الإحصاء

اعـداد

الدكتور

محمد كامل ريحــان

أستاذ الإحصاء والاقتصاد القياسي

كلية الزراعة – جامعة عين شمس

الدكتور

السعيد عبد الحميد البسيوني

أستاذ الإحصاء والاقتصاد القياسي

كلية الزراعة – جامعة عين شمس

وعميد المعهد العالى للتعاون الزراعى

مقدمة

في عصر العولمة وثورة تكنولوجيا المعلومات التي يعيشها العالم اليوم وبسبب ضخامة البيانات والمعلومات المتوفرة فإن الأمر يتطلب منا العمل الجاد لاستغلال تلك المعلومات في عملية تنظيم وتحليل لتلك المعلومات للوصول منها الي القرارات السليمة أو على الأقل التنبؤ واستخدام هذه التنبؤات في وضع إستراتيجيات للعمل خلال فترات زمنية لاحقة أو مستقبلية، ومما لاشك فيه فنحن في حاجة إلى مهارات خاصة للتعامل مع هذه المعلومات وهو ما يعرف بعلم الإحصاء فهو ذاك العلم الذي يبحث في كيفية جمع البيانات لظاهرة ما ومن ثم تصنيفها في جداول تمثلها بيانياً ومن ثم استخلاص النتائج التي لاتخاذ القرار المناسب، كما أن علم الإحصاء هو مجموعة من الأساليب والعمليات الخاصة لمعالجة البيانات المتوفرة سواء الكمية منها أو النوعية.

ومدلول كلمة إحصاء (Statistics) يراه البعض مجرد بيانات وأرقام بينما يراه آخرون طرق لجمع البيانات ووصفها وعرضها بصورة مبسطة في حين يراه البعض الآخر وسيلة لاتخاذ القرار حال عدم التأكد.

وينقسم علم الإحصاء الي قسمين أساسيين هما: الإحصاء الوصفي Descriptive Statistics والإحصاء الاستدلاليInferential Statistics

ويشير مفهوم الإحصاء الي مجموعة الطرق العلمية القياسية التي تهدف لجمع البيانات عن ظاهرة ما ومن ثم جدولتها وعرضها بصورة مختصرة وتقييمها للوصول لنتائج من خلال عينة صغيرة من مجتمع للتعرف على خصائص المجتمع بصورة تقريبية لخصائص العينة.

وتنقسم البيانات الاحصائية الي نوعين هما بيانات نوعية وبيانات كمية حيث تصف البيانات النوعية ظاهرة ما بصورة غير رقمية كالجنس (ذكر وأنثى) والتقدير (ممتاز و جيد و ...) .

أما البيانات الكمية (الرقمية) فتعرف بالبيانات المقيسة مثل الكيلوجرام للوزن والمتر للطول والجنيه للسعر، وهي تعبر عن ظاهرة وخاصية ما في المجتمع بصورة رقمية كإنتاج القطن بالطن.

ولقد أصبح استخدام الحاسب الآلي في الإحصاء أمراً ملحاً وناجحاً بدرجة كبيرة لتوفير الوقت ودقة النتائج ، وهناك الكثير من برامج الإحصاء المتخصصة في هذا المجال أهمها EXCEL ، SPSS, MINITAB فباستخدام أياً منهم يمكن الحصول على نتائج التحليل وغيره في وقت وجيز علاوة على أن الحاسب الآلي يمدنا بأمور أخرى كثيرة تساعدنا في عملية جمع البيانات أو البحث في شبكة الإنترنت مما يفيدنا في مجال البحوث.

ويحاول هذا الكتاب ان يقدم لطلبة الاقتصاد والعلوم الإدارية بصفة عامة وطلبة برنامج التعليم المفتوح بصفة خاصة توليفة متكاملة من أهم المفاهيم

الاحصـائية والمقـاييس الوصـفية والكميـة ، متضـمنة فـي ذلـك كـل مـن الاحصـاء الوصفي والاحصـاء الاسـتدلالي وطرق التقدير الاحصـائي، مـع أعطـاء العديد من الامثلـة التـي توضـح كيفيـة اسـتخدام هـذه المقـاييس وتلـك الطـرق والنظريـات الاحصائية ، ومناقشة المشاكل المتعلقة بها.

ولتحقيق الهدف الأساسي لهذا الكتاب فقد أحتوي علـي مجموعـة متكاملـة مـن المقـاييس الوصـفية ،متضـمنة مقـاييس النزعـة المركزيـة ومقـاييس التشـتتت والتشتت النسبي ، كما يحتوي الكتاب علي طرق قياس نماذج الانحدار الخطية وغير الخطية ، البسيطة منها والمتعددة.

كمـا يتعـرض الكتـاب لنظريـة الاحتمـالات وبعـض التوزيعـات الاحتماليـة المنفصلة والمتصلة ، كما يناقش الكتاب بعض الموضعات الاحصائية الهامة مثل تحليل التباين ومربع كا ، بالاضافة الي الأرقام الاقياسية ونظرية العينات.

وينتهي كل باب من ابواب الكتاب بملخص مختصر لاهم النقاط التي يعالجها تحت مسمي تذكر ، فضلاً عن عدد من الاسئلة والتمارين التي تمكن الطالب من أختبار مدي فهمه للمعلومات والأفكار التي وردت به.

ويشتمل هذا الكتاب علي ثمانية ابواب كانت موضوعاتها بالترتيب كما يلي:

الباب الأول: ـماهية علم الاحصاء ونشأته وتطوره.

الباب الثانى : المقاييس الإحصائية الوصفية.

الباب الثالث : أساسيات نظرية الاحتمالات والتوزيعات الاحتمالية.

الباب الرابع: تحليل الانحدار والارتباط .

الباب الخامس : العينات وتصميم التجارب.

الباب السادس : تحليل التباين .

الباب السابع : كا التربيعية.

الباب الثامن : الأرقام القياسية.

بالاضافة الي اهم المراجع التي تم الاستعانة بها لاخراج هذا الكتاب في هذا الشكل.

والله نسأل أن يكون هذا الكتاب مساهمة فعالة في طريق استخدام الأدوات الإحصائية لخدمة الدارسين الاقتصاديين بصفة عامة وفى مجال تكنولوجيا وإدارة المشروعات ببرنامج التعليم المفتوح بصفة خاصة .

والله الموفق،،،،،،

المؤلفان

أ.د / محمد كامل ابراهيم ريحان

أ.د / السعيد عبد الحميد البسيونى

المحتـــويـــات

الموضوع	الصفحة
المقدمة	
الباب الأول : ماهية علم الاحصاء ونشأته وتطوره	1
نشأة وتطور علم الإحصاء	2
أهمية استخدام الأساليب الإحصائية	3
الوظائف الأساسية للإحصاء	5
الإحصاء وعلاقته بالعلوم الأخرى	12
الباب الثاني :المقاييس الإحصائية الوصفية	24
أولاً: مقاييس النزعة المركزية(المتوسطات)	24
ثانياً: مقاييس التشتت والتشتت النسبي	40
تذكر	49
تمارين	52
الباب الثالث :أساسيات نظرية الاحتمالات والتوزيعات الاحتمالية	57
بعض المفاهيم الهامة فى الاحتمالات	57
التوزيعات الإحتمالية	70
أولاً: التوزيعات الإحتمالية المنفصلة	70
ثانيا : التوزيعات الاحتمالية المتصلة	77
تذكر	84
تمارين	87

الباب الأول
ماهية علم الاحصاء ونشأته وتطوره

يمكن تعريف الإحصاء لغوياً بأنه الإلمام بكل المفردات التي يشملها المجتمع المطلوب دراسته ودراسة كافة مفرداته رقمياً، ومما يؤكد هذا المفهوم هو ما ورد فى القرآن الكريم فى مواضيع كثيرة منها الآيات التالية:

" وكل شئ أحصيناه كتابا"

" وأحصينا كل شئ عددا"

ومما لا شك فيه أن تلك الآيات تؤكد المعنى الشمولي للإحصاء ولا تعطى مجالا للنسيان أو التحريف أو التشكيك.

والإحصاء علميا هو عبارة عن تصوير رقمي للواقع فى المجتمعات المطلوب دراستها سواء كانت بشرية أو غير بشرية عن طريق وصف مفردات المجتمع وصفا رقميا. والمقصود هنا بالواقع هو ما حدث بالفعل وقت أجراء عملية الإحصاء وليس ما يجب أن يكون أو الذي كان أو المطلوب أو المرغوب أو المنظور أن يكون فى المستقبل القريب أو البعيد.

هذا ويعرف علم الإحصاء بأنه العلم الذي يشمل "كل الطرق المستخدمة فى جمع وعرض وتبويب وتحليل البيانات الخاصة بظواهر محددة بغرض استخدامها فى اتخاذ القرارات وذلك فى وجود عنصر عدم التأكد.

وعلى ذلك فان عمل الإحصائى يتركز فى جمع البيانات وتقدير المقاييس أو العلاقات بينها واختيار هذه العلاقات للحكم على مدى دقتها فى تمثيل العلاقة الحقيقية بين الظواهر، أي اختيار ما إذا كانت العلاقات المفترضة نظريا تتفق مع البيانات الخاصة بالواقع العملي.

وعموما ينقسم علم الإحصاء إلى قسمين رئيسين، الأول يعرف بالإحصاء الوصفي ويختص بجمع وعرض البيانات وذلك بهدف توفير المعلومات عن الاتجاهات المختلفة للظواهر، وقد يصاحب هذا الاتجاه استخراج لبعض المؤشرات الإحصائية كالمتوسطات والتى تهدف إلى توضيح الطبيعة العامة للظواهر موضع الدراسة، أما القسم الثاني فيعرف بالإحصاء الرياضي ويستند هذا القسم على الكثير من المفاهيم والنظريات الرياضية والتى من أهمها نظرية الاحتمالات، كما يختص بمحاولة اتخاذ القرارات المناسبة فى المشاكل موضع الدراسة. وعلى سبيل المثال فانه يمكن القول أن الإحصائي يقوم بتقدير خواص المجتمع من واقع خواص العينة، أو تقدير سلوك ظاهرة ما فى المستقبل من واقع معرفة سلوكها فى

السنوات الماضية كأن يكون المطلوب مثلا تقدير متوسط عمر سكان مدينه القاهرة (المجتمع) وذلك بدراسة توزيع الأعمار لعينة مكونة من 1000 فرد سكان المدينة (العينة) أو قد يكون المطلوب تقدير حجم الناتج القومي أو السكان للسنوات القادمة من واقع البيانات المتوفرة عن الناتج القومي والسكان فى السنوات السابقة.

ومما سبق فإن الإحصاء هو العلم الذى يبحث فى:

(أ)جمع البيانات والمعلومات والحقائق الخاصة بمختلف الظواهر وتسجيلها فى صورة رقمية وتصنيفها فى جداول منظمة وتمثيلها بيانيا.

(ب) تحليل البيانات واستخلاص النتائج منها واتخاذ القرارات المناسبة

(ج) مقارنة الظواهر ببعضها ودراسة العلاقات بينها واستخدامها فى تفهم حقيقة الظواهر ومعرفة القوانين التى تسير تبعا لها.

ولا نريد الخوض فى ما إذا كان الإحصاء علما فقط أو الى جانب ذلك فنا أو طريقة أو أسلوب أو أداه بل لعله كل ذلك جميعا، فهو للمتخصصين قاعدة أساسية لانطلاق الفكر عبر النظريات الرياضية والتطبيقات الاقتصادية والاجتماعية والعلمية بحثا عن الحقيقة بصورة علمية منظمة .وهو لغير المتخصصين أداة تنسيق وتنظيم وتحليل بهدف استكشاف الحقيقة بطريقة علمية منظمة.

نشأة وتطور علم الإحصاء

استمر الإنسان في الاعتماد على تأملاته فترة طويلة في سبيل البحث عن الحقائق المحيطة به وكانت هذه التأملات الأساس الذي مهد الطريق إلى البحث العلمي ، حيث أنتقل الإنسان من بحثه عن طريق التأمل بالاستناد على منهاج الملاحظة ثم بدأ بالاعتماد على التجربة في العمل كمنهاج لبحثه عن الحقيقة إلى أن استطاع أن يتوصل إلى منهاج آخر يستعين به في الكشف عن الحقائق ذات العلاقة بالإنسان سواء كانت متعلقة بالنواحي الاجتماعية أو الاقتصادية والذي تمثل في انتهاج الأسلوب العلمي الإحصائي ، حيث تطور علم الإحصاء وتطبيقاته عبر سنوات طويلة بجهود ومشاركة كثير من العلماء من كافة أنحاء العالم العاملين في حقول وميادين مختلفة.

وتشير كثير من الدلائل على الاهتمام بالإحصاء واستخدامه منذ زمن بعيد (العصور القديمة) حيث أقتصر اهتمام الحكومات منذ القدم بالمعلومات الاجتماعية وذلك لأغراض التنظيم والتخطيط ، واستخدم الإحصاء في عصره الأول في جمع البيانات عن السكان وحصرهم من قبل الدولة لأهداف معينة تتمثل في استخدامهم

في الجيوش أو توجيههم لتنفيذ بعض المباني أو لغرض فرض الضرائب أو لتوزيع الأراضي الزراعية على السكان بطريقة عادلة ، ويعد قدماء المصريين أول من أستخدم هذا الأسلوب . وفي القرن السابع عشر والذي يمكن اعتباره العصر الإحصائي الثاني تم استخدام الطريقة الرقمية للدلالة على الظواهر موضوع البحث على اعتبار أن هذه الطريقة أدق وأقوى في التعبير عن هذه الظواهر وتركز الهدف من هذه الطريقة في معرفة عدد السكان وعدد المواليد وعدد الوفيات ومقدار الثروة والدخل ومقدار الضرائب المحصلة وكمية الناتج من المحاصيل الزراعية.

وباختصار نجد أن مجال الإحصاء قبل القرن العشرين كان مرتبطاً في الغالب بالمجالات الاقتصادية والاجتماعية المتمثلة بتعداد السكان ومعرفة خصائصهم الاجتماعية والاقتصادية ، وكانت الأساليب الإحصائية المستخدمة تمتاز بالبساطة بحيث لم توفر للإحصاء الأسس والمقومات الكافية ليصبح علماً.

ويمكن تحديد بداية العصر الإحصائي الثالث مع تطور علوم الرياضيات في القرن الثامن عشر وظهور بعض النظريات العلمية الهامة مثل نظرية الاحتمالات التي كان لها الدور الكبير في تطور هذا العلم واكتسابه أهمية كبرى بحيث أصبح علماً مستقلاً وانتشر استخدامه وبدأ الاهتمام من قبل العلماء في تطبيق النظريات والطرق والأساليب الإحصائية في الكثير من فروع العلم الحديث كالهندسة والطب والصيدلة والزراعة والصناعة والجغرافيا والفلك وعلم النفس باعتباره الطريقة الصحيحة والأسلوب الأمثل إتباعه في البحث العلمي.

وأخيراً فقد أدى ظهور الحاسبات الآلية وتطورها في وقتنا الحالي بأنواعها المختلفة وبقدرتها الفائقة ودقتها المتناهية إلى تمهيد الطريق لاستخدام وتطبيق الأساليب الإحصائية المختلفة في شتى المجالات والميادين.

أهمية استخدام الأساليب الإحصائية

يعد استخدام الأسلوب الإحصائي في أي دراسة الوسيلة المأمونة التي يمكن أن تضمن تحقيق الأهداف المرجوة من وراء تنفيذها سواء كان الهدف المقصود من الدراسة التعرف على نواحي معينة لبعض الظواهر الاجتماعية أو الاقتصادية أو لدراسة مشكلة معينة قائمة أو متوقعة ووضع الحلول المناسبة لها ويمكن للمنشآت سواء التابع منها للقطاع العام أو الخاص القيام بالأعمال والمهام والواجبات المنوطة بها على الوجه المطلوب إذا ما توافرت لها المعلومات والبيانات والمؤشرات الإحصائية وعلى درجة من الدقة والشمول ، فعلى سبيل المثال يمكن للمؤسسات العاملة في قطاع الخدمات الأمنية توزيع خدماتها على جميع نواحي

الدولة بشكل مناسب استنادا إلى البيانات المتوفرة عن التوزيع الجغرافي للسكان في هذه المناطق وطبيعتها الجغرافية ، كما ويمكن للقائمين على قطاع الخدمات التعليمية تلمس احتياجات المجتمع من المؤسسات التعليمية واحتياجاتها من المباني التعليمية والمدرسين والإدارات المدرسية في ضوء توفر بيانات ومعلومات مفصلة ودقيقة عن السكان وتوزيعهم العمري والنوعي ، كما أن التخطيط لإقامة مشاريع صناعية كانت أو تجارية تستلزم بالضرورة توفر بيانات عن مقومات قيام مثل هذه المشاريع ودراسة الجدوى الاقتصادية المأمولة من وراء إنشائها .

إن الأخذ بأساليب التخطيط التنموي ورسم السياسات التنموية لكل دولة يتطلب توفر بيانات ومعلومات ومؤشرات إحصائية مع ضمان دقتها وشمولها من أجل بلوغ الأهداف المرجوة من التخطيط وتمكين القائمين على التخطيط من متابعة تنفيذ جميع مراحل الخطط المرسومة والتأكد من سير هذه المراحل على الوجه المطلوب .ومن المعروف بأن استخدام الأساليب الإحصائية أصبح من الأعمدة الأساسية التي يركن إليها في التوصل للحلول المناسبة لكثير من المشاكل والقضايا التي تهم المجتمع كقضايا الصحة والتعليم والزراعة والصناعة والتجارة.

مما سبق يتضح بأن أهمية علم الإحصاء تكمن في أنه استطاع في الآونة الأخيرة أن يضع أساليبه العلمية ونظرياته موضع التطبيق بالإضافة إلى أهميته النظرية وفوائده التطبيقية الواسعة ، ويعكس ذلك الاتجاه الحديث للإحصاء واستخدامه بواسطة المنشآت على اختلاف أنواعها وأنشطتها في سبيل الوصول إلى قرارات حكيمة وبحيث أصبح من الممكن القول بأن الأساليب الإحصائية تستخدم غالباً في كل الدراسات والبحوث العلمية . ففي قطاع التجارة زاد الاهتمام باستخدام الأساليب الإحصائية لرسم سياسة المنشآت العاملة في هذا المجال في جميع عملياتها المختلفة بشكل يمكنها من اتخاذ قراراتها التجارية السليمة على أسس علمية ومراقبة عملياتها التجارية ورسم الخطط لعملياتها المستقبلية ، وبشكل عام يعتمد الاقتصاديون في وقتنا الحاضر اعتماداً كبيراً في رسم السياسات الاقتصادية على الأساليب الإحصائية من خلال دراستهم لعدد من المى لا لاى واضيع ذات العلاقة الوطيدة بالاقتصاد كإحصاءات الدخل القومي والإنفاق الاستهلاكي والتجارة الداخلية والخارجية والإنتاج الصناعي والزراعي والأرقام القياسية لأسعار السلع والخدمات وتكاليف المعيشة والإحصاءات المتعلقة بالبنوك والاستثمارات والمدخرات وإحصاءات القوى العاملة والإحصاءات السكانية والحيوية .

الوظائف الأساسية للإحصاء :

يتضمن علم الإحصاء الأسلوب العلمي اللازم لتقصي حقائق الظواهر واستخلاص النتائج عنها ، كما يتضمن أيضا النظرية اللازمة للقياس واتخاذ القرارات في كافة الميادين الاقتصادية والاجتماعية والسياسية وهو بذلك يعطي للباحثين والدارسين في تلك المجالات أدق أداة للبحث العلمي المبني على الأسلوب والنظرية ، ولعلم الإحصاء وظائف متعددة يمكن من خلالها استخلاص الكثير من الحقائق والنتائج الهامة والضرورية لوضع ورسم الخطط التنموية ، ومن هذه الوظائف ما يلي :

1 . وظيفة العد (الحصر) :

تعتبر وظيفة العد أو الحصر من أساسيات العمل الإحصائي بصرف النظر عن تطورات هذه الوظيفة في حد ذاتها ، فلقد بدأت انطلاقة العمل الإحصائي لعلم الإحصاء من هذه الوظيفة وعرف من خلالها وأرتبط بها ارتباطاً قوياً في الحقب القديمة من التاريخ ، ووصلت قوة هذا الارتباط إلى الدرجة التي عرف بها علم الإحصاء على أساس أنه علم العد أو الحصر أو التعدادات لقيم الظواهر المختلفة المحيطة والمؤثرة في النشاط اليومي للإنسان .

ولقد ظلت وظيفة عد الأشياء فترة طويلة من حقب التاريخ السابقة مسخرة لخدمة أهداف خاصة بالدولة ، وانحصرت الوظيفة في إطار هذه الأهداف الخاصة مما حد ذلك من التطور الوظيفي لعلم الإحصاء وأدى إلى تأخر ظهور الأساليب والنظريات الإحصائية في فترة مبكرة مثل باقي العلوم . فلقد انحصرت وظيفة حصر الأشياء في معرفة عدد الرجال لأي دولة مع مقارنة ذلك بما هو موجود في الدولة ممثلة في جيشها مما يساعد في اتخاذ قرارات الحروب ، كما استخدمت هذه الوظيفة في تحديد ما لدى الدولة من أموال حتى يكون ذلك مرشداً عند وضع السياسة الضريبية الحاضرة والمقبلة ، وإلى جانب ذلك فلقد عرفت التعدادات التجارية والزراعية والصناعية في صورة عامة إجمالية لغرض حصر الموارد الاقتصادية للبلاد ومقارنة ذلك بما هو موجود في الدول الأخرى .

غير أن التقدم التقني والذي فرض نفسه فجأة في جميع مجالات حياتنا اليومية كان له تأثيره في تغيير وجهة النظر الكلاسيكية تجاه وظيفة العد والإحصاء . فلم تعد عمليات التعدادات سواء ، كانت عن النواحي الديموغرافية أو الزراعية أو التجارية أو الصناعية ، عبارة عن عملية حصر إجمالي للأشياء وقيم الظواهر ، بل أصبحت هذه الوظيفة تعطي لنا المزيد من البيانات والمعلومات التفصيلية في

كل المجالات بأسلوب يخدم أغراض التخطيط والتنمية الاقتصادية للبلاد من خلال أسلوب يعتمد على النظريات الإحصائية في تفسير الاتجاهات وتحليل التغيرات وتفسير العلاقات بين المتغيرات وإيضاح أسبابها . زيادة على ذلك فإن تطور هذه الوظيفة كان من شأنه اقتحام ميادين جديدة لم تكن موجودة من قبل ، حيث لم تعد وظيفة الحصر قاصرة على تعداد السكان أو التعداد الزراعي أو التعداد الاقتصادي فحسب بل أصبح يوجد الآن إحصاءات خاصة بالقوى العاملة وإحصاءات تفصيلية للتجارة الخارجية وإحصاءات مالية ونقدية وإحصاءات المواصلات وإحصاءات الدخل وغير ذلك لما هو ضروري وأساسي في عملية التقدم والرقي .

2 . وظيفة جمع البيانات :

ثاني وظائف العمل الإحصائي ، يقدمه لنا الأسلوب الإحصائي لجمع البيانات عن مختلف الظواهر المحيطة بنا ، هذه الوظيفة لها وجود يمتد إلى فترة طويلة سابقة منذ الوقت الذي كان يعرف فيه العلم على أساس أنه علم جمع البيانات والحقائق وتستمد هذه الوظيفة أهميتها من خلال ضرورة توافر البيانات عن الظواهر والعوامل المحددة لها ، والمعلومات عن الظواهر موضع البحث بحيث يمكن دراسة وتحليل واستخلاص النتائج واتخاذ القرارات . فإذا ما أتبع أسلوب غير علمي وغير موضوعي في جمع البيانات وبطريقة غير دقيقة أدى ذلك إلى الحصول على حقائق عن الأشياء غير سليمة متحيزة وكان ذلك مصدراً في إفساد النتائج واتخاذ قرارات لها خطورتها وغير مأمونة العواقب والعكس صحيح إذا ما أتبع أسلوب علمي موضوعي غير متحيز في جمع البيانات أدى ذلك إلى الحصول على حقائق عن الظواهر بطريقة سليمة غير متحيزة وكان ذلك مصدراً أساسياً للوصول إلى نتائج دقيقة سليمة وإلى اتخاذ قرارات على درجة كبيرة من الكفاءة عند مستوى مرتفع من الثقة .

وبقدر قدم هذه الوظيفة الإحصائية إلا أنها وظيفة متطورة من حيث العمق والأتساع حيث أنها أصبحت تحوي أحسن وأدق وأحدث الطرق العلمية في جمع البيانات إلى جانب أنها لم تعد وظيفة جمع البيانات عن الظواهر التقليدية لتحديد قوة الدولة أو قدراتها على محاربة دولة مجاورة أو رغبتها في جباية الضرائب أو فرض نوعية جديدة منها ، بل امتدت عملية جمع البيانات لمعرفة أدق الحقائق عن الظواهر بمختلف أنواعها لتلبية احتياجات عملية التخطيط لكافة الأنشطة المختلفة للدولة العصرية من نشاط صناعي وتجاري وزراعي إلى نشاط اجتماعي ثقافي سواء كان ذلك على المستوى القومي أو الخاص .

وغني عن البيان فإن الأسلوب الإحصائي في إطاره الحديث وأسلوبه الجديد يقدم للباحث الطريقة العلمية لتجميع وجمع البيانات من مصادرها المختلفة بطرق موضوعية دون أي تحيز .

ويعتمد أسلوب جمع البيانات على الأسلوب العيني من واقع سحب عينة ممثلة لمجتمع ظاهرة البحث ومن واقع إطار إحصائي شامل .

3 . وظيفة التحليل البياني للمعلومات :

تعتبر هذه الوظيفة هي نقطة تحول أساسية في التطور الوظيفي لعلم الإحصاء وبداية لهذا التطور فبعد أن كانت العملية الإحصائية محصورة في مجرد إحصاء للبيانات من خلال وظيفتي العد وجمع البيانات أصبحت العملية الإحصائية تمتد إلى أبعد من ذلك وأعمق في وقتنا الحاضر وذلك على نحو ما سيأتي من خلال تتبع التطور الوظيفي للعلم . وفيما سبق كان الانطباع عن حقائق الظواهر يؤخذ بطريقة محدودة وسطحية غير دقيقة حيث أن وظيفتي العد وجمع المعلومات عن خصائص ظواهر المجتمع المختلفة لم تعد كافيه لتأسيس أخطر وأدق الحقائق عن الظواهر .وباستحداث أسلوب التحليل البياني أصبح سهلاً على الباحثين والدارسين تحديد أكبر عدد ممكن من خصائص الظواهر المحيطة وبطريقة علمية تهدف إلى إعطاء أشكال بيانية للظاهرة من خلال البيانات المتاحة مما يسهل ويبسط تحديد الخصائص والعلاقات والاتجاهات العامة للظاهرة وتحديد انتماء الشكل إلى بعض المجموعات الأساسية ذات الخصائص المحددة .

هذا الأسلوب في نطاق العمل الإحصائي هام ومفيد في مجال تحليل الظواهر بطريقة سهلة مبسطة فالشكل البياني هو أسهل الأدوات في الحكم والتعبير عن أهم الحقائق للظواهر موضع الدراسة .

4 . وظيفة التحليل الكمي للبيانات :

هذه الوظيفة تعد إضافة هائلة إلى أسلوب العمل الإحصائي في دراسة خصائص الظواهر بطريقة قياسية كمية أعطت للعلم قوة وأهمية ومكانة بين باقي العلوم الأخرى ظهرت في القرن السابع عشر وكانت نتيجة حتمية للتطور الهائل في استخدام العلوم والتكنولوجيا في كافة ميادين الحياة الحديثة .

ويعتمد هذا الأسلوب في البحث على استخدام المقاييس والمؤشرات الإحصائية بطريقة علمية وموضوعية سليمة في تقصي الحقائق وتحديد أدق الخصائص ومعرفة أسباب الحركة المستمرة لأهم ظواهر حياتنا اليومية . ونتيجة لاستخدام

الأسلوب الكمي في تحليل المعلومات أصبحت النتائج على درجة عالية من الدقة تصلح أساساً سليماً مطمئناً لاتخاذ القرارات .

5 . وظيفة وضع الفروض :

إن تعدد المشاكل في مختلف مجالات حياتنا المعاصرة ووجود الكثير من المتغيرات التي تحكم حركة هذه المشاكل وتعقد العلاقات المبادلة بين هذه المتغيرات وتشابكها وصعوبة تحديد العلاقات بينها بطريقة جعلت عملية البحث العلمي أكثر تعقيداً مما كانت عليه أدى ذلك إلى البحث عن الطريقة العلمية لتبسيط عملية التعامل مع هذه المتغيرات .

ويعتبر أسلوب العمل الإحصائي في تطوره الوظيفي من أدق وأحسن هذه الطرق ، حيث أن الأسلوب الإحصائي في شكله المعاصر يعطي للباحث الأسلوب العلمي لكيفية التعامل مع المتغيرات التي تحكم نظم التغير في الظواهر المختلفة ،ووظيفة وضع الفروض تهدف أساساً إلى تبسيط المشكلة موضع الدراسة والتحليل وذلك من خلال وضع فروض محددة من منطلق ما يتصوره وما يشعر به الباحث تجاه ما ينوي دراسته ووضع النتائج بصدد حل المشكلة موضع البحث . والأسلوب الإحصائي يعطي لنا تصور عام لطريقة وضع الفروض تمهيداً لاختبارها سواء كانت هذه الفروض على المستوى البسيط أو المعقد .ويعتبر أسلوب عزل بعض المتغيرات أي افتراض عدم تأثيرها على الظاهرة موضع الدراسة أحد الأساليب المستخدمة في تبسيط طرق معالجة المشاكل وتحديد الخصائص والتأكد من صحة بعض النظريات . فالدارس للمتغيرات المؤثرة في حجم مبيعات أحد السلع ويريد قياس مدى تأثير أحد هذه المتغيرات فإنه يفترض ثبات أثر العوامل العشوائية أو الدورية مثلاً حتى يستطيع بذلك تحديد درجة تأثير عامل الاتجاه العام أو الأثر الموسمي على حجم المبيعات

كما أن الباحث الاقتصادي عند وضع تصور عام عند بحث أحد المشاكل الاقتصادية إنما يحاول أن يضع المتغيرات المحددة لهذه المشكلة داخل إطار تصوره وذلك من خلال التفسير والافتراض ، فهو قد يفترض مثلاً رشد المستهلك أو رغبة المنتج في تعظيم دالة الربح أو تصغير دالة التكاليف وهو بذلك يكون قد عزل العديد من المتغيرات التي قد تتعارض مع هذه الفروض أو التي قد لا تفسر العلاقات المتبادلة بين متغيرات الظاهرة موضع البحث والدراسة .ويشير العمل الإحصائي من خلال هذه الوظيفة إلى العديد من الاعتبارات والضروريات التي يجب الاسترشاد بها عند وضع الفروض تمهيداً لاختيارها وللتأكد من صحتها أو

عدم صحتها . فعند إتباع أسلوب الإبعاد أو عزل المتغيرات أو عند وضع بعض الافتراضات السلوكية يجب ألا نتمادى في عزل العديد من المتغيرات حتى لا نفتقد الحقيقة وتثبيت عكسها بطريق مضلل نتيجة إفتراض هذا القبيل وعليه فيجب على الباحث وضع ترتيب منظم لدرجة تأثير وأهمية المتغيرات على حركة الظاهرة مع عدم إهمال إمكانية قياس التغير في هذه المتغيرات ومدى إمكانية استخدام القوانين والنظريات الإحصائية في ذلك .

وبصفة عامة فإننا يجب أن نحكم المنطق عند وضع الافتراضات والأوليات لدرجة تأثير المتغيرات على الظاهرة غير متجاهلين موقف هذه الافتراضات من الاختبارات الإحصائية .

6 . وظيفة الاختبارات الإحصائية :

هذه الوظيفة مكملة للوظيفة السابقة فاستخلاص النتائج واتخاذ القرارات لدراسات مبنية أساساً على وضع فروض محددة يجب ألا يتم إلا بعد اختبار صحة هذه الفروض وهنا نجد دوراً كبيراً للنظريات الإحصائية والتي خصصت لكيفية اختبار صحة هذه الفروض في ظل درجات ثقة عالية وأدنى درجات من الخطأ المسموح به .

والمعروف إحصائياً أن اختبار الفروض في مجال الدراسات الميدانية يكون أصعب منه في مجال الدراسات المعملية . فالدراسات الميدانية بحكم تغير ظواهرها والعديد من المتغيرات التي في كثير منها يصعب تحديدها عددياً أو قياسها كمياً وبالتالي فانه في هذه الحالة فإن الاختبار يتم من خلال المشاهدات المتكررة ومقارنة عملية التغير في الظاهرة وحقيقة هذا التغير بالفروض الموضوعة ويكون لنا قبول الفرض عن ملاحظة عدم وجود اختلافات جوهرية بين ما تم تسجيله من واقع المشاهدات وما تم افتراضه من واقع التصور وتفسير علاقات متغيرات للظاهرة ، ويعتبر الفرض صحيحاً إحصائياً ويمكن قبوله وذلك من خلال إتباع الأسلوب الإحصائي لاختبارات الفروض ، أما إذا وجدت اختلافات جوهرية فيجب علينا رفض الفرض وعدم قبوله لأنه بذلك يكون فرضاً غير صحيحاً لأن المشاهدات الواقعية لا تؤيد ما كان يتوقعه الباحث عند تفسيره للتغير في الظواهر ولم يكن موفقاً في ذلك ، بينما يتم اختبار الفروض في الدراسات المعملية من خلال تسجيل القراءات والقياسات نتيجة إجراء التجارب المعملية مع تطبيق بعض النظريات الإحصائية لاختبارات الفروض والتي سوف يتم التعرض لها فيما بعد لمعرفة درجة تطابق النتائج المعملية بما تصوره وتنبأ به الباحث من قبل حتى

يمكن قبول هذه الفروض أو رفضها فإذا تم التوصل إلى عدم وجود فرق جوهري بين القراءات وما تم التنبؤ به من قبل فيمكن قبول النظرية ويكون الفرض في هذه الحالة صحيحاً في حدود خطأ مسموح به عند مستوى معين ، وفي حالة التوصل إلى وجود فرق جوهري وحقيقي (معنوي) بين قياسات التجارب المعملية وما تم تصوره تجاه متغيرات الظاهرة سواء كان من خلال النظرية أو الفرض ففي هذه الحالة يتم رفض النظرية أو الفرض .

ولا ننسى هنا أن رفض الفرض لا يعني عدم صحته على الإطلاق ولكن هذا يعني أن الباحث لم يتوصل بعد من خلال مشاهداته الواقعية أو قياساته وقراءاته المعملية إلى درجة قبول هذا الفرض ، كما لا ننسى هنا إلى الإشارة بأن الخبرة الطويلة والخلفية السابقة في نطاق وضع الفروض واختبارها دور لا يمكن إهماله بأي حال من الأحوال في واقعية الفروض وقربها من الحقيقة وقبولنا هذه الفروض بعد اختبارها .

كما أن الإلمام بالطرق والأدوات الإحصائية والقوانين والنظريات المنظمة لأسلوب الاختبار الإحصائي يساعد إلى درجة كبيرة في استخلاص النتائج السليمة وإصدار القرار غير المتحيز بالنسبة لحل العديد من مشاكل وقتنا المعاصر .

7 . وظيفة استخلاص النتائج :

إن التطور الوظيفي لأسلوب العمل الإحصائي والذي ظهر بوضوح في نهاية القرن السابع عشر ومصاحبة هذا التطور بتطور في الطرق والنظريات واستخدام نظريات جديدة لها مجال تطبيقها الواسع الانتشار في العديد من نواحي الحياة المعاصرة المعقدة ، أدى ذلك إلى وجود الأسلوب العلمي في إطار إحصائي على درجة عالية من الكفاءة في استخلاص النتائج بطريقة موضوعية بعيدة عن أخطاء يمكن أن تقع نتيجة الاعتماد على الطرق العادية في استخلاص النتائج ولقد أصبحت النظرية الإحصائية في وقتنا المعاصر من أدق الأدوات للدراسات العلمية والتي يعتمد في تكوينها على فروض محددة وتأكد من صحة هذه الفروض واستخلاص النتائج .

8 . وظيفة اتخاذ القرارات :

أن أي دراسة علمية هادفة سليمة هي تلك التي تنتهي باتخاذ قرارات عملية صالحة للعمل بها . غير أن اتخاذ القرار السليم ليس بالمسألة السهلة وذلك لتشابك الأمور وتداخلها أو تعقد المتغيرات عن الظواهر وتأثيرها المتبادل في بعضها في ظل وجود العديد من البدائل لحل المشاكل وصعوبة تحديد البديل المناسب بسهولة

إلا أن الأسلوب الإحصائي وما يحمله في طياته من قوانين ونظريات إحصائية متطورة حديثة قد ساهم بقدر عظيم وخصوصاً بعد أن أخذت نظرية الاحتمالات والتوقع الرياضي نصيباً هائلاً من التطور في اتخاذ القرارات بدرجة من الثقة العالية وبنسب خطأ عند حدودها الدنيا .

لقد أصبحت وظيفة اتخاذ القرارات هي أساس العمل الإحصائي وعموده الفقري وأصبح علم الإحصاء في وقتنا المعاصر يفهم ويعرف من خلال وظيفة اتخاذ القرارات .

9 . وظيفة التنبؤ الاستدلالي :

من أهم وظائف واستخدامات الأسلوب والنظرية في علم الإحصاء وظيفة التنبؤ الاستدلالي بالخصائص والمؤثرات للعديد من متغيرات الظواهر في المجتمع . ومن خلال هذه الوظيفة وباستخدام طرق القياس والتحليل الإحصائي يمكن التوصل إلى اتجاه عام لما سيحدث في المستقبل للمتغيرات التي تتحكم في ظاهرة ما ، مثل التنبؤ بحجم الطلب الكلي أو التنبؤ بمعاملات المتغيرات المحددة لدالة الاستثمار القومي أو الدخل القومي إلى غير ذلك .

والتنبؤ في هذا الإطار خاص بالمستقبل وبتوضيح العلاقات بين متغيرات الظواهر لفترة مستقبلية . غير أن التنبؤ في مفهومة الاستدلالي أو التنبؤات الاستدلالية هي تلك التي تخص الماضي وليس المستقبل حيث يكون لها طابع الاستدلال أو الاكتشاف أو التأكد من وجود ظاهرة متكررة الحدوث دون ملاحظة سبب ذلك ويكون التنبؤ هنا لتأكيد وجود الظاهرة من خلال الملاحظة والقياس وتطبيق أسلوب العمل الإحصائي في تجميع البيانات وتسجيل الاتجاهات وتحديد الأسباب وتفسير التغيرات واستخلاص النتائج ، ففي هذا النوع من التنبؤ يقوم الباحث بوضع فروض محددة محاولاً بعد ذلك جمع البيانات مع الإطلاع على التقارير والسجلات عن الظاهرة موضع التنبؤ واختبار صحة هذه الفروض

10 . وظيفة البحث العلمي :

إن التطور الوظيفي لعلم الإحصاء في الإطار السابق عرضه إنما يعطي لنا أسلوباً علمياً وأداة حديثة تخدم أسلوب الدراسات العلمية سواء كانت ميدانية أو معملية. فإذا ما قمنا بأخذ الوظائف السابقة في ترتيبها المنطقي لوجدناها تصلح أساساً لخطوات تتبع في تنفيذ البحث العلمي . وعليه فإن العمل الإحصائي كالعملة لها وجهان الوجه الأول يعبر عنه بالوظائف الرئيسة لعلم الإحصاء أما الوجه الآخر فيعبر عنه بوظيفة البحث العلمي .

والباحث أو الدارس في استخدامه لهذه المراحل أو الوظائف في دراسته الميدانية أو المعملية ، يجب أن يدرك ويستوعب هذه المراحل ويعتبرها أحد طرق البحث العلمي ، كما يجب عليه أن يجيد الاختيار طبقاً لطبيعة دراسته ونوعية المتغيرات التي يتعامل معها ، وتحكيم كل من عنصري الزمان والمكان في ذلك .

وبصفة عامة فإن علم الإحصاء من خلال وظائفه المختلفة من اختيار موضوع البحث وتجميع المعلومات وتحليلها مع وضع الفروض واختبارها وأخيراً استخلاص النتائج واتخاذ القرارات إنما يصلح لأن يكون من أدق طرق البحث العلمي وإضافة حقيقية في هذا المجال.

<u>الإحصاء وعلاقته بالعلوم الأخرى</u>

علم الإحصاء هو علم العلاقات المتبادلة بالعلوم الأخرى فهو يؤثر ويتأثر بها في نطاق تطورها المستمر عبر التقدم التكنولوجي المعاصر حيث تحتل الطرق والنظريات الإحصائية مكانة مرموقة في العلوم الأخرى وتعتبر أساساً لتطورها ولاستحداث أبسط وأسرع الطرق في نطاق تطبيقها علمياً وذلك على نحو ما سيأتي :

1 . علاقة علم الإحصاء بمجموعة العلوم الإدارية :

يرتبط علم الإحصاء ارتباطا قوياً بمجموعة العلوم الإدارية وذلك على أساس أن وظائف علوم الإدارة تستند في القيام بها بطريقة موضوعية على العديد من الطرق والنظريات الإحصائية .فاتخاذ القرار ضروري وهام في علم الإدارة ويجب أن يؤخذ على أساس علمي غير متحيز ولكي يكون كذلك بفضل استخدام الأسلوب القياسي وهنا نجد أن نظرية الاحتمالات والتوقع الرياضي تقدم لنا هذا الأساس القياسي في اتخاذ القرار .

كما أن تخطيط عمليات الشراء أو البيع ودراسة طرق التخزين المتعددة وإدارة الإنتاج الصناعي وسياسات التسويق المختلفة والدراسات السلوكية المتعلقة بالمنتج والمستهلك وشؤون إدارة الأفراد وإدارة المؤسسات المتخصصة ودراسة الوقت والحركة كل هذا إنما يحتاج من دارسي العلوم الإدارية ومتخصصيها والباحثين في كافة هذه المجالات الإلمام بأحدث الطرق والأساليب الإحصائية وما تعطيه النظريات الإحصائية من تفسيرات وتحديداً للعلاقات بين متغيرات هذه العلوم وقدرة كبيرة على وضع الفروض واختبارها والتأكد من صحتها ومعرفة درجة صدق

المقياس المستخدم وثباته والقدرة على استخدام وتطبيق خرائط المراقبة الإحصائية لجودة الإنتاج وكلها أمور يستطيع الإداري الملم والمطلع بالجوانب المختلفة لعلم الإحصاء أن يتقن تنفيذها واستخدامها حيث أنها ضرورية ومستخدمة ومطبقة حالياً وأساسية في كافة علوم الإدارة .

2 . علاقة علم الإحصاء بأساليب بحوث العمليات :

تعتمد أساليب بحوث العمليات في عرضها واستخداماتها على العديد من المفاهيم والأساليب والقوانين الإحصائية ، مما يجعل من الضروري لمستخدمي أساليب بحوث العمليات الإلمام التام بالطرق الإحصائية ونظرياتها المختلفة .

وتحتل نظرية الاحتمالات والتوقع الرياضي والتوزيعات الاحتمالية وعلى الأخص التوزيع ذي الحدين وتوزيع بواسون في هذا الصدد مكانة مرموقة باعتبارها أساسية في وضع النماذج الرياضية المختلفة في حل المشاكل الإدارية والاقتصادية وتحديد تفسير العلاقات المتشابكة لمتغيرات كل نموذج ثم اتخاذ القرار اللازم لحل المشكلة والتأكد من صحة ذلك .

3 . علاقة علم الإحصاء بمجموعة العلوم المحاسبية :

إن استخدام طرق محاسبية جديدة في وقتنا المعاصر كان أساسه الطرق والنظريات الإحصائية في مجال مجموعة العلوم المحاسبية فالفضل يرجع إلى الأساليب الإحصائية والمبادئ والنظريات الحديثة لهذا العلم في تقدم مختلف العلوم المحاسبية حيث أصبحت النظم المحاسبية الحديثة هي التي تعتمد على النظرية الإحصائية في عرض الموضوعات بشكل مبسط غير متحيز ، فالمراجعة المستندية تعتمد وتستخدم أسلوب ونظرية العينات في عمليات المراجعة المختلفة في حدود درجات من الثقة المرتفعة دون تضحية بأخطاء لها ضررها على المراجع مع توفير الوقت والجهد والتكاليف تمشياً مع روح العصر في ضرورة السرعة في إتمام عملية المراجعة في ظل ظروف العمل الشاقة وكثرة العمليات المطلوب مراجعتها ، كما أن فكرة التكاليف المعيارية تعتمد أساساً على خصائص التوزيع المعتدل وعلى استخدام بعض المقاييس والمؤشرات الإحصائية .

وفي وقتنا الحالي أصبح الإحصاء جزء هام وضروري في دراسة المحاسبة الإدارية والنظم المحاسبية المعاصرة حيث أن اتخاذ قرار بين عدة بدائل لاختيار أنسب الطرق في التقدير والتنبؤ أصبح أساسه إحصائي قبل أن يكون محاسبي وذلك من خلال استخدام المقاييس والمؤشرات والنظريات والجداول الإحصائية

4 . علاقة علم الإحصاء بعلم الاقتصاد :

من الصعب فصل العمل الإحصائي عن العمل الاقتصادي فأي دراسة اقتصادية هدفها التخطيط أو التقدير أو التنبؤ سواء كان ذلك على مستوى المشروع الخاص أو الاقتصاد القومي يلزمها توفر البيانات والمعلومات عن كافة المتغيرات المحددة لهذه الدراسة والذي بدورة يمكن الحصول عليها باستخدام أسلوب العمل الإحصائي ، كما أن دراسة السوق لغرض معرفة وتحديد العوامل المؤثرة على طلب وعرض إحدى السلع أو الخدمات يكون من خلال الأسلوب العلمي للعمل الإحصائي ، ولا يمكن أن تخطط المدن وتحدد الأولويات العمرانية بين سكنى الأفراد وبناء المصانع وإقامة المحال التجارية وبناء المدارس والمستشفيات وتحديد التوزيع النوعي والعمري للسكان واللازمين لعملية التخطيط والبناء إلا بوجود بيانات ومعلومات كافية وشاملة عن ذلك ، وهنا نجد أن الإحصاءات الديموغرافية الغنية بالمقاييس والمؤشرات الإحصائية في هذا المجال وإحصاءات سوق العمل والإحصاءات الاقتصادية (تجارية ، صناعية ... الخ) وأيضاً الإحصاءات النقدية والمالية وإحصاءات المعاملات الخارجية ، كل هذه الإحصاءات تعتبر من أهم المصادر للمعلومات الضرورية للقيام بعملية التخطيط على كافة المستويات .

إن علم الإحصاء أصبح اليوم جزء أساسي وضروري للعمل الاقتصادي وتقدمه ، فأي دراسة اقتصادية إنما تعتمد على أسلوب العمل الإحصائي في تنفيذ ذلك ، كما أن المؤشرات والمقاييس الإحصائية أصبحت من الأدوات اللازمة في حقل العمل الاقتصادي سواء كان ذلك يتعلق بالأسعار أو الأجور أو بالاتجاهات العامة لكل من الادخار والاستثمار أو الاستهلاك أو أي متغير من متغيرات الاقتصاد القومي بصفة عامة سواء أكان ذلك لغرض التخطيط أو لعمل المقارنات

5 . علاقة علم الإحصاء بعلم الاقتصاد الرياضي :

إذا كان الاقتصاد الرياضي عبارة عن الطريقة التي تستخدم للتعبير عن العلاقات بين المتغيرات الاقتصادية للظواهر بشكل رياضي مستفيدين بذلك لما قد يعطيه لنا الأسلوب الرياضي عند التعبير عن المشاكل الاقتصادية من سهولة وتبسيط للعرض ودقة في الوصول إلى النتائج دون تحيز أو غموض في تفسير الحقائق، فإن الأسلوب الرياضي يعتمد إلى درجة كبيرة في عرضه لهذه المشاكل وصياغتها رياضياً إلى الأسلوب الإحصائي ونظرياته وخصوصاً عند تصميم النماذج الاقتصادية بأشكالها المختلفة وما يتضمنه ذلك من وضع للفروض وإجراء

للاختبارات الإحصائية واستخدام طرق التنبؤ الإحصائي للمتغيرات الاقتصادية موضع البحث وذلك لتحديد الاتجاهات وتعميم النتائج .
وفي الآونة الأخيرة وصلت العلاقة بين علم الإحصاء والاقتصاد الرياضي إلى درجة ملحوظة وخصوصاً من وجهة نظر الاقتصاديون الذين يرون ضرورة وضع بعض الفروض الواقعية عند تصميم النماذج الرياضية والتي تجعل الثغرة بين نظرية النماذج وواقعية تطبيقها أقل ما يمكن .

6 . علاقة علم الإحصاء بعلم الاقتصاد القياسي :

إذا تم التعبير عن المشاكل والنظريات الاقتصادية في صيغة رياضية وذلك من خلال النماذج الرياضية مع الأخذ في الاعتبار واقعية الفروض المحددة لشروط الصياغة الرياضية للنموذج أي الأخذ في الاعتبار العوامل الخارجية المؤثرة بدرجة معينة في طبيعة المشكلة الاقتصادية موضع البحث فإننا نكون في هذه الحالة بصدد أسلوب الاقتصاد القياسي ، والاقتصاد القياسي بهذا التصور يكون أقوى إلى درجة كبيرة في علاقته وتأثره بالأسلوب الإحصائي منه في الاقتصاد الرياضي .

وتقدم النظرية الإحصائية من خلال أسلوب العمل الإحصائي للعديد من احتياجات وأدوات العمل للاقتصاد القياسي من حيث العديد من المقاييس والمؤشرات الإحصائية وطرق قياس أثر المتغيرات المختلفة والمؤثرة في المشكلة الاقتصادية موضع الدراسة بكل تحديد ودقة إلى جانب استخدام طرق القياس الإحصائي في تخليص العديد من الظواهر الاقتصادية من أثر بعض المتغيرات مع عمل التطبيقات المختلفة بها من حيث إمكانية التحكم في قيامها والتنبؤ بما يمكن أن تكون عليه مستقبلاً ، وإذا كانت طريقة الاقتصاد القياسي في البحث نعتمد إلى حد كبير على ما يمكن مشاهدته عملياً من خلال الدراسات الميدانية فإن الطريقة الإحصائية هي خير طريقة يمكن استخدامها في هذا المجال

7 . علاقة علم الإحصاء بعلم الرياضة البحتة :

العلاقة بين علم الإحصاء وعلم الرياضة البحتة قوية وعظيمة فالعديد من النظريات الإحصائية تعتمد في صياغتها على الأسلوب الرياضي وفي تطورها واستخدام طرق بديلة للعرض والإثبات فالعديد من التوزيعات الاحتمالية توضع في شكل دوال رياضية بها العديد من المتغيرات ، كما أن المعالجة الرياضية لهذه الدوال باستخدام نظريات التفاضل والتكامل يعطي لنا أدق المقاييس والمؤشرات الإحصائية اللازمة لعملية التحليل والدراسة .

ولاشك في أن المتخصص الإحصائي يكون أكثر قدرة على استخدام الأسلوب الإحصائي إذا ما كانت لديه الخلفية الرياضية في حدود هذا الأسلوب وذلك عند المتخصص الإحصائي الذي لا يملك هذه الخلفية في مجال الرياضة البحتة .
ويجب ألا ننسى أن الأسلوب الإحصائي كطريقة علمية صالحة للتطبيق في مجال البحث العلمي كثيراً ما يستخدمه رجال الرياضة البحتة عند عرض مشاكلهم بأسلوب رياضي .

8 . علاقة علم الإحصاء بمجموعة العلوم الطبيعية :

تعتمد الآن معظم الدراسات المعملية على الأسلوب الإحصائي في تنفيذ التجارب وتصميمها ، وتلعب نظرية الاحتمالات والعينات دوراً كبيراً في هذا المجال سواء كان ذلك المجال كيميائي أو زراعي أو صيدلي أو طبي أو هندسي أو أي مجال يدخل في إطار مجموعة العلوم الطبيعية .
ويظهر هنا بوضوح الاستخدامات المختلفة لأسلوب العمل الإحصائي وذلك لغرض التقدير الإحصائي لخصائص الظواهر وتعميم النتائج على المجتمعات الأصلية وتفسير النتائج بأسلوب عام واختيار صحة هذه النتائج بدرجات ثقة يمكن التحكم فيها بحيث يمكن جعل الخطأ المسموح به عند أدنى حد ممكن .
كما أن أسلوب التأكد من صحة بعض النظريات في مجال العلوم العملية غالباً ما ينفذ من خلال إتباع الأسلوب الإحصائي من تسجيل للمشاهدات من واقع الظواهر العملي أو الحصول على القياسات والقراءات من واقع التجارب المعملية ثم إجراء المقارنات بين ما هو موجود نظرياً على أساس النظرية الإحصائية وما يمكن أن تعطيه للباحث من أدوات في مجال تصميم التجارب .

9 . علاقة علم الإحصاء بمجموعة العلوم الإنسانية :

بعد التطور التكنولوجي الهائل في كافة الميادين والذي فرض نفسه فجأة اصطحب هذا بتطور في كافة العلوم الإنسانية من حيث استحداث طرق جديدة لمعالجة الموضوعات الاجتماعية والفلسفية والنفسية وغير ذلك وأصبحت العلوم الطبيعية من أهم الموارد المساعدة في تنفيذ البحوث الاجتماعية .ولا يمكن إنكار دور علم الإحصاء في هذا التقدم ، فالعلم يحتل مكانة كبيرة ويعتبر جزأً غير بسيط من ضمن هذه العلوم إلى الحد الذي تجد فيه فرعاً من فروع علم الإحصاء يسمى بالإحصاء في مجال العلوم الاجتماعية أو الإحصاء الاجتماعي والطريقة الإحصائية والنظريات العلمية هي أهم أدوات الباحث في مجال العلوم الإنسانية .
فالطريقة الإحصائية هي أسلوب عمل لتنفيذ البحوث الاجتماعية ونظرية

الاحتمالات والنهاية المركزية وما يشمل ذلك من تطبيقات أساسية لها أهميتها في هذا المجال ، كما أن أسلوب إيجاد علاقة الارتباط سواء كان بسيطاً أو متعدداً للظواهر الاجتماعية والفلسفية وغير ذلك من الظواهر التي نفسرها وندرسها وتدخل في إطار العلوم الإنسانية ، وأيضاً تطبيق نظرية وضع الفروض والاختبارات الإحصائية وتحديد انتماء العديد من تلك الظواهر وتبعيتها لأحد التوزيعات الاحتمالية ، كل ذلك ضروري وهام في مجال العلوم الإنسانية ، وليس بالغريب القول بأن كل باحث متخصص في مجال العلوم الاجتماعية يجب عليه أن يكون ملماً عارفاً لأهم خطوات الطريقة الإحصائية والنظريات المختلفة لهذا العلم والمجالات التطبيقية المتعددة له إذا كان يريد أن يرقى بأبحاثه ومعلوماته إلى مستوى روح العصر .

وهكذا نستخلص من هذا العرض أن الإحصاء هو علم له طرقه العلمية ووظائفه المتطورة وقوانينه ونظرياته المتعددة والتي تعتبر أساساً للكثير من العلوم الأخرى ومنطلق لتطورها . وهو علم له علاقاته الممتدة عبر كل العلوم يؤثر فيها ويتأثر بها ويمثل جزء يكاد يكون عاماً ومشتركاً في كل العلوم تبدأ به وتنهل من طرقه ونظرياته مع اختلاف في درجة الامتداد والتشعب من علم إلى آخر ، كما أنه علم له وجوده في حياتنا العملية وأن أي تصرف أو سلوك شخصي أو غير شخصي يمكن أن تحكمه نظرية إحصائية أو أن يكون منطلقاً من أحد الطرق الإحصائية . إنه علم له العديد من الوظائف المتطورة مع التقدم والرقي في كافة الميادين وهي تشكل في إطارها العام أدق وأحسن أسلوب للبحث العلمي الخلاق وذلك على نحو ما تم إيضاحه

تذكـر

•المتغير الكمي (Quantitative variables)

هو المتغير الذي يعبر عنه بالمقدار مثل العمر، الدرجة لطالب، أيام الغياب، ... يمكن أن يأخذ صفة الترتيب

•المتغير النوعي (Qualitative variables)

الذي يصنف الأشياء كالجنس (ذكور وإناث) والتخصص (علمي – أدبي)، المرحلة (ابتدائي – إعدادي – ...)،... ويفقد صفة الترتيب

•المتغير المستقل والمتغير التابع:الذي يؤثر في النتائج أو الذي يتسبب فيها ويمكن التحكم به (تغييره) ويعرف بالمتغير التجريبي كطرق التدريس أو طرق العلاج والنتائج المترتبة على المتغير المستقل تعرف بالمتغير التابع فطرق التدريس كمتغير مستقل تؤثر في مستوى التحصيل الذي يعتبر متغير تابع (ناتج التجربة لطرق التدريس)، وكذلك نتائج العلاج على المريض تعتبر متغير تابع للمتغير المستقل طرق العلاج.

• تتعدد تعاريف علم الإحصاء من عدة نواحي كالتالي:

لغوياً الإلمام بكل المفردات التي يشملها المجتمع المطلوب دراسته ودراسة كافة مفرداته رقمياً.

علميا عبارة عن تصوير رقمي للواقع فى المجتمعات المطلوب دراستها سواء كانت بشرية أو غير بشرية عن طريق وصف مفردات المجتمع وصفا رقميا

التعريف الشامل الإحصاء هو العلم الذى يبحث فى:

(أ) جمع البيانات والمعلومات والحقائق الخاصة بمختلف الظواهر وتسجيلها فى صورة رقمية وتصنيفها فى جداول منظمة وتمثيلها بيانيا.

(ب) تحليل البيانات واستخلاص النتائج منها واتخاذ القرارات المناسبة

(ج) مقارنة الظواهر ببعضها ودراسة العلاقات بينها واستخدامها فى تفهم حقيقة الظواهر ومعرفة القوانين التي تسير تبعا لها.

• ينقسم علم الإحصاء إلى قسمين رئيسين، هما

الإحصاء الوصفي، ويختص بجمع وعرض البيانات وذلك بهدف توفير المعلومات عن الاتجاهات المختلفة للظواهر واستخراج بعض المؤشرات الإحصائية بهدف توضيح الطبيعة العامة للظواهر موضع الدراسة.
والإحصاء الرياضي، والذي يستند على الكثير من المفاهيم والنظريات الرياضية و من أهمها نظرية الاحتمالات، كما يختص بمحاولة اتخاذ القرارات المناسبة فى المشاكل موضع الدراسة.

- **للاحصاء وظائف متعددة من اهمها:**

العد (الحصر) حيث يعطي لنا المزيد من البيانات والمعلومات التفصيلية في كل المجالات بأسلوب يخدم أغراض التخطيط والتنمية الاقتصادية للبلاد من خلال أسلوب يعتمد على النظريات الإحصائية في تفسير الاتجاهات وتحليل التغيرات وتفسير العلاقات بين المتغيرات وإيضاح أسبابها. ولم تعد وظيفة الحصر قاصرة على تعداد السكان أو التعداد الزراعي أو التعداد الاقتصادي ولكنها امتدت الي العديد من المجالات الهامة.

جمع البيانات حيث من الضروري توافر البيانات عن الظواهر والعوامل المحددة لها ، والمعلومات عن الظواهر موضع البحث بحيث يمكن دراسة وتحليل واستخلاص النتائج واتخاذ القرارات . خاصة إذا ما أتبع أسلوب علمي وغير متحيز في جمع البيانات أدى ذلك إلى الحصول على حقائق عن الظواهر بطريقة سليمة غير متحيزة وكان ذلك مصدراً أساسياً للوصول إلى نتائج دقيقة سليمة وإلى اتخاذ قرارات على درجة كبيرة من الكفاءة .

التحليل البياني للمعلومات أن وظيفتي العد وجمع المعلومات عن خصائص ظواهر المجتمع المختلفة لم تعد كافيه ولذا استحداث أسلوب التحليل البياني ليسهلً على الباحثين تحديد أكبر عدد ممكن من خصائص الظواهر المحيطة وبطريقة علمية تهدف إلى إعطاء أشكال بيانية للظاهرة من خلال البيانات المتاحة مما يسهل ويبسط تحديد الخصائص والعلاقات والاتجاهات العامة للظاهرة ،

التحليل الكمي ان استخدام المقاييس والمؤشرات الإحصائية بطريقة علمية وموضوعية سليمة فى تقصى الحقائق وتحديد أدق الخصائص ومعرفة أسباب

للبيانات تغيير أهم ظواهر حياتنا اليومية . وباستخدام الأسلوب الكمي في تحليل المعلومات أصبحت النتائج على درجة عالية من الدقة وتصلح أساساً سليماً مطمئناً لاتخاذ القرارات .

وضع الفروض يمكن تبسيط المشكلة المدروسة من خلال وضع فروض محددة لما يتصوره وما يشعر به الباحث ووضع النتائج بصدد حل المشكلة موضع البحث . والأسلوب الإحصائي يعطي لنا تصور عام لطريقة وضع الفروض تمهيداً لاختبارها سواء كانت هذه الفروض على المستوى البسيط أو المعقد وكذلك التأكد من صحة بعض النظريات

الاختبارات الإحصائية وهي وظيفة مكملة للوظيفة السابقة حيث نجد دوراً كبيراً للنظريات الإحصائية والتي خصصت لكيفية اختبار صحة هذه الفروض في ظل درجات ثقة عالية وأدنى درجات من الخطأ المسموح به .كما ان الإلمام بالطرق والأدوات الإحصائية والقوانين والنظريات المنظمة لأسلوب الاختبار الإحصائي يساعد إلى درجة كبيرة في استخلاص النتائج السليمة وإصدار القرار غير المتحيز بالنسبة لحل العديد من مشاكل وقتنا المعاصر

استخلاص النتائج

اتخاذ القرارات لقد أصبحت وظيفة اتخاذ القرارات هي أساس العمل الإحصائي وعموده الفقري وأصبح علم الإحصاء في وقتنا المعاصر يفهم ويعرف من خلال وظيفة اتخاذ القرارات .

التنبؤ الاستدلالي باستخدام طرق القياس والتحليل الإحصائي يمكن التوصل إلى اتجاه عام لما سيحدث في المستقبل للمتغيرات التي تتحكم في ظاهرة ما ، مثل التنبؤ بحجم الطلب الكلي أو التنبؤ بمعاملات المتغيرات المحددة لدالة الاستثمار القومى أو الدخل القومى إلى غير ذلك .

البحث العلمي إن علم الإحصاء من خلال وظائفه المختلفة من اختيار موضوع البحث وتجميع المعلومات وتحليلها مع وضع الفروض واختبارها وأخيراً استخلاص النتائج واتخاذ القرارات إنما يصلح لأن يكون من أدق طرق البحث العلمى وإضافة حقيقية فى هذا المجال.

- **علاقة علم الاحصاء بالعلوم الاخري**

العلوم يرتبط علم الإحصاء ارتباطا قوياً بمجموعة العلوم الإدارية وذلك على

الإدارية أساس أن وظائف علوم الإدارة تستند في القيام بها بطريقة موضوعية على العديد من الطرق والنظريات الإحصائية .

بحوث العمليات تعتمد أساليب بحوث العمليات في عرضها واستخداماتها على العديد من المفاهيم والأساليب والقوانين الإحصائية ، مما يجعل من الضروري لمستخدمي أساليب بحوث العمليات الإلمام التام بالطرق الإحصائية ونظرياتها المختلفة

العلوم المحاسبية أصبح الإحصاء جزء هام وضروري في دراسة المحاسبة الإدارية والنظم المحاسبية المعاصرة حيث أن اتخاذ قرار بين عدة بدائل لاختيار أنسب الطرق في التقدير والتنبؤ أصبح أساسه إحصائي قبل أن يكون محاسبي وذلك من خلال استخدام المقاييس والمؤشرات والنظريات والجداول الإحصائية

علم الاقتصاد أي دراسة اقتصادية هدفها التخطيط أو التقدير أو التنبؤ يلزمها توفر البيانات والمعلومات عن كافة المتغيرات المحددة لهذه الدراسة والذي بدورة يمكن الحصول عليها باستخدام أسلوب العمل الإحصائي ، كما أن دراسة السوق لغرض معرفة وتحديد العوامل المؤثرة على طلب وعرض إحدى السلع أو الخدمات يعتمد على أسلوب العمل الإحصائي ، والمؤشرات والمقاييس الإحصائية أصبحت هامة في حقل العمل الاقتصادي سواء كان ذلك يتعلق بالأسعار أو الأجور أو بالاتجاهات العامة لكل من الادخار والاستثمار والاستهلاك أوأي متغير من متغيرات الاقتصاد القومى .

الاقتصاد ا لرياضي يري الاقتصاديون الذين يرون ضرورة وضع بعض الفروض الواقعية عند تصميم النماذج الرياضية والتي تجعل الثغرة بين نظرية النماذج وواقعية تطبيقها أقل ما يمكن

الاقتصاد القياسي اذا كانت طريقة الاقتصاد القياسي في البحث نعتمد إلى حد كبير على ما يمكن مشاهدته عملياً من خلال الدراسات الميدانية فإن الطريقة الإحصائية هي خير طريقة علمية يمكن استخدامها في هذا المجال .

