

جامعة محمد بوضياف – المسيلة

كلية العلوم الاقتصادية، العلوم التجارية وعلوم التسيير

السنة الأولى جذع مشترك LMD

مقياس: الإحصاء 1 2021/2020

حل سلسلة الإحصاء 1المفاهيم الأساسية

حل التمرين الأول:

1- التعريف بالمصطلحات الإحصائية التالية:

أ- الإحصاء: الإحصاء هو العلم الذي يبحث في طرق جمع البيانات الخاصة بمختلف الظواهر وعرضها وتحليلها للوصول إلى نتائج تساعد في اتخاذ القرارات المناسبة.

ب- المجتمع الإحصائي: هو مجموع الوحدات الإحصائية المراد دراستها والمعرفة بشكل دقيق والتي تشترك فيما بينها في الصفة الأساسية محل اهتمام الباحث، مثل: مجتمع من الطلبة، مجتمع من الأسر، مجتمع من المؤسسات.

ج- الوحدة الإحصائية: هي الكائن الواحد أو الخلية الأساسية التي تجرى عليه الدراسة الإحصائية، أي أن أسئلة الاستمارة تدور حوله، سواء أكان هذا الكائن إنسانا أو حيوانا أو شيئا، مثل: إنسان، بقرة، سيارة،.....إلخ.

د- المتغير الإحصائي: هو العنصر المشترك لكل الوحدات الإحصائية التي تشكل المجتمع الإحصائي، مثل: الطول، السن، مستوى التأهيل العلمي، الإنتاج،.... إلخ.

هـ- العينة: هي جزء من المجتمع الإحصائي، تستخرج بطرق إحصائية معينة حتى تكون ممثلة للمجتمع الإحصائي أحسن تمثيل، ويتم الاعتماد عليها في الدراسة بدل المجتمع نظرا لكبر حجم المجتمع، ربعا للوقت والجهد والمال، الفحص قد يكون مؤذيا أو متلفا للوحدات.

2- فروع علم الإحصاء وخصوصية كل فرع: ينقسم علم الإحصاء إلى:

أ- الإحصاء الوصفي: هو ذلك الجزء من الإحصاء الذي يهتم بتلخيص البيانات الإحصائية إلى عدد محدود من الأرقام تسمى مقاييس إحصائية أو في جدول إحصائي يسهل القراءة أو في رسوم بيانية، والغرض من كل ذلك هو إعطاء وصف أولي للظاهرة المدروسة بدون تحليل معمق.

ب- الإحصاء الاستدلالي: يستند على فكرة اختيار جزء من المجتمع يسمى العينة، وذلك بطريقة علمية مناسبة، بغرض استخدام بيانات هذه العينة في التوصل إلى نتائج يمكن تعميمها على مجتمع الدراسة، فنقول لقد استدللنا على خواص المجتمع على أساس خواص العينة.

3- أنواع المتغيرات الإحصائية: تنقسم المتغيرات الإحصائية إلى قسمين:

أ- متغيرات كمية: هي تلك المتغيرات التي لا يمكن قياسها، والتي تنقسم بدورها إلى قسمين:

1-1- متغيرات كمية قابلة للترتيب: مثل مستوى التأهيل العلمي،.... إلخ.

2-2- متغيرات كمية غير قابلة للترتيب: مثل الجنسية، الجنس، الحالة العائلية، اللون،.....إلخ.

ب- متغيرات كمية: هي تلك المتغيرات التي يمكن قياسها، وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام، والمتغيرات الكمية تنقسم بدورها إلى قسمين:

ب-1- متغيرات كمية منقطعة: هي تلك المتغيرات التي تأخذ قيما صحيحة لا يمكن تجزئتها، مثل عدد الأطفال في الأسرة الواحدة، عدد قطع الغيار المنتجة...إلخ.

ب-2- متغيرات كمية مستمرة: هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المنتهي لهذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات، مثال الطول، السن، الوزن،...إلخ.

