جامعة الموصل
كلية الزراعة والغابات
قسم الارشاد الزراعي ونقل التقنيات

الاحصاء الاجتماعي

لطلبة المرحلة الثانية

اعداد

م. رؤى محمد حامد

<u>المفردات:</u>

1. **مفهوم علم الإحصاء والإحصاء الاجتماعي وتطورهما**
2. **البيانات والرموز الإحصائية**
3. **المجتمع والعينة**
4. **عرض البيانات الإحصائية وتبويبها**
5. **مقاييس التمركز (التوسط)**

- **الوسط الحسابي(التوسط)**
- **الوسيط**
- **المنوال**

6. **مقاييس التشتت**

- **المدى**
- **الانحراف المتوسط**
- **التباين والانحراف القياسي**
- **الدرجة القياسية Z-Score**

7. **التوزيع ذي الحدين Binonial Distribution**
8. **التوزيع الطبيعي Normal Distribution**
9. **اختبار الفرضيات الاحصائية**

- **معامل ارتباط بيرسون**
- **اختبار (T- test)**
- **تحليل التباين (ذات الاتجاه الواحد) One Way Aqnalysis Of Variance**

10. **الطرق الاحصائية اللامعلمية Non-Parametric Methods**

- **الاختبارات اللامعلمية**
- **معامل ارتباط سبيرمان للرتب**
- **مربع كاي**
- **اختبار الإشارات :**

• **لعينة واحدة**

- **اختبار الوسيط**

المحاضرة الأولى

نشوء وتطور علم الإحصاء

علم الإحصاء: علم قديم كقدم المجتمع البشري حيث ارتبط منذ نشأته بعمليات العد التي كانت تجريها الدولة في العصور الوسطى لحساب أعداد جيوشها والضرائب التي تجبى من المزارعين وجمع المعلومات عن الأراضي التي تسيطر عليها الدولة وغيرها.

أصل كلمة الإحصاء statistics مشتقة إما

١- من اللاتينية status

٢- من الايطالية statista

٣- من الألمانية statistik

وكلها اشتقت من كلمة دولة state

تطور علم الإحصاء وتطبيقاته عبر سنوات طويلة ، وتم ذلك بجهود كثيرة من العلماء من دول مختلفة وكان التطور بطيئا إلى أن جاء القرن العشرين ليشهد معدلا هائلا للتطور فى النظريات الإحصائية فى مجالات كثيرة .

ويرجع الاهتمام بالإحصاء إلى عصور قديمه ، وان تعداد السكان عند القدماء المصرين وفي الصين أمثلة توضح اهتمام الحكومات منذ القدم بالمعلومات الاجتماعية وذلك لأغراض التنظيم والتخطيط فى أحوال السلم والحرب .

ويبدو أن كلمه إحصاء (statistics) قد ظهرت لأول مره عام١٧٤٩ وهى مشتقه من الكلمة اللاتينية (status) أو الايطالية (statista) وتعني كلاهما الدولة السياسية . ومن الطبيعي أن تكون الدولة أول من اهتم بجمع البيانات وذلك لإدارة شؤون البلاد خاصة عن السكان لأغراض حربية وضريبية ، وامتدت بعد ذلك لتشمل إحصاءات حجم السكان والمواليد والوفيات والإنتاج والاستهلاك والثروة ٠٠٠٠٠الخ . وهكذا بداء العلم وتطوره باعتباره علم الدولة أو علم الملوك .

وكلمة إحصاء (Statistics) لها ثلاث معانى :

(١) الإحصاءات أو البيانات : مثال ذلك إحصاءات السكان والمواليد والوفيات والإنتاج – الصادرات – الاستهلاك .

(٢) المؤشرات المحسوبة من عينة (العينة هى مجموعة جزئية من الوحدات محل الدراسة)

(٣) علم الإحصاء : وهو فرع من فروع الرياضيات يشمل النظريات والطرق الموجهة نحو جميع البيانات ووصف البيانات والاستقراء وصنع القرارات .

تعريف علم الإحصاء:

١. عرفه bodding ton بأنه علم التقديرات والاحتمالات

٢. عرفَه Lovitt بأنه العلم الذي يختص بجمع وتصنيف وتبويب الحقائق العددية كأساس لتفسير ووصف ومقارنة الظواهر

٣.عرفَه cowden croxton بأنه العلم الذي يختص بجمع وتحليل وتفسير البيانات العددية

٤. أما **المشهداني** فعرفه بأنه الطريقة العلمية التي تختص بجمع البيانات والحقائق بالشكل الذي يسهل عملية تحليلها وتفسيرها ومن ثم استخلاص النتائج واتخاذ القرار على ضوء ذلك

علاقة علم الإحصاء بالعلوم الاجتماعية

تأثرت العلوم الاجتماعية وخاصة علم الاجتماع وعلم النفس وعلم السياسة بالتطورات التي حققها علم الإحصاء ، واستعان العلماء الاجتماعيون بمنهج جديد فى دراساتهم . وهو المنهج الاحصائى الذى ينطوي علي نفس خطوات المنهج العلمي في البحث ، حيث يقدم علي عمليتين منطقيتين هما القياس و الاستنتاج .

ان حاجة البحوث الإنسانية أشد ما تكون إلى تطبيق هذه الوسائل . لذلك كانت البحوث النفسية والتربوية والاجتماعية من أصعب البحوث ، وتحتاج إلى حرص زائد ومهارة فائقة من الباحث .

ويمكن تلخيص أسباب ذلك فيما يلي :-

أ) السلوك البشرى فى تغير دائم، ومدى تغيره من فترة لأخرى أوسع مما نظن ، لدرجة تجعل من الصعوبة بمكان إعطاء تنبؤات علمية دقيقة عنه.

ب) السلوك البشرى كثيرا ما يخدع دارسة ، ذلك لان حقيقته قد تختلف كثيرا عما يبدوا علية ، فهو يحتاج إلى ضبط في البحث ودرجة كبيرة من الدقة الإحصائية .

ج) السلوك البشرى معقد تعقيدا كبيرا وتتدخل فيه عوامل قد تزيد أو تختلف عما يتوقعه الباحث .

د) البحوث الإنسانية يقوم بها إنسان . ذلك مما يسمح بتدخل العوامل الشخصية كثيرا في نواحي القياس والوصف بدرجة قد تكون كبيرة أو صغيرة حسب الطرق التى يستخدمها الباحث . وطرق الضبط الاحصائى خير وسيلة تعين الباحث على استبعاد هذه العوامل الشخصية ..

البيانات Data :

من الشائع فى مجال البحوث الاجتماعية توافر مجموعة من البيانات الإحصائية التى يحصل عليها الباحث باستخدام أدوات جمع بيانات مناسبة وعادة تتمثل تلك البيانات فى شكل أرقام تعتبر قياسا للمتغيرات تحت الدراسة ولما كانت تلك الأرقام تفتقر إلى الترتيب والتصنيف يطلق عليها البيانات الأولية أو البيانات الخام Raw Data.

وتعرف البيانات الإحصائية أنها كمية من المعلومات على هيئة أرقام وان تلك الأرقام إما أن تكون صحيحة Integers مثل ١٠ ، ٢٠ ، ٣٠ وهكذا أو تكون أرقاما عشرية أو حقيقية Real Numbers مثل ٨,٥ ، ١٠,٢٥ ، ١٥٠٥ وهكذا : ويتوقف حجم البيانات الخام على حجم المجتمع الأصلى فكلما ازداد حجم هذا المجتمع يتوقع مزيدا من الأرقام غير المرئية والتى يصعب مع كثرتها وعدم تصنيفها تفهم أو قياس متغير أو أكثر تحت الدراسة ومن ثم كان من الضرورى أن يقوم الباحث بتصنيف وتبويب تلك البيانات بالشكل أو بالأسلوب الذى

يخدم جيدا هدف الباحث من دراسة المتغيرات أو استنباط نوعية العلاقات أو المعلومات الهامة التى تتعلق بتلك المتغيرات .

ويقصد بتعبير البيانات " أى كمية من المعلومات فى صورة رقمية والصورة الرقمية للبيانات تبدو إما على شكل أرقام صحيحة مثل ١٠ ، ١١٢ ، ٤٦٤ . أو على شكل أرقام حقيقية مثل ٢٠,٤ ، ٦١,٨ ، ١٨٢,١ أى أنها الأرقام التى تحتوى على علامة عشرية . وتعتبر المعلومات الرقمية (البيانات) المادة الخام لأسلوب العمل الاحصائى كما أنها تلعب دورا كبيرا فى تطبيق الأساليب الإحصائية .

وتسمى البيانات المتاحة – المنشورة أو التى تم جمعها – تسمى بيانات خام أو أولية – ذلك أنها تكون غير مجهزة فهى لا تفصح إلا عن القليل من المعلومات . كما أنه يستحيل استخلاص المعلومات منها . وفى سبيل ذلك نستعين بأساليب ومقاييس وصف البيانات . وهذه الأساليب كثيرة ومتنوعة فهى تختلف حسب عوامل أهمها عدد المتغيرات ومستوى قياسها

ولعل ابسط الطرق الإحصائية لتنظيم وتلخيص البيانات طريقة التوزيع التكرارى Frequency Distribution، أو بمعنى ضمنى من التوزيع التكرارى يمكن استخدام وسيلة أو أكثر من الوسائل الثلاث التالية

والتى يمكن أن يتحول التوزيع إليها أو إلى أى منها .

أ) استخدام الجداول الإحصائية Statistical Tables فى عملية تصنيف وتبويب البيانات الخام .

ب) استخدام التمثيل البياني والخرائط فى عرض البيانات الإحصائية (تحويل التوزيع التكرارى إلى منحنيات تكرارية).

ج) استخدام مقياس أو أكثر من المقاييس الإحصائية مثل المتوسط الخام Mean الانحراف المعياري Standard Deviation ومعامل الارتباط Correlation Coefficient فى تلخيص البيانات الإحصائية فى صورة رقم أو نسبة مئوية ونرى أهمية الوقوف على نوعية البيانات الإحصائية من منظور مستويات القياس الاحصائى نظرا لأهمية تلك البيانات الإحصائية وفقا لمستويات القياس الاحصائى يرجع إلى أن المتغيرات التى تقاس كميا تنقسم من قيمتها العددية إلى المتغير المتصل والمتغير المتقطع .

المحاضرة الثانية

طبيعة البيانات الاحصائية: The nature of statistical data

عندما يريد الباحث دراسة ظاهرة معينة فأن أول عمل يقوم به هو جمع البيانات Collecting data حول تلك الظاهرة ومن ثم إيجاد طريقه لمعالجة تلك البيانات **والإحصاء** :هو علم يختص بطرق جمع وتنظيم وعرض وتحليل وتلخيص البيانات وذلك من اجل الوصول إلى نتائج مقبولة وقرارات سليمة، فعند جمع البيانات عن الظاهرة التي يريد الباحث دراستها فأنها تأخذ مفهوم **المتغير Variable** والذي يعرف بانه: هو أي ظاهرة تُظهر إختلافات بين مفرداتها ويرمز له بـ Xi أو Yi أو أي رمز والمتغيرات Variables تكون على نوعين هما:

أ- **المتغيرات الوصفية Qualitative Variables** مثل الصفات التي لا يمكن اخذ قراءة مباشرة عنها مثل اللون – السلوك – الكفاءة – الحالة الاجتماعية (غني ومتوسط الحال وفقير) أو ذكر وأنثى – المرحلة الدراسية (الأولى ، الثانية ، الثالثة والرابعة) وغيرها.

ب- **المتغيرات الكمية Quantitative Variables** وهي التي يمكن قياسها مباشرةً مثل الطول – الوزن – السرعة – الضغط – درجة الحرارة – المساحة وغيرها من القياسات الكمية الأخرى، وكذلك تقسم المتغيرات الكمية إلى قسمين وهما:

١- **المتغيرات الكمية المستمرة Continuous Quantitative Variables:** وتكون قيم مفرداتها ذات رقم صحيح أو كسر عشري مثل: الأوزان – الأطوال – المساحات – السرعة – درجة الحرارة – الضغط.....الخ فعلى سبيل المثال أن أطوال نباتات القطن تكون مثلا ٧٠ أو ٦٩,٤ أو ٦٥,٨ أو ٦٠ أو ٥٥,٩ سم أوأي رقم آخر.

٢- **المتغيرات الكمية المتقطعة Discrete Quantitative Variables:** فتكون متمثلةً بأعداد صحيحة فقط غير كسرية مثل عدد أفراد الأسرة – عدد السيارات المصنّعة لشركة معينة – عدد الثمار لمحصول معين – عدد أفراد فريق لنشاط رياضي معين .... الخ.

مصطلحات مهمّة:

الفئات Classes: تقسيم قيم متغير إلى مجاميع تأخذ مدى معين من قيم المتغير. فكل مجموعة بمدى معين تسمى بالفئة وعادة" يكون مدى الفئات متساوي فيما بينها.

حدود الفئات Class limits: لكل فئة هنالك حدين هما حدي الفئة الأدنى Lower Class limit والأعلى Upper Class limit، وغالباً ما تكون مكتوبة بصيغة الفئات الاعتيادية والتي قد يضطر الباحث إلى تحويلها إلى فئات حقيقية Class Boundaries or True Class limits.

طول الفئة Class length or Class width: يستحسن أن تكون أطوال الفئات متساوية وإذا كانت كذلك فان طول الفئة يحسب بحساب الفرق بين حدي أدنى أو حدي أعلى لأي فئتين متتاليتين، أو الفرق بين الحد الأعلى ناقصا حدها الأدنى إذا كانت من نوع الفئات الحقيقية.

مركز الفئة Class mark or Class Mid-point: وهو منتصف المدى بين حدي أي فئة يطلق علية بمركز الفئة ويرمز له بالرمز Xi or Yi ويحسب عادةً من جمع قيمة الحد الأدنى مع الحد الأعلى لنفس الفئة ثم قسمة المجموع على اثنين. ويمكن الاطلاع على الفئات ومراكز الفئات الموجودة في جدول (٤).

تكرار الفئة Class frequency: هو عدد القيم التي تقع ضمن حدود أي فئة ويرمز له بالرمز fi والمجموع التكراري يجب أن يكون مساويا لعدد قراءات أي ظاهره تدرس. والجدول رقم (٤) يبين التكرار بصيغة الحزمة والصيغة الثانية كعدد مكتوب

–Statistical symbols - بعض الرموز الإحصائية

N Factorial	n!	المجموع /"Sum of" or summation	$\sum$
مستوى المعنوية/ 0.05 statistical level	α	عدم التساوي /Not Equal	$\neq$
مستوى المعنوية0.01 statistical level	β	اكبر من/Greater than	$>$
القيمة المطلقة/Absolute value	$\lvert \quad \rvert$	اصغر من او اقل من /Less than	$<$
المتوسط البسيط/Sample mean	$\bar{x}$	متوسط المجتمع /Population mean	μ
مشاهدة/Individual observation	Xi	تباين/Variance	σ^2
		and standard deviation of the population وانحراف معياري للمجتمع	σ
المتغيرات/A variable	X or Y or Z	التباين Variance	S^2
المجموع الكليTotal sum of servation للمشاهدات	$\sum Xi$	الانحراف standard deviation of the sample المعياري	S
مربع المجموع الكلي Square total sum	$(\sum xi)^2$	حجم العينة/Sample size	n
Coefficient of variation معامل الاختلاف	C.V.	مجموع المربعات الكلي/Total sum of square	$\sum Xi^2$
معامل Sample regression coefficient الانحدار البسيط	*b*	الاختلاف/Difference (x1 –x2)	*d*
Degrees of freedom درجات الحرية	df	الارتباط البسيط/Sample correlation	R
متوسط المربعات Mean square	MS	مجموع المربعات/Sum of square	SS
النجمةSignificant and highly significant معنوي عند ٠,٠٥ ونجمتين معنوي عند ٠,٠١	(*) and (**)	غير معنوي/Not significant	ns
الفرضية البديلة/Alternative hypothesis	H1	فرضية العدم /Null hypothesis	H_o

The Population and the Sample : المجتمع والعينة

يمكن تعريف **المجتمع – Population** – : هو جميع المفردات أو القراءات التي يأخذها المتغير. ويكون مقسّم الى قسمين :

أ- **المجتمع المحدود Finite population** : هو الذي يمكن حصر عدد مفرداته مثل مجمع الجامعات العراقية – مجتمع طلبة جامعة معينة كأن تكون جامعة الموصل – عدد النباتات في متر مربع من الحقل – أعداد الأوراق في نبات زهرة الشمس.... الخ.

