جامعة مصراتة – كلية الهندسة

القســـــــــــــــم العام

المنهج الدراسي لمقرر

الإحصاء والاحتمالات

الفصل الدراسي: خريف (2024م–2025م)

الفصل الأول:

مقدمة في علم الإحصاء

تعريف علم الإحصاء :

هو مجموعة النظريات والطرق العلمية التي تبحث في جمع البيانات وعرضها وتحليلها واستخدام النتائج في التنبؤ او التقرير واتخاذ القرار.

ولكي نقوم بعملية البحث الاحصائي فيجب اتباع الخطوات التالية:

1- تحديد هدف الباحث.
2- تحديد مجتمع الدراسة.
3- تحديد مصادر البيانات والمعلومات وجمعها.
4- تصنيف البيانات وعرضها.
5- تحليل البيانات.
6- استخلاص النتائج وتفسيرها واتخاذ القرارات المناسبة لحل المشكلة موضوع البحث.

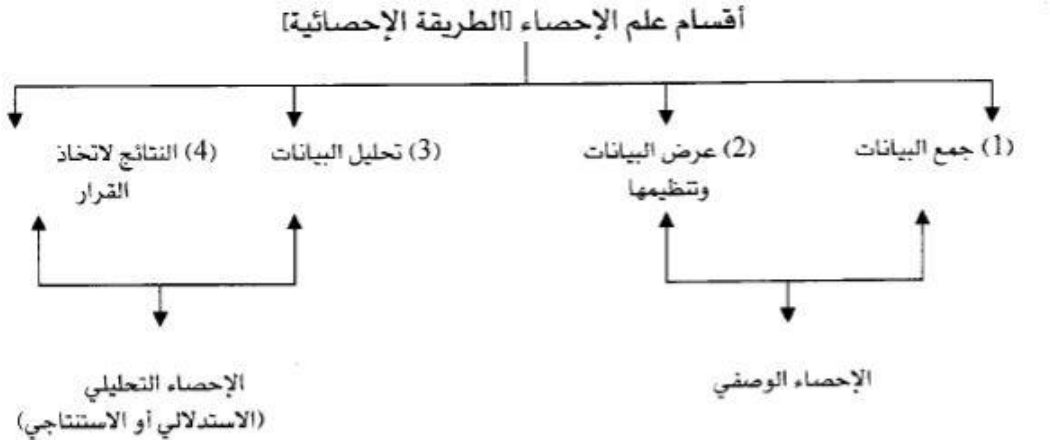
ينقسم علم الإحصاء الى نوعين:

1- الإحصاء الوصفي Descriptive Statistics:

وهو يختص بطريقة جمع ووصف وتلخيص البيانات وذلك باستخدام الجداول التكرارية والرسومات البيانية وبعض المقاييس الإحصائية.

2- الإحصاء التحليلي (الاستدلالي او الاستنتاجي) Inferential Statistical:

وهو يختص باستنتاج واتخاذ القرارات المناسبة للظاهرة قيد الدراسة مع حساب درجة الثقة المصاحبة لتلك القرارات والاستنتاجات.



أهداف الإحصاء :

1- تبسيط بيانات الظواهر الرياضية المعقدة، من خلال عرضها في جداول او رسومات بيانية او اشكال.
2- مقارنة المجموعات المختلفة وإيجاد العلاقة بينها.
3- وضع الحقائق في صورة عامة وواضحة (ارقام بدل جمل).
4- تهدف الطرق الإحصائية الى إيجاد أدوات مساعدة في عملية التنبؤ وكذلك في عملية التصنيف والتخطيط والتقويم الموضوعي والمناهج التربوية والتدريبية.

أهمية التوثيق وتوفير البيانات:

يعتمد علم الإحصاء في الوصول الاستنتاجات والقرارات المناسبة على تحليل البيانات التي تم تنظيمها وتلخيصها وعرضها بعد استخدام الطريقة المناسبة في جمع البيانات.

لذا تعتبر هذه البيانات هي الأساس الذي يعتمد عليه الباحث في عملية البحث والدراسة، ومن هنا تأتي أهمية توفر هذه البيانات وكذلك تعتبر صحة البيانات من الدعائم الأساسية لصحة الاستنتاجات والتقديرات والتنبؤات.

بعض المفاهيم الإحصائية العامة:

1- المجتمع الاحصائي: Statistical Population

وهو مجموعة من العناصر المشتركة في الصفة التي تهم الباحث، فمثلا عند دراسة ظاهرة الحوادث في مدينة ما فإن المجتمع الاحصائي هو جميع الافراد الذين تعرضوا او ارتكبوا حوادث في تلك المدينة.

وغالبا ما يكون المجتمع الاحصائي مجتمعا كبيرا وبالتالي دراسة جميع مفرداته قد يكون امر غير متيسر وعليه نلجأ لدراسة جزء من مفرداته يطلق عليه تسمية عينة Sample.

2- العينة: Sample

وهي مجموعة جزئية من مفردات المجتمع الاحصائي.

3- المعلمة (المؤثر): Parameter

وهي قيمة عددية تستخدم لوصف خاصية معينة في المجتمع الاحصائي، مثال على ذلك المتوسط الحسابي او الانحراف المعياري للمجتمع.

فمثلا متوسط دخل الفرد في بلد ما يعتبر معلمة وذلك لأنه يعكس المستوى المعيشي لأفراد تلك البلد.

4- الإحصاءه: Statistic

أي صيغة رياضية لا تعتمد على معلمات المجتمع وتستخدم لوصف خاصية معينة في العينة يطلق عليها تسمية احصاءه، مثال على ذلك الوسط الحسابي او الانحراف المعياري للعينة.

5- الظاهرة: Phenomena

هي صفة لعناصر المجتمع (او العينة) تختلف من عنصر لأخر في الشكل او النوع او الكمية، ويطلق عليها متغير Variable، مثل عدد السيارات المارة بإشارة ضوئية خلال فترة زمنية معينة.

مصادر جمع البيانات:

يمكن الحصول على المعلومات من مصدرين رئيسيين:

1- المصادر التاريخية:

وهي المصادر التي جرى سابقا توثيقها بعد جمع البيانات عنها وتلخيصها وتبويبها واستخلاص النتائج منها بحيث أصبحت مصدرا رئيسيا لكل مستفيد في المستقبل، ومن امثلة هذه المصادر:

- المصادر الموثقة في مصادر الدولة كدائرة الأحوال المدنية ودائرة الإحصاءات العامة.
- الإحصاءات التي وثقتها بحوث سابقة، ونجدها في الجامعات والمعاهد والمؤسسات التعليمية.
- البيانات الإحصائية المتوفرة في الشبكة المعلوماتية.

2- المصادر الميدانية:

وهي البيانات التي تستقى من الميدان مباشرة بإحدى طرق جمع البيانات كالتجربة او الاستمارة الإحصائية (الاستبيان) او التي جمعت بطرق أخرى لجمع البيانات كالمراسلة والاتصال الهاتفي والمقابلات الشخصية واستطلاعات الراي.

أنواع البيانات:

1- بيانات نوعية او وصفية: Qualitative Data

وهي التي يتم تصنيف مفرداتها وفقا لخاصية معينة في تلك البيانات، ومثالها لون الشعر (اسود، بني، ذهبي) او الحالة الاجتماعية (أعزب، متزوج، أرمل).

2- بيانات كمية (رقمية): Quantitative Data

وهي التي تكون مفرداتها مقاسه بمقياس كمي وقد تكون هذه البيانات متفصلة او متصلة، ومثالها (عدد الطلبة الرياضيين، الطول، الوزن، درجة الحرارة).

طرق جمع البيانات:

أولا: طريقة المسح الشامل: Census

وهي عملية جمع البيانات من جميع عناصر المجتمع الاحصائي، وتستخدم هذه الطريقة غالباً في الحالة التالية:

1- إذا كان المجتمع قيد الدراسة صغير.
2- إذا كان المطلوب الحصول على بيانات على مستوى عالي من الدقة.

مميزات المسح الشامل:

1- الدقة العالية. 2- الوضوح والتفصيل. 3- المصداقية.

عيوب المسح الشامل:

1- ارتفاع التكاليف. 2- الحاجة الى الوقت والجهد. 3- الحاجة الى عدد كبير من الباحثين.

ثانيا: طريقة المعاينة: Sampling Method

إذا تعذر استخدام طريقة المسح الشامل في الحصول على البيانات الإحصائية يتم اختيار جزء (عينة) من عناصر المجتمع قيد الدراسة بأسلوب علمي حيث نتأكد من تمثيل العينة للمجتمع تمثيلا سليما وهذا يتطلب منا تحديد هدف الدراسة ومجتمع الدراسة.

مميزات استخدام العينات في البحوث:

1- العينات تكتفي بعدد محدود من المفردات وليس جميعها، وذلك اقتصادي في الجهد والنفقات.
2- انها سريعة في إعطاء نتائج البحوث مقارنة بأسلوب الحصر الشامل.
3- تتيح للباحث التعميق في مصادر الاحكام واتخاذ القرارات.
4- تستخدم لأنها اقل عرضة للأخطاء مع الأساليب الأخرى.
5- تستخدم عندما يكون المجتمع الاحصائي لا نهائية.
6- انها طريقة مناسبة حيث إمكانية تحديد مدى الثقة في نتائجها، وكذلك نسبة تمثيلها للمجتمع.

عيوب استخدام العينة:

1- اخذ عينة من مصدر خاطئ، كأن تستخدم دليل الهاتف للحصول على عينة تمثل الرأي العام.
2- التحيز الشخصي، ويحدث ذلك حينما يأخذ الباحث عينته المختارة من فئة معينة لها خصائص مميزة عن المجتمع الكلي.
3- جمع بيانات ناقصة، فمثلا اهمال العامل الجغرافي عند دراسة المستوى الاقتصادي للسكان بتقسيم الاسر قيد البحث حسب دخلها.
4- خطأ الصدفة، يزداد احتمال ورود هذا الخطأ كلما صغر حجم العينة.

العينة وأساليب اختيارها:

المجتمع:

هو مجموعة من المشاهدات والمفردات والعناصر التي تمثل قيم ظاهرة مدروسة نمثلها احصائيا بالمتغير العشوائي كالطول والوزن والسعر والكمية المعروضة والجنس.

وينقسم المجتمع الى قسمين:

1- **مجتمع محدود:** وهو المجتمع الذي نستطيع حصر كل مفرداته مثل عدد الطلبة داخل الكلية.
2- **المجتمع غير المحدود:** وهو المجتمع الذي لا نستطيع حصر كل مفرداته مثل عدد جزيئات الماء في المحيط.

العينة:

وهي مجموعة من المفردات والمشاهدات التي يتم اختيارها من المجتمع الاحصائي بطريقة إحصائية معينة تتعلق بنوع العينة المسحوبة.

أنواع العينات:

أولا: العينات الاحتمالية: Probabilistic Sample

وهي العينات التي تسحب من المجتمع الاحصائي بحيث يكون لكل عنصر من عناصره فرصة او احتمال معروف لان يكون ضمن عناصر العينات المسحوبة.

1- العينة العشوائية البسيطة: Simple Random Sample

وهي العينة التي تسحب من مجتمع تكون لجميع مفرداته التي عددها n الفرصة نفسها بالاختيار، ولكي نحصل على عينة عشوائية فإننا بشكل عام نلجأ الى استخدام ما يسمى بجداول الأرقام العشوائية (Table of Random Numbers).

مثال1:

لاختيار عينة عشوائية تتألف من 5 أشخاص لتكوين لجنة من مجتمع يتكون من 1000 شخص، نستخدم جدول الأرقام العشوائية (موجود في نهاية الشيت) بالطريقة التالية:

نرقم عناصر المجتمع بالأرقام من 1 الى 1000 ثم نرجع الى احدى صفحات الجدول بشكل عشوائي (سنتعامل في هذا المثال مع الجدول المرفق)، ثم نختار سطر عشوائي وعمود عشوائي (وليكن السطر 5 والعمود 4) سنحصل على الرقم العشوائي 6446 ثم نستمر في نفس العمود لنحصل على الأرقام الأربعة المتبقية وهي الأرقام على التوالي: 1652، 9451، 3043، 1406، ولكن المجتمع في هذا المثال يتكون من 1000 مفردة (شخص) فقط فيمكن لنا اختيار اول ثلاث مراتب في اليمين او اليسار من الأرقام العشوائية التي اخترناها، وبذلك تكون اللجنة متألفة من الأشخاص الذين يحملون الأرقام التالية: 644، 165، 945، 304، 140.

وإذا كانت جداول الأرقام العشوائية غير متوفرة لذى الباحث يمكنه ان يلجأ لطريقة ثانية من طرق السحب العشوائي وهي طريقة صندوق القرعة (Chance Box)، حيث يضع الباحث أوراقا مطوية مرقمة من 1 الى N (حيث N تمثل حجم المجتمع، N=1000 في هذا المثال) داخل صندوق ما ويخلطها بشكل جيد ثم يسحب من الصندوق ورقة في كل مرة حتى يحدد كامل عناصر العينة.

جدول الأرقام العشوائية

51772	74640	42331	29044	46621	62898	93582	04186	19640	87056
24033	21491	83587	06568	21960	21387	76105	10863	97453	90581
45939	60173	52078	25424	11645	55870	56974	37428	93507	94271
30586	02133	75797	45406	31041	86707	12973	17169	88116	42187
03585	79353	81938	82322	96799	85659	36081	50884	14070	74950
64937	03355	95863	20790	65304	55189	00745	65258	11822	15804
15630	64759	51135	98527	62586	41889	25439	88036	24034	67283
09448	56301	57683	30277	94623	85418	68829	06652	41982	49159
21681	91157	77331	60710	52290	16835	48653	71590	16159	14676
91097	17480	29414	06829	87843	28195	27279	47152	35683	47280
50532	25496	95652	42457	78547	76552	50020	24819	52984	76168
07136	40876	79971	54195	25708	51817	36732	72484	94923	75986
27989	64728	10744	08396	56242	90985	28868	99431	50995	20507
86181	78949	86601	46258	00477	25234	09903	36574	72139	70185
54308	21154	97810	86764	82869	11785	55261	59009	38714	38723
65541	34371	09591	07889	58892	92843	72828	91341	84821	63886
08263	65952	85762	64236	39238	18776	84303	99247	46149	03229
39817	67906	48236	16057	81812	15815	63700	85915	19219	45943
62257	04077	79443	95203	02479	30763	92486	54083	23631	05325
53298	90276	62545	21944	16580	03878	07516	95715	02526	33537

2- العينة العشوائية المنتظمة (الدورية): Systematic Random Sample

وهي العينة التي يتم اختيارها من مجتمع يكون موزعا على أساس معين، كأن يكون موزع تصاعديا او تنازليا، ومن أمثلتها اختيار عدد من الصكوك المدفوعة من دفتر الصكوك المتسلسلة او اختيار عدد من المنازل المرقمة في محافظة ما، وتتم عملية الاختيار بتحديد الزيادة المنتظمة K ثم تحديد مفردة البداية التي يكون عادة اقل من K، ومن ثم يتم إضافة هذه الزيادة المنتظمة بشكل متسلسل.

وتعطى الزيادة المنتظمة K بالقانون: $K = \frac{N}{n}$

حيث: N: حجم المجتمع. n: حجم العينة.

مثال2:

إذا كان لدينا مجتمع احصائي يتكون من 36 مفردة، المطلوب اختيار عينة حجمها 6؟

1- نوجد مقدار الزيادة المنتظمة K:

$$K = \frac{N}{n} = \frac{36}{6} = 6$$

2- نختار مفردة البداية وهو رقم عشوائي من 1 الى 6.

والجدول التالي يوضح العينات التي يمكن سحبها من مجتمع الدراسة:

العينة الأولى	العينة الثانية	العينة الثالثة	العينة الرابعة	العينة الخامسة	العينة السادسة
1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36

مثال3:

إذا كان حجم مجتمع ما هو 1000 مفردة ويراد اختيار عينة حجمها 100 مفردة، أوجد الزيادة المنتظمة K؟

الحل: $K = \frac{N}{n} = \frac{1000}{100} = 10$

مثال4:

يراد اختيار عينة حجمها 100 من مجتمع عدد افراده 1260، اوجد مقدار الزيادة المنتظمة؟

الحل: $K = \frac{N}{n} = \frac{1260}{100} = 12.6$

يتم تقريبها للأصغر فيكون الجواب هو 12.

مثال5:

يراد اختيار عينة حجمها 200 مفردة من مجتمع حجمه 4000 مفردة، كيف يتم ذلك بطريقة العينة المنتظمة؟

الحل:

- نوجد مقدار الزيادة المنتظمة K: $K = \frac{N}{n} = \frac{4000}{200} = 20$
- نختار مفردة البداية وتكون اقل من 20 ولتكن 8.
- نضيف K بشكل متسلسل:

$$(8); (8 + 20); (8 + 20 + 20); (8 + 20 + 20 + 20); \dots\dots\dots$$
$$(8); (28); (48); (68); \dots\dots; (3988)$$

مثال6:

دائرة المحاسبة في شركة تود اجراء دراسة لعملاء الشركة من حيث طرق السداد وحجم الطلبات وكيفية الشحن، إذا تم اختيار العينات بطريقة العينة المنتظمة وكان عدد العملاء المسجلين 12500، وتم اختيار حجم العينة ليكون 350 اوجد العينة المطلوبة؟

الحل:

$$K = \frac{N}{n} = \frac{12500}{350} = 35.7 \cong 35$$

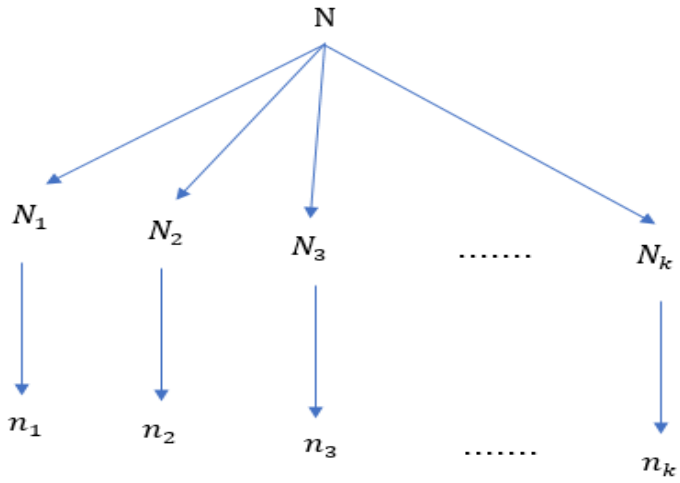
فتكون عينة العملاء المختارة هي كالاتي:

إذا اعتبرنا ان مفردة البداية هي (18):

$$18; 53; 88; 123; 141; 159; 177; 195; \dots; 12233$$

3- العينة العشوائية الطبقية: Stratified Random Sample

وهي الطريقة التي يكون فيها المجتمع مجزأ الى مجموعات (مجتمعات) جزئية (طبقية)، ونطبق على كل طبقة اختبار العينة العشوائية، والرسم التالي يبين كيفية اختبار العينة التي حجمها n من مجتمع طبقي مكون من k من الطبقات.



حيث N : حجم المجتمع الطبقي.

n : حجم العينة المطلوبة.

N_i : حجم الطبقة (i).

n_i : حجم العينة من الطبقة (i).

أي ان:

حجم المجتمع: $N = N_1 + N_2 + N_3 + \cdots + N_K$

حجم العينة: $n = n_1 + n_2 + n_3 + \cdots + n_K$

ونجد حجم العينة من الطبقة (i) الصيغة:

$$n_i = \frac{N_i}{N} n$$

ويتميز هذا النوع من العينات بما يلي:

1- ارتفاع مستوى المعاينة للمجتمع المستهدف بالدراسة.

2- يمكن الحصول على نتائج دقيقة من عينة حجمها صغير نسبيا.

3- أكثر كفاءة من غيرها في الأحوال التي تكون فيها مجموعات معينة من المجتمع ذات مواصفات خاصة هي السائدة.

4- تتطلب معرفة جيدة بالمجتمع المدروس ليتم تحديد الطبقات بشكل مناسب.

5- قد تتصف بشيء من التعقيد إذا كان عدد الطبقات كبيراً.

مثال7:

مجتمع حجمه 10000 مفردة ومكون من 4 طبقات حجم كل طبقة على التوالي 1000، 3500، 4000، 1500 يراد سحب عينة حجمها 400 مفردة من هذا المجتمع، كيف يتم ذلك بحيث تمثل هذه العينة تمثيلا سليماً؟

الحل:

$$N = 10000 \qquad n = 400$$

$$N_1 = 1000 \quad N_2 = 3500 \quad N_3 = 4000 \quad N_4 = 1500$$

حجم العينة من الطبقة رقم 1: $n_1 = \frac{N_1}{N} n = \frac{1000}{10000} * 400 = 40$

حجم العينة من الطبقة رقم 2: $n_2 = \frac{N_2}{N} n = \frac{3500}{10000} * 400 = 140$

حجم العينة من الطبقة رقم 3: $n_3 = \frac{N_3}{N} n = \frac{4000}{10000} * 400 = 160$

حجم العينة من الطبقة رقم 4: $n_4 = \frac{N_4}{N} n = \frac{1500}{10000} * 400 = 60$

ونلاحظ ان حجم العينة الكلي:

$$n = n_1 + n_2 + n_3 + n_4$$
$$= 40 + 140 + 160 + 60 \quad = 400$$

وهو حجم العينة المطلوبة.

4- العينة العشوائية العنقودية: Cluster Random Sample

وفيما يقسم المجتمع الى عناقيد (Clusters) متنافرة بحيث يحتوي كل عنقود أنواع العناصر الموجودة في المجتمع، ثم تسحب العينة عشوائية من هذه العناقيد، في حالة العينة الطبقية يتم على أساس التجانس، ولكن في حالة العينة العنقودية يتم على أساس تنافر العناصر لكي يكون كل عنقود ممثلا لكامل المجتمع.

5- العينة العشوائية متعددة المراحل: Multi-Stage Random Sample

تستخدم هذه الطريقة عندما يصعب الوصول مباشرة الى كافة عناصر المجتمع المستهدف بالدراسة.

مثال8:

يراد قياس المستوى التحصيلي لطلبة جامعة مصراتة في سنة ما، كيف تتم عملية الاختبار بطريقة سليمة؟

الحل: ننظر على مجتمع الطلاب بالجامعة على انه موزع على عدة كليات، ثم ان هذه الكليات مقسمة لعدة اقسام وهذه الأقسام بدورها تحتوي على عدة شعب دراسية، وبأخذ هذه المراحل بعين الاعتبار يتم اختيار العينة المطلوبة.

الجامعة ← كليات ← أقسام ← شعب دراسية

ثانيا: العينات غير الاحتمالية: Non-Probabilistic Sample

1- العينة العرضية (المختارة عن طريق الصدفة): Accidental Sample

وتعتمد في اختيارها على الصدفة المحضة.

2- عينة الحصة: Quota Sample

حيث تحدد حصة مقررة لكل مجموعة او طبقة من طبقات المجتمع المدروس، ثم نستخدم طريقة المصادفة في اختيار مفردات العينة.

حجم العينة: Sample Size

اهم الاعتبارات او العوامل التي تتحكم في اختيار حجم العينة:

1- التجانس (Homogeneity): كلما زاد تجانس (تماثل) عناصر المجتمع وقلت الفروقات بين عناصره كلما أمكن تصغير حجم العينة.
2- إجراءات او طريقة تحديد واختيار العينة (Sampling Procedures): حيث تؤثر نوعية العينة المختارة وطريقة اختيارها على حجم العينة.
3- الوقت والموارد المادية والبشرية (Time, Money and Personal) المتاحة للدراسة لها أثرها الكبير في تحديد حجم العينة، وكقاعدة عامة كلما ازداد حجم العينة المختارة بشكل صحيح كلما ازدادت دقة وموثوقية تعميم نتائجها على المجتمع المأخوذة منه.
4- درجة الخطأ المعياري المقبولة وحدود الثقة.

الفصل الثاني:

تصنيف البيانات وعرضها

تصنيف البيانات:

يقصد بتصنيف البيانات عرض هذه البيانات (البيانات الخام) في جداول مناسبة وذلك حتى يمكن تلخيصها وفهمها واستيعابها واستنباط النتائج منها ومقارنتها بغيرها من البيانات، كما يعتبر عرض وتصنيف البيانات الإحصائية الخطوة الثانية (بعد تجميع هذه البيانات الخام) في مفهوم التحليل الاحصائي.

أولا: العرض الجدولي للبيانات الإحصائية:

1- الجداول التكرارية البسيطة:

يستخدم هذا النوع من الجداول لوصف وتلخيص بيانات تتعلق بظاهرة واحدة فقط سواء كانت تلك الظاهرة كمية او وصفية وهي أسهل وابسط الجداول تركيباً ومفهوماً.

مثال1:

البيانات التالية هي درجات حصل عليها عشرون طالبا في مادة الإحصاء في امتحان نهاية العام:

12	11	15	14	12	10	15	13	12	10
14	10	13	12	15	13	12	10	12	15

والمطلوب تبويب هذه البيانات في جدول توزيع تكراري بسيط؟

الحل:

الدرجة	العلامات	التكرار
10	////	4
11	/	1
12	/ ////	6
13	///	3
14	//	2
15	////	4
المجموع $\sum$		20

مثال2:

البيانات التالية هي تقديرات 20 طالبا في مادة الإحصاء في العام الجامعي 2014/2013 والمطلوب هو وضع هذه البيانات في جدول بسيط؟

جيد	مقبول	جيد	جيد	مقبول	جيد	جيد جدا	مقبول	جيد	جيد جدا
ممتاز	جيد	ممتاز	جيد جدا	مقبول	جيد	ممتاز	جيد	جيد	مقبول

الحل:

التقدير	التكرار
مقبول	5
جيد	9
جيد حدا	3
ممتاز	3
المجموع	20

2- الجداول التكرارية ذات الفترات:

إذا كانت البيانات كثيرة وتكراراتها قليلة نجد ان الجدول التكراري البسيط لن يفي بالغرض المطلوب من حيث وصف وتلخيص البيانات.

خطوات بناء الجدول التكراري ذو الفترات:

1- حساب المدى = أكبر قيمة – أصغر قيمة.

2- حساب عدد الفترات (الفئات)، وهناك عدة طرق لتحديد عدد الفترات ومن اشهرها المعادلات الرياضية التي وضعها كل من "ستيرجس Sturges" و"يول Yole":

- معادلة ستيرجس: $k = \sqrt[4]{n} * 2.5$
- معادلة يول: $k = 1 + 3.322 \log_{10} n$

حيث ان: k: عدد الفترات (الفئات). n: عدد عناصر المجموعة.

3- حساب طول الفترة = $\frac{\text{المـــــــدى}}{\text{عدد الفـــــترات}}$

مع مراعاة ان: **(طول الفترة × عدد الفترات ≥ المـــدى)**.

4- اختيار بداية الفترة الأولى، أي الحد الأدنى لها مساوي لأقل قيمة موجودة بالبيانات او اقل بقليل منها، فمثلا تكون من الأرقام الصفرية لتسهيل الحسابات بعد ذلك.

5- بناء الجدول ووضع العلامات التي تمثل التكرار.

مركز الفترة:

$$\text{مركز الفترة} = \frac{\text{بداية الفترة + نهاية الفترة}}{2}$$

مثال3:

قام باحث بجمع بيانات تمثل درجات اختبار مادة الحاسب الآلي لخمسين طالبا من طلاب المرحلة الثانية من الثانوية العامة في الجدول التالي:

57	42	51	55	70
53	63	47	60	45
55	82	39	65	33
42	65	61	58	64
55	45	53	52	50
39	63	59	36	25
64	54	49	45	65
78	52	41	42	75
26	48	25	35	30
88	46	55	40	20

والمطلوب هو إعداد جدول توزيع تكراري ذو فترات للجدول السابق؟

الحل:

- المدى = أكبر قيمة – اصغر قيمة = $88 - 20 = 68$
- نوجد عدد الفترات: $k = \sqrt[4]{n} * 2.5 = \sqrt[4]{50} * 2.5 = 6.647$

$$k = 1 + 3.322 \log_{10} n = 1 + 3.322 \log_{10} 50 = 6.644$$

نقرب عدد الفترات لأقرب رقم صحيح فيكون عدد الفترات = 7

- طول الفترة = $\frac{\text{المـــــــدى}}{\text{عدد الفـــــترات}} = \frac{\mathbf{68}}{\mathbf{7}} = 9.714$

نقرب طول الفترة لأقرب رقم صحيح فيصبح طول الفترة = 10

حيث أن: طول الفترة × عدد الفترات ≤ المـــدى

$70=7*10 \leq 68$ ➔ OK

- نختار بداية الفترة الأولى وهو أصغر رقم = 20
- نبدأ في بناء الجدول كالتالي:

الفترات	العلامات	التكرار
20 –	///	4
30 –	// ///	6
40 –	/// /// //	12
50 –	/// /// ///	14
60 –	/// ///	9
70 –	///	3
88 – 80	//	2
المجموع ∑		50

3- الجداول التكرارية النسبية والمئوية:

هو تحويل تكرار كل فترة الى نسبة من مجموع التكرارات، أي ان:

$$\text{التكرار النسبي للفترة} = \frac{\text{تكرار الفترة}}{\text{مجموع التكرارات}}$$

التكراري المئوي للفترة = التكرار النسبي للفترة * 100

مثال4:

من نفس بيانات المثال السابق كون جدول التكرار النسبي؟

الحل:

بنفس الخطوات السابقة نكون جدول التوزيع التكراري ذو الفئات ومنه نكون جدول التوزيع التكراري النسبي والمئوي كالتالي:

حدود الفترات	التكرار النسبي للفترة = $\frac{\text{تكرار الفترة}}{\text{مجموع التكرارات}}$	التكراري المئوي للفترة=التكرار النسبي للفترة*100
20 −	$\frac{4}{50} = 0.08$	%8
30 −	$\frac{6}{50} = 0.12$	%12
40 −	$\frac{12}{50} = 0.24$	%24
50 −	$\frac{14}{50} = 0.28$	%28
60 −	$\frac{9}{50} = 0.18$	%18
70 −	$\frac{3}{50} = 0.06$	%6
88 − 80	$\frac{2}{50} = 0.04$	%4
المجموع $\sum$	1.000	% 100

4-الجداول التكرارية المتجمعة الصاعدة:

ويقصد بالتكرار المتجمع الصاعد هو تجميع تكرار كل فترة على جميع التكرارات السابقة لها بحيث يكون مجموع التكرار التصاعدي للفترة الأخيرة مساوي لمجموع التكرارات

مثال5: من نفس بيانات المثال السابق كون جدول التكرار للمتجمع الصاعد؟

الحل:

بنفس الخطوات السابقة نكون جدول التوزيع التكراري ذو الفترات ومنه نكون جدول التوزيع التكراري المتجمع الصاعد كالتالي:

حدود الفترات	التكرار المتجمع الصاعد
اقل من 20	صفر
اقل من 30	4
اقل من 40	10
اقل من 50	22
اقل من 60	36
اقل من 70	45
اقل من 80	48
اقل من 88	50

5- الجداول التكرارية المتجمعة الهابطة:

ويقصد بالتكرار المتجمع الهابط هو تجميع تكرار كل فئة على جميع التكرارات التالية لها بحيث يكون مجموع التكرار التنازلي للفئة الأولى مساوي لمجموع التكرارات.