حل التمرين الثاني:

العبرة	المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه
01	العدد الإجمالي للجالية المغربية بفرنسا	الفرد الواحد	الجنسية	كيفي غير قابل للترتيب
02	جميع السكنات في بلدية سطيف	المسكن الواحد	عدد الغرف	كمي متقطع
03	جميع سكان البلدية	الفرد الواحد	المستوى التعليمي	كيفي قابل للترتيب
04	65 عاملا في شركة متوسطة	العامل الواحد	الأجر الشهري	كمي مستمر

## التوزيعات التكرارية وتمثيلاتها البيانية - مقياس النزعة المركزية

حل التمرين الثالث:

1- عرض هذه البيانات في جدول تكراري، ثم حساب التكرارات المطلقة والنسبة المئوية الصاعدة والنازلة:

نقوم بترتيب البيانات كما يلي:

3	2	2	2	2	2	2	2	2	1
3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	3
5	5	5	5	5	4	4	4	4	4
8	8	7	7	7	6	6	6	6	5

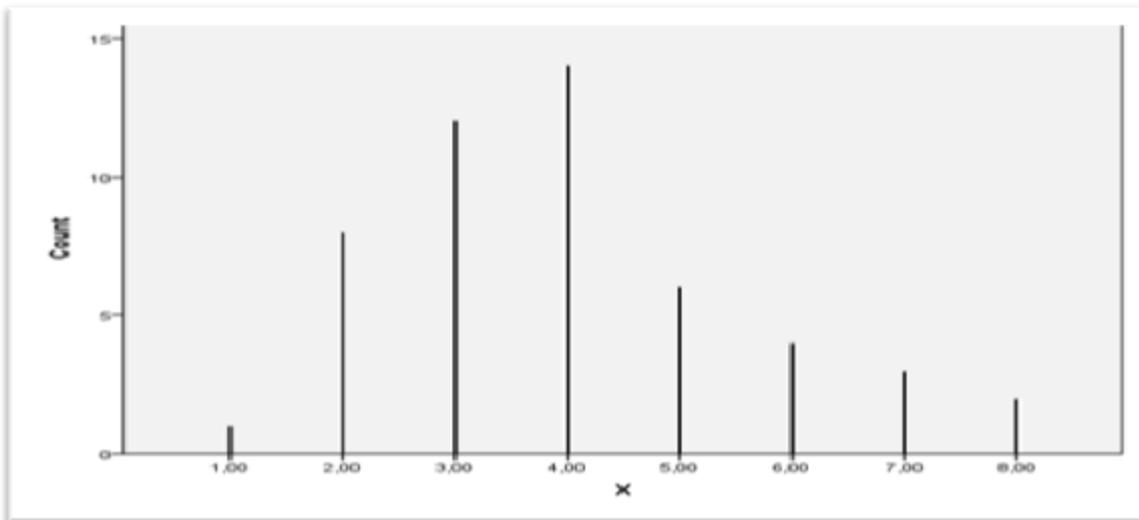
ثم نعرض البيانات في الجدول التكراري كما يلي:

$n_i \times X_i$	$F_i^{\downarrow}\%$	$F_i^{\uparrow}\%$	$f_i\%$	$N_i^{\downarrow}$	$N_i^{\uparrow}$	عدد السكنات $n_i$	عدد الغرف $X_i$
1	100	2	2	50	1	1	1
16	98	18	16	49	9	8	2
36	82	42	24	41	21	12	3
56	58	70	28	29	35	14	4
30	30	82	12	15	41	6	5
24	18	90	8	9	45	4	6
21	10	96	6	5	48	3	7
16	4	100	4	2	50	2	8
200	/	/	100		/	50	المجموع

2- التمثيل البياني:

بما أن المتغير الاحصائي عدد الغرف، متغير كمي منفصل، فإنه يمثل بواسطة الأعمدة، كما يلي:

الشكل 1: تمثيل بياني بواسطة الأعمدة لتوزيع عدد السكنات حسب عدد الغرف



3- حساب العدد المتوسط للغرف في المسكن الواحد:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{\sum_{i=1}^{50} X_i}{50} = \frac{4+2+4+6+\dots+3+3+4+2}{50} = \frac{200}{50} = 4 \quad \text{أ- على السلسلة:}$$

$$\bar{X} = \frac{\sum_{i=1}^k n_i X_i}{\sum n_i} = \frac{\sum_{i=1}^8 n_i X_i}{50} = \frac{200}{50} = 4 \quad \text{ب- على الجدول:}$$

4- حساب كلا من: المنوال، الوسيط، المتوسط الهندسي، المتوسط التوافقي والمتوسط التربيعي على الجدول فقط

أ- حساب المنوال: يعرف المنوال على أنه القيمة الأكثر تكراراً في التوزيع وبالتالي فإن المنوال يساوي 4 أي:  $M_o = 4$

- الشرح: السكنات التي تحتوي على 4 غرف هي الأكثر تكراراً في هذه الدراسة.

