جامعـــة الملــك عبــد العزيــز

مقدمـــة في الإحصـــاء

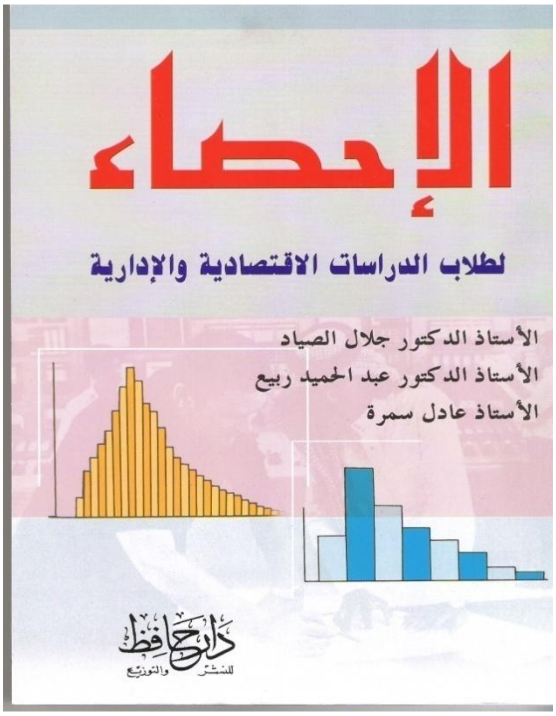
لطلاب الدراسات الاقتصادية والإدارية

قســم الإحصـــاء

1429/1428هـ

فهرس

الكتاب المقرر



الإحصاء لطلاب الدراسات

الاقتصادية والإدارية

تأليف

أ. د. جلال الصياد

أ.د. عبدالحميد ربيع

أ. عادل سمرة

مقدمـــــة

علم الإحصاء: هو العلم الذي يبحث في:

- طرق جمع البيانات الصحيحة والدقيقة عن ظاهرة ما، ثم تلخيص هذه البيانات (بيانياً أو جدولياً) في صورة مبسطة يسهل معها قراءتها واستخدامها.
- وصف هذه البيانات ثم تحليلها واستخراج النتائج منها، واتخاذ القرارات المناسبة.
- دراسة علاقة الظاهرة بباقي الظواهر وتقدير قيمة الظاهرة في المستقبل.

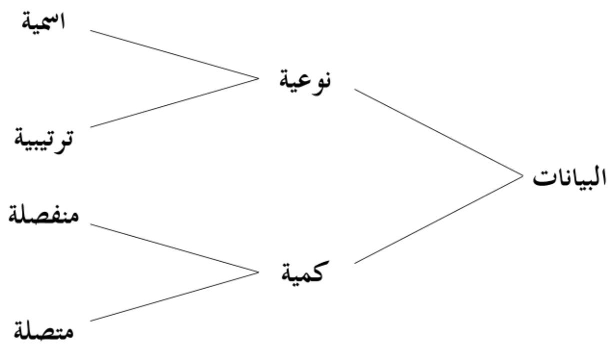
أقسام علم الإحصاء

- الإحصاء الوصفي: يختص بجمع البيانات وعرضها و وصفها وتلخيصها.
- الإحصاء الاستدلالي: يختص باستخلاص وتفسير النتائج واتخاذ القرارات.

مصطلحات إحصائية:

- **المجتمع**: هو مجموعة من المفردات أو المشاهدات أو الأشخاص والتي نرغب في دراسة وتحليل خصائصه، وهناك نوعين: مجتمع محدود أو نهائي و مجتمع غير محدود أو لا نهائي.
- **العينة**: هي أي مجموعة جزئية من المجتمع، وتعبر عنه أصدق تمثيل.
- **المعلمة**: عبارة عن قيمة تعبر عن بيانات المجتمع.

- **إحصاء العينة (المقياس):** عبارة عن قيمة تعبر عن بيانات العينة.
- **البيانات:** عبارة عن مجموعة القيم أو القياسات للمتغير الذي يرافق المفردات أو عناصر المجتمع. وقد تكون في شكل أرقام أو صفات أو رموز، ويمكن تصنيف البيانات على النحو الآتي:



البيانات الاسمية: عبارة عن اسم أو وصف لأي عنصر أو مفردة في المجتمع.

البيانات الترتيبية: عبارة عن اسم أو وصف يعبر عن التفضيل أو الترتيب لأي عنصر في المجتمع.

البيانات المنفصلة: عبارة عن قيم تدل على صفة يمكن عدها، وتأخذ قيم صحيحة فقط.

البيانات المتصلة: عبارة عن قيم تدل على صفة يمكن قياسها، وتأخذ جميع القيم الصحيحة والكسرية.

مثال: حدد نوع البيانات التالية:

البيان	بيانات اسمية	بيانات ترتيبية	بيانات منفصلة	بيانات متصلة
عدد أسهم شركة مساهمة المخصصة للفرد				
عدد الإداريين في إحدى الأقسام بالجامعة				
الدرجات العلمية لعضو هيئة التدريس				
رضا المستهلك عن منتج معين من إحدى الشركات				
كمية الحبر الموجودة في أحد أنواع الاقلام				
أنواع الألوان المستخدمة في طباعة كتاب معين				
عدد الغائبين من الموظفين في أحد الأيام				
معدلات الطلاب في إحدى الشعب بإحدى الجامعات				

الباب الأول: التوزيعات التكرارية

1.1. التوزيعات التكرارية : هي عبارة عن جداول تلخص فيها بيانات العينة إلى فئات، ولكل فئة تكرار تتحدد قيمته حسب البيانات. وتسمى البيانات بعد تلخيصها في جداول تكرارية ***بالبيانات المبوبة***. وعادة ما نلجأ لهذا الأسلوب إذا كانت لدينا عدد كبير من البيانات.

مثال (1.1) (مثال على البيانات الوصفية) البيانات الآتية تمثل المؤهلات العلمية لعينة من (35) موظف بإحدى الشركات، والمطلوب تلخيص هذه البيانات في جدول تكراري.

ثانوي	ثانوي	دكتوراه	ثانوي	جامعي	جامعي	جامعي
ابتدائي	ثانوي	جامعي	ثانوي	متوسط	جامعي	جامعي
ثانوي	ثانوي	متوسط	ثانوي	ثانوي	جامعي	دكتوراه
جامعي	جامعي	ثانوي	ثانوي	جامعي	ثانوي	جامعي
ثانوي	ثانوي	جامعي	ثانوي	ثانوي	جامعي	متوسط

الحل:

جدول تكراري ⟵ جدول تفريغي

عدد الموظفين (f)	المؤهل العلمي (c)
2	دكتوراه
13	جامعي
16	ثانوي
3	متوسط
1	ابتدائي
35	$\sum$

الفئات c	العلامات	التكرار f
دكتوراه	//	2
جامعي	~~////~~	13
ثانوي	~~////~~ ~~////~~ /	16
متوسط	///	3
ابتدائي	/	1
المجموع ($\sum$)		35

مثال (1.2): (مثال على البيانات المنفصلة) فيما يلي عدد الغيابات لعينة من (30) موظف بإحدى الشركات:

0	3	0	0	3	0
2	2	0	1	2	1
0	0	1	2	4	0
4	2	1	0	1	0
0	2	0	1	3	2

المطلوب: تلخيص بيانات العينة في جدول تكراري.

جدول تكراري ← جدول تفريغي

عدد الموظفين (f)	عدد الغيابات (c)
12	0
6	1
7	2
3	3
2	4
30	$\sum$

التكرار f	العلامات	الفئات c
12	𝍸 𝍸 //	0
6	𝍸 /	1
7	𝍸 //	2
3	///	3
2	//	4
30	$\sum$	

مثال (1.3): (مثال على البيانات المتصلة) البيانات التالية تمثل الأجور اليومية بالريال لعينة مكونة من (50) عامل بإحدى المصانع:

47	36	40	55	75	53	46	43	21	10
66	56	46	35	47	32	52	48	41	30
27	25	57	15	37	22	63	21	61	62
54	42	35	49	39	32	45	31	72	50
65	18	79	23	48	44	32	51	44	42

المطلوب: تلخيص بيانات العينة في جدول تكراري.

الحل: نتبع الخطوات التالية:

1. نحسب **المدى** (R) والذي يُعَرف بأنه الفرق بين أكبر قيمة وأقل قيمة

$$R = \max - \min = 79 - 10 = \boxed{69}$$

2. نوجد **عدد الفئات** (k)

$$k = 1+3.3\times\log n = 1+3.3\times\log(50)$$
$$= 1+(3.3\times1.69897) = 1+5.61 = 6.61 \approx \boxed{7}$$

3. نحدد **طول الفئة** (h)

$$h = \frac{R}{k} = \frac{69}{7} = 9.86 \approx \boxed{10}$$

4. نكون الجدول التفريغي، بأن نضع الفئات في العمود الأول من الجدول، مع ملاحظة أن الفئة الأولى لابد أن تبدأ أو تشمل أصغر قيمة، والفئة الأخيرة لابد أن تشمل أكبر قيمة، ثم نكمل الحل كالمعتاد.

جدول تفريغي ⟵ **جدول تكراري**

الفئات c	العلامات	التكرار f
10 -	///	3
20 -	~~////~~ /	6
30 -	~~////~~ ~~////~~	10
40 -	~~////~~ ~~////~~ ~~////~~	15
50 -	~~////~~ ///	8
60 -	~~////~~	5
70 - 80	///	3
$\sum$		50

الأجور (c)	عدد العمال (f)
10 -	3
20 -	6
30 -	10
40 -	15
50 -	8
60 -	5
70 - 80	3
$\sum$	50

ملاحظات هامة:

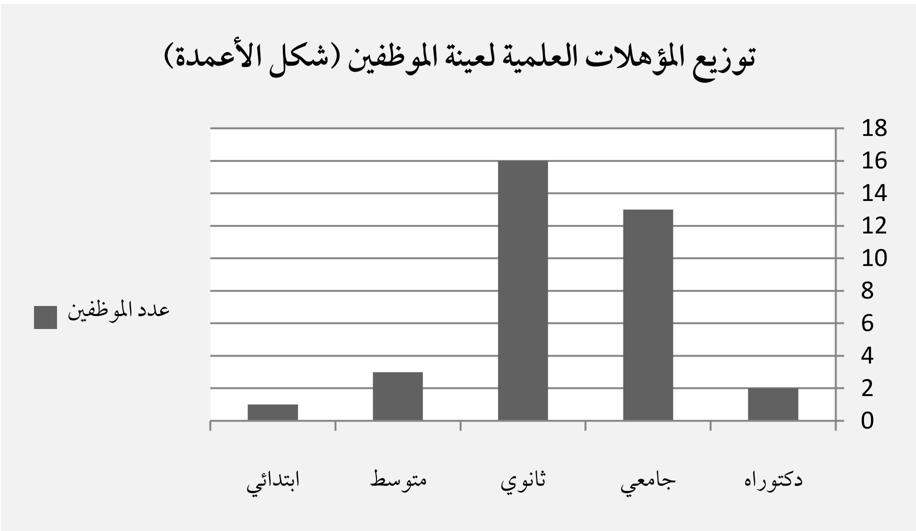
1. يرمز للفئات بالرمز c وللتكرار بالرمز f وحجم العينة بالرمز n، ويجب أن يكون مجموع التكرارات مساوي لحجم العينة، أي $\sum f = n$
2. في الفئة الأولى "- 10" المقصود بها من 10 إلى أقل من 20، كما أن القيمة 10 تمثل **الحد الأدنى للفئة الأولى**، وينطبق القول على بقية القيم.
3. **طول الفئة** هو الفرق بين الحد الأعلى للفئة والحد الأدنى للفئة.

1.2. التمثيل البياني للتوزيع التكراري : يمكن وصف البيانات النوعية والبيانات الكمية المنفصلة باستخدام شكل الأعمدة، أما البيانات الكمية يمكن تمثيلها بالمدرج التكراري أو بالمضلع التكراري أو بالمنحنى التكراري، وخطوات تكوين هذه الأشكال على النحو التالي:

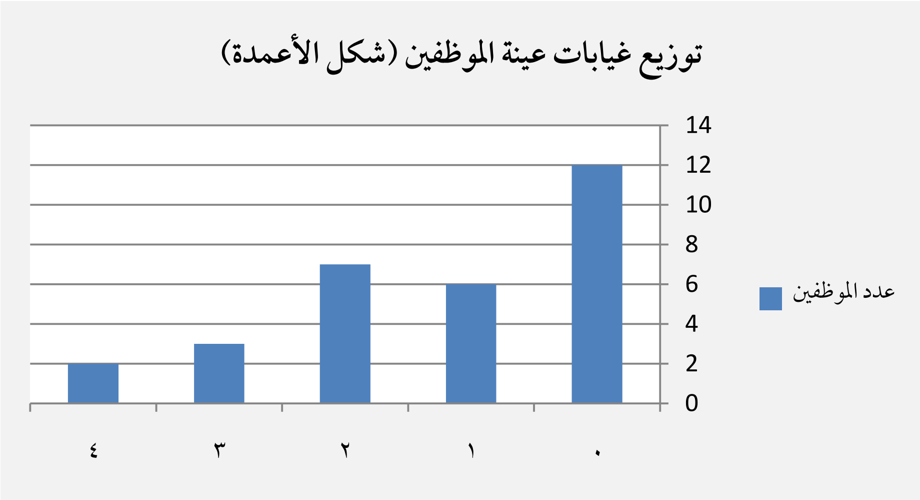
1. رسم محورين؛ الأفقي يمثل الفئات c والرأسي يمثل التكرار f.
2. رسم أعمدة طولها يعتمد على عدد التكرارات، وتكون غير متجاورة إذا كانت البيانات نوعية أو كمية منفصلة، ومتجاورة إذا كانت البيانات كمية ، ويسمى الشكل في هذه الحالة بالمدرج التكراري.
3. لتكوين المضلع التكراري نضع نقاط في منتصف أعمدة المدرج التكراري ثم نقوم بوصلها بخطوط مستقيمة، أما المنحنى التكراري فنقوم بوصل النقاط في منتصف الأعمدة يدوياً.

مثال (1.4): **مثل التوزيعات التكرارية في الأمثلة السابقة بما يناسبها من أشكال بيانية.**

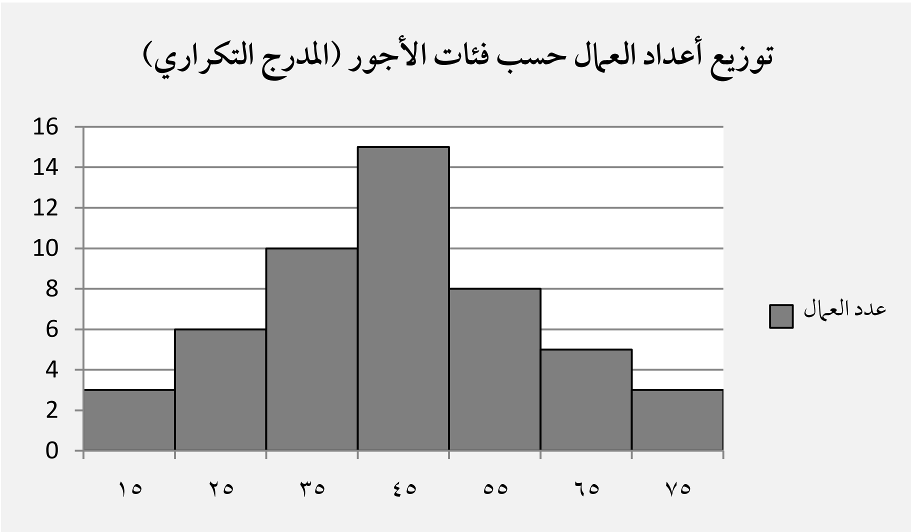
الحل:



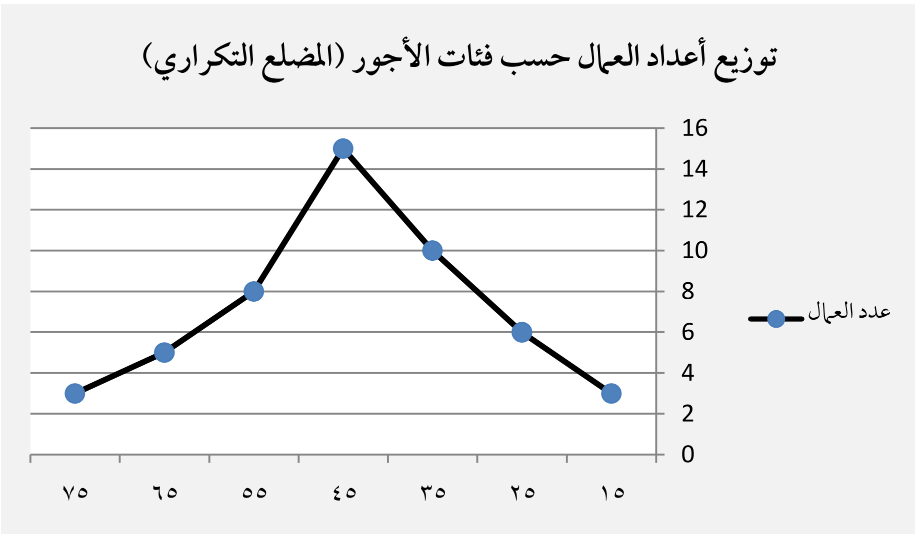
رسم توضيحي 1: توزيع أعداد الموظفين حسب مؤهلاتهم العلمية (مثال 1.1)



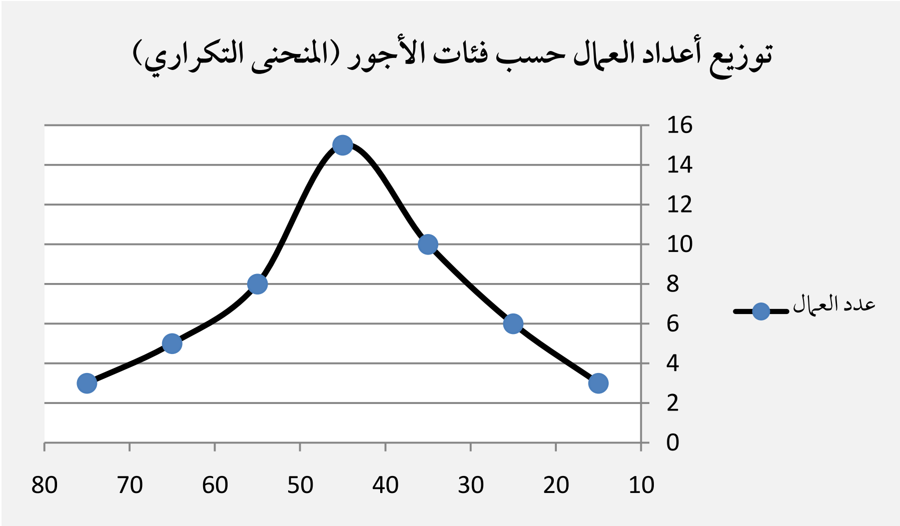
رسم توضيحي 2: توزيع أعداد الموظفين حسب الغياب (مثال 1.2)



رسم توضيحي 3: توزيع أعداد العمال حسب فئات الأجور (مثال 1.3)



رسم توضيحي 4: توزيع أعداد العمال حسب فئات الأجور (مثال 1.3)



رسم توضيحي 5: توزيع أعداد العمال حسب فئات الأجور (مثال 1.3)

1.3. الجداول التكرارية المتجمعة: الجدول التكراري البسيط يعطينا فقط عدد المفردات في كل فئة، ولكننا نحتاج إلى معرفة بيانات تجميعية لا نحصل عليها من الجداول البسيطة مباشرة، ولكن نحصل عليها من الجداول المتجمعة. وهي نوعان:

- **جدول متجمع صاعد**: وهو جدول ذو عمودين، عمود للفئات ويكتب "أقل من الحد الأعلى للفئة"، وعمود للتكرارات ويكتب "تكرار متجمع صاعد " (ت. م. ص.) ونحصل عليه بتجميع التكرارات في الجدول البسيط من بداية الجدول إلى نهايته.
- **(للإطلاع فقط) جدول متجمع نازل**: وهو جدول ذو عمودين، عمود للفئات ويكتب "الحد الأدنى للفئة فأكثر"، وعمود للتكرارات ويكتب "تكرار متجمع نازل" (ت. م. ن.) ونحصل عليه بتجميع التكرارات في الجدول البسيط من نهاية الجدول إلى بدايته.

مثال (1.5): كون الجدول المتجمع الصاعد والنازل لمثال أجور العمال (**1.3**)

الحل:

جدول متجمع نازل جدول متجمع صاعد جدول تكراري

ت. م. ن.	الحد الأدنى للفئة فأكثر	ت. م. ص.	أقل من الحد الأعلى للفئة	عدد العمال	الأجور
50	10 فأكثر	3	أقل من 20	3	10 -
47	20 فأكثر	9	أقل من 30	6	20 -
41	30 فأكثر	19	أقل من 40	10	30 -
31	40 فأكثر	34	أقل من 50	15	40 -
16	50 فأكثر	42	أقل من 60	8	50 -
8	60 فأكثر	47	أقل من 70	5	60 -
3	70 فأكثر	50	أقل من 80	3	70 - 80
				50	$\sum$

1.4. التمثيل البياني للتوزيع التكراري المتجمع

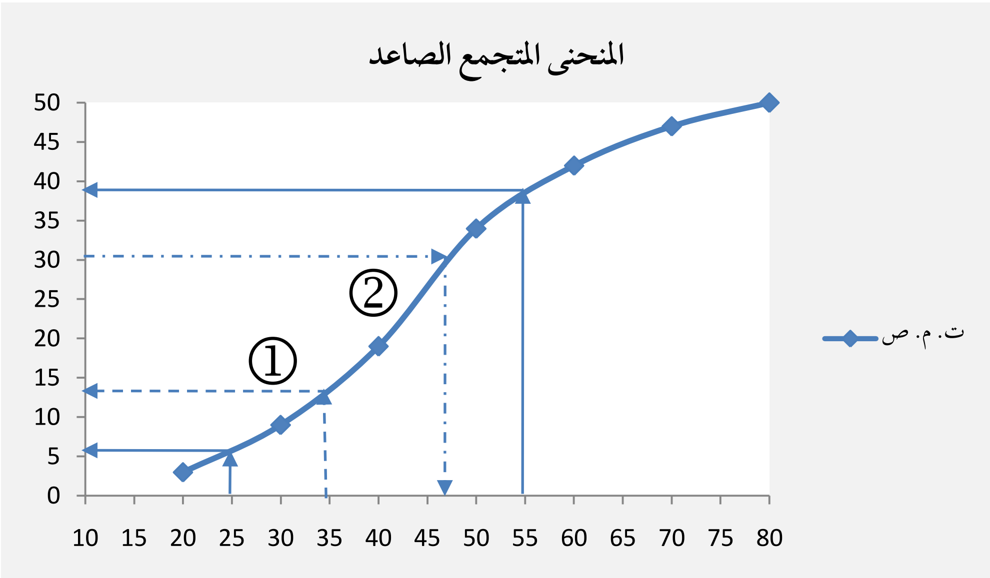
يتم تمثيل الجدول المتجمع الصاعد بيانيا بالمنحنى المتجمع الصاعد ، كذلك يتم تمثيل الجدول المتجمع النازل بيانياً بالمنحنى المتجمع النازل. وهذه المنحنيات تعطينا البيانات التي لا توفرها الجداول المتجمعة مباشرة، و لإنشاء المنحنى المتجمع، نتبع الخطوات التالية:

1. رسم محورين؛ الأفقي يمثل الحد الأعلى للفئة والرأسي يمثل ت. م. ص. في حالة التكرار المتجمع الصاعد، وفي حالة التكرار المتجمع النازل يمثل المحور الأفقي الحد الأدنى للفئة والرأسي يمثل ت. م. ن.
2. نحدد نقاط مناسبة حسب التكرار المتجمع، ثم نصلها يدوياً.

مثال (1.6): باستخدام الجدول المتجمع الصاعد في المثال (**1.5**)، ارسم المنحنى المتجمع الصاعد. ومن الرسم أوجد:

1. عدد العمال اللذين يقل أجورهم عن **35** ريال.
2. الحد الأعلى للأجور الذي بلغه **30** عامل.
3. نسبة عدد العمال اللذين تتراوح أجورهم اليومية بين **55** و**25** ريال.

الحل:



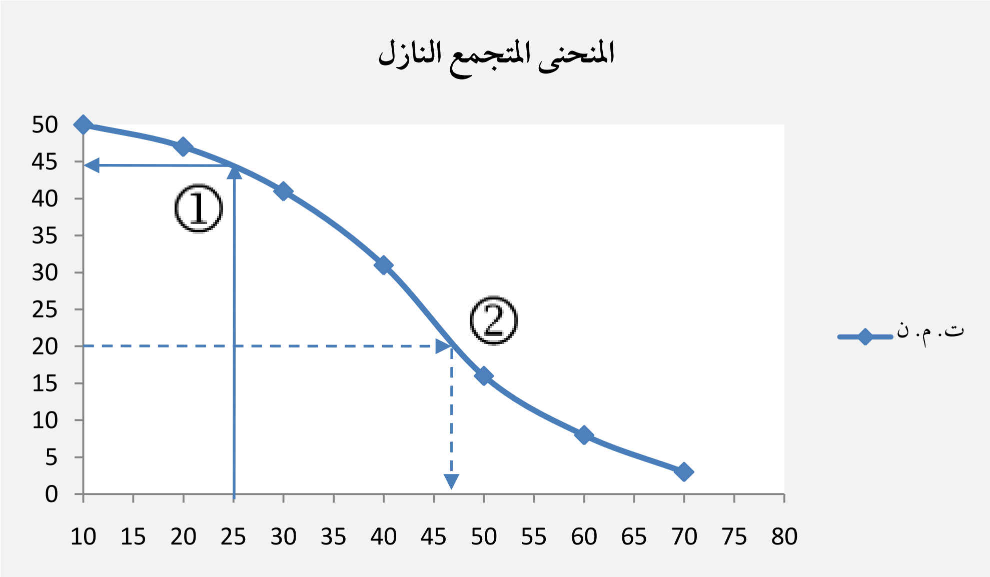
1. عدد العمال اللذين يقل أجورهم عن 35 ريال = 13 عامل
2. الحد الأعلى للأجور الذي بلغه 30 عامل = 46 ريال
3. نسبة عدد العمال اللذين تتراوح أجورهم اليومية بين 55 و25 ريال =

$$\{ (39 - 6) \div 50 \} \times 100 = 66\%$$

مثال (1.7) (للإطلاع فقط) باستخدام الجدول المتجمع النازل في المثال (1.5)، ارسم المنحنى المتجمع النازل. ومن الرسم أوجد:

1. **عدد العمال الذين تصل أجورهم 25 ريال فأكثر.**
2. **الحد الأدنى للأجور الذي بلغه 20 عامل.**

الحل:



1. عدد العمال الذين تصل أجورهم 25 ريال فأكثر = 44 عامل
2. الحد الأدنى للأجور الذي بلغه 20 عامل = 47 ريال

2.1. مقاييس النزعة المركزية

2.1. مقاييس النزعة المركزية : غالبا ما نلاحظ أن البيانات تميل إلى التركيز حول قيمة معينة. وفي هذه الحالة يمكن استخدام هذه "القيمة المركزية " لتمثيل هذه المجموعة من البيانات والمقاييس المستخدمة في التعرف على هذه القيمة المركزية تسمى "مقاييس النزعة المركزية" . ويجب أن يتوفر في هذه المقاييس الصفات الآتية لكي يكون المقياس جيداً:

- أن يعتمد المقياس في حسابه على كل المشاهدات.
- أن يكون المقياس سهل الحساب والفهم.
- أن يتوفر فيه القابلية للتعامل الجبري.
- ألا يتأثر بالقيم الشاذة أو المتطرفة.

