الإحصاء

I - مصطلحات و تعاريف

1- الساكنة الإحصائية :

الساكنة الإحصائية هي المجموعة التي تخضع لدراسة إحصائية
وكل عنصر من هذه المجموعة يسمى فردا أو وحدة إحصائية.

ميزة إحصائية أو المتغير الإحصائي:

ميزة إحصائية هي الخاصية موضوع الدرس,فهي كمية أو كيفية.

- **ميزة كمية** هي التي تترجم عدديا .
  أمثلة القامة- المحصول الفلاحي- استهلاك الماء.........
- **ميزة كيفية** هي التي لا تترجم إلى عدد .
  أمثلة فصيلة الدم - الجنس.............................

ملاحظة: الميزة الكمية فهي متقطعة فتأخذ قيما أو متصلة فيعبر عنها بالأصناف.

2- الحصيص و الحصيص المتراكم – التردد و التردد المتراكم

الحصيص:

الحصيص n_i الموافق لقيمة الميزة x_i (أو الموافق الصنف I_i) هو العدد المرات لتي تتكرر فيها القيمة x_i (أو هو عدد القيم التي تنتمي إلى الصنف I_i)

الحصيص المتراكم:

الحصيص المتراكم الموافق لقيمة الميزة x_i (أو الموافق الصنف I_i) هو العدد N_i
حيث $N_i=n_1+n_2+n_3+.................+n_i$
حيث n_1 و n_2 و.........n_i هي حصيصات القيم التي أصغر أو تساوي x_i

الحصيص الإجمالي:

الحصيص الإجمالي N هو مجموع جميع الحصيصات

التردد:

التردد f_i الموافق للقيمة الميزة x_i أو الصنف I_i هو العدد $f_i=\dfrac{n_i}{N}$

ملاحظة **مجموع جميع الترددات يساوي 1**

التردد المتراكم **F_i الموافق للقيمة الميزة x_i أو الصنف I_i هو** $F_i=f_1+f_2+......f_i$

النسبة المئوية:

النسبة المئوية P_i الموافق للقيمة الميزة x_i أو الصنف I_i هي $P_i=100f_i$ حيث f_i التردد الموافق لـ x_i أو I_i

- مجموعة الأزواج $(x_i;n_i)$ تسمى متسلسلة احصائية حيث n_i الحصيص الموافق للقيمة x_i

أ- ميزة كمية متقطعة

مثال2

نعتبر الكشف التالي الذي يعطينا معطيات إحصائية حول عددالغرف في منازل أحد الأحياء

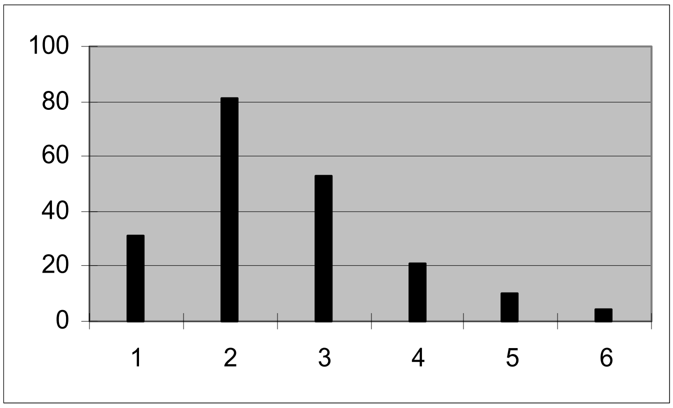
3	4	2	2	3	1	5	2	4	3
5	6	2	3	4	2	2	2	3	4
2	2	2	1	3	3	3	4	2	1
2	1	2	2	3	4	5	2	3	1
3	3	2	2	2	5	6	1	2	2
3	3	2	2	1	2	3	2	2	2
4	3	1	3	3	2	2	1	5	4
3	3	4	4	2	2	2	2	1	2
4	2	2	1	2	3	3	3	3	2
3	3	3	2	2	2	2	1	1	6
5	3	1	3	3	3	2	1	5	4
2	3	2	4	3	2	4	2	1	2
4	1	2	1	2	3	2	3	3	3
3	1	3	2	2	2	2	1	1	4
2	2	2	1	3	3	3	4	2	1
1	2	2	2	3	2	5	2	3	1
3	3	2	2	2	5	6	1	2	2
3	2	2	1	1	2	3	2	2	2
3	2	1	4	3	2	2	1	5	4
2	3	4	4	2	3	2	3	1	2

يعطينا هذا الكشف معلومات تهم ساكنة احصائية تتكون من **200** وحدة إحصائية.إذن الحصيص الإجمالي هو**200**
الميزة المدروسة هي عدد الغرف (ميزة كمية متقطعة)

نلاحظ أن العدد **1** يتكرر**31** مرة نقول إن **31** هو الحصيص الموافقللقيمة **1**
انطلاقا من هذا الكشف يمكن تكوين جدول إحصائي و ذلك بتنظيم المعلومات على الشكل التالي : يحتوي على قيم **x_i** مرتبة ترتيبا تزايديا و حصيصات موافقة لها، و ترددات موافق لها.

