اسس البحث العلمي

الفصل الاول

قسم الرياضيات

المرحلة الثانية

مدرسة المادة: اريج صلاح محمد

الفصل الاول : العلم و البحث العلمي

١- ١العلم و اهدافه

العلم : نشاط يهدف الى زيادة قدرة الانسان على السيطرة على الطبيعة من خلال فهم الظواهر وايجاد العلاقات التي تحكم هذه الظواهر.

أهداف العلم :

١. **الفهم** :وهو الغرض الاساسي للعلم،فالعلم كنشاط انساني يهدف الى فهم الظواهر المختلفة وتفسيرها ، اي فم الاسباب والعوامل التي ادت حدوث الظاهرة وليس الاكتفاء بتعداد صفاتها وخصائصها بل التعرف على الظاهرة بالظواهر الاخرى التي ادت الى وقوعها.
٢. **التنبؤ**: وهي قدرة الباحث على ان يستنتج من فهمه للظاهرة وقوانينها نتائج أخرى مرتبطة بهذا الفهم فالتبؤ هو تصور النتائج التي يمكن ان تحدث اذا طبقنا القوانين التي اكتشفناها على مواقف جديدة. وتزداد القدرة على التنبؤ بزيادة درجة التشابه بين الظاهرة التي درسناها وبين الظواهر التي سنطبق عليها فهمنا للظاهرة الاولى واي استنتاج لايعد صحيحا الا اذا تم اثباته تدريجيا .
٣. **الضبط والتحكم**: يساعد العلم والبحث العلمي في عملية الضبط والتحكم في الظواهر والاحداث والوقائع والسيطرة عليها وتوجيهها التوجيه المطلوب واستغلال النتائج لخدمة الانسانية ،وبذلك تمكن الانسان بفضل العلم من التحكم والضبط مثلا في مسار الانهار الكبرى و موارد المياه الاخرى واستغلال ذلك لخدمة البشرية،كما اصبح اليوم بفضل العلم امكانية التحكم بالامراض والسلوكيات البشرية وضبطها وتوجيهها نحو الخير.

١-٢ مفهوم البحث العلمي و اهميته

مفهومه : هو وسيلة للاستعلام و الاستقصاء المنضم الدقيق الذي يقوم به الباحث بهدف الوصول الى حلول لمشكلات معينة، و اكتشاف معلومات او علاقات جديدة، فضلا عن تطوير او تصحيح المعلومات الموجودة فعلا، على ان يتبع في هذا الفحص و الاستعلام خطوات المنهج العلمي.

و يمكن تعريف البحث العلمي بانه عرض مفصل او دراسة متعمقة تمثل كشفا لحقيقة جديدة او التاكيد على حقيقة قديمة سبق بحثها واضافة شيء جديد لها ، او حل لمشكلة كان قد تعهد بها شخص باحث بتقصيها و كشف حلها.

اهميته :يشكل الاهتمام بالبحث العلمي اتجاها عاما تاخذ به الدول المتقدمة على نطاق واسع وتسعى الدول النامية الى التوصل به لمجابهة مشكلاتها المختلفة وتطوير اوضاعها الاقتصادية والاجتماعية. وجاء اهتمام المتزايد بالبحث العلمي جزءا من الاتجاه العام، وان التربية هي التطبيق الاساسي لتحقيق القوة الذاتية لجميع أفراد المجتمع ، وان البحث العلمي هو وسيلة التربية لتحسين أساليبها والنهوض بمستقبلها ، ومواجهة المطالب المتعددة الملقاة عليها. ومن هنا تتسع مجالات البحث العلمي في التربية وتتعدد شاملة كل مدخلاتها ومخرجاتها وكل العوامل النفسية والاجتماعية والاقتصادية التي تؤثر في كفايتها وجودتها.ومن ذلك دراسة شخصيات الصغار والكبار وحالتهم والفروق بينهم وطرق تعلمهم ،والظروف البيئية التي تساعد على تحقيق تعلم اكثر ايجابية وأفضل أثرا،والبحث في صياغة الاهداف التربوية والوسائل التي تكفل تحقيقها دون الضياع أو الفقدان ،واخضاع المقررات الدراسية،والمناهج وطرائق التدريس والعمليات الاشرافية والادارية للتقويم والدراسة من اجل تطويرها والكشف عن الجديد فيها والبحث في نظم اعداد المعلم وتدرسه واقتصاديات التعليم وانواع المهام الدراسية وفي علاقة التعليم ككل بمطالب التنمية الاقتصادية والاجتماعية وفلسفة المجتمع وتطلعاته في المستقبل القريب و البعيد. و كل ذلك لغرض مشاركة العاملين في ميدان العملية التربوية و الحاجة الى تدريبهم على اساس البحث التربوي لكي يضطلعوا بدورهم المطلوب.

ان البحث العلمي ضرورة قائمة لكل انسان مهما كان علمه او مركزه لان مشكلات الحياة اليومية تتطلب تفكيرا ومنهجا علميا لحلها . وينفرد الانسان عن سائر أعضاء المملكة الحيوانية بأنه الكائن الوحيد في هذه المملكة الذي يتعامل مع الرموز –اللغة اساسا – مما ادى هذا العامل الى بناء حضارته وثقافته ونقلها عبر الاجيال ،كما ان قدراته مع المحسوس العياني والمجرد ينفرد بها عن باقي أفراد المملكة الحيوانية فضلا عن قدرته عن العمل.

يمكن تلخيص اهمية البحث العلمي بالنسبة للطلبة بالنقاط التالية:

١. أثراء معلومات الطالب في مواضيع معينة.
٢. تعود الطلبة على الاعتماد على النفس في دراسة المشكلات و اصدار الاحكام بشأنها.
٣. أتباع الاساليب و القواعد العلمية المعتمدة في كتابة البحوث.
٤. التعود على استعمال الوثائق و الكتب ومصادر المعلومات و الربط بينها للوصول الى نتائج جديدة.
٥. تدريب الطلاب على معالجة الموضوعات بنزاهة و موضوعية و نظام في العمل.

١-٣ خصائص البحث العلمي :

يتميز البحث العلمي بعدة خصائص اساسية اهمها :

١. الموضوعية : و تعني الموضوعية هنا، ان الباحث يلتزم في بحثه المقاييس العلمية الدقيقة ، ويقوم بأدراج الحقائق والوقائع التي تدعم وجهة نظره ،وكذلك الحقائق التي تتضارب مع منطلقاته وتصوراته ،فالنتيجة يجب ان تكون منطقية ومنسجمة مع الواقع ولاتناقضه.
٢. استخدام الطريقة الصحيحة والهادفة : ويقصد بذلك ان الباحث عندما يقوم بدراسة مشكلة او موضوع معين ويبحث عن حلها يجب ان يستعمل طريقة علمية صحيحة وهادفة للتوصل للنتائج المطلوبة.
٣. الاعتماد على القواعد العلمية : يتعين على الباحث الالتزام بتبني الاسلوب العلمي في البحث عن طريق احترام جميع القواعد العلمية المطلوبة لدراسة كل موضوع ،حيث ان تجاهل او اغفال اي عنصر من عناصر البحث العلمي يقود الى نتائج خاطئة او مخالفة للواقع .
٤. الانفتاح الفكري: ويقصد بذلك، انه يتعين على الباحث الحرص على التمسك بالروح العلمية والتطلع دائما الى معرفة الحقيقة فقط، والابتعاد قدر الامكان عن التزمت والتشبث بالرؤية الاحادية المتعلقة بالنتائج التي توصل اليها من خلال دراسته للمشكلة ويجب ان يكون ذهن الباحث منفتحا على كل تغير في النتائج المحصول عليها والاعتراف بالحقيقة وان كانت لاتخلو من مرارة .

٥. الابتعاد عن اصدار الاحكام النهائية: لاشك ان من خصائص البحث العلمي التي ينبغي على الباحث التقيد بها، هي ضرورة التأني وعدم أصدار الاحكام النهائية ،اذ يجب ان تصدر الاحكام استنادا الى البراهين والحجج والحقائق.

١-٤ صفات البحث الجيد :

١. العنوان الواضح و الشامل للبحث : ان الاختيار المناسب لعنوان البحث او الرسالة امر ضروري للتعريف بالبحث منذ الوهلة الاولى لقرائته من قبل الاخرين.

٢. تخطيط حدود البحث : ضرورة صياغة موضوع البحث ضمن حدود موضوعية و زمنية و مكانية واضحة المعالم ، و تجنب التخبط و المتاهة في امور لا تخص موضوع البحث .

٣. تركيز البحث على مجال معين : اذ لا يجوز البحث في اكثر من مجال من مجالات الحياة في ان واحد .

٤. الإسناد : ضرورة اعتماد الباحث في كتابة بحثه على الدراسات السابقة و الاراء الأصلية المسندة، و ان يكون دقيقا في سرد النصوص و ارجاعها لكاتبها الاصلي ، و الاطلاع على الاراء و الافكار المختلفة المتوفرة في مجال البحث لا سيما المذكورة في المصادر الحديثة. (مع الانتباه للامانة العلمية بالاقتباس و نقلها)

٥. وضوح الاسلوب : يجب ان يكون البحث الجيد مكتوب بأسلوب واضح و مقرؤ و مشوق، مع مراعاة السلامة اللغوية ، و ان تكون المصطلحات المستخدمة موحدة في متن البحث.

٦. الترابط بين اجزاء البحث : ضرورة ترابط اقسام البحث و اجزاءه المختلفة و انسجامها، كما يجب ان يكون هناك ترابط تسلسل منطقي ، تاريخي و موضوعي، يربط الفصول فيما بينها، ويكون هناك ايضا ترابط و تسلسل في المعلومات ما بين الفصول.

٧. الاسهام و الاضافة الى المعرفة في مجال تخصص الباحث : الباحث الجيد هو الذي يبدأ من حيث انتهى الاخرون بغرض مواصلة المسيرة البحثية و اضافة معلومات جديدة في المجال نفسه.

٨. ان لا يكون البحث محكوما عليه مسبقا : بمعنى ان لا يُقدر الباحث النتيجة و أن لا يبحث في شيء معروف مسبقا.

١-٥ صفات الباحث الجيد :

١. الرغبة في العمل : و ذلك بان تتوفر لدى الباحث الرغبة الاكيدة في القيام بعمل البحث و المواصلة في الموضوع الذي اختاره ، اذ ان الرغبة تعطي الحافز على العمل و تعطيه الطاقة من اجل مواصلة العمل.
٢. الصبر: و يعني المتابعة و المثابرة و عدم التذمر و مواجهة الصعاب ، فالصبر يوصل الباحث الى اكبر قدر ممكن من المعلومات.
٣. التتبع و حب الاطلاع و بذل الجهد و القراءة بفهم و عمق : و هذا يعني ان يلم الباحث بكل ما كتب عن موضوعه من بحوث و دراسات و أراء ، مهما تكون من اراء بسيطة اومخالفة .
٤. الدقة في فهم النصوص و اراء الغير : و ذلك بنقل العبارات بدقة و فهم و عدم تسرع لان التسرع و عدم الدقة تؤدي الى نتائج غير صحيحة.
٥. الثقة بالنفس : و ذلك بعدم الاستهانة بالكفاءة الشخصية و المهارات الذاتيةو تنميتها بالعمل و التمرين ، على ان تبنى هذه الثقة على العمل و الحفظ و الذكاء و التذكر.
٦. العقلية العلمية المنطقية المنظمة التي تعنى بترتيب الافكار و تحليلها و تنظيمها ، و تدريب العقل على النقد و الشك العلمي وعدم الاستسلام للبديهيات و الافكار العامة.
٧. الامانة العلمية : في النقل وعرض الافكار و عزوها الى اصحابها و الاشارة الى المصادر والمراجع التي انتفعت بها اوتعلمت منها .

١-٦ طريقة البحث العلمي :

١. تحديد مشكلة البحث : لا شك ان الخطوة الاولى في البحث العلمي هي ان يحدد الباحث المشكلة قيد الدراسة.
٢. جمع الحقائق و الملاحظات ذات العلاقة بالمشكلة .
٣. دراسة البيانات خلال طرح الاحتمالات لشرح العلاقات بين الحقائق و تقديم الحلول لهذه المشكلات.
٤. التفكير في النتائج على وفق الفروض .
٥. البحث عن ادلة لاثبات صحة او عدم صحة الفروض.

١-٧ التفكير العلمي : هو منهج او طريقة منظمة يمكن استعمالها في حياتنا اليومية او في اعمالنا و دراستنا ، فهو تفكير غير متخصص بموضوع معين بل يمكن ان يوجه في معالجة الموضوعات و القضايا و الاحداث التي تواجهنا .

١-٧-١ خصائص التفكير العلمي :

١. التراكمية : ينطلق التفكير العلمي من الواقع ، فالمعرفة فالمعرفة بناء يسهم فيه كل الباحثين و العلماء ، و ينطلق الباحث مما توصل اليه الاخرون.
٢. التنظيم :ان وسيلة العلم هي اتباع منهج علمي فالعلم معرفة منهجية تبدأ بالملاحظة و وضع الفروض و اختبارها عن طريق التجريب ثم الوصول الى النتائج.
٣. البحث عن الاسباب : يهدف العلم الى فهم الظواهر التي يدرسها، و لا يتم هذا الفهم من خلال الوصول الى المعلومات و الحقائق فقط ، بل لا بد من تفسير هذه الظواهر وتحليلها عن طريق معرفة اسبابها وعوامل نشئتها و تطورها.

٤. الشمولية و اليقين : يتصف التفكير العلمي بالشمول و اليقين، فالباحث ينطلق من دراسة المشكلة المحددة او الموقف الفردي للوصول الى نتائج و تعميمات تشمل الظواهر المشتركة او المواقف المشتركةمع موضوع دراسته .

٥. الدقة و التجريد : و هو ما يميز التفكير العلمي عن انماط التفكير الاخرى ، و لكي ينجح الباحث العلمي في حل مشكلاته بدقة فأنه يستعمل اللغة الرياضية المحددة ، التي تقوم على اساس القياس المنظم الدقيق و التحدث بلغة الارقام و الرموز و العلاقات الرياضية المحددة . استعمال هذه اللغة يؤدي الى فهم دقيق للظواهر .

١-٧-٢ عوائق التفكير العلمي :

١. انتشار الفكر الاسطوري: حيث الأسطورة أسلوب ناجح في تفسير الحياة والعلم في فترة كان التفكير الإنساني محدودا ولكن سيادة الفكر الأسطوري لم يكن إلا على حساب التصدي للعلم والتفكير العلمي. ونلحظ في مجتمعاتنا العربية والنامية ان التفكير الاحيائي والخرافي مازال قوياً ويقف موقفا معادياً للعلم والتفكير العلمي.

٢. الالتزام بالافكار الذائعة : إن الخضوع للأفكار الشائعة والمنتشرة بسبب صحتها وينظر الناس إليها نظرة تقديس واحترام ولا يقبل فيها مناقشة وكلما واجهت الإنسان ظروف ومصاعب زاد تمسكه بها.

٣. انكار قدرة العقل : حيث ظهرت اتهامات تستند إلى أساس ضعف العقل وعجزه عن فهم الكثير من الظواهر المرتبطة بظواهر الكون هذه الاتهامات زالت بعد تطور العقل الإنساني عن طريق المعرفة العلمية ومتابعتها.

١-٨ أخلاقيات البحث العلمي

هي مجموعة من المبادئ العامة والمبادئ التوجيهية التي تشكل و توجه الطريقة التي تصمم بها اي بحوث تشمل (الكائنات الحية)أي الناس والحيوانات وطريقة إجرائها وإدارتها واستخدامها ونشرها. وفي هذه

المبادئ التوجيهية، يُستخدم مصطلح "البحث" على نطاق واسع: فهو يشمل التحقيقات التشخيصية والاستقصائية للقضايا الاجتماعية ذات الأهمية.

١-٨-١ المبادئ الأساسية لأخلاقيات البحث العلمي

١- ينبغي أن تُحقق اقصى قدر ممكن من الفائدة للافراد و المجتمع، و ان تقلل المخاطر و الضرر الى ادنى حد.

٢- يجب إحترام حقوق الأفراد والجماعات وكرامتهم.

٣- لا بد من أن تكون المشاركة طوعية، حيثما أمكن، وأن تكون على القدر الملائم من العلم والاطلاع.

٤- ينبغي إجراء البحوث بنزاهة وشفافية.

٥- يجب تحديد خطوط المسؤولية والمساءلة بوضوح.

٦- لا بد من الحفاظ على استقلالية البحوث - وحيث يستحيل تجنب تضارب المصالح، ينبغي إبراز هذا التضارب.

اسس البحث العلمي

الفصل الثاني

الطريقة الاحصائية في البحث العلمي

جمع و تصنيف البيانات

مدرسة المادة : اريج صلاح محمد

٢ـ١ علم الاحصاء :

هو الطريقة العلمية التي تختص بجمع البيانات و الحقائق عن ظاهرة او فرضية (ظواهر او فرضيات) معينة و تنظيم و تبويب هذه البيانات والحقائق بالشكل الذي يسهل عملية تحليلها و تفسيرها و من ثم استخلاص النتائج و اتخاذ القرار على ضوء ذلك. وبشكل عام يميل معظم الاحصائيون للنظر الى هذا العلم على انه جمع لفرعين رئيسيين هما :

١. الاحصاء الوصفي Descriptive Statistic : يتضمن هذا الفرع الطرق و الاساليب المستخدمة في جمع البيانات و المعلومات عن ظاهرة معينه او مجموعة ظواهر و كيفية تنظيم و تصنيف و تبويب هذه البيانات مع امكانية عرضها في جداول و رسومات بيانية و حساب بعض المؤشرات الاحصائية منها.
٢. الاحصاء الاستدلالي Inferential Statistic : و هو الشطر الاخر من علم الاحصاء الذي يتم عادة بموضوعي التخمين و اختبار الفرضيات .

٢ـ٢ الطريقة الاحصائية في البحث العلمي :

ان استخدام الاسلوب الاحصائي في البحث العلمي يعني توفر البيانات و المعلومات عن ظاهرة من الظواهر المطلوب دراستها في ذلك البحث ، و هذا يعني ان امكانية تطبيق الطريقة الاحصائية مرهون بامكانية التعبير عن هذه الظاهرة او تلك تعبيرا كميا.

٢ـ٣ طبيعة البيانات الاحصائية :

عند جمع بيانات حول ظاهرة ما فأننا نرمز للظاهرة بالرمز (X) و كل مفردة او مشاهدة منها نرمز لها بالرمز (x_i) ، فمثلا عند دراسة درجات الطلبة في احدى الجامعات فأننا نرمز لصفة الدرجة بالرمز (X) و درجة اي طالب بالرمز (x_i) و تسمى المشاهدة او المفردة (observation) و ان قيمة (x_i) قد تختلف من طالب الى اخر و لهذا نقول بأن (X) هو متغير (Variable) ، و تقسم المتغيرات الى قسمين :

أ. متغيرات وصفية (نوعية) Qualitative Variable: و هي المتغيرات او الصفات التي لا يمكن قياسها بوسائل قياس مألوفة كالعد او القياس ، انما تكون صفات لذلك المتغير مثل لون العين كمتغير يمكن ان يكون (اسود ، ازرق ، عسلي) ، الجنس (ذكر ، انثى) ، الحالة الاجتماعية (اعزب ، متزوج ، مطلق ، ارمل) و غيرها .

ب. متغيرات كمية Quantitative Variable : و هي المتغيرات التي يمكن قياسها مباشرة بارقام عددية او وسائل قياس مألوفة مثل عدد الطلبة في مرحلة ما، عدد أشجار البرتقال في بستان ، طول الشخص بالسنتمتر ، و هذا القسم من المتغيرات يكون على نوعين :

١. متغيرات كمية متصلة Continuous Variable : و هي تلك المتغيرات التي تأخذ فيها المفردة او المشاهدة اي قيمة رقمية في مدى معين (فترة) مثل درجات الحرارة خلتال اليوم ، اطوال الطلبة في قسم ما .

٢. متغيرات كمية متقطعة Discrete Variable : و هي تلك المتغيرات التي تاخذ فيها المفردة او المشاهدة اعدادا صحيحة فقط ، مثل عدد الطلاب في قسم ما ، عدد افراد الاسرة.

٢-٤ اسلوب تصميم البحوث :

يجب مراعاة ما يلي عند تصميم البحث :

أ. تحديد الغرض من البحث : يجب ان يكون الهدف من البحث محدد بشكل واضح و دقيق بحيث يمكن التعرف على اوجه الاستفادة من نتائجه .

ب. تحديد امكانية التنفيذ الفعلي للبحث : من الضروري جدا تحديد المتطلبات التي تستلزمها عملية تنفيذ البحث و بشكل واضح و دقيق كالموارد المالية المطلوبة عند التنفيذ و الامكانيات البشرية المتاحة ، كذلك التأكد من مدى توفر البيانات و المعلومات الدقيقة عن مشكلة البحث.

ت. تحديد اطار البحث : ان احد الامور المهمة قبل البدء بتنفيذ البحث هو تحديد نوع و طبيعة مجال البحث .

ث. تحديد اسلوب جمع البيانات و المعلومات .

٢-٥ اساليب جمع البيانات :

ان اي بحث علمي يستند في تحليله الى الطريقة الاحصائية يحتاج الى بيانات و معلومات حول موضوع البحث قيد الدراسة . و يمكن للباحث الحصول على هذه البيانات و المعلومات من احد المصدرين الاتيين :

١. المصادر التاريخية Historical Sources :

و هي البيانلت و المعلومات المحفوظة و المجتمعة لدى اجهزة مؤسسات و دوائر الدولة المختلفة نتيجة لاستقصائات او مسوحات قامت بها هذه الجهات او هيئات معينة لاغراض خاصة بها او تجمعت لديها بحكم وظائفها الادارية و الفنية ، مثال على ذلك البيانات المتجمعة عن تعدادات السكان في العراق ، احصاءات الانتاج الزراعي و الصناعي ، احصائات التجارة الداخلية و الخارجية ، احصاءات الطلبة و الخريجين من الجامعات العراقية و غيرها.

٢. مصادر الميدان Field Sources :

و تمثل بيانات و معلومات يمكن الحصول عليها من مصادرها الاصلية بطريقة المراسلة او المواجهة او اي طريقة اتصال اخرى . مثال على ذلك نتائج التعداد العام لسكان العراق لعام ١٩٨٧ ، تسجيل حوادث الطرق خلال شهر معين ، تسجيل وقوعات الزواج و الطلاق خلال عام ٢٠٠٩ ، و غيرها. ان اختيار هذا المصدر دون ذاك في جمع البيانات و المعلومات يعتمد بالاساس على طبيعة البحث و النتائج المتوخاة منه ، و هنالك اسلوبان يمكن من خلالهما جمع البيانات و المعلومات ايا كان مصدرها ، هذان الاسلوبان هما اسلوب التسجيل الشامل و اسلوب العينات علما ان اختيار اي من هذين الاسلوبين يعتمد و بشكل اساس على طبيعة البحث او الدراسة و الدقة المطلوبة لنتائجها .

اولا. اسلوب التسجيل الشامل Census :

يقصد باسلوب التسجيل الشامل، جمع البيانات و المعلومات عن كافة المفردات التي تؤلف المجتمع الاحصائي للظاهرة او الظواهر قيد البحث، و في هذه الحالة يجب ان يكون هذا المجتمع محددا، اي يمكن ملاحظة و مواجهة كل مفردة من مفرداته. مثال على ذلك عملية التعداد العام للسكان في العراق لعام ١٩٨٧ حيث تم تسجيل البيانات و المعلومات عن كل فرد (مفردة احصائية) دون استثناء و يعتبر اسلوب التسجيل الشامل افضل اسلوب في جمع البيانات كونه يجهز الباحث ببيانات كاملة عن كافة مفردات مجتمع الدراسة الا انه يحتاج الى وقت و جهد و موارد مادية وبشرية كبيرة في انجاز مهمة جمع البيانات بالاضافة الى احتمال الوقوع في اخطاء نتيجة التعامل مع مفردات كثيرة.

ثانيا. اسلوب العينات Samples :

يقصد باسلوب العينات ، عملية جمع البيانات و المعلومات عن مجموعة معينة من مفردات مجتمع الدراسة، هذه المجموعة من المفردات تسمى (عينة sample) ، مثال على ذلك دراسة فاعلية دواء معين على بعض الاشخاص المصابين بمرض السرطان ، و غيرها من الامثلة. و يمتاز اسلوب العينات بأنه يحتاج وقت و جهد و موارد بشرية اقل مما يحتاجه اسلوب التسجيل الشامل.

٢-٦ المجتمع و العينة :

٢-٦-١ المجتمع Population:

هو عبارة عن جميع القيم او المفردات التي يمكن ان يأخذها المتغير، او هو مجموعة من الافراد او المشاهدات التي تشترك في صفة متغيرة واحدة او اكثر تميزه تماما عن بقية المجتمعات ، مثل قسم الرياضيات في كلية التربية للعلوم الصرفة ابن الهيثم، و يقسم المجتمع الى قسمين

١. مجتمع محدد Finite population :
   اي من الممكن جمع عدد مفرداته او عدد مشاهداته مثل عدد التدريسيين في كلية ما .
٢. مجتمع غير محدد Infinite population :
   اي من غير الممكن جمع عدد مفرداته او مشاهداته مثل عدد الاسماك في نهر دجلة.

يسعى الباحث الى اشتراك جميع افراد المجتمع، لكن الصعوبة تكمن في ان عدد افراد المجتمع قد يكون كبيرا، بحيث لا يستطيع الباحث اشراكهم جميع فيلجا الباحث في تلك الحالة الى اختيار مجموعة جزئية من مجتمع البحث.

٢-٦-٢ العينة Sample :

هي جزء من المجتمع يجري اختيارها وفق قواعد خاصة لكي تمثل المجتمع تمثيلا صحيحا، بحيث يمكن تعميم نتائج تلك العينة على المجتمع بأكمله وعمل استدلالات حول معالم المجتمع.

٢-٦-٣ اسباب استخدام العينات:

١- ان من فوائد استخدام العينات اختصار للوقت والجهد والتكاليف.
٢- الحاجة إلى الوصول إلى نتائج سريعة لاتخاذ قرارات ضرورية.
٣- الرغبة في الوصول على نتائج دقيقة وذات ثقة عالية بالاستدلال الإحصائي.

٤- نستخدم العينات إذا كانت الوحدات المدروسة ذات تشتت عالي بالنسبة للمتغيرات التي نرغب بها .

٥- إذا كان من الصعوبة أو الاستحالة بمكان الدخول إلى كامل المجتمع مثل: إذا كان المجتمع غير قابل للعد مثال على ذلك (مخزون بترول –مخزون ماء) .

٢-٦-٤ شروط اختيار عينة البحث

من أهم الشروط الواجب مراعاتها عند اختيار عينة للبحث:

١- ان تكون بعيدة عن الانحياز و المحاباة .

٢- ان تمثل العينة مجتمع البحث الاصلي بشكل صحيح ، و ان لا تمثل مجتمعا اخر.

٣- تحقيق التجانس بين مختلف مكونات مجتمع البحث الاصلي ، و في حال عدم القدرة على تحقيق ذلك و خصوصا في المجتمع غير المتجانس فيتوجب على الباحث تجزئته الى مجتمعات اصغر متجانسة.

٤- حصر مسبق لكافة مكوناتمجتمع البحث مع تجزئته الى وحدات معاينة ، و حصر كل وحدة منها داخل اطار احصائي خاص، ومثال على ذلك، فإنه عند دراسة سكان أحد المجتمعات، فإن وحدة المعاينة ستتمثل في الأسر أو الأفراد أو الجماعات، كما قد يتمثل في المجتمع الصغير عند دراسة المجتمعات الكبيرة.