ب- حساب الوسيط: بما أن عدد التكرارات هو 50 أي أنه عدد زوجي، فإن الوسيط هنا هو متوسط القيمة التي رتبها  $\frac{n}{2}$  والقيمة التي رتبها  $\frac{n}{2} + 1$  أي:

$$M_e = \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} = \frac{X_{(25)} + X_{(26)}}{2} = \frac{4+4}{2} = 4$$

- الشرح: 50% من السكنات عدد الغرف فيها أقل أو يساوي 4 غرف و50% من السكنات المتبقية عدد الغرف فيها أكبر أو يساوي 4 غرف.

$$\bar{X}_G = 10^{\left(\frac{1}{n} \sum_{i=1}^8 n_i \log(x_i)\right)} = 10^{\left(\frac{1}{50} 28,2\right)} = 10^{(0,564)} = 3,66 \quad \text{ج- حساب المتوسط الهندسي:}$$

$$\bar{X}_H = \frac{\sum_{i=1}^8 n_i}{\sum_{i=1}^8 \frac{n_i}{x_i}} = \frac{50}{15,05} = 3,32 \quad \text{د- حساب المتوسط التوافقي:}$$

$$\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^8 n_i x_i^2}{\sum_{i=1}^8 n_i}} = \sqrt{\frac{934}{50}} = 4,32 \quad \text{هـ- حساب المتوسط التربيعي:}$$

$n_i(X_i - 5)$	$n_i x_i^2$	$\frac{n_i}{x_i}$	$n_i \log(x_i)$	عدد السكنات $n_i$	عدد الغرف $X_i$
-4	1	1	0	1	1
-24	32	4	2,41	8	2
-24	108	4	5,72	12	3
-14	224	3,5	8,43	14	4
0	150	1,2	4,19	6	5
4	144	0,67	3,11	4	6
6	147	0,43	2,53	3	7
6	128	0,25	1,81	2	8
-50	934	15,05	28,2	50	المجموع

4- إيجاد العدد المتوسط الحقيقي للغرف إذا علمنا أن المتوسط الفرضي هو  $a = 5$ :

$$\bar{d} = \frac{\sum_{i=1}^8 n_i d_i}{n} = \frac{\sum_{i=1}^8 n_i d_i}{72} = \frac{-50}{50} = -1$$

$$\bar{X} = \bar{d} + a = -1 + 5 = 4 \quad \text{حساب المتوسط الحسابي الأصلي:}$$

إذن المتوسط الحقيقي لهذه البيانات هو: 4 غرف.

## حل التمرين الرابع:

1- عرض هذه البيانات في توزيع تكراري على شكل فئات باستخدام طريقة ستورجس *Sturges*:

أ- حساب المدى العام  $E$ : وهو الفرق بين أكبر قيمة وأقل قيمة، أي:  $E = X_{max} - X_{min} = 173 - 114 = 59$

ب- تحديد عدد الفئات  $K$ : فئات  $K = 1 + 3,322 \log(n) = 1 + 3,322 \log(32) = 6$

حيث:  $n$ : تمثل حجم العينة.

