Représentations graphiques

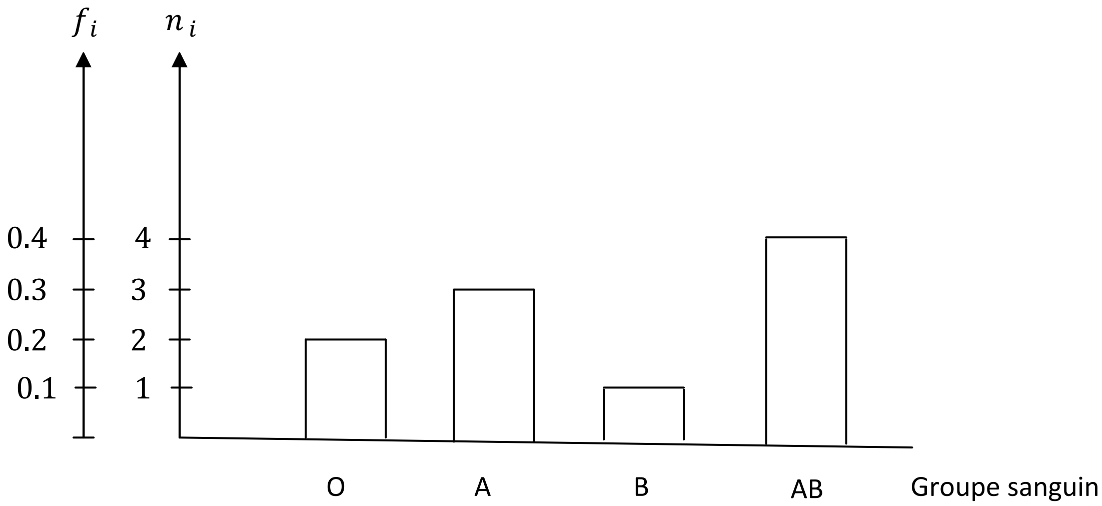
1. Caractère qualitatif

a. Diagramme à bandes (diagramme en tuyaux d'orgue)

Un diagramme à bandes est un graphique qui à chaque modalité de la variable (caractère) associe un rectangle (de base constante) dont la hauteur est proportionnelle à l'effectif (ou à la fréquence). Dans un système d'axes orthogonaux, les modalités du caractère sont placées sur l'axe horizontal (axe non orienté) et les effectifs (ou les fréquences) sont portés sur l'axe vertical (axe orienté).

Exemple : Groupe sanguin

Modalité x_i	Effectif n_i	Fréquence f_i
O	2	0.2
A	3	0.3
B	1	0.1
AB	4	0.4
$\sum$	10	1



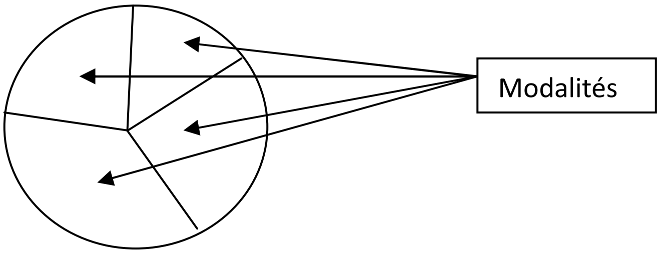
Remarque : L'ordre dans lequel sont présentés les tuyaux n'a pas d'importance.

b. Diagramme à secteurs circulaires (ou camemberts)

L'effectif total est représenté par un disque. Chaque modalité est représentée par un secteur circulaire dont la surface (l'angle au centre) est proportionnelle à l'effectif correspondant.