الرياضة البحتة تعتمد العديد من النظريات الإحصائية في صياغتها على الأسلوب الرياضي وفي تطورها واستخدام طرق بديلة للعرض والإثبات فالعديد من التوزيعات الاحتمالية توضع فى شكل دوال رياضية بها العديد من

المتغيرات ، كما أن المعالجة الرياضية لهذه الدوال باستخدام التفاضل والتكامل يعطي لنا أدق المقاييس والمؤشرات الإحصائية لعملية التحليل والدراسة

العلوم الطبيعية ان التقدير الإحصائي لخصائص الظواهر الطبيعية وتعميم النتائج على المجتمعات الأصلية وتفسير النتائج بأسلوب عام واختيار صحة هذه النتائج بدرجات ثقة والتأكد من صحة بعض النظريات في مجال العلوم العملية غالباً ما ينفذ من خلال إتباع الأسلوب الإحصائي من تسجيل للمشاهدات من واقع الظواهر العملي أو الحصول على القياسات والقراءات من واقع التجارب المعملية ثم إجراء المقارنات بين ما هو موجود نظرياً (النظرية الإحصائية) وما يمكن أن تعطيه للباحث من أدوات فى مجال تصميم التجارب

العلوم الإنسانية إيجاد علاقات الارتباط للظواهر الاجتماعية والفلسفية وغير ذلك من العلوم الإنسانية ، وأيضاً تطبيق نظرية وضع الفروض والاختبارات الإحصائية وتحديد انتماء العديد من تلك الظواهر وتبعيتها لأحد التوزيعات الاحتمالية ، كل ذلك ضروري وهام في مجال العلوم الإنسانية .

تمارين

س$_1$: "تتعدد تعاريف علم الإحصاء " ناقش ذلك باختصار؟

س$_2$: " تكلم بإيجاز عن ثلاث فقط من وظائف علم الإحصاء؟

س$_3$: ناقش بإيجاز علاقة علم الإحصاء بالعلوم الأخرى؟

الباب الثاني
المقاييس الإحصائية الوصفية

أولاً: مقاييس النزعة المركزية(المتوسطات)
Measures of central tendency

المتوسط هو القيمة النموذجية او الممثلة لمجموعة من البيانات ، وحيث ان مثل هذه القيمة النموذجية تميل الي الوقوع في مركز داخل مجموعة بيانات مرتبة حسب قيمتها ، فان المتوسطات تسمي ايضا بمقاييس النزعة المركزية . واهم هذه المتوسطات واكثرها شيوعا المتوسط الحسابي ، المتوسط التوافقي ، المتوسط الهندسي ، الوسيط والمنوال. وكل منها له مميزاته وعيوبه وهذا يعتمد علي البيانات والهدف من استخدامه ، فيما يناسب منها لنوع معين من البيانات قد لا يناسب نوعا اخر وسنوضح ذلك فما بعد بالتفصيل .

وعموما فان المتوسط الجيد يجب ان تتوافر فيه بعض الصفات والشروط منها ما يلي :

1. ان يكون معرفا تعريفا دقيقا ويتم حسابه علي اسس وقواعد محددة ولا يترك ذلك لتقدير الباحث وإلا اختلفت قيمته باختلاف الباحثين .
2. ان يكون المتوسط ممثلا جيدا لجميع القيم ، ولذلك يجب ان يأخذ جميع القيم في الاعتبار.
3. ان تكون طريقة حساب المتوسط سهلة وسريعة ، واذا ما تساوت المتوسطات في الخواص فيفضل الأسرع والأسهل في الحساب ، الا انه يجب عدم المغالاة في حساب الدقة .
4. الا يتاثر المتوسط كثيراً باختيار العينة (وهذا ما سيتناوله الجزء الخاص بالعينات فيما بعد بالتفصيل).
5. ان يكون من الممكن التعبير عنه ومعالجته جبريا .
6. يجب ان يكون معني المتوسط واضحا وان تكون خواصه بسيطة .

(1) <u>المتوسط الحسابي</u> : Arithmetic mean

هو ابسط المتوسطات جمعيا وأكثرها شيوعا ويمكن تعريفه بأنه القيمة التي إذا أعطيت لكل مفردة من المفردات المجموعة لكان مجموع هذه القيم الجديدة هو نفس مجموع القيم الاصلية ، وهو حاصل قسمة مجموع القيم علي عددها .

والوسط الحسابي لمجموعة N من الأرقام $X_1 , X_2 , X_3 ......X_N$ يرمز لـه بالرمز $\overline{X}$ ويحسب كما يلي :

$$\overline{X} = \frac{X_1 + X_2 + ... X_N}{N} = \frac{\sum X}{N}$$

مثال : المتوسط الحسابي للأرقام 6 , 9 , 5 , 3 , 2 هو :

$$\overline{X} = \frac{2 + 3 + 5 + 9 + 6}{5} = \frac{25}{5} = 5$$

ومن الملاحظ انه لو كانت الأعداد كبيرة وكثيرة فقد تكون عملية الجمع بالأعداد الكبيرة مملة وعرضة للأخطاء ، ولذلك فانه يفضل اختصار هذه العملية عن طريق الاستفادة من احدي خصائص المتوسط الحسابي والتي تتضح من المثال التالي :

مثال : أوجد الوسط الحسابي للقيم التالية :

106 , 109 , 105 , 103 , 102

وتكون خطوات الحل كما يلي :

1. نطرح عددا مشتركا من القيم (مثلا 100) فتصبح القيم :
2, 3, 5, 9, 6
2. نحسب المتوسط الحسابي للقيم الاصلية وفي هذه الحالة يساوي المتوسط الحسابي للقيم الجديدة مضافا اليه القيم السابق طرحها .

$$\overline{X} = \frac{2 + 3 + 5 + 9 + 6}{5} + 100 = 105$$

ولا شك ان لهذا الاختصار فائدة كبيرة خصوصا اذا كان المطلوب حساب المتوسط الحسابي لعدد كبير من القيم ، وغالبا ما يسمي العدد المشترك الذي نطرحه بالمتوسط الفرضي ويمكن اخذ أي قيمة كوسط فرضي ، الا انه بالخبرة يمكن ان نختار وسطا يسهل معه العمليات الحسابية .

وعلي لك فان المتوسط الحسابي لاي عدد من القيم يساوي أي قيمة اختيارية A (وسط فرضي) مضافا اليها الوسط الحسابي للقيم الجديدة الناتجة من طرح A من كل القيم الاصلية وتكون الصورة الجبرية كما يلي :

$$\overline{X} = A + \frac{\sum (X - A)}{5}$$

وبالإضافة الي الطريقة السابقة فانه يمكن اجراء عملية القسمة بدلا من الطرح ، ولنفرض مثلا اننا نريد ان نحسب متوسط القيم التالية :

200, 300, 500, 900, 600

فيمكنا قسمة كل قيمة علي العدد 100 ويكون المتوسط في هذه الحالة :

$$(\frac{2 + 3 + 5 + 9 + 6}{5}) X\ 100 = 500 = \overline{X}$$

كما انه يمكن استخدام الطريقتين معا وذلك بان نأخذ وسطا فرضيا ونطرحه من كل القيم ثم نقسم بواقي الطرح علي عدد معين من وحدات القياس ، فلحساب المتوسط الحسابي للقيم السابقة نأخذ وسطا فرضيا 100 مثلا ، وبلك تصبح الاعداد 500 ، 800 ، 400 ، 200 ، 100 وبقسمة كل منها علي 100 يكون الوسط الحسابي هو :

$$(\frac{1 + 2 + 4 + 8 + 5}{5}) = 4 = \overline{X}$$

وبذلك يكون المتوسط الحسابي للاعداد الاصلية :

$$X = 100 + (4 X 100) = 500$$

واذا كانت الارقام $X_1 , X_2 , \ldots\ldots X_K$ تحدث $F_1 , F_2 ,$ ......F_K مرة علي الترتيب (بمعني انها تحدث بتكرارات $F_1 , F_2 ,$ ......F_K) فان المتوسط الحسابي سيكون

$$\overline{X} = \frac{F_1 X_1 + F_2 X_2 + \ldots. F_K X_N}{F_1 + F_2 + \ldots.. F_K} = \frac{\sum FX}{F}$$

$$= \frac{\sum FX}{N}$$

حيث $N = \sum F$ هو مجموع التكرارات أي مجموع عدد الحالات .

مثال : اذا كانت 6 , 3 , 3 , 7 , 2 تحدث بتكرارات 3 , 4 , 2 , 1 علي الترتيب فان المتوسط الحسابي سيكون

$$\overline{X} = \frac{(1)(4) + (2)(7) + (4)(3) + (3)(6)}{1 + 2 + 4 + 3} = \frac{48}{10} = 4.8$$

خصائص المتوسط الحسابي :

1. المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفرا .

وهذا ما يتضح من طريقة استخدام وسط فرضي والتي سبق شرحها حيث بينا ان:

$$\overline{X} = A + \frac{\sum (X - A)}{5}$$

حيث A هي الوسط الفرضي ، (X-A) هي مجموع انحرافات القيم عن الوسط الفرضي ، فاذا اخذنا الوسط الفرضي مساويا للوسط الحسابي (أي اذا كانت $\bar{X} = A$) فان :

$$\overline{X} = \bar{X} + \frac{\sum (X - \bar{X})}{N}$$

وهذا لا يأتي إلا إذا كان :

$\frac{\sum (X - \bar{X})}{N} =$ صفرا أي ان :

$\sum \left(X - \bar{X} \right) =$ صفرا .

وهذا يعني ان مجموع الانحرافات عن الوسط الفرضي (وهو الوسط الحسابي في هذه الحالة) = صفرا .

مثال : انحرافات الأرقام 6 , 9 , 5 , 3 , 2 عن وسطها الحسابي 5 هي :

(2-5) , (3-5) , (5-5) , (9-5) , (6-5)

ومجموعها الجبري $= -3 - 2 + 0 + 4 + 1 = 0$

2. مجموع مربعات الانحرافات القيم عن وسطها الحسابي يكون اقل ما يمكن

وهذا يعني ان مجموع مربعات انحرافات مجموعة من الارقام X_i عن أي رقم A يكون اصغر ما يمكن في حالة واحدة فقط اذا كانت $A = \bar{X}$

والا فان $(X - A)^2 < (X - \bar{X})^2$

3. اذا كانت هناك مجموعة مكونة من N من المفردات وهذه المجموعة ذاتها مقسمة الى مجموعتين الاولى مكونة من مفردة والثانية من مفردة فان المتوسط الحسابى للمجموعة الكلية =عدد قيم المجموعة الاولى فى وسطها الحسابى + عدد قيم المجموعة الثانية فى وسطها الحسابى مقسومة على العدد الكلى للمجموعتين .
أي ان

$$\overline{X} = \frac{(\overline{X}_1)N_1 + (\overline{X}_2)N_2}{N_1 + N_2}$$

فاذا كان لدينا محل تجارى يتعامل مع كل من الرجال والسيدات وكان عدد المتعاملين من الرجال 200 رجل ومتوسط ومتوسط مشتريات الرجل 40 جنيه فى حين كان عدد المتعاملين من السيدات 300 سيدة ومتوسط ما تشترية السيدة الواحدة هو 60 جنية. فى هذه الحالة يكون متوسط مشتريات الفرد (المتوسط العام) هو :

$$\frac{(40)(200) + (60)(300)}{200 + 300} = 52$$

وبصفة عامة فان المتوسط العام هو :

$$\overline{x} = \frac{\overline{x}_1 N_1 + \overline{X}_2 N_2 + \ldots\ldots + \overline{X}_n N_n}{N_1 + N_2 + \ldots\ldots\ldots + N_n}$$

$$= \frac{\sum \overline{x}_i N}{\sum N}$$

<u>المتوسط الحسابى لتوزيع تكراري :</u>

عندما تعرض البيانات فى توزيع تكراري ،فان جميع القيم التى تقع داخل فئة معينة تعتبر انها مطابقة لمركز الفئة او منتصف مدى الفئة .

ولايجاد المتوسط الحسابى يوجد طريقتين احدهما مطولة والاخرى لاستخدام وسط فرضى وتسمى الطريقة المختصرة .

1. تعيين مراكز الفئات ، وذلك لجمع الحد الادنى والاعلى للفئة وقسمتها على (2) ويرمز لمركز الفئة X .
2. تضرب مراكز الفئات X_i فى تكرارتها المقابلة F_i

3. تجمع حواصل الضرب $\sum F_i X_i$

4. تقسم حواصل الجمع فى (3) على عدد التكرارات $\sum F_i$

ومنها يكون $\overline{X} = \dfrac{\sum F_i X_i}{\sum X_i}$

مثال : احسب المتوسط الحسابى للبيانات التالية مستخدما الطريقة المطولة مرة والطريقة المختصرة مرة اخرى .

الفئات	6 . 4	8 . 6	12 . 8	14. 12
التكرارات	2	4	3	1

الحل :

اولا : باستخدام الطريقة المطولة ، ويوضحة الجدول التالى :

الفئات	التكرار F_i	مراكز الفئات X_i	FX
4−6	2	5	10
6−8	4	7	28
8−12	3	10	30
12−14	1	13	13
	10		81

$$\overline{X} = \frac{\sum F_i X_i}{\sum X_i} = \frac{81}{10} = 8.1$$

ثانيا : باستخدام الطريقة المطولة (وسط فرضى) :

وتستخدم الصورة الجبرية التالية :

$$\overline{X} = A + \frac{\sum F_i (X_i - A)}{F_i}$$

حيث A وسط فرضى و F هى مجموع التكرارات، وعلى فرض اننا اخترنا قيمة اختيارية كوسط فرضى ولتكن 7 فيكون الحل :

الفئات	التكرار F	مراكز الفئات X	الانحرافات (X – A)	F (X – A)
4−6	2	5	2−	4−
6−8	4	7	0	0
8−12	3	10	3	9
12−14	1	13	6	6
	10			11

$$\overline{X} = 7 + \frac{11}{10} = 8.1$$

(2) المتوسط التوافقى : Harmonic Mean

المتوسط التوافقى لعدد معين من القيم هو المقلوب المتوسط الحسابى لمقلوبات تلك القيم. وبمعنى اخر فانة اذا كان لدينا القيم :

$X_1 , X_2 , X_3 \ldots\ldots X_n$

فان المتوسط التوافقى H لتلك القيم هو :

$$H = \frac{\text{عدد القيم}}{\text{مجموع مقلوبات القيم}}$$

$$H = \frac{N}{\frac{1}{X_1} + \frac{1}{X_2} + \ldots\ldots + \frac{1}{X_n}} = \frac{N}{\sum \frac{1}{X_i}}$$

وهذا يتضح من الصورة الجبرية السابقة ان المتوسط التوافقى يتطلب احتسابة بعض الجهد وهو اكثر تعقيدا من المتوسط الحسابى الا انة رغم ذلك يفضل المتوسط الحسابى فى بعض الحالات الخاصة التى سنوضحا فيما يلى بالامثلة المحلولة .

كما ان الوسط التوافقى يمكن ترجيحية بأوزان وإيجاد للتوزيعات التكرارية وفى هذة الحالة تكون الصورة الجبرية له هى :

$$H = \frac{\text{مجموع التكرارات}}{\text{مجموع حاصل ضرب مقلوب مراكز الفئات } X \text{ التكرارات}}$$

$$H = \frac{F_1 + F_2 + \ldots\ldots + F_n}{F_1\left(\frac{1}{X_1}\right) + F_2\left(\frac{1}{X_2}\right) + \ldots. + F_n\left(\frac{1}{X_n}\right)} = \frac{\sum F_i}{\sum\left(\frac{F_i}{X_i}\right)}$$

وتتمثل خطوات ايجاد المتوسط التوافقى للتوزيعات التكرارية فيما يلى :

1. حساب مراكز الفئات X_i
2. حساب مقلوبات مراكز الفئات $\frac{1}{X}$
3. ضرب مقلوبات مراكز الفئات فى فى تكرارتها المقابلة F_i

وحساب حاصل الجمع($\sum (F_i / X_i)$ ، و باستخدام الصورة الجبرية السابقة يمكن حساب المتوسط التوافقى H .

مثال : اوجد الوسط التوافقى H للقيم 12 , 10 , 7 , 6 , 6 , 5 , 3

$$H = N / \frac{1}{X} = \frac{7}{1/3 + 1/5 + 1/6 + 1/6 + 1/7 + 1/10 + 1/12}$$

$$= \frac{2940}{501} = 5.87$$

ومن الملاحظ ان المتوسط التوافقى لمجموعة من الارقام الموجبة والتى لا تتساوى كلها فى القيمة اقل من الوسط الهندسى والذى بدورة اقل من المتوسط الحسابى .

مثال: اذا انتقل شخص من A الى B بمتوسط سرعة 30 KM / H وعاد من B الى A مستخدما نفس الطريق بمتوسط سرعة 60 KM /H اوجد متوسط السرعة للمرحلة كلها .

الحل :

افتراض ان المسافة من A الى B هى 60 KM /H (على الرغم من انة يمكن فرض اى مسافة اخرى).......... وبهذا

وقت الذهاب من A الى B $= \frac{60\ KM}{30\ KM} = 2\,h$

وقت الذهاب من B الى A $= \frac{60\ KM}{60\ KM} = 1\,h$

$$\frac{\text{المسافة الكلية}}{\text{الوقت الكلي}} = \frac{120\,KM}{3h} = 40\,KM\ /\ h$$

= متوسط السرعة لرحلة ذهاب وإياب

والمتوسط السابق هو المتوسط التوافقى للرقمين H / 30 KM ، 60

وهذا يعنى $\frac{2}{1/30 + 1/60} = 40\ KM\ /\ h$ و اذا كان المسافات المقطوعة ليست كلها متساوية فانة يمكن استخدام المتوسط التوافقى المرجح للسرعات حيث تكون الاوزان هي المسافات .

ومن الملاحظ ان استخدام الوسط الحسابي للرقمين 30 وkm/h 60 هو 45km/h و خطأ وليس هو المقياس المناسب .

مثال : احسب المتوسط التوافقي للبيانات الواردة بالجدول التالي :

الفئات	التكرار f	مراكز الفئات x	مقلوب مركز الفئات $^1/_x$	$F\,(^1/_x)$
2−4	2	3	0.333	0.666
4−6	3	5	0.200	0.600
6−8	9	7	0.143	0.286
8−12	5	10	0.100	0.500
12−16	1	14	0.071	0.071

$$H = \frac{20}{3.124} = 6.41$$

(3) المتوسط الهندسي : Geometric Mean

المتوسط الهندسي للمجموعة من القيم هو الجذر النوني للحاصل ضرب هذه القيم حيث N هي عدد القيم ، بمعني اخر فانه اذا كان لدينا القيم :

$X_1\ ,\ X_2\ ,\ X_3 \ldots\ldots X_N$

فان المتوسط الهندسي (G) لتلك القيم هو :

$$G = \sqrt[N]{(X_1)(X_2)(X_3)\ldots\ldots(X_N)}$$

وباستخدام اللوغاريتمات لتسهيل عمليات الحساب يمكن الوصول إلي الصورة

$$\text{Log } G = \frac{1}{N}(\text{Log } x_1 + \text{Log } x_2 + \ldots\ldots \text{Log } x_n)$$

$$= \frac{\sum \log X_i}{N}$$

وهـذا يعنــي ان لوغــاريتم المتوســط الهندســي لمجموعــة معينة من القيم ما هو الا المتوسط الحسابي لمجموع اللوغاريتمات تلك القيم .

وفي حالة التوزيعات التكرارية او في حالة وجود أوزران للقيم . بأي صورة من الصور فان المتوسط الهندسي يكون علي الشكل التالي :

$$G = \sum f \sqrt{(X^{F1}{}_1)(X^{F2}{}_2).......(X^{FN}{}_N)}$$

$$\text{Log G} = \frac{1}{F}(F_1 \log x_1 + F_2 \log x_2 + ... F_N \log x_n)$$

$$\text{Log G} = \frac{\sum F \log x}{\sum F}$$

حيث F_1 , F_2 ,....F_N تمثل الاوزان او التكرارات للقيم X_1 , X_2 , X_3 ......X_N

مثال: اوجد المتوسط الهندسي والمتوسط الحسابي للارقام 3, 5, 6 , 6 ,7 , 10 , 12 وقارن بينها .

الحل :

ا) الوسط الهندسي $G = \sqrt{(3)(5)(6)(6)(7)(10)(12)} = \sqrt{453600}$

وباستخدام التحويل اللوغاريمي

Log G = 1/7 log 453600 = 1/7 (5.6567) = 0.808

ومنها G = 6.43

ب) الوسط الحسابي $\overline{X}$ = 1/7 (3+5 +6 +6 +7 + 10 + 12) = 7

وهذا يوضح ان المتوسط الهندسي مجموعة من القيم الموجبة غير متساوية اقل من وسطها الحسابي .

مثـال : اذا كانت نسبة سعر المتر من الصوف الي سعر المتر من الأقمشة القطنية هي 3:1 في سنة من السنين بينما بلغت هذه النسبة 2 : 1 في العام التالي

أ) أوجد المتوسط الحسابي لهذه النسب للفترة العامين

ب)أوجد المتوسط الحسابي للنسب أسعار القطن الي أسعار الصوف للفترة العامين .

ج) ناقش التوصية باستخدام الوسط الحسابي للحصول علي متوسط النسب .

د) ناقش ملائمة الوسط الهندسي للحصول علي متوسط النسب .

الحـل :

أ) متوسط نسبة سعر الصوف الي سعر القطن 2.5 = (2.0 + 3.0) 1/2

ب) بما ان نسبة سعر الصوف الي سعر القطن في السنة الاولي هي 3 : 1 فان نسبة سعر القطن الي سعر الصوف هي 0.333 = 1/3 كذلك فان نسبة سعر القطن الي سعر الصوف في السنة الثانية هي 1/2 = 0.50

وبهذا فان متوسط نسبة سعر القطن الي سعر الصوف هو

$$1/2\ (0.33 + 0.50\ ) = 0.417$$

ج) من الملائم ان نتوقع ان متوسط نسبة سعر الصوف الي سعر القطن هو مقلوب متوسط نسبة سعر القطن الي سعر الصوف

وذلك اذا كان المتوسط متوسطا ملائما . ولكن $1 / 0.417 = 2.4 \neq 2.5$

وهنا يظهر ان المتوسط الحسابي يعد متوسطا غير جيد عند استخدام النسب .

د) الوسط الهندسي لنسب سعر الصوف الي سعر القطن

$$\sqrt{(3)(2)} = \sqrt{6}$$

الوسط الهندسي لنسب سعر القطن الي سعر الصوف

$$\sqrt{(0.3)(0.5)} = \sqrt{0.017} = 1/\sqrt{6}$$

وبما ان هذه المتوسطات كل منها مقلوب الاخر ، فاننا نستنتج ان المتوسط الهندسي اكثر ملائمة من الوسط الحسابي للحصول علي وسط النسب في مثل هذا النوع من المسائل .

وعموما يلاحظ ان الوسط الهندسي يفضل استخدامه عن جميع مقاييس التوسط الاخري في حالة احتساب متوسط النسب او معدلات الزيادة او النقص سواء في الانتاج او السكان او النمو وكذلك في حالة احتساب الارقام القياسية .

(4) الوسيط Median

يعرف الوسيط علي انه قيمة التي تقسم مجموعة من القيم الي قسمين بحيث يكون عدد ما قبلة من القيم مساوي لعدد ما بعدها كذلك بعد ترتيب القيم تصاعديا او تنازليا ، وبذلك فهو يتوسط القيم من حيث الموقع ويكون الوسيط هو القيمة الوسطي اذا كان العدد فرديا ومتوسط القيمتين الوسيطتين اذا كان عدد القيم زوجيا . وعلي العموم فان كان لدينا عدد N من القيم وكانت N عددا فرديا فان الوسيط هو القيمة التي ترتبها $\frac{N+1}{2}$ بعد ترتيب القيم .

اما اذا كان عدد القيم زوجيا فان الوسيط هو متوسط قيمتين الاولي ترتبها $\frac{N}{2}$ الثانية ترتبها $\frac{N+2}{2}$

ومن الملاحظ ان الوسط لا يتأثر بالقيم المترفة اذا لا تدخل قيمتها في حسابة بينما كان المتوسط الحسابي يتحيز لهذه القيم ولذلك فاستخدام الوسيط في هذه الحالة افضل خاصة اذا كان عدد القيم كبير ، كما ان الوسيط ينصح باستخدامه اذا ما كانت بعض القيم الكبري او الصغري (او هما معا) غير معروفه . وكذلك في حالة الجد وال التكرارية المفتوحة .

مثال :

احسب الوسيط لكل مجموعة من المجموعات التالية :

المجموعة الاولي درجات طالب في سنة امتحانات كانت كالتالي :

84 , 91 , 72 , 68 ,87 , 78

المجموعة الثانية الأجر بالساعة لخمسة عاملين في مكتب :

2 ,3 ,4 ,5 , 6

الحـل :

المجموعة الأولي : بترتب الدراجات تصاعديا تصبح

68 , 72 , 78 , 84 , 91

وبما ان عدد الدراجات زوجي فان هناك قيمتين في المنتصف 78 , 84

وسطها الحسابي $\frac{78+84}{2}=81$ وهو الوسيط **المطلوب .**

المجموعة الثانية : الأجور مرتبة ، وبما ان هناك عددا فرديا من القيم فان هناك قيمة واحدة في المنتصف وهي 4 وهي الوسيط المطلوب .

مثـال :

أوجد الوسيط والمنوال للجدول التكرار التالي بطريقتين .

يمكن إيجاد الوسيط اما باستخدام الرسم البياني عن طريق إسقاط عمود من نقطة تقاطع المنحني المتجمع الصاعد بنظيره الهابط علي المحور الأفقي (محور الفئات) .وتكون القراءة مسقط علي المحور الأفقي هي قيمة الوسيط مباشرة ، وقد سبق توضيح الجزء الخاص بعرض البيانات .

اما الطريقة الحسابية فيوضحها الجدول التالي :

تكرار متجمع هابط	الحدود الدنيا	تكرار متجمع صاعد	الحدود العليا	التكرارات f	الفئات
15	اكبر من 2	2	اقل من 4	2	2−4
13	، ، 4	6	، ، 6	4	4−6
9	، ، 6	11	، ، 8	5	6−8
4	، ، 8	14	، ، 10	3	8−10
1	، ، 10	15	، ، 12	1	10−12

ولتعين موقع الوسيط تقسم مجموع التكرارات علي 2

$$\frac{\sum F}{2} = \frac{15}{2} = 7.5$$

وبالنظر الي الجدول السابق نجد ان القيمة التي ترتيبها 7.5 في الفئة الثالثة (8−6) أي ان الوسيط قيمة اكبر من 6 واقل من 8 ، وبفرض ان القيمة موزعة بانظام داخل الفئات فيمكن احتساب الوسيط كما يلي :

فرق في الفئات من (6−8) يقابلة فرق في التكرارات (6−11)

فرق في الفئات X يقابلة فرق في التكرارات (6−7.5)

$$X = \frac{(1.5)(2)}{5} = \frac{3}{5} = 0.6$$

.. الوسيط 6 + 0.6 = 6.6

حيث الوسيط الحد الادني للفئة الوسيطة + X

(5) المنوال Mode

المنوال لمجموعة من القيم هو القيمة التي تتكرر اكثر من غيرها أي القيمة الاكثر شيوعا .

فاذا كانت اجور عدد معين من العمال هي : 37 , 30 , 36 , 30 , 35 , 30 فان المنوال هو 30 .

الا انه اذا كانت مجموعة القيم صغيرة فاحيانا لا تتكرر احدي القيم عددا من المرات يسمح باعتارها منوالا . فمجموعة الاجور 36 , 50 , 30 , 45 , 40 ليس منوالا .

وعموما فان المنوال يمتاز عن باقي مقاييس التوسط بسهولة حسابة سواء بالرسم او بالحساب ويفضل حيثما كان المطلوب معرفة القيمة الشائعة كما في حالات التسويق حيث يكون المطلوب معرفة بيانات عن السلع والمواصفات الخاصة بالشائع منه او الاكثر مبيعا مثل اكثر مقاسات الاحذية والملابس مبيعا . وذلك حتي يمكن ان ينصح بانتاج ما يكفي منها لزيادة المبيعات الكلية وتلبية رغبات المستهلكين .

والمنوال مثل الوسيط لا يتاثر بالقيم الشاذة ويمكن حسابه للتوزيعات المفتوحة الا انه يجب تعديل التكرارات عند ايجاد المنوال اذا كانت اطوال الفئات غير متساوية .

ومن العيوب المنوال تعدد قيمة في بعض الاحوال ، فبعض التوزيعات التكرارية لها اكثر من قيمة منوالية واحدة .

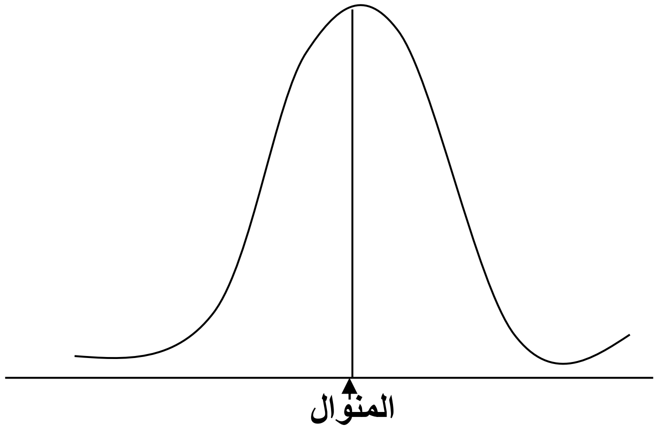
ويعتبر المنوال افضل المتوسطات لتمثيل البيانات الغير رقمية ، فاذا كان لدينا توزيع تقديرات 100 طالب في احدي المواد الامتحان كالاتي :

ضعيف جدا	ضعيف	مقبول	جيد	جيد جدا	ممتاز
2	9	42	27	16	4

فان المنوال هو تقدير مقبول لانه تقدير اكبر عددا من الطلبة ويجب ان يلاحظ ان المنوال لا ينصح باستخدامه في حالة التوزيعات الشديدة الالتواء لأنه في هه الحالة يبعد كثيرا عن وسط المجموعة ويصبح من القيم المتطرفة وليس من القيم المتوسطة .

<u>حساب المنوال بالرسم:</u>

1- عن طريق المنحني التكراري : يسقط عموداً من قمة المنحنى التكراري على المحور الأفقي فتكون نقطة تقاطع هذا العمود مع المحور الأفقي هى قيمة المنوال .

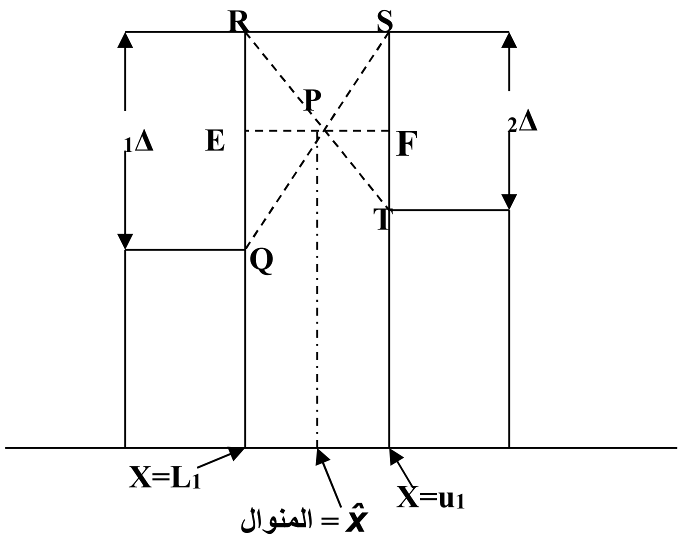


ويلاحظ أنـه فـى المنحنيـات المتماثلـة ينطبـق كـل مـن الوسـط الحسـابي والوسـيط والمنوال ا، فتكون العلاقة بينهم كما يلي:

$$\text{المتوسط الحسابي} = \frac{(3)\ \text{الوسيط} - \text{المنوال}}{2}$$

2- عن طريق المدرج التكراري :

ويكتفي فيه برسم المستطيل الذى يمثل أكبر التكرارات والمستطيلين المحيطين بـه، ثـم نصـل الـرأس الأيمـن العلـوي للمسـتطيل الأطـول بـالرأس السـمن العلـوي للمستطيل السابق لـه. (SQ فى الشكل). وكذلك نصل لرأس اليسر العلوي لأطول مستطيل بالرأس الأيسر العلوي للمستطيل التالي له (RT) فيتقاطع المستقيمان فى نقطـة P ، وإذا أسقطنا منها عمودا يقابل المحور الأفقي فى X فإن البعد السيني للنقطة X يكون هو قيمة المنوال .



ولايجاد قيمة المنوال حسابيا تستخدم الصيغة الجبرية التالية :

$$\text{المنوال} = X = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)C$$

حيث : L_1 = الحد الادني للفئة المنوالية .

Δ_1 = الفرق بين تكرار الفئة المنوالية والفئة التي قبلها .

Δ_2 = الفرق بين تكرار الفئة المنوالية والفئة التي تليها .

C = طول الفئة المنوالية

ويمكن اشتقاق الصيغة السابقة كما يلي :

يوضح الشكل السابق ثلاثة مستطيلات من المدرج التكراري ويمثل المستطيل الأوسط الفئة المنوالية ، وعلي فرض أن طول الفئات متساوية . وكما سبق توضيحه بيانيا فان المنوال هو النقطة $\hat{X}$ المحور السيني المقابلة للنقطة P وهي نقطة تقاطع الخطيين RT , QS إذا كانت $X = L_1$, $X = U_1$ تمثل الحدود الدنيا والعليا للفئة المنوالية و Δ_2 ، Δ_1 يمثلان علي الترتيب الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي علي يسارها والفئة التي علي يمينها ، فانه من المثلثات المتشابهة RST , PQR نجد

$$\frac{X - L_1}{\Delta_1} = \frac{U_1 - X}{\Delta_2} \quad \text{أو} \quad \frac{PF}{ST} = \frac{EP}{RQ}$$

$$\Delta_2(\hat{X} - L_1) = \Delta_1(U_1 - \hat{X}), \Delta_2\hat{X} - \Delta_2 L_1 = \Delta_1 U_1 + \Delta_1\Delta_2, (\Delta_1 + \Delta_2)X$$
$$= \Delta_1 U_1 + \Delta_2 L_1$$

إذن

$$\hat{X} = \frac{\Delta_1(L_1 + C) + \Delta_2 L_1}{\Delta_1 + \Delta_2} = \frac{(\Delta_1 + \Delta_2) L_1 + \Delta_1 C}{\Delta_1 + \Delta_2}$$

أو

$$\hat{X} = \frac{\Delta_1 U_1 + \Delta_2 L_1}{\Delta_1 + \Delta_2} \qquad = L_1 + (\frac{\Delta_1}{\Delta_1 + \Delta_2})C$$

مثال : اوجد القيمة المنوالية للبيانات التالية :

الفئات	التكرارات	التكرار المعدل
2−4	2	1
4−6	3	1.50
6−8	9	4.50
8−12	5	1.25
12−16	1	0.25

$$\text{يلاحظ أن التكرار المعدل} = \frac{\text{التكرار الاصلي}}{\text{طول الفئة}}$$

$$\text{المنوال} = 6 + (\frac{3}{3 + 3.25}) X 2 = 6.9$$

ثانياً: مقاييس التشتت والتشتت النسبي

Measures of dispersion and relative Dispersion

لاشك أن مقاييس التوسط تعطينا فكرة عن التوزيع التكراري إلا أن هذه الفكرة لا تكون كاملة إذ أن الوسط وحدة لا يكفي لإعطاء فكرة دقيقة عن المجموعة فلا يبين طبيعتها ولا كيفية توزيع مفرداتها. وكثيرا ما تتساوى القيمة الوسطية لمجموعتين بينما تختلف المجموعتان عن بعضها كل الاختلاف، فقد تكون مفردات إحدى المجموعتين متقاربة بعضها عن بعض (أى تتركز حول متوسطها) أو مبعثرة(مشتتة).

والتشتت هو الدرجة التي تتجه بها البيانات الرقمية للانتشار حول قيمة وسطى، وهناك عديد من مقاييس التشتت يمكن استخدامها وأهمها وأكثرها انتشارا المدى، الانحراف الربيعي (نصف المدى الربيعي)، الانحراف المتوسط والانحراف المعياري.

وأفضل مقاييس التشتت هو الذى تتوافر فيه بعض الخصائص مثل التي سبق ذكرها عند الكلام على المتوسطات ومنها سهولة فهم المقياس، وأن يأخذ فى الاعتبار جميع المفردات عند حسابه، وقلة تأثره باختيار العينة، وإمكان وضعه فى صورة جبرية تخضع للعمليات الرياضية المختلفة.

(1) **المدى Range:**

مدى مجموعة من الأرقام هو الفرق بين أكبر رقم وأقل رقم فى المجموعة .فإذا كانت لدينا القيم 15، 20، 30، 35، 40 فإن المدى 25=15−40 وهذا يعني أن قيم هذه المجموعة تتشتت فى مدى قدرة من 15 إلى 40 أى 25 بينما نجد أن المدى لمجموعة القيمة 30،35،40،45،50 هو 20 =30−50 أى أن قيم هذه المجموعة تتشتت فى مدى قدرة من 50 إلى 30 أى 20. وعلى هذا يمكن القول أن المجموعة الأولى أكثر تشتتا من المجموعة الثانية بمعنى أن مفردات المجموعة الأولى أكثر تبعثراً من مفردات المجموعة الثانية.

ويعيب المدى كمقياس من مقاييس التشتت رغم سهولة حسابه أنه يتوقف فى حساب قيمته على قيمتين طرفيتين أحدهما أو كلاهما يمكن أن تكون شاذة، وفى هذه الحالة يكون هذا المقياس مضللا ولا يعبر تماما عن درجة التشتت. كما يعيبه أيضا عدم إمكانية حسابه للتوزيعات التكرارية المفتوحة.

ويمكن التخلص من العيب السابق للمدى وذلك بأن نستبعد أصغر وأكبر قيمة من البيانات ثم نوجد المدى للقيم الباقية ويسمى المدى الناتج من هذه العملية بشبيه المدى Quasi-range فإذا كانت لدينا N من القيم وكانت مرتبة ترتيبا تصاعديا X_1،X_2،X_3....X_n فإن مدى البيانات ويسمى المدى الأول هو يساوي $X_{n-1}-X_2$ ويمكن بنفس الطريقة إيجاد المدى الثاني $X_{n-2}-X_3$ وهكذا .

(2) الإنحراف الربيعي: (نصف المدى بين الربيعين)

Semi Inter-Quartile Range

الإنحراف الربيعي هو نصف المدى بين الربيع الأعلى والربيع الأولى ، ويعرف كما يلي:

$$\text{الانحراف الربيعي} \quad Q = \frac{Q_3 - Q_1}{2}$$

حيث Q_1 هو الربيع الأول و Q_3 هو الربيع الثالث للبيانات. هذا ويستخدم المدى الربيعي Q_3-Q_1 فى بعض الأحيان بدلا من نصف المدى الربيعي كمقياس للتشتت والربيع الأول Q_1 هو القيمة التي يسبقها فى الترتيب ربع القيم ويليها 75% من القيم بغرض ترتيب هذه القيم ترتيبا تصاعديا. وترتيب الربيع الأول فى التوزيعات التكرارية هو $n/4$. ولاحتساب الربيع الأول للمثال الوارد فى مناقشة الوسيط يكون ترتيب.

$$3.75 = \frac{15}{4} = \frac{n}{4} = Q_1$$

ومن الواضح أن القيمة التي ترتيبها 3.75 تقع فى الفئة (6–4) وبنفس الطريقة يمكن إيجاد القيمة الوسيطية وهى القيمة التي ترتيبها $n/2$ ويمكن إيجاد قيمة الربيع الأول كالآتي:

وفرق فى التكرارات (3.75–2) بناظره فرق فى الفئات x

$$X = \frac{(2)(1.75)}{4} = \frac{3.50}{4} = 0.87$$

ويمكن الحصول على نفس النتيجة باستخدام المنحنى التكراري المتجمع الصاعد أو النازل.

الربيع الأعلى (Q_3) هو القيمة التي ترتيبها $\frac{3N}{4}$ وهو المثال السابق

$$\frac{(3)(15)}{4}=11.25$$

ويمكن إيجاد الربيع الثالث بنفس طريقة إيجاد الوسط (الربيع الثاني)

فرق فى التكرار (11−14) يقابله فرق فى الفئات (8−10)

فرق فى التكرار (11−11.25) يقابله فرق فى الفئات X

$$Q_3 = \frac{(2)(0.25)}{3} = 0.16$$

$$Q_3 = 8 + 0.16 = 8.16$$

وبذلك يمكن حساب الانحراف الربيعي كما يلي:

الانحراف الربيعي= $\boldsymbol{\frac{8.16 - 4.87}{2} = \frac{3.29}{2} = 1.65}$

والانحراف الربيعي يشبه إلى حد كبير شبيهات المدى حيث أنه لا يتأثر بالقيم الشاذة كما أنه مقياس تقريبي لتفتت المفردات وهو يفضل عن المدى المطلق لهذا السبب بالإضافة إلى أنه يمكن حسابه للتوزيعات التكرارية المفتوحة.

(3) الانحراف المتوسط: (متوسط القيم المطلقة للانحرافات)

Mean Absolute Deviation

الانحراف المتوسط هو متوسط انحرافات قيم المفردات عن وسطها الحسابي بغض النظر عن إشارات الانحرافات

والانحراف المتوسط لمجموعة من الأرقام X_1،X_2،X_3....X_n يعرف بما يلي:

الانحراف المتوسط= $M.D = \frac{\sum |X - \bar{X}|}{N}$

حيث $\bar{X}$ هو الوسط الحسابي للأرقام و $|X - \bar{X}|$ هو القيمة المطلقة لانحراف القيمة X عن $\bar{X}$ (القيمة المطلقة لرقم ما هو الرقم بدون الإشارة المرافقة له ويعبر عن ذلك بخطين راسيين يوضعان حول الرقم).

مثال: أوجد الانحراف المتوسط للأرقام 6،7 ،5 ،4 ،3 ويكون الحل كما يلي:

المتوسط الحسابي $= \overline{X} = \frac{3+4+5+6+7}{5} = 5$

الانحراف المتوسط $= M.D = \frac{|3-5|+|4-5|+|5-5|+|6-5|+|7-5|}{5}$

$= \frac{2+1+0+1+2}{5} = 1.2$

وفى حالة التوزيعات التكرارية فإن:

الانحراف المتوسط $= M.D = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$

حيث $\sum f_i$ هو مجموع التكرارات و x_i تمثل مراكز الفئات وفى بعض الأحيان يعرف الانحراف المتوسط بدلالة القيمة المطلقة للانحرافات عن الوسيط أو غيره من المتوسطات بدلا من المتوسط الحسابي وهناك تكون لهذا المقياس خاصية هامة وهى أن متوسط انحرافات القيم عن الوسيط أقل ما يمكن. وهذا يعني أن $|X_i - \delta|$ يكون أقل ما يمكن عندما تكون هى الوسيط.

\(4) الانحراف المعياري : Root Mean Square Deviation

هو الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابي ، ويسمى أحياناً جذر متوسط مربع الانحراف.

فإذا كانت لدينا مجموعة من القيم عددها n وهى: $X_1, X_2, X_3 .... X_n$ فإن الانحراف المعياري والمعبر عنه بالرمز يعرف بما يلي :

$$S = \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

وفى التوزيعات التكرارية يكون الانحراف المعياري كما يلي:

$$S = \sqrt{\frac{\sum f_i (x_i - \bar{x})}{\sum f_i}} = \sqrt{\frac{\sum f_i (x - \bar{x})^2}{N}}$$

حيث $N = \sum f_i$

وفى بعض الأحيان يعرف الانحراف المعياري لبيانات من عينة بالقسمة على (N-1) بدلا من N لأن هذا يؤدى للحصول على تقديرات أفضل للانحراف

المعياري للمجتمع الذى سحبت منه العينة. ولقيم N الكبيرة وهى غالبا أكبر من 30($N < 30$) فإنه من الناحية العملية لا يوجد فرق حقيقي بين التعريفين.

(5) التباين Variance

تباين مجموعة من البيانات يعرف بأنه مربع الانحراف المعياري وبهذا يعرف بـS2 وعندما يكون ضروريا التمييز بين الانحراف المعياري للمجتمع والانحراف المعياري لعينة مسحوبة من هذا المجتمع، فإننا نستخدم دائما الرمز s للأخير، σ للأول ، وبهذا فإن S2 ، σ^2 يمثلان تباين العينة وتباين المجتمع على الترتيب.

والانحراف المعياري يعتبر من أهم مقاييس التشتت وأكثرها انتشاراً وترجع أهميته إلى كونه كثيرا ما يدخل فى إحتساب عدد كبير من المقاييس الإحصائية الأخرى كمقاييس الالتواء ، والتفرطح. وهو أيضا من أدق مقاييس التشتت.
وعند حساب الانحراف المعياري يجرى أولا إحتساب المتوسط الحسابي ، ثم تعيين انحرافات قيم المتغير عن المتوسط الحسابي، ثم تربيع هذه الانحرافات والترجيح بالتكرارات المقابلة فى حالة التوزيعات التكرارية، ويلي ذلك جمع هذه المربعات وقسمتها على عدد القيم أو على عدد الأوزان، ثم نستخرج الجذر التربيعي لخارج القسمة فيكون هو الانحراف المعياري.
طريقة مختصرة لحساب الانحراف المعياري :وهى كما يلي:

$$S = \sqrt{\frac{\sum x_i^2}{N} - \left(\frac{\sum X_i}{N}\right)^2}$$

وفى حالة التوزيعات التكرارية

$$S = \sqrt{\frac{\sum f_i x_i^2}{N} - \left(\frac{\sum f_i X_i}{N}\right)^2}$$

وعن طريق إستخدام وسط فرضي (A) بدلا من المتوسط الحسابي فإذا كانت $d_i = X_i - A$ هى انحرافات X_i عن ثابت اختياري A فتصبح الصيغتين السابقتين كما يلي على الترتيب.

$$S = \sqrt{\frac{\sum d_i^2}{N} - \left(\frac{\sum d_i}{N}\right)^2}$$

وفى حالة التوزيعات التكرارية

$$S=\sqrt{\frac{\sum f_i d_i^2}{N}-\left(\frac{\sum f_i d_i}{N}\right)^2}$$

ومن خصائص الانحراف المعياري ما يلي:

$$\sqrt{\frac{\sum\left(X_i-\bar{x}\right)^2}{N}}\leq\sqrt{\frac{\sum\left(X_i-\delta\right)^2}{N}}$$

حيث δأى وسط فرضي وهذا يعني أن أصغر انحراف معياري يمكن الحصول عليه عندما $\delta=\overline{X}$. وهذا يرجع لإحدى خصائص المتوسط الحسابي السابق ذكرها.

(2) إذا افترضنا أن مجموعتين مكونتان من N_1 ,N_2 رقم (أو توزيعات تكراريان ومجموع تكراراتها هى N_1 ,N_2 ، وتباينها يعرف بـS_1 ,S_2 على الترتيب ولها نفس الوسط الحسابي $\overline{X}$ ، فإن التباين المشترك أو المجمع للمجموعتين (أو للتوزيعين التكراريين) هو

$$S=\frac{N_1 S_1^2+N_2 S_2^2}{N_1+N_2}$$

ويلاحظ أن هذا هو الوسط الحسابي المرجح للتباين . وهذه النتيجة يمكن تعميمها لحالة ثلاثة أو أكثر من التباينات.

وفيما يلي مثالين لاحتساب الانحراف المعياري لمجموعة من البيانات الأولية وآخر للتوزيعات التكرارية.

<u>مثال:</u>

أوجد الانحراف المعياري للقيم التالية بالطريقتين المطولة والمختصرة:

10،12،15،11،14،13،17،16،20

الحل :

المتوسط الحسابي $\overline{\mathrm{X}}=\frac{137}{10}=13.7$

والجدول التالي يوضح باقي خطوات الحل :

قيم المتغير X	الانحراف عن $\overline{X}$ (d)	(d^2)	قيم المتغير x	الانحراف عن $\overline{X}$ (d)	(d^2)
10	−3.7	13.69	4	0.3	0.09
12	−1.7	3.89	13	−0.7	0.49
15	1.3	1.69	17	3.3	10.89
9	−7.7	22.09	16	2.3	5.29
11	−2.7	7.29	20	6.3	39.69

$$\sqrt{\frac{\sum\left(X-\bar{X}\right)^2}{N}}=\sqrt{\frac{\sum d^2}{N}}$$

$$=\sqrt{\frac{104.1}{10}}=\sqrt{10.41}$$

ولتسهيل العمليات الحسابية يستخدم وسط فرضي وتحسب منه متوسط مربعات الانحرافات وكذلك قيمة الانحراف المعياري.

يوضح الجدول التالي خطوات حساب الانحراف المعياري لتوزيع تكراري

الفئات	التكرار f	مركزالفئة x	$\left(X-\bar{X}\right)$ d	d^2	F d^2
10−15	10	12.5	−17.79	316.48	3164.8
15−20	15	17.5	−12.79	163.58	2453.7
20−25	20	22.5	−7.79	60.68	1213.6
25−30	30	27.5	−2.79	7.78	233.4
30−35	40	32.5	2.21	4.88	195.2
35−40	35	37.5	7.21	51.98	1819.3
40−45	10	42.5	12.21	149.08	2236.2
45−50	5	47.5	17.21	296.18	1480.9
$\sum$	170				**136250.3**

بالطريقة المطولة وهى كما يلي:

$$\bar{X}=\frac{\sum fx}{f}=\frac{5150}{170}=30.29$$

$$\sigma=\sqrt{\frac{\sum f\left(X-\bar{X}\right)}{\sum f}}=\frac{13625.3}{170}=8.7$$

ولحساب الانحراف المعياري بالطريقة المختصرة تستخدم الصورة التالية عن طريقة إستخدام وسط فرضي.

$$\sigma=\sqrt{\frac{\sum fd^2}{\sum f}-\left(\frac{\sum fd}{\sum f}\right)^2}$$

حيث d=(X-A) وحيث A وسط فرضي

ومنها

$$\sigma = \sqrt{\frac{1362.5}{170} - \left(\frac{-375}{170}\right)} = 8.7$$

وهى نفس النتيجة السابق الحصول عليها.

(5) التشتت النسبي Relative Dispersion

إذا كان لدينا عينة من المزارع متوسط إنتاجها من القمح 5 إردب للفدان وانحراف المعياري لها هو أردبين بينما كانت هناك مجموعة من المزارع متوسط إنتاجها 525 أردب وانحرافها المعياري ثلاثة أردب. ولمقارنة تشتت هاتين المجموعتين يلزم إستخدام معامل التشتت النسبي وهو عبارة عن انحراف المعياري لأي مجموعة مقسوما على متوسطها الحسابي:

$$\text{معامل التشتت النسبي} = \frac{\text{الانحراف المعياري}}{\text{المتوسط الحسابي}} = \frac{\delta}{\overline{X}}$$

وباحتساب هذا المعامل بالنسبة للمجموعة الأولى فى المثال السابق نج أنه 4% بينما فى المجموعة الثانية 5.7% وفى هذه الحالة يمكن القول أن المجموعة الثانية أكثر تشتتا من المجموعة الأولى.

أما فى حالة التوزيعات التكرارية المفتوحة والتي يتعذر فيها إحتساب انحرافها المعياري فيستخدم معامل الاختلاف لقياس التشتت النسبي. ومعامل الاختلاف فى هذه الحالة يحسم بقسمة الانحراف الربيعي على الوسيط أى

$$\text{معامل الاختلاف} = \frac{\text{الانحراف الربيعى}}{\text{الوسيط}}$$

الإلتـواء

يعتبر التوزيع ملتويا إذا لم يكن متماثلا، والتوزيع المتماثل هو ذلك التوزيع الذى يقسمه وسطه الحسابي تقسيما متماثلا، أى أن الوسط الحسابي يقع على يمينه نصف المساحة الواقعة تحت المنحنى تماما كما يبدو ذلك واضحا فى الأشكال السابقة. ويقاس الالتواء إما باستخدام الوسط الحسابي والوسيط والانحراف المعياري كما يلي:

$$\text{الالتواء} = \frac{\text{(الوسط الحسابى – الوسيط)}}{\text{الانحراف المعياري}}$$

فإذا كان الناتج موجبا دل ذلك على أن التوزيع ملتوي إلى اليمين أما إذا كان سالبا فان هذا دليل على أن التوزيع سالب الالتواء، أو ملتوي إلى اليسار. أما إذا كان التوزيع متماثلا فان قيمته تكون صفراً. ومن الوجهة النظرية فانه لا يجوز أن تتجاوز قيمة معامل الالتواء ± 3 إلا أنه نادرا ما يصادفنا توزيعات يزيد أو يقل معامل التوائها عن ± 1

وفى التوزيعات التكرارية المفتوحة يستخدم المقياس التالي لقياس الالتواء:

$$\text{الالتواء} = \frac{\text{الربيع الأعلى + الربيع الأدنى} - (2)\ \text{الوسيط}}{\text{الربيع الأعلى – الربيع الأدنى}}$$

ومن الوجهة النظرية فان قيمة هذا المعامل لا تزيد عن ± 1 وبالمثل فإذا كانت قيمته صفراً فان التوزيع يعد متماثلا، أما أن كان موجبا فان ذلك يدل على أن التوزيع ملتوي إلى اليمين، وبالعكس فان كانت قيمته سالبة دل ذلك على أن التوزيع ملتوي إلى اليسار.

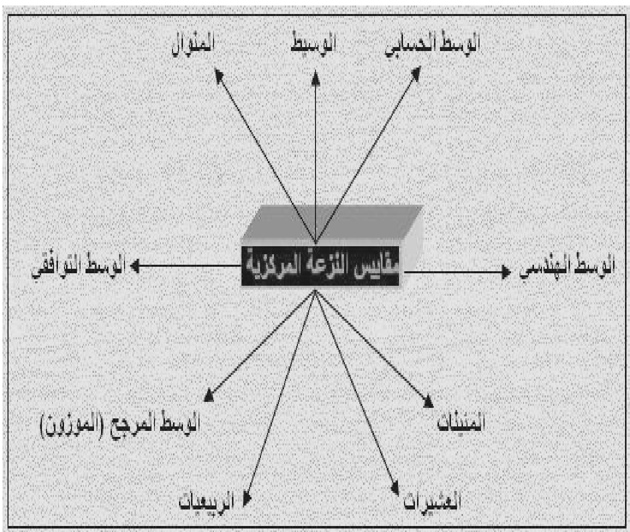
تذكر

ملخص مقاييس النزعة المركزية

مجموع مفردات العينة على عددها

القيمة أو القيمة أو القيمتان التي تتوسط المجموعة بعد ترتيب مفرداتها

القيم الأكثر شيوعاً



الجذر النوني لضرب قيم عددها ن

نقطة تقسم التوزيع إلى مئة قسم

مجموع مقلوبات القيم على عددها

ناتج قسمة مجموع كل قيمة مضروباً في وزنها على مجموع الأوزان)

النقاط التسع التي تقسم التوزيع لعشرة أقسام متساوية

النقاط الثلاث التي تقسم التوزيع لأربع أقسام متساوية

- من خواص المتوسط الحسابي
- يعتمد على جميع القيم والمشاهدات محل الدراسة
- هو نقطة اتزان المشاهدات

- انحراف الدرجة عن المتوسط يساوي بعدها عنه.
- المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفراً
- مربع الانحرافات اقل ما يمكن عن الوسط
- إضافة قيمة ثابتة إلى الدرجات أو طرحها منها أو ضربها فيه أو قسمتها عليه مما يجعل المتوسط الحسابي يزداد أو يقل بقيمة ثابتة.
- المتوسط الحسابي يتأثر بالدرجات القريبة منه تأثرا قليلا ، بينما يتأثر بالدرجات البعيدة عنه تأثرا كبيرا .
- اقل مقاييس النزعة المركزية تأثرا بالتقلبات العينية
- يتأثر بالقيم المتطرفة والقيم الشاذة لذا لا يصلح للتوزيعات الملتوية وخاصة لتوزيعات التكرارية شديدة الالتواء.

<u>لا يصلح المتوسط الحسابي في الحالات التالية</u>

1- عندما تكون قيم التوزيع متجمعة في طرف واحد أكثر من الطرف الأخر.
2- إذا كانت هناك قيم شاذة تتخلل التوزيع.
3- إذا كانت البيانات مبوبة في فئات وكان التوزيع مفتوحا في أحد طرفيه.
4- إذا كانت الفئات متباعدة نسبيا.
5- لا يمكن إيجاده من الرسم البياني كالوسيط والمنوال.
6. لا يمكن إيجاده من التوزيع التكراري المفتوح إلا في حال تقدير مراكز الفئات المفتوحة.
7. عدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.

مميزات الوسيط

أ) عدم تأثره بالقيم الشاذة.
ب) تعتمد قيمته على التكرارات.
ج) يوجد في كل البيانات التي تتصف بالترتيب

عيوب الوسيط

أ) يهتم بالقيم الوسطى للبيانات فقط

ب) لا يصلح لإعطاء فكرة عن النزعة المركزية للبيانات في فئات متباعدة

ج) لا يسهل التعامل معه في التحاليل الإحصائية في دراسة ظاهرة ما في عدة عينات.

تمارين

أولاً: تمارين علي مقاييس النزعة المركزية

1. احسب الوسط الحسابي والمنوال والوسيط للتوزيع التكراري الآتي :

الفئات	100 .	150 .	200 .	250 .	300 .	350 .	400 .	450
لتكرار	8	12	20	30	15	10	5	

2. الجدول الآتي يبين توزيعا تكراريا للأجور اليومية بالجنيه لعمال احد المصانع:

الأجر :	2.46 .	2.8 .	2.2 .	3.6 .	4.0 .	4.4 .	4.8 .	5.2
عدد العمال :	40	220	275	200	75	60	20	10

ارسم منحني لهذا التوزيع ومنه أوجد قيمة الوسيط والربيعين ثم حقق الناتج بالطرق الحسابية . استعمل القيم الناتجة في قياس درجة التواء المنحني التكراري للتوزيع المذكور .

3. الجدول الآتي يبين توزيع 600 منشاة تجارية موزعة حسب راس المال المستثمر فيها بالألف جنيه :

فئات راس المال:	0 .	50 .	100 .	200 .	500 .	1000 .	1000
عدد المنشات :	62	12	16	30	300	180	

والمطلوب :

أ)تمثيل هذا التوزيع بيانيا .

ب) حساب الوسط الحسابي والوسيط لرأس المال المستثمر وشرح سبب الفرق بين هذين المتوسطين .

ج) حساب مقدار الالتواء للتوزيع .

4. الجدول التالي يمثل متوسط الأجر ومتوسط ساعات العمل وعدد العمال وعدد المصانع في الجمهورية والمطلوب حساب المتوسط العام للاجور وساعات العمل

المناطق	متوسط	متوسط	عدد	عدد

	الأجر	ساعات العمل	العمال	المصانع
القاهرة	137	52	31.857	5.445
الإسكندرية	156	54	15.823	2.250
القنال	174	55	1.198	0.402
السويس	229	59	2700	0.217
دمياط	150	50	1900	145
الوجه البحري	103	49	46.896	6.409
الوجه القبلي	110	53	19.228	4.644
الحدود	150	50	2.123	10

5. فيما يلي عدد السكان في احدي الدول خلال الفترة من 75 . 1985
والمطلوب احتساب معدلات النمو السكاني فيما بين كل تعدادين ثم حساب متوسط هذه المعدلات بأفضل المتوسطات .

السنوات	75	76	77	78	79	80	81	82	83	84	85
السكان	7	9	11	13	14	15	17	20	22	25	30

(5): أحسب الوسط الهندسي للبيانات المبينة في الجدول الآتي:

ntervals (X)!	Frequency (f!)
20 – 39	20
40 – 59	35
60 – 79	80
80 – 99	40
100 – 119	15
120 or more	10
Total	200

6. الجدول التالي يمثل عدد وحدات البيض في 50 من محلات البقالة والمطلوب حساب متوسط سعر البيض مستخدما انسب مقياس .

عدد البيض / 25 قرش	2	3	4	5	6
عدد المحلات	5	7	18	11	4

7. قامت سيارة بقطع مسافات التالية بالسرعات المذكورة احسب متوسط سرعة

السيارة في الساعة بأنسب مقياس .

السرعة كيلو/ ساعة	30	40	60	55	70	80	85
مسافة بالكيلو متر	15	20	30	40	52	18	10

8. أوجد قيمة الوسيط للتوزيع التكراري التالي مستخدما في ذلك ثلاثة طرق :

فئات	7 . 5	9 . 7	11 . 9	13 . 11	15 . 13	17 . 15
لتكرار	3	5	12	15	4	2

9. الجدول التالي يبين توزيع مصانع مصر وفقا لعدد العمال والمطلوب حساب متوسط عدد العمال ثم تحقيق النتيجة بالرسم .

فئات العمال	15− 20	50 . 20	100 . 50	500 . 100	500 فأكثر
عدد المصانع	500	900	300	400	50

10. قامت مؤسسة لصناعة الملابس الجاهزة بجمع البيانات التالية عن مقاسات الملابس لمعرفة المقاس المناسب الذي يمكن ان تنتجه لزيادة مبيعاتها .
ما هو المتوسط المناسب الذي تنصح المؤسسة أخذه كأساس لتحقيق ها الهدف اذا أعطيت البيانات التالية :

المقاس	38	40	41	42	43	44	45	46
المبيعات	400	600	700	750	200	100	80	50

<u>ثانيا: تمارين علي مقاييس التشتت والتشتت النسبي</u>

1−أذكر الطرق التى تعرفها لقياس التتشتت والتشتت النسبي فى المجموعات الإحصائية ، مبينا باختصار أهم مميزات وعيوب كل منهما

2−أذكر بعض الخواص الهامة للمنحنى التكراري المتماثل ، وأذكر بعض الظواهر التى يمثلها منحنى تكرارى متماثل

3−مقاييس التشتت لا تصلح على علاقتها لمقارنة وأذكر بعض الظواهر التى يمثلها منحنى تكرارى متماثل . أشرح هذه العبارة بوضوح وبين كيف تقارن بين تشتت مجموعتين .

4− فيما يلي توزيع عدد من المصانع حسب ساعات العمل فيها إسبوعيا

ساعات العمل	30-	40-	60-	70 -	80-	90-	المجموع
عدد المصانع	30	330	315	90	70	35	790

المطلوب :

إيجاد الوسيط لساعات العمل الأسبوعية بالرسم من المنحنيين الصاعد والهابط . كذلك إيجاد كل من المتوسط الحسابى والانحراف المعيارى للتوزيع حسب التوائه.

5- ما يأتى يبين توزيع الأجور 150 عاملاً

الأجور	20 -25	25- 30	30 -35	35-40	40-45	45-50
التكرار	10	15	30	50	30	25

والمطلوب :

1- حساب الوسط الحسابي والوسيط والمنوال للأجور.
2- حساب الانحراف المعياري والانحراف الربيعي.
3- حساب معامل التشتت النسبي.
4- حساب مقياس الالتواء ودراسة تماثل التوزيع.

6- فيما يلى التوزيع التكراري لما تنفقه 350 أسرة بالجنيهات فى السنة على العلاج الطبى:-

الانفاق	50-60	60-70	70-80	80-90	90-100	100-110
الأسر	30	60	100	80	50	30

ارسم المنحنى التكراري لهذا التوزيع مهما بقدر الإمكان – وأوجد المنوال من الشكل ثم ارسم المنحنى منحنى التكرار المتجمع الصاعد وأوجد منه الوسيط والربيعين ومقياس تشتت ودرجه الالتواء

7- فيما يلى مبيعات عدد من متاجر التجزئة اليومى بالجنيهات ، والمطلوب حساب درجة تشتت هذا التوزيع :-

فئات المبيعات	15-20	20-25	25-30	30-35	35-40	40-45
التكرار	20	15	35	50	70	40

8- الجدول التالى يبين التوزيع التكرارى لعمال مصنعين حسب أجورهم اليومية بالجنيه .

والمطلوب:- مقارنه درجه التباين فى توزيع أجورهم فى المصنعين

فئات الأجر	4-2	6-4	8-6	10-8	12-10	14-12
عدد العمال بالمصنع الأول	20	40	160	60	30	10
عدد العمال بالمصنع الثانى	30	60	80	90	140	20

9- قارن بين درجه تشتت كل من التوزيعين التاليين باستخدام مقياس مناسب

التوزيع الاول

الفئـات	**أقل من 10**	**15-10**	**20-15**	**25-20**	**30-25**	**35-30**
التكرار	**10**	**15**	**30**	**20**	**10**	**5**

التوزيع الثانى

الفئـات	60-50	70-60	80-70	90-80	100-90
التكرار	15	60	90	70	30

10- اشـرح العلاقـة بـين الوسـط الحسـابي والوسـط الهندسـي والوسـط التـوافقي ، مستخدما القوانين ما امكن؟

الباب الثالث
أساسيات نظرية الاحتمالات
والتوزيعات الاحتمالية

تستخدم كلمة إحتمال كثيراً فى حياتنا اليومية فقد نقول أن هناك إحتمال أن تمطر اليوم أو نسأل عن إحتمال الفوز أو إحتمال الهزيمة واحتمال النجاح أو إحتمال الفشل واحتمال أن يعيش فرد أو إحتمال أن يموت أو إحتمال الإصابة بمرض معين. وفي جميع هذه الأمثلة نتكلم عادة عن حادثة فى المستقبل غير مؤكد أو مضمون حدوثها وفى بعض الأحيان نذكر عبارة تقريبية مثل " أن الفوز احتماله كبير" أو "أنه يوجد أمل قليل أن تنجح العملية الجراحية " وأحياناً نستخدم الأرقام فنقول" أن الأمل 100% أو أن فرصتنا فى الفوز هى نصف أو أن الرهان 1:2أو أن فريق معين سيفوز.

هذا وتعرف الاحتمالات أنها فرع من فروع الرياضة التطبيقية الذي يهتم بدراسة اثر عامل الصدفة.