من بين هذه المقاييس: **الوسط الحسابي (أو المتوسط) والوسيط والمنوال** ، ولا تتوفر كل الصفات السابق ذكرها في مقياس واح د، ولكن كل مقياس من هذه المقاييس يفضل استخدامه في حالات معينة ولا يفضل استخدامه في حالات أخرى، وفيما يلي مقارنة بين خصائص المتوسطات المذكورة:

الخاصية \ المقياس	المتوسط	الوسيط	المنوال
يعتمد المقياس في حسابه على كل المشاهدات	✓	✗	✗
المقياس سهل الحساب والفهم	✓	✓	✓
يتوفر فيه القابلية للتعامل الجبري	✓	✓	✓
لا يتأثر بالقيم الشاذة أو المتطرفة	✗	✓	✓

2.1.1. الوسط الحسابي (المتوسط)

تعريفه: يعرف الوسط الحسابي لمجموعة من القيم، بأنه مجموع قيم مفردات العينة مقسوماً على حجم العينة، ويرمز له بالرمز ($\bar{x}$).

طرق حسابه:

(أولاً) حالة البيانات الغير مبوبة: نستخدم القانون التالي:

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

حيث: $\sum x$: مجموع قيم مفردات العينة n : حجم العينة (عدد المفردات)

مثال (2.1): احسب الوسط الحسابي للأجور اليومية بالدولار للعينة التالية المكونة من خمس عمال بإحدى القطاعات: 50 70 80 90 60

الحل:

$$\bar{x} = \frac{\sum x}{n} = \frac{(50+70+80+90+60)}{5} = \frac{350}{5} = \boxed{70\$}$$

(ثانياً) حالة البيانات المبوبة: نتبع الخطوات التالية:

1. نضيف للجدول التكراري عمود يمثل مركز الفئة ونرمز له بالرمز x؛ ولحساب مراكز الفئات نتبع ما يلي:
   - **نحسب مركز الفئة الأولى = (الحد الأدنى للفئة الأولى + الحد الأدنى للفئة الثانية) ÷ 2**
   - **نحسب مركز الفئة الثانية = مركز الفئة الأولى + طول الفئة (h)**
   - **نحسب مركز الفئة الثالثة = مركز الفئة الثانية + h**، وهكذا حتى ننتهي من جميع الفئات.
2. نكون عمود آخر يحتوي على حاصل ضرب عمود التكرارات في عمود مركز الفئة.
3. نستخدم القانون التالي:

$$\bar{x} = \frac{\sum f \cdot x}{\sum f}$$

حيث: $\sum f \cdot x$: حاصل ضرب عمود التكرارات في عمود مركز الفئة

$\sum f$: مجموع التكرارات

مثال (2.2): استخدم الجدول التكراري الخاص بأجور العمال اليومية (مثال 1.3) لإيجاد ما يلي:

- **مراكز الفئات المختلفة**
- **حساب الوسط الحسابي (أو المتوسط)**

الحل: حيث أن طول الفئة h يساوي 10، فإن:

(1)	(2)	(3)	(4) = (2) × (3)
فئات الأجور (c)	عدد العمال (f)	مراكز الفئات (x)	$f \cdot x$
10 -	3	(10+20)÷2 = 15	45
20 -	6	15+h = 15+10 =25	150
30 -	10	35	350
40 -	15	45	675
50 -	8	55	440
60 -	5	65	325
70 - 80	3	75	225
$\sum$	**50** $\sum f$		**2210** $\sum f \cdot x$

$$\bar{x} = \frac{\sum f \cdot x}{\sum f} = \frac{2210}{50} = \boxed{44.2}$$ ريال

2.1.2. الوسيط

تعريفه: هو المفردة التي تقسم مفردات العينة بعد ترتيبها تصاعدياً (أو تنازلياً) إلى قسمين متساويين، بحيث يكون عدد المفردات الأصغر منها في القيمة مساوياً لعدد المفردات الأكبر منها في القيمة، ويرمز له بالرمز (m).

طرق حسابه:

(أولاً) حالة البيانات الغير مبوبة: نتبع الخطوات التالية:

1. نرتب مفردات العينة حسب قيمها إما تصاعدياً أو تنازلياً.
2. إذا كان حجم العينة فردي نستخدم تعريف الوسيط مباشرة، إما إذا كان حجم العينة زوجي فنقوم بإيجاد متوسط قيمة المفردتين التي تتوسط بقية المفردات.

مثال (2.3): **احسب وسيط الأجور اليومية بالدولار لعينتي العاملين بإحدى القطاعات التاليتين:**

العينة (1): 50 70 80 90 60

العينة (2): 50 70 80 90 60 100

الحل:

العينة (1): لحساب قيمة الوسيط، نرتب القيم تصاعديا فتصبح

50 60 70 80 90

من تعريف الوسيط، نجد أن قيمة الوسيط $\boldsymbol{m = 70\$}$

العينة (2): لحساب قيمة الوسيط، نرتب المفردات حسب قيمها تصاعديا فتصبح

50 60 70 80 90 100

نجد أن قيمة الوسيط

$$\boldsymbol{m = (70 + 80) \div 2 = 75\$}$$

(ثانياً) حالة البيانات المبوبة: نتبع الخطوات التالية:

1. نحدد ترتيب الوسيط والذي يحسب من العلاقة: $c_1 = \dfrac{\sum f}{2}$
2. نكون الجدول التكراري المتجمع الصاعد.
3. نحدد فئة الوسيط؛ بالبحث عن الفئتين التي تتراوح قيمة c_1 بين ت. م. ص. الخاص بهما، ثم اختيار الفئة ذات أكبر ت. م. ص. بينهما.
4. نستخدم العلاقة التالية:

$$m = L + \frac{c_1 - c_2}{c_3} \times h$$

حيث:

L : الحد الأدنى لفئة الوسيط

c_1: ترتيب الوسيط

c_2: ت. م. ص السابق لفئة الوسيط

c_3: التكرار الأصلي لفئة الوسيط

h : طول الفئة

مثال (2.4): احسب وسيط أجور العمال اليومية (مثال 1.3).

الحل:

الأجور	عدد العمال	أقل من الحد الأعلى للفئة	ت. م. ص.
10 -	3	أقل من 20	3
20 -	6	أقل من 30	9
30 -	10	أقل من 40	$c_2 = 19$
***L* = 40 -**	$c_3 = 15$	أقل من 50	34
50 -	8	أقل من 60	42
60 -	5	أقل من 70	47
70 - 80	3	أقل من 80	50

$c_1 = 25$

$$m = L + \frac{c_1 - c_2}{c_3} \times h = 40 + \frac{25-19}{15} \times 10 = \boxed{44}$$ ريال

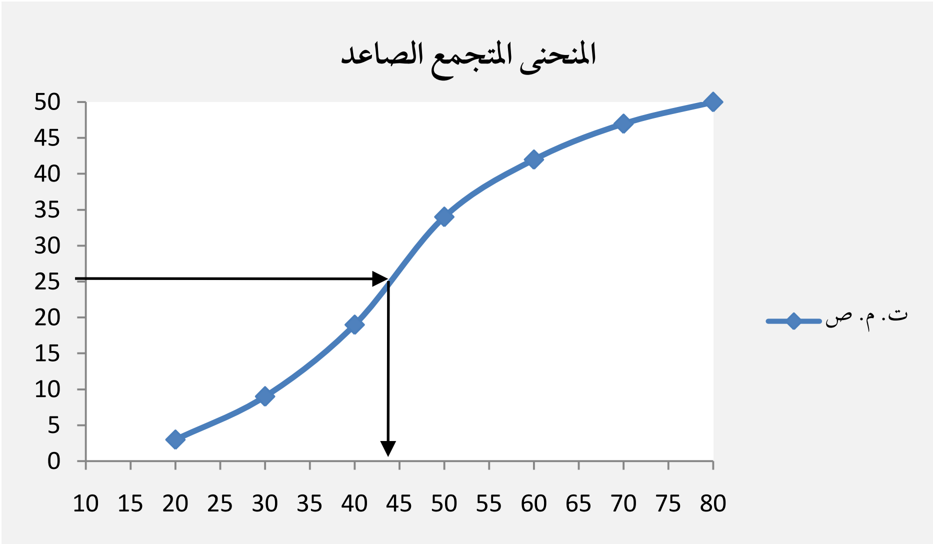
(ثالثاً) حالة البيانات المبوبة (طريقة الرسم): نتبع الخطوات التالية:

1. نحدد ترتيب الوسيط والتي تحسب من العلاقة: $c_1 = \frac{\Sigma f}{2}$
2. نرسم خط أفقي حتى يمس المنحنى المتجمع الصاعد ثم نسقط عمود على المحور الأفقي لنحصل على الوسيط مباشرة.

> **ملاحظة:** قيمة الوسيط بالرسم قد تختلف قليلا عن قيمته بالحساب. والمهم هو ألا تخرج قيمة الوسيط عن فئة الوسيط، أي لا تقل عن الحد الأدنى لفئة الوسيط ولا تزيد عن الحد الأعلى لها وفي مثالنا نجد أن قيمة الوسيط تنحصر بين 40 و50.

مثال (2.5): احسب وسيط أجور العمال اليومية (مثال **1.3**) بالرسم.

الحل: نعلم أن $c_1 = 25$، ومن الشكل التالي نجد أن $m \approx 44$.



2.1.3. المنوال

تعريفه: هو قيمة المفردة التي تتكرر أكثر من غيرها، أو هو المفردة ذات القيمة الأكثر **شيوعا** أو **تكراراً**، ويرمز له بالرمز (D).

ملاحظة: المنوال يمكن استخدامه كذلك في حالة البيانات الوصفية (الغير مبوبة فقط).

طرق حسابه:

(**أولاً**) **حالة البيانات الغير مبوبة**: باستخدام التعريف مباشرة.

مثال (2.6): احسب الأجر الشائع لعينتي العاملين بإحدى القطاعات التاليتين:

العينة (1):	**50**	**70**	**50**	**90**	**60**	
العينة (2):	**50**	**70**	**70**	**50**	**60**	**70**

الحل:

العينة (1): من تعريف المنوال، نجد أن $D = 50$

العينة (2): من تعريف المنوال، نجد أن $D = 70$

(ثانياً) حالة البيانات المبوبة: نتبع الخطوات التالية:

1. نحدد فئة المنوال؛ وهي الفئة ذات أكبر تكرار.
2. نستخدم العلاقة التالية:

$$D = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h$$

L : الحد الأدنى لفئة المنوال

Δ_1 : الفرق بين أكبر تكرار والسابق له

Δ_2 : الفرق بين اكبر تكرار واللاحق له

h : طول الفئة

مثال (2.7): احسب أجر العمال الشائع (مثال 1.3).

الحل:

الأجور (c)	عدد العمال (f)
10 -	3
20 -	6
30 -	**10**
***L* = 40 -**	**15**
50 -	**8**
60 -	5
70 - 80	3

$$D = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 40 + \frac{(15-10)}{(15-10)+(15-8)} \times 10 = 40 + \frac{5}{5+7} \times 10 = \boxed{44.17}$$ ريال

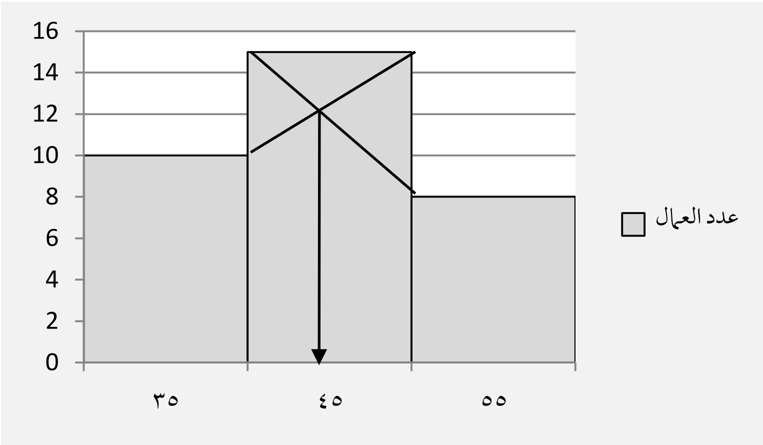
(ثالثاً) حالة البيانات المبوبة (طريقة الرسم): نتبع الخطوات التالية:

1. نرسم جزء من المدرج التكراري؛ مستطيل يمثل فئة المنوال (الفئة التي تقابل أكبر تكرار) ومستطيل يمثل الفئة السابقة وآخر يمثل الفئة اللاحقة لها.

2. نصل رؤوس المستطيلات ببعضها فتتقابل في نقطة نسقط منها عمود على المحور الأفقي، فتكون هي قيمة المنوال.

مثال (2.8): **احسب الأجر الشائع للعمال (مثال 1.3) بالرسم.**

الحل: $D \approx 44$



2.1.4. ملاحظات هامة على المتوسطات

1. درسنا الوسط الحسابي والوسيط والمنوال . نجد أن الوسط الحسابي أدق المتوسطات الثلاثة لأنه يأخذ في الاعتبار جميع فئات التوزيع، يليه الوسيط، وأخيرا المنوال لأن طرق حسابه تقريبية، وعلى ذلك إذا ذكر لفظ المتوسط دون تحديد فيقصد به الوسط الحسابي.
2. ذكر لفـظ "الشائع"، مثل الطـول الشائع ، العمر الشائع ......الخ، يقصد به المنوال.
3. عادة ما تكون قيم المتوسطات قريبة من بعضها البعض.

2.2. مقاييس التشتت

التشتت: يقصد بالتشتت دراسة مدى تقارب أو تباعد البيانات عن بعضها البعض أي عن وسطها الحسابي. فكلما كانت البيانات قريبة من بعضها البعض أي قريبة من الوسط الحسابي تكون **البيانات متجانسة**، والعكس كلما كانت البيانات بعيدة عن بعضها أي بعيدة عن الوسط الحسابي تكون **البيانات متباعدة أو مشتتة**.

يقاس التشتت بعدة مقاييس، منها: **المدى**[1]، **الانحراف المتوسط، الانحراف المعياري، الانحراف الربيعي (نصف المدى الربيعي)، معامل الاختلاف**. وسنكتفي بالانحراف المعياري، ومعامل الاختلاف.

2.2.1. الانحراف المعياري

تعريفه: هو أدق مقاييس التشتت وأكثرها استخداما، ويعرف بأنه الجذر التربيعي لمتوسط مربعات انحرافات (فروق) القيم عن وسطها الحسابي، ويرمز له بالرمز (s_x).

ملاحظات هامة:

1. متوسط مربعات انحرافات (فروق) القيم عن وسطها الحسابي يسمى **التباين**، أي أن الانحراف المعياري هو جذر التباين.
2. قيمة **التباين لا بد أن تكون موجبة أو تساوي الصفر**.
3. كلما اقتربت قيمة التباين من الصفر، أي كلما اقتربت قيمة الانحراف المعياري من الصفر كلما أصبحت البيانات قريبة من التجانس.
4. يتأثر الانحراف المعياري بالقيم الشاذة.

طرق حسابه:

(أولاً) حالة البيانات الغير مبوبة: نستخدم القانون التالي:

$$s_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

حيث: $\sum x^2$: مجموع مربعات قيم مفردات العينة n : حجم العينة

مثال (2.9): احسب الانحراف المعياري للأجور اليومية بالدولار للعينة التالية المكونة من خمس عمال بإحدى القطاعات: 50 70 80 90 60

الحل:

$$\bar{x} = \frac{350}{5} = \boxed{70\$}$$

$$\frac{\sum x^2}{n} = \frac{60^2 + 90^2 + 80^2 + 70^2 + 50^2}{5} = 5100$$

[1] سبق ذكره في الباب الأول.

$$s_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{5100 - 4900} = \sqrt{200} = \boxed{14.14\$}$$

(ثانياً) حالة البيانات المبوبة: نتبع الخطوات التالية:

1. نتبع خطوات حساب المتوسط الحسابي.
2. نضيف عمود جديد هو حاصل ضرب عمود مركز الفئة (عمود 3) في عمود حاصل ضرب عمود التكرارات في عمود مركز الفئة (عمود 4).
3. نستخدم القانون التالي:

$$s_x = \sqrt{\frac{\sum f \cdot x^2}{\sum f} - \bar{x}^2}$$

مثال (2.10): اوجد الانحراف المعياري لأجور العمال (مثال 1.3).

الحل: من مثال (2.2) نعلم أن:

$$\bar{x} = \frac{\sum f \cdot x}{\sum f} = \frac{2210}{50} = \boxed{44.2}$$

(1)	(2)	(3)	(4)	(5) = (3) × (4)
الأجور (c)	عدد العمال (f)	x	$f \cdot x$	$f \cdot x^2$
10 -	3	15	45	675
20 -	6	25	150	3750
30 -	10	35	350	12250
40 -	15	45	675	30375
50 -	8	55	440	24200
60 -	5	65	325	21125
70 - 80	3	75	225	16875
$\sum$	**50** $=\sum f$		**2210** $=\sum f \cdot x$	**109250** $=\sum f \cdot x^2$

$$s_x = \sqrt{\frac{\sum f \cdot x^2}{\sum f} - \bar{x}^2} = \sqrt{\frac{109250}{50} - 44.2^2} = \sqrt{2185 - 1953.64} = \sqrt{231.36} = \boxed{15.21} \text{ ريال}$$

2.2.2. معامل الاختلاف

لمقارنة تشتت مجموعتين أو أكثر من البيانات ، فإننا لا نستخدم مقاييس التشتت السابق ذكرها مباشرة ، وذلك بسبب:

- قد تكون البيانات ذات وحدات قياس مختلفة.
- قد تكون الأوساط الحسابية للظاهرتين مختلفة، لذلك نستخدم ما يسمى بمقاييس التشتت النسبي، وأهم هذه المقاييس هو ما يعرف بمعامل الاختلاف.

تعريف معامل الاختلاف: هو معامل نسبي يستخدم للمقارنة بين تشتت ظاهرتين أو أكثر مختلفتين أو حتى متشابهتين في وحدة القياس. والظاهرة التي معامل اختلافها أكبر تكون أكثر تشتتاً من الأخرى. ويستخدم في حسابه العلاقة الآتية:

$$c.v.(x)=\frac{s_x}{\bar{x}}\times 100$$

مثال (2.11): اوجد معامل الاختلاف لأجور العمال (مثال 1.3).

الحل: من مثال (2.2) و(2.10) نعلم أن:

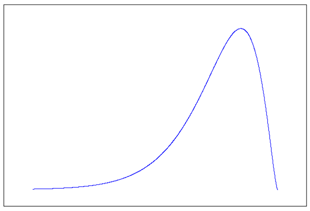
$$\bar{x}=44.2$$
$$s_x=15.21$$
$$c.v.(x)=\frac{s_x}{\bar{x}}\times 100=\frac{15.21}{44.2}\times 100=34.41\%$$

2.3. الالتواء

الالتواء: هو بعد المنحنى التكراري عن التماثل. ويقصد بالتماثل أنه إذا أسقطنا عموداً من قمة المنحنى التكراري وقسمه إلى قسمين منطبقين يكـون التوزيع متماثلاً. والعكس فيكون التوزيع غير متماثل أي ملتو إما إلى جهة اليمين أو إلى جهة اليسار.

ويقاس الالتواء بعدة مقاييس، منها:

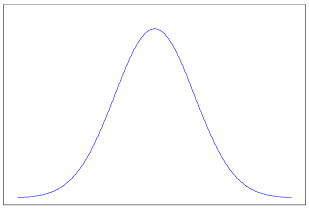
$$s.k.(I)=\frac{\bar{x}-D}{s_x}\qquad\qquad s.k.(II)=\frac{3(\bar{x}-m)}{s_x}$$



التوزيع غير متماثل

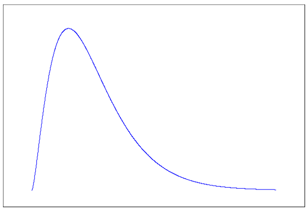
وملتو من جهة اليسار

$\bar{x} \neq m \neq D$



التوزيع متماثل

$\bar{x} = m = D$



التوزيع غير متماثل

وملتو من جهة اليمين

$\bar{x} \neq m \neq D$

ودلالة قيمة أي من المعاملين على النحو التالي:

$$s.k. \begin{cases} < 0 \quad (-) & \text{التوزيع ملتو لليسار} \\ = 0 & \text{التوزيع متماثل} \\ > 0 \quad (+) & \text{التوزيع ملتو لليمين} \end{cases}$$

مثال (2.12): ادرس التواء التوزيع التكراري لأجور العمال (مثال 1.3)، باستخدام:

الحل:

$$s.k.(I) = \frac{\bar{x} - D}{s_x} = \frac{44.2 - 44.17}{15.21} = 0.002$$

$$s.k.(II) = \frac{3(\bar{x} - m)}{s_x} = \frac{3(44.2 - 44)}{15.21} = 0.04$$

ملاحظات:

1. عادة ما نحصل على نتائج مختلفة لمعاملي الالتواء، وهذا لا يناقض بعضه إذ أن كل معامل يقيس الالتواء على أساس يخالف المعاملات الأخرى.
2. عند مقارنة التواء توزيعات مختلفة يجب استخدام نفس المعامل عند المقارنة.

2.4. مسائل محلولة

2.4.1. الجدول التالي يمثل التوزيع التكراري لفئات الدخل الشهري لعينة من الأسر (مئات الريالات) بإحدى المدن:

فئات الدخل	62 -	66 -	70 -	74 -	78 -	82 -	86 - 90
عدد الأسر	3	8	20	21	14	10	4

المطلوب:

1. حساب الوسط الحسابي، والانحراف المعياري.
2. حساب الوسيط والمنوال.
3. دراسة إلتواء التوزيع التكراري.

الحل:

فئات الدخل	عدد الأسر	x	$f\,x$	$f\,x^2$
62 -	3	64	192	12288
66 -	8	68	544	36992
70 -	20	72	1440	103680
74 -	21	76	1596	121296
78 -	14	80	1120	89600
82 -	10	84	840	70560
86 - 90	4	88	352	30976
$\sum$	**80**		**6084**	**465392**

1.

$$\bar{x} = \frac{\sum f \cdot x}{\sum f} = \frac{6084}{80} = \boxed{76.05} \quad \text{مئات الريالات}$$

$$s_x = \sqrt{\frac{\sum f \cdot x^2}{\sum f} - \bar{x}^2} = \sqrt{\frac{465392}{80} - 76.05^2} = \sqrt{33.8} = \boxed{5.8} \quad \text{مئات الريالات}$$

2.

فئات الدخل	عدد الأسر	أقل من الحد الأعلى للفئة	ت. م. ص.
62 -	3	أقل من 66	3
66 -	8	أقل من 70	11
70 -	**20**	أقل من 74	**31**
74 -	**21**	أقل من 78	52
78 -	**14**	أقل من 82	66
82 -	10	أقل من 86	76
86 - 90	4	أقل من 90	80
$\sum$	**80**		

$$m = L + \frac{c_1 - c_2}{c_3} \times h = 74 + \frac{40-31}{21} \times 4 = \boxed{75.71}$$ مئات الريالات

$$D = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 74 + \frac{1}{1+7} \times 4 = 74 + \frac{1}{8} \times 4 = \boxed{74.5}$$ مئات الريالات

3. نستخدم المنوال لدراسة الالتواء:

$$s.k.(I) = \frac{\bar{x} - D}{s_x} = \frac{76.05 - 74.5}{5.8} = 0.27$$

التوزيع غير متماثل وملتو لليمين.

2.4.2. إذا كانت y تمثل الاستهلاك الأسبوعي الشخصي بالريال، x تمثل الاستهلاك الأسبوعي لوقود السيارات (باللتر)، استخدم المعلومات التالية لتحديد أي الظاهرتين أكثر تشتتاً.

$\bar{y} = 75$ ريال , $s_y = 15$ ريال

- **التوزيع التكراري للاستهلاك الأسبوعي لوقود السيارات:**

فئات الاستهلاك	18 -	28 -	38 -	48 -	58 -
عدد المستهلكين	**26**	**59**	**105**	**79**	**31**

الحل: نقوم بحساب معامل الاختلاف للمتغير (أو الظاهرة) x ومن ثم نحسب معامل الاختلاف للمتغير (أو الظاهرة) y ثم نقارن بينهما.

فئات الاستهلاك	عدد السيارات	**x**	**$f\,x$**	**$f\,x^2$**
18 -	26	23	598	13754
28 -	59	33	1947	64251
38 -	105	43	4515	194145
48 -	79	53	4187	221911
58 - 68	31	63	1953	123039
$\sum$	**300**		**13200**	**617100**

$$\bar{x}=\frac{\sum f\cdot x}{\sum f}=\frac{13200}{300}=\boxed{44}$$ لتر

$$s_x=\sqrt{\frac{\sum f\cdot x^2}{\sum f}-\bar{x}^2}=\sqrt{\frac{617100}{300}-44^2}=\sqrt{121}=\boxed{11}$$ لتر

$$\therefore c.v.(x)=\frac{s_x}{\bar{x}}\times 100=\frac{11}{44}\times 100=25\%$$

كما أن:

$$\therefore c.v.(x)=\frac{s_y}{\bar{y}}\times 100=\frac{15}{75}\times 100=20\%$$

لذا يمكن القول أن ظاهرة الاستهلاك الأسبوعي لوقود السيارات أكثر تشتتاً من ظاهرة الاستهلاك الأسبوعي الشخصي.

2.5. تمارين

2.5.1. الجدول الآتي يوضح توزيع عينة من (100) موظف حسب فئات الزيادة التي حصلوا عليها في الراتب (بعشرات الريالات):

فئات الزيادة	30 -	40 -	50 -	60 -	70 -	80 -	90 -
عدد الموظفين	4	11	20	36	17	8	4

1. ارسم المدرج التكراري للتوزيع، ثم حدد منه قيمة المنوال.
2. ارسم المنحنى المتجمع الصاعد ومن الرسم أوجد:

أ. عدد الموظفين الذين حصلوا على الزيادة أقل من 750 ريال.

ب. الحد الأعلى الزيادة التي حصل عليها 30 موظف.

ج. وسيط الزيادة في الراتب.

2.5.2. الجدول الآتي يوضح توزيع عينة من (200) موظف بإحدى الوزارات حسب أعمارهم بالسنة.

فئات العمر	20 -	25 -	30 -	35 -	40 -	45 -	50 -	55 -
عدد الموظفين	10	17	24	43	34	30	23	19

1. احسب الوسط الحسابي والانحراف المعياري.
2. ارسم المنحنى المجتمع الصاعد ومن الرسم أوجد:

أ. نسبة عدد المشتغلين الذين يقل عمرهم عن 42 سنه.

ب. الحد الأعلى للعمر الذي بلغه 120 موظف.

ج. وسيط العمر.

3. ارسم المنحنى التكراري، هل المنحنى متماثل؟ دلل على إجابتك بحساب معامل الالتواء.

2.5.3. الجدول الآتي يوضح توزيع عينة من الموظفين بإحدى شركات القطاع الخاص حسب فئات الراتب (آلاف الريالات).

فئات الراتب	62 -	66 -	70 -	74 -	78 -	82 -	86 -
عدد الموظفين	3	8	20	21	14	10	4

1. احسب كل من الوسط الحسابي والانحراف المعياري.
2. الوسيط (بالحساب و بالرسم).
3. المنوال (بالحساب و بالرسم).
4. معامل الاختلاف.
5. معاملي الالتواء.

