الفصل الثاني

عرض و تمثيل البيانات الإحصائية

من الأفضل رسم السلسلة الإحصائية. حيث يوفر الرسم البياني نظرة عامة ، و ملخص لجميع البيانات المقاسة. كل هذا يكون أسهل إذا استخدمنا برامج الكمبيوتر R ، SPSS ، Exel ، Matlab ...

1.2 التمثيل البياني

الهدف من هذه العروض هو الحصول على ملخصات واضحة وموجزة ، مع الاحتفاظ بمعظم المعلومات الموجودة في البيانات الأولية ، واستخدام التقنيات الموضوعية التي تعطينا صورة مطابقة للواقع.

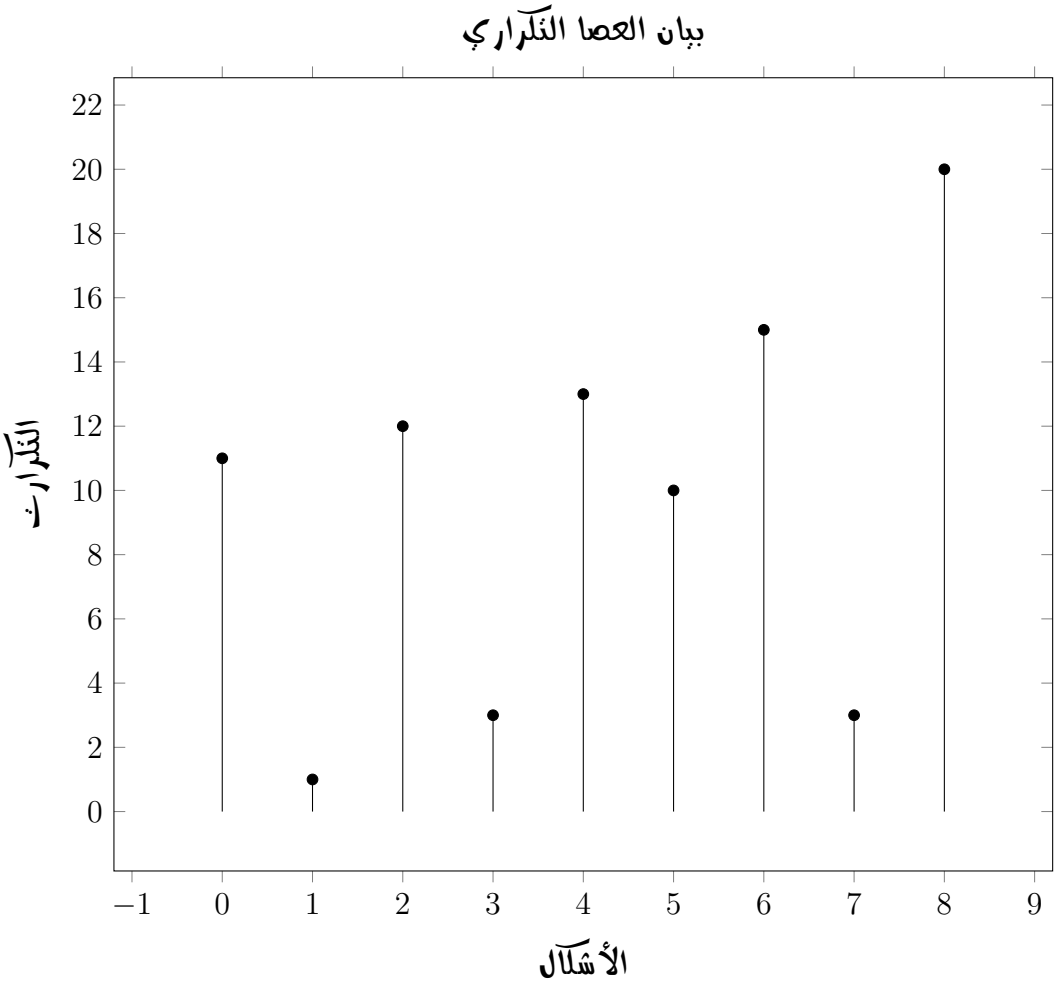
1.1.2. بيان العصا Diagramme en bâtons

يمثل الرسم بياني العصا توزيع المتغير الكمي المنفصل من قبل قطاعات الرأسية: يتم وضع القيم على محور الفواصل، والأرقام (أو التواترات) على محور التراتيب.

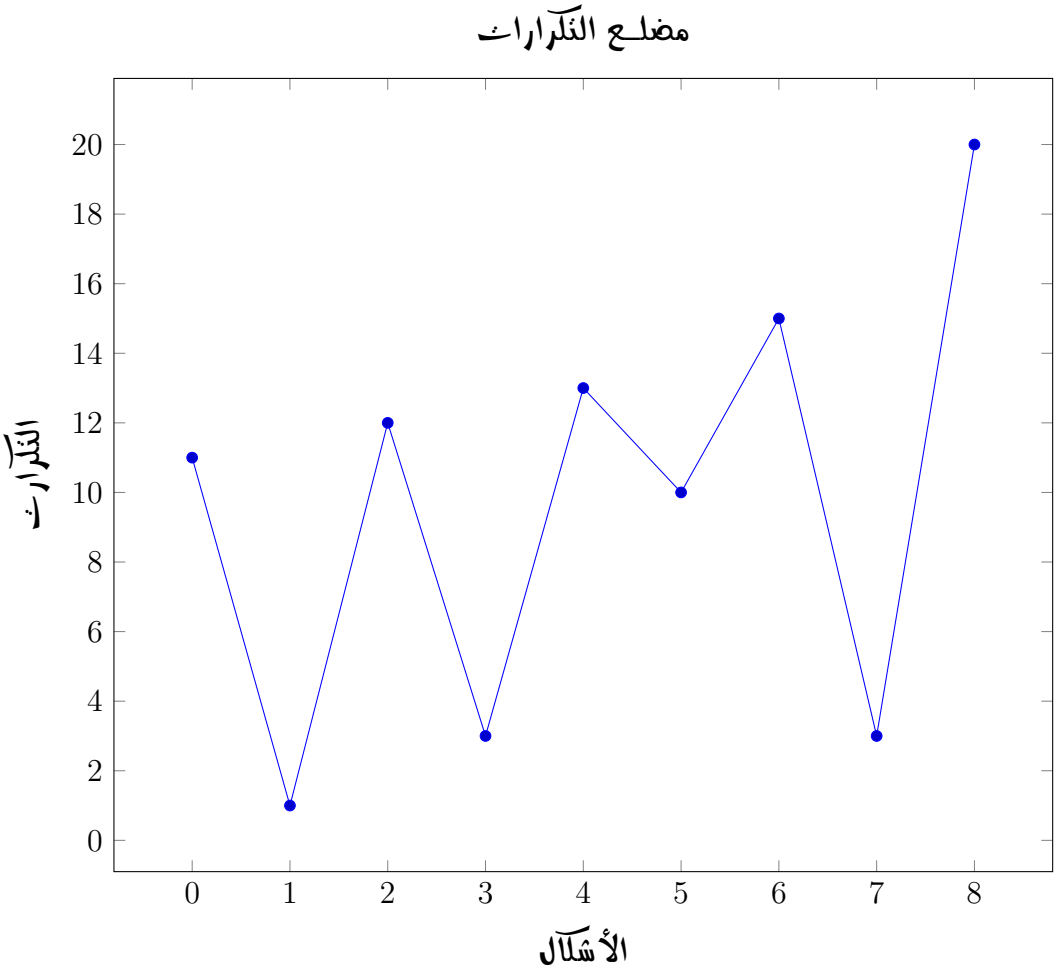
مثال 1 : لنكن S السلسلة الإحصائية التي يكون جدول توزيعها على النحو التالي:

الأشكال	0	1	2	3	4	5	6	7	8
التكرارت	11	1	12	3	13	10	15	3	20

رسم بياني العصا لسلسلة S يظهر أدناه.



