مفاهيم أساسية



1.1 مقدمة

مفاهيم أساسية – (1.1 مقدمة)

- **علم الإحصاء** (Statistics) : هو العلم الذي يبحث في تصميم أساليب جمع البيانات و التقنيات المختلفة لتنظيم و تصنيف وعرض هذه البيانات ، و تلخيص هذه البيانات في صورة مؤشرات رقمية لوصف و قياس خصائصها الأساسية ، و تحليلها بغرض اتخاذ القرارات المناسبة .
- **المجتمع** (Population) : هو المجموعة الكلية لمفردات الدراسة سواء كانت أفراد أو أشياء .
- العينة (Sample) : هي مجموعة جزئية من مفردات المجتمع محل الدراسة يتم اختيارها بحيث تكون ممثلة للمجتمع تمثيل صحيح .
- مثال : في دراسة لتحديد نسبة المتميزين بين طلاب جامعة الملك عبدالعزيز يكون *المجتمع* جميع طلاب الجامعة بينما اختيار 400 طالب بطريقة تضمن تمثيل كل كلية من كليات الجامعة بشكل عادل يعتبر *عينة*.

مفاهيم أساسية – (1.1 مقدمة)

- ينقسم علم الإحصاء إلى فرعين أساسيين :

1) **الإحصاء الوصفي** (Descriptive Statistics) : هو مجموعة الطرق و الأساليب التي تستخدم في تنظيم و عرض و تلخيص البيانات و استكشاف خصائصها الأساسية و تلخيصها في صورة مؤشرات رقمية .

مثال : عبارة "متوسط درجات الطلاب في مادة الرياضيات 87 في الفصل الأول".

2) **الإحصاء الاستدلالي** (Inferential Statistics) : هو مجموعة الطرق و الأساليب التي تستخدم في تعميم نتائج العينة على خصائص المجتمع الذي سحبت منه . و قياس العلاقات بين الخصائص المختلفة للمجتمع و التنبؤ بالقيم المستقبلية لهذه الخصائص .

مثال : عبارة " من المتوقع أن يزيد عدد سكان المملكة بعد خمس سنوات بمقدار 4 ملايين نسمة ".

مفاهيم أساسية – (1.1 مقدمة)

- يعتبر الأسلوب الإحصائي الوسيلة الأساسية في دراسة الظواهر الاقتصادية و قياس العلاقات بينهما، و هو وسيلة للتنبؤ بالقيم المستقبلية لهذه الظواهر.
- تتعدد استخدامات الإحصاء في مجال العلوم الاجتماعية فهناك الإحصاء الاجتماعي ، و الإحصاء النفسي ، و الإحصاء الجغرافي فعلى سبيل المثال هناك العديد من الأساليب الكمية التي تقيس الاتجاهات السلوكية و النفسية لدى الأفراد.

مفاهيم أساسية – (1.2 البيانات و أنواعها)

- **البيانات (Data)** : هي مجموعة القيم التي يتم جمعها من مفردات المجتمع أو العينة لخاصية (متغير) معينة.
- تنقسم البيانات إلى نوعين رئيسيين ، هما :

1) **البيانات النوعية (الوصفية) (Qualitative Data)** : هي البيانات الت يمكن حصرها في عدة أوجه وصفية و لا يمكن إجراء عمليات حسابية عليها كالجمع و الطرح .

مثال / الجنسية : (سعودي – كويتي- قطري – مصري – اماراتي - .....).

2) **البيانات الكمية (Quantitative Data)** : هي البيانات التي يمكن الحصول عليها في شكل أعداد و يمكن ترتيبها .

مثال / الرواتب الشهرية – درجات الحرارة -..... .

مفاهيم أساسية – (1.2 البيانات و أنواعها)

- **تنقسم البيانات الكمية إلى قسمين هما :**

1) **بيانات كمية منفصلة (Discrete Data) :** هي البيانات التي يمكن عدها حتى و لو لم تأخذ قيمًا صحيحة .

مثال / عدد الحجرات في المسكن – عدد أفراد الأسرة .

2) **بيانات كمية متصلة (Continuous Data) :** هي البيانات التي لا يمكن عدها و إنما يتم الحصول عليها عن طريق القياس و تأخذ أي قيمة داخل مدى معين سواء كانت صحيحة أو كسرية.

مثال / الدخل الشهري – أوزان الطلاب .

مفاهيم أساسية – (1.2 البيانات و أنواعها)

- **مستويات قياس البيانات :** يمكن تفصيل مستويات القياس المختلفة بأربع مستويات من الأدنى إلى الأعلى كما يلي :

1) **المستوى الاسمي (Nomial Level) :** مجموعة من الأوجه أو الصفات التي يأخذها المتغير الوصفي . و يمكن أن تعطى الصفات أرقام تعكس مدلول الصفة و لكن ليس لها معنى رياضي في مفهوم أكبر أو أصغر (مثل فصيلة الدم و الجنسية).

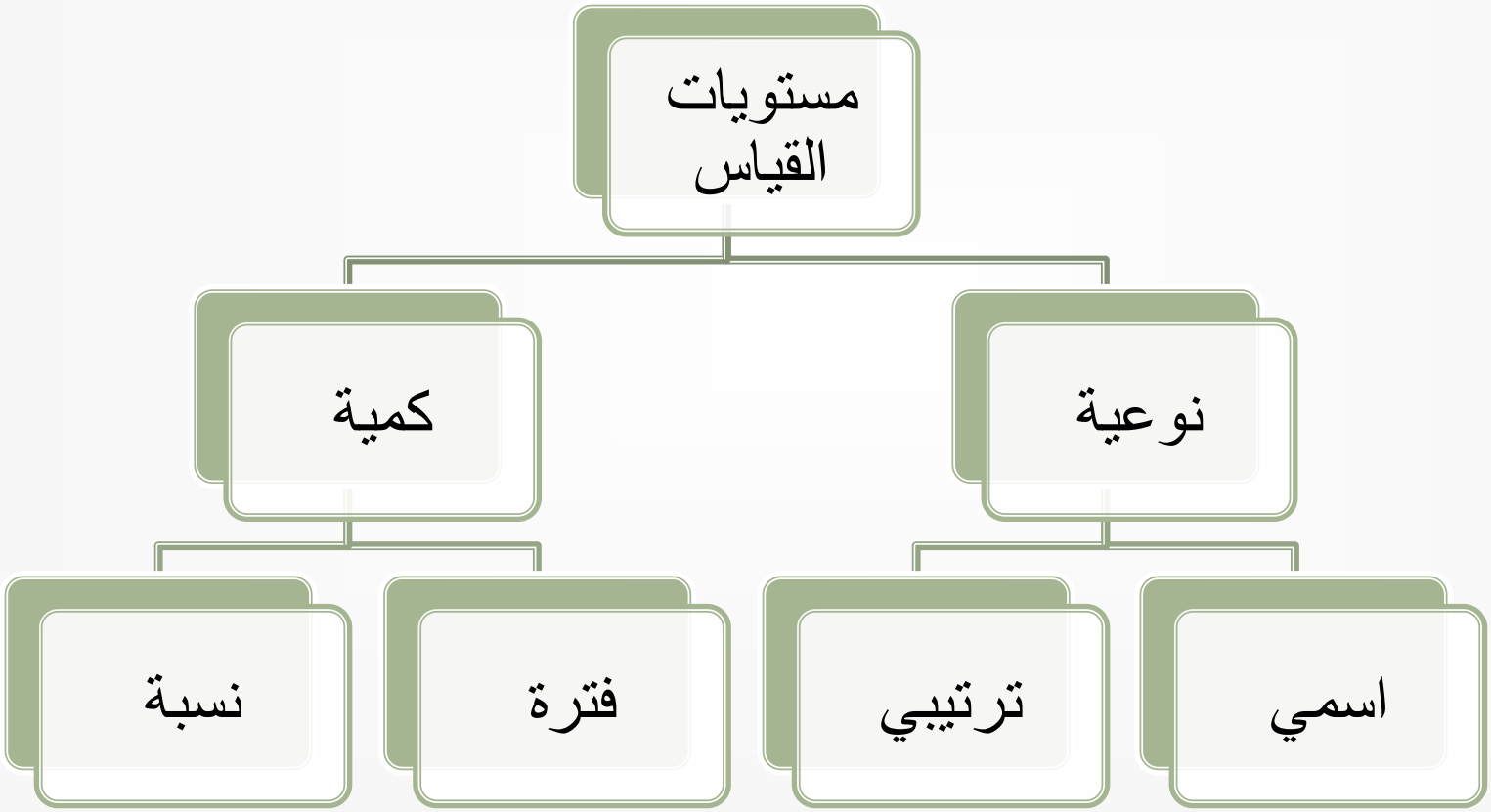
2) **المستوى الترتيبي (التفضيلي) (Ordinal Level) :** مجموعة من الأوجه أو الصفات التي يأخذها المتغير الوصفي مع إمكانية ترتيبها . و يمكن أن تعطى الصفات أرقام تعكس مدلول الصفة و لها معنى رياضي في مفهوم أكبر أو أصغر و لكن لا تعكس معنى حقيقي للفروق (مثل المستوى التعليمي و مدى الموافقة على رأي معين).

مفاهيم أساسية – (1.2 البيانات و أنواعها)

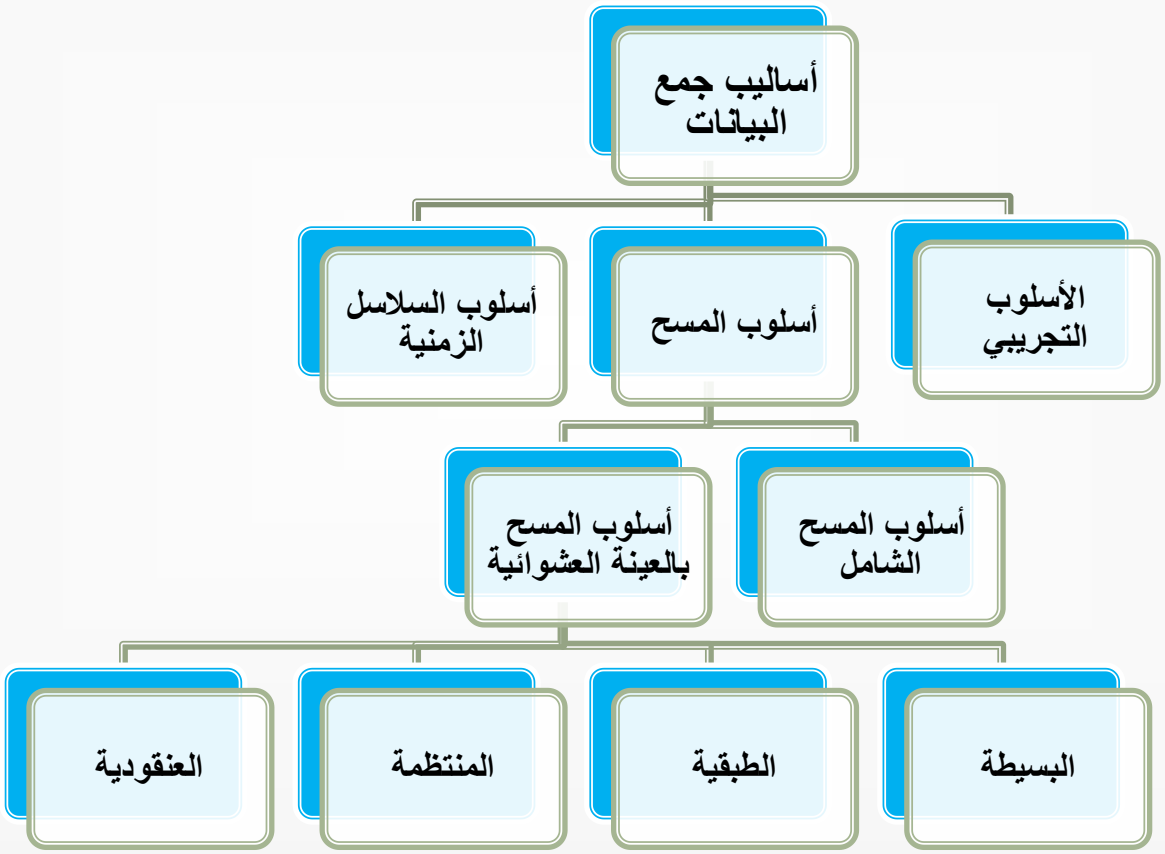
3) مستوى الفترة (Interval Level) : مجموعة من الأعداد أو القيم التي يأخذها المتغير الكمي و تعطى الصفات أرقام تعكس مدلول الصفة و لها معنى رياضي أكبر أو أصغر و تعكس معنى حقيقي للفروق ، و لكن ليس للصفر معنى حقيقي أي أن المتغير إذا كان مساويًا للصفر فلا يعني هذا انعدام الصفة (مثل درجة الحرارة و درجة امتحان الذكاء).

4) مستوى النسبة (Ratio Level) : مجموعة من الأعداد أو القيم التي يأخذها المتغير الكمي و تعطى الصفات أرقام تعكس مدلول الصفة و لها معنى رياضي أكبر أو أصغر و تعكس معنى حقيقي للفروق ، و الصفر له معنى حقيقي أي أن المتغير إذا كان مساويًا للصفر فهذا يعني انعدام الصفة (مثل الوزن و الطول).

مفاهيم أساسية – (1.2 البيانات و أنواعها)



مفاهيم أساسية – (1.3 منهجية علم الإحصاء)



مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

- **جمع البيانات** : من المهم أن يعلم الباحث أساليب جمع البيانات ، و هي كما يلي :

أولاً / الأسلوب التجريبي : يتم الحصول على البيانات عن طريق تصميم تجربة ، يتم فيها قياس تأثير العامل محل الاهتمام مع ثبات العوامل الأخرى ، حيث نحصل على البيانات في هذه الحالة عن طريق المشاهدة.

مثال/ (اقتصادي) الحصول على بيانات عن طريق تطبيق عدة سياسات تسويقية بهدف اختيار السياسة الأفضل.

ثانيًا/ أسلوب المسح : نحصل على البيانات في هذه الحالة من السجلات و التقارير و قواعد البيانات و الانترنت ، أو عن طريق الاستبيانات و المقابلات الشخصية .

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

- ينقسم أسلوب المسح إلى نوعين هما :

1) **أسلوب المسح الشامل** : يتم جمع البيانات من كل مفردات المجتمع محل الدراسة . مثل دراسة آراء جميع طلاب جامعة الملك عبدالعزيز حول موضوع معين.

2) **أسلوب المسح بالعينة العشوائية** : تجمع البيانات من بعض مفردات المجتمع محل الدراسة . مثل دراسة تأثير التلفاز على الأطفال.

- هناك العديد من أنواع العينة العشوائية ، منها :

- **العينة العشوائية البسيطة** : و هي التي تعطي كل مفردة من مفردات المجتمع نفس الفرصة في الاختيار .

مثال/ اختيار 5 موظفين من أصل 50 موظف بطريقة عشوائية ، حيث يتم ترقيم الموظفين ترقيمًا تسلسليًا من 1 إلى 50 و من ثم استخدام الحاسب الآلي في اختيار خمسة أرقام بشكل عشوائي.

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

- **العينة العشوائية الطبقية** : يتم تقسيم المجتمع محل الدراسة إلى مجموعات متجانسة و غير متداخلة تسمى طبقات مثل كليات أو محافظات أو النوع . ثم نقوم بسحب عينة عشوائية بسيطة من كل طبقة.

مثال/ عند اجراء دراسة عن مستوى الثقافة لدى طلاب جامعة الملك عبدالعزيز يتم تقسيم الجامعة إلى مجموعتين أو طبقتين و هما الكليات العلمية و الكليات الأدبية و يتم اختيار عينة عشوائية بسيطة من كل طبقة تتناسب مع حجم الطلاب داخل تلك الكلية.

- **العينة العشوائية المنتظمة** : يتم تقسيم مفردات المجتمع إلى مجموعات عددها مساوٍ لعدد مفردات العينة التي نريد اختيارها ، ثم نختار مفردة من المجموعة الأولى بشكل عشوائي فإذا كان الاختيار مثلاً وقع على المفردة الثالثة . فإننا نختار المفردة الثالثة من كل مجموعة حتى يكتمل حجم العينة تي نريدها.

مثال/ إذا كان لدينا مركز لبيع السيارات المستخدمة يتكون من 1000 سيارة و نرغب في اختيار عينة عشوائية منتظمة من 10 سيارات ، نقوم بما يلي:

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

1- إعطاء كل سيارات المركز أرقامًا متسلسلة.

2- يتم تقسيم المركز إلى 10 مجموعات ليتناسب مع حجم العينة المطلوب فتكون المجموعة الأولى تمثل المفردات من 1 إلى 100 ، و المجموعة الثانية تمثل المفردات من 101 إلى 200 ، و هكذا.

3- نختار رقمًا عشوائيًا باستخدام الحاسب الآلي أو باستخدام جداول الأرقام العشوائية من بين الأرقام من 1 إلى 100 و ليكن الرقم 48 .

4- في هذه الحالة تكون السيارة الأولى في العينة هي التي تحمل الرقم 48 ، و السيارة الثانية تحمل الرقم 148 و الثالثة 248 و هكذا حتى العاشرة التي تحمل الرقم 948 و بالتالي تتكون العينة من السيارات التي تحمل الأرقام :

48 ، 148 , 248 ، 348 ، 448 ، 548 ، 648 ، 748 ، 848 ، 948 .

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

- **العينة العشوائية العنقودية :** يمكن اختيار هذه العينة بإحدى الطرق الآتية حسب طبيعة المجتمع محل الدراسة :

1- على مرحلة واحدة و تسمى العينة العشوائية العنقودية ذات المرحلة الواحدة و ذلك ندما يكون المجتمع محل الدراسة على هيئة تجمعات أو عناقيد و كل تجمع أو عنقود يحتوي على عدد من مفردات المجتمع . و يتم اختيار العينة في هذه الحالة باختيار عدة عناقيد عشوائيا من بين كل العناقيد أو التجمعات التي يتكون منها المجتمع ثم ندرس جميع المفردات التي توجد في هذه المجموعات التي يتم اختيارها عشوائياً . فمثلاً إذا اردنا دراسة كفاءة الأداء لمستشفيات المملكة فيتم تقسيم المملكة إلى مناطق (كل منطقة تمثل عنقود) و يتم اختيار منطقتين كعينة عشوائية بسيطة ثم نقوم بدراسة المستشفيات الموجودة في المنطقتين .

