

پیشووا برآمد
عبداللہ

پیشووا کے
خاندان

الكلية الجامعية للعلوم والتكنولوجيا - خان يونس



القسم الأكاديمي

مبادئ الإحصاء والإحصاء الحيوي



إعداد
أ. إياد محمد الهوبي

راجعته أ. عماد الأغا أ. وائل الحنفي

الفهرس

الموضوع

رقم الصفحة

المقدمة

الفصل الأول: الإحصاء و الحياة

1	المجتمع والعينة.....
2	الإحصاء الوصفي والإحصاء الاستنتاجي.....
3	أساليب جمع البيانات.....
4	طرق جمع البيانات.....
5	أنواع العينات.....
6	أنواع البيانات.....
14	المتغيرات و الثوابت.....
15	تدريج القياس.....
18	تمارين.....
21	

الفصل الثاني: تبويب البيانات وعرضها بيانياً

25	تبويب البيانات النوعية.....
26	الأعمدة البيانية البسيطة.....
28	الخط المنكسر.....
30	الدوائر البيانية.....
32	الأعمدة البيانية المزدوجة.....
35	الخط المنكسر المتعدد.....
37	تبويب البيانات الكمية.....
39	المدرج التكراري.....
43	المنحنى التراكمي الصاعد.....
44	المنحنى التراكمي الهابط.....
46	تمارين.....
49	

الفصل الثالث: مقاييس النزعة المركزية

55	مقاييس النزعة المركزية للبيانات الأولية.....
56	الوسط الحسابي.....
56	الوسط الحسابي المشذب.....
60	الوسط الحسابي المرجح.....
61	الوسيط.....
62	المنوال.....
65	المئينات والربيعات.....
67	مقاييس النزعة المركزية للتوزيعات التكرارية.....
70	الوسط الحسابي.....
70	الوسيط.....
72	

74	 المنوال
76	 المئينات والربيعات
79	 الالتواء
81	 التحويلات الخطية ومقاييس النزعة المركزية
83	 تمارين
89	الفصل الرابع: مقاييس التشتت
90	 مقاييس التشتت للبيانات الأولية
90	 المدى
92	 الانحراف الربيعي
93	 الانحراف المتوسط
94	 التباين
95	 الانحراف المعياري
98	 التباين المرجح
99	 مقاييس التشتت للتوزيعات التكرارية
99	 المدى
100	 الانحراف الربيعي
102	 الانحراف المتوسط
103	 التباين
105	 معامل الاختلاف
107	 التحويلات الخطية ومقاييس التشتت
109	 تمارين
117	الفصل الخامس: الارتباط والانحدار الخطي البسيط
118	 لوحة الانتشار
122	 معامل ارتباط بيرسون
124	 معامل ارتباط سبيرمان
130	 الانحدار الخطي البسيط
134	 معامل التحديد
141	 تمارين
147	الفصل السادس: مقدمة في الاحتمالات
148	 فضاء العينة
149	 الحدث
156	 الحوادث المتنافية
157	 التكرار النسبي والاحتمال
161	 متمم الحدث
163	 الفرق بين حدثين
165	 اتحاد حدثين
170	 الحوادث المستقلة
176	 الاحتمال المشروط
182	 نظرية التجزئة

184	 نظرية بيز
186	 تمارين
193	الفصل السابع: المتغيرات العشوائية والتوزيعات الاحتمالية
197	 التوزيعات الاحتمالية المنفصلة
199	 التوزيع التراكمي للمتغير العشوائي المنفصل
201	 التوقع الرياضي للمتغير العشوائي المنفصل
203	 التباين للمتغير العشوائي المنفصل
206	 توزيع ذي الحدين
211	 توزيع بواسون
216	 التوزيعات الاحتمالية المتصلة
218	 التوزيع التراكمي للمتغير العشوائي المتصل
221	 التوقع الرياضي للمتغير العشوائي المتصل
222	 التباين للمتغير العشوائي المتصل
223	 التوزيع الطبيعي
237	 تمارين
245	الفصل الثامن: الإحصاءات الحيوية
246	 تعداد السكان
247	 إحصاءات المواليد
250	 إحصاءات الوفيات
254	 إحصاءات الهجرة
255	 إحصاءات الجنسية
255	 إحصاءات الحالة الزوجية
255	 إحصاءات الحالة الاقتصادية للسكان
257	 تقدير السكان بين سنوات التعداد
261	 تمارين

الملاحق والمراجع

مقدمة

الحمد لله الذي زين قلوب أوليائه بأنوار الوفاق، وسقى أسرار أحبائه شراباً لذيق المذاق، وألزم قلوب الخائفين الوجّل والإشفاق، فلا يعلم الإنسان في أي الدواوين كتب ولا في أيّ الفريقين يساق، فإنّ سامح فبفضله، وإنّ عاقب فبعدلّه، ولا اعتراض على الملك الخلاق. وأشهد أن لا إله إلا الله، وحده لا شريك له، له الملك وله الحمد، وهو على كل شيء قدير، إله عزّ من اعتز به فلا يضام، وذللّ من تكبر عن أمره ولقي الآثام.

وأشهد أن سيدنا وحبيبنا وشفيعنا محمداً عبد الله ورسوله، وصفيه من خلقه وحبيبه، خاتم أنبيائه، وسيد أصفیائه، المخصوص بالمقام المحمود، في اليوم المشهود، الذي جُمع فيه الأنبياء تحت لوائه ثمّ أما بعد.

يسعدني أن أقدم الطبعة الأولى من كتاب مبادئ الإحصاء والإحصاء الحيوي إلى أبنائي الطلبة في كل مكان، آملاً أن أكون موفقاً في عرض هذه المادة العلمية بأسلوب سهل ووضوح تام.

تضمن الكتاب تسعة فصول، تقدم كل فصل ذكراً للأهداف المرجوة واختتم الفصل بخلاصة ما فيه من قوانين وتعريفات. وفي نهاية كل فصل العديد من التمارين المتنوعة، وفي ثناياه العديد من الصور والرسومات التوضيحية والتي استخدمت برنامجي SPSS و R الإحصائيين للحصول عليها. وحرصت على أن يكون هذا الكتاب متطلباً سابقاً مناسباً لمساق الإحصاء التطبيقي من خلال التركيز على حساب المقاييس الإحصائية وتفسير النتائج التي نحصل عليها والتركيز على المتغيرات العشوائية وتوزيعاتها الاحتمالية ومفهوم التوقع.

يتحدث الفصل الأول عن الإحصاء والحياة، الفصل الثاني عن عرض البيانات بيانياً سواءً كانت أولية أو توزيعات تكرارية، الفصل الثالث يتحدث عن مقاييس النزعة المركزية، الفصل الرابع يتحدث عن مقاييس التشتت، الفصل الخامس يتناول موضوع الارتباط والانحدار بشكل مبسط، الفصل السادس يتحدث مقدمة في الاحتمال مع ذكر بعض القوانين الأساسية، الفصل السابع يتحدث عن المتغيرات العشوائية والتوزيعات الاحتمالية المنفصل منها والمتصل وعن مفهوم التوقع وبعض النظريات الخاصة به أما الفصل الثامن والأخير فتم تخصيصه لطلبة العلوم الطبية وهو يتحدث عن بعض الإحصاءات الحيوية الهامة.

أقدم خالص شكري للزميلين و الأخوين أ. عماد سعيد الأغا و أ. وائل محمد الحنفي لما قدماه لي من أمثلة لإثراء الكتاب ومراجعته القيمة و المفيدة.

أتمنى لجميع طلابنا وطالباتنا التفوق والنجاح سائلاً المولى عز و جل أن يهديهم سواء السبيل، إنه ولي ذلك و القادر عليه.

أ.أياد محمد الهوبي

ذو الحجة 1438هـ

الموافق أيلول 2017م

الإحصاء والحياة Statistics and life

1

في هذا الفصل سنتعرف على بعض المفاهيم والتعريفات الهامة في علم الإحصاء، وسنتحدث باختصار عن قسمي علم الإحصاء وأساليب جمع البيانات وبعض طرق جمع البيانات، ومن ثم سنتحدث عن العينات الاحتمالية والعينات غير الاحتمالية ونذكر بعض طرق المعاينة لهما.

سنتحدث كذلك عن المتغيرات والثوابت ونعرف بعض أنواع المتغيرات الهامة مع ضرب أمثلة على كل منها، ونختم الفصل بالتحدث عن تدرج القياس ونوضح كيفية تطبيق معايير خاصة لتصنيف المتغيرات وفق التدرج المناسب.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً على:

- تعريف الإحصاء، المجتمع، العينة.
- ذكر أساليب وطرق جمع البيانات مع التوضيح.
- ذكر أنواع العينات وطرق المعاينة لكل منها مع التوضيح.
- ذكر أنواع المتغيرات مع ذكر أمثلة على كل منها.
- تصنيف المتغيرات حسب تدرج القياس المناسب.

الإحصاء هو أحد فروع الرياضيات التطبيقية وهو العلم الذي يبحث في جمع البيانات والمعلومات والحقائق الخاصة بمختلف الظواهر وتسجيلها في صورة رقمية وتصنيفها في جداول منظمة، وتمثيلها بيانياً، وتحليل البيانات واستخلاص النتائج منها واتخاذ القرارات المناسبة.

ويطبق علم الإحصاء في مجالات واسعة وعديدة وفي جميع التخصصات الأدبية منها والعلمية، فلإحصاء تطبيقات في علم الاجتماع، التعليم، الطب، الاقتصاد، دراسات السكان، العلوم الطبيعية، علم النفس، الفلك وكثير من العلوم الأخرى، حيث يطبق على سبيل المثال لا الحصر في تحليل الارتباط والانحدار حيث ينصب الاهتمام هنا في دراسة العلاقة بين الظواهر المختلفة، وفي حالة وجود علاقة مناسبة بين هذه الظواهر يمكن العمل على إيجاد نماذج رياضية تربط هذه الظواهر معاً وبالتالي التنبؤ بقيمة ظاهرة إذا علمت قيم الظواهر المرتبطة بها.

المجتمع والعينة

من المفاهيم الهامة والرئيسية في علم الإحصاء مفهوم المجتمع **Population** ومفهوم العينة **Sample**، فالمجتمع الإحصائي يتكون من مجموعة العناصر أو الأفراد الذين ستجرى عليهم الدراسة الإحصائية وسنسمي هذه العناصر مفردات أو مشاهدات، وقد يكون عدد هذه المشاهدات محدود وقد يكون غير محدود بمعنى أن المجتمع يحتوي على عدد لانهائي من المشاهدات.

وقبل الشروع بعملية جمع البيانات لابد من تعريف المجتمع الإحصائي تعريفاً جيداً وتعيين الحدود الصريحة لعملية جمع البيانات. وأي مجموعة جزئية من المجتمع يطلق عليها اسم عينة، حيث يتم تحديد حجمها وكيفية سحبها من المجتمع من قبل الباحث حسب خواص المجتمع وما يراه الباحث مناسباً نظراً لظروف الدراسة والإمكانات المتاحة.

الإحصاء الوصفي والإحصاء الاستنتاجي

ينقسم علم الإحصاء إلى قسمين رئيسيين:

1- الإحصاء الوصفي Descriptive Statistics

فعندما يقوم مدير المدرسة بتسجيل عدد الطلبة في كل فصل بمدرسته ومن ثم عرض هذه الأعداد على شكل مستطيلات أو جدول فإنه يكون قد استعمل الإحصاء الوصفي في هذه الحالة.

2- الإحصاء الاستدلالي Inferential Statistics

هو مجموعة من الأساليب الإحصائية التي تستخدم لتحليل البيانات التي يتم الحصول عليها من عينة مسحوبة من مجتمع الدراسة وتفسيرها، حيث نكون قادرين بعد ذلك على اتخاذ القرارات المناسبة.

مما سبق يمكن تعريف علم الإحصاء تعريفاً شاملاً يوضح أقسامه والخطوات التي تمر بها عملية التحليل الإحصائي كما يلي:

تعريف

الإحصاء: هو رياضيات جمع البيانات وتنظيمها وتحليلها وتفسيرها

أساليب جمع البيانات

إذا أراد باحث جمع بيانات حول ظاهرة معينة فإنه يقوم بتحديد مجتمع الظاهرة تحديداً جيداً ثم يحدد أسلوباً لجمع البيانات ومن ثم يختار طريقة تناسب أهداف الدراسة وأهميتها والتكاليف والوقت المتاح. هناك أسلوبان لجمع البيانات نذكرهما بشيء من الاختصار وبعدها نتطرق لبعض طرق جمع البيانات.

1- أسلوب المسح الشامل Comprehensive Survey

حيث تجمع البيانات من جميع أفراد المجتمع. ويعتبر أسلوب المسح الشامل أفضل أسلوب لجمع البيانات وذلك لأنه يزودنا ببيانات كاملة تمثل كل المجتمع حول الظواهر التي تهم الباحث، لكن الصعوبات المالية والفنية واختيار الطريقة المناسبة والوقت الطويل تجعل من العسير استخدامها دائماً. وكذلك لا يمكن استخدام المسح الشامل عندما يكون المجتمع الإحصائي غير محدد أو كبيراً جداً مثل الثروات الطبيعية في البلد من غاز طبيعي وبتترول وأسماك.

وهناك حالات يتعذر فيها استخدام أسلوب المسح الشامل، فقد يؤدي استخدام هذا الأسلوب إلى هلاك المجتمع أو فساد، فمثلاً في حالات فحص إنتاج مزارع أو مصانع يؤدي هذا الأسلوب لفساد المجتمع، فلا يمكن فحص إنتاج مزرعة بيض فينتكسر كل البيض في المزرعة وإنما يكفي فحص عينة من هذا الإنتاج، ولا يمكن السحب من مريض كل دمه لفحصه لأن ذلك يؤدي لوفاته فنلجأ لسحب عينة من دمه وفحصها. ومن الأمثلة على ما يتعين فيه استخدام هذا الأسلوب دون غيره: التعداد السكاني في البلاد، متابعة الحالة الصحية في منطقة انتشار أوبئة.

2- أسلوب العينة Sampling

عند تعذر استخدام أسلوب المسح الشامل يلجأ الباحث لسحب عينة من مفردات المجتمع بطريقة عشوائية أو غير عشوائية وذلك حسب أهمية الدراسة ومستوى الدقة المطلوب والغرض من الدراسة. ويجتهد الباحث بأن تكون العينة ممثلة لهذا المجتمع تمثيلاً

جيداً حتى يعمم النتائج التي يحصل عليها على المجتمع بكل ثقة ويتخذ القرارات المناسبة دون تردد.

إن استخدام هذا الأسلوب من شأنه أن يوفر الوقت في جمع وتحليل البيانات وكذلك يوفر الجهد والتكلفة ويسهل عملية جمع البيانات.

طرق جمع البيانات

عند تحديد الأسلوب الأمثل لجمع البيانات يلجأ الباحث لاختيار طريقة تنفيذية مناسبة لجمع البيانات - سواءً اختار أسلوب المسح الشامل أو العينة - نذكر منها:

1- طريقة المقابلة الشخصية **Personal Interview**

وفيها يقوم الباحث أو أشخاص مدربون مؤهلون بإجراء مقابلات شخصية مع أفراد المجتمع أو العينة وطرح أسئلة معدة مسبقاً بدقة عليهم فتكون الإجابة عن هذه الأسئلة هي البيانات المطلوبة أو يتم استخراج البيانات المطلوبة منها.

2- طريقة المكالمات الهاتفية **By Phone**

وهي طريقة تشبه كثيراً طريقة المقابلة الشخصية ولكنها تتم باستعمال الهاتف، ولهذه الطريقة مزايا كثيرة حيث يمكن إجراء الكثير من المكالمات الهاتفية دون اضطراب الباحث لمغادرة مكتبه أو استدعاء الأفراد لمكتبه، ومما يعاب على هذه الطريقة غياب عنصر التفاعل بين الباحث والأفراد.

3- طريقة الاستبانة **Questionnaire**

تعتبر هذه الطريقة من أكثر الطرق استخداماً في جمع البيانات، والاستبانة عبارة عن ورقة يعدها الباحث بدقة ومهارة وتحتوي على مجموعة من الأسئلة تمت صياغتها بإحكام بحيث يكون لكل منها إجابة محددة، وتوزع الاستبانات على الفئة المستهدفة ويقوموا بالإجابة عنها، ثم يتم جمعها وترميزها وتفريغها وتحليلها بطرق خاصة من قبل المحلل الإحصائي. ويمكن استخدام الاستبانة البريدية في حال كون أفراد المجتمع

أو العينة موزعين توزيعاً جغرافياً يصعب على الباحث جمعهم في مكان واحد أو السفر إليهم جميعاً، حيث يقوم بإرسال الاستبانات بواسطة مكاتب البريد أو البريد الإلكتروني.

4- طريقة الملاحظة المباشرة Direct Observation

حيث يضع الباحث أفراد المجتمع أو العينة نصب عينيه بطريقة مباشرة أو غير مباشرة مثل المراقبة بالكاميرات أو أجهزة إلكترونية خاصة، حيث تجمع البيانات في هذه الحالة عن طريق مراقبة الأفراد وتسجيل الملاحظات.

5- طريقة المصادر الرسمية Official Sources

تعتبر هذه الطريقة من أهم الطرق وأكثرها ثقةً لجمع البيانات حيث يتم جمع البيانات من جهات رسمية مسئولة تتميز بحرصها الشديد على توفير مستوى عالٍ من الدقة في البيانات التي توثقها، كالسجلات الحكومية التي تجمع البيانات فيها وتحدث بشكل مستمر وتحت إشراف مختصين.

أنواع العينات

تنقسم العينات إلى نوعين رئيسيين هما: العينات الاحتمالية والعينات غير الاحتمالية. وكل نوع من هذين النوعين يحتوي على عدة طرق لسحب العينة وسنذكر هنا بعض الطرق لكل نوع على سبيل المثال لا الحصر.

1- العينات الاحتمالية Probability Samples

هذا النوع من العينات يتم اختيار جميع مفرداتها بأسلوب احتمالي معين، وهذا يعني أن كل مشاهدة في المجتمع لابد أن يكون لها احتمال لا يساوي صفر لاختيارها في العينة. وليتم سحب عينة احتمالية من مجتمع ما لابد أن يكون هذا المجتمع محدداً ومعروفاً للباحث بشكل دقيق ليتمكن من التعامل مع جميع مشاهداته.

وسنعرض أهم طرق سحب العينات الاحتمالية وأكثرها استخداماً وسنركز على الشروط الواجب توافرها في المجتمع لسحب كل نوع من العينات وآلية سحب كل نوع.

1.1 المعاينة العشوائية البسيطة Simple random sampling.

يتم سحب عينة بهذه الطريقة من المجتمعات المتجانسة حيث يتم اختيار مفرداتها على أساس تساوي الفرص أمام جميع مشاهدات المجتمع أي جميع احتمالات مشاهدات المجتمع متساوية. ويتم سحب العينة بطريقتين هما طريقة القرعة وطريقة جداول الأرقام العشوائية (جدول (1) والمرفق في ملحق الكتاب).

ففي حالة كون المجتمع صغير الحجم نستخدم طريقة القرعة حيث يتم إعطاء كل مشاهدة رقماً متسلسلاً، ويكتب كل رقم على ورقة صغيرة ويتم طي جميع الأوراق ووضعها في صندوق وتخلط جيداً ثم يتم سحب ورقة من الصندوق وفتحها لمعرفة المشاهدة التي تم اختيارها ثم تطوى مرة ثانية وتوضع في الصندوق وذلك للحفاظ على تساوي احتمالات جميع مشاهدات المجتمع ويخلط جيداً ثم نسحب ورقة ثانية ونكرر العملية حتى نحصل على العينة كاملة، وفي حال سحب ورقة مرتين يتم إعادتها للصندوق ثم نعيد الخلط والسحب.

وفي حالة المجتمعات المتجانسة الكبيرة نلجأ لطريقة الجداول العشوائية حيث تكون أسهل تنفيذاً من طريقة القرعة. فإذا أردنا اختيار عينة ذات حجم n من مجتمع متجانس حجمه N نقوم بإعطاء كل مشاهدة من مجتمع الدراسة رقماً متسلسلاً من صفر إلى $N - 1$ (أو من 1 إلى N) ونجعل هذه الأرقام مكونة من نفس العدد من المنازل، فمثلاً إذا كنا نستخدم ثلاث منازل فنكتب العدد 3 على الصورة 003 وهكذا، ثم نستخدم جدولاً للأرقام العشوائية مثل جدول (1) المرفق في ملحق الكتاب أو نستعمل الحاسوب لطباعة جدول أرقام عشوائية أعداده لها نفس عدد المنازل في الأرقام المتسلسلة للمجتمع وتقع في مداها، ونقرأ منه عمودياً من العمود الأول حتى نحصل على حجم العينة المطلوب فمثلاً إذا كان أول عدد في العمود الأول 131 فيتم اختيار المشاهدة التي ترتيبها 131 من المجتمع وتعتبر هذه المشاهدة أول مشاهدة في العينة وهكذا، فإذا انتهى الباحث من قراءة العمود الأول ولم يحصل على حجم العينة المطلوب ينتقل للعمود المجاور وهكذا وإذا حصلنا على رقم مكرر نتركه وننتقل للرقم الذي يليه إلى أن نحصل على حجم العينة المطلوب.

وفي حال كبر حجم المجتمع بشكل يصعب على الباحث التعامل معه بجدول الأرقام العشوائية يتم استخدام الحواسيب في عملة اختيار العينات من هذه المجتمعات.

مثال 1.1: من مجتمع حجمه 50 مشاهدة وضح كيفية سحب عينة منه بحجم 5 مشاهدات.
الحل/

نفرض أننا رقمنا المجتمع بأرقام متسلسلة من 1 إلى 50، نبدأ بالعمود الأول من جدول (1) ونبدأ بالخانتين الأولى والثانية فقط من اليمين وذلك لأن الأرقام المتسلسلة للمجتمع مكونة من خانتين فنجد أن الأعداد التي تقع في المجال المطلوب هي:

11 12 31 14 15

فنختار من المجتمع المشاهدات التي لها تلك الأرقام.

2.1 المعاينة العشوائية المنتظمة Systematic Sampling

يتم سحب عينة بهذه الطريقة من المجتمع بطريقة منتظمة وذلك عندما يتعذر توفر

قائمة بمشاهدات المجتمع المتجانس أو لكبر حجم المجتمع.

فنفرض أن حجم المجتمع N ونريد سحب عينة منتظمة منه حجمها n فيتم ذلك باتباع الخطوات التالية:

1- يتم إعطاء المشاهدات في المجتمع أرقاماً متسلسلة من 1 إلى N ، ولاحظ هنا أن كل رقم من هذه الأرقام لا يدل على مشاهدة معينة كما هو الحال في العينة العشوائية البسيطة.

2- نقسم الأرقام المتسلسلة إلى فترات متساوية الطول وعدد هذه الفترات يساوي حجم العينة حيث يحسب طول الفترة والذي سنرمز له بالرمز a من العلاقة :

$$a = \frac{N}{n}$$

3- يتم اختيار مشاهدة من كل فترة فنجمع بذلك مشاهدات تساوي حجم العينة.

ويتم اختيار المشاهدة من كل فترة كما يلي:

نختار أول مشاهدة من أول فترة عشوائياً بطريقة القرعة ولتكن رتبة هذه المشاهدة b ، ثم نختار من الفترة الثانية المشاهدة التي رتبته $b + a$ ، ثم نختار من الفترة الثالثة المشاهدة التي رتبته $b + 2a$ وهكذا حتى نختار العينة المطلوبة من كافة الفترات.

ويمكننا صياغة رتبة المشاهدة التي سنختارها من كل فترة كما يلي:

رتبة المشاهد من الفترة j تساوي $b + (j - 1)a$ حيث $j = 1, 2, 3, \dots, n$

وهذا يمكننا من القول بأن رتبة المشاهدة المختارة في أي فترة تساوي رتبة المشاهدة في الفترة التي تسبقها مضافاً إليها طول الفترة، والمثال التالي يوضح آلية سحب العينة المنتظمة:

مثال 2.1: تواجد في قسم القبول والتسجيل في إحدى الجامعات 100 بطاقة تسجيل مساقيات

دراسية لأحد الفصول في أول يوم من بدء التسجيل، فقرر قسم شئون الطلبة سحب عينة من 20 بطاقة لدراسة عدد البطاقات التي فيها أخطاء ليتمكن من ارشاد الطلبة لعملية التسجيل. وضح كيفية اختيار هذه البطاقات بطريقة المعاينة العشوائية المنتظمة.

الحل/ لاحظ هنا أن $N = 100$ و $n = 20$ ، وبالتالي يكون طول الفترة كما يلي:

$$a = \frac{N}{n} = \frac{100}{20} = 5$$

لذلك فإن طريقة المعاينة المنتظمة تقضي بأن تختار بطاقة عشوائية من كل خمس بطاقات. فإذا اخترنا عشوائياً (بطريقة القرعة) من أول خمس بطاقات البطاقة الرابعة وذلك على سبيل المثال بمعنى أن رتبة المشاهدة الأولى من الفترة الأولى هي $b = 4$ ، فتكون رتبة البطاقة التي تليها تساوي $b + a = 4 + 5 = 9$ أي أننا سنختار البطاقة التاسعة من الفترة الثانية و تكون المشاهدة التي تليها من الفترة الثالثة هي التي رتبته $b + 2a = 4 + 10 = 14$ ، أو يمكننا مباشرة قول أن رتبة المشاهدة المختارة من الفترة الثالثة يساوي رتبة المشاهدة المختارة من الفترة الثانية وهي 9 مضافاً إليها طول الفترة و الذي يساوي 5 فتكون رتبة المشاهدة تساوي 14 وهكذا نستمر في إضافة 5 لكل رتبة نحسبها لنحصل على رتبة المشاهدة المختارة في الفترة

التالية وذلك إلى أن نصل إلى آخر فترة حيث سنجد أن رتبة الملاحظة المختارة منها هي التي رتبناها 99 والجدول التالي يوضح العملية بشكل مبسط:

20	<u>19</u>	18	17	16	15	<u>14</u>	13	12	11	10	<u>9</u>	8	7	6	5	<u>4</u>	3	2	1
40	<u>39</u>	38	37	36	35	<u>34</u>	33	32	31	30	<u>29</u>	28	27	26	25	<u>24</u>	23	22	21
60	<u>59</u>	58	57	56	55	<u>54</u>	53	52	51	50	<u>49</u>	48	47	46	45	<u>44</u>	43	42	41
80	<u>79</u>	78	77	76	75	<u>74</u>	73	72	71	70	<u>69</u>	68	67	66	65	<u>64</u>	63	62	61
100	<u>99</u>	98	97	96	95	<u>94</u>	93	92	91	90	<u>89</u>	88	87	86	85	<u>84</u>	83	82	81

لاحظ أن كل خلية تمثل فترة من الفترات ومن كل فترة تم اختيار الملاحظة الرابعة وهذا سبب تسمية العينة بالعينة المنتظمة. واضح أن الفرق بين هذه الطريقة في المعاينة وطريقة المعاينة العشوائية البسيطة هو أن الأخيرة يتم سحب جميع مفرداتها بشكل عشوائي أما المنتظمة يتم سحب أو لمفردة بطريقة عشوائية وباقي المفردات بشكل منتظم.

3.1 المعاينة العشوائية الطبقية Stratified Random Sampling

تستخدم هذه الطريقة إذا كان المجتمع غير متجانس ويمكن تقسيمه إلى أجزاء متجانسة يطلق عليها اسم طبقات وفق معيار معين. فمثلاً في مجتمع السكان نلاحظ أنه من حيث الجنس مكون من طبقتي الإناث والذكور. كما أننا إذا نظرنا إلى المجتمع نفسه من جهة المستوى العلمي نجد أنه مكون من عدة طبقات تبدأ من الأمي وتنتهي بشهادة الدكتوراه. يتم سحب عينة طبقية من المجتمع بتقسيم المجتمع إلى طبقات متجانسة وسحب من كل طبقة جزءاً من العينة المطلوبة يتناسب مع حجم الطبقة بالنسبة للمجتمع، ويتم سحب ذلك الجزء بطريقة المعاينة العشوائية البسيطة أو المعاينة العشوائية المنتظمة السابق ذكرهما.

فإذا كان لدينا مجتمع حجمه N مكون من k طبقة متجانسة حجوماً $N_1, N_2, \dots, N_k$ على الترتيب وأردنا اختيار عينة حجمها n من المجتمع بحيث يكون n_1 الجزء من العينة المسحوب من الطبقة الأولى، n_2 الجزء المسحوب من الطبقة الثانية وهكذا إلى أن نصل n_k وهو الجزء من العينة المسحوب من الطبقة الأخيرة.

ولاحظ أن:

$$N = N_1 + N_2 + \dots + N_k$$

$$n = n_1 + n_2 + \dots + n_k$$

ولإيجاد الجزء من العينة المسحوب من كل طبقة نطبق العلاقة التالية:

$$n_i = n \times \frac{N_i}{N}, i = 1, 2, 3, \dots, k$$

والمثال التالي يوضح كيفية حساب الجزء المطلوب سحبه من كل طبقة.

مثال 3.1: إذا كان طلاب السنة الأولى في إحدى الجامعات موزعين حسب الكليات كما هو

مبين في الجدول التالي:

الآداب	العلوم	الاقتصاد	الهندسة
200	150	180	120

وضح كيف نختار عينة حجمها 130 طالباً بطريقة المعاينة العشوائية الطبقية.

الحل/

نفرض أن كليات الآداب، العلوم، الاقتصاد والهندسة تمثل الطبقات الأولى، الثانية، الثالثة

والرابعة على الترتيب ونرمز لحجومها كما هو مبين بالجدول التالي:

الآداب	العلوم	الاقتصاد	الهندسة
$N_1 = 200$	$N_2 = 150$	$N_3 = 180$	$N_4 = 120$

وبالتالي يكون حجم المجتمع:

$$N = N_1 + N_2 + N_3 + N_4 = 200 + 150 + 180 + 120 = 650$$

ولما كان حجم العينة $n = 130$ طالب ونحن نريد أن تكون جميع الطبقات ممثلة بما يتناسب مع حجمها فإن الجزء المسحوب من كل طبقة كالتالي :

$$n_1 = 130 \times \frac{200}{650} = 40 \quad \text{من كلية الآداب:}$$

$$n_2 = 130 \times \frac{150}{650} = 30 \quad \text{من كلية العلوم:}$$

$$n_3 = 130 \times \frac{180}{650} = 36 \quad \text{من كلية الاقتصاد:}$$

$$n_4 = 130 \times \frac{120}{650} = 24 \quad \text{من كلية الهندسة:}$$

ويتم سحب كل جزء من طبقته بطريقة المعاينة العشوائية البسيطة أو المنتظمة كما أوضحنا سابقاً.

4.1 المعاينة العنقودية Cluster Sampling

يستخدم هذا النوع من الطرق في حالة المجتمعات المتجانسة وغير المتجانسة في حالة كون المجتمع ذو حجم كبير، حيث يقسم المجتمع إلى عناقيد ويتم اختيار عدد من العناقيد عشوائياً تمثل الوحدات الأولية، ومن كل عنقود يتم اختيار عينة بإحدى الطرق سابقة الذكر حيث يكون اختيار مفردات كل عينة هو الاختيار النهائي في طريقة المعاينة، ويتم جمع كل العينات المختارة من العناقيد لتشكل العينة العنقودية المطلوبة، وتسمى هذه الطريقة بطريقة المعاينة ذات المرحلة الأولى.

وقد تتعدد المراحل بين عملية الاختيار الأولى للعناقيد وبين عملية الاختيار الأخيرة للمفردات حيث يطلق في هذه الحالة على طريقة المعاينة: **المعاينة العشوائية متعددة**

المراحل Multi-Stage Random Sampling

فمثلاً إذا أردنا إجراء دراسة حول مستوى المعيشة للأسر في دولة فيمكن كمرحلة أولى اختيار مجموعة من المحافظات عشوائياً وكمرحلة ثانية اختيار مجموعة من المدن والقرى من كل محافظة تم اختيارها في المرحلة الأولى عشوائياً وكمرحلة ثالثة يمكن اختيار مجموعة من الأحياء في كل مدينة عشوائياً ثم أخيراً اختيار مجموعة من الأسر في كل حي عشوائياً ليتم تشكل العينة متعددة المراحل.

2- العينات غير الاحتمالية Non-Probability Samples

هذا النوع من العينات يتم اختياره بطرق لا تأخذ بعين الاعتبار الأسس العلمية للاختيار ودون النظر لمستوى الدقة في البيانات حيث يضع الباحث معايير الخاصة في عملية الاختيار ويكون غير مهتم بتعميم النتائج على المجتمع قيد الدراسة وإنما همه الأساسي جمع معلومات بشكل سريع وبأقل التكلفة لتكوين فكره سريعة عن ظاهرة معينة. ونذكر من طرق المعاينة غير الاحتمالية ما يلي:

1.2 المعاينة العرضية Accidental sample

وهي الطريقة التي يتم اختيار مفردات العينة فيها بشكل عرضي بسبب سهولة الوصول لهذه المشاهدات، كأن يستطلع الباحث رأي المواطنين عن برنامج تليفزيوني معين فيستوقف أي مواطن يقابله في الشارع ويسأله عن رأيه. وتعتبر هذه الطريقة مفيدة في حالة رغبة الباحث تكوين فكرة أولية عن الظاهرة قيد الدراسة.

2.2 المعاينة الحصصية Quota sampling.

في هذه الحالة يتم تقسيم المجتمع إلى طبقات، مثل تقسيم الموظفين في مؤسسة معينة إلى طبقة ذكور وطبقة إناث، ويتم تحديد عدد المفردات الذي سيؤخذ من كل طبقة بما يتناسب مع حجمها، ومن ثم يتم اختيار المشاهدات من كل طبقة بشكل عَرَضي وواضح أن الفرق بين هذه الطريقة وطريقة المعاينة العشوائية الطبقية هو طريقة اختيار المفردات.

أنواع البيانات Types of Data

من تعريف الإحصاء وجدنا أن الإحصاء هو العلم الذي يهتم بجمع البيانات، والجدير بالذكر هنا أن نوع البيانات وطريقة قياسها من أهم الأشياء التي تحدد التحليل الإحصائي المستخدم، ويمكن تصنيف البيانات لنوعين رئيسيين هما:

1- البيانات النوعية Qualitative Data

وهي بيانات غير رقمية كجنس الطالب ذكر أم أنثى، أو بيانات رقمية مرتبة في شكل مستويات أو في شكل فئات رقمية مثل تحديد فئات مستوى لدخل الأسر في مدينة معينة وتحديد مستويات لتقدير الطلبة في الامتحان.

2- البيانات الكمية Quantitative Data

وهي بيانات يعبر عنها بأرقام عددية تمثل القيمة الفعلية لظاهرة ما ويمكن إجراء عمليات حسابية عليها كدرجة الحرارة المسجلة لمدينة في ساعة معينة خلال شهر أو دخل مجموعة من الأسر في مجتمع ما.

المتغيرات والثوابت Variables and Constants

نتعامل في حياتنا اليومية وفي دراساتنا مع العديد من الظواهر وقد نجد أن ظاهرة ما تتسم بالتغير في القيم التي تأخذها وأخرى تتسم بالثبات، بمعنى أن لها قيمة معينة لا تتغير. فالمتغير هو تلك السمة أو الصفة التي تتغير قيمتها من عنصر إلى آخر في المجتمع مثل طول طالب في فصل، ضغط الدم عند مجموعة من المرضى، مستوى السكر في الدم عند مريض بداء السكر، عدد الأطفال لدي عائلة في مجتمع معين.

أما الثوابت فهي السمات والخواص التي لا تتغير، مثلاً كأن تصف ماهية المواد في ظروف معينة مثل الكثافة النوعية لعنصر ما في ظرف محدد، قوة جاذبية الأرض، عدد أيام الأسبوع ودرجة غليان الماء أو تجمده تحت ظروف معينة.

وهناك ثوابت من نوع آخر تقع ضمن اهتمامات الإحصائي أو الباحث في الدراسات الاجتماعية أو الاقتصادية أو الطبية أو الصناعية أو الزراعية وغيرها وهي تلك الثوابت التي تصف المجتمعات مثل معدل العمر الافتراضي للمصابيح الكهربائية التي ينتجها مصنع معين علي مدى عام واحد ويعتبر هذا المعدل ثابت لكنه غير معلوم، ولو أخذنا عدة عينات من مصابيح ذلك المصنع وأردنا معرفة متوسط عمر الحياة لها لوجدنا أن هذا المتوسط متغير من عينة لأخرى.

تصنيف المتغيرات

يمكن تصنيف المتغيرات المستخدمة من قبل الباحث في الاختبارات التجريبية والدراسات التي يلزمها تحليل إحصائي إلي متغيرات منفصلة ومتغيرات متصلة ويمكن تصنيفها حسب تأثيرها إلى ثلاثة أنواع وهي متغيرات مستقلة ، متغيرات تابعة ومتغيرات وسيطة .

المتغيرات المتصلة والمنفصلة

يمكن تصنيف المتغيرات من حيث القيم التي يأخذها المتغير إلى ما يلي:

1- المتغير المنفصل Discrete variable

هو ذلك المتغير الذي يأخذ قيماً عبارة عن أعداد صحيحة، مثل عدد الأفراد لدى عائلة معينة، عدد الطلبة في شعب مساق معين، عدد الصور التي تظهر عند القاء قطعة نقد عدة مرات، عدد المكالمات الهاتفية التي تصل لمقسم الهاتف في الدقيقة الواحدة وعدد حوادث الطرق في طريق معين في الشهر وهكذا.

2- المتغير المتصل Continuous variable

هذا المتغير لا يأخذ أعداداً صحيحة فقط أو أعداداً معينة بذاتها وإنما يأخذ أي عدد حقيقي من مجموعة الأعداد الحقيقية أو أي مجموعة جزئية منها بمعنى: إذا كان مجال المتغير فترة أو عدة فترات من مجموعة الأعداد الحقيقية سمي المتغير متصلاً. ومن الأمثلة على المتغير المتصل: درجة الحرارة، الوزن، الطول، الزمن، المساحة والعمر وغيرها.

المتغيرات المستقلة والتابعة والوسيط

في دراستنا لبعض الظواهر الطبيعية أو الاجتماعية أو التربوية نجد أن هناك علاقات بين المتغيرات التي نقوم بدراستها ودراسة تأثيرها، فقد نجد أن متغيراً ما يتغير بصورة مستقلة ويؤثر في غيره من المتغيرات ونجد أن هناك متغيراً يتأثر بغيره من المتغيرات ولا تأثير له على غيره من المتغيرات وهناك متغيرات لا تحدث تأثيراً إلا إذا تأثرت بمتغير آخر، من هذا المنطلق يمكن تصنيف المتغيرات من حيث تأثيرها وتأثرها كما يلي:

1- المتغيرات المستقلة Independent Variables

المتغير المستقل هو المتغير الذي يؤثر في متغيرات أخرى (المتغيرات التابعة) ولا يتأثر بها، ويتحكم فيه الباحث. ففي دراسة العلاقة بين الدخل والاستهلاك نجد أن الدخل يؤثر في الاستهلاك وبالتالي يكون الدخل هو المتغير والمستقل، وكذلك عدد ساعات الدراسة تؤثر في درجة الطالب في الامتحان.

2- المتغيرات التابعة Dependent Variables

وهي المتغيرات التي تتأثر بغيرها من المتغيرات (المتغيرات المستقلة) حيث يرغب الباحث في دراسة وجود أثر لمتغير مستقل معين أو أكثر في متغيرات أخرى، فمثلاً عند دراسة تأثير عدد ساعات تعرض النبات لضوء الشمس في طول النبات يكون المتغير التابع هو طول النبات ويكون المتغير المستقل هو عدد ساعات تعرض النبات لضوء الشمس ولاحظ أن الباحث يمكنه التحكم في عدد ساعات تعرض النبات لضوء الشمس.

3- المتغيرات الوسيطة Mediator Variables

المتغير الوسيط هو الذي قد يكون له دور في التأثير على المتغير التابع، ولولا وجوده، لما استطاع المتغير المستقل التأثير في المتغير التابع، فإذا كان الباحث يدرس علاقة الطلاق بانحراف الصغار، يكون المتغير المستقل هنا هو الطلاق، ويكون المتغير التابع هنا هو انحراف الصغار، وقد يكون المتغير الوسيط هنا هو الخلافات الزوجية، أو غياب الأب أو غير ذلك.

في كثير من الأبحاث التي تجرى في المجالات المختلفة كالمجال التربوي يتعامل الباحثون مع متغيرات تأخذ قيماً محددة لا تتغير، هذا النوع من المتغيرات الهامة في علم الإحصاء والتي تلعب دوراً هاماً في عملية التحليل الإحصائي هو المتغير الفئوي حيث يلعب دوراً كبيراً في تصميم الاستبانات والتي تعتبر من الطرق الهامة لجمع البيانات.

المتغيرات الفئوية Categorical variables

هي متغيرات تأخذ قيماً من مجموعة محددة وثابتة من القيم أو من فئات متنافية لا تتقاطع مع بعضها البعض، فتعتبر فصيلة الدم متغير فئوي حيث أن فصيلة دم أي إنسان واحدة من المجموعة $\{O, AB, B, A\}$ ، ويمكن التعبير عن هذه القيم بالأرقام كأن نفرض أن $O=4$ ، $A=1$ ، $B=2$ ، $AB=3$ فنقول في هذه الحالة أن فصيلة الدم هي متغير فئوي يأخذ قيمه من المجموعة $\{4,3,2,1\}$.

من الممكن أن يحدد المعلم تقديرات الطلبة في امتحان معين لتأخذ قيمها من المجموعة {ممتاز، جيد جداً، جيد، مقبول، ضعيف} وبالتالي تعتبر تقديرات الطلبة في هذه الحالة متغيراً فئوياً، ويمكن التعبير عن هذه القيم بالأرقام بأن نفرض مثلاً أن "ضعيف"=1، "مقبول"=2، "جيد"=3، "جيد جداً"=4، "ممتاز"=5 ونقول في هذه الحالة أن المتغير يأخذ القيم من المجموعة $\{5,4,3,2,1\}$.

تدرج القياس Measurement scales

في حياتنا اليومية كثيراً ما نقوم بإعطاء قيم عددية للملاحظات أو الظواهر أو الحالات الاجتماعية فمثلاً نقول كانت علامة سعيد 38 في أحد الامتحانات أو كانت درجة حرارة المريض 38 درجة مئوية. وهنا نتساءل: هل هناك أي فروق في المقاييس التي نستعملها؟ هل للتدرج على مقياس الحرارة نفس معني التدرج على سلم درجات الطلاب؟

للإجابة على هذه التساؤلات سنقوم بدراسة وتوضيح مفهوم تدرج القياس، وسنقسم مستويات القياس إلى أربع مستويات حيث يختلف تدرج قياس عن تدرج قياس آخر من حيث المعايير التي تستخدم لتصنيف الظواهر حسب تدرج القياس المناسب، ويمكن صياغة هذه المعايير كما يلي:

- قابلية المقارنة لملاحظات للظاهرة.
- إمكانية ترتيب مشاهدات الظاهرة.
- إمكانية حساب فروق لملاحظات الظاهرة.

- إمكانية وجود الصفر كملاحظة للظاهرة ومعني الصفر في حالة وجوده هل هو صفر مطلق أم صفر اصطلاحي.

وسنذكر المستويات من أدنى مستوى إلى أعلى مستوى.

1- التدرج الاسمي Nominal scale

يستعمل التدرج الاسمي كمقياس لتحديد هوية الأفراد أو العناصر وبالتالي فهو يصنف عناصر الظاهرة التي تمثلها متغيرات نوعية لا كمية، وتندرج تحته الظواهر التي لا يمكن إجراء مقارنات بين مشاهداتها أو حتى المفاضلة بينها ولا يمكن ترتيب مشاهداتها أو إجراء عمليات حسابية عليها، مثل لون العيون، رقم الطالب الجامعي، رقم البطاقة الشخصية وهكذا. وكثيراً ما تستعمل الأعداد لتحديد هوية المفردات وفي هذه الحالة لا يكون للعدد ذلك المدلول الكمي الذي يفهم منه عادةً. فمثلاً يمكن استعمال العددين 1، 2 ليدلّا على التصنيف حسب الجنس فنستخدم العدد 1 ليدل على الذكر والعدد 2 ليدل على الأنثى ولا يحمل هذان العددان أي صفة رقمية وتمثل هذه الظواهر عادةً متغيرات فئوية.

2- التدرج الترتيبي Ordinal scale

يقع هذا التدرج في مستوى أعلى من مستوى التدرج الاسمي. إن هذا التدرج يسمح بالمفاضلة والمقارنة بين قيم الظاهرة وبالتالي ترتيب العناصر تصاعدياً أو تنازلياً، مثل الرتب العسكرية والتقدير الذي يحصل عليه خريج الجامعة. ومشاهدات الظواهر التي تقع وتصنف ضمن هذا التدرج لا يجرى عليها عمليات حسابية ويمثل هذه الظاهرة غالباً متغيرات فئوية.

3- تدرج الفترة Interval scale

يعطي هذا المستوى معنى للفروق بين المشاهدات وبذلك يقع في مستوى أعلى من مستوى التدرج الترتيبي وتكون وحدات القياس في هذا التدرج متساوية وتقاس قيم الظاهرة بمدى بعدها عن الصفر وهذا يعني أنه من الممكن أن تأخذ المشاهدات فيه القيمة صفر ولكن الصفر في هذه الحالة رمز اصطلاحي وليس هو الصفر المطلق الذي يعني العدم، مثل

القياس على ميزان الحرارة المئوي فهو اصطلاحى أى أننا اصطلاحنا أن نعين الدرجة الصفر مئوية لتدل على درجة تجمد الماء النقي تحت الضغط الجوى المعتاد، فذلك الصفر على ميزان الحرارة المئوي لا يعنى العدم بدليل شعورنا بدرجة حرارة الماء المتجمد وبالتالي فهو ليس الصفر المطلق. وفي هذا التدرج نستطيع إيجاد الفرق بين القيم، ويكون لهذه الفروق معنى واضح، ومن الأمثلة كذلك على المتغيرات التي تقع ضمن تدرج الفترة مقياس الضغط الجوى (البارومتر) وقياس معامل الذكاء والعلامات التي يحصل عليها الطلبة في الامتحانات.

4- التدرج النسبي Ratio scale

يقع هذا التدرج في أعلى مستوى من مستويات التدرج، ويتميز بكل خواص تدرج الفترة وأهم فرق بينه وبين تدرج الفترة أن التدرج النسبي يعطي معنى للصفر فيجعله الصفر المطلق، أي أن الصفر على هذا التدرج يعنى العدم، وبالتالي فهو ليس صفراً اصطلاحياً ومن ثم فإننا نستطيع إعطاء معنى للنسب بين القيم المقاسة بالتدرج النسبي، ومن أهم الأمثلة عليه قياسات الطول، المساحة، الحجم، الوزن والقوة.

تمارين:

1- صنف المتغيرات التالية من حيث كونها متصلة أو منفصلة:

- عدد السيارات التي تدخل حرم الكلية في اليوم.
- درجات الحرارة المسجلة الساعة الثانية عشر ظهرا في شهر آذار لعام 2013 في محافظة خان يونس.

- عدد المواليد الذكور الأحياء في المغازي شهرياً في سنة 2013.
- أوزان المرضى الذين يترددون على عيادة السكر في بيت حانون في أسبوع معين.
- أعمار الطلبة المشتركين في امتحان الشامل للعام 2014/2013.
- الوقت الذي يستغرقه الطالب في حل امتحان الإحصاء.

2- صنف ما يلي إلى متغيرات وثابت:

- درجة تجمد الماء النقي تحت الضغط الجوي العادي عند سطح البحر.
- دخل الفرد السنوي في قطاع غزة.
- العدد الذري للهيدروجين.
- عدد أفراد أسرة معينة.
- عدد أفراد الأسر في مدينة رفح.
- دخل العامل اليومي.

3- صنف المتغيرات التالية حسب تدرج القياس:

- علامات الطلبة في أحد الامتحانات.
- درجة الحرارة المئوية.
- طول قطعة مستقيمة.
- لون العين.
- رقم الطالب الجامعي.
- كمية الألبان التي تنتجها بقرة معينة في اليوم.

- وزن الأبقار في مزرعة لتربية الأبقار .
- الحالة الاجتماعية للموظف.
- أنواع التمور .
- المساحة المزروعة بالنخيل بالدونم.
- جنسية الطالب في إحدى الجامعات.
- المستوى التعليمي لربات البيوت.
- إنتاجية الدونم من القمح.

4- أراد طبيب دراسة تأثير العمر علي ضغط الدم لدى الرجال البالغين في مدينة النصيرات فدرس 30 حالة سجل فيها عمر الرجل وضغط الدم لديه. ما هو المتغير المستقل في هذه

الدراسة؟ ما هو المتغير التابع؟ ما هو مجتمع الدراسة وعينة الدراسة؟

5- هل يمكن استخدام أسلوب المسح الشامل فيما يلي:

- فحص انتاج مصنع لإطارات السيارات.
- فحص دم مريض لمعرفة نسبة الهيموجلوبين فيه.
- معرفة عدد المواليد في الدولة لتوفير التطعيمات اللازمة.
- فحص إنتاج مصنع من المشروبات الغازية.
- معرفة عدد المصابين بمرض معين في مدينة معينة.
- معرفة تعداد السكان للدول.

6- قامت إحدى المجلات باستطلاع الرأي العام للسكان في قطاع غزة حول ظاهرة استعمال حزام الأمان في السيارات، حيث قابل مندوبها 2000 شخص، فوجد أن 32% منهم يفضلون استعمال الحزام، بينما البقية لا تفضل ذلك.

- ما هو مجتمع هذه الدراسة وعينة الدراسة؟
- كم عدد الذين لا يفضلون استعمال حزام الأمان؟
- ما هي الطريقة المناسبة التي تعتقد أن المجلة استعملتها لإجراء الاستطلاع؟ لماذا؟

7- تبدأ عملية التسجيل في الجامعة بتسجيل طلبة السنة الرابعة أولاً ثم الثالثة ثم الثانية وأخيراً طلبة السنة الأولى، وتوضع بطاقات التسجيل لطلبة كل سنة على حدة. بلغ عد الطلبة المسجلين في السنة الرابعة 2000، الثالثة 2300، الثانية 2700، والأولى 2500 وأراد باحث دراسة الأخطاء في عملية التسجيل فقرر أخذ عينة حجمها 950 طالب بطريقة المعاينة الطبقية العشوائية. أوجد الحجم المأخوذ من كل طبقة.

8- وضح كيف يمكن سحب عينة بطريقة المعاينة العشوائية المنتظمة مع ذكر مثال على ذلك؟

9- أذكر طرق جمع البيانات وتحدث عن مزايا وعيوب كل طريقة.

10- عرف المتغيرات المنفصلة والمتصلة مع ذكر أمثلة لكل نوع.

11- أذكر أمثلة لمتغيرات مستقلة وتابعة ووسيطه في دراسات مقترحة.

12- أذكر أمثلة لمتغيرات فنوية.

13- اختر الإجابة الصحيحة فيما يلي:

• عند دراسة العلاقة بين درجة الطالب في الامتحان وعدد ساعات دراسته اليومية يكون المتغير المستقل هو:

أ) الطالب ب) درجة الطالب ج) عدد ساعات الدراسة د) الامتحان

• مساحة قطعة أرض يمتلكها أحد المواطنين تتبع التدرج:

أ) الاسمي ب) الترتيبي ج) الفترة د) النسبي

• يعتبر عدد الأبناء في أسرة:

أ) متغير منفصل ب) متغير متصل ج) ثابت د) متغير فنوي

• عند دراسة العلاقة بين دخل الفرد وإنفاقه في الشهر الواحد يكون المتغير التابع هو:

أ) الدخل ب) الفرد ج) الإنفاق د) الادخار

• يعتبر معامل الذكاء تابعاً للتدرج:

أ) الاسمي ب) الترتيبي ج) الفترة د) النسبي

الإحصاء: هو رياضيات جمع البيانات وتنظيمها وتحليلها وتفسيرها.
المجتمع: هو مجموعة العناصر أو الأفراد الذين ستجرى عليهم الدراسة الإحصائية وستعمم عليها نتائج الدراسة.
العينة: هي مجموعة جزئية من المجتمع ويتم اختيارها بإحدى طرق المعاينة.

