تابع للمحور الثاني: العرض الجدولي والبياني للبيانات

The second axis: Tabular and Graphical presentation of data

محتوى المحاضرة الرابعة

تمهيد: مفاهيم عامة

العرض الجدولي والبياني للبيانات الكمية المستمرة(مثال تطبيقي):

1- تحديد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي، نوعه
2- ترتيب البيانات تصاعديا
3- تحديد عدد الفئات:
4- حساب المدى
5- تحديد طول الفئات
6- تفريغ البيانات ورسم جدول التكرار المطلق والتكرار النسبي والتكرار المئوي والتكرار المتجمع الصاعد والتكرار المتجمع النازل :
7- العرض البياني للبيانات الكمية المستمرة
8- التمثيل البياني للتكرار المتجمع الصاعد والنازل للبيانات الكمية المستمرة
9- تعديل الفئات

تمهيد:

إذا كان المتغير المدروس متغير كمي مستمر فإن مجال الدراسة يضم مالا نهاية من القيم، ولصعوبة التعامل مع كل تلك القيم يتم تقسيم مجال الدراسة الى مجالات جزئية تسمى **فئات** والتي يتحدد عددها وطولها حسب حجم العينة، ثم تبويبها في جدول توزيع تكراري. وللقيام بذلك نتبع الخطوات التالية:

أ- **تحديد عدد الفئات(Number of classes):**

عادة يكون عدد الفئات محصورا بين 5 و15. وذلك لكي يكون الجدول مختصر وله معنى واضح ودقيق، فعدد فئات أقل من 5 يجعل الجدول مختصر جدا ويفقد دقة وتفاصيل البيانات، وإذا كان أكثر من 15 يصبح ضخما يصعب تحليله وقراءته. وقد اجتهد بعض العلماء في وضع معادلات متفق عليها لتحديد عدد الفئات أهمها:

- **معادلة ستورجس(Sturges rule):**

$$K_s = 1 + 3{,}322 \log(N)$$

- **معادلة يول(Yule rule):**

$$K_y = 2{,}5 \sqrt[4]{N}$$

حيث: N يمثل عدد القيم ، K يمثل عدد الفئات

ب- **المدى(Range):** وهو الفرق بين أكبر قيمة وأقل قيمة من بيانات العينة المدروسة، ويرمز له بالرمز: R، حيث:

$$R = X_{Max} - X_{Min}$$

ج- **تحديد طول الفئات(Class length):**

يحدد طول الفئات بالعلاقة التالية: $L = \frac{R}{K}$

- مع التأكد من تحقق الشرط: $K \cdot L \geq R$
- أما في حالة عدم تحقق الشرط فيجب إضافة واحد إلى عدد الفئات أو واحد إلى طول الفئة

د- **مراكز الفئات:**

ويرمز لها بالرمز mi، حيث: **(الحد الأعلى للفئة + الحد الأدنى للفئة) / 2 = mi**

مثال تطبيقي:

تمثل البيانات التالية قيمة المبيعات بآلاف الدينار الجزائري لـ 50 محل تجاري:

36	45	31	28	41	32	29	26	48	32
30	28	33	27	40	31	30	40	35	45
32	37	36	31	35	33	38	39	41	30
38	42	35	33	36	38	39	37	45	23
36	46	30	35	32	49	43	37	22	34

1- **تحديد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي، نوعه:**

- المجتمع الإحصائي: 50 محل تجاري
- الوحدة الإحصائية: المحل التجاري
- المتغير المدروس: قيمة المبيعات
- نوعه: كمي مستمر(متصل) (Quantitative continuous variable)

2- **ترتيب البيانات تصاعديا:**

$22_{(1)}, 23_{(1)}, 26_{(1)}, 27_{(1)}, 28_{(2)}, 29_{(1)}, 30_{(4)}, 31_{(3)}, 32_{(4)}, 33_{(3)}, 34_{(1)}, 35_{(4)}, 36_{(4)}, 37_{(3)}, 38_{(3)}, 39_{(2)},$

