جهاز نشر وتوزيع الكتاب الجامعى - جامعة حلوان

الإحصاء التطبيقي للتجاريين

الجزء الأول

الطبعة الرابعة

دكتورة
عفاف على حسن الدش
أستاذ الإحصاء ورئيس قسم الرياضة والتأمين والإحصاء التطبيقى
كلية التجارة - جامعة حلوان

القاهرة
٢٠٠٠

الإحصاء التطبيقى

الجزء الأول

الطبعة السادسة

الدكتورة

عفاف على حسن الدش

أستاذ الإحصاء **و**وكيل كلية التجارة **و**إدارة الأعمال

جامعة حلوان

الناشر

جهاز نشر وتوزيع الكتاب الجامعي

بجامعة حلوان

٢٠٠٥

الإحصاء التطبيقى

الطبعة السادسة

الدكتورة

عفاف على حسن الدش

أستاذ الإحصاء ووكيل كلية التجارة وإدارة الأعمال

جامعة حلوان

الطبعة الأولى : في سبتمبر 1992 برقم ايداع 1683/93

الترقيم الدولي : I.S.N.B : 977-00-4547-0

الطبعة الثانية : في سبتمبر 1994 برقم ايداع 1683/93

الترقيم الدولي :I.S.B.N : 977-00-4547-0

الطبعة الثالثة : في سبتمبر 1998 برقم ايداع 11721/98

الترقيم الدولي : I.S.B.N : 977-204-512-5

الطبعة الرابعة : في سبتمبر 2000 برقم ايداع 13225/2000

الترقيم الدولي:I.S.B.N : 977-5061-26-1

الطبعة الخامسة : في سبتمبر

الترقيم الدولي :I.S.B.N

الطبعة السادسة : في سبتمبر 2005 برقم ايداع

الترقيم الدوليI.S.B.N

الناشر

جهاز نشر وتوزيع الكتاب الجامعي

بجامعة حلوان

٢٠٠٥

بِسۡــــمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيــــمِ

"فَأَمَّا الزَّبَدُ فَيَذْهَبُ جُفَاءً وأَمَّا مَا يَنْفَعُ النَّاسَ فيَمْكُثُ فِي الأْرْض كَذَلِكَ يَضْرِبُ اللَّهُ الأْمْثَالَ"

سورة الرعد الآية

١٧

صدق الله العظيم

الفهرس

الباب الأول

أهمية دراسة الإحصاء

The Importance Of Studying Statistics

(١-١) مقدمـــة
Introduction

لقد شهد مطلع هذا القرن تطور عظيم في علوم وتطبيقات الاتصالات وتكنولوجيا المعلومات والحاسبات وانعكس هذا التطور على باقي العلوم الأخرى؛ بل على جميع مجالات الحياة مما جعلنا نشعر أننا محاطين من جميع الجوانب بسياج من الأرقام.
فعلى سبيل المثال نجد أن الجرائد اليومية والنشرات الإخبارية في التليفزيون ، الراديو ، ....، الخ ؛ تتضمن أرقام كدرجات الحرارة المتوقعة مثلاً أو الأرقام الخاصة بأسعار العملات ،الأرقام القياسية للأسعار ، معدلات البطالة ، معدلات المواليد، ....، الخ ؛ مما يتطلب ضرورة فهم وتفسير مدلول هذه البيانات الرقمية حتى نتمكن من تكوين حس رقمي Numerical Sense صحيح يمكننا من استخدام هذه البيانات الرقمية أفضل استخدام وهذا يتطلب ضرورة معرفة كيف يتم جمع هذه البيانات وتلخيصها وتحليلها وتخزينها.
وقد أدى وجود الحاسب الآلي Computer وبساطة استخدامه إلى إنجاز هذه العمليات "جمع البيانات – وتلخيصها - وتحليلها – وتخزينها بكفاءة وسرعة هائلة "*

(١-٢) صـــــــــــــــناعة القـــــــــــــــرار
Decision Making

كل فرد يعتبر متخذ قرار في جميع الأمور المتعلقة به ؛ فعلى سبيل المثال يعتبر الطالب متخذ قرار في تحديد مدة استذكار دروسه في اليوم وتوزيع هذه المدة على المقررات المحددة وإعطاء ترتيب لهذه المقررات في استذكارها ؛ كذلك يعتبر رب الأسرة متخذ قرار في توزيع دخله الشهري على أوجه الأنفاق المختلفة للأسرة بما يحقق لها مستوى المعيشة المناسب .
وفي معظم الأحيان تكون البيانات والمعلومات المطلوبة لاتخاذ القرار غير متاحة بمعنى أنها موجودة ولكن غير ممكن الحصول عليها بالنسبة لمتخذ القرار كما في حالة البيانات والمعلومات عن الشركات المتنافسة كذلك بالنسبة للأهداف الاقتصادية والعسكرية للدول ،....الخ.
وقد تكون البيانات المطلوبة لاتخاذ القرار غير موجودة بشكل مؤكد Certain Data كما في حالة البيانات في فترات زمنية مستقبلية فإن البيانات في هذه الحالة تكون غير موجودة وبالتالي تكون المعلومات المتعلقة بهذه الفترات المستقبلية معلومات غير مؤكدة Uncertain Information وبالتالي تصبح القرارات المبنية على هذه المعلومات غير المؤكدة قرارات تتضمن جزء من المخاطرةRisk

∗ Paul Newbold - William Carlson & Betty Thorne (2003) "Statistics For Business Economics"

وقد أدى التطور العظيم في علوم الرياضيات Mathematics Sciences والإحصاء Statistics إلى وجود مجموعة من الأساليب التي يطلق عليها أساليب الأمثلية Optimization Techniques حيث يمكن باستخدامها صناعة القرارات المثلي* في ظل وجود معلومات غير كاملة أو معلومات غير مؤكدة.

(١-٣) التفكير الإحصائي Statistical Thinking

صناعة القرار الأمثل يتطلب وجود مجموعة من العناصر الضرورية وأهم هذه العناصر هو توافر معلومات Information صحيحة ودقيقة والبيانات الصحيحة هي المادة الرئيسية لهذه المعلومات ؛ ومن ثم فأنه من الضروري أن تتوافر لدى متخذ القرار صورة رقمية صحيحة ودقيقة للمشكلة محل الدراسة.

ومن هنا تتضح أهمية التفكير الإحصائي في اتخاذ قرار لمشكلة ما ، وذلك من خلال تعريف المشكلة وتحديد البيانات المطلوبة وجمع هذه البيانات وعرضها بشكل يوضح الجوانب المختلفة للمشكلة ثم استخراج المعلومات المطلوبة من هذه البيانات لاتخاذ القرار ، ويمكن تلخيص التفكير الإحصائي بأنه رؤية المشكلة بشكل عام في صورة رقمية وتناولها من خلال البيانات المطلوبة.

(١-٤) علم الإحصاء The Science Of Statistics

منذ قديم الزمان استخدمت كلمة إحصاءStatistics للتعبير عن عملية العد (أو الحصر) للأشياء ، حيث استخدمت الإحصاء للتعبير عن حصر موارد الدولة من سكان وموارد ؛ فعلى سبيل المثال حصر عدد الذكور في سن التجنيد وعدد السكان بهدف تحديد قيمة الضرائب ، وحصر عدد المخزون من حبوب ومواشي ،...،الخ وذلك بهدف إعداد الجيوش وتكوين العتاد لهم.

وفي القرون السابقة عرف العلماء الإحصاء على أنه "العلم الذي يهتم بطرق وأساليب جمع وعرض البيانات"**.

ولكن مع التطور العظيم في جميع حقول المعرفة "اقتصادية ، إدارية ، عسكرية ، طبية ،...،الخ" وتطور مفهوم الدولة والتزايد الهائل في عدد السكان و التوسع في أنشطتهم و تخصصاتهم مما أدى إلى أهمية بل ضرورة إخضاع الأمور إلى القياس الكمي Quantitative Measurement.

حيث أصبحت لغة الأرقام Numerical Languageهي أكفأ اللغات Efficient Language في دراسة المشاكل سواء كانت مشاكل علمية Scientific Problems،أو المشاكل الإنسانية Human Problems ، أو المشاكل الاقتصادية Economic Problems أو الإدارية Managerial Problems أو اجتماعية Social Problems ...، الخ.

---

* عفاف الدش (٢٠٠٠) بحوث العمليات وصناعة القرارات – جهاز نشر وتوزيع الكتاب – جامعة حلوان.

** Heinz Kohler (1994) –"Statistics For Business &Economics"-Third Edition.

وقد أدى التطور الهائل في علوم الرياضيات Mathematics sciences وعلوم الحاسبات Computer Sciences والاتصالات Communication في النصف الثاني من القرن العشرين إلى تطوير وتطبيق علم الإحصاء على نطاق واسع في المجالات المختلفة فباستخدام الأساليب الإحصائية يمكن الحصول على المعلومات Information المطلوبة من البياناتData المتاحة.
ومن هنا برزت أهمية علم الإحصاء وأهمية تطويره والتوسع في تطبيقاته في المجالات المختلفة ليس فقط لجمع وعرض البيانات وإنما في الحصول على معلومات Information من البيانات التي تم جمعها؛ فأصبح علم الإحصاء بمفاهيمه الحالية هو العلم الذي يهتم بطرق وأساليب جمع وتحليل البيانات Colleting &Analyzing Data واستخراج المعلومات لتطبيق أساليب الأمثلية Optimizing Techniques التي تمكن متخذي القرارات Decisions Makers من الوصول إلى أفضل القرارات Best Decisions الممكنة في ظل الحقائق* غير الكاملةIncomplete Facts أو غير المتاحة لمتخذ القرار "كما في حالة دراسة الوضع المالي للشركات المنافسة للشركة محل الدراسة"، أو في ظروف عدم التأكد Uncertainty States "كما في حالة التنبؤ بقيم ظواهر معينة في المستقبل مثل الطلب في فترة مقبلة وتقدير الحجم الأمثل للعرض وفقاً لتقدير حجم الطلب"
ويمكن تلخيص ما سبق على النحو التالي : -

<u>تعريف</u>

علم الإحصاء "هو العلم الذي يهتم بطرق وأساليب جمع وعرض وتحليل البيانات كذلك بطرق وأساليب الأمثلية في حالة وجود معرفة غير كاملة أو ظروف عدم التأكد بالنسبة للظاهرة أو الظواهر محل الدراسة".
ووفقاً للتعريف السابق لعلم الإحصاء فإننا يمكننا تقسيم هذا العلم إلى ثلاثة فروع رئيسية على النحو التالي :-

أولاً: الإحصاء الوصفي **Descriptive Statistics

هو الفرع الذي يهتم بتقديم وتطوير وتطبيق طرق وأساليب جمع وعرض البيانات.
وفي الباب الثاني والثالث والرابع من هذا الكتاب سوف نتناول بالدراسة من الناحيتين النظرية و التطبيقية بعض أهم هذه الطرق والأساليب.

ثانياً: الإحصاء التحليلي Analytical Statistics

هو الفرع الذي يهتم بتقديم وتطوير وتطبيق طرق وأساليب تحليل البيانات الرقمية Numerical Data وعن خصائص وسلوك الظواهر محل الدراسة وأسبابها وعلاقتها بالظواهر الأخرى.

* Lawrence L.Lapin (1994)"Quantitative Methods For Business Decisions–With Cass"
** Mc Clave , Benzon & Sincich (1998) "Statistics For Business & Economics" .

وفي العديد من الكتابات يطلق على الإحصاء التحليلي* الاستدلال"أو الاستنتاج" الإحصائي inferential Statistics ؛ وذلك نظراً لأنه باستخدام هذه الطرق والأساليب يمكن الوصول إلى تقدير أو تنبؤ أو استنتاج معلومات معينة عن الظواهر محل الدراسة أو استنتاج أي معرفة رقمية عن مجتمع الدراسة من خلال بيانات عينة (أو عينات).

ثالثاَ: صناعة القرارات Decision's Making

ويختص هذا الفرع بتقديم وتطوير وتطبيق أساليب صناعة القرارات المثلى في ظل وجود معرفة غير كاملة أو في ظل ظروف عدم التأكد**.

كذلك يهتم بتقديم أساليب تقييم Evaluation Techniques مراحل تنفيذ هذه القرارات ، حيث تعتمد هذه الأساليب على البيانات المتعلقة بالظواهر محل الاعتبار كمادة أساسية في تطبيق هذه الأساليب ؛ فعلى قدر توافر هذه البيانات وشمولها ودقتها تتوقف صحة القرارات المبنية عليها.

والتقسيم السابق يعتبر تقسيم تصاعدي بمعنى أن إجراء التحليل الإحصائي لظاهرة ما يتطلب أولاَ جمع البيانات وعرضها في الصورة التي يتطلبها التحليل أي لابد أولاَ من استخدام أساليب الإحصاء الوصفي . كذلك استخدام طرق وأساليب اتخاذ القرارات ، لابد أن يسبقه أولاَ تحليل إحصائي لخصائص وسلوك الظواهر محل الدراسة والعلاقات بينها وأساليب التحليل تتطلب فيها أساليب وطرق التحليل الوصفي .

ويمكن تلخيص ذلك في شكل (١-١) الذي يوضح دور الإحصاء في المراحل المختلفة لصناعة القرار.

ونظراَ لأهمية الإحصاء في دراسة وحل المشاكل المختلفة أو التخطيط لأنظمة جديدة أو التوسع وتحديث الأنظمة القائمة ، لذا أهتم الإحصائيين بابتكار أساليب إحصائي خاصة بدراسة قطاعات معينة وعلى سبيل المثال:

- الإحصاء السكاني Demography Statistics
- الإحصاء الطبي Medical Statistics
- الإحصاء الزراعي Agricultural Statistics
- الإحصاء الاجتماعي Social Statistics
- الإحصاء الفيزيائي Physical Statistics
- الإحصاء الاقتصادي Economic Statistics

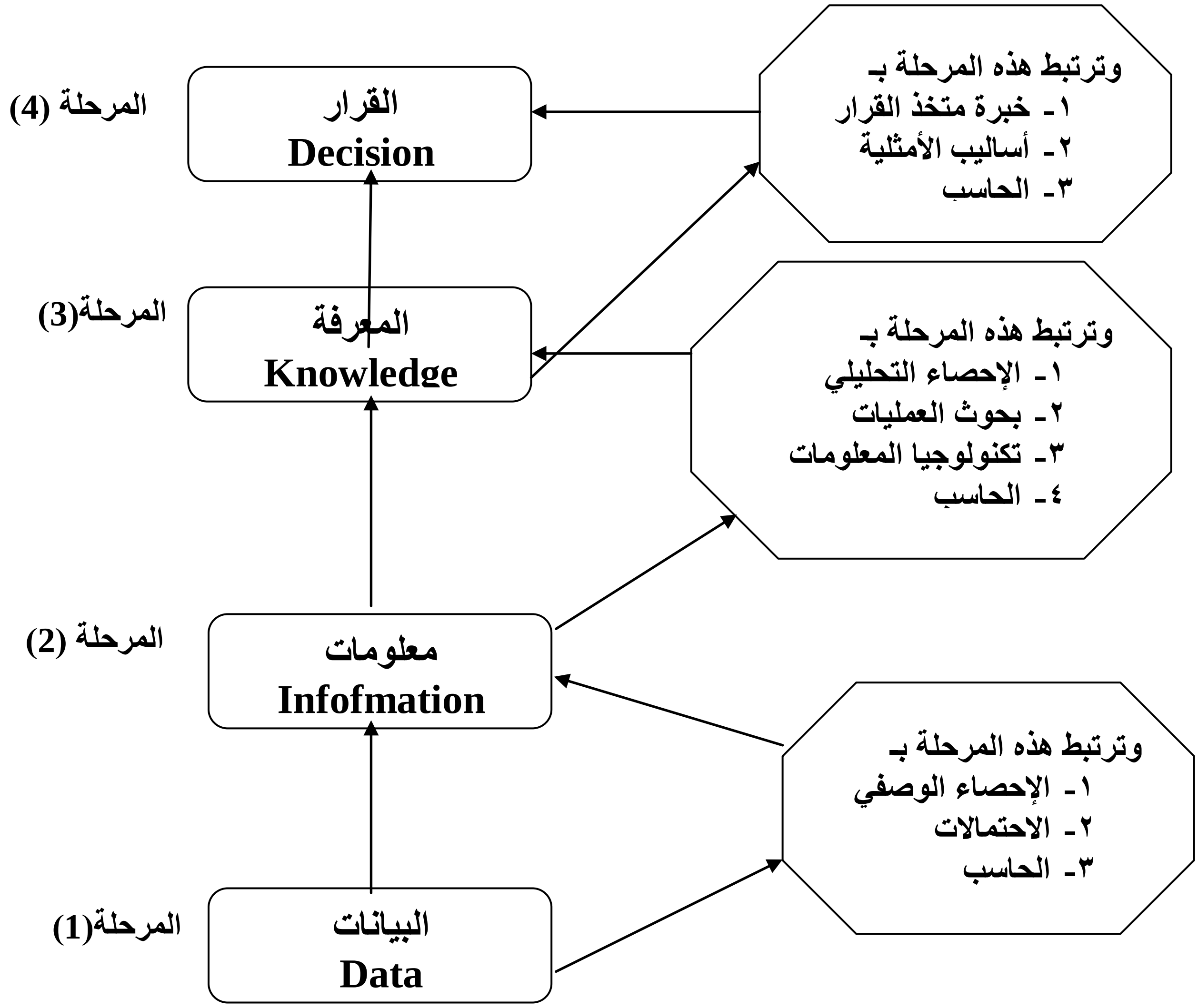
وكما سبق أن ذكرنا أن التطور العظيم في الرياضيات وعلوم الحاسب ونظم المعلومات أدى إلى تطوير هائل في علم الإحصاء وتطبيقاته العلمية والإنسانية .

ونظراَ للأهمية البالغة للإحصاء في حياتنا فقد أنشئت العديد من الأجهزة والمؤسسات الحكومية وغير الحكومية التي تختص بجمع وتحليل البيانات الخاصة بالقطاعات المختلفة ؛ فعلى سبيل المثال الجهاز المركزي للتعبئة العامة والإحصاء

* Mcclave J.T Benson P.G & Sincich ,T (1998)" Statistics For Business &Economics"

** أ.د عفاف الدش (١٩٨٧) "بحوث العمليات وصناعة القرارات "

مراحل صناعة القرار
والعلوم المرتبطة بكل مرحلة*



شكل (١-١)

وبنوك المعلومات كذلك أنشأت المعاهد والأقسام العلمية للإحصاء وتطبيقاتها بالجامعات والمراكز البحثية ، وأخيراً أدخلت الإحصاء كمادة للطلاب بالمرحلة الإعدادية بهدف خلق الوعي الإحصائي لدى الأفراد كذلك تكوين كوادر من الإحصائيين القادرين على تطويع الإحصاء لحل المشاكل الفعلية بالإضافة إلى تطوير علم الإحصاء وربطه بالعلوم والفنون الأخرى.

* Paul Newbold ,Willam L.Carlson;&Betty M . Thorne(2003) "Statistical For Business & Economic"

ومما سبق يتضح أن البيانات هي المادة الرئيسية في التحليل الإحصائي ؛ فعلى قدر توافرها وشمولها ودقتها تتوقف دقة النتائج المستخلصة من التحليل ، وبالتالي تتوقف صحة القرارات المبنية على هذه النتائج وكذلك يتوقف عليها مستوى المخاطرة المترتبة على هذه القرارات*.

(١-٥) مراحل الدراسة الإحصائية
The Stages Of Statistical Study

وبصفة عامة تتطلب الدراسة باستخدام الأساليب الإحصائية أربعة مراحل متتالية يمكن تلخيصها على النحو التالى:-

المرحلة الأولى:

وفي هذه المرحلة يتم تحديد عناصر المشكلة وأسبابها والأهداف المباشرة المراد الوصول إليها وفقاً لأولويتهافي صورة رقمية ووفقاً لعناصر المشكلة وأهدافها ويتم تحديد البيانات المطلوبة ومصدرها ووفقاً لذلك يتم تصميم استمارة جمع البيانات (في الباب التالى سوف نتعرض لتصميم الإستمارة)حيث يتم جمع البيانات اللازمة لدراسة الظواهر (المتغيرات) محل الإعتبار والتأكدمن صحة وشمول هذه البيانات باستخدام الأساليب الإحصائية الملائمة.

المرحلة الثانية:

وفي هذه المرحلة يتم عرض البيانات التى تم جمعها في المرحلة الأولى بأساليب ملائمة توضح ما تنطوي عليه هذه البيانات من خصائص للظواهر محل الدراسة والعلاقات بينها بصورة واضحة ومفهومة أي استخلاص أهم المعلومات المتعلقة بهذه الظاهرة بحيث تعطى معلومات رقمية عن الحالة تمكن متخذ القرار من تكوين رؤية رقمية محددة عن المشكلة محل الدراسة والأبعاد المختلفة لها.

المرحلة الثالثة:

ويتم فيها تحليل البيانات والمعلومات السابقة الحصول عليها في المرحلة الثانية بالأساليب الملائمة لطبيعة المشكلة أو الظاهرة ووفقاً للأهداف المراد تحقيقها للوصول إلى معرفة رقمية كمرحلة سابقة لتحديد الحلول الممكنة و البديلة للمشكلة.

المرحلة الرابعة:

وفي هذه المرحلة وبناء على المعرفة الرقمية التي تكونت في المراحل السابقة عن المشكلة أو الظواهر محل الراسة وفقاً للبيانات المتاحة يتم تحديد أساليب الأمثلية Optimization Techniques المناسبة لتطبيقها لإتخاذ أفضل القرارات

* أ.د. عفاف الدش (١٩٩٦) "الإحصاء وصناعة القرارات"

في ظل ظروف عدم التأكد كذلك قياس المخاطرة المترتبة على القرارات التي يتم التوصل إليها.

ملحوظة: تابع هذه المراحل من خلال الشكل (١-١)

ومن الملاحظ أن معظم الكتب الموجودة حالياً في الإحصاء بشكل عام والإحصاء التطبيقي بشكل خاص تركز على تقديم الطرق و الأساليب الإحصائية والخطوات الحسابية للحصول على النتائج مع عدم التركيز الكافى على كيف يمكن للمستخدم لهذه الأساليب والطرق أن يحدد الطريقة أو الأسلوب المناسب للمشكلة التي يتناولها في كل مرحلة حيث تتوقف طبيعة وأهمية المعرفة والمعلومات المستخلصة من البيانات على الطريقة والأسلوب الإحصائي المستخدم.

لذا يهدف هذا المرجع إلى تقديم المفاهيم والأساليب الإحصائية المطلوبة لتناول مراحل الدراسة الإحصائية في الجزء الأول وتناول المرحلتين الأولى والثانية والمقدمة الأساسية للمرحلة الثالثة من مراحل الدراسة الإحصائية السابق ذكرها مع التركيز على كيفية تحديد الطريقة أو الأسلوب المناسب للمشكلة محل الدراسة وذلك من خلال تقديم العديد من الأمثلة التطبيقية ويتناول الجزء الثاني من هذا المرجع المرحلتين الثالثة والرابعة بالتفصيل.

(١-٦) استخدام الحاسب Using Computer

يعتبر الحاسب الآليComputer أداة فعالة وبالغة الأهمية في استخدام وتطبيق الأساليب الإحصائية لما له من سعة كبيرة لحفظ وتخزين البيانات ولما له من سرعة فائقة في إجراء العمليات الحسابية ، ويوجد حزم لبرامج الحاسب لاستخدام الأساليب الإحصائيStatistical Software Package مثل:-

1- Microsoft Excel
2- Minitab
3- SPSS Windows

وسوف نستخدم في هذا المرجع برنامج Excel للحصول على النتائج لبعض التطبيقات وذلك لبساطة وسهولة استخدامه كما سوف نوضح فيما بعد.

ومما هو جدير بالذكر أنه لم يكن ممكن استخدام وتطبيق الأساليب الإحصائية على نطاق واسع وبالكفاءة والسرعة المطلوبة في حالة عدم وجود أجهزة حاسبات .

(١-٧) تمرينـات **Exercises**

(١-١)

لماذا تعتبر الإحصاء ذو أهمية في المجالات التالية :

١- المحاسبة	٢- التسويق
٣- الدراسات المالية	٤- تكنولوجيا المعلومات
٥- الدراسات الطبية	٦- الدراسات التعليمية

(١-٢)

ناقش أهمية وفوائد التفكير الإحصائي.

(١-٣)

حدد المتغيرات والعلاقات بينهم في كل حالة من الحالات التالية:

١- وصول ورحيل الطائرات على أحد خطوط الطيران من القاهرة إلى لندن.

٢- برامج إنقاص الوزن.

٣- برنامج لإستخراج شبكة الأنترنت.

٤- دورة تدريبية لمعاونى أعضاء هيئة التدريس بإحدى كليات التجارة.

(١-٤)

حدد مقالة أو أكثر في إحدى الجرائد اليومية

- يجب على القارئ عند الإطلاع عليها التفكير بأسلوب رقمي في أوجهها المختلفة.

الباب الثاني
جمع البيانات الإحصائية
Collecting Statistical Data

(٢-١) أنواع البيانات **Types Of Data**

(٢-٢) مصادر **Data Sources**

(٢-٣) أساليب الدراسة الاحصائية **Techniques Of Statistical Study**

(٢-٤) أساليب جمع البيانات **Techniques Of Collecting Data**

(٢-٥) تصميم استمارة جمع البيانات **Design Of Questionnaire**

(٢-٦) أمثلة تطبيقية **Applied Examples**

(٢-٧) تمرينات **Exercises**

(٢-١) أنواع البيانات   Types of Data

البيانات التي يتم جمعها قد تكون فى صورة رقمية مثل عمر الفرد (فيمكن أن يكون عمر الشخص 21 سنة مثلاً) أو دخل الأسرة (فيمكن أن يكون الدخل الشهرى للأسرة 1500 جنية مثلاً) وفى هذه الحالة تسمى البيانات بيانات كمية Quantitative Data ، وقد تكون البيانات المطلوب جمعها بيانات غير كمية مثل الحالة الإجتماعية للفرد (فيمكن أن تكون الحالة الإجتماعية : أعزب – متزوج – مطلق – أرمل) وكذلك تقدير الطالب فى المرحلة الجامعية (فيمكن أن يكون التقدير : مقبول – جيد – جيد جداً – إمتياز) فالبيانات فى هذه الحالة بيانات غير رقمية وتسمى بالبيانات الغير كمية Unquantitative Data أو بالبيانات الوصفية أو النوعية Qualitative Data .

والبيانات الغير كمية يمكن فى بعض الحالات تحويلها إلى بيانات كمية أى فى صورة رقمية وتسمى فى هذه الحالة بالبيانات الترتيبية Ranked Data ، فمثلاً تقدير الطالب يمكن الإشارة إليه كما يلى :

تقدير مقبول بالرقم (1)
تقدير جيد بالرقم (2)
تقدير جيد جداً بالرقم (3)
تقدير إمتياز بالرقم (4)

أى أنه تم الترتيب ترتباً تصاعدى ، ويمكن أن يكون الترتيب تنازلى وذلك على النحو التالى :

مقبول ⟵ (4)
جيد ⟵ (3)
جيد جداً ⟵ (2)
إمتياز ⟵ (1)

مثال (٢-١)
إذا كان لدينا عينة من العمال تم تحديد مستواهم المهنى وكان على النحو التالى :
جيد جداً ، ضعيف ، متوسط ، جيد ، جيد جداً ، جيد جداً

المطلوب

١- رتب البيانات ترتيب تصاعدى.
٢- أوجد المستوى المتوسط لهؤلاء العمال.

الحل

الترتيب التصاعدي على النحو التالي :

ضعيف ⟵ (1)

متوسط ⟵ (2)

جيد ⟵ (3)

جيد جداً ⟵ (4)

ومن ثم يصبح مستوى العمال على النحو:

4, 1, 2, 3, 4, 4

فإذا رمزنا إلى المستوى المتوسط بالرمز $\overline{X}$ فإن

$$\overline{X} = \frac{\text{مجموع المستويات}}{\text{عدد العمال}}$$

$$= \frac{4+1+2+3+4+4}{6} = \frac{18}{6} \qquad \overline{X} = 3$$

وبما أن 3 تشير إلى المستوى جيد فمن هنا يمكن القول أن المستوى المتوسط لهؤلاء العمال جيد.

وكما ذكرنا سابقاً أن بعض البيانات الوصفية (الغير كمية) يمكن تحويلها إلى بيانات كمية إذا كانت البيانات بيانات ترتيبية كما فى مثال (٢-١) ولكن قد تكون البيانات غير الكمية (بيانات غير ترتيبية) ، فمثلاً الحالة الإجتماعية للفرد (أعزب ، متزوج ، مطلق ، أرمل) فهنا لا يوجد ترتيب وإذا رتبت لتحويلها من بيانات غير كمية إلى بيانات كمية سوف تعطى نتائج مضلله كما سوف يتضح فى المثال التالى

مثال (٢-٢)

إذا كان المراد دراسة الحالة الإجتماعية وتم إعطاء الترتيب التالى:

يشار إلى أعزب بالرقم ⟵ (1)

يشار إلى متزوج بالرقم ⟵ (2)

يشار إلى مطلق بالرقم ⟵ (3)

يشار إلى أرمل بالرقم ⟵ (4)

واجريت الدراسة على عينة مكونة من 10 أفراد وبسؤال كل منهم عن حالته الإجتماعية فكانت إجابتهم على النحو التالى

1 ,1 ,3 ,4 ,1 ,1 ,2 ,3 ,1 ,3

المطلوب :

معرفة الحالة الإجتماعية فى المتوسط

الحل

إذا أشرنا للمتوسط بالرمز $\bar{y}$ فإن :

$$\bar{y} = \frac{1+1+3+4+1+1+2+3+1+3}{10} = \frac{20}{10} = 2$$

وحيث أن الرقم 2 يشير إلى متزوج فنجد أن الحالة الإجتماعية فى المتوسط متزوج ، وهنا يعتبر هذا الرقم 2 نتيجة مضللة لأن معظم الأفراد فى هذا المثال حالتهم أعزب ثم يليهم حالة مطلق ، أى أنه لم يكن مناسب استخدام المقياس الإحصائى (المتوسط) لوصف الحالة الإجتماعية ولكن من الأفضل استخدام مقياس آخر مثل النسبة مثلاً فنجد فى هذا المثال:

نسبة الحالات أعزب $= \frac{5}{10} \times 100 = 50\%$

نسبة الحالات متزوج $= \frac{1}{10} \times 100 = 10\%$

نسبة الحالات مطلق $= \frac{3}{10} \times 100 = 30\%$

نسبة الحالات أرمل $= \frac{1}{10} \times 100 = 10\%$

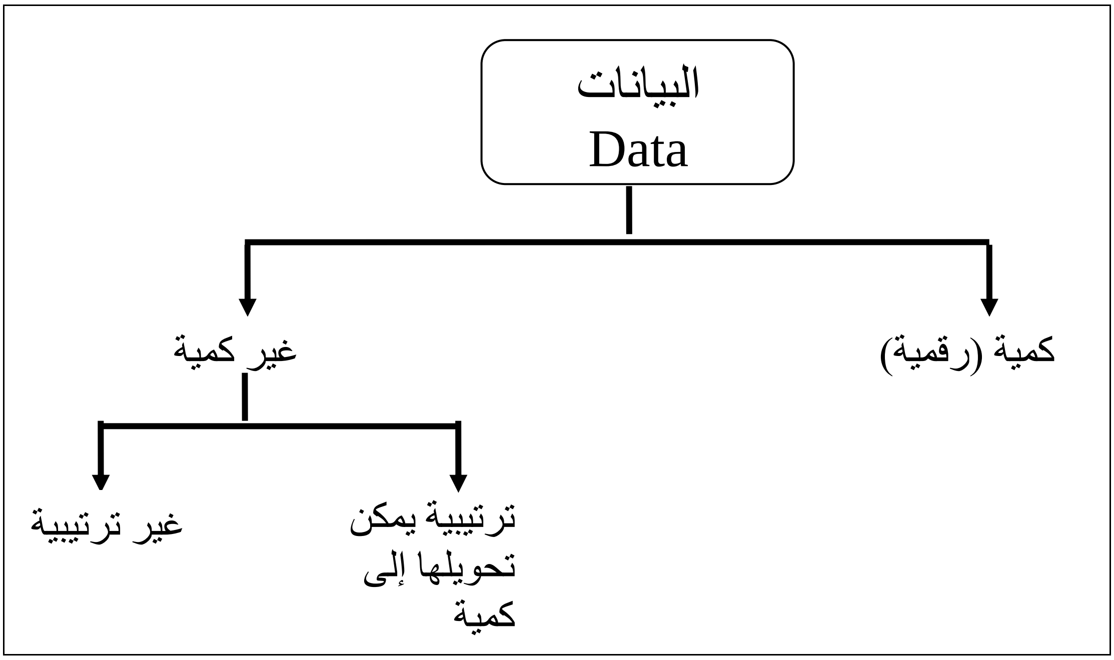
من ثم نجد أن مؤشر النسبة هنا أكفء من المتوسط للتعبير عن الظاهرة محل الدراسة . ومما سبق يمكن التاكيد على ما يلى :-

١- من المثال (٢-١) يتضح أن البيانات غير الكمية الترتيبية يمكن تحويلها إلى بيانات كمية ، ويمكن بالتالى تطبيق الأساليب الإحصائية للبيانات الكمية التى تعتمد على الترتيب مثل الوسيط (وسوف نتحدث عن الوسيط كمقياس فى الباب الثالث).

٢- أما إذا كانت البيانات غير الكمية ليست ترتيبية كما هو فى المثال (٢-٢) فلها الأساليب الإحصائية المناسبة مثل النسبة مثلاً وسوف نوضح ذلك من الناحيتين النظرية و التطبيقية في الأبواب التالية

٣- من المثال (٢-٢) يتضح أنه من الأهمية تحديد الأسلوب الإحصائي المناسب حتى لا نحصل على نتائج مضللة فنجد أن استخدام مقياس المتوسط كان غير مناسب وأعطى نتائج مضللة في حين استخدام أسلوب النسبة كان مناسب في هذه الحالة.

ومن هنا تبرز أهمية تحديد نوع البيانات هل هي كمية أم غير كمية لتحديد الأسلوب الإحصائي المناسب ، والشكل التالي يوضح تصنيف البيانات وهل هذه البيانات غير الكمية ترتيبية أم لا.



شكل (١-٢)

(٢-٢) مصادر البيانات Data Sources

من الباب السابق يتضح أن البيانات هى المادة الرئيسية فى أى دراسة إحصائية فعلى قدر توافر البيانات وشمولها ودقتها تتوقف دقة الدراسة وأهمية النتائج وصحة وفاعلية القرارات المبنية عليها .

وكما ذكرنا سابقاً أن تحديد الهدف من الدراسة يعتبر الخطوة الأولى فى أى دراسة إحصائية ، فبناءاً على هذا الهدف يتم تحديد ما يلى :

- المتغيرات محل الدراسة
- المفردات التى يجب جمع البيانات عنها
- المجتمع الإحصائى (المجتمع محل الدراسة)
- مصادر جمع البيانات
- أسلوب جمع البيانات
- نوع البيانات (كمية أو وصفية) االتى يتم جمعها

أولاً : المتغيرات Variables والمفردات أو الوحدات Items or Units

عـادة تهتم الدراسـة بـالأفراد أو الأشياء وخصائصـهم والعلاقات بينهم . والأفراد أو الأشــياء محــل الدراســة تســمى بــالمفردات أو الوحــدات وتســمى خصـــائص Characteristics هــذه المفــردات بــالمتغيرات Variables نظــراً لإخــتلاف الخصائص من مفردة إلى أخرى وكذلك إختلاف هذه الخصـائص لنفس المفردة من فترة لآخرى أو من مكان لآخر كما سوف نوضح ذلك فى المثال التالى

مثال (٢-٣)

في إحدى المدن الجامعية التابعة لإحدى الجامعات المصرية بمدينة القاهرة اجريت دراسة على 10 من الطلبة والطالبات فى العام الدراسى ٢٠٠٣/٢٠٠٤ للتعرف على المستوى الإجتماعى للطلبة والطالبات المقيمون بالمدينة الجامعية ، والجدول التالى يوضح حالة كل منهم

(1) الأسم	(2) الجنس	(3) الكلية	(4) التقدير	(5) موطنه	(6) عدد السنوات التى قضاها بالمدينة	(7) الدخل الشهرى بالجنية	(8) العمر بالسنوات
أحمد حسن محمد	ذكر	زراعة	إمتياز	قرية	2	150	19
محمود كامل حسين	ذكر	تجارة	مقبول	قرية	3	200	20
فاطمة على حسن	أنثى	طب	جيد	مدينة	4	300	22
إيمان منصور على	أنثى	علوم	جيد جداً	قرية	2	170	18
منصور جرجس لوقا	ذكر	هندسة	جيد	مدينة	3	220	21
عماد كمال أحمد	ذكر	تجارة	مقبول	مدينة	2	130	19
فهمى حسنى أحمد	ذكر	طب	جيد	قرية	3	270	20
نوال على أحمد	أنثى	حقوق	إمتياز	قرية	2	100	19
جمال يوسف حسن	ذكر	صيدلة	جيد	مدينة	3	250	21
نورا فتحى محمود	أنثى	تجارة	جيد جداً	قرية	2	120	20

ملحوظة : جميع الأسماء الواردة بالجدول افتراضية

من الجدول يتضح أن العمود الأول يمثل الطلبة والطالبات وكل واحد أو واحدة منهم يمثـل مفـردة أو وحـدة ، وتمثـل الأعمـدة مـن (2) –(8) خصـائص كـل مفـردة ، فكـل خاصية تمثل متغير من المتغيرات فعلى سبيل المثال يمثل العمود (2) متغير الجنس والعمود (7) يمثل متغير الدخل.

وبالإضافة نجد أن الأعمدة من (2) – (5) تمثل متغيرات معبر عنها بصورة لفظيـة (أى ليست رقمية) مثل الجنس (ذكر – أنثى) ، الكلية (طب – علوم -......) وتسمى المتغيـرات فــى صــورة لفظيــة بــالمتغيرات الوصــفية أو النوعيــة Qualitative Variables . أمــا الأعمــدة مــن (6) – (8) التــى تمثــل المتغيــرات الخاصــة بعــدد

السنوات التى قضاها الطالب أو الطالبة بالمدينة الجامعية ، دخله الشهرى بالجنية ، عمره بالسنوات على الترتيب فإنها تمثل متغيرات معبر عنها بصورة رقمية (أى ليست لفظية) وتسمى متغيرات كمية Quantitative Variables.
مما سبق يتضح أن خصائص المفردات محل الدراسة قد تقاس فى شكل متغيرات وصفية (نوعية) أى فى صورة لفظية ، أو فى شكل متغيرات كمية أى فى صورة رقمية.

ثانياً : المجتمع والعينة Population and Sample

في المثال السابق لدراسة المستوى الإجتماعى للطالب أو الطالبة بإحدى المدن الجامعية أجريت الدراسة على 10 من الطلبة والطالبات بهذه المدينة الجامعة فقط وهؤلاء الطلبة والطالبات يمثلون فئة جزئية Subset من جميع الطلبة والطالبات بهذه المدينة فى العام الدراسى٢٠٠٣/٢٠٠٤ وتسمى هذه الفئة الجزئية التى يتم باستخدامها دراسة المستوى الإجتماعى لجميع الطلبة والطالبات المقيمين بالمدينة فى العام الدراسى ٢٠٠٣/٢٠٠٤ بالعينة Sample . كذلك يسمى جميع الطلبة والطالبات بالمدينة الجامعية فى العام الدراسى ٢٠٠٣/٢٠٠٤ بالمجتمع الإحصائى Statistical Population وتسمى القائمة التى تتضمن جميع مفردات المجتمع الإحصائى بالإطار Frame وبالتالى فإن الإطار يحدد موقع كل مفردة من المفردات داخل المجتمع.

ومما سبق يمكن تعريف المجتمع الإحصائى بأنه جميع المفردات (أو الوحدات) محل الدراسة ، والعينة هى فئة جزئية من المجتمع الإحصائى يتم دراسة خصائص المجتمع عن طريقها . ويتناول الباب السابع من هذا الكتاب بالتفصيل الأنواع المختلفة من العينات وخصائص كل منها.

مصادر البيانات Data Source

يمكن تقسيم مصادر البيانات إلى قسمين :
١- مصادر داخلية Internal Data Sources
٢- مصادر خارجية External Data Sources

والمثال التالي يوضح الفرق بين المصادر الداخلية والخارجية للبيانات

مثال(٢-٤)
إذا أردنا جمع بيانات عن الوضع المالى لإحدى الشركات الصناعية فى ديسمبر ١٩٩٧ من خلال كل من الأرباح السنوية للشركة فى ١٩٩٧ والسنوات السابقة ، عدد العاملين (المهرة وغير المهرة) ، عدد الوحدات المنتجة ،..... الخ
فإنه يمكن أخذ هذه البيانات مباشرة من الشركة وفى هذه الحالة تعتبر الشركة مصدر داخلى للبيانات ، أما إذا قمنا بأخذ هذه البيانات من الجهاز المركزى للمحاسبات أو

أى جهة أخرى غير الشركة ففى هذه الحالة يسمى الجهاز المركزى أو الجهة الأخرى مصدر خارجى للبيانات.

ومما سبق يتضح أن المصادر الداخلية للبيانات هى الجهات التى تقوم بصناعة البيانات ، أما المصادر الخارجية هى الجهات التى يمكن أن نحصل منها على البيانات وفى نفس الوقت ليست الجهات التى قامت بصناعة البيانات.

ومما هو جدير بالذكر أنه توجد مؤسسات حكومية وغير حكومية كثيرة مثل بنوك المعلومات والجهاز المركزى للتعبئة العامة والإحصاء تمثل مصادر خارجية لجمع البيانات . وتحتل هذه المؤسسات مكانة كبيرة فى الدول ذات الإقتصاد الحر مثل الولايات المتحدة الأمريكية ، المملكة المتحدة البريطانية ، فرنسا ،..... الخ وذلك يرجع إلى الأهمية العظمى للبيانات فى الأسواق الحرة التى تقوم على المنافسة.

(٢-٣) أساليب الدراسة الإحصائية

Techniques of Statistical Study

سبق أن عرفنا العينة بأنها فئة جزئية من المجتمع الإحصائى محل الدراسة . وفى كثير من الدراسات يتم دراسة خصائص مفردات المجتمع عن طريق دراسة خصائص مفردات عينة مسحوبة من هذا المجتمع وممثلة له ، ومن ثم فأنه يتم دراسة خصائص مفردات المجتمع محل الدراسة بأحد الأسلوبين التاليين :

١- أسلوب الحصر الشامل Census Survey

٢- أسلوب المعاينة Sampling Survey

والأسلوب الأول (الحصر الشامل) يعنى دراسة خصائص مفردات المجتمع بجمع البيانات عن الخصائص محل الدراسة (المتغيرات محل الدراسة) من جميع مفردات المجتمع محل الدراسة.

أما الأسلوب الثاني (أسلوب المعاينة) يعنى دراسة خصائص مفردات المجتمع بجمع البيانات عن الخصائص محل الدراسة (المتغيرات محل الدراسة) من مفردات العينة المسحوبة من هذا المجتمع ، حيث يتم سحب هذه العينة بحيث تكون ممثلة لجميع مفردات المجتمع . ويتوقف حجم العينة Sample Size (عدد المفردات فى العينة) وطريقة سحبها على نوع الدراسة والهدف منها وأهميتها حيث توجد أساليب وطرق متعددة لتحديد ذلك ، والجزء الثانى من هذا المرجع يتناول الأساليب والطرق المختلفة لتحديد حجم العينة ، وكذلك أسلوب سحب مفردات العينة من المجتمع محل الدراسة.

وفيما يلى سوف نوضح أهم أسباب استخدام أسلوب العينة بدلاً من أسلوب الحصر الشامل فى الدراسة :-

١- تعتبر التكاليف Cost أحد الأسباب الهامة لإستخدام أسلوب العينة بدلاً من أسلوب الحصر الشامل.

فمثلاً إذا كان المطلوب عمل إستطلاع لآراء المواطنين الذين لهم حق الإنتخاب بالجمهورية لأعضاء مجلس الشعب بالنسبة للمرشحين ، فاجراء هذا الاستطلاع يتطلب جمع بيانات من جميع المواطنين الذين لهم حق الإنتخاب فى جميع محافظات الجمهورية مما يتطلب تكاليف باهظة لإجراء هذا الاستطلاع ، ولكن تكون التكلفة ممكنة إذا تم أخذ عينة يمكن باستخدامها تحديد الإتجاهات بالنسبة للمرشحين.

٢- يعتبر الوقت أو السرعة فى الحصول على النتائج من أهم أسباب استخدام أسلوب العينة بدلاً من الحصر الشامل.

فمثلاً فى المثال السابق إذا كان عدد المنتخبين فى جميع المحافظات ثلاثون مليون فرد فإن الزمن المطلوب لإجراء استطلاع آرائهم قد يتطلب عدة شهور ، فى حين إذا تم أخذ عينة مكونة من 500 فرد مثلاً فإن اجراء استطلاع آرائهم قد يتطلب أسبوع على الأكثر . وبالتالى فإن عنصر الوقت أو السرعة يعتبر من أهم أسباب استخدام أسلوب العينة بدلاً من الحصر الشامل.

٣- نظراً لأن العينة فئة جزئية من المجتمع فإن جميع البيانات لمفردات العينة تكون أدق من البيانات التى يتم جمعها من جميع مفردات المجتمع باستخدام أسلوب الحصر الشامل.

٤- فى بعض الحالات لا يمكن استخدام أسلوب الحصر الشامل ويكون الأسلوب الوحيد الممكن استخدامه أسلوب العينة ، ومن هذه الحالات الحالات التالية:

أ) عندما يكون عدد مفردات المجتمع عدد لا نهائى . فمثلاً عند دراسة أنواع معينة من الطفيليات الموجودة بإحدى الترع المصرية ، فنجد أن عدد الطفيليات من هذه الأنواع عدد لا نهائى لذا فعند إجراء هذه الدراسة نأخذ عينة من مياه هذه الترعة أى يتم أخذ عينة من أنواع الطفيليات المراد دراستها.

ب) عندما تفنى مفردة (وحدة) الدراسة عند دراستها . فمثلاً إذا كان الهدف معرفة كفاءة وتأثير وحدات معينة من المتفجرات ، ففى هذه الحالة تفنى الوحدة عند اختيارها لذا فلا يمكن أستخدام أسلوب الحصر الشامل لأنه سوف يؤدى إلى أهلاك مجتمع الدراسة.

كذلك عند إجراء فحص دم الفرد لتحديد نوع الميكروب المسبب للمرض حتى يمكن إعطاء الدواء المناسب فإنه يتم أخذ عينة من دم المريض ولا يمكن أخذ دم المريض كله إلا هلك الشخص.

(٢-٤) أساليب جمع البيانات

Techniques of Collecting Data

بصفة عامة يمكن تصنيف أساليب جمع البيانات إلى أسلوبين أساسيين (سواء كان الأسلوب المستخدم فى الدراسة الحصر الشامل أو أسلوب العينة) وهذان الأسلوبين هما :-

أولاً : أسلوب المسح Survey Technique

إذا تم جمع البيانات من المفردات محل الدراسة بدون إجراء أى تأثير على خصائص المفردات (أى بدون تأثير على قيم المتغيرات محل الدراسة) ، أو بعبارة أخرى بدون إخضاع المفردات لأى عوامل تحكمية Controlled Factors ففى هذه الحالة تسمى طريقة جمع البيانات بالمسح Survey أو بالدراسة المشاهدة Observational Study.

وقد يتم إجراء المسح بالنسبة لجميع مفردات المجتمع (أى عند استخدام أسلوب الحصر الشامل) وفى هذه الحالة يسمى بالمسح الشامل Census (or Complete Survey أو Population Survey ، وكذلك قد يتم إجراء المسح لمفردات العينة (أى عند استخدام أسلوب المعاينة) وفى هذه الحالة يسمى بمسح العينة أو المعاينة Sampling Survey . والمثال التالى يوضح ذلك

مثال (٢-٥)

إذا كان المطلوب معرفة رأى مشاهدى البرامج الإخبارية فى التليفزيون بمدينة القاهرة بالنسبة لبرنامج إخبارى اسبوعى معين.

ففى هذه الحالة يكون الهدف من الدراسة معرفة آراء مشاهدى البرامج الإخبارية بمدينة القاهرة فى هذا البرنامج ، وفى هذه الحالة إذا تم سؤال جميع مشاهدى البرامج الإخبارية فى التليفزيون بمدينة القاهرة عن آرائهم فى هذا البرنامج يكون هذا الأسلوب لجمع البيانات هو أسلوب المسح الشامل ، أما إذا اخذت عينة من هؤلاء المشاهدين وتم سؤال كل مفردة من مفردات العينة عن آرائهم فى هذا البرنامج الإخبارى فإن أسلوب جمع البيانات فى هذه الحالة يسمى بأسلوب مسح العينة Sample Survey أو مسح المعاينة Sampling Survey.

ثانياً : أسلوب التجربة Experimental Technique

مما سبق نجد أنه فى حالة المسح لا تخضع المفردات محل الدراسة لأى عوامل تحكمية من شأنها التأثير على خصائص المفردات ، أما إذا أخضعت المفردات كلها أو بعضها لعامل أو أكثر من العوامل التحكمية التى تؤثر على خصائص المفردات ففى هذه الحالة يسمى هذا النوع من الدراسة بالتجربة . والمثال التالى يوضح هذه الطريقة لجمع البيانات.

مثال (٢-٦)
إذا كان المطلوب معرفة تأثير زيادة جرعة فيتامين ج بنسبة 15% من الجرعة العادية بالنسبة لسرعة القضاء على الميكروب المسبب للإنفلوانزا عند الأطفال فى الفئة العمرية من 7 إلى 10 سنوات.

ففى هذه الحالة يكون الهدف من الدراسة تحديد تأثير نسبة الزيادة 15% من فيتامين ج على سرعة القضاء على الميكروب المسبب للإنفلوانزا للأطفال فى الفئة العمرية من 7 إلى 10 سنوات . وفى هذه الحالة تأخذ عينة من الأطفال المتماثلين تماما فى الفئة العمرية من 7 إلى 10 سنوات من المصابين بميكروب الإنفلوانزر ، ويتم إعطاء 50% منهم الجرعة العادية من فيتامين ج وكذلك يتم إعطاء 50% الأخرى الجرعة الزائدة بنسبة 15% من فيتامين ج فى نفس الوقت ، ويتم تسجيل الفترة التى يتم شفاء كل طفل من أطفال هذه العينة بعدها لتحديد هل يوجد فرق فى مدة شفاء الطفل الذى أخذ الجرعة العادية والطفل الذى أخذ الجرعة الزائدة . ففى هذه الحالة يسمى أسلوب جمع البيانات بأسلوب التجربة.

وكما ذكرنا سابقاً فإنه حالياً توجد العديد من المؤسسات والهيئات الحكومية وغير الحكومية المتخصصة فى جمع البيانات (سواء بأسلوب المسح أو التجربة) وتسويقها عن طريق إصدار نشرات دورية أو كتب بصفة دورية أيضاً للبيانات التى تم جمعها فى معظم القطاعات الاقتصادية ، الاجتماعية ،..... الخ ومن هذه الهيئات فى جمهورية مصر العربية الجهاز المركزى للتعبئة العامة والإحصاء ، البنك المركزى ، ..... الخ وبالتالى فإذا كانت البيانات المطلوب جمعها منشورة فى كتب أو نشرات فإنه يتم جمع البيانات من هذه الكتب أو النشرات مباشرة كمصدر من مصادر البيانات . وكذلك توجد هيئات دولية تقوم بإصدار نشرات للبيانات مثل اليونسكو.

ومما سبق يتضح أن اختيار أسلوب المسح أو أسلوب التجربة (الشامل أو العينة) أو جمع البيانات من مصادر منشورة Published Sources يتوقف على الهدف من الدراسة.

(٢-٥) تصميم استمارة البيانات Design of Questionnaire

مما سبق يتضح أن البيانات هى المادة الرئيسية فى أى دراسة إحصائية ، فعلى قدر توافر البيانات وشمولها ودقتها تتوقف دقة الدراسة وأهمية النتائج وصحة وفاعلية القرارات المبنية عليها . ويتم جمع البيانات من المفردات محل الدراسة فى استمارة تسمى باستمارة جمع البيانات Questionnaire.

أولاً :- تتكون استمارة جمع البيانات من عدد من الأسئلة بحيث توضع الأسئلة المرتبطة ببعضها البعض فى مجموعات بحيث ترتب الأسئلة داخل المجموعة الواحدة بترتيب منطقى وكذلك يترك لكل سؤال مكان خالى كافى لوضع الإجابة فيه. وبالنسبة للأسئلة عن المتغيرات غير الكمية (الوصفية) يفضل غالباً أن يصاحب كل سؤال قائمة بالإجابات الممكنة ليختار منها المبحوث.

فمثلاً السؤال عن تحديد مستوى جودة المنتج فيمكن أن تكون الإجابة:

ٯ ردى ٯ جيد ٯ ممتاز

أو السؤال عن تقدير الطالب فيمكن أن يختار المبحوث أحد الإجابات التالية:

ٯ ضعيف جداً ٯ ضعيف ٯ مقبول ٯ جيد ٯ جيد جداً ٯ امتياز

ثانياً :- تشمل الصفحة الأولى من الاستمارة على ما يلي :

- الجهة القائمة بالبحث
- عنوان البحث
- فترة إجراء البحث
- ما يفيد أن البيانات سرية وتستخدم لأغراض البحث فقط

ثالثاً :- يجب مراعاة الاعتبارات التالية عند وضع الأسئلة

- يوجد تناسق عكسى بين عدد الأسئلة ومعدل استجابة المبحوث للإجابة ، وبالتالى يكون من الضرورى مراعاة عدد الأسئلة المناسب بالاستمارة.
- توضع الأسئلة بتعبيرات بسيطة وواضحة بحيث تكون صياغة السؤال بطريقة سهلة وواضحة ولا تحمل إلا معنى واحد.
- عند تصميم الاستمارة لابد من مراعاة الطريقة التى يتم جمع البيانات بها ، هل سيتم جمع البيانات عن طريق المقابلة الشخصية ، أم عن طريق ارسال استمارة جمع البيانات بالبريد، .....
- يجب تحاشى الأسئلة المحرجة.

ومما هو جدير بالذكر أن تصميم استمارة جمع البيانات يعتمد إلى حد كبير على خبرة المصمم ومدى المامه بالجوانب المختلفة للمشكلة التى يتم جمع البيانات بشأنها.

مثال (٢-٧)

تقوم الشركة (أ) بإنتاج ثلاثه أنواع من المشروبات المثلجة I , II , III تباع فى عبوات نصف لتر . وللتخطيط للإنتاج فى السنة القادمة قامت الشركة بإجراء دراسة ميدانية على عينة من العملاء لمعرفة :

- نوعية العملاء مستهلكى منتجات الشركة ، وتفضيل العميل لأى نوع من المنتجات I , II , III .
- المناطق الأكثر إقبال على المنتجات.

- كذلك تقدير عدد الوحدات المستهلكة أسبوعياً لكل عميل من كل نوع بالإضافة إلى معرفة المناطق الأكثر مبيعات لكل نوع .

فى هذه الحالة ممكن تصميم استمارة جمع البيانات على النحو التالي :
الشركة (أ) للمثلجات

بحث عن تفضيل المستهلك لمنتجات الشركة
خلال الفترة(٢٠٠٥/٦/١ - ٢٠٠٥/١٠/١)

المجموعة الأولى

(١-١) جنس المستهلك : ذكر ☐ أنثى ☐

ضع علامة (✓) أمام الإجابة الصحيحة

(١-٢) تاريخ الميلاد : ....

(١-٣) عنوان السكن : حى ☐ .... مدينة ☐ ....

قرية ☐ .... مركز ☐ .... محافظة ☐ ....

(١-٤) الوظيفة الحالية : ....

(١-٥) عنوان العمل : حى ☐ .... مدينة ☐ ....

قرية ☐ .... مركز ☐ .... محافظة ☐ ....

المجموعة الثانية

(٢-١) هل تتناول أى نوع من منتجات الشركة (أ) : نعم ☐ لا ☐

ضع علامة (✓) أما الإجابة الصحيحة :

(٢-٢) أى نوع تفضل : I ☐ II ☐ III ☐

(٢-٣) عدد الوحدات المستهلكة فى الأسبوع من :

النوع I :
النوع II :
النوع III :

(٢-٦) أمثلة تطبيقية Examples

تطبيق (٢-١)

تقوم إحدى شركات إنتاج السمن الصناعى بإنتاج نوع معين ، وزن العبوة منه 5 كيلو جرام حيث تنتج الشركة 1000 وحدة يومياً . وفى الفترة الأخيرة تلقت إدارة الشركة العديد من شكاوى العملاء من مستهلكى إنتاج الشركة من وجود نقص فى وزن العبوات المطروحة فى السوق ، لذا قامت الشركة بأخذ عينة حجمها 75 عبوة من إنتاج يوم معين لفحصها للتأكد من صحة الشكاوى المقدمة .

المطلوب :

١- أوصف المجتمع الإحصائى للدراسة
٢- أوصف المتغير محل الإهتمام
٣- أوصف عينة الدراسة
٤- أوصف الاستدلال (الاستنتاج) الإحصائى فى هذه الحالة

الحل :

١- المجتمع الإحصائى للدراسة هو جميع الوحدات المنتجة فى اليوم المعين المسحوب منها العينة وعدد مفرداته (حجمه) 1000 وحدة (مفردة أو عبوة) .

٢- بما أن الشركة ترغب فى معرفة وجود نقص فى وزن العبوات عن 5 كيلو جرام بالنسبة لكل عبوة فإن المتغير محل الإهتمام هو وزن العبوة مقاس بالكيلو جرام مقرب لأقرب جرام (أى لأقرب ثلاثة أرقام عشرية) .

٣- عينة البحث حجمها 75 وحدة من إنتاج اليوم محل الدراسة .

٤- دراسة الوزن فى العينة بهدف تعميم النتائج التى يتم التوصل إليها من العينة على المجتمع ، ويمكن أن يتم ذلك من خلال :-

أ) تحديد متوسط وزن العبوة فى العينة وباستخدام المتوسط فى العينة يتم تقدير المتوسط في المجتمع .

ب) باستخدام المتوسط فى العينة يتم إختبار افتراض أن متوسط وزن العبوة فى المجتمع 5 كيلو جرامات .

تطبيق (٢-٢)

في إحدى الأعوام قام التليفزيون بإجراء استطلاع لآراء المشاهدين بالنسبة لتطوير البرامج فى ذلك العام بالنسبة للعام السابق له . فاخذت عينة من مشاهدى البرامج التليفزيونية بمدينة القاهرة فى الفئة العمرية (20-40) حجمها 1000 مشاهد ، وتم سؤال كل منهم عن حدوث تطوير فى البرامج فى ذلك العام أم لا .

المطلوب :

١- عرف مجتمع الدراسة
٢- عرف المتغير محل الدراسة
٣- أوصف عينة الدراسة
٤- اوصف الاستدلال (الاستنتاج) الإحصائى فى هذه الحالة

الحل :

١- مجتمع الدراسة هو جميع مشاهدى البرامج التليفزيونية فى مدينة القاهرة فى الفئة العمرية (20-40) سنة فى هذا العام .

٢- بما أن التليفزيون يرغب فى معرفة رأى المشاهدين فى مدينة القاهرة فى الفئة العمرية (20-40) سنة بالنسبة لحدوث تطور فى هذا العام بالنسبة للبرامج أم لا فإن المتغير محل الإهتمام هو رأى المشاهد ، والرأى هنا متغير وصفى إما أن يكون نعم أم لا .

٣- العينة حجمها 1000 مشاهد فى الفئة العمرية (20-40) سنة بمدينة القاهرة

٤- دراسة رأى المشاهدين فى العينة بهدف تعميم النتائج التى تم التوصل إليها من العينة على المجتمع ، ويمكن أن يتم ذلك من خلال :

أ) تحديد نسبة المؤيدين لحدوث تطور فى العينة وباستخدام هذه النسبة يتم تقدير نسبة المؤيدين لحدوث تطور فى المجتمع الإحصائى

ب) باستخدام هذه النسبة فى العينة يتم إختيار افتراض أن النسبة فى المجتمع الإحصائى للمؤيدين تزيد عن 50%

تطبيق (٢-٣)

عادة يتم إجراء صيانة دورية كل 6 شهور للكبارى العلوية بإحدى الدول ، حيث يتم الفحص من خلال عدة متغيرات تبين مدى كفاءة الكبوبرى . وفيما يلى بعض هذه المتغيرات :

- أقصى اتساع للكوبرى (بالمتر)
- عدد الممرات للمركبات بالكوبرى
- متوسط الكثافة المرورية اليومية
- حالة أرضية الكوبرى (جيد – متوسط – سيئة)
- طول المنعطفات الجانبية للدخول أو الخروج (بالمتر)
- نوع الطريق (داخل حى – يربط بين أكثر من حى – يربط بين محافظتين أو أكثر – يربط بين الدولة ودولة أخرى أو أكثر)

المطلوب

تحديد أى المتغيرات السابقة يعتبر متغير وصفى وأيهم كمى

الحل
المتغيرات (1) ، (2) ، (3) ، (5) متغيرات كمية . والمتغيرات (4) ، (6) متغيرات وصفية

تطبيق (٢-٤)

تقوم إحدى شركات إنتاج المياه الغازية بجمهورية مصر العربية بإنتاج ثلاثة أحجام من العبوات : زجاجة عبوة نصف لتر ، وزجاجة عبوة لنر ، وزجاجة عبوة لتر ونصف . وتقوم الشركة بتسويق منتجاتها عن طريق 20000 مركز توزيع ، حيث يتم وضع كل 12 زجاجة من نفس الحجم في صندوق ثم يتم توزيع الصناديق على مراكز التوزيع . ووجدت إدارة التسويق بالشركة أن الطلب على إنتاج الشركة يتزايد في فصل الصيف وبصفة خاصة في المحافظات الساحلية ، لذا رأت الشركة من الضروري إجراء دراسة إحصائية عن طريق عينة حجمها 2000 مركز توزيع لتقدير حجم المبيعات في الفصول المختلفة عام ١٩٩٩ م وذلك بهدف تقدير حجم المبيعات في الفصول المختلفة لعام ٢٠٠١ م من كل حجم من أحجام العبوات ، وتحديد المدن والقرى (السياحية وغير السياحية) الأكثر استهلاكاً.

المطلوب : حدد كل من
١) مجتمع الدراسة
٢) مفردات البحث
٣) العينة التي يجب استخدامها

الحل
١) مجتمع الدراسة في هذه الحالة هو مجتمع مراكز التوزيع للشركة وعددها 20000 مركز.

٢) مفردات البحث هى مراكز التوزيع للشركة وعددها 2000 مركز.

٣) بما أن الكميات المستهلكة من المياه الغازية في الفصول المختلفة من السنة (صيف – خريف – شتاء - ربيع) يختلف من المحافظات السياحية عن المحافظات غير السياحية بالإضافة إلى اختلاف الكميات في المحافظة الواحدة من فصل إلى أخر ، فمثلاً عادة تكون الكميات المستهلكة في الصيف أكبر من الكميات المستهلكة في الشتاء مثلاً وبالتالي لابد أن تشمل العينة على المراكز بالمحافظات السياحية كذلك أن تحتوى على مراكز التوزيع بالمحافظات غير السياحية.

تطبيق (٢-٥)

أ) تكلم باختصار عن أنواع البيانات – مع ذكر مثال لكل نوع.

ب) ترغب إحدى شركات الاتصال بشبكة المعلومات (الانترنت) بمدينة القاهرة في إجراء دراسة على 100 عميل من الأشخاص المشتركين ، لتحديد خصائص العملاء الأكثر إقبال على الاشتراك في شبكة الانترنت وأسباب ذلك في عام ١٩٩٩.

١- حدد كل من :

- مجتمع الدراسة.
- عينة البحث.
- أسلوب الدراسة المستخدم في هذه الحالة.
- مفردة البحث.

٢- عرف المتغيرات محل الدراسة ونوع كل متغير

٣- صمم استمارة جمع بيانات مناسبة

الحل

أ) يوجد لدينا ثلاث أنواع من البيانات :

١- بيانات كمية أي في شكل رقمي مثل الوزن (بالكيلوجرام) ، الطول (بالسنتيمتر) ، ...... الخ.

٢- وبيانات غير كمية تسمى أحياناً بيانات وصفية مثل الحالة الاجتماعية (أعزب ، متزوج ، مطلق ، ارمل).

٣- توجد بيانات غير كمية أي وصفية مثل التقديرات (ضعيف ، مقبول ، جيد ، ممتاز) يمكن تحويلها إلى بيانات رقمية في صورة رقمية وتسمى بيانات ترتيبيه Ranked Data.

ب) ١-

- مجتمع الدراسة هو جميع عملاء الشركة المشتركين في عام ١٩٩٩م
- عينة البحث هي 1000 عميل من العملاء المشتركين في عام ١٩٩٩م
- أسلوب الدراسة المستخدم في هذه الحالة هو أسلوب العينة
- مفردة البحث هي العميل المشترك في الشركة في عام ١٩٩٩م

٢- المتغيرات محل الدراسة هي :

- الجنس - متغير وصفي (أي غير كمي).
- العمر - متغير كمي (أي رقمي).
- الحالة الاجتماعية - متغير وصفي.
- محل الإقامة - متغير وصفي.
- الحالة الاقتصادية - متغير كمي.

- الحالة التعليمية - متغير وصفي.
- الوظيفة الحالية - متغير وصفي.
- الغرض الأساسي من الاشتراك في الشبكة - متغير وصفي.

٣- بما أن هدف الشركة هو تحديد خصائص العملاء الأكثر إقبال على الاشتراك في الشركة ، وكما ذكرنا سابقاً أنه عند تصميم الاستمارة لابد أولاً من تحديد المتغيرات محل الدراسة وتحديد نوعية كل متغير وهو ما تم إجرائه في (٢) . ويمكن تصميم الاستمارة كما يلي :

<u>المجموعة الأولى</u>

(١-١) الجنس: ☐ **ذكر** ☐ **أنثى**

ضع علامة (✓) أمام الإجابة الصحيحة

(١-٢) تاريخ الميلاد : ....

(١-٣) أخر مؤهل دراسي: ☐ **ابتدائي** ☐ **إعدادي** ☐ **ثانوي** ☐ **جامعي** ☐ **ماجستير** ☐ **دكتوراه**

ضع علامة (✓) أمام الإجابة الصحيحة

(١-٤) محل الإقامة: ☐ **حى** .... ☐ **قسم** .... ☐ **محافظة** ....

<u>المجموعة الثانية</u>

(٢-١) الوظيفة الحالية : ....

(٢-٢) متوسط الدخل الشهري بالجنية : ....

(٢-٣) التكلفة الشهرية الإجمالية للاشتراك (بالجنية) : ....

(٢-٤) الفرض الأساسي من الاشتراك: ☐ **إجراء بحوث ودراسات** ☐ **ممارسة بعض الأنشطة والأعمال** ☐ **المعرفة والثقافة**

ضع علامة (✓) أما الإجابة الصحيحة :

## (٢-٧) تمرينـات	Exercises

(٢-١) إذا اعتبرنا طلاب الفرقة الثانية بكلية التجارة وإدارة الأعمال بجامعة حلوان الذين أدوا امتحان مادة مبادئ الإحصاء في العام الدراسي ٢٠٠٠/ ٢٠٠١– وبافتراض اهتمام الأستاذ القائم بتدريس هذه المادة بمتوسط درجة الطالب (أو الطالبة)

١- أوصف المجتمع الإحصائي محل الدراسة.
٢- عرف المتغير محل الدراسة.
٣- إذا تم تسجيل درجة كل طالب (أو طالبة) فهل هذا يمثل أسلوب الحصر الشامل أم أسلوب المعاينة في الدراسة.
٤- إذا تم سحب عينة تمثل درجات 20 طالب وطالبة ثم تم حساب متوسط الدرجة في العينة .
هل متوسط الدرجة في العينة يساوي متوسط الدرجة لهؤلاء الطلاب بصفة عامة؟

(٢-٢) تنظم إحدى الهيئات التعليمية برنامج لإعداد القادة للأفراد المتقدمين لشغل وظائف قيادية ؛ حيث تقدم متكامل لهؤلاء الأفراد في شكل دورات مدة الدورة ثلاثة شهور ، وتقوم الهيئة بتطوير هذا البرنامج بصفة مستمرة وذلك من خلال أخذ آراء هؤلاء الأفراد في نهاية كل دورة حيث تف\قدم استمارة لكل من أدى الدورة تحتوي على مجموعة من الأسئلة التالية:-

١ - عمر الفرد
٢- مهنته
٣- مدى الاستفادة من الدورة (إفادة كاملة – متوسطة – غير مفيدة).
٤- نسبة الحضور للدورة
٥ - نسبة المعلومات الجديدة التي تم الحصول عليها من متابعة البرنامج

حدد نوع كل متغير (كمي - وصفي) من المتغيرات محل الاهتمام في الاستمارة

(٢-٣) في إحدى المؤسسات وجد عجز شديد في مدرسي المدارس الابتدائية مما أدى إلى الإعلان عن احتياج وزارة التربية والتعليم إلى500 مدرس ومدرسة من خريجي كلية رياض الأطفال ، كلية التربية ، دبلوم معلمين ومعلمات ، شهادات أخرى مع الحصول على أحد الدبلومات في العلوم التربوية حيث تكون الأولوية للمتقدمين من خريجي رياض الأطفال ثم يليهم في الأولوية خريجي كلية التربية ثم خريجي دبلوم المعلمين

والمعلمات ثم الحاصلين على شهادات أخرى، وتقدم لهذه الوظائف 2400 فرد لاختيار 500 فرد منهم

المطلوب:

١- أوصف مجتمع الدراسة.
٢- أوصف المتغير محل الاهتمام.

(٢-٤) قامت محافظة القاهرة بإجراء دراسة عن استخدام الفرد لشبكة الانترنت في مدينة القاهرة بهدف دراسة الفئات العمريه الأكثر استخداماً ، ارتباط الدخل باستخدام الشبكة ، ..... الخ.

المطلوب:

١- تحديد المتغيرات المؤثرة في استخدام الفرد لشبكة الانترنت (مثل مستوى التعليم ، الوظيفة ، السن ،...الخ).
٢- تصميم استمارة بيانات يمكن باستخدامها تحديد:
أ- الفئة العمرية الأكثر استخداماً للشبكة
ب- فئات الدخل الأكثر استخداماً
جـ- مستوى التعليم الأكثر شيوعاً للمستخدمين للشبكة
د- متوسط الدخل للفئة الأكثر استخداماً للشبكة

(٢-٥) تهدف إحدى كليات التجارة في إستطلاع آراء طلابها في مستوى الخدمات التي تقدم للطلاب.

المطلوب :

١- حدد المتغيرات محل الدراسة.
٢- صمم إستمارة لجمع البيانات.

(٢-٦) ترغب إحدى الشركات في إستطلاع آراء العاملين بها ومستوى رضائهم عن وجودهم بالشركة من حيث :-
١- الأجور. ٢- الحوافز. ٣- الأنشطة.

المطلوب :

صمم إستمارة لجمع البيانات.

(٢-٧) تقدم إحدى المكتبات المركزية بأحدى الجامعات بتقديم الخدمات التالية :-
١- الإستعارة ٢- الإطلاع
٣- التصوير ٤- استخدام شبكة الاتصال
٥- عروض للأنشطة الثقافية والتعليمية
فإذا أخذت عينة عشوائية من المترددين على المكتبة .

المطلوب :

أ- حدد المتغيرات محل الدراسة ، وحدد نوع كل منهم.

ب- صمم إستمارة مناسبة لجمع البيانات مع تحديد كل من :
١) مفردات البحث ٢) مجتمع الدراسة

الباب الثالث

عرض البيانات الكمية

The Presentation Of Quantitative Data

(٣-١) التوزيع التكراري البسيط
Simple Frequency Distribution

في البابين السابقين تناولنا بالتفصيل أساليب الدراسة الإحصائية والأساليب المختلفة لجمع البيانات من المفردات محل الدراسة. وتسمى البيانات التي يتم جمعها دون إجراء أي عملية حسابية عليها لعرضها بطريقة منظمة بالبيانات الخام Raw Data

وعادة لا يمكن استخدام البيانات الخام مباشرة في دراسة خصائص الظاهرة (أي دراسة المتغيرات) واستنتاج العلاقات بينهم وفقاً للهدف المطلوب تحقيقه.

لذلك تصبح الخطوة الأولى في أي دراسة هي عرض البيانات الخام بصورة توضح خصائص المتغيرات واتجاه العلاقات بين المتغيرات ، وفقاً للهدف المحدد للدراسة.

والمثال التالي يوضح أنه لا يمكن استخلاص آي معلومات من البيانات الخام إلا عند عرضها بصورة يسهل منها استخلاص المعلومات وبالتالي المعرفة المطلوبة من هذه البيانات

مثال (٣-١)

إذا أخذت عينة بسيطة مكونة من 42 طالباً بالصف الثالث بكلية التجارة بجامعة حلوان وسجلت درجاتهم في مادة الإحصاء التطبيقي فكانت على النحو التالي (النهاية العظمى للدرجة 20)

13.5	18	5.5	18	13	9.5	15
14	15	19	6.5	12	17	8.5
18	2	11	7.5	11.5	13	10
20	12	4	0	3	10	12.5
5	17	10	7	17	12	18
15	6.5	12	17	8.5	5.5	10

وتعتبر البيانات السابقة بيانات خام لا يمكن منها مباشرة تحديد المستويات المختلفة للطلاب أو تحديد نسبة الناجحين أو الراسبين أو أعداد الراسبين....الخ ، وتصبح المشكلة أكثر تعقيداً كلما زاد عدد المفردات بالعينة.

ولكن إذا وضعت درجات الطلاب في جدول يوضح توزيع الطلاب وفقاً لمستوياتهم في المادة على النحو الموضح في الجدول التالي :-

جدول (٣-١)

عدد الطلاب	درجة الطالب	المستوى
4	0 -	ضعيف جداً
10	5 -	ضعيف
14	10 -	مقبول
4	14 -	جيد
4	16 -	جيد جداً
6	18 – 20	امتياز
42	المجموع	

فنجد أن جدول (٣-١) يوضح توزيع 42 طالباً وفقاً لدرجاتهم في مادة الإحصاء ، ومن الجدول نجد أن عدد الطلاب الذين يحصل كل منهم على درجة أقل من 5 (أي مستوى كل منهم ضعيف جداً) يساوي 4 طلاب ، وبالمثل نجد أن عدد الطلاب الذين يحصل كل منهم على 5 درجات فأكثر وأقل من 10 درجات (أي مستوى كل منهم ضعيف) يساوي 10طلاب ، وبالمثل بالنسبة لباقي الطلاب فنجد أن 14 طالباً يحصل كل منهم على 10 درجات إلى أقل من 14 درجة (أي مستوى مقبول)....وهكذا بالمثل بالنسبة لباقي المستويات ، ونجد أن الجدول قسم الطلاب وفقاً لدرجاتهم (مستوياتهم) إلى 6 مجموعات (أو فئات).
ويسمى جدول (٣-١) بجدول التوزيع التكراري البسيط للطلاب في الصف الثالث بكلية التجارة وفقاً لدرجاتهم في مادة الإحصاء.
وعادة يتكون الجدول التكراري Frequency Distribution Table من عمودين (أو صفين) حيث يشتمل العمود الأول على فئات المتغير (أي المستويات المختلفة للمتغير) والعمود الثاني (أو الصف الثاني) يشتمل على التكرارات (عدد الطلاب أي عدد المفردات) المناظرة لكل فئة من فئات المتغير؛ ويسمى جدول التوزيع التكراري السابق بجدول التوزيع التكراري البسيط .

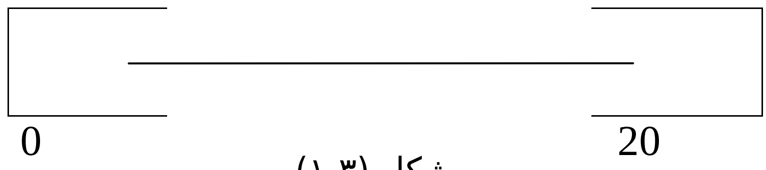
- ويمكن منه اشتقاق كل من :-
  - جدول التوزيع التكراري النسبي
  - جدول التوزيع التكراري التراكمي (الصاعد أو الهابط)

وفي الفصول الثلاثة التالية سوف نتناول بالدراسة كيفية بناء جداول التوزيعات التكرارية السابق ذكرها من البيانات الخام ، وكذلك نتناول طريقة عرض كل منهم بيانياً بالإضافة إلى تناول أهمية كل منهم في التحليل الإحصائي.
مما سبق يتضح أن جدول التوزيع التكراري هو عبارة عن جدول مكون عمودين (أو صفين) العمود الأول (أو الصف الأول) يمثل فئات المتغير محل الدراسة والعمود الثاني (أو الصف الثاني) يمثل عدد المفردات (أو التكرارات)المناظرة لكل فئة أو بعبارة أخرى عدد المفردات التي تأخذ كل منها قيمة تقع داخل فئة المتغير المناظرة لها.

وفيما يلي سوف نقدم بعض التعريفات الضرورية لبناء جدول توزيع تكراري بسيط.

١ـ المدى

هو الفترة التي يأخذ المتغير محل الدراسة بعض أو كل القيمة داخلها.
ففي المثال السابق نجد أن المتغير محل الدراسة هو درجة الطلاب وسوف نشير إليه بالرمز(x) فنجد الحقيقة أن $0 \le x \le 20$ ، وبالتالي فإن المدى في هذه الحالة هو فئة الأعداد من0 إلى 20 كما هو موضح بالشكل التالي :-



شكل (٣-١)

وحسابياً يعتبر المدى هو الفرق بين أكبر قيمة وأصغر قيمة للمتغير أي أن :-

أصغر قيمة - أكبر قيمة = المدى (3-1)

٢ـ حدود الفئة

لكل فئة من الفئات التي يتم تقسيم المتغير إليها حد أدنى Lower Limit (وهو أقل قيمة تقع داخل الفئة) وحد أعلى Upper Limit (وهو أكبر قيمة داخل الفئة)

مثال (٣ـ٢)

الجدولين التاليين يوضحان توزيع 100 طفل وفقاً للوزن (بالكيلو جرام) والعمر بالسنوات

جدول (٣-٢):التوزيع التكراري للأطفال وفقاً لأوزانهم بالكيلو جرام

الوزن بالكيلو جرام(x)	عدد الأطفال (التكرارات)
1 - 4	8
5 – 9	50
10 -14	30
15 - 20	12
المجموع	100

جدول (٣-٣):التوزيع التكراري للأطفال وفقاً لأعمارهم بالسنوات

العمر بالسنوات(x)	عدد الأطفال (التكرارات)
0 -	55
2 -	30
4 -	12
6 -	3
المجموع	100

فإذا نظرنا إلى جدول (٣-٢) نجد أن الفئة الأولى (1 - 4) والتكرار المناظر لها 8 أطفال ، وبالتالي فإن الحد الأدنى للفئة الأولى هو (1) والحد الأعلى للفئة هو (4) وبالمثل لباقي فئات هذا الجدول نجد لكل فئة من فئات الوزن حد أدنى محدد وحد أعلى محدد أيضاَ . لذا يسمى هذا النوع من الفئات بالفئات المغلقة Closed Classes وبالتالي فإن الفئة المغلقة هي الفئة التي لها حد أدنى محدد وحد أعلى محدد أيضاَ ، ونلاحظ أن جميع فئات الجدول (٣-٢) مغلقة بمعنى أن لكل فئة حد أدنى وحد أعلى محددين.

أما بالنسبة لجدول (٣-٣) فنجد أن الثلاثة فئات الأولى في العمود الأول لكل فئة حد أدنى محدد ولا يوجد حد أعلى محدد ، فعلى سبيل المثال الفئة الأولى[2 – 0] تعني أن الحد الأدنى صفر والحد الأعلى غير محدد ولكن لابد أن يكون أقل من 2 حيث 2 تمثل الحد الأدنى للفئة الثانية ، كذلك الفئة الثانية [4 – 2] فنجد أن حدها الأدنى 2 وحدها الأعلى أقل من 4 وهكذا بالنسبة للفئة الثالثة ، أما الفئة الرابعة فحدها الأدنى 6 وحدها الأعلى غير محدد ولذا تسمى بالفئة المفتوحة من أعلى (أي حدها الأعلى غير محدد) ، وقد توجد بعض الفئات (الأولى والأخيرة غالباَ) في الجداول التكرارية البسيطة مفتوحة من أعلى (أي لا يوجد حد أعلى) أو مفتوحة من أدنى (أي لا يوجد حد أدنى)

٣- طول الفئة

بالنسبة للفئة المغلقة فإن طول الفئة (وأحياناً تسمى بفترة الفئة) Class Interval هو الفرق بين الحد الأعلى والحد الأدنى للفئة أي أن:-

الحد الأدنى - الحد الأعلى = طول الفئة (3-2)

أما بالنسبة للفئة المفتوحة من أعلى (أي الحد الأعلى للفئة غير محدد) ففي هذه الحالة يستخدم الحد الأدنى للفئة التالية لهذه الفئة كحد أعلى لها ، أي أن:-

الحد الأدنى للفئة – الحد الأدنى للفئة التالية = طول الفئة (3-3)

أما إذا كانت الفئة مفتوحة من أدنى فإن الحد الأعلى للفئة السابقة لها يستخدم كحد أدنى لها أي أن:-

(3-4) الحد الأعلى للفئة السابقة لها – الحد الأعلى للفئة = طول الفئة

وسوف نوضح ذلك من خلال المثال التالي :

مثال (٣-٣)

الجداول التكرارية التالية توضح توزيع 100 شركة من شركات القطاع الخاص(للأغذية) وفقاً للربح الشهري بالجنيه ، كذلك توزيع 120 شركة من شركات قطاع العام لإنتاج الكيماويات وفقاً للخسارة السنوية بالجنيه

والمطلوب : تحديد طول الفئات لكل توزيع من التوزيعات التالية.

جدول (٣- ٤):التوزيع التكراري ل100 شركة من شركات الأغذية وفقاً للربح الشهري بالألف جنيه

عدد الشركات	ربح الشركة (بالألف جنيه)
3	5 – 10
5	11-
17	15-
25	20-
40	23-
10	25 – 35
100	المجموع

جدول (٣-٥): التوزيع التكراري لـ 120شركة من شركات إنتاج الكيماويات وفقاً للخسارة السنوية بالألف جنيه

عدد الشركات	ربح الشركة (بالألف جنيه)
48	أقل من 3
32	3-
25	4-
10	5-
5	6 فأكثر
120	المجموع

أولاً : بالنسبة للتوزيع التكراري بجدول (٣-٤) نجد أن:
١- الفئة الأولى فئة مغلقة وطولها يساوي 5 (5-10) أنظر المعادلة (٣-٢)
٢- الفئة رقم 5,4,3,2 فئات مفتوحة من أعلى فيستخدم الحد الأدنى لكل فئة تالية كحد أعلى للفئة المراد إيجاد طولها باستخدام المعادلة (٣-٣) كما هو موضح بجدول (٣-٦)
٣- الفئة الأخيرة نجدها فئة مغلقة وطولها يساوي 10 (25-35) ونلاحظ أن أطوال الفئات بالنسبة لهذا التوزيع غير متساوية.

جدول (٣-٦)

مراكز الفئات	أطوال الفئات	التكرارات	الفئات
(10+5) / 2 = 7.5	10 – 5 = 5	3	5 - 10
(11+15) / 2 =13	15 - 14 =4	5	11-
(20+15) / 2 = 17.5	20 – 15 = 5	17	15-
(23+20) / 2 = 21.5	23 - 20 =3	25	20-
(25+23) / 2 = 24	25 – 23 = 2	40	23-
(35+25) / 2 = 30	35 – 25 = 10	10	25 – 35
		100	المجموع

ثانياً : بالنسبة للتوزيع التكراري بجدول(٣- ٥)
١- الفئة الأولى هي فئة مفتوحة من أسفل أي حدها الأدنى غير محدد ويمكن اعتبار الحد الأدنى لها على النحو :
طول الفئة اللاحقة لنفس الفئة – الحد الأعلى للفئة = طول الفئة

أي

2= 1 – 3 = الحد الأدنى

ونظراً لأن طول الفئة الثانية يساوي 1حيث 1= 2-3 كما هو موضح بجدول(٣-٧)
٢ – الفئات رقم 4,3,2 فكل منها فئة مفتوحة من أعلى ويستخدم الحد الأدنى لكل من الفئات الثالثة والرابعة والخامسة كحدود عليا للفئات الثانية والثالثة والرابعة على الترتيب عند حساب طول كل منها كما هو موضح بجدول (٣-٧)
٣ – أما الفئة الأخيرة (6 فأكثر) فهي فئة أخيرة ومفتوحة من أعلى ، ويمكن اعتبار الحد الأعلى لها حيث:
طول الفئة السابقة لها مباشرة + الحد الأدنى = الحد الأعلى
= 6 + 1 = 7

جدول (٣-٧)

الفئات	التكرارات	أطوال الفئات	مراكز الفئات
أقل من 3	48	3 – 2 = 1	(2 + 3) / 2 = 2.5
3 -	32	4 – 3 = 1	(3 + 4) / 2 = 3.5
4 -	25	5 – 4 = 1	(4 + 5) / 2 = 4.5
5 -	10	6 – 5 = 1	(5 + 6) / 2 = 5.5
6 فأكثر	5	7 – 6 = 1	(6 + 7) / 2 = 6.5
المجموع	120		

ملاحظات:

* بالنسبة لفئة الأولى (أقل من 3) فإنه يمكن افتراض أن طولها يساوي طول الفئة اللاحقة لها، وبالنسبة للفئة الأخيرة (6 فأكثر) فإنه يمكن افتراض أن طولها يساوي طول الفئة السابقة لها ؛ وهذه ليست قاعدة فافتراض أن الحد الأعلى للفئة الأخيرة (أو الحد الأدنى للفئة الأولى في بعض الحالات) يتوقف على طبيعة البيانات والهدف من الجدول التكراري ، وبالتالي فإن غلق الفئة الأخيرة من أعلى (أو الفئة الأولى من أسفل) متروك للإحصائي لتحديد ذلك وفقاَ لطبيعة المشكلة والهدف المراد الوصول إليه.
* افترضنا أن طول كل من الفئتين الأولى والأخيرة مساوي لطول الفئة اللاحقة والسابقة نظرا لأن جميع أطوال الفئات لهذا الجدول متساوية.

٤ – مركز الفئة Class Mark Or Midpoint

مركز الفئة هو النقطة التي تتوسط الفئة ويتم حساب مركز الفئة بالنسبة للفئات المغلقة أو المفتوحة من أعلى أو من أسفل على النحو التالي:

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{2} \qquad (3\text{-}5)$$

أو

$$\text{مركز الفئة} = \text{الحد الأدنى للفئة} + \frac{1}{2}\,(\text{طول الفئة}) \qquad (3\text{-}6)$$

أو

$$\text{مركز الفئة} = \text{الحد الأعلى للفئة} - \frac{1}{2}\,(\text{طول الفئة}) \qquad (3\text{-}7)$$

وجدولي (٣-٦)، (٣-٧) يوضحان مراكز الفئات للتوزيعات في المثال السابق.

ولبناء جدول توزيع تكراري بسيط من بيانات خام تتبع الخطوات الآتية :-
١ – تحديد أكبر قيمة وأصغر قيمة ثم تحديد المدى ؛ ويمكن تحديد المدى بافتراض قيمة أقل من أصغر قيمة وافتراض قيمة أكبر كمن أكبر قيمة في البيانات.
٢ – افتراض عدد مناسب من الفئات وفقاً لطبيعة البيانات.
٣ – حساب طول كل فئة باستخدام العلاقة التالية:

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} \qquad (3\text{-}8)$$

وفي هذه الحالة تكون أطوال الفئات بالجداول متساوية ، وهذه ليست قاعدة فقد تتطلب طبيعة البيانات أو الهدف من عرضها تكوين جدول تكراري ذو فئات غير متساوية فهنا يتم تحديد أطوال الفئات من خلال طبيعة البيانات أو الهدف من استخدام الجدول ؛ وسوف يتضح ذلك من خلال الأمثلة والتمرينات التالية:-

٤ - نكون جدول من ثلاثة أعمدة على النحو التالي:
أ- العمود الأول توضع به الفئات.
ب- العمود الثاني يتم فيه تفريغ البيانات الخام (أي حصر المفردات التي قيمة كل منها تقع في الفئة المناظرة لها) كما سوف يتضح من المثال التالي
جـ - العمود الثالث توضع به عدد المفردات (التكرارات)المناظرة لكل فئة

٥ – يأخذ العمود الأول والعمود الثالث معاً فيكونان جدول التوزيع التكراري وتسمى الجداول ذات الفئات المتساوية الأطوال بالجداول التكرارية المنتظمة.

مثال (٣-٤)

فيما يلي بيانات عن قيمة الأرباح في إحدى السنوات لـ 100 شركة من شركات المنتجات الغذائية (بـ 100 ألف جنيه):-

-1	-2.1	4.1	5.5	-1.1	-2.1	2.1	-1	2	2.5	3.1	4.1	6.2
-3	5.4	2.3	3.2	-0.7	2.7	3.4	5	-0.5	0	2.5	2.7	6.3
0.5	1.2	5.1	-2.9	0.5	-0.3	-1.5	2	3	4	3.1	3.9	4.1
1.3	2.7	1.2	-0.4	0	-0.2	0	-2	3.2	7	-1.9	0	6.2
1.4	3.1	1.4	0.5	0.4	1	5.2	0	2.1	-2	-1.1	5	4.1
1.9	1.3	1	4.7	2	6	1.2	3.1	6.1	-1	2.1	4.5	5
-1.3	2.8	0	6.5	5	3	2.1	4.6	7	-1.9	2.8	4.7	7
0.5	5	4	6.3	1	3	2.7	5	4.9				

المطلوب

تكوين جدول توزيع تكراري مناسب يوضح توزيع الشركات حسب قيمة الأرباح.

الحل

نتبع الخطوات السابق ذكرها لتكوين الجدول على النحو التالي:-

١ – بما أن أقل قيمة 3-* وأكبر قيمة 7 إذن:

المدى $=7-(-3)=10$

٢ - نفترض أن عدد الفئات 5 فئات**

٣ – ثم حساب طول الفئة على النحو

طول الفئة $= \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{10}{5} = 2$

٤ – تكوين الجدول التالي ؛ حيث تتم عملية التفريغ Tally Process بالعمود الثاني

جدول (٣-٨) : جدول تفريغ البيانات

الفئات	التفريغ	التكرارات
- (-3)	~~////~~ ~~////~~ //	12
- (-1)	~~////~~ ~~////~~ ~~////~~ ////	19
1-	~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ //	27
3-	~~////~~ ~~////~~ ~~////~~ ~~////~~ //	22
5 – 7	~~////~~ ~~////~~ ~~////~~ ~~////~~	20
المجموع		100

جدول : (٣-٧): جدول توزيع تكراري لـ 100 شركة حسب فئات الربح السنوي (بـ 100 ألف جنيه)

الفئات	التكرارات
- (-3)	12
- (-1)	19
1-	27
3-	22
5 – 7	20
المجموع	100

* القيمة السالبة للربح تمثل الخسارة

** يفضل افتراض عدد مناسب من الفئات بحيث تكون أطوال الفئات أطوال متساوية مناسبة من حيث البيانات بالإضافة إلى سهولة تناولها حسابياً

ملحوظة:

الخطوات السابقة لتكوين جدول التوزيع التكراري البسيط من بيانات خام خطوات عامة ؛ بمعنى انه يمكن تكوين جدول التوزيع التكراري عن طريق عملية التفريغ بحيث تكون :-.