2- على أكثر من مرحلة و تسمى العينة العشوائية العنقودية متعددة المراحل و ذلك عندما يكون المجتمع محل الدراسة على هيئة مجموعات أو عناقيد تشكل كل مجموعة عنقوداً يتفرع منه أيضاً العديد من المجموعات أو العناقيد . فمثلاً إذا أردنا تقدير متوسط دخل الأسرة في المملكة فإنه يستلزم تقسيم المملكة إلى مجموعة من المحافظات و تقسم كل محافظة إلى مجموعة من المدن و أيضاً يتم تقسيم كل مدينة إلى مجمعة من الأحياء السكنية. فيتم اختيار عينة عشوائية من المحافظات كمرحلة أولى ثم في المرحلة الثانية يتم اختيار عينة عشوائية من المدن داخل كل محافظة تم اختيارها في المرحلة الأولى ثم يتم اختيار عينة عشوائية من الأحياء السكنية داخل كل مدينة تم اختيارها في المرحلة الثانية .

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

ثالثًا / أسلوب السلاسل الزمنية : يتم الحصول على البيانات عن طريق رصد البيانات التي تعبر عن ظاهرة ما عند نقاط زمنية متتالية . مثل كمية الصادرات السنوية ، حجم التعاملات الربع سنوي في البورصة ، عدد المرضى الشهري في العيادات النفسية.

- يمكن أن تتعرض البيانات إلى نوعين من الأخطاء عند جمعها هما :

1. **خطأ التحيز** : هو الخطأ الذي يحدث عند جمع البيانات و مصدر هذا الخطأ إما من الباحث أو من مفردات المجتمع محل الدراسة.
2. **خطأ المعاينة العشوائية** : هو الخطأ الذي يحدث عند إجراء الدراسة الإحصائية بأسلوب العينة العشوائية و يرجع فقط ألى الصدفة و ليس لأخطاء من الباحث أو المبحوث.

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

- **بعض البرامج (الحزم) الإحصائية :** أصبحت البرامج الإحصائية من أهم الأدوات المستخدمة لإجراء التحليلات الإحصائية خصوصًا إننا في زمن تطور فيع العلم، مما أدى إلى زيادة أحجام البيانات التي تحتاج إلى معالجة، و تمتاز البرامج الإحصائية بدقتها العالية في تحليل البيانات و سرعة معالجتها ، و توجد العديد من البرامج الإحصائية مثل : MINITAB , SAS , SPSS , R , S-PLUS.
- برنامج مايكروسوفت إكسل يوفر حزمة بسيطة من البرامج الإحصائية التي تساعد في إجراء التحليلات الإحصائية الأساسية.

مفاهيم أساسية – (1.3 منهجية علم الإحصاء)

- **تدريبات :**

1) ما هو علم الإحصاء ؟

2) اعط مثال لبيانات وصفية و أخرى كمية ؟

3) ما هي أساليب جمع البيانات ؟

4) ما الفرق بين أسلوب المسح الشامل و أسلوب المسح بعينة عشوائية ؟

عرض و تنظيم البيانات



1.1 التوزيعات التكرارية

عرض و تنظيم البيانات – (1.1 التوزيعات التكرارية)

- **التوزيعات التكرارية** (Frequency Distribution) : هي عبارة عن جداول لجميع الأوجه أو القيم التي يمكن أن يأخذها المتغير موضع الدراسة و عدد المفردات التي تمثل تكرارات مناظرة لكل وجه أو قيمة.

مثال 1 (البيانات الكمية المنفصلة) : اختار أحد مرشدي الطلاب بإحدى المدارس الثانوية عينة عشوائية مكونة من ثلاثين طالباً ، لمعرفة مستوى الغياب خلال ثاني أسبوع في رمضان عام 1427 هـ لدى الطلاب في هذه المدرسة . و المطلوب تبويب البيانات التي حصل عليها في توزيع تكراري و نسبي ، إذا كانت عدد أيام الغياب كالتالي :

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

عرض و تنظيم البيانات – (1.1 التوزيعات التكرارية)

الحل /

عدد أيام الغياب الفئات (c)	العلامات	عدد الطلاب التكرار (f)	التكرار النسبي (p)	نسبة الطلاب
0	𝍸 𝍸 //	12	0.400	40.0%
1	𝍸 /	6	0.200	20.0%
2	𝍸 //	7	0.233	23.3%
3	///	3	0.100	10.0%
4	//	2	0.067	6.7%
$(\sum)$ المجموع		30	1	100%

عرض و تنظيم البيانات – (1.1 التوزيعات التكرارية)

مثال 2 (البيانات الكمية المتصلة) : زودك مدير الأجور اليومية بالريال لعينة مكونة من (50) عامل بإحدى المصانع ، كما يلي :

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 توزيع تكراري . و إيجاد التوزيع التكراري النسبي

عرض و تنظيم البيانات – (1.1 التوزيعات التكرارية)

الحل /

1 – نحسب المدى (R) و يعرف بأنه الفرق بين أكبر قيمة و أصغر قيمة ، كما يلي :

$$R = \max - \min = 79 - 10 = 69.$$

2 – نوجد عدد الفئات (k) ، كما يلي :

$$k = 1 + 3.3 \log n = 1 + 3.3 \log(50) = 1 + (3.3 \times 1.69897) = 6.61 \approx 7.$$

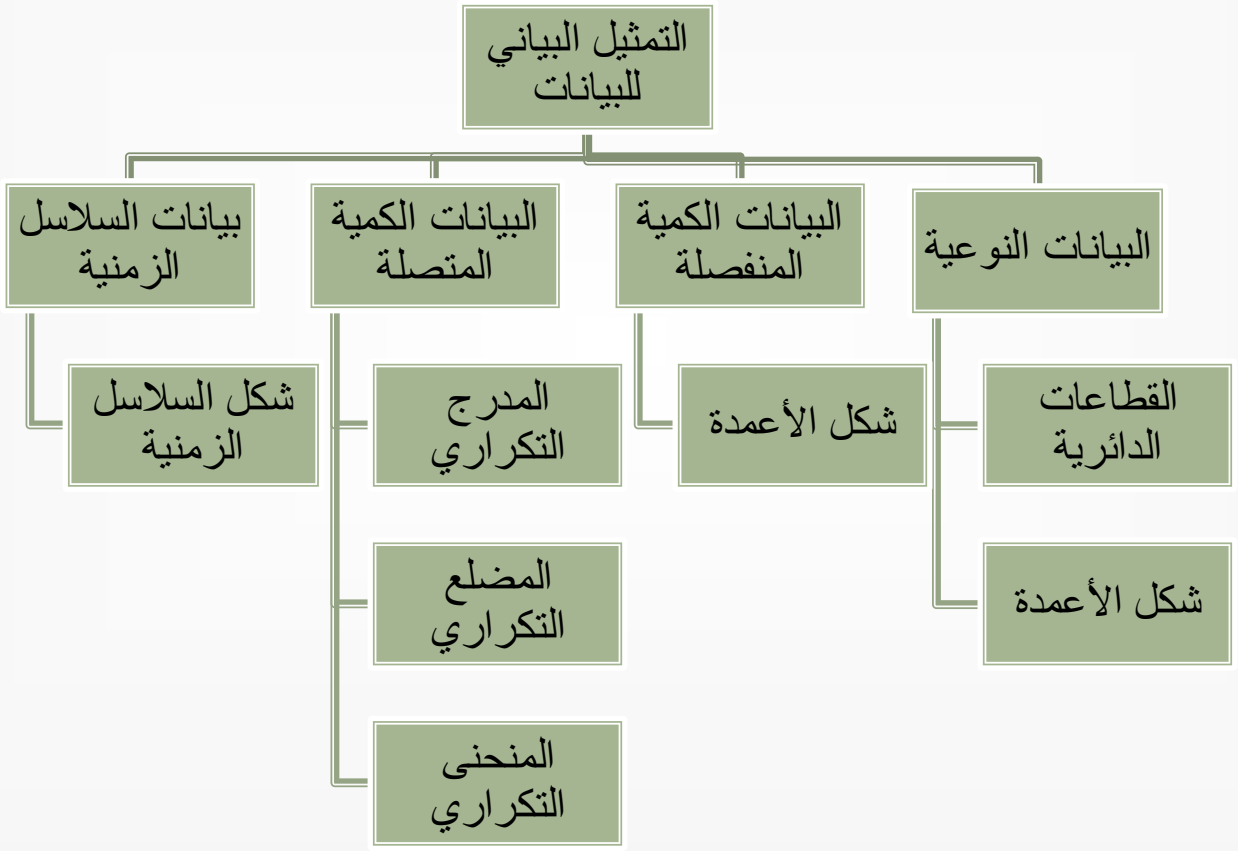
3 – نحدد طول الفترة (h) : $h = \frac{R}{k} = \frac{69}{7} = 9.86 \approx 10$.

4 – نكون الجدول التكراري :

عرض و تنظيم البيانات – (1.1 التوزيعات التكرارية)

فئات الأجور الفئات (c)	العلامات	عدد العمال التكرار (f)	التكرار النسبي (p)	نسبة العمال
10 -	///	3	0.06	6%
20 -	𝍸 /	6	0.12	12%
30 -	𝍸 𝍸	10	0.20	20%
40 -	𝍸 𝍸 𝍸	15	0.30	30%
50 -	𝍸 ///	8	0.16	16%
60 -	𝍸	5	0.10	10%
70 - 80	///	3	0.06	6%
$\sum$		50	1	100%

عرض و تنظيم البيانات – (1.2 التمثيل البياني للبيانات)

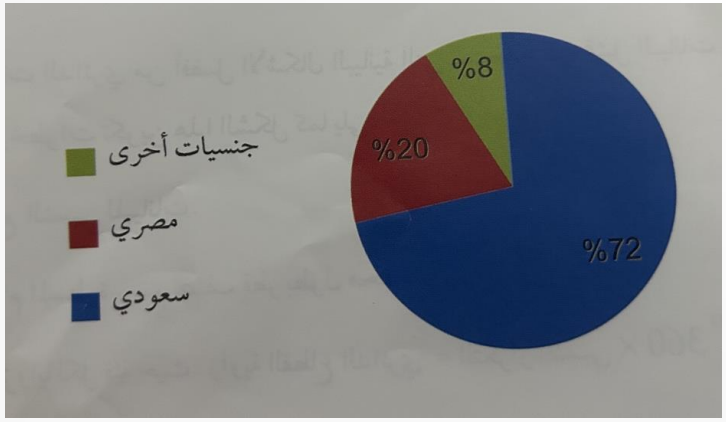


عرض و تنظيم البيانات – (1.2 التمثيل البياني للبيانات)

- **القطاعات الدائرية :**

مثال / لتحديد نسبة السعودة في إحدى شركات القطاع الخاص ، جمعت بيانات عينة مكونة من (1250) موظف من هذه الشركة . تم تلخيص هذه البيانات في توزيع تكراري و نسبي ، كما يلي :

الجنسية	عدد الموظفين	نسبة الموظفين
سعودي	900	72%
مصري	250	20%
جنسيات أخرى	100	8%
الإجمالي	1250	100%



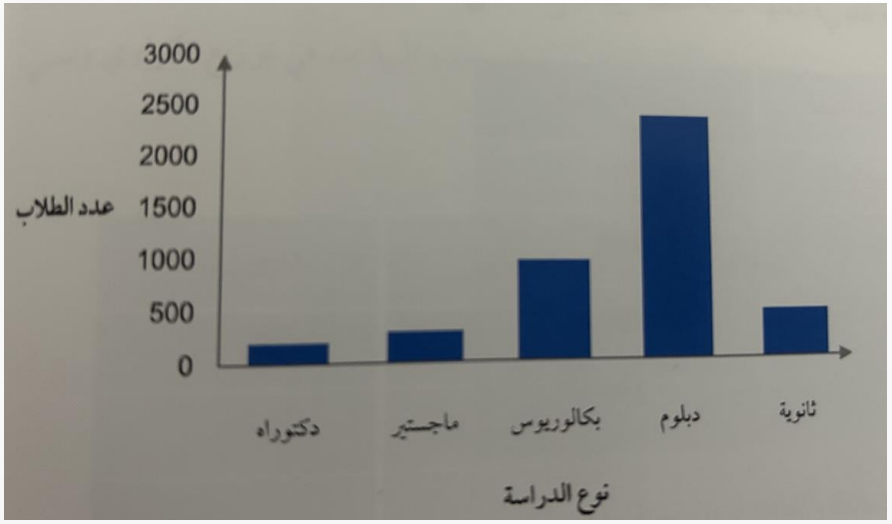
زاوية القطاع الأول (سعودي) = $0.72 \times 360° \approx 259°$

عرض و تنظيم البيانات – (1.2 التمثيل البياني للبيانات)

- **شكل الأعمدة :**

مثال / الجدول التالي يمثل توزيع تكراري لعينة من (4500) طالب سعودي يدرسون في أحد الدول الغربية ، كما يلي :

عدد الطلاب	نوع الدراسة
200	دكتوراة
300	ماجستير
1000	بكالوريوس
2500	دبلوم
500	ثانوية
4500	الإجمالي

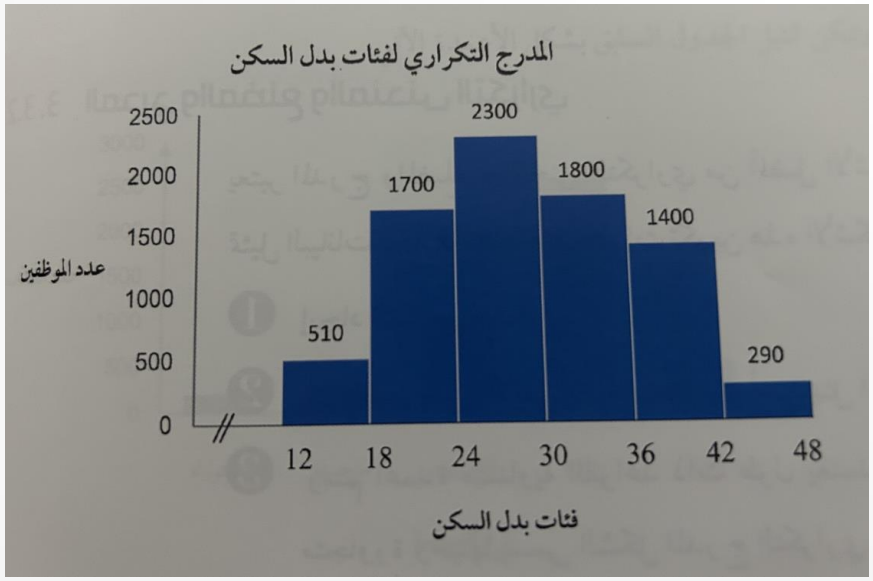


عرض و تنظيم البيانات – (1.2 التمثيل البياني للبيانات)

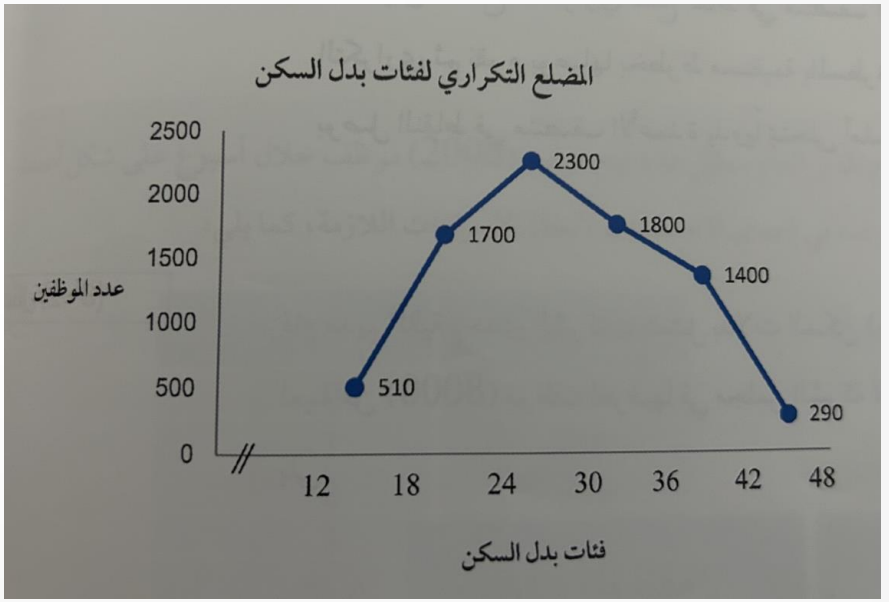
- **المدرج و المضلع و المنحنى التكراري :**

مثال / قام مدير المالية بإحدى الشركات بتمثيل بدلات السكن (بآلاف الريالات) المعطاة سنويًا لعينة من (8000) موظف لعرضها في مجلس الشركة لاتخاذ الإجراءات اللازمة ، كما يلي :

عدد الموظفين	فئات البدل
510	12 -
1700	18 -
2300	24 -
1800	30 -
1400	36 -
290	42 – 48
8000	∑



عرض و تنظيم البيانات – (1.2 التمثيل البياني للبيانات)