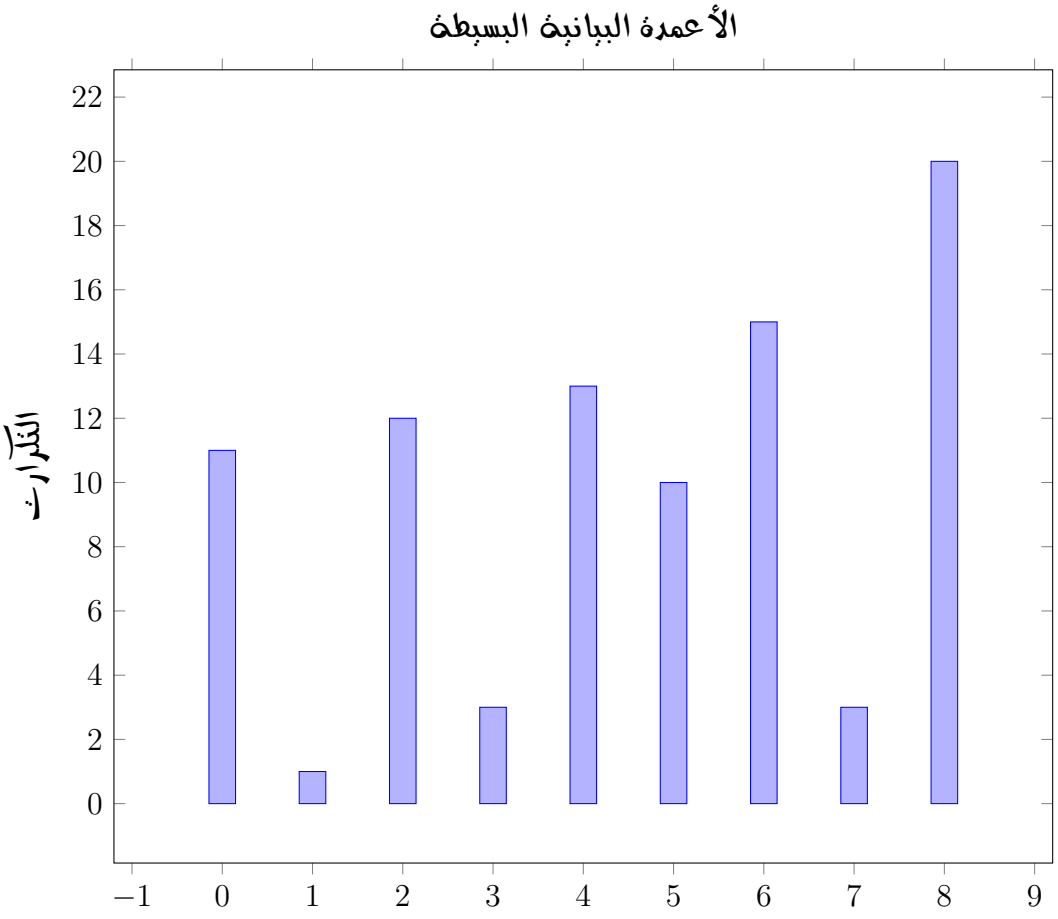
يطلق على المنحنى الذي يصل بين رؤوس العصي مضلع التردد المطلق *polygone des fréquences absolues* .



2.1.2. الأعمدة البيانية البسيطة Diagramme en barres

الأعمدة البيانية البسيطة (المستطيلات) متطابقة وفقاً لعُرضها الأساسي ، حيث يتركز كل من المستطيلات على النقاط الممثلة على محور الفواصل.

مثال 2 : يمثل الرسم البياني التالي بيانات المثال *1:* بواسطة مخطط الأعمدة البيانية البسيطة.



3.1.2. المدرج التكراري Histogramme

المدرج التكراري هو التمثيل البياني الذي يوضح توزيع السلسلة ، ولكن قبل المضي قدما ، من الضروري إعطاء تعريف لسعة فئة l'amplitude C_i.

تعريف 1.1.2 : نسمي سعة *Amplitude* للفئة $C_i =]a_i, a_{i+1}]$ العدد :

$$h_i = a_{i+1} - a_i.$$

ولإنشاء المدرج التكراري نتبع الخطوات التالية :

1. نضع على محورالفواصل النهايات الحدية المختلفة a_i للفئات C_i.

2. نقوم برسم ، لكل فئة، مستطيل موازٍ لمحور التراتيب، بحيث يكون للجزء الموازي لمحور الفواصل طولاً مطابقا لسعة الفئة وأن مساحة المستطيل تتناسب مع حجم الفئة، وهذا لتمثيل أهمية كل فئة.

3. نضع قيمة التكرارات n_i أعلى المستطيل الذي يتوافق مع الفئة C_i.

ملاحظة 1 : .

1. ارتفاع كل مستطيل يتوافق مع كثافة (أو كثافة التردد) للفئة C_i، $i = 1, 2, ..., k$ يتم حسابها بواسطة

$$d_i = \frac{f_i}{h_i},$$

حيث f_i هو التردد الجزئي *la fréquence partielle* و h_i سعة الفئة C_i. في هذه الحالة فإن مساحة سطح هذا المستطيل S_i هي

$$S_i = h_i H_i,$$

حيث H_i هي ارتفاع المستطيل.

2. يجب أن تكون الأسطح k متناسبة مع التراكيز n_i، أو يجب أن تكون ارتفاعات H_i متناسبة مع الكثافة d_i، وهذا يعني أن هناك ثابت c بحيث من أجل كل i.

