

عدد القرى (تكرار المشاهد)	النسبة المئوية للتكوين الجيولوجي	التكرار المتوقع	(تكرار مشاهد − تكرار متوقع)^2	(تكرار مشاهد − تكرار متوقع)^2 / تكرار متوقع
30	25%	15	225	15
15	30	18	9	0.5
6	20	12	36	3
8	20	12	16	1.33
1	5	3	4	1.33
60	100%			21.16

القيمة الجدولية 9.48 وتقارن بالمحتسبة = 21.16

χ^2_c المحتسبة $> \chi^2_t$ الجدولية $\Leftarrow 21.16 > 9.48$

∴ نرفض الفرضية الصفرية اي ان هناك فروق معنوية مابين التكرارات المشاهدة والمتوقعة.

نسبة التقاطع :-

وهي النسبة التي تقيس اشتراك متغيرين وتستخدم عادة في الموضات والاستبيانات الميدانية ويتم الاعتماد على الجدول الآتي في استخراج نسبة التقاطع ويسمى بالجدول التوافقي وهو نفسه يستخدم في ايجاد معامل يول ، معامل فاي .
حيث يتم الاعتماد هنا على متغيرين اثنين فقط وكل متغير اتجاهين (موضح في المثال) :

Y \ X	①	②	المجموع
①	a	b	a+b
②	c	d	c+d
المجموع	a+c	b+d	

$$\text{نسبة التقاطع} = \frac{(a)\cdot(d)}{(b)\cdot(c)}$$

فاذا
فاذا كانت النسبة مساوية الى الواحد (1) فهذا يعني عدم وجود اشتراك بين المتغيرين .
وكلما ازدادت النسبة كلما دل هذا على وجود علاقة متآزرة بينهما .

مثال :- جد نسبة التقاطع بين متغيري (الحضر) و (امتلاك مكتبة شخصية) ——— كما موضح في الجدول التوافقي :-

الحضر \ امتلاك مكتبة	وجود مكتبة	عدم وجود مكتبة	المجموع
حضر	70	28	98
ريف	50	52	102
المجموع	120	80	200

الحل :-

$a = 70$ ، $b = 28$

$c = 50$ ، $d = 52$

$$\text{نسبة التقاطع} = \frac{(70)(52)}{(28)(50)} = \frac{3640}{1400}$$

$$= 26$$

هذه النسبة اكبر من (1) بكثير ، يدل ذلك على ان العلاقة الموجودة بين المتغيرين قوية .

الخطوة الاولى : الفرضية الصفرية :

عدم وجود فروقات مابين التكرارات المشاهدة (عدد القرى) و Ho :
التكرارات المتوقعة

الفرضية البديلة :-

وجود فروقات مابين التكرارات المشاهدة والتكرارات المتوقعة H1 :

الخطوة الثانية : عدد القرى تعتبر تكرارات مشاهدة ، اذاً لابد من ايجاد التكرارات المتوقعة وكالاتي :-

التكرار المتوقع Ei = مجموع التكرارات المشاهدة Oi X النسبة المئوية للتلوث الجيولوجي

فمثلاً التكرار المتوقع لاول E_1

$$E_1 = 60 \times \frac{25}{100} = 15$$

الخطوة الثالثة : ايجاد مربع الفرق مابين التكرار المشاهد والتكرار المتوقع ومن ثم قسمة الناتج على التكرار المتوقع المقابل لنتيجة الفرق وكما هو موضح في العامودين الرابع والخامس .

فمثلاً القيمة الاولى :-

$$(O_1 - E_1)^2 = (30 - 15)^2 = (15)^2 = 225$$

$$\frac{225}{15} = 15$$

القيمة الثانية

$$(O_2 - E_2)^2 = (15 - 18)^2 = (-3)^2 = 9 \Rightarrow \frac{9}{18} = 0.5$$

وهكذا ...

الخطوة الرابعة : جمع العامود الاخير (الخامس) والذي يكون نتيجة الصيغة الآتية :-

مجموع (تكرار مشاهد - تكرار متوقع)^2 / التكرار المتوقع

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{O_i}$$

قيمة مربع كاي المحسوبة

الخطوة الخامسة : تقارن قيمة مربع كاي المحسوبة بالقيمة الجدولية لمربع كاي اعتماداً على درجة الحرية العلوية اي $(N-1)$: القيمة الجدولية عند مستوى معنوية $\chi^2_{N-1,\alpha}$

$\alpha = 0.95$ وعدد المجاميع $= N$.