ج- تحديد أطوال الفئات  $C$ :  $C = \frac{E}{K} = \frac{59}{6} = 9,83 \approx 10$

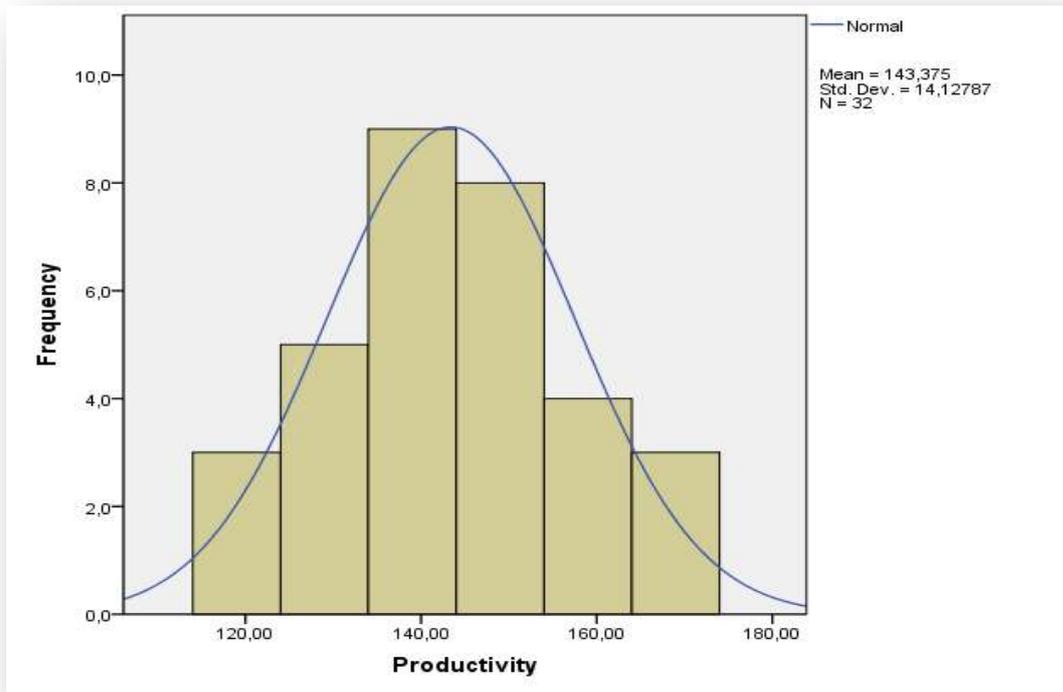
أي طول كل فئة يساوي 10 ساعات، ومنه تكون الفئات هي:  $[114 - 124]$ ،  $[124 - 134]$ ، .....،  $[164 - 174]$ .

نقوم بتفريغ البيانات في الفئات المذكورة، فنحصل على الجدول الموالي:

$N_i^{\downarrow}$	$N_i^{\uparrow}$	$n_i \times C_i$	$C_i$	عدد العمال $n_i$	متوسط إنتاجية العمال $X_i$
32	3	357	119	3	$[124 - 114]$
29	8	645	129	5	$[134 - 124]$
24	17	1251	139	9	$[144 - 134]$
15	25	1192	149	8	$[154 - 144]$
7	29	636	159	4	$[164 - 154]$
3	32	507	169	3	$[174 - 164]$
/	/	4588	/	$\sum n_i = 32$	المجموع

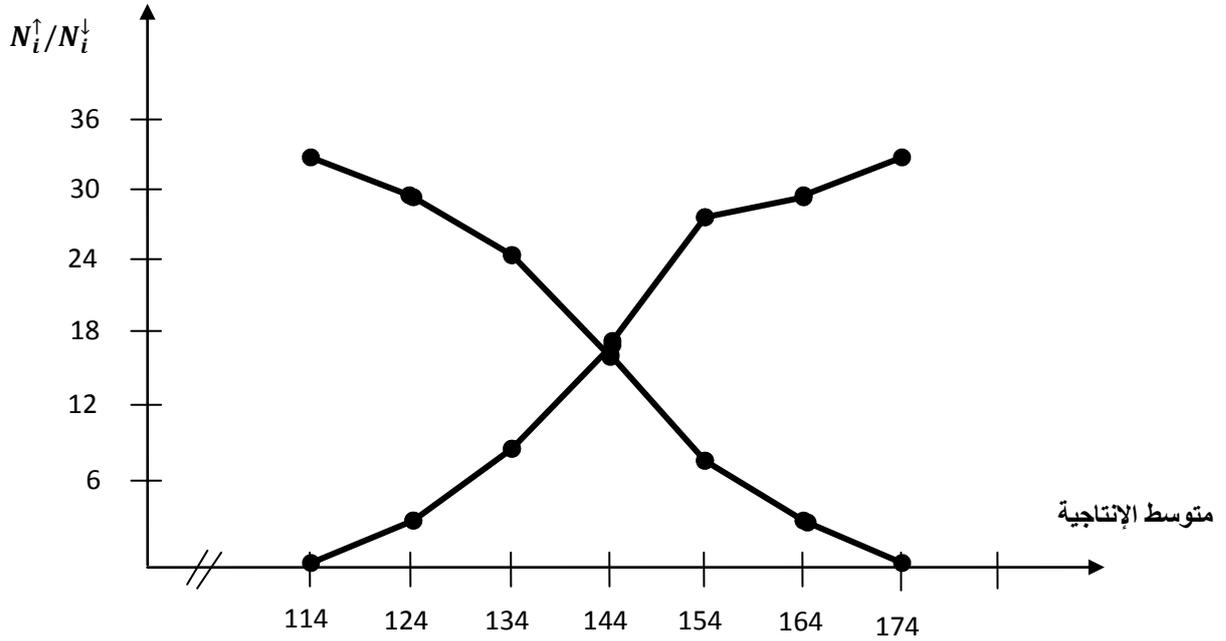
2- التمثيل البياني: بواسطة المدرج التكراري، ويمكن رسم المنحنى التكراري والمضلع التكراري على نفس الشكل.