2.5.4. الجدول الآتي يوضح توزيع الدخل السنوي (بآلاف الريالات) لعينة من الأسر بإحدى المدن:

فئات الدخل	40-	48-	56-	64-	72-	80-	88-
عدد الأسر	8	16	24	36	30	18	8

1. الانحراف المعياري للتوزيع.
2. الوسيط.
3. الدخل الشائع لهذه الأسر.
4. معامل الاختلاف.
5. هل هذا التوزيع متماثل؟ علك إجابتك.

2.5.5. الجدول الآتي يوضح توزيع عينة من **(500)** موظف حسب الساعات الإضافية التي حققها أسبوعياً بإحدى الشركات:

الساعات الإضافية	2 -	3 -	4 -	5 -	6 -	7 -	8 -	9 -
عدد الموظفين	98	118	101	95	50	20	10	8

احسب المتوسط والتباين لعدد الساعات الإضافية، ثم احسب معامل الاختلاف.

2.5.6. البيانات التالية تمثل أسعار منتجات معينة (بالريال) مختارة من ثلاثة مراكز تجارية مختلفة:

عينة (1):	7	2	1	3	6	5
عينة (2):	8	4	2	1	3	6
عينة (3):	1	4	2	5	1	3

احسب معامل الاختلاف لكل عينة، أي العينات أكثر تشتتاً ؟

2.5.7. الجدول الآتي يوضح توزيع عينة من المرضى بمرض معين بإحدى المستشفيات حسب الساعات التي قضوها حتى تماثلوا للشفاء:

الساعات	15 -	19 -	23 -	27 -	31 -
عدد المرضى	6	14	42	10	8

1. **ارسم المدرج و المضلع التكراري للتوزيع.**
2. **احسب الوسط الحسابي والانحراف المعياري للتوزيع.**
3. **أوجد منوال التوزيع.**
4. **احسب معامل الاختلاف.**
5. **هل هذا التوزيع متماثل؟ علل إجابتك.**

2.5.8. الجدول الآتي يوضح توزيع أسعار عينة من العقارات بمدينة جدة حسب سعرها (بمئات الآلاف ريال)

فئات الأسعار	15 -	17 -	19 -	21 -	23 -
عدد العقارات	3	7	21	5	4

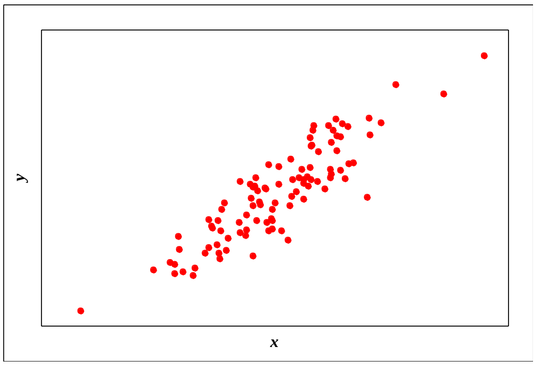
1. **احسب الانحراف المعياري للتوزيع.**
2. **احسب وسيط التوزيع، واستخدمه في دراسة تماثل التوزيع.**

إذا كان الوسط الحسابي والانحراف المعياري لمساحة نفس العقارات هو 600، 150 (متر مربع) على التوالي. فأيهما أكثر تشتتاً؟ السعر أم المساحة.

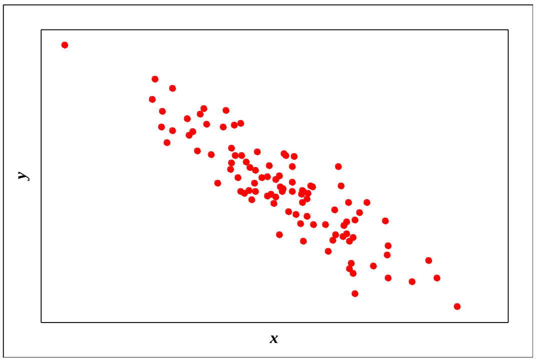
3.1. الارتباط

الارتباط: هو علاقة بين متغيرين (x, y) بمعنى أنه إذا تغير أحد المتغيرين فإن الآخر قد يتبعه. إما في نفس الاتجاه فيكون الارتباط طردي، كعلاقة الصادرات بالميزان التجاري **(رسم توضيحي 6)**. أو في الاتجاه المضاد فيكون الارتباط عكسي، كعلاقة الواردات بالميزان التجاري **(رسم توضيحي 7)**. ويقال أن المتغيرين مستقلتين عندما ينعدم الارتباط، كعلاقة دخل الفرد بوزنه **(رسم توضيحي 8)**.

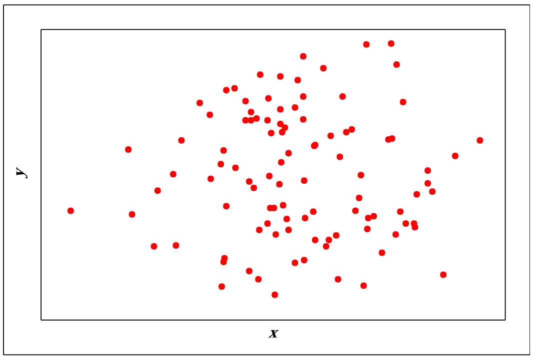
ملاحظة: الرسوم التوضيحية 6، 7، 8 تسمى أشكال الانتشار.



رسم توضيحي 6: الارتباط الطردي



رسم توضيحي 7: الارتباط العكسي



رسم توضيحي 8: لا يوجد ارتباط (استقلال)

ملاحظة: الارتباط لا يدل على السببية، حيث ليس شرطاً أن يتغير أحد المتغيرين دائماً بتغير أحدهما.

معامل الارتباط: والذي يرمز له بالرمز r، عبارة عن مقياس كمي نسبي يقيس قوة الارتباط بين متغيرين، حيث تتراوح قيمته بين 1 و -1، أي أن $-1 \le r \le 1$. يقال أن الارتباط **طردي تام** إذا كان معامل الارتباط $r = +1$، ويقال أن الارتباط **عكسي تام** إذا كان معامل الارتباط $r = -1$. يقال عن المتغيرين أنهما **مستقلان** إذا كان $r = 0$.

> **ملاحظة:** كلما اقتربت قيمة معامل الارتباط إلى 1 كلما كان الارتباط الطردي قوياً بين الظاهرتين (أو المتغيرين) وكلما اقتربت قيمته إلى الصفر كلما كان الارتباط الطردي ضعيفاً. ونفس القول ينطبق على الارتباط العكسي.

3.1.1. معامل الارتباط الخطي

يقيس العلاقة بين متغيرين كميين، ويسمى أيضاً **بمعامل ارتباط بيرسون**، ويحسب من العلاقة التالية:

$$r_p = \frac{\dfrac{\sum xy}{n} - \bar{x}\ \bar{y}}{s_x\ s_y}$$

حيث:

$\sum xy$: مجموع حاصل ضرب x في y

$\bar{x}$: متوسط المتغير (أو الظاهرة) x

$\bar{y}$: متوسط المتغير (أو الظاهرة) y

s_x : الانحراف المعياري للمتغير (أو الظاهرة) x

s_y : الانحراف المعياري للمتغير (أو الظاهرة) y

> **ملاحظة:** قيمة معامل الارتباط الخطي التي تساوي الصفر تدل على عدم وجود **علاقة ارتباط خطية فقط** بين المتغيرين محل الدراسة.

مثال (3.1): لدراسة علاقة الصادرات بالميزان التجاري خلال عدة سنوات، أخذنا عشر قراءات تقريبية لقيمة صادرات المملكة العربية السعودية (x) وقيمة الميزان التجاري (y) بعشرات المليارات ريال كما يلي:

y	1	3	8	7	6	5	7	8	12	12
x	9	11	17	18	19	16	16	19	23	23

مصدر البيانات الأصلية: موقع مصلحة الإحصاءات العامة على شبكة الإنترنت

هل توجد علاقة ارتباط خطية؟ ما نوعها وما مدى قوتها؟

الحل:

	x	y	xy	x^2	y^2
	9	1	9	81	1
	11	3	33	121	9
	17	8	136	289	64
	18	7	126	324	49
	19	6	114	361	36
	16	5	80	256	25
	16	7	112	256	49
	19	8	152	361	64
	23	12	276	529	144
	23	12	276	529	144
$\sum$	171	69	1314	3107	585
	$=\sum x$	$=\sum y$	$=\sum xy$	$=\sum x^2$	$=\sum y^2$

$$\bar{x}=\frac{\sum x}{n}=\frac{171}{10}=17.1 \quad , \quad \bar{y}=\frac{\sum y}{n}=\frac{69}{10}=6.9$$

$$s_x=\sqrt{\frac{\sum x^2}{n}-\bar{x}^2}=\sqrt{\frac{3107}{10}-(17.1)^2}=\sqrt{310.7-292.41}=\sqrt{18.29}$$

$$s_y=\sqrt{\frac{\sum y^2}{n}-\bar{y}^2}=\sqrt{\frac{585}{10}-(6.9)^2}=\sqrt{58.5-47.61}=\sqrt{10.89}$$

$$r_p=\frac{\frac{\sum xy}{n}-\bar{x}\;\bar{y}}{s_x\,s_y}=\frac{\frac{1314}{10}-(17.1)(6.9)}{\sqrt{(18.29)(10.89)}}=\frac{131.4-117.99}{14.11}=\frac{13.41}{14.11}\approx\boxed{0.95}$$

من الملاحظ أن علاقة الارتباط الخطي بين قيمة صادرات المملكة العربية السعودية وقيمة الميزان التجاري موجودة وهي علاقة ارتباط طردية قوية.

مثال (3.2): سُجلت ست قراءات تقريبية لحجم الإنتاج وحجم صادرات النفط الخام بالمملكة العربية السعودية (بالمليار برميل) خلال عدة سنوات كما يلي:

حجم الصادرات (y)	2	2	2	1	1	1
حجم الإنتاج (x)	3	4	2	2	2	2

مصدر البيانات الأصلية: موقع وزارة البترول والثروة المعدنية على شبكة الإنترنت

ادرس وجود علاقة ارتباط خطية بين حجم الإنتاج وحجم صادرات النفط الخام.

الحل:

	x	y	xy	x^2	y^2
	3	2	6	9	4
	4	2	8	16	4
	2	2	4	4	4
	2	1	2	4	1
	2	1	2	4	1
	2	1	2	4	1
$\sum$	15	9	24	41	15
	$=\sum x$	$=\sum y$	$=\sum xy$	$=\sum x^2$	$=\sum y^2$

$$\bar{x}=\frac{\sum x}{n}=\frac{15}{6}=2.5 \quad , \quad \bar{y}=\frac{\sum y}{n}=\frac{9}{6}=1.5$$

$$s_x=\sqrt{\frac{\sum x^2}{n}-\bar{x}^2}=\sqrt{\frac{41}{6}-(2.5)^2}=\sqrt{6.83-6.25}=\sqrt{0.58}$$

$$s_y=\sqrt{\frac{\sum y^2}{n}-\bar{y}^2}=\sqrt{\frac{15}{6}-(1.5)^2}=\sqrt{2.5-2.25}=\sqrt{0.25}$$

$$r_p=\frac{\frac{\sum xy}{n}-\bar{x}\,\bar{y}}{s_x\,s_y}=\frac{\frac{24}{6}-(1.5)(2.5)}{\sqrt{(0.58)(0.25)}}=\frac{4-3.75}{\sqrt{0.145}}=\frac{0.25}{0.38}\approx\boxed{0.66}$$

من الملاحظ أن علاقة الارتباط الخطي بين حجم الإنتاج وحجم صادرات النفط الخام علاقة طردية متوسطة.

3.1.2. معامل ارتباط الرتب

يسمى أيضاً **بمعامل ارتباط سبيرمان** ، ويستخدم هذا المعامل عندما يكون كلا المتغيرين ترتيبيين، أو أحدهم كمي والآخر ترتيبي، ويحسب من العلاقة التالية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث:

d الفرق بين رتب (ترتيب) x، ورتب (ترتيب) y، $\sum \boldsymbol{d} = \mathbf{0}$

مثال (3.3): لدراسة علاقة ارتباط تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات، اخترنا ثمان طلاب وكانت تقديراتهم كما يلي:

D	**C**	**B**	**C**	**D**	**C**	**A**	**F**	**تقديرات الإحصاء (x)**
F	**D**	**A**	**D**	**F**	**B**	**C**	**D**	**تقديرات الرياضيات (y)**

هل توجد علاقة ارتباط؟ ما نوعها ومدى قوتها؟

الحل:

x	y	رتب x	رتب y	d	d^2
F	D	2.5	1.5	1	1
D	C	5	4	1	1
A	B	7	8	-1	1
D	C	5	4	1	1
F	D	2.5	1.5	1	1
B	C	5	7	-2	4
C	A	8	6	2	4
D	F	1	4	-3	9
			$\sum$	0	**22**
				$\sum d$	$\sum d^2$

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{(6)(22)}{8(64-1)} = 1 - \frac{132}{504} = 1 - 0.26 = 0.74$$

نلاحظ وجود علاقة ارتباط طردية متوسطة بين تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات.

مثال (3.4): في دراسة لمعرفة العلاقة بين عدد الحقول المكتشفة وطول الأنابيب (بالكيلومتر) الناقلة للنفط الخام بالمملكة العربية السعودية خلال عدة سنوات، سجلت سبع قراءات على النحو التالي:

عدد الحقول (x)	**67**	**63**	**62**	**61**	**56**	**54**	**54**
طول الأنابيب (y)	**23120**	**23120**	**23020**	**23008**	**23006**	**22027**	**21960**

مصدر البيانات: موقع وزارة البترول والثروة المعدنية على شبكة الإنترنت

هل توجد علاقة ارتباط بين عدد الحقول وطول الأنابيب؟

الحل:

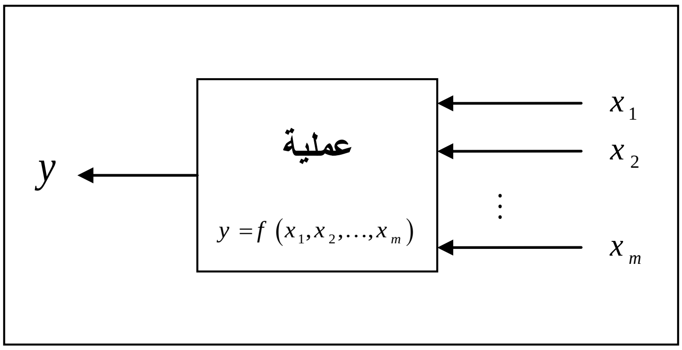
x	y	رتب x	رتب y	d	d^2
54	21960	1.5	1	0.5	0.25
54	22027	1.5	2	-0.5	0.25
56	23006	3	3	0	0
61	23008	4	4	0	0
62	23020	5	5	0	0
63	23120	6	6.5	-0.5	0.25
67	23120	7	6.5	0.5	0.25
			$\sum$	0.0	**1**
				$\sum d$	$\sum d^2$

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{(6)(1)}{7(49-1)} = 1 - \frac{6}{336} = 1 - 0.02 = 0.98$$

نلاحظ وجود علاقة ارتباط طردية قوية بين عدد الحقول المكتشفة وطول الأنابيب الناقلة للنفط الخام.

3.2. تحليل الانحدار الخطي البسيط

تحليل الانحدار: عبارة عن أسلوب إحصائي يقوم بصياغة دالة رياضية لعملية ذات عوامل مؤثرة عدة $x_1,x_2,\ldots,x_m$ لوصف متغير ناتج y من هذه العملية والتحكم به وتوقع قيم غير معروفة له. ويكون شكل الدالة الرياضية تكون على الصورة $y=f(x_1,x_2,\ldots,x_m)$



تسمى المتغيرات $x_1,x_2,\ldots,x_m$ **بالمتغيرات المستقلة**، والمتغير الناتج y **بالمتغير التابع**. تُعرف الدالة السابقة بدالة الانحدار، وأبسط حالة لهذه الدالة عندما يكون للعملية متغير مستقل واحد فقط يرتبط مع المتغير التابع بعلاقة خط مستقيم، أي أن: $(*)$ $\hat{y}=b_0+b_1x$

حيث:

b_0: الجزء المقطوع من محور y $\qquad$ b_1: ميل الخط المستقيم أو معامل انحدار x على y (أو y/x)

وتحسب القيمتين b_0 و b_1 من العلاقتين التاليتين:

$$b_1=\frac{\dfrac{\sum xy}{n}-\bar{x}\,\bar{y}}{s_x^2} \qquad\qquad b_0=\bar{y}-b_1\bar{x}$$

ولإيجاد أي قيمة مقدرة جديدة $\hat{y}_h$، نعوض بقيمة معلومة للمتغير المستقل ولتكن x_h في المعادلة $(*)$.

ملاحظة: قيمة معامل الانحدار تدل على نوع الارتباط. (ألا تظن أن هناك علاقة بين معامل الانحدار ومعامل الارتباط الخطي؟)

مثال (3.5): لدراسة علاقة الاستهلاك المحلي (y) بالإنتاج (x) لمادة الإسفلت (بالمليون برميل) خلال عدة سنوات، أخذنا عشر قراءات تقريبية كما يلي:

y	6	8	9	8	7	6	5	6	5	5
x	10	13	15	14	9	7	6	6	5	5

مصدر البيانات الأصلية: موقع وزارة البترول والثروة المعدنية على شبكة الإنترنت

أوجد معادلة الانحدار الخطي البسيط، وتوقع قيمة الاستهلاك المحلي عندما يصل إنتاج 16000000 برميل.

الحل:

	x	y	xy	x^2
	10	6	60	100
	13	8	104	169
	15	9	135	225
	14	8	112	196
	9	7	63	81
	7	6	42	49
	6	5	30	36
	6	6	36	36
	5	5	25	25
	5	5	25	25
$\sum$	90	65	632	942
	$=\sum x$	$=\sum y$	$=\sum xy$	$=\sum x^2$

$$\bar{x}=\frac{\sum x}{n}=\frac{90}{10}=9 \quad , \quad \bar{y}=\frac{\sum y}{n}=\frac{65}{10}=6.5$$

$$s_x^2=\frac{\sum x^2}{n}-\bar{x}^2=\frac{942}{10}-(9)^2=94.2-81=13.2$$

$$b_1=\frac{\frac{\sum xy}{n}-\bar{x}\,\bar{y}}{s_x^2}=\frac{\frac{632}{10}-(9)(6.5)}{13.2}=\frac{63.2-58.5}{13.2}\approx\boxed{0.36}$$

$$b_0=\bar{y}-b_1\bar{x}=6.5-(0.36)9=6.5-3.24\approx\boxed{3.26}$$

$\therefore$ معادلة خط الانحدار البسيط في هذه الحالة: $\hat{y}=3.26+0.36x$ ، ولتوقع قيمة الاستهلاك المحلي عندما يصل الإنتاج 16000000 برميل، نحول وحدة هذه القيمة من برميل إلى مليون برميل بالقسمة على مليون أي أن القيمة المستخدمة في توقع الاستهلاك هي $x_h=16$ ، وبالتعويض في المعادلة السابقة نجد أن:

$$\hat{y}_h=b_0+b_1x_h=\hat{y}=3.26+0.36(16)=\boxed{9.02}$$

أي أن الاستهلاك المحلي قد يصل إلى 9.02 مليون برميل، أي ما يعادل 9020000 برميل خلال السنة.

3.3. مسائل محلولة

3.3.1. لدراسة العلاقة بين الدخل (x) والاستهلاك (y) بآلاف الريالات، كانت لدينا النتائج الآتية:

$$\sum x = 120 \qquad \sum y = 100 \qquad \sum xy = 516$$
$$\sum x^2 = 720 \qquad \sum y^2 = 410 \qquad n = 40$$

1. احسب معامل الارتباط الخطي بين الظاهرتين. ما نوع الارتباط؟ وما مدى قوته؟
2. معادلة (خط) انحدار الاستهلاك على الدخل.
3. تقدير الاستهلاك عندما يصل الدخل إلى **(10000)** ريال.

الحل:

.1

$$\bar{x} = \frac{\sum x}{n} = \frac{120}{40} = 3 \quad , \quad S_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{720}{40} - (3)^2} = \sqrt{18-9} = \sqrt{9}$$

$$\bar{y} = \frac{\sum y}{n} = \frac{100}{40} = 2.5 \quad , \quad S_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2} = \sqrt{\frac{410}{40} - (2.5)^2} = \sqrt{10.25-6.25} = \sqrt{4}$$

$$r_p = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{S_x\, S_y} = \frac{\frac{516}{40} - (3)(2.5)}{\sqrt{9}\,\sqrt{4}} = \frac{12.9-7.5}{\sqrt{36}} = \frac{5.4}{6} = \boxed{0.9}$$

من الملاحظ أن الارتباط طردي قوي بين الدخل والاستهلاك.

.2

$$b_1 = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{S_x^2} = \frac{5.4}{9} = \boxed{0.6}$$

$$b_0 = \bar{y} - b_1\bar{x} = 2.5 - (0.6)3 = 2.5 - 1.8 = \boxed{0.7}$$

$$\Rightarrow \boxed{\hat{y} = 0.7 + 0.6x}$$

3. نلاحظ أن وحدة القياس هي آلاف الريالات لذلك فإن قيمة الدخل 10000 ريال ستحول إلى 10 آلاف الريالات وبالتالي $x_h = 10$، أي أن:

$$\hat{y}_h = b_0 + b_1 x_h = 0.7 + 0.6(10) = 0.7 + 6 = 6.7$$

أي أن قيمة الاستهلاك المقدرة تساوي 6700 ريال.

3.3.2. البيانات التالية تمثل أعمار عينة من الأزواج وزوجاتهم بالسنوات:

عمر الزوج (y)	50	60	24	30	25	35	44	56	37	30
عمر الزوجة (x)	40	37	20	25	19	25	25	42	30	20

1. احسب معامل الارتباط الخطي بين الظاهرتين بين أعمار الزوج والزوجة.
2. احسب معامل ارتباط الرتب بين أعمار الزوج والزوجة.
3. حدد نوع وقوة الارتباط من خلال معاملي الارتباط.

الحل:

1.

	x	y	xy	x^2	y^2
	40	50	2000	1600	2500
	37	60	2220	1369	3600
	20	24	480	400	576
	25	30	750	625	900
	19	25	475	361	625
	25	35	875	625	1225
	25	44	1100	625	1936
	42	56	2352	1764	3136
	30	37	1110	900	1369
	20	30	600	400	900
$\sum$	283	391	11962	8669	16767
	$=\sum x$	$=\sum y$	$=\sum xy$	$=\sum x^2$	$=\sum y^2$

$$\bar{x}=\frac{\sum x}{n}=\frac{283}{10}=28.3 \quad , \quad \bar{y}=\frac{\sum y}{n}=\frac{391}{10}=39.1$$

$$S_x=\sqrt{\frac{\sum x^2}{n}-\bar{x}^2}=\sqrt{\frac{8669}{10}-(28.3)^2}=\sqrt{866.9-800.89}=\sqrt{66.01}$$

$$S_y=\sqrt{\frac{\sum y^2}{n}-\bar{y}^2}=\sqrt{\frac{16767}{10}-(39.1)^2}=\sqrt{1676.7-1528.81}=\sqrt{149.89}$$

$$r_p=\frac{\frac{\sum xy}{n}-\bar{x}\,\bar{y}}{S_x\,S_y}=\frac{\frac{11962}{10}-(28.3)(39.1)}{\sqrt{(66.01)(149.89)}}=\frac{1196.2-1106.53}{\sqrt{9894.2389}}=\frac{89.67}{99.47}=\boxed{0.90}$$

.2

x	y	رتب x	رتب y	d	d^2
50	40	8	9	-1	1
60	37	10	8	2	4
24	20	1	2.5	-1.5	2.25
30	25	3.5	5	-1.5	2.25
25	19	2	1	1	1
35	25	5	5	0	0
44	25	7	5	2	4
56	42	9	10	-1	1
37	30	6	7	-1	1
30	20	3.5	2.5	1	1
			$\sum$	0.0	**17.5**
				$=\sum d$	$=\sum d^2$

$$r_s = 1-\frac{6\sum d^2}{n(n^2-1)} = 1-\frac{6(17.5)}{10(100-1)} = 1-\frac{105}{990} = 1-0.11 = 0.89$$

.3 نلاحظ أن كلا المعاملين قيمتهما تدل على أن الارتباط طردي قوي بين أعمار عينة الأزواج وزوجاتهم.

3.4. تمارين

3.4.1. لدراسة العلاقة بين الدخل (x) والاستهلاك (y) بمئات الريالات في مدينه ما، أخذت عينة من الأسر فأعطت النتائج الآتية:

x	5	4	5	6	9	10	9	12	11	9
y	5	4	5	5	8	6	8	11	10	8

1. احسب معامل ارتباط بيرسون بين الظاهرتين.
2. احسب معامل ارتباط سبيرمان بين الظاهرتين.
3. أوجد قيمة الاستهلاك عندما يصل الدخل **(800)** ريال.

3.4.2. البيانات الآتية توضح المبالغ المنصرفة على الدعاية (x) (بالآلف الريالات) لإحدى المؤسسات في عدة مناطق وحجم المبيعات (y) (بالآلف الريالات) في تلك المناطق:

x	5	3	3	7	6	3	1
y	20	15	10	25	12	18	5

1. احسب معامل ارتباط بيرسون بين الظاهرتين علماً بأن:

$$\sum y^2 = 1843 \;,\; \sum xy = 481$$

2. قدر حجم المبيعات عندما يصل المنصرف على الدعاية **4500** ريال.

3.4.3. البيانات الآتية تمثل التكلفة الحدية (x) لإنتاج السلعة بالريال وإجمالي المنتج السنوي (y) بالمليون طن لثمانية مؤسسات:

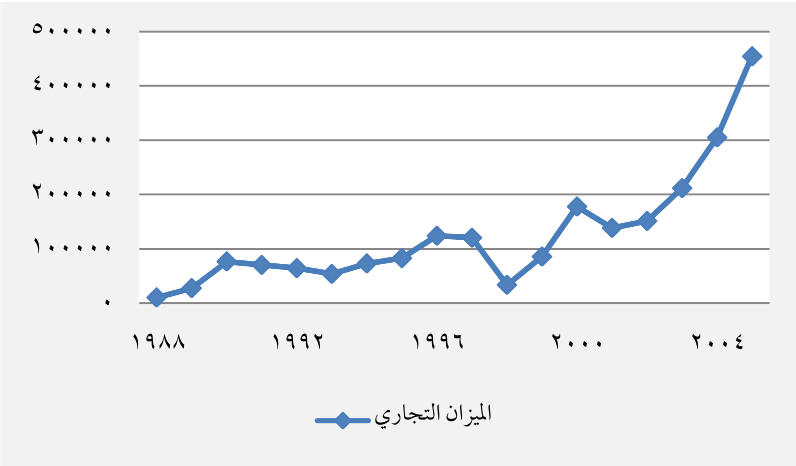
x	5	3	3	7	6	3	1
y	20	15	10	25	12	18	5

1. احسب معامل ارتباط بيرسون بين الظاهرتين.
2. قدر حجم المبيعات عندما يصل المنصرف على الدعاية **4500** ريالاً.