الإحصاء الوصفي: هو مجموعة من الأساليب الإحصائية التي تستخدم لجمع البيانات وتنظيمها وتصنيفها وعرضها عن طريق الجداول أو الرسوم البيانية وغيرها وحساب مقاييس إحصائية.

الإحصاء الاستدلالي: هو مجموعة من الأساليب الإحصائية التي تستخدم لتحليل البيانات التي يتم الحصول عليها من عينة مسحوبة من مجتمع الدراسة وتفسيرها.

أساليب جمع البيانات: أسلوب المسح الشامل - أسلوب العينة.
طرق جمع البيانات: المقابلة الشخصية - المكالمات الهاتفية - الاستبانة - الملاحظة المباشرة المصادر الرسمية.

العينات الاحتمالية: العشوائية البسيطة - العشوائية المنتظمة - العشوائية الطبقية - العشوائية العنقودية (العشوائية متعددة المراحل).
العينات غير الاحتمالية: العرضية - الحصصية.

تصنيف المتغيرات من حيث القيم التي يأخذها: متغيرات منفصلة - متغيرات متصلة.
أنواع المتغيرات من حيث التأثير والتأثر: متغيرات مستقلة - متغيرات تابعة - متغيرات وسيطة - متغيرات فئوية.

مستويات تدرج القياس: تدرج اسمي - تدرج ترتيبي - تدرج فترة - تدرج نسبي.

تبويب البيانات وعرضها بيانياً Data Tabulation & Charts Graphs 2

عرفنا في الفصل الأول الإحصاء وقلنا إنه رياضيات جمع البيانات وتنظيمها وتحليلها وتفسيرها، وتعرضنا فيه لأساليب وطرق جمع البيانات، وفي هذا الفصل سنعرض كيف ننظم البيانات والذي نقصد به كيفية تبويب البيانات أي وضعها في جداول وتوزيعات تكرارية وكيفية عرض هذه الجداول والتوزيعات التكرارية برسوم بيانية مناسبة.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً على:

- تبويب البيانات النوعية في جداول تكرارية.
- تعرض الجداول التكرارية بطريق الأعمدة البيانية البسيطة والمزدوجة، الخطوط المنكسرة، الدوائر البيانية.
- تبويب البيانات الكمية في توزيعات تكرارية.
- تعرض التوزيعات التكرارية بطريقة المدرج التكراري.
- ترسم المنحنى التراكمي الصاعد والهابط.
- تحدد الوسيط بيانياً.

إن أول خطوة في عملية التحليل الإحصائي هي عملية جمع البيانات وقد تحدثنا عنها في الفصل الأول. هذه البيانات التي يتم جمعها تسمى البيانات الخام Raw Data والتي من الصعب التعامل معها مباشرة في عملية التحليل الإحصائي وذلك لصعوبة الحصول منها على المعلومات وعدم المقدرة على وصف المجتمع والتعرف على صفاته وخصائصه. من هنا نلجأ لتنظيم البيانات الخام ليسهل علينا الحصول على المعلومات وتحليل البيانات بشكل يمكننا من تحقيق أهدافنا.

تبويب البيانات أي وضعها في جداول تكرارية أو توزيعات تكرارية تعتبر الخطوة الأولى في عملية تنظيم البيانات، حيث يتم بعد ذلك عرضها بيانياً بطريقة مناسبة تسهل الحصول على مزيد من المعلومات وإجراء المقارنات ومعرفة بعض صفات وخصائص المجتمع. وسنبدأ الفصل بتبويب البيانات النوعية وعرضها بيانياً ثم تليها البيانات الكمية.

أولاً: تبويب البيانات النوعية وعرضها بيانياً

من المخططات الهامة التي تعرض الجداول التكرارية للبيانات النوعية الأعمدة البيانية، الخطوط المنكسرة والدوائر البيانية.

إذا كان لدينا مجموعة كبيرة من البيانات فإننا سوف نواجه مشقة التحديق في العشرات وربما المئات منها عند الرغبة في فهم هذه البيانات أو التعرف على خصائصها. ولكن يمكن استنباط الصورة العامة لهذه البيانات بتلخيصها عن طريق تجميع تلك البيانات في جداول. تسمى هذه العملية تبويب البيانات، والجداول التي نحصل عليها سنطلق عليها اسم الجداول التكرارية، وقد يمثل الجدول متغيراً فئوياً واحداً أو متغيرين فئويين.

وفيما يلي تلخيص لخطوات تبويب البيانات النوعية:

1- نكون جدولاً من ثلاثة أعمدة نطلق عليه اسم جدول التفريغ، حيث نضع في العمود الأول القيم المختلفة لفئات المتغير الفئوي والعمود الثاني نفرغ فيه البيانات والعمود الثالث هو عمود التكرارات.

2- نبدأ بعد عدد مرات تكرار قيم أول فئة وفي العمود الثاني نضع شرطة مائلة (/) عن كل قيمة في البيانات تقع في هذه الفئة وعند وجود قيمة خامسة نضع لها شرطة مائلة عكسية (\\) لتشكّل حزمة تعبر عن خمسة تكرارات وهكذا لجميع الفئات.

3- نعد التكرارات في عمود التفريغ ونكتب قيمة التكرار في العمود الثالث، حيث يكون مجموع التكرارات في العمود الثالث يمثل عدد البيانات الخام.

4- نأخذ العمود الأول والثالث معاً لنحصل على الجدول التكراري المطلوب.

مثال 1.2: تقدم 30 طالباً لامتحان الإحصاء وكانت تقديراتهم كما يلي:

جيد	ممتاز	ممتاز	جيد جداً	ضعيف	جيد	ممتاز	جيد جداً	مقبول	جيد
مقبول	ضعيف	جيد جداً	ممتاز	جيد	ممتاز	جيد	جيد	مقبول	جيد
ممتاز	ضعيف	جيد جداً	ممتاز	جيد	مقبول	جيد	جيد	جيد	مقبول

ضع البيانات في جدول تكراري مناسب.

الحل/

لاحظ هنا أن البيانات نوعية وفئات المتغير الفئوي المختلفة هي ضعيف، مقبول، جيد، جيد جداً وممتاز، فنكون جدول التفريغ التالي ونقوم بتفريغ جميع المشاهدات فيه وذلك بعدد مرات تكرار كل قيمة وذلك كما هو موضح في جدول التفريغ التالي:

Categories الفئات	التفريغ	الترار Frequency
ضعيف	///	3
مقبول	////	5
جيد	/ ////	11
جيد جدا	////	4
ممتاز	// ////	7
مجموع التكرارات		30

ويمكن الاستغناء عن العمود الأوسط فيكون الجدول التكراري المطلوب كما يلي:

الفئة	ضعيف	مقبول	جيد	جيد جدا	ممتاز
التكرار	3	5	11	4	7

عرض الجداول التكرارية بيانياً

نقدم فيما يلي بعض طرق عرض الجداول التكرارية للبيانات النوعية بيانياً ونوضح باختصار كيفية الحصول على بعض المعلومات منها.

1- طريقة الأعمدة البيانية البسيطة Simple Bar Charts

تعتبر الأعمدة البيانية من الرسوم البيانية الهامة والواسعة الاستخدام وذلك لسهولة رسمها وقراءتها والقدرة على استخدامها بسهولة لإجراء المقارنات بين الفئات المختلفة.

ونوضح فيما يلي خطوات رسم مخطط الأعمدة البيانية البسيطة:

1- نرسم محورين متعامدين ونجعل المحور الأفقي لفئات الجدول التكراري والمحور الرأسي للتكرار.

2- نقسم كل محور لأجزاء متساوية حيث نأخذ للفئة الأولى جزء من المحور الأفقي مبتدئين من اليسار نحو اليمين ونترك الجزء الثاني ونأخذ الثالث للفئة الثانية ونترك الرابع وهكذا، ثم نرقم محور التكرارات بطريقة منتظمة مناسبة تبدأ بالصفر وذلك حسب قيمة أكبر تكرار.

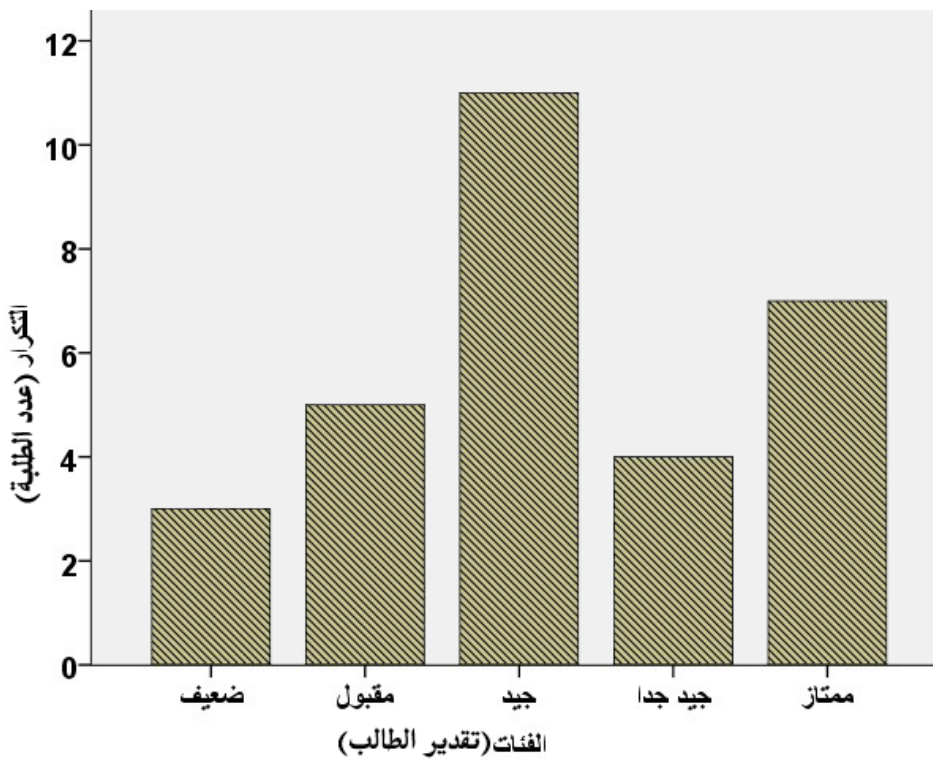
3- نرسم على كل فئة مستطيلاً رأسياً يتناسب طوله مع تكرار هذه الفئة ومن الممكن تظليله أو تلوينه بلون معين أو تهشيره برسم معين لتمييزه عن غيره من المستطيلات. ولكن في حالة الجداول التكرارية التي تمثل متغيراً فئوياً واحداً يفضل استخدام تظليل أو لون أو تهشير واحد لجميع الأعمدة.

مثال 2.2: الجدول التكراري التالي يبين تقديرات 30 طالباً في امتحان. اعرض الجدول بطريقة الأعمدة البيانية البسيطة:

الفئة	ضعيف	مقبول	جيد	جيد جداً	ممتاز
التكرار	3	5	11	4	7

الحل/

نرسم محورين متعامدين بحيث يمثل المحور الأفقي الفئات والتي تمثل تقديرات الطلبة، والمحور الرأسي للتكرار والذي يمثل عدد الطلبة، ثم نرسم على كل فئة مستطيلاً ارتفاعه يساوي تكرار الفئة كما هو موضح في الرسم.



ويمكن الاستنباط من الرسم أن فئة الطلبة الحاصلين على تقدير جيد هي أكبر فئة، وفئة الطلبة الحاصلين على تقدير ضعيف هي أقل فئة، كما يمكننا ملاحظة أن فئة الطلبة الحاصلين على تقدير مقبول أكبر من فئة الطلبة الحاصلين على تقدير جيد جداً وهكذا يمكننا إجراء مقارنات بين الفئات المختلفة بسهولة.

2- طريقة الخط المنكسر البسيط Line Charts

عند الرغبة في عرض الجدول التكراري بطريقة الخط المنكسر نتبع الخطوات التالية:

- 1- يتم تقسيم محور الفئات الأفقي إلى نقاط بينها مسافات متساوية.
- 2- من اليسار اليمين يتم تخصيص نقطة لكل فئة تصاعدياً.
- 3- نرسم نقطة في المستوى فوق كل فئة مقابل تكرارها.
- 4- نصل بين النقاط بخط منكسر باستخدام المسطرة.

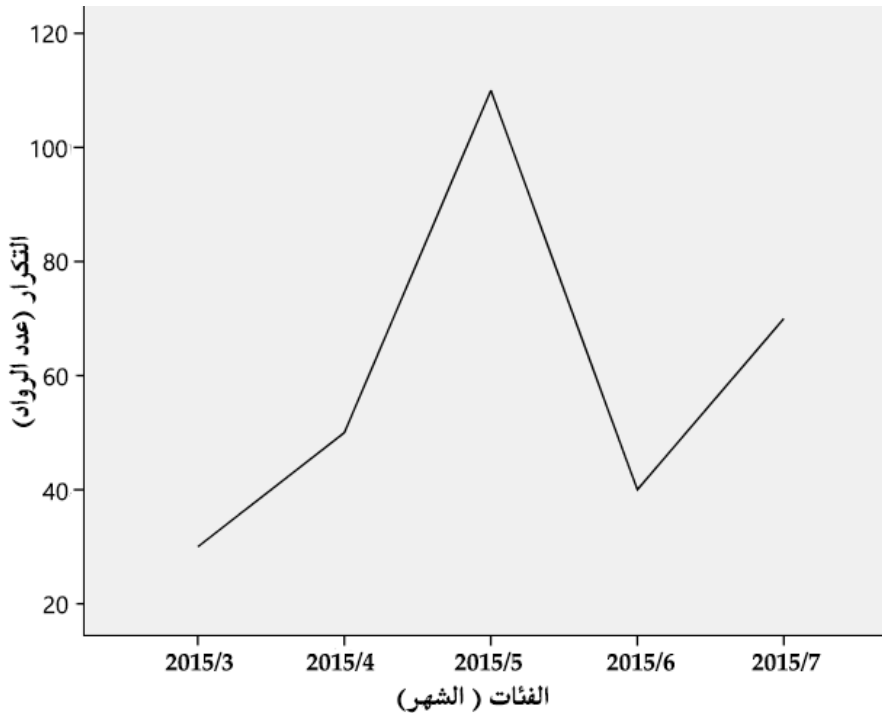
وتستخدم هذه الطريقة لمعرفة الاتجاه العام للظاهرة قيد البحث لمتغير فئوي تمثل فئاته فترات زمنية متتابعة.

مثال 3.2: الجدول التكراري التالي يبين أعداد المرضى الذين دخلوا في قسم الباطنة في مشفى المدينة في خمسة شهور لسنة 2015. والمطلوب عرض الجدول بطريقة الخط المنكسر البسيط.

الشهر	3	4	5	6	7
عدد المرضى	30	50	110	40	70

الحل/

نرسم محورين متعامدين بحيث يمثل المحور الأفقي الفئات حيث لكل فئة نقطة واحدة على المحور، وتمثل الأشهر من 3 إلى 7 للعام 2015، ويمثل المحور الرأسي التكرار ويمثل عدد المرضى، ثم نرسم نقطة في المستوى رأسياً فوق كل فئة تبعد عن المحور الأفقي مسافة تساوي تكرار الفئة كما هو موضح في الرسم التالي:



ويوضح الرسم أن عدد المرضى يزداد في الفترة من شهر 3 إلى شهر 5 ثم ينخفض في شهر 6 ليبدأ في الارتفاع مرة أخرى في شهر 7.

3- طريقة الدوائر البيانية Pi Charts

من أهم استعمالات هذه الطريقة هو تقسيم مجتمع الظاهرة ككل إلى أجزائه، فيمثل المجموع الكلي بدائرة كاملة ويمثل كل جزء بقطاع دائري تكون قياس زاويته هي:

$$\text{زاوية القطاع الدائري} = \frac{\text{حجم القطاع}}{\text{الحجم الكلي}} \times 360$$



ولرسم الدائرة البيانية نتبع الخطوات التالية:

- 1- حساب زاوية القطاع لكل قطاع من قطاعات المجتمع ويجب أن نتأكد أن مجموع الزوايا مجتمعة 360 درجة.
- 2- نرسم دائرة بنصف قطر مناسب.
- 3- نرسم نصف قطر وليكن أفقياً لنبدأ منه رسم القطاعات المختلفة.
- 4- من نصف القطر هذا نرسم أول قطاع برسم زاويته بالمنقلة.
- 5- من نصف القطر الذي يمثل نهاية القطاع الأول نرسم القطاع الثاني وهكذا.
- 6- نلون كل قطاع بلون خاص أو نعمل تهشير خاص لكل قطاع ونبين ذلك في مفتاح الرسم.

مثال 4.2: الجدول التالي يوضح أعداد الطلبة الملتحقين في دورة اللغة الانجليزية موزعين حسب القسم:

القسم	العلوم الإدارية	العلوم الطبية	علوم الحاسوب	الهندسية
عدد الطلبة	30	20	15	7

أعرض الجدول التكراري بطريقة الدوائر البيانية.

الحل/ نمثل هنا أن كل قسم بقطاع دائري، وسنحسب الآن زاوية كل قطاع باستخدام العلاقة السابقة وذلك كما يلي:

$$\text{الحجم الكلي} = 7 + 15 + 20 + 30 = 72$$

$$\text{قياس زاوية قطاع العلوم الإدارية} = \frac{30}{72} \times 360 = 150 \text{ درجة}$$

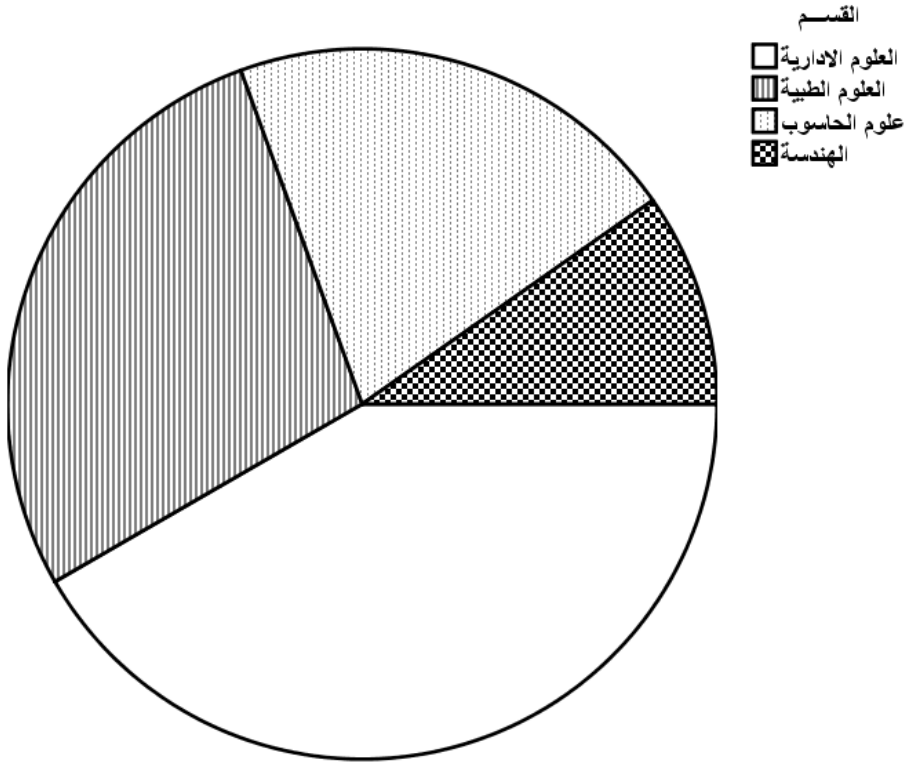
$$\text{قياس زاوية قطاع العلوم الطبية} = \frac{20}{72} \times 360 = 100 \text{ درجة}$$

$$\text{قياس زاوية قطاع علوم الحاسوب} = \frac{15}{72} \times 360 = 75 \text{ درجة}$$

$$\text{قياس زاوية قطاع الهندسية} = \frac{7}{72} \times 360 = 35 \text{ درجة}$$

نرسم دائرة بنصف قطر مناسب ونرسم نصف قطر أفقي على سبيل المثال كخط بداية لرسم زوايا القطاعات، ثم نرسم زاوية أحد القطاعات منطلقين منه وليكن قطاع الهندسية، ومن خط النهاية نبدأ بقياس زاوية قطاع آخر وليكن علوم الحاسوب، ومن خط نهاية هذا القطاع نبدأ بقياس زاوية قطاع آخر وليكن العلوم الطبية وبعد قياس زاويته يكون آخر قطاع قد تم تحديده.

ويتم تظليل كل قطاع بطريقة معينة لتمييزه ولتسهيل الحصول على المعلومات من الرسم وإجراء المقارنات. ومن الممكن استخدام الألوان في التظليل أو التهشير. والرسم التالي يوضح العرض المطلوب:



يمكن من الرسم الحصول على معلومات حول كل قسم وإجراء مقارنات بين الأقسام من حيث أكبر الأقسام وأصغرها، حيث يوضح أن قسم العلوم الإدارية هو أكبر الأقسام من حيث عدد الطلبة بينما قسم العلوم الهندسية أصغرها.

في بعض الدراسات والأبحاث قد نحصل على جداول تكرارية تمثل متغيرين فئويين وفي هذه الحالة يتم تمثيلها بيانياً بطرق خاصة نذكر منها طريقة الأعمدة البيانية المزدوجة وطريقة الخط المنكسر المتعدد.

1- طريقة الأعمدة البيانية المزدوجة Clustered Bar Charts

تستخدم هذه الطريقة للمقارنة بين ظاهرتين أو أكثر حسب الصفات التي تمثلها. ونوضح فيما يلي خطوات رسم مخطط الأعمدة البيانية المزدوجة:

1- نرسم محورين متعامدين ونجعل المحور الأفقي لفئات أحد متغيري الجدول التكراري والمحور الرأسي للتكرار.

2- نقسم كل محور لأجزاء متساوية ونبدأ بالمحور الأفقي حيث نأخذ الفئة الأولى لأحد المتغيرين جزءاً من هذه الأجزاء مبتدئين من اليسار نحو اليمين ونقسمها لأجزاء تساوي عدد فئات المتغير الفئوي الثاني ونجعل لكل فئة جزءاً منها ونترك الجزء الثاني ونأخذ الثالث للفئة الثانية للمتغير الأول ونقسمها كما فعلنا في الفئة السابقة وهكذا.

3- نرسم مستطيلاً فوق كل جزء من الفئة الأولى ارتفاعه يتناسب مع تكرار الفئة المقابل بالنسبة للمتغير الثاني فتكون الفئة الأولى للمتغير الفئوي الأول قد أقيم عليها مستطيلات بعدد فئات المتغير الثاني وتكون جميعها متلاصقة.

4- نعطي كل فئة لوناً أو تهشيراً معيناً ويكرر نفسه في كافة المستطيلات المقامة على فئات المتغير الثاني والمرسومة على كل فئة للمتغير الأول.

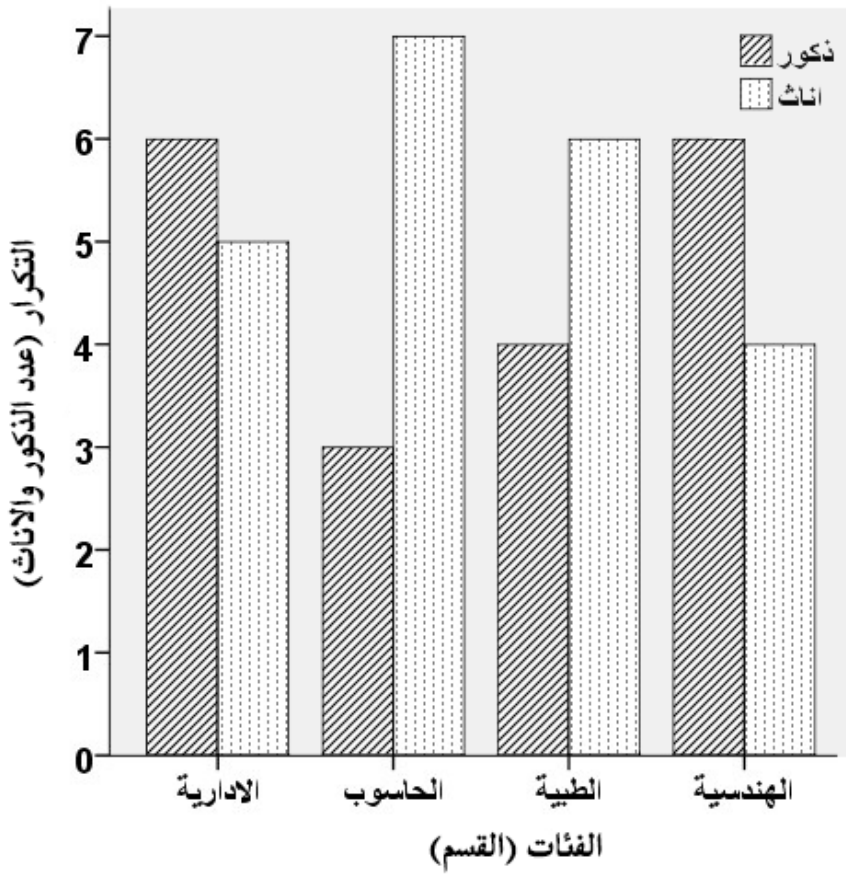
5- نضيف مفتاح للرسم ليسهل قراءتها وإجراء المقارنات المطلوبة.

مثال 5.2: الجدول التكراري التالي يوضح أعداد الطلبة الملتحقين في دورة اللغة الانجليزية موزعين حسب القسم والجنس:

القسم		الإدارية	الحاسوب	الطبية	الهندسية
عدد الطلبة	ذكور	6	3	4	6
	إناث	5	7	6	4

مثل الجدول بطريقة الاعمدة البيانية المزدوجة.

نرسم محورين متعامدين بحيث يمثل المحور الأفقي الفئات والتي تعبر عن القسم، والمحور الرأسي التكرار والذي يمثل عدد الطلبة، ثم نقسم كل فئة لجزئين يمثلان عدد فئات المتغير الثاني (الجنس) وهما فئتان، فئة الذكور وفئة الإناث، ثم نرسم على كل فئة مستطيلين ارتفاع أحدهما يساوي تكرار الذكور والآخر يمثل تكرار الإناث ونفَرِّق بينهما بتهشير أعمدة الذكور بطريقة معينة وأعمدة الإناث بتهشير مختلف، أو استخدام الألوان بدل التهشير، والرسم التالي يوضح ذلك:



2- طريقة الخط المنكسر المتعدد Multiple Line Charts

في حالة وجود جدول تكراري يعبر عن أكثر من متغير فئوي أحدها تعبر فئاته عن فترات زمنية، يمكن استخدام طريقة الخط المنكسر المتعدد حيث يوضح الرسم الاتجاه العام لفئات المتغير الثاني خلال الفترات الزمنية. ولرسم الجدول التكراري بهذه الطريقة نتبع الخطوات التالية:

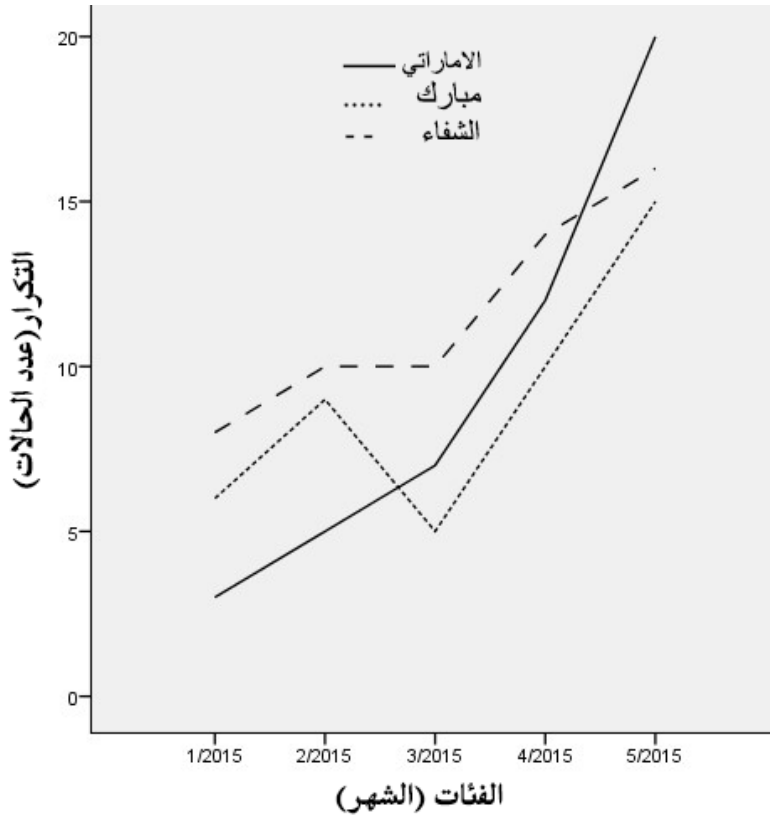
- 1- يتم تقسيم محور الفئات الأفقي إلى نقاط بينها مسافات متساوية.
- 2- من اليسار لليمين يتم تخصيص نقطة لكل فئة زمنية تصاعدياً.
- 3- نرسم نقطة في المستوى رأسياً فوق كل فئة زمنية تقابل تكرار فئة معينة للمتغير الثاني.
- 4- نصل بين النقاط بخط منكسر باستخدام المسطرة ويتم تقطيعه بطريقة معينة لتمييزه عن غيره من الخطوط التي تعبر عن الخطوط المنكسرة لفئات الأخرى.
- 5- نعمل مفتاح للخطوط لسهولة قراءتها وتتبع الاتجاه العام.

مثال 6.2: الجدول التكراري التالي يبين عدد حالات الولادة القيصرية في ثلاث مشافي أول خمسة أشهر للعام 2015:

الشهر المشفى	1	2	3	4	5
الإماراتي	3	5	7	12	20
مبارك	6	9	5	10	14
الشفاء	8	10	10	14	16

مثل الجدول التكراري بطريقة الخط المنكسر المتعدد.

نرسم محورين متعامدين ونضع على المحور الأفقي فئات المتغير الفئوي الزمني وعلى المحور الرأسي التكرار والذي يمثل عدد حالات الولادة، ثم نرسم خطاً منكسراً بتقاطع معين لكل فئة من فئات المتغير الثاني والذي يمثل المشفى، والرسم التالي يوضح ذلك:



ومن الرسم يمكن استنباط أن أعداد الولادة في تزايد مستمر في مشفى الإماراتي، أما في مشفى مبارك فهناك زيادة في أول شهرين ثم نقص في الثالث يليه زيادة في باقي الشهور، أما في مشفى الشفاء فهناك تزايد في الأعداد مع ثبات واضح في شهري 2 و3.

ثانياً: تبويب البيانات الكمية وعرضها بيانياً

عند الحصول على بيانات كمية كبيرة نسبياً خلال عملية جمع البيانات فمن المنطقي تنظيم هذه البيانات حتى يمكن التعامل معها وفهمها ودراسة خصائص المجتمع وصفاته. ويتم ذلك بتبويبها في جداول تسمى توزيعات تكرارية. وتتم عملية تبويب هذه البيانات بعدة خطوات نذكرها فيما يلي:

1- نحدد عدد الفئات، ويراعى أن يكون عدد الفئات بشرط ألا يقل عن 5 ولا يزيد عن 12.

2- نحسب المدى والذي قيمته تساوي أكبر مشاهدة مطروحاً منها أصغر مشاهدة، حيث يُعرف المدى بأنه القيمة التي تعبر عن المسافة التي تنتشر فيها المشاهدات.

3- نجد طول الفئة وذلك بقسمة المدى على عدد الفئات وتقريب الناتج لأقرب عدد صحيح لأعلى دائماً، حيث:

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

4- نكون الفئات حيث تبدأ الفئة الأولى بأصغر مشاهدة أو بقيمة تقل عنها بقليل، وتسمى هذه القيمة بالحد الأدنى للفئة الأولى ونحصل على ما يسمى الحد الأعلى للفئة بإضافة (طول الفئة-1) للحد الأدنى آخذين في الاعتبار ضرورة أن تتضمن الفئة الأولى أصغر مشاهدة والفئة الأخيرة أكبر مشاهدة.

5- نكون جدول التفريغ ثم نحسب تكرار كل فئة كما وضعنا في أول هذا الفصل.

ومن الجدير ذكره هنا أنه يجب أن يكون طول الفئة مناسباً بحيث يحافظ على معالم المجتمع، حيث أن كبر طول الفئة يجعل عدد كبير من المشاهدات تقع ضمنها وقد تكون المشاهدات متباينة فتضيع معالمها، وإذا كان طول الفئة صغيراً يكون عددها كبيراً وبالتالي نكون قد أضعنا الفائدة من عملية التبويب وهي تسهيل دراسة خواص المجتمع.

لذلك إذا وجد الباحث أن عدد الفئات وطول الفئة المحسوبين غير مناسبين، فيمكنه إجراء التعديلات المناسبة للحصول على أفضل عدد للفئات وأفضل طول فئة. ونحن خلال هذا الكتاب سنلتزم بالخطوات المحددة سابقاً، والمثال التالي يوضح عملية التبويب:

مثال 7.2: إليك درجات 50 من طلبة كلية التجارة بالجامعة الإسلامية بغزة في مادة مبادئ الإحصاء في الفصل الأول للعام الجامعي 1995 - 1996:

18	11	16	23	16	28	14	23	24	18
13	19	16	28	14	24	22	22	16	22
9	12	17	25	28	14	16	18	20	15
16	24	18	21	17	15	28	8	14	15
11	8	18	17	30	17	9	17	22	12

والمطلوب تبويب هذه البيانات في توزيع تكراري من 6 فئات.

أولاً نحسب المدى وذلك كما يلي:

$$\text{المدى} = \text{أكبر مشاهدة} - \text{أصغر مشاهدة} = 30 - 8 = 22$$

وبالتالي يكون طول الفئة مساوياً

$$4 \cong \frac{22}{6} = \frac{\text{المدى}}{\text{عدد الفئات}} = \text{طول الفئة}$$

حيث نقرب الناتج لأقرب عدد صحيح لأعلى دائماً.

نبدأ بتكوين الفئات مبتدئين بالفئة الأولى حيث تبدأ كل فئة بما يسمى **الحد الأدنى للفئة (lower cut point)** وتنتهي بما يسمى **الحد الأعلى للفئة (upper cut point)**، ونأخذ الحد الأدنى للفئة الأولى مساوياً أصغر مشاهدة أو عدد قريب منه ونكتب أول فئة كالتالي:

" 8 - 11 " حيث تحوي الأرقام من 8 الى 11 وعددها 4، وتبدأ الفئة الثانية بالقيمة التالية مباشرة وهي 12 ونضيف إليها (طول الفئة - 1) لنحصل على الفئة الثانية وهي " 12 - 15 " وهكذا.

وحتى لا نقع في حرج عند وجود قيمة تقع بين القيمة 11 والقيمة 12 سنكتب الفئة الأولى على الصورة " 8 - " والفئة الثانية على الصورة " 12 - " وبنفس السياق تكون الفئة الثالثة " 16 - " والفئة الرابعة " 20 - " والفئة الخامسة " 24 - " أما الفئة الأخيرة فتكتب على الصورة " 28 - 32 " ويكون الحد الأدنى لكل فئة هو الأعلى للفئة التي تسبقها.

ولاحظ هنا أن أصغر مشاهدة لابد أن تقع في الفئة الأولى وأكبر مشاهدة لابد أن تقع في الفئة الأخيرة.

نكون الآن جدول تفريغ البيانات كالتالي:

الفئات	التفريغ	التكرار
-8		6
-12		10
-16		17
-20		8
-24		4
32-28		5
مجموع التكرارات		50

ونستغني في الصورة النهائية للجدول عن المساحة المخصصة للتفريغ، فيظهر التوزيع التكراري كما يلي:

الفئات	-8	-12	-16	-20	-24	32-28
التكرار	6	10	17	8	4	5

ويمكن كتابة التوزيع التكراري بشكل عمودي وتكون هذه الطريقة مناسبة عند حساب بعض المقاييس والمعاملات الإحصائية كما سنرى في الفصول القادمة.

عرض التوزيعات التكرارية بيانياً

تُعرض التوزيعات التكرارية بطرق بيانية خاصة وهامة يمكن من خلالها التعرف على الكثير من خصائص وصفات المجتمع، ويمكن حساب بعض المقاييس الإحصائية باستخدامها. ونعرض من هذه الطرق ما يلي:

1-طريقة المدرج التكراري Frequency Histogram

يتم عرض التوزيعات التكرارية بطريقة المدرج التكراري وذلك باتباع الخطوات التالية:

- 1- رسم محورين متعامدين وتقسيمهما بمقياس رسم مناسب.
- 2- يمثل المحور الأفقي حدود الفئات والجدول الرأسي التكرار.
- 3- نرسم فوق كل فئة مستطيلاً ارتفاعه يساوي التكرار المقابل لهذه الفئة.

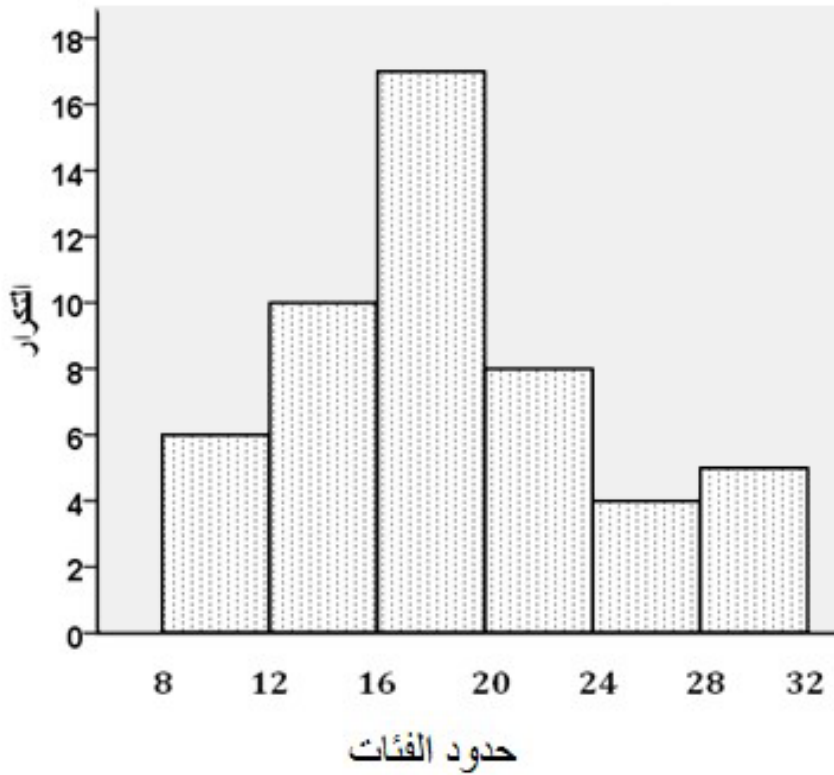
ولاحظ هنا أن المستطيلات تكون متلاصقة مع بعضها البعض وذلك بخلاف طريقة الأعمدة البيانية: ويمكن التعرف على خصائص هامة للتوزيع من المدرج التكراري كتماثل التوزيع والتوائه.

مثال 8.2: اعرض التوزيع التكراري التالي بطريقة المدرج التكراري:

32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

/الحل/

باتباع الخطوات التي ذكرناها أعلاه نجد أن المدرج التكراري يأخذ الشكل التالي:



2- طريقة المنحنى التراكمي الصاعد

Less than Cumulative Frequency Curve

يعرض التوزيع التراكمي بهذه الطريقة لأسباب هامة كثيرة منها حساب بعض المقاييس الإحصائية كالوسيط والمنئينات والربيعات والمدى الربيعي ومعرفة خصائص وصفات هامة للتوزيع وحساب الاحتمالات ولرسم المنحنى نتبع الخطوات التالية:

- 1- نكون الجدول التراكمي الصاعد وذلك بإضافة عمود للتوزيع يحتوي على الحدود العليا للفئات ومقابل كل حد نضع التكرار التراكمي والذي يساوي تكرار الفئة المقابلة مضافاً اليه كل تكرارات الفئات السابقة لها.

2- نرسم محورين متعامدين ونخصص المحور الأفقي للحدود العليا للفئات والمحور الرأسي للتكرار التراكمي.

3- نرسم في المستوى نقطة فوق كل حد أعلى ويقابل رأسياً التكرار التراكمي الصاعد.

4- نصل بين النقاط بمنحنى.

مثال 9.2: اعرض التوزيع التكراري التالي بطريقة المنحنى التراكمي الصاعد:

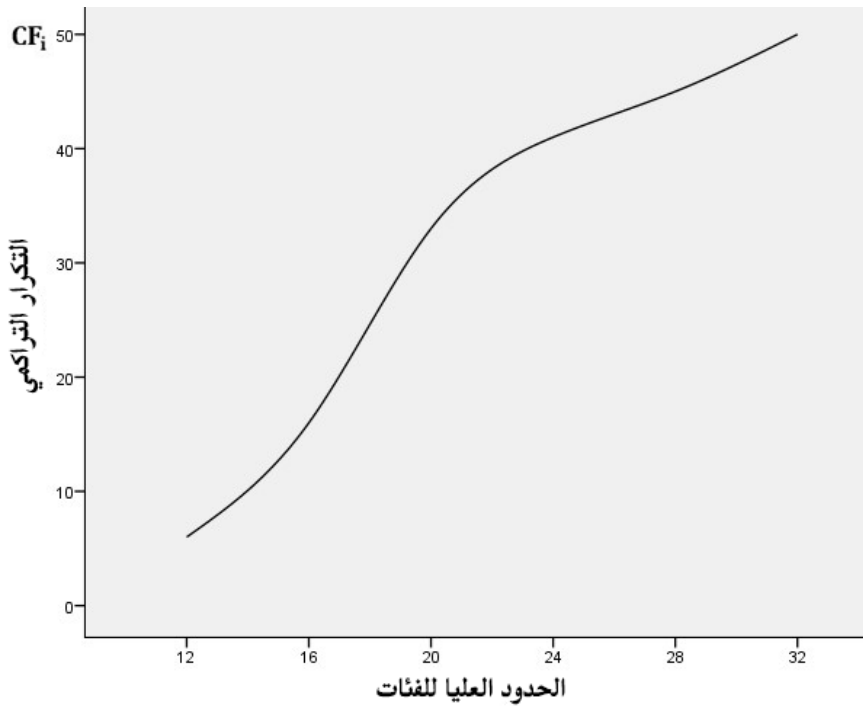
32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

الحل/

نكون الجدول التراكمي الصاعد كما يلي:

الفئات	التكرار	أقل من الحد الأعلى للفئة	التكرار المتجمع الصاعد CF_i
-8	6	أقل من 12	6
-12	10	أقل من 16	16
-16	17	أقل من 20	33
-20	8	أقل من 24	41
-24	4	أقل من 28	45
32-28	5	أقل من 32	50

ونستخدم العمود الثالث والرابع في رسم المنحنى ليظهر كما يلي:



3- طريقة المنحنى التراكمي الهابط

More than Cumulative Frequency Curve

لرسم المنحنى نتبع الخطوات التالية:

- 1- نكون الجدول التراكمي الهابط وذلك بإضافة عمود للتوزيع يحتوي على الحدود الدنيا للفئات ومقابل أول حد نضع مجموع التكرارات كلها ومقابل الحد الثاني نضع مجموع التكرارات مطروحاً منه تكرار الفئة الأولى ومقابل الحد الثالث نطرح من الناتج تكرار الفئة السابق لها وهكذا.
- 2- نرسم محورين متعامدين ونخصص المحور الأفقي للحدود الدنيا للفئات والمحور الرأسي للتكرار التراكمي.
- 3- نرسم في المستوى نقطة فوق كل حد أدنى ويقابل رأسياً التكرار التراكمي الهابط.
- 4- نصل بين النقاط بمنحنى.

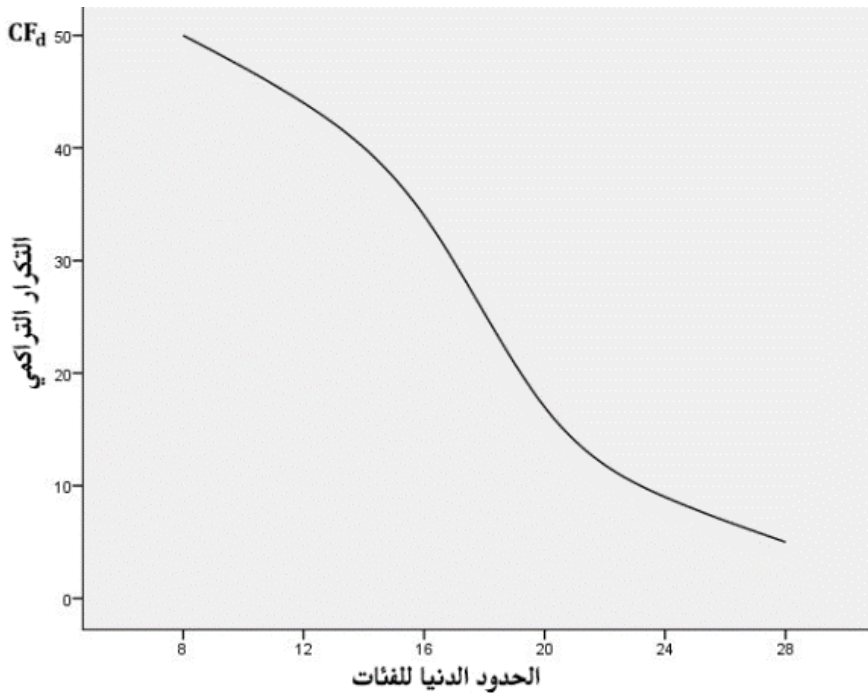
مثال 10.2: اعرض التوزيع التكراري التالي بطريقة المنحنى التراكمي الهابط:

32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

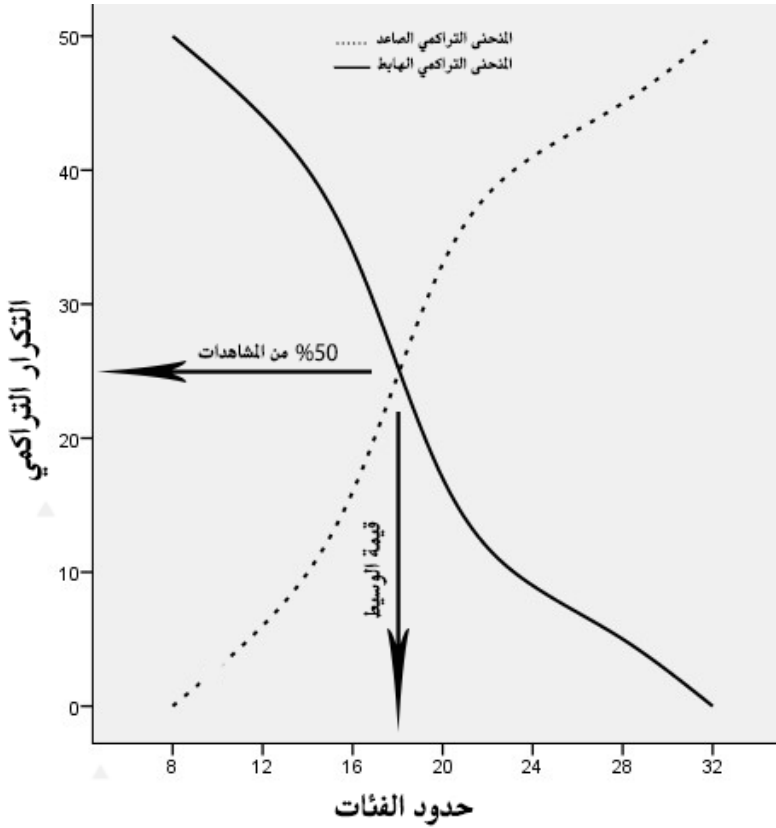
نكون الجدول التراكمي الهابط كما يلي:

الفئات	التكرار	أكثر من الحد الأدنى للفئة	التكرار المتجمع الهابط CF_d
-8	6	أكثر من 8	50
-12	10	أكثر من 12	44
-16	17	أكثر من 16	34
-20	8	أكثر من 20	17
-24	4	أكثر من 24	9
32-28	5	أكثر من 28	5

ونستخدم العمود الثالث والرابع في رسم المنحنى ليظهر كما يلي:



عند رسم المنحنيين التراكميين الصاعد والهابط في نفس الرسمة فإنهما يتقاطعان في نقطة هامة حيث يكون الإحداثي الأفقي لها يعبر عن قيمة الوسيط والذي يعرف بأنه القيمة التي تتوسط مشاهدات المجتمع، ويعبر الإحداثي الأفقي لها عن القيمة 50% من مشاهدات المجتمع والرسم التالي يوضح ذلك.



من الرسم نلاحظ أن السهم الأفقي يشير إلى 25 والتي تمثل نسبة 50% من المشاهدات، حيث أن العدد الكلي للمشاهدات يساوي 50.

والسهم الرأسي يشير لقيمة الوسيط والتي تساوي تقريباً 18 والتي تعني أن نصف المشاهدات قيمتها أقل من 18 والنصف الآخر قيمته أكبر من 18.

تمارين:

1- الجدول التالي يمثل أعداد الخريجين لكلية الهندسة في إحدى الجامعات في السنوات من 2000 إلى 2004:

السنة	2000	2001	2002	2003	2004
عدد الخريجين	110	135	150	150	175

مثل الجدول التكراري بالطرق التالية:

طريقة الأعمدة البيانية البسيطة - طريقة الخط المنكسر - طريقة الدائرة البيانية.

2- إذا كان لديك التوزيع التكراري التالي:

الفئات	-25	-30	-35	-40	50-45
التكرار	14	19	7	3	7

اعرض التوزيع التكراري بالطرق التالية:

- المدرج التكراري.
- المنحنى التراكمي الصاعد.
- المنحنى التراكمي الهابط.
- قدر قيمة الوسيط.

3- ثلاثة مصانع A,B,C للألبسة. إذا كان مجموع إنتاجها في عام 1997 يساوي 10000 قطعة. مُثِّلَ إنتاج هذه المصانع باستخدام طريقة القطاعات الدائرية وكانت زاوية قطاع إنتاج المصنع B هي 162 درجة. ما هو حجم إنتاج هذا المصنع في ذلك العام؟

4- المشاهدات التالية تمثل رواتب 40 موظفاً في دائرة حكومية:

117 200 137 145 113 117 115 109

225 230 113 145 115 225 250 113

185 180 175 113 200 113 250 117

137 148 248 237 245 240 195 190

219 213 209 173 167 195 194 188

كون توزيعاً تكرارياً ذي 6 فئات لهذه البيانات، ثم مثل هذه البيانات بالطرق التالية:

- المدرج التكراري.
- المنحنى التراكمي الصاعد.
- المنحنى التراكمي الهابط.
- قدر قيمة الوسيط.

5- الجدول التالي يوضح أعداد الطلبة المقيدين في أقسام الكلية المختلفة في الفصل الدراسي

الثاني للعام الجامعي 2014:

الهندسية	الحاسوب	الطبية	الإدارية	القسم	
150	350	670	850	ذكور	عدد الطلبة
100	200	750	700	اناث	

مثل الجدول بطريقة الأعمدة البيانية المزدوجة.

6- في سوق عمان المالي إذا كان عدد الأسهم المباعة في أحد الأيام ممثلة بالجدول التالي:

القطاع	عدد الأسهم
البنوك	29000
الخدمات	12000
التأمين	6000
الصناعة	123000
المجموع	170000

اعرض البيانات بطريقة الدائرة.

7- إذا كان متوسط الدخل الشهري بالدينار لخمسة موظفين في إحدى الجامعات كما يلي:

الموظف	المبلغ
أستاذ دكتور	3000
دكتور	2000
ماجستير	1400
حارس	200
آذن	600

اعرض الجدول بالطرق التالية:

- الدوائر البيانية.
- الأعمدة البيانية البسيطة.
- الخط المنكسر.