$40_{(2)}, 41_{(2)}, 42_{(1)}, 43_{(1)}, 45_{(3)}, 46_{(1)}, 48_{(1)}, 49_{(1)}$

3- **تحديد عدد الفئات:**

$$K_s = 1 + 3{,}322 \log(N)$$

$$K_s = 1 + 3.322 \log(50) = 6.64 \approx 7$$

أو:

$$K_y = 2{,}5 \sqrt[4]{N}$$

$$K_y = 2.5 \sqrt[4]{50} = 2.5\ (2.65) = 6.64 \approx 7$$

4- **حساب المدى(Range):**

$$R = X_{Max} - X_{Min}$$

$$R = 49 - 22 = 27$$

5- **تحديد طول الفئات:**

$$L = \frac{R}{K} = \frac{27}{7} = 3.85 \approx 4$$

التحقق من الشرط: $K \cdot L \geq R$

$K \cdot L = 4 \cdot 7 = 28 > R = 27$ ⇒ الشرط محقق

6- **تفريغ البيانات ورسم جدول التكرار المطلق والتكرار النسبي والتكرار المئوي والتكرار المتجمع الصاعد والتكرار المتجمع النازل:**

الجدول: التكرار المطلق والتكرار النسبي والتكرار المئوي، التكرار المتجمع الصاعد والتكرار المتجمع الصاعد النسبي والتكرار المتجمع الصاعد المئوي، التكرار المتجمع النازل والتكرار المتجمع النازل النسبي والتكرار المتجمع النازل المئوي

x_i قيمة المبيعات	f_i عدد المحلات التجارية	rf_i	pf_i	$cf_i\uparrow$	$rcf_i\uparrow$	$pcf_i\uparrow$	$cf_i\downarrow$	$rcf_i\downarrow$	$pcf_i\downarrow$
[22 26[	2	0,04	4%	2	0,04	4%	50	1	100%
[26 30[	5	0,1	10%	7	0,14	14%	48	0,96	96%
[30 34[	14	0,28	28%	21	0,42	42%	43	0,86	86%
[34 38[	12	0,24	25%	33	0,66	66%	29	0,58	58%
[38 42[	9	0,18	18%	42	0,84	84%	17	0,34	34%
[42 46[	5	0,1	10%	47	0,94	94%	8	0,16	16%
[46 50[	3	0,06	6%	50	1	100%	3	0,06	6%
Σ	50	1	100%	/	/	/	/	/	/

المصدر: معطيات فرضية

- **تفسير بعض قيم جداول التوزيع التكراري:**

أ- قيم جداول التكرار المطلق والتكرار النسبي والتكرار المئوي:

- **الفئة الأولى** [22 26[: قيمة المبيعات **أكبر من أو يساوي 22** ألف دينار جزائري و**أقل تماما من 26 ألف دينار جزائري.**
- $f_2 = 5$: يوجد 5 محلات تجارية مبيعاتها **أكبر من أو يساوي 26** ألف دينار جزائري و**أقل تماما من 30 ألف دينار جزائري .**
- $f_4 = 12$: يوجد 12 محل تجاري مبيعاتهم **أكبر من أو يساوي 34** ألف دينار جزائري و**أقل تماما من 38 ألف دينار جزائري.**
- $pf_3 = 28\%$: 28 بالمئة من إجمالي المحلات التجارية مبيعاتهم **أكبر من أو يساوي 30** ألف دينار جزائري و**أقل تماما من 34 ألف دينار جزائري .**

- $pf_5 = 18\%$: 18 بالمئة من إجمالي المحلات التجارية مبيعاتهم **أكبر من أو يساوي 38** ألف دينار جزائري و**أقل تماما من 42 ألف دينار جزائري** .

ب- قيم جداول التكرار المتجمع الصاعد والتكرار المتجمع الصاعد المئوي، التكرار المتجمع النازل والتكرار المتجمع النازل المئوي:

- **7** = $cf_2\uparrow$: هناك7 محلات تجارية من بين 50 مبيعاتهم **أقل تماما من 30 ألف دينار جزائري(وأكبر من أو يساوي 22** ألف دينار جزائري).
- **47** = $cf_6\uparrow$: هناك47 محل تجاري من بين 50 مبيعاتهم **أقل تماما من 46 ألف دينار جزائري(وأكبر من أو يساوي 22** ألف دينار جزائري).
- $pcf_3\uparrow = 42\%$: 42 **%** من إجمالي المحلات التجاري مبيعاتهم **أقل تماما من 34 ألف دينار جزائري(وأكبر من أو يساوي 22** ألف دينار جزائري).
- 43 = $cf_3\downarrow$: هناك43 محل تجاري من بين 50 مبيعاتهم **أكبر من أو يساوي 30 ألف دينار جزائري(وأقل من أو يساوي 50** ألف دينار جزائري).
- $pcf_4\downarrow = 58\%$: 58 **%** من إجمالي المحلات التجاري مبيعاتهم **أكبر من أو يساوي 34 ألف دينار جزائري(وأقل من أو يساوي 50** ألف دينار جزائري).