4.1. مفاهيم أساسية

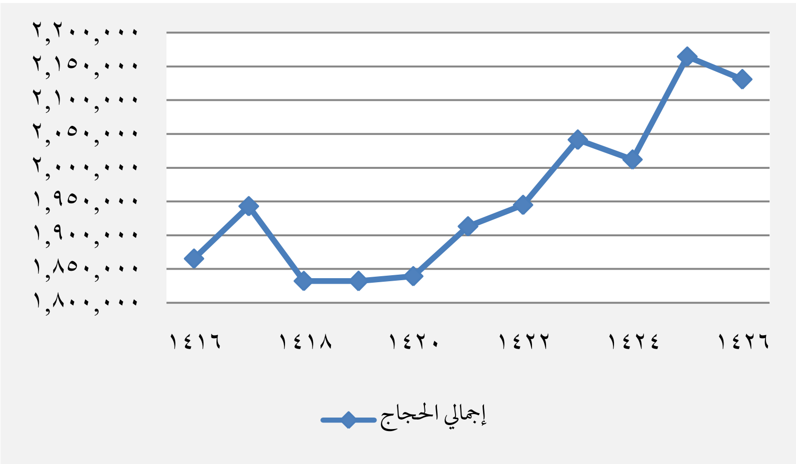
السلسلة الزمنية: هي مجموعة القراءات التي تأخذها ظاهرة (أو متغير) ما خلال فترات زمنية غالباً تكون متساوية وتختلف هذه الفترات حسب طبيعة الظاهرة.

مثال (4.1): في ما يلي مجموعة من أشكال السلاسل الزمنية لبعض الظواهر:



رسم توضيحي 9: قيمة الميزان التجاري خلال الفترة 1988م إلى 2005م (القيمة بملايين الريالات)

(المصدر: موقع مصلحة الإحصاءات العامة على شبكة الانترنت – وزارة الاقتصاد والتخطيط – المملكة العربية السعودية)

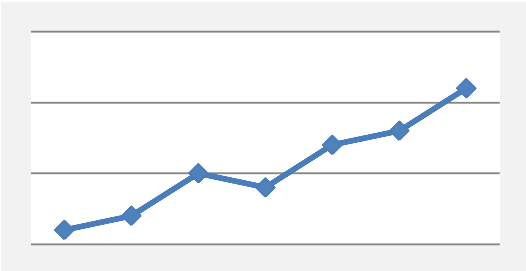


رسم توضيحي 10: إجمالي أعداد الحجاج من 1416م إلى 1426هـ

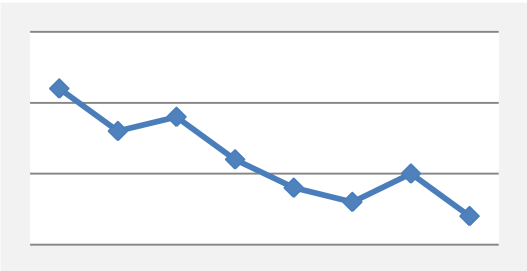
(المصدر: موقع وزارة الحج على شبكة الانترنت – المملكة العربية السعودية)

تتكون السلسلة الزمنية لأي ظاهرة عادة من العناصر الآتية:

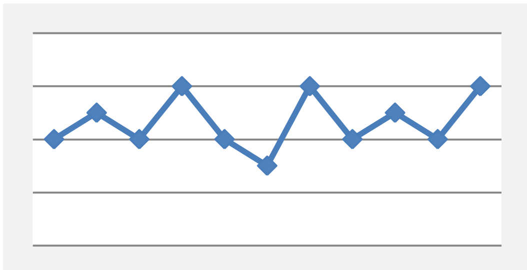
- **الاتجاه العام**: وهو اتجاه التطور الذي تأخذه السلسلة الزمنية خلال فترة طويلة من الزمن بالرغم من التذبذبات الموجودة بها، ويكون التطور إما بالزيادة أو بالنقصان، وبعض السلاسل لا يوجد لها اتجاه.
- **التغيرات الموسمية**: وهي التغيرات التي تتكرر بانتظام خلال فترة زمنية أقل من السنة، عادة في المواسم.
- **التغيرات الدورية**: وهي التغيرات التي تحدث في فترات زمنية أكثر من سنة، وعادة كل خمس أو عشر سنوات.
- **التغيرات العرضية**: وهي التغيرات التي تحدث نتيجة حوادث فجائية غير متوقعة مثل الفيضانات والأعاصير والحروب...الخ.



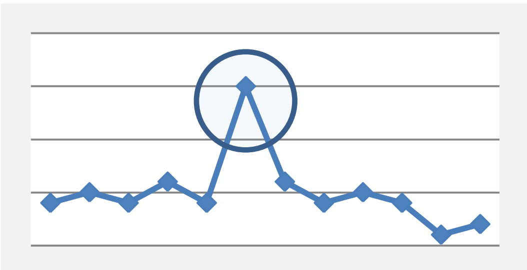
سلسلة ذات اتجاه عام بالزيادة



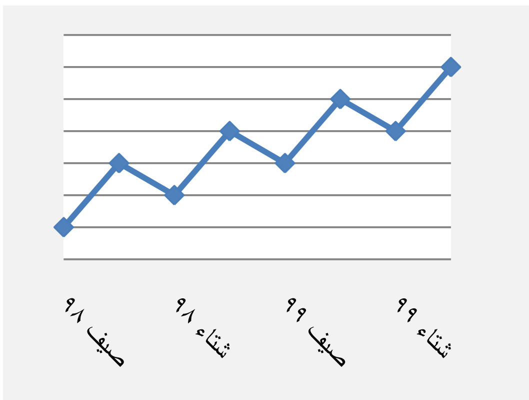
سلسلة ذات اتجاه عام بالنقصان



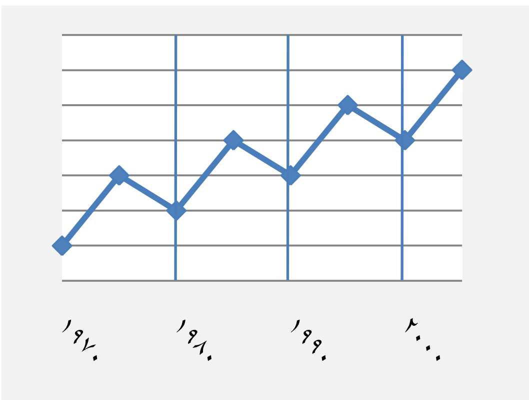
سلسلة ليس لها اتجاه عام



سلسلة ذات عامل عرضي



سلسلة ذات اتجاه زيادة وتغيرات موسمية



سلسلة ذات اتجاه زيادة تغيرات دورية

4.2. تعيين الاتجاه العام

يتطلب تحليل السلاسل الزمنية عادة تحليل المكونات الأربعة السالف ذكرها، ولكن في الأغلب تخلو السلاسل الزمنية من التغيرات الموسمية والدورية والعرضية، لذلك سنتعرض لتعيين الاتجاه العام، وسنكتفي بالسلاسل ذات الاتجاه العام الخطي. يجري تعيين الاتجاه العام الخطي بأسلوب الانحدار الخطي البسيط السالف ذكره في الباب السابق، باعتبار أن الزمن (السنوات، الشهور،...الخ) متغير مستقل (x)، والمتغير التابع (y) هو الظاهرة (أو المتغير) محل الدراسة، وتُقدر معادلة الاتجاه العام الخطي على الصورة:

$$\hat{y} = b_0 + b_1 x \quad (*)$$

حيث: b_0: الجزء المقطوع من محور y

b_1: ميل الخط المستقيم أو معامل انحدار x على y (أو y/x)

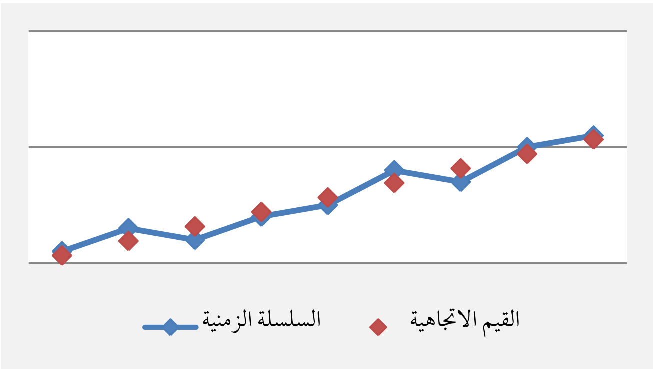
وتحسب القيمتين b_0 و b_1 من العلاقتين التاليتين:

$$b_1 = \frac{\dfrac{\sum xy}{n} - \bar{x}\,\bar{y}}{s_x^2} \qquad\qquad b_0 = \bar{y} - b_1\bar{x}$$

ملاحظات:

1. تُعين للمتغير المستقل القيم $x = 0, 1, 2, \ldots$ لتمثل وحدة الزمن.
2. تدل قيمة b_1 على الاتجاه العام.

ولإيجاد أي قيمة مقدرة جديدة $\hat{y}_h$، نعوض بقيمة معلومة للمتغير المستقل ولتكن x_h في المعادلة $(*)$، وإذا عوضنا بقيم السلسلة المشاهدة في معادلة نحصل على ما يسمى بالقيم الاتجاهية وسنرمز لها بالرمز y'.



رسم توضيحي 11: السلسلة الزمنية والقيم الاتجاهية

مثال (4.1): البيانات التالية تمثل عدد الحقول المكتشفة (y) خلال الأعوام 1991م إلى 2000م:

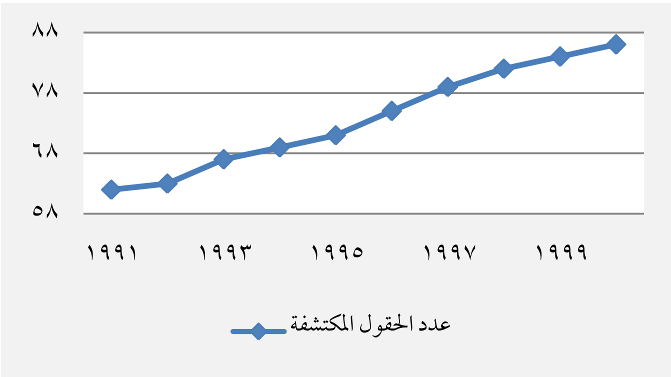
السنة	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
y	62	63	67	69	70	75	79	82	84	86

مصدر البيانات: موقع وزارة البترول والثروة المعدنية على شبكة الإنترنت – المملكة العربية السعودية

1. مثل السلسلة الزمنية بيانياً، على ماذا يدل الاتجاه العام للظاهرة، على الزيادة أم النقصان؟
2. قدر معادلة الاتجاه العام الخطي ثم توقع عدد الحقول المكتشفة عام 2002م.
3. أوجد القيمة الاتجاهية لعدد الحقول المكتشفة لعام 1994م.

الحل:

.1



يدل الاتجاه العام على الزيادة في قيمة عدد الحقول المكتشفة.

.2

السنة	x	y	$x\,y$	x^2
1991	0	62	0	0
1992	1	63	63	1
1993	2	67	134	4
1994	3	69	207	9
1995	4	70	280	16
1996	5	75	375	25
1997	6	79	474	36
1998	7	82	574	49
1999	8	84	672	64
2000	9	86	774	81
$\sum$	45	737	3553	285
	$=\sum x$	$=\sum y$	$=\sum x\,y$	$=\sum x^2$

$$\bar{x} = \frac{\sum x}{n} = \frac{45}{10} = 4.5 \quad , \quad \bar{y} = \frac{\sum y}{n} = \frac{737}{10} = 73.7$$

$$s_x^2 = \frac{\sum x^2}{n} - \bar{x}^2 = \frac{285}{10} - (4.5)^2 = 28.5 - 20.25 = 8.25$$

$$b_1 = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{s_x^2} = \frac{\frac{3553}{10} - (4.5)(73.7)}{8.25} = \frac{355.3 - 331.65}{8.25} \approx \boxed{2.87}$$

$$b_0 = \bar{y} - b_1\bar{x} = 73.7 - (2.87)4.5 \approx \boxed{60.79}$$

$\therefore$ معادلة الاتجاه العام الخطي في هذه المثال $\hat{y} = 60.79 + 2.87x$ ، ولتوقع عدد الحقول المتوقع اكتشافها عام 2002م نعوض بقيمة تدل على هذا الزمن؛ حيث أن 2000م $\leftarrow x = 9$ إذن 2002م $\leftarrow x_h = 11$ ، وبالتعويض في معادلة الاتجاه العام نجد أن:

$$\hat{y}_h = 60.79 + 2.87x_h = 60.79 + 2.87(11) = 92.36 \approx 92$$ حقل

3. القيمة الاتجاهية لعدد الحقول المكتشفة عام 1994م نحصل عليها بالتعويض بقيمة x المقابلة لهذه السنة:

$$y'_{x=3} = 60.79 + 2.87x = 60.79 + 2.87(3) = 69.4 \approx 69$$

القيمة الاتجاهية تساوي 69 حقل.

4.3. تمارين

4.3.1. البيانات التقريبية التالية تمثل الاستهلاك المحلي السنوي لمادة الإسفلت (y) بملايين البراميل خلال الأعوام **1995**م إلى **2002**م:

السنة	1995	1996	1997	1998	1999	2000	2001	2002
الاستهلاك	4.5	5	6	6.5	7	8	9	11

مصدر البيانات الأصلية: موقع وزارة البترول والثروة المعدنية على شبكة الإنترنت – المملكة العربية السعودية

1. مثل السلسلة الزمنية بيانياً، ثم علق على الرسم.
2. قدر معادلة الاتجاه العام الخطي ثم توقع الاستهلاك المحلي لمادة الإسفلت عام **2004**م.
3. أوجد القيمة الاتجاهية للاستهلاك المحلي لمادة الإسفلت لعام **1999**م.

4.3.2. البيانات التقريبية التالية تمثل متوسط سعر سلة أوبك (دولار/ برميل) للنفط الخام (y) خلال الأعوام **1991**م إلى **2000**م:

السنة	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
السعر	17	20	19	12	17	28	23	24	28	36

مصدر البيانات الأصلية: موقع وزارة البترول والثروة المعدنية على شبكة الإنترنت – المملكة العربية السعودية

1. حدد الاتجاه العام للظاهرة، عن طريق تمثيل السلسلة بيانياً.
2. قدر معادلة الاتجاه العام الخطي ثم توقع سعر سلة أوبك للنفط الخام عام **2005**م.
3. أوجد القيمة الاتجاهية لسعر سلة أوبك للنفط الخام لعام **2000**م.

4.3.3. البيانات التالية تمثل إجمالي الوفيات السنوي نتيجة الحوادث المرورية بمدينة الرياض خلال الأعوام **1423**هـ إلى **1427**هـ:

السنة	1423	1424	1425	1426	1427
الوفيات	451	479	430	408	353

مصدر البيانات: تقرير حوادث المرور لمدينة الرياض لعام **1427**هـ – موقع إدارة مرور منطقة الرياض – المملكة العربية السعودية

1. ارسم السلسلة الزمنية وعلق على اتجاه الظاهرة.
2. قدر معادلة الاتجاه العام الخطي ثم توقع إجمالي الوفيات عام **1428**هـ.
3. أوجد القيمة الاتجاهية لإجمالي الوفيات لعام **1425**هـ.

5.1. مفاهيم أساسية

الرقم القياسي للأسعار: هو رقم نسبي يقيس التغير الذي يطرأ على أسعار سلعة واحدة أو أكثر، عادة من سنة – تسمى سنة الأساس – لأخرى – تسمى سنة المقارنة –.

رموز مستخدمة:

P_0 : الأسعار في فترة الأساس. P_1 : الأسعار في فترة المقارنة.

Q_0: الكميات في فترة الأساس. Q_1: الكميات في فترة المقارنة.

دلالة الرقم القياسي: سنرمز للرقم القياسي بالرمز I

- إذا كان الرقم القياسي $(I) < 100$، فذلك يدل على النقصان في الأسعار بمقدار:

$$(100 - I) \times 100 = \%$$

- إذا كان الرقم القياسي $(I) > 100$، فذلك يدل على الزيادة في الأسعار بمقدار:

$$(I - 100) \times 100 = \%$$

5.2. الأرقام القياسية للأسعار

(أولاً) الرقم القياسي البسيط: يحسب من العلاقة التالية:

$$I_S = \frac{\sum P_1}{\sum P_0} \times 100$$

(ثانياً) الرقم القياسي المرجح بكميات الأساس (لاسبير): يحسب من العلاقة التالية:

$$I_L = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100$$

(ثالثاً) الرقم القياسي المرجح بكميات المقارنة (باشي): يحسب من العلاقة التالية:

$$I_P = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100$$

(رابعاً) الرقم القياسي الأمثل (فيشر): يحسب من العلاقة التالية:

$$I_F = \sqrt{I_L \times I_P}$$

مثال (5.1): الجدول التالي يوضح سعر كميات معينة لبعض مشتقات النفط:

السنة / المنتج	عام 1425 هـ		عام 1427 هـ	
	السعر اللتر (بالريال)	الكمية (باللتر)	السعر اللتر (بالريال)	الكمية (باللتر)
البنزين	0.9	10	0.6	11
الديزل	0.4	11	0.3	12

باعتبار أن سنة 1425 هـ سنة الأساس، ناقش التغير الحاصل في الأسعار بحساب:

- **الرقم البسيط للأسعار.**
- **الرقم القياسي للأسعار المرجح بكميات الأساس (رقم لاسبير).**
- **الرقم القياسي للأسعار المرجح بكميات المقارنة (رقم باشي).**
- **الرقم القياسي الأمثل للأسعار (رقم فيشر).**

الحل:

السنة / المنتج	عام 1425 هـ		عام 1427 هـ		P_1Q_0	P_0Q_0	P_1Q_1	P_0Q_1
	P_0	Q_0	P_1	Q_1				
البنزين	0.9	10	0.6	11	6	9	6.6	9.9
الديزل	0.4	11	0.3	12	3.3	4.4	3.6	4.8
المجموع	**1.3** $\sum P_0$		**0.9** $\sum P_1$		**9.3** $\sum P_1Q_0$	**13.4** $\sum P_0Q_0$	**10.2** $\sum P_1Q_1$	**14.7** $\sum P_0Q_1$

- **الرقم البسيط للأسعار**

$$I_S = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{0.9}{1.3} \times 100 = 69.23$$

أي أن الأسعار أن انخفضت بمقدار 30.77%

- **الرقم القياسي للأسعار المرجح بكميات الأساس (رقم لاسبير)**

$$I_L = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 = \frac{9.3}{13.4} \times 100 = 69.40$$

أي أن الأسعار أن انخفضت بمقدار 30.6%

- **الرقم القياسي للأسعار المرجح بكميات المقارنة (رقم باشي)**

$$I_P = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100 = \frac{10.2}{14.7} \times 100 = 69.39$$

أي أن الأسعار أن انخفضت بمقدار 30.61%

- **الرقم القياسي الأمثل للأسعار (رقم فيشر)**

$$I_F = \sqrt{I_L \times I_P} = \sqrt{69.40 \times 69.39} = 69.39$$

أي أن الأسعار أن انخفضت بمقدار 30.61%

مثال (5.2): الجدول الآتي يوضح أسعار ثلاث سلع والكميات المستهلكة منها عامي 1424هـ، 1427هـ.

السلعة	1424		1427	
	السعر	الكمية	السعر	الكمية
أ	5	30	8	40
ب	8	10	12	20
ج	7	20	10	30

باعتبار أن عام 1424هـ سنة أساس، قيم التغير الحاصل في الأسعار بحساب:

- **الرقم البسيط للأسعار**
- **الرقم القياسي الأمثل للأسعار.**

الحل:

السنة / السلعة	1424		1427		P_1Q_0	P_0Q_0	P_1Q_1	P_0Q_1
	P_0	Q_0	P_1	Q_1				
أ	5	30	8	40	240	150	320	200
ب	8	10	12	20	120	80	240	160
ج	7	20	10	30	200	140	300	210
المجموع	**20**		**30**		**560**	**370**	**860**	**570**

- **الرقم البسيط للأسعار**

$$I_S = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{30}{20} \times 100 = 150$$

أي أن الأسعار أن زادت بمقدار 50%

- **الرقم القياسي الأمثل للأسعار (رقم فيشر)**

$$I_F = \sqrt{I_L \times I_P} = \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1}} \times 100 = \sqrt{\frac{560}{370} \times \frac{860}{570}} \times 100$$

$$= \sqrt{1.51 \times 1.51} \times 100 = 151$$

أي أن الأسعار أن زادت بمقدار 51%

5.3. تمارين

5.3.1. الجدول الآتي يوضح أسعار ثلاث سلع والكميات المستهلكة منها عامي **1426**هـ، **1427**هـ.

السلع	كميات 1427	كميات 1426	أسعار 1427	أسعار 1426
A	**300**	**200**	**20**	**25**
B	**100**	**150**	**40**	**60**
C	**40**	**20**	**30**	**35**

باعتبار ان عام **1426**هـ سنة أساس، قيم التغير الحاصل في الأسعار بحساب:

- الرقم البسيط للأسعار
- الرقم القياسي الأمثل للأسعار.

5.3.2. الجدول الآتي يوضح القيم التقريبية للسلع المستوردة (بالمليار ريال) والأوزان (بمائة ألف طن) لعامي **2005**م و**2006**م.

السنة / السلع المستوردة	عام 2006م		عام 2005م	
	الوزن (مئة ألف طن)	القيمة (مليار ريال)	الوزن (مئة ألف طن)	القيمة (مليار ريال)
الآلات والأجهزة ونحوها	**16**	**64**	**14**	**54**
معدات النقل	**14**	**48**	**13**	**47**
معادن عادية ومصنوعاتها	**81**	**38**	**58**	**34**
مواد غذائية	**11**	**31**	**15**	**33**
بقية السلع	**18**	**68**	**18**	**65**

مصدر البيانات الأصلية: موقع مصلحة الإحصاءات العامة على شبكة الانترنت – المملكة العربية السعودية

باعتبار ان عام **2005**م سنة أساس، ناقش التغير الحاصل في الأسعار بحساب:

- الرقم البسيط للأسعار.
- الرقم القياسي للأسعار المرجح بكميات الأساس (رقم لاسبير).
- الرقم القياسي للأسعار المرجح بكميات المقارنة (رقم باشي).
- الرقم القياسي الأمثل للأسعار (رقم فيشر).

6.1. مقدمة عن الاحتمالات

6.1.1. التجارب العشوائية

تلعب الاحتمالات دوراً هاماً في حياتنا اليومية، لأننا نستخدمها في قياس عدم التأكد. والاحتمالات أحد فروع الرياضيات الذي يهتم بدراسة **التجارب العشوائية**. وتسمى التجربة عشوائية إذا كانت نتائجها غير مؤكدة، أي لا نستطيع التنبؤ بها مسبقاً.

نتائج التجارب تنقسم من وجهة نظر الاحتمالات إلى ثلاث أنواع هي:

- **نتائج أو حوادث مؤكدة:** وهي نتائج أو حوادث لابد من وقوعها أو حدوثها. فمثلاً، إذا ألقيت قطعة نقود في الهواء، فإنها لابد وأن تسقط على الأرض. وإذا كانت الحادثة مؤكدة الوقوع فإنه يقال أن احتمال وقوعها يساوي **الواحد**.
- **نتائج أو حوادث مستحيلة:** وهي نتائج أو حوادث يستحيل وقوعها مثل، طرح أسهم لشركة مساهمة كبرى ثم لا يتم تخصيص أي سهم لأي مكتتب . وإذا كانت الحادثة مستحيلة الوقوع، فإن احتمال وقوعها يساوي **الصفر**.
- **نتائج أو حوادث محتملة (ممكنة / غير مؤكدة)**: وهي نتائج التجارب العشوائية والتي لا نستطيع التنبؤ بوقوعها مسبقاً ، ولكننا نستطيع باستخدام تعريف الاحتمالات أن نحسب احتمال وقوعها. وإذا كانت الحادثة محتملة، فإن احتمال وقوعها ينحصر بين **الصفر والواحد**.

6.1.2. تعريف الاحتمال

إذا كان لدينا تجربة ما تقع بطرق عددها (N) طريقة، وكان من بينها حدث معين (A) مثلاً يقع بطرق عددها (m) طريقة. فإن احتمال وقوع الحدث (A) ويرمز له بالرمز $P(A)$ هو:

$$P(A) = \frac{\text{(A) عدد مرات ظهور الحدث}}{\text{عدد الحالات الكلية للتجربة}} = \frac{m}{N} \qquad m \le N$$

من خلال مناقشتنا لتعريف التجارب العشوائية:

إذا كانت الحادثة A ...	فإن احتمالها يساوي ...
مؤكدة	$\frac{m}{N}=1$
محتملة	$0<\frac{m}{N}<1$
مستحيلة	$\frac{m}{N}=0$

إذا كان احتمال وقوع حدث A يعبر عنه بالرمز p، وكان احتمال عدم وقوع الحدث نفسه – نرمز له بالرمز A' – يعبر بالرمز q، فإن:

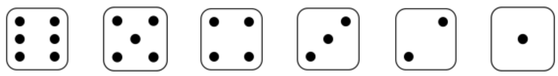
احتمال وقوع الحادثة A + احتمال عدم وقوعها (احتمال الحادثة A') = 1

$$p+q=1$$

مثال (6.1): ألقيت زهرة نرد (طاولة) متزنة مرة واحدة، ما هو احتمال ظهور عدد زوجي؟

الحل:

الحدث A  $\Leftarrow m=3$

حالات التجربة الكلية  $\Leftarrow N=6$

$$\therefore P(A)=\frac{m}{N}=\frac{3}{6}=\boxed{0.5}$$

مثال (6.2): سحبت ورقة من أوراق اللعب (الكوتشينة) ما احتمال أن:

- **تحمل الورقة الرقم سبعة؟**
- **أن تحمل الورقة صورة؟**

(تلميح: أوراق اللعب تتكون من أربع مجوعات كل مجموعة تحتوى على ثلاثة عشر ورقة ثلاث منها صور)

الحل:

- احتمال الحصول على الرقم سبعة

الحدث B 7♠ 7♣ 7♥ 7♦ $\Leftarrow m=4$

حالات التجربة ♠ × 13 ورقة ♣ × 13 ورقة ♥ × 13 ورقة ♦ × 13 ورقة $\Leftarrow N=52$

$$\therefore P(B)=\frac{m}{N}=\boxed{\frac{4}{52}}$$

- احتمال الحصول على الرقم صورة

الحدث C: ♠ صورة $\times 3$ ♣ صورة $\times 3$ ♥ صورة $\times 3$ ♦ صورة $\times 3$ $\Leftarrow$ $m = 12$

حالات التجربة: ♠ $13 \times$ ورقة ♣ $13 \times$ ورقة ♥ $13 \times$ ورقة ♦ $13 \times$ ورقة $\Leftarrow$ $N = 52$

$$\therefore P(B)=\frac{m}{N}=\boxed{\frac{12}{52}}$$

مثال (6.3): مصنع للمصابيح الكهربائية، من كل (1000) مصباح منتج يوجد (50) مصباح رديء، اخترنا أحد المصابيح من إنتاج المصنع:

- **ما احتمال الحصول على مصباح جيد ؟**
- **ما احتمال الحصول على مصباح رديء؟**

الحل: إنتاج المصنع هو العدد الكلي للحالات أي أن $N = 1000$.

- احتمال الحصول على مصباح جيد.

الحدث G: مصباح جيد $\Leftarrow m = 950$

$$\therefore P(G)=\frac{m}{N}=\frac{950}{1000}=\boxed{0.95}$$

- احتمال الحصول على مصباح رديء.