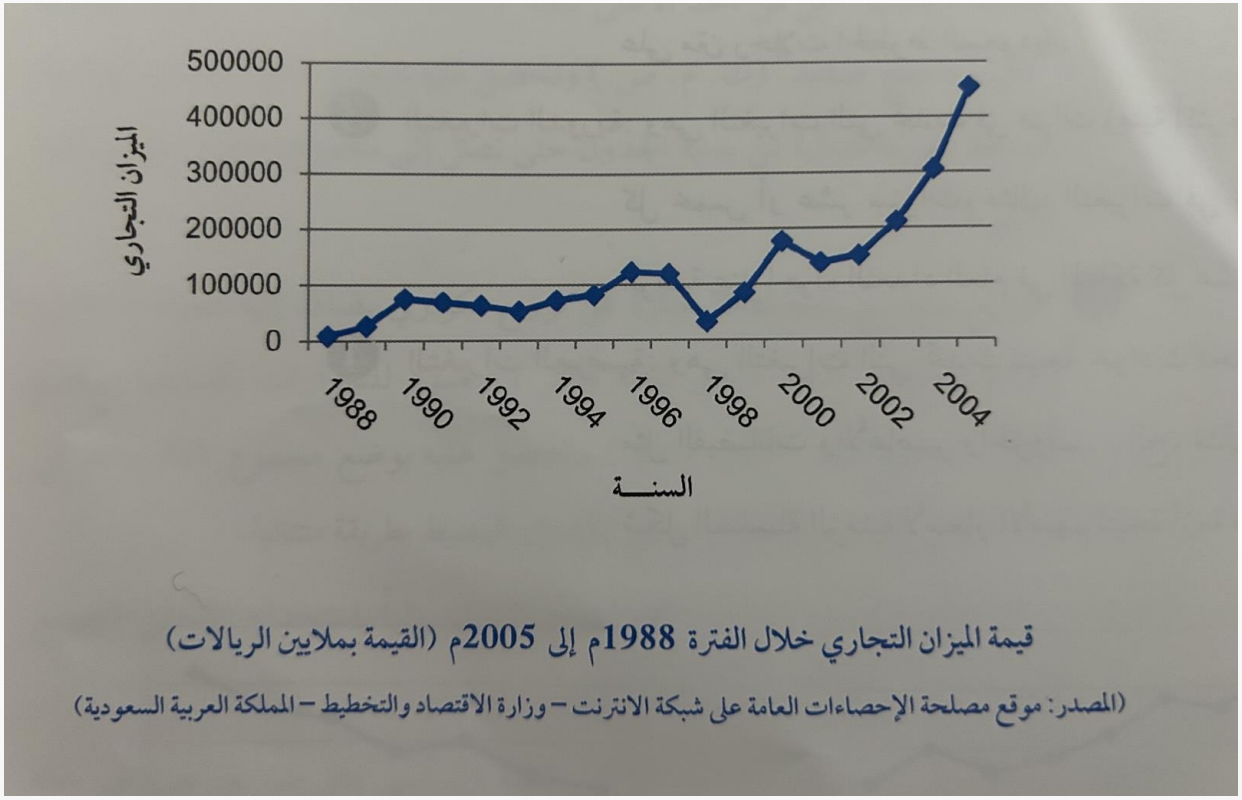




عرض و تنظيم البيانات – (1.2 التمثيل البياني للبيانات)

- **شكل السلسلة الزمنية :**

مثال /



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)

عرض و تنظيم البيانات – (1.3 التوزيعات التكرارية المتجمعة)

- **توزيع متجمع صاعد :** و هو عبارة عن توزيع تكراري يضاف إليه عمودان ، عمود للفئات و يكتب " أقل من الحد الأعلى للفئة " ، و عمود للتكرارت المتجمعة و يكتب " تكرار متجمع صاعد" (ت.م.ص.) و نحصل عليه بتجميع التكرارات في الجدول البسيط بطريقة متتالية من بداية الجدول حتى نصل إلى مجموع التكرارات في الفئة الأخيرة .

- **توزيع متجمع نازل :** و هو عبارة عن توزيع تكراري يضاف إليه عمودان ، عمود للفئات و يكتب " أقل من الحد الأدنى للفئة" ، و عمود للتكرارت المتجمعة و يكتب " تكرار متجمع نازل" (ت.م.ن.) و نحصل عليه بوضع مجموع التكرارات في الفئة الأولى ثم طرح التكرارات في الجدول البسيط بطريقة متتالية.

عرض و تنظيم البيانات – (1.3 التوزيعات التكرارية المتجمعة)

مثال / مستخدمًا البيانات المعطاة في مثال 2 أجب عن الأسئلة التالية :

- تكوين التوزيع التكراري المتجمع الصاعد و النازل ، و من ثم تكوين المنحنى المتجمع الصاعد و النازل.
- من المنحنى المتجمع الصاعد ، حدد ما يلي : عدد العمال التي تقل أجورهم عن 35 ريال ، الحد الأعلى للأجور الذي بلغه 30 عامل ، نسبة العمال الذين تتراوح أجورهم اليومية بين 55 و 25 ريال.
- من المنحنى المتجمع النازل ، حدد ما يلي : عدد العمال التي تصل أجورهم 25 ريال فأكثر، الحد الأدنى للأجور الذي بلغه 20 عامل.

عرض و تنظيم البيانات – (1.3 التوزيعات التكرارية المتجمعة)

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	فئات 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	∑

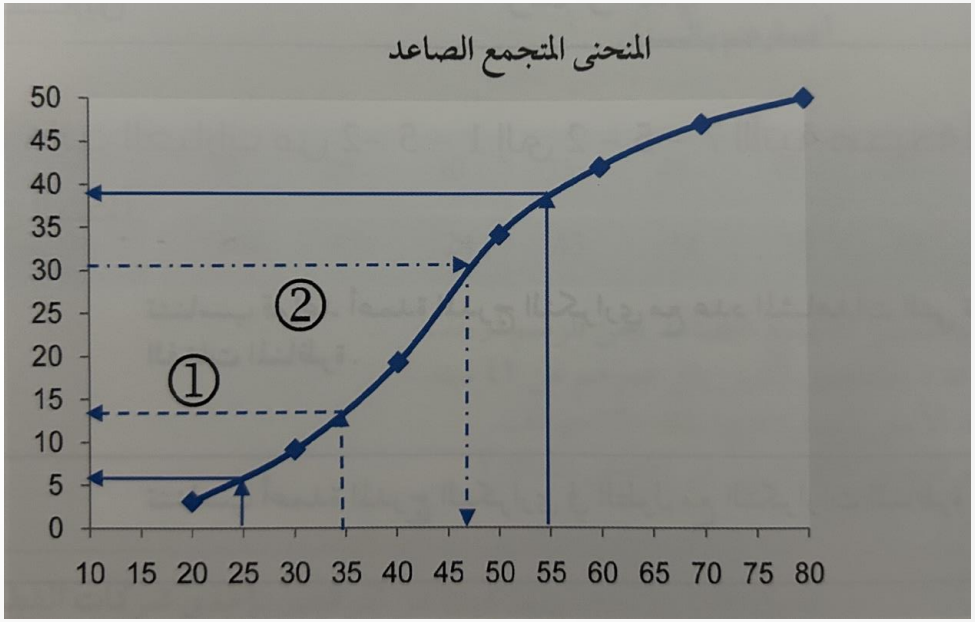
عرض و تنظيم البيانات – (1.3 التوزيعات التكرارية المتجمعة)

1- عدد العمال الذين تقل أجورهم عن 35 ريال = 13 عامل.

2- الحد الأعلى للأجور الذي بلغه 30 عامل = 46 ريال

3- نسبة العمال الذين تتراوح أجورهم اليومية بين 55 و 25 ريال =

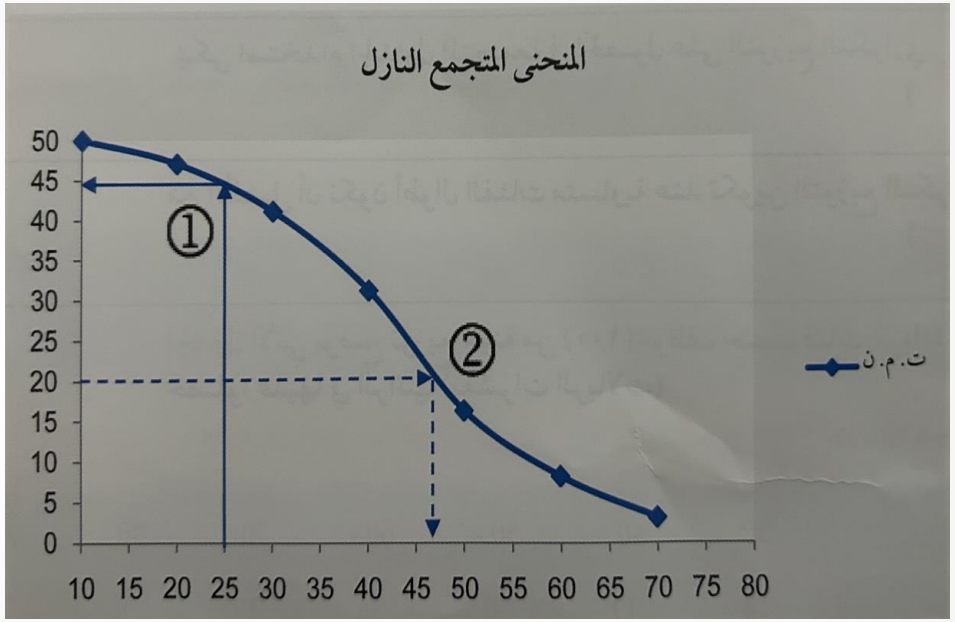
$\{(39-6) \div 50\} \times 100 = 66\%$



عرض و تنظيم البيانات – (1.3 التوزيعات التكرارية المتجمعة)

1- عدد العمال الذين تصل أجورهم 25 ريال فأكثر = 44 عامل.

2- الحد الأدنى للأجور الذي بلغه 20 عامل = 47 ريال.



عرض و تنظيم البيانات – (1.3 تدريبات)

- حل تمارين الكتاب من 2-5-1 إلى 2-5-8 صفحة 56.
- حل تمرين الكتاب 16.5.2 صفحة 58 .

المقاييس الإحصائية الوصفية



1.1 مقاييس النزعة المركزية (المتوسطات)

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- يقصد بمقاييس النزعة المركزية أو المتوسطات ، المقاييس التي تتمركز حولها معظم البيانات. و بعض هذه المقاييس : الوسط الحسابي ، الوسيط ، المنوال ، الوسط الهندسي ، و الوسط التوافقي.
- الوسط الحسابي (Arithmetic Mean) : الوسط الحسابي لمجوعة بيانات هو حاصل جمعها مقسومًا على عددها، و يمز له بالرمز μ ليمثل متوسط المجتمع ، و $\overline{x}$ ليمثل متوسط العينة.
- حساب الوسط الحسابي في حالة البيانات غير المبوبة :

• إذا كانت $x_1, x_2, \dots, x_N$ تمثل N من بيانات المجتمع ، فإن الوسط الحسابي (μ) يحسب كالتالي :

$$\mu_x = \frac{\sum x}{N}$$

حيث $\sum x$ مجموع بيانات المجتمع ، N عدد بيانات المجتمع .

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- إذا كانت $x_1, x_2, \ldots, x_n$ تمثل n من بيانات العينة ، فإن الوسط الحسابي ($\bar{x}$) يحسب كالتالي :

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

حيث $\sum x$ مجموع بيانات العينة ، n عدد بيانات العينة .

مثال 1 / احسب الوسط الحسابي للأجور اليومية بالدولار للعينة التالية المكونة من خمس عمال بإحدى القطاعات : 50 70 80 90 60 ؟

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

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

مثال 2/ شركة لديها (6) مصانع موزعة في مناطق مختلفة لإنتاج نفس المنتج و تبلغ السعة الإنتاجية للوحدات من هذا المنتج في هذه المصانع كما يلي : 1200, 2500, 3500, 1000, 2000, 3000.

المطلوب : حساب معدل (متوسط) إنتاج الشركة الكلي من هذا المنتج .

الحل/

$$\mu_x = \frac{\sum x}{N} = \frac{(1200 + 2500 + 3500 + 1000 + 2000 + 3000)}{6} = \frac{13200}{6}$$
$$= 2200 \text{ وحدة}$$

تدريب / أوجد الوسط الحسابي للبيانات : 37, 31, 35, 40 , 32 ؟

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- حساب الوسط الحسابي في حالة البيانات المبوبة :

إذا كانت لدينا k من الفئات ذات المراكز $x_1, x_2, \dots, x_k$ ، و لها تكرارات مقابلة $f_1, f_2, \dots, f_k$ على الترتيب فإننا نتبع الخطوات التالية لحساب الوسط الحسابي $\bar{x}$:

أولاً / نضيف للجدول التكراري عمود يمثل مركز الفئة و نرمز له بالرمز x ، و لحساب مراكز الفئات نتبع ما يلي :

- نحسب طول الفترة (h) = الحد الأعلى للفئة – الحد الأدنى لنفس الفئة.
- نحسب مركز الفئة الأولى = (الحد الأدنى للفئة الأولى + الحد الأعلى للفئة الأولى) ÷ 2 .
- نحسب مركز الفئة الثانية = مركز الفئة الأولى + طول الفترة (h) .
- نحسب مركز الفئة الثالثة = مركز الفئة الثانية + h ، و هكذا حتى الانتهاء من جميع الفئات.

ثانيًا / نكون عمود أخر يحتوي على حاصل ضرب عمود التكرارات في عمود مركز الفئة.

ثالثًا / نستخدم القانون التالي :

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_k f_k}{f_1 + f_2 + \cdots + f_k} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i}$$

حيث $\sum_{i=1}^{k} x_i f_i$: مجموع حاصل ضرب عمود التكرارات في عمود مركز الفئة .

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

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

مثال 3 / يبين الجدول التالي توزيع نتائج الامتحان النهائي (40) درجة لـ (100) طالب في مادة الإحصاء كما يلي :

درجات الامتحان النهائي (الفئات)	0 -	8 -	16 -	24 -	32 - 40
عدد الطلاب (التكرار)	5	10	30	40	15

الحل / حيث أن طول الفترة h يساوي 8 ، فإن :

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

درجات الامتحان النهائي (c)	عدد الطلاب (f_i)	مراكز الفئات (x_i)	$x_i f_i$
0 -	5	$(0+8) \div 2 = 4$	20
8 -	10	$4 + h = 4 + 8 = 12$	120
16 -	30	20	600
24 -	40	28	1120
32 -40	15	36	540
$\sum$	100		2400
	$\sum_{i=1}^{5} f_i$		$\sum_{i=1}^{5} x_i f_i$

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i} = \frac{2400}{100} = 24 \text{ درجة}$$

- **الوسيط** (Median) : هو القيمة العددية التي تقسم البيانات إلى قسمين متساويين بعد ترتيبها تصاعديًا أو تنازليًا ، و يرمز له بالرمز m و يعرف أيضًا بأنه مقياس الموقع لأن قيمته تعتمد على موقعه الذي يتوسط البيانات.

- حساب الوسيط في حالة البيانات غير المبوبة : إذا كانت $x_1, x_2, \dots, x_n$ تمثل n من بيانات العينة و كانت n فردية فإن الوسيط يكون القيمة التي ترتيبها $\frac{n+1}{2}$ بعد ترتيب البيانات تصاعديًا أو تنازليًا ، أما إذا كانت n زوجية فإن الوسيط يكون حاصل قسمة مجموع القيمتين اللتان ترتيبهما $\frac{n}{2}$ ، $\frac{n}{2} + 1$ على 2 بعد ترتيب البيانات تصاعديًا أو تنازليًا.

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

مثال 4 / احسب وسيط الأجور اليومية بالدولار للبيانات الآتية و التي تمثل عينتين من العمال مختارتين من شركتين مختلفتين : العينة (1) : 50 70 80 90 60

العينة (2) : 50 70 80 90 60 100؟

الحل / العينة (1) : نرتب البيانات تصاعديًا لتصبح : 90 80 70 60 50 و بالتالي الوسيط هو

$$m = 70\ \$$$

العينة (2) : نرتب البيانات تصاعديًا لتصبح : 100 90 80 70 60 50 و بالتالي الوسيط هو

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

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- حساب الوسيط في حالة البيانات المبوبة :

إذا كانت لدينا k من الفئات ، لها تكرارات مقابلة $f_1, f_2, \dots, f_k$ على التوالي. فإننا نتبع الخطوات التالية لحساب الوسيط m :

أولاً / نحدد ترتيب الوسيط من العلاقة : $c_1 = \frac{\sum_{i=1}^{k} f_i}{2}$.

ثانيًا / نكون الدول التكراري المتجمع الصاعد.

ثالثًا / نحدد فئة الوسيط ؛ بالبحث عن الفئتين تتراوح قيمة c_1 بين ت. م. ص. الخاص بهما ، ثم اختيار الفئة ذات ت. م. ص. أكبر من c_1 .

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

رابعًا / 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.

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

مثال 5 / أجرى معلم اختبار مستوى (100) درجة في اللغة الإنجليزية لـ (28) طالب في الصف الأول ثانوي بإحدى المدارس و كانت درجاتهم موزعة على النحو التالي :

درجات الطلاب	4 -	20 -	36 -	52 -	68 -	84 – 100
عدد الطلاب	1	2	6	10	7	2

المطلوب : حساب الوسيط لدرجات الطلاب في مادة اللغة الإنجليزية.

الحل / نلاحظ ترتيب الوسيط يساوي $c_1 = \frac{28}{2} = 14$ ، و طول الفئة $h = 16$ ، و بالتالي :

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

درجات الطلاب	عدد الطلاب	أقل من الحد الأعلى للفئة	ت. م. ص.
4 -	1	أقل من 20	1
20 -	2	أقل من36	3
36 -	6	أقل من 52	$c_2 = 9$
$L = 52$ –	$c_3 = 10$	أقل من 68	19
68 -	7	أقل من 84	26
84 – 100	2	أقل من 100	28
$\sum$	$\sum_{i=1}^{k} f_i = 28$		

$c_1 = 14$

و بالتالي ، قيمة الوسيط هي :

$$m = L + \frac{c_1 - c_2}{c_3} \times h = 52 + \frac{14 - 9}{10} \times 16 = 60 \text{ درجة}$$

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- **المنوال** (Mode) : هو قيمة المفردة التي تتكرر أكثر من غيرها ، أو هو المفردة ذات القيمة الأكثر شيوعًا أو تكراراً ، و يرمز له بالرمز (D) ، و يمكن استخدامه كذلك في حالة البيانات الوصفية .