قيمة الميزة x_i	**1**	**2**	**3**	**4**	**5**	**6**
الحصيص n_i	**31**	**81**	**53**	**21**	**10**	**4**
الحصيص المتراكم N_i	**31**	**112**	**165**	**186**	**196**	**200**
التردد f_i	**0,155**	**0,405**	**0,265**	**0,105**	**0,05**	**0,02**
التردد المتراكمF_i	**0,155**	**0,56**	**0,825**	**0,93**	**0,98**	**1**

رغم ما تمتاز به الجداول من الدقة فإنها لا تعطينا فكرة واضحة و سريعة عن الظاهرة التي نحن بصدد دراستها.
لذا نعمد إلى تمثيل الجداول الإحصائية مبيانيا
التمثيل المبياني للحصيص



<u>مخطط عصوي للحصيص</u>

بنفس الطريقة نمثل الحصيص المتراكم و التردد و التردد المتراكم

<u>ب- ميزة كمية متصلة</u>

<u>مثال1</u>

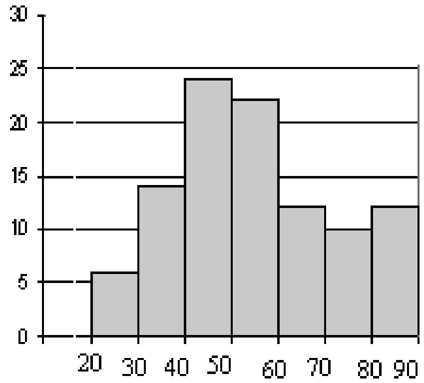
الكشف التالي يتضمن معطيات إحصائية تتعلق بثمن نفس الكمية من منتوج فلاحي (بالدرهم) في نقط مختلفة للبيع.

45	80,5	46	41,5	41	51	20	40	84	43
41	32,5	54	43	21,5	69	61,5	37,5	82	67
48	84	56	70,5	58	25	44	70	32,5	43
64	68	51	75	43	81	50	48	86	60,5
29	48	59	74	48	30,5	56	58	49,5	33,5
34	53	53	42	28	59	67	72	77	45
60	55,5	33	63	44,5	34,5	38,5	56,5	44	51
53	78,5	38	38	25,5	62,5	77,5	57	67	47
34	55	67	69	31	37	44	47	51,5	58
55	49	34	44	37,5	74	56	37	72,5	67

يعطينا هذا الكشف معلومات عن ساكنة إحصائية تتكون من 100 وحدة إحصائية . الميزة المدروسة ثمن المنتوج الفلاحي
نلاحظ أنه ليس هناك تكرار كبير للمعلومات
لتبسيط الدراسة نعمد إلى تجميع المعلومات في مجالات لها نفس السعة تسمى **أصنافا**.
و بذل دراسة جميع قيم الميزة نختار في كل صنف قيمة وحيدة هي مركز الصنف و تسمى **قيمة الصنف.**
قيمة الصنف $[a;b[$ هي $\frac{a+b}{2}$
في المثال الذي لدينا يمكن تجميع المعلومات في مجالات سعته10
فنحصل مثلا على الصنف $[20;30[$ قيمة هذا الصنف هي 25
نقول في هذه الحالة ان الميزة المدروسة **ميزة كمية متصلة**

التردد f_i	الحصيص المتراكم N_i	الحصيص n_i	قيمة الصنف x_i	الصنف $[a_{i-1};a_i[$
0,06	6	6	25	$[20;30[$
0,14	20	14	35	$[30;40[$
0,24	44	24	45	$[40;50[$
0,22	66	22	55	$[50;60[$
0,12	78	12	65	$[60;70[$
0,10	88	10	75	$[70;80[$
0,12	100	12	85	$[80;90[$

التمثيل المبياني للحصيص



مدراج للحصيص

بالمثل نمثل التردد و الحصيص المتراكم ....