١ – كل أو بعض الفئات مغلقة أو مفتوحة من أعلى أو من أسفل

٢ –أطوال الفئات متساوية أو غير متساوية

مثال(٣-٥)

فيما يلي بيانات عن 30 عامل بأحد المصانع بالكيلو جرام:

70	60	76	63	68	70	95
79	77	88	75	69	82	73
80	85	93	75	72	78	65
82	75	88	89	80	70	82
73	92					

والمطلوب:

تكوين جدول توزيع تكراري ذي فئات بأطوال متساوية لأوزان العاملين بالمصنع.

الحل:-

١ – بما أن أكبر وزن هو95 كيلو جرام وأقل وزن هو 60 فإن المدى يساوي :

المدى = 95 – 60 = 35 كيلو جرام

٢ - نفترض أن عدد الفئات 7 فئات وبالتالي فإن

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{35}{7} = 5$$

٣ – تكوين جدول التفريغ التالي :

جدول (٣-١٠): جدول تفريغ البيانات

الفئات	التفريغ	التكرارات
60-	//	2
65-	///	3
70-	~~////~~ /	6
75-	~~////~~ //	7
80-	~~////~~	5
85-	////	4
90 – 95	//	3
المجموع		100

ومن الجدول السابق نكون جدول التوزيع التكراري على النحو التالي:-

جدول(٣-١١) :جدول التوزيع التكراري لأوزان 30 عامل من العاملين بالمصنع بالكيلو جرام

فئات الوزن بالكيلو جرام(المتغير)	60-	65-	70-	75-	80-	85-	90-95	المجموع
عدد العمال (التكرارات)	2	3	6	7	5	4	3	30

التمثيل البياني للجدول التكراري البسيط
Graphic Representation Of Simple Frequency Table

في هذا الفصل تناولنا سابقاً كيفية بناء (أو تكوين) جدول توزيع تكراري بسيط من بيانات خام تكرار الجدول الظاهر (المتغير) في كل فئة من الفئات للجدول ، وسوف نوضح فيما يلي التمثيل البياني لجدول التوزيع التكراري البسيط حيث يوضح العرض البياني للجدول التكراري توزيع المفردات وفقاً للمتغير محل الدراسة ويعتبر التمثيل البياني وسيلة إيضاح هامة وبصفة خاصة بالنسبة للأفراد غير المتخصصين.

ويمكن تمثيل جدول التوزيع التكراري البسيط بيانياً بأحد الأشكال التالية:-

١. المدرج التكراري Histogram
٢. المضلع التكراري Frequency Polygon
٣. المنحنى التكراري Frequency Curve

أولاً : المدرج التكراري

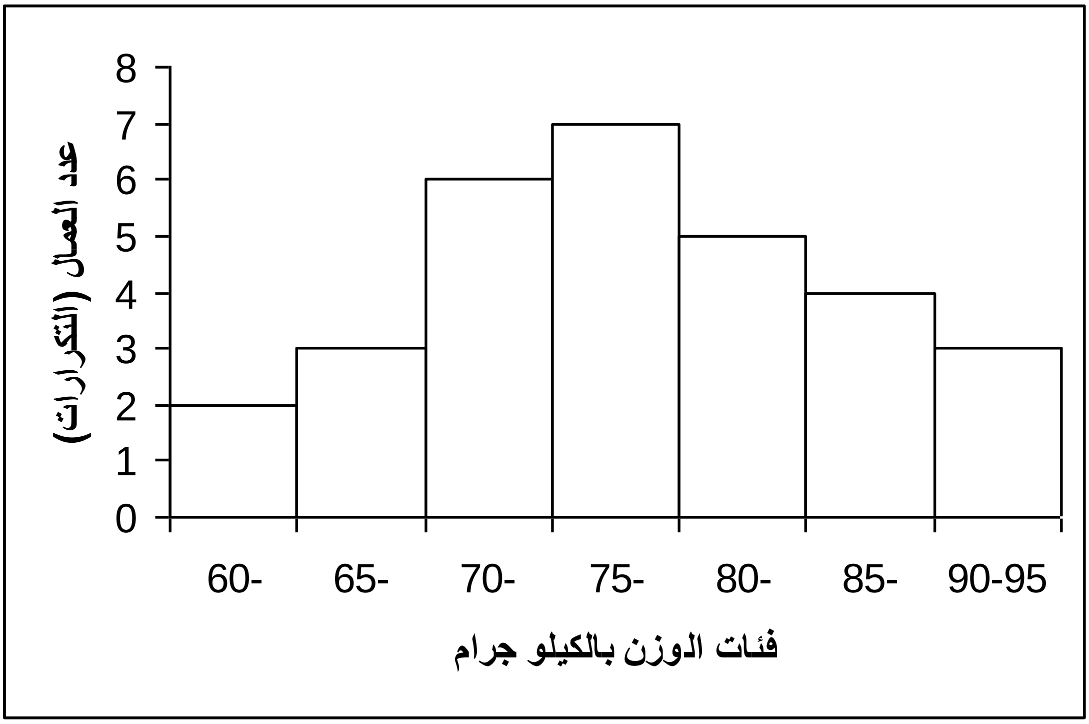
المدرج التكراري هو مجموعة من الأعمدة الملتصقة كل عمود من أعمدة هذه المجموعة تمثل قاعدته فئة وارتفاعه هو التكرار المناظر لهذه الفئة ، حيث تتناسب مساحة المدرج التكراري مع مجموع التكرارات

ويتم رسم المدرج التكراري في حالة التوزيع التكراري ذو الفئات المتساوية الأطوال وذلك بأن تحدد الفئات على المحور الأفقي ويقسم المحور الرأسي ليمثل التكرارات(بحيث يستوعب أكبر تكرار) ثم يقام على المحور الأفقي عند كل فئة

عمود ارتفاعه هو التكرار المناظر لتلك الفئة وقاعدته طول الفئة ، وسوف نوضح ذلك من خلال المثال التالي :-

مثال (٣-٦)

إذا اعتبرنا التوزيع التكراري لأوزان العاملين بالكيلو جرام بجدول(٣-١١) فإن المدرج التكراري الذي يمثل هذا التوزيع كما هو موضح بالشكل التالي :-



شكل (٣-١) : مدرج تكراري يوضح توزيع 30عامل حسب فئات الوزن بالكيلو جرام

ومن الشكل يتضح أن أكبر عدد من العمال يقع في فئة الوزن (80 - 75) كذلك أقل عدد من العمال يقع في فئة الوزن (65 - 60)

أما إذا كانت أطوال الفئات في جدول التوزيع التكراري غير متساوية ويراد تمثيلها بيانيا بمدرج تكراري فلابد أولاً من إيجاد التكرارات المعدلة حيث إن :-

$$\text{التكرار المعدل} = \frac{\text{التكرار الأصلي للفئة}}{\text{طول الفئة}} \qquad (3 - 9)$$

ويمثل التكرارات المعدلة على المحور الرأسي بدلا من التكرارات الأصلية في هذه الحالة وذلك حتى تتناسب مساحة الدرج التكراري مع مجموع التكرارات وسوف نوضح ذلك من خلال المثال التالي :-

مثال (٣-٧)

الجدول التالي يوضح توزيع 1000 عامل حسب فئات الأجر اليومي بالجنيهات على النحو التالي :-

جدول (٣-١٢)

الأجر اليومي بالجنيهات	5-	10-	20-	25-	40-50	المجموع
عدد العمال	125	500	75	195	105	1000

والمطلوب :-

١. أحسب التكرارات المعدلة
٢. ارسم المدرج التكراري الذي يمثل التوزيع التكراري بجدول(٣-١٢)

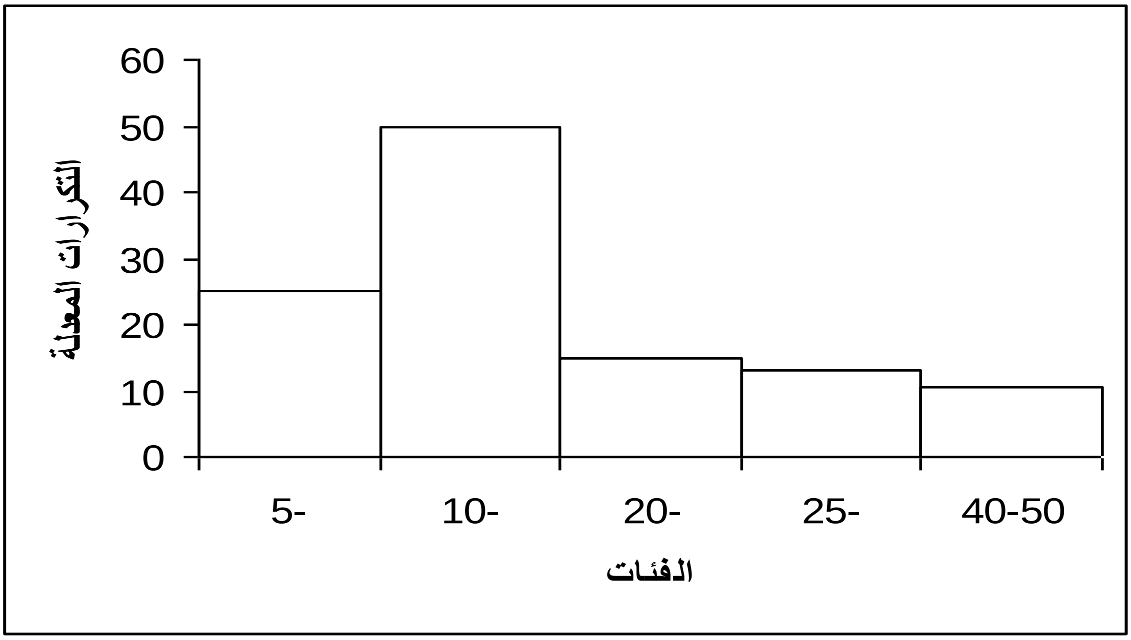
الحل

١ – تحسب التكرارات المعدلة باستخدام المعادلة(٣-٩)

جدول (٣-١٣)

الفئات	التكرارات	أطوال الفئات	التكرارات المعدلة
5-	125	5	125 / 5 = 25
10-	500	10	500 / 10 = 50
20-	75	5	75 / 5 = 15
25-	195	15	195 /15 = 13
40-50	105	10	105 / 10 = 10.5
المجموع	1000		

٢ – نرسم المدرج التكراري ونحدد على المحور الأفقي فئات الأجر اليومي بالجنيهات وعلى المحور الرأسي التكرارات المعدلة كما هو موضح بالشكل التالي



شكل (٣-٢): مدرج تكراري يوضح توزيع1000 حسب فئات الأجر اليومي بالجنيهات

ثانياً : المضلع التكراري

المضلع التكراري هو شكل مضلع قاعدته المحور الأفقي ويتناسب مساحة المضلع التكراري مع مجموعة التكرارات ، ويرسم المضلع التكراري بأن تحد مراكز الفئات على المحور الأفقي وتحدد على المحور الرأسي التكرارات (أو التكرارات المعدلة في حالة أطوال فئات غير متساوية)

مثال (٣-٨)

١ – احسب مراكز الفئات للتوزيع التكراري في الجدول (٣-١١) ثم ارسم المضلع التكراري الذي يوضح التوزيع

٢ – احسب التكرارات المعدلة للتوزيع التكراري بجدول(٣-١٢)ثم ارسم المضلع التكراري الذي يوضح التوزيع

الحل :-

الجدول التالي يوضح مراكز الفئات للتوزيع بجدول (٣-١١)

جدول(٣-١٤)

مراكز الفئات	أطوال الفئات	التكرارات	الفئات
(55+60)/2 = 57.5	5	0	55-
(60+65)/ 2 = 62.5	5	2	60-
(65+70)/ 2 = 67.5	5	3	65-
(70+75)/2 = 72.5	5	6	70-
(75+80)/ 2 = 77.5	5	7	75-
(80+85)2 = 82.5	5	5	80-
(85+90)/ 2 = 87.5	5	4	85-
(90+95)/ 2 = 92.5	5	3	90-
(95+100)/ 2 = 97.5	5	0	95-100
		30	المجموع

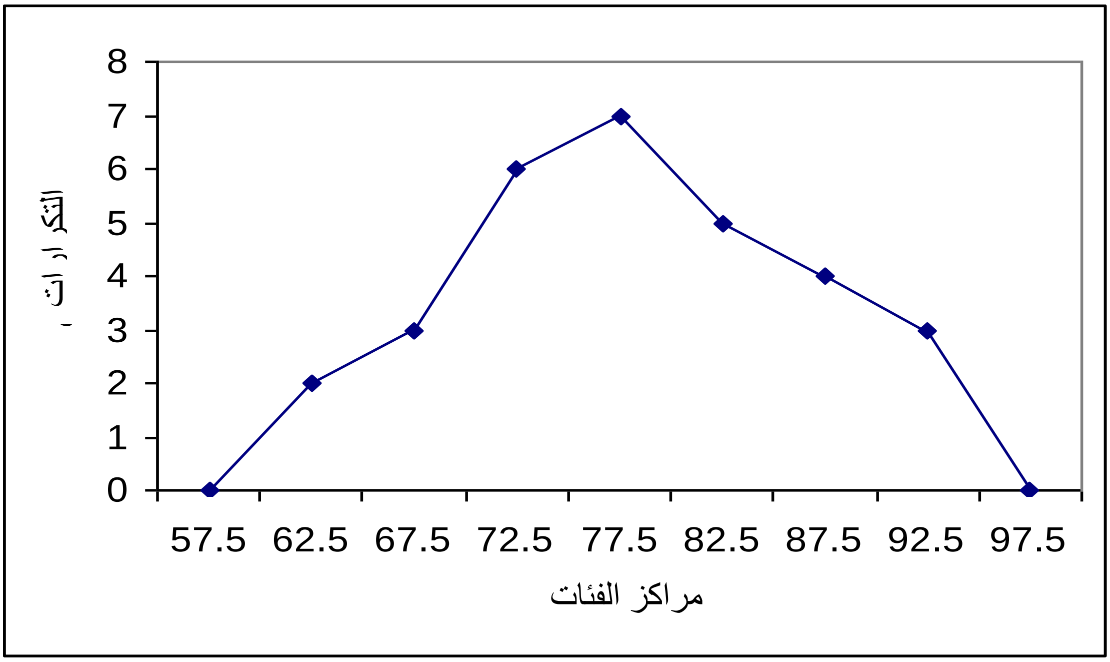
ملحوظة :

تم إضافة الفئتين (55-60) ,(95-100) قبل الفئة الأولى وبعد الفئة الأخيرة على الترتيب ويناظر كل منهما تكرارات تساوي صفر وذلك :

١- حتى تصبح قاعدة المضلع تتناسب مع مجموع التكرارات.

٢- حتى تظل مساحة المضلع تتناسب مع مجموع التكرارات.

والشكل التالي يوضح المضلع التكراري لجدول (٣-١٠)



شكل(٣-٣)

٣- الجدول التالي يوضح أطوال الفئات والتكرارات المعدلة للتوزيع التكراري بجدول (٣-١٢)

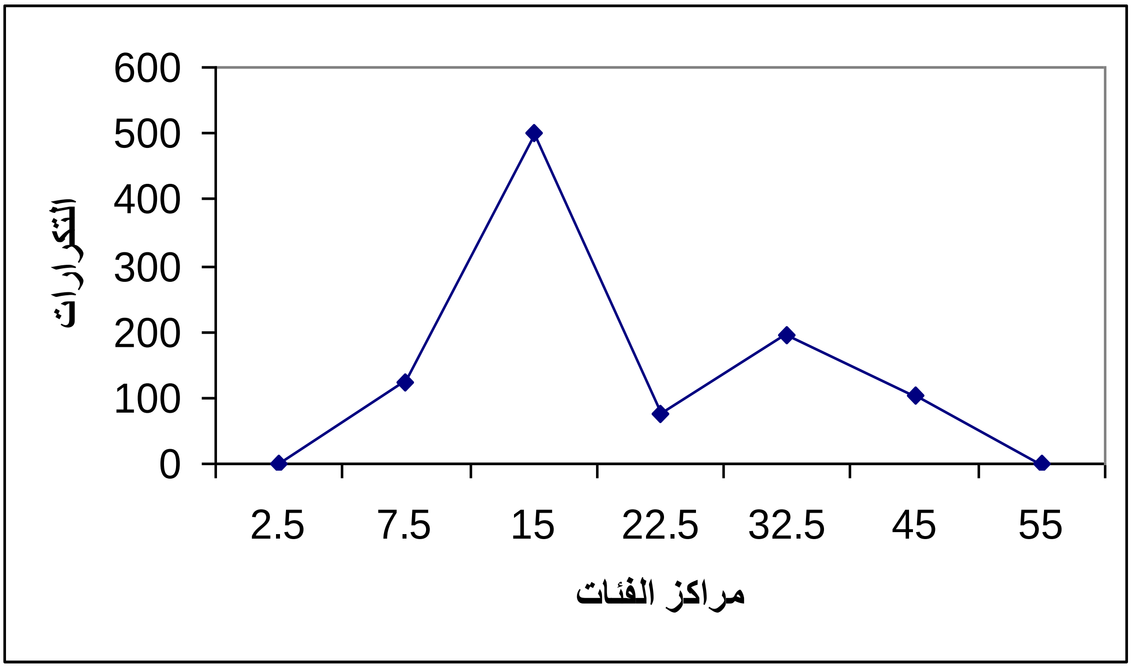
جدول(٣-١٥)

الفئات	التكرارات	أطوال الفئات	التكرارات المعدلة	مراكز الفئات
* 0-	0	5	0	2.5
5-	125	5	25	7.5
10-	500	10	50	15
20-	75	5	15	22.5
25-	195	15	13	32.5
40	105	10	10.5	45
50-60*	0	10	0	55
المجموع	1000		113.5	

* الفئتين [0-5] ، [50-60] تمثلين فئتين أفتراضيتين

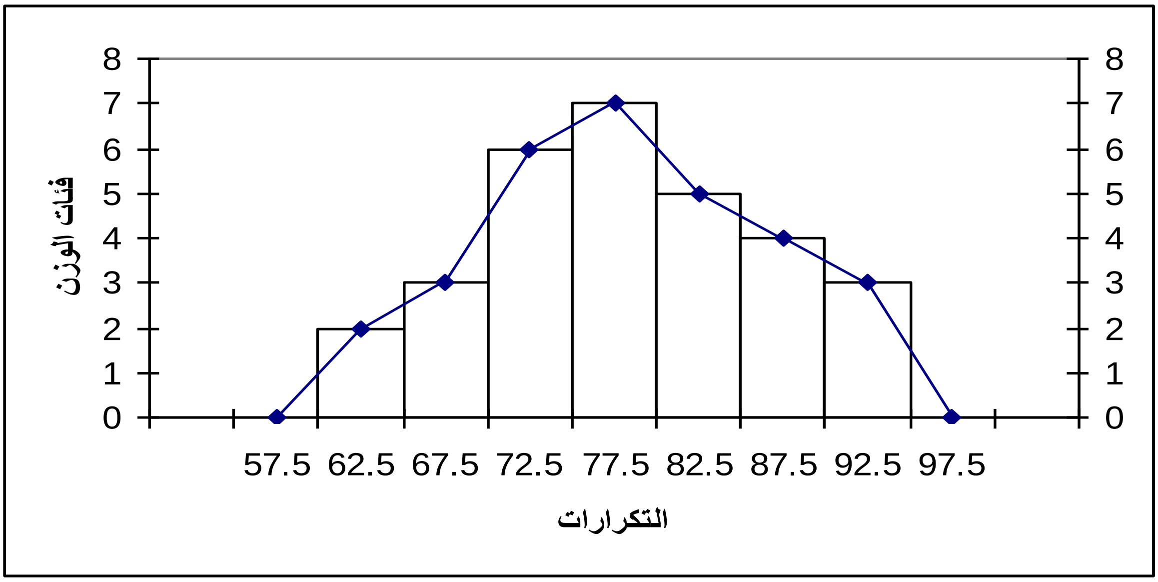
ملحوظة:
بالنسبة للفئتين الإضافيتين(الأولى قبل الفئة الأولى والثانية بعد الفئة الأخيرة) نظراً لأن الفئات غير متساوية الطول فحتى تظل مساحة المضلع تتناسب مع مجموع التكرارات المعدلة ؛ فإن الفئة الإضافية الأولى تسبق الفئة الأولى بنفس طول الفئة الأولى كذلك بالنسبة للفئة الإضافية التي تلحق الفئة الأخيرة يكون لها نفس طول الفئة الأخيرة كما هو موضح بجدول (٣-١٥)

والشكل التالي يوضح المضلع التكراري لجدول التوزيع (٣-١٢):

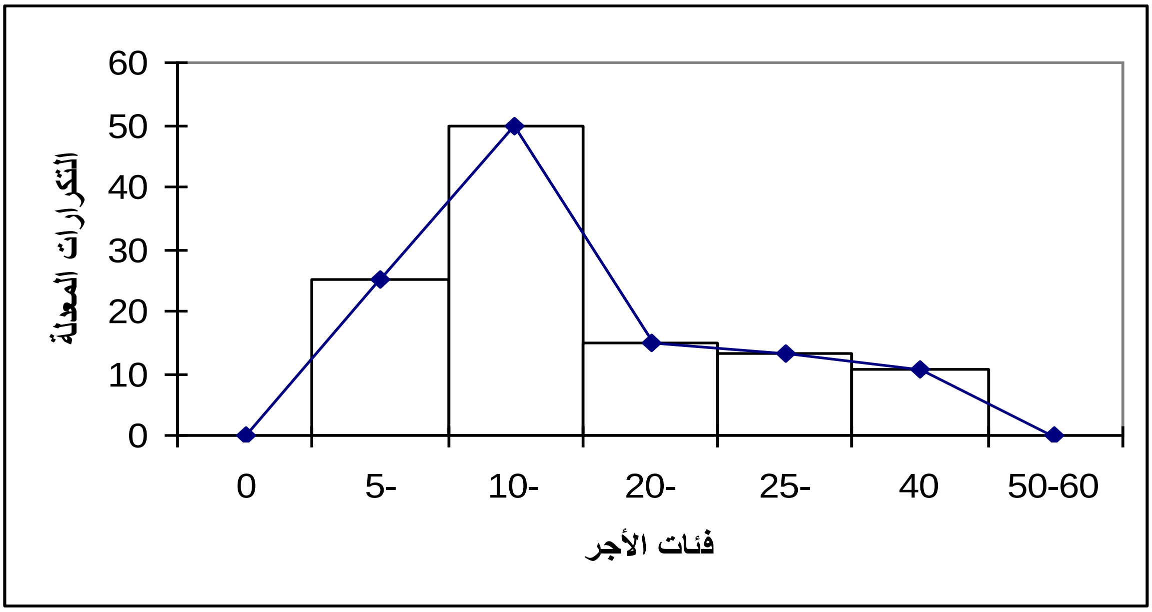


شكل (٣-٤): مضلع تكراري يوضح توزيع 1000عامل حسب فئات الأجر اليومي بالجنيهات

ويمكن رسم المدرج التكراري والمضلع التكراري في شكل واحد كما هو موضح بالشكلين التاليين:
فشكل (٣-٥) يوضح المدرج التكراري والمضلع التكراري في شكلي(٣-١)،(٣-٣) على الترتيب ، كذلك بالمثل يمكن تمثيل المدرج التكراري والمضلع التكراري في شكلي (٣-٢) ،(٣-٤) لجدول التوزيع التكراري (٣-١٢)في شكل واحد كما هو موضح بالشكل(٣-٦)



شكل (٣-٥)



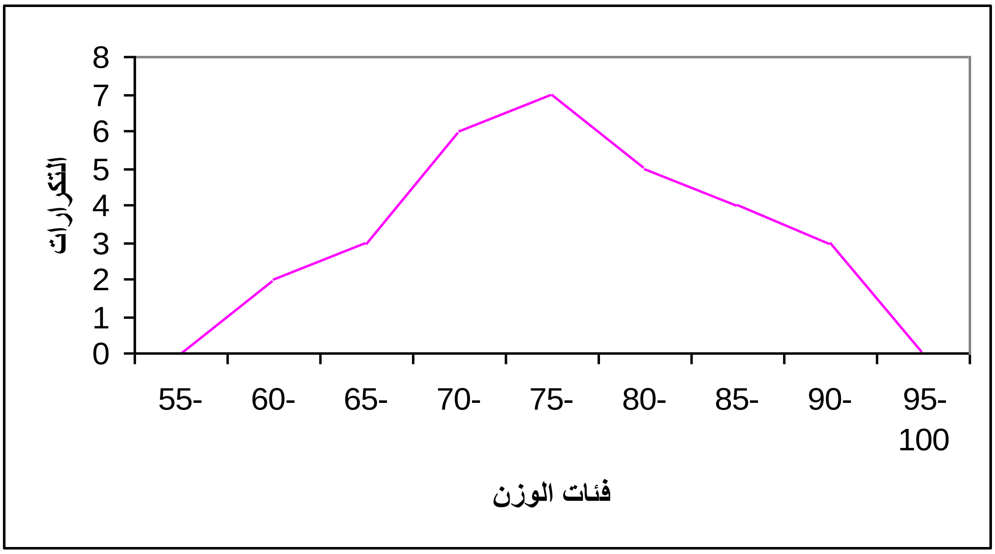
شكل (٣-٦)

ومن شكلي (٣-٥)،(٣-٦) يتضح أن مساحة المدرج التكراري تكافئ مساحة المضلع التكراري ويمكن إثبات ذلك رياضياَ بسهولة.

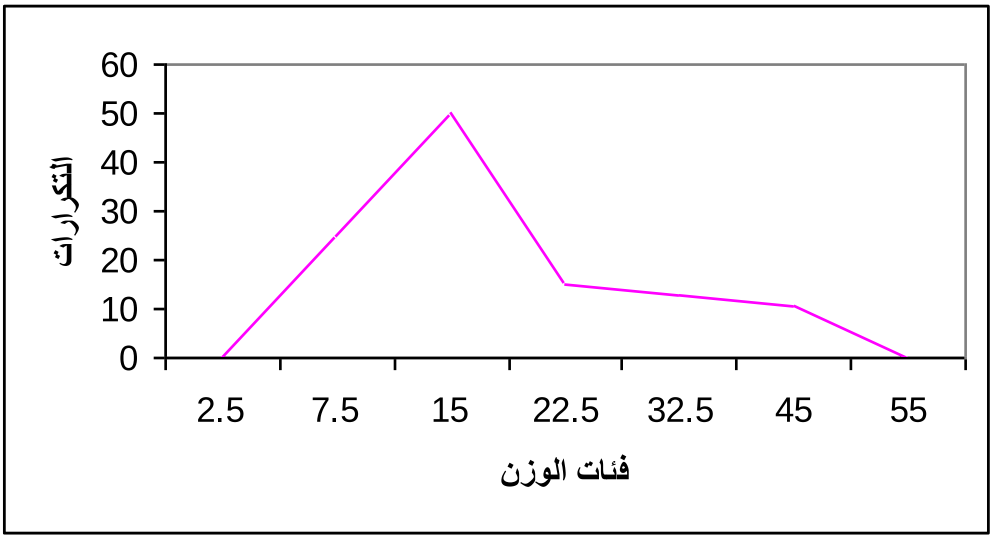
ثالثاَ : المنحنى التكراري

مما سبق اتضح أن المضلع التكراري ينشأ من توصيل النقط التي تمثل التكرارات (أو التكرارات المعدلة) المناظرة لمراكز الفئات بالمسطرة ، ولكن إذا تم التوصيل بين معظم هذه النقط باليد فإنها تعطي منحنى يسمى" المنحنى التكراري" وحيث أن عملية التمهيد تختلف من شخص إلى أخر وبالتالي فالمساحة تحت المنحنى التكراري

تكون تقريبية بمعنى أنها تتناسب تقريباً مع مجموع التكرارات وبالتالي فإن المساحة تحت المنحنى تكون غير متساوية تماماً للمساحة تحت المدرج أو المضلع.
والشكلين التاليين يوضحان المنحنى التكراري لكل من التوزيعين(٣-١١)،(٣-١٢) على التوالي:-



شكل(٣-٧): المنحنى التكراري لتوزيع 30عامل وفقاً لأوزانهم بالكيلو جرام



شكل(٣-٨): المنحنى التكراري لتوزيع 1000عامل وفقاً لأجورهم بالجنيه

(٣-٢) التوزيعات التكرارية التراكمية
Cumulative Frequency Distribution

من التوزيع التكراري البسيط يمكن اشتقاق توزيعات أخرى متعددة توضح خصائص الظاهرة محل الدراسة ، ومن هذه التوزيعات ما يسمى بالتوزيعات التراكمية (أو التجميعية).

والتوزيعات التكرارية التراكمية نوعان هما :-

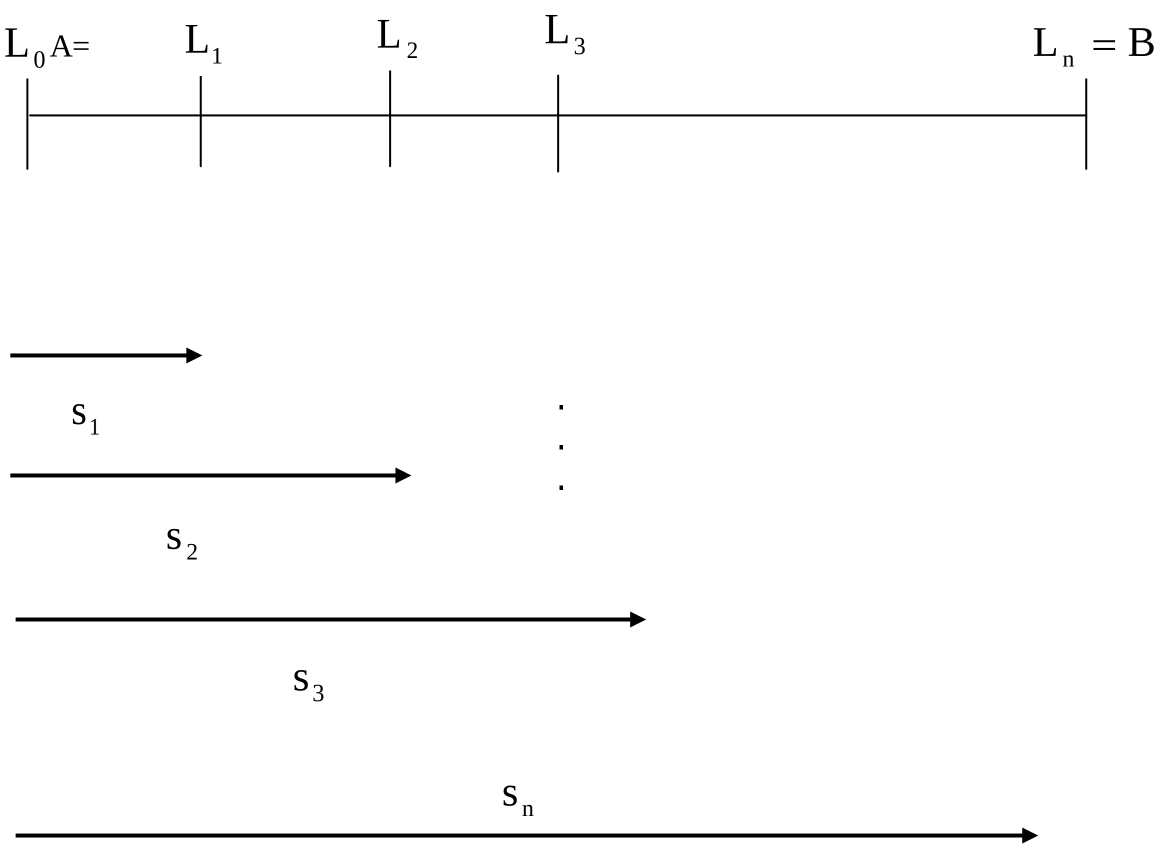
- التوزيع المتجمع الصاعد (المتزايد).
- التوزيع المتجمع الهابط (المتناقص).

أولاَ: التوزيع المتجمع الصاعد Cumulative Increasing Distribution

إذا كان المتغير محل الدراسة x بحيث $A \le x \le B$ وتم تقسيم الفترة [A,B] التي تقع فيها القيم التي يأخذها المتغير إلى فئات جزئية تراكمية (بمعنى أن كل فئة هي فئة جزئية من الفئة التالية لها) $S_1, S_2, \ldots\ldots\ldots, S_n$ بحيث تشير إلى الفئات 1,2,……,n فإذا اعتبرنا

$$S_1 \subset S_2 \subset S_3 \subset \ldots\ldots \subset S_n \qquad (3\text{-}10)$$

فإذا كان $L_1, L_2, \ldots\ldots, L_N$ هي الحدود العليا للفئات السابقة كما هو موضح بالشكل التالي:-



شكل (٢-٩)

حيث L_0 هو الحد الأدنى للفئة الأولى وبالتالي فإن $L_0 = A \;,\; L_n = B$ وتمثل التكرارات $s_1, s_2, \ldots\ldots, s_n$ بالتكرارات التراكمية الصاعدة وسوف نشير لها بالرموز $F_1, F_2, \ldots, F_n$ وتعتبر $F_1, F_2, \ldots, F_n$ تكرارات تصاعدية أي متزايدة حيث:-

$$F_1 \subset F_2 \subset F_3 \subset \ldots \subset F_n \qquad (3\text{-}11)$$

حيث نلاحظ أن F_n هي مجموع التكرارات

والجدول التكراري الذي يحتوي على الحدود $L_1, L_2, \ldots, L_n$ والتكرارات التراكمية التصاعدية المناظرة لها تسمى بجدول التوزيع التكراري التراكمي الصاعد (المتزايد) كما هو موضح بالجدول التالي:

جدول (٣-١٦) : جدول التوزيع المتجمع الصاعد

حدود الفئات العليا	التكرارات الصاعدة
أقل من L_1	0
أقل من L_2	F_1
أقل من L_3	F_2
.	.
.	.
أقل من أو يساوي L_n	F_n

ويلاحظ أن F_n هي عبارة عن مجموع التكرارات

ويمكن تكوين جدول التوزيع المتجمع الصاعد من جدول التوزيع التكراري البسيط على النحو التالي:

١- إيجاد الحدود العليا للفئات بجدول التوزيع التكراري البسيط لتكوين العمود الأول بجدول التوزيع المتجمع الصاعد

٢- إذا فرضنا أن $f_1, f_2, \ldots, f_n$ هي التكرارات المناظرة للفئات في جدول التوزيع التكراري البسيط أن

$$F_r = \sum_{r=1}^{n} f_r \qquad (3\text{-}12)$$

$$F_0 = 0 \qquad , \qquad F_n = \sum_{i=1}^{n} f_i$$

ملحوظة : $\sum$ ترمز لإشارة المجموع "أنظر ملحق (١)"

ويمثل التوزيع المتجمع الصاعد بيانياً عادة بمنحنى يسمى منحنى المتجمع الصاعدCumulative Increasing Curve

والمثال التالي يوضح كيفية بناء توزيع متجمع صاعد باستخدام التوزيع التكراري البسيط ،كذلك يمكن باستخدام المنحنى المتجمع الصاعد للتقدير كما سوف يتضح من المثال التالي:-

مثال(٣-٩)

الجدول التالي يوضح توزيع 1000 شركة حسب قيمة رأس المال لكل شركة (بالألف جنيه)عند بدء نشاط كل منها

جدول (٣-١٧)

فئات رأس المال (بالألف جنيه)	100-	120-	140-	160-	180-200	المجموع
عدد الشركات	50	150	400	300	100	1000

المطلوب

١- كون التوزيع المتجمع الصاعد من جدول التوزيع التكراري السابق

٢- ارسم المنحنى المتجمع الصاعد ومنه قدر عدد الشركات :

أ- التي يقل رأس مال كل منها عن 130 ألف جنيه

ب- عدد الشركات التي ينحصر رأس مال كل منها بين 155-145 ألف جنيه

جـ-عدد الشركات التي يزيد رأس مال كل منها عن 125 ألف جنيه

الحل

١- الجدول التالي يوضح التوزيع المتجمع الصاعد وفقاً لرأس مال الشركات عند بدء نشاط كل شركة

جدول (٣ - ١٨)

الحدود العليا	التكرارات الصاعدة
أقل من 100	0
أقل من 120	50
أقل من 140	50+150=200
أقل من 160	50+150+400=600
أقل من 180	50+150+400+300=900
أقل أو يساوي 200	50+150+400+300+100=1000

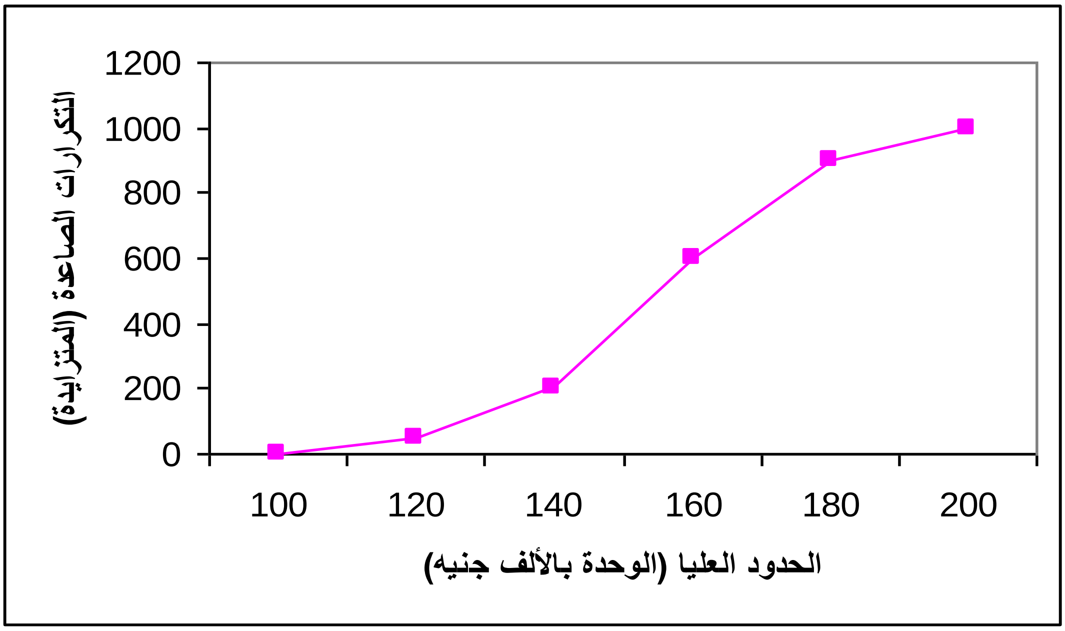
ملاحظات:

١- يبدأ العمود الثاني من الجدول المتجمع الصاعد بالصفر وينتهي بمجموع التكرارات

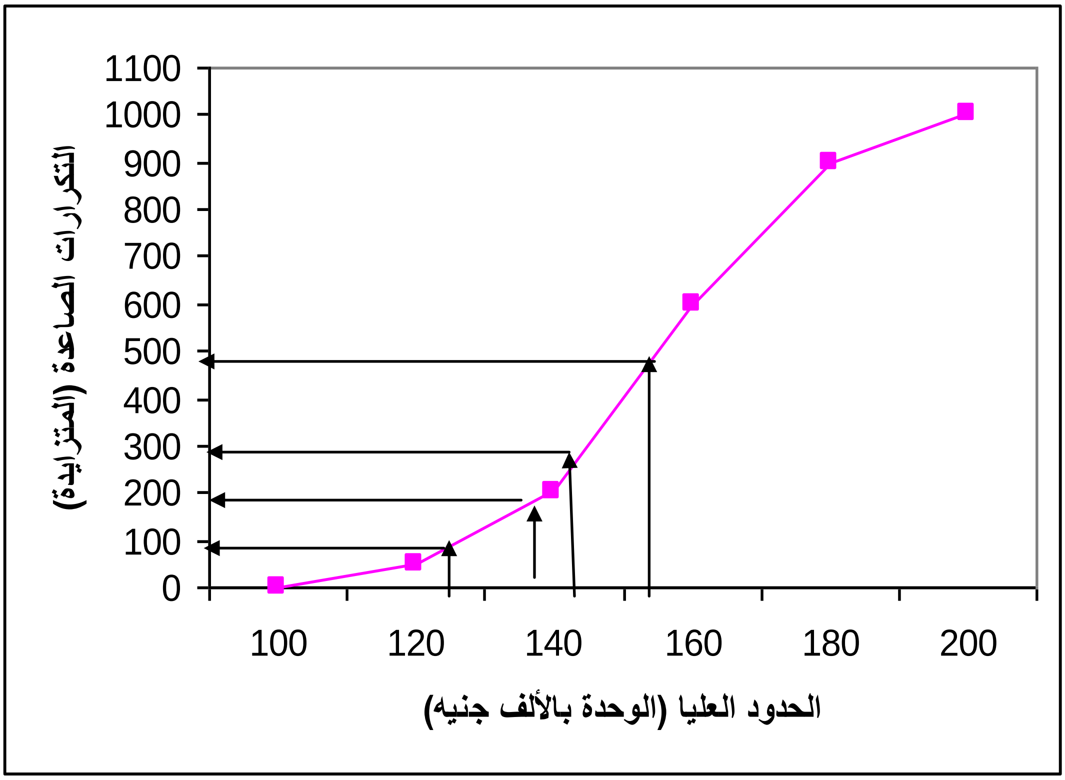
٢- التكرارات في العمود الثاني متزايدة كلما اتجهنا لأسفل

٣- جميع الحدود بالعمود الأول أقل من باستثناء الحد الأخير فهو أقل من أو يساوي

٤- ولرسم المنحنى المتجمع الصاعد تحدد الحدود العليا على المحور الأفقي والتكرارات المتجمعة على المحور الرأسي ثم توصل باليد فتعطي المنحنىcurve أما إذا وصلت بالقلم والمسطرة فإنها تعطي خط منكسر Ogive والشكل التالي يوضح المنحنى المتجمع الصاعد:-



شكل (٣- ١٠): المنحنى المتجمع الصاعد



شكل (٣- ١١

وباستخدام المنحنى المتجمع الصاعد يمكن حساب التقديرات المطلوبة ، ومن الشكل (٣-١١) نجد أن :-

أ- عدد الشركات التي يقل رأس مال كل منها عن 130 ألف جنيه ≈ 125 شركة

ب-عدد الشركات التي ينحصر رأس مال كل منها بين 145-155 ألف جدنيه

شركة $500-300=200 \approx$

جـ- وبالتالي فإن عدد لاشركات التي يزيد رأس مال كل منها عن125 ألف جنيه

شركة $1000-88=912 \approx$

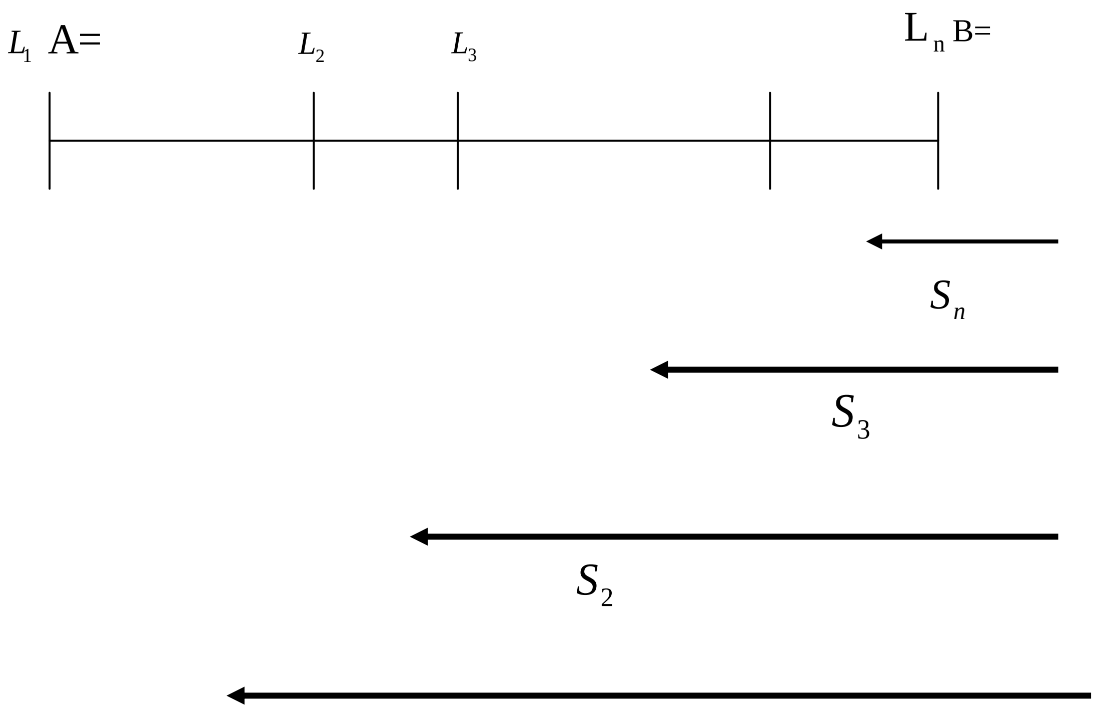
ثانياَ : التوزيع المتجمع الهابط(النازل)

Cumulative Decreasing Distribution

التوزيع المتجمع الهابط هو التوزيع العكسي للتوزيع المتجمع الصاعد ، فإذا كان المتغير محل الدراسة x بحيث$A \le x \le B$ وتم تقسيم الفترة [A,B] التي يقع فيها القيم التي يأخذها المتغير x إلى فئات جزئية تراكمية بحيث :

$$S_n \subset S_{n-1} \subset S_{n-2} \subset .....S_2 \subset S_1$$

فإذا كان $L_1, L_2, ..., L_n$ هي الحدود الدنيا للفئات $S_1, S_2, ..., S_n$ كما هو موضح بالشكل التالي:-



شكل(٣-١٢)

حيث $L_1 = A, L_n = B$.

وتمثل التكرارات المناظرة للفئات $S_1 \quad S_2,....S_n$ كرارات التراكمية الهابطة فإذا أشرنا لها بالرموز $^{/}F_1, ^{/}F_2,....., ^{/}F_n$ حيث:

$$^{/}F_n \subset ^{/}F_{n-1} \subset .....\subset ^{/}F_2 \subset ^{/}F_1$$

والجدول التكراري الذي يحتوي على الحدود الدنيا $L_1, L_2,......, L_n$ والتكرارات التراكمية الهابطة المناظرة لها يسمى بجدول التوزيع التكراري التراكمي الهابط كما هو موضح بالجدول التالي:

جدول (٣-١٩): جدول التوزيع المتجمع الهابط

التكرارات الهابطة	الحدود الدنيا
$^{/}F_1$	أكبر من أو يساوي L_1
$^{/}F_2$	أكبر من أو يساوي L_2
$^{/}F_3$	أكبر من أو يساوي L_3
.	.
.	.
$^{/}F_{n-1}$	أكبر من أو يساوي L_{n-1}
$^{/}F_n$	أكبر من L_n

ويمكن أيضاً تكوين جدول التوزيع المتجمع الهابط من جدول التوزيع التكراري على النحو التالي:

١- نأخذ الحدود الدنيا بجدول التوزيع التكراري البسيط لتكوين العمود الأول بجدول التوزيع المتجمع الهابط

٢- يتم حساب التكرارات المتجمع الهابط باستخدام المعادلة الآتية:-

$$^{/}F_r = \sum_{i=1}^{n} F_i - \sum_{i=1}^{n-1} F_i \qquad ,r=1,2,.....,n \qquad (3\text{-}15)$$

ويمثل التوزيع المتجمع الهابط بمنحنى يسمى بالمنحنى المتجمع الهابط ويمكن استخدام المنحنى المتجمع الهابط في التقدير كما هو في حالة استخدام المنحنى المتجمع الصاعد

وسوف نوضح ذلك من خلال المثال التالي:-

مثال(٣-١٠) : اعتبر مثال (٣-٩)

١- كون التوزيع المتجمع الهابط من جدول التوزيع التكراري (٣-١٧)

٢- ارسم المنحنى المتجمع الهابط ومنه:

أ- قدر عدد الشركات التي يزيد رأس مال كل منها عن145 ألف جنيه.

ب-قدر عدد الشركات التي يقل رأس مال كل منها عن185 ألف جنيه.

الحل

١- جدول (٣-٢٠): الجدول التالي يوضح التوزيع المتجمع الهابط

الحدود الدنيا	التكرارات الهابطة
أكبر من أو يساوي 100	1000
أكبر من أو يساوي 120	1000 – 50 = 950
أكبر من أو يساوي 140	1000 – (50+150) = 800
أكبر من أو يساوي 160	1000 – (50+150+400) = 400
أكبر من أو يساوي 180	1000 – (50+150+400+300) = 100
أكبر من 200	1000-(50+150+400+300+100) = 0

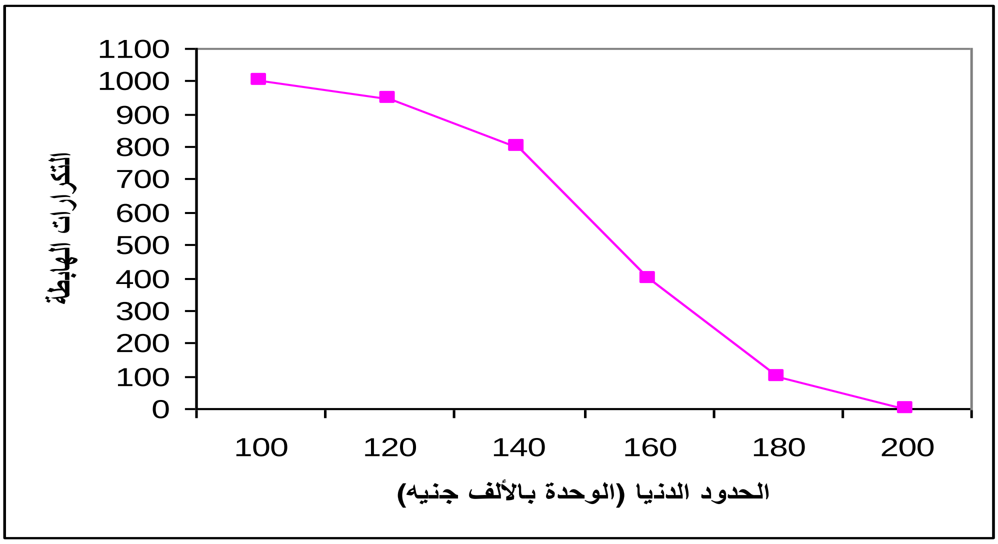
ملاحظات:

أ- يبدأ العمود الثاني في الجدول السابق بمجموع التكرارات (1000) وينتهي بالصفر

ب-التكرارات في العمود الثاني تكرارات متناقصة (هابطة)

جـ- جميع الحدود في العمود الأول أكبر من أو يساوي باستثناء الحد الأخير فهو أكبر مكن فقط

٣- ولرسم المنحنى المتجمع الهابط تحدد الحدود الدنيا على المحور الأفقي والتكرارات الهابطة على المحور الرأسي كما هو موضح بالشكل التالي:



شكل (٣-١٣)

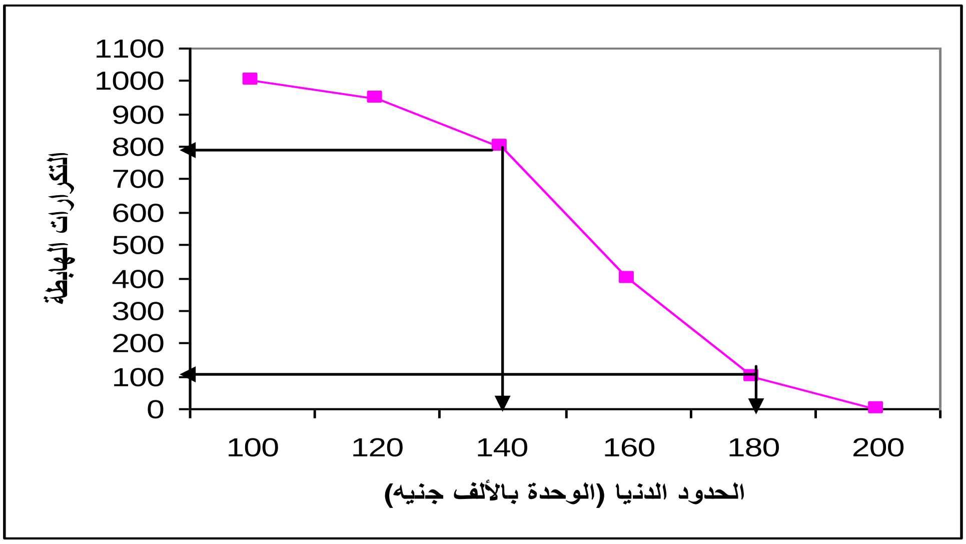
:المنحنى المتجمع الهابط هوا لمنحنى العكسي للمنحنى المتجمع الصاعد.

وباستخدام المنحنى المتجمع الهابط يمكن حساب التقديرات المطلوبة ، فمن شكل (٣-١٤) نجد أن:

أ- عدد الشركات التي يزيد رأس مال كل منها عن 145 ألف جنيه ≅ 700شركة

ب- بما أن عدد الشـركات التـي يزيـد رأس مـال كل منها عـن 185ألـف جنيـه ≅ 75شركة وبالتالي يكون عدد الشركات التي يقل رأس مال كل منهم عن185 ألف جنيه

$\cong 1000 - 75 = 92$ شركة



شكل (٢-١٤)

(٣-٣) التوظزيعات التكرارية النسبية
Relative Frequency Distributions

من الجدول التكراري البسيط يمكن تكوين التوزيع التكراري النسبي ، والتوزيع التكراري النسبي عبارة عن جدول مكون من عمودين ، العمود الأول يمثل الفئات والعمود الثاني يمثل نسبة التكرارات داخل كل فئة بالنسبة لمجموع التكرارات حيث :

$$\text{التكرار النسبي المناظر للفئة} = \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}} \times 100 \qquad (3.16)$$

ويمثل التوزيع التكراري النسبي بيانياً بمدرج أو مضلع أو منحنى تكراري نسبى وسوف نوضح ذلك من خلال المثال التالي

مثال (٣-٩)

الجدول التالي يوضح توزيع 200 فرد حسب فئات العمر بالسنوات

فئات العمر بالسنوات	7-	10-	13-	16-	19-	22-25	المجموع
عدد الأفراد	14	46	54	36	30	20	200

المطلوب

١- أوجد التوزيع التكراري النسبي

٢- أرسم المدرج والمضلع التكراري النسبي

الحل

١- نحسب التكرارات النسبية المناظرة لكل فئة على النحو التالي:

التكرار النسبي المناظر للفئة الأولى $= \frac{14}{200} \times 100 = 7\%$

التكرار النسبي المناظر للفئة الثانية $= \frac{46}{200} \times 100 = 23\%$

التكرار النسبي المناظر للفئة الثالثة $= \frac{54}{200} \times 100 = 27\%$

التكرار النسبي المناظر للفئة الرابعة $= \frac{36}{200} \times 100 = 18\%$

التكرار النسبي المناظر للفئة الخامسة $= \frac{30}{200} \times 100 = 15\%$

التكرار النسبي المناظر للفئة السادسة $= \frac{20}{200} \times 100 = 10\%$

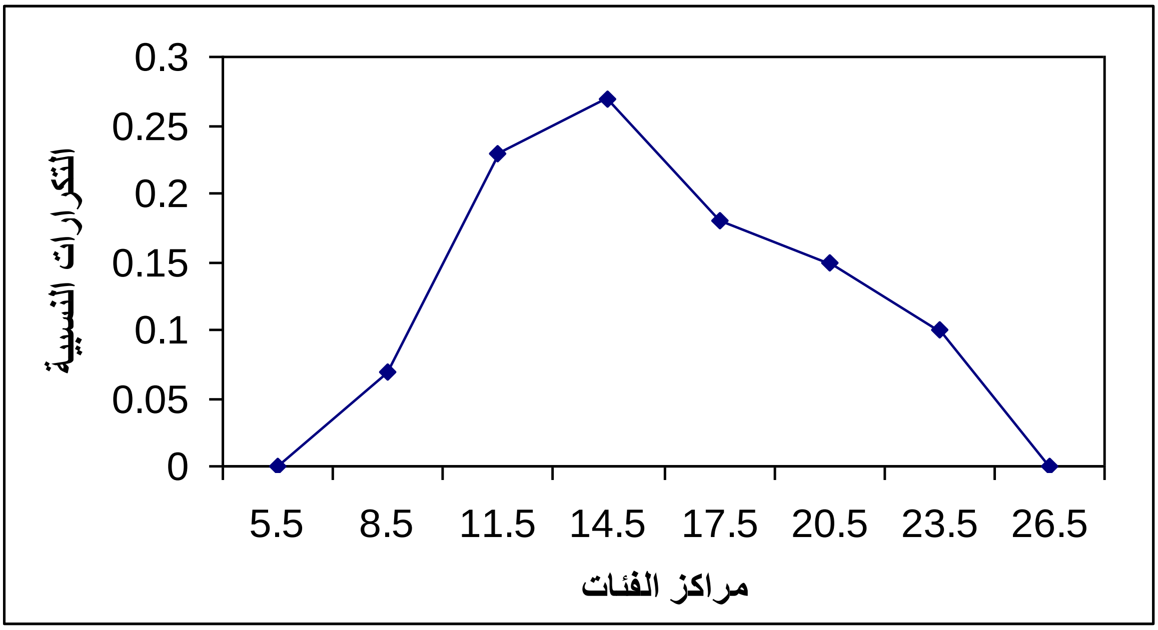
والجدول التالي يعطى التوزيع التكراري النسبي لفئات العمر بالسنوات

جدول (٣-٢١)

المجموع	22-25	19-	16-	13-	10-	7-	الفئات
100%	10%	15%	18%	27%	23%	7%	التكرارات النسبية

ومن الجدول نجد أن أكبر نسبة للأفراد تقع في فئة العمر [13-16) ثم تليها الفئة [10-13) .

٢- ويمكن تمثيل جدول التوزيع النسبي السابق بمدرج أو مضلع تكراري نسبي على النحو الموضح بالشكل التالي



شكل (٣-١٥)

ملحوظة : عندما تكون أطوال الفئات غير متساوية فإنه يجب حساب التكرارات النسبية المعدلة لرسم كل من المضلع أو المنحنى التكراري النسبي حيث

$$\text{التكرار النسبي المعدل للفئة} = \frac{\text{تكرار الفئة}}{\text{طول الفئة} \times \text{مجموع التكرارات}}$$

المقارنة بين توزيعين تكراريين

عند مقارنة توزيعين تكراريين لابد وأن تكون مجموع التكرارات في التوزيعين متساوية لأنه في حالة عدم تساوي مجموع التكرارات في التوزيعين تكون المساحتين تحت المنحنيين التكراريين أو المضلعين التكراريين غير متساوية وللتغلب على ذلك نستخدم التوزيعات التكرارية النسبية في المقارنة بين التوزيعات التكرارية حيث أن مجموع التكرارات النسبية في كل توزيع تكون متساوية (تساوي100%) كذلك نستخدم المضلعات التكرارية النسبية أو المنحنيات التكرارية النسبية لتوضيح المقارنة بيانياً بين التوزيعات.

والمثال التالي سوف يوضح ذلك.

مثال (٣- ١٠):

الجدول التالي يوضح العاملين في مصنعين مختلفين وفقاً للأجر الشهري بالجنيه للعامل:

جدول (٣-٢٢)

فئات الأجر الشهري بالجنيه	80-	100-	120-	140-	160-180	المجموع
عدد العمال في المصنع الأول	15	25	150	60	50	300
عدد العمال في المصنع الثاني	100	120	140	80	60	500

المطلوب

قارن بين التوزيعين ووضح ذلك بيانياً.

الحل

من الجدول السابق يتضح أن مجموع التكرارات في التوزيع الأول يساوي 300 عامل وفي التوزيع الثاني يساوي 500 عامل أي أن مجموع التكرارات في التوزيعين مختلفين ، لذلك سوف توجد التوزيعات التكرارية النسبية لتوحيد مجموع التكرارات في التوزيعين (100%)

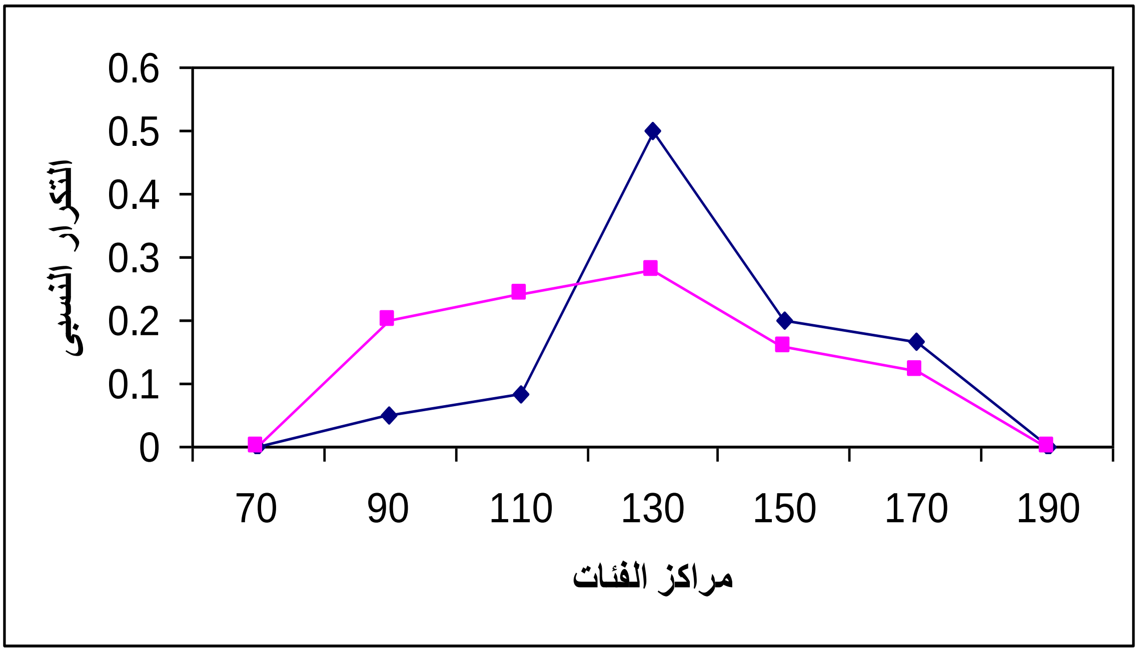
والجدول التالي يوضح التوزيعات التكرارية النسبي

جدول (٣-٢٣)

التوزيع الأول		التوزيع الثاني		الفئات
التكرار الأصلي	التكرار النسبي	التكرار الأصلي	التكرار النسبي	
15	5%	100	20%	80-
25	8.33%	120	24%	100-
150	50%	140	28%	120-
60	20%	80	16%	140-
50	16.67%	60	12%	160-180
300	100%	500	100%	المجموع

ومن الجدول يتضح أن أكبر نسبة من العاملين في المصنعين تقع في فئة الأجر[120-140] ولكن نسبة العاملين في هذه الفئة في المصنع الأول (50%) أكبر من نسبة العاملين في هذه الفئة في المصنع الثاني (28%) ، وتعتبر أقل نسبة من العاملين في المصنع الأول (5%) تقع في فئة الأجر [80-100] في حين أن أقل نسبة عاملين في المصنع الثاني (12%) تقع في فئة الأجر [160-180]

والشكل التالي يوضح الفرق بين التوزيعين وفقاً لفئات الأجر الشهري بالجنيه



شكل (٣-١٦)

(٣-٤)التوزيع التكراري المزدوج
Double Frequency Distribution

في دراسة المشاكل الفعلية نجد أن كثير من الظواهر (أو المتغيرات) مرتبطة ببعضها البعض ، فعلى سبيل المثال عادة ترتبط أوزان الأفراد بأطوالهم وهنا نجد أن الطول يمثل المتغير الأساسي والوزن يمثل المتغير التابع له ، كذلك نجد أن عدد الأطفال في الأسرة الواحدة مرتبطة بمستوى تعليم الأم والأب معاَ ودخل الأسرة ،...:الخ . وبالتالي يصبح من الأهمية دراسة العلاقة بين هذه المتغيرات .

وتعتبر الخطوة الأولى في دراسة العلاقة بين متغيرين أو ظاهرتين هي عرض البيانات بأسلوب يوضح توزيع المفردات محل الدراسة وفقاّ للمستويات المختلفة للمتغيرين معاَ ويعتبر الجدول التكراري المزدوج أحد أساليب عرض البيانات بحيث يمكن باستخدامه دراسة العلاقة بين المتغيرين بالإضافة إلى أنه يمكن باستخدام هذا الجدول مقارنة توزيع المفردات وفقاَ للمتغيرين بسهولة

وفيما يلي سوف نوضح خطوات بناء جدول التوزيع التكراري المزدوج من خلال المثال التالي .

مثال(٣-١١)

فيما يلي بيانات عن درجات 24 طالب وطالبة في مادتي الرياضة والإحصاء بالصف الثاني بكلية التجارة وإدارة الأعمال في إحدى السنوات الدراسية.

جدول (٣-٢٤)

رقم الطالب	درجة الرياضة	درجة الإحصاء	رقم الطالب	درجة الرياضة	درجة الإحصاء
1	10	9	13	4	6
2	20	18	14	2	5
3	19	15	15	0	2
4	0	3	16	5	7
5	5	2	17	10	12
6	10	7	18	13	15
7	11	10	19	19	18
8	15	14	20	13	15
9	14	14	21	9	8
10	20	18	22	7	8
11	15	16	23	5	3
12	13	10	24	15	15

ولتكوين جدول التوزيع التكراري المزدوج نتبع الخطوات التالية:

<u>الخطوة الأولى</u>

١- نحدد أقل وأكبر درجة فى مادة الرياضيات (المتغير الأول) فتكون 0 ، 20 على الترتيب وبالتالى يكون المدى فى الرياضة

20-0 = 20 درجة

٢- نفترض عدد منسب من الفئات للمتغير الأول ولتكن 5 فئات وبالتالى يصبح طول الفئة على النحو :

$$\text{طول الفئة} = \frac{\text{عدد الفئات}}{\text{المدى}} = \frac{20}{5} = 4 \text{ درجات}$$

٣- نحدد أقل وأكبر درجة فى الإحصاء (المتغير الثانى) فتكون 2, 18 على الترتيب

٤- نفترض عدد مناسب من الفئات للمتغير الثانى وليكن 4 فئات وبالتالى فإن :

$$\text{طول الفئة}^{*} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{16}{4} = 4 \text{ درجات}$$

<u>الخطوة الثانية</u>

* ليس بالضرورة أن تكون طول الفئة للمتغير الأول يساوى طول الفئة للمتغير الثانى ولكن فى أغلب الأحيان تكون أطوال الفئات مختلفة للمتغيرين .

١- نكون جدول تفريغ مزدوج كما هو موضح فى الجدول التالى ، حيث تمثل فئات المتغير الأول (درجات الرياضة) فى العمود الأول ، وفئات المتغير الثانى (درجات الإحصاء) فى الصف الأول

٢- نبدأ فى تفريغ قيم المفردات داخل الجدول رقم (٣-٢٥) على النحو التالى :
بالنسبة للطالب الأول نجد أن درجته فى الرياضة 10 وبالتالى فإنه يقع فى فئة [8-12] فى العمود الأول الذى يمثل درجة الرياضة أى فى الصف الثالث فى الجدول ، وبما أن درجته فى الإحصاء 9 أى تقع فى الفئة [6-10] فى الصف الذى يمثل درجة الإحصاء أى فى العمود الثانى بالجدول ، وبالتالى فإن موقع هذا الطالب هو الخانة الناتجة من تقاطع الصف الثالث بالجدول (8-12) مع العمود الثانى (6-10) فيوضح فى هذه الخانة خطة مائل يمثل هذا الطالب.
وبالمثل بالنسبة للطالب رقم 20 نجد أن درجته فى الرياضة 13 وفى الإحصاء 15 وبالتالى يصبح موقعه فى الصف الرابع والعمود الرابع ، وهكذا بالنسبة لباقى الطلاب .

وبإجراء عملية التفريغ لباقى الطلاب نحصل على جدول التفريغ

جدول (٣-٢٥)

درجة الإحصاء / درجة الرياضة	2-	6-	10-	14-18
0-	///			
4-	//	///		
8-		///	//	
12-			/	/ ////
16-20				////

ومن الجدول السابق نكون جدول التوزيع التكراري المزدوج التالى

جدول (٣-٢٦) : جدول تكراري مزدوج يمثل توزيع الطلاب وفقاً لدرجاتهم فى مادتي الرياضة والإحصاء

درجة الإحصاء / درجة الرياضة	2-	6-	10-	14-18	المجموع
0-	3				3
4-	2	3			5
8-		3	2		5
12-			1	6	7
16-20				4	4
المجموع	5	6	3	10	24

ومن التوزيع التكراري المزدوج السابق يمكن إشتقاق التوزيع الهامشى لدرجة الطالب فى الرياضة من العمود الأول والعمود الأخير بالجدول المزدوج كما هو موضح بجدول (٣-٢٧) ، وبالمثل يمكن تكوين التوزيع الهامشى لدرجة الطالب فى الإحصاء من الصف الأول والصف الأخير بالجدول المزدوج كما هو موضح بجدول (٣-٢٨) .

جدول (٣-٢٧) : جدول التوزيع الهامشي لدرجة الطالب فى الرياضة

درجة الرياضة	0-	4-	8-	12-	16-20	المجموع
عدد الطلاب	3	5	5	7	4	24

جدول (٣-٢٨) : جدول التوزيع الهامشي لدرجة الطالب فى الإحصاء

درجة الإحصاء	2-	6-	10-	14-18	المجموع
عدد الطلاب	5	6	3	10	24

ملاحظات:

١- عند تكوين (أو بناء) جدول التوزيع التكراري المزدوج ممكن أن تكون أطوال الفئات للمتغير الأول متساوية وتساوى أطوال الفئات للمتغير الثانى ، وممكن أن تكون مختلفة وتختلف عن أطوال الفئات للمتغير الثانى.

٢- ممكن أن يختلف عدد الفئات للمتغير الأول عن عدد الفئات للمتغير الثانى.

٣- مجموع التكرارات للتوزيعات الهامشية المشتقة من التوزيع المزدوج لابد أن تكون متساوية.

٤- ليس ضرورى أن تكون وحدة القياس للمتغير الأول هى نفس وحدة القياس للمتغير الثانى ، بمعنى ممكن أن تكون وحدة القياس للمتغير الأول الجنية ووحدة القياس للمتغير الثانى كيلو جرام . والمثال التالى يوضح ذلك

مثال (٣-١٢)

فيما يلى بيانات عن المستوى التعليمي للأم وعدد أفراد الأسرة لكل من الأمهات محل الدراسة

جدول (٣-٢٩)

رقم الأم	مستوى تعليم الأم	عدد أفراد الأسرة	رقم الأم	مستوى تعليم الأم	عدد أفراد الأسرة
(1)	أمى	10	(13)	جامعى	3
(2)	أمى	8	(14)	ثانوى	2
(3)	ثانوى	4	(15)	إعدادى	5
(4)	ثانوى	6	(16)	أمى	9
(5)	إبتدائى	6	(17)	فوق جامعى	3
(6)	إبتدائى	7	(18)	أمى	8
(7)	إعدادى	4	(19)	جامعى	4
(8)	إعدادى	5	(20)	إعدادى	5
(9)	جامعى	3	(21)	إبتدائى	6
(10)	جامعى	4	(22)	أمى	6
(11)	فوق جامعى	2	(23)	جامعى	4
(12)	فوق جامعى	2			

المطلوب :

١- حدد مستويات التعليم المختلفة للأم.
٢- حدد المدى لعدد أفراد الأسرة.
٣- كون جدول توزيع تكرارى مزدوج مناسب.
٤- كون التوزيع الهامشى لمستوى تعليم الأم.
٥- كون التوزيع الهامشى لعدد الأطفال فى الأسرة.

الحل

١- من واقع البيانات نجد أن المستويات المختلفة لتعليم الأم هى : أمى ، إبتدائى ، إعدادى ، ثانوى ، جامعى ، فوق الجامعى.

٢- بما أن أقل عدد فى الأسرة هو 2 وأكبر عدد هو 10 أفراد بالتالى فإن :

المدى $= 10 - 2 = 8$ أفراد

٣- وبافتراض أن مستوى تعليم الأم يمثل المتغير الأول (متغير وصفى) ، وعدد الأفراد فى الأسرة يمثل المتغير الثانى (متغير كمى)

نفترض أن عدد الفئات للمتغير الثانى 4 فئات وبالتالى فإن طول الفئة يساوى

$$\frac{10-2}{4} = \frac{8}{4} = 2$$

ونكون جدول التفريغ التالى:

جدول (٣-٣٠)

عدد أفراد الأسرة مستوى تعليم الأم	2-	4-	6-	8-10
أمى			/	////
إبتدائى			///	
إعدادى		////		
ثانوى	/	/	/	
جامعى	//	///		
فوق الجامعى	///			

جدول (٣-٣١) جدول التوزيع التكرارى لـ 23 أم وفقاً لمستوى تعليم الأم وعدد أفراد أسرتها

عدد أفراد الأسرة مستوى تعليم الأم	2-	4-	6-	8-10	المجموع
أمى			1	4	5
إبتدائى			3		3
إعدادى		4			4
ثانوى	1	1	1		3
جامعى	2	3			5
فوق جامعى	3				3
المجموع	6	8	5	4	23

جدول (٣-٣٢) : التوزيع الهامشى لمستوى تعليم الأم

مستوى تعليم الأم	أمى	إبتدائى	إعدادى	ثانوى	جامعى	فوق الجامعى	المجموع
عدد الأمهات	5	3	4	3	5	3	23

جدول (٣-٣٣) : التوزيع الهامشى لعدد الأفراد فى الأسرة

عدد الأفراد فى الأسرة	2-	4-	6-	8-10	المجموع
عدد الأمهات	6	8	5	4	23

(٣-٥) أمثلة تطبيقية Applied Examples

تطبيق (٣-١)

عادة ما يتم إجراء صيانة دورية كل 6 شهور للكباري العلوية بإحدى الدول حيث يتم الفحص من خلال عدة متغيرات تبين مدى كفاءة الكوبري ، وفيما يلي بعض هذه المتغيرات:-

١- أقصى اتساع للكوبري (بالمتر).
٢- عدد الممرات للمركبات بالكوبري.
٣- متوسط الكثافة المرورية.
٤- حالة أرضية الكوبري (جيدة – متوسطة - سيئة).
٥- طول المنعطفات الجانبية للدخول أو الخروج (بالمتر).
٦- نوع الطريق (داخل حي ، يربط بين أكثر من حي ، يربط بين محافظتين أو أكثر ، يربط بين دولة ودولة أخرى أو أكثر).

المطلوب

تحديد أي من المتغيرات السابقة يعتبر متغير وصفي وأيهم يعتبر متغير كمي

الحل

المتغيرات (١) ، (٢) ، (٣) ، (٥) متغيرات كمية ، والمتغيرات (٤) ، (٦) متغيرات وصفية.

تطبيق (٣-٢)

فيما يلي بيانات عن رأس مال 40 شركة من الشركات الصغيرة (بالألف جنيه)

11.5	22	19	22	6	15	15
18	25	21	23	9	12	7.5
20	28	15.5	11	15	25	8
22	21	19.1	15	17	30	10
20	19	15	13	11.5	17	13
		18	21	29	28	21

المطلوب

١- كون جدول توزيع تكراري مناسب يوضح توزيع الشركات حسب فئات رأس المال
٢- ارسم المنحنى التكراري
٣- أوجد التوزيع المتجمع الصاعد ووضح ذلك بيانياً
٤- من (٣) قدر عدد الشركات التي يقل رأس مال كل منهم عن 17ألف جنيه

٥- من(٣)قدر عدد الشركات التي يتراوح رأس مال كل منهم بين12-17 ألف جنيه

الحل

١-

(أ) منت البيانات يتضح أن أكبر قيمة بالألف تساوي 3 وأصغر قيمة تساوي 6 ومن ثم يصبح المدى (24=6-30) ، فإذا فرضنا أن طول الفئة يساوي 4 فإن:-

$$\text{عدد الفئات} = \frac{\text{المدى}}{\text{طول الفئة}} = \frac{24}{4} = 6 \ \text{فئات}$$

(ب)تكون جدول تفريغ البيانات على النحو الموضح بالجدول التالي:-

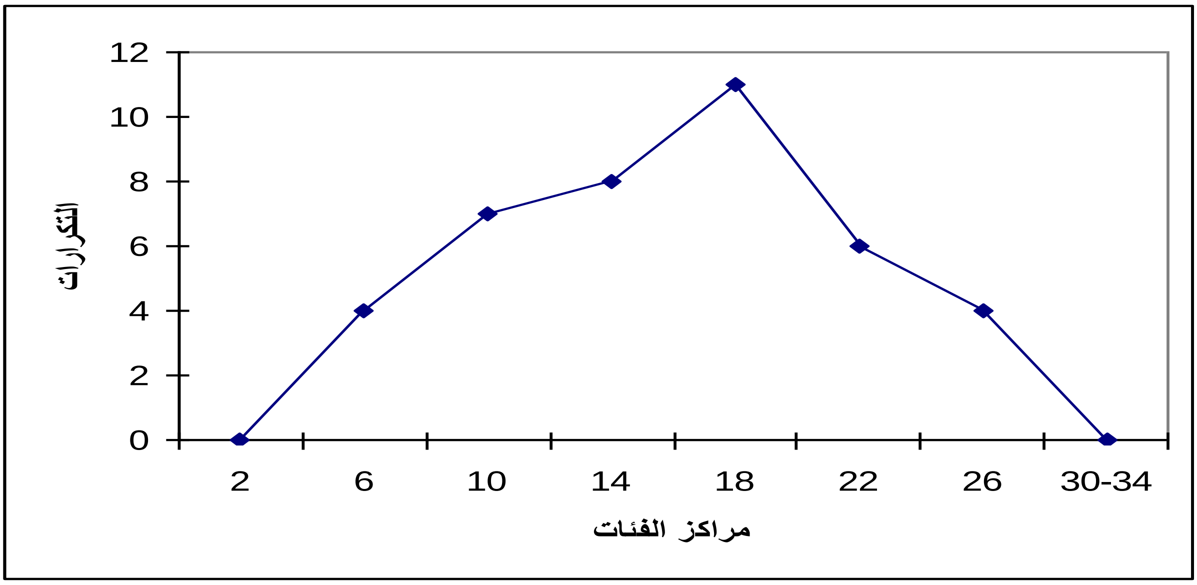
جدول(٣-٣٤)

الفئات	التفريغ	التكرارات
6-	////	4
10-	//// //	7
14-	//// ///	8
18-	~~////~~ //// /	11
22-	//// /	6
90 – 95	////	4
المجموع		40

من الجدول السابق نكون جدول التوزيع التكراري التالي

جدول (٣-٣٥)

الفئات	التكرارات	مراكز الفئات
2-	0	(2+6) ÷ 2 = 4
6-	4	(6+10) ÷ 2 = 8
10-	7	(10+14) ÷ 2 = 12
14-	8	(14+18) ÷ 2 = 16
18-	11	(22+18) ÷ 2 = 20
22-	6	(22+26) ÷ 2 = 24
26-	4	(26+30) ÷ 2 = 28
30-34	0	(30+34) ÷ 2 = 32

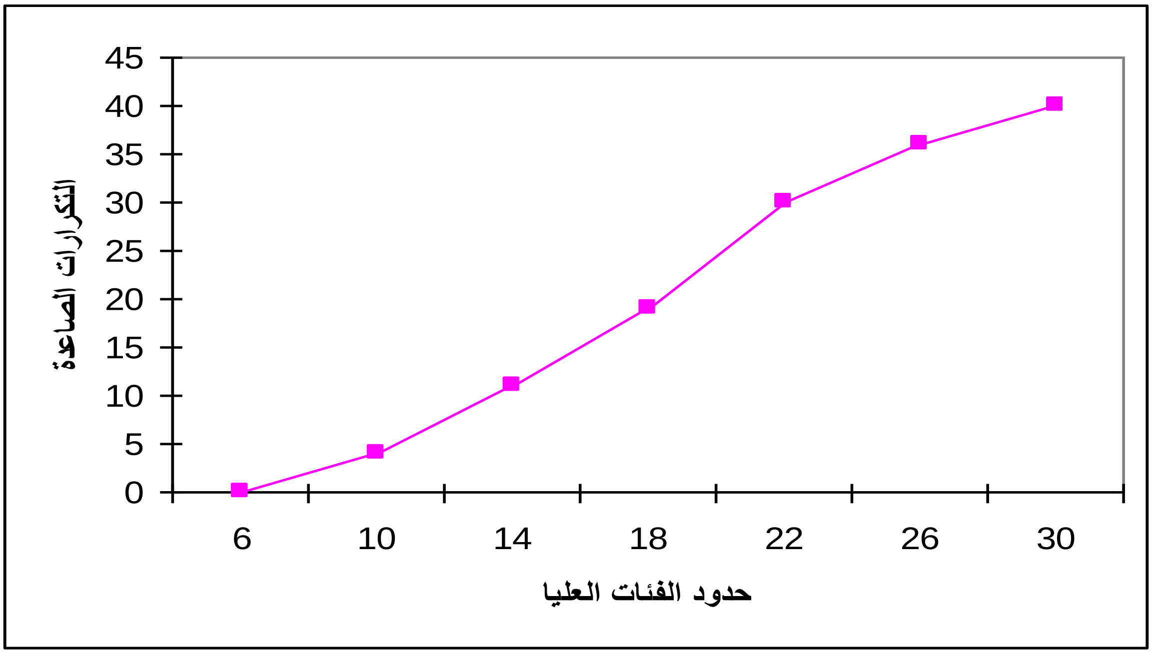


شكل (٣-١٧) : المكنحنى التكراري لتوزيع الشركات وفقاَ لرأس المال (بالألف جنيه)

٣- الجدول التالي يوضح التوزيع المتجمع الصاعد .

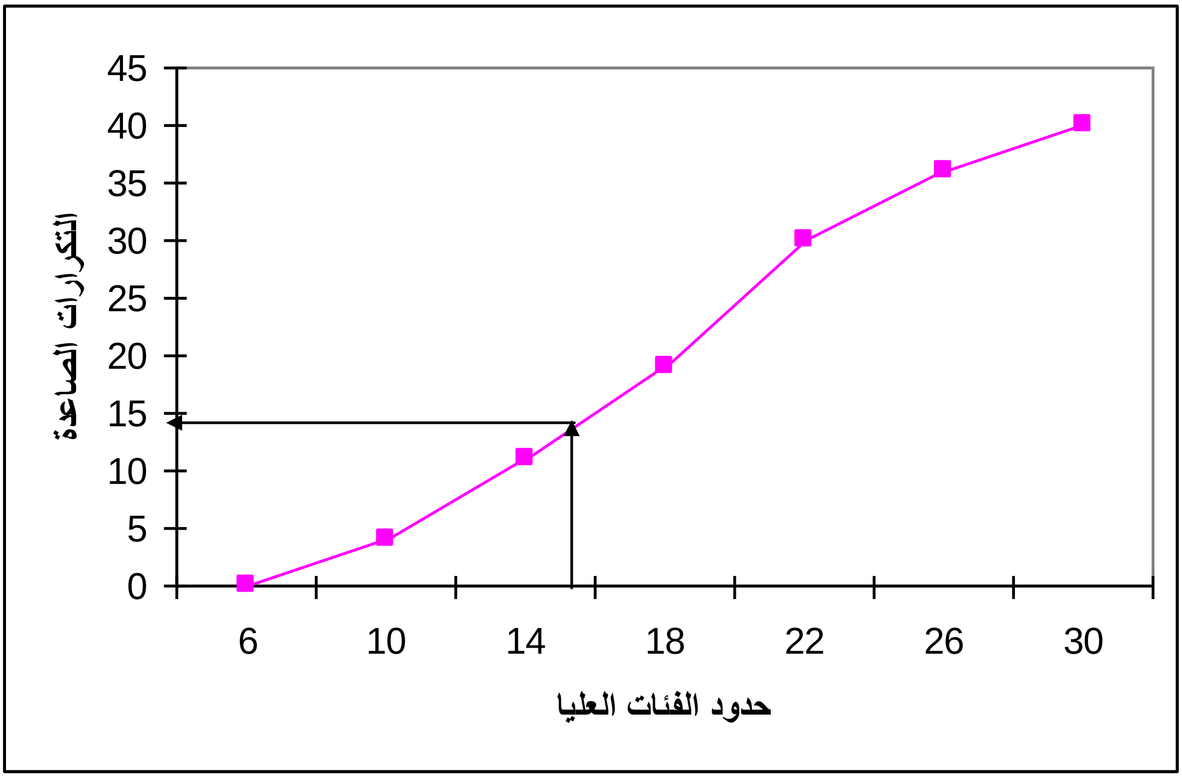
جدول (٣-٣٦)

الحدود العليا للفئات	التكرارات الصاعدة
أقل من 6	0
أقل من 10	4
أقل من 14	4+7=11
أقل من 18	4+7+8=19
أقل من 22	4+7+8+11=30
أقل من 26	4+7+8+11+6=36
أقل من أو يساوي 30	4+7+8+11+6+4=40



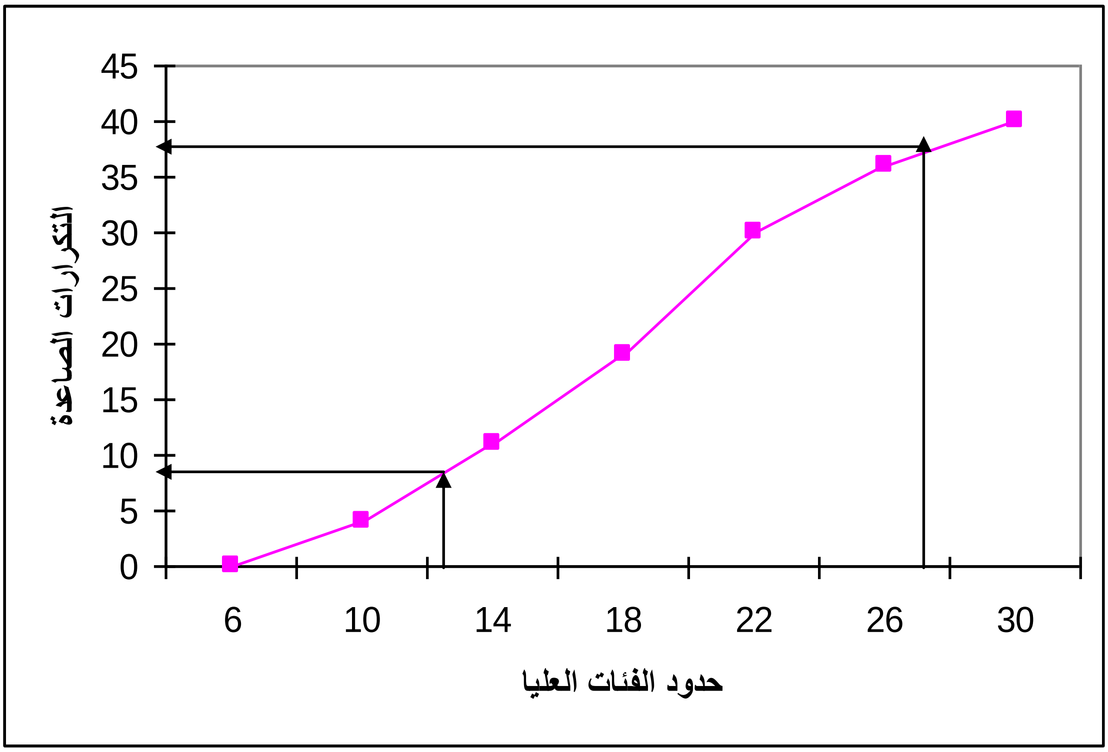
الشكل (٣-١٨) : المنحنى المتجمع الصاعد لعدد الشركات وفقاَ لرأس المال (بالألف جنيه)

٤- باستخدام المنحنى المتجمع الصاعد نجد أن عدد الشركات التي يقل رأس مال كل منهم عن 7 ألاف جنيه يساوي 17 شركة تقريباَ كما هو موضح بالشكل التالي:-



شكل (٣-١٩)

٥- من الشكل التالي نجد أن عدد الشركات التي رأس مال كل منهم أقل من 27 ألف جنيه يساوي 37 شركةتقريباً ، كذلك عدد الشركات التي رأس مال كل منهم أقل من 12 ألفجنيه يساوي 8شركة تقريباً ، وبالتالي فإن عدد الشركات التي رأس مال كل منهم بين 27-12 ألف جنيه يساوي 29=8-37 شركة



شكل (٣-٢٠)

(٣-٦)تمرينات Exercises

(٣-١)

الجدول التالي يوضح توزيع 120 أسرة حسب عدد الأفراد في كل أسرة

جدول (٣-٣٧)

عددالأفراد في الأسرة	2 - 4	5 - 7	8 - 10	11 - 13	المجموع
التكرارات	80	20	15	5	120

المطلوب

١- حدد الحدود العليا والدنيا لكل فئة
٢- أوجد أطوال الفئات ، ثم أوجد مراكز الفئات
٣- أوجد التوزيع التكراري النسبي ثم ارسم المنحنى التكراري النسبي

(٣-٢)

الجدول التالي يوضح توزيع 500 عامل حسب فئات الأجر الأسبوعي بالجنيه

جدول (٣-٣٨)

الأجر الأسبوعي بالجنيه	50-	100-	150-	200-	250-300	المجموع
عدد العمال	80	120	200	75	25	500

المطلوب

١- ارسم كل من الددرج التكراري والمضلع التكراري في شكل واحد ، ثم اثبت أن مساحة المدرج التكراري تكافئ مساحة المضلع التكراري
٢- قدر عدد العمال الذين يزيد الأجر الأسبوعي لكل منهم عن 220 جنيه أسبوعياً

(٣-٣)

الجدولين التاليين يوضحان توزيع إحدى الدفعات الدراسية بإحدى كليات التجارة في الصف الأول والثاني وفقاً لدرجاتهم في مادة الرياضة

جدول (٣-٣٩) توزيع الطلاب في الصف الأول في مادة الرياضيات

الفئات (درجة)	0-	5-	10-	15-20	المجموع
التكرارات (عدد الطلاب)	400	600	1000	500	2500

جدول (٣-٤٠) توزيع الطلاب في الصف الثاني في مادة الرياضيات

الفئات (درجة)	0-	5-	10-	15-20	المجموع
التكرارات (عدد الطلاب)	350	500	900	100	1850

المطلوب :

قارن توزيع الطلاب وفقاً لدرجاتهم في مادة الرياضة في الصف الأول والثاني.

(٣-٤)

فيما يلي بيانات عن مستوى تعليم الأم في كل أسرة :-

أمي	متوسط	متوسط	متوسط	أمي	عالي	أمي
أمي	أمي	متوسط	عالي	أمي	أمي	أمي
متوسط	أمي	متوسط	أمي	متوسط	أمي	أمي
عالي	متوسط	أمي	أمي	متوسط	أمي	أمي
أمي	عالي	أمي	أمي	أمي	أمي	عالي
أمي	أمي	أمي	متوسط	أمي	عالي	أمي
عالي	عالي	متوسط	متوسط	أمي	أمي	أمي

المطلوب :

١- حدد المستويات المختلفة لتعليم الأم في هذه العينة.
٢- حول مستوى تعليم الأم من متغير وصفي إلى متغير كمي.
٣- كون جدول توزيع تكراري يوضح توزيع الأسر وفقاً لمستوى تعليم الأم في الأسرة.
٤- قدر نسبة الأسر التي يكون مستوى تعليم الأم فيها أقل من المتوسط.

(٣-٥)

أخذت عينة من محدودي الدخل مكونة من 20 فرد ، وفيما يلي بيانات عن متوسط الدخل الشهري للفرد بالجنية (المتغير الأساسي) ومتوسط الاستهلاك الشهري لكل فرد بالجنية (المتغير التابع).