٥- مناسبة حجم العينة ونوعها مع الأهداف الأساسية للبحث، ومع طبيعة مجتمع البحث، وطبيعة المشكلة المراد دراستها من خلال هذا البحث.

اسس البحث العلمي

الفصل الثالث

مخطط البحث

اعداد مسودة البحث

مدرسة المادة : اريج صلاح محمد

المُسَوَدَةُ لغة :

الصحيفة أو الصحائف تكتب أول كتابة ثم تُنقح وتُحرَر وتُبيض.

اما في اصطلاح كتاب البحوث والرسائل فهي من الاوليات التي لا يستغني عنها الباحث في ترتيب افكار واراء نسخت في ذهنه .

وتعد المسودة التجربة الاولى لكتابة البحث وغالبا ما يعتريها بعض الضعف ونقص في المعلومات او عدم الدقة في كتابة العرض للموضوع وهي خطوة مهمة لاخراج البحث من حيز التفكير الى حيز الوجود .

يحتاج الباحث الى تدوين المعلومات التي حصل عليها سواء اكانت معلومات جمعها من المصادر او الوثائق، كما ويحدد الباحث في مسودات البحث اماكن الهوامش والمصادر وبعض المعلومات مثل المؤلف، العنوان ، النشر فضلا عن فائدة الاضافة والتعقيب ، وتترك فراغات مناسبة بين السطور والهوامش في مسودة البحث .

اهمية كتابة المسودة بالنسبة للباحث :

١- اعطاء صورة تقريبية للبحث بشكله النهائي.

٢- ان يدرك الباحث من خلال المسودة ما هو ناقص في بحثه او زائد وان يعمل على الموازنة في ذلك.

٣- الاستعانة والاقتباس من النصوص او مواد مأخوذة من مصادر اخرى ، وما يحتاج اليه الباحث ، وترتيب اسلوبه بما يلزم عمله .

٤- محاولة ترتيب مواقع المقاطع او الفصول المختلفة.

مخطط البحث :

هو مشروع عمل او خطة منظمة تجمع عناصر التفكير المسبق اللازمة لتحقيق الغرض من الدراسة ، و يهدف المخطط تحقيق ثلاثة اغراض اساسية هي :

- يصف اجراءات القيام بالدراسة و متطلباتها.
- يوجه خطوات الدراسة و مراحل تنفيذها.
- يشكل اطارا لتقويم الدراسة بعد انتهائها.

عناصر البحث :

تختلف عناصر البحث باختلاف المؤسسة التي تشرف على البحث ، لكن هناك عناصر اساسية يجب توافرها و هي :

اولا. العنوان : يجب ان يكون محدد بدلالة البحث و ان يشير الى موضوع الدراسة ، و ان تكون الغة المستعملة في العنوان لغة مهنية عادية و ليست لغة صحفية استعراضية ، يفضل ان لا يزيد عدد كلمات النوان عن خمس عشرة كلمة.

ثانيا. مشكلة البحث : يجب ان تكون المشكلة التي جرى اختيارها قابلة للبحث. تتنوع المصادر التي ياخذ منها الباحث مشكلة بحثه ، فقد يتطوع الباحث للبحث في مشكلة جرى تحدديها من قبل اخرين ، و قد يقوم اكثر من باحث بدراسة مشكلة معينة يتناول كل منهم جانب محدد من جوانبها.

ثالثا. فرضيات البحث : عندما تتضح مشكلة البحث ، يحتاج الباحث الى تحديد المعيار الذي سوف يجمع على ضوئه المعلومات ، و هذا المعيار يكون اما اسئلة البحث او فرضياته .

فرضيات البحث : هي الحلول الممنكنة التي يفرضها الباحث للمشكلة .

معايير الفروض الجيدة :

أ. ان تصور ما يتوقع الباحث فعلا عند حل المشكلة.

ب. ان تستخدم اسس نظرية و براهين علمية تؤكد جدوى اختبارها .

ت. ان تكون قابلة للاختبار ليست فيها عمومية.

ث. ان تكون مختصرة و واضحة.

الفروض الاحصائية : هي تعبير عن واحد او اكثر من مقاييس المجتمع التي سحبت منها العينة و فرض العدم و الفرض البديل هما شكلان من الفروض الاحصائية .

١. فرض العدم (الصفري) Null Hypothesis: يفترض الباحث بان العلاقة بين المتغيرات المدروسة صفر (اي فرق او مقدار العلاقة ما هو الا صدفة) ، وتطبق المعالجات الاحصائية التي تساعد الباحث على قبول الفرض او عدمه.

٢. الفرض البديل Alternative Hypotheses: يفترض الباحث ان العلاقة بين المتغيرات المدروسة ليست صفرا اي انه يميل الى جهة ما .

رابعا. الاطار النظري و الدراسة النظرية و الدراسات السابقة: الاطار النظري هو الخلفية العلمية النظرية التي يحتاجها الباحث ليستطيع اعداد بحث علمي له اهداف . اما الدراسة النظرية يقصد بها الوثائق المنشورة كتبا كانت او غيرها ليستنتج منها الادلة و البراهين. مصطلح الدراسات السابقة يراد به مراجعة الدراسات السابقة التي تناولت الموضوع او بعض جوانبه .

خامسا. خلفية الدراسة و اهميتها : توثيق مواقف الباحثين الاخرين لنفس المشكلة في البحوث المنشورة ، و الاشارة في بعض الاحيان الى طول الفترة الزمنية التي انقضت بين الدراسات السابقة و بين هذه الدراسة و مع تطور الظروف و التقنيات الامر الذي يقتضي تحديث الدراسات السابقة.

سادسا تعريف المصطلحات: من المهم توضيح المصطلحات المستعملة بالبحث حتى لا يساء فهمها او تفهم بدلالة غير الدلالة الواردة في الدراسة .

سابعا محددات الدراسة: هي الصعوبات و العوامل التي تواجه الباحث و تعيق امكانية تعميم نتائج البحث .

ثامنا طريقة الدراسة و اجرائاتها: وهي الطريقة التي يجيب فيها الباحث عن اسئلة الدراسة او يختبر فرضياتها بشكل تفصيلي.

تاسعا المراجع: يجب ان تتضمن قائمة المراجع على الاقل المراجع التي قادت الباحث الى اختيار مشكلته ، و المراجع الضرورية لفهم المشكلة .

عاشرا ملاحق البحث: يمكن ان يتضمن البحث بعض المواد التي لا يكون مناسبا عرضها في صلب البحث لكن وجودها يعرض مزيد من التوضيح و التفصيل .

اعداد مسودة البحث :

تدوين البحث:

ان اسس كتابة البحث العلمي المتفق عليها تقضي بان يدون البحث مقسما الى الاقسام الاتية حسب تسلسلها

عنوان البحث – الموجز (الملخص) – المقدمة – المواد و الطرق – النتائج – المناقشة – الخاتمة (المختصر)- الشكر – قائمة المصادر – الملاحق (الجدول ، الخطوط البيانية ، الصور التوضيحية ، ....)

و قبل ان يبدء الباحث في تدوين مسودة بحثه عليه ان يراعي النقاط التالية قدر المستطاع

١. تحديد المصادر: على الباحث تحديد مصادره و التي تكون غالبا بحوث منشورة تتناول نفس موضوع البحث او قريبة منه.

2. لغة البحث : عادة ما تنشر البحوث باحدى اللغات الحية كاللغة الانجليزية مثلا.
3. عنوان البحث: يتضمن البحث عدة عناوين عنوان البحث و عناوين فرعية تمثل اقسام البحث .
4. الاقتباس: يضطر الباحث احيانا لاقتباس سطور او فقرات منشورة لباحث اخر .
5. الهوامش: لا تستخدم الهوامش في البحوث العلمية الا للضرورة ، و توضع الارقام (١،٢،....) فوق اواخر الكلمات او الجمل التي يريد الباحث كتابة هوامش لها .
6. الاسماء العلمية : يجب كتابة الاسماء العلمية كما متفق عليها علميا .
7. الرموزو المصطلحات: بعض الرموزو المصطلحات المستخدمة في الكتابة العلمية باللغة الاجنبية هي في حقيقتها مختصرات لكلمات لاتينية او يونانية اوانجليزية ، تدل على معاني محددة ينبغي على الباحث ان يكون ملما بها.
8. اختيار المجلة الناشرة: قبل ان يدون الباحث مسودة بحثه عليه ان يحدد المجلة التي سيختارها لنشر بحثه ليستطيع تدوينه وفقا لتعليمات المجلة ، توفيرا للوقت و الجهد.

هيكلية البحث العلمي او التقرير :

١. العنوان Title:

يجب ان يدل العنوان دلالة واضحة على المشكلة المدروسة ، يتبعه اسم الباحث (او اسماء الباحثين) بدون القاب علمية او اجتماعية مثل (الدكتور ، الاستاذ، السيد،...) ، بعده يذكر اسم المؤسسة .

٢. الملخص (الموجز) Abstract:

يلخص فيه الباحث بصورة موجزة مشكلة البحث والطرق التي استخدمها في معالجتها، على ان لا يتجاوز نصف صفحة.

٣. المقدمة Introduction :

تعتبر المقدمة هي المدخل للبحث ، و فيها يعرض الباحث طبيعة المشكلة و يلخص الدراسات المرتبطة بها ، و يقدم فروض البحث و الافتراضات التي تستند اليها .

٤. المواد و الطرق Materials and Methods :

في هذا القسم يذكر الباحث بايجاز شديد المواد التي استخدمها في اجراء بحثه ، و طرق استخدامها و المكان و الزمان اللذين استخدم فيها تلك المواد.

٥. النتائج و المناقشة Results and discussion :

تتضمن النظر في النتائج الحاصلة و مناقشتها و ينبغي مقارنة النتائج الحاصلة مع نتائج البحوث السابقة .

٦. الخلاصة Conclusion :

و فيها توفير للوقت من اجل القارئ الذي يريد الاطلاع على البحث اطلاعا سريعا ، حيث يدون فيها الباحث النتائج النهائية و المعلومات الجديدة التي استخلصها من بحثه ، و يمكن ان يدون توصيات مكتوبة بشكل جمل قصيرة .

٧. الشكر Acknowledgement :

في هذا القسم يشكر الباحث الاشخاص و المؤسسات الذين ساعدوه بصورة او باخرى في البحث.

٨. قائمة المصادر References :

يجب ذكر قائمة كاملة للمصادر (المراجع) التي استعان بها الباحث ، و يكون ترتيبها حسب الهيكلية التي تطلبها المجلة ، في الاغلب تترتب حسب ضهورها في البحث و احيانا حسب الاحرف الابجدية لاسماء المؤلفين ، و لا يجوز ذكر مصادر لم يتم الاستعانة بها حتى و ان كانت تتعلق بموضوع البحث.

الفصل الرابع // التوزيعات التكرارية واساليب عرض البيانات

Frequency distribution

المقدمة :

عند جمع البيانات الاولية الخاصة بدراسة ظاهرة معينة فانه عادة لا يمكن الاستفادة منها وهي بهذه الصورة ، لذلك فهي غالباً ما توضع في جداول مبسطة او يعبر عنها في صورة اشكال او رسوم بيانية لكي يسهل من عملية دراستها وتحليلها وتقيمها ، وسنركز الاهتمام في هذا الفصل على اساليب تبويب البيانات في جداول خاصة تدعى بجدول التوزيعات التكرارية وكذلك استعراض لأساليب عرض البيانات هندسياً :

(٤-١) جدول التوزيع التكراري Frequency Distribution or Frequency Table

بعد اتمام عملية جمع البيانات وتدقيقها وتصنيفها نحصل على مجموعة من الاعداد التي تمثل احد الظواهر الاحصائية وهذه البيانات تكون غير مرتبة وغير مفهومة وحتى يصبح لهذه البيانات معنى ويمكن الاستفادة منها نقوم بتوزيع هذه البيانات او المفردات الاحصائية الى مجموعات متجانسة بحسب كل صفة من الصفات المميزة وترتب هذه البيانات تصاعدياً او تنازلياً وتسمى كل مجموعة من هذه المجموعات بالفئة (Class) ويسمى هذا التوزيع بحسب الفئات بالتوزيع التكراري ، بعدها يتم ايجاد عدد الحالات او المفردات التي تقع داخل حدود كل فئة لإيجاد ما يسمى بالتكرار ، وغالبا توضع البيانات او التوزيعات في جدول يسمى جدول التوزيع التكراري وقد تكون الفئات (فئات التوزيع) متساوية بالطول ام غير متساوية وذلك يعتمد على طبيعة الدراسة ومتطلباتها ولكن الشائع في الاستعمال هو الفئات المتساوية لكونها اسهل في العمليات الحسابية ، ويمكن تعريف التوزيع التكراري على انه عبارة عن تلخيص وترتيب وفرز وتفريغ لبيانات المتغيرات التي سبق وان جمعت وصنفت .

(٤-٢) الخطوات العامة في انشاء التوزيعات التكرارية General Rule for Construction Frequency Distribution

(٤-٢-١) المدى الكلي للتوزيع Total Range :

يعرف المدى بانه الفرق بين اكبر قيمة واصغر قيمة في المجموعة مضافا له واحد وكما يأتي :

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

حيث ان :

$T.R.$: يمثل المدى الكلي للتوزيع

X_L : تمثل اكبر قيمة في المجموعة

X_S : تمثل اصغر قيمة في المجموعة

(٤-٢-٢) عدد فئات التوزيع Number of Classes :

تمثل عدد المجاميع التي يتألف منها التوزيع التكراري وهناك صيغ تقريبية يمكن من خلالها تحديد عدد فئات التوزيع اهمها :

أ- صيغة يول (Yule) وهي :

$$m = 2.5\sqrt[4]{n}$$

ب- صيغة سترجس (Sturges) وهي :

$$m = 1 + 3.322\ Log_{10}(n)$$

حيث ان :

m : تمثل عدد الفئات

n : تمثل عدد المفردات

ملاحظة : عند التطبيق يتم تقريب الناتج لأقرب عدد صحيح .

(٤-٢-٣) طول الفئة Length of a Class :

ويسمى احيانا المدى الفئوي ويمثل سعة الفئة ، اي مقدار المسافة ما بين الحد الادنى للفئة والحد الاعلى للفئة . ان طول الفئة يتناسب عكسيا مع عدد الفئات في التوزيع فكلما كبر طول الفئة قل عدد الفئات والعكس صحيح ، ويمكن تحديد طول الفئة في حالة التوزيعات ذات الفئات المتساوية من خلال الصيغة الاتية :

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

حيث ان L يمثل طول الفئة ، اما في حالة كون الفئات غير متساوية فيمكن تحديد طول الفئة وكما يأتي :

طول الفئة = الفرق بين ادنى حدين او (اعلى حدين) لفئتين متتاليتين

او

طول الفئة = الحد الاعلى للفئة – الحد الادنى للفئة

(٤-٢-٤) الحد الادنى للفئة والحد الاعلى للفئة Lower&Upper Bound of a Class

لكل فئة من فئات التوزيع التكراري نهاية وبداية ، فالبداية تعني الحد الادنى للفئة والنهاية تعني الحد الاعلى لها . وسوف نرمز للحد الادنى للفئة بالرمز $(L.C)$ ، والحد الاعلى للفئة بالرمز $(U.C)$ ، اما مركز الفئة فسنرمز له بالرمز (X) وهو عبارة عن قيم المتغير العشوائي التي تتوسط المسافة بين الحد الادنى والحد الاعلى كما يأتي :

$$X = \frac{L.C. + U.C.}{2}$$

(٤-٢-٥) تكرار الفئة Class Frequency :

يمثل تكرار الفئة جزءاً من مفردات العينة التي تتصف بكونها تقع من حيث القيمة العددية ما بين حدي الفئة بحيث ان مجموع هذه الاجزاء يشكل عدد مفردات العينة n ، فاذا رمزنا لتكرارات الفئات بالرمز $f_1, f_2, \dots, f_m$ وأن

$\sum_{i=1}^{m} f_i = f_1 + f_2 + \dots + f_m = n$ ، عندئذ فان الشكل العام لجدول التوزيع التكراري ذات فئات متساوية الاطوال هو كما يأتي :

تسلسل الفئة	الحد الادنى للفئة	الحد الاعلى للفئة	طول الفئة (L)	مركز الفئة (X)	تكرار الفئة (f)
1	X_S	$X_S + L$	L	$X_S + (1/2)L$	f_1
2	$X_S + L$	$X_S + 2L$	L	$X_S + (3/2)L$	f_2
3	$X_S + 2L$	$X_S + 3L$	L	$X_S + (5/2)L$	f_3
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
m	$X_S + (m-1)L$	$X_S + mL$	L	$X_S + (m-(1/2))L$	f_m
المجموع					n

مع ملاحظة انه ليس من الضروري عند التطبيق ان يكون الحد الادنى للفئة الاولى مساوياً تماماً لأصغر قيمة في المجموعة بل قد يكون اقل منها لاعتبارات تتعلق بتسهيل العمليات الحسابية اللاحقة . كذلك فأن التوزيع التكراري قد يكون توزيع مغلق او توزيع مفتوح وذلك يعتمد على طبيعة الدراسة ، ويقصد بالتوزيع المغلق بانه ذلك التوزيع الذي يمتلك حداً أدنى للفئة الاولى وحداً أعلى للفئة الاخيرة ، في حين يقصد بالتوزيع المفتوح بأنه ذلك التوزيع الذي لا يمتلك حداً للفئة الاولى أو حداً أعلى للفئة الأخيرة أو كليهما معاً .

وبشكل عام يفضل ان يمتاز التوزيع التكراري بما يأتي :

أ- ان تكون فئات التوزيع متساوية الطول .

ب- ان يكون التوزيع التكراري توزيع مغلق ، وفي ذلك اهميه كبيرة في تسهيل عملية حساب بعض المؤشرات الاحصائية (كالوسط الحسابي ، الانحراف المعياري ، وغيرها) التي يعتمد حسابها على تحديد مراكز الفئات التي صعب تحديدها اذا كان التوزيع مفتوحاً .

ت- ان تبدأ فئات التوزيع وتنتهي (حدود الفئات) بأعداد صحيحة لما لذلك من أثر في تسهيل العمليات الحسابية .

ث- ان لا يقل عدد فئات التوزيع عن خمسة ولا يزيد عن خمسة عشر فئة ، فاذا قل عدد فئات التوزيع عن خمس فئات فأن عملية التبويب قد تؤدي الى عدم كشف الصفات الاساسية للمجتمع الاحصائي ، اما اذا زاد عددها عن خمسة عشر فئة فان في ذلك بعض الصعوبة في اجراء العمليات الحسابية لبعض المؤشرات الاحصائية .

ج- ان تكون حدود الفئات محددة بشكل واضح بحيث ان كل قيمة من قيم المفردات يجب ان تبوب (تضمن) في فئة واحدة فقط من فئات التوزيع .

(٤-٣) كتابة حدود الفئات للتوزيع التكراري استناداً الى نوع المتغير العشوائي :

(٤-٣-١) في حالة كون المتغيرات متقطعة :

في هذه الحالة نكتب حدود الفئات حسب التوزيع التكراري الاتي بحيث نضمن ان كل قيمة من قيم البيانات ضمن فئة واحدة من فئات التوزيع دون اي تكرار قد يحصل في هذه الفئة او تلك وبحيث ان طول الفئة (L) مساوياً للفرق ما بين الحد الاعلى للفئة وحدها الادنى مضافاً للفرق العدد واحد كما موضح ادناه :

تكرار الفئة (f)	حدود الفئات	الحد الاعلى للفئة (الى)	الحد الادنى للفئة (من)	تسلسل الفئة
f_1	$(X_S) - (X_S + L - 1)$	$X_S + L - 1$	X_S	1
f_2	$(X_S + L) - (X_S + 2L - 1)$	$X_S + 2L - 1$	$X_S + L$	2
f_3	$(X_S + 2L) - (X_S + 3L - 1)$	$X_S + 3L - 1$	$X_S + 2L$	3
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
f_m	$(X_S + (m-1)L) - (X_S + mL - 1)$	$X_S + mL - 1$	$X_S + (m-1)L$	m
n				المجموع

مثال (١) : البيانات الاتية تمثل الدرجات التي حصل عليها (30) طالب في مادة الاحصاء ،يطلب تفريغ وتبويب البيانات (درجات الطلاب) في جدول توزيع تكراري .

61	42	73	79	55	63	52	68	65	56
44	73	43	68	82	76	53	46	70	72
47	82	87	70	63	48	45	53	75	60

الحل : من الواضح ان المتغير العشوائي هو متغير متقطع (درجات الطالب) ولأجل تكوين جدول توزيع تكراري يتم حساب ما يلي :

1) $\boldsymbol{T.R. = X_L - X_S + 1}$

$\boldsymbol{T.R. = 87 - 42 + 1 = 46}$

2) $\boldsymbol{m = 2.5\sqrt[4]{n} = 2.5\sqrt[4]{30} = 5.8508 \approx 6}$

3) $\boldsymbol{L = \frac{T.R.}{m} = \frac{46}{6} = 7.67 \approx 8}$

وعليه فان جدول التوزيع التكراري سيكون كالاتي :

تسلسل الفئة	الحد الادنى للفئة (من)	الحد الاعلى للفئة (الى)	حدود الفئات	تكرار الفئة (f)
1	42	42+08 - 1	42 – 49	7
2	42+08	42+16 - 1	50 – 57	5
3	42+16	42+24 - 1	58 – 65	5
4	42+24	42+32 - 1	66 – 73	7
5	42+32	42+40 - 1	74 – 81	3
6	42+40	42+48 - 1	82 – 89	3
المجموع				30

ومن الممكن اختصار الجدول اعلاه بجدول التوزيع التكراري بوجود مراكز الفئات وكالآتي :

تسلسل الفئة	حدود الفئات	تكرار الفئة (f)	مركز الفئات (X)
1	*42 – 49*	7	45.5
2	50 – 57	5	*53.5*
3	58 – 65	5	*61.5*
4	66 – 73	7	*69.5*
5	74 – 81	3	*77.5*
6	82 – 89	3	*85.5*
المجموع		30	

حيث ان : مركز الفئة $(X) = \left(\text{الحد الادنى للفئة} + \text{الحد الاعلى للفئة}\right)/2$

(٤-٣-٢) في حالة كون المتغيرات مستمرة (متصلة) :

في هذه الحالة تكتب حدود الفئات حسب التوزيع التكراري الاتي بحيث تضمن ان كل قيمة من قيم البيانات ضمن فئة واحدة من فئات التوزيع دون اي تكرار قد يحصل في هذه الفئة او تلك وبحيث ان طول الفئة L مساوياً للفرق ما بين الحد الاعلى للفئة وحدها الادنى كما موضح في ادناه :

تسلسل الفئة	الحد الادنى للفئة (من)	الحد الاعلى للفئة (الى)	حدود الفئات	تكرار الفئة (f)
1	X_S	$X_S + L$	$(X_S) - < (X_S + L)$	f_1
2	$X_S + L$	$X_S + 2L$	$(X_S + L) - < (X_S + 2L)$	f_2
3	$X_S + 2L$	$X_S + 3L$	$(X_S + 2L) - < (X_S + 3L)$	f_3
.	.	.	.	.
.	.	.	.	.
m	$X_S + (m-1)L$	$X_S + mL$	$(X_S + (m-1)L) - < (X_S + mL)$	f_m
المجموع				n

مثال (٢) : البيانات ادناه تمثل اوزان (كغم) لعينة من طلبة احدى الكليات قوامها (16) طالب ، يطلب تفريغ هذه البيانات في جدول توزيع تكراري ومن ثم حساب مراكز الفئات في حالة المتغير مستمر .

90	73.5	88.5	66.5	55.5	46.5	46	80.5
87.5	69.5	77.3	60.8	48.2	51.3	49.2	49.5

الحل : من الواضح ان المتغير العشوائي (اوزان الطلبة) هو متغير متصل ولأجل تكوين جدول توزيع تكراري يتم حساب ما يأتي :

1) $\boldsymbol{T.R. = X_L - X_S + 1 = 90 - 46 + 1 = 45}$

2) $\boldsymbol{m = 2.5\sqrt[4]{n} = 2.5\sqrt[4]{16} = 5}$

3) $\boldsymbol{L = \frac{T.R.}{m} = \frac{45}{5} = 9}$

وعليه فأن جدول التوزيع التكراري سيكون كالآتي :

مركز الفئات (X)	تكرار الفئة (f)	حدود الفئات	تسلسل الفئة
50.5	6	46 – < 55	1
59.5	2	55 – < 64	2
68.5	2	64 – < 73	3
77.5	3	73 – < 82	4
86.5	3	82 – < 91	5
	16		المجموع

(٤-٤) التوزيع التكراري النسبي A Relative Frequency Distribution

ان التوزيع التكراري النسبي هو نفسه التوزيع التكراري الذي سبق شرحه الا ان التكرارات تكون على شكل نسب مئوية يمكن الحصول عليها من خلال قسمة تكرار كل فئة على مجموعة التكرارات الكلية . فعلى افتراض ان هناك توزيع تكراري عدد فئاته $\boldsymbol{m}$ وان التكرارات المقابلة لهذه الفئات حسب تسلسلها هي $f_1 , f_2 , f_3 , \dots , f_m$ عندئذ فأن التكرارات النسبية المقابلة لهذه الفئات هي $f_1^* , f_2^* , f_3^* , \dots , f_m^*$ حيث ان :

$$f_i^* = \frac{f_i}{n} * 100$$

حيث أن :

$n =$ تمثل مجموع التكرارات الكلية .

مثال (٣) : الجدول ادناه يوضح التوزيع التكراري لدرجات (*90*) طالب وطالبة في امتحان معين ، يطلب انشاء توزيع تكراري نسبي :

الفئات	*10-*	*20-*	*30-*	*40-*	*50-*	*60-*	*70-*	*80-90*
التكرار (عدد الطلبة)	*2*	*3*	*5*	*10*	*18*	*25*	*15*	*12*

الحل : بهدف الحصول على التوزيع التكراري النسبي نقسم تكرار الفئة على مجموع التكرارات الكلية والبالغة *90* وكما موضح في الجدول الاتي :

عدد الفئات	**عدد الطلبة f_i**	**التكرار النسبي f_i^***
10-	2	(2/90)*100=2.222%
20-	3	(3/90)*100=3.333%
30-	5	(5/90)*100=5.556%
40-	10	(10/90)*100=11.111%
50-	18	(18/90)*100=20%
60-	25	(25/90)*100=27.778%
70-	15	(15/90)*100=17.778%
80-90	12	(12/90)*100=14.444%

وهذا يعني ان نسبة *2.222%* من الطلبة تتراوح درجاتهم ما بين *10* الى اقل من *20* درجة . وان نسبة *27.778%* من الطلبة تتراوح درجاتهم ما بين *60* الى اقل من *70* درجة وهكذا .

(٤-٥) التوزيع التكراري المتجمع Cumulative Frequency Distribution

ان جدول التوزيع التكراري الذي سبق شرحه يبين توزيع قيم المتغير على الفئات المختلفة ، ولكن في بعض الاحيان قد يكون حاجة الى معرفة عدد القيم او المفردات التي تقل او تزيد عند قيمة معينة ، بمعنى اخر هو التوزيع الذي يبين كمية التكرار المتجمع عند قيمة معينة من قيم المتغير العشوائي . هذا التوزيع على نوعين :

(٤-٥-١) التوزيع التكراري المتجمع الصاعد Ascending Cumulative Frequency Distribution

وهو التوزيع الذي يبين تراكم التكرارات ابتداءً من الفئة الاولى في التوزيع وانتهاءً بالفئة الاخيرة منه ، ويتم حساب التكرارات المتجمعة على اساس الحدود العليا للفئات وسوف نرمز له بالرمز (**Fv**).