- في حالة تكرار قيمة واحدة في البيانات أكثر من غيرها من القيم فإن هذه البيانات يطلق عليها وحيدة المنوال ، أما في حالة تساوي قيمتين أو أكثر في العدد الأكبر من التكرارات فإن هذه البيانات تعتبر ثنائية المنوال أو متعددة المنوال ، أما في حالة عدم وجود أي قيمة متكررة فإن البيانات تعتبر عديمة المنوال أو لا يوجد منوال لها.

- حساب المنوال في حالة البيانات غير المبوبة يتضح في المثال التالي :

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

مثال 6 / احسب الأجر اليومي الشائع للبيانات الآتية و التي تمثل عينتين من الموظفين مختارتين من شركتين مختلفتين :

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 .

تدريب / أوجد المنوال للقيم التالية : 12 , 16 , 17 ,14, 16 , 13 ؟

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- حساب المنوال في حالة البيانات المبوبة : إذا كانت البيانات كمية منفصلة مبوبة فيطبق التعريف مباشرة أما إذا كانت كمية متصلة فهناك عدة طرق تقريبية من أفضلها طريقة بيرسون و يمكن تلخيصها كما يلي :

أولاً / نحدد فئة المنوال ؛ و هي الفئة ذات أكبر تكرار.

ثانيًا / نستخدم العلاقة التالية:

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

L : الحد الأدنى لفئة المنوال ، Δ_1 : الفرق بين أكبر تكرار و السابق له.

Δ_2 : الفرق بين أكبر تكرار و اللاحق له , h : طول الفئة.

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

مثال 7 / حساب المنوال لعدد الساعات الأسبوعية التي قضاها (33) متطوعًا في العمل التطوعي في العمل:

عدد المتطوعين(التكرار)	الساعات الأسبوعية (الفئات)
3	1 -
9	3 -
10	L =5 -
7	7 -
3	9 -
1	11 - 13

$$\Delta_1 = 10 - 9 = 1$$

$$\Delta_2 = 10 - 7 = 3$$

$$D = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 5 + \frac{1}{1+3} \times 2 = 5.5 \text{ ساعة}$$

المقاييس الإحصائية الوصفية



1.1 مقاييس النزعة المركزية (المتوسطات)

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- **الوسط المرجح (Weighted Mean) :** الوسط المرجح لمجموعة من القيم ، هو مجموع حواصل ضرب قيم مفردات العينة في أوزان مخصصة لكل منها ، مقسومًا على مجموع هذه الأوزان ، و يرمز له بالرمز $\bar{x}_w$ ، و يحسب بالقانون التالي :

$$\bar{x}_w = \frac{w_1x_1 + w_2x_2 + \cdots + w_mx_m}{w_1 + w_2 + \cdots + w_m} = \frac{\sum_{i=1}^{m} w_ix_i}{\sum_{i=1}^{m} w_i}$$

حيث $x_1, x_2, \ldots, x_m$ مفردات العينة ، و التي لها الأوزان $w_1, w_2, \ldots, w_m$.

مثال 8 / أوجد الدرجة الموزونة (المتوسط المرجح) لدرجات أحد الطلاب في ثلاث مقررات بأحد الفصول الدراسة حيث كانت درجاته : 50, 70, 40 و كانت الساعات الدراسية المعتمدة هي 4, 3, 2 على الترتيب.

الحل /

$$\bar{x}_w = \frac{\sum_{i=1}^{m} w_ix_i}{\sum_{i=1}^{m} w_i} = \frac{(2)(40)+(3)(70)+(4)(50)}{9} = 54.4 \text{ درجة}$$

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- **مزايا الوسط الحسابي :** - إذا كانت البيانات موزعة بشكل طبيعي أو معتدل تقريبًا ، فإن الوسط الحسابي يعتبر أفضل المقاييس للنزعة المركزية - سهولة حسابه و التعامل معه جبريًا - تدخل جميع القيم في حسابه – لا يحتاج في حسابه إلى ترتيب البيانات – يعتبر الأساس في معظم عمليات الإحصاء الاستدلالي.

- **عيوب الوسط الحسابي :** لا يمكن إيجاده للبيانات الوصفية – قد لا يساوي عددًا صحيحًا أو أي من القيم الداخلة فيه – يصعب حسابه للجداول التكرارية المفتوحة – لا يمكن إيجاده بالرسم _ يتأثر بالقيم الشاذة (المتطرفة).

- **مزايا الوسيط :** سهولة حسابه أو إيجاده – لا يتأثر بالقيم الشاذة – يمكن حسابه في الجداول التكرارية المفتوحة – يمكن أحيانًا إيجاده للبيانات الوصفية التفضيلية – يمكن ايجاده بالرسم .

- **عيوب الوسيط :** لا تدخل جميع القيم في حسابه أو إيجاده و إنما يتحدد الوسيط عن طريق قراءة واحدة إذا كانت عدد القيم فردي أو قراءتين إذا كان عدد القيم زوجي – قد يصعب استخدامه ف الإحصاء الاستدلالي لصعوبة إمكانية معالجته بالطرق الجبرية.

المقاييس الإحصائية الوصفية – (3.1 مقاييس النزعة المركزية (المتوسطات))

- **مزايا المنوال :** سهولة حسابه – لا يتأثر بالقيم الشاذة – يمكن حسابه للجداول التكرارية المفتوحة – يمكن إيجاده لجميع أنواع البيانات سواء الوصفية أو الكمية (ما عدا عديمة المنوال) – يعتبر المقياس الوحيد للنزعة المركزية الذي يمكن إيجاده للبيانات الاسمية – يمكن إيجاده بالرسم.

- **عيوب المنوال :** عدم دخول جميع القيم في حسابه أو إيجاده – يصعب التعامل معه في الاستدلال الإحصائي , لأنه قد يكون للبيانات أكثر من منوال واحد أو تكون عديمة المنوال – قد لا يقع المنوال في مركز البيانات بل في الأطراف بالأخص في الجداول التكرارية – تتغير قيمة المنوال بحسب تنظيم البيانات أو طريقة اختيار الفئات التكرارية – معرف بصيغة غير قابلة للمعالجة الجبرية.

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

- **المدى (Range) :** في حالة البيانات غير المبوبة هو الفرق بين أكبر قيمة و أقل قيمة من البيانات ، و في حالة البيانات المبوبة هو الفرق بين الحد الأعلى للفئة الأخيرة و الحد الأدنى للفئة الأولى.

مثال 9 / البيانات التالية تمثل درجة الرضا الوظيفي لعينة من الموظفين بإحدى الشركات :

60 51 56 50 68 59 أوجد المدى ؟

الحل / درجة $R = 68 - 50 = 18$.

مثال 10 / الجدول التالي يوضح توزيع (100) شخص حسب أوزانهم بالكيلوجرامات ، و المطلوب حساب مدى الوزن لهؤلاء الأشخاص :

فئات الوزن	50 -	58 -	66 -	74 -	82 -	90 - 98
عدد الأشخاص	3	10	24	40	15	8

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

الحل / يتضح أن هناك تفاوتًا بين أوزان الأشخاص لأن مدى الأوزان يساوي (48) كجم :

$$R = 98 - 50 = 48$$ كيلوجرام

- **بعض مزايا المدى :** سهولة حسابه – يعطي فكرة سريعة عن تفاوت البيانات.
- **بعض عيوب المدى :** لا يدخل في حسابه إلا قراءتين (العظمى و الصغرى) و لربما تكون إحداهما أو كلاهما قيم متطرفة، لذلك لا يعتمد عليه كثيرًا – يصعب حسابه في البيانات الوصفية أو الجداول التكرارية المفتوحة – يتأثر بالقيم الشاذة .

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

- **التباين و الانحراف المعياري :** الانحراف المعياري (Standard Deviation) هو الجذر التربيعي الموجب للتباين و التباين (Variance) لبيانات المجتمع هو عبارة عن متوسط مربعات انحرافات القيم عن وسطها الحسابي .
- كلما اقتربت قيمة التباين من الصفر ، كلما أصبحت البيانات قريبة من التجانس.
- حساب التباين و الانحراف المعياري في حالة البيانات غير المبوبة :

- إذا كانت $x_1, x_2, \dots, x_N$ تمثل N من بيانات المجتمع ، بمتوسط حسابي (μ_X) ، و كانت هذه المشاهدات تعبر عن جميع بيانات المجتمع تحت الدراسة ، فإن التباين و الانحراف المعياري لهذا المجتمع يحسبان عن طريق الصيغتين التاليتين لى التوالي :

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

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

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

- إذا كانت $x_1, x_2, \ldots, x_n$ تمثل n من بيانات العينة ، بمتوسط حسابي ($\bar{x}$) ، و كانت هذه المشاهدات تعبر عن عينة مأخوذة من مجتمع تحت الدراسة ، فإن التباين و الانحراف المعياري لهذا المجتمع يحسبان عن طريق الصيغتين التاليتين لى التوالي :

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

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

هناك طريقة مختصرة لحساب الانحراف المعياري كالتالي :

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

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

مثال 11 / أوجد التباين و الانحراف المعياري للأجور اليومية بالدولار للعينة العشوائية التالية المكونة من ثلاثة عمال بإحدى القاطاعات : 50 70 90 ؟

الحل /

i	1	2	3	$\sum$
x_i	50	70	90	$\sum x_i = 210$
x_i^2	2500	4900	8100	$\sum x_i^2 = 15500$

$$S = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \left[\frac{(\sum_{i=1}^{n} x_i)^2}{n}\right]}{n-1}} = \sqrt{\frac{15500 - \left[\frac{(210)^2}{3}\right]}{3-1}} = \sqrt{\frac{800}{2}} = \sqrt{400} = 20\ \$$$

تدريب / ما هي قيمة التباين في المثال السابق ؟

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

- حساب التباين و الانحراف المعياري في حالة البيانات المبوبة :

- إذا تم أخذ عينة عشوائية من مجتمع محل الدراسة ، و كان لدينا k من الفئات ذات المراكز $x_1, x_2, \dots, x_k$ ، و لها تكرارات مقابلة $f_1, f_2, \dots, f_k$ على الترتيب ، فإن التباين و الانحراف المعياري لهذا العينة يحسبان عن طريق الصيغتين التاليتين على الترتيب :

$$S^2 = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^2}{\sum_{i=1}^{k} f_i - 1}$$

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

- هناك طريقة مختصرة لحساب الانحراف المعياري كالتالي :

$$S = \sqrt{\frac{\sum_{i=1}^{k} f_i x_i^2 - \left[\left(\sum_{i=1}^{k} f_i x_i \right)^2 \Big/ \sum_{i=1}^{k} f_i \right]}{\sum_{i=1}^{k} f_i - 1}}$$

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

مثال 12 / الجدول أدناه يوضح توزيع عينة عشوائية مكونة من (36) عاملاً بإحدى الشركات حسب درجة التكيف الاجتماعي:

درجة التكيف الاجتماعي	30 -	36 -	42 -	48 -	54 -	60 - 66
عدد العمال	1	4	8	12	9	2

المطلوب : حساب التباين و الانحراف المعياري لدرجات التكيف الاجتماعي لعمال هذه الشركة .

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

الحل /

درجة التكيف الاجتماعي (c)	عدد العمال (f_i)	مراكز الفئات (x_i)	$x_i f_i$	$x_i^2 f_i$
30 -	1	33	33	1089
36 –	4	39	156	6084
42 -	8	45	360	16200
48 -	12	51	612	31212
54 -	9	57	513	29241
60 – 66	2	63	126	7938
$\sum$	36		1800	91764
	$\sum_{i=1}^{6} f_i$		$\sum_{i=1}^{6} x_i f_i$	$\sum_{i=1}^{6} x_i^2 f_i$

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

نحسب التباين كما يلي :

$$S^2 = \frac{\sum_{i=1}^{k} f_i x_i^2 - \left[\left(\sum_{i=1}^{k} f_i x_i \right)^2 \Big/ \sum_{i=1}^{k} f_i \right]}{\sum_{i=1}^{k} f_i - 1} = \frac{91764 - \left[\frac{(1800)^2}{36} \right]}{36 - 1} = \frac{1764}{35} = 50.4$$

أما الانحراف المعياري فيساوي :

$$S = \sqrt{S^2} = \sqrt{50.4} = 7.1 \text{ درجة}$$

المقاييس الإحصائية الوصفية – (4.1 مقاييس التشتت)

- **خواص التباين و الانحراف المعياري** : لا تتأثر قيمتا التباين و الانحراف المعياري بطرح أو إضافة مقدار ثابت لجميع القيم – تتأثر قيمتا التباين و الانحراف المعياري بضرب جميع القيم في مقدار ثابت .
- **بعض مميزات الانحراف المعياري** : سهولة حسابه و التعامل معه جبريًا – تدخل جميع القيم في حسابه – و لذلك يعتبر من أدق مقاييس التشتت – له نفس وحدة القياس الظاهرة محل الدراسة .
- **بعض عيوب الانحراف المعياري** : تأثره بالقيم الشاذة – لا يمكن حسابه للبيانات الوصفية – يصعب حسابه للجداول التكرارية المفتوحة .

المقاييس الإحصائية الوصفية – (5.1 العلاقة بين المتوسطات و مقاييس التشتت)

- **معامل الاختلاف (Coefficient of Variation) :** هو معامل نسبي يستخدم للمقارنة بين تشتت ظاهرتين أو أكثر مختلفتين أو حتى متشابهتين في وحدة القياس. و الظاهرة التي معامل اختلافها أكبر تكون أكثر تشتتًا من الأخرى. و يرمز له بالرمز $c.v.(x)$.

و تستخدم الصيغة التالية عند حسابه من بيانات العينة :

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

بينما يتم حسابه من بيانات المجتمع عن طريق الصيغة التالية :

$$c.v.(x) = \frac{\sigma}{\mu} \times 100\ \%$$

المقاييس الإحصائية الوصفية – (5.1 العلاقة بين المتوسطات و مقاييس التشتت)

مثال 13 / إذا أراد شخص الاستثمار في إحدى الشركتين (B أو A) . فأيهما يختار؟ إذا كانت لديه المعلومات التالية :

العائد على أسهم الشركتين خلال السنوات الخمس الماضية		
	الوسط الحسابي	الانحراف المعياري
الشركة (A)	10	2
الشركة (B)	12	5

الحل / $c.v.(x) = \frac{5}{12} \times 100\ \% = 41.67\ \%$, $c.v.(x) = \frac{2}{10} \times 100\ \% = 20\ \%$

مما سبق يتضح أن العائد المتوقع من سهم الشركة (B) أعلى إلا أنه استثماراً ذا مخاطر أكبر من الاستثمار في الشركة (A) .

لذا الاستثمار الأفضل في الشركة (A) .

المقاييس الإحصائية الوصفية – (5.1 العلاقة بين المتوسطات و مقاييس التشتت)

- **معامل الالتواء :** الالتواء (Skewness) هو درجة بعد المنحنى التكراري عن التماثل و يرمز له بالرمز s.k. .
- يكون التوزيع ملتو لليسار إذا كان $s.k. < 0$ ، و ملتو لليمين إذا كان $s.k. > 0$ ، و متماثل إذا كان $s.k. = 0$.
- يقاس معامل الالتواء بإحدى الصيغتين التاليتين :
  - معامل الالتواء الأول (يحسب عن طريق المنوال) :

$$s.k.(I) = \frac{\bar{x} - D}{S}$$

  - معامل الالتواء الثاني (يحسب عن طريق الوسيط) :

$$s.k.(II) = \frac{3(\bar{x} - m)}{S}$$

المقاييس الإحصائية الوصفية – (5.1 العلاقة بين المتوسطات و مقاييس التشتت)

مثال 14 / الجدول التالي يعطي بعض المقاييس الوصفية لمبالغ الاستثمارات (بالمليون ريال) لـ(40) شركة، و المطلوب قياس معامل الإلتواء المناسب لهذه البيانات :

الانحراف المعياري	المنوال	الوسط الحسابي
10.43	153	152

الحل /

$$s.k.(I) = \frac{\bar{x} - D}{S} = \frac{152 - 153}{10.43} = -0.096$$

شكل توزيع مبالغ الاستثمارات لهذه الشركات ملتو قليلاً جهة اليسار.

تدريب / حل تمرين 17.5.3 صفحة 113 .

الارتباط و الانحدار الخطي البسيط



4.1 مفهوم الارتباط

الارتباط و الانحدار الخطي البسيط – (4.1 مفهوم و قياس الارتباط)

- **الارتباط :** هو تعيين طبيعة و قوة العلاقة بين متغيرين أو عدمها.
- يعرف **الارتباط الموجب** (الطردي) (Positive Correlation) بأنه علاقة بين متغيرين بحيث إذا تغير أحدهما فإن الآخر يتبعه في نفس الاتجاه ، أما **الارتباط السالب** (العكسي) (Negative Correlation) بأنه علاقة بين متغيرين بحيث إذا تغير أحدهما فإن الآخر يتبعه في الاتجاه المضاد .
- يعرف **معامل الارتباط** و الذي يرمز له بالرمز r ، بأنه عبارة عن مقياس رقمي يقيس قوة و نوع الارتباط بين متغيرين ، حيث تتراوح قيمته بين (1+) و (1-) ، أي أن $-1 \leq r \leq +1$. و تدل إشارة المعامل الموجبة على العلاقة الطردية ، بينما السالبة على العكسية.
- يعرف معامل المشاركة (Association Coefficient) ، بأنه عبارة عن مقياس رقمي يقيس قوة العلاقة بين متغيرين أحدهما على الأقل له مستوى قياس اسمي ، حيث تتراوح قيمته بين (1+) و (0) . و هو يعكس مدى اقتران أو ارتباط المتغيرين بعضهما ببعض .