الشكل 2: تمثيل بياني بواسطة المدرج التكراري والمنحنى التكراري لتوزيع العمال حسب متوسط إنتاجيتهم الشهرية



## 3- حساب التكرارات المطلقة الصاعدة والنازلة، وتمثيلها بيانياً:

الشكل 3: تمثيل بياني بواسطة المنحنى التجميعي الصاعد والنازل لتوزيع العمال حسب متوسط إنتاجيتهم الشهرية



## 4- حساب كلا من المتوسط الحسابي، الوسيط والمنوال:

$$\bar{X} = \frac{\sum_{i=1}^6 n_i \times C_i}{\sum_{i=1}^6 n_i} = \frac{4588}{32} = 143,375 h \quad \text{أ- المتوسط الحسابي:}$$

ب- المنوال:

- الفئة المنوالية، التي تقابل أكبر تكرار هي:  $[144 - 134]$

وبالتالي فإن:  $\Delta_1 = 9 - 5 = 4$  ،  $\Delta_2 = 9 - 8 = 1$

$$M_o = Lim_{M_o} + \left[ \frac{\Delta_1}{\Delta_1 + \Delta_2} \right] \times A_{M_o} \quad \text{ومنه:}$$

$$M_o = 134 + \left[ \frac{4}{4+1} \right] \times 10 = 142 h$$

ج- الوسيط:

- تحديد الفئة الوسيطة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي  $\frac{n}{2}$  ، أي:  $N_{M_e}^1 \geq \left( \frac{n}{2} = 16 \right)$

ومنه الفئة الوسيطة هي:  $[144 - 134]$

- حساب الوسيط بطريقة المد الداخلي:

$$M_e = Lim_{M_e} + \left[ \frac{\frac{n}{2} - N_{M_e}^1 - 1}{n_{M_e}} \right] \times A_{M_e} = 134 + \left[ \frac{16-8}{9} \right] \times 10 = 142,83 h$$

## مقاييس التشتت والشكل

حل التمرين الخامس:

1- حساب كلا من المدى، الانحراف المتوسط، التباين والانحراف المعياري على بيانات المجموعتين:

أ- المجموعة A:

$$R = \text{Max}(X_i) - \text{Min}(X_i) = 13 - 2 = 11 \quad \text{طن} \quad \text{المدى } R:$$

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{\sum_{i=1}^7 X_i}{7} = \frac{12+6+13+5+11+7+2}{7} = \frac{56}{7} = 8 \quad \text{طن} \quad \text{الانحراف المتوسط } MD:$$

$$MD_A = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n} = \frac{\sum_{i=1}^7 |X_i - \bar{X}|}{7} = \frac{|12-8|+|6-8|+\dots+|2-8|}{7} = \frac{4+2+\dots+6}{7} = \frac{24}{7} = 3,43 \quad \text{طن}$$

$$V(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^7 (X_i - \bar{X})^2}{7} = \frac{(12-8)^2+(6-8)^2+\dots+(2-8)^2}{7} = \frac{16+4+\dots+36}{7} = \frac{100}{7} = 14,28 \quad \text{التباين } V(X):$$

$$\sigma(X) = \sqrt{V(X)} = \sqrt{14,28} = 3,78 \quad \text{طن} \quad \text{الانحراف المعياري } \sigma(X):$$

ب- المجموعة B:

$$R = \text{Max}(Y_i) - \text{Min}(Y_i) = 11 - 2 = 9 \quad \text{طن} \quad \text{المدى } R:$$