$$S_i = cn_i.$$

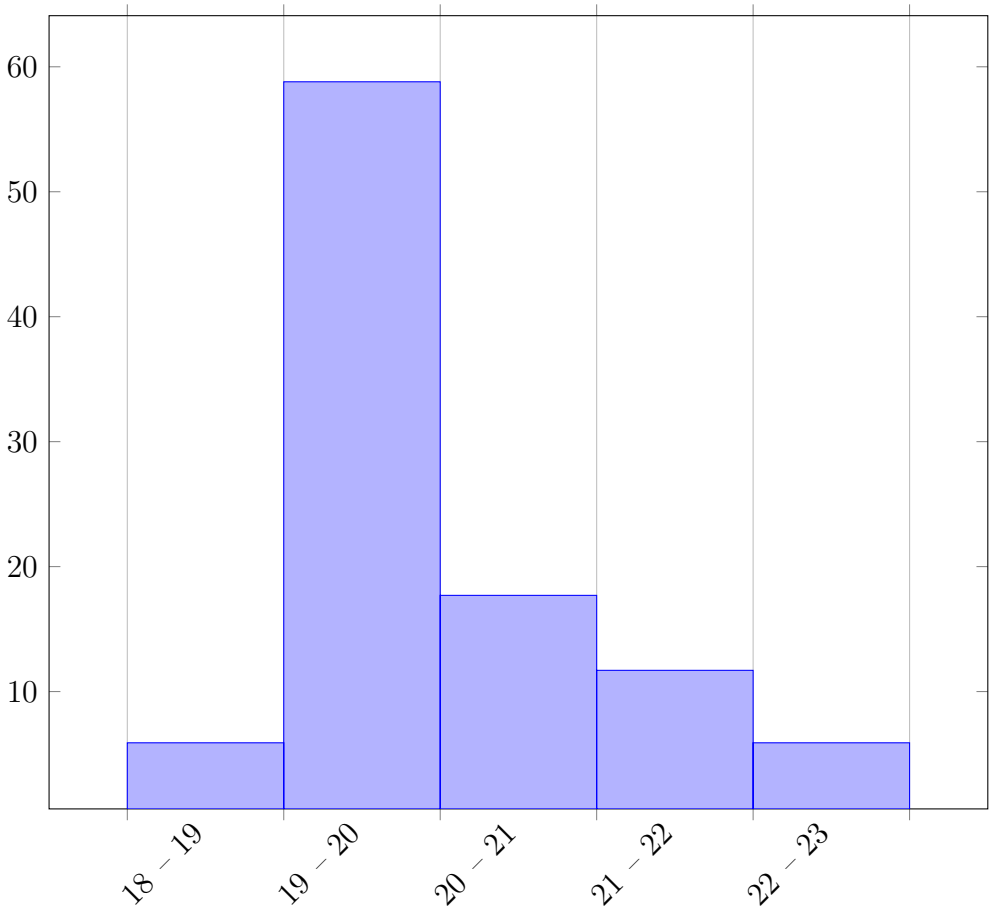
هذا يكافئ

$$h_i H_i = cnf_i H_i = cnd_i$$

مثال 3 : والجدول البيانات التالية حيث السطر الثالث يلخص الكثافة:

العمر C_i	$[18-19]$	$]19-20]$	$]20-21]$	$]21-22]$	$]22-23]$
التكرارات n_i	1	10	3	2	1
السعة h_i	1	1	1	1	1
الكثافة $d_i\%$	5.9	58.8	17.7	11.7	5.9

الرسم البياني لسلسلة البيانات هذه هو:



ملاحظة 2 : هنا، جميع السعات h_i مساوية لـ 1 في هذه الحالة ، يتناسب ارتفاع كل مستطيل مع حجم الفئة المقابلة له.

تمرين 1 : الجدول التالي يمثل توزيع القوى العاملة في دولة ما حسب السن (بالآلاف)

فئات الأعمار	15/10	20/15	25/20	30/25	40/30	50/40	60/50	65/60	70/65
عدد العمال	10	70	105	90	130	80	50	15	15

- أوجد التكرار، المعدل (المصحح) لهذا الجدول.
- مثل بيانيا المدرج التكراري في هذه الحالة.

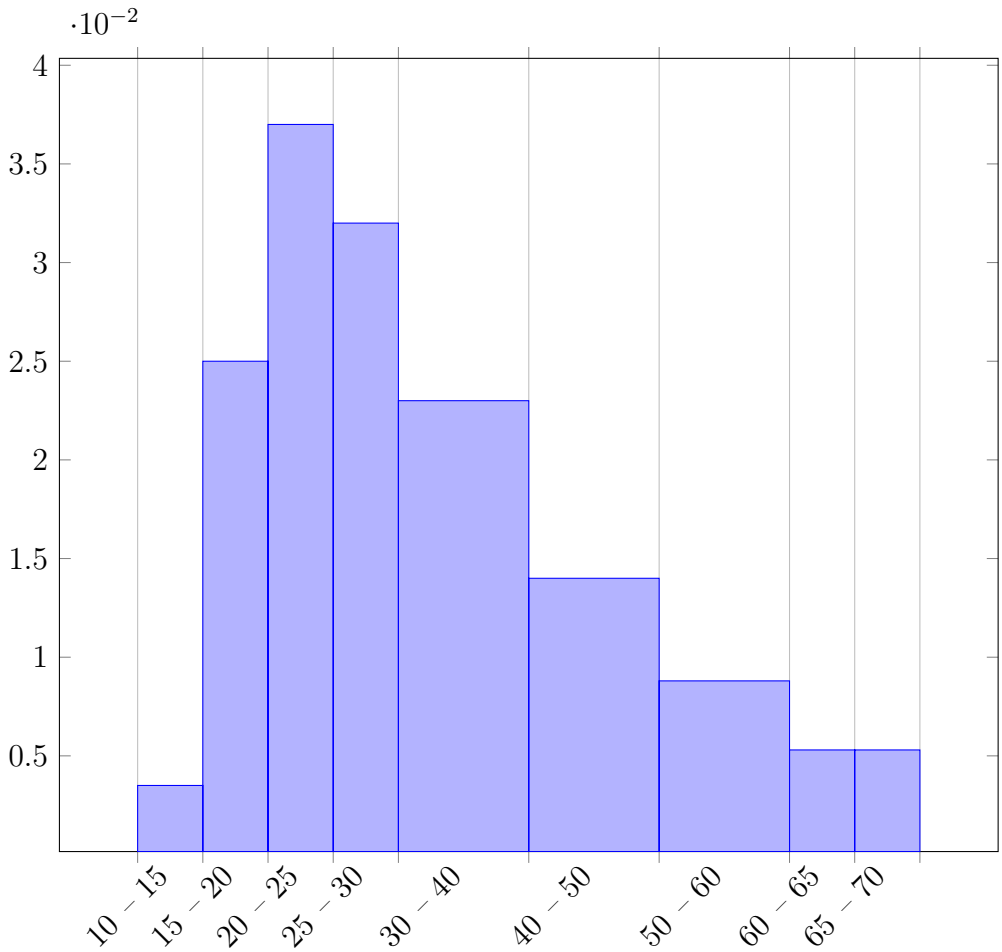
الحـــل

يمكن مالحظة من البداية بأن أطوال الفئات غير متساوية، وهذا ما يدفعنا لتصحيح التكرارات

الأصلية من أجل رسم المدرج التكراري أو حساب المنوال.

فئات الأعمار	15/10	20/15	25/20	30/25	40/30	50/40	60/50	65/60	70/65
التكرار الأصلي	10	70	105	90	130	80	50	15	15
طول الفئة h_i	5	5	5	5	10	10	10	5	5
التكرار المصحح	2	14	21	18	13	8	5	3	3
الكثافة $\frac{f_i}{h_i}$	0.0035	0.025	0.037	0.032	0.023	0.014	0.0088	0.0053	0.0053

يمكن ملاحظة في هذه الحالة أن أكبر تكرار كان (قبل التصحيح) 130 وهو يوافق الفئة $30 - 40$ بينما أصبح (بعد التصحيح) 21 وهو يوافق الفئة $20 - 25$.