اولاً: في حالة كون المتغير العشوائي من النوع المتقطع

ان التوزيع التكراري المتجمع الصاعد في هذه الحالة سوف يأخذ الشكل الاتي :

الحدود العليا للفئات	عبارة التجميع	التكرار الاصلي	التكرار المتجمع الصاعد (Fv)
$X_S + L - 1$ $X_S + 2L - 1$ $X_S + 3L - 1$ $\vdots$ $X_S + mL - 1$	اقل من او يساوي قيمة الحد الاعلى للفئة	f_1 f_2 f_3 $\vdots$ f_m	$Fv_1 = f_1$ $Fv_2 = f_1 + f_2 = Fv_1 + f_2$ $Fv_3 = f_1 + f_2 + f_3 = Fv_2 + f_3$ $\vdots$ $Fv_m = f_1 + f_2 + \cdots + f_m = Fv_{m-1} + f_m = n$
المجموع		n	

ويمكن حساب التوزيع المتجمع الصاعد النسبي والذي سنرمز له بالرمز ***Fv** ، حيث ان :

$$\mathbf{Fv}^* = \frac{\mathbf{Fv}}{\boldsymbol{n}} \cdot \mathbf{100}$$

مثال (٥): الاتي جدول توزيع تكراري يمثل توزيع *60* عائلة فلاحية حسب ملكيتها من عدد اشجار البرتقال .

عدد الاشجار	*60-74*	*75-89*	*90-104*	*105-119*	*120-134*	*135-149*	*150-164*
عدد العوائل	*4*	*5*	*10*	*12*	*16*	*7*	*6*

المطلوب : تكوين توزيع تكراري متجمع صاعد ، وتوزيع تكراري متجمع صاعد نسبي .

الحل:

الفئات	الحدود العليا للفئات	التكرار f_i	ت.م.ص Fv	ت.م.ص.ن Fv^*
60-74	Less than or equal 74	4	4	(4/60)100=6.67%
75-89	Less than or equal 89	5	9	15%
90-104	Less than or equal 104	10	19	31.67%
105-119	Less than or equal 119	12	31	51.67%
120-134	Less than or equal 134	16	47	78.33%
135-149	Less than or equal 149	7	54	90%
150-164	Less than or equal 164	6	60	100%
المجموع		60		

ثانياً: في حالة كون المتغير العشوائي من النوع المستمر

ان التوزيع التكراري في هذه الحالة سيكون بالشكل الاتي :

الحدود العليا للفئات	عبارة التجميع	التكرار الاصلي	التكرار المتجمع الصاعد (**Fv**)
$X_S + L$ $X_S + 2L$ $X_S + 3L$ $\vdots$ $X_S + mL$	اقل من او يساوي قيمة الحد الاعلى للفئة	f_1 f_2 f_3 $\vdots$ f_m	$Fv_1 = f_1$ $Fv_2 = f_1 + f_2 = Fv_1 + f_2$ $Fv_3 = f_1 + f_2 + f_3 = Fv_2 + f_3$ $\vdots$ $Fv_m = f_1 + f_2 + \cdots + f_m = Fv_{m-1} + f_m = n$
المجموع		n	

مثال (٦): الاتي جدول توزيع تكراري لأوزان عينه من طلبة احدى الكليات قوامها (100) طالب .

الفئات	46-	53-	60-	67-	74-	81-	88-	95- <102
التكرارات	7	15	27	21	14	8	5	3

المطلوب : تكوين توزيع تكراري متجمع صاعد ، وتوزيع تكراري متجمع صاعد نسبي .

الحل :

الفئات	الحدود العليا للفئات	التكرار f_i	ت.م.ص Fv	ت.م.ص.ن Fv^*
46-	Less than 53	07	07	7%
53-	Less than 60	15	22	22%
60-	Less than 67	27	49	49%
67-	Less than 74	21	70	70%
74-	Less than 81	14	84	84%
81-	Less than 88	08	92	92%
88-	Less than 95	05	97	97%
95- <102	Less than 102	03	100	100%
المجموع		*100*		

وهذا يعني ان *22* طالب تقل اوزانهم عن *60* كغم وان *70* طالب تقل اوزانهم عن *74* كغم كذلك فان نسبة عدد الطلبة الذين تقل اوزانهم عن *95* كغم هي *97%* وهكذا …

(٤-٥-٢)التوزيع التكراري المتجمع النازل Descending Cumulative Frequency Distribution

وهو التوزيع الذي يبين تناقص التكرارات ابتداءً من الفئة الاولى في التوزيع وانتهاءً بالفئة الاخيرة منه ، ويتم حساب التكرارات المتجمعة على اساس الحدود الدنيا للفئات وسوف نرمز له بالرمز (***FF***) .

في حالة كون المتغير العشوائي من النوع المتقطع أوالمستمر

ان التوزيع التكراري المتجمع النازل في هذه الحالة سوف يأخذ الشكل الاتي :

الحدود العليا للفئات	عبارة التجميع	التكرار الاصلي	التكرار المتجمع النازل (FF)
X_S X_S+L X_S+2L $\vdots$ $X_S+(m-1)L$	اكبر من او يساوي قيمة الحد الادنى للفئة	f_1 f_2 f_3 $\vdots$ f_m	$FF_1=f_1+f_2+\cdots+f_m=n$ $FF_2=n-f_1=FF_1-f_1$ $FF_3=n-f_1-f_2=FF_2-f_2$ $\vdots$ $FF_m=n-f_1-f_2-\cdots-f_{m-1}=FF_{m-1}+f_{m-1}=f_m$
المجموع		n	

ويمكن حساب التوزيع المتجمع النازل النسبي والذي سنرمز له بالرمز ***FF****** ، حيث ان :

$$FF^* = \frac{FF_i}{n} \cdot 100$$

مثال (٧) : بالرجوع الى المثال رقم (٥) ، المطلوب تكوين توزيع تكراري متجمع نازل ، وتوزيع تكراري متجمع نازل نسبي .

الفئات	الحدود العليا للفئات	التكرار f_i	ت.م.ن FF	ت.م.ن.ن FF^*
60-74	اكبر من او يساوي*60*	4	60	100%
75-89	اكبر من او يساوي 75	5	56	93.33%
90-104	اكبر من او يساوي *90*	10	51	85%
105-119	اكبر من او يساوي *105*	12	41	68.33%
120-134	اكبر من او يساوي *120*	16	29	48.33%
135-149	اكبر من او يساوي 135	7	13	21.67%
150-164	اكبر من او يساوي *150*	6	6	10%
المجموع		*60*		

مثال (٨) : بالرجوع الى مثال (٦) ، المطلوب : تكوين توزيع تكراري متجمع نازل ، وتوزيع تكراري متجمع نازل نسبي .

الفئات	الحدود الدنيا للفئات	التكرار f_i	ت.م.ن FF	ت.م.ن.ن FF^*
46-	اكبر من او يساوي *46*	07	100	100%
53-	اكبر من او يساوي *53*	15	93	93%
60-	اكبر من او يساوي *60*	27	78	78%
67-	اكبر من او يساوي *67*	21	51	51%
74-	اكبر من او يساوي *74*	14	30	30%
81-	اكبر من او يساوي *81*	08	16	16%
88-	اكبر من او يساوي *88*	05	08	8%
95- <102	اكبر من او يساوي *95*	03	03	3%
المجموع		*100*		

(٤-٦) العرض الهندسي للبيانات

ان الرسوم والاشكال الهندسية ما هي الا تعبير وتوضيح للبيانات بطريقة جذابة وسهلة وفعالة تساعد القارئ على فهم واستيعاب قيم الظاهرة ومقارنتها مع بعضها . ان وسائل التمثيل البياني كثيرة ومتنوعة وسنكتفي بشرح العرض البياني للتوزيعات التكرارية فقط . وعادة يخصص المحور الافقي او المحور السيني (*X-axis*) ليمثل قيم او فئات المتغير ، بينما يخصص المحور العمودي او المحور الصادي (*Y-axis*) ليمثل تكرارات هذا المتغير ، وهذه الاشكال والرسوم هي :

(٤-٦-١) المدرج والمضلع والمنحني التكراري

وهي عبارة عن اشكال هندسية الهدف منها عرض البيانات المبوبة في جداول التوزيع ولهذا النوع من الاشكال اهمية كبيرة في دراسة خصائص التوزيعات التكرارية وكذلك في حساب بعض المؤشرات الاحصائية التي سنأتي لدراستها في الفصول القادمة .

اولاً: المدرج التكراري Histogram

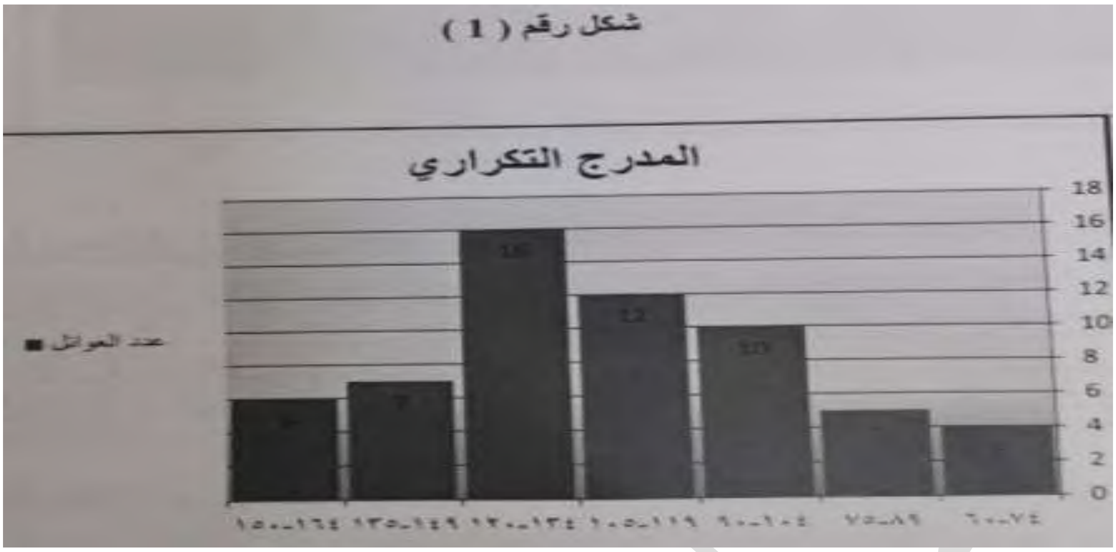
عبارة عن مجموعة من المستطيلات قاعدة كل منها تمثل طول الفئة في التوزيع التكراري وارتفاع كل منها يمثل قيمة التكرار المقابل لتلك الفئة ، هذه المستطيلات تكون منفصلة عن بعضها في حالة المتغيرات المتقطعة وتكون متصلة مع بعضها في حالة المتغيرات المستمرة وحسب تسلسل فئات التوزيع .

وفي حالة اختلاف اطوال الفئات في ذلك التوزيع عندئذ وقبل رسم المدرج التكراري يستوجب الامر تعديل التكرارات من خلال قسمة التكرار الاصلي على طول تلك الفئة ، فاذا رمزنا للتكرار المعدل بالرمز f° فأن :

$$f_i^{\circ} = \frac{f_i}{L_i} \qquad , i = 1,2,3,\dots,n$$

مثال (١) في حالة المتغير المتقطع : الاتي توزيع تكراري يمثل توزيع *60* عائلة فلاحية حسب ملكيتها من عدد اشجار البرتقال .

عدد الاشجار	60-74	75-89	90-104	105-119	120-134	135-149	150-164
عدد العوائل	4	5	10	12	16	7	6

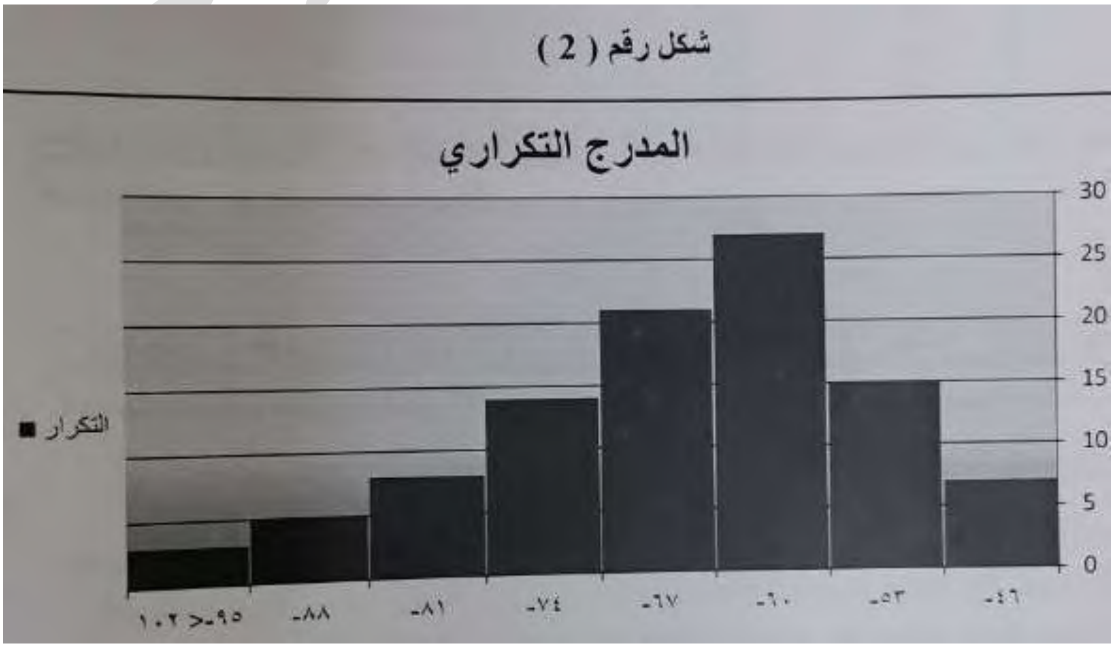


الواضح وحسب البيانات الموزعة في جدول التوزيع التكراري ان المتغير العشوائي من النوع المتقطع وعليه فان المدرج التكراري سيأخذ الشكل اعلاه ، بعد اختيار مناسب لقياس طول الفئة.

مثال (٢) في حالة المتغير المستمر : الاتي جدول توزيع تكراري لأوزان عينه من طلبة احدى الكليات قوامها *100* طالب ، يطلب رسم المدرج التكراري لهذا التوزيع .

الفئات	*46-*	*53-*	*60-*	*67-*	*74-*	*81-*	*88-*	*95- <102*
التكرارات	*7*	*15*	*27*	*21*	*14*	*8*	*5*	*3*

من الواضح ان المتغير العشوائي من النوع المستمر ، وعليه فان المدرج التكراري سيأخذ الشكل الاتي ، بعد اختيار مناسب لقياس طول الفئة .



مثال (٣) في حالة كون اطوال الفئات غير متساوية : لجدول التوزيع التكراري الاتي ، يطلب رسم المدرج التكراري لهذا التوزيع .

75- <100	60-	40-	30-	15-	10-	الفئات
20	15	30	25	9	5	التكرارات

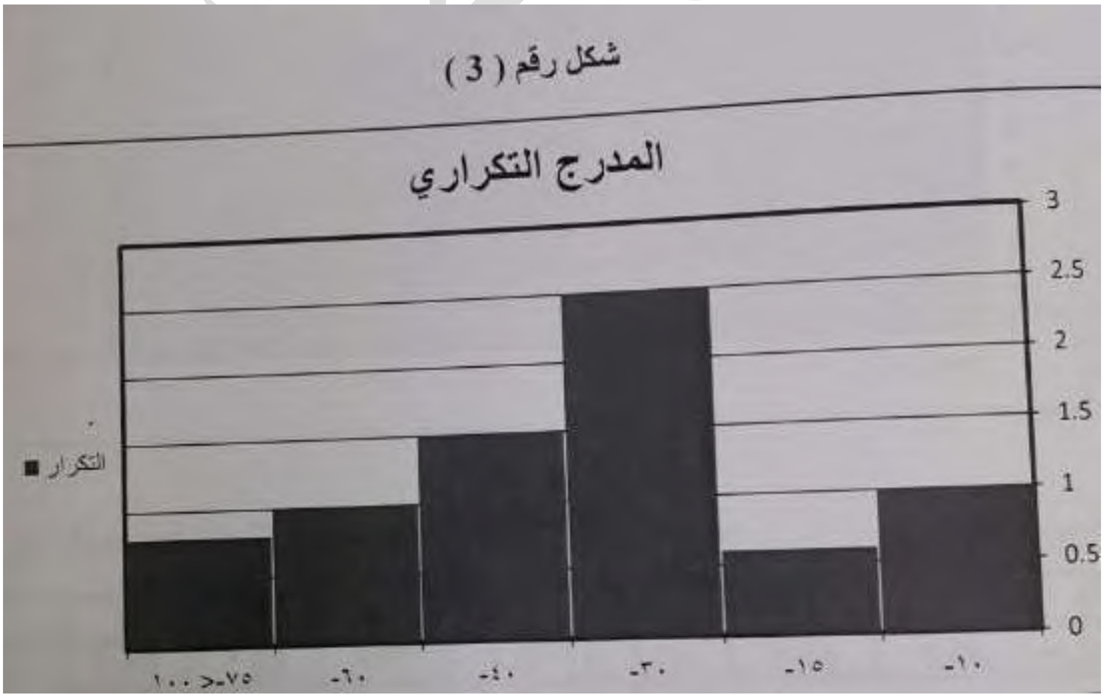
بما ان اطوال الفئات غير متساوية ، لذا يستوجب الامر تعديل التكرارات قبل رسم المدرج وعلى النحو الاتي :

75- <100	60-	40-	30-	15-	10-	الفئات
20	15	30	25	9	5	التكرار f_i
25	15	20	10	15	5	طول الفئة
0.8	1	1.5	2.5	0.6	1	التكرار المعدل f_i°

علما بان : التكرار المعدل f_i° = التكرار / طول الفئة = $\frac{f_i}{L_i}$

$$0.8 = \frac{25}{20} = f_6^{\circ}, \dots\dots\dots\dots, 0.6 = 15/9 = f_2^{\circ}, 1 = 5/5 = f_1^{\circ}$$

بعد ذلك يتم رسم المدرج التكراري المطلوب على اساس التكرارات المعدلة وعلى النحو الموضح في الشكل الاتي :



ثانياً : المضلع التكراري Frequency polygon

عبارة عن عدد من المستقيمات المتصلة مع بعضها على شكل سلسلة ونقطة اتصال المستقيم بالآخر تقابل مركز الفئة ، وهذا يعني انه عند رسم مضلع تكراري يستوجب الامر ايجاد مراكز الفئات ومن ثم رسم المضلع على اساس ازواج القيم (مركز الفئة والتكرار) وعادة يفضل غلق المضلع التكراري مع المحور السيني وذلك باختيار مركز فئة وهمي قبل مركز الفئة الاولى وآخر بعد مركز الفئة الاخيرة ونفترض ان تكرار هذين المركزين مساوٍ للصفر وتتم عملية الغلق بخط منقط وكما موضح في الامثلة التالية :

مثال(٤) : بالرجوع الى المثال (١) يطلب رسم المضلع التكراري .

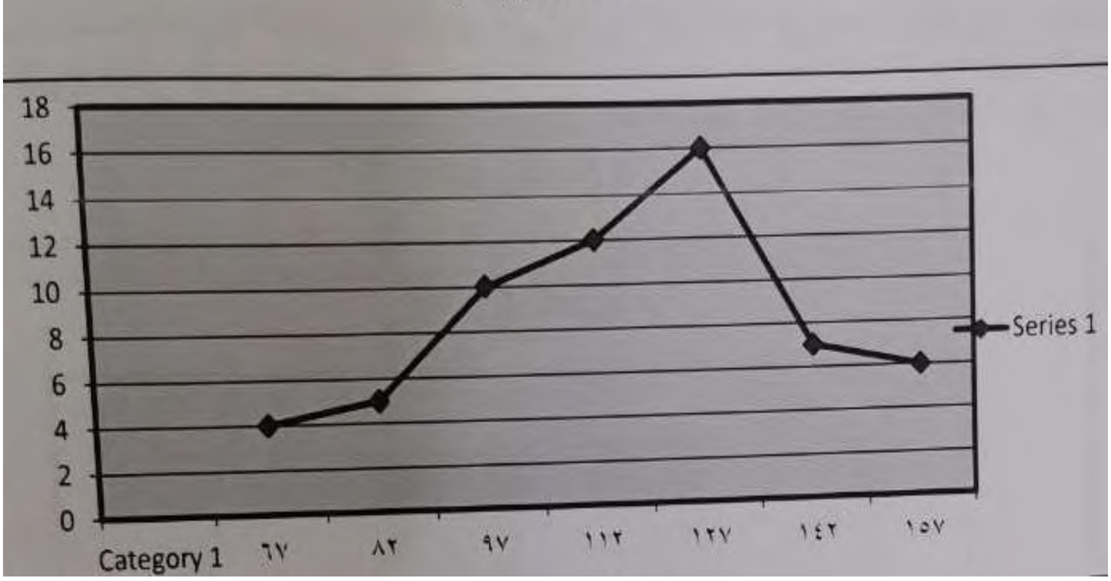
الحل/ يجب اولاً ان نستخرج مراكز الفئات ، حيث ان :

مركز الفئة = (الحد الادنى للفئة + الحد الاعلى للفئة) /2

التكرار	4	5	10	12	16	7	6
مراكز الفئات	67	82	97	112	127	142	157

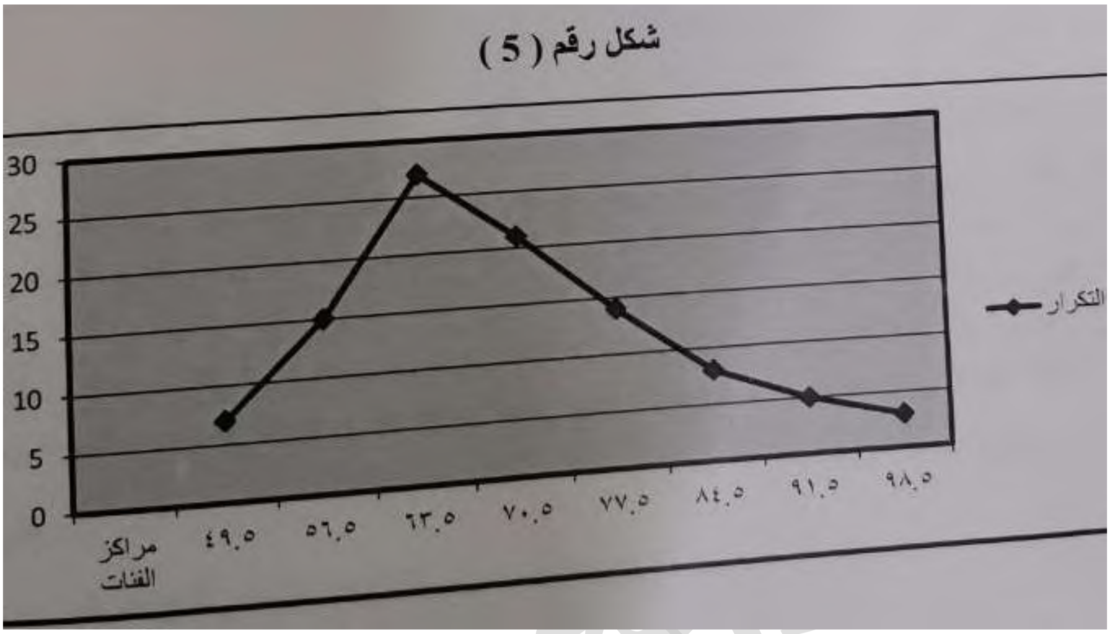
نحدد نقاط مراكز الفئات على المحور السيني ونحدد نقاط التكرار على المحور الصادي ومن ثم نحدد نقاط تقاطع مركز الفئة مع تكرارها ونوصل النقاط بخطوط مستقيمة .

شكل رقم (4)



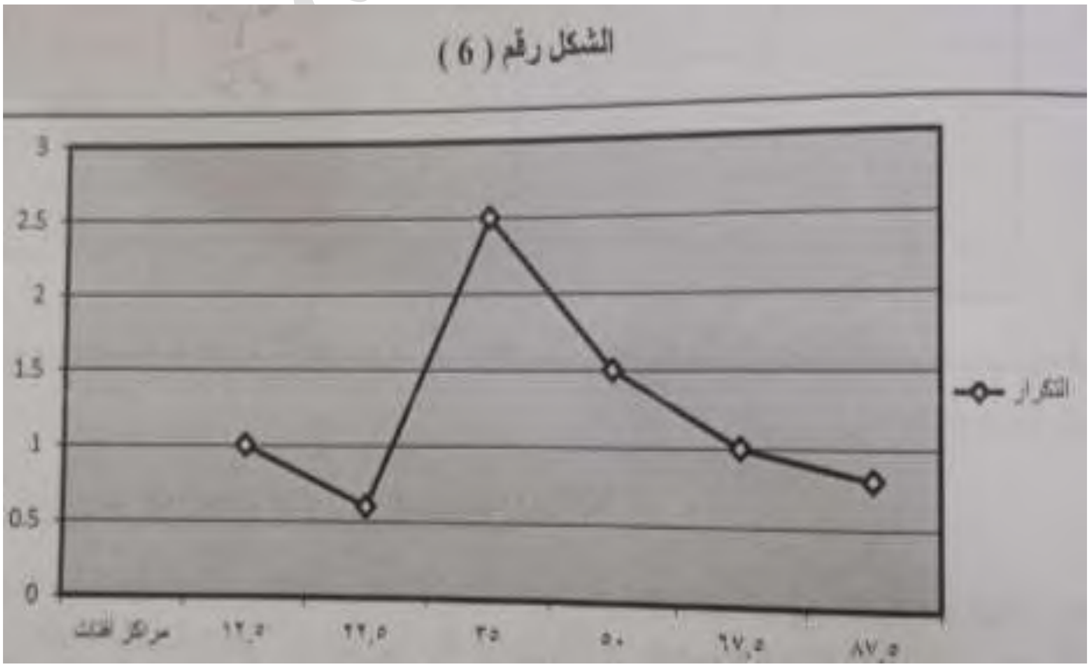
مثال (٥) : بالرجوع الى المثال (٢) ، يطلب رسم المضلع التكراري .

التكرار	7	15	27	21	14	8	5	3
مراكز الفئات	49.5	56.5	63.5	70.5	77.5	84.5	91.5	89.5



مثال (٦) : بالرجوع الى المثال (٣) ، يطلب رسم المضلع التكراري .

التكرار	1	0.6	2.5	1.5	1	0.8
مراكز الفئات	12.5	22.5	35	50	67.5	87.5



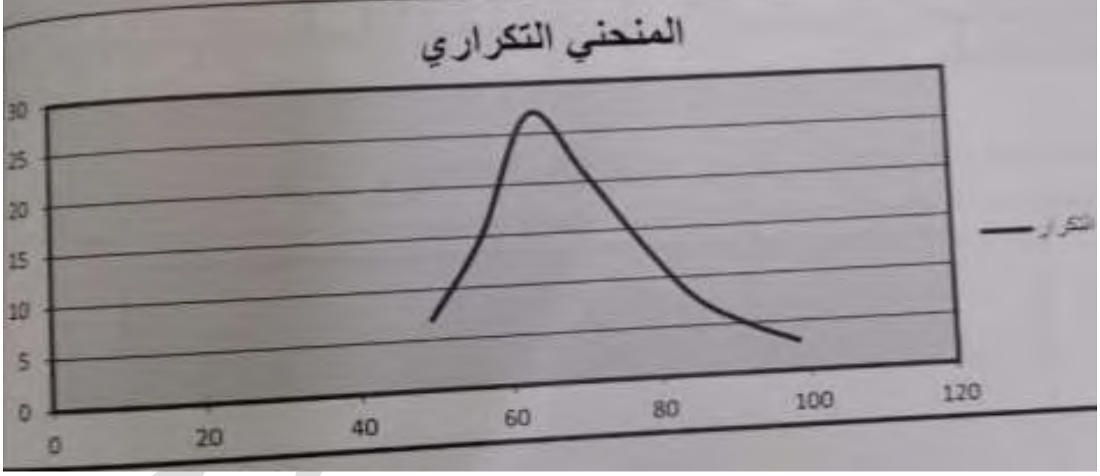
ثالثاً : المنحني التكراري Frequency Curve

لا تختلف فكرة رسم المنحني التكراري عن المضلع التكراري من حيث الاسلوب لكن الفرق الوحيد ما بينهما هو انه بدلا من توصيل النقاط (مركز الفئة ، التكرار) بمستقيمات فانه يتم تمرير منحني ما بين هذه النقاط ، هذا المنحني يمثل المنحني التكراري لذلك التوزيع ، ان المنحني التكراري يتم رسمه للتوزيعات التكرارية الخاصة بالمتغيرات من النوع المستمر فقط .