قيمة معامل الارتباط	0	0.01 – 0.49	0.50 – 0.69	0.70 – 0.99	+1
المعنى	لا يوجد ارتباط	ارتباط طردي ضعيف	ارتباط طردي متوسط	ارتباط طردي قوي	ارتباط طردي تام

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

- **معامل بيرسون للارتباط الخطي :**

مثال 1 / بفرض أنه لدينا ثلاثة أزواج من القيم لمتغيرين كميين كما في الجدول المجاور:

x	1	2	3
y	2	5	6

ادرس وجود علاقة ارتباط خطية بين x و y ؟

الحل /

	x	y	xy	x^2	y^2
	1	2	2	1	4
	2	5	10	4	25
	3	6	18	9	36
$\sum$	$6 = \sum x$	$13= \sum y$	$30= \sum xy$	$14= \sum x^2$	$65= \sum y^2$

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

حيث أن عدد أزواج القيم هو ثلاثة فإن قيمة n=3 و عليه فإن :

$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

$$= \frac{(3)(30) - (6)(13)}{\sqrt{[(3)(14) - (6)^2][(3)(65) - (13)^2]}}$$

$$= \frac{90 - 78}{\sqrt{(42 - 36)(195 - 169)}} = \frac{12}{\sqrt{(6)(26)}} = 0.961$$

و هو يعبر عن وجود علاقة ارتباط خطية طردية عالية .

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

- **معامل سبيرمان لارتباط الرتب :**

مثال 2 / لدارسة علاقة ارتباط تقديرات الطلاب في مادة الإحصاء و تقديراتهم في مادة الرياضيات ، اخترنا خمسة طلاب و كانت تقديراتهم كما يلي :

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

هل توجد علاقة ارتباط ؟ ما نوعها و مدى قوتها ؟

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

الحل /

x	y	رتب x	رتب y	d	d^2
F	D	1	2	-1	1
A	C	5	3	2	4
C	B	3	4	-1	1
D	F	2	1	1	1
B	A	4	5	-1	1
			$\sum$	$0=\sum d$	$8=\sum d^2$

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{(6)(8)}{120} = 1 - 0.4 = 0.6$$

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

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

- **معامل بوينت بايسيريال للارتباط :** يستخدم لقياس علاقة الارتباط بين متغير كمي (X) و متغير اسمي (Y) (ذي مستويين) كالإجابة (نعم – لا) ، أو الجنس (ذكر – أنثى)... إلخ .

مثال 3 / أوجد قيمة معامل الارتباط بين الإجابة على السؤال الإجباري x (حيث 1 تعني الإجابة على السؤال الإجباري و 0 تعني عدم الإجابة على السؤال الإجباري) و الدرجة الإجمالية Y لعدد خمسة من الطلبة حسب البيانات التالية :

زوج القراءات	0	1	2	3	4
Y	1	1	1	0	0
X	18	14	19	11	7
	X_1			X_2	

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

الحل / $n = 5, n_1 = 3, n_2 = 2$

$$\bar{x}_1 = \frac{18 + 14 + 19}{3} = 17, \bar{x}_2 = \frac{11 + 7}{2} = 9$$

$$S_X = \sqrt{\frac{\sum x^2 - \left[(\sum x)^2 / n\right]}{n - 1}}$$

$$= \sqrt{\frac{18^2 + 14^2 + 19^2 + 11^2 + 7^2 - \left[(18 + 14 + 19 + 11 + 7)^2 / 5\right]}{4}} = \sqrt{\frac{98.8}{4}} = \sqrt{24.7} = 4.97$$

$$r_{pb} = \frac{\bar{x}_1 - \bar{x}_2}{S_X} \sqrt{\frac{n_1 n_2}{n(n-1)}} = \frac{17-9}{4.97} \sqrt{\frac{3 \times 2}{5 \times 4}} \approx 0.88$$

أي أن قيمة معامل الارتباط تدل على أن هناك ارتباط قوي بين الإجابة على السؤال الإجباري و الدرجة الإجمالية

و هذا يعني أن السؤال الإجباري يقيس ما يقيسه الاختبار ككل بشكل كبير.

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

- **معامل الاقتران (فاي) :**

مثال 4 / أوجد قيمة معامل الاقتران بين النوع x (ذكر / أنثى) و الإصابة بمرض الاكتئاب y (مصاب / غير مصاب) حسب البيانات التالية:

مرض الاكتئاب / النوع	مصاب	غير مصاب
ذكر	12	8
أنثى	4	6

الارتباط و الانحدار الخطي البسيط – (4.2 بعض معاملات الارتباط)

الحل /

مرض الاكتئاب / النوع	مصاب	غير مصاب	المجموع
ذكر	12	8	20
أنثى	4	6	10
المجموع	16	14	30

و عليه فإن :

$$a = 12\ , b = 8, c = 4, d = 6$$

$$r_{\emptyset} = \frac{ad\ - bc}{\sqrt{(a+b\ )(\ c+d\ )(\ a+c)(b+d)}} = \frac{(12)(6) - (8)(4)}{\sqrt{(20)(10)(16)(14)}} = \frac{72-32}{\sqrt{44800}} \approx 0.19$$

أي أنه توجد علاقة ضعيفة بين النوع و الإصابة بمرض الاكتئاب .

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)

- **مفهوم التنبؤ (التوقع) :** التنبؤ (التوقع) (Prediction) في أبسط معنى له هو تقدير القيمة المستقبلية لمتغير واحد بناء على معرفة قيم متغير آخر ، و من الممكن استخدام مفهوم الارتباط في عملية التنبؤ .

- تحتاج عملية التنبؤ إلى خطوتين أساسيتين هما :

  - تعيين خط الانحدار ، و هي معادلة رياضية .
  - استخدام هذه المعادلة الرياضية في التنبؤ الحقيقي للقراءات .

- **تحليل الانحدار (Regression Analysis) :** عبارة عن أسلوب إحصائي يقوم بصياغة أثر عدة ظواهر مؤثرة $x_1, x_2, \dots, x_m$ على ظاهرة معينة y على شكل دالة رياضية بغرض توقع قيم غير معروفة لها. و تسمى الدالة الرياضية المصاغة باسم **دالة الانحدار** و تكون على الصورة $y = f(x_1, x_2, \dots, x_m)$.

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)

- تسمى المتغيرات (الظواهر) $x_1, x_2, \ldots, x_m$ بالمتغيرات المستقلة ، و المتغير الناتج y بالمتغير التابع .
- أبسط حالة لدالة الانحدار عندما يكون هناك متغير مستقل واحد يرتبط بالمتغير التابع أي أن :

$$\hat{y} = a + bx$$

حيث a : ثابت الانحدار أو الجزء المقطوع من محور y ، b : ميل الخط المستقيم أو معامل انحدار y على x (أو y/x)

و تحسب القيمتان a و b من العلاقتين :

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n}$$

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)

مثال 5 / بفرض أنه لدينا ثلاثة أزواج من القيم لمتغيرين كميين كما هو في الجدول التالي :

(x)	1	2	3
(y)	2	5	6

أوجد معادلة خط الانحدار بين (x) و (y) ، ثم أوجد قيمة $\hat{y}$ إذا كانت قيمة $x = 2.5$ ؟

الحل /

x	y	xy	x^2
1	2	2	1
2	5	10	4
3	6	18	9
$6 = \sum x$	$13 = \sum y$	$30 = \sum xy$	$14 = \sum x^2$

حيث أن عدد أزواج القيم هو ثلاثة فإن قيمة $n = 3$ و عليه فإن :

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} = \frac{(3)(30) - (6)(13)}{(3)(14) - 6^2}$$

$$= \frac{90 - 78}{42 - 36} = \frac{12}{6} = 2$$

$$a = \frac{\sum y - b\sum x}{n} = \frac{13 - (2)(6)}{3} = \frac{13 - 12}{3} = \frac{1}{3} = 0.333$$

مما سبق يتضح أن معادلة خط الانحدار هي : $\hat{y} = a + bx = 0.333 + 2x$

و إذا كانت $x = 2.5$ فإن قيمة y المقدرة تكون : $\hat{y} = 0.333 + (2)(2.5) = 5.333$

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)

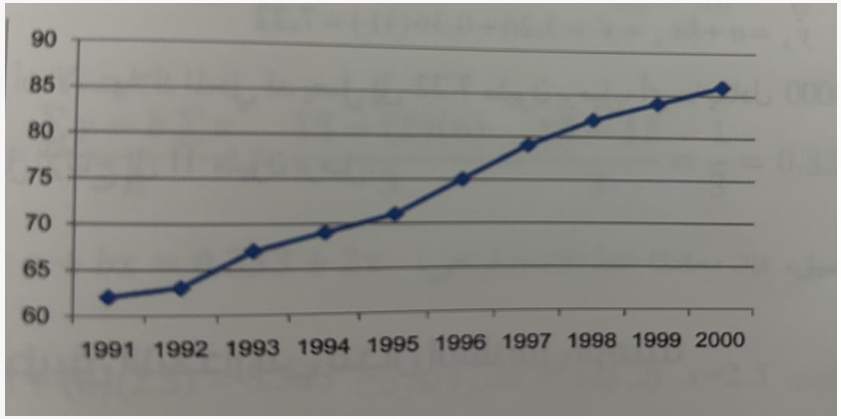
- **تطبيق الانحدار في مجال السلاسل الزمنية :**

مثال 6 / البيانات التالية تمثل عدد الحقول المكتشفة (y) خلال الأعوام من 1991 م إلى 2000 م :

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

حدد نوع الاتجاه العام ، ثم قدر معادلة الاتجاه العام الخطي ، ثم توقع عدد الحقول المكتشفة عام 2002 م ؟

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)



الحل / يدل الاتجاه العام على الزيادة في قيمة عدد الحقول المكتشفة.

السنة	x	y	xy	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
	$45 = \sum x$	$737 = \sum y$	$3553 = \sum xy$	$285 = \sum x^2$

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} = \frac{35530 - (45)(737)}{2850 - 45^2} = \frac{2365}{825} = 2.87$$

$$a = \frac{\sum y - b\sum x}{n} = \frac{737 - (2.87)(45)}{10} = 60.79$$

مما سبق يتضح أن معادلة الاتجاه العام الخطية هي : $\hat{y} = a + bx = 60.79 + 2.87x$ و لتوقع عدد الحقول المتوقع اكتشافها عام 2002 م نعوض بقيمة تدل على هذا الزمن ؛ حيث أن $x = 9$ تقابل العام 2000 م إذن $x_h = 11$ تقابل العام 2002 م ، و بالتعويض في معادلة الاتجاه العام نجد أن :

و إذا كانت $x = 2.5$ فإن قيمة y المقدرة تكون :

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

الارتباط و الانحدار الخطي البسيط – (4.3 الانحدار الخطي البسيط)

تدريب / البيانات التالية تمثل التكلفة الحدية (x) لإنتاج السلعة بالريال و إجمالي المنتج السنوي (y) بالمليون طن لسبع مؤسسات :

X	5	2	3	7	6	2	1
Y	20	15	10	25	12	14	5

- احسب معامل ارتباط بيرسون بين الظاهرتين .
- قدر حجم إجمالي المنتج السنوي عندما تكون التكلفة الحدية (4) ريالات.

نماذج تطبيقية لاستخدامات الإحصاء الوصفي

5.1 تعداد التحليل الإحصائي للبيانات السكانية



نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- إن مجموعة الطرق الخاصة بتحليل البيانات السكانية تندرج تحت تخصص يسمى الإحصاء السكاني (Demography)،

و لعل أهم مصادر هذا النوع من البيانات : تعداد السكان ، المسوح السكانية البيئية ، الإحصاءات الحيوية.

- **تعداد السكان (Population Census) :** يعرف تعداد السكان بأنه تسجيل لعدد الأشخاص الموجودين على قيد الحياة عند نقطة زمنية محددة و كذلك تسجيل خصائصهم الحيوية و الاقتصادية و الاجتماعية في تلك النقطة . و التعداد عملية كلية لجمع و تجهيز و تحليل و نشر البيانات السكانية المتعلقة بكل الأفراد في وقت معين داخل حدود معروفة .

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- يوجد أساسان لإجراء التعداد هما :

- الأساس الفعلي" الواقعي " : و يتميز بالسهولة حيث يتم حصر الأشخاص في مكان وجودهم وقت التعداد بصرف النظر عن كونهم من سكان هذا المكان أصلاً أو زائرين بصفة مؤقتة، و من الدول التي تستخدم هذا الأساس إنجلترا و مصر و المملكة العربية السعودية.

- الأساس النظري" الحقيقي" : و هو يعطي صورة صحيحة عن السكان الدائمين بكل منطقة حيث يتم حصر الأشخاص حسب محل إقامتهم الأصلية الحقيقية بصرف النظر عن أماكن تواجدهم وقت التعداد ، و من الدول التي تستخدم هذا الأساس كندا و الولايات المتحدة و ألمانيا .

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- **المسوحات السكانية البيئية** : و يقصد بها المسوح المتخصصة في جانب معين بالخصوبة أو الجوانب الاقتصادية أو السكانية أو التعليمية أو الصحية . أو مسوح عامة تشتمل جوانب عديدة مثل : مستوى الدخل ، و مستوى المعيشة ، و الجوانب الإسكانية و التعليمية و الصحية.

- **الإحصاءات الحيوية** (Vital Statistics) : تعرف بأنها تلك الإحصاءات التي تتناول الوقائع المتعلقة بحياة الفرد منذ ولادته و حتى وفاته . فهي بذلك تشمل كافة ما يتعلق بحالة السكان و تكوينهم و حركتهم و الحوادث الهامة التي تقع لهم ، و هذا يتمثل في تعدادات السكان و إحصاءات المواليد و الوفيات و الزواج و الطلاق و الهجرة و إحصاءات الأمراض و أسبابها. فهي تشمل كافة ما يتعلق بحالة السكان من حيث التعداد و النمو و التصنيف و التغيرات التي تحدث لهم و الأنشطة و الوظائف التي يمارسونها .

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- **تخدم الإحصاءات الحيوية عدة أغراض أهمها :**

- التخطيط في جميع المجالات التعليمية و الصحية و الاقتصادية و الاجتماعية .
- تنظيم و تحسين الخدمات العامة و الخاصة .
- قياس المستوى العلمي و الحضري و الثقافي للمجتمع .
- البحث العلمي بجميع فروعه خاصة في ميادين البيئة ، الطب ، الاجتماع و التعليم .
- المقارنات المحلية و الدولية .

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

1 – عدد السكان و توزيعهم الجغرافي : يساعد في معرفة الآتي :

- حساب مقياس يدل على درجة ازدحام الدولة بالسكان و هو : كثافة السكان = $\frac{\text{عدد السكان في الدولة}}{\text{مساحة الدولة بالكيلومتر المربع}}$

- حساب مقياس يوضح درجة الازدحام داخل المسكن و هو : كثافة السكن = $\frac{\text{عدد السكان في الدولة}}{\text{عدد حجرات المساكن}}$

- حساب مقياس يساعد على تقدير عدد السكان في غير سنوات التعداد و ذلك بحساب معدل زيادة السكان من تعداد لآخر :

معدل الزيادة في عدد السكان = $\frac{\text{عدد السكان في سنة المقارنة – عدد السكان في سنة الأساس}}{\text{عدد السنوات}}$

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

مثال 1 / بفرض أن تعداد السكان في إحدى الدول 50 مليون نسمة في منتصف 1420هـ و كانت مساحة هذه الدولة 4 مليون كيلومتر مربع و عدد حجرات المساكن 25 مليون حجرة . 1 – احسب كلاً من كثافة السكان و كثافة السكن .

2 – بفرض أن تعداد السكان لهذه الدولة في منتصف عام 1425هـ هو 60 مليون نسمة فما هو معدل الزيادة السنوية للسكان.

الحل /

1 – كثافة السكان = $\frac{50}{4}$ = 12.5 شخص لكل كم ^{2} .

كثافة السكن = $\frac{50}{25}$ = 2 شخص لكل حجرة.

2 – تسمى سنة 1420 هـ بسنة الأساس و سنة 1425هـ بسنة المقارنة و بالتالي فإن :

معدل الزيادة السنوية في عدد السكان = $\frac{50-60}{5}$ = 2 مليون نسمة.

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

2 – التركيب العمري و النوعي للسكان : يمكن تلخيص فوائد دراسة التركيب العمري للسكان كالآتي :

- معرفة اتجاه التغير في عدد السكان و تقدير التغيرات المستقبلية.
- دراسة الخصوبة و اتجاهاتها.
- حساب معدلات الوفاة و جداول الحياة .
- تقدير حجم القوى العاملة اللازمة لعمليات التنمية و الدفاع .
- الوقوف على مشكلات التعليم و التأمين الصحي و الضمان الاجتماعي .
- تقدير أعباء الإعالة.

❖ التركيب النوعي للسكان : على سبيل المثال تساعد معرفة الجنسية من خلال التعداد على معرفة عدد الأجانب و خصائصهم العمرية و أسباب تواجدهم و معرفة المجالات التي تحتاج الدولة فيها للعنصر الأجنبي.

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- **إحصاءات المواليد و الخصوبة و الوفيات :** تعتبر من أهم المواضيع لمعرفة حركة السكان زيادة و نقصانًا . إذ أن معرفة هذه المعدلات تساعد على التخطيط لإنشاء مدارس و خلق فرص عمل و بناء المستشفيات و توفير الخدمات العامة .

1 – معدلات خاصة بالمواليد :

- معدل المواليد الخام $= \frac{\text{عدد المواليد الأحياء خلال عام}}{\text{عدد السكان منتصف العام}} \times 1000$
- معدل الخصوبة العام $= \frac{\text{عدد المواليد الأحياء خلال عام}}{\text{عدد النساء في سن الحمل}} \times 1000$
- معدل التوالد $= \frac{\text{عدد المواليد في بلد الأحياء خلال عام}}{\text{عدد النساء المتزوجات في سن الحمل}} \times 1000$

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

مثال 2 / إذا كان عدد المواليد أحياء بإحدى الدول عام 1425 هـ بالمليون هو 1.1 و عدد النساء في سن الحمل 6.6 و عدد المتزوجات منهن 4.4 فما هو معدل الخصوبة العام و ما هو معدل معدل التوالد ؟

الحل / معدل الخصوبة العام $= \frac{1.1}{6.6} \times 1000 = 166.67$ في الألف.