ويشير الإحتمال عادة على فرصة حدوث حدث معين وفى هذا المجال يتعدد مفهوم الإحتمال فهناك الإحتمال التقليدي أو المسبق وهناك الإحتمال التجريبي ، ثم الإحتمال الشخصي أو الذاتي وعموماً يمكن تعريف الإحتمال تعريفاً مبسطاً بأنه التكرار النسبي لحدوث حدث معين وهو ما يعنيه الإحتمال التجريبي، حيث يستخدم التكرارات النسبية التي حدثت لظاهرة معينة فى الماضي لتعطي فكرة عن احتمالات ظهورها فى المستقبل. فى حين أن الإحتمال الكلاسيكي فهو يعنى تحديد فرصة حدوث حدث معين مسبقاً أو دون تجريب ، والمثال هو تحديد إحتمال ظهور رقم معين عند إلقاء زهرة طاولة أو إحتمال ظهور وجه معين عند إلقاء قطعة . وأن كان هذا ممكنا فى هذه الحالة فانه غالباً ما يصعب استخدامه للتحديد المسبق لفرض حدوث حدث معين فى المشاكل الواقعية والأعمال الحقيقية ولذلك يفضل عنه فى الغالب الإحتمال التجريبي. أما الإحتمال الشخصي أو الذاتي فيشير إلى درجة الاعتقاد لدى شخص ما بأنه حدث معينا سوف يحدث، تأسيساً على أدلة متوفرة لديه.

<u>1 – بعض المفاهيم الهامة فى الاحتمالات:</u>

عند عرض نظرية الاحتمالات يجب الإشارة بإيجاز إلى بعض المفاهيم والتعاريف الأساسية التي يلزم الإلمام بها.وفيما يلي عرض موجز لهذه المفاهيم:

1- **<u>الحادثة Event</u>:** تعني الحادثة بأنها ظهور حدث معين فعلى سبيل المثال عند قذف زهرة طاولة واحدة (محاولة) وظهور رقم معين وليكن 4 فمعنى ذلك يعبر عنه بحادثة ما وكذلك ظهور وجه تحمل الصورة عند إلقاء قطعة نقود يمكن أن نطلق عليه اصطلاح الحادثة.
2- **<u>المحاولة Trial</u>**: وهى عملية يرجي منها ملاحظة حدث معين، فعلى سبيل المثال إلقاء قطعة نقود أو قذف قطعة نقود أو زهرة طاولة أو سحب ورق من كتشينة تعتبر محاولة.

3- **التجربـــة Experiment:** وتعـرف التجربــة علــى أنهــا مجموعــة مــن المحاولات مثل قذف زهرة طاولة عدد من المرات أو إلقاء عدد من قطع النقود مرة واحدة.

4- **الحالات الممكنـة:** هى جميع الحادثات البسيطة الممكن الحصول عليها من إجراء المحاولة فى تجربـة معينـة، فعلى سبيل المثال عند رمي زهرة طاولة توجد ست حالات ممكنة هى أوجه الزهرة الست وكذلك عند سحب ورقة من كتشينة فإنه توجد 52 حالة ممكنة هى أوراق اللعب مجتمعة.

وتتعـدد الحادثـات فـى حياتنـا اليوميـة وبصـفة عامـة يمكـن تقسـيم هـذه الحادثات على النحو التالي:

أ- من حيث التماثل: الحادثات المتماثلة وغير المتماثلة:

ويقصـد بالحادثـات المتماثلـة بأنهـا تلـك الحادثـات التـي لهـا نفـس فـرص الحدث، فعلى سلبيل المثال عند قذف قطعة نقود فإن فرصة ظهور أى وجه من وجهيها تساوي 2 . أما الحادثات غير المتماثلة فتعني أن فرصة حدوث حادثة معينة تختلف عن فرصة حدوث الحادثة الثانية، فعلي سبيل المثال إذ وضعت خمس كرات حمراء اللون مع ثلاث كرات بيضاء5 اللون فى صندوق وسحبت كرة واحدة فإن ف3رصة ظهور كرة حمراء تساوي 8 فى حين فرصة ظهور كرة بيضاء تساوي 8 .

ب – من حيث التعارض أو التنافي: الحادثات المتنافية وغير المتنافية:

يقصـد بالحادثـات المتنافيـة أنهـا الحادثـات التـي لا يمكـن أن تحـدث مـع بعضـها البعض، بمعنى أن ظهور حادثة فيها ينفي ظهور الحادثات الأخرى، فعلى سبيل المثال عند إلقاء قطعة نقود فإن ظهور الوجه الذي يحمل الشعار تنفى بالقطع فرصة ظهور الوجه الأخر وذلك فى المحاولة الواحدة.

جـ – من حيث الارتباط أو الاستقلال الإحصائي: الحادثات المستقلة وغير المستقلة :

يطلق على الحادثة أنها مستقلة إذا كانت فرصة ظهورها لا تؤثر على فرصة ظهور الحادثات الأخرى، فالحصول مثلا على وجه زهرة الطاولة الذي يحمل رقم (2) فى محاولتين متكررتين بغير حادثتين مستقلين ذلك لأن إحتمال ظهـور هـذا الوجـه فـى المحاولـة الأولـى والـذي يعـادل 6 لا يـؤثر ولا يتـأثر باحتمال ظهور نفس الوجه فى المحاولة الثانية. ومثال آخر فرض أن هناك خمس كرات حمراء وثلاثـة بيضـاء وضـعت فـى صـندوق وسـحبت كرتـان مـن الصندوق فيمكننا ملاحظة التالي:

إذا كـان السحب بـالإحلال : بمعنـى أنـه سحبت الكرة الأولـى ثم أعيدت إلـى الصندوق مرة أخرى وسحبت الكرة الثانية، ففي هذه الحالـة يكون إحتمال الحادثة الأولى والتي تعبر عن نتيجة سحب الكرة الأولى) وليكن هو ظهور كرة حمراء تعادل $^{5}/_{8}$ ، واحتمال الحادثة الثانية ولتكن أيضاً كرة حمراء تعادل أيضاً $^{5}/_{8}$ ، وفى هذه الحالة فإن الحادثتين مستقلين عن بعضهما إحصائياً.

إذا كان السحب بدون إحلال: بمعنى عدم إعادة الكرة الأولى إلى الصندوق قبل سحب الكرة الثانية وذلك فى المثال السابق، فإن إحتمال ظهور الكرة الأولى حمراء يعادل $^5/_8$ فى حين أن إحتمال ظهور الكرة الثانية حمراء أيضاً أما أن يكون $^4/_7$ (فى حالة ظهور الكرة الأولى حمراء بالفعل) أو يكون $^5/_7$ (فى حالة ظهور الكرة الأولى بيضاء بالفعل) وفى مثل هذه الحالة يطلق على هذه الحادثات بأنها حادثات غير مستقلة إحصائية، الأمر الذي يعني أن ظهور حادثة معينة يؤثر بالقطع على إحتمال الحادثات التالية.

2- <u>خصائص الإحتمال:</u>

سبق الإشارة على أن التعريف البسيط لإحتمال حدوث حادثة ما يعني أنه التكرار النسبي لحدوث هذه الحادثة.

ولتوضيح ذلك يفترض قذف زهرة طاولة عشر مرات متتالية ودونت النتائج المتحصل عليها فى جدول تكراري بسيط على النحو التالي:

القيم	1	2	3	4	5	6
عدد مرات الظهور (التكرار)	2	2	1	3	2	0
التكرار النسبي للاحتمال	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{0}{0}$

ومن هذا المثال المبسط يمكننا استنتاج بعض خصائص الاحتمال على النحو التالي:

1- إن قيمة الإحتمال تتراوح بين حد أدنى قدره صفر وحد أعلى قدره الواحد الصحيح:

فبفرض أن الحدث يمكن أن يحدث بطرق عددها h ومن بين نواتج متساوية الفرصة فى الوقوع عددها N فإن إحتمال الحدث E يعرف كما يلي:

$$P(E) = \frac{h}{N}$$

ومن المثال السابق يتبين أن :

$$P \geq P(E) \geq \text{صفر}$$

2- إن مجموع الاحتمالات الممكنة لظهره ما لابد وأن تساوى الواحد الصحيح:

فكما هو موضح بالمثال السابق فإن مجموع الاحتمالات أو التكرارات النسبية يساوي $\frac{10}{10}$=1 وهذا ما يعبر عنه كالآتي:

$$\sum P(E) = PE_1 + PE_2 + PE_3 + \ldots\ldots. + PE_n = 1$$

3- <u>جمع الاحتمالات</u> : <u>Additive Law of Probability</u>

يعنى جمع الاحتمالات تحقق حادثتين أو أكثر فى نفس الوقت، ويعبر عنه بحدوث الحادثة الأولى أو الحادثة الثانية وتعنى كلمة أو هنا تحقق أى من الحادثتين أو كليهما معاً.

ولتوضيح قاعدة جمع الاحتمالات يلزم التعرف على طبيعة الحادثات محل الجمع والتمييز فيما إذا كانت هذه الحادثات متنافية أم غير متنافية.

<u>3-1 جمع الاحتمالات فى حالة الحادثات المتنافية:</u>

إذا فرض أن A، B عبارة عن حادثتين متنافيتين فإن :

$$P(A+B+C+D)=F(A)+(B)+(C)+(D)$$

<u>مثال رقم (1):</u>

ما هو إحتمال الحصول على أقل من 3 فى رمية واحدة لزهرة طاولة.

<u>الحل</u>

أقل من 3 يعنى ظهور الرقم 1 أو 2 ولما كان هذين الحدثين متنافيين أى يمنع حدوث أحدهما حدوث الأخر فنطبق عليهما قاعدة الجمع للحادثات المتنافية أى أن :

$$P(A\,or\,B)=P(A+B)=P(A)+P(B)$$

$$P(1or2)=p(1)+p(2)=\frac{1}{6}+\frac{1}{6}=\frac{2}{6}=\frac{1}{3}$$

<u>مثال رقم (2):</u>

ما هو إحتمال ظهور كرة حمراء أو زرقاء عند سحب كرة من وعاء يحتوى على خمس كرات حمراء ، وثلاثة زرقاء واثنان خضراء.

<u>الحــل:</u>

بفرض أن إحتمال ظهور كرة حمراء =p(A)

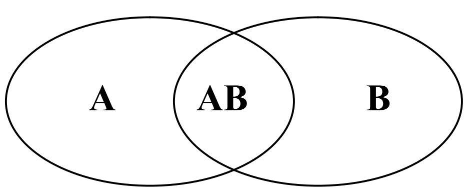
واحتمال ظهور كرة زرقاء = P(B)

فإن $P(A\,or\,b)=p(A+B)=P(A)+P(B)$

$$=\frac{5}{10}+\frac{3}{10}=\frac{8}{10}$$

<u>3-2 : جمع الاحتمالات فى حالة الحادثات غير المتنافية:</u>

بفرض أن A، B حادثتين غير متنافيتين فإن هذا يعني ببساطة إمكانية حدوث الحادثة A وحدها أو الحادثة B وحدها أو حدوثها معاً أى AB وبتوقيع هذه الحالات الممكنة نحصل على ما يطلق عليه فراغ العينة وهو الموضح بالشكل التالي المسمى بشكل فني.



ومن الشكل يمكن استخلاص الأتي:

$$P(A)=\frac{A+AB}{N} \quad , \qquad P(B)\frac{AB+B}{N}$$

$$P(AB)=\frac{ANB}{N}$$

أى أن إحتمال حدوث A يعني إحتمال ظهور A وحدها أو معها B، وكذلك إحتمال حدوث يعني إحتمال ظهور B وحدها أو معها A. أما إحتمال AB فيعني بالضرورة إحتمال ظهور A ومعها ظهور B.

<u>مثال رقم (3):</u>

بفرض أن عشرة من الدارسين دخلوا امتحان مادتين الأولى A والثانية B وقد نجح فى امتحان المادة الأولى فقط ثلاثة من الدارسين ، وفى الثانية أربعة من الدارسين فى حين أن الباقي قد نجح فى المادتين معا، وهذا يمكن توضيحه بفراغ العينة التالي:

حيث إحتمال النجاح فى المادة A سواء وحدها = P(A) أو مع غيرها

$$P\ (A)=\frac{A+AB}{N}=\frac{3+3}{10}=\frac{6}{10}$$

$$P\ (B)=\frac{B+AB}{N}=\frac{4+3}{10}=\frac{7}{10}$$

ولإثبات ذلك تستخدم ما سبق إيضاحه كالتالي:

يعني إحتمال النجاح فى A أو B أو كلاهما = P(A+B)

=1

P (A) +P (B)-P (AB) =

$$\frac{A+AB}{N}+\frac{B+AB}{N}-\frac{AB}{N}=\frac{A+B+AB}{N}=\frac{N}{N}=1$$

الطرفان متساويان وبذلك تكون فى عدة جمع الاحتمالات للحادثات غير المتنافية كما سبق إيضاحه وهى كالتالي:

P(A+B) =P (A) +P (B)-P (AB) =

ويمكن بنفس الأسلوب إثبات الأتي:

$$P\ (A+B+C+D)=P\ (A)+(B)+(C)-P(AB)-P(AC)$$
$$-P(BC)+P(ABC)$$

مثال رقم (4):

ما هو إحتمال الحصول على أس أو سبات عند سحب ورقة واحدة من مجموعة من أوراق اللعب.

الحـل:

باعتبار أن أس أو سباتي حدثين غير متنافيين حيث انه يمكننا أو نحصل على ورقة أس وفى نفس الوقت فهي سباتي وبذلك تنطبق قاعدة الجمع فى حالة الحادثات غير المتنافية.

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

باعتبار أن:

احتمال الحصول على ورقة أس = $P(A)$

احتمال الحصول على ورقة سباتي = $P(B)$

احتمال الحصول على ورقة أس أو سباتي فى نفس الوقت = $P(AB)$

وبذلك يكون :

$$P(A+B) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} - \frac{4}{31}$$

مثال رقم (5):

الحوادث هنا متنافية أى أن هناك إحتمال مشترك وبفرض أن الحدث الأول E_1 والثاني E_2 فإن:

الإحتمال المطلوب هو $P(E_1+E_2)=P(E_1)+P(E_2)+P(E_2)-P(E_1E_2)$

=P(رقم زوجي)+P(3 رقم قابل لقسمة على) – P(6)

$$= \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

مثال رقم (6):

مجموعة تتكون من 120 طالبا يدرس 60 منهم اللغة الفرنسية، ويدرس 50 منهم اللغة العربية ، بينما يدرس 20 فقط اللغتين الفرنسية والعربية معاً. فإذا اختير طالبا عشوائياً ما هو إحتمال أن تكون:

أ- دارساً للفرنسية أو العربية.

ب- لا يدرس الفرنسية أو العربية.

الحـل:

الحوادث هنا واضح أنها غير متنافية

أ- بفرض أن دارس الفرنسية يعبر عنه بالرمز E_1 ، ودارس العربية يعبر عنه بالرمز E_2 .

إحتمال أن يكون الطالب دارساً للفرنسية أو العربية هو:

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

$$= \frac{60}{120}+\frac{50}{120}-\frac{20}{120}=\frac{90}{120}$$

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

ب – احتمال أن يكون الطالب لا يدرس أى من اللغتين الفرنسية أو العربية هو:

P(لا يدرس أى من اللغتين)= 1− (احتمال دراسة الفرنسية أو العربية)

$$= 1-P(E_1+E_2)$$

$$= 1-\frac{3}{4}=\frac{1}{4}$$

القواعد العامة لضرب الاحتمالات:

Multiplicative law of Probability:

قبل تطبيق أو استخراج قاعدة ضرب الاحتمالات يلزم التمييز مرة أخرى بين الحادثات فيما إذا كانت هذه الحادثات مستقلة إحصائياً أم غير مستقلة إحصائياً.

أولاً: ضرب الاحتمالات فى حالة الحادثات المستقلة إحصائياً:

تسمى الأحداث مستقلة إذا كان حدوث أى منها ليس له أى تأثير على احتمالات حدوث أى من الأحداث الأخرى.

وكمثال للأحداث المستقلة افترض أننا قذفنا زهرة طاولة فى نفس الوقت الذي رمينا فيه قطعة نقود . فإن ما يحدث لقطعة النقود ليس له صلة بما يحدث لزهرة الطاولة ومن ثم فإن الحدثين "صورة فى قطعة النقود" ، و"5 فى زهرة الطاولة" سوف يكونان حدثين مستقلين. وإذا كانت الأحداث مستقلة فإن إحتمال حدوثها جميعا يكون حاصل ضرب احتمالاتها الفردية.

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

وفى المثال السابق فإنه بإستخدام هذه القاعدة نحصل على الإحتمال التالي:

P(5 فى زهرة الطاولة وصورة فى قطعة النقود)

= P (صورة فى قطعة النقود) × p(5 فى زهرة الطاولة)

$$= \frac{1}{2}\times\frac{1}{6}=\frac{1}{12}$$

مثال رقم (7):

ما هو إحتمال الحصول على ما مجموعة يساوي عشرة أو أكثر عند رمي زهرتي طاولة متوازنتين معا مرة واحدة.

الحل:

إحتمال الحصول على ما مجموعة يساوي (10) أو أكثر

عشرة أو أكثر تعني 10i و11و12

أولاً: إحتمال الحصول على 10 هو مجموع الاحتمالات التالية:

(6 و4) (4 و6)i، (5 و 5)

$$=\left(\frac{1}{6}\times\frac{1}{6}\right)+\left(\frac{1}{6}\times\frac{1}{6}\right)+\left(\frac{1}{6}\times\frac{1}{6}\right)=\frac{3}{36}$$

ثانياً: إحتمال الحصول على 11 هو مجموع الاحتمالات التالية:

(6 و5)i، (5 و6)

$$=\left(\frac{1}{6}\times\frac{1}{6}\right)+\left(\frac{1}{6}\times\frac{1}{6}\right)=\frac{2}{36}$$

ثالثاً: إحتمال الحصول على 12 هو :

(6 و6) فقط

$$=\left(\frac{1}{6}\times\frac{1}{6}\right)=\frac{1}{36}$$

الإحتمال المطلوب= $\frac{3}{36}+\frac{2}{36}+\frac{1}{36}=\frac{6}{36}=\frac{1}{6}$

ثانياً : ضرب الاحتمالات فى حالة الحادثات غير المستقلة إحصائياً:

قبل الوصول إلى قاعدة ضرب الاحتمالات فى حالة الحادثات غير المستقلة يلزم التعرف على مفهوم الإحتمال الشرطي Conditional Probability والذي يعني تحقق حادثة ما بشرط تحقق حادثة أخرى . فإذا فرض أن:

إحتمال حدوث الحادثة A = P(A)

إحتمال حدوث الحادثة B = P(B)

إحتمال حدوث الحادثة A،B معاً = P(AB)

فإن :

إحتمال حدوث A حدوث B= $P\left(\frac{A}{B}\right)$

وفى هذه الحالة يكون

$$P\left(\frac{A}{B}\right)=\frac{P(AB)}{P(B)}$$

وكذلك

$$P\left(\frac{B}{A}\right)=\frac{P(AB)}{P(A)}$$

ويمكن إثبات هذه القاعدة كالتالي:

$$P\left(\frac{A}{B}\right)=\frac{AB}{B+AB}$$

$$P(B)=\frac{B+AB}{N}$$

$$P(AB)=\frac{AB}{N}$$

$$\frac{P(AB)}{P(B)}=\frac{AB}{N}\bullet\frac{N}{B+AN}=\frac{AB}{B+AB}$$

ومن القاعدة السابقة يمكن استنتاج قاعدة ضرب الاحتمالات فى حالة الحادثات غير المستقلة كالآتي:

$$P\left(\frac{A}{B}\right)=\frac{P(AB)}{P(B)}$$

$$P(AB)=P(B)P\left(\frac{A}{B}\right)$$

$$=P(A)P\left(\frac{B}{A}\right)$$

مثال رقم (8):

إذا فرض أنه وضعت فى سلة عشر علب فاصوليا منها أربعة تالفة فإذا سحبت علبة وتبين أنها تالفة ، فما هو إحتمال أن تكون الثانية تالفة أيضاً.

الحل:

إحتمال أن تكون الأولى تالفة $P(A)=\frac{4}{10}$

إحتمال أن تكون الثانية تالفة بشرط أن تكون الأولى تالفة $P(AB)=\frac{3}{9}$

إحتمال أن تكون الأولى والثانية تالفة

$$P(AB)=P(A)\cdot P\left(\frac{B}{A}\right)=\frac{4}{10}\cdot\frac{3}{9}=\frac{12}{90}$$

التباديل والتوافيق: Permutations and Combinations

يعنـي تبـديل مجموعـة مـن الحادثـات علـى أنـه ترتيـب معـين لحـدوث هـذه الحادثات، وعندما يتغير هذا الترتيب يحصل على تبديل آخر. وعلى سبيل المثال إذا كان لدينا ثلاث حروف هى A,B,C فإنه يمكن ترتيبها فى صورة ست تباديل كالأتي:

ABC	BAC	CAB
ACB	BCA	CBA

وبطبيعة الحال يزداد عدد التباديل زيادة كبيرة كما زاد عدد الحادثات ، ولذلك تستخدم القاعدة التالية لمعرفة عدد التباديل

$$R_x^n = \frac{ni}{(n-x)i}$$

حيث عدد التباديل = R_x^n

العدد الكلي للمفردات أو الحادثات = n

عدد المفردات أو الحادثات المكون للتبديلة = x

وبتطبيق هذه القاعدة على المثال السابق يتبين التالي:

n=3 x=3

$$R_x^n = \frac{3i}{(3-3)i} = 3i = 3 \times 2 \times 1 = 6$$

ويتبين من المثال السابق

أن التباديل هى عبارة عن ترتيب للحادثات ، فإذا اختلف هذا الترتيب أدى ذلك إلى ظهور تباديل جديدة ، على الرغم من أن هذه التباديل قد تحتوى على نفس المكونات ولكن بترتيب مختلف، أما التوافيق فإنها تعنى للحادثات بغض النظر عن ترتيبها. ففي المثال السابق لدينا ستة تباديل كل منها مكون من الثلاثة حروف ، وعلى ذلك يكون لدينا توفيقه واحدة.

ولتوضيح ذلك نفترض أن لدينا حروف A,B,C والمطلوب حساب عدد التباديل، والتوافيق المكونة من حرفين من هذه الحروف

عدد التباديل هو

$$R_x^n = \frac{ni}{(n-x)i} = \frac{3i}{(3-2)i} = \frac{3i}{1} = 6$$

عدد التوافيق هو ثلاثة كالتالي:

AB	AC	BC

ويمكـــــــن الوصـول إلــى هـذا العـــدد بإســـتخدام القاعدة التالية:

$$R_x^n = \frac{ni}{xi(n-x)i} = \frac{3i}{2i(3-2)i} = \frac{3 \times 2i}{2i} = 3$$

حيث أن:

عدد التوافيق لعدد n كل منها مكون من x = C_x^n
العدد الكلى للحادثات = n
عدد المفردات أو الحادثات فى التوفيقة = x

القيم المتوقعة: Expected Values

يصف التوزيع الاحتمالي لمتغير عشوائي ما السلوك الاحتمالي لذلك المتغير ، حيث يمكن تبديل مجموعة من المعالم Parameters للتوزيع الاحتمالي تساعد توصيفه وفى تحديد القيم الاحتمالية لقياسات معينة تنطوي ضمن هذا التوزيع الاحتمالي، ويطلق على هذه المعالم تعبير القيم المتوقعةExpected Values أو العزوم Moments.

وتنقسم القيم المتوقعة أو العزوم إلى قسمين رئيسين أولهما المركزية Central Moments وثانيهما العزوم التشتتية Dispersion Moments.

القيم المتوقعة أو العزوم المركزية: Central Moments

أولاً: العزوم المركزية للتوزيعات الاحتمالية المنفصلة:

تتعدد العزوم المركزية للتوزيعات الاحتمالية المنفصلة فهناك العزم المركزي الأول (التوقع الأول)، وهناك العزم المركزي الثاني، والثالث وهكذا، وفيما يلي عرضا لهذه العزوم . فإذا فرض أنX متغير عشوائي موزع توزيعاً احتماليا منفصلاً فإن:

العزم الأول = $E(x) = \sum x_i P(x)$

وهذه القيمة المتوقعة تعادل المتوسط الحسابي $\overline{X}$حيث:

$$E(x) = \sum \frac{x \quad f_i}{\sum f_i} = \frac{\sum fx}{\sum f_i} = \overline{X}$$

حيث تكرار حدوث القيمة x = fX

مجموع التكرارات = $\sum f$

كما تبين المعادلات التالية العزم المركزي الثاني وهكذا

العزم المركزي الثاني = $E_2(X) = \sum X^2 P(X)$

العزم المركزي الثالث = $E_3(X) = \sum X^3 P(X)$

العزم رقم n = $E_n(X) = \sum X^n P(X)$

ثانياً: العزوم المركزية للتوزيعات الاحتمالية المستمرة:

تتنـاول القواعـد الإحصـائية الخاصـة بـالعزوم المركزيـة للتوزيعـات الاحتماليـة المستمرة مـع نظيرتها الخاصـة بـالعزوم المنفصـلة، ماعدا بعض التغيرات المرتبطـة بتطبيق التوزيعات المستمرة ، وذلك على النحو التالي:

$$E(X) = \int^{+\infty} XP(X)D(X) = \overline{X}$$

وكمـا يلاحـظ هنـا فقـد اسـتبدلنا إشـارة الجمـع $\sum$ بإشـارة التكامـل والتـي تتناسب $\int D(X)$ مع طبيعة التوزيعات المستمرة.

وبنفس الطريقة يمكن استخراج العزوم المركزية الأخرى كالتالي:

العزم المركزي الثاني =

$$E_2(X) = \int_{-\infty}^{+\infty} X^2 P(X) D(X)$$

العزم رقم n =

$$E_n(X) = \int_{-\infty}^{+\infty} X^n P(X) D(X)$$

القيم المتوقعة أو العزوم التشتتية: Dispersion Moments

وتبين هذه العزوم تغير أو تشتت التوزيـع الاحتمالي حول متوسطه الحسابي وفيما يلي عرض لهذه العزوم فى كل من التوزيعات المنفصلة والمستمرة.

أولاً: العزوم التشتتية للتوزيعات الاحتمالية المستمرة:

العزم التشتتي الأول = $E_1'(X) = \sum [X - E_1(X)] P(X)$

وهذا العزم يعني ببساطة مجموع انحرافات قيم التوزيع الاحتمالي عن وسطها الحسابي (العزم المركزي الأول) وهو بذلك لابد وأن يساوي الصفر ويمكن إثبات ذلك كالتالي:

$$\begin{aligned} E_1'(x) &= \sum [X - E_1(X)] P(X) \\ &= \sum XP(X) - \sum EXP(X) = \sum XP(X) - E(X) \sum P(X) \\ &= EP(X) - E_1(X) = 0 \end{aligned}$$

حيث أن $\sum P(X) = 1$

كما يعتبر العزم التشتتي الثاني من أهم العزوم حيث يطلق عليه التباين Variance ويرمز له بالرمز σ^2 وهو كالتالي:

$$E_2'(X) = \sigma^2(X) = \sum [X - E(X)]^2 P(X)$$

ويمكن استخراج قيمة هذا العزم التشتتي الثاني σ^2 بمعرفة العزمين المركزين الأول والثاني وذلك من خلال العلاقة التالية:

$E_2'(X) = \sigma^2(X) = E_2(X) - [E_1(X)]^2$

ويمكن إثبات هذه القاعدة كالتالي:

$$E_2^{/}(X) = \sigma^2(X) = \sum[X_2 - E(X)]^2 P(X)$$

$$= \sum\left[X^2 + E(X)^2 - 2XE(X)\right] P(X)$$

$$= \sum X^2 P(X) + \sum(E(X))^2 P(X) - \sum 2XE(X)P(X)$$

$$= E_2(X) + (E(X))^2 \sum P(X) - 2E(X)\sum XP(X)$$

$$= E_2(X) + (E(X))^2 - 2E(X)E(X)$$

$$= E_2(X) + (E(X))^2 - 2(E(X))^2$$

$$= E_2(X) - (E_1(X))^2$$

<u>ثانياً</u>: <u>العزوم التشتتية للتوزيعات الاحتمالية المنفصلة</u>:

بنفس الطريقة يمكن إستباط قواعد العزوم التشتتية المستمرة وذلك على النحو التالي:

$$E_1'(X) = \int_{-\infty}^{+\infty} X - E_1(X)P(X)D(X) = 0$$

$$E_2(X) = \int_{-\infty}^{+\infty} X - E_1(X)^2 P(X)D(X) = \sigma^2(X)$$

$$= E_2(X) - (E_1(X))^2$$

<u>مثال رقم (9)</u>:

بفرض أن المتغيرين X، Y متغيرين منفصلين فى صورة دالة خطية:

Y=A+BX

بإعتبار أن A،B ما هى إلا ثوابت الدالة الخطية فإثبت أن:

$E_1(y) = A + BE_1(X)$

$E_2(Y) = \sigma^2\, Y = B^2E^2(X)$

<u>الحـل</u>:

Y=A+BX

$$E(Y) = E_1\ A+BX = \sum(A+BX)P(X)$$

$$= \sum AP(X) + \sum BXP(X)$$

$$= A\sum P(X) + B\sum XP(X)$$

$$= A + BE_1(X)$$

حيث :

$$E_2(Y) = \sum(Y-E)_1(Y)^2 P(Y)$$

$$= \sum\left[(A+BX)-(A+BE_1(X))P(X)\right]^2 P(X)$$

$E_1(Y)$ = حيث أنه قد ثبت أن

$A+BE_1(X)$

$$= \sum\left[(A+BX-A+BE_1(X))\right]^2 P(X)$$

$$= \sum\left[(BX-A+BE_1(X))\right]^2 P(X)$$

$$= B^2\sum\left[(X-A+B_1(X))\right]^2 P(X)$$

$$= B^2 E_2(X)$$

$E_2(X)$ = حيث أن

$$\sum\left[(X-E_1(X))\right]^2 P(X)$$

التوزيعات الإحتمالية

أولاً : التوزيعات الإحتمالية المنفصلة

(أ) : التوزيع البرنيولي: Bernoulli Distribution

بفرض أن المتغير العشوائي x فى تجربة مكونة من محاولة واحدة فقط أما أن يأخذ قيمة واحد أو قيمة الصفر، فإذا نجحت المحاولة أخذ قيمة واحد وإذا فشلت أخذ قيمة صفر ،ويمكن التعبير عن ذلك كالتالي:

قيمة المتغير x	الإحتمال p(x)
1	p
0	q

ولما كانت التجربة مكونة من محاولة واحدة فقط وأنه ليس هناك أية نتائج أخرى غير النجاح أو الفشل فبذلك يكون الأتي:

$$\sum P(X) = P + q = 1$$

$$P = 1 - q$$
$$P = 1 - p$$

وتتبع هذه الظاهرة بمواصفاتها السابقة أبسط التوزيعات المنفصلة والمسمى بالتوزيع البرنو لي ومن أمثلة ذلك قذف زهرة نرد فإذا ظهر الوجه الذي يحمل 3 مثلاً اعتبرت المحاولة ناجحة وإذا لم تظهر هذا الوجه اعتبرت المحاولة فاشلة.وبذلك يمكننا أن نقول أن المتغير x يأخذ القيمة 1 إذا نجحت المحاولة ويكون احتماله هو p. وأن المتغير x يأخذ القيمة صفر إذ فشلت المحاولة ويكون احتماله هو q.

وتجدر الإشارة إلى أن قواعد العزوم السابق عرضها سواء كانت المركزية أو التشتتية يمكن تطبيقها على هذا التوزيع وغيره من التوزيعات وذلك بهدف استخراج قواعد خاصة لزوم كل توزيع تتناسب طبيعته.

وبتطبيق قواعد العزوم يمكننا استخراج عزوم التوزيع البرنو لي كالتالي:

$$E_1(X)=\sum X\,P(X)$$
$$=1\cdot P+0\cdot q$$

أى أن العزم المركزي الأول للتوزيع = x= إحتمال النجاح(p) وكذلك يمكننا استخراج العزم المركزي الثاني كالتالي:

$$E_2(x)=\sum X^2 P(X)$$

إحتمال النجاح أيضاً = $=1^2 . P+ \ 0^2 . q= P$

وبمعرفة العزمين المركزين الأول والثاني يمكننا استخراج العزم التشتتي الثاني σ^2 (X) كالتالي:

$$E_2^{/}(X) = \sigma^2 X = E_2(X) - E_1(X)^2$$
$$=P-P^2$$
$$=P(1-P)=Pq$$

أى حاصل ضرب إحتمال النجاح وإحتمال الفشل

<u>(ب) : توزيع ذو الحدين</u> : Binominal Distribution

يعتبر التوزيع ذو الحدين أحد التوزيعات الاحتمالية المنفصلة وهو باختصار تكرار للتوزيع البرنولى، أى أنه يتكون من مجموعة من المحاولات كل منهما إما أن تنجح المحاولة باحتمال p ويأخذ المتغير x القيمة واحد أو تفشل باحتمال q ويأخذ المتغير القيمة صفر وبصفة أكثر تحديداً يكون التوزيع الاحتمالي لظاهرة ما تابعا للتوزيع ذو الحدين إذ توفرت الشروط التالية:

1- أن تتكون التجربة من عدد من المحالات المستقلة عن بعضها البعض.
2- أن يكون لكل محاولة ناتجان ممكنا ومتنافيات لكل محاولة .
3- إن إحتمال وقوع الحادث المطلوب (النجاح ولا يتغير من محاولة لأخرى).
4- ولاستخراج الدالة الاحتمالية لتوزيع ذو الحدين نفترض تجربة مكونة من خمس محاولات تقذف زهرة نرد، وبفرض أن ظهور الوجه الذى يحمل الرقم هو نجاح المحاولة وعدم ظهور هذا الوجه يمثل فشل المحاولة وبفرض أن المحاولتين الأولى والثانية نجحها والمحاولات الثلاث التالية فشلت فما هو إحتمال حدوث ذلك.
5- وبتطبيق قواعد ضرب الإحتمالات ، حيث أننا نريد إحتمال أن تكون المحاولة الأولى والثانية ناجحة والثالث والرابعة والخامسة فاشلة.

ويمكننا صياغة ذلك كالتالي:

إحتمال فشل المحاولة = $q=\dfrac{5}{6}$

إحتمال نجاح المحاولة = $P=\dfrac{1}{6}$

الإحتمال المطلوب = $\frac{1}{6}.\frac{1}{6}.\frac{5}{6}.\frac{5}{6}\frac{5}{6}=P^2q^3$

وبفرض أن العدد الكلي للمحاولات هو n وأن عدد مرات النجاح x فيمكننا حساب عدد الإحتمالات بصورة أخرى كالتالي:

$$P^x q^{n-x} = p^x(1-p)^{n-x}$$

يلاحظ من السابق أننا افترضنا تدريباً معيناً لحدوث مرات النجاح أو الفشل، غلا أنه بتطبيق الحال يمكن الحصول على عدة طرق لتحقيق الهدف ومن ثم يمكن صياغة الدالة الاحتمالية للمتغير x للحصول على عدد n من مرات النجاح كالتالي:

$$P(X)=C_x^n p(1-P)^{n-x}$$

$$P(X)=\frac{ni}{xi(n-x)i}P^X(1-P)^{n-X}$$

وبتطبيق هذه القاعدة على المثال السابق يمكننا حساب احتمال الحصول على مرتين نجاح فى خمس محاولات كالتالي:

$$P(X)=C_5^2\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$$

$$=\frac{5i}{2i.3i}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2-\frac{5\times 4\times 3i}{-2i\times 3i}\left(1/6\right)^2\left(5/6\right)$$

$$=\frac{5\times 4}{2\times 1}\left(1/6\right)^2\left(5/6^2\right)=10\left(1/5\right)^2\left(5/6\right)^2$$

ويمكن استخراج القواعد الخاصة بمتوسط وتباين توزيع ذو الحدين كالتالي:

$$u_1=E_1(X)=\overline{X}=\sum_{x=0}^{n}XP(X)$$

$$=\sum_{x=0}^{n}XC_x^n P^X(1-P)^{n-x}$$

$$=0.P^n+1C_1^n Pq^{n-1}+2c_2^n P^2q^{n-2}+........+nP$$

$$=nP(q^{n-1}+c_2^{n-1}P^2q^{n-2}+Pc_2^{n-1}q^{n-3}P+........+PP+P^{n-1}$$

$$=np(q+P)^{n-1}=np$$

$$u2=E2(X)=\sum_{x=0}^{n}X^2P(X)=\sum_{x=0}^{n}X^2C_x^n P^X(q)^{n-x}$$

$$=\sum_{x=0}^{n}[x+x(x-1)]c_2^n P^x q^{n-x}$$

$$= \sum_{x=0}^{n} X^2 C_x^n P^{X} (q)^{n-x} + \sum_{x=0}^{n} x(x-1) c_x^n p^x q^{n-x}$$

$$= nP + n(N-1) \sum_{x=2}^{n} c_{x-2}^{n-2} Pxq^{n-x}$$

$$= nP + n(5n-1)P^2 (q+p)^{n-2}$$

$$= nP[1 + (n-1)p] = np(q + np)$$

$$= nPq + n^2 p^2$$

$$E_2(X) = \sigma^2 X = E_2(X) - E_1(X)^2$$

$$= nPq + n^2 p^2 - (np)^2$$

$$= nPq$$

<u>مثال رقم (10):</u>

إذا فرض أن إحتمال حصول أبوين على طفل أشقر الشعر هو فإذا كان فى الأسرة 6 أطفال.

أ- فما هو إحتمال أن نصفهم ذو شعر أشقر.

ب-ما هو إحتمال أن يكون اثنان منهم على الأكثر ذو شعر أشقر .

جـ- قدر متوسط وتباين التوزيع.

<u>الحـل:</u>

أ- احتمال أن نصف الأطفال أى ثلاثة منهم من ذوى الشعر الأشقر

$p = \frac{3}{4}$ n=6 x=3

$$P(X) = \frac{ni}{xi(n-x)i} P^X q^{n-X}$$

$$P(X-3) = \frac{6i}{3i(6-3)i} \left(\tfrac{1}{4}\right)^3 \left(\tfrac{3}{4}\right)^{6-3}$$

$$= \frac{5.6.4.3.2.1}{3.2.1.3.2.1} . \tfrac{1}{6} . 4 . \tfrac{27}{64}$$

ب- إحتمال أن يكون إثنان منهم على الأكثر ذو شعر أشقر

X=2 or x=1 or =0

$$P(X=0) = \frac{6i}{6i} \left(\tfrac{1}{4}\right)^0 \left(\tfrac{3}{4}\right)^6 = \left(\frac{3}{4}\right)^6 = 0.17798$$

$$P(X=1) = \frac{6i}{5i} \left(\tfrac{1}{4}\right)^0 \left(\tfrac{3}{4}\right)^5 = 6 \times 0.25 \times 0.2373 = 0.35595$$

$$P(X=2)=\frac{6i}{2i.6i}\left(\frac{1}{4}\right)2\left(\frac{3}{4}\right)4$$
$$=\frac{6\times5\times4i}{2\times1\times4i}=\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^4$$
$$=15\times0.0625\times0.31641=0.29663$$
$$p(x\le2)=p(x=0)+p(x=1)+p(x=2)$$
$$=0.17798+0.35595+0.29663$$
$$=0.83056$$

جـ- متوسط وتباين التوزيع.

$$E_1(X)=np=6\times0.25=1.5$$
$$E_2(x)=(x)^2=npq=6\times0.25\times0.75=1.125$$

مثال رقم (11):

إذا علمت أن إحتمال حصول أسرة ما على ولد يساوى إحتمال حصولها على بنت أوجد الإحتمالات التالية:

(أ) إحتمال الحصول على ولد على الأقل (ب) على ولد وبنت على الأقل لعائلة لديها أربعة أطفال.

الحل الأول:

إحتمال الحصول على بنت يساوى احتمال الحصول على ولد إذن غحتمال الحصول على اى منهم $=\frac{1}{2}$

(أ) p(ولد على الأقل) $=1-$ p(عدم وجود ولد)

p(ولد على الأقل) $= 1-\left(\frac{1}{2}\right)^4=\frac{15}{16}$

(ب) p(ولد وبنت على الأقل) $=1-$ p(عدم وجود ولد) $-$ p(عدم وجود بنت).

$$=1-\frac{1}{16}-\frac{1}{16}=\frac{14}{16}=\frac{7}{8}$$

حل آخر باستخدام توزيع ذو الحدين : (p) ولد على الأقل

(1) p (ولد) $=c_1^4\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3$

(2) p(ولدين) $=c_2^4\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2$

(3) p(ثلاث أولاد) $=c_3^4\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)1$

(4) p (أربع أولاد) $=c_4^4\left(\frac{1}{2}\right)4\left(\frac{1}{2}\right)^0$

$$p(\text{ولد على الأقل}) = (1)+(2)+(3)+\left(4 = \frac{15}{16}\right)$$

(ب) ولد وبنت على الأقل

$$p(\text{عدم وجود ولد}) = c_0^4\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

$$p(\text{عدم وجود بنت}) = c_0^4\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

$$p(\text{ولد وبنت على الأقل}) = 1-\left(\frac{1}{16}\right)-\left(\frac{1}{16}\right) = \frac{14}{16}$$

<u>(جـ) : التوزيع فوق الهندسي:</u>

فى حالة عدم الاستقلال الإحصائي كما هو الحال عند اخذ عنة من مجتمع محدود بدون إحلال ، فإنه لا يمكن إستخدام توزيع ذي الحدين لأن الأحداث تكون غير مستقلة وفى هذه الحالة يستخدم التوزيع فوق الهندسي وقانونه كالأتي:

$$P_H = \frac{C_r^A \quad C_{n-r}^B}{N_n^N}$$

ويقتبس إحتمال عدد نجاحات r ن فى عينة حجمها n مأخوذة عشوائياً بدون إحلال من مجتمع حجمه N.

حيث A هى عدد الوحدات التي تتحقق فيها الخاصية التي سميت نجاحا B هى فرصة عدم ظهور (A) أى فرض الفشل وهى تمثل العدد الباقي فى المجتمع. ومنها فإن N=A+B ، r هى عدد مرات النجاح المطلوبة فى العينة المسحوبة.

<u>مثال رقم (12):</u>

أحسب إحتمال 2 من الرجال فى عينة من 6 أفراد اختيرت عشوائياً بدون إحلال من مجموعة من 10 أفراد منها 5 رجال ن ثم وضح كيف تكون النتيجة لو استخدمنا (خطأ) توزيع ذي الحدين؟

<u>الحل</u>

من المعطيات يمكن الوصول إلى ما ياتي:

N=10 A=5 B=5 n=6 r=2

$$P(x) = \frac{C_2^5 \quad C_4^5}{N_6^{10}}$$

$$= \frac{5i}{2i \quad 3i} \times \frac{5i}{4i \quad 1i} / \frac{10i}{6i \quad 4i}$$

$$= \frac{(5) \quad (10)}{210} = 0.24$$

ثانياً : فى حالة إستخدام توزيع ذى الحدين فإن النتيجة كانت ستكون كما يلي:

$$p = \frac{1}{2} \quad , q = \frac{1}{2} \quad n = 6 \quad r = 2$$

$$P(X=2) = \frac{ni}{xi(n-x)i} P^r q^{n-r}$$

$$= \frac{6i}{2i \quad 4i} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4 = \frac{15}{64} = 0.23$$

على أننا نلاحظ أنه عندما يكون حجم العينة صغيراً جداً بالنسبة لحجم المجتمع (اقل من 5% مثلاً من المجتمع)، فإن المعاينة أو السحب بدون إحلال يكون تأثيره صغيراً على إحتمال النجاح فى كل محاولة ويكون توزيع ذي الحدين (الأسهل فى الاستخدام) تقريباً جيدا للتوزيع فوق الهندسي.

مثال رقم (13):

كيس به 4 كور بيضاء و5كور زرقاء سحبت ثلاثة كور على التوالي فوجد احتمال الحصول على كرة على الأقل بيضاء.

الحل

واضح أن السحب هنا بدون إحلال أى أن الحوادث غير مستقلة

ومن معطيات المثال السابق يتضح أن

N=q A=4 B=5 n=3 $r \geq 1$

كرة على الأقل بيضاء تعني ما يلي:

$$p(x \geq 1) = p(x = 1، 2، 3)$$

$$p(x=1) = \frac{C_1^4 \quad C_2^5}{C_3^q} \qquad (1)$$

$$p(x=2) = \frac{C_1^4 \quad C_1^5}{C_3^q} \qquad (2)$$

$$p(x=3) = \frac{C_3^4 \quad C_0^5}{C_3^q} \qquad (1)$$

p(على الأقل كرة بيضاء) = (1)+(2)+(3)

هناك حل أخر

على الأقل كرة واحدة بيضاء يعني (إحتمال أن جميع الكور وليست بيضاء)-حيث أن إحتمال أن جميع الكور ليست بيضاء يمكن الحصول عليه كما يلي:

$$\frac{C_3^5 \qquad C_0^4}{N_3^q}$$

الإحتمال المطلوب هو

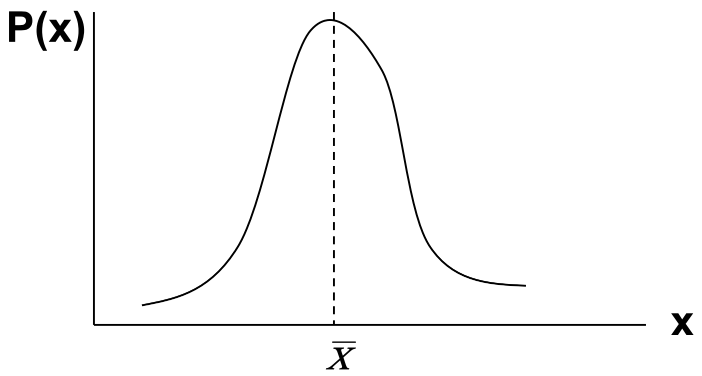
$$p(x \geq 1) = 1 - p(x = 0)$$

ثانياً : التوزيعات الاحتمالية المتصلة

التوزيع الطبيعي : Normal Distribution

يعد التوزيع الطبيعي من أهم التوزيعات الاحتمالية وأكثرها انتشارا إذا يتبع هذا التوزيع الغالبية العظمى من الظواهر الطبيعية ويشير هذا الفصل ٔإلى الخصائص الرياضية لهذا التوزيع إلى كل من لابلاس Piere Laplace ، ودي موافر Abrahamde Moiver ، وجاوس Garl Guass

.



ويمكن تمثيل التوزيع الطبيعي بالشكل المرفق ن حيث يلاحظ أن القيمة الاحتمالية تزداد لدى القيم المتوسطة لتوزيع الاحتمالي وتنقص تدريجياً بالتحرك لدى طرفي التوزيع أى باتجاه نحو القيم الصغيرة أو القيم الكبيرة للمتغير X وتنطبق هذه الخواص على العديد من الظواهر الطبيعية فعلى سبيل المثال إذ قيس أطوال عدد كبير من الأفراد فى سن الشباب فيلاحظ أن الغالبية العظمى منهم ذوى طول متوسط أو يقترب من المتوسط فى حين أن العدد يقل تدريجياً كلما إزداد الطول أو قصر عن هذا المتوسط ، وهذا يعني أن التكرار النسبي أى الإحتمال للقيم الوسطى أعلى تأثيراً من القيم الصغيرة أو القيم الكبيرة.

وتبين الدالة الاحتمالية للتوزيع الطبيعي.

$$0-\frac{1}{2}\,\frac{(x-u)^{2}}{2}$$

$$p(x)=\frac{1}{\sqrt{2\pi\sigma^{2}}}$$

حيث:

ارتفاع المنحنى الطبيعي أو القيمة الإحتمالية للمتغير p(x)=x

قيمة ثابتة وتعادل $\pi=\frac{22}{7}=3.1416$

أساس اللوغاريتم الطبيعي وهى تعادل e=2.7183

الوسط الحسابي للتوزيع = u

الانحراف المعياري للتوزيع= σ

ومن الدالة السابقة يتضح أنه بمعرفة المتوسط الحسابي للتوزيع وبيانه يمكننا تحديد القيم الاحتمالية للمتغير X كما يمكن تحديد الإحتمال التجميعي عن طريق إجراء عمليات التكامل للدالة الاحتمالية للتوزيع بين حد أدنى وحد أقصى.

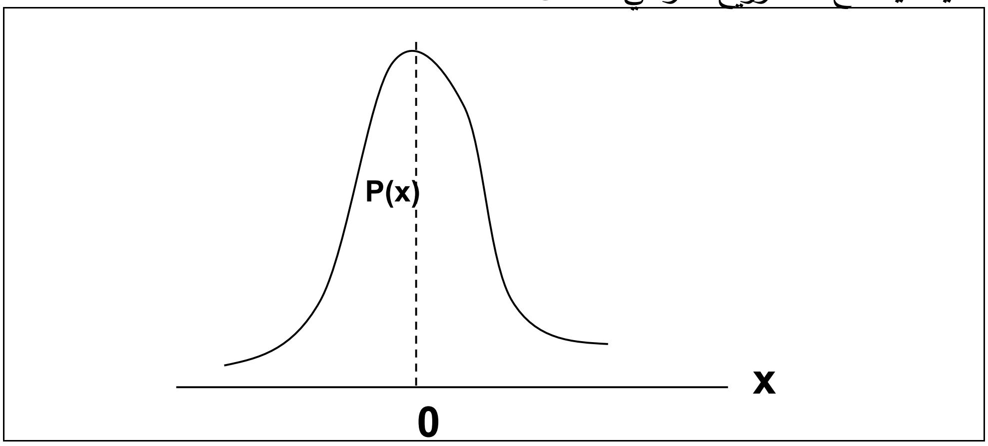
ولأهمية هذا التوزيع وانطباق خواصه على العديد من الظواهر الطبيعية ، وبفرض تسهيل عمليات الحساب فقد افترض توزيعاً طبيعياً ذو خواص محددة أطلق عليه التوزيع الطبيعي القياسي Standard Normal Distribution وقد قدرت الاحتمالات التجميعية تحت هذا التوزيع القياسي وسجلت فى جدول. ثم وضعت طريقة لتحويل قيم التوزيع الطبيعي إلى قيم التوزيع القياسي ومن ثم إستخدام هذا الجدول فى تحديد الإحتمال.

ويعد التوزيع الطبيعي توزيعاً قياسياً إذا توافرت فيه الشروط التالية:

المتوسط الحسابي =u=0

التبايـن=1= σ^2

ومن ثم فإن الشكل التالي يوضح التوزيع القياسي بخصائصه السابق بيانها، حيث يتضح أنه توزيع ناقوسي متماثل تماماً.



ويمكن استخراج الدالة الاحتمالية للتوزيع القياسي من الدالة الاحتمالية للتوزيع الطبيعي على النحو التالي.

$$e^{-\frac{1}{2}} \frac{(x-u)}{2}$$

$$p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}$$

وبالتعويض عن قيم u، σ^2 فى التوزيع القياسي:

$$p(x)=\frac{1}{\sqrt{2\pi}} \qquad e^{-\frac{1}{2}x^2}$$

ويمكن تحويل قيم التوزيع الطبيعي إلى قيم التوزيع القياسي فى حالة معرفة متوسط وتباين التوزيع الطبيعي حيث أن

القيم القياسية= z= $\frac{x-u}{\sigma}$ حيث

القيم الطبيعية = X

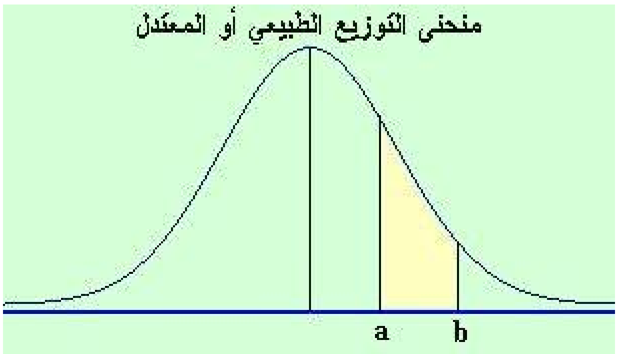
متوسط التوزيع الطبيعي =u

الانحراف المعياري للتوزيع الطبيعي= σ

خصائص التوزيع المعتدل:

1) منحنى التوزيع المعتدل متصل (مستمر) منحناه (Normal Curve) يشبه شكل الجرس ويمتد ذراعه من – ∞ إلى ∞ .

معادلته الرياضية في الفترة] – ∞ ، ∞ [هي:



$$Y=\frac{e^{-\frac{(X-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} , -\infty < X < \infty$$

بإجراء تكامل Y على الفترة] – ∞ ، ∞ [نحصل على المساحة تحت المنحنى وفوق المحور الأفقي، والتمثيل البياني له كما مبين بالشكل المقابل وكل نقطة من نقاط المنحنى تمثل قيمة لدالة تعرف بدالة كثافة الاحتمال (function Probability density) عند هذه النقطة والاحتمال هنا أي في التوزيع المستمر هو قيمة المساحة تحت منحنى دالة الكثافة المناظرة لفترة وليس لنقطة فالمساحة باللون الأصفر والمحصورة بين المنحنى والمحور الأفقي والمستقيمان $x_1 = a$, $x_2 = b$ تساوي احتمال المتغير العشوائي المستمر X أي

قيمة في الفترة [a , b] هذا وأن المساحة الكلية الواقعة بين منحى التوزيع المعتدل والخط الأفقي تساوي الواحد الصحيح وهي ما تعرف بالمساحة تحت المنحنى = 1

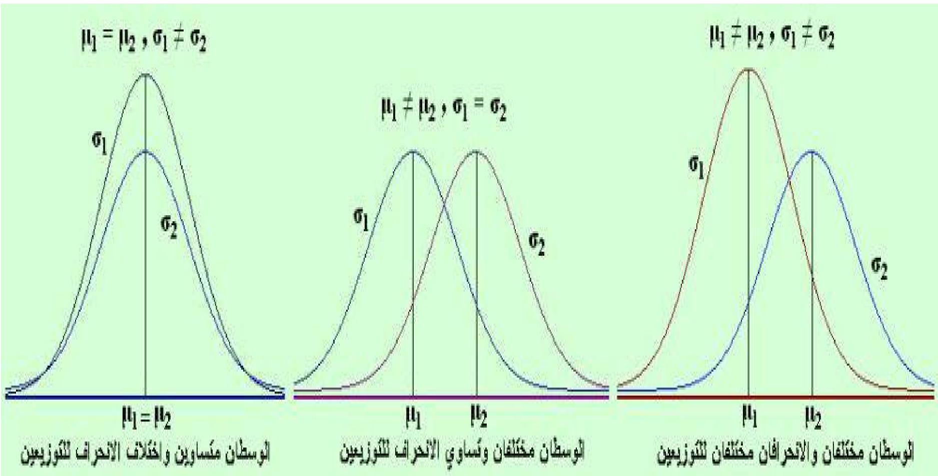
ولمعرفة احتمال وقوع x بين x_1 و x_2 نحسب تكامل الدالة السابقة من x_1 إلى x_2 ، مع ملاحظة أن احتمال أي حدث P(A) يقع بين الصفر والواحد الصحيح أي أن: 1 >P(A) > 0

2) المنحنى متماثل حول الخط الرأسي (العمود النازل من أعلى نقطة للمنحنى على الخط الأفقي) وإن التماثل يعني بأن صورة الشكل على أحد جانبي محور التماثل هي الجزء الواقع على الجانب الأخر وموقع العمود على الخط الأفقي يمثل قيمة الوسط الحسابي أي أن المنحنى متماثل حول وسطه الحسابي أو حول المستقيم $x = \mu$ ، وان μ هي

القيمة المتوقعة ويصل المنحنى لقيمته العظمى عند $X = \mu$

3) المنحنى ممتد من $-\infty$ إلى $+\infty$ ولا يلتقي بالمحور الأفقي.

4) للمنحنى المعتدل معلمتين هما الوسط الحسابي والانحراف المعياري معتمد كلياً عليهم فاختلاف الوسط أو الانحراف المعياري لتوزيعين معتدلين يعني اختلاف في الشكل أو اختلاف في المركز كما مبين بالشكل الآتي ولكل زوج (μ ، σ) للوسط والانحراف المعياري منحنى توزيع مختلف وبالتالي تختلف المساحة تحت المنحنى لكل منحنى ولذا أخذنا (0 ، 1) كتوزيع معياري يسمى التوزيع الطبيعي المعياري متغيره العشوائي هو Z السابق ذكرها



5) للمنحنى قمة واحدة أي له منوال واحد وبالتالي فالمنحني وحيد المنوال

6) المتوسطات الثلاثة متساوية (الوسط والوسيط والمنوال) بالنسبة للمتغير العشوائي المعتاد.

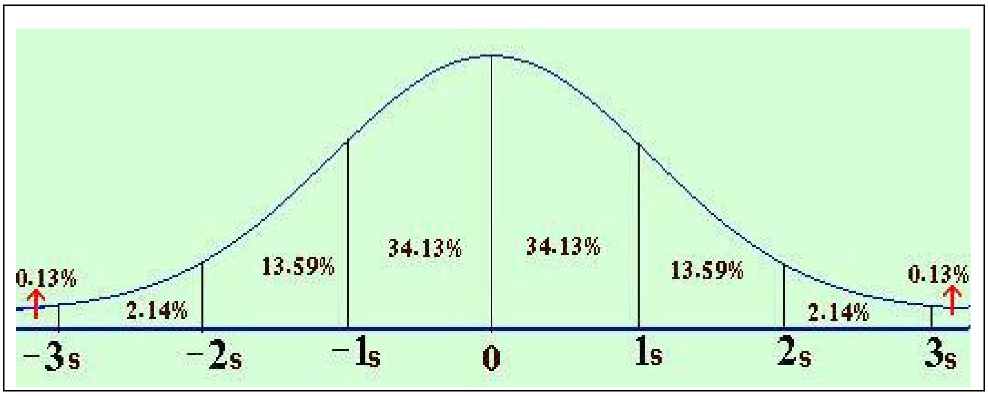
7) المساحة الواقعة تحت المنحنى والمحصورة بالمستقيمين:

$x = \mu - \sigma$ و $x = \mu + \sigma$ تساوي 68.26% تقريباً من المساحة الكلية تحت المنحنى أي 68.26% من قيم المتغير العشوائي المعتاد تقع في $[\mu + \sigma ، \mu - \sigma]$

$x = \mu - 2\sigma$ و $x = \mu + 2\sigma$ تساوي 95.45% تقريباً من المساحة الكلية تحت المنحنى أي 95.45% من قيم المتغير العشوائي المعتاد تقع في $[\mu + 2\sigma$ ، $\mu - 2\sigma]$

$x = \mu - 3\sigma$ و $x = \mu + 3\sigma$ تساوي 99.73% تقريباً من المساحة الكلية تحت المنحنى أي 99.73% من قيم المتغير العشوائي المعتاد تقع في $[\mu + 2\sigma$ ، $\mu - 2\sigma]$

أي أن وقوع أي مفردة على بعد 1، 2، 3 انحرافات معيارية (s, 2s, 3s1) من الوسط الحسابي هي القيم السابقة كما مبين بالشكل الآتي:



لاحظ أن 34.19% من المساحة تحت المنحنى التي تساوي الواحد الصحيح أي 0.3413 ، وبجمع القيم المبينة في الرسم أعلاه نجد أنها تساوي الواحد الصحيح تقريباً.

إن هذه القيم ما هي إلا احتمالات للقيم كمساحة تحت المنحنى ولأي دالة احتمال يكون مجموع احتمالاتها البسيطة يساوي الواحد الصحيح ونقصد في الأصل المساحة هنا لمساحة الأعمدة للقيم ولكن من الصعب رسم كل الأعمدة وعرض احتمال كل منها ولذا استعضنا عنها باحتمالاتها.

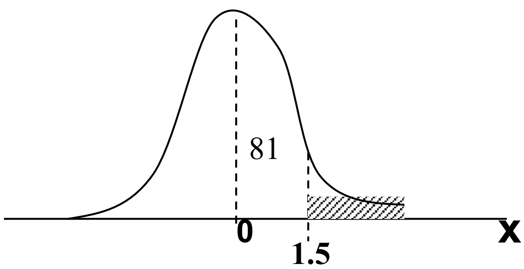
$$0.0013 + 0.0214 + 0.1359 + 0.3413 + 0.3413 + 0.1359 +$$
$$0.0214 + 0.0013 = 0.9998 \approx 1$$

والتوزيع الطبيعي المعياري (Standard Normal Distribution) الذي وسطه صفر وانحرافه المعياري 1 متغيره العشوائي المعياري Z بالصيغة السابق ذكرها، ومنحناه كما مبين أعلاه ويمكن حذف s من القيم على الخط الأفقي وقد نضع قيم x والمناظرة لها Z على الخط الأفقي إن دعت الحاجة.

<u>مثـال :</u>

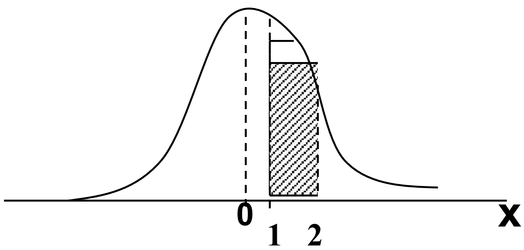
إذا فـرض أن xموزعـة توزيعـاً طبيعيـاً قياسـياً فأوجـد إحتمـال 1.5 ≤ x واحتمال 1≤x≤2.

<u>الحـل:</u> إحتمـال1.5 ≤ x هـو عبـارة عـن إجمـالي المسـاحة المظللـة فـى الشـكل المرفق.



وبالكشف فى جدول (z) عن القيم 2 تبين أن المساحة من القيمة 0 إلى القيمة 2 تعادل 0.4332 وهذا يعني أن المساحة المظللة $0.5-0.4332=0.0668$ أى أن $P(x \ge 1.5)=0.0668$

وبنفس الأسلوب يمكن تقدير $p(2 \ge x \ge 1)$



بالكشف فى جدول (z) عن القيمة الاحتمالية المقابلة للرقم2 والقيمة الاحتمالية للرقم 1 ثم طرح الثانية من الأولى كما هو موضح فى الشكل المرفق.

حيث :

$$p(2 \ge x \ge 1)=0.4772-0.3413$$
$$= 0.1359$$

مثال :

إذا فرض أن عمر المصابيح الكهربائية التي تنتجها شركة ما موزع توزيعاً طبيعياً بمتوسط قده 100 ساعة اختراق وتباين قدره 64 فإذا اختير مصباح واحد عشوائياً فما هو إحتمال أن يكون المصباح عمره بين 110، 120 ساعة اختراق.