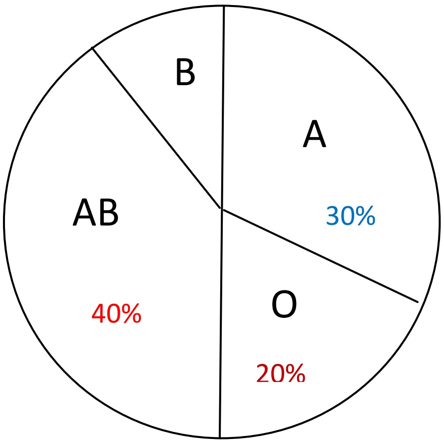
L'angle au centre $\theta_i = f_i * 360^{\circ}$

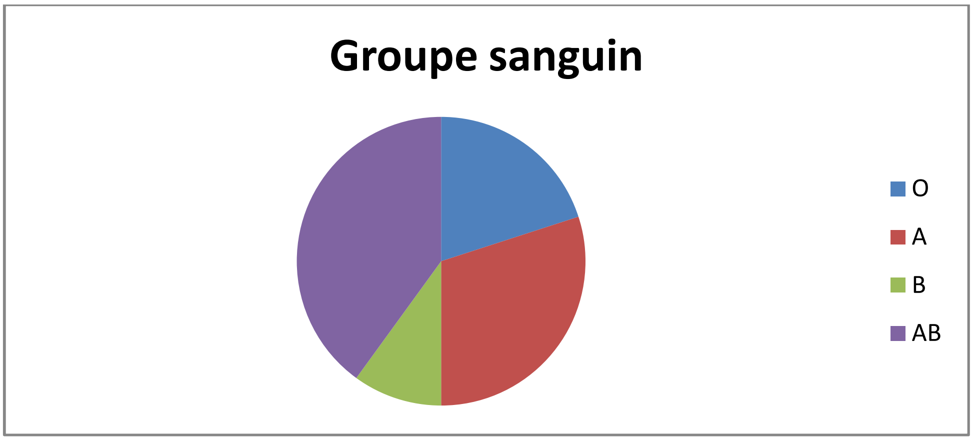
$$n \to 360^{\circ}$$
$$n_i \to \theta_i$$



Exemple : Groupe sanguin

Groupe sanguin	f_i	$\theta_i = f_i * 360^{\circ}$
O	0.2	72°
A	0.3	108°
B	0.1	36°
AB	0.4	144°





2. Caractère quantitatif

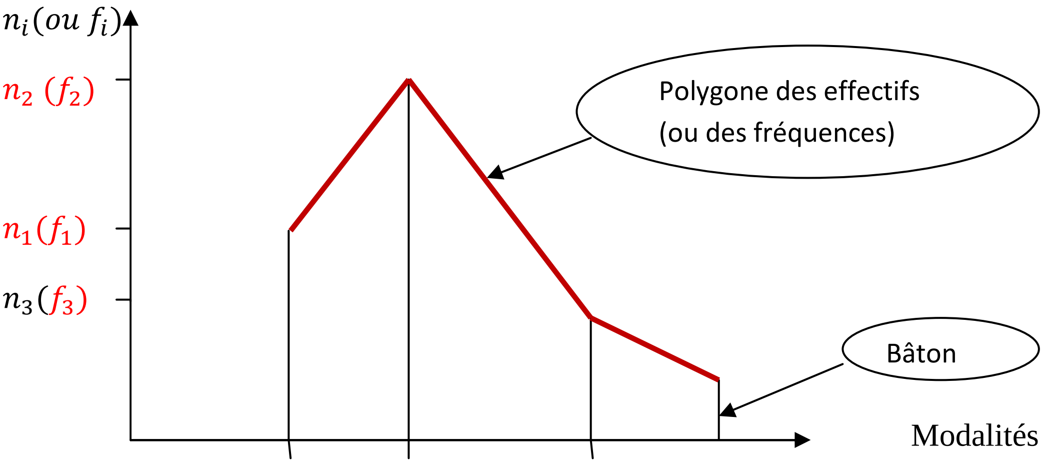
Il existe deux types de représentations graphiques : le diagramme différentiel et le diagramme cumulatif (courbe cumulative).