معدل التوالد $= \frac{1.1}{4.4} \times 1000 = 250$ في الألف.

تدريب / إذا كان عدد المواليد أحياء في دولة ماء عام 1410 هـ هو 200000 نسمة مولوداً حيًا و كان عدد النساء اللاتي في سن الحمل 900000 أوجد معدل الخصوبة العام ؟

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

2 – معدلات خاصة بإحصاءات الوفيات :

- معدل الوفيات العام = $\frac{\text{عدد الوفيات خلال العام}}{\text{عدد السكان في منتصف العام}} \times 1000$
- معدل الزيادة الطبيعية = معدل المواليد الخام – معدل الوفيات الخام .
- معدل الوفيات الرضع = $\frac{\text{عدد الوفيات في الأطفال الذين تقل أعمارهم عن سنة واحدة}}{\text{عدد الأطفال المولودين أحياء في نفس العام}} \times 1000$
- معدل الوفيات لفئة عمرية معينة = $\frac{\text{عدد الوفيات خلال السنة من تلك الفئة العمرية في الدولة}}{\text{عدد السكان في منتصف السنة من تلك العمرية الفئة}} \times 1000$

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- الخصائص التعليمية للسكان :

❑ المعدل الخام للمسجلين في المراحل التعليمية = $\frac{\text{عدد المسجلين في المراحل التعليمية المختلفة}}{\text{عدد السكان الكلي}} \times 1000$

❑ المعدل العمري للتسجيل = $\frac{\text{عد المسجلين في المؤسسات التعليمية من فئة عمرية معينة}}{\text{عدد السكان في تلك الفئة العمرية}} \times 1000$

❑ معدل الأمية الخام = $\frac{\text{عدد الأميين من السكان}}{\text{عدد السكان}} \times 1000$

❑معدل الأمية العمري = $\frac{\text{عدد الأميين من السكان في فئة عمرية معينة}}{\text{عدد السكان في تلك الفئة العمرية المعينة}} \times 1000$

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

- الخصائص الاقتصادية للسكان :

❑ معدل النشاط الاقتصادي الخام $= \frac{\text{عدد الأفراد الناشطين اقتصاديًا}}{\text{عدد السكان الكلي}} \times 1000.$

❑ معدل النشاط الاقتصادي العام $= \frac{\text{عدد الأفراد الناشطين اقتصاديًا}}{\text{عدد السكان في سن العمل}} \times 1000.$

❑ معدل الإعالة $= \frac{\text{عدد السكان غير الناشطين اقتصاديًا}}{\text{عدد السكان الناشطين اقتصاديًا}} \times 1000.$

- مقاييس الهجرة : تنقسم الهجرة إلى قسمين داخلية و خارجية و من مقاييسها الهامة :

❑ معدل الهجرة الوافدة لمنطقة معينة $= \frac{\text{عدد المهاجرين الوافدين في منطقة معينة}}{\text{عدد السكان الكلي}} \times 1000.$

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.1 التحليل الإحصائي للبيانات السكانية)

❑ معدل الهجرة الصافية = $\frac{\text{عدد المهاجرين الوافدين في منطقة معينة- المهاجرين المغادرين لمنطقة معينة}}{\text{عدد السكان الكلي}} \times 1000$.

مثال 3 / إذا كان عدد سكان منطقة ما في منتصف عام 1425 هـ هو مليون نسمة و كان عدد المهاجرين من المنطقة خلال العام 10000 نسمة بينما كان عدد المهاجرين الوافدين للمنطقة هو 15000 نسمة خلال نفس العام . فاحسب معدلي الهجرة الوافدة و الصافية خلال عام 1425 هـ ؟

الحل / معدل الهجرة الوافدة = $\frac{15000}{1000000} \times 1000 = 15$ في الألف .

معدل الهجرة الصافية = $\frac{10000-15000}{1000000} \times 1000 = 5$ في الألف .

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.2 الأرقام القياسية للأسعار)

- الرقم القياسي للأسعار(Price Index Number) : هو رقم نسبي يقيس التغير الذي يطرأ على أسعار سلعة واحدة أو أكثر، عادة من سنة (سنة الأساس) إلى سنة أخرى (سنة المقارنة).
- نستخدم الرموز : P_0 : الأسعار في فترة الأساس ، P_1 : الأسعار في فترة المقارنة ،

Q_0 : الكميات في فترة الأساس ، Q_1 : الكميات في فترة المقارنة ، I : الرقم القياسي .

- للرقم القياسي دلالة على النحو التالي :
  - إذا كان الرقم القياسي $(I) < 100$، فذلك يدل على النقصان في الأسعار بمقدار : % (100 - I) .
  - إذا كان الرقم القياسي $(I) > 100$، فذلك يدل على الزيادة في الأسعار بمقدار : % (I -100) .

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.2 الأرقام القياسية للأسعار)

مثال 4 / الجدول التالي يمثل سعر بعض مشتقات النفط :

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

أوجد ما يلي :

1 - الرقم البسيط للأسعار .

2 – الرقم القياسي للأسعار المرجح بكميات الأساس (رقم لاسبير) .

3 – الرقم القياسي للأسعار المرجح بكميات المقارنة (رقم باشي) .

4 – الرقم القياسي الأمثل للأسعار (رقم فيشر).

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.2 الأرقام القياسية للأسعار)

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		0.9		9.3	13.4	10.2	14.7
	$\sum P_0$		$\sum P_1$		$\sum P_1Q_0$	$\sum P_0Q_0$	$\sum P_1Q_1$	$\sum P_0Q_1$

نماذج تطبيقية لاستخدامات الإحصاء الوصفي — (5.2 الأرقام القياسية للأسعار)

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 انخفضت بمقدار 30.77% .

2 – 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 انخفضت بمقدار 30.6% .

3 – 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 انخفضت بمقدار 30.61% .

4 – 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 انخفضت بمقدار 30.61% .

مقدمة في الاحتمالات و التوزيعات الاحتمالية



6.1 تعاريف أساسية

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.1 تعاريف أساسية)

- **التجربة العشوائية** (Random Experiment) : هي أي إجراء نعلم مسبقًا جميع النواتج الممكنة له و إن كنا لا نستطيع أن نتنبأ بأي من هذه النتائج سيتحقق فعلاً. حيث من الواضح أننا لا نستطيع أن نتنبأ بنتيجة التجربة العشوائية إلا إننا نستطيع حساب احتمال ظهور أي نتيجة و ذلك باستخدام الاحتمالات .

- **فراغ العينة** (Sample Space) : يعرف بأنه المجموعة المكونة من النتائج الممكنة من تجربة عشوائية . و يرمز له بالرمز (S) . كما يطلق عليه الحدث الشامل أو الحالات الممكنة.

مثال 1 / عند إلقاء قطعة نقود متوازنة مرة واحدة فما هو فراغ العينة ؟

الحل / القطعة النقدية لها وجهين أحدهما صورة و يرمز له بالرمز H و الآخر كتابة و يرمز له بالرمز T ،

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.1 تعاريف أساسية)

لذلك فإن مجموعة النتائج الممكنة لهذه التجربة و التي هي فراغ العينة تكون : S= { H, T } .

مثال 2 / عند إلقاء قطعتي نقود مرة واحدة أو (قطعة نقود واحدة مرتين) ما هو فراغ العينة ؟

الحل / S = { HH, HT, TH, TT} .



مثال 3 / عند إلقاء زهرة نرد متزنة مرة واحدة ، فما هو فراغ العينة ؟

الحل / S = { 1, 2, 3, 4, 5, 6 } .

مثال 4 / صندوق به 4 كرات حمراء و 3 كرات بيضاء ، فما هو فراغ العينة التي ستسحب حتى الحصول على أول كرة حمراء ؟

الحل / S = { 1, 2, 3, 4 } .

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.1 تعاريف أساسية)

- **الحادثة** (Event) : هي مجموعة جزئية من فراغ العينة.
- **الحادثة البسيطة** (Simple Event) : هي الحادثة التي تتكون من عنصر واحد من عناصر فراغ العينة.
- **الحادثة المركبة** (Composite Event) : هي الحادثة التي تحتوي على أكثر من عنصر من عناصر فراغ العينة.

مثال 5 / عند إلقاء قطعتي نقود مرة واحدة ، حدد ما إذا كانت الحوادث الآتية حوادث بسيطة أم لا ، إذا علمت أن فراغ العينة هو S = { HH, HT, TH, TT} :

الحادثة	وصفها	نوع الحاثة
A = { TT}	حادثة تمثل ظهور كتابتين	حادثة بسيطة
B = { HH, HT, TH }	حادثة تمثل ظهور صورة واحدة على الأقل	حادثة مركبة

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.1 تعاريف أساسية)

- **الحادثة المؤكدة** (Sure Event) : هي الحادثة التي لابد من وقوعها . فمثلاً عند إلقاء زهرة نرد متزنة مرة واحدة فإن الحادثة المؤكدة ظهور أي وجه من 1 إلى 6 .

- **الحادثة المستحيلة** (Impossible Event) : عندما لا يكون للتجربة أي نواتج متعلقة بالحادثة المذكورة ، و يرمز لها بالرمز ∅ . فمثلا ظهور الوجه 8 عند إلقاء زهرة نرد متزنة لها ستة أوجه من 1 إلى 6 .

- **الحادثة المتماثلة** (Equally Likely Event) : هي تلك الحوادث التي لها نفس فرصة الحدوث. فمثلا عند إلقاء زهرة النرد متزنة فإن فرص الحصول على إحدى الأوجه تكون متساوية.

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.1 تعاريف أساسية)

- **الأحداث المتنافية (المانعة) بالتبادل** (Mutually Exclusive Event) : إذا كان هناك حدثان A, B ، و كان وقوع الحادثين معًا حدثًا مستحيلاً و هذا يعني أن الحدثين لا يمكن أن يقعا معًا أو وقوع أحدهما يمنع وقوع الآخر فإنه يقال A, B حدثان مانعان أو متنافيان . فمثلاً عند رمي قطعة نقود مرة واحدة ظهور الكتابة يمنع حدث ظهور الصورة .

- الأحداث المستقلة (Independent Events) : لو أن لدينا حدثان و كان وقوع أحدهما لا يؤثر في وقوع الآخر فيقال أن الحدثين مستقلان . فمثلاً عند إلقاء قطعتي نرد فإن ظهور رقم على الزهرة الأولى لا يؤثر و لا يتأثر بما ظهر على القطعة الثانية .

- **طرق العد (التوفيق) :** عدد الطرق التي يمكن بها اختيار (x) من الأشياء من بين (n) من هذه الأشياء (الترتيب غير مهم) هو : $\binom{n}{x}$

$$\binom{n}{x} = C_x^n = \frac{n!}{x!(n-x)!},\ x \le n$$ ، حيث :

$n! = n\times(n-1)\times(n-2)\ldots\times2\times1$ فمثلا : $3! = 3\times2\times1 = 6$.

مثال 6 / بكم طريقة يمكن اختيار رجلين من بين 4 رجال ؟

الحل / n=4 ، $x = 2$ و بالتالي :

$$\binom{n}{x} = C_x^n = \frac{n!}{x!\,(n-x)!} = \frac{4!}{2!\,(4-2)!} = \frac{4\times3\times2\times1}{(2\times1)(2\times1)} = \frac{24}{4} = 6$$ طرق

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.2 تعريف الاحتمال)

- **التعريف الكلاسيكي للاحتمال** (Classical Definition of Probability) : إذا كان عدد الطرق التي يمكن أت تظهر بها نتائج تجربة ما هو n طريقة و كانت هذه النتائج لها نفس فرصة الظهور و كان من بينها m طريقة تظهر بها حادثة ما (A) . فإنه يقال إن احتمال وقوع الحادثة هو $\frac{m}{n}$. $P(A) = \frac{m}{n}, m \leq n$

مثال 7 / إذا كانت لديك عشر بطاقات مرقمة من الرقم (1) إلى (10) موضوعة على طاولة بشكل عشوائي و مقلوبة ، ثم سحبت إحدى هذه البطاقات : 1 – ما احتمال الحصول على بطاقة تحمل الرقم (4) ؟

2 – ما احتمال الحصول على بطاقة تحمل رقم يقبل القسمة على (3) ؟

الحل / عدد عناصر فراغ العينة n=10

1 – لنرمز لهذه الحادثة بالرمز A و بالتالي : $P(A) = \frac{m}{n} = \frac{1}{10} = 0.1$.

2 – لنرمز لهذه الحادثة بالرمز B و بالتالي : $P(B) = \frac{m}{n} = \frac{3}{10} = 0.3$.

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.2 تعريف الاحتمال)

- **التعريف التجريبي للاحتمال** (Experimental Definition of Probability) : إذا كان عدد المرات التي أجريت بها تجربة ما تحت نفس الظروف هي n و كان عدد المرات (من بينها) التي لوحظ فيها حدث معين (A) مثلاً هي m فإن احتمال وقوع الحدث A هو $P(A) = \lim_{n\to\infty} \frac{m}{n}$ ، حيث $\frac{m}{n}$ هي التكرار النسبي .

- **التعريف الرياضي للاحتمال** (Mathematical Definition of Probability) : يبنى التعريف الرياضي للاحتمال على أساس افتراض مسلمات (بديهيات) نظرية الاحتمالات .

- مسلمات نظرية الاحتمالات :

❖ يرافق كل حادثة A عدد معين $P(A)$ يسمى احتمال A و يحقق : $P(A) \geq 0$.

❖ احتمال وقوع حادثة مؤكدة يساوي واحد أي أن $P(S) = 1$ و $P(A) \leq 1$.

❖ إذا كانت A,B حادثتين متنافيتان فإن : $\mathrm{P}(\mathrm{A}\ or\ B) = P(A) + P(B)$.

مقدمة في الاحتمالات و التوزيعات الاحتمالية — (6.2 تعريف الاحتمال)

- **الحادثة المكملة :** إذا كانت $\bar{A}$ هي مكملة المجموعة A فإن : $P(\bar{A}) = 1 - P(A)$.

مثال 8 / إذا كان احتمال نجاح عملية تصحيح النظر هو 85% . فما هو احتمال عدم نجاح تلك العملية ؟

الحل / $P(A) = 0.85$ و بالتالي : $P(\bar{A}) = 1 - P(A) = 1 - 0.85 = 0.15$

إذاً احتمال عدم نجاح العملية هو : 15% .



- احتمال وقوع حدثين مانعين : $\mathrm{P(A}\ or\ B) = P(A) + P(B)$.
- احتمال وقوع حدثين غير مانعين : $\mathrm{P(A}\ or\ B) = P(A) + P(B) - P(A\ and\ B)$.

مثال 9 / سحبت ورقة من أوراق اللعب (الكوتشينة) ما احتمال أن تحمل الورقة الرقم 10 أو صورة ؟

الحل / لنفرض أن : A الورقة تحمل الرقم 10 ، B الورقة تحمل صورة

A, B حدثان مانعان و بالتالي : $\mathrm{P}(\mathrm{A}\ or\ B) = P(A) + P(B) = \frac{4}{52} + \frac{12}{52} = \frac{16}{52} = \frac{4}{13}$.

- إذا كان A, B حدثان مستقلان فإن : $\mathrm{P}(\mathrm{A}\ and\ B) = P(A)P(B)$.
- إذا كان A, B حدثان غير مستقلان فإن : $\mathrm{P}(\mathrm{A}\ and\ B) = P(A)P(B|A)$

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

- **المتغير العشوائي** (Random Variable) : هو المقدار الذي يأخذ قيما رقمية مختلفة و التي تعبر عن نتائج التجربة العشوائية.
- **المتغير العشوائي المنفصل** (أو المتقطع) (Discrete Random Variable) : يقال أن المتغير العشوائي (X) متغير عشوائي منفصل إذا كان يأخذ قيماً تنتمي إلى مجموعة محدودة أو معدودة حتى و لو لم تكن هذه القيم قيمًا صحيحة.
- **المتغير العشوائي المستمر** (أو المتصل) (Continuous Random Variable) : يقال أن المتغير العشوائي (X) متغير عشوائي متصل إذا كان يأخذ جميع القيم الصحيحة و الكسرية في مدى تغيره ، أو كان ينتمي إلى مجموعة غير محدودة أو معدودة.
- في تجربة إلقاء زهرة النرد نجد أن :

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

- **التوزيعات الاحتمالية المنفصلة :** باعتبار أن X متغير عشوائي منفصل يأخذ القيم $x_1, x_2, \dots, x_n$ و لكل قيمة احتمالات معينة كالتالي : $P(x_1), P(x_2), \dots, P(x_n)$ ، يقال أن للمتغير العشوائي المنفصل X توزيعاً احتماليا منفصلاً $P(X)$ إذا حقق هذا التوزيع الشروط التالية :

1- $P(X) \geq 0$ لجميع قيم X

2- $\sum P(X) = 1$ (مجمع الاحتمالات يساوي واحد)

- **خصائص أساسية للتوزيع الاحتمالي المنفصل :**

- توقع التوزيع (التوقع الرياضي أو متوسط التوزيع) : $E(X) = \mu = \sum x\, P(x)$.
- تباين التوزيع : $var(X) = \sigma^2 = \sum x^2 P(x) - \mu^2$.
- الانحراف المعياري : $\sqrt{var(X)} = \sigma = \sqrt{\sigma^2}$.