الحل:

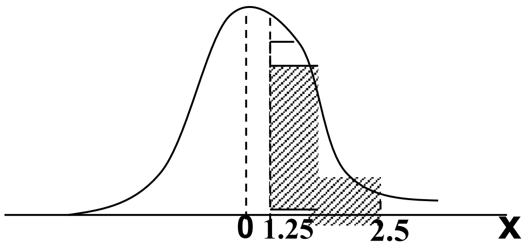
أولاً : تحويل قيمة التوزيع الطبيعي إلى توزيع طبيعي قياسي.

$$p(120 \ge x \ge 110) = \frac{p(120-110)}{8} \ge Z \ge \frac{110-100}{8}$$

$$= \frac{P(20)}{8} \ge Z \ge \frac{110}{8} = P(2.5 \ge Z \ge 1.25)$$

ثانياً الكشف فى جدول z عن القيمتين 2.5، 1.25.وطرح القيمتين تحصل على المطلوب

$$p(2.5 \ge z \ge 1.25)=0.4938-0.3944$$
$$=0.0994$$



0
1.25
2.5
x

تذكر

- تعرف الاحتمالات أنها فرع من فروع الرياضة التطبيقية الذي يهتم بدراسة اثر عامل الصدفة.
- يشير الإحتمال عادة على فرصة حدوث حدث معين
- **الحادثة** تعني ظهور حدث معين
- **المحاولة** عملية يرجي منها ملاحظة حدث معين
- **التجربة** تعرف على أنها مجموعة من المحاولات
- **الحالات الممكنة:** هى جميع الحادثات البسيطة الممكن الحصول عليها من إجراء المحاولة فى تجربة معينة
- يمكن تقسيم الحادثات التي تتعدد في حياتنا اليومية الي حادثات متماثلة وحادثات متنافية :
- يقصد بالحادثات المتماثلة بأنها تلك الحادثات التي لها نفس فرص الحدث
- يقصد بالحادثات المتنافية أنها الحادثات التي لا يمكن أن تحدث مع بعضها البعض، بمعنى أن ظهور حادثة فيها ينفي ظهور الحادثات الأخرى
- يطلق على الحادثة أنها مستقلة إذا كانت فرصة ظهورها لا تؤثر على فرصة ظهور الحادثات الأخرى
- **من خصائص الإحتمال:** إن قيمة الإحتمال تتراوح بين حد أدنى قدره صفر وحد أعلى قدره الواحد الصحيح ، مجموع الاحتمالات الممكنة لظهره ما لابد وأن تساوى الواحد الصحيح
- **يعنى جمع** الاحتمالات تحقق حادثتين أو أكثر فى نفس الوقت، ويعبر عنه بحدوث الحادثة الأولى أو الحادثة الثانية وتعنى كلمة أو هنا تحقق أى من الحادثتين أو كليهما معاً.
- تسمى الأحداث مستقلة إذا كان حدوث أى منها ليس له أى تأثير على احتمالات حدوث أى من الأحداث الأخرى.
- الأحداث مستقلة يكون إحتمال حدوثها جميعا يساوي حاصل ضرب احتمالاتها الفردية.
- $P(E_1E_2)=P(E_1)P(E_2)$
- الإحتمال الشرطي يعني تحقق حادثة ما بشرط تحقق حادثة أخرى
- التباديل تعني تبديل مجموعة من الحادثات على أنه ترتيب معين لحدوث هذه الحادثات، وعندما يتغير هذا الترتيب يحصل على تبديل آخر وتستخدم القاعدة التالية لمعرفة عدد التباديل

- أي أن التباديل هى عبارة عن ترتيب للحادثات ، فإذا اختلف هذا الترتيب أدى ذلك إلى ظهور تباديل جديدة ، على الرغم من أن هذه التباديل قد تحتوى على نفس المكونات ولكن بترتيب مختلف، أما التوافيق فإنها تعنى للحادثات بغض النظر عن ترتيبها. ففي المثال السابق لدينا ستة تباديل كل منها مكون من الثلاثة حروف ، وعلى ذلك يكون لدينا توفيقه واحدة.
- تنقسم القيم المتوقعة أو العزوم إلى قسمين رئيسين أولهما المركزية وثانيهما العزوم التشتتية
- وتتعدد العزوم المركزية للتوزيعات الاحتمالية المنفصلة فهناك العزم المركزي الأول (التوقع الأول)، وهناك العزم المركزي الثاني، والثالث وهكذا
- **العزم الأول** = $E(x)=\sum x_i P(x)$ **وهذه القيمة المتوقعة تعادل المتوسط الحسابي**

العزم المركزي الثاني = $E_2(X)=\sum X^2 P(X)$

العزم المركزي الثالث = $E_3(X)=\sum X^3 P(X)$

العزم رقم n = $E_n(X)=\sum X^n P(X)$

- العزوم التشتتية تعني **تغير أو تشتت التوزيع الاحتمالي حول متوسطه الحسابي**
- العزم التشتتي الأول = $E_1'(X)=\sum[X-E(X)]P(X)$
- وهذا العزم يعني ببساطة مجموع انحرافات قيم التوزيع الاحتمالي عن وسطها الحسابي (العزم المركزي الأول) وهو بذلك لابد وأن يساوي الصفر ويمكن إثبات ذلك كالتالي:
- كما يعتبر العزم التشتتي الثاني من أهم العزوم حيث يطلق عليه التباين Variance ويرمز له بالرمز σ^2 وهو كالتالي:
- ويمكن استخراج قيمة هذا العزم التشتتي الثاني σ^2 بمعرفة العزمين المركزين الأول والثاني.
- العزم المركزي الأول للتوزيع = $\overline{X}$ = إحتمال النجاح(p) وكذلك يمكننا استخراج العزم المركزي الثاني كالتالي:
- $E_2(x)=\sum X^2 P(X)$
- بمعرفة العزمين المركزين الأول والثاني يمكننا استخراج العزم التشتتي الثاني (X) σ^2 (ويساوي حاصل ضرب احتمال النجاح في احتمال الفشل) وياخذ الصيغة التالية :
- $E_2'(X)=\sigma^2 X=E_2(X)-\{E_1(X)\}^2=$ Pq

- يعتبر توزيع ذو الحدين أحد التوزيعات الاحتمالية المنفصلة وهو باختصار تكرار للتوزيع البرنولى، أى أنه يتكون من مجموعة من المحاولات كل منهما إما أن تنجح المحاولة باحتمال p ويأخذ المتغير x القيمة واحد أو تفشل باحتمال q ويأخذ المتغير القيمة صفر
- **خصائص أو شروط توزيع ذو الحدين:**
- أن تتكون التجربة من عدد من المحالات المستقلة عن بعضها البعض.
- أن يكون لكل محاولة ناتجان ممكنا ومتنافيات لكل محاولة .
- إن إحتمال وقوع الحادث المطلوب (النجاح ولا يتغير من محاولة إلى أخرى).
- يستخدم التوزيع فوق الهندسي فى حالة عدم الاستقلال الإحصائي كما هو الحال عند اخذ عنة من مجتمع محدود بدون إحلال ، حيث لا يمكن إستخدام توزيع ذي الحدين لأن الأحداث تكون غير مستقلة وقانونه كالأتي:

$$P_H = \frac{C_r^A \qquad C_{n-r}^B}{N_n^N}$$

- عندما يكون حجم العينة صغيراً جداً بالنسبة لحجم المجتمع (اقل من 5% مثلاً من المجتمع)، فإن المعاينة أو السحب بدون إحلال يكون تأثيره صغيراً على إحتمال النجاح فى كل محاولة ويكون توزيع ذي الحدين (الأسهل فى الاستخدام) تقريباً جيدا للتوزيع فوق الهندسي
- يعد التوزيع الطبيعي من أهم التوزيعات الاحتمالية وأكثرها انتشارا إذا يتبع هذا التوزيع الغالبية العظمى من الظواهر الطبيعية
- يمكن تحويل قيم التوزيع الطبيعي إلى قيم التوزيع القياسي فى حالة معرفة متوسط وتباين التوزيع الطبيعي

تمارين

(1) – يصاب أشجار البرتقال فى حديقة من الحدائق بنوعين من الأمراض A، B فإذا كان إحتمال أن تصاب شجرة من الأشجار بالمرض A هو $\frac{1}{5}$ واحتمال أن تصاب بالمرض B هو $^{1}/_{4}$ وذلك بغض النظر إذا كانت مصابة بالمرض A. احسب إحتمال أن الشجرة مصابة بـ:

(أ) كلا المرضين A، B.

(ب) مرض واحد فقط A أو B.

(جـ) غير مصابة على الإطلاق.

ما هى العلاقة البسيطة الموجودة بين إجابتك لـ(أ) و(ب)و(جـ).

(2)– إذا رميت قطعة نقود متوازنة 6 مرات أوجد إحتمال ظهور

(أ) 1 (ب) 2 (جـ) 3 (د) 4 (هـ) 5 صورة

(3)– فى رمية واحدة لست عملات غير متحيزة أوجد إحتمال ظهور

(أ) 2 أو أكثر صورة (ب) أقل من 4 صور

(4)– إذا كانت X تعبر عن عدد الصور لرمية واحدة لأربع عملات متوازنة أوجد

(أ)$Pr(X=3)$ (ب)($X<2$)

(جـ) $Pr(X\leq 3)$ (د)($Pr(1\leq X\leq 3)$)

(5)– فى 800 عائلة بكل منها 5 أطفال ، ما عدد الأسر المتوقع أن يكون بها

(أ) 3 أولاد (ب) 5 بنات (جـ) 2 أو 3 أولاد

مفترضا أن إحتمال بنت أو ولدا إحتمال متساوي.

(6)– أوجد احتمال الحصول على ما مجموعه 11.

(أ) مرة واحدة (ب) مرتان، فى رميتين لزهرتين متوازنتين.

(7)– أوجد احتمال تخمين الإجابة الصحيحة على 6 أسئلة على الأقل من 10 أسئلة فى امتحان "خطأ – صواب".

(8)– مندوب تأمين باع بوالص تأمين إلى 5 أشخاص ، جميعهم فى نفس العمر وفى صحة جيدة طبقاً لجداول التأمين قارن إحتمال بقاء شخص على قيد الحياة له هذه المواصفات 30 عاما تماما هو $^{2}/_{3}$ أوجد إحتمال أنه فى خلال 30 عاما يبقى على قيد الحياة.

(أ) كل الـ5 رجال (ب) على الأقل 3 رجال (جـ) رجلان فقط،

(د) على الأقل رجل واحد

(9)- ما هو إحتمال إختيار (أ) سيدتين فى عينة من 5 أفراد مسحوبة عشوائياً من مجموعة 9 أفراد منهم 4 سيدات.

(10)- فى امتحان مادة الإحصاء كان المتوسط الحسابي لدرجات الطلبة 78 والانحراف المعياري 10.

(أ) اوجد الدرجات المعيارية لطالبين درجاتهما 62،93.

(ب) أوجد درجات طالبين درجاتهما المعيارية 0.6-1.2.

(11)- أوجد (ا) الوسط الحسابي (ب) الانحراف المعياري فى امتحان كانت الدرجات به 70، 88 مقابلة للدرجات المعيارية 0.6- 1.4 على الترتيب.

(12)- أوجد المساحة تحت المنحنى الطبيعي بين

(أ) Z=1.20، Z=2.40 (ب) Z=1.23، Z=1.87 (ج) -Z=2.35،Z=-0.50-

(13)- إذا كانت أوزان 400 طالبا تتوزع توزيعاً طبيعياً متوسطه 70.0 KG وإنحرافه المعيارى هو 4.0 KG كم عدد الطلبة الذين تكون أوزانهم

(أ) أكبر من 75 KG (ب) أقل من أو يساوى 65 KG (جـ) بين 65، 75 kg

(د) مساوية 70 kg مفترضاً أن القياسات مقربة إلى أقرب كيلو جرام.

الباب الرابع

الانحدار والارتباط

Linear Regression and Correlation

تمهيد:

يعد الانحدار والارتباط من الموضوعات ذات الأهمية الخاصة فى الدراسات الإحصائية من الناحيتين النظرية والتطبيقية. وبرغم مناقشة هذا الفصل لبعض الجوانب النظرية فان الاهتمام الأساسي يقتصر على الجانب التطبيقي، وبشكل اكثر تحديدا يهدف هذا الفصل إلي تمكين الطالب من استخدام مفاهيم وأدوات الانحدار والارتباط فى تقدير معامل العلاقة بين المتغيرات واستخدام هذه المعالم فى التنبؤ بقيم المتغيرات فى المستقبل، وقياس درجة قوة هذه العلاقة مع أجراء الاختبارات اللازمة للتأكد من معنوية هذه المعالم، وان كانت المعالجة التفصيلية للموضوع الأخير (اختبارات صحة الفرض) ستكون محل مناقشة فى فصل آخر من هذه المذكرات.

والفكرة الأساسية فى موضوع الانحدار والارتباط هى وجود متغيرين (أو اكثر) تعتمد قيمة أحدهما على الآخر. ويشار إلي أحدهما بالمتغير التابع (Y)، والثاني بالمتغير المستقل (X).

ومن أمثلة هذه العلاقات فى المجال الزراعي: العلاقة بين الكمية المستخدمة من السماد ومتوسط محصول الفدان من محصول معين، أو العلاقة بين كمية الأعلاف المستخدمة والزيادة فى وزن الحيوان، أو العلاقة بين حجم المزرعة والدخل الناتج من الأنشطة المزرعية بها.

وفى هذا الإطار يقصد بالانحدار كمية التغير فى المتغير التابع المرتبطة بتغير قدرة الوحدة فى المتغير المستقل، بينما يحدد الارتباط مدى قوة العلاقة بين هذين المتغيرين.

ويناقش هذا الباب عدد من الموضوعات الأساسية فى تحليل الانحدار والارتباط منها: النموذج الاحتمالي الخطى البسيط، طريقة المربعات الدنيا، حساب تباين الخطأ ومعامل الارتباط ومعامل التحديد والتحويلات الخطية لبعض العلاقات غير الخطية بالإضافة إلي مجموعة من الأمثلة التطبيقية والتمارين.

النموذج الاحتمالي الخطى البسيط:Simple Linear Probabilistic Model

يهتم هذا الجزء بالدوال الخطية، ويرجع ذلك إلي انه بالرغم من بساطة الصورة الخطية الا أنها منتشرة للغاية فى الواقع وفى الدراسات التطبيقية، هذا بالإضافة إلي أن الكثير من المشاكل غير الخطية يمكن تحويلها بسهولة لصورة خطية، كما أن العلاقات الخطية يمكن أن تمثل على الأقل جزء من العلاقات غير الخطية. ولتبسيط عرض هذا النموذج يتم الاعتماد على المثال التالى:

مثال (1)

تحتاج شركة من شركات المنتجات الغذائية إلي التنبؤ بحجم مبيعاتها الشهرية. وبافتراض أن المتغير الممكن استخدامه للتنبؤ بهذا المتغير التابع (حجم المبيعات الشهرية) هو حجم الأنفاق على الدعاية خلال الشهر. أذن يمكن تحديد هدف هذه الشركة كما يلي:

أ- تحديد طبيعة العلاقة بين الأنفاق على الدعاية وحجم المبيعات الشهرية، أو بعبارة أخرى تحديد هل هناك علاقة بين الأنفاق على الدعاية وحجم المبيعات الشهرية أم لا؟.

ب- الحصول على معادلة لاستخدامها فى التنبؤ بحجم المبيعات. ولتحقيق هذا الهدف تم الحصول على عينه من الأنفاق على الدعاية (X) وما يصاحبها من مبيعات خلال عشرة شهور مختارة عشوائيا، وتم تسجيل هذه البيانات فى الجدول رقم (1).

والخطوة الأولى فى تحليل بيانات الجدول هو رسم شكل انتشارى (Scatter Diagram). وفى هذا الشكل تعتبر المبيعات الشهرية متغيرا تابع (X) والأنفاق على الدعاية متغيرا مستقلا (X) ومن الشكل رقم (1) يتبين أن (Y) تزداد بزيادة (X).

وللحصول على معادلة تفيد فى التنبؤ بقيم المتغير (Y) بمعلومية قيم المتغير يمكن (تحت افتراض أن العلاقة خطية) رسم خط مستقيم يمر خلال نقاط الشكل رقم (1) ويوفر افضل توفيق لهذه البيانات (Best fit) ويمكن للخط المشار إليه أن يوفر حلا لمشكلة التنبؤ بالمبيعات الشهرية بها الشركة كدالة لانفاقها على الدعاية خلال نفس الشهر.

ويلاحظ أن الخطوات السابقة تضمنت اختيار نموذج رياضي يعبر عن العلاقة الدالية المفترضة بين (Y, X)، وهذا النموذج هو الخط المستقيم ومن المعلوم أي المعادلة الرياضية للخط المستقيم.

$$\mathbf{Y= A + B\ X}$$

حيث:

الجزء المقطوع من محور A = Intercept = Y

ميل الخط المستقيم = B= Slope

ويسمى هذا النموذج الرياضي المحدد لأنه بالتعويض بقيم (X) فى المعادلة الأخيرة نحصل على قيم (Y) المقابلة لها بدون احتمال للخطأ.

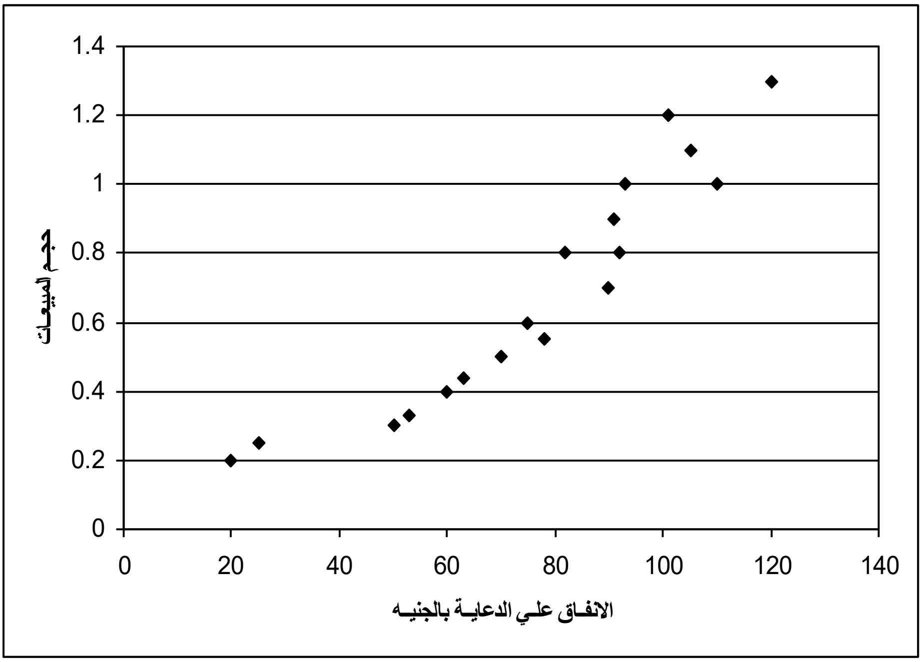
ولكن فى الواقع فان اغلب الدراسات العلمية تعتمد على نوع آخر من النماذج هو النموذج الرياضي الاحتمالي وهذا النوع من النماذج يحتوى على واحد أو اكثر من العناصر العشوائية والتى لها توزيع احتمالي معين.

جدول رقم (1)

الأنفاق على الدعاية وحجم المبيعات للشركة خلال عشرة شهور مختارة عشوائيا

الشهر	الأنفاق على الدعاية (X) الف جنيه	حجم المبيعات (Y) الف جنيه
1	1.2	101
2	0.8	92
3	1.0	110
4	1.3	120
5	0.7	90
6	0.8	82
7	1.0	93
8	0.6	75
9	0.9	91
10	1.1	105

شكل رقم (1)
الشكل الانتشارى لبيانات جدول رقم (1)



وبالنسبة للمثال السابق يمكن الربط بين حجم المبيعات وحجم الأنفاق على الدعاية بالمعادلة التالية:

$$Y= A + B X + E$$

حيث :

متغير عشوائي = الخطأ العشوائي = E

القيمة المتوقعة للخطأ العشوائي = E (E) = 0

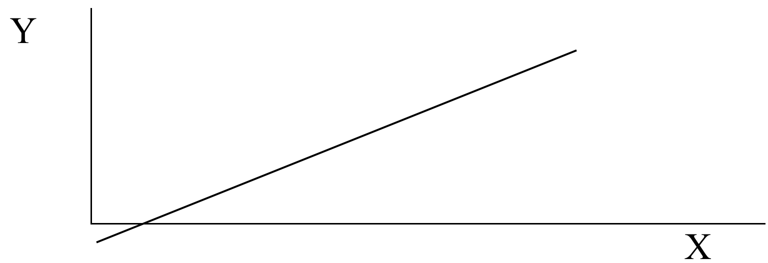
$$\sigma^2 E= \sigma^2$$

كما يفترض أن أي زوجين من الأخطاء العشوائية (E_i E_j) والمرتبطين بالملاحظتين (Y_i Y_i) مستقلين. بعبارة أخرى يفترض هذا النموذج أن المتوسط أو القيمة المتوقعة للمتغير (Y) ترتبط خطيا بالمتغير (X)، وان القيم المشاهدة من (Y) ستختلف أو تنحرف عن هذا الخط (لأعلى أو لأسفل) بكمية عشوائية (E)،

ويفرض أيضا أن توزيع الأخطاء حول الخط سيكون متماثلا بغض النظر عن قيمة (X)، وان هذه الأخطاء مستقلة عن بعضها البعض.

ويبين الشكل رقم (2) الخط المتوقع لتوفيق نقاط تمثل القيم المتوقعة لـ (Y) والمصاحبة لقيم معينة من (X) ويوضح نفس الشكل الأخير التوزيع الاحتمالي للخطأ العشوائي (E) لقيم مختلفة من (X) ويلاحظ أن (σ) تقيس انتشار هذه التوزيعات المتماثلة للخطأ العشوائي.

شكل رقم (2) النموذج الاحتمالي الخطى



طريقة المربعات الدنيا The Method of Least Squares

من الشكل الانتشارى لبيانات المثال السابق يتضح انه لا يوجد الخط المستقيم الذي يمكن أن يمر بجميع النقاط الممثلة للبيانات، وبالتالي فان اختيار أى خط مستقيم لابد وان يترتب عليه أن بعض النقاط تنحرف عنه بمقدار عشوائي (E). وقد سبق الإشارة فى الباب الثاني من هذ1 الكتاب إلي أن مجموع مربعات الانحرافات عن متوسط مجموعة من الملاحظات يمثل مقياسا للتشتت داخل هذه المجموعة، أو التباين حيث تم تعريف هذه الكمية كما يلي:

$$\sigma^2 X = \sum\left(X - \overline{X}\right)^2 / N$$

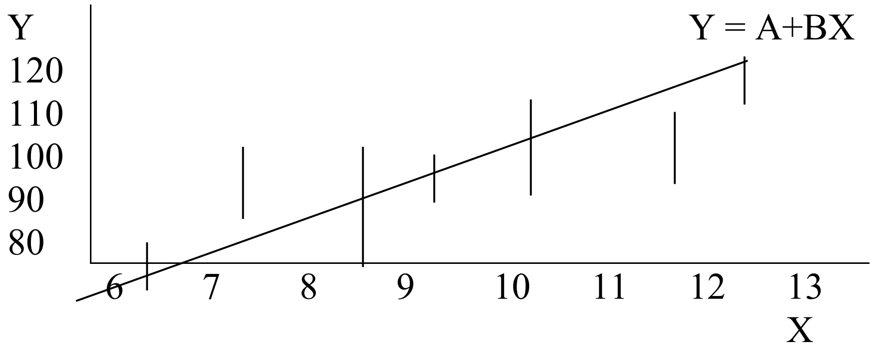
لذلك يبدو منطقيا أن يقاس التشتت أو الانحراف حول الخط المستقيم بحساب مجموع مربعات الانحرافات عن هذا الخط. وباعتبار هذا المقدار مقياسا لدرجة جودة التوفيق يتم اختيار الخط المستقيم الذي يرتبط به اقل قدر من مجموع مربعات الانحرافات. ويسمى هذا الأسلوب فى اختيار الخط المستقيم المناسب لتمثيل مجموعة من الملاحظات أسلوب المربعات الدنيا.

وهدف طريقة المربعات الدنيا هو التوصل للنهاية الدنيا لدالة مجموع مربعات الانحرافات عن الخط المستقيم ولتوضيح هذه الفكرة يفترض أن القيمة المتوقعة لـ (Y) والتى تم التوصل إليها من توفيق الخط المستقيم = (Y).

ومعادلة التنبؤ = $\hat{Y} = \hat{A} + BX$

حيث : تقديرات للقيم الحقيقية $(A, B) = \hat{A}, B$

شكل رقم (3) معادلة التنبؤ الخطية



ويبين الشكل رقم (3) الخط الممثل لمعادلة التنبؤ وانحرافات النقاط عن هذا الخط حيث تمثل الخطوط الرأسية من خط التنبؤ لكل نقطة انحراف هذه النقطة عن القيم المتوقعة من (Y) وعلى هذا فان انحراف النقطة (i) يعادل المقدار $\left(Y_i - \overline{Y_i}\right)$ وعند اختيار افضل الخطوط المستقيمة تمثيلا للعلاقة محل الاهتمام يعتمد على مبدأ المربعات الدنيا الذي يمكن صياغته كما يلي:

"افضل خط مستقيم أو الخط الذي يحقق افضل توفيق للبيانات الممثلة للعلاقة بين المتغيرين (Y, X) هو الخط الذي يتحقق به تدنية مجموع مربعات انحرافات القيم المشاهدة لـ (Y_i) عن القيم المتوقعة (Y_i) لاقل ما يمكن":

والمقدار المشار إليه أو الدالة المطلوبة تدنيتها يسمى أيضا مجموع مربعات الأخطاء = S S E = Sum Squares of errors

ويمكن تعريف : $S\,S\,E = \Sigma(Y_i - \hat{Y}_i)$

والتعويض عن $\hat{Y}_i$ فى هذه الصورة

$$SSE = \sum_{i=1}^{n} \left(Y_i - \left(A + BX\right)\right)^2$$

وبإيجاد النهاية الصغرى للدالة (S SE) ويمكن التوصل لتقديرات لمعالم الخط المستقيم (A, B) تحقق شرط المربعات الدنيا كما يلي

$$L = \sum_{i=1}^{n} \left[Y_i - \left(A + BX_i \right) \right]^2$$

$$\Delta L / \Delta A = -2 \sum \left[Y_i - (A + BX_i) \right]^2 = 0$$

$$\Delta L / \Delta B = -2 \sum_{xi} \left[Y_i - (A + BX_i) \right]^2 = 0$$

وبهذا يمكن الحصول على المعادلتين التاليتين (معادلات المربعات الدنيا)

$$\sum_{i=1}^{n} Y_i = nA + B \sum_{i=1}^{n} X_i$$

$$\sum_{i=1}^{n} Y_i X_i = A \sum_{i=1}^{n} X_i + B \sum_{i=1}^{n} X_i^2$$

وحل هاتين المعادلتين آنيا يوفر التقديرات المطلوبة (A, B) كما يمكن التوصل للصور التالية لهذه التقديرات

$$\hat{A} = \overline{Y} - \hat{B} \, \overline{X}$$

$$\hat{B} = \frac{\sum X_i Y_i - n \overline{X} \, \overline{Y}}{\sum X_i^2 - n \overline{X}^2}$$

وبهذا يمكن استخدام الصور الأخيرة لحساب كل من (Â ,B) للمثال الخاص بأثر الدعاية على المبيعات وبالاعتماد على بيانات جدول رقم (2) يتم الحصول على ما يلي:

$$\hat{B} = \frac{\sum X_i Y_i - n\overline{X}\overline{Y}}{\sum X_i^2 - n\overline{X}^2}$$

$$\hat{B} = \frac{924.8 - (10)(0.94)(95.9)}{9.28 - (10)(0.94)^2} = 52.567$$

وعلى ذلك وطبقا لمبدأ المربعات الدنيا فان افضل الخطوط المستقيمة توفيقا للبيانات الخاصة بمتغيرات الأنفاق على الدعاية وحجم المبيعات هو الخط الذي تمثله المعادلة:

$$\hat{Y} = \hat{A} + \hat{B} X = 46.49 + 52.57$$

والشكل رقم (3) يوضح هذا الخط ويمكن شرح هذه المعادلة كما يلي:

ابتداء من حجم مبيعات يعادل 46.49 الف جنيه فان زيادة الأنفاق على الدعاية بمقدار وحدة واحدة (الف جنيه) فى الشهر يصحبها زيادة فى حجم المبيعات الشهرية تقدر بحوالي 52.57 وحدة (الف جنيه)

جدول (2) حساب بيانات فى جدول رقم (1)

Y_1^2	X_1Y_1	X_1^2	X_1	Y_1
10.201	121.2	1.44	1.2	101
8.464	73.6	0.64	0.8	92
12.100	110.0	1.00	1.0	110
14.400	156.0	1.69	1.3	120
8.100	63.0	0.49	0.7	90
6.724	65.0	0.64	0.8	82
8.649	93.0	1.00	1.0	93
5.625	45.0	0.36	0.6	75
8.281	81.9	0.81	0.9	91
11.025	115.5	1.21	1.1	105
93.569	924.8	9.28	9.4	Sum 959

بالنسبة لهدف التنبؤ بقيم المتغير التابع (Y) بمعلومية قيمة معينة للمتغير المستقل (X) من الممكن الرجوع للشكل رقم (3) أو التعويض فى معادلة التنبؤ السابقة، فإذا قررت الشركة أن تتفق 1000 ج (الف جنيه) على الدعاية فى الشهر فان حجم المبيعات المتوقع يكون

$$\hat{Y} = \hat{A} + B\,X = 46.49 + 52.57(1.0)$$

$$\hat{Y} = 99.060$$

أو بعبارة أخرى فان القيمة المتوقعة من المبيعات الشهرية عند هذا المستوى من الأنفاق على الدعاية تعادل 99.060 الف جنيه.

ملاحظات على الانحدار الخطى:

1) عندما يكون المطلوب هو تقدير قيمة للمتغير Y المقابلة لقيمة معطاة للمتغير x وذلك باستخدام بيانات عينة ويتم ذلك بتقدير قيمة Y من منحنى أو خط المربعات الدنيا، والخط الناتج يسمى خط انحدار Y على X حيث أن Y تقدر من X.

2) إذا كان المطلوب هو تقدير قيمة (X) من قيمة معطاة لـ (Y) فان المنحنى أو الخط المستقيم المستخدم يكون خط انحدار (X) على (Y).

وهذا يتضمن تبديل المتغيرات فى الشكل الانتشارى بحيث تكون (X) هى التابع و (Y) هى المتغير المستقل. وتصبح صورة معامل العلاقة كما يلي:

$$A = \bar{X} - B\bar{Y}$$

$$\hat{B}_1 = \frac{\sum Y_i X_i - n\bar{X}\bar{Y}}{\sum Y_i^2 - n\bar{Y}^2}$$

$$\hat{B}_1 = \frac{\sum \left(X_i - \bar{X}\right) - \left(Y_i - \bar{Y}\right)}{\sum \left(Y_i - \bar{Y}\right)}$$

3) المشاكل التي تتضمن أكثر من متغيرين تسمى بمشاكل الانحدار المتعدد. وطريقة المربعات الدنيا صالحة للتعميم لمثل هذه الحالات. ولكن فى هذه المذكرات يقتصر العرض على حالة الانحدار البسيط.

أمثلة تطبيقية:

1- بمعلومة القيمة التالية للمتغيرين (X, Y) أوجد خط المربعات الدنيا فى حالة إذا كانت :

أ – X كمتغير مستقل **ب** –X كمتغير تابع.

X	1	3	4	6	8	9	11	14
Y	1	2	4	4	5	7	8	9

الحـل.

أ- باستخدام البيانات الخاصة بقيمة (Y , X) يمكن التوصل للكميات التالية:

$\bar{X} = 7$, $\Sigma X = 56$ $\Sigma Y = 40$, $\Sigma X^2 = 524$

$\bar{Y} = 5$ $\Sigma XY = 364$ $\Sigma Y^2 = 256$ $n=8$

وباستخدام الصورة التالية لكل من (A, b) يمكن التوصل للقيم التالية للحالة (أ).

$$\hat{B}_{y.x}=\frac{\sum X_iY_i-n\overline{X}\overline{Y}}{\sum X_i^2-n\overline{X}^2}=\frac{364-(8)-(5)(7)}{524-(8)(49)}$$

$$\hat{B}_{y.x}=0.636$$

$$\dot{A}_{y.x}=\overline{Y}-\hat{B}\overline{X}=5-(0.636)(7)=0.545$$

وبذلك يكون الخط المطلوب فى حالة (X) متغير مستقل ، (Y) متغير تابع:

$$Y= 0.545 + 0.636\ X$$

ب- وفى حالة كون (X) متغيرا تابعا ن (Y) متغيرا مستقلا لنفس الأزواج من الملاحظات تكون القيم المقدرة لكل من (A, B) كما يلي:

$$\hat{B}_{x.y}=\frac{\sum X_iY_i-n\overline{X}\overline{Y}}{\sum X_i^2-n\overline{X}^2}=\frac{364-(8)-(5)(7)}{524-(8)(25)}$$

$$\hat{B}_{x.y}=\frac{364-280}{256-200}=\frac{84}{56}=1.5$$

$$A_{x.y}=\overline{X}-\hat{B}\overline{Y}=7-(1.5)(5)=-0.5$$

أن تكون معادلة خط المربعات الدنيا فى حالة ما تكون (X) متغير تباع:

$$X= -0.50 + 1.5Y$$

مثال (2)

بين أن خطى المربعات الدنيا فى المثال السابق يتقاطعان فى النقطة (X,Y) ثم أوجد قيمة (Y) عندما تكون قيمة (X= 12) وقدر قيمة (X) عندما تكون قيمة (Y=3).

الحل : معادلتى الخطين هما

$$Y= (6/11) + (7/11)X$$

$$X=\left(\frac{1}{2}\right)+\left(\frac{3}{2}\right)Y$$

بحل المعادلتين انيا نجد أن $(x= 7, Y=5)$ وبهذا يثبت أن الخطين يتقاطعان فى النقطة (7.5).

* ويوضع $(X= 12)$ فى خط انحدار (Y على X) فان قيمة (Y) المطلوبة=

$$Y= 0.545 + 0.636\ (12) = 8.2$$

* ويوضع $(Y= 3)$ فى خط انحدار (Xعلى Y) فان قيمة (X) المطلوبة =

$$X = -0.50 + 1.50\ (3) = 4.0$$

مثال (3): إذا كانت

$$Y^{*}= Y + C \qquad , \quad X^{*}= X +D$$

حيث C, D ثوابت اثبت أن

$$\hat{B}_{x.y} = \frac{\sum\left(X_i - \overline{X}\right)-\left(Y_i - \overline{Y}\right)}{\sum\left(X_i - \overline{X}\right)^2} = \frac{\sum\left(X_i^* - \overline{X}^*\right)\left(Y_i^* - \overline{Y}^*\right)}{\sum\left(X_i^* - \overline{X}^*\right)^2}$$

$$\left(X_i^* - \overline{X}^*\right) = \left[\left(X_i + D\right) - \left(X_i + D\right)\right] = \left(X_i - \overline{X}_i\right)$$

$$\left(Y_i^* - \overline{Y}^*\right) = \left[\left(Y_i + C\right) - \left(Y_i + C\right)\right] = \left(Y_i - \overline{Y}_i\right)$$

وعلى هذا الأساس فان المقداران متساويان. وأهمية هذه النتيجة أنها تمكن من تبسيط الحسابات بطرح ثوابت اختيارية من المتغيرات (X, Y).

مثال (4)

بافتراض أن البيانات التالية تمثل الكميات المعروضة من الطماطم فى سوق الجملة بالقاهرة وأسعارها خلال الفترة (1990–1999).

أ– ارسم شكلا انتشاريا لهذه البيانات باعتبارها الكميات المعروضة من الطماطم متغيرا مستقلا، والأسعار متغيرا تابعا (Y).

ب– أوجد معادلة المربعات الدنيا (معادلة الخط المستقيم) لهذه البيانات.

ج– باستخدام المعادلة المتحصل عليها فى 0ب) أوجد القيم المتوقعة (Y) عند مستويات (x) المختلفة أوجد انحرافات هذه القيم عن القيم الأصلية المشاهدة لـ (Y) ومربعات هذه الانحرافات.

سعر الكيلو من الطماطم بالقرش (Y)	الكميات المعروضة من الطماطم بالطن (X)	السنة
18.0	73	1970
20.0	79	1971
17.8	80	1972
21.4	69	1973
21.6	66	1974
15.0	75	1975
14.4	78	1976
17.8	74	1977
19.6	74	1978
14.1	84	1979

الحـل

أ- الشكل الانتشارى للبيانات - راجع في ذلك الشكل الانتشاري رقم (3)

ب- من الجدول السابق يمكن أيجاد الكميات التالية

$\Sigma X = 752, \quad \bar{X} = (752/10) = 75.2$

$\Sigma Y = 179.7, \quad \bar{Y} = 17.97$

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) = -93.04$$

$$\sum (X_i - \bar{X})^2 = 253.6$$

$$\hat{B} = \frac{\sum (X_i - \bar{X})(Y_I - \bar{Y})}{\sum (X_I - \bar{X})^2} = \frac{-93.04}{253.6}$$

$$\hat{B} = 0.367$$

$$\hat{A} = \bar{Y} - \hat{B}\bar{X} = 17.97 - (0.367)(75.2) =$$

$$\hat{A} = 45.57$$

وبهذا فان معادلة الخط المستقيم لهذه البيانات

$$\hat{Y} = 45.57 - 0.367 X$$

$$\sum_{i-1} (Y_i - \hat{Y})^2 = 34.26$$

X	Ŷ	Y	Y- Ŷ	(Y- Ŷ)^2
73	18.0	18.8	-0.8	0.64
79	20.0	16.6	3.4	11.54
80	17.8	16.2	1.6	2.56
69	21.4	20.2	1.2	1.44
66	21.6	21.4	0.2	0.04
75	15.0	18.1	-3.1	9.61
78	14.4	16.9	-2.3	6.25
74	17.8	18.4	-0.6	0.36
74	19.6	18.4	1.2	1.44
84	14.1	17.7	-0.6	0.36
	Total		0.0	34.26

حساب تباين الخطأ Caiculationg S^2 an Estimator of σ^2

فى النموذج الاحتمالي السابق الإشارة إليه (Y= A+ BX+E) حيث (E) تمثل خطأ عشوائي له متوسط = صفر = 0 = (E) E وتباين = σ^2 ، ويلاحظ أن (Y) تتعرض لخطأ عشوائي (E) ويظل هذا الخطأ فى حسابات (A,B) وبالتالي يلحق بهذه التقديرات أخطاء، وكذلك عند استعمال الانحدار (خط المربعات الدنيا) Ŷ = A+B X لأغراض التنبؤ بقيمة (Y) فى المستقبل فان الخطأ العشوائي يؤثر على هذه التنبؤات، وبهذا يتضح أن تقلبات الخطأ العشوائي (E) أو تباين الملاحظات حول خط الانحدار الحقيقي مقاسه فى صورة σ^2 تلعب دورا هاما عند استخدام خط المربعات الدنيا لأغراض التقدير أو التنبؤ.

ولتحديد قيمة تباين الخطأ العشوائي (E) يتم استخدام مجموع مربعات الخطأ أو مجموع مربعات الانحرافات حول خط التنبؤ (SSE) وبهذا يمكن تحديد تقدير جيد غير متحيز لهذا التباين ويقوم على درجات حرية (n-2) ويتمثل فى المقدار:

$$\sigma^2_{\hat{Y}.X} = S^2_{\hat{Y}.X} = (1/n-2)SSE = (1/n-2)\sum\left(Y_i - \hat{Y}_I\right)^2$$

ولكن هناك صيغا أخرى اكثر سهولة من الناحية الحسابية ومنها:

$$\sigma^2_{y.x}S^2_{y.x} = \left(\frac{1}{n-2}\right)\left(\sum(Y_i - \bar{Y})\right)\frac{\left(\sum\left(X_i - \bar{X}\right)\left(Y_I - \bar{Y}\right)\right)}{\sum\left(X_i - \bar{X}\right)^2}$$

مثال (5)

يمكن باستخدام المثال السابق رقم (5–ج) التوصل لتقدير لهذه الكمية حيث يبين حل هذا المثال أن قيمة المقدار.

$\left[\left(\sum\left(Y_i-\bar{Y}\right)^2=34.06\right)\right]$ وكانت n =10 إذن

$$\sigma^2_{y.x}=S^2_{y.x}=\frac{\sum(Y_i-\bar{Y})^2}{n-2}$$

$$S^2_{Y.X}=\frac{43.26}{8}=4.283$$

ملاحظات:

* يسمى الجذر التربيعي للكمية (S^2) بالخطأ المعياري للتقدير أو الانحراف المعياري عن خط الانحدار S_{YX}

* إذا كانت الانحرافات عن خط الانحدار تتوزع توزيعا طبيعيا فان B تتوزع أيضا توزيعا طبيعيا بمتوسط = E (B) = B

وتباين =

$$E\left(B_i-\bar{B}\right)^2=\sigma^2 B_i=\frac{\sigma^2\left(Y-X\right)}{\sum\left(X_i-\bar{X}\right)^2}$$

مثال (6)

استمرارا مع بيانات مثال (4–ج) ونتائج المثال السابق رقم (5) يمكن تحديد قيمة تباين التقدير $\sigma^2(B)$ بالتطبيق فى القانون المذكور فى الملاحظة الأخيرة كما يلي:

$$\sigma^2 B_i=\frac{\sigma^2\left(Y-X\right)}{\sum\left(X_i-\bar{X}\right)^2}=\frac{S^2}{\sum\left(X_i-\bar{X}\right)^2}$$

التباين الكلى والتباين المشروح والتباين غير المشروح

Total sum of squares of Deviations and its components.

يقصد بالتباين الكلى (TSS) المجموع الكلى لمربعات الانحرافات $TSS=\sum\left(Y_i-\bar{Y}\right)^2$ وهذا المقدار يمكن تقسيمه إلي جزأين. أحدهما هو (SSE) أو مجموعة مربعات انحرافات قيم (Y) حول خط الانحدار المقدر. والجزء الثاني يمثل النقص فى مجموع مربعات الانحرافات الكلى الناجم عن إدخال التغير (X) فى

هذه العلاقة أو ما يسمى بـ (SSR) أو مجموع مربعات الانحرافات الناجمة عن الانحدار، وبالنسبة لكل ملاحظة فان.

$$\left(Y_i-\overline{Y}\right)=\left(Y_i-\hat{Y}_I\right)+\left(\hat{Y}_i-\overline{Y}_I\right)$$

<u>وهذا التقسيم يوضحه الشكل (4) ويأخذ مجموع مربعات الانحرافات فوق كل الملاحظات لهذا التقسيم المشار إليه يتضح أن</u>

$$\sum_{i-1}^{n}=(Y_i-\overline{Y})^2=\sum_{i-1}^{n}(Y_i-\hat{Y})^2+\sum_{i-1}^{n}(Y_i-\overline{Y}_i)^2$$

$$TSS=SSR+SSE$$

$$SSR=\sum_{i=1}^{n}\left(\hat{Y}_i-\overline{Y}\right)^2$$

مجموع مربعات الانحرافات الناجمة عن الانحدار = SSR أو كمية الانحرافات الكلية التى يشرحها المتغير (X)

$$SSE=\sum_{I=1}^{n}\left(Y_i-\hat{Y}_i\right)^2$$

مجموع مربعات الخطأ (SSE) أو كمية الانحرافات الكلية التى لا يشرحها المتغير (X) وتسلك سلوكا عشوائيا لا يمكن التنبؤ به.

<u>وأهمية هذا التقسيم ترجع للسببين التاليين:</u>

1) تقليل الجهد الحسابي اللازم فى تحليل الانحدار، حيث بحساب اثنين يمكن معرفة الثالث.

2) يساعد هذا التقسيم على شرح إسهام المتغير (X) فى توفير معلومات لأغراض التنبؤ بقيم المتغير (Y).

وعلى أساس من المناقشة السابقة يمكن تعريف معامل التحديد (r^2) أو مربع معامل الارتباط كما يلي:

معامل التحديد = $r^2=\dfrac{TSS-SSE}{TSS}=\dfrac{SSR}{TSS}$

وفيما يلي مناقشة اكثر تفصيلا لمعامل التحديد ومعامل الارتباط.

معامل الارتباط A Coeficient of Correlation

معامل الارتباط الخطى يمثل مؤشرا لمدى قوة العلاقة الخطية بين المتغيرين (X,Y) وهذا المؤشر يتميز بعدم اعتماده على وحدات القياس ولا على اختيار نقط الأصل. ويسمى هذا المؤشر بمعامل ارتباط (X,Y) ويرمز لهذه الكمية بالرمز (r) ويستخدم لحسابها اكثر من صورة من أهمها:

$$r=\frac{\sum\left(X_i-\bar{X}\right)-\left(Y_i-\bar{Y}\right)}{\sqrt{\sum\left(X_i-\bar{X}\right)^2\sum\left(Y_i-\bar{Y}\right)^2}}=\frac{\sum X_I y_i-n\overline{X}\overline{Y}}{\sqrt{(\sum X_i^{\ 2}-n(\bar{X})^2)(\sum\left(Y_i^2-n(\bar{Y}\right)^2)}}$$

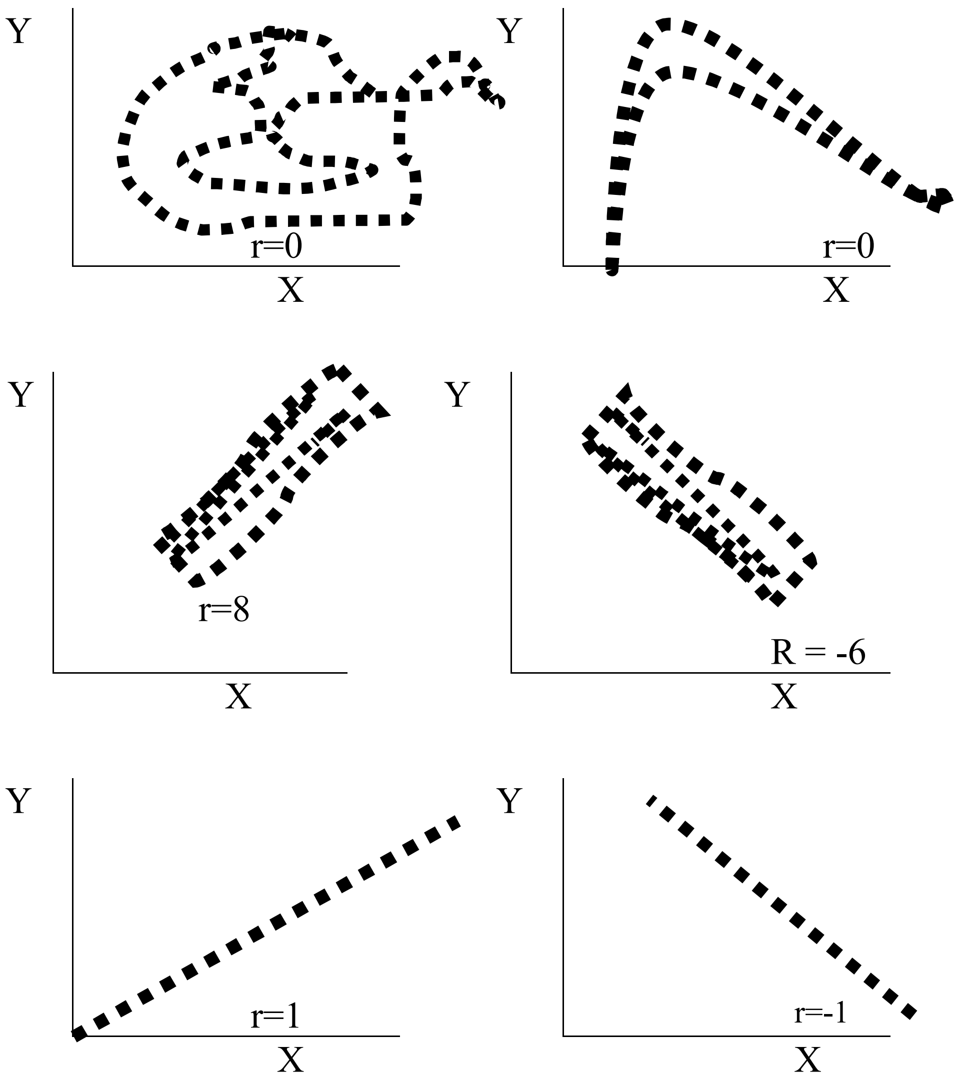
وملاحظـة الصـورة المـذكورة أعـلاء لمعامـل الارتبـاط توضـح أن المقـام فـى كلتـا الحـالتين لا بـد وان يكـون موجبا والبسط يتطـابق مـع بسط الصـورة المستخدمة فـى ايجـاد(B) ، ومـن ثـم يتضـح أن إشـارة معامـل الارتبـاط تتطـابق مـع إشـارة (B) ويساوى معامل الارتباط (صفر) عندما تكون (B=0)

ومعامل التحديد هو مربع معامل الارتباط وتتراوح قيمته بين صفر وواحد صحيح $(0 \leq r^2 \leq 1)$

ويوضح الشكل رقم (4) بعض الأشكال الانتشارية ومعامل الارتباط فى كل حالة. ويوضح ذلك الشكل انه عندما تكون (r=1) أو (r-i) فان كل النقاط تقع على الخط المستقيم. وعندما تكون (r=0) تكون النقاط مبعثرة ولا تعطى أي دليل على علاقة خطية، وكل قيم (r) الأخرى تبين درجة قوة العلاقة الخطية بين المتغيرين.

ومن الضروري ملاحظة أن قيمة (r=0) تعنى عدم وجود ارتباط خطى بين (X,Y)، ولكنها لا تعنى عدم وجود ارتباط على الإطلاق، حيث وجود علاقة غير خطية بين (X,Y) قد يسبب وجود قيمة صفرية لمعامل الارتباط الخطى بينهما (لاحظ الشكل رقم 4).

الشكل رقم(4) بعض الأشكال الانتشارية للبيانات ومعاملات الارتباط فى كل حالة



ملاحظات على معامل الارتباط (r) ومعامل التحديد r^2

* سبق الإشارة إلي أن معامل انحدار Y على X فى حالة اعتبار Y متغير تابعا و X متغير مستقلا:

$$\hat{B}_{y.x} = \frac{\sum\left(X_i - \overline{X}\right) - \left(Y_i - \overline{Y}\right)}{\sum\left(Y_i - \overline{Y}\right)^2}$$

وفى الحالة العكسية حتى تكون (X) متغيرا تابعا، (Y) مستقلا

$$\hat{B}_{y.x}=\frac{\sum\left(X_i-\bar{X}\right)-\left(Y_i-\bar{Y}\right)}{\sum\left(X_i-\bar{X}\right)^2}$$

وبمقارنة هاتين الصورتين والصورة المذكورة لمعامل التحديد (r2) يلاحظ أن هناك العلاقة التالية:

$$r^2=\left(B_{y.x}\right)*\left(B_{x.y}\right)$$

$$r^2=\frac{\sum\left(X_i-\bar{X}\right)-\left(Y_i-\bar{Y}\right)}{\sum\left(X_i-\bar{X}\right)^2}*\frac{\sum\left(X_i-\bar{X}\right)-\left(Y_i-\bar{Y}\right)}{\sum\left(Y_i-\bar{Y}\right)^2}$$

وهـذا يبـين العلاقـة بـين معـاملات الانحـدار ومعامـل الارتبـاط ومعامـل التحديد.

* يلاحظ أن الاسم الكامل لهذا المعامل هو معامل الارتباط الخطى، وان تعريف الارتباط لا يشير إلي أن العلاقة بين المتغيرين هى علاقة سبب ونتيجة. ومعامل الارتباط لا يوفر إجابة على السؤال الخاص بأي المتغيرين التابع وايهما المستقل.
* انخفاض قيمة معامل الارتباط لا تعنى فى كل الحالات عدم وجود علاقة بين المتغيرات محل الاهتمام فقد تكون هناك علاقة غير خطية ولهذا يكون معامل الارتباط الخطى منخفضا.

مثال (7)

أوجد معامل الارتباط (r) ومعامل التحديد (r2) لحجم المبيعات والأنفاق على الدعاية فى المثال السابق رقم (1).

الحـل

$$r=\frac{\sum X_iY_i-n\bar{X}\bar{Y}}{\sqrt{\sum X_i^{\,2}-n\bar{X}^2\sum Y_i^2n\bar{Y}^2}}$$

$$r=\frac{924.8-(10)(0.94)(95.9)}{\sqrt{(9.28-(10)(0.94)^2)(9356-(10)(95.9)^2)}}$$

r= 0.88 $\qquad$ r^2= 0.77

ويوضح ذلك أن 77% من التغيرات فى المتغير التابع (Y) حجم المبيعات ترجع أو تعزى إلي المتغير المستقل (X) الأنفاق على الدعاية.

مثال (8)

أوجد قيمة معامل التحديد (r^2) للمتغيرين (X, Y) فى المثال رقم (2) بمعلومية معامل انحدار (Y) على (X) ومعامل انحدار X على Y .

الحل

من حل مثال رقم (2) يتضح أن

$B_{Y\text{-}X}= 0.636$ M $B_{XY}= 1.5$

$r^2= B_{Y\text{-}X} , B_{XY}$

$r^2= (0.636). (1.5)$

$r^2= 0.954$

تحويل العلاقات غير الخطية إلي علاقات خطية:

سبق الإشارة إلي أن معامل الارتباط والانحدار التى تم التوصل إليها تفترض وجود علاقة خطية بين متغيرين. وهذه العلاقة الخطية هى اكثر العلاقات بساطه وانتشارا، ولكن الواقع قد يحتوى علاقات غير خطية أيضا وفى مثل هذه الحالات (المنحنيات) يكون من الممكن (غالبا) التحويل لعلاقة خطية وبالتالي يمكن استخدام نفس الصور والصيغ السابق ذكرها فى المعاملات الانحدار الارتباط وفى المناقشة التالية يتم استعراض بعض أنواع العلاقات غير الخطية الشائعة الانتشار بالذات فى البيانات الاقتصادية والحيوية مع شرح كيفية تحويل هذه العلاقات إلي الصورة الخطية.

أ- التحويل باستخدام الصورة اللوغاريتمية:

إذا تم تحويل كل من المتغيرين العشوائيين (Y,X) إلي لوغاريتمات (Log Y, Log X) ، ثم تم رسم شكل انتشارى لهذه البيانات (بعد التحويل) وتبين من هذا الشكل الانتشارى أن النقاط تقترب بشكل واضح من توفيق خط مستقيم اذن يمكن القول بان المنحنى الذي يصف هذه العلاقة من الصورة

$Y= AX^{B}$

بأخذ الوغاريتم تصبح الصورة كالتالي : $\text{Log } Y= \text{Log } A+ B \text{ Log}X$

وكما ذكرنا من قبل يتم تحويل المتغيرات الأصلية (Y,X) إلي متغيرات جديدة Log (Y) ، Log (X) ثم يستمر العمل باستخدام نفس الصيغ الإحصائية التى

استخدمت فى حالات الانحدار والارتباط الخطى لإيجاد معادلة الانحدار للخط المستقيم. $Y= A + B X$

ب–التحويل باستخدام الصور النصف لوغاريتمية

إذا تم تحويل (Y) إلي لوغاريتمات وتم رسم الشكل الانتشارى{ Log (Y), X)) وتبين أن البيانات تكون خطأ مستقيما اذن المنحنى المناسب يكون من الصورة $\mathbf{Y= A\ B^{X}}$

ويأخذ لوغاريتم طرفي المعادلة الأخيرة:

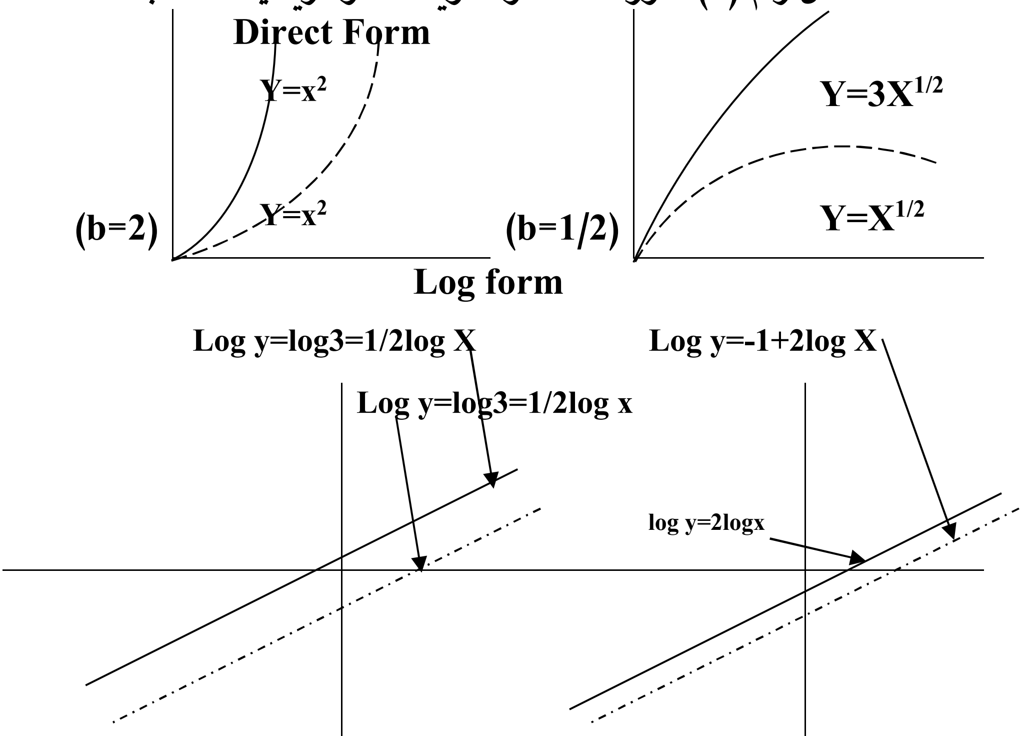
$$Log\ (Y) = Log\ (A) + Log\ (B)\ X$$

والحالات التى تناسب هذا النوع من المنحنيات عادة هى التى يكون فيها المتغير (Y) يزيد أو ينقص بمعدلات ثابته. والتحليل يتم تحويل (Y) إلي Y,=Log(Y) ثم يستمر تحليل الانحدار كالمعتاد للمعادلة التالية.

$$Y = A + B X$$

وهذا تسمى الصورة نصف اللوغاريتمية، راجع الشكل رقم (5)

الشكل رقم (5) صور مختلفة والتحويلات اللوغاريتمية المناسبة.



تذكــر

** **الانحـدار البسـيط** :هو دراسـة العلاقـة بـين متغيـرين فقـط أحداهمـا تـابع والأخـر مستقل .

** الفكرة الأساسية فى موضوع الانحدار والارتباط هى وجود متغيرين (أو اكثر) تعتمد قيمة أحدهما على الآخر. ويشار إلي أحدهما بالمتغير التابع (Y)، والثاني بالمتغير المستقل (X).

** هدف طريقة المربعات الدنيا هو التوصل للنهاية الدنيا لدالة مجموع مربعات الانحرافات عن الخط المستقيم .

** تعتمد طريقة المربعات الدنيا على المبدأ التالي:

"ان افضل خط مستقيم أو الخط الذي يحقق افضل توفيق للبيانات الممثلة للعلاقة بين المتغيرين (Y, X) هو الخط الذي يتحقق به تدنية مجموع مربعات انحرافات القيم المشاهدة لـ (Y_i) عن القيم المتوقعة (Y_i) لاقل ما يمكن":

** يمكن كتابه نموذج الانحدار البسيط كالأتي:

$$\mathbf{Y_i = b_o + b_i X_i}$$

وتحسب ثوابت معادلة الانحدار البسيط باستخدام الصيغ التالية

$$\hat{A} = \overline{Y} - \hat{B}\,\overline{X}$$

$$\hat{B} = \frac{\sum X_i Y_i - n\,\overline{X}\,\overline{Y}}{\sum X_i^2 - n\,\overline{X}^2}$$

** من خصائص معامل الارتباط عدم اعتماده على القيم نفسها بل على تباعدها عن بعضها، حيث لا تتغير قيمة معامل الارتباط بالعمليات الحسابية الأربع الجمع والطرح والقسمة والضرب مع عدد ثابت بالنسبة لقيم x , y .

** يمكن ايجاد معامل الارتباط (r) بمعلومية معامل الانحدار (B) من الصيغة الرياضية التالية :

$$R = \hat{B} * \frac{\sqrt{\sum (X_i - \overline{X})^2}}{\sqrt{\sum (Y - \overline{Y})^2}}$$

اي ان :

$$R = \hat{B} * \frac{\sigma \ Y}{\sigma \ X}$$

وكذلك فان معامل التحديد

$$r^2 = \hat{B}_{x,y} * \hat{B}_{y,x}$$

**يجب ملاحظة ان قيمة معامل التحديد تنحصر بين قيمتين هما (صفر) عند حده الادني ، و(واحد صحيح) عند حده الاقصي.

اما معامل الارتباط فقيمته تنحصر بين الصفر ، ± 1، ويكون الارتباط تام عندما تصل قيمته الي الواحد الصحيح سواء كانت القيمة موجبة أو سالبة، فيسمي ارتباط تام سالب او ارتباط تام موجب ، وكلما اقتربت قيمة معامل الارتباط كان ارتباطا ضعيفا في حين كلما اقتربت من الواحد الصحيح بغض النظر كان ارتباط قوياً.

تمارين

تمرين (1)

الجدول التالى يوضح درجات عشرة طلاب فى امتحان مادتى الاقتصاد والنبات.

أ- ارسم خط الشكل الانتشارى لهذه البيانات

ب- أوجد المربعات الدنيا الذي يوفق هذه البيانات باعتبار X متغير مستقل

ج- أوجد خط المربعات الدنيا الذي يوافق هذه البيانات باعتبار Y متغير مستقل.

د- إذا حصل طالب على 79 درجة فى الاقتصاد ما هى الدرجة المتوقع أن يحصل عليها فى النبات.

هـ- إذا حصل طالب على 95 درجة فى النبات ما هى الدرجة المتوقع أن يحصل عليها فى الاقتصاد.

الاقتصاد (Y)	75	80	93	65	87	71	98	68	84	77
النبات (X)	82	78	86	72	91	80	95	72	89	74

تمرين (2)

يوضح الجدول التالى عدد البكتريا (Y) الموجودة فى وحدة حجم معين فى مزرعة بكتريا بعدد (X) ساعة.

عدد الساعات (X)	0	1	2	3	4	5	6
عدد البكتريا فى وحدة حجم (Y)	32	47	65	82	132	190	275

أ- ارسم هذه البيانات باستخدام ورق رسم بيانى نصف لوغاريتمى

ب- أوجد منحنى المربعات الدنيا على الصورة $Y= AB^X$

ج- أوجد قيمة Y عند (X=7)

تمرين (3)

الجدول التالى يوضح درجات عشرة طلاب فى امتحان مادتى الاقتصاد والنبات.

أ- ارسم خط الشكل الانتشارى لهذه البيانات

ب- أوجد المربعات الدنيا الذي يوفق هذه البيانات باعتبار X متغير مستقل

ج- أوجد خط المربعات الدنيا الذي يوافق هذه البيانات باعتبار Y متغير مستقل.

د- إذا حصل طالب على 79 درجة فى الاقتصاد ما هى الدرجة المتوقع أن يحصل عليها فى النبات.

هـ- إذا حصل طالب على 95 درجة فى النبات ما هى الدرجة المتوقع أن يحصل عليها فى الاقتصاد.

الاقتصاد (Y)	75	80	93	65	87	71	98	68	84	77
النبات (X)	82	78	86	72	91	80	95	72	89	74

تمرين (4)

أوجد معادلة انحدار المربعات الصغرى للمتغير Y على X1 , X2

ب- اختبر معنوية تقديرات معالم الانحدار .

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y	6	8	8	7	7	12	9	8	9	10	10	11	9	10	11
X1	9	10	8	7	10	4	5	5	6	8	7	4	9	5	8
X2	8	13	11	10	12	16	10	10	12	14	12	16	14	10	12

تمرين (5)

بيانات الجدول التالى توضح الكمية المطلوبة من سلعة ما y ، وسعرها x1 ، ودخل المستهلك x2 من عام 1980 إلى 2000

السنة	Y	X1	X2
1990	60	7	800
1991	70	6	900
1992	65	6	1000
1993	65	8	1100
1994	75	5	1200
1995	75	5	1300
1996	80	5	1400
1997	100	3	1500
1998	90	4	1600
1999	95	3	1700
2000	85	4	1800

والمطلوب :

أ – تقدير معادلة الانحدار لهذه المشاهدات .

ب – اختبر عند مستوى معنوى 5% المعنوية الإحصائية لمعامل الانحدار .

جـ – أوجد معامل التحديد ومعامل التحديد المعدل .

د – اختبر المعنوية الكلية للانحدار

هـ – أوجد معاملات الارتباط الجزئى أوجد أى متغير مستقل يساهم أكثر فى مقدرة النموذج لتفسيرية

تمرين (6)

ارسم شكل انتشاري لبيانات الجدول التالي وحدد بالنظر ما إذا كانت توجد علاقة خطية تقريبية بين إجمالي الإنفاق الاستهلاكي (Y) ، وإجمالي الدخل المتاح بالدولار الأمريكي (X) وذلك خلال الفترة 1990 - 1999 .

السنة	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Yi	40	44	46	48	52	58	60	68	74	80
Xi	6	10	12	14	16	18	22	24	26	32

المطلوب :–

أ – قدر معادلة الانحدار Y على X .

ب – اختبر معنوية تقديرات المعالم .

ج – احسب قيمة معامل الارتباط البسيط ومعامل التحديد .

تمرين (7)

– الجدول التالي يوضح الأسعار المزرعية للطن من فول الصويا بالجنية خلال الفترة (1980 – 1992) .

السنة	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
السعر	260	285	285	375	425	500	800	800	850	810

تمرين (8)

يوضح الجدول التالي المساحة المزروعة من البرتقال بالفدان Yiوالناتجة عن كميات مختلفة من السماد بالكجم Xi في إحدى المزارع خلال 10 سنوات من 1991 إلي 2000 .

والمطلوب : تقدير العلاقة بين المتغيرين(Yi , Xi) بطريقة المربعات الصغرى.

السنة	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
N	1	2	3	4	5	6	7	8	9	10
Yi	40	44	46	48	52	58	60	68	74	80
Xi	6	10	12	14	16	18	22	24	26	32

الباب الخامس

العينات وتصميم التجارب

البيانات الإحصائية هي الأساس للتخطيط الاقتصادي والاجتماعي ولكل البرامج الإنمائية ولمتخذي القرار. وبدخول عصر العولمة ومع الوضع الراهن للدول النامية أصبحت هناك ضرورة مُلحة ومتزايدة للإحصاءات بوجه عام وللبيانات الاقتصادية والاجتماعية بوجه خاص. واستجابة لهذه الحاجة تسعي، كثيراً من دول العالم النامي إلى النهوض بالعمل الإحصائي إلى المستوى اللازم للوفاء باحتياجات المسئولين عن التخطيط للتنمية الاقتصادية والاجتماعية. كما تبذل جُهوداً كبرى في تدريب الكوادر الوطنية القادرة على القيام بإجراء التعدادات والمسوحات وغيرها من نشاطات جمع البيانات وإجراء التحليل بشكل فعّال

ويمكن تقسيم الدراسات والبحوث من حيث المجال أي من حيث درجة الشمول لمفردات المجتمع الأصلي إلى بحوث شاملة وبحوث بطريقة العينات .فالبحث الشامل هو الذي ندرس فيه حاله جميع أفراد المجتمع موضوع البحث بهذه الطريقة إذا كان الغرض منه هو الحصر وذلك مثل تعداد السكان التعداد الزراعي..الخ. وهذا يتطلب تكلفة كبيرة من الوقت والمال والجهد. إن البحث بطريقة العينة فهو الذي نبحث فيه حاله جزء معين (أو نسبة معينة) من أفراد المجتمع الأصلي ثم نقوم بعد ذلك بتعميم نتائج الدراسة على المجتمع كله بتكلفة أقل كثيراً من البحث الشامل .

<u>أساليب جمع البيانات:</u>

<u>أولاً : عمليات المسح (الحصر الشامل)</u>:هو عد جميع الأشخاص الموجودين على قيد الحياة داخل الحدود السياسية للجمهورية ليلة التعداد سواء كانوا مصريين أو أجانب.

و تعد عمليات الاستطلاع والمسح من الأدوات القيمة التي تستخدم كوسائل لإضافة البنية وتسهيل القياسات في الأبحاث الاقتصادية والاجتماعية. إلا أن عمليات الاستطلاع والمسح تفقد قيمتها عند مرحلة محددة في معظم استقصاءات الأبحاث.

