الدرس ⑦ : الإحصاء

I - تذكير

1) الدراسة الإحصائية: هي دراسة ظاهرة أو خاصية يتميز بها أفراد مجموعة.

2) الساكنة الإحصائية: هي العينة أو المجموعة التي تخضع للدراسة الإحصائية وكل عنصر منها يسمى فردا أو وحدة إحصائية

3) الميزة: هي الظاهرة التي تتم دراستها وهي خاصية يمكن ملاحظتها أو قياسها وهي نوعان:

أ- ميزة كمية: ميزة يمكن التعبير عنها بأعداد (عدد الأطفال، نقط التلاميذ، العمر، الوزن، الطول...

ب- ميزة كيفية (نوعية): لا يمكن التعبير عنها بأعداد (الجنس، فصيلة الدم، اللون، نوع السيارة...)

4) الحصيص: هو عدد الوحدات التي تأخذها كل قيمة من قيم الميزة ونرمز له بالرمز n_i

5) الحصيص الإجمالي: هو مجموع الحصيصات نرمز له N

6) الحصيص المتراكم: هو مجموع حصيصات القيم التي تصغر أو تساوي هذه القيمة.

7) التردد: تردد قيمة هو خارج حصيصها على الحصيص الإجمالي $f_i = \frac{n_i}{N}$

8) التردد المتراكم: لقيمة هو خارج حصيصها المتراكم على الحصيص الإجمالي.

9) النسبة المئوية: $p_i = \frac{\text{الحصيص}}{\text{الحصيص الإجمالي}} \times 100 = f_i \times 100$

II - جدول الحصيصات والحصيصات المتراكمة والترددات والترددات المتراكمة:

1) متسلسلة بالقيم

إذا كانت الميزة كمية وعدد قيمها قليل، نرتب هذه القيم تصاعديا.

* تطبيق ①

يمثل الجدول التالي متسلسلة إحصائية تعبر عن توزيع 24 منخرطا بإحدى الأندية الرياضية حسب أعمارهم

الميزة	12	13	14	15	16
الحصيص	5	6	x	8	4
الحصيص المتراكم		11			24
التردد					
التردد المتراكم					

1) الساكنة الإحصائية هي 24 منخرطا بإحدى الأندية

2) الوحدة الإحصائية هي منخرط

3) الميزة المدروسة هي عمر المنخرط وهي ميزة كمية متقطعة

4) عدد المنخرطين x الذين عمرهم 14 سنة

لدينا الحصيص الإجمالي هو $N=24$

إذن: $5+6+x+8+4=24$ إذن: $23+x=24$

ومنه: $x=24-23$; $\boxed{x=1}$

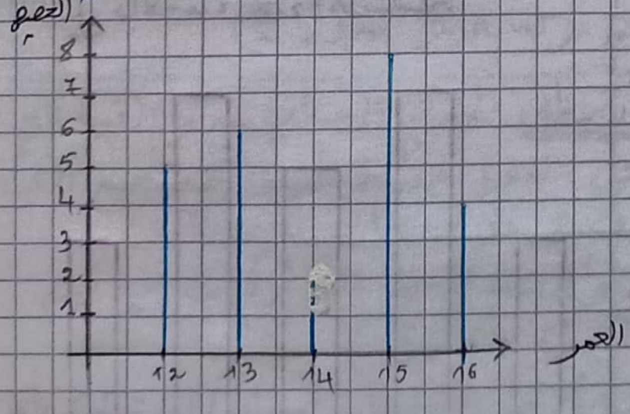
5) لنتمم الجدول السابق

الميزة	12	13	14	15	16
الحصيص	5	6	1	8	4
الحصيص المتراكم	5	11	12	20	24
التردد	$\frac{5}{24}=0,21$	$\frac{6}{24}=0,25$	$\frac{1}{24}=0,04$	0,33	0,17
التردد المتراكم	0,21	0,46	0,50	0,83	1

6) لدينا: $0,21+0,25+0,04+0,33+0,17=1$

نستنتج أن مجموع الترددات يساوي 1

7) لنمثل هذه المتسلسلة بمبيان عصوي



تطبيق ②:

أجريت دراسة إحصائية حول عدد الأطفال بـ 20 أسرة وأعطت النتائج التالية:

-2_1_1_2_3_4_0_3_4_3_2

1_0_3_1_4_3_2_0_1

1) إعط جدول الحصيص والحصيص المتراكمة لهذه المتسلسلة

2) أحسب التردد الموافق لقيمة الميزة 0

3) أحسب النسبة المئوية الموافقة لقيمة الميزة 0

4) أحسب النسبة المئوية لعدد الأسر التي يفوق بها عدد الأطفال طفلين.