ب- **المجتمع اللامحدود Infinite population** هو الذي لا يمكن حصره بسبب كثرة مفرداته أو انه قابل للتغير بسرعة مثل حساب أعداد البكتريا في طبق بيتري الذي يحتوي على عدد أكثر من عشرة ملايين خلية وأثناء العد سيتغير العدد أو أن مفرداته متحركة صعوبة حصرها وعدها مثل مجتمع الأسماك في بحيرة أو نهر وكذلك عدد الحشرات في أحد حقول الشعير في شمال العراق خلال موسم نمو المحصول.... الخ.

المحاضرة الثالثة

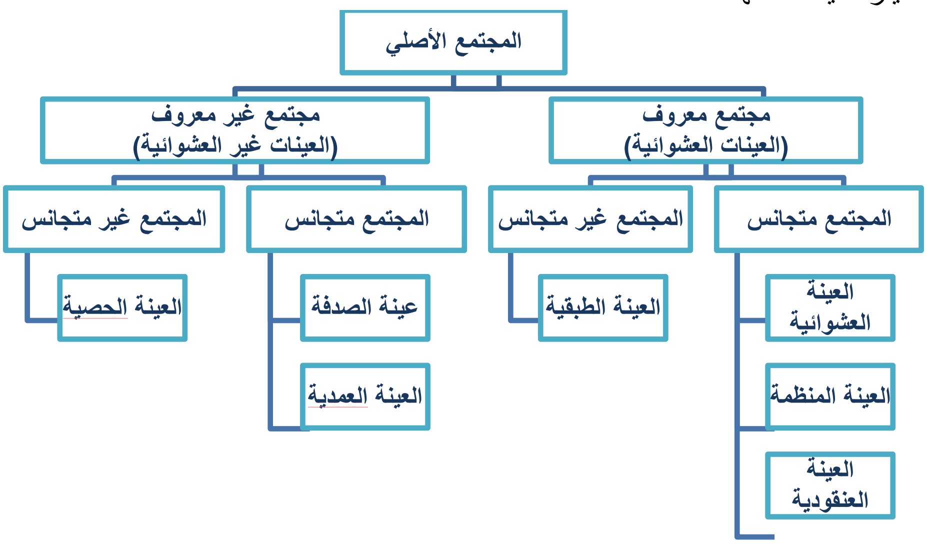
العينة

هى جزء أو شريحة من المجتمع تتضمن خصائص المجتمع الأصلي الذي نرغب فى التعرف على خصائصه ويجب أن تكون تلك العينة ممثلة لجميع مفردات هذا المجتمع تمثيلا صحيحا

والعينة هى جزء من المجتمع ونقوم بدراستها للتعرف على خصائص المجتمع التى سحبت منه هذه العينة – ولكى تصلح النتائج التى نحصل عليها للتعبير عن المتجمع لا بد وان تكون العينة ممثلة للمجتمع (أي جميع المفردات المراد بحثها) تمثيلا صحيحا .

أسلوب اختيار العينة

هناك أساليب مختلفة لاختيار العينات ولكن نوع العينة وإجراءات سحبها من المجتمع الإحصائي تختلف من موقف لآخر والاعتبار الجوهري الذى يراعيه الباحث هو الحصول على عينة مناسبة . ويوجد عدة أساليب يعتمد عليها الباحث لاختيار العينات منها



أ ـ العينات الاحتمالية:

العينة العشوائية	جميع عناصر المجتمع لها نفس الفرصة في الظهور في العينة
العينة الطبقية	يقسم المجتمع إلى طبقتين على الأقل ثم نختار العينة من كل منهما
العينة المنتظمة	نختار نقطة بداية من المجتمع ثم نختار العنصر الموجود على بعد ثابت من هذه النقطة
العينة العنقودية	يقسم المجتمع إلى مساحات أو أجزاء ثم نختار عشوائيا بعض هذه المساحات، ثم نختار جميع عناصرها بالعينة.

ب ـ العينات غير الاحتمالية:

عينة الصدفة	يتم اختيارها عن طريق الصدفة
العينة العمدية (القصدية)	يتم اختيار أفراد العينة تحت شروط معينة لتحقيق الهدف من التجربة
العينة الحصية	يقسم المجتمع إلى أجزاء ثم نختار العينة من كل جزء من أجزاء المجتمع وفقا للنسب المحددة

أخطاء البيانات الإحصائية:

تتعرض البيانات الإحصائية التي يتم جمعها إلى نوعين من الأخطاء:

1. خطأ التميز أو التحيز: وهو ذلك الخطأ الناتج عن مصادر متعددة، منها أخطاء في تصميم البحث أو التجربة أو أخطاء فنيه أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المتجمعة.. أخطاء التميز تزداد بازدياد الفروق بين الإمكانيات (المادية والفنية) اللازم توافرها لضمان أقصى درجة دقة ممكنه وبين الإمكانيات الفعلية المتاحة للباحث.

اذا سحبنا عدة عينات من مجتمع ما وحسبنا المتوسط الحسابي لكل عينة من هذه العينات ثم حسبنا المتوسط الحسابي لهذه المتوسطات فهذا المتوسط يجب أن يساوي المتوسط الحسابي الحقيقي للمجتمع المسحوبة منه هذه العينات، وفي حال وجود فرق بين المتوسطين فإن هذا الفرق يسمى بخطأ التميز أو التحيز

أسباب خطأ التميز أوالتحيز:

- الاختيار غير العشوائي للعينة: تعتمد بعض طرق الاختيار للعينة على خاصية معينة كالاعتماد على دليل الهاتف (عند دراسة الدخل والانفاق).
- التحيز المقصود (تعمد إدخال بعض الوحدات)
- استبدال وحدة بوحدة أخرى غير مدرجة ضمن الإطار العام للدراسة

كيفية التقليل من أخطاء التميز أوالتحيز:

- اختيار جميع وحدات العينة عشوائياً باستخدام إحدى طرق الاختيار العشوائي
- عدم استبدال أية وحدة تم اختيارها بوحدة أخرى
- تدريب الباحثين بشكل جيد على جمع البيانات والتقيد بالتعليمات

٢. خطأ المعاينة العشوائية أو خطأ الصدفة: وهو الخطأ الناتج عن فروق الصدفة بين مفردات المجتمع التي دخلت العينة وبين تلك المفردات التي لم تشأ الصدفة أن تدخل العينة .

عند اختيار العينة العشوائية هناك خطأ ينتج عن الاختلاف أو التشتت Variation بين قيم الوحدات التي تتكون منها العينة وتلك الوحدات التي لم تشأ الصدفة أن تدخلها في العينة وهذا الخطأ يسمى بخطأ المعاينة العشوائي .

كيف نقلل من خطأ المعاينة العشوائي:

- زيادة حجم العينة
- طريقة الاختيار المناسب التي تقلل من اختلاف قيم الوحدات الإحصائية (كالأسلوب الطبقي أو العينة المنتظمة...الخ).

عرض وتبويب البيانات الاحصائية :

تبويب البيانات :

يقصد بتبويب البيانات عرض هذه البيانات (البيانات الخام) فى جداول مناسبة وذلك حتى يمكن تلخيصها وفهمها واستيعابها واستنتاج النتائج منها ومقارنتها بغيرها من البيانات ، كما يسهل الرجوع إليها فى صورة جداول دون الاطلاع على الاستمارات الأصلية التى قد تحمل أسماء أصحابها مما يخل بمبدأ سرية البيانات الإحصائية .
كما يعتبر عرض وتبويب البيانات الإحصائية الخطوة الثانية (بعد تجميع هذه البيانات الخام) في مفهوم التحليل الإحصائي، ويلجأ الباحث إلى حصر وتصنيف هذه البيانات وعرضها بطريقة مختصرة تساعد على فهمها وتحليلها إحصائياً للتعرف عليها ووصفها ومقارنتها بغيرها من الظواهر ، والخروج ببعض المدلولات الإحصائية عن مجتمع الدراسة

يمكن عرض البيانات في صورة جدول تكراري، ويختلف شكل الجدول طبقا لنوع البيانات، وحسب عدد المتغيرات، وفيما يلي عرض بيانات متغير (وصفي أو كمي) في شكل جدول تكراري بسيط.

عرض بيانات المتغير الوصفي في شكل جدول تكراري بسيط

إذا كنا بصدد دراسة ظاهرة ما تحتوي على متغير وصفي واحد، فإنه يمكن عرض بياناته في شكل جدول تكراري بسيط، وهو جدول يتكون من عمودين، أحدهما به مستويات (مجموعات) المتغير، والثاني به عدد المفردات (التكرارات) لكل مستوى (مجموعة).
والمثال التالي يبين لنا كيف يمكن تبويب البيانات الوصفية الخام في شكل جدول تكراري.

مثال

فيما يلي بيانات عينة من **40** مزرعة عن نوع التمر الذي تنتجه المزرعة.

سكري	خلاص	برحي	خلاص	برحي	خلاص	صقعي	خلاص
برحي	سكري	برحي	صقعي	خلاص	برحي	نبوت سيف	برحي
صقعي	برحي	سكري	خلاص	برحي	برحي	صقعي	خلاص
برحي	خلاص	برحي	سكري	نبوت سيف	صقعي	نبوت سيف	صقعي
خلاص	برحي	صقعي	نبوت سيف	سكري	برحي	صقعي	خلاص

والمطلوب:
١- ما هو نوع المتغير؟، وما هو المعيار المستخدم في قياس البيانات؟.
٢- اعرض البيانات في شكل جدول تكراري.
٣- كون التوزيع التكراري النسبي.
٤- علق على النتائج.

<u>الحل</u>

١- نوع التمر (سكري – خلاص – برحي – صقعي – نبوت سيف) متغير وصفي، تقاس بياناته بمعيار اسمي.
٢- لعرض البيانات في شكل جدول تكراري ، يتم إتباع الآتي:
- تكوين جدول تفريغ البيانات:

وهو جدول يحتوي على علامات إحصائية، كل علامة تعبر عن تكرار للمجموعة التي ينتمي إليها نوع التمر الذي تنتجه المزرعة، وكل خمس علامات تكون حزمة إحصائية، كما هو مبين بالجدول التالي:

جدول تفريغ البيانات

نوع التمر	العلامات الإحصائية	عدد المزارع (التكرارات)
سكري	~~||||~~	5
خلاص	~~||||~~ ~~||||~~	10
برحي	~~||||~~ ~~||||~~ |||	13
صقعي	~~||||~~ |||	8
نبوت سيف	||||	4
Sum		**40**

- تكوين الجدول التكراري.

وهو نفس الجدول السابق، باستثناء العود الثاني، ويأخذ الصورة التالية:

جدول يوضح التوزيع التكراري لعينة حجمها 40 مزرعة حسب نوع التمر الذي تنتجه

نوع التمر	عدد المزارع (التكرارات) (f)	التوزيع التكراري النسبي
سكري	**5**	$\left(\frac{5}{40}\right) = 0.125$
خلاص	**10**	$\left(\frac{10}{40}\right) = 0.25$
برحي	**13**	$\left(\frac{13}{40}\right) = 0.325$
صقعي	**8**	$\left(\frac{8}{40}\right) = 0.20$
نبوت سيف	**4**	$\left(\frac{4}{40}\right) = 0.10$
Sum	**40**	**1.00**

المصدر: بيانات افتراضية.

٣- التوزيع التكراري النسبي:
يحسب التكرار النسبي بقسمة تكرار المجموعة على مجموع التكرارات، أي أن:

$$\text{التكرار النسبي} = \frac{\text{تكرار المجموعة}}{\text{مجموع التكرارات}\ (n)} = \left(\frac{f}{\sum f}\right)$$

والعمود الثالث في الجدول اعلاه يعرض التكرار النسبي للمزارعين حسب نوع التمر.

٤- التعليق: من الجدول اعلاه يلاحظ أن نسبة المزارع التي تنتج النوع "برحي" في العينة هي **32.5%** وهي أكبر نسبة مما يدل على أن النمط الشائع في إنتاج التمور هو ذلك النوع، بينما نجد أن نسبة المزارع التي تنتج النوع "نبوت سيف" حوالي **10.0%** وهي أقل نسبة.

عرض بيانات المتغير الكمي في شكل جدول تكراري بسيط

بنفس الأسلوب السابق المتبع في تكوين جدول تكراري، يمكن أيضا عرض بيانات المتغير الكمي في شكل جدول تكراري بسيط، ويتكون هذا الجدول من عمودين، الأول يحتوي على فئات تصاعدية للقراءات التي يأخذها المتغير، والثاني يشمل التكرارات أو عدد المفردات التي تنتمي قراءاتها للفئة المناسبة لها، والمثال التالي يبين كيف يمكن عرض البيانات الكمية بيانيا.

مثال

فيما يلي بيانات درجات **70** طالب في الاختبار النهائي لمقرر مادة الإحصاء التطبيقي.

56	**65**	**70**	**65**	**55**	**60**	**66**	**70**	**75**	**56**
60	**70**	**61**	**67**	**61**	**71**	**67**	**62**	**71**	**66**
68	**72**	**57**	**68**	**72**	**69**	**57**	**71**	**69**	**75**
72	**62**	**67**	**73**	**58**	**63**	**66**	**73**	**63**	**65**
58	**73**	**74**	**76**	**74**	**80**	**81**	**60**	**74**	**58**
76	**82**	**77**	**83**	**77**	**85**	**91**	**78**	**94**	**72**
79	**64**	**57**	**79**	**55**	**87**	**64**	**88**	**78**	**62**

والمطلوب:

١- كون التوزيع التكراري لدرجات الطلاب.
٢- كون التوزيع التكراري النسبي.
٣- ما هو نسبة الطلاب الحاصلين على درجة ما بين **70**إلى أقل من **80**؟
٤- ما هو نسبة الطلاب الحاصلين على درجة أقل من **70** درجة؟
٥- ما هو نسبة الطلاب الحاصلين على درجة **80** أو أكثر ؟

الحل

١- تكوين التوزيع التكراري:
درجة الطالب في الاختبار متغير كمي مستمر، ولكي يتم تبويب البيانات في شكل جدول تكراري، يتم اتباع الآتي:

- حساب المدى **Range(R)**

Range = Maximum – Minimum
R = 94 - 55 = 39

- تحديد عدد الفئات **Classes(C)**:

تتحدد عدد الفئات وفقا لاعتبارات منها: رأي الباحث، والهدف من البحث، وحجم البيانات، ويرى كثيرا من الباحثين أن أفضل عدد للفئات يجب أن يتراوح بين **5** إلى **15** ، بفرض أن عدد الفئات هو **8** فئات، أي أن: **(C=8)**.

- حساب طول الفئة **Length(L)**:

$$L = \frac{Range}{Classes} = \frac{R}{C} = \frac{39}{8} = 4.875 \approx 5$$

- تحديد الفئات:

الفئة تبدأ بقيمة تسمي الحد الأدنى، وتنتهي بقيمة تسمي الحد الأعلى، ومن ثم نجد أن :
- الحد الأدنى للفئة الأولى هو أقل قراءة (درجة) أي أن الحد الأدنى للفئة الأولى = **55**
الحد الأعلى للفئة الأولى = الحد الأدنى + طول الفئة = **55 + L = 55+5=60**
إذا الفئة الأولى هي: **"55 to les than 60"** وتقرأ " من **55** إلى أقل من **60** "
_ الحد الأدنى للفئة الثانية = الحد الأعلى للفئة الأولى = **60**
الحد الأعلى للفئة الثانية = الحد الأدنى للفئة + طول الفئة = **65 = 60 + 5**
إذا الفئة الثانية هي: **"60 to less than 65"** وتقرأ "من **60** إلى أقل من **65**"
- وبنفس الطريقة يتم تكوين حدود الفئات الأخرى، وهي:

الفئة الثالثة : **65 to less than 70** الفئة الرابعة : **70 to less than 75**
الفئة الخامسة: **75 to less than 80** الفئة السادسة: **80 to less than 85**
الفئة السابعة: **85 to less than 90** الفئة الثامنة: **90 to less than 95**

ويمكن كتابة الفئات بأشكال مختلفة كما هو مبين بجدول تفريغ البيانات:

- تكوين جدول تفريغ البيانات:

جدول تفريغ البيانات

الدرجة			**العلامات الإحصائية**	**عدد الطلاب (التكرارات)**
فئات	**فئات**	**فئات**		
55 to less than 60	55 – 60	55-	~~////~~ ~~////~~	10
60 to less than 65	60 – 65	60-	~~////~~ ~~////~~ //	12
65 to less than 70	65 – 70	65-	~~////~~ ~~////~~ ///	13
70 to less than 75	70 – 75	70-	~~////~~ ~~////~~ ~~////~~ /	16
75 to less than 80	75 – 80	75-	~~////~~ ~~////~~	10
80 to less than 85	80 – 85	80-	////	4
85 to less than 90	85 – 90	85-	///	3
90 to less than 95	90 - 95	90-95	//	2
Sum				70

- تكوين الجدول التكراري:

جدول التوزيع التكراري لعدد **70** طالب حسب درجاتهم في اختبار مقرر الإحصاء

فئات الدرجة	**عدد الطلاب (التكرارات)** (f)	**التكرار النسبي**
55 – 60	10	0.143
60 – 65	12	0.171
65 – 70	13	0.186
70 – 75	16	0.229
75 – 80	10	0.143
80 – 85	4	0.057
85 – 90	3	0.043
90 – 95	2	0.028
Sum	70	1.00

المصدر: بيانات نتيجة العام ١٤٢٦هـ

٢- التوزيع التكراري النسبي:

$$\text{التكرار النسبي} = \frac{f}{n}$$

والعمود الثالث في الجدول رقم (٢-٣) يبين التكرار النسبي.

