

Pr. Oubrahim
Pr. Madkour

TD DE STATISTIQUE DESCRIPTIVE

Série 1 : Introduction à la statistique.

Ex. 1 — Pour garantir l'égalité des chances dans un concours, les villes de provenance des candidats ont été codées de la façon suivante : Agadir : 1, Casablanca : 2, Fès : 3, Marrakech : 4, Rabat : 5 et les autres villes : 6. La répartition des candidats selon leurs villes de provenance est présentée dans le tableau suivant :

x_i	1	2	3	4	5	6
n_i	27	52	48	37	68	62

1. Quelle est la population étudiée ? Quelles sont ses unités statistiques ?
2. Quel est le caractère étudié ? Quelle est sa nature ?
3. Que représente chacune des lignes du tableau ? Que représente le couple $(x_i ; n_i)$?
4. Quel est l'effectif total de cette population ?
5. Construire le tableau statistique associé à cette distribution des données.
6. Représenter graphiquement les effectifs et les fréquences.
7. Quelle est la ville la plus représentée par les candidats ? Quelle est la ville la moins représentée ?

Answer (Ex. 1) —

1. La population étudiée est « l'ensemble des candidats à un concours ». Chaque candidat est une unité statistique.
2. Le caractère étudié est « la ville de provenance des candidats ». Il s'agit d'un caractère qualitatif nominal : il est qualitatif parce qu'une ville n'est pas mesurable et même si, dans cet exercice, les villes sont représentées par des nombres, ces derniers n'ont pas de signification mathématique car $1+2=3$ par exemple ne signifie pas que la ville de Fès est la somme des villes d'Agadir et de Casablanca. Le caractère est nominal puisqu'il n'existe aucun classement objectif entre les villes.
3. La première ligne du tableau représente les modalités du caractère « ville de provenance des candidats », et la deuxième ligne représente les effectifs des modalités. Le couple $(x_i ; n_i)$ associé à chaque modalité x_i son effectif n_i , il s'agit donc d'une distribution des effectifs.
4. L'effectif total de cette population est :

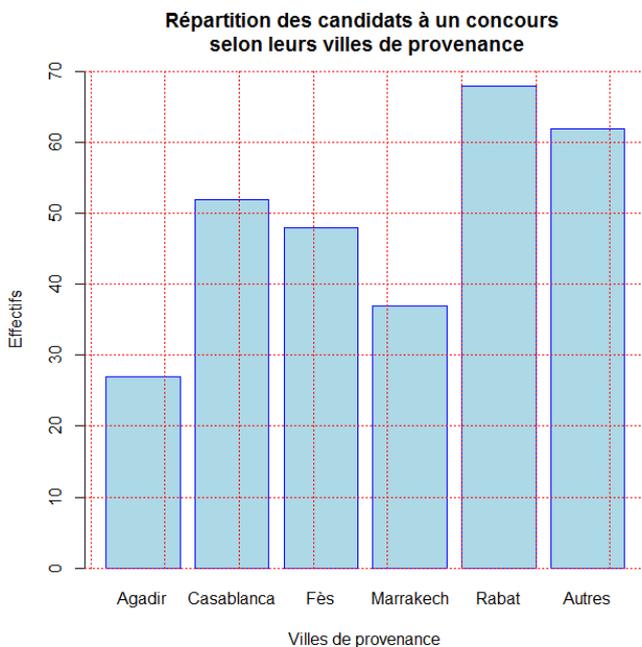
$$\begin{aligned}n &= \sum_{i=1}^6 n_i \\ &= 27 + 52 + 48 + 37 + 68 + 62 \\ &= 294\end{aligned}$$

Le tableau statistique associé à cette distribution des données :