5) مثل هذه المتسلسلة الإحصائية بالأعمدة.

حل التطبيق ②:

1) جدول الحصيص والحصيص المتراكمة

الميزة: عدد الأطفال	0	1	2	3	4
الحصيص: عدد الأسر	3	5	4	5	3
الحصيص المتراكم	3	8	12	17	20

2) ليكن f تردد قيمة الميزة 0 إذن:

$$f = \frac{n}{N} = \frac{3}{20} = 0,15$$

3) النسبة المئوية الموافقة لقيمة الميزة 0 هي:

$$p = f \times 100 = 0,15 \times 100 = 15\%$$

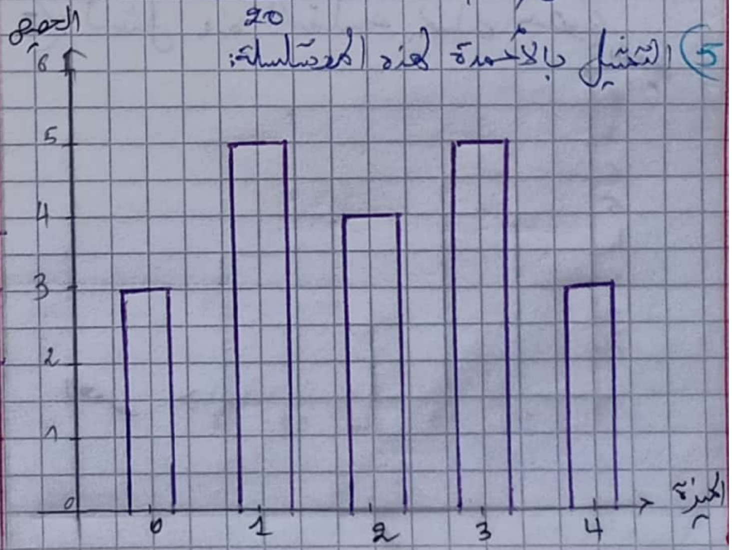
4) لدينا عدد الأسر التي يفوق بها عدد الأطفال طفلين هو: $n = 5+3 = 8$

إذن النسبة المئوية هي:

$$p = f \times 100 = \frac{n}{N} \times 100$$

$$= \frac{8}{20} \times 100 = 40\%$$

5) التمثيل بالأعمدة لهذه المتسلسلة:



2) عرض متسلسلة بالأصناف

إذا كانت الميزة كمية وعدد قيمها كبير فبدل دراسة جميع قيم الميزة، نلجأ إلى حصرها في مجالات $[a,b[$ لها نفس السعة تسمى أصنافا ونسمي مركز الصنف العدد $\frac{a+b}{2}$

* تطبيق ③:

يعطي الكشف التالي توزيع أعمار العاملين في إحدى الضيعات الفلاحية

29_27_36_45_22_17_34_26_16

16_35_33_21_42_32_18_19_25

38_29_27_36_38_22_17_34_26

39_30_32_18_13

1) حدد الساكنة الإحصائية لهذه المتسلسلة

2) حدد الميزة الإحصائية وحدد نوعها

3) أنقل الجدول ثم أكمله

العمر بالسنوات	[10,20[	[20,30[	[30,40[	[40,50[
مركز الصنف				
الحصيص عدد العمال				

4) كم عدد عمال الضيعة الفلاحية.

5) أحسب نسبة العمال الذين أعمارهم أقل من 20 سنة

6) أحسب تردد الصنف $[30,40[$

7) أنشئ مدراج لتوزيع عمال الضيعة الفلاحية حسب أصناف أعمارهم.

* حل التطبيق ③:

1) الساكنة الإحصائية هي عمال الضيعة الفلاحية

2) الميزة الإحصائية هي عمر العمال وهي ميزة كمية متصلة.