وعمليات المسح هي طريقة لدراسة الخصائص الاجتماعية، والعلاقات، والسلوكيات. وتوفر عملية الاستطلاع والمسح التجانس، حيث أن أفراد المجموعة يجيبون على نفس السؤال، وتعتبر هذه الطريقة خالية من الانحراف والتحيز الذي يسببه التفاعل مع الباحث، كما أنها تسمح بانعكاس الخصوصية والوقت على السؤال المطروح. إلا أن النتائج التي نحصل عليها من عمليات المسح الاستبياني قد تؤدي إلى إجابات سطحية أحياناً.

ويلخص الجدول التالي مزايا وعيوب جمع البيانات عن طريق الاستطلاع والمسح:

المزايا	العيوب
1.تمكن من جمع فعال للبيانات من عدد كبير من الأفراد 2.تسمح بإجراء مقارنات دقيقة بين إجابات الأفراد المستجوبين المختلفة.	1.احياناً تكون المادة التي تم جمعها سطحية، فالاستبيانات معيارية لحد كبير، وهذا يؤدي لتجاوز العديد من الفروقات الهامة 2.قد تكون الإجابات هي ما يتظاهر الناس بالاعتقاد به، أكثر منها ما يعتقدون به حقاً.

ويمكن لعمليات الاستطلاع والمسح أيضاً أن تقلل التحيز الذي قد يشكل صعوبة في المقابلات الشخصية والملاحظات العينية. فقد تؤدي المقابلات إلى التحيز أو الانحراف بسبب عدة عوامل، كتغيير الأسئلة أو إعادة ترتيبها، بالإضافة إلى العديد من العوامل الموضوعية كالتوقيت، ومكان المقابلة، وظروف المقابلة، التي تؤثر جميعها على طريقة استجابة الفرد مع الشخص الذي يجري المقابلة.

وقد لا تخلو الاستبيانات من مثل هذا التحيز. حيث تؤدي عمليات المسح الاستبياني إلى تحيز خاص بها لعدد من الأسباب: كالأسئلة السيئة، والعينات السيئة، والاستجابات السيئة، أو التحليل غير المناسب للبيانات. لذا، يجب أن يصمم الاستبيان ويدار بعناية وحذر إذا ما أريد له أن يكون فعالاً في جمع البيانات المطلوبة.

وهناك العديد من المخاطر التي تواجه إدارة عمليات الاستطلاع والمسح، منها مشاكل تصميم، وإدارة عمليات الاستبيان، ومتوسطات الاستجابة المنخفضة للاستبيانات البريدية. لذا، ينبغي التحقق من الاستبيانات دائماً قبل استخدامها، ويجب أن يتم تدوين الملاحظات بعناية، وخاصة من الردود على هذه الاستبيانات.

وتعد الاستبيانات طريقة لتحصيل وتجميع البيانات الرئيسية في عمليات عمليات الاستطلاع والمسح، وغالباً ما تكون الاستبيانات ورقية أو توزع بريدياً.

على أن هناك عدداً من البدائل. فمثلا، يمكن تجميع البيانات بوساطة المقابلات المباشرة وجهاً –لوجه، حيث يتم سؤال المشاركين، ويقوم الشخص الذي يجري المقابلة بتسجيل إجاباتهم. وبالمثل، يمكن ان يتم سؤال الأسئلة هاتفياً.

وقد شهدت السنوات الأخيرة ارتفاعاً في متوسط استخدام الطرق الإلكترونية، كالبريد الإلكتروني والانترنت، حيث يتم إجراء الاستبيانات مباشرة على الويب. تعرف مجموعة الطرق الإلكترونية هذه "بطرق جمع المعلومات إحصائياً بمساعدة الحاسوب"

يوفر استخدام التكنولوجيا عدداً من الفوائد للباحث وأحد أهم هذه الفوائد التخفيض الهام التي توفره التكنولوجيا في كمية العمل، والتكلفة المالية المتعلقة بتحليل البيانات. وأحد الفوائد غير المتوقعة للإحصاءات التي يتم إجراؤها مباشرة على الويب هي أن هذا النوع من الإحصاءات يتمتع بقابلية أعلى لإثارة عدد أكبر من الردود الأكثر تفصيلاً، وخاصة لتلك الأسئلة ذات النهايات المفتوحة، ويعد هذا عاملاً هاماً يساهم بشكل كبير في أغناء البيانات.

غير أن استخدام التكنولوجيا يحد من أنماط الأشخاص الذين يمكن الوصول إليهم، ويحد بذلك من تمثيلية العينة، فلا يمكننا معرفة المعدات أو البرمجيات أو الأجهزة الحاسوبية التي سيستخدمها الأفراد الذين يقومون بالاستجابة لنا. لذلك، يجب بذل المزيد من الانتباه في تصميم واختبار الاستبيانات. يتوجب أيضاً وضع تعليمات دقيقة الأفراد الذين يقومون بالاستجابة بما يخص كيفية إتمام الاستبيان. إضافة، قد يكون متوسط الاستجابة للاستبيانات على الانترنيت منخفضاً جداً، وتشير بعض الأدلة إلى أن هذه المتوسطات تنخفض مع ازدياد كمية البريد الإلكتروني وارتفاع نسبة استخدام الإحصاءات المباشرة على الويب.

ثانياً :المعاينة:

تستخدم كلمة العينة كثيرا في حياتنا اليومية ، اذ عندما يمرض الشخص يطلب الطبيب فحص عينه من دمه أي بجزء منه . كذلك عندما نريد شراء سلعة معينة كالحبوب (رز ، قمح) نختار جزاء من هذه السلعة للتأكد من جودتها ، حتى نتخذ قرار بشرائها او عدم شرائها . ان عملية الاختيار قد تكون جيدة ومناسبة بحيث تمكننا من الوصول الى القرار السليم ، وقد تكون خاطئة تعطي نتائج مضلله .

وأسلوب العينة هو البديل الوحيد فى كثير من الأحيان للحصر الشامل الذي يستحيل إجراءه. وليس القصد من أخذ العينة هو دراستها لذاتها لنخرج بنتائج تنطبق على العينة فقط ، ولكن الغاية الأولى والأساسية من دراسة عينة ما هو إلا تعميم نتائجها على المجتمع الأصلي الذي اشتقت منه . فإذا تمكنا من الحصول على مقاييس إحصائية معينة تصف العينة ، فإن هذه المقاييس نفسها يجب أن تكون قابلة للتعميم على الإطار العام وذلك فى حدود معينة سنتكلم عنها فيما بعد.

ولا يتبادر أيضاً إلى الذهن أن المقاييس الإحصائية التي نحصل عليها من فحص عينة ما أقل فى الدقة من نتائج الحصر الشامل ، إذا أثبتت التجارب عكس ذلك ، وثبت أن خطا المعاينة قد يكون أقل من خطأ المعاينة قد يكون أقل من خطأ الحصر الشامل ، بالإضافة إلى خطأ المعاينة يمكن التحكم فيه كما يمكن التنبؤ به . ويمكن للباحث إيجاد كافة الوسائل والضمانات التى تقلل من أخطاء المعاينة ، وعليه فيمكن القول أن عملية المعاينة فى البحوث تأتى لنا بأصدق النتائج فى أقصر وقت ممكن وبأقل مجهود ممكن وبأدنى التكاليف . وكل هذه مبررات كافية لاستخدام أسلوب المعاينة. وبالإضافة إلى كل ما سبق فأن المشرفين على البحث يمكنهم تدريب المساعدين تدريباً كافياً ومراقبتهم مراقبة دقيقة وبذلك تكون عملية البحث تحت إشراف دقيق.

وليس معنى كل ما ذكرناه أن نضع أكاليل على رأس عملية المعاينة ونحرم طريقة الحصر الشامل من المزايا، فالطريقة الأخيرة لها استعمالاتها ولها مبرراتها فى أحوال كثيرة . فمثلاً إذا أراد باحث ما أن يقوم بدراسة عن تأثير مهنة معينة على الحالة الاجتماعية والصحية للعاملين فى تلك المهنة ، وكان عدد العمال الذين يشتغلون بهذه المهنة لا يزيد عن طاقة البحث وليكن 1000 عاملاً مثلاً، فإن الباحث يستطيع دراسة كل هذا المجتمع بالحصر الشامل إذا أراد ، وله ان يستخدم أسلوب المعاينة إذا أراد أيضاً.

وتستخدم عملية المعاينة فى كثير من البحوث فى شتى المجالات ، خاصة عندما تكون الظاهرة المراد بحثها متوافرة فى الإطار الأصلي، أما إذا كانت هذه الظاهرة نادرة الحدوث (مرض نادر جداً أو شذوذ نادر جداً) فأن عملية المعاينة تكون غير ذات موضوع فى هذه الحالة ، إذ أنه يتحتم فحص عينة كبيرة جداً قد تصل إلى حد فحص المجموع الكلى للإطار العام للحصول على عدد يمكن دراسته من هذه الظاهرة.

ومعظم العينات التى تستخدم تفترض تجانس مفردات الإطار أو مفردات الإطار أو مفردات المجتمع الأصلي، الإ أنه يمكن استخدام أسلوب المعاينة بطريقة محورة فى المجتمعات غير المتجانسة وذلك بتقسيمها الى
مجتمعات صغيرة متجانسة، ثم أخذ العينات من هذه المجتمعات الصغيرة. وسيأتى الكلام عن هذا الموضوع عند دراسة أنواع العينات.

وتعرف العينة بأنها " جزء من المجتمع يتم اختياره لتمثيل المجتمع " اما المعاينة فتعرف " عملية اختيار جزء من المجتمع الإحصائي للاستدلال على خواص المجتمع باكلمه عن طريق تعميم نتائج العينة ".
نفرض اننا نريد دراسة مستوى الرضا الوظيفي لموظفي احد المؤسسات ، ونظرا لضخامة عدد موظفي هذه فقد تقرر اختيار عدد من الموظفين يمثلون المجتمع . ان الموظفين الذين تم اختيارهم هم العينة . اذ يشكلون جزءاً من المجتمع يتضمن خصائصه . اما عملية اختيار العينة وتعميم النتائج للاستدلال على خصائص المجتمع فتسمى " معاينة " .
وتبرز الحاجة إلى جمع البيانات عند استقصاء مشكلة ما. والسؤال الأول الذي ينبغي طرحه هو: ما كمية البيانات التي ينبغي جمعها؟ فعند القيام بالإحصاء السكاني الذي يستقصي تعداد السكان الذين يعيشون في بلد ما، والعوامل التي تؤثر على مستوى معيشتهم، من الضروري جمع بيانات حول كل فرد في ذلك البلد. ويعدّ هذا عملاً ضخماً. والمعاينة هي أسلوب لتخفيض الجهد المبذول في عملية جمع البيانات.
فمثلاً، عند إجراء دراسة لأنماط تسوق الأفراد الذين يزورون أحد مواقع التجارة الإلكترونية، فقد يكون من غير الملائم طرح سؤال على كل منهم إذا كان عدد هؤلاء الزوار كبيراً جداً، غير أننا قد نحصل على نتائج جيدة بسؤال واحد من كل عشرة أفراد فقط. ويدعى هذا "بالمسح عن طريق العينة". وقد يكون من الضروري اتخاذ بعض التدابير الوقائية لضمان أن الأفراد الذين سيتم سؤالهم يشكلون عينة عشوائية، ولا يشكلون عينة متحيزة، فهذا سيؤدي بدوره إلى نتائج متحيزة. فمثلاً، قد تقتصر العينة المتحيزة على مجموعة الأفراد الذين يقومون بشراء البضائع الكترونياً، وقد يكون سؤال الأفراد الذين اشتروا عدداً من البضائع المختلفة أقل تحيزاً.

ومن أمثلة أهم البحوث بالعينة التي تجري على أرض الواقع تلك البحوث التي تستخدم مسوح ميزانية الأسرة وبُحوث القوى العاملة والتي عادة ما تجريها الحكومات أو المؤسسات الدولية أو الإقليمية. كما تشمل مسوحات التجارة والصناعة والمساكن وأبحاث استطلاع الرأي.

إن إحدى المهام الأولى للمسوح هي إجراء التعدادات ومنها السكانية والصحية ....الخ ، والتعداد السكاني هو جمع المعلومات عن مجموعة من الأفراد أو العناصر. وغالباً ما يؤدي حجم هذه المجموعة إلى استحالة جمع بيانات حول كافة أفرادها أو عناصرها. وتعرف العينة بأنها أي شيء يقل عن مسح أو استطلاع مجموعة السكان بكاملها. وغالباً ما ينظر إليها على أنها جزء صغير من إجمالي عدد السكان.

إطار المعاينة : تسمي القائمة التي تحتوي علي كافة العناصر في التعداد "بإطار العينة". وهذا الإطار ضروري لكي يتمتع كل عنصر في التعداد بالفرصة لكي يصبح جزءا من العينة المختارة. فمثلا، إذا ما تألف التعداد من طلاب مدرسة ما، فإن "إطار العينة" في هذه الحالة هو السجل المدرسي الذي يحوي أسماء كافة طلاب المدرسة. والأمر الأكثر أهمية حول إطار العينة هو انه يجب ان يمثل كافة التعداد، لتدنية مقدار التحيز أكثر ما يمكن.

ويعرف اطار العينة بانه قائمة (وأحيانا خريطة) تضم جميع وحدات المجتمع الذي ستتم دراسته والذي سنختار منها لعينة. ويتضمن إطار المعاينة أيضا بيانات تمكن من الوصول للوحدات التي سيتم اختيارها (كالعنوان البر يد ي ، رقم التلفون أو البر يد الإلكتروني أو الموقع على خريطة .. الخ). وفي حالات كثيرة ، قد يكون الإطار مكونا من عدة قوائم مكملة لبعضها.

من السياق أعلاه، يتضح أن تكوين إطار المعاينة ليس دائما بالأمر السهل. لماذ؟

لأن تكوين الإطار يتطلب حصر لجميع وحدات المجتمع المستهدف. وهل هذا ممكن؟

ليس في جميع الحالات. خاصة في المجتمعات غير المحدودة أو في حالة المجتمعات التي يصعب الوصول لبعض وحداتها. وما لعمل؟

أحيانا، نميز بين مجتمعين: المجتمع النظري (المستهدف) والمجتمع الذي تتم دراسته (المجتمع المعاين).

وقد يواجه الباحث بعض المشاكل في تكوين أطر المعاينة من أهمها:

* تغطية ناقصة (القائمة أو الإطار لا يشمل جميع وحدات المجتمع).

* تغطية زائدة (القائمة أو الإطار تضم وحدات لا يفترض أن تدخل ضمن المجتمع الذي نرغب في دراسته.

* أيضا التكرارات.نفس الوحدة تتكرر في القائمة أو الإطار أكثر من مرة.

والحل، تحديث الإطار قبل الشروع في استخدامه وبما يتوافق مع المجتمع الذي ستتم دراسته. أو تقييد مجتمع الدراسة بما يتوافق مع الإطار المتاح (حل غير عملي ولكن قد لا يكون هناك مخرج آخر). وبما أن تكوين أطر المعاينة مرتبط بالعينات الاحتمالية (العشوائية) لأنه شرط ضروري لإتمامها فقد يعمد البعض لاستخدام العينات غير الاحتمالية أو يلجأ للعينة المنتظمة (عينة احتمالية).

<u>أمثلة لإطارات المعاينة:</u>

* قائمة أعضاء منتديات الإحصائيون العرب.
* دليل الهاتف.
* القائمة البر يد ية للمشتركين في إحدى النشرات على الإنترنت.
* مشتركي مجلة أو جريدة أو مطبوعة.
* سجل الطلبة لمادة مبادئ الإحصاء.

هناك صعوبة في الحصول على إطار للمعاينة في بعض الحالات منها الحالات التالية:

* إجراء استفتاء لزوار موقع إلكتروني.
* إجراء دراسة عن الفقراء في بلد كالسعودية مثلا.
* إجراء دراسة حول تفضيلات المستهلكين لسلعة معينة.

والعامل المشترك بين كل ما سبق تتمثل في صعوبة حصر مفردات المجتمع لتكوين الإطار.

<u>أخطاء المعاينة :</u>

إن المعاينة عرضة لنوعين من الأخطاء هما: خطأ التحيز ، خطأ المعاينة

(1) **خطأ التحيز** : ينشأ التحيز بسبب تأثر المشاهدات بطريقة معينة، تؤدي لإنطباعاً خاطئاً. **ومن أمثلة التحيز:**

- اختيار عينة من الأسماء من دليل الهاتف، يؤدي إلى تحيز بسبب أن جميع الأشخاص الذين سيتم اختيارهم هم ممن يملكون خطاً هاتفياً. أما الأشخاص الذين لا يملكون اشتراكات هاتفية، فسيكون من غير الممكن اختيارهم.

- عند سؤال مديري الأقسام عن التكلفة التي يتوقعونها، فإنهم غالباً ما ينحون نحو المبالغة والتضخيم الزائدين، وبذلك تنحاز توقعات التكلفة نحو الأعلى دوما.
- عندما تقرب مجموعة من الأرقام إلى أقرب ألف، تنشأ مجموعة من أخطاء التحيز.
- عندما يتم سؤال الأفراد عن أعمارهم وقت عيد ميلادهم، فإن مجموع أعمار المجموعة الناتج سيكون اقل من مجموع أعمار المجموعة الحقيقي؛ يعرف الخطأ المطلق بأنه الفرق بين القيمة التقديرية أو التقريبية وبين العدد الحقيقي. وغالباً ما يعبر عن الخطأ النسبي بالنسبة المئوية الناتجة من قسمة الخطأ المطلق على القيمة التقديرية.

(2) خطأ المعاينة : يعرف خطأ المعاينة بأنه الفرق بين قيمة المشاهدات التي تستنتج من العينة، وبين قيمة نفس المشاهدات عند الحصول عليها من كامل التعداد. إلا أن قياس كامل التعداد قد لا يكون ممكناً وبالتالي قد لا يكون من الممكن قياس خطأ المعاينة. عادة ما تؤدي زيادة حجم العينة إلى تقليل نسبة هذا الخطأ، على أنه تنبغي ملاحظة أن التحيز قد ينشا في العينة، وقد لا يمكن إزالته أو تقليله بزيادة حجم العينة. يحدث التحيز، وبالتالي تقع أخطاء في المعاينة من:

- إطار خاطيء للعينة.
- عدم الاستجابة لعملية المسح.
- العينة مسحوبة بطريقة خطأ.
- صياغة الأسئلة غير صالحة او تحتوي علي توجيه للأفراد في اتجاه معين.
- وحدة العينة غير واضحة او غير معرفة بدقة.

وعلي وجه العموم اذا كان المجتمع المسحوب منه العينة كبيرا وكانت العينة كبيرة نسبيا فان:

خطأ المتوسط المسحوب من العينة **<u>الانحراف المعياري المسحوب من العينة</u>**

√ **حجم العينة**

وعليه اذا كان **المتوسط المسحوب من العينة** = $\overline{X}$

الانحراف المعياري المسحوب من العينة = σ

حجـم العينــة $n =$ فان

$\bar{x} \pm (1) \frac{\sigma}{\sqrt{n}}$ تعطي ثقة في التقدير في حدود 68%

$\bar{x} \pm (1) \frac{\sigma}{\sqrt{n}}$ تعطي ثقة في التقدير في حدود 95%

$\bar{x} \pm (3) \frac{\sigma}{\sqrt{n}}$ تعطي ثقة في التقدير في حدود 99.7%

طرق المعاينة

بمجرد أن يتم تنسيق إطار المعاينة، ستصبحون في حاجة لاختيار طريقة المعاينة.وتصنف طرق المعاينة إلى نمطين هما: الطرق العشوائية (العينات الاحتمالية) ، والطرق غير العشوائية (العينات غير الاحتمالية) ، ويعتبر استخدام أسلوب العينات من بين الأساليب الشائعة الاستخدام فى الحياة العملية وذلك لدراسة المجتمعات المختلفة وذلك لما تتصف به دراسات العينات من عدة مميزات تتعلق بالوفر فى التكاليف والجهد والوقت، ذلك بالإضافة إلى إمكانية تحقق قدر كبير من الدقة والكفاية عند جمع البيانات وتبويبها بالمقارنة بما يمكن حدوثه عند دراسة جميع مفردات المجتمع ، ويوضح الجدول التالي اعم الفروق بين نوعي العينات الاحتمالية وغير الاحتمالية.

الفرق بين العينة الاحتمالية والغير احتمالية.

	العينة الاحتمالية	العينة غير الاحتمالية
المجتمع	معروف	غير معروف
السجلات الرسمية	موثقة	غير موثقة
القوانين	تخضع لقوانين	تخضع لعمليات حسابية

	الرياضيات	
النوع	عشوائية	عمدية
الإطار	وجود إطار (المجتمع المعاين) لسحب مفرداتها	لا يوجد إطار للسحب منه
معلومية الاحتمال	احتمال مفرداتها معلوماً ولا يكون مستحيل (الاحتمال ≠ الصفر)	لا يتحقق هنا
الاستدلال	يمكن تطبيق أساليب الاستدلال الإحصائي	لا تطبق أساليب الاستدلال الإحصائي

وسنتناول بشيء من التفصيل كل من نوعي المعاينة.

أولا :الطرق العشوائية لاختيار العينات :

تتميز الطرق العشوائية بشيوع استخدامها وسهولة إجرائها.وفيها تكون لكل مفردة من مفردات المجتمع فرصة متكافئة بالنسبة لاختيارها من بين مفردات العينة. ويتطلب ذلك اختيار مفردات العينة بدون تدخل الباحث فى هذا الاختيار وحتى يمكن تحقيق ذلك الاختيار بالوسائل العشوائية ومثال ذلك اختيار مائة طالب من بين أسماء طلبه السنة الثانية بالكلية وذلك لأجراء دراسة معينه ثم تصميم النتائج بعد ذلك على مجتمع طلبة السنة الثانية، وهناك عدة طرق عشوائية نذكر منها.ويطلق علي هذا النوع من العينات اسم العينات الاحتمالية

وأهم طرق المعاينة العشوائية هي:
1- الطريقة العشوائية البسيطة. 2- الطريقة العشوائية الطبقية. 3- الطريقة العشوائية المنتظمة.

انواع العينات العشوائية (الاحتمالية)

أنواع العينة	الغرض
1.العشوائية البسيطة	احتمالات أفراد العينة متساو . تستخدم جداول الأرقام العشوائية لاختيار العينة
2.المنتظمة	ثبات البعد بين مفرداتها . تتعين من N مفردات المجتمع ويحدد البعد **K حيث $K \leq N/n$**

3.الطبقية	تتعين من طبقات المجتمع (الريف . المدينة . (... عشوائياً لضمان التمثيل للطبقات
4.العنقودية	تنتج من تقسيم المجتمع لمجموعات جزئية يعرف كل منها بالعنقود
5.المتعددة المراحل	تتكون من عدة مراحل حيث تصمم باستخدام أي من العينات السابقة

(1) العينة العشوائية البسيطة: Simple Random Sample

تعتبر المعاينة بهذه الطريقة ابسط الطرق الاحتمالية والعشوائية وبها تكون لكل عينه من حجم معين ممكن سحبها من المجتمع الأصلي فرصة متكافئة ومعلومة المقدار فيما يخص احتمال اختيارها كعينه وبها أيضا تكون لكل مفرده من مفردات المجتمع فرصة متكافئة ومعلومة المقدار فيما يتعلق باحتمال اختيارها ضمن العينة فعند اختيار عينه قوامها مفردتان من مجتمع به ثماني مفردات مثلا فان ذلك يعطى لاى زوج من هذه المفردات فرصة متكافئة مقدارها 28 لتكوين العينة، انه يمكن تكوين ثمان وعشرين عينه مختلفة تضم كل عينه مفردتين من هذا المجتمع وفى هذه الحالة أيضا تكون لكل مفردة فرصة متكافئة للاختيار ضمن العينة ومقدارها 8/2 أي (4/1) وان تكرار سحب عينه من مفردتين مئات المرات ينتج عنه تساوى جميع العينات المسحوبة- تقريبا – فى تكرار سحب كل منها وبالتالي تكرار سحب آيه مفردة تشترك فى تكوين العينة.

ويكون اختيار العينة من شأنه ضمان الفرصة المتكافئة والمعلومة المقدار عينه من حجم معين تسحب من مجتمع البحث، وكذلك لأية مفردة، فلا بد من إتباع وسيلة من الوسائل الآلية أو العشوائية للاختيار. ومن ابسط هذه الوسائل تخصيص بطاقة لكل مفردة من مفردات مجتمع البحث وبعد ذلك تخلط البطاقات خلطا جيدا ثم يسحب منها آليا عدد بقدر حجم العينة المطلوبة فتكون البطاقات المسحوبة شاملة للمفردات التى تتكون منها العينة. بهذه الوسيلة تتميز كل مفردة من مفردات مجتمع البحث برقم معين وبعد ذلك تأتى بكرات متشابهة الشكل والحجم والوزن تبلغ عددها مجموع مفردات مجتمع البحث وتحمل أرقامه على كل كرة رقما. وبعد أن تخلط الكرات بعضها بالبعض الآخر خلطا جيدا يتم الاختيار الصدفي لكرات

عددها حجم العينة المطلوبة ولهذا تتكون العينة من مفردات المجتمع التى تحمل الأرقام المقابلة لتلك المبينة على الكرات المختارة.

ولا تصلح هاتان الوسيلتان أو ما شابهها فى اختيار العينة عندما يضم مجتمع مفردات كثيرة جدا، وهذا هو الشائع لذلك يستعين الباحث بجداول الأرقام العشوائية فى اختيار العينة.

جداول الأرقام العشوائية :تتكون جدول الأرقام العشوائية من قوائم بالالف من الأرقام جاء ترتيبها على أساس من الصدفة وذلك وفقا لما أثبتته كافة الاختبارات الإحصائية والعملية الممكنة، وعند سحب العينة باستخدام جداول الأرقام العشوائية تتبع الخطوات التالية:

1- يخصص لكل مفردة من مفردات مجتمع البحث رقم مسلسل من (1) إلي ن (مجموع المفردات).

2- عن طريق الصدفة تتخذ نقطة بدء على جدول الأرقام الممثلة للمفردات الممثلة للمفردات التى تتكون منها العينة المطلوبة، ويكون اختيار الأرقام على التوالي أفقيا أو رأسيا، وعندما يكون مجتمع البحث شاملا لعشر مفردات أو اقل فان كل رقم من الأرقام الممثلة لمفردات العينة يتكون من عدد واحد ويخصص الصفر للمفردات التى تحمل الرقم 10، ويتكون من عدديين أذا بلغ عدد مفردات مجتمع البحث ما بين 11، 100 مفردة مع تمييز المفردة التى يخصصها الرقم 100 بصفرين، ويتكون من ثلاثة أعداد إذا بلغ عدد مفردات مجتمع البحث ما بين 101، 1000 مفردة مع تمييز المفردة التي يخصها الرقم 1000 بثلاثة أصفار وهكذا.

3- يستبعد تكرار الأرقام التى سبق اختيارها وكذلك يهمل كل رقم يزيد عن مجموع مفردات مجتمع البحث.

تقدير القيم من نتائج العينة.

كثيرا ما يتخذ مقادير القيم محل القياس شكل الوسط الحسابي أو النسبة المئوية وعند الاعتماد على عينه مختارة بالطريقة العشوائية المطلقة من اجل تقديم قيم محل قياس فانه يهم الباحث معرفة مدى تقارب مقادير القيم المتحصل عليها بالعينة من المقادير الحقيقية لهذه القيم فى مجتمع البحث. فنادرا ما يكون مقدار الوسط الحسابي منه كقيمة معينه والمتحصل عليه من عينه مساويا تماما لمقدار الوسط الحسابي لنفس القيمة فى مجتمع البحث وذلك بسبب خطأ العينة.

على انه إذا قمنا باختبار كافة العينات الممكن اختيارها بحجم معين من مجتمع البحث وذلك بالطريقة العشوائية المطلقة ثم حسبنا الوسط الحسابي لقيمة

معينة فى كل عينه من هذه العينات، فان الوسط الحسابي للأوساط الحسابية المتحصل عليها يعادل الوسط الحسابي لنفس القيمة فى مجتمع البحث. فإذا فرضنا مجتمعا مكونا من أربعة أفراد أ ، ب ، جـ ، د ، دخلهم السنوي على التوالي 100، 300، 400 ، 800 جنيه فان الوسط الحسابي لدخل الفرد فى هذا المجتمع يساوى 400 جنيه، وإذا سحبنا من هذا المجتمع عينات ثلاثية بالطريقة العشوائية المطلقة فان جميع التشكيلات المحتمل تكوين العينة منها تشمل أ ب جـ ، أ جـ د، ب جـ د والأوساط الحسابية لدخل الفرد فى هذا العينات تبلغ على التوالي 267، 400، 433 ، 500 جنيه ويبلغ الوسط الحسابي لأوساط العينات والسابق حسابها حوالى 405 وهو يعادل الوسط الحسابي لدخل الفرد فى المجتمع الأصلي، أما الوسط الحسابي المتحصل عليه من أي عينه فانه يزيد أو يقل عادة عن الوسط الفعلي لمجتمع البحث ويتقارب الوسطان بصفة عامة كلما كبر حجم العينة. وينطبق ذلك أيضا بطبيعة الحال على مقادير القيم التى تتخذ شكل النسبة المئوية.

وعلى ذلك فان أي عينه من حجم معين تسحب بالطريقة العشوائية المطلقة من مجتمع معين تختلف عادة عن زميلاتها من نفس الحجم فى تقدير محل القياس ويختلف هذا التقدير عن القيمة الفعلية فى مجتمع البحث بسبب خطأ العينة أي بسبب أن العينة تمثل المجتمع تمثيلا لا يبلغ حد الكمال، ولذلك فانه يهم الباحث التعرف على صحة التقديرات التى توصل إليها من بحث قائم على عينه مسحوبة بالطريقة العشوائية المطلقة وذلك بمعرفة الحدود التى تقع بينها المقادير الحقيقية للقيم فى مجتمع البحاث نفسه. وهذه المعرفة ممكنة فى جميع الحالات التى يتبع فيها البحث طريقة احتمالية عند سحب العينة – وتعتبر هذه الناحية أهم مزايا الطرق العشوائية وهذه المعرفة اكثر يسرا عند اتباع الطريقة الاحتمالية المطلقة.

الخطأ المعياري للوسط الحسابي :بالنسبة لقيمة معينة فى مجتمع معين – كالسن أو الدخل تستخدم العلاقة أساليب لحساب انحراف الوسط الحسابي لهذه القيمة والمتحصل عليه من عينه احتمالية مطلقة عن الوسط الحسابي لهذه القيمة فى مجتمع البحث كله، ويسمى هذا الانحراف بالخطأ المعياري للوسط الحسابي.

$$\sigma x = (\sigma x / n)$$

حيث σx الخطأ المعياري للوسط الحسابي للقيمة محل القياس.

σx الانحراف المعياري لمقادير القيمة فى مجتمع البحث.

n عدد مفردات العينة.

وكما سبق واشرنا الي ان الانحراف المعياري يعتبر من أهم مقاييس التشتت وأكثرها استعمالات، وهو مقياس للتشتت حول الوسط الحسابي ويساوى الجذر التربيعي لمتوسط مربعات انحرافات القيم عن الوسط الحسابي.

ولتوضيح ذلك تطبيقا، نفرض أنه كان معروفا لدينا الانحراف المعياري لمقدار المخزون من صنف معين فى متاجر التجزئة ويبلغ 90 وحدة. فإذا أخذنا عينة من المتاجر مكونه من 100 متجر مثلا فان الخطأ المعياري للوسط الحسابي يساوى:

$$\frac{90}{\sqrt{100}} = \frac{9}{\sqrt{10}}$$

= $\pm$ 9 وحدات وذلك بمعامل ثقة 68%

ومعنى ذلك انه إذا أسفرت العينة المكونة من 100 متجر عن أن الوسط الحسابي للمخزون من الصنف بتاجر التجزئة يبلغ 500 وحدة مثلا فان الوسط الفعلي فى مجتمع البحث يقع فى فترة ثقة حداها 491 ، 509 وحدة وذلك بمعامل ثقة 68% والمقصود المبسط بمعامل ثقة 68% هو أن ثقتنا فى وقع الوسط الفعلي ما بين 491 ، 509 وحدة تبلغ 68% من جميع العينات الممكن سحبها من حجم معين بالطريقة العشوائية المختلفة يقع الوسط الفعلي للقيمة محل القياس فى حدود فترة الثقة المقدرة بالوسط المتحصل عليه من العينة + σ ، وبمعامل ثقة 95% يقع الوسط الفعلي للقيمة محل القياس فى حدود فترة الثقة المقدرة بالوسط المتحصل عليه من العينة + 2σ (فى المثال السابق 500 وحدة + 2σ = 482 إلي 518 وحدة وبمعامل ثقة 99.7% يقع الوسط الفعلي للقيمة محل القياس فى حدود فترة الثقة المقدرة بالوسط المتحصل عليه من العينة + 3σ (فى المثال السابق 500 وحدة + 3σ = 473 إلي 527 وحدة)

ويلاحظ فى المثال السابق أن حساب فترة الثقة قام على أساس المعرفة السابقة للانحراف المعياري للقيمة فى المجتمع الأصلي للبحث. مع انه عند التطبيق العملي فى بحوث التسويق والإعلان لا يكون الباحث على علم عادة بمقدار الانحراف (σ)، ولذلك يستخدم بدلا منه الانحراف المعياري للقيمة فى العينة (σ). وهكذا يستند الباحث على ثلاث حقائق مستمدة من العينة عند حساب فترة الثقة وهذه تشمل الوسط الحسابي والانحراف المعياري وحجم العينة.

والاعتماد على الانحراف المعياري للعينة كبديل للانحراف المعياري للمجتمع لا يؤثر كثيرا فى دقة الخطأ المعياري عندما يكون حجم العينة كبيرا نسبيا (اكثر من 30 مفردة) أما إذا ضمت العينة عددا قليلا من المفردات فان الاعتماد على الانحراف المعياري للعينة كبديل للانحراف المعياري للمجتمع يميل إلي تخفيض مقدار الخطأ المعياري، ويمكن علاج هذه الحالة باستخدام العلاقة التالية فى حساب الخطأ المعياري للوسط الحسابي.

σ x = (σ x/ n1) هذا ويلاحظ أن حساب الخطأ المعياري للوسط الحسابي على النحو السالف لا يأخذ فى الاعتبار علاقة حجم العينة بحجم مجتمع البحث حيث يقوم الحساب أصلا على أساس تأثير هذا الخطأ بتباين مفردات المجتمع فى القيمة محل القياس وبحجم العينة. والواقع أن الخطأ المعياري يتأثر مقداره أيضا بالعلاقة بين حجم العينة وحجم مجتمع البحث بدرجة طفيفة تجعل من غير الضروري آخذها فى الحسبان الا إذا بلغ حجم العينة 5% من مجتمع البحث أو اكثر وحينئذ تستخدم العلاقة التالية:

$$\sigma \ x \quad \frac{\sigma \quad x}{n} * \sqrt{1 \ - \ 6}$$

σ x = (σ x/ n) (1-n/N)

حيث أن عدد مفردات مجتمع البحث (N) وعدد مفردات العينة = n ، (1-p) معامل التصحيح $\frac{n}{N}$ = كسر المعاينة.

**** أهم خصائصها:**

- تستخدم عينات عشوائية (مسحوبة بطريقة عشوائية)
- يتم نقلها بدون تبديل، وبذلك لا يظهر العنصر الواحد أكثر من مرة واحدة في العينة.
- طريقة بسيطة وتعرف بالمعاينة الاحتمالية البسيطة.
- مناسبة للحالات التي يكون فيها التعداد صغيرا نسبيا وإطار المعاينة كامل.

(2)- المعاينة العشوائية المنتظمة:Systematic Sample

وهي شكل من اشكال المعاينة العشوائية التي تتضمن نظاماً معيناً. بشرط ان يتم اختيار العنصر الأول بشكل عشوائي من إطار المعاينة، ومن ثم يتم اختيار كل عنصر تالي بالقوة **n** من قائمة(الإطار) ، بحيث تتم معاينة النسبة المئوية

للتعداد ككل. فمثلا ، إذا ما كان ينبغي اختيار عشرة أسماء من قائمة تتألف من 100 اسم، فيمكن اختيار الأسماء ذوي الأرقام
1 ، 11 ، 31، 41، 51، 61، 71، 81، و 91 من القائمة.
ويلاحظ أن اختيار عينة بالطريقة الاحتمالية البسيطة أو بالطريقة العشوائية الطبقية يقوم على البسيطة – لا يكون لأية تشكيله من حجم معين ممكن سحبها من مجتمع البحث فرصة مكافئة لتكوين العينة.

ولعل أهم مزايا الطريقة المنتظمة هو أن اختيار العينة يتم بسهولة وبسرعة إذا لا تحتاج هذه الطريقة إلي استخدام جداول الأرقام العشوائية ذلك الأمر الذي يقتضيه عادة اتباع الباحث للطريقة العشوائية المطلقة. ثم أنها تكون أكثر كفاية من الطريقة العشوائية المطلقة فى حالة ما إذا كان هناك تشابه فى القيم محل القياس من المفردات المتتالية بكشف بيان مفردات المجتمع وتباين فى هذه القيمة بين المفردات المتباعدة لان فى ذلك ضمانا اكبر لتمثيل مفردات مجتمع البحث.

** **أهم خصائصها:**

- عشوائية بما يكفي للحصول على تقدير لمقدار الخطأ في العينة.
- تسهل عملية اختيار واحدة العينة.
- ليست عشوائية بشكل كامل ولذلك يبقى هناك احتمال لظهور التحيز في العينات المختارة.
- قد يؤدي وجود خصائص معينة في قوائم الأشخاص أو العناصر والتي تبرز في فترات منتظمة إلى ظهور التحيز في العينات المختارة.

<u>(3)– **المعاينة العشوائية الطبقية**</u> Stratified Sample

عند سحب عينه من مجتمع البحث بالطريقة الطبقية يقسم الباحث هذا المجتمع إلي قطاعات حسب خصائص معينة ذات ارتباط بالقيمة محل القياس ثم يختار من كل قطاع عينه بالطريقة العشوائية المطلقة. وسحب العينة بالطريقة الطبقية يمكن الباحث من الوصول إلي تقديرات للقيم محل القياس أدق مما يمكن الحصول عليه بالطريقة العشوائية المطلقة وذلك لأنه عند اتباع الطريقة العشوائية المطلقة فى اختيار عينة من مجتمع معين يلاحظ أن الباحث يختار عينات من مفردات مجتمع البحث فى مجموعة– وهنا ينبع الخطأ المعياري للقيمة محل القياس من مصدرين:–

المصدر الأول: مرجعة احتمال عدم تمثيل العينة لمختلف القطاعات التى يتكون منها مجتمع البحث تمثيلا مطابقا تماما للأصل، المصدر الثاني: مرجعة أن الوحدات المختارة فى العينة التابعة لاى قطاع يحتمل لا تمثل مفردات القطاع تمثيلا يبلغ حد الكمال وفى حالة اتباع الطريقة الطبقية فان الباحث يضمن تمثيل العينة لقطاعات المجتمع ذات الأهمية تمثيلا مطابقا تماما لاصل المجتمع ويكون المصدر الوحيد للخطأ المعياري هو احتمال عدم تمثيل المفردات المختارة من كل قطاع للقطاع نفسه تمثيلا كاملا. ولتوضيح ذلك نفرض أن باحثا كان بصدد قياس متوسط المخزون من صنف معين فى بضع الاف من متاجر التجزئة التى تتعامل فى هذا الصنف وكان من بين هذه المتاجر عدد قليل ذات حجم كبير وتبلغ نسبتها 5% من مجموع المتاجر محل البحث فإذا كان كبر حجم المتجر يستتبعه كبر المخزون من الصنف وقام الباحث بسحب عينة صغيرة من مائة متجر بالطريقة العشوائية المطلقة لقياس متوسط المخزون فان الخطأ المعياري لهذا المتوسط ينبع من احتمال عدم التمثيل النسبي الصحيح لكل من قطاعى المتاجر الكبيرة والصغيرة فى العينة ومن احتمال عدم تمثيل الوحدات المختارة من قطاع المتاجر الكبيرة ومن قطاع المتاجر الصغيرة لهذين القطاعين تمثيلا مطابقا تماما للأصل.

وعندما يتبع الباحث الطريقة الطبقية فى سحب عينة من مجموع المتاجر محل البحث فانه يقسمها إلي قطاعين – أو اكثر من ذلك حسب الحجم من كل قطاع عينه صغيرة بالطريقة العشوائية. وبعد حساب مقادير المخزون من الصنف فى كل عينة يصير الترجيع على أساس أن نسبة عدد المتاجر فى كل قطاع إلي المجموع الأصلي للمتاجر من اجل استخراج رقم يمثل متوسط المخزون من الصنف فى جميع المتاجر وبهذه الطريقة يضمن الباحث التمثيل النسبي الصحيح لكل قطاع ويكون المصدر الوحيد للخطأ المعياري للعينة هو احتمال عدم تمثيل الوحدات المختارة من كل قطاع لمجموع مفردات القطاع تمثيلا تماما وبالتالي تكون القيمة المستخرجة لمتوسط المخزون فى جميع المتاجر اكثر دقة مما لو اعتمد الباحث على الطريقة العشوائية المطلقة. ويمثل الجذر التربيعي لمجموع المقادير المتحصل عليها من الخطوة السابقة الخطأ المعياري للوسط الحسابي.

أساس اختيار مفردات مجتمع البحث، وتتمتع كل مفردة من مفردات مجتمع البحث بفرصة معلومة المقدار فيما يتعلق باحتمال اختيارها ضمن العينة وعلى خلاف الطريقة العشوائية البسيطة فانه لا يمكن سحب عينه من حجم معين تضم أي

تشكيلة ممكنة من مفردات مجتمع البحث وبمعنى آخر فانه ليس لجميع التشكيلات من حجم معين – الممكن سحبها من مفردات مجتمع البحث فرصا متكافئة لتكوين العينة المطلوبة.

ولتوضيح ذلك نفرض أن الباحث يريد اختيار عينه قوامها 100 مفرده من مجتمع البحث يضم 1000 مفردة . ولنفرض أن نقطة البدء فى سحب العينة قد تحددت عشوائيا بالرقم 10، هكذا يصير تكوين العينة بحيث تضم المفردات التى تحمل الأرقام 10 و 20 و 30 و 40 و 50 .... إلي 1000 وإذا تحددت نقطة البدء فى سحب العينة عشوائيا بالرقم 5 فان العينة تضم المفردات التى تحمل الأرقام 5 و 15 و 35و 45.... إلي 995 وهكذا وفى جميع هذه الظروف يلاحظ أن لكل مفردة فى مبدأ الأمر فرصة متكافئة ومعلومة المقدار فيما يتعلق باحتمال– اختيارها ضمن العينة وهى 1 / 10 ولكن التشكيلات أو مجموعات فقط. فلا يمكن فى مثالنا هذا تكوين عينه تشتمل على المفردات التى تحمل الأرقام 2 و 13 و 19 و 27 و 43...الخ مثلا. ويكون احتمال سحب عينة بهذا التكوين صفرا . أي مستحيلا بينما احتمال سحب تشكيلة أخرى من المفردات حددتها نقطة البدء الصدفي فى سحب العينة يبلغ 10/1 كما هو الحال بالنسبة للعينات التى تشتمل على المفردات 10 و 20و 30 الخ أو التى تشتمل على المفردات 3 و 13 و 33.... الخ، وهذا هو المقصود بأنه فى حالة اتباع الطريقة المنتظمة – وعلى خلاف الطريقة العشوائية، وتمتاز الطريقة العشوائية الطبقية فى تكلفتها ودقتها عندما يكون التشابه كبيرا جدا بين مفردات كل قطاع والتباين كبيرا جدا بين قطاع وآخر. وذلك فيما يتعلق بالقيمة محل القياس. فالتشابه الكبير بين مفردات كل طبقة من شأنه الاكتفاء بعينة صغيرة وضآلة الانحراف المعياري فى الطبقة مما ينتج عنه تقدير دقيق للقيمة محل القياس. واما التباين الكبير بين مختلف الطبقات فيما يتعلق بالقيمة محل القياس فيعتبر عاملا أساسيا لتفضيل الباحث باتباع الطريقة العشوائية الطبقية فى اختبار العينة لأنه إذا لم تتفاوت كثيرا الأوساط الحسابية محل القياس مثلا فيما بين الطبقات لما كانت هناك ضرورة لاختيار العينة بهذه الطريقة وكان من الأنسب اختيارها بالطريقة العشوائية المطلقة. هذا ومن المزايا الهامة للطريقة العشوائية الطبقات أنها تمكن الباحث من استخراج تقديرات جديدة للقيم محل القياس وذلك فى كل قطاع من قطاعات مجتمع البحث وكثيرا ما يهم الباحث معرفة ذلك فقد يهمه مثلا معرفة متوسط استهلاك الأسرة من صنف معين فى كل قطاع من قطاعات مجتمع البحث حسب فئات الدخل إلي جانب معرفة متوسط استهلاك

الأسرة من نفس الصنف فى مجتمع البحث كله ويكون ذلك للباحث عند اختيار العينة بالطريقة العشوائية الطبقية.

وفي هذا النوع من العينات يتم تقسيم جميع الأشخاص أو العناصر في إطار العينة إلى عدة مستويات محددة (مجموعات أو أصناف)، أي أنها تختلف فيما بينها عن بعضها البعض، وتغطي مع بعضها البعض كامل التعداد، مثال ذلك: مجموعات الأعمار، الفئات المهنية، والأقاليم الطبوغرافية. يتم اختيار عينة عشوائية بسيطة أو عينة نظامية في كل طبقة، وعندما توضع المستويات مع بعضها البعض فإنها تشكل كامل التعداد.

**** أهم خصائصها:**

- إن العينات عشوائية ضمن الطبقة الواحدة.
- تكون مفيدة بشكل خاص عندما يكون المجتمع به طبقات واضحة.
- تظهر المشاكل عندما تكون الطبقات غير واضحة.

<u>(4)- المعاينة العنقودية</u> Cluster Sample

تتشكل العناقيد عبر تجزئة المساحة المراد مسحها إلى مساحات أصغر، أو "عناقيد"، يتم بعد ذلك اختيار عدد من هذه العناقيد بشكل عشوائي ليتم مسحها، ومن ثم يتم اختيار الأفراد أو العناصر عشوائيا في العناقيد المختارة.

أهم خصائصها:

- ليست طريقة عشوائية بالكامل، على الرغم من أنه يتم اختيار بعض العناصر بشكل عشوائي
- تعتبر طريقة مفيدة عندما يكون التعداد مشتتا بشكل كبير
- قد تكون الطريقة الوحيدة الممكن استخدامها عندما لا يتوافر إطار معاينة مناسب وينبغي توفير الوقت والمال مبذولين لأجل السفر بين مختلف المواضع والمستجوبين قدر الإمكان.

<u>(5)- المعاينة متعددة المراحل</u> Multistage Sample

تتألف هذه الطريقة من سلسلة من العينات المأخوذة عبر مراحل متتالية، وقد تكون أقاليم جغرافية، أو مدن، أو مناطق، أو أفراد، أو أسر.

أهم خصائصها:

- ليست طريقة عشوائية بالكامل، على الرغم من أنه يتم اختيار بعض العناصر بشكل عشوائي ، حيث يمكن استخدام طرق المعاينة العشوائية في كل مرحلة بمجرد أن يتم التقسيم على المناطق والفروق الأخرى
- قد تكون الطريقة العملية الوحيدة في معاينة التعداد السكاني لبلد ما.

ثانيا الطرق الغير عشوائية (الغير احتمالية)

من المهم ان نعرف متي نستخدم عينة احتمالية تمكننا من قياس الاخطاء والتحكم فيها وعينة عمدية لا نحتاج فيها الا لتقديرات تقريبية لا يبني عليها أي اجراءات لاحقة أو دراسات متعمقة ففي مثل هذه الحالة قد يكون التحيز الناتج اقل خطرا من أخطاء المعاينة الناتجة من عينة صغيرة.

(1) العينة العمدية Purposive Sample

أحيانا يكون من المفيد اختيار عينة لا تخضع لنظرية الاحتمالات ،خاصة اذا ما أريد اختيار عينة صغيرة الحجم من مجتمع كبير فعند اذ يعطي أسلوب الاختيار العمدي نتيجة أفضل من اختيار عشوائي لعينة صغيرة الحجم ، وهذا النوع من العينات يتم اختيار مفرداته بطريقة متعمدة او مقصودة بغض النظر عن التحيز في الاختيار ، وغالبا ما يلجأ الباحثون الي هذا النوع من العينات في دراسات الحالة ، او وجود عدد قليل من الأفراد داخل المجنمع المدروس (محدودية المجتمع) او عدم توافر الإمكانيات اللازمة للعينة الاحتمالية ، ولكن يعيب هذا النوع من العينات ان النتائج المتحصل عليها لا يمكن تعميمها او الأخذ بها علي مستوي المجتمع كما هو الحال في العينات تامة العشوائية.

وتعتمد هذه العينة علي البيانات السابقة باعتقاد الباحث بتحقيقها تمثيل جيد للمجتمع

(2) –المعاينة الحصصية Quota Sample

قد يطلب من الباحث إجراء الاستبيان مع أشخاص لهم صفات اجتماعية واقتصادية معينة داخل منطقة معينة .. وهنا لا يتم الاختيار عشوائيا ولكن يتم بأسلوب الحصص والغرض من هذه الطريقة هو استخدام أسلوب التقسيم الطبقي باقل جهد وتكاليف أي بدون الحاجة الي متخصص في مجال تصميم العينات أو الي معرفة أخطاء التقدير.

وغالباً ما يتم تقسيم الحصص النسبية إلى عدة أنواع من الأشخاص، مثلاً، وفق العمر، أو الجنس، أو العمل، أو الفئة الاجتماعية. وتستخدم هذه الطريقة على نطاق واسع في أبحاث التسويق، حيث لا يتم إعطاء الشخص المكلف بإجراء المقابلات أسماء الأفراد الذين يجب عليه مقابلتهم، بل يطلب منه أن يتبع حدسه، أو أن يتصرف على طبيعته. وتستخدم في دراسات التسويق وذلك بالتحكم في مفرداتها كالتقسيم ألعمري لمفرداتها.

أهم خصائصها:

- ليست طريقة للمعاينة عشوائية، لذا فمن غير الممكن تقدير نسبة خطأ المعاينة.
- تعاني من مشكلات في التحكم والتحقق.
- يتم اختيار النسب أو الحصص بشكل تعكس فيه العينة الكلية بدقة خصائص التعداد في عدة أوجه.
- قد تكون الطريقة الوحيدة الممكن استخدامها عندما يراد تحصيل النتائج بسرعة ولا يتوافر إطار معاينة مناسب للعينة العشوائية.

<u>(3) المعاينة الحصصية النسبية</u> Quota Convenient Sample

وهي نوع من العينات غير الاحتمالية يتم فيها التوصل الي العينة المراد اختيارها عن طريق اختيار عناصر العينة بحيث تكون نسبة عناصر العينة التي تملك بعض السمات الخاصة هي نفس نسبة العناصر المميزة في الاشخاص المختارين ويتم فيها اختيار أسهل المفردات وتتميز بسهولة الحصول عليها كقربها من مكان إقامة الباحث ، وفيه تؤخذ نسب معينة من الحصص المختلفة.

<u>(4) المعاينة المساحية</u> Area Sample

وهي العينة التي يتم اختيارها من خلال اختيار الاماكن اولا ثم تختار الافراد ثانياً أي الاعتماد علي الاطار الجغرافي في الاساس ، مثلا دراسة اقتصادية لريف محافظة معينة ، ومركز معين ، وحدة معينة ....الخ.

<u>مثال:</u>

نفترض ان لدينا مجتمعا من الوحدات السكنية ايجارها كالاتي :

ترتيب الوحدات السكنية	القيمة الايجارية	ترتيب الوحدات السكنية	القيمة الايجارية

(1)	120	(7)	420
(2)	400	(8)	380
(3)	350	(9)	110
(4)	140	(10)	230
(5)	250	(11)	125
(6)	220	(12)	225

ونريد تقدير القيمة الايجارية من واقع عينة حجمها ثلاث وحدات سكنية :

أولاً : التقدير باستخدام المعاينة العشوائية البسيطة وباتخاذ رقم (2) كبداية عشوائية فتكون مفردات العينة هي :-

الترتيب	2	6	10
القيمة الايجارية	400	220	230

ويكون تقدير القيمة الإيجارية (400 + 220 + 230) ÷ 3 = 283.3

ثانيا التقدير بأسلوب العينة الطبقية العشوائية وأيضا باتخاذ المفردة الثانية من كل طبقة أساس الاختيار:

الطبقة	ترتيب مفردات الطبقة			القيم الايجارية لمفردات الطبقة				
الاولى	1	4	9	11	120	140	110	125
الثانية	5	6	10	12	250	220	230	225
الثالثة	2	3	7	8	400	350	420	380

وتكون القيمة الايجارية = (140+ 220 + 350) ÷3 = 236.7 جنيه.

ثالثاً : التقدير باسلوب العينة الطبقية العشوائية ولكن بعد ترتيب القيم تصاعديا داخل كل طبقة فتكون الطبقات كالاتي:

الطبقة الاولي	110	120	125	140
الطبقة الثانية	220	225	330	250
الطبقة الثالثة	250	380	400	420

وباختيار القيمة الثالثة في كل طبقة باعتبارها القيمة الوسطي تكون مفردات العينة هي : 125 ، 230 ، 400.

وتكون القيمة الإيجارية : (125 +230 +400) ÷ 3 = 251.7

وبتقدير القيمة الإيجارية من مجتمع الدراسة نجد انها :

247.5 = 12 ÷ (225 + 350 + 400+ 120)

وبفحص التقديرات الثلاث للقيمة الإيجارية نجد ان أقربها للواقع هو التقدير بأسلوب المعاينة الطبقية العشوائية بعد إعادة ترتيب مفردات كل طبقة .. وللتقريب للأذهان نجد أن أخطاء التقدير في الأساليب الثلاثة هي :

- الأسلوب الأول يكون الخطأ في حدود 14.5 %
- الأسلوب الثاني يكون الخطأ في حدود 4.4 %
- الأسلوب الثالث يكون الخطأ في حدود 1.7 %

التخطيط للاستبيانات وتصميمها

تفقد عمليات المسح أهميتها وقيمتها عند مرحلة محددة في معظم استقصاءات الأبحاث، فعمليات المسح توفر طريقة لدراسة الظروف الاجتماعية ، والعلاقات، والسلوك.

"يجب أن ينظر الباحث إلى عمليات المسح على أنها طريقة، وأسلوب هام ومفيد للغاية في اكتشاف المجال، وفي التجميع المباشر وغير المباشر للبيانات المتعلقة بالموضوع قيد الدراسة، حيث تسمح بالتركيز على المشكلة كما تقترح النقاط التي تستحق عناء التعقب والدراسة

ويجب مراعاة النقاط التالية حين استخدام الاستبيانات:

- ما هي الأسئلة التي سيتم طرحها؟
- من هم الأشخاص الذين سيتم سؤالهم؟
- كيف سيتم طرح الأسئلة؟
- كيف، وأين، ومتى سيتم استخدامها؟
- ما هو عدد الأشخاص الذين نحتاج لسؤالهم؟
- كيف ستتم عملية عد وتسجيل البيانات؟

خطوات إجراء المعاينة

يمكن تقسيم عملية المعاينة هذه إلى أربعة خطوات أساسية:

الخطوة الأولى: التخطيط للبحث المعتمد على الاستبيان

- تقرير ماهية البيانات المراد جمعها.
- بحث مزايا وعيوب أسلوب الاستبيان في جمع البيانات بعين الاعتبار.

- أخذ إمكانية تطبيق أسلوب الاستبيان على السكان، أو على العينة بعين الاعتبار.
- تحديد نوعية الاستبيان المراد استخدامها.
- تحديد طريقة جمع البيان، مثال: بريدياً، أو هاتفياً، أو شخصياً، أو الكترونياً كالبريد الإلكتروني أو طرحه علي مواقع يشيكة الانترنت (على الويب).
- تحديد جدول مواعيد إنجاز العمل.

الخطوة الثانية: اختيار العينة وتتم كما يلي :

- تحديد السكان المراد سؤالهم (إطار ومفردات العينة).
- تحديد نمط وحجم العينة المراد استطلاعها
- يجب ملاحظة أن استخدام عينات كبيرة لا يؤدي بالضرورة للحصول عينات خالية من التحيز.
- بناء قاعدة بيانات بالأسماء أو عناوين أو حيازات أو مساحات.

الخطوة الثالثة: تصميم الاستبيان

- وضع الأسئلة
- تنظيم الأسئلة وفق ترتيب منطقي، والقيام بأية عمليات تجميع أو تنظيم ممكنة لتسهيل إتمام الاستبيان
- تقسيم الأسئلة لمجموعات وإضافة العناوين، والعناوين الفرعية لكل مجموعة.
- إضافة تعليمات حول كيفية إتمام الاستبيان
- يجب التوضيح في نهاية الاستبيان ما ينبغي على المستجيب للاستبيان فعله بعد ذلك
- يجب تحضير واختبار الملفات المطلوبة حين استخدام الوسائل الإلكترونية
- تجربة الاستبيان على مجموعة صغيرة من المستجوبين
- مراجعة الإجابات ، والتحقق من أن الأسئلة تستخرج النمط المطلوب من الردود، وتحديد الزمن المطلوب لإتمام الاستبيان.

الخطوة الرابعة: إدارة الاستبيان والإشراف عليه

- كتابة رسالة مرافقة للاستبيان تشرح الهدف منه
- طباعة وثائق تدعو لموافقة الأفراد إذا ما دعت الحاجة لذلك
- إرسال الاستبيانات؛ فإذا ما تم إرسال الاستبيانات بالبريد، ينبغي تجنب ظهورها بمظهر يشابه البريد التسويقي

• يتوجب دراسة ما إذا كان هناك حاجة لتضمين جوائز أو مكافآت رمزية مع الاستبيان
• إذا ما كانت الاستبيانات ستدار بصورة شخصية، فينبغي تحضير تعليمات مكتوبة للأشخاص الذين سيقومون بإجراء المقابلات
• تحضير وإرسال رسائل تذكير إذا ما تطلب الأمر زيادة متوسط الردود على الاستبيان
• ينبغي اخذ الأفراد الذين تم الاتصال بهم ولم يستجيبوا للاستبيان، وذلك للحصول على معلومات حول طبيعة التحيز.

أنماط أسئلة الاستبيان :

هناك عدة أنماط شائعة الاستخدام للأسئلة وتتضمن الاتي :

• أسئلة "نعم" أو "لا"؛ والتي قد تترافق أحيانا مع خيار "ربما" أو خيار "لا أعرف".
• الأسئلة الاختيارية؛ والتي تتضمن إما اختيار جواب واحد أو عدة أجوبة ممكنة.
• الأسئلة التقييمية ذات المقاييس المختلفة؛ حيث ينبغي الأخذ بعين الاعتبار عدد النقاط من كل مقياس، وخاصة فيما إذا كان يحتوي نقطة وسطية.
• الأسئلة المجمعة.
• الأسئلة التي تتطلب كتابة نص حر.

وتقسم الأسئلة إلى نوعين، وهما: الأسئلة المغلقة والأسئلة المفتوحة. السؤال المغلق هو السؤال الذي ينحصر جوابه ضمن مجال محدد من الإجابات المتوقعة، كأسئلة "نعم" أو "لا"، وأسئلة المتوسط التقييمية.

ويشجع السؤال المفتوح الأشخاص الذين سيجيبون على الاستبيان على إضافة آرائهم الخاصة، ومشاعرهم، ومواقفهم، حيث يمكن القائمين على الاستبيان من استخدامها في جمع البيانات النوعية.

وقد أشار بعض الاحصائيون في مجال المسح والاستبيان الي أن الأسئلة المغلقة تنتج "بيانات أفضل" من الأسئلة المفتوحة، ولكن يجب الاشارة الي وجود ثلاث فوائد هامة وحساسة للأسئلة ذات الإجابة التي تستدعي كتابة نص وهذه الفوائد هي :

أولاً: تتيح الأسئلة المفتوحة للباحثين الفرصة في الحصول على أجوبة غير متوقعة.
ثانياً: تصف بعمق أكبر وجهات النظر الحقيقية الأشخاص الذين يجيبون على الاستبيان.

ثالثاً: إن الأفراد الذين يستجيبون للاستبيانات يرغبون بتوافر الإمكانية لإجابتهم على بعض الأسئلة بكلماتهم الخاصة.

إرشادات وتوجيهات عامة لتصميم الاستبيانات

يجب مراعاة النقاط التالية عند القيام بتصميم الاستبيانات:

- اختصار الاستبيانات قدر الإمكان.
- استخدام اللغة البسيطة أي اللغة الأكثر تكراراً.
- إيجاز الأسئلة.
- تجنب استخدام المصطلحات.
- استخدام أشكال بسيطة للردود، مثل "نعم" أو "لا"، والخيارات المتعددة.
- تضمين خيار "ربما" أو "لا أعرف" في الأماكن الملائمة.
- تجنب طرح الأسئلة الشخصية.
- تجنب طرح الأسئلة المرشدة نحو إجابة معينة.
- تجنب طرح الأسئلة التي تتطلب إجراء حسابات ذهنية، أو التي تعتمد على ذاكرة المستجيب.
- طرح سؤال واحد فقط في الفقرة.
- تجنب جعل صفحة الاستبيانات تبدو فوضوية، أو غير منتظمة.
- ترك مساحة كافية للإجابة.
- طرح الأسئلة وفق ترتيب منطقي معين.
- وضع مساحة في الاستبيان، كافية لكي يضع الأفراد ملاحظاتهم الشخصية.
- أخيراً، قوموا بتوضيح أين، ومتى تنبغي إعادة الاستبيان.

اختبارات المعنوية:-

بعد الانتهاء من عملية جمع البيانات، يتم تحليلها. ويسمح الاستدلال الإحصائي للمحللين الإحصائيين بمعرفة ما إذا كانت هناك أية علاقة بين متغيرين للحصول على أدلة من التعداد الذي تم أخذ العينة منه.

وتعرف طرق الاستدلال التي تستخدم لقبول أو رفض الأدلة والادعاءات استناداً إلى عينة البيانات باختبارات الأهمية أو الدلالة.

ولا يمكن بأي حال من الأحوال التأكد (بنسبة 100%) من أن هناك علاقة بين متغيرين. وهناك العديد من مصادر الخطأ التي يجب التحكم بها ، كأخطاء العينة، وتحيز الباحثين ، ومشاكل الثقة والصلاحية، والأخطاء البسيطة، إلخ...

وباستخدام نظرية الاحتمالات، والمنحني الطبيعي، والافتراض بأن هناك علاقة ، يصبح بإمكاننا تقدير احتمالات الخطأ ، فإذا كان احتمال أن نكون مخطئين قليلاً ، فيمكن القول أن ملاحظاتنا حول العلاقة هامة ، وأنها ذات دلالة إحصائية. تعني الدلالة أو الأهمية الإحصائية بأن هناك احتمال كبير للصواب في اكتشاف العلاقة الموجودة بين متغيرين.

ونظرا لوجود فروق بين البيانات، نحتاج لتحديد مستوي مسبق يوضح ما اذا كانت الفروق بينها كبيرة جداً بدرجة "هامة أو ذات دلالة إحصائية" (فروق معنوية).والحد الأكثر شيوعاً لمستويات المعنوية هو 5% ، حيث يفترض أنه يمكننا تكرار عملية المسح /95/ مرة من أصل /100/ والحصول على نتائج متماثلة في كل مرة. وتكون الخمس مرات التي نتلقى فيها نتائج مختلفة هي نسبة الخطأ المقبول.

ومن الصعب على الباحث ان يتأكد بنسبة 100% من أي شيء. ولكي تكون عملية البحث والاستبيان ممثلة للواقع بنسبة 100%، نحتاج إلى معاينة كل فرد من التعداد المطلوب. تدعى هذه العملية بالمسح الإحصائي. ولكننا نادراً ما نكون قادرين على مسح كافة الأفراد في تعداد كبير. وحتى عندما نستطيع القيام بذلك، فإننا نعاين نقطة واحدة من الزمن فقط. مثلاً، الأسبوع الذي يحتاجه الأشخاص لإتمام عملية البحث والاستبيان.

وبدلاً من ذلك، نجعل مستوى الأهمية (مستوى الثقة) 95%، ونقبل بحقيقة أننا سنحصل على نسبة خطأ صغيرة، ونقوم بمسح مجموعة أصغر (عينة).

ويتم حساب حجم المجموعة الأصغر من المعادلة التي تعطينا حجم العينة الذي نحتاجه لنضمن التأكد بنسبة 95% أن عينتنا تمثل التعداد.

خطوات اختبارات المعنوية أو الدلالة الإحصائية :

1. تكوين وصياغة فرضية أن توزيع المعاينة الملاحظ لا يختلف عن التوزيع النظري
2. تحديد مستوى المعنوية أو الدلالة، وليكن مثلاً 5%.
3. اختيار وحساب اختبار المعنوية أو الدلالة الإحصائية
4. تحليل وتفسير النتائج.حيث يتم كتابة النتائج بمصطلحات المعالم للمجتمع كالمتوسط وقد تكون الفرضيات البديلة وحيدة الجانب(أحادية الذيل) أو ثنائية الجوانب(ثنائية الذيل). وتدعي الفرضية وحيدة الجانبان بأن المعامل يكون إما أكبر أو أصغر من القيمة المعطاة بالفرضية الفارغة.

وتدعى الفرضية ثنائية الجوانب أن المعامل ببساطة لا يساوي القيمة المعطاة بالفرضية الفارغة؛ أي أن الاتجاه لا يهم.