مثال6: من نفس بيانات المثال السابق كون جدول التكرار المتجمع الهابط؟

الحل:

بنفس الخطوات السابقة نكون جدول التوزيع التكراري ذو الفئات ومنه نكون جدول التوزيع التكراري المتجمع الهابط كالتالي:

حدود الفتات	التكرار المتجمع الهابط
20 فأكثر	50
30 فأكثر	46
40 فأكثر	40
50 فأكثر	28
60 فأكثر	14
70 فأكثر	5
80 فأكثر	2
88 فأكثر	صفر

6- الجداول التكرارية المقفلة والمفتوحة:

الجداول المقفلة: وهي الجداول التكرارية التي تكون بها الفترة الأولى والأخيرة محدودة.

الجداول المفتوحة: وهي تنقسم الى قسمين:

أ- الجداول المفتوحة من الاعلى:

أي ان نهاية الفترة الأخيرة غير محدد، كما في المثال التالي:

الفترات	التكرار
4-6	
7-9	
أكثر من 9	

ب-الجداول المفتوحة من الاسفل:

أي ان بداية الفترة الاولى غير محدد،

كما في المثال التالي:

الفترات	التكرار
اقل من 7	
9-7	
12-9	

7-الجداول التكرارية المنتظمة وغير المنتظمة:

وهو يعتمد على طول الفئة.

أ- الجداول التكرارية المنتظمة:

أي ان جميع الفئات متساوية في الطول،

كما في المثال التالي:

الفترات	التكرار
6-4	7
9-7	5
12-10	10

ت-الجداول التكرارية الغير منتظمة

أي ان جميع الفئات غير متساوية في الطول،

كما في المثال التالي:

الفترات	التكرار
5-2	3
11-5	12
15-11	8

8-الجداول التكرارية المزدوجة:

يستخدم هذا النوع من الجداول في وصف وتلخيص البيانات المتعلقة بدراسة ظاهرتين في آن واحد.

أ- جدول تكراري مزدوج كمي:

الطول / الوزن	120 -	140 -	160 - 180	المجموع ∑
20 -	10	8	6	24
40 -	15	22	10	47
80 - 100	12	17	4	33
المجموع ∑	45	53	22	120

ب-جدول تكراري مزدوج وصفي:

التدخين / الاصابة	مدخن	غير مدخن	المجموع ∑
مصاب	300	40	340
سليم	10	250	260
المجموع ∑	310	290	600

ت-جدول تكراري مزدوج خليط (كمي ووصفي):

عدد الاطفال / المستوى التعليمي لرب الاسرة	0 -	2 -	4 -	6 -	8 – 10	المجموع ∑
أمي	5	6	10	4	9	34
اساسي	3	8	6	12	10	39
متوسط	1	12	14	9	11	47
عالي	6	10	11	14	3	44
المجموع ∑	15	36	41	39	33	164

9- الجداول التكرارية ذات الفترات المتصلة والمنفصلة:

أ- جدول التوزيع التكراري ذو الفترات المتصلة: وذلك إذا كانت بداية كل الفترة من ضمن بياناتها بينما نهايتها من ضمن بيانات الفترة التي تليها، ولهذا السبب في بعض الأحيان لا نكتب نهاية الفترة وذلك لتفادي اللبس الذي قد يحدث عند القراءة، ما عدا الفترة الأخيرة بالجدول تكتب فتكتب بدايتها ونهايتها (كما هو واضح في الجداول السابقة).

ب-جدول التوزيع التكراري ذو الفترات المنفصلة: وذلك إذا كانت بداية ونهاية كل فترة تنتمي للفترة نفسها.

مثال6:

بفرض ان البيانات التالية تمثل اجمالي ما أنفقه 75 شخص خلال أسبوع:

73	67	73	74	65	62	66	78	68	82	73	53	68	72
62	78	97	69	62	79	71	75	79	60	83	63	63	81
74	69	77	84	73	82	75	66	73	75	78	82	65	76
61	75	83	79	71	90	60	75	85	71	79	78	96	60
90	57	73	93	76	89	79	80	91	85	74	87	75	65
83	76	68	88	62									

المطلوب:

أ- وضع البيانات في جدول توزيع تكراري ذي فترات متصلة واخر ذي فترات منفصلة؟

ب-إيجاد التكراري النسبي والتكرار المئوي؟

الحل:

1- جدول التوزيع التكراري بفترات متصلة

الفترات	التكرار (f_i)	الحدود الحقيقية للفترات
50 -	1	50 – اقل من 56
56 -	5	56 – اقل من 62
62 -	12	62 – اقل من 68
68 -	15	68 – اقل من 74
74 -	22	74 – اقل من 80
80 -	11	80 – اقل من 86
86 -	6	86 – اقل من 92
92 - 98	3	92 – اقل من 98
المجموع ∑	75	

2- جدول التوزيع التكراري بفترات منفصلة

الفترات	التكرار (f_i)	الحدود الحقيقية للفترات
55 – 50	1	55.5 – 49.5
61 – 56	5	61.5 – 55.5
67 – 62	12	67.5 – 61.5
73 - 68	15	73.5 – 67.5
79 – 74	22	79.5 – 73.5
85 – 80	11	83.5 – 79.5
91 - 86	6	91.5 – 83.5
98 - 92	3	97.5 – 91.5
المجموع ∑	75	

- جدول التوزيع التكرار النسبي والتكرار المئوي:

الفترات	التكرار (f_i)	التكرار النسبي	التكرار المئوي
56 – 50	1	$\frac{1}{75} = 0.013$	0.013*100 = 1.3%
62 – 56	5	$\frac{5}{75} = 0.067$	6.7 %
68 – 62	12	$\frac{12}{75} = 0.160$	16 %
74 – 68	15	$\frac{15}{75} = 0.200$	20 %
80 – 74	22	$\frac{22}{75} = 0.293$	29.3 %
86 – 80	11	$\frac{11}{75} = 0.147$	14.7 %
92 – 86	6	$\frac{6}{75} = 0.080$	8 %
98 – 92	3	$\frac{3}{75} = 0.040$	4 %
المجموع $\sum$	75	1.000	100 %

ثانيا: العرض البياني للبيانات الإحصائية:

يعتبر العرض البياني للبيانات الإحصائية بمثابة تلخيص للبيانات الإحصائية في شكل يسهل استيعاب خصائص موضوع الدراسة، وتختلف طرق عرض البيانات باختلاف نوع البيانات والغرض من الرسم.

1. العرض البياني للبيانات الكمية:

أ- المدرج التكراري:

يستخدم المدرج التكراري في وصف وتلخيص البيانات الكمية بيانيا وذلك بعد وضعها في جدول تكراري ذي فترات حيث يتم تمثيل كل فترة بمستطيل يكون أحد اضلاعه طول الفترة والضلع الاخر تكرارها.

ولكن يجب مراعاة النقاط التالية عند رسم المدرج التكراري:

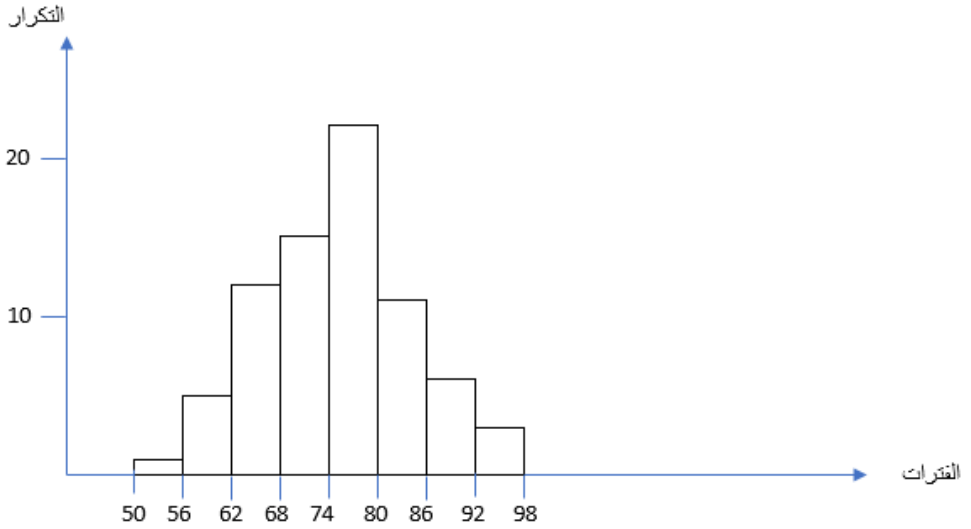
1- ان تكون الفترات متصلة.

2- ان تكون الفترات ذات اطوال متساوية.

مثال7:

مثل بيانات المثال السابق بيانياً باستخدام المدرج التكراري؟

الحل:



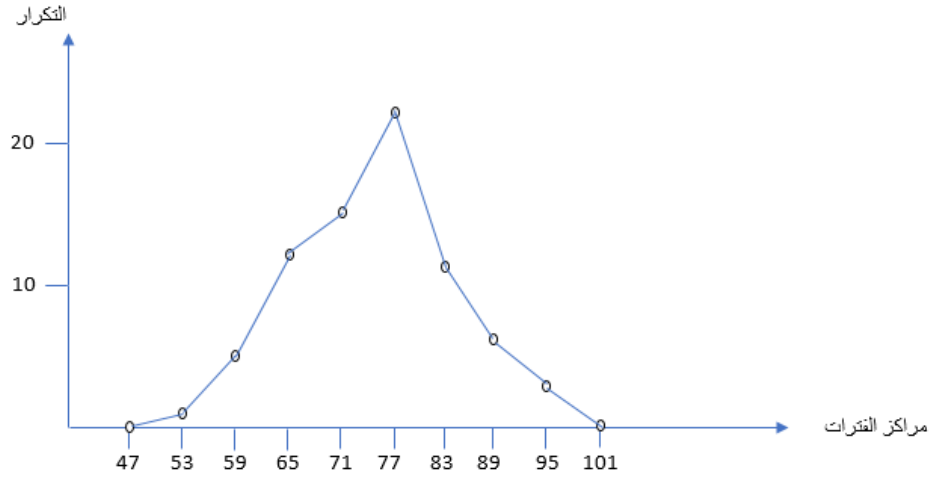
ب-المضلع التكراري:

حيث يتم تحديد مراكز الفترات مع التكرارات المناظرة لها.

مثال8: مثل بيانات المثال السابق باستخدام المضلع التكراري؟

الحل:

مراكز الفترات	التكرار (f_i)	الفترات
$\frac{44+50}{2}=47$	0	50 – 44
$\frac{50+56}{2}=53$	1	56 – 50
59	5	62 – 56
65	12	68 – 62
71	15	74 – 68
77	22	80 – 74
83	11	86 – 80
89	6	92 – 86
95	3	98 – 92
101	0	104 – 98

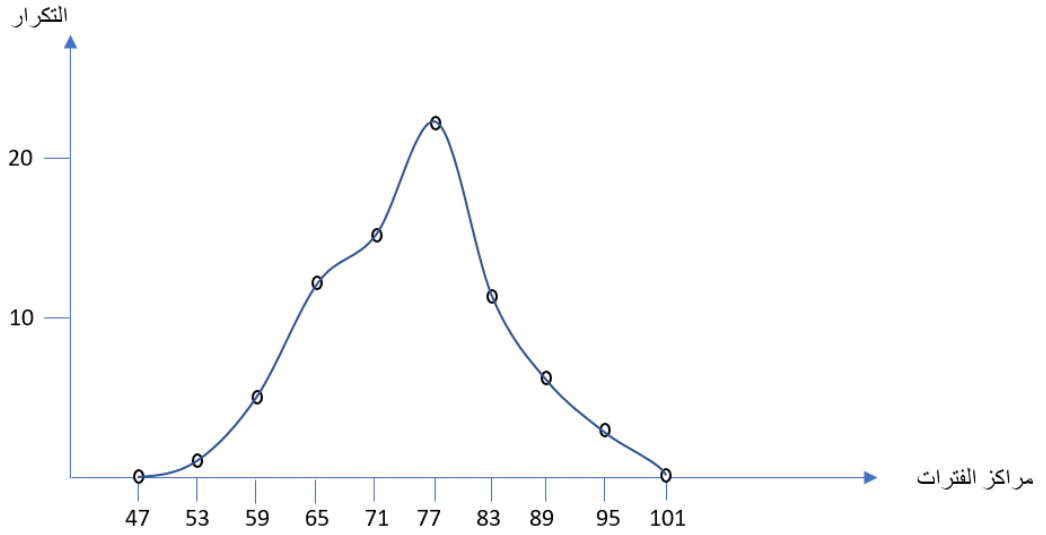


ت- المنحنى التكراري:

عندما تكون البيانات كثيرة وعدد الفترات كبير نجد ان اطوالها قصيرة وفي هذه الحالة نلجأ الى المنحنى التكراري.

مثال9:

مثل بيانات المثال السابق باستخدام المنحنى التكراري؟

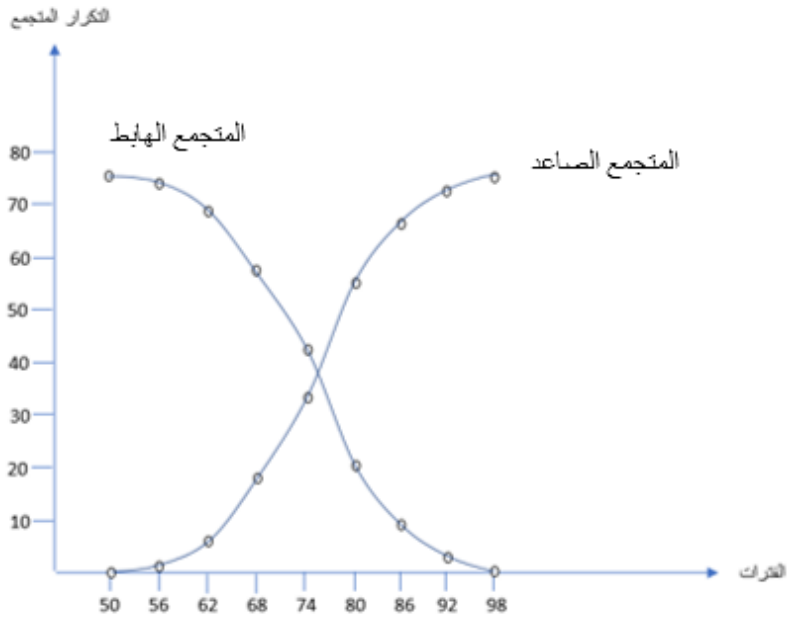


ث-المنحنى التكراري المتجمع الصاعد والهابط:

المثالي التالي يوضح فكرة هذه المنحنيات.

مثال: مثل بيانات المثال السابق باستخدام المنحنى التكراري المتجمع الصاعد وكذلك المنحنى التكراري المتجمع الهابط؟

		المنحنى التكراري المتجمع الصاعد		المنحنى التكراري المتجمع الهابط	
الفترات	التكرار (f_i)	الحدود العليا للفترات	التكرار المتجمع الصاعد	الحدود السفلى للفترات	التكرار المتجمع الهابط
56 – 50	1	اقل من 56	1	50 فأكثر	75
62 – 56	5	اقل من 62	6	56 فأكثر	74
68 – 62	12	اقل من68	18	62 فأكثر	69
74 – 68	15	اقل من 74	33	68 فأكثر	57
80 – 74	22	اقل من 80	55	74 فأكثر	42
86 – 80	11	اقل من 86	66	80 فأكثر	20
92 – 86	6	اقل من 92	72	86 فأكثر	9
98 – 92	3	اقل من 98	75	92 فأكثر	3



ملاحظة:

لتمثيل البيانات باستخدام المدرج التكراري او المنحنى التكراري المتجمع الصاعد او الهابط يجب ان يكون جدول التوزيع التكراري ذي فترات متصلة وان لم يكن الجدول التكراري ذي فترات متصلة نلجأ الى الحدود الحقيقية للفترات.

اما في حالة تمثيل البيانات باستخدام المنحنى التكراري والمضلع التكراري فليس من الضروري ان تكون البيانات متصلة لأننا نستخدم مراكز الفترات (مفترضين ان البيانات تتوزع بطريقة منتظمة).

مثال10:

بفرض ان البيانات التالية تمثل الفترة الزمنية التي عمرتها عينة عشوائية من النضائد السائلة المنتجة من قبل أحد المصانع قبل ان تستهلك بالكامل:

1.7	1.9	2.0	1.4	1.8	4.3	2.5	0.3	2.1	1.3	1.2	2.0	5.9
3.5	2.6	1.5	2.3	3.1	3.7	2.8	3.9	2.6	1.8	3.4	2.3	1.3
2.8	1.1	0.2	2.1	2.5	6.3	3.9	0.4	2.4	2.1	0.9	3.5	2.9
0.7	2.4	1.6	3.2	4.6	0.4	1.8	2.7	1.7	5.3	1.2		

المطلوب:

أ- وضع البيانات في جدول توزيع تكراري ذي فترات منفصلة ثم اوجد الحدود الحقيقية للفترات والتكرار النسبي والمئوي.

ب-مثل البيانات بيانياً باستخدام المدرج التكراري والمنحنى التكراري والمنحنى المتجمع الصاعد.

الحل:

أ- وضع البيانات في جدول توزيع تكراري:

1- عدد الفترات المطلوبة: $k = 2.5\sqrt[4]{n} = 2.5\sqrt[4]{50} = 6.648 \approx 7.0$

2- المدى = الحد الأعلى – الحد الأدنى = 6.3 – 0.2 = 6.1

3- طول الفترة $= \frac{\text{المدى}}{\text{عدد الفترات}} = \frac{6.1}{7} = 0.8711 \approx 0.9$ (تحقيق: $7 * 0.9 = 6.3 > 6.1$)

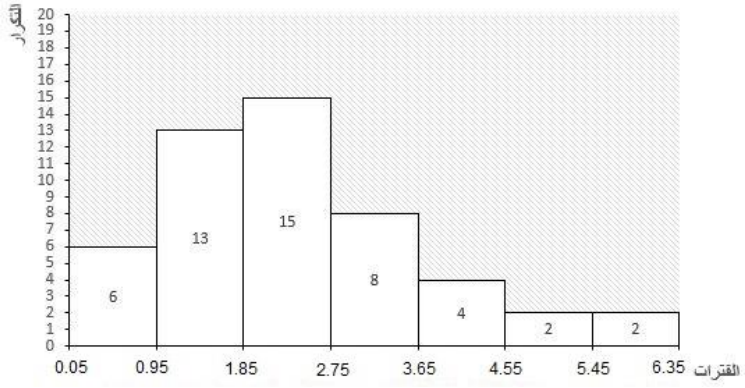
نكون جدول التوزيع التكراري ذي الفترات المنفصلة:

الفترات	التكرار (f_i)	الحدود الحقيقية	التكرار النسبي	التكرار المئوي %
0.9 – 0.1	6	0.95 – 0.05	$\frac{6}{50} = 0.12$	0.12*100=12%
1.8 – 1.0	13	1.85 – 0.95	$\frac{13}{50} = 0.26$	26%
2.7 – 1.9	15	2.75 – 1.85	0.30	30%
3.6 – 2.8	8	3.65 – 2.75	0.16	16%
4.5 – 3.7	4	4.55 – 3.65	0.08	8%
5.4 – 4.6	2	5.45 – 4.55	0.04	4%
6.3 – 5.5	2	6.35 – 5.45	0.04	4%

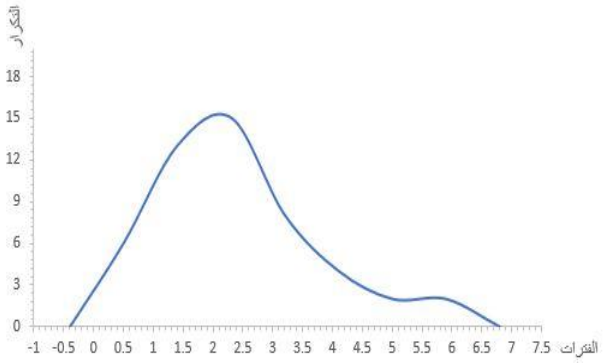
المجموع $\sum$	50		1.00	100%

ب-تمثيل البيانات بيانياً:

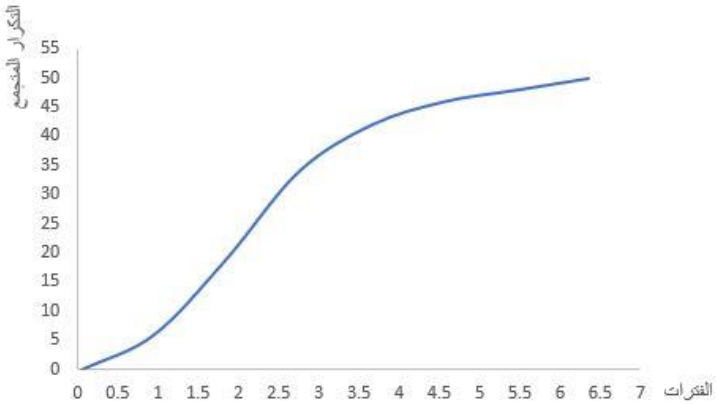
المضلع التكراري والمنحنى التكراري			المنحنى المتجمع الصاعد	
حدود الفترات	مراكز الفترات	التكرار (f_i)	حدود الفترات	التكرار المتجمع الصاعد
0.95– 0.05	0.5	6	اقل من 0.95	6
1.85 – 0.95	1.4	13	اقل من 1.85	19
2.75 – 1.85	2.3	15	اقل من 2.75	34
3.65 – 2.75	3.2	8	اقل من 3.65	42
4.55 – 3.65	4.1	4	اقل من 4.55	46
5.45 – 4.55	5.0	2	اقل من 5.45	48
6.35 – 5.45	5.9	2	اقل من 6.35	50



المدرج التكراري



المنحنى التكراري



المنحنى التكراري المتجمع الصاعد

العرض البياني في حالة الفترات غير متساوية الاطوال:

في بعض الأحيان قد يستحيل من الناحية العملية تكوين جدول توزيع تكراري ذي فترات متساوية.

وبالتالي يتم عمل فترات غير متساوية، ويجب تعديل التكرار بكل فترة وفقا لطول تلك الفترة.

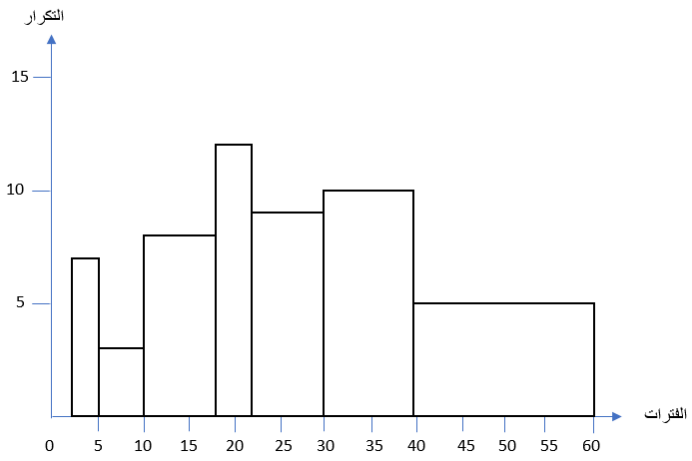
$$\text{التكرار المعدل لأي فترة} = \frac{\text{التكرار بتلك الفترة}}{\text{طول الفترة نفسها}}$$

مثال11: مثل البيانات الاتية بيانيا باستخدام المدرج التكراري؟

الفترات	2 –	5 –	10 –	18 –	22 –	30 –	40 – 60
التكرار f	21	15	64	48	72	100	100

الحل:

الفترات	التكرار f	طول الفترة	التكرار المعدل
2 – 5	21	3	$\frac{21}{3} = 7$
5 – 10	15	5	3
10 – 18	64	8	8
18 – 22	48	4	12
22 – 30	72	8	9
30 – 40	100	10	10
40 – 60	100	20	5



المدرج التكراري

2. التمثيل البياني للبيانات النوعية:

أ- الاعمدة البيانية:

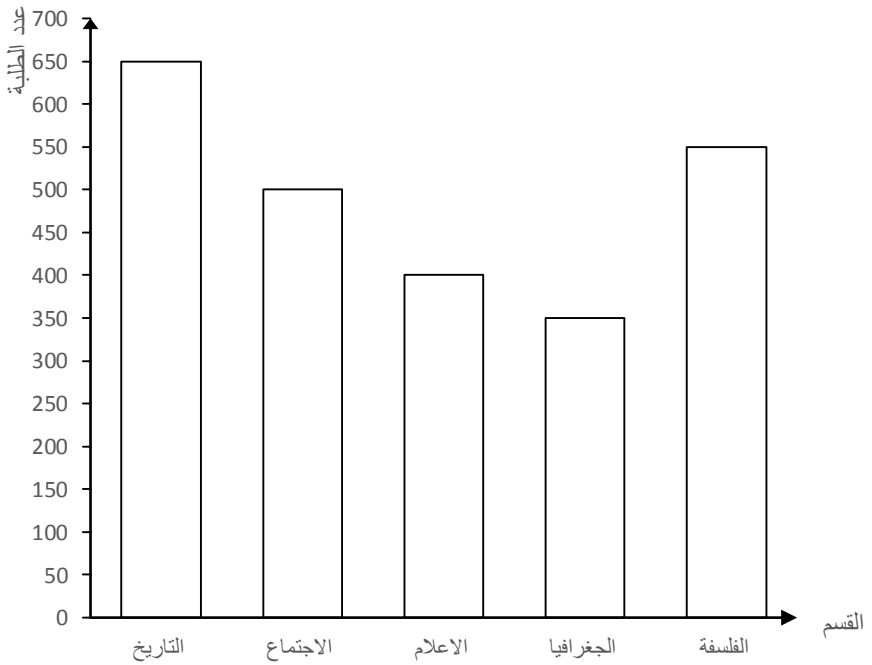
يتم تحديد المحور الافقي لصفات او أنواع او أزمنة الظاهرة قيد الدراسة، ويحدد المحور الرأسي التكرارات المناظرة لها.

مثال12:

الجدول التالي يوضح اعداد الطلاب ببعض اقسام كلية الآداب جامعة مصراتة، والمطلوب عرض هذه البيانات باستخدام طريقة الاعمدة البيانية؟

القسم	التاريخ	الاجتماع	الاعلام	الجغرافيا	الفلسفة
عدد الطلبة	650	500	400	350	550

الحل:



الاعمدة البيانية

ب-القطاعات الدائرية:

يتم رسم دائرة بنصف قطر معين ويتم تمثيل كل نوع من أنواع الظاهرة قيد الدراسة بجزء من تلك الدائرة، حيث ان:

زاوية القطاع (لكل نوع) = التكرار النسبي للقطاع (النوع) * 360

مثال13:

لنفس المثال السابق، المطلوب عرض البيانات بيانياً باستخدام طريقة القطاعات الدائرية؟

الحل:

نكون جدول التوزيع التكراري النسبي ومنه نوجد القطاعات الدائرية لكل نوع:

الفئة	التكرار	التكرار النسبي	زاوية القطاع الدائري
التاريخ	650	0.265	$0.265 * 360 = 95.4°$
الاجتماع	500	0.204	$0.204 * 360 = 73.5°$
الاعلام	400	0.163	$0.163 * 360 = 58.7°$
الجغرافيا	350	0.144	$0.144 * 360 = 51.8°$
الفلسفة	550	0.224	$0.224 * 360 = 80.6°$
المجموع $\sum$	2450	1.000	$360°$

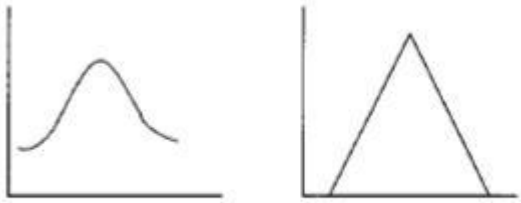
القطاعات الدائرية

أنواع المنحنيات التكرارية:

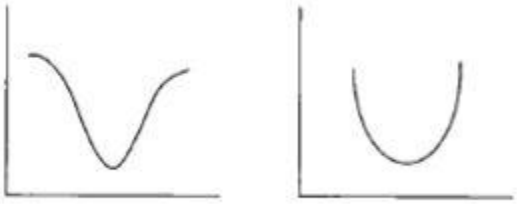
أولا: المنحنيات المتماثلة:

تتوزع قيمتها بشكل متماثل على خط المنتصف.

أ- المنحنى الطبيعي (الجرسي).



ب- المنحنى شكل حرف (U).

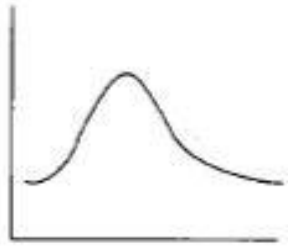


ثانياً: المنحنيات غير المتماثلة (الملتوية):

وهي المنحنيات التي تكون أحد أطرافها أطول من الطرف الاخر.

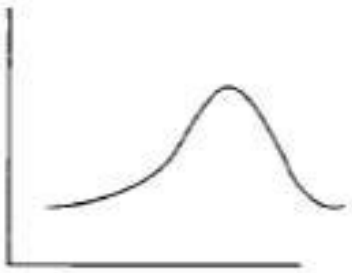
أ- ملتوية نحو اليمين (التواء موجب).

أي ان الطرف الطويل يقع جهة اليمين.

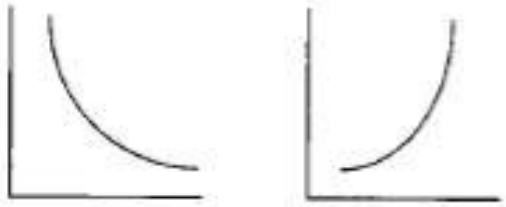


ب-ملتوية نحو اليسار (التواء سالب).

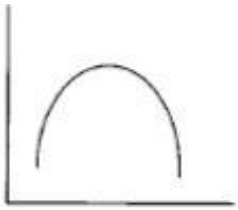
أي ان الطرف الطويل يقع جهة اليسار.



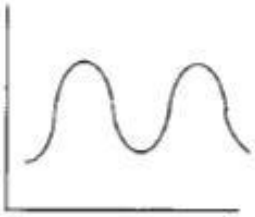
ج- التواء شديد لليمين او اليسار.



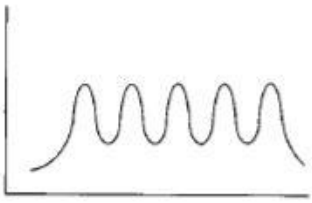
ثالثاً: منحنيات متعددة القمم:



أ- منحنى قمة واحدة (منوال واحد).



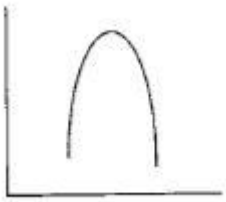
ب-منحنى قمتان (منوالان).



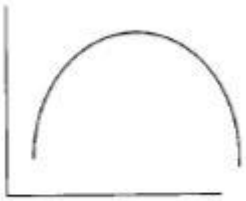
ج- منحنى متعدد القمم (متعدد المنوالات).