3)

العمر بالسنوات	[10,20[	[20,30[	[30,40[	[40,50[
مركز الصنف	$\frac{10+20}{2}=15$	25	35	45
الحصيص عدد العمال	8	10	12	2

4) عدد عمال الضيعة الفلاحية هو الحصيص الإجمالي

$N = 8 + 10 + 12 + 2 = 32$

عدد عمال الضيعة الفلاحية هو 32 عاملا

5) أحسب نسبة العمال الذين أعمارهم أقل من 20 سنة

لدينا عدد العمال الذين تقل أعمارهم عن 20 سنة هو $n = 8$

$p = f \times 100 = \frac{n}{N} \times 100 = \frac{8}{32} \times 100 = 25\%$

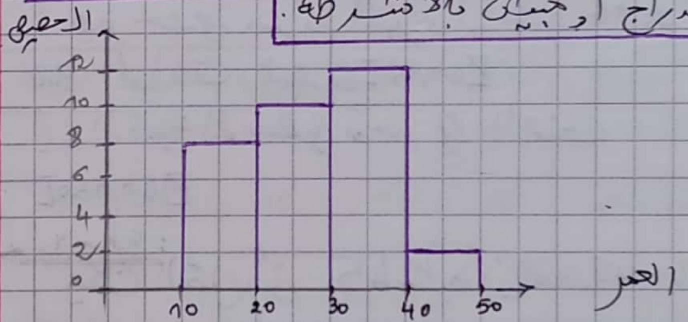
6) تردد الصنف $[30,40[$

ليكن f تردد الصنف $[30,40[$ إذن:

$f = \frac{n}{N} = \frac{12}{32} = 0,375$

7) مدراج لتوزيع عمال الضيعة الفلاحية حسب أصناف أعمارهم

إذا كانت المتسلسلة عبارة عن أصناف $[a,b[$ أو مجالات $a < x < b$ فنستعمل مبيان يسمى مدراج أو مبيان بالأشرطة.



III- وسيطات الوضع:

1) المنوال:

أ- تعريف:

المنوال هو قيمة الميزة التي لها أكبر حصيص

ب- أمثلة:

* متسلسلة بالقيم:

مثال 1: أنظر التطبيق 1

الميزة	12	13	14	15	16
الحصيص	5	6	1	8	4

لدينا أكبر حصيص هو 8 وميزته هي 15

إذن المنوال هو 15

مثال 2:

الميزة (المعاملات)	1	2	3	5
الحصيص (عدد المواد)	3	2	2	3

لدينا أكبر حصيص هو 3، ونلاحظ أن هناك قيمتين للميزة توافقتين لهذا الحصيص هما 1 و 5

إذن لهذه المتسلسلة منوالين هما 1 و 5

مثال 3: أطوال 8 تلاميذ بالسنتمتر هي

70, 59, 64, 66, 73, 67, 68, 61

هل توجد قيمة تكررت أكثر من غيرها في هذه البيانات؟

الجواب لا. إذن لا يوجد منوال لهذه المتسلسلة

* متسلسلة بالأصناف

نعتبر المتسلسلة الإحصائية التالية:

الصنف	$[120,130[$	$[130,140[$	$[140,150[$	$[150,160[$
الحصيص	9	11	12	18

لدينا أكبر حصيص هو 18 الموافق للصنف $[150,160[$

إذن الصنف المنوالي لهذه المتسلسلة هو الصنف $[150,160[$

* ملاحظة: يمكن لمتسلسلة إحصائية أن لا يكون لها منوال، كما يمكن أن يكون لها أكثر من منوال واحد.

2) المعدل الحسابي:

أ- تعريف:

المعدل الحسابي لمتسلسلة إحصائية هو خارج جداءات قيم الميزة (أو مراكز الأصناف) في الحصيصات الموافقة لها على الحصيص الإجمالي ونرمز له بالرمز m (أو $\bar{x}$)

* ملاحظة: المعدل الحسابي m هو القيمة التي يمكن الحصول عليها عن طريق جمع جميع القيم وقسمتها على عددها

(أي هو القيمة المحصل عليها لو كانت جميع قيم الميزة متساوية)

- يمكن أن نسمي المعدل الحسابي أيضا القيمة المتوسطة.

ب - أمثلة:

*** متسلسلة بالقيم:**

← مثال 1: أنظر التطبيق 2

لدينا:
$$m = \frac{(0\times3)+(1\times5)+(2\times4)+(3\times5)+(4\times3)}{20}$$
$$= \frac{0+5+8+15+12}{20} = 2$$

إذن 2 هو معدل عدد الأطفال في كل أسرة.