الحدث B: مصباح رديء $\Leftarrow m = 50$

$$\therefore P(B)=\frac{m}{N}=\frac{50}{1000}=\boxed{0.05}$$

6.2. بعض طرق العد (التوافيق)

عدد الطرق التي يمكن بها اختيار (x) من الأشياء من بين (n) من هذه الأشياء (الترتيب غير مهم) هو:

$$\binom{n}{x}=C_x^n=\frac{n!}{x!(n-x)!}$$

حيث: $n!=n\times(n-1)\times(n-2)\times\cdots\times2\times1$

فمثلاً: $4! = 4 \times 3 \times 2 \times 1 = 24$

مثال (6.3): بكم طريقة يمكن اختيار رجلين من بين أربع رجال؟

الحل: حيث أن $x = 2$ من الأشياء من بين $n = 4$

$$\binom{4}{2} = C_2^4 = \frac{4!}{2!(4-2)!} = \frac{4!}{2!\,2!} = \frac{4\times3\times2\times1}{(2\times1)\times(2\times1)} = \frac{24}{4} = 6$$

كتوضيح لهذه المسألة: افترض الأرقام التالية تدل على مكان الرجل: 1 2 3 4، طرق اختيار رجلين من هؤلاء الأربعة: 1,2 1,3 1,4 2,3 2,4 3,4، إذن عدد طرق الاختيار يساوي 6 (لاحظ أن الحالة 2,1 أهملت باعتبارها هي نفسها الحالة 1,2 وهكذا بالنسبة لباقي الحالات)

يمكن استخدام التوافيق لحساب عدد مرات إجراء تجربتين معاً كما يلي: إذا كان لدينا تجربة ما تقع بطرق عددها (H) طريقة ، تجربة أخرى تقع بطرق عددها (L) طريقة . فإن عدد مرات إجراء التجربتين معاً يساوي: $H \times L$

مثال (6.4): أعلنت إحدى الشركات عن توفر ثلاث وظائف شاغرة للرجال ووظيفتين للنساء، بكم طريقة يمكن الاختيار إذا كان عدد المتقدمين ست رجال وخمس نساء.

الحل:

عدد طرق اختيار الرجال:

$$H = C_3^6 = \frac{6!}{3!\,3!} = 20 \text{ طريقة}$$

عدد طرق اختيار النساء:

$$L = C_2^5 = \frac{5!}{2!\,3!} = 10 \text{ طريقة}$$

$\therefore$ عدد طرق الاختيار:

$$H \times L = 20 \times 10 = 200 \text{ طريقة}$$

ملاحظات هامة:

(i) $a^0 = 1$ $\quad$ a = عدد لا يساوي الصفر

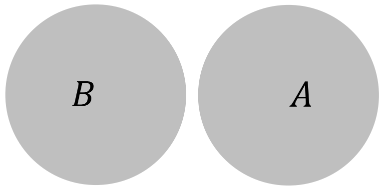
(ii) $0! = 1$

(iii) $C_0^n = 1$ $\quad$ (iv) $C_n^n = 1$ $\quad$ (v) مجموع الاحتمالات لأي حدث يساوي 1

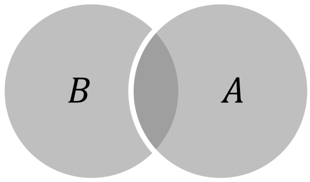
6.3. بعض قوانين الاحتمالات

6.3.1. الأحداث المانعة والغير مانعة

يقال أن الحدثين A وB حدثان مانعان أو متعارضان أو متنافيان إذا كان وقوع أحدهما يمنع وقوع الآخر (أي لا يقعان معاً). والعكس في حالة الأحداث الغير مانعة.



رسم توضيحي 12: حدثان مانعان



رسم توضيحي 13: حدثان غير مانعان

لحساب احتمال وقوع حدثان مانعان نستخدم القانون التالي:

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

لحساب احتمال وقوع حدثان غير مانعان نستخدم القانون التالي:

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

6.3.2. الأحداث المستقلة والغير مستقلة

يقال أن الحدثين A وB حدثان مستقلان إذا كان وقوع أحدهما لا يتأثر بوقوع أو عدم وقوع الآخر. والعكس في حالة الأحداث الغير مستقلة.

لحساب احتمال وقوع حدثان مستقلان نستخدم القانون التالي:

$$P(A \cap B) = P(A)P(B)$$

لحساب احتمال وقوع حدثان غير مستقلان نستخدم القانون التالي:

$$P(A \cap B) = P(A)P(B|A)$$

حيث: $P(B \mid A)$ يسمى الاحتمال الشرطي، بمعنى وقوع الحدث B بشرط أن يكون الحدث A وقع فعلاً.

ملاحظة هامة: في مسائل الاحتمالات، الحرف "أو" يدل على أن الأحداث إما مانعة أو غير مانعة، والحرف "و" يدل على أن الأحداث إما مستقلة أو غير مستقلة.

6.3.3. مسائل محلولة

أولاً: حالة الأحداث المانعة

مثال (6.5): سحبت ورقة من أوراق اللعب (الكوتشينة) ما احتمال أن تحمل الرقم 3 أو صورة؟

الحل: مما سبق في مثال (6.2)، حالات التجربة $N = 52$، كما أن:

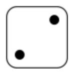
A: ورقة تحمل الرقم 3 $\qquad$ B: ورقة تحمل صورة

الحدثان A وB مانعان.

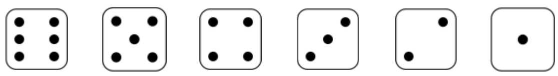
$$\therefore P(A \cup B) = P(A) + P(B) = \frac{4}{52} + \frac{12}{52} = \boxed{\frac{16}{52}}$$

مثال (6.6): ألقيت زهرة طاولة متزنة مرة واحدة ما احتمال ظهور العدد 2 أو عدد فردي؟

الحل:

الحدث A  $\Leftarrow \quad m = 1$

الحدث B  $\Leftarrow \quad m = 3$

حالات التجربة الكلية  $\Leftarrow \quad N = 6$

الحدثان A وB مانعان.

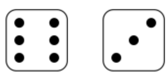
$$\therefore P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{3}{6} = \boxed{\frac{4}{6}}$$

ثانياً: حالة الأحداث الغير المانعة

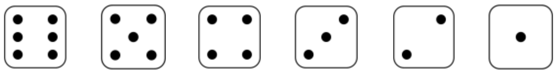
مثال (6.7): ألقيت زهرة طاولة متزنة مرة واحدة ما احتمال ظهور عدد يقبل القسمة على 2 أو على 3؟

الحل:

الحدث A  $\Leftarrow \quad m = 3$

الحدث B  $\Leftarrow \quad m = 2$

$A \cap B$  $\Leftarrow \quad m = 1$

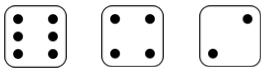
حالات التجربة الكلية  $\Leftarrow$ $N = 6$

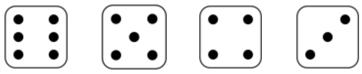
الحدثان A وB غير مانعان.

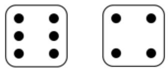
$$\therefore P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \boxed{\frac{4}{6}}$$

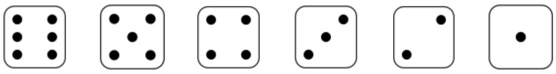
مثال (6.8): ألقيت زهرة طاولة متزنة مرة واحدة ما احتمال ظهور عدد زوجي أو عدد أكبر من 2 ؟

الحل:

الحدث A  $\Leftarrow$ $m = 3$

الحدث B  $\Leftarrow$ $m = 4$

$A \cap B$  $\Leftarrow$ $m = 2$

حالات التجربة الكلية  $\Leftarrow$ $N = 6$

الحدثان A وB غير مانعان.

$$\therefore P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{4}{6} - \frac{2}{6} = \boxed{\frac{5}{6}}$$

ثالثاً: حالة الأحداث المستقلة

مثال (6.9): إذا كان احتمال ارتفاع مؤشر سوق الأسهم بالدولة (A) يساوي (0.8) واحتمال ارتفاع مؤشر سوق الأسهم في دولة أخرى (B) يساوي (0.7)، ما احتمال أن يرتفع مؤشر سوقي أسهم الدولتين AوB ؟

الحل:

A: ارتفاع مؤشر سوق أسهم الدولة (A) B: ارتفاع مؤشر سوق أسهم الدولة (B)

الحدثان A وB مستقلان.

$$\therefore P(A \cap B) = P(A)P(B) = 0.8 \times 0.7 = \boxed{0.56}$$

رابعاً: حالة الأحداث الغير المستقلة

مثال (6.10): دولاب يحتوي على ثلاث ملفات حمراء وخمس ملفات سوداء، سحب منه ملفين عشوائياً على التوالي (بدون إرجاع)، ما احتمال:

- **أن يكون الملفين أسودين.**
- **أن يكون الملف الأول أسود والثاني أحمر**
- **أن يكون أحد الملفين أحمر والآخر أسود.**

الحل: الحدثان A وB غير مستقلان، لأن احتمال سحب ملف ثاني يعتمد على احتمال سحب الملف الأول.

- A: الملف الأول أسود (black) B: الملف الثاني أسود

$$P(A\cap B)=P(A)\cdot P(B\mid A)=P\left(\text{black}\right)\cdot P\left(\text{black}\mid\text{black}\right)$$
$$=\frac{5}{8}\cdot\frac{4}{7}=\frac{20}{56}$$

- A: الملف الأول أسود B: الملف الثاني أحمر (red)

$$P(A\cap B)=P(A)\cdot P(B\mid A)=P\left(\text{black}\right)\cdot P\left(\text{red}\mid\text{black}\right)$$
$$=\frac{5}{8}\cdot\frac{3}{7}=\frac{15}{56}$$

- A: الملف أسود B: الملف أحمر 1: الملف الأول 2: الملف الثاني

أحد الملفين أحمر والآخر أسود = (الملف الأول أسود والثاني أحمر) **أو** (الملف الأول أحمر والثاني أسود)

$$P(A\cap B)=P(A_1)\cdot P(B_2\mid A_1)+P(A_2)\cdot P(B_1\mid A_2)$$
$$=P\left(1^{\text{st}}\text{ black}\right)\cdot P\left(2^{\text{nd}}\text{ red}\mid 1^{\text{st}}\text{ black}\right)+P\left(1^{\text{st}}\text{ red}\right)\cdot P\left(2^{\text{nd}}\text{ black}\mid 1^{\text{st}}\text{ red}\right)$$
$$=\frac{5}{8}\cdot\frac{3}{7}+\frac{3}{8}\cdot\frac{5}{7}=\frac{30}{56}$$

مثال (6.11): دولاب يحتوي على ثلاث ملفات حمراء وخمس ملفات سوداء، سحب منه ثلاث ملفات عشوائياً على التوالي (بدون إرجاع)، ما احتمال:

- **أن يكون من بين الثلاث ملفات ملفين أسودين.**
- **أن يكون من بين الثلاث ملفات ملفين أسودين على الأقل.**
- **أن يكون من بين الثلاث ملفات ملفين أسودين على الأكثر.**

الحل:

- A: الملف أسود B: الملف أحمر   1: الملف الأول   2: الملف الثاني   3: الملف الثالث

وحيث لم يرد في المسألة ترتيب لعملية السحب ، فسنجد أن:

من بين الثلاث ملفات ملفين أسود = (الملف الأول أسود والثاني أسود والثالث أحمر) **أو** (الملف الأول أسود والثاني أحمر والثالث أسود) **أو** (الملف الأول أحمر والثاني أسود والثالث أسود)

وعند حساب احتمال أي طريقة (أي ترتيب) نجد أنه متساوي بالنسبة للطرق الثلاثة. وعلى ذلك يكون الحل بطريقة مختصرة باستخدام التوافيق، كما يلي :

$P(A \cap B) =$ [عدد طرق الاختيار] $\times$ [احتمال أي ترتيب]

$$\begin{aligned} P(A \cap B) &= C_2^3 \times P(A_1) \cdot P(A_2 / A_1) \cdot P(B_3 / A_2 A_1) \\ &= 3\left[\frac{5}{8} \cdot \frac{4}{7} \cdot \frac{3}{6}\right] = \frac{180}{336} \end{aligned}$$

من بين الثلاث ملفات ملفين أسودين على الأقل = من بين الثلاث ملفات ملفين أسود **(الفقرة السابقة)** **أو** جميع الملفات الثلاث سوداء

$$\begin{aligned} P(A \cap B) &= C_2^3 \times P(A_1) \cdot P(A_2 / A_1) \cdot P(B_3 / A_2 A_1) + P(A_1) \cdot P(A_2 / A_2) \cdot P(A_3 / A_2 A_1) \\ &= 3\left[\frac{5}{8} \cdot \frac{4}{7} \cdot \frac{3}{6}\right] + \frac{5}{8} \cdot \frac{4}{7} \cdot \frac{3}{6} = \frac{180}{336} + \frac{60}{336} = \frac{240}{336} \end{aligned}$$

من بين الثلاث ملفات ملفين أسودين على الأكثر = من بين الثلاث ملفات ملفين أسود **(الفقرة السابقة)** **أو** من بين الثلاث ملفات ملف أسود **أو** من بين الثلاث ملفات ولا ملف أسود

$$\begin{aligned} P(A \cap B) &= C_0^3 \times P(B_1) \cdot P(B_2 / B_1) \cdot P(B_3 / B_2 B_1) \\ &+ C_1^3 P(A_1) \cdot P(B_2 / A_1) \cdot P(B_3 / B_2 A_1) \\ &+ C_2^3 P(A_1) \cdot P(A_2 / A_1) \cdot P(B_3 / A_2 A_1) \end{aligned}$$

يمكن حساب هذا الاحتمال مباشرة ولكن من الأسهل في هذه الحالة استخدام تعريف الاحتمال بأن احتمال وقوع حدث + عد وقوعه يساوي الواحد، بمعنى:

من بين الثلاث ملفات ملفين أسودين على الأكثر + احتمال جميع الثلاث الملفات سوداء = 1

$\Leftarrow$ احتمال من بين الثلاث ملفات ملفين أسودين على الأكثر = 1 – احتمال جميع الثلاث الملفات سوداء

$$\Rightarrow P(A \cap B) = 1 - P(A_1) \cdot P(A_2 / A_2) \cdot P(A_3 / A_2 A_1) = 1 - \frac{60}{336} = \frac{276}{336}$$

ملاحظة هامة: إذا لم يُذكر الترتيب في المسألة فيمكن استخدام التوافيق لتبسيط الحل.

6.4. تمارين

6.4.1. سحبت ورقه من أوراق اللعب (الكوتشينة)، ما هو احتمال أن تحمل الورقة المسحوبة الرقم **(2)** أو صورة؟

6.4.2. ألقيت زهرة نرد (طاوله) مرة واحدة ما هو احتمال الحصول على عدد زوجي أو عدد أكبر من **4**؟

6.4.3. صندوق به **(6)** كرات حمراء، **(4)** بيضاء سحبنا منه عشوائيا **(3)** كرات على التوالي ما هو احتمال الحصول على:

1. كرتين حمراء
2. كرتين على الأقل حمراء.
3. كرتين على الأكثر حمراء.

6.4.4. في التمرين السابق إذا سحبنا كرتين على التوالي ما هو احتمال أن تكون الكرتان:

1. من لون واحد .
2. من لونين مختلفين.

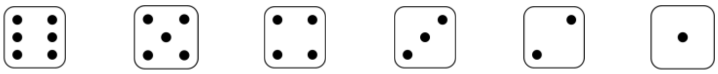
6.4.5. ألقيت قطعة نقود معدنية ثلاث مرات، ما احتمال:

1. ظهور شعار على قطعتين من بين الثلاث القطع النقدية.
2. ظهور كتابة على قطعتين من بين الثلاث القطع النقدية.
3. ظهور شعار واحد على الأقل على قطعة من بين الثلاث القطع النقدية.
4. ظهور كتابتين على الأكثر على قطعتين من بين الثلاث القطع النقدية.

7.1. المتغير العشوائي

7.1.1. مقدمة

يصاحب نتائج التجربة العشوائية مقدار يسمى "المتغير العشوائي"وهذا المقدار يأخذ قيماً مختلفة يعبر عن نتائج التجربة العشوائية. فمثلاً عند إلقاء زهرة طاولة مرة واحدة. التجربة هنا عشوائية ونتائجها هي:

نتائج التجربة 

X	**1**	**2**	**3**	**4**	**5**	**6**

المقدار X الذي رافق نتائج هذه التجربة يمكن أن يأخذ القيم: $1, 2, 3, 4, 5, 6$ ويسمى المتغير العشوائي **متغيراً** لأنه يأخذ قيماً مختلفة ويسمى **عشوائياً** لأنه يرافق نتائج التجربة العشوائية. رمزنا للمتغير العشوائي بالرمز X، ويمكن استخدام رموز أخرى للتعبير عن المتغير العشوائي.

7.1.2. أنواع المتغيرات العشوائية

(أولاً): المتغير العشوائي المنفصل (أو المتقطع)

يقال أن المتغير العشوائي (X) متغير منفصل إذا كان يأخذ قيماً صحيحة فقط تنتمي إلى مجموعة محدودة أو معدودة، من الأمثلة على هذا المتغير:

- عدد الأسهم المخصصة للفرد المكتتب في شركة مساهمة.
- عدد حوادث الشهرية على الطرق السريعة.

(ثانياً) المتغير العشوائي المستمر (أو المتصل)

يقال أن المتغير العشوائي (X) متغير مستمر إذا كان يأخذ جميع القيم الصحيحة والكسرية في مدى تغيره، أو كان ينتمي إلى مجموعة غير محدودة أو معدودة.

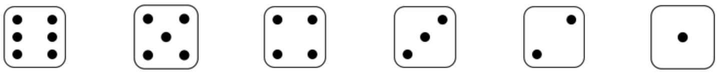
- أسعار المنتجات المختلفة.
- أجور العمال بإحدى الشركات.

7.2. التوزيعات الاحتمالية

7.2.1. مقدمة

التوزيع الاحتمالي ل لمتغير العشوائي عبارة عن دالة توضح احتمالات معينة لقيم المتغير العشوائي المختلفة، وهذه الدالة يعبر عنها بجدول أو صيغة رياضية تبين قيم المتغير والاحتمالات المقابلة لكل منها. فمثلاً، لتجربة إلقاء زهرة الطاولة نجد أن:

نتائج التجربة



X	1	2	3	4	5	6
$P(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

7.2.2. التوزيعات الاحتمالية المنفصلة

إذا كان X متغير عشوائي منفصل يأخذ القيم:

$$x_1, x_2, \ldots, x_n$$

ولكل قيمة احتمالات معينة كالتالي

$$P(x_1), P(x_2), \ldots, P(x_n)$$

فيقال أن للمتغير العشوائي المنفصل X **توزيعاً احتمالياً منفصلاً** $P(x)$ إذا حقق هذا التوزيع الشروط التالية:

$(1)\ P(x) \geq 0$ لجميع قيم X

$(2)\ \sum P(x) = 1$ (بمعنى أن مجموع الاحتمالات يساوي الواحد)

خصائص أساسية للتوزيع الاحتمالي المنفصل

1. **توقع التوزيع (التوقع الرياضي أو متوسط التوزيع)**

$$E(x) = \mu = \sum x\, P(x)$$

2. **تباين التوزيع**

$$\operatorname{var}(x) = \sigma^2 = \sum x^2 P(x) - \mu^2$$

3. **الانحراف المعياري**

$$\sqrt{\operatorname{var}(x)} = \sigma = \sqrt{\sigma^2}$$

مثال (7.1) (مثال على الأحداث الغير مستقلة): أعلنت إحدى شركات القطاع الخاص عن توفر ثلاث وظائف إدارية فتقدم لها أربع رجال وثلاث نساء. أوجد:

- **التوزيع الاحتمالي لعدد الرجال المختارين.**
- **احتمال اختيار رجلين على الأقل.**
- **احسب احتمال اختيار رجل واحد على الأكثر.**
- **خصائص التوزيع الاحتمالي.**

الحل:

X: عدد الرجال المختارين ← X: متغير عشوائي يأخذ القيم 0, 1, 2, 3

الحالة	$P(X=x)$
لم يتم اختيار أي رجل (تم اختيار ثلاث نساء)	$P(X=0)=C_0^3\left(\frac{3}{7}\cdot\frac{2}{6}\cdot\frac{1}{5}\right)=\frac{6}{210}$
تم اختيار رجل واحد (تم اختيار اثنين من النساء)	$P(X=1)=C_1^3\left(\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\right)=\frac{72}{210}$
تم اختيار رجلين (تم اختيار واحدة من النساء)	$P(X=2)=C_2^3\left(\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5}\right)=\frac{108}{210}$
تم اختيار ثلاث رجال (لم يتم اختيار أي واحدة من النساء)	$P(X=3)=C_3^3\left(\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\right)=\frac{24}{210}$

- التوزيع الاحتمالي لعدد الرجال المختارين:

X	0	1	2	3	$\sum$
P(x)	$\frac{6}{210}$	$\frac{72}{210}$	$\frac{108}{210}$	$\frac{24}{210}$	$\frac{210}{210}=1$

- احتمال اختيار رجلين على الأقل

$$P(X\geq 2)=\frac{108}{210}+\frac{24}{210}=\frac{132}{210}$$

- احتمال اختيار رجل واحد على الأكثر

$$P(X\leq 1)=\frac{6}{210}+\frac{72}{210}=\frac{78}{210}$$

> **ملاحظة هامة:** نستطيع استنتاج التوزيع الاحتمال لعدد النساء المختارات بسهولة، فمن الملاحظ أن احتمال عدم اختيار أي رجل هو نفسه احتمال اختيار ثلاث نسوة.

- خصائص التوزيع الاحتمالي:

X	0	1	2	3	$\sum$
$P(x)$	$\frac{6}{210}$	$\frac{72}{210}$	$\frac{108}{210}$	$\frac{24}{210}$	$\frac{210}{210}=1$
$x \cdot P(x)$	0	$\frac{72}{210}$	$\frac{216}{210}$	$\frac{72}{210}$	$\frac{360}{210}$
$x^2 \cdot P(x)$	0	$\frac{72}{210}$	$\frac{432}{210}$	$\frac{216}{210}$	$\frac{720}{210}$

$\mu = E(x) = \sum x \cdot P(x) = \frac{360}{210} = 1.71$ رجل

$$\sigma^2 = \text{var}(x) = \sum x^2 \cdot P(x) - \mu^2 = \frac{720}{210} - 1.71^2 = 3.43 - 2.92 = 0.51$$

$\sigma = \sqrt{\text{var}(x)} = \sqrt{0.51} = 0.71$ رجل

7.2.3. بعض التوزيعات الاحتمالية المنفصلة (حالة الأحداث المستقلة)

(أولاً) توزيع ذو الحدين: إذا كان لدينا تجربة ما تتكرر (n) مرة، وكان احتمال ظهور حدث ما مرة واحدة (النجاح) هو p، واحتمال عدم ظهور الحدث مرة واحدة (الفشل) هو q، فإن احتمال ظهور الحدث (X) مرة من بين التكرار n ، يتبع توزيع ذي الحدين الذي دالته الاحتمالية:

$$P(x) = C_x^n p^x q^{n-x} \quad ; X = 0,1,2,\dots,n$$

مواصفات التوزيع:

- توزيع منفصل يستخدم في حالة الأحداث المستقلة. ويتوقف على قيمة الاحتمال.
- $\sum P(x) = \sum C_x^n p^x q^{n-x} = 1$
- $p + q = 1$

خصائص التوزيع: يمكن إيجاد الخصائص مباشرة من المسألة حتى قبل حساب التوزيع الاحتمالي، وهذه الخصائص على النحو التالي:

- المتوسط: $\mu = np$
- التباين: $\sigma^2 = npq$
- الانحراف المعياري: $\sigma = \sqrt{npq}$

مثال (7.2): إذا كان احتمال ارتفاع مؤشر سوق الأسهم هو (3/4) اخترت ثلاث دول. أوجد:

- **التوزيع الاحتمالي لعدد الدول التي يرتفع مؤشر سوق أسهمها.**
- **متوسط التوزيع وتباينه وانحرافه المعياري.**
- **احتمال ارتفاع مؤشر سوق الأسهم لدولتين على الأقل.**

الحل:

$$n=3 \qquad p=\frac{3}{4} \qquad q=1-p=1-\frac{3}{4}=\frac{1}{4}$$

X (الحدث): ارتفاع مؤشر سوق الأسهم بدولة.

X: متغير عشوائي منفصل يأخذ القيم 0, 1, 2, 3، ويتبع توزيع ذي الحدين الذي دالته الاحتمالية:

$$P(x)=C_x^n p^x q^{n-x}=C_x^3\left(\frac{3}{4}\right)^x\left(\frac{1}{4}\right)^{3-x} \;;\; X=0,1,2,3$$

- التوزيع الاحتمالي لعدد الدول التي يرتفع مؤشر سوق أسهمها

$$P(X=0)=P(0)=C_0^3\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)^3=1\times1\times\frac{1}{64}=\frac{1}{64}$$

$$P(X=1)=P(1)=C_1^3\left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right)^2=3\times\frac{3}{4}\times\frac{1}{16}=\frac{9}{64}$$

$$P(X=2)=P(2)=C_2^3\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^1=3\times\frac{9}{16}\times\frac{1}{4}=\frac{27}{64}$$

$$P(X=3)=P(3)=C_3^3\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^0=1\times\frac{27}{64}\times1=\frac{27}{64}$$

$$\sum P(x)=\frac{1}{64}+\frac{9}{64}+\frac{27}{64}+\frac{27}{64}=\frac{64}{64}=1$$

- متوسط التوزيع وتباينه وانحرافه المعياري

$$\mu=np=3\times\frac{3}{4}=2.25$$

$$\sigma^2=npq=3\times\frac{3}{4}\times\frac{1}{4}=\frac{9}{16} \quad , \quad \sigma=\sqrt{\sigma^2}=\sqrt{\frac{9}{16}}=0.75$$

- احتمال ارتفاع مؤشر سوق الأسهم لدولتين على الأقل ($X = 2$ أو $X = 3$)

$$P(X\ge 2)=P(X=2)+P(X=3)=\frac{27}{64}+\frac{27}{64}=\frac{54}{64}$$

(ثانياً) توزيع بواسون: هو توزيع منفصل يستخدم في حالة الأحداث المستقلة ويهتم بالحوادث النادرة، مثل : الحرائق في إحدى المدن، الزلازل، الحوادث المرورية على إحدى الطرق، الأخطاء المطبعية في إحدى صفحات كتاب، المكالمات التي يتلقاها سنترال ما في ثانية واحدة ونحو ذلك، ويعطينا احتمال عدد مرات ظهور هذه الحوادث النادرة، فإذا كانت (X) ترمز لعدد مرات ظهور حادثة نادرة، فإن الدالة الاحتمالية للتوزيع تكون:

$$P(x)=\frac{e^{-\lambda}\lambda^{x}}{x!} \quad ; X =0,1,2,...$$

حيث: $e=2.7$ مقدار ثابت، وλ متوسط التوزيع.

مواصفات التوزيع:

- توزيع منفصل يستخدم في حالة الأحداث المستقلة النادرة.
- $\sum P(x)=\sum\frac{e^{-\lambda}\lambda^{x}}{x!}=1$
- متوسط هذا التوزيع يساوي تباينه

خصائص التوزيع: الخصائص على النحو التالي:

- المتوسط: $\mu=\lambda$
- التباين: $\sigma^{2}=\lambda$
- الانحراف المعياري: $\sigma=\sqrt{\lambda}$

مثال (7.3): إذا كان متوسط عدد الأخطاء المطبعية خلال صفحات أحدى الكتب هو 3 أخطاء. أوجد:

(ملاحظة $e^{-3}=0.05$)

- **احتمال عدم ظهور أي خطأ.**
- **احتمال ظهور خطأين.**
- **احتمال ظهور خطأين على الأكثر.**
- **احتمال ظهور خطأين على الأقل.**

الحل: نلاحظ أن $\lambda=3$

X (الحدث): عدد الأخطاء المطبعية.