8- فيما يلي تقديرات 20 طالباً في أحد الامتحانات:

جيد	جيد جداً	ممتاز	جيد	ممتاز	مقبول	ضعيف	جيد جداً
مقبول	جيد جداً	جيد	ضعيف	جيد	مقبول	جيد جداً	مقبول
مقبول	ضعيف	مقبول	جيد				

- بوب البيانات في جدول تكراري مناسب.
- مثل الجدول التكراري الناتج بطريقة مناسبة.

9- الجدول التالي يبين توزيع درجات الحرارة في إحدى المدن في 120 يوماً بالدرجة المئوية:

درجة الحرارة	-8	-12	-16	-20	-24	-28	-32	40-36
عدد الايام	6	12	20	30	24	16	8	4

والمطلوب ما يلي:

- عرض التوزيع بطريقة المدرج التكراري.
- عرض التوزيع بطريقة المنحنى التراكمي الصاعد.
- عرض التوزيع بطريقة المنحنى التراكمي الهابط.
- قدر قيمة الوسيط.

10- الجدول التالي يبين دخل 80 أسرة في عمان بمئات الدنانير:

فئات الدخل	-16	-22	-28	-34	-40	46	58-52
عدد الأسر	4	10	16	24	14	8	4

مثل الجدول بالطرق التالية:

- المنحنى التراكمي الصاعد والهابط على نفس المستوى.
- قدر الدخل الذي تحصل 40 أسرة على دخل أقل منه.

11- الجدول التالي يمثل الواردات والصادرات لبلد ما في السنوات من 2007 إلى

2016 بملايين الدولارات:

السنة	7	8	9	10	11	12	13	14	15	16
الواردات	22	23	25	30	34	45	51	60	54	70
الصادرات	30	34	45	52	59	62	67	76	80	85

اعرض الجدول التكراري بالطرق التالية:

- الأعمدة البيانية المزدوجة.
- الخطوط المنكسرة المتعددة.
- اشرح الاتجاه العام للصادرات والواردات.

12- القراءات التالية تمثل درجات الحرارة التي تم رصدها في مدينة في شهر تموز:

34	33	31	28	29	29	32	32	30	31
34	31	30	31	30	29	26	27	27	28
30	29	30	34	35	37	36	35	34	34

- بوب البيانات في توزيع تكراري مناسب.
- اعرض التوزيع بطريقتي المنحنى التراكمي الصاعد والهابط في نفس المستوى.
- قدر قيمة الوسيط.

الخلاصة

- تعلمنا في هذا الفصل أن تنظيم البيانات عبر تبويبها وعرضها برسومات بيانية خطوة هامة جداً في التحليل الإحصائي لأنه بدونها يكون من الصعب الحصول على المعلومات ولا نستطيع التعرف على صفات وخصائص المجتمع بشكل جيد وتضيق الكثير من المعلومات الهامة.
- يتم تبويب البيانات النوعية في جداول تكرارية وتعرض بطرق كثيرة منها:
 - الأعمدة البيانية البسيطة والمزدوجة.
 - الخطوط المنكسرة.
 - الدوائر البيانية.
- يتم تبويب البيانات الكمية في توزيعات تكرارية وتعرض بطرق كثيرة منها:
 - المدرج التكراري.
 - المنحنى التراكمي الصاعد والهابط.
- يمكن معرفة قيمة الوسيط برسم المنحنيين التراكميين الصاعد والهابط معاً وتحديد نقطة تقاطعهما.

مقاييس النزعة المركزية Central Tendency Measures

3

من المهم جداً بعد جمع البيانات وتنظيمها البدء بعملية التحليل الإحصائي واستخراج المعلومات المطلوبة ووصف المجتمع والتعرف على خصائصه وإجراء المقارنات بين المجموعات، وهذا كله لا يمكن أن يتم بسهولة وبمستوى مرضٍ من الدقة دون اللجوء لحساب بعض المقاييس الإحصائية التي تساعد على تحقيق ذلك، ومن المقاييس الإحصائية الهامة مقاييس النزعة.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً

على:

- حساب الوسط الحسابي، الوسيط، المنوال للبيانات الأولية والتوزيعات التكرارية.
- حساب المئينات والربيعات للبيانات الأولية والتوزيعات التكرارية.
- تحديد تماثل التوزيعات من التوائها برسم المدرج التكراري ومقارنة مقاييس النزعة المركزية للمشاهدات.
- حساب مقاييس النزعة المركزية بعد تأثر المشاهدات بالتحويلات الخطية.

في هذا الفصل سنتطرق لمقاييس النزعة المركزية والتي لها نفس وحدات قياس المشاهدات، وسنتحدث فيه عن الوسط الحسابي، الوسط المشذب، الوسط الموزون، الوسيط، المئينات والربيعات وكيفية حسابها في حالة البيانات الخام غير المبوبة ومن ثم حسابها في حالة التوزيعات التكرارية.

أولاً: حساب مقاييس النزعة المركزية للمشاهدات الأولية

1-الوسط الحسابي (Arithmetic Mean (Average

يُعرّف الوسط الحسابي بأنه المشاهدة التي تتمركز حولها المشاهدات، وهي القيمة التي تنزع المشاهدات لأن تساويها. فمثلاً عندما ينتج مصنع مصابيح ضوئية ويكون العمر الافتراضي الذي حدده المصنع لهذه المصابيح وفق مواصفات معينة هو 1000 ساعة، فيكون هذا هو معدل أو متوسط الساعات التي سيعيش بها كل مصباح، وفي الواقع كل مصباح ينزع لأن يلتزم بهذه القيمة بمعنى أن يضيء 1000 ساعة ولكن قد يحول بينه وبين ذلك عوامل كثيرة.

يرمز للوسط الحسابي للمجتمع بالرمز μ وهو حرف لاتيني يقرأ (ميو) ويرمز للوسط الحسابي للعينات بأحد الأحرف الكبيرة الإنجليزية وعليها شرطة أفقية مثل $\bar{X}$ والذي تقرأ (إكس بار).

فبفرض أن لدينا المشاهدات التالية والتي عددها n :

$$x_1, x_2, \dots, x_n$$

يُحسب الوسط الحسابي لها (ويرمز له بالرمز $\bar{X}$) من العلاقة:

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

الوسط الحسابي للبيانات الأولية



مثال 1.3: أوجد الوسط الحسابي للملاحظات: 3، 7، 6، 5، 9.

الحل/ بالتطبيق المباشر لقانون الوسط الحسابي نجد أن:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{3 + 7 + 6 + 5 + 9}{5} = \frac{30}{5} = 6$$

وهذا يعني أن الملاحظات تتمركز حول القيمة 6 والتي تمثل الوسط الحسابي.

خواص ومزايا الوسط الحسابي

1- المتوسط الحسابي سهل التعريف والحساب.

2- المتوسط الحسابي وحيد القيمة.

3- المتوسط الحسابي تدخل في حسابه جميع المفردات لذلك يعتبر ممثلاً جيداً للملاحظات.

4- يتأثر الوسط الحسابي بضرب الملاحظات في أي عدد حقيقي أو إضافة أي عدد حقيقي

لها، بمعنى أنه لو أثّرنا على ملاحظات يمثلها المتغير X ووسطها $\bar{X}$ بالتحويل الخطي

$Y = aX + b$ فإن الوسط الحسابي للملاحظات الجديدة $\bar{Y}$ والتي يمثلها المتغير Y

يساوي: $\bar{Y} = a\bar{X} + b$.

في المثال (1.3) لو ضربنا الملاحظات بالعدد 2 فإننا سنحصل على الملاحظات

الجديدة وهي: Y : 6، 14، 12، 10، 18 ويكون الوسط الحسابي الجديد:

$$\bar{Y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{6 + 14 + 12 + 10 + 18}{5} = \frac{60}{5} = 12$$

لاحظ هنا أن الوسط الحسابي الجديد يساوي الوسط القديم مضروباً في 2 أي:

$$\bar{Y} = 2\bar{X}$$

وهذا يعني أن الوسط الحسابي يتأثر بضرب الملاحظات بعدد حقيقي وبنفس التأثير.

وبنفس الطريقة عند إضافة القيمة 3 لملاحظات مثال 1.3 فإننا نحصل على الملاحظات الجديدة التالية: $Y : 12, 8, 9, 10, 6$ ويساوي الوسط الحسابي الجديد:

$$\bar{Y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{6 + 10 + 9 + 8 + 12}{5} = \frac{45}{5} = 9$$

ولاحظ أن الوسط الجديد يساوي الوسط القديم مضافاً إليه 3 أي:

$$\bar{Y} = \bar{X} + 3$$

وهذا يعني أن الوسط الحسابي يتأثر بإضافة عدد حقيقي للملاحظات وبنفس التأثير.

مثال 2.3: تقدم 30 طالباً لامتحان في الكيمياء العضوية وكان متوسط درجاتهم 15 حيث كانت النهاية العظمى للامتحان 50، فلجأ المدرس لتعديل الدرجات بضربها في 0.6 وإضافة 20 درجة لكل منها. أوجد متوسط الدرجات بعد التعديل؟

الحل/

بفرض أن المتغير X يمثل الدرجات قبل التعديل والمتغير Y يمثل الدرجات بعد التعديل فيكون التحويل الخطي المستخدم هو: $Y = 0.6X + 20$ وبالتالي يكون الوسط الحسابي للدرجات بعد التعديل يساوي:

$$\bar{Y} = 0.6 \times 15 + 20 = 29$$

5- مجموع انحرافات الملاحظات عن وسطها الحسابي يساوي صفراً. أي أن:

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

حيث نعني بانحرافات القيم عن وسطها الحسابي الفروق بين القيم ووسطها الحسابي. ولتوضيح هذه الخاصية نوجد مجموع انحرافات القيم عن وسطها الحسابي في مثال 1.3 وهو كما يلي:

الفروق بين القيم ووسطها الحسابي هي:

$$3 - 6 = -3, 7 - 6 = 1, 6 - 6 = 0, 5 - 6 = -1, 9 - 6 = 3$$

ومجموع الفروق يساوي:

$$\sum_{i=1}^n (x_i - 6) = 3 + (-1) + 0 + 1 + (-3) = 0$$

6- مجموع مربعات انحرافات القيم عن وسطها الحسابي أصغر ما يمكن، أي:

$$\sum_{i=1}^n (x_i - \bar{X})^2 < \sum_{i=1}^n (x_i - a)^2$$

حيث a أي عدد حقيقي لا يساوي $\bar{X}$.

فمثلاً عند حساب مربعات انحرافات مشاهدات مثال 1.3 عن وسطها الحسابي 6 نجد أنها تساوي:

$$\sum_{i=1}^n (x_i - 6)^2 = 9 + 1 + 0 + 1 + 9 = 20$$

وفي المقابل لنحسب مربع انحرافات المشاهدات نفسها عن القيمة 2 لنجد أنها تساوي:

$$3 - 2 = 1, 7 - 2 = 5, 6 - 2 = 4, 5 - 2 = 3, 9 - 2 = 7$$

$$\sum_{i=1}^n (x_i - 2)^2 = (1)^2 + (5)^2 + (4)^2 + (3)^2 + (7)^2 = 100$$

وواضح الفرق بين المجموعين.

عيوب الوسط الحسابي

1- يتأثر الوسط الحسابي بحجم العينة، وكلما كان حجم العينة أكبر كلما كانت قيمته أقرب من وسط المجتمع.

2- لا يمكن حساب الوسط الحسابي ببيانياً.

3- الوسط الحسابي يتأثر بالقيم الشاذة والمتطرفة.

فقد يظهر في المشاهدات عند جمع البيانات قيم تصغر أو تكبر بشكل ملحوظ عن باقي القيم وتظهر مثل هذه القيم عند ترتيب المشاهدات تصاعدياً أو تنازلياً وتتواجد في الطرفين بعد الترتيب. هذه القيم تؤثر في الوسط الحسابي عند حسابه بحيث تجعله يتجه نحوها وبالتالي سيؤثر في استنتاجاتنا وتحليلنا للبيانات. ولتوضيح ذلك تأمل المشاهدات:

500 4 3 6 5 2

حيث نلاحظ أن المشاهدة 500 تعتبر مشاهدة متطرفة، وبحساب الوسط الحسابي للمشاهدات نجد أنه يساوي:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{500 + 4 + 3 + 6 + 5 + 2}{6} = \frac{520}{6} = 86.6$$

وواضح أن المشاهدات المذكورة لا تتمركز حول هذه القيمة وبالتالي لا يمكن استخدام الوسط الحسابي في وصف هذه المشاهدات. ولتلافي هذا العيب يمكن حساب الوسط الحسابي المشذب.

2- الوسط الحسابي المشذب Trimmed Mean

يحسب الوسط الحسابي المشذب بنفس الطريقة التي حسبنا فيها الوسط الحسابي ولكن هنا يتم استبعاد 5% من المشاهدات من الطرفين وذلك بعد ترتيب المشاهدات تصاعدياً وذلك لضمان حذف المشاهدات الشاذة والمتطرفة.

مثال 3.3:

أوجد الوسط الحسابي المشذب للملاحظات التالية:

26 26 25 21 20 19 15 13 10 9 8 6 6 5 4 3 2
70 27 27

/الحل/

من الملاحظ أن الملاحظة 70 تبعد عن باقي القيم بمسافة كبيرة لذلك نظرياً يمكن اعتبارها مشاهدة شاذة أو متطرفة لذلك سنلجأ لحساب الوسط الحسابي المشذب. ولاحظ كذلك أن عدد المشاهدات هو 20 مشاهدة وهي مرتبة تصاعدياً، ويكون عدد المشاهدات المستبعد من الطرفين هو $1 = 20 \times 0.05$ وهذا يعني أن نستبعد أول مشاهدة والتي قيمتها 2 وآخر مشاهدة والتي قيمتها 70 وبالتالي يكون الوسط الحسابي المشذب يساوي:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{270}{18} = 15$$

3-الوسط الحسابي المرجح Weighted Arithmetic Mean

بفرض أنه لدينا مجموعتين من المشاهدات، الأولى وسطها $\bar{X}_1$ وعددها n_1 والثانية وسطها $\bar{X}_2$ وعددها n_2 ، فإذا دمجت المجموعتان معاً في مجموعة واحدة فإن الوسط الحسابي لها والذي يعرف بالوسط الحسابي المرجح أو الوسط الحسابي الموزون يعطى بالعلاقة:

$$\bar{X} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2}{n_1 + n_2}$$

الوسط الحسابي المرجح



ويمكن التعميم لأكثر من مجموعتين.

مثال 4.3: تقدم ثلاثة شعب لامتحان مبادئ الإحصاء. الشعبة الأولى فيها 15 طالباً متوسط درجاتهم 15، والثانية فيها 30 طالباً متوسط درجاتهم 20، والثالثة فيها 25 طالب متوسط درجاتهم 14. احسب متوسط درجات الطلبة في الشعب الثلاثة.

الحل/

من السؤال نجد أن:

$$n_1 = 15, \bar{X}_1 = 15, n_2 = 30, \bar{X}_2 = 20, n_3 = 25, \bar{X}_3 = 14$$

وبالتطبيق المباشر للقانون نحسب قيمة الوسط الحسابي المرجح:

$$\begin{aligned}\bar{X} &= \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + n_3\bar{X}_3}{n_1 + n_2 + n_3} = \frac{15 \times 15 + 30 \times 20 + 25 \times 14}{15 + 30 + 25} \\ &= \frac{1175}{70} = 16.8\end{aligned}$$

4-الوسيط Median

يعرف الوسيط بأنه القيمة التي تتوسط المشاهدات، حيث يكون نصف المشاهدات أقل منها أو يساويها والنصف الآخر أكبر منها وذلك بعد ترتيب المشاهدات تصاعدياً أو تنازلياً، و سنرمز للوسيط بالرمز **Me** .

لحساب الوسيط نرتب المشاهدات تصاعدياً من اليمين إلى اليسار ثم نحسب ترتيب الوسيط من العلاقة:

$$O_{Me} = \frac{(n + 1)}{2}$$

لا تنسى ترتيب الوسيط للبيانات الأولية



فإذا كان ترتيب الوسيط عدداً صحيحاً نعد المشاهدات المرتبة وتكون المشاهدة صاحبة الترتيب هي الوسيط أما إذا كان ترتيب الوسيط يحتوي على كسر فيتم حساب الوسيط بطريقة سنطلق عليها طريقة الميزان. كما سنوضح في المثالين التاليين:

مثال 5.3: أوجد الوسيط للمشاهدات 1 7 9 16 7 10 18
الحل/ بترتيب المشاهدات تصاعدياً:

1 7 7 9 10 16 18

نجد ترتيب الوسيط كما يلي:

$$O_{Me} = \frac{(n + 1)}{2} = \frac{8}{2} = 4$$

أي أن الوسيط هو المشاهدة الرابعة والتي قيمتها 9 وبالتالي يكون:

$$Me = 9$$

مثال 6.3: أوجد الوسيط للمشاهدات:

3 6 1 11 8 7 1 5

الحل/

نرتب المشاهدات تصاعدياً كما يلي:

1 1 3 5 6 7 8 11

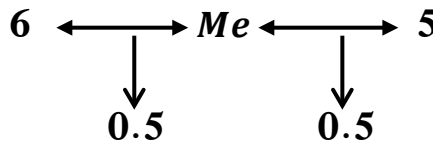
نجد ترتيب الوسيط كما يلي:

$$O_{Me} = \frac{(n + 1)}{2} = \frac{9}{2} = 4.5$$

وبالتالي تكون قيمة الوسيط واقعة بين المشاهدة الرابعة وهي قيمة العدد الصحيح في ترتيب الوسيط والمشاهدة الخامسة وهي التي تليها مباشرة.

11 8 7 6 5 3 1 1

نكون الميزان حيث نضع على الطرفين المشاهدة الرابعة والخامسة وتكون قيمة الوسيط واقعة في منتصف المسافة بين المشاهدين وتبعد عن كل مشاهدة مسافة مقدارها 0.5



ثم نحسب قيمة الوسيط بطريقة الميزان وذلك كما يلي:

$$Me = 5 \times 0.5 + 6 \times 0.5 = 5.5$$

لاحظ هنا أن الوسيط يساوي الوسط الحسابي للمشاهدين الرابعة والخامسة.

خواص الوسيط

- 1- الوسيط سهل التعريف والحساب.
 - 2- الوسيط وحيد القيمة.
 - 3- لا يتأثر بالقيم الشاذة.
 - 4- يمكن حسابه بيانياً (برسم المنحنى التراكمي الصاعد أو الهابط أو كليهما).
- ومن عيوب الوسيط أنه لا يعتبر ممثلاً جيداً للقيم حيث أنه يركز على القيم التي تقع في المنتصف فقط ولا تدخل في حسابه جميع القيم.

5-المنوال Mode

سنرمز للمنوال بالرمز **Mo**، ويُعرّف بأنه الملاحظة الأكثر تكراراً. وعند عرض البيانات بيانياً نلاحظ أن المنوال يُمثّل قِمّة.

مثال 7.3: أوجد المنوال للملاحظات 3 2 4 5 7 2

الحل/

واضح من القيم أن الملاحظة التي قيمتها 2 هي الأكثر تكراراً بين الملاحظات لذلك يكون:

$$Mo = 2$$

ملاحظة/

إذا كانت جميع المفردات لها نفس التكرار فنقول إنه لا يوجد منوال.

خواص المنوال

- 1- يعتبر أسهل مقاييس النزعة المركزية حساباً.
- 2- يكون لبعض البيانات أكثر من منوال.
- 3- لا يتأثر بالقيم الشاذة.

عيوب المنوال

- 1- ليس له معنى إذا كان عدد القيم قليلاً.
- 2- أنه من أضعف مقاييس النزعة المركزية.
- 3- يتأثر بتكرار القيم وليس بالقيم نفسها.

العلاقة الخطية بين الوسط والوسيط والمنوال

في التوزيعات أحادية المنوال أُسْتُتِجَ من الملاحظة والتجربة أن هناك علاقة خطية تربط مقياس النزعة المركزية وهذه العلاقة هي:

لا تنسى

العلاقة الخطية بين مقاييس النزعة المركزية

$$Mo = 3Me - 2\bar{X}$$



وتستخدم هذه العلاقة لإيجاد أحد المقاييس إذا علم الاثنان الآخران.

مثال 8.3: إذا كانت الدرجة الأكثر شيوعاً لطلبة العلوم الإدارية هي 80 والدرجة التي يقل عنها 50% من الدرجات هي 75، فما قيمة الوسط الحسابي لهذه الدرجات؟

الحل/ من السؤال نلاحظ أن $Mo = 80$ و $Me = 75$ لذلك و بالتطبيق المباشر للعلاقة السابقة نجد أن:

$$Mo = 3Me - 2\bar{X}$$

$$80 = 3 \times 75 - 2\bar{X}$$

$$2\bar{X} = 145$$

$$\bar{X} = 72.5$$

6-المئينات والربيعات Percentiles and Quartiles

لمجموعة من المشاهدات يُعرّف المئين k ويرمز له بالرمز P_k بأنه القيمة التي يقل عنها أو يساويها $k\%$ من المشاهدات وتسمى k رتبة المئين.

هناك ثلاثة مئينات خاصة نطلق عليه اسم ربيعات وهي:

الربيع الأول (الربيع الأدنى) ويرمز له بالرمز Q_1 ويعرف بأنه المئين الخامس

$$Q_1 = P_{25}$$

الربيع الثاني (الربيع الأوسط) ويرمز له بالرمز Q_2 ويعرف بأنه المئين الخمسون أي:

$$Q_2 = P_{50}$$

الربيع الثالث (الربيع الأعلى) ويرمز له بالرمز Q_3 ويعرف بأنه المئين الخامس

$$Q_3 = P_{75}$$

ونلاحظ من التعريف أن الربيع الثاني هو الوسيط وبالتالي:

$$Q_2 = P_{50} = Me$$

ويتم حساب المئينات والربيعات كما يلي:

1- نرتب المشاهدات تصاعدياً من اليمين إلى اليسار.

2- نحسب ترتيب المئين من العلاقة التالية حيث k رتبة المئين، n عدد المشاهدات:

$$O_{P_k} = (n + 1) \times \frac{k}{100}$$

ترتيب المئين



3- إذا كان الترتيب عدداً صحيحاً كانت المشاهدة صاحبة هذا الترتيب من اليمين هي المئين المطلوب.

4- إذا كان الترتيب عدداً كسرياً نعمل ميزاناً بنفس الطريقة التي حسبنا فيها الوسيط كما سنوضح في المثال التالي:

مثال 9.3: تم رصد درجة الحرارة في مدينة رفح لعشرة أيام متتالية في شهر تموز وكانت النتائج كما يلي:

30 35 35 30 32 33 28 27 31 33

أحسب P_{40} , Q_1 , Q_3 .

الحل/

نرتب البيانات تصاعدياً من اليمين إلى اليسار:

35 35 33 33 32 31 30 30 28 27

أولاً: إيجاد P_{40}

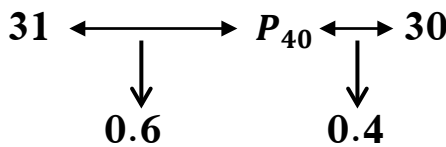
نحسب ترتيب المئين:

$$O_{P_{40}} = (n + 1) \times \frac{k}{100} = 11 \times \frac{40}{100} = 4.4$$

وبالتالي تكون قيمة المئين واقعة بين المشاهدة الرابعة والمشاهدة الخامسة

35 35 33 33 32 31 30 30 28 27

ويبعد عن المشاهدة الرابعة مسافة 0.4 وعن المشاهدة الخامسة مسافة 0.6



ثم نحسب قيمة المئين بطريقة الميزان وذلك كما يلي:

$$P_{40} = 30 \times 0.6 + 31 \times 0.4 = 30.4$$

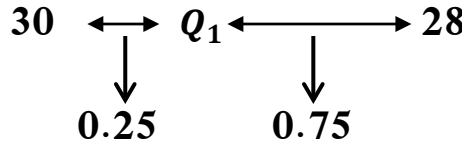
ثانياً: حساب الربع الأول Q_1 حيث نعلم أن $Q_1 = P_{25}$

$$O_{P_{25}} = (n + 1) \times \frac{k}{100} = 11 \times \frac{25}{100} = 2.75$$

وبالتالي تكون قيمة الربع الأول واقعة بين المشاهدة الثانية والمشاهدة الثالثة

35 35 33 33 32 31 30 30 28 27

ويبعد عن المشاهدة الثانية مسافة 0.75 وعن المشاهدة الثالثة مسافة 0.25



ثم نحسب قيمة الربع الأول بطريقة الميزان وذلك كما يلي:

$$Q_1 = 28 \times 0.25 + 30 \times 0.75 = 29.5$$

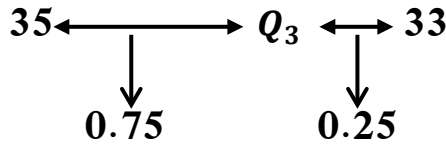
ثالثاً: حساب الربع الثالث Q_3 حيث نعلم أن $Q_3 = P_{75}$

$$O_{P_{75}} = (n + 1) \times \frac{k}{2} = 11 \times \frac{75}{100} = 8.25$$

وبالتالي تكون قيمة الربع الثالث واقعة بين المشاهدة الثامنة والمشاهدة التاسعة

35 35 33 33 32 31 30 30 28 27

ويبعد عن الملاحظة الثامنة مسافة 0.25 وعن الملاحظة التاسعة مسافة 0.75



ثم نحسب قيمة الربع الثالث بطريقة الميزان وذلك كما يلي:

$$Q_3 = 33 \times 0.75 + 35 \times 0.25 = 33.5$$

ثانياً: حساب مقاييس النزعة المركزية للتوزيعات التكرارية

1- الوسط الحسابي

في حالة التوزيعات التكرارية يحسب الوسط الحسابي من العلاقة:

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

الوسط الحسابي للتوزيعات التكرارية



حيث f_i تكرار الفئة، n مجموع التكرارات، أي $n = \sum f_i$ و x_i مركز الفئة i حيث يعرف مركز الفئة بأنه الوسط الحسابي لحدي الفئة بمعنى:

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{2}$$



مثال 10.3: أوجد الوسط الحسابي للتوزيع التكراري التالي:

40-36	-32	-28	-24	-20	الفئة
5	9	20	17	9	التكرار (f_i)

الحل/

أولاً نكون جدول الحساب والذي نضع في أول عمود فيه الفئات والعمود الثاني التكرار والعمود الثالث مراكز الفئات (x_i) أما الأعمدة الباقية فتكون حسب المجاميع الموجودة في قوانين المقاييس الإحصائية.

جدول الحساب

$x_i f_i$	x_i	f_i	الفئات
198	22	9	-20
442	26	17	-24
600	30	20	-28
306	34	9	-32
190	38	5	40-36
1736		60	المجموع

ويكون الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^n x_i f_i}{n} = \frac{1736}{60} = 28.9$$

2- الوسيط

لحساب الوسيط في حالة التوزيعات التكرارية نتبع الخطوات التالية:

- 1- نكون الجدول التراكمي المتجمع الصاعد للتوزيع التكراري.
- 2- نحسب التكرار التراكمي للوسيط من العلاقة التالية حيث n مجموع التكرارات .

التكرار التراكمي للوسيط

$$CF = \frac{n}{2}$$



- 3- نحدد الفئة الوسيطة وهي أول فئة تكرارها التراكمي أكبر أو يساوي التكرار التراكمي للوسيط.

- 4- نحسب قيمة الوسيط من العلاقة التالية:

الوسيط للتوزيعات التكرارية

$$Me = a + \frac{CF - n_1}{f_m} \times c$$



حيث:

a : الحد الأدنى للفئة الوسيطة.

n_1 : التكرار التراكمي للفئة قبل الوسيطة.

f_m : تكرار الفئة الوسيطة.

c : طول الفئة الوسيطة.

مثال 11.3: أوجد الوسيط للتوزيع التكراري التالي:

32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

الحل/

من التوزيع التكراري نجد أن مجموع التكرارات يساوي 50 وبالتالي يكون التكرار التراكمي

للوسيط يساوي:

$$CF = \frac{n}{2} = \frac{50}{2} = 25$$

نكون الجدول التراكمي للتوزيع التكراري كما يلي:

الفئات	التكرار	أقل من الحدود العليا للفئات	التكرار التراكمي الصاعد
-8	6	أقل من 12	6
-12	10	أقل من 16	16
-16	17	أقل من 20	33
-20	8	أقل من 24	41
-24	4	أقل من 28	45
32-28	5	أقل من 32	50

ونلاحظ من الجدول أن الفئة الوسيطة هي الفئة الثالثة لأنها أول فئة تكرارها التراكمي

أكبر من أو يساوي التكراري التراكمي للوسيط، وبالتالي يكون:

$$16 = a \quad \text{وهي الحد الأدنى للفئة الوسيطة.}$$

$$16 = n_1 \quad \text{وهي التكرار التراكمي للفئة قبل الوسيطة.}$$

$$17 = f_m \quad \text{وهي تكرار الفئة الوسيطة.}$$

$$4 = c \quad \text{وهي طول الفئة الوسيطة (طول الفئة = حدها الأعلى - حدها الأدنى).}$$

وبالتعويض المباشر في قانون الوسيط نجد أن:

$$Me = a + \frac{CF - n_1}{f_m} \times c = 16 + \frac{25 - 16}{17} \times 4 = 18.12$$

لاحظ أن قيمة الوسيط تقع ضمن الفئة الثالثة وهي الفئة الوسيطة.

ملاحظة/

يمكن حساب الوسيط بطريقة التناسب كما يلي:

$$\begin{array}{ccc} & 16 & \\ 17 \left[\begin{array}{c} \\ 25 \\ \\ \end{array} \right. & \begin{array}{c} \\ 9 \\ \end{array} & \\ & 33 & \end{array} \quad \begin{array}{ccc} \underline{16} & & \\ & Me & \\ & 20 & \end{array} \left[\begin{array}{c} \\ 4 \\ \end{array} \right.$$

حيث الصف العلوي يمثل الحد الأعلى للفئة قبل الوسيطة وتكرارها التراكمي والصف السفلي الحد الأعلى للفئة الوسيطة وتكرارها التراكمي وبينهما نكتب الوسيط بين الحدود العليا والتكرار التراكمي للوسيط بين التكرارين التراكميين ثم نحسب الفروق المبينة في التناسب ونحسب قيمة الوسيط كما يلي:

$$Me = \underline{16} + \frac{9}{17} \times 4 = 18.12$$

3- المنوال

هناك طرق عديدة لحساب المنوال للتوزيعات التكرارية مثل طريقة الرافعة، طريقة الفروق وهناك من يعتبر أن المنوال هو مركز الفئة التي يقابلها أكبر تكرار.

وسنذكر هنا طريقة الفروق، حيث نحدد في البداية الفئة المنوالية وهي الفئة التي تقابل أكبر تكرار ثم نحسب قيمة المنوال من العلاقة:

لا تنسى المنوال للتوزيعات التكرارية



$$Mo = a + \frac{f_m - f_b}{(f_m - f_b) + (f_m - f_a)} \times c$$

حيث:

a : الحد الأدنى للفئة المنوالية.

c : طول الفئة المنوالية.

f_m : تكرار الفئة المنوالية.

f_a : تكرار الفئة بعد المنوالية مباشرة.

f_b : تكرار الفئة قبل المنوالية مباشرة.

مثال 12.3: أوجد المنوال للتوزيع التكراري التالي:

32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

الحل/

لاحظ أن أكبر تكرار هو 17 ويقابل الفئة " -16 " والتي سنعتبرها الفئة المنوالية وبالتالي يكون:

$$10 = f_b , 8 = f_a , 17 = f_m , 4 = c , 16 = a$$

وبالتالي تكون قيمة المنوال:

$$Mo = a + \frac{f_m - f_b}{(f_m - f_b) + (f_m - f_a)} \times c$$
$$= 16 + \frac{17 - 10}{(17 - 10) + (17 - 8)} \times 4 = 17.75$$

عند حساب المنوال بطريقة مركز الفئة المنوالية نجد أن المنوال يساوي 18 ولاحظ اختلاف قيمته عن القيمة المحسوبة بالطريقة السابقة، لذلك ننوه إلى أن الطرق المستخدمة لحساب المنوال هي طرق تقريبية.

4- المئينات والربيعات

ذكرنا خلال هذا الفصل أن الوسيط = الربيع الثاني = المئين الخمسون وبالتالي لحساب المئينات والربيعات للتوزيع التكراري نطبق نفس القانون المستخدم لحساب الوسيط مع مراعاة أن التكرار التراكمي للمئين P_k يحسب من العلاقة:

$$CF(P_k) = n \times \frac{k}{100}$$

التكرار التراكمي للمئين



حيث n مجموع التكرارات و k رتبة المئين.

ويحسب المئين للتوزيعات التكرارية من القانون:

لا تنسى المئين للتوزيعات التكرارية



$$P_k = a + \frac{CF(P_k) - n_1}{f_m} \times c$$

مثال 13.3: للتوزيع التكراري التالي:

32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

أحسب:

1- المئين السبعون (P_{70}).

2- الربع الثالث.

الحل/

أولاً: حساب المئين السبعون P_{70} :

من التوزيع التكراري نجد أن مجموع التكرارات يساوي 50 وبالتالي يكون التكرار

التراكمي للمئين السبعون يساوي:

$$CF(P_{70}) = n \times \frac{k}{100} = 50 \times \frac{70}{100} = 35$$

والجدول التراكمي الصاعد كما يلي:

الفئات	التكرار	أقل من الحدود العليا للفئات	التكرار التراكمي الصاعد
-8	6	أقل من 12	6
-12	10	أقل من 16	16
-16	17	أقل من 20	33
-20	8	أقل من 24	41
-24	4	أقل من 28	45
32-28	5	أقل من 32	50

ونلاحظ من الجدول أن الفئة المئينية هي الفئة الرابعة لأن تكرارها التراكمي هو أول تكرار تراكمي أكبر من أو يساوي التكراري التراكمي للمئين السبعون، وبالتالي يكون:

$$20 = a \quad \text{وهي الحد الأدنى للفئة المئينية.}$$

$$33 = n_1 \quad \text{وهي التكرار التراكمي للفئة قبل المئينية.}$$

$$8 = f_m \quad \text{وهي تكرار الفئة المئينية.}$$

$$4 = c \quad \text{وهي طول الفئة المئينية.}$$

وبالتعويض المباشر في قانون المئين نجد أن:

$$P_{70} = a + \frac{CF(P_{70}) - n_1}{f_m} \times c = 20 + \frac{35 - 33}{8} \times 4 = 21$$

ثانياً: حساب الربع الثالث $Q_3 = P_{75}$:

التكرار التراكمي للربع الثالث يساوي:

$$CF(P_{75}) = n \times \frac{k}{100} = 50 \times \frac{75}{100} = 37.5$$

من الجدول التراكمي الصاعد التالي نجد أن أول فئة تكرارها التراكمي أكبر من 37.5 هي الفئة الرابعة لذلك تكون هي الفئة المطلوبة.

الفئات	التكرار	أقل من الحدود العليا للفئات	التكرار التراكمي الصاعد
-8	6	أقل من 12	6
-12	10	أقل من 16	16
-16	17	أقل من 20	33
-20	8	أقل من 24	41
-24	4	أقل من 28	45
32-28	5	أقل من 32	50

من الجدول التكراري الصاعد نجد أن:

$$20 = a \quad \text{وهي الحد الأدنى للفئة المئينية.}$$

$$33 = n_1 \quad \text{وهي التكرار التراكمي للفئة قبل المئينية.}$$

$$8 = f_m \quad \text{وهي تكرار الفئة المئينية.}$$

$$4 = c \quad \text{وهي طول الفئة المئينية.}$$

وبالتالي:

$$Q_3 = a + \frac{CF(P_{75}) - n_1}{f_m} \times c = 20 + \frac{37.5 - 33}{8} \times 4 = 22.25$$

الالتواء Skewness

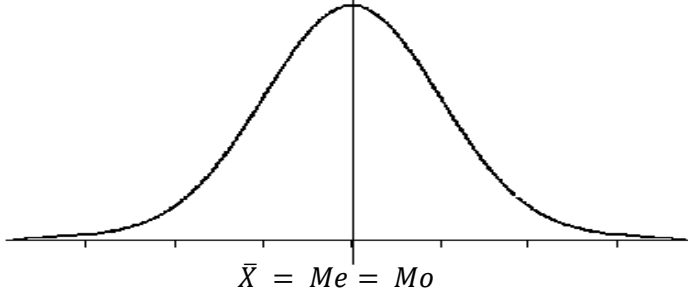
من المفاهيم الهامة التي تصف التوزيعات وتوضح بعض صفات المجتمعات مفهوم

الالتواء فقد يكون التوزيع غير ملتوٍ (أي متماثلاً) أو ملتوياً نحو اليمين وقد يكون ملتوياً نحو اليسار،

وسنوضح مفهوم الالتواء من خلال الرسومات التالية ولن نتطرق لحسابه:

التوزيعات المتماثلة Symmetric

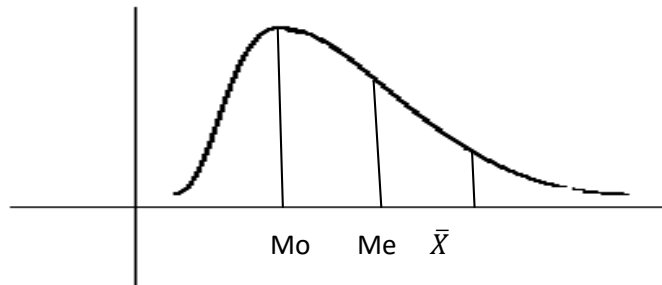
في حالة التوزيعات المتماثلة يكون معامل الالتواء يساوي صفراً ونقول إن التوزيع غير ملتو ويكون التوزيع متماثلاً حول الوسط الحسابي:



وفي هذه الحالة يكون: $Mo = Me = \bar{X}$

التوزيع الموجب الالتواء (ملتو نحو اليمين) Skewed to the right

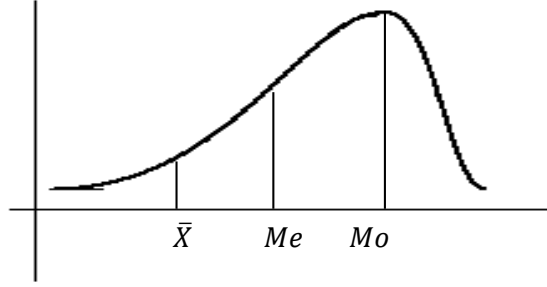
حيث لا يتماثل التوزيع حول الوسط ويكون هناك زيادة في المشاهدات التي قيمتها أقل من الوسط الحسابي كأن يحصل معظم طلبة شعبة على درجات منخفضة أقل من الوسط الحسابي في الامتحان كما هو موضح في الشكل التالي:



ولاحظ في هذه الحالة أن: $Mo \leq Me \leq \bar{X}$

سالب الالتواء (ملتوي نحو اليسار) Skewed to the left

حيث لا يتماثل التوزيع حول الوسط ويكون هناك زيادة في المشاهدات التي قيمتها أكبر من الوسط الحسابي كأن يحصل معظم طلبة شعبة على درجات أكبر من الوسط الحسابي في الامتحان كما هو موضح في الشكل التالي:



ولاحظ في هذه الحالة أن: $\bar{X} \leq Me \leq Mo$

وكلما كان معامل الالتواء مقترباً من الصفر كلما كان التوزيع أقرب للتماثل.

التحويلات الخطية ومقاييس النزعة المركزية:

يُعرَّف التحويل الخطي بالعلاقة $Y = aX + b$ والذي يحول مشاهدة x_i الى مشاهدة الجديدة y_i عن طريق ضرب مشاهدة القديمة x_i في a وإضافة b الى الناتج.

وقد ذكرنا في خواص الوسط الحسابي تأثره بالتحويلات الخطية، والسؤال الذي يطرح نفسه هل تتأثر بقية مقاييس النزعة المركزية بالتحويل الخطي المذكور؟ الإجابة نعم. والنظرية التالية توضح ذلك.

نظرية 1.3



بفرض أنه لدينا المشاهدات $x_n, \dots, x_2, x_1$ و أثرنا عليها
بالتحويل الخطي $Y = aX + b$ لنحصل على المشاهدات
الجديدة $y_n, \dots, y_2, y_1$ على الترتيب فإن:

$$\bar{Y} = a \times \bar{X} + b$$

$$Me_Y = a \times Me_X + b$$

$$Mo_Y = a \times Mo_X + b$$

حيث $\bar{X}$ ، Me_X ، Mo_X تمثل الوسط، الوسيط والمنوال للملاحظات القديمة
 $x_n, \dots, x_2, x_1$ على الترتيب و $\bar{Y}$ ، Me_Y ، Mo_Y تمثل الوسط، الوسيط والمنوال
للملاحظات الجديدة $y_n, \dots, y_2, y_1$ على الترتيب.

مثال 14.3: إذا حسبنا مقاييس النزعة المركزية لمجموعة من المشاهدات فكانت كالتالي:

الوسط الحسابي يساوي 5 ، الوسيط يساوي 7 ، المنوال يساوي 10 وتم تعديل
الملاحظات حسب العلاقة $Y = 3 - 2X$ حيث X تمثل القيمة القديمة ، Y تمثل القيمة
الجديدة بعد التعديل، فاحسب الوسط الحسابي، الوسيط و المنوال بعد التعديل.

الحل/ معلوم لدينا في السؤال ما يلي:

$$Mo_X = 10 , Me_X = 7 , \bar{X} = 5$$

وبالتطبيق المباشر للنظرية نحصل على المقاييس للملاحظات الجديدة كما يلي:

$$\bar{Y} = 3 - 2\bar{X} = 3 - 2 \times 5 = -7$$

$$Me_Y = 3 - 2Me_X = 3 - 2 \times 7 = -11$$

$$Mo_Y = 3 - 2Mo_X = 3 - 2 \times 10 = -17$$

تمارين:

1- إذا اخترنا عينة عشوائية من الأسر حجمها 10 وسُجِّل دخل كل أسرة بالدينار وكانت كما يلي:

150 170 200 250 130 120 130 100 110 190

فأوجد:

- الوسط الحسابي لدخل الأسرة في هذه العينة.
- الوسيط.
- المنوال.
- الربيعين الأول والثالث.

2- الجدول التالي يبين فئات أطوال 45 طالباً في إحدى المدارس الأساسية

الفئة	-100	-110	-120	-130	150-140
التكرار	1	2	7	20	15

أوجد:

- الوسط الحسابي.
- الوسيط.
- المنوال.
- الربيعين الأدنى والأعلى.
- المئين 60.

3- إذا كان عدد الطلاب المتقدمين لامتحان مادة مبادئ الإحصاء 20 طالباً، وعدد الطالبات المتقدمات لنفس الامتحان 30 طالبة، فإذا علمت أن الوسط الحسابي لدرجات الطلاب 70 والوسط الحسابي لدرجات الطالبات 75 فاحسب الوسط الحسابي لجميع المتقدمين؟

4- إذا كان الوسط الحسابي لخمسين طالباً هو 70 فراجع المعلم طالبان فزادت علامة الأول بمقدار 10 ونقصت علامة الآخر بمقدار 5، احسب الوسط الحسابي لجميع الطلبة بعد عملية المراجعة.

5- إذا كان الوسط الحسابي لتوزيع أحادي المنوال هو 50 وكان المنوال هو 40 فأوجد الوسيط.

6- إذا كانت مبيعات أحد المتاجر لشهر نيسان لعام 1998 بمئات الدنانير هي:

4	7	5	9	12	17
4	16	11	18	7	19
12	10	14	20	8	13
4	20	13	14	11	6
9	8	19	18	13	15

• بوب البيانات في توزيع تكراري ذي 5 فئات ومنه أوجد الوسط الحسابي والوسيط والمنوال.

• أوجد المئين 80 لمبيعات هذا المتجر.

7- أوجد قيمة k إذا كانت انحرافات خمس قيم عن وسطها الحسابي هي:

$$k, 8 - 2k, 2 - k, 4k, 3k$$

أوجد أقل قيمة يمكن أن يأخذها مجموع مربعات الانحرافات.

8- مستشفى فيه ثلاثون ممرضاً وعشرة أطباء، إذا كان معدل رواتب الممرضين 200 ديناراً ومعدل رواتب الأطباء 700 ديناراً وقررت إدارة المستشفى زيادة رواتب الأطباء عشرين ديناراً وزيادة رواتب الممرضين 10% من رواتبهم احسب:

- معدل رواتب كل من الأطباء والممرضين على حده بعد الزيادة.
- معدل رواتب الأطباء والممرضين معاً قبل وبعد الزيادة.

9- للتوزيع التكراري التالي:

الفئة	-10	-20	-30	-40	60-50
التكرار	7	13	30	19	1

أوجد الربيعات الثلاثة واعرض التوزيع بطريقة المدرج التكراري ومنه صف التواء التوزيع.

10- إذا كان الوسط الحسابي للملاحظات: 0، 4، a، 3، -2 يساوي 5 فما قيمة a؟

11- قام باحث بحساب أوزان 10 طلبة بالكيلوجرام فكانت كما يلي:

55 60 78 85 89 69 70 64 84 90

احسب ما يلي:

- الوسط الحسابي والوسيط لأوزان الطلبة.
- إذا أضاف الباحث أوزان طالبين للمجموعة وهي 77، 68 كيلوجرام، احسب الوسط الحسابي والوسيط لأوزان الطلبة بعد الإضافة.
- إذا اكتشف الباحث خطأ في رصد وزن أحد الطلبة وقيده 90 بدلاً من 93، احسب الوسط الحسابي والوسيط بعد تعديل الوزن.
- إذا عدل الباحث جميع الأوزان بهدف الدراسة مستخدماً التحويل الخطي التالي: $Y = 2X - 100$ حيث X هي الأوزان قبل التعديل، احسب الوسط الحسابي والوسيط للأوزان بعد التعديل؟

12- اختر الإجابة الصحيحة فيما يلي:

• الوسيط هو المئين:

أ) 30 ب) 40 ج) 50 د) 60

• إذا كان الوسط لعشرة مشاهدات هو 7 وتم تعديل هذه المشاهدات حسب العلاقة

$$Y = 2X - 3$$

فإن الوسط الجديد بعد التعديل يساوي:

أ) 11 ب) 12 ج) 13 د) 14

• إذا كان $Me = 22$, $Mo = 25$, $\bar{X} = 20$ فإن التوزيع يكون:

أ) موجب الالتواء ب) سالب الالتواء ج) التوزيع متماثل د) منفصل

• إذا كان مجموع عدة مشاهدات 17 ووسطها الحسابي 3.4 فإن عدد المشاهدات

يساوي:

أ) 3 ب) 5 ج) 7 د) 9

• عدد المنوالات للملاحظات 0،4،9،8،1،2،3 يساوي:

أ) 1 ب) 2 ج) 3 د) لا يوجد منوال

• في توزيع تكراري إذا كان: $\sum f_i = 40$ فإن ترتيب الربيع الثالث يساوي:

أ) 10 ب) 20 ج) 30 د) 40

• إذا كان المنوال لعدة مشاهدات هو 7 وتم ضرب المشاهدات جميعاً في 3 فإن المنوال

للملاحظات بعد التعديل يساوي:

أ) 3.5 ب) 4 ج) 10 د) 21

تعتبر مقاييس النزعة المركزية من المقاييس الهامة في التحليل الإحصائي، وتعلمنا خلال الفصل الثالث كيفية حسابها للبيانات الأولية وللتوزيعات التكرارية.

الوسط الحسابي:

الوسط الحسابي يعتبر من أهم مقاييس النزعة المركزية وأكثرها استخداماً ويحسب من العلاقات التالية:

للملاحظات الأولية: $\bar{X} = \frac{\sum x_i}{n}$ وللتوزيعات التكرارية: $\bar{X} = \frac{\sum x_i f_i}{n}$. ويحسب الوسط الحسابي المشذب بنفس الطريقة مع استبعاد 5% من المشاهدات من الطرفين بعد ترتيبهما تصاعدياً أو تنازلياً.

الوسط الحسابي المرجح يحسب من العلاقة: $\bar{X} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2}{n_1 + n_2}$ ويمكن تعميمه لأكثر من مجموعتين.

الوسيط:

يحسب الوسيط للبيانات الأولية بتحديد ترتيبه من العلاقة: $O_{Me} = \frac{(n+1)}{2}$ ويتم حسابه بطريقة الميزان. وفي حالة التوزيعات التكرارية يتم حساب التكرار التراكمي للوسيط من العلاقة:

$CF = \frac{n}{2}$ وحساب الوسيط من القانون: $Me = a + \frac{CF - n_1}{f_m} \times c$ وذلك بعد حساب الجدول التراكمي الصاعد.

المنوال:

يعتبر من أضعف مقاييس النزعة المركزية وهو المشاهدة الأكثر شيوعاً في المشاهدات ويحسب بطرق تقريبية منها طريقة مركز الفئة المنوالية وطريقة الفروق حيث القانون المستخدم:

$$Mo = a + \frac{f_m - f_b}{(f_m - f_b) + (f_m - f_a)} \times c$$

وكل مقياس من مقاييس النزعة المركزية له مزايا وعيوب توضح خصائصه.

المئينات والربيعات:

تحسب المئينات للمشاهدات الأولية بحساب ترتيب المئين من العلاقة:

$$O_{P_k} = (n + 1) \times \frac{k}{100}$$

ثم بتطبيق طريقة الميزان كما في الوسيط تماماً. وفي حالة التوزيعات التكرارية يتم حساب التكرار التراكمي للمئين من العلاقة: $CF(P_k) = n \times \frac{k}{100}$ ويحسب المئين من القانون:

$$P_k = a + \frac{CF - n_1}{f_m} \times c$$

من الخصائص الهامة للتوزيعات والتي نحرص على معرفتها تماثل التوزيع أو التواءه، ويمكن بحث الالتواء بطرق كثيرة منها رسم المدرج التكراري للتوزيع ومقارنة الوسط والوسيط والمنوال له.

مقاييس التشتت Measures of Dispersion

4

إن حساب مقاييس النزعة المركزية للتعرف على خصائص وصفات المجتمع بعد عملية جمع البيانات وتنظيمها لا يكفي، حيث أن هذه المقاييس تصف جانباً معيناً والذي له علاقة بالمتوسطات ولا تصف تشتت المشاهدات عن بعضها البعض أو عن وسطها الحسابي. فقد يتساوى الوسط الحسابي لمجتمعين ويكون أحدهما أكثر تجانساً من الآخر. في هذا الفصل سنعرض بعض مقاييس التشتت للمشاهدات الأولية والتوزيعات التكرارية ونبين كيف تتأثر بالتحويلات الخطية.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً

على:

- حساب المدى، الانحراف الربيعي، الانحراف المتوسط، التباين والانحراف المعياري للمشاهدات الأولية وللتوزيعات التكرارية.
- حساب التباين المرجح.
- حساب معامل الاختلاف.
- معرفة تأثير التحويلات الخطية على مقاييس التشتت وكيفية حساب المقاييس للمشاهدات الجديدة.

تعتبر المتوسطات كالوسط الحسابي والوسيط من أهم المقاييس الإحصائية الوصفية التي نحتاجها لوصف البيانات وصفاً كمياً، إلا أن إيجاد أحد المتوسطات لا يعطي وصفاً كاملاً للبيانات، لأنه على سبيل المثال يبين القيمة التي تتمركز حولها البيانات كالوسط الحسابي أو القيمة التي تتوسط القيم كالوسيط دون أن يعطي أي فكرة عن مدى تقارب أو تباعد البيانات عن بعضها البعض أو عن قيمة المتوسط ذاتها، أي أن قيمة المتوسط لا تبين مدى تجانس البيانات أو تشتتها.

لذلك نحتاج إلى مقاييس أخرى تبين مدى تشتت البيانات، وهذه المقاييس تسمى **مقاييس التشتت**. وتصنف لمقاييس التشتت المطلق (وسنذكر منها المدى، الانحراف المتوسط، التباين والانحراف المعياري والتي لها نفس وحدات قياس المشاهدات عدا التباين) ومقاييس التشتت النسبي (وسنذكر منه معامل الاختلاف والذي لا يوجد له وحدات قياس).