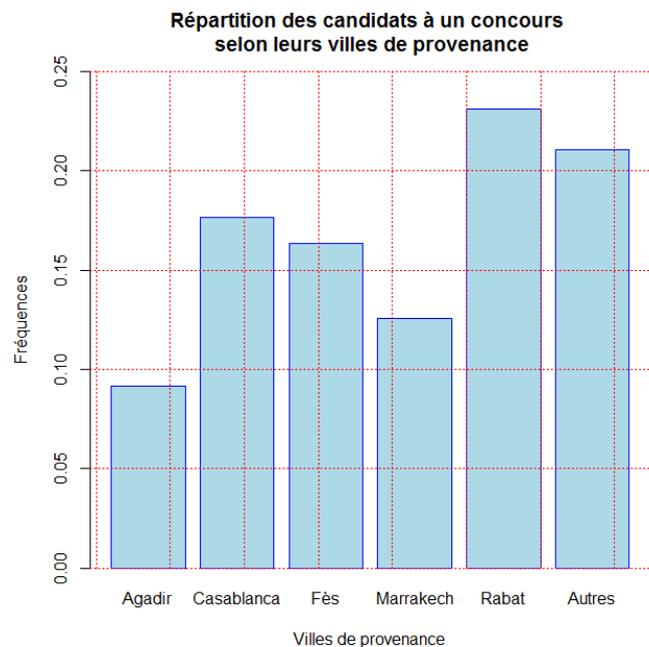
x_i	n_i	f_i
1 : Agadir	27	0.09
2 : Casablanca	52	0.18
3 : Fès	48	0.16
4 : Marrakech	37	0.13
5 : Rabat	68	0.23
6 : Autres	62	0.21
Total	294	1

Remarque : le caractère étudié étant qualitatif nominal, on ne peut pas calculer les effectifs et les fréquences cumulés.

6. Représentations graphiques des effectifs et des fréquences (Diagrammes en barres) :



(a) Effectifs



(b) Fréquences

7. D'après le tableau statistique et les diagrammes en barres, la ville la plus représentée par les candidats est Rabat et la ville la moins représentée est Agadir.

Ex. 2 — La répartition des étudiants d'une licence selon la mention obtenue est donnée dans le tableau suivant :

Mention obtenue	Passable	Assez bien	Bien	Très bien	Excellent
Nombre d'étudiants	127	57	48	7	1

1. Quelle est la population étudiée ? Quelles sont ses unités statistiques ?

2. Quel est le caractère étudié ? Quelle est sa nature ?
3. Que représente chacune des lignes du tableau ?
4. Quel est l'effectif total de cette population ?
5. Construire le tableau statistique associé à cette distribution des données.
6. Représenter graphiquement les effectifs et les fréquences.
7. Quelle est la mention dominante dans cette promotion d'étudiants ?

Answer (Ex. 2) —

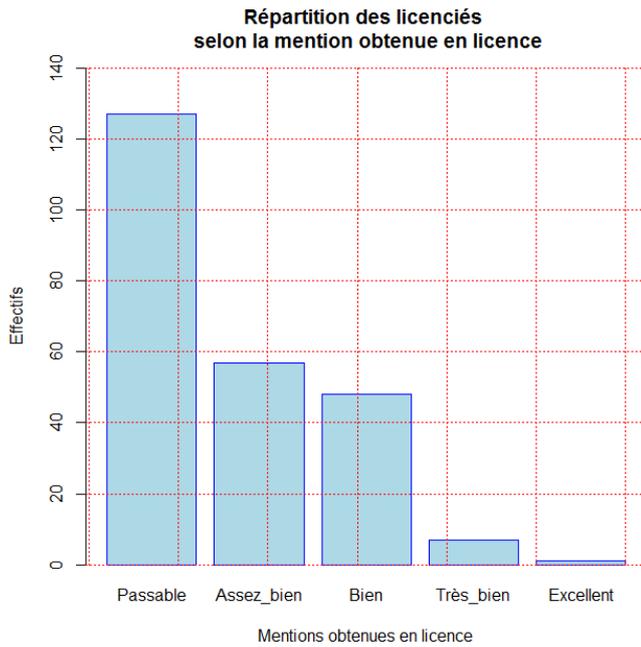
1. La population étudiée est « l'ensemble des étudiants licenciés ». Chaque étudiant licencié est une unité statistique.
2. Le caractère étudié est « la mention obtenue en licence ». Il s'agit d'un caractère qualitatif ordinal : il est qualitatif car « une mention » n'est ni mesurable ni représentable par un nombre ayant un sens mathématique. Il est ordinal car on peut classer les mentions, la mention « Bien » par exemple est toujours supérieure à la mention « Passable ».
3. La première ligne du tableau représente les modalités du caractère « mention obtenue en licence » et la deuxième ligne correspond aux effectifs des ces modalités.
4. L'effectif total de cette population est :

$$\begin{aligned}
 n &= \sum_{i=1}^5 n_i \\
 &= 127 + 57 + 48 + 7 + 1 \\
 &= 240
 \end{aligned}$$