٣- نسبة الطلاب الحاصلين على درجات ما بين **70** إلى أقل من **80** هو مجموع التكرارين النسبيين للفئتين الرابعة والخامسة:

نسبة الطلاب الحاصلين على درجات ما بين **(70 , 80)** = $0.229 + 0.143 = 0.372$

أي حوالي **37.2%** من الطلاب حصلوا على درجات ما بين **(70 , 80)** .

٤- نسبة الطلاب الحاصلين على درجات أقل من **70**، هو مجموع التكرارات النسبية للفئات الأولى والثانية، والثالثة:

نسبة الطلاب الحاصلين على درجة أقل من **70** = $0.143 + 0.171 + 0.186 = 0.5$

أي أن حوالي **50%** من الطلاب حصلو على درجة أقل من **70** درجة

٥- نسبة الطلاب الحاصلين على درجة **80** أو أكثر، هو مجموع التكرارات النسبية للفئات الثلاث الأخيرة:

$0.057+0.043+0.028=0.128$ = نسبة الطلاب الحاصلين على درجات **80**أو أكثر

أي أن حوالي **12.8%** من الطلاب حصلوا على درجة **80** أو أكثر.

المحاضرة الرابعة

العرض البياني للبيانات الوصفية

يمكن عرض البيانات الخاصة بمتغير وصفي في شكل دائرة بيانية أو أعمدة بيانية، يمكن من خلاله وصف ومقارنة مجموعات أو مستويات هذا المتغير.

الدائرة البيانية

لعرض بيانات المتغير الوصفي في شكل دائرة، يتم توزيع الـ $\mathbf{360^o}$ درجة حسب التكرار النسبي لمجموعات المتغير، حيث تحدد مقدار الزاوية الخاصة بالمجموعة رقم **r** بتطبيق المعادلة التالية:

مقدار الزاوية $= \mathbf{360^o} \times$ التكرار النسبي للمجموعة

مثال

الجدول التكراري التالي يبين توزيع عينة حجمها **500** أسرة حسب المنطقة التي تنتمي إليها.

المنطقة	**الرياض**	**الشرقية**	**القصيم**	**الغربية**	sum
عدد الأسر	150	130	50	170	500

مثل البيانات أعلاه في شكل دائرة بيانية.

الحل:

١- تحديد مقدار الزاوية المخصصة لكل منطقة، بتطبيق المعادلة:

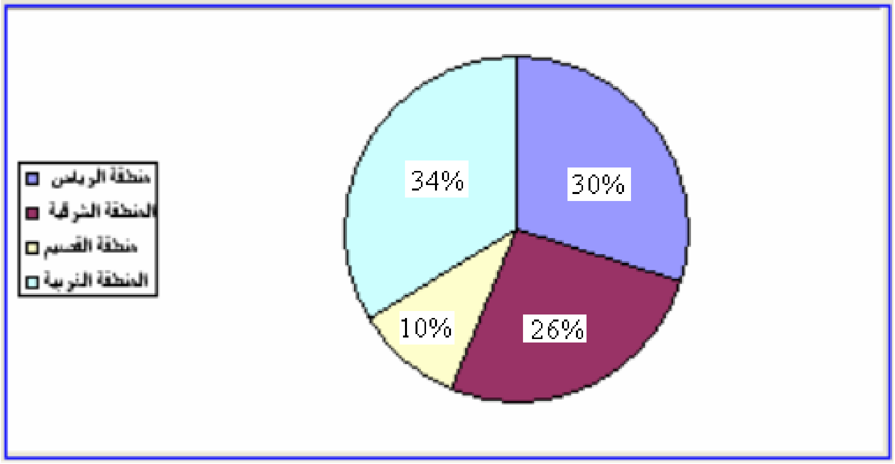
مقدار الزاوية المخصص للمنطقة $= \mathbf{360^o} \times$ التكرار النسبي للمنطقة

المنطقة	**عدد الأسر**	**التكرار النسبي**	**مقدار الزاوية**
الرياض	150	0.30	$360 \times 0.30 = 108^o$
الشرقية	130	0.26	$360 \times 0.26 = 93.6^o$
القصيم	50	0.10	$360 \times 0.10 = 36^o$
الغربية	170	0.34	$360 \times 0.30 = 122.4^o$
Sum	500	1.00	360^o

٢- رسم الدائرة

يتم رسم دائرة وتقسيمها إلى أربع أجزاء لكل منطقة جزء يتناسب مع مقدار الزاوية المخصصة له، كما هو مبين في الشكل التالي:

شكل يوضح الدائرة البيانية لعينة حجمها **500** أسرة موزعة حسب المنطقة



ومن الشكل أعلاه يلاحظ أن نسبة الأسر التي تنتمي للمنطقة الغربية حوالي **34%** وهي أكبر نسبة في العينة، بينما يكون نسبة الأسر في منطقة القصيم حوالي **10%** وهي أقل نسبة في العينة.

المحاضرة الخامسة

العرض البياني للبيانات الكمية

العرض البياني للبيانات، هو أحد طرق التي يمكن استخدامها في وصف البيانات، من حيث شكل التوزيع ومدى تمركز البيانات، وفي كثير من النواحي التطبيقية يكون العرض البياني أسهل وأسرع في وصف الظاهرة محل الدراسة، وتختلف طرق عرض البيانات بيانيا حسب نوع البيانات المبوبة في شكل جدول تكراري

المدرج التكراري Histogram

المدرج التكراري هو التمثيل البياني للجدول التكراري البسيط الخاص بالبيانات الكمية المتصلة، وهو عبارة عن أعمدة بيانية متلاصقة، حيث تمثل التكرارات على المحور الرأسي، بينما تمثل قيم المتغير (حدود الفئات) على المحور الأفقي، ويتم تمثيل كل فئة بعمود، ارتفاعه هو تكرار الفئة، وطول قاعدته هو طول الفئة.

مثال

فيما يلي التوزيع التكراري لأوزان عينة من الدواجن بالجرام، حجمها **100** اختيرت من أحد المزارع بعد **45** يوم.

الوزن	600-	620-	640-	660-	680-	700-720	Sum
عدد الدجاج	10	15	20	25	20	10	100

والمطلوب:

١- ما هو طول الفئة؟
٢- ارسم المدرج التكراري.
٣- ارسم المدرج التكراري النسبي، ثم علق على الرسم.

الحل

١- طول الفئة **(L)**

طول الفئة = الحد الأعلى للفئة – الحد الأدنى للفئة

$$L = upper - Lower$$

$$L = 620 - 600 = 640 - 620 = \ldots = 720 - 700 = 20$$

إذا طول الفئة = **20**

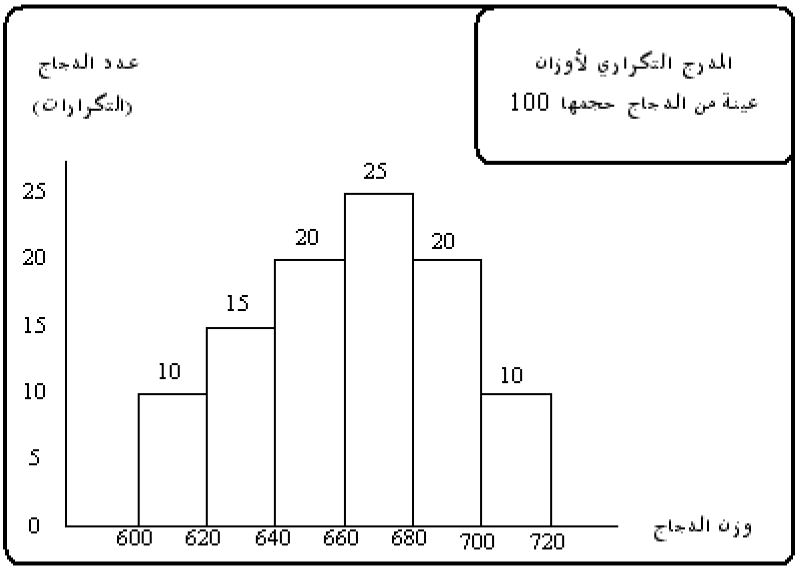
٢- رسم المدرج التكراري.

لرسم المدرج التكراري يتم إتباع الخطوات التالية:

- رسم محوران متعامدان، الرأسي ويمثل التكرارات، الأفقي ويمثل الأوزان.
- كل فئة تمثل بعمود ارتفاعه هو تكرار الفئة، وطول قاعدته هو طول الفئة.
- كل عمود يبدأ من حيث انتهى به عمود الفئة السابقة.

والشكل التالي يبين المدرج التكراري لأوزان الدجاج.

شكل يوضح المدرج التكراري لأوزان عينة من الدجاج حجمها **100** دجاجة

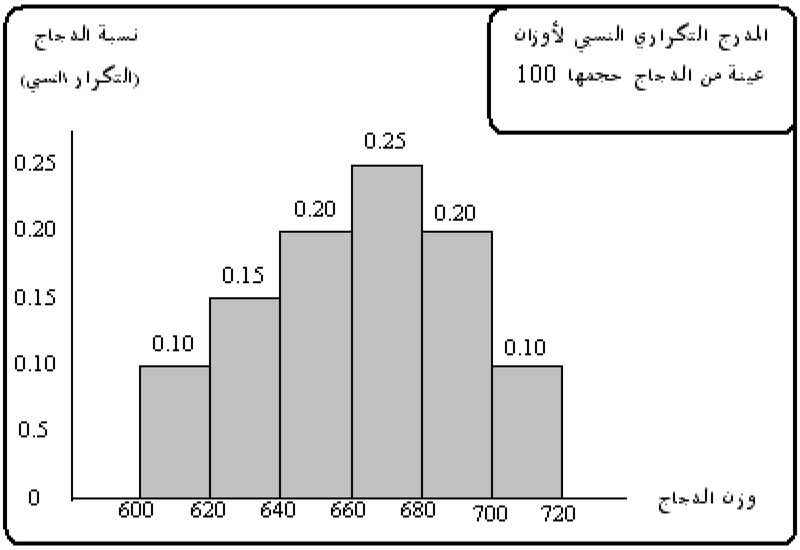


٣- رسم المدرج التكراري النسبي: لرسم المدرج التكراري النسبي يتم إجراء الآتي:
- حساب التكرارات النسبية.

الوزن	600-	620-	640-	660-	680-	700-720	Sum
عدد الدجاج	10	15	20	25	20	10	100
التكرار النسبي	0.10	0.15	0.20	0.25	0.20	0.10	1.00

- بإتباع نفس الخطوات السابقة عند رسم المدرج التكراري، يتم رسم المدرج التكراري النسبي، بإحلال التكرارات النسبية محل التكرارات المطلقة على المحور الرأسي، كما هو مبين في الشكل التالي:

شكل يوضح المدرج التكراري النسبي لأوزان عينة من الدجاج حجمها **100** دجاجة

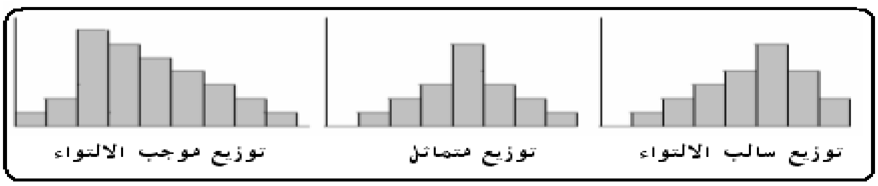


ومن الشكل أعلاه يلاحظ الآتي:
- أن **25%** من الدجاج يتراوح وزنه بين **660** ، **680** جرام وهي أكبر نسبة.
- أن الشكل ملتوي جهة اليسار، مما يدل على أن توزيع أوزان الدجاج سالب الالتواء.

ملاحظات على شكل المدرج التكراري

أ- أن المساحة أسفل المدرج التكراري تساوي مجموع التكرارات **(n)**.
ب- أما المساحة أسفل المدرج التكراري النسبي، فهي تعبر عن مجموع التكرارات النسبية، وهي تساوي الواحد الصحيح.
ت- يمكن تقدير القيم الشائعة، وهي القيم التي يناظرها أكبر ارتفاع، ففي الشكلين السابقين، نجد أن الوزن الشائع يقع في الفئة **(660-680)** ويطلق عليه المنوال.
ث- يمكن معرفة شكل توزيع البيانات، كما هو مبين بالأشكال الثلاث التالية:

شكل



<u>تغيير صيغ جدول التوزيع التكراري:</u>

يمكن تحويل جدول التوزيع التكراري الاعتيادي الى جدول تكراري تصاعدي باستخدام عبارة "أقل من" بدلاً من الحد الأدنى للفئة أي بمعني سيطلق على الفئة الأولى في المثال التالي أقل من ١٢٨ وتكرارها = ٤ ثم اقل من ١٣٨ وتكرارها ٤+٧=١١ وهكذا (أي بمعنى سيكون الحذف للحدود الدنيا للفئات) كما يبينها الجدول التالي.

جدول يوضح : يبين طريقة تحويل الجدول الاعتيادي إلى جدول تكراري صاعد أو تصاعدي.

التكرار	الفئات التصاعدية
٤	أقل من ١٢٨
١١	أقل من ١٣٨
٢٤	أقل من ١٤٨
٣٣	أقل من ١٥٨
٣٨	أقل من ١٦٨
٤٠	أقل من ١٧٨

وهناك جدول آخر يسمّى بجدول التوزيع التكراري التنازلي إذ تستخدم العبارة "أكثر من" بدلاً من الحد الأعلى للفئة ففي المثال (٢-١) سيطلق على الفئة الأولى أكثر من ١١٩ وتكرارها يساوي ٤٠، وأكثر من ١٢٩ ويساوي ٤٠-٤=٣٦ (أي بمعنى الحذف سيكون للحدود العليا للفئات)، والجدول التالي يبين ذلك.

جدول يوضح يبين طريقة تحويل الجدول الاعتيادي إلى جدول تكراري نازل أو تنازلي.

التكرار	الفئات التنازلية
٤٠	أكثر من ١١٩
٣٦	أكثر من ١٢٩
٢٩	أكثر من ١٣٩
١٦	أكثر من ١٤٩
٧	أكثر من ١٥٩
٢	أكثر من ١٦٩

كما يمكن تشكيل جدول التوزيع التكراري النسبي من الجدول الاعتيادي بكتابة نسبة ما يشكله تكرار الفئة الى المجموع التكراري فالتكرار النسبي للفئة الأولى ناتج من قسمة التكرار لتلك الفئة على المجموع مضروباً ×١٠٠ (أي أن التكرار النسبي للفئة الأولى = (٤ /٤٠) × ١٠٠) وهكذا بالنسبة لبقية التكرارات. كما مبين في الجدول التالي.

جدول يبين طريقة تحويل تكرار الفئات الى تكرار نسبي.

التكرار النسبي (%)	فئات التكرار النسبي
١٠	١١٩ ــ ١٢٨
١٧,٥	١٢٩ ــ ١٣٨
٣٢,٥	١٣٩ ــ ١٤٨
٢٢,٥	١٤٩ ــ ١٥٨
١٢,٥	١٥٩ ــ ١٦٨
٥	١٦٩ ــ ١٧٨

المحاضرة السادسة

مقاييس النزعة المركزية: Measures of Central Tendency

وهي المقاييس التي تبحث عن قيمة تتمركز حولها البيانات، وهي مقاييس متعددة أهمها **الوسط الحسابي** أو **المتوسط** The Arithmetic Mean ومن المقاييس الأخرى **الوسيط** و**المنوال** و**الوسط الهندسي** و**الوسط التوافقي** و**الوسط ألتربيعي. لكن سنتطرق للاتي:**

الوسط الحسابي أو (المتوسط): The Arithmetic Mean

وهو من أهم الأوساط وأكثرها استخداماً والذي يرمز له بالرمز $\bar{X}$ أو $\bar{Y}$.والذي يمثل مجموع القيم مقسوماً على عددها، أي أن

$$\bar{X} = \frac{\sum Xi}{n} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

مثال : قام أحد الباحثين بقياس أطوال تسعة نباتات من القطن (n = 9) وكانت نتائج القياس بالأمتار وكما يلي :

١,٢٣	١,٤	١,٣٦	١,٤٧	١,٥	١,٢٩	١,٣٢	١,٣٣	١,٤

المطلوب: احسب متوسط طول نبات القطن.

