الدرس ⑦ : الإحصاء

I - تذكير

1) الدراسة الإحصائية: هي دراسة لظاهرة 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 نقوم بدراستها 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 البياض...)

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

5) الحصيص الإجمالي: هو مجموع الحصيصات ونرمز 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 حصيص 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 وعددها قيمها قليل، نرتب هذه القيم تصاعديا:

تطبيق ①

يمثل الجدول التالي متسلسلة إحصائية تعبر عن توزيع 24 متخرط واحد الأندية الرياضية حسب أعمارهم

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

1) الساكنة الإحصائية هي 24 متخرط واحد الأندية

2) الوحدة الإحصائية هي متخرط

3) الميزة المدروسة هي عمر المتخرط وهي ميزة كمية متقطعة

4) عدد المتخرطين x الذين عمرهم 14 سنة

لدينا الحصيص الإجمالي هو $N=24$

أي: $5+6+x+8+4=24$ ق! $23+x=24$

وعليه، $x=24-23$ ؛ $\boxed{x=1}$

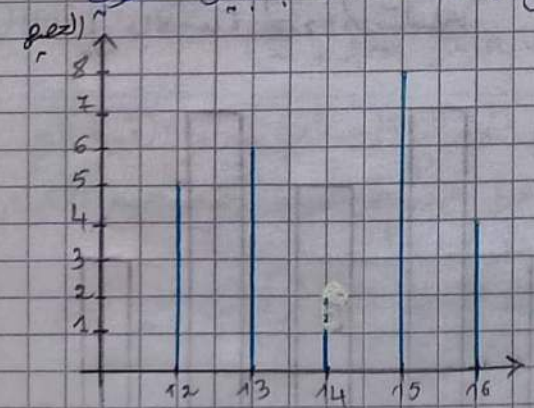
5) لنتمم الجدول السابق

العمر	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) لتمثيل هذه المتسلسلة بيانيا نختار



تطبيق ②:

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

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

1) أعط جدول الحصص والحصص المتراكمة لهذه السلسلة

2) احسب التردد الموافق لقيمة الميزة 0

3) احسب النسبة المئوية الموافقة لقيمة الميزة 0

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

5) مثل هذه السلسلة الإحصائية بالأعمدة.

حل التطبيق ②:

1) جدول الحصص والحصص المتراكمة

الميزة: عدد الأطفال	0	1	2	3	4
الحصص: عدد الأسر	3	5	4	5	3
الحصص 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:

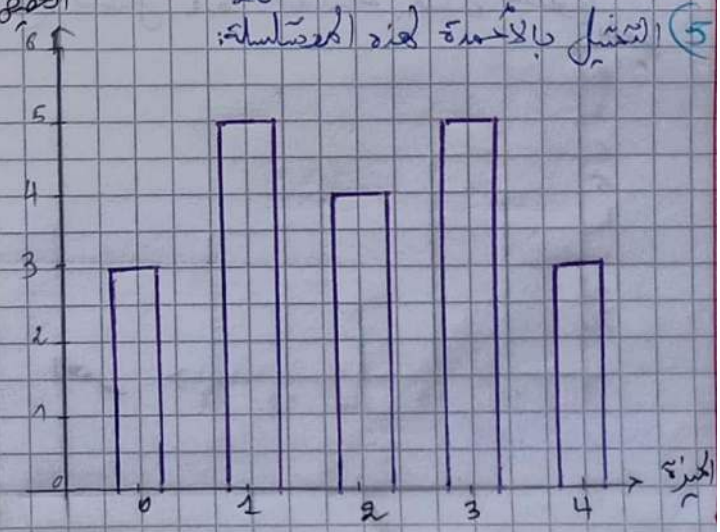
$$p = f \times 100 = 0,15 \times 100 = 15\%$$

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 السلسلة:



2) سلسلة بالأقران:

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

تطبيق ③:

يعطي الجدول 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 السلسلة

2) حدد الميزة الإحصائية وحدد نوعها

3) أنقل الجدول ثم أكمله

العمر بالسنوات	$[10,20[$	$[20,30[$	$[30,40[$	$[40,50[$
مركز الصنف				
الحصص: عدد العمال				

4) كم عدد عمال الضيعة الفلاحية.