سنحدث في بداية هذا الفصل عن حساب هذه المقاييس للبيانات الأولية ومن ثم للتوزيعات التكرارية.

أولاً: مقاييس التشتت للبيانات الأولية

1- المدى The Range

يعتبر المدى من أبسط مقاييس التشتت وأسرعها حساباً ويُعرّف بأنه الفرق بين أكبر مشاهدة وأصغر مشاهدة، وسنرمز له بالرمز R ، ولو رمزنا لأكبر مشاهدة بالرمز X_{max} ولأصغر مشاهدة بالرمز X_{min} فإن المدى يحسب من العلاقة:

$$R = X_{max} - X_{min}$$

المدى للملاحظات الأولية



يقيس هذا المدى تشتت المشاهدات عن بعضها البعض ولا علاقة له بالوسط الحسابي وعند استخدامه لمقارنة تشتت مجتمعين يشترط سحب عينتين بنفس الحجم من كل مجتمع.

مثال 1.4: احسب المدى للمشاهدات: 12 5 7 9 8

/الحل/

لاحظ أن: $X_{max} = 12$ ، $X_{min} = 5$ وبالتالي تكون قيمة المدى:

$$R = X_{max} - X_{min}$$

$$R = 12 - 5 = 7$$

وهذا يعني أن جميع المشاهدات تقع في مسافة طولها 7 والفرق بين أي مشاهدين يكون أقل من أو يساوي 7 وواضح أنه كلما صغرت قيمة المدى فإن الفروق بين المشاهدات تقل، أي يقل تشتت المشاهدات عن بعضها البعض وتكون أكثر تجانساً.

مثال 2.4: مستخدماً المدى حدد أي المجموعتين التاليتين أقل تشتتاً:

المجموعة الأولى: 17 13 20 14 15 18

المجموعة الثانية: 21 23 19 20 22 20

/الحل/ نحسب المدى لكل مجموعة كما يلي:

$$R_1 = X_{1max} - X_{1min}$$

$$R_1 = 20 - 14 = 6$$

$$R_2 = X_{2max} - X_{2min}$$

$$R_2 = 23 - 19 = 4$$

وحيث أن مدى المجموعة الثانية أصغر من مدى المجموعة الأولى فإننا نستطيع القول أن المجموعة الثانية أقل تشتتاً من المجموعة الأولى.

مثال 3.4: احسب المدى للملاحظات: 1 2 3 4 5 500

الحل/

$$R = X_{max} - X_{min}$$

$$R = 500 - 1 = 499$$

وهذا يعني أن المشاهدات تتوزع في مسافة مقدارها 499، ونلاحظ أن المشاهدات عددها صغير والخمسة مشاهدات الأولى منها متقاربة من بعضها البعض والمدى يخبرنا من ناحية أخرى أن المشاهدات مشتتة بشكل كبير جداً. إن السبب في جعل المدى يفسر المشاهدات تفسيراً غير دقيق هو وجود مشاهدة متطرفة وهي المشاهدة التي قيمتها 500 وهذا يعني تأثر المدى بالمشاهدات الشاذة والمتطرفة مما يجعله مقياساً غير دقيق، وللتخلص من هذه المشكلة ممكن حساب ما يُعرَف بالانحراف الربيعي أو حذف نسبة معينة من طرفي المشاهدات بعد ترتيبها تصاعدياً.

2- الانحراف الربيعي (SIR) Simi-Interquartile Range

تحدثنا في الفصل الثالث عن الربيعات ووضحنا كيفية حسابها للمشاهدات الأولية وللتوزيعات التكرارية، ونعرف هنا الانحراف الربيعي ويرمز له بالرمز SIR بأنه (نصف الفرق بين الربيع الثالث والربيع الأول) أي أن:

$$SIR = \frac{Q_3 - Q_1}{2}$$

الانحراف الربيعي



مثال 4.4: تم رصد درجة الحرارة في مدينة رفح لعشرة أيام متتالية في شهر تموز وكانت النتائج كما يلي:

30 35 35 30 32 33 28 27 31 33

أحسب الانحراف الربيعي لدرجات الحرارة.

الحل/

تم حساب الربيعين الأول والثالث لهذه المشاهدات في مثال (9.3) وكانت النتائج كما يلي:

$$33.5 = Q_3, 29.5 = Q_1$$

وبالتالي يكون:

$$SIR = \frac{Q_3 - Q_1}{2} = \frac{33.5 - 29.5}{2} = \frac{4}{2} = 2$$

3- الانحراف المتوسط Mean Deviation

إذا كان لدينا المشاهدات: $x_1, x_2, \dots, x_n$ وكان وسطها الحسابي $\bar{X}$ ، فإننا نعرّف الانحراف المتوسط لهذه المشاهدات بأنه الوسط الحسابي لانحراف المشاهدات عن وسطها الحسابي و يرمز له بالرمز $M.D.$ ويحسب من العلاقة:

الانحراف المتوسط للمشاهدات الأولية

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



ويعتبر الانحراف المتوسط مقدار تشتت المشاهدات عن وسطها الحسابي، والمثال التالي يوضح كيفية حساب الانحراف المتوسط للبيانات الأولية:

مثال 5.4: أوجد الانحراف المتوسط للملاحظات: 2 5 7 4 2

الحل/ نحسب أولاً الوسط الحسابي للملاحظات:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{2 + 4 + 7 + 5 + 2}{5} = \frac{20}{5} = 4$$

فيكون الانحراف المتوسط مساوياً:

$$\begin{aligned} M.D &= \frac{\sum_{i=1}^n |x_i - \bar{X}|}{n} = \frac{\sum_{i=1}^5 |x_i - 4|}{5} \\ &= \frac{|2 - 4| + |4 - 4| + |7 - 4| + |5 - 4| + |2 - 4|}{5} = \frac{8}{5} = 1.6 \end{aligned}$$

وهذا يعني أن متوسط انحراف المشاهدات عن وسطها الحسابي يساوي 1.6 وحدة. وكلما كان الانحراف المتوسط صغيراً كلما كانت المشاهدات أقرب لوسطها الحسابي؛ بمعنى أن المشاهدات أكثر تجانساً وبالتالي أقل تشتتاً.

4-التباين (S^2) The Variance

يُعتبر التباين من أهم مقاييس التشتت وأكثرها استخداماً ويعرف بأنه الوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي، وقد ذكرنا في الفصل الثالث أن مجموع مربعات هذه الانحرافات هو أقل ما يمكن، فإذا كان لدينا المشاهدات: $x_1, x_2, \dots, x_n$ والتي وسطها الحسابي $\bar{X}$ ، فإن التباين في هذه الحالة يحسب من العلاقة:

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

التباين للملاحظات الأولية



ووحدة قياس التباين هي مربع وحدة قياس المشاهدات، فمثلاً إذا كانت وحدة قياس المشاهدات هي (متر) فإن وحدة قياس التباين هي (متر)².

ويوجد صورة أخرى لحساب التباين في حالة المشاهدات الأولية هي:

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

5- الانحراف المعياري (S) The Standard Deviation

يُعرّف الانحراف المعياري بأنه الجذر التربيعي الموجب للتباين أي أن:

$$S = \sqrt{s^2}$$

الانحراف المعياري



ووحدة قياس الانحراف المعياري هي نفسها وحدة قياس المشاهدات.

ويعتبر الانحراف المعياري من أكثر مقاييس التشتت استخداماً حيث أنه يدخل في تعريف توزيعات احتمالية هامة كالتوزيع الطبيعي ويدخل في حساب الكثير من المقاييس الإحصائية واختبارات الفرضيات.

مثال 6.4: أوجد التباين والانحراف المعياري للمشاهدات: 1 2 3 4 5

الحل/

نحسب أولاً الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1 + 2 + 3 + 4 + 5}{5} = 3$$

فيكون التباين:

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^n x_i^2 - n\bar{X}^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - 5 \times 3^2}{5-1} \\ &= \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 - 5 \times 9}{4} \\ &= \frac{1 + 4 + 9 + 16 + 25 - 45}{4} = \frac{10}{4} = 2.5 \end{aligned}$$

وبالتالي يكون الانحراف المعياري:

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

خواص ومزايا التباين والانحراف المعياري

- 1- يعتبر التباين والانحراف المعياري من أهم مقاييس التشتت.
- 2- يدخل في حسابهما جميع القيم.
- 3- يعتبران أدق مقاييس التشتت المعروفة.
- 4- لا يتأثر كلا من التباين والانحراف المعياري بإضافة عدد حقيقي للملاحظات.

ففي المثال (6.4) لو أضفنا للملاحظات العدد 2 فإننا سنحصل على المشاهدات الجديدة وهي: $Y: 3, 4, 5, 6, 7$ ويكون الوسط الحسابي الجديد:

$$\bar{Y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{3 + 4 + 5 + 6 + 7}{5} = \frac{25}{5} = 5$$

ويكون التباين الجديد:

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^n y_i^2 - n\bar{Y}^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - 5 \times 5^2}{5-1} \\ &= \frac{3^2 + 4^2 + 5^2 + 6^2 + 7^2 - 5 \times 25}{4} \\ &= \frac{9 + 16 + 25 + 36 + 49 - 125}{4} = \frac{10}{4} = 2.5 \end{aligned}$$

وبالتالي يكون الانحراف المعياري الجديد:

$$S_Y = \sqrt{S_Y^2} = \sqrt{2.5} = 1.58$$

ونلاحظ أن قيم التباين والانحراف المعياري لم تتغير وهذا يدل على عدم تأثرهما بالإضافة.

5- يتأثر كل من التباين والانحراف المعياري بضرب المشاهدات في عدد حقيقي.
ففي المثال (6.4) لو ضربنا المشاهدات بالعدد 2 فإننا سنحصل على المشاهدات الجديدة وهي: $Y : 2, 4, 6, 8, 10$ ويكون الوسط الحسابي الجديد:

$$\bar{Y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{2 + 4 + 6 + 8 + 10}{5} = \frac{30}{5} = 6$$

ويكون التباين الجديد:

$$\begin{aligned} S_Y^2 &= \frac{\sum_{i=1}^n y_i^2 - n\bar{Y}^2}{n-1} = \frac{\sum_{i=1}^n y_i^2 - 5 \times 6^2}{5-1} \\ &= \frac{2^2 + 4^2 + 6^2 + 8^2 + 10^2 - 5 \times 36}{4} \\ &= \frac{4 + 16 + 36 + 64 + 100 - 180}{4} = \frac{40}{4} = 10 \end{aligned}$$

وبالتالي يكون الانحراف المعياري الجديد:

$$S_Y = \sqrt{S_Y^2} = \sqrt{10} = 3.16$$

ونلاحظ هنا تغير قيمة كلاً من التباين والانحراف المعياري، وسنذكر في آخر هذا الفصل تأثير التحويلات الخطية في بعض مقاييس التشتت. ويعاب على التباين والانحراف المعياري تأثرهما بالقيم الشاذة والمتطرفة.

6- التباين المرجح The Weighted Variance

بفرض أنه لدينا مجموعتين من المشاهدات، الأولى تباينها S_1^2 وعددها n_1 والثانية تباينها S_2^2 وعددها n_2 ، فإذا دمجت المجموعتان معاً في مجموعة واحدة فإن التباين لها - والذي يعرف بالتباين المرجح أو التباين الموزون - يعطى بالعلاقة:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

التباين المرجح



ويمكن التعميم لأكثر من مجموعتين.

مثال 7.4: أُخِذَت عِينَتَانِ مِنْ مَجْتَمَعَيْنِ فَكَانَتِ النَتَائِجُ كَمَا هُوَ مَبِينٌ فِي الْجَدُولِ التَّالِي:

المجموعة الثانية	المجموعة الأولى
$\sum_{i=1}^{40} y_i = 280, \sum_{i=1}^{40} y_i^2 = 2100$	$\sum_{i=1}^{50} x_i = 300, \sum_{i=1}^{50} x_i^2 = 1950$

1- أوجد تباين كل مجموعة.

2- إذا دُمِجَتِ المجموعتان في مجموعة واحدة، احسب تباين المجموعة الجديدة.

1- نحسب التباين لكل مجموعة كما يلي:

المجموعة الأولى: $n_1 = 50$

$$\bar{X} = \frac{\sum_{i=1}^{50} x_i}{n_1} = \frac{300}{50} = 6$$

$$S_X^2 = \frac{\sum_{i=1}^{50} x_i^2 - n_1 \bar{X}^2}{n_1 - 1} = \frac{1950 - 50 \times 6^2}{50 - 1} = 3.06$$

المجموعة الثانية: $n_2 = 40$

$$\bar{Y} = \frac{\sum_{i=1}^{40} y_i}{n_2} = \frac{280}{40} = 7$$

$$S_Y^2 = \frac{\sum_{i=1}^{40} y_i^2 - n_2 \bar{Y}^2}{n_2 - 1} = \frac{2100 - 40 \times 7^2}{40 - 1} = 3.59$$

2- عند دمج المجموعتين معاً نحصل على مجموعة جديدة تباينها هو التباين المرجح ويساوي:

$$S_p^2 = \frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2}$$

$$= \frac{(50 - 1) \times 3.06 + (40 - 1) \times 3.59}{50 + 40 - 2} = 3.29$$

ثانياً: مقاييس التشتت للتوزيعات التكرارية

1- المدى

يحسب المدى للتوزيعات التكرارية بنفس العلاقة المستخدمة لحسابه في حالة المشاهدات

الأولية، إلا أننا نعتبر في هذه الحالة ما يلي:

X_{max} : الحد الأعلى لآخر فئة في التوزيع التكراري.

X_{min} : الحد الأدنى لأول فئة في التوزيع التكراري.

مثال 8.4: أوجد المدى للتوزيع التكراري التالي:

30-25	-20	-15	-10	-5	الفئة
4	6	5	2	3	التكرار

الحل/

لاحظ أن:

$$X_{max} = 30 , \quad X_{min} = 5$$

وبالتالي يكون:

$$R = X_{max} - X_{min} = 30 - 5 = 25$$

2- الانحراف الربيعي

تعلمنا في الفصل الثالث كيفية حساب المئينات والربيعات للتوزيعات التكرارية،

والمثال التالي يوضح كيفية حساب الانحراف الربيعي لها.

مثال 9.4: أوجد الانحراف الربيعي للتوزيع التكراري التالي:

32-28	-24	-20	-16	-12	-8	الفئات
5	4	8	17	10	6	التكرار

الحل/

في مثال (13.3) تم حساب الربع الثالث للتوزيع وكانت قيمته تساوي: $Q_3 = 22.25$

حساب الربع الأول $Q_1 = P_{25}$:

التكرار التراكمي للربع الأول يساوي:

$$CF(P_{25}) = n \times \frac{k}{100} = 50 \times \frac{25}{100} = 12.5$$

من الجدول التراكمي الصاعد التالي نجد أن أول فئة تكرارها التراكمي أكبر من 12.5 هي الفئة الثانية لذلك تكون هي الفئة المئينية المطلوبة.

الفئات	التكرار	أقل من الحدود العليا للفئات	التكرار التراكمي الصاعد
-8	6	أقل من 12	6
-12	10	أقل من 16	16
-16	17	أقل من 20	33
-20	8	أقل من 24	41
-24	4	أقل من 28	45
32-28	5	أقل من 32	50

من الجدول التراكمي الصاعد نجد أن:

$$12 = a \quad \text{وهي الحد الأدنى للفئة المئينية.}$$

$$6 = n_1 \quad \text{وهي التكرار التراكمي للفئة قبل المئينية.}$$

$$10 = f_m \quad \text{وهي تكرار الفئة المئينية.}$$

$$4 = c \quad \text{وهي طول الفئة المئينية.}$$

وبالتالي:

$$Q_1 = a + \frac{CF - n_1}{f_m} \times c = 12 + \frac{12.5 - 6}{10} \times 4 = 14.6$$

أو نكون التناسب التالي كما تعلمنا في الفصل الثالث:

$$10 \left[\begin{array}{c} 6 \\ 12.5 \\ 16 \end{array} \right] 6.5 \qquad \underline{\underline{12}} \left[\begin{array}{c} Q_1 \\ 16 \end{array} \right] 4$$

ومن التناسب يكون الربيع الأول يساوي:

$$Q_1 = 12 + \frac{6.5}{10} \times 4 = 14.6$$

وبالتالي يكون الانحراف الربيعي:

$$SIR = \frac{Q_3 - Q_1}{2} = \frac{22.25 - 14.6}{2} = 3.83$$

3- الانحراف المتوسط

إذا كان لدينا توزيع تكراري مكون من k فئة ومراكز فئاته على الترتيب هي كما يلي:

$$x_k, \dots, x_2, x_1$$

وكانت تكرارات الفئات على الترتيب هي:

$$f_k, \dots, f_2, f_1$$

وكان مجموع التكرارات:

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

وكان الوسط الحسابي للتوزيع التكراري $\bar{X}$ فإن الانحراف المتوسط يحسب من العلاقة:

الانحراف المتوسط للتوزيعات لتكرارية

$$M.D. = \frac{\sum |x_i - \bar{X}| f_i}{n}$$



مثال 10.4: أوجد الانحراف المتوسط للتوزيع التكراري في مثال (7.4)

الحل/ نكون جدول الحساب التالي:

جدول الحساب

الفئات	f_i	x_i	$x_i f_i$	$ x_i - 19 $	$ x_i - 19 f_i$
-5	3	7.5	22.5	11.5	34.5
-10	2	12.5	25	6.5	13
-15	5	17.5	87.5	1.5	7.5
-20	6	22.5	135	3.5	21
30-25	4	27.5	110	8.5	34
المجموع	20		380		110

من الجدول يكون الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^n x_i f_i}{n} = \frac{380}{20} = 19$$

وبالتالي يكون الانحراف المتوسط:

$$M.D. = \frac{\sum |x_i - \bar{X}| f_i}{n} = \frac{110}{20} = 5.5$$

4-التباين

إذا كان لدينا توزيع تكراري مكون من k فئة ومراكز فئاته على الترتيب كما يلي:

$$x_k, \dots, x_2, x_1$$

وكانت تكرارات الفئات على الترتيب هي:

$$f_k, \dots, f_2, f_1$$

وكان مجموع التكرارات:

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

وكان الوسط الحسابي للتوزيع التكراري $\bar{X}$ فإن تباين التوزيع التكراري يحسب من العلاقة:

$$s^2 = \frac{\sum_{i=1}^k x_i^2 f_i - n(\bar{X})^2}{n - 1}$$



التباين للتوزيعات لتكرارية

مثال 11.4: أوجد التباين والانحراف المعياري للتوزيع التكراري التالي:

35-30	-25	-20	-15	-10	الفئات
4	7	6	2	1	التكرار

الحل/

نكون جدول الحساب التالي:

$x_i^2 f_i$	$x_i f_i$	x_i	f_i	الفئات
156.25	12.5	12.5	1	-10
612.5	35	17.5	2	-15
3037.5	135	22.5	6	-20
5293.75	192.5	27.5	7	-25
4225	130	32.5	4	35 - 30
13325	505		20	المجموع

نحسب أولاً الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^n x_i f_i}{n} = \frac{505}{20} = 25.25$$

فيكون التباين:

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^k x_i^2 f_i - n(\bar{X})^2}{n - 1} \\ &= \frac{13325 - 20 \times (25.25)^2}{19} = 30.2 \end{aligned}$$

ويكون الانحراف المعياري:

$$S = \sqrt{S^2} = \sqrt{30.2} = 5.5$$

5- معامل الاختلاف The Coefficient of Variation

يعتبر معامل الاختلاف مقياساً من مقاييس التشتت النسبي ولا يوجد له وحدة قياس، ويستخدم لمقارنة تشتت مجتمعين مختلفين. مثل مقارنة تشتت مجتمع أطوال مع مجتمع أوزان وهكذا.

إذا كان $\bar{X}$, S_X يمثلان الوسط الحسابي والانحراف المعياري لمجموعة من المشاهدات على الترتيب، فإن معامل الاختلاف لهذه المجموعة يعطى بالعلاقة:

$$CV = \frac{S_X}{\bar{X}} \times 100\%$$

معامل الاختلاف



مثال 12.4: احسب معامل الاختلاف للتوزيع التكراري في مثال (11.4).

الحل/ من المثال (11.4) نجد أن: $\bar{X} = 25.25$ ، $S_X = 5.5$ وبالتالي يكون معامل الاختلاف للتوزيع التكراري يساوي:

$$CV = \frac{S_X}{\bar{X}} \times 100\% = \frac{5.5}{25.25} \times 100\% = 21.78\%$$

مثال 13.4: مصنع لإنتاج المصابيح الكهربائية ينتج نوعين من المصابيح، النوع الأول A ومتوسط أعمارها 1400 ساعة بانحراف معياري 238 ساعة، والنوع الثاني B ومتوسط أعمارها 1800 ساعة بانحراف معياري 333 ساعة. حدد أي النوعين الذي له أقل تشتت نسبي.

الحل/

النوع الأول A :

من السؤال نجد أن: $\bar{X}_A = 1400$, $S_{X_A} = 238$ وبالتالي يكون معامل الاختلاف للنوع الأول من المصابيح يساوي:

$$CV_A = \frac{S_{X_A}}{\bar{X}_A} \times 100\% = \frac{238}{1400} \times 100\% = 17\%$$

النوع الثاني B :

من السؤال نجد أن: $\bar{X}_B = 1800$, $S_{X_B} = 333$ وبالتالي يكون معامل الاختلاف للنوع الثاني من المصابيح يساوي:

$$CV_B = \frac{S_{X_B}}{\bar{X}_B} \times 100\% = \frac{333}{1800} \times 100\% = 18.5\%$$

وبمقارنة المعاملين نجد أن: $CV_A < CV_B$ وهذا يدل على أن النوع الأول أقل تشتتاً من النوع الثاني.

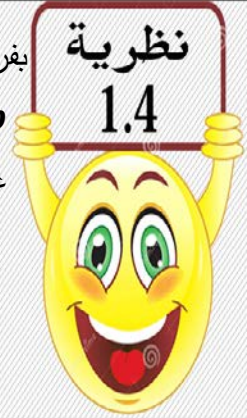
التحويلات الخطية ومقاييس التشتت:

عرفنا في الفصل الثالث التحويل الخطي ووضحنا تأثيره في مقاييس النزعة المركزية، وذكرنا من ضمن خواص ومزايا التباين والانحراف المعياري في هذا الفصل أنهما لا يتأثران بالإضافة، والسؤال الذي يطرح نفسه هنا هل تتأثر مقاييس التشتت للملاحظات القديمة بالتحويل الخطي سابق الذكر عامة؟ بصورة أخرى نتساءل: هل مقاييس التشتت للملاحظات الجديدة تختلف عنها للملاحظات القديمة؟

الإجابة نعم. والنظرية التالية توضح العلاقة بين مقاييس التشتت للملاحظات الجديدة ومقاييس التشتت للملاحظات القديمة.

بفرض أنه لدينا الملاحظات $x_1, x_2, \dots, x_n$ و أثنا عليها بالتحويل الخطي $Y = aX + b$ لنحصل على الملاحظات الجديدة $y_1, y_2, \dots, y_n$ على الترتيب فإن:

$$\begin{aligned} s_Y^2 &= a^2 \cdot s_X^2 \\ s_Y &= |a| \cdot s_X \\ R_Y &= |a| \cdot R_X \end{aligned}$$



حيث s_X^2 ، s_X ، R_X . تمثل التباين، الانحراف المعياري والمدى للملاحظات القديمة $x_1, x_2, \dots, x_n$ على الترتيب و s_Y^2 ، s_Y ، R_Y تمثل التباين، الانحراف المعياري، والمدى للملاحظات الجديدة $y_1, y_2, \dots, y_n$ على الترتيب.

ملاحظة: نلاحظ من النظرية السابقة أن مقاييس التشتت تتأثر بضرب المشاهدات في عدد حقيقي ولا تتأثر بإضافة عدد حقيقي للمشاهدات.

مثال 14.4: تم حساب التباين والمدى لمجموعة من المشاهدات فكان:

التباين يساوي 16 ، المدى يساوي 25 ، وتم تعديل المشاهدات حسب العلاقة: $Y = 3 - 2X$ حيث X تمثل القيمة القديمة ، Y تمثل القيمة الجديدة بعد التعديل، فاحسب التباين، الانحراف المعياري و المدى بعد التعديل؟

الحل/

معلوم لدينا في السؤال ما يلي:

$$R_X = 25 \quad , \quad S_X^2 = 16$$

وبالتطبيق المباشر لنظرية (1.4) نحصل على المقاييس للمشاهدات الجديدة كما يلي:

$$S_Y^2 = (-2)^2 S_X^2 = 4 \times 16 = 64$$

$$S_Y = \sqrt{64} = 8$$

$$R_Y = |a|R_X = |-2|.25 = 50$$

تمارين:

1- للملاحظات التالية احسب المدى، الانحراف الربيعي، الانحراف المتوسط، التباين والانحراف المعياري:

3 10 3 9 5 1 4 -1 0 6

2- للتوزيع التكراري التالي:

50-40	-30	-20	-10	-0	الفئات
6	10	8	9	7	التكرار

احسب: المدى، الانحراف الربيعي، الانحراف المتوسط، التباين والانحراف المعياري.

3- إذا كان لدينا مجموعة من الملاحظات المفردة، انحرافها المعياري 5 و عدلت جميع الملاحظات المفردة بالعلاقة $Y = 6 - 2X$ فأوجد:

- الانحراف المعياري بعد التعديل.
- التباين قبل التعديل.
- التباين بعد التعديل.

4- إذا كان متوسط أطوال 100 طالب يساوي 65 بوصة، وكان معامل الاختلاف للأطوال يساوي 6.6%، أوجد تباين الأطول.

5- إذا كانت انحرافات 6 ملاحظات عن وسطها الحسابي هي:

-4 5 -2 7 -6 0

أوجد الانحراف المتوسط والانحراف المعياري لهذه الملاحظات.

6- رُصِدَت درجات 9 طلاب في مساق الإحصاء التطبيقي فوجد أن الوسط الحسابي للدرجات يساوي 20 درجة، ثم أضيفت درجة طالب جديد للدرجات فوجد أن الوسط الحسابي قد زاد بمقدار 0.8 أوجد درجة الطالب الجديد.

7- الجدول التالي يبين توزيع وفيات الأطفال دون سن السنة (بالألف) في إحدى البلاد حسب أعمارهم بالشهر:

الفئات بالشهر	-0	-2	-4	-6	-8	10-12
عدد الوفيات	27	22	20	23	19	9

- احسب متوسط الأعمار.
- احسب تباين الأعمار.
- احسب معامل الاختلاف.

8- إذا كان الانحراف المعياري والوسط الحسابي لرواتب مجموعة من الموظفين في شركة لصناعة الأدوية بالدينار 12، 180 على الترتيب. فإذا قرر مدير الشركة زيادة الرواتب بمقدار عشرة دنانير لكل موظف. احسب الانحراف المعياري والوسط الحسابي بعد التعديل؟

9- إذا كان لدينا توزيع تكراري ذو 9 فئات لها نفس الطول، وكان الحدان الأدنى والأعلى للفئة الأولى هما 10، 20 على الترتيب، وكان التكرار التراكمي الصاعد للفئتين الخامسة والسادسة يساوي 45، 78 على الترتيب، وكان: $\sum f_i = 101$. احسب قيمة الوسيط.

- 10- الدخل الأسبوعي لعينة من 101 شخصاً في مدينة A له وسط حسابي 100 دينار، وسيط 110 دنانير وانحراف معياري 30 دينار. وفي مدينة B يمثل التوزيع التكراري التالي الدخل الأسبوعي لمجموعة من الأشخاص:

100-90	-80	-70	-60	-50	-40	الدخل الأسبوعي
6	20	30	10	9	7	عدد الأشخاص

بين في أي مدينة يعتبر الدخل أكثر تغيراً.

- 11- إذا علمت أن $\bar{Z} = 45$ ، $S_Z^2 = 20$ ، $Z = 2Y + 15$ أوجد الوسط الحسابي والتباين والانحراف المعياري للمتغير Y .

- 12- إذا كان لدينا المجموعتين التاليتين:

$$X: 2, 8, 10, 12 \quad , \quad Y: 6, 7, 5, 11, 17$$

- احسب الوسط الحسابي والانحراف المعياري لكل مجموعة.
- احسب معامل الاختلاف لكل مجموعة وقارن بينهما.
- إذا دُمجت المجموعتان معاً في مجموعة واحدة، احسب الوسط الحسابي والانحراف المعياري للمجموعة الجديدة.

- 13- إذا علمت أن $\bar{Y} = 12$ ، $S_Y^2 = 5$ أوجد الوسط الحسابي والتباين والانحراف المعياري للمتغير X إذا علمت أن:

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

14- الجدول التالي يمثل دخل 250 عاملاً يومياً بالجنيه:

110-100	-90	-80	-70	-60	-50	-40	الدخل اليومي
14	44	79	75	27	8	3	عدد العمال

- احسب الوسط الحسابي والوسيط للدخل اليومي.
- احسب الوسط الحسابي للدخل الأسبوعي.

15- إذا كان لدينا البيانات التالية لعينتين مستقلتين:

$$\sum_{i=1}^{10} x_i = \sum_{i=1}^{20} y_i = 100, \sum_{i=1}^{100} x_i^2 = 1225, \sum_{i=1}^{20} y_i^2 = 671$$

- احسب الوسط الحسابي والانحراف المعياري لكل عينة.
- أي العينتين أكبر تشتتاً.
- إذا دمجت العينتان معاً احسب الوسط الحسابي والانحراف المعياري للعينة الجديدة.

16- في عينة مؤلفة من 25 من الذكور، 15 من الإناث، تبين أن متوسط رواتب الذكور

2260 جنيهه بانحراف معياري 450 جنيهه، بينما متوسط رواتب الإناث 2400 جنيهه

بانحراف معياري 200 جنيهه.

- احسب متوسط رواتب الذكور والإناث معاً.
- احسب الانحراف المعياري لرواتب الإناث والذكور معاً.

- 17- بدراسة توزيعين عن ظاهرتين مختلفتين وجد أن الوسط الحسابي للظاهرة الأولى 75 بانحراف معياري 15، أما الظاهرة الثانية فتم تنظيم مشاهداتها في التوزيع التكراري التالي:

68-58	-48	-38	-28	-18	الفئات
31	79	105	59	26	التكرار

بين أي الظاهرتين أكثر تغيراً.

- 18- مؤسسة بها 125 موظفاً تدفع مكافأة سنوية قدرها 50 دينار لكل موظف من موظفي الدرجة الأولى وعددهم 10، وتدفع 30 دينار لكل موظف من موظفي الدرجة الثانية وعددهم 25، وتدفع 20 دينار لكل موظف من موظفي الدرجة الثالثة وعددهم 40، وتدفع 10 دنانير لكل موظف من باقي الموظفين. أوجد الوسط الحسابي الموزون للأجور التي تدفعها الشركة.

- 19- اختر الإجابة الصحيحة:

- في توزيع تكراري إذا كانت الفئة الأولى (8-) والفئة الأخيرة (32-38) فإن المدى يساوي:

أ) 14 ب) 30 ج) 35 د) 40

- الانحراف المتوسط للملاحظات k, k, k, k, k يساوي:

أ) 0 ب) 5 ج) k د) $5k$

- مجموعة من المشاهدات انحرافها المعياري 6 ومعامل الاختلاف 20% يكون وسطها الحسابي:

أ) 10 ب) 20 ج) 30 د) 120

- إذا كان الانحراف المعياري لعدة مشاهدات هو 1.5 وتم تعديل المشاهدات

حسب العلاقة: $Y = 3 - aX$ فكان الانحراف المعياري الجديد هو 6 فإن

قيمة a تساوي:

- أ) 4 ب) -4 ج) $4 \mp$ د) $9 \mp$

- مجموعة من المشاهدات مداها 17 وأصغر مشاهدة هي -3 فإن أكبر مشاهدة

تساوي:

- أ) 11 ب) 14 ج) 20 د) 10

- إذا كان تباين عدة مشاهدات هو 20 وتم تعديل المشاهدات حسب العلاقة:

$Y = X - 20$ فإن التباين الجديد يساوي:

- أ) 0 ب) 1 ج) $\sqrt{20}$ د) 20

تعتبر مقاييس التشتت من المقاييس الهامة في التحليل الإحصائي حيث يعتبر التباين والانحراف المعياري أكثرها دقة وأهمية وأكثرها استخداماً.

المدى:

في حالة المشاهدات الأولية هو الفرق بين أكبر مشاهدة وأصغر مشاهدة، أما في حالة التوزيعات التكرارية فهو الفرق بين أعلى حد في التوزيع وأصغر حد فيه:

$$R = X_{max} - X_{min}$$

الانحراف الربيعي:

$$SIR = \frac{Q_3 - Q_1}{2} \text{ هو نصف الفرق بين الربيعين الثالث والأول:}$$

الانحراف المتوسط:

هو الوسط الحسابي للفروق المطلقة بين المشاهدات ووسطها الحسابي

$$M.D = \frac{\sum_{i=1}^n |x_i - \bar{X}|}{n} \text{ والانحراف المتوسط للملاحظات الأولية يساوي:}$$

$$M.D = \frac{\sum_{i=1}^n |x_i - \bar{X}| f_i}{n} \text{ وللتوزيعات التكرارية يساوي:}$$

التباين:

هو متوسط مربعات الفروق بين المشاهدات ووسطها الحسابي وللملاحظات

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1} \text{ أو } S^2 = \frac{\sum_{i=1}^n x_i^2 - n\bar{X}^2}{n-1} \text{ الأولية يساوي:}$$

$$S^2 = \frac{\sum_{i=1}^n x_i^2 f_i - n\bar{X}^2}{n-1} \text{ وللتوزيعات التكرارية يساوي:}$$

الانحراف المعياري:

هو الجذر التربيعي الموجب للتباين: $S = \sqrt{S^2}$

التباين المرجح:

عند دمج مجموعتين أو أكثر فإن المجموعة الناتجة لها تباين مختلف عن تبايني المجموعتين الأساسيتين، ويطلق على هذا التباين اسم **التباين المرجح** ويساوي:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

معامل الاختلاف:

عند الرغبة في إجراء مقارنة بين تشنت مجتمعين لهما وحدات قياس مختلفة نلجأ لمقياس تشنت نسبي مثل معامل الاختلاف والذي يساوي:

$$CV = \frac{S_x}{\bar{X}} \times 100\%$$

الارتباط والانحدار الخطي البسيط Correlation and Simple Linear Regression

5

يعتبر مفهوم الارتباط من المفاهيم الهامة التي يلعب دوراً كبيراً في التحليل الإحصائي حيث من خلاله نبحث وجود علاقة بين المتغيرات ولا تخلو منه الكثير من التطبيقات العملية والدراسات البحثية في الكثير من المجالات. فوجود علاقة مناسبة بين المتغيرات قيد الدراسة يفتح المجال أمام الباحثين للتساؤل عن أفضل معادلة انحدار تربط بين المتغيرات وتستخدم للتنبؤ بقيمة المتغير التابع إذا علمت قيم المتغيرات المستقلة. في هذا الفصل سنتحدث عن الارتباط الخطي وكيفية حساب معامل ارتباط بيرسون ومعامل ارتباط سبيرمان وكيفية إيجاد معادلة خط انحدار المتغير التابع على المتغير المستقل قيد الدراسة ونُعرِّف معامل التحديد وكيف نفسره.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون

قادراً على:

- حساب معامل ارتباط بيرسون ومعامل ارتباط سبيرمان.
- حساب معامل التحديد وتفسيره.
- حساب معاملي الانحدار.
- إيجاد معادلة خط انحدار المتغير التابع على المستقل.

تحدثنا في الفصل الأول عن المتغيرات والثوابت وعرفنا المتغير المستقل والمتغير التابع وتعرفنا على بعض المقاييس الإحصائية الخاصة بوصف مشاهدات ظاهرة معينة (متغير) دون الأخذ بعين الاعتبار ارتباط هذه الظاهرة بمتغيرات أخرى. وفي هذا الفصل سنبحث وجود علاقة (ارتباط) بين متغيرين أحدهما تابع والآخر مستقل كما هو الحال في معظم التطبيقات النظرية كدراسة العلاقة بين الوزن (متغير مستقل) وضغط الدم (متغير تابع) ودراسة العلاقة بين عدد ساعات اللعب (متغير مستقل) والتحصيل في الامتحانات (متغير تابع). في هذين المثالين إذا وجد ارتباط بين المتغيرين نطلق عليه اسم ارتباط خطي بسيط وهو موضوعنا في هذا الكتاب. وقد يهتم الباحث في دراسته ببحث وجود علاقة بين أكثر من متغيرين كمتغير تابع واحد ومتغيرين مستقلين أو أكثر، كدراسة العلاقة بين العمر والوزن (متغيران مستقلان) وضغط الدم (متغير تابع).

إن بحث وجود علاقة بين المتغيرات قد يتم باستخدام ما يسمى بلوحة الانتشار والتي يمكن من خلالها معرفة اتجاه العلاقة بين المتغيرين وتعطينا وصفاً مبدئياً لمدى قوة العلاقة بينهما ولكن لوحات الانتشار لا تعطينا وصفاً كمياً لقوة العلاقة، لذلك نلجأ لحساب ما يسمى معامل الارتباط الخطي كما سنوضح لاحقاً في هذا الفصل.

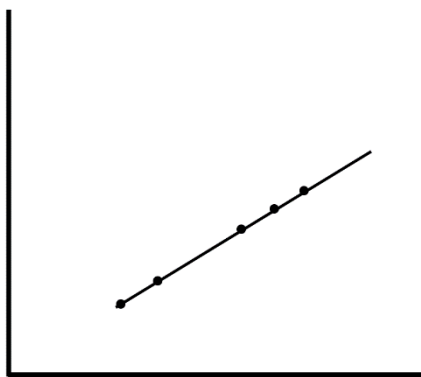
إذا ثبت وجود علاقة مناسبة بين المتغيرات فإن هذا يقودنا إلى البحث عن نماذج رياضية تربط بين المتغيرات والتي نستخدمها للتنبؤ بقيمة المتغير التابع عند القيم المختلفة التي يمكن أن تأخذها المتغيرات المستقلة، وهذا هو موضوع الانحدار الخطي البسيط والذي سنتحدث عنه في هذا الفصل.

Scatter Plot لوحة الانتشار

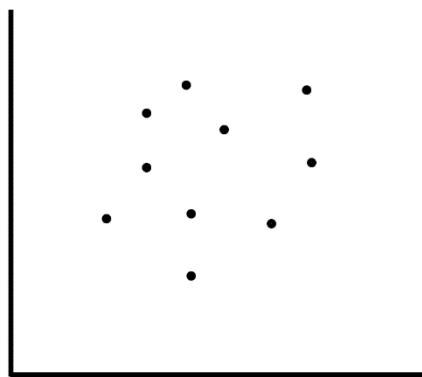
الانتشار هو التمثيل البياني للعلاقة بين متغيرين، ويتم رسم لوحة الانتشار برسم محورين متعامدين وتقسيمهما بمقياس رسم مناسب ثم تمثيل قيم أحد المتغيرين على أحد المحورين وقيم المتغير الآخر على المحور الثاني ونعتبر كل قيمة للمتغير الأول والقيمة التي تقابلها من المتغير

الثاني تمثّلان زوجاً مرتباً والذي يُمثّل بنقطة في المستوى، ومن ثم يتم تحديد ورسم جميع النقاط على المستوى دون التوصيل بينها فنحصل على لوحة الانتشار.

ومن خلال لوحة الانتشار يمكن تحديد وجود علاقة بين المتغيرين من عدمه، وإذا وجدت علاقة بينهما فيمكننا معرفة اتجاه العلاقة بينهما إذا ما كانت طردية أم عكسية، ولكن مثل هذه اللوحات لا يزودنا بقوة العلاقة بين المتغيرين، والرسومات التالية توضح كيفية استنتاج خواص الارتباط من لوحات الانتشار.



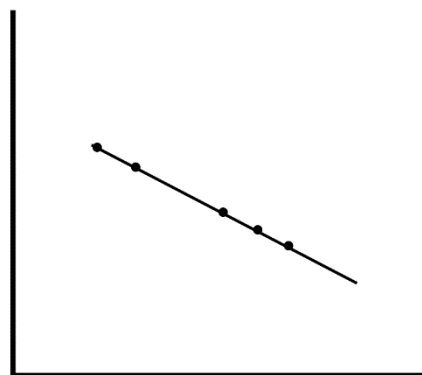
شكل (2)



شكل (1)



شكل (4)



شكل (3)

شكل (1) يبين عدم وجود علاقة بين المتغيرين، حيث تنتشر النقاط في المستوى بشكل عشوائي ولا يوجد أي نمط معين لسلوكها. شكل (2) يبين وجود علاقة تامة حيث تقع جميع النقاط على خط مستقيم واحد وهذه العلاقة لها اتجاه طردي بينما شكل (3) يبين وجود علاقة تامة عكسية. أما شكل (4) فيبين وجود علاقة عكسية بين المتغيرين ولكنها غير تامة لوقوع بعض النقاط خارج الخط المستقيم المحدد، وفي هذه الحالة تكون العلاقة أكثر قوة كلما كانت النقاط الخارجية أقرب للخط المستقيم.

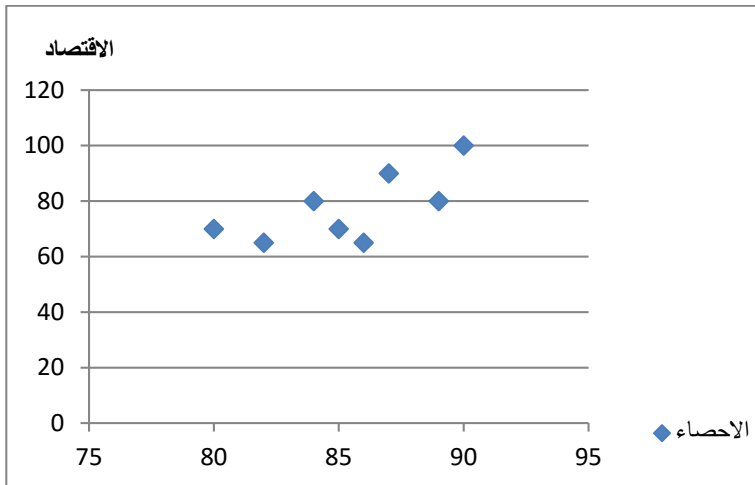
مثال 1.5: في دراسة أجريت لقياس مدى العلاقة بين التحصيل في مادة الإحصاء ومادة الاقتصاد، أجري امتحان في المادتين على ثمانية طلاب وكانت النتائج كما في الجدول التالي:

الإحصاء	89	87	85	82	80	84	86	90
الاقتصاد	80	90	70	65	70	80	65	100

ارسم لوحة الانتشار وبين اتجاه العلاقة بين المتغيرين.

/الحل/

باعتبار أن المحور الأفقي يمثل درجات الإحصاء والمحور الرأسي يمثل درجات الاقتصاد، وبرسم كل درجتين متقابلتين في الجدول كنقطة تمثل كزوج مرتب في المستوى نجد أن لوحة الانتشار تأخذ الشكل التالي:



ونرى من خلال لوحة الانتشار أن الارتباط طردي غير تام. (وواضح أنه لا يوجد خط مستقيم واحد يمكن أن تمر به جميع النقاط).

معامل الارتباط الخطي البسيط Simple Linear Correlation Coefficient

معامل الارتباط الخطي البسيط هو المقياس الرقمي لقوة الارتباط بين المتغيرات و سنرمز له بالرمز r ، وقيمة هذا الرقم تحدد قوة الارتباط أما إشارته فتحدد اتجاه العلاقة، فإن كانت قيمة معامل الارتباط موجبة كانت العلاقة طردية وإذا كانت قيمة معامل الارتباط سالبة كانت العلاقة عكسية.

خصائص معامل الارتباط:

$$-1 \leq r \leq 1 \quad \checkmark$$

✓ إذا كانت $r = 0$ فان الارتباط منعدم ولا توجد علاقة بين المتغيرات.

✓ إذا كانت $r = 1$ فان العلاقة طردية تامة وهي أقوى ارتباط طردي ممكن بين متغيرين

كالعلاقة بين نصف القطر ومساحة الدائرة. و إذا كانت $r = -1$ فان العلاقة عكسية

تامة وهي أقوى ارتباط عكسي كالعلاقة بين طول الشمعة ومدة اشتعالها. وفي التطبيقات

العملية نادراً ما نجد علاقة تامة بين متغيرين.

✓ إذا كانت r موجبة فان العلاقة طردية وإذا كانت سالبة فان العلاقة عكسية.

✓ تزداد قوة العلاقة كلما اقتربنا من 1، -1 وتقل كلما اقتربنا من الصفر.

خلال دراستنا لموضوع الارتباط سنستخدم المعيار التالي لتحديد قوة الارتباط:

- إذا كان $0 < r \leq 0.5$ فإن الارتباط طردي ضعيف.
- إذا كان $0.5 < r \leq 0.75$ فإن الارتباط طردي متوسط.
- إذا كان $0.75 < r < 1$ فإن الارتباط طردي قوي.
- إذا كان $-0.5 \leq r < 0$ فإن الارتباط عكسي ضعيف.

• إذا كان $-0.75 \leq r < -0.5$ فإن الارتباط عكسي متوسط.

• إذا كان $-1 < r < -0.75$ فإن الارتباط عكسي قوي.

وستحدث فيما يلي عن معاملين للارتباط هما معامل ارتباط بيرسون ومعامل ارتباط الرتب لسبيرمان.

1-معامل ارتباط بيرسون Person's Correlation Coefficient

يستخدم هذا المعامل لمعرفة اتجاه وقوة العلاقة بين المتغيرين، ويعد من أفضل معاملات الارتباط وأكثرها شيوعاً. فبفرض أن هناك متغيران Y, X مستقلان وكل منها طبيعي التوزيع ويأخذ كل منها n قيمة وهي على الترتيب:

$$y_1, y_2, \dots, y_n \text{ و } x_1, x_2, \dots, x_n$$

فإن معامل ارتباط بيرسون بين المتغيرين Y, X يعطى بالعلاقة:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{nS_X S_Y}$$

حيث $\bar{X}$ هو الوسط الحساب لمشاهدات المتغير X ، و $\bar{Y}$ هو الوسط الحسابي لمشاهدات المتغير Y وكذلك S_X هو الانحراف المعياري لمشاهدات المتغير X ، و S_Y هو الانحراف المعياري لمشاهدات المتغير Y .

ويوجد صورة أخرى للقانون تعتبر الأسهل في الحسابات وهي:

$$r = \frac{\sum x_i y_i - n\bar{X}\bar{Y}}{\sqrt{\sum x_i^2 - n\bar{X}^2} \sqrt{\sum y_i^2 - n\bar{Y}^2}}$$

معامل ارتباط بيرسون



مثال 2.5: احسب معامل ارتباط بيرسون بين علامات عشرة طلاب في مساق التحاليل الطبية

(X) و التشریح (Y) من الجدول التالي:

80	60	55	40	75	85	70	60	80	90	X
75	65	60	50	70	90	70	55	80	85	Y

الحل/

نكون جدول الحساب وهو كما يلي:

جدول الحساب

X^2	Y^2	XY	Y	X
6400	5625	6000	75	80
3600	4225	3900	65	60
3025	3600	3300	60	55
1600	2500	2000	50	40
5625	4900	5250	70	75
7225	8100	7650	90	85
4900	4900	4900	70	70
3600	3025	3300	55	60
6400	6400	6400	80	80
8100	7225	7650	85	90
50475	50500	50350	700	695

المجموع

ومن الجدول نحصل على:

$$\bar{X} = \frac{\sum x}{n} = \frac{695}{10} = 69.5 \quad , \quad \bar{Y} = \frac{\sum y}{n} = \frac{700}{10} = 70$$

$$\begin{aligned}
 r &= \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sqrt{\sum x_i^2 - n \bar{X}^2} \sqrt{\sum y_i^2 - n \bar{Y}^2}} \\
 &= \frac{50350 - 10 \times 69.5 \times 70}{\sqrt{5047 - 10(69.5)^2} \sqrt{50500 - 10(70)^2}} \\
 &= \frac{1700}{\sqrt{2172.5} \sqrt{1500}} = \frac{1700}{1805.2} = 0.94
 \end{aligned}$$

ومن الواضح أن الارتباط طردي قوي.

2- معامل ارتباط لسبيرمان Spearman's Correlation Coefficient

يستخدم هذا المعامل إذا كانت هناك صعوبة في استخدام معامل ارتباط بيرسون مثل كون المشاهدات عبارة عن بيانات نوعية وليست كمية ويعتمد على رتب المشاهدات وليس على قيمها حتى لو كانت كمية، وكذلك إذا توفرت رتب القيم فقط أو عدم خضوع أحد المتغيرين للتوزيع الطبيعي وهو شرط أساسي لتطبيق قانون بيرسون وكذلك عدم استقلالية المتغيرين.

ويحسب معامل ارتباط سبيرمان من العلاقة:

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

معامل ارتباط سبيرمان



حيث n عدد المشاهدات لكل متغير، d_i فرق الرتب بين قيمتين متناظرتين للمتغيرين.

وبفرض رتب X هي O_X ورتب Y هي O_Y فان:

$$d_i = O_{x_i} - O_{y_i}$$

وفيما يلي خطوات من المفيد اتباعها بالترتيب نوضح من خلالها كيفية حساب معامل ارتباط الرتب لسبيرمان:

خطوات حل المسألة:

1- نحدد رتب قيم X كما يلي:

- نرتب قيم X تنازلياً (أو تصاعدياً).
- نعطي قيم X أرقاماً متسلسلة مبتدئاً بـ 1 الى آخر قيمة لـ X .
- قيم X غير المكررة تكون رتبها هي نفس ارقامها المتسلسلة أما قيم X المكررة تكون رتبها هي الوسط الحسابي لأرقامها المتسلسلة.

2- بنفس الطريقة السابقة نحدد رتب قيم Y .

3- عند بناء الجدول نضع قيم X و قيم Y مرتبة كما وردت في المسألة بدون تغيير ثم نضع أمام كل قيمة رتبها التي حسبناها من قبل.

ملاحظة/ يجب ترتيب قيم المتغيرين عند حساب الرتب بنفس الطريقة إما تصاعدياً معاً وإما تنازلياً معاً.

مثال 3.5: تقدم سبعة طلاب لاختبارين في مادتي الإحصاء والرياضيات وكانت تقديراتهم كما هو مبين في الجدول التالي:

الاحصاء X	جيد	ممتاز	جيد	جيد جداً	ضعيف	مقبول	ممتاز
الرياضيات Y	مقبول	ممتاز	جيد	جيد	ضعيف	ضعيف	جيد جداً

احسب معامل ارتباط سبيرمان.

نكون جدول الرتب كما يلي:

جدول حساب الرتب

X	Y	O_X	O_Y	d_i	d_i^2
جيد	مقبول	3.5	3	0.5	0.25
ممتاز	ممتاز	6.5	7	-0.5	0.25
جيد	جيد	3.5	4.5	-1	1
جيد جدا	جيد	5	4.5	0.5	0.25
ضعيف	ضعيف	1	1.5	-0.5	0.25
مقبول	ضعيف	2	1.5	0.5	0.25
ممتاز	جيد جدا	6.5	6	0.5	0.25
					2.5

لاحظ أن $n = 7$ ، ومن الجدول وبالتعويض المباشر في قانون معامل ارتباط سبيرمان

نحسب معامل الارتباط كما يلي:

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \times 2.5}{7(49 - 1)}$$

$$= 0.96$$

ومن الواضح أن الارتباط طردي قوي.

مثال 4.5: أوجد معامل ارتباط الرتب لقيم المتغيرين المبينة في الجدول التالي والذي يمثل تكاليف

الدعاية لنوع من البضائع وقيمة المبيعات بالدنانير:

9	11	5	13	12	4	6	10	8	التكاليف X
150	160	120	180	165	130	150	160	150	المبيعات Y

الحل/ نكون جدول الرتب كما يلي:

جدول حساب الرتب

X	Y	O_X	O_Y	d_i	d_i^2
8	150	6	6	0	0
10	160	4	3.5	0.5	0.25
6	150	7	6	1	1
4	130	9	8	1	1
12	165	2	2	0	0
13	180	1	1	0	0
5	120	8	9	-1	1
11	160	3	3.5	-0.5	0.25
9	150	5	6	-1	1
					4.5

لاحظ أن $n = 9$ ، ومن الجدول وبالتعويض المباشر في قانون معامل ارتباط سبيرمان

نحسب معامل الارتباط كما يلي:

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \times 4.5}{9(81 - 1)}$$

$$= 1 - 0.0375 = 0.96$$

وهو ارتباط طردي قوي.