جدول (٣-٤١)

رقم الفرد	متوسط الدخل	متوسط الإستهلاك	رقم الفرد	متوسط الدخل	متوسط الإستهلاك
(1)	300	280	(11)	170	175
(2)	400	350	(12)	180	190
(3)	220	230	(13)	200	200
(4)	250	240	(14)	150	180
(5)	310	300	(15)	190	200
(6)	270	290	(16)	303	300
(7)	210	215	(17)	400	370
(8)	270	295	(18)	390	350
(9)	300	330	(19)	250	260
(10)	280	300	(20)	380	300

المطلوب :

١- كون جدول توزيع تكراري مزدوج يوضح توزيع الأفراد وفقاً لمتوسط الدخل والإستهلاك الشهري بالجنية.

٢- كون التوزيع الهامشي لكل من متوسط من متوسط الدخل ومتوسط الإستهلاك.

٣- كون التوزيع المتجمع الصاعد لكل من متوسط متوسط الدخل ومتوسط الإستهلاك.

٤- قدر عدد الأفراد الذين يقل دخلهم عن 300 جنية ، والذين يقل إستهلاكهم عن 300 جنية أيضاً.

٥- قارن بين توزيع الأفراد وفقاً لمتوسط الدخل ومتوسط الإستهلاك.

(٣-٦)

نظراً للطلب المتزايد على زجاجات المياه المعدنية (حجم لتر) أرادت إحدى المؤسسات إنتاج هذه الزجاجات التأكد من صحة حجم المياه المعبأة وذلك لزيادة حجم الإنتاج ، فأخذت عينة من 50 زجاجة حجم المياه المعبأة بها كما يلي :

1.02	1.01	1.00	0.85	0.99	0.91	0.88	1.00	1.04
0.98	0.98	1.00	0.89	0.97	0.92	0.89	1.03	1.05
0.99	1.02	0.99	0.98	0.95	0.95	0.89	1.05	1.03
1.00	0.99	0.97	0.88	0.94	0.91	1.00	1.10	1.02
1.01	0.98	0.99	0.98	0.90	0.93	0.99	1.03	0.99
1.00	1.02	1.04	1.03	1.02				

المطلوب :

١- أوجد أقل وأكبر حجم للمياه بالزجاجة.
٢- كون جدول تكراري مناسب.
٣- أرسم مدرج تكراري يوضح توزيع الزجاجات حسب حجم المياه الموجودة بالزجاجة.
٤- كون التوزيع المتجمع الهابط ، ثم أرسم المنحني الهابط ومنه قدر عدد الزجاجات التي حجم كل منها أكبر من (1.00).
٥- قدر عدد الزجاجات التي حجم كل منها أقل من (100).

(٣-٧)

فيما يلي بيانات عن درجات عينة من 70 دارس في مقرر الإحصاء :-

100	71	12	99	11	22	61
82	82	25	79	25	71	72
55	17	37	66	16	53	53
49	11	45	55	66	61	75
23	22	89	64	65	65	99
50	25	69	62	59	99	81
11	77	77	88	66	42	92
61	99	89	63	98	45	77
82	82	99	55	77	59	99
91	71	72	66	91	59	11

المطلوب :

١- كون توزيع تكراري مناسب يوضح توزيع الـ 70 دارس وفقاً لدرجات الدارس.
٢- أرسم المدرج التكراري الذي يوضح التوزيع في (١).
٣- قدر عدد الطلاب الذين حصلوا على 80 درجة فأكثر.
٤- قدر عدد الطلاب الذين حصلوا على 60 درجة فأكثر.

الباب الرابع

بعض مقاييس الموضع والتشتت

Some Measures of Location and Desperation

(٤-١) الحصول على المعلومات من البيانات

Finding an information from data

(٤-٢) الوسط الحسابي **Arithmetic Mean**

(٤-٣) الوسيط **Median**

(٤-٤) المنوال **Mode**

(٤-٥) نصف المدى الربيعى **Semi- Interquarter Range**

(٤-٦) التباين والانحراف المعياري

Variance and Standard Deviation

(٤-٧) معاملات الاختلاف

Coefficients of Variation

(٤-٨) أمثلة تطبيقية **Applied Examples**

(٤-٩) تمرينات **Exercises**

(٤-١) الحصول على المعلومات من البيانات
Finding an information from data

في الباب الأول وضحنا بالتفصيل أن البيانات هي العنصر الرئيسي في صناعة القرار . ولكن لا يمكن استخدام البيانات في صناعة القرار إلا بعد تحويلها إلى معلومات Information ، ثم معرفة Knowledge ، كما سبق توضيح ذلك في شكل (١-١) صفحة ١٥
وفي الباب السابق تناولنا كيفية تصنيف مجموعة من البيانات في جدول تكراري بسيط وعرضها بيانياً في صورة مدرج تكرارى أو مضلع تكرارى أو منحنى تكرارى .

والجدول التكراري يوضح توزيع المفردات وفقاً لقيم المتغير (الظاهرة) محل الدراسة ، إلا أنه قد لا يكون من اليسير استخدام الجداول التكرارية في إبراز خصائص المتغير (الظاهرة) أو المقارنة بين هذا المتغير ومتغير أخر أو عدة متغيرات أخرى . لذا كان من الضروري والمفيد استخراج معلومات رقمية لتلخيص خصائص الظاهرة وتتمثل هذه المعلومات عادة في صورة مقاييس (مؤشرات) عددية Measures يمكن استخدامها لتحديد خصائص الظاهرة بالإضافة إلى استخدامها في المقارنة بين عدة ظواهر (متغيرات) .

وتعتبر مقاييس الموضع Measures of Location وقد تسمى بمقاييس النزعة المركزية Measures of Central of Tendency أو بالمتوسطات Averages أحياناً ، وكذلك مقاييس التشتت Measures of Variation من أهم المؤشرات التي تعطي معلومات ذات أهمية في دراسة خصائص الظاهرة (المتغير) .

أولاً : مقاييس الموضع

بالنسبة لمعظم المتغيرات (أو الظواهر) نجد أن قيم المتغير تقع بين نهاية عظمى ونهاية صغرى ، وبين هاتين النهايتين تتوزع القيم التي تأخذها المفردات وعادة تتوزع المفردات (التكرارات) بحيث إذا انتقلنا من أصغر القيم إلى أكبرها فإن التكرارات المناظرة لهذه القيم تتزايد حتى تصل إلى أكبر تكرار ثم تبدأ التكرارات في التناقص التدريجي مع تزايد القيم للمتغير . ومن هنا فأنه توجد نقطة ما يميل أكبر عدد من المفردات إلى التركيز عندها ، ولما كانت هذه النقطة تقع غالباً عند مركز التوزيع فإن هذا الميل إلى التركيز حول نقطة يسمى بالنزعة المركزية Central of Tendency وتسمى الأساليب المختلفة لتحديد القيم التي تتركز حولها المفردات بمقاييس النزعة المركزية أو بمقاييس الموضع أو بالمتوسطات .

والهدف من حساب هذه المقاييس (أو المؤشرات) هو تحديد قيمة واحدة تمثل مجموعة كبيرة من قيم المفردات ، حيث يمكن استخدام هذه المؤشرات في التعرف على خصائص الظاهرة كذلك في المقارنة بينها وبين ظواهر أخرى .

ويوجد العديد من مقاييس الموضع (المتوسطات) مثل الوسط الحسابي Arithematic Mean ، الوسيط Median ، المنوال Mode ، الوسط الهندسي Geometric Mean ، الوسط المرجح Weighted Mean ، الوسط التوافقي Harmonic Mean .
وفي الفصول الثلاثة التالية سوف نتناول بالدراسة بعض أهم هذه المؤشرات

ثانياً : مقاييس التشتت

مما سبق يتضح أن مقاييس الموضع (النزعة المركزية) تعطى القيمة التي تتركز حولها معظم المفردات دون أن تعطى أى إيضاح عن مدى تقارب أو تباعد قيم هذه المفردات عن هذه القيمة أو عن بعضها البعض ، أو بعبارة أخرى لا تعطى مؤشرات الموضع أى معلومات عن مدى تجانس أو تشتت قيم المفردات .
لذلك فإنه يكون من الضروري وجود مؤشرات أخرى تصف تشتت (أو تقارب وتباعد) قيم المفردات مع بعضها البعض .

ويوجد العديد من المؤشرات التي تصف مدى تشتت أو تجانس القيم مثل المدى Range ، نصف المدى الربيعي Semi - Interquarter Range ، التباين والانحراف المعياري Variance and Standard Deviation ، الانحراف المتوسط Mean Deviation ، .... الخ .
وفي الفصليين (٤-٥) (٤-٦) سوف نتناول بعض هذه المؤشرات.

أولاً : مقاييس الموضع

وفيما يلي سوف نتناول في الفصول (٤-٢) - (٤-٤) أهم مقاييس النزعة المركزية "الوسط الحسابي ، الوسيط ، المنوال" على الترتيب

(٤-٢) الوسط الحسابي Arithmetic Mean

أولاً : الوسط الحسابي من بيانات غير مبوبة

إذا كان الاستهلاك اليومي (X) لعشرة أفراد بالجنية على النحو التالي

X : 12 , 18 , 10 , 13 , 5 , 7 , 4 , 6 , 5 , 10

وبالتالي فإن مجموع الاستهلاك اليومي للأفراد العشر يصبح

$$X = 12+ 18+ 10+ 13+ 5+ 7+ 4+ 6+ 5+ 10 = 80 \text{ جنية} \qquad (4.1)$$

وإذا قسمنا المجموع (80) على عدد الأفراد (10) نجد أن

$$\frac{80}{10} = 8 \text{ جنية} \qquad (4.2)$$

أى أنه إذا استهلك كل فرد من الأفراد العشر (8) جنيهات يومياً فإن مجموع استهلاك الأفراد العشر يكون (80) جنية – أى نفس مجموع استهلاكهم الأصلي في (4.1) ، وتسمى القيمة (8) بالوسط الحسابي أو المتوسط لاستهلاك الفرد اليومي .

تعريف (٤-١)

ومما سبق يمكن تعريف الوسط الحسابي بأنه القيمة التي لو أعطيت لكل مفردة لكان مجموع القيم يساوى مجموع القيم الأصلية .

وبصفة عامة إذا كان لدينا N من المشاهدات (المفردات) تأخذ القيم $X_1 , X_2 , X_3 , \ldots\ldots\ldots , X_N$ ، وأشرنا للوسط الحسابي في المجتمع بالرمز μ فإن :-

$$\mu = X_1 + X_2 + X_3 + \ldots\ldots\ldots\ldots + X_N = \frac{\sum_{i=N}^{N} X_i}{N} \qquad (4.3)$$

وعادة يشار إلى الوسط الحسابي للمتغير عندما يتم حساب الوسط الحسابي من بيانات عينة حجمها n بالرمز $\bar{x}$ فإن

$$\bar{X} = \qquad (4.4)\frac{\sum_{i=1}^{n} X_i}{n}$$

خصائص الوسط الحسابي

١- إذا فرضنا أن x , y متغيرات بحيث $\bar{x}, \bar{y}$ هما الوسط الحسابي للمتغيرين x ,y على الترتيب . فإذا كان :

$$y_i = x_i - a \qquad (4.5)$$

حيث أن a مقدار ثابت (يسمى بالوسط الافتراضي) ويسمى المتغير y بالانحرافات البسيطة عن المتغير x بمقدار a . فإنه يمكن إثبات أن :

$$\bar{X} = \bar{y} + a \qquad (4.6)$$

أى أن الوسط الحسابي للقيم الأصلية x_i يساوي الوسط الحسابي للانحرافات البسيطة y_i مضاف إليه المقدار الثابت a.

الأثبات : أنظر ملحق (١) صفحة ٣٣٩

٢- إذا فرضنا أن المقدار الثابت a بحيث $a = \bar{x}$ فتصبح الانحرافات البسيطة

$$y_i = x_i - \bar{X} \qquad (4.7)$$

ويمكن إثبات أن :

$$\sum_{i=1}^{n} y_i = 0 \quad , \quad (4.8) \quad \bar{y} = 0$$

أى أن مجموع انحرافات القيم عن وسطها الحسابي يساوى صفر ، وبالتالي متوسط انحرافات هذه القيم يساوى صفر أيضاً

٣- إذا فرضنا أن

$$h_i = \frac{x_i - a}{b} \qquad (4.9)$$

حيث a,b مقادير ثابتة ، $b \neq 0$ ، وتسمى h_i انحرافات مختصرة وسطها الحسابي $\bar{h}$ فإنه يمكن إثبات أن :

$$\bar{X} = b\,\bar{h} + a \qquad (4.10)$$

وتستخدم جميع العناصر السابقة في تبسيط الحسابات لحساب الوسط الحسابي كما سوف يتضح في الأمثلة التالية :

مثال (٤-١)

إذا كان الدخل الشهري لـ 10 عمال بالجنية على النحو التالي :

$$x_i = 199\ ,199\ ,197,\ 192,\ 198,\ 208,\ 201,\ 213,\ 195,\ 203$$

المطلوب

١- أوجد الوسط الحسابي بالطريقة المباشرة

٢- أوجد الوسط الحسابي باستخدام الانحرافات البسيطة

الحل

١- إذا فرضنا أن $\bar{x}$ تشير إلى الوسط الحسابي فإن :

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{n}$$

$$= \frac{199 + 199 + 197 + 192 + 198 + 208 + 201 + 213 + 195 + 203}{10}$$

$$= \frac{2005}{10} = 200.5 \quad \text{جنية}$$

٢- إذا فرضنا أن المقدار الثابت a بحيث $a = 200$ ، y_i تمثل الانحرافات البسيطة بحيث

$$y_i = x_i - a = x_i - 200$$

فإن :

$$y_1 = 199 - 200 = -1 \qquad y_2 = 199 - 200 = -1$$
$$y_3 = 197 - 200 = -3 \qquad y_4 = 192 - 200 = -8$$
$$y_5 = 198 - 200 = -2 \qquad y_6 = 108 - 200 = 8$$
$$y_7 = 201 - 200 = 1 \qquad y_8 = 213 - 200 = 13$$
$$y_9 = 195 - 200 = -5 \qquad y_{10} = 203 - 200 = 3$$

وبما أن :

$$\bar{y} = \frac{\sum_{i=1}^{10} y_i}{10}$$
$$= \frac{(-1) + (-1) + (-3) + ...+ (3)}{10} = \frac{5}{10} = 0.5$$

ومن الخاصية (١) نجد أن

$\bar{x} = \bar{y} + a$ = 0.5 +200 = 200.5 جنية

ومما سبق يتضح أن قيمة $\bar{x}$ باستخدام الانحرافات البسيطة y_i هي نفس القيمة باستخدام الحسابات المباشرة.

مثال (٤-٢)
فيما يلي أطوال 7 أشخاص بالسنتمترات :

83 ,98 ,113 ,128 ,143 ,158 ,173

أحسب الوسط الحسابي للطول باستخدام طريقة الانحرافات المختصرة

الحل

إذا فرضنا أن x_i تشير إلى طول الفرد رقم i ، h_i هى الانحرافات المختصرة بحيث :

$$h_i = \frac{x_i - a}{b} = \frac{x_i - 68}{15}$$

نجد أن :

$$h_1 = \frac{83-68}{15} = 1 \qquad h_2 = \frac{98-68}{15} = 22$$

$$h_3 = \frac{113-68}{15} = 3 \qquad h_4 = \frac{128-68}{15} = 4$$

$$h_5 = \frac{143-68}{15} = 5 \qquad h_6 = \frac{158-68}{15} = 6$$

$$h_7 = \frac{173-68}{15} = 7$$

وبالتالي فإن :

$$h = \frac{\sum_{i=1}^{n} h_i}{n} = \frac{1+2+3+.....+7}{7} = 4$$

$$\therefore \overline{X} = b\,\overline{h} + a = 15(4) + 68 = 128 \text{ سم}$$

ثانياً إيجاد الوسط الحسابي من بيانات مبوبة

إذا كانت البيانات مبوبة في جدول تكراري بسيط بحيث أن x_i, f_i تمثل تكرار الفئة i ومركز الفئة i على الترتيب ، فإذا كان عدد الفئات يساوى n حيث $i = 1,2,3,....,n$ فإنه يمكن حساب الوسط الحسابي μ للمتغير محل الدراسة في المجتمع على النحو التالي :

$$\mu = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + ...+ x_n f_n}{f_1 + f_2 + f_3 + ...+ f_n} = \frac{\sum x_i f_i}{\sum f_i} \qquad (4.11)$$

أما إذا كانت البيانات تمثل بيانات عينة فإن الوسط الحسابي $\overline{x}$ للمتغير في العينة يصبح على النحو التالي :

$$\overline{X} = \frac{\sum x_i f_i}{\sum f_i}$$

حيث f_i, x_i تمثل مراكز الفئات والتكرارات للتوزيع التكراري للعينة.

مثال (٤-٣)

الجدول التالي يوضح توزيع 100 أسرة حسب عدد الأطفال في الأسرة الواحدة بإحدى قرى جمهورية مصر العربية.

جدول (٤-١)

عدد الأطفال في الأسرة الواحدة	0-	2-	4-	6-	8-10	المجموع
عدد الأسر	5	25	45	15	10	100

المطلوب :

أحسب متوسط (أو الوسط الحسابي) لعدد الأطفال في الأسرة الواحدة.

الحل

لحساب الوسط الحسابي (المتوسط) من بيانات مبوبة نكون الجدول التالي:-

جدول (٤-٢)

الفئات	التكرارات f_i	مراكز الفئات x_i	$x_i f_i$
0-	5	1	5
2-	25	3	75
4-	45	5	225
6-	15	7	105
8-10	10	9	90
المجموع	$\sum f_i = 100$		$\sum x_i f_i = 500$

حيث تم حساب مراكز الفئات (x_i) في العمود الثالث ، ثم تم ضرب كل عنصر في العمود الثالث في العنصر المناظر له في العمود الثاني فنحصل على قيم العمود الرابع $(x_i f_i)$

وبما أن

$$\overline{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{500}{100} = 5 \text{ أطفال}$$

وتنطبق خصائص الوسط الحسابي في حالة البيانات غير المبوبة في حالة البيانات المبوبة أيضاً كما يتضح فيما يلى :

١- إذا كان y_i انحرافات بسيطة بحيث

$$y_i = x_i - a \qquad (4.12)$$

حيث (x_i) مركز الفئة i ، a مقدار ثابت فإن

$$\overline{X} = \overline{Y} + a \qquad (4.13)$$

حيث أن

$$\bar{Y} = \frac{\sum y_i f_i}{\sum f_i}, \qquad \bar{X} = \frac{\sum x_i f_i}{\sum f_i}$$

٢- كذلك إذا كان h_i انحرافات مختصرة بحيث

$$h_i = \frac{x_i - a}{b} \qquad (4.14)$$

فإن :

$$\bar{X} = b\ h_i + a \qquad (4.15)$$

حيث أن :

$$\bar{h} = \frac{\sum h_i f_i}{\sum f_i}$$

مثال (٤-٤)

الجدول التالي يوضح توزيع 1280 عامل بإحدى المصانع وفقاً للدخل الشهري بالجنية

جدول (٤-٣)

فئات الدخل بالجنية	213-	216-	219-	222-	225-228	المجموع
عدد العمال	127	321	352	281	199	1280

المطلوب :

أحسب الوسط الحسابي للأجر الشهري للعامل باستخدام :-

١- طريقة الانحرافات البسيطة ٢- طريقة الانحرافات المختصرة

الحل

١- لحساب الوسط الحسابي باستخدام طريقة الانحرافات البسيطة تتبع الخطوات التالية :

أ- حساب مراكز الفئات x_i كما هو موضح بالجدول التالي

ب- نفترض وسط فرضي a ويفضل أن يكون إحدى قيم مراكز الفئات كذلك يفضل أن تكون القيمة المناظرة لأكبر تكرار أو القيمة التي تقع في منتصف عمود مراكز الفئات ، ويرجع هذا التفضيل إلى تبسيط الحسابات فقط . ولكن ممكن أن يكون الوسط الفرضي أى قيمة أخرى وفي هذه الحالة سوف نفترض أن $a = 220.5$ ثم نكون عمود يتم فيه

حساب الانحرافات البسيطة y_i حيث $(y_i = x_i - a)$ كما هو موضح بالعمود رقم (4) بالجدول التالي

ج- نكون العمود رقم (5) وهو عبارة عن حاصل ضرب الانحرافات البسيطة y_i في التكرارات المناظرة لها f_i ، أى نقوم بحساب قيم $y_i f_i$

جدول (٤-٤)

(1) الفئات	(2) f_i	(3) x_i	(4) y_i	(5) $y_i f_i$
213-	127	214.5	214.5-220.5=-6	127=-762×-6
216-	321	217.5	217.5-220.5=-3	321=-963×-3
219-	352	220.5	220.5-220.5=0	352=0×0
222-	281	223.5	223.5-220.5=3	281=843×3
225-228	199	226.5	226.5-220.5=6	199=1194×6
المجموع	$\sum f_i = 1280$			$\sum y_i f_i = 312$

د- من الجدول السابق نجد أن :

$$\bar{Y} = \frac{\sum y_i f_i}{\sum f_i} = \frac{312}{1280} = 0.25$$

وبما أن $\bar{X} = \bar{Y} + a$ فإن :

$$\bar{X} = 0..25 + 220.5 = 220.75 \text{ جنية}$$

٢- من العمود رقم (4) بالجدول السابق نجد أن الانحرافات البسيطة y_i يوجد بينها عامل مشترك يساوى 3 ، لذا يمكن تكوين الانحرافات المختصرة h_i حيث يمكن افتراض أن المقدار الثابت b يساوى العامل المشترك وبالتالي يمكن افتراض أن الانحرافات المختصرة h_i على النحو التالي :

$$h_i = \frac{x_i - a}{b} = \frac{x_i - 220.5}{3}$$

ثم نكون الجدول التالي :

جدول (٤-٥)

الفئات	f_i	x_i	h_i	$h_i f_i$
213-	127	214.5	(214.5-220.5)/3=-2	-254
216-	321	217.5	(217.5-220.5)/3=-1	-321
219-	352	220.5	(220.5-220.5)/3=0	0
222-	281	223.5	(223.5-220.5)/3=1	281
225-228	199	226.5	(226.5-220.5)/3=2	398
المجموع	1280			$\sum h_i f_i = 104$

وبما أن :

$$\bar{h} = \frac{\sum h_i f_i}{\sum f_i} = \frac{104}{1280} = 0.082$$

كذلك :

$$\bar{x} = b\,\bar{h} + a = 3(0.082) + 220.5 = 220.75 \text{ جنية}$$

(٤-٣) الوسيط Median

في الفصل السابق تناولنا الوسط الحسابي كأحد المؤشرات الهامة لمؤشرات الموضع (النزعة المركزية) وفي هذا الفصل سوف نتناول مؤشر أخر وهو الوسيط.

تعريف (٤-٢)

الوسيط هو القيمة التي تتوسط القيم إذا رتبت القيم ترتيباً تصاعدياً أو تنازلياً وعادة يشار للوسيط بالرمز M_2

<u>أولاً : الوسيط من بيانات غير مبوبة</u>

إذا كان لدينا مجموعة من المفردات قيمها على النحو التالي:

8, 6, 4, 2, 25, 25, 30, 1, 15, 13

فإنه إذا تم ترتيبها ترتيباً تصاعدياً على النحو التالي

1 ,2 ,4 ,6 ,8 ,13 ,15 ,25 ,30

فنجد أن القيمة (8) تتوسط القيم ، بمعنى أن أربعة مفردات تأخذ قيم أقل من (8) وكذلك أربعة مفردات تأخذ قيم أكبر من (8) . وبالتالي تصبح القيمة (8) هى الوسيط لهذه القيم أى أن :

$$M_2 = 8$$

ويمكن تعميم ذلك لأى عدد من القيم فإذا كان عدد المفردات يساوى n ، حيث n عدد فردى فإن ترتيب الوسيط بعد ترتيب القيم تصاعدياً أو تنازلياً يساوى

$$\frac{n+1}{2} \qquad (4.16)$$

وإذا كان n عدد زوجي فنجد أن الوسيط هو المتوسط للقيمتين اللتين ترتيبيهما :

$$\frac{n}{2} \quad ، \quad \frac{n}{2}+1 \qquad (4.17)$$

كما سوف يتضح في المثال التالي

مثال (٤-٦)

أوجد القيمة الوسيطة لمجموعة القيم التالية :

382, 388, 390, 369, 311, 402, 410, 395, 352, 375

الحل

١- ترتيب البيانات تصاعدياً (أو تنازلياً) على النحو التالي:

410, 402, 395, 390, 388, 382, 375, 369, 352, 311

٢- بما أن عدد المفردات n =10 ، أى عدد زوجي فبالتالي يكون الوسيط هو متوسط القيمتين اللتين لهما الترتيب

$$\frac{10}{2}=5 \quad ، \quad \frac{10}{2}+1=6$$

أى أن الوسيط (M_2) بحيث

$$M_2=\frac{382+388}{2}=\frac{775}{2}=385$$

فنجد أن خمسة مفردات تاخذ قيمة أقل من (285) وخمسة مفردات الأخرى تأخذ قيمة أكبر من (385)

ثانياً : الوسيط من بيانات مبوبة

لإيجاد الوسيط من بيانات مبوبة في جدول توزيع تكراري يتطلب تكوين التوزيع المتجمع الصاعد (أو الهابط) ثم إيجاد قيمة الوسيط بيانياً باستخدام المنحنى المتجمع الصاعد (أو الهابط).

فوفقاً لتعريف الوسيط يكون ترتيب الوسيط يساوى $\frac{\sum f_i}{2}$ ، حيث يحدد ترتيب

الوسيط على المحور الراسي وتكون قيمة الوسيط هى القيمة المناظرة لهذا الترتيب على المحور الأفقي كذلك يمكن تحديد قيمة الوسيط جبرياً من العلاقات التالية:

$$M_2 = M_2 \text{ بداية فئة } + \frac{\text{التكرار الصاعد السابق لترتيب } M_2 \text{ - ترتيب } M_2}{\text{تكرار فئة } M_2} \times M_2 \text{ طول فئة} \quad (4.18)$$

إذا استخدمنا التوزيع المتجمع الصاعد كذلك يمكن حساب الوسيط جبرياً باستخدام التوزيع المتجمع الهابط من العلاقة التالية:

$$M_2 = M_2 \text{ نهاية فئة } - \frac{\text{التكرار الهابط اللاحق لترتيب } M_2 \text{ - ترتيب } M_2}{\text{تكرار فئة } M_2} \times M_2 \text{ طول فئة} \quad (4.19)$$

وسوف نوضح ذلك في المثال التالي

مثال (٤-٧)
الجدول التالي يوضح توزيع 200 أسرة من محدودى الدخل حسب الدخل الشهرى للأسرة بالجنية

جدول (٤-٦)

الدخل الشهرى بالجنية	200-	220-	240-	260-	300-280	المجموع
عدد الأسر	20	40	100	30	10	200

المطلوب

١- كون التوزيع المتجمع الصاعد (أو الهابط)
٢- أرسم المنحنى المتجمع الصاعد ثم قدر باستخدامه قيمة الوسيط
٣- أوجد قيمة الوسيط جبرياً

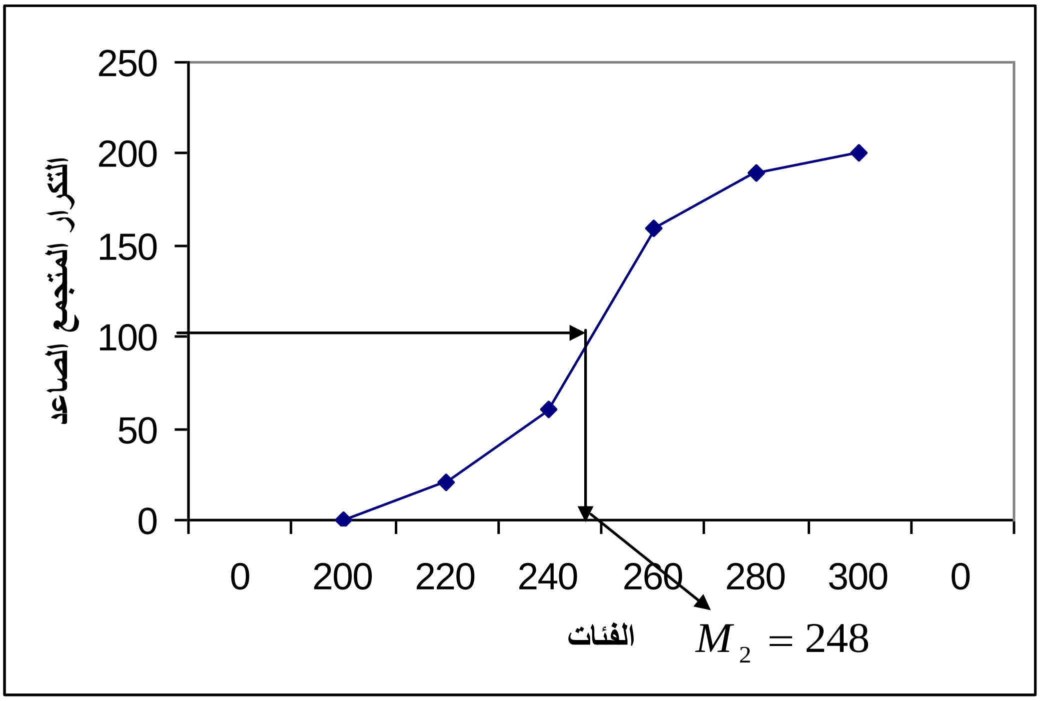
الحل

١- نكون التوزيع المتجمع الصاعد كما هو موضح بالجدول التالي :

جدول (٤-٧)

التكرارات المتجمعة الصاعدة	الحدود العليا للفئات
0	أقل من 200
20	أقل من 220
60	أقل من 240
160	أقل من 260
190	أقل من 280
200	أقل من أو يساوى 300

٢- والشكل التالي يوضح المنحنى المتجمع الصاعد



شكل (٤-١)

وبما أن

$$\text{ترتيب الوسيط} = \frac{\sum f_i}{2} = \frac{200}{2} = 100$$

وبتحديد ترتيب الوسيط على المحور الراسى ثم رسم مستقيم من هذه النقط موازى للمحور اللأفقى فتكون نقطة تقاطع هذا العمود مع المحور الأفقى هى قيمة الوسيط ، كما هو موضح بالرسم ، حيث نجد أن هذا العمود يلاقى المحور اللأفقى عند 248 وبالتالي نجد أن :

$$M_2 = 248 \text{ جنية}$$

أي أن 50% من الأسر (100 أسرة) يكون دخل كل منها أقل من أو يساوى 248 ، وبالتالي الـ 50% الأخرى (100 أسرة الأخرى) من الأسر يكون دخل كل منها أكبر من أو يساوى 248 جنية شهرياً

٣- من العلاقة (4.18) نجد أن :

$$M_2 = \text{بداية فئة } M_2 + \frac{\text{ترتيب } M_2 - \text{التكرار الصاعد السابق لترتيب } M_2}{\text{تكرار فئة } M_2} \times \text{طول فئة } M_2$$

أ- بما أن ترتيب الوسيط 100 فنجد أن هذا الترتيب يقع في العمود الثاني في جدول (٤-٧) بين (160 - 60) ، ونجد أن حدود الفئات المناظرة لكل من 60 ، 160 هى 240 ، 260 على الترتيب وبالتالي فإن الفئة (260- 240) هى فئة الوسيط ، أى الفئة التي يقع فيها قيمة الوسيط . وبالتالي نجد أن :

$$\text{بداية فئة } M_2 = 240$$

$$\text{ترتيب الوسيط } M_2 = 100$$

$$\text{التكرار المتجمع السابق} = 60$$

$$\text{تكرار فئة } M_2 = 100$$

$$\text{طول الفئة } M_2 = 260 - 240 = 20$$

وبالتعويض في المعادلة السابقة نجد أن :

$$M_2 = 240 + \left[\frac{100-60}{100}\right] \times 20 = 240 + 8 = 248 \text{ جنية}$$

<u>حل أخر</u>

يمكن تقدير قيمة الوسيط باستخدام المنحنى المتجمع الهابط على النحو التالي :

١- نكون التوزيع المتجمع الهابط كما هو موضح بالجدول التالي

جدول (٤-٨)

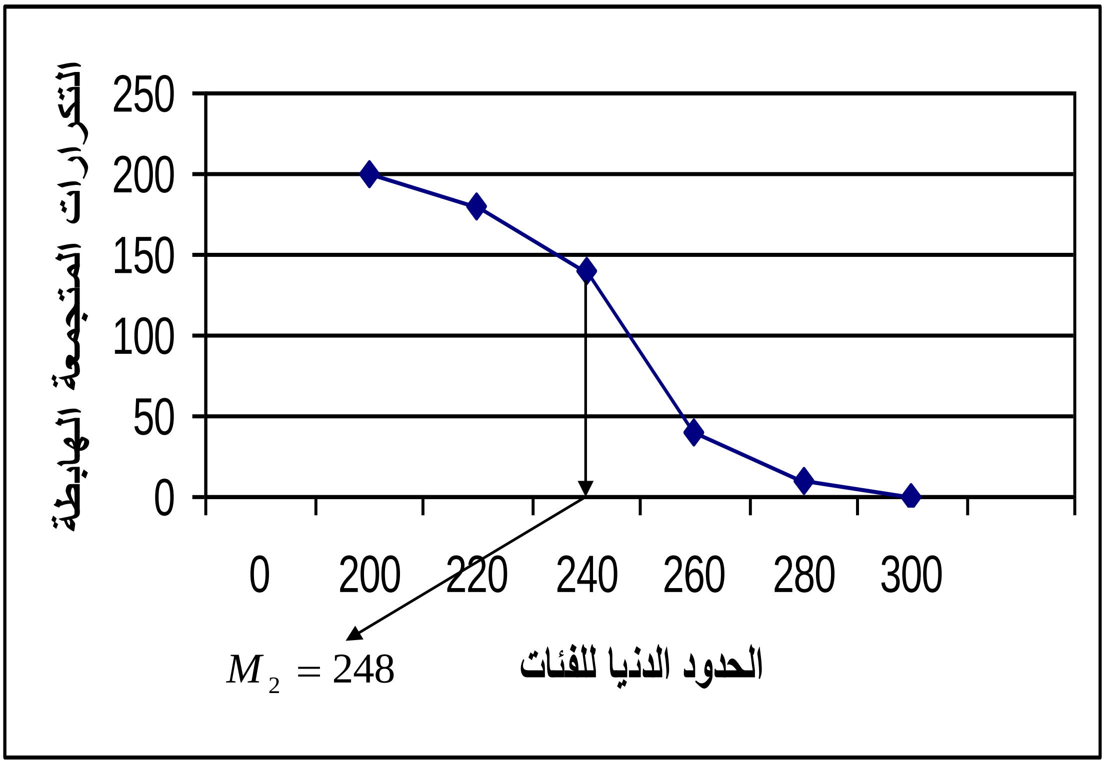
الحدود الدنيا للفئات	التكرارات المتجمعة الهابطة
أكبر من أو يساوى 200	200
أكبر من 220	180
أكبر من 240	140
أكبر من 260	40
أكبر من 280	10
أكبر من 300	0

بما أن ترتيب الوسيط يساوى

$$\text{ترتيب الوسيط} = \frac{\sum f_i}{2} = \frac{200}{2} = 100$$

٢- إذن ترتيب الوسيط يقع بين التكراران الهابطين 40, 140 على الترتيب في العمود الثاني بجدول (٤-٨) وبالتالي تصبح القيمة الوسيطة هى (260 - 240)

٣- برسم المنحنى المتجمع الهابط كما هو موضح بالشكل التالي :



شكل (٤-٢)

ومن الشكل نجد أن قيمة الوسيط تساوى 248 أى نفس القيمة التي تم تقديرها باستخدام المنحنى المتجمع الصاعد في الشكل (٤-١)

٤- ويمكن حساب الوسيط جبرياً باستخدام العلاقة (4.19) عن طريق التوزيع المتجمع الهابط ، فمن التوزيع بجدول (٤-٨) نجد أن :

ترتيب الوسيط $M_2 = 100$

الحد الأدنى للفئة $M_2 = 240$

الحد الأعلى للفئة $M_2 = 260$

طول فئة الوسيط $M_2 = 20$

التكرار الهابط السابق $= 14$

التكرار الهابط اللاحق $= 40$

وبالتالي فإن :
$$M_2 = 260 - \left[\frac{100-40}{100}\right] \times 20 = 248 \text{ جنية}$$

(٤-٤)المنوال **Mode**

في كثير من الظواهر يكون من الأهمية تحديد قيمة الظاهرة (التغير) التي تتكرر أكثر من أى قيمة أخرى القيمة الأكثر شيوعاً (أى الأكثر تكراراً)

تعريف (٤-٣)
المنوال هو القيمة الأكثر شيوعاً (أى الأكثر تكراراً).

<u>أولاً : المنوال من بيانات غير مبوبة</u>

أحسب القيمة المنوالية للبيانات التالية :

3, 10, 5, 3, 1, 4, 20, 3, 5, 12, 9, 3

بالنظر إلى قيم المفردات السابقة نجد أن القيمة 3 تتكرر أربع مرات وجميع القيم الأخرى تكرار كل منها أقل من 4 . وبالتالي تصبح القيمة 3 هى القيمة التي لها أكبر تكرار أى القيمة 3 هى القيمة المنوالية

<u>ثانياً : المنوال من بيانات مبوبة</u>

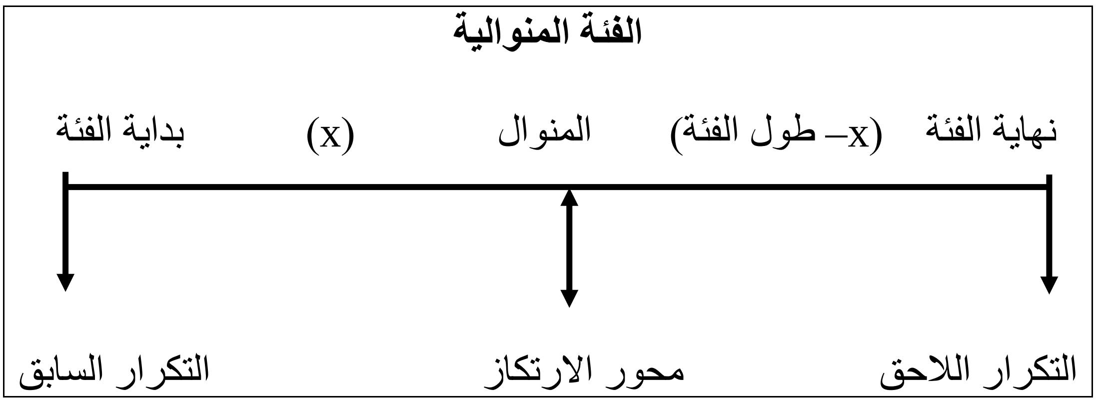
إذا كانت البيانات مبوبة في جدول تكراري فإنه يوجد طريقتين لإيجاد القيمة المنوالية وهما :

١- طريقة الرافعة ، ٢- طريقة التناسب

<u>أولاً : طريقة الرافعة</u>

تعتمد هذه الطريقة على تحديد الفئة التي لها أكبر تكرار فتكون هى الفئة المنوالية أى الفئة التي تقع فيها قيمة المنوال.

وتعتبر هذه الطريقة أن الفئة المنوالية رافعة من النوع الأول تؤثر عليها قوتين هما التكرار السابق لفئة المنوال والتكرار اللأحق لفئة المنوال.



شكل (٤-٣)

وأن المنوال يمثل محور الارتكاز وفي وضع التوازن يكون على بعد (x) من بداية الفئة كما هو موضح بالشكل (٤-٣) . وباستخدام قانوا الروافع نجد أن :

المقاومة × ذراعها= القوة × ذراعها

التكرار السابق × (x) = التكرار اللاحق × (طول الفئة - x)

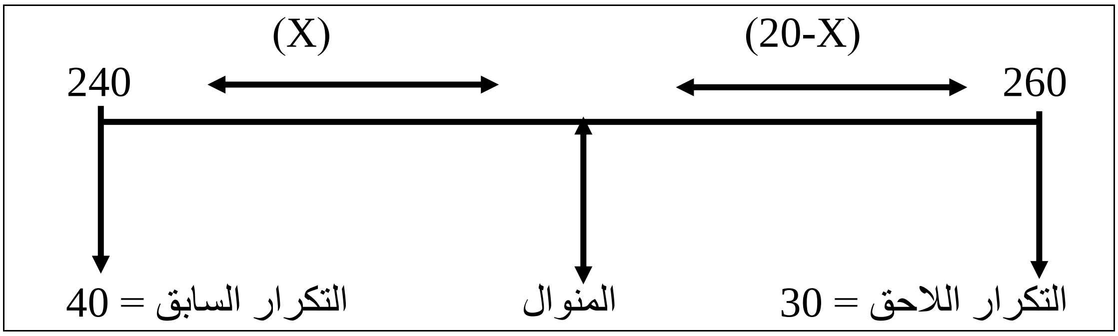
$$\text{المنوال} = \text{بداية الفئة المنوالية} + (x) \qquad (4.20)$$

حيث يتم حساب (x) من العلاقة (4.20) كما سوف يتضح في المثال التالي

مثال (٤-٩)
أحسب القيمة المنوالية للدخل الشهري للأسر في مثال (٤-٧)

الحل

من جدول (٤-٦) نجد أن أكبر تكرار يساوى 100 ومناظر للفئة (260 - 240) وبالتالي تصبح الفئة (260 - 240) الفئة المنوالية ، والشكل التالي يوضح الفئة المنوالية كرافعة من النوع الأول



شكل (٣-٤)

وبما أن : المقاومة × ذراعها= القوة × ذراعها

$$40\,(X) = 30\,(20 - X)$$
$$40\,X = 600 - 30\,X$$
$$70\,X = 600 \longrightarrow X=8.58$$

وبما أن :

$$\text{المنوال} = \text{بداية الفئة المنوالية} + (x)$$
$$= 240 - 8.58 = 248.58$$

ثانياً طريقة التناسب

وتعتمد هذه الطريقة على قانون التناسب التالي :

$$\frac{\text{بعد المنوال عن بداية الفئة } (X)}{\text{بعد المنوال عن نهاية الفئة (طول الفئة - } X)} = \frac{\text{التكرار السابق – تكرار الفئة المنوالية}}{\text{التكرار اللاحق – تكرار الفئة المنوالية}}$$

$$\text{المنوال} = \text{بداية الفئة المنوالية} + (x) \qquad (4.21)$$

فإذا اعتبرنا المثال السابق ، فنجد أن

40 = التكرار السابق
30 = التكرار اللاحق
100 = التكرار لفئة المنوال
20 = طول الفئة المنوالية

وباستخدام العلاقة (4.21) نجد أن

$$\frac{100-40}{100-30} = \frac{x}{20-x}$$

$$\frac{60}{70} = \frac{x}{20-x}$$

$$70x = 1200 - 60x$$

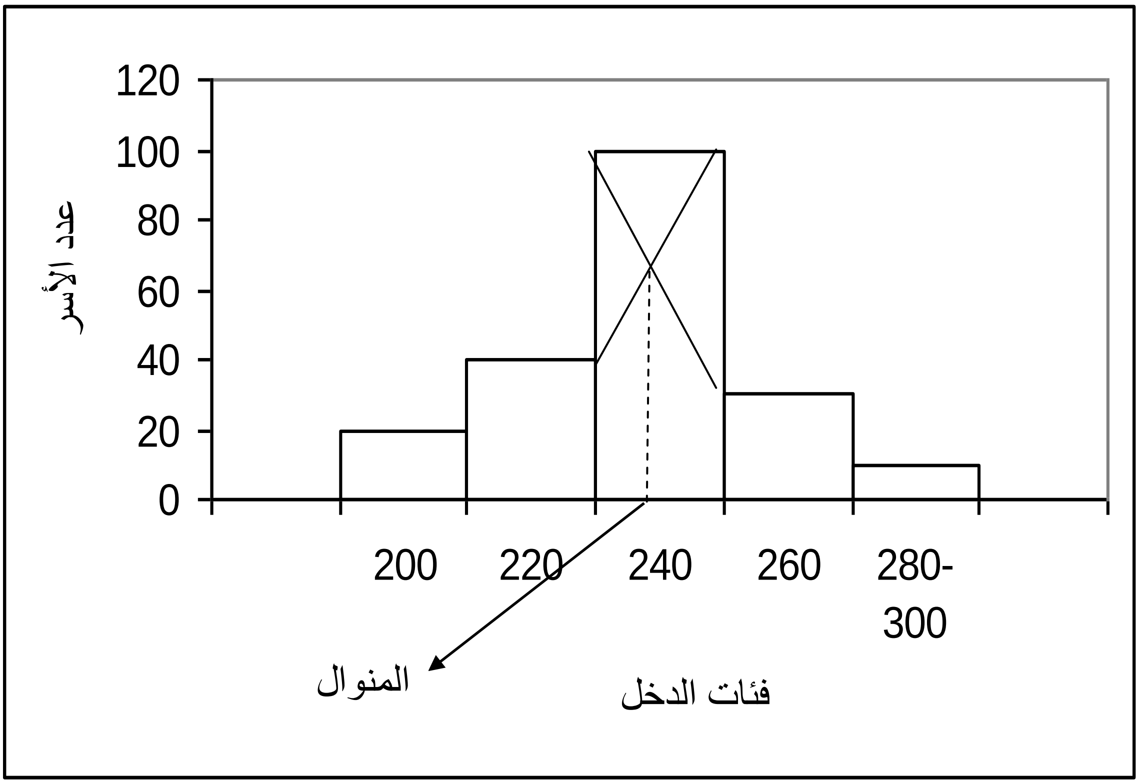
$$70x + 60x = 1200 \qquad \rightarrow 130x = 1200$$

$$\therefore x = \frac{1200}{130} = 9.23$$

وبالتالي فإن

المنوال = 240 + 9.23 = 249.23 جنية

ويمكن الحصول على نفس القيمة المنوالية التي حصلنا عليها بطريقة التناسب بيانياً ، وذلك برسم المدرج التكراري وتحديد الفئة المنوالية وهى تمثل قاعدة أطول عمود (أكبر تكرار) وبتوصيل الحد الأدنى للفئة المنوالية بالحد الأدنى للفئة اللاحقة وتوصيل الحد الأعلى للفئة المنوالية بالحد الأعلى للفئة السابقة ، كما هو موضح بالشكل التالي . فنقطة تقاطع العمود مع المحور الافقى هى القيمة المنوالية وتساوى 249.5 تقريباً كما هو موضح بالشكل



شكل (٤-٥)

ثالثاً : المنوال من جداول غير متساوية أطوال الفئات

بما أن التكرار المناظر لكل فئة يعتمد على كل من طول الفئة* وتكرارات القيمة الواحدة داخل كل فئة ، لذلك فإنه لحساب قيمة المنوال من جدول تكراري ذات فئات غير متساوية فإنه يجب أولاً التخلص من اثر إختلاف أطوال الفئات وذلك بإيجاد التكرارات المعدلة حيث أن :

$$\text{التكرار المعدل للفئة} = \frac{\text{التكرار الأصلي للفئة}}{\text{طول الفئة}}$$

(كما ذكرنا سابقاً في الباب السابق) تم استخدام التكرارات المعدلة بدلاً من الأصلية إيجاد قيمة المنوال بأى طريقة من الطرق السابق ذكرها أعلاه ، وسوف نوضح ذلك من خلال المثال التالي

* أ. د. نادية مكاري (١٩٧١) : مبادئ الإحصاء – كلية الاقتصاد والعلوم السياسية – جامعة القاهرة

مثال (٤-١٠)
الجدول التالي يوضح توزيع 148 طالب وفقاً لدرجاتهم في مادة الإحصاء

جدول (٤-٩)

درجة الطالب	2-	5-	10-	12-	15-20	المجموع
عدد الطلاب	18	40	40	30	20	148

أحسب القيمة المنوالية لدرجة الطالب في الإحصاء باستخدام طريقة التناسب ووضح ذلك بيانياً

الحل

١- من الجدول يتضح أن أطوال الفئات غير متساوية لذلك نحسب التكرارات المعدلة كما هو موضح بالجدول التالي :

جدول (٤-١٠)

الفئات	التكرارات	أطوال الفئات	التكرارات المعدلة
2-	18	3	18 / 3 = 6
5-	40	5	40 / 5 = 8
10-	40	2	40 / 2 = 20
12-	30	3	30 / 3 = 10
15-20	20	5	20 / 5 = 4

٢- وبما أن أكبر تكرار معدل يساوى 20 مناظر للفئة (10-12) أى أن الفئة المنوالية هى الفئة [10-12]، ويكون التكرار السابق المعدل يساوى 8 ، والتكرار اللاحق المعدل يساوى 10

وباستخدام العلاقة (4.21) نجد أن

$$\frac{20-8}{20-10} = \frac{x}{2-x}$$

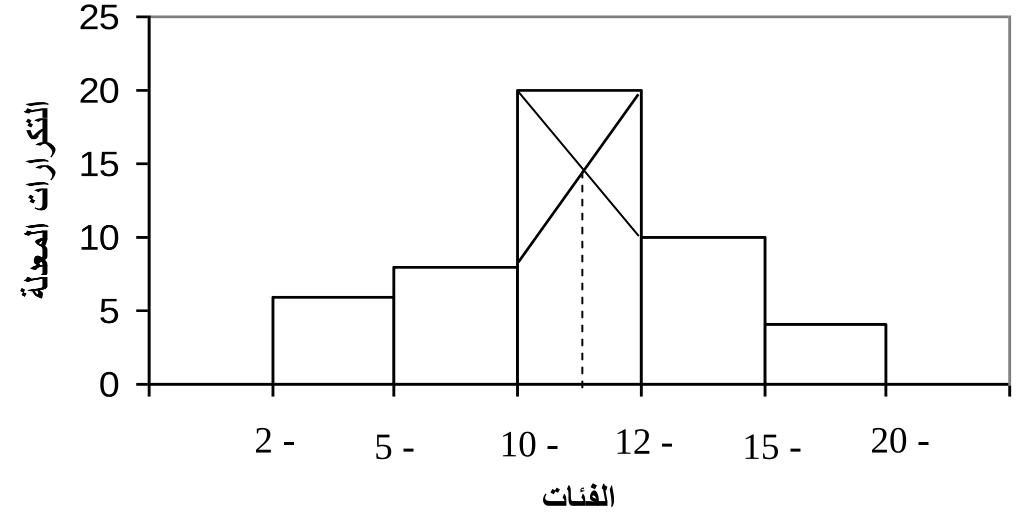
$$\frac{12}{10} = \frac{x}{2-x}$$

$$10x = 24 - 12x$$

$$22x = 24 \Rightarrow \therefore x = \frac{24}{22} = 1.1$$

المنوال $= 10 + 1.1 = 11.1 \; 11 \cong$ درجة

والشكل التالي يوضح القيمة المنوالية بيانياً



شكل (٤-٦)

ثانياً مقاييس التشتت

وكما ذكرنا سابقاً أن مقاييس التشتت هى المقاييس التي تقيس مدى تجانس قيم المتغير (الظاهرة) بعضها عن بعض أو مدى قرب أو بعد قيم المتغير عن قيمة أحد مؤشرات الموضع ، وسوف نعرض في الفصول التالية بعض هذه المؤشرات

(٤-٥) نصف المدى الربيعى **Semi- Interquarter Range**

في الفصول الثلاثة السابقة (٤-٢) – (٤-٤) تناولنا بالدراسة ثلاثة من أهم مقاييس (مؤشرات) الموضع وهم الوسط الحسابي ، الوسيط ، المنوال . وفي هذا الفصل والفصل التالي سوف نتناول بالدراسة بعض أهم مؤشرات مقاييس (مؤشرات) التشتت ، ففي هذا الفصل سوف نتناول نصف المدى الربيعي كمؤشرات لقياس التشتت بين القيم المختلفة للمتغير محل الدراسة .

تعريف (٤-٤)

المدى هو الفرق بين أكبر قيمة وأصغر قيمة للمتغير محل الدراسة

فإذا كان لدينا مجموعتين من الطلاب ودرجاتهم في مادة الإحصاء على النحو التالي :-

المجموعة الأولى : 0 ,7 ,12 ,15 ,20 ,14 ,17 ,10 ,15 ,11 ,10

المجموعة الثانية : 15 ,12 ,10 ,15 ,17 ,7 ,11 ,20 ,10 ,10 ,14

فنجد أن :

المدى بالنسبة للمجموعة الأولى $= 20 - 0 = 20$ درجة

درجات $10 = 10 - 20 =$ المدى بالنسبة للمجموعة الثانية

وبما أن المدى في المجموعة الأولى يساوى 20 درجة ، وفي المجموعة الثانية يساوى 10 درجات ، فهذا يدل على أن التشتت في المجموعة الأولى أكبر من التشتت في المجموعة الثانية.

ورغم أن المدى يعتبر مقياس للتشتت سهل الحساب إلا أنه يعتمد في تكوينه على قيمتين فقط هما أكبر قيمة وأصغر قيمة ويتجاهل باقي القيم ولاياخذ في الاعتبار باقي القيم ، بالإضافة إلى أن القيمتين المعتمد عليهما المدى في الحساب تعتبر قيم متطرفة (شاذة) وبالتالي فإن المدى مقياس لا يعكس التشتت الفعلى للقيم .

ففي المثال السابق يتضح أن الدرجات للمجموعتين واحدة باستثناء قيمة واحدة التي تمثل الحد الأدنى لكل مجموعة ، فالحد الأدنى في المجموعة الأولى (صفر) وفي المجموعة الثانية (10) ، فأدى ذلك إلى أن المدى في المجموعة الأولى يساوى ضعف المدى في المجموعة الثانية.

لذا يعتبر اعتماد المدى على قيمتين فقط وتجاهل باقى القيم بالإضافة إلى أن القيمتين التي يعتمد عليهما قيم متطرفة من العيوب الرئيسية التي أدت إلى البحث عن مقياس أخر لقياس التشتت لتلافي هذه العيوب

وفيما يلى سوف نقدم مقياس أخر للتشتت يعتمد على قيمتين فقط ولكنه يعتبر مقياس أفضل من المدى يسمى " نصف المدى الربيعي" حيث أنه يعتمد على قيمتين ولكنهما قيم غير شاذة حيث يتم استبعاد الربع الأول والأخير من القيم (بعد ترتيبها تصاعدياً أو تنازلياً) ثم مقارنة أصغر وأكبر قيمة من باقي قيم المجموعة . وبذلك نحصل على مقياس يسمى بالمدى الربيعي ومنتصف هذا المدى يسمى بنصف المدى الربيعي ويأخذ كمقياس للتشتت.

تعريف (٤-٥)
الربيع الأول هو القيمة التي 25% من القيم تأخذ قيم أقل منها أو تساويها ، وبالتالي 75% من القيم الباقية تأخذ قيم أكبر منها أو تساويها ، وعادة يشار إليها بالرمز (M_1) .

تعريف (٤-٦)
الربيع الثالث هو القيمة التي 75% من القيم تأخذ قيم أقل منها أو تساويها ، وبالتالي 25% من القيم الباقية تأخذ قيم أكبر منها أو تساويها ، وعادة يشار إليها بالرمز (M_3) .

ومن تعريف الوسيط (تعريف ٤-٢) والتعريفين السابقين ، نجد أن M_2 الوسيط هو الربيع الثاني .

أولاً : إيجاد نصف المدى الربيعي من بيانات غير مبوبة

لإيجاد نصف المدى الربيعي من بيانات غير مبوبة لابد من ترتيبها أولاً تصاعدياً (أو تنازلياً) ، فمثلاً إذا كانت لدينا القيم التالية :

6 ,5 ,7 ,10 ,0 ,8 ,13 ,15 ,20 ,3 ,11 ,14

فإن الترتيب التصاعدي لهذه القيم يصبح على النحو التالي :

0 ,3 ,5 ,6 ,7 ,8 ,10 ,11 ,13 ,14 ,15 ,20

ثم نحدد القيمة التي تقع في نهاية الربع الأول ولنرمز لها بالرمز (M_1) ، ثم نحدد القيمة التي تقع في نهاية الربع الثالث ولنرمز لها بالرمز (M_3)

ففي المثال السابق نجد أن M_1 تقع بين القيمتين التي ترتيبها ثلاثة والقيمة التي ترتيبها أربعة ، وبالتالي تكون قيمة M_1 متوسط القيمة التي ترتيبها ثلاثة والقيمة التي ترتيبها أربعة ، أى أن :

$$M_1 = \frac{5+6}{2} = 5.5$$

كذلك M_3 تقع بين القيمة التي ترتيبها تسعة والقيمة التي ترتيبها عشرة ، وبالتالي فإن :

$$M_3 = \frac{13-14}{2} = \frac{27}{2} = 13.5$$

ويكون المدى الربيعي هو الفرق بين M_1, M_3 ، أى أن :

$$\text{المدى الربيعي} = M_3 - M_1 \tag{4.22}$$

وبالتالي يكون

$$\text{نصف المدى الربيعي} = \frac{M_3 - M_1}{2} \tag{4.23}$$

وفي هذا المثال نجد أن :

$$\text{نصف المدى الربيعي} = \frac{13.5-5.5}{2} = \frac{8}{2} = 4$$

ويمكن تعميم الخطوات السابقة لحساب كل من M_1, M_3 في حالة وجود عدد n من القيم على النحو التالي :

١- ترتيب المفردات تصاعدياً (أو تنازلياً)

٢- إيجاد ترتيب الربيع الأول M_1 والربيع الثالث M_3 ، حيث أن

$$M_1 \text{ ترتيب } = \frac{\mathrm{n}+1}{4} \qquad (4.24)$$

إذا كان الترتيب تصاعدي

$$M_1 \text{ ترتيب } = \frac{(\mathrm{n}+1)\times 3}{4} \qquad (4.25)$$

إذا كان الترتيب تنازلي ، كذلك

$$M_3 \text{ ترتيب } = \frac{(\mathrm{n}+1)\times 3}{4} \qquad (4.26)$$

إذا كان الترتيب تصاعدي

$$M_3 \text{ ترتيب } = \frac{n+1}{4} \qquad (4.27)$$

إذا كان الترتيب تنازلي

٣- تحديد قيمة كل من M_1, M_3 من البيانات

٤- حساب نصف المدى الربيعي من العلاقة التالية

$$\text{نصف المدى الربيعي} = \frac{M_3 - M_1}{2} \qquad (4.28)$$

ثانياً : إيجاد نصف المدى الربيعي من بيانات مبوبة

يمكن إيجاد قيمة نصف المدى الربيعي بيانياً عن طريق تحديد قيمة كل من M_1, M_3 باستخدام المنحنى المتجمع الصاعد أو الهابط أو جبرياً بطريقة مشابهة لإيجاد قيمة الوسيط (M_2) بيانياً أو جبرياً حيث أن :

$$M_1 \text{ ترتيب } = \frac{\sum_{i=1}^{n} f_i}{4} \qquad (4.29)$$

عند استخدام التوزيع المتجمع الصاعد

$$\text{ترتيب } M_1 = \frac{3(\sum_{i=1}^{n} f_i)}{4} \qquad (4.30)$$

عند استخدام التوزيع المتجمع الهابط

وبالمثل :

$$\text{ترتيب } M_3 = \frac{3(\sum_{i=1}^{n} f_i)}{4} \qquad (4.31)$$

عند استخدام التوزيع المتجمع الصاعد

$$\text{ترتيب } M_3 = \frac{\sum_{i=1}^{n} f_i}{4} \qquad (4.32)$$

عند استخدام التوزيع المتجمع الهابط

ويتم حساب قيم كل من M_1, M_3 جبرياً باستخدام التوزيع المتجمع الصاعد على النحو التالي :

$$M_1 = \text{بداية فئة } M_1 + \frac{\text{ترتيب } M_1 - \text{التكرار الصاعد السابق لترتيب } M_1}{\text{تكرار فئة } M_1} \times \text{طول فئة } M_1 \qquad (4.33)$$

$$M_3 = \text{بداية فئة } M_3 + \frac{\text{ترتيب } M_3 - \text{التكرار الصاعد السابق لترتيب } M_3}{\text{تكرار فئة } M_3} \times \text{طول فئة } M_3 \qquad (4.34)$$

وعند استخدام التوزيع المتجمع الهابط فإنه يمكن حساب كل من M_1, M_3 باستخدام العلاقات التالية ·

$$M_1 = \text{نهاية فئة } M_1 + \frac{\text{ترتيب } M_1 - \text{التكرار الهابط اللاحق لترتيب } M_1}{\text{تكرار فئة } M_1} \times \text{طول فئة } M_1 \qquad (4.35)$$

$$M_3 = \text{نهاية فئة } M_3 + \frac{\text{ترتيب } M_3 - \text{التكرار الهابط اللاحق لترتيب } M_3}{\text{تكرار فئة } M_3} \times \text{طول فئة } M_3 \qquad (4.36)$$

وسوف نوضح ذلك من خلال الأمثلة التالية

مثال (٤- ١١)

الجدول التالي يوضح الإنتاج اليومي من إحدى السلع في 200 مصنع من المصانع التي تقوم بإنتاج هذه السلعة .

جدول (٤-١١)

عدد الوحدات المنتجة يومياً	100-	200-	300-	400-	500-600	المجموع
عدد المصانع	30	50	90	20	10	200

المطلوب

١- كون التوزيع المتجمع الصاعد ثم أرسم المنحنى المتجمع الصاعد .

٢- من الرسم أوجد قيمة M_1, M_3 ثم أحسب نصف المدى الربيعي .

٣- من (١) أحسب قيمة M_1, M_3 جبرياً ثم أحسب نصف المدى الربيعي .

٤- كون التوزيع المتجمع الهابط ثم أرسم المنحنى المتجمع الهابط .

٥- من (٤) أوجد قيمة M_1, M_3 ثم أحسب نصف المدى الربيعي .

٦- من (٤) أحسب قيمة M_1, M_3 جبرياً ثم أحسب قيمة نصف المدى الربيعي .

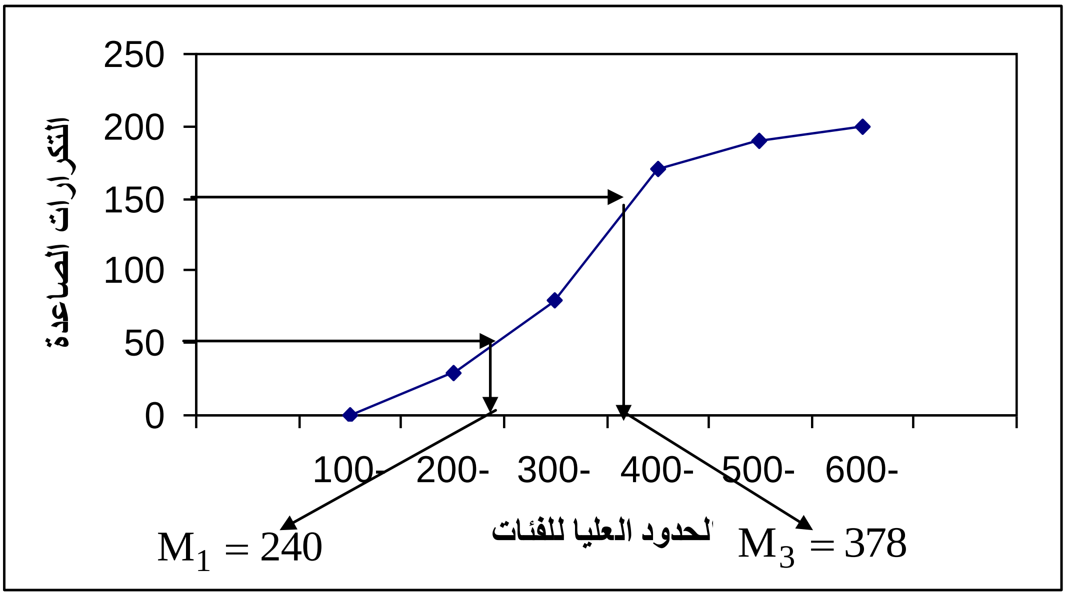
الحل

١- الجدول التالي يوضح التوزيع المتجمع الصاعد لعدد المصانع وفقاً لعدد الوحدات المنتجة يومياً .

جدول (٤-١٢)

الحدود العليا للفئات	التكرارات الصاعدة
أقل من 100	0
أقل من 200	30
أقل من 300	80
أقل من 400	170
أقل من 500	190
أقل من أو يساوي 600	200

والشكل التالي يوضح المنحنى المتجمع الصاعد للجدول السابق



شكل (٤-٧)

٢- بما أن ترتيب M_1س يساوى

$$\text{ترتيب } M_1 = \frac{\sum_{i=1}^{n} f_i}{4} = \frac{200}{4} = 50$$

وكذلك ترتيب M_3 يساوى

$$\text{ترتيب } M_3 = \frac{3(\sum_{i=1}^{n} f_i)}{4} = \frac{3(200)}{4} = 150$$

وبتحديد ترتيب كل من M_1, M_3 على المحور الرأسي في الشكل السابق ثم رسم مستقيمين موازيين للمحور الأفقي من هاتان النقطتان فيلاقيان المنحنى المتجمع الصاعد في نقطتين من هاتان النقطتان نسقط عمودين على المحور الأفقي ، فيكون تقاطع العمود الأول مع المحور الأفقي يعطى نقطة تمثل قيمة M_1 ، كذلك تقاطع العمود الثاني مع المحور الأفقي يعطى نقطة تمثل قيمة M_3 كما هو موضح في شكل (٤-٧) حيث نجد :

$$M_1 = 240 \qquad , \qquad M_3 = 378$$

٣- بما أن (ترتيب $M_1 = 50$) فإن فئة M_1 هي الفئة [300-200] والتكرار السابق يساوى 30 والتكرار اللاحق يساوى 80 حيث تكرار فئة $M_1 = 50$ ، وبالتعويض في العلاقة (4.33) , (4.34) نجد أن :

$$M_1 = 200 + \frac{50-30}{50} \times 100$$

$$= 200 + \frac{20}{50} \times 100 = 200 + 40 = 240 \quad \text{وحدة}$$

وبالمثل

$$M_3 = 300 + \frac{150-80}{90} \times 100$$

$$= 300 + \frac{70}{90} \times 100 = 377.8 \approx 378 \quad \text{وحدة}$$

وبالتالي فإن :

$$\text{نصف المدى الربيعي} = \frac{\mu_3 - \mu_1}{2} = \frac{377.8 - 240}{2} = 68.9 \approx 69 \quad \text{وحدة}$$

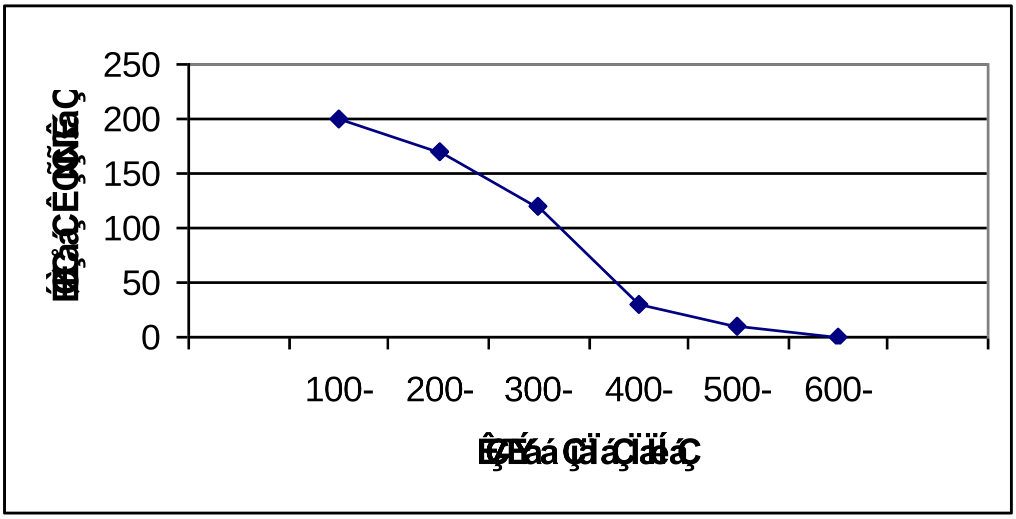
(4.37)

٤- الجدول التالي يوضح التوزيع المتجمع الهابط

جدول (٤-١٣)

التكرارات الصاعدة	الحدود العليا للفئات
200	أكبر من أو يساوى 100
170	أكبر من 200
120	أكبر من 300
30	أكبر من 400
10	أكبر من 500
0	أكبر من 600

والشكل التالي يوضح المنحنى المتجمع الهابط



شكل (٤-٨)

٥- بما أن

$$M_1 \text{ ترتيب } = \frac{3\left(\sum f_i\right)}{4} = \frac{3(200)}{4} = 150$$

$$M_3 \text{ ترتيب } = \frac{\sum f_i}{4} = \frac{200}{4} = 50$$

$$M_1 = 240 \qquad , \qquad M_3 = 378$$

وبالتالي نجد أن :

$$\text{نصف المدى الربيعي} = \frac{M_3 - M_1}{2} = \frac{378-240}{2} = \frac{138}{2} = 69 \text{ وحدة}$$

وباستخدام العلاقتين (4.35) ، (4.36) نجد أن :

$$M_1 = 300 - \left[\frac{150-120}{50}\right] \times 100$$

$$= 300 - \left[\frac{30}{50}\right] \times 100 = 240 \text{ وحدة}$$

$$M_3 = 400 - \left[\frac{50-30}{90}\right] \times 100$$

$$= 400 - 22.22 = 377.8 \approx 378 \quad \text{وحدة}$$

وبالتالي يصبح :

$$\frac{M_3 - M_1}{2} = \frac{377.8 - 240}{2} = 68.9 \approx 69 \quad \text{وحدة تقريباً}$$

ملاحظات :

١- قيم كل من M_3, M_1 التي يتم إيجادهم باستخدام التوزيع المتجمع الصاعد هى نفس القيم التي يتم إيجادها باستخدام التوزيع المتجمع الهابط.

٢- قيمة M_3 أكبر من قيمة M_1 .

(٤-٦) التباين والانحراف المعياري
Variance and Standard Deviation

تناولنا في الفصل السابق المدى ونصف المدى الربيعي كمقياسين من مقاييس التشتت حيث أتضح أن نصف المدى الربيعي مقياس للتشتت أفضل من المدى حيث أنه لا يعتمد على القيم المتطرفة (الشاذة) ولكنة ما زال يعتمد على قيمتين فقط هما M_3, M_1 ولا يأخذ في الاعتبار باقي قيم الظاهرة (المتغير).

وفي هذا الفصل سوف نتناول التباين ونشتق منه الانحراف المعياري كمقياس للتشتت يعتبر أفضل من نصف المدى الربيعي ، حيث أنه يأخذ في الاعتبار جميع قيم الظاهرة (المتغير) بالإضافة إلى مجموعة أخرى من الخصائص الرياضية Mathematical Properties التي يتميز بها هذا المقياس عن المقاييس الأخرى للتشتت.

أولاً : إيجاد التباين والانحراف المعياري من بيانات غير مبوبة

إذا كان لدينا مجتمع مكون من N من المفردات قيمتها على النحو التالي :

$$x_1, x_2, x_3, x_4, \ldots\ldots, x_N$$

فإن التباين لقيم المتغير x ونرمز له بالرمز σ^2 هو عبارة عن متوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي ، ويمكن حسابه من العلاقة التالية :

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x-\mu)^2}{N} \qquad (4.37)$$

كذلك يمكن إثبات أن :

$$\sigma^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} - (\mu)^2 \qquad (4.38)$$

أما إذا كانت البيانات تمثل بيانات عينة حجمها n فإننا نرمز للتباين المحسوب من بيانات عينة بالرمز S^2 حيث :

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \qquad (4.39)$$

كذلك يمكن إثبات أن :

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} \qquad (4.40)$$

ويعرف الانحراف المعياري كمقياس للتشتت بأنه الجذر التربيعي الموجب للتباين أى أن σ هي الانحراف المعياري للظاهرة في المجتمع كذلك S هي الانحراف المعياري للظاهرة في العينة.

مثال (٤-١٢)

إذا كان لدينا مجتمع مكون من 10 مفردات قيمتها :

2 , 7 , 11 , 10 , 15 , 5 , 12 , 8 , 13 , 17

فلحساب التباين σ^2 نتبع الخطوات التالية :

١- نوجد الوسط الحسابي لهذه القيم μ حيث

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{2+7+11+10+15+5+12+8+13+17}{10} = 10$$

٢- نحسب انحرافات القيم عن وسطها الحسابي μ أى نحسب $(x_i - \mu)$ ثم نحسب مربعات هذه الانحرافات على النحو التالي :

جدول (٤-١٤)

x_i	$(x_i-\mu)$	$(x_i-\mu)^2$
2	2 -10 = -8	64
7	7 - 10 = -3	9
11	11 – 10 = 1	1
10	10 – 10 = 0	0
15	15 – 10 = 5	25
5	5 – 10 = 5	25
12	12 – 10 = 2	4
8	8 - 10 = -2	4
13	13 – 10 = 3	9
17	17 – 10 = 7	49
$\sum x_i = 100$	$\sum(x_i-\mu)=0$	$\sum(x_i-\mu)^2 = 190$

$$\therefore \sigma^2 = \frac{\sum_{i=1}^{N}(x-\mu)^2}{N} = \frac{190}{10} = 19$$

وبالتالي فإن :

$$\text{الانحراف المعياري} = \sigma = \sqrt{\sigma^2} = \sqrt{19} = 4.3$$

ثانياً : التباين والانحراف المعياري من بيانات مبوبة

إذا كـان x_i ، f_i تشـير إلـى مركـز الفئـة رقـم (i) والتكـرار المنـاظر لهـا حيث $i = 1,2,3,\ldots,n$ ، فإن التبـاين σ^2 للمتغير محل الدراسة في المجتمع يصبح على النحو التالي :

$$\sigma^2 = \frac{\sum\{(x_i-\mu)^2 f_i\}}{\sum f_i} \qquad (4.41)$$

كذلك يمكن إثبات أن :

$$\sigma^2 = \frac{\sum x_i^2 f_i}{\sum f_i} - (\mu)^2 \tag{4.42}$$

وإذا كانت البيانات بيانات عينة فإن تباين المتغير في العينة يحسب على النحو التالي :

$$S^2 = \frac{\sum\{(x_i - \bar{x})^2 f_i\}}{(\sum f_i) - 1} \tag{4.43}$$

كذلك يمكن إثبات أن :

$$S^2 = \frac{\sum x_i^2 f_i - \frac{(\sum x_i f_i)^2}{\sum f_i}}{\sum f_i - 1} \tag{4.44}$$

وسوف نوضح ذلك من خلال الأمثلة التالية :

مثال (٤-١٣)

في دراسة عن معرفة التشتت بالنسبة للطول (بالسنتيمتر) بالنسبة لـ 100 طالب من المتقدمين لإحدى الكليات العسكرية ، ووجد أن توزيع الطلاب وفقاً للطول كما هو موضح بالجدول التالي :

جدول (٤-١٥)

فئات الطول	150-	160-	170-	180-	190-200	المجموع
عدد المتقدمين	5	15	50	20	10	100

أحسب الانحراف المعياري للطول

الحل

١- نحسب مراكز الفئات x_i

٢- نوجد الوسط الحسابي μ حيث أن ($\mu = 176.5$)

٣- نكون الجدول التالي :

جدول (٤-١٦)

الفئات	التكرارات (f_i)	مراكز الفئات (x_i)	$(x_i - \mu)^2$	$(x_i - \mu)^2 f_i$
150-	5	155	462.25	2311.25
160-	15	165	132.25	1983.75
170-	50	175	2.25	112.50
180-	20	185	27.25	1445.00

190-200	10	195	342.25	3422.50
المجموع	100			9375.00

وباستخدام العلاقة (3.41) نجد أن :

$$\sigma^2 = \frac{\sum\{(x_i-\mu)^2 f_i\}}{\sum f_i} = \frac{9375}{100} = 93.75$$

وبالتالي فإن :

$$\text{الانحراف المعياري} = \sigma = \sqrt{\sigma^2} = \sqrt{93.75} = 9.62$$

مثال (٤-١٤)

أخذت عينة مكونة من 1000 طفل من الأطفال من سن شهر إلى سن عام وتم توزيعهم وفقاً لفئات الوزن بالكيلوجرام .

جدول (٤-١٧)

فئات الوزن بالكيلوجرام	2-	4-	6-	8-	10-12	المجموع
عدد الأطفال	20	180	500	200	100	1000

المطلوب

١- أوجد متوسط وزن الطفل

٢- أحسب الانحراف المعياري لوزن الطفل

الحل

نكون الجدول التالي :

جدول (٤-١٨)

الفئات	f_i	x_i	$x_i f_i$	$x_i^2 f_i$
2-	20	3	60	180
4-	180	5	900	4500
6-	500	7	3500	24500
8-	200	9	1800	16200
10-12	100	11	1100	12100
المجموع	1000		7360	57480

وبالتعويض من الجدول في العلاقة (4.44) نجد أن :

$$S^2 = \frac{\sum x_i^2 f_i - \frac{\left(\sum x_i f_i\right)^2}{\sum f_i}}{\sum f_i - 1}$$

$$= \frac{57480 - \frac{(7360)^2}{1000}}{1000 - 1} = \frac{57480 - 54169.6}{999} = \frac{3310.4}{999} = 3.31$$

وبالتالي فإن :

الانحراف المعياري $= S = \sqrt{S^2} = \sqrt{3.31} = 1.82$ كيلوجرام

(٤-٧) معاملات الاختلاف **Coefficients of Variation**

في الفصلين السابقين تناولنا المدى ونصف المدى الربيعي والانحراف المعياري كمقاييس للتشتت ولاحظنا أن كل مؤشر من هذه المؤشرات له وحدة قياس هى نفسها وحدة قياس المتغير محل الدراسة .

وتسمى مؤشرات التشتت السابقة بالمؤشرات المطلقة (أو المقاييس المطلقة) ، وعند استخدامها في مقارنة تشتت مجموعتين لابد أن يكون للمجموعتين نفس وحدة القياس . فمثلاً إذا كان لدينا مجموعتين من العاملين وأخذت أوزانهم كمتغير فكان الانحراف المعياري للوزن في المجموعة الأولى يساوى 5 كيلوجرام ، والانحراف المعياري للوزن في المجموعة الثانية 3 كيلوجرام فإنه يتضح أن التشتت في الوزن في المجموعة الأولى أكبر منه في المجموعة الثانية.
ومقاييس التشتت المطلقة لا يمكن استخدامها لمقارنة تشتتي مجموعتين لهما وحدات قياس مختلفة ، فمثلاً مقارنة التشتت للدخول الشهرية لمجموعة من الأفراد بالجنية يتشتت أوزانهم بالكيلوجرام ، فوحدة القياس للمتغير الذى يمثل الدخل الشهري هى الجنية ووحدة القياس للمتغير الذى يمثل الوزن هى الكيلوجرام .
ومن هنا نجد أن وحدات القياس مختلفة ، وبالتالي عند مقارنة تشتتي المتغيرين فإنه لا تصلح مقاييس التشتت المطلقة لإجراء المقارنة . وفي هذه الحالة نستخدم مقاييس أخرى تسمى معاملات الاختلاف.
وفيما يلي سوف نقدم مقياسين من المقاييس المستخدمة لقياس التشتت وإجراء المقارنات وهما :-

١- معامل الاختلاف المعياري Relative Standard Deviation
٢- معامل الاختلاف الربيعي Coefficient of Quartile Deviation

أولا : معامل الاختلاف المعياري Relative Standard Deviation

إذا رمزنا لمعامل الاختلاف المعياري بالرمز RS فإن :

$$RS = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100 \qquad (4.45)$$

أى أن :

بالنسبة للمجتمع
$$RS = \frac{\sigma}{\mu} \times 100$$

أو

بالنسبة للعينة
$$RS = \frac{S}{\bar{x}} \times 100$$

ثانياً : معامل الاختلاف الربيعي Coefficient of Quartile Deviation

إذا رمزنا لمعامل الاختلاف الربيعي بالرمز QD فإن :

$$QD = \frac{\text{نصف المدى الربيعي}}{\text{الوسيط}} \times 100 \qquad (4.47)$$

ويمكن إثبات أن :

$$QD = \frac{M_3 - M_1}{M_3 + M_1} \times 100 \qquad (4.48)$$

ملحوظة :

معاملات الاختلاف RS ، QD هي نسب مئوية وليس لها تمييز.

مثال (٤-١٥)

إذا كان متوسط الأجر الشهري لـ 100 عامل في القطاع العام 500 جنية ، بانحراف معياري 70 جنية ، ومتوسط الأجر الشهري لـ 100 عامل آخرين بإحدى الشركات الأجنبية 400 دولار بانحراف معياري 20 دولار.

المطلوب :

قارن بين التشتت الأجرين

الحل

في هذه الحالة نجد أن وحدة القياس للأجر للمجموعة الأولى بالجنية ، وللمجموعة الثانية بالدولار . بالتالي لا يمكن مقارنة التشتت باستخدام الانحراف المعياري كمقياس مطلق .

ولذا نستخدم معامل الاختلاف المعياري فنجد أن معامل الاختلاف المعياري للمجموعة الأولى $(RS)_1$ حيث :

$$(RS)_1 = \frac{S}{\bar{x}} \times 100 = \frac{70}{500} \times 100 = 14\%$$

ومعامل الاختلاف المعياري للمجموعة الثانية $(RS)_2$ حيث :

$$(RS)_2 = \frac{S}{\bar{x}} \times 100 = \frac{20}{400} \times 100 = 5\%$$

وبالتالي نجد أن أجور العمال في المجموعة الثانية (أى عمال الشركات الأجنبية) أكثر تجانس من أجور العمال بالقطاع العام ، أو بعبارة أخرى أجور العمال في القطاع العام أكثر تشتت من العاملين في الشركة الأجنبية .

مثال (٤-١٦)

إذا كان الربيع الأول والثالث لأوزان مجموعة من الأفراد هما$(M_1 = 20)$ ، $(M_3 = 45)$ بالكيلوجرام . كذلك تم حساب الربيع الأولى والثالث لأطوال هؤلاء الأفراد فكانا $M_1 = 150$ ، $M_3 = 165$.

المطلوب :

قارن بين الوزن والطول لهؤلاء الأفراد من حيث التشتت .

الحل

إذا فرضنا أن $(QD)_1, (QD)_2$ هما معامل الاختلاف الربيعي للوزن ومعامل الاختلاف الربيعي للطول فنجد أن :

$$(QD)_1 = \frac{M_3 - M_1}{M_3 + M_1} \times 100 = \frac{45 - 20}{45 + 20} \times 100 = 38.46\%$$

$$(QD)_2 = \frac{M_3 - M_1}{M_3 + M_1} \times 100 = \frac{165 - 150}{165 + 150} \times 100 = 4.76\%$$

وبمقارنة $(QD)_1, (QD)_2$ نجد أن أوزان هؤلاء الأفراد أكثر تشتت من أطوالهم.

(٤-٨) أمثلة تطبيقية Applied Examples

تطبيق (٤-١)

فيما يلي بيانات عن سعر الوحدة الواحدة من إحدى السلع الاستهلاكية في مدينتي القاهرة والإسكندرية خلال عام ١٩٩٧ (بالجنية)

جدول (٤-١٩)

الشهر	سعر الوحدة بمدينة الإسكندرية	سعر الوحدة بمدينة القاهرة
يناير	12.5	10.1
فبراير	13.0	12.5
مارس	13.5	11.6
إبريل	11.5	14.0
مايو	9.2	8.0
يونيو	10.0	13.6
يوليو	12.5	13.0
أغسطس	14.0	12.5
سبتمبر	17.0	13.0
أكتوبر	15.0	14.0
نوفمبر	9.0	9.4
ديسمبر	8.0	9.0

المطلوب :

١- أحسب الوسط الحسابي لسعر الوحدة في كل من المدينتين.
٢- أحسب الانحراف المعياري للسعر في المدينتين ، ثم قارن التشتت للسعر في المدينتين.

الحل

١- إذا رمزنا للوسط الحسابي للسعر في الإسكندرية والقاهرة بالرمز $\bar{x}_2, \bar{x}_1$ على الترتيب فإن :

$$\bar{x}_1 = \frac{12.5+13+13.5+11.5+9.2+10+12.5+14+17+15+9+8}{12}$$

$$= \frac{145.2}{12} = 12.1 \text{ جنية}$$

$$\bar{x}_2 = \frac{10.1+12.5+11.6+14+8+13.6+13+12.5+13+14+9.4+9}{12}$$

$$= \frac{140.7}{12} = 11.73 \quad \text{جنية}$$

وبما أن $\bar{x}_1 = 12.1$, $\bar{x}_2 = 11.73$ ، وبالتالي نجد أن متوسط سعر الوحدة في مدينة القاهرة على مدار هذا العام أقل منه في مدينة الإسكندرية.

٢- إذا فرضنا أن σ_1^2 , σ_2^2 هما تباين السعر في الإسكندرية والقاهرة على الترتيب ، وبما أن :

$$\sigma_1^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x}_1)^2}{N} = \frac{\sum_{i=1}^{N} x_1^2}{N} - (\bar{x}_1)^2$$

$$= \frac{1835.64}{12} - (12.1)^2 = 152.97 - 146.41 = 6.56$$

وبالتالي فإن :

$$\sigma_1 = \sqrt{\sigma_1^2} = \sqrt{6.56} = 2.6 \quad \text{جنية}$$

كذلك

$$\sigma_2^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x}_2)^2}{N} = \frac{\sum_{i=1}^{N} x_2^2}{N} - (\bar{x}_2)^2$$

$$= \frac{1697.39}{12} - (11.73)^2 = 141.45 - 137.59 = 3.86$$

وبالتالي فإن :

$$\sigma_2 = \sqrt{\sigma_2^2} = \sqrt{3.86} = 1.96 \quad \text{جنية}$$

وبما أن $\sigma_1 = 2.6$, $\sigma_2 = 1.96$ ، إذن سعر الوحدة من السلعة على مدار العام أكثر تجانس في مدينة القاهرة عنه قي مدينة الإسكندرية.

تطبيق (٤-٢)

الجدول التالي يوضح قيمة فاتورة التليفون في يناير ١٩٩٨ لـ 1000 أسرة بإحدى أحياء القاهرة.

جدول (٤-٢٠)

قيمة الفاتورة بالجنية	عدد الأسر
أقل من 50	0
أقل من 70	150
أقل من 90	370
أقل من 110	800
أقل من 130	950
أقل من أو يساوى 150	1000

المطلوب :

١- أحسب متوسط قيمة الفاتورة للأسرة.
٢- أحسب التباين والانحراف المعياري ، ثم أحسب معامل الاختلاف.

الحل

من الجدول السابق يمكن تكوين جدول التوزيع التكراري لقيمة الفاتورة على النحو التالي :

جدول (٤-٢١)

قيمة الفاتورة	50-	70-	90-	110-	130-150	المجموع
عدد الأسر	150	220	430	150	50	1000

ولحساب الوسط الحسابي (متوسط) الفاتورة والانحراف المعياري نكون الجدول التالي :

جدول (٤-٢٢)

الفئات	التكرارات (f_i)	مراكز الفئات (x_i)	$x_i f_i$	$x_i^2 f_i$
50-	150	60	9000	540000
70-	220	80	17600	1408000
90-	430	100	43000	4300000
110-	150	120	18000	2160000
130-150	50	140	7000	980000
المجموع	1000		94600	9388000

$$1-\bar{x}=\frac{\sum x_i f_i}{\sum f_i}=\frac{94600}{1000}=94.6$$

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

$$=\frac{9388000}{1000}-(94.6)^2=9388-8949.16=438.84$$

$$\therefore \sigma=\sqrt{\sigma^2}=\sqrt{438.84}=20.95 \text{ جنية}$$

وبالتالي فإن :

$$\text{معامل الاختلاف المعياري}=\frac{\sigma}{\bar{x}}\times 100=\frac{20.95}{94.6}\times 100=22.15\%$$

تطبيق (٤-٣)

إذا كان متوسط الأجر الشهري لعمال أحد المصانع يساوى 300 جنية شهرياً بانحراف معياري 15 جنية ، ومتوسط عمر العامل في هذا المصنع 30 سنة بانحراف معياري 15 سنة.

المطلوب :

هل الاختلاف النسبي Relative Variation في الأجر أكبر من الاختلاف النسبي في أعمار العمال بهذا المصنع.

الحل

إذا فرضنا أن $\bar{x}_2, \bar{x}_1$ هما متوسط الأجر والعمر للعامل على الترتيب ، وأن σ_2, σ_1 هما الانحراف المعياري للأجر والعمر على الترتيب أيضاً.

وبما أن $\sigma_1 = \sigma_2 = 15$ ، فهذا لا يعني أن التشتت في الأجر مكافئ للتشتت في أعمار العاملين . لذلك نحسب التشتت النسبي لكل من الأجر والعمر باستخدام معامل الاختلاف للأجر ومعامل الاختلاف للعمر على النحو التالي :

$$\text{معامل الاختلاف المعياري للأجر}=\frac{\sigma_1}{\bar{x}_1}\times 100=\frac{15}{300}\times 100=5\%$$

$$\text{معامل الاختلاف المعياري للعمر}=\frac{\sigma_2}{\bar{x}_2}\times 100=\frac{15}{30}\times 100=50\%$$

وبما أن معامل الاختلاف للأجر يساوى 5% ، ومعامل الاختلاف للعمر يساوى 50% فمن هنا يتضح أن التشتت في أعمار العمال أكبر من التشتت في أجورهم.

تطبيق (٤-٤)

فيما يلي بيانات عن قيمة الاستهلاك الشهري لـ 10 أسر بالجنية :

x_i: 610 , 600 , 650 , 612 , 618 , 638 , 622 , 610 , 640 , 1500

المطلوب :

١- أوجد متوسط الاستهلاك الشهري للأسرة الواحدة ، وهل يمكن القول بأن هذا المتوسط ممثل جيد للاستهلاك في هذه الحالة.

٢- أستبعد $x_{10} = 1500$ ، ثم أحسب متوسط الاستهلاك الأسرة وقارنه بالمتوسط في (١) ـ ثم عقب على الناتج.

٣- أحسب القيمة المنوالية للاستهلاك.

الحل

$$\bar{x} = \frac{\sum x_i}{n}$$

$$= \frac{610 + 600 + 650 + 612 + 618 + 638 + 622 + 610 + 640 + 1500}{10}$$

$$= \frac{7100}{10} = 710 \quad \text{جنية}$$

وبما أن قيمة المتوسط $x_i = 710$ فهي لا تمثل استهلاك الأسر في هذه الحالة.

٢- إذا تم استبعاد القيمة $x_{10} = 1500$ ، ونحسب الوسط الحسابي في هذه الحالة نجد أن :

$$\bar{x} = \frac{\sum x_i}{n}$$

$$= \frac{610 + 600 + 650 + 612 + 618 + 638 + 622 + 610 + 640}{9}$$

$$= \frac{5600}{9} = 622.2 \quad \text{جنية}$$

وبما أن $\bar{x} = 622.2$ جنية ، فهي قيمة تتوسط جميع قيم المفردات من $x = x_1$ إلى $x = x_9$ وتمثل الاستهلاك تمثيل جيد.

من (١) ، (٢) يتضح أن المتوسط $\bar{x}$ في حالة أخذ القيم الشاذة (المتطرفة) $x_{10} = 1500$ في الاعتبار وجدنا أن قيمة المتوسط لا تعبر عن قيم المتغير تعبير

جيد ، في حين عند استبعاد القيمة الشاذة $x_{10} = 1500$ ثم حساب المتوسط بعد استبعادها فكانت قيمة المتوسط $\bar{x} = 622.2$ تعبر عن قيم المتغير (الاستهلاك) تعبير جيد . ومن هنا يمكن القول بأن المتوسط (الوسط الحسابي) يتأثر بالقيم الشاذة ، وبالتالي في حالة وجود قيم شاذة فإنه يفضل استخدام مقياس أخر غير الوسط الحسابي مثل الوسيط أو حساب المتوسط بعد استبعاد هذه القيم المتطرفة.

٣- بالرجوع إلى القيم x_i نجد أنه لا يوجد أي قيمة لها تكرار أكبر من القيم الأخرى ، وبالتالي في هذه الحالة فإنه لا يوجد قيمة منوالية للاستهلاك.