مثال 10 / ألقيت ثلاث قطع من العملة المعدنية المتوازنة . أجب عما يلي :

1 – ما هو المتغير العشوائي الذي يمثل عدد الصور التي تظهر على الوجه العلوي ؟

2 – أوجد التوزيع الاحتمالي لعدد الصور التي تظهر على السطح العلوي ؟

3 – أوجد خصائص توزيع الصور التي تظهر على السطح العلوي ؟

الحل / 1- فراغ العينة هو :

$$S = \{ (T,T,T), (T,T,H), (T,H,T), (H,T,T), (H,H,T),(H,T,H), (T,H,H), (H,H,H)\}$$

لنفرض أن X يمثل عدد الصور التي تظهر على السطح العلوي و الذي يمثل المتغير العشوائي ، وعلى ذلك فإن فراغ العينة يرافقه هو X = { 0, 1, 2, 3 } حيث العدد 0 يعني عدم ظهور صورة و العدد 2 ظهور صورتين و هكذا أي أن المتغير العشوائي يأخذ القيم 0, 1, 2 , 3 .

2 – التوزيع الاحتمالي :

$$P(X=0) = P(TTT)= \frac{1}{8} \geq 0 \text{ ، } P(X=1) = P(HTT)+ P(THT) + P(TTH)= \frac{3}{8} \geq 0 \text{ ،}$$

$$P(X=2) = P(HHT)+ P(HTH) + P(THH)= \frac{3}{8} \geq 0 \text{ ،} P(X=3) = P(HHH)= \frac{1}{8} \geq 0$$

مما سبق نلاحظ أن كل احتمال أكبر من الصفر و أقل من واحد و مجموع الاحتمالات يساوي واحد. $\frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$

3 – التوزيع الاحتمالي :

X	0	1	2	3	$\sum$
$P(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{8}{8}=1$
$x\,P(x)$	0	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{3}{8}$	$\frac{12}{8}$
$x^2P(x)$	0	$\frac{3}{8}$	$\frac{12}{8}$	$\frac{9}{8}$	$\frac{24}{8}$

$$E(X)=\mu=\sum x\,P(x)=\frac{12}{8}=1.5$$ صورة

$$var(X)=\sigma^2=\sum x^2P(x)-\mu^2$$

$$=\frac{24}{8}-(1.5)^2=0.75$$

$$\sigma=\sqrt{\sigma^2}=\sqrt{0.75}=0.87$$ صورة

- **التوزيعات الاحتمالية المتصلة :** يقال أن للمتغير العشوائي المتصل X توزيعاً احتماليا متصلاً $f(X)$ (يسمى دالة كثافة الاحتمال) إذا حقق الشروط التالية :

$$1\text{-}f(X) \geq 0$$

$$2\text{-} \int f(x) = 1 \qquad \text{(مجمع الاحتمالات يساوي واحد)}$$

تدريب / صمم مجموعة من المبرمجين 1000 برنامج ، فنجحوا في 900 برنامج . سحب برنامج عشوائياً فما احتمال عدم نجاحه ؟

التوزيع الطبيعي



7.1 مبادئ أساسية

التوزيع الطبيعي — (7.1 مبادئ أساسية)

- يعتبر التوزيع الطبيعي (المعتدل) من أهم و أشهر التوزيعات الاحتمالية لأنه يصف معظم الظواهر الطبيعية مثل درجات الحرارة ، درجات الذكاء ، الوزن ، الطول ....الخ . فعلى سبيل المثال استخدام التوزيع في مجال التربية و التعليم يظهر في نواحي عدة منها :

- إيجاد معايير الاختبارات المختلفة .

- تحديد النسبة المئوية لعدد الطلبة الحاصلين على درجات أقل من أو أكثر من درجة معينة أو نسبة الطلبة الحاصلين على درجات تنحصر بين درجات معينة .

- تحديد الصعوبة النسبية لأسئلة الاختبارات و بنودها.

التوزيع الطبيعي — (7.1 مبادئ أساسية)

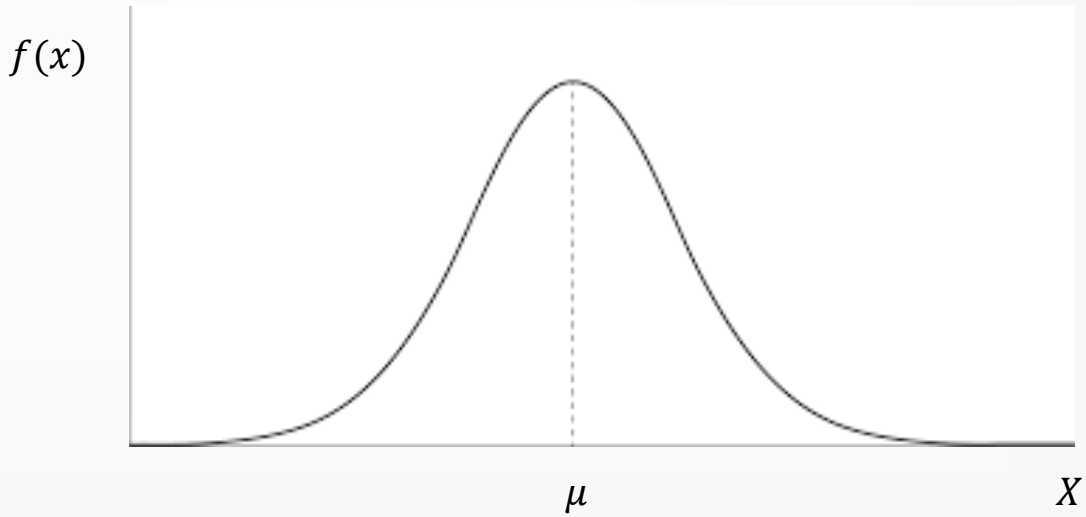
- **دالة التوزيع الطبيعي** : يقال أن المتغير العشوائي X يتبع توزيعاً طبيعيًا بالمعلمتين (μ ,σ^2) إذا كانت دالة كثافته الإحتمالية هي :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty$$

حيث μ هي متوسط التوزيع و σ^2 هي تباين التوزيع.

التوزيع الطبيعي — (7.1 مبادئ أساسية)

- **منحنى التوزيع الطبيعي** : يكون منحنى التوزيع الطبيعي متماثلاً حول محور $x = \mu$ كما في الشكل التالي :



التوزيع الطبيعي — (7.1 مبادئ أساسية)

- يمثل المتغير القياسي (المعياري) $z = \frac{X-\mu}{\sigma}$ الذي متوسطه صفر و تباينه واحد الصورة القياسية للتوزيع الطبيعي و الذي يطلق عليه اسم التوزيع الطبيعي المعياري أو القياسي (Standard Normal Distribution) و غالباً ما يكتب بالصورة

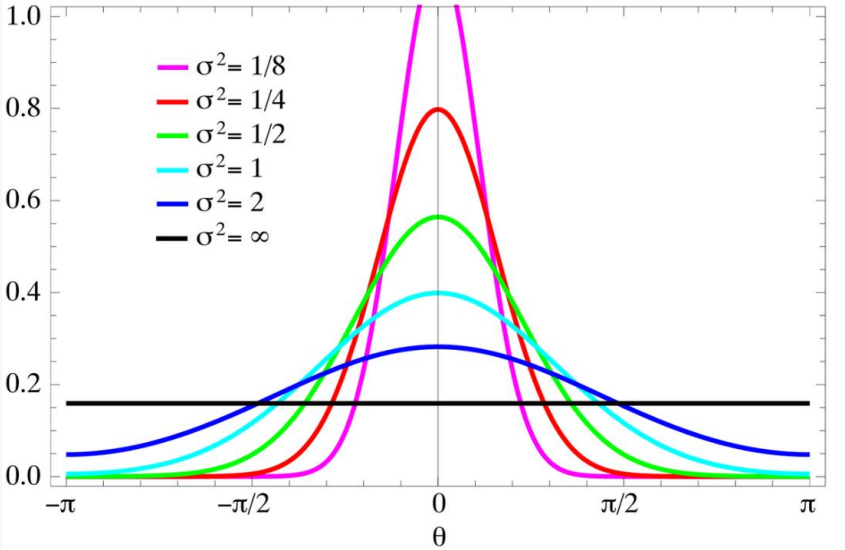
$$Z \sim N(0,1)$$

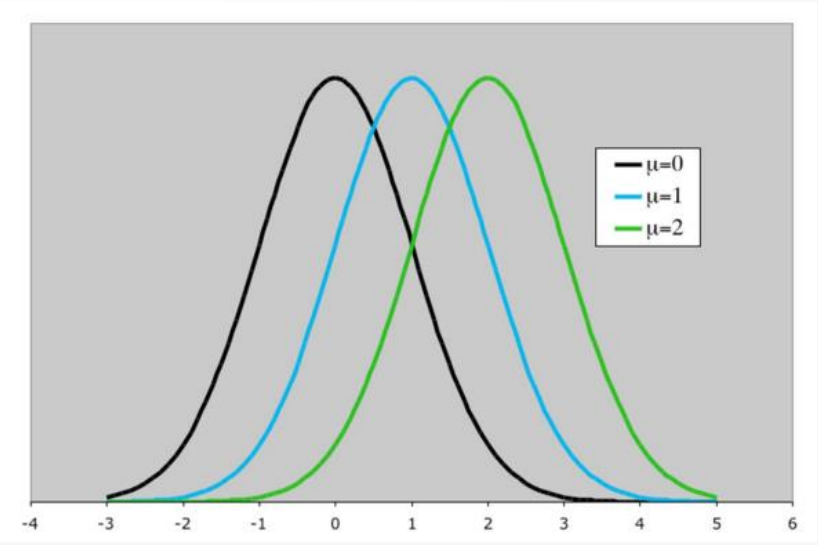
و دالة كثافته الاحتمالية هي :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \qquad -\infty < z < \infty$$

التوزيع الطبيعي — (7.1 مبادئ أساسية)

- معالم التوزيع الطبيعي : يعتمد التوزيع الطبيعي على معلمتين الأولى معلمة الموقع μ و هي التي تحدد موقع التوزيع و الثانية هي معلمة القياس σ و التي تحدد مدى تباعد (تشتت) المنحنى حول المركز.





التوزيع الطبيعي — (7.1 مبادئ أساسية)

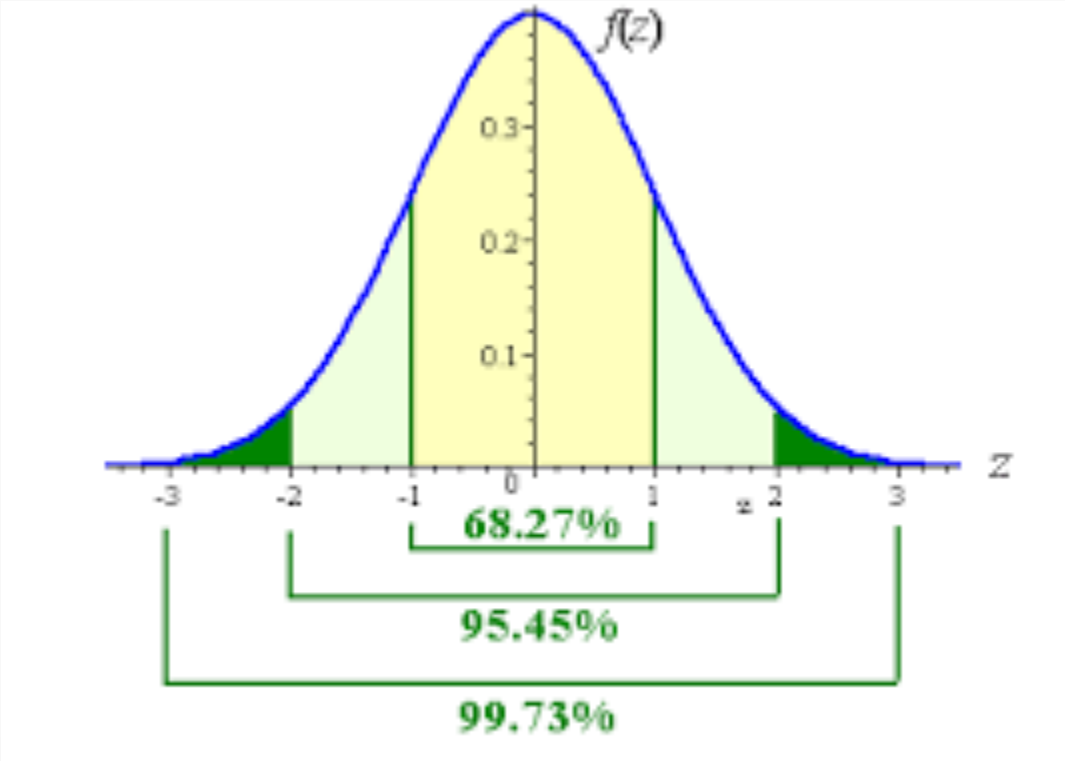
- القانون التجريبي للتوزيع الطبيعي : تنقسم المساحة تحت المنحنى الطبيعي حسب القانون التجريبي (Empirical Rule)

إلى ثلاثة أقسام كما هو موضح بالجدول التالي :

المساحة التقريبية	تنحصر بين ($\mu - k\sigma, \mu + k\sigma$)	
تحت المنحنى	الطبيعي	الطبيعي القياسي
68% من المساحة الكلية	$(\mu - \sigma, \mu + \sigma)$	(-1,1)
95% من المساحة الكلية	$(\mu - 2\sigma, \mu + 2\sigma)$	(-2,2)
99.7% من المساحة الكلية	$(\mu - 3\sigma, \mu + 3\sigma)$	(-3,3)

التوزيع الطبيعي — (7.1 مبادئ أساسية)

و تتضح هذه النسب في الشكل التالي:



مثال 1 / إذا علمت أن درجات الطلاب في مادة الإحصاء تتبع توزيع طبيعي بمتوسط 82 درجة و انحراف معياري 6 فهذا

يعني أن 68% من الطلاب درجاتهم محصورة بين 76 درجة ، و 88 درجة لأن :

$$\mu - \sigma = 82 - 6 = 76$$

$$\mu + \sigma = 82 + 6 = 88$$

التوزيع الطبيعي — (7.2 تطبيقات على الدرجة المعيارية و التوزيع الطبيعي)

- **الدرجة المعيارية** (Z-Score) : هي درجة معيارية في توزيع متوسطه الحسابي يكون (صفراً) و انحرافه المعياري يساوي (1) و بذلك يمكن مقارنة الدرجة مع غيرها من الدرجات التي تنتمي لنفس التوزيع ، و ذلك بحساب الدرجة المعيارية لكل منها . و من ثم يكون التفضيل في ضوء مقدار قيمة الدرجة المعيارية . و يرمز للدرجة المعيارية بالرمز Z و للدرجة الخام بالرمز X و يتم تحويل X إلى Z كما ذكرنا سابقًا بالتحويلة :

$$Z = \frac{X - \mu}{\sigma}$$

التوزيع الطبيعي — (7.2 تطبيقات على الدرجة المعيارية و التوزيع الطبيعي)

مثال 2 / إذا كان الوسط الحسابي لمجموعة من المفردات (50) و الانحراف المعياري (10) فأوجد :

1 – الدرجة المعيارية المناظرة للدرجة الخام 60 .

2 – الدرجة المعيارية المناظرة للدرجة الخام 45 .

3 – الدرجة الخام المناظرة للدرجة المعيارية 1.5.

الحل / حسب المعطيات فإن : $\mu = 50$,$\sigma = 10$ و التحويلة التي تستخدم هنا هي : $Z = \frac{X-50}{10}$

1 – الدرجة المعيارية المناظرة للدرجة الخام $X = 60$ هي : $Z = \frac{60-50}{10} = 1$.

2 – الدرجة المعيارية المناظرة للدرجة الخام $X = 45$ هي : $Z = \frac{45-50}{10} = \frac{-5}{10} = -0.5$.

3 – الدرجة الخام المناظرة للدرجة المعيارية $Z = 1.5$ هي :

$$1.5 = \frac{X-50}{10} \Rightarrow X - 50 = 15 \Rightarrow X = 65$$

مثال 3 / إذا كان دخل (1000) أسرة في مدينة ما يتبع توزيعاً طبيعياً متوسطه μ هو 1800 ريال و انحرافه المعياري σ هو 300 ريال . أوجد :

1 – احتمال الحصول على دخل أسرة أكبر من 2400 ريال.

2 – احتمال الحصول على دخل أسرة أكبر من 1500 ريال .