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} = \frac{\sum_{i=1}^8 Y_i}{8} = \frac{11+2+\dots+7}{8} = \frac{52}{8} = 6,5 \quad \text{طن} \quad \text{الانحراف المتوسط } MD:$$

$$MD_B = \frac{\sum_{i=1}^n |Y_i - \bar{Y}|}{n} = \frac{\sum_{i=1}^8 |Y_i - \bar{Y}|}{8} = \frac{|11-6,5|+|2-6,5|+\dots+|7-6,5|}{8} = \frac{4,5+4,5+\dots+0,5}{8} = \frac{18}{8} = 2,25 \quad \text{طن}$$

$$V(X) = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n} = \frac{\sum_{i=1}^8 (Y_i - \bar{Y})^2}{8} = \frac{(11-6,5)^2+\dots+(7-6,5)^2}{8} = \frac{20,25+\dots+0,25}{8} = \frac{58}{8} = 7,25 \quad \text{التباين } V(X):$$

$$\sigma(X) = \sqrt{V(X)} = \sqrt{7,25} = 2,69 \quad \text{طن} \quad \text{الانحراف المعياري } \sigma(X):$$

2- المقارنة بين مستوى وتشتت مبيعات الشركتين:

أ- المقارنة بين مستوى: بما أن  $\bar{X} > \bar{Y}$ ، فإننا نقول أنه في المتوسط، المبيعات اليومية للمجموعة A أكبر من المبيعات اليومية للمجموعة B.

ب- مقارنة التشتت: بما أن  $\bar{X} \neq \bar{Y}$ ، فإننا نستخدم معامل الاختلاف CV للمقارنة بين المجموعتين من حيث التشتت.

$$CV_A = \frac{\sigma(X)}{\bar{X}} \times 100 = \frac{3,78}{8} \times 100 = 47,25\%$$

$$CV_B = \frac{\sigma(Y)}{\bar{Y}} \times 100 = \frac{2,69}{6,5} \times 100 = 41,38\%$$

بما أن  $CV_A > CV_B$ ، فإننا نقول أن المبيعات اليومية للمجموعة A أكثر تشتتاً من المبيعات اليومية للمجموعة B. أي أن مبيعات المجموعة B أكثر تجانساً.

3- تحديد شكل التوزيع للمجموعة A من ناحية الالتواء، باستخدام معامل بيرسون:

$$SK = \frac{3(\bar{X} - M_e)}{\sigma(X)}$$

- حساب قيمة الوسيط:

- ترتيب القيم: 2، 5، 6، 7، 11، 12، 13

بما أن عدد القيم هو 7 أي أنه عدد فردي، فإن الوسيط هنا هو متوسط القيمة التي رتبها  $\frac{n+1}{2}$ ، أي:

$$M_e = X_{\left(\frac{n+1}{2}\right)} = X_{\left(\frac{7+1}{2}\right)} = X_{(4)} = 7 \quad \text{طن}$$

$$SK = \frac{3(8-7)}{3,78} = 0,79$$

بما أن  $SK > 0$ : فإن التوزيع ملتوي نحو اليمين وإشارته موجبة.

**4- تحديد شكل التوزيع من ناحية التفرطح لمبيعات المجموعة A:**

$$\mu_4 = \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{n} \quad \text{حيث:} \quad \beta_F = \frac{\mu_4}{[\delta(X)]^4} - 3$$

$$\mu_4 = \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{n} = \frac{\sum_{i=1}^7 (X_i - \bar{X})^4}{7} = \frac{(12-8)^4 + (6-8)^4 + \dots + (2-8)^4}{7} = \frac{256+16+\dots+1296}{7} = \frac{2356}{7} = 336,57$$

$$\beta_F = \frac{336,57}{[3,78]^4} - 3 = \frac{336,57}{204,16} - 3 = -1,35$$

بما أن  $\beta_F < 0$ : فإن التوزيع مفرطح.

حل التمرين السادس:

$n_i(C_i - \bar{X})^4$	$n_i(C_i - \bar{X})^3$	$n_i(C_i - \bar{X})^2$	$n_i C_i - \bar{X} $	$N_i^\uparrow$	$n_i \times C_i$	$C_i$	عدد المبحوثين $n_i$	الفئات $X_i$
54700,816	-6360,56	739,6	86	10	180	18	10	[20 - 16]
6716,184	-1460,04	317,4	69	25	330	22	15	[24 - 20]
5,184	-8,64	14,4	24	65	1040	26	40	[28 - 24]
2672,672	786,08	231,2	68	85	600	30	20	[32 - 28]
44979,864	6078,36	821,4	111	100	510	34	15	[36 - 32]
109074,72	-964,8	2124	358	/	2660	/	$\sum n_i = 100$	المجموع