تطبيق (٤-٥)

ترغب إحدى شركات الاتصالات التليفونية في دراسة جدوى التوسع في الخدمات الخارجية للعملاء فقامت الشركة بإجراء دراسة على عينة من 80 مشترك في الخدمة الخارجية من خلال تسجيل قيمة أول فاتورة لاشتراك كل منهم (مدة الفاتورة ستة شهور) بالجنيه المصري . فكانت البيانات على النحو التالي :

800	210	510	910	218	200	900	200	212	240
710	270	620	700	511	210	910	311	315	250.75
510	910	830	620	370	710	581	375	310	278
370	600	570	510	680	510	811	400	450	1070
300	250	601	411	750	211	920	410	611	560.25
400	200	805	270	910	310	368	1000	920	1200
470	280	850	380	1000	1000	273	1002	815	270
500	300	400	610	500	1100	485	515	700	515

فإذا كان عدد المشتركين في الشركة 60 الف مشترك منهم 4 ألاف مشترك في الخدمات الخارجية.

المطلوب :

١- أوصف المجتمع الإحصائي محل الدراسة.
٢- أوصف العينة.
٣- أوصف المتغير محل الدراسة.
٤- أعرض البيانات السابقة في جدول تكراري بسيط مناسب.
٥- أرسم المدرج التكراري ومن الرسم أوجد القيمة المنوالية.
٦- أوجد كل من : الوسط الحسابي والانحراف المعياري.
٧- أحسب معامل الاختلاف ثم عقب على الناتج.

الحل

١- المجتمع الإحصائي محل الدراسة هو العملاء المشتركين في الخدمات الخارجية في الشركة وعددهم 4000 مشترك

٢- العينة هي 80 عميل من المشتركين في الخدمات الخارجية بالشركة

٣- المتغير محل الاهتمام هو : قيمة الفاتورة الأولى (أول ست شهور للاشتراك في المكالمات الخارجية) بالجنية المصري.

٤- لتكوين الجدول التكراري البسيط نتبع الخطوات التالية:

- تحديد أقل قيمة (200 جنية) ، وتحديد أكبر قيمة (1200 جنية).
- نحسب المدى حيث :

أقل قيمة – أكبر قيمة = المدى

جنية 1000 = 200 – 1200 =

- نفرض أن عدد الفئات يساوى 10 بالتالي فإن :

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{1000}{10} = 100 \text{ جنية}$$

- نكون جدول التفريغ التالي :

جدول(٤-٢٣) يوضح تفريغ البيانات لتكوين جدول توزيع تكراري

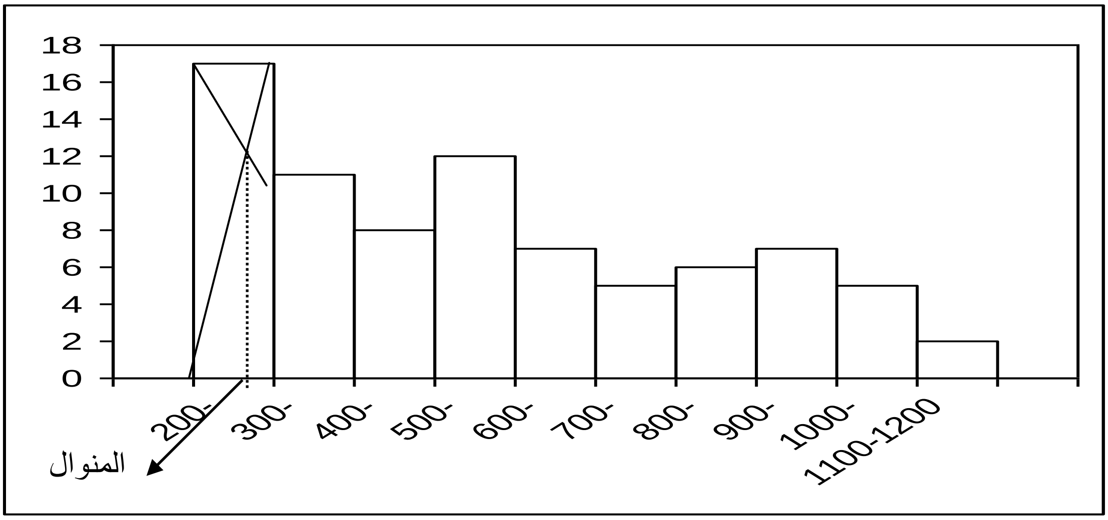
الفئات	عملية التفريغ	التكرارات
200-	// ~~////~~ ~~////~~ ~~////~~	17
300-	/ ~~////~~ ~~////~~	11
400-	/// ~~////~~	8
500-	// ~~////~~ ~~////~~	12
600-	// ~~////~~	7
700-	~~////~~	5
800-	/ ~~////~~	6
900-	// ~~////~~	7
1000-	~~////~~	5
1100-1200	//	2
المجموع		80

ومن الجدول السابق نكون الجدول التكراري التالي :

جدول (٤-٢٤) يوضح توزيع 80 عميل وفقاً لقيمة أول فاتورة
بالجنية المصري للمكالمات الخارجية

قيمة الفاتورة	200-	300-	400-	500-	600-	700-	800-	900-	1000-	1100-1200
عدد المشتركين	17	11	8	12	7	5	6	7	5	2

٥- الشكل التالي يوضح المدرج التكراري للتوزيع التكراري السابق



٦- لحساب كل من الوسط الحسابي والانحراف المعياري نكون الجدول التالي :

جدول (٤-٢٥)

الفئات	التكرارات f	مراكز الفئات x	$y = x - 750$	$h = \frac{y}{100}$	$f\ h$	$f\ h^2$
200-	17	250	-500	-5	-85	425
300-	11	350	-400	-4	-44	176
400-	8	450	-300	-3	-24	72
500-	12	550	-200	-2	-24	48
600-	7	650	-100	-1	-7	7
700-	5	750	0	0	0	0
800-	6	850	100	1	6	6
900-	7	950	200	2	14	28
1000-	5	1050	300	3	15	45
1100-1200	2	1150	400	4	8	32
المجموع	80				-141	839

$$\bar{h} = \frac{\sum f\ h}{\sum f - 1} = \frac{-141}{80-1} = \frac{-141}{79} = -1.77$$

وبما أن :

$$\bar{X} = a + b\bar{h}$$

فإن :

$$\bar{X} = 750 + 100\bar{h} = 750 + 100(-1.77)$$
$$= 750 - 177 = 573 \text{ جنية}$$

وبما أن :

$$S_h^2 = \frac{\sum h^2 f - \frac{\left(\sum hf\right)^2}{\sum f}}{\sum f - 1} = \frac{839 - \frac{(-141)^2}{80}}{79} = 7.5$$

فإن :

$$\therefore S_h = \sqrt{7.5} = 2.74$$

وبالتالي فإن :

الانحراف المعياري للمتغير X $= bh = 100 \times 2.74 = 274$ جنية

٧- وبما أن :

معامل الاختلاف المعياري $= \dfrac{bS_h}{\bar{X}} \times 100 = \dfrac{(100)2.74}{573.75} \times 100 = 47.35\%$

تطبيق (٤-٦)

تقوم إحدى شركات إنتاج اللمبات الكهربائيه بإنتاج 1000 وحدة يومياً . فإذا قامت الشركة بدراسة لعدد الوحدات المعيبة اليومية ، فقامت بجمع بيانات عن عدد الوحدات المعيبة في اليوم لمدة شهر (٣٠ يوم) فكانت البيانات على النحو التالي:

6	3	14	$\boxed{20}$	4	9	7	5	12	7
9	5	16	13	7	8	$\boxed{0}$	2	15	10
2	3	4	9	5	11	15	14	0	12

المطلوب :

١- كون جدول توزيع تكراري مناسب.
٢- كون التوزيع التكراري النسبي.
٣- أوجد متوسط عدد الوحدات المعيبة في اليوم.
٤- أوجد محدد الوحدات المعيبة الأكثر شيوعاً.
٥- أوجد احتمال وجود أقل من 4 لمبات معيبة في الانتاج اليومي كذلك احتمال وجود أكثر من أو يساوى 8 لمبات معيبة في الإنتاج اليومي.

الحل

١- أقل عدد = 0 ، أكبر عدد = 20 وبالتالي فإن :

المدى = 20 – 0 = 20 لمبة

وبفرض أن عدد الفئات = 5

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{20}{4} = 4$$

جدول (٤-٢٦) تفريغ البيانات

الفئات	التفريغ	التكرارات
0-	/ ~~////~~	6
4-	//// ~~////~~	9
8-	/ ~~////~~	6
12-	// ~~////~~	7
16-20	//	2
المجموع		30

جدول (٤-٢٧) يوضح توزيع 30 يوما وفقاً لعدد الوحدات المعيبة في اليوم الواحد

عدد الوحدات المعيبة في اليوم الواحد	0-	4-	8-	12-	16-20	المجموع
عدد الأيام	6	9	6	7	2	30

٢- والجدول التالي يوضح التوزيع التكراري النسبي لعدد الوحدات المعيبة في اليوم الواحد.

جدول (٤-٢٨)

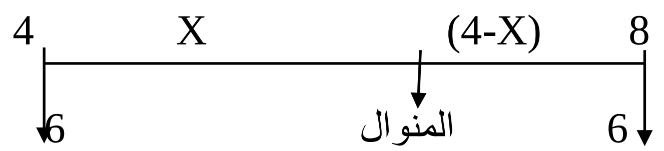
الفئات	التكرارات f	التكرار النسبي	مراكز الفئات x	$f\ x$
0-	6	$\frac{6}{30}\times 100 = 20\%$	2	12
4-	9	$\frac{9}{30}\times 100 = 30\%$	6	54
8-	6	$\frac{6}{30}\times 100 = 20\%$	10	60
12-	7	$\frac{7}{30}\times 100 = 23.33\%$	14	98
16-20	2	$\frac{2}{30}\times 100 = 6.67\%$	18	36
المجموع	30	100%		260

٣- متوسط عدد اللمبات المعيبة في اليوم الواحد $\bar{X}$ حيث :

$$\bar{X} = \frac{\sum f\ x}{\sum f} = \frac{260}{30} = 8.67 \approx 9$$ لمبات معيبة يومياً

٤- من الجدول التكراري يتضح أن الفئة المنوالية هي [4-8]

نفرض أن المنوال يقع على بعد x من بداية الفئة المنوالية



$$6\times(x) = 6\times(4-x)$$

$$6x = 24 - 6x$$

$$12x = 24 \quad \rightarrow \therefore\ x = \frac{24}{12} = 2$$

وبالتالي فإن :

المنوال = 4 + x = 4 + 2 = 6 لمبات معيبة يوميا

٥- احتمال وجود أقل من 4 لمبات معيبة = 0.20

واحتمال وجود أكثر من أو يساوى 8 لمبات معيبة = 0.667+0.2333 = 0.50

تطبيق (٤-٧)

يرغب أحد المدربين في اختيار لاعب من لاعبين لكرة السلة للاشتراك في مباراة دولية ، فأجرى اختبار لكل منهما حيث يقوم كل مهما بإجراء 10 محاولات لتصويب هدف على بعد معين وفي كل محاولة لكل منهما قام بتسجيل المسافات التي يقطعها كل منهما تجاه الهدف بالياردة فكانت على النحو التالي :

المسافة التي حققها اللاعب الأول : 41,55,30,38,50,42,39,25,28,52
المسافة التي حققها اللاعب الثاني : 39,42,38,42,44,40,41,38,36,40

ومن البيانات يتضح أن متوسط المسافة للاعبين متساوية وتساوى 40 ياردة
المطلوب :
أي اللاعبين يتم اختياره.

الحل
يعتبر اللاعب الأفضل هو اللاعب الأقل تشتت في تصوبيه للهدف وبالتالي يتم اختيار اللاعب بحيث يكون مقياس التشتت للمسافات له أقل . ومن البيانات نجد أن متوسط المسافة لكل من اللاعبين الأول والثاني هما μ_2, μ_1 على النحو التالي :

$$\mu_1 = \frac{41+55+30+38+50+42+39+25+28+52}{10}$$

$$= \frac{400}{10} = 40 \text{ ياردة}$$

$$\mu_2 = \frac{39+42+38+42+44+40+41+38+36+40}{10}$$

$$= \frac{400}{10} = 40 \text{ ياردة}$$

وبالتالي فإن :

$$\sigma_1^2 = \frac{\sum_{i=1}^{n}(x-\mu_1)^2}{n}$$

$$= \frac{(41-40)^2 + (55-40)^2 + (30-40)^2 + ...+ (52-40)^2}{10}$$

$$= \frac{1+225+100+4+100+4+1+225+144+144}{10}$$

$$= \frac{948}{10} = 94.8$$

$\therefore \sigma_1 = \sqrt{94.8} = 9.74$ ياردة (1)

وبالمثل

$$\sigma_2^2 = \frac{\sum_{i=1}^{n}(x-\mu_2)^2}{n}$$

$$= \frac{(39-40)^2 + (42-40)^2 + (38-40)^2 + ...+ (40-40)^2}{10}$$

$$= \frac{1+4+4+4+16+0+1+4+16+0}{10} = \frac{49}{10} = 4.9$$

$\therefore \sigma_2 = \sqrt{4.9} = 2.21$ ياردة (2)

من (1) ، (2) نجد أن التشتت في التصويب للهدف للاعب الثاني أقل من التشتت في التصويب للهدف للاعب الأول . وبالتالي يكون اختيار اللاعب الثاني هو الاختيار الأفضل.

تطبيق (٤-٨)

فيما يلي بيانات عن صافي الربح السنوي لكل من الشركة (١) ، والشركة (٢) " بـ 10 ألاف جنية" في الفترة : ١٩٩٠ – ١٩٩٩ م .

جدول (٤-٢٩)

السنة	١٩٩٠	١٩٩١	١٩٩٢	١٩٩٣	١٩٩٤	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩
صافي ربح الشركة (١)	48	52	51	49	50	53	55	47	45	50
صافي ربح الشركة (٢)	50	-60	0	155	170	0	19	51	50	65

فإذا طرحت كل من الشركتين (١) ، (٢) أسهم لهما في السوق . أي شركة من الشركتين يفضل شراء الأسهم منها مع ذكر السبب.

الحل

بحساب متوسط صافي الربح في الشركتين خلال الفترة (١٩٩٠ - ١٩٩٩) نجد أن : $\mu_1 = \mu_1 = 500$ ألف جنية سنوياً . لذا يتم التفضل في شراء الأسهم من الشركة لبتي تتمتع باستقرار نسبي في صافي الربح ، ومن ثم يفضل الشراء من الشركة التي يكون لها أقل تشتت في صافي الربح.

متوسط ربح الشركة (١) هو μ_1 حيث :

$$\mu_1 = \frac{48+52+51+49+50+53+55+47+45+50}{10}$$

$$= \frac{500}{10} = 50 = 500 \text{ حنية ألف} \qquad (1)$$

متوسط ربح الشركة (٢) هو μ_2 حيث :

$$\mu_2 = \frac{50-60+0+155+170+0+19+51+50+65}{10}$$

$$= \frac{500}{10} = 50 = 500 \text{ حنية ألف} \qquad (2)$$

وبالتالي فإن :

$$\sigma_1^2 = \frac{\sum_{i=1}^{n}(x_i - \mu_1)^2}{n}$$

$$= \frac{(48-50)^2 + (52-50)^2 + (51-50)^2 + ...+ (50-50)^2}{10}$$

$$= \frac{4+4+1+1+0+9+25+9+25+0}{10} = \frac{78}{10} = 7.8$$

$$\therefore \sigma_1 = \sqrt{7.8} = 2.79285 = 2792.85 \text{ جنية} \qquad (3)$$

$$\sigma_2^2 = \frac{\sum (x-\mu_2)^2}{n}$$

$$= \frac{(50-50)^2 + (60-50)^2 + (0-50)^2 + ... + (56-50)^2}{10}$$

$$= \frac{0+12100+2500+11025+14400+2500+961+1+0+225}{10}$$

$$= \frac{43711}{10} = 4371.1$$

$$\therefore \sigma_2 = \sqrt{4371.1} = 66.1143 = 66114.30 \text{ جنية} \qquad (4)$$

من (3) ، (4) يتضح أن صافي الربح في الشركة (٢) أكثر تشتت (اختلاف) من صافي الربح للشركة (١) ، وبالتالي يفضل شراء الأسهم من أسهم الشركة (١).

تطبيق (٤-٩)

تجرى إحدى الشركات المساهمة الاستثمارية الكبرى دراسة عن أعمار المساهمين بالشركة فأخذت عينة مكونه من 32 مساهم وسجلت أعمارهم بالسنوات فكانت على النحو التالي :

47	60	38	31	56	30	48	6
52	62	55	60	59	35	52	62
54	43	57	62	47	55	40	37
48	57	30	41	49	38	57	63

المطلوب :

١- كون جدول توزيع تكراري مناسب يوضح توزيع المساهمين وفقاً للسن.
٢- كون التوزيع المتجمع الصاعد ثم وضح ذلك بيانياً.
٣- من (٢) قدر عدد المساهمين الذين يقل عمر كل منهم عن 57 سنة ، ثم قدر عدد المساهمين الذين يزيد عمر كل منهم عن 57 سنة.
٤- أحسب الوسط الحسابي والانحراف المعياري للعمر في العينة.

الحل

١- من البيانات نجد أن أقل قيمة تساوى 30 سنة وأكبر قيمة تساوى 65 سنة وبالتالي فإن :

المدى = 65 – 30 = 35 سنة

نفرض أن عدد الفئات يساوى 7 فئات بالتالي فإن :

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{35}{7} = 5 \text{ سنوات}$$

ونكون جدول التفريغ التالي :

جدول (٤-٣٠)

التكرارات	التفريغ	الفئات
3	///	30-
4	////	35-
3	///	40-
5	/////	45-
3	///	50-
7	///// //	55-
7	///// //	60-65
32		المجموع

جدول (٤-٣١) توزيع المساهمين في العينة وفقاً لفئات العمر المختلفة

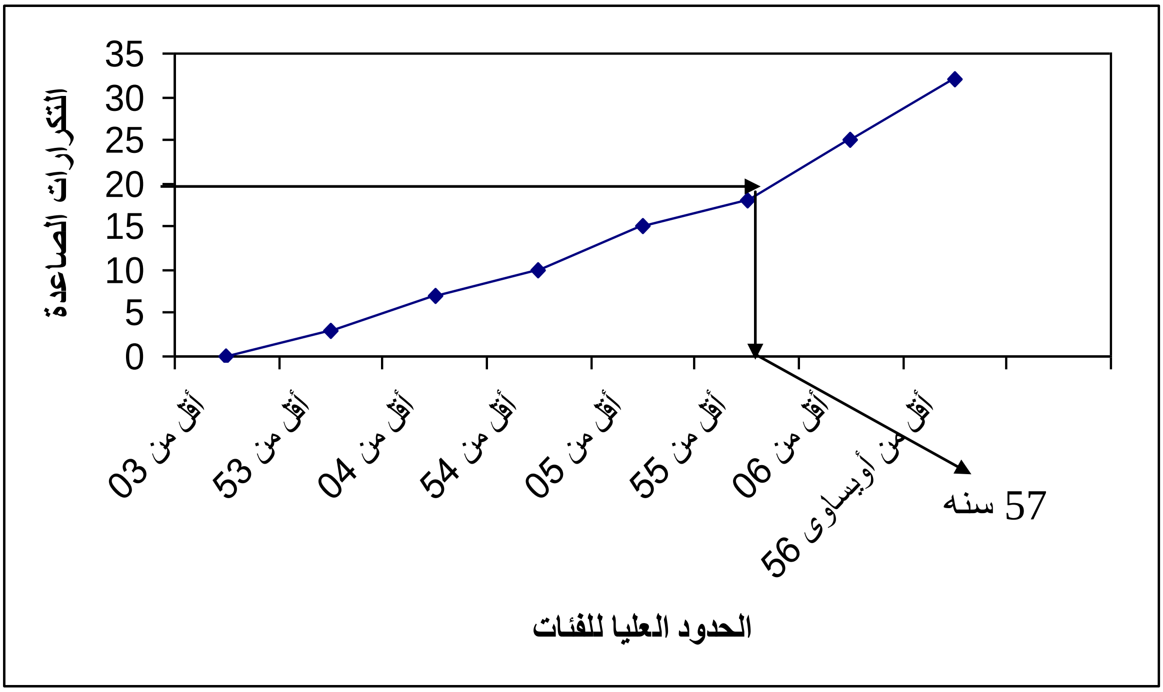
المجموع	60-65	55-	50-	45-	40-	35-	30-	الفئات
32	7	7	3	5	3	4	3	التكرارات

٢- الجدول التالي يوضح التوزيع المتجمع الصاعد للتوزيع التكراري السابق

جدول (٤-٣٢)

أقل من أو يساوى 65	**أقل من 60**	**أقل من 55**	**أقل من 50**	**أقل من 45**	**أقل من 40**	**أقل من 35**	**أقل من 30**	الحدود العليا للفئات
32	25	18	15	10	7	3	0	التكرارات الصاعدة

والشكل التالي يوضح المنحنى الصاعد الذي يمثل التوزيع المتجمع الصاعد السابق.



٣- ومن الرسم يتضح أن :

- عدد المساهمين الذين عمر كل منهم يقل عن 57 سنة = 21 مساهم تقريباً
- عدد المساهمين الذين يزيد عمر كل منهم عن 57 سنة = 21 - 32 = 11 مساهم تقريباً

٤- لحساب الوسط الحسابي والانحراف المعياري نكون الجدول التالي :

جدول (٤-٣٣)

الفئات	التكرارات f_i	مراكز الفئات x_i	$y_i = x_i - 47.5$	$h_i = \frac{y_i}{5}$	$f_i\ h_i$	$f_i\ h_i^2$
30-	3	32.5	-15	-3	-9	27
35-	4	73.5	-10	-2	-8	16
40-	3	42.5	-5	-1	-3	3
45-	5	47.5	0	0	0	0
50-	3	52.5	5	1	3	3
55-	7	57.5	10	2	14	28
60-65	7	62.5	15	3	21	63
المجموع	32				18	140

$$\bar{h} = \left(\frac{\sum_{i=1}^{n} f_i \ h_i}{\sum_{i=1}^{n} f_i} \right) \times 5 + 47.5$$

$$= \left(\frac{18}{32} \right) \times 5 + 47.5 = 2.813 + 47.5 = 50.31 \text{ سنة}$$

وبالتالي فإن :

$$\sigma_h^2 = \frac{\sum_{i=1}^{n} f_i \ h_i^2}{\sum_{i=1}^{n} f_i} - (\bar{h})^2$$

$$= \frac{140}{32} - \left(\frac{18}{32} \right)^2 = 4.375 - 0.3164 = 4.0586$$

وكذلك فإن :

$$\sigma_x^2 = \sigma_h^2 \times 25 = 4.0586 \times 25 = 101.465$$

$$\therefore \sigma_x = \sqrt{101.465} = 10.07 \text{ سنة}$$

(٤-٩) تمرينات Exercises

(٤-١)

في إحدى الشركات يوجد 150 عامل ، فإذا كان متوسط الإنتاجية اليومية لـ 100 عامل منهم 15 وحدة في الساعة ، 50 عامل الآخرين متوسط الإنتاجية اليومية لكل منهم 10 وحدات في الساعة.

المطلوب :

أ – أوجد متوسط الإنتاجية في الساعة.

ب- هل الإجابة في (أ) لن تتغير إذا كان متوسط الإنتاجية لـ 100 هو 15 وحدة ، الـ 50 الآخرين متوسط إنتاجيتهم 10 وحدات.

(٤-٢)

إذا كان x_i تمثل مراكز الفئات لإحدى التوزيعات التكرارية البسيطة ، y_j , h_j تمثل الانحرافات البسيطة والانحرافات المختصرة لـ x_i بحيث $y_j = x_j - a$ ، $h_j = \dfrac{x_j - a}{b}$ حيث a,b مقادير ثابتة

أثبت أن :

$$1)\ \bar{y} = \bar{x} - a \quad , \quad \bar{x} = b\bar{h} + a$$

$$2)\ \sigma_x^2 = \sigma_y^2 \quad , \quad \sigma_y^2 = b^2 \sigma_h^2$$

(٤-٣)

الجدول التالي يوضح توزيع عينة مكونة من 270 شركة من إنتاج المنتجات الغذائية وفقاً لراس مال كل شركة (الوحدة بالألف جنية)

مئات راس المال (بالألف جنية)	130-	140-	150-	160-	170-	180-190	المجموع
عدد الشركات	10	33	100	50	70	7	270

المطلوب :

١- أحسب متوسط راس مال الشركة.

(٤-٤)

فيما يلي بيانات عن كمية المخزون من سلعة ما (بالألف وحدة) في 40 مخزن بمحافظتي القاهرة والجيزة :

1.1	6.2	7.3	5.5	7.2	0.75	4	7.5
3.5	5.5	8.5	5.6	3.3	2.5	3	3
4.2	5	2.1	7.8	11.5	10.5	7.5	1.5
6.5	13	4.4	11	12.4	13.1	11.2	15
7.2	10	5.1	12.3	1.7	2	13	1

المطلوب :

١- حدد أكبر كمية توزيع من المخزون في المخازن المختلفة.
٢- كون جدول توزيع تكراري مناسب.
٣- أوجد مراكز الفئات في (٢) ثم أرسم المضلع التكراري.

(٤-٥)

في التمرين السابق أوجد ما يلي :

١- التوزيع التكراري الصاعد النسبي ووضح ذلك بيانياً.
٢- التوزيع التكراري الهابط النسبي ووضح ذلك بيانياً.
٣- قدر نسبة المخازن التي يقل فيها المخزون عن 7 آلاف وحدة.
٤- قدر نسبة المخازن التي يزيد فيها المخزون عن 5 آلاف وحدة ويقل عن 9 آلاف وحدة.

(٤-٦)

الجدول التالي يوضح توزيع 1500 عامل حسب فئات الأجر الشهرية بالجنية

فئات الأجر بالجنية	70-	100-	150-	190-	200-300	المجموع
عدد العمال	120	300	520	60	500	1500

المطلوب :

١- أوجد التكرارات المعدلة ومراكز الفئات.
٢- أرسم كل من المدرج والمضلع التكراري.
٣- قدر عدد العمال الذين تقل أجورهم عن 220 جنية.

(٤-٧)

فيما يلي بيانات عن الدخل والإنفاق في إحدى الشهور لـ 20 أسرة بمحافظة القاهرة

رقم الأسرة	الدخل	الإنفاق	رقم الأسرة	الدخل	الإنفاق
1	400	370	11	800	750
2	500	350	12	600	560
3	1000	750	13	470	450
4	500	525	14	300	320
5	300	300	15	720	700
6	250	270	16	850	800
7	460	500	17	790	750
8	270	300	18	230	250
9	300	310	19	320	340
10	450	450	20	510	500

المطلوب :

١- كون جدول توزيع تكراري مناسب يوضح توزيع الأسر وفقاً للدخل والإنفاق الشهري معاً.

٢- قدر بيانياً عدد الأسر التي يقل دخل كل منها عن 500 جنية.

٣- قدر بيانياً عدد الأسر التي يزيد إنفاق كل منها عن 650 جنية.

(٤-٨)

الجدول التالي يوضح توزيع العاملين في المصنعين A , B وفقاً لفئات العمر المختلفة بالسنوات

فئات العمر بالسنوات	أقل من 20	20-	22-	25-	35-	40 فأكثر	المجموع
عدد العمال في المصنع A	18	50	270	700	300	162	1500
عدد العمال في المصنع B	100	300	1000	300	200	1000	2000

قارن بين توزيع العاملين في المصنعين وفقاً لأعمارهم ووضح ذلك بيانياً

(٤-٩)

الجدول التالي يوضح توزيع 120 أسرة حسب عدد الأفراد في كل أسرة

عدد الأفراد في الأسرة	عدد الأسر
2-5	25
6-9	50
10-13	35
14-17	10
المجموع	120

المطلوب :

١- حدد الحدود العليا والدنيا لكل فئة.
٢- أوجد أطوال الفئات.
٣- أوجد مراكز الفئات.
٤- أرسم شكل بياني يوضح توزيع الأسر حسب عدد الأفراد.
٥- أوجد التوزيع التكراري النسبي للأسرة.

(٤-١٠)

الجدول التالي يوضح توزيع 500 عامل حسب فئات الأجر الأسبوعي بالجنية

الأجر بالجنية	عدد العمال
50-	80
100-	120
150-	200
200-	75
250-300	25
المجموع	500

المطلوب :

١- أرسم مدرج تكراري يوضح توزيع العمال وفقاً للأجر.
٢- أرسم مضلع تكراري يوضح توزيع العمال وفقاً للأجر أيضاً
٣- أثبت أن مساحة المدرج التكراري تكافئ مساحة المضلع التكراري.

(٤-١١)

فيما يلي أوزان 40 طالب وطالبة بالكيلوجرام بالمرحلة الابتدائية :

17	15.5	40	31	18.5	30	31
21	18.5	22	35	20.5	40	22
30	30	23	28	24	42	25
35	25	15	25	25	49	27
32	26	43	27.5	25.5	30	31
		31	41.5	48	21	31

المطلوب :

١- كون توزيع تكراري مناسب يمثل توزيعات الطلاب وفقاً لأوزانهم.
٢- قدر عدد الطلاب الذين تزيد أوزانهم عن 44 كيلوجرام.
٣- قدر عدد الطلاب الذين تقل أوزانهم عن 23 كيلوجرام.

(٤-١٢)

في التمرين السابق :

١- أوجد نسبة الطلاب الذين تقل أوزانهم عن 22 كيلوجرام.
٢- أوجد نسبة الطلاب الذين تزيد أوزانهم عن 35 كيلوجرام.
٣- قدر عدد الطلاب الذين تنحصر أوزانهم بين (45-35) كيلوجرام.

(٤-١٣)

الجدولان التاليان يوضحان توزيع الطلاب في الصفين الأول والثاني في مادة الرياضيات بإحدى كليات التجارة.

توزيع الطلاب في الصف الأول

الفئات	التكرارات
0-	40
5-	60
10-	100
15-20	50
المجموع	250

توزيع الطلاب في الصف الثاني

الفئات	التكرارات
0-	35
5-	50
10-	90
15-20	10
المجموع	185

المطلوب :

مقارنة توزيع الطلاب في الصف الأول والصف الثاني في الرياضيات

(٤-١٤)

فيما يلي بيانات عن راس مال 40 شركة من الشركات الصغيرة (بالألف جنية).

10	15	7	22	19	22	11.5
7.5	12	9	23	21	25	18
8	25	15	11	15.5	28	20
10	30	17	15	19.1	21	22
13	17	11.5	13	15	91	20
		18	21	29	28	21

المطلوب :

١- كون جدول توزيع تكراري مناسب يوضح توزيع الشركات حسب فئات راس المال.
٢- أرسم المنحني التكراري.
٣- أوجد التوزيع التكراري النسبي.
٤- أوجد التوزيع التكراري النسبي الصاعد.
٥- أوجد التوزيع التكراري النسبي الهابط.
٦- قدر نسبة الشركات التي يزيد راس مالها عن 15 ألف جنية.
٧- قدر نسبة الشركات التي يقل راس مالها عن 10 آلاف جنية.

(٤-١٥)

قامت إحدى شركات التامين بإجراء دراسة على عينة عشوائية مكونة من 50 عميل من العملاء المأمن عليهم بشأن الشكاوى التي قدمت من العملاء خلال العامين الماضيين . والجدول التالي يعطى عدد الشكاوى المقدمة من العميل خلال العام وعدد العملاء المناظر :

عدد الشكاوى المقدمة من العميل الواحد	0	1	2	3	4	5	6
عدد العملاء	20	14	6	3	3	2	2

المطلوب :

١- أوجد متوسط عدد الشكاوى اليومى للشركة.
٢- أوجد التباين والانحراف المعياري لعدد الشكاوى في العينة.
٣- أوجد القيمة الوسيطة للشكاوى ، كذلك الوسط الحسابي للشكاوي ثم قارن بينهما.

(٤-١٦)

الجدول التالي يوضح توزيع 2000 أسرة من محافظتين مختلفتين وفقاً لمتوسط عدد مكالمات التليفون في أحد الشهور

عدد المكالمات	0-	20-	40-	60-	80-	100 فأكثر
عدد الأسر في المحافظة الأولي	15	135	350	220	180	100
عدد الأسر في المحافظة الثانية	25	175	200	220	230	50

المطلوب :

١- أرسم المضلع التكراري لكل توزيع في شكل واحد.

٢- أوجد التوزيع المتجمع الصاعد لكل توزيع ، ثم قدر من الرسم عدد الأسر في كل محافظة التي تزيد عدد مكالمتها الشهرية عن 45 مكالمة.

٣- أوجد الانحراف المعياري لعدد المكالمات في كمحافظة ، ثم قارن بين تشتت التوزيعين.

٤- أوجد الوسط الحسابي ، الوسيط والمنوال لعدد المكالمات ووضح ذلك على الرسم ثم قارن بين المؤشرات الثلاثة :-

أ – بالنسبة لكل توزيع على حدة.

ب- بالنسبة للتوزيعين معاً

الباب الخامس

التوزيعات الاحتمالية المتقطعة

Discrete Probability Distribution

(٥-١) التوزيع الاحتمالي المتقطع
Discrete Probability Distribution

(٥-٢) عملية برنولي **Bernoulli Process**

(٥-٣) توزيع ذات الحدين **Binomial Distribution**

(٥-٤) توزيع الهيبروجومترك
Hyper Geometric Distribution

(٥-٥) توزيع بواسون **Poisson Distribution**

(٥-٦) أمثلة تطبيقية **Applied Examples**

(٥-٧) تمرينات **Exercises**

(٥-١) أهمية دراسة الاحتمالات

The Importance of Studying Probabilities

علم الاحتمالات من العلوم الهامة والأساسية في الدراسات النظرية والتطبيقية ، ففي جميع الدراسات التي تتناول ظاهرة أو أكثر في ظروف عدم التأكد Uncertainty States فإن دراسة هذه الظواهر وعلاقتها ببعض دراسة علمية تتطلب إتباع نظرية الاحتمالات.

(٥-٢) تعريفات **Definitions**

في هذا الفصل نتناول بالدراسة أهم المفاهيم الأساسية والضرورية لنظرية الاحتمالات Probability Theory ، ومجموعة القواعد Rules المرتبطة بها والتي يجب الإلمام بها عند دراسة الظواهر التي لا تتوافر معرفة كاملة عن خصائصها ، هذا من ناحية ومن ناحية أخري فإن محتويات هذا الباب تمثل عنصر أساسي في دراسة الأبواب التالية سواء من الناحية النظرية أو الناحية التطبيقية.

كذلك يتضمن هذا الباب مجموعة من الأمثلة التطبيقية التي توضح استخدام نظرية الاحتمالات في حل العديد من المشاكل الفعلية وكيفية الوصول إلي قرار صحيح بشأنها• في ظروف عدم التأكد Uncertain States.

وفي الفصول التالية سوف نقدم بعض المفاهيم Concepts الضرورية لدراسة الاحتمالات ، ومما هو جدير بالذكر أن دراسة نظرية الاحتمالات مرتبطة ارتباط وثيق بنظرية الفئات•• "Set's Theory" . لذلك سوف نقدم أولاً بعض التعريفات الضرورية المتعلقة بنظرية الفئات.

تعريف (٥-١)

المتغير Variable هو الشيء الذي يمكن أن يأخذ قيم مختلفة في الظروف المختلفة (زمنية ، مكانية ، سياسية ، اقتصادية ، ... الخ) .

فمثلاً سعر كيلو الدقيق يختلف من سوق إلي أخر في نفس الوقت ، ويختلف من عام إلى عام أخر ، .... الخ . كذلك وزن الفرد يمثل متغير يتوقف على عمر الفرد ، وطوله ، ... الخ . وعادة يرمز للمتغير الذي يأخذ عدد (n) من القيم بالرمز X ، حيث X تأخذ القيم X_r ، $r = 1,2,3,....,n$.

كذلك الشيء الثابت Constant هو مقدار لا يتغير بتغير الظروف وعادة يرمز له بالرمز C .

وإذا كان لدينا متغيران x, y تربط بينهما العلاقة (الدالة) f بحيث أن :

• أ.د. عفاف الدش : "الإحصاء وصناعة القرارات – الجزء الثاني" – الطبعة الأولى – مكتبة عين شمس ١٩٩٦ – القاهرة برقم إيداع ٩٦/٩٣٦٤ .

•• أ.د. عفاف الدش "الرياضيات وصناعة القرارات" – مكتبة عين شمس ١٩٩٤ ، القاهرة برقم إيداع ٩٤/٨٦٨٢

$$y = f(x) \qquad (5.1)$$

فإن المتغير x يمثل نطاق الدالة f ويسمى بالمتغير المستقل Independent Variable ويسمى المتغير y بمدى الدالة أو المتغير التابع Dependent Variable.

والمتغير الذي يمكن أن يأخذ أى قيمة بين قيمتين معينتين يسمى متغير متصل Continuous Variable ، وخلاف ذلك يسمى المتغير بالمتغير المتقطع Discrete Variable .

مثال (٥-١)

إذا كان x متغير يمثل عدد الأطفال الممكن إنجابهم في إحدى الأسر ، فإن x ممكن أن تأخذ القيم 10, ..., 4, 3, 2, 1, 0 مثلاً وتكتب على النحو :

$$x = 0 ,1 ,2 ,3 ,4 ,.... ,10$$

ولا يمكن أن تأخذ x القيمة 1.5 , 2.31 ، وبالتالي فإن x في هذه الحالة يمثل متغير متقطع .

مثال (٥-٢)

إذا كان المتغير y يمثل عمر الفرد فمن الممكن أن تأخذ y القيمة 15 سنة ، 15.2 سنة ، 15.25 سنة ، وبالتالي فإن المتغير y يمكن أن يأخذ أى قيمة في فترة زمنية معينة وبالتالي فإن y يمثل متغير متصل .

مثال (٥-٣)

إذا كان x متغير بحيث أن :

$$x = \{x \mid x \in I \ , \ 1 \le x \le 10\} \qquad (5.2)$$

حيث I تمثل فئة الأعداد الصحيحة ، فإن :

$x = 0 ,1 ,2 ,3 ,4 ,5 ,6 ,7 ,8 ,9 ,10$ ، أى أن x متغير متقطع .

أما إذا كان المتغير H بحيث أن :

$$H = \{h \mid h \in R \ , \ 1 \le h \le 10\} \qquad (5.3)$$

حيث R تمثل فئة الأعداد الحقيقة ، وبالتالي فإن المتغير H يمكن أن يأخذ أى قيمة بين (1 ,10) مثل 1.65 ، 2.4 ، ..... ، الخ ، وبالتالي فإن H يمثل متغير متصل .

تعريف (٥-٢)
التجربة العشوائية Random Experiment هى التجربة التى يعرف مقدماً جميع النتائج الممكنة لها ، ولكن لا يمكن أن يعرف مقدماً ترتيب حدوث هذه النتائج . وتمثل نتائج Outcomes التجربة العشوائية متغير يسمى بالمتغير العشوائي Random Variable ، حيث أن الصدفة (الصدفة عبارة عن مجموعة من العوامل التى لا يمكن التحكم فيها) هى التى تحدد نتائج التجربة .

فمثلاً عند إلقاء قطعة عملة متوازنة فإننا نعرف قبل الرمي أن نتيجة هذه التجربة إما صورة ونرمز لها بالرمز x_1 ، وإما كتابة ونرمز لها بالرمز x_2 ، ولكن أيهما سيقع اولاً لا يمكن معرفة ذلك قبل الرمي . وإذا رمزنا لنتائج التجربة السابقة بالرمز x ، فإن x تمثل متغير عشوائي يأخذ القيم x_1, x_2 ، أى أن :

$$x = x_1, x_2 \tag{5.4}$$

كذلك عند إلقاء زهرة نرد (طاولة) متوازنة فإننا نعرف قبل الرمي أن نتائج هذه التجربة ، إما أن تكون النتيجة 6, 5, 4, 3, 2, 1 .
وبذلك يعتبر إلقاء هذه الزهرة تجربة عشوائية ، فإذا رمزنا لنتائج هذه التجربة بالرمز x ، فإن x يمثل متغير عشوائي بحيث :

$$x = 1 ,2 ,3 ,4 ,5 ,6 \tag{5.5}$$

تعريف (٥-٣)
فراغ المعاينة Sampling Space هو الفئة التى تتكون من جميع النتائج الممكنة للتجربة العشوائية . حيث يمثل كل عنصر في هذه الفئة نتيجة من نتائج التجربة العشوائية ، وأحياناً تسمى هذه الفئة بفراغ النتائج • Outcome Space أو فراغ الإحتمال Probability Space . ويرمز عادة لفراغ المعاينة بالرمز S ، فمثلاً فراغ المعاينة لرمي قطعة عملة في (5.4) هي الفئة S حيث :

$$S = \{x_1, x_2\} \tag{5.6}$$

كذلك فراغ المعاينة لرمي زهرة النرد في (5.5) هي الفئة S حيث :

$$S = \{1,2,3,4,5,6\} \tag{5.7}$$

وبالتالي فإن فراغ المعاينة هو الفئة الشاملة Universal Set لنتائج التجربة العشوائية .

•Heinz Hohler (1994) : "Statistics for Business and Economics" . Harper Collins publishers . New York

تعريف (٥-٤)
الحدث العشوائي Random Event هو فئة جزئية Subset من فراغ المعاينة ، فقد يكون الحدث فئة مكونة من عنصر واحد (أى نتيجة واحدة) أو أكثر من عنصر.

فعند رمي قطعة العملة ، إذا كان الحدث A يمثل ظهور صورة فإن :

$$A = \{x_1\} \tag{5.8}$$

كذلك عند رمي زهرة النرد ، فإذا كان الحدث B يمثل ظهور رقم فردى فإن :

$$B = \{1,3,5\} \tag{5.9}$$

فنجد أن الحدث A يمثل فئة جزئية من فراغ المعاينة S في العلاقة (5.8) هو فئة تحتوى على عنصر واحد كذلك الحدث B في العلاقة (5.9) هو فئة تحتوى على ثلاثة عناصر وهو فئة جزئة من فراغ المعاينة S في العلاقة (5.7) . ويسمى كل من الحدث A , والحدث B بالأحداث البسيطة Simple Events ، حيث العناصر (النتائج) الموجودة في كل حدث تعتبر عناصر (نتائج) بسيطة .

ويوجد نوع أخر من الأحداث العشوائية تسمى بالأحداث المركبة Compound Events حيث تتكون هذه الأحداث من عناصر (أو نتائج) مركبة كما سوف يتضح من المثال التالي :

مثال (٥-٤)
إذا رميت قطعتي عملة متوازنتين ، بحيث تشير x_1 للصور ، وتشير x_2 للكتابة . فإن فراغ المعاينة لهذه التجربة S يصبح :

$$\{(x_1,x_1),(x_1,x_2),(x_2,x_1),(x_2,x_2)\}\text{S} = \tag{5.10}$$

فنجد أن كل عنصر في الفئة S في العلاقة (5.10) عنصر (نتيجة) مركبة ، فالعنصر الواحد مكون من نتيجتين ، نتيجة على القطعة الأولى وأخرى على القطعة الثانية فمثلاً العنصر (x_1,x_1) يشير إلى أن النتيجة على القطعة الأولى صورة والنتيجة على القطعة الثانية صورة أيضاً ، في حين أن العنصر (x_2,x_1) يشير إلى ظهور الكتابة على القطعة الأولى والصورة على القطعة الثانية .

ويمكن أن يتكون العنصر الواحد (أو النتيجة الواحدة) المركب من أكثر من نتيجتين

مثال (٥-٥)

فـي إحـدى شـركات إنتـاج المعلبـات الغذائيـة تـم فحـص 3 وحـدات لتحديـد أى عبـوة صـالحة للاستخدام وأى عبوة غير صالحة

المطلوب :

١- حدد فراغ المعاينة ووضح ذلك من خلال شكل "فان"

٢- أوجد الأحداث التالية :

أ- جميع الوحدات الصالحة

ب- وحدة واحدة على الأكثر فاسدة

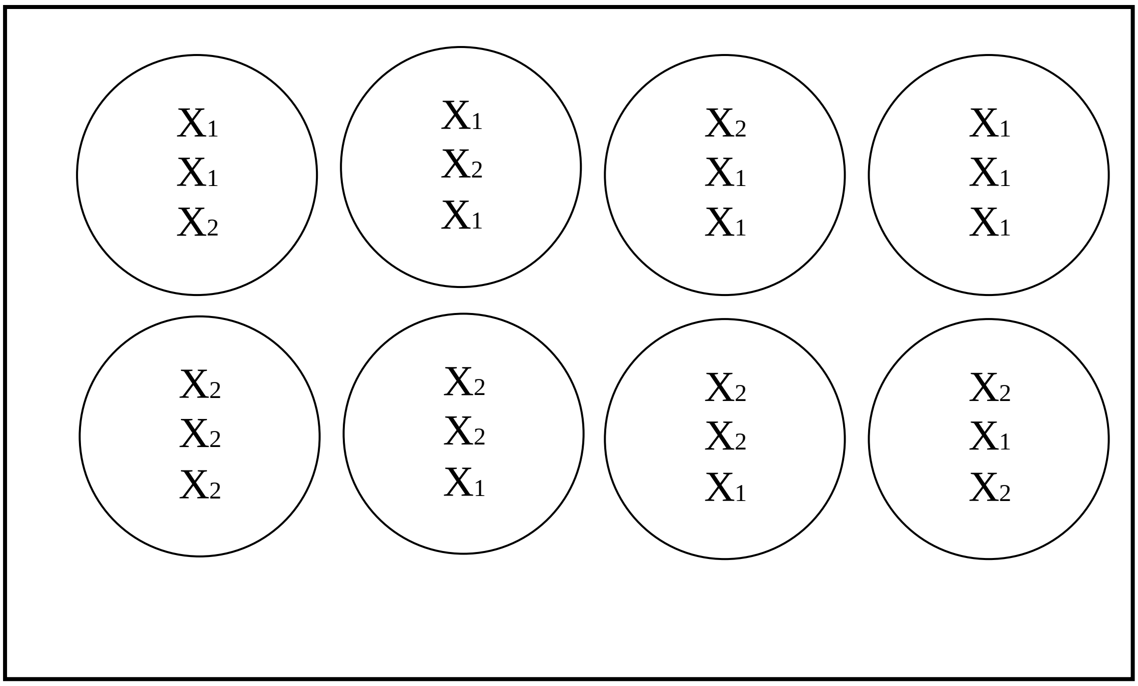
ج- جميع الوحدات فاسدة

الحل

١- إذا افترضـنا أن x_2 , x_1 تشـير إلـى العبـوة الصـالحة والعبـوة غيـر الصـالحة علـى الترتيب ، كذلك S تشير إلى فراغ المعاينـة . وفـي هـذه الحالـة نجـد أن عـدد العناصـر في فراغ المعاينة $= 2^3 = 8$ عناصر مركبة على النحو التالي :

$$S = \{(x_1, x_1, x_1), (x_1, x_1, x_2), (x_1, x_2, x_1), (x_2, x_1, x_1), (x_2, x_1, x_2), (x_2, x_2, x_1), (x_1, x_2, x_2), (x_2, x_2, x_2)\} \tag{5.11}$$

وشكل فان التالي يوضح فراغ المعاينة S



شكل (٥-١)

٢-

أ- نفرض أن الحدث A يشير إلى " الوحدات الصالحه" وبالتالي فإن :

$$\mathrm{A} = \{x_1, x_1, x_1\} \qquad (5.12)$$

فنجد أن الفئة A تتكون من عنصر مركب واحد.

ب- كذلك إذا فرضنا أن B تشير إلى الحدث "وجود وحدة واحدة على الأقل فاسدة" فإن :

$$\mathrm{B} = \{(x_1, x_1, x_1), (x_1, x_2, x_1), (x_1, x_1, x_2), (x_2, x_1, x_1)\} \qquad (5.13)$$

فنجد أن الحدث B يتكون من 4 عناصر مركبة كل عنصر مركب من ثلاثة عناصر بسيطة.

ج- بالمثل إذا فرضنا أن C تشير إلى الحدث "جميع الوحدات فاسدة" فإن :

$$\mathrm{C} = \{x_2, x_2, x_2\} \qquad (5.14)$$

ووقع الحدث العشوائى يعنى ظهور أحد عناصر فئة الحدث كناتج للتجربة العشوائية كما أن عدم وقوع الحدث يعنى عدم ظهور أى عنصر من عناصر فئة الحدث كنتيجة للتجربة وبالتالي فهو يعنى وقوع الحدث المكمل له Complement Event.
فمثلاً عدم وقوع الحدث B يعنـى وقوع الحدث B' (أى وجود وحدة صالحة على الأكثر) حيث :

$$B' = \{(x_2, x_2, x_1), (x_2, x_1, x_2), (x_1, x_2, x_2), (x_2, x_2, x_2)\} \qquad (5.15)$$

ومما سبق يتضح أنه إذا كانت $(A_1, A_2, ..., A_n)$ هي عناصر فئة الحدث A فإن وقوع الحدث A يعنى وقوع واحد على الأقل من الأحداث الجزئية $(\{A_1\}, \{A_2\}, ..., \{A_n\})$ ، أو بعبارة أخرى :

$$A = \{A_1\} \cup \{A_2\} \cup .... \cup \{A_n\} = \cup_{i=1}^{m} \{A_i\} \qquad (5.16)$$

تعريف (٥-٥)

يقال أن الحدثين A ، B حدثين متنافيين Two Exclusive Events إذا كان وقوع أحداهما يمنع وقوع الأخر . فمثلاً عند رمى قطعة عملة متوازنة ، إذا ظهر الوجة الذي علية الصورة فإن هذا يؤدى إلى منع ظهور الوجة الذي عليه الكتابة ، وبالتالي فإن حدث ظهور الصورة وحدث ظهور الكتابة يمثلان حدثين متنافيين . كذلك وجود شخص ما في مكان معين في لحظة زمنية معينة يمنع وجوده في مكان أخر في نفس اللحظة الزمنية ، وبالتالي فوجود الشخص في مكانيين مختلفين في نفس اللحظة الزمنية يمثل حدثين متنافيين . وبالتالي يكون الحدثين A ، B حدثين متنافيين إذا كان :

$$A \cap B = \phi \tag{5.17}$$

حيث تمثل ϕ الفئة الخالية.

تعريف (٥-٦)
إذا كان لدينا الحدثين A ، B بحيث وقوع الحدث A لا يمنع ولا يؤثر على وقوع الحدث B ، كذلك وقوع الحدث B لا يمنع ولا يؤثر على وقوع الحدث A . فإنه يقال أن الحدثين A ، B حدثين مستقلين Two Independent Events.

ويمكن تعميم ذلك ، بمعنى إذا كان لدينا الحدث A_i ، $i = 1,2,...,n$ بحيث وقوع أى حدث A_i لا يمنع ولا يؤثر على وقوع أى حدث أخر A_j بحيث $i \neq j$ ، $i = 1,2,...,n$ ، $j = 1,2,...,n$ فإنه يقال أن الأحداث A_i ، $i = 1,2,...,n$ أحداث مستقلة.

ومن التعريقيين (٥-٥) ، (٥-٦) يتضح أن مفهوم الحوادث المتنافية يختلف عن مفهوم الحوادث المستقلة.

تعريف (٥-٧)
إذا كانت m تمثل عدد مرات وقوع الحدث A في n من المحاولات (التجارب) المتماثلة ، وعرفنا الدالة $p_r(A)$ ، حيث $p_r(A)$ تشير إلى إحتمال وقوع الحدث A فإن :

$$p_r(A) = \lim_{n \to \infty} \frac{m}{n} \tag{5.18}$$

والمقصود بكلمة $\lim\limits_{n\to\infty}$* ليس المعنى الرياضى المعروف ولكن المقصود هو أن قيمة هذا الإحتمال سوف تقترب من الصحة إذا كان عدد المحاولات كبير جداً .
ويسمى الإحتمال المعرف في (5.18) بالاحتمال التجريبي Experimental Probability.

مثال (٥-٦)

إذا قذفت عملة متوازنة عشوائياً 10000 مرة ، وكانت الأحداث x_2, x_1 تشير إلى ظهور الصورة والكتابة على الترتيب ، فإذا كانت نتيجة القذف 5720 مرة ظهور الصورة ، 4270 مرة ظهور الكتابة .
أوجد احتمال ظهور الصورة $p_r(x_1)$ ، واحتمال ظهور الكتابة $p_r(x_2)$.

الحل

$$1)\, p_r(x_1) = \frac{5720}{10000} = 0.572$$

$$2)\, p_r(x_2) = \frac{4380}{10000} = 0.428$$

كذلك إذا أجريت تجربة عشوائية وكان عدد النتائج الممكنة لهذه التجربة n (عدد عناصر فراغ المعاينة) بحيث كان لكل نتيجة للتجربة نفس الفرصة في الظهور ، فإذا كان m من هذه النتائج تؤدى إلى وقوع الحدث A فإنه يمكن حساب احتمال وقوع الحدث A كما يلي :

$$p_r(A) = \frac{m}{n} \qquad (4.20)$$

وبعبارة أخرى فإن احتمال وقوع الحدث A هو $p_r(A)$ حيث :

$$p_r(A) = \frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}} \qquad (5.21)$$

حيث أن الحالات المواتية هي النتائج المؤدية إلى وقوع الحدث A كذلك الحالات الممكنة هي جميع النتائج الممكنة أي عناصر فراغ العينة .

* أ.د. سمير كامل عاشور (١٩٨٦) : "الإحصاء التحليلى" ، معهد الدراسات والبحوث الإحصائية – جامعة القاهرة .

مثال (٥-٧)
إذا رميت زهرة طاولة (نرد) متوازنة عشوائياً . أحسب احتمال ظهور رقم فردي ، وكذلك احتمال ظهور رقم زوجي.

الحل
بما أن فراغ المعاينة في هذه الحالة هو :

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

وبالتالي فإن $n = 6$ ، فإذا كان A , B تشيران إلى حدث ظهور رقم فردي وحدث ظهور رقم زوجي على الترتيب ، فإن :

$$A = \{1 , 3 , 5\} \quad , \quad B = \{2 , 4 , 6\}$$

وبالتالي فإن :

$$p_r(A) = \frac{3}{6} = \frac{1}{2} \quad ، \quad p_r(B) = \frac{3}{6} = \frac{1}{2}$$

مثال (٥-٨)
الجدول التالي يوضح توزيع 5000 أسرة من ذو الدخل المتوسط حسب عدد الغرف التي تسكنها الأسرة الواحدة.

جدول (٥-١)

عدد الغرف التي تسكنها الأسرة الواحدة	1	2	3	4	المجموع
عدد الأسر	500	3000	1000	500	5000

الحل
إذا فرضنا أن x تشير إلى عدد الغرف التي تسكنها الأسرة الواحدة أي أن :
$$x = 1 , 2 , 3 , 4$$

وبما أن عدد الأسر الكلي يساوى 5000 حالة ، وبالتالي فإن احتمال أن تسكن الأسرة غرفة واحدة هو :

$$p_r(x=1) = \frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}} = \frac{500}{5000} = 0.1$$

وبالمثل احتمال أن تسكن الأسرة غرفتين هو $p_r(x=2)$ حيث :

$$p_r(x=2)=\frac{3000}{5000}=0.60$$

واحتمال أن تسكن الأسرة ثلاثة غرف هو $p_r(x=3)$ حيث :

$$p_r(x=3)=\frac{1000}{5000}=0.2$$

واحتمال أن تسكن الأسرة أربعة غرف هو $p_r(x=4)$ حيث :

$$p_r(x=4)=\frac{500}{5000}=0.10$$

مثال (٥-٩)

إذا ألقي زهرة نرد عشوائياً ، أوجد ما يلي :

١- فراغ المعاينة لتجربة الرمي.

٢- احتمال أن يكون الرقم على الزهرة الأولي يساوى الرقم على الزهرة الثانية.

٣- احتمال أن يكون الرقمين على الزهرتين مساوي 4 ، كذلك احتمال أن يكون المجموع أقل من 4 .

٤- احتمال أن يكون مجموع الرقمين على الزهرتين أكبر من 9.

الحل

الجدول التالي يوضح جميع النتائج الممكنة لرمي زهرتي النرد :

جدول (٢-٥)

نتائج الزهرة الأولي / نتائج الزهرة الثانية	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

من الجدول يتضح أن عدد النتائج الممكنة n = 36 نتيجة (أو حالة) وهي عبارة عن نتائج (أو حالات) مركبة كل نتيجة مكونة من نتيجتين (النتيجة على الزهرة الأولي ، والنتيجة على الزهرة الثانية) ، وبالتالي فإن فراغ المعاينة S يأخذ الشكل التالي :

$$S = \{(1,1)\ ,\ (1,2)\ ,\ .....\ ,\ (2,1)\ ,\ ....\ ,\ (6,6)\} \qquad (5.23)$$

٢- إذا فرضنا أن الحدث A يمثل الحدث أن تكون نتيجة الرمي (الرقم على الزهرة الأولي يساوي الرقم على الزهرة الثانية) فإن :

$$A = \{(1,1)\ ,\ (2,2)\ ,\ (3,3)\ ,\ (4,4)\ ,\ (5,5)\ ,\ (6,6)\} \qquad (5.24)$$

ونلاحظ أن عدد العناصر في الفئة A يساوي عدد الحالات المواتية(m) يساوي 6 عناصر ، وبالتالي فإن احتمال أن يكون الرقم على الزهرة الأولي مساوي للرقم على الزهري الثانية يساوي $p_r(A)$ حيث :

$$p_r(A) = \frac{m}{n} = \frac{6}{36} = \frac{1}{6}$$

٣- إذا فرضنا أن الحدث (B) هو أن يكون مجموع الرقمين على الزهرتين مساوي4 فإن :

$$B = \{(1,3)\ ,\ (2,2)\ ,\ (3,1)\}$$

وبالتالي فإن احتمال أن يكون مجموع الرقمين على الزهرتين مساوي 4 هو $p_r(B)$ حيث :

$$p_r(B) = \frac{m}{n} = \frac{3}{36} = \frac{1}{12}$$

كذلك إذا فرضنا أن الحدث C هو أن يكون مجموع الرقمين على الزهرتين أقل من أو يساوي 4 فإن :

$$C = \{(1,1)\ ,\ (2,1)\ ,\ (1,2)\ ,\ (1,3)\ ,\ (3,1)\ ,\ (2,2)\}$$

وبالتالي فإن احتمال أن يكون مجموع الرقمين على الزهرتين أقل من أو يساوي 4 هو $p_r(C)$ حيث :

$$p_r(C) = \frac{m}{n} = \frac{6}{36} = \frac{1}{6}$$

٤- بالمثل إذا كان الحدث H هو أن يكون مجموع الرقمين على الزهرتين أكبر من 9 فإن :

$$H = \{(5,5) , (5,6) , (6,5) , (6,6)\}$$

وبالتـالي فـإن احتمـال أن يكـون مجمـوع الـرقمين علـى الزهـرتين أكبـر مـن 9 هو $p_r(H)$ حيث :

$$p_r(H) = \frac{m}{n} = \frac{4}{36} = \frac{1}{9}$$

٥- إذا فرضنا أن الحدث K هو أن يكون حاصل ضرب الرقمين على الزهرتين رقم فردي أقل من 10 فإن :

$$K =\{(1,1) , (3,1) , (1,3) , (5,1) , (1,5) , (7,1) , (1,7) , (9,1) , (1,9)\}$$

وبالتالي فإن احتمال أن يكون حاصل ضرب الرقمين على الزهرتين رقم فردي أقل من 10 هو $p_r(K)$ حيث :

$$p_r(K) = \frac{m}{n} = \frac{9}{36} = \frac{1}{4}$$

(٥-٣) قوانين الاحتمالات **Laws of Probabilities**

من التعريف السابق للاحتمال أتضح أنه يمكن إيجاد قيمة احتمال وقوع الحدث بقسمة عدد العناصر في فئة الحدث محل الدراسة (m) على عدد العناصر في فئة فراغ المعاينة (n) .

ومن العمليات على الفئات Set's Operation فقد يكون الحدث محل الاعتبار ناتج من عملية على حدثين أو أكثر ، فعلى سبيل المثال قد يكون الحدث محل الاعتبار هو A عبارة عن اتحاد Union الحدثين B ، C أى أن :

$$A = \{B \cup C\}$$

أو قد يكون الحدث A هو حدث نـاتج من تقاطع Intersection الحدثين B ، C أى أن :

$$A = \{B \cap C\}$$

أو قد يكون الحدث A هو حدث نـاتج من الفرق Difference بين الحـدثين B ، C أى أن :

$$A = \{B - C\}$$

وفي أواخر القرن السابع عشر والنصف الأول من القرن الثامن عشر الميلادي اشتق عـالم الرياضـيات أبراهـام دو مـوفر Abraham De Mover (١٦٦٧- ١٧٥٤) الصياغات الأولى لقوانين الإضـافة والضـرب Addition and Multiplication Laws . وكان لاشتقاق هذه القـوانين وتطويرهـا أهميـة بـالغـة فـي تطـوير وتنميـة علـم الاحتمالات بصفة خاصة وعلوم الإحصاء والرياضيات التطبيقية بصفة عامة.

وفي هذا الفصل سوف نتناول أهم هذه القوانين التى يمكن باستخدامها الحصول علـى احتمال لحدث ناتج من عمليـة أو أكثـر علـى حدثين أو أكثـر من خـلال مجموعـة من النظريات الأساسية لعلم الاحتمالات بالإضافة إلى بعض المسلمات Axioms المبنية عليهـا هـذه النظريـات بالإضــافة إلــى تقـديم العديـد مـن الأمثلـة علــى اسـتخدام هـذه القوانين .

مسلمة (١) : إذا كان الحدث A فئة جزئية من فراغ المعاينة S أى أن :

$$A \subseteq S$$

فإن :

$$p_r(A) \geq 0 \tag{5.24}$$

مسلمة (٢) : إذا كانت S هى فراغ المعاينة فإن :

$$p_r(S) = 1 \tag{5.25}$$

مسلمة (٣) : إذا كانت الأحداث $A_1, A_2, A_3, \ldots\ldots, A_n$ سلسلة من الأحداث المتنافية Sequence of Mutually Exclusive Events في فراغ العينة S بمعنى :

$$A_i \cap A_j = \phi \qquad i \neq j \tag{5.26}$$

فإن :

$$p_r(A_1 \cup A_2 \cup \ldots\ldots \cup A_n) = p_r(A_1) + p_r(A_2) + \ldots\ldots + p_r(A_n)$$

أو

$$p_r(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} p_r(A_i) \tag{5.27}$$

ومن المسلمات الثلاثة السابقة تمكن العلماء من استنتاج النظريات التالية :

نظرية (٥-١)
إذا كان الحدث A' هو مكمل الحدث A في فراغ المعاينة S فإن

$$p_r(A') = 1 - p_r(A) \qquad (5.28)$$

الإثبات :-
من نظرية الفئات نجد أنه :

$$(A' \cap A) = \phi \quad ، \quad (A' \cup A) = S \qquad (5.29)$$

وبالتالي فإن

$$p_r(A' \cup A) = p_r(S) = 1 \qquad (5.30)$$

وبما أن الحدثين A, A' حدثين متنافيين ، فمن المسلمة (٣) نجد أن :

$$p_r(A' \cup A) = p_r(A') + p_r(A) \qquad (5.31)$$

وبالتعويض في (5.31) بقيمة $p_r(A' \cup A)$ في (5.30) نجد أن :

$$p_r(A') + p_r(A) = 1 \;\; \rightarrow \therefore \; p_r(A') = 1 - p_r(A)$$

نظرية (٥-٢)
إذا كان $A \subseteq S$
فإن :

$$0 \le p_r(A) \le 1 \qquad (5.32)$$

الإثبات :-
من المسلمة (١) نجد أن

$$p_r(A) \ge 0 \quad ، \quad p_r(A') \ge 0 \qquad (5.33)$$

ومن النظرية السابقة نجد أن

$$p_r(A) = 1 - p_r(A') \qquad (5.34)$$

وبما أن $p_r(A) \ge 0$ فإن :

$$p_r(A) \le 1 \;\; \rightarrow \therefore \; 0 \le p_r(A) \le 1 \qquad (5.35)$$

نظرية (٥-٣)

إذا كان الحدث A فئة خالية Null Set أى أن

$$A=\phi$$

فإن

$$p_r(A)=p_r(\phi)=0 \tag{5.36}$$

الإثبات :-

بما أن $A=\phi \;\; \rightarrow \therefore A=S=\phi$

وبالتالي فإن :

$$p_r(A)=p_r(S')$$

$$p_r(S \cup S')=p_r(S)+p_r(S') \tag{5.37}$$

$$=1+p_r(A) \tag{5.38}$$

وبما أن $(S \cup S')=S$

وبالتالي فإن

$$p_r(S \cup S')=p_r(S) \tag{5.39}$$

بالتعويض في الطرف الأيمن من المعادلة (5.38) بـ $p_r(S \cup S')$ من المعادلة (5.39) نجد أن :

$$p_r(S)=1+p_r(A) \;\; \rightarrow \therefore 1=1+p_r(A) \;\; \rightarrow p_r(A)=0$$

نظرية (٥-٤)

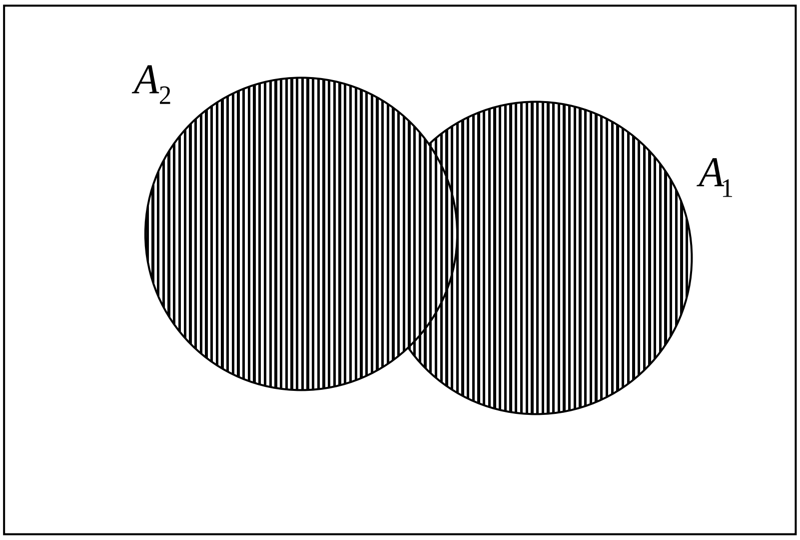
إذا كان الحدثين $A_1, A_2 \subseteq S$

فإن :

$$p_r(A_1 \cup A_2)=p_r(A_1)+p_r(A_2)-p_r(A_1 \cap A_2) \tag{5.40}$$

الإثبات :-

بما أن $(A_1 \cup A_2)=A_1 \cup (A_1' \cap A_2)$ كما هو موضح بالشكل التالي



شكل (٥-٢)

وبما أن الحدثين $[(A_1' \cap A_2)\ ,\ (A_1)]$ حدثين متنافيين ، فمن المسلمة (٢) نجد أن :

$$\begin{aligned} p_r(A_1 \cup A_2) &= p_r[A_1 \cup (A_1' \cap A_2)] \\ &= p_r(A_1) + p_r(A_1' \cap A_2) \qquad (5.41) \end{aligned}$$

وبما أن

$$A_1 = (A_1 \cap A_2) \cup (A_1' \cap A_2) \qquad (4.42)$$

وحيث أن الحدثين $[(A_1' \cap A_2)\ ,\ (A_1 \cap A_2)]$ حدثين متنافيين فإن :

$$\begin{aligned} p_r(A_2) &= p_r(A_1 \cap A_2) + p_r(A_1' \cap A_2) \\ p_r(A_1 \cap A_2) &= p_r(A_2) - p_r(A_1 \cap A_2) \qquad (4.43) \end{aligned}$$

وبالتعويض في الطرف الأيسر من المعادلة (5.41) بالطرف الأيسر في المعادلة (5.43) نجد أن :

$$p_r(A_1 \cap A_2) = p_r(A_1) + p_r(A_2) - p_r(A_1 \cap A_2) \qquad (4.44)$$

نتيجة

إذا كان الحدثين A_1 , A_2 حدثين متنافيين ، أى أن $p_r\left(A_1 \cap A_2\right) = \phi$. ومن نظرية (٥-٣) نجد أن :

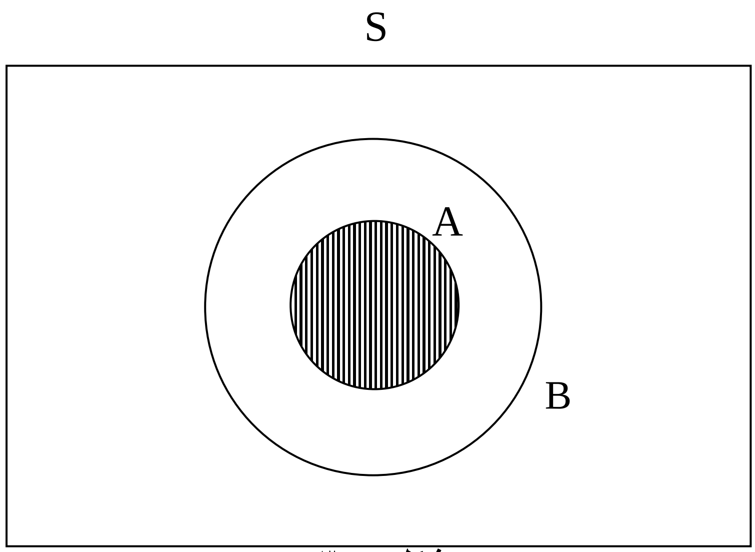
$$p_r\left(A_1 \cap A_2\right) = p_r(\phi) = 0 \qquad (4.45)$$

وبالتعويض بقيمة $p_r\left(A_1 \cap A_2\right)$ من المعادلة (5.45) في الطرف الأيسر من المعادلة (4.44) نجد أن :

$$p_r\left(A_1 \cap A_2\right) = p_r\left(A_1\right) + p_r\left(A_2\right) \qquad (4.46)$$

نظرية (٥-٥)

إذا كان الحدثين A ، B بحيث $(A , B \subseteq S)$ ، $(A \subseteq B)$ كما هو موضح بالشكل التالي :



شكل (٥-٣)

فإن :

$$p_r(A) \le p_r(B) \qquad (5.47)$$

الإثبات :-

بما أن $B \subseteq A$ إذن :

$$(B) = (A) \cup \left(A' \cap B\right) \qquad (5.48)$$

وبما أن الفئتين (A) ، $(A'\cap B)$ فئتين متنافيتين ، فمن النتيجة السابقة نجد أن :

$$p_r(B) = p_r(A) + p_r(A'\cap B) \tag{5.49}$$

ومن المسلمة (١) نجد أن :

$$p_r(A'\cap B) \geq 0 \tag{5.50}$$

ومن المعادلة (5.49) والمتباينة (5.50) نجد أن :

$$p_r(A) \leq p_r(B)$$

مثال (٥-١٠)

إذا كانت الفئة S هي فراغ المعاينة ، والحدثين A ، B بحيث $A , B \subseteq S$. بحيث

$$p_r(A) = 0.3\ , p_r(B) = 0.5\ , p_r(A\cap B) = 0.1$$

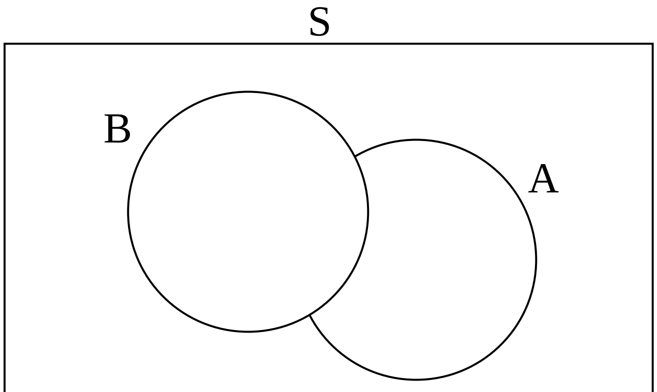
أحسب الاحتمالات التالية :

١- $p_r(A')\ ,\ \ p_r(B')$

٢- $p_r(A'\cap B)\ ,\ \ p_r(A'\cap B')$

الحل

الشكل التالي يوضح الحدثين A , B



شكل (٥-٤)

١- من نظرية (٥-١) نجد أن :

$$p_r(A') = 1 - p_r(A) = 1 - 0.3 = 0.7$$
$$p_r(B') = 1 - p_r(B) = 1 - 0.5 = 0.5$$

٢- ومن الشكل نجد أن :

$$(B)=(A'\cap B)\cup(A\cap B)$$
$$p_r(B)=p_r(A'\cap B)+p_r(A\cap B)$$
$$p_r(A'\cap B)=p_r(B)-p_r(A\cap B)$$
$$=0.5-0.1=0.4$$

كذلك

$$(A'\cap B')\cup(A\cap B)'=S-(A\cap B)$$
$$p_r(A'\cap B')=p_r(S)-p_r(A\cup B)$$
$$=1-[0.3+(0.5-0.1)]=1-(0.7)=0.3$$

ملحوظة

إذا كان وقوع الحدثين A ، B معا هو $(A\cap B)$ فإن احتمال وقوع الحدثين A ، B معا هو $p_r(A\cap B)$ يسمى بالاحتمال المشترك Joint Probability .

(٥-٤) الاحتمالات الشرطية **Conditional Probabilities**

إذا كان A , B حدثين في فراغ العينة S ، أى أن $A , B \subseteq S$. فإذا كان حدوث الحدث A يتوقف على حدوث الحدث B ، فإن احتمال وقوع الحدث A بشرط وقوع الحدث B يسمى بالاحتمال الشرطي Conditional Probability ويرمز له بالرمز $p_r(A|B)$ ويقرأ "احتمال وقوع الحدث A بشرط وقوع الحدث B" ويسمى الحدث B بالحدث القبلي Prior Event ويسمى الحدث (A|B) بالحدث البعدي Posterior Event .

ويتم حساب قيمة $p_r(A|B)$ من المعادلة التالية :

$$p_r(A|B)=\frac{p_r(A\cap B)}{p_r(B)} \quad , \quad p_r(B)>0 \qquad (5.51)$$

وبالمثل

$$p_r(B\,|\,A)=\frac{p_r(A \cap B)}{p_r(A)} \quad , \quad p_r(A)>0 \qquad (5.52)$$

من المعادلتين (5.51) ، (5.52) نجد أن :

$$p_r(A \cap B)= p_r(A\,|\,B)\;p_r(B) \qquad (5.53)$$

$$= p_r(B\,|\,A)\;p_r(A) \qquad (5.54)$$

والمثال التالي يوضح مفهوم الاحتمال الشرطي الذي يمكن حسابه من المعادلتين (5.51) ، (5.52) .

مثال (٥-١٢)

الجدول التالي يوضح عد العمال المهرة وغير المهرة في ثلاثة أقسام بأحد المصانع

جدول (٥-٣)

القسم / نوع العمل	الأول (A_1)	الثاني (A_2)	الثالث (A_3)	المجموع
ماهر (B)	370	410	55	835
غير ماهر (B')	30	90	45	165
المجموع	400	500	100	1000

أحسب الاحتمالات التالية :

١- احتمال اختيار عامل من القسم الأول ، احتمال اختيار عامل من القسم الثاني ، احتمال اختيار عامل من القسم الثالث

٢- احتمال اختيار عامل ماهر بشرط أن يكون من القسم الأول

٣- احتمال اختيار عامل من القسم الأول بشرط أن يكون ماهر

٤- احتمال اختيار عامل غبر ماهر بشرط أن يكون من القسم الثالث

الحل

١- إذا فرضنا أن A_1, A_2, A_3 تمثل اختيار عامل من القسم الأول ، الثاني ، الثالث على الترتيب فإن :

$$p_r(A_1)=\frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}}=\frac{400}{1000}=0.4$$

وبالمثل :

$$p_r(A_2)=\frac{500}{1000}=0.5$$

$$p_r(A_3)=\frac{100}{1000}=0.1$$

٢- إذا فرضنا أن الحدث B يمثل اختيار عامل ماهر ، فإن الحدث $(B \mid A_1)$ يمثل اختيار عامل ماهر بشرط أن يكون من القسم الأول . ففي هذه الحالة يكون الحدث القبلي هو (A_1) أى اختيار عامل من القسم الأول ، والحدث البعدي هو $(B \mid A_1)$ أى اختيار عامل ماهر من عمال القسم الأول . أى عند اختيار العامل يكون الاختيار من عمال القسم الأول فقط وليس من عمال المصنع ككل ، وبالتالي فإن :

$$p_r(B \mid A_1)=\frac{p_r(B \cap A_1)}{p_r(A_1)}=\frac{370/1000}{400/1000}=\frac{37}{40}$$

٣- كذلك احتمال اختيار عامل من القسم الأول بشرط أن يكون ماهر هو $p_r(A_1 \mid B)$ أى أن الحدث القبلي هو(B) أى اختيار عامل ماهر ، والحدث البعدي هو $(A_1 \mid B)$ أى اختيار عامل من العمال المهرة بشرط أن يكون من القسم الأول ، أى عند اختيار العامل يكون الاختيار من العمال المهرة فقط وليس من عمال المصنع ككل . وبالتالي فإن :

$$p_r(A_1 \mid B)=\frac{p_r(A_1 \cap B)}{p_r(B)}=\frac{370/1000}{835/1000}=\frac{37}{167}$$

ونلاحظ أن

$$p_r(A_1 \mid B)\neq p_r(B \mid A_1)$$

٤- بالمثل احتمال اختيار عامل غير ماهر بشرط أن يكون من القسم الثالث هو $p_r(B' \mid A_3)$ حيث أن :

$$p_r(B' \mid A_3)=\frac{p_r(B' \cap A_3)}{p_r(A_3)}=\frac{45/1000}{100/1000}=\frac{45}{100}=0.45$$

مما سبق يتضح أن دراسة الاحتمال الشرطي لحدث معين في تجربة ما يمكننا من دراسة مفردات الحدث القبلي فقط بدلاً من دراسة جميع مفردات فراغ المعاينة ، أو بعبارة أخرى يؤدى إلى التعامل مع جزء من عناصر فراغ المعاينة وليس الكل.

مثال (٥-١١)

إذا ألقيت قطعتي عملة متوازنتين عشوائياً . ما هو احتمال ظهور صورة على القطعة الأولى بشرط ظهور كتابة على القطعة الثانية .

الحل

الجدول التالي يوضح النتائج الممكنة للتجربة حيث H تشير إلى الصورة ، W تشير إلى الكتابة.

جدول (٥-٤)

نتائج القطعة الثانية / نتائج القطعة الأولى	H	W
H	(H,H)	(H,W)
W	(W,H)	(W,W)

من الجدول يتضح أن عدد الحالات الممكنة n = 4 حيث :

فراغ المعاينة = S = {(H,H) , (H,W) , (W,H) , (W,W)}

فإذا فرضنا أن الحدث A يشير إلى ظهور الصورة على القطعة الأولى ، والحدث B يشير إلى ظهور الكتابة على القطعة الثانية ، فإن :

$$(A) = \{(H,H)\ ,\ (H,W)\} \qquad \rightarrow p_r(A) = \frac{2}{4} = \frac{1}{2} \qquad (5.53)$$

$$(B) = \{(H,W)\ ,\ (W,W)\} \qquad \rightarrow p_r(B) = \frac{2}{4} = \frac{1}{2} \qquad (5.54)$$

كذلك

$$(A \cap B) = (H,W) \quad \rightarrow p_r(A \cap B) = \frac{1}{4}$$

كذلك إذا فرضنا أن الحدث (A | B) هو ظهور الصورة على القطعة الأولى بشرط ظهور الكتابة على الثانية ، فنجد أن :

$$(A \mid B) = \{(H,W)\} \qquad (5.55)$$

ومن تعريف الاحتمال نجد أن :

$$p_r(A \mid B) = \frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}} = \frac{1}{2} \qquad (5.56)$$

حيث أن عدد الحالات المواتية حالة واحدة الموجودة في المعادلة (5.55) وعدد الحالات الممكنة في هذه الحالة يساوى حالتين كما هو موضح في المعادلة (5.54) ، وفي المعادلة (5.56) تم حساب الاحتمال الشرطي عن طريق تعريف الاحتمال.
والآن سوف نقوم بحساب $p_r\left(A\,|\,B\right)$ باستخدام قاعدة الاحتمال الشرطي في (5.57) على النحو التالي :

$$p_r\left(A\,|\,B\right)=\frac{p_r\left(A\cap B\right)}{p_r\left(B\right)}=\frac{1/4}{2/4}=\frac{1}{2} \qquad (5.57)$$

ونلاحظ أن قيمة $p_r\left(A\,|\,B\right)$ باستخدام قاعدة الاحتمال الشرطي في (5.57) هى نفس القيمة باستخدام تعريف الاحتمال في (5.56)

مثال (٥-١٣)
أجريت دراسة على 1000 أسرة من الأسر ذات الدخل المحدود لمعرفة العلاقة بين مستوى تعليم الأم وعدد الأطفال في الأسرة . والجدول التالي يوضح توزيع 1000 أسرة حسب مستوى تعليم الأم وعدد الأطفال في الأسرة الواحدة

جدول (٥-٥)

عدد الأطفال في الأسرة / مستوى تعليم الأم	1 (B_1)	2 (B_2)	3 (B_3)	4 (B_4)	المجموع (Σ)
A_1 "أقل من المتوسط"	1	7	70	300	379
A_2 "متوسط"	3	23	60	205	291
A_3 "عالي"	125	30	174	1	330
المجموع (Σ)	130	60	304	306	1000

المطلوب :

١- أحسب احتمال وجود طفلين في الأسرة من الأسر ذات الدخل المحدود وفي نفس الوقت يكون مستوى تعليم الأم أقل من المتوسط.
٢- أحسب احتمال وجود أكثر من طفلين في الأسرة وفي نفس الوقت مستوى تعليم الأم عالي.
٣- أحسب احتمال وجود 3 أطفال في الأسرة وفي نفس الوقت مستوى تعليم الأم متوسط.
٤- أحسب احتمال أن يكون عدد الأطفال في الأسرة 3 أطفال بشرط أن يكون مستوى تعليم الأم متوسط.
٥- أحسب احتمال أن يكوم عدد الأطفال في الأسرة أكثر من طفلين بشرط أن يكون مستوى تعليم الأم عالي.

٦- أحسب احتمال أن يكوم عدد الأطفال في الأسرة أقل من 3 أطفال بشرط أن يكون مستوى تعليم الأم أقل من عالي.

الحل

١- نفرض أن الحدث A_i يشير إلى مستوى تعليم الأم i ، حيث $i = 1, 2, 3$
والحدث B_j يشير إلى عدد الأطفال j في الأسرة الواحدة حيث $j = 1, 2, 3, 4$,
وبالتالي فإن احتمال وجود طفلين في الأسرة وفي نفس الوقت يكون مستوى تعليم الأم في هذه الأسرة أقل من المتوسط هو $p_r(A_1 \cap B_2)$ حيث :

$$p_r(A_1 \cap B_2) = \frac{7}{1000} = 0.007 \qquad (5.58)$$

٢- احتمال وجود أكثر من طفلين في الأسرة وفي نفس الوقت يكون مستوى تعليم الأم عالي هو $[(A_3 \cap B_3) \cup (A_3 \cap B_4)]$ وبما أن الحدثين $(A_3 \cap B_3)$, $(A_3 \cap B_4)$ حدثين متنافيين فإن :

$$p_r[(A_3 \cap B_3) \cup (A_3 \cap B_4)] = p_r(A_3 \cap B_3) + p_r(A_3 \cap B_4)$$
$$= \frac{174}{1000} + \frac{1}{1000} = \frac{175}{1000} = 0.175$$

٣- احتمال وجود ثلاثة أطفال في الأسرة ومستوى تعليم الأم متوسط هو $p_r(A_2 \cap B_3)$ حيث :

$$p_r(A_2 \cap B_3) = \frac{60}{1000} = 0.060$$

٤- احتمال أن يكون عدد الأطفال في الأسرة 3 أطفال بشرط أن يكون مستوى تعليم الأم متوسط هو $p_r(B_3\ A_2)$ حيث :

$$p_r(B_3 | A_2) = \frac{p_r(B_3 \cap A_2)}{p_r(A_2)} = \frac{60/1000}{291/1000} = \frac{60}{291}$$

٥- احتمال أن يكوم عدد الأطفال في الأسرة أكثر من طفلين بشرط أن يكون مستوى تعليم الأم عالي هو $p_r[(B_3 | A_3) \cup (B_4 | A_3)]$ وبما أن

الحدثين $(B_3 \mid A_3),(B_4 \mid A_3)$ حدثين متنافيين وبالتالي :

$$p_r[(B_3 \mid A_3) \cup (B_4 \mid A_3)] = p_r(B_3 \mid A_3) + p_r(B_4 \mid A_3)$$

$$= \frac{p_r(B_3 \cap A_3)}{p_r(A_3)} + \frac{p_r(B_4 \cap A_3)}{p_r(A_3)}$$

$$= \frac{174/1000}{330/1000} + \frac{1/1000}{330/1000} = \frac{175}{330}$$

٦- احتمال أن يكوم عدد الأطفال في الأسرة أقل من 3 أطفال بشرط أن يكون مستوى تعلــيم الأم أقــل مــن عــالي هــو $p_r[(B_3 \mid A_1) \cup (B_3 \mid A_2)]$ ، وحيــث أن الحدثين $(B_3 \mid A_1),(B_3 \mid A_2)$ أحداث متنافية ، إذن :

$$p_r[(B_3 \mid A_1) \cup (B_3 \mid A_2)] = p_r(B_3 \mid A_1) + p_r(B_3 \mid A_2)$$

$$= \frac{70/1000}{379/1000} + \frac{60/1000}{291/1000} = \frac{70}{379} + \frac{60}{291} = \frac{43110}{110289} = 0.39$$

(٥-٥) شجرة الاحتمالات وصناعة القرار

Probability Tree for Decision's Making

فـي الفصـلين السـابقين تناولنـا بالتفصـيل مفهـوم الاحتمـال وقـانوني جمـع وضـرب الاحتمالات ، كذلك تناولنا مفهوم الاحتمال الشرطي . وفي هذا الفصل سوف نتنـاول ما يسمى بشجرة الاحتمالات Probability Tree وهو شكل يوضـح تطبيـق قـانوني جمـع وضـرب الاحتمـالات معـاً ، كـذلك سـوف نتنـاول فـي هـذا الفصـل نظريـة بيـز Bayes's Theorem للاحتمالات الشرطية وكيفية استخدامها في صناعة القرارات مع توضيح ذلك من خلال بعض الأمثلة التطبيقية.

أولاً : شجرة الاحتمالات

إذا كان الحدثين A ، B حدثين متنافيين فإن :

$$p_r(A \cup B) = p_r(A) + p_r(B) \qquad \text{(قانون الجمع)} \qquad (5.59)$$

وكذلك

$$p_r(A \cap B) = p_r(A \mid B)\, p_r(B) = p_r(B \mid A)\, p_r(A)$$

(قانون الضرب) (5.60)

فإن شجرة الاحتمالات هو شكل بياني يوضح العلاقة بين الأحداث المتنافية والاحتمالات الشرطية والاحتمالات المشتركة ، أى يوضح العلاقتين في (5.59) ، (5.60).

وفيما يلي سوف نوضح كيفية بناء شجرة الاحتمالات من خلال المثال التالي

مثال (٥-١٤)

اعتبر مثال (٥-١١) أن نقطة البداية يمكن أن تكون تقسيم عمال المصنع إلى عمال مهرة (الحدث B) وعمال غير مهرة (الحدث B') حيث نجد أن الفئتين B ، B' فئتين متنافيتين . وبالتالي فإن :

$$p_r(B \cup B') = \frac{835}{1000} + \frac{165}{1000} = 1$$

كذلك نجد أن العمال المهرة (الفئة B) يتوزعون على الأقسام الثلاثة بالمصنع ، أى يمكن النظر إلى الفئة B على إنها إتحاد لثلاث فئات متنافية (فئة العمال المهرة في القسم الأول ، فئة العمال المهرة في القسم الثاني ، فئة العمال المهرة في القسم الثالث) هي الفئات $(A_1 \mid B), (A_2 \mid B), (A_3 \mid B)$ على الترتيب.

بالمثل بالنسبة لفئة العمال غير المهرة (B') فيمكن تقسيمها إلى ثلاث فئات متنافية أيضاً هي الفئات $(A_1 \mid B'), (A_2 \mid B'), (A_3 \mid B')$ على الترتيب.

وشكل (٥-٥) يعطى شجرة الاحتمالات للمثال محل الدراسة حيث يمكن باستخدامها الحصول على الاحتمالات المشتركة باستخدام الاحتمالات الشرطية وغير الشرطية على النحو الموضح بالشكل التالى :

نقطة البداية

$$A_1 \to p_r(A_1 \cap B) = p_r(A_1 \mid B)\,p_r(B)$$
$$= \frac{370}{835} \times \frac{835}{1000} = 0.375$$

$$A_2 \to p_r(A_2 \cap B) = p_r(A_2 \mid B)\,p_r(B)$$
$$= \frac{410}{835} \times \frac{835}{1000} = 0.410$$

$$A_3 \to p_r(A_3 \cap B) = p_r(A_3 \mid B)\,p_r(B)$$
$$= \frac{55}{835} \times \frac{835}{1000} = 0.055$$

$$A_1 \to p_r(A_1 \cap B) = p_r(A_1 \mid B)\,p_r(B)$$
$$= \frac{30}{165} \times \frac{165}{1000} = 0.030$$

$$A_2 \to p_r(A_2 \cap B) = p_r(A_2 \mid B)\,p_r(B)$$
$$= \frac{90}{165} \times \frac{165}{1000} = 0.090$$

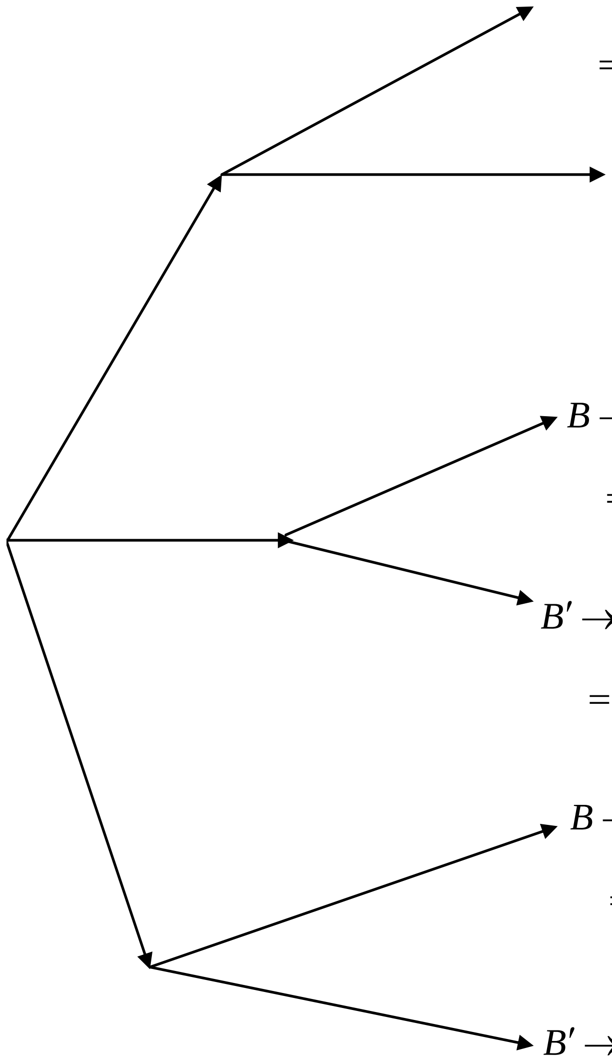
$$A_3 \to p_r(A_3 \cap B) = p_r(A_3 \mid B)\,p_r(B)$$
$$= \frac{45}{165} \times \frac{165}{1000} = 0.045$$

شكل(٥-٥)

ومـن الشـكل يتضـح أنــه يمكـن الحصـول علـى الاحتمـالات المشـتركة باسـتخدام الاحتمالات الشرطية وغير الشرطية للأحداث محل الدراسة.