ملاحظات:

1- إذا حسبنا لمثال (5.5) معامل ارتباط بيرسون فإنه سيكون أدق من معامل ارتباط سبيرمان،

وذلك لأن معامل ارتباط بيرسون يتعامل مع القيم الحقيقية للمتغيرين ولكن معامل الرتب

يتعامل مع رتب هذه القيم.

2- في حالة البيانات الكمية لا يعتمد معامل ارتباط سبيرمان على القيم نفسها بل على تباعدها عن بعضها.

3- إذا كان r معامل الارتباط بين المتغيرين X, Y وأجرينا تعديلاً على كل من قيم المتغيرين بالتحويلين الخطيين التاليين: $L = aX + b$, $Z = cY + d$ حيث a, b, c, d هي أعداد حقيقية و $a \neq 0, c \neq 0$ ، فإن معامل الارتباط بين المتغيرين L, Z لا يتأثر إذا اتفقت a و c في الإشارة و تتغير إشارته فقط إذا اختلفتا في الإشارة.

فإذا كان r^* هو معامل الارتباط بعد التعديل فيمكن تلخيص ما سبق كما يلي:

$$r^* = \begin{cases} r, & ac > 0 \\ -r, & ac < 0 \end{cases}$$

مثال 6.5: أعد حل مثال (2.5) بالتأثير على المتغيرين بالتحويلين الخطيين التاليين:

$$L = 0.1X + 1, \quad Z = 0.2Y - 1$$

/الحل

نكون جدول الحساب كما يلي:

Z^2	L^2	LZ	Z	L	Y	X
196	81	126	14	9	75	80
144	49	84	12	7	65	60
121	42.25	71.5	11	6.5	60	55
81	25	45	9	5	50	40
169	72.25	110.5	13	8.5	70	75
289	90.25	161.5	17	9.5	90	85
169	64	104	13	8	70	70
100	49	70	10	7	55	60
225	81	135	15	9	80	80
256	100	160	16	10	85	90
1750	653.75	1067.5	130	79.5	المجموع	

ومن الجدول نحصل على:

$$\bar{L} = \frac{\sum l}{10} = \frac{79.5}{10} = 7.95$$

$$\bar{Z} = \frac{\sum z}{10} = \frac{130}{10} = 13$$

$$\begin{aligned} r &= \frac{\sum lz - n\bar{L}\bar{Z}}{\sqrt{\sum l^2 - n\bar{L}^2} \sqrt{\sum z^2 - n\bar{Z}^2}} \\ &= \frac{1067.5 - 10 \times 7.95 \times 13}{\sqrt{653.75 - 10(7.95)^2} \sqrt{1750 - 10(13)^2}} = 0.94 \end{aligned}$$

وواضح هنا أن معامل الارتباط لم يتغير.

4- يمكن استخدام التحويلات التالية لإيجاد معامل ارتباط بيرسون:

$$L = X - \bar{X}, \quad Z = Y - \bar{Y}$$

والتي تبسط الحسابات بشكل كبير حيث يصبح معامل ارتباط بيرسون كما يلي:

$$r = \frac{\sum l_i z_i}{\sqrt{\sum l_i^2} \sqrt{\sum z_i^2}}$$

الانحدار الخطي البسيط Simple Linear Regression

إذا وجد الباحث أن المتغيرين قيد الدراسة مرتبطتين معاً بقوة مناسبة فمن الممكن التفكير بإيجاد نموذج رياضي يمثل معادلة خط مستقيم يطلق عليه خط الانحدار، حيث يستخدم للتنبؤ بقيمة المتغير التابع إذا عُلِمَت قيمة المتغير المستقل. وتكون معادلة خط انحدار المتغير التابع Y على المتغير المستقل X كما يلي:

معادلة خط انحدار Y على X

$$\hat{Y} = aX + b$$



حيث a تمثل معامل خط انحدار Y على X وهي تعبر عن ميل خط الانحدار والذي يفسر بأنه التغير في المتغير التابع الناتج عن تغير المتغير المستقل بوحدة واحدة، و b تعبر عن الجزء المقطوع من محور الصادات والذي يمثل قيمة المتغير التابع في حالة مساواة المتغير المستقل للصفر (والذي يمثل محور السينات).

فإذا أعطينا قيمة ما للمتغير X تقع في مداه في المعادلة السابقة نحصل علي قيمة مناظرة للمتغير Y والتي تعتبر القيمة المُقدَّرة للمتغير التابع وهي التي نرسم لها بالرمز $\hat{Y}$ والتي نقرأها " Y هات" لنميزها عن القيم الحقيقية للمتغير Y .

يطلق على الفرق بين القيمة المقدرة والقيمة الحقيقية للمتغير التابع اسم الخطأ في التقدير ويرمز له بالرمز e (اختصاراً لكلمة **error** والتي تعني خطأ).

الخطأ في التقدير

$$e = Y - \hat{Y}$$



ملاحظات:

1- إذا كانت a موجبة فإن علاقة المتغير التابع Y بالمتغير المستقل X تكون

علاقة طردية ويصنع خط الانحدار $\hat{Y} = aX + b$ زاوية حادة مع الاتجاه الموجب لمحور السينات.

2- إذا كانت $a = 0$ فينعدم الارتباط الخطي بين المتغيرين ويكون خط الانحدار ثابت $\hat{Y} = b$.

3- إذا كانت a سالبة فإن علاقة المتغير التابع Y بالمتغير المستقل X تكون

علاقة عكسية ويصنع خط الانحدار $\hat{Y} = aX + b$ زاوية منفرجة مع الاتجاه الموجب لمحور السينات.

4- إذا كانت $b = 0$ فإن خط الانحدار يمر بنقطة الأصل.

5- إذا عُلمت زاوية ميل خط الانحدار θ مع الاتجاه الموجب لمحور السينات فإن:

$$a = \tan \theta$$

6- إذا كان الارتباط بين المتغيرين تاماً فإن الخطأ في التقدير يساوي صفراً.

ويجدر القول هنا أن اهتمامنا يكون على العلاقة نفسها بين Y, X وليس على سبب

وجود العلاقة أو الظروف المحيطة بها فهناك علاقة سببية بين كمية السماد وكمية الناتج

الزراعي للقمح مثلاً ولكننا لا يمكن أن نجزم بوجود علاقة سببية لعدد المساكن وميزانية الدولة.

لإيجاد معادلة خط الانحدار $\hat{Y} = aX + b$ نحسب قيمة كل من a و b كما يلي:

لا تنسى ميل خط انحدار Y على X بمعلومية الانحراف المعياري ومعامل الارتباط

$$a = \frac{S_Y}{S_X} r$$



حيث S_X الانحراف المعياري لملاحظات المتغير X ، S_Y الانحراف المعياري لملاحظات المتغير Y و r معامل الارتباط بين المتغيرين.

ويوجد طريقة أخرى لحساب معامل خط الانحدار وهي طريقة المربعات الصغرى وتستخدم عند معرفة قيم المتغيرات والمجاميع المطلوبة حيث نحسب a كما يلي:

لا تنسى ميل خط انحدار Y على X بمعلومية أوساط المتغيرين ومجاميع خاصة

$$a = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum x_i^2 - n \bar{X}^2}$$



أما قيمة b فيتم إيجادها من العلاقة :

الجزء المقطوع من محور المتغير التابع لخط انحدار Y على X

$$b = \bar{Y} - a \bar{X}$$



ونذكر بأن معادلة خط الانحدار $\hat{Y} = aX + b$ تستخدم للتنبؤ بقيمة Y إذا علمت قيمة X وسنكتفي في هذا الكتاب بحساب هذه المعادلة.

ويمكن إيجاد معادلة خط انحدار المتغير X على المتغير Y حيث المعادلة هي:

$$\hat{X} = cY + d$$

حيث c تمثل معامل خط انحدار X على Y وهي تعبر عن ميل خط الانحدار، و d تعبر عن الجزء المقطوع من محور المتغير X و $\hat{X}$ القيمة المقدرة للمتغير X .

لإيجاد معادلة خط الانحدار $\hat{X} = cY + d$ نحسب قيمة كل من c و d كما يلي:

لا تنسى ميل خط انحدار X على Y بمعلومية الانحراف المعياري ومعامل الارتباط

$$c = \frac{S_X}{S_Y} r$$



حيث S_X الانحراف المعياري لملاحظات المتغير X ، S_Y الانحراف المعياري لملاحظات المتغير Y و r معامل الارتباط بين المتغيرين.

ويوجد طريقة أخرى لحساب معامل الانحدار وهي طريقة المربعات الصغرى وتستخدم عند معرفة قيم المتغيرات والمجاميع المطلوبة حيث نحسب c كما يلي:

لا تنسى ميل خط انحدار X على Y بمعلومية أوساط المتغيرين ومجاميع خاصة

$$c = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum y_i^2 - n \bar{Y}^2}$$



أما قيمة d فيتم إيجادها من العلاقة :

الجزء المقطوع من محور المتغير التابع لخط انحدار X على Y

$$d = \bar{X} - c \bar{Y}$$



ملاحظة: إذا رسمنا خطي الانحدار على نفس المستوى الديكارتي فإن الخطين يتقاطعان عند النقطة $(\bar{X}, \bar{Y})$.

معامل التحديد Coefficient of Determination

يمثل معامل التحديد النسبة التي تفسرها معادلة خط الانحدار في التغير في المتغير التابع ويرمز له بالرمز r^2 وهي نفسها قيمة مربع معامل ارتباط بيرسون وتساوي أيضاً مربع حاصل ضرب معاملي الانحدار.

معامل التحديد

$$r^2 = ac$$



ملاحظات:

- $0 \leq r^2 \leq 1$ بمعنى أن معامل التحديد قيمته موجبة دائماً.
- إذا كانت اشارتا معاملي الانحدار موجبة فالارتباط طردي ومعامل ارتباط بيرسون يساوي: $r = +\sqrt{a \times c}$
- إذا كانت اشارتا معاملي الانحدار سالبة فالارتباط عكسي ومعامل ارتباط بيرسون يساوي: $r = -\sqrt{a \times c}$
- إذا كان: $r^2 = 0$ فإن معامل التحديد لا يفسر شيئاً من التباين في المتغير التابع.

مثال 7.5: في شركة لتجارة السيارات يمثل الجدول التالي عدد السيارات المباعة X في السنوات 2000-2004 و الربح Y بآلاف الدنانير:

X	50	40	45	55	60
Y	75	63	50	72	80

أوجد:

- 1- معادلة خط انحدار Y على X .
- 2- فسر معامل الانحدار والجزء المقطوع من المتغير التابع.
- 3- لو افترضنا أن الشركة ستبيع 50 سيارة في العام القادم فما الربح المتوقع.
- 4- احسب الخطأ في التقدير.
- 5- احسب معامل التحديد وفسره.

الحل/

نكون جدول الحساب التالي:

جدول الحساب

X	Y	XY	X^2	Y^2	
50	75	3750	2500	5625	
40	63	2520	1600	3969	
45	50	2250	2025	2500	
55	72	3960	3025	5184	
60	80	4800	3600	6400	
250	340	17280	12750	23678	المجموع

نحسب الوسط الحسابي لكل من المتغيرين:

$$\bar{X} = \frac{\sum x}{n} = \frac{250}{5} = 50$$
$$\bar{Y} = \frac{\sum y}{n} = \frac{340}{5} = 68$$

1- نحسب معامل الانحدار a :

$$a = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum x_i^2 - n \bar{X}^2} = \frac{17280 - 5 \times 50 \times 68}{12750 - 5 \times (50)^2} = \frac{280}{250} = 1.12$$

نحسب الجزء المقطوع من محور الصادات:

$$b = \bar{Y} - a \bar{X} = 68 - 1.12 \times 50 = 12$$

فتكون معادلة خط الانحدار هي:

$$\hat{Y} = 1.12X + 12$$

2- يخبرنا معامل الانحدار بأنه لو تم بيع سيارة واحدة في السنة فإن الربح المتوقع سيكون 1120 دينار، ويخبرنا الجزء المقطوع بأن الربح سيكون 12000 دينار في حال عدم انتاج سيارات وهذا تفسير رياضي.

3- الربح المتوقع للشركة إذا باعت 50 سيارة:

في هذه الحالة تكون قيمة $X = 50$ ، لذلك بالتعويض المباشر في معادلة خط الانحدار نحصل على :

$$\hat{Y} = 1.12 \times 50 + 12 = 68$$

وبالتالي يكون الربح 68×1000 دينار = 68000 دينار.

4- الخطأ في التقدير يساوي في هذه الحالة:

$$e = Y - \hat{Y} = 75 - 68 = 7$$

حيث القيمة الحقيقية للمتغير التابع هي تلك التي في الجدول وتقابل القيمة 50 للمتغير المستقل وقد تم الإشارة اليهما بالتظليل في الجدول، والخطأ في التقدير يساوي 7000 دينار.

5- لحساب معامل التحديد نحسب معامل انحدار X على Y :

$$c = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum y_i^2 - n \bar{Y}^2} = \frac{17280 - 5 \times 50 \times 68}{23678 - 5 \times (68)^2} = \frac{280}{558} = 0.502$$

ويكون معامل التحديد يساوي:

$$r^2 = ac = 1.12 \times 0.502 = 0.5622$$

وهذا يعني أن خط الانحدار يفسر ما نسبته 56.22% من التباين في المتغير

التابع.

مثال 8.5: إذا علمت أن $\bar{X} = 10, \bar{Y} = 20, S_X = 1.5, S_Y = 2, r = 0.6$ أوجد:

1- معادلة انحدار Y على X .

2- معامل انحدار X على Y .

3- معامل التحديد وفسره.

الحل/

1- نوجد أولاً معامل انحدار Y على X :

$$a = \frac{S_Y}{S_X} r = \frac{2}{1.5} \times 0.6 = 0.8$$

ثم الجزء المقطوع من محور المتغير التابع:

$$b = \bar{Y} - a\bar{X} = 20 - 0.8 \times 10 = 12$$

فتكون معادلة خط الانحدار المطلوبة هي:

$$\hat{Y} = 0.8X + 12$$

2- إيجاد معامل انحدار X على Y :

$$c = \frac{S_X}{S_Y} r = \frac{1.5}{2} \times 0.6 = 0.45$$

3- معامل التحديد:

$$r^2 = (0.6)^2 = 0.36$$

أو بطريقة أخرى:

$$r^2 = ac = 0.8 \times 0.45 = 0.36$$

وهذا يعنى أن خط الانحدار يفسر ما نسبته 36% من تباين المتغير التابع.

مثال 9.5: إذا علمت أن الانحراف المعياري لقيم X , Y يساوي 0.61 ، 1.22 على الترتيب فاحسب معامل الارتباط بين المتغيرين إذا علمت أن معادلة خط انحدار Y على X هي: $\hat{Y} = 0.84X + 21.28$.

الحل/

من معادلة الانحدار نجد أن معامل انحدار Y على X يساوي: $a = 0.84$ وبالتالي:

$$a = \frac{S_Y}{S_X} r$$

$$0.84 = \frac{1.22}{0.61} \times r$$

$$r = \frac{0.84 \times 0.61}{1.22} = 0.42$$

مثال 10.5: إذا علمت أن معادلتى انحدار Y على X و X على Y هما:

$$\hat{X} = 0.288Y + 16.792 \quad , \quad \hat{Y} = 0.87X + 6.75$$

على الترتيب، وكان الانحراف المعياري لقيم المتغير X يساوي 0.8

أوجد:

1- معامل ارتباط بيرسون.

2- معامل التحديد.

3- الانحراف المعياري لقيم المتغير Y .

الحل/

1- من معادلتى خط الانحدار نجد أن معاملى الانحدار يساويا:

$$a = 0.87 \quad , \quad c = 0.288$$

وبالتالى يكون معامل ارتباط بيرسون يساوي:

$$r = +\sqrt{a \times c} = \sqrt{0.87 \times 0.288} = 0.5$$

2- معامل التحديد:

$$r^2 = (0.5)^2 = 0.25$$

3- الانحراف المعياري لقيم المتغير Y :

$$a = \frac{S_Y}{S_X} r$$

$$0.87 = \frac{S_Y}{0.8} \times 0.5$$

$$0.5 \times S_Y = 0.8 \times 0.87$$

$$S_Y = \frac{0.8 \times 0.87}{0.5} = 1.392$$

مثال 11.5: في دراسة ما تم الحصول على البيانات التالية:

$$\sum_{i=1}^6 x_i = 48 , \quad \sum_{i=1}^6 y_i = 240 , \quad \sum_{i=1}^6 x_i y_i = 2016$$

$$\sum_{i=1}^6 x_i^2 = 432 , \quad \sum_{i=1}^6 y_i^2 = 10000$$

أوجد معادلة خط انحدار Y على X ثم استنتج قيمة معامل الارتباط.

الحل/

$$\bar{X} = \frac{\sum x}{n} = \frac{48}{6} = 8$$

$$\bar{Y} = \frac{\sum y}{n} = \frac{240}{6} = 40$$

نحسب معامل الانحدار a :

$$a = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum x_i^2 - n \bar{X}^2} = \frac{2016 - 6 \times 8 \times 40}{432 - 6 \times (8)^2} = 2$$

نحسب الجزء المقطوع من محور الصادات:

$$b = \bar{Y} - a \bar{X} = 40 - 2 \times 8 = 24$$

فتكون معادلة خط الانحدار هي:

$$\hat{Y} = 2X + 24$$

لإيجاد قيمة معامل الارتباط نحسب معامل انحدار X على Y :

$$c = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum y_i^2 - n \bar{Y}^2} = \frac{2016 - 6 \times 8 \times 40}{10000 - 6 \times (40)^2} = 0.24$$

وبالتالي:

$$r = +\sqrt{a \times c} = \sqrt{2 \times 0.24} = 0.69$$

تمارين:

1- لمثال (2.5) احسب معامل ارتباط سبيرمان وقارن بينه وبين معامل ارتباط بيرسون.

2- أعد حل مثال (2.5) للقيم الناتجة من التحويلات الخطية التالية:

$$Z = 2Y - 8, L = X - 5$$

$$Z = -Y + 3, L = 2X - 7$$

$$Z = Y - \bar{Y}, L = X - \bar{X}$$

3- يمثل الجدول التالي الأطوال (X) بالسنتيمتر والأوزان (Y) بالكيلوجرام لستة طلاب في إحدى كليات المجتمع:

170	172	165	175	168	180	X
70	72	73	72	70	77	Y

• ارسم لوحة الانتشار.

• احسب معامل ارتباط بيرسون بين أطوال وأوزان الطلبة.

• احسب معامل ارتباط سبيرمان بين الأطوال والأوزان.

• أوجد معادلة خط انحدار Y على X .

4- إذا كان معامل الارتباط بين Y, X هو $r = 0.7$ و عُدَّت البيانات حسب المعادلتين:

$$L = -2.5X + 7, Z = 3Y - 6$$

فما هو معامل التحديد بين Z, L ؟

5- للجدول التالي:

2	4	6	8	10	X
15	19	12	6	3	Y

• جد معامل ارتباط بيرسون بين قيم X و Y .

• معادلة خط انحدار Y على X .

• احسب الخطأ في التنبؤ بقيمة Y عندما تكون: $X = 8$.

6- تقدم ثمانية طلاب لدورتين تدريبيتين وكانت تقديراتهما كما هو موضح في الجدول التالي:

الدورة الأولى	جيد	ممتاز	جيد	جيد	جيد جدا	ضعيف	مقبول	ممتاز
الدورة الثانية	مقبول	ممتاز	ضعيف	جيد	جيد	ضعيف	ضعيف	جيد جدا

احسب معامل الارتباط بين تقديرات الدورتين.

7- إذا كانت معادلة خط انحدار Y على X هي: $\hat{Y} = -15 + 2X$ وكانت $r = 0.8$ و الوسط الحسابي للمتغير X يساوي 50، اوجد قيمة الوسط الحسابي للمتغير Y ومعامل التحديد وفسره .

8- إذا كانت معادلة خط انحدار Y على X هي: $\hat{Y} = -2 + 1.2X$ و معادلة خط انحدار X على Y هي: $\hat{X} = 11 + 0.7Y$ اوجد الوسط الحسابي لكل من قيم Y, X ؟

9- في امتحان تحصيلي لستة طلاب في مادتي الرياضيات (X) و الإحصاء (Y) كانت النتائج كالتالي:

60	75	50	40	63	72	X
80	83	55	70	60	78	Y

أوجد:

- معادلة خط انحدار Y على X .
- إذا حصل طالب على علامة 62 في الرياضيات فكم تكون علامته في الإحصاء.
- قدر درجة الطالب في الإحصاء إذا كانت درجة الرياضيات 60 ثم احسب الخطأ في التقدير.

10- لقيم المتغيرين في الجدول التالي:

8	5	9	7	6	X
6	3	2	10	4	Y

اوجد:

- معامل ارتباط بيرسون بين المتغيرين.
- معادلة خط انحدار Y على X . ماذا تستنتج؟

11- لقيم المتغيرين في الجدول التالي:

1	2	3	4	5	X
2	4	6	8	10	Y

أوجد:

- معامل ارتباط بيرسون بين المتغيرين.
 - معادلة خط انحدار Y على X. ماذا تستنتج؟
- 12- إذا كانت زاوية ميل خط انحدار Y على X مع الاتجاه الموجب لمحور السينات تساوي 45 درجة، وكان خط الانحدار يمر بنقطة الأصل، أوجد معادلة خط الانحدار. ما العلاقة بين متوسطي المتغيرين؟

- 13- أياً من القيم التالية تصلح لأن تكون معاملي انحدار:
- (10 ، 12) ، (-0.3 ، 1.5) ، (-0.9 ، -0.5)

- 14- البيانات التالية خاصة بالسعر والكمية المطلوبة من سلعة معينة:

5	7	3	5	4	5	6	السعر
11	5	6	10	9	7	8	الكمية

- احسب معامل الارتباط بين السعر والكمية.
 - احسب معامل التحديد.
 - أوجد معادلة خط انحدار السعر على الكمية.
- 15- في دراسة معينة على عينة مكونة من 12 مشاهدة لمتغيرين Y و X وجد أن:

$$\sum_{i=1}^{12} x_i = 725 , \quad \sum_{i=1}^{12} x_i y_i = 61685$$

$$\sum_{i=1}^{12} x_i^2 = 44475 , \quad \bar{Y} = 84.25 , \quad r = 0.75$$

بين أي المتغيرين أكثر تبايناً من الآخر.

16- إذا علمت أن معادلة خط انحدار Y على X هي: $\hat{Y} = 12 + 0.8X$ ومعادلة خط انحدار X على Y هي: $\hat{X} = 1 + 0.45Y$ والانحراف المعياري لقيم المتغير X يساوي 1.5

- احسب قيمة معامل الارتباط بين المتغيرين.
- أي المتغيرين أكثر تبايناً.
- احسب معامل التحديد.

17- اختر الإجابة الصحيحة فيما يلي:

- إذا كان معامل الارتباط بين متغيرين X , Y يساوي 0.068 فإن الارتباط بين المتغيرين يكون:

- أ) طري ضعيف ب) عكسي ضعيف ج) طردي قوي د) عكسي قوي
- أقوى معامل ارتباط عكسي فيما يلي هو:

- أ) 0.9 ب) -0.73 ج) -0.85 د) -1.2

- إذا كان معامل انحدار y على X يساوي 0.9 ومعامل انحدار X على Y يساوي 0.4 فإن معامل التحديد يساوي:

- أ) 0.35 ب) 0.36 ج) 0.81 د) 0.16

- إذا كانت معادلة خط انحدار y على X هي $Y = 0.3X + 2$ فإن معامل خط الانحدار يساوي:

- أ) 2 ب) 0.6 ج) 2.3 د) 0.3

- عند حساب معامل ارتباط سبيرمان بين قيم X , Y إذا كان $\frac{6\sum d_i^2}{n(n^2-1)} = 2$ فإن قيمة معامل الارتباط بين المتغيرين تساوي:

- أ) 1 ب) -1 ج) 0 د) لا يمكن حسابه

- أي القيم تصلح لأن تكون معاملي انحدار:

- أ) 0.2 ، 1.19 ب) 0.25 ، 1 ج) 1.5 ، 3.7 د) 2 ، 1

الخلاصة

معامل ارتباط بيرسون:

يعتبر معامل ارتباط بيرسون من أكثر معاملات الارتباط استخداماً عند توفر الشروط المناسبة وهو يستخدم للبيانات الكمية حيث يحسب من العلاقة:

$$r = \frac{\sum x_i y_i - n\bar{X}\bar{Y}}{\sqrt{\sum x_i^2 - n\bar{X}^2} \sqrt{\sum y_i^2 - n\bar{Y}^2}}$$

معامل ارتباط سبيرمان:

وهو للبيانات النوعية ويستخدم رتب البيانات عند حسابه ويحسب من العلاقة:

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

معادلة خط انحدار Y على X :

$$\hat{Y} = aX + b$$

حيث:

$$a = \frac{\sum x_i y_i - n\bar{X}\bar{Y}}{\sum x_i^2 - n\bar{X}^2}, \quad a = \frac{S_Y}{S_X} r, \quad b = \bar{Y} - a\bar{X}$$

معادلة خط انحدار X على Y :

$$\hat{X} = aY + b$$

حيث:

$$c = \frac{\sum x_i y_i - n \bar{X} \bar{Y}}{\sum y_i^2 - n \bar{Y}^2}, \quad c = \frac{S_X}{S_Y} r, \quad d = \bar{X} - a \bar{Y}$$

معامل التحديد:

يحدد معامل التحديد النسبة التي يفسرها خط الانحدار للمتغير التابع.

$$r^2 = ac$$

الخطأ في التقدير:

$$e = Y - \hat{Y}$$

مقدمة في الاحتمالات Introduction to Probability

6

يعتبر مفهوم الاحتمال من المفاهيم الهامة في الإحصاء، وهو مفهوم يُعنى بفرص وقوع حدث معين. في هذا الفصل سنتحدث باختصار عن بعض المفاهيم الإحصائية: التجربة العشوائية، فضاء العينة، الحدث، وقوع الحدث، الاحتمال، مسلمات الاحتمال، متمم الحدث، الفرق بين حدثين، اتحاد حدثين والاحتمال الشرطي. ونتحدث في نهاية الفصل عن نظرية التجزئة ونظرية بيز.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً على:

- إيجاد فضاء العينة لتجربة عشوائية.
- إيجاد حوادث معينة من فضاء عينة محدد.
- حساب احتمالات حوادث معينة بتطبيق قوانين الاحتمال المذكورة في الفصل.
- حساب الاحتمالات الشرطية.
- تطبيق نظرية التجزئة وقانون بيز.

الاحتمال لغةً هو أحد الخيارات المتاحة أمام تجربة أو حادثة غير محسومة النتيجة. وفي الرياضيات تعبر كلمة الاحتمال عن قيمة عددية تدل على مدى أو نسبة تكرارية هذا الخيار عند تطبيق التجربة لمرات عديدة. ففي العلوم التجريبية ليس بالضرورة الحصول على نفس النتيجة في حال تكرار التجربة أكثر من مرة، لذلك من المستحسن دائماً تكرار التجربة عدداً كبيراً من المرات لعلنا نستطيع استنباط علاقة بين المتغيرات التي ندرسها.

سننتظر في هذا الفصل لبعض المفاهيم الهامة في علم الاحتمال كمفهوم التجربة العشوائية، فضاء العينة والحدث، وسنوضح كيفية إيجاد فضاء العينة لبعض التجارب العشوائية بالإضافة إلى حساب احتمالات حوادث معينة من فضاء العينة.

تعريف

Random Experiment التجربة العشوائية

هي التجربة التي يمكن حصر جميع نتائجها الممكنة قبل اجراءها ولا نستطيع تحديد النتيجة التي سنحصل عليها بالفعل قبل إجراءها.

تعريف

Sample Space فضاء العينة

هو مجموعة كل النتائج الممكنة للتجربة العشوائية، وسنرمز له بالرمز Ω و التي تقرأ أوميغا (Omega).

الحدث Event

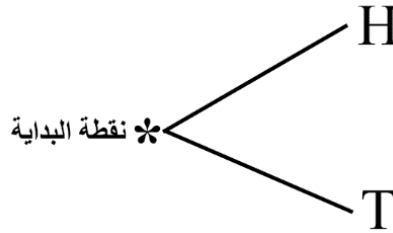
هو مجموعة جزئية من فضاء العينة.

سنذكر الآن بعض الأمثلة على إيجاد فضاءات العينة لبعض التجارب العشوائية عن طريق رسم ما يسمى بالشجرة البيانية.

مثال 1.6: اكتب فضاء العينة لتجربة القاء قطعة نقد مرة واحدة.

الحل/

أولاً نرسم الشجرة البيانية والتي نبدأها بنقطة البداية ثم نرسم منها خطوطاً مستقيمة بعدد نواتج التجربة ونكتب في نهاية كل خط إحدى هذه النواتج وذلك كما يلي:



حيث تشير H إلى ظهور صورة (H: Head) و T إلى ظهور كتابة (T: Tail) ويكون فضاء العينة كما يلي:

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

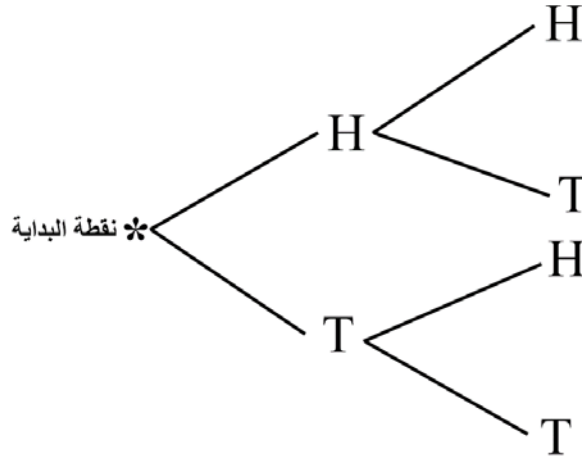
لاحظ أن فضاء العينة يحتوي على نتيجتين فقط، ونعبر عن ذلك بكتابة: $\#\Omega = 2$ والتي تُقرأ "عدد عناصر Ω يساوي 2".

مثال 2.6: اكتب فضاء العينة لتجربة إلقاء قطعة نقد مرتين.

الحل/

نرسم الشجرة البيانية مع ملاحظة أننا نبدأ من نقطة البداية برسم نتائج الرمية الأولى

ومن ثم نرسم نتائج الرمية الثانية عند جميع نواتج الرمية الأولى وذلك كما يلي:



ومن الرسم نجد أن:

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

ويكون عدد عناصر فضاء العينة: $\#\Omega = 4$

ملاحظات:

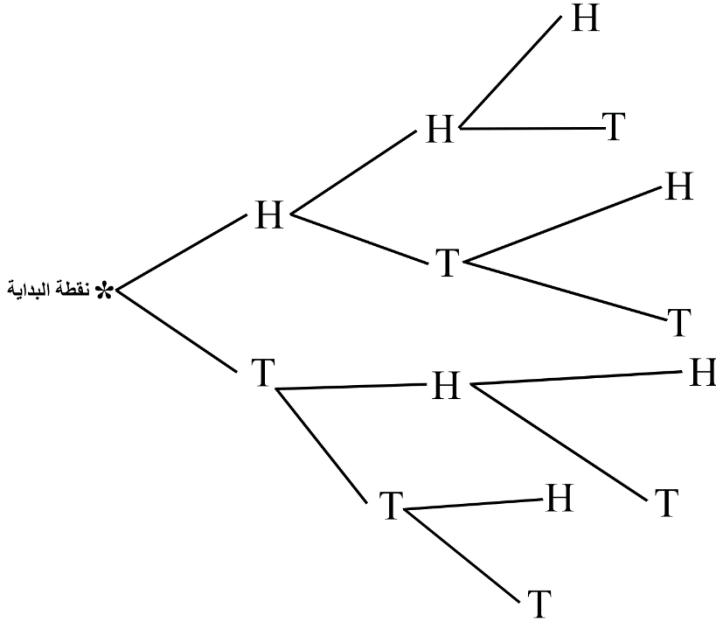
1- عند قولنا إلقاء قطعتي نقد متميزتين (مختلفتين) مرة واحدة فإن ذلك يكافئ قولنا إلقاء قطعة نقد مرتين.

2- في المثال السابق يكون الرمز الأول من اليسار في كل نتيجة هو ناتج الرمية الأولى والرمز الذي يليه في هو ناتج الرمية الثانية وهكذا.

مثال 3.6: اكتب فضاء العينة لتجربة إلقاء قطعة نقد ثلاث مرات.

الحل/

نرسم الشجرة البيانية مع ملاحظة أننا نرسم أولاً نتائج الرمية الأولى، ومن ثم نرسم نتائج الرمية الثانية عند جميع نواتج الرمية الأولى ونرسم أخيراً نواتج الرمية الثالثة عند جميع نواتج الرمية الثانية وذلك كما يلي:



ومن الرسم نجد أن:

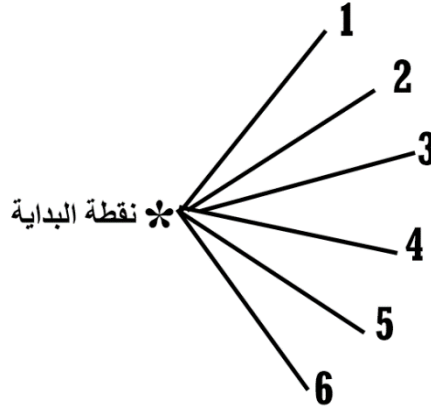
$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

ويكون عدد عناصر فضاء العينة: $\#\Omega = 8$.

مثال 4.6: اكتب فضاء العينة لتجربة إلقاء حجر نرد منتظم مرة واحدة.

الحل/

الشجرة البيانية للتجربة هي:



ومن الرسم نجد أن:

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

ويكون عدد عناصر فضاء العينة: $\#\Omega = 6$

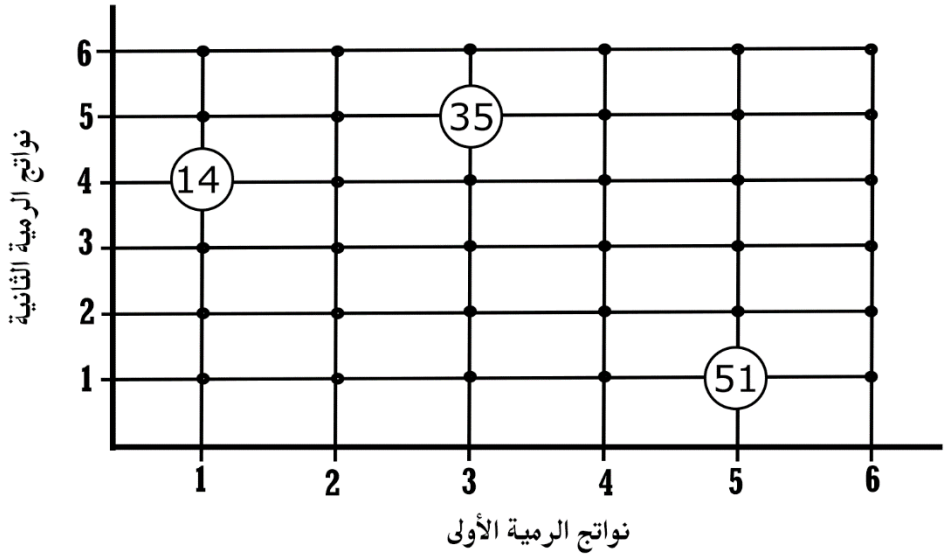
مثال 5.6: اكتب فضاء العينة لتجربة إلقاء حجر نرد مرتين.

الحل/

عند رسم الشجرة البيانية للتجربة نجد أنها كبيرة وسيصعب علينا التعامل معها، لذلك

سنلجأ في هذه الحالة لرسم الشبكة البيانية للتجربة العشوائية وذلك كما يلي:

نرسم خطين متعامدين أحدهما أفقي ونضع عليه نواتج الرمية الأولى والآخر عمودي ونضع عليه نواتج الرمية الثانية ومن عند كل نتيجة نرسم خط عمودي على المحور فتكون كل نقطة تقاطع تمثل أحد نواتج التجربة العشوائية كما هو مبين في الرسم التالي:



ويمكن الاستغناء بهذه الشبكة عن كتابة فضاء العينة. ويكون: $\#\Omega = 36$. لاحظ في الشبكة أن كل نقطة تمثل نتيجة ويظهر في الشبكة ثلاث نتائج معينة محاطة بدوائر.

ملاحظة: إذا أُجريت تجربة عشوائية n مرة، وكان عدد النواتج التي يمكن أن نحصل عليها في كل مرة متساوٍ وليكن m فإن:

$$\#\Omega = m^n$$

فمثلاً في تجربة إلقاء قطعة نقد خمس مرات يكون:

$$\#\Omega = 2^5 = 32$$

وفي تجربة إلقاء حجر نرد ثلاث مرات يكون:

$$\#\Omega = 6^3 = 216$$

تعرفنا فيما سبق على طريقة إيجاد فضاء العينة لبعض التجارب العشوائية باستخدام الشجرة والشبكة البيانية، والآن سنتعلم كيفية إيجاد حدث معين من فضاء العينة، حيث نصف عناصره بصفة مميزة تستخدم في انتقاء العناصر الخاصة به من فضاء العينة.

مثال 6.6: في تجربة إلقاء قطعة نقد ثلاث مرات أوجد الحوادث التالية:

- 1- الحصول على صورتين بالضبط.
- 2- الحصول على كتابة في الرمية الثالثة.
- 3- الحصول على صورة على الأكثر.
- 4- الحصول على كتابتين على الأقل.
- 5- الحصول على نفس النتيجة في الرميات الثلاث.

الحل/

نعطي كل حدث حرفاً كبيراً من أحرف اللغة الإنجليزية ليبدل عليه. ونعلم من مثال (3.6) أن فضاء العينة للتجربة العشوائية هو:

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

وبالتالي:

1- A : حدث الحصول على صورتين بالضبط.

نختار من فضاء العينة كل النواتج التي فيها صورتين بالضبط دون زيادة او نقصان فيكون الحدث:

$$A = \{HHT, HTH, THH, \}$$

2- B : حدث الحصول على كتابة في الرمية الثالثة.

نختار من فضاء العينة كل النواتج التي فيها كتابة (T) في الرمية الثالثة بغض النظر عن ناتج الرمية الأولى وناتج الرمية الثانية فيكون الحدث:

$$B = \{HHT, HTT, THT, TTT\}$$

3- C : حدث الحصول على صورة على الأكثر.

نختار من فضاء العينة كل النواتج التي تحتوي على صورة (H) واحدة فأقل، أي النواتج التي يكون عدد الصور فيها يساوي 1 أو صفر، فيكون الحدث:

$$C = \{HTT, THT, TTH, TTT\}$$

4- D : حدث الحصول على كتابتين على الأقل.

نختار من فضاء العينة النواتج التي يكون فيها عدد الكتابات أكبر من أو يساوي 2 أي عدد كتابات يساوي 2 أو 3 فيكون الحدث:

$$D = \{HTT, THT, TTH, TTT\}$$

5- E : حدث الحصول على نفس النتيجة في الرميات الثلاث.

واضح أن الحدث هو:

$$E = \{HHH, TTT\}$$

ملاحظات:

- 1- تعتبر Ω نفسها حدث حيث أن أي مجموعة تعتبر مجموعة جزئية من نفسها.
- 2- تعتبر المجموعة الخالية $\emptyset$ حدث وذلك لأنها تعتبر مجموعة جزئية من أي مجموعة.

وقوع الحدث:

يقال لحدث ما أنه وقع إذا كان ناتج التجربة العشوائية هو أحد عناصره. وبناءً عليه سنطلق على فضاء العينة Ω اسم **الحدث المؤكد** حيث أنه من المؤكد وقوع ناتج التجربة العشوائية فيه مهما كان، وكذلك سيطلق على المجموعة الخالية $\emptyset$ اسم **الحدث المستحيل** وذلك لاستحالة وقوع ناتج التجربة العشوائية فيه. والمثال التالي يوضح معنى الإجابة على هذا التساؤل.

مثال 7.6: في مثال (6.6) بفرض أننا أجرينا التجربة العشوائية لإلقاء قطعة نقد ثلاث مرات وحصلنا على النتيجة " HTH " فأی الحوادث السابقة الذكر في المثال نقول أنها وقعت.

الحل/

واضح أن الحدث Ω وكذلك الحدث A قد وقعا وذلك لوجود النتيجة "HTH" فيهما، أما باقي الحوادث فإنها لم تقع.

ملاحظات:

1- إذا كان الحدث مكوناً من عنصر واحد فإنه يسمى حدثاً بسيطاً Simple event.

2- إذا كان الحدث يحتوي على أكثر من عنصر فإنه يسمى حدثاً مركباً

Compound event.

الأحداث المتنافية The Disjoint events

تعريف

يقال لحدثين أنهما متنافيان إذا كان وقوع أحدهما ينفي وقوع الآخر، وبصورة أخرى:

يكون الحدثان A, B متنافيين إذا وفقط إذا كان $A \cap B = \emptyset$

مثال 8.6: في تجربة إلقاء حجر نرد مرة واحدة إذا كان:

$$E_1 = \{2,4\}, \quad E_2 = \{1,3,5\}, \quad E_3 = \{1,2,3\}$$

ثلاثة أحداث بين أي منها متنافية.

نجري عملية التقاطع بين الأحداث لنجد أن:

$$E_1 \cap E_2 = \emptyset, \quad E_1 \cap E_3 = \{2\}, \quad E_3 \cap E_2 = \{1,3\}$$

ومن نتائج التقاطع نلاحظ أن الحدثين E_2 , E_1 متافيان.

الترار النسبي والاحتمال Relative Frequency and Probability

إذا كان n عدد مرات إجراء تجربة عشوائية، m عدد مرات الحصول على الحدث E فإن التكرار النسبي للحدث E (أو يسمى الاحتمال التجريبي للحدث E) يساوي النسبة $\frac{m}{n}$. وعندما تصبح n كبيرة جداً فإن التكرار النسبي يقترب من قيمة محددة سنرمز لها بالرمز $P(E)$ وتسمى الاحتمال النظري للحدث E .

الفضاء المنتظم Regular Sample Space

هو فضاء العينة الذي تتساوى فيه احتمالات كل حوادثه البسيطة، وفي هذه الحالة يكون الاحتمال النظري للحدث E يساوي:

$$P(E) = \frac{\#(E)}{\#(\Omega)}$$

والمثال التالي يوضح كيفية حساب احتمال وقوع حدث:

مثال 9.6: في مثال (6.6) أوجد احتمال وقوع الحوادث A, B, C, D, E .

الحل/

لاحظ أن $\#(\Omega) = 8$ ، $\#(A) = 3$ لذلك يكون: $P(A) = \frac{3}{8}$.

وحيث أن $\#(B) = 4$ فيكون: $P(B) = \frac{4}{8}$.

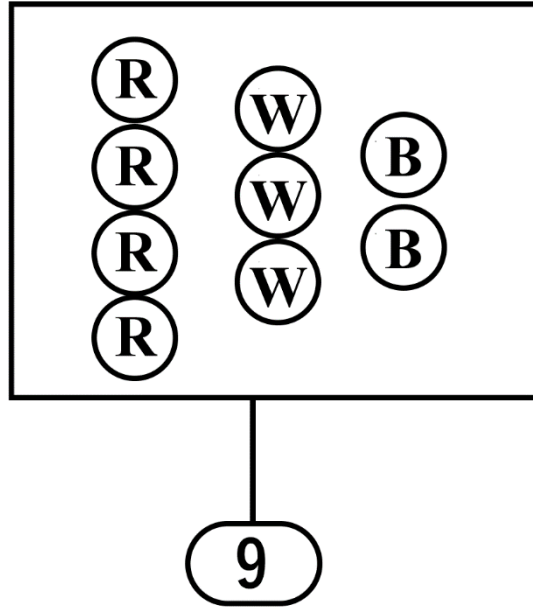
وبنفس الطريقة يمكن حساب باقي الاحتمالات لنجد أن:

$$P(C) = P(D) = \frac{4}{8} , \quad P(D) = \frac{4}{8} \quad P(E) = \frac{2}{8}$$

مثال 10.6: صندوق يحتوي على 4 كرات حمراء (R) وثلاث كرات بيضاء (W) وكرتين سوداوين (B). إذا سحبنا كرة عشوائياً احسب احتمال أن تكون الكرة المسحوبة بيضاء، حمراء، سوداء.

الحل/

نرسم الشكل التالي ليعبر عن الصندوق وبه الكرات وليدل على فضاء العينة:



كل كرة من الكرات تمثل ناتجاً ممكن الحدوث عند سحب كرة من الصندوق عشوائياً، وبالتالي يكون عدد عناصر فضاء العينة يساوي مجموع الكرات في الصندوق، وكل مجموعة من الكرات تمثل لوناً معيناً ننظر إليها كحدث الحصول على كرة من هذا اللون وبالتالي لدينا ثلاثة حوادث، حدث الحصول على كرة لونها أحمر وليكن R وعدد نواتجه 4 وهي عدد الكرات الحمراء، وحدث الحصول على كرة لونها

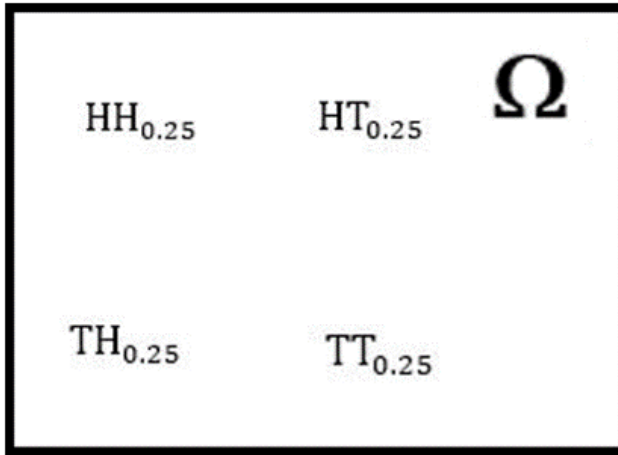
أبيض وليكن W وعدد نواتجه 3 وهو عدد الكرات البيضاء، وحدث الحصول على كرة لونها أسود وليكن B وعدد نواتجه 2 وهو عدد الكرات السوداء.

وبالتالي تكون الاحتمالات المطلوبة هي:

$$P(R) = \frac{4}{9}, \quad P(W) = \frac{3}{9} = \frac{1}{3}, \quad P(B) = \frac{2}{9}$$

ملاحظة: بصفة عامة نستطيع القول أن احتمال الحدث يساوي مجموع احتمالات حوادثه البسيطة.

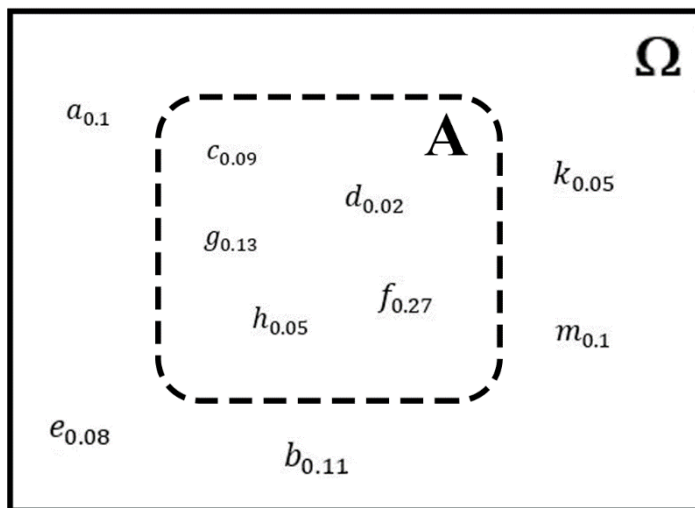
في بعض الأحيان يُعرض فضاء العينة باستخدام أشكال فن (Ven)، وهي أشكال مفيدة في فهم العلاقة بين الحوادث وفي فهم قوانين الاحتمال، فمثلاً يمكن عرض فضاء العينة لتجربة إلقاء قطعة نقد مرتين كما يلي:



حيث تكتب جميع عناصر فضاء العينة داخل المربع والذي يعبر عن فضاء العينة، وفي حالة معرفة احتمالات الحوادث البسيطة نكتبها كدليل سفلي لكل حدث بسيط حيث

يساعد هذا في حساب احتمالات الحوادث المركبة المختلفة. وإذا تساوت هذه الاحتمالات سُمِّيَ الفضاء منتظماً وإلا أطلق عليه اسم فضاء غير منتظم.

مثال 11.6: شكل فن التالي يمثل فضاء عينة لتجربة عشوائية واحتمالات الحوادث البسيطة، والحدث A من هذا الفضاء.



إن احتمال الحدث A يساوي مجموع احتمالات الحوادث البسيطة المكونة له وبالتالي يكون:

$$A = \{c, f, d, g, h\}$$

$$P(A) = 0.09 + 0.27 + 0.02 + 0.13 + 0.05 = 0.56$$

مسلمات الاحتمال Axioms of Probability

1- لأي حدث $A \subset \Omega$ يكون : $P(A) > 0$

2- $P(\emptyset) = 0$ ، $P(\Omega) = 1$ وعليه فإن : $0 \leq P(A) \leq 1$

3- لأي حدثين متنافيين A, B يكون:

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

متعم الحدث The Complement of the Event

إذا كان E حدثاً في فضاء العينة Ω ، فإن متعم الحدث E هو حدث يعبر عن عدم وقوع الحدث E ويرمز له بالرمز $\bar{E}$.

ونفهم من هذا أن متعم الحدث E هو الحدث الذي يحتوي على كل النواتج غير الموجودة في E .