رابعا: منحنيات مفرطحة (مدببة القمم او معتدلة القمم):

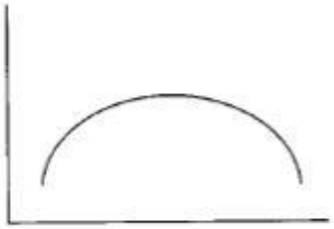
أ- منحنى مدبب



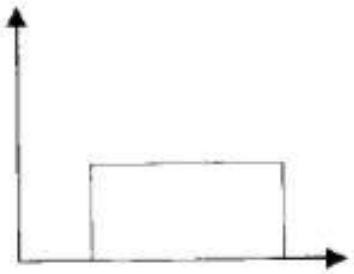
ب- منحنى معتدل



ج- منحنى مفلطح



خامساً: المنحنى المتجانس:



الفصل الثالث

مقاييس النزعة المركزية ومقاييس التشتت

مقاييس النزعة المركزية: Measures of Central Tendency

يقصد بمقاييس النزعة المركزية الطرق الإحصائية التي تقوم بحساب القيمة التي تتمركز حولها معظم المشاهدات، ومن اهم هذا المقاييس:

أولا: الوسط الحسابي (المتوسط): $(\overline{x})$

أ- طريقة حساب المتوسط الحسابي من البيانات الأولية:

1- الطريقة المباشرة: Direct Method

$$\bar{x} = \frac{\sum x}{n}$$

مثال1:

احسب الوسط الحسابي لدرجات 8 طلاب في مادة الإحصاء والتي كانت بياناتهم كالتالي:

9 8 8 7 6 5 3 2

الحل:

$$\bar{X} = \frac{2+3+5+6+7+8+8+9}{8} = \frac{48}{8} = 6$$

2- طريقة الانحرافات البسيطة (طريقة المتوسط الفرضي): Normal Deviation Method

يتم فرض قيمة عددية A تمثل المتوسط الفرضي (يفضل ان تكون القيمة A مساوية لإحدى المفردات التي بالظاهرة وان تكون واقعة في المنتصف من حيث القيمة).

نفرض ان: $d_i = x_i - A$

حيث ان d_i تمثل انحراف المفردات عن وسطها الافتراضي.

وبالتالي يمكن حساب الوسط الحسابي كما يلي:

$$\bar{x} = A + \frac{\sum_{i=1}^{n} d_i}{n}$$

مثال2: اوجد المتوسط الحسابي باستخدام الوسط الفرضي للبيانات الاتية:

10 9 12 11 8 13 14 10 15

الحل:

نختار وسط فرضي: A= 11، وعليه فإن:

$$\sum_{i=1}^{9} d_i = \sum_{i=1}^{9}(x_i - A)$$

$$= (10-11) + (9-11) + (12-11) + (11-11) + (8-11) + (13-11) + (14-11) + (10-11) + (15-11)$$

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

$$\bar{x} = A + \frac{\sum_{i=1}^{n} d_i}{n} = 11 + \frac{3}{9} = 11.333$$

3- طريقة الانحرافات المختصرة: Short-Cut Deviation

ويتم حساب المتوسط الحسابي في هذه الحالة كما يلي:

$$d_i = x_i - A \qquad u_i = \frac{d_i}{c} \qquad \bar{u} = \frac{\sum_{i=1}^{n} u_i}{n} \qquad \bar{x} = A + c * \bar{u}$$

حيث ان:

A: الوسط الافتراضي (كما تم تعريفه سابقا).

c: ثابت القسمة، وهو اكبر قيمة عددية تقبل كل القيم (المفردات) القسمة عليه بدون باقي (كسر) كلما أمكن ذلك من أجل تبسيط العمليات الحسابية.

مثال3:

إذا كانت البيانات التالية تمثل درجات 5 طلاب في امتحان، اوجد متوسط درجات هذا الامتحان باستخدام طريقة الانحرافات المختصرة؟

30 15 10 20 25

الحل:

باختيار وسط افتراضي: A=15 وثابت قسمة: c=5

Values	25	20	10	15	30	
$d_i = x_i - 15$	10	5	−5	0	15	
$u_i = \frac{d_i}{5}$	2	1	−1	0	3	$\sum = 5$

$$\bar{u} = \frac{\sum_{i=1}^{n} u_i}{n} = \frac{5}{5} = 1$$

$$\bar{x} = A + c * \bar{u} = 15 + 5 * 1 = 20$$

ب-طريقة حساب المتوسط الحسابي من الجدول التكرارية:

1- الطريقة المباشرة: Direct Method

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

حيث ان:

x_i : قيمة المفردة i (او مركز الفترة i).

f_i : تكرار المفردة i.

k: عدد الفترات.

مثال4:

يوجد بأحد المستشفيات موظفا يتقاضى رتبا شهريا قدره 2000 دينار وخمسة موظفين يتقاضى كل منهم 1800 دينار شهريا و10 موظفين يتقاضى كل منهم 1400 دينار شهريا، والمطلوب إيجاد متوسط الرواتب الشهرية في هذه المستشفى؟

الحل:

بفرض ان x ترمز الرواتب المختلفة وf ترمز لعدد الموظفين، فإنه بالإمكان تلخيص العمليات الحسابية في الجدول التالي:

الراتب (x_i)	عدد الموظفين (التكرار f_i)	$F_i * x_i$
2000	1	2000
1800	5	9000
1400	10	14000
المجموع $\sum$	16	25000

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i \, x_i}{\sum_{i=1}^{k} f_i} = \frac{25000}{16} = 1562.5$$

ملاحظة:

يمكن حساب المتوسط الحسابي في حالة اختلاف الأهمية النسبية للبيانات حيث يكون لكل مفردة وزنا يختلف عن بقية المفردات، يطلق على هذا المتوسط تسمية المتوسط الحسابي المرجح (الموزون) (Weighted Mean)، حيث ان:

$$\overline{W} = \frac{\sum_{i=1}^{k} w_i x_i}{\sum_{i=1}^{k} w_i}$$

حيث ان:

$\overline{W}$: المتوسط الحسابي المرجح.

x_i : قيمة المفردة i والتي وزنها w_i.

مثال5:

الجدول التالي يمثل المقررات الدراسية وعدد الوحدات لكل منها والدرجة المتحصل عليها طالب ما في احدى الفصول الدراسية، والمطلوب حساب المتوسط الحسابي لدرجات هذا الطالب خلال هذا الفصل؟

رمز المقرر	عدد الوحدات	الدرجة
ST 403	3	50
ST 402	4	62
ST 410	4	80
ST 405	4	55
ST 492	2	70

الحل

رمز المقرر	عدد الوحدات w_i	الدرجة x_i	$w_i x_i$
ST 403	3	50	150
ST 402	4	62	248
ST 410	4	80	320
ST 405	4	55	220
ST 492	2	70	140
المجموع $\sum$	17		1078

$$\overline{W} = \frac{\sum_{i=1}^{k} w_i x_i}{\sum_{i=1}^{k} w_i} = \frac{1078}{17} = 63.412$$

مثال6:

اوجد المتوسط الحسابي للبيانات التالية:

الفترات	التكرار (f_i)	مركز الفترات (x_i)	$f_i * x_i$
56 – 50	1	53	53
62 – 56	5	59	295
68 – 62	12	65	780
74 – 68	15	71	1065
80 – 74	22	77	1694
86 – 80	11	83	913
92 – 86	6	89	534
98 – 92	3	95	285
المجموع ∑	75		5619

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{5619}{75} = 74.92$$

2- طريقة الانحرافات البسيطة:

$$x_i = A + \frac{\sum_{i=1}^{n} f_i * d_i}{n} \qquad d_i = x_i - A$$

حيث ان:

A: الوسط الفرضي (كما تم تعريفه سابقا).

يفضل ان تكون قيمة الوسط الفرضي قيمة المفردة ذات أكبر تكرار في حالة البيانات المكررة، وتكون قيمته مركز الفترة التي تقع في وسط الفترات وان يكون لتلك الفترة أكبر تكرار ان أمكن ذلك.

مثال7:

اوجد المتوسط الحسابي باستخدام الوسط الفرضي للبيانات في المثال السابق؟

الحل:

نفرض وسط افتراضي مساويا لمركز الفترة التي يقابلها اكبر تكرار، أي ان (A =77)

الفترات	التكرار (f_i)	مركز الفترات (x_i)	$d_i = x_i - A$	$f_i * d_i$
56 – 50	1	53	53−77= −24	−24
62 – 56	5	59	59−77= −18	−90
68 – 62	12	65	65−77= −12	−144
74 – 68	15	71	71−77 = −6	−90
80 – 74	22	77	77−77 = 0	0
86 – 80	11	83	83−77 = 6	66
92 – 86	6	89	89 – 77 = 12	72
98 – 92	3	95	95−77 = 18	54
المجموع $\sum$	75			−156

$$x_i = A + \frac{\sum_{i=1}^{n} f_i * d_i}{n} = 77 + \frac{-156}{75} = 74.92$$

نلاحظ ان قيمة المتوسط $x_i = 74.92$ هي نفس القيمة في الطريقة السابقة.

3- طريقة الانحرافات المختصرة: Short-Cut Deviation

ويتم حساب المتوسط الحسابي في هذه الحالة كما يلي:

$$d_i = x_i - A \qquad u_i = \frac{f_i * d_i}{c} \qquad \bar{u} = \frac{\sum_{i=1}^{n} u_i}{n} \qquad \bar{x} = A + c * \bar{u}$$

حيث ان:

A: الوسط الافتراضي (كما تم تعريفه سابقا).

c: ثابت القسمة، ويفضل ان تكون قيمته مساوية لأكبر قيمة تقبل المفردات القسمة عليه بدون باقي (كسر) ان امكن ذلك في البيانات المتكررة، كما بفضل ان تكون قيمته مساوية لطول الفترة في حالة الجداول التكرارية ذات الاطوال المتساوية.

مثال8:

من بيانات المثال السابق، اوجد المتوسط الحسابي باستخدام طريقة الانحرافات المختصرة؟

الحل:

حيث ان طول الفترة يساوي 6 وعليه فإن: (c=6)، وباختيار (A=77) وذلك لأنها تقابل أكبر تكرار، نجد ان:

الفترات	التكرار (f_i)	مركز الفترات (x_i)	$d_i = x_i - A$	$u_i = \frac{f_i * d_i}{c}$
56 – 50	1	53	53−77= −24	−4
62 – 56	5	59	59−77= −18	−15
68 – 62	12	65	65−77= −12	−24
74 – 68	15	71	71−77 = −6	−15
80 – 74	22	77	77−77 = 0	0
86 – 80	11	83	83−77 = 6	11
92 – 86	6	89	89 – 77 = 12	12
98 – 92	3	95	95−77 = 18	9
المجموع $\sum$	75			−26

$$\bar{u} = \frac{\sum_{i=1}^{n} u_i}{n} = \frac{-26}{75} = -0.34667$$

$$\bar{x} = A + c * \bar{u} = 77 + 6 * (-0.34667) = 74.92$$

خصائص الوسط الحسابي:

1- مجموع انحرافات القيم عن الوسط الحسابي يساوي **صفر**. $\sum(x - \bar{X}) = 0$

2- الوسط الحسابي يتأثر بالقيم المتطرفة (الشاذة).

3- مجموع مربعات انحرافات القيم عن الوسط الحسابي أقل من مجموع مربعات انحرافات القيم عن أي قيمة أخرى.

4- الوسط الحسابي يتأثر بالعمليات الحسابية الأربعة.

5- لا يمكن حسابه بيانياً.

6- لا يمكن حسابه من الجداول التكرارية المفتوحة.

7- عند حساب المتوسط الحسابي باستخدام المتوسط الفرضي يجب استخدامه التكرار المعدل إذا كانت الفترات ذات أطوال غير متساوية.

8- اذا كانت الظاهرة x تأخذ القيم $x_1, x_2, x_3, \ldots, x_m$ و $\bar{x}$ يمثل المتوسط الحسابي لها، وكانت الظاهرة y تأخذ القيم $y_1, y_2, y_3, \ldots, y_n$ و $\bar{y}$ يمثل المتوسط الحسابي لها، فإن المتوسط الحسابي للظاهرتين معا يمكن أن نرمز له بالرمز $\bar{\bar{x}}$ ومعرف كما يلي:

$$\bar{\bar{x}} = \frac{m\bar{x}+n\bar{y}}{m+n}$$

9- يعد المتوسط الحسابي اهم مقاييس النزعة المركزية وأكثرها استخداماً.

مميزات الوسط الحسابي:

1- سهول حسابه.

2- مشاركة جميع قيم المفردات الظاهرة في حسابه.

عيوبه الوسط الحسابي:

1- تأثره بالقيم المتطرفة.

2- عدم إمكانية حسابه في حالة البيانات الوصفية.

ثانياً: الوسط الهندسي: Geometric Mean

هو الجذر النوني لحاصل ضرب الاعداد ببعضها، ورمزه (G).

أ- حساب المتوسط الهندسي في حالة البيانات الأولية:

$$G = \sqrt[n]{x_1 * x_2 * x_3 * \dots \dots * x_n}$$

مثال9:

احسب الوسط الهندسي للقيم التالية: 2 3 5 7 ؟

الحل:

حيث ان عدد القيم (n=4) فان الوسط الهندسي لها هو الجذر الرابع لحاصل ضرب هذه القيم:

$$G = \sqrt[4]{2 * 3 * 5 * 7} = \sqrt[4]{210} = 3.8$$

مثال10:

احسب الوسط الهندسي للقيم التالية: 4 6 3 8 9 ؟

الحل: $G = \sqrt[5]{9 * 8 * 3 * 6 * 4} = \sqrt[5]{5184} = 5.53$

مثال11:

بفرض ان انتاج مصنع القره بوللي للدائن ارتفع بنسبة 2% خلال الفترة من 1995م الى 1996م، كما ارتفع بنسبة 5% خلال الفترة من 1996م الى 1997م، وأيضا ارتفع بنسبة 10% خلال الفترة من 1997م الى 1998م، والمطلوب إيجاد معدل الزيادة في الإنتاج خلال الفترة من 1995م الى 1998م؟

الحل: $G = \sqrt[3]{\frac{102}{100} \times \frac{105}{100} \times \frac{110}{100}} = \sqrt[3]{1.1781} = 1.056$

وعليه فإن معدل الزيادة في الإنتاج خلال الفترة من 1995م الى 1998 يساوي 5.6%.

ب-حساب المتوسط الهندسي في حالة البيانات المكررة والبوبة:

اذا كانت $x_1, x_2, x_3, \dots, x_k$ تمثل المفردات المختلفة للظاهرة x (او مركز الفترات) و $f_1, f_2, f_3, \dots, f_k$ تمثل التكرارات المناظرة لها فإن المتوسط الهندسي لهذه البيانات يعرف كما يلي:

$$G = \sqrt[n]{x_1^{f_1} * x_2^{f_2} * x_3^{f_3} * \dots \dots * x_k^{f_k}}$$

مثل12:

اوجد المتوسط الهندسي لجدول التوزيع التكراري التالي:

الفترات	2 −	4 −	6 −	8 −	10 − 12
التكرار f_i	6	3	5	4	2

الحل:

الفترات	التكرار f_i	مركز الفترات x_i	$x_i^{f_i}$
2 −	6	3	3^6=729
4 −	3	5	5^3=125
6 −	5	7	7^5=16807
8 −	4	9	9^4=6561
10 − 12	2	11	11^2=121
المجموع $\sum$	20		

$$G = \sqrt[20]{729 * 125 * 16807 * 6561 * 121} = 5.679$$

بعض خواص الوسط الهندسي:

1- يدخل في حسابه جميع القيم.

2- لحساب الوسط الهندسي لمجموعة من القيم لا يصح ان تكون احدى القيم (او بعضها) مساويا للصفر.

3- لحساب الوسط الهندسي لمجموعة من القيم يجب ان تكون جميعها موجبة.

4- لا يتأثر كثيرا بالقيم المتطرفة (الشاذة).

5- لا يمكن حسابه من الجداول التكرارية المفتوحة، لأنه يعتمد في حسابه على مراكز الفترات.

6- قابلة للعملية الجبرية (الضرب والقسمة).

7- المتوسط الهندسي أصغر من المتوسط الحسابي لأي مجموعة من البيانات الموجبة وغير المتساوية.

ثالثاً: الوسط التوافقي: Harmonic Mean

مقلوب الوسط الحسابي لمقلوبات القيم، ويرمز له بالرمز (H).

ويفضل استخدامه في بعض الحالات مثل حساب متوسط السرعة اذ انها تعطى في العادة بدلالة وحدة الزمن.

أ- طريقة حساب المتوسط التوافقي من البيانات الأولية:

اذا كانت $x_1, x_2, x_3, \ldots, x_n$ تمثل قيم لظاهرة ما فإن المتوسط التوافقي لهذه البيانات يعرف كما يلي:

$$\frac{1}{H} = \frac{1}{n}\sum \frac{1}{x_i}$$

مثال13:

احسب الوسط التوافقي للقيم التالية: 5 4 2 ؟

الحل:

نطبق في القانون:

$$\frac{1}{H} = \frac{1}{n}\sum \frac{1}{x_i} = \frac{1}{3} * \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{5}\right) = \frac{1}{3} * \left(\frac{10+5+4}{20}\right) = \frac{1}{3} * \frac{19}{20} = \frac{19}{60}$$

$$\therefore H = \frac{60}{19} = 3.15$$

ب- طريقة حساب المتوسط التوافقي من الجداول التكرارية:

اذا كانت $x_1, x_2, x_3, \ldots, x_n$ تمثل المفردات المختلفة للظاهرة x (او مركز الفترات) و $f_1, f_2, f_3, \ldots, f_n$ تمثل التكرارات المناظرة لها فإن المتوسط التوافقي لهذه البيانات يعرف كما يلي:

$$\frac{1}{H} = \frac{1}{\sum f_i}\sum \frac{f_i}{x_i}$$

مثال14: اوجد المتوسط التوافقي من الجدول الاتي:

الفترات	-3	-7	-11	-15	-19	$27-23$
التكرار f_i	8	12	10	9	6	5

الحل:

الفترات	التكرار f_i	مركز الفترات x_i	$\frac{f_i}{x_i}$
$3-$	8	5	$\frac{8}{5}=1.6$
$7-$	12	9	$\frac{12}{9}=1.333$
$11-$	10	13	0.7682
$15-$	9	17	0.5294
$19-$	6	21	0.2857
$27-23$	5	25	0.2
المجموع $\sum$	50		4.7176

$$\frac{1}{H}=\frac{1}{\sum f_i}\sum\frac{f_i}{x_i}=\frac{1}{50}*4.7176\ =0.0943$$

$$H=\frac{1}{0.0943}=10.5986$$

بعض خواص المتوسط التوافقي:

1- يتأثر بالقيم المتطرفة (الشاذة).

2- لا يمكن حسابه من الجداول التكرارية المفتوحة.

3- قابل للعمليات الجيرية (الضرب والقسمة).

4- يكون $H \leq G \leq \bar{x}$ ويحدث التساوي في حالة ما تكون جميع القيم متساوية.

رابعاً: الوسيط: $(\tilde{x})$ *Median*

يعرف الوسيط على انه القيمة التي تتوسط مجموعة من القيم إذا رتبت ترتيبا تصاعديا او تنازليا.

أ- طريقة حساب الوسيط من البيانات الأولية:

لحساب الوسيط من البيانات الأولية يجب أولا ترتيب البيانات ترتيبا تصاعديا او تنازليا، وهناك حالتين:

1- اذا كان عدد المفردات فردي (n فردية):

يوجد رقم واحد يمثل الوسيط ويحسب ترتيبه من العلاقة: $\frac{n+1}{2}$

مثال15:

احسب الوسيط من البيانات التالية:

20 12 15 10 40 80 61

الحل:

نرتب البيانات تصاعديا أولا:

10 12 15 **20** 40 61 80

نحسب ترتيب الوسيط من العلاقة السابقة = $\frac{7+1}{2} = 4$

يعني ان ترتيب الوسيط هو الرقم الرابع، وبالتالي فالوسيط = 20.

2- إذا كان عدد المفردات زوجي (n زوجي):

يوجد رقمين يمثلان الوسط ويحسب عن طريق إيجاد الوسط الحسابي لهما ويحسب ترتيبه من العلاقة:

$$\frac{n}{2} \quad \text{و} \quad \frac{n}{2}+1$$

ويحسب الوسيط بأخذ المتوسط الحسابي لهما.

مثال16:

احسب الوسيط من البيانات التالية:

15 12 15 14 18 20 33 40

الحل:

نرتب تصاعدياً أولا:

12 14 15 15 20 33 40

نحسب ترتيب الوسيط من العلاقة السابقة= $(\frac{8}{2}\ ،\ \frac{8}{2}+1)$ = (4 ، 5)

ترتيب الوسيط هو الرابع والخامس وقيمة الوسيط هي المتوسط الحسابي للقيمتين اللتان ترتيبهما الرابع والخامس.

$$\text{الوسيط} = \frac{15+18}{2} = 16.5$$

ب-طريقة حساب الوسيط من الجداول التكرارية:

1- الطريقة الحسابية:

وذلك باتباع الخطوات التالية:

1- حساب التكرار المتجمع الصاعد.

2- حساب رتبة الوسيط: حيث رتبة الوسيط = $\frac{\sum_{i=1}^{k} f_i}{2}$ بغض النظر فيما اذا كان مجموع التكرارات فردي او زوجي.

3- تحديد الفترة التي يقع فيها الوسيط: وهي الفترة التي يقابلها تكرار متجمع صاعد يساوي رتبة الوسيط او أكثر منها.

4- تحديد طول الفترة الوسيطة وكذلك تكرارها.

5- نستخدم القانون:

$$\tilde{x} = L + \frac{\frac{n}{2} - F}{f} * c$$

حيث ان:

$\tilde{x}$: الوسيط.

L: الحد الأدنى للفترة الوسيطة.

n: مجموع التكرارات $n = \sum_{i=1}^{k} f_i$.

F : القيمة السابقة للتكرار المتجمع الصاعد للفترة الوسيطة.

f : تكرار فترة الوسيطة.

c : طول الفترة الوسيطة.

مع ملاحظة ان يكون جدول التوزيع التكراري ذي بيانات متصلة وان لم يكن كذلك فيجب حساب الحدود الحقيقية للفترات.

مثال17:

الجدول التالي يوضح العلاقة بين فترات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط؟

فئات الدخل	30-20	40-30	50-40	60-50	70-60
عدد العمال	20	40	100	30	10

الحل:

نكون الجدول التالي:

الفترات	f	الحدود العليا للفترات	المتجمع الصاعد
30-20	20	اقل من 30	20
40-30	40	اقل من 40	60
50-40	100	اقل من 50	160
60-50	30	اقل من 60	190
70-60	10	اقل من 70	200
المجموع $\sum$	200		

f

القيمة السابقة F

الفترة الوسيطة

- نحسب ترتيب الفترة الوسيطة = $\frac{n}{2} = \frac{200}{2} = 100$
- نبحث داخل عمود (المتجمع الصاعد) عن القيمتين التي ينحصر بيمهما ترتيب الوسيط فنجد ان قيمة ترتيب الوسيط = 100 محصورة بين (60 – 160).

فنجد ان:

$$c = 10 \qquad L = 40 \qquad F = 60 \qquad f = 100$$

بالتطبيق في القانون:

$$\tilde{x} = L + \frac{\frac{n}{2} - F}{f} * c = 40 + \frac{\frac{200}{2} - 60}{100} * 10 = 44$$

2- الطريقة البيانية:

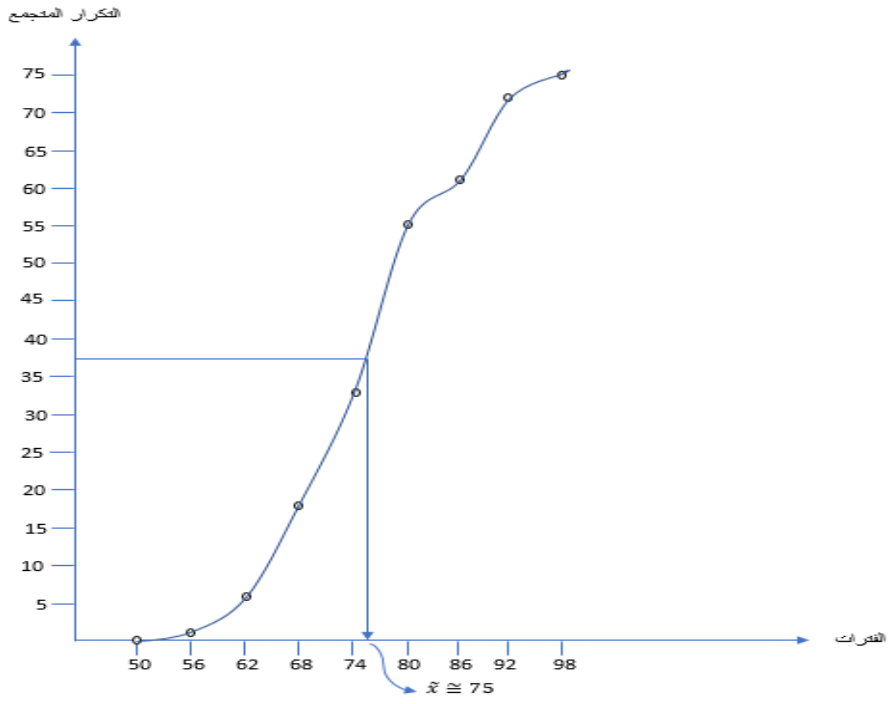
يمكن إيجاد الوسيط بيانيا باستخدام منحنى التكرار المتجمع الصاعد او الهابط او كلاهما معاً.

مثال18:

للبيانات التالية، اوجد الوسيط بيانيا؟

الفترات	التكرار f_i	الحدود العليا للفترات	التكرار المتجمع الصاعد
56 – 50	1	أقل من 56	1
62 – 56	5	أقل من 62	6
68 – 62	12	أقل من 68	18
74 – 68	15	أقل من 74	33
80 – 74	22	أقل من 80	55
86 – 80	11	أقل من 86	66
92 – 86	6	أقل من 92	72
98 – 92	3	أقل من 98	75
المجموع $\sum$	75		

رتبة الوسيط = $\frac{75}{2}$ = 37.5 فترة الوسيط: 80 – 74



المنحنى المتجمع الصاعد $(\tilde{x} \cong 75)$

بعض خواص الوسيط:

1- لا يتأثر بالقيم المتطرفة (الشاذة).

2- يمكن حسابه بيانياً.

3- يمكن حسابه من الجداول التكرارية المفتوحة.

4- يمكن حسابه من البيانات الوصفية القابلة للترتيب وعددها فردي.

5- لا يدخل في حسابه جميع القيم.

6- لا يحتاج لتعديل التكرارات إذا كانت أطوال الفترات غير متساوية.

خامساً: المنوال: (*m*) ***Mode***

المنوال هو القيمة او الصفة الأكثر تكرارا مقارنة بباقي القيم او الصفات.

يستخدم هذا المقياس إذا كان الهدف من الدراسة الحصول على مقياس سريع لظاهرة النزعة المركزية بصرف النظر عن الدقة في القياس.

أ- طريقة حساب المنوال من البيانات الأولية:

مثال19:

أوجد المنوال للبيانات التالية:

30 25 20 26 28 24 28 27

الحل:

حيث ان المفردة 28 متكررة مرتين وهي الوحيدة المتكررة وعليه فإن المنوال: *m= 28*

مثال20:

احسب المنوال في كل من الحالات التالية:

7 8 9 8 10 12	المنوال = 8
10 12 10 15 12 10	المنوال = 10
15 16 15 20 16 30	المنوال = 15، 16
20 30 40 140 50 60	المنوال = لا يوجد

ب-طريقة حساب المنوال من الجداول التكرارية:

1- الطريقة الحسابية:

يمكن حساب المنوال في حالة الجداول التكرارية على النحو التالي:

$$m = L + \frac{d_1}{d_1+d_2} * c$$

حيث ان:

m: المنوال.

L: الحد الأدنى للفترة المقابلة لأعلى تكرار.

d_1 : حاصل طرح اعلى تكرار من التكرار السابق له (اعلى تكرار – التكرار السابق له).

d_2 : حاصل طرح اعلى تكرار من التكرار اللاحق له (اعلى تكرار – التكرار اللاحق له).

c : طول الفترة.

مثال21:

اوجد المنوال من الجدول التالي:

فئات الدخل	20−10	30−20	40−30	50−40	60−50	70−60	80−70
عدد العمال	5	12	22	38	22	12	5

الحل:

أكبر تكرار هو 38.

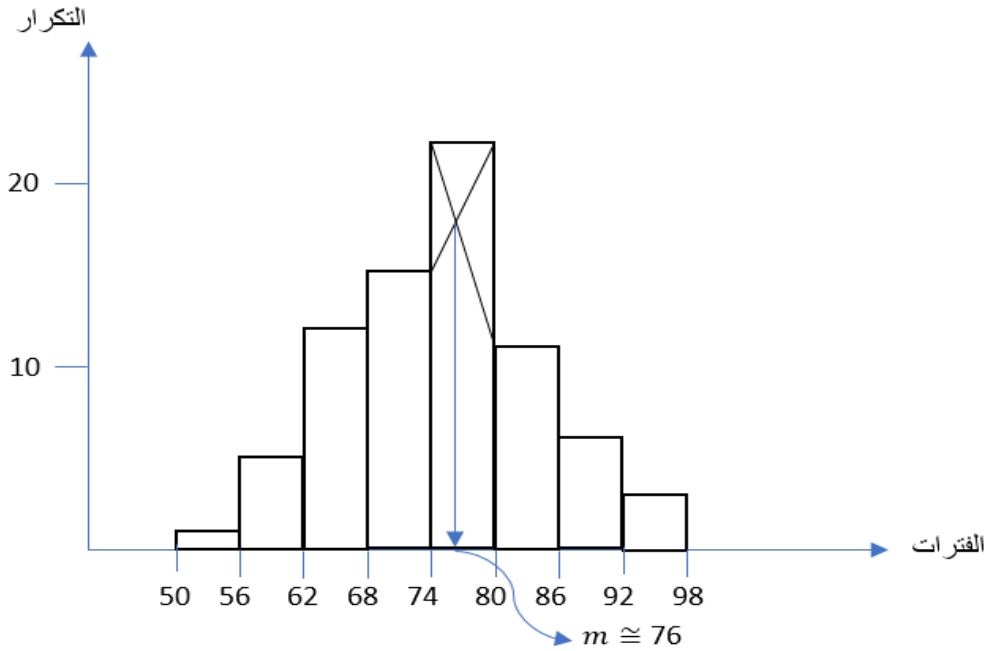
$$L = 40 \qquad d_1 = 38 - 22 = 16 \qquad d_2 = 38 - 22 = 16 \qquad c = 10$$

بالتطبيق في القانون:

$$D = L + \frac{d_1}{d_1+d_2} * c = 40 + \frac{16}{16+16} * 10 = 45$$

2- الطريقة البيانية:

يمكن إيجاد المنوال بيانيا باستخدام المدرج التكراري حيث يتم توصيل نهاية المستطيل الذي يمثل الفترة المنوالية من الناحية اليسرى بنهاية المستطيل للفترة السابقة لها من الناحية اليسرى، وكذلك نوصل نهاية المستطيل للفترة المنوالية من الناحية اليمنى بنهاية المستطيل للفترة اللاحقة لها من الناحية اليمنى، ومن نقطة التقاطع نسقط عمود على المحور الأفقي ونقطة التقائه مع هذا المحور تعطي قيمة المنوال.