مثال : بالرجوع الى المثال (٥) ، يطلب رسم المنحني التكرار .

الحل : نحدد نقاط مراكز الفئات على المحور السيني ونحدد نقاط التكرار على المحور الصادي ويتم توصيل نقاط التقاطع مركز الفئة مع التكرار بمنحني وكما موضح في الشكل الاتي :

الشكل رقم (7)



- تمارين -

س١/ البيانات الاتية تمثل عدد افراد عينه من الاسر قوامها 70 اسرة ، يطلب :

١. حدد نوع المتغير العشوائي للبيانات .
٢. تفريغ البيانات في جدول توزيع تكراري .
٣. تكوين جدول توزيع تكراري متجمع صاعد ومتجمع صاعد نسبي .
٤. تكوين جدول توزيع تكراري متجمع نازل ومتجمع نازل نسبي .
٥. رسم المدرج التكراري للتوزيع والمضلع التكراري للتوزيع .

5	9	2	2	6	6	8	4	13	9	6	4	3	2
15	14	16	14	12	13	11	6	22	20	8	7	6	6
9	13	12	14	15	11	8	10	9	15	12	10	9	18
17	18	10	13	15	8	9	12	15	17	18	9	9	8
20	11	12	8	10	14	15	8	10	8	9	9	11	12

س٢/ البيانات الاتية تمثل معدلات مجموعة من الطلبة خريجي الدراسة الاعدادية مقربة لمرتبة عشرية واحدة . يطلب :

١. حدد نوع المتغير العشوائي للبيانات .
٢. تفريغ البيانات في جدول توزيع تكراري .
٣. تكوين جدول توزيع تكراري متجمع صاعد ومتجمع صاعد نسبي .
٤. تكوين جدول توزيع تكراري متجمع نازل ومتجمع نازل نسبي .
٥. رسم المنحني التكراري للتوزيع .

70.8	68	64.2	53	82.3	85.9	95	79	68.1	55.2
58.3	89	92.1	55	71.3	70.7	89	86	63.2	75.1
63.1	66.2	81.6	82.5	94.1	93	80	64.1	71.3	62.2
86.6	82.2	83.1	86.9	91	55.3	54	69.2	74.2	71.8
81.7	69	62.9	60.3	60.1	87.7	62.5	66.1	76.3	74

الفصل الخامس // مقاييس النزعة المركزية – مقاييس التوسط او التمركز-

(٥-١) رموز ومصطلحات احصائية

سنتعامل مع الرموز والمعادلات اللاتينية كما هي بدون تغيير لكونها رموزا عالمية من جهة ولسهولة الاستفادة والاستنارة بالمراجع الاجنبية .

سوف نرمز للمتغير بالرمز X ولكل قيمة له بالرمز x_i ، فلو كان اعمار 5 طلاب كالاتي : 16,22,24,18,20 سنة ، فنكتب :

$x_i = 16,22,24,18,20$

اي ان $x_1 = 16$ تعني القيمة الاولى للمتغير، $x_2 = 22$ القيمة الثانية للمتغير وهكذا ... الى $x_5 = 20$ اي القيمة الاخيرة $n = 5$ ، ويرمز عادة لمجموع قيم المتغير بالرمز

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

فالرمز $\sum$ هو حرف اغريقي يسمى *Sigma* اي مجموع او *Summation of* والرقمان *1* و *n* هما حدا المجموع .

- مجموع قيم عناصر السلسلة :

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

- مجموع مربعات قيم عناصر السلسلة :

$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

(٥-٢) بعض القواعد المفيدة في عملية الجمع

١. اذا كانت a اي عدد ثابت فأن :

$$\sum_{i=1} a = na$$

proof:

$$\sum_{i=1}^{n} a = a + a + \cdots + a = na$$

٢. اذا كانت a اي عدد ثابت فأن :

$$\sum_{i=1}^{n} a\,x_i = a\sum_{i=1}^{n} x_i$$

Proof :

$$\sum_{i=1}^{n} a\,x_i = ax_1 + ax_2 + \cdots + ax_n$$

$$= a(x_1 + x_2 + \cdots + x_n)$$

$$= a\sum_{i=1}^{n} x_i$$

٣. جمع (طرح) قيم متغيرين او اكثر هو مجموع جمعهم اي :

$$\sum_{i=1}^{n} (X_i \pm Y_i) = \sum_{i=1}^{n} X_i \pm \sum_{i=1}^{n} Y_i$$

Proof:

$$\sum_{i=1}^{n} (X_i \pm Y_i) = (X_1 \pm Y_1) + (X_2 \pm Y_2) + \cdots + (X_n \pm Y_n)$$

$$= (X_1 + X_2 + \cdots + X_n) \pm (Y_1 + Y_2 + \cdots + Y_n)$$

$$= \sum_{i=1}^{n} X_i \pm \sum_{i=1}^{n} Y_i$$

٤. مربع مجموع قيم عناصر السلسلة :

$$\left(\sum_{i=1}^{n} X_i\right)^2 = (X_1 + X_2 + \cdots + X_n)^2$$

٥. مجموع حاصل ضرب قيم متغيرين X و Y :

$$\sum_{i=1}^{n}(X_i.Y_i) = (X_1.Y_1) + (X_2.Y_2) + \cdots + (X_n.Y_n)$$

٦. حاصل ضرب مجموعين لقيم متغيرين :

$$\sum_{i=1}^{n} X_i . \sum_{i=1}^{n} Y_i = (X_1 + X_2 + \cdots + X_n).(Y_1 + Y_2 + \cdots + Y_n)$$

٧. مجموع مقلوب عناصر السلسلة :

$$\sum_{i=1}^{n}\frac{1}{X_i} = \frac{1}{X_1} + \frac{1}{X_2} + \cdots + \frac{1}{X_n}$$

٨. مقلوب مجموع عناصر السلسلة :

$$\frac{1}{\sum_{i=1}^{n} X_i} = \frac{1}{X_1 + X_2 + \cdots + X_n}$$

٩. مجموع لوغاريتمات عناصر السلسلة :

$$\sum_{i=1}^{n} \log X_i = \log X_1 + \log X_2 + \cdots + \log X_n$$

١٠. مجموع جذر عناصر السلسلة :

$$\sum_{i=1}^{n}\sqrt{X_i} = \sqrt{X_1} + \sqrt{X_2} + \cdots + \sqrt{X_n}$$

١١. مجموع حاصل اضافة او طرح الثابت a الى عناصر السلسلة X_i يمثل مجموع عناصر السلسلة X_i مضافاً له او مطروحاً منه حاصل ضرب الثابت a بعدد العناصر وكما يأتي :

$$\sum_{i=1}^{n}(X_i \pm a) = \sum_{i=1}^{n} X_i \pm na$$

Proof:

$$\sum_{i=1}^{n}(X_i \pm a) = (X_1 \pm a) + (X_2 \pm a) + \cdots + (X_n \pm a)$$

$$= (X_1 + X_2 + \cdots + X_n) \pm (a + a + \cdots + a)$$

$$= \sum_{i=1}^{n} X_i \pm na$$

ملاحظات يجب مراعاتها :

١.

$$\sum_{i=1}^{n}(X_i . Y_i) \neq \sum_{i=1}^{n} X_i . \sum_{i=1}^{n} Y_i$$

٢.

$$\sum_{i=1}^{n} X_i^2 \neq \left(\sum_{i=1}^{n} X_i\right)^2$$

٣.

$$\sum_{i=1}^{n} \frac{X_i}{Y_i} \neq \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} Y_i}$$

مثال : اذا علمت بأن قيم كل من المتغيرين Y و X هي كالاتي :

$$X_i = 2,6,3,1$$

$$Y_i = 3,9,6,2$$

جد قيمة كل من :

a.

$$\sum_{i=1}^{n}(Y_i - X_i)^2$$

b.

$$\sum_{i=1}^{n}(X_i - 3)(Y_i - 5)$$

c.

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

d.

$$\sum_{i=1}^{n} X_i Y_i - \frac{(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i)}{n}$$

(٥-٣) رمزالضرب $\prod$

ان الرمز $\prod$ هو دليل لوجود عملية ضرب مجموعة من الكميات فعلى افتراض ان $X_1, X_2, \dots, X_n$ تمثل عناصر سلسلة من الكميات ، فان حاصل ضرب هذه الكميات هو $X_1 * X_2 * \dots * X_n$ وباستخدام رمز الضرب فان التعبير عن هذه العملية هو كالاتي :

$$\prod_{i=1}^{n} X = X_1 * X_2 * \dots * X_n$$

وفيما يلي خصائص هذا الرمز وبافتراض $X_1, X_2, \dots, X_n$ تمثل عناصر سلسلة من البيانات عن المتغير (X)، وان $y_1, y_2, \dots, y_n$ تمثل عناصر سلسلة من البيانات عن المتغير (y) وان b, a ثابتين حقيقين ، فأن :

1. $$\prod_{i=1}^{n} a = a^n$$

Proof:

$$\prod_{i=1}^{n} a = a * a * \dots * a = a^n$$

2. $$\prod_{i=1}^{n} a\, X_i = a^n \prod_{i=1}^{n} X_i$$

proof:

$$\prod_{i=1}^{n} a\, X_i = aX_1 * aX_2 * \dots * aX_n = a^n \prod_{i=1}^{n} X_i$$

3. $$\prod_{i=1}^{n} a\, X_i bY_i = a^n b^n \prod_{i=1}^{n} X_i \prod_{i=1}^{n} Y_i$$

Proof:

$$\prod_{i=1}^{n} a\, X_i bY_i = (a\, X_1 * bY_1) * (a\, X_2 * bY_2) * \dots * (a\, X_n * bY_n)$$

$$= (ab * ab * \dots * ab)(x_1 y_1 * x_2 y_2 * \dots * x_n y_n)$$

$$= a^n b^n \prod_{i=1}^{n} X_i \prod_{i=1}^{n} Y_i$$

4. $$\prod_{i=1}^{n} \frac{1}{X_i} = \frac{1}{\prod_{i=1}^{n} X_i}$$

5. $$\prod_{i=1}^{n} \frac{a}{X_i} = \frac{a^n}{\prod_{i=1}^{n} X_i}$$

6. $$\prod_{i=1}^{n} \sqrt{X_i} = \sqrt{X_1} * \sqrt{X_2} * \dots * \sqrt{X_n}$$

$$= \sqrt{X_1 * X_2 * \dots * X_n} = \sqrt{\prod_{i=1}^{n} X_i}$$

7. $$\prod_{i=1}^{n} (X_i)^n = \left(\prod_{i=1}^{n} X_i\right)^n$$

8. $$\prod_{i=1}^{n} \frac{X_i}{Y_i} = \frac{\prod_{i=1}^{n} X_i}{\prod_{i=1}^{n} Y_i}$$

(٥-٤) مقاييس النزعة المركزية – مقاييس التوسط او التمركز -

في الفصل السابق استعرضنا اهم اساليب جمع وتصنيف وتبويب البيانات ، وفي هذا الفصل سوف نتطرق للحديث عن كيفية تمثيل مجموعة من البيانات بقيمة واحدة فقط من خلال مقياس النزعة المركزية .

ان معظم القيم لمختلف الظواهر الطبيعية تتمركز عادة في الوسط او القريب منه ، ومقاييس التمركز او التوسط لأي مجموعة من البيانات لظاهرة ما ، هي تلك المقاييس التي تبحث في تقدير قيمه تتمركز حولها اغلبية هذه البيانات ، ومن اهم مقاييس التوسط هي :

(٥-٥) الوسط الحسابي Arithmetic mean

ويسمى في بعض الاحيان (الوسط) او (المتوسط) او (المعدل الحسابي *Average*) وهو احد اهم مقاييس النزعة المركزية لما يمتاز بخصائص وسهوله في الحساب ويرمز له بـ $(\bar{X})$ وهو القيمة الناتجة من قسمة مجموع تلك القيم على عددها ، سوف نتطرق في هذه الفقرة

الى طرق حساب الوسط الحسابي لبيانات غير مبوبة ولبيانات مبوبة في جداول التوزيع التكراري .

(٥-٥-١) احتساب الوسط الحسابي لبيانات غير مبوبة :

لنفرض ان $X_1, X_2, \dots, X_n$ عينة من القيم او المشاهدات التي تساوي (n) ، فان الوسط الحسابي لهذه القيم هو :

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad i = 1,2, \dots, n$$

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

حيث ان : $\bar{X}$ = الوسط الحسابي

وتجدر الاشارة هنا الى ان $\bar{X}$ هو تقدير للوسط الحسابي لقياسات مفردات العينة . ان متوسط قياسات مفردات المجتمع غالبا ما يرمز له بالرمز μ ، معطى بالصيغة الاتية في حالة كون مجتمع الدراسة مجتمع محدود .

$$\mu = \frac{\sum_{i=1}^{n} X_i}{N} \qquad \dots (1)$$

مثال (١): جد الوسط الحسابي لكل من القيم الاتية 400,380,450,350,520

الحل :

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{400 + 380 + 450 + 350 + 520}{5} = 420$$

مثال (٢): البيانات الاتية تمثل عدد افراد عينة من الاسر قوامها (١٢) اسرة ، يطلب ايجاد متوسط عدد افراد الاسرة . 3 ,4 ,7 ,8 ,10 ,9 ,2 ,5 ,6 ,9 ,7 ,5

الحل:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{3 + 4 + 7 + 8 + 10 + 9 + 2 + 5 + 6 + 9 + 7 + 5}{12}$$
$$= 6.25$$

وحيث ان عدد افراد الاسرة متغير متقطع لذا يتم تقريب الناتج الى اقرب عدد صحيح لانه لا يوجد قياس لجزء من الفرد ، وعليه فان متوسط عدد افراد الاسرة في هذه العينة هو تقريبا ستة افراد . $\bar{X} \approx 6$

(٥-٥-٢) احتساب الوسط الحسابي لبيانات مبوبة :

لتكن $X_1, X_2, \dots, X_n$ تمثل مراكز فئات توزيع تكراري عدد فئاته m وان $f_1, f_2, \dots, f_n$ تمثل التكرارات المقابلة لهذه الفئات . عندئذ يستخرج الوسط الحسابي لهذا التوزيع وفق ما يأتي سواء كانت اطوال الفئات متساوية ام غير ذلك :

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i X_i}{\sum_{i=1}^{n} f_i} \qquad \dots (2)$$

ان الوسط الحسابي في هذه الحالة يتطلب ايجاد :

١. مركز الفئة (X_i) .

٢. ضرب كل مركز فئة بمقدار تكرارها (f_i) .

٣. يتم تطبيق الصيغة (2) التي تمثل قسمة (مجموع حاصل ضرب مركز كل فئة X تكرارها) على مجموع التكرارات .

مثال (١) : الاتي توزيع التكراري لدرجات الحرارة في مدينة معينة المسجلة لمدة 99 يوما متتاليا ، يطلب حساب متوسط درجة الحرارة في هذه المدينة خلال هذه الفترة .

درجات الحرارة	1-	2-	3-	4-	5-	6-	7-	8-9
عدد الايام	4	8	12	16	20	25	6	8

الحل : نكون الجدول الاتي الذي يبين خطوات حساب الوسط الحسابي

الفئات	التكرار f_i	مركز الفئات X_i	$f_i X_i$
1-	4	1.5	6
2-	8	2.5	20
3-	12	3.5	42
4-	16	4.5	72
5-	20	5.5	110
6-	25	6.5	162.5
7-	6	7.5	45
8-9	8	8.5	68
المجموع	99		525.5

حيث ان مركز الفئة = (الحد الادنى للفئة + الحد الاعلى للفئة)/2

لذا فان مركز الفئة الاولى $= (1+2)/2 = 1.5$

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i X_i}{\sum_{i=1}^{n} f_i} = \frac{525.5}{99} = 5.308$$

وهذا يعني ان متوسط درجة الحرارة في هذه المدينة خلال تلك الفترة كان 5.308 درجة مئوية.

مثال (٢): *Find the arithmetic mean for the following frequency distribution*

جد الوسط الحسابي للتوزيع التكراري الاتي :

Classes	31-40	41-50	51-60	61-70	71-80	81-90	91-100
Freq.	1	2	5	15	25	20	12

الحل:

نستخرج اولا مراكز الفئات ولحساب الوسط الحسابي نستخرج حاصل ضرب كل مركز فئة مع تكرارها وكما موضح في الجدول الاتي :

الفئات *classes*	التكرار f_i	مركز الفئة X_i	$f_i X_i$
31-40	1	35.5	35.5
41-50	2	45.5	91
51-60	5	55.5	277.5
61-70	15	65.5	982.5
71-80	25	75.5	1887.5
81-90	20	85.5	1710
91-100	12	95.5	1146
المجموع	80		6130

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i X_i}{\sum_{i=1}^{n} f_i} = \frac{6130}{80} = 76.62$$

(٦-٥) بعض خصائص الوسط الحسابي some properties of the mean

١. ان مجموع انحرافات قيم المتغير (X_i) عن وسطها الحسابي $\bar{X}$ الذي احتسب منها يكون مساويا للصفر .

a) في حالة البيانات غير المبوبة :

$$\sum_{i=1}^{n}(X_i-\bar{X})=0$$

Proof:

$\sum_{i=1}^{n}(X_i-\bar{X})=\sum_{i=1}^{n}X_i-n\bar{X}$ (لان $\bar{X}$ تعتبر ثابت)

$=n\bar{X}-n\bar{X}=0$ since $\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}\Longrightarrow\sum_{i=1}^{n}x_i=n\bar{X}$

b) في حالة البيانات المبوبة :

$$\sum_{i=1}^{n}f_i(X_i-\bar{X})=0$$

Proof:

$\sum_{i=1}^{n}f_i(X_i-\bar{X})=\sum_{i=1}^{n}X_if_i-\bar{X}\sum_{i=1}^{n}f_i$

$=\bar{X}\sum_{i=1}^{n}f_i-\bar{X}\sum_{i=1}^{n}f_i=0$ since $\bar{x}=\frac{\sum_{i=1}^{n}X_if_i}{\sum_{i=1}^{n}f_i}\Longrightarrow\sum_{i=1}^{n}X_if_i=\bar{X}\sum_{i=1}^{n}f_i$

٢. حاصل ضرب كل قيمة من قيم المشاهدات في قيمة الثابت (k) فان الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الاصلية × العدد الثابت (k) .

$$y_i=kx_i\Longrightarrow\bar{y}=k\bar{x}$$

Proof:

$y_i=kx_i$ (بأخذ المجموع للطرفين)

$\sum_{i=1}^{n}y_i=k\sum_{i=1}^{n}x_i$ (بقسمة الطرفين على n للحصول على الوسط الحسابي)

$\frac{\sum_{i=1}^{n}y_i}{n}=k\frac{\sum_{i=1}^{n}x_i}{n}\Longrightarrow\bar{y}=k\bar{x}$

٣. عند اضافة الثابت (k) الى قيمة من قيم المشاهدات فان الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الاصلية + الحد الثابت (k) .

$$y_i = x_i + k \Longrightarrow \bar{y} = \bar{x} + k$$

Proof:

$y_i = x_i + k$ (بأخذ المجموع للطرفين)

$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} (x_i + k)$

$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i + nk$ (بقسمة الطرفين على n للحصول على الوسط الحسابي)

$\frac{\sum_{i=1}^{n} y_i}{n} = \frac{\sum_{i=1}^{n} x_i}{n} + k \Longrightarrow \bar{y} = \bar{x} + k$

(٥-٧) الوسط الحسابي الموزون (المرجح) *Weighted mean*

i. في حالة البيانات غير المبوبة :

لو افترضنا ان $x_1, x_2, \dots, x_n$ تمثل قياسات عينة من المفردات قوامها n وان $w_1, w_2, \dots, w_n$ تمثل اوزان هذه المفردات ، عندئذ يعرف الوسط الحسابي المرجح $\bar{x}_w$ بهذه الاوزان على النحو الاتي :

$$\bar{x}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} = \frac{x_1 w_1 + x_2 w_2 + \cdots + x_n w_n}{w_1 + w_2 + \cdots + w_n}$$

ii. في حالة البيانات المبوبة :

ان الوسط الحسابي المرجح في هذه الحالة يتطلب ايجاد :

١. مركز الفئة x_i
٢. ضرب كل مركز فئة بمقدار تكرارها (f_i) وبمقدار وزنها (w_i)
٣. تطبيق الصيغة الاتية لإيجاد الوسط الحسابي المرجح في حالة البيانات المبوبة :

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i f_i x_i}{\sum_{i=1}^{n} w_i f_i} = \frac{w_1 f_1 x_1 + w_2 f_2 x_2 + \cdots + w_n f_n x_n}{w_1 f_1 + w_2 f_2 + \cdots + w_n f_n}$$

مثال (١) جد الوسط الحسابي المرجح للبيانات الاتية ((درجات طالب لعدة مواد دراسية (x_i) وعدد الساعات المعطاة لتلك المواد (w_i)))

الدرجات x_i	62	80	75	88	84	86	90
عدد الساعات w_i	2	2	2	3	3	3	3

الحل :

$$\bar{x}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}$$

$$= \frac{(62*2)+(80*2)+(75*2)+(88*3)+(84*3)+(86*3)+(90*3)}{2+2+2+3+3+3+3}$$

$$= \frac{1478}{18} = 82.11$$

مثال (٢) جد الوسط الحسابي المرجح لجدول التوزيع التكراري الاتي :

الفئات	2-<4	4-<6	6-<8	8-<10	10- <12
التكرار f_i	4	5	6	3	2
الوزن w_i	6	5	6	4	4

الحل : اولا نستخرج مركز الفئة (x_i) ولحساب المقاييس المطلوبة لإيجاد الوسط الحسابي المرجح نكون الجدول الاتي :

الفئات	f_i(التكرارات)	W_i(الوزن)	X_i(مركز الفئات)	$W_i f_i$	$X_i w_i f_i$
2-<4	4	6	3	24	72
4-<6	5	5	5	25	125
6-<8	6	6	7	36	252
8-<10	3	4	9	12	108
10- <12	2	4	11	8	88
المجموع	20			105	745

$$\bar{x}_w = \frac{\sum_{i=1}^{n} x_i f_i w_i}{\sum_{i=1}^{n} f_i\, w_i} = \frac{745}{105} = 7.095$$

(٥-٨) الوسط التوافقي Harmonic mean

i. في حالة البيانات غير مبوبة : يعتبر الوسط التوافقي من احد مقاييس التوسط ذات الاستخدامات القليلة في التطبيقات الاحصائية ، ويعرف الوسط التوافقي بانه مقلوب الوسط الحسابي ويرمز له (H) ويحسب وفق الصيغة الاتية :

$$H=\frac{1}{\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)\Big/_{n}}=\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}\qquad ,x_i\neq 0$$

مثال: جد الوسط التوافقي لكل من القيم الاتية : *6, 7, 6, 5, 3, 10, 12*

الحل:

$$H=\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}=\frac{7}{\frac{1}{12}+\frac{1}{10}+\frac{1}{3}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{6}}=5.87$$

ii. في حالة البيانات المبوبة : ان الوسط التوافقي في هذه الحالة يتطلب ايجاد :

a) مركز الفئة (x_i)

b) قسمة كل تكرار فئة (f_i) على مركز الفئة (x_i)

يتم تطبيق القانون الاتي الذي يمثل مجموع قسمة (مجموع التكرارات) على (مجموع حاصل قسمة التكرارات على مراكز فئاتها) وكما يأتي :

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

ملاحظة:

- يمكن ايجاد قيمة الوسط التوافقي سواء كانت اطوال الفئات للتوزيع التكراري متساوية ام غير متساوية .
- لا يمكن تحديد قيمة الوسط التوافقي اذا كانت احدى قيم المتغير العشوائي مساوية للصفر او ان احد مراكز فئات التوزيع التكراري كان مساويا للصفر .

مثال: استخراج الوسط التوافقي لجدول التوزيع التكراري الاتي :

الفئات	50-	60-	70-	80-	90-	100-	110-120
التكرار f_i	8	10	16	14	10	5	5

الحل:

اولا استخراج مراكز الفئات وكما تم توضيحه سابقاً ، ثم نستخرج حاصل قسمة كل تكرار فئة على مركز تلك الفئة .

الفئات	f_i	X_i	f_i/x_i
50-< 60	8	55	0.145
60-< 70	10	65	0.154
70-< 80	16	75	0.213
80-< 90	14	85	0.165
90-<100	10	95	0.105
100-<110	5	105	0.048
110-< 120	5	115	0.043
المجموع	68		0.873

$$H = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{x_i}} = \frac{68}{0.873} = 77.892$$

(٥-٩) الوسط التربيعي *Quadratic mean*

يعرف الوسط التربيعي بانه الجذر التربيعي الموجب لمتوسط مربعات قيم المتغير العشوائي (X) ، ويلاحظ من هذا التعريف ان قيمة الوسط التربيعي قيمة موجبة دائماً وذلك بسبب عملية التربيع ويرمز له بالرمز (Q) ،وفيما يأتي طرق حساب هذا المقياس الذي يطبق بكثرة في العلوم الفيزيائية .

i. في حالة البيانات غير المبوبة : لتكن $x_1, x_2, \dots, x_n$ تمثل قياسات مفردات عينة قوامها *n* عندئذ يمكن حساب الوسط التربيعي استنادا الى الصيغة الاتية :

$$Q = \sqrt[2]{\frac{\sum_{i=1}^{n} {x_i}^2}{n}}$$

مثال: جد الوسط التربيعي للبيانات الاتية : *2, 3, 4, 5, 6*

الحل: نجد مجموع مربعات هذه القيم :

$$\sum_{i=1}^{n} {x_i}^2 = 2^2 + 3^2 + \cdots + 6^2 = 90$$

$$Q = \sqrt[2]{\frac{\sum_{i=1}^{n} {x_i}^2}{n}} = \sqrt{\frac{90}{5}} = 4.243$$

ii. في حالة البيانات المبوبة : ان الوسط التربيعي في هذه الحالة يتطلب ايجاد :

- مركز الفئة (x_i) ، ومن ثم ايجاد مربعات مراكز الفئات .
- حاصل ضرب كل تكرار (f_i) في مربعات مراكز الفئات .

يتم تطبيق الصيغة الاتية التي تمثل الجذر التربيعي لمجموع حاصل ضرب (مربعات مراكز الفئات في تكرارها)مقسومة على (مجموع التكرارات) :

$$Q = \sqrt{\frac{\sum_{i=1}^{n} f_i {x_i}^2}{\sum_{i=1}^{n} f_i}}$$

مثال : الاتي توزيع تكراري لأعمار مجموعة من المرضى الوافدين في احدى المستشفيات ، يطلب حساب الوسط التربيعي لعمر المريض في هذا التوزيع .