صفة عامة

عندما تأخذ الميزة الإحصائية عددا كبيرا من القيم فإننا نغطي مجموع هذه القيم بمجالات تسمى أصنافا $I_1=[a_0;a_1[\quad I_2=[a_1;a_2[\quad ....\quad I_n=[a_{n-1};a_n[$

و يرمز له بـ n_i **الحصيص هو عدد الوحدات التي تأخذ فيا الميزة قيمة تنتمي إلى الصنف** I_i

مجموعة الأزواج $(I_i;n_i)$ **تسمى متسلسلة معبر عنها بالأصناف.**

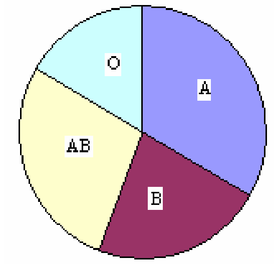
ج – ميزة كيفية

مثال3 نعتبر الكشف التالي الذي يحتوي على فصيلة الدم لـ 180 فردا
كما يلي60 فرد الفصيلة A و 40 فصيلة B و 50 فصيلةAB و 30 فصيلة O
الجدول الإحصائي

O	AB	B	A	الميزة
30	50	40	60	الحصيص
60°	100°	80°	120°	α_i

$$\alpha_i = n_i \frac{360}{180}$$

المخطط الدائري



II- وسيطات الوضع

1- المنوال

تعريف

منوال متسلسلة إحصائية هو كل قيمة أو صنف أو نوع له أكبر حصيص.

أمثلة

في المثال 1 السابق : 2 منوال للمتسلسلة الإحصائية

في المثال 2 السابق : $[40;50[$ منوال للمتسلسلة الإحصائية

في المثال 3 السابق : الفصيلة A منوال للمتسلسلة الإحصائية

2- القيمة الوسطية

أ- تعريف

لتكن متسلسلة ذات ميزة كمية و M عدد حقيقي يحقق الخاصية التالية : نصف وحدات الساكنة الإحصائية على الأقل تأخذ فيها الميزة قيمة أصغر من أو تساوي M و نصف وحدات الساكنة الإحصائية على الأقل تأخذ فيها الميزة قيمة أكبر من أو تساوي M

مثال

الجدول التالي يعطي النقط التي حصل عليها تلاميذ أحد الأقسام

النقطة	2	7	8	10	11	12	16
الحصيص	3	10	5	4	5	2	1
الحصيص المتراكم	3	13	18	22	27	29	30

نلاحظ أكثر من نصف عدد التلاميذ حصلوا على نقطة أصغر من أو تساوي 8. و أكثر من نصف عدد التلاميذ حصلوا على نقطة أكبر من أو تساوي 8

إذن العدد **8 قيمة وسطية** لهذه المتسلسلة الإحصائية.

ب- مبرهنة

أصغر قيم الميزة التي حصيصها المتراكم أكبر من أو يساوي نصف الحصيص الإجمالي هي قيمة وسطية في متسلسلة غير معبر عنها بالأصناف.

مثال

في المثال السابق لدينا $\frac{N}{2}=\frac{30}{2}=15$ و أصغر قيم الميزة التي حصيصها المتراكم أكبر من أو يساوي 15 هي8

إذن العدد **8 قيمة وسطية**

ج- مبرهنة

لتكن $([a_{i-1};a_i[;n_i)$ متسلسلة معبر عنها بالأصناف و N_i الحصيص المتراكم الموافق لصنف $[a_{i-1};a_i[$.

القيمة الوسطية لهذه المتسلسلة الإحصائية هي القيمة M

المحددة ب $M=(a_k-a_{k-1})\dfrac{\frac{N}{2}-N_{k-1}}{n_k}+a_{k-1}$

حيث k هو العدد الصحيح الطبيعي الذي يحقق $N_{k-1}\le\frac{N}{2}\prec N_k$ **(نأخذ $N_0=0$)**

ملاحظة N_k **يوافق** $[a_{k-1},a_k[$ n_k **يوافق** $[a_{k-1},a_k[$

مثال

الصنف $[a_{i-1};a_i[$	الحصيص n_i	الحصيص المتراكم N_i
$[20;30[$	**6**	**6**
$[30;40[$	**14**	**20**
$[40;50[$	**24**	**44**
$[50;60[$	**22**	**66**
$[60;70[$	**12**	**78**
$[70;80[$	**10**	**88**
$[80;90[$	**12**	**100**

لدينا $\frac{N}{2}=\frac{100}{2}=50$ و $44 \leq 50 \prec 66$ $(N_k=66 \quad N_{k-1}=44)$