الحل: يتم تطبيق القانون التالي للحصول على النتيجة:

$$\bar{X} = \frac{\sum Xi}{n} = \frac{1.4 + 1.33 + 1.32 + 1.29 + 1.5 + 1.47 + 1.36 + 1.4 + 1.23}{9}$$

=**1.366** Metres.

ملاحظة: عند حساب الوسط الحسابي من البيانات الموجودة في جدول توزيع تكراري (البيانات المبوبة) فتكون كما يوضحها المثال

مثال : تم حساب أوزان ١٠٠ طالب من قبل أحد الباحثين وتم وضع النتائج في جدول توزيع تكراري هو على النحو التالي:

الفئات (Classes)	التكرار (Fi) Frequency	مركز الفئات Xi	التكرار × مركز الفئات $Fi.Xi$
٦٠ ـ ٦٢	٥	٦١	٣٠٥
٦٣ ـ ٦٥	١٨	٦٤	١١٥٢
٦٦ ـ ٦٨	٤٢	٦٧	٢٨١٤
٦٩ ـ ٧١	٢٧	٧٠	١٨٩٠
٧٢ ـ ٧٤	٨	٧٣	٥٨٤
المجموع	١٠٠		$\sum Fi.Xi =$٦٧٤٥

الوسط الحسابي= $\dfrac{\text{مجموع حاصل ضرب التكرار × مراكز الفئات}}{\text{المجموع التكراري}}$

أي أن متوسط أوزان الطلبة هو:

$$\bar{X} = \frac{\sum Fi.Xi}{\sum Fi} = \frac{6745}{100} = \mathbf{67.45}$$

مثال : البيانات التالية تمثل نتائج قياس أطوال ثلاثين نبات من الذرة الصفراء (الأطوال مقاسة بالمتر)، **المطلوب**: شكل جدول توزيع تكراري ثم احسب الوسط الحسابي .

١,٦٢	١,٥	١,٨٦	١,٧٢	١,٦٩	١,٥٩
١,٨	١,٧٦	١,٧٥	١,٦٦	١,٧١	١,٦٢
١,٦	١,٦٨	١,٨٦	١,٨٩	١,٧٥	١,٧٧
١,٥٩	١,٥٦	١,٩	١,٩١	١,٦٤	١,٦٤
١,٧٦	١,٧	١,٧٧	١,٩٢	١,٦	١,٨٣

الحل: لتنظيم جدول توزيع تكرار يجب تحديد القيمة الصغرى والعظمي لتحديد المدى ومن ثُمَّ تحديد عدد الفئات وطولها كما مبين في الخطوات التالية:

١- (أقل قيمة – أعلى قيمة – المدى) هي (١,٥ – ١,٩٢ – ٠,٤٢) على التوالي.
٢- عدد الفئات = ١ + (٣,٣ × لو (٣٠)) = ٥,٨٧٤ ≈ ٦.
٣- طول الفئة = المدى ÷ عدد الفئات = ٠,٤٢ ÷ ٦ = ٠,٠٧.
٤- تنظيم الجدول كالآتي:

جدول يبين جدول التوزيع التكراري لأطوال نباتات الذرة الصفراء للمثال .

الفئات	التكرار رمزاً	التكرار كتابةً	مركز الفئات	التكرار × مركز الفئات
١,٥٠ – ١,٥٦	//	٢	١,٥٣	٣,٠٦
١,٥٧ – ١,٦٣	/////	٥	١,٦٠	٨
١,٦٤ – ١,٧٠	///////	٧	١,٦٧	١١,٦٩
١,٧١ – ١,٧٧	////////	٨	١,٧٤	١٣,٩٢
١,٧٨ – ١,٨٤	//	٢	١,٨١	٣,٦٢
١,٨٥ – فأكثر	//////	٦	١,٨٩	١١,٣٤
المجموع		٣٠		٥١,٦٣

٥- حساب الوسط الحسابي من المعادلة التالية:

$$\text{الوسط الحسابي} = \frac{\text{مجموع حاصل ضرب التكرار × مراكز الفئات}}{\text{المجموع التكراري}}$$

= **١,٧٢** متراً.

الوسط الحسابي الموزون Weighing mean :

وهو يمثل حاصل قسمة مجموع حاصل ضرب القراءات في وزنها مقسوم على مجموع الأوزان، ويرمز له بالرمز$(\overline{WX})$ ويتم حسابه من المعادلة التالية :

$$\overline{WX} = \frac{\sum wixi}{\sum wi}$$

حيث أن:

Xi = the values (القيم)
Wi = the weight of the value (وزن القيمة)

مثال : البيانات التالية تمثّل معدل احد الطلبة تبعا للمرحلة، والمطلوب حساب معدله النهائي (أي إيجاد الوسط الحسابي الموزون).

المرحلة	معدل الطالب (القيم Xi)	وزن كل مرحلة (الوزن Wi)	القيم × الوزن ($Xi.Wi$)
الأولى	٦٠	١٠ %	٦٠٠
الثانية	٧٠	٢٠ %	١٤٠٠
الثالثة	٨٠	٣٠ %	٢٤٠٠
الرابعة	٧٥	٤٠ %	٣٠٠٠
المجموع		١٠٠ %	٧٤٠٠

$$\text{الوسط الحسابي الموزون} = \frac{\text{مجموع حاصل ضرب القيم} \times \text{مجموع أوزانها}}{\text{مجموع الأوزان}}$$

$$= \frac{٧٤٠٠}{١٠٠} = ٧٤ \ \%$$

الوسيط: The Median($\overline{Me}$)

هي قيمة تحسب أو تستخرج مباشرة بعد ترتيب البيانات تصاعدياً أو تنازلياً، فإذا كان عدد القيم فردي فبعد ترتيب البيانات تصاعدياً أو تنازلياً فان ترتيب الوسيط هو عدد القيم + ١ مقسوماً على ٢، أما اذا كان عدد القيم زوجياً فإن قيمة الوسيط ستكون قيمتين إذ تكون القيمة الأولى من خلال قسمة عدد القيم على ٢ والثانية هي القيمة التي تليها أي بعد إضافة (١)، أي بمعنى:

أ- في حالة عدد القيم فردياً:

$$\text{الوسيط} = \frac{\text{عدد القيم} + ١}{٢}$$

ب- في حالة عدد القيم زوجياً: نتبع الخطوات التالية:

- $\text{القيمة الأولى} = \frac{\text{عدد القيم}}{٢}$
- $\text{القيمة الثانية} = \frac{\text{عدد القيم}}{٢} + ١$
- $\text{الوسيط} = \frac{\text{القيمة الأولى} + \text{القيمة الثانية}}{٢}$

مثال: جد وسيط القيم التالية:

١- [١١، ١٢، ٨، ١٠، ١٦، ١٥، ٩].
٢- [١٠، ١٥، ١٤، ١٣، ٨، ١٥، ١٠، ١١].

الحل:

١- في الحالة الأولى يتم ترتيب القيم تصاعدياً أو تنازلياً ليكون شكل القيم كما يلي:
[٨، ٩، ١٠، ١١، ١٢، ١٥، ١٦]، من الملاحظ أن عدد القيم هو (٧) وهو عدد فردي، فتكون قيمة الوسيط هي التي ترتيبها (٧+١) ÷ ٢ =٤ والقيمة الرابعة مقدارها = ١١ وهي تمثل قيمة الوسيط.

٢- في الحالة الثانية وبعد ترتيب القيم تصاعدياً أو تنازلياً سيكون شكل القيم كالآتي:
[٨، ١٠، ١٠، ١١، ١٣، ١٤، ١٥، ١٥]، فمن الملاحظ أن عدد القيم = ٨ وهو عدد زوجي، فتكون قيمة الوسيط كالآتي: متوسط قيمتين الاولى (عدد القيم ÷ ٢) والثانية ((القيمة الأولى) + ١) وهذا يعني القيمة الرابعة والخامسة ومن ثم حساب متوسطيهما أي بمعنى سيكون الوسيط = (١١+١٣) ÷ ٢ =١٢.

المحاضرة السادسة

المنوال ($\overline{Mo}$) The Mode:

وهي القيمة الأكثر تكراراً وتسمى أيضاً بالقيمة المنوالية، وقد تكون هنالك أكثر من قيمة منوالية واحدة أو أكثر من قيمتين منواليتين أو قد لا يكون هنالك قيمة منوالية والمثال التالي يوضّح ذلك.

مثال : تم الحصول على البيانات التالية والتي تمثل عدد الأيام من الزراعة وحتى بزوغ ٨٠% البذور لثلاثة أصناف من القطن وكما يلي:

جدول : يبين عدد الأيام من الزراعة وحتى بزوغ ٨٠% لثلاثة أصناف من محصول القطن.

الصنف الأول	الصنف الثاني	الصنف الثالث
١٢	١٥	١٠
١٤	١٢	١٠
١٥	١٦	١٤
١٦	١٣	١٦
٩	١٥	١٤
١٧	١٥	١٣

١٨	١٠	١٧
١٣	١٥	١٠
١١	١٣	١٤
٧	١٥	٩
٢٠	٨	١٢

المطلوب: ماهي القيمة المنوالية لكل صنف؟

الحل:

في الصنف الأول: لا توجد قيمة منوالية.
في الصنف الثاني: القيمة المنوالية هي (١٥).
في الصنف الثالث: هنالك قيمتين منواليتين هما (١٠) و (١٤).

المحاضرة السابعة

مقاييس التشتت: Measures Of Dispersion Or Variation

هي المقاييس التي تبحث عن قيمة تدل على مقدار تشتت او اختلاف او تباعد المشاهدات (قيم المتغير) عن وسطها الحسابي. فكلما كان مقدار التشتت كبير دل ذلك على عدم التجانس بين القيم، أي أن القيم متباينة. وتقسم مقاييس التشتت إلى :

١- الدرجة القياسية (Standardized Scores).
٢- مقاييس التشتت النسبي
٣- مقاييس التشتت المطلق: هي مقاييس تاخذ وحدة قياس نفس وحدة المتغير وتقسم إلى:
  أ- المدى (Rang).
  ب- الانحراف المتوسط (Mean Deviation).
  ج- التباين (Variance).
  د- الانحراف القياسي (Standard deviation).
  ه- الخطأ القياسي (Standard Error).

المدى (R) The Range: وهو الفرق بين اكبر قيمة واقل قيمة، أي أن

$$R = max\,X - min\,X$$

مثال : البيانات التالية تمثل تواجد الحشرات على نباتات صنفين من نبات السلق:

الصنف الأول	٩	١٧	١٣	٦	١٩	٢١	٥	٢٤	٢٦	١٠
الصنف الثاني	١٠	١٩	١٧	١٤	١٣	١١	١٤	٣١	١٢	١٣

المطلوب: إيجاد المدى لكل صنف. ....الحل/

المدى للصنف الأول = ٢٦ - ٥ = ٢١.

المدى للصنف الثاني = ٣١ - ١٠= ٢١.

الانحراف المتوسط: The mean absolute deviation

هو المجموع المطلق لإنحرافات القيم عن متوسطها مقسوما على عدد القيم ويرمز له بالرمز (MAD)، أي أن:

$$MAD = \frac{\sum |X_i - \bar{X}\,|}{n}$$

مثال : البيانات التالية تمثّل أطوال محصول الشعير عند مرحلة الأشطاء، المطلوب إيجاد الإنحراف المتوسط للأطوال.

أطوال النباتات	٥٤	٦٢	٦٣	٦٥	٦٨	٧١	٧٣	٧٨	٨٢	٨٤

الحل: للتوضيح سوف يتم عمل جدول بخطوات الحل، وكما يلي:

	أطوال النباتات	الإنحراف عن المتوسط $(\bar{X})$ $(X_i - \bar{X})$	الإنحراف المطلق $\lvert X_i - \bar{X} \rvert$
	٥٤	-١٦	١٦
	٦٢	-٨	٨
	٦٣	-٧	٧
	٦٥	-٥	٥
	٦٨	-٢	٢
	٧١	١	١
	٧٣	٣	٣
	٧٨	٨	٨
	٨٢	١٢	١٢
	٨٤	١٤	١٤
المجموع	٧٠٠	صفر	٧٦
المتوسط	٧٠		

ومن خلال تطبيق معادلة الإنحراف المتوسط المذكورة آنفاً:

$$MAD = \frac{\sum |X_i - \bar{X}\ |}{n}$$

$$= \frac{76}{10} = \mathbf{7.6}$$

المحاضر
الثامنة

التباين والانحراف القياسي (Variance and Standard Deviation)

إن من اكثر مقاييس التشتت المطلق استخداماً هما التباين والانحراف القياسي، إذ يرمز لتباين العينة بالرمز $\boldsymbol{S}^2$ ويرمز للانحراف القياسي للعينة $\boldsymbol{S}$، وهذا يعني أن الانحراف القياسي هو الجذر التربيعي للتباين، فالتباين مقياس مطلق ويأخذ نفس وحدة قياس البيانات ألا أنها بصفة التربيع فمثلاً عند قياس أطوال طلبة بوحدة (سم) تكون قيمة التباين (سم^2) ولا تفهم هكذا لأنها تحولت إلى وحدة مساحة بدلاً من وحدة طول ولذلك يعتبر الانحراف القياسي هو المقياس الأمثل لحل هذه المشكلة لان الوحدة التي يأخذها هي نفس وحدة القيم (سم) مثلاً لأطوال الطلبة. ومن اجل حساب الانحراف القياسي يتطلب أولاً حساب التباين ثم جذر القيمة لتصبح انحرافاً قياسياً.

طرق حساب التباين:

١- طريقة انحرافات القيم.
٢- طريقة تربيع القيم **وهي المفضلة**.