5) احسب نسبة العمال الذين أعمارهم أقل من 20 سنة

6) احسب تردد الصنف $[30,40[$

7) أنشئ مدرج 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
الحصص: عدد العمال	8	10	12	2

4) عدد عمال الضيعة الفلاحية هو الحصى 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 أقل أعمارهم عن 20 سنة هو $n = 8$

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

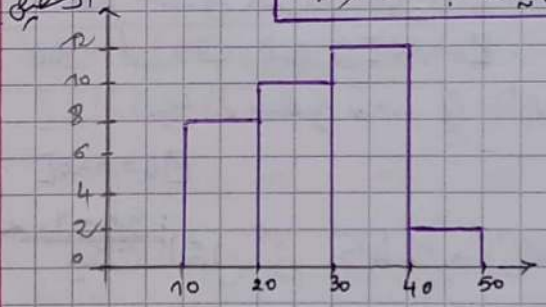
6) حدد الصنف $[30,40[$

ليكن f تردد الصنف $[30,40[$ إذن:

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

7) مدرج لتوزيع عمال الضيعة الفلاحية حسب أعمارهم

إذا كانت الحدسلة عبارة عن أصناف $[a,b[$ أو مجالات $a \leq x < b$ فنشعل مبيان يسمى مدراج أو مبيان بالأعمدة.



III - وسيطات الوضع:

1) المنوال:

أ- تعريف:

المنوال هو قيمة الحيزة التي لها أكبر حصى

ب- أمثلة:

* حدسلة بالحيز:

مثال ①: أنظر التطبيق ①

الحيزة	12	13	14	15	16
الحصى	5	6	1	8	4

لدينا أكبر حصى هو 8 وحيزته هي 15

إذن المنوال هو 15

مثال ②:

الحيزة (العاملات)	1	2	3	5
الحصى (عدد الولد)	3	2	2	3

لدينا أكبر حصى هو 3، نلاحظ أن هناك قيمتين للحيزة حوافقتين لهذا الحصى هما 1 و 5

إذن لهذه الحدسلة منوالين هما 1 و 5

مثال ③: أطوال 8 تلاميذ بالسنتمتر هي

70، 59، 64، 66، 73، 67، 68، 61

هو يوجد قيمة تتكرر أكثر من غيرها في هذه البيانات؟ الجواب لا. إذن لا يوجد منوال لهذه الحدسلة

* حدسلة بالأصناف

نعتبر الحدسلة الإحصائية التالية:

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

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

إذن الصنف المنوال لهذه الحدسلة هو الصنف $[150,160[$

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

2) العمل الحسابي:

أ- تعريف:

العمل الحسابي لحدسلة إحصائية هو خارج جداء قيم الحيزة (أو مراكز الأصناف) في الحصيات الموافقة لها على الحصى الإجمالي، ويرمز له بالرمز m (أو $\bar{x}$)

* ملاحظة: العمل 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 الحيزة متساوية)

- يمكن أن نسمي العمل الحسابي أيضا القيمة المتوسطة.

ب - أمثلة:

* متسلسلة بالقيم:

مثال ①: أنظر التطبيق ②

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

إذن 2 هو معدل عدد الأطفال في كل أسرة.

مثال ②: نعتبر المتسلسلة التالية:

القيمة (العائلات)	1	2	3	5
الحصيص (عدد الجواد)	3	2	2	3

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

$$m = 2,8$$

* 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 لعمال نفس العمر سيكون عمر كل عامل 27,5 سنة

3) القيمة الوسطية:

أ - تعريف:

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

ب - أمثلة:

* متسلسلة بالقيم:

أنظر التطبيق ①

المميزة	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 المميزة التي حصيصها المتراكم أكبر أو يساوي 12 هي 14

إذن القيمة الوسطية هي 14

* متسلسلة بالأصناف

الصنف	[130,140[	[140,150[	[150,160[
الحصيص	9	11	12	1
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 قيم مميزة كما هي مرتبة ترتيبا تصاعديا أو تنازليا هي قيمة المميزة التي تقسم هذه المتسلسلة إلى جزأين لهما نفس الحصيص

← الحالة ①: عدد قيم المميزة عدد فردي

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

نرتب عدد الغيابات

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

إذن القيمة الوسطية هي 2

← الحالة ② : عدد قيم السلسلة الإحصائية عدد زوجي

* مثال : عدد الغيابات المسجلة خلال 8 أيام لعمال شركة كان كالتالي : 4 - 1 - 2 - 4 - 0 - 5 - 1 - 3

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 لهما نفس المعدل الحسابي m.

نقول أن السلسلة الأقل تشتتا هي السلسلة التي قيم متغيرتها الأكثر قربا إلى المعدل الحسابي m هي قيم المتغيرة الأخرى.

2) مثال :

نعتبر الجدول التالي :

الفرق	الفرق 1	الفرق 2	الفرق 3	الفرق 4	الفرق 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 12

نلاحظ أن نقط حسن أقرب إلى المعدل 12 من نقط خالد

نقول أن : نقط حسن أقل تشتتا من نقط خالد.