الحل / نفرض أن (X) هو دخل الأسرة . المتغير العشوائي (X) متغير عشوائي مستمر يتبع توزيع طبيعي عادي بمتوسط (μ) يساوي 1800 ريال ، و انحراف معياري (σ) يساوي 300 ريال . عند حساب الاحتمالات نحول المتغير العادي X إلى المتغير القياسي Z بالتحويلة :

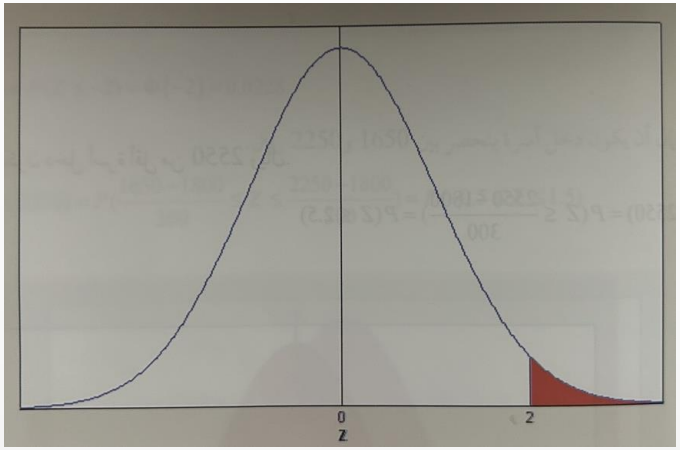
$$Z = \frac{X - \mu}{\sigma} = \frac{X - 1800}{300}$$

1- احتمال أن يكون دخل أسرة أكبر من 2400 ريال ، أي أن المطلوب إيجاد $P(X \geq 2400)$.

$$X = 2400 \Rightarrow Z = \frac{X - 1800}{300} = \frac{2400 - 1800}{300} = 2$$

و بالتالي فإنه من جدول التوزيع الطبيعي القياسي نجد أن :

$$P(X \geq 2400) = P(Z \geq 2) = 1 - P(Z < 2) = 1 - \emptyset(2) = 1 - 0.9772 = 0.0228$$

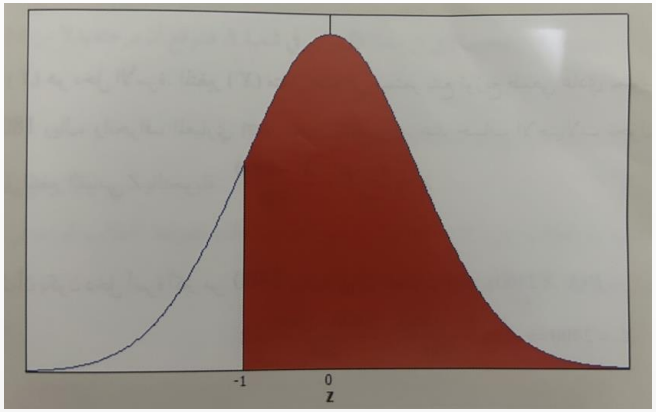


2- احتمال أن يكون دخل أسرة أكبر من 1500 ريال ، أي أن المطلوب إيجاد $P(X \geq 1500)$.

$$X = 1500 \Rightarrow Z = \frac{1500 - 1800}{300} = -1$$

و بالتالي فإنه من جدول التوزيع الطبيعي القياسي نجد أن :

$$P(X \geq 1500) = P(Z \geq -1) = 1 - P(Z < -1) = 1 - \emptyset(-1) = 1 - 0.1587 = 0.8413$$



التوزيع الطبيعي — (7.3 توزيعات المعاينة)

- **نظرية النهاية المركزية :** إذا كان لدينا مجتمع غير محدود مفرداته (X) تتبع توزيعاً احتمالياً متوسطه μ و انحرافه المعياري σ . سحبنا من هذا المجتمع عينات عشوائية حجم كل منها (n) [و كانت (n) كبيرة الحجم $(n \geq 30)$] .

فإن الوسط الحسابي لهذه العينات $\bar{X}$ يتبع تقريبًا توزيعًا طبيعيًا له الخصائص التالية :

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

التوزيع الطبيعي — (7.3 توزيعات المعاينة)

تدريب / إذا كان دخل (1000) أسرة يتبع توزيع طبيعي متوسطه (16000) ريال و انحرافه المعياري (2000) ريال :

1 – ما احتمال الحصول على دخل أقل من 15000 ريال.

2 – ما احتمال الحصول على دخل أكبر من 18000 ريال.

التقدير و اختبارات الفروض



8.1 تقدير متوسط المجتمع

التقدير و اختبارات الفروض — (8.1 تقدير متوسط المجتمع)

- **المعلمة** (Parameter) : هي عبارة عن خاصية أو مقياس يتم حسابها من المجتمع محل الدراسة . مثل نسبة البطالة في السعودية – متوسط العمر الافتراضي لجهاز معين .

- **الإحصاءة** (Statistic) : هي عبارة عن خاصية أو مقياس يتم حسابها من العينة المسحوبة من المجتمع محل الدراسة .

- ينقسم الإحصاء الاستدلالي أو (الاستدلال الإحصائي) إلى موضوعين رئيسين هما :

  - تقدير معالم المجتمع (Estimation) .
  - اختبارات فروض بشأن صحة قيم معالم المجتمع (Testing of Hypothesis) .

التقدير و اختبارات الفروض — (8.1 تقدير متوسط المجتمع)

- **التقدير** (Estimation) : هو أسلوب إحصائي مبني على نظريات إحصائية ، يستخدم لتقدير معلمة ما محل الاهتمام عن طريق استخدام مقاييس العينة . نستخدم متوسط (الوسط الحسابي) $\bar{x}$ للعينة كتقدير نقطة لمتوسط المجتمع μ .

- **الخطأ المعياري** (Standard Error) : الخطأ المعياري للمتوسط هو الانحراف المعياري لتوزيع مجتمع متوسطات العينات . بمعنى انحراف متوسطات العينات عن متوسط مجتمعها . و يرمز له بالرمز $\sigma_{\bar{X}}$ حيث :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} .$$

غالبًا ما تكون قيمة الانحراف المعياري للمجتمع (σ) غير معلومة ، لذا نقدرها بقيمة الانحراف المعياري للعينة (S) .

فيصبح لدينا تقدير للخطأ المعياري للمتوسط باستخدام الانحراف المعياري للعينة بالصيغة التالية : $\hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}}$.

مثال 1 / أوجد الخطأ المعياري للوسط الحسابي لعينة حجمها (49) تم سحبها من مجتمع له انحراف معياري يساوي (14) ؟

الحل / الخطأ المعياري لمتوسط العينة يساوي $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{14}{\sqrt{49}} = 2$.

مثال 2 / تم قياس مستوى الأداء على مهارة معينة لأفراد عينة حجمها (36) ، فكان الوسط الحسابي و الانحراف المعياري للعينة 30 و 9 على الترتيب . أوجد الخطأ المعياري لمتوسط العينة.

الحل / لأن قيمة الانحراف المعياري للمجتمع (σ) مجهولة ، لذا نقدرها بقيمة الانحراف المعياري للعينة (S) .

فيصبح لدينا تقدير للخطأ المعياري للوسط الحسابي للعينة يساوي :

$$\hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{9}{\sqrt{36}} = 1.5$$

التقدير و اختبارات الفروض — (8.1 تقدير متوسط المجتمع)

- **تقدير الفترة لمتوسط المجتمع :** تسمى الفترة التي تحتوي على مجموعة من القيم تتضمن فيما بينها قيمة معلمة المجتمع بتقدير (فترة) الثقة . نرمز لفترة الثقة الرمز $(1-\alpha)$ ، و مكمل هذه القيمة يسمى مستوى المعنوية و يركز له بالرمز α ،

فمثلاً إذا كانت درجة الثقة 95% يكون مستوى المعنوية يساوي 0.05 .

و لتقدير فترة الثقة لمتوسط المجتمع μ باستخدام متوسط العينة $\bar{x}$ و في حالة حجم العينة الكبير $(n \geq 30)$ نستخدم العلاقة :

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

حيث تعتبر $\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ الحد الأنى لفترة الثقة ، و $\bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ الحد الأعلى لفترة الثقة.

- نحدد قيمة $Z_{\frac{\alpha}{2}}$ حسب درجة الثقة (تبعاً لمستوى المعنوية (α) كما يلي :

مستوى المعنوية α	درجة الثقة	قيمة Z الجدولية
0.10	90%	$Z_{\frac{\alpha}{2}} = 1.65$
0.05	95%	$Z_{\frac{\alpha}{2}} = 1.96$
0.01	99%	$Z_{\frac{\alpha}{2}} = 2.58$

عندما تكون فترة الثقة بنسبة $100(1-\alpha)\%$ حول μ فإننا عادة ما نعبر عن هذه الفترة بعبارة مثل : " لدينا ثقة بنسبة $100(1-\alpha)\%$ أن المتوسط μ للمجتمع يقع في هذه الفترة " .

مثال 3 / مصنع لإنتاج المصابيح الكهربائية ، اختير من إنتاجه عينة حجمها (100) مصباح لتقييم جودة الإنتاج ، حيث حدد مدير المصنع الجودة أن يتراوح عمر المصباح بين (1000) و (1100) . إذا كان الوسط الحسابي لعمر المصباح في العينة المختارة (1200) ساعة و انحرافه المعياري (250) ساعة. قدر بدرجة ثقة 95% متوسط عمر المصباح من إنتاج المصنع كله؟

الحل / حيث أن الانحراف المعياري للمجتمع σ مجهول نقدره باستخدام الانحراف المعياري للعينة .

$$n = 100, \qquad \bar{x} = 1200, \qquad S = 250, \qquad \alpha = 0.05, \qquad Z_{\frac{\alpha}{2}} = 1.96$$

و على ذلك فإن :

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$1200 - 1.96 \times \frac{250}{\sqrt{100}} \leq \mu \leq 1200 + 1.96 \times \frac{250}{\sqrt{100}}$$

$$1200 - 49 \leq \mu \leq 1200 + 49$$

$$\therefore 1151 \leq \mu \leq 1249$$

إذاً متوسط عمر المصباح من إنتاج المصنع كله يتراوح بين 1151 و 1249 ساعة و ذلك بدرجة ثقة 95% . أي أن الحد الأدنى لفترة الثقة يساوي 1151 ، بينما الحد الأعلى لفترة الثقة يساوي 1249 . و يدل هذا التقدير أن جودة المصابيح عالية من هذا المصنع .

التقدير و اختبارات الفروض — (8.2 اختبار الفروض حول متوسط المجتمع)

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

I. صياغة فرضين يسميان فرض العدم H_0 و فرض البديل H_1 حول معلمة (أو خاصية) في مجتمع الدراسة.

II. حساب بعض الإحصاءات كالمتوسط ، و الانحراف المعياري ...الخ.

III. نحسب من قيم الإحصاءات إحصاء الاختبار .

IV. نتخذ القرار برفض أو عدم رفض فرض العدم .

التقدير و اختبارات الفروض — (8.2 اختبار الفروض حول متوسط المجتمع)

- **فرض العدم** (Null Hypothesis) : هو ادعاء (أو عبارة) عن معلمة مجتمع يفترض صحته حتى يثبت عكس ذلك .

- **الفرض البديل** (Alternative Hypothesis) : هو ادعاء عن معلمة مجتمع سوف يكون صحيحًا إذا كان فرض العدم غير صحيح .

- **إحصاء الاختبار** (Test Statistics) : أسلوب أو طريقة لتحديد قاعدة ترفض فرض العدم .

- **اختبار الفرض حول متوسط المجتمع من جانبين في حالة العينات الكبيرة $(n \geq 30)$:**

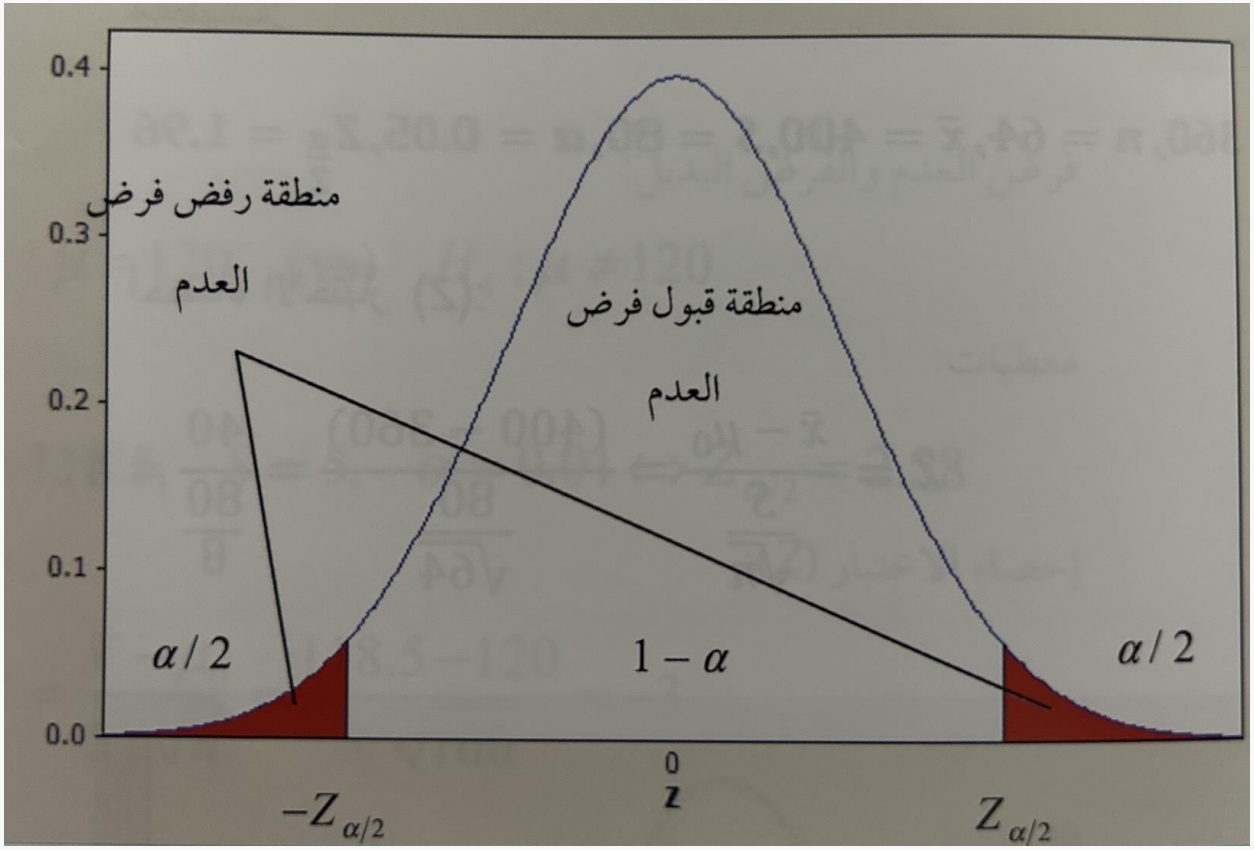
فرض العدم و فرض البديل : $H_1: \mu \neq \mu_0 \quad (vs) \quad H_0: \mu = \mu_0$

إحصاء الاختبار (Z) : $Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

نحدد مكان وقوع قيمة إحصاء الاختبار (Z) كما في الشكل .

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

التقدير و اختبارات الفروض — (8.2 اختبار الفروض حول متوسط المجتمع)



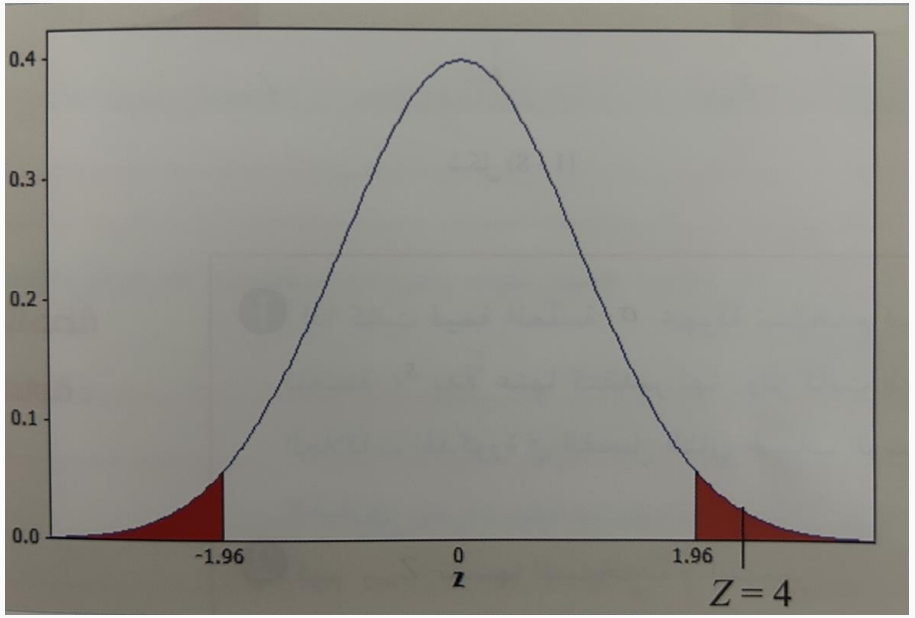
مثال 4 / إذا كان متوسط الزيادة في أجور العاملين في إحدى المؤسسات عام 2009 م هو (360) ريال ، و في عام 2012 م أخذت عينة من (64) فردًا من العاملين في هذه المؤسسة ، فوجد أن الوسط الحسابي للزيادة في أجورهم (400) ريال و الانحراف المعياري (80) ريال. هل يدل ذلك على أن متوسط الزيادة في أجور العاملين في المؤسسة عام 2012 م قد اختلف عن متوسط الزيادة في الأجور عام 2009 م ؟ و ذلك عند مستوى معنوية $\alpha = 0.05$.

الحل / فرض العدم و فرض البديل : $H_1: \mu \neq 360 \quad (vs) \quad H_0: \mu = 360$

معطيات : $\mu_0 = 360, \quad n = 64, \quad \bar{x} = 400, \quad S = 80, \quad \alpha = 0.05, \quad Z_{\frac{\alpha}{2}} = 1.96$

إحصاء الاختبار (Z) :

$$Z = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{(400 - 360)}{\frac{80}{\sqrt{64}}} = \frac{40}{\frac{80}{8}} = 4$$



نرفض فرض العدم H_0 و نقبل الفرض البديل H_1 ، أي أن متوسط الزيادة في الأجور عام 2012 م قد اختلف عن متوسط الزيادة في الأجور عام 2009 م . و ذلك عن مستوى معنوية $\alpha = 0.05$.

- استخدام قيمة (P) لاختبار الفروض الإحصائية :

قيمة (P) (P-Value) هي أصغر قيمة لمستوى المعنوية (α) يمكن عندها رفض فرض العدم .

- حيث إذا كانت $\alpha < P - Value$ فإننا نقبل فرض العدم .
- حيث إذا كانت $\alpha \geq P - Value$ فإننا نرفض فرض العدم .

❖ أغلب البرامج الإحصائية توفر هذه القيمة مباشرة عند إجراء الاستدلال الإحصائي على مجموعة من البيانات ، و تكون تحت مسمى " P-Value " أو " P " أو "Sig.".

التقدير و اختبارات الفروض — (8.2 اختبار الفروض حول متوسط المجتمع)

تدريب / الجدول الآتي يوضح توزيع عينة من (100) طالب حسب فئات الذكاء :

فئات الذكاء	50 -	60 -	70 -	80 -	90 - 100
عدد الطلبة	8	12	30	40	10

1 . قدر متوسط الذكاء في مجتمع الطلاب بدرجة ثقة 99% .

2 . اختبر الفرض القائل أن متوسط الذكاء في مجتمع الطلاب لا يساوي (76) درجة عند مستوى معنوية $\alpha = 0.01$.