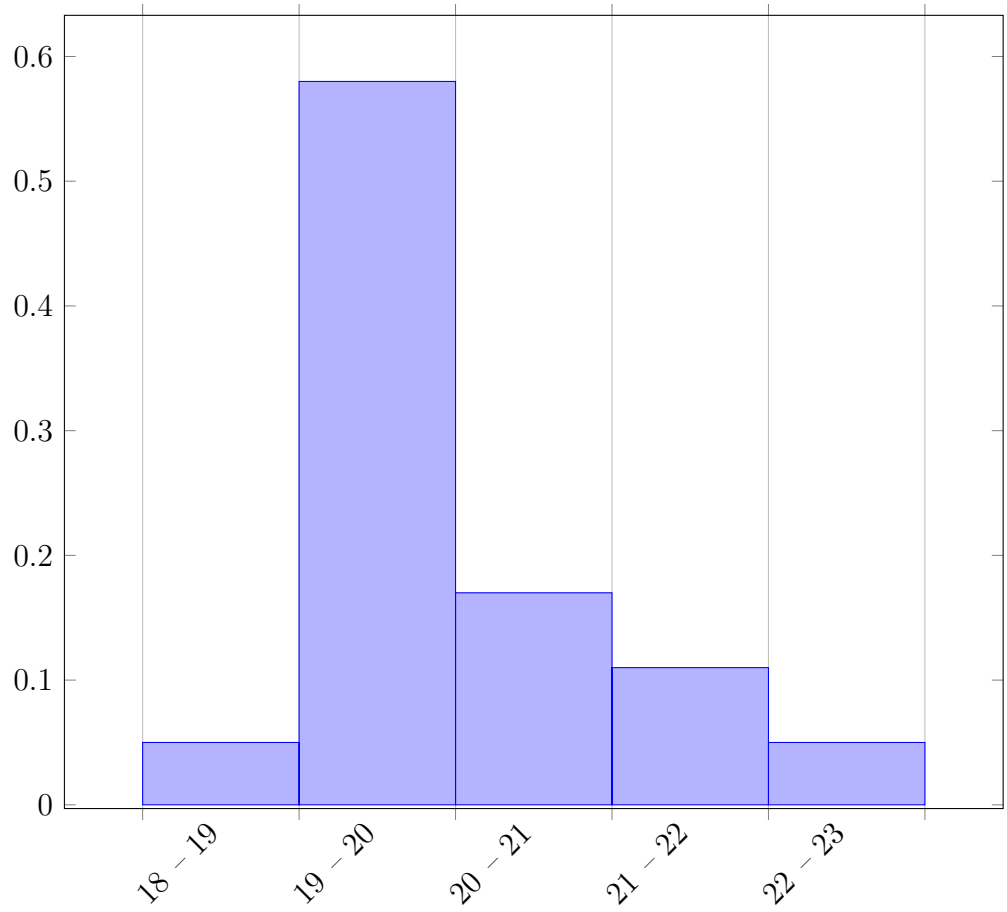


4.1.2. المخطط الجزئي Diagramme Partiel

يتم إنشاء المخطط الجزئي Diagramme Partiel **بنفس طريقة الرسم البياني، باستثناء أن التكرارات** n_i **يتم تغييرها بواسطة الترددات الجزئية** f_i **للفئة** C_i **، والغرض منها هو تمثيل جدول التواترات الجزئية.**

مثال 4 : **الرسم البياني الجزئي لسلسلة البيانات في المثال** *3:* **السابق**

C_i العمر	$[18-19]$	$]19-20]$	$]20-21]$	$]21-22]$	$]22-23]$
f_i التواترات الجزئية	0.05	0.58	0.17	0.11	0.05



5.1.2. المخطط المتكامل أو دالة التوزيع

تعريف 2.1.2 : نسمي دالة التوزيع للمتغير X *Fonction de répartition* الذي يرمز لها بالرمز F ، هو التطبيق من مجموعة الأعداد الحقيقية $\mathbb{R}$ نحو المجال $[0,1]$ ، والذي يرفق بكل قيمة معينة x من $\mathbb{R}$ ، نسبة الأفراد المنتمين إلى المجموعة الإحصائية التي تقل قيمة المتغير فيها عن قيمة x :

$$F(x) = P(X \leq x).$$

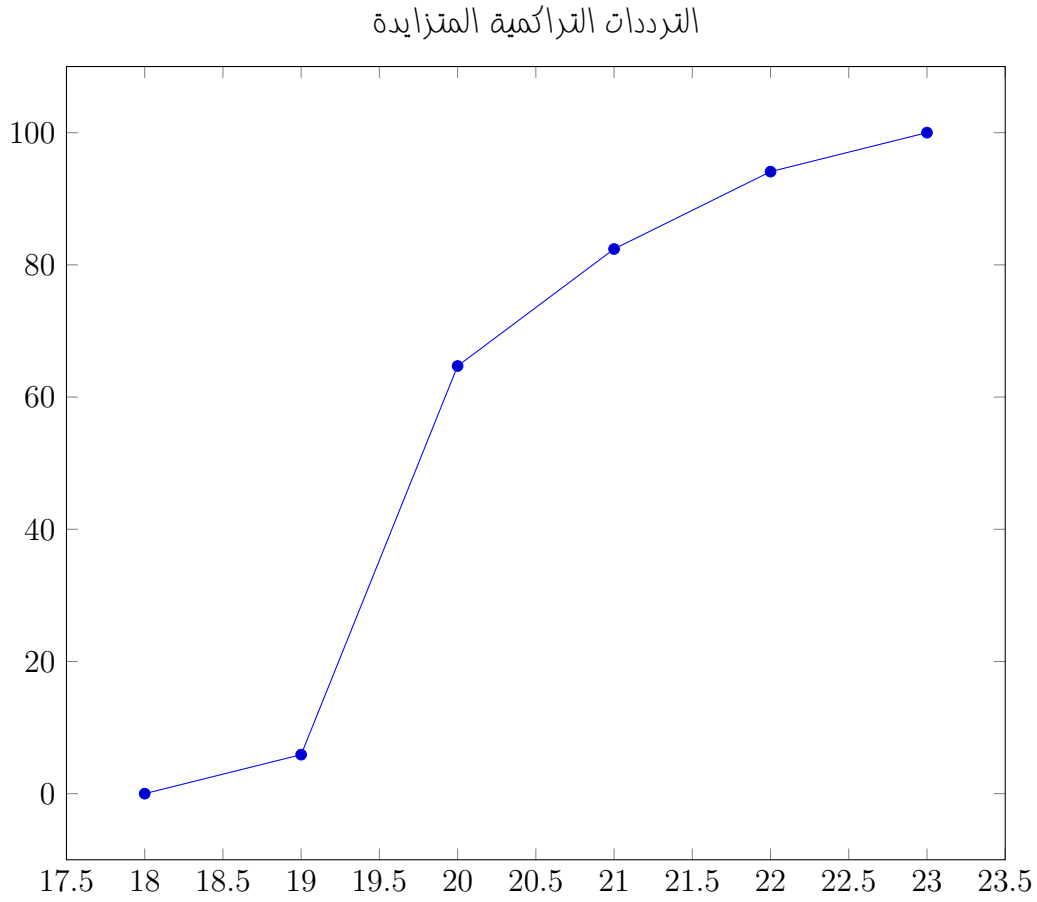
وبالمثل ، فإن إحتمال أن يكون X بين a و b، حيث $(b > a)$ يحقق

$$P(a < X < b) = F(b) - F(a).$$

الرسم البياني المتكامل أو دالة التوزيع هو الرسم البياني الذي يستخدم لتمثيل جدول زيادة التردادات التراكمية. ولإنشائه في حالة وجود متغير مستمر ، نضع على محور الفواصل، الأطراف لa_i لفئات C_i وعلى المحور التراتيب التردادات التراكمية les fréquences cumulées F_i بالطريقة التالية:

1. من أجل كل فئة $C_i =]a_i, a_{i+1}]$ نعين النقطة $M_i = (a_{i+1}, F_i)$, حيث $i = 1, 2, ..., k$ الموافقة لها
2. نعين النقطة $M'_0 = (a_0, 0)$.
3. نربط بين كل النقاط بقطع مستقيمة
4. نمدد بـ 0 على اليسار وبالـ 1 على اليمين لنحصل على دالة مستمرة.