**1- حساب كلا من: المتوسط الحسابي، الوسيط والمنوال، والمقارنة بينها لتحديد شكل التوزيع:**

$$\bar{X} = \frac{\sum_{i=1}^5 n_i \times C_i}{\sum_{i=1}^5 n_i} = \frac{2660}{100} = 26,6 \quad \text{أ- المتوسط الحسابي:}$$

**ب- المنوال:**

الفئة المنوالية، التي تقابل أكبر تكرار هي: [28 - 24]

$$\Delta_1 = 40 - 15 = 25 \quad , \quad \Delta_2 = 40 - 20 = 20 \quad \text{وبالتالي فإن:}$$

$$M_o = \text{Lim}_{M_o} + \left[ \frac{\Delta_1}{\Delta_1 + \Delta_2} \right] \times A_{M_o} \quad \text{ومنه:}$$

$$M_o = 24 + \left[ \frac{25}{25+20} \right] \times 4 = 26,22$$

**ج- الوسيط:**

تحديد الفئة الوسيطة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي  $\frac{n}{2}$ ، أي:  $N_{M_e}^\uparrow \geq \left( \frac{n}{2} = 50 \right)$

ومنه الفئة الوسيطة هي: [28 - 24]

حساب الوسيط بطريقة المد الداخلي:

$$M_e = \text{Lim}_{M_e} + \left[ \frac{\frac{n}{2} - N_{M_e-1}^\uparrow}{n_{M_e}} \right] \times A_{M_e} = 24 + \left[ \frac{50-25}{40} \right] \times 4 = 26,5$$

بما أن:  $M_o < M_e < \bar{X}$  فإن التوزيع ملتوي نحو اليمين، أي موجب الالتواء.

2- حساب كلا من: الانحراف المتوسط، التباين والانحراف المعياري ومعامل الاختلاف:

$$\bar{X} = 26,6 \quad \text{- الانحراف المتوسط } MD:$$

$$MD = \frac{\sum_{i=1}^k n_i |C_i - \bar{X}|}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^5 n_i |C_i - \bar{X}|}{100} = \frac{358}{100} = 3,58$$

$$V(X) = \frac{\sum_{i=1}^k n_i (C_i - \bar{X})^2}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^5 n_i (C_i - \bar{X})^2}{100} = \frac{2124}{100} = 21,24 \quad \text{- التباين:}$$

$$\sigma(X) = \sqrt{V(X)} = \sqrt{21,24} = 4,61 \quad \text{- الانحراف المعياري } \sigma(X):$$

$$CV = \frac{\sigma(X)}{\bar{X}} \times 100 = \frac{4,61}{26,6} \times 100 = 17,33\% \quad \text{- معامل الاختلاف } CV:$$

3- تحديد شكل التوزيع من ناحية الالتواء وإشارته باستخدام معامل فيشر للالتواء:

$$\mu_3 = \frac{\sum_{i=1}^k n_i (C_i - \bar{X})^3}{\sum_{i=1}^k n_i} \quad \text{حيث:} \quad \alpha_F = \frac{\mu_3}{[\sigma(X)]^3}$$

$$\mu_3 = \frac{\sum_{i=1}^k n_i (C_i - \bar{X})^3}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^5 n_i (C_i - \bar{X})^3}{100} = \frac{-964,8}{100} = -9,648$$

$$\alpha_F = \frac{-9,648}{[4,61]^3} = -0,098$$

بما أن  $\alpha_F < 0$ : فإن التوزيع ملتوي نحو اليسار، أي سالب الالتواء.

4- تحديد شكل التوزيع من ناحية التفرطح:

$$\mu_4 = \frac{\sum_{i=1}^k n_i (C_i - \bar{X})^4}{\sum_{i=1}^k n_i} \quad \text{حيث:} \quad \beta_F = \frac{\mu_4}{[\sigma(X)]^4} - 3$$

$$\mu_4 = \frac{\sum_{i=1}^k n_i (C_i - \bar{X})^4}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^5 n_i (C_i - \bar{X})^4}{100} = \frac{109074,72}{100} = 1090,7472$$

$$\beta_F = \frac{1090,7472}{[4,61]^4} - 3 = 2,41 - 3 = -0,59$$

بما أن  $\beta_F < 0$ : فإن التوزيع مفرطح.