- Le diagramme différentiel met en évidence les différences d'effectifs (ou de fréquences) entre les différentes modalités ou classes.
- Le diagramme cumulatif permet de répondre aux questions du style combien d'individus ont pris une valeur inférieure (ou supérieure) à tant.

2.1. Caractère quantitatif discret

a. Diagramme en bâtons

C'est un graphique qui associe à chaque valeur du caractère un segment (bâton) dont la hauteur est proportionnelle à l'effectif ou à la fréquence. Dans un système d'axes orthogonaux, les valeurs isolées du caractère sont portées sur l'axe des abscisses et les effectifs (ou les fréquences) sont portés en ordonnées



b. Polygone des effectifs (ou de fréquences)

Il est obtenu en joignant les sommets des bâtons par une ligne brisée.

c. Diagramme cumulatif

Définition :

- On appelle fonction de répartition d'une variable X, et on note F, l'application définie sur $\mathbb{R}$ à valeurs dans l'intervalle$[0,1]$, qui pour une valeur donnée x, précise la proportion des individus pour lesquels la valeur de la variable est inférieure ou égale à x

$$F: \quad \mathbb{R} \rightarrow [0,1]$$
$$x \mapsto \quad F(x) = proportion\ des\ individus\ tq\ X \leq x.$$

- On appelle fonction cumulative des effectifs d'une variable X, et on note N, l'application définie sur $\mathbb{R}$ à valeurs dans $\{0,1,\dots,n\}$, qui pour une valeur donnée x, précise le nombre d'individus pour lesquels la valeur de la variable est inférieure ou égale à x.

$$\begin{array}{llc} N: & \mathbb{R} \to & \{0,1,\dots,n\} \\ & x \mapsto & N(x) = nombre\ d'individus\ tq\ X \leq x \end{array}$$

Si X est une variable quantitative discrète à valeurs dans $\{x_1, x_2, \dots, x_k\}$ alors F est donnée par

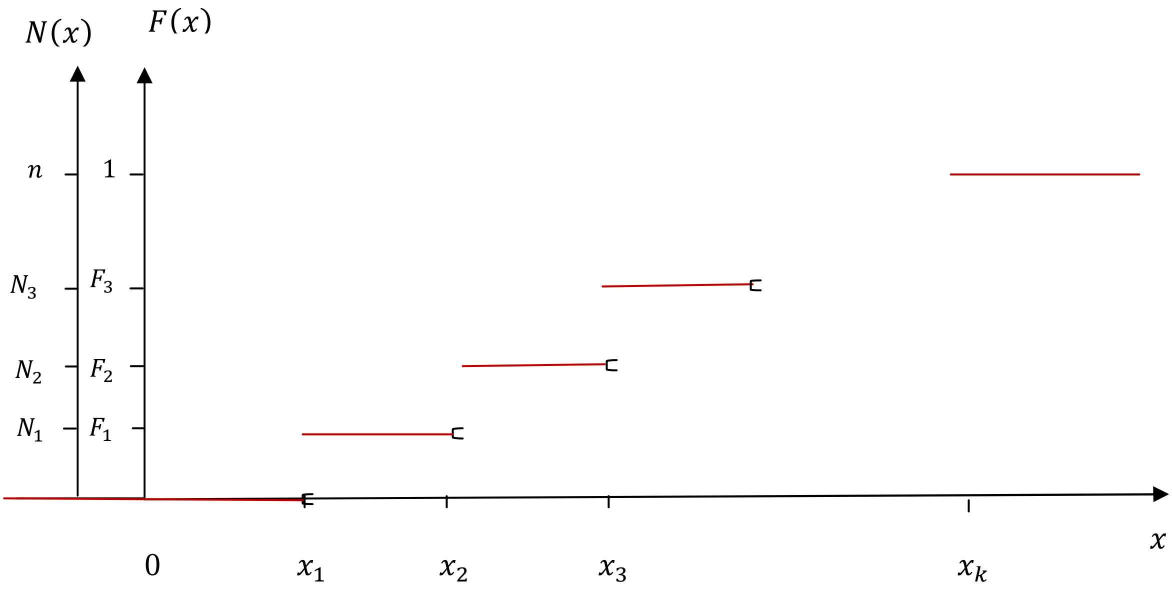
$$F(x) = \begin{cases} 0 & si\ \ x < x_1 \\ F_1 & si\ \ x_1 \leq x < x_2 \\ F_2 & si\ \ x_2 \leq x < x_3 \\ \vdots & \\ 1 & si\ x \geq x_k \end{cases}$$