X: متغير عشوائي منفصل يأخذ القيم $0, 1, 2,...$، ويتبع توزيع بواسون الذي دالته الاحتمالية:

$$P(x)=\frac{e^{-3}3^{x}}{x!} \quad ; X=0,1,2,\dots$$

- احتمال عدم ظهور أي خطأ $(X=0)$

$$P(X=0)=\frac{(0.05)\times 3^{0}}{0!}=0.05$$

- احتمال ظهور خطأين $(X=2)$

$$P(X=2)=\frac{(0.05)\times 3^{2}}{2!}=\frac{(0.05)\times 9}{2x1}=0.225$$

- احتمال ظهور خطأين على الأكثر ($X=0$ أو $X=1$ أو $X=2$)

$$P(X\le 2)=P(X=2)+P(X=1)+P(X=0)=0.225+\frac{(0.05)\times 3^{1}}{1!}+0.05$$
$$=0.225+0.15+0.05=0.425$$

- احتمال ظهور خطأين على الأقل ($X=2$ أو $X=3$ أو $X=4$ أو...الخ)

$$P(X\ge 2)=P(X=2)+P(X=3)+P(X=4)+\cdots$$
$$=1-\{P(X=1)+P(X=0)\}=1-\{0.15+0.5\}=1-0.20=0.8$$

7.2.3. التوزيعات الاحتمالية المتصلة (أو المستمرة)

عند دراستنا للمتغير العشوائي المنفصل، ذكرنا أن المتغير (X) يأخذ قيم صحيحة فقط وينتمي إلى مجموعة معدودة. أما في حالة المتغير العشوائي المستمر، فإن المتغير (X) يأخذ جميع القيم الصحيحة والكسرية في مدى تغيره ، و التوزيع الاحتمالي للمتغير المستمر يعطى في صورة دالة، تسمى **دالة كثافة الاحتمال** ويرمز لها بالرمز $f(x)$، ولا بد أن يتوفر فيها الشرطين الآتية لجميع قيم X:

$$[1] \quad f(x)\ge 0 \qquad\qquad [2] \quad \int f(x)dx=1$$

خصائص أساسية للتوزيع الاحتمالي المتصل (المستمر)

1. **توقع التوزيع (التوقع الرياضي أو متوسط التوزيع)**

$$E(x)=\mu=\int x\cdot f(x)dx$$

2. **تباين التوزيع**

$$\operatorname{var}(x)=\sigma^{2}=\int x^{2}\cdot f(x)dx-\mu^{2}$$

3. **الانحراف المعياري**

$$\sqrt{\operatorname{var}(x)}=\sigma=\sqrt{\sigma^{2}}$$

7.3. التوزيع الطبيعي

7.3.1. مقدمة

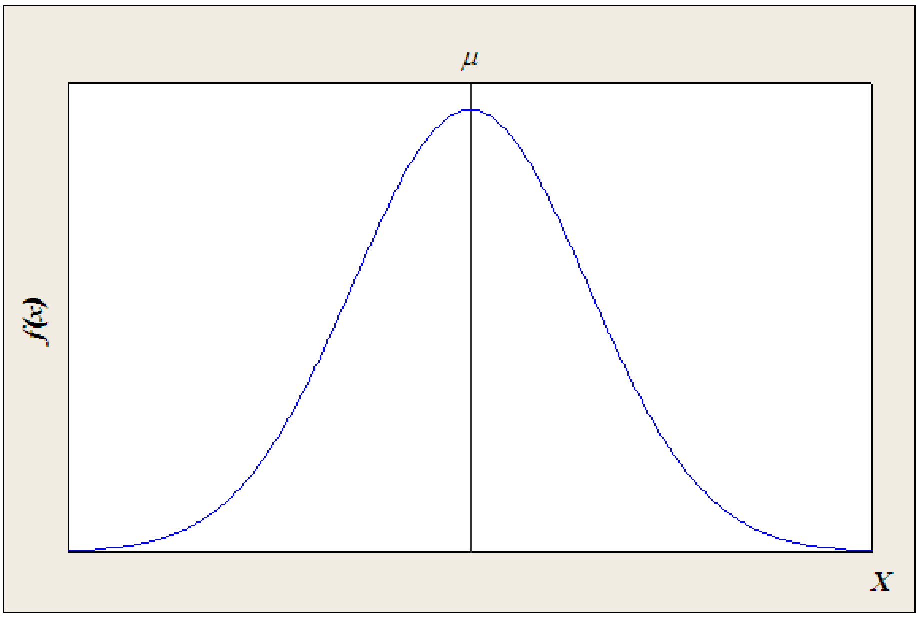
هو من أهم التوزيعات المستمرة وتأخذ دالة كثافته الاحتمالية شكل منحنى متماثل ذو قمة واحدة، ويمتد طرفاه إلى ما لا نهاية (∞) وقد وجد أن معظم التوزيعات الاحتمالية الأوزان، الأطوال، الأعمار ونحوها تأخذ شكلاً قريبا من المنحنى الطبيعي (رسم توضيحي 14). إذا كانت (X) متغير عشوائي متصل (مستمر) دالة كثافته الاحتمالية هي:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ;-\infty\le X\le\infty$$

فإنه يقال إن X يتبع توزيع طبيعي عادي متوسطه μ وانحرافه المعياري σ ونكتب باختصار

$$X \sim N(\mu,\sigma)$$

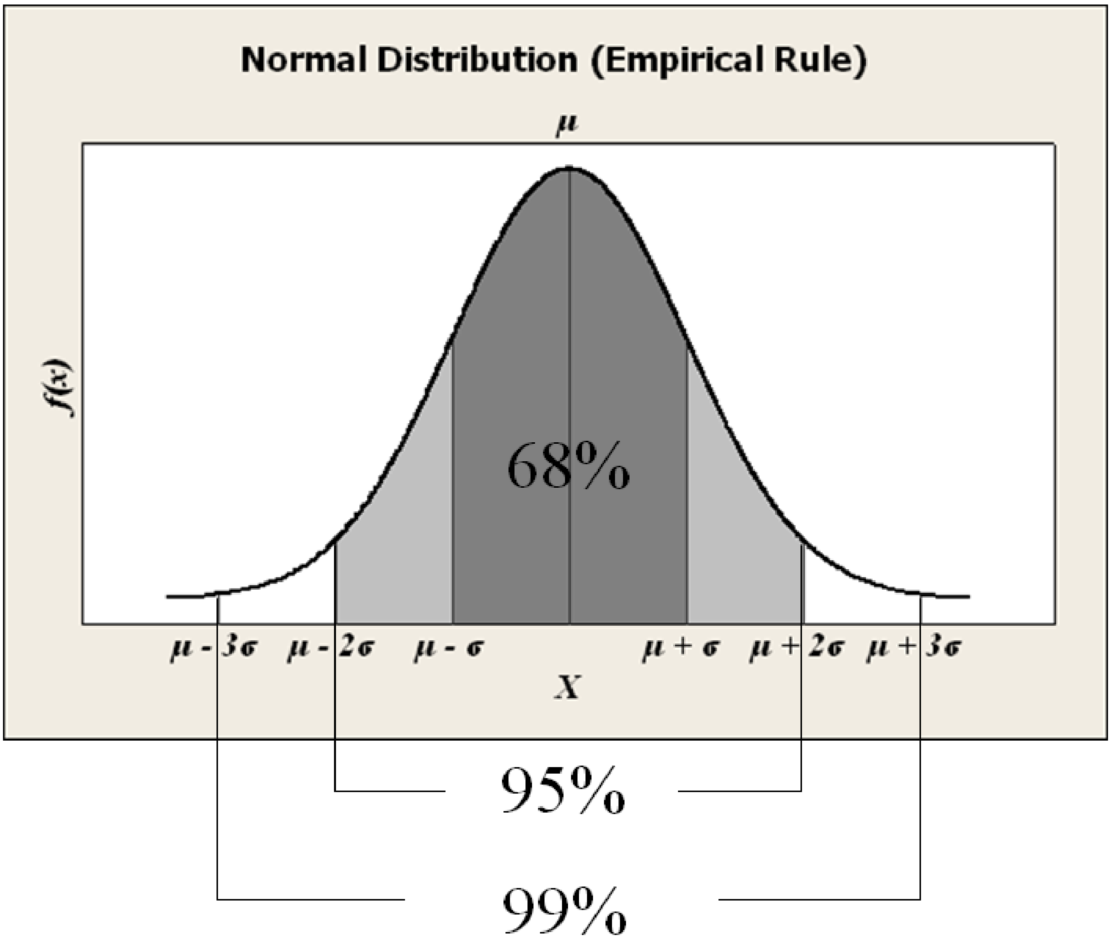
من خصائص التوزيع الطبيعي أنه متماثل حول متوسطه، كما أن المساحة تحت منحنى التوزيع تساوي الواحد وحيث أن التوزيع متماثل فإن المساحة تحت نصف المنحنى تساوي 0.5.



رسم توضيحي 14: التوزيع الطبيعي

القانون التجريبي للتوزيع الطبيعي (رسم توضيحي15):

- أي أن 68٪ من البيانات المتغير تنحصر بين $\mu\pm\sigma$.
- أي أن 95٪ من البيانات المتغير تنحصر بين $\mu\pm2\sigma$.
- أي أن 99٪ من البيانات المتغير تنحصر بين $\mu\pm3\sigma$.



رسم توضيحي 15: القانون التجريبي للتوزيع الطبيعي

7.3.2. التوزيع الطبيعي القياسي

يقال أن المتغير العشوائي المتصل Z يتبع توزيع طبيعي قياسي متوسطه $\mu = 0$ وانحرافه المعياري $\sigma = 1$ ونكتب باختصار

$$Z \sim N(0,1)$$

إذا كان المتغير العشوائي المتصل Z يعرف كما يلي: $Z = \dfrac{X - \mu}{\sigma}$، ومن خلال التوزيع الطبيعي القياسي يمكن حساب قيمة أي احتمال لمتغير عشوائي يتبع توزيع طبيعي عادي، كما سيتضح من المثال. ولقد خصصت جداول توضح قيمة الاحتمال تحت منحنى التوزيع الطبيعي القياسي.

ملاحظة هامة: بسبب تماثل التوزيع الطبيعي القياسي فإن:

$$P(0 < X < a) = P(-a < X < 0)$$

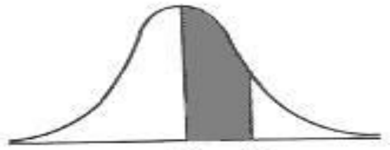
شروط استخدام جدول (Z):

1. تكون المسألة في التوزيع الطبيعي العادي، بمعنى $\mu \neq 0, \sigma \neq 0$ ، نحول التوزيع الطبيعي العادي إلى قياسي باستخدام تعريف Z.
2. جميع قيم الاحتمالات في الجدول محسوبة للنصف الموجب وما يقال على النصف الموجب ينطبق على النصف السالب بسبب تماثل التوزيع.
3. الاحتمالات محسوبة كلها على الصورة $P(0 < Z < a)$.

١ - جدول التوزيع الطبيعي القياسي

$Z \sim N(0, 1)$

المساحة المظللة تمثل $P(0<Z<z)$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.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	.2703	.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

مثال (7.4): إذا كان دخل **(1000)** أسرة في مدينة ما يتبع توزيع طبيعي متوسطه **(μ) 1800** ريال وانحرافه المعياري **(σ) 300** ريال. أوجد:

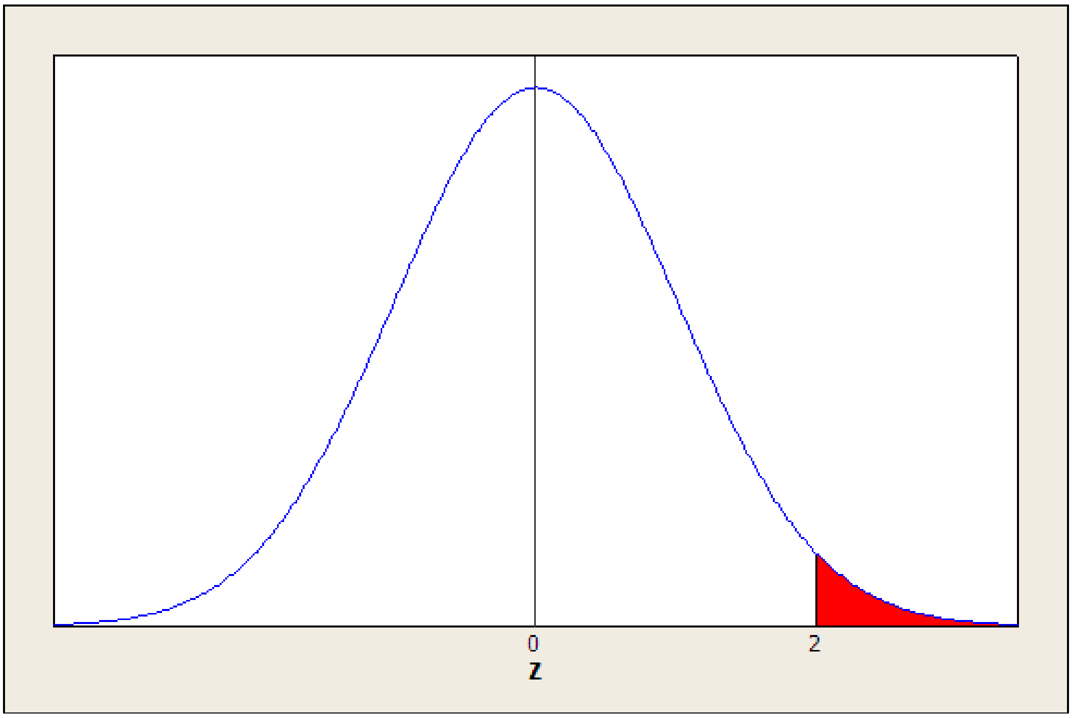
1. احتمال الحصول على دخل أسرة أكبر من **2400** ريال.
2. احتمال الحصول على دخل أسرة أكبر من **1500** ريال.
3. احتمال الحصول على دخل أسرة أقل من **2550** ريال.
4. احتمال الحصول على دخل أسرة أقل من **1200** ريال.
5. احتمال الحصول على دخل أسرة ينحصر بين **1650** و**2250** ريال.
6. احتمال الحصول على دخل أسرة ينحصر بين **2400** و**2550** ريال.
7. أوجد عدد الأسر التي يزيد دخلها عن **1500** ريال.

الحل: X متغير عشوائي مستمر يتبع توزيع طبيعي عادي متوسط $\mu = 1800$ ريال، وانحرافه المعياري $\sigma = 300$ ريال.

1. احتمال أن يكون دخل أسرة أكبر من 2400 ريال.

$$P(X \geq 2400) = P\left(\boxed{\frac{X-\mu}{\sigma}} \geq \frac{2400-\mu}{\sigma}\right) = P\left(Z \geq \frac{2400-\mu}{\sigma}\right)$$

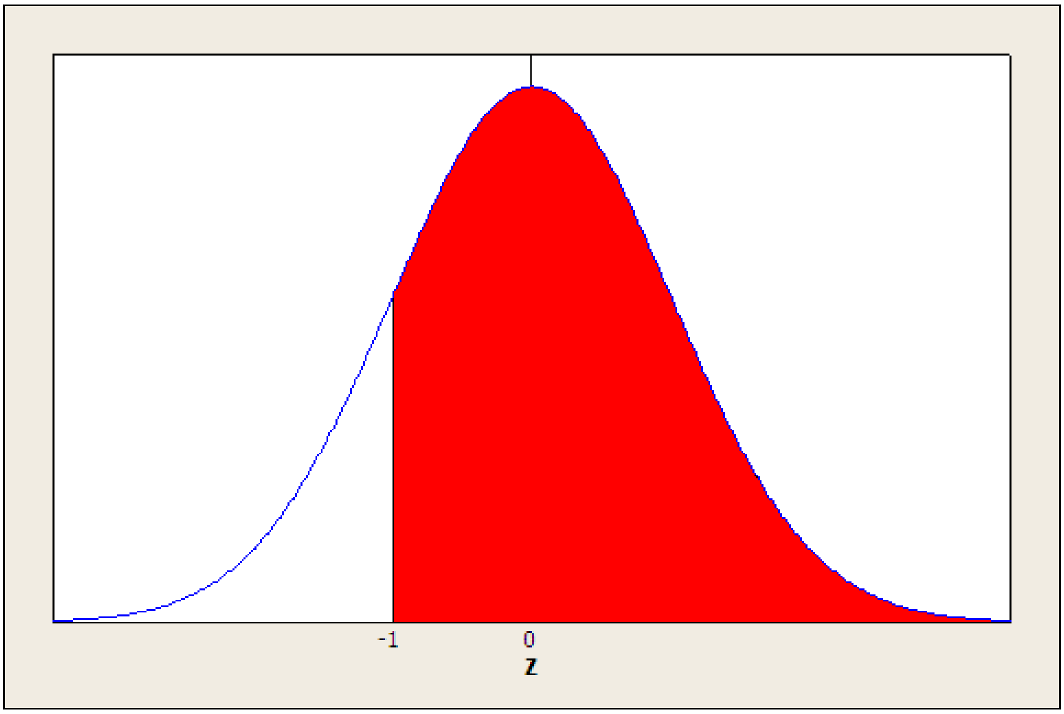
$$= P\left(Z \geq \frac{2400-1800}{300}\right) = P(Z \geq 2)$$



$$P(X \geq 2400) = P(Z \geq 2) = 0.5 - \underbrace{P(0 \leq Z \leq 2)}_{\text{من جدو ل } Z} = 0.5 - 0.4772 = 0.0228$$

2. احتمال أن يكون دخل أسرة أكبر من 1500 ريال.

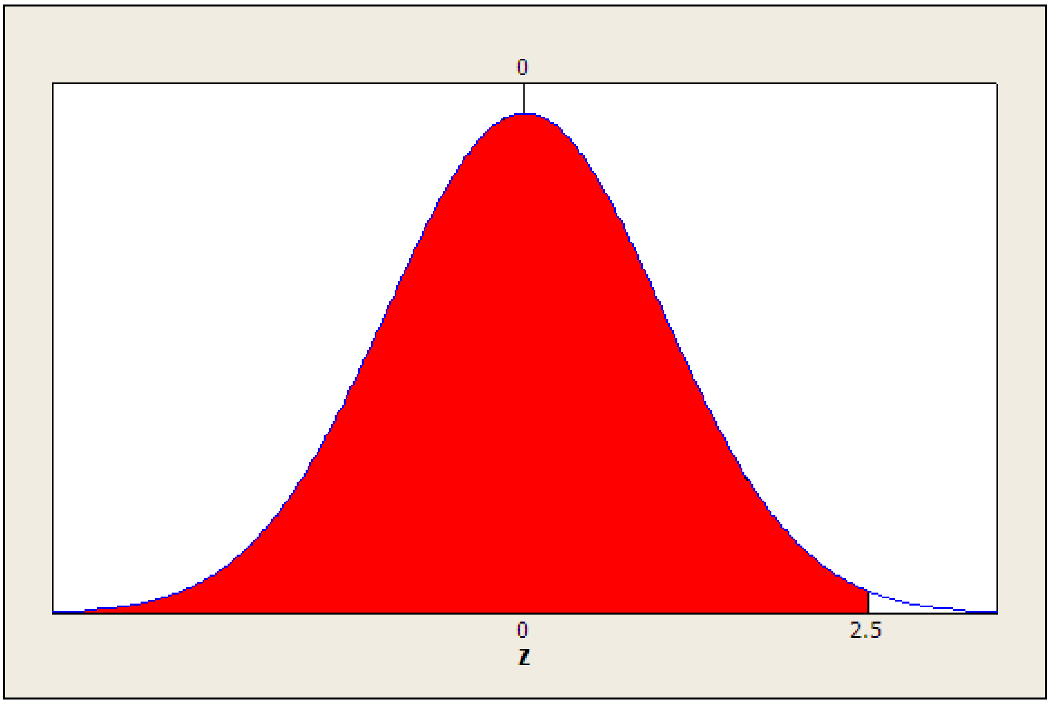
$$P(X \geq 1500) = P(Z \geq \frac{1500-\mu}{\sigma}) = P(Z \geq \frac{1500-1800}{300}) = P(Z \geq -1)$$



$$P(X \geq 1500) = P(Z \geq -1) = 0.5 + P\left(-1 \leq Z \leq 0\right) = 0.5 + P\left(0 \leq Z \leq 1\right)$$
$$= 0.5 + 0.3413 = 0.8413$$

3. احتمال أن يكون دخل أسرة أقل من 2550 ريال.

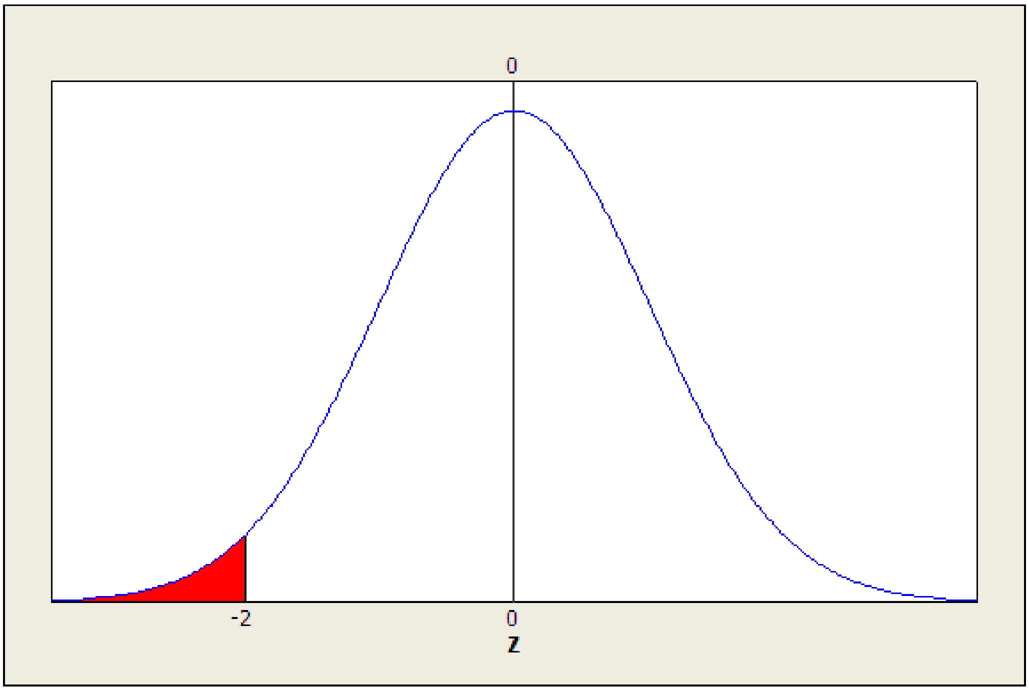
$$P(X \leq 2550) = P(Z \leq \frac{2550-1800}{300}) = P(Z \leq 2.5)$$



$$P(X \le 2550) = P(Z \le 2.5) = 0.5 + P\left(0 \le Z \le 2.5\right) = 0.5 + 0.4938 = 0.9938$$

4. احتمال أن يكون دخل أسرة أقل من 1200 ريال.

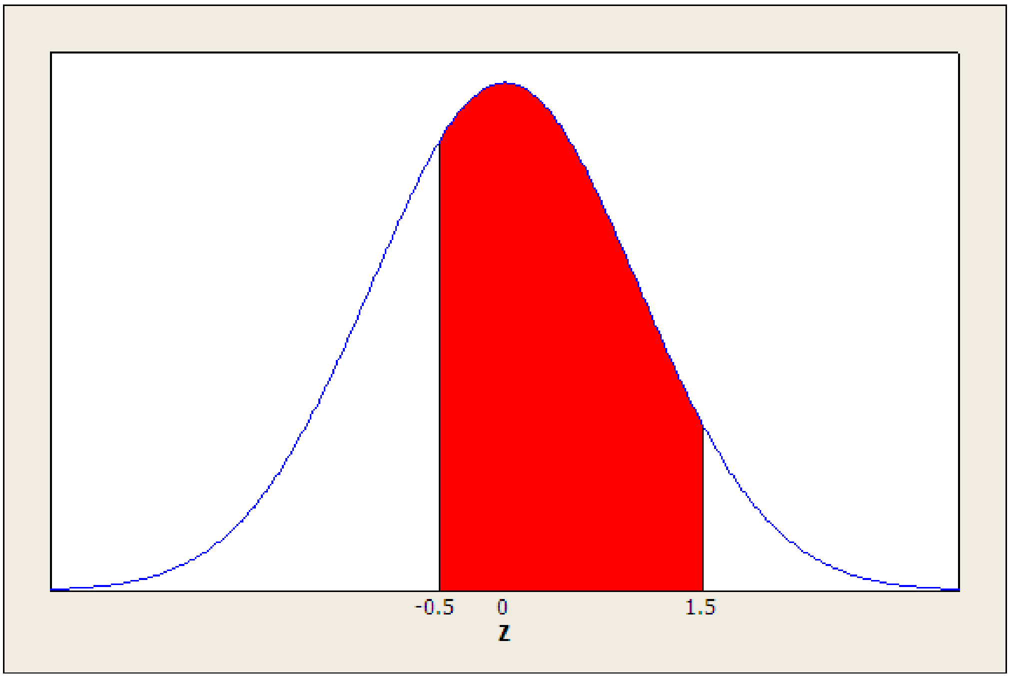
$$P(X \le 1200) = P(Z \le \frac{1200-1800}{300}) = P(Z \le -2)$$



$$P(X \le 1200) = P(Z \le -2) = 0.5 - P\left(0 \le Z \le 2\right) = 0.5 + 0.4772 = 0.0228$$

5. احتمال أن يكون دخل أسرة ينحصر بين 1650 و2250 ريال.

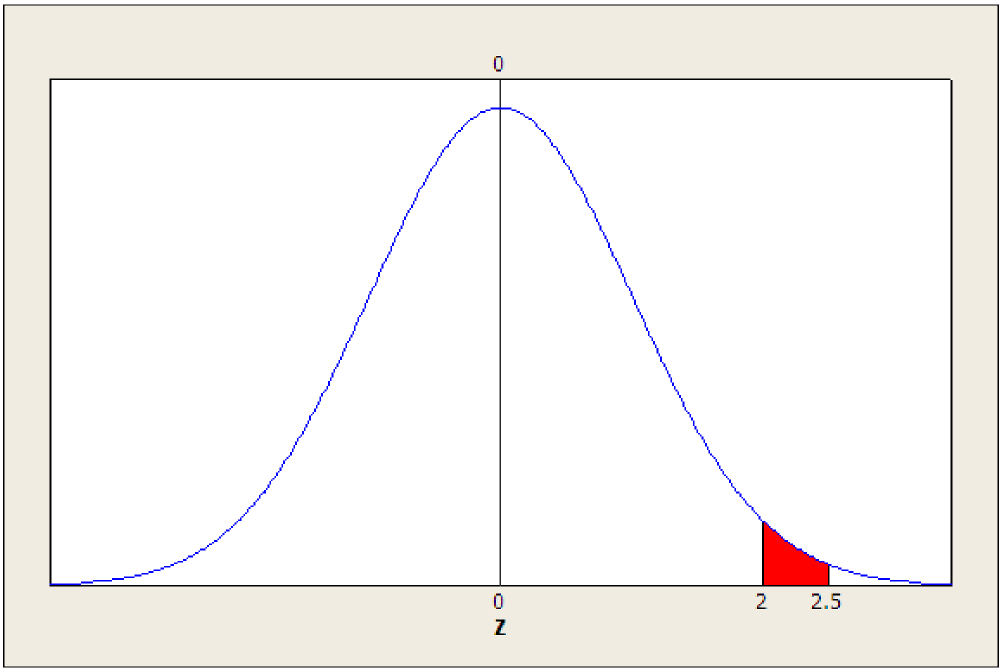
$$P(1650 \le X \le 2250) = P(\frac{1650-1800}{300} \le Z \le \frac{2250-1800}{300}) = P(-0.5 \le Z \le 1.5)$$



$$P(1650 \le X \le 2250) = P(-0.5 \le Z \le 1.5)$$
$$= P(0 \le Z \le 0.5) + P(0 \le Z \le 1.5) = 0.1915 + 0.4332 = 0.6247$$

6. احتمال أن يكون دخل أسرة ينحصر بين 2400 و2550 ريال.

$$P(2400 \le X \le 2550) = P(\frac{2400-1800}{300} \le Z \le \frac{2550-1800}{300}) = P(2 \le Z \le 2.5)$$



$$P(2400 \le X \le 2550) = P(2 \le Z \le 2.5)$$
$$= P(0 \le Z \le 2.5) - P(0 \le Z \le 2) = 0.4938 - 0.4772 = 0.0166$$

7. عدد الأسر التي يزيد دخلها عن 1500 ريال.