اهم اختبارات المعنوية

نستخدم اختبارات المعنوية للتعرف علي الأهمية أو الدلالة. فيما يلي بعض هذه الاختبارات:

- اختبار **t** (ولقد تم الإشارة إليه في باب الانحدار)
- تحليل التباين (ويدعى أيضا **ANOVA**، أو الاختبار **F**) وقد افرد له الباب السادس من هذا الكتاب.
- مربع كاي **Chi-square** (ويرمز له بالرمز χ^2).ولقد تم شرح هذا الاختبار بشيء من التفصيل في الباب السابع من الكتاب
-

ولكن فيما يلي موجزا للاختبارات الثلاث واستخداماتها:

اختبار **t**	نستخدم هذا الاختبار عندما نريد مقارنة متوسطي مجموعتين أو متغيرين فقط. فمثلاً، تقرير فيما إذا باعت منطقة ما أكثر من أخرى، أو إذا ما كانت استجابة الرجال تحتلف عن النساء. يحدد الاختبار **t** فيما إذا ما كانت متوسطات النتائج المختلفة كبيرة بما يكفي لاعتبارها معنوية أو ذات دلالة.
تحليل التباين (**ANOVA**، أو اختبار **F**)	يماثل هذا الاختبار الاختبار **t** على أنه يقارن بين متوسطات ثلاث مجموعات أو أكثر، كعدد الكتب التي يقرأها المواطنون عادة في محافظات ثلاث هي القاهرة والاسكندرية وأسوان.
مربع **Chi-square** (يرمز له بالرمز χ^2	وهو احد مقاييس المعنوية ويعتمد على الفروق بين التكرارات المتوقعة والمشاهدة فعلاً. تشمل استخداماتها التجارب الجينية، واحتمالات قذف قطع النقود، والتحقق من تكرار الأحداث، ومقارنة التكرارت، أو أصوات الإقتراع. فمثلا يمكنها أن تشير فيما إذا حصل مرشح ما على أصوات أكثر بكثير من مرشح آخر.

تصميم التجارب

التجربة:

تعنى التجربة القيام بعمل يمكن بواسطته معرفة حقيقة الأشياء التى نهتم بها وفى الحياة العملية والعلمية، فقد يلاحظ الباحث مثلا تغيرا فى ظاهرة ما، ويكون المطلوب فى هذه الحالة دراسة أسباب هذا التغير، ويقوم الباحث بوضع الفروض لتفسير هذه التغيرات ثم يجرى التجربة لاختبار صحة هذه الفروض.

مثال:

إذا أراد أحد الباحثين أن يدرس تأثير السماد فى زيادة المحصول بسبب ما لاحظه فى تباين المحصول فى قطعتين متجاورتين من الأرض.

لأجراء هذه التجربة سيقوم الباحث باختيار قطعتين من الأرض متشابهتين تماما من الرقعة ونوع التربة والرى والصرف وكذلك يلزم توحيد كافة العمليات الزراعية لكلا القطعتين من حيث طريقة الزراعة وكمية ونوع التقاوي. وهكذا... ويكون الاختلاف فقط فيما يسمى بالمعاملات التجريبية وهى هنا السماد.

العشوائية:

يقصد بالعشوائية فى تصميم التجارب توزيع المعالجات المطلوب دراستها على وحدات التجربة بطريقة لا تؤثر فيها الظروف الخاصة من حيث اختلاف القائمين بالتجربة أو اختلاف المكان أو الزمان أو غير ذلك.

ويمكن تحقيق العشوائية بطرق عدة. فيفرض وجود 4 معالجات أرقاما مسلسلة 1، 2، 3، 4 . ثم نختار أي رقم منها وذلك عن طريق وضع 4 كروت فى سلة كل كرت يحمل رقم من الأرقام السابقة ثم يسحب الكرت الأول الرقم المسحوب هو المخصص للقطعة الأولى وهكذا... ولتبسيط هذه العملية نلجأ إلي جداول الأرقام العشوائية.

ويفيد تطبيق مبدأ العشوائية فى التخلص من عوامل التحيز عند تخصيص المعالجات للوحدات التجريبية والتى من شأنها محاباة إحدى المعالجات على حساب الأخرى والعكس صحيح.

أغراض إجراء التجربة:

يلزم على الباحث باستمرار وقبل البدء فى أجراء التجربة تحديد المشكلة الراغب دراستها وكذلك تحديد الأغراض التى من اجلها تجرى التجربة. كما يجب

أن يقوم الباحث كذلك بمسح للدراسات والتجارب التى أجريت فى موضوع المشكلة محل الدراسة للاستفادة من النتائج المتحصل عليها وكذلك لتفادى الأخطاء والمشاكل التى اعترضتها. ويجب على الباحث بعد ذلك عمل وصف كامل للتجربة المزمع القيام بها من ناحية الوحدات التجريبية والمعاملات أو المعالجات وحجم التجربة والتصميم المقترح أو المناسب. ويمكن حصر أغراض إجراء التجارب فى النقاط التالية:

1-اختبار درجة تأثير المعالجات أو المعاملات المختلفة.

2- تقدير حجم المحصول أو الناتج الناشئ كنتيجة لإحدى المعالجات.

3- تقدير الفروق بين تأثير المعالجات المختلفة.

4- تقدير كفاية التجربة بالنسبة لنفس التجربة فيما لو صممت بطرق أخرى.

5- أجراء اختبارات للفروض وكذلك استخلاص النتائج.

وحتى يمكن الحصول على نتائج من التجربة يمكن وصفها بالدقة وعدم التحيز فان هناك بعض الاعتبارات الواجب الأخذ بها عند بدء وتنفيذ وأثناء أجراء التجربة تحاشيا لاخطاء عديدة يكن أن تتعرض لها التجربة وتؤثر على النتائج بطريقة قد تؤدى إلي تشوية الحقيقة.

1- العناية باختبار مواد التجربة ووحداتها التجريبية.
2- تحسين الأساليب التجريبية المستخدمة.
3- التقسيم المتجانس بقدر الإمكان لوحدات التجربة.
4- ملاحظة القائمين بالعمل ومتابعتهم باستمرار.
5- تكبيد حجم التجربة بزيادة عدد التكرارات.

مما سبق يتبين أن الباحث يواجه بمجموعة من المصاعب يلزم مواجهتها باستخدام طرق إحصائية يمكن بواسطتها تنظيم وتصميم التجربة بهدف تصغير وتبسيط الخطأ الناتج إلي اقل درجة ممكنة.

ويعتبر فيشر فى مقدمة العلماء الذين وضعوا الأسس والمبادئ وصمموا التجارب اللازمة لذلك. ويقصد بتصميم التجارب فى هذا الشأن البحث عن طرق لتوزيع المعاملات أو المعالجات على الوحدات التجريبية مع تصغير للخطأ الناتج

إلـي اقـل درجـة ممكنـة والحصـول علـى تقـديرات صـحيحة لا تـؤثر العوامـل المـراد دراسة تأثيرها.

ويتوقـف اختبـار طريقـة معينـه مـن طـرق تصـميم التجـارب علـى عـدد مـن العوامل نذكر منها.

1- بساطة التصميم وسهولة التحليل الإحصائي.
2- اختبـار التصـميم الـذي يتناسـب وأهميـة التـأثيرات التـى يجـب تخليصـها مـن اثر الإدماج.
3- مراعاة التكاليف عند اختبار التصميم المناسب.
4- إمكانية حساب الخطأ التجريبي.
5- سهولة وإمكانيـة تحليل النتائج الإحصـائية عند فقد إحدى النواتج للوحدات التجريبية أو مجموعة منها.

وسيتناول الجزء التالى بالدراسة بعض الطرق الشائعة فى تصميم التجارب.

التصميم التام العشوائية

بفرض أن المطلـوب هـو دراسـة تـأثير أربـع معـاملات (أ، ب، ج، د) لاربـع مقـادير مختلفـة مـن السـماد علـى النـاتج فـى محصـول معين، أو بفـرض أن المطلـوب هـو دراسة أربع معاملات من الدواء لعلاج مرض معين.

ويمكن أجراء تجربة لدراسة تأثير المعالجات الأربع بتكرار التجربة أربع مرات وبذلك نحصـل علـى قطـاع متجـانس تمامـا مقسـم إلـي 16 قطعـة ويلاحـظ أن توزيـع المعالجات علي القطاع يتم بطريقة عشوائية وذلك باخذ فى معالجة احتمالا متساويا فى التوزيع على أي من القطع 16. ويمكن باستخدام جداول الأرقام العشوائية أو باستخدام طريقة الكروت الحصول على الترتيب التالى للقطع المختلفة فى الحقل مثلا.

د	ج	د	د
ج	د	أ	ب
ب	أ	ج	أ
ج	ب	ب	أ

مزايا التصميم التام العشوائية

1) يمكن استخدام المعالجات أو التكرارات بالاعداد التى يرغبها الباحث.
2) سهولة تحليل النتائج مهما تغير عدد التكرارات لكل معالجة.
3) فقد نتائج إحدى القطع لا يؤثر على سهولة التحليل الإحصائي.

ويعاب عل هذا التصميم ما يعاب عادة على العشوائية المطلقة بصفة عامة حيث لا يمكن ضمان أن تكون الوحدات التى تقع تحت إحدى المعالجات مشابهة تماما لمثيلتها الواقعة تحت معالجة أخرى. فيلاحظ مثلا أن المعالجة (د) لم تجرب فى السطرين الثالث والرابع وكذلك المعالجة (ب) لم تجرب فى السطر الأول. ولذلك فعند حساب متوسط النتائج الأربعة لكل معالجة فان الفروق بين المتوسطات ستتضمن كذلك الفروق بين السطور المختلفة.

كما يعاب على هذا التصميم صعوبة تحقيق التجانس فى القطاع محل التجربة ولذلك لا يستخدم كثيرا فى التجارب الزراعية الا انه يكثر استخدامه فى المعامل حيث يمكن تقسيم المادة المراد دراسة تأثير المعالجات عليها إلي أقسام أو عينات صغيرة ومتجانسة إلي درجة بعيدة بالمقارنة ومثيلتها فى التجارب الزراعية.

مثال:

قطعة من الأرض مقسمة الى ستة أقسام متساوية وكل قسم إلي أربع قطع متساوية. وزعت عليها ستة أنواع من القمح توزيعا عشوائيا والجدول التالى يبين النتائج من الأنواع المختلفة فى القطع المختلفة.

والمطلوب دراسة درجة الاختلاف بين الأنواع المختلفة على المحصول.

و	194	ب	196	أ	182	ج	203
هـ	171	ب	202	د	198	ج	212
أ	214	د	203	ب	208	و	218
هـ	192	و	223	أ	216	ب	224
ج	221	ج	242	د	207	هـ	197
د	222	هـ	204	أ	231	و	232

يمكن كتابة النتائج السابقة فى الجدول التالى:

النوع	النتائج 1	2	3	4
أ	182	214	216	231
ب	196	202	208	224
ج	203	212	221	242
د	198	203	207	222
هـ	171	192	197	204
و	194	218	223	232

ويمكن تبسيط العمليات الحسابية وذلك بطرح 200 من كل رقم واجراء الحسابات التالية كما فى الجدول التالى:

النوع الناتج (−200)					مجمع المربعات	متوسط مجموع المربعات	متوسط النوع − 200
أ	18−	14	16	31	43	1737	462.25	10.75
ب	4−	2	8	24	30	660	225.00	7.5
ج	3	12	21	42	78	2358	1521.00	19.5
د	2−	3	7	22	30	546	225.00	7.5
هـ	29−	8−	3−	4	36−	930	324.00	9.0
و	6−	18	23	32	67	1913	1122.25	16.75
المجموع					222	8144	3879.5	

$2(222)$

.. معامل التصحيح = ـــ = 2053.5

24

ويمكن تكوين جدول تحليل التباين على النحو التالى

مصدر التباين	درجات الحرية	مجموع المربعات	م مجموع المربعات	ف
بين الأنواع	5	2053.5−3879.5 182.60 =	365.2	
داخل الأنواع	18	3879.5−8144.0 4264.5=	236.9	1.541
المجموع	23	2053.5−8144.0 6090.5 =		

ومن الجدول (ف) وبدرجات حرية 5 ، 18 وعند مستوى معنوية 0.05

.. ف 0.05 (5 ، 18) = 2.77

وبمقارنة ف الجدولية بقيمة ف المحسوبة يتبين أن ف المحسوبة اقل من الجدولية وهذا يعنى أن الفروق بين الأنواع المختلفة غير معنوية ويمكن إهمالها ولا تختلف اختلافا جوهريا عن الصفر.

ولاختبار الفرق بين كل متوسطين يحسب الخطأ المعياري للفرق وهو:

1 1

= ع2 (-) + (-)

ن1 ن2

حيث ن1 = عدد المفردات فى المجموعة (1).

ن2= عدد المفردات

وبتساوى عدد المفردات فان الخطأ المعياري=

2

= ع2 (-)

ن

... خ . م = 203.9 × 4/2 = 10.1

وبالكشف عن قيمة (ت) من الجدول عند درجات حرية 18، 5 ومستوى 0.05ويستوى معنوية تبين أنها تساوى 2.101.

معنى ذلك الفرق بين المتوسطين يساوى أو يزيد عن

2.101 × 10.1 = 21.22

فهو فرق معنوي ولا يمكن إهماله.

وبالبحث فى متوسطات الأنواع المختلفة يتبين أن متوسطات الأنواع ج، و تختلف اختلافا جوهريا عن متوسطات الأنواع أ ، ب ، د ، هـ.

القطاعات الكاملة العشوائية

تبين من دراسة التصميمات التامة العشوائية انه يلزم أن تكون جميع القطع متجانسة تماما، وفى حالة عدم تجانسها تماما فان الفرق بين هذه القطع سيدخل ضمن ما يسمى بالخطأ التجريبي ومن ثم يقلل من كفاءة التجربة موضع الدراسة ولما كان افتراض تجانس كافة القطاعات أمر مستبعد لدرجة بعيدة فان الأمر يستلزم ضرورة اختبار طريقة أخرى من شانها تقسيم التجربة إلي مجموعات متجانسة فيما بينها من الوحدات التجريبية وتسمى كل مجموعة منها بالقطاع ويمثل كل قطاع واحد الجميع المعالجة. ومما لا شك فيه أن مثل هذا التجميع فى قطاعات سيؤدى الى تصغير الخطأ التجريبي.

وفى القطاعات الكاملة العشوائية يمكن مقارنة المعالجات التى تدخل فى قطاع واحد مهما واحد مهما اختلفت وحدات التجربة بين قطاع وآخر طالما أنها متجانسة داخل القطاع الواحد.

مثال :

إذا ما كان المطلوب هو تقسيم قطعة من الأرض إلي اربعة أقسام وكانت خصوبة الأرض تتزايد بالاتجاه من أعلى إلي اسفل أو من اليمين إلي الشمال ويتم ذلك بتقسيم الأرض إلي أربعة قطع توزع عليها المعالجات المختلفة بطريقة عشوائية بشرط أن تظهر كل معالجة فى كل قطاع مره واحدة فقط وذلك ما فى الشكلين الاتيين:

قطاع (1)	قطاع (2)	قطاع (3)	قطاع (4)
أ	ب	ب	د
ب	ج	د	أ
ج	د	ج	ب
د	أ	أ	ج

قطاعات رأسية

قطاع (1)	ب	أ	ج	د
قطاع (2)	د	ج	ب	أ
قطاع (3)	د	أ	ب	ج
قطاع (4)	د	ج	أ	ب

قطاعات أفقية

مميزات هذا التصميم

1) دقة النتائج بالمقارنة والتصميم التام العشوائية.
2) سهولة وبساطه التحليل الإحصائي عند فقد بعض قطع أو نتائج قطاع كامل.
3) إمكانية استخدام أي عدد من المعالجات والتكرارات.

مثال :

تجربة اجريت بين 6 أصناف من محصول معين (أ، ب ، ج، د، هـ، و) وكانت التجربة فى قطاعات كاملة كل منها يتكون من 6 قطع، وقد أخذت 4 قطاعات وكان التوزيع العشوائي للتجربة كما يلي:

قطاع (1)

أ	هـ	ج
ب	و	د

قطاع (2)

و	د	هـ
أ	ب	ج

قطاع (3)

أ	ب	أ
و	د	جـ

قطاع (4)

و	أ	د
هـ	ب	ج

وكانت النتائج المتحصل عليها وفقا للجدول التالى:

القطاع	1	2	3	4	المجموع	المتوسط
أ	2.1	2.2	3.4	2.5	9.2	2.30
ب	2.5	2.4	2.7	3.0	10.6	2.65
ج	2.4	2.6	2.8	3.0	10.8	2.70
د	3.0	3.1	3.2	3.1	12.4	3.10
هـ	3.0	2.8	3.2	3.2	12.2	3.05
و	3.2	3.3	2.8	3.5	12.8	3.20
المجموع	16.2	16.4	17.1	18.3	68.0	2.83

وعلى ذلك يمكن القول أن:

1) ناتج القطعة الواحدة = المتوسط العام + تأثير الأصناف

+ تغير القطاعات + خطأ تجريبي

2) المجموع الكلى للمربعات

$$= (2.1)^2 + (2.2)^2 + (2.4)^2 + (2.5)^2$$

$$+ (2.5)^2 + (2.4)^2 + (2.7)^2 + (3.0)^2$$
$$+ (2.4)^2 + (2.6)^2 + (2.8) + (3.0)^2$$
$$+ (3.0)^2 + (3.1)^2 + (3.2)^2 + (3.1)^2$$
$$+ (3.0)^2 + (2.8)^2 + (3.2)^2 + (3.2)^2$$
$$+ (3.2)^2 + (3.3)^2 + (2.8)^2 + (3.5)^2$$
$$= 195.92$$

3) معامل التصحيح = $\frac{(68.0)^2}{24} = 192.67$

4) مجموع المربعات بعد التصحيح

$$= 195.92 - 192.67 = 3.25$$

5) مجموع مربعات الأصناف

$$= 1/6\ (9.2)^2 + (10.6)^2 + (10.8)^2$$
$$+ (12.4)^2 + (12.2)^2 + (12.8)^2$$
$$= (195.02 - 192.67) = 2.35$$

6) مجموع مربعات القطاعات

$$= ¼\ (16.2)^2 + (16.4)^2 + (17.1)^2 + (18.3)^2$$
$$= (193.12 - 192.67) = 0.45$$

7) مجموع مربعات الخطأ

= المجموع الكلى للمربعات – مجموع المربعات للاصناف – مجموع مربعات القطاعات.

$$= 3.25 - 2.35 - 0.45 = 0.45$$

ويمكن تصوير النتائج فى جدول تحليل التباين التالى:

مصدر التباين	درجات الحرية	مجموع المربعات	متوسط المربعات	ف
بين القطاعات	4−1=3	0.45	0.15	5.00
بين الأصناف	6−1= 5	2.35	0.47	15.66
الخطأ التجريبي	(4−1) (6−1) = 15	0.45	0.03	
المجموع	23	3.25		

بمقارنة قيمة ف المحسوبة لكل من القطاعات والأصناف بقيمة ف من الجدول يلاحظ معنوية تأثير كل من القطاعات والأصناف عند مستويات المعنوية العالية.

المربع اللاتيني

بفرض أن المطلوب دراسته هو تأثير أربعة أسمدة أ، ب، ج، د على كمية الناتج من محصول معين، ويتم ذلك عن طريق اختبار قطعة من الأرض ثم تقسم إلي أقسام متساوية فى المساحة عن طريق صفوف أربعة أعمدة. ويلزم استخدام كل نوع من أنواع السماد فى أربعة أقسام تختار عشوائيا بحيث لا يستخدم فى سماد سوى مرة واحدة فى أي صنف وكذلك فى اى عمود كما يتضح ذلك من الشكل التالى:

ا	ب	أ	د	د
2	د	ج	أ	ب
3	أ	د	ب	ج
4	ج	ب	د	أ

ويستخدم المربع اللاتيني عادة فى التجارب الزراعية حيث أن خصوبة الصفوف والأعمدة لا تدخل فى مقارنة المعالجات المختلفة.

ويفرض أن المطلوب توزيع أربع معالجات (أسمدة) فى مربع لا تينى 4×4 وانه تم اختبار المربع اللاتيني السابق عشوائيا، يلي ذلك اختيار الأعمدة عشوائيا وبفرض أنها 2، 1، 4، 3 فتكتسب هكذا.

أ	ب	ج	ج
ج	د	ب	أ
د	أ	ج	ب
ب	ج	أ	د

وبعد ذلك تختار الصفوف عشوائيا أيضا بفرض أنها 3، 4، 1 ،2

د	أ	ج	ب
ب	ج	أ	د
أ	ب	د	ج
ج	د	ب	أ

وباختيار الرموز عشوائيا يمكن مثلا الحصول على د، أ، ج ، ب. معنى ذلك توضع د مكان أ ، أ مكان ب وهكذا وذلك على النحو التالى:

ب	د	ج	أ
أ	ج	د	ب
د	أ	ب	ج
ج	ب	أ	د

مثال : بفرض أن المطلوب هو دراسة تأثير خمسة أنواع من الأسمدة علي كمية الإنتاج فى محصول معين. والمربع اللاتيني التالى يبين كميات الإنتاج:

أ	13	ب	99	ج	21	د	7	هـ	6
د	9	هـ	8	أ	15	ب	7	ج	16

ب	11	ج	17	د	8	هـ	10	أ	17
هـ	8	أ	15	ب	7	ج	10	د	7
ج	11	د	9	هـ	8	أ	15	ب	11

ولتسهيل العمليات الحسابية يمكن طرح 10 من كل رقم وذلك كما يتضح من المربع اللاتيني التالى:

.. مجموع الأسمدة كما يلي:

أ= 25 ب= -5 ج= 25

د= -10 هـ= 10

مجموع المربعات الكلية = $413 - (25)^2/25 = 388$

مجموع المربعات بين الصفوف

$= (6)^2 + (5)^2 + (13)^2 + (-3)^2 + (4)^2 - 5(25)^2 = 26.0$

مجموع المربعات بين الأسمدة

$= (2)^2 + (8)^2 + (9)^2 + (-1)^2 + (7)^2 - 5(25)^2 = 14.8$

مجموع المربعات بين الأسمدة

$= (35)^2 + (-5)^2 + (25)^2 + (-10)^2$

$+ (10)^2 - 5(25)^2 = 27.0$

.. البواقى = $388 - (260 \times 14.8) + 270 = 77.2$

ويمكن تكوين جدول تحليل التباين على النحو التالى:

متوسطات المربعات	مجموع المربعات	درجة الحرية	مصدر التباين
6.5	26.0	4	بين الصفوف
3.7	14.8	4	بين الأعمدة
67.5	270.0	4	بين الأسمدة
6.43	77.2	12	البواقى
	388.0	24	المجموع

وبتطبيق اختبار ف بالنسبة للصفوف

∴ ف = (6.43/6.5) = 1.01.

وبمقارنة قيمة ف المحسوبة بمثيلتها الجدولية تبين انه لا يوجد فروق معنوية نتيجة خصوبة التربة من صف إلي آخر. وبتطبيق اختبار ف بالنسبة للأعمدة.

ف= (6.43/3.7) = 0.57

معنى ذلك أيضا انه لا توجد فروق معنوية نتيجة خصوبة التربة من عمود إلي آخر. أما بين الأسمدة فان قيمة ف المحسوبة هى:

ف= (6.43/67.5)= 10.5

معنى ذلك انه توجد فروق معنوية ومؤكدة بين أنواع الأسمدة المختلفة، وواضح أن السماد أ ، ج لهما اثر كبير على الإنتاجية دون غيرهما.

تذكـر

- المسح (الحصر الشامل):هو عد جميع الأشخاص الموجودين على قيد الحياة داخل الحدود السياسية للجمهورية ليلة التعداد .
- تتميز المسوح بأنها تمكن من جمع بيانات من عدد كبير من الأفراد ، كما أنها تسمح بإجراء مقارنات دقيقة بين إجابات المبحوثين المختلفة.

ومن عيوبها أحيانا ان المادة التي تم جمعها سطحية، وقد تكون الإجابات هي ما يتظاهر الناس بالاعتقاد به، أكثر منها ما يعتقدون به حقاً.

- أسلوب العينة هو البديل الوحيد فى كثير من الأحيان للحصر الشامل والغاية الأولى والأساسية من دراسة عينة ما هو إلا تعميم نتائجها على المجتمع الأصلي الذي اشتقت منه .
- تعرف العينة بأنها " جزء من المجتمع يتم اختياره لتمثيل المجتمع " أما المعاينة فتعرف " عملية اختيار جزء من المجتمع الإحصائي للاستدلال على خواص المجتمع بأكمله عن طريق تعميم نتائج العينة ".
- من أمثلة البحوث بالعينة التي تجري على أرض الواقع تلك البحوث التي تستخدم مسوح ميزانية الأسرة وبُحوث القوى العاملة والتي عادة ما تجريها الحكومات أو المؤسسات. كما تشمل مسوحات التجارة والصناعة وأبحاث استطلاع الرأي.
- تسمي القائمة التي تحتوي علي كافة العناصر في التعداد "بإطار العينة". وهذا الإطار ضروري لكي يتمتع كل عنصر في التعداد بالفرصة لكي يصبح جزءا من العينة المختارة.
- يعرف إطار العينة بأنه قائمة (وأحيانا خريطة) التي تضم جميع وحدات المجتمع الذي ستتم دراسته والذي سنختار منها لعينة. ويتضمن إطار المعاينة أيضا بيانات تمكن من الوصول للوحدات التي سيتم اختيارها (كالعنوان البريد ي ، رقم التلفون أو البر يد الإلكتروني أو الموقع على خريطة .. الخ). وفي حالات

كثيرة ، قد يكون الإطار مكونا من عدة قوائم مكملة لبعضها.

- المعاينة عرضة لنوعين من الأخطاء هما: خطأ التحيز وينشأ بسبب تأثر المشاهدات بطريقة معينة، تؤدي لانطباع خاطئ ، خطأ المعاينة ويعرف بأنه الفرق بين قيمة المشاهدات بالعينة، وبين قيمة نفس المشاهدات من كامل التعداد. إلا أن قياس كامل التعداد قد لا يكون ممكناً وبالتالي يصعب قياس خطأ المعاينة.
- عادة ما يؤدي تكبير حجم العينة إلى تقليل نسبة هذا الخطأ، على أنه تنبغي ملاحظة أن التحيز قد ينشا في العينة
- خطأ المعاينة قد لا يمكن إزالته أو تقليله بزيادة حجم العينة. يحدث التحيز ، وبالتالي تقع أخطاء في المعاينة بسبب إطار خاطئ للعينة، او عدم الاستجابة لعملية المسح ، او ان تكون العينة مسحوبة بطريقة خطأ، او صياغة الأسئلة غير صالحة او تحتوي علي توجيه للأفراد في اتجاه معين ، او وحدة العينة غير معرفة بدقة.
- العينة الاحتمالية تتصف بان المجتمع معروف ، السجلات موثقة ، تخضع لقوانين الرياضيات، عشوائية ، وجود إطار (المجتمع المعاين) لسحب مفرداتها، احتمال مفرداتها معلوماً ولا يكون مستحيل (الاحتمال ≠ الصفر) ،يمكن تطبيق أساليب الاستدلال الإحصائي
- العينة غير الاحتمالية المجتمع غير معروف، السجلات الرسمية غير موثقة ، تخضع لعمليات حسابية، عمدية ، لا يوجد إطار للسحب منه، لا يتحقق هنا، لا تطبق أساليب الاستدلال الإحصائي
- **تتنوع العينات الي :**
- **عشوائية بسيطة:** وفيها احتمالات أفراد العينة متساو وتستخدم جداول الأرقام العشوائية لاختيارها.
- **منتظمة** : وتتميز ثبات البعد بين مفرداتها . تتعين من N مفردات المجتمع

ويحدد البعد K حيث $K \le N/n$

- **طبقية** : وتتعين من طبقات المجتمع (الريف والمدينة) والاختيار بين الطبقات عشوائي لضمان التمثيل للطبقات
- تتكون جدول الأرقام العشوائية من الألف الأرقام مرتبة على أساس من الصدفة وذلك وفقا لما أثبتته كافة الاختبارات الإحصائية والعملية الممكنة
- يحسب الخطأ المعياري للوسط الحسابي بالقانون $б x = (\sigma x/ n)$
- حيث σx الخطأ المعياري للوسط الحسابي للقيمة محل القياس.
- الاعتماد على الانحراف المعياري للعينة كبديل للانحراف المعياري للمجتمع لا يؤثر كثيرا فى دقة الخطأ المعياري عندما يكون حجم العينة كبيرا نسبيا (اكثر من 30 مفردة) أما إذا ضمت العينة عددا قليلا من المفردات فان الاعتماد على الانحراف المعياري للعينة كبديل للانحراف المعياري للمجتمع يميل إلي تخفيض مقدار الخطأ المعياري.
- لا يأخذ فى الاعتبار علاقة حجم العينة بحجم مجتمع البحث حيث يقوم الحساب أصلا على أساس تأثير هذا الخطأ بتباين مفردات المجتمع فى القيمة محل القياس وبحجم العينة. والواقع أن الخطأ المعياري يتأثر مقداره أيضا بالعلاقة بين حجم العينة وحجم مجتمع البحث بدرجة طفيفة تجعل من غير الضروري آخذها فى الحسبان الا إذا بلغ حجم العينة 5% من مجتمع البحث أو أكثر $x = (\sigma x/ n) (1-n/N)$
- تقسم أسئلة الاستبيان إلى نوعين، وهما: الأسئلة المغلقة والأسئلة المفتوحة. السؤال المغلق هو السؤال الذي ينحصر جوابه ضمن مجال محدد من الإجابات المتوقعة، كأسئلة "نعم" أو "لا"، وأسئلة المتوسط التقييمية.
- يشجع السؤال المفتوح الأشخاص الذين سيجيبون على الاستبيان على إضافة آرائهم الخاصة، ومشاعرهم، ومواقفهم، حيث يمكن القائمين على الاستبيان من استخدامها في جمع البيانات النوعية.

1. تعرف طرق الاستدلال التي تستخدم لقبول أو رفض الأدلة والادعاءات استناداً إلى عينة البيانات باختبارات الأهمية أو الدلالة(اختبارات المعنوية)

- نستخدم اختبارات المعنوية للتعرف علي الأهمية أو الدلالة ومن اهمها اختبار **t** لمقارنة متوسطي مجموعتين أو متغيرين فقط. ، تحليل التباين (اختبار **F**) ويقارن بين متوسطات ثلاث مجموعات أو أكثر،مربع كاي (χ^2) ويعتمد على الفروق بين التكرارات المتوقعة والمشاهدة فعلاً
- **التجربة:**تعنى القيام بعمل يمكن بواسطته معرفة حقيقة الأشياء التى نهتم بها وفى الحياة العملية والعلمية، فقد يلاحظ الباحث مثلا تغيرا فى ظاهرة ما، ويكون المطلوب فى هذه الحالة دراسة أسباب هذا التغير، ويقوم الباحث بوضع الفروض لتفسير هذه التغيرات ثم يجرى التجربة لاختبار صحة هذه الفروض.
- يقصد بالعشوائية فى تصميم التجارب توزيع المعالجات المطلوب دراستها على وحدات التجربة بطريقة لا تؤثر فيها الظروف الخاصة من حيث اختلاف القائمين بالتجربة أو اختلاف المكان أو الزمان أو غير ذلك.
- من مزايا التصميم التام العشوائية انه يمكن استخدام المعالجات أو التكرارات بالإعداد التي يرغبها الباحث، وسهولة تحليل النتائج مهما تغير عدد التكرارات لكل معالجة، كما ان فقد نتائج إحدى القطع لا يؤثر على سهولة التحليل بينما يعاب عليه انه لا يمكن ضمان العشوائية التامة للمعالجات .
- فى القطاعات الكاملة العشوائية يمكن مقارنة المعالجات التى تدخل فى قطاع واحد مهما واحد مهما اختلفت وحدات التجربة بين قطاع وآخر طالما أنها متجانسة داخل القطاع الواحد.
- يستخدم المربع اللاتيني عادة فى التجارب الزراعية حيث أن خصوبة الصفوف والأعمدة لا تدخل فى مقارنة المعالجات المختلفة.

تماريــن

س1 : اذكر أساليب جمع البيانات ثم وضح مزايا وعيوب المسوح الميدانية موضحا مبررات اللجوء إلي استخدام أسلوب العينات

س2 : عرف كل من المصطلحات الاتية :

المعاينــة–خطــأ المعاينــة –خطــأ التحيــز –العينــة الاحتماليــة التجربــة العشــوائية– التصميم التام العشوائية–القطاعات الكاملة العشوائية

س 3: تكلم بإيجاز عن أنواع طرق المعاينة – ثم اشرح بالتفصيل خصائص نوعين فقط منها ؟

س 4 : اذكر فقط خطوات التخطيط للاستبيانات وتصميمها ؟

س 5 : وضح كيف يمكن ان تؤثر أسئلة الاستبيان علي كفاءة المعاينة

س6 : وضــح الفــرق بــين اختبــارات المعنويــة الــثلاث اختبــار t – تحليــل التبــاين (اختبار F) –مربع كاي ($\chi 2$)

الباب السادس

تحليل التباين (ANOVA - analysis of variance)

تمهيد

تحليل التباين هو مجموعة من النماذج الإحصائية (statistical model) يتم من خلالها مقارنة المتوسطات لمجتمعات إحصائية مختلفة عن طريق تقسيم التباين variance الكلي الملاحظ بينهم إلى أجزاء مختلفة.

وأول طرق تحليل التباين تم وضعها من قبل الإحصائي رونالد فيشر في العشرينات والثلاثينات من القرن العشرين لذلك تعرف أحيانا بتحليل فيشر للتباين.

قد سبق دراسة اختبارات الفروض لتساوي متوسطي مجتمعين ولكن هناك دراسات لتساوي متوسطات ثلاث مجتمعات أو أكثر، وهل يمكن التعميم لأكثر من مجتمعين، نعم قد يكون ذلك ولكن وجود ثلاث عقبات رئيسية تجعلنا البحث عن طريقة أخرى وهذه الثلاث عقبات هي:

1) الجهد المبذول في المقارنة بين كل مجتمعين وخاصة إذا كثر عدد المجموعات الثنائية والتي عددها يحدد من $^{ن}ق_{ر}$ = ن(ن − 1) ÷ 2

2) إذا كان لدينا العديد من المستويات فالمقارنة الثنائية بينهم تفقد الكثير من المعلومات المتوفرة لدينا عن المجتمع محل الدراسة وهو ما ينقص من دقة تقدير معالم المجتمع.

3) إن كثرة المستويات ينقص بشكل ملحوظ القيمة $(1 - \alpha)^{ن}$ ما يزيد في قيمة P حيث $P = 1 - (1 - \alpha)^n$ حيث α الخطأ من النوع الأول.

فإن كنا بصدد اختبار تساوي متوسطات لخمس مجتمعات بمستوى معنوية 0.05 فاحتمال الحصول على قرار صحيح بعدم وجود فرق معنوي واحد لكل اختبار هو 0.95 وعليه يكون احتمال الحصول إلى قرار صحيح بالنسبة لكل الاختبارات وعددها 10 من $^{5}ق_{2}$ = (5 × 4) ÷ (2 × 1) = 10 يساوي $(0.95)^{10} = 0.4013$ مما يؤدى لاحتمال القرار الخاطئ (α) يساوي $1 - (0.95)^{10} = 0.4013$ ويزداد بزيادة عدد المجتمعات (احتمال الوقوع في الخطأ من النوع الأول) لا بد من أسلوب آخر لاختبار تساوي المتوسطات يعرف بتحليل التباين (Analysis of variance) أو ANOVA الذي قدمه العالم فيشر (Ronald A. Fisher) كأسلوب لتحليل البيانات للتجارب المختلفة وهو عبارة

عن مجموعة من الطرق الإحصائية المساعدة لاختبارات الفروض أبسطها one-way ANOVA.

تحليل التباين الاحادي (في اتجاه واحد)

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

1) العينات عشوائية ومستقلة.
2) مجتمعات هذه العينات تتبع توزيع طبيعي.
3) تساوي تباين المجتمعات التي أخذت منها العينات العشوائية المستقلة.

مثال

ولتوضيح ما سبق بمقارنة متوسطات ثلاث مجتمعات باستخدام ثلاث عينات (تحقق فيها الشروط الثلاثة السابقة) موضحة بالجدول الآتي:

العينة الأولى	العينة الثانية	العينة الثالثة
40	27	33
41	28	32
40.5	26.5	33.5
38.5	26.5	31.5
$\bar{X}_1 = 40$	$\bar{X}_2 = 27$	$\bar{X}_3 = 32.5$
$S_1 = 1.08$	$S_2 = 0.71$	$S_3 = 0.91$

وبالنظر في البيانات لتحديد ما اذا كان هناك فرق بين المتوسطات يتضح انه يوجد فرق نعم ويتضح بمجرد النظر ، فالتشتت (أو التباين) ظاهر بين هذه الارقام 40، 27، 32.5 (المتوسطات) وبمقارنته بالتشتت بين العينات مفرداتها (40 ، 41 ، 40.5، 38.5) لا يبدو ذلك التشتت مطلقاً .

إذا أخذنا البيانات الآتية:

العينة الأولى	العينة الثانية	العينة الثالثة

10 60 27.5	50 20 11	40 15 65
$\bar{X}_3 = 32.5$ **$S_3 = 25.4$**	**$\bar{X}_2 = 27$** **$S_2 = 20.4$**	**$\bar{X}_1 = 40$** **$S_1 = 25$**

فالبيانات هنا لها نفس المتوسطات في البيانات السابقة ولكن التشتت (داخل لعينات) كبيراً بما هو عليه في المتوسطات.
فالدليل على وجود الفرق بين متوسطات الجدول الأول واضح في حين لا يظهر ذلك بوضوح في بيانات الجدول الثاني بالرغم من تساوي المتوسطات في الحالتين ولذا يتبين لنا ان القصد من تحليل التباين والذي يهتم بالفرق بين المتوسطات والذي يقاس بالتشتت داخل البيانات.

<u>اختبار تساوي أكثر من متوسطين</u>

ليكن لدينا الاختبار التالي:

$H_o : \mu_1 = \mu_2 = \mu_3 = \dots \mu_k$, H_1 : الأقل عدم تساوي متوسطين على

وبفرض أن العينات مأخوذة من مجتمعات طبيعية ولها نفس التباين، فلاختبار يرتكز على مقارنة التباين داخل العينات وبينها بتقدير التباين المشترك بطريقتين فالأولى لا تعتمد على صحة أو عدم صحة الفرض الصفري بينما تتأثر الطريقة الثانية بالفرض الصفري فإن تبين خلاف معنوي بين الطريقتين (في التقدير) أخذنا بعدم صحة الفرض الصفري لأن عدم تساوي المتوسطات قد أثر على التقدير الثاني فتسبب في تجاوزه التقدير الأول فرفض H_o ونفصل ذلك بتقديرين للتباين σ^2 كالآتي:

للتبسيط لنأخذ عينات من المجتمعات محل الدراسة لها نفس الحجم وحيث أن التباين في المجتمعات متماثل فنقدر التباين σ^2 بمتوسط التباينات في العينات أي أن:

$$\hat{\sigma}^2 = \sum_{i=1}^{k} \frac{S_i^2}{k}$$

ويرمز لهذا التقدير بالرمز S_w^2 لكونه يمثل التباين داخل المجموعات (Within Group) أي:

$$S_w^2 = \sum_{i=1}^{k} \frac{S_i^2}{k}$$

وفي حالة تساوي حجم العينات. ويلاحظ عدم اعتماد هذا التقدير على صحة أو عدم صحة H_o لأن كل تباين S_i^2 محسوب بطريقة مستقلة عن الآخرين وبافتراض صحة H_o فيعني أن العينات مأخوذة من مجتمع واحد، ونعلم تباين المتوسطات مأخوذة من مجتمع تباينه σ^2 ويساوي σ^2 / n وتقديره:

$$S_{\bar{x}}^2 = \frac{\sum_{i=1}^{k} \left(\bar{x}_i - \bar{\bar{x}}\right)^2}{k-1}$$

حيث يمثل $\bar{\bar{x}}$ المتوسط الكلي للمتوسطات وعليه فيمكن تقدير σ^2 بضرب S_{x}^2 بالحجم المشترك للعينات n' أي:

$$\hat{\sigma}^2 = n' S_{\bar{x}}^2 \quad \text{ويرمز له بالرمز} \quad S_{\beta}^2$$

وهذا يمثل التباين بين المجموعات (Between Group) وهذان التقديران للتباين المشترك σ^2 أحدهم لا يعتمد على صحة أو عدم صحة H_0 في حين الآخر يجب صحة H_0 أي أن جميع العينات المأخوذة يجب أن تكون من نفس المجتمع فتطابق التقديرين يعني صحة H_0 وإلا تعارضت البيانات مع H_0 ويجب أن نعلم أن اختلاف حجم العينات يجعل قيمة التقدير الأول S_w^2 كالآتي:

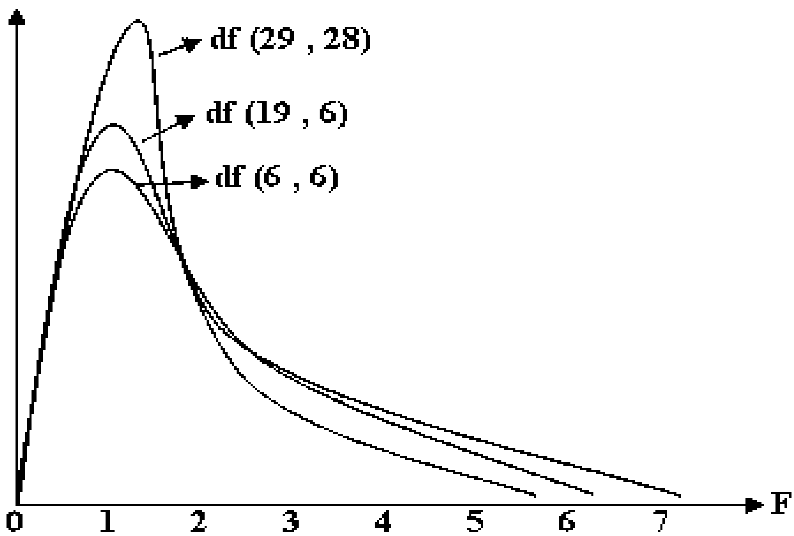
$$S_w^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2 + \ldots + (n_k - 1) S_k^2}{n_1 + n_2 + \ldots + n_k - k}$$

وهو امتداد لتقدير المجتمع واستخدم للاستدلال الإحصائي لمتوسطين حال تساوي تباين المجتمعين ويكون التقدير الثاني كالآتي:

$$S_B^2 = \frac{\sum_{i=1}^{k} n_i(\overline{X}_i - \overline{\overline{X}})^2}{k-1}$$

ويمكن استخدم الصيغ التالية:

النسبة بين التقديرين S_B^2 ، S_W^2 <u>تعرف بتوزيع F</u> وهو توزيع ملتو جهة اليمين بمعلمتين تتمثلان بدرجتي حرية (البسط ، المقام) وهما $k-1$ للبسط ، $n-k$ للمقام حيث n مجموع إحجام العينات، فإذا كان لدينا اختبار لقياس معنوية الفرق بين التقديرين (F) نوجد F_α حيث α مستوى المعنوية المستخدم للفرضية H_0 التي ترفض إذا كان $F > F_\alpha$ وإلا نؤكد بوجود الاختلاف بين المتوسطات، والشكل التالي يبين توزيع F.



منحنى توزيع F حسب درجات الحرية

<u>جدول تحليل التباين</u>

إذا أردنا إجراء اختبار فروض بين متوسطات عددها k من العينات العشوائية المستقلة وبفرض n_1 عدد مفردات العينة الأولى ، n_2 عدد مفردات العينة الثانية ، ...، n_k عدد مفردات العينة k وأن X_{ji} للقيمة المفردة الموجودة i في العينة j وسنضع ذلك في الجدول التالي لبيانات العينات في تحليل التباين:

Sample k	Sample ...	Sample 2	Sample 1	المشاهدات العينة
X_{k1} X_{k2} : X_{knk}	: : : :	X_{21} X_{22} : X_{2n2}	X_{11} X_{12} : X_{1n1}	1 2 : N
T_k		T_2	T_1	العينة مجموع مفردات
$N = n_1 + n_2 + ... + n_k$				N = عدد المشاهدات الكلى
$T = T_1 + T_2 + ... + T_k$				المجموع الكلى (العام)

من الجدول يتبين لنا:

1) الانحراف بين قيمة المشاهدة والمتوسط الحسابي العام وهو الانحراف الكلي ويرمز له ($X_{ji}..$ – `X) حيث i = 1، 2، ....، k (للعينات) ، j = 1، 2، ....، n (للمشاهدات).

2) الاختلاف بين المتوسط الحسابي بكل عينة والمتوسط الحسابي العام ($X_{ji}..$ – `X) وهو الانحراف بين العينات ويرجع لأسباب عشوائية حال عدم الاختلاف للعامل المؤثربين العينات وإلا فالاختلاف يرجع للأسباب العشوائية بجانب تأثير هذا العامل.

3) الاختلاف بين قيمة كل مشاهدة داخل العينة والمتوسط الحسابي (X_{ji} – `Xi) وهو الانحراف داخل العينات ويرجع هذا الاختلاف لأسباب عشوائية بحتة.

بناء على ما سبق يمكن النظر للجدول التالي (جدول تحليل التباين) الذي يبين الخطوات اللازمة لحساب F (قيمة إحصائية الاختبار) حيث K عدد مستويات المتغير المستقل:

F (الجدولية) Tabulated (Sig.)	F (المحسوبة) Calculated	متوسط مجموع المربعات أو التباين Mean squares (MS)	درجات الحرية df	مجموع المربعات Sum of squares (SS)	مصدر التباين Source of Variance
$F_{\alpha\,(K-1),\,(N-K)}$	S_B^2 / S_W^2	S_B^2	K – 1	SS_B	بين المجموعات Between Groups
		S_W^2	N – K	SS_W	داخل المجموعات Within Groups (Error)
			N – 1	$SS_T = SS_B + SS_W$	المجموع Total

مثال:

فـي دراسـة لتـأثير وجـود الطـلاب فـي الصـفوف علـى تحصـيلهم فـي مـادة الإحصـاء، قام أستاذ الإحصـاء بأخذ عينـات عشوائية ومستقلة من ثلاثة صفوف (يقوم بتدريسـها) كل منهـا مكـون مـن خمسـة طـلاب وقـام الأسـتاذ برصـد درجـاتهم والجدول التـالي يبينهـا. بمسـتوى معنويـة $\alpha = 0.05$ اختبـر مـا إذا كـان متوسـط النتائج في اختبارات الأداء يختلف في تحصيل الطلاب.

Class 1	Class 2	Class 3
66	96	58
65	87	62
88	66	77
92	55	90
60	78	80

الحل الاختبـار: $H_0 : \mu_1 = \mu_2 = \mu_3$ ، متوسطان علـى الأقل غيـر متساويين : H_1

نستكمل الجدول كالآتي:

Class 3		Class 2		Class 1	
X_3^2	X_3	X_2^2	X_2	X_1^2	X_1
3364	58	9216	96	4356	66
3844	62	7569	87	4225	65
5929	77	4356	66	7744	88
8100	90	3025	55	8464	92
6400	80	6084	78	3600	60
27637	**$T_3 = 367$, $T_3^2 = 134589$**	**30250**	**$T_2 = 382$, $T_2^2 = 145924$**	**28389**	**$T_1 = 371$, $T_1^2 = 137641$**

$T = T_1 + T_2 + T_3 = 371 + 382 + 367 = 1120$, $T^2 = 1254400$, $n_1 = n_2 = n_3 = 5$, $N = 15$

$$SSB = \sum_{k=1}^{K}\left(\frac{T_k^2}{n_k}\right) - \frac{T^2}{N} \quad , \quad SSW = \sum_{k=1}^{K}\left(\sum_{i=1}^{n_k} X_{ki}^2\right) - \sum_{k=1}^{K}\frac{T_k^2}{n_k}$$

$$S_B^2 = \frac{SSB}{K-1} \quad , \quad S_W^2 = \frac{SSW}{N-K}$$

$SSB = 137641 / 5 + 145924 / 5 + 134689 / 5 – 1254400 / 15$

$= 418254 / 5 – 1254400 / 15$

$= 83650.8 – 83626.7$

$= 24.1$

$SSW = \sum X_1^2 + \sum X_2^2 + \sum X_3^2 – 83650.5$

$= 28389 + 30250 + 27637 – 83650.5$

$= 86276 – 83650.5$

$= 2625.5$

$S_B^2 = 24.1 / (3 – 1) = 12.1$

$S_W^2 = 2625.5 / (15 – 3) = 218.8$

$F = S_B^2 / S_W^2$

$F = 12.05 / 218.8$

$F = 0.055 < 3.89 = F_{\alpha(2 , 12)}$

جدول النتائج

F (الجدولية) Tabulated	F (المحسوبة) Calculated	متوسط مجموع المربعات أو التباين Mean squares (MS)	درجات الحرية df	مجموع المربعات Sum of squares (SS)	مصدر التباين Source of Variance
$F_{\alpha\,(K-1),\ (N-K)}$ 3.89	S_B^2 / S_W^2 12.05/218.5 0.055	$S_B^2 = 24.1/2 = 12.05$	$K – 1 = 3 – 1 = 2$	$SS_B = 24.1$	بين المجموعات Between Groups
		$S_W^2 = 2625.5/12 = 218.8$	$N – K = 15 – 3 = 12$	$SS_W = 2625.5$	داخل المجموعات Within Groups (Error)
			$N – 1 = 15 – 1 = 14$	$SS_T = SS_B + SS_W = 2649.6$	المجموع Total

إن قيمة F المحسوبة أقل من قيمة F الجدولية ولذا نقبل الفرضية الصفرية عند $\alpha = 0.05$ بعدم وجود اختلاف بين المتوسطات.

تحليل التباين في اتجاهين

يهدف هذا النوع من التحليل لنوعين من التأثير (أساسي ومشترك) فالتأثير الأساسي للمتغير المستقل يقصد به تأثيره على المتغير التابع بصرف النظر عن مستويات المتغير المستقل الثاني، في حين التأثير المشترك للمتغيرين وهو يعرف أيضاً بالتأثير التفاعلي فهو ناتج عن تفاعل المتغيرين المستقلين ويكون هناك تفاعل بينهم حال اختلاف تأثير أي منهم يختلف باختلاف مستويات الآخر.

لقياس هذا التحليل سنتبع نفس الطريقة في تحليل التباين الأحادي من خلال تقدير التباين في المجتمع بأربع تقديرات معتمدة على أربع مكونات للتباين عبارة عن مجموع المربعات الخاصة بها بتقسيم المجموع الكلي للمربعات إلى كل من تأثير العامل الأول والثاني . العاملين . (SS_W ، SS_B) والتفاعلي والتباين داخل المجموعات أي المجموع الكلي إلى SS_B ، SS_W ، مجموع المربعات بين المجموعات لكل من مجموع مربعات المتغير المستقل الأول والثاني ومجموع المربعات للتفاعل بين المتغيرين المستقلين والجدول التالي يبين تفاصيل ذلك مع ملاحظة أن $SS_T = {}^2S_B + {}^2S_W$ ، ${}^2S_B = SS_A + SS_B + SS_{AB}$

جدول تحليل التباين في اتجاهين Two-Ways Analysis of Variance

F	MS	d f	SS	Source Variance
MS_A / MS_W	MS_A	$n_A - 1$	SS_A	العامل الأول A
MS_B / MS_W	MS_B	$n_B - 1$	SS_B	العامل الثاني B
MS_{AB} / MS_W	MS_{AB}	$(n_A - 1)(n_B - 1)$	SS_{AB}	التفاعل بين المتغيرين
	MS_W	$n - n_A n_B$	SS_W	داخل المجموعات
		$n - 1$	SS_T	المجموع

أو:

مجموع مربعات التباين الكلي يتضمن المركبات الأربع التالية وهي مستقلة فيما بينها وتخضع للتوزيع المعتدل بمتوسط صفر وتباين قدره 1 والمركبات الأربع هي:

$$SST = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n} X_{ijk}^2 - \frac{T_{...}^2}{r\,c\,n}$$

$$SSR = \frac{\sum_{i=1}^{r} T_{i..}^2}{c\,n} - \frac{T_{...}^2}{r\,c\,n}$$

$$SSC = \frac{\sum_{J=1}^{c} T_{.j.}^2}{r\,n} - \frac{T_{...}^2}{r\,c\,n}$$

$$SS(RC) = \frac{\sum_{i=1}^{r}\sum_{J=1}^{c} T_{ij.}^2}{n} - \frac{\sum_{i=1}^{r} T_{i..}^2}{c\,n} - \frac{\sum_{J=1}^{c} T_{.j.}^2}{r\,n} + \frac{T_{...}^2}{r\,c\,n}$$

1) A مركبة بين الأعمدة
2) B مركبة بين الصفوف
3) B, A مركبة التفاعل بين الأعمدة والصفوف
4) e_{ijk} مركبة الخطأ

حيث أن:

i يشير لعدد الصفوف $i = 1, 2, 3, \ldots, r$ و r عدد الصفوف

j يشير لعدد الصفوف $j = 1, 2, 3, \ldots, c$ و c عدد الأعمدة

k يشير لعدد الصفوف $k = 1, 2, 3, \ldots, n$ و n عدد المشاهدات في الخلية

نستخدم تحليل التباين هذا مع الفرضيات الآتية واستخدام اختبار F:

1) تأثير جميع مستويات العامل A متساوٍ
1) تأثير جميع مستويات العاملB متساوٍ
1) عدم وجود تفاعل بين مستويات العاملين A , B

نستخدم الصيغ الآتية لحسابات تحليل التباين الثنائي مع وجود التفاعل الداخلي:

$$SSE = SST - SSR - SSC - SS(RC)$$

مع ملاحظة أن:

SST مجموع المربعات الكلي وهو: تربيع كل درجة خام في كل خلية في الجدول ثم نجمع كل المربعات الناتجة (تربيع كل مشاهدة ثم الجمع) ، N عدد الدرجات أو المشاهدات،

ثم نعوض في المعادلة أعلاه للحصول على المطلوب.

SSR مجموع مربعات الصفوف وهو: إن كان لدينا مجموعتان للمقارنة مع ثلاث مستويات مثلاً فلكل مجموعة مقارنة ثلاثة أعمدة(عدد المستويات) لدينا ثلاثة مجاميع للأعمدة الثلاثة بجمعها نحصل على مجموع الصف للمجموعة الأولى وبالمثل نحصل على مجموع الصف (مجاميع الأعمدة للمجموعة الثانية) وبالتالي يكون لدينا قيمتان مجموع الصف الأول للممجوعة الأولى ومجموع الصف الثاني للمجموعة الثانية بتربيع كل منهم وجمع المربعين لنحصل على مجموع مربعات الصفوف (مع ملاحظة الجدول هنا لأربع مستويات أي أربع أعمدة في كل مجموعة)،

ثم نعوض في المعادلة أعلاه للحصول على المطلوب

SSC مجموع مربعات الأعمدة لدينا مجموعتان للمقارنة وأربعة مستويات يعني وجود 4 × 2 = 8 أعمدة أربعة لكل مجموعة وبجمع كل مستوى للمجموعتين نحصل على مجموع الأعمدة الأربع في الصف الأخير. فنربع كل منها (4 قيم) ثم نجمع المربعات الناتجة، ثم نعوض في المعادلة أعلاه للحصول على المطلوب.

SS (RC) مجموع مربعات الصفوف والأعمدة وهي 8 مجاميع المبينة في الجدول نربع كل منها ونجمع النواتج، ثم نعوض في المعادلة أعلاه للحصول على المطلوب.

SSE في الجدول هنا 8 مجاميع نربع كل منه ونجمع النواتج، ثم نعوض في المعادلة مجموع المربعات الكلي – المجموع الناتج ÷ عدد الدرجات في العمود (للمجموعة الواحدة) هنا يساوي 4 أو من:

i(SS(RC − SSC − SSR − SSE = SST

درجات الحرية لتباين الصفوف = عددها – 1

درجات الحرية لتباين الأعمدة = عددها – 1

درجات الحرية لتباين التفاعل = (عدد الصفوف – 1)(عدد الأعمدة – 1)

درجات الحرية لتباين الخطأ = عدد الصفوف × عدد الأعمدة × (عدد درجات المجموعة الواحدة – 1) ، في جدولنا هنا توجد 8 مجاميع كل منها يضم 4 درجات

<u>جدول تحليل التباين</u>

F		MS	d f	SS	Source Variance
$F_1 =$	$\frac{S_1^2}{S_3^2}$	$S_1^2 = \frac{SSR}{r-1}$	$r-1$	SSR	بين الصفوف
$F_2 =$	$\frac{S_2^2}{S_3^2}$	$S_2^2 = \frac{SSC}{c-1}$	$c-1$	SSC	بين الأعمدة
$F_3 =$	$\frac{{}_3^2}{S_3^2}$	$S_3^2 = \frac{SS(RC)}{(r-1)(c-1)}$	$(r-1)(c-1)$	SS(RC)	التفاعل
		$S_4^2 = \frac{SSE}{rc(r-1)}$	$rc(n-1)$	SSE	الخطأ
		$rcn-1$		SST	الكلي المجموع

مثال:

في تقيم لمعرفة تأثير كل من المستوى والفترة الزمنية على أداء الطلاب في مادة الرياضيات فجرى اختيار أربع مستويات أخذ من كل منها 8 طلاب نصفهم للفترة الصباحية والنصف الآخر للفترة المسائية وبصورة عشوائية والمطلوب تحليل التباين ذي اتجاهين لمعرفة الإجابة على التأثير المطلوب حيث كانت نتائجهم في الامتحان النهائي مبينة في الجدول الآتي:

Levels				Time
L4	L3	L2	L1	
80	77	85	60	A M
85	85	80	75	
86	67	79	80	
77	90	66	67	
85	88	78	90	P M
66	80	59	90	
67	78	67	88	
84	77	85	77	

الحل : نقوم هنا بإجراء ثلاثة اختبارات وهي التأثير الأساسي للعامل الأول (المستوى) والثاني للفترة الزمنية والثالث للتأثير التفاعلي كما يلي:

1) التأثير الأساسي للعامل الأول (المستوى) نرمز له بالرمز A وفرضيتي العدم والبديل هما:

$H_0 : \mu_{A1}=\mu_{A2}=\mu_{A3}=\mu_{A4}$, H_1 : غير متوسطين على الأقل متساويين

2) التأثير الأساسي للفترة الزمنية ولنرمز له بالرمز B وفرضيتي العدم والبديل هما:

$H_0 : \mu_{B1}=\mu_{B2}$, $H_1 : \mu_{B1}\neq\mu_{B2}$

3) التأثير التفاعلي وهو: H_0 : ليس هناك تأثير تفاعلي بين A , B على المتغير التابع ، H_1 : هناك تأثير تفاعلي بين A , B على المتغير التابع

نكون الجدول الآتي:

				Levels					Time
L_4^2	L_3^2	L_2^2	L_1^2	Total	L_4	L_3	L_2	L_1	
6400	5929	7225	3600	302	80	77	85	60	A M
7225	7225	6400	5625	325	85	85	80	75	
7396	4489	6241	6400	312	86	67	79	80	
5929	8100	4356	4489	300	77	90	66	67	
				1239	328	319	310	282	∑(AM)
7225	7744	6084	8100	341	85	88	78	90	P M
4356	6400	3481	8100	295	66	80	59	90	
4489	6084	4489	7744	300	67	78	67	88	
7056	5929	7225	5929	323	84	77	85	77	
				1259	302	323	289	345	∑(PM)
50076	51900	45501	49987	2498	630	642	599	627	Total

باستخدام الصيغ:

$$SST = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{n} X_{ijk}^2 - \frac{T_{...}^2}{r\,c\,n}$$

$$SSR = \frac{\sum_{i=1}^{r} T_{i..}^2}{c\,n} - \frac{T_{...}^2}{r\,c\,n}$$

$$SSC = \frac{\sum_{J=1}^{c} T_{.j.}^2}{r\,n} - \frac{T_{...}^2}{r\,c\,n}$$

$$SS(RC) = \frac{\sum_{i=1}^{r}\sum_{J=1}^{c} T_{ij.}^2}{n} - \frac{\sum_{i=1}^{r} T_{i..}^2}{c\,n} - \frac{\sum_{J=1}^{c} T_{.j.}^2}{r\,n} + \frac{T_{...}^2}{r\,c\,n}$$

$$SST = (60)^2 + (75)^2 + ... + (67)^2 + (84)^2 - [(2498)^2] / [2\times4\times4]$$
$$= 197464 - 195000.13 \qquad = \mathbf{2463.88}$$

$$SSR = [(1239)^2 + (1259)^2] / [16] - [(2498)^2] / [2\times4\times4]$$
$$= 195012.63 - 195000.13 \qquad = \mathbf{12.50}$$

$$SSC = [(627)^2 + (599)^2 + (642)^2 + (630)^2] / [\ 8\ ] - [(2498)^2] / [2\times4\times4]$$
$$= 195124.250 - 195000.125 \qquad = \mathbf{124.13}$$

$$SS(RC) = [(282)^2 +...+ (302)^2] / 4 - 195012.63 - 195124.250 + 195000.13$$
$$= 195762 - 195136.75 \qquad \mathbf{= 625.25}$$

$$SSE = SST - SSR - SSC - SS(RC)$$
$$SSE = 2463.88 - 12.50 - 124.13 - 625.25 = \mathbf{1702}$$

<u>من جدول f</u> نجد أن:

$F_1 > f_{0.05 ، 1 ، 24}$ = 26.4 و $F_2 > f_{0.05 ، 3 ، 24}$ = 3.01 و $F_3 > f_{0.05 ، 3 ، 24}$ = 3.01

جدول تحليل التباين

F من الجدول	**F**	**MS**	**d f**	**SS**	**Source Variance**
4.26	**$F_1 = 12.5 \div 70.92 = 0.18$**	**$S_1^2 = 12.5 / 1 = 12.5$**	**2 – 1 =1**	**12.50**	**بين الصفوف**
3.01	**$F_2 = 41.4 \div 70.92 = 0.58$**	**$S_2^2 = 124.13 / 3 = 41.4$**	**4– 1 = 3**	**124.13**	**بين الأعمدة**
3.01	**$F_3 = 208.42 \div 70.92 = 2.94$**	**$S_3^2 = 625.25 / 3 = 208.42$**	**(2 – 1)(4 – 1) = 3**	**625.25**	**التفاعل**
		$S_4^2 = 1702 / 24 = 70.92$	**(2×4) (4 – 1) = 24**	**1702.00**	**الخطأ**

		2×4×4 – 1 = 31	**246388**	**المجموع الكلي**

بناء على قيم F نجد أن:

1) قيمة F المحتسبة أقل من قيم F الجدولية (0.18 > 4.26) فنقبل فرضية العدم أي عدم وجود اختلاف

2) قيمة F المحتسبة أقل من قيم F الجدولية (0.58 > 3.01) فنقبل فرضية العدم أي عدم وجود اختلاف

3) قيمة F المحتسبة أقل من قيم F الجدولية (2.94 > 3.01) فنقبل فرضية العدم أي عدم وجود تفاعل بين العاملين

<u>تحليل التباين بمعيارين مع عدم وجود تفاعل داخلي</u>

<u>Analysis of Variance without internal interaction Two-Ways</u>

المقصود بعدم وجود تفاعل بين المتغيرين (معياري التصنيف) هو لا يتفاعلان ببقاء تأثير الأعمدة (r) هو نفسه لكل نوع أو صنف والصيغ المستخدمة هنا هي:

$$SST = \sum_{i=1}^{r}\sum_{j=1}^{c} X_{ij}^2 - \frac{T_{..}^2}{r\,c}$$

$$SSR = \frac{\sum_{i=1}^{r} T_{i.}^2}{c} - \frac{T_{..}^2}{r\,c}$$

$$SSC = \frac{\sum_{J=1}^{c} T_{.J}^2}{r} - \frac{T_{..}^2}{r\,c}$$

والصيغ المستخدمة هي: SSE = SST – SSR – SSC

Source Variance	SS	d f	MS	F
بين الصفوف	SSR	**r – 1**	**S_1^2 =SSR / (r – 1)**	$F_1 = S_1^2 / S_3^2$
بين الأعمدة	SSC	**c – 1**	**S_2^2 =SSC / (c – 1)**	$F_2 = S_2^2 / S_3^2$
الخطأ	SSE	**(r –1) (c – 1)**	**S_3^2 =SSE / (r – 1)(c – 1)**	

المجموع	SST	**r c – 1**		

مثال:

من أجل تحسين إنتاج مدرس الرياضيات في استخدام العمليات الحسابية قامت أحدى المدارس بإجراء تجربة لمعرفة تأثير ثلاثة طرق من التدريس أ استخدام الآلة الحاسبة ، ب استخدام الحاسب الآلي ، حـ استخدام القوانين وقدمت أربعة أنواع من الحوافز المادي ، المعنوي ، العقاب ، الثواب. اختبر وجود فروق جوهرية بين متوسطات الإنتاج للمدرس وكذلك بين متوسطات الإنتاج واختلاف نوع الحافز عند مستوى معنوية 0.05 من خلال جدول الإنتاج للمدرسين مع الحوافز الآتي:

الحوافز	الآلة الحاسبة	الحاسب الآلى	القوانين
مادي	60	75	70
معنوي	69	80	60
عقاب	75	77	55
ثواب	58	69	57

الحل:

الفرضيات:

(1) انعدام تأثيرات الصف $H_0 : a_1 = a_2 = a_3 = a_4 = 0$

انعدام تأثيرات العمود $H_0 : b_1 = b_2 = b_3 = 0$

(2) على اقل أحداها لا يساوي الصفر : H_0

على اقل أحداها لا يساوي الصفر : H_1

بتطبيق الصيغ الآتية:

$$SST = \sum_{i=1}^{r}\sum_{j=1}^{c} X_{ij}^2 - \frac{T_{..}^2}{r\,c}$$

$$SSR = \frac{\sum_{i=1}^{r} T_{i.}^2}{c} - \frac{T_{..}^2}{r\,c}$$

$$SSC = \frac{\sum_{J=1}^{c} T_{.J}^2}{r} - \frac{T_{..}^2}{r\,c}$$

$SST = [(60)^2 + (69)^2 + ... + (55)^2 + (57)^2] – (805)^2 / 4×3$

$= 54839 – (805)^2 / 4×3$

$= 54839 – 54002.08$

$= 836.92$

$SSR = [(205)^2 + (209)^2 + (207)^2 + (184)^2] / 3 – (805)^2 / 4×3$

$= 54137 – 54002.08 = \quad = 134.92$

$SSC = [(262)^2 + (301)^2 + (242)^2] / 4 – (805)^2 / 4×3$

$= 54452.25 – 54002.08 = \quad = 450.17$

$SSE = SST – SSR – SSC \qquad SSE = 836.92 – 134.92 – 450.17$

$= 251.83$

وجدول تحليل التباين هو:

Source Variance	SS	d f	MS	F
بين الصفوف	134.92	4 – 1 = 3	S_1^2 =134.92/ 3 = 44.97	F_1 = 44.97 / 41.97 = 1.071
بين الأعمدة	450.17	3 – 1 = 2	S_2^2 =450.17 / 2 = 225.08	F_2 = 225.08 / 41.97 = 5.36
الخطأ	251.83	(4 – 1)(3 – 1) = 6	S_3^2 =251.83 / 6 = 41.97	
المجموع	836.92	4 × 3 – 1 = 11		

عند مستوى معنوية $\alpha = 0.05$ نجد أن:

$F_1 < f_{0.05,3,6} = 4.76$ قبول الفرضية الصفرية أي لا يوجد اختلاف في متوسط إنتاجية المدرسين عندما تختلف الحوافز

$F_2 > f_{0.05,2,6} = 5.14$ عدم قبول الفرضية الصفرية أي يوجد اختلاف في متوسط إنتاجية المدرسين عندما تختلف الحوافز

تحليل التباين في اتجاه واحد عند عدم تساوي حجم العينات

بنفس الطريقة السابقة (تساوي حجم العينات) حسب الصيغ الآتية:

مثال 1:

يمثل الجدول التالي نتائج طلبة تم تدريسهم مادة الإحصاء بأربعة طرق تدريس مختلفة أختبر الفرضية القائلة بتساوي متوسطات الطلبة وفق الاختبارات الأربعة عند مستوى معنوية 0.10

M4	M3	M2	M1
94	59	75	65
89	78	69	87
80	67	83	73
88	62	81	79
90	83	72	81
	76	79	96
		90	

الحل:

نضيف للجدول بعض القيم الأخرى لنحصل على

M4	M3	M2	M1
94	59	75	65
89	78	69	87
80	67	83	73

88	62	81	79
90	83	72	81
	76	79	96
		90	
441	425	549	481

الفرضية: $H_o : \mu_1 = \mu_2 = \mu$ ، $H_1 : \mu_1 \neq \mu_2 \neq \mu$

$$1)\ \sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij} = 65 + 87 + \ldots + 90 = 1896$$

$$2)\ \sum^{k}\sum^{n} x^2 = (65)^2 + (87)^2 + \ldots + (90)^2 = 152066$$

$$3)\ \sum^{k} \frac{(\sum^{n} x)^2}{n_i} = \frac{(481)^2}{6} + \frac{(549)^2}{7} + \frac{(425)^2}{6} + \frac{(441)^2}{5} = 150617.8$$

$$4)\ \frac{(\sum^{k}\sum^{n} x)^2}{\sum n_i} = \frac{(1896)^2}{24} = 149784$$

$$5)\ SST = \sum_{i=1}^{k}\sum_{j=1}^{n} X_{ij}^2 - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij})^2}{\sum n_i} = 152066 - 149784 = 2282$$

$$6)\ SSB = \sum^{k} \frac{(\sum^{n} x)^2}{n_i} - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij})^2}{\sum n_i} = 150617.8 - 149784 = 833.819$$

$$7)\ SSW = SST - SSB = 2282 - 833.819 = 1448.181$$

نوجد كل مما يأتي حسب الترتيب:

1) مجموع العناصر

2) مجموع مربعات العناصر

3) مجموع مربعات العناصر مقسوماً على حجم العينة

4)معامل التصحيح

5) مجموع مربعات التباين (SST)

6) مربعات التباين بين المجموعات (SSB)

7) مربعات التباين ضمن المجموعات (SSW)

وحسب الخطوات السبعة أعلاه نجد الآتي:

$$1)\ \sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}$$

$$1)\ \sum^{k}\sum^{n} x^{2}$$

$$3)\ \sum^{k}\frac{(\sum^{n} x)^{2}}{n_i}$$

$$4)\ \frac{(\sum^{k}\sum^{n} x)^{2}}{\sum n_i}$$

$$5)\ SST=\sum_{i=1}^{k}\sum_{j=1}^{n} X_{ij}^{2}-\frac{\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij})^{2}}{\sum n_i}$$

$$6)\ SSB=\sum^{k}\frac{(\sum^{n} x)^{2}}{n_i}-\frac{(\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij})^{2}}{\sum n_i}$$

$$7)\ SSW = SST - SSB$$

وتكون نتائج التحليل كالآتي:

F (الجدولية) Tabulated (Sig.)	F (المحسوبة) Calculated	متوسط مجموع المربعات أو التباين Mean squares (MS)	درجات الحرية df	مجموع المربعات Sum of squares (SS)	مصدر التباين Source of Variance
$F_{\alpha\,(K-1),\,(N-K)}$ $F_{\alpha(3,20)}=$ 2.38	S_B^2 / S_W^2 277.940/72.409 = 3.839	$S_B^2 =$ 833.819/3 = 277.940	K – 1 = 4 – 1 = 3	$SS_B =$ 833.819	بين المجموعات Between Groups
		$S_W^2 =$ 1448.181/20 = 72.409	N – K = 24 – 4 = 20	$SS_W =$ 1448.181	داخل المجموعات Within Groups (Error)
			N – 1 = 24 – 1 = 23	$SS_T = SS_B + SS_W$ = 2282	المجموع Total

وعند مستوى معنوية 0.10 نجد أن F الجدولية ($F_{\alpha(3,20)}$ = 2.38) أقل من F المحتسبة فنرفض H_0 مما يعني وجود فروق بين المتوسطات.