فيما يلي نوضح منحنى دالة التوزيع للمثال 3 السابق

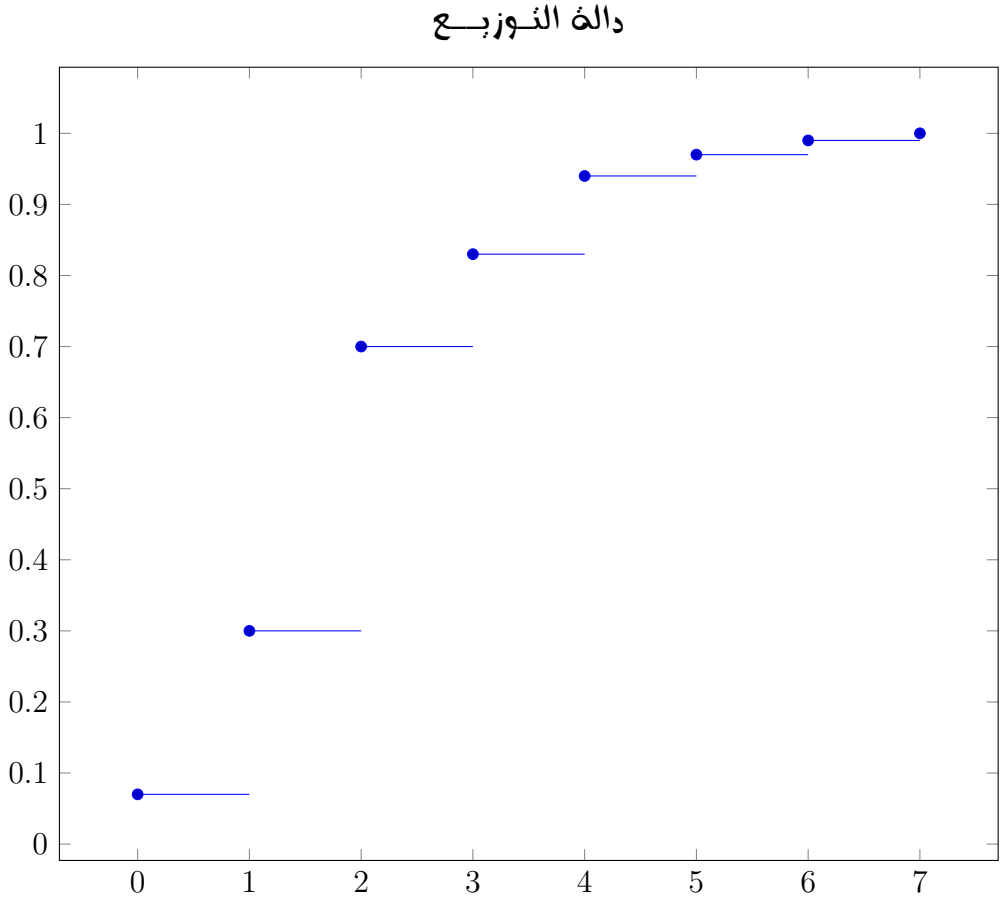


بالنسبة للمتغيرات الكمية المنفصلة، يتم وضع الترددات التراكمية المتزايدة les fréquences cumulées croissantes **على محور التراتيب، ثم يتم تمديد كل نقطة بقطعة أفقية للحصول على دالة درجية** fonction en escalier .

مثال 5 : لنكن السلسلة الإحصائية البسيطة التالية التي تمثل عدد الأطفال دون سن 15 لـ 100 أسرة في قرية ما:

عدد الأطفال	0	1	2	3	4	5	6	7
عدد التكرارات	7	23	40	13	11	3	2	1

يسمى هذا الجدول توزيع السلسلة الإحصائية *Distribution de la série statistique*. دالة التوزيع للسلسلة المذكورة في هذا المثال هي:



6.1.2. التمثيل بالدائرة البيانية Diagramme circulaire

هذا التمثيل غير فعال بصريا (قراءة الزاوية الصغيرة بصريا صعب)، ولا يسمح بإجراء مقارنة دقيقة لتردد كل فئة.

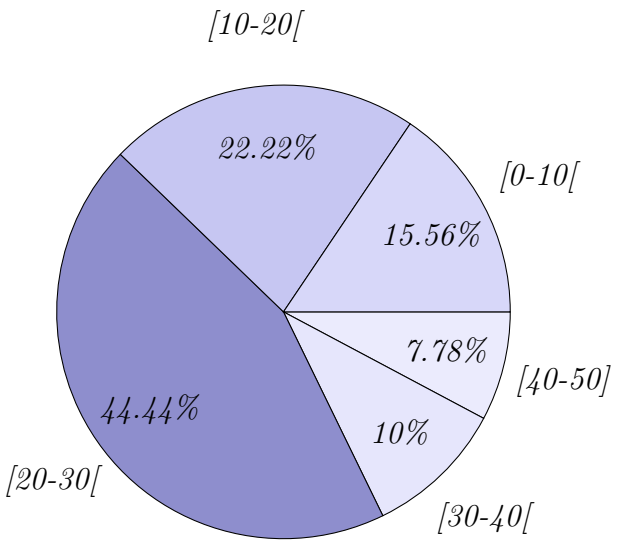
لإنشاء مخطط دائري ، ما عليك سوى مطابقة كل تكرار أو نسبة مئوية أو تردد ، بقياس زاوية متناسبة.

إذا كان مجموع التكرارات هو n وكانت للسلسلة الإحصائية عدد k من الفئات وكان لكل فئة $i = 1, 2, ..., k$ العدد التكراري n_i. فإنه يتم حساب الزوايا An_i من المخطط الدائري بالقاعدة التالية:

$$An_i = \frac{n_i}{n} \cdot 360^\circ$$

مثال 6 : في شركة لنقل و تسليم البضائع، نقوم بتصنيف 90 عملية تسليم ، اعتمادًا على حجم البضائع المنقولة. بإستعمال قاعدة حساب الزوايا نضع النتائج المتحصل عليها في الجدول التالي:

الحجم m^3	عدد التوصيلات	التكافة % بتقريب 0.01 %	الزاوية بالدرجة الزاوية بتقريب 0.01	الزاوية بالتراكم
$[0;10[$	14	15.56	56.02	56.02
$[10;20[$	20	22.22	79.99	136.00
$[20;30[$	40	44.44	159.98	295.99
$[30;40[$	9	10.00	36.00	331.99
$[40;50]$	7	7.78	28.01	360.00
المجمـــوع	90	100.00	360.00	



7.1.2. مخطط الصندوق Boîtes à moustaches

إن مخطط الصندوق أو الرسم الصندوقي Boîtes à moustaches، **وهي ترجمة لـ** Box & Whiskers Plot ، **هي اختراع لـ** TUKEY (1977) **لتمثيل التوزيع بشكل تخطيطي. يمكن أن يكون هذا الرسم البياني وسيلة لمقاربة المفاهيم المجردة للإحصاءات.**

تعريف 3.1.2 : مخطط الصندوق هو طريقة للتمثيل البياني لمجموعة من القيم العددية لعينة إحصائية من خلال تمثيل القيم الإحصائية الخمس المحددة للعينة وهي: القيمة الصغرى ، الرُبيع

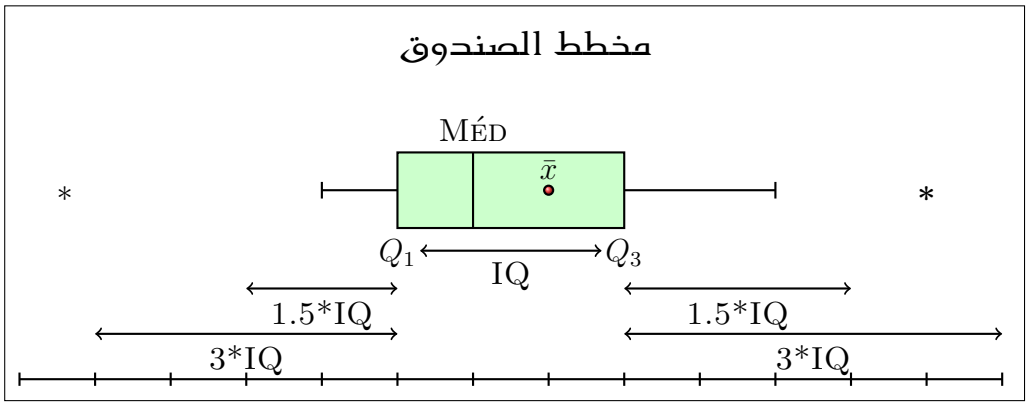
الأدنى Q_1، الوسيط Q_2، الرُبيع الأعلى Q_3، والقيمة العظمى. ويمكن لمخطط الصندوق أن يشير أيضاً إلى قراءات العينة التي تم اعتبارها قيماً شاذّة.