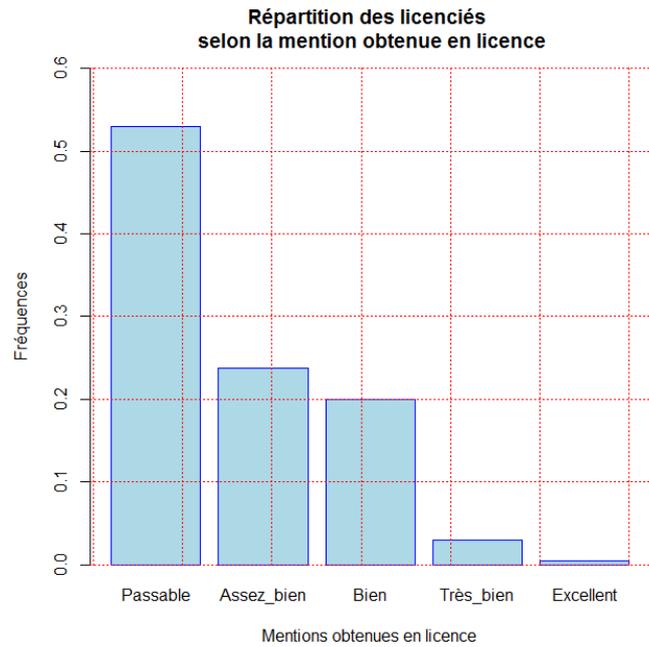
5. Le tableau statistique associé à cette distribution des données :

Mention	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
Passable	127	0.529	127	240	0.529	1
Assez bien	57	0.238	184	113	0.768	0.471
Bien	48	0.200	232	56	0.967	0.233
Très bien	7	0.029	239	8	0.996	0.033
Excellent	1	0.004	240	1	1	0.004
Total	240	1	–	–	–	–

6. Représentations graphiques des effectifs et des fréquences (Diagrammes en barres) :



(a) Effectifs



(b) Fréquences

7. C'est la mention « Passable » qui est dominante dans cette promotion d'étudiants.

Ex. 3 — La répartition des salaires horaires en dirhams d'ouvriers du secteur industriel est la suivante :

Salaires horaires (en DH)	[12.85 ; 13.00[[13.00 ; 13.25[[13.25 ; 13.75[[13.75 ; 14.00[[14.00 ; 15.00[
Nombre d'ouvriers	210	184	116	92	76

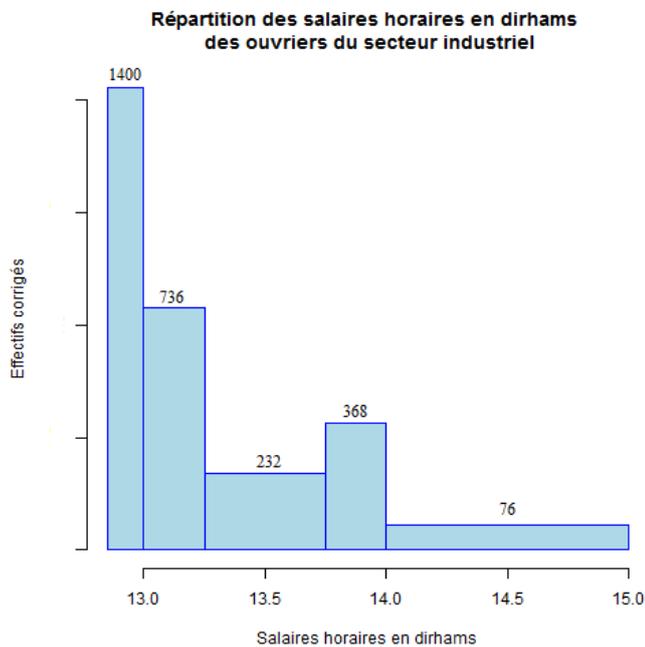
1. Quelle est la nature de la variable statistique « salaires horaires » ?
2. Construire le tableau statistique.
3. Représenter graphiquement les effectifs et les effectifs cumulés.
4. Quel est le nombre d'ouvriers qui touchent au moins 13.25 dirhams par heure ?
5. Quel est le nombre d'ouvriers qui touchent au plus 13 dirhams par heure ?