مثال: ما هو تباين والانحراف القياسي للبيانات التالية:

Xi	١٧	٢٠	٢٣	١٨	١٩	٢١	٢٢

- حساب التباين (بالطريقة الأولى) وهي انحرافات القيم:

أ- حساب قيمة المتوسط $\overline{Xi} = \dfrac{17+20+23+18+19+21+22}{7} = 20$

ب- تشكيل جدول الانحراف ومربعات الانحراف على النحو التالي :

Xi	$Xi - \bar{X}$	$(Xi - \bar{X})^2$
17	17-20=-3	$(-3)^2=9$
20	20-20=0	$(0)^2=0$
23	23-20=3	$(3)^2=9$
18	19-20=-2	$(-2)^2=4$

$(-1)^2=1$	19-20=-1	19
$(1)^2=1$	21-20=1	21
$(2)^2=4$	22-20=2	22
$\sum(X_i - \bar{X})^2 = 28$	$\sum Xi-\bar{X}$=Zero	140

ويكون حساب التباين، حسب المعادلة التالية:

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{28}{7-1} = 4.67$$

ومن خلال وضع الناتج تحت الجذر التربيعي يتم الحصول على الإنحراف القياسي، وكما يلي:

$$S = \sqrt{S^2} = \sqrt{4.67} = 2.16$$

- حساب التباين بالطريقة الثانية (تربيع القيم)، إذ لانحتاج هذه الطريقة الى حساب الوسط الحسابي كما أن في الطريقة الأولى لو كانت قيمة متوسط البيانات قيمه دوريه فسيكون التباين المحسوب بها غير دقيق بسبب أن جميع القيم تطرح من الوسط الحسابي الذي نضطر الى تقريب قيمته .
-

Xi	Xi^2
17	$17^2=289$
20	$20^2=400$
23	$23^2=529$
18	$18^2=324$
19	$19^2=361$
21	$21^2=441$
22	$22^2=464$
$\sum Xi= 140$	$\sum Xi^2=2828$

وعليه فإن التباين يتم حسابه من المعادلة التالية:

$$S^2 = \frac{\sum {X_i}^2 - \frac{(\sum X_i)^2}{n}}{n-1} = \frac{2828 - \frac{(140)}{7}}{7-1} = 4.67$$

وعليه فإن قيمة الإنحراف القياسي هوجذر التباين $S = \sqrt{S^2}$

$$S = 2.16$$

حساب التباين من البيانات المبوبة (جدول توزيع تكراري) :

حساب التباين والانحراف القياسي من جدول التوزيع التكراري للمثال الذي تناول اوزان الطلبة (مثال ٣ – ٢) في الفصل الثالث مقاييس التمركز، والجدول التالي يوضح ذلك:

الفئات (Classes)	التكرار (Fi) Frequency	مركز الفئات Xi	Xi^2	التكرار × مركز الفئات $Fi.Xi$	$Fi.Xi^2$
٦٠ – ٦٢	٥	٦١	٣٧٢١	٣٠٥	١٨٦٠٥
٦٣ – ٦٥	١٨	٦٤	٤٠٩٦	١١٥٢	٧٣٧٢٨
٦٦ – ٦٨	٤٢	٦٧	٤٤٨٩	٢٨١٤	١٨٨٥٣٨
٦٩ – ٧١	٢٧	٧٠	٤٩٠٠	١٨٩٠	١٣٢٣٠٠
٧٢ – ٧٤	٨	٧٣	٥٣٢٩	٥٨٤	٤٢٦٣٢
المجموع	١٠٠			$\sum Fi.Xi =$ ٦٧٤٥	$\sum Fi.Xi^2 =$ ٤٥٥٨٠٣

من الملاحظ أن الجدول أعلاه قد تم إضافة عمودين جديدين هما (Xi^2 و $Fi.Xi^2$)، إذ سيتم الاستفادة من هذين العمودين في إستخراج أحد مكونات معادلة حساب التباين وكما يلي:

فمن خلال تطبيق معادلة الطريقة الأولى لحساب التباين تكون خطوات الحل كما يلي:

$$\sum(X_i - \bar{X})^2 = \sum F_i.X_i{}^2 - \frac{(\sum F_i.X_i)^2}{\sum F_i} = 455803 - \frac{(6745)^2}{100} = \mathbf{852.75}$$

وعند تطبيق معادلة التباين (الطريقة الأولى)، تكون النتيجة كما يلي:

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{852.75}{99} = \mathbf{8.61}$$

وتكون قيمة الإنحراف القياسي هي :

$$S = \sqrt{S^2} = \sqrt{8.61} = \mathbf{2.9}$$

مثال : احسب التباين والانحراف القياسي من خلال الجدول (٤ – ١).

الجدول يبين التوزيع التكراري لبيانات إحدى القرى التابعة لإحدى محافظات العراق.

حدود الفئات	مركز الفئة (Xi)	التكرار (Fi)	التكرار×مركز الفئة ($Fi.Xi$)	مربع مركز الفئة (Xi^2)	التكرار×مربع مركز الفئة ($Fi.Xi^2$)
١٠ – ١٩	١٤,٥	٥	٧٢,٥	٢١٠,٢٥	١٠٥١,٢٥
٢٠ – ٢٩	٢٤,٥	١٩	٤٥٦,٥	٦٠٠,٢٥	١١٤٠٤,٧٥
٣٠ – ٣٩	٣٤,٥	١٠	٣٤٥,٠	١١٩٠,٢٥	١١٩٠٢,٥
٤٠ – ٤٩	٤٤,٥	١٣	٥٧٨,٥	١٩٨٠,٢٥	٢٥٧٤٣,٢٥
٥٠ – ٥٩	٥٤,٥	٤	٢١٨,٠	٢٩٧٠,٢٥	١١٨٨١,٠
٦٠ – ٦٩	٦٤,٥	٤	٢٥٨,٠	٤١٦٠,٢٥	١٦٦٤١,٠
٧٠ – ٧٩	٧٤,٥	٢	١٤٩,٠	٥٥٥٠,٢٥	١١١٠٠,٥
المجموع	٣١١,٥	٥٧	٢٠٨٩,٥	١٦٦٦١,٧٥	٨٩٧٢٤,٢٥
المتوسط	٣٦,٦				

الحل:

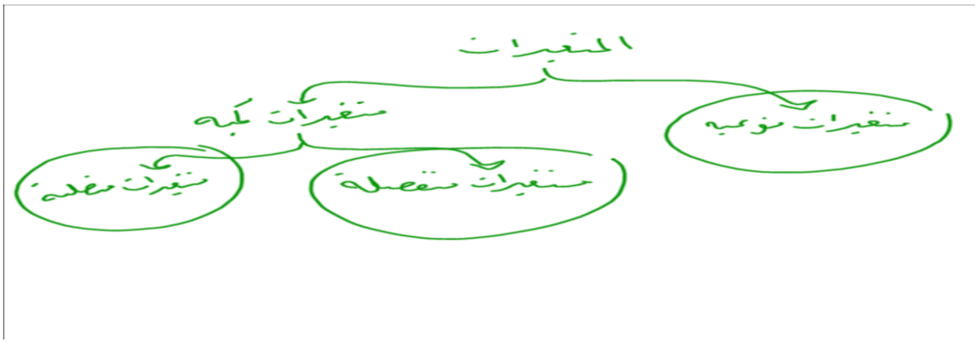
$$\text{التباين} = \frac{\text{مجموع حاصل ضرب التكرار} \times \text{مربع مركز الفئة} - \dfrac{(\text{التكرار} \times \text{مركز الفئة})^{٢}}{\text{مجموع التكرار}}}{\text{مجموع التكرار - ١}}$$

أي أن:

$$S^2 = \frac{\Sigma FiXi^2 - \dfrac{(\Sigma FiXi)^2}{\Sigma Fi}}{\Sigma Fi - 1} = \frac{89724.25 - \dfrac{(2089.5)^2}{57}}{57 - 1} = 234.4211 \quad S = \sqrt{234.4211} = 15.31$$

<u>التوزيع الاحتمالي للمتغير المستمر</u> <u>Continuous Probability</u>

عند تمثيل بيانات المتغير الكمي المستمر في شكل مدرج تكراري النسبي، نجد أن شكل هذا المدرج هو أقرب وصف لمنحنى التوزيع الاحتمالي للمتغير المستمر، وكلما ضاقت الفترات بين مراكز الفئات، يمكن الحصول على رسم دقيق للمنحنى الخاص بدالة احتمال المتغير المستمر.



المحاضر

التاسعة

<u>التوزيعات الاحتمالية للمتغيرات المنفصلة:</u>

المتغير العشوائي المنفصل (مقارنة بالمتصل) هو المتغير الذى يمكن أن يأخذ قيما محدودة ومتميزة، وتسمى مجموعة كل القيم الممكنة للمتغير العشوائي واحتمالاتها المناظرة **بالتوزيع الاحتمالي،** ويكون مجموع الاحتمالات = 1

أ- **<u>توزيع ذي الحدين:</u>**

يستخدم هذا التوزيع في الحالات التي يكون للظاهرة محل الدراسة نتيجتان فقط متنافيتان، النتيجة محل الاهتمام وتسمى بحالة النجاح، والأخرى تسمى بحالة الفشل،

إذا فتوزيع ذو الحدين هو أحد التوزيعات الاحتمالية المنفصلة، ويستخدم لإيجاد احتمال وقوع حدث معين (نجاح) عدداً من المرات مقداره **X** من بين **n** من المحاولات لنفس التجربة (ونرمز لهذا الاحتمال بالرمز **P(X)** وذلك عندما تتحقق الشروط التالية :

- هناك ناتجان ممكنان فقط ومتنافيان لكل محاولة
- المحاولات وعددها **n مستقلة** عن بعضها البعض

- احتمال وقوع الحدث المعين فى كل محاولة (النجاح) **P** ثابت ولا يتغير من محاولة لأخرى **<u>تحديد</u>**

<u>شكل التوزيع:</u>

يتحدد شكل التوزيع ثنائي الحدين وفقا لقيمة احتمال النجاح كما يلي:

- إذا كان $p = 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون متماثل.
- إذا كان $p < 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون موجب الالتواء.
- إذا كان $p > 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون سالب الالتواء.

$p > 0.5$ **الاحتمال أكبر من ٠,٠٥**

<u>التوزيع الطبيعى</u>

هو أفضل وأكثر التوزيعات الاحتمالية المتصلة استخداماً في النواحي التطبيقية، ومنها الاستدلال الإحصائي شاملا التقدير، واختبارات الفروض، كما أن معظم التوزيعات يمكن تقريبها إلى هذا التوزيع، وفيما يلي عرض لهذا التوزيع.

ويكمن السبب في ذلك :

- أن توزيعات كثيرة لمتغيرات مثل الطول والوزن تتبع توزيعات طبيعية.
- النتيجة الرياضية التي تسمى بنظرية النهاية المركزية.

والتوزيع الطبيعي هو توزيع احتمالي متصل ويعد من أكثر التوزيعات استخداماً فى التحليل الإحصائى، وهو جرسي الشكل ومتماثل حول الوسط الحسابي، ويمتد إلى مالا نهاية فى الاتجاهين، ولكن معظم المساحة (الاحتمال) تتركز حول الوسط الحسابى.

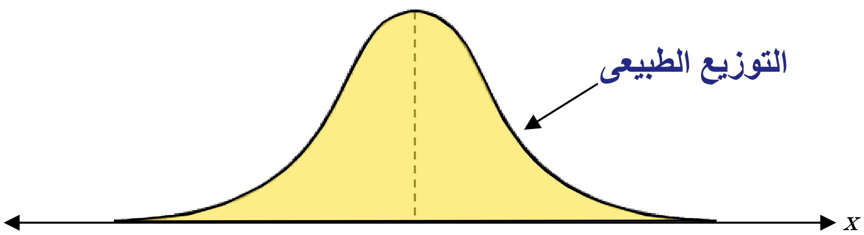
ولرسم التوزيع الطبيعى نستخدم المنحنى الاحتمالي للتوزيع الطبيعي

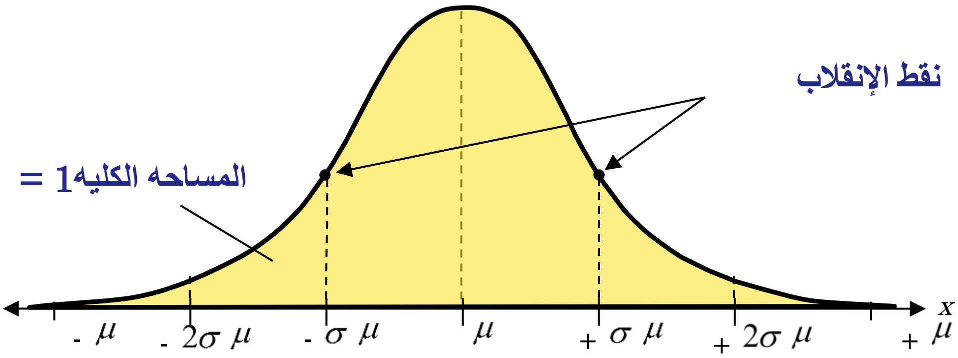
إذا كانت x متغيراً عشوائياً متصلاً يتبع توزيعاً طبيعياً متوسطه μ و تباينه σ^2 يمكن رسم المنحنى الاحتمالي للتوزيع الطبيعي بإستخدام المعادله

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \quad e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad ; \; -\infty \le x \le \infty$$

فإنه يقال إن x تتبع توزيعاً طبيعياً متوسطه μ و تباينه σ^2. و تكتب اختصاراً

$$X \sim N(\mu , \sigma^2)$$





خصائص المنحنى الإحتمالى للتوزيع الطبيعى

١. طرفا التوزيع تمتد من $-\infty$ إلى ∞
٢. متماثل حول العمود الذى يمر بقمته أى عند $X=\mu$ لذلك فإن هذا العمود يجزئ المنحنى الطبيعى القياسى إلى قسمين متماثلين فى الشكل والمساحه.
٣. يشبه الناقوس من حيث الشكل.
٤. المنحنى له نقطتا إنقلاب عند $X=\mu\pm\sigma$
٥. المساحه الكليه تحت المنحنى تساوى الواحد الصحيح .
٦. نظرا لتطابق جانبيه وتوسط تفرطحه فإن مقاييس النزعه المركزيه الثلاثه (الوسط الحسابى والوسيط والمنوال) تلتقى فى نقطه واحده (فى منتصف التوزيع) أى أن الوسط الحسابى = الوسيط = المنوال
٧. نلاحظ أن ٩٩.7% من قيم المتغير تقع تقريبا فى الفتره $(\mu-3\sigma\ ,\ \mu+3\sigma)$ أى أنه نادرا ما نجد قيمه من قيم X خارج هذه الفتره. كما أن ٩٥.5% من قيم المتغير X تقع فى الفتره $(\mu-2\sigma\ ,\ \mu+2\sigma)$ وأن% ٦٨ منها تقع فى الفتره$(\mu-\sigma\ ,\ \mu+\sigma)$

المحاضر
العاشرة

اختبارات الفروض الإحصائية

يتعرض الإنسان في كثير من الحالات وفي مجالات العمل الى مواقف معينة تتطلب منه اتخاذ قرار بناء على معلومات محسوبة من عينه ، وعليه يجب اتخاذ هذا القرار بأقل قدر ممكن من الخطأ.

مثال: نفرض أن باحثا اجتماعيا ادعى ان متوسط اعمار طلاب الجامعة لايختلف عن متوسط أعمار الطالبات. للتأكد من ذلك فإن الشئ الطبيعي أن نقوم بحصر اعمار الطلاب والطالبات ومنها نحسب المتوسط لكل منهما ثم نقرر من منهما اكبر..ولكن عملية الحصر صعبة ومجهدة لذلك نضطر الى اختيار عينة عشوائية من بين الطلاب وعينة عشوائية من بين الطالبات ونحسب متوسط العمر في كل عينة منهما،،،،فإذا كان متوسط عمر الطالب هو ٢٤ وكان متوسط عمر الطالبة هو ٢٢ فهل يعني ذلك ان متوسط عمر الطالب اكبر من متوسط عمر الطالبة ؟؟ هل الفرق راجع لمجرد الصدفة ؟؟ متى يكون الفرق نتيجة للصدفة؟؟ متى يكون الفرق دالا على وجود اختلاف حقيقي أو جوهري بين متوسطي المجتمعين الأصليين ..

الفرض الاحصائى statistical hypothesis

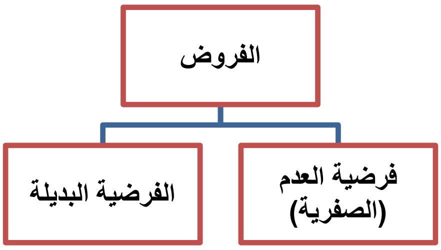
هو عبارة عن إدعاء او تخمين معين حول معلمة من معالم المجتمع ويكون المطلوب اختبار صحة هذا الادعاء أو التخمين ...هناك نوعين من الفروض :

- **فرض العدم (null hypothesis)** ويرمز له بالرمز H_0 ويصاغ في صورة عدم وجود فرق أو عدم وجود علاقة أو عدم وجود تغير – مثال : في مثال أعمار الطلاب وطالبات الجامعة فإن فرض العدم هو

H_0 : نفترض عدم وجود اختلاف بين متوسطي اعمار الطلاب والطالبات

- **الفرض البديل (alternative hypothesis)** ويرمز له بالرمز H_1 وهو الفرض الذي يجب أن يكون صحيحا اذا كان فرض العدم غير صحيح – مثال : في مثال أعمار الطلاب وطالبات الجامعة فإن الفرض البديل هو

H_1: يوجد اختلاف حقيقي وليس ظاهري بين متوسط اعمار الطلاب والطالبات .



مستوى المعنوية α ودرجة الثقة (1 – α):

- إن القرار الذي سوف نتخذه بناء على الاختبار الإحصائي لا يمكن اعتباره صحيح % 100 فهناك مقدار من الخطأ لأن المعلومات التي نتخذ قرارنا بناء عليها بيانات مأخوذة من عينة وليس من المجتمع الأصلي
- في اختبار فرض معين، فإن مقدار ثقتنا في القرار المتخذ بالرفض أو القبول يسمى بدرجة الثقة ويرمز له بالرمز (1 – α) كما وأن مقدار عدم الثقة أو مقدار الخطأ يسمى بمستوى المعنوية ويرمز له بالرمز α

وعادة يحدد الباحث مستوى المعنوية أو درجة الثقة قبل البدء في عملية الاختبار.