من التعريف نستنتج أن الحدث ومتعمه حدثان متنافيان واتحادهما هو فضاء العينة بمعنى:

$$E \cap \bar{E} = \emptyset , \quad E \cup \bar{E} = \Omega$$

لاحظ أيضاً من مسلمات الاحتمال أن:

$$P(E \cup \bar{E}) = P(\Omega) = 1$$

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

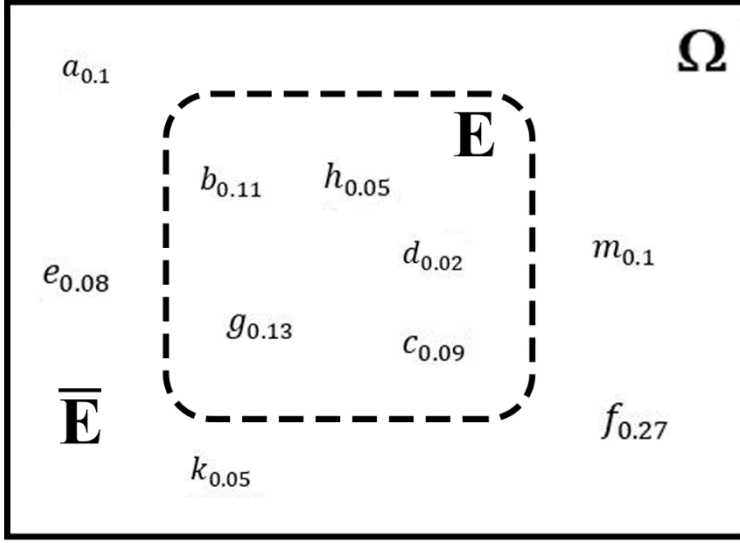
وبالتالي:

احتمال متعم الحدث

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



مثال 12.6: مستخدماً شكل فن التالي أوجد : $\bar{E}$, $P(E)$, $P(\bar{E})$



في شكل فن المبين أعلاه نجد أن الحدث E ومتممته $\bar{E}$ هما:

$$\bar{E} = \{a, e, f, k, m\} , \quad E = \{b, c, d, g, h\}$$

وبالتالي يكون احتمال كل من الحدث E ومتممته على الترتيب تساوي:

$$P(E) = 0.09 + 0.11 + 0.02 + 0.13 + 0.05 = 0.4$$

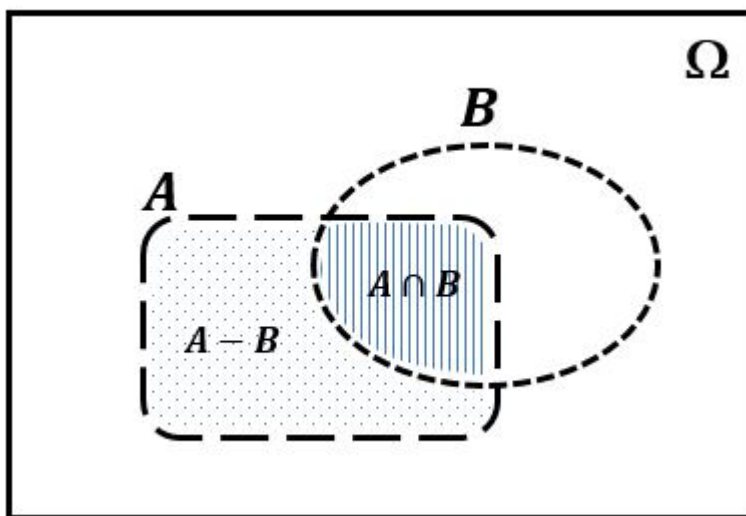
$$P(\bar{E}) = 0.1 + 0.08 + 0.27 + 0.05 + 0.1 = 0.6$$

ومن الممكن حساب احتمال $\bar{E}$ بمعلومية احتمال الحدث E كما يلي:

$$P(\bar{E}) = 1 - P(E) = 1 - 0.4 = 0.6$$

الفرق بين حدثين The Difference Between Two Events

بفرض أن A ، B حدثان من فضاء عينة Ω ، نُعرِّف الحدث $A - B$ "ونقرأه (A فرق B)" بأنه الحدث الذي يحوي الحوادث البسيطة الموجودة في A وغير موجودة في B أو هو مجموعة عناصر A غير الموجودة في B . ومن شكل فن التالي سنستنتج قانون ايجاد احتمال الفرق بين حدثين:



لاحظ أن الحدثين $A \cap B$ ، $A - B$ هما حدثان متتافيان ولاحظ أن:

$$A = (A - B) \cup (A \cap B)$$

ومن مسلمات الاحتمال نجد أن:

$$P(A) = P[(A - B) \cup (A \cap B)] = P(A - B) + P(A \cap B)$$

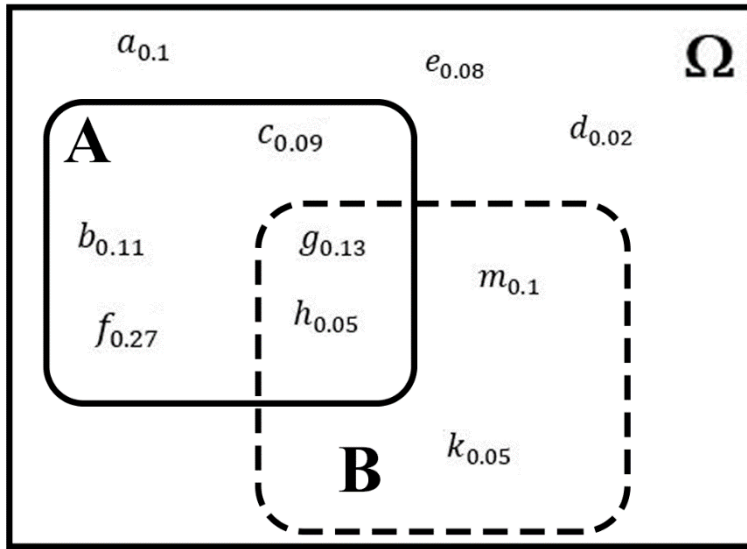
وبالتالي:

احتمال الفرق بين حدثين

$$P(A - B) = P(A) - P(A \cap B)$$



مثال 13.6: مستعيناً بالشكل التالي أوجد $P(A - B)$.



الحل/

من شكل فن المعطى نجد أن:

$$A - B = \{c, b, f\}$$

$$P(A - B) = 0.11 + 0.27 + 0.09 = 0.47$$

وبطريقة أخرى نجد أن:

$$A \cap B = \{g, h\} \quad , \quad A = \{f, b, c, h, g\}$$

$$P(A) = 0.11 + 0.27 + 0.09 + 0.05 + 0.13 = 0.65$$

$$P(A \cap B) = 0.05 + 0.13 = 0.18$$

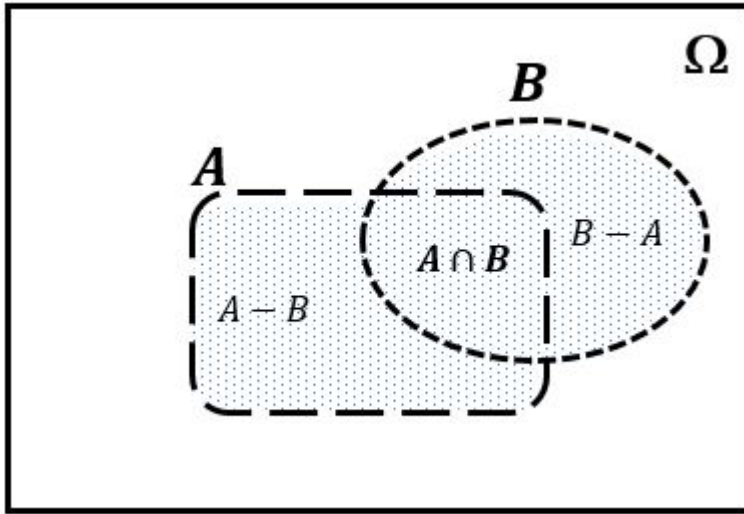
وبالتالي:

$$P(A - B) = P(A) - P(A \cap B) = 0.65 - 0.18 = 0.47$$

اتحاد حدثين The Union of Two Events

بفرض أن A و B حدثان من فضاء عينة Ω . نعرف الحدث $A \cup B$ "ونقرأه $A \cup B$ "
اتحاد B " بأنه الحدث الذي يحوي الحوادث البسيطة الموجودة في A وفي B .

ويمكن تمثيل $(A \cup B)$ بشكل فن كما يلي:



لاحظ أن الحدثين $A - B$, B حدثان متنافيان وأن:

$$A \cup B = (A - B) \cup B$$

ومن مسلمات الاحتمال نجد أن:

$$P(A \cup B) = P(A - B) + P(B)$$

وبالتالي:

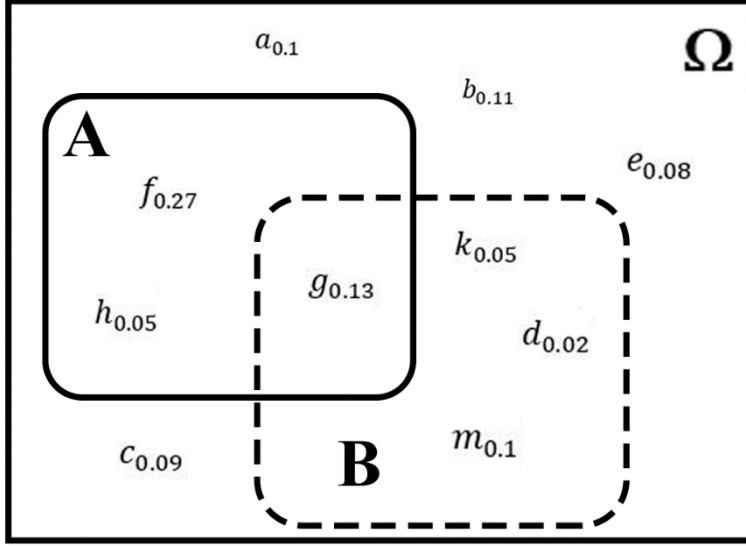
احتمال اتحاد حدثين

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



لا تنسى

مثال 14.6: مستعيناً بالشكل التالي أوجد $P(A \cup B)$.



الحل/

مباشرة من الشكل نجد أن:

$$A \cup B = \{f, h, g, k, m, d\}$$

$$P(A \cup B) = 0.27 + 0.05 + 0.13 + 0.05 + 0.1 + 0.02 = 0.62$$

وبطريقة ثانية نجد أن:

$$A \cap B = \{g\} \quad , \quad B = \{g, k, m, d\} \quad , \quad A = \{f, h, g\}$$

$$P(A) = 0.27 + 0.05 + 0.13 = 0.45$$

$$P(B) = 0.13 + 0.05 + 0.1 + 0.02 = 0.3$$

$$P(A \cap B) = 0.13$$

وبالتالي فإن:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.45 + 0.3 - 0.13 = 0.62$$

ملاحظة:

هناك بعض العلاقات الهامة والمفيدة في حساب الاحتمالات وسنذكر بعضها بدون برهان.

لأي حدثين A, B يكون:

$$A \cap \bar{B} = A - B \quad (1)$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \quad (2)$$

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad (3)$$

حيث تسمى العلاقات (2) و (3) بقوانين دي مورجان.

مثال 15.6 : إذا كان $P(A \cap B) = 0.35$, $P(B) = 0.8$, $P(A) = 0.4$

أوجد:

$$P(A \cup B) \quad (1)$$

$$P(B - A) \quad (2)$$

$$P(\bar{A} - B) \quad (3)$$

الحل/

بالتطبيق المباشر لقوانين الاحتمال نجد أن:

$$1- P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.8 - 0.35 = 0.85$$

$$2- P(B - A) = P(B) - P(B \cap A) = 0.8 - 0.35 = 0.45$$

$$3- P(\bar{A} - B) = P(\bar{A}) - P(\bar{A} \cap B)$$

$$= 1 - P(A) - P(B - A) = 1 - 0.4 - 0.45 = 0.15$$

أحياناً يتم التعبير عن الأحداث البسيطة بجمل خبرية ويُرَكَّب منها أحداثاً مركبة باستخدام بعض أدوات الربط مثل "و" والتي تعني في نظرية المجموعات "∩"، و أداة الربط "أو" والتي تعني في نظرية المجموعات "∪"، وأحياناً ننفي حدثاً باستخدام أداة نفي أو بالنقيض اللغوي حيث يقابلها في نظرية المجموعات المتممة.

وهناك تعبيرات لغوية أخرى تستخدم في ربط الأحداث البسيطة مثل (فقط - على الأقل - على الأكثر - معاً) والمثال التالي يوضح كيف نتعامل مع مثل هذه الأحداث.

مثال 16.6: إذا كانت نسبة الأشخاص الذين عيونهم سوداء في مدينة عمان 70% وكانت نسبة الأشخاص الذين شعرهم أسود 40% ونسبة الأشخاص الذين عيونهم سوداء وشعرهم أسود 30% فإذا اخترنا شخصاً عشوائياً من مدينة عمان، فأوجد احتمال الحوادث التالية:

1- أن تكون عيونه سوداء أو شعره أسود.

2- ألا يكون ذو عيون سوداء.

3- أن يكون ذو شعر أسود وعيونه ليست سوداء.

الحل/ نفرض أن الحدث A يعبر عن "الشخص عيونه سوداء"، وأن الحدث B يعبر عن "الشخص ذو شعر أسود" لذلك يكون لدينا من معطيات السؤال الاحتمالات التالية:

$$P(A \cap B) = 0.3, \quad P(B) = 0.4, \quad P(A) = 0.7$$

1- نعبّر أولاً عن الحدث "عيونه سوداء أو شعره أسود" بدلالة الحدثين A و B بالتعبير

$A \cup B$ حيث يُعبّر عن كلمة "أو" برمز الاتحاد $\cup$ وبالتالي يكون:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.7 + 0.4 - 0.3 = 0.8$$

2- نعبر عن الحدث " عيونه ليست سوداء " بالحدث $\bar{A}$ ، حيث نعبر دائماً عن نفي الحدث بأداة نفي أو بالنقيض اللغوي بمتمة هذا الحدث، (كأن نقول الحدث "نجح" متمته "لم ينجح" أو نقول "رسب")، وبالتالي يكون:

$$P(\bar{A}) = 1 - P(A) = 1 - 0.7 = 0.3$$

3- نعبر عن الحدث "ذو شعر أسود و عيونه ليست سوداء " بالتعبير $B \cap \bar{A}$ وبالتالي يكون:

$$P(B \cap \bar{A}) = P(B - A) = P(B) - P(B \cap A) = 0.4 - 0.3 = 0.1$$

مثال 17.6: صندوق يحوي 7 كرات سوداء (B)، 3 كرات بيضاء (W). سُحِبَت من الصندوق كرة واحدة عشوائياً أوجد:

1- احتمال أن تكون الكرة المسحوبة سوداء.

2- احتمال أن تكون الكرة المسحوبة بيضاء.

3- احتمال أن تكون الكرة المسحوبة حمراء (R).

الحل/ لاحظ أن مجموع الكرات الكلي في الصندوق يساوي 10 والذي يعبر عن عدد الحوادث البسيطة في فضاء العينة بمعنى $\#(\Omega) = 10$ ، ونعبر عن عدد الكرات البيضاء بالتعبير $\#(W) = 3$ وعن عدد الكرات السوداء بالتعبير $\#(B) = 7$ وبالتالي تكون احتمالات الحوادث المطلوبة كما يلي:

$$P(B) = \frac{\#(B)}{\#(\Omega)} = \frac{7}{10}$$

$$P(W) = \frac{\#(W)}{\#(\Omega)} = \frac{3}{10}$$

$$P(R) = 0$$

تعريف

إذا كان A, B حدثين في فضاء عينة Ω ، فإننا نقول أن الحدثين مستقلان إذا كان وقوع أحدهما لا يؤثر في وقوع الآخر. وبصيغة رياضية إذا تحقق الشرط:

$$P(A \cap B) = P(A) \times P(B)$$

مثال 18.6: إذا كان A, B حدثين مستقلين في فضاء عينة Ω وكان:

$$P(A) = 0.6 \quad , \quad P(B) = 0.5$$

فأوجد: $P(A \cup B)$, $P(A \cap B)$

الحل/

من استقلالية الحدثين نجد أن:

$$P(A \cap B) = P(A) \times P(B) = 0.6 \times 0.5 = 0.3$$

وبالتالي:

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.6 + 0.5 - 0.3 = 0.8 \end{aligned}$$

مثال 19.6: أطلق صيادان سهمين معاً نحو هدف ما، بحيث أطلق كل واحد منهما سهماً. فإذا كان احتمال إصابة الصياد الأول للهدف يساوي 0.8 واحتمال إصابة الصياد الثاني للهدف يساوي 0.6 فأوجد الاحتمالات التالية:

1- إصابة الصيادان للهدف معاً.

2- إصابة الهدف.

/الحل/

نفرض أن حدث "الصياد الأول أصاب الهدف" هو الحدث A ، وأن حدث "الصياد الثاني أصاب الهدف" هو الحدث B ، فتكون احتمالات هذين الحدثين من معطيات السؤال هي

$$P(A) = 0.8 , P(B) = 0.6$$

1- نفهم من صيغة المطلوب الأول أن "الصياد الأول أصاب الهدف و الصياد الثاني

أصاب الهدف" ونعبر عن ذلك بالرموز بالتعبير $A \cap B$ ، فيكون الاحتمال المطلوب:

$$P(A \cap B) = P(A) \times P(B) = 0.8 \times 0.6 = 0.48$$

2- نفهم من صيغة المطلوب الثاني أن المطلوب احتمال اصابة الهدف بغض

النظر الصياد الذي أصابه، لذلك نستطيع القول أننا نعتبر أن الهدف أصيب إذا

"أصاب الصياد الأول الهدف أو أصاب الصياد الثاني الهدف" ونعبر عن ذلك بالتعبير

$A \cup B$ فيكون:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= 0.8 + 0.6 - 0.48 = 0.92$$

نذكر النظرية التالية بدون برهان والتي توضح أن استقلالية حدثين صفة تورث لمتغيرات هذين الحدثين.

إذا كان A, B حدثين مستقلين من فضاء عينة Ω فإن:

$P(A \cap \bar{B}) = P(A) \times P(\bar{B})$ أي $A, \bar{B}$ مستقلان ،

$P(\bar{A} \cap B) = P(\bar{A}) \times P(B)$ أي $\bar{A}, B$ مستقلان ،

$P(\bar{A} \cap \bar{B}) = P(\bar{A}) \times P(\bar{B})$ أي $\bar{A}, \bar{B}$ مستقلان ،



مثال 20.6: إذا كان A, B حدثين مستقلين في فضاء عينة Ω وكان:

$$P(A) = 0.3 , P(B) = 0.4$$

فأوجد: $P(\bar{A} \cup B)$ ، $P(\bar{A} \cap B)$ ، $P(\bar{A} \cap \bar{B})$ ، $P(A \cap B)$

الحل/ نحسب في البداية احتمالات متممات الحدثين.

$$P(\bar{A}) = 1 - P(A) = 1 - 0.3 = 0.7$$

$$P(\bar{B}) = 1 - P(B) = 1 - 0.4 = 0.6$$

وبالتطبيق المباشر للنظرية نجد أن:

$$P(A \cap B) = P(A)P(B) = 0.3 \times 0.4 = 0.12$$

$$P(\bar{A} \cap \bar{B}) = P(\bar{A})P(\bar{B}) = 0.7 \times 0.6 = 0.42$$

$$P(\bar{A} \cap B) = P(\bar{A})P(B) = 0.7 \times 0.4 = 0.28$$

$$P(\bar{A} \cup B) = P(\bar{A}) + P(B) - P(\bar{A} \cap B) = 0.7 + 0.4 - 0.28 = 0.82$$

مثال 21.6: صندوق يحوي 6 كرات سوداء (B) و 4 كرات حمراء (R). سحب من

الصندوق كرتان على التوالي مع الإرجاع. أوجد احتمال:

1- أن تكون الكرتان سوداوين.

2- أن تكون الكرة الأولى سوداء والثانية حمراء.

3- أن تكون إحدى الكرتين سوداء.

4- أن تكون الكرتان من نفس اللون.

الحل/

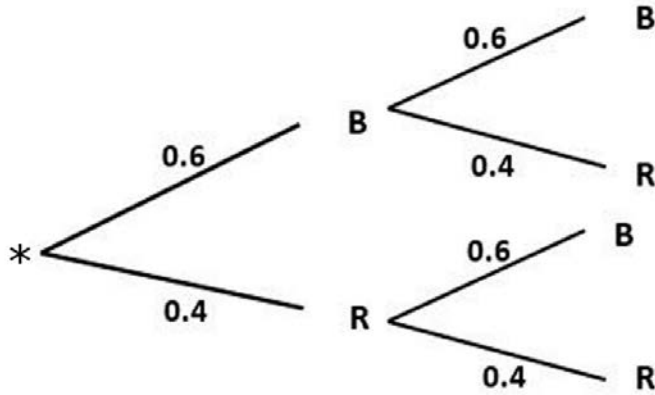
يتميز السحب مع الإرجاع بأنه يحافظ على عدد الكرات لكل لون وعدد الكرات الكلي في الصندوق، وبالتالي لا يختلف احتمال سحب أي كرة ومن أي لون في أي سحبة.

ويفضل في مثل هذه الحالة رسم الشجرة البيانية للتجربة العشوائية وكتابة الاحتمالات عليها لنحصل على الشجرة الاحتمالية ثم نستغل استقلالية الحوادث في إيجاد الاحتمالات المطلوبة حيث أن ناتج كل سحبة مستقل عن ناتج السحبة التي تليها وذلك كما يلي:

لاحظ أن مجموع الكرات الكلي في الصندوق يساوي 10 والذي يعبر عن عدد الحوادث البسيطة في فضاء العينة بمعنى $\#(\Omega) = 10$ ، ونعبر عن عدد الكرات الحمراء بالتعبير $\#(R) = 4$ وعن عدد الكرات السوداء بالتعبير $\#(B) = 6$ وبالتالي تكون احتمالات سحب كرة من كل لون كما يلي:

$$P(B) = \frac{\#(B)}{\#(\Omega)} = \frac{6}{10} , \quad P(R) = \frac{\#(R)}{\#(\Omega)} = \frac{4}{10}$$

وتكون الشجرة الاحتمالية كما يلي:



وتكون الاحتمالات المطلوبة:

$$1- P(BB) = P(B)P(B) = 0.6 \times 0.6 = 0.36$$

$$2- P(B, R) = P(B)P(R) = 0.6 \times 0.4 = 0.24$$

$$\begin{aligned} 3- P((B, R) \cup (R, B)) &= P(B, R) + P(R, B) = P(B)P(R) + P(R)P(B) \\ &= 0.6 \times 0.4 + 0.4 \times 0.6 = 0.48 \end{aligned}$$

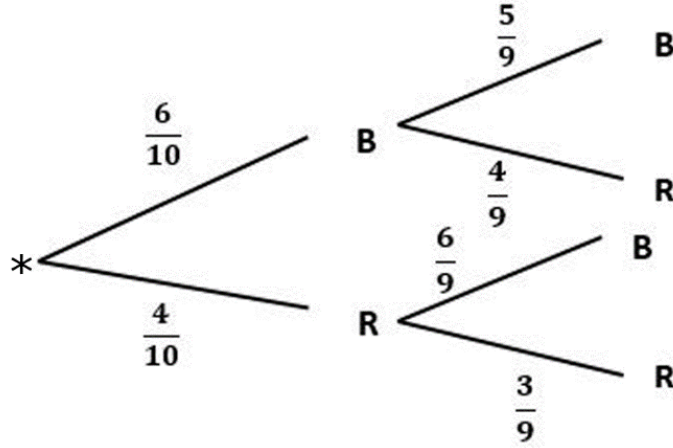
$$\begin{aligned} 4- P((B, B) \cup (R, R)) &= P(B, B) + P(R, R) = P(B)P(B) + P(R)P(R) \\ &= 0.6 \times 0.6 + 0.4 \times 0.4 = 0.52 \end{aligned}$$

ملاحظة:

إذا كان السحب بدون إرجاع فإن عدد الكرات الكلي في الصندوق ينقص بمقدار 1 في كل سحبة وكذلك اللون الذي سحبت منه الكرة ينقص بمقدار 1 وبالتالي فإن الاحتمالات على الشجرة الاحتمالية تختلف.

مثال 22.6: أعد حل مثال (21.6) في حالة السحب بدون إرجاع.

الحل/ في هذه الحالة تكون الشجرة الاحتمالية كما يلي:



حيث نلاحظ اختلاف الاحتمالات لكافة الالوان في السحبة الثانية، فقد تم حسابها على اعتبار أن عدد الكرات في الصندوق يساوي 9 وذلك لعدم ارجاع الكرة المسحوبة في السحبة الأولى. وتكون الاحتمالات المطلوبة:

$$1- P(BB) = P(B)P(B) = \frac{6}{10} \times \frac{5}{9} = \frac{30}{90}$$

$$2- P(B, R) = P(B)P(R) = \frac{6}{10} \times \frac{4}{9} = \frac{24}{90}$$

$$\begin{aligned} 3- P((B, R) \cup (R, B)) &= P(B, R) + P(R, B) \\ &= P(B)P(R) + P(R)P(B) \\ &= \frac{6}{10} \times \frac{4}{9} + \frac{4}{10} \times \frac{6}{9} = \frac{48}{90} \end{aligned}$$

$$\begin{aligned} 4- P((B, B) \cup (R, R)) &= P(B, B) + P(R, R) \\ &= P(B)P(B) + P(R)P(R) \\ &= \frac{6}{10} \times \frac{5}{9} + \frac{4}{10} \times \frac{3}{9} = \frac{42}{90} \end{aligned}$$

تعريف

إذا كان A, B حدثين من فضاء عينة Ω فإن احتمال حدوث A بشرط حدوث B يحسب من العلاقة:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

والجدير بالذكر هنا أن $P(A/B)$ يختلف عن $P(A)$ وذلك لأن الاحتمال المشروط يتضمن توفر معلومات قبل حساب الاحتمال المطلوب مما يقلص فضاء العينة.

مثال 23.6: إذا كان $P(A) = 0.7$ ، $P(B) = 0.8$ ، $P(A \cap B) = 0.6$

أوجد: $P(A/B)$ ، $P(\bar{A}/B)$

الحل/

بالتطبيق المباشر لقانون الاحتمال المشروط نجد أن:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.6}{0.8} = \frac{3}{4}$$

$$P(\bar{A}/B) = \frac{P(\bar{A} \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)} = \frac{0.8 - 0.6}{0.8} = \frac{1}{4}$$

قاعدة الضرب للاحتمال المشروط

نعلم أن:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

وبإجراء المساواة لحواصل ضرب الطرفين والوسطيين في العلاقتين نحصل على:

$$\begin{aligned} P(A \cap B) &= P(B) \times P(A/B) \\ &= P(A) \times P(B/A) \end{aligned}$$

وهما تمثلان قاعدة الضرب للاحتمال المشروط، حيث تُستَخدم هذه القاعدة لحساب احتمال وقوع حدثين معاً إذ علم احتمال وقوع أحدهما واحتمال وقوع الآخر بشرط وقوع الأول.

مثال 24.6: إذا كان $P(B) = 0.45$, $P(A/B) = 0.3$ أوجد:

$$P(A \cap B) - 1$$

$$P(\bar{A} \cup \bar{B}) - 2$$

الحل/ بالتطبيق المباشر لقاعدة الضرب نجد أن:

$$P(A \cap B) = P(B) \times P(A/B) = 0.3 \times 0.45 = 0.135$$

وبتطبيق قانون ديمو رجان نحصل على الاحتمال الثاني حيث:

$$P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - 0.135 = 0.865$$

إذا كان A, B حدثين مستقلين فإن:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$$

أي أن للحوادث المستقلة A, B يكون:

$$P(A/B) = P(A) \quad , \quad P(B/A) = P(B)$$

مثال 25.6 : صندوق يحوي 6 كرات بيضاء (W) و 4 كرات سوداء (B)، سحب من

الصندوق كرتان على التوالي دون إرجاع احسب.

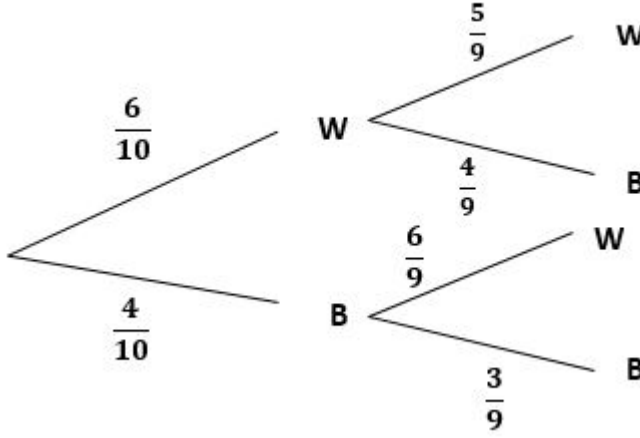
1- احتمال أن تكون الكرة الثانية سوداء إذا كانت الكرة الأولى بيضاء.

2- احتمال أن تكون الكرة الأولى بيضاء والكرة الثانية سوداء.

3- احتمال أن تكون الكرة الثانية بيضاء إذا عُلِمَ أن الكرة الأولى سوداء.

4- احتمال أن تكون الكرتان مختلفتين في اللون.

الحل/ الشجرة الاحتمالية للتجربة هي:



في المطلوب الأول المطلوب حساب احتمال أن تكون الكرة الثانية سوداء إذا عُلِمَ (بشرط) أن الكرة الأولى بيضاء، وطالما أن هناك شرط يجب تحقيقه لحساب الاحتمال، يكون المطلوب هو احتمال شرطي.

لإيجاد المطلوب الأول نعلم من الشجرة البيانية أن:

$$P(WB) = \frac{6}{10} \times \frac{4}{9} = \frac{24}{90}$$

ومباشرة من قانون الاحتمال المشروط نجد أن:

$$P\left(\frac{\text{الكرة الثانية B}}{\text{الأولى بيضاء W}}\right) = \frac{P(WB)}{P(W)} = \frac{\frac{24}{90}}{\frac{6}{10}} = \frac{4}{9}$$

حيث: $WB = \text{"الكرة الأولى بيضاء"} \cap \text{"الكرة الثانية سوداء"}$

لإيجاد المطلوب الثاني يمكننا استخدام الشجرة الاحتمالية مباشرة لنجد أن:

$$P(WB) = \frac{6}{10} \times \frac{4}{9} = \frac{24}{90}$$

والمطلوب الثالث يأتي مباشرة من الشجرة الاحتمالية أيضاً، ويمكننا تطبيق قانون الاحتمال الشرطي كالتالي:

$$P(WB) = \frac{6}{10} \times \frac{4}{9} = \frac{24}{90}$$

$$P\left(\frac{W}{B}\right) = \frac{P(WB)}{P(B)} = \frac{\frac{24}{90}}{\frac{4}{10}} = \frac{6}{9}$$

ولإيجاد المطلوب الرابع نجد من الشجرة الاحتمالية أن:

$$P((WB) \cup (BW)) = P(WB) + P(BW)$$

$$= \frac{24}{90} + \frac{4}{10} \times \frac{6}{9} = \frac{8}{15}$$

مثال 26.6: إذا كان A, B حدثين من فضاء عينة Ω بحيث $P(A) = 0.7$ و

$$P(A/B) = 0.3 \quad , \quad P(B) = 0.4$$

أوجد : $P(A \cup B)$, $P(A \cap B)$

الحل/ مباشرة من قاعدة الضرب نجد أن:

$$P(A \cap B) = P(B) \times P(A/B) = 0.4 \times 0.3 = 0.12$$

وبتطبيق قانون احتمال اتحاد حدثين نجد أن:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.7 + 0.4 - 0.12 = 0.98$$

تجزئة فضاء العينة Partition the Sample Space

إذا كانت $E_1, E_2, \dots, E_n$ حوادث في Ω بحيث تحقق:

$$\Omega = E_1 \cup E_2 \cup \dots \cup E_n \quad (1)$$

(2) تتنافى الحوادث $E_1, E_2, \dots, E_n$ متشئ متشئ. أي: $E_i \cap E_j = \emptyset, \forall i \neq j$

فإن هذه الحوادث تسمى تجزئة لفضاء العينة.

مثال 27.6: في تجربة رمي حجر نرد مرة واحدة، إذا كان:

$$E_1 = \{1,6\}, \quad E_2 = \{2,4,5\}, \quad E_3 = \{3\}$$

هل تمثل هذه الحوادث تجزئة لفضاء العينة Ω ؟

الحل/

لاحظ أن فضاء العينة للتجربة المذكورة في السؤال هو:

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

$$E_1 \cup E_2 \cup E_3 = \{1,2,3,4,5,6\} = \Omega \quad \text{وكذلك}$$

وبإجراء التقاطعات بين الأحداث متشئ متشئ نجد أن:

$$E_1 \cap E_2 = \emptyset, \quad E_1 \cap E_3 = \emptyset, \quad E_3 \cap E_2 = \emptyset$$

وهذا يعني أن مجموعة الحوادث المذكورة تشكل تجزئة لفضاء Ω .

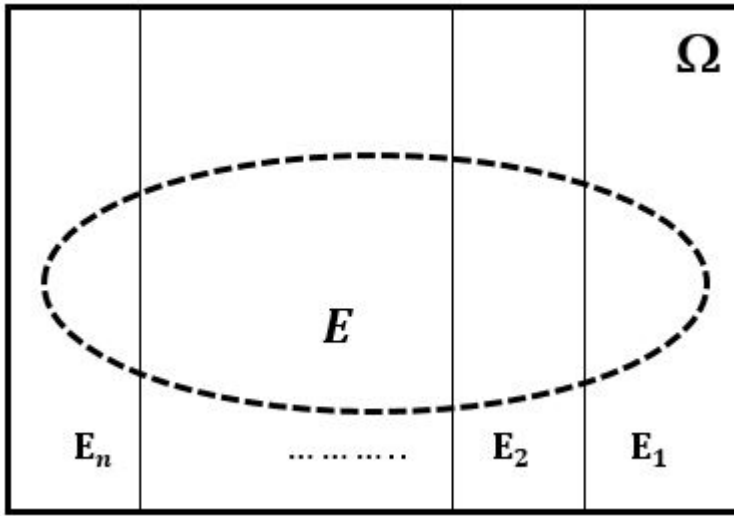
في بعض الأحيان يرتبط وقوع حدث معين بتجربتين عشوائيتين أو أكثر، وحساب الاحتمالات في هذه الحالة يتطلب معرفة طرق وقوانين مناسبة، وسنتحدث هنا عن حالة ارتباط الحدث بتجربتين عشوائيتين فقط وسنذكر النظرية التالية والتي تمدنا بطريقة لحساب الاحتمالات في هذه الحالات.

نظرية التجزئة

إذا كانت $E_1, E_2, \dots, E_n$ مجموعة من الحوادث تشكل تجزئة لفضاء عينة Ω ، وكان E حدث يمكن أن يقع مع أي حدث من هذه المجموعة فإن احتمال وقوع الحدث E يعطى بالعلاقة:

$$P(E) = P(E_1)P(E/E_1) + P(E_2)P(E/E_2) + \dots + P(E_n)P(E/E_n)$$

سنستخدم شكل فن التالي لاستنتاج العلاقة السابقة:



من الواضح أن:

$$E = (E \cap E_1) \cup (E \cap E_2) \cup \dots \cup (E \cap E_n)$$

وبحساب الاحتمال للطرفين وبملاحظة أن الطرف الأيمن هو اتحاد لحوادث متنافية نجد أن:

$$P(E) = P(E \cap E_1) + P(E \cap E_2) + \dots + P(E \cap E_n)$$

وبتطبيق قاعدة ضرب الاحتمالات نجد أن:

$$P(E) = P(E_1)P(E/E_1) + P(E_2)P(E/E_2) + \dots + P(E_n)P(E/E_n)$$

مثال 28.6: صندوق A يحتوي على 6 مصابيح صالحة و 4 مصابيح معيبة، و صندوق B يحتوي على 8 مصابيح صالحة و 2 معيبة. اختير أحد الصندوقين عشوائياً ثم سحب منه مصباح واحد عشوائياً. احسب احتمال أن يكون هذا المصباح صالحاً.

الحل/

لاحظ هنا اننا سنجري عمليتين، الأولى اختيار صندوق من الصندوقين، والثانية اختيار مصباح واحد من هذا الصندوق، وكذلك الحدث المشترك هو "المصباح صالح" وإذا رمزنا لهذا الحدث بالرمز E فيكون المطلوب ايجاد $P(E)$ حيث قانون نظرية التجزئة في هذه الحالة هو:

$$P(E) = P(A)P(E/A) + P(B)P(E/B)$$

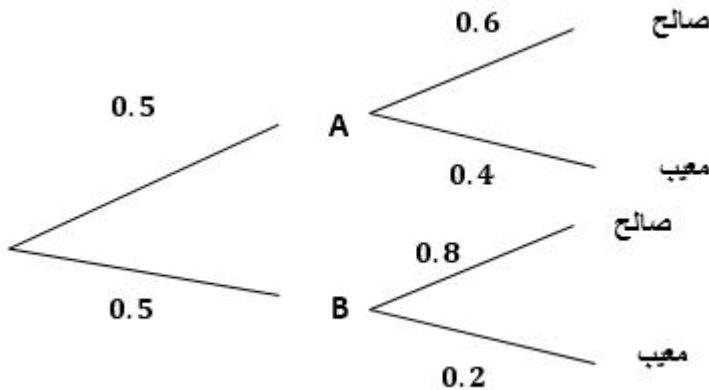
ومن معطيات المثال نجد أن:

$$P(E/B) = 0.8 , P(E/A) = 0.6 , P(B) = 0.5 , P(A) = 0.5$$

وبالتعويض المباشر في قانون نظرية التجزئة نحصل على الاحتمال المطلوب وذلك كما يلي:

$$P(E) = P(A)P(E/A) + P(B)P(E/B) = 0.5 \times 0.6 + 0.5 \times 0.8 = 0.7$$

ويمكن حل السؤال باستخدام الشجرة الاحتمالية وذلك كما يلي:



ويكون احتمال أن يكون المصباح صالحاً هو احتمال أن يكون صالحاً ومن الصندوق الأول +
احتمال أن يكون صالحاً ومن الصندوق الثاني أي:

$$0.5 \times 0.6 + 0.5 \times 0.8 = 0.7$$

في مثال (28.6): لو فرضنا أننا سحبنا مصباحاً عشوائياً من أحد الصندوقين ووجدنا أنه صالح، فالسؤال الذي يطرح نفسه هو: ما احتمال أن يكون المصباح من الصندوق الأول مثلاً؟ للإجابة على هذا السؤال نذكر نص النظرية الهامة التالية والتي تسمى نظرية بييز.

إذا كانت $E_1, E_2, \dots, E_n$ مجموعة من الحوادث تشكل تجزئة لفضاء عينة Ω ، بحيث $P(E_i) \neq 0$ لجميع قيم $\{i = 1, 2, \dots, n\}$ وكان E حدث يمكن أن يقع مع أي حدث من هذه المجموعة فإن احتمال وقوع E_i بشرط وقوع E يعطى بالعلاقة:

$$P(E_i/E) = \frac{P(E_i) \times P(E/E_i)}{P(E)}$$

نظرية
2.6



حيث

$$P(E) = P(E_1)P(E/E_1) + P(E_2)P(E/E_2) + \dots + P(E_n)P(E/E_n)$$

مثال 29.6: في مثال (28.6) بفرض أننا اخترنا صندوقاً عشوائياً ثم سحبنا منه كرة عشوائياً

ووجدنا أنها صالحة، احسب احتمال أن تكون هذه الكرة سُحِبَت من الصندوق B .

الحل/

من معطيات المثال المذكور وجدنا أن:

$$P(E/B) = 0.8 \quad , \quad P(E/A) = 0.6 \quad , \quad P(B) = 0.5 \quad , \quad P(A) = 0.5$$

وكنا قد حسبنا احتمال أن يكون المصباح صالحاً ووجدنا أنه يساوي $P(E) = 0.7$

والمطلوب في هذا المثال حساب $P(B/E)$ ، وبتطبيق نظرية بييز مباشرة نجد أن:

$$P(B/E) = \frac{P(B) \times P(E/B)}{P(E)} = \frac{0.5 \times 0.8}{0.7} = \frac{4}{7}$$

تمارين:

1- إذا كان $P(A) = 0.6$ و كان $2P(B) = 0$ ، أوجد احتمال وقوع (A أو B) اذا علمت انهما متنافيان؟

2- إذا كان A, B حدثين مستقلين، و كان $P(A) = 0.15$ و $P(B) = 0.4$ فأوجد:

$$P(A \cap B) , P(A \cup B) , P(\bar{A} \cap \bar{B}) , P(\bar{A} \cup \bar{B})$$

3- كيس يحتوي 9 كرات سوداء، 6 كرات حمراء، سحب من الكيس كرة واحدة عشوائياً، احسب:

- احتمال أن تكون حمراء.
- احتمال أن تكون سوداء.
- احتمال أن تكون بيضاء.

4- كيس يحوي 9 كرات بيضاء، 11 كرة حمراء. سحب من الكيس كرتان على التوالي دون ارجاع عشوائياً، احسب احتمال أن تكون:

- الكرتان حمراوين.
- الكرتان مختلفتين في اللون.

5- تقدم موظفان لأحد البنوك. إذا كان احتمال قبول الأول 0.7 واحتمال قبول الثاني 0.6 احسب احتمال:

- قبول الاثنين معاً.
- قبول الأول أو الثاني.
- قبول الأول وعدم قبول الثاني.
- عدم قبول الاثنين.
- عدم قبول الاثنين معاً.

6- صندوق يحوي 9 كرات حمراء، 6 كرات سوداء، 5 كرات بيضاء، سحب من الكيس

كرتان على التوالي بشكل عشوائي مع الإرجاع، ما احتمال:

- أن تكون الكرة الأولى حمراء والثانية بيضاء.

- أن تكون الكرتان من نفس اللون.

- أن تكون الكرتان مختلفتين في اللون.

- أن تكون إحدى الكرتين ليست سوداء.

7- أعد حل السؤال السادس إذا كان السحب بدون إرجاع.

8- تقدم طالب لامتحانين في الرياضيات والعلوم. فإذا كان احتمال نجاحه في الرياضيات

0.7 واحتمال نجاحه في العلوم إذا نجح في الرياضيات 0.8 احسب احتمال نجاحه في

المادتين معاً.

9- إذا كان $P(A) = 0.65$ ، $P(B) = 0.8$ ، $P(A \cap B) = 0.55$

احسب:

$$P(\bar{A} / \bar{B}) , P(A / A \cup B) , P(\bar{A} / B) , P(B / A) , P(A / B)$$

10- صندوقان يحوي الأول 9 بطاقات مرقمة من 1 إلى 9 والثاني يحوي 5 بطاقات

مرقمة من 1 إلى 5، فإذا اختير أحد الصندوقين عشوائياً ثم سحبت منه بطاقة واحدة،

أوجد احتمال أن يكون العدد المكتوب على هذه البطاقة عدداً زوجياً. وإذا عُلِمَ أن الرقم

المكتوب على البطاقة زوجي فما احتمال أن تكون قد سُحِبَت من الصندوق الأول.

11- ثلاثة صناديق A، B، C يحتوي A على 5 كرات حمراء و 3 كرات بيضاء، ويحتوي الصندوق B على 1 حمراء و 2 بيضاء، ويحتوي الصندوق C على 3 حمراء و 2 بيضاء. إذا اختير صندوق من الثلاثة عشوائياً ثم سُحِبَت منه كرة واحدة عشوائياً، احسب احتمال أن تكون الكرة المسحوبة بيضاء، وإذا عُلِمَ أن الكرة المسحوبة بيضاء احسب احتمال أن تكون سُحِبَت من الصندوق الثاني.

12- يعمل في مكتب للترجمة مترجمان، الأول يطبع 60% من التقارير وتكون 80% من الصفحات التي يطبعها دون أخطاء، والثاني يطبع باقي التقارير وتكون 85% من الصفحات التي يطبعها دون أخطاء. اختار مدير المكتب صفحة من التقارير عشوائياً. احسب احتمال أن تكون بلا أخطاء. وإذا علم أن الورقة المسحوبة بلا أخطاء فاحسب احتمال أن تكون طُبِعَت من قبل المترجم الأول ثم احتمال أن تكون طُبِعَت من قبل المترجم الثاني.

13- إذا كان A, B حدثين بحيث :

$$P(A/B) = 0.3, P(B) = 0.4, P(A) = 0.5$$

بين ما إذا كان الحدثان مستقلين أو لا ثم أوجد $P(A \cap B)$

14- إذا كان A, B حدثين بحيث :

$$P(A/B) = \frac{2}{3}, P(A \cap B) = \frac{1}{3}$$

احسب $P(B)$.

15- إذا كان $P(A) = 0.3$, $P(A \cup B) = 0.75$ أوجد $P(B)$ إذا كان الحدثان A, B متنافيين مرة ومستقلين مرة أخرى.

16- اختر الإجابة الصحيحة فيما يلي:

• عند القاء حجر نرد منتظم مرة واحدة يكون احتمال الحصول على عدد فردي أكبر من 5 هو:

أ) 1 ب) $\{7\}$ ج) $\emptyset$ د) 0

• عند القاء حجر نرد منظم مرة واحدة يكون حدث الحصول على عدد فردي أكبر من 5 هو:

أ) 1 ب) $\{7\}$ ج) $\emptyset$ د) 0

• إذا كان A, B حدثين مستقلين، وكان $P(A) = 0.6$, $P(A \cap B) = 0.24$ فإن $P(B)$ يساوي:

أ) 0.4 ب) 0.6 ج) 0.24 د) 0.76

• إذا كان A, B حدثين متنافيين، وكان $P(A) = 0.5$, $P(B) = 0.2$ فإن $P(A \cup B)$ يساوي:

أ) 0.1 ب) 0.3 ج) 0.7 د) 0.6

• عند إلقاء قطعة نقود متزنة 4 مرات فإن عدد عناصر فضاء العينة يساوي:

أ) 8 ب) 16 ج) 32 د) 20

• إذا كان A, B حدثين مستقلين، وكان $P(A) = 0.6$, $P(B/A) = 0.7$ فإن $P(B)$ يساوي:

أ) 0.4 ب) 0.7 ج) 0.42 د) 1.3

• إذا كان: $P(\bar{A} \cup \bar{B}) = 0.8$ فإن $P(A \cap B)$ يساوي:

أ) 0.2 ب) 0.4 ج) 0.6 د) 0.8

• إذا كان $P(\bar{A}) = 0.3$ فإن $P(\Omega - A)$ يساوي:

أ) 0.7 ب) 0.3 ج) 1 د) 1.3

• إذا كان $P(A \cup B) = 0.8$ و $P(A \cap B) = 0.4$ فإن $P(A)$ يمكن أن

يساوي:

أ) 0.9 ب) 0.3 ج) 0.6 د) 0.2

الخلاصة

فضاء العينة: هو مجموعة كل النواتج الممكنة للتجربة العشوائية.

الحدث: هو مجموعة جزئية من فضاء العينة.

الحدثان المتنافيان:

يكون الحدثان A, B متنافيين إذا وفقط إذا كان $A \cap B = \emptyset$

قانون احتمال متمم الحدث:

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

قانون احتمال الفرق بين حدثين:

$$P(A - B) = P(A) - P(A \cap B)$$

قانون احتمال اتحاد حدثين:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

الحوادث المستقلة:

يكون الحدثان مستقلين إذا كان احتمال حدوثهما معاً يساوي حاصل ضرب احتماليهما أي:

$$P(A \cap B) = P(A) \times P(B)$$

الاحتمال الشرطي:

إذا كان A, B حدثين من فضاء عينة Ω فإن احتمال حدوث A بشرط حدوث B يحسب من العلاقة:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

قانون احتمال اتحاد حدثين:

إذا كان A, B حدثين مستقلين فإن:

$$P(A/B) = P(A), \quad P(B/A) = P(B)$$

نظرية التجزئة:

إذا كانت $E_1, E_2, \dots, E_n$ مجموعة من الحوادث تشكل تجزئة لفضاء عينة Ω ، وكان E حدث يمكن أن يقع مع أي حدث من هذه المجموعة فإن احتمال وقوع الحدث E يعطى بالعلاقة:

$$P(E) = P(E_1)P(E/E_1) + P(E_2)P(E/E_2) + \dots + P(E_n)P(E/E_n)$$

نظرية بييز:

إذا كانت $E_1, E_2, \dots, E_n$ مجموعة من الحوادث تشكل تجزئة لفضاء عينة Ω ، بحيث $P(E_i) \neq 0$ لجميع قيم $i \in \{1, 2, \dots, n\}$ وكان E حدث يمكن أن يقع مع أي حدث من هذه المجموعة فإن احتمال وقوع E_i بشرط وقوع E يعطى بالعلاقة:

$$P(E_i/E) = \frac{P(E_i) \times P(E/E_i)}{P(E)}$$

7

المتغيرات العشوائية والتوزيعات الاحتمالية Random Variables and Probability Distributions

تكلّمنا في الفصل السادس عن بعض مفاهيم الاحتمالات والتجارب العشوائية، وفي كثير من الأحيان نرغب في التعامل مع قيم عددية مرتبطة بنتائج التجربة العشوائية وليس بالنواتج نفسها لأن النواتج تكون في بعض الأحيان عبارة عن صفات أو مسميات. إن الطريقة التي يتم بها تحويل هذه النواتج لقيم عددية هي تعريف ما يسمى بالمتغيرات العشوائية على فضاءات العينة، وهذا ما سنوضحه خلال هذا الفصل.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً على:

- حساب احتمالات لمتغيرات عشوائية منفصلة ومتصلة.
- إيجاد التوزيعات الاحتمالية التراكمية المنفصلة والمتصلة.
- حساب التوقع والتباين للمتغيرات العشوائية المنفصلة والمتصلة.
- حساب الاحتمالات لتوزيعات ذي الحدين، بواسون والتوزيع الطبيعي.
- حساب التوقع والتباين لتوزيع ذي الحدين وتوزيع بواسون.
- حساب الاحتمالات لمتغير عشوائي يتبع توزيعاً طبيعياً غير معياري.

عند إجراء تجربة عشوائية قد لا يكون اهتمام الباحث منصباً على نواتج التجربة العشوائية نفسها، وإنما على أرقام مرتبطة بهذه النواتج، هذه الأرقام تعتبر أرقاماً عشوائية، وذلك لأن نواتج التجربة العشوائية هي أصلاً نواتج عشوائية. وهذا يعني أننا يمكن أن نقابل كل نتيجة من نواتج التجربة العشوائية بقيمة عددية، مما يعني أنه يمكن تعريف اقتران من فضاء العينة إلى مجموعة الأعداد الحقيقية. مثل هذه الاقترانات يطلق عليها اسم متغيرات عشوائية.

المتغير العشوائي The Random Variable

المتغير العشوائي هو اقتران حقيقي مجاله Ω ومجاله المقابل مجموعة الأعداد الحقيقية R . ويرمز للمتغير العشوائي بأحد أحرف اللغة الانجليزية الكبيرة مثل Z, Y, X ويرمز للقيم التي يأخذها بحرف اللغة الانجليزية المقابل الصغير وهو z, y, x على الترتيب. فإذا كان X متغيراً عشوائياً فإننا نعبر عن ذلك رياضياً بما يلي:

$$X: \Omega \rightarrow R$$

مثال 1.7: في تجربة إلقاء قطعة نقد ثلاث مرات، إذا كان المتغير العشوائي X يعبر عن عدد الصور التي تظهر في الرميات الثلاث فإن القيم التي يمكن أن يأخذها X هي:

$$x \in \{0, 1, 2, 3\}$$

حيث x تعبر عن عدد الصور التي تظهر.

والجدير بالذكر هنا أن كل قيمة تقابل حدثاً من فضاء العينة، فالرقم صفر يعبر عن حدث الحصول على ثلاث كتابات أي يحوي جميع النتائج التي لا تحتوي على أية صورة، والرقم واحد يعبر عن حدث يحوي جميع النتائج التي تحتوي على صورة واحدة وبغض النظر عن ترتيبها، فقد تكون هذه الصورة ناتج الرمية الأولى وقد تكون ناتج الرمية الثانية وقد تكون ناتج الرمية الثالثة وهكذا.

ونبين هذه المقابلة في الجدول التالي:

الحدث	x
$\{TTT\}$	0
$\{HTT, THT, TTH\}$	1
$\{HHT, HTH, THH\}$	2
$\{HHH\}$	3

مثال 2.7: من دراسات سابقة اتضح أن أوزان الأرناب عند عمر شهرين تتراوح بين 1250 جرام و1560 جرام. إذا عرفنا متغيراً عشوائياً X يعبر عن وزن الأرناب عند عمر شهرين نجد أن القيم التي يمكن أن يأخذها هذا المتغير العشوائي هي:

$$x \in [1250, 1560]$$

المتغير العشوائي المنفصل Discrete Random Variable

لاحظ في مثال (1.7) أن القيم التي يأخذها المتغير العشوائي هي قيم قابلة للعد. مثل هذه المتغيرات يطلق عليها اسم متغيرات عشوائية منفصلة، فالمتغير العشوائي المنفصل هو المتغير العشوائي الذي يأخذ قيماً قابلة للعد حتى لو كانت غير منتهية. وفي المسائل العملية يقيس هذا النوع من المتغيرات العشوائية كميات عددية قابلة للعد، مثل عدد المصابيح الصالحة، عدد الحوادث في الشهر في طريق معين، عدد رواد عيادة العيون في اليوم وعدد الطلبة المسجلين في شعب تخصص معين.