يظهر التمثيل البياني بالصندوق وطرفيه الاختلافات بين جمهرات مختلفة من دون الأخذ بأي من فرضيات التوزع الاحتمالي لهذه الجمهرات بمعنى آخر هو تمثيل للقيم الحقيقية للعينة وليس تمثيلا لقيم مستنتجة من توزيع الاحتمال.

عند تمثيل مجموعة القيم الرقمية بهذه الطريقة، تظهر المسافات المختلفة في الصندوق الخاص بالمجموعة مميزات هذه المجموعة مثل درجة التشتت أو النزعة المركزية للقيم، ودرجة الالتواء في العينة، كما تمكن من تحديد القيم الشاذة. يمكن رسم الصندوق بشكل أفقي أو عامودي.

يتم تشكيل مخطط الصندوق بإتباع الخطوات التالية :

1. نرسم مربع طوله $Q_3 - Q_1$.
2. نقسم المربع عموديا في نقطة المتوسط الحسابي La médiane
3. نمثل بخط أفقي أكبر وأصغر قيمة في مجموعة القيم واللتان تقعان ضمن المجال $Q_3 + 1.5(Q_3 - Q_1)$.
4. و تمثيل رأس الشارب العلوي هو الحد الأقصى لقيمة البيانات التي تقل عن قيمة الحد العلوي للمجال $Q_3 - 1.5(Q_3 - Q_1)$.
5. إذا كان بعض الأفراد خارج مجال شعيرات شوارب الصندوق ، فيتم تمثيلهم بواسطة *.



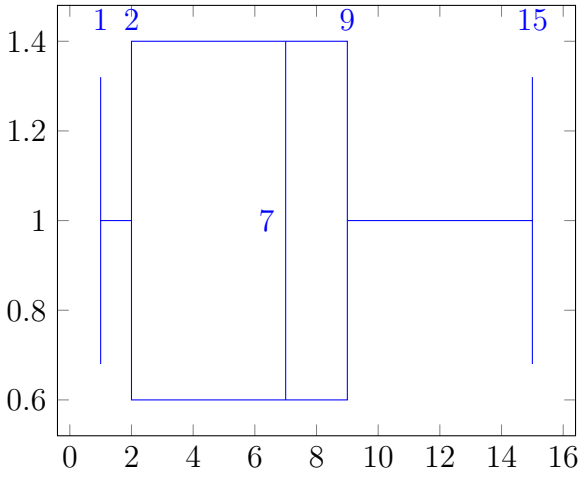
مثال 7 : لننشئ مخطط الصندوق للسلسلة الإحصائية الثالية

1	2	1	5	4	10	7	15	9	8	9	9

أولا نتحصل على الملخص التالي

القيمة الصغرى	1
Q_1	3
الوسيط الحسابي	7.5
المتوسط الحسابي	6.25
Q_3	9
القيمة القصوى	15
IQ	6
$Q_3 + 1.5IQ$	18
$Q_3 - 1.5IQ$	0
الشارب العلوي	15
الشارب السفلي	1

و منه نستطيع رسم البيان كما في الشكل



ملاحظة 3 : تعطي مقارنة المخططات الصندوقية لكل مجموعة من البيانات فكرة جيدة عن تشتت البيانات ، مع تصور الوسيط (الذي يعتبر في الغالب أفضل من المتوسط *La moyenne*).

تمارين مفتوحة

2.2 سلسلة التمارين رقم 2

تمرين 1 : في منطقة صغيرة ، أخذنا عدد الغرف لكل شقة فتحصلنا على الجدول التالي:

عدد الغرف	1	2	3	4	5	6
عدد الشقق	51	70	98	62	40	26

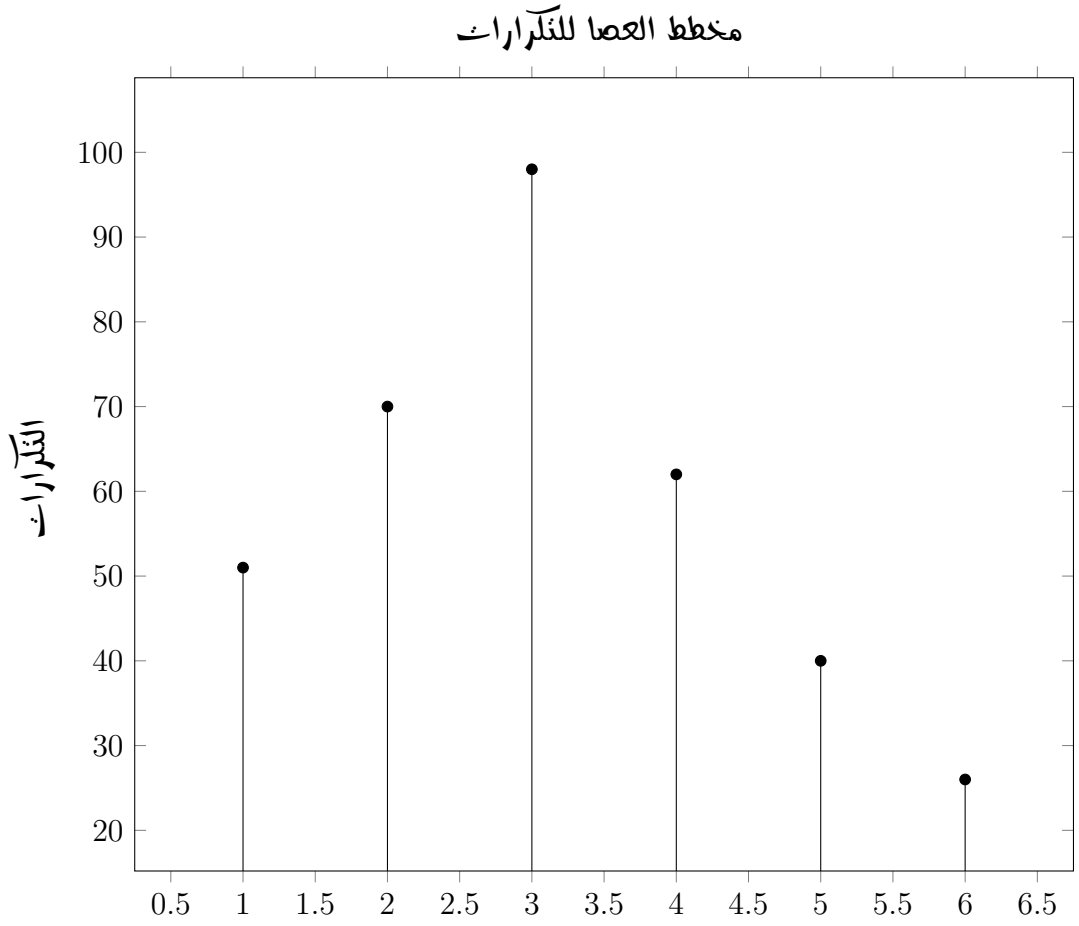
1- مثل بيانيا المعطيات السابقة باستخدام بيان العصا.