حساب المنوال بيانيا ($m \cong 76$)

بعض خواص المنوال:

1- أسهل مقاييس النزعة المركزية.

2- يمكن حسابه بيانيا كما يمكن حسابه من الجداول التكرارية المفتوحة.

3- ليس له معنى إذا كانت التكرارات قليلة.

4- أفضل مقاييس النزعة المركزية لوصف الظواهر النوعية.

5- لا يتأثر بالقيم المتطرفة (الشاذة).

6- يجب استخدام التكرار المعدل إذا كانت أطول الفترات غير متساوية.

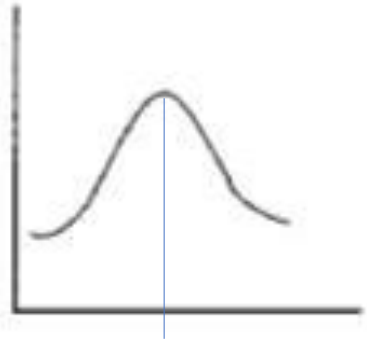
7- لا يدخل في حسابه جميع البيانات المعطاة.

8- غير دقيق ويمكن وجود أكثر من منوال لنفس المجموعة من البيانات.

العلاقة بين المتوسط الحسابي والوسيط والمنوال:

إذا كان لمجموع البيانات منوال واحد فإن المتوسط والوسيط والمنوال ترتبط مع بعضها بإحدى العلاقات التالية:

1- إذا كان التوزيع التكراري متماثل فإن المنحنى سيكون متماثل وله قمة واحدة وشكله يشبه شكل الجرس ويكون المتوسط الحسابي مساويا للوسيط مساويا للمنوال ويساوي مركز الفترة التي يقابلها أكبر تكرار.

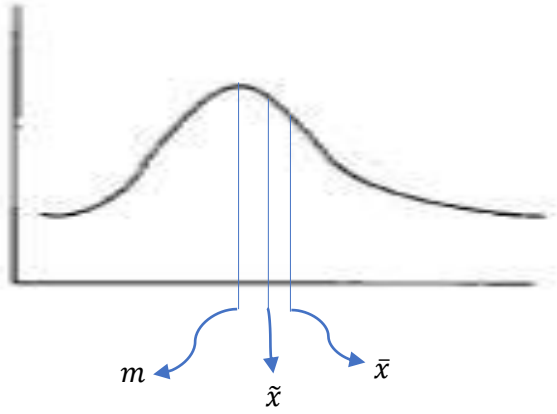


$\bar{x} = \tilde{x} = m$

2- إذا كان المتوسط الحسابي أكبر من الوسيط والمنوال فإن المنحنى ملتوي ناحية اليمين.

$\bar{x} > \tilde{x}$

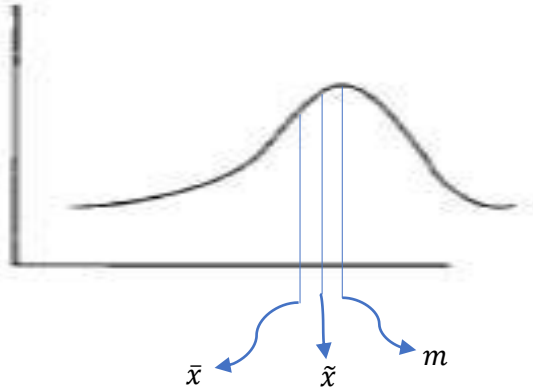
$\bar{x} > m$



3- إذا كان المتوسط الحسابي أقل من الوسيط والمنوال فإن المنحنى ملتوي ناحية اليسار.

$\bar{x} < \tilde{x}$

$\bar{x} < m$



الربيعيات والعشريات والمئويات:

أولا: الربيعيات: Quartiles

تقسيم النقاط الى أربعة اقسام متساوية تسمى الربيعيات وعددها ثلاثة هي من اليسار الى اليمين، الربيع الأول (الأدنى) Q_1، الربيع الثاني Q_2، الربيع الثالث (الأعلى) Q_3.

الربيع الأول (الأدنى) (Lower quartile) Q_1: وهي القيمة التي يسبقها ربع البيانات ويليها ثلاثة ارباع البيانات.

ويتم حسابه من الجداول التكرارية كما يلي: $$Q_1 = L + \frac{\frac{n}{4} - F}{f} * c$$

حيث أن: L) الحد الأدنى لفترة الربيع الأدنى، (F) التكرار المتجمع الصاعد للفترة التي تسبق فترة الربيع الأدنى، (f) تكرار فترة الربيع الأدنى، (c) طول فترة الربيع الأدنى، (n) مجموع التكرارات.

الربيع الثاني Q_2: وهي القيمة التي يسبقها نصف البيانات ويليها نصف البيانات وهي مساوية لقيمة الوسيط.

الربيع الثالث (الأعلى) (Upper quartile) Q_3: وهي القيمة التي يسبقها ثلاثة ارباع البيانات ويليها ربع البيانات.

ويتم حسابه من الجداول التكرارية كما يلي: $$Q_3 = L + \frac{\frac{3n}{4} - F}{f} * c$$

حيث أن: L) الحد الأدنى لفترة الربيع الأعلى، (F) التكرار المتجمع الصاعد للفترة التي تسبق فترة الربيع الأعلى، (f) تكرار فترة الربيع الأعلى، (c) طول فترة الربيع الأعلى، (n) مجموع التكرارات.

ثانيا: العشريات (D): Deciles

تقسيم النقاط الى عشرة اقسام متساوية في المساحة تسمى العشريات، فالعشير الأول هو القيمة التي يسبقها $(\frac{1}{10})$ من البيانات ويليها $(\frac{9}{10})$ من البيانات على فرض ان القيم مرتبة ترتيبا تصاعديا، وتعرف العشريات الأخرى بنفس الطريقة، ونلاحظ ان العشير الخامس هو الوسيط، ويكون هناك تسع عشريات ويرمز لها بالرموز $(D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9)$.

ويتم حسابه من الجداول التكرارية كما يلي:

$$D_i = L + \frac{\frac{i*n}{10} - F}{f} * c \qquad i = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

حيث أن: (L) الحد الأدنى لفترة العشير i، (F) التكرار المتجمع الصاعد للفترة التي تسبق فترة العشير i، (f) تكرار فترة العشير i، (c) طول فترة العشير i، (n) مجموع التكرارات.

ثالثاً: المئويات (P): Percentiles

تقسيم النقاط الى مائة قسم متساوٍ في المساحة تسمى المئويات، فالمئوي الأول هو القيمة التي يسبقها $(\frac{1}{100})$ من البيانات ويليها $\left(\frac{99}{100}\right)$ من البيانات المرتبة ترتيبا تصاعديا، ويرمز لها بالرمز P_1.

ويلاحظ ان المئوي الخمسين (P_{50}) هو الوسيط (M) وهو الربيع الثاني (Q_2) وهو العشير الخامس (D_5)، أي ان:

$$P_{50} = \mathrm{M} = Q_2 = D_5$$

حساب المئويات:

حساب المئويات للبيانات الأولية:

1- ترتب البيانات تصاعديا حسب قيمتها، وليكن عددها (n).

2- لإيجاد ترتيب المئوي K نطبق في المعادلة التالية: $P_K = \frac{K}{100} * n$

اذا كانت القيمة كسرا نقربه الى اعلى، ويكون P_K مساويا للعدد الذي ترتيبه القيمة المقربة.

مثال22:

اوجد المئوي الخامس والعشرون للبيانات التالية:

16	40	31	27	3	22	17	12	5	7

الحل:

1- نرتب البيانات تصاعديا (من اليسار الى اليمين):

3	5	7	12	16	17	22	27	31	40

2- نوجد ترتيب المئوي الخامس والعشرون P_{25} من القانون:

$$P_K = \frac{K}{100} * n \quad \rightarrow \quad P_{25} = \frac{25}{100} * 10 = 2.5 \cong 3.0$$

اذا ترتيب المئوي الخامس والعشرون هو الرقم الثالث، وبالتالي يكون: $P_{25} = 7$

حساب المئويات من الجداول التكرارية:

يتم حسابه باستخدام القانون التالي:

$$P_i = L + \frac{\frac{i*n}{100} - F}{f} * c \qquad i = 1, 2, 3, \ \dots \ , 99$$

حيث أن: (L) الحد الأدنى لفترة المئوي i، (F) التكرار المتجمع الصاعد للفترة التي تسبق فترة المئوي i، (f) تكرار فترة المئوي i، (c) طول فترة المئوي i، (n) مجموع التكرارات.

ملاحظة:

يجب أن يكون جدول التوزيع التكراري ذي فترات متصلة وان لم يكن كذلك فيجب حساب الحدود الحقيقية للفترات عند حساب أي مقياس من المقاييس السابقة.

مثال23:

أوجد كلا من Q_1 و Q_3 و P_{10} و P_{90} من البيانات التالية:

الفترات	10 −	20 −	30 −	40 −	50 −	60 −	70 −	90 − 80
التكرار f_i	22	18	15	12	11	10	16	8

الحل:

نوجد أولا التكرار المتجمع الصاعد:

الفترات	التكرار f_i	التكرار المتجمع الصاعد
20 − 10	22	22
30 − 20	18	40
40 − 30	15	55
50 − 40	12	67
60 − 50	11	78
70 − 60	10	88
80 − 70	16	104
90 − 80	8	112
المجموع $\sum$	112	

1- الربيع الأول (الأدنى) Q_1:

رتبة الربيع الأدنى = $\frac{\sum f_i}{4} = \frac{112}{4}$ = 28 وعليه فإن فترة الربيع الأدنى هي (20 − 30).

$L = 20 \quad F = 22 \quad f = 18 \quad c = 30 - 20 = 10 \quad n = 112$

$$Q_1 = L + \frac{\frac{n}{4} - F}{f} * c = 20 + \frac{\frac{112}{4} - 22}{18} * 10 = 23.333$$

2- الربيع الثالث (الأعلى) Q_3:

رتبة الربيع الأدنى = $\frac{3*\sum f_i}{4} = \frac{3*112}{4}$ = 84 وعليه فإن فترة الربيع الأدنى هي (60 − 70).

$L = 60 \quad F = 78 \quad f = 10 \quad c = 70 - 60 = 10 \quad n = 112$

$$Q_1 = L + \frac{\frac{n}{4} - F}{f} * c = 60 + \frac{\frac{3*112}{4} - 78}{10} * 10 = 66$$

3- المئوي العاشر P_{10}:

رتبة المئوي العاشر = $\frac{10*\sum f_i}{100} = \frac{10*112}{100}$ = 11.2 وعليه فإن فترة الربيع الأدنى هي (10 − 20).

$L = 10 \quad F = 0 \quad f = 22 \quad c = 20 - 10 = 10 \quad n = 112$

$$P_{10} = L + \frac{\frac{i * n}{100} - F}{f} * c = 10 + \frac{\frac{10 * 112}{100} - 0}{22} * 10 = 15.091$$

4- المئوي التسعون P_{90}:

رتبة المئوي التسعون = $\frac{90*\sum f_i}{100} = \frac{90*112}{100}$ = 100.8 وعليه فإن فترة الربيع الأدنى هي (70 − 80).

$L = 70 \quad F = 88 \quad f = 16 \quad c = 80 - 70 = 10 \quad n = 112$

$$P_{90} = L + \frac{\frac{i*n}{100} - F}{f} * c = 70 + \frac{\frac{90*112}{100} - 88}{16} * 10 = 78$$

مقاييس التشتت:

لا تعتبر مقاييس النزعة المركزية كافية لوصف مجموعة من البيانات وصفا كاملا فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات)، فالعينات التالية ذات وسط حسابي واحد (8) ولكنها لا شك تختلف عن بعضها.

عينة 1	8	8	8	8	8
عينة 2	4	3	6	16	11

تعريف مفهوم التشتت: إذا كانت مجموعة البيانات متباعدة او متباينة عن بعضها يقال انها مشتتة، اما إذا كانت البيانات متجانسة وغير متباعدة فيقال انها غير مشتتة.

فالوسط الحسابي يمثل مركز البيانات لكنه لا يبين مدى التفاف او بعثرة البيانات حول هذا الوسط، ولهذا لا بد من وجود مقاييس اخر مع المقاييس المركزية لقياس درجة التجانس او التشتت في داخل هذه البيانات.

ان الدرجة التي تتجه بها البيانات الرقمية للانتشار حول القيمة الوسطية تسمى تشتت او توزيع البيانات.

ويمكن تقسيم مقاييس التشتت الى مجموعتين:

1- مقاييس تقيس تقارب او تباعد القيم عن بعضها البعض وهي: المدى والانحراف الربيعي (او نصف المدى الربيعي).
2- مقاييس تقيس قرب او بعد القيم من قيمة معينة كالمتوسط الحسابي مثلا وهي: الانحراف المتوسط المطلق والانحراف المعياري.

اولاً: المدى (R):

يستخدم هذا المقياس عندما يكون الهدف هو الحصول على مقياس سريع لمدى تشتت المفردات دون الاهتمام الكبير بالدقة في القياس او حين ما يكون للمفردات المتطرفة أهمية خاصة.

ويعرف المدى على انه الفرق بين أكبر قيمة وأصغر قيمة.

أ- طريقة حساب المدى من البيانات الأولية:

المدى = أكبر قيمة – أصغر قيمة

$$R = x_n - x_1$$

مثال24: احسب المدى للبيانات التالية:

80 350 100 150 90 110 300 250 200 95

الحل:

حيث ان أكبر قيمة في البيانات هي 350 وأصغر قيمة في البيانات هي 80 وعليه فإن:

$$R = 350 - 80 = 270$$

ب-طريقة حساب المدى من الجداول التكرارية:

ويتم حساب المدى بطريقتين:

1- المدى (R) = الحد الاعلى للفترة الأخيرة – الحد الأدنى للفترة الأولى.

$$R = U_k - L_1$$

حيث ان: U_k: نهاية الفترة الأخيرة، L_1: بداية الفترة الأولى.

2- المدى (R) = مركز الفترة الأخيرة – مركز للفترة الاولى

$$R = \mathrm{M}_k - M_1$$

حيث ان: M_k: مركز الفترة الأخيرة، M_1: مركز الفترة الأولى.

مثال25: احسب المدى للجدول التالي:

الفئات	20-16	24-20	28-24	32-28	36-32
عدد المبعوثين f_i	10	15	40	20	15

الحل:

من الصيغة (1): $R = 36 - 16 = 20$

من الصيغة (2): $R = 34 - 18 = 16$

بعض خواص المدى:

1- إذا كانت جميع المفردات متساوية فإن المدى يساوي صفر، أي انه لا يوجد تشتت.
2- مقياس مضلل في حالة وجود قيم متطرفة (شاذة).
3- لا يمكن حسابه من الجداول التكرارية المفتوحة.
4- بسيط في الحساب وسهل في المفهوم ولذلك فهو كثير الاستخدام في الأوساط العامة.
5- لا يعتمد في حسابه على كل البيانات.

ثانيا: الانحراف الربيعي (Q.D.) Quartile Deviation

يستخدم هذا الانحراف للتخلص من بعض عيوب المدى والتي من أهمها تأثره بالقيم المتطرفة وعدم إمكانية حسابه في حالة الجداول التكرارية المفتوحة.

فكرة هذا المقياس هي اهمال الربع الأول والأخير من البيانات المرتبة ترتيبا تصاعدياً.

ويتم حساب هذا المقياس سواء من البيانات الأولية او الجداول التكرارية كما يلي:

$$\text{Q. D} = Q_3 - Q_1$$

مثال26: من البيانات التالية، أوجد الانحراف الربيعي؟

الفترات	10 –	20 –	30 –	40 –	50 –	60 –	70 –	80 – 90
التكرار f_i	22	18	15	12	11	10	16	8

الحل:

حيث أن: $Q_1 = 23.333$ و $Q_3 = 66$ (راجع المثال في درس الربيعيات) وعليه فإن:

$$\text{Q. D} = Q_3 - Q_1 = 66 - 23.333 = 42.667$$

بعض خواص الانحراف الربيعي:

1- لا يتأثر بالقيم المتطرفة (الشاذة).
2- يمكن الاستعانة بالطريقة البيانية في حسابه.
3- يستخدم كمقياس للتشتت في التوزيعات التكرارية شديدة الالتواء.
4- يمكن حسابه في حالة الجداول التكرارية المفتوحة.
5- تتحدد بعدد البيانات وليس بقيمتها.
6- هو عبارة عن فترة تحتوي على 50% من البيانات.

ثالثاً: الانحراف المتوسط Mean Deviation (M.D)

الفكرة الأساسية في هذا المقياس هي قياس مدى تباعد (انحراف) القيم عن وسطها الحسابي بغض النظر فيما إذا كان ذلك الانحراف سلبيا او إيجابيا، فكلما كانت تلك القيم قريبة من وسطها دل ذلك على تجانسها والعكس صحيح.

أ- حساب الانحراف المتوسط في حالة البيانات الأولية:

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

مثال27:

اوجد الانحراف المتوسط للبيانات التالية:

10 7 5 3 2

الحل:

يمكن توضيح خطوات الحل في الجدول التالي:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{27}{5} = 5.4$$

$\lvert x_i - \bar{x} \rvert$	$x_i - \bar{x}$	x_i	
3.4	−3.4	2	
2.4	−2.4	3	
0.4	−0.4	5	
1.6	1.6	7	
4.6	4.6	10	
12.4	0	27	المجموع ∑

مجموع انحرافات القيم عن وسطها الحسابى = صفر

$$\text{M. D} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} = \frac{12.4}{5} = 2.48$$

ب-طريقة حساب الانحراف المتوسط من الجداول التكرارية:

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n} \qquad \text{M. D} = \frac{\sum_{i=1}^{k} f_i |x_i - \bar{x}|}{n}$$

مثال28:

اوجد الانحراف المتوسط للبيانات التالية:

الفترات	0 −	2 −	4 −	6 −	8 −10
التكرار f_i	2	5	4	8	1

الحل:

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n} = \frac{102}{20} = 5.1$$

الفترات	التكرار f_i	مراكز الفترات x_i	$f_i\, x_i$	$\lvert x_i - \bar{x}\rvert$	$f_i \lvert x_i - \bar{x}\rvert$
2 – 0	2	1	2	4.1	8.2
4 – 2	5	3	15	2.1	10.5
6 – 4	4	5	20	0.1	0.4
8– 6	8	7	56	1.9	15.2
10 – 8	1	9	9	3.9	3.9
المجموع ∑	20		102		38.2

$$\text{M. D} = \frac{\sum_{i=1}^{k} f_i \lvert x_i - \bar{x}\rvert}{n} = \frac{38.2}{20} = 1.91$$

بعض خواص الانحراف المتوسط:

1- يتأثر بالقيم المتطرفة (الشاذة).

2- لا يمكن حسابه في حالة الجداول التكرارية المفتوحة.

3- يعتمد في حسابه على كل القيم.

4- قليل الاستخدام في الإحصاء الاستنتاجي لعدم قابليته للعمليات الجبرية.

5- يمكن حسابه عن طريق الانحرافات عن الوسيط مع ملاحظة ان الانحراف المتوسط عن الوسيط أقل من الانحراف المتوسط عن الوسط الحسابي.

رابعاً: التباين والانحراف المعياري: Variance Standard Deviation (S.D)

مما سبق يتضح ان الانحراف المتوسط أفضل من كل من المدى والانحراف الربيعي وذلك لأنه يعتمد في حسابه على كل البيانات، ولكن من عيوبه عدم قابليته للعمليات الجبرية، لذلك وجد مقياس آخر لظاهرة التشتت وهو ما يسمى الانحراف المعياري.

يعرف متوسط مربعات انحراف البيانات عن وسطها الحسابي على انه تباين مجموعة هذه البيانات (ويرمز له بالرمز S^2)، ويعرف الجذر التربيعي للتباين على انه الانحراف المعياري للبيانات (ويرمز له بالرمز S.D).

1) حساب التباين والانحراف المعياري للمجتمع:

إذا توفرت لدينا جميع البيانات للمجتمع قيد الدراسة، فإن التباين والانحراف المعياري لهذا المجتمع يمكن حسابه كما يلي:

أ- حساب التباين والانحراف المعياري في حالة البيانات الأولية (للمجتمع):

$$s^2 = \frac{\sum(x_i-\bar{X})^2}{n} \quad \text{or} \quad s^2 = \frac{\sum x_i{}^2-n\bar{X}^2}{n}$$

$$S.D = \sqrt{\frac{\sum(x-\bar{X})^2}{n}} = \sqrt{\frac{\sum x_i{}^2-n\bar{X}^2}{n}} \qquad \mathrm{S.D} = \sqrt{s^2}$$

مثال29: احسب الانحراف المعياري للقيم التالية:

35	17	22	32	19
48	13	19	20	

الحل:

- نحسب المتوسط:

$$\bar{X} = \frac{\sum x}{n} = \frac{35+17+22+32+19+48+13+19+20}{9} = \frac{225}{9} = 25$$

يمكن حساب التباين:

$$S^2 = \frac{\sum(x-\bar{X})^2}{n} = \frac{992}{9} = 110.222$$

ومنه نوجد الانحراف المعياري:

$$\mathrm{S.D} = \sqrt{S^2} = \sqrt{110.222} = 10.499$$

x_i	$x_i - \bar{X}$	$(x_i - \bar{X})^2$
35	10	100
17	−8	64
22	−3	9
32	7	49
19	−6	36
48	23	529
13	−12	144
19	−6	36
20	−5	25
225	--	992

ب-طريقة حساب التباين والانحراف المعياري من الجداول التكرارية (للمجتمع):

$$s^2 = \frac{\sum f_i(x_i-\bar{X})^2}{n} \quad \text{or} \quad s^2 = \frac{\sum f_i {x_i}^2 - n\bar{X}^2}{n}$$

$$\text{S. D} = \sqrt{s^2}$$

مثال30: احسب الانحراف المعياري من الجدول التالي (بيانات المجتمع بالكامل):

فئات الدخل	20-10	30-20	40-30	50-40	60-50	70-60	80-70
عدد العمال f_i	5	12	22	38	22	12	5

الحل:

- نحسب المتوسط: $\bar{X} = \frac{\sum xf}{\sum f} = \frac{5220}{116} = 45$

الفئات	f	x	xf	x^2f
20-10	5	15	75	1125
30-20	12	25	300	7500
40-30	22	35	770	26950
50-40	38	45	1710	76950
60-50	22	55	1210	66550
70-60	12	65	780	50700
80-70	5	75	375	28125
$\sum$	116		5220	257900

ومنه نحسب التباين:

$$s^2 = \frac{\sum f_i {x_i}^2 - n\bar{X}^2}{n} = \frac{257900 - 116*45^2}{116} = 198.276$$

ومنه نوجد الانحراف المعياري:

$$\text{S. D} = \sqrt{s^2} = \sqrt{198.276} = 14.081$$

2) حساب التباين والانحراف المعياري للعينة:

إذا لم تتوفر لدينا جميع البيانات للمجتمع قيد الدراسة فيتم اخذ عينة من هذا المجتمع، ويتم حساب التباين والانحراف المعياري لهذا العينة وتعمم النتائج على المجتمع، ويمكن حساب التباين والانحراف المعياري للعينة كما يلي:

أ- حساب التباين والانحراف المعياري في حالة البيانات الأولية (للعينة):

$$s^2 = \frac{\sum(x_i-\bar{X})^2}{n-1} \quad \text{or} \quad s^2 = \frac{\sum x_i{}^2-n\bar{X}^2}{n-1}$$

$$\text{S. D} = \sqrt{s^2} \qquad S.D = \sqrt{\frac{\sum(x-\bar{X})^2}{n-1}} = \sqrt{\frac{\sum x_i{}^2-n\bar{X}^2}{n-1}}$$

ب-طريقة حساب التباين والانحراف المعياري من الجداول التكرارية (للعينة):

$$s^2 = \frac{\sum f_i(x_i-\bar{X})^2}{n-1} \quad \text{or} \quad s^2 = \frac{\sum f_i x_i{}^2-n\bar{X}^2}{n-1}$$

$$\text{S. D} = \sqrt{s^2}$$

بعض خواص الانحراف المعياري:

1- لا يمكن حسابه من الجداول التكرارية المفتوحة.

2- يتأثر بالقيم المتطرفة وذلك لأنه يعتمد في حسابه على المتوسط الحسابي الذي بدوره يتأثر بها.

3- قابل للعمليات الجبرية.

4- اذا كانت الظاهرة X تأخذ القيم $x_1, x_2, x_3, \dots, x_m$ و $\bar{x}$ تمثل المتوسط الحسابي و $s_x{}^2$ تمثل التباين لها، وكانت الظاهرة Y تأخذ القيم $y_1, y_2, y_3, \dots, y_n$ و $\bar{y}$ تمثل المتوسط الحسابي و $s_y{}^2$ تمثل التباين لها، فإن التباين المشترك للظاهرتين معاً نرمز له بالرمز $s_p{}^2$ (*Pooled Variance*)، معرف كما يلي:

$$s_p{}^2 = \frac{1}{m+n}\{ms_x{}^2 + ns_y{}^2 + \frac{m*n}{m+n}(\bar{x}-\bar{y})^2\}$$

وإذا كانت $\bar{x} = \bar{y}$ فإن:

$$s_p{}^2 = \frac{1}{m+n}\{ms_x{}^2 + ns_y{}^2\}$$

حيث أن:

$$s_x{}^2 = \frac{\sum_{i=1}^{m}(x_i-\bar{x})^2}{m} \qquad s_y{}^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n}$$

5- يعد أهم المقاييس التشتت وأكثرها استخداماً.

خامساً: معامل الاختلاف Coefficient of variation (c.v)

معامل الاختلاف هو عبارة عن النسبة المئوية للانحراف المعياري منسوبا الى المتوسط الحسابي.

حيث انه كلما كانت قيمة هذا المعامل كبيرة دل ذلك على وجود تشتت كبير بين مفردات التوزيع والعكس صحيح.

$$c.v = \frac{s.d}{\bar{x}} * 100$$

حيث ان:

$c.v$: التشتت المطلق او النسبي.

$s.d$: الانحراف المعياري.

$\bar{x}$: الوسط الحسابي.

مثال31:

إذا علمت أن متوسط درجة الحرارة في مدينة طرابلس في شهر مايو السنة الماضية كانت 24°C والانحراف المعياري 3°C. بينما كان متوسط درجة الحرارة في نفس الشهر في السنة الماضية في مدينة سبها هو 34°C والانحراف المعياري 4°C، فأي المدينتين أقل تشتتا (أكثر تجانساً) من حيث درجة الحرارة.

الحل:

1- معامل الاختلاف لدرجات الحرارة في مدينة طرابلس في شهر مايو السنة الماضية هو:

$$c.v = \frac{s.d}{\bar{x}} * 100 = \frac{3}{24} * 100 = 12.5\ \%$$

2- معامل الاختلاف لدرجات الحرارة في مدينة سبها في شهر مايو السنة الماضية هو:

$$c.v = \frac{s.d}{\bar{x}} * 100 = \frac{4}{34} * 100 = 11.77\ \%$$

وعليه فإن درجات الحرارة في مدينة سبها أقل اختلافاً (أكثر تجانساً) من درجات الحرارة في مدينة طرابلس.

بعض خواص معامل الاختلاف:

1- يقيس الاختلاف النسبي دون وحدة تمييز.

2- ليس له معنى إذا كانت المتوسطات الحسابية تساوي صفر.

3- يستخدم في مقارنة مجموعتين أو أكثر من البيانات من حيث التشتت، وخاصة إذا اختلفت المتوسطات الحسابية.

سادساً: العزوم (m') Moment

إذا كانت الظاهرة X تأخذ القيم $x_1, x_2, x_3, \dots, x_m$ و $\bar{x}$ تمثل المتوسط الحسابي لها فإن العزم الرائي حول المتوسط الحسابي يمكن حسابه كما يلي:

أ- في حالة البيانات الأولية معرف كما يلي:

$$m'_r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^r}{n} \qquad , r = 1, 2, 3, 4, \dots$$

وعليه فإن:

$$r = 1 \quad \rightarrow \quad m'_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})}{n} = 0$$
$$r = 2 \quad \rightarrow \quad m'_2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = S^2$$
$$\vdots$$
$$r = k \quad \rightarrow \quad m'_k = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^k}{n}$$

وإذا كانت $\bar{x} = 0$ فإن العزم الرائي حول نقطة الأصل (الصفر) يكون كالآتي:

$$m_r = \frac{\sum_{i=1}^{n} x_i^r}{n} \qquad , r = 1, 2, 3, 4, \dots$$

ملاحظة:

ويمكن حساب العزم حول أي قيمة أخرى خلاف المتوسط الحسابي وذلك باستبدال المتوسط الحسابي بتلك القيمة في الصيغة أعلاه.

ب-في حالة الجدول التكرارية معرفة كما يلي:

$$m'_r = \frac{\sum_{i=1}^{n} f_i(x_i - \bar{x})^r}{n} \qquad , r = 1, 2, 3, 4, \dots$$

وإذا كانت $\bar{x} = 0$ فإن العزم الرائي حول نقطة الأصل (الصفر) يكون كالآتي:

$$m_r = \frac{\sum_{i=1}^{n} f_i x_i^r}{n} \qquad , r = 1, 2, 3, 4, \dots$$

سابعاً: الالتواء Skewness

الالتواء هو درجة توزيع البيانات حول نقطة التمركز، وينقسم الى نوعين أساسيين:

أ- إذا كان المتوسط الحسابي يمثل نقطة التمركز في التوزيع:

إذا كان التوزيع التكراري متماثلا فإن المتوسط الحسابي والوسيط والمنوال تتساوى جميعها عند النقطة المقابلة لقمة المنحنى وكلما بعد التوزيع عن التماثل كلما اختلفت هذه المتوسطات الثلاث عن بعضها.

وبدلك اقترح بيرسون مقاييس الالتواء عرفه كالاتي:

$$\alpha_1 = \frac{\bar{x}-m}{s.d}$$

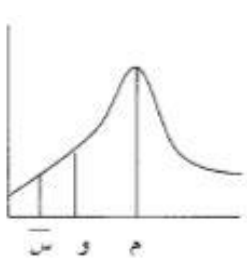
حيث ان:

$\bar{x}$: المتوسط الحسابي.

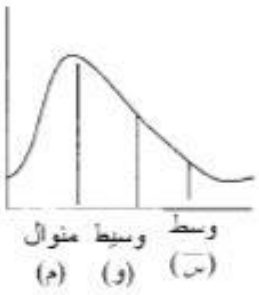
m: المنوال.

$s.d$: الانحراف المعياري.

- الالتواء = صفر المنحنى متماثل.
- الالتواء > الصفر الالتواء موجب.
- الالتواء < الصفر الالتواء سالب.