عند اختبار فرض العدم H_o ضد الفرض البديل H_1 نجد أننا امام احدى الحالات الاربع الاتية :

	H0 صحيح	H0 خطأ
قبول H0	قرار سليم	خطأ من النوع الثاني
رفض H0	خطأ من النوع الاول	قرار سليم

١) أن يكون فرض العدم صحيحا ويكون القرار بقبوله ....وهذا قرار سليم
٢) أن يكون فرض العدم صحيحا ويكون القرار برفضه .... وهذا قرار خاطئ
(الخطأ من النوع الأول : رفض H_o عندما يكون H_o صحيحا ويرمز لحجم هذا الخطأ بالرمز α)
٣) أن يكون فرض العدم خطأ ويكون القرار برفضه وهذا قرار سليم
٤) أن يكون فرض العدم خطأ ويكون القرار بقبوله ..وهذا قرار خاطئ (الخطأ من النوع الثاني : قبول H_o عندما يكون H_o خاطئ ويرمز لحجم هذا الخطأ بالرمز β)

- احتمال الوقوع في الخطأ من النوع الأول يسمى مستوى المعنوية ويرمزله بالرمز α أي ان α = احتمال رفض فرض العدم H_o عندما يكون صحيح = مستوى المعنوية

- احتمال الوقوع في خطأ من النوع الثاني يرمز له بالرمز β أي أن β = احتمال قبول فرض العدم H_o عندما يكون خطأ

المحاضرة الحادية عشر

الاختبارات المعلمية Parametric Test

تستخدم في حال خضوع البيانات للتوزيع الطبيعي ومنها اختبار تي تست والارتباط البسيط وتحليل التباين

توزيـــــــــع T- Distribution. T

يعد اختبار "ت" من أكثر اختبارات الدلالة شيوعاً في الأبحاث النفسية والاجتماعية والتربوية ، وترجع نشأته الأولى إلى أبحاث العالم "ستودنت" ولهذا سمى الاختبار بأكثر الحروف تكراراً في اسمه وهو حرف التاء .
ومن أهم المجالات التى يستخدم فيها هذا الاختبار الكشف عن الفروق بين تحصيل الذكور والإناث فى مادة دراسية ما وذلك عن طريق حساب دلالة فرق متوسط تحصيل الذكور عن متوسط تحصيل الإناث .
ويمكن القول أن اختبار "T" يستخدم لقياس دلالة فروق المتوسطات غير المرتبطة والمرتبطة للعينات المتساوية والغير متساوية .كما يستخدم اختبار ت اذا كان قيمة n اقل من 30 اكثر من 30 فيستخدم اختبار z وبنفس صيغة القانون T

١- في حالة تساوى العدد في العينين يستخدم القانون التالى

$$T = \frac{\bar{X}1 - \bar{X}2}{\sqrt{\frac{{S_1}^2 + {S_2}^2}{N-1}}}$$

$\bar{X}1$ = المتوسط الحسابي للعينة الاولى
$\bar{X}2$ = المتوسط الحسابي للعينة الثانية
${S_1}^2$ = تباين العينة الاولى
${S_2}^2$ = تباين العينة الثانية
N= عدد افراد احدى العينتين (لان العدد متساوي)

مثال:

اراد باحث ارشادي ان يتعرف على درجة المعرفة الزراعية لمجموعتين من المزارعين عدد كل منها يساوي ٦ وكانت درجات المعرفة في الاختبار كما ياتي علما ان تباين العينة الاولى 10= ${S_1}^2$ وتباين العينة الثانية 30 = ${S_2}^2$

درجات الحرية DF=(N-1) =(6-1) =5
قيمة T الجدولية =2.57

الحل:
نكتب الفرضيات
لاتوجد فروق معنوية بين متوسطي المعرفة بين مجموعتي المزارعين $Ho : \mu1 = \mu2$
توجد فروق معنوية بين متوسطي المعرفة بين مجموعتي المزارعين $H1 : \mu1 \neq \mu2$

.

$$T = \frac{5 - 6}{\sqrt{\frac{10 + 30}{6-1}}} = \frac{-1}{\sqrt{\frac{40}{5}}} = \frac{-1}{\sqrt{8}} = \frac{-1}{2.828} = 0.353$$

القرار:

بما ان t المحسوبة هي اقل من t الجدولية لمستوى الاحتمال 0.05 اذن لاتوجد فروق معنوية بين متوسطي العينتين اذن نقبل فرضية العدم ونرفض البديلة

.

٢- في حالة عدم تساوي العدد في العينين يستخدم القانون التالى:

$$t = \frac{\bar{X}1 - \bar{X}2}{\sqrt{\left(\frac{N_1{S_1}^2 + N_2{S_2}^2}{N_1 + N_2 - 2}\right)\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

$\bar{X}1$ = المتوسط الحسابي للعينة الاولى
$\bar{X}2$ = المتوسط الحسابي للعينة الثانية
${S_1}^2$ = تباين العينة الاولى
${S_2}^2$ = تباين العينة الثانية
N_1= عدد افراد العينة الاولى
N_2= عدد افراد العينة الثانية

مثال:

اجريت دراسة على 6 فلاحين و5 فلاحات لمعرفة الفروق في مستوى المعلومات الزراعية بين الفلاحين والفلاحات وكانت القيم التي تم الحصول عليها كما يلي علما ان تباين العينة الاولى هو 10وتباين العينة الثانية هو 14

الحل:

نكتب الفرضيات اولا:

لاتوجد فروق معنوية بين متوسطي الفلاحين والفلاحات $Ho : \mu 1 = \mu 2$

لاتوجد فروق معنوية بين متوسطي الفلاحين والفلاحات $H1 : \mu 1 \neq \mu 2$

$$t = \frac{\bar{X}1 - \bar{X}2}{\sqrt{\left(\frac{N_1 {S_1}^2 + N_2 {S_2}^2}{N_1 + N_2 - 2}\right)\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

العينة الثانية	العينة الاولى
15	5
19	10
16	8
10	4
10	2
----	1
$\bar{X}2=\frac{70}{6}=14$	$\bar{X}1=\frac{30}{6}=5$

$$t = \frac{5 - 14}{\sqrt{\left(\frac{6 * 10 + 5 * 14}{6 + 5 - 2}\right)\left(\frac{1}{6} + \frac{1}{5}\right)}} = \frac{-9}{\sqrt{\left(\frac{60 + 70}{9}\right)\left(\frac{1}{6} + \frac{1}{5}\right)}} = -3.947$$

درجات الحرية$DF= N_1 + N_2 - 2$=(6+5-2) =9

قيمة Tالجدولية= 2.26

القرار:

بما ان t المحسوبة هي اكبر من tالجدولية لمستوى الاحتمال 0.05 اذن توجد فروق معنوية بين متوسطي العينتين اذن نقبل فرضية البديلة ونرفض العدم

المحاضرة الثانية عشر

تحليل التباين

هو طريقة لاختبار معنوية الفرق بين المتوسطات لعدة عينات بمقارنة واحدة، ويعرف أيضاً بطريقة تؤدي لتقسيم الاختلافات الكلية لمجموعة من المشاهدات التجريبية لعدة أجزاء للتعرف على مصدر الاختلاف بينها ولذا فالهدف هنا فحص تباين المجتمع لمعرفة مدى تساوى متوسطات المجتمع ولكن لا بد من تحقيق ثلاثة أمور قبل استخدامه وهي:

١) العينات عشوائية ومستقلة.

٢) مجتمعات هذه العينات كلً لها توزيع طبيعي.

٣) تساوي تباين المجتمعات التي أخذت منها العينات العشوائية المستقلة.

التصنيف الأحادي (في حالة تساوي أحجام العينات) :

يعتبر هذا التصنيف هو أبسط أنواع تحليل التباين، حيث تصنف المشاهدات إلى عدة مجموعات على أساس متغير واحد أو خاصية واحدة.

والافتراضات الأساسية لهذا التحليل هي ما يلي :

١- نفترض أن عدد المجتمعات K وأنها جميعاً مستقلة.

٢- أنها جميعاً تخضع للتوزيع الطبيعي بمتوسطات تساوي $\mu_1, \mu_2, \dots, \mu_k$

٣- أن لها جميعاً التباين نفسه σ^2 أي أن التباين لكل المجتمعات ثابت ويساوي σ^2

٤- يتم سحب عينة عشوائية من كل من هذه المجتمعات وأن أحجام هذه العينات كلها متساوية وتساوي n ويمكن بكل بساطة افتراض عدم تساوي أحجام العينات ولن يختلف أسلوب التحليل على الإطلاق إلا في أشياء بسيطة جداً.

مثال/ أراد احد الباحثين أن يدرس تأثير ثلاثة اصناف من الحنطه على حاصل الحنطه بالطن للدونم والاصناف هي ابو غريب وماكسيباك وشام (١) وحصل على النتائج التالية علما" بان.

t1= ابو غريب   t2= ماكسيباك   t3 = شام (١)

observation							Total Xi.	Mean $X\bar{i}.$
t1	1.59	1.64	1.60	1.63	1.60	1.64	9.70	1.62
t2	1.58	1.63	1.65	1.66	1.63	1.65	9.80	1.63
t3	1.71	1.75	1.69	1.68	1.73	1.74	10.30	1.72
							29.80 X.. or ∑Xij	1.66 $\overline{X}..$

خطوات الحل

١/ حساب معامل التصحيح C.F

$$C.F. = \frac{(\Sigma Xij)^2}{t.r} = \frac{(29.8)^2}{3*6} = 49.34$$

المجموع العام تربيع $\left(\sum (X_{ij})^2\right)$

$(t.r)$ يمثل عدد المشاهدات أو ما يسمى N

٢/ حساب مجموع مربعات الانحرافات الكلية (TSS) (Total sum of square)

TSS =∑Xij^2 - C.F ={ 1.59^2+1.64^2+…..+1.73^2+1.74^2} -49.34

TSS =49.38 -49.34 =0.04

٣/ حساب مجموع مربعات انحرافات المعاملات (Between groups) (sst)(treatment sum of square)

$$sst = \frac{\Sigma xi.^2}{r} - C.F. = \frac{(9.72^2 + 9.8^2 + 10.3^2)}{6} - 49.34 = 0.03$$

r = k

٤/ حساب مجموع مربعات انحرافات الخطأ التجريبي (Within groups) (sse) Error sum of square

sse = TSS -sst = 0.04 -0.03 =0.01

٥/ تشكيل جدول تحليل التباين (Table of Analysis of variance) ويسمى مختصراً (ANOVA Table) ويتكون من ستة حقول توضيحيه الحقل

الاول يوضح مصادر الاختلاف او مصادر التباين (Sources of Variation) وتعرف مختصراً (S.O.V)

الثاني فهو لدرجات الحريه (df) (Degrees of Freedom)

الثالث فهو يبين مجاميع المربعات والمحسوبه بالخطوات السابقه في الخطوات ٢ و٣ و٤ وتعرف مختصرا" (SS) وتعني (Sum of Square)

الرابع والذي هو اهم هذه الحقول فهو حقل متوسط المربعات او حقل التباين Variance ثم يأتي حقل القيمه الفائيه المحسوبه (F) وقد يكون هنالك حقل يوضح (F) الجدوليه للمستويين الاحتماليين ٠,٠٥ و٠,٠١) واما من يقوم بتحليل بياناته من خلال احد البرامج الاحصائيه مثال برنامج Spss فانها تعطي جدول ANOVA و F المحسوبه ولا

تعطي F الجدوليه وبدلا عنها تعطي قيمة P أي معنويه لحد مستوى p التي يحدده الاختبار. والجدول والنتائج بالتحليل اليدوي او الحاسبه على النحو التالي:

S.O.V	(df)	SS	Ms التباين	F calculated	F tabulated	
Treatment (بين المعاملات) (المجموعات)	(t-1)3-1=2	0.03	0.015	22.39^{**}	0.05	0.01
Error الخطأ التجريبي داخل (المجموعات)	(t(r-1), 3(5)=15	0.01	0.00067		3.68	6.36
Total المجموع	rt-1 18-1=17	0.04				

٦/ القرار: بما أن قيمة F المحسوبة هي أكبر من الجدولية تحت مستوى ٠,٠٥ و ٠,٠١ وهذا يعني وجود فروقات عالية المعنوية بين الأصناف.

<u>التصنيف الأحادي (في حالة عدم تساوي أحجام العينات)</u>

مثال٢/ اختبر احد الباحثين تأثير نوع المغذي على ارتفاع نبات الحنطة وقد اجرى التجربة في ظروف تحكمية (بيت زجاجي) واستخدم مرشه يدويه صغيره لرش المغذيات وحصل على النتائج التالية التي وضعها بالجدول المرتب وعلى النحو التالي:

مغذي أ	مغذي ب	مغذي ج	مغذي د	
59	64	80	61	
64	60	61	60	
77	58	63	66	
81	60	70	64	
77	52	75		
84	51			
	50			
	55			
442	450	349	251	المجموع
73.7	56.3	69.8	62.8	المتوسط

(General Total)GT = 1492 =المجموع العام

$$C.F. = \frac{GT^2}{r_1 + r_2 + r_3 + r_4} = \frac{(1492)^2}{6+8+5+4} = 96785.4$$

ملاحظة: كل من r_1 و r_2 و r_3 و r_4 تمثل أعداد القيم للمغذيات (أي أن المغذي "أ" له ستة قيم و المغذي "ب" له ثمانية قيم و المغذي "ج" له خمسة قيم والمغذي "د" له أربعة قيم).

$TSS = \sum Xij^2 - C.F = \{ 59^2+64^2+\ldots\ldots+66^2+64^2\} - 96785.4$

TSS= 98930-96785.4=2144.6

$$SSt = (\frac{(442)^2}{6} + \frac{(450)^2}{8} + \frac{(349)^2}{5} + \frac{(251)^2}{4}) - C.F.$$

SSt= { 32568.7+25312.5+24360.2+15750.3}-96785.4

SSt = 97991.65 -96785.4 =1206.3

Sse = TSS - SSt =2144.6 -1206.3 =938.3

ANOVA Table

S.O.V	df	SS	Ms	F Calculated	F Tabulated 0.05	0.01
treatment	3 (عدد المغذيات-١)	1206.3	402.1	**8.149	3.13	5.01
error	$\sum r_i - t = 19$	938.3	49.38			
total	$\sum r_i - 1 = 22$	2144.6				

القرار: بما ان F المحسوبة هي اكبر من ألجدوليه للمستويين الاحتماليين هذا يعني وجود اختلافات عالية المعنوي (الدلاله) بين انواع المغذيات.

المحاضرة الثالثة عشر

الارتباط الخطي البسيط Simple Correlation

الغرض من تحليل الارتباط الخطي البسيط هو تحديد نوع وقوة العلاقة بين متغيرين، ويرمز له في حالة المجتمع بالرمز ρ (رو)، وفي حالة العينة بالرمز r ، وحيث أننا في كثير من النواحي التطبيقية نتعامل مع بيانات عينة مسحوبة من المجتمع، سوف نهتم بحساب معامل الارتباط في العينة r كتقدير لمعامل الارتباط في المجتمع، ومن التحديد السابق للغرض من معامل الارتباط، نجد أنه يركز على نقطتين هما

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

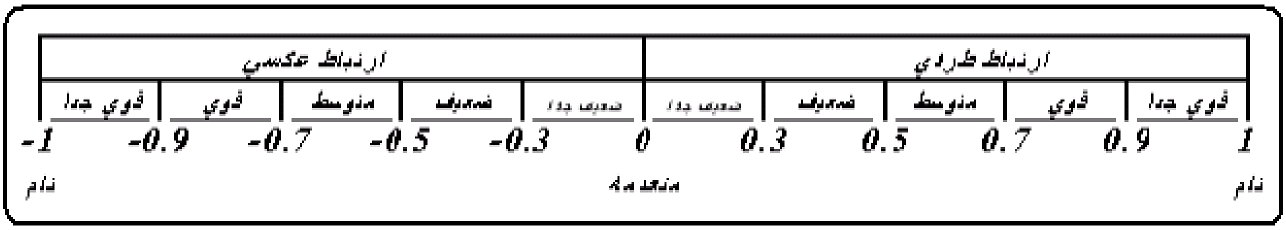
١- إذا كانت إشارة معامل الارتباط سالبة $(r < 0)$ توجد علاقة عكسية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه انخفاض في المتغير الثاني، والعكس.

٢- إذا كانت إشارة معامل الارتباط موجبة $(r > 0)$توجد علاقة طردية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني، والعكس .

٣- إذا كان معامل الارتباط قيمته صفرا $(r = 0)$ دل ذلك على انعدام العلاقة بين المتغيرين.