كذلك يمكن بناء شجرة الاحتمال للمثال السابق بأسلوب أخر . وسوف نصل إلى نفس قيم الاحتمالات المشتركة التى وصلنا إليها باستخدام الشجرة السابقة.

فإذا اعتبرنا أن عمال المصنع موزعون على الأقسام الثلاثة ، أى يوجد لدينا ثلاث فئـات متناسـقة هـي A_1, A_2, A_3 وكـل فئـة مـن هـذه الفئـات تتكـون مـن فئتين متنـافيتين تمثلان العمال المهرة (B) أو غير المهرة (B') فتكـون نقطـة البدايـة فـي هـذه الحالـة هي تقسيم العمال على الأقسام كما هو موضح بالشكل التالي :



$$B \rightarrow p_r(B \cap A_1) = p_r(B \mid A_1)\,p_r(A_1) = \frac{370}{400} \times \frac{400}{1000} = \frac{37}{100} = 0.37$$

$$B' \rightarrow p_r(B' \cap A_1) = p_r(B' \mid A_1)\,p_r(A_1) = \frac{30}{400} \times \frac{400}{1000} = \frac{3}{100} = 0.030$$

$$B \rightarrow p_r(B \cap A_2) = p_r(B \mid A_2)\,p_r(A_2) = \frac{410}{500} \times \frac{500}{1000} = 0.410$$

$$B' \rightarrow p_r(B' \cap A_2) = p_r(B' \mid A_2)\,p_r(A_2) = \frac{90}{500} \times \frac{500}{1000} = 0.090$$

$$B \rightarrow p_r(B \cap A_3) = p_r(B \mid A_3)\,p_r(A_3) = \frac{55}{100} \times \frac{100}{1000} = 0.055$$

$$B' \rightarrow p_r(B' \cap A_3) = p_r(B' \mid A_3)\,p_r(A_3) = \frac{45}{100} \times \frac{100}{1000} = 0.045$$

شكل (٥-٦)

ثانياً : نظرية بيز

في القرن الثامن عشر الميلادي توصل عالم الرياضيات البريطاني* توماس بيز Thomas Bayes (سنة ١٧٠٢- سنة ١٧٦١) إلى صياغة جديدة للاحتمال الشرطي $p_r(A|B)$ ، أي الاحتمال البعدي Posterior Probability للحدث محل الدراسة (وليكن الحدث A) باستخدام الاحتمال القبلي Prior Probability لنفس الحدث A ومعلومات إضافية عن الحدث A ممثلة في $p_r(B|A)$.

وفيما يلي سوف نقدم نظرية بيز Bayes's Theorem وتعميمها من خلال النظريتين التاليتين.

نظرية (٥-٨)

$$p_r(A|B) = \frac{p_r(A)\,p_r(B|A)}{p_r(A)\,p_r(B|A) + p_r(A')\,p_r(B|A')} \qquad (5.61)$$

الإثبات :

بما أن :

$$p_r(A|B) = \frac{p_r(A \cap B)}{p_r(B)} \qquad (5.62)$$

ومن العلاقة (5.54) نجد أن :

$$p_r(A \cap B) = p_r(A)\,p_r(B|A) \qquad (5.63)$$

كذلك بما أن :

$$B = (A \cap B) \cup (A' \cap B)$$

وبما أن الفئتين $(A \cap B)$ ، $(A' \cap B)$ فئتين متنافيتين فإن :

$$\begin{aligned} p_r(B) &= p_r(A \cap B) + p_r(A' \cap B) \\ &= p_r(A)\,p_r(B|A) + p_r(A')\,p_r(B|A') \end{aligned} \qquad (5.64)$$

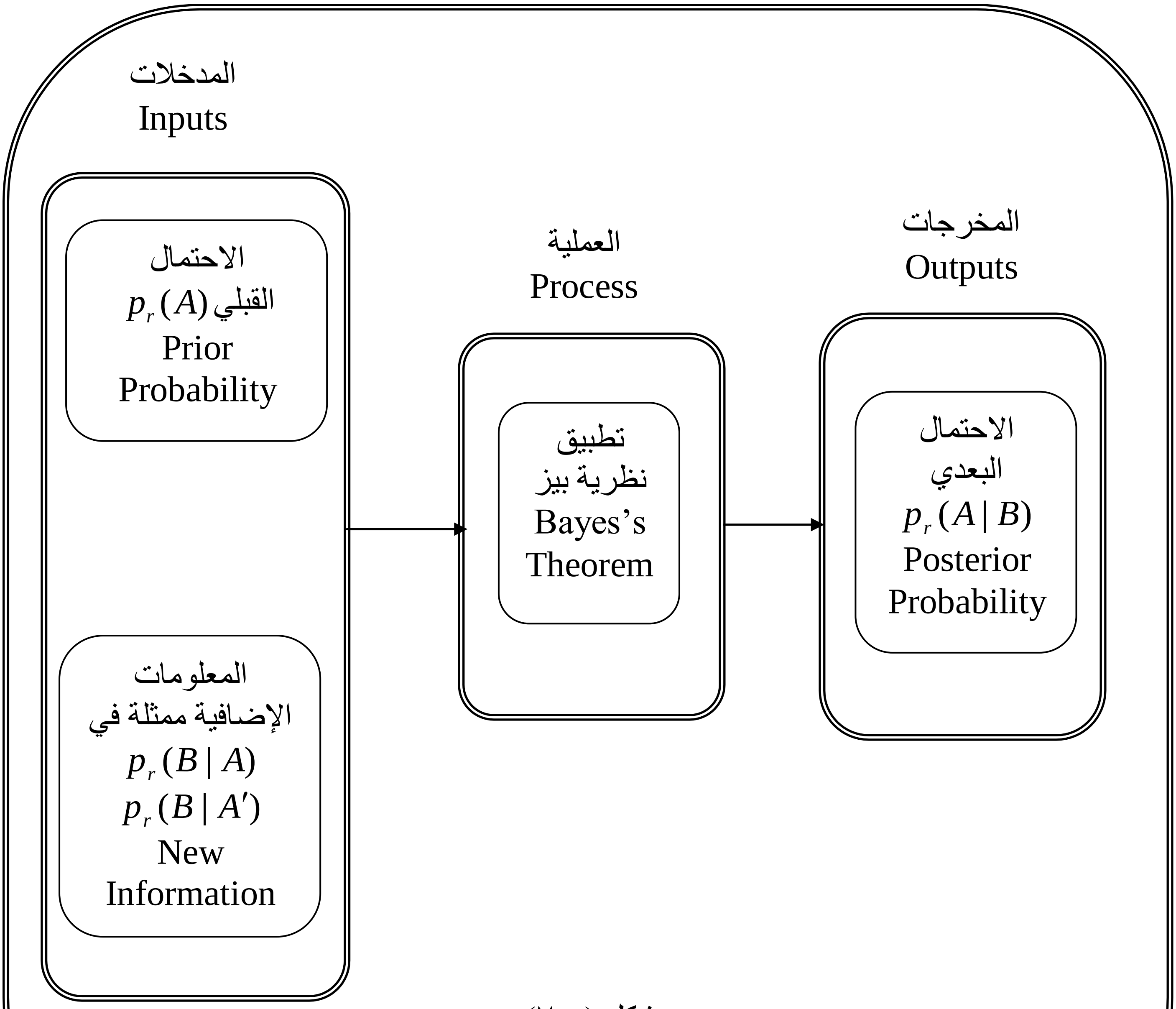
بالتعويض في (5.62) ، (5.63) ، (5.64) نجد أن :

$$p_r(A|B) = \frac{p_r(A)\,p_r(B|A)}{p_r(A)\,p_r(B|A) + p_r(A')\,p_r(B|A')}$$

---

* Heinz Kohler (1994) " Statistics for Business and Economics" , Harper Collins Clooeger Publishers , U.S.A

وتعتبر نظرية بيز من النظريات الهامة في صناعة القرارات حيث يمكن باستخدامها تحديد احتمالات وقوع الأحداث محل الاعتبار في ظل وجود معلومات جديدة إضافية مرتبطة بهذه الأحداث ، وبالتالي فإن استخدام نظرية بيز يمكننا من حساب احتمال المخاطرة Risk الفعلية للقرار محل الاعتبار . ويمكن تصوير ذلك في الشكل التالي :



شكل (٥-٧)

مثال (٥-١٥)

تحتكر إحدى الشركات للتنقيب والبحث عن البترول والغاز الطبيعي بالصحراء الغربية بجمهورية مصر العربية البحث والتنقيب في هذه المنطقة ، فإذا أثبتت الدراسات أن احتمال وجود بترول وغاز طبيعي بالصحراء الغربية يساوى 0.5 فإذا استحدثت معلومات تفيد ارتباط وجود مياه جوفية بالمنطقة بوجود بترول وغاز طبيعي ، حيث أفاد بأن احتمال وجود مياه جوفية بشرط وجود بترول وغاز طبيعي يساوى 0.7

فإذا فرضنا أن الحدث A يشير إلى وجود بترول وغاز طبيعي بالمنطقة ، كذلك A' يشير إلى عدم وجود بترول وغاز طبيعي بالمنطقة ، كذلك الحدث B يشير إلى وجود مياه جوفية . فإن :-

$$p_r(A) = 0.5 \tag{5.65}$$

$$p_r(A') = 0.5 \tag{5.66}$$

كذلك :

$$p_r(B \mid A) = 0.7 \tag{5.67}$$

فإذا فرضنا أن احتمال وجود مياه جوفية بشرط عدم وجود بترول وغاز طبيعي يساوى $p_r(B \mid A')$ حيث :

$$p_r(B \mid A') = 0.3 \tag{5.68}$$

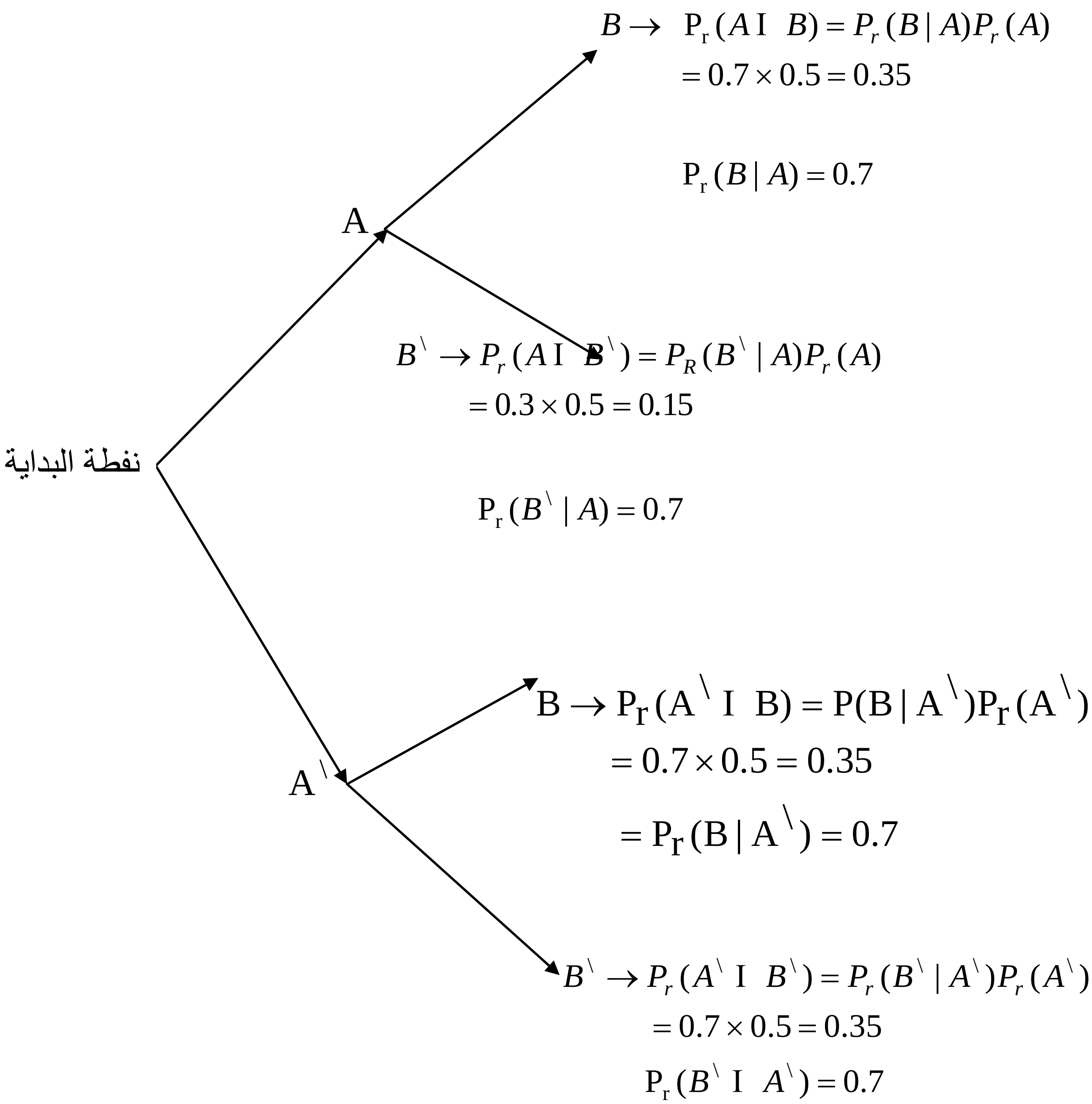
وبالتالي فإن احتمال وجود بترول وغاز طبيعي في المناطق التي يوجد بها مياه جوفية هو $p_r(A \mid B)$ حيث :

$$p_r(A \mid B) = \frac{p_r(A)\,p_r(B \mid A)}{p_r(A)\,p_r(B \mid A) + p_r(A')\,p_r(B \mid A')}$$

$$= \frac{0.5(0.7)}{0.5(0.7) + 0.5(0.3)} = \frac{0.35}{0.35 + 0.15} = \frac{0.35}{0.50} = 0.70 \tag{5.69}$$

ومن المعادلتين (5.69) ، (5.66) يتضح أن احتمال وجود البترول والغاز الطبيعي في المناطق التي بها مياه جوفية يساوى 0.7 ، أي يزيد عن احتمال وجود بترول وغاز في الصحراء الغربية دون الأخذ في الاعتبار وجود مياه جوفية ، وبالتالي يصبح القرار الأمثل التنقيب في المناطق التي يوجد بها مياه جوفية وتصبح المخاطرة Risk في هذه الحالة 0.3 بدلاً من 0.5 حيث احتمال وجود مياه بشرط عدم وجود بترول وغاز يساوى "$p_r(B \mid A') = 0.3$" في حيث احتمال الحفر وعدم وجود بترول أو غاز يساوى "$p_r(A) = 0.5$"

والشكل التالي يوضح شجرة القرارات لهذه المشكلة :



ويمكن تعميم نظرية بيز في حالة وجود أكثر من حدث محل الاعتبار في فراغ المعاينة ، والنظرية التالية تعطى تعميم لنظرية بيز.

نظريه (٥-٩)

إذا كان S فراغ المعاينة ، بحيث $A_1, A_2, A_3, \ldots, A_n \subseteq S$ ، والأحداث $A_1, A_2, A_3, \ldots, A_n$ أحداث متنافية بحيث $\bigcup_{i=1}^{n} A_i = S$. فإن :

$$p_r(A_j \mid B) = \frac{p_r(A_j)\, p_r(B \mid A_j)}{\sum_{j=1}^{n} p_r(A_j)\, p_r(B \mid A_j)}$$

الإثبات : بنفس طريقة إثبات نظرية (٥-٨).

(٥-٧) أمثلة تطبيقية Applied Examples

تطبيق (٥-١)

في استطلاع للرأي حول مدى متابعة أحد برامج المنوعات الأسبوعية بتليفزيون جمهورية مصر العربية ، فأغذت عينة من 500 مشاهد للتليفزيون من 4 مناطق مختلفة بمدينة القاهرة (A , B , C , D) وتم تسجيل إجابة كل مشاهد في استمارة خاصة به توضح مستوى متابعته للبرنامج . والجدول التالي يوضح توزيع وفقاً للمنطقة ومستوى المنطقة :

مستوى المتابعة / المنطقة	دائماً I	أحياناً II	عدم مشاهدة البرنامج III	المجموع
A	100	20	5	125
B	115	5	5	125
C	50	60	15	125
D	30	50	40	125
المجموع	300	135	65	500

فإذا تم سحب استمارة عشوائياً . أحسب الاحتمالات التالية :

١- احتمال أن تكون الاستمارة لشخص يقيم بالمنطقة B ودائم على متابعة البرنامج.

٢- احتمال أن تكون الاستمارة لشخص بالمنطقة D وأحياناً يتابع البرنامج.

٣- احتمال أن تكون الاستمارة لشخص مقيم بالمنطقة B بشرط أن يكون غير متابع للبرنامج.

٤- احتمال أن تكون الاستمارة لشخص غير متابع للبرنامج بشرط أن يكون مقيم بالمنطقة C.

الحل

إذا فرضنا أن الأحداث D , C , B , A تشير إلى تكون الاستمارة لفرد من المنطقة A ، فرد من المنطقة B ، ..... الخ على الترتيب.

كذلك الأحداث III , II , I تشير إلى أن تكون الاستمارة المسحوبة لشخص دائم على متابعة البرنامج ، لشخص أحياناً يتابع البرنامج ، لشخص غير متابع للبرنامج على الترتيب . وبالتالي فإن :

١- احتمال أن تكون الاستمارة لشخص يقيم بالمنطقة B ودائم على متابعة البرنامج هو :

$$p_r(B \cap I) = \frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}} = \frac{115}{500} = 0.230$$

٢- احتمال أن تكون الاستمارة لشخص بالمنطقة D وأحياناً يتابع البرنامج هو $p_r(D \cap II)$ حيث :

$$p_r(D \cap II) = \frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}} = \frac{50}{500} = 0.10$$

٣- احتمال أن تكون الاستمارة لشخص مقيم بالمنطقة B بشرط أن يكون غير متابع للبرنامج هو $p_r(B \mid III)$ حيث :

$$p_r(B \mid III) = \frac{p_r(B \cap III)}{p_r(III)} = \frac{5/500}{65/500} = \frac{5}{65} = \frac{1}{13}$$

٤- احتمال أن تكون الاستمارة لشخص غير متابع للبرنامج بشرط أن يكون مقيم بالمنطقة C هو $p_r(III \mid C)$ حيث :

$$p_r(III \mid C) = \frac{p_r(III \cap C)}{p_r(C)} = \frac{15/500}{125/500} = \frac{15}{125} = \frac{3}{25}$$

تطبيق (٥-٢)

في إحدى الدراسات عن التوزيع اليومي للجرائد في إحدى المدن الكبرى وجد أن 90% من سكانها يشترون إحدى الجرائد الصباحية يومياً ، كذلك 70% من سكانها يشترون إحدى الجرائد المسائية يومياً ، وبافتراض أن شراء الفرد للجريدة الصباحية مستقل عن شرائه للجريدة المسائية.

المطلوب :

ما هو احتمال شراء أحد سكان هذه المدينة للجريدة الصباحية والجريدة المسائية في نفس اليوم.

الحل

إذا فرضنا أن الحدث A يمثل شراء الجريدة الصباحية ، والحدث B يمثل شراء الجريدة المسائية ، وبالتالي فإن :

$$p_r(A) = 0.90$$
$$p_r(B) = 0.70$$

كذلك احتمال شراء فرد للجريدة الصباحية والمسائية في نفس الوقت (نفس اليوم) هو $p_r(A \cap B)$. وبما أن شراء الجريدة الصباحية مستقل عن شراء الجريدة المسائية فإن :

$$p_r(A \cap B) = p_r(A)\,p_r(B) = 0.90 \times 0.70 = 0.63 = 63\%$$

أي أن 63% من سكان هذه المدينة يشترون جريدة صباحية وجريدة مسائية يومياً.

تطبيق (٥-٣)

تقوم إحدى الشركات بإنتاج نوعين من المنتجات A , B بحيث أن كل منتج 3 مستويات للجودة I , II , III . والجدول التالي يوضح عدد الوحدات المنتجة يومياً من كل نوع وفقاً لمستوى جودتها.

مستوى الجودة نوع المنتج	I	II	III	المجموع
A	4000	9500	2500	16000
B	1000	2500	500	4000
المجموع	5000	12000	3000	20000

المطلوب :

أ – فإذا سحبت وحدة واحدة من الإنتاج اليومي ، ما هو احتمال :

١- أن تكون الوحدة من إنتاج النوع A.

٢- أن تكون الوحدة من إنتاج النوع B.

٣- أن يكون مستوى جودتها III.

٤- أن يكون من النوع B وبمستوى جودة II.

٥- أن تكون من المنتج A بشرط أن تكون من درجة جودة I.

٦- أن تكون درجة جودتها III بشرط أن تكون من المنتج B.

ب- أبني شجرة الاحتمالات للجدول السابق ووضح عليها الاحتمالات القبلية والبعدية.

الحل

١- إذا فرضنا أن $p_r(A)$ هو احتمال أن تكون الوحدة المسحوبة من النوع A حيث :

$$p_r(A) = \frac{16000}{20000} = \frac{16}{20} = \frac{4}{5}$$

٢- وبالمثل إذا فرضنا أن $p_r(B)$ هو احتمال أن تكون الوحدة المسحوبة من النوع B حيث :

$$p_r(B) = \frac{4000}{20000} = \frac{4}{20} = \frac{1}{5}$$

٣- إذا فرضنا أن $p_r(III)$ هو احتمال أن تكون الوحدة بمستوى جودة III ، حيث :

$$p_r(III) = \frac{3000}{20000} = \frac{3}{20}$$

٤- إذا فرضنا أن $p_r(B \cap II)$ هو احتمال أن تكون الوحدة من النوع B وبمستوى جودة II ، حيث :

$$p_r(B \cap II) = \frac{2500}{20000} = \frac{25}{200} = \frac{1}{8}$$

٥- إذا فرضنا أن $p_r(A | I)$ هو احتمال أن تكون الوحدة من النوع A وبمستوى جودة I ، حيث :

$$p_r(A | I) = \frac{p_r(A \cap I)}{p_r(I)} = \frac{4000/20000}{5000/20000} = \frac{4}{5}$$

٦- إذا فرضنا أن $p_r(III | B)$ هو احتمال أن تكون الوحدة بمستوى جودة III ومن النوع B، حيث :

$$p_r(III | B) = \frac{p_r(III \cap B)}{p_r(B)} = \frac{500/20000}{4000/20000} = \frac{500}{4000} = \frac{1}{8}$$

تطبيق (٥-٤)

الجدول التالي يوضح توزيع 2000 حادثة للسيارات الخاصة وفقاً لجنس السائق (ذكر - أنثى)ووفقاً للمدة الزمنية التي تمثل خبرة مرتكب الحادث في قيادة السيارة:-

الخبرة(B) / الجنس(A)	مدة الخبرة في قيادة السيارة بالسنوات				
	أقل من سنة	1-	5-	10 فأكثر	المجموع
ذكر	70	700	580	10	1360
أنثى	90	200	295	55	640
المجموع	160	900	875	65	2000

المطلوب :-

١-احسب كل من الاحتمالات التالية :-

أ- احتمال وقوع حادثة يكون السائق ذكر وخبرته 10 سنوات فأكثر.

ب- احتمال وقوع حادثة يكون السائق أنثى وخبرتها أقل من 10 سنوات .

جـ- احتمال وقوع حادثة يكون قائد السيارة فيها ذكر بشرط أن تكون مدة خبرته أقل من سنة.

د- احتمال وقوع حادثة تكون فيها مدة خبرة قائد السيارة 10 سنوات فأكثر.

هـ- احتمال وقوع حادثة يكون فيها مدة خبرة قائد السيارة أقل من 56 سنوات بشرط أن يكون قائد السيارة أنثى .

٢- ارسم شجرة الاحتمالات للجدول السابق ووضح عليها الاحتمالات القبلية والبعدية.

الحل:-

إذا فرضنا أن الأحداث A_1, A_2 تشير إلى جنس قائد السيارة ذكر أو أنثى على الترتيب كذلك إذا فرضنا أن الأحداث B_1, B_2, B_3, B_4 تشير إلى مدة خبرة قائد السيارة بالسنوات على الترتيب.

(١)

أ- احتمال وقوع حادثة يكون السائق ذكر وخبرته 10سنوات فأكثر

$$= p_r(A_1 \cap B_4) = \frac{10}{2000} = 0.005$$

ب- احتمال وقوع حادثة يكون السائق أنثى وخبرتها أقل من 10 سنوات

$$= P_r(A_2 \, I \; B_1) + P_r(A_2 \, I \; B_2) + P_r(A_2 + B_3)$$

$$= \frac{90}{2000} + \frac{200}{2000} + \frac{295}{2000} = \frac{585}{2000} = 0.2925$$

جـ- بما أن قائدي السيارات التي مدة خبرة كل منهم أقل من 10 سنوات

=160+900+875=1935 سائق

كذلك عدد السائقات الأنثى اللاتي خبرة كل منهن أقل من 10 سنوات

=90+200+295=580 سائق

احتمال وقوع حادثة يكون قائد السيارة فيها ذكر بشرط أن تكون مدة خبرته أقل من 10 سنوات

$$P_r(A_2 \mid B_1 \, Y \, B_2 \, Y \, B_3) = \frac{P_r[(A_2 \, I \; (B_1 \, Y \, B_2 \, Y \, B_3)]}{P_r(B_1 \, Y \, B_2 \, Y \, B_3)}$$

$$= \frac{580}{1935} 0.29974$$

د- احتمال وقوع حادثة تكون فيها مدة خبرة قائد السيارة 10سنوات فأكثر

$$P_r(B_4) = \frac{65}{2000} = 0.0325$$

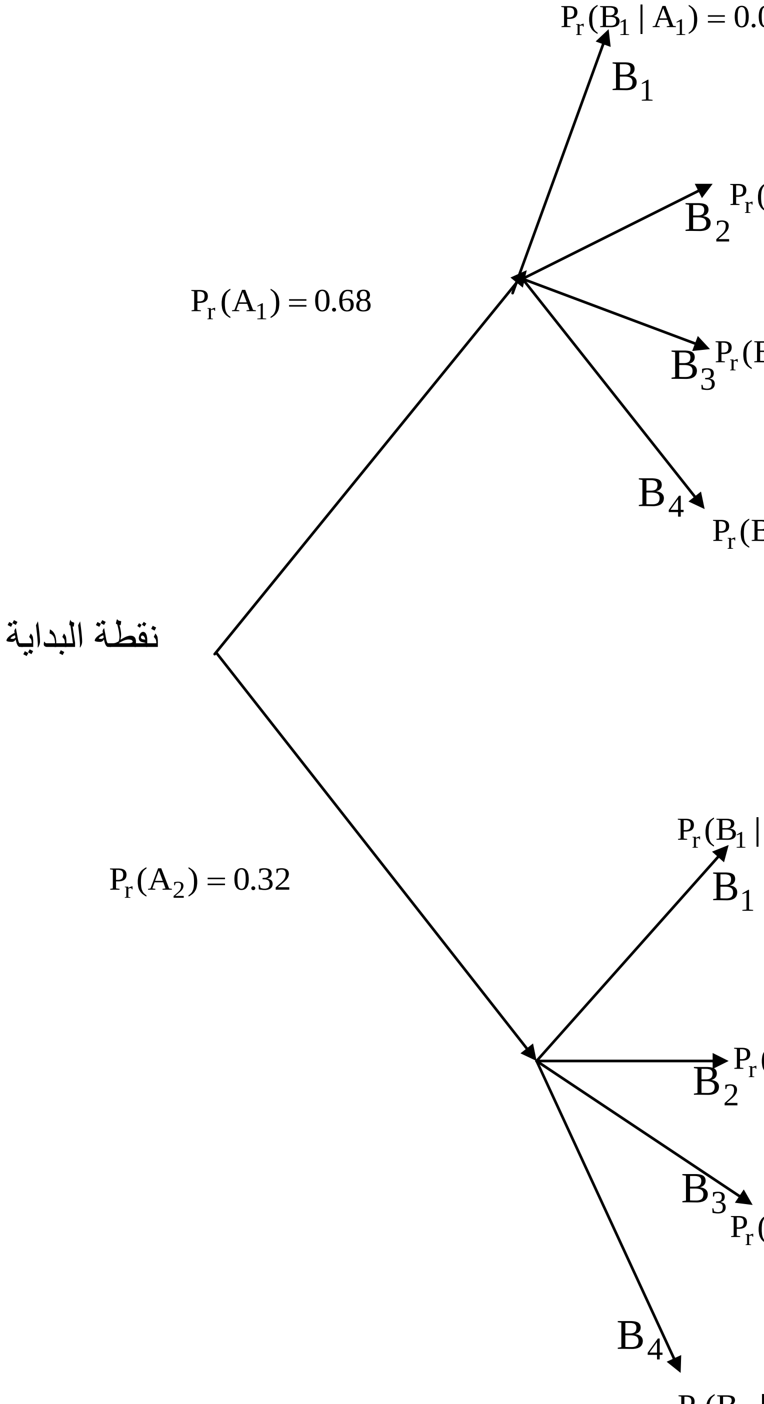
هـ- احتمال وقوع حادثة تكون فيها مدة خبرة قائد السيارة أقل من 5 سنوات بشرط أن يكون قائد السيارة أنثى

$$= P_r[(B_1 \, Y \, B_2) \mid A_2] = \frac{[P_r[(B_1 \, Y \, B_2) \, I \;\; A_2]}{P_r(A_2)}$$

$$= \frac{[P_r(B_1 \, I \;\; B_2) \, Y \, B_2 \, I \;\; A_2]}{P_r(A_2)}$$

$$= \frac{\frac{90}{2000} + \frac{200}{2000}}{\frac{640}{2000}} = \frac{290}{640} = 0.453$$

٢- الشكل التالي يوضح شجرة الاحتمالات موضح عليها الاحتمالات القبلية والبعدية



$$P_r(B_1 \mid A_1) = 0.052 \rightarrow P_r(A_1 \cap B_1) = P_r(A_1)P_r(B_1 \mid A_1) = 0.68 \times 0.052 = 0.035$$

$$P_r(B_2 \mid A_1) = 0.515 \rightarrow P_r(A_1 \cap B_2) = P_r(A_1)P_r(B_2 \mid A_1) = 0.68 \times 0.515 = 0.35$$

$$P_r(B_3 \mid A_1) = 0.427 \rightarrow P_r(A_1 \cap B_3) = P_r(A_1)P_r(B_3 \mid A_1) = 0.68 \times 0.427 = 0.290$$

$$P_r(B_4 \mid A_1) = 0.007 \rightarrow P_r(A_1 \cap B_4) = P_r(A_1)P_r(B_4 \mid A_1) = 0.68 \times 0.007 = 0.005$$

$$P_r(B_1 \mid A_2) = 0.141 \rightarrow P_r(A_2 \cap B_1) = P_r(A_2)P_r(B_1 \mid A_2) = 0.32 \times 0.141 = 0.0451$$

$$P_r(B_2 \mid A_2) = 0.313 \rightarrow P_r(A_2 \cap B_2) = P_r(A_2)P_r(B_2 \mid A_2) = 0.32 \times 0.313 = 0.100$$

$$P_r(B_3 \mid A_2) = 0.461 \rightarrow P_r(A_2 \cap B_3) = P_r(A_2)P_r(B_3 \mid A_2) = 0.32 \times 0.461 = 0.148$$

$$P_r(B_4 \mid A_2) = 0.086 \rightarrow P_r(A_2 \cap B_4) = P_r(A_2)P_r(B_4 \mid A_2) = 0.32 \times 0.086 = 0.028$$

تطبيق (٥-٥)

ترغب إحدى الشركات في تحديث الجهاز الإداري بها فقامت بعمل دورة تدريبية للإداريين بالشركة على استخدام الحاسب الآلي بدلاً من العمل اليدوي وفي نهاية الدورة تم تصنيف العاملين بالجهاز الإداري إلى:-

١- إداريين لهم القدرة على التكيف مع التطوير

٢- إداريين لهم القدرة على التكيف مع التطوير وذلك وفقاً لمدة العمل لكل منهم- كما هو موضح بالجدول التالي:-

المجموع	15 سنة فأكثر (B_4)	10- (B_3)	5- (B_2)	أقل من 5 سنوات (B_1)	مدة الخدمة(B) / القدرة على التكيف(A)
100	8	15	17	60	قادر على التكيف (A_1)
100	50	25	20	5	غير قادر على التكيف (A_2)
200	58	40	37	65	المجموع

فإذا تم اختيار أحد الإداريين عشوائياً.

١- احسب كل من الاحتمالات التالية :-

أ- احتمال قدرة الإدارة على التكيف مع التطوير وأن تكون مدة خدمته 15 سنة فأكثر.

ب-احتمال قدرة الإداري على التكيف مع التطوير بشرط أن تكون مدة خدمته 15 سنة فأكثر .

جـ- احتمال عدم قدرة الإدارة على التكيف مع التطوير.

د- احتمال أن تكون مدة خدمته 5 سنوات فأكثر وان يكون غير قادر على التكيف.

٢- ارسم شجرة الاحتمالات موضح عليها الاحتمالات القبلية والبعدية

الحل:-

أ- احتمال قدرة الإداري علي التكيف مع التطوير و أن تكون مدة خدمته 15 سنة فأكثر

$$= P_r(A_1 \cap B_4) = \frac{8}{200} = 0.04$$

ب-احتمال قدرة الإداري علي التكيف مع التطوير بشرط أن تكون مدة خدمته 15 سنة فأكثر

$$= P_r(A_1 \mid B_4) = \frac{8}{58} = 0.138$$

جـ- احتمال عدم قدرة الإداري علي التكيف مع التطوير

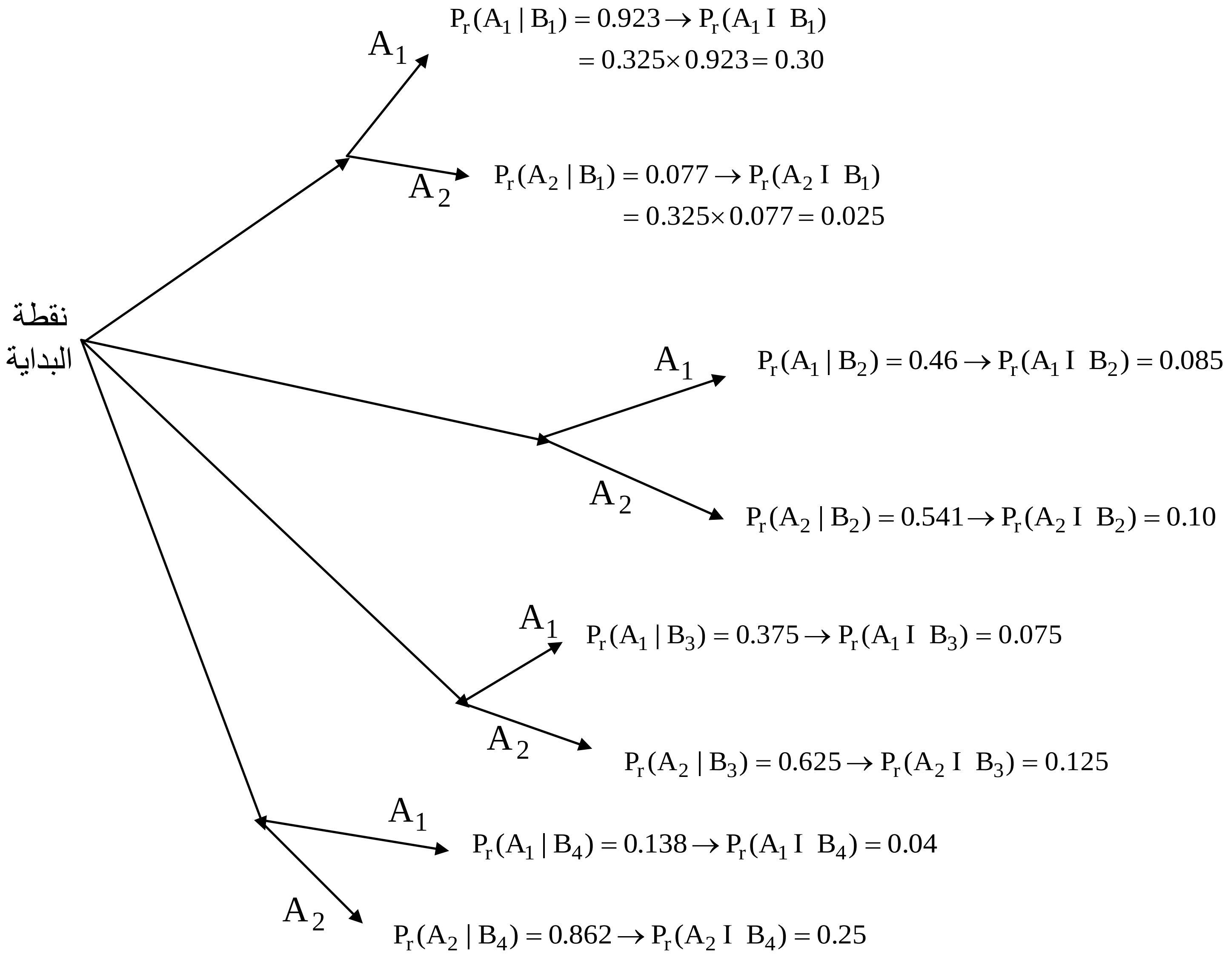
$$= P_r(A_2) = \frac{100}{200} = 0.50$$

د- احتمال أن تكون مدة خدمته 5 سنوات فأكثر وأن يكون غير قادر على التكيف

$$= P_r(A_2 \cap B_2) + P_R(A_2 \cap B_3) + P_r(A_2 \cap B_4)$$

$$= \frac{20}{200} + \frac{25}{200} + \frac{50}{200} + \frac{90}{200} = 0.475$$

٢- الشكل التالي يوضح شجرة الاحتمالات للجدول السابق موضح عليها الاحتمالات القبلية والبعدية:-



## (٥-٨) تمرينات		Exercises

(٥-١)

إذا كان A , B حدثين في فراغ المعاينة S بحيث :

$p_r(A) = 0.3$ ، $A, B \subseteq S$

$p_r(B) = 0.5$ ، $p_r(A \cap B) = 0.1$

المطلوب :

أحسب الاحتمالات التالية:

$i)\, p_r(A)$		$ii)\, p_r(A \cup B)$		$iii)\, p_r(A \cap B)'$

(٥-٢)

الجدول التالي يوضح توزيع 500 سرير بأحد المستشفيات وفقاً لمستوى الخدمة المقدمة للمريض بهذه المستشفى :

مستوى الخدمة المقدمة	متوسط	فوق المتوسط	ممتاز	المجموع
عدد الأسر	100	280	120	500

المطلوب :

١- احتمال أن يكون مستوى الخدمة متوسط.
٢- احتمال أن يكون مستوى الخدمة ممتاز.
٣- احتمال أن يكون مستوى الخدمة أقل من المتوسط.

(٥-٣)

تنتج إحدى شركات الأغذية المحفوظة نوع معين من الخضروات المحفوظة ، حيث يستمر الإنتاج اليومي خلال ورديتين ، الوردية الأولي تبدأ من الساعة الثامنة صباحاً إلى الساعة الثالثة ظهراً ، والوردية الثانية تبدأ من الساعة الثالثة ظهراً حتى الساعة العاشرة مساء . فإذا كان إنتاج الوردية الأولي يمثل 60% من الإنتاج اليومي ، ومن دراسة المنتج في أحد أيام الأسبوع وجد أن 3% من إنتاج الوردية الأولي معيب ، كذلك 5% من إنتاج الوردية الثانية معيب.

فإذا سحبت إحدى العبوات المنتجة عشوائياً ، أحسب الاحتمالات التالية :

١- أن تكون الوحدة المسحوبة سليمة ومن إنتاج الوردية الأولي.
٢- أن تكون العبوة المسحوبة معيبة ومن إنتاج الوردية الثانية.
٣- احتمال أن تكون العبوة سليمة بشرط أن تكون من إنتاج الوردية الثانية.
٤- احتمال أن تكون العبوة المسحوبة من إنتاج الوردية الثانية بشرط أن تكون سليمة.

(٥-٤)

الجدول التالي يوضح توزيع 1000 طالب في السنة النهائية بالمرحلة الثانوية وفقاً لمستوى الطالب والمحافظة التي يقيم فيها.

المحافظة التي يقيم فيها / مستوى الطالب	القاهرة	الإسكندرية	الفيوم	المجموع
أقل من المتوسط	50	50	100	200
متوسط	150	175	75	400
عالي	300	75	25	400
المجموع	500	300	200	1000

فإذا تم اختيار طالب عشوائياً . احسب الاحتمالات التالية :

١- احتمال أن يكون مستوى الطالب عالي.
٢- احتمال أن يكون مستوى الطالب عالي وأن يكون من القاهرة.
٣- احتمال أن يكون مستوى الطالب عالي بشرط أن يكون من القاهرة.
٤- احتمال أن يكون من الفيوم بشرط أن يكون مستواه متوسط.
٥- احتمال أن يكون من الإسكندرية ومستواه أقل من المتوسط.

(٥-٥)

في استطلاع للرأي عن إمكانية إحداث تغير اقتصادي واجتماعي خلال الخمسة أعوام القادمة . أخذت عينة مكونة من 5000 فرد في فئات عمريه مختلفة ، وسئل كل منهم عن إمكانية التغيير . والجدول التالي يوضح إجابات أفراد العينة وفقاً لفئات أعمارهم المختلفة :

إجابة الفرد / عمر الفرد بالسنوات	ممكن جداً	ممكن	غير ممكن	المجموع
20-29	500	1300	150	2000
30-39	350	900	250	1500
40 فأكثر	100	300	1100	1500
المجموع	1000	2500	1500	5000

فإذا اختير أحد أفراد العينة عشوائياً . أحسب الاحتمالات التالية :

١- احتمال أن يكون الفرد عمره أقل من ثلاثين عام ويري أنه غير ممكن إحداث تغيير اقتصادي واجتماعي خلال الخمسة أعوام القادمة.
٢- احتمال أن يري أنه ممكن جداً إحداث تغيير بشرط أن يكون عمره 40 عام فأكثر.
٣- احتمال أن يكون عمره أقل من 40 عام ويري عدم إمكانية التغيير.

(٥-٦)

إذا كانت الأحداث $A, B, C, D \subseteq S$ بحيث :

$p_r(A) = 0.45$ ، $p_r(B) = 0.32$ ، $p_r(C \mid A) = 0.5$

$p_r(D \mid B) = 0.72$ ، $p_r(C \mid B) = 0.28$ ، $p_r(D \mid A) = 0.38$

أحسب الاحتمالات التالية :

$i)\, p_r(A \cap C)$ ، $ii)\, p_r(A \cap D)$

$iii)\, p_r(B \cap C)$ ، $iv)\, p_r(B \cap D)$

(٥-٧)

إذا كانت الأحداث $A, B, C \subseteq S$ بحيث :

$p_r(B \mid A) = 0.6$ ، $p_r(A \cap B \cap C) = 0.017$ ، $p_r(C \mid [A \cap B]) = 0.2$

المطلوب :

أحسب $p_r(A)$

(٥-٨)

إذا كان الحدثين A , B حدثين مستقلين بحيث :

$p_r(A) = 0.42$ ، $p_r(B) = 0.55$

المطلوب :

أحسب كل من :

$p_r(A \cap C)$ ، $p_r(A \cup C)$ ، $p_r(A \mid C)$

(٥-٩)

تقوم إحدى المؤسسات العلاجية باستثمار أرباحها في إحدى المشروعات والجدول التالي يعطي معدل العائد والاحتمال المناظر له :

معدل الفائدة	أقل من 0%	0%-9%	10%-15%	16%-20%	المجموع
الاحتمال	0.04	0.15	0.70	0.11	1.0

فإذا فرضنا أن الحدث A يشير إلى معد العائد أكبر من 10% ، والحدث B يشير إلى معد العائد سالب.

المطلوب :

١- أوجد احتمال الحدث A.
٢- أوجد احتمال الحدث B.
٣- أوصف الأحداث التالية : B', A' .
٤- أوجد كل من الأحداث B', A' .
٥- أوصف الحدث A تقاطع B.
٦- أوصف الحدث A إتحاد B.
٧- أوجد احتمال A إتحاد B.
٨- أوجد احتمال A تقاطع B.
٩- أوجد الاحتمال المكمل لاحتمال $A \cap B$.

(٥-١٠)

أجرى أحد السوبر ماركت دراسة على نوع العملاء (المترددين وغير المترددين) عن آرائهم في استخدام أحد منتجات الألبان . وتم تسجيل نسبة العملاء وفقاً لآرائهم في الجدول التالي :

آرائهم في المنتج / نوع العميل	آراء العملاء في استخدام المنتج		
	غالباً	أحياناً	غير مستخدمين
متردد	0.12	0.48	0.19
غير متردد	0.07	0.06	0.08

فإذا تم اختيار أحد العملاء أوجد :

١- احتمال أن يكون من العملاء المترددين على السوبر ماركت وغالباً يستخدم المنتج.
٢- احتمال أن يكون من العملاء المترددين على السوبر ماركت بشرط أن يكون من المستخدمين دائماً للمنتج.
٣- احتمال أن يكون غير مستخدم ومن المترددين على السوبر ماركت.
٤- احتمال أن يكون غير مستخدم بشرط أن يكون من المترددين على السوبر ماركت.
٥- احتمال أن يكون من المترددين بشرط أن يكون غير مستخدم للمنتج.
٦- أرسم شجرة الاحتمالات ووضح عليها الاحتمالات الشرطية واحتمالات التقاطع.

(٥-١١)

الجدول التـالي يوضـح توزيـع 1000 فرد ممـا يجيـدون استخدام الحاسـب والعـاملين بالجهاز الحكومي - تم تصنيفهم وفقاً للحالة الاجتماعية ووفقاً لاستمرارهم فـي عملهم أو تركه إلى عمل أخر.

حالة العمل / الحالة الاجتماعية	الاستمرار في نفس العمل	الانتقال إلى عمل أخر
متزوج	350	180
أعزب	150	420

أ – فإذا تم اختيار أحد العاملين في هذا الجهاز ويجيد استخدام الحاسب . أوجد :

١- احتمال أن يكون متزوج وانتقل إلى عمل أخر.

٢- احتمال أن يكون كتزوج بشرط انتقاله إلى عمل أخر.

٣- احتمال انتقاله إلى عمل أخر بشرط أن يكون متزوج.

٤- احتمال أن يكون أعزب.

٥- احتمال استمراره في نفس العمل.

ب- أرسم شجرة الاحتمالات التـي تصـف الجـدول السـابق موضـح عليهـا الاحتمـالات الشرطية واحتمالات التقاطع.

صفحة (١٧٩)

$$p_r(B) = \frac{835}{1000}$$

$$p_r(A_1 \mid B) \qquad p_r(A_2 \mid B) \qquad p_r(A_3 \mid B)$$

$$p_r(B') = \frac{165}{1000}$$

$$p_r(A_1 \mid B') \qquad p_r(A_2 \mid B') \qquad p_r(A_3 \mid B')$$

صفحة (١٨٠)

$$P_r = (B \mid A_1) = \frac{370}{400}$$

الباب السادس

عرض البيانات الكمية

The Presentation Of Quantitative Data

(٣-١)التوزيع التكراري البسيط

Simple Frequency Distribution

(٣-٢) التوزيعات التكرارية التراكمية

Cumulative Frequency Distribution

(٣-٣)التوزيعات التكرارية النسبية

Relative Frequency Distribution

(٣-٤)التوزيع التكراري المزدوج

Double Frequency Distribution

(٣-٥)أمثلة تطبيقية

Applied Examples

(٣-٦) تمرينـــات

Exercises

(٦-١) التوزيع الاحتمالي المتقطع
Discrete Probability Distribution

من الباب السابق يتضح أن النتائج الممكنة للتجربة العشوائية (أو الأحداث) تمثل متغير عشوائي Random Variable . فإذا كانت x تشير إلى النتائج الممكنة للتجربة فإن x تأخذ قيم مختلفة وفقاً لتعريف التجربة العشوائية في الباب السابق وبالتالي فإن x متغير كذلك عشوائي لأنه لا يمكن معرفة قيمة x قبل إجراء التجربة ، حيث أن العامل العشوائي هو الذي يتحكم في تحديد قيمة x - وبالتالي فإن نتائج التجربة x هي متغير عشوائي أي متغير تتحكم الصدفة فقط في تحديد قيمته.

وفي هذا الباب والباب التالي سوف نتناول بالدراسة المتغيرات العشوائية (أي النتائج الممكنة للتجارب) وتوزيعاتها الاحتمالية . فالمتغير العشوائي إما أن يكون متغير متقطع (منفصل) Discrete Random Variable أو يكون متغير متصل (مستمر) Continuous Random Variable ، حيث يسمى التوزيع الاحتمالي للمتغير المتقطع بالتوزيع المتقطع ، كذلك يسمى التوزيع الاحتمالي للمتغير المتصل بالتوزيع المتصل للاختصار فقط.
وسوف نتناول في هذا الباب تعريف التوزيع الاحتمالي المتقطع ، كذلك نقدم بعض التوزيعات الاحتمالية المتقطعة النظرية الأكثر استخداماً مثل (توزيع ذو الحدين ، توزيع بواسون ، ...... ، الخ) وأخيراً نقدم أمثلة تطبيقية لهذه التوزيعات المتقطعة المقدمة في هذا الباب.

التوزيع الاحتمالي المتقطع Discrete Probability Distribution

إذا كانت x تشير إلى النتائج الممكنة للتجربة حيث يمكن عدها Countable ، فإن x تسمى بالمتغير المتقطع (المنفصل) حيث :

$$X = x_1, x_2, x_3, ... \qquad (6.1)$$

فإذا كان عدد هذه النتائج محدود Finite بعدد n بمعنى :

$$X = x_1, x_2, x_3, ..., x_n \qquad (6.2)$$

أو غير محدود Infinite بمعنى :

$$X = x_1, x_2, x_3, ... \qquad (6.3)$$

مثال (٦-١)
يعتبر المتغير العشوائي الذي يمثل النتائج الممكنة لرمى زهرة طاولة (نرد) متغير عشوائي متقطع محدود بالرقم 6 حيث :

$$X = 1 , 2 , 3 , 4 , 5 , 6 \tag{6.4}$$

كذلك عدد الميكروبات في لتر ماء من إحدى الترع يعتبر متغير عشوائي متقطع ولكن غير محدود حيث :

$$X = 1 , 2 , 3 , 4 , \ldots\ldots. \tag{6.5}$$

دالة الاحتمال
إذا كانت الدالة $f(x)^*$ حيث :

$$f(x) = \begin{cases} p_r(x_j) & \text{for all (j)} \\ 0 & \text{فيما عدا ذلك} \end{cases} \tag{6.6}$$

حيث تسمى الدالة f(x) بدالة الاحتمال للمتغير x عندما $X = x_j$.

مثال (٦-٢)
إذا رميت زهرة طاولة (نرد) عشوائياً فإن النتائج الممكنة للرمي تمثل متغير عشوائي له دالة الاحتمال f(x) التالية :

$$f(x_j) = \begin{cases} \frac{1}{6} & j=1,2,3,4,5,6 \\ 0 & \text{فيما عدا ذلك} \end{cases} \tag{6.7}$$

وتسمى العلاقة (6.6) بالتوزيع الاحتمالي للمتغير X ، وبالتالي فالتوزيع الاحتمالي للمتغير هو عبارة عن القيم المختلفة للمتغير والاحتمالات المناظرة لها ، ويمكن أن يكتب التوزيع في صورة دالة احتمال والقيم المختلفة للمتغير X المناظرة لها (أى مجال أو نطاق الدالة) ويمكن التعبير أيضاً عن التوزيع الاحتمالي في صورة جدول يسمى بجدول التوزيع الاحتمالي.

* أ.د. عفاف الدش : " الرياضيات وصنع القرار" – الباب الثاني ، مكتبة عين شمس (١٩٩٤) القاهرة ، رقم إيداع ٩٤/٨٦٨٢.

وجدول التوزيع الاحتمالي هو جدول مكون من عموديين (أو صفين) العمود (أو الصف) الأول يحتوى على جميع القيم الممكنة للمتغير (أو جميع النتائج الممكنة للتجربة) والعمود أو الصف الثاني يحتوى على الاحتمالات المناظرة لقيم المتغير في العمود (أو الصف) الأول.

فالتوزيع الاحتمالي للمتغير X في المثال السابق في صورة جدول توزيع احتمالي كما هو موضح بالجدول (٥-١) التالي :

جدول (٦-١)

قيم x_j	$p_r(x_j)$
$x_1 = 1$	1/6
$x_2 = 2$	1/6
$x_3 = 3$	1/6
$x_4 = 4$	1/6
$x_5 = 5$	1/6
$x_6 = 6$	1/6
المجموع (Σ)	1

مثال (٦-٣)
إذا رميت ثلاثة قطع عملة متوازنة عشوائياً . أوجد التوزيع الاحتمالي لعدد الصور التي تظهر عند الرمي :

١- في صورة جدول توزيع احتمالي.
٢- في صورة دالة احتمال.

الحل
إذا فرضنا أن X تمثل عدد الصور التي تظهر عند الرمي فإن :

$$X = 0 , 1 , 2 , 3$$

١- والجدول التالي يوضح جميع النتائج الممكنة والقيم المختلفة للمتغير x والاحتمالات المناظرة لهذه القيم.

جدول (٦-٢)

النتائج الممكنة	قيم المتغير x_j	$p_r(x_j)$
(H,H,H)	3	1/8
(H,W,H)	2	1/8
(H,H,W)	2	1/8
(W,H,H)	2	1/8
(W,W,W)	0	1/8
(W,H,W)	1	1/8
(H,W,W)	1	1/8
(W,W,H)	1	1/8
المجموع(Σ)		8/8=1

ومن الجدول السابق يمكن تكوين جدول التوزيع الاحتمالي التالي للمتغير (x)

جدول (٦-٣)

قيم (x)	0	1	2	3	المجموع(Σ)
f(x)	1/8	3/8	3/8	1/8	1

٢- ويمكن التعبير عن التوزيع الاحتمالي السابق في صورة دالة رياضية f(x) على النحو التالي :

$$f(x) = \begin{cases} \frac{1}{8} C_x^3 & x=0,1,2,3 \\ 0 & \text{فيما عدا ذلك} \end{cases} \qquad (6.8)$$

فنجد أن :

$$f(x=0) = \frac{1}{8} C_0^3 = \frac{1}{8}(1) = \frac{1}{8}$$

$$f(x=1) = \frac{1}{8} C_1^3 = \frac{1}{8}(3) = \frac{3}{8}$$

$$f(x=2) = \frac{1}{8} C_2^3 = \frac{1}{8}(3) = \frac{3}{8}$$

$$f(x=3) = \frac{1}{8} C_3^3 = \frac{1}{8}(1) = \frac{1}{8}$$

خصائص دالة الاحتمال

فيما يلي سوف نلخص أهم خصائص دالة الاحتمال f(x) للمتغير المتقطع X :

١- تنحصر قيمة $f(x_i)$ لجميع قيم i في الفترة [0 , 1]
أى أن :

$$0 \le f(x_i) \le 1 \qquad , \mathrm{i} = 1,2,3,\ldots\ldots,\mathrm{n} \qquad (6.9)$$

٢- مجموع قيم $f(x_i)$ لجميع قيم $\mathrm{i} = 1,2,3,\ldots\ldots,\mathrm{n}$ تساوى الواحد الصحيح
أي أن :

$$\sum_{i=1}^{n} f(x_i) = 1 \qquad (6.10)$$

مثال (٦-٤)
الجدول التالي يوضح توزيع 500 أسرة حسب متوسط الدخل الشهري بالجنية للأسرة.

جدول (٦-٤)

متوسط دخل الأسرة* بالجنية (x)	150	200	300	400	500	المجموع
عدد الأسر	30	70	300	80	20	500

فإذا اخترت أسرة واحدة عشوائياً من هذه الأسر أوجد :
١- احتمال أن تكون هذه الأسرة لها متوسط الدخل 200 جنية.
٢- احتمال أن تكون هذه الأسرة لها متوسط دخل أقل من أو يساوى 400 جنية.

الحل

إذا فرضنا أن X متغير يشير إلى متوسط دخل الأسرة ، فالجدول التالي يوضح التوزيع الاحتمالي للمتغير X .

* الصف الأول بالجدول الذي يمثل الفئات نجد أن كل فئة فيه مكونة من عنصر واحد وهو يعتبر مركز الفئة.

جدول (٦-٥)

x	عدد الأسر	f (x)
150	30	30/500 = 0.06
200	70	70/500 = 0.14
300	300	300/500 =0.60
400	80	80/500 = 0.16
500	20	20/500 = 0.04
المجموع	500	500/500 = 1

١- احتمال أن يكون متوسط دخل الأسرة 200 جنية يساوى $f\,(x = 200)$ حيث أن :

$$f\,(x = 200) = \frac{70}{500} = 0.14$$

٢- احتمال أن يكون متوسط دخل الأسرة أقل من أو يساوى 400 جنية هو :

$$f\,(x \le 400) = f\,(x = 150) + f\,(x = 200) + f\,(x = 300) + f\,(x = 400)$$

$$= \frac{30}{500} + \frac{70}{500} + \frac{300}{500} + \frac{80}{500} = \frac{480}{500} = \frac{48}{50} = 0.96$$

أو

$$f\,(x \le 400) = 1 - f\,(x = 500) = 1 - \frac{20}{500} = 1 - 0.04 = 0.96$$

(٦-٢) عملية برنولي **Bernoulli Process**

في هذا الباب سوف نتناول بالتفصيل بعض التوزيعات الاحتمالية المتقطعة (المنفصلة) الهامة مثل توزيع ذات الحدين ، والتوزيع الهيبروجومترك ، وتوزيع بواسون نظراً لأهمية هذه التوزيعات من الناحيتين النظرية والتطبيقية.

فمن الناحية النظرية نجد أن التوزيعات المذكورة أعلاه معروف لكل منها الصياغة الرياضية Mathematical Formulation لدالة الاحتمال ، كذلك الصياغات الرياضية للمؤشرات التي تصف الخصائص الرياضيةMathematical Features للتوزيع كما سوف يتضح في الفصول التالية.

أما من الناحية التطبيقية فسوف يتضح من الأمثلة التطبيقية المقدمة أن كثير من المتغيرات العشوائية التي تمثل عناصر أساسية في المشاكل الفعلية Real Problems الهامة كثير منها يتبع التوزيعات المذكورة أعلاه.

في القرن السابع عشر الميلادي قدم عالم الرياضيات السويدي* جيمس برنولي James Bernoulli تجربة عشوائية Random Experiment تتمثل في وجود عدد n من المحاولات المتماثلة Identical Trails (المحاولات المتماثلة هي المحاولات التي لكل منها نفس الظروف ونفس النتائج للمحاولات الأخرى).

حيث لكل محاولة نتيجتين متنافيتين (أي ين متنافيين) فقط أطلق عليهما نجاح باحتمال ثابت وليكن (p) وفشل باحتمال ثابت يساوى (p - 1) ، واعتبر أن كل محاولة من المحاولات n مستقلة عن المحاولة الأخرى ، ونسمي المتغير العشوائي الذي يمثل عدد مرات النجاح في n محاولة من المحاولات المتماثلة بمتغير ذات الحدين (حيث أن للمحاولة الواحدة نتيجتين فقط) Binomial Random Variable.

ومما سبق يمكن تعريف عملية برنولي (نسبة إلى العالم الذي قدمها وحدد خصائصها) بأنها n من المحاولات المتماثلة التي تتميز بالآتي :
١- محاولة متماثلة مستقلة ، لكل محاولة نتيجتين متنافيتين فقط
٢- احتمال النجاح (p) واحتمال الفشل (p - 1) ثابت لا يتغير من محاولة لأخرى

مثال (٦-٥)
إذا رميت قطعة عملة متوازنة عشوائياً n من المرات (أو رميت n من قطع العملة المتوازنة عشوائياً) فهذه العملية تعتبر عملية برنولي حيث يتوفر فيها خصائص عملية برنولي على النحو التالي :

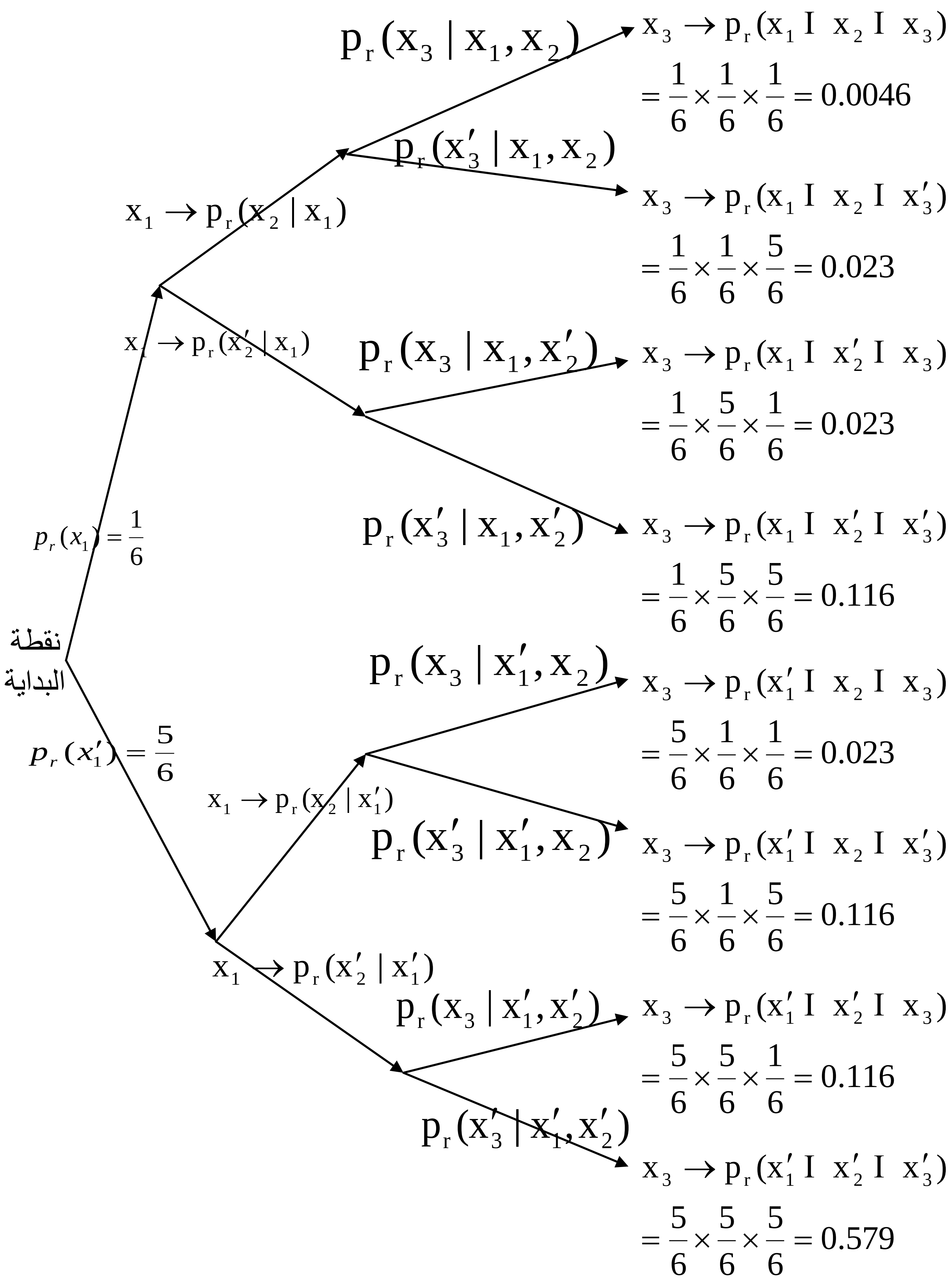
١- عملية رمي قطعة العملة يمثل محاولة وكل رمية مستقلة عن الرمية الأخرى (أو كل قطعة عملة مستقلة عن القطعة الأخرى في حالة وجود n من قطع العملة) ، ولكل رمية نتيجتين متنافيتين فقط هما صورة (نجاح) وكتابة (فشل).
٢- احتمال ظهور صورة في المحاولة الواحدة (أي احتمال النجاح) يساوى (1/2) واحتمال الكتابة (أي احتمال الفشل) يساوى (1/2) ، وهذان الاحتمالان لا يتغيران من محاولة لأخرى ، بمعني ثبات احتمال النجاح واحتمال الفشل في جميع المحاولات.

* المرجع السابق صفحة (٢٠٤)

مثال (٦-٦)
إذا رميت عشوائياً ثلاثة زهرات نرد (طاولة) متوازنة ، وإذا فرضنا أن النتيجة محل الاهتمام لدينا هو ظهور الرقم 4 على كل زهرة من الزهرات الثلاثة.
فتجد أن هذه العملية تمثل عملية برنولي حيث أن :

١- كل زهرة تمثل محاولة ونعتبر أن النتيجة محل الاهتمام وهي ظهور الرقم 4 هي النجاح وظهور أي رقم أخر غير الـ 4 يمثل الفشل ، وبالتالي فإن احتمال النجاح يساوى (1/6) واحتمال الفشل يساوى (5/6) . وتجد أن كل محاولة مستقلة عن المحاولة الأخرى (أي ظهور 4 على الزهرة الأولى مستقل عن ظهور 4 على الزهرة الثانية والثالثة ، كذلك ظهور الرقم 4 على الثانية مستقل عن ظهور 4 على الأولى والثالثة وبالمثل ظهور الرقم 4 على الثالثة مستقل عن ظهور الرقم 4 على الأولى والثانية).
٢- احتمال ظهور الرقم 4 على الزهرة الأولى يساوى احتمال ظهور 4 على الثانية يساوى احتمال ظهور الرقم 4 على الثالثة يساوى (1/6).
٣- يمكن تصوير هذه التجربة بشجرة الاحتمالات التالية ، فإذا فرضنا أن ظهور الرقم 4 على الزهرة رقم j هو الحدث x_j حيث j = 1,2,3 ، كذلك الحدث x'_j هو ظهور أي رقم أخر غير الرقم 4.

ويمكن توضيح هذه العملية باستخدام شجرة الاحتمالات كما هو موضح في الشكل التالي:


نقطة البداية
$p_r(x_1) = \frac{1}{6}$
$p_r(x_1') = \frac{5}{6}$
$x_1 \rightarrow p_r(x_2 \mid x_1)$
$x_1 \rightarrow p_r(x_2' \mid x_1)$
$x_1 \rightarrow p_r(x_2 \mid x_1')$
$x_1 \rightarrow p_r(x_2' \mid x_1')$
$p_r(x_3 \mid x_1, x_2)$
$p_r(x_3' \mid x_1, x_2)$
$p_r(x_3 \mid x_1, x_2')$
$p_r(x_3' \mid x_1, x_2')$
$p_r(x_3 \mid x_1', x_2)$
$p_r(x_3' \mid x_1', x_2)$
$p_r(x_3 \mid x_1', x_2')$
$p_r(x_3' \mid x_1', x_2')$
$x_3 \rightarrow p_r(x_1 \cap x_2 \cap x_3) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = 0.0046$
$x_3 \rightarrow p_r(x_1 \cap x_2 \cap x_3') = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} = 0.023$
$x_3 \rightarrow p_r(x_1 \cap x_2' \cap x_3) = \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} = 0.023$
$x_3 \rightarrow p_r(x_1 \cap x_2' \cap x_3') = \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} = 0.116$
$x_3 \rightarrow p_r(x_1' \cap x_2 \cap x_3) = \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6} = 0.023$
$x_3 \rightarrow p_r(x_1' \cap x_2 \cap x_3') = \frac{5}{6} \times \frac{1}{6} \times \frac{5}{6} = 0.116$
$x_3 \rightarrow p_r(x_1' \cap x_2' \cap x_3) = \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6} = 0.116$
$x_3 \rightarrow p_r(x_1' \cap x_2' \cap x_3') = \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = 0.579$

شكل (٦-١)

(٦-٣) توزيع ذات الحدين **Binominal Distribution**

في الفصل السابق عرفنا متغير " ذات الحدين" هو عبارة عن عدد مرات النجاح في n محاولة من محاولات برنولي . وفي هذا الفصل سوف نتناول التوزيع الاحتمالي للمتغير ذات الحدين وأهم الخصائص المميزة له.

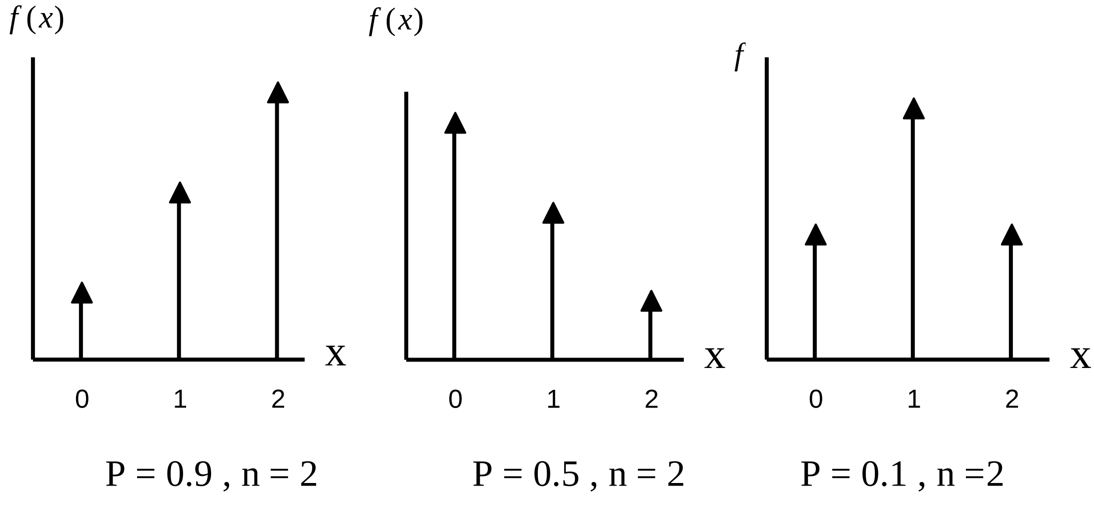
تعريف (٦-١)

إذا كان X متغير عشوائي يمثل عدد مرات النجاح في n من التجارب المتماثلة والمستقلة حيث احتمال النجاح في التجربة الواحدة (p) واحتمال الفشل (1 - p) . فإن دالة الاحتمال للمتغير X هي f(x) حيث :

$$f(x) = \begin{cases} C_x^n (P)^x (1-P)^{n-x} & x=0,1,2,3,\ldots,n \\ 0 & \text{فيما عدا ذلك} \end{cases} \qquad (6.11)$$

حيث أن : $0 < P < 1$ ، $C_x^n = \dfrac{n!}{x!(n-x)!}$

والأشكال التالية توضح شكل التوزيع الاحتمالي لمتغير ذات الحدين عند قيم مختلفة لكل من المعلمتين P , n .



شكل (٦-٢)

ونظراً لأهمية توزيع ذات الحدين في دراسة العديد من المشاكل الفعلية في كثير من المجالات الاقتصادية ، التجارية ، الطبية ، الموسيقية ، ...... ، الخ. والتي يتناولها الكثير من غير المتخصصين كذلك لتوفير الوقت والجهد للمتخصصين ، فقد صممت ١٩٤٩* جداول باستخدام الحاسب الآلي Computer تعطي مباشرة قيم الاحتمالات $f(x = j)$ عند القيم المحددة لكل من n , p , j وتسمى بجداول احتمالات ذات الحدين Binomial Probability Tables ، وفي ملحق رقم (٣) نقدم جزء من هذه الجداول.

خصائص دالة الاحتمال

لتوضيح خصائص دالة الاحتمال لتوزيع ذات الحدين سوف نقدم المثال التالي

مثال (٦-٧)

إذا كان $n = 2,4,6$. أحسب قيم $f(x)$ عندما $P = (0.5)$ أو $P = (0.1)$ أو $P = (0.9)$
ثم أرسم التوزيع في كل حالة من الحالات السابقة.

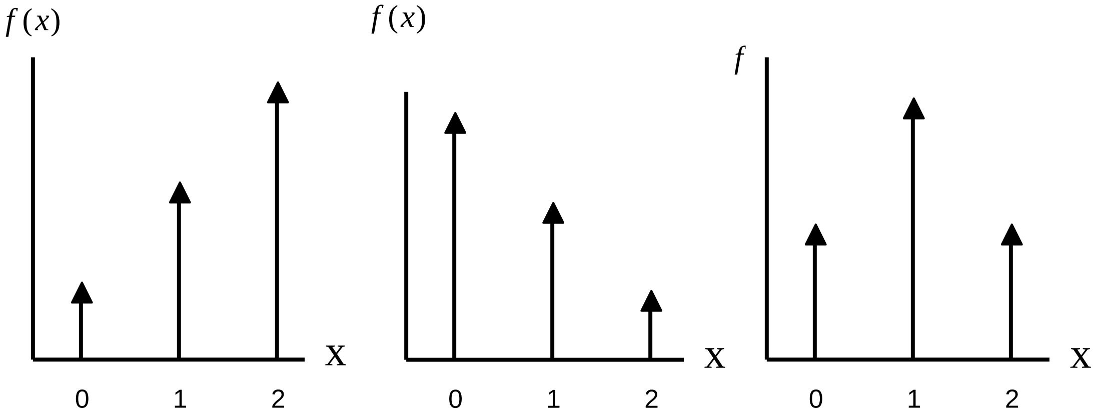
بما أن كل معلمة من معالم التوزيع n , p تأخذ لكل منهما ثلاثة قيم حيث $n = 2,4,6$ و $P = 0.1,0.5,0.9$ وبالتالي يصبح لدينا 9 حالات أو توزيعات . وفي الجدول التالي سوف نقوم بحساب قيم $f(x)$ في كل حالة من الحالات السابقة :

جدول (٦-٦)

n	p / J	0.1	0.5	0.9
2	0	$C_0^2(0.1)^0(0.9)^2 = 0.81$	$C_0^2(0.5)^0(0.5)^2 = 0.25$	$C_0^2(0.9)^0(0.1)^2 = 0.01$
	1	$C_1^2(0.1)\ (0.9) = 0.18$	$C_1^2(0.5)\ (0.5) = 0.50$	$C_1^2(0.9)\ (0.1) = 0.18$
	2	$C_2^2(0.1)^2(0.9)^0 = 0.01$	$C_2^2(0.5)^2(0.5) = 0.25$	$C_2^2(0.9)^2(0.1)^0 = 0.81$
4	0	$C_0^4(0.1)^0(0.9)^4 = 0.956$	$C_0^4(0.5)^0(0.5)^4 = 0.0625$	$C_0^4(0.9)^0(0.1)^4 = 0.0001$
	1	$C_1^4(0.1)\ (0.9)^3 = 0.292$	$C_1^4(0.5)\ (0.5)^3 = 0.250$	$C_1^4(0.9)\ (0.1)^3 = 0.003$
	2	$C_2^4(0.1)^2(0.9)^2 = 0.049$	$C_2^4(0.5)^2(0.5)^2 = 0.375$	$C_2^4(0.9)^2(0.1)^2 = 0.049$
	3	$C_3^4(0.1)^3(0.9) = 0.004$	$C_3^4(0.5)^3(0.5) = 0.250$	$C_3^4(0.9)^3(0.1) = 0.292$
	4	$C_4^4(0.1)^4(0.9) = 0.0001$	$C_4^4(0.5)^4(0.5)^0 = 0.0625$	$C_4^4(0.9)^4(0.1)^0 = 0.656$
5	0	$C_0^5(0.1)^0(0.9)^5 = 0.591$	$C_0^5(0.5)^0(0.5)^5 = 0.031$	$C_0^5(0.9)^0(0.1)^5 = 0.00001$
	1	$C_1^5(0.1)\ (0.9)^4 = 0.328$	$C_1^5(0.5)\ (0.5)^4 = 0.155$	$C_1^5(0.9)\ (0.1)^4 = 0.0005$
	2	$C_2^5(0.1)^2(0.9)^3 = 0.073$	$C_2^5(0.5)^2(0.5)^3 = 0.310$	$C_2^5(0.9)^2(0.1)^3 = 0.008$
	3	$C_3^5(0.1)^3(0.9)^2 = 0.008$	$C_3^5(0.5)^3(0.5)^2 = 0.310$	$C_3^5(0.9)^3(0.1)^2 = 0.073$
	4	$C_4^5(0.1)^4(0.9) = 0.0005$	$C_4^5(0.5)^4(0.5) = 0.155$	$C_4^5(0.9)^4(0.1) = 0.328$
	5	$C_5^5(0.1)^5(0.9)^0 = 0.00001$	$C_5^5(0.5)^5(0.5)^0 = 0.031$	$C_5^5(0.9)^5(0.1)^0 = 0.591$

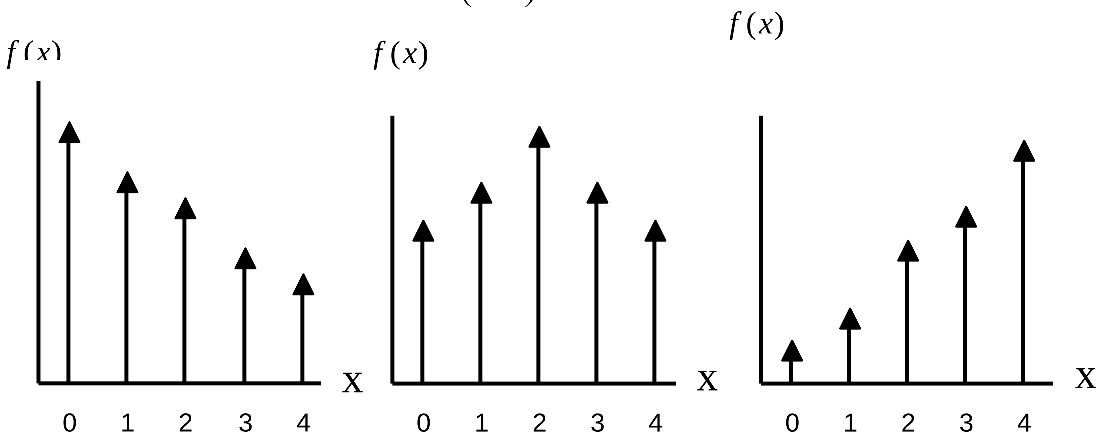
Applied Mathematics Series , No 6,1949,U.S.A.*

والأشكال (٦-٣) ، (٦-٤) ، (٦-٥) يوضح كل منها التوزيعات الاحتمالية عندما $n = 2 , 4 , 5$ على الترتيب عندما $P = 0.1 , 0.5 , 0.4$



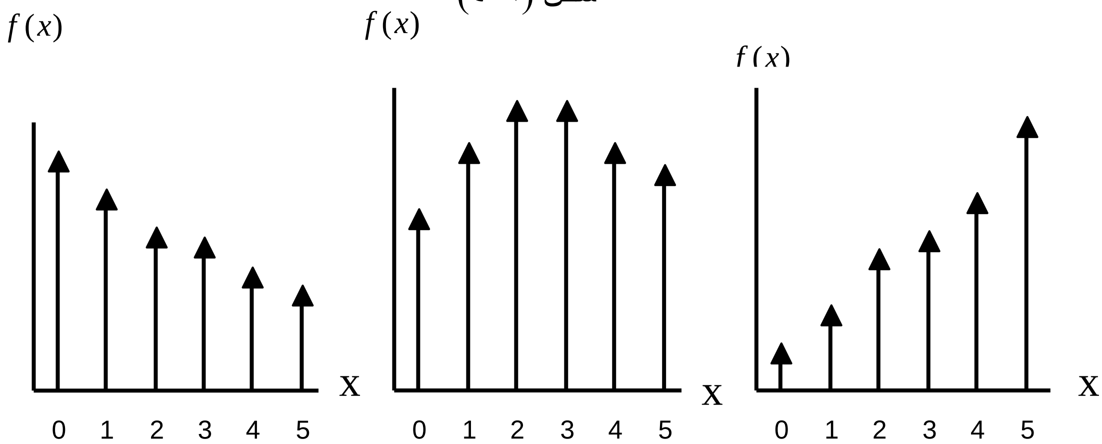
P = 0.9 , n = 2 P = 0.5 , n = 2 P = 0.1 , n =2

شكل (٦-٣)



P = 0.9 , n = 4 P = 0.5 , n = 4 P = 0.1 , n = 4

شكل (٦-٤)



P = 0.9 , n = 5 P = 0.5 , n = 5 P = 0.1 , n = 5

شكل (٦-٥)

من الأشكال السابقة يتضح ما يلي :

١- عندما تكون قيمة $P < 0.5$ يكون التوزيع ملتوي ناحية اليمين (التواء موجب) ، وعندما تكون $P > 0.5$ يكون التوزيع ملتوي ناحية اليسار (التواء سالب) ، وعندما تكون $P = 0.5$ يكون التوزيع متماثل حول القيمة المتوقعة للمتغير X وهي تساوى (np) وذلك عند جميع القيم المختلفة للمعلمة n.

٢- تقل قيمة النهاية العظمى للدالة $f(x)$ بزيادة قيمة n.

٣- عندما $P = 0.5$ فنجد أن التفرطح Kurtsis للتوزيع تزداد بزيادة قيمة n.

والنتائج الثلاثة السابقة نتائج عامة بالنسبة لتوزيع ذات الحدين ويمكن إثباتها رياضياً*

نظرية (٦-١)

إذا كان X متغير عشوائي يتبع توزيع ذات الحدين بمعلمتين P , n فإن القيمة المتوقعة للمتغير X وتباين المتغير X هما nP , nP(1-P) على الترتيب أي أن :

$$E(x) = \mu = \sum_{j=1}^{n} x_j \, f(x_j) = np \qquad (6.12)$$

$$\operatorname{var}(x) = \sigma^2 = \sum_{j=1}^{n} (x_j - \mu)^2 \, f(x_j) = np(1-p) \qquad (6.13)$$

الإثبات

$$\mu = np\left[\sum_{x=1}^{n-1} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1}(1-p)^{n-x}\right]$$

$$= np\left[\sum_{x-1=0}^{n-1} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1}(1-p)^{n-x}\right]$$

$$= \sum_{x-1=0}^{n-1} C_{x-1}^{n-1} p^{x-1}(1-p)^{n-x} = np(1)^{x-1} = np$$

حيث يمثل المقدار $\sum_{x-1=0}^{n-1} C_{x-1}^{n-1} p^{x-1}(1-p)^{n-x}$ مجموع الاحتمالات للمتغير (X-1) الذي يتبع توزيع ذات الحدين أيضاً بمعلمتين p , (n-1).

وبما أن :

$$\sigma^2 = \sum_{x=0}^{n} (x-\mu)^2 \, f(x) = \sum x^2 f(x) - (\mu)^2 \qquad (6.14)$$

* Kendall and Stuart (1977) "The Advanced Theory of Statistics" Vol.1, Charles Griffin & Company Limited , London.

وبما أن :

$$\sum_{x=0}^{n} x^2 f(x) = \sum_{x=0}^{n} [x(x-1)+x] f(x)$$

$$= \sum_{x=0}^{n} x(x-1) f(x) + \sum_{x=0}^{n} x f(x)$$

$$= \sum_{x=0}^{n} x(x-1) f(x) + \mu \qquad (6.15)$$

وبما أن

$$\sum_{x=1}^{n-1} x(x-1) f(x) = \sum_{x=1}^{n-1} x(x-1) C_x^n P^x (1-P)^{n-x}$$

$$= nP \sum_{x=2}^{n-2} \frac{n!}{(x-2)!(n-x)!} P^{x-1} (1-P)^{n-x}$$

$$= P^2 n(n-1) \sum_{x=2}^{n-2} \frac{(n-2)!}{(x-2)!(n-x)!} P^{x-2} (1-P)^{n-x}$$

$$= P^2 n(n-1) \sum_{x=2}^{n-2} C_{x-2}^{n-2} P^{x-2} (1-P)^{n-x}$$

$$= P^2 n(n-1)(1) = P^2 n(n-1) \qquad (6.16)$$

بالتعويض بكل من $\mu = nP$ والطرف الأيسر في المعادلة (6.6) في الطرف الأيسر في المعادلة (6.5) نجد أن :

$$\sum x^2 f(x) = P^2 n(n-1) + np \qquad (6.17)$$

وبالتعويض بالطرف الأيسر في المعادلة (6.17) في الطرف الأيسر في المعادلة (6.14) نجد أن :

$$\sigma^2 = P^2 n(n-1) + P - (nP)^2 = P^2 n^2 - P^2 n + Pn - P^2 n^2$$
$$= Pn(1-P) = nP(1-P)$$

مثال (٦-٨)

إذا رميت قطعة عملة متوازنة 5 مرات متتالية ، وإذا اشرنا لعدد الصور على القطعة في الخمسة مرات بالرمز X.

المطلوب :

١- أوجد دالة الاحتمال للمتغير X ، ووضح ذلك بيانياً.

٢- أوجد التباين للمتغير X ثم أحسب الانحراف المعياري.

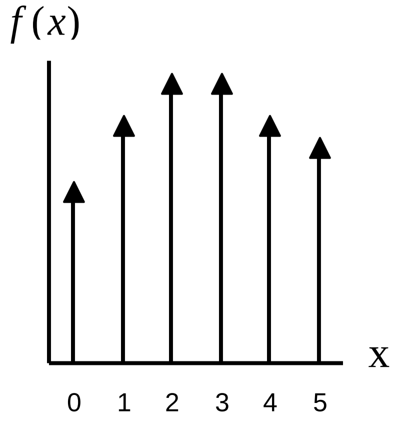
٣- أوجد الاحتمالات عندما X = 0 , 1 , 2 , … , n ، ثم أوجد الاحتمالات التراكمية ووضح ذلك بيانياً.

الحل

١- بما أن احتمال ظهور الصورة في الرمية الواحدة يساوى $\frac{1}{2}$ أي أن $P = \frac{1}{2}$ وكذلك n = 5 فإن :

$$f(x) = C_x^5 (\frac{1}{2})^x (\frac{1}{2})^{5-x} \qquad , \qquad x = 0,1,2,3,4,5$$

وباستخدام ملحق (٣) نجد أن التوزيع الاحتمالي للمتغير X كما هو موضح بالشكل التالي :



شكل (٦-٦)

ملحوظة :

بما أن $P = \frac{1}{2}$ فنجد ان التوزيع متماثل حول القيمة المتوقعة للمتغير X حيث :

$$\mu = nP = 5(\frac{1}{2}) = 2.5$$

٢- $$\sigma^2 = nP(1-P) = 5(\frac{1}{2})(1-\frac{1}{2}) = 5\times\frac{1}{4} = 1.25$$

وبالتالي فإن الانحراف المعياري :

$$\sigma = \sqrt{1.25} = 1.118$$

٣- بما أن دالة التوزيع التراكمية $F(x_j) = p_r(X \le x_j)$ فنجد أن :

$$p_r(x=0) = C_0^5(\frac{1}{2})^0(\frac{1}{2})^5 = \frac{1}{32}$$

$$p_r(x=1) = C_1^5(\frac{1}{2})^1(\frac{1}{2})^4 = \frac{5}{32}$$

$$p_r(x=2) = C_2^5(\frac{1}{2})^2(\frac{1}{2})^3 = \frac{10}{32}$$

$$p_r(x=3) = C_3^5(\frac{1}{2})^3(\frac{1}{2})^2 = \frac{10}{32}$$

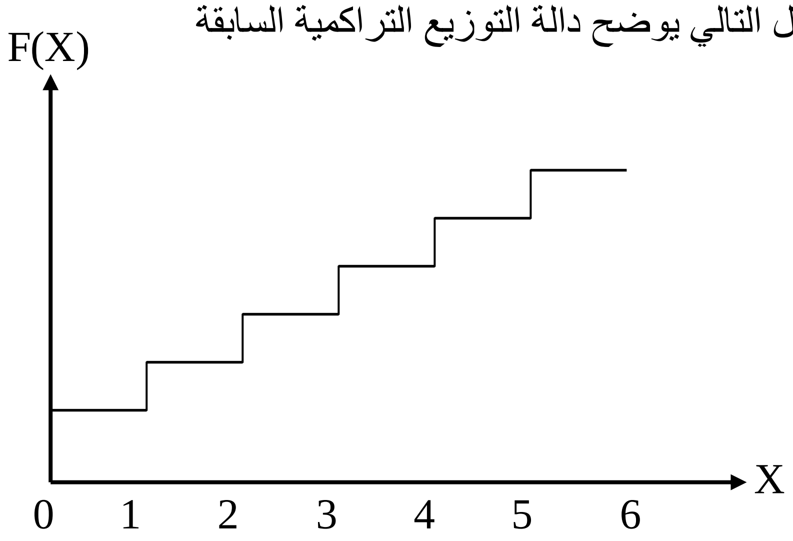
$$p_r(x=4) = C_4^5(\frac{1}{2})^4(\frac{1}{2})^1 = \frac{5}{32}$$

$$p_r(x=5) = C_5^5(\frac{1}{2})^5(\frac{1}{2})^0 = \frac{1}{32}$$

وبالتالي فإن الاحتمالات التراكمية سوف نشير إليها بالرمز F(x) فتصبح على النحو التالي :

$$F(x) = \begin{cases} \frac{1}{32} & x \le 0 \\ \frac{6}{32} & x \le 1 \\ \frac{16}{32} & x \le 2 \\ \frac{26}{32} & x \le 3 \\ \frac{31}{32} & x \le 4 \\ \frac{32}{32} & x \le 5 \end{cases}$$

والشكل التالي يوضح دالة التوزيع التراكمية السابقة



شكل (٦-٧)

مثال (٦-٩)

في إحدى المخابز الخاصة ، تتم عملية مراقبة جودة المنتج من حيث مطابقته للمواصفات عن طريق وزارة التموين فإذا كان احتمال سحب رغيف غير مطابق للمواصفات في هذا المخبز يساوى 0.01 فإذا قامت لجنة الفحص بسحب عينة مكونة من 10 أرغفة عشوائياً من عدد كبير جداً من إنتاج أحد الأيام ، فإذا كان المتغير X يمثل عدد الأرغفة غير المطابقة للمواصفات التي تم سحبها من بين الـ 10 أرغفة المسحوبة.

المطلوب :

١- أوجد التوزيع الاحتمالي للمتغير X.

٢- أوجد احتمال عدم وجود أي رغيف غير مطابق للمواصفات في العينة المسحوبة (أي احتمال أن تكون جميع الأرغفة مطابقة للمواصفات).

٣- احتمال أن تكون جميع الأرغفة المسحوبة غير مطابقة للمواصفات.

٤- أوجد التوقع والتباين للمتغير X ثم أحسب الانحراف المعياري.

الحل

١- بما أن : n = 10 , P = 0.01

$$f(x) = C_x^{10}(0.01)^x(0.99)^{10-x} \qquad , \qquad x = 0,1,2,3,\dots,10$$

٢- احتمال عدم وجود أي رغيف غير مطابق للمواصفات يساوى $f(x=0)$ حيث :

$$f(x=0) = C_0^{10}(0.01)^0(0.99)^{10} = 0.904$$

٣- احتمال أن تكون جميع الأرغفة المسحوبة غير مطابقة للمواصفات يساوى $f(x=10)$ حيث :

$$f(x=10) = C_{10}^{10}(0.01)^{10}(0.99)^{0} = (0.01)^{10}$$

٤- توقع المتغير X هو μ حيث :

$$\mu = \sum_{x=0}^{10} x p_r(x) = nP = 10 \times 0.01 = 0.1$$

والتباين σ^2 حيث :

$$\sigma^2 = \sum_{x=0}^{10}(x-\mu)^2 p_r(x) = nP(1-P)$$
$$= 10 \times 0.01 \times 0.99 = 0.099$$

وبالتالي فإن الانحراف المعياري

$$\sigma = \sqrt{0.099} = 0.315$$

مثال (٦-١٠)

صندوق يحتوى على 6 كرات بيضاء ، 5 كرات حمراء . فإذا سحبت 4 كرات عشوائياً على التوالي من الصندوق حيث تعاد الكرة المسحوبة إلى الصندوق فيل السحبة التالية لها.