ملتو نحو اليسار (سالب)



ملتو نحو اليمين (موجب)

معتدل (توزيع طبيعي)

ملاحظة:

المعادلة السابقة تعتمد على المنوال (وهو مقياس غير دقيق) وبالتالي لا يجب الاعتماد عليه في قياس الالتواء، ويفضل استخدام المعادلة التالية في حالة التوزيعات القريبة من التماثل (المعتدلة):

$$\alpha_2 = \frac{3(\bar{x} - \tilde{x})}{s.d}$$

حيث ان:

$\tilde{x}$: الوسيط.

$\bar{x}$: المتوسط الحسابي.

$s.d$: الانحراف المعياري.

مثال32:

إذا علمت ان $\bar{x} = 74.92$ و $m = 76.33$ و $\tilde{x} = 75.227$ و $s.d = 22.16996$

الحل:

التواء سالب $\alpha_1 = \frac{\bar{x}-m}{s.d} = \frac{74.92-76.333}{22.16996} = -0.0637$

التواء سالب $\alpha_2 = \frac{3(\bar{x}-\tilde{x})}{s.d} = \frac{3*(74.92-75.227)}{22.16996} = -0.0415$

نلاحظ ان القيمة تختلف في الحالتين وذلك لان كل معادلة لها مفهوم مختلف، ولكن نلاحظ ان الخلاصة هي نفسها (الالتواء سالب).

ب-إذا كان الوسيط يمثل نقطة التمركز في التوزيع:

إذا كان التوزيع مفتوحا فإن لالتواء في هذه الحالة يسمى معامل الالتواء الربيعي، يقاس بالعلاقة التالية:

$$\alpha_3 = \frac{(Q_3-\tilde{x})-(\tilde{x}-Q_1)}{Q_3-Q_1}$$

حيث ان:

Q_3: الربيع الثالث.

Q_1: الربيع الأول.

$\tilde{x}$: الوسيط.

ثامناً: التفرطح Kurtosis

إذا كانت البيانات أكثر تجمعا حول المنوال أي ان قاعدة التوزيع ضيقة وطرفاه مرتفعين فإن التوزيع مدبب.

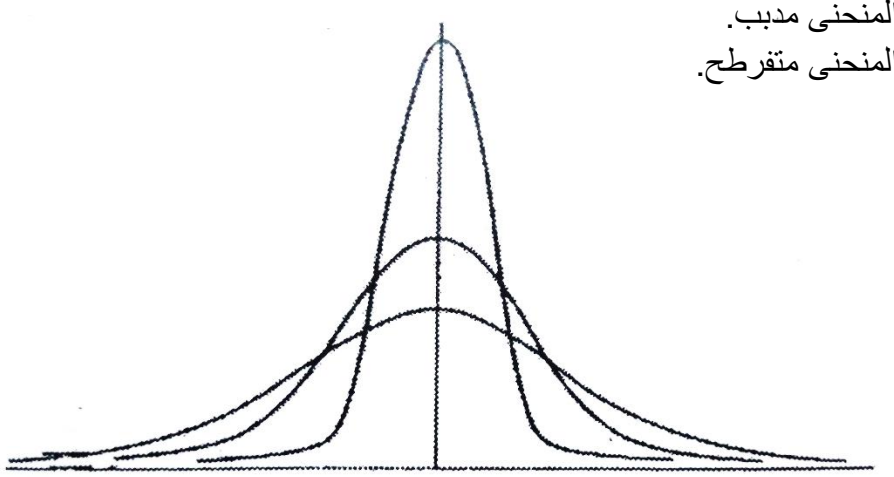
اما إذا كانت البيانات كثيرة على الجانبين أي ان قاعدة المنحنى واسعة وطرفاه منخفضين فإن المنحنى متفرطح.

ويمكن قياسه كما يلي:

$$\gamma = \frac{m'_4}{(s.d)^4}$$

حيث ان: m'_4: هو العزم الرابع حول المتوسط الحسابي.

- $\gamma = 3$ فإن المنحنى معتدل التفرطح.
- $\gamma > 3$ فإن المنحنى مدبب.
- $\gamma < 3$ فإن المنحنى متفرطح.



ويمكن قياس التفرطح بمقياس أخر يسمى معامل التفرطح المئوي (β):

$$\beta = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

- $\beta = 0.263$ فإن المنحنى معتدل التفرطح.
- $\beta < 0.263$ فإن المنحنى مدبب.
- $\beta > 0.263$ فإن المنحنى متفرطح.

مثال33:

إذا علمت ان $P_{10} = 15.0909$ و $P_{90} = 78$ و $Q_1 = 23.333$ و $Q_3 = 66$

اوجد معامل التفرطح المئوي؟

الحل:

$$\beta = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} = \frac{66 - 23.333}{2*(78 - 15.0909)} = \frac{42.667}{125.8182} = 0.3391$$

حيث ان $\beta = 0.3391 > 0.263$ فإن المنحنى متفرطح

الفصل الرابع:

الاحتمـــــــالات

مقدمة:

يعرف الاحتمال على انه مقدار ثقتنا في إمكانية حدوث شيء غير مؤكد الوقوع.

ان فهم القواعد الأساسية لنظرية الاحتمالات يساعد في زيادة الدور الذي تلعبه الاحتمالات في مجال الإحصاء التحليلي وتمكن من تقييم مدى دقة الاستنتاجات الإحصائية او مدى الثقة التي يمكن وضعها في هذه النتائج.

أهمية الاحتمالات:

تستخدم الاحتمالات كأداة لاتخاذ القرارات في جميع المجالات الغير مؤكدة الحدوث، لذلك فإن الاحتمالات تظهر تطبيقاتها في المجالات الهندسية بمختلف أنواعها وخاصة في مجالي الهندسة الصناعية وهندسة الحاسوب والاتصالات.

التجربة العشوائية: Random Experiments

من المعروف أهمية التجارب في مجال العلوم الهندسية، وبفرض ان اجراء التجارب تحت شروط متقاربة سوف يعطي نتائج متساوية، وفي هذه الظروف نستطيع تحديد قيم المتغيرات التي تؤثر على نتائج التجربة.

في بعض التجارب لا نتمكن من تحديد قيم بعض المتغيرات وبالتالي سوف تتغير النتائج من تجربة الى أخرى مع ان معظم الشروط تظل كما هي، توصف هذه التجارب بالتجارب العشوائية.

1- عند رمي زهرة نرد مرة واحدة، فالنتائج التي سوف تحدث تكون أحد الأرقام التالية:

$$\{1,2,3,4,5,6\}$$

2- عند رمي قطعة نقود مرة واحدة، فالنتائج التي سوف تحدث تكون أحد الاحتمالين:

{صورة , كتابة}

3- عند فحص صندوق به 100 جهاز كهربائي تم اختياره من بين مجموعة من الصناديق التي بها أجهزة متشابهة يمكن تحديد احتمال عدم وجود أكثر من جهاز واحد به عطل في ذلك الصندوق.

نظرية المجموعات: Set Theory

تعرف المجموعة على انها أي تجمع من الأشياء التي تشترك في صفة معينة، في العادة ما يرمز للمجموعة بأحد الحروف الهجائية الكبيرة مثل A او B او C او . . . ..

ويسمى كل عضو من أعضاء المجموعة عنصراً ويرمز له غالبا بأحد الحروف الهجائية الصغيرة مثل a او b او c او . . . ..

ويتم حصر عناصر المجموعة بقوسين من النوع { }.

يرمز لانتماء عنصر ما الى مجموعة معينة الرمز (∈) وعدم الانتماء بالرمز (∉).

يمكن ان تعرف المجموعة بذكر الخواص التي تحقق عناصرها، فمثلاً إذا كانت C هي مجموعة الاعداد الفردية فإن المجموعة C يمكن التعبير عنها كالتالي:

$$C = \{x: \text{عدد فردي } x\}$$

وتقرأ: المجموعة C تساوي العنصر x حيث العنصر x عدد فردي.

بعض التعريفات:

1- المجموعات الشاملة Universal Set

هي المجموعة التي تحتوي على جميع عناصر الظاهرة قيد الدراسة، ويرمز لها بالرمز S، فمثلا اذا كانت المجموعة الشاملة هي مجموعة الاعداد الحقيقية فإن:

$$S = \{x: \text{عدد حقيقي } x\}$$

2- المجموعة الخالية Empty Set

هي المجموعة التي لا تحتوي على عناصر، ويرمز لها بالرمز { } أو ∅.

3- المجموعة الجزئية Subset

هي المجموعة التي تكون جميع عناصرها موجودة في مجموعة أخرى، فمثلا اذا كانت $A = \{1,2,3,4,5,6\}$ وكانت $B = \{1,3,5\}$ فإن المجموعة B جزء من المجموعة A وتكتب $B \subset A$ ، كما نلاحظ ان المجموعة الخالية هي جزء من أي مجموعة أخرى.

وبصفة عامة اذا كان عدد عناصر المجموعة هو n فإن عدد المجموعات الجزئية لهذه المجموعة هو 2^n، فمثلا اذا كانت $A = \{1,2,3\}$ فإن عدد المجموعات الجزئية للمجموعة A هو $(2^n=2^3=8)$، وهذه المجموعات الجزئية هو: $\{\ \}, \{3\}, \{2\}, \{1\}, \{2,3\}, \{1,3\}, \{1,2\}, \{1,2,3\}$

4- المجموعة المكملة Complement Set

هي المجموعة التي تحتوي جميع العناصر الموجودة في المجموعة الشاملة S وغير موجودة في المجموعة الجزئية. فمثلا إذا كانت المجموعة الشاملة {1,2,3,4,5,6,7,8,9} =S والمجموعة الجزئية {1,5,7}=A فإن المجموعة المكملة للمجموعة A في المجموعة الشاملة S هي A' او A^c

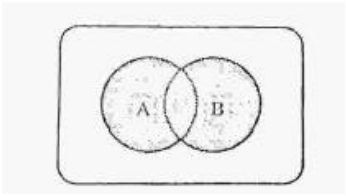
حيث ان: $A' = \{2,3,4,6,8,9\}$.

العمليات على المجموعات:

1- الاتحاد The Union

إذا كانت $A_1, A_2, A_3, \cdots, A_n$ مجموعات جزئية في S فإن اتحاد هذه المجموعات يمكن كتابته في الصورة التالية:

$$\cup_{i=1}^{n} A_i = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n$$

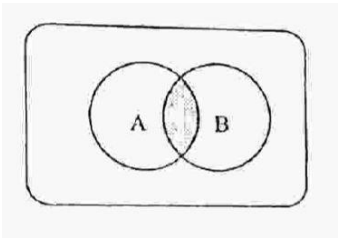


$A \cup B$

2- التقاطع The Intersection

إذا كانت $A_1, A_2, A_3, \cdots, A_n$ مجموعات جزئية في S فإن تقاطع هذه المجموعات يمكن كتابته في الصورة التالية:

$$\cap_{i=1}^{n} A_i = A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n$$

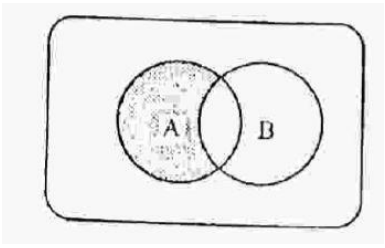


$A \cap B$

3- الفرق The Difference

إذا كانت A و B مجموعتين جزئيتين في S فإن الفرق بينهما يرمز له بالرمز (A – B) ويعرف على انه جميع العناصر الموجودة في المجموعة A باستثناء العناصر المشترك مع B، ويمكن كتابه على الصورة:

$$A - B = \{x : x \in A, x \notin B\}$$



$A - B$

العلاقات بين المجموعات:

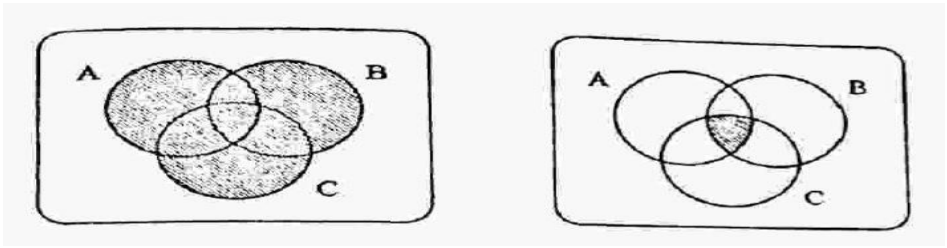
إذا كانت A و B و C ثلاث مجموعات جزئية من المجموعة الشاملة S فإن:

أ- قانون التبديل Commutative Law

$$A \cap B = B \cap A \qquad A \cup B = B \cup A$$

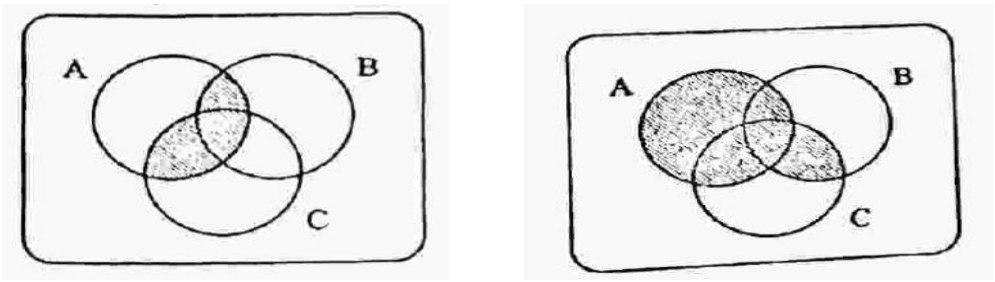
ب- قانون التنسيق Associative Law

$$(A \cup B) \cup C = A \cup (B \cup C) \qquad (A \cap B) \cap C = A \cap (B \cap C)$$



ت- قانون التوزيع Distributor Law

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \qquad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$



$A \cap (B \cup C)$ $A \cup (B \cap C)$

ث- قانون المكملة Complementary Law

$$A \cup A' = S \qquad A \cap A' = \emptyset \qquad (A')' = A$$

$$A \cup S = S \qquad A \cap S = A \qquad A \cup \emptyset = A \qquad A \cap \emptyset = \emptyset$$

ج- قانون الفرق Difference Law

$$A - B = A \cap B'$$

$$A - B = A - (A \cap B) = (A \cup B) - B$$

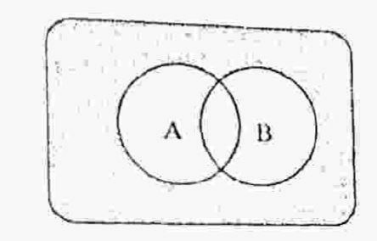
$$(A \cup B) - C = (A - C) \cup (B - C)$$

$$(A \cap B) \cup (A - B) = A \qquad (A \cap B) \cap (A - B) = \emptyset$$

ح- قانون دى مورجان De Morgan's Law

$$(A \cup B)' = A' \cap B'$$

$$(A \cap B)' = A' \cup B'$$



$(A \cup B)'$

5- المجموعة القابلة للعد *Countable Set*

هي المجموعة التي يمكن عد عناصرها، فمثلا $A= \{1,2,3,\dots\}$ مجموعة قابلة للعد، بينما المجموعة $B= \{x: 0 < x < 10\}$ غير قابلة للعد.

6- المجموعة المحدودة (المنتهية) *Finite Set*

هي المجموعة التي تحتوي على عدد محدود من العناصر، فمثلا $A=\{1,2,3, \dots , n\}$ مجموعة عدد عناصرها n عنصر، بينما المجموعة $C=\{1,2,3, \dots \}$ و $C=\{x: 0<x\}$ مجموعات غير منتهية.

فراغات العينة والأحداث: Sample Spaces and Events

فراغ العينة:

المجموعة التي تتكون من كل النتائج الممكنة للتجربة العشوائية يطلق عليها فراغ العين(Sample Space)، ويرمز له بالرمز (Ω) (Omega)، ويطلق على كل نتيجة نقطة من العينة Sample Point.

مثال1:

إذا ألقيت قطعة نقود معدنية متزنة مرتين فإن فراغ العينة يكون: $\Omega = \{HH, HT, TH, TT\}$

حيث ان: T: ترمز للكتابة H: ترمز للصورة.

مثال2:

إذا تم اختيار نقطة من الفترة [1,2]، فإن فراغ العينة في هذه الحالة هو: $\Omega = \{x: 1 \leq x \leq 2\}$ وهو يحتوي على عدد من النقاط غير القابلة للعد.

مثال3:

اذا كانت التجربة تتضمن إلقاء قطعة نقود معدنية متزنة حتى تظهر صورة H ففي هذه الحالة يكون فراغ العينة هو: $\Omega = \{H, TH, TTH, TTTH, \dots\}$ وهي مجموعة لا نهائية من النتائج وقابلة للعد.

مثال4:

إذا تم إلقاء مكعب (زهرة) نرد مرة واحدة فإن فراغ العينة: $\Omega = \{1,2,3,4,5,6\}$

مميزات فراغ العينة:

1- ان فراغ العينة Ω مناظر للمجموعة الشاملة، وبمجرد اختياره يبقى ثابتا وسيكون كل النقاش متعلق بهذا الفراغ.

2- من الممكن ان تكون النقاط او النتائج التي تحتويها Ω عددية او وصفية.

3- من الممكن ان يحتوي فراغ العينة على عدد محدود او غير محدود من العناصر.

طرق عد عناصر فراغ (فضاء) العينة:

1- قاعدة الضرب:

بفرض ان لدينا تجربتين عشوائيتين A ولديها عدد n من النتائج المختلفة و b ولديها عدد m من النتائج المختلفة، فيمكن حساب عدد النتائج الممكنة من التجربتين معا عن طريق ضرب عدد النتائج ببعضها (n*m).

مثال5:

لنفرض وجود ثلاثة طرق مختلفة تؤدي من المدينة A الى المدينة B، وتوجد خمسة طرق مختلفة تؤدي من المدينة B الى المدينة C، ما هو عدد الطرق التي يمكن استخدامها للتوجه من المدينة A الى المدينة C مروراً بالمدينة B؟

الحل:

عدد الطرق التي تؤدي من المدينة A الى المدينة C مرورا بالمدينة B تكون:

$$n * m = 3 * 5 = 15$$

مثال6:

كم عدد النقاط بفضاء العينة إذا تم إلقاء مكعبي (زهرتي) نرد متزنين مرة واحدة؟

الحل:

$$n * m = 6 * 6 = 36$$

ملاحظة: يمكن تعميم قاعدة الضرب لتشمل عدد k من التجارب العشوائية، حيث أن:

$$\prod_{i=1}^{k} n_i = n_1 * n_2 * n_3 * \ \dots \ * \ n_k$$

مثال7: إذا تم إلقاء قطعة نقود معدنية متزنة 5 مرات، فما هو عدد فراغ العينة؟

الحل:

$$n_1 * n_2 * n_3 * n_4 * n_5 = 2 * 2 * 2 * 2 * 2 = 2^5 = 32$$

2- قاعدة الجمع:

إذا كانت التجربة A تحدث في n من النتائج الممكنة والتجربة B تحدث في m من النتائج المحتملة وكانت التجربتين متنافيتين فإن عدد النتائج الممكنة من التجربة A او التجربة B هو (m+n) من النتائج.

مثال8: لنفرض أن لدينا ثلاثة طرق تؤدي من المدينة A الى المدينة B، وخمسة طرق تؤدي من المدينة A الى المدينة C، وستة طرق تؤدي من المدينة A الى المدينة D، فأوجد عدد الطرق التي تؤدي من المدينة A الى إحدى المدن الثلاث؟

الحل: $n_1 + n_2 + n_3 = 3 + 5 + 6 = 14$

الحـــــدث: Event

الحدث هو مجموعة جزئية Subset A من مجموعة فراغ العينة Ω، ويرمز له في العادة بالرموز A او B او C ...، فاذا كان ناتج التجربة العشوائية يمثل عنصرا في الحدث A فإننا نقول ان الحدث A يتحقق، والحدث الذي يتكون من نقطة واحدة من فراغ العينة S يسمى الحدث البسيط Simple Event او الحدث الابتدائي Elementary Event وحينما يكون الحدث هو S نفسها يمثل الحدث المؤكد Sure or Certain Event حيث انه من الضروري ان يحدث عنصر من S، بينما المجموعة الفارغة Empty set $\emptyset$ والتي تسمى بالحدث المستحيل Impossible Event لان أي عنصر في $\emptyset$ لا يمكن حدوثه.

وباستخدام عمليات المجموعات على الاحداث في فراغ العينة Ω، فإننا نحصل على احداث أخرى في نفس فراغ العينة Ω، فاذا كان لدينا A,B تمثل احداثا فان:

1- الحدث المكمل Complement Event

هو الحدث الذي يحتوي على جميع نتائج التجربة العشوائية (عناصر فراغ العينة) التي لا يشملها الحدث A، ويرمز له بالرمز A^c او A'.

2- الحدث المؤكد Sure Event

هو الحدث الذي يحتوي على جميع عناصر فراغ العينة، حيث ان $A = \Omega$.

3- الحدث المستحيل Impossible Event

هو الحدث الذي لا يحتوي أي نتيجة من نتائج فراغ العينة.

4- الاحداث المتنافية Mutually Exclusive Event

هي الاحداث المانعة لبعضها البعض، أي انه إذا حدث الحدث A يمنع حدوث الحدث B، وبالتالي فإن:

$$A \cap B = \emptyset$$

5- الاحداث المستقلة Independent Event

هي الاحداث التي إذا حدث أحدها فإنه لا يؤثر على حدوث الآخر.

مثال9:

إذا ألقيت قطعة نقود معدنية متزنة ومكعب (زهرة) نرد معا مرة واحدة فإن فراغ العينة يكون:

$$\Omega = \{(H,1),(H,2),(H,3),(H,4),(H,5),(H,6),(T,1),(T,2),(T,3),(T,4),(T,5),(T,6)\}$$

المسلمات الاحتمالية Axioms of Probability

إذا تم اجراء تجربة عشوائية بفراغ عينة Ω وكان A يمثل حدث معرف على هذا الفراغ، فإذا تم إعادة اجراء نفس التجربة عدد n من المرات وكان m عدد المرات التي يظهر فيها الحدث A فإن احتمال حدوث الحدث A عندما تقترب n من ∞ هو:

$$P_A = \lim_{n\to\infty} \frac{m}{n}$$

حيث ان:

P_A: احتمال حدوث الحدث A.

n: عدد مرات اجراء التجربة العشوائية.

m: عدد المرات التي يظهر فيها الحدث A من مرات إعادة التجربة العشوائية n.

عند اجراء تجربة عشوائية بفراغ عينة Ω، فإن الدالة $\mathrm{P}(\mathrm{A})$ تسمى دالة احتمالية إذا تحققت الشروط التالية:

1- $0 \le \mathrm{P}(\mathrm{A}) \le 1$ لأي حدث A ينتمي لفراغ العينة Ω.

2- $\mathrm{P}(\Omega) = 1$ و $\mathrm{P}(\emptyset) = 0$

3- إذا كانت $\{A_n\}$ تمثل متوالية من الأحداث المتنافية $\{A_i \cap A_j = \emptyset \ , i \neq j\}$ فإن:

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i)$$

هذه الشروط الثلاثة تسمى مسلمات الاحتمال التي من خلالها يمكن تحديد مفهوم الاحتمال.

مثال10:

إذا ألقيت قطعة نقود معدنية متزنة ثلاث مرات، حدد عناصر فراغ العينة ثم أوجد احتمال حدوث الأحداث التالية:

1- الحصول على صورة واحدة على الأقل.

2- الحصول على ثلاثة وجوه متشابهة.

3- الحصول على صورتين على الأقل.

الحل:

عدد عناصر فراغ العينة: $n(\Omega) = 2^3 = 8$

وتكون مجموعة فراغ العينة هي: $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT, \}$

1- الحدث A هو الحصول على صورة واحدة على الأقل، فإن الحدث A يكون:

$$A = \{HHH, HHT, HTH, HTT, THH, TTH, THT\}$$

وبالتالي يكون عدد عناصر الحدث A هو $n(A) = 7$

وعليه فإن احتمال حدوث الحدث A هو:

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{7}{8}$$

2- الحدث B هو الحصول على ثلاثة وجوه متشابهة، فإن الحدث B يكون: $B = \{HHH, TTT\}$

وبالتالي يكون عدد عناصر الحدث B هو $n(B) = 2$

وعليه فإن احتمال حدوث الحدث B هو: $P(B) = \frac{n(B)}{n(\Omega)} = \frac{2}{8}$

3- الحدث C هو الحصول على صورتين على الأكثر، فإن الحدث C يكون:

$$A = \{HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

وبالتالي يكون عدد عناصر الحدث C هو $n(C) = 7$

وعليه فإن احتمال حدوث الحدث C هو: $P(C) = \frac{n(C)}{n(\Omega)} = \frac{7}{8}$

مثال11:

عند إلقاء مكعب (زهرة) نرد متزنة مرة واحدة، حدد عناصر فراغ العينة ثم أوجد احتمال حدوث الأحداث التالية:

1- الحصول على عدد زوجي.

2- الحصول على عدد أكبر من 2 وأقل من أو يساوي 4.

3- الحصول على عدد فردي.

4- الحصول على عدد على الأقل يساوي 2 وأقل من 6.

الحل:

عدد عناصر فراغ العينة يساوي 6 وهي: $\Omega = \{1,2,3,4,5,6\}$

بفرض أن:

- A حدث الحصول على عدد زوجي.
- B حدث الحصول على عدد أكبر من 2 وأقل من أو يساوي 4.
- C حدث الحصول على عدد فردي.
- D حدث الحصول على عدد على الأقل يساوي 2 وأقل من 6.

وعليه فإن: $A = \{2,4,6\} \quad \rightarrow \quad n(A) = 3$

$$B = \{3,4\} \quad \rightarrow \quad n(B) = 2$$
$$C = \{1,3,5\} \quad \rightarrow \quad n(C) = 3$$
$$D = \{2,3,4,5\} \quad \rightarrow \quad n(D) = 4$$

وبالتالي يمكن حساب احتمال حدوث كل حدث من الأحداث السابقة كما يلي:

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6} = 0.5$$
$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{2}{6} = \frac{1}{3}$$
$$P(C) = \frac{n(C)}{n(\Omega)} = \frac{3}{6} = 0.5$$
$$P(D) = \frac{n(D)}{n(\Omega)} = \frac{4}{6} = \frac{2}{3}$$

بعض نظريات الاحتمال:

نظرية (1): لأي حدث A يكون: $P(A') = 1 - P(A)$

المبرهنة:

حيث أن: $P(A \cap A') = \emptyset$ و $P(A \cup A') = 1$

$$P(\Omega) = P(A \cup A') = P(A) + P(A')$$
$$P(A') = 1 - P(A)$$

نظرية (2): لأي حدثين A و B يكون:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

نظرية (3): لأي حدثين A و B يكون:

$$P(A \cup B) \leq P(A) + P(B)$$

نظرية (4): إذا كان $A \subset B$ (A جزء من B) فإن:

$$P(A' \cap B) = P(B) - P(A) \quad \text{و} \quad P(A) \leq P(B)$$

نظرية (5): إذا كان A و B حدثين فإن:

$$P(A \cap B) \geq 1 - P(A') - P(B')$$

مثال12:

إذا كان A و B حدثان متنافيان وكان P(A)=0.10 و P(B)=0.20، فأوجد الاحتمالات التالية:

$$P(A') \ , \ P(B') \ , \ P(A \cup B) \ , \ P(A \cap B) \ , \ P(A' \cap B')$$

الحل:

$$P(A') = 1 - P(A) = 1 - 0.10 = 0.90$$
$$P(B') = 1 - P(B) = 1 - 0.20 = 0.80$$
$$P(A \cup B) = \mathrm{P(A)} + \mathrm{P(B)} = 0.1 + 0.2 = 0.30$$
$$P(A \cap B) = 0$$
$$P(A' \cap B') = P((A \cup B)') = 1 - P(A \cup B) = 1 - 0.30 = 0.70$$

مثال13:

أخذت عينة عشوائية من طلبة قسم الأحياء بكلية العلوم جامعة طرابلس حجمها 60 طالباً، فوجد أن 15 طالب مسجل في مقرر ST305 و12 طالباً مسجل في مقرر ST401 و6 طلبة مسجلين في المقررين معاً، فإذا تم اختيار أحد الطلبة عشوائيا من العينة المختارة.

فأجد احتمال حدوث الأحداث التالية:

1- أن يكون الطالب مسجل في أحد المقررين على الأقل.

2- أن يكون الطالب غير مسجل في أياً من المقررين.

3- أن يكون الطالب مسجل في مقرر ST305 فقط.

الحل:

نفرض أن: A حدث اختيار طالب مسجل في المقرر ST305

B حدث اختيار طالب مسجل في مقرر ST401

وعليه فإن:

$$P(A) = \frac{15}{60} \qquad P(B) = \frac{12}{60} \qquad P(A \cap B) = \frac{6}{60}$$

1- أن يكون الطالب مسجل في أحد المقررين على الأقل:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{15}{60} + \frac{12}{60} - \frac{6}{60} = \frac{21}{60} = 0.35$$

2- أن يكون الطالب غير مسجل في أياً من المقررين:

$$P((A \cup B)') = 1 - P(A \cup B) = 1 - 0.35 = 0.65$$

3- أن يكون الطالب مسجل في مقرر ST305 فقط:

$$P(A \cap B') = P(A) - P(A \cap B) = \frac{15}{60} - \frac{6}{60} = \frac{9}{60} = 0.15$$

مثال14:

بفرض أن 90% من الجرائم التي تحدث في مدينة ما تحدث في منطقة معينة داخل تلك المدينة، وأن 80% من الجرائم التي تحدث في هذه المدينة تحدث ليلاً، فما احتمال ان تكون الجرائم التي تحدث في تلك المنطقة تحدث ليلاً؟

الحل:

بفرض أن: الحدث A الجرائم التي تحدث ليلاً.

الحدث B الجرائم التي تحدث في المنطقة المحددة.

$$P(A) = 0.80 \qquad P(A') = 0.20 \quad , \quad P(B) = 0.90 \qquad P(B') = 0.10$$

$$P(A \cap B) \geq 1 - P(A') - P(B')$$

$$P(A \cap B) \geq 1 - 0.20 - 0.10$$

$$P(A \cap B) \geq \ 0.70$$

هذا يعني احتمال ان الجريمة تحدث ليلاً وفي المنطقة المحددة يساوي على الأقل 70%.

قوانين الاحتمال:

أولا: قانون الجمع:

يستخدم لتحديد احتمال حدوث حدث واحد من حدثين أو أكثر ويتوقف حساب الاحتمال على طبيعة العلاقات بين الاحداث (متنافية، غير متنافية).

1- الاحداث المتنافية:

لأي حدثين متنافيين A وB فإن احتمال حدوث الحدث A او الحدث B يمكن حسابه من العلاقة التالية:

$$P(A \cup B) = P(A) + P(B)$$

ويمكن تعميم المعادلة السابقة لأي مجموعة من الاحداث المتنافية A وB وC ....

فإن احتمال حدوث أحد هذه الاحداث (A أو B أو C ...) يمكن حسابه من العلاقة التالية:

$$P(A \cup B \cup C \cup \ \dots \ \cup Z) = P(A) + P(B) + P(C) + \ \dots + P(Z)$$

مثال15: وعاء يحتوي على كرات ملونة منها أربع كرات بيضاء، وثلاث كرات صفراء، وكرتان سوداء، وكرة واحدة حمراء، فما هو احتمال:

1) سحب كرة بيضاء أو صفراء؟
2) سحب كرة سوداء أو صفراء؟

الحل:

نفرض أن: الحدث W يمثل سحب كرة بيضاء.