احتمال الأسر التي يزيد دخلها عن 1500 ريال (فقرة 2) × عدد الأسر الكلي (معطى)

= 1000 × 0.8413

= 841.3 ≈ 841 أسرة

7.4. تمارين

7.4.1. أعلنت الجامعة عن **(4)** وظائف إدارية شاغرة، فتقدم لها **(6)** رجال، **(4)** نساء. وقررت الاختيار عشوائيا أوجد:

1. التوزيع الاحتمالي لعدد النساء المختارات.
2. الانحراف المعياري للتوزيع.
3. احتمال اختيار امرأة على الأكثر.
4. احتمال اختيار رجل على الأكثر.

7.4.2. إذا كان احتمال أن يكون المصباح جيداً من بين مجموعة من المصابيح هو **(0.8)** اخترنا **(4)** مصابيح عشوائياً، أوجد:

1. التوزيع الاحتمالي لعدد المصابيح الرديئة.
2. متوسط عدد المصابيح الجيدة.
3. الانحراف المعياري للتوزيع.
4. احتمال اختيار مصباحين على الأكثر جيده.
5. احتمال اختيار مصباح رديء على الأكثر.

7.4.3. شركة لتعبئه المنتجات الزراعية، احتمال أن يكون أحد الصناديق المعبأة به سلع تالفة هو **(0.3)** اخترنا عيىنة من **(4)** صناديق. وكان التوزيع الاحتمالي لعدد الصناديق السليمة **(X)** كما هو موضح في الجدول الآتي:

عدد الصناديق السليمة X	0	1	2	3	4
الاحتمال $P(x)$	0.0081	?	0.2646	?	0.2401

1. أذكر اسم التوزيع الاحتمالي للمتغير العشوائي **(X)** و اكتب دالته الاحتمالية.
2. استكمل البيانات الناقصة في الجدول.
3. أحسب متوسط التوزيع و تباينه.
4. أحسب احتمال الحصول على **(3)** صناديق على الأقل بها سلع تالفة.

7.4.4. إذا كان متوسط عدد الحوادث الشهرية على إحدى الطرق السريعة هو **(4)** حوادث، احسب:

($e^{-4} = 0.018$)

1. احتمال عدم وقوع أي حادثة.
2. احتمال وقوع حادثين على الأقل.
3. احتمال وقوع حادثين على الأكثر.
4. تباين التوزيع.

7.4.5. إذا كانت أعمار المصابيح الكهربائية تتبع توزيع طبيعي متوسطه، ساعة $\mu = 100$ وانحرافه المعياري، ساعات $\sigma = 8$، اخترنا مصباح عشوائيا، ما هو احتمال أن:

أ. يزيد عمره عن **116** ساعة.

ب. يقل عمره عن **108** ساعة.

ج. يتراوح عمره بين **90** و**120** ساعة.

د. يتراوح عمره بين **110** و**120** ساعة.

7.4.6. إذا كان دخل **(1000)** أسرة يتبع توزيع طبيعي متوسطه **(16000)** ريال و انحرافه المعياري **(2000)** ريال:

أ. ما احتمال الحصول على دخل ينحصر بين **15000** و**18000** ريال.

ب. ما احتمال الحصول على دخل أقل من **15000** ريال.

ج. ما احتمال الحصول على دخل أكبر من **18000** ريال.

د. كم عدد الأسر التي يزيد دخلها عن **20000** ريال.

8.1. مقدمة

نعلم أن المجتمع الإحصائي يتكون من مجموعة من المفردات تكون محل اهتمام دراسة الباحث. وأن المجتمع له بعض المعالم أو الخصائص مثل **متوسط المجتمع** μ **وانحرافه المعياري** σ **ونسبة صفة معينة في المجتمع P**، وبدلاً من دراسة جميع مفردات المجتمع، فإننا نختار **عينة عشوائية** تكون ممثلة للمجتمع، ثم نقوم بدراسة مفردات العينة وحساب بعض المقاييس (الإحصاءات) منها مثل **الوسط الحسابي للعينة** $\bar{x}$ **والانحراف المعياري للعينة** s_x **ونسبة صفة معينة في العينة** r. ويمكن استخدام القيمة المحسوبة من العينة كتقدير للمعلمة المناظرة.

في هذا الباب نستعرض مبادئ أساسية تمكننا من استخدام العينات العشوائية في التعرف على خواص المجتمعات التي اختيرت منها هذه العينات.

8.2. توزيعات المعاينة

نفرض أن لدينا مجتمع، مفرداته تتبع توزيعاً احتمالياً معيناً. ونريد اختيار عينة عشوائية حجمها (n) مفردة من هذا المجتمع، كما يلي: بفرض أننا اخترنا عينة عشوائية أولى حجمها (n) مفردة من هذا المجتمع، وحسبنا وسطها الحسابي فكان $\bar{x}_1$، ثم اخترنا عينة عشوائية ثانية لها نفس الحجم (n) مفردة، وحسبنا وسطها الحسابي $\bar{x}_2$، ثم اخترنا عينة عشوائية ثالثة لها نفس الحجم (n) مفردة، وحسبنا وسطها الحسابي $\bar{x}_3$، وكررنا هذه العملية لجميع العينات التي يمكن سحبها من هذا المجتمع. فإنه سيتوفر لدينا عدد كبير من القيم للوسط الحسابي للعينات، لا نتوقع بالطبع أن تكون كلها متساويـة. وهذه القيم تكون مجتمعاً آخر عدد مفرداته أكبر بكثير من مفردات المجتمع الأصلي ، وعلى ذلك يمكن النظر إلى هذا المقياس (الوسط الحسابي) على أنه متغير عشوائي يأخذ قيماً مختلفة، ويتبع توزيعاً احتماليا معيناً قد يختلف عن توزيع المجتمع الأصلي. ويسمى هذا المجتمع الجديد **بمجتمع المتوسطات الحسابية بتوزيع المعاينة للوسط الحسابي**، وبهذا يمكن أن نعرف توزيع المعاينة لمقياس ما على أنه التوزيع الاحتمالي لمجتمع هذا المقياس.

8.2.1. مجتمع المتوسطات الحسابية للعينات

يعتمد شكل ونوع التوزيع الاحتمالي لمجتمع المتوسطات الحسابية للعينات العشوائية على توزيع المجتمع الأصلي الذي اختيرت منه هذه العينات العشوائية، والنظرية الآتية تعطي التوزيع الاحتمالي لمجتمع المتوسطات الحسابية للعينات العشوائية الكبيرة.

نظرية النهاية المركزية :

إذا كان لدينا مجتمع – غير محدود – مفرداته (X) تتبع توزيعاً احتمالياً متوسطه μ وانحرافه المعياري σ. سحبنا من هذا المجتمع عينات عشوائية حجم كل منها (n)، وكانت (n) كبيرة الحجم $(n \geq 30)$. فإن الوسط الحسابي لهذه العينات $\bar{X}$ يتبع **تقريباً** توزيعاً طبيعياً له الخصائص التالية:

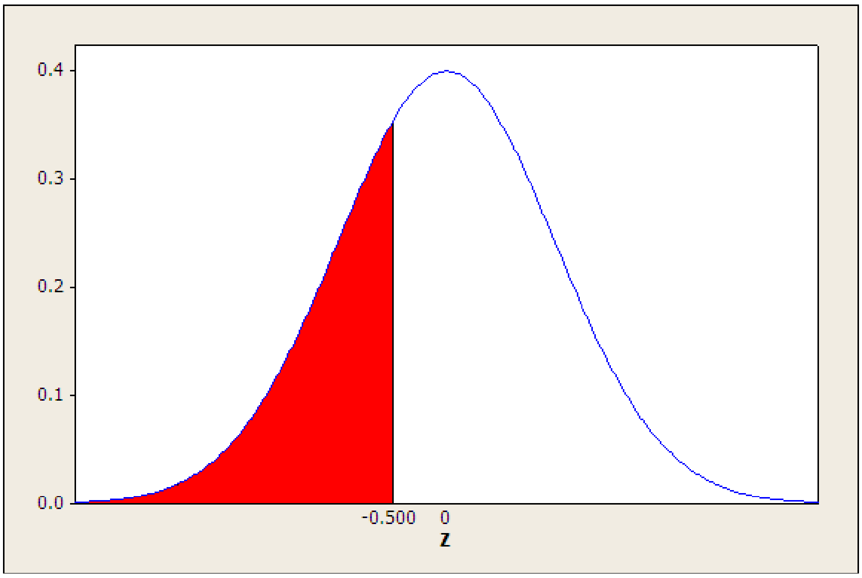
$$\mu(\bar{X}) = \mu, \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

مثال (8.1): **إذا كان بدل السكن المعطى للموظف بإحدى الشركات الكبرى يتبع توزيع طبيعي متوسطه μ (170) مئة ريال وانحرافه σ (8) مئة ريال.**

1. **اخترنا موظفاً عشوائياً فما احتمال أن يقل بدل سكنه عن 166 مئة ريال؟**
2. **سحبت عينة من (64) موظف. فما هو احتمال أن يكون متوسط بدل سكنهم أكبر من 172 مئة ريال؟**

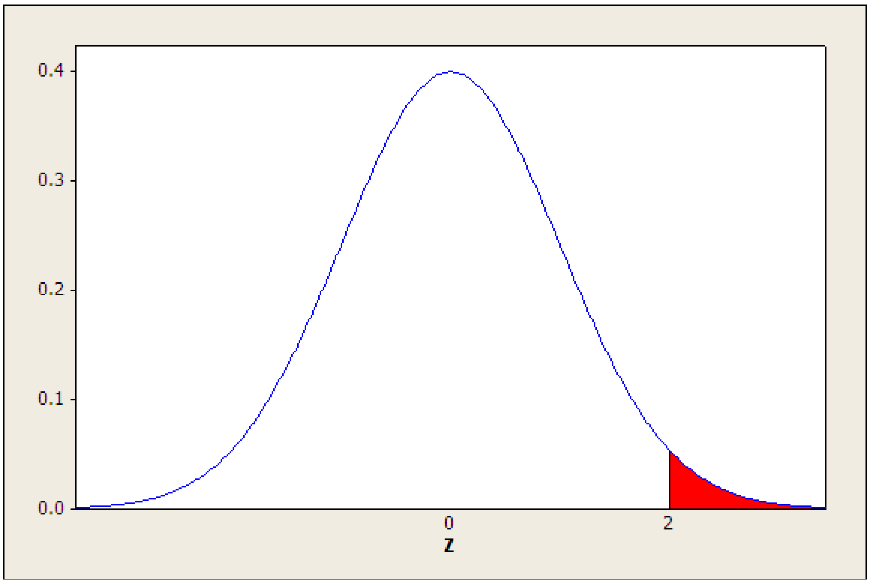
الحل:

$$1. \quad X < 166 \Rightarrow Z < \frac{166-170}{8} \Rightarrow Z < -0.5$$



$$P\left(X<166\right)=P\left(Z<-0.5\right)=0.5-0.1915=\boxed{0.3085}$$

2. $\bar{X}>172 \Rightarrow Z>\dfrac{172-\mu\left(\bar{X}\right)}{\sigma\left(\bar{X}\right)} \Rightarrow Z>\dfrac{172-\mu}{\sigma/\sqrt{n}} \Rightarrow Z>\dfrac{172-170}{8/\sqrt{64}} \Rightarrow Z>2$



$$P\left(\bar{X}>172\right)=P\left(Z>2\right)=0.5-0.4772=\boxed{0.0228}$$

مثال (**8.2**): إذا كان عدد الحوادث الأسبوعية على إحدى الطرق تتبع توزيع بواسون بمتوسط **(4)** حوادث.

1. اختير أحد الأسابيع عشوائيا فما احتمال حصول حادثين على الأكثر.
2. خلال **(64)** أسبوع، ما هو احتمال أن يكون متوسط عدد الحوادث فيها يزيد عن **(4.2)** حادثاً؟

الحل:

1. احتمال حدوث حادثين على الأكثر:

$$P(x)=\frac{e^{-\lambda}\lambda^{x}}{x!}=\frac{e^{-4}4^{x}}{x!}=\frac{(0.018)4^{x}}{x!}$$

$$\Rightarrow P\left(X\leq 2\right)=P\left(X=0\right)+P\left(X=1\right)+P\left(X=2\right)$$

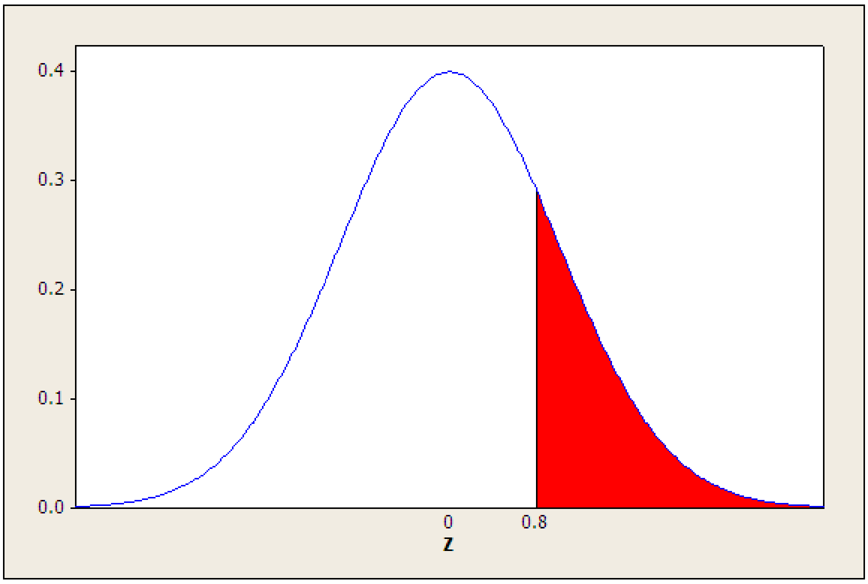
$$=0.018+0.072+0.144=\boxed{0.234}$$

2. خلال (64) أسبوع، احتمال أن يكون متوسط عدد الحوادث فيها يزيد عن (4.2) حادثاً

$\mu = \lambda = 4$

$\sigma = \sqrt{\lambda} = 2$

$$\bar{X} > 4.2 \Rightarrow Z > \frac{\bar{X} - \mu}{\sigma\sqrt{n}} \Rightarrow Z > \frac{4.2-4}{2/\sqrt{36}} \Rightarrow Z > \frac{0.2}{0.25} \Rightarrow Z > 0.8$$



$$P\left(\bar{X} > 4.2\right) = P\left(Z > 0.8\right) = 0.5 - 0.2881 = \boxed{0.2119}$$

8.2.2. مجتمع نسب الصفة في العينات

إذا كانت نسبة وجود صفة معينة في المجتمع هي P، كنسبة الأميين في إحدى المدن، نسبة المتزوجين في مجتمع ما، نسبة التالف في إنتاج أحد المصانع...الخ. واخترنا عينة عشوائية حجمها (n) مفردة، ووجدنا أن نسبة الصفة في العينة هي r. وإذا اخترنا عينات أخرى متساوية في الحجم، وحجم كل منها (n) مفردة. وحسبنا نسبة الصفة في كل عينة r، فإننا نجدها تتغير من عينة لأخرى، بمعنى أنها متغير عشوائي له توزيع معاينة، تحدده النظرية الآتية :

نظرية:

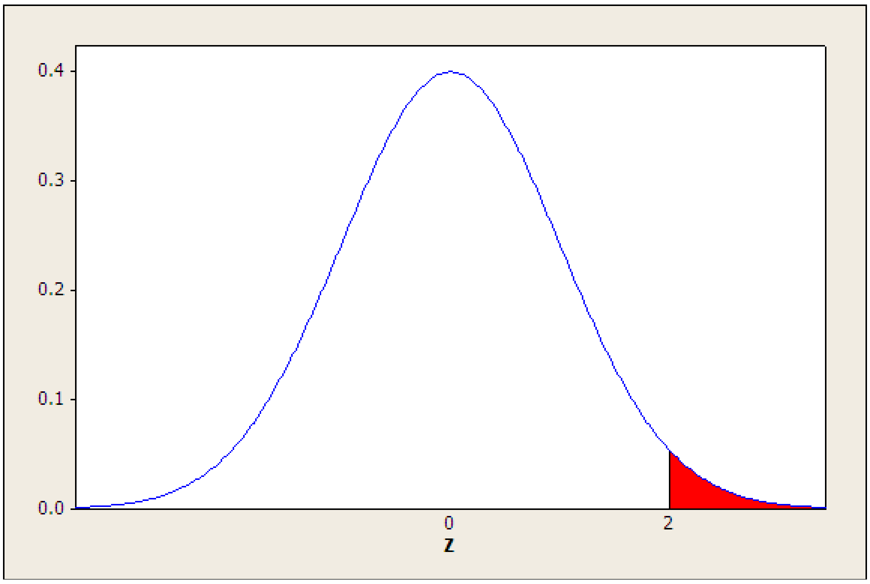
إذا كانت P تمثل نسبة صفة في مجتمع – غير محدود – وسحبنا من هذا المجتمع عينات عشوائية حجم كل منها (n)، وكانت (n) كبيرة الحجم $(n \geq 30)$. فإن r التي تمثل نسبة الصفة في العينة العشوائية، تعتبر متغيراً عشوائياً يتبع **تقريباً** توزيعاً طبيعياً له الخصائص التالية:

$$\mu(r) = \mathrm{P}, \quad \sigma(r) = \sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{n}}$$

مثال (8.3): إذا كان نسبة المصابيح الكهربائية التالفة التي ينتجها أحد المصانع هي **3%**، اشترى شخص **400** مصباح كهربائي من هذا المصنع، ما احتمال أن يجد من بينها **20** مصباح على الأقل تالفة؟

الحل: المعطيات من المسألة:

$$P = 0.03, \quad n = 400$$
$$\Rightarrow \mu(r) = P = 0.03$$
$$\Rightarrow \sigma(r) = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{0.03\times 0.97}{400}} \approx 0.01$$
$$X > 20 \overset{r=\frac{X}{n}}{\Rightarrow} r > \frac{20}{400} \Rightarrow r > 0.05 \Rightarrow Z > \frac{0.05-0.03}{0.01} \Rightarrow Z > 2$$

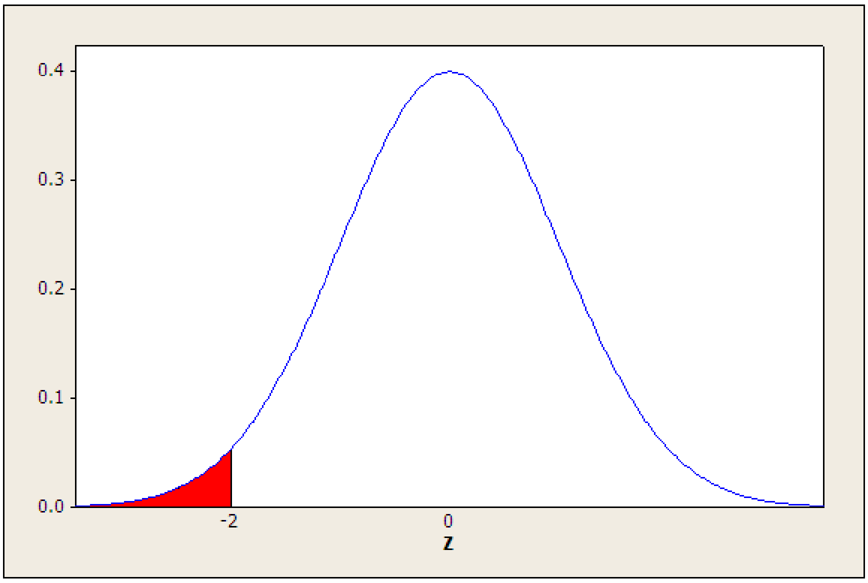


$$\Rightarrow P\left(r > 0.05\right) = P\left(Z > 2\right) = 0.5 - 0.4772 = \boxed{0.0228}$$

مثال (8.4): إذا كان نسبة الموظفين الذين حصلوا على زيادة في الراتب بإحدى الشركات هي **91%**، اخترنا **1000** موظف عشوائياً ، ما احتمال أن يجد من بينه م **70** على الأكثر لم يحصلوا على زيادة في الراتب؟

الحل:

$$P = 1 - 0.91 = 0.09, \quad n = 1000$$
$$\Rightarrow \mu(r) = P = 0.09$$
$$\Rightarrow \sigma(r) = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{0.09\times 0.91}{400}} \approx 0.01$$
$$X < 70 \overset{r=\frac{X}{n}}{\Rightarrow} r < \frac{70}{1000} \Rightarrow r < 0.07 \Rightarrow Z < \frac{0.07-0.09}{0.01} \Rightarrow Z < -2$$



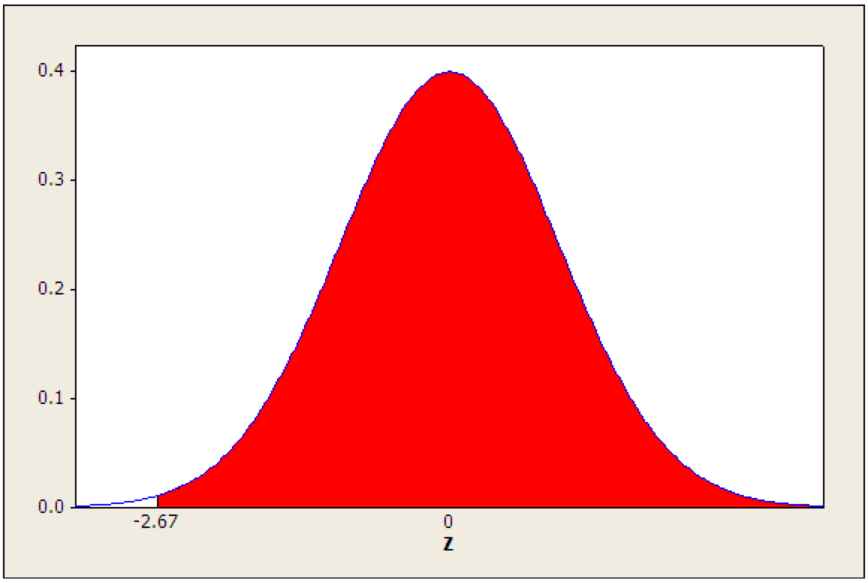
$$\Rightarrow P\left(r < 0.07\right) = P\left(Z < -2\right) = 0.5 - 0.4772 = \boxed{0.0228}$$

8.3. مسائل محلولة

8.3.1. إذا كان الأجر اليومي بالريال لعمال إحدى شركات القطاع الخاص يتبع توزيع طبيعي متوسطه μ **(90)** وانحرافه المعياري σ **(9)**. سحبت عينة من **(36)** عاملاً. ما هو احتمال أن يكون متوسط الأجر اليومي في العينة أكبر من **86**؟

الحل:

$$\bar{X} > 86 \Rightarrow Z > \frac{86-90}{9/\sqrt{36}} \Rightarrow Z > -\frac{4}{1.5} \Rightarrow Z > -2.67$$



$$P\left(\bar{X} > 86\right) = P\left(Z > -2.67\right) = 0.5 + 0.4962 = \boxed{0.9962}$$

8.3.2. إذا علم أن نسبة غياب الموظفين بإحدى الشركات هي **3%**، اخترنا عشوائياً **(400)** موظف من هذه الشركة. ما هو احتمال أن من بينهم **(16)** موظف على الأكثر غائب؟

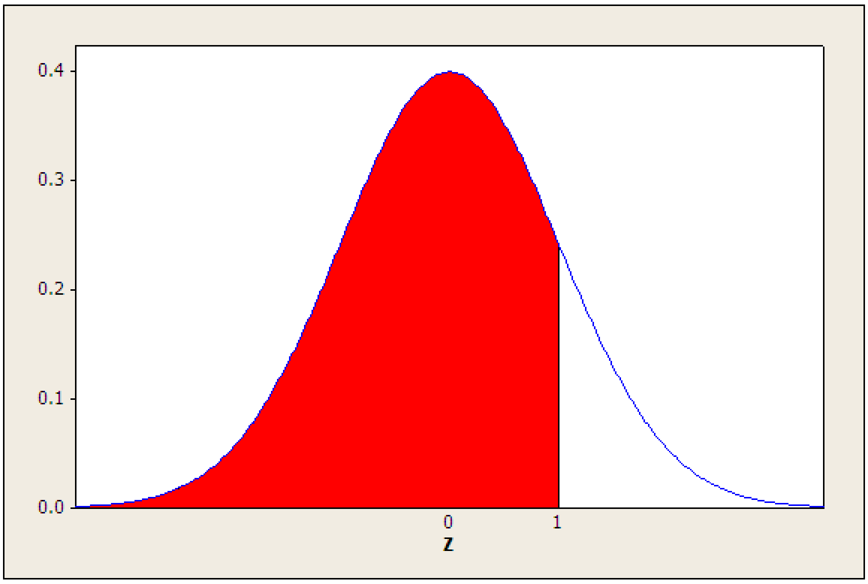
الحل:

$$\mathrm{P}=0.03,\quad n=400$$

$$\Rightarrow \mu(r)=\mathrm{P}=0.03$$

$$\Rightarrow \sigma(r)=\sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{n}}=\sqrt{\frac{0.03\times 0.97}{400}}\approx 0.01$$

$$X<16 \overset{r=\frac{X}{n}}{\Rightarrow} r<\frac{16}{400} \Rightarrow r<0.4 \Rightarrow Z<\frac{0.04-0.03}{0.01} \Rightarrow Z<1$$



$$\Rightarrow P\left(r<0.04\right)=P\left(Z<1\right)=0.5+0.3413=\boxed{0.8413}$$

8.4. تمارين

8.4.1. مصنع للمصابيح الكهربائية، متوسط عمر المصباح به **(800)** ساعة وانحرافه المعياري **(60)** ساعة. سحبت عينه من **(64)** مصباح، ما هو احتمال أن يكون متوسط عمر المصباح في العينة:

1. ينحصر بين **790** و**810** ساعة.
2. يقل عن **785** ساعة.
3. يزيد عن **820** ساعة.