Answer (Ex. 3) —

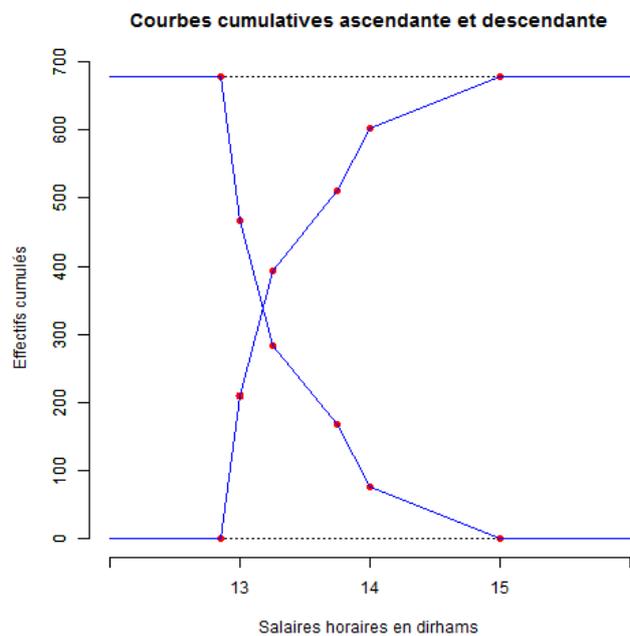
1. La variable statistique « salaires horaires » est quantitative continue : elle est quantitative parce qu'elle est mesurable en dirhams et elle est continue parce qu'elle peut prendre toutes les valeurs possible d'un intervalle.
2. Le tableau statistique :

Salaires	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-	a_i	n_i^c
[12.85; 13.00[210	0.31	210	678	0.31	1	0.15	1400
[13.00; 13.25[184	0.27	394	468	0.58	0.69	0.25	736
[13.25; 13.75[116	0.17	510	284	0.75	0.42	0.50	232
[13.75; 14.00[92	0.14	602	168	0.89	0.25	0.25	368
[14.00; 15.00[76	0.11	678	76	1	0.11	1	76
Total	678	1	-	-	-	-	-	-

3. Représentations graphiques des effectifs et des effectifs cumulés :



(a) Effectifs



(b) Effectifs cumulés

- D'après la colonne des effectifs cumulés décroissants (N_i^-), le nombre d'ouvriers qui touchent au moins 13.25 dirhams par heure est 284.
- D'après la colonne des effectifs cumulés croissants (N_i^+), le nombre d'ouvriers qui touchent au plus 13 dirhams par heure est 210.

Ex. 4 — Les notes sur 20, arrondies à l'entier supérieur, d'une classe de master obtenues dans une épreuve écrite sont les suivantes :

14	10	12	4	16	14	12	14	6	18
20	8	10	10	8	12	14	8	16	4
6	10	12	14	18	16	4	8	14	18
8	12	14	16	6	12	14	18	20	10

1. Quel est le caractère étudié ? Quelle est sa nature ?
2. Que représente la suite de ces notes ?
3. Construire le tableau statistique associé à cette distribution des notes.
4. Représenter graphiquement les effectifs et les effectifs cumulés.
5. Quel est le nombre d'étudiants qui ont obtenu une note égale à 14 ?
6. Quel est le nombre d'étudiants qui ont obtenu une note au plus égale à 20 ?
7. Quel est le nombre d'étudiants qui ont obtenu une note au moins égale à 18 ?