الحدث Y يمثل سحب كرة صفراء.

الحدث B يمثل سحب كرة سوداء.

الحدث R يمثل سحب كرة حمراء.

وبالتالي يكن احتمال سحب كرة من الوعاء كما يلي:

$$P(W) = \frac{4}{10} \qquad P(Y) = \frac{3}{10} \qquad P(B) = \frac{2}{10} \qquad P(R) = \frac{1}{10}$$

وحيث أن هذه الاحداث أحداث مستقلة، فإن:

1) $P(W \cup Y) = P(W) + P(Y) = \frac{4}{10} + \frac{3}{10} = \frac{7}{10}$
2) $P(B \cup Y) = P(B) + P(Y) = \frac{2}{10} + \frac{3}{10} = \frac{5}{10} = \frac{1}{2} = 0.5$

مثال16:

استخدم الجدول التالي في تحديد الاحتمالات الأتية:

الحدث	العمر	التكرار	التكرار النسبي (الاحتمال)
A	17	2	0.02
B	18	35	0.35
C	19	45	0.45
D	20	13	0.13
E	21	5	0.05
		100	1.000

1) احتمال طالب عمره 17 سنة أو طالب عمره 18 سنة؟
2) احتمال طالب عمره 18 سنة أو أكثر؟

الحل:

1) $P(A \cup B) = P(A) + P(B) = 0.02 + 0.35 = 0.37$

2) $P(B \cup C \cup D \cup E) = P(B) + P(C) + P(D) + P(E) = 0.35 + 0.45 + 0.13 + 0.05 = 0.98$

ويمكن يمكن حسابه بطريقة اخرى:

$$P(A') = 1 - P(A) = 1 - 0.02 = 0.98$$

2-الاحداث غير المتنافية:

إذا كان لدينا حدثين غير متنافيين A و B فإن احتمال حدوث أحدهما على الأقل (أي احتمال حدوث الحدث A أو حدوث الحدث B او حدوث كلاهما معاً) سيكون على النحو التالي:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

وإذا كان لدينا ثلاث أحداث غير متنافية A و B و C فإن احتمال حدوث أحدها على الأقل هو:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

مثال17:

إذا أعطيت مشكلة إحصائية لثلاثة طلاب احمد وإبراهيم وعلي، فإذا كان احتمال نجاحهم في حل هذه المشكلة هو $\frac{1}{2}$ و $\frac{3}{4}$ و $\frac{1}{4}$ على التوالي، فما هو احتمال ان المشكلة سوف تحل؟

الحل:

بفرض ان:

الحدث A يمثل نجاح حل المشكلة للطالب احمد.

الحدث B يمثل نجاح حل المشكلة للطالب ابراهيم.

الحدث C يمثل نجاح حل المشكلة للطالب علي.

$$P(A \cup B \cup C) = \frac{1}{2} + \frac{3}{4} + \frac{1}{4} - \left(\frac{1}{2} * \frac{3}{4}\right) - \left(\frac{1}{2} * \frac{1}{4}\right) - \left(\frac{3}{4} * \frac{1}{4}\right) + \left(\frac{1}{2} * \frac{3}{4} * \frac{1}{4}\right) = 0.90625$$

مثال18:

في أحد المناطق الزراعية وجد أن 40% من المزارع تستخدم الري بالتنقيط و10% منها تستخدم البيوت البلاستيكية وهي نشاطات غير متنافية، فإذا اختيرت إحدى المزارع، فما احتمال ان هذه المزرعة تستخدم الري بالتنقيط او البيوت البلاستيكية او كلاهما معاً؟

الحل:

بفرض ان الحدث A هو المزرعة تستخدم الري بالتنقيط، والحدث B هو المزرعة تستخدم البيوت البلاستيكية، فإن:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.1 - (0.4 * 0.1) = 0.46$$

مثال19:

في رميتين لزهرة نرد غير متحيزة، أوجد احتمال ظهور الرقم 4 مرة واحدة على الأقل؟

الحل: بفرض أن: الحدث A هو ظهور الرقم 4 في الرمية الأولى.

الحدث B هو ظهور الرقم 4 في الرمية الثانية.

$$P(A) = \frac{1}{6} \quad , \quad P(B) = \frac{1}{6}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{1}{6} - \left(\frac{1}{6} * \frac{1}{6}\right) = 0.3055$$

ثانيا: قانون الضرب:

يستخدم لتحديد احتمال حدوث حدثين او أكثر معاً، ويتوقف الاحتمال على ما إذا كانت الاحداث مستقلة أو غير مستقلة للأحداث الغير متنافية، أما الاحداث المتنافية فاحتمال حدوثها معا يساوي **صفر**.

1- الاحداث المستقلة:

إذا كانت الاحداث A وB وC و .... Z أحداثاً مستقلة فإن احتمال وقوعها معاً هو حاصل ضرب احتمالاتها، ويكتب رياضياً كالتالي:

$$P(A \cap B \cap C \cap \dots \cap Z) = P(A) * P(B) * P(C) * \dots * P(Z)$$

يقال بأن الحدثين A وB مستقلان إذا تحقق أحد الشروط الأتية:

1- $P(A \cap B) = P(A) * P(B)$
2- $P(A|B) = P(A) \quad , P(B) > 0$
3- $P(B|A) = P(B) \quad , P(A) > 0$

مثال20:

بفرض انه ألقيت قطعة نقود معدنية متزنة مرتين، وعرّفنا الحدث A بأنه حدث الحصول على وجهين متشابهين، والحدث B حدث الحصول على صورة واحدة على الأقل، والحدث C حدث الحصول على صورة في الرمية الأولى، حدد عناصر فراغ العينة وكذلك الأحداث، ثم بين فيما إذا كانت مستقلة ثنائياً؟

الحل:

عدد عناصر فراغ العينة = $4=2^2$ وهي: $\Omega = \{HH, HT, TH, TT\}$

والاحداث: $A = \{HH, TT\}$, $B = \{HH, HT, TH\}$, $C = \{HH, HT\}$

ويمكن حساب احتمال حدوث كل حدث من الأحداث السابقة كما يلي:

$$P(A) = \frac{2}{4} = 0.5 \quad , \quad P(B) = \frac{3}{4} = 0.75 \quad , \quad P(C) = \frac{2}{4} = 0.5$$

التأكد من استقلالية الأحداث:

$$A \cap B = \{HH\} \rightarrow P(A \cap B) = \frac{1}{4} = 0.25 \quad , \quad P(A) * P(B) = 0.5 * 0.75 = 0.375$$

$$P(A \cap B) \neq P(A) * P(B)$$

أي أن A و B حدثان غير مستقلان.

$$A \cap C = \{HH\} \rightarrow P(A \cap C) = \frac{1}{4} = 0.25 \quad , \quad P(A) * P(C) = 0.5 * 0.5 = 0.25$$

$$P(A \cap C) = P(A) * P(C)$$

أي أن A و C حدثان مستقلان.

$$B \cap C = \{HH, HT\} \rightarrow P(B \cap C) = \frac{2}{4} = 0.5 \quad , \quad P(B) * P(C) = 0.75 * 0.5 = 0.375$$

$$P(B \cap C) \neq P(B) * P(C)$$

أي أن B و C حدثان غير مستقلان.

- **تعريف الاستقلالية بصورة عامة:**

يُقال بأن الاحداث $A_1, A_2, A_3, \dots , A_n$ مستقلة عن بعضها البعض إذا تحققت الشروط الاتية:

1- $P(A_i \cap A_j) = P(A_i) * P(A_j) \quad , 1 \leq i \leq j \leq n$

2- $P(A_i \cap A_j \cap A_k) = P(A_i) * P(A_j) * P(A_k) \quad , 1 \leq i \leq j \leq k \leq n$

3- $P(\cap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$

مثال21:

بفرض ان $\Omega = \{s_1, s_2, s_3, s_4\}$، و $P(s_i) = \frac{1}{4}$ حيث i=1,2,3,4، وبفرض ان الحدث A هو $A = \{s_1, s_2\}$ والحدث B هو $B = \{s_1, s_3\}$ والحدث C هو $C = \{s_1, s_4\}$ فإن:

$$A \cap \mathrm{B} = \mathrm{A} \cap \mathrm{C} = B \cap C = \{s_1\} \quad , \quad A \cap \mathrm{B} \cap C = \{s_1\}$$

وعليه فإن:

$$P(A \cap \mathrm{B}) = \mathrm{P}(\mathrm{A} \cap \mathrm{C}) = \mathrm{P}(B \cap C) = P(A \cap \mathrm{B} \cap C) = \frac{1}{4} = 0.25$$

$$P(A \cap \mathrm{B}) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(A) * P(B)$$

$$P(A \cap \mathrm{C}) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(A) * P(C)$$

$$P(B \cap \mathrm{C}) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(B) * P(C)$$

وبذلك تكون الاحداث A و B و C مستقلة ثنائياً.

ولكن: $P(A \cap \mathrm{B} \cap C) = \frac{1}{4} \neq \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = P(A) * P(B) * P(C)$

وبذلك تكون الاحداث A و B و C غير مستقلة مع بعضها البعض.

مثال22:

صندوق يحتوي على كرات ملونة منها أربع كرات بيضاء وثلاث كرات صفراء وكرتان سوداء وكرة واحدة حمراء، فإذا علمت أن **السحب مع الارجاع** احسب الاتي:

1) احتمال سحب كرة بيضاء ثم سحب كرة صفراء؟
2) احتمال سحب كرة صفراء ثم سحب كرة سوداء؟
3) سحب كرة صفراء ثم سحب كرة سوداء ثم سحب كرة حمراء؟
4) سحب كرة بيضاء ثم سحب كرة صفراء ثم سحب كرة سوداء ثم سحب كرة حمراء؟

الحل:

نفرض أن: الحدث W يمثل سحب كرة بيضاء.

الحدث Y يمثل سحب كرة صفراء.

الحدث B يمثل سحب كرة سوداء.

الحدث R يمثل سحب كرة حمراء.

وبالتالي يكن احتمال سحب كرة من الوعاء كما يلي:

$$P(W) = \frac{4}{10} \qquad P(Y) = \frac{3}{10} \qquad P(B) = \frac{2}{10} \qquad P(R) = \frac{1}{10}$$

وبالتالي فإن:

1) $P(W \cap Y) = P(W) * P(Y) = 0.4 * 0.3 = 0.12$
2) $P(Y \cap B) = P(Y) * P(B) = 0.3 * 0.2 = 0.06$
3) $P(Y \cap B \cap R) = P(Y) * P(B) * P(R) = 0.3 * 0.2 * 0.1 = 0.006$
4) $P(W \cap Y \cap B \cap R) = P(W) * P(Y) * P(B) * P(R) = 0.4 * 0.3 * 0.2 * 0.1 = 0.0024$

مثال23:

يبيع محل الازهار عبوات تحمل بذور أزهار، أربع بذور حمراء، وثلاث بذور صفراء، وبذرتين بنفسجيتين، وبذرة واحدة برتقالية، فإذا كانت احتمال انبات كل بذرة 0.6، فأوجد الاحتمالات التالية:

1) احتمال أن تنبت جميع البذور الصفراء؟
2) احتمال عدم انبات أي بذرة حمراء؟
3) احمال أن تنبت جميع البذور ماعدا البذور البرتقالية؟

الحل:

حيث أن احتمال انبات أي بذرة هو $P(x)=0.6$.

1) $P(Y \cap Y \cap Y) = P(x) * P(x) * P(x) = 0.6 * 0.6 * 0.6 = 0.216$
2) $P(x') = 1 - 0.6 = 0.4$

$P(R' \cap R' \cap R') = P(x') * P(x') * P(x') * P(x') = 0.4 * 0.4 * 0.4 * 0.4 = 0.0256$

3) $P(Y \cap Y \cap Y \cap R \cap R \cap R \cap R \cap M \cap M \cap L') = (0.6)^9 * 0.4 = 0.00403$

2- الاحداث غير المستقلة:

إذا كانت الاحداث غير مستقلة (وغير متنافية)، وهي ما يسمى بالاحتمال الشرطي، فإن احتمال وقوع حدثين معاً يكون كالتالي:

$$P(AB) = P(A \cap B) = P(A) * P(B/A)$$

واحتمال حدوث ثلاثة أحداث هو:

$$P(ABC) = P(A \cap B \cap C) = P(A) * P(B/A) * P(C/AB)$$

مثال24:

إذا كان وعاء يحتوي على كرات ملونة، أربع كرات بيضاء، وثلاث كرات صفراء، وكرتان سوداء، ووحدة حمراء، فكم تكون الاحتمالات التالية إذا كان **السحب بدون إرجاع:**

1) سحب كرة بيضاء ثم سحب كرة صفراء؟
2) سحب كرة صفراء ثم سحب كرة سوداء؟
3) سحب كرة صفراء ثم سحب كرة سوداء ثم سحب كرة حمراء؟

الحل:

1) $P(WY) = P(W \cap Y) = P(W) * P(Y/W) = \frac{4}{10} * \frac{3}{9} = 0.1333$
2) $P(YB) = P(Y \cap B) = P(Y) * P(B/Y) = \frac{3}{10} * \frac{2}{9} = 0.06667$
3) $P(YBR) = P(Y \cap B \cap R) = P(Y) * P(B/Y) * P(R/(Y \cap B)) = \frac{3}{10} * \frac{2}{9} * \frac{1}{8} = 0.008333$

- **التباديل Permutations**

التباديل هو عدد الطرق المختلفة التي يمكننا بها اختيار k عنصر من n من العناصر حيث $n \geq k$

مع **مراعاة الترتيب** في كل حالة اختيار.

فإذا كان لدينا عدد n من العناصر فإنه يمكن اختيار عدد k من هذه العناصر (مع مراعاة الترتيب) بعدد معين من الاختيارات يمكن حسابه من العلاقة التالية:

$$_nP_k = \frac{n!}{(n-k)!}$$

حيث ان: الرمز (!) يدل على مضروب العدد.

مثال25:

بكم طريقة يمكن اختيار مدير لإدارة ومساعده، إذا علمت أن عدد الموظفين بتلك الإدارة 12 موظفاً.

الحل:

حيث انه يراعى الترتيب والسحب بدون إرجاع، وعليه باستخدام التباديل نجد ان عدد الطرائق التي يمكن بها اختيارهما هو:

$$_{12}P_2 = \frac{12!}{(12-2)!} = 132$$

نتيجة1:

عدد تباديل n من العناصر المميزة المأخوذة سوياً هو:

$$_nP_n = n(n-1)(n-2)(n-3)\ldots\ (2)(1) = n!$$

مع ملاحظة ان: 0!=1 1!=1 2!=2

وإذا طبقنا المعادلة الأولى سنحصل على نفس النتيجة:

$$_nP_k = \frac{n!}{(n-k)!} \quad \rightarrow \quad _nP_n = \frac{n!}{(n-n)!} = n!$$

مثال26:

بكم طريقة يمكن تنظيم ستة كتب مختلفة في رف مكتبة؟

الحل:

$$_6P_6 = 6! = 720$$

مثال27:

لنفرض أن لدينا مجموعة بها عدد k شخصاً ونود تحديد احتمال ان على الأقل اثنين منهم مولودان في نفس اليوم والشهر، ولكن ليس بالضرورة في نفس السنة.

الحل:

بفرض انه لا توجد علاقة بين تواريخ الميلاد لهذه المجموعة (أي لا يوجد توائم) وان لكل يوم من أيام السنة (على افتراض ان السنة بها 365 يوم) نفس الفرصة بأن يكون تاريخ ميلاد أي شخص في هذه المجموعة، أي انه سيكون هناك 365 تاريخ ميلاد ممكن لكل شخص في هذه المجموعة وعليه فإن فراغ العينة يحتوي على 365^k نتيجة ممكنة جميعها لها نفس الفرصة في الظهور، علاوة على ذلك فإن عدد نتائج فراغ العينة التي تمثل ان جميع تواريخ الميلاد K ستكون مختلفة هو $_{365}P_k$.

بفرض ان $P(A)$ هو احتمال ان كل شخص مولود في يوم مختلف عن الشخص الاخر.

$P(B)$ هو احتمال انه على الأقل اثنين منهم مولودان في نفس اليوم والشهر.

وبالتالي فإن: $P(B) = 1 - P(A)$

$$P(A) = \frac{_{365}P_k}{365^k} \quad \rightarrow \quad P(B) = 1 - \frac{_{365}P_k}{365^k}$$

والجدول التالي يوضح قيما عددية للاحتمال $P(B)$ لقيم مختلفة من k:

K	5	15	25
$P(B)$	0.0271	0.2529	0.5687

- **التوافيق Combinations**

التوافيق هو عدد الطرق المختلفة التي يمكننا بها اختيار عدد k عنصر من n من العناصر حيث $n \geq k$ مع **عدم مراعاة الترتيب** في كل حالة اختيار.

فإذا كان لدينا عدد n من العناصر فإنه يمكن اختيار عدد k من هذه العناصر (مع عدم مراعاة الترتيب) بعدد معين من الاختيارات يمكن حسابه من العلاقة التالية:

$$\binom{n}{k} = C_k^n = \quad {}_nC_k = \frac{n!}{k!(n-k)!}$$

مثال28:

ما هو عدد العينات التي يمكن اختيارها لعدد كرتين من بين أربع كرات، صفراء وحمراء وبيضاء وسوداء، بدون إرجاع وبدون مراعاة الترتيب.

الحل:

$$\binom{4}{2} = C_2^4 = \quad {}_4C_2 = \frac{4!}{2!(4-2)!} = 6$$

نظرية6:

إذا كانت $n_1, n_2, n_3, \dots, n_k$ أعداد صحيحة غير سالبة بحيث أن $\sum_{i=1}^{k} n_i = n$ فإن عدد الطرائق التي يمكن بها تجزئة المجموعة المكونة من n عنصراً الى k مجموعة جزئية بحيث تحتوي المجموعة i على n_i من العناصر المتشابهة لجميع قيم i حيث i=1,2, … , k هو:

$$\binom{n}{n_1,\ n_2,\ n_3,\dots,\ n_k} = \frac{n!}{n_1!n_2!n_3!\dots n_k!}$$

مثال29:

ما هي عدد الطرائق التي يمكن بها تقسيم 15 طالباً بحيث يكون هناك ثلاثة طلاب تقديرهم A وطالبان تقديرهم B وستة طلبة تقديرهم C وأربعة طلبة تقديرهم D؟

الحل:

$$\binom{15}{3,2,6,4} = \frac{15!}{3!2!6!4!} = 6306300$$

مثال30:

في أحد دوريات كرة القدم كان عدد الفرق 24 فريقاً فتم تقسيم هذه الفرق المشاركة الى أربع مجموعات كل مجموعة تحتوي على 6 فرق وذلك بإجراء مباريات الدوري الأول، فكم يكون عدد مباريات الدوري الأول؟

الحل:

عدد المباريات للمجموعة الواحدة: مبارة $15 = \frac{6!}{2!(6-2)!} = {}_6C_2 = \quad C_2^6 = \binom{6}{2}$

عدد المباريات الكلية: مبارة $4 * 15 = 60$

مثال31:

إذا تم إلقاء 12 مكعب نرد متزن معا مرة واحدة، ما احتمال ظهور كل رقم من الأرقام سرتين؟

الحل:

حيث ان: $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 2$ وعدد النقاط بفراغ العينة هو 6^{12} وان عدد الطرائق التي يمكن ان تظهر فيها أي رقم مرتين من هو $\frac{12!}{(2!)^6}$ وعليه فإن الاحتمال المطلوب هو:

$$P(\ ) = \frac{\frac{12!}{(2!)^6}}{6^{12}} = 0.003438$$

مثال32: كيس يحتوي على أربع كرات بيضاء وكرتان سوداء، وكيس آخر يحتوي على ثلاث كرات بيضاء وخمس كرات سوداء، فإذا سحبت كرة من كل كيس، فأوجد اجتمال:

1) كلتا الكرتين لونهما أسود؟ 2) كلتا الكرتين لونهما أبيض؟ 3) كرة بيضاء وكرة سوداء (بغض النظر عن الأكياس)؟

الحل:

1) $P(B_1 \cap B_2) = \frac{2}{6} * \frac{5}{8} = \frac{5}{24}$
2) $P(W_1 \cap W_2) = \frac{4}{6} * \frac{3}{8} = \frac{1}{4} = 0.25$
3) $P(B_1 \cap W_2) + P(B_2 \cap W_1) = \left(\frac{2}{6} * \frac{3}{8}\right) + \left(\frac{4}{6} * \frac{5}{8}\right) = \frac{13}{24}$

مثال33:

صندوق يحتوي على ثماني كرات حمراء وثلاثة كرات بيضاء وستة كرات زرقاء، فإذا سحبت ثلاث كرات عشوائياً بدون ارجاع، فأوجد الاحتمالات التالية:

1- الكرات الثلاث حمراء؟

$$P(R \cap R \cap R) = \frac{8}{17} * \frac{7}{16} * \frac{6}{15} = \mathbf{0.08235}$$

2- الكرات الثلاث بيضاء؟

$$P(W \cap W \cap W) = \frac{3}{17} * \frac{2}{16} * \frac{1}{15} = \mathbf{0.001471}$$

3- كرتان حمراء وواحدة بيضاء؟

$$P(R \cap R \cap W) = \left(\frac{8}{17} * \frac{7}{16} * \frac{3}{15}\right) * 3 = \mathbf{0.1474}$$

4- على الأقل كرة بيضاء؟

$$P(X \geq 1) = 1 - P(X = 0)$$

$$= 1 - [P(R \cap R \cap R) + P(B \cap B \cap B) + P(R \cap R \cap B) * 3 + P(B \cap B \cap R) * 3] = \frac{\mathbf{79}}{\mathbf{170}}$$

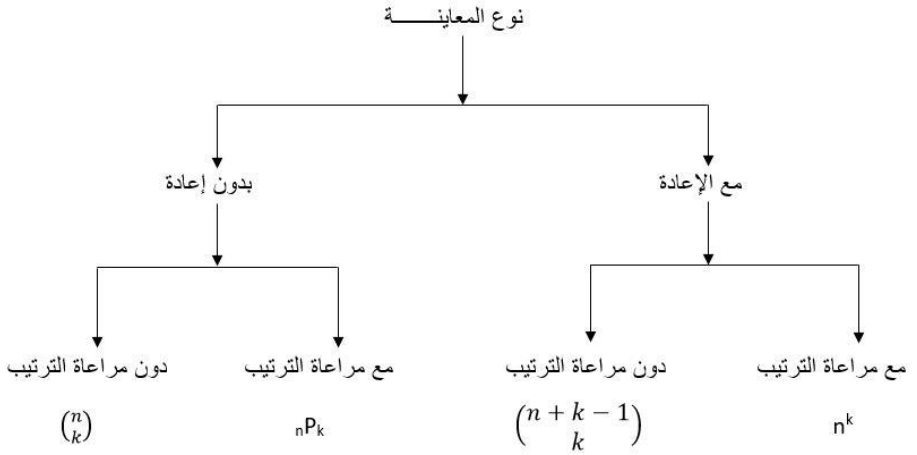
5- كرة من كل لون؟

$$P(R \cap W \cap B) * 6 = \left(\frac{8}{17} * \frac{3}{16} * \frac{6}{15}\right) * 6 = \frac{\mathbf{18}}{\mathbf{85}}$$

6- الكرات سحبت بالترتيب (حمراء، ثم بيضاء، ثم زرقاء)؟

$$P(R \cap W \cap B) = \left(\frac{8}{17} * \frac{3}{16} * \frac{6}{15}\right) = \frac{\mathbf{3}}{\mathbf{85}}$$

ينمكن تلخيص طرق عد عناصر فراغ العينة كما يلي:



الاحتمال الشرطي Conditional Probability

الاحتمال الشرطي هو احتمال حدوث حدث معين A بمعلومية حدوث حدث آخر B.

إذا كان A وB حدثين فإن الاحتمال الشرطي للحدث A إذا علم حدوث الحدث B هو:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad ,P(B) > 0$$

وكذلك يكون الاحتمال الشرطي للحدث B إذا علم حدوث الحدث A هو:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad ,P(A) > 0$$

حيث ان:

$P(A|B)$: هو احتمال حدوث الحدث A بمعلومية حدوث الحدث B.

$P(B|A)$: هو احتمال حدوث الحدث B بمعلومية حدوث الحدث A.

تبرير مسلمات الاحتمال:

1- من تعريف الاحتمال الشرطي انجد ان:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad ,P(B) > 0$$

$$0 \leq P(A \cap B) \leq P(B)$$

$$0 \leq P(A|B) \leq 1$$

2-

$$P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

مثال34:

إذا تم ألقاء مكعب نرد منتظم مرتين ولوحظ أن مجموع الرقمين T (المجموع) عدد فردي، ما احتمال ان تكون T أقل من 8؟

الحل:

لنفرض ان الحدث A هو أن يكون $(T < 8)$ والحدث B هو أن يكون T عدد فردي، وبالتالي نجد أن:

نكون جدول للنتائج المحتملة في التجربة:

نتيجة الرمية الأولى / نتيجة الرمية الثانية	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$P(A \cap B) = \frac{12}{36} = \frac{1}{3}$$

$$P(B) = \frac{18}{36} = \frac{1}{2}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

ملاحظة:

- إذا توفرت معلومات عن وقوع الحدث A فهذا لا يعني بالضرورة أن فرصة حدوث الحدث B ستكون كبيرة، وفي الواقع إذا كان $(A \cap B = \emptyset)$ فإن:

$$P(B|A) = 0 \leq \mathrm{P(B)} \qquad ,P(A) \geq 0$$

- إذا كانت $B \subset A$ فإن:

$$P(B|A) = \frac{P(B)}{P(A)} \geq P(B)$$

- إذا كانت $B \supset A$ فإن:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

نتيجة2:

$$P(\emptyset|B) = \frac{P(\emptyset \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = 0$$

نتيجة3:

إذا كان A وB حدثين فإن:

$$P(A'|B) = 1 - P(A|B)$$

نتيجة4:

إذا كان A_1, A_2 حدثين بحيث $A_1 \subset A_2$ فإن:

$$P(A_1) \leq P(A_2)$$

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} \quad \leq \quad \frac{P(A_2 \cap B)}{P(B)} = P(A_2|B)$$

مثال35:

في اختبار نهاية الفصل الدراسي وجد ان 40% من الطلبة نجحوا في مادة الإحصاء و25% نجحوا في مادة الحاسوب و20% نجحوا في مادتي الإحصاء والحاسوب، فإذا تم اختيار أحد الطلاب عشوائياً وكان الحدث A يمثل نجاح الطالب في الإحصاء، والحدث B يمثل نجاحه في الحاسوب، فأوجد احتمال حدوث الاحداث التالية:

1) نجاح الطالب في الحاسوب إذا علمنا انه نجح في الاحصاء؟
2) نجاح الطالب في الإحصاء إذا علمنا انه نجح في الحاسوب؟
3) نجاح الطالب في الإحصاء إذا علمنا رسوبه في الحاسوب؟
4) رسوب الطالب في الحاسوب بشرط رسوبه في الإحصاء؟

الحل:

حيث ان: $P(A) = 0.40 \ , \ P(B) = 0.25 \ , \ P(A \cap B) = 0.20$

وبالتالي فإن:

1) $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.20}{0.40} = 0.50$
2) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.20}{0.25} = 0.80$
3) $P(A|B') = \frac{P(A \cap B')}{P(B')} = \frac{P(A) - P(A \cap B)}{1 - P(B)} = \frac{0.40 - 0.20}{1 - 0.25} = \frac{0.20}{0.75}$
4) $P(B'|A') = \frac{P(A' \cap B')}{P(A')} = \frac{P((A \cup B)')}{P(A')} = \frac{1 - P(A \cup B)}{1 - P(A)} = \frac{1 - [P(A) + P(B) - P(A \cap B)]}{1 - P(A)}$
$= \frac{1 - (0.4 + 0.25 - 0.2)}{1 - 0.40} = 0.9167$

قاعد الضرب للاحتمال الشرطي:

من خلال تعريف الاحتمال الشرطي نجد أن:

$$P(A \cap B) = \begin{cases} P(A) * P(B|A) & , P(A) > 0 \\ P(B) * P(A|B) & , P(B) > 0 \end{cases}$$

مثال36:

يحتوي صندوق على عشرة مصابيح كهربائية من بينها أربع مصابيح معيبة (غير صالحة)، فإذا سحب مصباحان عشوائياً الواحد تلو الأخر وبدون إعادة، فاحسب احتمال حدوث الأحداث التالية:

1) أن يكون المصباحان معيبان.
2) أن يكون المصباحان صالحان.
3) أن يكون الأول صالحاً والثاني معيباً.
4) أن يكون أحدهما على الأقل صالح.

الحل:

نفرض أن: الحدث A المصباح الأول معيب.

الحدث B المصباح الثاني معيب.