7- العرض البياني للبيانات الكمية المستمرة (Quantitative continuous variables)

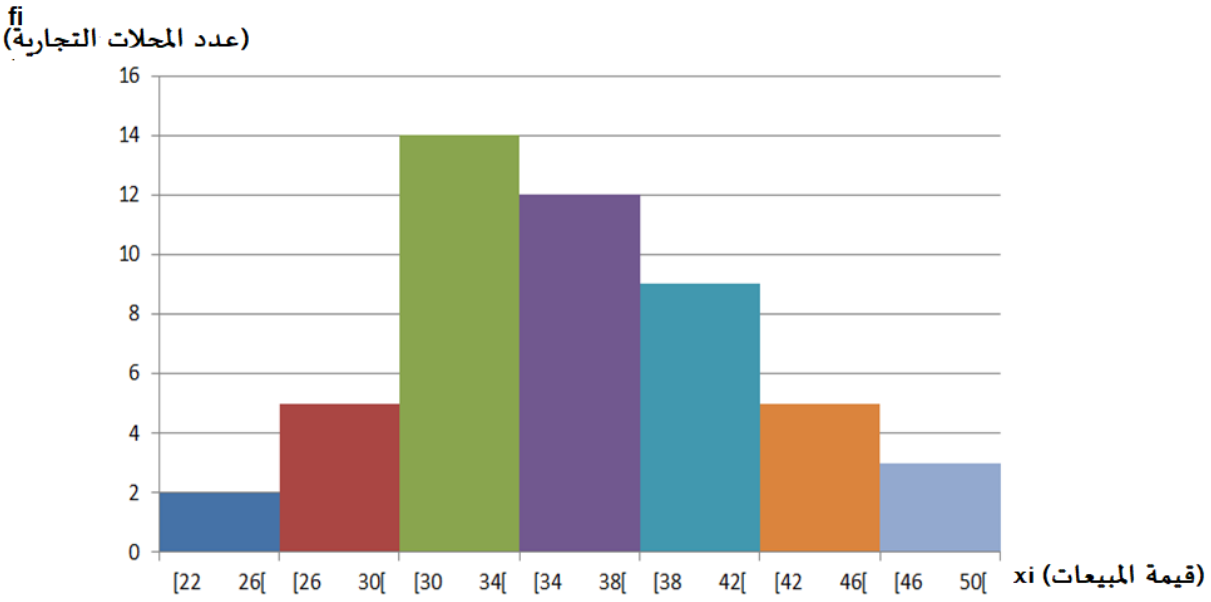
التمثيل البياني المناسب للبيانات الكمية المستمرة يكون كما يلي:

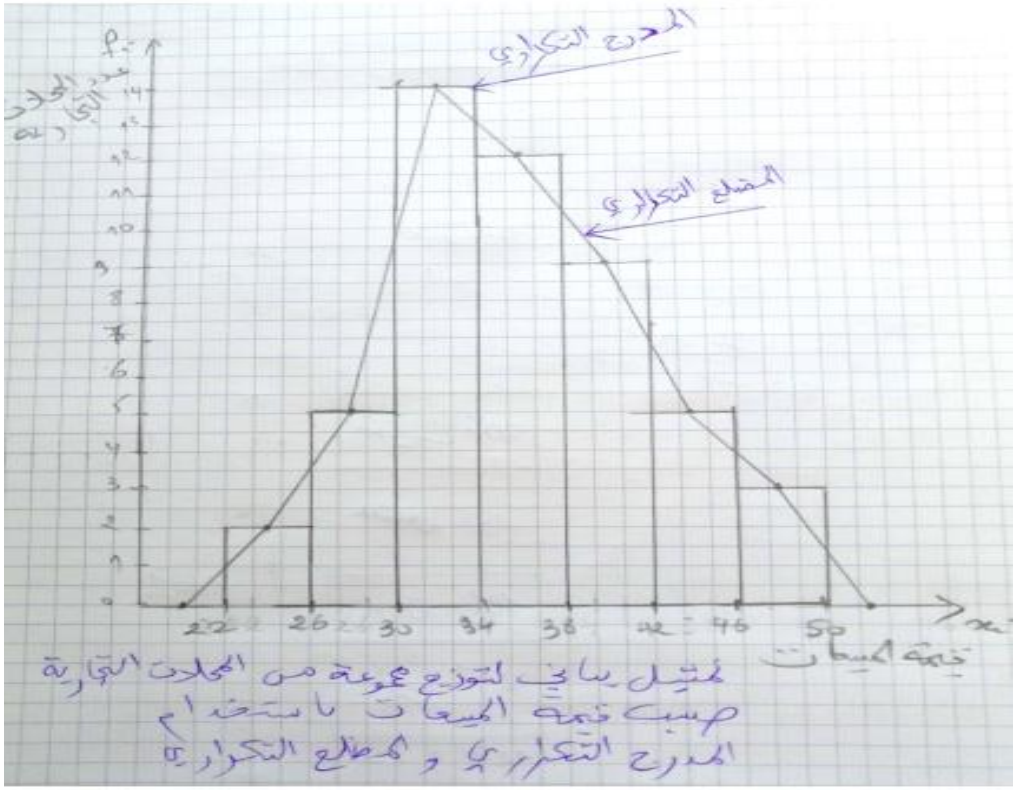
- **المدرج التكراري:**

وهو عبارة عن مستطيلات **متلاصقة** ترسم على المحور الأفقي والذي يمثل الفئات، أما المحور العمودي فيمثل التكرارات العادية المقابلة لكل فئة.

- **المضلع التكراري:**

ويرسم بخط منكسر يصل بين نقاط إحداثياتها (x,y)، حيث: **x**: مركز الفئة التي تمثلها و **y** : التكرار المقابل للفئة. ويجب أن يغلق المضلع التكراري من طرفيه على المحور الافقي، ويتم ذلك بإضافة فئتين وهميتين قبل الفئة الاولى وبعد الفئة الاخيرة يكون تكرارهما **معدوما** وذلك حتى تكون المساحة داخل المدرج التكراري مساوية للمساحة داخل المضلع التكراري.





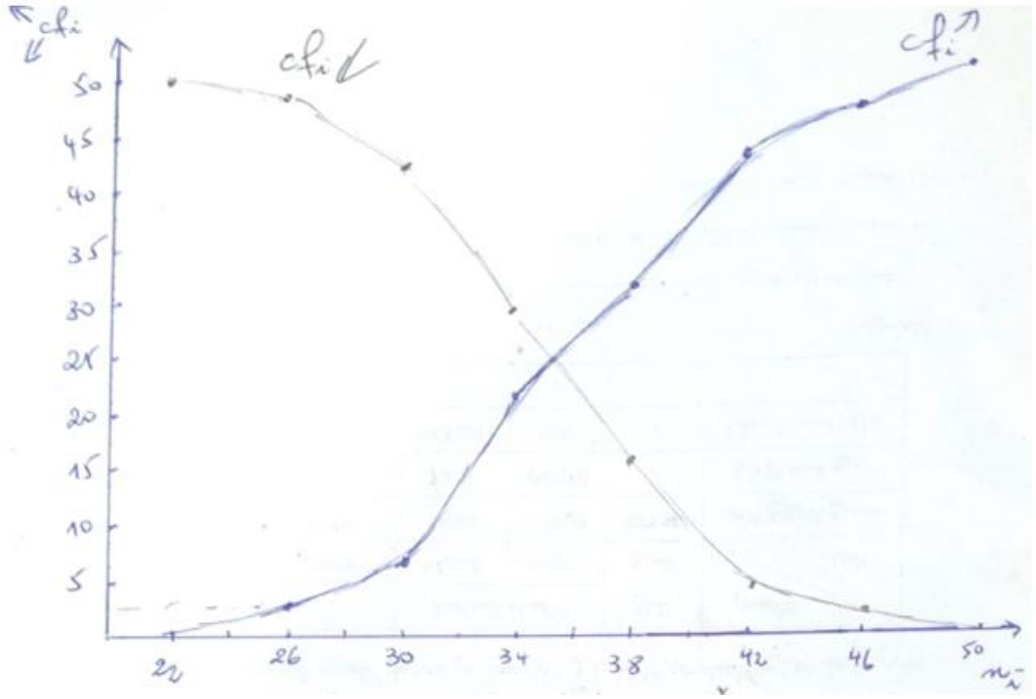
ملاحظة:

من التمثيل البياني للمضلع التكراري نلاحظ أن القيمتين 20 و 52 غير موجودتين في البيانات، وقد قمنا بإضافتهما لاستكمال رسم المضلع وحتى تتساوى مساحة المضلع التكراري مع مساحة المدرج التكراري. فالقيمة 20 هي مركز فئة وهمية]22 18] وتكرارها يساوي 0، و القيمة 52 هي مركز فئة وهمية] 54 50] وتكرارها يساوي 0.