الحصيص المتراكم 66 موافق لصنف $[50;60[$

الحصيص 22 موافق لصنف $[50;60[$

إذن $M=(60-50)\frac{50-44}{22}+50=\frac{580}{11}$

3- المعدل الحسابي

تعريف **لتكن $(x_1;n_1);(x_2;n_2);.........;(x_p;n_p)$ متسلسلة إحصائية حيث x_i هو قيمة الميزة (أو قيمة الصنف I_i) و n_i هو الحصيص الموافق لـ x_i.**
الوسط أو المعدل الحسابي هو العدد

$$\overline{X}=\frac{x_1n_1+x_2n_2+x_3n_3+.............+x_pn_p}{n_1+n_2+n_3+.......................+n_p}$$

أمثلة (نأخذ الأمثلة السابقة)

خاصية

لتكن $\overline{x}$ المعدل الحسابي لمتسلسلة حصيصها الاجمالي N و $\overline{x'}$ المعدل الحسابي لمتسلسلة أخرى حصيصها الاجمالي N'

المعدل الحسابي للمتسلسلة المكونة من تجميع المتسلسلتين هو $\frac{N\overline{x}+n'\overline{x'}}{N+N'}$

III – وسيطات التشتت

1- نشاط تمهيدي

يعطي الجدولان التاليان نقط 20 تلميذا في مادة الرياضيات و الفرنسية.

الرياضيات

النقطة	8	9	10	11	12	13	14	15
الحصيص	1	1	3	5	2	2	2	4

الفرنسية

النقطة	2	5	7	8	10	11	12	14	15	16	17	18	19	20
الحصيص	1	2	1	2	1	3	1	2	1	1	2	1	1	1

حدد وسيطات الوضع(المنوال – القيمة الوسطية – المعدل الحسابي)
لاحظ أن لهما نفس وسيطات الوضع
أنجز مخططا عصويا لكل منهما

رغم أن لهذين المتسلسلتين نفس وسيطات الوضع إلا أنهما يختلفان جذريا. فالنقط التي حصل عليها التلاميذ في الرياضيات تتجمع حول القيمة 11 في حين نلاحظ تشتت نقط الفرنسية بين 2 و 20

يبين هذا أن وسيطات الوضع غير كافية لإعطاء نظرة كاملة علىمتسلسلة إحصائية ، وهذا ما يتطلب أخرى تسمى

وسيطات التشتت

2- الانحراف المتوسط

تعريف

الانحراف المتوسط لمتسلسلة إحصائية $(x_i ; n_i)_{1\le i\le p}$ **هو العدد**

$$\rho = \frac{\sum_{i=1}^{i=p} n_i \left|x_i - \bar{x}\right|}{N}$$

$$\rho = \frac{n_1\left|x_1-\bar{x}\right| + n_2\left|x_2-\bar{x}\right| + \ldots\ldots\ldots\ldots n_p\left|x_p-\bar{x}\right|}{N}$$

حيث $\bar{x}$ **المعدل الحسابي و** N **الحصيص الإجمالي.**

مثال نأخذ النشاط السابق

الرياضيات

النقطة x_i	8	9	10	11	12	13	14	15
الحصيص n_i	1	1	3	5	2	2	2	4
$\left\|x_i-\bar{x}\right\|$	4	3	2	1	0	1	2	3

$$\rho_M = \frac{4+3+6+5+0+2+4+12}{20} = 1,8$$

بالمثل بالنسبة الفرنسية نحصل $\rho_F = 4,2$

نلاحظ $\rho_M \prec \rho_F$ و هذا يبين أن النقط الرياضيات أقل تشتتا من نقط الفرنسية

3- الانحراف الطرازي و المغايرة

تعريف

مغايرة متسلسلة إحصائية $(x_i ; n_i)_{1\le i\le p}$ **هو العدد** $v = \frac{1}{N}\sum_{i=1}^{p} n_i \left(x_i - \bar{x}\right)^2$

الانحراف الطرازي لهذه المتسلسلة هو $\sigma = \sqrt{v}$

ملاحظة

* $v = \left(\frac{1}{N}\sum_{i=1}^{p} n_i {x_i}^2\right) - \bar{x}^2$

*** إذا كانت المتسلسلة معبرا عنها بالأصناف فنعتبر** x_i **قيمةالصنف.**

مثال

المثال السابق

الرياضيات

النقطة x_i	8	9	10	11	12	13	14	15
الحصيص n_i	1	1	3	5	2	2	2	4
$(x_i-\bar{x})^2$	16	9	4	1	0	1	4	9

$$\sigma_M = 2\sqrt{1,1} \quad ; \quad v_M = 4,4$$