1) $P(A \cap B) = P(A) * P(B/A) = \frac{4}{10} * \frac{3}{9} = \mathbf{\frac{12}{90}}$
2) $P(A' \cap B') = P(A') * P(B'/A') = \frac{6}{10} * \frac{5}{9} = \frac{30}{90} = \mathbf{\frac{1}{3}}$

$$P(A' \cap B') = P(A \cup B)' = 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(A' \cap B')]$$
$$= 1 - \left[\frac{4}{10} + \frac{4}{10} - \frac{12}{90}\right] = \mathbf{\frac{1}{3}}$$

3) $P(A' \cap B) = P(C) * P(B/C) = \frac{6}{10} * \frac{4}{9} = \mathbf{\frac{24}{90}}$

$$= P(B) - P(A \cap B) = \frac{4}{10} - \frac{12}{90} = \mathbf{\frac{24}{90}}$$

4) $P(A' \cup B') = P(A \cap B)' = 1 - P(A \cap B) = 1 - \frac{12}{90} = \mathbf{\frac{78}{90}}$

$$= P(A') + P(B') - P(A' \cap B') = \frac{6}{10} + \frac{6}{10} - \frac{1}{3} = \mathbf{\frac{78}{90}}$$

ويمكن حساب الاحتمالات السابقة بطريقة أخرى:

1) $P(A \cap B) = \frac{{}_4C_2}{{}_{10}C_2} = \frac{6}{45} = \mathbf{\frac{12}{90}}$
2) $P(A' \cap B') = \frac{{}_6C_2}{{}_{10}C_2} = \frac{15}{45} = \frac{30}{90} = \frac{1}{3}$
3) $P(A' \cap B) = \frac{{}_6P_1 * {}_4P_1}{{}_{10}P_2} = \frac{24}{90}$

نتيجة5:

إذا كانت $A_1,\ A_2,\ A_3, \dots\ ,\ A_n$ تمثل احداثاً بحيث أن $P\left(\cap_{i=1}^{k} A_i\right) > 0$ و

$k = 1,2,3, \dots, (n-1)$، فإن:

$$P\left(\cap_{i=1}^{n} A_i\right) = P(A_1) * P(A_2|A_1) * P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

مثال37:

سحبت أربع كرات وبدون إعادة من صندوق به r كرة حمراء و b كرة زرقاء، قما احتمال الحصول على متوالية من النتائج التالية:

الكرة الأولى حمراء والثانية زرقاء والثالثة حمراء والرابعة زرقاء؟

الحل:

بفرض ان R_i هو حدث الحصول على كرة حمرة في السحبة (i=1, 3)، وان B_j هو حدث الحصول على كرة زرقاء في السحبة (j=2, 4)، وبالتالي يمكن حساب الاحتمال المطلوب:

$$P(R_1 \cap B_2 \cap R_3 \cap B_4) = P(R_1) * P(B_2/R_1) * P(R_3/R_1 \cap B_2) * P(B_4/R_1 \cap B_2 \cap R_3)$$

$$= \frac{r}{r+b} * \frac{b}{r+b-1} * \frac{r-1}{r+b-2} * \frac{b-1}{r+b-3}$$

مثال38:

يحتوي صندوق على 8 كرات حمراء و6 كرات بيضاء، فإذا تم اختيار ثلاث كرات بطريقة عشوائية، ما احتمال أن تكون الكرات جميعها حمراء؟

الحل: الطريقة الأولى:

يتم سحب الكرات الواحدة وراء الأخرى، وبفراض ان الحدث A يمثل ان الكرة الأولى حمراء، والحدث B يمثل ان الكرة الثانية حمراء، والحدث C يمثل ان الكرة الثالثة حمراء، فإن الاحتمال المطلوب هو:

$$P(A \cap B \cap C) = P(A) * P(B/A) * P(C/(A \cap B)) = \frac{8}{14} * \frac{7}{13} * \frac{6}{12} = \mathbf{\frac{2}{13}}$$

الطريقة الثانية: يتم سحب الكرات الثلاثة مع بعض مرة واحدة، فإن الاحتمال المطلوب هو:

$$P = \frac{\binom{8}{3}}{\binom{14}{3}} = \frac{56}{364} = \mathbf{\frac{2}{13}}$$

تجزئة فراغ العينة: Partition of Sample Space

يقال لمجموعة من الأحداث $A_1, A_2, A_3, \dots, A_n$ انها تجزئة لفراغ العينة Ω إذا تحقق كلا الشرطين التاليين:

1- $\cup_{i=1}^{n} A_i = \Omega$

2- $A_i \cap A_j = \emptyset \quad , i = 1,2,3, \dots \quad j = 1,2,3, \dots \quad i \neq j$

مثال39: أجريت تجربة عشوائية برمي زهرة نرد مرة واحدة، عندئذ يكون: $\Omega = \{1,2,3,4,5,6\}$

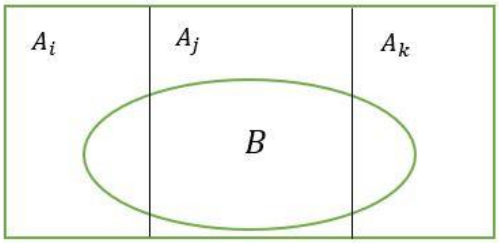
فإن المجموعة {{1},{2,3},{4,5,6}} تمثل تجزئة لفراغ العينة وذلك لان:

$$\cup_{i=1}^{n} A_i = \{1\} \cup \{2,3\} \cup \{4,5,6\} = \{1,2,3,4,5,6\} = \Omega$$

$$\{1\} \cap \{2,3\} = \emptyset \quad , \{1\} \cap \{4,5,6\} = \emptyset \quad , \quad \{2,3\} \cap \{4,5,6\} = \emptyset$$

نظرية الاحتمال الكلية:

إذا كانت الأحداث A_n , ... $A_1,\ A_2,\ A_3,$ تجزئة لفراغ العينة Ω وكان $P(A_i)$ لجميع قيم i وإن B حدثاً ينتمي لنفس فراغ العينة فإن:



$$P(B) = \sum_{i=1}^{\infty} P(A_i) * P(B|A_i)$$

مثال40:

إذا علمت بأنه في أحد المصانع ثلاث خطوط للإنتاج ينتج الخط الأول 50% من إنتاج المصنع، وينتج الخط الثاني 30% من الإنتاج والباقي يقوم بإنتاجه الخط الثالث، ما احتمال إنتاج وحدة معيبة (defective) في المصنع ككل؟ علماً بأن نسب الإنتاج المعيب في الخطوط الثلاثة هي 2% و3% و5% على الترتيب.

الحل:

بفرض ان: A_i: حدث أن الإنتاج كان من قبل الخط i، حيث ان i= 1، 2، 3.

D: حدث إنتاج وحدة معيبة.

وعليه فإن:

$$P(A_1) = 0.50 \ , \ P(A_2) = 0.30 \ , \ P(A_3) = 0.20$$

وحيث ان:

$$P(D/A_1) = 0.02 \ , \ P(D/A_2) = 0.03 \ , \quad P(D/A_3) = 0.05$$

فإن الاحتمال المطلوب هو:

$$P(D) = \sum_{i=1}^{3} P(D|A_i)\ P(A_i) = P(D|A_1)\ P(A_1) + P(D|A_2)\ P(A_2) + P(D|A_3)\ P(A_3)$$
$$= 0.02 * 0.50 + 0.03 * 0.30 + 0.05 * 0.20 = 0.029$$

نتيجة6:

لأي حدثين A و B فإن:

$$P(A) = P(A|B) * P(B) + P(A|B') * \ P(B')$$

مثال41:

يحتوي الصندوق "أ" على كرتين بيضاء وأربع كرات حمراء، بينما يحتوي الصندوق "ب" على كرة واحدة بيضاء وأخرى حمراء، فإذا سحبت كرة من الصندوق "أ" ووضعت في الصندوق "ب"، ثم سحبت كرة من الصندوق "ب"، فما احتمال أن تكون الكروة المسحوبة من الصندوق "ب" بيضاء؟

الحل:

بفرض ان: الحدث B يمثل الكرة المنقولة من الصندوق الأول الى الصندوق الثاني بيضاء.

والحدث 'B يمثل الكرة المنقولة من الصندوق الأول الى الصندوق الثاني حمراء.

والحدث A يمثل الكرة المسحوبة من الصندوق الثاني بيضاء.

وبالتالي يمكن حساب الاحتمال المطلوب كما يلي:

$$P(A) = P(A|B) * P(B) + P(A|B') * P(B') = \frac{2}{3} * \frac{2}{6} + \frac{1}{3} * \frac{4}{6} = \mathbf{\frac{4}{9}}$$

نظرية بييز: Baye's Theorem

إذا كانت الأحداث $A_1,\ A_2,\ A_3, \dots\ ,\ A_n$ تجزئة لفراغ العينة Ω بحيث $P(A_i) > 0$ لجميع قيم i وكان B أي حدث بشرط أن $P(B) > 0$ فإن:

$$P(A_i|B) = \frac{P(A_i) * P(B|A_i)}{\sum_{j=1}^{n} P(A_j) * P(B|A_i)} \qquad ,i = 1,2,3, \dots\ ,n$$

مثال42:

يحتوي صندوق على ثلاث بطاقات متشابهة، الأولى لونها أحمر من الجهتين، والثانية لونها أبيض من الجهتين، والأخيرة لونها أحمر من جهة وأبيض من الجهة الأخرى، وبفرض انه تم اختيار بطاقة بطريقة عشوائية من ذلك الصندوق، وكانت الجهة العلوية لها حمراء اللون، ما احتمال ان تكون الجهة الأخرى لونها أبيض؟

الحل:

بفرض أن: الحدث RR يمثل البطاقة التي لونها أحمر من الجهتين.

والحدث WW يمثل البطاقة التي لونها أبيض من الجهتين.

والحدث RW يمثل البطاقة التي لونها أحمر من جهة وأبيض من الجهة الأخرى.

والحدث R يمثل ان البطاقة التي تم اختيارها كانت جهتها العلوية حمراء.

وبالتالي يكون الاحتمال المطلوب هو:

$$P(\mathrm{RW}|R) = \frac{P(\mathrm{RW}) * P(R|\mathrm{RW})}{P(RR) * P(R/RR) + P(RW) * P(R/RW) + P(WW) * P(R/WW)} = \frac{\frac{1}{2} * (\frac{1}{3})}{1 * (\frac{1}{3}) + \frac{1}{2} * (\frac{1}{3}) + 0 * (\frac{1}{3})} = \frac{1}{3}$$

التوقع الرياضي:

إذا كان B تمثل احتمال حصول شخص على كمية من النقود، فإن التوقع الرياضي أو ما يسمى التوقع يمكن توضيحه في المثال التالي:

مثال43:

إذا كان احتمال ان يكسب شخص جائزة قيمتها 10,000 دولار هو $\frac{1}{5}$، احسب التوقع الرياضي للمبلغ الذي سيحصل عليه؟

الحل:

يمكن حساب التوقع الرياضي كما يلي:

$$\text{التوقع الرياضي} = 10{,}000 * \frac{1}{5} = 2000$$

مثال44:

في تجارة معينة تتضمن مخاطرة يمكن أن يكسب شخص 300 دينار باحتمال 60% أو يتكبد خسارة قدرها 100 دينار باحتمال 40% حدد القيمة المتوقعة بالنسبة له؟

الحل:

$$\text{التوقع الرياضي} = 300 * 0.6 - 100 * 0.4 = 140$$

الفصل الخامس:

التوزيعات الاحتمالية المنفصلة (المتقطعة)

المتغير العشوائي:

في كثير من الأحيان عند اجراء التجارب العملية تنتهي كل منها بالحصول على نتائج، ولكن من طبيعة الأشياء انه لا يمكن التنبؤ بهذه النتائج بشكل مؤكد قبل اجراء التجربة، وعندما تتباين النتائج او القيم رقمياً لمتغير من مشاهدة الى أخرى نتيجة عوامل ترتبط بالصدفة فإن المتغير يسمى متغير عشوائي.

فمثلاً إذا أُلقيت قطعة نقود معدنية متزنة فإن فراغ العينة للنتائج الممكنة هو $\Omega = \{H, T\}$، وهذه القيمة غير رقمية، والمتغير العشوائي في هذه الحالة إما عدد الصور $(0,1)$ او عدد الكتابة $(0,1)$ وهذه تعتبر قيم عددية، وبالمثل إذا أُلقيت قطعتي نقود فإن فراغ العينة يكون $\Omega = \{HH, HT, TH, TT\}$ وهذه القيم غير رقمية، والمتغير العشوائي في هذه الحالة يمثل عدد مرات ظهور الصورة $(0,1,2)$ أو عدد مرات ظهور الكتابة $(0,1,2)$.

وإذا ألقينا ثلاثة قطع نقود معدنية متزنة فإن فراغ العينة للنتائج الممكنة هو:

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

واحتمال حدوث أي حدث أولي في فراغ العينة السابقة هو $\frac{1}{8}$، فإذا كان اهتمامنا يتركز على عدد مرات ظهور الصور (x) فإن الفراغ العيني x يصبح $\{0,1,2,3\}$ ويمكننا القول بأن:

- عندما $x=3$ يكون الحدث {HHH} واحتماله $\frac{1}{8}$.
- عندما $x=2$ يكون الحدث {HHT,HTH,THH} واحتمال كل منها $\frac{1}{8}$ وبالتالي يكون احتماله $\frac{3}{8}$.
- عندما $x=1$ يكون الحدث {HTT,THT,TTH} واحتمال كل منها $\frac{1}{8}$ وبالتالي يكون احتماله $\frac{3}{8}$.
- عندما $x=0$ يكون الحدث {TTT} واحتماله $\frac{1}{8}$.

وبذلك يمكن القول ان المتغير x هو متغير بأعداد حقيقية تقترن قيمته بالاحتمالات المقابلة وهو متغير عشوائي.

التوزيعات الاحتمالية المنفصلة (المتقطعة):

إذا كان المتغير x يمكن ان يأخذ مجموعة من القيم المتقطعة $x_1, x_2, x_3, \dots, x_k$ باحتمالات $P_1, P_2, P_3, \dots, P_k$ على الترتيب، حيث أن $P_1 + P_2 + P_3 + \dots + P_k = 1$ فإنه يمكن القول إن هذا التوزيع يعد تعريفاً للتوزيع الاحتمالي المنفصل (متقطع) للمتغير x.

الدلة $P(x)$ والتي تأخذ القيم $P_1, P_2, P_3, \dots, P_k$ لقيم $x_1, x_2, x_3, \dots, x_k$ تسمى دالة الاحتمال (دالة الكتلة) (mass function).

فمثلاً إذا ألقينا زهرتي نرد متزنتين (غير متحيزتين) وافترضنا أن x تعبر عن مجموع النقاط التي نتحصل عليها فإن التوزيع الاحتمالي سيكون كما هو موضح بالجدول:

x	2	3	4	5	6	7	8	9	10	11	12
P(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

وبتجميع الاحتمالات نتحصل على دالة التوزيع الاحتمالي التراكمي والمقابلة للتوزيع التكراري المتجمع النسبي، والدالة المرتبطة بهذا التوزيع تسمى أحياناً دالة التوزيع التراكمي (C.D.F) (Cumulative Distribution Function) ويرمز لها بالرمز F(x)، ويمكن تكوين دالة التوزيع التراكمي للمثال السابق كما يلي:

x	≤ 2	≤ 3	≤ 4	≤ 5	≤ 6	≤ 7	≤ 8	≤ 9	≤ 10	≤ 11	≤ 12
F(x)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}=1$

مثال1:

إذا كان المطلوب اختيار طالبين بطريقة عشوائية من بين 3 طلاب و3 طالبات، وكان المتغير العشوائي X يمثل عدد الطالبات اللواتي سيتم اختيارهن، فأوجد دالة التوزيع الاحتمالي لهذا المتغير؟

الحل:

باستخدام التوافيق (نستخدم التوافيق لان الترتيب غير مهم في هذه الحالة) يمكن حساب عدد الطرائق التي يمكن بها اختيار طالبين من بين 6 طلاب كما يلي:

$$\binom{6}{2} = \quad {}_6C_2 = \frac{6!}{2!\,(6-2)!} = 15$$

أي ان فضاء العينة يتكون من 15 نقطة ولكل عنصر نفس الفرصة في الظهور وذلك لان المعاينة عشوائية، وإن القيم الممكنة للمتغير العشوائي X هي $(0،1،2)$، ويمكن حساب احتمال حدوث كل قيمة من قيم المتغير العشوائي X كما يلي:

- X=0 ➔ $P_X(0) = P(X=0) = \frac{\binom{3}{0}\binom{3}{2}}{\binom{6}{2}} = \frac{3}{15} = \frac{1}{5} = 0.20$
- X=1 ➔ $P_X(1) = P(X=1) = \frac{\binom{3}{1}\binom{3}{1}}{\binom{6}{2}} = \frac{9}{15} = \frac{3}{5} = 0.60$
- X=2 ➔ $P_X(2) = P(X=2) = \frac{\binom{3}{2}\binom{3}{0}}{\binom{6}{2}} = \frac{3}{15} = \frac{1}{5} = 0.20$

ويمكن كتابة التوزيع الاحتمالي المنفصل في صورة رياضية تسمى **دالة كتلة الاحتمال** كما يلي:

$$P_X(X) = P(X=x) = \begin{cases} \frac{\binom{3}{x}\binom{3}{2-x}}{\binom{6}{2}} & , \ x = 0,1,2 \\ 0 & , \quad ow \end{cases}$$

توزيع ذي الحدين:

إذا كان p احتمال وقوع حدث ما في محاولة وحيدة ويسمى نجاحه، وq هو احتمال عدم وقوع الحدث في أي محاولة ويسمى احتمال الفشل حيث أن (q=1−p) فإن احتمال وقوع الحدث مرات عددها x بالضبط في N محاولة أي حدوث x مرة من النجاح و (N−x) مرة من الفشل يعطى بالمعادلة التالية:

$$P(x) = \quad {}_N C_x p^x q^{N-x} = \frac{\mathrm{N!}}{x!\,(N-x)!}\, p^x q^{N-x}$$

حيث أن: $x = 0,1,2,3, \dots \ , \ N$

مثال2:

أوجد احتمال الحصول على صورتين بالضبط في 6 رميات لعملة غير متحيزة، ثم أوجد احتمال الحصول على أربع صور؟

الحل:

$$N = 6 \qquad , \qquad p = q = \frac{1}{2}$$

$$x = 2 \quad \rightarrow \quad P(2) = \frac{6!}{2!\,(6-2)!} * 0.5^2 * 0.5^4 = 0.234375$$

$$x = 4 \quad \rightarrow \quad P(2) = \frac{6!}{4!\,(6-4)!} * 0.5^4 * 0.5^2 = 0.234375$$

بعض خصائص توزيع ذي الحدين:

المقياس	المعادلة
الوسط $\bar{x}$	N*p
التباين	N*p*q
الانحراف المعياري	$\sqrt{N * p * q}$
معامل الالتواء باستخدام العزوم α_1	$\frac{q-p}{\sqrt{N*p*q}}$
معامل التفرطح باستخدام العزم γ	$3 + \frac{1-6p*q}{N*p*q}$

مثال3:

في مائة رمية لعملة معدنية غير متحيزة، أوجد متوسط ظهور الصورة والانحراف المعياري ومعامل الالتواء ومعامل التفرطح لهذا التوزيع؟

الحل:

$$N = 100 \quad , \quad p = q = \frac{1}{2}$$

$$\bar{x} = N * p = 100 * \frac{1}{2} = 50$$

$$S.D = \sqrt{N * p * q} = \sqrt{100 * \frac{1}{2} * \frac{1}{2}} = 5$$

$$\alpha_1 = \frac{q - p}{\sqrt{N*p*q}} = \frac{\frac{1}{2} - \frac{1}{2}}{\sqrt{100*\frac{1}{2}*\frac{1}{2}}} = 0$$

$$\gamma = 3 + \frac{1-6p*q}{N*p*q} = 3 + \frac{1-6*\frac{1}{2}*\frac{1}{2}}{100*\frac{1}{2}*\frac{1}{2}} = 2.98$$

مثال4:

عند رمي عملة معدنية متوازنة 3 مرات، أوجد احتمال ظهور الاتي:

1) 3 صور. 2) صورتين وكتابة. 3) كتابتين وصورة. 4) ثلاث كتابات.

الحل:

نفرض ان المتغير العشوائي x هو ظهور صورة.

$$x = 0,1,2,3 \qquad N = 3 \qquad p = q = \frac{1}{2}$$

1) $P(x = 3) = \quad {}_3C_3 * (\frac{1}{2})^3 * (\frac{1}{2})^0 = 0.125$
2) $P(x = 2) = \quad {}_3C_2 * (\frac{1}{2})^2 * (\frac{1}{2})^1 = 0.375$
3) $P(x = 1) = \quad {}_3C_1 * (\frac{1}{2})^1 * (\frac{1}{2})^2 = 0.375$
4) $P(x = 0) = \quad {}_3C_0 * (\frac{1}{2})^0 * (\frac{1}{2})^3 = 0.125$

جدول (3) : التوزيع الاحتمالي لتوزيع بواسون لقيم مختلفة بالنسبة الى λ ، حيث

$$p_X(x) = P(X = x) = \frac{e^{-\lambda}\,\lambda^x}{x!} \quad , x = 0,1,2,3,\cdots$$

	λ									
x	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.0905	.1637	.2222	.2681	.3033	.3293	.3476	.3595	.3659	.3679
2	.0045	.0164	.0333	.0536	.0758	.0988	.1217	.1438	.1647	.1839
3	.0002	.0011	.0033	.0072	.0126	.0198	.0284	.0383	.0494	.0613
4	.0000	.0001	.0002	.0007	.0016	.0030	.0050	.0077	.0111	.0153
5	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0012	.0020	.0031
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

	λ									
x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.3662	.3614	.3543	.3452	.3347	.3230	.3106	.2975	.2842	.2707
2	.2014	.2169	.2303	.2417	.2510	.2584	.2640	.2678	.2700	.2707
3	.0738	.0867	.0998	.1128	.1255	.1378	.1496	.1607	.1710	.1804
4	.0203	.0260	.0324	.0395	.0471	.0551	.0636	.0723	.0812	.0902
5	.0045	.0062	.0084	.0111	.0141	.0176	.0216	.0260	.0309	.0361
6	.0008	.0012	.0018	.0026	.0035	.0047	.0061	.0078	.0098	.0120
7	.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0020	.0027	.0034
8	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0005	.0006	.0009
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

	λ									
x	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

	λ									
x	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	.0450	.0408	.0369	.0334	.0302	.0273	.0247	.0224	.0202	.0183
1	.1397	.1304	.1217	.1135	.1057	.0984	.0915	.0850	.0789	.0733
2	.2165	.2087	.2008	.1929	.1850	.1771	.1692	.1615	.1539	.1465
3	.2237	.2226	.2209	.2186	.2158	.2125	.2087	.2046	.2001	.1954
4	.1734	.1781	.1823	.1858	.1888	.1912	.1931	.1944	.1951	.1954
5	.1075	.1140	.1203	.1264	.1322	.1377	.1429	.1477	.1522	.1563
6	.0555	.0608	.0662	.0716	.0771	.0826	.0881	.0936	.0989	.1042
7	.0246	.0278	.0312	.0348	.0385	.0425	.0466	.0508	.0551	.0595
8	.0095	.0111	.0129	.0148	.0169	.0191	.0215	.0241	.0269	.0298
9	.0033	.0040	.0047	.0056	.0066	.0076	.0089	.0102	.0116	.0132
10	.0010	.0013	.0016	.0019	.0023	.0028	.0033	.0039	.0045	.0053
11	.0003	.0004	.0005	.0006	.0007	.0009	.0011	.0013	.0016	.0019
12	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006
13	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002
										.0001

المصدر :

Handbook of Probability and Statistics With Tables 2nd ed. McGraw Hill Inc.1970.

توزيع بواسون:

لأي توزيع احتمالي متقطع (x=0, 1, 2, 3, …) فإنه يمكن حساب احتمال حدوث x كما يلي:

$$P(x) = \frac{\lambda^x * e^{-\lambda}}{x!} \quad , \quad e = 2.71828$$

حيث ان: λ: ثابت معطى.

خصائص توزيع بواسون:

المقياس	المعادلة
الوسط $\bar{x}$	λ
التباين	λ
الانحراف المعياري	$\sqrt{\lambda}$
معامل الالتواء باستخدام العزوم α_1	$\frac{1}{\sqrt{\lambda}}$
معامل التفرطح باستخدام العزم γ	$3 + \frac{1}{\sqrt{\lambda}}$

مثال5:

إذا علمت أن عدد حوادث السيارات التي تحدث في الأسبوع بمدينة معينة تتبع توزيع بواسون بمتوسط يساوي 0.70، فما احتمال وقوع ثلاثة حوادث على الأقل خلال الأسبوع؟

الحل:

بفرض ان المتغير العشوائي X يمثل عدد حوادث السيارات التي تحدث في الأسبوع، وان $\lambda = 0.70$.

$$P(X \geq 3) = \sum_{X=3}^{\infty} \frac{0.7^X * e^{-0.7}}{x!}$$

$$= 1 - P(X < 3) = 1 - P(X \leq 2)$$

$$= 1 - \sum_{X=0}^{2} \frac{0.7^X * e^{-0.7}}{x!}$$

$$= 1 - [0.4966 + 0.3476 + 0.1217] = 0.0341$$

مثال6:

إذا علمت بان معدل العمليات الجراحية بأحد الأقسام بمركز طرابلس الطبي هو 3 عمليات في اليوم، احسب احتمال وقوع الاحداث التالية:

أ) عدم اجراء أي عملية جراحية في يوم معين؟

ب) اجراء عملية جراحية واحدة على الأقل في يوم معين؟

ت) اجراء عمليتين جراحيتين على الأكثر في يوم معين؟

ث) اجراء من 4 الى 6 عمليات جراحية خلال يومين؟

الحل:

بفرض ان x هو عدد العمليات الجراحية التي تجرى في اليوم.

$$\lambda = 3 \quad , \quad x = (0,1,2,3, \dots)$$

1) $P(X = 0) = \frac{3^0 * e^{-3}}{0!} = 0.04979$

2) $P(X \geq 1) = 1 - P(X = 0) = 1 - 0.04979 = 0.9502$

3) $P(x \leq 2) = \sum_{x=0}^{2} \frac{3^x * e^{-3}}{x!} = [0.04979 + 0.1494 + 0.2240] = 0.4232$

4) حيث ان العمليات خلال يومين، فإن المتوسط في هذه الحالة ($\lambda = 2 * 3 = 6$)

$$P(4 \leq x \leq 6) = \sum_{X=4}^{6} \frac{6^x * e^{-6}}{x!} = [0.1339 + 0.1606 + 0.1606] = 0.4551$$

تقريب توزيع ذي الحدين بتوزيع بواسون:

يمكن تقريب توزيع بواسون من توزيع ذي الحدين وذلك كحالة تقريبية تحت شروط معينة.

يعتبر توزيع بواسون حالة خاصة من توزيع ذي الحدين عندما يكون احتمال النجاح صغير جداً $(P(x) \to 0)$، ويكون عدد المحاولات n كبير جداً $(n \to \infty)$ بحيث تبقى $(\lambda = n * P(x))$.

فإذا كان المتغير العشوائي X يتوزع وفق توزيع ذي الحدين وتتوفر فيه الشروط التالية:

1) عدد المحاولات (n) كبيرا، غالباً $(n \geq 50)$.
2) احتمال النجاح P(x) صغير جداً، في الغالب $(P(x) \leq 0.10)$.
3) حاصل ضرب $(n * P(x) \leq 5)$.

فإن هذا المتغير يتوزع توزيعاً قريباً من توزيع بواسون، أي أن:

$$P(X=x) = \binom{n}{x} * p^x * q^{n-x} \cong \frac{e^{-\lambda} * \lambda^x}{x!} \quad , \quad \lambda = n * p$$

مثال7:

10% من الأدوات المنتجة في عملية صناعية معينة هي مواد تالفة، أوجد احتمال ان يكون في 10 من هذه الوحدات وحدتان تالفتان بالضبط باستخدام:

أ) توزيع ذي الحدين؟

ب) تقريب توزيع بواسن لتوزيع ذي الحدين؟

الحل:

$$x = 2 \qquad N = 10 \qquad p = 0.10 \qquad q = 0.90$$

أ) توزيع ذي الحدي:

$$x = 2 \quad \rightarrow \quad P(2) = \frac{10!}{2!\,(10-2)!} * 010^2 * 0.90^8 = 0.1937$$

ب) تقريب توزيع بواسن لتوزيع ذي الحدين:

$$\lambda = N * p = 10 * 0.10 = 1 \quad < \quad 5 \qquad OK$$

$$x = 2 \quad \rightarrow \quad P(2) = \frac{\lambda^x * e^{-\lambda}}{x!} = \frac{1^2 * e^{-1}}{2!} = 0.1839$$

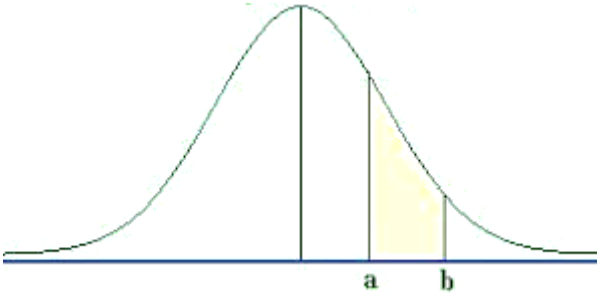
الفصل السادس:

التوزيعات الاحتمالية المتصلة (المستمرة)

مقدمة:

لقد تعرفنا في الفصل السابق عن المتغير العشوائي X وبينا انه يأخذ قيماً محددة متقطعة كما في حالة عدد مرات ظهور الصورة او الكتابة عند رمي العملة المعدنية عدد معين من المرات، فإن هذا التوزيع يكون توزيع متقطع، أما إذا اخذ المتغير X مجموعة من القيم المتصلة فإنه يسمى متغير متصل.

يتم التعبير عن دالة التوزيع المتصلة بمنحنى تكون معادلته: $y = P(x)$



منحنى التوزيع الطبيعي

المساحة الكلية تحت المنحنى تساوي 1، والمساحة تحت المنحنى التي تقع بين الخطوط $x=a$ و $x=b$ (المضللة في الشكل) تعطي احتمال x بين a و b أي ان $(a \le x \le b)$ والتي يمكن التعبير عنها كالتالي:

$$Pr(a \le x \le b)$$

تسمى P(x) دالة كثافة الاحتمال، وإذا تحصلنا على مثل هذه الدالة فإنه يمكن القول ان هذا تعريفاً للتوزيع الاحتمالي المتصل للمتغير x، ويسمى المتغير x متغير عشوائي متصل.

ولاي دالة $f_x(X)$ معرفة على خط الاعداد الحقيقية R، بحيث انه لاي فترة $A \subseteq R$ يكون:

$$P(X \in A) = \int_A f_x(X)\,dx \qquad , \quad A \subseteq R$$

ويمكن القول إن الدالة هي دالة كثافة احتمالية إذا تحقق الشرطين التالين:

$$f_x(x) \ge 0 \qquad , \quad x \in R \quad -1$$
$$\int_{-\infty}^{\infty} f_x(x) = P(R) = P(\Omega) = 1 \quad -2$$

وإن دالة التوزيع التراكمي لهذا المتغير x تكون متصلة على خط الاعداد الحقيقية R بالكامل وتعرف كما يلي:

$$F_X(X) = P(X \le x) = \int_{-\infty}^{x} f_x(u)\,du \qquad , \quad A \subseteq R$$

مثال1:

بفرض ان X متغير عشوائي بدالة كثافة احتمال معرفة كما يلي:

$$f_x(x) = \begin{cases} \frac{1}{2} & , \quad 5 \le x \le 7 \\ 0 & , \quad ow \end{cases}$$

اوجد ما يلي:

1) $P(5.75 \le X \le 6.25)$
2) $P(X \le 6)$
3) $P(X \le t)$ حيث ان t ما بين 5 و 7.