المطلوب :

أحسب الاحتمالات التالية :

١- احتمال أن تكون الكرة المسحوبة حمراء.
٢- احتمال أن تكون الكرة المسحوبة بيضاء.
٣- احتمال سحب كرة حمراء وثلاثة كرات بيضاء.
٤- احتمال سحب كرتين حمراء وكرتين بيضاء.

الحل

بما أنه يتم إعادة الكرة قبل السحبة التالية ، بالتالي فإن عدد الكرات في الصندوق لا يتغير قبل كل سحبة وبالتالي احتمال سحب كرة حمراء أو سحب كرة بيضاء لا يتغير من سحبة إلى أخرى . فإذا فرضنا أن P يشير إلى احتمال سحب كرة حمراء وبالتالي فإن :

١- احتمال سحب كرة حمراء يساوى :

$$P = \frac{5}{11}$$

٢- احتمال سحب كرة بيضاء يساوى :

$$(1-P)=\frac{6}{11}$$

وإذا تم سحب أربعة كرات وإذا فرضنا أن المتغير X يشير إلى عدد الكرات الحمراء المسحوبة ، فإن :

٣- احتمال سحب كرة حمراء وثلاثة كرات بيضاء يصبح $p_r(x=1)$ حيث :

$$p_r(x=1)=C_1^4(\frac{5}{11})(\frac{6}{11})^3=4\times(\frac{5}{11})\times(\frac{6}{11})^3=0.295$$

٤- احتمال سحب كرتين حمراء وكرتين بيضاء يساوى $p_r(x=2)$ حيث :

$$p_r(x=2)=C_2^4(\frac{5}{11})^2(\frac{6}{11})^2=6\times(0.0415)=0.249$$

(٦-٤) توزيع الهيبروجومترك

Hyper geometric Distribution

في الفصل السابق تناولنا بالتفصيل توزيع ذات الحدين حيث يوجد لدينا n من المحاولات المتماثلة المستقلة ولكل محاولة نتيجتين متنافيتين فقط هما نجاح باحتمال (P) وفشل باحتمال (1-P) . واحتمال النجاح لا يتغير من محاولة لأخرى ، أي احتمال النجاح ثابت في جميع المحاولات وقيمته تساوي P.

ولكن في كثير من الحالات إذا كان لدينا n من المحاولات المتماثلة Identical Trails ولكل محاولة نتيجتين متنافيتين فقط هما " نجاح وفشل" ، ولكن احتمال النجاح في أي محاولة من الـ n محاولة يعتمد على احتمالات النجاح والفشل في المحاولات السابقة لها ، أي يختلف احتمال النجاح من محاولة لأخرى . وفي هذه الحالة تكون المحاولات غير مستقلة Dependent حيث يعتمد احتمال النجاح أو الفشل في أي محاولة على احتمالات النجاح أو الفشل في المحاولات السابقة لها . وسوف نوضح ذلك من خلال المثال التالي.

مثال (٦-١١)
إذا كان لدينا صندوق يحتوى على 10 كرات منها 4 كرات صفراء ، 6 كرات بيضاء . فإذا سحبت من الصندوق كرتين على التوالي حيث يتم السحب بدون إرجاع Without Replacement . فإذا أشرنا إلى ظهور كرة صفراء في السحبة رقم j ، j = 1 , 2 بالرمز x_j ، وإلى ظهور كرة بيضاء في السحبة رقم j بالرمز y_j .

الحل

صفراء (x)	بيضاء (y)
4	6

فعند سحب الكرة الأولى نجد أن :

$$p_r(x_1) = \frac{4}{10}$$ = احتمال أن تكون الكرة صفراء

$$p_r(y_1) = \frac{6}{10}$$ = احتمال أن تكون الكرة بيضاء

وعند سحب الكرة الثانية نجد أن :

- احتمال أن تكون الكرة صفراء بشرط أن تكون الكرة في السحبة الأولي صفراء يساوى $p_r(x_2 \mid x_1)$ حيث :

$$p_r(x_2 \mid x_1) = \frac{3}{9}$$

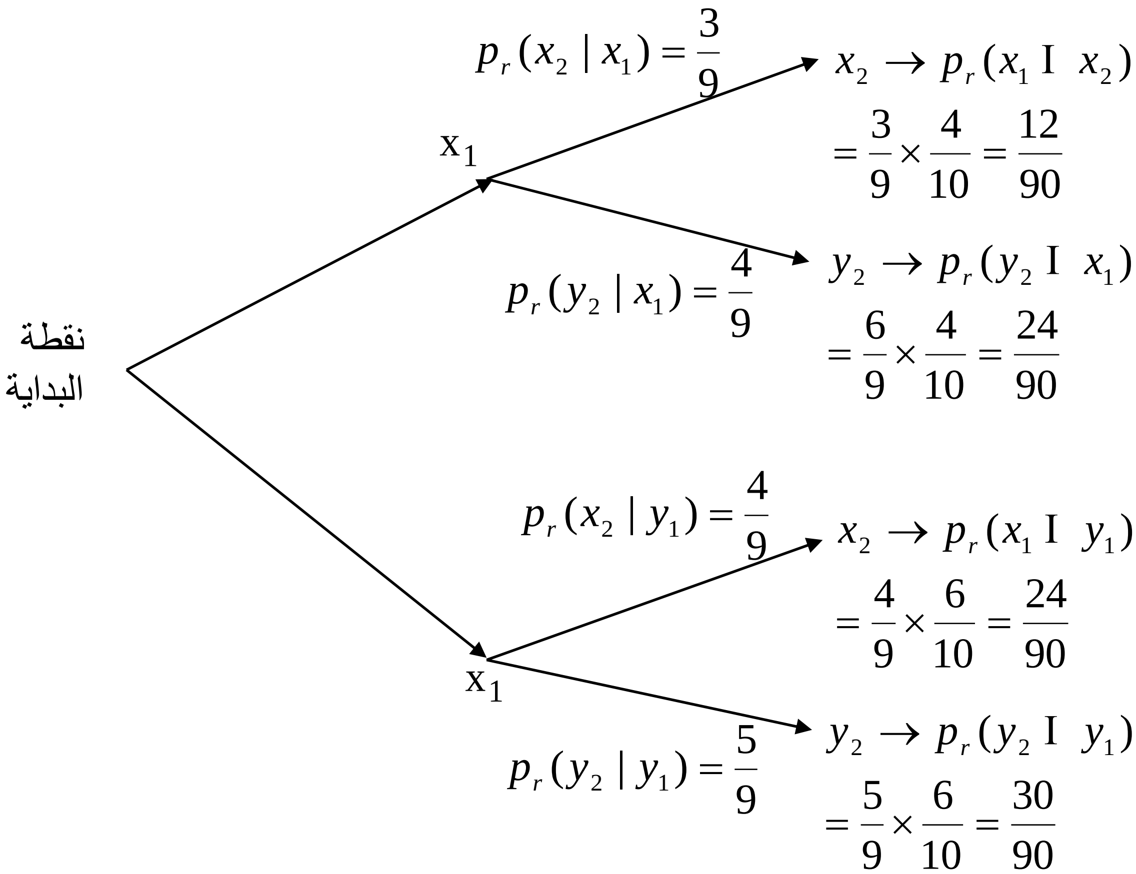
- واحتمال أن تكون الكرة صفراء بشرط أن تكون الكرة في السحبة الأولى بيضاء يساوى $p_r(x_2 \mid y_1)$ حيث :

$$p_r(x_2 \mid y_1) = \frac{4}{9}$$

- كذلك احتمال أن تكون الكرة بيضاء بشرط أن تكون الكرة في السحبة الأولي صفراء يساوى $p_r(y_2 \mid x_1)$ حيث :

$$p_r(y_2 \mid x_1) = \frac{6}{9}$$

وشجرة الاحتمالات التالية توضح ذلك . الاحتمالات الشرطية وغير الشرطية السابقة كذلك الحصول منها على الاحتمالات المشتركة.



شكل (٦-٨)

وبالتالي احتمال أن تكون الكرتين المسحوبتان صفراء يساوى

$$p_r(x_2 \cap x_1) = \frac{12}{90}$$

كذلك احتمال أن تكون الكرتين المسحوبتان بيضاء يساوى

$$p_r(y_2 \cap y_1) = \frac{30}{90}$$

واحتمال أن تكون الكرتين مختلفتين يساوى

$$p_r(y_2 \cap x_1) \cup p_r(x_2 \cap y_1) = \frac{24}{90} + \frac{24}{90} = \frac{48}{90}$$

ملحوظة :

إذا كان السحب مع الإرجاع فإن احتمال ظهور كرة صفراء يساوى $\frac{4}{10}$ ، واحتمال ظهور كرة بيضاء يساوى $\frac{6}{10}$ في كل من السحبة الأولي والثانية (أي الاحتمال ثابت).

مما سبق يمكن القول أنه إذا كان لدينا مجتمع محل الدراسة حجمه n مفردة منها n_1 مفردة لها خاصية معينة (سوف نشير لها بالنجاح مثلاً) ، وبالتالي يكون $(n-n_1)$ مفردة ليس لها هذه الخاصية (سوف نشير لها بالفشل مثلاً) ، وتم سحب j مفردة من هذا المجتمع بحيث يتم السحب بدون إرجاع . فإذا فرضنا أن المتغير العشوائي X يشير إلى عدد المفردات التي لها الخاصية المعينة في الـ j مفردة المسحوبة (أي عدد مرات النجاح) في العينة التي حجمها j فإنه يقال أن المتغير X يتبع التوزيع الهيبروجومترك.

تعريف (٦-٢)

إذا كان المتغير X يتبع التوزيع الهيبروجومترك فإن دالة الاحتمال للمتغير X أي $f(x)$ تأخذ الصيغة التالية :

$$f(x)=\begin{cases}\dfrac{C_x^{n_1}C_{(j-x)}^{(n-n_i)}}{C_j^n} & x=0,1,2,\dots,\ j \\ & j-x\le n-n_1 \quad ,n_1\le n \\ 0 & \text{فيما عدا ذلك}\end{cases} \tag{6.18}$$

ومن تعريف توزيع الهيبروجومترك فإنه يمكن استخدامه في المثال السابق.

فإذا فرضنا أن المتغير X يشير إلى عدد الكرات الصفراء التي يتم سحبها(سحب بدون إرجاع) فإننا نجد أن حجم المجتمع n=10 ، كذلك عدد الكرات الصفراءn_1=4 وبالتالي إذا كان المطلوب :

١- احتمال سحب ثلاثة كرات (سحب بدون إرجاع) ويكونوا اثنين صفراء وواحدة بيضاء.

٢- احتمال سحب ثلاثة كرات (سحب بدون إرجاع) صفراء.

٣- احتمال سحب خمسة كرات (سحب بدون إرجاع) ويكونوا أربعة بيضاء وواحدة صفراء.

الحل

نجد أن المتغير X يتبع التوزيع الهيبروجومترك حيث :

$$n = 10 \quad , \quad n_1 = 4 \quad , \quad \mathrm{n - n_1} = 6$$

وباستخدام الدالة (6.8) نجد أن :

١ - احتمال سحب ثلاثة كرات (بدون إرجاع) اثنين صفراء وواحدة بيضاء يساوى $p_r(x=2)$ حيث :

$$p_r(x=2) = \frac{C_2^4 C_{(3-2)}^{(10-4)}}{C_3^{10}} = \frac{C_2^4 C_1^6}{C_3^{10}}$$

$$= \frac{\dfrac{4!\times 6!}{5!\times 1!\times 2!\times 2!}}{\dfrac{10!}{3!\times 7!}} = \frac{36}{120} = 0.30$$

٢ - احتمال سحب ثلاثة كرات صفراء يساوى $p_r(x=3)$ حيث :

$$p_r(x=3) = \frac{C_3^4 C_{(3-3)}^{(10-4)}}{C_3^{10}} = \frac{C_3^4 C_0^6}{C_3^{10}}$$

$$= \frac{\dfrac{4!}{3!\times 1!}}{\dfrac{10!}{3!\times 7!}} = \frac{4}{120} = \frac{1}{30}$$

٣ - إذا قمنا بسحب 5 كرات فإن n = 5 ، وبالتالي احتمال سحب 5 كرات ويكون أربعة منهم بيضاء وواحدة صفراء يساوى $p_r(x=1)$ حيث :

$$p_r(x=1) = \frac{C_1^4 C_{(5-1)}^{(10-4)}}{C_5^{10}} = \frac{C_1^4 C_4^6}{C_5^{10}} = \frac{4\times 15}{18\times 14} = \frac{5}{12}$$

نظرية (٦-٢)
إذا كان المتغير العشوائي x يتبع التوزيع الهيبروجومترك في مجتمع حجمه n وله دالة الاحتمال $f(x)$ حيث :

$$f(x)=\begin{cases}\dfrac{C_x^{n_1}C_{(j-x)}^{(n-n_i)}}{C_j^n} & x=0,1,2,\ldots,\ j \\ & j-x\le n-n_1 \quad ,n_1\le n \\ 0 & \text{فيما عدا ذلك}\end{cases}$$

فإن توقع وتباين المتغير X هما μ,σ^2 على الترتيب على النحو التالي :

$$\mu=j(\frac{n_1}{n}) \qquad (6.19)$$

$$\sigma^2=j(\frac{n_1}{n})(\frac{n-n_1}{n})(\frac{n-j}{n-1}) \qquad (6.20)$$

الإثبات :
١- بما أن

$$\mu=\sum_{x=0}^{j}xp_r(x)=\frac{\sum\limits_{x=0}^{j}xC_x^{n_1}C_{(j-x)}^{(n-n_1)}}{C_j^n}$$

$$=\frac{j\dfrac{n_1}{n}\sum\limits_{x=1}^{j}C_{(x-1)}^{(n_1-1)}C_{(j-x)}^{(n-n_1)}}{C_{(j-1)}^{(n-1)}}$$

فإذا فرضنا أن y = x-1 فإن :

$$\mu=j\frac{n_1}{n}\sum_{y=0}^{j-1}\frac{C_y^{(n_1-1)}C_{(j-1-y)}^{(n-j-n_1+1)}}{C_{(j-1)}^{(n-1)}} \qquad (6.21)$$

وتمثل الدالة بين القوسين في (6.11) دالة الاحتمال للمتغير العشوائي y الذي يتبع التوزيع الهيبروجومترك أيضاً ، وبالتالي فإن :

$$\sum_{y=0}^{j-1}\frac{C_y^{(n_1-1)}C_{(j-1-y)}^{(n-j-n_1+1)}}{C_{(j-1)}^{(n-1)}}=1 \qquad (6.22)$$

وبالتعويض في (6.11) بـ (6.12) نجد أن :

$$\mu = j\frac{n_1}{n}(1) = j\frac{n_1}{n}$$

٢- كذلك نجد أن :

$$\sigma^2 = \sum_{x=0}^{j}(x-\mu)^2 f(x) = \sum_{x=0}^{j} x^2 f(x) - (\mu)^2 \qquad (6.23)$$

وبما أن :

$$\sum_{x=0}^{j} x^2 f(x) = \sum_{x=0}^{j}[x(x-1) f(x)] + \mu \qquad (6.24)$$

حيث نجد أن :

$$\sum_{x=0}^{j}[x(x-1)f(x)] = \sum_{x=0}^{j} x(x-1)\frac{C_x^{n_1} C_{(j-x)}^{(n-n_1)}}{C_j^{(n-2)}}$$

$$= j(j-1)\frac{n_1(n_1-1)}{n(n-1)}\sum_{x=2}^{j-2}\frac{C_{(x-2)}^{(n_1-2)} C_{(j-x)}^{(n-n_1)}}{C_{(j-2)}^{(n-2)}}$$

$$= j(j-1)\frac{n_1(n_1-1)}{n(n-1)}(1) \qquad (6.25)$$

بالتعويض في (6.23) بـ (6.24) ، (6.25) ، وقيمة μ نجد أن :

$$\sigma^2 = j(j-1)\frac{n_1(n_1-1)}{n(n-1)} + j - \frac{n_1}{n} - [j - \frac{n_1}{n}]^2$$

$$= j\frac{n_1}{n}\times\left[\frac{(j-1)(n_1-1)}{(n_1-1)} + 1 - j\frac{n_1}{n}\right]$$

$$= j\frac{n_1}{n}\times\left[\frac{n(j-1)(n_1-1) + n(n-1) - jn_1(n-1)}{n(n-1)}\right]$$

$$= j\frac{n_1}{n}\times\left[\frac{n(n-n_1) - j(n-n_1)}{n(n-1)}\right] = j\frac{n_1}{n}\times\left[\frac{(n-n_1)-(n-j)}{n(n-1)}\right]$$

ملاحظات :

١- إذا فرضنا أن $P = \frac{n_1}{n}$ حيث P هو احتمال النجاح لمتغير ذات الحدين فنجد أن :
توقع توزيع ذات الحدين يساوى توقع توزيع الهيبروجومترك كذلك يصبح في هذه الحالة تباين المتغير التابع لتوزيع الهيبروجومترك يساوى تباين المتغير التابع لتوزيع ذات الحدين مضروب في المقدار الثابت $\frac{n-j}{n-1}$.

٢- عندما يكون حجم المجتمع n كبير جداً فإن المقدار $\frac{n-j}{n-1}$ يؤول إلى الواحد بمعنى أن $\lim_{n\to\infty}\frac{n-j}{n-1} = 1$

وفي هذه الحالة فإن $P = \frac{n_1}{n}$ ونجد أن التوقع والتباين للمتغير الذي يتبع التوزيع الهيبروجومترك يساوى التوقع والتباين للمتغير الذي يتبع توزيع ذات الحدين.

٣- من (١) ، (٢) يمكن استخدام توزيع ذات الحدين كتقريب جيد لتوزيع الهيبروجومترك ، حيث يمكن رياضياً إثبات أن دالة الاحتمال للتوزيع الهيبروجومترك تؤول إلى دالة* ذات الحدين عندما $n \to \infty$ ، فإن $P = \frac{n_1}{n}$

## (٦-٥) توزيع بواسون		Poisson Distribution

في القرن التاسع عشر الميلادي عرف عالم الرياضيات سيمون بواسون** (١٧١٨-١٨٤٠) متغير عشوائي سمي بمتغير بواسون نسبة إلى العالم سيمون بواسون حيث اعتبر بواسون وجود عدد غير محدود Infinite من المحاولات المستقلة ولكل محاولة نتيجتين متنافيتين فقط سوف نشير لها نجاح وفشل باحتمالي P , (1-P) على الترتيب ، وعرف متغير بواسون X بأنه عدد مرات النجاح في عدد غي محدود من المحاولات.

---

* Mood , Graybill , and Boes (1974) : " Introduction to The Theory of Statistics ", MeGraw- Hill International Book Company , London.

** نفس المرجع السابق صفحة (٢١٥)

ونجد أن متغير بواسون موجود بالفعل في كثير من المشاكل الفعلية التي يقاس فيها المتغير في فترة زمنية متصلة Continuous Time Interval معينة أو دراسة المتغير في فراغ معاينة متصل Continuous Sampling Space ، فعلى سبيل المثال :

١- في دراسة عدد الميكروبات والجراثيم من أنواع معينة في لتر ماء من إحدى الترع بجمهورية مصر العربية ، فنجد أن عدد الميكروبات والجراثيم في لتر واحد (فراغ معاينة متصل Continuous Sampling Space) وليكن X يمثل متغير عشوائي يأخذ القيم $0 , 1 , 2 , 3 , \ldots , \infty$.

٢- في دراسة الكميات المطلوبة والمعروضة من سلعة ما (أو خدمة ما) في إحدى الأسواق المصرية في شهر يوليو ، فنجد أن عدد العملاء الذين يطلبون هذه السلعة (أو الخدمة) في هذا الشهر يمثل متغير بواسون حيث أن عدد هؤلاء العملاء عدد غير محدود (قد يكون بعضهم مصريين ، وبعضهم عرب ، وبعضهم أجانب) وبالتالي فإن $X = 0,1,2,3,\ldots,\infty$.

ويوجد العديد من الأمثلة الأخرى التي يكون فيها كل من عدد المحاولات n وعدد مرات النجاح X ممكن أن يكون عدد غير محدود ، ويعتبر هذا هو الفرق بين المتغير الذي يتبع توزيع ذات الحدين والمتغير الذي يتبع توزيع بواسون.

واشتق العالم سيمون بواسون دالة الاحتمالات لمتغير بواسون كذلك قدم أهم خصائص التوزيع . وفيما يلي سوف نقدم دالة الاحتمال وأهم خصائص التوزيع.

تعريف (٦-٣)

إذا كان المتغير العشوائي X يتبع توزيع بواسون ، فإن دالة الاحتمال للمتغير X أي $f(x)$ تأخذ الصيغة التالية :

$$f(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & x = 0,1,2,3,\ldots,\infty \\ 0 & \text{فيما عدا ذلك} \end{cases} \tag{6.26}$$

حيث λ, e مقادير ثابتة ، e هي الأساس الطبيعي حيث $e \cong 2.7828$ ، λ هي عبارة عن عدد مرات النجاح المتوقعة في وحدة القياس . كما سوف يتضح في الأمثلة التالية :

وكما في حالة توزيع ذات الحدين فإنه توجد جداول توزيع بواسون ، ففي سنة ١٩٥٣ وباستخدام الحاسب الآلي تم حساب قيمة $f(x)$ عند القيم المختلفة لكل من λ، X في جداول تيسر لغير المتخصصين كذلك توفر الوقت والجهد للمتخصصين لاستخدام توزيع بواسون في الدراسات النظرية والتطبيقية ، وفى ملحق رقم (٤) نقدم جزء من هذه الجداول.

مثال (٦-١٢)

قام أحد البنوك المصرية بفتح فروع له في إحدى ضواحي القاهرة ، فإذا كان البنك يرغب في تحديد عدد الشبابيك التي يجب فتحها (وبالتالي عدد الموظفين العاملين على هذه الشبابيك) لخدمة عملاء هذا الفرع ، وبدراسة المنطقة التي يقع فيها الفرع وجد أن العدد المتوقع للعملاء على هذا الفرع في يوم العمل (7 ساعات) 21 عميل.

ويرغب البنك في تحديد التوزيع الاحتمالي لعدد العملاء الذين يصلون إلى هذا الفرع في ساعة العمل.

الحل

إذا فرضنا أن المتغير X يمثل عدد العملاء الذين يصلون إلى البنك في أحد أيام العمل الأسبوعية خلال ساعة العمل ، كذلك عدد العملاء المتوقع وصولهم خلال ساعة عمل يساوى λ حيث $\lambda =$ (21/7)3 عملاء في الساعة.

وبالتعويض في الدالة (6.16) فإن التوزيع الاحتمالي لعدد العملاء الذين يصلون إلى فرع البنك خلال ساعة عمل على النحو التالي :

$$f(x) = \frac{e^{-3}(3)^x}{x!} \qquad x = 0,1,2,3,\ldots,\infty$$

وبالتعويض عن قيم X المختلفة في الدالة $f(x)$ نحصل على قيم $f(x)$ كما هو موضح بالجدول التالي :

جدول (٦-٧)

عدد العملاء خلال ساعة العمل	$e^{-\lambda}$	$\lambda^x = 3^x$	X!	$f(x)$
0	$e^{-6} = 0.0498$	1	1	0.0498
1		3	1	0.1494
2		9	2	0.2241
3		27	6	0.2241
4		81	24	0.1681
5		243	120	0.1009
6		729	720	0.0504
7		2187	5040	0.0216
8		6561	40320	0.0081
9		19683	362880	0.0027
10		59049	3628800	0.0008
11		177147	39916800	0.0002
12		531441	479001600	0.0001
13		1594323	6227020800	0.0000

ومن الجدول يتضح أن أقصي عدد من العملاء يمكن أن يصل إلى فرع البنك خلال ساعة عمل لا يزيد عن 12 عميل ، وفي ضوء هذا التوزيع الاحتمالي يستطيع البنك تحديد العدد الأمثل* من الشبابيك التي يجب فتحها بالفرع.

نظرية (٦-٣)
إذا كان X متغير عشوائي يتبع توزيع بواسون فإن كل من التوقع والتباين للمتغير X يساوى المعلمة λ أي :

$$\mu=\lambda \qquad , \qquad \sigma^2=\lambda \tag{6.27}$$

الإثبات :
بما أن :

$$\begin{aligned}\mu &= \sum_{x=0}^{\infty} x f(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\ &= e^{-\lambda}\lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\ &= e^{-\lambda}\lambda \left[1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+.....\right] \\ &= e^{-\lambda}\lambda e^{\lambda} = \lambda e^0 = \lambda(1) = \lambda\end{aligned}$$

وبما أن :

$$\sigma^2 = \sum_{x=0}^{\infty} x^2 f(x) - \mu^2 = \sum_{x=0}^{\infty} x^2 f(x) - \lambda^2 \tag{6.28}$$

وبما أن :

$$\sum_{x=0}^{\infty} x^2 f(x) = \sum_{x=0}^{\infty} \left[x(x-1) f(x) + \mu\right] \tag{6.29}$$

* أ.د. عفاف الدش (١٩٨٧) : " بحوث العمليات واتخاذ القرارات" ، مكتبة عين شمس – القاهرة - رقم الإيداع ١٩٨٧/٨٨٨٧ – الترقيم الدولي ٠- ٠١١٣- ٠٧- ٩٧٧.

وحيث أن :

$$\sum_{x=0}^{\infty}\left[x(x-1)\,f(x)\right]=\sum_{x=0}^{\infty}x(x-1)\,\frac{e^{-\lambda}\lambda^{x}}{x(x-1)(x-2)!}$$

$$=\lambda^{2}e^{-\lambda}\sum_{x=2}^{\infty}x(x-1)\,\frac{\lambda}{x(x-1)(x-2)!}$$

$$=\lambda^{2}e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda}{(x-2)!}=\lambda^{2}e^{-\lambda}\left[1+\frac{\lambda}{1!}+\frac{\lambda^{2}}{2!}+\ldots\ldots\right]$$

$$=\lambda^{2}e^{-\lambda}e^{\lambda}=\lambda^{2} \qquad (6.30)$$

وبالتعويض في (6.29) بـ (6.30) نجد أن :

$$\sum_{x=0}^{\infty}x^{2}f(x)=\lambda^{2}+\lambda \qquad (6.31)$$

وبالتعويض في (6.28) ، (6.31) نجد أن :

$$\sigma^{2}=\lambda^{2}+\lambda-\lambda^{2}=\lambda$$

مثال (٦-١٣)

١- أرسم شكل توزيع بواسون عندما :

$\lambda = 0.4$ $\lambda = 0.5$ $\lambda = 1$ $\lambda = 2$

$\lambda = 3$ $\lambda = 4$ $\lambda = 5$ $\lambda = 6$ $\lambda = 7$

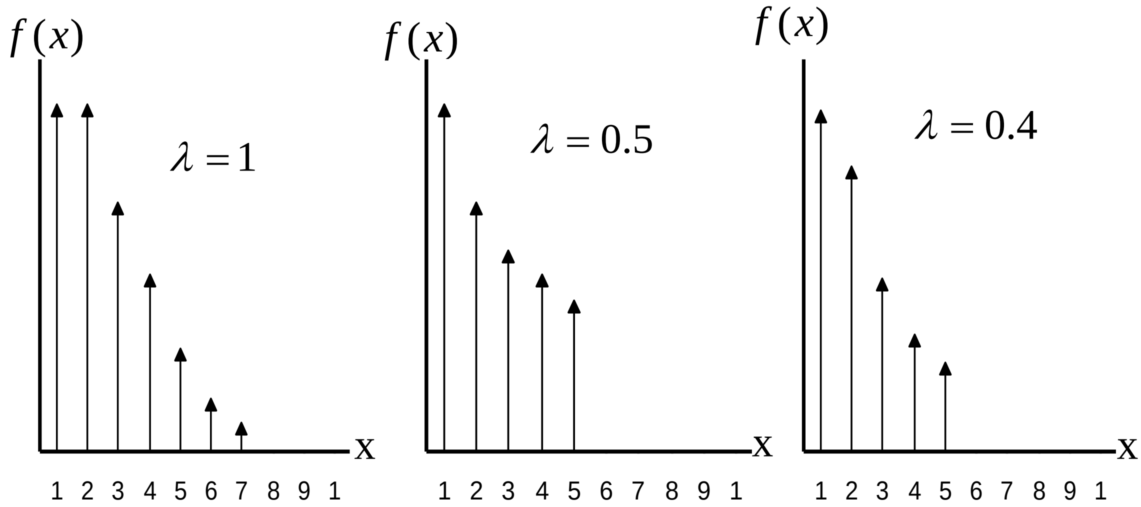
٢- أستنتج من الرسم أهم خصائص توزيع بواسون.

الحل

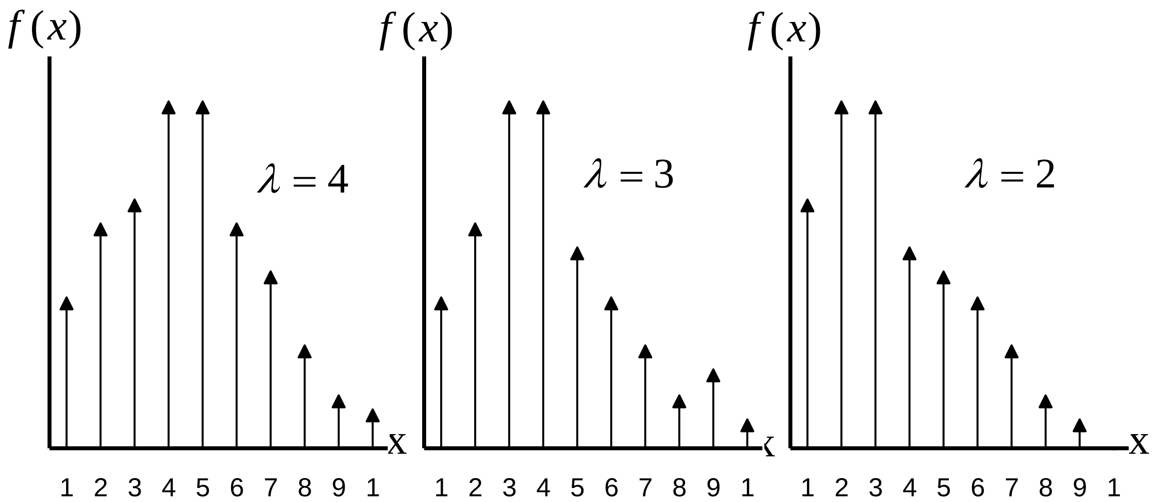
باستخدام ملحق رقم (٤) نجد أن التوزيعات الاحتمالية عندما $\lambda = 0.4 , 0.5 , 1 , 2 , 3 , 4 , 5 , 6 , 7$. كما هو موضح بجدول (٦-٨) ، وأشكال (٦-٩) – (٦-١١) توضح هذه التوزيعات.

جدول (٦-٨) : يعطي قيم دالة الاحتمال لتوزيع بواسون عند القيمة المعطاة للمعلمة λ

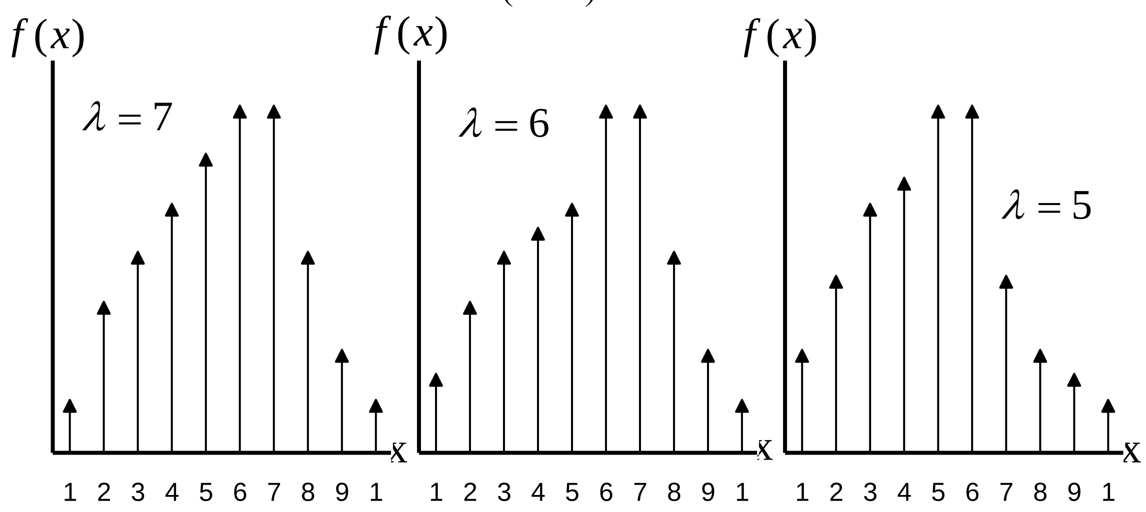
X	λ =0.4	λ =0.5	λ =1	λ =2	λ =3	λ =4	λ =5	λ =6	λ =7
0	0.6703	0.6065	0.3679	0.1353	0.0498	0.0183	0.0067	0.0025	0.0009
1	0.2681	0.3033	0.3679	0.2707	0.1494	0.0733	0.0337	0.0149	0.0064
2	0.0536	0.0758	0.1839	0.2707	0.2240	0.1465	0.0842	0.0446	0.0223
3	0.0072	0.0126	0.0613	0.1804	0.2240	0.1954	0.1404	0.0892	0.0521
4	0.0007	0.0016	0.0153	0.0902	0.1680	0.1954	0.1755	0.1339	0.0912
5	0.0001	0.0002	0.0031	0.0361	0.1008	0.1563	0.1755	0.1606	0.1277
6	0.0000	0.0000	0.0005	0.0120	0.0504	0.1042	0.1462	0.1606	0.1490
7			0.0001	0.0034	0.0216	0.0595	0.0653	0.1377	0.1490
8			0.0000	0.0009	0.0081	0.0298	0.0363	0.1033	0.1304
9				0.0002	0.0027	0.0132	0.0181	0.0688	0.1014
10				0.0000	0.0008	0.0053	0.0082	0.0413	0.0710
11					0.0002	0.0019	0.0034	0.0225	0.0452
12					0.0001	0.0006	0.0013	0.0113	0.0264
13					0.0000	0.0002	0.0005	0.0052	0.0142
14						0.0001	0.0002	0.002	0.0071
15						0.0000	0.0000	0.0009	0.0033
16								0.0003	0.0014
17								0.0001	0.0006
18								0.0000	0.0002
19									0.0001
20									0.0000



شكل (٦-٩)



شكل (٦-١٠)



شكل (٦-١١)

من الأشكال السابقة يمكن استنتاج ما يلي :
من الشكل (٦-٩) نجد أن :
١- يكون التوزيع ملتوي ناحية اليمين (أي التواء موجب) عندما $1 \le \lambda$
٢- عندما $1 \le \lambda$ تكون النهاية عظمي[1] للدالة $f(x)$ عندما $X = 0$ أو بعبارة أخرى تكون القيمة المنوالية للمتغير X تساوى صفر عندما $1 \le \lambda$.
٣- تتناقص قيمة النهاية العظمي للدالة $f(x)$ عندما $1 \to \lambda$.

ومن الشكلين (٦-١٠) ، (٦-١١) نجد أن :
١- عندما $1 \ge \lambda$ ، النهاية العظمي للدالة $f(x)$ تتناقص بزيادة قيمة λ
٢- عندما $1 < \lambda$ تكون النهاية العظمي للدالة $f(x)$ عندما $x = \lambda$ أي أنه عندما $1 < \lambda$ فإن القيمة المتوقعة للمتغير X تساوى القيمة المنوالية تساوى λ.
٣- يقترب التوزيع من التماثل حول القيمة المتوقعة (المنوالية أيضاً) كلما زادت قيمة λ.

والنتائج السابقة نتائج عامة ويمكن إثباتها رياضياً*.

ويعتبر توزيع بواسون تقريب جيد لتوزيع ذات الحدين عندما تكون عدد المحاولات كبير أي عندما $n \to \infty$ ، وتكون قيمة احتمال النجاح P قيمة صغيرة جداً (أي عندما $P \to 0$) ويمكن إثبات ذلك رياضياً . وعندما يستخدم توزيع بواسون كتقريب لتوزيع ذات الحدين فإن $\lambda = nP$. وبالرجوع إلى ملحق (٣) عندما تأخذ n قيم كبيرة وP قيم صغيرة جداً فنجد قيمة دالة الاحتمال $f(x)$ عند هذه القيم تساوى تقريباً قيم $f(x)$ بالنسبة لتوزيع بواسون بملحق (٤) عندما $\lambda = nP$.

مثال (٦-١٤)
إذا كان احتمال أن تنجب سيدة مولود ذكر في إحدى القرى يساوى 0.1 ، وبالتالي احتمال أن تنجب أنثي يساوى 0.9 ، فإذا وجد 20 سيدة بهذه القرية حوامل . وبافتراض أن المتغير X يشير إلى عدد المواليد الذكور للسيدات الحوامل بهذه القرية.

المطلوب :
١- أوجد التوزيع الاحتمالي للمتغير X ووضح ذلك بيانياً.
٢- أوجد التوزيع الاحتمالي التقريبي للمتغير X ووضح ذلك بيانياً.

---
[1] نفس المرجع المذكور صفحة (٢٣١)
* نفس المرجع المذكور صفحة (٢١٥)

الحل

١- إذا فرضنا أن دالة الاحتمال للمتغير X هي $f(x)$ ، كذلك P قيمة احتمال إنجاب طفل ذكر حيث P = 0.1 وبالتالي فإن :

$$f(x) = C_x^{20}(0.1)^x(0.9)^{20-x} \qquad x = 0,1,2,3,\ldots,20$$

٢- التوزيع الاحتمالي التقريبي للمتغير X هو توزيع بواسون حيث $\lambda = nP$ ، أى

$\lambda = 20\ (0.1) = 2$

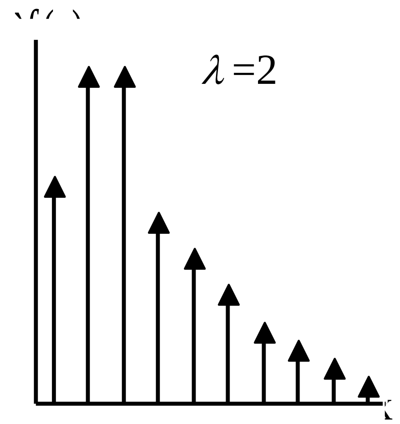
فإذا فرضنا أن $`f(x)$ هى دالة الاحتمال التقريبية للمتغير X فإن :

$$`f(x) = \frac{e^{-2}(2)^x}{x!} \qquad x = 0,1,2,\ldots..$$

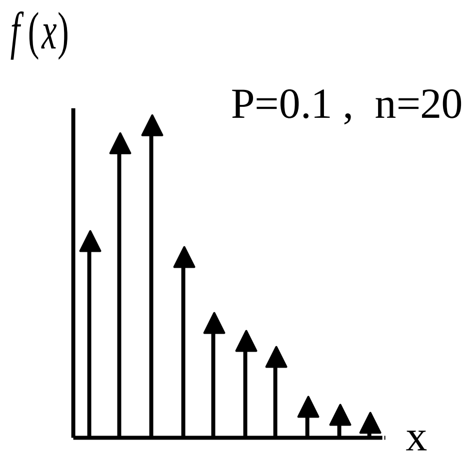
ومن ملحقي (٣) ، (٤) نكون الجدول التالي الذي يوضح قيم $f(x)$ ، $`f(x)$ عند قيم X الممكنة.

جدول (٦-٩)

X	$f(x)$	$`f(x)$
0	0.1216	0.1353
1	0.2702	0.2707
2	0.2852	0.2707
3	0.1901	0.1804
4	0.0898	0.0902
5	0.0319	0.0361
6	0.0089	0.0120
7	0.0020	0.0034
8	0.0004	0.0009
9	0.0001	0.0002
10	0.0000	0.0000
المجموع	1.000	1.000



دالة الاحتمال التقريبي $\grave{f}(x)$
شكل (٦-١٣)



دالة الاحتمال الصحيحة $f(x)$
شكل (٦-١٢)

ملحوظة :

من جدول (٦-٤) يتضح أن احتمال إنجاب عدد من الأطفال الذكور أكثر من 9 أطفال ذكور يساوى $f(x>9)=0$ أيضاً ، كذلك الاحتمال التقريبي أيضاً لإنجاب أكثر من 9 أطفال ذكور يساوى $\grave{f}(x>9)=0$. ويتضح من جدول (٦-٤) كذلك من الشكلين (٦-١٢) ، (٦-١٣) مدى التقارب بين التوزيع الاحتمالي للمتغير X (أي توزيع ذات الحدين) والتوزيع التقريبي للمتغير X (أي توزيع بواسون) ويرجع ذلك إلى أن عدد المحاولات كبير نسبياً (حيث n = 20) كذلك احتمال إنجاب طفل ذكر صغير نسبياً حيث (p = 0.1).

ومن دراستنا في هذا الفصل والفصول السابقة بهذا الباب يتضح أنه توجد بعض أوجه الشبه بين توزيع ذات الحدين ، وتوزيع الهيبروجومترك ، وتوزيع بواسون كذلك توجد بعض أوجه الاختلاف بين هذه التوزيعات الثلاثة المذكورة . لذلك سوف نقدم في الجدول التالي مقارنة مختصرة بين التوزيعات الثلاثة.

جدول (٦-١٠)

التوزيع / وجه المقارنة	ذات الحدين	الهيبروجومترك	بواسون
المحاولات (n)	- متماثلة ومستقلة ولكل محاولة نتيجتين متنافيتين نجاح أو فشل . احتمال النجاح والفشل ثابت لا يتغير من محاولة لأخرى.	- متماثلة وغير مستقلة ولكل محاولة نتيجتين متنافيتين نجاح أو فشل . احتمال النجاح والفشل يتغير من محاولة لأخرى.	- متماثلة ومستقلة ولكل محاولة نتيجتين متنافيتين نجاح أو فشل . احتمال النجاح والفشل ثابت لا يتغير من محاولة لأخرى.
حجم المجتمع محل الدراسة (n)	- لا يتطلب معرفة حجم المجتمع الذي يتم إجراء المحاولات على مفرداته. - قد يصل حجم المجتمع إلى عدد غير محدود Infinite	- يتطلب معرفة حجم المجتمع الذي يتم إجراء المحاولات على مفرداته. - يجب أن يكون هذا الحجم عدد محدود - كذلك يعتمد التوزيع على معرفة عدد مرات النجاح في n ولتكن n_1 وبالتالي معرفة عدد مرات الفشل $(n-n_1)$	- قد يصل حجم المجتمع إلى عدد غير محدود Infinite
المتغير (x)	- متغير متقطع يأخذ قيم صحيحة غير سالبة بحيث X = 0,1,2,…,n	- متغير متقطع يأخذ قيم صحيحة غير سالبة بحيث X = 0,1,2,…,n	- متغير متقطع يأخذ قيم صحيحة غير سالبة بحيث $X= 0,1,2,\ldots,\infty$
دالة الاحتمال $f(x)$	$f(x) = C_x^n P^x (1-p)^{n-x}$	$f(x) = \dfrac{C_x^{n_1} C_{j-x}^{n-n_1}}{C_j^n}$	$f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$
التوقع (μ)	nP	$j\dfrac{n_1}{n}$	λ

التوزيع / وجه المقارنة	ذات الحدين	الهيبروجومترك	بواسون
التباين (σ^2)	$nP(1-p)$	$j\frac{n_1}{n}\left[\frac{(n-n_1)(n-j)}{n(n-1)}\right]$	λ
التقريب	عندما يكون عدد المحاولات كبير واحتمال النجاح صغير يستخدم توزيع بواسون كتقريب جيد لتوزيع ذات الحدين	عندما يكون حجم المجتمع n كبير يستخدم توزيع ذات الحدين كتقريب جيد للتوزيع الهيبروجومترك	———

(٦-٦) أمثلة تطبيقية Applied Examples

تطبيق (٦-١)

في إحدى الدراسات عن الحوادث بإحدى الطرق الصحراوية بجمهورية مصر العربية وجد أن احتمال وقوع حادث بهذا الطريق ويكون السائق سيدة يساوى 0.15 واحتمال وقوع حادث بالطريق ويكون السائق رجل يساوى 0.85 فإذا تم تسجيل 5 حوادث بهذا الطريق خلال أحد الأشهر.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد الحوادث التي يكون السائق فيها سيدة ثم أوجد العدد المتوقع لهذه الحوادث.

٢- أوجد التوزيع الاحتمالي لعدد الحوادث التي يكون السائق فيها رجل ثم أوجد العدد المتوقع لهذه الحوادث.

٣- أحسب احتمال وقوع حادثين يكونا فيهما السائق رجل.

٤- أحسب احتمال وقوع جميع الحوادث ويكون السائق رجل.

٥- أحسب احتمال وقوع جميع الحوادث ويكون السائق سيدة.

الحل

إذا فرضنا أن X تشير إلى عدد الحوادث التي يكون فيها السائق رجل ، y تشير إلى عدد الحوادث التي يكون فيها السائق سيدة.

١-

$$f(x) = C_x^5 (0.85)^x (0.15)^{5-x} \qquad \text{x} = 0,1,2,3,4,5$$

٢-

$$f(y) = C_y^5 (0.15)^y (0.85)^{5-y} \qquad \text{y} = 0,1,2,3,4,5$$

$$\mu_{\text{x}} = 5(0.85) = 4.25 \approx 4 \text{ حوادث}$$

$$\mu_{\text{غ}} = 5(0.15) = 0.75 \approx 1 \text{ حادث}$$

٣- احتمال وقوع حادثين يكونا فيهما السائق رجل يساوى $f(x=2)$ أو يساوى $f(x=3)$ حيث :

$$f(x=2) = C_2^5 (0.85)^2 (0.15)^3 = 10(0.7225)(0.0034) = 0.025$$

أو

$$f(x=3) = C_3^5 (0.15)^3 (0.85)^2 = 10(0.0034)(0.7225) = 0.025$$

٤- احتمال وقوع جميع الحوادث ويكون السائق رجل يساوى $f(x=5)$ أو يساوى $f(y=0)$ حيث :

$$f(x=5) = C_5^5 (0.85)^5 (0.15)^0 = 0.445$$

أو

$$f(y=0) = C_0^5 (0.15)^0 (0.85)^5 = 0.445$$

٥- احتمال وقوع جميع الحوادث ويكون السائق فيها سيدة يساوى $f(y=5)$ أو يساوى $f(x=0)$ حيث :

$$f(y=5) = C_5^5 (0.15)^5 (0.85)^0 = 0.00008$$

أو

$$f(x=0) = C_0^5 (0.85)^0 (0.15)^5 = 0.00008$$

تطبيق (٦-٢)

في إحدى الشركات الإنتاجية أظهرت دراسة جودة المنتج للشركة أن 15% من الإنتاج اليومي النهائي تكون الوحدات المنتجة غير مطابقة للمواصفات . فإذا أخذت عينة مكونة من 4 وحدات من الإنتاج اليومي للشركة.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد الوحدات غير المطابقة.
٢- أوجد احتمال وجود وحدة واحدة على الأكثر غير مطابقة.
٣- أوجد احتمال أن تكون جميع الوحدات غير مطابقة.
٤- أوجد العدد المتوقع والانحراف المعياري للوحدات غير المطابقة للمواصفات.

الحل

١- بفرض أن المتغير X يشير إلى عدد الوحدات غير المطابقة للمواصفات، $f(x)$ هي دالة الاحتمال للمتغير X .

من المشكلة يتضح أن المتغير X متغير يتبع توزيع ذات الحدين حيث عدد المحاولات n = 4 واحتمال وجود وحدة غير مطابقة P = 0.15 وبالتالي فإن :

$$f(x) = C_x^4 (0.15)^x (0.85)^{4-x} \qquad x = 0,1,2,3,4$$

٢- احتمال وجود وحدة واحدة على الأكثر غير مطابقة يساوى $f(x \le 1)$ حيث :

$$\begin{aligned} f(x \le 1) &= f(x=0) + f(x=1) \\ &= C_0^4 (0.15)^0 (0.85)^4 + C_1^4 (0.15)\ (0.85)^3 \\ &= 0.552 + 0.368 = 0.89 \end{aligned}$$

٣- احتمال أن تكون جميع الوحدات غير مطابقة يساوى $f(x=4)$ حيث :

$$f(x=4)=C_4^4(0.15)^4(0.85)^0=(1)(0.0005)(1)=0.0005$$

٤-

$$\mu_x=nP=4(0.15)=0.6\approx 1 \text{ وحدة}$$

$$\sigma_x^2=nP(1-P)=4(0.15)(0.85)=0.51\approx 1 \text{ وحدة}$$

$$\therefore \sigma_\mathrm{x}=\sqrt{\sigma_x^2}=\sqrt{1}=1 \text{ وحدة}$$

تطبيق (٦-٣)

أعلنت إحدى المؤسسات الاستثمارية عن وجود 3 وظائف خالية بالمؤسسة فتقدم للمؤسسة 15 فرد تنطبق عليهم شروط التعيين منهم عشرة رجال وخمسة سيدات.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد السيدات التي يمكن أن تشغل هذه الوظيفة.
٢- أوجد احتمال أن تشغل الوظائف بسيدتين ورجل.
٣- أوجد احتمال أن تشغل هذه الوظائف بالرجال فقط.
٤- أوجد احتمال أن تشغل هذه الوظائف بالسيدات فقط.

الحل

إذا فرضنا أن المتغير X يمثل عدد الوظائف التي يتم شغلها بالسيدات ، فنجد أن المتغير X يتبع توزيع الهيبروجومترك (أنظر الفصل (٦-٤)) بحيث :

$$n=15 \quad , \quad \mathrm{n}_1=5 \quad , \quad \mathrm{j}=3 \quad , \quad \mathrm{x}=0,1,2,3$$

فإن :

١-

$$f(x)=\frac{C_x^{n_1}C_{j-x}^{n-n_1}}{C_j^n} \qquad , \quad \mathrm{x}=0,1,2,3$$

$$=\frac{C_x^5 C_{3-x}^{15-5}}{C_3^{15}} \qquad , \quad \mathrm{x}=0,1,2,3$$

وبالتعويض بقيم X المختلفة في الدالة $f(x)$ نجد أن :

٢- احتمال أن تشغل الوظائف بسيدتين ورجل يساوى $f(x=2)$ حيث :

$$f(x=2)=\frac{C_2^5 C_1^{10}}{C_3^{15}}=\frac{(5!/2!3!)(10!/1!9!)}{5!/2!3!}$$

$$=\frac{10\times 10}{130}=\frac{10}{13}=0.769$$

٣- احتمال أن تشغل جميع الوظائف بالرجال فقط يساوى $f(x=0)$ حيث :

$$f(x=0)=\frac{C_0^5 C_0^{10}}{C_5^{15}}=\frac{(5!/2!3!)(10!/1!9!)}{5!/2!3!}$$

$$=\frac{10\times 10}{130}=\frac{10}{13}=0.769$$

٤- احتمال أن تشغل جميع الوظائف بالسيدات فقط يساوى $f(x=3)$ حيث :

$$f(x=3)=\frac{C_3^5 C_0^{10}}{C_3^{15}}=\frac{10\times 1}{130}=\frac{1}{13}=0.077$$

تطبيق (٦-٤)

دلت الأبحاث الطبية أن احتمال إصابة الفرد بمرض الإيدز بالدول الإسلامية يساوى 0.0001 ، فإذا أخذت عينة مكونة من 10000 فرد من إحدى المدن الإسلامية.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد المصابين بالمرض في هذه العينة.
٢- أحسب احتمال عدم وجود أي فرد مصاب.
٣- أحسب العدد المتوقع والانحراف المعياري لعدد المصابين.

الحل

إذا فرضنا أن X يمثل عدد المصابين بالإيدز بهذه العينة ، فنجد أن المتغير X يتبع توزيع بواسون حيث :

$$\lambda=nP=10000\times 0.0001=1$$

وبالتالي فإن :

$$f(x)=\frac{e^{-\lambda}\lambda^x}{x!}=\frac{e^{-1}(1)^x}{x!}\qquad ,\ \ x=0,1,2,.....$$

٢- احتمال عدم وجود أي فرد مصاب بهذه المدينة يساوى $f(x=0)$ حيث :

$$f(x=0)=\frac{e^{-1}(1)^0}{1!}=e^{-1}=\frac{1}{e}=0.38$$

٣- احتمال وجود فردين مصابين بالمرض على الأكثر يساوى $f(x \le 2)$ حيث :

$$f(x \le 2)= f(x=0)+ f(x=1)+ f(x=2)$$
$$=e^{-1}+\frac{e^{-1}(1)}{1!}+\frac{e^{-1}(1)^2}{2!}=e^{-1}(2+0.5)$$
$$=2.5e^{-1}=2.5(0.368)=0.92$$

٤- العدد المتوقع للمصابين يساوى μ حيث :

$\mu=\lambda=1$ مصاب

وبالتالي فإن الانحراف المعياري لعدد المصابين σ حيث :

$\sigma=\sqrt{\lambda}=\sqrt{1}=1$ مصاب

تطبيق (٦-٥)

في إحدى الدراسات عن سلوك النزلاء بإحدى مؤسسات الأحداث بالقاهرة بعد ترك المؤسسة وجد أن 35% من النزلاء يعودون للمؤسسة مرة ثانية بعد انتهاء مدتهم . فإذا أخذت عينة مكونة من 10 نزلاء من هذه المؤسسة.

المطلوب :

١- أوجد احتمال وجود نزيلين أو أكثر يمكن عودتهم للمؤسسة مرة ثانية.
٢- أحسب احتمال عدم عودة أي نزيل من الـ 10 نزلاء محل الدراسة.
٣- أحسب العدد المتوقع للنزلاء الممكن عودتهم مرة ثانية.

الحل

إذا فرضنا أن X هو عدد النزلاء الذين يعودون إلى المؤسسة مرة ثانية بعد انتهاء مدتهم بها . فنجد أن X متغير عشوائي يتبع توزيع ذات الحدين حيث :

١-

$$f(x)=C_x^{10}(0.35)^x(0.65)^{10-x} \qquad \text{x} = 0,1,2,...,10$$

وبما أن احتمال وجود نزيلين أو أكثر يمكن عودتهم للمؤسسة مرة ثانية يساوي $f(x \ge 2)$ حيث :

$$f(x \geq 2) = 1 - \left[f(x=0) + f(x=1) \right]$$
$$= 1 - \left[C_0^{10} (0.35)^0 (0.65)^{10} + C_1^{10} (0.35)\ (0.65)^9 \right]$$
$$= 1 - \left[(1)(1)(0.014) + (10)(0.35)(0.021) \right]$$
$$= 1 - \left[0.014 + 0.0735 \right] = 1 - \left[0.0875 \right] = 0.913$$

٢- احتمال عدم عودة أي نزيل من الـ 10 نزلاء محل الدراسة يساوى $f(x=0)$ حيث :

$$f(x=0) = C_0^{10} (0.35)^0 (0.65)^{10} = 0.014$$

٣-

$$\mu_x = \sum_{x=0}^{10} x f(x) = nP = 10(0.35) = 3.5 \approx 4 \text{ نزلات}$$

تطبيق (٦-٦)

في انتخابات مجلس الشعب في إحدى دوائر الوجه البحري تم ترشيح عدد من الأفراد عن الفلاحين من الحزب الوطني وحزب الوفد لشغل ثلاثة مقاعد بالمجلس ، فإذا فرضنا أن المتغير X يشير إلى عدد المرشحين الفائزين من الحزب الوطني عن الفلاحين بهذه الدائرة.

المطلوب :

١- أحسب التوزيع الاحتمالي للمتغير X.
٢- أحسب التوقع لعدد الفائزين من الحزب الوطني.
٣- أحسب الانحراف المعياري لعدد الفائزين من الحزب الوطني.

الحل

إذا فرضنا أن a يشير إلى فوز مرشح من الحزب الوطني ، وb تشير إلى فوز مرشح من حزب الوفد ، فإن فراغ المعاينة للمنتخبين S يصبح على النحو التالي :

S = {(a,a,a) , (a,b,a) , (a,a,b) , (b,a,a) , (b,b,a) ,
(b,a,b) , (a,b,b) , (b,b,b)}

وبالتالي فإن المتغير X = 0,1,2,3 والتوزيع الاحتمالي للمتغير X في هذه الحالة يكون على النحو الموضح بالجدول التالي :

جدول (٦-١١)

X	0	1	2	3	المجموع
$f(x)$	1/8	3/8	3/8	1/8	1

ولحساب القيمة المتوقعة والانحراف المعياري للمتغير X نكون الجدول التالي

جدول (٦-١٢)

X	$p_r(x)$	$xp_r(x)$	$(x-\mu)$	$(x-\mu)^2$	$p_r(x)(x-\mu)^2$
0	$\frac{1}{8}$	0	0-1.5 = -1.5	$\frac{9}{4}$	$\frac{9}{4}\left(\frac{1}{8}\right)=\frac{9}{32}$
1	$\frac{3}{8}$	$\frac{3}{8}$	1-1.5 = -0.5	$\frac{1}{4}$	$\frac{1}{4}\left(\frac{3}{8}\right)=\frac{3}{32}$
2	$\frac{3}{8}$	$\frac{6}{8}$	2-1.5 = 0.5	$\frac{1}{4}$	$\frac{1}{4}\left(\frac{3}{8}\right)=\frac{3}{32}$
3	$\frac{1}{8}$	$\frac{3}{8}$	3-1.5 = 1.5	$\frac{9}{4}$	$\frac{9}{4}\left(\frac{1}{8}\right)=\frac{9}{32}$
المجموع	1	$\frac{12}{8}=1\frac{1}{2}$	0		$\frac{24}{32}=\frac{3}{4}$

١- من الجدول يتضح أن القيمة المتوقعة $\mu=1.5$ حيث :

$$\mu=\sum xp_r(x)=1.5$$

٢- ومن الجدول نجد أن تباين X ، أي $\sigma^2=0.75$ حيث :

$$\sigma^2=\sum(x-\mu)^2\,p_r(x)=\frac{3}{4}=0.75$$

٣- الانحراف المعياري يساوى σ حيث :

$$\sigma=\sqrt{\sigma^2}=\sqrt{0.75}=0.87$$

تطبيق (٦-٧)

تقوم إحدى الشركات الإنتاجية بإنتاج أحد الأجهزة الكهربائية ، فإذا كان الجهاز الواحد المنتج يتكون من 5 أجزاء ، ويعتبر الجهاز غير صالح للاستعمال إذا حدث عطل في أحد أجزئه.

فإذا كان المتغير X يمثل عدد الأجهزة الممكن أن يحدث بها عطل خلال فترة الصلاحية للجهاز ، فإن :

$$X = 0,1,2,3,4,5$$

والجدول التالي يعطى التوزيع الاحتمالي للمتغير X.

جدول (٦-١٣)

X	0	1	2	3	4	5	المجموع
$f(x)$	0.04	0.12	0.18	0.30	0.30	0.06	1

المطلوب :

١- ما هو العدد المتوقع للأجزاء التي يمكن أن يحدث بها أعطال أثناء فترة الصلاحية.

٢- أحسب الانحراف المعياري للأجزاء التي يمكن أن يحدث بها عطل.

الحل

نكون الجدول التالي لحساب القيمة المتوقعة والانحراف المعياري للمتغير X.

جدول (٦-١٤)

X	$p_r(x)$	$xp_r(x)$	$(x-\mu)^2$	$(x-\mu)^2 p_r(x)$
0	0.04	0	$(0-2.88)^2 = 8.29$	0.332
1	0.12	0.12	$(1-2.88)^2 = 3.53$	0.424
2	0.18	0.36	$(2-2.88)^2 = 0.77$	0.139
3	0.30	0.90	$(3-2.88)^2 = 0.01$	0.003
4	0.30	1.20	$(4-2.88)^2 = 1.25$	0.375
5	0.06	0.30	$(5-2.88)^2 = 4.49$	0.369
المجموع	1	2.88		0.542

١- العدد المتوقع للأجزاء التي يمكن أن يحدث بها أعطال أثناء فترة الصلاحية يساوى μ حيث :

$$\mu = \sum xp_r(x) = 2.88 \approx 3$$ أجزاء

٢- الانحراف المعياري يساوى σ حيث :

$$\sigma^2 = \sum (x-\mu)^2 p_r(x) = 1.542$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.542} = 1.24 \approx 1 \text{ جزء}$$

تطبيق (٦-٨)

في إحدى المؤسسات التجارية الكبرى قامت إدارة التسويق بدراسة عن الكميات المطلوبة والمعروضة من إحدى السلع التي تقوم بتسويقها ، وذلك لوضع سياسة التسويق والتخزين لهذه السلعة في العام القادم ، وتم التوصل من خلال البيانات عن الكميات المطلوبة والمعروضة إلى التوزيعات الاحتمالية المتوقعة للكميات المطلوبة والمعروضة كما هو موضح بالجدول التالي :

جدول (٦-١٥)

الكميات	1000	2000	3000	4000	المجموع
احتمال الطلب $f_1(x)$	0.20	0.15	0.50	0.02	1.00
احتمال العرض $f_2(x)$	0.25	0.25	0.30	0.20	1.00

المطلوب :

١- أحسب القيمة المتوقعة للكمية المطلوبة في العام القادم.

٢- أحسب القيمة المتوقعة للعرض في العام القادم.

٣- أوجد التباين والانحراف المعياري لكل من الكمية المطلوبة والكمية المعروضة في العام القادم.

الحل

نفرض أن μ_1, μ_2 هما القيمة المتوقعة للكمية المطلوبة والكمية المعروضة على الترتيب ، كذلك σ_1, σ_2 هما الانحراف المعياري للكميات المطلوبة والمعروضة على التوالي.

-١

$$\mu_1 = \sum x f_1(x)$$
$$= 1000(0.2) + 2000(0.15) + 3000(0.5) + 4000(0.2)$$
$$= 200 + 300 + 1500 + 80 = 2080 \text{ وحدة}$$

٢-

$$\mu_2 = \sum x f_2(x)$$
$$= 1000(0.25) + 2000(0.25) + 3000(0.3) + 4000(0.2)$$
$$= 250 + 500 + 900 + 800 = 2450$$ وحدة

٣- لحساب التباين والانحراف المعياري للكميات المطلوبة والمعروضة نكون الجدول التالي :

جدول (٦-١٦)

X	$(x-\mu_1)$	$(x-\mu_1)^2 f_1(x)$	$(x-\mu_2)$	$(x-\mu_2)^2 f_2(x)$
1000	-1080	233280	-1450	525625
2000	-80	960	-450	50625
3000	920	423200	550	90750
4000	1920	73728	1550	480500
المجموع		731168		1147500

من الجدول نجد أن :

تباين الكمية المطلوبة يساوى σ_1^2 حيث :

$$\sigma_1^2 = \sum (x-\mu_1)^2 f_1(x) = 731168$$

وبالتالي الانحراف المعياري للكمية المطلوبة يساوى σ_1 حيث :

$$\sigma_1 = \sqrt{731168} = 855.03 \approx 855$$ وحدة

وبالمثل :

تباين الكمية المطلوبة يساوى σ_2^2 حيث :

$$\sigma_2^2 = \sum (x-\mu_2)^2 f_2(x) = 1147500$$

وبالتالي الانحراف المعياري للكمية المطلوبة يساوى σ_2 حيث :

$$\sigma_2 = \sqrt{1147500} = 1071.21 \approx 1071$$ وحدة

تطبيق (٦-٩)

في إحدى الشركات التي تقوم بإنتاج منتج معين تم سحب عينة مكونة من ثلاث وحدات من المنتج لاختبار جودة المنتج حيث تم فحص الوحدات الثلاثة لتحديد أيهم سليم Acceptable وأيهم معيب Defective.

المطلوب :

١- تحديد فراغ المعاينة لتجربة السحب.

٢- حدد الأحداث التالية ثم أحسب احتمال كل منهما :

- أن تكون الوحدات الثلاثة سليمة.
- أن تكون وحدتان فقط سليمتين.
- أن تكون الوحدات الثلاثة معيبة.

الحل

إذا فرضنا أن x تشير إلى الوحدة السليمة ، h تشير إلى الوحدة المعيبة فإن :

١-

فراغ المعاينة يساوي S حيث :

$$S = \{(x,x,x) , (x,x,h) , (x,h,x) , (h,x,x) , (h,h,x) , (h,x,h) , (x,h,h) , (h,h,h)\}$$

٢-

- نفرض أن الحدث D يشير إلى أن تكون الوحدات الثلاثة سليمة ، فإن :

$$D = \{(x,x,x)\} \quad \rightarrow \therefore\ p_r(D) = \frac{1}{8}$$

- نفرض أن الحدث E يشير إلى أن تكون وحدتين فقط سليمتين ، فإن :

$$E = \{(x,h,x) , (x,x,h) , (h,x,x)\} \quad \rightarrow \therefore\ p_r(E) = \frac{3}{8}$$

- نفرض أن الحدث K يشير إلى أن تكون الوحدات الثلاثة معيبة ، فإن :

$$K = \{(h,h,h)\} \quad \rightarrow \therefore\ p_r(K) = \frac{1}{8}$$

تطبيق (٦-١٠)

في إحدى الدراسات عن توزيع 1000 عامل بأحد المصانع وجد أن متوسط الأجر الشهري 750 جنيه وأن الأجر الأكثر شيوعاً 800 جنيه والقيمة الوسيطة للأجر 775 جنيه والانحراف المعياري 50 جنيه.

احسب معامل الالتواء ثم عقب على النتائج

الحل

$$\text{معامل الالتواء} = \frac{3\ (\text{الوسط الحسابي - الوسيط})}{\text{الانحراف المعياري}}$$

$$= \frac{3(750 - 775)}{50} = \frac{3(-25)}{50}$$

$$= -1.5$$

بما أن معامل الالتواء سالب فهذا يعني أن أكثر من 50% من العمال يأخذون أجور أقل من المتوسط (750 جنيه)

تطبيق (٦-١١)

أخذت عينة حجمها 1000 من السائقين للسيارات الخاصة وسجلت عدد المخالفات التي قيدت لكل منهم خلال سنة1999 فكانت كما هو موضح بالجدول التالي:-

جدول (٦-١٧)

المجموع	5	4	3	2	1	0	عدد المخالفات
1000	25	75	150	500	150	100	عدد السائقين

المطلوب:-

١- أوجد التوزيع الاحتمالي لعدد المخالفات ثم ارسم التوزيع.

٢- احسب القيمة المتوقعة والانحراف المعياري لعدد المخالفات للسائق الواحد.

٣- احسب احتمال أن يسجل للسائق مخالفتين على الأقل.

الحل:-

١- الجدول التالي يوضح التوزيع الاحتمالي لعدد المخالفات

جدول (٦-١٨)

عدد المخالفات x	0	1	2	3	4	5	المجموع
الاحتمال P(x)	0.1	0.150	0.500	0.150	0.075	0.025	1

٢- لحساب التوقع والانحراف المعياري نكون الجدول التالي:-

جدول (٦-١٩)

x	$P_r(x)$	$xP_r(x)$	$x^2P_r(x)$
0	0.1	0	0
1	0.150	0.150	0.150
2	0.500	1	2
3	0.150	0.45	1.350
4	0.075	0.3	1.2
5	0.025	0.125	0.625
المجموع	1	2.025	5.325

$$\mu = \sum_{i=1}^{5} x_i P_r(x_i)$$
$$= 2.025 \approx 2$$
$$\sigma^2 = \sum_{i=1}^{5} x_i^2 P_r(x_i) - \mu^2$$
$$= 5.325 - (2.025)^2$$
$$= 5.325 - 4.101 = 1.224$$
$$\sigma = \sqrt{\sigma^2} = \sqrt{1.224} = 1.11$$

٣- احتمال أن يسجل للسائق مخالفتين على الأقل

$$= P_r(x=2) + P_r(x=3) + P_r(x=4) + P_r(x=5)$$
$$= 1 - P_r(x=0) + P_r(x=1)$$
$$= 1 - [0.1 + 0.15]$$
$$= 1 - 0.25 = 0.75$$

تطبيق (٦-١٢)

أشارت إحدى الدراسات أن نسبة غياب العمال الذين يعملون بمصانع الحديد والصلب في الأيام شديدة الحرارة تساوي 12% . فإذا تم اختيار عينة مكونة من 6 عمال في أحد الأيام شديدة الحرارة .

١- أوجد احتمال غياب عامل واحد من 6 عمال في هذا اليوم.
٢- أوجد احتمال غياب عامل واحد من 6 عمال في هذا اليوم على الأقل.
٣- أوجد احتمال عدم غياب أي عامل من 6 عمال في هذا اليوم.
٤- أوجد العدد المتوقع للعمال الذين يتم غيابهم في هذا اليوم كذلك الانحراف المعياري.
٥- ارسم التوزيع الاحتمالي لعدد العمال الذين يتم غيابهم ثم عقب على الرسم.

الحل :

n = 6 , p = 12% , (1-p) = 88%

إذا فرضنا أن x تشير إلى عدد العمال الذين يتم غيابهم في هذا اليوم

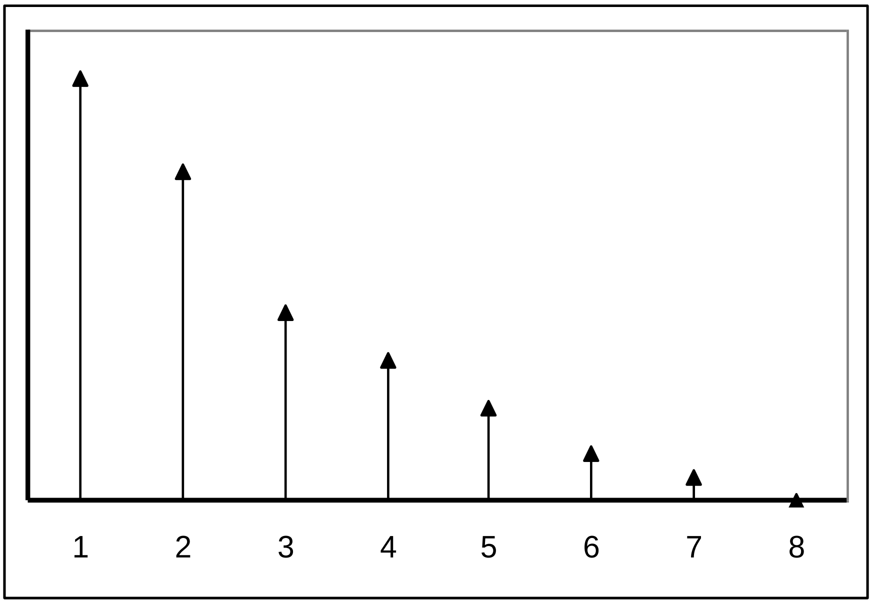
$$\mathbf{P_r(x) = C_x^n\, p^x (1-p)^{n-x}}$$

$$1)\ P_r(x=1) = C_1^6 (0.12)(0.88)^5 = 0.38$$
$$2)\ P_r(x \le 1) = 1 - P_r(x=0)$$
$$= 1 - C_0^6 (0.12)^0 (0.88)^6 = 1 - 0.644$$
$$= 0.536$$
$$3)\ P_r(x=0) = C_0^6 (0.12)^0 (0.88)^6 = 0.464$$
$$4)\ \mu = np = 6 \times 0.12 = 0.72 \approx 1$$
$$\sigma^2 = np(1-p) = 6 \times 0.12 \times 0.88 = 0.6336$$
$$\therefore \sigma = \sqrt{0.6336} = 0.8 \approx 1$$

٥- الجدول التالي يوضح التوزيع الاحتمالي لعدد العمال الذين يتم غيابهم

جدول (٦-٢٠)

x	$P_r(x)$
0	$C_0^6(0.12)^0(0.88)^6 = 0.464$
1	$C_1^6(0.12)\ (0.88)^5 = 0.38$
2	$C_2^6(0.12)^2(0.88)^4 = 15\times0.0144\times0.6 = 0.13$
3	$C_3^6(0.12)^3(0.88)^3 = 20\times0.0017\times0.682 = 0.023$
4	$C_4^6(0.12)^4(0.88)^2 = 15\times0.0002\times0.7744 = 0.0023$
5	$C_5^6(0.12)^5(0.88) = 6\times0.00005\times0.88 = 0.00013$
6	$C_6^6(0.12)^6(0.88)^0 = 0.000003$
$\sum$	1



شكل (٦-١٤)

يوضح التوزيع الاحتمالي لعدد العمال الذين يتم غيابهم

تطبيق(٦-١٣)

في إحدى الدراسات الحديثة عن نسبة قائدي السيارات في مصر الذين يستخدمون حزام المقعد وفقاً للتعليمات المرورية الحديثة وجد أن نسبة الذين يستخدمون الحزام 17% فإذا أخذت عينة عشوائية مكونة من 5 أشخاص من قائدي السيارات.

١- أوجد احتمال عدم استخدام أي شخص في العينة للحزام.
٢- احتمال وجود شخص واحد يستخدم الحزام.
٣- احتمال استخدام جميع الأشخاص للحزام.
٤- أوجد العدد المتوقع للأشخاص الذين يستخدمون الحزام كذلك الانحراف .المعياري.

الحل

احتمال أن يستخدم قائد السيارة الحزام = 0.17

احتمال عدم يستخدم قائد السيارة للحزام = 0.83

إذا فرضنا أن x تشير إلى عدد قائدي السيارات المستخدمين للحزام :

$$1)\ P_r(x=0)=C_0^5(0.17)^0(0.83)^5=0.394$$

$$2)\ P_r(x=1)=C_1^5(0.17)\ \ (0.83)^4=0.403$$

$$3)\ P_r(x=5)=C_5^5(0.17)^5(0.83)^0=0.00014$$

$$4)\ \mu=np=5\times 0.17=0.85\approx 1$$

$$\sigma^2=np(1-p)=5\times 0.17\times 0.83=0.7055$$

$$\sigma=\sqrt{0.7055}=0.84$$

تطبيق (٦-١٤)

في دراسة قام بها أحد البنوك التجارية وجدوا أن العدد المتوقع لطلبات القروض التي يزيد قيمة كل قرض منها عن مليون جنيه يساوي 5 طلبات سنوياً

١- أوجد التوزيع الاحتمالي لعدد الطلبات ثم وضح ذلك بيانياً
٢- أوجد احتمال عدم تقديم أي طلبات للبنك
٣- أوجد احتمال تقديم أكثر من 3 طلبات
٤- أوجد احتمال تقديم أكثر من 6طلبات

الحل

١- إذا أشرنا للعدد المتوقع لطلبات القروض في السنة بالرمز λ فإن:-

$$\lambda=5$$

كذلك إذا أشرنا إلى عدد الطلبات بالرمز x فإن :-

$$P_r(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \frac{e^{-5}5^x}{x!} \qquad x = 0,1,2,3,.. ......$$

٢-احتمال عدم تقديم أي طلب للبنك $= P_r(x=0)$

$$P_r(x) = \frac{e^{-5}(5)^0}{0!} = 0.0067$$

٣- احتمال تقديم أكثر من 3 طلبات $= P_r(x=4) + P_r(x=5) + ...$

$$= 1 - [P_r(x=0) + P_r(x=1) + P_r(x=2) + P_r(x=3)$$

$$= 1 - [\frac{e^{-5}(5)^0}{0!} + \frac{e^{-5}(5)}{1!} + \frac{e^{-5}(5)}{2!} + \frac{e^{-5}(5)^3}{3!}$$

$$= 1 - [0.0067 + 0.0335 + 0.08375 + 0.1396]$$

$$= 1 - [0.26355] = 0.73645 \approx 0.74$$

٤- $= P_r(x>6) = 1 - [P_r(x=0) + P_r(x=1) + P_r(x=2)$

$$+ P_r(x=3) + P_r(x=4) + P_r(x=5) + P\ (x=6)]$$

$$= 1 - [0.0067 + 0.0335 + 0.08375 + 0.1396$$

$$+ 0.1745 + 0.1745 + 0.1454]$$

$$= 1 - 0.75795 = 0.24205 \approx 0.24$$

تطبيق (٦-١٥)

إذا كان عدد السيارات التي تصل إلى إحدى محطات الخدمة متغير يتبع توزيع بواسون بمعدل متوقع 5 سيارات في الساعة

المطلوب:-

١- أوجد دالة الاحتمال لعدد السيارات التي تصل إلى المحطة في الساعة.
٢- أوجد احتمال عدم وصول أي سيارة للخدمة في المحطة في ساعة معينة.
٣- أوجد احتمال وصول 3 سيارات على الأقل للخدمة في ساعة معينة.
٤- أوجد دالة الاحتمال لعدد السيارات التي تصل إلى المحطة في الدقيقة.
٥- أوجد احتمال عدم وصول أي سيارة في الدقيقة.
٦- أوجد احتمال وصول 2 سيارة في الدقيقة على الأكثر.

الحل:-

١- إذا أشرنا إلى عدد السيارات التي تصل إلى محطة الخدمة في الساعة بالرمز x فإن:-

$$P_r(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad , x = 0,1,2,3,\dots\dots$$

$$= \frac{e^{-5}(5^x)}{x!} \quad , x = 0,1,2,3,\dots\dots$$

٢- احتمال عدم وصول أي سيارة للخدمة في المحطة في ساعة معينة

$$= P_r(x = 0) = \frac{e^{-5}(5)^0}{0!} = e^{-5} = 0.00674$$

٣- $P_r(x \geq 3) = 1 - [P_r(x = 0) + P_r(x = 1) + P_r(x = 2)]$

$$= 1 - [\frac{e^{-5}(5)^0}{0!} + \frac{e^{-5}(5)}{1!} + \frac{e^{-5}(5)^2}{2!}]$$

$$= 1 - e^{-5}[1 + 5 + 12.5] = 1 - 0.00674(18.5)$$

$$= 1 - 0.1245 = 0.8755$$

٤- إذا فرضنا أن λ' هي المعدل المتوقع لعدد السيارات التي تصل إلى المحطة في الدقيقة ، فإن :

$\lambda' =$ المعدل المتوقع لعدد السيارات في الساعة $\times \frac{1}{6}$

$$\lambda' = \lambda \times \frac{1}{6} = 5 \times \frac{1}{6} = 0.0833$$

وإذا فرضنا أن y تشير إلى عدد السيارات التي تصل إلى المحطة في الدقيقة فإن :

$$p_r(y) = \frac{e^{\lambda'}(\lambda')^y}{y!} \quad , \quad y = 0,1,2,\dots\dots$$

٥- احتمال عدم وصول أي سيارة في دقيقة معينة

$$p_r(y=0)=\frac{e^{-0.0833}(0.0833)^0}{0!}=e^{-0.0833}$$

٦- احتمال وصول 2 سيارة في الدقيقة على الأكثر

$$p_r(y\le 2)=p_r(y=0)+p_r(y=1)+p_r(y=2)$$
$$=\frac{e^{-0.0833}(0.0833)^0}{0!}+\frac{e^{-0.0833}(0.0833)^1}{1!}+\frac{e^{-0.0833}(0.0833)^2}{2!}$$
$$=e^{-0.0833}\left[1+0.0833+0.00347\right]=0.920075(1.08677)=0.9999$$

(٦-٦) تمرينات Exercises

(٦-١) :

إذا رميت 4 قطع عملة متوازنة عشوائياً :

١- أوجد التوزيع الاحتمالي لعدد الصور التي تظهر على السطح العلوي لكل القطع.

٢- أوجد احتمال ظهور صورة واحدة على الأكثر.

٣- أوجد احتمال ظهور صورة واحدة على الأقل.

٤- أوجد احتمال ظهور الكتابة على قطعة واحدة.

(٦-٢) :

إذا كان احتمال تخرج الطالب الذي يلتحق بإحدى الكليات العسكرية (أي اجتيازه مدة الدراسة بالكلية) يساوى 0.8 ، فإذا سحبت عينة مكونه من 10 طلاب بإحدى الكليات العسكرية من الطلاب المستجدين.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد الطلاب الذين يتم تخرجهم بهذه الكلية.

٢- أوجد التوزيع الاحتمالي التقريبي لعدد الطلاب الذين يتم تخرجهم بهذه الكلية وقارنه بالتوزيع في (١).

٣- أوجد احتمال تخرج طالب واحد على الأكثر.

٤- أوجد العدد المتوقع للطلاب الخريجين من العينة.

(٦-٣) :

تقوم إحدى شركات التأمين على حياة العاملين بإحدى شركات الدخان الكبرى فإذا كان احتمال وفاة العامل في هذه الشركة في سن أقل من 40 سنة من خلال التاريخ الطبي للعمال بالشركة يساوى 0.2 . فإذا قامت الشركة بالتأمين على 10 من العاملين.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد المتوفين في الشركة في عمر أقل من 40 سنة.

٢- أوجد التوزيع الاحتمالي التقريبي لعدد المتوفين في الشركة في عمر أقل من 40 سنة وقارنه بالتوزيع في (١).

٣- أوجد احتمال وفاة عاملين على الأقل من المؤمن عليهم في عمر أقل من 40 سنة.

٤- أوجد احتمال وفاة جميع العمال المؤمن عليهم.

٥- أوجد العدد المتوقع والانحراف المعياري للمتوفين في عمر أقل من 40 سنة ومؤمن عليهم.

(٦-٤) :

تقوم إحدى الشركات الإنتاجية بإنتاج أحد المنتجات الغذائية في وحدات معينة في أكياس . فإذا كان احتمال وجود كيس غير مطابق للمواصفات من إنتاج هذه الشركة يساوي 0.002 ، فإذا سحبت عينة مكونة من 100 كيس.

المطلوب

١- أحسب احتمال وجود كيس غير مطابق للمواصفات.
٢- أحسب احتمال وجود كيس واحد على الأقل غير مطابق للمواصفات.
٣- أحسب احتمال وجود كيسين على الأكثر غير مطابقين للمواصفات.
٤- أوجد العدد المتوقع والانحراف المعياري لعدد الأكياس غير المطابقة للمواصفات.

(٦-٥) :

إذا كان احتمال حدوث تصادم في الطريق الصحراوي بين القاهرة والإسكندرية لأي سيارة تمر بهذا الطريق يساوى 0.0005 ، فإذا تم عمل إحصائية عن السيارات التي تمر في هذا الطريق في فترة زمنية معينة فكان عددها 4000 سيارة.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد السيارات التي يمكن أن تحدث تصادم بهذا الطريق خلال الفترة محل الاعتبار.
٢- أوجد احتمال تصادم أكثر من سيارة بهذا الطريق خلال الفترة محل الاعتبار.
٣- أوجد العدد المتوقع للسيارات التي يمكن أن تحدث تصادم بالطريق.
٤- أوجد دالة التوزيع التراكمية لعدد السيارات التي يمكن أن تحدث تصادم بالطريق.

(٦-٦) :

صندوق يحتوى على 7 كرات خضراء . فإذا سحبت 5 كرات عشوائياً (سحب بدون إرجاع).

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد الكرات البيضاء ووضح ذلك بيانياً.
٢- أوجد التوزيع الاحتمالي لعدد الكرات الخضراء ووضح ذلك بيانياً.
٣- أوجد احتمال ظهور كرة واحدة على الأقل بيضاء.

(٦-٧) :

في إحدى القرى المصرية كانت نسبة الأميين بها 44% ، فإذا أخذت عينة مكونة من 5 أفراد من هذه القرية.

المطلوب :

١- أوجد احتمال وجود واحد على الأقل من أمي.
٢- أوجد احتمال أن يكون جميع أفراد العينة أميين.
٣- أوجد العدد المتوقع للأميين بالعينة.

(٦-٨) :

يقوم أحد المصانع بإنتاج منتج معين ، وعند اختبارات جودة الإنتاج وجد أن 0.01% من المنتجات معيبة . فإذا أخذت عينة مكونة من 10000 وحدة من المنتج وأجريت الاختبارات عليها.

المطلوب :

١- أحسب احتمال أن تكون جميع الوحدات سليمة.
٢- أحسب احتمال أن تكون جميع الوحدات معيبة.
٣- أحسب احتمال وجود وحدة واحدة معيبة على الأقل.
٤- أوجد العدد المتوقع للوحدات المعيبة.

(٦-٩) :

في إحدى الدراسات السكانية عن جنس المولود (ذكر أو أنثي) في إحدى القرى المصرية وجد أن احتمال أن تنجب الأم مولود ذكر 0.45 واحتمال أن تنجب الأم مولود أنثي 0.55 ، فإذا تم اختيار أحد الأسر بهذه القرية وكان لديها أربع أطفال.

المطلوب :

١- أوجد التوزيع الاحتمالي لعدد الأطفال الذكور بالأسرة.
٢- أوجد التوزيع الاحتمالي لعدد الأطفال الإناث بالأسرة.
٣- أوجد العدد المتوقع للأطفال الذكور بالأسرة كذلك العدد المتوقع للأطفال الإناث بالأسرة.

(٦-١٠) :

فيما يلي التوزيع الاحتمالي لعدد الوحدات غير المطابقة للمواصفات في الإنتاج اليومي لأحد المصانع.

جدول (٦-٢١)

عدد الوحدات المعيبة (x)	0	1	2	3	4	5	6	المجموع
دالة الاحتمال $f(x)$	0.80	0.10	0.075	0.006	0.008	0.007	0.004	1

المطلو ب:

١- وضح بيانياً التوزيع الاحتمالي لعدد الوحدات المعيبة.

٢- أحسب الاحتمالات التالية :

$p_r(x+2 \geq 3)$　　$p_r(x \geq 2)$

$p_r(x-5<1)$　　$p_r(x+1>2)$

٣- أحسب التوقع والانحراف المعياري لعدد الوحدات المعيبة.

جدول (٦-٢٢)

X	0	1	2	3	المجموع
$f(x)$	$\frac{1}{10}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{10}$	1

(٦-١١) :

في كل حالة من الحالات التالية حدد أي المتغيرات متقطعة وأيهم متصلة:-

١- عدد الخطابات التي ترد يومياً إلى أحد مراكز البريد بمدينة القاهرة.

٢- عدد السيارات التي تصل إلى أحد مراكز الإصلاح والصيانة للسيارات.

٣- عدد السيارات التي تنتجها إحدى شركات انتاج السيارات في السنة.

٤- عدد العملاء اليومي المسافرون على أحد\ الخطوط الجوية المصرية.

(٦-١٢) :

في احدى مصانع انتاج السيارات وفي مرحلة اختبار السيارة قامت الشركة بتحديد نسبة السيارات التي يعاد اختبارها أكثر من مرة نتيجة وجود أحد العيوب في التصنيع أو التجميع ، والجدول التالي يوضح عدد مرات اعادة الإختبار ونسبة السيارات التي يعاد اختبارها

جدول (٦-٢٣)

عدد مرات اعادة الاختبار	0	1	2	3	4	المجموع
نسبة السيارات	0.27	0.37	0.23	0.08	0.05	1

المطلوب :

١- ارسم دالة الاحتمال ثم عقب على الرسم
٢- أوجد دالة التوزيع التراكمية ووضح ذلك بيانياً
٣- أوجد العدد المتوقع لعدد مرات إعادة الفحص ثم ؟اوجد التباين والانحراف المعياري
٤- اوجد احتمال أن تكون السيارة ليس بها عيوب
٥- أوجد احتمال أن يكون بالسيارة عيب أو أكثر

(٦-١٣) :

في دراسة عن مدة الاستخدام اليومي بالساعات لشبكة الإتصالات (الإنترنت) في أحد الأحياء بمحافظة القاهرة والجدول التالي يعطي نسبة المستخدمين في هذا الحي.

جدول (٦-٢٤)

عدد الساعات اليومية	0	0.5	1	1.5	2	2.5	3	3.5	4	المجموع
نسبة المستخدمين	0.5	0.1	0.05	0.05	0.11	0.1	0.06	0.03	0.01	1

المطلو ب:

١- ارسم دالة الاحتمال ثكم عقب على الرسم
٢- أوجد فترة استخدام الإنترنت المتوقعة لهذا الحي ثم أوجد التباين والانحراف المعياري
٣- إذا تم اختيار احد الأفراد في هذا الحي أوجد:-
أ- احتمال عدم استخدامه للشبكة
ب-احتمال استخدامه للشبكة ساعتين فأكثر
جـ- احتمال استخدامه أقل من 3.5ساعة يومياً

صفحة ٢٢٣

$$p_r(x_1) = \frac{1}{10}$$

$$p_r(y_1) = \frac{6}{10}$$

الباب السابع

التوزيعات الاحتمالية المتصلة

Continuous Probability Distribution

(٧-١) التوزيع الاحتمالي المتصل

Continuous Probability Distribution

(٧-٢)التوزيع الأسي

Exponential Distribution

(٧-٣)التوزيع المعتاد القياسي

Standard Normal Distribution

(٧-٤)**التوزيع المعتاد**

Normal Distribution

(٧-٥) استخدام التوزيع المعتاد كتقريب لتوزيع ذي الحدين

Approximation A Binomial Distribution With A Normal Distribution

(٧-٦)أمثلة تطبيقية

Applied Examples

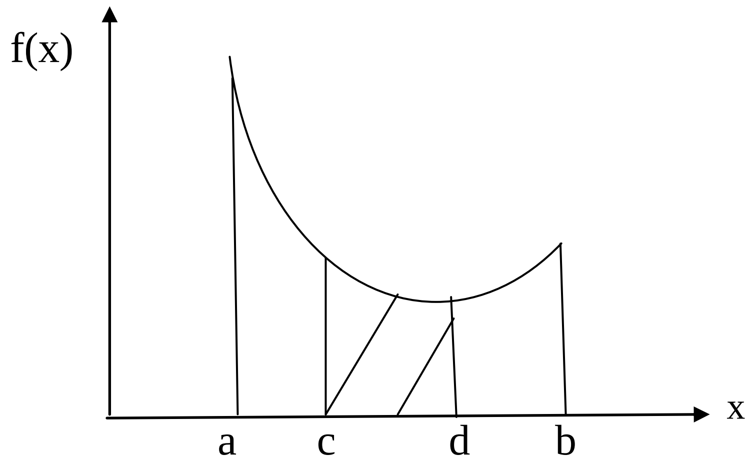
(٧-٧)تمرينات

Exercises

(٧-١)التوزيع الاحتمالي المتصل (المستمر)
Continuous Probability Distribution

إذا كانت نتائج التجربة العشوائية أو(الأحداث) لا يمكن عدها Non – Countable فإن المتغير العشوائي الذي يمثل هذه النتائج أو(الأحداث) يسمى متغير عشوائي متصل (أو مستمر)Continuous Random Variable ويمثل المتغير العشوائي المتصل بفترة Interval على المحور الأفقي وليس بنقط كما في حالة المتغير المتقطع.