المقارنات المتعددة Multiple Comparisons

هي طريقة لإجراء عدد من الاختبارات الأولية لتحديد الفروق المعنوية للمتوسطات حال رفض فرضية العدم.

حال رفض فرضية العدم عند مستوى معنوية فلا دليل على وجود فروق معنوية بين المتوسطات ولا نعرف أي منها يختلف عن متوسطات (العينات) ومن حيث قيمة F معنوية ففرق المشاهدات بين المتوسطات معنوي أيضاً وعلى العموم توجد عدة طرق إحصائية لعمل اختبار بهذا الخصوص مثل:

Scheffe Test, Student-Newman Keuls, Tukey's Procdure, Least Significant difference, Duncans New Multiple range

سنتعرض كل منها بصورة مبسطة مع مثال.

طريقة difference Least Significant للمقارنة بين متوسطين(L.S.D)

يرمز لها LSD وتحدد حال استخدام برنامج SPSS وتعرف Fisher's LSD وتختبر كل الأزواج فإن تمَ رفض فرضية العدم (تساوي المتوسطات) مقابل الفرضية البديلة (عدم تساوي المتوسطات) فإذا كان الفرق أكبر من LSD فنرفض فرض العدم، والصيغة الرياضية لحساب LSD حيث MSE متوسط مجموع المربعات (التباين) داخل المجموعات) هي:

A	B	C	D
8	3	5	6
7	4	3	5
9	6	4	6
5	5	5	4
6	2	3	3
7	7	2	4
42	27	22	28

جدول تحليل التباين

Source of Variance	SS	d f	MS	F	$f_{0.05,3,20}$
SSB	SST – SSW = 36.791	K – 1 = 4 – 1 = 3	12.264	5.818	3.10
SSW	42.167	K(n – 1) = 4(6 – 1) = 20	2.108		
TOTAL	78.958	Kn – 1 = 24 – 1 = 23			

قيمة F المحسوبة أكبر من F الجدولية (7.64 < 3.1) فالمتوسطات غير متساوية، وبإجراء اختبار LSD نجد أن

$$\mathbf{LSD = t_{2n-2,\alpha/2}\sqrt{MSE \times \frac{2}{n}}}$$

$MSE \times 2 / n = 3 \times 2 \div 6 = 1$, $t_{2n-2,\alpha/2} = t_{10,\ 0.025} = 2.228$

$LSD = 2.228 \times 1$

$= 2.228$

جدول متوسطات الفروق حيث وجود *على العدد يعني وجود فرق معنوي بين المتوسطين ولدينا أربع متوسطات (7، 4.5، 3.7، 4.7) تقارن كل منها بقيمة 2.228 = LSD

$\bar{X}_4 = 4.667$	$\bar{X}_3 = 3.667$	$\bar{X}_2 = 4.5$	$\bar{X}_1 = 7$	
2.333^*	3.333^*	2.5^*	0	$\bar{X}_1 = 7$
– 0.167	0.833	0	–	$\bar{X}_2 = 4.5$
– 1	0	–	–	$\bar{X}_3 = 3.667$
0	–	–	–	$\bar{X}_4 = 4.667$

تذكـر

يقصـد بـالاختلاف الكلـى (T S S) المجمـوع الكلـى لمربعـات الانحرافـات ($TSS= \Sigma$ Yi- Y)2) وهذا المقدار يمكـن تقسـيمه إلـي جـزأين. أحـدهما هـو (SSE) أو مجموعـة مربعـات انحرافـات قـيم (Y) حـول خـط الانحـدار المقـدر. والجزء الثاني يمثل النقص فى مجموع مربعات الانحرافات الكلى الناجم عن إدخال التغير (x) فى هذه العلاقة أو ما يسمى بـ (SSR) أو مجموع مربعات الانحرافات الناجمة عن الانحدار، وبالنسبة لكل ملاحظة فان.

$$(Yi- Y) = (Yi- \hat{Y}i)+ (\hat{Y}i-Y).$$

مجموع مربعات الانحرافات الناجمة عن الانحدار = SSR أو كمية الانحرافات الكلية التى يشرحها المتغير (X)

$$SSE= \Sigma_{i=1}^{n} (\hat{Y}i-\hat{Y}i)^2$$

مجموع مربعات الخطأ (SSE) أو كمية الانحرافات الكلية التى لا يشرحها المتغير (X) وتسلك سلوكا عشوائيا لا يمكن التنبؤ به.

$$r^2= \frac{\Sigma (X- X) (Y- Y)}{\Sigma (Y-Y)^2} * \frac{\Sigma (X-X) (Y-Y)}{\Sigma(X-X)^2}$$

وهـذا يبـين العلاقـة بـين معـاملات الانحـدار ومعامـل الارتبـاط ومعامـل التحديد.

مجموع مربعات للانحدار هو عبارة عن مقياس التشتت للقيم الفعلية حول خط الانحدار ويعرف بمجموعات مربعات البواقي.

$SSE = \sum (Y - \bar{Y})^2$ وهي مجموع مربعات الانحرافات لقيم Y الفعلية عن قيم Y المقدرة.

تباين الخطأ العشوائي:هو تقدير تباين خط انحدار Y على x ويحسب من الصيغة الرياضية($S^2_{Y/x}$) ويساوي:

$$S^2_{Y/x} = \frac{SSE}{(n-2)} = \frac{\sum (Y - \bar{Y})^2}{(n-2)s}$$

الخطأ المعياري لتقدير خط انحدار Y على x هو قيمة الجذر التربيعي لهذا التقدير.

مجموع الانحرافات = الانحراف المفسر + الانحراف غير المفسر

$$SSE = SST - SSR$$

خطأ التقدير هو الفرق بين القيم المتنبئة y' والقيم الأصلية ويقل الفرق بقوة العلاقة بين المتغيرين ويزداد كلما ضعفت العلاقة الارتباطين بين المتغيرين ، والخطأ المعياري للتقدير (SEE) يقيس دقة معادلة الانحدار في تقدير العلاقة بين المتغيرين فيصغر SEE كلما اقتربت قوة العلاقة من $1 \pm$.

تماريـن

(1) كون جدول تحليل التباين للانحدار من الجدول الآتـي:

الجدول الآتـي يبين إنتاج محصول الذرة Y من المساحة المزروعة به X . اختبر معنوية معامل الانحدار عند مستوى معنوية 0.05

المنطقة	X المساحة المزروعة بالهكتار	Y إنتاج الذرة بالآف الكيلوجرام
1	50	140
2	200	500
3	110	400
4	80	300
5	120	356
6	74.5	240.5
7	88.9	200.6
8	5.7	33.5
9	11	69.8
10	3.2	18.7

(2) احسب الخطأ المعياري للتقدير لقيم y من قيم x من البيانات التالية:

x	10	12	15	11	9	14
y	2	3	5	6	1	4

(3) إذا علمت أن البيانات التالية للمتغيرين x , y

$\bar{x} = 12$, $\bar{y} = 4$, $S_x = 2.098$, $S_y = 1.414$, $S_{y'} = 0.477$

أوجد كل من: معامل الارتباط ، الخطأ المعياري للتقدير ، معادلة التنبؤ (نموذج الانحدار) بقيم y من قيمx

الباب السابع

كـا ٢ (كاي تربيع) Chi Square Test

يتنـاول هذا الجزء بعض الأسـاليب التـي تسـتخدم فـي التحليـل الإحصـائي ويتمثل في **كا ٢**

وتستخدم توزيعات **كـا ٢** في اختبـارات صـحة نظريـة معينـة ويسـتخدم ذلك فـي اختبارات الاستقلال بين المتغيرات أو الصفات المختلفة.

يمكننا اختبار مربع–**Chi** بمعرفة فيما إذا كانت القيم الملاحظة للبيانات في جدولنا تختلف بشكل كبير عن تلك المتوقعة أو لا في حال لم يكن هناك أي عوامل تحتية تؤثر على نموذج هذه القيم.

ويمكننا مقارنة قيمنا الجدولية (القيم الملاحظة) مع القيم التي توقعنا الحصول عليها (القيم المتوقعة) إذا لم يكن هناك أية علاقة تحتية.

وتتبع إحصائيات الاختبار توزيعاً يدعى توزيع مربع–**Chi** **(Chi-Squared)** ، أو توزيع χ^2 (يستخدم التربيع للتأكيد على أن نتائج الإحصائيات لا يمكن أن تكون سالبة). ويتعلق هذا التوزيع بالتوزيع الطبيعي، ويعتمد على معامل يدعى "عدد درجات الحرية" **(df)**، وهو معامل يحدد توزيع χ^2 المحدد. ويميل هذا التوزيع دوما إلى اليمين وقيمة المتوسط الخاص به تساوي عدد درجات الحرية. كما يتجه التوزيع للقيم الكبيرة للمعامل **df** باتجاه التوزيع الطبيعي.

وبعد حساب إحصائيات الاختبار، تتم مقارنتها مـع القيمـة الحدية χ^2 التـي تحسب مـن جـداول χ^2 الإحصـائية (يمكـن إيجـاد هـذه الجـداول فـي الكتـب التعليميـة الإحصائية).

<u>خطوات حساب كا ٢</u>

1–نضع الفرض الصفري (العدمي) ويقصد به هنا فرض العدم التغير ويتم وضعه حسب النظرية أو البيانات المراد التحقق منها.

2– حساب التكرارات المتوقعة من خلال النظرية أو التوزيع المطلوب.

3– نحدد مستوي المعنوية α بمعني تحديد القيمة الحرجة التي عندها يمكن رفض الفرض الصفري أو قبوله أي تحديد درجة المخاطرة التي يحددها الباحث لنفسه عند رفض أو قبول فروض بحثه الصفية وقبول الفرض البديل.

4– نستخدم جدول توزيع **كا٢** بدرجات حرية مناسبة Φ حيث ان Φ = ن –1 ، ن عدد الاختيارات المختلفة اما اذا كان التوزيع ياخذ شكل المصفوفة فان :

Φ = (عدد الأعمدة –1) × (عدد الصفوف –1)

5–نحصل علي **كا٢** المحسوبة من العلاقة :

كا٢ المحسوبة = مجـ (التكرار المشاهد – الكرار المتوقع)

التكرار المتوقع

كا٢ المحسوبة = $\dfrac{(F_O - F_e)^2}{F_e}$

6–اذا كان **كا٢**المحسوبة اقل من او يساوي **كا٢**الجدولية (العشوائية) بدرجات حرية Φ ومستوي معنوية α نقبل الفرض الصفري (العدمي) أما إذا كانت **كا٢** المحسوبة اكبر من او تساوي **كا٢** الجدولية فإننا نرفض الفرض الصفري (العدمي).

7–لابد أن نلاحظ أن مستوي الثقة 1 – α = 95% مثلا يعني مستوي الثقة 0.95 مقابل 0.05 عدم ثقة (شك) كما أن **كـا٢**الجدولية يستخدم فيها مستوي المعنوية α دائماً.

مثال : فيما يلي استجابات مجموعة من الطلاب نحو قضية ما

نعم	لا	المجموع
60	20	80

احسب **كا٢** عند مستوي معنوية 5%.

الحـل :

(1) الفرض الصفري أن استجابات الموافقين تماثل غير الموافقين أي لا توجد فروق بين من يوافق ومن لم يوافق.

(2) التكرار المتوقع = 80 ÷ 2 = 40

(3) **كا**^{٢} المحسوبة = $\dfrac{(f_1 - f_2)^2}{f_2}$

حيث f_1 = التكرار المشاهد أو التجريبي

f_2 = التكرار المتوقع

كا^{٢} المحسوبة = $\dfrac{(60 - 40)^2}{40} + \dfrac{(60 - 40)^2}{40}$

$= \dfrac{400}{40} + \dfrac{400}{40}$

(4) درجات الحرية $1 = 2-1 = \Phi$

(5) **كا**^{٢} الجدولية = **كا**^{٢} (1 ، 0.05) = 3.84146

(6) **كا**^{٢} المحسوبة > **كا**^{٢} الجدولية

∴ يرفض الفرض الصفري ويقبل البديل

مثال: فيما يلي استجابات مجموعة من الافراد نحو قضية ما :

موافق	لا ادري	ارفض	مجموع
12	2	16	30

احسب كا^{٢} عند مستوي ثقة 95 %.

الحل

(1) الفرض الصفري H_0 يعني ان الفروق بين استجابات الموافقين والمعارضين ليست فروقا معنوية ذات دلالة حقيقية تؤخذ في الاعتبار عند اتخاذ القرارات اللازمة.

(2) التكرار المتوقع = 30 / 20 = 10

كـا^{٢} المحسوبة = $\dfrac{(12 - 10)^2}{10} + \dfrac{(10 - 2)^2}{10} + \dfrac{(10 - 16)^2}{10}$

$$\text{كا}^2 \text{ المحسوبة} = \frac{4}{10} + \frac{64}{10} + \frac{36}{10} + \frac{104}{10} = 10.4$$

(4) درجات الحرية $\Phi = 3-1 = 2$

(5) **كا**٢ الجدولية = **كا**٢(2 ، 5 %) =5.99147

(6) **كا**٢ المحسوبة > **كا**٢ الجدولية ∴ نرفض الفرض الصفري

بمعني ان هناك فوقا ذات معنوية عند (0.05).

مثال : فيما يلي استجابات عينة من الطلاب نحو موضوع معين

أشارك بدرجة كبيرة	بدرجة متوسطة	بدرجة ضعيفة	محايد
60	20	15	5

أحسب **كا**٢ عند درجة ثقة 99 %.

الحل

(1) الفرض الصفري ان الفروق بين استجابات المشاركين وغير المشاركين ليست فروقا معنوية أي ليست ذات دلالة حقيقية.

(2) التكرار المتوقع = 100 / 4 = 25

(3) كا٢ المحسوبة = $\frac{(f_1 - f_2)^2}{f_2}$

حيث T_1= التكرار المشاهد

T_2= التكرار المتوقع

كا٢ المحسوبة =

$$\frac{(60-25)^2}{25} + \frac{(60-25)2}{25} + \frac{(60-25)^2}{25} + \frac{(60-25)^2}{25}$$

= 49 + 1 + 4 + 16 = 70

(4) درجة الثقة 99 % ∴ مستوي المعنوية 1 % أي = α 1%

(5) درجات الحرية Φ = (4 − 1) = 3

(6) كا^2 الجدولية = **كا^2** (3 ، 0.01) = 3.11.345

(7) ∵ **كا^2** المحسوبة > **كا^2** الجدولية

∴ يرفض الفرض الصفري ويقبل البديل

ملاحظات:

(1) ان هناك جداول عشوائية لحساب **كا^2** الجدولية تستخدم درجات الثقة مباشرة ولا تحتاج فيها الحصول عل مستوي المعنوية.

(2) ان هناك شرطان أساسيان عند استخدام **كا^2** في اختبارات الفروض وهما:

(أ) يجب ان تكون التكرارات الكلية كبيرة بحيث لا تقل عن 50 مشاهد.

(ب) يجب الا يقل تكرار أي خلية عن 5 تكرارات ويمكننا التغلب علي ذلك بإدماج خلايا التكرارات الصغيرة وذلك بدمج أكثر من صف أو أكثر من عمود.

مثال : فيما يلي استجابات عينة من الطلاب نحو استطلاع راي ما :

موافق	موافق الي حد ما	لا ادري	ارفض	ارفض الي حد ما
37	3	20	28	2

احسب **كا^2** عند مستوي معنوية 5 %

الحـل

نظرا لأن استجابات موافق الي حد ما = 3 فيمكن إضافتها إلي خانة موافق فتصبح 40 ، ونفس الكلام بالنسبة إلي خانة ارفض إلي حد ما تضاف إلي خانة ارفض فتصبح = 30

ومن ثم تأخذ الاستجابات الشكل التالي:

موافق	لا ادري	ارفض
40	20	30

(1) التكرار المتوقع = 10 ÷ 3 = 30

(2) كا^2 المحسوبة = $\frac{(40-30)^2}{30} + \frac{(20-30)^2}{30} + \frac{(30-30)^2}{30}$

$$= \frac{100}{30} + \frac{100}{30} + \text{صفر} = 6.7$$

(2) درجات الحرية = 3 − 1 = 2

(3) **كا**٢ الجدولية = **كا**٢ (2 ، 5 %) = 5.99147

(4) ∵ **كا**٢ المحسوبة > **كا**٢ الجدولية

∴ نرفض الفرض الصفري ونقبل البديل

مثال : فيما يلي استجابات عينة من الشباب نحو قضية المشاركة السياسية

	نعم	**لا**	المجموع
ذكور	100	20	120
اناث	65	15	80
المجموع	165	35	200

احسب **كا**٢ عند مستوي معنوية 5 %

الحـل

(1) سنقوم حساب التكرار المتوقع لكل خلية علي حدا كالاتي :

– التكرار المتوقع للذكور (نعم) = $\frac{120 \times 165}{200} = 99$

– **كا**٢ المحسوبة للذكور (نعم) = $\frac{(100-99)^2}{200} = 0.01$

– التكرار المتوقع للذكور (لا) = $\frac{120 \times 35}{200} = 21$

– **كا**٢ المحسوبة للذكور (لا) = $\frac{(20-21)^2}{21} = 0.05$

– التكرار المتوقع للإناث (نعم) = $80 \times 165 = 66$

$$\overline{200}$$

- كا٢ المحسوبة للإناث (نعم) = $\frac{(65-66)^2}{66} = 0.02$

- التكرار المتوقع للإناث (نعم) = $\frac{80 \times 35}{200} = 99$

- كا٢ المحسوبة للإناث (نعم) = $\frac{(14-15)^2}{14} = 0.07$

(2) درجات الحرية = (عدد الصفوف – 1) × (عدد الاعمدة – 1)

= (2 – 1) × (2 – 1) = 1

(3) كا٢ الجدولية = كا٢ (1 ، 5 %) = 2.841

(4) ∵ كا٢ المحسوبة < كا٢ الجدولية ∴ يقبل الفرض الصفري ويرفض الفرض البديل

الطريقة المباشرة لايجاد كا٢ في حالة متغيرين 2 × 2 :

اذا كان التوزيع التكراري للظاهرتين ياخذ الشكل :

الظاهرة الثانية / الظاهرة الاولى	Y^1	Y^2	المجموع
X^1	A	B	A + B
X^2	C	D	C + D
المجموع	C+A	D+B	

فان كان كا٢ = $\frac{(AD - BC)^2 \times N}{\text{حاصل ضرب المجاميع الهامشية}}$

كا٢ = $\frac{(AD - BC)^2 \times N}{}$

$$(D + C)\ (B + A)\ (D + B)\ (C + A)$$

مثال : نفس المثال السابق

	نعم	لا	المجموع
ذكور	100	20	120
اناث	65	15	80
المجموع	165	35	200

احسب **كا^{٢}** عند مستوي معنوية 5% بالطريقة المباشرة

الحـل

(1) **كـا^{٢}** المحسوبة = ن × (أ د – ب جـ) (أ + جـ) (ب + د) (أ + ب) (جـ + د)

= 200× (100 × 15− 20 × 65) (165 × 35 × 120 ×80)

= 200 × (1500 −1300) (165 × 35 × 120 ×80)

= 200 × 40000 (165 × 35 × 120 ×80) = 0.144

(2) درجات الحرية = (عدد الصفوف – 1) (عدد الأعمدة – 1)

(3) = (2 – 1) (2 – 1) = 1

(4) **كـا^{٢}** الجدولية = **كـا^{٢}** (1 ، 5%) = 2.841

(5) ∵ **كـا^{٢}** المحسوبة < **كـا^{٢}** الجدولية

∴ نقبل الفرض الصفري وهي نفس النتيجة التي توصلنا اليها الطريقة السابقة.

ملحوظة هامة:

الطريقة الهامة التي تم بها حل المثال السابق تستخدم في حالة ما يكون التوزيع التكراري للصفين يأخذ شكل المصفوفة المربعة 2 × 2 فقط ولكن في حالة ما يكون التوزيع أكثر من ذلك تتبع الطريقة العامة في الحل

مثال : قام الأخصائي الاجتماعي بالمدينة الجامعية لجامعة الأزهر بدراسة رأي مجموعة من الطلاب نحو قضية الاغتراب داخل المدينة الجامعية للبنين والبنات وحصل علي البيانات التالية:

النوع / الرأي	أوافق	متردد	أعارض	المجموع
بنين	130	50	20	200
بنات	30	10	60	100
المجموع	160	60	80	300

هل توجد علاقة بين اغتراب البنين والبنات ، ناقش ذلك عند مستوي = 1%.

الحـل :التكرار المتوقع للبنين (موافق) = $\frac{160 \times 200}{300} = 106.7$

كا^2 المحسوبة للبنين (موافق) = $\frac{(106.7 - 130)^2}{106.7} = 5.09$

التكرار المتوقع للبنين (متردد) = $\frac{60 \times 200}{300} = 40$

كا^2 المحسوبة للبنين (متردد) = $\frac{(40 - 50)^2}{40} = 2.5$

التكرارالمتوقع للبنين (اعارض) = $\frac{80 \times 200}{300} = 53.3$

كا^2 المحسوبة للبنين (اعارض) = $\frac{(53.2 - 20)^2}{53.2} = 20.8$

التكرارالمتوقع للبنات (موافق) = $\frac{160 \times 100}{300} = 53.3$

كا² المحسوبة للبنات (موافق) = $\frac{(53.3 - 30)^2}{53.2} = 10.2$

التكرارالمتوقع للبنات (متردد) = $\frac{60 \times 100}{300} = 20$

كا² المحسوبة للبنات (متردد) = $\frac{(20 - 10)^2}{40} = 5$

التكرارالمتوقع للبنات (اعارض) = $\frac{80 \times 100}{300} = 26.7$

كا² المحسوبة للبنات (أعارض) = $\frac{(26.7 - 60)^2}{26.7} = 41.5$

(1) **كــا²** المحسـوبة للجـدول ككـل = 5.09 + 2.5 + 20.8 + 10.2 + 5 + 41.5 = 85.09

(2) درجات الحرية = (عدد الصفوف – 1) × (عدد الأعمدة – 1) = (2 – 1) × (3 –1) = 2 × 1 = 2

(3) ∴ Q **مستوي المعنوية 1 = α %**

(4) **كا²** الجدولية = **كا²** (2 ، 1 %) = 9.210

(5) ∴ **كا²** المحسوبة > **كا²** الجدولية العشوائية لأن 85ز09 > 9.210

∴ نرفض الفرض الصفري بمعني انه لا توجد علاقة بين درجة اغتراب البنين والبنات نحو تعايشهم داخل المدن الجامعية.

حساب كا² لتوزيع تكراري

عند حساب اختبار كا² لتوزيع تكراري ذو فئات نتبع ما يلي :

(1) نحصل علي مراكز الفئات ومن ثم نحصل علي قيمة كل من XF
، XF2.

(2) نحصل علي المتوسط الحسابي

$$\bar{X} = \frac{\sum F X}{\sum F}$$

(3) نحصل علي التباين من العلاقة :

$$\sigma^2 = \frac{1}{N} (\sum FX)^2 - \frac{(\sum FX)^2}{N}$$

(4) نحصل علي الانحراف المعياري σ وهو يساوي الجذر التربيعي للتباين.

(5) نحصل علي قيمة $x^{/}$

(6) نحصل علي الدرجة المعيارية D من العلاقة $(x^{/} - x) / \sigma$

(7) ومن جدولي المنحني الاعتدالي المعياري نحصل علي قيمة الارتفاع المعياري المقابل للدرجات المعيارية التي حصلنا عليها مسبقاً.

(8) نحسب التكرار المعدل من العلاقة :

$$\text{الارتفاع المعياري (الاعتدالي)} \times \frac{\text{مجموع التكرار}}{\text{الانحراف المعياري}} \times \text{طول الفئة}$$

(9) نحسب كا٢ المحسوبة من العلاقة:

$$\text{كا}^{٢} = \text{مجـ} \frac{(\text{التكرار التجريبي} - \text{التكرار المعدل})^2}{\text{التكرار المعدل}}$$

(10) نحدد درجات الحرية ø حيث درجات الحرية = عدد الفئات – 1

(11) نحدد مستوي المعنوية المطلوب α

(12) نحسب كا٢ الجدولية عند α المطلوبة

كا٢ (α ، ø)

(13) نقارن بين كا٢ المحسوبة ، كا٢ الجدولية.

(14) نتخذ القرار بقبول أو رفض الفرض الصفري.

مثال:

من جدول التوزيع التكراري التالي :

الفئات	10−	20−	30−	40−	50−	المجموع
التكرار	3	6	12	6	3	30

أحسب كا٢ عند مستوي معنوية **α** 5%.

الحل:

المتوسط الحسابي =

$$\dot{X} = \frac{\sum F X}{\sum F} = \frac{1050}{30} = 35$$

التباين

$$б = \frac{1}{\sum F} = \sum F X^2 - \frac{\sum F X}{\sum F}$$

$$б = \frac{1}{30} = 40350 - \frac{(1050)^2}{30}$$

$$б = \frac{1}{30} = 40350 - 36750 \quad \frac{1}{30} \quad X\ 3600 = 120$$

الانحراف المعياري σ =

$$\sqrt{\text{التباين}} = \sqrt{120} = \boxed{\text{11 تقريبا}}$$

لاحظ ان التكرار المعدل = الارتفاع الاعتدالي × مج ك÷ σ × طول الفئة

مثلا بالنسبة للفئة الاولي = $(30 \times 0.0761 / 11) \times 10 = 2.08$ هكذا

(1) كا2 المحسوبة = 1.3231

(2) درجات الحرية = عدد الفئات − 1 = 5 − 1 =4

(3) مستوي المعنوية $\alpha = 50\%$

(4) نحسب كا^{2} الجدولية = كا^{2} (4 ، 5 %) = 9.488

(5) كا^{2} المحسوبة 1.3231 < كا^{2} الجدولية 9.488

∴ يقبل الفرض الصفري بمعني ان الفروق غير معنوية أي ينطبق التوزيع التجريبي علي التوزيع الطبيعي

تذكـر.

- تستخدم توزيعات **كـا**^{٢} فـي اختبـارات صـحة نظريـة معينـة ويسـتخدم ذلـك فـي اختبارات الاستقلال بين المتغيرات أو الصفات المختلفة.
- نستخدم جدول توزيع **كا**^{٢} بدرجات حرية مناسبة Φ حيث ان

Φ = (عدد الأعمدة – 1) × (عدد الصفوف – 1)

- **كا**^{٢} المحسوبة = مجـ (التكرار المشاهد – الكرار المتوقع)

التكرار المتوقع

- **كا**^{٢} المحسوبة = $\dfrac{(F_O - F_e)^2}{F_e}$

- اذا كان **كا**^{٢}المحسوبة اقل من او يساوي **كا**^{٢}الجدولية بدرجات حرية Φ ومستوي معنوية α نقبل الفرض الصفري (العدمي) ، أما إذا كانت **كا**^{٢} المحسوبة اكبر من او تساوي **كا**^{٢} الجدولية فإننا نرفض الفرض الصفري (العدمي).
- في حالة وجود جداول عشوائية لحساب **كـا**^{٢} الجدوليـة تستخدم درجات الثقة مباشرة ولا تحتاج فيها الحصول عل مستوي المعنوية.
- عند استخدام **كا**^{٢} في اختبارات الفروض يجب ان تكون التكرارات الكلية كبيرة بحيـث لا تقل عـن 50 مشـاهد ، كمـا يجـب الا يقـل تكـرار أي خليـة عـن 5 تكرارات ويمكننا التغلب علي ذلك بإدماج خلايا التكرارات الصغيرة وذلك بدمج أكثر من صف أو أكثر من عمود.
- لايجاد كا^{٢} في حالة متغيرين 2 × 2 :

فان كان **كا**^{٢} = $\dfrac{(AD - BC)^2 \times N}{\text{حاصل ضرب المجاميع الهامشية}}$

$(D + C)\ (B + A)\ (D + B)\ (C + A)$

- عند حساب اختبار كا٢ لتوزيع تكراري ذو فئات نتبع ما يلي :

- نحصل علي مراكز الفئات ومن ثم نحصل علي قيمة كل من XF ، XF2.
- نحصل علي المتوسط الحسابي
- نحصل علي التباين من العلاقة :

$$^{2}б = \frac{1}{N} \quad (\sum FX)^2 - \frac{(\sum FX)^2}{N}$$

- نحصل علي الانحراف المعياري б وهو يساوي الجذر التربيعي للتباين.
- نحصل علي قيمة $x^{/}$
- نحصل علي الدرجة المعيارية D من العلاقة **($x^{/}$ - x) / б**
- ومن جدولي المنحني الاعتدالي المعياري نحصل علي قيمة الارتفاع المعياري المقابل للدرجات المعيارية التي حصلنا عليها مسبقاً.
- نحسب التكرار المعدل من العلاقة :

$$\text{الارتفاع المعياري (الاعتدالي)} \times \frac{\text{مجموع التكرار}}{\text{الانحراف المعياري}} \times \text{طول الفئة}$$

- ثم نحسب كا٢ المحسوبة من العلاقة:

$$\text{كا}^{٢} = \text{مجـ} \frac{(\text{التكرار التجريبي} - \text{التكرار المعدل})^2}{\text{التكرار المعدل}}$$

- نحدد درجات الحرية ø حيث درجات الحرية = عدد الفئات – 1
- نحدد مستوي المعنوية المطلوب α
- نحسب كا٢ الجدولية عند α المطلوبة
- نقارن بين كا٢ المحسوبة ، كا٢ الجدولية.
- نتخذ القرار بقبول أو رفض الفرض الصفري.

<u>تمارين</u>

تمرين (1) فيما يلي استجابات مجموعة من الطلاب نحو قضية ما

نعم	لا	المجموع
110	40	150

احسب **كا**^{2} عند مستوي معنوية 5%.

تمرين (2) فيما يلي أراء مجموعة من الأفراد نحو برامج تلفزيونية معينة

ممتازة	جيدة	معقولة	رديئة	رديئة جدا	مجموع
22	4	26	11	8	71

احسب كا^{2} عند مستوي ثقة 95 %.

تمرين (3) فيما يلي استجابات عينة من الطلاب نحو المشاركة في الانشطة الطلابية

أشارك كثيرا	اشارك بدرجة متوسطة	اشارك قليلا	لا أشارك
24	43	33	17

أحسب **كا**^{2} عند درجة ثقة 99 %.

تمرين (4) فيما يلي استجابات عينة من الطلاب نحو استطلاع راي ما :

موافق	موافق الي حد ما	لا ادري	ارفض	ارفض الي حد ما
37	3	20	28	2

احسب **كا**^{2} عند مستوي معنوية 5 %

تمرين(5): فيما يلي استجابات عينة من الطلاب نحو رايهم في قضايا التنمية بمصر

	نعم	**الي حد ما**	**لا**	**المجموع**
ذكور	300	40	60	400
اناث	400	20	80	500
المجموع	700	60	140	900

أحسب التكرار المتوقع – واحسب **كا**^{٢} عند مستوي معنوية 5 %

تمرين(6): الجدول التالي يبين استجابات عينة ما ذات ثقافات مختلفة واتجاهات نحو المشاركة الشعبية لمشروعات التنمية المحلية:

المؤهل\ الاتجاه	**موافق**	**محايد**	**غير موافق**
اقل من متوسط	70	30	40
مؤهل متوسط	100	20	50
مؤهل عال	300	10	100

ناقش وجو علاقة بين درجة التعليم والاتجاه نحو المشاركة الشعبية ، وذلك عند مستوي معنوية 1%.

الباب الثامن
الأرقام القياسية
Index Numbers

الأرقام القياسية عبارة عن مقاييس للتغير النسبي فى قيم الظواهر أو مجموعة من الظواهر سواء من زمن لأخر أو من منطقة لأخرى، أى أنها أداة إحصائية لقياس الاختلافات أو التغيرات النسبية فى القيم وقد تكون تلك التغيرات فى أسعار أو كميات الإنتاج أو الكميات المسوقة أو المستهلكة من السلع المختلفة.

ويمكن تلخيص أهم الأرقام القياسية واستخداماتها فيما يلي:

1- الأرقام القياسية للأسعار: وهى تقيس التغير فى مستوى الأسعار خلال فترة زمنية معينة وبدراسة الأرقام القياسية للأسعار يمكن معرفة أسباب التغيرات السعرية أو تأثيرها على اقتصاديات المنطقة المدروسة أو تستخدم لاتخاذ خطوات أو إجراءات معينة للتحكم فى بعض تلك الأسعار.

2- الأرقام القياسية لمقارنة التغير فى كميات الإنتاج أو التجارة أو المبيعات أو المخزون أو المصدر أو المستورد وغيره من المعاملات وهى مفيدة فى دراسة التغيرات التاريخية للسلاسل الزمنية واستنتاج الدورات أو الإتجاه العام أو التغيرات الموسمية وهى ضرورية للمهتمين بالعمليات الإنتاجية والتسويقية.

3- الأرقام القياسية للتنبؤ: وتستخدم فى تحديد الاتجاهات المستقبلية بناء على ما تم من معاملات.

مفهوم الرقم القياسي:

هو عبارة عن رقم نسبي يمكن الحصول عليه بنسبة متغير أو قيمة ظاهرة فى فترة زمنية إلى نفس المتغير أو الظاهرة أو الكمية فى فترة زمنية أخرى.

حيث تسمى الفترة الأولى بفترة الأساس وتسمى الفترة الثانية بفترة المقارنة وكما يسمى المكان الذي تنسب إليه بمكان الأساس والمكان الأخر بمكان المقارنة. وإذا كان الرقم أكبر من 100% يعني ذلك أن الظاهرة فى تزايد والعكس صحيح وكذلك فالفرق بين الرقم القياسي ورقم الأساس يعطي معدل الزيادة أو النقص لقيم الظواهر.

خطوات تركيب الأرقام القياسية:-

هناك عدة طرق لتركيب الأرقام القياسية إلا أنها مهما اختلفت فى طريقة التركيب فإنها تشترك فى الخطوات التالية:

1- إختيار السلع أو القيم التي تدخل فى تركيب الرقم القياسي واختيار المصادر التي تستقي منها أسعار هذه السلع.

2- إختيار سنة أو فترة الأساس.

3- تحديد الأوزان لمختلف السلع مع مراعاة أن تناسب وزن كل سلعة مع أهميتها.

4- تلخيص البيانات فى شكل رقم واحد يمثل متوسط التغير النسبي فى الرقم القياسي.

أولاً- إختيار السلع والمصادر التي تستقي منها الأسعار:

باستعراض قائمة السلع التي يتداولها أفراد مجتمع معين يتبين لنا أنه يكاد يكون من المستحيل عملياً أن نحسب الرقم القياسي لكافة أنواع السلع جميعها ولكن هناك عدد محدود من السلع الهامة تستنفذ الجزء الأكبر من دخول أفراد المجتمع ولذلك نأخذ عينة تمثل جميع أنواع السلع.

أما عن المصادر التي يرجع إليها لجمع البيانات عن أسعار هذه السلع العامة فالمفروض أن أسعار السلعة الواحدة تتساوى فى السوق الواحدة تحت تأثير المنافسة الكاملة، أما فروق الأسعار التي نلاحظها بين مؤسسة تجارية وأخرى فإنها عادة ما يمكن التغلب عليها بإيجاد متوسط عدد من أسعار كل سلعة.

ثانياً- فترة أو سنة الأساس:

يشترط فى سنة الأساس أن تكون سنة طبيعية خالية من الأزمات الاقتصادية التي تنتج عن الحروب وغيرها من الظروف الدورية أو الشاذة.

ومن الضرورى أن نراعي طول الفترة بين سنة الأساس وسنة المقارنة.

ثالثاً: إختيار الأوزان:

من الواضح أن الأهمية النسبية للسلع المختارة لتركيب الأرقام القياسية تختلف فيما بينها. فمثلاً عند تركيب الرقم القياسي للأسعار المزرعية، نجد أن القطن والبامية يكونان ضمن السلع المختارة. وتأثير أسعار القطن على دخل المزارعين أكبر بكثير من اثر أسعار البامية. ولذا فإنه من الخطأ أن تحسب متوسطات أسعار السلع كما هى بل لابد من الأخذ فى الاعتبار أهميتها النسبية، هذا بالإضافة إلى أن وحدات

قياس السلع تختلف من سلعة إلى أخرى. ومن الضروري أن يعبر عن الأهمية النسبية لكل سلعة ويتم ذلك بترجيح أسعار السلع المختلفة بأوزان مناسبة تعكس الأهمية النسبية لكل منها. والطريقة الأكثر شيوعاً تعرف بالأوزان التجميعية Weighted Aggregate Method.

طرق تركيب الأرقام القياسية:

يتم تركيب الأرقام القياسية بإحدى الطريقتين الآتيتين:

أولاً:- الرقم القياسي بالسعر النسبي Relative Price index

ويقصد بذلك معدل التغير السنوي فى سعر سلعة ما فى سنة ما بالنسبة لسعرها فى سنة الأساس ويحسب بإحدى الطرق الآتية:

أ- السعر النسبي البسيط Simple Relative Price

ويتحصل عليه بقسمة سعر السلعة فى سنة المقارنة على سعرها فى سنة الأساس وضرب الناتج فى 100 أى أن:

$$X = \frac{X_i}{X_o} \times 100$$

حيث أن: X السعر النسبي البسيط للسلعة.

X_i سعر السلعة فى سنة المقارنة.

X_o سعر السلعة فى سنة الأساس.

مثـال:

يبين الجدول التالي أسعار إحدى السلع فى الفترة ما بين 1980 سنة 1985 كما أنه يبين أسعارها النسبية بفرض أن سنة الأساس 1980 وقد حسبت الأسعار النسبية بقسمة سعر السلعة فى السنة المطلوبة على 2 وهو سعر السلعة فى سنة 1980.

السنة	السعر بالجنيه	السعر النسبي البسيط

1980	2	$X = \frac{x_o}{x_o} \times 100 = 100$
1981	2.5	$X = \frac{X_1}{X_o} \times 100 = \frac{2.5}{2} \times 100 = 125$
1982	2.8	$X = \frac{X_2}{X_o} \times 100 = \frac{2.8}{2} \times 100 = 140$
1983	4.2	$X = \frac{X_3}{X_o} \times 100 = \frac{4.2}{2} \times 100 = 210$
1984	4.6	$X = \frac{X_5}{X_o} \times 100 = \frac{4.6}{2} \times 100 = 230$
1985	4.8	$X = \frac{X_6}{X_o} \times 100 = \frac{4.8}{2} \times 100 = 240$

ب- الرقم القياسي التجميعي البسيط للأسعار: Aggregate Simple Price index

الرقم القياسي التجميعي للأسعار $= \frac{\sum X_i}{\sum x_o} \times 100$

حيث :

إجمالي أسعار مجموعة من السلع فى سنة المقارنة $\sum X_i$

إجمالي أسعار نفس السلع فى سنة الأساس $\sum X_o$

مثـال:

يبين الجدول التالي أسعار بعض المحاصيل الحقلية فى سنة 1980 (سنة الأساس) وسنة 1985 (سنة المقارنة) كما أنه يبين السعر النسبي لكل سلعة محسوباً بالمعادلة السابقة.

المحصول	سعر سنة المقارنة X_O بالجنيه	سعر سنة المقارنة X_i بالجنيه	السعر النسبي البسيط
القمح	4.0	4.5	112.5
الشعير	3.5	4.1	117.1
الفول	3.8	4.2	110.5
القطن	9.0	10.2	113.3
$\sum$	20.3	23	453.4

ويكون الرقم القياسي التجميعي البسيط للأسعار:

$$\frac{23}{20.3}\times 100 = 113.3$$

ويعاب على هذه الطريقة أنها لا تعطي أى وزن للأهمية النسبية للسلع الداخلة فى تركيب الرقم القياسي.

(جـ) الوسط الحسابي البسيط للأسعار النسبية:

$$\text{الوسط الحسابي البسيط للأسعار النسبية} = \frac{\text{مجموع الأسعار النسبية للسلع}}{\text{عدد السلع}}$$

ومن المثال السابق يتضح أن:

$$\text{الوسط الحسابي للأسعار النسبية} = \frac{452.4}{4} = 113.1$$

ويعاب على هذه الطريقة أنها لا تأخذ فى الاعتبار الأهمية النسبية للسلع الداخلة فى حساب الرقم القياسي.

وبالمثل يمكن حساب الوسط الهندسي البسيط للأسعار النسبية .

ثانياً:- الرقم القياسي التجميعي المرجح Aggregate Weighted Price Index

فى هذه الطريقة تستخدم الكميات المنتجة أو المباعة أو المشتراه كأوزان لترجيح أسعار السلع عند حساب الأرقام القياسية للكميات المستخدمة كأوزان إما أن تكون كميات سنة الأساس أو كميات سنة المقارنة وفى حالة تركيب الأرقام القياسية للكميات المنتجة أو المباعة فتستخدم أسعار السلع كأوزان لترجيح الكميات المختلفة من السلع.

وفى هذه الطريقة تظهر الأهمية النسبية للسلع الداخلة فى تركيب الرقم القياسي.

يوجد أربعة أرقام قياسية تجميعية مرجحة تعرف بأسماء علماء الإحصاء الذين تواصلوا إليهم

1- رقم لاسبير Laspayar's Index

تستخدم الكميات المباعة أو المشتراه فى سنة الأساس لكل سلعة من السلع الداخلة فى حساب الرقم القياسي كأوزان ترجيح أسعارها ويمكن حسابه بإستخدام المعادلة:

$$\text{رقم لاسبير} = \frac{\sum X_i q_o}{\sum X_o q_o} \times 100$$

حيث أن $X_i q_o$ سعر سنة المقارنة مرجح بكمية الأساس

$X_o q_o$ سعر سنة الأساس مرجح بكمية الأساس

واستخدام كمية الأساس للترجيح لا يعكس بالضرورة الأهمية النسبية للسلع فى سنة المقارنة حيث أن أنماط الإنتاج أو استهلاك هذه السلع تختلف من سنة إلى أخرى.

2- رقم باش Pasch's index

تستخدم الكميات المنتجة أو المباعة أو المشتراه فى سنة المقارنة لكل سلعة من السلع الداخلة فى حساب الرقم القياسي كأوزان لترجيح أسعارها.

$$\text{رقم باش} = \frac{\sum X_i q_i}{\sum X_o q_i} \times 100$$

حيث أن $\sum X_i q_i$ سعر سنة المقارنة مرجح بكمية سنة المقارنة

$\sum X_o q_i$ سعر سنة الأساس مرجح بكمية سنة المقارنة

مثـال:

يقوم أحد المستهلكين بإنفاق مبلغ مائة جنيه سنوياً على كل من السلعتين (أ)،(ب) على مدار السنة، لذا فإن الكمية المستهلكة من كل سلعة فى أى سنة يتوقف على سعرها، ويبين الجدول التالي أسعار هاتين السلعتين وبالتالي الكمية المستهلكة منهما فى سنة الأساس صفر وسنتي المقارنة (1 ، 2) . ويلاحظ أن سعر السلعة (أ) قد انخفض فى سنة المقارنة (1) وانه قد زاد فى سنة المقارنة (2) يعكس ما حدث فى سعر السلعة (ب) والمطلوب تركيب رقم لاسبير ورقم باش لسنتي2،1.

الإنفاق السنوي بالجنيه	سعر الوحدة بالجنيه	الكمية المستهلكة بالوحدة
	سنة الأساس صفر	

إجمالي ما ينفق على سلعة (أ)	100	5	20
إجمالي ما ينفق على سلعة (ب)	100	2.5	40
إجمالي الإنفاق	200		
سنة المقارنة(1)			
إجمالي ما ينفق على سلعة (أ)	100	2.5	40
إجمالي ما ينفق على سلعة (ب)	100	3	33.33
إجمالي الإنفاق	200		
سنة المقارنة(2)			
إجمالي ما ينفق على سلعة (أ)	100	10	10
إجمالي ما ينفق على سلعة (ب)	100	2	50
إجمالي الإنفاق	200		

<u>الحـل :</u> سنة المقارنة (1):

$$\text{رقم لاسبير} = \frac{\sum X_i q_o}{\sum X_o q_o} \times 100$$

السلعة	$X_o\ q_o$	$X_i\ q_o$
أ	100	2.5×20=50
ب	100	3×40=120
$\sum$	**200**	**170**

$$\text{رقم لاسبير} = \frac{170}{200} \times 100 = 85$$

$$\text{رقم باش} = \frac{\sum X_i q_i}{\sum x_o q_i} \times 100$$

السلعة	$X_i\ q_i$	$X_o\ q_i$
أ	100	200
ب	99.99	83.33
$\sum$	**199.99**	**283.33**

$$\text{رقم باش} = \frac{199-99}{283-33} \times 100 = 70.6$$

سنة المقارنة (2):

$$\text{رقم لاسبير} = \frac{\sum X_i q_o}{\sum X_o q_o} \times 100$$

السلعة	$X_i q_o$	$X_o\ q_o$
أ	10×20=200	100
ب	2×40=80	100
∑	**280**	**200**

$$\text{رقم لاسبير} = \frac{170}{200} \times 100 = 85$$

$$\text{رقم باش} = \frac{\sum X_1 q_1}{\sum x_o q_1} \times 100$$

السلعة	$X_i\ q_i$	$X_o\ q_i$
أ	100	200
ب	99.99	83.33
∑	**199.99**	**283.33**

$$\text{رقم لاسبير} = \frac{280}{200} \times 100 = 140$$

$$\text{رقم باش} = \frac{\sum X_1 q_1}{\sum x_o q_1} \times 100$$

السلعة	$X_i q_i$	$X_o\ q_i$
أ	100	50
ب	100	125
∑	**200**	**175**

$$\text{رقم باش} = \frac{200}{175} \times 100 = 114.29$$

<u>الفرق بين رقم لاسبير ورقم باش:</u>

نلاحظ بأن لاسبير يرجح بكميات الأساس بينما يرجح باش بكميات المقارنة اعتقادا بأن نمط الإستهلاك عند الناس ثابت وبالتالي فكميات الاستهلاك

من المواد تقريباً ثابتة إلا أن باش يخالفه الرآى ويرى بأن النمط الاستهلاكي عند الناس يتغير مع الزمن ولا يمكن أن تبقى نفس الكميات ثابتة أو لا يمكن أن يبقى تفضيل الناس للمواد ثابت، فمواد كثيرة تكون مهمة فى سنة ما تصبح غير مهمة بعد عدة سنين أو العكس مادة تكون غير مهمة تصبح مهمة، وعليه يرى باش ضرورة الترجيح بالكميات المستهلكة فى سنة المقارنة وليس سنة الأساس، وهذا صحيح وخاصة إذا طالت الفترة بين سنتي الأساس والمقارنة. ولكن وبالرغم من هذا فإن رقم لاسبير هو الأكثر إستخداماً وشيوعاً وذلك لأنه يعتمد على بيانات سنة الأساس ولا يحتاج إلى بيانات جديدة كل سنة بينما رقم باش يحتاج إلى تجديد البيانات مما يتطلب إجراء مسوح سنوية لنفقات ودخل الأسرة، وهذا أمر ليس بالسهل على الدول لا سيما وأن تكلفة المسح عالية . ولكن إذا توفرت المسوح الحديثة وبالتالي البيانات الحديثة فمن الممكن حساب الرقمين بسهولة

(3) رقم مارشال – إدجورث: Marshall- Edgaworth Index

للتخلص من التحيز الناتج عن إستخدام كميات سنة الأساس أو كميات سنة المقارنة للترجيح فى كل من الرقمين السابقين لجأ مارشال إدجورث فى أبحاثهما المشتركة إلى إستخدام متوسط الكميات المنتجة أو المستهلكة لسنتي الأساس والمقارنة لترجيح أسعار السلع الداخلة فى تركيب الرقم القياسي للأسعار.

رقم مارشال – إدجورث $= \dfrac{\sum X_i(q_o+q_i)}{\sum x_o(q_o+q_i)} \times 100$

ويمكن حساب رقم مارشال – إدجورث لسنتي المقارنة فى المثال بالنسبة لسنة المقارنة (1)

السلعة	X_i	X_o	q_o+q_i	$X_i(q_o+q_i)$	$X_o(q_o+q_o)$
أ	2.5	5	60	150	200
ب	3	2.5	73.33	220	183.33
$\sum$				**370**	**483.33**

رقم مارشال– إدجورث $= \dfrac{370}{483.33} \times 100 = 76.55$

(4) رقم فيشر الأمثل Fisher's ideal index

هو عبارة عن الوسط الهندسي لرقمي لاسبير وباش ويحسب بإستخدام المعادلة

$$\text{رقم فيشر الأمثل} = \sqrt{\frac{x_i\, q_o}{x_o q_o} \times \frac{x_i q_i}{x_o q_i}}$$

فى المثال السابق لسنة المقارنة (1)

$$\text{رقم فيشر الأمثل} = \sqrt{\frac{170}{200} \times \frac{200}{283.33}} \times 100 = 75$$

مميزات الأرقام القياسية:

1- يعتبر بعض الإحصائيين أن الرقم القياسي الأمثل هو من الأرقام غير المفضلة وذلك لصعوبة تحديد ما يقيس هذا الرقم حيث أن كل ما يقيسه هو التغير فى المستوى العام للأسعار مثلاً. بينما الأرقام التجميعية تقارن النفقات بين الفترات المختلفة فهي أشمل وأفضل.

2- اعترض بعض الاقتصاديين على بعض الأرقام القياسية فقد اعترض كينز على الرقم القياسي للاسبير على أنه يفترض ثبات أذواق المستهلكين وكذلك ثبات كمية الاستهلاك بينما الحقيقة هى غير ذلك وهو بذلك رقم متحيز نحو الأعلى.

3- كذلك بالنسبة للرقم القياسي لباش فقد ظهر متحيزاً نحو الأسفل لأنه ليس من المعقول أن يكون المستهلك قد اشترى نفس الكمية المستهلكة فى سنة المقارنة بعد ارتفاع الأسعار عن فترة المقارنة.

4- ليس من الضرورى أن يكون رقم مارشال الأمثل متحيز.

5- وقد يرى كينز أن أفضل رقم قياسي للأسعار هو الذى يقيس التغير فى قيمة النقود ولهذا لابد من تركيب رقم قياسي يقيس المنفعة المتغيرة لجميع السلع التي تعطي نفس المنفعة لمجاميع متشابهة من الأفراد خلال فترة المقارنة وفترة الأساس.

ثالثاً:الرقم القياسي المتتابع:

يتم فيه إستخدام أساس متحرك بصورة مستمرة وعدم الإعتماد على فترة أساس واحدة ثابتة يقارن بها جميع الفترات الأخرى فإذا استخدمت إحدى الفترات كأساس لفترة ثانية تستخدم الفترة الثانية كأساس للفترة الثالثة وهكذا.

ومن مزايا هذا الرقم انه إذا استخدم لحساب سلسلة من الأرقام القياسية لأسعار مجموعة من السلع فإنه يمكن إضافة سلع جديدة أو التخلص من بعض السلع التي تفقد أهميتها بالنسبة على السلع الداخلة فى حساب الرقم القياسي.

فإذا كان المطلوب مقارنة السعر النسبي لسلعة ما فى سنوات مختلفة اعتبارا من عام 1975 يعتبر السعر فى ذلك العام مساوياً 100 كأساس ويحسب السعر النسبي لعام 1976 بالنسبة لسنة الأساس 1975 ثم تعتبر سنة 1976 سنة أساس لعام 1977 وهكذا.

$$\text{السعر النسبي للسلعة فى 1976} = \frac{\text{سعر سنة 1976}}{\text{سعر سنة 1975}} \times 100$$

$$\text{السعر النسبي للسلعة فى 1977} = \frac{\text{سعر سنة 1977}}{\text{سعر سنة 1976}} \times 100$$

وعموماَ تستخدم الأرقام القياسية للأسعار لإزالة أثر التغير فى القوة الشرائية للنقود علاوة على استخدامها فى قياس التغيرات التي تحدث فى أسعار سلعة ما أو مجموعة من السلع. كما أنها تستخدم فى تركيب مقاييس أخرى مثل الأرقام القياسية للقوة الشرائية وأسعار المساواة والأسعار المعدلة وكلها مقاييس تستخدم فى فهم وشرح العلاقات الاقتصادية السائدة بين الأنشطة الاقتصادية المختلفة.

رابعاً:الأرقام القياسية للقوة الشرائية:

يمكن حساب الرقم القياسي للقوة الشرائية Purchasing Power Index لمجموعة من الأفراد بإيجاد النسبة بين أسعار السلع والخدمات التي يقومون بإنتاجها وبيعها وأسعار السلع والخدمات التي يتحصلون عليها والأسعار التي يدفعونها.

$$\textbf{الرقم القياسي للقوة الشرائية للمزارعين} = \frac{\textbf{الرقم القياسي للأسعار التي تدفع للمزارعين}}{\textbf{الرقم القياسي للأسعار التي يدفعها المزارعون}}$$

ويلاحظ إستخدام رقمين قياسين للأسعار أحدهما يمثل دخل المزارعين والأخر يمثل إنفاقهم ولذلك يجب توحيد سنة الأساس المستخدمة فى تركيب كل منهما وكذلك سنة المقارنة .

إلا أن هذا الرقم يعطي فكرة عامة من القوة الشرائية للمزارعين لأنه مركب من أسعار عديدة من المحاصيل الزراعية وبالتالي يتأثر بها كما أن هذا الرقم لا يأخذ فى الاعتبار التغيرات فى تكاليف الإنتاج ويعكس فقط التغيرات التي تطرأ على الأسعار.

الأسعـار المعدلـة

يستخدم السعر المعدل Adjusted Price لسلعة، فى تحليل الأسعار وينحصر الغرض من استخدامه فى التخلص من اثر المستوى العام للأسعار حتى يمكن عمل المقارنة بين سعري سلعتين ويمكن حسابه كالآتي:

$$\text{السعر المعدل سنة 1990} = \frac{\text{سعر السلعة فى سنة 1990}}{\text{الرقم القياسي العام لأسعار سنة 1990}} \times 100$$

ويمكن التخلص من اثر المستوى العام لأسعار عند حساب الرقم القياسي للقوة الشرائية لسلعة ما كالأتي:

$$\text{السلع} = \frac{\text{السعر النسبى للسلعة}}{\text{الرقم القياسي لأسعار السلع المماثلة}}$$

$$\text{القوة الشرائية للجنيه فى سنة المقارنة} = \frac{\text{الرقم القياسي للأسعار فى سنة الأساس}}{\text{الرقم القياسي للأسعار فى سنة المقارنة}} \times 100$$

مثـال:

الجدول التالي يبين الكميات المباعة وأسعار سنة الأساس وسنة المقارنة لأربع من السلع الزراعية أ،ب،جـ،د كما يبين الرقم القياسي لأسعار السلع الزراعية والرقم القياسي للأسعار التي يدفعها المزارعون لسلع المشتراه والمطلوب حساب الرقم القياسي لأسعار السلع الزراعية الأربع.

1- القوة الشرائية للجنيه.
2- القوة الشرائية للأربع سلع الزراعية بالنسبة للسلع التي يشتريها المزارعين.
3- السعر المعدل للسلعة بالنسبة للتغير فى المستوى العام للأسعار.
4- السعر النسبي.
5- الرقم القياسي للقوة الشرائية للسلعة بالنسبة لباقي المحاصيل الزراعية.

السلعة	متوسط الكمية المباعة	سعر سنة الأساس بالجنيه (صفر)	سعر سنة المقارنة بالجنيه
أ	80 إردب	8	7
ب	160 قنطار	11	10
جـ	800 كيله	1.2	0.6
د	7500 وحدة	0.2	0.15
الرقم القياسي لأسعار السلع الزراعية		**100**	**130**
الرقم القياسي لأسعار السلع المشتراه		**100**	**140**

<u>الحــل</u>

يجب التخلص من أثر اختلاف وحدات القياس بإستخدام القيمة النقدية للكميات المباعة أو المشتراه من السلع الداخلة فى حساب رقم القياس للأسعار كما يلي:

	القيمة النقدية لسنة الأساس	القيمة النقدية لسنة المقارنة
السلعة أ	640=80×8	560=80×7
السلعة ب	1760=160×11	1600=160×10
السلعة جـ	960=800×1.2	720=800×1.9
السلعة د	1500=7500×0.2	1125=7500×0.15
إجمــالي	**4860**	**4005**

أ- الرقم القياسي لأسعار السلع الزراعية الأربع $= \frac{4005}{4860} \times 100 = 82.4$

أى أن أسعار السلع الزراعية انخفضت بمقدار 17.6%

ب- القوة الشرائية للجنيه فى سنة المقارنة $= \frac{100}{130} \times 100 = 76.9$

أى أن القوة الشرائية لجنيه قد انخفضت بمقدار 23.1%

جـ- القوة الشرائية للسلع $= \frac{82.4}{140} \times 100 = 59$

أى أن القوة الشرائية للسلعة الزراعية الأربعة قد انخفضت بحوالي 41%

د- السعر المعدل للسلعة أ بالنسبة للتغير فى المستوى العام لأسعار

$= \frac{\text{سعر السلعة فى سنة المقارنة}}{\text{الرقم القياسى العام للأسعار فى سنة المقارنة}} \times 100 = \frac{7}{130} \times 100 = 5.38$ جنيهاً

هـ- السعر النسبي للسلعة أ $= \frac{7}{8} \times 100 = 87.5$

- السعر النسبي للسلعة ب $= \frac{0.15}{0.20} \times 100 = 75$

و- الرقم القياسي للقوة الشرائية للسلعة د بالنسبة للمحاصيل الزراعية الأربعة

$= \frac{\text{السعر النسبى للسلعة}}{} \times 100 = \frac{75}{} \times 100 = 91$

$$\frac{\text{الرقم القياسي العام لأسعار السلع المماثلة}}{82.4}$$

الرقم القياسي للقوة الشرائية للسلعة أ بالنسبة للمحاصيل الزراعية الأربعة

$$= \frac{87.5}{82.4} \times 100 = 106.19$$

سعر المساواة

يعرف سعر المساواة Parity Price لسلعة ما بأنه سعر الوحدة من السلعة التي يمكن المنتج من الحصول على نفس الكمية من السلع والخدمات التي كان عليها فى فترة الأساس مقابل وحدة واحدة من إنتاجه.

وهو مقياس لحساب السعر العادل لسلع بعض المنتجين بالنسبة لأسعار السلع التي يشترونها.

وتنحصر خطوات حساب سعر المساواة فى الأتي:

1- حساب الرقم القياسي للمساواة لفترة الأساس المختارة وهو عبارة عن رقم قياسي للأسعار التي يدفعها المزارعون ويشتمل على تكاليف الإنتاج المزرعى وأسعار السلع والخدمات التي يشتريها المزارعون وأوجر العمال والفائدة على القروض الزراعية.

2- حساب الرقم القياسي للأسعار التي يحصل عليها المزارعون لجميع السلع والمنتجات الزراعية بالإضافة على الإعانات الحكومية التي قد تدفع للمزارعين.

3- قسمة الرقم القياسي للأسعار المتحصل عليها (خطوة (2)) على الرقم القياسي للأسعار المدفوعة (خطوة(1)) تنتج نسبة المساواة والتي تعتبر مقياسا للقوة الشرائية للمنتجات التي يبيعها المزارعون بالنسبة للسلع والخدمات التي يشترونها بالنسبة لفترة الأساس المختارة.