فئات العمر	10-	20-	30-	40-	50-	60-< 70
عدد المرضى f_i	2	3	4	10	36	14

الحل:

الفئات	*التكرار* f_i	X_i	${X_i}^2$	$f_i{X_i}^2$
10-	2	15	225	450
20-	3	25	625	1875
30-	4	35	1225	4900
40-	10	45	2025	20250
50-	36	55	3025	108900
60-< 70	14	65	4225	59150
المجموع	69			195525

$$Q = \sqrt{\frac{\sum_{i=1}^{n} f_i {x_i}^2}{\sum_{i=1}^{n} f_i}} = \sqrt{\frac{195525}{69}} = 53.232$$

(٥-١٠) الوسط الهندسي Geometric mean

يعتبر الوسط الهندسي احد مقاييس النزعة المركزية المهمة جدا في الدراسات السكانية وخصوصا عند حساب معدلات نمو السكان وكذلك في تكوين الارقام القياسية. ويعرف الوسط الهندسي لمجموعة قياسات متغير عشوائي بأنه الجذر ذي المرتبة n لحاصل ضرب قياسات هذه المجموعة ببعضها البالغ عددها n ، ويلاحظ من هذا التعريف ان قيمة الوسط الهندسي هو عدد موجب ويرمز له بالرمز (G) . وفيما يلي طرق حساب هذا المقياس :

i. في حالة البيانات غير المبوبة : لتكن $x_1, x_2, \dots, x_n$ تمثل قياسات عينة من المفردات قوامها n ، بحيث ان $x_i > 0$ لجميع قيم i ، يمكن حساب الوسط الهندسي وفق الصيغة الاتية :

$$G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

ويمكن ايجاد قيمة الوسط الهندسي باستخدام اللوغاريتم وذلك بأخذ اللوغاريتم للطرفين وكما يأتي :

$$\log G = [\log(x_1) + \log(x_2) + \cdots + \log(x_n)]/n$$

مثال : جد الوسط الهندسي للبيانات الاتية : *10, 20, 30, 40, 50, 60*

الحل:

$$G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} = (10 * 20 * 30 * 40 * 50 * 60)^{\frac{1}{6}}$$

$$= (720000000)^{\frac{1}{6}} = 29.938$$

ii. في حالة البيانات المبوبة : ان الوسط الهندسي في هذه الحالة يتطلب ايجاد :

- مركز الفئة (x_i).
- حاصل ضرب مركز الفئات المرفوع الى تكرارها .
- تطبيق الصيغة الاتية :

$$G = \sqrt[\Sigma f_i]{\prod_{i=1}^{n} x_i^{f_i}} = \left(\prod_{i=1}^{n} x_i^{f_i}\right)^{1/\Sigma f_i}$$

وباستخدام اللوغاريتم فان الصيغة تكون بالشكل الاتي :

$$\log G = \frac{\sum f_i\,(log x_i)}{\sum f_i} = \frac{f_1 log x_1 + f_2 log x_2 + \cdots + f_n log x_n}{f_1 + f_2 + \cdots + f_n}$$

مثال: جد الوسط الهندسي لجدول التوزيع التكراري الاتي :

الفئات	10-	20-	30-	40-	50-	60-	70-< 80
التكرار	3	5	7	8	6	5	2

الحل: اولا / نستخرج اولا مراكز الفئات .

الفئات	10-	20-	30-	40-	50-	60-	70-< 80
التكرار	3	5	7	8	6	5	2
مركز الفئة X_i	15	25	35	45	55	65	75

ثانيا/ نطبق الصيغة الاتية ، علما ان مجموع التكرارات =36 :

$$G = \sqrt[\Sigma f_i]{\prod_{i=1}^{n} x_i^{f_i}}$$

$$G = \sqrt[36]{15^3 * 25^5 * 35^7 * 45^8 * 55^6 * 65^5 * 75^2} = 40.35$$

(٥-١١) المنوال *The Mode*

يعرف المنوال بانه تلك القيمة التي تتكرر اكثر من غيرها من بين مجموعة من القيم وانها القيمة الشائعة من بين مجموعة من القيم .

i. ايجاد المنوال لبيانات غير مبوبة : لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات عينة من المفردات قوامها *n* ، وافرض ان x_i قيمة من قيم هذه المجموعة لوحظ انها تكررت اكثر من غيرها ، عندئذ وحسب تعريف المنوال فان x_i تمثل المنوال لهذه المجموعة ، ويرمز له بـ(*Mo*) .

مثال : للبيانات الاتية جد المنوال : 2, 3, 2, 4, 2, 5, 4, 4, 5, 4, 6, 8, 9, 4, 7, 3, 7, 6

الحل:

واضح من هذه المجموعة ان العدد 4 قد تكرر خمس مرات وهو اكبر من تكرار اي عدد آخر ، وعليه فان المنوال لهذه المجموعة هو 4 . $\rightarrow \mathrm{Mo} = 4$

مثال : جد المنوال للبيانات الاتية : 2, 4, 3, 6, 8, 7, 10, 12

الحل :

واضح من هذه المجموعة انه لا يوجد عدد متكرر اكثر من غيره ، وعليه فانه لا يوجد منوال لهذه المجموعة .

ملاحظة :

قد يكون هناك منوالا واحداً (قيمة واحدة) لهذه المفردات وعندها يسمى التوزيع وحيد القيمة ، او يكون لها منوالان (قيمتان) عندها يسمى التوزيع ذو قيمتين ، وقد يكون لها اكثر من منوالين ، كما انه قد لا يكون هناك منوال للمجموعة .

ii. ايجاد المنوال لبيانات مبوبة :

اولاً: في حالة المتغير العشوائي من النوع المتقطع :

لتكن $x_1, x_2, \dots, x_m$ تمثل مراكز فئات توزيع تكراري عدد فئاته m ، وان $f_1, f_2, \dots, f_m$ تمثل التكرارات المقابلة لفئات هذا التوزيع ، عندئذ فان المنوال لهذا التوزيع يمثل قيمة مركز تلك الفئة التي تقابل اكبر تكرار في التوزيع .

المنوال = مركز الفئة التي تقابل اكبر تكرار في التوزيع

مثال : جد المنوال للتوزيع التكراري ادناه .

الفئات	60-74	75-89	90-104	105-119	120-134
التكرار f_i	2	6	14	10	8

الحل: اكبر تكرار موجود في هذا التوزيع هو (14) المقابل للفئة الثالثة (90-104) وعليه فان المنوال هو مركز الفئة الثالثة

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

ثانياً: في حالة المتغير العشوائي من النوع المستمر :

اذا كانت القيم $x_1, x_2, \dots, x_n$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها $f_1, f_2, \dots, f_n$ على التوالي ، فان المنوال :

$$Mo = L_1 + \left(\frac{d_1}{d_1 + d_2}\right) w$$

حيث ان : فئة المنوال : تلك الفئة التي تملك اكبر التكرارات ، وان :

L_1 : الحد الادنى لفئة المنوال .

d_1 : الفرق بين تكرار فئة المنوال والفئة السابقة لها .

d_2 : الفرق بين تكرار فئة المنوال وتكرار الفئة اللاحقة لها .

w : طول الفئة .

مثال : جد المنوال لجدول التوزيع التكراري الاتي :

المجموع	72-75	-69	-66	-63	-60	الفئات
100	8	27	42	18	5	التكرار f_i

الحل : فئة المنوال = الفئة الثالثة (66-69) والتي لها اكبر تكرار (42)

$$L_1 = 66 \quad , d_1 = 42 - 18 = 24 \quad , d_2 = 42 - 27 = 15 \quad , w = 3$$

$$Mo = L_1 + \left(\frac{d_1}{d_1 + d_2}\right) w = 66 + \left(\frac{24}{24 + 15}\right) 3 = 67.8$$

(٥-١٢) الوسيط *The Median*

يعتبر الوسيط احد مقاييس النزعة المركزية المهمة في التطبيقات الاحصائية ، ويعرف الوسيط بانه تلك القيمة من قيم المتغير العشوائي X التي تقسم مجموعة من قيم المتغير الى قسمين متساويين . اي انها قيمة X التي تجعل عدد القيم التي قبلها مساوٍ لعدد القيم بعدها ، ويرمز له بالرمز (Me) .

***i* . ايجاد الوسيط لبيانات غير مبوبة :**

لتكن $X_1, X_2, \dots, X_n$ تمثل قياسات مفردات عينة قوامها n مفردة ، اولا نرتب هذه القياسات تصاعدياً او تنازلياً عندئذ :

١. اذا كان عدد القيم n عدد فردي *odd number* ، فان قيمة الوسيط تمثل قيمة X بعد الترتيب (التصاعدي او التنازلي) والتي تسلسلها هو $\left[\frac{n+1}{2}\right]$.

مثال : الاتي درجات عينة من الطلبة قوامها (9) طلاب في امتحان معين ، جد الوسيط لهذه المجموعة . *55, 62, 53, 70, 68, 65, 63, 79, 80*

الحل: اولا / نرتب القيم تصاعدياً وكما يأتي : *53, 55, 62, 63, 65, 68, 70, 79, 80*

ثانيا/ نجد ترتيب الوسيط $\leftarrow \quad \frac{n+1}{2} = \frac{10}{2} = 5$

وهذا يعني ان القيمة الخامسة هي قيمة الوسيط اي الدرجة (65)

$$\rightarrow Me = 65$$

٢. اذا كان عدد القيم n عدد زوجي *even number* ، عندئذ فان قيمة الوسيط تمثل الوسط الحسابي (المعدل) لقيمتي X بعد الترتيب واللتين تسلسلهما هو: $\left[\frac{n}{2} + 1\right]$ و $\left[\frac{n}{2}\right]$.

مثال : الاتي اعمار عينة من الافراد قوامها (12) فرد ، جد الوسيط لعمر الفرد في هذه المجموعة . *20, 22, 19.5, 26, 24.5, 27, 28, 29, 18, 20, 23, 25*

الحل : اولا/ نرتب القيم تنازلياً *29, 28, 27, 26, 25, 24.5, 23, 22, 20, 20, 19.5, 18*

ثانيا/ نجد ترتيب الوسيط وعليه فان تسلسل القيمتان اللتان تحددان الوسيط هو :

$$\frac{n}{2} = \frac{12}{2} = 6 \qquad , \qquad \frac{n}{2} + 1 = \frac{12}{2} + 1 = 7$$

اذن القيمتين اللتين تسلسلهما هو 6 و 7 بعد ترتيب القيم تنازليا هما (24.5, 23) وبذلك فان الوسيط لهذه المجموعة يمثل الوسط الحسابي لهاتين القيمتين ، اي قيمة الوسيط هي :

$$\therefore Me = \frac{24.5+23}{2} = 23.75$$

ii . ايجاد الوسيط لبيانات مبوبة :

اولا : في حالة المتغير متقطع :

في هذه الحالة يجب ان يكون الوسيط مساوياً لمركز الفئة الوسيطية ، ويحسب وفق الخطوات الاتية :

١. تحديد ترتيب الوسيط وهو نصف مجموع التكرارات ، اي أن :

$$\text{ترتيب الوسيط} = \frac{\sum_{i=1}^{n} f_i}{2}$$

٢. ايجاد التكرار المتجمع الصاعد (F)

٣. مقارنة ترتيب الوسيط مع التكرار المتجمع الصاعد (Fv) فاذا كان :

$$F_k \leq \frac{\sum f_i}{2} \leq F_{k+1}$$

عندئذ يقال ان فئة الوسيط هي الفئة التي تسلسلها هو (k+1) ، وبذلك فان قيمة الوسيط تمثل مركز هذه الفئة .

مثال : الاتي توزيع تكراري لعينة من الاسر قوامها (80) اسرة حسب عدد افراد الاسرة ، يطلب حساب الوسيط لعدد افراد الاسرة .

20-22	17-19	14-16	11-13	8-10	5-7	2-4	عدد الافراد
8	11	14	20	12	9	6	عدد الاسر (f_i)

الحل: اولا / نجد التكرار المتجمع الصاعد للتوزيع اعلاه :

20-22	17-19	14-16	11-13	8-10	5-7	2-4	عدد الافراد
8	11	14	20	12	9	6	عدد الاسر (f_i)
22	19	16	13	10	7	4	الحدود العليا للفئات
80	72	61	47	27	15	6	التكرار المتجمع الصاعد F_i

ثانيا/ نستخرج ترتيب الوسيط $= \frac{\sum_{i=1}^{n} f_i}{2} = \frac{80}{2} = 40$

ثالثا/ نقارن ترتيب الوسيط مع التكرار المتجمع الصاعد ، ونلاحظ بان ترتيب الوسيط يقع بين $27 \le 40 \le 47$ ، وعليه فان الفئة الوسيطية هي الفئة الرابعة من التوزيع وهي *(11-13)* وعليه فان قيمة الوسيط لهذا التوزيع يمثل مركز الفئة الرابعة اي ان :

$$Me = \frac{11 + 13}{2} = 12$$

ملاحظة :

اذا كان الناتج قيمة كسرية يتم التقريب الى اقرب عدد صحيح لان المتغير في السؤال هو عدد افراد الاسرة .

ثانيا: في حالة المتغير المستمر :

لو فرضنا وجود توزيع تكراري عدد فئاته m وان $f_1, f_2, \ldots, f_m$ تمثل التكرارات المقابلة لفئات التوزيع ، وان $F_1, F_2, \ldots, F_m$ تمثل التكرارات المتجمعة الصاعدة المقابلة للحدود العليا لفئات التوزيع .

وليكن $\frac{\sum f_i}{2}$ يمثل ترتيب الوسيط لهذا التوزيع ، فاذا كان :

$$F_{k-1} \le \frac{\sum f_i}{2} \le F_k$$

عندئذ يتم حساب قيمة الوسيط وفق الصيغة الاتية :

$$Me = L_k + \frac{h_k}{f_k}\left(\frac{\sum f_i}{2} - F_{k-1}\right)$$

حيث ان :

فئة الوسيط هي الفئة التي تسلسلها (K) .

L_k : الحد الادنى لفئة الوسيط

f_k : تكرار فئة الوسيط

h_k : طول فئة الوسيط

F_{k-1} : التكرار المتجمع الصاعد السابق لفئة الوسيط

مثال : استخرج الوسيط لجدول التوزيع التكراري الاتي :

الفئات	100-	120	140-	160-	180-	200-	220-<240
التكرار (f_i)	3	7	14	20	18	12	6

الحل: نجد اولا/ التكرار المتجمع الصاعد .

الفئات	التكرار (f_i)	الحدود العليا للفئات	التكرار المتجمع الصاعد F
100-	3	120	3
120-	7	140	10
140-	14	160	24
160-	20	180	44
180-	18	200	62
200-	12	220	74
220-<240	6	240	80

ثانيا/ نحسب ترتيب الوسيط = $\frac{\sum f_i}{2} = \frac{80}{2} = 40$

ثالثا/ نقارن ترتيب الوسيط مع التكرار المتجمع الصاعد فنلاحظ ان $24 < 40 < 44$ وعليه فان فئة الوسيط هي الفئة الرابعة من التوزيع اي الفئة (160-180) وبذلك فان :

$$L_4 = 160\ ,\ h_4 = 20\ ,\ f_4 = 20\ ,\quad F_{4-1} = 24$$

رابعا/ لإيجاد الوسيط نطبق الصيغة :

$$Me = L_k + \frac{h_k}{f_k}\left(\frac{\sum f_i}{2} - F_{k-1}\right)$$

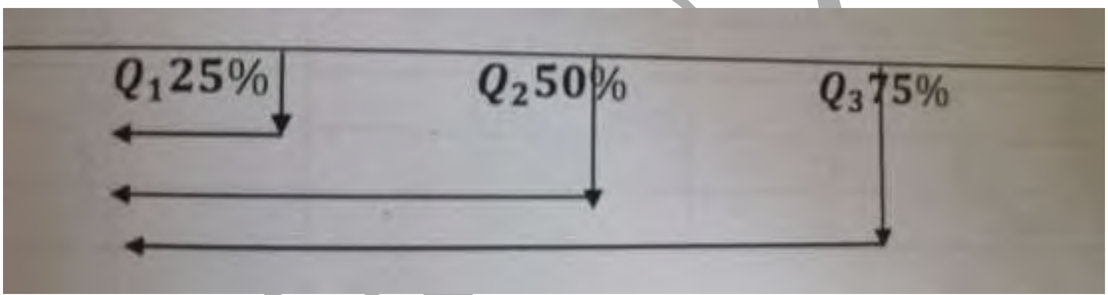
$$Me = 160 + \frac{20}{20}(40 - 24) = 160 + 16 = 176$$

(٥-١٣) المقاييس التجزيئية *Partition measures*

هي المقاييس التي تقسم مجموعة القيم الى عدد من الاجزاء المتساوية من حيث عدد القيم ، هذه المقاييس على انواع كثيرة اهمها ما يأتي :

(٥-١٣-١) الرُبيعات *The Quartiles*

عند قسمة مجموعة من القيم المرتبة على نحو تصاعدي او تنازلي الى اربعة اجزاء متساوية فان القيم الثلاث التي تقسم هذه المجموعة تسمى وعلى التوالي : الرُبيع الاول Q_1 الذي يقسم المجموعة الى قسمين بحيث تسبقه *25%* من القيم وتليه *75%* منها ، الرُبيع الثاني Q_2 الذي يقسم المجموعة الى قسمين بحيث تسبقه *50%* من القيم وتليه *50%* منها (وهذا يعني ان $Q_2 = Me$ ، الوسيط يساوي الرُبيع الثاني) ، الرُبيع الثالث Q_3 الذي يقسم المجموعة الى قسمين بحيث تسبقه *75%* وتليه *25%* منها ، وكما موضح في الشكل الاتي :



i. ايجاد قيمة الرُبيعات لبيانات غير المبوبة :

لنفرض ان $X_1, X_2, \dots, X_n$ تمثل قياسات عينة من المفردات قوامها *n* و بعد ترتيبها تصاعديا (او تنازليا) ، عندئذ فان :

Q_1 : تمثل القيمة التي ترتيبها $= \frac{n}{4}$

Q_2 : تمثل القيمة التي ترتيبها $= \frac{n}{2}$

Q_3 : تمثل القيمة التي ترتيبها $= \frac{3n}{4}$

مثال : للبيانات الاتية جد قيم الرُبيعات Quartiles .

2, 7, 3, 5, 8, 6, 10, 12, 9, 11, 5, 1, 6, 13, 16, 14

الحل : اولا/ نرتب القيم تصاعديا 1, 2, 3,5,5,6,6,7,8,9,10,11,12,13,14, 16

ثانيا/ نجد ترتيب الرُبيعات لتحديد قيم الرُبيعات

ترتيب الرُبيع الاول $= \frac{n}{4} = \frac{16}{4} = 4$

اذن قيمة الرُبيع الاول هي القيمة التي ترتيبها (*4*)

$\therefore Q_1 = 5$

ترتيب الرُبيع الثاني $= \frac{n}{2} = \frac{16}{2} = 8$

اذن قيمة الرُبيع الاول هي القيمة التي ترتيبها (*8*)

$\therefore Q_2 = 7$

ترتيب الرُبيع الثاني $= \frac{3n}{4} = \frac{3*16}{4} = 12$

اذن قيمة الرُبيع الثالث هي القيمة التي ترتيبها (*12*)

$\therefore Q_3 = 11$

ملاحظة: (اذا كان عدد القيم فردي فالربيع الاول و الثالث تكون بين قيمتين فنأخذ المعدل للقيمتين، بينما الربيع الثاني هو نفسه الوسيط)

ii. ايجاد الرُبيعات لبيانات مبوبة :

اولاً: في حالة المتغيرات المتقطعة :

افرض وجود توزيع تكراري لبيانات متغير عشوائي متقطع عدد فئاته *m* وان $f_1, f_2, \dots, f_m$ تمثل التكرارات المقابلة لفئات هذا التوزيع وان $F_1, F_2, \dots, F_m$ يمثل التكرار المتجمع الصاعد المقابل للحدود العليا لفئات التوزيع ، وان

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

عندئذ اذا كانت $1 \le i \le k \le j \le m$ تمثل اعداد صحيحة موجبة ، نقارن ترتيب الرُبيعات مع التكرار المتجمع الصاعد للتوزيع ، فاذا كان :

$F_{i-1} \le \frac{n}{4} \le F_i$ عندئذ فان قيمة الرُبيع الاول Q_1 يمثل مركز الفئة i

$F_{k-1} \le \frac{n}{2} \le F_k$ عندئذ فان قيمة الرُبيع الثاني Q_2 يمثل مركز الفئة k

$F_{j-1} \le \frac{3n}{4} \le F_j$ عندئذ فان قيمة الرُبيع الثالث Q_3 يمثل مركز الفئة j

مثال: الاتي توزيع تكراري لعينة من الاسر قوامها *124* اسرة حسب عدد افراد الاسرة ، يطلب حساب قيم الرُبيعات .

الفئات عدد الافراد	2-4	5-7	8-10	11-13	14-16	17-19	20-22
عدد الاسر f_i	4	15	22	36	28	14	5

الحل : اولا/ نجد التكرار المتجمع الصاعد للتوزيع اعلاه

الفئات عدد الافراد	2-4	5-7	8-10	11-13	14-16	17-19	20-22
عدد الاسر f_i	4	15	22	36	28	14	5
الحدود العليا للفئات	4	7	10	13	16	19	22
التكرار المتجمع الصاعد F_i	4	19	41	77	105	119	124

ثانيا/ نجد ترتيب كل رُبيع وتحديد فئته لإيجاد قيم الرُبيعات :

ترتيب الرُبيع الاول $Q_1 = \frac{124}{4} = 31$ $\qquad \because 19 < 31 < 41$

$\therefore$ Q_1 يمثل مركز الفئة الثالثة $(8-10) = 9$ افراد

ترتيب الرُبيع الثاني $Q_2 = \frac{124}{2} = 62$ $\qquad \because 41 < 62 < 77$

$\therefore$ Q_2 يمثل مركز الفئة الرابعة $(11-13) = 12$ فرد

ترتيب الرُبيع الثالث $Q_3 = \frac{3*124}{4} = 93$ $\qquad \because 77 < 93 < 105$

$\therefore$ Q_3 يمثل مركز الفئة الخامسة $(14-16) = 15$ فرد

ثانياً: في حالة المتغيرات المستمرة

افرض ان هنالك توزيعا تكراريا لبيانات متغير عشوائي مستمر عدد فئاته m وان $f_1, f_2, \dots, f_m$ تمثل التكرارات المقابلة لفئات هذا التوزيع وان $F_1, F_2, \dots, F_m$ يمثل التكرار المتجمع المقابلة للحدود العليا لفئات هذا التوزيع ، عندئذ يمكن ايجاد قيم الرُبيعات من خلال الصيغة الاتية :

$$Q_j = L_j + \frac{h_j}{f_j}\left(\frac{n*j}{4} - F_{j-1}\right) \qquad ,j = 1,2,3$$

حيث ان :

Q_j : يمثل الربيع المطلوب ايجاد قيمته

L_j : الحد الادنى لفئة الربيع

h_j : طول فئة الربيع

f_j : تكرار فئة الربيع

$\frac{n*j}{4}$: ترتيب فئة الربيع ، علما ان $n = \sum_{i=1}^{n} f_i$

F_{j-1} : التكرار المتجمع الصاعد السابق لترتيب الربيع

مثال (١) : الاتي توزيع تكراري لإعمار تلاميذ احدى المدارس ، يطلب ايجاد قيم الربيعات .

الفئات	5-	6-	7-	8-	9-	10-	11-	12-<13
التكرار	16	120	131	145	122	115	101	22

الحل : نكون اولا/ جدول توزيع تكراري متجمع صاعد

الفئات	5-	6-	7-	8-	9-	10-	11-	12-<13
التكرار	16	120	131	145	122	115	101	22
الحدود العليا للفئات	6	7	8	9	10	11	12	13
التكرار المتجمع الصاعد	16	136	267	412	534	649	750	772

ترتيب الرُبيع الاول $Q_1 = \frac{772}{4} = 193$ $\qquad$ $\because 136 < 193 < 267$

$\therefore$ فئة الرُبيع الاول Q_1 هي الفئة الثالثة $(7-< 8)$

$$Q_j = L_j + \frac{h_j}{f_j}\left(\frac{n*j}{4} - F_{j-1}\right) \rightarrow Q_1 = 7 + \frac{1}{131}(193 - 136) = 7.457$$

ترتيب الرُبيع الثاني $Q_2 = \frac{772}{2} = 386$ $\qquad$ $\because 267 < 386 < 412$

$\therefore$ فئة الرُبيع الثاني Q_2 هي الفئة الرابعة $(8 -< 9)$

$$Q_2 = 8 + \frac{1}{145}(386 - 267) = 8.821$$

ترتيب الرُبيع الثالث $Q_3 = \frac{3*772}{4} = 579$ $\qquad$ $\because 534 < 579 < 649$

$\therefore$ فئة الرُبيع الثالث Q_3 هي الفئة السادسة $(10 -< 11)$

$$Q_3 = 10 + \frac{1}{115}(579 - 534) = 10.391$$

مثال (٢) : لجدول التوزيع التكراري الاتي ، يطلب حساب قيم الرُبيعات .

الفئات	2-< 4	4-< 6	6-< 8	8-< 10	10-< 12	12-< 14
التكرار f_i	3	5	8	11	4	9

الحل : اولا/ يتم حساب التكرار المتجمع الصاعد للتوزيع اعلاه .

الفئات	2-< 4	4-< 6	6-< 8	8-< 10	10-< 12	12-< 14
التكرار f_i	3	5	8	11	4	9
التكرار المتجمع الصاعد	3	8	16	27	31	40

ثانيا/ نجد ترتيب كل رُبيع ومن ثم نحدد فئته

ترتيب الرُبيع الاول $Q_1 = \frac{n}{4} = \frac{40}{4} = 10$ $\qquad$ $\because 8 < 10 < 16$

نستنتج بان فئة الرُبيع الاول تمثل الفئة الثالثة وهي $(6 -< 8)$ وعليه فان قيمة الرُبيع الاول هي :

$$Q_j = L_j + \frac{h_j}{f_j}\left(\frac{n*j}{4} - F_{j-1}\right) \rightarrow Q_1 = 6 + \frac{2}{8}(10 - 8) = 6.5$$

ترتيب الرُبيع الثاني $Q_2 = \frac{40}{2} = 20$ $\qquad \because 16 < 20 < 27$

نستنتج بان فئة الرُبيع الثاني تمثل الفئة الرابعة وهي (*10 >-8*) وعليه فان قيمة الرُبيع الثاني هي :

$$Q_2 = 8 + \frac{2}{11}(20 - 16) = 8.72$$

ترتيب الرُبيع الثالث $Q_3 = \frac{3*40}{4} = 30$ $\qquad \because 27 < 30 < 31$

نستنتج بان فئة الرُبيع الثالث تمثل الفئة الخامسة وهي (*12 >-10*) وعليه فان قيمة الرُبيع الثالث هي :

$$Q_3 = 10 + \frac{2}{4}(30 - 27) = 11.5$$

- التمارين -

السؤال الاول// جد الوسط الحسابي والوسط التوافقي لمجموعة البيانات الاتية :

50, 52, 60, 59, 62, 63, 65, 54, 57

السؤال الثاني // الاتي توزيع تكراري لأعمار مجموعة من الطلبة حسب المرحلة الدراسية ، يطلب حساب

- متوسط عمر الطالب في هذا التوزيع
- الوسط التوافقي للتوزيع
- ارسم المنحني التكراري للتوزيع وبين نوعه (متماثل ام غير متماثل)

الفئات العمر	6-	12-	15-	18-	24-	26-<30
عدد الطلبة	30	55	80	90	66	14

السؤال الثالث // جد الوسط التربيعي والوسط الهندسي لمجموعة البيانات الاتية :

10, 15, 18, 22, 19, 14, 11, 21, 20, 19, 13

السؤال الرابع // الاتي توزيع تكراري لعينة من الطلبة موزعين حسب فئات الدرجات التي حصلوا عليها في امتحان معين ، يطلب حساب

- الوسط التربيعي للدرجة
- الوسط الهندسي للدرجة

الفئات	10-	20-	30-	40-	50-	60-	70-	80-	90-<100
عدد الطلبة	2	3	6	8	24	20	26	14	10

السؤال الخامس // لمجاميع البيانات الاتية جد :

- المنوال
- الوسيط

2, 5, 3, 5, 25, 3.3, 2.52, 3.4, 3.29, 3.6

السؤال السادس // الاتي توزيع تكراري لعدد من الموظفين في دائرة معينة حسب فئات الراتب الاسمي ، يطلب :

- ايجاد الراتب الاسمي الشائع في هذه الدائرة
- جد الوسيط للراتب الاسمي

الفئات	400-	420-	440-	460-	480-	500-<520
العدد	14	20	36	45	26	15

السابع // الاتي توزيع تكراري لإنتاج مصنع معين بالـ(طن) من سلعة معينة لاحد الايام موزع حسب عدد المكائن العاملة في ذلك اليوم وعدد ساعات العمل المحددة لاشتغال كل ماكنة حسب مواصفات المنشأ ، يطلب حساب متوسط انتاجية الماكنة الواحدة في هذا المصنع .