المتغير العشوائي المتصل Continuous Random Variable

نلاحظ في مثال (2.7) أن القيم التي يأخذها المتغير العشوائي هي قيم غير قابلة للعد وتمثل فترة من مجموعة الأعداد الحقيقية. مثل هذه المتغيرات يطلق عليها اسم متغيرات عشوائية متصلة، فالمتغير العشوائي المتصل هو المتغير العشوائي الذي يأخذ قيماً غير قابلة للعد تنتمي

لمجموعة الأعداد الحقيقة كلها أو أي مجموعة جزئية منها، وفي المسائل العملية يقيس هذا النوع من المتغيرات العشوائية كميات عددية غير قابلة للعد مثل الوزن، الطول، درجة الحرارة والعمر. عرفنا أن كل قيمة يأخذها المتغير العشوائي تعبر عن حدث من فضاء العينة، وتعلمنا كيف نحسب الاحتمالات لحوادث مأخوذة من فضاء العينة وبالتالي يمكننا الآن مقابلة كل قيمة من القيم التي يأخذها المتغير العشوائي باحتمال الحدث المقابل.

هذه المقابلة تمكنا من تعريف اقتران مجاله هو مجموعة قيم المتغير العشوائي (مدى المتغير العشوائي) ومجاله المقابل هو الفترة الاحتمالية $[0, 1]$. مثل هذه الاقترانات يطلق عليها اسم توزيعات احتمالية أو اقترانات احتمالية.

في حالة المتغيرات العشوائية المنفصلة يسمى هذا الاقتران: توزيع احتمالي منفصل أو اقتران الكتلة الاحتمالية **Probability mass function (pmf)**.

وفي حالة المتغيرات العشوائية المتصلة يسمى: توزيع احتمالي متصل أو اقتران كثافة احتمالية **Probability density function (pdf)**.

ولهذه الاقترانات الاحتمالية يمكن حساب ما يسمى بالتوزيع الاحتمالي التراكمي والذي يلعب دوراً هاماً في الإحصاء الاستدلالي حيث يستخدم التوزيع التراكمي لحساب الاحتمالات لقيم أقل من أو تساوي قيمة معينة من قيم المتغير العشوائي. وسنتحدث في هذا الفصل أولاً عن المتغيرات العشوائية المنفصلة وتوزيعاتها الاحتمالية ومثالين على توزيعين منفصلين، وثانياً عن المتغيرات العشوائية المتصلة وتوزيعاتها الاحتمالية وعن التوزيع الطبيعي كمثال على التوزيعات المتصلة الخاصة.

أولاً: التوزيعات الاحتمالية المنفصلة

Discrete Probability Distributions

بفرض أن X متغير عشوائي منفصل معرف على فضاء عينة Ω ومداه $X(\Omega)$ ،
يسمى الاقتران:

$$f: X(\Omega) \rightarrow [0, 1]$$

اقتران كتلة احتمالية إذا تحقق الشرطان:

$$1) f(x) \geq 0, \quad \forall x \in X(\Omega)$$

$$2) \sum_{x \in X(\Omega)} f(x) = 1$$

ويكون:

$$P(X = x) = f(x), \quad \forall x \in X(\Omega)$$

مثال 3.7: إذا كان X متغيراً عشوائياً اقتران الكتلة الاحتمالية له يعطى بالعلاقة:

$$f(x) = \frac{1}{6}x, \quad x = 1, 2, 3$$

أوجد احتمالات قيم المتغير العشوائي المختلفة.

الحل/

بالتعويض المباشر نجد أن:

$$P(X = 1) = f(1) = \frac{1}{6} \times 1 = \frac{1}{6}$$

$$P(X = 2) = f(2) = \frac{1}{6} \times 2 = \frac{1}{3}$$

$$P(X = 3) = f(3) = \frac{1}{6} \times 3 = \frac{1}{2}$$

ويمكن التعبير عن اقتران الكتلة الاحتمالية بجدول يحوي قيم المتغير العشوائي واحتمالاتها وذلك كما يلي:

x	1	2	3
$P(X = x) = f(x)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

مثال 4.7: إذا كان X متغيراً عشوائياً يُعطى اقتران الكتلة الاحتمالية له بالجدول التالي:

x	-1	0	1	2
$P(x)$	0.12	k	0.35	0.25

أوجد قيمة k ، ثم احسب:

$$P(-1 < X \leq 1) \quad -1$$

$$P(X \leq 1) \quad -2$$

/الحل/

من شروط تعريف اقتران الكتلة الاحتمالية نجد أن:

$$\sum P(x) = 1$$

$$P(-1) + P(0) + P(1) + P(2) = 1$$

$$0.12 + k + 0.35 + 0.25 = 1$$

$$k = 1 - 0.72 = 0.28$$

فيكون جدول التوزيع الاحتمالي كما يلي:

x	-1	0	1	2
$P(x)$	0.12	0.28	0.35	0.25

وتكون الاحتمالات المطلوبة:

$$1- P(-1 < X \leq 1) = P(0) + P(1) = 0.28 + 0.35 = 0.63$$

$$2- P(X \leq 1) = P(-1) + P(0) + P(1) = 0.12 + 0.28 + 0.35 = 0.75$$

ويمكن حساب هذا الاحتمال باستخدام قانون المتممة حيث يكون:

$$P(X \leq 1) = 1 - P(x > 1) = 1 - P(2) = 1 - 0.25 = 0.75$$

التوزيع الاحتمالي التراكمي (cdf) Cumulative Distribution function

بفرض أن X متغير عشوائي منفصل معرف على فضاء عينة Ω ومداه $X(\Omega)$ ، و له اقتران كتلة احتمالي f . نُعرّف التوزيع التراكمي للمتغير العشوائي بأنه الاقتران F الذي يحقق:

$$F(x) = P(X \leq x), \quad -\infty < x < \infty$$

وبفرض أن $X(\Omega) = \{x_1, x_2, \dots, x_n\}$ فإن اقتران التوزيع الاحتمالي التراكمي يُعرّف كما يلي:

تعريف

التوزيع التراكمي للمتغير العشوائي المنفصل

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ 1 & x_n \leq x < \infty \end{cases}$$

مثال 5.7: لمثال (4.7) أوجد التوزيع التراكمي للمتغير العشوائي X ثم احسب:
 $P(X \leq 1.3)$

الحل/

التوزيع الاحتمالي في مثال (4.7) هو:

x	-1	0	1	2
$P(x)$	0.12	0.28	0.35	0.25

ومن تعريف التوزيع التراكمي مباشرة نجد أن:

$$F(x) = \begin{cases} 0 & -\infty < x < -1 \\ 0.12 & -1 \leq x < 0 \\ 0.12 + 0.28 & 0 \leq x < 1 \\ 0.12 + 0.28 + 0.35 & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

$$F(x) = \begin{cases} 0 & x < -1 \\ 0.12 & -1 \leq x < 0 \\ 0.4 & 0 \leq x < 1 \\ 0.75 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

ويكون الاحتمال المطلوب:

$$P(X \leq 1.3) = F(1.3) = 0.75$$

التوقع الرياضي Mathematical Expectation

ليكن X متغيراً عشوائياً منفصلاً. يُعرّف التوقع الرياضي (ويسمى أيضاً الوسط الحسابي) للمتغير العشوائي بأنه:

تعريف

التوقع الرياضي

$$\mu_X = E(X) = \sum_{x \in X(\Omega)} x P(x)$$

ويمثل التوقع الرياضي القيمة المتوسطة التي يأخذها المتغير العشوائي، أي القيمة المتوسطة للمدى، ولاحظ أنه إذا كانت الاحتمالات متساوية أي:

$$P(x) = \frac{1}{n}, \quad \forall x \in \{x_1, x_2, x_3, \dots, x_n\}$$

فإن:

$$\mu_X = E(X) = \sum_{i=1}^n x_i \frac{1}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

وهو تعريف الوسط الحسابي الشائع الاستخدام الذي سبق دراسته في الفصل الثالث والذي يساوي مجموع الق مقسوماً على عددها.

مثال 6.7: إذا كان X متغيراً عشوائياً يُعطى اقترانه الاحتمالي بالجدول التالي:

x	0	1
$p(x)$	0.6	0.4

أوجد توقع X .

مباشرة من تعريف التوقع نجد أن:

$$\mu_X = E(X) = \sum_x x P(x) = 0 \times 0.6 + 1 \times 0.4 = 0.4$$

والآن سنذكر بعض النظريات الهامة في التوقع الرياضي بدون برهان:

إذا كان X متغيراً عشوائياً، c ثابتاً اختيارياً فإن:

$$E(c) = c$$

$$E(cX) = cE(X)$$



من هذه النظرية نجد أن: $E(0) = 0$ ، $E(-2) = -2$ ، $E(5) = 5$

وإذا كان: $E(X) = 3$ فإن:

$$E(2X) = 2 \times E(X) = 2 \times 3 = 6$$

إذا كان X, Y متغيرين عشوائيين، فإن:

$$E(X \pm Y) = E(X) \pm E(Y)$$



فإذا كان: $E(X) = 3$ ، $E(Y) = 4$ فإن:

$$E(2X - 3Y) = E(2X) + E(-3Y)$$

$$= 2 \times E(X) - 3 \times E(Y) = 2 \times 3 - 3 \times 4 = 6 - 12 = -6$$



إذا كان Y, X متغيرين عشوائيين مستقلين، فإن :

$$E(XY) = E(X)E(Y)$$

فإذا كان:

$E(Y) = -3$, $E(X) = 5$ وكان المتغيران العشوائيان Y, X مستقلين فإن:

$$E(2XY) = 2E(XY) = 2 \times E(X) \times E(Y) = 2 \times 5 \times -3 = -30$$

ملاحظة:

إذا كان X متغيراً عشوائياً منفصلاً وكان $Y = g(X)$ فإن Y متغير عشوائي من نفس

نوع X ويكون توقع Y كما يلي:

$$E(Y) = E(g(X)) = \sum_{x \in X(\Omega)} g(x) P(x)$$

يُعرّف تباين المتغير العشوائي X ويرمز له بالرمز σ_x^2 أو بالرمز $Var(X)$ بأنه :

تعريف

تباين المتغير العشوائي المنفصل

$$Var(X) = \sigma_x^2 = E(X^2) - [E(X)]^2$$

حيث:

$$E(X^2) = \sum_{x \in X(\Omega)} x^2 P(x)$$

إذا كان X متغيراً عشوائياً، c ثابتاً اختياريّاً فإن:

$$Var(c) = 0$$

$$Var(cX) = c^2 Var(X)$$



من هذه النظرية نجد أن:

$$Var(5) = 0 \quad , \quad Var(-2) = 0 \quad , \quad Var(0) = 0$$

وإذا كان: $Var(X) = 2$ فإن:

$$Var(-3X) = (-3)^2 \times Var(X) = 9 \times 2 = 18$$

إذا كان X, Y متغيرين عشوائيين مستقلين، فإن:

$$Var(X \pm Y) = Var(X) + Var(Y)$$



فإذا كان: $Var(X) = 3$, $Var(Y) = 4$ فإن:

$$Var(2X - 3Y) = Var(2X) + Var(-3Y)$$

$$= (2)^2 \times Var(X) + (-3)^2 \times Var(Y)$$

$$= 4 \times 3 + 9 \times 4 = 12 + 36 = 48$$

مثال 7.7: ليكن X متغيراً عشوائياً اقترانه الاحتمالي معطى بالجدول:

x	0	1	2	3
$p(x)$	0.2	0.3	0.4	0.1

أوجد:

$$E(X) \quad (1)$$

$$\sigma_X^2 \quad (2)$$

$$\sigma_X \quad (3)$$

$$E(X^2 + 2X) \quad (4)$$

$$Var(-2X) \quad (5)$$

الحل/

بالتطبيق المباشر لتعريف التوقع والتباين نجد أن:

$$1) E(X) = \sum xP(x) = 0 \times 0.2 + 1 \times 0.3 + 2 \times 0.4 + 3 \times 0.1 = 1.4$$

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

$$= 0 \times 0.2 + 1 \times 0.3 + 4 \times 0.4 + 9 \times 0.1 = 2.8$$

$$\sigma_X^2 = Var(X) = E(X^2) - (E(X))^2 = 2.8 - (1.4)^2 = 0.84$$

$$3) \sigma_X = \sqrt{\sigma_X^2} = \sqrt{0.84} = 0.92$$

$$4) E(X^2 + 2X) = \sum (x^2 + 2x)P(x)$$

$$= 0 \times 0.2 + 3 \times 0.3 + 8 \times 0.4 + 15 \times 0.1 = 5.6$$

$$5) Var(-2X) = (-2)^2 \times Var(X) = 4 \times 0.84 = 3.36$$

ويمكن ايجاد المطلوب الرابع باستخدام خواص التوقع كما يلي:

$$E(X^2 + 2X) = E(X^2) + 2E(X) = 2.8 + 2 \times 1.4 = 5.6$$

توزيعات احتمالية خاصة Special probability distributions

التوزيعات الاحتمالية الخاصة هي عبارة عن توزيعات احتمالية لها قاعدة معروفة، وهي تنتج عن تجارب عشوائية لوحظ أنها تأخذ نمطاً معيناً في تغييرها مما مكن المختصين من التوصل إلى صيغ رياضية تصف تغيرات نتائج تلك التجارب العشوائية.

وستحدث عن توزيع ذي الحدين وتوزيع بواسون كأمثلة على المتغير العشوائي المنفصل وعن التوزيع الطبيعي كمثال على المتغير العشوائي المتصل، حيث سنعرف التوزيعات الاحتمالية لها ونذكر بعض خواصها.

توزيع ذي الحدين Binomial distribution

في كثير من التجارب العشوائية تكون نتيجة التجربة أحد أمرين، إما نجاح وإما فشل. وتتكرر هذه التجارب عدة مرات بحيث تكون نتيجة كل إجراء للتجربة مستقلة عن نتيجة أي إجراء آخر. فمثلاً عند رمي قطعة نقود عدة مرات فإن النتيجة تكون إما ظهور صورة أو كتابة، وتكون نتيجة أي رمية مستقلة عن نتيجة أي رمية أخرى، فإذا كانت نتيجة الرمية الأولى ظهور الصورة لأعلى، فإن ذلك لا يؤثر في نتيجة الرمية الثانية وهكذا. إن مثل هذه التجربة تسمى تجربة ذي الحدين.

وبصورة عامة كل تجربة تحقق الشروط التالية تسمى تجربة ذي الحدين:

- 1- تتألف التجربة من عدد معين من المحاولات و ليكن n .
- 2- نتيجة كل محاولة أحد ناتجين، نسمي أحدهما نجاحاً والآخر فشلاً.
- 3- نتيجة كل محاولة مستقل عن نتيجة أي محاولة أخرى.
- 4- احتمال النجاح في كل محاولة ثابت و ليكن p .

بفرض أن X متغير عشوائي يعبر عن عدد مرات النجاح في تجربة ذي، وكان احتمال النجاح p وكان n عدد مرات تكرار التجربة، فنقول أن X يتبع توزيع ذي الحدين ونعبر عن ذلك بكتابة:

$$X \sim b(n, p)$$

إذا كان اقتران الكتلة الاحتمالي له يعطى بالعلاقة:

لا تنسى اقتران الكتلة الاحتمالي لتوزيع ذي الحدين



$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

حيث:

x تمثل عدد مرات النجاح و

$$\binom{n}{x} = \frac{n!}{x! (n - x)!}$$

وتسمى p و n بمعلمتي التوزيع .

ونحصل على توقع وتباين المتغير العشوائي الذي يتبع توزيع ذي الحدين من العلاقتين التاليتين:

توقع وتباين المتغير العشوائي لتوزيع ذي الحدين على الترتيب:



$$E(X) = np$$

$$\sigma^2 = np(1 - p)$$

مثال 8.7: إذا كان X متغيراً عشوائياً بحيث $X \sim b(4, 0.2)$ أوجد ما يلي:

(1) $p(x)$

(2) $p(2)$

(3) $P(X \geq 1)$

(4) التوقع والتباين للمتغير العشوائي.

الحل/ من الواضح أن $n = 4$, $p = 0.2$ وبالتالي:

1) $P(x) = \binom{4}{x}(0.2)^x(0.8)^{4-x}$, $x = 0, 1, 2, 3, 4$

ويمكن التعبير عن اقتران الكتلة الاحتمالي بالجدول التالي:

x	0	1	2	3	4
$P(x)$	0.4096	0.4096	0.1536	0.0256	0.0016

وتكون الاحتمالات المطلوبة كما يلي:

2) $P(2) = \binom{4}{2}(0.2)^2(0.8)^2 = 0.1536$

3) $P(X \geq 1) = P(1) + P(2) + P(3) + P(4)$

وحيث أن الحسابات في هذه الحالة طويلة، من الأفضل حساب الاحتمال باستخدام

متمم الحدث وذلك كما يلي:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(0)$$

$$= 1 - \binom{4}{0}(0.2)^0(0.8)^4 = 0.5904$$

4) $E(X) = np = 4 \times 0.2 = 0.8$

$$\sigma^2 = np(1 - p) = 4 \times 0.2 \times 0.8 = 0.64$$

مثال 9.7: إذا كان احتمال نجاح عملية جراحية هو 0.9 وأجريت العملية لعشرة مرضى فاحسب ما يلي:

- 1- احتمال نجاح العملية لسبعة مرضى.
- 2- احتمال نجاح العملية لجميع المرضى.
- 3- احتمال نجاح العملية لثمانية مرضى على الأقل.
- 4- احتمال نجاح العملية لثمانية مرضى على الأكثر.
- 5- توقع عدد المرضى الذين سيجرون العملية.
- 6- تباين عدد المرضى الذين سيجرون العملية.

الحل/

لاحظ هنا أن التجربة هي تجربة ذي الحدين حيث $n = 10$ و $p = 0.9$ وبفرض أن X متغير عشوائي يعبر عن عدد العمليات الناجحة فيكون:

$$X \sim b(10, 0.9)$$

ويكون اقتران الكتلة الاحتمالي للمتغير العشوائي X :

$$P(x) = \binom{10}{x} (0.9)^x (0.1)^{10-x}, \quad x = 0, 1, 2, \dots, 10$$

وبالتالي:

$$1) \quad P(7) = \binom{10}{7} (0.9)^7 (0.1)^3 = 0.0574$$

$$2) \quad P(10) = \binom{10}{10} (0.9)^{10} (0.1)^0 = 0.3487$$

$$3) \quad P(X \geq 8) = P(8) + P(9) + P(10)$$

$$= \binom{10}{8} (0.9)^8 (0.1)^2 + \binom{10}{9} (0.9)^9 (0.1)^1 + \binom{10}{10} (0.9)^{10} (0.1)^0$$

$$= 0.9298$$

$$\begin{aligned}
4) \quad P(X \leq 8) &= P(0) + P(1) + \cdots + P(8) \\
&= 1 - P(x > 8) \\
&= 1 - [P(9) + P(10)] \\
&= 1 - \left[\binom{10}{9} (0.9)^9 (0.1)^1 + \binom{10}{10} (0.9)^{10} (0.1)^0 \right] = 0.2639
\end{aligned}$$

$$5) E(X) = np = 10 \times 0.9 = 9$$

$$6) \sigma^2 = np(1 - p) = 10 \times 0.9 \times 0.1 = 0.9$$

مثال 10.7: رميت قطعة نقود متزنة أربع مرات. جد التوزيع الاحتمالي للمتغير العشوائي الذي يعبر عن عدد مرات ظهور الصورة في هذه التجربة، وأوجد احتمال عدم الحصول على أي صورة. ثم احسب توقع وتباين المتغير العشوائي.

الحل/ لاحظ هنا أن التجربة هي تجربة ذات الحدين حيث $n = 4$ و $p = 0.5$ وبفرض أن X متغير عشوائي يعبر عن عدد الصور التي تظهر فيكون:

$$X \sim b(4, 0.5)$$

ويكون اقتران الكتلة الاحتمالي للمتغير العشوائي X :

$$P(x) = \binom{4}{x} (0.5)^x (0.5)^{4-x}, \quad x = 0, 1, 2, 3, 4$$

وا احتمال عدم الحصول على أي صورة:

$$P(0) = \binom{4}{0} (0.5)^0 (0.5)^4 = 0.0625$$

أما توقع وتباين المتغير العشوائي:

$$E(X) = np = 4 \times 0.5 = 2$$

$$\sigma^2 = np(1 - p) = 4 \times 0.5 \times 0.5 = 1$$

توزيع بواسون Poisson distribution

توزيع بواسون يُعنى بالمتغير المنفصل الذي يمثل عدد النجاحات في وحدة الزمن أو الفترة المكانية ويمكن أن تكون وحدة الزمن هنا ثانية، دقيقة، ساعة أو يوماً إلى غير ذلك. أما وحدة الفترة المكانية فيمكن أن تكون وحدة طول، مساحة، حجم، أو غير ذلك.

ويعرف هذا التوزيع بتوزيع الحوادث النادرة، حيث أنه يصلح للحوادث نادرة الحدوث مثل عدد حوادث سقوط الطائرات وعدد وصول رسائل بالخطأ لبريد المدينة. ويمثل هذا التوزيع حالة خاصة من توزيع ذي الحدين وذلك عندما يكون احتمال النجاح صغيراً جداً مقابل عدد تكرارات كبير جداً.

وهناك أمثلة كثيرة يمكن دراستها كتطبيقات على توزيع بواسون ومنها:

- في مقسم الهاتف في إحدى الجامعات إذا كان احتمال وصول مكالمات هاتفية في الدقيقة ثابتاً، فإن عدد المكالمات الهاتفية في الدقيقة يعتبر متغيراً عشوائياً خاضعاً لتوزيع بواسون.

- عدد الأخطاء المطبعية في الصفحة الواحدة في كتاب ما يعتبر متغيراً عشوائياً خاضعاً لتوزيع بواسون.

- عدد الجرذان في المتر المربع في حقل قمح.

- عدد السمك الذي يصطاد في منطقة محددة في اليوم الواحد.

- عدد حوادث السيارات على طريق معين في الأسبوع.

- عدد الرسائل المفقودة في مكتب بريد في اليوم.

ولكي تكون تجربة عشوائية هي تجربة بواسون لا بد أن تحقق ما يلي:

1- معدل عدد النجاحات التي تحدث في الفترة المحددة ثابت ومعلوم. وسنعتبر عن ذلك

المعدل بالرمز λ ، ولاحظ هنا أن عدد النجاحات يعني عدد مرات حدوث الظاهرة التي

نحن بصدد دراستها في الفترة المحددة .

2- احتمال حدوث نجاح واحد في فترة صغيرة جداً يتناسب مع طول تلك الفترة.

3- إذا اعتبرنا عدد فترات منفصلة عن بعضها البعض فإن حدوث النجاحات في أي فترة مستقل عن حدوث النجاحات في أي فترة أخرى.

بفرض X هو متغير عشوائي يعبر عن عدد مرات النجاح في فترة زمنية أو مكانية محددة وكان معدل عدد النجاحات يساوي λ ، فنقول أن X يتبع توزيع بواسون بِمَعْلَمَة λ ونعبر عن ذلك بكتابة:

$$X \sim \text{PO}(\lambda)$$

إذا كان اقتران الكتلة الاحتمالي للمتغير العشوائي يعطى بالعلاقة:

لا تنسى! اقتران الكتلة الاحتمالي لتوزيع بواسون:

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$



حيث x تمثل عدد النجاحات في الفترة المحددة.

ويكون توقع المتغير العشوائي X وتباينه :

لا تنسى! توقع وتباين المتغير العشوائي لتوزيع بواسون على الترتيب:

$$E(X) = \lambda$$

$$\sigma_X^2 = \lambda$$



مثال 11.7 : إذا كان X متغيراً عشوائياً بحيث $X \sim PO(4)$ ، احسب:

$$P(X > 1) \quad (1)$$

(2) توقع وتباين المتغير العشوائي.

الحل/

(1) التوزيع الاحتمالي للمتغير العشوائي يعطى بالعلاقة:

$$P(x) = \frac{e^{-4} 4^x}{x!}, \quad x = 0, 1, 2, \dots \dots \dots$$

وبالتالي:

$$P(X > 1) = 1 - P(X \leq 1) = 1 - \left\{ \frac{e^{-4} 4^0}{0!} + \frac{e^{-4} 4^1}{1!} \right\} = 0.91$$

(2) التوقع والتباين للمتغير العشوائي:

$$E(X) = \sigma^2 = 4$$

مثال 12.7: إذا كان متوسط عدد الأيام التي تمطر فيها في شهر شباط هي ثلاثة أيام في الأسبوع. فما احتمال أن تمطر خمسة أيام في الأسبوع في ذلك الشهر؟ وأوجد التوقع والتباين.

الحل/

لاحظ أن التوزيع هو توزيع بواسون حيث الفترة المحددة هي زمنية ومقدارها أسبوع وأعطى المعدل بمقدار 3 أيام أسبوعياً وبالتالي تكون $\lambda=3$ ، و بفرض أن X هو متغير عشوائي يعبر عن عدد الأيام الماطرة في الأسبوع فيكون:

$$X \sim PO(3)$$

ويكون التوزيع الاحتمالي للمتغير العشوائي يعطى بالعلاقة:

$$P(x) = \frac{e^{-3} 3^x}{x!}, \quad x = 0, 1, 2, 3, \dots$$

وبالتالي يكون الاحتمال المطلوب:

$$P(5) = \frac{e^{-3} 3^5}{5!} = \frac{(0.05)(243)}{120} = 0.101$$

ويكون التوقع والتباين للمتغير العشوائي:

$$E(X) = \sigma^2 = 3$$

مثال 13.7: إذا كان معدل حوادث السيارات على طريق معين هو 10 حوادث في الأسبوع.

ما احتمال عدم حدوث أي حادث على ذلك الطريق في أسبوع معين؟ ما احتمال حدوث 6

حوادث في أسبوع معين؟

الحل/

لاحظ أن التوزيع هو توزيع بواسون حيث الفترة المحددة هي زمنية ومقدارها أسبوع وأعطى المعدل بمقدار 10 حوادث اسبوعياً وبالتالي تكون $\lambda = 10$. وبفرض أن X هو متغير عشوائي يعبر عن عدد الحوادث في الأسبوع فيكون:

$$X \sim PO(10)$$

ويكون التوزيع الاحتمالي للمتغير العشوائي يعطى بالعلاقة:

$$P(x) = \frac{e^{-10} 10^x}{x!}, \quad x = 0, 1, 2, 3, \dots$$

فيكون احتمال عدم حدوث أي حادث على الطريق في أسبوع معين هو:

$$P(0) = \frac{e^{-10}10^0}{0!} = 0.00004$$

وا احتمال حدوث 6 حوادث في اسبوع معين هو:

$$P(6) = \frac{e^{-10}10^6}{6!} = 0.063$$

ثانياً: التوزيعات الاحتمالية المتصلة

Continuous Probability Distributions

بفرض أن X متغير عشوائي متصل معرف على فضاء عينة Ω ومداه $[a, b]$ ،

يسمى الاقتران:

$$f: [a, b] \rightarrow [0, 1]$$

اقتران كثافة احتمالية إذا تحقق الشرطان:

$$1- f(x) \geq 0, \quad \forall x \in [a, b]$$

$$2- \int_a^b f(x) dx = 1$$

ويعرف احتمال وقوع قيمة المتغير العشوائي X في الفترة $[c, d]$ بأنه:

$$P(c < X < d) = \int_a^b f(x) dx$$

ولاحظ هنا أن الاحتمال عبارة عن المساحة المحصورة أسفل منحنى التوزيع الاحتمالي وفوق

محور X خلال الفترة المطلوب حساب احتمال وقوع قيمة المتغير العشوائي فيها وذلك من تعريف التكامل المحدود للاقترانات غير السالبة، وبالتالي يكون:

$$P(X = x) = 0, \quad \forall x \in [a, b]$$

—

وبناءً عليه فإن الاحتمالات التالية متساوية:

$$P(c < X < d) = P(c \leq X < d) = P(c < X \leq d) = P(c \leq X \leq d)$$

وسنكتفي في دراستنا على اقترانات الكثافة الاحتمالية التي لها قاعدة واحدة وتمثل كثيرة حدود من الدرجة الثانية على الأكثر.

مثال 14.7 : بين أن الاقتران $f(x)$ يمثل اقتران كثافة احتمالية للمتغير العشوائي X ثم احسب $P(1 < X < 2)$ حيث:

$$f(x) = \begin{cases} \frac{1}{9}x^2, & 0 < x < 3 \\ 0 & \text{في غير ذلك} \end{cases}$$

الحل/

أولاً: من الواضح أن الاقتران موجب لجميع قيم x بمعنى أن:

$$f(x) \geq 0, \quad \forall x \in (0, 3)$$

وكذلك لاحظ أن:

$$\int_0^3 \frac{1}{9}x^2 dx = \frac{1}{27}x^3 \Big|_0^3 = \frac{27}{27} - 0 = 1$$

وبالتالي فإن الاقتران المذكور يمثل اقتران كثافة احتمالية للمتغير العشوائي X .

ثانياً:

$$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2 dx = \frac{1}{27}x^3 \Big|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

مثال 15.7 : إذا كان X متغيراً عشوائياً يُعطى اقتران الكثافة الاحتمالية له بالعلاقة:

$$f(x) = \begin{cases} cx, & 0 < x < 2 \\ 0 & \text{في غير ذلك} \end{cases}$$

أوجد قيمة الثابت c ثم احسب $P(-1 < X < 1)$

أولاً: من شروط تعريف اقتران الكثافة الاحتمالية للمتغير العشوائي نجد أن:

$$\int_0^2 cx \, dx = 1$$

$$\left. \frac{cx^2}{2} \right|_0^2 = 1 \quad \Rightarrow \quad c \left[\frac{4}{2} - 0 \right] = 1$$

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

فيصبح التوزيع الاحتمالي كما يلي:

$$f(x) = \begin{cases} \frac{1}{2}x, & 0 < x < 2 \\ 0 & \text{في غير ذلك} \end{cases}$$

ثانياً:

$$\begin{aligned} P(-1 < X < 1) &= P(-1 < X \leq 0) + P(0 < X < 1) \\ &= 0 + P(0 \leq X < 1) \\ &= \int_0^1 \frac{1}{2}x \, dx = \left. \frac{1}{4}x^2 \right|_0^1 = \frac{1}{4} \end{aligned}$$

التوزيع الاحتمالي التراكمي (cdf) Cumulative Distribution function

بفرض أن X متغير عشوائي متصل معرف على فضاء عينة Ω ومداه $[a, b]$ ، وله

اقتران كثافة احتمالية f . نعرف التوزيع التراكمي للمتغير العشوائي X بأنه الاقتران F الذي يحقق:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) \, dt, \quad -\infty < x < \infty$$

وفي حالة كون اقتران الكثافة الاحتمالية بقاعدة واحدة يمكننا القول بأن:

تعريف

$$F(x) = \begin{cases} 0, & x \leq a \\ \int_a^x f(t) dt, & a < x < b \\ 1, & x \geq b \end{cases}$$

وكنتائج مباشرة على هذا التعريف يمكن استنتاج ما يلي:

$$1) F(-\infty) = F(a) = 0$$

$$2) F(\infty) = F(b^+) = 1$$

$$3) f(x) = \frac{dF(x)}{dx}$$

$$4) P(c < X < d) = F(d) - F(c)$$

حيث ترمز $F(b^+)$ للنهاية اليمنى للاقتران التراكمي عند النقطة b .

مثال 16.7 : إذا كان X متغيراً عشوائياً اقتران كثافته الاحتمالية يعطى بالعلاقة:

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{عدا ذلك} \end{cases}$$

أوجد التوزيع التراكمي للمتغير العشوائي X ثم استخدمه في حساب $P(0 < X \leq 1)$.

الحل/ من تعريف التوزيع الاحتمالي نجد أن:

$$F(x) = 0 , \quad -\infty < x \leq -1$$

ولقيم x في الفترة $[-1, 2]$ نجد أن:

$$F(x) = \int_{-\infty}^x \frac{t^2}{3} dt = \int_{-1}^x \frac{t^2}{3} dt = \frac{t^3}{9} \Big|_{-1}^x = \frac{x^3}{9} + \frac{1}{9} = \frac{x^3 + 1}{9}$$

فيكون التوزيع التراكمي كما يلي:

$$F(x) = \begin{cases} 0 & x \leq -1 \\ \frac{x^3 + 1}{9} & -1 < x < 2 \\ 1 & x \geq 2 \end{cases}$$

والاحتمال المطلوب:

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{1+1}{9} - \frac{0+1}{9} = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

التوقع الرياضي Mathematical Expectation

ليكن X متغيراً عشوائياً متصلاً مداه $[a, b]$ ، يُعرّف التوقع الرياضي أو الوسط الحسابي

للمتغير العشوائي بأنه :

توقع المتغير العشوائي المتصل

$$\mu_X = E(X) = \int_a^b x f(x) dx$$

مثال 17.7 أوجد توقع المتغير العشوائي المتصل X إذا كان اقتران كثافته الاحتمالية معطى بالعلاقة:

$$f(x) = \begin{cases} x, & 0 < x < \sqrt{2} \\ 0, & \text{في غير ذلك} \end{cases}$$

الحل/ بالتطبيق المباشر لتعريف التوقع نجد أن:

$$\begin{aligned} \mu_X = E(X) &= \int_0^{\sqrt{2}} x f(x) dx = \int_0^{\sqrt{2}} x \cdot x dx \\ &= \int_0^{\sqrt{2}} x^2 dx = \frac{x^3}{3} \Big|_0^{\sqrt{2}} = \frac{2\sqrt{2}}{3} \end{aligned}$$

ملاحظة: إذا كان X متغيراً عشوائياً متصلاً مداه $[a, b]$ وكان $Y = g(X)$ فإن Y متغير عشوائي من نفس نوع X ويكون توقع Y كما يلي:

$$E(Y) = E(g(X)) = \int_a^b g(x) f(x) dx$$

ويُعرّف تباين المتغير العشوائي المتصل X ويرمز له بالرمز σ_X^2 أو بالرمز $Var(X)$ بأنه :

التباين

$$Var(X) = \sigma_X^2 = E(X^2) - [E(X)]^2$$

حيث:

$$E(X^2) = \int_a^b x^2 f(x) dx$$

مثال 18.7: أوجد التوقع و الانحراف المعياري للمتغير العشوائي X إذا كان توزيعه الاحتمالي معطى بالعلاقة:

$$f(x) = \begin{cases} \frac{x}{2}, & 0 < x < 2, \\ 0, & \text{في غير ذلك} \end{cases}$$

ثم اوجد $E(X + 1)$

الحل/

بالتطبيق المباشر لتعريف التوقع والتباين نجد أن:

$$E(X) = \int_0^2 x \cdot \frac{x}{2} dx = \int_0^2 \frac{x^2}{2} dx = \frac{x^3}{6} \Big|_0^2 = \frac{8}{6}$$

$$E(X^2) = \int_0^2 x^2 \cdot \frac{x}{2} dx = \int_0^2 \frac{x^3}{2} dx = \frac{x^4}{8} \Big|_0^2 = \frac{16}{8} = 2$$

$$\sigma_X^2 = E(X^2) - (E(X))^2 = 2 - \left(\frac{8}{6}\right)^2 = 0.22$$

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{0.22} = 0.47$$

$$E(X + 1) = \int_0^2 (x + 1) \cdot \frac{x}{2} dx = \int_0^2 \left(\frac{x^2}{2} + \frac{x}{2}\right) dx = \left(\frac{x^3}{6} + \frac{x^2}{4}\right) \Big|_0^2 = \frac{7}{3}$$

ويمكن ايجاد المطلوب الأخير باستخدام خواص التوقع كما يلي:

$$E(X + 1) = E(X) + 1 = \frac{8}{6} + 1 = \frac{7}{3}$$

التوزيع الطبيعي The Normal distribution

يعتبر هذا التوزيع من أهم التوزيعات الاحتمالية المتصلة وأكثرها استخداماً، ويلاحظ أن أغلب الظواهر الطبيعية تتبع في توزيعها التوزيع الطبيعي، مثل أعمار الكائنات الحية، أوزان المنتجات الزراعية والصناعية، ذكاء البشر وأطوال الأطفال في سن معين والكثير من الظواهر الطبيعية الأخرى.

فإذا كان X متغيراً عشوائياً له توزيع طبيعي وسطه الحسابي μ وتباينه σ^2 فإن اقتران كثافته الاحتمالية هو:

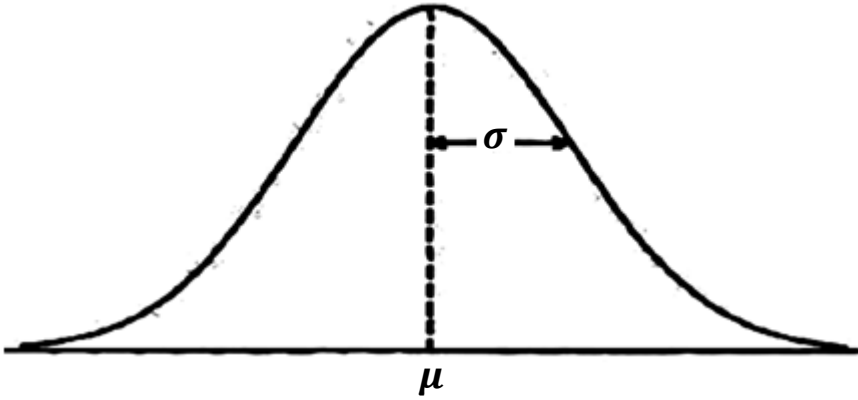
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

ونعبر عن ذلك بالصيغة

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

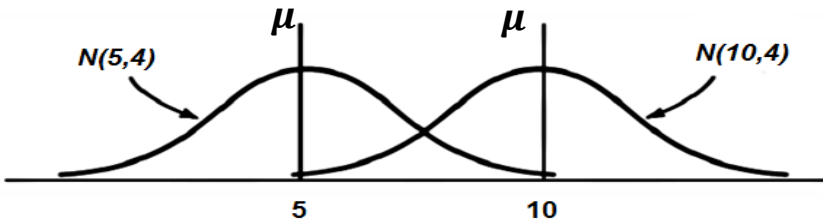
والتي تُقرأ " X تتبع التوزيع الطبيعي بوسط μ وتباين σ^2 "

والرسم التالي يوضح منحنى التوزيع:

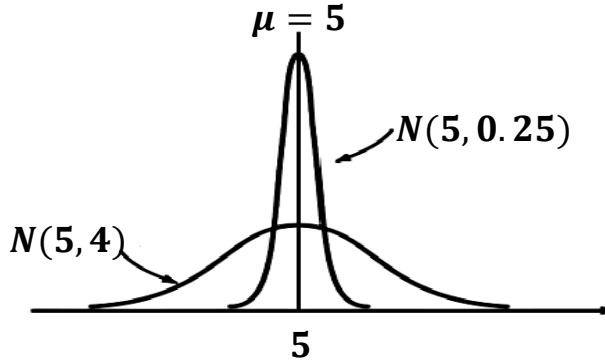


خصائص التوزيع الطبيعي:

- 1- منحناه يأخذ شكل الناقوس المقلوب و يمتد من طرفيه الى $-\infty$, ∞ .
- 2- المساحة المحصورة بين المنحنى ومحور السينات تساوي وحدة مربعة واحدة.
- 3- يكون المنحنى متماثلاً حول المستقيم $x = \mu$.
- 4- التوزيع وحيد المنوال.
- 5- $\mu = Me = Mo$
- 6- بتغير μ و ثبات σ يتحرك المنحنى افقياً يميناً أو يساراً كما هو موضح بالشكل التالي:



7- كلما ازدادت σ و بقيت μ ثابتة فإن المنحنى يبتعد أكثر عن الوسط الحسابي من الجهتين كما هو موضح بالشكل التالي:



وفي الحالة الخاصة التي يكون فيها المتغير العشوائي وسطه الحسابي $\mu = 0$ وتباينه $\sigma^2 = 1$ يرمز له بالرمز Z ويطلق على التوزيع في هذه الحالة اسم التوزيع الطبيعي المعياري **Standard Normal Distribution** ونعبر عن ذلك بالصيغة:

$$Z \sim N(0, 1)$$

ويكون التوزيع الاحتمالي في هذه الحالة معطى بالعلاقة:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty$$

ملاحظة: سنرمز للتوزيع التراكمي بالرمز الشائع $\Phi(z)$ بدلاً من $F(z)$ وبالتالي يكون:

$$\Phi(z) = P(Z \leq z)$$

ولاحظ أنه من تماثل منحنى التوزيع الطبيعي حول وسطه الحسابي يكون:

$$\Phi(z) + \Phi(-z) = 1$$

نذكر الآن النظرية الهامة التالية بدون برهان والتي تمكننا من تحويل أي متغير عشوائي

يتبع توزيع طبيعي غير معياري لمتغير عشوائي يتبع توزيع طبيعي معياري:

إذا كان:

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

فإن:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$



حيث تسمى: $Z = \frac{X - \mu}{\sigma}$ الدرجة المعيارية. وهي قيم خالية من وحدات القياس.

وحيث أن التوزيع من النوع المتصل فتحسب الاحتمالات له بحساب المساحات أسفل المنحنى وفوق محور السينات وذلك خلال الفترة التي نريد حساب احتمال وقوع قيمة المتغير العشوائي فيها.

ومرفق في ملحق الكتاب جدول (2) والخاص بالتوزيع الطبيعي المعياري التراكمي لقيم المتغير العشوائي Z الموجبة و السالبة والذي سنستخدمه في حساب الاحتمالات حيث يعطي الجدول المساحة الواقعة علي يسار قيمة المتغير العشوائي Z أي قيمة $\Phi(z)$.

حيث يعبر العمود أقصى اليسار في الجدول على العدد الصحيح وأول منزله عشرية من قيمة المتغير العشوائي " $\#.\#$ " والصف العلوي في الجدول يعبر عن المنزلة العشرية الثانية أي الجزء من مائة " $\#0.0$ " والقيم التي تقع داخل الجدول تعبر عن المساحات.

والمثال التالي يوضح كيفية استخدام الجدول في حساب المساحات:

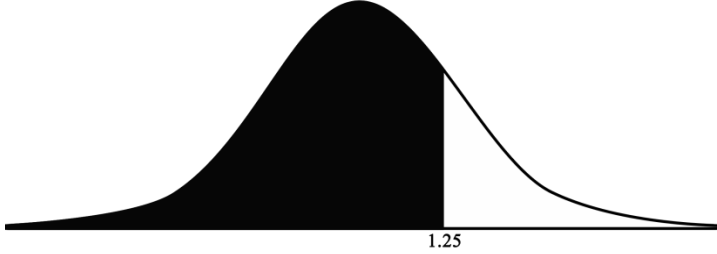
مثال 19.7: أوجد الاحتمالات التالية:

$$P(Z < 1.25) - 1$$

$$P(Z < -1) - 2$$

1- الرسم التالي يوضح المساحة التي تعبر عن الاحتمال المطلوب ولاحظ أن الاحتمال المطلوب يعبر مباشرة عن قيمة التوزيع التراكمي مباشرة أي:

$$P(Z < 1.25) = \Phi(1.25)$$

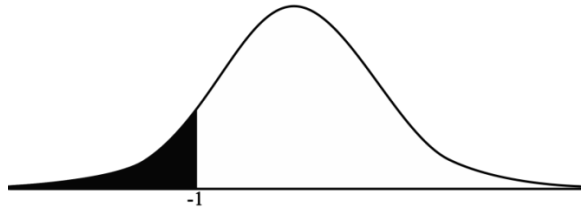


وحيث أن قيمة Z المعطاة موجبة، سنبحث عن المساحة في جدول القيم الموجبة بحيث نختار من العمود أقصى اليسار القيمة 1.2 وسنبحث عن المساحة المطلوبة في صف المساحات التي يقع على يمينها، ومن الصف العلوي نختار القيمة 0.05 وسنبحث عن المساحة المطلوبة في عمود المساحات الذي يقع أسفلها، وتكون المساحة المطلوبة هي التي تقع في تقاطع الصف والعمود المختارين وبالتالي نجد أن الاحتمال المطلوب هو:

$$P(Z < 1.25) = \Phi(1.25) = 0.8944$$

2- حيث أن قيمة Z المعطاة سالبة، سنبحث عن المساحة في جدول القيم السالبة بحيث نختار من العمود أقصى اليسار القيمة -1 وسنبحث عن المساحة المطلوبة في صف المساحات التي يقع على يمينها، ومن الصف العلوي نختار القيمة 0.00 وسنبحث عن المساحة المطلوبة في عمود المساحات الذي يقع أسفلها.

وتكون المساحة المطلوبة هي التي تقع في تقاطع الصف والعمود المختارين وبالتالي نجد أن الاحتمال المطلوب هو:



$$P(Z < -1) = \Phi(-1) = 0.1587$$

مثال 20.7: أوجد الاحتمالات التالية:

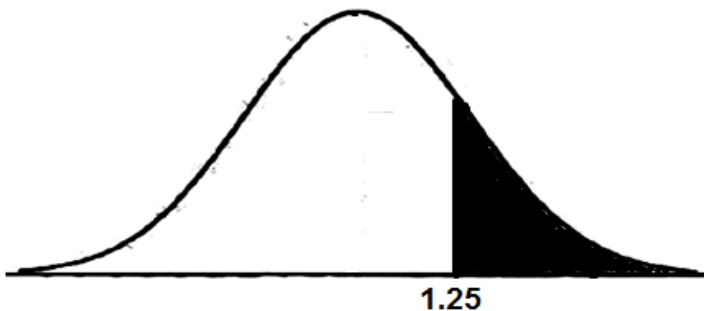
$$P(Z > 1.25) \quad -1$$

$$P(|Z| < 2) \quad -2$$

الحل/

1- لاحظ أن المطلوب المساحة التي تقع على يمين قيمة المتغير العشوائي $Z = 1.25$ وجدول (2) لا يحسب مثل هذه الاحتمالات بل يحسب المساحات التي تقع على يسار قيم المتغير العشوائي، وفي نفس الوقت نجد أن هذه المساحة تساوي:

المساحة الكلية الواقعة أسفل المنحنى وهي تساوي "1" مطروحاً منها المساحة الواقعة على يسار قيمة المتغير العشوائي 1.25 والتي تساوي $\Phi(1.25)$.



$$P(Z > 1.25) = 1 - \Phi(1.25) = 1 - 0.8944 = 0.1056$$

ومن خاصية التماثل يمكننا حساب الاحتمال كما يلي:

$$P(Z > 1.25) = P(Z < -1.25) = \Phi(-1.25) = 0.1056$$

وبصفة عامة يمكننا القول أن:

$$P(Z > z) = 1 - \Phi(z)$$

$$P(Z > z) = \Phi(-z)$$



2- نعلم أن $|Z| < a$ تكافئ $-a < Z < a$ وبالتالي:

$$P(|Z| < 2) = P(-2 < Z < 2)$$

$$= \Phi(2) - \Phi(-2) = 0.9772 - 0.0228 = 0.9544$$

مثال 21.7: أوجد الاحتمالات التالية:

$$P(0 < Z < 1.25) \text{ -1}$$

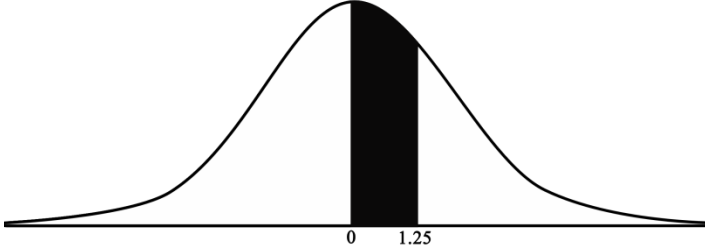
$$P(-2 < Z < 0) \text{ -2}$$

$$P(0.5 < Z < 1.5) \text{ -3}$$

$$P(-2 < Z < -1.5) \text{ -4}$$

$$P(-1 < Z < 1.5) \text{ -5}$$

1- حيث أن قيمة Z المعطاة تقع بين قيمتين، نجد أن المساحة المطلوبة تنحصر بين هاتين القيمتين كما هو موضح بالرسم المصاحب، ومن خواص التوزيع التراكمي نجد أن الاحتمال المطلوب هو:



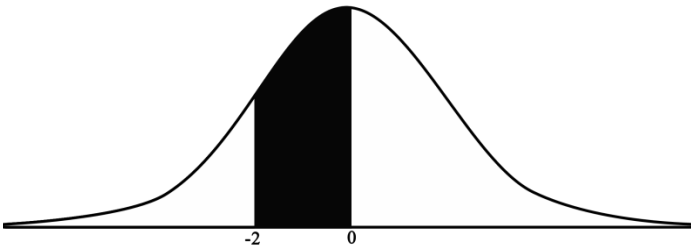
$$P(0 < Z < 1.25) = \Phi(1.25) - \Phi(0) = 0.8944 - .05 = 0.3944$$

وبصفة عامة يمكننا القول أن:

$$P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$$

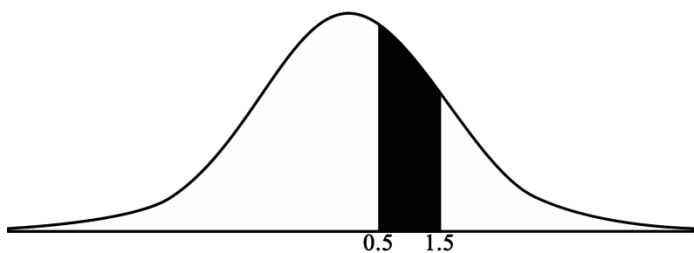


-2



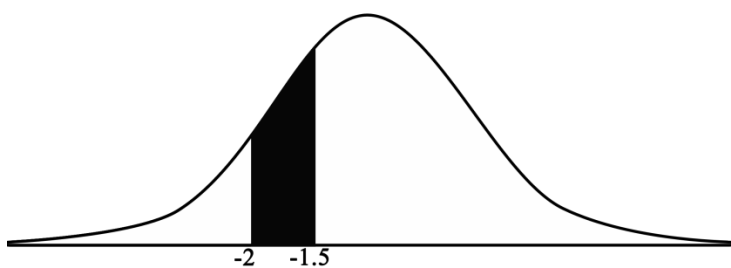
$$P(-2 < Z < 0) = \Phi(0) - \Phi(-2) = 0.5 - 0.0228 = 0.4772$$

-3



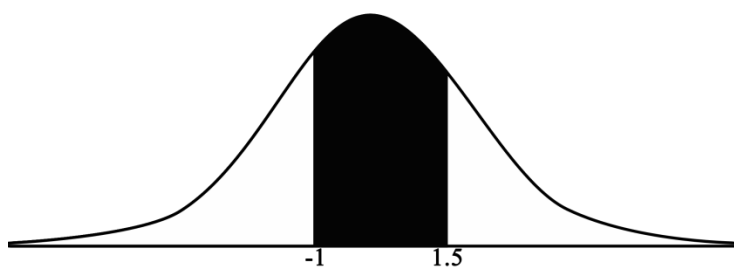
$$\begin{aligned} P(0.5 < Z < 1.5) &= \Phi(1.5) - \Phi(0.5) \\ &= 0.9332 - 0.6915 = 0.2417 \end{aligned}$$

-4



$$\begin{aligned} P(-2 < Z < -1.5) &= \Phi(-1.5) - \Phi(-2) \\ &= 0.0668 - 0.0228 = 0.044 \end{aligned}$$

-5



$$\begin{aligned} P(-1 < Z < 1.5) &= \Phi(1.5) - \Phi(-1) \\ &= 0.9332 - 0.1587 = 0.7745 \end{aligned}$$

مثال 22.7: أوجد قيمة a التي تحقق:

$$1 - P(Z < a) = 0.8413$$

$$2 - P(Z > a) = 0.937$$

الحل/

1- لاحظ هنا أن المساحة معلومة والمطلوب إيجاد قيمة المتغير العشوائي التي تقع على يساره مساحة مقدارها 0.8413 ، أي قيمة a التي تحقق:

$$\Phi(a) = 0.8413$$

ولاحظ هنا أن قيمة المساحة المعطاة أكبر من 0.5 لذلك نتوقع أن تكون قيمة المتغير العشوائي موجبة وبالتالي سنستخدم الجدول الموجب لإيجاد قيمة المتغير العشوائي وذلك بالبحث في المساحات داخل جدول (2) الموجب عن المساحة 0.8413 وإذا لم نجدها نفسها في الجدول نأخذ أقرب قيمة لها ومن ثم ننظر إلى القيمة التي تقابلها في العمود أقصى اليسار و القيمة التي تقابلها في الصف العلوي نجد أن $a = 1$

2- لاحظ أن:

$$P(Z > a) = 1 - \Phi(a)$$

وبالتالي:

$$1 - \Phi(a) = 0.937$$

$$\Phi(a) = 1 - 0.937 = 0.063$$

ولاحظ أن المساحة 0.063 أقل من 0.5 لذلك سنبحث في الجدول السالب عن قيمة

$$a = -1.53 \text{ المطلوبة لنجد أن:}$$

مثال 23.7: أوجد قيمة a بحيث أن $P(0 < Z < a) = 0.4251$

الحل/

$$P(0 < z < a) = \Phi(a) - \Phi(0)$$

$$\Phi(a) - 0.5 = 0.4251$$

$$\Phi(a) = 0.9251$$

وبالبحث في جدول (2) الموجب نجد أن $a = 1.44$.