وفي مثالنا تكون القيمة الجدولية $\chi^2_{4,0.95}$ فنأخذ القيمة المناسبة .

من ملاحظة الجدول نعتمد على درجة الحرية العلوية اي $N-1$ حيث ان N = عدد المجاميع وفي مثالنا عدد المجاميع = 5 اذاً درجة الحرية $5-1=4$

ومستوى المعنوية $0.95 = 95\%$ فنختار

جدول القيم الجدولة لقيمة χ^2

درجات الحرية	مستوى المعنوية 0.95
2	5.99
3	7.81
4	9.48
5	11.07
8	15.50

ثالثاً:- معامل يول :-

تعتمد هذه الطريقة على الجدول التوافقي ايضاً، اي على متغيرين وبمراتين (اتجاهين) لكل متغير، حيث يستخدم العلاقة التالية :-

$$\text{معامل يول} = \frac{(a \cdot d) - (b \cdot c)}{(a \cdot d) + (b \cdot c)}$$

ويتم الاعتماد على مضامين قيمة الارتباط التي سبق ان ذكرناها في تفسير نتائج هذا المعامل .

مثال :- في المثال السابق (في شعبة القاطع) جد قيمة معامل يول .

$$\text{معامل يول} = \frac{3640 - 1400}{3640 + 1400} = \frac{2240}{5040} = 0.44$$

$$\because \quad -1 \leq R \leq +1$$

$\therefore$ 0.44 تقع ضمن الفترة المحددة نستنتج من ملاحظتنا القيمة ان قيمة معامل يول ضعيفة وباتجاه طردي اي ان العلاقة مابين المتغيرين امتلاك مكتبة و التحضر ضعيفة ولكنها باتجاه طردي .

رابعاً:- معامل فاي :-

نستند لايجاد هذا المعامل الجدول التوافقي وتطبيق الصيغة الآتية :-

$$\text{معامل فاي} = \frac{(a \cdot d) - (c \cdot b)}{\sqrt{(a+c) \cdot (b+d) \cdot (a+b) \cdot (c+d)}}$$

ثم نجد القيمة المحتسبة من خلال الآتي :-

القيمة المحتسبة = المجموع الكلي للجدول التوافقي X مربع معامل فاي

وتقارن القيمة المحتسبة مع القيمة المجدولة لاختبار مربع كاي بدرجة حرية مساوية الى (عدد مجاميع المتغيرات - 1) اي ان درجة الحرية تساوي (ا-ن) ، ن عدد المجاميع

وبمستوى معنوية 99٪ ونختبر الفرضية الصفرية :-

H_0 : عدم وجود علاقة بين المتغيرين

H_1 : وجود علاقة بين المتغيرين

المثال الذي تم ذكره في نسبة التقاطع ، اختبر فيما اذا كان هناك علاقة ما بين المتغيرين باستخدام معامل فاي .

الحل :-

مكتبة شخصية / نوع العينة	وجود مكتبة	عدم وجود مكتبة	المجموع
حضر	70	28	98
ريف	50	52	102
المجموع	120	80	200

H_0 : عدم وجود علاقة بين الحضر واقتناء مكتبة شخصية

H_1 : وجود علاقة بين الحضر واقتناء مكتبة شخصية

$$\text{معامل فاي} = \frac{(70 \cdot 52) - (28 \cdot 50)}{\sqrt{(120) \cdot (80) \cdot (98) \cdot (102)}}$$

$$= \frac{3640 - 1400}{\sqrt{95961600}} = \frac{2240}{9795.99} = 0.228 \simeq 0.23$$

$$\text{القيمة المحسوبة} = 200 \times (0.23)^2 = 200 \times 0.053 = 10.6$$

تقارن القيمة المحسوبة (10.6) مع الجدولة (جداول اختبار مربع كاي) وبدرجة حرية (1) ومستوى معنوية 0.99 وهي تساوي (6.63) .

القيمة المحسوبة > الجدولية

لذا نرفض الفرضية القائلة بعدم وجود علاقة ما بين المتغيرين اي ان هناك علاقة احصائية ما بين التحضر واقتناء مكتبة شخصية .