فئات الانتاج (طن)	2-	4-	6-	8-	10-<12
عدد المكائن العاملة	4	5	6	3	2
ساعات العمل المقررة لعمل كل ماكنة	6	5	6	4	4

الفصل السادس // مقاييس التشتت (الاختلاف)

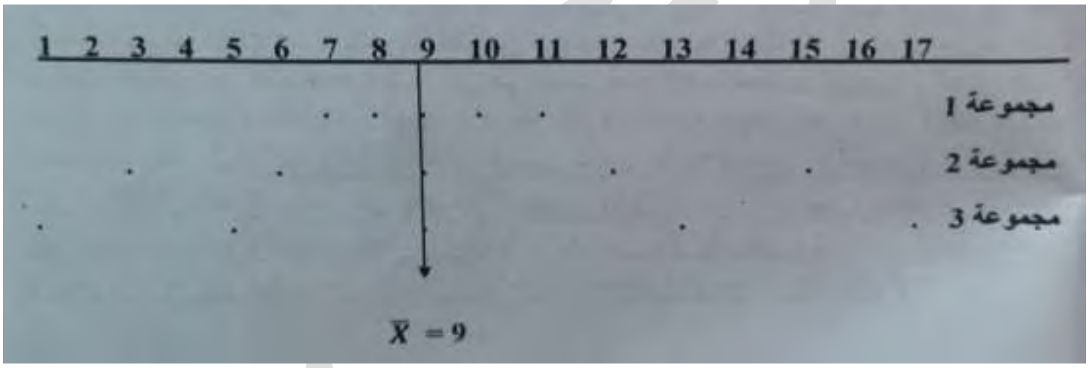
(٦-١) المقدمة

يعرف التشتت بانه تباعد او انتشار قيم مجموعة من المفردات عن بعضها البعض او عن قيمة معينة ثابته (كالوسط الحسابي مثلاً) ، ان الهدف من دراسة التشتت هو تكوين فكرة عن مدى تجانس قيم مجموعة من المفردات وهذا يعني ان دراسة التشتت هو مفيد في اجراء المقارنة بين قيم مجموعتين او اكثر من البيانات عن ظاهرة معينة ، فعلى سبيل المثال ، يلاحظ ان الوسط الحسابي لكل مجموعة من المجموعات الاتية هو $(\bar{X} = 9)$:

المجموعة الاولى : 7,8,9,10,11

المجموعة الثانية : 3,6,9,12,15

المجموعة الثالثة : 1,5,9,13,17



يلاحظ ان المجموعة الاولى اكثر تجانساً (اقل انتشاراً) من المجموعتين الثانية والثالثة ، كذلك فان المجموعة الثانية اكثر تجانساً من المجموعة الثالثة ، يلاحظ ان المقارنة بين هذه المجاميع على اساس الوسط الحسابي عديمة الفائدة .

مثال (١) : في قياس لدرجات الحرارة في ثلاثة محافظات لمدة اسبوع معين من احد الاشهر كان كالاتي :

محافظة ١: -2,-1,0,1,2,3,4

محافظة ٢: -3,-2,-1,0,2,4,7

محافظة ٣: -1,0,0,1,2,2,3

م// في اية محافظة كان الجو اكثر استقراراً .

يلاحظ ان متوسط درجة الحرارة في كل محافظة خلال هذا الاسبوع كانت درجة مئوية واحدة ، وهذا امر غير نافع في تحديد المطلوب ، الا انه يلاحظ ان الجو في المحافظة الثالثة كان اكثر استقراراً وذلك بسبب تجانس درجات الحرارة فيها اكثر من درجة تجانسها في المحافظتين الاولى والثانية .

مثال (٢): الاتي دراسة لثلاث مجموعات مختلفة من الطلاب X, Y, Z وكانت الدرجات كالاتي :

X	56 , 61 , 62 , 58 , 60
Y	50 , 60 , 66 , 54 , 70
Z	39 , 65 , 46 , 78 , 72

وبحساب الوسط الحسابي للمجموعات الثلاث نجده يساوي 60 درجة لكل منها ، ولكن عند النظر لدرجات المجموعة الاولى نجدها متقاربة اكثر من درجات المجموعة الثانية ، ودرجات المجموعة الثالثة اقل تقارباً من درجات المجموعة الثانية . اي ان المجموعات الثلاث مختلفة التجانس رغم ان الوسط الحسابي لهم متساوٍ ، وبذلك تكون مقاييس النزعة المركزية غير كافية للمقارنة بين طبيعة البيانات الاحصائية ، لذلك نشأت الحاجة الى ايجاد مقاييس تقيس درجة التجانس (تقارب) او تشتت (تباعد) مفردات البيانات عن بعضها البعض ، وتعرف هذه المقاييس بمقاييس التشتت Measures of Dispersion ، وسوف نستعرض منها كل من (المدى ، نصف المدى الربيعي ، الانحراف المتوسط ، التباين ، الانحراف المعياري ومعامل الاختلاف) .

(٦-٢) المدى Range

يعتبر المدى ابسط انواع مقاييس التشتت المطلقة ، ويعرف المدى بانه الفرق ما بين اكبر قيمة في مجموعة بيانات واصغر قيمة فيها .

i. حساب المدى في حالة البيانات الغير مبوبة

لتكن $X_1, X_2, \dots, X_n$ تمثل قياسات عينة من المفردات قوامها n ، فان المدى لهذه المجموعة هو :

$$R = X_L - X_S$$

حيث ان :

X_L : اكبر قيمة في المجموعة .

X_S : اصغر قيمة في المجموعة .

مثال : جد المدى للبيانات الاتية : 15 , 12 , 9 , 10 , 7 , 8 , 3 , 5 , 2

الحل :

$$X_L = 15 \qquad , \quad X_S = 2$$

$$R = X_L - X_S = 15 - 2 = 13$$

ii. في حالة البيانات المبوبة

لنفرض وجود توزيع تكراري، عندئذ فان المدى يُعرف في هذه الحالة بـ :

المدى = الفرق بين مركزي الفئة العليا والفئة الدنيا

أو

المدى = الحد الاعلى للفئة العليا – الحد الادنى للفئة الدنيا

$$R = U.L_{L.C} - L.L_{F.C}$$

حيث ان :

$U.L_{L.C}$: الحد الاعلى للفئة العليا (الاخيرة)(Upper limit last class)

$L.L_{F.C}$: الحد الادنى للفئة الاولى (Lower limit first class)

مثال : جد المدى (R) لدرجات مجموعة من الطلاب معطاة بالجدول الاتي :

الفئات	40-49	50-59	60-69	70-79	80-89	90-99
عدد الطلاب	2	9	15	11	2	1

الحل : يمكن ايجاد المدى بطريقتين :

الطريقة الاولى // نقوم بحساب مركز الفئة العليا = $(90 + 99)/2 = 94.5$

ومركز الفئة الدنيا = $(40 + 49)/2 = 44.5$

$$\therefore R = 94.5 - 44.5 = 50 \qquad \text{degree}$$

الطريقة الثانية // الحد الاعلى للفئة العليا = 99 ، الحد الادنى للفئة الدنيا = 40

$$\therefore R = 99 - 40 = 59 \qquad \text{degree}$$

ونلاحظ اختلاف كل من الطريقتين في حساب قيمة المدى وغالباً ما تستخدم الطريقة الاولى في ايجاد المدى .

مميزات المدى :

١. سهل الحساب .
٢. يعطي فكرة سريعة عن طبيعة البيانات ويستخدم كثيراً في مراقبة جودة الانتاج وكذلك في وصف طبيعة الاحوال الجوية .

عيوب المدى :

١. يعتمد في حسابه على قيمتين فقط من البيانات مع اهمال باقي القيم .
٢. يتأثر كثيرا بالقيم الشاذة (المتطرفة) لذلك فهو مقياس تقريبي لا يعتمد عليه .

مثال : جد المدى لمستوى الهيموكلوبين في الدم لعينة مكونة من خمسين شخصاً تم تشخيص الهيموكلوبين لهم كما موضح في الجدول التكراري الاتي :

مركز الفئة X_i	التكرار f_i	مستوى الهيموكلوبين
13.45	3	12.95 – 13.95
14.45	5	13.95 – 14.95
15.45	15	14.95 – 15.95
16.45	16	15.95 – 16.95
17.45	10	16.95 – 17.95
18.45	1	17.95 – 18.95

الحل :

مركز الفئة العليا $X_{max} = 18.45$

مركز الفئة الدنيا $X_{min} = 13.45$

$$Range = X_{max} - X_{min} = 18.45 - 13.45 = 5$$

(٦-٣) الانحراف الرُبيعي Quartile deviation

يعرف الانحراف الرُبيعي بانه متوسط الفرق بين الربيع الثالث Q_3 والربيع الاول Q_1 لمجموعة من البيانات سواء كانت غير مبوبة ام مبوبة في توزيع تكراري ، فان الانحراف الربيعي .

$$Q.D = \frac{Q_3 - Q_1}{2}$$

مثال : من جدول توزيع تكراري لإعمار عينة من تلاميذ احدى المدارس لوحظ ان قيمة الرُبيع الاول Q_1 كانت *7.433* سنة وان قيمة الرُبيع الثالث Q_3 كانت *10.391* سنة ، جد الانحراف الرُبيعي للعمر في هذا التوزيع .

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{10.391 - 7.433}{2} = 1.47 \text{ سنة}$$

ملاحظة // ان الانحراف الرُبيعي افضل من المدى كونه يستخدم *50%* من البيانات المتاحة فقط ويهمل *50%* منها ، ان هذا المقياس هو الاخر قليل الاستخدام .

(٦-٤) الانحراف المتوسط Mean deviation

i. البيانات غير المبوبة

اذا كانت $x_1, x_2, \dots, x_n$ تمثل قياسات عينة من المفردات قوامها n ، فان الانحراف المتوسط هو متوسط مجموع الانحرافات المطلقة عن وسطها الحسابي ويرمز له بالرمز *M.D* ، اي ان :

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} \quad \dots((1))$$

مثال : جد الانحراف المتوسط للبيانات الاتية :

$$X_i = 9, 8, 6, 5, 7$$

الحل : نجد اولاً الوسط الحسابي للبيانات

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{35}{5} = 7$$

ولاستخراج الانحراف المتوسط نستخرج المقاييس الموضحة في الجدول ادناه :

x_i	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$
9	2	2
8	1	1
6	-1	1
5	-2	2
7	0	0
		6

وبعدها نطبق الصيغة الاتية :

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} = \frac{6}{5} = 1.2$$

ii. حساب الانحراف المتوسط لبيانات مبوبة

لتكن $X_1, X_2, \dots, X_m$ تمثل مراكز الفئات لتوزيع تكراري عدد فئاته m وان $f_1, f_2, \dots, f_m$ تمثل التكرارات المقابلة لهذه الفئات ، فان الانحراف المتوسط يمكن حسابه وفق الصيغة الاتية :

$$M.D = \frac{\sum_{i=1}^{m} f_i |x_i - \bar{x}|}{\sum_{i=1}^{m} f_i} \quad \dots ((2))$$

مثال : الاتي توزيع تكراري لدرجات مجموعة من الطلبة لامتحان معين . يطلب حساب الانحراف المتوسط .

الفئات	0-	10-	20-	30-	40-	50-	60-	70-	80-	90-100
التكرار	2	4	8	16	25	60	42	35	18	10

الحل: لحساب الانحراف المتوسط ، نستخرج المقاييس الموضحة في الجدول ادناه ، ثم نطبق الصيغة *((2))* .

الفئات	f_i	x_i	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i\|x_i - \bar{x}\|$
0-	2	5	10	54.5	109
10-	4	15	60	44.5	178
20-	8	25	200	34.5	276
30-	16	35	560	24.5	392
40-	25	45	1125	14.5	362.5
50-	60	55	3300	4.5	270
60-	42	65	2730	5.5	231
70-	35	75	2625	15.5	542.5
80-	18	85	1530	25.5	459
90-100	10	95	950	35.5	355
المجموع	220		13090		3175

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i x_i}{\sum_{i=1}^{m} f_i} = \frac{13090}{220} = 59.5$$

$$M.D = \frac{\sum_{i=1}^{m} f_i |x_i - \bar{x}|}{\sum_{i=1}^{m} f_i} = \frac{3175}{220} = 14.4318$$

(٦-٥) الانحراف المعياري Standard deviation

ويسمى في بعض الاحيان بالانحراف القياسي ، ويعتبر هذا المقياس من افضل مقاييس التشتت لما يمتاز به من مميزات مثلى جعلته يقف في مقدمتها عند التطبيق .

ويعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي .

i. حساب الانحراف المعياري لبيانات غير مبوبة

لتكن $x_1, x_2, \ldots, x_n$ تمثل قياسات مفردات عينة قوامها n ، وليكن $\bar{x}$ يمثل الوسط الحسابي لهذه القياسات ، عندئذ وحسب التعريف اعلاه فان الانحراف المعياري (S) لهذه المجموعة هو :

$$S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \quad \ldots(((1)))$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} {x_i}^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2}{n-1}}$$

وحيث ان $\sum_{i=1}^{n} x_i = n\bar{x}$ عندئذ فان الانحراف المعياري يمكن ان يحسب وفق الصيغة الاتية :

$$S = \sqrt{\frac{\sum_{i=1}^{n} {x_i}^2 - n\bar{x}^2}{n-1}} \quad \ldots(((2)))$$

او يحسب وفق الصيغة الاتية :

$$S = \sqrt{\frac{\sum_{i=1}^{n} {x_i}^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}} \quad \ldots(((3)))$$

مثال : جد الانحراف المعياري لمجموعة الاوزان (كغم) الاتية :

7.1 , 2.5 , 2.5 , 5.4 , 8.3

الحل : البيانات اعلاه غير مبوبة و لإيجاد الانحراف المعياري نحسب اولا الوسط الحسابي وبعدها نطرح الوسط الحسابي من كل قيمة ، والجدول الاتي يوضح المقاييس المطلوب حسابها:

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	${x_i}^2$
7.1	1.94	3.7636	50.41
2.5	-2.66	7.0756	6.25
2.5	-2.66	7.0756	6.25
5.4	0.24	0.0576	29.16
8.3	3.14	9.8596	68.89
25.8	0.00	27.832	160.96

حيث ان :

$$n = 5 \qquad \sum x_i = 25.8 \qquad \sum_{i=1}^{n}(x_i - \bar{x})^2 = 27.832$$

وان متوسط العينة (الوسط الحسابي للعينة) هو :

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{25.8}{5} = 5.16$$

ولايجاد الانحراف المعياري نستخدم الصيغة (((1))) :

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{27.832}{5-1}} = 2.6378 \quad \text{كيلوغرام}$$

تمرين // جد الانحراف المعياري للمثال اعلاه وفق الصيغة (((2))) والصيغة (((3))) .

مثال : البيانات الاتية تمثل اوزان عينة من الطلبة قوامها عشرة طلاب ، يطلب حساب قيمة الانحراف المعياري .

$$x_i = 56\,,62\,,69\,,71\,,68\,,65\,,63\,,72\,,68\,,56$$

الحل: نحسب اولا الوسط الحسابي لأوزان الطلبة

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 65 \text{ كغم}$$

x_i:	56	62	69	71	68	65	63	72	68	65
$x_i - \bar{x}$:	-9	-3	4	6	3	0	-2	7	3	-9
$(x_i - \bar{x})^2$:	81	9	49	9	0	4	36	16	9	81

بعد ان يتم طرح متوسط اوزان الطلبة من كل وزن طالب ثم تربيع القيمة يتم حساب الانحراف المعياري وفق الصيغة الاتية :

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{294}{10-1}} = 5.2599$$

ii. حساب الانحراف المعياري لبيانات مبوبة

لتكن $X_1, X_2, \dots, X_m$ تمثل مراكز الفئات لتوزيع تكراري عدد فئاته m وان $f_1, f_2, \dots, f_m$ تمثل التكرارات المقابلة لهذه الفئات ، عندئذ وحسب التعريف لهذا يمكن حساب قيمته وفق الصيغة الاتية :

$$S = \sqrt{\frac{\sum_{i=1}^{m} f_i (x_i - \bar{x})^2}{n-1}} \qquad \dots [1]$$

حيث ان : $\bar{x} = \frac{\sum_{i=1}^{m} f_i\, x_i}{\sum_{i=1}^{m} f_i}$, $n = \sum_{i=1}^{m} f_i$

$$S = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{m} f_i\, x_i^2 - \frac{\left(\sum_{i=1}^{m} f_i x_i\right)^2}{n}\right)} \qquad \dots \quad [2]$$

علما ان : $n = \sum_{i=1}^{m} f_i$

مثال : جد الانحراف المعياري لدرجات الطلاب المعطاة (القيم المظللة) بالجدول الاتي :

الحل : اولا يتم حساب مركز الفئة ثم استخراج الوسط الحسابي ومن ثم حساب القيم الموضحة في الجدول ادناه لإيجاد الانحراف المعياري :

الفئات (درجات الطلاب)	40-49	50-59	60-69	70-79	80-89	90-99
عدد الطلاب f_i	2	9	15	11	2	1
مركز الفئة x_i	44.5	54.5	64.5	74.5	84.5	94.5
$f_i x_i$	89	490.5	967.5	819.5	169	94.5
$x_i - \bar{x}$	-21.25	-11.25	-1.25	8.75	18.75	28.75
$(x_i - \bar{x})^2$	451.5625	126.5625	1.5625	76.56	351.56	826.56
$f_i(x_i - \bar{x})^2$	903.13	1139.06	23.44	842.16	703.12	826.56

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i\, x_i}{\sum_{i=1}^{m} f_i} = \frac{2630}{40} = 65.75$$

$$S = \sqrt{\frac{\sum_{i=1}^{m} f_i (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{4437.5}{40-1}} = 10.667$$

تمرين // جد الانحراف المعياري للمثال اعلاه باستخدام الصيغة [2] .

مثال : الاتي توزيع تكراري لدرجات مجموعة من الطلبة في امتحان معين { القيم المظللة في الجدول ادناه } . يطلب حساب الانحراف المعياري (S) Standard deviation .

الحل: اولا يتم حساب مركز الفئة ثم حساب القيم الموضحة في الجدول ادناه ، وباستخدام الصيغة [2] يتم ايجاد الانحراف المعياري .

الفئات	f_i	مركز الفئة x_i	$f_i x_i$	$f_i x_i^2$
0-	2	5	10	50
10-	4	15	60	900
20-	8	25	200	5000
30-	16	35	560	19600
40-	25	45	1125	50625
50-	60	55	3300	181500
60-	42	65	2730	177450
70-	35	75	2625	196875
80-	18	85	1530	130050
90-100	10	95	950	90250
المجموع	220		13090	852300

$$S = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{m} f_i\, x_i^2 - \frac{\left(\sum_{i=1}^{m} f_i x_i\right)^2}{n}\right)}$$

$$S = \sqrt{\frac{1}{220-1}\left(852300 - \frac{(13090)^2}{220}\right)} = 18.313$$

(٦-٦) خصائص الانحراف المعياري

a. ان $S \geq 0$ هذا يعني ان قيمة الانحراف المعياري هي دائماً قيمة موجبة وتساوي صفر في حالة خاصة .

b. اذا كانت a كمية ثابته حقيقية وان $y_i = ax_i$ عندئذ فان $S_y = |a|S_x$

c. اذا كانت a كمية ثابته حقيقية وان $y_i = a \mp x_i$ عندئذ فان $S_y = S_x$

(٦-٧) التباين ***The Variance***

i. حساب التباين لبيانات غير مبوبة :

لتكن $x_1, x_2, \dots, x_n$ تمثل قياسات عينة من المفردات قوامها n ، وليكن $\bar{X}$ يمثل الوسط الحسابي لهذه القياسات عندئذ يعرف التباين بانه متوسط مجموع مربعات انحرافات قيم المتغير العشوائي x_i عن وسطها الحسابي .

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \qquad \dots(1)$$

هذا يعني ان التباين ما هو الا مربع الانحراف المعياري لتلك المجموعة من القيم وعليه فان الرمز S^2 يشير الى تباين قيم العينة .

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} {x_i}^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right) \qquad \dots(2)$$

مثال : احسب التباين لأعمار مجموعة من الطلاب في المرحلة الابتدائية :

$$8\,,9\,,7\,,6\,,5$$

الحل :

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{8+9+7+6+5}{5}=\frac{35}{5}=7$$

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

$$S^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)=\frac{1}{5-1}(10)=2.5$$

ii. حساب التباين لبيانات المبوبة

اذا كان لدينا عدد m من الفئات ذات المراكز $X_1,X_2,\dots,X_m$ ولها التكرارات $f_1,f_2,\dots,f_m$ على الترتيب فان التباين يحسب وفق الصيغ الاتية :

$$S^2=\frac{1}{n-1}\left(\sum_{i=1}^{m}f_i(x_i-\bar{x})^2\right)\quad \dots(3)$$

$$S^2=\frac{1}{n-1}\left(\sum_{i=1}^{m}f_i{x_i}^2-\frac{\left(\sum_{i=1}^{m}f_ix_i\right)^2}{n}\right)\qquad \dots(4)$$

مثال : جد التباين S^2 لدرجات مجموعة من الطلاب معطاة بالجدول الاتي :

90-99	80-89	70-79	60-69	50-59	40-49	*الفئات*
1	2	11	15	9	2	*عدد الطلاب f*

الحل :

الفئات	40-49	50-59	60-69	70-79	80-89	90-99
عدد الطلاب f	2	9	15	11	2	1
x_i	44.5	54.5	64.5	74.5	84.5	94.5
$f_i x_i$	89	490.5	967.5	819.5	169	94.5
$x_i-\bar{x}$	-21.25	-11.25	-1.25	8.75	18.75	28.75
$(x_i-\bar{x})^2$	451.5625	126.5625	1.5625	76.56	351.56	826.56
$f_i(x_i-\bar{x})^2$	903.13	1139.06	23.44	842.16	703.12	826.56

علما ان الوسط الحسابي يساوي :

$$\bar{x}=\frac{\sum_{i=1}^{m} f_i x_i}{n}=\frac{2630}{40}=65.75$$

$$S^2=\frac{1}{n-1}\left(\sum_{i=1}^{m} f_i(x_i-\bar{x})^2\right)=\frac{1}{40-1}(4437.5)=113.78$$

تمرين : حل المثال اعلاه حسب الصيغة (*4*) .

(٨-٦) معامل الاختلاف Coefficient of Variation(C.V)

افرض ان $\bar{X}$ يمثل الوسط الحسابي لمجموعة قيم وان S يمثل الانحراف المعياري لها ، عندئذ يعرف معامل الاختلاف $C.V$ على النحو الاتي :

$$C.V=\frac{S}{\bar{X}}\cdot 100$$

معامل الاختلاف: هو معامل نسبي يستخدم للمقارنة بين تشتت بيانات ظاهرتين أو اكثر مختلفتين في وحدات القياس او متفقتين، او مختلفتين في القيمة المتوسطة لهما.

ان معامل الاختلاف يعتبر بحق افضل انواع المعاملات التشتت الانفة الذكر كونه يعتمد على افضل مقياس نزعة مركزية وافضل مقياس تشتت ، ان هذا المعامل يوضح نسبة كل وحدة من وحدات الوسط الحسابي من الانحراف المعياري ، ويعتبر معامل الاختلاف من مقاييس التشتت النسبي . وعليه وعند اجراء مقارنه بين قيم مجموعتين تتم مقارنة معامل الاختلاف المجموعة الاولى مع معامل اختلاف المجموعة الثانية وعندئذ يقال عن المجموعة بانها اكثر تجانساً اذا كان معامل اختلافها اقل من المجموعة الاخرى .

مثال : كان متوسط درجات طلبة لأحدى المراحل في امتحان الرياضيات *69* درجة بانحراف معياري قدره *19.3* في حين كان متوسط درجاتهم في امتحان الاحصاء *75* درجة بانحراف معياري قدره *25.5* ، في اي من الامتحانين كان مستوى اداء الطلبة اكثر تقارباً ؟

الحل:

نجد معامل الاختلاف لكل امتحان ...

$$C.V(stat.) = \frac{S}{\overline{\overline{X}}} \cdot 100 = \frac{25.5}{75} \cdot 100 = 34\%$$

$$C.V(math) = \frac{S}{\overline{\overline{X}}} \cdot 100 = \frac{19.3}{69} \cdot 100 = 28\%$$

وحيث ان معامل الاختلاف في امتحان الرياضيات اقل من معامل اختلاف امتحان الاحصاء عليه فان مستوى الطلبة في امتحان الرياضيات كان اكثر تقارباً (تجانساً) .

- التمارين -

Q_1// احسب المدى والانحراف المتوسط والانحراف المعياري والتباين للبيانات الاتية :

6 , 3 , 5 , 5 , 9 , 4 , 6 , 7 , 1 , 2 , 4 , 8

Q_2 // فيما يلي اوزان *50* طالب من طلاب جامعة بغداد ، جد :

- مدى اوزان الطلبة
- الانحراف المتوسط والانحراف المعياري والتباين لأوزان الطلبة

73-75	70-72	67-69	64-66	61-63	58-60	*فئات الوزن*
4	8	15	14	7	2	*عدد الطلاب*

Q_3 // الجدول ادناه يتضمن بيانات احدى الدراسات التي طبقت على خمسة اشخاص لقياس الوزن (بالكيلو غرام) والطول (بالسنتمتر) ، اي البيانات اكثر تشتتاً (اقل تجانساً) بيانات الاوزان ام بيانات الاطوال ؟

تسلسل الاشخاص	*1*	*2*	*3*	*4*	*5*
الوزن *Kg*	*69*	*59*	*65*	*67*	*65*
الطول *cm*	*164*	*162*	*155*	*165*	*158*

Q_4 // الاتي توزيع تكراري لأعمار عينة من الافراد ، جد الانحراف المعياري والتباين للعمر ، ثم جد قيمة تقريبية للانحراف المتوسط

80-90	70-	60-	50-	40-	30-	20-	*الفئات*
2	51	120	133	120	61	13	*التكرار*

مقدمة:

تقسم العينات الى نوعين ، الاول هو العينات الاحتمالية Probabilistic Samples وهو النوع الذي يعتمد على نظرية الاحتمالات ، اي ان عملية الاختيار تجري بحيث يكون لكل مفردة في المجتمع احتمال معلوم ، وان احتمال ظهور تلك المفردة في العينة غير مساوي للصفر ، و هذا الاحتمال يمكن ان يكون متساويا لكافة المفردات كما يحصل ذلك في العينة العشوائية البسيطة ، التي هي احد انواع العينات الاحتمالية ، او يكون غير متساوي كما في الانواع الاخرى من العينات الاحتمالية ، كالطبقية ، المنتظمة ، العنقودية .... الخ .