4- يحسب سعر الأساس المعدل بإيجاد النسبة بين متوسط أسعار السلعة أو الخدمة التي يتحصل عليها المزارعون لفترة عشر سنوات سابقة لتاريخ حساب سعر المساواة ومتوسط الأرقام القياسية للأسعار التي يتحصل عليها المزارعون لكل السلع الزراعية لفترة العشر سنوات نفسها أى أن:

$$\textbf{سعر الأساس المعدل لسلعة ما} = \frac{\textbf{متوسط سعر السلعة فى فترة عشر سنوات}}{\textbf{متوسط الأرقام القياسية لأسعار السلع الزراعية فى نفس الفترة}} \times \mathbf{100}$$

5- يكون سعر المساواة لسلعة ما = سعر الأساس المعدل لها × الرقم القياسي للمساواة

مثـال:

بفرض أن متوسط السعر الذى حصل عليه المزارعون للأردب القمح لفترته 120 شهراً ابتداء من يناير 1950 إلى ديسمبر 1959 كان 5.4 جنيهاً وأن متوسط الأرقام القياسية لأسعار كل المنتجات الزراعية خلال نفس الفترة كان 180 بالنسبة لفترة الأساس (1930–1934) وان الرقم القياسي للأسعار التي يدفعها المزارعون والذي يشمل الفائدة على القروض الزراعية والضرائب ومعدلات أجور العمال الزراعيين (100) أى الرقم القياسي للمساواة لشهر إبريل 1960 كان مساوياً بالنسبة لفترة الأساس المستخدمة والمطلوب حساب الأتي:

(1) سعر المساواة للقمح فى إبريل سنة 1960.

(2) نسبة المساواة.

الحـل:

سعر الأساس المعدل للقمح $= \frac{4.5}{180} \times 100 = 2.5$ جنيهاً

سعر المساواة للقمح فى شهر إبريل1960 $= 2.5 \times \frac{220}{100} = 5.50$ جنيهاً

أى أن سعر المساواة للقمح فى شهر إبريل 1960 يجب أن يكون 5.50 جنيهاً، حتى يتسنى للمزارع أن يتحصل على نفس الكمية من السلع والخدمات التي كان يتحصل عليها فترة الأساس مقابل أردب واحد من القمح.

نسبة المساواة =

$$\frac{\text{الرقم القياسي للأسعار التي يتحصل عليها المزارعون}}{\text{الرقم القياسي للأسعار التي يدفعها المزارعون}} \times 100$$

$$= \frac{180}{220} \times 100 = 81.81$$

أى أن القوة الشرائية للمنتجات الزراعية قد انخفضت عما كانت عليه

اختبار الأرقام القياسية

توجد عدة اختبارات تجرى على الأرقام القياسية المقدرة بالطرق المختلفة لمعرفة مدى كفاءة تلك الأرقام وهى.

1- اختبار الانعكاس الزماني أو المكاني.
2- اختبار الانعكاس فى المعامل.

(1) اختبار الانعكاس الزماني: Time Reversal Test

أول من استعمل هذا الاختبار هو فيشر، ويحقق الرقم القياسي اختبار الانعكاس الزماني إذا كان حاصل ضرب قيمة (بإستخدام الفترة الزمنية a كأساس الفترة الزمنية b للمقارنة) فى قيمة الرقم القياسي للانعكاس(بإستخدام الفترة الزمنية b كأساس والفترة الزمنية a للمقارنة) يساوى الواحد الصحيح.

بفرض أن الرقم القياسي البسيط (السعر النسبي) لإحدى السلع لسنة المقارنة b إلى سعرها لسنة الأساس a هو:

ويلاحظ أن رقم لاسبير وباش لا يحثث هذا الاختبار ، أى أن النتائج المتحصل عليها من إستخدام رقم لاسبير وباش تكون متحيزة بعض الشئ.

ويلاحظ أن رقم فيشر يحقق هذا الاختبار ويشترك معه الرقم القياسي التجميعي النسبي البسيط.

(2) – اختبار الانعكاس المكاني:Place Reversal Test

هو الاختبار المستخدم عند مقارنة التغير بين المناطق المختلفة وبذلك تستخدم إحدى المناطق كأساس للمقارنة وتقارن بها باقي المناطق وينطبق على هذا الاختبار ما قيل عن اختبار الانعكاس الزماني.

(3)– اختبار الانعكاس المعامل:Factor Reversal Test f

وفيشر أول من استخدم هذا الاختبار وتبني فكرته على أساس أن

القيمة = الكمية ×السعر

الرقم القياسي للأسعار × الرقم القياسي للكميات = الرقم القياسي للقيمة

ورقم فيشر هو المحقق أيضاً لهذا الاختبار

ويمكن إثبات هذا بالمعادلات

الرقم القياسي للأسعار × الرقم القياسي للكميات=

$$\sqrt{\frac{\sum p_i q_i}{\sum p_o q_o} \times \frac{\sum p_i q_i}{\sum p_o q_i}} \times \sqrt{\frac{\sum p_i q_o}{\sum p_o q_o} \times \frac{\sum p_i q_i}{\sum p_o q_i}}$$

$$= \sqrt{\frac{(p_i q_i)^2}{(p_o q_o)^2}} = \frac{p_i q_i}{p_o q_o}$$

الرقم القياسي الأمثل للقيمة =

بعض الأرقام القياسية الهامة:

(1) الرقم القياسي لنفقة المعيشة Cost of Living Index

ويعرف أيضاً باسم الرقم القياسي لأسعار المستهلك أو الرقم القياسي لأسعار التجزئة، وهو يعبر عن التغير فى أسعار السلع والخدمات التي تستهلكها الأسرة خلال فترة زمنية محددة.

لهذا الرقم أهمية خاصة حيث يتخذ كدليل فى تحديد الأجور وفى مراقبة الأسعار، هذا ويجب التمييز بين نفقة المعيشة ومستوى المعيشة، حيث أن نفقة المعيشة تعبر عن ثمن السلع والخدمات المستهلكة خلال فترة زمنية محددة بينما مستوى المعيشة يعبر عن مقدار السلع والخدمات المستهلكة خلال نفس الفترة.

فارتفاع نفقة المعيشة يعني انخفاض الدخل الحقيقي للفرد بينما مستوى المعيشة يعبر عن ارتفاع الدخل الحقيقي.

مثال:

نفرض أن أحد بحوث نفقة المعيشة لعدد من أسر صغار الموظفين والمستخدمين قد دل على أن دخلهم يتوزع على أقسام الأنفاق السابقة على النحو التالي:

المواد الغذائية	50%	الدخان	3%
المحروقات والماء والكهرباء	6%	نفقات أخرى	9%
المسكن	15%		
الملبس	10%		
التجهيزات المنزلية	7%		

لتركيب الرقم القياسي يمكن اتخاذ هذه النسب كأوزان للترجيح ويتم تحديد فترة الأساس ويحسب رقم قياسي لكل قسم من الأقسام السبعة حيث يتركب كل قسم من عدة سلع يتم تحديدها حسب أهميتها أو تماثلها ويتم تحديد سعرها ويضرب فى الكمية المناظرة فيكون المجموع الناتج من ضرب الأسعار فى فترة الأساس فى نفس الكميات فيكون الناتج هو الرقم القياسي لأسعار القسم.

بعد حساب الرقم القياسي لكل قسم يضرب فى النسبة المئوية المذكورة أعلاه المناظرة للقسم وتجمع النتائج ويقسم المجموع على 100 فنحصل على الرقم القياسي لنفقة المعيشة.

ومعنى هذا أن الرقم القياسي لنفقة المعيشة يساوى الوسط الحسابي المرجح للأرقام السبعة حيث تعتبر النسب السابقة لتوزيع نفقة المعيشة كأوزان للترجيح.
هذا ويجب التنبيه إلى أنه كلما ارتفاع مستوى المعيشة كلما قلت نسبة الإنفاق على المواد الغذائية وهكذا فمثالنا أعلاه يدل على مستوى منخفض للمعيشة.

(2) الرقم القياسي لأسعار الجملة: The Wholesale Price Index

الرقم القياسي لأسعار الجملة هو أداة إحصائية لقياس متوسط التغير فى أسعار مجموعة معينة من مواد الجملة التي يتم تبادلها خلال فترة معينة من الزمن.
ويمكن للرقم القياسي لأسعار الجملة أن يمثل جميع المواد التي تدخل فى مبادلات الجملة، أو أن يقتصر على فئة كأن يكون رقماً قياسياً لأسعار الجملة للمواد الغذائية مثلا أو للمواد الأولية أو للمواد المستوردة أو المصدرة.
ويمكنا حساب الرقم القياسي لأسعار الجملة بإستخدام أى صيغة من صيغ الأرقام القياسية المرجحة التي درسناها وأكثر هذه الصيغ استعمالا هى الصيغ التالية:

(1) صيغة لاسبير

$$\text{رقم لاسبير} = \frac{\sum X_i q_o}{\sum X_o q_o} \times 100$$

(2) صيغة باش

$$\text{رقم باش} = \frac{\sum X_i q_i}{\sum X_o q_i} \times 100$$

(3) صيغة الرقم القياسي المرجح بقسم سنة الأساس

$$= \frac{\sum \frac{x_i}{x_o} X_o q_o}{\sum X_o q_o} \times 100$$

والصيغة الأخيرة هذه تقود إلى نتائج مماثلة لصيغة لاسبير هذا وتعتبر صيغة لاسبير أكثر شيوعاً واستعمالا من غيرها وذلك لسهولتها ووضح معناها. تليها الصيغة الثالثة أى صيغة الرقم القياسي بقسم سنة الأساس.

(3) الأرقام القياسية للتجارة الخارجية:

تقوم معظم دول العالم بنشر إحصاءات التجارة الخارجية التي تمثل كميات وقيم الصادرات والواردات سنوياً. فعندما نرغب فى معرفة التطورات التي طرأت

على الصادرات مثلا فى بلد معين بلغت فيه الصادرات فى سنة معينة 300 مليون جنيه ثم ارتفعت إلى 360 مليون جنيه فى سنة أخرى أى أن الرقم القياسي للقيم كما يلي$\frac{360}{300}\times100=120$

أى أن التصدير قد ازداد بنسبة 20% ولكن هذه الزيادة يمكن أن تكون نتيجة لأحد **الاحتمالات التالية:**

(1) ازدياد كمية الصادرات وارتفاع الأسعار.

(2) ازدياد كمية الصادرات وبقاء الأسعار ثابتة.

(3) ازدياد كمية الصادرات وهبوط الأسعار.

(4) ثبات كمية الصادرات وارتفاع الأسعار.

(5) تناقص كمية الصادرات وارتفاع الأسعار.

كذلك لو زادت قيمة الواردات من 400 مليون جنيه إلى 500 مليون جنيه أى بنسبة 25% فان الزيادة يمكن أن تكون نتيجة لأحد **الاحتمالات التالية:**

(1) ازدياد كمية الواردات وارتفاع الأسعار.

(2) ازدياد كمية الواردات وبقاء الأسعار ثابتة.

(3) ازدياد كمية الواردات وهبوط الأسعار.

(4) ثبات كمية الواردات وارتفاع الأسعار.

(5) تناقص كمية الواردات وارتفاع الأسعار.

ولتحديد أى الاحتمالات السابقة كان وراء هذا التغير فى الصادرات أو الواردات فإنه لابد من حساب أثر كل من الأسعار والكميات على قيمة الصادرات أو الواردات مع باء المتغيرات الأخرى ثابتة.

ومن أجل ذلك يجب حساب رقمين قياسيين للتجارة الخارجية هما:

- الرقم القياسي لوحدة الكمية Quantum Index number.
- والرقم القياسي لسعر الوحدة Unit value index number.

وفيما يلي مثال على ذلك:

مثـــال:

لنفرض أن بيانات الصادرات لأحد الدول كانت كما يلي:

	سنة 1990	سنة 2000
كمية الصادرات (بالطن)	2.6	3
قيمة الصادرات (بالجنيه)	450	550

ادرس تطور صادرات هذه الدولة.

الحـــل:

الرقم القياسي لكمية الصادرات يساوي : $\frac{3}{2.6}\times 100=115.4\%$

أى أن كمية الصادرات زادت بنسبة 15.4%.

الرقم القياسي لقيمة الصادرات يساوى: $\frac{550}{450}\times 100=122.2$

أى أن قيم الصادرات زادت بنسبة 22.2%

وهذا يدل على أن الزيادة فى القيمة كانت أكبر من الزيادة فى الكمية الأمر الذى يوضح أن الرقم القياسي للقيمة = الرقم القياسي للكمية × الرقم القياسي للأسعار

$$\text{الرقم القياسي للأسعار} = \frac{\text{الرقم القياسي للقيمة}}{\text{الرقم القياسي لكمية}}\times 100$$

$$= \frac{1222}{1154}\times 100=1059$$

مما يستدل منه على أن الأسعار قد تحسنت بمقدار 5.9% خلال الفترة 1990-2000.

هذا ومن الممكن استعمال أى صيغة من صيغ الأرقام القياسية التي درسناها لحساب كل من الرقم القياسي لوحدة الكمية والرقم القياسي لسعر الوحدة. غير أنه من الشائع لدى معظم دول العالم استخدام صيغة لاسبير كما يلي

(ا) الرقم القياسي لوحدة الكمية يساوي

$$\frac{\sum q_i X_o}{\sum q_o X_o}\times 100$$

(4) الرقم القياسي للإنتاج: The Production Index Number

بالنسبة للرقم القياسي للإنتاج تختلف وحدات الإنتاج فيما بينها كالطن ، والمتر ، والذراع واللتر وغيرها من وحدات قياس الإنتاج. وعند حسابنا للرقم القياسي للإنتاج لابد من إيجاد عامل مشترك بين هذه الوحدات المختلفة وهذا العامل فى الواقع هو أوزان الترجيح التي تعطي لكل نوع من أنواع الإنتاج وهناك أربعة عوامل يمكن أن نرجح بها وهى:

(1) الإشباع أو الفائدة التي تعطيها كل وحدة إنتاج.

(2) أسعار الوحدة المنتجة.

(3) ساعات العمل الإنساني اللازمة لإنتاج وحدة واحدة.

(4) قيم الإنتاج.

وغالباً ما تستخدم صيغة لاسبير لحساب الرقم القياسي للإنتاج. وهكذا تصبح صيغ لاسبير المستخدمة للعوامل الأربعة السابقة كما يلي:

(1) إذا رجحنا بالفائدة أو الإشباع الذى نحصل عليه من كل وحدة فى سنة الأساس فان رقم لاسبير للإنتاج هو

$$= \frac{\sum q_i f_o}{\sum q_o f_o} \times 100$$

حيث: q_i كمية الإنتاج فى السنة المدروسة (سنة المقارنة).

q_o كمية الإنتاج سنة الأساس.

f_o الفائدة أو الإشباع فى سنة الأساس.

إلا أن هذه المعادلة غير ممكنة التطبيق لصعوبة تقدير قيمة f_o.

(2) أما إذا رجحنا بالسعار فان صيغة لاسبير للرقم القياسي للإنتاج تكون :

$$= \frac{\sum q_i x_o}{\sum q_o x_o} \times 100$$

حيث p_o أسعار سنة الأساس.

(3)وإذا رجحنا بساعات العمل اللازمة لإنتاج وحدة واحدة فى سنة الأساس تكون صيغة لاسبير كما يلي:

$$= \frac{\sum q_i h_o}{\sum q_o h_o} \times 100$$

حيث h_o هى ساعات العمل اللازمة لإنتاج وحدة إنتاج واحدة فى سنة الأساس.

(4) وأخيراً إذا رجحنا بقيمة الإنتاج فى سنة الأساس فان صيغة لاسبير تكون:

$$= \frac{\sum \frac{q_i}{q_o} X_o q_o}{\sum X_o q_o} = \frac{\sum \frac{q_i}{q_o} v_o}{\sum v_o}$$

حيث v_o هى قيمة الوحدة الإنتاجية فى سنة الأساس وتساوي $q_o\,p_o$

(5) القوة الشرائية للعملة:Purchasing Power of Money

لقد ناقشنا حتى الآن العديد من الصيغ لحساب الأرقام القياسية، وفى هذا الجزء سنناقش كيفية استخدام الأرقام القياسية فى تعديل الأسعار والدخول.

إن الهدف الأساس للرقم القياسي للأسعار هو قياس التغيرات فى الأسعار خلال الفترة المدروسة.أو بعبارة أخرى هو قياس القوة الشرائية للعملة.

من المعروف أن القوة الشرائية تتناقص كلما تزايدت الأسعار. أو بمعنى أخر طالما أن الرقم القياسي لسنة الأساس هو 100 فإن القوة الشرائية للعملة فى سنة المقارنة يساوي تقسم الرقم القياسي للأسعار فى سنة المقارنة أى أن:

$$\text{القوة الشرائية للعملة} = \frac{100}{\text{الرقم القياسي للأسعار}} \times 100$$

مثـــال:

نفرض أن الرقم القياسـي لأسـعار فـى الأردن عـام 2000 بالنسبة لعـام 1990 هـو 120% فـان القـوة الشـرائية للـدينار الأردنـي عـام 2000 مقارنـة بعـام 1990 تساوي $\frac{100}{120} \times 100 = 0.83$

وتكون القوة الشرائية للدولار فى الأردن تساوى $\frac{100}{120} \times 0.71 = 0.59$

حيث الدولار =0.71 دينار.

مثـــال:

إذا كان الرقم القياسي لأسعار المستهلكة للسنوات 1988–1994 كما هو مبين فى الجدول التالي ، احسب القوة الشرائية للسنوات المتعاقبة.

السنة	الرقم القياسي لأسعار المستهلك 1988=100
1988	100

1989	103
1990	105
1991	108
1992	112
1993	114
1994	115

<u>الحـل:</u>

نحسب القوة الشرائية للعملة فى كل سنة من العلاقة التالية:

$$\text{القوة الشرائية للعملة} = \frac{100}{\text{الرقم القياسي للأسعار}} \times 100$$

والنتائج مبينة فى الجدول التالي:

السنة	القوة الشرائية للعملة 1988 سنة أساس
1988	1.00
1989	0.97
1990	0.95
1991	0.93
1992	0.89
1993	0.88
1994	0.87

النتيجـة 0.87 المنـاظرة لسـنة 1994 تعتنـي أن القـوة الشـرائية للعملـة عـام 1994 تعادل 0.87 من قوتها عام 1988.

بإستخدام الرقم القياسي لسعار المستهلك نستطيع أيضا أن نحدد قيمة الدخل الحقيقي للشركة أو الأجر الحقيقي للعملة أو السعر الحقيقي للسلعة.

بمـا أن الرقم القياسـي لسـعار المسـتهلك يسـاوى 100 فـى سنة الأسـاس فـان الأجـر الحقيقـي (أو الـدخل الحقيقـي أو السـعر الحقيقـي) هـو ببسـاطة يسـاوي الأجـر (أو الدخل أو السعر) فى سنة المقارنة مضروبا بالقوة الشرائية للعملة.

الدخل الحقيقي = الدخل فى سنة المقارنة × القوة الشرائية للعملة

وهذا ينطبق على الأجر أو السعر الحقيقي.

أو بعبارة أخرى

$$\text{الدخل الحقيقي} = \text{الدخل فى سنة المقارنة} \times \frac{100}{\text{الرقم القياسي للأسعار}}$$

سنوضح هذا الانكماش (Deflation) فى الدخل أو الأجر أو السعار الحقيقية بالمثال التالي:

<u>مثـال:</u>

إذا كان متوسط الدخل الشهري لأسرة فى دولة ما والرقم القياسي لسعر المستهلك خلال الفترة 1990–1994 هى كما فى الجدول . أوجد متوسط الدخول للسنوات المتعاقبة بعملة سنة الأساس 1990.

الرقم القياسي للأسعار	متوسط الدخل الشهري	السنة
100	80	**1990**
103	100	**1991**
108	120	**1992**
112	130	**1993**
115	150	**1994**

<u>الحـل:</u>

يمكننا أن نحسب المتوسط الشهري للدخل بعملة سنة الأساس 1990 بقسمة متوسط الدخل الشهري لسنة معينة على الرقم القياسي المناظر لتلك النسبة وضرب الناتج فى 100 والنتائج مبينة فى الجدول التالي

متوسط الدخل الشهري بعملة 1990	الرقم القياسي للأسعار	متوسط الدخل الشهري	السنة
80	100	80	**1990**
97	103	100	**1991**
111	108	120	**1992**
116	112	130	**1993**
130	115	150	**1994**

<u>تـذكر</u>

الأرقام القياسية عبارة عن مقاييس للتغير النسبي فى قيم الظواهر أو مجموعة من الظواهر سواء من زمن لأخر أو من منطقة لأخرى، أى أنها أداة إحصائية لقياس الاختلافات أو التغيرات النسبية فى القيم وقد تكون تلك التغيرات فى أسعار أو كميات الإنتاج أو الكميات المسوقة أو المستهلكة من السلع المختلفة.

الرقم القياسي:هو عبارة عن رقم نسبي يمكن الحصول عليه بنسبة متغير أو قيمة ظاهرة فى فترة زمنية إلى نفس المتغير أو الظاهرة أو الكمية فى فترة زمنية أخرى.

يمكن تلخيص أهم الأرقام القياسية واستخداماتها فيما يلي:

- **الأرقام القياسية للأسعار**: وهى تقيس التغير فى مستوى الأسعار خلال فترة زمنية معينة
- **الأرقام القياسية لمقارنة التغير** فى كميات الإنتاج أو التجارة أو المبيعات أو المخزون أو المصدر أو المستورد وغيره من المعاملات
- **الأرقام القياسية للتنبؤ:** وتستخدم فى تحديد الاتجاهات المستقبلية بناء على ما تم من معاملات.

يتم تركيب الأرقام القياسية بإحدى الطريقتين الآتيتين:

أولاً:- الرقم القياسي بالسعر النسبي Relative Price index

ويقصد بذلك معدل التغير السنوي فى سعر سلعة ما فى سنة ما بالنسبة لسعرها فى سنة الأساس ويحسب بإحدى الطرق الآتية:

أ- السعر النسبي البسيط $\mathrm{X} = \frac{X_i}{X_o} \times 100$

حيث أن: $\mathbf{X}$ السعر النسبي البسيط للسلعة.

$\mathbf{X_i}$ سعر السلعة فى سنة المقارنة.

$\mathbf{X_O}$ سعر السلعة فى سنة الأساس.

ب- الرقم القياسي التجميعي البسيط للأسعار

$$\frac{\sum X_i}{\sum x_o} \times 100$$

$\sum X_i$ إجمالي أسعار السلع فى سنة المقارنة

$\sum X_o$ إجمالي أسعار نفس السلع فى سنة الأساس

الرقم القياسي التجميعي المرجح

فى هذه الطريقة تستخدم الكميات المنتجة أو المباعة أو المشتراه كأوزان لترجيح أسعار السلع عند حساب الأرقام القياسية للكميات المستخدمة كأوزان إما أن تكون

كميات سنة الأساس أو كميات سنة المقارنة وفى حالة تركيب الأرقام القياسية للكميات المنتجة أو المباعة فتستخدم أسعار السلع كأوزان لترجيح الكميات المختلفة من السلع.

يوجد أربعة أرقام قياسية تجميعية مرجحة هي:

رقم لاسبير $= \frac{\sum X_i q_o}{\sum X_o q_o} \times 100$

حيث أن

$X_i\, q_o$ سعر سنة المقارنة مرجح بكمية الأساس

$X_o q_o$ سعر سنة الأساس مرجح بكمية الأساس

2- رقم باش Pasch's index

رقم باش $= \frac{\sum X_i q_i}{\sum X_o q_i} \times 100$

حيث أن

$\sum X_i q_i$ سعر سنة المقارنة مرجح بكمية سنة المقارنة

$\sum X_o q_i$ سعر سنة الأساس مرجح بكمية سنة المقارنة

(3) **رقم مارشال – إدجورث** $= \frac{\sum X_i (q_o + q_i)}{\sum x_o (q_o + q_i)} \times 100$

(4) رقم فيشر الأمثل Fisher's ideal index

رقم فيشر الأمثل $= \sqrt{\frac{x_i\, q_o}{x_o q_o} \times \frac{x_i q_i}{x_o q_i}}$

<u>ثالثاً:الرقم القياسي المتتابع:</u>

يتم فيه استخدام أساس متحرك بصورة مستمرة وعدم الإعتماد على فترة أساس واحدة ثابتة يقارن بها جميع الفترات الأخرى فإذا استخدمت إحدى الفترات كأساس لفترة ثانية تستخدم الفترة الثانية كأساس للفترة الثالثة وهكذا.

ومن مزايا هذا الرقم انه إذا استخدم لحساب سلسلة من الأرقام القياسية لأسعار مجموعة من السلع فإنه يمكن إضافة سلع جديدة أو التخلص من بعض السلع التي تفقد أهميتها بالنسبة على السلع الداخلة فى حساب الرقم القياسي.

$$\text{الرقم القياسي للقوة الشرائية للمزارعين} = \frac{\text{الرقم القياسي للأسعار التي تدفع للمزارعين}}{\text{الرقم القياسي للأسعار التي يدفعها المزارعون}}$$

يستخدم السعر المعدل Adjusted Price لسلعة، فى تحليل الأسعار وينحصر الغرض من استخدامه فى التخلص من اثر المستوى العام للأسعار حتى يمكن عمل المقارنة بين سعري سلعتين ويمكن حسابه كالآتي:

$$\text{السعر المعدل سنة } 1990 = \frac{\text{سعر السلعة في سنة } 1990}{\text{الرقم القياسي العام لأسعار سنة } 1990} \times \mathbf{100}$$

$$\text{الرقم القياسي للقوة الشرائية لسلعة ما} = \frac{\text{السعر النسبي للسلعة}}{\text{الرقم القياسي لأسعار السلع المماثلة}}$$

$$\text{القوة الشرائية للجنيه فى سنة المقارنة} = \frac{\text{الرقم القياسي للأسعار فى سنة الأساس}}{\text{الرقم القياسي للأسعار فى سنة المقارنة}} \times 100$$

سعر المساواة

يعرف سعر المساواة Parity Price لسلعة ما بأنه سعر الوحدة من السلعة التي يمكن المنتج من الحصول على نفس الكمية من السلع والخدمات التي كان عليها فى فترة الأساس مقابل وحدة واحدة من إنتاجه.

سعر الأساس المعدل لسلعة ما= $\dfrac{\text{متوسط سعر السلعة فى فترة عشر سنوات}}{\text{متوسط الأرقام القياسية لأسعار السلع الزراعية فى نفس الفترة}} \times 100$

توجد عدة اختبارات تجرى على الأرقام القياسية المقدرة بالطرق المختلفة لمعرفة مدى كفاءة تلك الأرقام وهى.

(1) اختبار الانعكاس الزماني قق الرقم القياسي اختبار الانعكاس الزماني إذا كان حاصل ضرب قيمة (باستخدام الفترة الزمنية a كأساس الفترة الزمنية b للمقارنة) فى قيمة الرقم القياسي للانعكاس(باستخدام الفترة الزمنية b كأساس والفترة الزمنية a للمقارنة) يساوى الواحد الصحيح.

(2) – اختبار الانعكاس المكاني هوالاختبار المستخدم عند مقارنة التغير بين المناطق المختلفة وبذلك تستخدم إحدى المناطق كأساس للمقارنة وتقارن بها باقي المناطق وينطبق على هذا الاختبار ما قيل عن اختبار الانعكاس الزماني.

(3)– اختبار الانعكاس المعامل: وفيشر أول من استخدم هذا الاختبار وتبني فكرته على أساس أن

القيمة = الكمية ×السعر

الرقم القياسي للأسعار × الرقم القياسي للكميات = الرقم القياسي للقيمة

ورقم فيشر هو المحقق أيضاً لهذا الاختبار

(1) الرقم القياسي لنفقة المعيشة

ويعرف أيضاً باسم الرقم القياسي لأسعار المستهلك أو الرقم القياسي لأسعار التجزئة، وهو يعبر عن التغير فى أسعار السلع والخدمات التي تستهلكها الأسرة خلال فترة زمنية محددة.

(2) الرقم القياسي لأسعار الجملة الرقم القياسي لأسعار الجملة هو أداة إحصائية لقياس متوسط التغير فى أسعار مجموعة معينة من مواد الجملة التي يتم تبادلها خلال فترة معينة من الزمن.

(1) صيغة لاسبير

$$\text{رقم لاسبير}=\frac{\sum X_i q_o}{\sum X_o q_o}\times 100$$

(2) صيغة باش

$$\text{رقم باش}=\frac{\sum X_i q_i}{\sum X_o q_i}\times 100$$

(3) صيغة الرقم القياسي المرجح بقيم سنة الأساس

$$=\frac{\sum \frac{x_i}{x_o} X_o q_o}{\sum X_o q_o}\times 100$$

(4) الــــرقم القياســـــي للأسعار $= \frac{\text{الرقم القياسي للقيمة}}{\text{الرقم القياسي لكمية}} \times 100$

(5) **الرقم القياسي للإنتاج** بالنسبة للرقم القياسي للإنتاج تختلف وحدات الإنتاج فيما بينها كالطن ، والمتر، والذراع واللتر وغيرها من وحدات قياس الإنتاج. وعند حسابنا للرقم القياسي للإنتاج لابد من إيجاد عامل مشترك بين هذه الوحدات المختلفة وهذا العامل فى الواقع هو أوزان الترجيح التي تعطي لكل نوع من أنواع الإنتاج

القوة الشرائية للعملة $= \frac{100}{\text{الرقم القياسي للأسعار}} \times 100$

الدخل الحقيقي $=$ الدخل فى سنة المقارنة $\times \frac{100}{\text{الرقم القياسي للأسعار}}$

تمارين

1 . إذا كانت لديك البيانات التالية:

السلعة	سعر الوحدة	الكمية المستهلكة بملايين الوحدات

	1990	2000	1990	2000
أ	3	5	2	3
ب	2	4	2	3
جـ	5	8	1	4
د	7	10	3	5
هـ	2	6	5	6

إحسب كلاً مما يلي:

1- الرقم القياسي التجمعي البسيط للاسعار.

2- الرقم القياسي التجمعي للأسعار المرجح بكميات سنة الأساس.

3- الرقم القياسي التجمعي للأسعار المرجح بكميات سنة المقارنة.

4- الرقم القياسي الأمثل للأسعار.

5- الرقم القياسي الأمثل للكميات.

6- فى هذا التمرين هل حل يحقق الرقم الأمثل خاصية الإنعكاس فى المعامل.

2 . يبين الجدول التالي أسعار وكميات أربع سلع غذائية فى كلاً من عامي 1990، 1994.

السلعة	**1990**		**1994**	
	السعر	الكيمة	السعر	الكيمة
أ	12	400	14	600
ب	14	280	15	420
جـ	20	800	19	640
د	40	250	32	300

المطلوب حساب الأرقام القياسية التالية للأسعار على إعتبار أن عام 1990 سنة اساس .

1- رقم لاسبير. 2- رقم باش.

3- رقم مارشال- إدجورث .

4- الرقم القياسي مرجحاً بالقيمة فى سنة الأساس باسعار سنة الأساس.

2- الرقم القياسي للمناسيب مرجحاً بالقيمة فى سنة المقارنة بأسعار سنة المقارنة.

3 . **إذا كان لديك البيانات التالية عن أسعار وكميات ثلاث سلع .**

السلعة	**1990**		**2000**	
	السعر	الكيمة	السعر	الكيمة
أ	2	5	3	7
ب	4	9	5	10
جـ	5	8	8	12

المطلوب حساب:

1- الرقم القياسي النسبي البسيط لأسعار 2000 باعتبار 1990 سنة اساس وبالطرق المختلفة.

2- الرقم القياسي النسبي لأسعار 2000 باعتبار 1990 سنة اساس مرجحاً بالكميات فى سنة المقارنة .

3-الرقم القياسي الامثل لفيشر

4 . إذا كان لدينا اسعار وكميات مبيعات ثلاث سلع فى عامى 1990، 1994 كما يلي:

السلعة	**السعر**		**الكمية**	
	1990	1994	1990	1994
أ	20	25	1200	1500
ب	15	18	1500	2000
جـ	10	12	2000	2700

المطلوب حساب ما يلي باعتبار أن 1990 سنة أساس .

1- رقم لاسبير للأسعار. 2- رقم باش للأسعار.

3- رقم لاسبير للكميات . 4- رقم باش للكميات.

4- الرقم القياسي للقيمة.

5 . **إذا كان متوسط الأجر فى الساعة فى احد المصانع الرقم القياسي لأسعار المستهلك خلال الفترة 1995 ، 1990 :**

السنة	الأجر بالساعة	الرقم القياسي للأسعار المستهلكة
1995	3.25	100
1996	3.75	104
1997	4.15	107
1998	4.50	110
1999	4.75	112

6 . الجدول التالي يوضح اسعار ثلاث سلع والكميات المستهلكة منها فى عامي 1990، 2000.

السلعة	1990		2000	
	السعر	الكيمة	السعر	الكيمة
أ	25	200	20	300
ب	60	150	40	100
جـ	35	20	30	40

المطلوب:باعتبار أن سنة 1990 هى سنة الأساس أحسب:

1- الرقم البسيط للأسعار.

2- الرقم القياسي الأمثل للأسعار.

7 . الجدول التالي يوضح القيم التكرارية للسلع المستوردة(بالمليون جنيه) والأوزان(بالألف طن) لعامة 2005، 2006.

السنة / السلع المستوردة	2005		2006	
	القيمة	الوزن	القيمة	الوزن
الآلات والأجهزة	54	14	64	16
معدات النقل	47	13	48	14
معادن عادية ومصنوعاتها	34	58	38	81
مواد غذائية	33	15	31	11
باقية السلع	65	18	68	18

المطلوب : باعتبار أن عام 2005 هو سنة الأساس إحسب:

1- الرقم البسيط للاسعار .

2- الرقم القياسي للأسعار المرجح بكميات الأساس (رقم لاسبير).

3- الرقم القياسي المرجح بكميات المقارنة (رقم باش).

4- الرقم القياسي الأمثل للأسعار (رقم فيشر

F Table for alpha=0.0

INF	120.00	60.00	40.00	30.00	24.00	20.00	15.00	12.00	10.0	9.00	8.00	7.00	6.00	5.00	4.00	3.00	2.00	1.00	df_2/df_1
63.33	63.06	62.79	62.53	62.26	62.00	61.74	61.22	60.71	60.2	59.86	59.44	58.91	58.20	57.24	55.83	53.59	49.50	39.8	1.00
9.49	9.48	9.47	9.47	9.46	9.45	9.44	9.42	9.41	9.39	9.38	9.37	9.35	9.33	9.29	9.24	9.16	9.00	8.53	2.00
5.13	5.14	5.15	5.16	5.17	5.18	5.18	5.20	5.22	5.23	5.24	5.25	5.27	5.28	5.31	5.34	5.39	5.46	5.54	3.00
3.76	3.78	3.79	3.80	3.82	3.83	3.84	3.87	3.90	3.92	3.94	3.95	3.98	4.01	4.05	4.11	4.19	4.32	4.54	4.00
3.11	3.12	3.14	3.16	3.17	3.19	3.21	3.24	3.27	3.30	3.32	3.34	3.37	3.40	3.45	3.52	3.62	3.78	4.06	5.00
2.72	2.74	2.76	2.78	2.80	2.82	2.84	2.87	2.90	2.94	2.96	2.98	3.01	3.05	3.11	3.18	3.29	3.46	3.78	6.00
2.47	2.49	2.51	2.54	2.56	2.58	2.59	2.63	2.67	2.70	2.72	2.75	2.78	2.83	2.88	2.96	3.07	3.26	3.59	7.00
2.29	2.32	2.34	2.36	2.38	2.40	2.42	2.46	2.50	2.54	2.56	2.59	2.62	2.67	2.73	2.81	2.92	3.11	3.46	8.00
2.16	2.18	2.21	2.23	2.25	2.28	2.30	2.34	2.38	2.42	2.44	2.47	2.51	2.55	2.61	2.69	2.81	3.01	3.36	9.00
2.06	2.08	2.11	2.13	2.16	2.18	2.20	2.24	2.28	2.32	2.35	2.38	2.41	2.46	2.52	2.61	2.73	2.92	3.29	10.0
1.97	2.00	2.03	2.05	2.08	2.10	2.12	2.17	2.21	2.25	2.27	2.30	2.34	2.39	2.45	2.54	2.66	2.8	3.23	11.0
1.90	1.93	1.96	1.99	2.01	2.04	2.06	2.10	2.15	2.19	2.21	2.24	2.28	2.33	2.39	2.48	2.61	2.81	3.18	12.0
1.85	1.88	1.90	1.93	1.96	1.98	2.01	2.05	2.10	2.14	2.16	2.20	2.23	2.28	2.35	2.43	2.56	2.76	3.14	13.0
1.80	1.83	1.86	1.89	1.91	1.94	1.96	2.01	2.05	2.10	2.12	2.15	2.19	2.24	2.31	2.39	2.52	2.73	3.10	14.0
1.76	1.79	1.82	1.85	1.87	1.90	1.92	1.97	2.02	2.06	2.09	2.12	2.16	2.21	2.27	2.36	2.49	2.70	3.07	15.0
1.72	1.75	1.78	1.81	1.84	1.87	1.89	1.94	1.99	2.03	2.06	2.09	2.13	2.18	2.24	2.33	2.46	2.67	3.05	16.0
1.69	1.72	1.75	1.78	1.81	1.84	1.86	1.91	1.96	2.00	2.03	2.06	2.10	2.15	2.22	2.31	2.44	2.64	3.03	17.0
1.66	1.69	1.72	1.75	1.78	1.81	1.84	1.89	1.93	1.98	2.00	2.04	2.08	2.13	2.20	2.29	2.42	2.62	3.01	18.0
1.63	1.67	1.70	1.73	1.76	1.79	1.81	1.86	1.91	1.96	1.98	2.02	2.06	2.11	2.18	2.27	2.40	2.61	2.99	19.0
1.61	1.64	1.68	1.71	1.74	1.77	1.79	1.84	1.89	1.94	1.96	2.00	2.04	2.09	2.16	2.25	2.38	2.59	2.97	20.0
1.59	1.62	1.66	1.69	1.72	1.75	1.78	1.83	1.87	1.92	1.95	1.98	2.02	2.08	2.14	2.23	2.36	2.57	2.96	21.0
1.57	1.60	1.64	1.67	1.70	1.73	1.76	1.81	1.86	1.90	1.93	1.97	2.01	2.06	2.13	2.22	2.35	2.56	2.95	22.0
1.55	1.59	1.62	1.66	1.69	1.72	1.74	1.80	1.84	1.89	1.92	1.95	1.99	2.05	2.11	2.21	2.34	2.55	2.94	23.0
1.53	1.57	1.61	1.64	1.67	1.70	1.73	1.78	1.83	1.88	1.91	1.94	1.98	2.04	2.10	2.19	2.33	2.54	2.93	24.0
1.52	1.56	1.59	1.63	1.66	1.69	1.72	1.77	1.82	1.87	1.89	1.93	1.97	2.02	2.09	2.18	2.32	2.53	2.92	25.0
1.50	1.54	1.58	1.61	1.65	1.68	1.71	1.76	1.81	1.86	1.88	1.92	1.96	2.01	2.08	2.17	2.31	2.52	2.91	26.0
1.49	1.53	1.57	1.60	1.64	1.67	1.70	1.75	1.80	1.85	1.87	1.91	1.95	2.00	2.07	2.17	2.30	2.51	2.90	27.0
1.48	1.52	1.56	1.59	1.63	1.66	1.69	1.74	1.79	1.84	1.87	1.90	1.94	2.00	2.06	2.16	2.29	2.50	2.89	28.0
1.47	1.51	1.55	1.58	1.62	1.65	1.68	1.73	1.78	1.83	1.86	1.89	1.93	1.99	2.06	2.15	2.28	2.50	2.89	29.0
1.46	1.50	1.54	1.57	1.61	1.64	1.67	1.72	1.77	1.82	1.85	1.88	1.93	1.98	2.05	2.14	2.28	2.49	2.88	30.0
1.38	1.42	1.47	1.51	1.54	1.57	1.61	1.66	1.71	1.76	1.79	1.83	1.87	1.93	2.00	2.09	2.23	2.44	2.84	40.0
1.29	1.35	1.40	1.44	1.48	1.51	1.54	1.60	1.66	1.71	1.74	1.77	1.82	1.87	1.95	2.04	2.18	2.39	2.79	60.0
1.19	1.26	1.32	1.37	1.41	1.45	1.48	1.55	1.60	1.65	1.68	1.72	1.77	1.82	1.90	1.99	2.13	2.35	2.75	120.

F Table for alpha=.01

INF	120	60	40	30	24	20	15	12	10	9	8	7	6	5	4	3	2	1	df2/df1
63.33	63.06	62.79	62.53	62.26	62.00	61.74	61.22	60.705	60.19	59.858	59.44	58.91	58.20	57.24	55.83	53.59	49.5	39.86	1
63.33	**63.06**	**62.79**	**62.53**	**62.26**	**62.00**	**61.74**	**61.22**	**60.71**	**60.19**	**59.86**	**59.44**	**58.91**	**58.20**	**57.24**	**55.83**	**53.59**	**49.50**	**39.86**	**1.00**
9.49	**9.48**	**9.47**	**9.47**	**9.46**	**9.45**	**9.44**	**9.42**	**9.41**	**9.39**	**9.38**	**9.37**	**9.35**	**9.33**	**9.29**	**9.24**	**9.16**	**9.00**	**8.53**	**2.00**
5.13	**5.14**	**5.15**	**5.16**	**5.17**	**5.18**	**5.18**	**5.20**	**5.22**	**5.23**	**5.24**	**5.25**	**5.27**	**5.28**	**5.31**	**5.34**	**5.39**	**5.46**	**5.54**	**3.00**
3.76	**3.78**	**3.79**	**3.80**	**3.82**	**3.83**	**3.84**	**3.87**	**3.90**	**3.92**	**3.94**	**3.95**	**3.98**	**4.01**	**4.05**	**4.11**	**4.19**	**4.32**	**4.54**	**4.00**
3.11	**3.12**	**3.14**	**3.16**	**3.17**	**3.19**	**3.21**	**3.24**	**3.27**	**3.30**	**3.32**	**3.34**	**3.37**	**3.40**	**3.45**	**3.52**	**3.62**	**3.78**	**4.06**	**5.00**
2.72	**2.74**	**2.76**	**2.78**	**2.80**	**2.82**	**2.84**	**2.87**	**2.90**	**2.94**	**2.96**	**2.98**	**3.01**	**3.05**	**3.11**	**3.18**	**3.29**	**3.46**	**3.78**	**6.00**
2.47	**2.49**	**2.51**	**2.54**	**2.56**	**2.58**	**2.59**	**2.63**	**2.67**	**2.70**	**2.72**	**2.75**	**2.78**	**2.83**	**2.88**	**2.96**	**3.07**	**3.26**	**3.59**	**7.00**
2.29	**2.32**	**2.34**	**2.36**	**2.38**	**2.40**	**2.42**	**2.46**	**2.50**	**2.54**	**2.56**	**2.59**	**2.62**	**2.67**	**2.73**	**2.81**	**2.92**	**3.11**	**3.46**	**8.00**
2.16	**2.18**	**2.21**	**2.23**	**2.25**	**2.28**	**2.30**	**2.34**	**2.38**	**2.42**	**2.44**	**2.47**	**2.51**	**2.55**	**2.61**	**2.69**	**2.81**	**3.01**	**3.36**	**9.00**
2.06	**2.08**	**2.11**	**2.13**	**2.16**	**2.18**	**2.20**	**2.24**	**2.28**	**2.32**	**2.35**	**2.38**	**2.41**	**2.46**	**2.52**	**2.61**	**2.73**	**2.92**	**3.29**	**10.00**
1.97	**2.00**	**2.03**	**2.05**	**2.08**	**2.10**	**2.12**	**2.17**	**2.21**	**2.25**	**2.27**	**2.30**	**2.34**	**2.39**	**2.45**	**2.54**	**2.66**	**2.86**	**3.23**	**11.00**
1.90	**1.93**	**1.96**	**1.99**	**2.01**	**2.04**	**2.06**	**2.10**	**2.15**	**2.19**	**2.21**	**2.24**	**2.28**	**2.33**	**2.39**	**2.48**	**2.61**	**2.81**	**3.18**	**12.00**
1.85	**1.88**	**1.90**	**1.93**	**1.96**	**1.98**	**2.01**	**2.05**	**2.10**	**2.14**	**2.16**	**2.20**	**2.23**	**2.28**	**2.35**	**2.43**	**2.56**	**2.76**	**3.14**	**13.00**
1.80	**1.83**	**1.86**	**1.89**	**1.91**	**1.94**	**1.96**	**2.01**	**2.05**	**2.10**	**2.12**	**2.15**	**2.19**	**2.24**	**2.31**	**2.39**	**2.52**	**2.73**	**3.10**	**14.00**
1.76	**1.79**	**1.82**	**1.85**	**1.87**	**1.90**	**1.92**	**1.97**	**2.02**	**2.06**	**2.09**	**2.12**	**2.16**	**2.21**	**2.27**	**2.36**	**2.49**	**2.70**	**3.07**	**15.00**
1.72	**1.75**	**1.78**	**1.81**	**1.84**	**1.87**	**1.89**	**1.94**	**1.99**	**2.03**	**2.06**	**2.09**	**2.13**	**2.18**	**2.24**	**2.33**	**2.46**	**2.67**	**3.05**	**16.00**
1.69	**1.72**	**1.75**	**1.78**	**1.81**	**1.84**	**1.86**	**1.91**	**1.96**	**2.00**	**2.03**	**2.06**	**2.10**	**2.15**	**2.22**	**2.31**	**2.44**	**2.64**	**3.03**	**17.00**
1.66	**1.69**	**1.72**	**1.75**	**1.78**	**1.81**	**1.84**	**1.89**	**1.93**	**1.98**	**2.00**	**2.04**	**2.08**	**2.13**	**2.20**	**2.29**	**2.42**	**2.62**	**3.01**	**18.00**
1.63	**1.67**	**1.70**	**1.73**	**1.76**	**1.79**	**1.81**	**1.86**	**1.91**	**1.96**	**1.98**	**2.02**	**2.06**	**2.11**	**2.18**	**2.27**	**2.40**	**2.61**	**2.99**	**19.00**
1.61	**1.64**	**1.68**	**1.71**	**1.74**	**1.77**	**1.79**	**1.84**	**1.89**	**1.94**	**1.96**	**2.00**	**2.04**	**2.09**	**2.16**	**2.25**	**2.38**	**2.59**	**2.97**	**20.00**
1.59	**1.62**	**1.66**	**1.69**	**1.72**	**1.75**	**1.78**	**1.83**	**1.87**	**1.92**	**1.95**	**1.98**	**2.02**	**2.08**	**2.14**	**2.23**	**2.36**	**2.57**	**2.96**	**21.00**
1.57	**1.60**	**1.64**	**1.67**	**1.70**	**1.73**	**1.76**	**1.81**	**1.86**	**1.90**	**1.93**	**1.97**	**2.01**	**2.06**	**2.13**	**2.22**	**2.35**	**2.56**	**2.95**	**22.00**
1.55	**1.59**	**1.62**	**1.66**	**1.69**	**1.72**	**1.74**	**1.80**	**1.84**	**1.89**	**1.92**	**1.95**	**1.99**	**2.05**	**2.11**	**2.21**	**2.34**	**2.55**	**2.94**	**23.00**
1.53	**1.57**	**1.61**	**1.64**	**1.67**	**1.70**	**1.73**	**1.78**	**1.83**	**1.88**	**1.91**	**1.94**	**1.98**	**2.04**	**2.10**	**2.19**	**2.33**	**2.54**	**2.93**	**24.00**
1.52	**1.56**	**1.59**	**1.63**	**1.66**	**1.69**	**1.72**	**1.77**	**1.82**	**1.87**	**1.89**	**1.93**	**1.97**	**2.02**	**2.09**	**2.18**	**2.32**	**2.53**	**2.92**	**25.00**
1.50	**1.54**	**1.58**	**1.61**	**1.65**	**1.68**	**1.71**	**1.76**	**1.81**	**1.86**	**1.88**	**1.92**	**1.96**	**2.01**	**2.08**	**2.17**	**2.31**	**2.52**	**2.91**	**26.00**
1.49	**1.53**	**1.57**	**1.60**	**1.64**	**1.67**	**1.70**	**1.75**	**1.80**	**1.85**	**1.87**	**1.91**	**1.95**	**2.00**	**2.07**	**2.17**	**2.30**	**2.51**	**2.90**	**27.00**
1.48	**1.52**	**1.56**	**1.59**	**1.63**	**1.66**	**1.69**	**1.74**	**1.79**	**1.84**	**1.87**	**1.90**	**1.94**	**2.00**	**2.06**	**2.16**	**2.29**	**2.50**	**2.89**	**28.00**
1.47	**1.51**	**1.55**	**1.58**	**1.62**	**1.65**	**1.68**	**1.73**	**1.78**	**1.83**	**1.86**	**1.89**	**1.93**	**1.99**	**2.06**	**2.15**	**2.28**	**2.50**	**2.89**	**29.00**
1.46	**1.50**	**1.54**	**1.57**	**1.61**	**1.64**	**1.67**	**1.72**	**1.77**	**1.82**	**1.85**	**1.88**	**1.93**	**1.98**	**2.05**	**2.14**	**2.28**	**2.49**	**2.88**	**30.00**
1.38	**1.42**	**1.47**	**1.51**	**1.54**	**1.57**	**1.61**	**1.66**	**1.71**	**1.76**	**1.79**	**1.83**	**1.87**	**1.93**	**2.00**	**2.09**	**2.23**	**2.44**	**2.84**	**40.00**
1.29	**1.35**	**1.40**	**1.44**	**1.48**	**1.51**	**1.54**	**1.60**	**1.66**	**1.71**	**1.74**	**1.77**	**1.82**	**1.87**	**1.95**	**2.04**	**2.18**	**2.39**	**2.79**	**60.00**
1.19	**1.26**	**1.32**	**1.37**	**1.41**	**1.45**	**1.48**	**1.55**	**1.60**	**1.65**	**1.68**	**1.72**	**1.77**	**1.82**	**1.90**	**1.99**	**2.13**	**2.35**	**2.75**	**120.0**

F Table for alpha=.05

INF	120	60	40	30	24	20	15	12	10	9	8	7	6	5	4	3	2	1	df2/df1
254.3144	253.2529	252.1957	251.1432	250.0951	249.0518	248.0131	245.9499	243.9060	241.8817	240.5433	238.8827	236.7684	233.9860	230.1619	224.5832	215.7073	199.5000	161.4476	1
19.4957	19.4874	19.4791	19.4707	19.4624	19.4541	19.4458	19.4291	19.4125	19.3959	19.3848	19.3710	19.3532	19.3295	19.2964	19.2468	19.1643	19.0000	18.5128	2
8.5264	8.5494	8.5720	8.5944	8.6166	8.6385	8.6602	8.7029	8.7446	8.7855	8.8123	8.8452	8.8867	8.9406	9.0135	9.1172	9.2766	9.5521	10.1280	3
5.6281	5.6581	5.6877	5.7170	5.7459	5.7744	5.8025	5.8578	5.9117	5.9644	5.9988	6.0410	6.0942	6.1631	6.2561	6.3882	6.5914	6.9443	7.7086	4
4.3650	4.3985	4.4314	4.4638	4.4957	4.5272	4.5581	4.6188	4.6777	4.7351	4.7725	4.8183	4.8759	4.9503	5.0503	5.1922	5.4095	5.7861	6.6079	5
3.6689	3.7047	3.7398	3.7743	3.8082	3.8415	3.8742	3.9381	3.9999	4.0600	4.0990	4.1468	4.2067	4.2839	4.3874	4.5337	4.7571	5.1433	5.9874	6
3.2298	3.2674	3.3043	3.3404	3.3758	3.4105	3.4445	3.5107	3.5747	3.6365	3.6767	3.7257	3.7870	3.8660	3.9715	4.1203	4.3468	4.7374	5.5914	7
2.9276	2.9669	3.0053	3.0428	3.0794	3.1152	3.1503	3.2184	3.2839	3.3472	3.3881	3.4381	3.5005	3.5806	3.6875	3.8379	4.0662	4.4590	5.3177	8
2.7067	2.7475	2.7872	2.8259	2.8637	2.9005	2.9365	3.0061	3.0729	3.1373	3.1789	3.2296	3.2927	3.3738	3.4817	3.6331	3.8625	4.2565	5.1174	9
2.5379	2.5801	2.6211	2.6609	2.6996	2.7372	2.7740	2.8450	2.9130	2.9782	3.0204	3.0717	3.1355	3.2172	3.3258	3.4780	3.7083	4.1028	4.9646	10
2.4045	2.4480	2.4901	2.5309	2.5705	2.6090	2.6464	2.7186	2.7876	2.8536	2.8962	2.9480	3.0123	3.0946	3.2039	3.3567	3.5874	3.9823	4.8443	11
2.2962	2.3410	2.3842	2.4259	2.4663	2.5055	2.5436	2.6169	2.6866	2.7534	2.7964	2.8486	2.9134	2.9961	3.1059	3.2592	3.4903	3.8853	4.7472	12
2.2064	2.2524	2.2966	2.3392	2.3803	2.4202	2.4589	2.5331	2.6037	2.6710	2.7144	2.7669	2.8321	2.9153	3.0254	3.1791	3.4105	3.8056	4.6672	13
2.1307	2.1778	2.2229	2.2664	2.3082	2.3487	2.3879	2.4630	2.5342	2.6022	2.6458	2.6987	2.7642	2.8477	2.9582	3.1122	3.3439	3.7389	4.6001	14
2.0658	2.1141	2.1601	2.2043	2.2468	2.2878	2.3275	2.4034	2.4753	2.5437	2.5876	2.6408	2.7066	2.7905	2.9013	3.0556	3.2874	3.6823	4.5431	15
2.0096	2.0589	2.1058	2.1507	2.1938	2.2354	2.2756	2.3522	2.4247	2.4935	2.5377	2.5911	2.6572	2.7413	2.8524	3.0069	3.2389	3.6337	4.4940	16
1.9604	2.0107	2.0584	2.1040	2.1477	2.1898	2.2304	2.3077	2.3807	2.4499	2.4943	2.5480	2.6143	2.6987	2.8100	2.9647	3.1968	3.5915	4.4513	17
1.9168	1.9681	2.0166	2.0629	2.1071	2.1497	2.1906	2.2686	2.3421	2.4117	2.4563	2.5102	2.5767	2.6613	2.7729	2.9277	3.1599	3.5546	4.4139	18

1.8780	1.9302	1.9795	2.0264	2.0712	2.1141	2.1555	2.2341	2.3080	2.3779	2.4227	2.4768	2.5435	2.6283	2.7401	2.8951	3.1274	3.5219	4.3807	19
1.8432	1.8963	1.9464	1.9938	2.0391	2.0825	2.1242	2.2033	2.2776	2.3479	2.3928	2.4471	2.5140	2.5990	2.7109	2.8661	3.0984	3.4928	4.3512	20
1.8117	1.8657	1.9165	1.9645	2.0102	2.0540	2.0960	2.1757	2.2504	2.3210	2.3660	2.4205	2.4876	2.5727	2.6848	2.8401	3.0725	3.4668	4.3248	21
1.7831	1.8380	1.8894	1.9380	1.9842	2.0283	2.0707	2.1508	2.2258	2.2967	2.3419	2.3965	2.4638	2.5491	2.6613	2.8167	3.0491	3.4434	4.3009	22
1.7570	1.8128	1.8648	1.9139	1.9605	2.0050	2.0476	2.1282	2.2036	2.2747	2.3201	2.3748	2.4422	2.5277	2.6400	2.7955	3.0280	3.4221	4.2793	23
1.7330	1.7896	1.8424	1.8920	1.9390	1.9838	2.0267	2.1077	2.1834	2.2547	2.3002	2.3551	2.4226	2.5082	2.6207	2.7763	3.0088	3.4028	4.2597	24
1.7110	1.7684	1.8217	1.8718	1.9192	1.9643	2.0075	2.0889	2.1649	2.2365	2.2821	2.3371	2.4047	2.4904	2.6030	2.7587	2.9912	3.3852	4.2417	25
1.6906	1.7488	1.8027	1.8533	1.9010	1.9464	1.9898	2.0716	2.1479	2.2197	2.2655	2.3205	2.3883	2.4741	2.5868	2.7426	2.9752	3.3690	4.2252	26
1.6717	1.7306	1.7851	1.8361	1.8842	1.9299	1.9736	2.0558	2.1323	2.2043	2.2501	2.3053	2.3732	2.4591	2.5719	2.7278	2.9604	3.3541	4.2100	27
1.6541	1.7138	1.7689	1.8203	1.8687	1.9147	1.9586	2.0411	2.1179	2.1900	2.2360	2.2913	2.3593	2.4453	2.5581	2.7141	2.9467	3.3404	4.1960	28
1.6376	1.6981	1.7537	1.8055	1.8543	1.9005	1.9446	2.0275	2.1045	2.1768	2.2229	2.2783	2.3463	2.4324	2.5454	2.7014	2.9340	3.3277	4.1830	29
1.6223	1.6835	1.7396	1.7918	1.8409	1.8874	1.9317	2.0148	2.0921	2.1646	2.2107	2.2662	2.3343	2.4205	2.5336	2.6896	2.9223	3.3158	4.1709	30
1.5089	1.5766	1.6373	1.6928	1.7444	1.7929	1.8389	1.9245	2.0035	2.0772	2.1240	2.1802	2.2490	2.3359	2.4495	2.6060	2.8387	3.2317	4.0847	40
1.3893	1.4673	1.5343	1.5943	1.6491	1.7001	1.7480	1.8364	1.9174	1.9926	2.0401	2.0970	2.1665	2.2541	2.3683	2.5252	2.7581	3.1504	4.0012	60
1.2539	1.3519	1.4290	1.4952	1.5543	1.6084	1.6587	1.7505	1.8337	1.9105	1.9588	2.0164	2.0868	2.1750	2.2899	2.4472	2.6802	3.0718	3.9201	120
1.0000	1.2214	1.3180	1.3940	1.4591	1.5173	1.5705	1.6664	1.7522	1.8307	1.8799	1.9384	2.0096	2.0986	2.2141	2.3719	2.6049	2.9957	3.8415	inf

جدول Z

0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	Z
0.5359	0.5319	0.5279	0.5239	0.5199	0.5160	0.5120	0.5080	0.5040	0.5000	0.0
0.5753	0.5714	0.5675	0.5636	0.5596	0.5557	0.5517	0.5478	0.5438	0.5398	0.1
0.6141	0.6103	0.6064	0.6026	0.5987	0.5948	0.5910	0.5871	0.5832	0.5793	0.2
0.6517	0.6480	0.6443	0.6406	0.6368	0.6331	0.6293	0.6255	0.6217	0.6179	0.3
0.6879	0.6844	0.6808	0.6772	0.6736	0.6700	0.6664	0.6628	0.6591	0.6554	0.4
0.7224	0.7190	0.7157	0.7123	0.7088	0.7054	0.7019	0.6985	0.6950	0.6915	0.5
0.7549	0.7517	0.7486	0.7454	0.7422	0.7389	0.7357	0.7324	0.7291	0.7257	0.6
0.7852	0.7823	0.7794	0.7764	0.7734	0.7704	0.7673	0.7642	0.7611	0.7580	0.7
0.8133	0.8106	0.8078	0.8051	0.8023	0.7995	0.7967	0.7939	0.7910	0.7881	0.8
0.8389	0.8365	0.8340	0.8315	0.8289	0.8264	0.8238	0.8212	0.8186	0.8159	0.9
0.8621	0.8599	0.8577	0.8554	0.8531	0.8508	0.8485	0.8461	0.8438	0.8413	1.0
0.8830	0.8810	0.8790	0.8770	0.8749	0.8729	0.8708	0.8686	0.8665	0.8643	1.1
0.9015	0.8997	0.8980	0.8962	0.8944	0.8925	0.8907	0.8888	0.8869	0.8849	1.2
0.9177	0.9162	0.9147	0.9131	0.9115	0.9099	0.9082	0.9066	0.9049	0.9032	1.3
0.9319	0.9306	0.9292	0.9279	0.9265	0.9251	0.9236	0.9222	0.9207	0.9192	1.4
0.9441	0.9429	0.9418	0.9406	0.9394	0.9382	0.9370	0.9357	0.9345	0.9332	1.5
0.9545	0.9535	0.9525	0.9515	0.9505	0.9495	0.9484	0.9474	0.9463	0.9452	1.6
0.9633	0.9625	0.9616	0.9608	0.9599	0.9591	0.9582	0.9573	0.9564	0.9554	1.7
0.9706	0.9699	0.9693	0.9686	0.9678	0.9671	0.9664	0.9656	0.9649	0.9641	1.8
0.9767	0.9761	0.9756	0.9750	0.9744	0.9738	0.9732	0.9726	0.9719	0.9713	1.9
0.9817	0.9812	0.9808	0.9803	0.9798	0.9793	0.9788	0.9783	0.9778	0.9772	2.0
0.9857	0.9854	0.9850	0.9846	0.9842	0.9838	0.9834	0.9830	0.9826	0.9821	2.1
0.9890	0.9887	0.9884	0.9881	0.9878	0.9875	0.9871	0.9868	0.9864	0.9861	2.2

0.9916	0.9913	0.9911	0.9909	0.9906	0.9904	0.9901	0.9898	0.9896	0.9893	2.3
0.9936	0.9934	0.9932	0.9931	0.9929	0.9927	0.9925	0.9922	0.9920	0.9918	2.4
0.9952	0.9951	0.9949	0.9948	0.9946	0.9945	0.9943	0.9941	0.9940	0.9938	2.5
0.9964	0.9963	0.9962	0.9961	0.9960	0.9959	0.9957	0.9956	0.9955	0.9953	2.6
0.9974	0.9973	0.9972	0.9971	0.9970	0.9969	0.9968	0.9967	0.9966	0.9965	2.7
0.9981	0.9980	0.9979	0.9979	0.9978	0.9977	0.9977	0.9976	0.9975	0.9974	2.8
0.9986	0.9986	0.9985	0.9985	0.9984	0.9984	0.9983	0.9982	0.9982	0.9981	2.9
0.9990	0.9990	0.9989	0.9989	0.9989	0.9988	0.9988	0.9987	0.9987	0.9987	3.0

المراجع

أولاً المراجع العربية :

1. احمد عبادة سرحان (دكتور) ، صلاح الدين طلبة (دكتور) ، أسس الأحصاء ، دار الكتب الجامعية ، الطبعة الاولي ، 1968.
2. السعيد عبد الحميد البسيوني (دكتور)، محمد كامل ريحان (دكتور) ، مذكرات في الإحصاء ، كلية الزراعة ، جامعة عين شمس، سنوات مختلفة.
3. دومينيك سالفاتور ، نظريات ومسائل في الأحصاء والأقتصاد القياسي ، سلسلة ملخصات شوم ، ترجمة (د.سعدية حافظ) ، دار ماكجروهيل للنشر ، 1982.
4. دومينيك سالفاتور ، نظريات ومسائل في الأحصاء ، سلسلة ملخصات شوم ، ترجمة (د.شعبان عبد الحميد شعبان) ، دار ماكجروهيل للنشر ، 1981.
5. عباس السيد (دكتور) ، الاقتصاد القياسي ، دار الجامعات المصرية ، الأسكندرية ، 1988.
6. عبد القادر محمد عبد القادر (دكتور) ، طرق قياس العلاقات الاقتصادية – مع تطبيقات الحاسب الألكتروني ، دار الجامعات المصرية ، الأسكندرية ، 1990.
7. فريد الحسيني عبد البديع (دكتور) ، المدخل في نظرية التقديرات وتطبيقاتها ، مكتبة عين شمس ، 1978.
8. مجدي الشوربجي (دكتور) ، الاقتصاد القياسي – النظرية والتطبيق ، كلية التجارة وإدارة الأعمال ، جامعة حلوان.
9. محمد كامل ريحان (دكتور) ، السعيد عبد الحميد البسيوني (دكتور)، مذكرات في الاقتصاد القياسي ، كلية الزراعة ، جامعة عين شمس، سنوات مختلفة.
10. محمد كامل ريحان (دكتور) ، وآخر – الطرق الكمية في العلوم الاقتصادية والادارية ، جامعة الإمارات العربية ، 1983.
11. محمد كامل ريحان ، السعيد عبد الحميد البسيوني – الاقتصاد القياسي– برنامج التعليم المفتوح– كلية الزراعة– جامعة عين شمس– 2007.

12. محمد كامل ريحان ، السعيد عبد الحميد البسيوني – مباديء علم الاحصاء- مذكرات دراسية – قسم الاقتصاد الزراعي- كلية الزراعة- جامعة عين شمس- 2007.

ثانياً : المراجع الاجنبية :

1. Dunn, W.N. : Public Policy Analysis : An Introduction, Prentice- Hall Inc., Englewood Cliffs, New Jersey, 1981.
2. Gujarati , Damodar , Basic Econometrics, McGraw-Hell , International Book Company , London , 1983.
3. Johnston , J. Econometric Methods , Ins., Tokyo , 1984.
4. Johnston, A.C., Johnston, M.B. And Buse , R.C: Econometrics : Basic And Applied, Macmillan Publishing Company , New York, 1987.
5. Kementa, J. Elements Of Econometrics, The Macmillan Company, New York, 1971.
6. Maddala, G.S. : Econometrics, Mcgraw- Hill Book Company, New York, 1977.
7. Pindych , R & Rubin field , D., Econometric Models And Economic Forecasts, (Second Edition) , New York: McGraw-Hell , 1983.
8. Thomas E. Kurtz, Basic Statistics, Prentice – hall of India, New Delhi, 1964 .
9. William L. Hays and Robert L. Winker, Statistics , Holt, Rinehart and Winston, Inc, New York , 1970 .
10. Norman R. Kurtz , Introduction to Social Statistics, International Book Company , 1983.
11. Sendecer G.W. & W. G. Cochran, Statistical methods, Taw state University Press, UAS , 1967.