ثانياً : معامل كاما :-

يستخدم هذا المعامل عندما تكون البيانات محتوية على أكثر من متغيرين او أكثر من قراءة وتكون الاستخدام كالاتي :-

مثلاً لو كان لدينا متغيرين وكل متغير ثلاث قراءات وبترقيم المربعات بالارقام (1, 2, ... ,9) فلتحديد العلاقة مابين المتغيرين (قوة العلاقة) نتبع الخطوات الآتية :-

	المتغير الأول		
	3	2	1
المتغير الثاني	6	5	4
	9	8	7

الخطوة الأولى : نبدأ من اليمين لايجاد قيمة (P)

- نضرب قيمة المربع (1) X مجموع قيم المربعات الباقية بعد حذف المربعات التي تقع في صف اوعامود قيمة المربع (1).

(1) X (9+8+7+5)

- نضرب قيمة المربع (2) X مجموع قيم المربعات الباقية بعد حذف المربعات التي تقع في صف اوعامود القيمة (2) وتحذف اول عامود.

(2) X (9 + 6)

- نستمر بالنسبة لقيم المربعين 4 و5 ونتبع نفس الأسلوب السابق

(4) X (9 + 8)

(5) X (9)

- نجمع نواتج حواصل الضرب اعلاه والتي تساوي قيمة (P).

الخطوة الثانية :-

الآن نبدأ بايجاد قيمة (ب)

نفس الطريقة السابقة ولكن من اليسار فيكون لدينا :-

(3) X (4+5+7+8)

(2) X (4 + 7)

(6) X (7 + 8)

(5) X (7)

نجمع النواتج ونساويها بالقيمة (ب) .

الخطوة الثالثة :- نطبق الصيغة الآتية لايجاد معامل گاما

$$\text{معامل گاما} = \frac{\text{ا}-\text{ب}}{\text{ب}+\text{ا}}$$

مثال :- حدد درجة العلاقة بين المستوى التعليمي والطبقة الاجتماعية للبيانات الآتية :-

المستوى التعليمي / الطبقة الاجتماعية	جامعي	ثانوي	ابتدائي
الطبقة العليا	12	6	2
الطبقة المتوسطة	8	10	2
الطبقة الفقيرة	0	4	16

الحل :-

1- ايجاد قيمة (ا)

$384 = (16+4+2+10) \times 2$

$108 = (16+2) \times 6$

$160 = (16+4) \times 8$

$160 = 16 \times 10$

$\text{ا} = 384 + 108 + 160 + 160 = 812$

2- ايجاد قيمة (ب)

$44 = (4+0+10+8) \times 2$

$48 = (0+8) \times 6$

$8 = (0+4) \times 2$

$0 = (0) \times 10$

$\text{ب} = 44 + 48 + 8 + 0 = 100$

3- معامل گاما $= \frac{\text{ا}-\text{ب}}{\text{ب}+\text{ا}} = \frac{812-100}{812+100} = \frac{712}{912} = 0.78$

نستدل من القيمة الناتجة ان هناك علاقة قوية بين المستوى التعليمي والطبقة الاجتماعية حيث ان $-1 \leq R \leq +1$.

سادساً: معامل سبيرمان: يستخدم هذا المعامل لقياس درجة العلاقة مابين متغيرين قد يكونان وصفيان وليس كميان اي ان قيمهما ليست كمية وانما وصفية والمثال الآتي يوضح لنا كيف نستخدم هذا المعامل:-

مثال:- تم تصنيف نوعية بذور محصنة تم استلامها من شركتين X و Y وكما مبين في ادناه حيث تم التصنيف بدرجات ممتاز، جيد جداً، جيد، متوسط، ضعيف وحسب الوجبات الستة التي تم استلامها، جد درجة العلاقة بين نوعيات بذور الشركتين.

نوعية بذور الشركة Y	نوعية بذور الشركة X
جيد	ضعيف
ضعيف	جيد
ممتاز	ممتاز
ممتاز	جيد
ضعيف	جيد
متوسط	جيد جداً

الحل:- اولاً نرتب القيم تصاعدياً ثم نعطيها رتبة، وعند تكرار نفس القيمة مثلاً (ضعيف، ضعيف) فأننا نجمع رتبها ونقسمها على عددها مثل الرتبة [2، 3، 4] في X اخذت الرتبة 3 ومثل الرتبة [1، 2]، [5، 6] في Y اخذت الاولى والثانية وقسمت على عددها فأصبحت رتبتهما (1.5) وهكذا بالنسبة للخامسة والسادسة اخذت الرتبة (5.5).