اما النوع الثاني فهي العينات غير الاحتمالية ، و في هذا النوع يتم اختيار مفردات العينة عمديا و من امثلتها العينة العمدية و العينة الحصصية و عينات المجتمعات المتحركة . ان الشائع في الاستخدام العملي هي العينات الاحتمالية نظرا لما تتمتع به من دقة و اعتمادية و عليه فأننا سنركز على هذا النوع .

٧-١ العينة معناها و أسباب اختيارها Sample and Reasons for Sampling

اننا نعيش في عصر يصعب الحكم فيه على صحة الاشياء بدون الادلة الموضوعية و البحث الواسع ، عصر مليء بالمشاكل العلمية المختلفة و الظواهر المتعددة . و لدراسة اية مشكلة علمية نحتاج الى جمع كل ما يتعلق بتلك المشكلة من معلومات . و تسمى مجموعة العناصر المتعلقة بتلك المشكلة المجتمع الاحصائي . من هنا تبدأ دراسة احصائية بالبيانات الخام المتوفرة عن الدراسة و التي يتم جمعها بأحدى الطرق الاتية :

١- طريقة المسح الشامل : وفيها تجمع البيانات من جميع افراد المجتمع الاحصائي ، فمثلا اذا اردنا التعرف على مستوى الطلاب في جامعة لاقسام الرياضيات نقوم بجمع علامات جميع الطلاب في كليات الجامعة التي تضم هذا القسم .
٢- طريقة العينات : هناك عدة حالات يتعذر فيها المسح الشامل و عندها نلجأ الى دراسة جزء من المجتمع الاحصائي يسمى العينة Sample ، ويعرف حجم العينة بأنه عدد عناصرها .

٧-٢ انواع العينات الأحتمالية :

١- العينة العشوائية البسيطة Simple Random Sample :

ان مسوح العينة تتعامل مع عينة مختارة من مجتمع فيه عدد محدد من الوحدات المختلفة و لنقل N و اذا كان بالامكان التمييز بين واحدة و اخرى من هذه الوحدات ، فأن عدد العينات المختلفة من حجم (n) التي يمكن سحبها من المجتمع الذي عدد وحداته (N) يمكن ايجادها باستخدام قاعدة التوافيق الاتية :

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

فلو كان لدينا مجتمع يتكون من الوحدات (A,B,C,D) اي ان N=4 ، فأن هناك ست عينات مختلفة من حجم n =4 ممكن سحبها كالاتي :

$$C_2^4 = \frac{4!}{2!(4-2)!} = \frac{4\times3\times2\times1}{2\times1(2\times1)} = 6$$

و اذا اردنا ترجمة ذلك العدد الى صيغة الحروف فأ ن العينات الممكنة و البالغ عددها (6) هي كالاتي : AB, AC,AD,BC,BD, CD.

و يمكن ملاحظة ان (نفس المفردة) لا تظهر مرتين في عينة واحدة من العينات الممكنة.

ان العينة العشوائية البسيطة هي طريقة لاختيار (n) من الوحدات من مجتمع حجمه (N) من الوحدات بحيث يكون لكل مفردة من المفردات المختارة نفس الاحتمال في الظهور في العينة .

تعتبر هذه الطريقة من اسهل طرق المعاينة الاحتمالية ، ولكن ليس اكثرها استعمالا بل انها اقل انواع العينات استخداما في الحياة العملية ، بالرغم من سهولتها و دقتها عندما يكون المجتمع متجانس و في حالة توفر التجانس فان الامر يغنينا عن استعمال انواع اخرى من العينات.

تصنف اساليب المعاينة من حيث طريقة التعامل مع المفردات التي تسحب من العينة ، الى نوعين :

أ- المعاينة بدون اعادة : اي ان المفردة التي تختار ضمن العينة في اي مرحلة من مراحلها، لا تعاد الى المجتمع كي لا تعطى فرصة جديدة في الظهور ثانية في العينة .

ب- المعاينة مع الاعادة : اي ان المفردة التي تختار ضمن العينة تعطى فرصة اخرى للظهور ثانية في العينة ضمن مراحلها المتعاقبة ، و هذا النوع لا يستخدم الا لظروف الخاصة و حالات خاصة .

٢- العينة العشوائية الطبقية Stratified Random Sample:

اشرنا سابقا الى ان استخدام العينة العشوائية البسيطة يعتمد على حقيقة ان المجتمع متجانس، اي ان مفرداته متجانسة من حيث الخاصية او (الخواص) التي يراد معاينتها، و عند عدم توفر التجانس يتم استخدام الاسلوب الطبقي (اي تقسيم المجتمع الى طبقات (اجزاء)) بحيث تكون مفردات كل جزء او طبقة متجانسة فيما بينها و مختلفة عن الطبقات الاخرى، و يتم التعامل مع كل طبقة و كأنها مجتمع مستقل تسحب منه عينة عشوائية ذات حجم معين .

وفقا لما سبق اعلاه فان حجم المجتمع البالغ عدد مفرداته (N) سيمثل مجموع عدد المفردات في الطبقات كافة ، و ان عدد المفردات في كل طبقة سنرمز له بالرمز (N_h) .

حيث ان h=1,2,…,L ، اي ان عدد الطبقات في المجتمع = L ،

$$N = \sum_{h=1}^{L} N_h = N_1 + N_2 + \cdots + N_L$$

و ان $N_1, N_2, \dots, N_L$ تسمى طبقات (strata) و ان الجزء الواحد يسمى طبقة (stratum)، اما حجم العينة التي نسحبها هو (n)، اي ان :

n=عدد المفردات في العينة.

حجم العينة الكلي $n=n_1+n_2+\dots+n_L$

٣- العينة المنتظمة Systematic Sample :

تشير تسمية هذا النوع من العينلت الى انه يتبع اسلوبا منتظما لاختيار وحدات المجتمع. فأذا كان لدينا مجتمع يتكون من (N) من الوحدات و رغبنا ان نأخذ عينة حجمها (n) من هذا المجتمع فان الاسلوب المنتظم يكون بالشكل التالي :

أ- نستخرج قيمة (k) و التي تساوي ($\frac{N}{n}$) و بذلك نقسم المجتمع الى مجاميع عددها (k) و كل مجموعة تضم (n) من المفردات .

ب- نختار وحدة واحدة بشكل عشوائي من اول مجموعة .

ت- بعد ان يتم تحديد الوحدة الاولى يمكن ان تؤخذ الوحدات المتبقية ، حيث يكون تسلسل الوحدة الثانية بعد (k) من الوحدات (ابتداءا من الوحدة المختارة من المجموعة الاولى) فاذا كان لدينا مجتمع حجمه (N=200) و اردنا اختيار عينة حجمها (n=10) فان قيمة k هي :

$$k = \frac{N}{n} \qquad \rightarrow \qquad k = \frac{200}{10} = 20$$

نقوم باختيار وحدة واحدة عشوائيا من اول (عشرين) وحدة، و لنقل انها كانت الوحدة رقم (6) بهذا فان الوحدة الثانية ستكون (26) و الثالثة (46) ثم (66) , (86), (106) ، (126)، (146)، (166)، (186) و بهذا تكون العينة المطلوبة التي حجمها n=10 قد تحققت.

تستخدم العينة المنتظمة في المجتمعات الكبيرة و في المشاهدات المشاهدات الميدانية، علما ان العينة المنتظمة اكثر تمثيلا من العينة العشوائية البسيطة حيث يجري تقسيم المجتمع الى طبقات عددها (k) و كل طبقة حجمها (n) يتم اختيار واحدة منها، و كذلك ان العينة المنتظمة تتوزع على المجتمع توزيعا شاملا اكثر مما يحدث في العينة العشوائية البسيطة.

مثال: نفرض لدينا مجتمع حجمه (N=20) و رغبنا بسحب عينة حجمها (n=5) و عليه فان :

$$k = \frac{N}{n} \qquad \rightarrow \qquad k = \frac{20}{5} = 4$$

اي ان عدد العينات الممكنة هي (4) و كما ياتي :

العينة الاولى	العينة الثانية	العينة الثالثة	العينة الرابعة
1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20

تمارين :

١- مجتمع مكون من (50) طالب ، ما هو عدد العينات العشوائية البسيطة الممكن سحبها من هذا المجتمع بحجم (10) طلاب .

٢- في احدى المناطق يوجد (40) مخزن، ما هو عدد العينات العشوائية المنتظمة الممكن سحبها من هذا المجتمع ، علما ان حجم العينة المراد سحبها هة (8) مخازن.

جامعة بغداد

كلية التربية للعلوم الصرفة / ابن الهيثم

منهج بحث

قسم الرياضيات

المرحلة الثانية

نظرية الاحتمال Probability Theory

1-8 المقدمة :

تعد نظرية الاحتمال، أو قوانين المصادفة Law of chances ، من النظريات الرياضية المهمة التي تتخصص في دراسة الحوادث والمتغيرات والظواهر التي تتميز بعدم تأكد حدوثها.

ولنظرية الاحتمال تاريخ حافل يبدأ من منتصف القرن السابع عشر عندما أقترح النبيل الفرنسي Chevalier De Mere مسألة معينة في لعب الورق على العالم الفرنسي باسكال Blaise Pascal ومن خلال مراسلة باسكال الرياضي الفرنسي فرمات Pierre De Fermat وضعت الاسس الاولى لنظرية الاحتمال.

في حياتنا اليومية نستعمل بكثرة كلمة إحتمال، فيقول طالب إحتمال ان انجح في الامتحان % 60 ، ويقول شخص احتمال سقوط الامطار غدا 20% ، ويقول الطبيب احتمال ان تنجح العملية الجراحية 97% ، ان الحوادث التي تمر علينا يوميا يمكننا ان نصنفها الى ثلاثة اصناف:

1- حوادث مؤكدة الحدوث Certain events : وهي حوادث نجزم بحدوثها جزما مؤكدا، والسبب ان هذه الحوادث قد حدثت بالتأكيد في جميع الفترات الزمنية ، مثل حادثة " موت الانسان " هي حادثة مؤكدة.

2- حوادث مستحيلة الحدوث Impossible events : وهي حوادث نجزم بعدم حدوثها ولانتوقع ان تحدث على الاطلاق، مثل حادثة " شروق الشمس ليلا في مدينة بغداد " هي حادثة مستحيلة، وكذلك حادثة " ان تعيش سمكة خارج الماء" هي حادثة مستحيلة.

3- حوادث محتملة الحدوث (غير مؤكدة) Probable (uncertain) events : وهي حوادث لايمكننا ان نجزم جزما مؤكدا بحدوثها أو عدم حدوثها، أي انها حوادث غير مؤكدة الحدوث ولامستحيلة الحدوث، مثال ذلك " سقوط الامطار يوم غد " " ارتفاع اسعار النفط بعد خمس سنوات من الآن ".
الاسبوع القادم

ولما كنا نعرف سلفاً نتائج الحوادث المؤكدة والمستحيلة فاننا لانهتم كثيرا بدراسة مثل تلك الحوادث، اما الحوادث المحتملة فاننا لايمكننا معرفة نتائجها معرفة قاطعة قبل حدوثها، لهذا فاننا في نظرية الاحتمال نركز اهتماننا على دراسة الحوادث المحتملة لكي نتمكن (بشكل احتمالي) من التكهن بنتائجها ولمعرفة القوانين الاحتماليه المتحكمة بحدوثها. وبعبارة اخرى فان نظرية الاحتمال هي نظرية رياضية تتركز على دراسة الحوادث المحتملة وقياس مقدار احتمال حدوثها.

8 – 2 التجربة العشوائيـــة Random Experiment :

يمكننا تعريف التجربة بانها حالة خاصة يمكننا من خلالها ملاحظة نتيجة ما، هنالك نوعان أساسيان من التجارب : محددة Deterministic وغير محددة يطلق عليها غالبا عشوائية Stochastic . وتكون التجربة محددة اذا كانت نتائجها ثابته ومستقرة ولاتخضع للمصادفة، وإذا اعيدت التجربة أكثر من مرة واحدة وبنفس الظروف في كل مرة فانها تعطي دائما نفس النتيجة، (مثال ذلك اذا كان لديك ميزان حساس وطلب منك وزن كتله معينه، ولو أعيدت التجربة (وزن الكتله) وتحت نفس الظروف ولأكثر من مرة واحدة فان النتيجة ستكون نفسها).

أما في التجارب العشوائية فان النتيجة تخضع دائماً لتأثيرات عشوائية(ولهذا سميت تجارب عشوائية)، فاذا أعيدت التجربة العشوائية أكثر من مرة واحدة فانها في كل مرة سوف تعطي نتائج مختلفة ولو بقدر بسيط عن المرات السابقة (مثال ذلك إذا رميت قطعة نقود أكثر من مرة فانك سوف لاتحصل دائما على نفس الوجه).

ان نظرية الاحتمال تتخصص بدراسة التجارب العشوائية، وباختصار فان التجربة العشوائية هي عملية تحقيق شروط معينه بحيث يمكنها اعطاء ناتج واحد من بين أكثر من نتيجة ممكنة.

<u>مثال (1) :</u> ان عملية رمي قطعة نقود هي تجربة عشوائية وان نتائج هذه التجربة هي أما ظهور الصورة H أو الكتابة T (Head or Teal) .

Tail

<u>مثال (2) :</u> ان عملية رمي قطعة زهر النرد هي تجربة عشوائية وان النتائج الممكنة لهذه التجربة ان يكون عدد النقاط على الوجه العلوي لقطعة النرد (6 , 5 , 4 , 3 , 2 , 1) .

<u>مثال (3):</u>ان فحص فصيلة الدم هي تجربة عشوائية نتائجها أحد الأصناف (A , B , AB , O)

<u>8 – 3 فضاء العينه Sample Space :</u> ان فضاء العينه S المتعلق بالتجربة العشوائية هو المجموعة الجامعة المؤلفة من جميع النتائج الممكننة للتجربة.

- في المثال (1) يتكون فضاء العينه من العنصرين H و T أي ان : S={ H ,T }
- في المثال (2) ان فضاء العينه المتعلق بالتجربة العشوائية هو: S={ 1,2,3,4,5,6 }
- في المثال (3) ان فضاء العينه المتعلق بالتجربة هو : S= { A , B , AB , O }

8 – 4 الحادث (أو الحدث) The event :

الحادث هو نقطة أو عدة نقاط في فضاء العينه ويرمز له بـ (Ei) . والحادث قد يكون بسيطا Simple event اذا تكون من نقطة واحدة في فضاء العينه (أي حالة واحدة من الحالات التي تظهر نتيجة التجربة) أو يكون حادثا مركبا Compound event اذاشمل حالتين أو أكثر من الحالات التي تظهر نتيجة التجربة .

مثال : الحصول على (عدد زوجي) في رمية الزار (النرد) يسمى حادثا وهو يتكون من النقاط 2,4 , 6) من مجموعة نقاط فضاء العينه { 1 , 2, 3 , 4 , 5 , 6 } .

مثال : الحصول على الصورة (H) من رمي قطعة نقود معدنية مرة واحدة يسمى حادثا وهو يتكون من نقطة واحدة { H } من مجموعة نقاط فضاء العينه { H , T } .

الحوادث المتنافيه (المتناقضه) Mutually events :

يقال عن الحادثين E_1 , E_2 انهما متنافيان اذا استحال حدوثهما معا.

مثال : عند رمي قطعة نقود فمن المستحيل الحصول على (صورة وكتابة) في نفس الوقت .

الحوادث المستقلة Independent events :

الحوادث المستقلة هي الحوادث التي اذا وقع أحدها لايمنع أو يؤثر على وقوع الحوادث الاخرى.

مثال : عند رمي قطعتي نقود فالحصول على صورة في القطعة الاولى مثلا لايؤثر في نتيجة القطعة الثانية.

مثال : صندوق فيه عدد معين من الكرات البيضاء والسوداء (المتماثلة وزنا وحجما)، فعند سحب كرتين بحيث تعاد الاولى قبل سحب الثانية فان نتيجة السحبة الاولى لاتؤثر في نتيجة السحبة الثانية لذا فالحادثان مستقلان .

الحوادث غير المستقلة Non Independent events :

الحوادث غير المستقلة هي الحوادث التي اذا وقع أحدها يؤثر في وقوع الحوادث الاخرى.

مثال: حالة الصندوق وفيه الكرات البيضاء والسوداء فعند سحب كرتان على التوالي بحيث لاتعاد الكرة الاولى فان نتيجة السحبة الثانية تتأثر بنتيجة السحبة الاولى لذلا فالحادثان غير مستقلين.

الحالات الممكنة Possible Cases :

الحالات الممكنة هي جميع الحالات المختلفة التي يمكن ان تظهر في تجربة معينه.

مثال: عند رمي قطعة النقود فعدد الحالات الممكنة هنا حالتين (صورة Head أو كتابة Tail).

مثال : عند رمي زهر النرد عدد الحالات الممكنة هي 6 حالات ، أما عند رمي زهري نرد فعدد الحالات الممكنة هي $36 = 6 * 6$ حالة ، من ذلك نرى بأن الحالات الممكنة هي نفسها فضاء العينة.

الحالات المواتية Favorable Cases :

هي الحالات التي تحقق ظهور الحادث المراد دراسته وتسمى ايضا بحالات النجاح.

مثال: عند رمي زهر النرد فاذا كان الحادث هو الحصول على عدد زوجي فالحالات التي تحقق ظهور هذا الحادث هي الحصول على (2) أو (4) أو (6) فهذه الحالات الثلاثة تسمى الحالات المواتية.

الحالات المتماثلة Equally Likely Cases :

هي الحالات المتكافئة والمتساوية في امكانية حدوثها.

مثال : عند رمي قطعة النقود فان الظروف المهيأه للحصول على أي وجه (صورة أو كتابة) تكون متكافئة، فيقال بان الحالتين التي تنتج عن تجربة رمي قطعة النقود حالتان متماثلتان.

8 – 5 المجمــــوعات Sets :

المجموعة هي تجمع أي عدد من العناصر وسنمثل المجموعات بحروف لاتينية كبيرة مثل A,B,C,... وعناصر المجموعة بحروف صغيرة a,b,c,.... ، المجموعة $A = \{ a , b , c \}$ فيقال ان $a \in A$

(العنص a ينتمي الى المجموعة A) وان $k \notin A$ (العنصر k لاينتمي الى المجموعة A).

المجموعة الخالية Empty set : هي المجموعة التي لايوجد فيها عناصر مطلقا ويعبر عنها بالرمز ϕ أو $\{ \}$.

المجموعة الجزئية Subset : نقول ان Aمجموعة جزئية من B ونرمز لها $A \subset B$ اذا كان كل عنصر في A منتميا للمجموعة B .

التساوي Equality : المجموعة A تساوي المجموعة B أي $A = B$ اذا كانت A , B تحتويان على نفس العناصر، ويكون ذلك اذا تحقق الشرطان $A \subset B$ و $B \subset A$.

المجموعة الكلية Universal set : هي المجموعة التي تدرس جميع المجموعات تحت البحث باعتبارها مجموعات جزئية منها ويرمز لها (S) .

المجموعة الكلية Universal set : هي المجموعة التي تدرس جميع المجموعات تحت البحث باعتبارها مجموعات جزئية منها ويرمز لها (S) .

الاتحاد Union : اتحاد المجموعتين A و B أي A UB هو مجموعة العناصر الموجودة في A أو في B أو في كليهما وعبارة أخرى:

$$A \cup B = \{ x / x \in A \text{ or } x \in B \}$$

التقاطع Intersection : تقاطع المجموعتين B وA أي A∩B هي مجموعة العناصر المشتركة بين A و B أي هو مجموعة العناصر المنتمية الى A والى B في نفس الوقت، بعبارة أخرى:

$$A \cap B = \{ x / x \in B \text{ and } x \in A \}$$

أحيانا يستخدم الرمز AB للدلالة على A∩B

المتممة Complement : المتممة A (بالنسبة الى المجموعة الكلية S) هي مجموعة العناصر المنتمية الى المجموعة الكلية S ولاتنتمي الى A ويرمز لها A^c ، بعبارة أخرى:

$$A^C = \{ x/x \in s \ , \ x \notin A \}$$

التباديل Permutation : يقصد بالتباديل بأنها عدد طرق الأختيار المرتب التي يمكن تكوينها من عدة أشياء يأخذها كلها او بعضها ويرمز له ((nPr)) أي تباديل r من n وقانونه هو :

$$nPr = n! / (n-r)!$$

حيث ان :

$$n! = n (n-1) (n-2) \ .... \ 3*2*1$$

$$0! = 1 \quad , \quad 1! = 1$$

مثال : إذا كان لدينا أربعة حروف A , B , C , D وأختير منها حرفان، فما هي عدد طرق الاختيار المرتب التي يمكن بها أختيار هذين الحرفين ؟

$$nPr = n! / (n-r)!$$

$$4 P2 = 4 ! / (4-2)! = (4*3*2*1) / 2*1 = 12$$

أي ان عدد الطرق = 12 وهذه الطرق هي :

AB	AC	AD	BC	BD	CD
BA	CA	DA	CB	DB	DC

حيث ان كلاً منهما يمثل ترتيباً مختلفاً للحرفين

التوافيق Combinations :

يقصد بالتوافيق بأنها عدد طرق الاختيار غير المرتب التي يمكن تكوينها من عدة أشياء بأخذها كلها أو بعضها ويرمز له $\binom{n}{r}$ أو C_r^n وقانونه هو :

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

حيث ان :

$$n! = n(n-1)\,(n-2)\ .....\ 3\ 2\ 1$$

$$0! = 1 \quad , \quad 1! = 1$$

<u>مثال :</u> ماعدد طرق الاختيار التي يمكن الحصول عليها لاختيار لجنة مؤلفة من 5 أشخاص من مجموع 9 أشخاص ؟

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

$$= 9! \,/\, 5!\,(9-5)! = 9! \,/\, 5! * 4! = 126$$

8 – 6 قاعدتان أساسيتان يعتمد عليهما كل من التباديل والتوافيق :

1- إذا كان عدد الطرق الممكنة لوقوع الحادث E_1 هو n وان عدد الطرق الممكنة لوقوع الحادث E_2 هو m ، وان E_1 و E_2 حادثان متنافيان فإن عدد الطرق الممكنة لوقوع الحادث E_1 أو E_2 هو (n + m) من الطرق.

(n+m) عشوائية

ما هو عدد الطرق الممكنة لسحب ورقة من أوراق اللعب بحيث تكون الورقة المسحوبة من ورق السنك أو القلب؟

<u>مثال :</u> عدد أوراق اللعب هو 52 ورقة وان ورقة (السنك ♣) يمكن أن تحدث 13 طريقة وان ورقة (القلب ♥) يمكن أن تحدث 13 طريقة ايضاً فعند سحب ورقة عشوائياً من أوراق اللعب فإن عدد الطرق الممكنة لإختيار السنك أو القلب = 13 + 13 = 26 طريقة

2- إذا كان عدد الطرق الممكنة لوقوع الحادث E_1 هو n وان عدد الطرق الممكنة لوقوع E_2 هو m وكان E_1 و E_2 حادثان مستقلان فان عدد الطرق المكنة لوقوع الحادثان E_1 و E_2 هو (n m) من الطرق.

<u>مثال :</u> اذا سحبت ورقتان من مجموعة أوراق اللعب بحيث ان احداهما تكون (سنك) والاخرى (قلب) فان هناك $13 * 13 = 169$ طريقة لعمل ذلك .

<u>مثال :</u> صندوق فيه 6 كرات حمراء و 4 سوداء و 2 بيضاء فبكم طريقة يمكن أختيار 5 كرات بحيث تكون 3 منها حمراء و 2 سوداء ؟

الحل : عدد طرق اختيار 3 كرات حمراء هو $C_3^6 = 20$

عدد طرق أختيار 2 كرة سوداء هو $C_2^4 = 6$

∴ عدد الطرق لاختيار 3 كرات حمراء و 2 سوداء هو $C_3^6\ C_2^4 = 20 * 6 = 120$

8 – 7 قياس الاحتمال Measuring of Probability :

أفرض إن هنالك محاولة يمكن ان تقع بعدد من الحالات الكلية المتنافية (أي أنه لايمكن وقوع حادثتين أو أكثر في آن واحد لنفس المحاولة، مثال ذلك عند رمي قطعة نقود فان نتيجة الرمية هي أما صورة أو كتابة ولايمكن ظهور الاثنين معا) التي تملك نفس الفرصة في الظهور وأفرض إن هذا العدد هو n ، وأفرض إن m من هذهِ الحالات $m \le n$ ممكنة لوقوع حادثة معينة لتكن E ، عندئذ فإن إحتمال وقوع E يعرَف على النحو الآتي:

$$\text{Probability } (E) = Pr(E) = \frac{\text{الحالات الممكنة لوقوع الحادثة } E}{\text{الحالات الكلية لـ } E}$$

$$P = m / n$$

بعبارة أخرى لتكن E حدث يمكن أن يحدث بـ m طريقة وكانت n عدد جميع الحالات الكلية والتي لها نفس الفرصة في الظهور ، وبهذا فإن إحتمال حدوث الحدث (E) ⇦ Pr (E)

$$Pr (E) = m / n = P$$

إن الاحتمال (P) غالباً مايطلق عليه بأحتمال نجاح الحادثة E . إن عـدد الحالات الممكنة لعدم وقوع الحادثة E هو n – m وعندئذ فإن احتمال عدم وقوع E هو :

$$q = (n - m) / n$$

وغالباً مايطلق عليه باحتمال فشل الحادثة E وعليه فان :

$$q = (n - m) / n = 1 - m /n = 1 - p \Rightarrow p + q =1$$

8 – 8 بعض الملاحظات التي تساهم في توضيح فكرة ومفهوم الاحتمال :

1- ان كل من p و q كميات غير سالبة non – negative quantities ولايمكن ان تزيد قيمة أي منها عن الواحد وهذا يعني ان :

$$0 \le P \le 1 \quad , \quad 0 \le q \le 1$$

2- اذا كان $Pr (E) = 1$ فذلك يعني ان الحادثة E ((مؤكدة)) الوقوع، (مثال ذلك ماهو إحتمال وفاة شخص يوما ما؟ إن هذا الاحتمال مساوٍ للواحد لانه مؤكد وفاة أي شخص يوماً ما عاجلاً أو آجلاً).

3- إذا كان $Pr (E) = 0$ فذلك يعني ان الحادثة E ((مستحيلة)) الوقوع (مثال ذلك ماهو أحتمال ان يعيش شخص توقف قلبه عن العمل ؟ واضح إن هذا الاحتمال مساوٍ للصفر).

4- بشكل عام فإن إحتمال وقوع الحادثة E محصور في الفترة (0 , 1) أي ان $0 \le Pr(E) \le 1$

وعليه يمكن القول إنه إذا كانت $Pr (E) > ½$ فذلك يعني ان عدد الحالات الممكنة لوقوع E أكبر من عدد حالات عدم وقوعها، في حين إذا كانت $Pr (E) < ½$ فذلك يعني ان عدد الحالات الممكنة لوقوع E أقل من عدد حالات عدم وقوعها.

5- إن صغر أو كبر إحتمال وقوع الحادثة E يعتمد بطبيعة الحال على عدد الحالات الممكنة لوقوع E وعلى عدد الحالات الكلية المتاحة.

مثال 1 : سحبت قطرة دم من شخص ، ماهو إحتمال أن يكون صنف دم هذا الشخص هو AB ؟

الحل : ان عدد الحالات الكلية لعملية فحص الدم هي أربعة حالات وإن عدد الحالات الممكنة للصنف AB هي حالة واحدة، وعليه فإن إحتمال أن يكون صنف دم هذا الشخص AB هو:

$$Pr(AB) = 1/4 = 0.25$$

مثال 2: رمي زهرا نرد مرة واحدة،ماهو احتمال أن يكون مجموع النقاط على الوجهين الظاهرين :

A – مساوِ الى 9 B-مساوِالى 7 D- – أقل من 8 E- – أكبر من أو يساوي 6 .