← مثال 2: نعتبر المتسلسلة التالية:

الميزة (المعاملات)	1	2	3	5
الحصيص (عدد المواد)	3	2	2	3

$$m = \frac{(1\times3)+(2\times2)+(3\times2)+(5\times3)}{10}$$
$$= \frac{3+4+6+15}{10} = \frac{28}{10}$$
$$\boxed{m = 2,8}$$

*** متسلسلة بالأصناف:**

قاعدة: إذا كان $a \leq x < b$ صنف لمتسلسلة إحصائية فإن مركزه هو العدد $\frac{a+b}{2}$

* مثال: أنظر التطبيق 3

العمر بالسنوات	[10,20[	[20,30[	[30,40[	[40,50[
مركز الصنف	$\frac{10+20}{2}=15$	25	35	45
الحصيص عدد العمال	8	10	12	2

$$m = \frac{(15\times8)+(25\times10)+(35\times12)+(45\times2)}{32}$$
$$m = 27,5$$

إذن معدل أعمار العمال هو 27,5 وهذا يعني أنه إذا افترضنا أن للعمال نفس العمر سيكون عمر كل عامل 27,5 سنة.

3) القيمة الوسطية:

أ- تعريف:

> القيمة الوسطية لمتسلسلة إحصائية هي أصغر قيم الميزة التي حصيصها المتراكم أكبر من أو يساوي نصف الحصيص الإجمالي.

ب - أمثلة:

*** متسلسلة بالقيم:**

أنظر التطبيق 1

الميزة	12	13	14	15	16
الحصيص	5	6	1	8	4
الحصيص المتراكم	5	11	12	20	24

- نصف الحصيص الإجمالي هو: $\frac{24}{2} = 12$
- أصغر قيم الميزة التي حصيصها المتراكم أكبر أو يساوي 12 هي 14

إذن القيمة الوسطية هي 14

*** متسلسلة بالأصناف:**

الصنف	[120,130[	[130,140[	[140,150[	[150,160[
الحصيص	9	11	12	18
الحصيص المتراكم	9	20	32	50

- نصف الحصيص الإجمالي هو: $\frac{50}{2} = 25$
- أصغر حصيص متراكم أكبر أو يساوي 25 هو 32 الموافق للصنف [140,150[
- إذن القيمة الوسطية توجد في الصنف [140,150[

*** ملاحظة:** يمكن القول أن 145 (مركز الصنف [140,150[) هي القيمة الوسطية لهذه المتسلسلة الإحصائية

*** تعريف آخر:**

> القيمة الوسطية لمتسلسلة إحصائية تجمع ميزتها مرتبة ترتيبا تصاعديا أو تنازليا هي قيمة الميزة التي تقسم هذه المتسلسلة إلى جزأين لهما نفس الحصيص

← الحالة 1: عدد قيم الميزة عدد فردي

- مثال: عدد الغيابات المسجلة خلال 7 أيام لعمال شركة كان كالتالي: 3 - 1 - 2 - 0 - 4 - 2 - 3

نرتب عدد الغيابات

$$\underbrace{4 - 3 - 3}_{\text{3 قيم}} - 2 - \underbrace{2 - 1 - 0}_{\text{3 قيم}}$$

إذن القيمة الوسطية هي 2

← الحالة ② : عدد حجم السلسلة الإحصائية عدد زوجي

* مثال : عدد الغيابات المسجلة خلال 8 أيام لعمال

شركة كان كالتالي : 3 - 1 - 5 - 0 - 4 - 2 - 1 - 4

نرتب عدد الغيابات : $\underbrace{0 - 1 - 1 - 2}_{4\ \text{قيم}} - \underbrace{3 - 4 - 4 - 5}_{4\ \text{قيم}}$

إذن القيمة الوسطية هي القيمة المحصورة بين

2 و 3 أي $\frac{2+3}{2} = 2{,}5$

إذن القيمة الوسطية هي 2,5

VI - التشتت :

1) تعريف :

نعتبر سلسلتين إحصائيتين x_i و y_i لهما

نفس المعدل الحسابي m.

نقول أن x_i أقل تشتتا من y_i يعني أن

قيم المميزة x_i أقرب إلى المعدل الحسابي m

من قيم المميزة y_i

2) مثال :

نعتبر الجدول التالي :

الفروض	الفرض 1	الفرض 2	الفرض 3	الفرض 4	الفرض 5
نقط حسن	9	14	10	13	14
نقط خالد	8	16	10	17	9

معدل حسن هو : $m_1 = \frac{9+14+10+13+14}{5}$

$= \frac{60}{5} = 12$

معدل خالد هو : $m_2 = \frac{8+16+10+17+9}{5}$

$= \frac{60}{5} = 12$

إذن $m_1 = m_2$

أي أن التلميذين حسن وخالد لهما نفس المعدل

نلاحظ أن نقط حسن أقرب إلى المعدل 12

من نقط خالد

نقول أن : نقط حسن أقل تشتتا من نقط خالد.