الحل:

1) $P(5.75 \le X \le 6.25) = \int_{5.75}^{6.25} \left(\frac{1}{2}\right) dx = 0.25.$

2) $P(X \le 6) = \int_{-\infty}^{6} (f_x(x)) dx = \int_{-\infty}^{5} f_x(x) dx + \int_{5}^{6} f_x(x) dx = 0 + \int_{5}^{6} \left(\frac{1}{2}\right) dx = 0 + 0.5 = 0.5$

3) $P(X \le t) = \int_{-\infty}^{t} (f_x(x)) dx = \int_{-\infty}^{5} f_x(x) dx + \int_{5}^{t} f_x(x) dx = 0 + \int_{5}^{t} \left(\frac{1}{2}\right) dx = 0 + [\frac{1}{2}x]_5^t = \mathbf{\frac{t-5}{2}}$

مثال2:

إذا كان X متغيراً عشوائياً بدالة كثافة احتمالية معرفة كما يلي:

$$f_x(x) = \begin{cases} ax & , \quad 0 \le x \le 1 \\ a & , \quad 1 \le x \le 2 \\ -ax + 3a & , \quad 2 \le x \le 3 \\ 0 & , \quad ow \end{cases}$$

حيث ان a مقدار ثابت، أوجد قيمة الثابت a؟ ثم أوجد $P(X \le 1.5)$؟

الحل:

لكي تكون الدالة $f_x(x)$ دالة كثافة احتمال يجب تحقق الشرط:

$$\int_{-\infty}^{\infty} f_x(x) = 1$$

وعليه فإن:

$$= \int_{-\infty}^{0} f_x(x) dx + \int_{0}^{1} f_x(x) dx + \int_{1}^{2} f_x(x) dx + \int_{2}^{3} f_x(x) dx + \int_{3}^{\infty} f_x(x) dx = 1$$

$$= \int_{-\infty}^{0} 0 \; dx + \int_{0}^{1} \mathrm{ax}\, dx + \int_{1}^{2} \mathrm{a}\, dx + \int_{2}^{3} (-\mathrm{ax} + 3\mathrm{a})\, dx + \int_{3}^{\infty} 0 \; dx = 1$$

$$= 0 + \frac{a}{2} + a + \frac{a}{2} = 1 \qquad \rightarrow \quad a = 0.5$$

وعليه فإن دالة كثافة الاحتمال $(p.d.f)$ للمتغير العشوائي X تكون كالتالي:

$$f_x(x) = \begin{cases} 0.5x & , 0 \le x \le 1 \\ 0.5 & , 1 \le x \le 2 \\ -0.5x + 1.5 & , 2 \le x \le 3 \\ 0 & , ow \end{cases}$$

$$P(X \le 1.5) = \int_{-\infty}^{1.5} f_x(x)dx = \int_{-\infty}^{0} f_x(x)dx + \int_{0}^{1} f_x(x)dx + \int_{1}^{1.5} f_x(x)dx$$

$$= \int_{-\infty}^{0} 0 \; dx + \int_{0}^{1} 0.5x \, dx + \int_{1}^{1.5} 0.5 \; dx = 0 + 0.25 + 0.25 = 0.50$$

مثال3:

بفرض ان X متغير عشوائي بدالة معرفة كما يلي:

$$f_x(x) = \begin{cases} x & , 0 < x \le 1 \\ 2 - x & , 1 < x \le 2 \\ 0 & , ow \end{cases}$$

المطلوب:

1) أثبت أن الدالة دالة احتمالية ومثلها بيانياً؟
2) أوجد دالة التوزيع التراكمي $F_x(x)$ ومثلها بيانياً؟
3) أوجد $P(0.3 < X \le 1.5)$؟

الحل: $\int_0^1 x \; dx + \int_1^2 (2 - x)dx = 1$

الدالة التراكمية:

$$F_x(x) = \int_{-\infty}^{x} f_x(x)$$

$$x < 0 \qquad F_x(x) = \int_{-\infty}^{x} 0 \; dx = 0$$

$$0 \le x < 1 \qquad F_x(x) = \int_{-\infty}^{0} 0 \; dx + \int_{0}^{x} x \; dx = \frac{x^2}{2}$$

$$1 \le x < 2 \qquad F_x(x) = \int_{-\infty}^{0} 0 \; dx + \int_{0}^{1} x \; dx + \int_{1}^{x} (2 - x)dx = \left(\frac{x^2}{2}\right)_0^1 + \left(2x - \frac{x^2}{2}\right)_1^x$$

$$F_x(x) = 2x - \frac{x^2}{2} - 1$$

$$2 < x \qquad F_x(x) = \int_{-\infty}^{0} 0 \; dx + \int_{0}^{1} x \; dx + \int_{1}^{2} (2 - x)dx + \int_{2}^{+\infty} 0 \; dx = 1$$

$$\mathrm{F_x}\,(x) = \begin{cases} 0 & 0 > x \\ \frac{x^2}{2} & 0 \le x < 1 \\ 2x - \frac{x^2}{2} - 1 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

$$P(0.3 < X \le 1.5) = \int_{0.3}^{1} x\,dx + \int_{1}^{1.5} (2 - x)dx = \mathbf{0.83}$$

$$P(0.3 < X \le 1.5) = \mathrm{F_x}\,(1.5) - \mathrm{F_x}\,(0.3) = \left(2 * 1.5 - \frac{1.5^2}{2} - 1\right) - \frac{0.3^2}{2} = \mathbf{0.83}$$

مثال4:

اذا كان X متغيراً عشوائياً بدلة توزيع تراكمي $(c.d.f)$ معرفة كما يلي:

$$\mathrm{F_x}\,(x) = \begin{cases} 0 & , \quad x \le 0 \\ \frac{x^3}{8} & , \quad 0 < x \le 2 \\ 1 & , \quad x > 2 \end{cases}$$

المطلوب:

أوجد دالة كثافة الاحتمال (p.d.f)؟

الحل:

$$f_x(x) = \begin{cases} \frac{3x^2}{8} & 0 < x < 2 \\ 0 & ow \end{cases}$$

التوزيع الطبيعي:

إن التوزيع الطبيعي يعد من اهم التوزيعات الاحتمالية في الإحصاء ويرجع ذلك للأسباب التالية:

- إن كثيرا من الظواهر التي تظهر في التجارب العملية تتوزع توزيعاً طبيعياً.
- في بعض الأحيان قد لا يكون المتغير العشوائي موزعاً توزيعاً طبيعياً ولكن يمكن تحويله الى متغير عشوائي يتوزع توزيعاً طبيعياً وذلك تحت شروط معينة.
- قد توجد توزيعات معقدة وبالتالي يمكن ان يستخدم التوزيع الطبيعي تقريباً لها.

إذا كان X متغيراً عشوائياً يتوزع وفق التوزيع الطبيعي فإن دالة الكثافة الاحتمالية لهذا المتغير تصاغ كالتالي:

$$f_x(x) = \frac{1}{\sqrt{2\pi S^2}}\ e^{-\frac{1}{2S^2}(x-\bar{x})^2} \qquad -\infty < x < +\infty$$

حيث ان:

$\bar{x}$: المتوسط الحسابي $(-\infty < \bar{x} < +\infty)$، S^2 : التباين $(S^2 > 0)$، e= 2.71828

ويرمز للتوزيع الطبيعي في العادة بالرمز $N\ (\bar{x}, S^2)$.

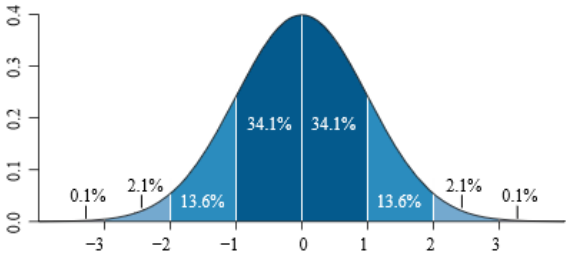
حيث ان المساحة الكلية تحت المنحنى تساوي 1، ولهذا فإن المساحة تحت المنحنى والمحصورة بين الاحداثيات $x=a$ و $x=b$ حيث ان ($a < b$) تمثل احتمال ان تقع x بين a و b ويعبر عنها $P_r(a \le x \le b)$.

وعند التعبير عن المتغير x بدلالة الوحدات المعيارية فإن: $z = \frac{x-\bar{x}}{s.d}$

فإن معادلة التوزيع الطبيعي أعلاه تستبدل بما يسمى بالصورة القياسية او المعيارية:

$$f_x(x) = y = \frac{1}{s.d*\sqrt{2\pi}}\ e^{-\frac{1}{2}z^2}$$

وفي هذه الحالة يقال أن z تتوزع توزيع معتدل متوسطه (صفر) وتباينه (1)، والشكل البياني التالي يبين المنحنى الطبيعي.



خصائص التوزيع الطبيعي:

1- المتوسط الحسابي = 0

2- التباين = الانحراف المعياري = 1

3- معامل الالتواء = 0

4- معامل التفرطح = 3

5- الانحراف المتوسط = $s.d * \sqrt{\frac{2}{\pi}} = 0.79788$

مثال5:

في أحد الامتحانات النهائية لمادة الإحصاء كان المتوسط الحسابي لدرجات الطلبة هو 72 درجة، والانحراف المعياري هو 15 درجة، أوجد الدرجات المعيارية (z) للطلبة الحاصلين على الدرجات التالية:

1) 60 درجة. 2) 93 درجة. 3) 72 درجة.

الحل:

1) $z = \frac{x-\bar{x}}{s.d} = \frac{60-72}{15} = -0.8$
2) $\frac{93-72}{15} = 1.4$
3) $\frac{72-72}{15} = 0$

مثال6:

اوجد المساحة تحت المنحنى الطبيعي في كل من الحالات الاتية باستخدام جدول التوزيع الطبيعي:

1) بين $0 \leq z \leq 1.2$
2) بين $z= -0.68$ و $z=0$
3) بين $z= -0.46$ و $z=2.21$
4) بين $z= 0.81$ و $z=1.94$
5) الى يسار $z=-0.6$
6) الى يمين $z= -1.28$
7) الى يمين $z= 2.05$ والى يسار $z= -1.44$

الحل:

1) $0 \leq z \leq 1.2$ z=0 ➔ A= 0.00

A_T= 0.3849 + 0 = 0.3849

z= 1.2 ➔ A = 0.3849

2) $-0.68 \leq z \leq 0$ z=0 ➔ A= 0.00

A_T= 0.2517 + 0 = 0.2517

z= −0.68 ➔ A = 0.2517

3) $-0.46 \leq z \leq 2.21$ z=−0.46 ➔ A= 0.1772

A_T= 0.1772 + 0.4868 = 0.6640

z= 2.21 ➔ A = 0.4868

4) $0.81 \leq z \leq 1.94$ z=0.81 ➔ A= 0.4738

A_T= 0.4738 − 0.2910= 0.1828

z= 2.94 ➔ A = 0.2910

5) $z \leq -0.6$ z=−0.6 ➔ A= 0.2257

A_T= 0.5 − 0.2257 = 0.2743

6) $z \geq -1.28$ z=−1.28 ➔ A= 0.3997

A_T= 0.5 + 0.3997 = 0.8997

7) $2.05 \leq z \leq -1.44$ z=−1.44 ➔ A= 0.4251

A_T= (0.5− 0.4251) + (0.5 − 0.4798) = 0.0951

z= 2.05 ➔ A = 0.4798

التوزيع الطبيعي كتقريب لتوزيع ذي الحدين:

يمكن تقريب توزيع ذي الحدين باستخدام التوزيع الطبيعي اذا كان N كبيره وكل من p و q ليسا قريبين من الصفر وبالطبع فإن التقريب يصبح أكثر دقة كلما ازدادت قيمة N ومن الناحية العملية يعتبر التقريب جيداً إذا كان كل من Np و N_q أكبر من خمسة.

مثال7:

أوجد احتمال الحصول على ما بين 3 و 6 صور في 10 رميات لعملة معدنية متوازنة متضمناً 3 و 6.

1) باستخدام توزيع ذي الحدين مباشرة.
2) باستخدام التوزيع الطبيعي كتقريب لتوزيع ذي الحدين.

الحل:

1) $p = 0.5 \qquad q = 0.5 \qquad N = 10$

$$P(3 \le X \le 6) = \sum_{X=3}^{6} {}_{10}C_x\, p^x q^{10-x}$$

$$= [0.1171875 + 0.2051 + 0.2461 + 0.2051] = \mathbf{0.7735}$$

2) $\bar{x} = Np = 10 * 0.5 = 5 \qquad s.d = \sqrt{Npq} = \sqrt{10 * 0.5 * 0.5} = 1.5811$

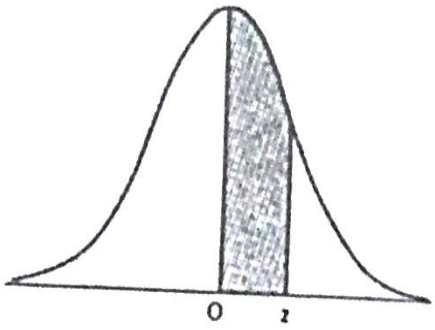
$$x_1 = 2.5 \qquad x_2 = 6.5$$

$$z_1 = \frac{2.5 - 5}{1.5811} = -1.58 \quad \rightarrow \quad \text{A} = 0.4429$$

$$z_2 = \frac{6.5 - 5}{1.5811} = 0.95 \quad \rightarrow \quad \text{A} = 0.3289$$

$$A_T = 0.4429 + 0.3289 = \mathbf{0.7718}$$

جدول (4) : جدول التوزيع الطبيعي المعياري ، الذى يعطي القيم التى تنحصر بين 0 و z الموجبة ، أما القيم السالبة فيتم الحصول عليها بالتماثل .



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0 6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000									

المصدر :

2nd ed., (1977)

الفصل السابع:

الارتباط والانحدار

Correlation and Regression

الارتباط Correlation:

يدرس الارتباط طبيعة العلاقة بين عدد من الظواهر ونوعها أي بين متغيرين أو أكثر وينقسم الى ثلاثة أنواع هي سببي وتبادلي ووهمي كما ينقسم حسب المتغيرات الى ارتباط بسيط ويربط بين متغيرين فقط او ارتباط متعدد يربط بين عدة متغيرات او ارتباط جزئي الذي يتم فيه تثبيت كل المتغيرات عدا متغير واحد لدارسة أثره على متغير الدراسة.

أولاً: الارتباط البسيط:

هو الارتباط الذي يدرس العلاقة بين متغيرين وهذه العلاقة اما ان تكون خطية او غير خطية.

الارتباط الخطي:

قياس الارتباط بين البيانات غير المبوبة:

ان الارتباط بين البيانات قد يكون ارتباط خطي موجب وقد يكون ارتباط خطي سالب وقد لا يكون هناك ارتباط مطلقاً، كما ان الارتباط الخطي سواء كان موجب او سالب يمكن ان يكون قوياً او ضعيفاً.

1- معامل ارتباط بيرسون (r):

يتم التأكيد في البداية على ان الارتباط خطي تقريباً وذلك برسم شكل الانتشار، فإذا كان غير خطي فإن هذا العامل لا يمكن استخدامه.

تكون قيمة هذا العامل ما بين [1+ ، 1−] وكلما اقتربت قيمة المعامل من 1 دل ذلك على قوة الارتباط الخطي بين المتغيرين، بينما تدل الإشارة على طبيعة العلاقة اما موجبة او سالبة.

يمكن إيجاد معامل ارتباط بيرسون بالعلاقة التالية:

$$r = \frac{\sum xy}{NS_xS_y}$$

حيث ان: x: انحراف القيمة x عن المتوسط الحسابي.

y: انحراف القيمة y عن المتوسط الحسابي.

N: عدد المشاهدات.

S_x : الانحراف المعياري للمتغير X.

S_y : الانحراف المعياري للمتغير Y .

مثال1:

في احدى التجارب تحصل باحث على قراءات عن التيار y والجهد x والمطلوب حساب درجة ونوع الارتباط الخطي بين المتغيرين وذلك بناءاً على البيانات التالية:

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

الحل:

$$\bar{x} = \frac{\sum x}{N} = \frac{56}{8} = \mathbf{7} \quad , \quad \bar{y} = \frac{\sum y}{N} = \frac{40}{8} = \mathbf{5}$$

	X	Y	x	y	xy	x^2	y^2
	1	1	−6	−4	24	36	16
	3	2	−4	−3	12	16	9
	4	4	−3	−1	3	9	1
	6	4	−1	−1	1	1	1
	8	5	+1	0	0	1	0
	9	7	+2	+2	4	4	4
	11	8	+4	+3	12	16	9
	14	9	+7	+4	28	49	16
∑	56	40	0	0	84	132	56

$$S_x = \sqrt{\frac{x^2}{N}} = \sqrt{\frac{132}{8}} = \mathbf{4.06} \quad , \quad S_y = \sqrt{\frac{y^2}{N}} = \sqrt{\frac{56}{8}} = \mathbf{2.65}$$

$$r = \frac{\sum xy}{NS_xS_y} = \frac{84}{8 * 4.06 * 2.65} = \mathbf{0.976}$$

ارتباط موجب قوي.

- **طريقة حساب معامل بيرسون باستخدام وسط فرضي:**

يمكن حساب معامل بيرسون باستخدام الانحرافات عن الوسط الفرضي باستخدام العلاقة التالية:

$$r = \frac{\sum x'y' - \frac{\sum x' * \sum y'}{N}}{NS_xS_y}$$

حيث ان:

x': انحراف القيمة x عن وسط فرضي.

y' : انحراف القيمة y عن وسط فرضي.

$$S_x = \sqrt{\frac{\sum x'^2}{N} - (\frac{\sum x'}{N})^2} \qquad S_y = \sqrt{\frac{\sum y'^2}{N} - (\frac{\sum y'}{N})^2}$$

مثال2:

لنفس المثال السابق، احسب درجة ونوع الارتباط الخطي بين المتغيرين باستخدام وسط فرضي؟

الحل:

نفرض ان الوسط الفرضي للمتغير x هو 8، والوسط الفرضي للمتغير y هو 7.

X	Y	x'	y'	x' y'	x'^{2}	y'^{2}
1	1	−7	−6	42	49	36
3	2	−5	−5	25	25	25
4	4	−4	−3	12	16	9
6	4	−2	−3	6	4	9
8	5	0	−2	0	0	4
9	7	+1	0	0	1	0
11	8	+3	+1	3	9	1
14	9	+6	+2	12	36	4
∑	56	40	−8	−16	100	140	88

$$S_x = \sqrt{\frac{\sum x'^2}{N} - (\frac{\sum x'}{N})^2} = \sqrt{\frac{140}{8} - (\frac{-8}{8})^2} = \mathbf{4.06} \qquad S_y = \sqrt{\frac{\sum y'^2}{N} - (\frac{\sum y'}{N})^2} = \sqrt{\frac{88}{8} - (\frac{-16}{8})^2} = \mathbf{2.65}$$

$$r = \frac{\sum x'y' - \frac{\sum x' * \sum y'}{N}}{NS_xS_y} = \frac{100 - (\frac{(-8) * (-16)}{8})}{8 * 4.06 * 2.65} = \mathbf{0.976} \leftarrow \text{نفس الحل}$$

- **طريقة حساب معامل بيرسون باستخدام الأرقام الخام:**

يمكن استخدام الأرقام الخام مباشرة خاصة إذا كان عدد المشاهدات قليل وقيمتها منخفضة وذلك باستخدام العلاقة التالية:

$$r = \frac{\sum xy - N\,\bar{X}\,\bar{Y}}{N S_x S_y}$$

$$S_x = \sqrt{\frac{\sum x^2}{N} - (\frac{\sum x}{N})^2} \qquad S_y = \sqrt{\frac{\sum y^2}{N} - (\frac{\sum y}{N})^2}$$

مثال3:

لنفس المثال السابق، احسب درجة ونوع الارتباط الخطي بين المتغيرين باستخدام الأرقام الخام؟

الحل:

نفرض ان الوسط الفرضي للمتغير x هو 8، والوسط الفرضي للمتغير y هو 7.

	X	Y	x y	x^2	y^2
	1	1	1	1	1
	3	2	6	9	4
	4	4	16	16	16
	6	4	24	36	16
	8	5	40	64	25
	9	7	63	81	49
	11	8	88	121	64
	14	9	126	196	81
∑	56	40	364	524	256

$$\bar{X} = \frac{\sum x}{N} = \frac{56}{8} = \mathbf{7}$$

$$\bar{Y} = \frac{\sum y}{N} = \frac{40}{8} = \mathbf{5}$$

$$S_x = \sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2} = \sqrt{\frac{524}{8} - \left(\frac{56}{8}\right)^2} = \mathbf{4.06}$$

$$S_y = \sqrt{\frac{\sum y^2}{N} - (\frac{\sum y}{N})^2} = \sqrt{\frac{256}{8} - (\frac{40}{8})^2} = \mathbf{2.65}$$

$$r = \frac{\sum xy - N\,\bar{X}\,\bar{Y}}{N S_x S_y} = \frac{364 - 8 * 7 * 5}{8 * 4.06 * 2.65} = \mathbf{0.976}$$ ← نفس الحل

2- معامل ارتباط سبيرمان: (ارتباط الرتب)

هذا المعامل سهل الحساب ولكنه يعطي قيمة تقريبية أقل دقة من معامل بيرسون لأنه يعتمد على تراتيب القيم بينما معامل بيرسون يعتمد على القيوم نفسها.

ومن مزايا هذا المقياس هي إمكانية دراسة الارتباط بين البيانات الوصفية (غير الرقمية).

يَستخدِم قانون سبيرمان رتب القيم او المفردات بدلا عن القيم نفسها، وتُستخدَم في حسابه العلاقة التالية:

$$r = 1 - \frac{6 * \sum D^2}{N\,(N^2 - 1)}$$

حيث ان:

D: الفرق المطلق بين رتب المتغيرين.

N: عدد المشاهدات.

مثال4: احسب معامل ارتباط الرتب بين المتغيرين في الجدول التالي:

x	5	7	9	10	8	6
y	1	3	2	6	5	4

الحل:

نقوم بترتيب البيانات تصاعدياً او تنازلياً: نقوم بتكوين الجدول التالي بناءاً على الرتب في الجدول السابق:

x	y	رُتب x	رُتب y	D	D^2
5	1	6	6	0	0
7	3	4	4	0	0
9	2	2	5	3	9
10	6	1	1	0	0
8	5	3	2	1	1
6	4	5	3	2	4
				$\sum$	14

x	y
10	6
9	5
8	4
7	3
6	2
5	1

$$r = 1 - \frac{6 * \sum D^2}{N\,(N^2 - 1)} = 1 - \frac{6 * 14}{6\,(6^2 - 1)} = 0.600$$

مثال5:

احسب درجة الارتباط بين درجات الطلبة في مقرر الإحصاء ودرجة الذكاء، علماً بأن العينة التي تم الحصول عليها كانت على النحو التالي:

الذكاء x	عالي	عالي	متوسط	ضعيف	ضعيف جدا	عالي	فوق المتوسط	تحت المتوسط
درجة الإحصاء y	90	95	90	59	52	90	89	76

الحل:

ملاحظة: عند تكرر القيمة فإنها تأخذ نفس الترتيب، فمثلا الذكاء عالي سيأخذ الرتب 1 و2 و3 لانه متكرر 3 مرات وتكون الرتبة تساوي متوسط الرتب أي $\frac{1+2+3}{3} = 2$ ، وبالمثل فالدرجة 90 ستأخذ الرتب 2 و3 و4 وبالتالي فإن الرتبة هي متوسط الرتب أي $\frac{2+3+4}{3} = 3$ مع ملاحظة ان الرتب يمكن ان تأخذ كسور عشرية.

نقوم بترتيب البيانات تصاعدياً او تنازلياً:

y	x	الرتب
عالي	95	1
عالي	90	2
عالي	90	3
فوق المتوسط	90	4
متوسط	89	5
تحت المتوسط	76	6
ضعيف	59	7
ضعيف جداً	52	8

نقوم بتكوين الجدول التالي بناءاً على الرتب في الجدول السابق:

x	y	رُتب x	رُتبy	D	D^2
عالي	90	2	3	1	1
عالي	95	2	1	1	1
متوسط	90	5	3	2	4
ضعيف	59	7	7	0	0
ضعيف جداً	52	8	8	0	0
عالي	90	2	3	1	1
فوق المتوسط	89	4	5	1	1
تحت المتوسط	76	6	6	0	0
				$\sum$	8

$$r = 1 - \frac{6 * \sum D^2}{N\,(N^2 - 1)} = 1 - \frac{6 * 8}{8\,(8^2 - 1)} = 0.905$$

الانحدار الخطي:

الانحدار هو تمثيل العلاقات المتوسطة بين المتغيرات، فإذا وجدت العلاقة بين المتغيرات المطلوب دراستها يمكن توفيق معادلة لمنحنى او الخط يحدد طبيعة تلك العلاقة ومن تم يمكن استخدام هذه المعادلات للتقدير، وتسمى المنحنيات او المعادلات التي تربط بين المتغيرات بمنحنيات او معادلات الانحدار.

يعتبر ابسط صور الانحدار على الاطلاق هو الانحدار الخطي او الانحدار المستقيم وهناك عدة طرق للحصول على المعادلة المطلوبة:

أولاً: طريقة الرسم البياني:

في البداية نرسم شكل الانتشار تم نحاول ان نمهد باليد الخط المستقيم الذي يمر بين أكبر عدد من النقاط، وبالتالي فإن هذا الخط يصف العلاقة المتوسطة بين المتغيرين.

وإذا كان الارتباط قوياً بين المتغيرين فإن أغلب النقاط ستقع على الخط او قريبة منه.

وإذا اعتبرنا ان y هو المتغير التابع وx هو المتغير المستقل فإن معادلة خط الانحدار يمكن تمثيلها بالمعادلة التالية:

$$y = A + Bx$$

حيث ان:

A و B ثوابت يمكن إيجادها من خلال الرسم، حيث يمثل الثابت A الجزء المقطوع من المحور y والثابت B ميل الخط المستقيل، وبإيجاد قيمة الثوابت نجد المعادلة التي تربط بين الظاهرتين.

ثانيا: طريقة المربعات الصغرى:

في هذه الطريقة يراعى ان يكون مجموع مربعات انحرافات قيم المشاهدات (الفعلية) عن خط الانحدار أصغر ما يمكن وكذلك مجموع الانحرافات يساوي صفر.

$$\sum(y - \hat{y}) = 0 \qquad , \quad \sum(y - \hat{y})^2 = \text{أقل ما يمكن}$$

حيث ان: $\hat{y}$: تمثل القيمة المقدرة باستخدام دالة الانحدار.

y: تمثل القيم الفعلية للمتغير.

وبالتعويض عن y في معادلة الانحدار في الشرط الثاني واستخدام التفاضل الجزئي يمكن إيضاح أنه لكي يكون الشرط الثاني نهاية صغرى يلزم ان يكون:

$$\sum y = N\ A + B\ \sum x \qquad \rightarrow \quad (1)$$

$$\sum xy = \ A\ \sum x + B\ \sum x^2 \qquad \rightarrow (2)$$

حيث ان: N: عدد القراءات للعينة (عدد المفردات).

$\sum y$: مجموع قيم المتغير التابع y.

$\sum x$: مجموع قيم المتغير المستقل x.

$\sum xy$: مجموع حاصل ضرب أزواج القيم المتناظرة للمتغيرين.

وبحل المعادلتين (1) و (2) آنياً نجد ان:

$$\hat{y} = A + B\ x$$

$$B = \frac{N\ \sum xy - \sum x\ \sum y}{N\ \sum x^2 - (\sum x)^2} \qquad , \qquad A = \bar{y} - B\ \bar{x}$$

وبالمثل يمكن إيجاد معادلة انحدار y على x $(\frac{y}{x})$ وتكون المعادلة على النحو التالي:

$$\hat{x} = H + W\ y$$

$$W = \frac{N\ \sum xy - \sum x\ \sum y}{N\ \sum y^2 - (\sum y)^2} \qquad , \qquad H = \bar{x} - W\ \bar{y}$$

مثال6: احسب معادلتي الانحدار (x على y) و (y على x) من البيانات التالية:

x	1.2	0.8	1	1.3	0.7	0.8	1	0.6	0.9	1.1
y	101	92	110	120	90	82	93	75	91	105

الحل:

	x	y	xy	X^2	Y^2
	1.2	101	121.2	1.44	10201
	0.8	92	73.6	0.64	8464
	1	110	110	1	12100
	1.3	120	156	1.69	14400
	0.7	90	63	0.49	8100
	0.8	82	65.6	0.64	6724
	1	93	93	1	8649
	0.6	75	45	0.36	5625
	0.9	91	81.9	0.81	8281
	1.1	105	115.5	1.21	11025
$\sum$	9.4	959	924.8	9.28	93569

أولا: معادلة انحدار $\frac{x}{y}$:

$$B = \frac{10 * (924.4) - (9.4)(959)}{(10)(9.28) - (9.4)^2} = 51.67$$

$$A = \bar{y} - B\,\bar{x} = \frac{959}{10} - 51.67 * \frac{9.4}{10} = 47.33$$

$$\hat{y} = 47.33 + 51.67\,x$$

ثانياً: معادلة انحدار $\frac{y}{x}$:

$$\hat{x} = H + W\,y$$

$$W = 0.014$$

$$H = -0.403$$

$$\hat{x} = -0.403 + 0.014\,y$$

ملاحظة:

يمكن استخدام المعادلات التي توصلنا اليها للتنبؤ بقيمة المتغير التابع عندما يأخذ المتغير المستقل قيمة محدد ضمن البيانات الموجودة بالعينة او خارجها.

مثال7:

أوجد قيمة y ($\hat{y}$) التقديرية في المثال السابق عندما x=2, x=1؟

$$x = 1 \quad \rightarrow \quad \hat{y} = 47.33 + 51.67\,(1) = 99$$

$$x = 2 \quad \rightarrow \quad \hat{y} = 47.33 + 51.67\,(2) = 150.67$$

ملاحظة:

يمكن حساب معادلة $\frac{x}{y}$ و $\frac{y}{x}$ باستخدام العلاقات التالية:

$$\hat{y} = \bar{y} + r\,\frac{s_y}{s_x}\,(x - \bar{x})$$

$$\hat{x} = \bar{x} + r\,\frac{s_x}{s_y}\,(y - \bar{y})$$

مثال8:

احسب معادلة انحدار $\frac{y}{x}$ ومعادلة انحدار $\frac{x}{y}$ من البيانات التي سبق حساب معامل ارتباط بيرسون لها؟ (اول مثال)

الحل:

بالرجوع للمثال نجد ان:

$$\bar{y} = 5 \qquad \bar{x} = 7 \qquad s_x = 4.06 \qquad s_y = 2.65 \qquad r = 0.976$$

وبالتعويض في العلاقات السابقة نجد ان:

$$\hat{y} = \bar{y} + r\,\frac{s_y}{s_x}(x - \bar{x}) = 5 + 0.976 * \frac{2.65}{4.06} * (x - 7)$$

$$\hat{y} = 0.54 + 0.64\,x$$

$$\hat{x} = \bar{x} + r\,\frac{s_x}{s_y}(y - \bar{y}) = 7 + 0.976 * \frac{4.06}{2.65} * (y - 5)$$

$$\hat{x} = -0.477 + 1.495\,y$$

ملاحظة:

يمكن حساب معامل الارتباط إذا ما تحصلنا على معادلتي خط الانحدار مباشرة وذلك بأخذ الجدر التربيعي لحصل ضرب معامل x في معادلة انحدار $\frac{x}{y}$ مضروب في معامل y في معادلة انحدار $\frac{y}{x}$.

مثال9:

إذا علمت ان معادلتي انحدار بين متغيرين كالتالي:

$$\frac{x}{y} \quad \rightarrow \quad \hat{y} = 4 + 0.5\,x$$

$$\frac{y}{x} \quad \rightarrow \quad \hat{x} = 2.4 + 0.61\,y$$

$$r = \sqrt{0.5 * 0.61} = 0.5523$$

المراجـــــــــــــع:

1- د. علي عبد السلام العماري، د. علي حسن العجيلي، الإحصاء والاحتمالات النظرية والتطبيق، منشورات ELGA، 2000.

2- د. أحمد عبد السميع طبيّة، مبادئ الإحصاء، الطبعة الأولى، دار البداية، 2008.

3- د. شبيجل، د. شيلر، د. سرينيقاسار، ترجمة: د. محمود علي أبو النصر، د. مصطفى جلال مصطفى، الاحتمالات والإحصاء، الطبعة الأولى، 2004.

4- Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, Probability & Statistics for Engineers & Scientists, 9th Edition.