أحياناً نتعامل مع متغير عشوائي X يتبع توزيع طبيعي غير معياري بمعنى:

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

ونرغب في حساب احتمالات للمتغير العشوائي فنلجأ لتطبيق نظرية (6.7). فإذا كان X متغيراً عشوائياً بحيث:

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

فإن:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

وبالتالي:

$$P(a < X < b) = P\left(\frac{a - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{b - \mu}{\sigma}\right)$$

$$= P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right)$$

$$= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

والأمثلة التالية توضح ذلك:

مثال 24.7: إذا كان $X \sim N(3,4)$ أوجد $P(2 < X < 5.4)$

الحل/لاحظ أن التوزيع الطبيعي هنا غير معياري حيث $\mu = 3$ و $\sigma^2 = 4$ وبتطبيق نظرية (6.7) نحصل على:

$$Z = \frac{X - 3}{2} \sim N(0,1)$$

ويكون الاحتمال المطلوب:

$$\begin{aligned} P(2 < X < 5.4) &= P\left(\frac{2 - 3}{2} < \frac{X - 3}{2} < \frac{5.4 - 3}{2}\right) \\ &= \Phi(1.2) - \Phi(-0.5) \\ &= 0.8849 - 0.3085 = 0.5764 \end{aligned}$$

مثال 25.7: تتخذ أطوال 1000 طالب توزيعاً طبيعياً وسطه الحسابي 160 سم وانحرافه المعياري 10 اختير طالب عشوائياً أوجد:

1- احتمال أن يكون طوله أقل من 170 سم.

2- النسبة المئوية للطلبة الذين تتراوح أطوالهم بين 175 سم و 165 سم.

3- عدد الطلبة الذين تزيد أطوالهم عن 175 سم.

الحل/ نفرض أن X متغيراً عشوائياً يمثل أطوال الطلاب ويتبع توزيعاً طبيعياً وسطه $\mu = 160$ وانحرافه المعياري $\sigma = 10$ ومن الواضح في هذه الحالة أنه غير معياري.

وبتطبيق نظرية (6.7) نحصل على:

$$Z = \frac{X - 160}{10} \sim N(0,1)$$

وتكون الاحتمالات المطلوبة كما يلي:

$$1- P(X < 170) = P\left(\frac{X-160}{10} < \frac{170-160}{10}\right)$$

$$= P(Z < 1) = \Phi(1) = 0.8413$$

$$2- P(165 < X < 175) = P\left(\frac{165-160}{10} < Z < \frac{175-160}{10}\right)$$

$$= P(0.5 < Z < 1.5)$$

$$= \Phi(1.5) - \Phi(0.5)$$

$$= 0.9332 - 0.6915 = 0.2417$$

فتكون النسبة المطلوبة = $0.2417 \times 100\% = 24.17\%$

$$3- P(X > 175) = P\left(\frac{x-160}{10} > \frac{175-160}{10}\right)$$

$$= P(Z > 1.5) = 1 - \Phi(1.5)$$

$$= 1 - 0.9332 = 0.0668$$

فيكون عدد الطلبة = عدد الطلبة الكلي $\times$ احتمال الحدث

$$= 0.0668 \times 1000 = 66.8 \cong 67 \text{ طالب}$$

مثال 26.7: إذا كانت علامات طلبة الثانوية العامة في أحد الأعوام تتخذ توزيعاً طبيعياً

وسطه الحسابي 65 وانحرافه المعياري 15، فإذا علمت أن الجامعات الرسمية قبلت الطلبة

على المعدل التنافسي 10% من هؤلاء الطلبة. فما هو أدنى معدل قبل في الجامعات

الرسمية حسب المعدل التنافسي؟

نفرض أن X متغير عشوائي يمثل علامات الطلبة وعليه فإن X يتبع توزيعاً طبيعياً
وسطه $\mu = 65$ وانحرافه المعياري $\sigma = 15$ ومن الواضح أنه غير معياري،
وبتطبيق نظرية (6.7) نحصل على:

$$Z = \frac{X - 65}{15} \sim N(0,1)$$

نفرض أن أدنى علامة معيارية قُبِلَتْ في الجامعات الرسمية تساوي a فيكون 10% من
الطلبة حصلوا على درجات أكبر أو تساوي a وبالتالي نستطيع القول أن:

$$P(Z \geq a) = 0.10$$

وبناءً عليه يكون:

$$P(z \geq a) = 1 - \Phi(a) = 0.10$$

$$\Phi(a) = 0.9$$

و من جدول (2) للقيم الموجبة نجد أن $a = 1.28$ وهذه أدنى علامة معيارية قُبِلَتْ
في الجامعات الرسمية، و لتحويلها إلى علامة خام نستخدم علاقة التحويل:

$$a = \frac{X - \mu}{\sigma}$$

لنحصل على المطلوب:

$$a = \frac{X - \mu}{\sigma} \Rightarrow 1.28 = \frac{X - 65}{15}$$

$$X - 65 = 1.28 \times 15$$

$$X = 84.2$$

تمارين:

1- الجدول التالي يمثل توزيعاً احتمالياً لمتغير عشوائي X :

x	2	4	6	8
$p(x)$	0.1	k	$2k$	0.3

والمطلوب ايجاد ما يلي:

- قيمة الثابت k .
- تباين المتغير العشوائي.
- $P(X < 5)$

2- إذا كان X متغيراً عشوائياً له توزيع احتمالي معطى بالعلاقة:
 $f(x) = k(2x + 1), \quad x = 0, 1, \dots, 6$

أوجد ما يلي:

1- قيمة الثابت k .

2- $F(x)$

3- $P(0 \leq X \leq 4)$

4- $P(X > 3)$

5- التباين.

3- يرمي شخص قطعة نقد غير متزنة صُمِّمَتْ بحيث يكون احتمال ظهور صورة على

الوجه العلوي يساوي 0.8، أُلْقِيَت القطعة مرتين، فإذا كان الناتج صورتين يسجل 15

نقطة وإذا كان الناتج كتابتين يسجل - 30 نقطة وإذا كان الناتج وجهين مختلفين يسجل

- 10 نقاط، احسب توقع النقاط المسجلة.

4- إذا كان عدد الطلاب الذين يجيدون استخدام الحاسوب يساوي 7 أضعاف عدد الطلاب

الذين لا يجيدون استخدام الحاسوب، فإذا اختير 10 طلاب عشوائياً فما احتمال أن

يكون منهم 7 يجيدون استخدام الحاسوب.

5- إذا كان $X \sim b(4, 0.3)$ احسب ما يلي:

- توقع وتباين X .
- $P(0 < X \leq 3)$

6- تفخر شركة أن نسبة المعيب 0.01 من انتاجها اليومي الذي يبلغ 25 قطعة. أوجد ما يلي:

- التوزيع الاحتمالي للمتغير العشوائي X الذي يعبر عن عدد القطع المعيبة.
- توقع وتباين المتغير العشوائي.
- $P(X > 8)$

7- معدل عدد حوادث السيارات على طريق صحراوي 5 في الشهر. ما احتمال عدم حدوث أي حادث على ذلك الطريق في شهر معين؟ ما احتمال حدوث 4 حوادث أو أقل في شهر معين؟

8- حجر نرد كتب على ثلاثة وجوه منه الرقم 2 وعلى وجهين الرقم 3 وعلى الوجه الأخير الرقم 6، فإذا أُلقي هذا الحجر 20 مرة أوجد احتمال ظهور الرقم 3 خمس مرات وأوجد توقع ظهور الرقم 2.

9- تسحب كرات الواحدة تلو الأخرى مع الإرجاع من صندوق يحتوي على كرتين بيضاوين وكرة سوداء، أوجد أقل عدد من الكرات يجب سحبه من الصندوق حتى يكون احتمال ظهور كرة واحدة سوداء على الأقل يزيد عن 0.79

10- بفرض أن عدد الأخطاء في قطعة الكترونية يتبع توزيع بواسون بمعدل 3 أخطاء في القطعة، احسب احتمال عدم وجود أخطاء في قطعة معينة، ثم احسب احتمال أن هذه القطعة تحتوي على ثلاثة أخطاء على الأقل.

11- قطعنا نقد احتمال ظهور الصورة في كل منهما 0.6، إذا أُلقيت القطعتان معاً 25 مرة أوجد احتمال أن تظهر كتابتان 5 مرات فقط. ثم أوجد احتمال الحصول على صورة وكتابة 7 مرات فقط.

- 12- تستقبل غرفة الطوارئ ما معدله 49 حالة كسر أسبوعياً. كم معدل حالات الكسر اليومية التي تصل غرفة الطوارئ، وإذا عُلِمَ أن غرفة الطوارئ مجهزة لاستقبال 10 حالات كسر يومياً فما احتمال أن تمر غرفة الطوارئ بأزمة؟
- 13- تقدم طالب لامتحان اختيار من متعدد مكون من 20 فقرة ولكل فقرة خمسة خيارات، فإذا أجاب الطالب على الامتحان عشوائياً أوجد احتمال أن يجيب على 12 فقرة. وإذا كان لكل فقرة ثلاث درجات أوجد احتمال أن يحصل الطالب على درجة 30.

14- إذا كان X متغيراً عشوائياً متصلاً له توزيع احتمالي معطى بالعلاقة:

$$f(x) = \frac{x}{c}, \quad 0 < X < 2$$

أوجد قيمة c وتوقع المتغير العشوائي. ما هي القيمة المتوقعة لمربعات قيم X ؟

15- إذا كان X متغيراً عشوائياً له توزيع تراكمي معطى بالعلاقة :

$$F(x) = \frac{(x-1)^2}{k}, \quad 1 < x < 3$$

أوجد قيمة k ثم اوجد اقتران الكثافة الاحتمالي للمتغير العشوائي.

- 16- إذا كانت أوزان أكياس الطحين حسب تعليمات وزارة التموين تخضع للتوزيع الطبيعي بوسط 50 كغم وانحراف معياري 4 كغم. إذا اختير كيس عشوائياً، فما احتمال أن يزيد وزنه عن 53 كغم؟ وما احتمال أن يكون وزنه أقل من 48 كغم؟

17- إذا كان $Z \sim N(0,1)$ أوجد:

$$P(-1 < Z < 1.5), \quad P(Z > 2.5)$$

$$P(1.25 < Z < 3), \quad P(-3 < Z < -1)$$

18- إذا كان $Z \sim N(0,1)$ أوجد b في كل مما يلي:

$$P(Z < b) = 0.90, \quad P(-b < Z < b) = 0.9, \quad P(Z > b) = 0.10$$

19- إذا كان $X \sim N(70, 625)$ أوجد:

• $P(X > 73)$, $P(65 < X < 72)$

• قيمة a في الحالات التالية:

$P(X < a) = 0.9$, $P(X > a) = 0.05$

20- إذا كانت أوزان منتج في مصنع تتبع التوزيع الطبيعي، وكان 5% من المنتج يزيد وزنه عن 85 جرام، وكان 10% من المنتج يقل وزنه عن 25 جرام. أوجد الوسط الحسابي والتباين لأوزان المنتج.

21- اختر الإجابة الصحيحة فيما يلي:

• عند القاء قطعة نقود متزنة ثلاث مرات فإن عدد مرات ظهور الصورة في الرميات الثلاث يعتبر:

أ) ثابت ب) متغير منفصل ج) متغير متصل د) ليس مما سبق

• إذا كان التوزيع الاحتمالي للمتغير العشوائي X يعطى بالعلاقة:

$f(x) = kx + 1$, $x = 1, 2$ فإن قيمة k تساوي:

أ) 2- ب) 0.5- ج) 3- د) $\frac{1}{3}$ -

• إذا كان التوزيع الاحتمالي للمتغير العشوائي X يعطى بالعلاقة:

$f(x) = \frac{1-\frac{1}{8}x}{3}$, $x = -1, 1, a, 5$ فإن قيمة a تساوي:

أ) 2 ب) 3 ج) 4 د) 4.5

• إذا كان توقع ذي الحدين يساوي 2.8 واحتمال النجاح في المحاولة الواحدة يساوي

0.7 فإن عدد مرات تكرار التجربة n يساوي:

أ) 4 ب) 5 ج) 6 د) 7

- إذا كان X متغيراً عشوائياً وكان تباين $(3X)$ يساوي 18 فإن تباين X يساوي:

أ) 6 ب) 4 ج) 2 د) 0

- إذا كان Y, X متغيران عشوائيان بحيث وكان $E(X) = E(y)$ فإن $E(X - Y)$ يساوي:

أ) 4 ب) 0.25 ج) صفر د) 2

- إذا كان $Var(X) = 1$, $Var(3Y) = 9$ فإن $Var(X + Y)$ يساوي:

أ) 18 ب) 8 ج) 2 د) 12

- إذا كان Ω فضاء عينة لتجربة عشوائية فإن $P(\Omega)$ يساوي:

أ) 10 ب) 7 ج) 1 د) 5

- إذا كان $X \sim PO(3)$ فإن $E(X)$ يساوي:

أ) 2 ب) صفر ج) 3 د) 6

- إذا كان $X \sim N(4,1)$ فإن الانحراف المعياري للتوزيع يساوي:

أ) 4 ب) 10 ج) 1 د) 5

المتغير العشوائي:

هو اقتران مجاله فضاء العينة ومجاله المقابل مجموعة الأعداد الحقيقية.

التوزيع الاحتمالي:

هو اقتران مجاله مدى المتغير العشوائي ومجاله المقابل $[0,1]$

التوزيع التراكمي للمتغير العشوائي المنفصل:

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ 1 & x_n \leq x < \infty \end{cases}$$

التوقع الرياضي للمتغير العشوائي المنفصل:

$$\mu_X = E(X) = \sum_{x \in X(\Omega)} x P(x)$$

نظريات على التوقع الرياضي:

$$(c) = c \quad , \quad E(cX) = cE(X) \quad , \quad E(X \pm Y) = E(X) \pm E(Y)$$

إذا كان X, Y متغيران عشوائيان مستقلان، فإن: $E(XY) = E(X)E(Y)$

تباين المتغير العشوائي المنفصل:

$$Var(X) = \sigma_X^2 = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum_{x \in X(\Omega)} x^2 P(x) \quad \text{حيث:}$$

نظريات على التباين:

إذا كان X متغيراً عشوائياً، c ثابت اختياري فإن:

$$Var(c) = 0, \quad Var(cX) = c^2 Var(X)$$

إذا كان X, Y متغيرين عشوائيين مستقلين، فإن:

$$Var(X \pm Y) = Var(X) + Var(Y)$$

توزيع ذي الحدين:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

$$E(X) = np, \quad \sigma^2 = np(1 - p)$$

توزيع بواسون:

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

$$E(X) = \lambda, \quad \sigma^2 = \lambda$$

التوزيع التراكمي للمتغير العشوائي المتصل:

$$F(x) = \begin{cases} 0, & x \leq a \\ \int_a^x f(t) dt, & a < x < b \\ 1, & x \geq b \end{cases}$$

توقع المتغير العشوائي المتصل:

$$\mu_X = E(X) = \int_a^b x f(x) dx$$

تباين المتغير العشوائي المتصل:

$$Var(X) = \sigma_X^2 = E(X^2) - [E(X)]^2$$

حيث:

$$E(X^2) = \int_a^b x^2 f(x) dx$$

نظرية تحويل المتغير العشوائي غير المعياري لمتغير عشوائي معياري:

إذا كان:

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

فإن:

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

حيث يسمى المقدار

$$Z = \frac{X - \mu}{\sigma}$$

بالدرجة المعيارية

الإحصاءات الحيوية Vital Statistics

8

يعتبر الاهتمام بالإحصاءات الحيوية من أهم ما يميز الدول المتقدمة عن غيرها من الدول، فهي البداية للتنمية الحقيقية والتقدم في المجتمعات. سنعرض خلال هذا الفصل بعض الإحصاءات الحيوية الهامة مثل تعداد السكان، إحصاءات المواليد والوفيات، الهجرة، الجنسية الحالة الزوجية والحالة الاقتصادية وسنذكر بعض المعدلات المرتبطة بها.

الأهداف

ينتظر منك عزيزي الطالب بعد الانتهاء من دراسة هذا الفصل أن تكون قادراً على:

- ذكر بعض الإحصاءات الحيوية الهامة.
- التحدث باختصار عن تلك الإحصاءات الحيوية.
- حساب معدلات التوالد والوفيات والنشاط الاقتصادي.
- حساب عدد السكان بين سنوات التعداد.

الاحصاءات الحيوية من الموضوعات الهامة التي تهتم بها معظم الدول، حيث تنشئ الدولة هيئات وأقسام خاصة تختص بجمع وتحليل البيانات التي تشمل كل الظواهر التي تحدث للبشر كتعداد السكان، الولادة، الوفاة، الهجرة، الزواج، الطلاق، المرض والامية. وهي تعتمد عادة على وثائق مسجلة بحكم القانون وخاصة إحصاءات المواليد والوفيات والتبليغ عن الأمراض المعدية وإحصاءات الزواج والطلاق.

ومن أهم أهداف الاحصاءات الحيوية:

- 1- تنظيم وتحسين الخدمات العامة.
- 2- قياس المستوى الصحي للمجتمع.
- 3- التخطيط في جميع المجالات التعليمية والصحية والاقتصادية.
- 4- البحث العلمي في جميع الميادين كالاقتصاد والطب والتعليم.
- 5- المقارنة الدولية وخاصة في المجال الصحي.

وسنتحدث عن بعض الاحصاءات الحيوية بشيء من الاختصار.

تعداد السكان Population

يعرف تعداد السكان بأنه عملية حصر جميع الأفراد في مجتمع معين وذلك في كل أماكن تواجد هؤلاء الأفراد خلال فترة زمنية معينة. ويتم تجميع البيانات الاحصائية عادة من كل فرد من هؤلاء الأفراد وذلك لمعرفة بعض الصفات الاساسية المهمة التي يتم دراستها والاهتمام بها من قبل الدولة والجهات المختصة.

طرق التعداد السكاني

يتم التعداد السكاني عادة بإحدى الطريقتين التاليتين:

- 1- **طريقة التعداد الواقعي:** حيث يتم حصر الأفراد في مكان إقامتهم في اللحظة المحددة للتعداد سواء كانوا من سكان هذا المكان بصفة دائمة أو بصفة مؤقتة. مثل نزلاء دور الرعاية والمستشفيات، ويسمى هذا النوع من التعداد بالتعداد الواقعي أو الفعلي.
- 2- **طريقة التعداد النظري:** وفي هذه الطريقة يتم حصر الأفراد حسب مكان إقامتهم الدائمة بغض النظر عن مكان وجودهم في اللحظة المحددة للتعداد وتسمى هذه الطريقة بالتعداد النظري أو الاعتيادي.

إحصاءات المواليد Statistics of births

تعتبر إحصاءات المواليد إحدى الركائز الأساسية في مجال الإحصاءات الحيوية إذ تمثل عنصراً هاماً ورئيساً في معرفة حركة السكان زيادةً أو نقصاً. كما تدخل كعنصر أساسي في تقدير عدد السكان في سنوات ما بين التعداد. ولذلك تحتم قوانين غالبية الدول على ضرورة تسجيل المواليد رسمياً خلال مدة معينة من وقت الولادة. وتقوم الجهات المسؤولة عن الإحصاء في الدولة بإعداد الإحصاءات الخاصة بالمواليد إجمالاً على مستوى الدولة وتفصيلاً على مستويات المناطق الجغرافية وعلى مستوى النوع (ذكور وإناث) وفي شهور السنة لمعرفة أي الشهور تكثر فيها المواليد عن غيرها.

والبيانات التي يتم تسجيلها عن كل مولود تختلف من دولة لأخرى حسب درجة انتشار الوعي الإحصائي والتعليم والثقافة والتقاليد والعادات، وإن كانت البيانات الأساسية اللازم تسجيلها عن كل مولود لا تخرج عن: تاريخ الميلاد، اسم المولود، الجنس، الجنسية، محل الميلاد، المولود حي أم ميت، اسم الأب ومهنته، اسم الأم، الديانة للأب وللأم.

وإحصاءات المواليد تستخدم في حساب الكثير من المعدلات الهامة التي تستخدم كمقاييس إحصائية في مجالات الدراسات السكانية نذكر منها على سبيل المثال المعدلات الإجمالية الآتية:

أ - معدل المواليد الخام: وهو معدل أولى وتقريبي ويساوي:

$$1000 \times \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد السكان في منتصف السنة}}$$

= معدل المواليد الخام



ويتم الحصول علي بيان عدد المواليد أحياء أثناء السنة من واقع إحصاءات المواليد أما عدد السكان فيتم الحصول عليه إما من تعداد السكان أو بإحدى طرق تقدير السكان بين سنوات التعداد (والتي سيجيئ شرحها فيما بعد).

ويعاب على هذا المعدل أنه يأخذ في حسابه جميع السكان دون اعتبار النوع أو السن وهو بذلك لا يعطي مقياساً حقيقياً وصحيحاً لحالة الخصوبة الحقيقية للسكان، إذ من المعروف أن عدد المواليد في أي بلد لا يتوقف على عدد الأطفال والشيوخ بل على عدد النساء اللاتي في سن الحمل دون باقي السكان.

ب- **معدل الخصوبة العام:** وهذه المعدل يدخل في حسابه عدد النساء في سن الحمل من 15 إلى 50 سنة فقط دون باقي السكان وهو بذلك يتلافى العيوب التي تؤخذ على معدل المواليد الخام ويساوي:

$$\text{معدل الخصوبة الخام} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد النساء في سن الحمل (15-50 سنة)}} \times 1000$$



إلا أن هذا المعدل يؤخذ عليه أيضاً أن النساء في سن الحمل لسن كلهن متزوجات لذلك من الأفضل استبعاد النساء غير المتزوجات وفي سن الحمل من حساب المعدل.

ج - **معدل التوالد:** وهذا المعدل يأخذ في اعتباره عدد النساء في سن الحمل والمتزوجات فقط دون باقي النساء وهو بذلك يتلافى العيوب التي تؤخذ على معدل الخصوبة العام ويساوي:

$$\text{معدل التوالد} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد النساء المتزوجات في سن الحمل}} \times 1000$$



ومما لا شك فيه أن هذا المعدل أدق من معدل الخصوبة العام وأكثر واقعياً ومنطقاً. ويلاحظ أننا أخذنا في حساب المعدلات السابقة عدد المواليد أحياء دون المواليد الموتي وهذا أمر طبيعي إذ أن المواليد الموتي لا تؤثر على نمو السكان وإن كانت تعطى مؤشراً هاماً عن الحالة الصحية للأمهات ومقدار الرعاية الصحية والطبية التي تُقدّم لهن أثناء مدة الحمل.

مثال 1.8: البيانات التالية خاصة بإحدى الدول عام 1960 (بالمليون).

عدد المواليد أحياء	عدد النساء في سن الحمل	عدد النساء المتزوجات في سن الحمل	تقدير عدد السكان في منتصف السنة
1.1	5.3	4.2	26.1

احسب:

- 1- معدل المواليد الخام.
- 2- معدل الخصوبة العام.
- 3- معدل التوالد.

$$\text{معدل المواليد الخام} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد السكان في منتصف السنة}} \times 1000$$

$$42.15 = 1000 \times \frac{1.1}{26.1} =$$

$$\text{معدل الخصوبة الخام} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد النساء في سن الحمل (15-50 سنة)}} \times 1000$$

$$207.55 = 1000 \times \frac{1.1}{5.3} =$$

$$\text{معدل التوالد} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد النساء المتزوجات سن في الحمل}} \times \frac{1.1}{4.2} = 1000 \times 261.9$$

إحصاءات الوفيات Mortality statistics

تعتبر إحصاءات الوفيات عنصراً هاماً في الإحصاءات الحيوية فهي علاوة على أنها تعتبر من العوامل المؤثرة في تقدير عدد السكان بين سنوات التعداد، فإنها تُكوّن مع إحصاءات المواليد الركائز الأساسية في حركة السكان.

وكما تحتم القوانين ضرورة تسجيل المواليد فإنها تحتم كذلك ضرورة تسجيل الوفيات رسمياً خلال مدة معينة من لحظة الوفاة. وهذه المدة تتفاوت من دولة لأخرى. وتقوم الجهات المعنية بالدولة بإعداد الإحصاءات الخاصة بالوفيات حسب التقسيمات المختلفة.

والبيانات الأساسية التي يتم تسجيلها في حالة الوفاة هي: تاريخ الوفاة، اسم المتوفى، الجنس، عمر المتوفى، المهنة، سبب الوفاة، مكان الوفاة ومحل الإقامة المعتاد.

وإحصاءات الوفيات تستخدم في حساب بعض المعدلات الهامة التي تستخدم كمقاييس إحصائية في مجال الدراسات الإحصائية في مجال الدراسات السكانية نذكر منها على سبيل المثال المعدلات التالية:

أ - معدل الوفيات الخام ويساوي:

$$\text{معدل الوفيات الخام} = \frac{\text{عدد الوفيات في البلد أثناء السنة}}{\text{عدد السكان في منتصف السنة}} \times 1000$$



ويتم الحصول على بيان عدد الوفيات من واقع إحصاءات الوفيات. أما عدد السكان فيتم الحصول عليه إما من تعدادات السكان أو بإحدى طرق تقدير السكان بين سنوات التعداد.

ويعتبر هذا المعدل من أهم المقاييس التي تستخدم لدراسة الوفيات ويستخدم لمعرفة التغير الذي يطرأ على الحالة الصحية في بلد ما خلال فترات زمنية متقاربة. ولكن يؤخذ على هذا المعدل انه لا يُعطى نمطاً صحيحاً ودقيقاً للوفاة بالنسبة لفئات السن المختلفة وبالنسبة للجنس في المجتمع الواحد كذلك لا يُعطى مؤشراً دقيقاً للمقارنة بين الحالة الصحية للمجتمعات المختلفة في التركيب العمري والنوعي للسكان.

ب - المعدل النوعي للوفيات

وهذا المعدل يمكن من الحصول على معدل الوفيات لكل فئة من فئات السن أو لكل فئة من فئات الجنس كلاً على حدة، أو لكل جنس بفئة عمرية محددة. فإذا توافرت البيانات اللازمة لتكوين هذا المعدل وهي الوفيات مبوبة على سبيل المثال حسب فئات السن لكل من الذكور

والإناث وعدد السكان لفئات السن المختلفة لكل من الذكور والإناث أمكن الحصول على معدل وفيات لكل فئة من فئات السن للذكور والإناث كل على حدة.

$$\text{المعدل النوعي للوفيات بالنسبة للسن} = \frac{\text{عدد الوفيات في البلد خلال السنة لكل فئة من فئات السن}}{\text{عدد السكان في البلد في منتصف السنة لكل فئة من فئات السن}} \times 1000$$

فمثلاً معدل وفيات الذكور في فئة السن (20-25) سنة يتم حسابه من العلاقة التالية:

$$\text{المعدل النوعي للوفيات بالنسبة للسن} = \frac{\text{عدد الوفيات الذكور في البلد خلال السنة في فئة السن (20-25)}}{\text{عدد السكان في البلد في منتصف السنة في فئة السن (20-25)}} \times 1000$$

وهذا المعدل يُحسَب لجميع فئات السن عدا فئة السن أقل من سنة فيحسب لها معدل خاص سنذكره فيما بعد.

ووجود معدلات نوعية للوفيات بهذا الشكل يُمكن من معرفة نمط الوفيات ودراسة أي فئة عمرية معرضة للوفاة أكثر من غيرها وهكذا، ويربط هذه المعدلات ببعض المتغيرات الأخرى كأسباب الوفاة والمهنة وغيرها من المتغيرات الاقتصادية والاجتماعية الأخرى يمكن الوقوف على الحالة الصحية ومعرفة الأسباب الحقيقية ووضع خطط مناسبة لتحسين الحالة الصحية للسكان تتضمن المشروعات الصحية والوقائية وغيرها. علاوة على ذلك المعدل النوعي للوفيات يُمكن من المقارنة السليمة والدقيقة بين الحالة الصحية للدول المختلفة.

ج- معدل الوفيات الأطفال الرضع

قبل شرح هذا المعدل وأهميته لابد من إعطاء تعريف واضح لوفيات الأطفال الرضع حيث تُعرف وفيات الأطفال الرضع بأنها: "وفيات الأطفال الذين لم يبلغوا السنة الأولى من العمر"، وتمثل هذه الوفيات نسبة كبيرة من الوفيات في البلاد النامية. ويحسب هذا المعدل من العلاقة التالية:

لا تنسى



$$\text{معدل وفيات الأطفال الرضع} = \frac{\text{عدد وفيات الأطفال الرضع في البلد خلال السنة}}{\text{عدد المواليد أحياء في البلد خلال نفس السنة}} \times 1000$$

وهذا المعدل له أهمية كبيرة، إذ يُعْتَبَر مقياساً حساساً للحالة الصحية ومؤشراً دقيقاً لحالة البيئة من حيث توفر الظروف الصحية المناسبة والتغذية السليمة الخ لأن الأطفال الرضع يتأثرون بشدة بأي مرض نظراً لضعف احتمالهم لمقاومة الأمراض. لذلك فدراسة وتتبع هذا المعدل لسلسلة من السنوات له أهمية كبيرة في الوقوف على المستوي الصحي للسكان في الدولة وكذلك في المناطق المختلفة داخل الدولة الواحدة كالمناطق الصناعية وغيرها وكذلك لمعرفة أسباب الوفاة بين الأطفال في هذه السن ومحاولة التغلب عليها.

مثال 2.8: الجدول التالي يبين البيانات الخاصة بإحدى الدول في سنة من السنوات بالآلاف والمطلوب إيجاد كلاً من:

- معدل الوفيات الخام.
- معدل وفيات الأطفال في سن الرضاعة.

42000	عدد السكان في منتصف السنة
600	عدد الوفيات خلال العام
100	عدد وفيات الاطفال الرضع
1500	عدد المواليد الأحياء

الحل/

$$142.86 = 1000 \times \frac{600}{4200} = 1000 \times \frac{\text{عدد الوفيات في البلد أثناء السنة}}{\text{عدد السكان في منتصف السنة}} = \text{معدل الوفيات الخام}$$

$$1000 \times \frac{\text{عدد وفيات الأطفال الرضع في البلد خلال السنة}}{\text{عدد المواليد أحياء في البلد خلال نفس السنة}} = \text{معدل وفيات الأطفال الرضع}$$

$$66.76 = 1000 \times \frac{100}{1500} =$$

احصاءات الهجرة Migration statistics

تشمل هذه الاحصاءات البيانات الخاصة بالأفراد الذين ينتمون للبلد (المواطنون) والذين يهاجرون من البلد بشكل نهائي وكذلك الأجانب القادمون للبلد من بلدان أخرى لغرض الإقامة لفترة معينة أو للإقامة الدائمة. حيث تقوم إدارة خاصة تشكلها الدولة لمتابعة حركة الهجرة من وإلى البلد، حيث يتم تسجيل بيانات عن المهاجرين مثل الجنس والعمل والمؤهل العلمي والحالة الاجتماعية والديانة، والجدير بالذكر أن هناك نوع من الهجرة يسمى هجرة العقول ويكثر هذا النوع في البلاد النامية، الفقيرة والمتخلفة أو البلاد التي لا تشهد استقراراً أمنياً حيث يهاجر العلماء والخبراء لبلاد يجدوا فيها من يهتم بهم ومستوى معيشة أفضل، وهذا النوع من الهجرة يؤثر سلباً في البلد المهاجر منها وإيجاباً في البلد المهاجر إليها.

وقد يهاجر من البلد العاطلون عن العمل بسبب عدم توفر وظائف لهم ويعملون في بلاد غنية ثم يعودون بالأموال لبلدهم الأم من أجل عمل مشاريع خاصة بهم مما يعود بالفائدة على بلادهم.

احصاءات الجنسية Statistics of citizenship

يفيد بيان الجنسية وتحديد أعداد الأجانب وجنسياتهم وأعمارهم ومؤهلاتهم العلمية وعملهم الدولة في عدة نواحي أهمها:

- معرفة عدد الأجانب وخصائصهم من حيث الجنس والسن والحالة التعليمية ومعرفة الهجرة الى البلاد وفي أي مرحلة من مراحل السنة تكون هذه الهجرة ونوعية المهاجرين من وإلى الدولة.
- معرفة مدى مساهمة المهاجرين في قوة العمل وفي أي الأنشطة تكون هذه المساهمة أكبر من غيرها.
- دراسة مسائل الأمن الداخلي وتنظيم رعاية الأجانب.
- تتبع التغير في عدد وأوضاع الأجانب من تعداد لآخر.

احصاءات الحالة الزوجية Statistics of Marital Status

بيانات الحالة الزوجية هامة بالنسبة للدولة لأنها تساعد على معرفة اتجاهات الخصوبة ودرجة النمو السكاني وكذلك مدى الإقبال أو الإحجام على الزواج في فئات السن المختلفة. وكذلك معرفة مدى التغير في متوسط سن الزواج. كما يفيد هذا البيان في تقدير الاحتياجات من المساكن والسلع الاستهلاكية والخدمات الأخرى هذا علاوة على أن معرفة عدد المطلقات والأرامل يؤدي إلى التعرف على المشاكل الاجتماعية لهن ويساعد على وضع الحلول المناسبة لها.

احصاءات الحالة الاقتصادية للسكان

ويقصد بها النشاط الاقتصادي والحالة العملية للسكان. بصفة عامة يمكن تقسيم السكان من الناحية الاقتصادية الى قسمين رئيسين:

أ - القوة البشرية.

ب - خارج القوى البشرية.

القوة البشرية: هي ذلك الجزء من السكان الذين يمكن استغلاله في النشاط الاقتصادي ويشمل السكان من سن 15 الى 65 سنة وتختلف هذه الفترة من دولة لأخرى، والقوة البشرية تنقسم بدورها الى قوة العمل وخارج قوة العمل.

خارج القوة البشرية: فهم الأفراد الغير قادرين على العمل مثل الأطفال والعجزة الذين لا يمكنهم أداء عمل مثمر بسبب عاهات مقعدة او أمراض مزمنة أو كبر السن.

وتفيد بيانات الحالة الاقتصادية في النواحي الآتية:

- تعطي صورة واضحة لهيكل النشاط الاقتصادي للدولة.
- معرفة القوى العاملة وتوزيعها على المهن المختلفة في المناطق الجغرافية وكذلك معرفة عدد المشتغلين والعاطلين عن العمل في الأنشطة المختلفة.
- دراسة تطور الصناعة ومقوماتها.
- تحديد نسبة القوى العاملة المشتغلة في قطاع معين كالزراعة مثلاً إلى مثيلاتها المشتغلة في القطاعات الأخرى.
- استخراج بعض المؤشرات والمعدلات الهامة، مثل معدل النشاط الاقتصادي وهو:

$$\text{معدل النشاط الاقتصادي} = \frac{\text{عدد المكتسبين}}{\text{عدد السكان}} \times 1000$$



حيث يُعرَّف المُكْتَسِب بأنه الشخص الذي له نشاط اقتصادي أو الشخص الموجود داخل سوق العمل أو الشخص الذي يبذل مجهوداً في سبيل إنتاج سلعة أو خدمة داخل سوق

العمل. وهذا المعدل يختلف حسب القطاعات الاقتصادية والأعمار والنوع ومستوى التقدم الصناعي.

مما سبق نجد أن تعداد السكان يعتبر من أهم عناصر الإحصاءات السكانية حيث يعتبر نقطة الارتكاز في هذه الإحصاءات فهو يصور السكان من حيث العدد والخصائص المختلفة في لحظة معينة، وحيث أن السكان يتصفون بالحركة المستمرة نجد أن بيانات التعداد تمثل اللحظة التي تُجمع فيها البيانات فقط. بمعنى أنه بعد تاريخ التعداد تحدث ولادات ووفيات وزواج وطلاق وتغير في المهن والأنشطة الاقتصادية والحالة التعليمية والعملية الخ. أي باختصار تختلف صورة وتكوين المجتمع، وعلى ذلك إذا كنا سنعتمد على تعدادات السكان فقط فهذا يعني أننا سننتظر عشر سنوات أخرى حتى يحل موعد إجراء تعداد آخر، لذلك نجد أن المسؤولين عن التخطيط في الدول يعتمدون على التعدادات كنقطة ارتكاز لمعرفة صورة واضحة على المجتمع في لحظة التعداد ثم يعتمد في فترة بعد سنوات التعداد على العديد من الإحصاءات ذات الدورة القصيرة والتي غالباً ما تكون سنوية لمعرفة التغيرات التي تطرأ على السكان منذ تاريخ التعداد.

تقدير السكان بين سنوات التعداد

أهمية التقديرات السكانية:

للتقديرات السكانية وإسقاطاتها في المستقبل أهمية خاصة في تخطيط السياسة العامة للدولة في مختلف الميادين، فلا بد من معرفة أعداد السكان عند تصميم البرامج والمشروعات المختلفة في مجال الصناعة والتخطيط لاستيعاب القوى العاملة في المستقبل ومعرفة الكميات التي يحتاجها السكان من السلع والخدمات المختلفة.

وتختلف الاحتياجات البشرية في المجتمعات السكانية باختلاف فئات أعمار الأفراد وتتفاوت تبعاً لذلك أعباء الدول وسياساتها الداخلية لمواجهة تلك الاحتياجات بتوجيه جهود خاصة نحو توفير سلع استهلاكية أو خدمات من أنواع معينة. فمرحلة الطفولة المبكرة، تحتاج إلى

خدمات صحية وقائية علاجية وهو أمر تشترك فيه كل المجتمعات على السواء وإن كانت حاجة المجتمعات النامية إلى هذه الخدمات أشد بسبب ارتفاع عدد المواليد فيها. لذلك فإن تقدير أعداد المواليد في كل عام لمدة طويلة أمر ضروري للتخطيط لمثل هذه الخدمات الصحية.

ويعقب هذه المرحلة مرحلة التعليم، حيث يحتاج التخطيط التربوي ضمن إطار التخطيط العام إلى بيانات مفصلة عن أعداد الأطفال المطلوب توفير الخدمات التعليمية لهم، حيث يتم تحديد أعداد المدارس والفصول والمدرسين اللازمين على أساس هذه الأعداد.

والمرحلة التالية لذلك هي مرحلة الدخول في قوة العمل، فتحديد أعداد الأفراد الذين يدخلون سوق العمل سنوياً أمر ضروري لتهيئة فرص العمل المثمر لهم في حدود احتياجات القطاعات المختلفة وعلي أساس تحقيق التوازن بين العرض والطلب في هذا السوق.

مصادر البيانات اللازمة لتقدير السكان:

تعتبر نتائج تعدادات السكان المصدر الرئيسي للتوزيعات السكانية حسب الخصائص المختلفة، وأبسط المقاييس التي تستخرج من تعدادات السكان هي معدلات النمو السنوي للسكان وهذه المعدلات تسمح بالحصول على التقديرات الدقيقة إلى حد كبير لأعداد السكان في المستقبل القريب، وخصوصاً إذا كان الغرض هو التقدير الإجمالي للسكان، وجداول الحياة التي تنشأ في سنوات التعداد تعتبر الأداة الأساسية للتحليل الديموغرافي وتستخدم لتقدير أعداد السكان في المستقبل لدرجة كبيرة من الدقة.

إن جميع المصادر اللازمة للتقديرات السكانية تعتمد أساساً على تعدادات السكان وحيث أن عملية تعداد السكان تتطلب الكثير من التكاليف الباهظة والمجهودات المضنية فلا يعقل القيام بهذه العملية سنوياً. وللتغلب على ذلك تستخدم نماذج رياضية لتقدير عدد السكان بين سنوات التعداد مع ملاحظة أن هذه التقديرات تمثل عدد السكان في منتصف السنة التي يجري فيها التقدير. فإذا توفر لدينا عدد السكان في تعدادين متتاليين فإنه يمكن تقدير عدد السكان في

أي تاريخ بين هذين التعدادين أو خارجهما بإحدى طرق التقدير مثل طريقة ثبات مقدار الزيادة، طريقة ثبات معدل الزيادة وطريقة المنحنى الأسّي وسنتحدث هنا عن طريقة ثبات مقدار الزيادة.

طريقة ثبات مقدار الزيادة:

هذه الطريقة تفترض ثبات مقدار الزيادة السنوية في عدد السكان من سنة إلى أخرى أي

أن عدد السكان يتزايد بمقدار ثابت سنوياً، وبفرض أن:

عدد السكان في التعداد الأول x_0

عدد السكان في التعداد الثاني x_1

الفترة الزمنية بين التعدادين t

فإن:

مقدار الزيادة السنوية في السكان

$$y = \frac{x_1 - x_0}{t}$$



ويكون عدد السكان x بعد الفترة n سنة من التعداد الأول معطى بالعلاقة:

عدد السكان بعد n سنة من التعداد الأول

$$x = x_0 + ny$$



وفيما يلي مثال يوضح هذه الطريقة.

مثال 3.8: إذا كان تعداد السكان في إحدى البلاد في مارس 1947 هو 19 مليون نسمة وفي سبتمبر 1960 هو 26 مليون. فأوجد عدد السكان التقديري في مايو 1966.

الحل/

من السؤال نجد أن:

$$x_0 = 19 \quad , \quad x_1 = 26$$

الفترة الزمنية بين تاريخي التعدادين = سبتمبر 1960 - مارس 1947 = 13.5 سنة
مقدار الزيادة السنوية للسكان بين التعدادين:

$$y = \frac{x_1 - x_0}{t} = \frac{26 - 19}{13.5} = 0.519$$

$$n = \text{مايو 1966} - \text{مارس 1947} = 19 \text{ سنة} + 2 \text{ شهر}$$

و بالتالي يكون عدد السكان التقديري x في مايو 1966 يساوي:

$$x = x_0 + ny$$

$$x = 19 + 19 \frac{2}{12} \times 0.519 = 19 + 9.948 = 28.948$$

وبالتالي يقدر عدد السكان في مايو 1966 حوالي 28.948 مليون نسمة.

تمارين:

- 1- ماذا يقصد بتعداد السكان، وما هي أغراضه؟
- 2- ما هي كثافة السكان، وكيف يتم حسابها؟ (بحث)
- 3- ما أهمية التركيب العمري والنوعي للسكان؟
- 4- ماذا يقصد بالحالة الاقتصادية للسكان، وما هي أوجه الاستفادة منها؟
- 5- تكلم باختصار عن معدلات المواليد، مع ذكر مزايا وعيوب كل منها؟
- 6- تكلم باختصار عن معدلات الوفيات، مع ذكر مزايا وعيوب كل منها؟
- 7- ما أهمية تقدير السكان بين سنوات التعداد؟
- 8- بلغ تعداد السكان في إحدى الدول 60 مليون نسمة منتصف عام 1970، بينما كان تعداد هذه الدولة 52 مليون نسمة في منتصف عام 1965. قدر عدد السكان في هذه الدولة في يوليو 1972.

9- البيانات التالية خاصة بإحدى البلاد عام 1980 وهي مقدرة بالمليون:

- عدد المواليد أحياء خلال السنة = 1.8
 - تقدير عدد السكان في منتصف السنة = 26
 - عدد الوفيات أثناء السنة = 0.5
 - عدد وفيات الأطفال الرضع خلال السنة = 0.1
 - عدد النساء في سن الحمل خلال السنة = 7.4
 - عدد النساء المتزوجات في سن الحمل خلال السنة = 5.8
- احسب في ضوء ما سبق من بيانات ما يلي:

- معدل المواليد الخام.
- معدل الخصوبة العام.

- معدل التوالد.
- معدل الوفيات الخام.
- معدل وفيات الأطفال الرضع.

10- إذا كان عدد السكان في مدينة ما سنة 1970 هو 570000 نسمة فما هو عدد السكان المتوقع عام 1976 علماً بأن الزيادة السنوية ثابتة من سنة لأخرى بمقدار 1500 نسمة سنوياً.

11- إذا كان عدد السكان في مدينة ما سنة 2000 هو 640000 نسمة وفي عام 2010 أصبح عدد السكان 820000 نسمة فما هو عدد السكان المتوقع عام 2015.

الخلاصة

معدلات إحصاءات المواليد

$$\text{معدل المواليد الخام} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد السكان في منتصف السنة}} \times 1000$$

$$\text{معدل الخصوبة الخام} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد النساء في سن الحمل (15-50 سنة)}} \times 1000$$

$$\text{معدل التوالد} = \frac{\text{عدد المواليد أحياء في البلد أثناء السنة}}{\text{عدد النساء المتزوجات في سن الحمل}} \times 1000$$

معدلات إحصاءات الوفيات

$$\text{معدل الوفيات الخام} = \frac{\text{عدد الوفيات في البلد أثناء السنة}}{\text{عدد السكان في منتصف السنة}} \times 1000$$

$$\text{معدل وفيات الأطفال الرضع} = \frac{\text{عدد وفيات الأطفال الرضع في البلد خلال السنة}}{\text{عدد المواليد أحياء في البلد خلال نفس السنة}} \times 1000$$

معدل الحالة الاقتصادية

$$\text{معدل النشاط الاقتصادي} = \frac{\text{عدد المكتسبين}}{\text{عدد السكان}} \times 1000$$

الملاحق والمراجع

جدول (1) الأعداد العشوائية

3784	5736	5619	7307	6525	8124	0830	6852	5505	4008	6141	8389	1785
3629	5407	1045	5953	8533	7468	9650	3312	8005	5480	9355	0901	2587
6576	6717	5418	2581	9206	5803	5797	1441	6440	9873	5489	3444	0611
0741	4461	2153	3575	9982	5743	4540	9197	4240	2135	4901	0284	2951
5687	8134	7280	8940	8733	8469	6842	3961	7413	3255	7976	9550	8612
1489	6983	0101	7194	5558	8485	8784	7587	8547	0118	2109	7054	2370
0573	4831	1739	8149	5705	5052	4741	6276	3720	6778	9361	2510	8594
2598	1664	8074	1481	9499	1267	1797	5805	3543	2216	4373	3675	0489
1756	3251	4385	5670	9782	3863	4717	3956	4603	1278	2609	4482	9231
0219	9453	9494	6480	4834	2584	1076	7053	4459	6499	9194	4259	1191
0470	7752	8932	6935	2064	0371	1244	4939	3379	9864	7626	7552	8287
0959	9647	5057	9443	6617	9036	9920	8246	1229	8031	2050	2110	0295
1196	4895	8276	1287	2302	4794	0946	1672	8351	9548	3159	6180	2511
5927	0197	7508	7621	8894	8576	4999	8724	9015	7292	3635	7422	9214
9482	7723	4984	6113	0651	4773	9819	9368	2136	2206	9481	2941	3670
0758	7810	4554	4677	7674	8289	2733	4365	5146	6544	3515	2382	3481
9419	5897	9827	7630	9438	1647	1636	5389	2617	1438	0220	1035	4212
3855	4624	2366	9717	0661	0765	2311	8351	9367	6545	1240	7161	5277
5701	1642	2622	4678	8239	6666	8472	7203	2593	6235	5450	5906	2086
8486	8137	1550	0506	0379	0812	3824	3343	8137	7061	6519	0561	0174
5112	1203	0672	7109	9748	0084	4332	2251	3147	7474	1929	4937	2215
9245	4644	5426	2493	2461	2881	6196	8919	5321	2364	5375	1455	9927
2981	4243	7969	2256	6870	4799	6367	9628	6837	1352	6381	0648	3756
3506	0035	3574	6718	0476	4403	1759	5867	0013	0852	6530	0380	6513
8858	7232	2101	6483	9232	9972	3770	2621	9968	3551	2185	0311	2341
8026	8602	0191	4566	1102	4108	0069	0896	4591	9500	8249	4021	2533
3610	4439	3487	0994	2494	6051	0880	1334	1501	5932	1968	0010	1993
6317	4643	3039	1048	0597	7219	6474	8180	2949	2378	8726	3831	8530
2359	6006	1825	3241	8209	4401	5983	8385	1894	2136	6368	4296	0242
7030	1741	2975	6042	7888	0804	9041	7237	5968	4900	5925	4865	3661
4575	5225	5439	0462	9218	9624	6687	8754	4910	2980	3031	2531	1541
2575	3315	0172	7918	9494	3341	0168	7492	5656	9018	1031	2296	4930
9556	0178	9397	5889	1468	6240	3388	6346	1597	1617	9637	1292	5989
0687	0452	3852	5486	2672	3920	7616	3818	2042	4061	9536	8948	7344
6435	2440	0630	4570	3536	7060	7166	6529	8779	9559	5709	8059	5374

جدول (2) المساحات للتوزيع الطبيعي المعياري التراكمي لقيم المتغير العشوائي السالبة $\Phi(z) = P(Z \leq z)$

z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3829
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

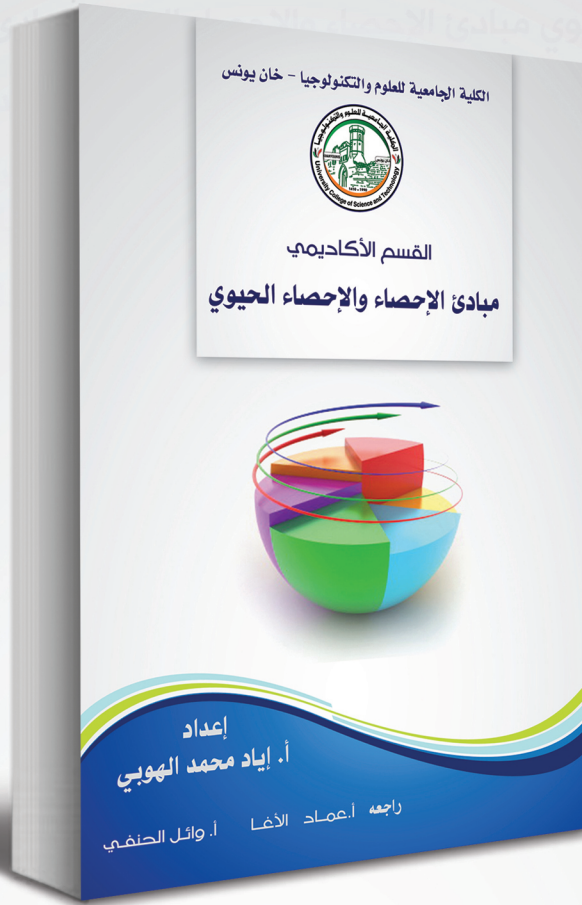
المراجع

المراجع العربية:

1. الإحصاء ، كامل فليفل، فتحي حمدان ، عمان الأردن 2005.
2. مبادئ الإحصاء لبرنامج الأعمال الإدارية و المالية في كليات المجتمع، عبد الرحمن عدس الأردن 1989.
3. مبادئ الإحصاء، جلال الصياد، عادل سمرة، المملكة العربية السعودية الطبعة الثانية 1983.
4. مقدمة في الإحصاء، محمد صبحي أبو صالح، عدنان محمد عوض، الأردن 1990.
5. مبادئ الإحصاء، محمود خالد عكاشة، فلسطين-غزة 1999.
6. كتاب مبادئ الإحصاء المقرر لجامعة القدس المفتوحة فلسطين.

المراجع الأجنبية:

- 1- Ronald E. Walpole , Raymond H. Myers, Sharon L. Myers, *Probability and Statistics for Engineers and Scientists*, Eighth Edition.
- 2- Robert V. Hogg, Allen T. Craig, *Introduction to Mathematical Statistics*, Fifth Edition.
- 3- Murray R. Spiegel, John Schiller, R. Alu Srinivasan, *Probability and Statistics*, Third Edition.



الكلية الجامعية للعلوم و التكنولوجيا

الطبعة الأولى

2017