- قوة العلاقة:- ويمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن (± 1)، حيث أن قيمة معامل الارتباط تقع في المدى$(-1 < r < 1)$، وقد صنف بعض الإحصائيين درجات لقوة العلاقة يمكن تمثيلها على الشكل التالي:

شكل يوضح
درجات قوة معامل الارتباط



معامل الارتباط الخطي البسيط " لبيرسون" Pearson

في حالة جمع بيانات عن متغيرين كميين $(y\ ,\ x)$، يمكن قياس الارتباط بينهما، باستخدام طريقة "بيرسون" Pearson، ومن الأمثلة على ذلك: قياس العلاقة بين الوزن والطول، والعلاقة بين الإنتاج والتكلفة، والعلاقة بين الإنفاق الاستهلاكي والدخل، والعلاقة بن الدرجة التي حصل عليها الطالب وعدد ساعات الاستذكار، وهكذا الأمثلة على ذلك كثيرة.

ولحساب معامل الارتباط في العينة ، تستخدم صيغة " بيرسون" التالية :

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}}$$

حيث :

$\sum XY$ تعني مجموع حاصل ضرب كل قيمة من X في Y.

$(\sum X)$ تعني مجموع قيم المتغير X .

$(\sum Y)$ تعني مجموع قيم المتغير Y .

$\sum X^2$ تعني مجموع مربع قيم المتغير X .

$(\sum X)^2$ تعني مربع مجموع قيم المتغير X .

$\sum Y^2$ تعني مجموع مربع قيم المتغير Y .

$(\sum Y)^2$ تعني مربع مجموع قيم المتغير Y .

n عدد قيم الدراسة (عدد الأزواج المطلوب حساب الارتباط بينها) .

مثـــال

فيما يلي المساحة المنزرعة بالأعلاف الخضراء بالألف هكتار، وإجمالي إنتاج اللحوم بالألف طن، خلال الفترة من 1995 حتى عام 2002.

السنة	1995	1996	1997	1998	1999	2000	2001	2002
المساحة	305	313	297	289	233	214	240	217
الكمية	592	603	662	607	635	699	719	747

والمطلوب: حساب معامل الارتباط بين المساحة والكمية، وما هو مدلوله ؟

الحــل

بفرض أن (x) هي المساحة المنزرعة، (y) هي الكمية، ولحساب معامل الارتباط بين $(x \ , \ y)$ يتم تطبيق المعادلة وذلك على النحو التالي:

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

- **حساب المجاميع:**

x	y	xy	x^2	y^2
305	592	180560	93025	350464
313	603	188739	97969	363609
297	662	196614	88209	438244
289	607	175423	83521	368449
233	635	147955	54289	403225
214	699	149586	45796	488601
240	719	172560	57600	516961
217	747	162099	47089	558009
2108	5264	1373536	567498	3487562

المجاميع المطلوبة
$\sum x=2108\ ,\ \sum y=5264$
$\sum xy=1373536$
$\sum x^2=567498$
$\sum y^2=3487562$

- **حساب معامل الارتباط:**

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

$$=\frac{1373536-\frac{(2108)(5264)}{8}}{\sqrt{\left(567498-\frac{(2108)^2}{8}\right)\left(3487562-\frac{(5264)^2}{8}\right)}}$$

$$=\frac{-13528}{\sqrt{(12040)(23850)}}=\frac{-13528}{16945.619}=-0.798$$

مدلول معامل الارتباط :

بما أن $r=0.798$ ، ويدل ذلك على وجود ارتباط طردي قوي بين بين المساحة والكمية

المحاضرة الرابعة عشر

الاختبارات الغير معلمية Non Parametric Test

تستخدم في حال عدم خضوع البيانات للتوزيع الطبيعي وعندما يكون حجم العينة أقل من ٣٠

معامل ارتباط الرتب (اسبيرمان) Spearman

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبين، ومثال على ذلك قياس العلاقة بين تقديرات الطلبة في مادتين ، أو العلاقة بين درجة تفضيل المستهلك لسلعة معينة ، ومستوى الدخل، فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية ، ويطلق على هذا المعامل " معامل ارتباط سبيرمان " Spearman و يستخدم هذا الاختبار الاحصائي اذا كان حجم العينة صغير ويمكن استخدام صيغة معامل ارتباط "اسبيرمان" في حساب الارتباط بين متغيرين كميين ايضا

، ويعبر عنه بالمعادلة التالية :

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

حيث أن d هي الفرق بين رتب مستويات المتغير الأول x ، ورتب مستويات المتغير الثاني y ، أي أن : $d = R_x - R_y$.

مثـــال

اراد باحث ان يتعرف على العلاقة الارتباطية بين حجم اسرة المزارع وكفاءته الانتاجية وكان حجم العينة ٥ والجدول التالي يبين ادناه البيانات

x	y	رتبة x	رتبة y	d(x-y)	$d(x-y)^2$
5	4	1	2	-1	1
2	1	4	5	-1	1
4	3	2	3	-1	1
3	5	3	1	2	4
1	2	5	4	1	1
				∑d=0	∑d=8

المطلوب:

١- احسب معامل الارتباط بين حجم اسرة المزارع وكفاءته الانتاجية.
٢- وما هو مدلوله ؟

الحـــل

١- ترتيب قيم المتغير x تنازليا ويعطي الرقم الاول رقم ١
٢- ترتيب قيم المتغير y بنفس الطريقة السابقة ايضا يعطى الرقم الاول رقم ١
٣- نحسب الفرق بين رتبة x ورتبة y وتمثل الناتج d
٤- نرتب الفرق بين الرتبتين ثم نجمع هذا التربيع ثم نطبق القانون التالي

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

$$rs = 1 - \frac{6 \times 8}{5(25-1)}$$

$$rs = 1 - \frac{48}{120} = 1 - 0.4 = 0.6$$

١- مدلول معامل الارتباط :

بما أن $r = 0.60$ ، ويدل ذلك على وجود ارتباط طردي متوسط بين حجم اسرة المزارع وكفاءته الانتاجية

المحاضرة الخامسة عشر

اختبارمربع كاي X^2

يستخدم هذا الاختبار الاحصائي لايجاد مدى الاستقلالية بين متغيرين احدهما مستقل(X) والاخر تابع (Y) اي مدى العلاقة الارتباطية بين متغيرين.

مثال

تم توجيه سؤال الى ٢٠ فلاح منتمي الى جمعية تعاونية (١٢) فلاح غير منتمي حول رضا هم عن الخدمات التي يقدمها الارشاد الزرعي وكانت اجاباتهم كما موضحة في الجدول ادناه.

تابع Y / مستقل X	منتمون للجمعية	غير منتمين للجمعية	المجموع
راضين عن العمل الارشادي	١٢	٦	١٨
غير راضين عن العمل الارشادي	٨	٦	١٤
المجموع	٢٠	١٢	٣٢/ المجموع الكلي

الحل:

لحساب قيمة مربع كاي نتبع الخطوات التالية

١- نضع فرضية العدم والفرضية البديلة

HO:لاتوجد علاقة بين المتغيرين

H1:توجد علاقة بين المتغيرين

نحسب التكرار المتوقع لكل خلية ونرمز للتكرار المتوقع ب E

$$Ei = \frac{\text{مجموع العمود} \times \text{مجموع الصف}}{\text{مجموع الكلي}}$$

$$E1 = \frac{18 \times 20}{32} = 11.25$$

$$E2 = \frac{18 \times 12}{32} = 6.75$$

$$E3= \frac{20 \times 14}{32} = 8.75$$

$$E4= \frac{12 \times 14}{32} = 5.25$$

$$X^2 = \sum \frac{(Oi - Ei)^2}{Ei} == \sum \frac{(12-11.25)^2}{11.25} + \frac{(6-6.75)^2}{6.75} + \frac{(8-8.75)^2}{8.75} + \frac{(6-6.25)^2}{6.25} =$$

$$\sum \frac{(0.75)^2}{11.25} + \frac{(-0.75)^2}{6.75} + \frac{(-0.75)^2}{8.75} + \frac{(0.25)^2}{6.25}$$

$$= \mathbf{0.05 + 0.83 + 0.064 + 0.010 = 0.954}$$

نجد قيمة كاي سكوير الجدولية عند مستوى معنوية 0.05 بدرجة حرية = (عدد الصفوف -١)(عدد الاعمدة-١)=١ والتي تساوي *3.84*

نقارن قيمة كاي سكوير مع القيمة الجدولية ،فان قيمة كاي سكوير المحسوبة اقل من الجدولية عند مستوى *0.05* اي ان المتغيرين مستقلين اي لاتوجد علاقة ارتباط بين المتغيرين X , Y

اختبار الإشارة للعينة الواحدة : One sample sign test

يستخدم اختبار الإشارة للعينة الواحدة عادة للاستدلال على وسيط المجتمع خاصة عندما يكون توزيع المجتمع غير متماثل ، وفي حالة تماثل المجتمع فإن الاستدلال على الوسيط هو استدلال على المتوسط. ويشكل الاختبار في هذه الحالة بديلاً لا معلمياً لاختبار t للمتوسط والاختبار الطبيعي للمتوسط في الحالة التي يكون فيها التباين معروفاً.

تلخيص خطوات اختبار الإشارة للعينة الواحدة :

في ضوء ما سبق يمكننا أن نلخص خطوات اختبار الإشارة للعينة الواحدة للاستدلال على وسيط المجتمع m فيما يلي (ترتيب الخطوات اختياري) :

١) التأكد من أن البيانات ناتجة عن عينة عشوائية مستقلة ، وأنها مقاسه بمقياس ترتيبي على الأقل حيث الترتيب بالنسبة لـ m_o ، أو اسمي مثل "فوق" و "تحت" حيث " فوق " تعني أن القيمة أكبر من m_o و "تحت" تعني أنها أصغر من m_o .

٢) تحديد فرض العدم

$$H_o : m = m_o$$

والفرض البديل والذي يأخذ أحد الأشكال

أ- $H_1 : m > m_o$

ب- $H_1 : m < m_o$

ج- $H_1 : m \neq m_o$

لاحظ أنه في حالة أ) و ب) نستبدل علامة المساواة في H_o بـ $\leq$ و $\geq$ بالترتيب ليشمل الفرضان كل البدائل.

ويعتمد الاختيار بين (أ) ، (ب) و (ج) كما ذكرنا على توقعاتنا حول قيمة m بالنسبة لـ m_o في حالة عدم صحه H_o ونستهدي في ذلك بعدد الإشارات الموجبة أو السالبة المشاهدة في العينة. فإذا كنا نتوقع أن تكون m أكبر من m_o وكان عدد الإشارات الموجبة أكبر بكثير من العدد المتوقع $\frac{n}{2}$ بحيث يدعم توقعاتنا، فإننا نختار (أ) ، وإذا كنا نتوقع أن يكون m أقل من m_o وكان عدد الإشارات الموجبة أقل بكثير من $\frac{n}{2}$ نختار (ب) . وفي حالة عدم وجود اتجاه معين متوقع نختار (ج).

٣) نحدد مستوى معنوية معين α مثلاً $\alpha = ٠,٠٥$ أو $\alpha = ٠,٠١$.

٤) نحسب الفروقات $x_i - m_o$ ($i = 1,\dots, n^*$) حيث n^* حجم العينة و $x_1, x_2, \dots, x_{n^*}$ قيم العينة . من الآن فصاعداً نستخدم X_i الكبيرة ليشير للمتغير العشوائي الذي يرمز للقيمة في السحب i و x_i الصغيرة للقيمة المشاهدة فعلا في السحب i .

٥) نستبدل كل فرق بالإشارة "+" إذا كان أكبر من الصفر وبـ "-" إذا كان أقل من الصفر. وإذا كان هناك فرقاً يساوي صفر نقول أن ربطه tie قد حدثت وتحذف المشاهدة المؤدية للربطه ويتم تطبيق اختبار الإشارة على المشاهدات في العينة التي ليس بها ربطات فقط ما لم تكن نسبة الربطات (الأصفار) كبيرة بالنسبة لـ n^* (مثلا اكبر من 5%). في مثل هذه الحالة يمكن – كحل – إعطاء العلامة "+" لنصف الأصفار و "-" للنصف الآخر وتحديد موقع الإشارة عشوائياً باستخدام قطعة نقود متوازنة مثلا.

٦) نحسب عدد الإشارات الموجبة S_+ وعدد الإشارات السالبة S_- . ونلاحظ هنا أن $S_- = n - S_+$ حيث n عدد الإشارات الموجبة والسالبة أي حجم العينة بعد استبعاد الأصفار.

٧) إذا كان الفرض البديل من النوع (أ) بخطوة (٢) نستخدم* S_+ كإحصائية اختبار في اختبار ذو طرف أيمن . ومن جدول (١) نوجد عند n و $S= S_+$ وتحت العمود المعنون " يمين " قيمة p^* - احتمال الحصول على قيمة تساوي أو تزيد عن S_+ إذا كان H_o صحيحاً. ويكون القرار رفض H_o إذا كانت $p^* \le \alpha$ وقيمة p^* في هذه الحالة هي قيمة P المستوى الحرج.

وجدول (١) هو التوزيع التراكمي لتوزيع ذو الحدين بمعالم n و $p = 0.5$ وقيمة p^* المعطاة هي احتمال قيمة تساوي أو تقل عن S عندما تكون $S \le 0.5n$ ونستخدم الطرف الأيسر ، بينما هي احتمال قيمة تساوي أو تزيد عن S عندما تكون $S \ge 0.5n$ ونستخدم الطرف الأيمن . و S يمكن أن تكون S_+ أو S_- . ونتيجة لتماثل توزيع ذو الحدين فإن احتمال قيمة تساوي أو تقل عن قيمة معينة لـ S مثلاً S_* ، يساوي احتمال قيمة تساوي أو تزيد عن $n - S_*$.

٨) إذا كان الفرض البديل من النوع (ب) نستخدم S_- كإحصائية اختبار في اختبار ذو طرف أيمن أيضا (أو S_+ في اختبار ذو طرف أيسر*) ومن جدول (١) نوجد قيمة p^* ونقارنها بمستوى المعنوية كما في خطوة (7) . أيضاً p^* في هذه الحالة هي P .

٩) إذا كان H_1 من النوع "ج" أي الاختبار ذو طرفين فمن الأفضل أن نستخدم الأكبر من بين S_+ و S_- كإحصائية اختبار. فإذا كانت الأكبر هي S_+ ، مثلا ، نوجد من جدول (١) احتمال قيمة تساوي أو تزيد عن S_+ أي p^* . وبما

---

* في هذه الحالة تكون قيمة p هي احتمال قيمة تساوي أو تقل عن التي نحدد قيمتها في العمود " يسار " .

أن الاختبار ذو طرفين فإن P هي احتمال الحصول على قيمة تساوي أو تزيد عن S_+ أو تساوي أو تقل عن n - S_+ لهذا فإن $P = 2p^*$

نرفض H_o إذا كانت $P \le \alpha$ (أو إذا كانت $P^* \le \frac{\alpha}{2}$).

مثال

في مصنع لتعبئة الشاي في أكياس صممت الماكينة بحيث يكون وزن الكيس ٢ جرام. بعد فترة أخذت عينة عشوائية من ١٥ كيساً ، وكانت أوزانها كما يلي :-

1.98, 1.98, 2.10, 1.99, 1.97, 1.98, 2.00, 2.12, 2.09, 1.99, 1.98,1.99, 2.03, 1.96, 1.95

والمطلوب اختبار الفرض بأن وسيط الوزن لازال ٢ مقابل الفرض بأنه تغير. أي اختبار

$$H_o: m = 2$$

مقابل

$$H_1: m \neq 2$$

الحل: هنا $m_o = 2$, $n^* = 15$

أولاً نحسب الفرق بين كل قيمة والوسيط المفترض ٢ ثم نستبدل كل فرق موجب بـ "+" وكل فرق سالب بـ "-" بينما تحذف حالات الصفر. الجدول التالي يوضح ذلك :

جدول ١

X	X-2	الإشارة	X	2–X	الإشارة
١,٩٨	-٠,٠٢	-	٢,٠٩	٠,٠٩	+
١,٩٨	-٠,٠٢	-	١,٩٩	-٠,٠١	-
٢,١٠	٠,١٠	+	١,٩٨	-٠,٠٢	-
١,٩٩	-٠,٠١	-	١,٩٩	-٠,٠١	-
١,٩٧	-٠,٠٣	-	٢,٠٣	٠,٠٣	+
١,٩٨	-٠,٠٢	-	١,٩٦	-٠,٠٤	-
٢,٠٠	٠,٠٠				
٢,١٢	٠,١٢	+	١,٩٥	-٠,٠٥	-

من جدول (٢,١) نجد $S_+ = 4$. وبما أن هناك ربطه واحدة عند $X = 2$ فإننا نحذف هذه القيمة ويصبح حجم العينة $n = 15-1 = 14$

وبما أن عدد الإشارات السالبة أكبر من الموجبة فمن الأنسب أن نستخدم S_- كإحصائية اختبار. ومن جدول (١) وباستخدام $S = 10$ في العمود " يمين " و $n = 14$ نجد أن احتمال الحصول على قيمة أكبر من أو تساوي ١٠ هو ٠,٠٨٩٨ وهو نفسه احتمال الحصول على قيمة تساوي أو تقل عن ٤ (قيمة S_+). وبما أن الاختبار ذو طرفين فأن قيمة P تساوي $2 \times 0.0898 = 0.1796$. وهي قيمة كبيرة إذا استخدمنا مستوى معنوية $\alpha=0.05$ مثلا. لهذا فلا نرفض H_o بأن الوسيط يساوي ٢.