رتبة Y	Y بعد الترتيب	رتبة X	X بعد الترتيب تصاعدياً
1.5 (1+2/2)	ضعيف	1	ضعيف
1.5	ضعيف	3 (2+3+4/3)	جيد
3	متوسط	3	جيد
4	جيد	3	جيد
5.5 (11/2)	ممتاز	5	جيد جداً
5.5	ممتاز	6	ممتاز

ثانياً :- نرجع الى البيانات الأصلية ونعطيها الرتبة لكل قيمة كما يأتي :-

نوعية بذور X	الرتب	نوعية بذور Y	الرتب	$d = X - Y$	d^2
ضعيف	1	جيد	4	$1-4=-3$	9
جيد	3	ضعيف	1.5	$3-1.5=1.5$	2.25
ممتاز	6	ممتاز	5.5	$6-5.5=0.5$	0.25
جيد	3	ممتاز	5.5	$3-5.5=-2.5$	6.25
جيد	3	ضعيف	1.5	$3-1.5=1.5$	2.25
جيد جداً	5	متوسط	3	$5-3=2$	4
					$\Sigma d^2 = 22$

ثالثاً :- نجد قيم الرقم بين X و Y ثم نرتبها هذه القيم ونجمعها وكما هو واضح في العامودين الخامس والسادس اعلاه ثم نطبق الصيغة الآتية :-

$$R = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6(22)}{6(6^2-1)} = 1 - \frac{132}{6(36-1)} = 1 - \frac{132}{210} = 0.37$$

رابعاً :- معامل بيرسون :- يستخدم لإيجاد مدى العلاقة مابين متغيرين تكون قيمهما عبارة عن بيانات رقمية (كمية) والصيغة الخاصة به هي :-

$$R = \frac{n\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

ملاحظات مهمة :- الفرق بين معامل بيرسون ومعامل سبيرمان :-

1) - ان معامل بيرسون يستخدم للظواهر او المتغيرات التي تحتوي قيمها على قيم رقمية فقط ففي حالة وجود احد المتغيرات وهي (اي انه قيمة لاتعبر بارقام) فلا يستخدم هذا المعامل .

ثانياً: تحليل الانحدار Rigression Analysis :-

ان طريقة الانحدار هي نموذج بصورة تقريبة يساهم في تلخيص العلاقات المعقدة للعالم الواقعي ويبحث في ظاهرتين او متغيرين او اكثر (لكن فيما تخص دراستنا سنهتم بدراسة متغيرين فقط).

فاذا كان الجغرافي يدرس ظاهرة معينة ولتكن (Y) فان اهتمامه ينصب على مدى تأثر هذه الظاهرة بظاهرة اخرى ولتكن (X) اي ان المتغيرات التي تحدث في الظاهرة (Y) تتبع وتعتمد على المتغيرات التي تحدث في الظاهرة (X) وبذلك تسمى الظاهرة (Y) او المتغير المعتمد او التابع والظاهرة (X) بالمتغير المستقل، ويتم التعبير عنها من خلال نموذج بشكل رمزي يعبر عن العلاقة بين المتغيرات وبشكل معادلة ويتوضح ذلك من المثال التالي :-

مثال :- تحدث مزارع A, B, C كمية انتاجها للقمح ومقدار الأمطار الساقطة عليها لسنة معينة كالآتي :-

المزارع	انتاج الدونم الواحد (كغم)	كمية الامطار على الدونم الواحد (ملم)
A	30	3
B	20	2
C	10	1

يتضح لنا أن تغير وحدة واحدة في كمية تنتج عنه نسبة مماثلة من التغير في انتاج الدونم وان زيادة مقياس واحد من الامطار يؤدي الى زيادة (10) عشرة مقاييس من انتاج القمح .

اذاً هناك نسبة ثابتة لتغير انتاجية القمح من الأمطار، اي ان انتاجية القمح تعتمد على الامطار في زيادة الانتاج، اذاً المتغير المعتمد هنا هو (الانتاج) والمتغير المستقل هو (كمية الامطار) .