مما سبق يتضح أنه في حالة المتغير المتصل لا يمكن وضع التوزيع الاحتمالي لهذا المتغير في صورة جدول توزيع احتمالي ، ولكن يمكن صياغة التوزيع الاحتمالي للمتغير المتصل X باستخدام دالة تسمى بدالة كثافة الاحتمال Probability Density Function وعادة يرمز لها بالرمز f(x) وبالنسبة للمتغير المتصل X الواقع في الفترة [a , b] بمعنى أن $a \le x \le b$ فإن احتمال وقوع X في الفترة [c , d] هو المساحة المحصورة أسفل منحنى دالة كثافة الاحتمال f(x) والمحور X في الفترة [c , d] كما هو موضح بالشكل التالي:



شكل(٧-١)

وبما أن احتمال وقوع المتغير X في الفترة [c , d] يمثل دالة في المتغير X ويرمز لها بالرمز $P_r(c \le x \le d)$ ويتم حساب قيمة احتمال وقوع المتغير X خلال الفترة [c , d] على النحو التالي:

$$P_r(c \le x \le d) = \int_c^d f(x) \quad dx \qquad \textbf{(7-1)}$$

مثال (٧-١)

إذا كان X متغير عشوائي متصل له دالة كثافة الاحتمال f(x) حيث :

$$f(x) = \begin{cases} \frac{1}{16}x & 2 \le x \le 6 \\ 0 & \text{فيما عدا ذلك} \end{cases} \qquad (2\text{-}7)$$

احسب الاحتمالات التالية :-

1) $P_r\,(2 \le x \le 6)$

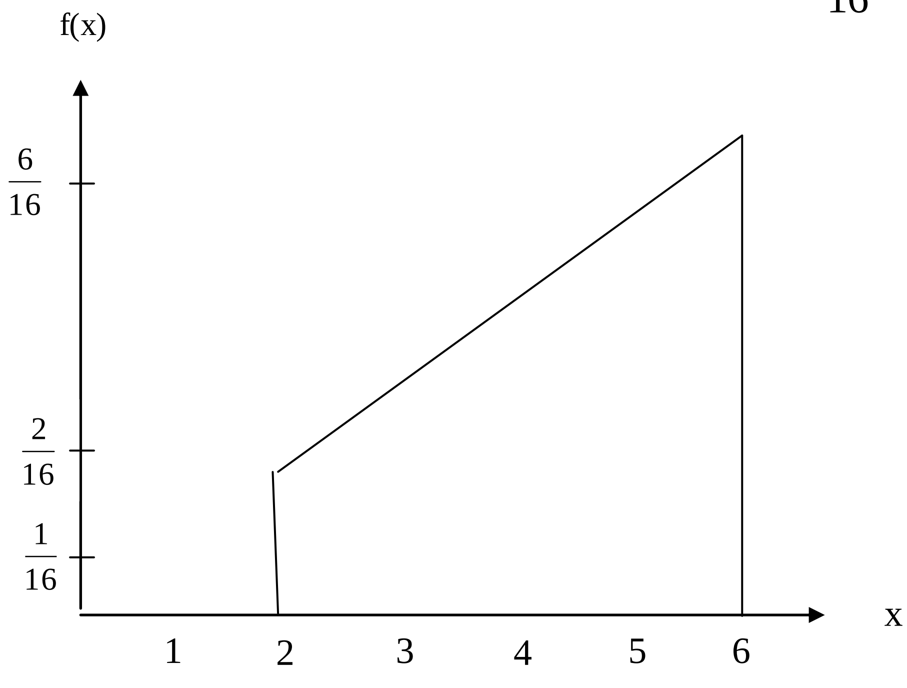
2) $P_r\,(x \le 4)$

3) $P_r\,(3 \le x \le 5)$

4) $P_r\,(0 \le x \le 3.5)$

الحل :

بما أن $f(x) = \frac{x}{16}$ أي دالة خطية معرفة في الفترة [2 , 6] والشكل التالي يوضح ذلك :



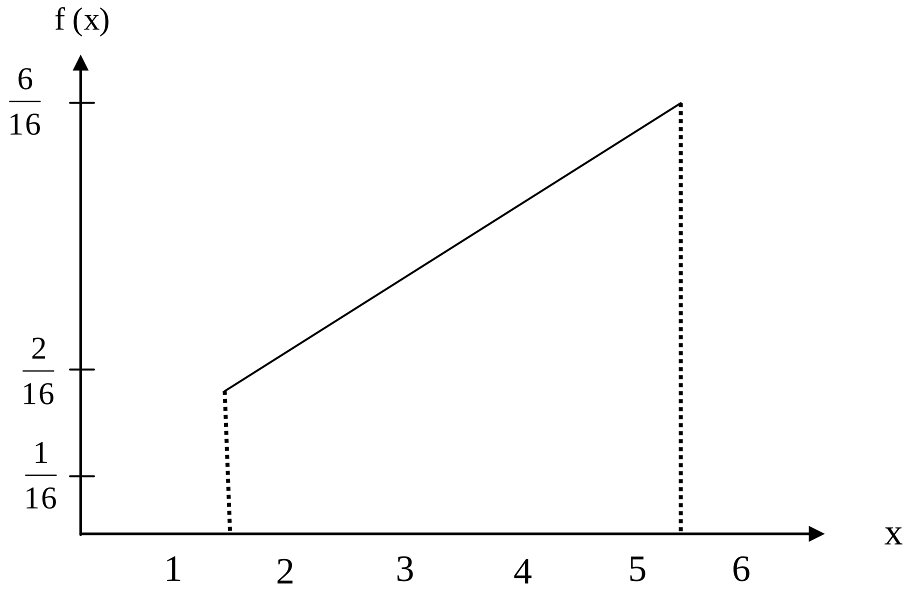
شكل (٧-٢)

$$1)\quad P_r(2 \le x \le 6) = \int_2^6 f(x)\quad dx = \int_2^6 \frac{1}{16}x \quad dx$$

$$= \frac{1}{16}[\frac{x^2}{2}]_2^6 = \frac{1}{16}[\frac{6^2}{2} - \frac{2^2}{2}]$$

$$= \frac{1}{16}[\frac{36}{2} - \frac{4}{2}] = \frac{1}{16}[\frac{32}{2}] = 1$$

كما هو وضح بالشكل التالي :



شكل (٧-٣)

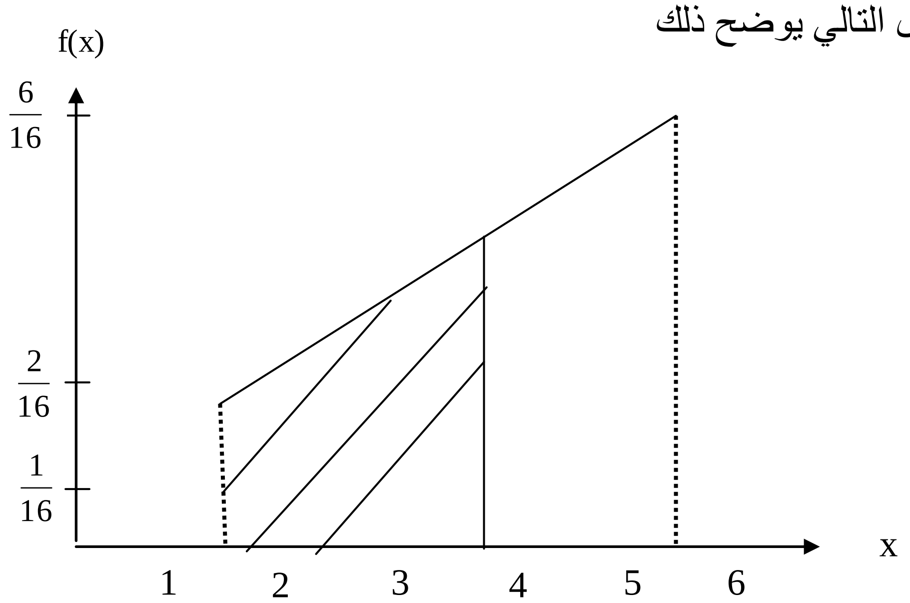
2)

$$P_r(x \le 4) = \int_2^4 f(x)\ dx = \int_2^4 \frac{1}{16}\ x\ dx$$

$$= \frac{1}{16}[\frac{x^2}{2}]_2^4 = \frac{1}{16}[\frac{4^2}{2} - \frac{2^2}{2}]$$

$$= \frac{1}{16}[8 - 2] = \frac{6}{16}$$

والشكل التالي يوضح ذلك



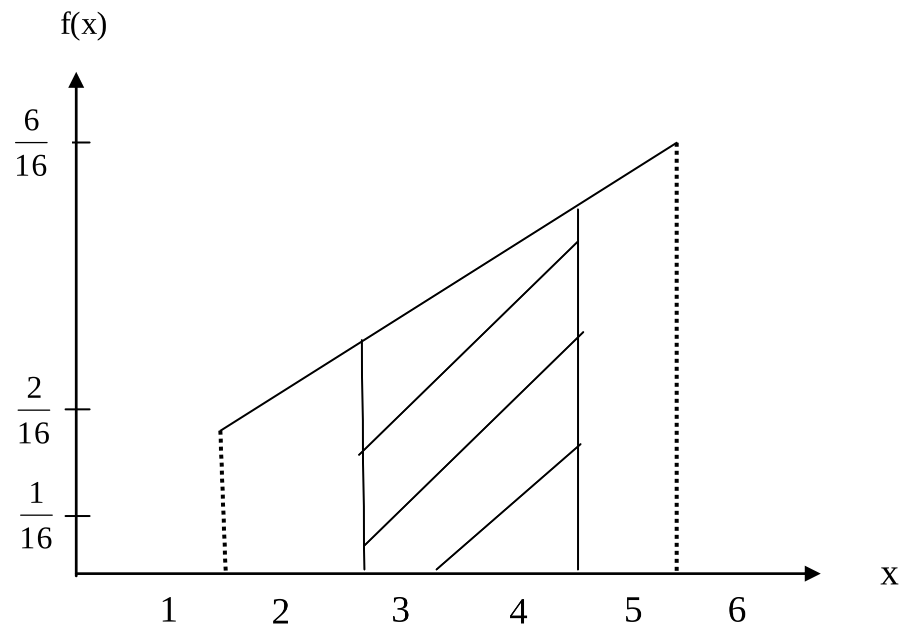
شكل (٧-٤)

$$3)\quad P_r(3 \le x \le 5) = \int_3^5 \frac{1}{16}\, x \;\; dx \;=\frac{1}{16}\Big[\frac{x^2}{2}\Big]_3^5$$

$$= \frac{1}{16}\Big[\frac{5^2}{2} - \frac{3^2}{2}\Big] = \frac{1}{16}\Big[\frac{25}{2} - \frac{9}{2}\Big]$$

$$= \frac{1}{16}\Big[\frac{16}{2}\Big] = \frac{1}{2}$$

والشكل التالي يوضح ذلك :



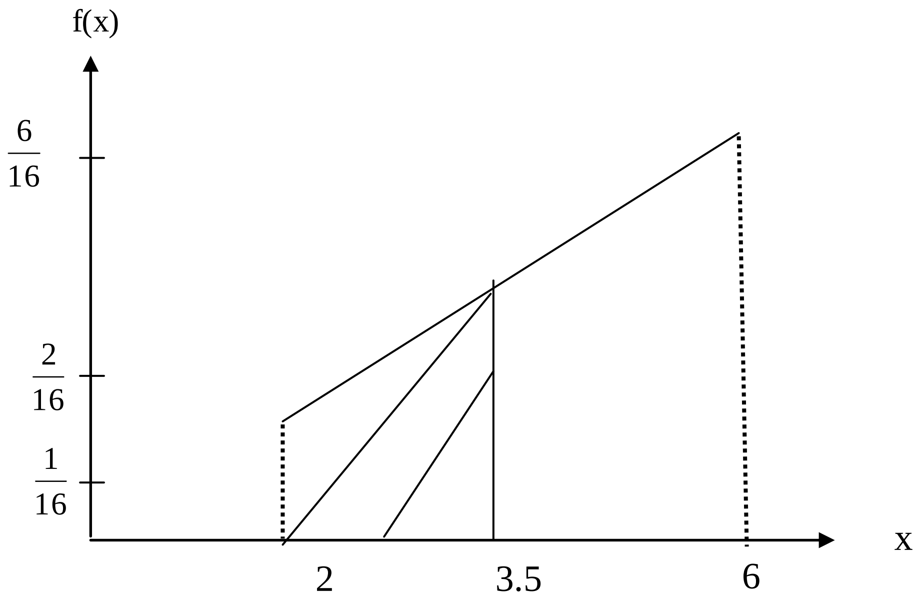
4) شكل (٧-٥)

$$P_r(2 \le x \le 3.5) = \int_{2}^{3.5} \frac{1}{16}\, x \, dx$$

$$= \frac{1}{16}\left[\frac{3.5^2}{2} - \frac{2^2}{2}\right] = \left[\frac{12.25}{2} - \frac{4}{2}\right]$$

$$= \frac{1}{16}\left[\frac{825}{2}\right] = \frac{825}{3200} = 0.2078$$

والشكل التالي يوضح ذلك:



شكل (٧-٦)

مثال (٧-٢)

إذا كان X متغير عشوائي متصل بحيث $0 \le x \le 4$ ودالة كثافة احتمال f(x) تأخذ الصياغة التالية :

$$f(x) = \begin{cases} \frac{1}{28}(x+5) & , 0 \le x \le 4 \\ 0 & \text{فيما عدا ذلك} \end{cases} \qquad (7\text{-}3)$$

المطلوب :

(١) وضح بيانياً دالة كثافة الاحتمال

(٢) احسب الاحتمالات التالية

i) $P_r(2 \le x \le 4)$

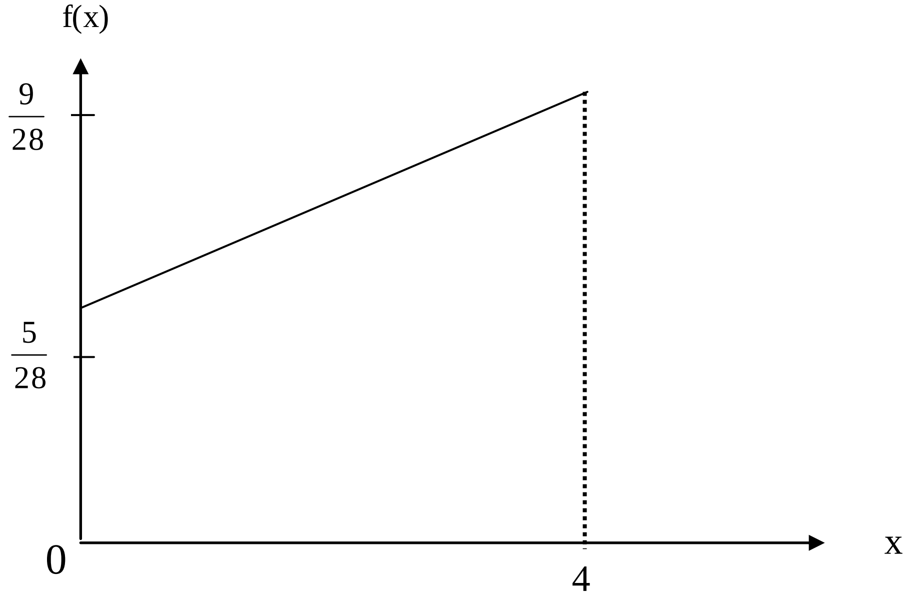
ii) $P_r(x \le 4)$

iii) $P_r(x \le 2.5)$

الحل:

(١) بما أن $f(x) = \frac{1}{28}(x+5)$ في الفترة $0 \le x \le 4$ ، أي أن :

f(x) دالة خطية متزايدة في X ، فهي تمثل بخط مستقيم كما هو موضح بالشكل التالي:



شكل(٧-٧)

2-

$$i)P_r(2 \le x \le 4) = \int_2^4 f(x)\ dx$$

$$= \frac{1}{28}[\frac{x^2}{2} + 5x\ ]_2^4$$

$$= \frac{1}{28}[(\frac{16}{2} + 20) - (\frac{4}{2} + 10)]$$

$$= \frac{1}{28}(16) = \frac{4}{7}$$

$$ii)P_r(x \le 4) = \int_0^4 f(x)\quad dx = \frac{1}{28}[\frac{x^2}{2} + 5x]_0^4$$

$$= \frac{1}{28}[(\frac{16}{2} + 20) - (\frac{0}{2} + 0)\ ]$$

$$= \frac{1}{28}(28) = 1$$

$$iii)P_r(x \le 2.5) = \int_{2.5}^4 \frac{1}{28}(x + 5)\ \ dx$$

$$= \frac{1}{28}[\frac{x^2}{2} + 5x]_{2.5}^4$$

$$= \frac{1}{28}[(\frac{16}{2} + 20) - (\frac{6.25}{2} + 12.5)]$$

$$= \frac{1}{28}[\ 28 - 15.625] = \frac{1}{28}(12.375)$$

$$= 0.442$$

خصائص دالة كثافة الاحتمال:-

إذا كانت f(x) دالة كثافة احتمال للمتغير x حيث :

$$-\infty \le x \le +\infty$$

فإن :

(١) الدالة f(x) دالة غير سالبة ، بمعنى أن :

$$f(x) \ge 0 \quad , -\infty \le x \le +\infty \qquad (7\text{-}4)$$

(٢)تكامل الدالة f(x) على الفترة التي يقع فيها المتغير x يساوي الواحد الصحيح بمعنى:

$$\int_{-\infty}^{+\infty} f(x) \quad dx = 1 \qquad (7\text{-}5)$$

أي أن المساحة المحصورة بين منحنى الدالة f(x) ومحور x تساوي الواحد الصحيح (حيث أن مجموع الاحتمالات يساوي واحد) ويترتب على الخصائص (١) ، (٢) النتائج التالية:

١- إذا كان $a \le x \le b$ فإن:

$$P_r(a \le x \le b) = \int_a^b f(x)\, dx \qquad (7\text{-}6)$$

كذلك إذا كان $a \le x \le b$ فإن :

$$P_r(a \le x \le b) = \int_a^b f(x)\, dx \qquad (7\text{-}7)$$

من (7-5)، (7-6) نجد أن المساحة أسفل المنحنى f(x) لا تتأثر بنقطتي طرفي الفترة المعرف عليها x أي لا تتأثر بالنقطتين a , b .

٢- إذا كان $a \le x \le b$ فإن

$$P_r(x < a) = \int_{-\infty}^{a} f(x)\, dx = 0 \qquad (7\text{-}8)$$

كذلك:

$$P_r(x > b) = \int_b^{\infty} f(x)\, dx = 0 \qquad (7\text{-}9)$$

حيث أن الدالة f(x) غير معرفة في الفترات $(b, +\infty), (-\infty, a)$

مثال(٧-٣)

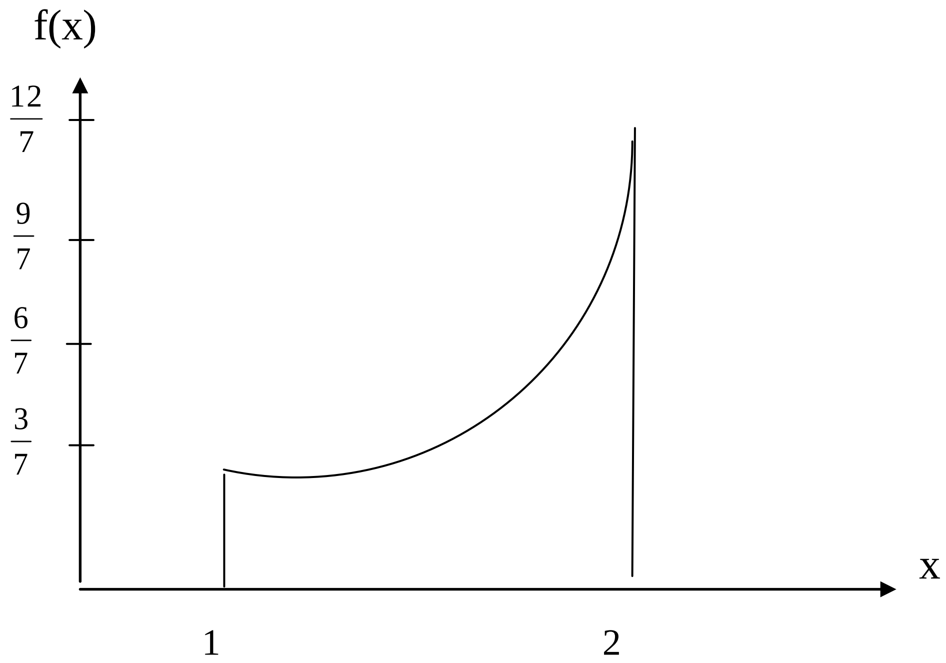
أثبت أن الدالة $f(x)$ دالة كثافة احتمال ، حيث :

$$f(x)=\begin{cases}\frac{3}{7}x^2 & 1\le x\le 2\\ 0 & \text{فيما عدا ذلك}\end{cases}$$

ووضح ذلك بيانياً

الحل

بما أن الدالة $f(x)=\frac{3}{7}x^2$ دالة من الدرجة الثانية أي تمثل بمنحنى كما هو موضح في الشكل التالي:



شكل (٧-٨)

كذلك نجد أن :

$$f(x)=\frac{3}{7}x^2>0 \qquad (7-10)$$

في الفترة $1\le x\le 2$

$$\int_1^2 \frac{3}{7}x^2\ dx=\frac{3}{7}\left[\frac{x^3}{3}\right]_1^2=\frac{3}{7}\left[\frac{8}{3}-\frac{1}{3}\right]=1 \qquad (7-11)$$

من (7-10) ، (7-11) نجد أن الدالة $f(x)$ دالة كثافة احتمال

مثال (٧-٤)

إذا كانت f(x) دالة كثافة احتمال للمتغير x حيث:

$$2 \le x \le 4$$

$$, f(x) = \begin{cases} c(x+2) & ,2 \le x \le 4 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

المطلوب

احسب قيمة الثابت c ، ثم احسب $P_r(2.5 \le x \le 3.5)$ ووضح ذلك بيانياً.

الحل

بما أن f(x) دالة كثافة احتمال، إذن :

$$P_r(2 \le x \le 4) = \int_2^4 f(x)\ dx$$

$$= \int_2^4 c(x+2)\ dx = 1$$

$$= c\,[\frac{x^2}{2} + 2x\ ]_2^4 = 1$$

$$c\,[16 - 6] = 1$$

$$\therefore c = \frac{1}{10}$$

وبالتالي فإن

$$f(x) = \frac{1}{10}(x+2) = \frac{1}{10}x + \frac{1}{5}$$

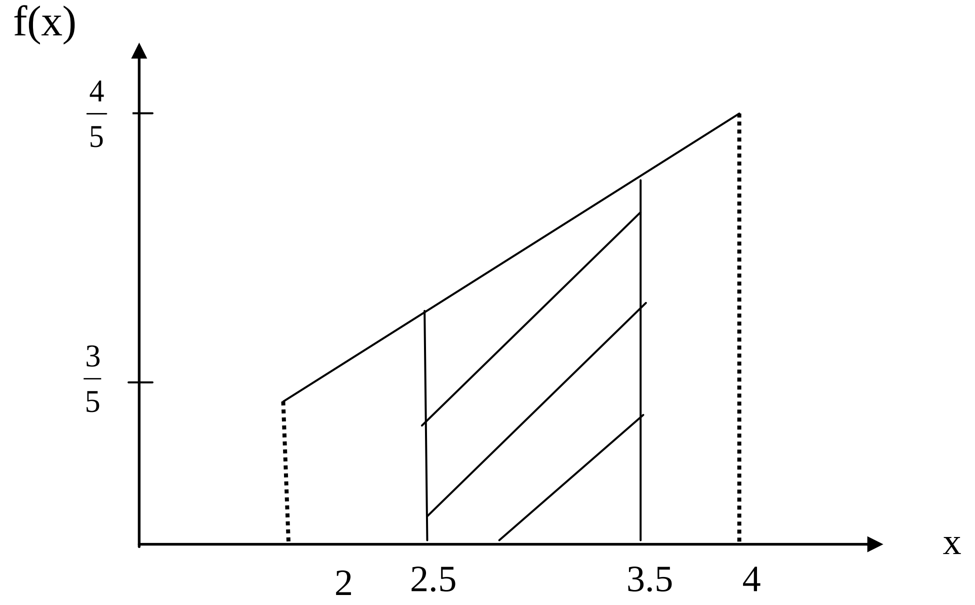
كذلك :

$$P_r(2.5 \le x \le 3.5) = \int_{2.5}^{3.5} \frac{1}{10}(x+2)\ dx$$

$$= \frac{1}{10}[(\frac{x^2}{2} + 2x)]_{2.5}^{3.5} = \frac{1}{10}[(\frac{3.25}{2} + 7) - (\frac{2.25}{2} + 5)]$$

$$= 1.3125 - 0.8125 = 0.5$$

والشكل التالي يوضح ذلك:



شكل (٧-٩)

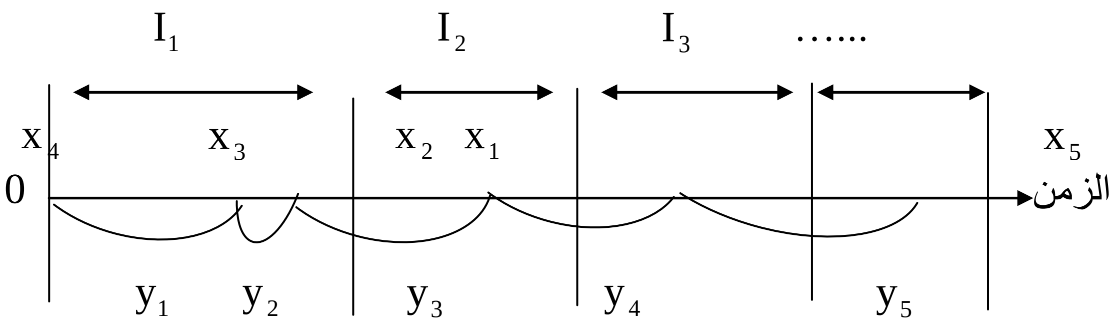
(٧-٢)التوزيع الآسي السالب

Negative Exponential Distribution

في الفصل (٦-٥) من الباب السابق تناولنا بالدراسة متغير بواسون والتوزيع الاحتمالي له من الجانبين النظري والتطبيقي ويرتبط بمتغير بواسون متغير عشوائي أخر يسمى بالمتغير الآسي Exponential Random Variable ومما سبق وجدنا أن متغير بواسون x يمثل عدد مرات النجاح في n من المحاولات المتماثلة المستقلة ولكل محاولة نتيجتين متنافيتين هما نجاح باحتمال النجاح p وفشل باحتمال(1-p) بحيث $n \longrightarrow \infty, p \longrightarrow 0$ حيث $\lambda = np$ أو بعبارة أخرى فالمتغير x هو عدد مرات النجاح في فترة زمنية معينة وبالتالي تصبح λ هو العدد المتوقع خلال تلك الفترة

وتمثل الفترة الزمنية بين وقوع نجاحين (حدثين) متتالين متغير عشوائي أيضاً يسمى بالمتغير الآسي وسوف نرمز له بالرمز y

والشكل التالي يوضح العلاقة بين متغير بواسون x والمتغير الآسي y فإذا فرضنا أن x تمثل النجاح رقم(i) حيث $i=1,2,3,\ldots$ والمتغير y_i يمثل الفترة الزمنية بين وقوع النجاح رقمI والنجاح التالي له مباشرة (أي النجاح رقم (i+1) حيث الزمن مقسم إلى فترات متساوية $i_1, i_2, \ldots\ldots$ كما هو موضح بالشكل التالي:-



شكل (٣-١٠)

ونظراً لأن المتغير الآسي y يمثل الفترة الزمنية Interval Time بين وقوع حدثين متتالين فهو متغير غير سالب ومتغير متصل أيضاً.

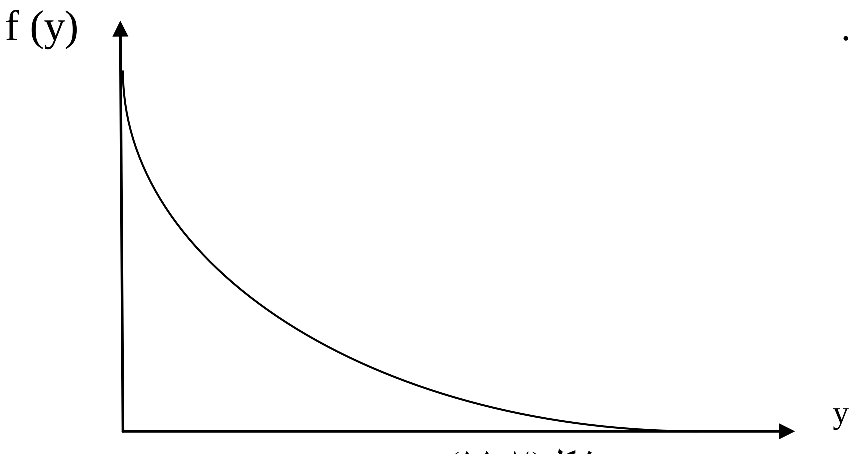
تعريف(٧-١)

إذا كان المتغير العشوائي y يتبع التوزيع الآسي السالب فإن دالة كثافة الاحتمال للمتغير y تأخذ الصياغة التالية:-

$$f(x) = \begin{cases} \lambda e^{-\lambda y} & y \geq 0 \;,\; \lambda \geq 0 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

حيث e هي الأساس الطبيعي ، e =2.71828 ، المعلمة λ تمثل العدد المتوقع لعدد مرات النجاح (أي توقع المتغير y)

والشكل التالي يوضح منحنى دالة كثافة الاحتمال f(y) ويتضح من الشكل أن النهاية العظمى للدالة f(y) تساوي λ عندما y = 0 أي القيمة المنوالية للمتغير y تساوي صفر.



شكل(٧-١١)

وكما في توزيع ذات الحدين وتوزيع بواسون توجد جداول تعطي قيمة دالة الاحتمال أيضاً بالنسبة للتوزيع الآسي توجد جداول تعطي المساحة أسفل المنحنى f(y) من y = 0 إلى $y = \infty$ يساوي قيمة معينة ، أي تعطي احتمال y أقل من أو تساوي القيمة المعينة وملحق رقم (٥) يقدم جزء من هذا الجدول.

نظرية (٧-١):

إذا كان y متغير عشوائي يتبع التوزيع الآسي وله دالة كثافة الاحتمال

$$f(y) = \lambda e^{-\lambda y} \qquad , y \geq 0 \ , \lambda \geq 0$$

فإن

$$\mu_y = \frac{1}{\lambda} \qquad (7-12)$$

$$\sigma_y^2 = \frac{1}{\lambda^2} \qquad (7-13)$$

الإثبات :

$$\mu = \int_0^\infty y\ f(y)\ dy$$

$$= \frac{1}{\lambda} \int_0^\infty (\lambda y)\ e^{-\lambda y} \quad d\lambda y$$

حيث يمثل المقدار $[\int_0^\infty (\lambda y) e^{-\lambda y} \quad dy]$ تكامل دالة كثافة احتمال المتغير (λy) وبالتالي فهو يساوي الواحد "من خصائص دالة كثافة الاحتمال انظر الفصل الأول من هذا الباب

$$\sigma^2 \quad = [\int_0^\infty y^2 f(y)\ dy] - \mu^2 \qquad (7-14)$$

$$= \int_0^\infty y^2 \lambda e^{-\lambda y}\ dy = \int_0^\infty y^2 \quad d(e^{-\lambda y}) \qquad (7-15)$$

$$= -[y^2 e^{-\lambda y} |_0^\infty + \int_0^\infty e^{-\lambda y}\ d(y^2)$$

$$= 0 + 2\int_0^\infty y\ e^{-\lambda y}\ dy$$

$$= \frac{2}{\lambda} \int_0^\infty y \lambda e^{-\lambda y}\ dy = \frac{2}{\lambda}(\frac{1}{\lambda}) = \frac{2}{\lambda^2} \qquad (7-16)$$

وبالتعويض في (7-15) بالطرف الأيسر في (7-16) وقيمة μ في(7-14) فنجد أن:-

$$\sigma^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

مثال (٧-٥)

إذا كانت سفن التفريغ تصل إلى ميناء الإسكندرية وفقاً لتوزيع بواسون بمعدل 5سفن يومياً

المطلوب :

١- أوجد التوزيع الاحتمالي للفترة الزمنية بين وصول سفينتين على التوالي ، ثم أوجد الفترة الزمنية المتوقعة بين وصول سفينتين تفريغ على التوالي.

٢- أوجد احتمال أن تكون الفترة الزمنية بين وصول سفينتي تفريغ على التوالي أقل من 4.8 ساعة ووضح ذلك بيانياً.

٣- أوجد احتمال أن تكون الفترة بين وصول سفينتي تفريغ على التوالي أكبر من 5 ساعات ووضح ذلك بيانياً.

الحل

١- إذا فرضنا أن المتغير العشوائي x يمثل عدد سفن التفريغ التي تصل إلى الميناء كل ساعة فإن $\lambda = \frac{5}{24} \approx 0.21$ سفينة /ساعة ، فإن دالة كثافة الاحتمال للمتغير x تأخذ الصياغة التالية

$$f(x) = \frac{e^{-0.21}(0.21)^{x}}{x!} \qquad , x = 0, 1, 2, 3, ...$$

كذلك إذا فرضنا أن المتغير العشوائي y يمثل الفترة الزمنية بين وصول سفينتي تفريغ ، فإن y متغير يتبع التوزيع الآسي ، أي أن دالة كثافة الاحتمال f(y) على النحو التالي:

$$f(y) = \lambda e^{-\lambda y} = 0.21e^{-0.21y} \qquad , y \geq 0$$

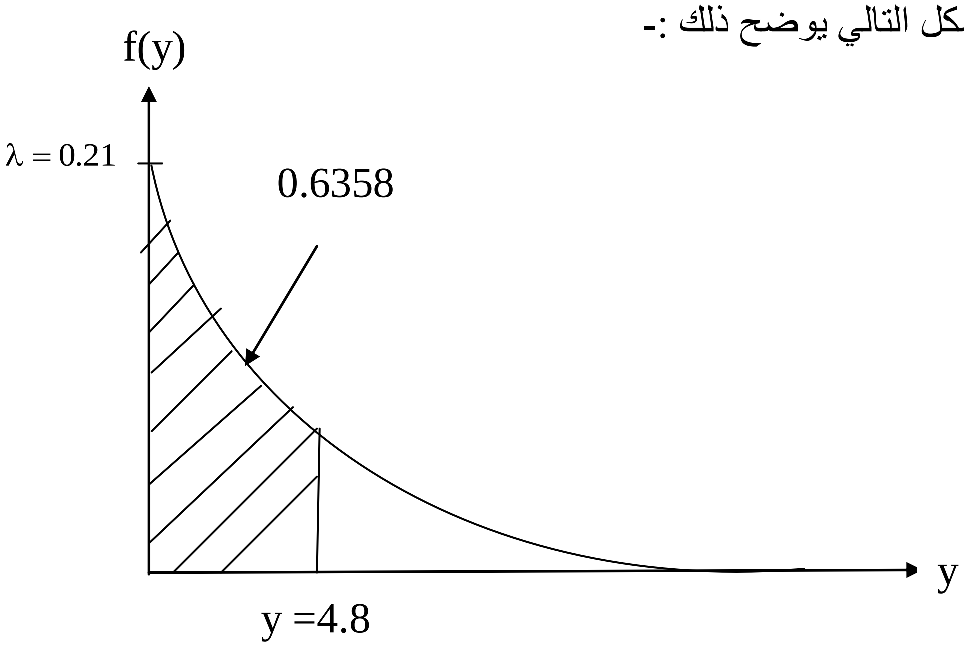
$$\mu = \frac{1}{\lambda} = \frac{1}{0.21} = 4.8$$

وباستخدام ملحق رقم (٥) سوف نقوم بحساب الاحتمالات التالية:

٢- احتمال أن تكون الفترة الزمنية بين وصول سفينتي تفريغ على التوالي أقل من 4.8 ساعة يساوي $P_r(y < 4.8)$ حيث :

$$P_r\ (y < 4.8) = \int_{0}^{\infty} 0.21\, e^{-0.21y}\, dy$$

والشكل التالي يوضح ذلك :-



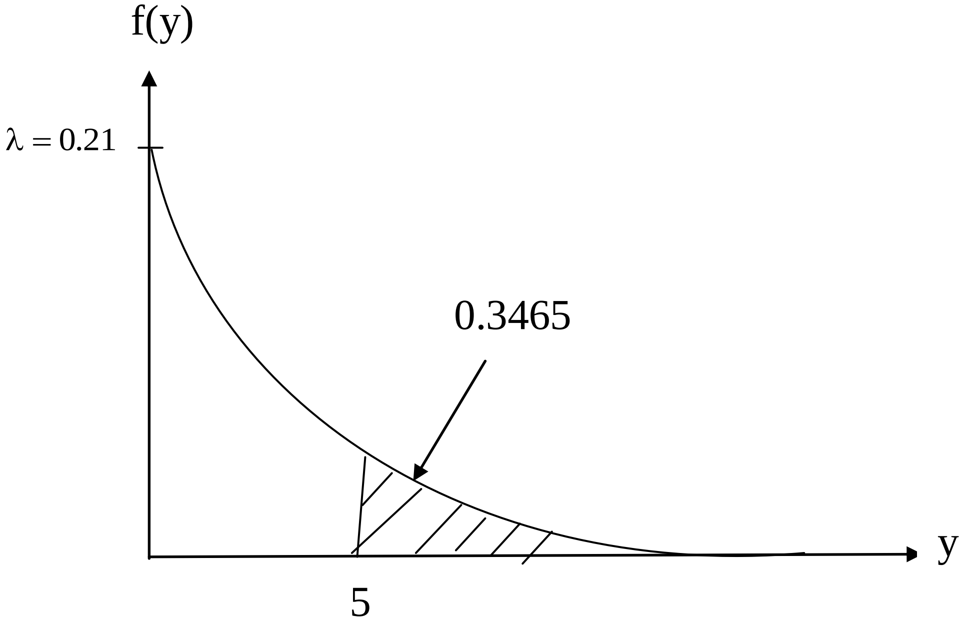
شكل (٧-١٢)

٣- احتمال أن تكون الفترة الزمنية بين وصول سفينتين تفريغ متتاليتين أكبر من 5 ساعات تساوي $P_r(y < 5)$ حيث:

$$P_r(y > 5) = \int_5^{\infty} 0.21 e^{-0.21y} dy$$

$$= 1 - \int_0^5 0.21 e^{-0.21y} dy$$

$$= 1 - 0.6535 = 0.3465$$



شكل (٧-١٣)

مثال (٧-٦)

إذا كـان عـدد المرضـى الـذين يـتم خدمتهم فـي العيـادة الخارجيـة بإحدى المستشفيات الجامعية يمثل متغير يتبع توزيع بواسون بمعدل متوقع 8 مرضى في الساعة.

المطلوب:-

١- أوجد عدد المرضـى المتوقع خدمتهم بالعيادة الخارجيـة فـي يوم عمـل (حيث يوم العمل بالعيادة الخارجية 6 ساعات).

٢- أوجد التوزيع الاحتمالي لزمن خدمة المريض بالعيادة الخارجية.

٣- أوجد احتمال أن يكون زمن خدمة المريض أكبر من 15 دقيقة.

٤- أوجد احتمـال أن يكـون زمـن خدمـة المـريض أقل مـن 5 دقـائق ووضـح ذلـك بيانياً.

الحل

١- بما أن عدد المرضى المتوقع خدمتهم في الساعة λ حيث $\lambda = 8$ مريض / ساعة ، بالتالي عدد المرضى المتوقع خدمتهم في يوم عمل بالعيادة الخارجية

$= \lambda \times 6 = 8 \times 6 = 48$ مريض / يوم

٢- إذا فرضنا أن y يمثل زمن خدمة المريض بالعيـادة الخارجيـة ، وبمـا أن عدد المرضى الذين يتم خدمتهم يتبع توزيع بواسون ، بالتالي فإن زمن الخدمة يتبع التوزيع الآسي ، حيث y تأخذ الشكل التالي:

$$f(y) = \lambda e^{-\lambda y} = 8e^{-8y} \qquad , y \geq 0$$

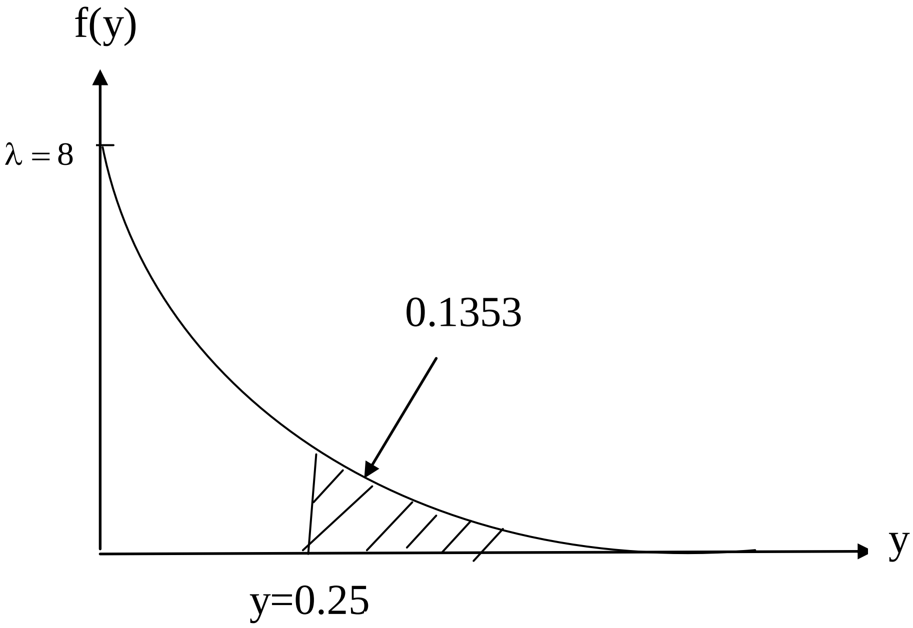
٣- بمـا أن 15دقيقـة تسـاوي (سـاعة $\frac{15}{60} = \frac{1}{4} = 0.25$) ، وبالتـالي فـإن احتمـال أن يكون زمن خدمة المريض أكبر من 15 دقيقة يساوي :

حيث $P_r(y > 0.25)$

$$P_r(y > 0.25) = \int_{0.25}^{\infty} 8e^{-8y} dy$$

$$= 1 - 0.8647 = 0.1353$$

والشكل التالي يوضح ذلك:-

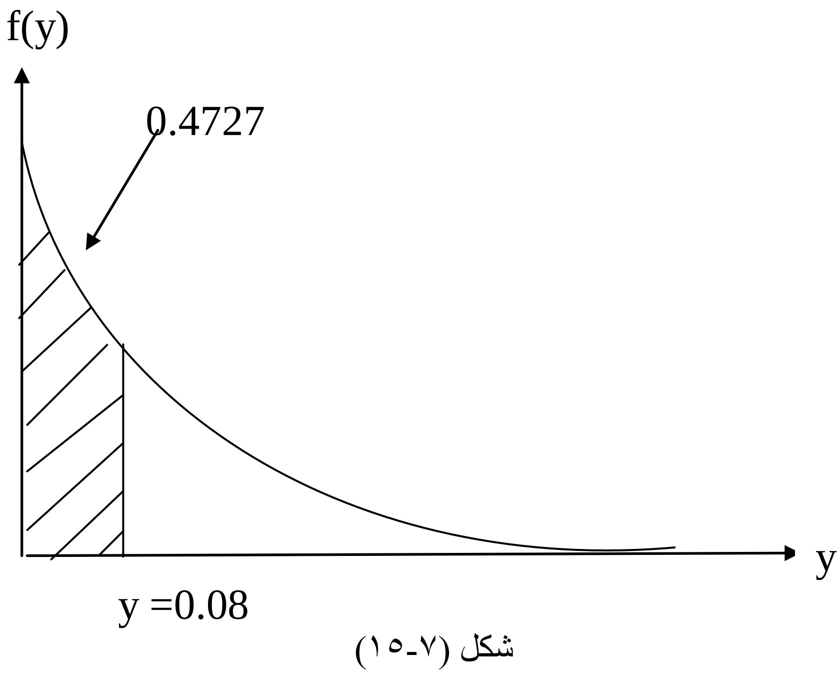


شكل (٧-١٤)

٤- وبالمثل احتمال أن يكون زمن خدمة المريض أقل من 5 دقائق يساوي $P_r(y < 0.08)$ حيث:

$$P_r(y < 0.08) = \int_0^{0.08} 8e^{-8y}dy = 0.4727$$

كما هو موضح بالشكل التالي:



شكل (٧-١٥)

(٧-٣)التوزيع المنتظم المتصل
Continuous Uniform Distribution

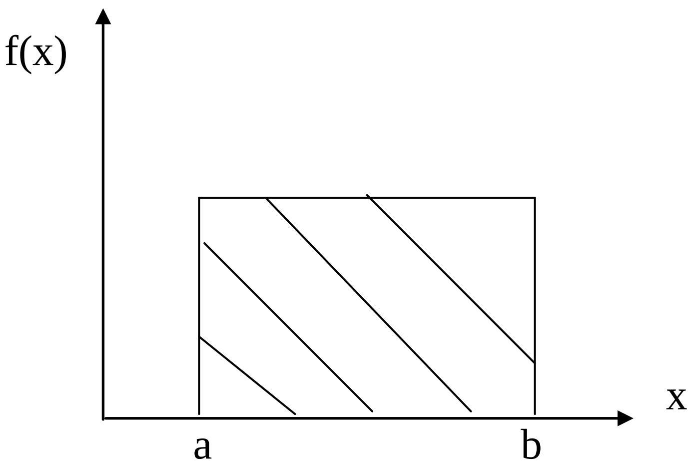
إذا كان المتغير العشوائي x متغير متصل بحيث أن احتمالات وقوع المتغير في فترات متساوية تكون احتمالات متساوية ، أي أن تكون الفرص متساوية لوقوع الحدث في فترات متساوية ، فإن المتغير x يسمى بالمتغير العشوائي المنتظم Uniform Random Variable

تعريف (٧-٢)

إذا كان المتغير x متغير عشوائي منتظم فإن دالة كثافة الاحتمال f(x) تصاغ على النحو التالي:-

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

والشكل التالي يوضح الدالة f(x) :-



شكل (٧-١٦)

مثال (٧-٧)

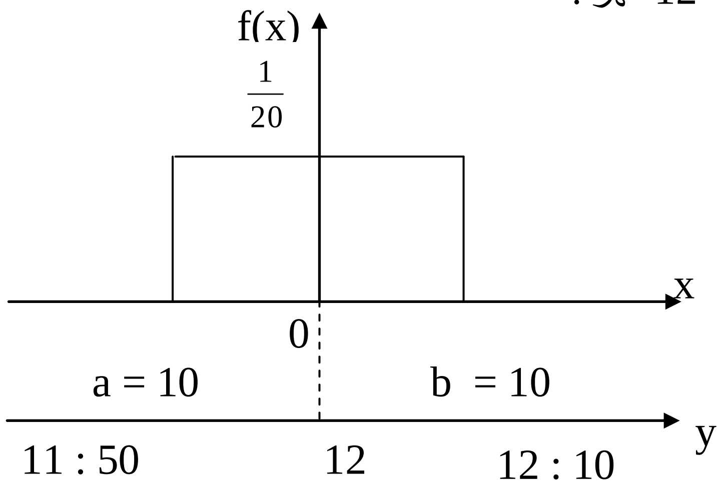
في محطة السكة الحديد برمسيس ووفقاً لجداول تشغيل القطارات على الخط من القاهرة إلى الإسكندرية وجد أنه إذا قام قطار من القاهرة متجهاً إلى الإسكندرية الساعة9 صباحاً فإنه يتوقع وصوله إلى الإسكندرية الساعة 12 ظهراً ولكن لظروف عديدة ممكن أن يصل القطار إلى الإسكندرية قبل الساعة 12 بفترة زمنية لا تزيد عن 10 دقائق أو يصل بعد الساعة 12 بفترة زمنية لا تزيد عن 10 دقائق (أي يوجد 10دقائق تقديم أو تأخير).

المطلوب

١- أوجد التوزيع الاحتمالي لزمن وصول القطار إلى الإسكندرية .
٢- احتمال أن يصل القطار بعد موعده على الأكثر بـ 5 دقائق ووضح ذلك بيانياً.
٣- احتمال أن يصل القطار قبل موعده على الأقل بـ 2 دقائق ووضح ذلك بيانياً.

الحل:-

١- إذا فرضنا أن المتغير y يمثل لحظة وصول القطار إلى الإسكندرية حيث y=12+x والمتغير x يمثل الفرق بين موعد وصول القطار إلى الإسكندرية والساعة12ظهراً.



شكل (٧-١٧)

فنجد أن x=0 إذا وصل القطار 12 ظهرا، كذلك نجد أن المتغير x يتبع التوزيع المنتظم المتصل حيث أن القطار ممكن أن يصل في xأي لحظة في الفترة من 11:50 إلى 12:10 وبالتالي فإن $-10 \le x \le 10$ بحيث دالة كثافة الاحتمال للمتغير x تأخذ الشكل التالي:

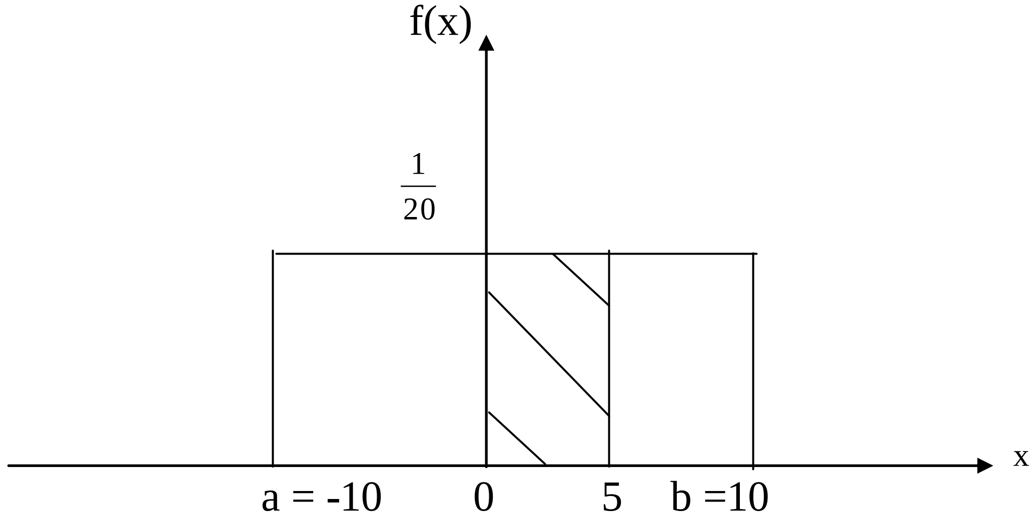
$$f(x) = \frac{1}{b-a} = \frac{1}{10-(-10)} = \frac{1}{20}$$

كما هو موضح بشكل (٧-١٧)

٢-احتمال أن يصل القطار بعد موعده على الأكثر بـ5 دقائق يساوي $P_r(x \le 5)$ حيث :

$$P_r(x \le 5) = \int_0^5 \frac{1}{20}\ dx = \frac{1}{20}[x]_0^5 = \frac{1}{20}[5-0] = \frac{5}{20} = \frac{1}{4}$$

كما هو موضح بالشكل التالي :



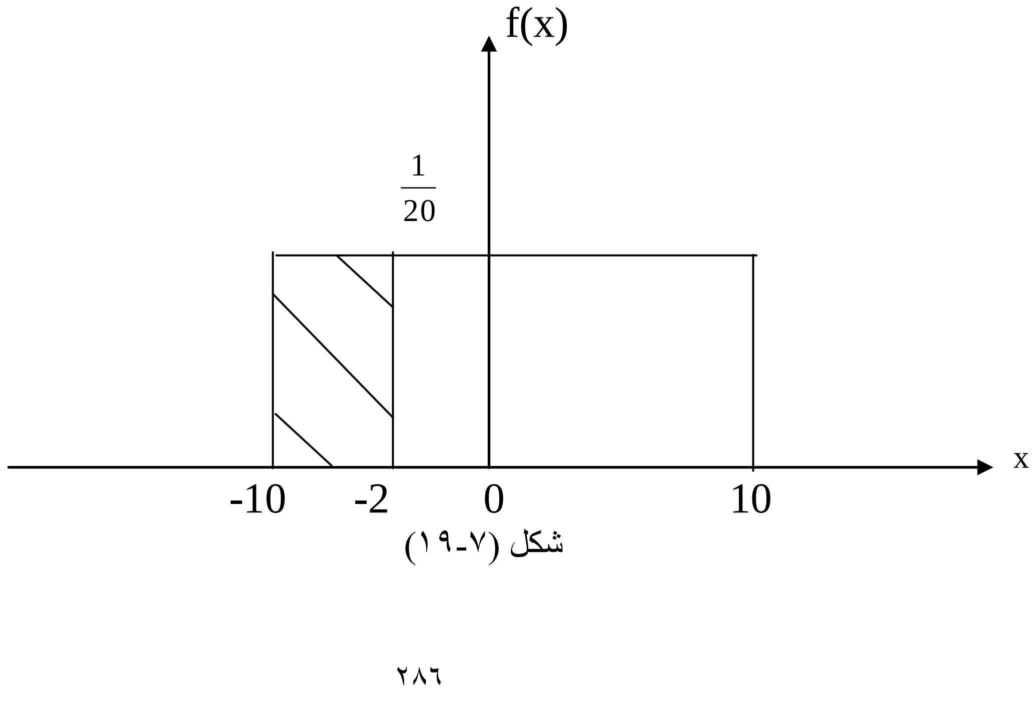
شكل (٧-١٨)

٣- احتمــال أن يصــل القطــار قبــل موعــده علــى الأقــل بــ2 دقيقــة يســاوي $P_r(-10 \le x \le -2)$ بحيث :

$$P_r(-10 \le x \le -2) = \int_{-10}^{-2} \frac{1}{20}\ dx = \frac{1}{20}[x]_{-10}^{-2}$$

$$= \frac{1}{20}[-2+10] = \frac{8}{20} = \frac{2}{5}$$

كما هو موضح بالشكل التالي



شكل (٧-١٩)

نظرية (٧-٢)

إذا كان x متغير منتظم ومتصل وله دالة كثافة الاحتمال f(x) بحيث :

$$f(x) = \frac{1}{b-a} \qquad a \le x \le b$$

$$\mu = \frac{1}{2}(a+b) \qquad (7-17)$$

$$\sigma^2 = \frac{1}{12}(b-a)^2 \qquad (7-18)$$

الإثبات

$$\mu = \int_a^b \frac{x}{b-a} \quad dx = \frac{1}{b-a}[\frac{x^2}{2}]_a^b$$

$$= \frac{1}{b-a}[\frac{b^2-a^2}{2}] = \frac{1}{b-a}[\frac{(b-a)(b+a)}{2}]$$

$$= \frac{b+a}{2} \qquad (7-19)$$

$$\sigma^2 = [\int_a^b x^2 (\frac{1}{b-a}) \quad dx] - \mu^2$$

$$= \frac{1}{b-a}[\frac{x^3}{3}]_a^b - \mu^2$$

$$= \frac{1}{3(b-a)}[b^3 - a^3] - \mu^2$$

$$= \frac{1}{3\{b-a)} - (b-a)(b^2 - ab - a^2) - [\frac{a+b}{2}]^2$$

$$= \frac{(b^2 - ab - a^2)}{3} - \frac{(a^2 + 2ab + b^2)}{4}$$

$$= \frac{4b^2 + 4ab + 4a^2 - 3a^2 - 6ab - 3b^2}{12}$$

$$= \frac{b^2 - a^2 - 2ab}{12} = \frac{(b-a)^2}{12} \qquad (7-20)$$

مثال (٧-٨)

إذا كانت الدالة $f(x) = 2$ بحيث: $5 \le x \le 10$

المطلوب

١- أثبت أن الدالة $f(x)$ دالة كثافة احتمال.

٢- أوجد الاحتمالات التالية :-

$$i)P_r(6 \le x \le 9)$$

$$ii)P_r(4 \le x \le 8)$$

٣- أوجد التوقع والانحراف المعياري للمتغير x.

الحل

١- بما أن

$$f(x) = 0.20>0 \qquad (7\text{-}21)$$

، كذلك

$$\int_5^{10} 0.2 \ dx = 0.2[x]_5^{10} = 0.2[10-5] = 1 \qquad (7\text{-}22)$$

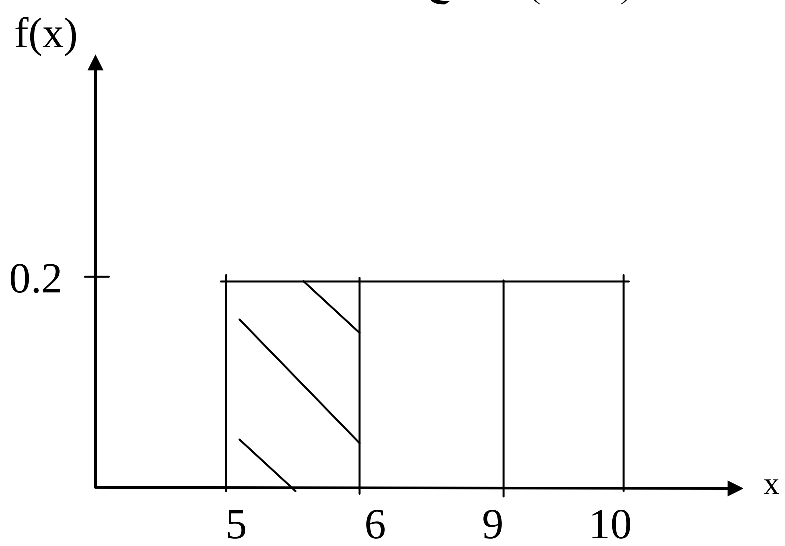
من (7-21) ، (7-22) فإن $f(x)$ تعتبر دالة كثافة احتمال

-٢

$$i) \ P_r(6 \le x \le 9) = \int_6^9 0.2 \ dx = 0.2[x]_6^9$$

$$= 0.2[9-6] = 0.6$$

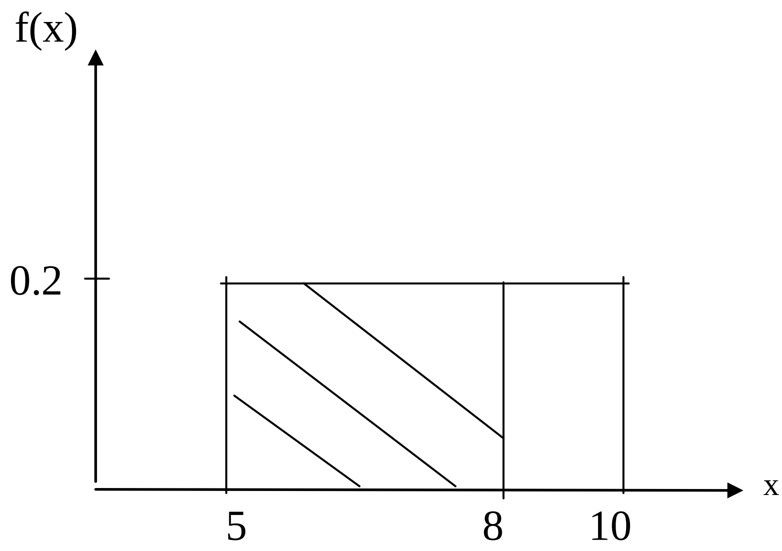
وشكل (7-20) يوضح ذلك



شكل (٧-٢٠)

$$ii)P_r(4 \le x \le 8) = \int_4^8 0.2dx = 0.2[x]_4^8 = 0.2[8-4] = 0.2(3) = 0.6$$

وشكل (7-21) يوضح ذلك



شكل (٧-٢١)

-٣

$$\mu = \frac{1}{2}(a+b) = \frac{1}{2}(5+10) = \frac{15}{2} = 7.5$$

$$\sigma^2 = \frac{1}{12}(a-b)^2 = \frac{1}{12}(10-5)^2 = \frac{25}{12} = 2.08$$

$$\therefore \sigma = \sqrt{2.08} = 1.443$$

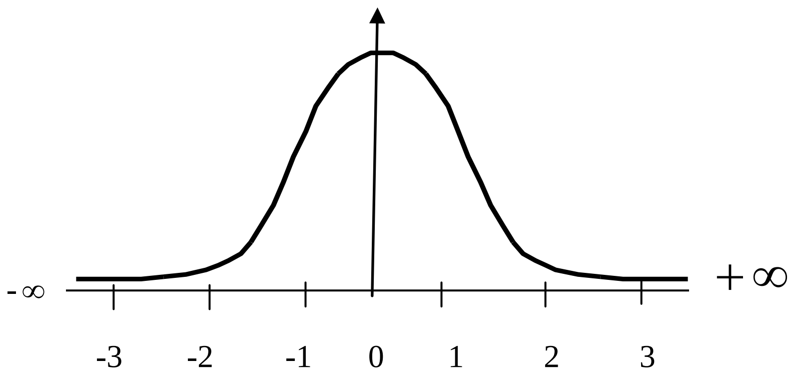
(٧-٤)التوزيع المعتاد القياسي
Standard Normal Distribution

تعريف (٧-٣)

إذا كان x متغير عشوائي متصل له دالة كثافة الاحتمال f(x) حيث :

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2}} & -\infty \le x \le +\infty \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

والشكل التالي يوضح منحنى الدالة f(x):-



شكل (٧-٢٢)

فإنه يقال أن المتغير x يتبع التوزيع المعتاد القياسي.

خصائص الدالة f(x) :-

١- الدالة f(x) دالة متماثلة حول النقطة x = 0 أي أن :

$$a)\ \int_{-\infty}^{0} f(x)\ dx = \int_{0}^{\infty} f(x)\ dx = \frac{1}{2} \qquad (7-23)$$

b) إذا كان j>i فإن

$$\int_{i}^{j} f(x)\ dx = \int_{-j}^{-i} f(x)\ dx \qquad (7-24)$$

٢- القيمة المتوقعة expected Value للمتغير x تساوي القيمة المنوالية Mode تساوي القيمة الوسيطة Median تساوي صفر.

وفي عام 1799 الميلادي قدم عالم الرياضيات والفيزياء الفرنسي كارتن كرامب Chretien Kramp جداول تسمى بجداول التوزيع المعتاد القياسي ، باستخدامها يمكن حساب المساحة المحصورة بين منحنى الدالة f(x) ومحور x عند أي نقطتين للمتغير x حيث تعطي الجداول قيمة $P_r(O \le x \le x_i)$ حيث x تنتمي إلى فئة الأعداد الحقيقية الموجبة فباستخدامها واستخدام خصائص الدالة f(x) السابق ذكرها أعلاه يمكن الحصول على قيمة أي احتمال مطلوب كما سوف يتضح من الأمثلة التالية

وبعد اختراع الحاسب الآلي تم تطوير جداول التوزيع المعتاد القياسي التي قدمها كارتن كرامب "وسوف نقدم جزء من هذه الجداول بملحق رقم (٦)"

نظرية (٧-٣)

إذا كان المتغير x يتبع التوزيع المعتاد القياسي فإن :

$$\mu_x = 0 \qquad (7-25)$$

$$\sigma_x^2 = 1 \qquad (7-26)$$

مثال(٧-٩)

إذا كان x متغير عشوائي يتبع التوزيع المعتاد القياسي ، باستخدام جداول التوزيع المعتاد القياسي" بملحق رقم (٦)" أوجد ما يلي:-

$a)P_r(0 < x < 2)$ $\qquad$ $b)P_r(-2 < x < 2)$

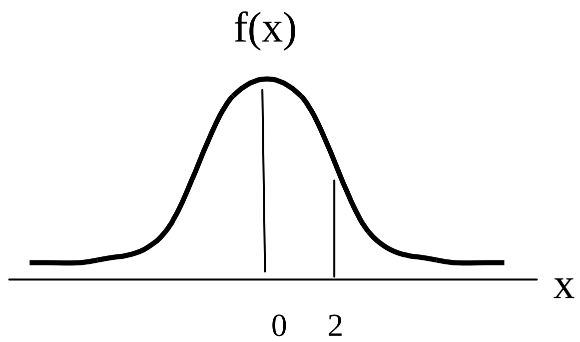
$c)\ P_r(x < 1.51)$ $\qquad$ $d)P_r(-2.5 < x < -1.56)$

$e)P_r(x > 1.53)$ $\qquad$ $f)P_r(x > -1.138)$

الحل

ا- المساحة المظللة في الشكل (٧-٢٣) توضح $P_r(0 < x < 2)$ حيث

$$P_r(0 < x < 2) = 0.4772$$

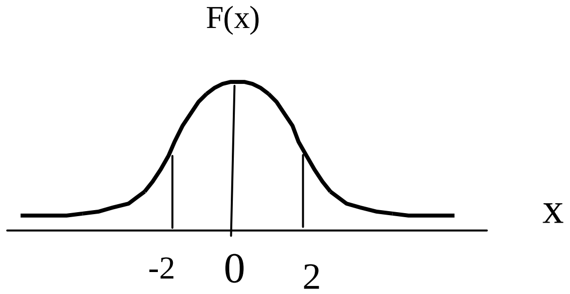


شكل (٧-٢٣)

٢- المساحة المظللة في الشكل (٧-٢٤) توضح $P_r(-2 < x < 2)$ حيث

$$P_r(-2 < x < 2) = 2P_r[(0 < x < 2)]$$

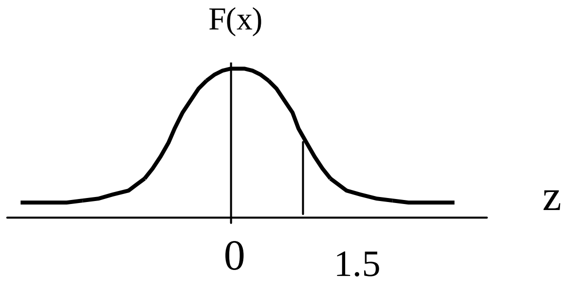
$$= 2(0.4772)0.9544$$



شكل (٧-٢٤)

٣- المساحة المظللة في الشكل (٧-٢٥) توضح $P_r(x < 1.51)$ حيث

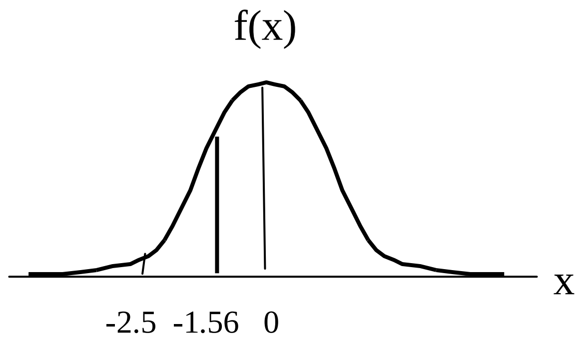
$$P_r(x < 1.51) = 0.5 + P_r(0 < x < 1.51)$$
$$= 0.5 + 0.4345 = 0.9345$$



شكل (٧-٢٥)

٤-المساحة المظللة في الشكل (٧-٢٦) توضح $P_r(-2.5 < x < -1.56)$ حيث

$$P_r(-2.5 < x < 1.56) = P_r(1.56 < x < 2.5)$$
$$= 0.4938 - 0.4406 = 0.0532$$

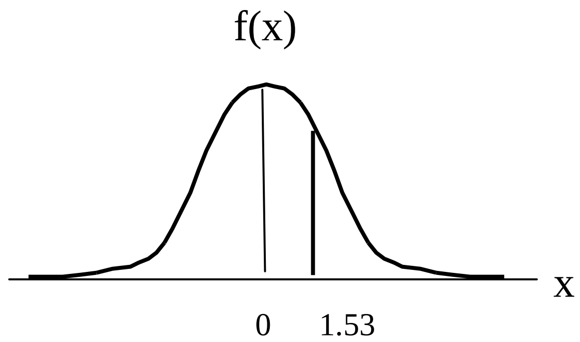


شكل (٧-٢٦)

٢

٥- المساحة المظللة في الشكل (٧-٢٧) توضح $P_r(x > 1.53)$ حيث

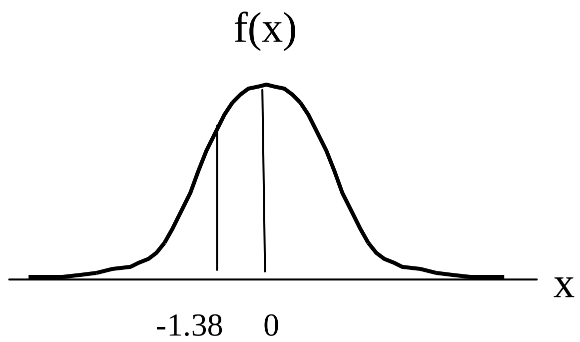
$$P_r(x > 1.53) = 0.5 - P_r(0 < x < 1.51)$$
$$= 0.5 + 0.4370 = 0.0630$$



شكل (٧-٢٧)

٦ - المساحة المظللة في الشكل (٧-٢٨) توضح $P_r(x > -1.38)$ حيث

$$P_r(x > -1.38) = 0.5 + P_r(0 < x < 1.38)$$
$$= 0.5 + 0.4262 = 0.9162$$



شكل (٧-٢٨)

(٧-٥)التوزيع المعتاد (الطبيعي)
Normal Distribution

يعتبر التوزيع المعتاد (الطبيعي)الذي سوف نقوم بدراسته في هذا الفصل من أهم التوزيعات الاحتمالية المتصلة التي سوف نقوم بدراستها للأسباب التالية:

١- كثير من المشاكل الفعلية تتضمن متغيرات تتبع التوزيع المعتاد.
٢- كثير من الأساليب الإحصائية في التحليل الإحصائي تعتمد علي هذا التوزيع مثل التقديرات واختبارات الفروض.
٣ – كثير من التوزيعات الاحتمالية مثل توزيع بواسون, أو توزيع ذات الحدين...الخ,يمكن أن يستخدم التوزيع المعتاد كتقريب جيد لهذه التوزيعات.

تعريف (٧-٤)
إذا كان x متغير عشوائي متصل له دالة كثافة الاحتمال $f(x)$ بحيث

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & -\infty \le x \le +\infty \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

فانه يقال أن المتغير x يتبع التوزيع الطبيعي بتوقع μ, وتباين σ^2 والشكل التالي يوضح دالة كثافة الاحتمال للمتغير المعتاد. ونلاحظ تزايد التفرطح للدالة $f(x)$ كلما زادت قيمة الانحراف المعياري σ او بعبارة أخري يزداد تدبب المنحني $f(x)$ كلما نقصت قيمة σ والشكل التالي يوضح ذلك:

شكل (٧-٢٩)

خصائص الدالة $f(x)$:

١- دالة كثافة الاحتمال للمتغير المعتاد $f(x)$ دالة متماثلة حول النقطة $\mu = x$ أي أن:

$$\text{i)}\ \int_{-\infty}^{\mu} f(x)\,dx = \int_{\mu}^{\infty} f(x)\,dx = \frac{1}{2} \qquad (7-27)$$

$$\text{ii)}\ \int_{j}^{i} f(x)\,dx = \int_{-i}^{-j} f(x)\,dx \qquad (7-28)$$

٣- القيمة المتوقعة = القيمة المتوالية = القيمة الوسيطة

وأول من قدم التوزيع المعتاد عام ١٧٣٣ ميلادية عالم الرياضيات الفرنسي ابراهام دي موافر Abraham de moivre, وبعده قام العالم الفرنسي ادولف كيتيليه adolphe quetelet(١٧٩٦-١٨٧٤) بدراسة وتعميم أهم خصائص التوزيع .ولكن يعتبر عالم الفيزياء كارل جاوس Carl Gauss(١٧٧٧-١٨٥٥) أول من استخدم هذا التوزيع في الدراسات التطبيقية في العلوم الطبيعية لذا يسمي أحياناً التوزيع المعتاد بالتوزيع الطبيعي أو توزيع الملكية جاوس , فقد استخدم جاوس التوزيع المعتاد في وصف الأخطاء في القياسات astronomy وفي دراسة كثير من الظواهر الطبيعية الاخري.

نظرية (٧-٤):

إذا كان x متغير معتاد توقعه μ وتباينه σ^2 فإذا كان المتغير العشوائي بحيث:

$$z = \frac{x-\mu}{\sigma} \qquad -\infty \le z \le +\infty \qquad (7-29)$$

فان المتغير z يتبع التوزيع المعتاد القياسي بدالة احتمال $f(z)$ حيث :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

فباستخدام هذه النظرية يمكن تحويل المتغير المعتاد إلي متغير معتاد قياسي وتعتبر هذه النظرية علي جانب كبير من الأهمية فباستخدام هذه النظرية يمكن حساب المساحات المختلفة تحت المنحني المعتاد باستخدام المساحات المناظرة تحت المنحني المعتاد القياسي والتي يمكن حسابها مباشرة باستخدام التوزيع المعتاد القياسي بملحق () وسوف نوضح عملية التحويل باستخدام هذه النظرية من خلال الأمثلة التالية :

مثال(٧-١٠):

إذا كان x متغير معتاد توقعه يساوي5 $\mu =$ وتباينه يساوي16 $\sigma^2 =$. احسب الاحتمالات التالية باستخدام جداول التوزيع المعتاد القياسي بملحق رقم (٦):

$$a)\ P_r(x < 3)$$
$$b)\ P_r(x < 11)$$
$$c)\ P_r(10 < x < 12)$$
$$d)\ P_r(x > 9)$$

الحل:

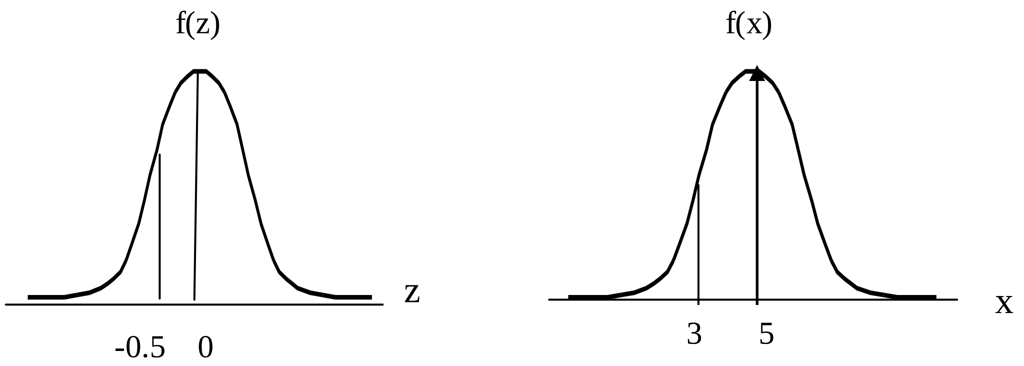
إذا فرضنا أن σ^2 متغير معتاد قياسي له دالة كثافة الاحتمال$f(z)$ حيث :

$$z = \frac{x-\mu}{\sigma} = \frac{x-5}{4}$$

فإن:

$$a) P_r(x < 3) = P_r(\frac{x-5}{4} - \frac{3-5}{4}) = P_r(z < -0.5)$$

والشكل التالي يوضح المساحة تحت المنحني $f(x)$ التي تمثل الاحتمال المطلوب والمساحة المساوية لها تحت المنحني $f(z)$.



شكل (٧-٣٠)

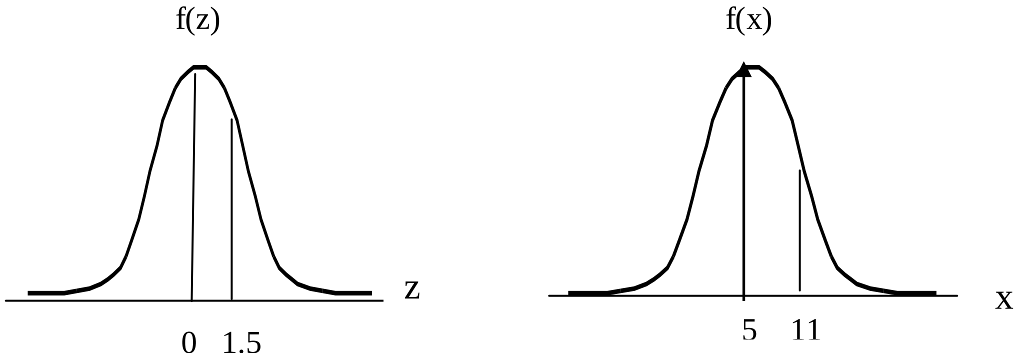
ومن جدول التوزيع المعتاد القياس نجد أن :

$$P_r(x < 3) = P_r(z < -0.5) = 0.5 - P_r(0 < z < 0.5)$$
$$= 0.5 - 0.1544 = 0.3446$$

بالمثل:

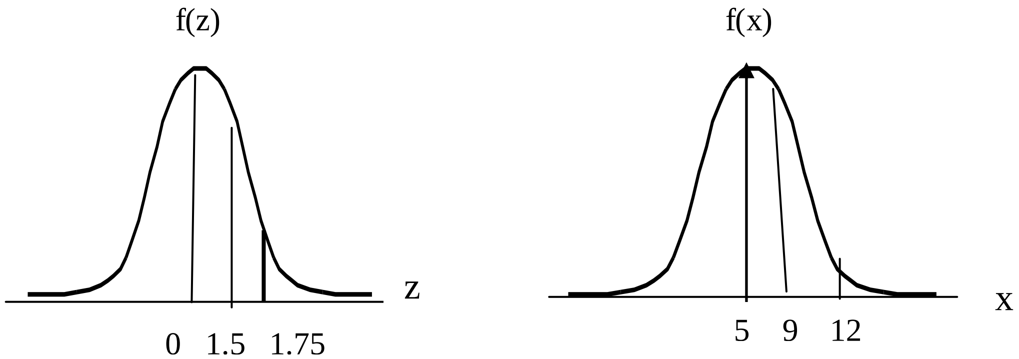
$$b) P_r(x < 11) = P_r(z < \frac{11-5}{4}) = P_r(z < 1.5)$$
$$= 0.5 + 0.4332 = 0.9332$$

والشكل التالي يوضح المساحة تحت المنحني f(x) التي تمثل الاحتمال المطلوب والمساحة المساوية لها تحت المنحني f(z) :



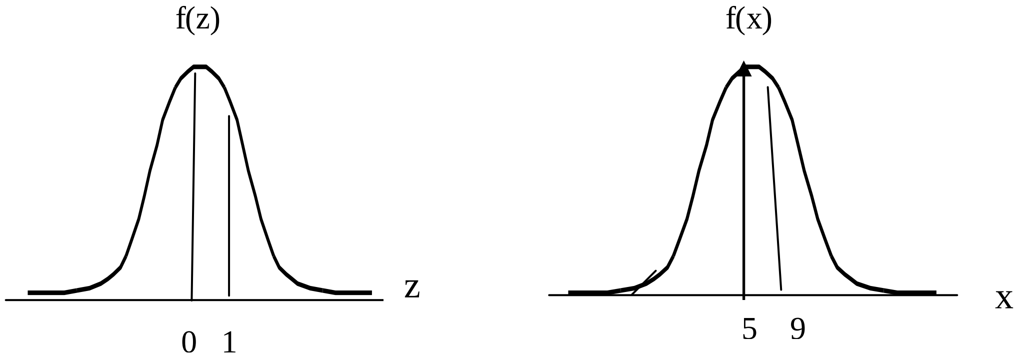
شكل (٧-٣١)

$$\text{c)}P_r(10 < x < 12) = P_r(1.25 < z < 1.75)$$
$$= P_r(0 < z < 1.75) - P_r(0 < z < 1.25)$$
$$= 0.4599 - 0.3944 = 0.655$$



شكل (٧-٣٣)

$$\text{d)}P_r(x > 9) = P_r(z > 1)$$
$$= 0.5 - P_r(0 < z < 1)$$
$$= 0.5 - 0.3413 = 0.1587$$



شكل (٧-٣٣)

مثال (٧-١١)

قامت إحدى شركات تصنيع التليفزيونات بدراسة فترة الصلاحية للتليفزيون الذي تقوم الشركة بتصنيعه ، فوجدت أن الفترة العمرية للتليفزيونات تتبع التوزيع المعتاد بتوقع 10سنوات وانحراف معياري سنتين.

المطلوب

١- ما هو احتمال أن تكون الفترة الصلاحية للتلفزيون من إنتاج هذه الشركة أكبر من 10سنوات

٢- ما هو احتمال أن تكون فترة الصلاحية للتلفزيون من إنتاج هذه الشركة أقل من 6 سنوات

الحل

إذا فرضنا أن المتغير x تمثل فترة الصلاحية للتلفزيون ، كذلك

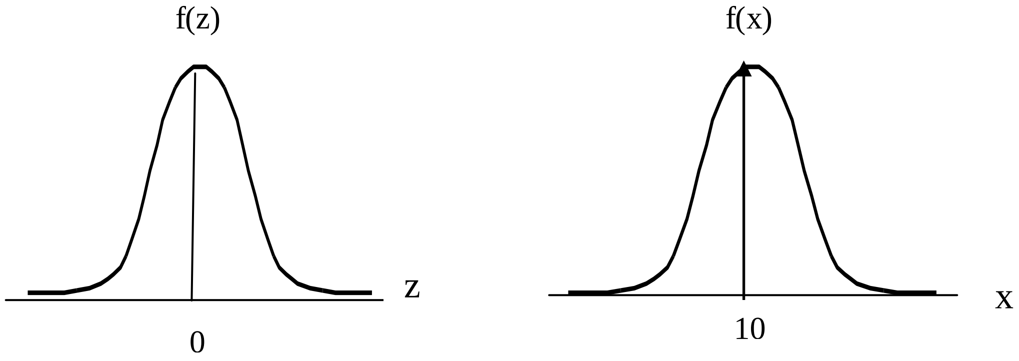
$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{x - 10}{2}$$

١

-احتمال أن تكون فترة الصلاحية للتلفزيون أكبر من 10سنوات يساوي:

$$P_r(x > 10) = P_r(z > \frac{10 - 10}{2}) = P_r(z > 0) = 0.5$$



شكل (٧- ٣٤)

مثال(٧-١٢)

في إحدى السنوات الدراسية تقدم للالتحاق بكلية التربية الرياضية 3000 طالب وكانت أطوالهم موزعة توزيعا معتادا بتوقع 160 سم وانحراف معياري 8 سم , فإذا كانت الكلية تقبل الطلاب الذين تتراوح أطوالهم بين 150-180سم ,

المطلوب :

١-تحديد نسبة وعدد الطلاب المتوقع أن تزيد أطوالهم عن 170سم.

٢-تحديد نسبة الطلاب الذين تتراوح أطوالهم بين 175-195 سم , وتحديد عددهم المتوقع.

٣- تحديد نسبة الطلاب الذين تقل أطوالهم عن 155 سم , وتحدد عددهم المتوقع .

الحل :

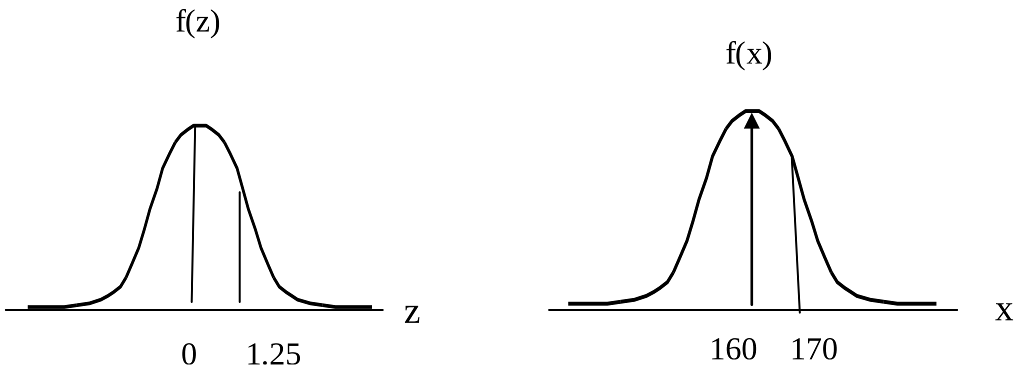
إذا كان x يمثل طول الطالب , $z = \frac{x-160}{8}$ فإن :

١- احتمال أن يكون الطالب يزيد عن 170سم يساوي

$$P_r(x > 170) = P_r(z > \frac{170-160}{8}) = P_r(z > 1.25)$$

$$= 0.5 - P_{r(}(0 < z < 1.25)$$

$$= 0.5 - 0.3944 = 0.1056$$



شكل (٧-٣٥)

وبالتالي تكون نسبة الطلاب المتوقع أن تزيد أطوالهم عن 170 سم

$$= 0.1056 \times 100 = 10.56\%$$

عدد الطلاب المتوقع أن تكون أطوالهم أزيد من 170سم

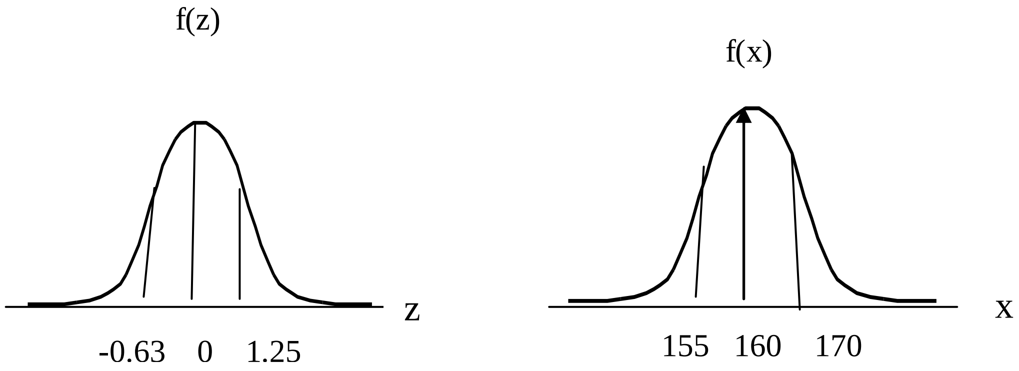
= العدد الكلي للطلاب المتقدمين × نسبة الطلاب الذين تزيد أطوالهم عن 170سم

$$= 3000 \times \frac{10.56}{100} = 316.80 \approx 317$$ طالب

بالمثل :

٢-احتمال أن يتراوح طول الطالب بين 170-155 سم يساوي $P_r(155 < x < 170)$

حيث :



شكل (٧-٣٦)

$$P_r(155 < x < 170) = P_r\left(\frac{155-160}{8} < z < \frac{160-170}{8}\right)$$
$$= P_r(-0.63 < z < 0) + P_r(0 < z1.5)$$
$$= P_r(0 < z < 0.63) + P_r(0 < z1.5)$$
$$= 0.2357 + 0.3944 = 0.6301$$

كما هو موضح في شكل (٧-٣٦)

وبالتالي نسبة هؤلاء الطلاب

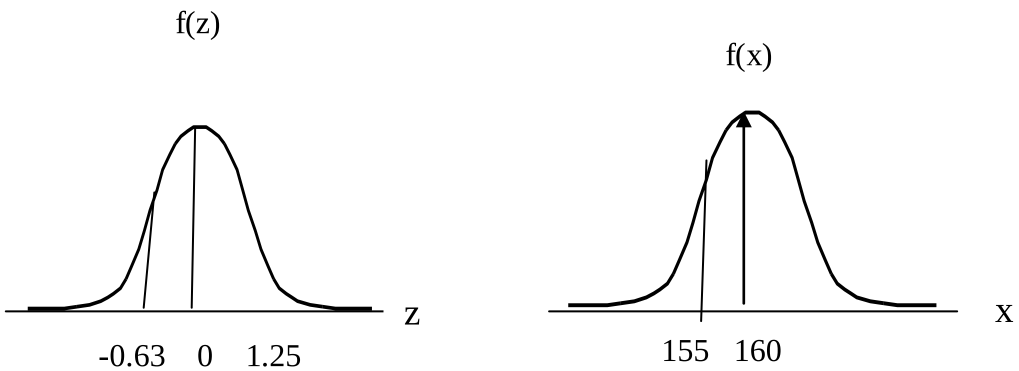
$$= 0.6301 \times 100 = 63.01\%$$

وعددهم المتوقع

$$= 3000 \times \frac{63.01}{100} = 1890.3 = 1890$$ طالب

٣- احتمال أن يقل طول المتقدم عن 155 سم يساوي $P_r(x < 155)$ حيث:

$$P_r(x < 155) = P_r\left(z < \frac{155-160}{8}\right) = P_r(z < -0.63)$$
$$= 0.5 - P_r(0 < z < 0.63) = 0.2634$$



شكل (٧-٣٧)

وبالتالي نسبة هؤلاء الطلاب

$$= 0.2643 \times 100 = 26.43\%$$

وعددهم المتوقع

$$= 3000 \times \frac{26.43}{100} = 792.9 \approx 793$$ طالب

كما هو موضح بشكل (٧-٣٧)

التوزيع المعتاد كتقريب جيد لتوزيعات احتمالية أخرى :

يمكن استخدام التوزيع المعتاد كتقريب جيد لبعض التوزيعات الاحتمالية المتقطعة والمتصلة ، فعلى سبيل المثال:

١- يمكن استخدام التوزيع المعتاد كتقريب لتوزيع ذات الحدين (أنظر الفصل ٥-٢) عندما تكون عدد المحاولات n كبيرة ، يمكن إثبات ذلك رياضيا حيث أن دالة الاحتمال لتوزيع ذات الحدين تؤول إلى دالة كثافة الاحتمال للتوزيع المعتاد عندما $n \to \infty , P = \frac{1}{2}$

٢- يمكن استخدام التوزيع المعتاد كتقريب للتوزيع الهيبروجومترك [انظر الفصل (٦-٤)] ، ويمكن إثبات ذلك رياضياً حيث أن دالة الاحتمال للتوزيع الهيبروجومترك تؤول إلى دالة كثافة الاحتمال للتوزيع المعتاد عندما $\frac{n_1}{n} \to \frac{1}{2} \quad , n \to \infty$

٣- يمكن استخدام التوزيع المعتاد كتقريب لتوزيع بواسون* ويمكن إثبات ذلك رياضياً حيث أن دالة كثافة الاحتمال لتوزيع بواسون تؤول إلى دالة كثافة

٤-

٥- الاحتمال للتوزيع المعتاد عندما $\lambda \longrightarrow np$

مثال(٧-١٣)

إذا كان المتغير x يتبع توزيع ذات الحدين حيث $n = 20 , p = 0.5$

المطلوب:

١- $P_r(x < 15), P_r(x < 14), P_r(x < 13)$

٢- استخدم التوزيع المعتاد كتقريب لتوزيع ذات الحدين ، وحدد كل من μ , σ

٣- استخدم القيم التقريبية للاحتمالات في(١) باستخدام (٢)

الحل:

من جداول توزيع ذات الحدين (بملحق رقم) نجد أن قيمة الاحتمالات المطلوبة عندما $n = 20 , p = 0.5$ كما هو موضح بالجدول التالي:

جدول (٧-١)

القيمة التقريبية باستخدام التوزيع المعتاد	القيمة الفعلية باستخدام توزيع ذات الحدين	الاحتمالات
$0.9971 \approx 0.99$	$0.9941 \approx 0.99$	$P_r(x < 15)$
$0.9633 \approx 0.96$	$0.9763 \approx 0.98$	$P_r(x < 14)$
$0.9236 \approx 0.92$	$0.9423 \approx 0.94$	$P_r(x < 13)$

* *نفس المرجع في صفحة( )

٢-إذا استخدمنا التوزيع المعتاد كتقريب لتوزيع ذات الحدين فإن:

$$\mu = np = 20(0.5) = 10$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{20(0.5)(0.5)} = 2.24$$

فإذا كانت y متغير معتاد قياسي فإن:-

$$P_r(x < 15) = P_r(z < \frac{15-10}{2.24}) = P_r(z < 2.23)$$

$$= 0.5 + P_r(0 < z < 2.23)$$

$$= 0.9871$$

وبالمثل

$$P_r(x < 14) = P_r(z < \frac{14-10}{2.24}) = P_r(z < 1.79)$$

$$= 0.5 + 0.4633 = 0.9633$$

$$P_r(x < 13) = P_r(z < \frac{13-10}{2.24}) = P_r(z < 1.34) = 0.9236$$

ملحوظة : يلاحظ أنه كلما

٣ زادت قيمة x تقترب القيمة الفعلية للاحتمال من القيمة التقريبية

(٧-٦) استخدام التوزيع المعتاد كتقريب لتوزيع ذي الحدين

Approximation A Binomial Distribution With A Normal Distribution

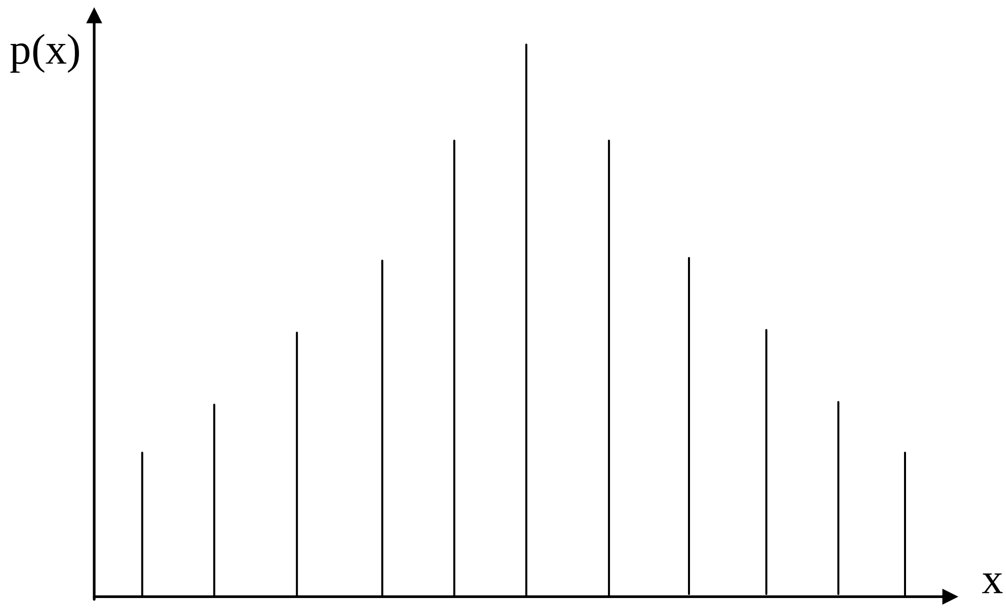
يستخدم التوزيع الطبيعي كتقريب للعديد من التوزيعات الاحتمالية الأخرى المتقطعة والمتصلة ؛ وقد يكون هذا التقريب جيد في بعض الحالات وغير ذلك في حالات أخرى .