8.4.2. إذا كان متوسط الدخل الشهري للأسر في إحدى المدن هو **(6000)** ريال و انحرافه المعياري **(1000)** ريال. اختيرت عينه حجمها **(100)** أسره من هذه المدينة. أوجد احتمال:

1. أن يقل متوسط دخل الأسرة في العينة عن **5800** ريال.
2. أن يزيد متوسط دخل الأسرة في العينة عن **6150** ريال.
3. أن يتراوح متوسط دخل الأسرة في العينة بين **5900** و**6200** ريال.

8.4.3. إذا كانت أعمار إحدى السلع بالأسابيع تتبع توزيعاً طبيعياً متوسطه **50** أسبوعاً و تباينه **25**.

1. اختيرت إحدى السلع عشوائيا، ما هو احتمال أن يزيد عمرها عن **60** أسبوعاً؟
2. اختيرت عينه من **36** سلعه، من هو احتمال أن يزيد متوسط أعمار السلع في العينة عن **52** أسبوعاً؟

8.4.4. إذا كان عدد الحوادث التي تقع شهرياً على إحدى الطرق السريعة تتبع توزيع بواسون بمتوسط **(9.25)**، اختيرت عينه من **(36)** شهراً. ما هو احتمال أن نجد في المتوسط ما لا يقل عن **(10.5)** حادثة؟

8.4.5. إذا كان احتمال الحصول على طفل مريض بمرض وراثي في إحدى البلاد هو **(0.3)**:

1. اخترنا **(5)** أطفال عشوائياً، أوجد:

   أ. داله التوزيع الاحتمالي لعدد الأطفال غير المرضى.

   ب. احتمال الحصول على طفل واحد على الأكثر مريض.

2. إذا تم اختيار عينة من (100) طفل، ما هو احتمال أن يكون من بينهم (25) طفل مريض على الأقل.

8.4.6. إذا كانت نسبة التأييد لرأي معين (0.7)

أ. اخترنا (4) أشخاص عشوائياً أوجد التوزيع الاحتمالي لعدد المؤيدين للرأي كذلك متوسط التوزيع و تباينه و انحرافه المعياري.

ب. إذا تم اختيار (50) شخص ما هو احتمال أن نجد من بينهم (20) شخص على الأكثر غير مؤيدين لهذا الرأي؟

8.4.7. إذا كانت نسبه المتعلمين بإحدى القرى هي % 51، ما هو احتمال الحصول على:

أ. أقل من % 45 متعلمين.

ب. ما بين 45% و54% أميين.

وذلك من بين (200) شخص مختارين عشوائياً من هذه القرية.

8.4.8. إذا كانت نسبه الموظفين السعوديين بإحدى الشركات 80%:

أ. اخترنا عشوائيا (3) موظفين، اوجد التوزيع الاحتمالي لعدد الموظفين الغير سعوديين.

ب. اوجد متوسط عدد الموظفين السعوديين.

ج. اخترنا عينه من (100) موظف، ما هو احتمال أن نجد من بينهم (90) موظف سعودي؟

9.1. التقدير

الهدف الأساسي من دراسة أي مجتمع هو إيجاد أو تقدير كاستدلال لبعض خصائصه أو معالمه، كمتوسط المجتمع μ وانحرافه المعياري σ ونسبة صفة معينة في المجتمع P وهذه تعتبر أهم معالم المجتمع. **وهذه المعالم غالباً ما تكون مجهولة ونرغب في تقديرها**. وحيث أن العينة تعتبر صورة مصغرة من المجتمع، فإننا نلجأ إلى حساب ما يقابل معالم المجتمع في العينة من مقاييس وهي على التوالي $r, s_x, \bar{x}$ وذلك من بيانات العينة، ويمكن استخدام قيمة المقياس كتقدير للمعلمة المناظر ة له في المجتمع، فمثلاً يمكن استخدام $\bar{x}$ لتقدير μ وهكذا.

كما نعلم فمن العينة يمكن حساب قيمة وحيدة هي التي تستخدم لتقدير المعلمة المقابلة لها في المجتمع، وهو ما يعرف **بتقدير النقطة**. ونعلم أيضا أن العينات المختلفة والتي لها نفس الحجم تعطينا تقديرات مختلفة الأمر الذي يترتب عليه وجود فرق بين المعلمة والتقدير، ويسمى هذا الفرق: **خطأ التقدير**.

مما سبق نجد أن تقدير النقطة نادراً ما يساوي المعلمة التي نرغب في تقديرها، لذلك فإننا نحدد فترة تحتوي على مجموعة من القيم تتضمن فيما بينها قيمة معلمة المجتمع، وتسمى هذه الفترة **بفترة الثقة**، واحتمال وقوع المعلمة في هذه الفترة يسمى **درجة الثقة**، ونرمز لدرجة الثقة بالرمز $(1 - \alpha)$، ومكمل هذه القيمة يسمى مستوى المعنوية ويرمز له بالرمز α، **فمثلاً** إذا كانت درجة الثقة 95% يكون مستوى المعنوية يساوي 0.05.

9.1.1. تقدير متوسط المجتمع (حالة العينات الكبيرة)

نستخدم العلاقة الآتية:

$$P\left[\bar{x} - Z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}\ \le \mu \le \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1 - \alpha$$

حيث تتحدد قيمة $Z_{\alpha/2}$ حسب درجة الثقة (أو مستوى المعنوية) كما يلي: إذا كانت درجة الثقة 90% فإن $Z_{\alpha/2} = 1.65$، وإذا كانت درجة الثقة 95% فإن $Z_{\alpha/2} = 1.96$، وإذا كانت درجة الثقة 99% فإن $Z_{\alpha/2} = 2.58$.

ملاحظات هامة:

1. إذا كانت قيمة المعلمة σ مجهولة نستخدم قيمة الانحراف المعياري للعينة s_x بدلاً عنها كتقدير لها.
2. في حالة أن البيانات مبوبة نستخدم العلاقات المذكورة في الفصل الثاني لحساب الوسط الحسابي والانحراف المعياري.
3. قيم $Z_{\alpha/2}$ ستستخدم في كثير من العلاقات تحت موضوع التقدير واختبارات الفروض.

مثال (9.1): مصنع لإنتاج المصابيح الكهربائية، اختير من إنتاجه عينة حجمها (100) مصباح، فإذا كان الوسط الحسابي لعمر المصباح في العينة (1200) ساعة وانحرافه المعياري (250) ساعة. قدر بدرجة ثقة 95% متوسط عمر المصباح من إنتاج المصنع كله.

الحل: حيث أن الانحراف المعياري للمجتمع σ مجهول سنستخدم الانحراف المعياري للعينة.

$$n = 100, \quad \bar{x} = 1200, \quad s_x = 250, \quad 1-\alpha = 0.95 \Rightarrow Z_{\alpha/2} = 1.96$$

$$P\left[\bar{x} - Z_{\alpha/2}\frac{s_x}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{s_x}{\sqrt{n}}\right] = 1-\alpha$$

$$P\left[1200 - (1.96)\frac{250}{\sqrt{100}} \le \mu \le 1200 + (1.96)\frac{250}{\sqrt{100}}\right] = 0.95$$

$$P\left[1200 - 49 \le \mu \le 1200 + 49\right] = 0.95$$

$$P\left[1151 \le \mu \le 1249\right] = 0.95$$

$\therefore$ متوسط عمر المصباح من إنتاج المصنع كله يتراوح بين 1151 و1249 ساعة بدرجة ثقة 95%.

مثال (9.2): فيما يلي توزيع (100) موظف بإحدى الشركات الكبرى، حسب الزيادة السنوية التي حصلوا عليها في الراتب.

الزيادة	20 -	30 -	40 -	50 -	60 -	70 -
عدد الموظفين	10	15	30	22	14	9

قدر متوسط الزيادة في الراتب للموظف في الشركة كلها بدرجة ثقة 95%.

الحل: كل من قيمة $\bar{x}, s_x$ غير معروفة ولابد من حسابها، كما يلي:

c	f	x	f x	f x²
20-	10	25	250	6250
30-	15	35	525	18375
40-	30	45	1350	60750
50-	22	55	1210	66550
60-	14	65	910	59150
70-	9	75	675	50625
Σ	**100**		**4920**	**261700**

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{4920}{100} = 49.2$$

$$s_x = \sqrt{\frac{\sum fx^2}{\sum f} - \bar{x}^2} = \sqrt{\frac{261700}{100} - (49.2)^2}$$

$$= \sqrt{196.36} \approx 14$$

$$n = 100, \quad 1-\alpha = 0.95 \Rightarrow Z_{\alpha/2} = 1.96$$

$$P\left[\bar{x} - Z_{\alpha/2}\frac{s_x}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{s_x}{\sqrt{n}}\right] = 1-\alpha$$

$$P\left[49.2 - (1.96)\frac{14}{\sqrt{100}} \le \mu \le 49.2 + (1.96)\frac{14}{\sqrt{100}}\right] = 0.95$$

$$P\left[49.2 - 2.74 \le \mu \le 49.2 + 2.74\right] = 0.95$$

$$P\left[46.46 \le \mu \le 51.94\right] = 0.95$$

أي أن متوسط الزيادة في الراتب للموظف في الشركة كلها يتراوح بين 46.46 و51.94 ريال بدرجة ثقة 95%.

9.1.2. تقدير نسبة صفة في المجتمع

نستخدم العلاقة الآتية:

$$P\left[r - Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}} \le \mathrm{P} \le r + Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\right] = 1-\alpha$$

مثال (9.3): أخذت عينة من مصنع لإنتاج قطع غيار السيارات، اختير من إنتاجه عينة حجمها (500) قطعة، وجد من بينها 100 قطعة معيبة، أوجد بدرجة ثقة 95% نسبة القطع المعيبة في إنتاج المصنع كله.

الحل:

$$r=\frac{100}{500}=0.2$$

$$P\left[r-Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\le \text{P}\le r+Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\right]=1-\alpha$$

$$P\left[0.2-(1.96)\sqrt{\frac{0.2\times 0.8}{500}}\le P\le 0.2+(1.96)\sqrt{\frac{0.2\times 0.8}{500}}\right]=0.95$$

$$P\left[0.2-0.04\le \text{P}\le 0.2+0.04\right]=0.95$$

$$P\left[0.16\le \text{P}\le 0.24\right]=0.95$$

أي أن نسبة قطع الغيار المعيبة في إنتاج المصنع كله تتراوح بين 16% و24% وذلك بدرجة ثقة 95%.

مثال (9.4): استخدم توزيع الموظفين في مثال (9.2)، لحساب ما يلي:

أ. قدر نسبة العمال الذين يقل أجرهم عن 40 ريال في المصنع كلها بدرجة ثقة 95%

ب. قدر نسبة العمال الذين يصل أجرهم 50 ريال فأكثر في المصنع كلها بدرجة ثقة 99%

الحل: الجدول التكراري كما يلي:

c	f		
20-	10	$r=\frac{10+15}{100}=0.25$	**نسبة العمال اللذين يقل أجرهم عن 40 ريال في العينة**
30-	15		
40-	30		
50-	22	$r=\frac{22+14+9}{100}=0.45$	**نسبة العمال اللذين يصل أجرهم 50 ريال فأكثر في العينة**
60-	14		
70-	9		
Σ	**100**		

أ.

$$r=0.25$$

$$P\left[r-Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\le \text{P}\le r+Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\right]=1-\alpha$$

$$P\left[0.25-(1.96)\sqrt{\frac{(0.25)(0.75)}{100}}\le \text{P}\le 0.25+(1.96)\sqrt{\frac{(0.25)(0.75)}{100}}\right]=0.95$$

$$P\left[0.17\le \text{P}\le 0.33\right]=0.95$$

أي أن نسبة العمال اللذين يقل أجرهم عن 40 ريال تتراوح بين 17% و33% بدرجة ثقة 95%.

ب.

$$r = 0.45$$

$$P\left[r - Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}} \le \mathrm{P} \le r + Z_{\alpha/2}\sqrt{\frac{r(1-r)}{n}}\right] = 1-\alpha$$

$$P\left[0.45-(2.58)\sqrt{\frac{(0.45)(0.55)}{100}} \le \mathrm{P} \le 0.45+(2.58)\sqrt{\frac{(0.45)(0.55)}{100}}\right] = 0.99$$

$$P\left[0.32 \le \mathrm{P} \le 0.58\right] = 0.99$$

أي أن نسبة العمال اللذين يصل أجرهم 50 ريال فأكثر تتراوح بين 32% و58% بدرجة ثقة 99%.

9.2. اختبار الفروض الإحصائية

يتعرض الإنسان في كثير من الحالات وفي مجالات العمل المختلفة إلى مواقف معينة تتطلب منه اتخاذ قرار بناء على معلومات محسوبة من العينة. ومن الطبيعي أن يتخذ هذا القرار بشيء من الحكمة وبأقل قدر من المخاطر. فمثلاً: قد يدعي باحث أن متوسط الدخل الشهري للأسرة في مدينة ما هو (6000) ريال. لاختبار هذا الإدعاء نختار عينة عشوائية من سكان هذه المدينة ونحسب الوسط الحسابي للدخل الشهري في العينة $\bar{x}$، ولنفرض أنه بلغ (6200) ريال. فهل الفرق بين متوسط العينة (6200) وادعاء الباحث (6000) يرجع إلى مجرد الصدفة أم أن متوسط دخل الأسرة في المدينة يختلف عن 6000 ريال؟ للوصول إلى اتخاذ قرار في هذا الموضوع، نتبع الخطوات الآتية سواء بالنسبة لاختبار فرض حول متوسط المجتمع μ أو اختبار فرض حول النسبة في المجتمع P.

تتمحور خطوات إجراء أي اختبار للفروض الإحصائية بشكل عام كما يلي:

- صياغة فرضان يسميان **فرض العدم** H_0 **والفرض البديل** H_1 حول معلمة (أو خاصية) في مجتمع الدراسة.
- حساب بعض الإحصاءات كالمتوسط، والانحراف المعياري...الخ.
- نحسب من قيم الإحصاءات **إحصاء الاختبار**.
- نتخذ القرار برفض أو عدم رفض فرض العدم.

9.2.1. اختبار فرض حول متوسط مجتمع من جانبين (حالة العينات الكبيرة)

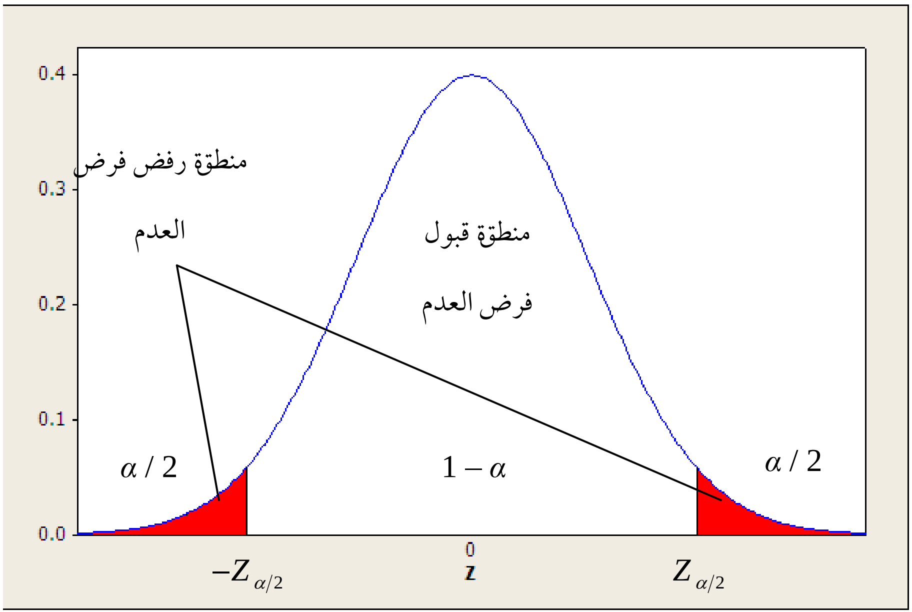
فرض العدم والفرض البديل:

$$H_0 : \mu = \mu_0 \quad (\text{vs}) \quad H_1 : \mu \neq \mu_0$$

إحصاء الاختبار (Z):

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

نحدد مكان وقوع قيمة إحصاء الاختبار (Z) في الرسم التوضيحي 16.



رسم توضيحي 16: منطقة قبول ورفض فرض العدم

إذا وقعت قيمة (Z) في منطقة القبول $(1 - \alpha)$ نقبل فرض العدم H_0 . وإذا وقعت قيمة (Z) في منطقة الرفض نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 .

ملاحظات هامة:

1. إذا كانت قيمة المعلمة σ مجهولة نستخدم قيمة الانحراف المعياري للعينة s_x بدلاً عنها كتقدير لها. ولو كانت البيانات مبوبة نستخدم العلاقات المذكورة في الفصل الثاني لحساب الوسط الحسابي والانحراف المعياري.
2. قيم $Z_{\alpha/2}$ نفسها المستخدمة في التقدير.

مثال (**9.5**): **إذا كان متوسط الزيادة في أجور العاملين في إحدى المؤسسات عام 1998 هو (36) عشرات الريالات، وفي عام 2001م أخذت عينة من (64) فرداً من العاملين في هذه المؤسسة، فوجد أن الوسط الحسابي للزيادة في أجورهم (40) عشرات الريالات والانحراف المعياري (8) عشرات الريالات. هل يدل ذلك على أن متوسط الزيادة في أجور العاملين في المؤسسة عام 2001م قد اختلف عن متوسط الزيادة في الأجور عام 1998م؟ وذلك عند مستوى معنوية $\alpha = 0.05$.**

الحل:

فرض العدم والفرض البديل:

$$H_0 : \mu = 36 \quad \text{(vs)} \quad H_1 : \mu \neq 36$$

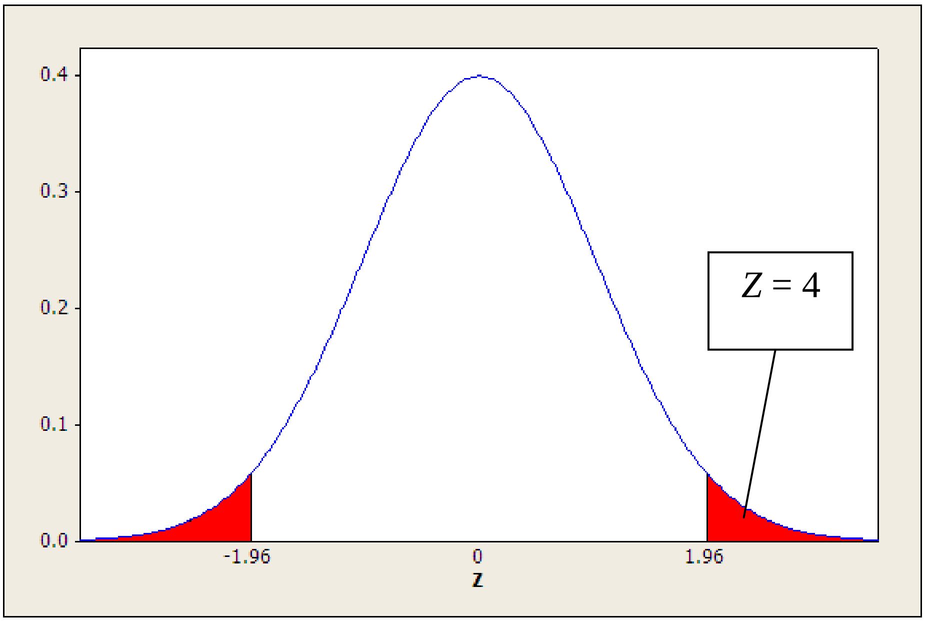
معطيات:

$$n = 64 \qquad \bar{x} = 40 \qquad s_x = 8$$

إحصاء الاختبار (Z):

$$Z = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}} = \frac{40 - 36}{8 / \sqrt{64}} = 4$$

القرار:



نرفض فرض العدم H_0 ونقبل الفرض البديل H_1؛ أي أن متوسط الزيادة في الأجور عام 2001م قد اختلف عن متوسط الزيادة في الأجور عام 1998م. وذلك عند مستوى معنوية $\alpha = 0.05$.

9.2.2. اختبار فرض حول نسبة صفة في المجتمع من جانبين

فرض العدم والفرض البديل:

$$H_0 : \mathrm{P} = \mathrm{P}_0 \quad \text{(vs)} \quad H_1 : \mathrm{P} \neq \mathrm{P}_0$$

إحصاء الاختبار (Z):

$$Z = \frac{r - \mathrm{P}_0}{\sqrt{\frac{\mathrm{P}_0\left(1 - \mathrm{P}_0\right)}{n}}}$$

نحدد مكان وقوع قيمة إحصاء الاختبار (Z) في الرسم التوضيحي. حيث إذا وقعت قيمة (Z) في منطقة القبول نقبل فرض العدم H_0. وإذا وقعت قيمة (Z) في منطقة الرفض نرفض فرض العدم H_0 ونقبل الفرض البديل H_1.

مثال (9.6): من بين (900) شخص وجد أن عدد المؤيدين منهم لرأي معين هو (738) شخص. اختبر الفرض أن نسبة المؤيدين لذلك الرأي في المجتمع هو (0.8)، وذلك عند مستوى معنوية $\alpha = 0.01$.

الحل:

فرض العدم والفرض البديل:

$$H_0 : \mathrm{P} = 0.8 \quad \text{(vs)} \quad H_1 : \mathrm{P} \neq 0.8$$

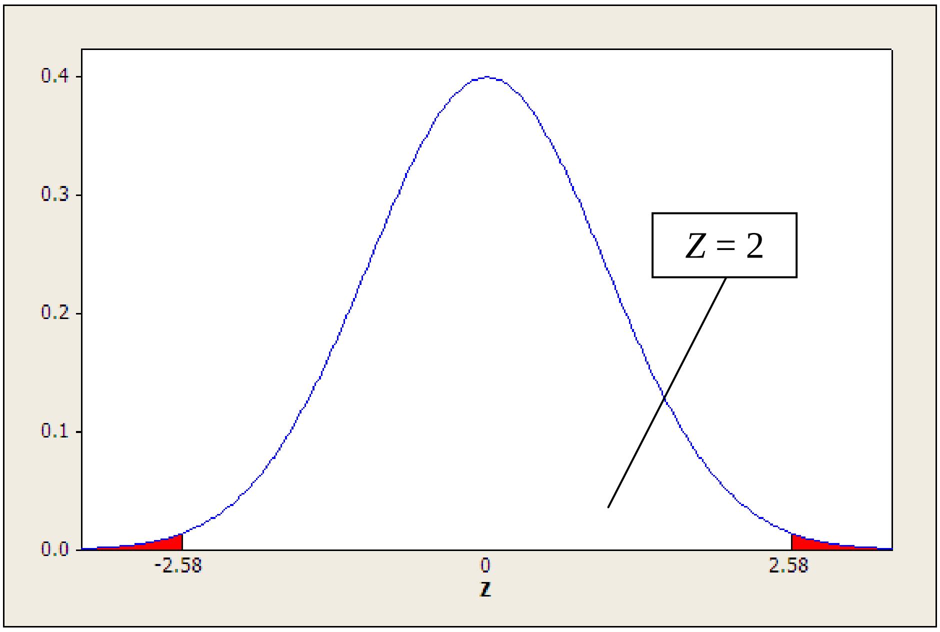
معطيات:

$$n = 900 \qquad\qquad r = \frac{738}{900} = 0.82$$

إحصاء الاختبار (Z):

$$Z = \frac{r - \mathrm{P}_0}{\sqrt{\frac{\mathrm{P}_0(1 - \mathrm{P}_0)}{n}}} = \frac{0.82 - 0.8}{\sqrt{\frac{0.8 \times 0.2}{900}}} = \frac{0.02}{0.01} = 2$$

القرار:



تقبل فرض العدم H_0 ؛ نسبة المؤيدين لذلك الرأي في المجتمع هو (0.8)، وذلك عند مستوى معنوية $\alpha = 0.01$.

9.3. تمارين

9.3.1. الجدول الآتي يوضح توزيع عينه من (100) طالب حسب فئات الذكاء:

فئات الذكاء	50-	60-	70-	80-	90-
عدد الطلبة	8	12	30	40	10

1. **قدر متوسط الذكاء في مجتمع الطلاب بدرجه ثقة 99%.**
2. **أختبر الفرض القائل أن متوسط الذكاء في مجتمع الطلاب لا يساوي (76) درجة عند مستوى معنوية $\alpha = 0.01$.**
3. **قدر نسبة عدد الطلاب الذين يقل ذكائهم عن (70) درجه، بدرجه ثقة 95%.**
4. **اختبر الفرض القائل أن نسبة عدد الطلاب الذين يصل ذكائهم (80) درجه فأكثر، لا تساوي 55%، وذلك عند مستوى معنوية $\alpha = 0.05$.**

9.3.2. الجدول الآتي يوضح توزيع عينه من (100) أسره حسب فئات الدخل السنوي بآلاف الريالات في مدينه ما:

فئات الدخل	10-	14-	18-	22-	26-	30-
عدد الأسر	12	16	30	25	12	5

1. قدر نسبه عدد الأسر التي يصل دخلها (22000) ريال فأكثر في المدينة بدرجة ثقة 99%.
2. أختبر الفرض القائل أن نسبه عدد الأسر التي يقل دخلها عن (18000) ريال لا تساوي 25%، وذلك عند مستوى معنوية $\alpha = 0.05$.

9.3.3. الجدول الآتي يوضح توزيع عينة من (60) موظفاً بإحدى المؤسسات حسب فئات العمر

فئات العمر	20-	24-	28-	32-	36-	40-
عدد الموظفين	3	9	15	18	10	5

1. أختبر الفرض القائل أن متوسط أعمار العاملين بالمؤسسة لا يساوي (30) سنة، عند مستوى معنوية $\alpha = 0.05$.
2. قدر نسبه عدد الموظفين بالمؤسسة الذين تقل أعمارهم عن (28) سنه، بدرجة ثقة 99%.

9.3.4. الجدول الآتي يوضح توزيع عينه من العقارات بإحدى المدن حسب المساحة بعشرات الأمتار المربعة:

فئات المساحة	15-	19-	23-	27-	31-
عدد العقارات	6	14	42	10	8

1. قدر متوسط مساحة العقارات في المدينة بدرجة ثقة 99%.
2. أختبر الفرض القائل أن نسبه العقارات التي تقل مساحتها عن (230) متر مربع في المدينة لا تساوي 17%، وذلك عند مستوى معنوية $\alpha = 0.05$.

9.3.5. إذا توفر لدينا البيانات الآتية عن سعر عينة من (100) منتج من نوع ما بالريال:

$$\sum f \cdot x = 1028 \quad , \quad \sum f \cdot x^2 = 11120$$

عدد المنتجات التي يزيد سعرها عن (13) ريال يساوي (15) سلعة.

1. أختبر الفرض القائل أن متوسط أسعار المنتجات لهذا النوع يختلف عن **(10)** ريال، وذلك عند مستوى معنوية $\alpha = 0.01$.
2. اختبر الفرض القائل أن نسبة عدد المنتجات التي يقل طولها عن **(13)** ريال لا تساوي **80%** من المنتجات وذلك عند مستوى معنوية $\alpha = 0.05$.

9.3.6. إذا كان متوسط أعمار المصابيح التي ينتجها أحد المصانع هو **(950)** ساعة مع انحراف معياري **(120)** ساعة. يدعى مدير المصنع أنه أدخل تعديلات على وسائل الإنتاج أدت إلى اختلاف أعمار المصابيح. لاختبار صحة هذا الادعاء ، أخذت عينه من **(81)** مصباح وأضيئت حتى احترقت كلها وحسب متوسط أعمارها فكان **(1100)** ساعة. فهل يمكنك تأييد ادعاء المدير؟ وذلك عند مستوى معنوية $\alpha = 0.05$.