الحل : بهدف التوضيح سنكون الجدول الآتي الذي يبين الحلات الكلية عند رمي زهرا النرد:

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

من الواضح ان عدد الحالات الكلية الممكنة هي 36 حالة:

A – إن الحالات الممكنة التي فيها مجموع النقاط مساوٍ الى 9 هي أربع حالات وعليه فإن :

$$Pr(E = 9) = 4 / 36 = 0.111$$

B – الحالات الممكنة التي فيها مجموع النقاط مساوٍ الى 7 هي ست حالات وعليه فان :

$$Pr(E = 7) = 6 / 36 = 0.167$$

D – إن عدد الحالات التي يكون فيها المجموع أقل من ثمانية هو 21 وعليه فان :

$$Pr(E < 8) = 21 / 36 = 0.583$$

E – عدد الحالات التي يكون فيها المجموع أكبر من أو يساوي 6 هو 26 عليه فان :

$Pr(E \geq 6) = 26/36 = 0.722$

مثال :

صندوق يحتوي على 12 كرة، ثلاث منها حمراء وأربعة بيضاء وخمس سوداء، أختيرت ثلاث كرات من هذا الصندوق بشكل عشوائي، ماهو أحتمال أن تكون :

-1- جميعها حمراء -2- جميعها بيضاء -3- جميعها سوداء -4- أثنين منها حمراء والاخرى بيضاء -5- واحدة بيضاء والبقية سوداء.

الحل : ان عدد الحالات الكلية لاختيار ثلاث كرات من هذا الصندوق هو :

$$c_3^{12} = \frac{12!}{3!(12-3)!} = 220$$

-1- عـدد الحالات الممكنة لان تكون الكرات الثلاث المختارة حمراء هو :

$$c_3^3\ c_0^4\ c_0^5 = 1$$

∴ إحتمال أن تكون الكرات الثلاث حمراء هو :

Pr (3 balls are red) = 1 / 220 = 0.0045

-2- عـدد الحالات الممكنة لأن تكون الكرات الثلاث المختارة بيضاء هو :

$$c_0^3\ c_3^4\ c_0^5 = 4$$

∴ احتمال ان تكون الكرات الثلاث بيضاء هو : Pr (3balls are white)= 4/220 = 0.0182

-3- عدد الحالات الممكنة لأن تكون الكرات الثلاث المختارة سوداء هو :

$$c_0^3\ c_0^4\ c_3^5 = 10$$

∴ احتمال ان تكون الكرات الثلاث سوداء هو: Pr(3balls are black) = 10/220 = 0.045

-4- عدد الحالات الممكنة لان تكون كرتين منها حمراء والثالثة بيضاء:

$$c_2^3\ c_1^4\ c_0^5 = 12$$

∴ احتمال ان تكون هنالك كرتين حمراء والثالثة بيضاء هو:

Pr(2 red balls and the other white) = 12 / 220 = 0.055

-5- عدد الحالات الممكنة لان تكون كرة بيضاء والبقية سوداء هو :

$$c_0^3 \; c_1^4 \; c_2^5 = 40$$

∴ احتمال ان تكون هنالك كرة بيضاء والبقية سوداء هو:

Pr(one white ball and the others black) = 40/220 = 0.182

8 – 9 قواعد عامة في نظرية الاحتمالات General rules in probability theory

أولاً : قاعدة الجمع The addition rule :

أ-عندما تكون الحوادث متنافية: لتكن E_1 ، E_2 حادثتين متنافيتين في محاولة معينة، عندئذ فان احتمال حدوث E_1 أو E_2 يمثل حاصل جمع احتمال E_1 مع احتمال E_2 ، أي أن :

$$Pr (E_1 \text{ or } E_2) = Pr (E_1) + Pr (E_2)$$

بشكل عام اذا كانت E_1 , E_2 ,........, E_k تمثل K من الحوادث المتنافية في محاولة معينة، عندئذ فان احتمال حدوث E_1 أو E_2 أو ...... E_k يمثل حاصل جمع احتمالات هذه الحوادث ، أي أن:

$$Pr (E_1 \text{ or } E_2 \text{ or } \ldots\ldots \text{ or } E_k) = \sum_{i=1}^{k} Pr\,(E_i)$$

مثال : تأمل تجربة رمي زهر النرد ، واضح إن إحتمال ظهور الوجه الذي يجعل عدد النقاط 3 أو الوجه الذي يجعل عدد النقاط 5 أو الذي عدد النقاط 6 هو :

$$Pr (3 \text{ or } 5 \text{ or } 6) = Pr (3) + Pr (5) + Pr (6) = 1/6 + 1/6 + 1/6$$

$$= 3/6 = 0.5$$ طالما إن هذه الحوادث متنافية

مثال : سحبت قطرة دم من شخص يرغب في فحص دمه ، ماهو إحتمال ان يكون دم هذا الشخص من

صنف غير الصنف O .

$$Pr(A \text{ or } B \text{ or } AB) = Pr(A) + Pr(B) + Pr(AB)$$

$$= 1/4 + 1/4 + 1/4 = 3/4 = 0.75$$

Or $Pr(A \text{ or } B \text{ or } AB) = 1 - Pr(O) = 1 - 1/4 = 3/4 = 0.75$

ب – *عنـــدما تكون الحوادث غير متنافية* :

لتكن E_1 و E_2 حادثتين ممكنتي الوقوع معاً في محاولة معينة، عندئذ فان احتمال وقوع E_1 أو E_2 هو حاصل جمع احتمال وقوع E_1 مع احتمال وقوع E_2 مطروحا من ذلك احتمال وقوعهما معاً، أي :

$$Pr(E_1 \text{ or } E_2) = Pr(E_1) + Pr(E_2) - Pr(E_1E_2)$$

<u>مثال :</u> إذا علمت إن احتمال ســقوط المطر في احد أيام فصل الربيع هو 0.52 وان إحتمال كون الطقس مشمس هو 0.36 وإن إحتمال كون الطقس مشمس وممطر في آن واحد هو 0.18 ، ماهو إحتمال ان يكون الطقس مشمس أو ممطر.

الحل : لنفرض ان E_1 تمثل حادثة سقوط المطر في ذلك اليوم ، وان E_2 تمثل حادثة كون الطقس مشمس في ذلك اليوم.

$$\therefore\ Pr(E_1 \text{ or } E_2) = Pr(E_1) + Pr(E_2) - Pr(E_1E_2)$$

$$= 0.52 + 0.36 - 0.18 = 0.70$$

<u>ثانياً : قاعـــدة الضرب The multiplication rule :</u>

أ- *عندما تكون الحوادث مستقلة* : لتكن E_1 و E_2 حادثتين مستقلتين، عندئذ فان إحتمال حدوث

E_1 و E_2 معاً هو :

$$Pr (E_1 \text{ and } E_2) = Pr(E_1) . Pr(E_2)$$

بشكل عام اذا كانت E_1 , E_2 ,……. E_k تمثل K من الحوادث لمحاولة معينة وإن هذه الحوادث مستقلة عن بعضها ، عندئذ فان احتمال حدوثها معا هو :

$$Pr (E_1 \text{ and } E_2 \text{ and } .... \text{ and } E_k) = \prod_{i=1}^{k} Pr (E_i)$$

مثال : تأمل تجربة رمي قطعتي نقود معدنية، وأفرض إن E_1 تمثل ظهور الوجه الذي يحمل صورة للقطعة الأولى، وإن E_2 تمثل ظهور الوجه الذي يحمل كتابة للقطعة الثانية.

الحل : واضح ان احتمال E_1 هو 0.5 وإن احتمال E_2 هو 0.5 ، عليه فان احتمال ظهور صورة في القطعة الاولى وكتابة في القطعة الثانية هو :

$$Pr (E_1 \text{ and } E_2) = Pr (E_1) . Pr (E_2) = 0.5 * 0.5 = 0.25$$

طالما إن الحادثان مستقلان

ب - *عندما تكون الحوادث مشـــروطة* ***Conditional events:***

لتكن E_1 حادثة في محاولة معينة وإن E_2 حادثة أخرى وقوعها مشروط بوقوع الحادثة E_1 ، عندئذ فان احتمال وقوع الحادثتين في آن واحد مســاوٍ لحاصل ضرب إحتمال وقوع E_1 في احتمال وقوع E_2 المشروطة بوقوع E_1 ، ويرمز لاحتمال وقوع E_2 المشروطه بوقوع E_1 بالشكل [$Pr (E_2/E_1)$] وعندئذ فان :

$$Pr (E_1 \text{ and } E_2) = Pr (E_1 . E_2) = Pr(E_1) . Pr (E_2 / E_1)$$ تصحيح

وهذا يعني إن :

(احتمال وقوع الحادثة E_2 المشروط بوقوع الحادثة E_1) $Pr(E_2/E_1) = Pr(E_1.E_2)/Pr(E_1)$

كذلك فان :

(احتمال وقوع الحادثة E_1 المشروط بوقوع الحادثة E_2) $Pr(E_1/E_2) = Pr(E_1.E_2)/Pr(E_2)$

<u>مثال :</u> تأمل تجربة رمي زهري النرد ولتكن E_1 حادثة تتمثل في أن يكون مجموع النقاط على وجهي الزهرين مساوٍ الى 6 ، ولتكن الحادثة E_2 تتمثل في أن يكون وجه أحد الزهرين يحمل نقاط عددها 2 ، جـد إحتمال حدوث E_2 المشروط بوقوع الحدث E_1 .

الحل: واضح ان عدد الحالات الممكنة لحدوث E_1 هي خمس حالات وهي :

(1,5),(5,1),(2,4),(4,2),(3,3)

عدد الحالات الكلية عند رمي زهري نرد = 36

∴ احتمال حدوث E_1 = عدد الحالات الممكنة / عددالحالات الكلية

$$\therefore\ Pr(E_1) = 5/36$$

∵ إن عدد الحالات الممكنة لحدوث E_2 بحيث يكون المجموع مساوٍ الى 6 هو حالتين فقط هما (2,4) و (4,2) ، عندئذ فان احتمال حدوث E_1 و E_2 معاً هو:

$$Pr(E_1 . E_2) = 2/36$$

عليه فإن إحتمال حدوث E_2 مشـروط بحدوث E_1 ، لذا فان :

$$Pr(E_2/E_1) = Pr(E_1.E_2)/Pr(E_1) = (2/36)/(5/36) = 2/5 = 0.4$$

8- 10 بعض التوزيعات الاحتمالية Some Probability distribution

أستعرضنا في الفقرات السابقة مفهوم الاحتمال وكيفية قياسه وبعض القواعد الهامة المتعلقه به، في هذه الفقرة سوف نتطرق الى دراسة موجزة لبعض التوزيعات ا لاحتمالية الشائعة الاستخدام في التطبيقات الاحصائية من خلال عرض مبسط لكل توزيع منها دون الدخول في التفاصيل.

أولاً : توزيع ذي الحدين The Binomial distribution

أفرض ان هنالك تجربة يمكن تكرارها بعدد معين من المحاولات المستقلة ليكن n ، بحيث ان نتائج كل محاولة تتمثل بنتيجتين ممكنتين فقط هما نجاح المحاولة أو فشلها، ولنفرض ان P يمثل احتمال نجاح المحاولة وان $q = 1-p$ يمثل احتمال فشلها، عندئذ فان احتمال الحصول على X من المحاولات الناجحة $(X \le n)$ من أجمالي هذه المحاولات هو:

$$Pr (X) = C_x^n \, p^x q^{n-x} \qquad \ldots\ldots\ldots \quad x = 0,1,2, \ \ldots\ldots \ ,n$$

يلاحظ من هذا الا ستنتاج ان هنالك متغير عشوائي من النوع المتقطع x (يشير الى المحاولات الناجحة)

وهنالك دالة بدلالة هذا المتغير هي p (x) هذه الـدالة تـعرف بـدالة الكتلة الاحتمالية للمتغير x Probability mass function وتعرف هذه الدالة بأسم توزيع ذي الحدين وغالباً مايتم الاصطلاح لهذا التوزيع بالشكل $x \sim b(n,p)$ ، ويمكن صياغة التعريف أعلاه بالشكل، حيث يقال ان المتغير العشوائي x يتوزع وفق توزيع ذي الحدين إذا كانت دالة الكتلة الاحتمالية تأخذ الشكل الآتي :

$$P (x) = \begin{cases} C_x^n \, p^x q^{n-x} & x = 0,1,2,\ldots\ldots,n \\ 0 & other\ wise \end{cases}$$

علما ان $0 < p < 1$ وان $q = 1 - p$

الجدول الآتي يبين بعض المقاييس الخاصة بهذا التوزيع :

المقياس	الصيغة
1- الوسط الحسابي لقيم x في هذا التوزيع هو	$\mu_x = np$
2- التباين لقيم x في هذا التوزيع هـــو	$\sigma_x^2 = npq$
3- الانحراف المعياري لقيم x في هذا التوزيع هو	$\sigma_x = \sqrt{npq}$

خصائص الدالة :

1- انها دالة وحيدة القيمة Single valued function.

2- انها دالة موجبة تتراوح بين $0 < pr(x) < 1$.

3- مجموع القيم الاحتمالية (الكتل الاحتمالية) المقترنة بكافة قيم المتغير x يجب ان يكون مســاوٍ للواحد، أي ان :

$$\sum_{x=0}^{n} C_x^n P^x \; q^{n-x} = 1$$

مثال : أفرض ان $x \sim b(5, 0.8)$ ، جـد مايأتي :

1- دالة الكتلة الاحتمالية p.m.f الى x .

2- رســم دالة الكتلة الاحتمالية لـ x .

3- حساب الوسط الحسابي والتباين لـ x .

4- $Pr(x=5)$, $Pr(x=2)$, $Pr(x \le 4)$, $Pr(x>2)$

الحل:

1- بما ان $x \sim b\,(5,0.8)$ لذا فان دالة الكتلة الاحتمالية الى x هي :

$$P\,(x) = C_x^5\,(0.8)^x(0.2)^{5-x}$$

2- لغرض رسم هذه الدالة سنوجد قيم x علما ان $x=0,1,2,3,4,5$

$P(x=0) = C_0^5\; 0.8^0\; 0.2^{5-0} = 0.00032$

$P(x=1) = C_1^5\; 0.8^1 0.2^{5-1} = 0.0064$

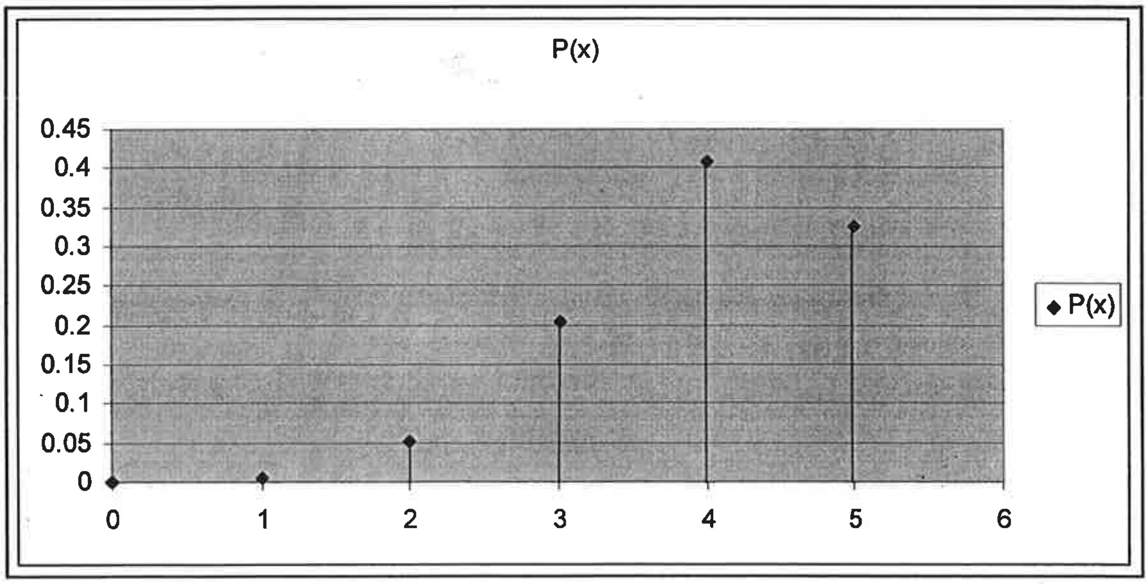
$P(x=2) = C_2^5\; 0.8^2\; 0.2^{5-2} = 0.0512$

$P(x=3) = C_3^5\; 0.8^3\; 0.2^{5-3} = 0.2048$

$P(x=4) = C_4^5\; 0.8^4\; 0.2^{5-4} = 0.409$

$P(x=5) = C_5^5\; 0.8^5\; 0.2^{5-5} = 0.327$

ثم تعيين احداثيات النقاط (x , P(x)) وكما موضح في الشكل :



3- $\mu_x = n\,p = 5 * 0.8 = 4$

$$\sigma_x^{\,2} = n\,p\,q = 5 * 0.8 * 0.2 = 0.8$$

4- من الجدول الســـابق نلاحظ ان

$$Pr\,(X = 5) = P\,(5) = 0.3276$$

$$Pr = (X = 2) = P\,(2) = 0.0512$$

$$Pr\,(X \le 4) = P(0) + P(1) + P(2) + P(3) + P(4)$$

$$= 1 - Pr\,(x = 5) = 1 - 0.32768 = 0.67232$$

$$Pr\,(x > 2) = P(3) + P(4) + P(5) = 0.94208$$

ثانياً : توزيع بواسون Poisson distribution

يعتبر توزيع بواسون واحد من التوزيعات المتقطعة ذات أهمية كبيرة في التطبيقات الاحصائية. ويمثل هذا التوزيع حالة خاصة من توزيع ذي الحدين عندما يكون احتمال نجاح المحاولة P صغيراً جـداً (نظرياً يقترب من الصفر) وان عدد المحاولات n كبيرجداً (نظرياً يقترب من ∞) بحث ان nP وفق هذين الشرطين لتستقر نحو عدد ثابت مثل λ ، عليه ووفق هذه الشروط فان توزيع بواسون هو توزيع تقاربي من توزيع ذي الحـــدين ، أي ان :

$$\lim_{\substack{p \to 0 \\ n \to \infty}} C_x^n\, p^x q^{n-x} = \frac{\lambda^x e^{-\lambda}}{x!} \quad , \quad x = 0, 1, 2, \ldots\ldots$$

يلاحظ من هذا الاستنتاج ان هنالك متغير عشوائي متقطع X وهنالك دالة بدلالة هذا المتغير هي:

$P\,(X) = \dfrac{\lambda^x e^{-\lambda}}{x!}$ ، هذه الدالة تعرف بأســـم دالة توزيع بواسون ، والتي تعبر عن احتمال الحصول على x من المحاولات الناجحة.

ان المتغير العشوائي x يتوزع وفق دالة توزيع بواسون اذا كانت دالة الكتلة الاحتمالية لهذا المتغير تأخذ الشكل التالي :

$$P(x) = \begin{cases} \dfrac{\lambda^{x}\, e^{-\lambda}}{x!} & , x = 0, 1, 2, \dots\dots\dots \\ 0 & \text{other wise} \end{cases}$$

حيث ان :-

$$e = 2.71828 \qquad , \quad \lambda > 0$$

أي أن $x \sim Po(\lambda)$ ، ويتم تشخيص التوزيع من خلال تحديد قيمة λ . ان هذه الدالة تتميز بنفس خصائص دالة توزيع ذي الحدين من حيث كونها دالة وحيدة القيمة وانها دالة موجبة وان :

$$\sum \frac{\lambda^{x}\; e^{-\lambda}}{x!} = 1$$

غالباً مايطلق على هذا التوزيع تسمية توزيع الحوادث النادرة كونه يتعامل مع حالات تطبيقية تتصف بكون ان احتمال نجاح المحاولة صغيرة جداً، على سبيل المثال حوادث سقوط الطائرات. الجدول الآتي يبين بعض المقاييس الخاصة بهذا التوزيع:

المقياس	الصيغة
1- الوسط الحسابي لقيم x في هذا التوزيع هو	λ
2- التباين لقيم x في هذا التوزيع هـو	λ
3- الانحراف المعياري لقيم x في هذا التوزيع هو	$\sqrt{\lambda}$

مثال : لوحظ في منطقة معينة ان من بين كل 1000 شخص هنالك شخص واحد مصاب بمرض معين، ماهو احتمال ان من بين 2000 شـــخص هنالك :

1- ثلاثة أشخاص مصابين بهذا المرض.

2- أكثر من شخصين مصابين بهذا المرض.

الحل : (طالما ان احتمال نجاح المحاولة صغير جدا وان عدد المحاولات كبير نستخدم توزيع بواسون)

ان احتمال الاصابة بهذا المرض هو 0.001 (احتمال نجاح المحاولة) ، عليه فان احتمال اصابة X شخص بهذا المرض هو :

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

وان $\lambda = np = 2000 * 0.001 = 2$

1- احتمال ان هنالك ثلاث أشخاص مصابين بهذا المرض هو :

$$P(x=3) = \frac{2^3 e^{-2}}{3!} = 0.18045$$

2- احتمال وجود أكثر من شخصين مصابين بهذا المرض هو:

$$Pr(x>2) = 1 - Pr(x \le 2) = 1 - [P(0) + P(1) + P(2)]$$

$$= 1 - 0.67668 = 0.32332$$

مثال : اذا علمت ان $x \sim Po(3)$ ، جـد الوسط الحسابي والتباين، ثم أرســم دالة هذا التوزيع مع حساب الاحتمالات الآتية: $Pr(x=2)$, $Pr(1 \le x \le 3)$, $Pr(x \le 2)$

الحل : حيث ان $\lambda = 3$

$$P(x) = \frac{3^x \ e^{-3}}{x!}$$

- الوسط الحسابي لقيم x هو ⇦ $\mu_x = \lambda = 3$
- تبايــن قيــم x هو ⇦ $\sigma_x^2 = \lambda = 3$

لغرض رســم دالة هذا التوزيع نعمل الجدول الآتي :

X :	0	1	2	3	4	5	6	7	8	...
P(x) :	0.05	0.15	0.22	0.22	0.17	0.1	0.05	0.03	0.01	....

Where $P(\ x=0\ ) = \frac{\lambda^x \ e^{-\lambda}}{x!} = \frac{3^0 \ e^{-3}}{0!} = \frac{1}{(\ 2.71828\ )^3} = \frac{1}{20.0855} = 0.05$

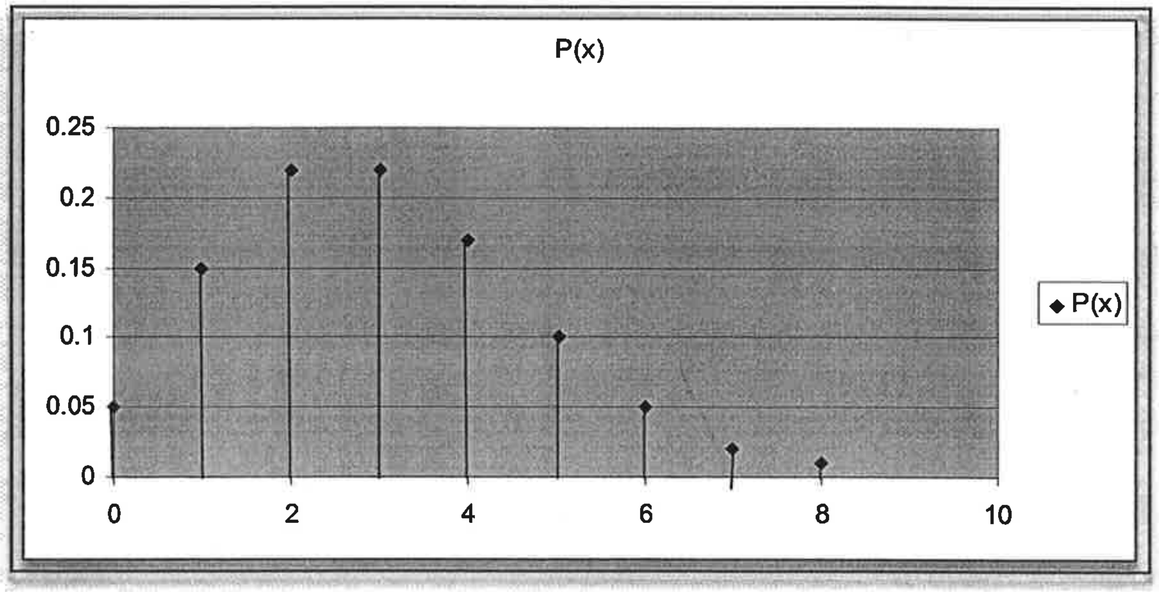
$$P\ (\ x=1\ ) = \frac{3^1 \ \ e^{-3}}{1!} = \frac{3}{(2.71828\ )^3} = 0.15$$

$$P(\ x=2\ ) = \frac{3^2 \ e^{-3}}{2!} = \frac{9}{2*(\ 2.71828\ )^3} = 0.22$$

$$P(\ x=3\ ) = \frac{3^3 \ e^{-3}}{3!} = \frac{27}{(3*2*1)*(2.71828)} = 0.22$$

وهكذا يتم التعويض بنقاط المتغير(x) الى 8 = x أو أكثر ليتسنى لنا رسم دالة التوزيع .....

ومن ثم نعين احداثيات النقاط (x , P(x)) كما موضح في الشكل الآتي :



$$Pr\ (1 \le x \le 3) = p(1) + p(2) + p(3)$$

$$= \frac{3^1 e^{-3}}{1!} + \frac{3^2 e^{-3}}{2!} + \frac{3^3\ e^{-3}}{3!} = 0.59$$

$$Pr\ (x \le 2) = p\ (0) + p\ (1) + p(2)$$

$$= \frac{3^0\ e^{-3}}{0!} + \frac{3^1\ e^{-3}}{1!} + \frac{3^2\ e^{-3}}{2!} = 0.42$$

ثالثاً : التوزيع الطبيعي The Normal distribution :

يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية الشائعة الاستخدام في النظرية الأحصائية، حيث ان غالبية الظواهر الطبيعية تسلك وفق هذا التوزيع، وهو من التوزيعات الاحتمالية المستمرة.

يقال ان المتغير العشوائي المستمر (x) هو ذا توزيع طبيعي اذا كانت الدالة الاحتمالية لهذا التوزيع تأخذ الشكل الآتي:

$$f\ (x) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}}\ e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)} \qquad -\infty < x < \infty$$

$$0 < \sigma < \infty \quad , \quad -\infty < \mu < \infty$$

حيث ان $\pi = 3.14159$, $e = 2.71828$

إن μ تمثل الوسط الحسابي لقيم (x) في التوزيع، σ الانحراف المعياري لقيم x ، وتسميان بمعلمتي التوزيع، وغالباً مايتم اعتماد الشكل الآتي للدلالة على هذا التوزيع $x \sim N(\mu, \sigma^2)$ الذي يعني ان المتغير x يسلك (~) وفق دالة توزيع طبيعي (N) بوسط قدره (μ) وتباين مقداره (σ^2) ، وحيث ان دالة هذا التوزيع هي دالة احتمالية فانها يجب أن تتصف بما يلي:

1- ان f(x) دالة وحيدة القيمة، أي مانعنيه ان لأية من قيم x المعوضة في f(x) تعطي قيمة واحدة فقط الى f(x) .

2- ان f(x) دالة موجبة دائما مهما كانت قيمة x سالبة أم موجبة.

3- ان المساحة تحت منحني الداله f(x) ضمن الفترة ($-\infty,\infty$) تساوي واحـــد، أي ان :