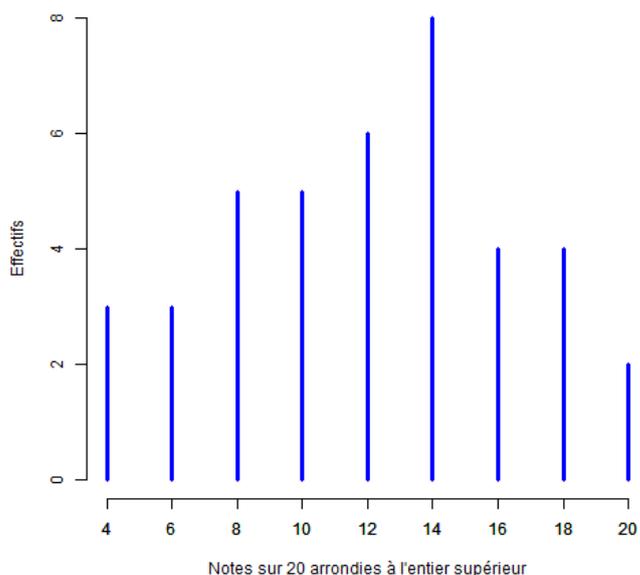
Answer (Ex. 4) —

1. Le caractère étudié est « la note sur 20 obtenue dans une épreuve écrite ». Il s'agit d'un caractère quantitatif car une note est un nombre au sens mathématique du terme sur lequel on peut appliquer toute sorte d'opérations arithmétiques (addition, soustraction...etc). Comme les notes sont arrondies à l'entier supérieur, le caractère, qui est à l'origine continu, s'est transformé en caractère discret.
2. La suite de ces notes est une série statistique brute puisqu'il s'agit d'une suite d'observations non ordonnées.
3. Le tableau statistique associé à cette distribution des notes :

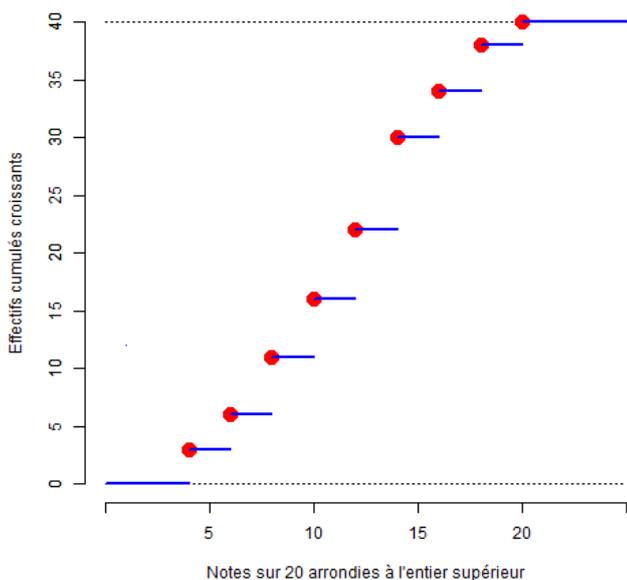
Notes	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
4	3	0.075	3	40	0.750	1
6	3	0.075	6	37	0.150	0.925
8	5	0.125	11	34	0.275	0.850
10	5	0.125	16	29	0.400	0.725
12	6	0.150	22	24	0.550	0.600
14	8	0.200	30	18	0.750	0.450
16	4	0.100	34	10	0.850	0.250
18	4	0.100	38	6	0.950	0.150
20	2	0.050	40	2	1	0.050
Total	40	1	—	—	—	—

4. Représentations graphiques des effectifs et des effectifs cumulés :

Répartition des notes sur 20, arrondies à l'entier supérieur, d'une classe de master obtenues dans une épreuve écrite

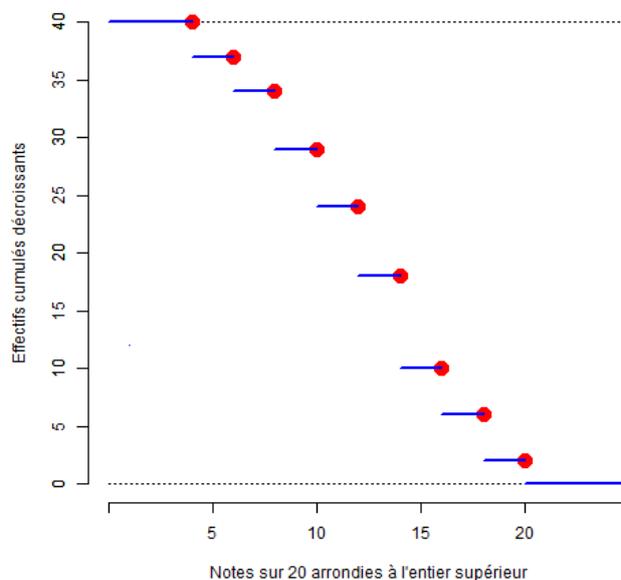


Courbe cumulative ascendante



(a) Effectifs cumulés croissants

Courbe cumulative descendante



(b) Effectifs cumulés décroissants

5. 8 étudiants ont obtenu une note égale à 14.
6. 40 étudiants ont obtenu une note au plus égale à 20.
7. 6 étudiants ont obtenu une note au moins égale à 18.