2- أحسب التكرارات التراكمية ثم أرسم منحنى دالة التوزيع للتكرارات.

3- ارسم مخطط العصا للتواترات ومنحنى دالة توزيع التواترات.

الحـــل

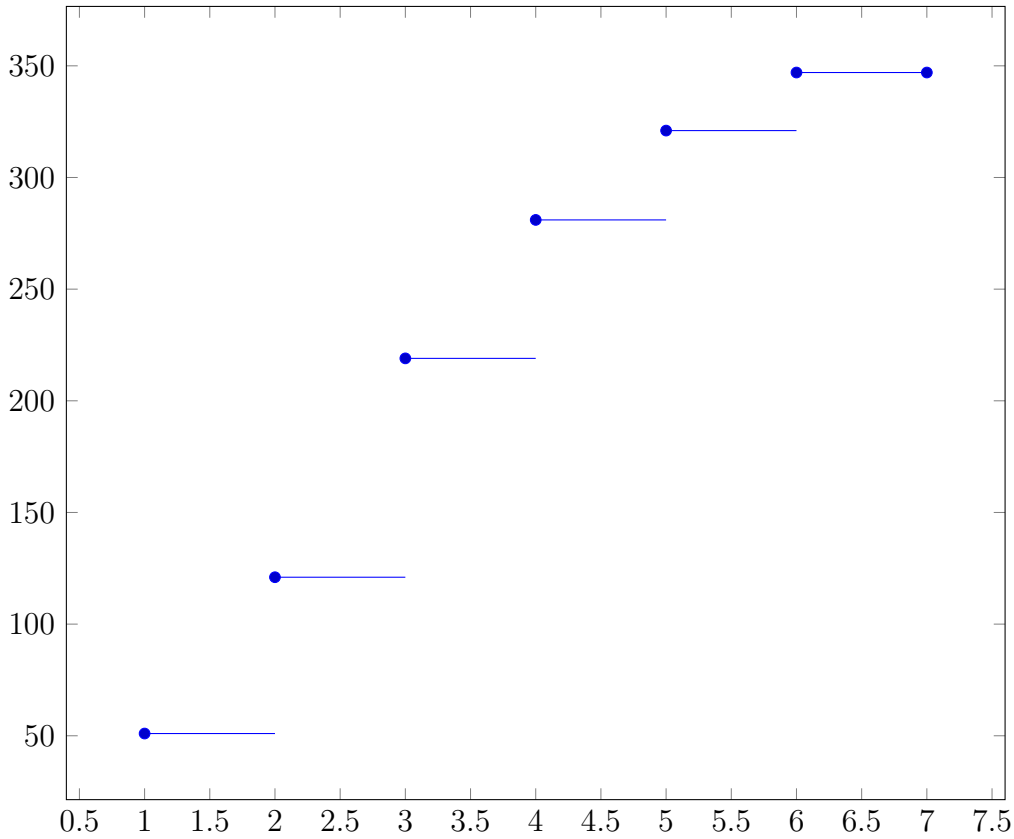
1- التمثيل البياني لمخطط العصا



2- حساب التكرارات التراكمية

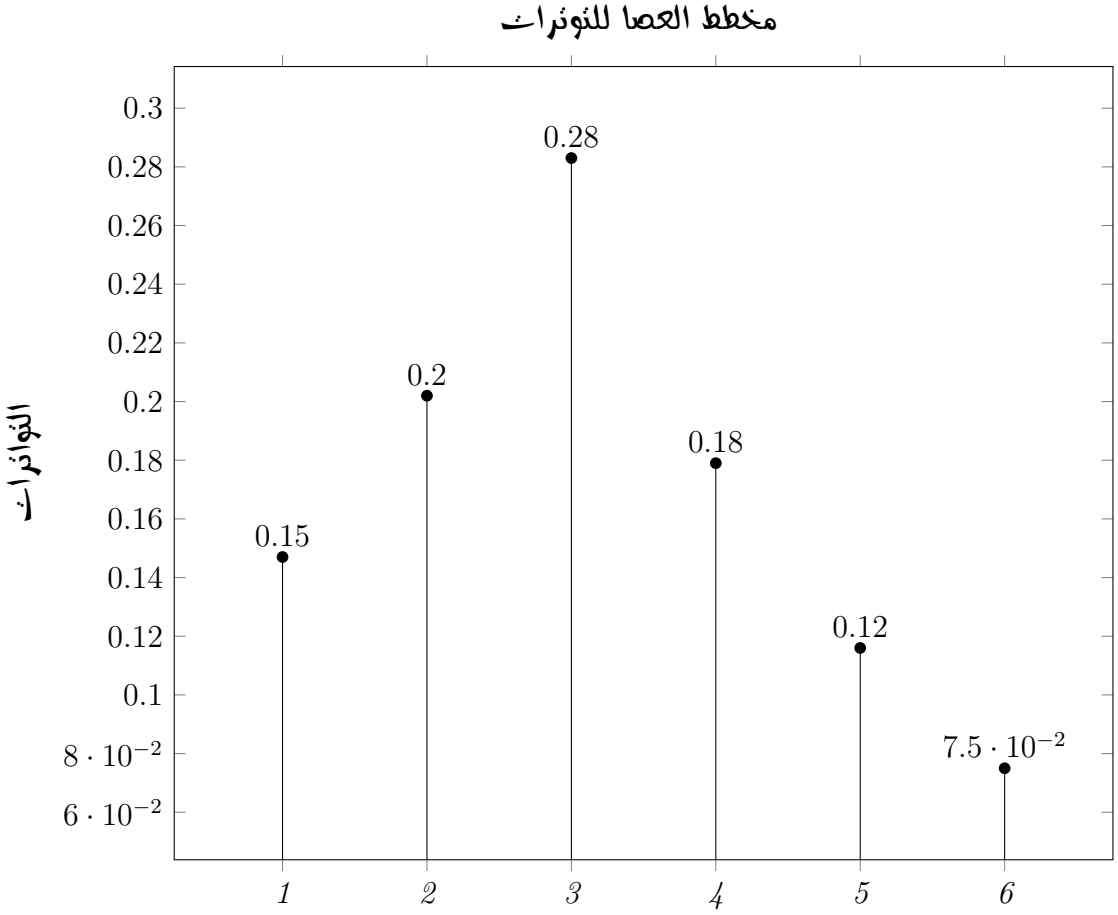
6	5	4	3	2	1	الشكل *Modalités*
26	40	62	98	70	51	التكرارات
347	321	281	219	121	51	التكرارات التراكمية