وفي هذا الفصل سوف نوضح كيفية استخدام التوزيع المعتاد كتقريب لتوزيع ذات الحدين .

من الفصل (٦-٣) بالباب السادس – إذا فرضنا أن x_c متغير يتبع توزيع ذات الحدين فإن

$$E(x) = np \qquad Var(x) = np(1-p)$$

وعندما p=0.5 أو $p \to 0.5$ نجد أن التوزيع متماثل كما هو في الشكل التالي:-



شكل (٧-٣٨)

ويقترب شكل التوزيع الاحتمالي لتوزيع بواسون من التوزيع المعتاد كلما زاد حجم العينة وكلما اقترب احتمال النجاح p من 0.5 أي عندما $p \to 0.5$

ويعتبر التوزيع المعتاد تقريب جيد لتوزيع ذات الحدين عندما $np(1-p)>q$ ويعتبر غير ذلك إذا كان $np(1-p)<q$ وسوف نوضح ذلك في الأمثلة التالية.

ونظراً لأن متغير ذات الحدين متغير متقطع والمتغير المعتاد متغير متصل في حالة إذا كان عدد المحاولات عدد كبير فإنه يمكن استخدام معامل تصحيح الاتصال * Continuity Correction Factor وذلك بطرح 0.5 من الحد الأدنى للمتغير وإضافة 0.5 للحد الأعلى للمتغير ويعتبر استخدام التوزيع الطبيعي كتقريب لتوزيع ذات الحدين عندما يكون عدد المحاولات n كبيرة وغير موجودة بالجداول فإن جداول التوزيع المعتاد القياسي تستخدم في هذه الحالات ، وسوف نوضح ذلك من خلال المثال التالي.

مثال (٧-١٤)

إذا فرضنا أن xمتغير يتبع توزيع ذات الحدين بحيث $n = 40$, $p = 0.5$ فإن :-

$$E(x) = np = 40\,(0.5) = 20$$

$$Var\,(x) = np(1-p) = 40(0.5)(0.5) = 10$$

*Paul Newbold , Willam L. Carlson ;&Betty M . Thorne(2003) "Statistical For Business & Economic"

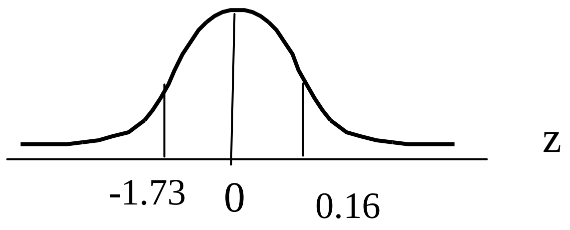
أوجد:-

$$1)P_r(15 \le x \le 20)$$
$$2)P_r(x \ge 20)$$
$$3)P_r(x \le 18)$$

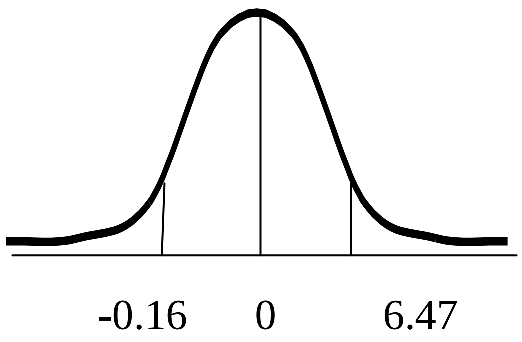
الحل:-

أولاً : لابد من أن نستخدم معامل تصحيح الاتصال بإضافة 0.5 للحد الأعلى وطرح 0.5 من الحد الأدنى

$$\begin{aligned} P_r(15 \le x \le 20) &= P_r(15-0.5 \le x \le 20+0.5) \\ &= P_r(14.5 \le x \le 20.5) \\ &= P_r\left(\frac{14.5-20}{\sqrt{10}} \le \frac{x-e(x)}{\sqrt{Var(x)}} \le \frac{20.5-20}{\sqrt{10}}\right) \\ &= P_r\left(\frac{-5.5}{3.17} \le z \le \frac{0.5}{3.17}\right) \\ &= P_r(-1.73 \le z \le 0.16) \end{aligned}$$



$$
\begin{aligned}
P_r(x \geq 20) = P_r(20 \leq x \leq 40) \\
&= P_r(20 - 0.5 \leq x \leq 40 + 0.5) \\
&= P_r(19.5 \leq x \leq 40.5) \\
&= P_r\left(\frac{19.5 - 20}{3.17} \leq z \leq \frac{40.5 - 20}{3.17}\right) \\
&= P_r\left(\frac{0.5}{3.17} \leq z \leq \frac{20.5}{3.17}\right) \\
&= P_r(-0.16 \leq z \leq 6.47) \\
&= P_r(0 \leq z \leq 0.16) + 0..5 \\
&= 0.0636 + 0.5 = 0.5636
\end{aligned}
$$



(٧-٧) أمثلة تطبيقية Applied Examples

تطبيق (٦-١):-

وفقا لجداول إقلاع ووصول الطائرات من والي مطار القاهرة الدولي , إذا أقلعت الطائرة من القاهرة الساعة 10 صباحاَ فإنها تصل إلي مطار أسوان الساعة 12 ظهراَ ,ولكن لعديد من الظروف الممكن حدوثها فإنه من الممكن أن تصل الطائرة إلي مطار أسوان قبل أو بعد الساعة 12 ظهرا في 15 دقيقة علي الأكثر (أي يوجد 15 دقيقة تقديم أو تأخير).

المطلوب:-

١- أوجد التوزيع الاحتمالي لزمن وصول الطائرة إلي مطار أسوان إذا أقلعت من القاهرة الساعة 10 صباحا .

٢- أوجد احتمال أن تصل الطائرة قبل موعدها بـ10 دقائق أي تصل الساعة 11:50.

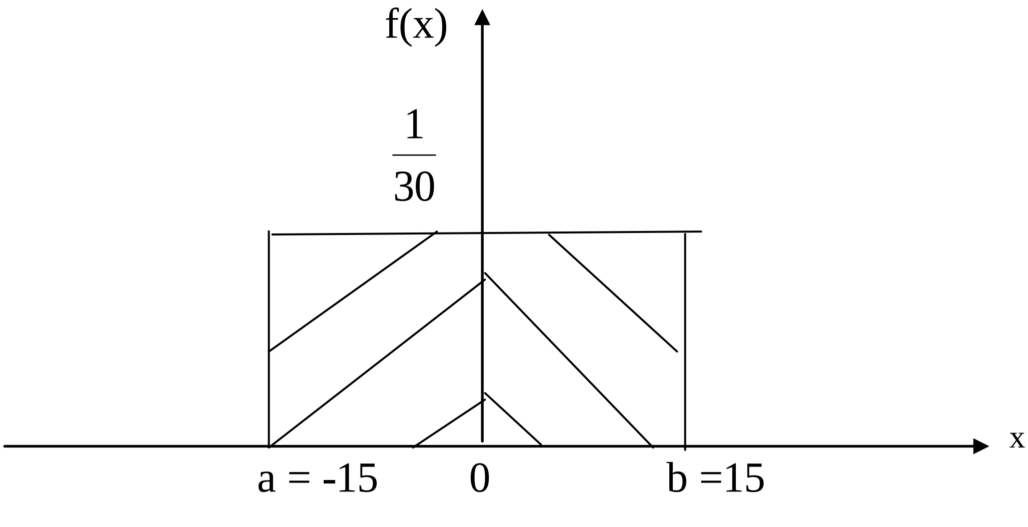
٣- احتمال أن تصل الطائرة بعد موعدها ب ١٥ دقيقة أي تصل الساعة 12:15 .

الحل :-

إذا فرضنا أن المتغير x يمثل لحظة وصول مطار أسوان فإن $y = x+12$, حيث المتغيرx يمثل الفرق بين موعد وصول الطائرة إلي مطار أسوان و الساعة 12 ظهرا . فنجد أن $x=0$ إذا وصلت الطائرة الساعة 12 ظهرا , كذلك نجد أن المتغير x يتبع التوزيع المنتظم حيث أن الطائرة ممكن أن تصل في أي لحظة في الفترة من 11:45 إلي 12:15 , وبالتالي فإن $-15 \le x \le 15$ بحيث أن دالة كثافة الاحتمال للمتغير x تأخذ الصياغة التالية :

$$f(x) = \frac{1}{b-a} = \frac{1}{15-(15)} = \frac{1}{30}$$

كما هو موضح بالشكل التالي :

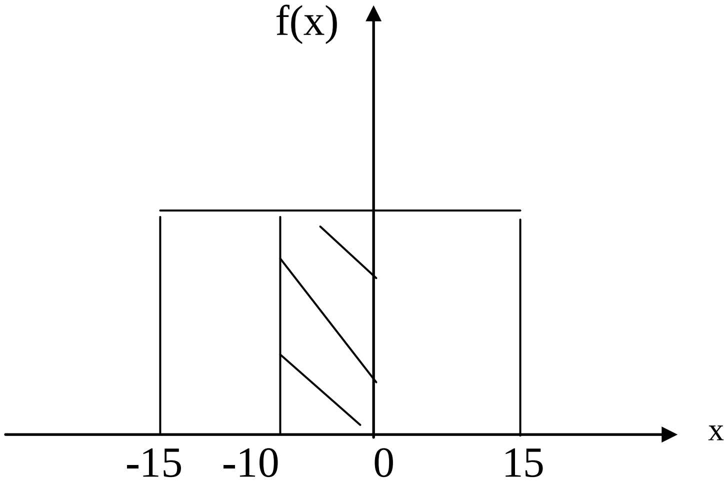


٢- احتمال أن تصل الطائرة قبل موعدها ب10 دقائق

$$P_r(11:50 \le y \le 12) = P_r(-10 \le x \le 0)$$

$$= \int_{-10}^{0} \frac{1}{30}\,dx = \frac{1}{30}[x]_{-10}^{0}$$

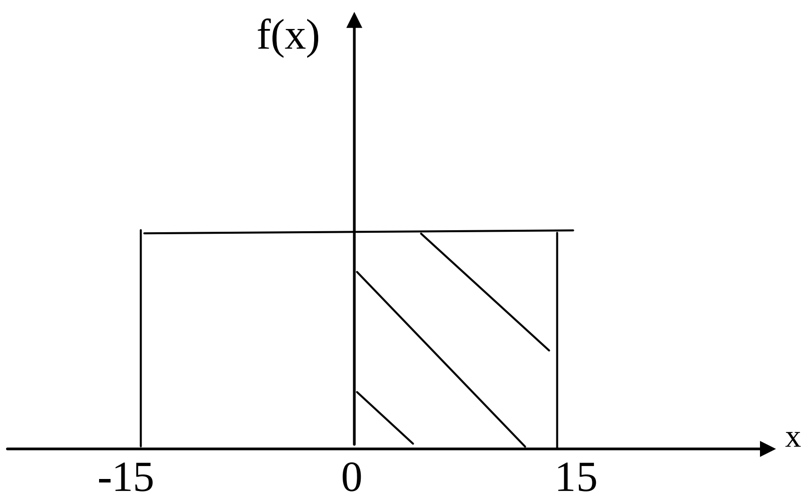
$$= \frac{1}{30}[0-(-10)] = \frac{10}{30}$$



٣- احتمال أن تصل الطائرة بعد موعدها بـ15 دقيقة

$$P_r(12 \le y \le 12:15) = P_r(0 \le x \le 15)$$

$$= \frac{1}{30}(15-0) = \frac{15}{30} = \frac{1}{2}$$



تطبيق (٧-٢):

تقوم إحدى شركات إصلاح الكمبيوتر بجمهورية مصر العربية بدراسة الطلب علي الخدمات التي تقوم الشركة بتقديمها حتي تتمكن من توفير الخبرات والمستلزمات التي تغطي هذا الطلب في السوق , فإذا أثبتت الدراسة أن الطلبات علي خدمات الشركة تصل وفقا لتوزيع يواسون بتوقع 40 طلب في يوم العمل (10 ساعات يوميا).

المطلوب:

١- أوجد التوزيع الاحتمالي لعدد الطلبات التي تصل إلي الشركة في الساعة .
٢- أوجد احتمال عدم وصول أي طلب في الساعة .
٣- أوجد احتمال وصول أكثر من طلب في الساعة .
٤- أوجد التوزيع الاحتمالي للفترة الزمنية بين وصول طلبين إلي الشركة , ثم أوجد توقع وتباين هذه الفترة .
٥- أوجد احتمال أن تكون الفترة الزمنية بين وصول طلبين إلي الشركة , ثم أوجد توقع وتباين هذه الفترة ..
٦- أوجد احتمال أن تكون الفترة الزمنية بين وصول طلبين علي التوالي أقل من 15 دقيقة.

الحل:-

إذا فرضنا أن x , y متغيران عشوائيين يمثلان عدد الطلبات التي تصل إلي الشركة في الساعة , والفترة الزمنية بين طلب وأخر علي الترتيب , وبالتالي فإن عدد الطلبات المتوقع خلال ساعة عمل = λ حيث

$$\lambda = \frac{40}{10} = 4 \text{ طلبات / ساعة}$$

١- دالة الاحتمال للمتغير x هي$P(x)$ حيث:-

$$P_r(x) = \frac{e^{-4}(4)^x}{x!} \quad , x = 0,1,2,\ldots\ldots$$

٢- احتمال عدم وصول أي طلب في الساعة يساوي $P_r(x=0)$ حيث:

$$P_r(x=0) = \frac{e^{-4}(4)^0}{0!} = 0.018$$

٣- احتمال وصول أكثر من طلب في الساعة يساوي $p_r(x>1)$, وبما أن :

$$p_r(x>1)=1-\left[p_r(x=0)+p_r(x=1)\right]$$
$$=1-[0.018+\frac{e^{-4}(4)}{1!}]$$
$$=1-(0.018+0.072)$$
$$=1-0.09=0.91$$

٤- بما أن المتغير x يتبع توزيع بواسون فإن المتغير y يتبع التوزيع الآسي وبالتالي دالة كثافة الاحتمال للفترة الزمنية بين وصول طلبين إلي الشركة f(y) حيث :

$$f(y)=\lambda e^{-\lambda y}=4e^{-4y} \quad y\ge 0$$
$$\mu_y=\frac{1}{\lambda}=\frac{1}{4} \text{ ساعة } =15 \text{دقيقة}$$
$$\sigma_y^2=\frac{1}{\lambda^2}=\frac{1}{16}\to\sigma=\frac{1}{4}=15 \text{ دقيقة}$$
$$P_r(0<y<10)=0.92$$

التطبيق (٧-٣):

أجريت دراسة علي الدخل اليومي للعاملين بإحدى القطاعات بجمهورية مصر العربية , فوجد أن الدخل الشهري للعاملين بهذا القطاع يتبع التوزيع المعتاد وبتوقع 9.8 جنيه , والانحراف المعياري 1.6 جنيه , فإذا تم اختيار أحد العاملين بهذا القطاع عشوائيا . أوجد ما يلي :

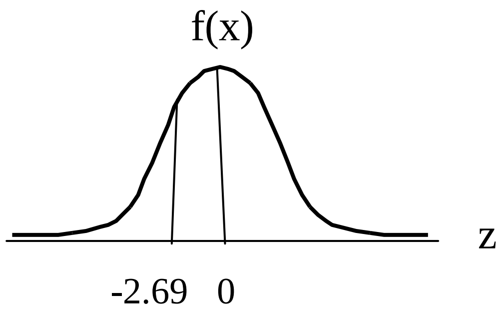
١- احتمال أن يكون دخل هذا العامل اليومي أقل من 5.5 جنيها.
٢- احتمال أن يكون دخل هذا العامل أكبر 10 جنيهات .
٣- احتمال أن يكون دخل هذا العامل أكبر من 11.4 وأقل من 13 جنيه .

الحل:-

إذا فرضنا أن x يمثل الدخل اليومي , فإن المتغير z حيث :

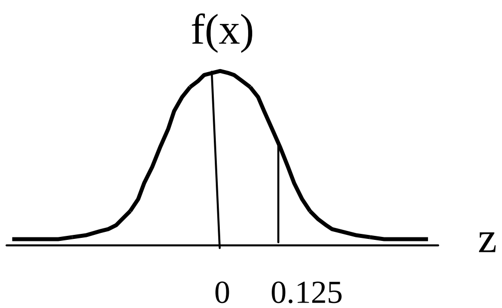
$z=\frac{x-\mu}{\sigma}=\frac{x-9.8}{106}$ يتبع التوزيع المعتاد القياسي وبالتالي فإن :

١- $p_r(x<5.5)=p_r(z<\frac{5.5-9.8}{1.6})=p_r(y<-2.69)$



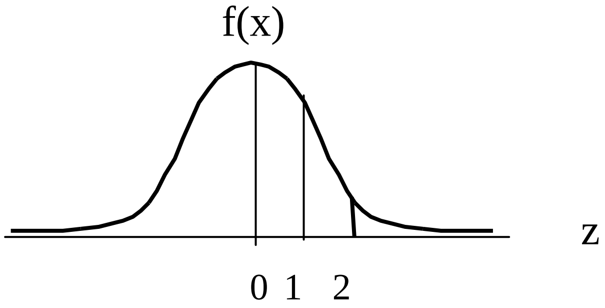
-٢

$$p_r(x > 10) = p_r(z > \frac{10-9.8}{1.6}) = p_r(z > 0.125)$$
$$= 0.5 - P_r(0 < z < 0.125)$$
$$= 0.4483$$



-٣

$$P_r(11.4 \le x \le 13) = P_r(\frac{11.4-9.8}{1.6} \le z \le \frac{13-9.8}{1.6})$$
$$= P_r(1 \le z \le 2)$$
$$= P_r(0 \le z \le 2) - p_r(0 \le z \le 1)$$
$$= 0.4772 - 0.3413 = 0.1359$$



تطبيق(٧-٤) :-

إذا كانت الكمية المستهلكة من الأكسجين (باللتر) في الساعة للفرد يتوقف علي وزن وطول وجنس الفرد . فإذا وجد أن هذه الكمية تتبع التوزيع المعتاد , فإذا أخذت عينة عشوائية مكونة من 9 أفراد ووجد أن الكمية المستهلكة من الأكسجين لأفراد هذه العينة تمثل متغير يتبع ستيودنت بدرجات حرية 0.8 , أحسب الاحتمالات التالية:

١- أن يكون استهلاك الفرد في العينة العشوائية أكثر من2.3 لتر في الساعة .

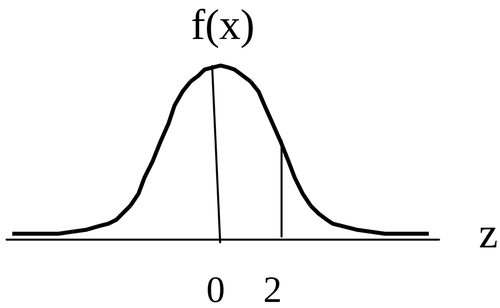
٢- أن يكون استهلاك الفرد في العينة من (3.4 - 2.9) لتر في الساعة , ووضح ذلك بيانيا .

الحل:

إذا فرضنا أن المتغير x يمثل استهلاك الفرد في العينة من الأكسجين في الساعة .

١- احتمال أن يستهلك الفرد أكثر من 2.3 لتر يساوي

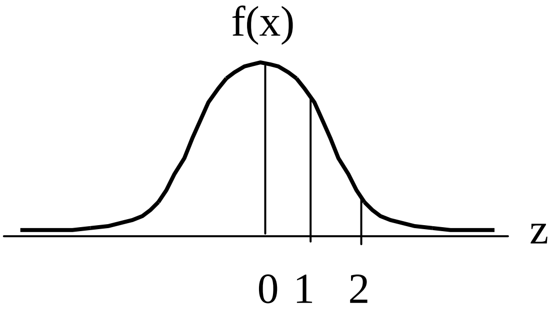
$$p_r\,(x \le 2.3) = 0.975$$



٢- احتمال أن يستهلك الفرد من 2.9 إلى 3.4 يساوي

$$P_r\,(2.9 \le x \le 3.4) = P_r\,(-\infty \le x \le 3.4) - P_r\,(-\infty \le x \le 2.9)$$

$$= 0.995\text{-}0.990\text{=}0.005$$



تطبيق (٧-٥) :

إذا كان استهلاك إحدى المؤسسات الصناعية الشهري للكهرباء يتوزع توزيع معتاد بتوقع 5000 كيلووات شهريا وانحراف معياري 250 كيلووات , أوجد :

١- احتمال أن يكون استهلاك المؤسسة أكثر من5700 كيلووات .

٢- احتمال أن يكون استهلاك المؤسسة أقل من 4500 كيلووات شهريا .

٣- احتمال أن يزيد استهلاك المؤسسة عن 5500 كيلووات شهريا .

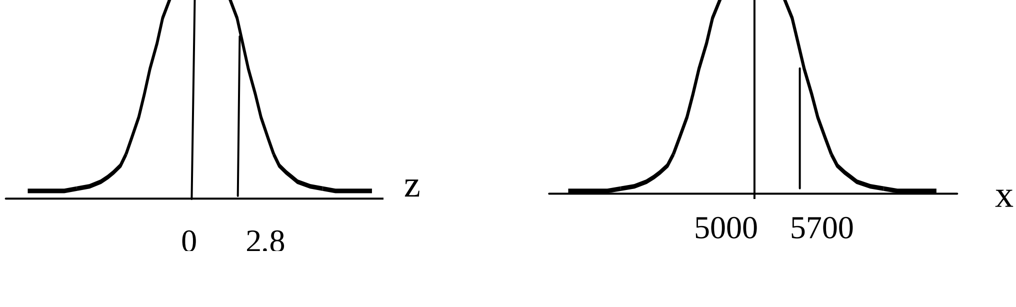
الحل :

إذا فرضنا أن المتغير x يمثل استهلاك المؤسسة الشهري من الكهرباء , فإن المتغير $z = \dfrac{x - 500}{250}$ يتبع التوزيع المعتاد القياسي ، وبالتالي فإن:-

١-

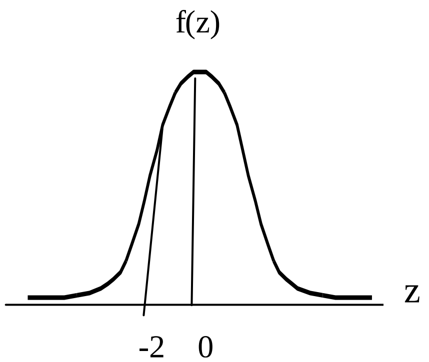
$$\begin{aligned} P_r(x > 5700) &= P_r(z > \frac{5700 - 5000}{250} \\ &= P_r(z > 2.8) = 0.5 - 0.4983 \\ &= 0.0062 \end{aligned}$$

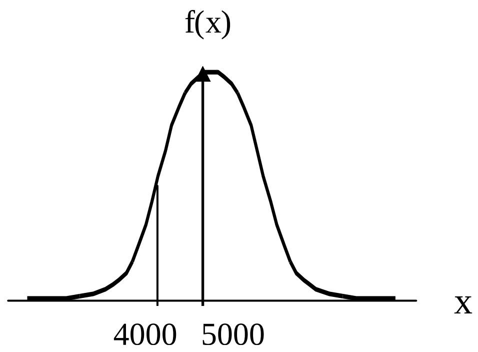
والشكل التالي يوضح ذلك:



-٢

$$P_r(x < 4000) = P_r(z < \frac{5000 - 4500}{250})$$
$$= P_r(z < -2) = 0.5 - P_r(0 < z < 2)$$
$$= 0.5 - 0.4772 = 0.0228$$

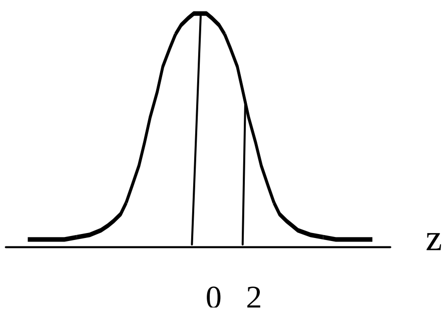


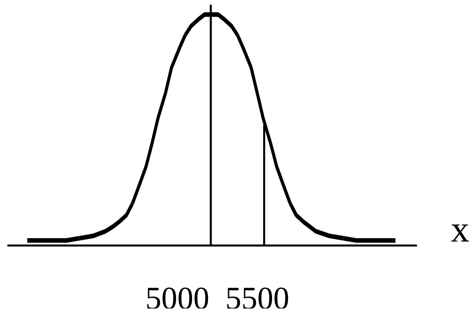


-٣

$$P_r(x < 5500) = P_r(z \frac{5500 - 5000}{250})$$
$$= P_r(z > 2) = 0.5 - 0.4772$$
$$= 0.0228$$

f(z) f(x)





تطبيق (٧-٦)

أثبتت إحدى الدراسات أن الأجر الشهري للعامل متغير يتبع التوزيع الطبيعي بتوقع 550جنيه وانحراف معياري 50 جنيه –فإذا تم اختيار أحد العمال عشوائياً

المطلوب :

١- أوجد احتمال أن يكون أجره أقل من 5.25 جنيه ثم أوجد نسبة العمال الذين يقل أجرهم عن 525 جنيه

٢- أوجد احتمال أن يكون أجره أكبر من 670جنيه ثم أوجد نسبة العمال الذين يزيد أجرهم عن 670 جنيه

٣- أوجد احتمال أن يكون أجر العامل بين 600-500 ثم أوجد نسبة العاملين الذين تتراوح أجورهم من 600-500جنيه

الحل:

$$\mu = 550 \quad , \sigma = 50$$

نفرض أن xتشير إلى أجر العامل ، بالتالي فإن x تتبع التوزيع المعتاد كذلك

تتبع التوزيع المعتاد القياسي $y = \frac{x\mu}{\sigma}$

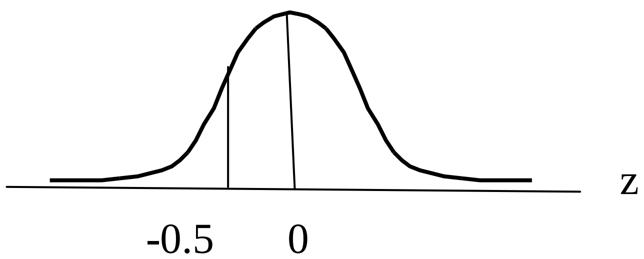
١- احتمال أن يكون أجر العامل أقل من 525 جنيه = $P_r(x < 525)$

-٢

$$= P_r(y < \frac{525-550}{50}) = P_r(y < -0.5)$$

$$= 0.5 - P_r(0 < y < 0.5)$$

$$= 0.5 - 0.1915 = 0.3085$$



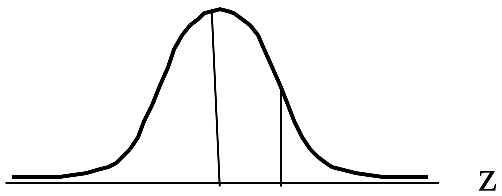
نسبة العاملين الذين يقل أجر كل منهم عن 550 جنيه

$$= P_r(x < 550) \times 100$$

$$= 0.3085 \times 100 = 30.85\%$$

٢-احتمال أن يكون أجر العامل أكبر من 670 جنيه $= P_r(x > 670)$

$$= P_r(y > \frac{670-550}{50}) = P_r(y > 2.4)$$
$$= 0.5 - P_r(0 > y > 2.4)$$
$$= 0.5 - 0.4918 = 0.0082$$



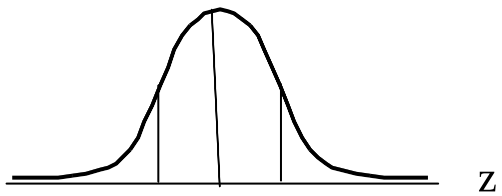
نسبة العمال الذين يزيد أجر كل منهم عن 670جنيه

$$= P_r(x > 670) \times 100$$
$$= 0.0082 \times 100 = 0.82\% \approx 1\%$$

٣-احتمال أن يكون أجر العمال الشهري بين 500-600جنيه

$(P_r\, 500 \le x \le 600) =$

$$= P_r(\frac{500-550}{50} \le y \le \frac{600-550}{50})$$
$$= P_r(-1 \le y \le 1) = 2P_r(0 \le y \le 1)$$
$$= 2 \times 0.3413 = 0.6826$$



نسبة العاملين الذين تتراوح أجورهم بين 500-600جنيه

$$= P_r(500 \le x \le 600) \times 100$$
$$= 0.6826 \times 100 = 68.26\% \approx 68\%$$

تطبيق (٧-٧)

في دراسة أجرتها إحدى محطات البنزين عن كمية البنزين التي تطلبها السيارات التي تطلب الخدمة من المحطة فوجد أن كمية البنزين تمثل متغير عشوائي يتبع التوزيع الطبيعي بتوقع 20 لتراً وانحراف معياري 5 لتر . أوجد الاحتمالات التالية:

١- أن تطلب السيارة كمية بنزين أقل من 10لتر

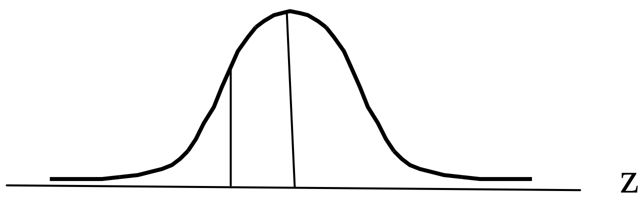
٢- أن تطلب السيارة كمية أكبر من 35 لتر

٣- أن تطلب السيارة كمية من 35-25لتر

الحل:

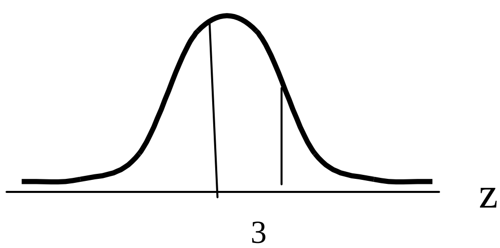
١-

$$= P_r(x < 10) = P_r(y < \frac{10-20}{5})$$
$$= P_r(y < -2) = 0.5 - P_r(0 < y < 2)$$
$$= 0.5 - 0.4861 = 0.0139$$



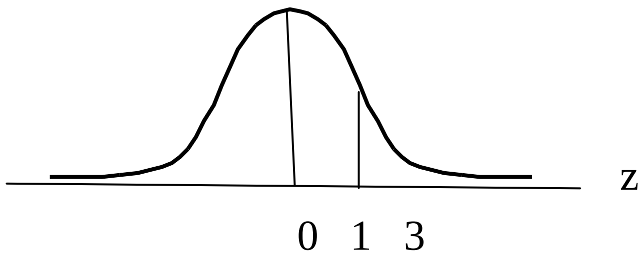
٢-

$$= P_r(x > 40) = P_r(y > \frac{35-20}{5})$$
$$= P_r(y > 3) = 0.5 - P_r(0 < y < 3)$$
$$= 0.5 - 0.4987 = 0.0013$$



-٣

$$= P_r(25 \le x \le 35) = P_r(\frac{25-20}{5}) \le y \le \frac{35-20}{5})$$
$$= P_r(1 \le y \le) = P_r(0 \le y \le 3) - P_r(0 \le y \le 1)$$
$$= 0.4987 - 0.33413 = 0.1574$$



تطبيق(٧-٨)

قامت إحدى شركات الإتصال بعمل دراسة عن متوسط عدد الساعات اليومية لإستخدام شبكة الإنترنت للعميل بالنسبة للعملاء المشتركين في الشبكة فوجد أن متوسط عدد الساعات للعميل يمثل متغير معتاد بتوقع 3 ساعات يومياً وانحراف معياري ساعة واحدة ونصف . أوجد الاحتملات التالية:-

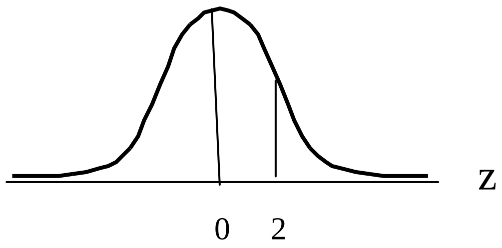
١- احتمال أن يكون متوسط عدد الساعات اليومية أكبر من 5 ساعات.
٢- احتمال أن يكون متوسط عدد الساعات اليومية أقل من 2 ساعة.
٣- احتمال أن يكون متوسط عدد الساعات اليومية بين 6-4 ساعات يومية.

الحل:-

إذا فرضنا أن x تشير إلى متوسط عدد الساعات اليومية لاستخدام شبكة الإنترنت فإن

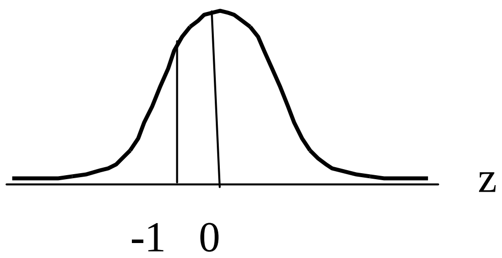
١- احتمال أن يكون متوسط عدد الساعات اليومية أكبر من 5 ساعات

$$= P_r(x > 5) = P_r(y > \frac{5-3}{1}) = P_r(y > 2)$$
$$= 0.5 - P_r(0 < y < 2)$$
$$= 0.5 - 0.4772 = 0.0228$$



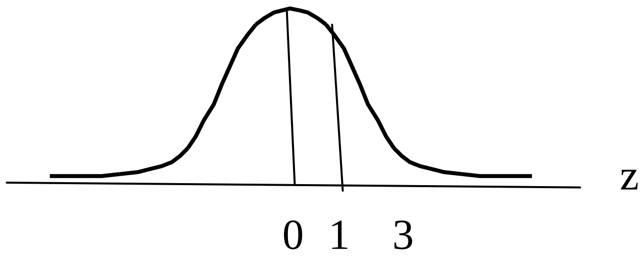
٢- احتمال أن يكون منوسط عدد الساعات اليومية أقل من ساعتين

$$= P_r(x < 2) = P_r(y < \frac{2-3}{1}) = P_r(y < -1)$$
$$= 0.5 - P_r(0 < y < 1)$$
$$= 0.5 - 0.3413 = 0.1587$$



٣-احتمال أن يكون متوسط عدد الساعات اليومية بين 4-6 ساعات

$$= P_r(4 < x < 6) = P_r\ ?(\frac{4-3}{1} < y < \frac{6-3}{1})$$
$$= P_r(1 < y < 3)$$
$$= P_r(0 < y < 3) - P_r(1 < y < 1)$$
$$= 0.4987 - 0.3413 = 0.1574$$



تطبيق (٧-٩)

إذا كـان الأجـر الشـهري للعامـل يمثـل متغيـر معتـاد توقعـه 570 جنيـه وانحرافـه المعيـاري 100 جنيـه فـي إحـد المؤسسـات الصـناعية فـإذا تـم اختيـار أحـد العمـال عشوائياً

المطلوب:-

١- أوجد احتمال أن يكون أجره الشهري أكبر من 800 جنيه
٢- أوجد احتمال أن يكون أجره الشهري أقل من 500 جنيه
٣- أوجد احتمال أن يكون أجره الشهري بين 600-700 جنيه
٤- أوجد نسبة العمال الذين يزيد أجر كل منهم 800 جنيه
٥- أوجد نسبة العمال الذين يقل أجر كل منهم 750 جنيه
٦- أوجد المبلغ الذي 60% من العمال يأخذ كل منهم أجر أقل منه
٧- أوجد المبلغ الذي 75% من العمال يأخذ كل منهم أجر أكبر منه

الحل:-

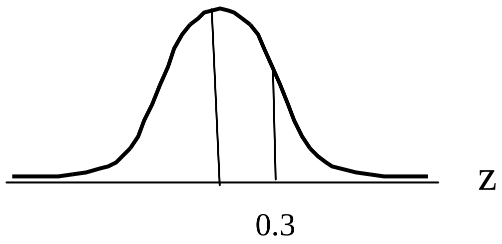
إذا فرضنا أن x تشير ؟إلى الأجر الشهري للعامل فإن :

١- احتمال أن يكون أجر العامل الشهري أكبر من 800 جنيه

$$= P_r(x > 800) = P_r(y > \frac{600-700}{100}) = P_r(y > 0.30)$$

$$= 0.5 - P_r(0 < y < 0.3)$$

$$= 0.5 - 0.1179 = 0.3821$$



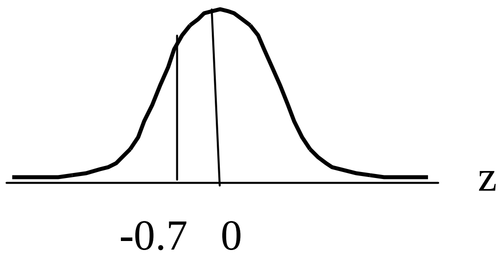
٢- احتمال أن يكون أجر العامل الشهري أقل من 500 جنيه

$$= P_r(x < 500) = P_r(y < \frac{500-570}{100}) = P_r(y < -0.7)$$

$$= 0.5 - P_r(0 < y < 0.7)$$

$$= 0.5 - P_r(0 < y < 0.7)$$

$$= 0.5 - 0.2580 = 0.242$$

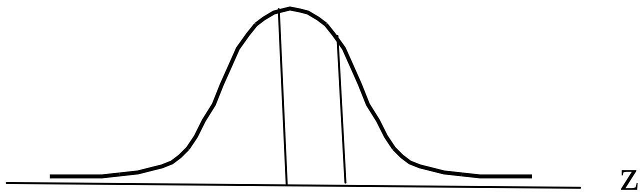


٣-احتمال أ\ن يكون أجر العامل بين 700-600 جنيه

$$= P_r(600 < x < 700) = P_r(\frac{600-570}{100} < y < \frac{700-570}{100})$$

$$= P_r(0.3 < y < 1.3) = P_r(0 < y < 1.3) - P_r(0 < y < 0.3)$$

$$= 0.4032 - 0.1179 = 0.2853$$

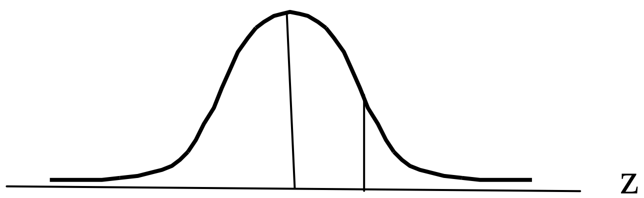


٤- نسبة العمال الذين يزيد أجر كل منهم عن 800 جنيه

$$= P_r(x > 800)\times 100 = P_r(y > \frac{800-570}{100})\times 100$$

$$= P_r(y > 2.3)\times 100 \;= [0.5 - 0.4893]\times 100$$

$$= 0.0107\times 100 = 1.07\%$$



٥- نسبة العمال الذين يقل أجر كل منهم عن 570 جنيه

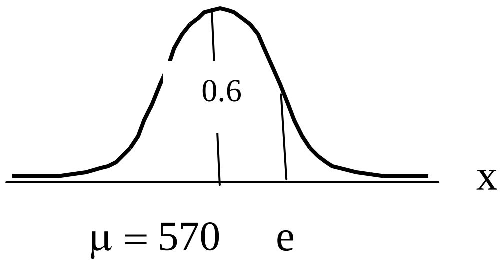
$$= P_r(x < 570)\times 100 = P_r(y < 0)\times 100$$

$$= 0.50\times 100 = 50\%$$

٦- إذا فرضنا أن المبلغ الذي 60% من العمال يأخذ كل منهم أجر أقل منه يساوي e

$$= P_r(x < e) = 0.60 = P_r(\frac{e-570}{100}) = 0.60$$

$$= 0.5 + P_r(0 < y < \frac{e-570}{100}) = 0.1$$

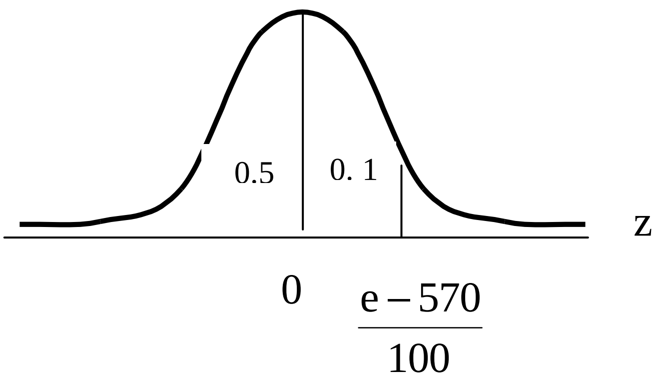


من جدول التوزيع المعتاد القياسي نجد أن :

$$\frac{e-570}{100} = 0.25$$

$$e - 570 = 0.25 \times 100 = 25$$

$$e = 570 + 25 = -595$$



٧- إذا فرضتنا أن المبلغ الذي 75% من العمال يأخذ كل منهم أجر أكبر منه يساوي a

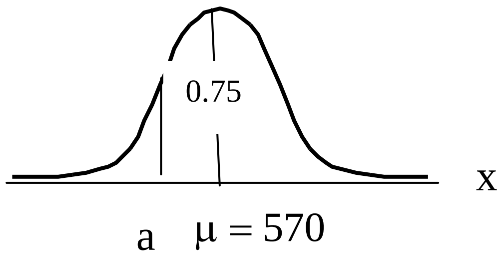
$$P_r(x > a) = 0.75$$

$$= P_r(y > \frac{a-570}{100}) = 0.75$$

$$= 0.5 + (a < x < 570) = 0.75$$

$$= P_r(a < x < 570) = 0.25$$

$$= P_r(\frac{a-570}{100} < y < 0) = 0.25$$

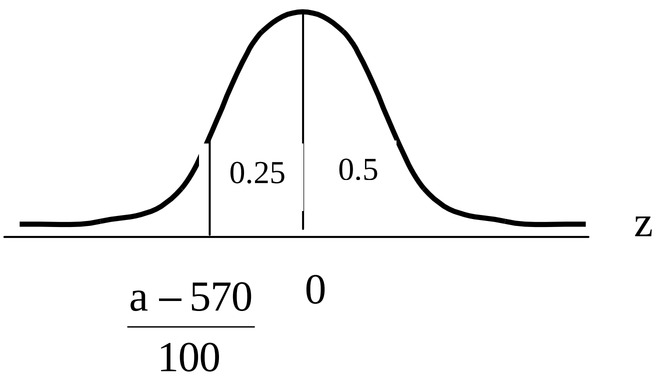


من جدول التوزيع المعتاد القياسي نجد أن :

$$\frac{a-570}{100} = -0.67$$

$$a - 570 = -67$$

$$a = -67 + 570 = 503 \quad \text{جنيه}$$



تطبيق (٧-١٠)

إذا كان وصول المرضى إلى إحدى العيادات العلاجية وفقاً لتوزيع بواسون بمعدل 3 مرضى في الساعة

المطلوب:

١- أوجد التوزيع الاحتمالي لعدد المرضى في الساعة.

٢- أوجد احتمال وصول 5 مرضى للعيادة في الساعة .

٣- أوجد التوزيع الاحتمالي للفترة الزمنية بين وصول إثتين من المرضى على التوالي.
٤- أوجد احتمال أن تكون الفترة الزمنية بين وصول إثنين من المرضى على التوالي أقل من 15 دقيقة ووضح ذلك بيانياَ.
٥- أوجد احتمال أن تكون الفترة الزمنية بين وصول إثنين من المرضى على التوالي أكبر من 30 دقيقة ووضح ذلك بيانيا.

الحل:

١- إذا فرضنا أن x تشير إلى عدد المرضى تصل إلى العيادة في الساعة فإن $\lambda = 5$ مريض/ساعة

$$P_r(x) = \frac{e^{-5}(5)^x}{x!} \quad , x = 0,.1,2,.....$$

٢- إذا فرضنا أن y هى الفترة الزمنية بين وصول إثنين من المرضى على التوالي فإن:

$$\lambda = \frac{5}{60} = 0.0833$$

$$\begin{aligned} P_r(y) &= \lambda e^{-\lambda x} \quad , y \geq 0 \\ &= 0.0833 e^{-0.0833y} \quad , y \geq 0 \end{aligned}$$

٤- احتمال أن تكون الفنرة الزمنية بين وصول إثنين من المرضى على التوالي أقل من 15 دقيقة يساوي:

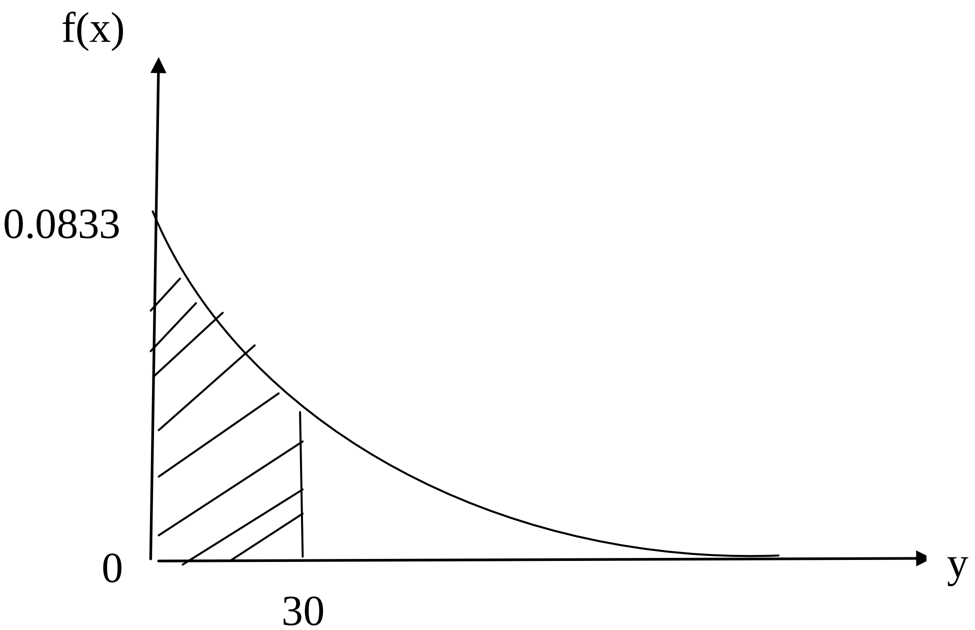
$$\begin{aligned} P_r(y < 15) &= \int_0^{15} 0.0833 e^{-0.0833} \quad dy \\ &= [-e^{-0.0833y}]_0^{15} \\ &= [e^{-0.033(15)} - 1] \\ &= 1 - 0.286648 = 0.713352 \end{aligned}$$

$$P_r(y > 30) = 1 - P_r(y < 30)$$

$$= 1 - \int_0^{30} 0.0833 e^{-0.0833\,y} \; dy$$ ـه

$$= 1 - \left\{ [-e^{-0.0833y}]_0^3 \right\}$$

$$= e^{-2.499} = 0.0822$$



(٧-٨) تمرينات Exercises

(٧-١)

إذا كان متوسط المعدل المتوقع لوصول العملاء لأحد البنوك التجارية لطلب الخدمة 120 عميل ، فإذا كان عدد العملاء الذين يطلبون الخدمة يتبع توزيع بواسون:

أوجد :

١- أوجد التوزيع الاحتمالي لعدد العملاء في الدقيقة.

٢- احتمال وصول أكثر من 100 عميل في الدقيقة.

٣- احتمال وصول أقل من 10 عملاء في الدقيقة .

٤- أوجد التوزيع الاحتمالي للفترة الزمنية بين وصول عميلين علي التوالي .

٥- أوجد احتمال أن تكون الفترة الزمنية بين وصول عميلين علي التوالي أكثر من 6 ثواني وأقل من 30 ثانية .

(٧-٢)

في دراسة بأحدي محطات البنزين عن الفترة الزمنية التي يستغرقها العامل في تموين السيارة وجد أن زمن خدمة السيارة بالمحطة يتبع التوزيع الأسي بتوقع 7 دقائق لخدمة السيارة .

المطلوب

١- أوجد التوزيع الاحتمالي لفترة خدمة السيارة بالمحطة ووضح ذلك بيانيا .

٢- أوجد احتمال أن فترة خدمة السيارة بالمحطة أكبر من 5 دقائق وأقل من 10 دقائق , ووضح ذلك بيانيا .

٣- أوجد التوزيع الاحتمالي لعدد السيارات التي تصل إلي المحطة لطلب الخدمة في الدقيقة .

٤- أوجد احتمال وصول أكثر من 10 سيارات للمحطة لطلب الخدمة في الدقيقة .

(٧-٣)

إذا كانت قدرة الفرد علي تذكر مجموعة من الأحداث(أو المعلومات) المرتبطة يعتمد علي طول الفترة بين وقوع هذه الأحداث (أو المعلومات) والسؤال عنها . فإذا كان المتغير x يمثل طول الفترة الزمنية بين وقوع الأحداث و السؤال عنها ، حيث x يتبع التوزيع الأسي بتوقع 3 أيام .

أوجد ما يلي :

١- التوزيع الاحتمالي للمتغير x ووضح ذلك بيانيا .

٢- أحسب احتمال أن يتذكر الفرد المعلومات بعد أكثر من 3 أيام .

٣- احتمال أن يتذكر الفرد في فترة من يومين إلي 5 أيام .

(٧-٤)

إذا كان x متغير عشوائي متصل له دالة كثافة الاحتمال f(x) حيث :

$$f(x) = \begin{cases} 5c & , \ 2 < x < 10 \\ 0 & \text{فيما عداغ ذلك} \end{cases}$$

١- أوجد قيمة الثابت c , ثم وضح بيانيا الدالة f(x) .

٢- أحسب الاحتمالات التالية ووضح ذلك بيانيا

$$P_r(x < 7) , P_r(x > 5) , P_r(3 < x < 8)$$

(٧-٥)

إذا كان مستوي تحصيل الطالب في المرحلة الثانوية في مادة الرياضيات يتبع التوزيع المعتاد بتوقع 75 درجة وتباين 25

المطلوب

١- ما هو احتمال أن يكون مستوى تحصيل الطالب في الرياضيات أكبر من 80 درجة

٢- ما هو احتمال أن يكون مستوى تحصيل الطالب ينحصر بين 65-50 درجة

٣- ما هو احتمال أن يكون مستوى تحصيل الطالب أقل من 55 درجة

٤- احسب نسبة الطلاب في (١)، (٢)، (٣)

(٧-٦)

إذا كان وزن الطفل في السن من 10-5 سنوات يمثل متغير معتاد توقعه 45 كيلو جرام وتباينه 9

المطلوب

١- ماهو احتمال وجود طفل وزنه أكبر من 40 كيلو جرام

٢- ما هو احتمال وجود طفل وزنه أقل من 18 كيلو جرام

٣- ما هي نسبة الأطفال الذين تتراوح أوزانهم بين 30-25 كيلو جرام

(٧-٧)

تقوم احدى الشركات الصناعية الصناعية بإنتاج محركات الطائرات ، فإذا كان عدد ساعات التشغيل المنوقعة للمحرك بدون حدوث أي عطلة تسلوي 10000 ساعة تشغيل.

فإذا كان أقصى عدد ساعات تشغيل للمحرك بدون حدوث أي عطل تساوي 1500 ساعة تشغيل ، أقل عدد ساعات تشغيل بدون حدوث عطل للمحرك يساوي 5000 ساعة تشغيل.

المطلوب

١- أوجد التوزيع الاحتمالي لعدد ساعات تشغيل المحرك بدون عطل ووضح ذلك بيانياً .
٢- أوجد القيمة المتوقعة والانحراف المعياري لساعات تشغيل المحرك.
٣- أوجد احتمال أن تكون ساعات تشغيل المحرك أكثر من 8000 ساعة تشغيل.
٤- أوجد احتمال أن تكون ساعات تشغيل المحرك أقل من 5700 ساعة تشغيل.

(٧-٨)

إذا كان متغير متصل له دالة كثافة الاحتمال $f(x)$ حيث:-

$$f(x) = \begin{cases} 2e^{-2x} & 0 \le x \le \infty \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

المطلوب

١- أثبت أن $f(x)$ دالة كثافة احتمال .
٢- احسب:

$$i)P_r(x < 5)$$
$$ii)P_r(10 < x < 20)$$

ووضح ذلك بيانياً

(٧-٩)

إذا كان x متغير عشوائي له دالة الاحتمال $f(x)$ بحيث :-

$$f(x) = \begin{cases} \frac{1}{12} & 8<x<20 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

المطلوب

١- أوجد :

$$i)P_r(x \le 15)$$
$$ii)P_r(10 \le x \le 15)$$

ووضح ذلك بيانياً.
٢- أوجد التوقع والتباين للمتغير x

(٧-١٠)

في إحدى المؤسسات العلاجية وجد أن أعمار المترددين عليها في أحد التخصصات يمثل متغير يتبع التوزيع المكنتظم من 35-65 سنة.
فإذا كان المتغير x يشير إلى عمر المتردد بالسنوات فإن دالة كثافة الاحتمال f(x) بحيث :

$$f(x) = \begin{cases} \frac{1}{30} & 35 \le x \le 65 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

المطلوب

١- ارسم دالة كثافة الاحتمال .
٢- أوجد دالة التوزيع التراكمية للمتغير x ووضح ذلك بيانياً.
٣- احسب احتمال اختيار أحد المترددين يكون عمره من 55-60 سنة
٤- أوجد العمر المتوقع للمتردد

(٧-١١)

إذا كانت f(x) دالة للمتغير x بحيث:-

$$f(x) = \begin{cases} x & 0<x<1 \\ 2x & 1<x<3 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

المطلوب

١- ارسم الدالة f(x).
٢- أثبت أن f(x) دالة كثافة احتمال.
٣- أوجد احتمال أن تكون x أكبر من 0.5 وأقل من 1.5.
٤- أوجد توقع وتباين x

(٧-١٢)

استثمر أحد الأشخاص مبلغ 3000 جنيه ف أحد المشاريع الذي يعطي عائد سنوي متوقع 10% ، كذلك استثمر مبلغ 2000 جنبه بمعدل سنوي متوقع 16% وانحراف معياري 8% في مشروع أخر .

المطلوب

١- أوجد القيمة المتوقعة للمبلغين بعد عام .

٢- أوجد الانحراف المعياري لإجمالي المبلغ بعد عام.

أهم المصطلحات الإحصائية

باللغة العربية	باللغة الإنجليزية	رقم الصفحة المحتوية للمصطلح
الوسط الحسابي	Arithmetic Means	٩٠
نظرية بيز	Bayes′s Theorem	١٨١
توزيع برنولي	Bernoulli Distribution	٢٠٨
توزيع ذات الحدين	Binomial Distribution	٢١٢
معاملات الاختلاف	Coefficient of Variation	١٢٣
مكمل	Complement	١٥٨
احتمال شرطي	Conditional Probability	١٧١
متغير متصل (مستمر)	Continuous Variable	٢٠٣
التوزيع التكراري التراكمي	Cumulative Frequency Distribution	٤٣
بيانات	Data	٢١
قرار	Decision	١١
صناعة القرار	Decision′s Making	١١
الإحصاء الوصفي	Description Statistics	١٣
متغير متقطع (منفصل)	Discrete Variable	٢٠٣
حدث	Event	٢٠٩
قيمة متوقعة	Expected Value	٢٠٩
التوزيع الآسي	Exponential Distribution	٢٧٧
توزيع تكراري	Frequency Distribution	٤٣
الوسط الهندسي	Geometric Means	٩٠
مدرج	Histogram	٥٣
التوزيع الهندسي الزائد	Hyper Geometric Distribution	٢٢١
استقلال	Independence	١٥٤

باللغة العربية	باللغة الإنجليزية	رقم الصفحة المحتوية للمصطلح
أرقام قياسية	Index Numbers	٣٤٣
الاستدلال الإحصائي	Inferential Statistics	١٤
معلومات	Information	١٣
تقاطع	Intersection	٧٤
معرفة	Knowledge	١٤
الوسيط	Median	٩٨
المنوال	Mode	١٠٤
التوزيع المعتاد	Normal Distribution	٢٩٤
توزيع بواسون	Poisson Distribution	٢٢٨
مضلع	Polygon	٥٦
مجتمع	Population	٢٦
احتمال	Probability	١٥٣
دالة كثافة الاحتمال	Probability Function	٢٠٤
شجرة الاحتمال	Probability Tree	١٧٧
عملية	Process	١٦٤
بيانات وصفية	Qualitative Data	٢٤
بيانات كمية	Quantitive Data	٢٤
عينة عشوائية	Random Sample	٣٩
توزيع نسبي	Relative Distribution	٦٨
عينة	Sample	٢٦
فراغ المعاينة	Sampling Space	١٥٥
نصف المدى الربيعي	Semi_Inter quartile	١٠٩
انحراف معياري	Standard Deviation	١١٨
التوزيع المعتاد القياسي	Standard Normal Distribution	٢٩٠

رقم الصفحة المحتوية للمصطلح	باللغة الإنجليزية	باللغة العربية
١٥٤	Statistical Independent	الاستقلال الإحصائي
١٢	Statistical Thinking	التفكير الإحصائي
١٢	Statistics	الإحصاء
١١٨	Variance	تباين
١٦٤	Union	إتحاد

الملاحق

ملحق (١) : إشارة المجموع Σ **Notation**

ملحق (٢) : الأرقام القياسية **Index Numbers**

ملحق (٣) : توزيع ذات الحدين

Binomial Probability Distribution

ملحق (٤) : توزيع بواسون

Poisson Distribution

ملحق (٥) : التوزيع الأسى السالب

Negative Exponential Distribution

ملحق (٦) : التوزيع المعتاد القياسى

Standard Normal Probability Distribution

ملحق (١) إشارة المجموع (Σ) وخصائصها

Summation Notation (Σ) and its Properties

إذا كان x متغير يأخذ قيماً مختلفة عددها n على النحو التالي $x_1, x_2, x_3, ..., x_n$ فإنه يمكن الإشارة إلى هذه القيم بالرمز x_i ، $i = 1,2,3,....., n$. فإذا كان C هو مجموع هذه القيم حيث :

$$C = x_1 + x_2 + x_3 + ... + x_n \qquad (1)$$

فإنه يمكن كتابة هذا المجموع بطريقة مختصرة باستخدام الإشارة (Σ) على النحو التالى :

$$C = \sum_{i=1}^{n} x_i \qquad (2)$$

حيث (Σ) تشير إلى إجراء عملية الجمع بالنسبة لحدود x_i ابتداء من $i = 1$ في صورة متتالية عددية ، أى $i = 1 \;,\; i = 2 \;,\; i = 3 \;,\; ..... \;,\; i = n$.

خصائص إشارة المجموع

١- إذا كان a مقدار ثابت فإن :

$$\sum_{i=1}^{n} a = \underbrace{a + a + a + ... + a}_{n} = na \qquad (3)$$

مثال (١)

$$\sum_{i=1}^{5} 100 = 100 + 100 + 100 + 100 + 100 = 5(100) = 500$$

٢- إذا كان a مقدار ثابت ، x_i متغير فإن :

$$\sum_{i=1}^{n} ax_i = ax_1 + ax_2 + ax_3 + ... + ax_n$$

$$= a(x_1 + x_2 + x_3 + ... + x_n) = a\sum_{i=1}^{n} x_i \qquad (4)$$

وهذا يعنى أنه يمكن إخراج المقدار الثابت خارج إشارة المجموع (Σ)

مثال (٢)

$$\sum_{i=1}^{n} 5y_i = 5y_1 + 5y_2 + 5y_3 + \dots\dots + 5y_n$$

$$= 5(y_1 + y_2 + y_3 + \dots\dots + y_n) = 5\sum_{i=1}^{n} y_i$$

٣- إذا كان $x_i \;, y_i$ متغيران فإن :

$$\sum_{i=1}^{n} (x_i \pm y_i) = (x_1 \pm y_1) + (x_2 \pm y_2) + \dots\dots + (x_n \pm y_n)$$

$$= (x_1 + x_2 + \dots\dots + x_n) \pm (y_1 + y_2 + \dots\dots + y_n) = \sum_{i=1}^{n} x_i \pm \sum_{i=1}^{n} y_i \quad (5)$$

مثال (٣)

$$\sum_{j=1}^{10} (2x_j - 3y_j) = \sum_{j=1}^{10} 2x_j - \sum_{j=1}^{10} 3y_j = 2\sum_{j=1}^{10} x_j - 3\sum_{j=1}^{10} y_j$$

٤- إذا كان $x_j \;, y_j$ متغيران فإن :

$$\sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \dots\dots + x_n y_n \neq (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i) \qquad (6)$$

٥- إذا كان $x_i \;, y_i$ متغيران فإن :

$$\sum_{i=1}^{n} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \frac{x_3}{y_3} + \dots\dots + \frac{x_n}{y_n} \neq \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} \qquad (7)$$

مثال (٤)

أوجد مفكوك المجاميع التالية :

1 - $\sum_{j=1}^{4} (x_j - 3)^2$ ، 2 - $\sum_{i=1}^{10} (y_i - 5)$

الحل

$$1-\sum_{j=1}^{4}(x_j-3)^2=(x_1-3)^2+(x_2-3)^2+(x_3-3)^2+(x_4-3)^2$$

$$=(x_1^2-6x_1+9)+(x_2^2-6x_2+9)+(x_3^2-6x_3+9)+(x_4^2-6x_4+9)$$

$$=(x_1^2+x_2^2+x_3^2+x_4^2)-6(x_1+x_2+x_3+x_4)+36$$

$$=\sum_{j=1}^{4}x_j^2-6\sum_{j=1}^{4}x_j+36$$

وبالمثل :

$$2-\sum_{i=1}^{10}(y_i-5)=(y_1-5)+(y_2-5)+(y_3-5)......+(y_{10}-5)$$

$$=(y_1+y_2+y_3+.....+y_{10})-(5+5+5+.....+5)$$

$$=\sum_{i=1}^{10}y_i-10(5)=(\sum_{i=1}^{10}y_i)-50$$

حل أخر

$$1-\sum_{j=1}^{4}(x_j-3)^2=\sum_{j=1}^{4}(x_j^2-6x_1+9)$$

$$=\sum_{j=1}^{4}x_j^2-6\sum_{j=1}^{4}x_j+\sum_{j=1}^{4}9$$

$$2-\sum_{i=1}^{10}(y_i-5)=\sum_{i=1}^{10}y_i-10(5)=(\sum_{i=1}^{10}y_i)-50$$

مثال (٥)

استخدم إشارة المجموع للتعبير عن المجاميع التالية :

$$1-x_1^2+x_2^2+.......+x_{30}^2$$

$$2-z_1x_1^2+z_2x_2^2+.......+z_{25}x_{25}^2$$

$$3-\frac{(x_1-3)}{(y_1-5)}+\frac{(x_2-3)}{(y_2-5)}+.......+\frac{(x_{30}-3)}{(y_{30}-5)}$$

الحل

$$1-x_1^2+x_2^2+.......+x_{30}^2=\sum_{j=1}^{30}x_j^2$$

$$2-z_1x_1^2+z_2x_2^2+.......+z_{25}x_{25}^2=\sum_{j=1}^{25}z_jx_j^2$$

$$3-\frac{(x_1-3)}{(y_1-5)}+\frac{(x_2-3)}{(y_2-5)}+.......+\frac{(x_{30}-3)}{(y_{30}-5)}=\sum_{i=1}^{30}\frac{(x_i-3)}{(y_i-5)}$$

Exercises تمرينات

١- إذا كان

$$\sum_{i=1}^{7}x_i^3=-4\ ,\ \sum_{i=1}^{7}x_i^3=10\ ,\ \sum_{i=1}^{7}x_i^3=8$$

أوجد ما يلي :

$$1-\sum_{i=1}^{7}(5x_i+3)\qquad,\qquad 2-\sum_{i=1}^{7}(5x_i+3)^2$$

$$3-\sum_{i=1}^{7}(x_i)(x_i+2)\qquad,\qquad 3-\sum_{i=1}^{7}(x_i+11)^2$$

٢- اثبت أن :

$$\sum_{i=1}^{n}(5x_i+10y_i+8z_i)=5\sum_{i=1}^{n}x_i+10\sum_{i=1}^{n}y_i+8\sum_{i=1}^{n}z_i$$

٣- استخدم أشارة المجموع للتعبير عن المجاميع التالية :

$$1-x_1^2+x_2^2+x_3^2+.......+x_{10}^2$$

$$2-z_1x_1+z_1^2x_2+z_1^3x_3+.......+z_1^{20}x_{20}$$

$$3-z_1x_1+y_1^2+z_2x_2+y_2^2+......+z_{20}x_{20}+y_{20}^2$$

$$4-\frac{(x_1^2)}{(y_1)}+\frac{(x_2^2)}{(y_2)}+.......+\frac{(x_7^2)}{(y_7)}$$

ملحق (٢) الأرقام القياسية **Index Number**

في الباب الرابع من هذا الكتاب تناولنا العديد من مقاييس الموضع والتشتت التي تصف الظواهر وخصائصها ، ولكن قد تتطلب دراسة بعض الظواهر وبصفة خاصة الاقتصادية والتجارية تتبع التغيرات التي تطرأ على قيم الظاهرة وفقاً للزمان أو المكان

والرقم القياسي هو مؤشر إحصائي يستخدم لقياس التغيرات التي تطرأ على ظاهرة معينة من فترة زمنية لفترة زمنية أخري أو من منطقة لمنطقة أخري*.

فمثلاً إذا كان سعر الكيلو الواحد من السكر سنة ١٩٦٠ هو 20 قرش وأصبح سعره سنة ١٩٩٠ هو 100 قرش ، كذلك سعر كيلو اللبن في ١٩٦٠ كان 80 قرش فأصبح سنة ١٩٩٠ هو 200 قرش . فنجد أن كيلو السكر زاد بمقدار 80 قرش سنة ١٩٩٠ عما كان عليه في سنة ١٩٦٠ ، ويتضح أن الزيادة المطلقة في سعر كيلو اللبن أكبر من الزيادة المطلقة في سعر كيلو السكر.

ولكن إذا تتبعنا التغير النسبي في سعر السكر في سنة ١٩٩٠ عنه في سنة ١٩٦٠ نجد أن نسبة الزيادة في سعر الكيلو وصلت 400% في حين الزيادة في سعر كيلو اللبن في سنة ١٩٩٠ عنه في سنة ١٩٦٠ وصات إلى 150% فقط .
ومن ذلك يتضح أن التغير النسبي في سعر كيلو اللبن (150%) أقل من بكثير من التغير النسبي في سعر كيلو السكر (400%).

ويستخدم الرقم القياسي في تحديد التغير النسبي في قيمة الظاهرة باعتباره أكثر أهمية من التغير المطلق الذي قد لا يعكس الاختلاف الحقيقي في قيمة الظاهرة – خاصة في حالة تعدد وحدات القياس المستخدمة.

وقد بدأت دراسة الأرقام القياسية في القرن الثامن عشر نتيجة الحاجة إلى مقاييس تساعد في دراسة وحل المشاكل الاقتصادية . وفيما يلي سوف نتناول الطرق المختلفة لتكوين الرقم القياسي ، ونظراً لأن الهدف من حساب الرقم القياسي هو تحديد التغيرات التي تطرأ على ظاهرة معينة بين فترتين زمنيتين (أو منطقتين) وبالتالي فإن تكوين الرقم القياسي يتطلب :

١- تحديد الفترة (أو المنطقة) التي يبدأ فيها قياس التغير وتسمى بفترة الأساس Base Period ، أما الفترة الثانية فتسمى فترة المقارنة Given Period وعند تحديد فترة الأساس فإنه من الضروري تجنب فترات عدم الاستقرار من كافة النواحي الاقتصادية والسياسية و..... الخ.
٢- توفير بيانات كاملة عن الظاهرة محل الدراسة في فترتي الأساس والمقارنة.

وفيما يلي سوف نتناول طرق تكوين الرقم القياسي

أولاً : الأرقام القياسية باستخدام المناسيب

---

* أ.د. نادية مكارى جرجس (١٩٧٣) : " مذكرات في الإحصاء الاقتصادي" – كلية الاقتصاد والعلوم السياسية – جامعة القاهرة.

إذا كان j,i ترمز لسنة الأساس وسنة المقارنة على الترتيب ، وقيمة الظاهرة محل الدراسة في سنة الأساس هي x_i وفي سنة المقارنة x_j . ففي هذه الحالة يمكن أعتبار الرقم القياسي لقيمة الظاهرة هو منسوب التغير في قيمة الظاهرة ، فإذا رمزنا للرقم القياسي لسنة المقارنة j بالنسبة لسنة الأساس i بالرمز $I_{j,i}$ حيث :

$$I_{i,j} = \frac{x_j}{x_i} \times 100 \qquad (1)$$

مثال (١)

إذا كان سعر كيلو السكر 20 قرش سنة ١٩٧٥ وأرتفع ليصل إلى 60 قرش سنة ١٩٨٠ . أوجد الرقم القياسي لسعر كيلو السكر في سنة ١٩٨٠ بالنسبة لسنة ١٩٧٥ ثم عقب على النتائج.

الحل :

في هذه الحالة نجد أن سنة ١٩٧٥ هي سنة الأساس وبالتالي فإن $x_i = 25$ قرش ، كذلك سنة ١٩٨٠ هي سنة المقارنة وبالتالي $x_j = 60$ وبالتالي فإن الرقم القياسي في سنة المقارنة سنة ١٩٨٠ بالنسبة لسنة الأساس ١٩٧٥ هو :

$$I_{1980,1975} = \frac{x_{1980}}{x_{1975}} \times 100$$

$$= \frac{60}{25} \times 100 = 240\%$$

وبما أن الرقم القياسي 240% ، أى أن سعر كيلو السكر في سنة ١٩٨٠ زاد عن سعره سنه ١٩٧٥ بنسبة 140% من سعره سنة ١٩٧٥ .

ثانياً : الرقم التجميعي البسيط

إذا كان لدينا عدد n من السلع (أو الأشياء) وكان سعر (أو قيمة) السلعة رقم k سنة المقارنة هو x_{jk} وسعرها سنة الأساس هو x_{ik} بحيث $k = 1,2,3,\ldots,n$ ، فإن الرقم التجميعي البسيط S يعرف على النحو التالي :

$$S = \frac{\sum_{k=1}^{n} x_{jk}}{\sum_{k=1}^{n} x_{ik}} \times 100 \qquad (2)$$

مثال (٢)

فيما يلي أسعار 5 أنواع من السلع في سنتي ١٩٦٠ ، ١٩٧٠

الرقم	السلعة	سعر الوحدة سنة ١٩٧٠	سعر الوحدة سنة ١٩٦٠

1	القطن (قنطار)	40	10
2	القمح (إردب)	18	8
3	الصوف (قنطار)	28	8
4	الورق (طن)	33	13
5	الخضراوات (طن)	25	15

أوجد الرقم القياسي التجميعي البسيط للمجموعة السلعية المذكورة أعلاه.

الحل :

بما أن الرقم التجميعي البسيط S حيث :

$$S = \frac{\sum_{k=1}^{n} x_{jk}}{\sum_{k=1}^{n} x_{ik}} \times 100$$

$$= \frac{40+18+28+33+25}{10+8+8+13+15} \times 100$$

$$= \frac{144}{54} \times 100 = 266.67\%$$

ومن أهم عيوب الرقم التجميعي البسيط أنه يتأثر بوحدة القياس المحددة للسعر ، فمثلاً نجد أن وحدة قياس السلعة رقم (5) الخضراوات بالطن ، فإذا استخدمنا سعر الكيلو بدلاً من الطن فإن الرقم القياسي سوف يتغير ، وهذا يعني أن الرقم التجميعي البسيط هو في الواقع رقم مرجح* بأوزان لا تعكس مطلقاً الأهمية الفعلية للسلع من حيث مدى تأثيرها على مستوى الأسعار.

ثالثاً : الوسط البسيط المناسب

إذا فرضنا أن s_k هو منسوب السلعة رقم k بحيث k = 1,2,3,…,n من السلع حيث أن :

$$S_k = \frac{\sum_{k=1}^{n} x_{jk}}{\sum_{k=1}^{n} x_{ik}} \times 100 \qquad (3)$$

فإن s_k يقيس التغير في سعر السلعة k ، وبالتالي فإن استخدام أي متوسط لمناسيب أسعار السلع المختلفة يجب أن يعطي مقياساً للتغير الكلي في أسعار هذه السلع بين فترتي الأساس والمقارنة ومن أهم المتوسطات للمناسيب :

---

* نفس المرجع السابق.

١- الوسط الحسابي للمناسيب وسوف نرمز له بالرمز S حيث :

$$S = \frac{\sum_{k=1}^{n} s_k}{n} \times 100 \qquad (4)$$

٢- الوسط الهندسي البسيط للمناسيب وسوف نشير له بالرمز S' حيث :

$$S' = \sqrt[n]{s_1 \times s_2 \times ....\times s_n} \qquad (5)$$

مثال (٣)

فيما يلي بيانات عن أسعار ثلاثة سلع في سنتي الأساس والمقارنة

السعر سنة المقارنة x_{jk}	السعر سنة الأساس x_{ik}	رقم السلعة
2	1	(1)
5	4	(2)
3	3	(3)

المطلوب :

١- أحسب الوسط الحسابي للمناسيب.

٢- أحسب الوسط الهندسي للمناسيب.

الحل

نوجد مناسيب الأسعار في سنة المقارنة بالنسبة لسنة الأساس على النحو التالي :

$$S_1 = \frac{2}{1} = 2 \quad , \quad S_2 = \frac{5}{4} = 1.25 \quad , \quad S_3 = \frac{3}{3} = 1$$

وبالتالي فإن الوسط الحسابي للمناسيب S يصبح على النحو التالي :

$$S = \frac{\sum_{k=1}^{3} s_k}{3} \times 100$$

$$= \frac{2 + 1.25 + 1}{3} \times 100 = \frac{4.25}{3} \times 100 = 141.66\%$$

٢- كذلك الوسط الهندسي للمناسيب S' يصبح على النحو التالي :

$$S' = \sqrt[3]{s_1 \times s_2 \times s_3} = \sqrt[3]{2 \times 1.25 \times 1}$$

$$= 1.357 \times 100 = 135.7\%$$

ومما سبق يتضح أن الرقم التجميعي البسيط يتأثر بوحدة قياس السعر ، بالإضافة أنه يلاحظ أن الرقم التجميعي البسيط ما هو إلا وسط حسابي مرجح لمناسيب أسعار السلع مرجحة بأسعار السلعة في سنة الأساس.

وبما أن

$$S_k = \frac{x_{jk}}{x_{ik}} \qquad (6)$$

$$S_k \times (x_{ik}) = x_{jk} \qquad (7)$$

وبما أن الرقم التجميعي البسيط

$$S_k = \frac{\sum_{k=1}^{n} x_{jk}}{\sum_{k=1}^{n} x_{ik}} \times 100 \qquad (8)$$

وبالتعويض في أسعار السلع في سنة المقارنة x_{jk} في المعادلة (8) بقيمتها في المعادلة (7) نجد أن الرقم التجميعي البسيط هو :

$$S_k = \frac{\sum_{k=1}^{n} s_k x_{jk}}{\sum_{k=1}^{n} x_{ik}} \times 100 = \frac{1}{\sum_{k=1}^{n} x_{ik}} \left\{ \sum_{k=1}^{n} s_k x_{jk} \right\} \times 100 \qquad (9)$$

من المعادلة (9) يتضح أن الرقم التجميعي البسيط ما هو إلا وسط حسابي للمناسيب مرجح بأسعار السلع في سنة الأساس (i) ، وبالتالي فإن أسعار سنة الأساس سوف يتحكم في تحديد قيمة الرقم القياسي في هذه الحالة ولتلافي ذلك بالإضافة إلى أهمية حساب مقاييس تعكس الأهمية النسبية لكل سلعة داخل المجموعة ، لذلك كان الاتجاه إلى تكوين أرقام قياسية تجميعية مرجحة ، ونظراً لأن الكمية المستهلكة (أو المنتجة) من السلعة تتناسب مع الأهمية النسبية للسلعة كذلك تبين مدى تأثير سعر هذه السلعة على باقي سلع المجموعة ، ومن هنا كان الاتجاه إلى استخدام الكميات كأوزان للترجيح.

فإذا فرضنا أن Q_{jk} هي الكميات المستهلكة من السلعة k في سنة المقارنة (j) ، Q_{ik} هي الكمية المستهلكة من السلعة k في سنة الأساس (i).

في سنة ١٨٦٤ استخدم العالم لاسبير Laspeyre الكميات في سنة الأساس كأوزان للترجيح ، وعرف الرقم القياسي التجميعي المرجح بكميات الأساس (وأحياناً يسمى برقم لاسبير نسبة إلى العالم لاسبير) على النحو التالي :

$$L = \frac{\sum_k X_{jk} Q_{ik}}{\sum_k X_{ik} Q_{ik}} \times 100 \qquad (10)$$

ويلاحظ أن رقم لاسبير في (10) متحيز للكميات في سنة الأساس لذلك في سنة ١٨٧٤ استخدم باش Paasche الكميات في سنة المقارنة لأوزان الترجيح ، وعرف الرقم التجميعي المرجح بكميات المقارنة (وأحياناً يسمى برقم باش نسبة إلى العالم باش) على النحو التالي :

$$P = \frac{\sum_k X_{jk} Q_{jk}}{\sum_k X_{ik} Q_{jk}} \times 100 \qquad (11)$$

ويلاحظ أن رقم باش في (11) متحيز للكميات في سنة المقارنة.

وللتغلب على التحيز في رقمي لاسبير وباش قدم العالم فيشر Fisher في سنة ١٨٩٢ رقم قياسي يسمى بالرقم القياسي الأمثل ، وهذا الرقم عبارة عن الوسط الهندسي لرقمي لاسبير وباش ، وعرف هذا الرقم على النحو التالي :

$$F = \sqrt{\frac{\sum_k X_{jk} Q_{ik}}{\sum_k X_{ik} Q_{ik}} \times \frac{\sum_k X_{jk} Q_{jk}}{\sum_k X_{jk} Q_{jk}}} \times 100 \qquad (12)$$

مثال (٤)

الجدول التالي يوضح أسعار وكميات 4 سلع تقوم بإنتاجها إحدى الشركات في عام ١٩٨٠ ، ١٩٩٠ .

رقم السلعة	سنة ١٩٨٠		سنة ١٩٩٠	
	الأسعار x_i	الكميات Q_i	الأسعار x_j	الكميات Q_j
(1)	50	100	100	100
(2)	80	120	120	200
(3)	90	300	110	500
(4)	180	400	150	300

باعتبار أن سنة ١٩٩٠ هي سنة المقارنة ، وسنة ١٩٨٠ هي سنة الأساس.

المطلوب :

١- أحسب الرقم القياسي المرجح للأسعار مرجح بكميات سنة الأساس (رقم لاسبير).
٢- أحسب الرقم القياسي المرجح للأسعار مرجح بكميات سنة المقارنة (رقم باش).
٣- أحسب الرقم القياسي الأمثل (رقم فيشر).
٤- وضح العلاقة بين الأرقام الثلاثة السابقة.

الحل

١- بما أن رقم لاسبير يساوى L حيث :

$$L = \frac{\sum_k X_{jk} Q_{ik}}{\sum_k X_{ik} Q_{ik}} \times 100$$

$$= \frac{(100\times100)+(120\times120)+(110\times300)+(150\times400)}{(50\times100)+(80\times120)+(90\times300)+(180\times400)} \times 100$$

$$= \frac{117400}{113600} \times 100 = 103.35\%$$

٢- وبما أن رقم باش يساوى P حيث :

$$P = \frac{\sum_k X_{jk} Q_{jk}}{\sum_k X_{ik} Q_{jk}} \times 100$$

$$= \frac{(100\times100)+(120\times200)+(110\times500)+(150\times300)}{(50\times100)+(80\times200)+(90\times500)+(180\times300)} \times 100$$

$$= \frac{164000}{135000} \times 100 = 121.48\%$$

٣- بما أن فيشر يساوى F حيث :

$$F = \sqrt{\frac{\sum_k X_{jk} Q_{ik}}{\sum_k X_{ik} Q_{ik}} \times \frac{\sum_k X_{jk} Q_{jk}}{\sum_k X_{jk} Q_{jk}}} \times 100$$

$$= \sqrt{\frac{117400}{113600} \times \frac{164000}{135000}} \times 100 = 112.05\%$$

٤- من (١) ، (٢) , (٣) يتضح أن :

$$L \le F \le P$$

أي أن رقم فيشر يعتبر رقم أكبر من رقم لاسبير وأقل من رقم باش

تمرينات **Exercises**

١- الجدول التالي يوضح أسعار 4 سلع في سنتي ١٩٨٥، ١٩٩.

سعر الوحدة في سنة ١٩٩٥	سعر الوحدة في سنة ١٩٨٥	رقم السلعة
5	2	(1)
7	3	(2)
13	10	(3)
5	10	(4)

المطلوب :

١- أوجد الرقم القياسي لكل سلعة في سنة ١٩٩٥ بالنسبة لسنة ١٩٨٥ ثم عقب على الناتج.

٢- أوجد الرقم التجميعي البسيط ثم عقب على الناتج.

٣- أوجد كل من الوسط الحسابي والوسط الهندسي البسيط للمناسيب ثم عقب على الناتج.

٢- الجدول التالي يوضح أسعار وكميات 3 سلع مطروحة في السوق في إحدى المدن في سنتي ١٩٩٢، ١٩٩٧.

سنة ١٩٩٧		سنة ١٩٩٢		السلعة
الكميات (بالوحدة)	سعر الوحدة (بالجنية)	الكميات (بالوحدة)	سعر الوحدة (بالجنية)	
40000	30	50000	25	القميص
100000	300	290000	250	البذلة
225000	97	300000	55	الحذاء

باعتبار أن سنة ١٩٩٢ هي سنة الأساس وسنة ١٩٩٧ هي سنة المقارنة ، أوجد :

١- الرقم القياسي المرجح للأسعار بكميات سنة الأساس (رقم لاسبير).

٢- الرقم القياسي المرجح للأسعار بكميات سنة المقارنة (رقم باش).

٣- الرقم القياسي الأمثل (رقم فيشر).

ملحق (٣)
توزيع ذات الحدين

الجدول التالي يعطى قيم دالـة الإحتمـال $p_{r(X=x)}$ عنـد عدد n مـن المحـاولات وبأحتمـال نجـاح فـى المحاولـة يساوى Π حيث أن :

$$p_r\left(X=x\right)=C_x^n\Pi^x\left(1-\Pi\right)^{n-x}$$

تابع توزيع ذات الحدين

الجدول التالى يعطى قيم الدالـة التراكميـة $p_r\left(X\le x\right)$ لتوزيـع ذات الحدين عند n من المحـاولات وبأحتمـال نجاح فى المحاولة يساوى Π حيث أن :

$$p_r\left(X\le x_y\right)=\sum_{i=1}^{y}C_{x_i}^n\Pi^{x_y}\left(1-\Pi\right)^{n-x_i}$$

ملحق (٤)
توزيع بواسون

الجدول التالي يعطى قيم دالة الإحتمال $p_{r(X=x)}$ عند القيم المختلفة لـ μ حيث أن :

$$p_r\left(X=x\right)=\frac{e^{-\mu}\left(\mu\right)^x}{x!}$$

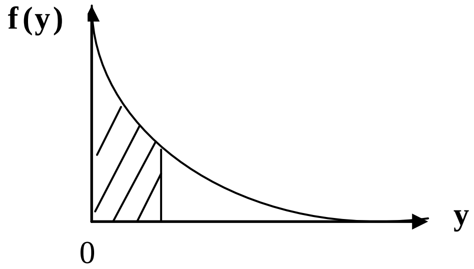
تابع توزيع بواسون

الجدول التالى يعطى قيم الدالة التراكمية $p_r\left(X\le x\right)$ لتوزيع بواسون حيث أن :

$$p_r\left(X=x_y\right)=\sum_{i=1}^{y}\frac{e^{-\mu}\left(\mu\right)^{x_i}}{x_i!}$$

ملحق (٥)
التوزيع الأسي السالب

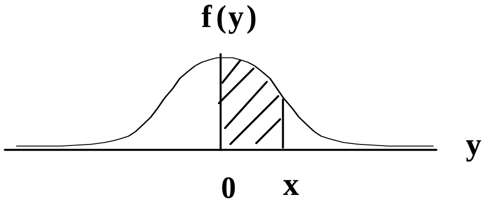
الجدول التالى يعطى قيم الاحتمالات المختلفة من $y = 0$ إلى $y = x$ عند القيم المختلفة للمعلمة λ حيث أن :



$$p_r(0 \le y \le x) = \int_0^x \lambda e^{-y}$$

ملحق (٦)
التوزيع المعتاد القياسي

الجدول التالى يعطى قيم الاحتمالات المختلفة من $y = -\infty$ إلى $y = z$ حيث أن y متغير يتبع التوزيع المعتاد القياسي:



$$p_r(-\infty < y < z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

قائمة المراجع

أولاَ:المراجع العربية

١- د.ربيع زكي عامر (١٩٨٩) : "تحليل الانحدار – أساليبه وتطبيقاته العملية باستخدام البرنامج الجاهزSPSS/Pct".معهد الدراسات والبحوث الاحصائية - جامعة القاهرة.

٢- أ.د سمير عاشور ، أ.د سامية أبو الفتوح (١٩٩٤) : "مقدمة لنظرية العينات" . معهد الدراسات والبحوث الاحصائية - جامعة القاهرة.

٣- أ.د سمير عاشور (١٩٨٦): "مقدمة في الاحصاء التحليلي".معهد الدراسات والبحوث الاحصائية - جامعة القاهرة.

٤- أ.د عبد الله توفيق الهلباوي (١٩٨٧): "الاحصاء التطبيقي". مكتبة عين شمس – القاهرة.

٥- أ.د عفاف على الدش (١٩٩٤): "الاحصاء التطبيقي للتجاريين" . الطبعة الخامسة - مكتبة عين شمس – القاهرة – رقم الايداع ٢٠٠٠/١٣٢٢٥.

٦- أ.د عفاف على الدش (١٩٨٧): "بحوث العمليات واتخاذ القرار". مكتبة عين شمس – القاهرة – رقم الايداع ١٩٨٧/٨٨٨٧.

٧- أ.د عفاف الدش (١٩٩٦) : "الاحصاء وصناعة القرارات"- مكتبة عين شمس – القاهرة – رقم الايداع ٩٦/٩٣٦٤.

٨- أ.د عفاف الدش (١٩٩٤) : "الرياضيات وصناعة القرارات"- مكتبة عين شمس – القاهرة – رقم الايداع ٩٤/٨٦٨٢.

٩- أ.د محمد عبد السميع عناني (١٩٩٣) : "مبادئ الاقتصاد القياسي النظري والتطبيقي" . الطبعة الثانية – جامعة الزقازيق – جمهورية مصر العربية.

١٠- مواري د.شيجل (١٩٨٩) : "الاحصاء"، سلسلة ملخصات شوم – نظريات ومسائل الدار الدولية للنشر والتوزيع- القاهرة.

١١- أ.د فتحي محمد على (١٩٨٤) : "الاحصاء المتقدم" . مكتبة عين شمس – القاهرة

ثانياً : المراجع الأجنبية

12- Anderson , T.W.(1971) : "An Introduction to Multivariate Statistical Analysis" , John Wiley & Sons , Inc. , New York.

13- David , J.B. , Andrew , F.F. and Sally I.M.C. (1991) : "Statistical Technique for Manpower planning" , Second Edition , John Wiley & Sons , New York.

14- Frank , S. (1993) : "Applied Mathematics for Business , Economic and The Social Sciences" , McGraw-Hill International Edition , New York.

15- Forgiel , Director (1968) : "The Statistics Problems Solver" , Research and Education Association , New York.

16- Gallant , A.R. (1987) : "Nonlinear Statistical Models" , John Wiley & Sons , Inc. , New York.

17- Hamburg , M. (1970) : "Statistical Analysis for Decision Making" , Harsart , Brace & World , Inc. , New York.

18- Harry Frank and Steven Althoes (1994) : "Statistics – Concepts and Applications" , Cambridge University Press.

19- Heinz Kohler (1994) : "Statistics for Business and Economics" – Third Edition , Harper Collins College Publishers , U.S.A.

20- Lawrence L. lapin (1994) : "Quantitative Methods for Business Decision With Cases" , Sixth Edition , The Dryden press , New York.

21- Kendal , M.G. and Studart , A. (1982) : "The Advanced Theory of Statistics" , Vol. 1 , 2 , Charles Griggin & Company Limited , London.

22- Keller and Warrack (2000) : "Statistics for Management and Economics" Duxbury , New York.

23- Mcallister , H.E. (1975) : "Elements of Business and Economic Statistics : learning by Objectives" , John Wiley & Sons Inc. , New York.

24- Mood , A.S. and Graybill , F.A. (1963) : "Introduction to The Theory of Statistics" , McGraw-Hill Book Company , New York.

25- Prem S. Mann (1995) : "Statistics for Business and Economics" , John Wiley & Sons , Inc. , New York.

26- Spiegel , M.R. (1972) : "Schaums Outline of Theory and Problem of Statistics" , McGraw–Hill , International Book Company , New York.

27- Shao & Shao (1990) : "Mathematical for Management and Finance" , Sixth Edition , South–Western Publishing Co. , West Chicago.

28- William A. Spurr and Charles P. Bonini (1973) : "Statistical Analyses for Business Decisions" . Richard D. Irwin , Inc. , Landon.

29- Robert D. Mason and Douglas A. Lind (1996) : "Statistical Technique in Business and Economics" , Ninth Edition , IRWWN , Landon.

30- Paul Newhold , Willam L. Carlson , and Betty M. Thorme (2003) : "Statistics for Business and Economics" , Inc , New Jersey.

كتب للمؤلفة

- الإحصاء التطبيقي للتجاريين – الجزء الأول – الطبعة الرابعة (٢٠٠٠م) – جهاز نشر وتوزيع الكتاب الجامعي – جامعة حلوان – القاهرة.
- رياضيات الاستثمار (٢٠٠٠) - جهاز نشر وتوزيع الكتاب الجامعي – جامعة حلوان – القاهرة.
- رياضيات الأعمال – الطبعة الثانية (١٩٩٩م) – جهاز نشر وتوزيع الكتاب الجامعي – جامعة حلوان – القاهرة.
- الإحصاء وصناعة القرارات – الجزء الثاني (١٩٩٦م) – مكتبة عين شمس – القاهرة.
- مقدمة في الرياضيات وتطبيقاتها الإنسانية (١٩٩٦م) – مكتبة عين شمس – القاهرة.
- الرياضيات وصناعة القرارات (١٩٩٤م) – مكتبة عين شمس – القاهرة.
- نماذج الانحدار (١٩٩٠م) – مكتبة عين شمس – القاهرة.
- بحوث العمليات واتخاذ القرارات (١٩٨٧م) – مكتبة عين شمس – القاهرة.

رقم الإيداع :

الترقيم الدولي :