Ex. 5 — La facture téléphonique mensuelle d'un abonné à un réseau mobile fait apparaître, entre autres, la durée en minutes de chaque appel passé par cet abonné. Le tableau suivant donne la répartition du nombre d'appels selon leurs durées :

Durée de l'appel	[0; 5[[5; 10[[10; 15[[15; 20[[20; 25[[25; 30[
Nombre d'appels	26	16	10	0	2	2

1. Quelle est la variable statistique étudiée ? Quelle est sa nature ?
2. Construire le tableau statistique.
3. Représenter graphiquement les fréquences et les fréquences cumulées.
4. Quelle est la proportion des durées d'appels les plus longues ?
5. Quelle est la proportion des durées d'appels au plus égales à 10 minutes ?
6. Quelle est la proportion des durées d'appels au moins égales à 20 minutes ?

Answer (Ex. 5) —

1. La variable statistique étudiée est « la durée en minutes des appels téléphoniques » émis par un abonné à un réseau mobile. Il s'agit d'une variable quantitative continue. Elle est quantitative puisque la durée d'un appel téléphonique est mesurée en minutes et elle est continue car elle peut prendre toutes les valeurs possibles d'un intervalle.
2. Le tableau statistique :

Salaires	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-	a_i	f_i^c
[0; 5[26	0.4643	26	56	0.4643	1	5	0.09
[5; 10[16	0.2857	42	30	0.7500	0.5357	5	0.06
[10; 15[10	0.1786	52	14	0.9286	0.2500	5	0.04
[15; 20[0	0	52	4	0.9286	0.0714	5	0
[20; 25[2	0.0357	54	4	0.9643	0.0714	5	0.01
[25; 30[2	0.0357	56	2	1	0.0357	5	0.01
Total	56	1	-	-	-	-	-	-

3. Représentations graphiques des fréquences et des fréquences cumulées : (cf. page 1 du document ci-joint)
4. La proportion des durées d'appels les plus longues est de 0.0357.
5. La proportion des durées d'appels au plus égales à 10 minutes est de 0.75.
6. La proportion des durées d'appels au moins égales à 20 minutes de 0.0714.

Ex. 6 — La facture téléphonique de l'exercice précédent fait apparaître également le nombre d'appels téléphoniques passés par jour par le client :

Nombre d'appels par jour	0	1	2	3	4	5	6
Nombre de jours	2	14	8	2	1	2	1

1. Quelle est la variable statistique étudiée ? Quelle est sa nature ?
2. Construire le tableau statistique.

3. Représenter graphiquement des fréquences et des fréquences cumulées.
4. Quelle est la proportion des jours pendant lesquels le client a passé 2 appels téléphoniques ?
5. Quelle est la proportion des jours pendant lesquels le client a passé au moins 2 appels ?
6. Quelle est la proportion des jours pendant lesquels le client a passé au plus 5 appels ?

Answer (Ex. 6) —

1. La variable statistique étudiée est « le nombre d'appels téléphoniques passés par jour » par le client. Il s'agit d'une variable quantitative discrète : elle est quantitative car on peut compter le nombre d'appels et elle est discrète puisqu'elle ne peut prendre que des valeurs entières et isolées.
2. Le tableau statistique :

Notes	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
0	2	0.0667	2	30	0.0667	1
1	14	0.4667	16	28	0.5333	0.9333
2	8	0.2667	24	14	0.8000	0.4667
3	2	0.0667	26	6	0.8667	0.2000
4	1	0.0333	27	4	0.9000	0.1333
5	2	0.0667	29	3	0.9667	0.1000
6	1	0.0333	30	1	1	0.0333
Total	30	1	–	–	–	–

3. Représentations graphiques des fréquences et des fréquences cumulées : (cf. page 2 du document ci-joint)
4. La proportion des jours pendant lesquels le client a passé 2 appels téléphoniques est 0.2667.
5. La proportion des jours pendant lesquels le client a passé un nombre d'appels au moins égal à 2 est 0.4667.
6. La proportion des jours pendant lesquels le client a passé un nombre d'appels au plus égal à 5 est 0.9667.