دالة توزيع التكرارات



3- لرسم مخطط العصا للتواترات: نحسب التواترات أولا

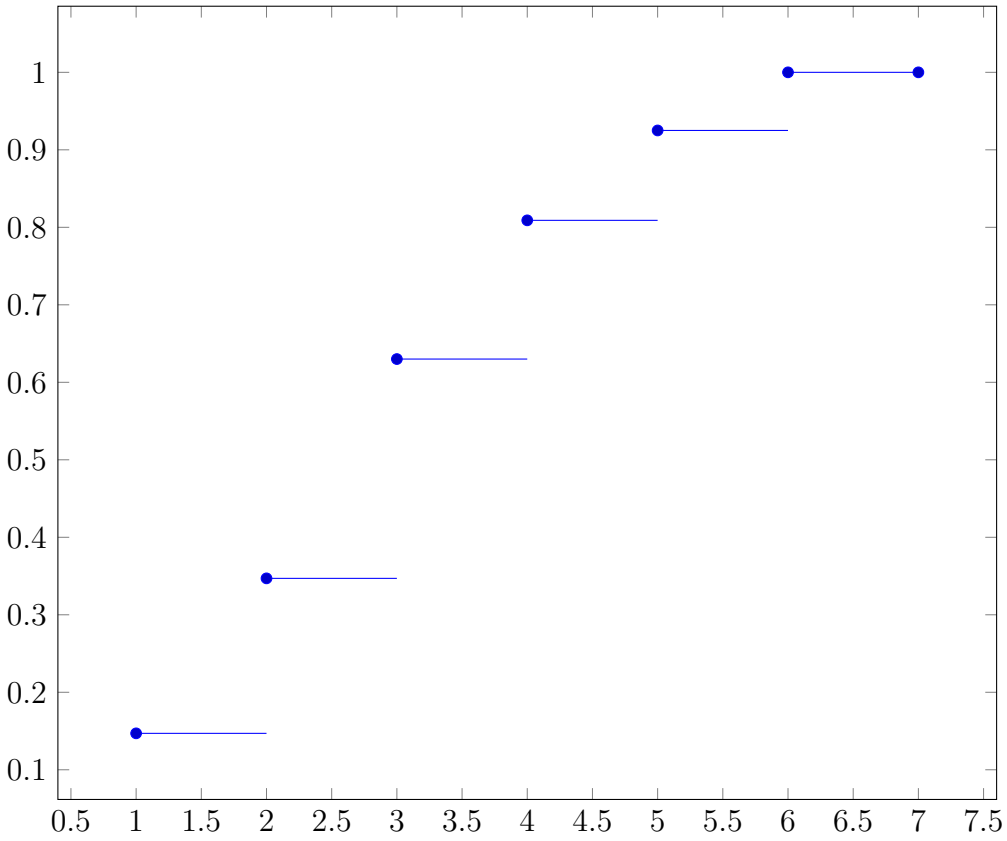
الشكل	1	2	3	4	5	6
التكرارات	51	70	98	62	40	26
التواترات	0.147	0.202	0.283	0.179	0.116	0.075



دالة توزيع التواترات: نحسب التواترات التراكمية

الشكل	1	2	3	4	5	6
التكرارات	51	70	98	62	40	26
التواترات	0.147	0.202	0.283	0.179	0.116	0.075
التواترات التراكمية	0.147	0.347	0.63	0.809	0.925	1

دالة تـوزيـــع التواتر



▼

تمرين 2 : خلال دراسة استقصائية للتسويق ، ندرس الحالة الزوجية لـ 15 مشتري لسيارة *Symbol Algeria*. نحصل على السلسلة التالية:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_i	C	M	D	V	V	M	D	M	M	C	D	M	D	M	C

$C = Célibataire$, $M = Mariée$ $D = Divorcé$ et $V = Veuf$.

1- ما هو المتغير قيد الدراسة؟ ما هي أشكاله؟

2- أوجد توزيع التكرارات وتوزيع التواتر لهذا المتغير.

3- مثل بيانياً توزيع التواتر التراكمية المتنازلة باستخدام مخطط العصا.

الحـــل

1- المتغير المدروس هو الحالة الزوجية للمشتريين ، وتتكون من أربع أشكال: أعزب ، متزوج ، مطلق و أرمل.

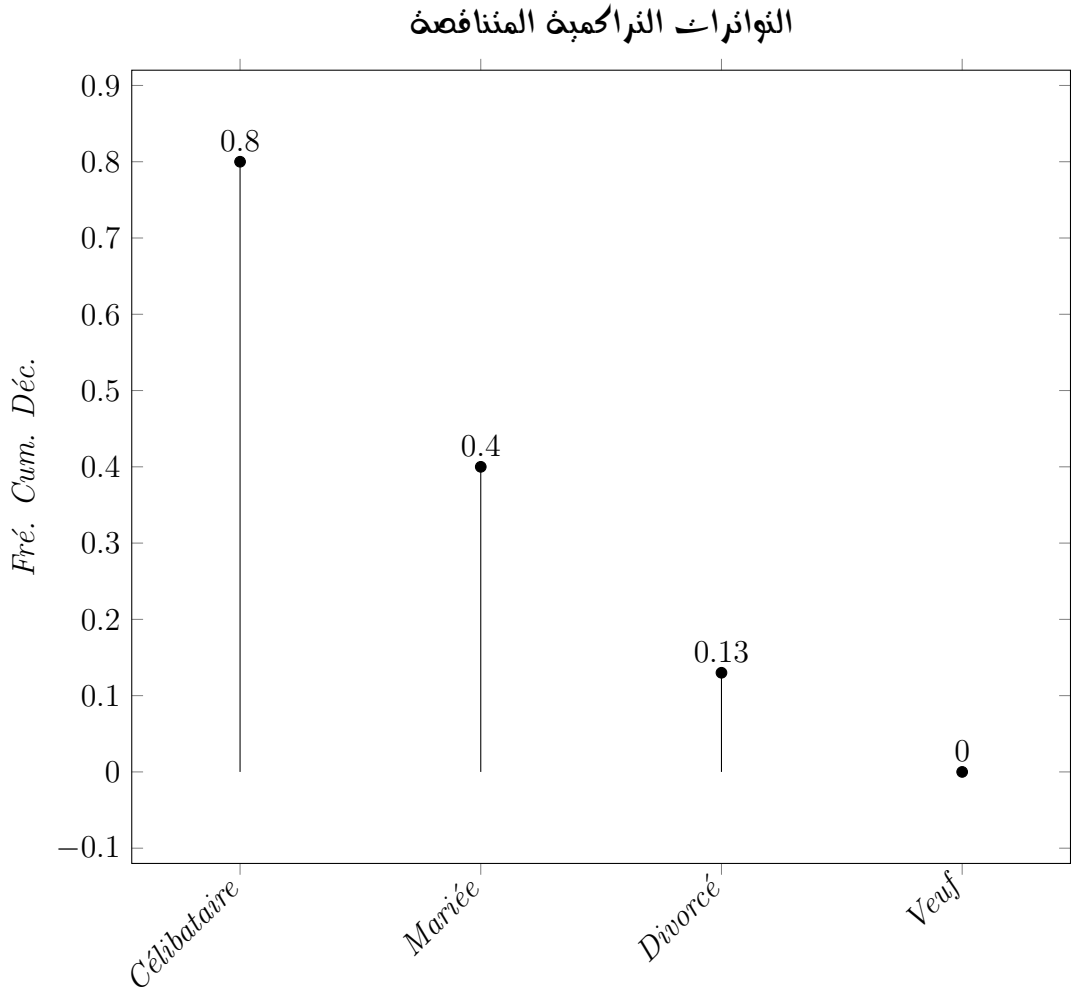
2- أ) توزيع التكرارات

الشكل	C	M	D	V
التكرارات n_i	3	6	4	2

ب) توزيع التواترات

الشكل	C	M	D	V
التواترات f_i	0.2	0.4	0.27	0.13

3- التمثيل بواسطة مخطط العصا لتوزيع التواترات التراكمية المتنافصة.


التواترات التراكمية المتناقصة
Fré. Cum. Déc.
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
0.8
0.4
0.13
0
Célibataire
Mariée
Divorcé
Veuf

▼