8- التمثيل البياني للتكرار المتجمع الصاعد(More than Cumulative Frequency curve) والنازل(Less than Cumulative Frequency curve):

يرسم منحنى التكرار المتجمع الصاعد عن طريق إيصال مجموعة النقاط التي إحداثياتها الحدود العليا للفئات والتكرار المتجمع الصاعد المقابل لها، ويرسم منحنى التكرار المتجمع النازل بإيصال مجموعة النقاط التي إحداثياتها الحدود الدنيا للفئات والتكرار المتجمع النازل مقابل لها.

التمثيل البياني:



تمثيل بياني للتكرارات المتجمعة الصاعدة والنازلة لتوزيع مجموعة من المحلات التجارية حسب قيمة المبيعات

9- تعديل الفئات:

عندما يكون جدول التوزيع التكراري غير منتظم أي فئاته غير متساوية الطول يجب تعديل التكرارات قبل رسم التمثيل البياني، ويكون ذلك بحساب **التكرار المعدل** (Af_i) كما يلي:

$$Af_i = L^* . \left(\frac{f_i}{L_i}\right)$$

حيث: L^* : الطول الشائع للفئات

f_i: التكرار العادي

L_i: طول الفئة

مثال:

يبين الجدول التكراري التالي توزيع الإنفاق الشهري لعينة من العائلات(الوحدة: 1000دج)

x_i	[10 15[	[15 25[	[25 35[	[35 50[	[50 65[	[65 70[	Σ
f_i	3	10	24	30	18	4	179

المطلوب: مثل التوزيع التكراري بيانيا

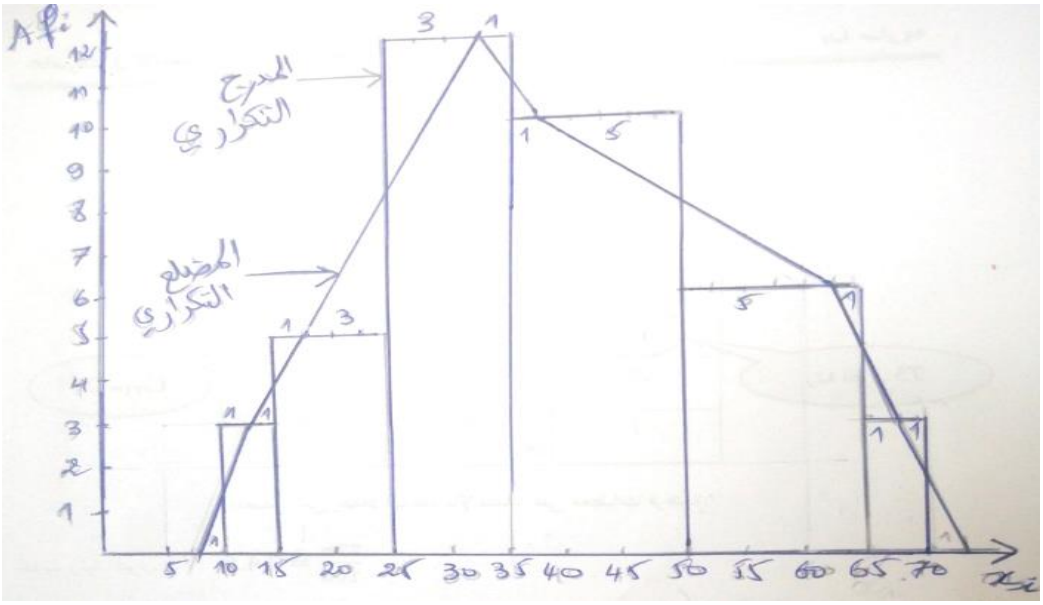
الحل:

بما أن فئات التوزيع التكراري غير منتظمة نقوم بتعديل التكرارات ونأخذ طول الفئة الشائع ($L^*=5$) كأساس لتعديل التكرارات.

x_i	f_i	L_i	$Af_i = L^*.\left(\frac{f_i}{L_i}\right)$
[10 15[	3	5	$5.\frac{3}{5}=3$
[15 25[	10	10	$5.\frac{10}{10}=5$
[25 35[	24	10	$5.\frac{24}{10}=12$
[35 50[	30	15	$5.\frac{30}{15}=10$
[50 65[	18	15	$5.\frac{18}{15}=6$
[65 70[	4	5	$5.\frac{4}{5}=4$
Σ	179	/	/

التمثيل البياني بعد التعديل:

يتم رسم المدرج التكراري بالإعتماد على الفئات في محور الفواصل والتكرار المعدل في محور التراتيب.



تمثيل بياني لتوزيع مجموعة من العائلات حسب الإنفاق الشهري باستخدام المدرج التكراري والمضلع التكراري