اختبار الوسيط The median test

يستخدم هذا الاختبار لمقارنة وسيطي عينتين مسحوبتين عشوائيا من مجتمعين .

خطوات تطبيق الاختبار :

١- نضع فرضية الاختبار

H_o : جميع المجتمعات لها نفس الوسيط

H_1 : أثنين على الأقل من المجتمعات ليس لها نفس الوسيط

٢- ندمج جميع العينات في مجموعة واحدة ونرتبها تصاعدياً ثم نوجد وسيطها M_g والذي يسمى الوسيط الكلي .

٣- نكون جدولاً ونقوم بحصر وعد القيم التي هي اكبر من الوسيط وكذلك القيم التي تساوي او اقل من الوسيط

٤- نفرغ البيانات في الجدول حسب الارشادات كما يلي

رقم العينة / الوضع بالنسبة للوسيط الكلي	1 2 . . . j . . . k	**المجموع**
(أكبر من الوسيط الكلي)	O_{11} O_{12} . . . O_{1j} . . . O_{1k}	A
(يساوي أو يقل عن الوسيط الكلي)	O_{21} O_{22} . . . O_{2j} . . . O_{2k}	B
المجموع	n_1 n_2 . . . n_j . . . n_k	n

وفي الجدول تمثل A عدد القيم التي تزيد عن الوسيط الكلي في جميع العينات و B عدد القيم التي تساويه أو تقل عنه. أما n فهي عدد المشاهدات من جميع العينات.

٥- نطبق قانون كاي سكوير ونستخرج قيمتها المحسوبة ونقارنها مع قيمة كاي سكوير الجدولية عند مستوى معنوية 0.05, 0.01 وبدرجات حرية ،(عدد الاعمدة-١)(عدد الصفوف-١)=df

٦- اتخاذ القرار: اذا كانت قيمة كاي سكوير المحسوبة اقل من قيمة كاي سكوير الجدولية اذن نقبل فرضية العدم أي لايوجد اختلاف بين وسيطي العينتين ،واذا كانت المحسوبة اكبر من الجدولية نقبل البديلة أي يوجد اختلاف بين وسيطي العينتين

مثال

تم اختيار عينتين عشوائيتين من الطلبة احدهما من مدارس المدينة تتكون من ١٢ طالب والاخرى من مدارس الريف تتكون من ٩ طلاب واجري اختبار لقياس مستوى الصحة عندهم والجدول الاتي يوضح نتائج قياس اوزانهم بالكيلو غرام/المطلوب هل يوجد فرق بين وسيطي المجتمعين الذين تم سحب العينتين منها عند مستوى معنوية 0.05

الحل/

HO:مجتمعان متماثلان في الوسيط

H1:مجتمعان غير متماثلان في الوسيط

العينة الاولى=33,33,33,35,38,40,42,45,48,49,40,43

العينة الثانية =29,30,31,31,35,39,42,42,45

نرتب البيانات تصاعديا بعد دمج العينتين

29,30,31,31,33,33,33,35,35,38,39,40,40,42,42,43,45,48,49

نستخرج الوسيط في حالة الفردي $= \frac{n}{2}+1$ =21/2+1=11

(اذن ناخذ الرقم الذي يتوسط الرتبة١١ =٣٩)

ملاحظة مهمة (الوسيط في حالة الفردي= $\frac{n}{2}+1$ ،واذا كان عدد القيم زوجي يمكن أن نأخذ الوسيط الكلي على أنه متوسط القيمتين اللتين في الوسط أي القيمة ذات الرتبة $\frac{n}{2}$ و $\frac{n}{2}+1$)

رقم العينة / العلاقة مع الوسيط	عينة ١	عينة ٢	المجموع
اكبر من الوسيط	7	3	10
اقل او يساوي الوسيط	5	6	11
المجموع	12	9	21

نطبق قانون كاي سكوير :

$$X^2 = \sum \frac{(Oi - Ei)^2}{Ei}$$

$$Ei = \frac{\text{مجموع العمود} \times \text{مجموع الصف}}{\text{مجموع الكلي}}$$

X^2= قيمة مربع كآي المحسوبة

Oi = القيم أو البيانات الملاحظة.

Ei =القيم أو البيانات المتوقعة.

$$E1 = \frac{10 \times 12}{21} = 5.7$$

$$E2 = \frac{10 \times 9}{21} = 4.3$$

$$E3 = \frac{11 \times 12}{21} = 6.3$$

$$E4 = \frac{11 \times 9}{21} = 4.7$$

$$X^2 = \sum \frac{(Oi - Ei)^2}{Ei} = \sum \frac{(7-5.7)^2}{5.7} + \frac{(3-4.3)^2}{4.3} + \frac{(5-6.3)^2}{6.3} + \frac{(6-4.7)^2}{4.7} = 1.316$$

كاي سكوير المحسوبة

<u>القرار</u>

الجدولية = 3.84عند مستوى معنوية 0.05 ودرجات حرية ١

القرار : بما ان قيمة كاي سكوير الحسوبة اقل من الجدولية اذن نقبل فرضية العدم اي ان المجتمعان متماثلان في الوسيط اي لايوجد اختلاف بين وسيطي العينتين.

جدول كا٢

درجة الحرية	مستوى الدلالة أو الثقة او المعنوية		
	0.05	0.01	0.001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.13
9	16.92	21.67	27.88
10	18.31	23.21	29.59
11	19.68	24.73	31.26
12	21.03	26.22	32.91
13	22.36	27.69	34.53
14	23.69	29.14	36.12
15	25.00	30.58	37.70
16	26.30	32.00	39.25
17	27.59	33.41	40.79
18	28.87	34.81	42.31
19	30.14	36.19	43.82
20	31.41	37.57	45.32
21	32.67	38.93	46.80
22	33.92	40.29	48.27
23	35.17	41.64	49.73
24	36.42	42.98	51.18
25	37.65	44.31	52.62
26	38.89	45.64	54.05
27	40.11	46.96	55.48
28	41.34	48.28	56.89
29	42.56	49.59	58.30
30	43.77	50.89	59.70
31	44.99	52.19	61.10
32	46.19	53.49	62.49
33	47.40	54.78	63.87
34	48.60	56.06	65.25
35	49.80	57.34	66.62
36	51.00	58.62	67.99
37	52.19	59.89	69.35
38	53.38	61.16	70.71
39	54.57	62.43	72.06
40	55.76	63.69	73.41
41	56.94	64.95	74.75
42	58.12	66.21	76.09
43	59.30	67.46	77.42

44	60.48	68.71	78.75
45	61.66	69.96	80.08
46	62.83	71.20	81.40
47	64.00	72.44	82.72
48	65.17	73.68	84.03
49	66.34	74.92	85.35
50	67.51	76.15	86.66
51	68.67	77.39	87.97
52	69.83	78.62	89.27
53	70.99	79.84	90.57
54	72.15	81.07	91.88
55	73.31	82.29	93.17
56	74.47	83.52	94.47
57	75.62	84.73	95.75
58	76.78	85.95	97.03
59	77.93	87.17	98.34
60	79.08	88.38	99.62
61	80.23	89.59	100.88
62	81.38	90.80	102.15
63	82.53	92.01	103.46
64	83.68	93.22	104.72
65	84.82	94.42	105.97
66	85.97	95.63	107.26
67	87.11	96.83	108.54
68	88.25	98.03	109.79
69	89.39	99.23	111.06
70	90.53	100.42	112.31
71	91.67	101.62	113.56
72	92.81	102.82	114.84
73	93.95	104.01	116.08
74	95.08	105.20	117.35
75	96.22	106.39	118.60
76	97.35	107.58	119.85
77	98.49	108.77	121.11
78	99.62	109.96	122.36
79	100.75	111.15	123.60
80	101.88	112.33	124.84
81	103.01	113.51	126.09
82	104.14	114.70	127.33
83	105.27	115.88	128.57
84	106.40	117.06	129.80
85	107.52	118.24	131.04
86	108.65	119.41	132.28
87	109.77	120.59	133.51
88	110.90	121.77	134.74
89	112.02	122.94	135.96
90	113.15	124.12	137.19

91	114.27	125.29	138.45
92	115.39	126.46	139.66
93	116.51	127.63	140.90
94	117.63	128.80	142.12
95	118.75	129.97	143.32
96	119.87	131.14	144.55
97	120.99	132.31	145.78
98	122.11	133.47	146.99
99	123.23	134.64	148.21
100	124.34	135.81	149

جدول ت-t

درجة الحرية	مستوى الدلالة							
طرف واحد	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001	0.00005
طرفين	0.2	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001
2	1.89	2.92	4.30	9.92	14.09	31.60	44.70	100.14
3	1.64	2.35	3.18	5.84	7.45	12.92	16.33	28.01
4	1.53	2.13	2.78	4.60	5.60	8.61	10.31	15.53
5	1.48	2.02	2.57	4.03	4.77	6.87	7.98	11.18
6	1.44	1.94	2.45	3.71	4.32	5.96	6.79	9.08
7	1.41	1.89	2.36	3.50	4.03	5.41	6.08	7.89
8	1.40	1.86	2.31	3.36	3.83	5.04	5.62	7.12
9	1.38	1.83	2.26	3.25	3.69	4.78	5.29	6.59
10	1.37	1.81	2.23	3.17	3.58	4.59	5.05	6.21
11	1.36	1.80	2.20	3.11	3.50	4.44	4.86	5.92
12	1.36	1.78	2.18	3.05	3.43	4.32	4.72	5.70
13	1.35	1.77	2.16	3.01	3.37	4.22	4.60	5.51
14	1.35	1.76	2.14	2.98	3.33	4.14	4.50	5.36
15	1.34	1.75	2.13	2.95	3.29	4.07	4.42	5.24
16	1.34	1.75	2.12	2.92	3.25	4.01	4.35	5.13
17	1.33	1.74	2.11	2.90	3.22	3.97	4.29	5.04
18	1.33	1.73	2.10	2.88	3.20	3.92	4.23	4.97
19	1.33	1.73	2.09	2.86	3.17	3.88	4.19	4.90
20	1.33	1.72	2.09	2.85	3.15	3.85	4.15	4.84
21	1.32	1.72	2.08	2.83	3.14	3.82	4.11	4.78
22	1.32	1.72	2.07	2.82	3.12	3.79	4.08	4.74
23	1.32	1.71	2.07	2.81	3.10	3.77	4.05	4.69
24	1.32	1.71	2.06	2.80	3.09	3.75	4.02	4.65
25	1.32	1.71	2.06	2.79	3.08	3.73	4.00	4.62
26	1.31	1.71	2.06	2.78	3.07	3.71	3.97	4.59
27	1.31	1.70	2.05	2.77	3.06	3.69	3.95	4.56
28	1.31	1.70	2.05	2.76	3.05	3.67	3.93	4.53
29	1.31	1.70	2.05	2.76	3.04	3.66	3.92	4.51
30	1.31	1.70	2.04	2.75	3.03	3.65	3.90	4.48
35	1.31	1.69	2.03	2.72	3.00	3.59	3.84	4.39

40	1.30	1.68	2.02	2.70	2.97	3.55	3.79	4.32
45	1.30	1.68	2.01	2.69	2.95	3.52	3.75	4.27
50	1.30	1.68	2.01	2.68	2.94	3.50	3.72	4.23
55	1.30	1.67	2.00	2.67	2.92	3.48	3.70	4.20
60	1.30	1.67	2.00	2.66	2.91	3.46	3.68	4.17
65	1.29	1.67	2.00	2.65	2.91	3.45	3.66	4.15
70	1.29	1.67	1.99	2.65	2.90	3.43	3.65	4.13
75	1.29	1.67	1.99	2.64	2.89	3.42	3.64	4.11
80	1.29	1.66	1.99	2.64	2.89	3.42	3.63	4.10
85	1.29	1.66	1.99	2.63	2.88	3.41	3.62	4.08
90	1.29	1.66	1.99	2.63	2.88	3.40	3.61	4.07
95	1.29	1.66	1.99	2.63	2.87	3.40	3.60	4.06
100	1.29	1.66	1.98	2.63	2.87	3.39	3.60	4.05
200	1.29	1.65	1.97	2.60	2.84	3.34	3.54	3.97
500	1.28	1.65	1.96	2.59	2.82	3.31	3.50	3.92
1000	1.28	1.65	1.96	2.58	2.81	3.30	3.49	3.91
∞	1.28	1.64	1.96	2.58	2.81	3.29	3.48	3.89

جدول ف

درجة حرية التباين الصغير للخطا التجريبي	درجة حرية التباين الكبير للمعاملات								
	1	2	3	4	5	6	8	12	مالانهاية
1	161	200	216	225	230	234	239	244	254
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.8	8.7	8.5
4	7.7	6.9	6.6	6.4	6.3	6.2	6.0	5.9	5.6
5	6.6	5.8	5.4	5.2	5.1	5.0	4.8	4.7	4.4
6	6.0	5.1	4.8	4.5	4.4	4.3	4.2	4.0	3.7
7	5.6	4.7	4.4	4.1	4.0	3.9	3.7	3.6	3.2
8	5.3	4.5	4.1	3.8	3.7	3.6	3.4	3.3	2.9
9	5.1	4.3	3.9	3.6	3.5	3.4	3.2	3.1	2.7
10	5.0	4.1	3.7	3.5	3.3	3.2	3.1	2.9	2.5
11	4.8	4.0	3.6	3.4	3.2	3.1	3.0	2.8	2.4
12	4.8	3.9	3.5	3.3	3.1	3.0	2.9	2.7	2.3
13	4.7	3.8	3.4	3.2	3.0	2.9	2.8	2.6	2.2
14	4.6	3.7	3.3	3.1	3.0	2.9	2.7	2.5	2.1
15	4.5	3,7	3.3	3.1	2.9	2.8	2.6	2.5	2.1
16	4.5	3.6	3.2	3.0	2.9	2.7	2.6	2.4	2.0
17	4.5	3.6	3.2	3.0	2.8	2.7	2.6	2.4	2.0
18	4.4	3.6	3.2	2.9	2.8	2.7	2.5	2.3	1.9
19	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.9
20	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.8
21	4.3	3.5	3.1	2.8	2.7	2.6	2.4	2.3	1.8
22	4.3	3.4	3.1	2.8	2.7	2.6	2.4	2.2	1.8
23	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.8
24	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.7
25	4.2	3.4	3.0	2.8	2.6	2.5	2.3	2.2	1.7
26	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.2	1.7
27	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.1	1.7
28	4.2	3.3	3.0	2.7	2.6	2.4	2.3	2.1	1.7
29	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6
30	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6
40	4.1	3.2	2.8	2.6	2.5	2.3	2.2	2.0	1.5
60	4.0	3.2	2.8	2.5	2.4	2.3	2.1	1.9	1.4
120	3.9	3.1	2.7	2.5	2.3	2.2	2.0	1.8	1.3
مالا نهاية	3.8	3.0	2.6	2.4	2.2	2.1	1.9	1.8	1.0

المراجع

١- القصاص، د.مهدي محمد، ٢٠٠٧مبادئ الاحصاء والقياس الاجتماعي،كلية الاداب،جامعة المنصورة،مصر.

٢- حسن،د.خالد حامد، ٢٠١١محاضرات مبادئ علم الاحصاء،كلية الزراعة،جامعة ديالى علي،د.هيثم عبد السلام ود.ميسر محمد عزيز،الاحصاء،كتاب قيد التاليف.

٣- صيني،د.سعيد اسماعيل٢٠١٠،قواعد اساسية في البحث العلمي،المدينة المنورة،الطبعة الثانية.

٤- الجوني،ابو جنى،٢٠١٤،ملخص التحليل الاحصائي،محاضرات ادارة اعمال ،جامعة الملك فيصل

تم بحمد الله وعونه