٢) - ان معامل سبيرمان يمكن استخدامه في حالة كون احد المتغيرات كمي والآخر وصفي اي انه يستخدم اذا كان المتغيرين وصفيان او في حالة احدهما يحتوي على ارقام والآخر وصفي .

مثال : جد درجة العلاقة بين العمر (متغير Y) وعدد اللقاحات التي تم اخذها (متغير X) باستخدام معامل الارتباط بيرسون .

الحل :

Y	X	XY	X^2	Y^2
5	9	5X9=45	81	25
4	12	48	144	16
7	11	77	121	49
3	8	24	64	9
10	24	240	576	100
$\Sigma Y=29$	$\Sigma X=64$	$\Sigma XY=434$	$\Sigma X^2=986$	$\Sigma Y^2=199$

$$R = \frac{n\Sigma XY - \Sigma X \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\ \sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$R = \frac{5(434) - (64)(29)}{\sqrt{5(986) - (64)^2}\ \sqrt{5(199) - (21)^2}}$$

$$R = \frac{2170 - 1856}{\sqrt{4930 - 4096}\sqrt{995 - 841}} = \frac{314}{358} = 0.87$$

اذاً العلاقة بين متغير العمر ومتغير اللقاحات قوية وذات اتجاه طردي موجب.

معادلة الأنحدار (نموذج الانحدار) :-

يمكن كتابة نموذج الانحدار لمتغير معتمد واحد ومتغير مستقل واحد كالآتي :-

$$Y = \alpha + \beta X + E$$

حيث أن :-

- α : الفا
- β : بيتا

$\Leftarrow$ معالم (ثوابت) يتم تقديرها بواسطة طريقة المربعات الصغرى الاعتيادية

X : المتغير المستقل

Y : المتغير المعتمد (المتوقع)

E : الخطأ العشوائي والذي يكون غير معهود من قبل الباحث.

تقدير المعالم α و β :-

$$\hat{\beta} = \frac{N\Sigma XY - \Sigma X \cdot \Sigma Y}{N\Sigma X^2 - (\Sigma X)^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X}$$

X	Y	X^2	XY
.	.	.	.
.	.	.	.
.	.	.	.
ΣX	ΣY	ΣX^2	ΣXY

$$\bar{Y} = \frac{\Sigma Y}{n} \quad , \quad \bar{X} = \frac{\Sigma X}{n}$$

مثال :- اوجد معادلة خط الانحدار للبيانات التالية :-

$\Sigma X = 10$, $\Sigma Y = 50$, $\Sigma XY = 120$, $N = 5$

$$Y = \alpha + \beta X_i + E$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

$$\hat{\beta} = \frac{N\Sigma XY - \Sigma X \Sigma Y}{N\Sigma X^2 - (\Sigma X)^2}$$

$$= \frac{5(120) - (10)(50)}{5(60) - (10)^2} = \frac{600 - 500}{300 - 100} = 0.5$$

$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{10}{5} = 2$$

$$\bar{Y} = \frac{\Sigma Y_i}{n} = \frac{50}{5} = 10$$

$$\therefore \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

$$= 10 - (0.5)(2) = 9$$

$\hat{\beta} = 0.5$, $\hat{\alpha} = 9$ ، القيم المقدرة :-

∴ معادلة الانحدار التقديرية هي :-

$$\hat{Y} = \hat{\alpha} + \hat{\beta} X$$

$$= 9 + 0.5 X$$

فعندما $X = 3$ فإن $\hat{Y}$ ستكون :-

$$\hat{Y} = 9 + 0.5(3)$$

$$= 9 + 1.5 = 10.5$$

التنبؤ :-

ان الاهمية العملية لمثل هذه العلاقة الوظيفية التي تتضمن معادلة الانحدار تتمثل على امكانية استخدامها في التنبؤ لبعض قيم غير معروفة لمتغير من قيم معروفة لمتغير آخر مما لا يمكن تحقيقه عند استخدام مفهوم الارتباط .

فمثلاً لو أردنا معرفة كمية انتاج الدونم الواحد لزراعة ما اذا كانت كمية المطر على الدونم الواحد (0.5 ملم) وقيم الثابت السابق فيمكن تقديرها كالآتي :-

$$\hat{Y} = \hat{\alpha} + \hat{\beta} X$$

$$\hat{Y} = \hat{\alpha} + \hat{\beta} (0.5)$$