$$F(x) = \begin{cases} 0 & si\ \ x < x_1 \\ & \\ F_i & si\ \ x_i \leq x < x_{i+1}, i = 1, \dots, k-1 \\ \vdots & \\ 1 & si\ x \geq x_k \end{cases}$$

et N est donnée par

$$N(x) = \begin{cases} 0 & si\ \ x < x_1 \\ N_1 & si\ \ x_1 \leq x < x_2 \\ N_2 & si\ \ x_2 \leq x < x_3 \\ \vdots & \\ n & si\ x \geq x_k \end{cases}$$

$$N(x) = \begin{cases} 0 & si\ \ x < x_1 \\ & \\ N_i & si\ \ x_i \leq x < x_{i+1}\,, i = 1, \dots, k-1 \\ \vdots & \\ n & si\ x \geq x_k. \end{cases}$$



Courbe cumulative

Exemple : Nombre d'absences

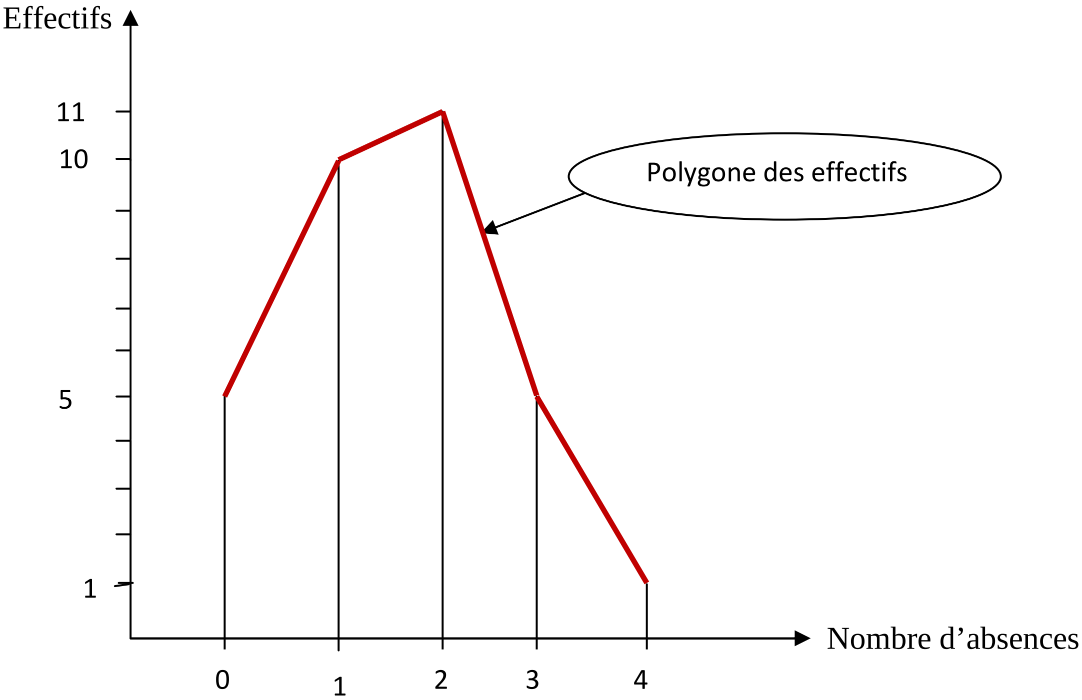
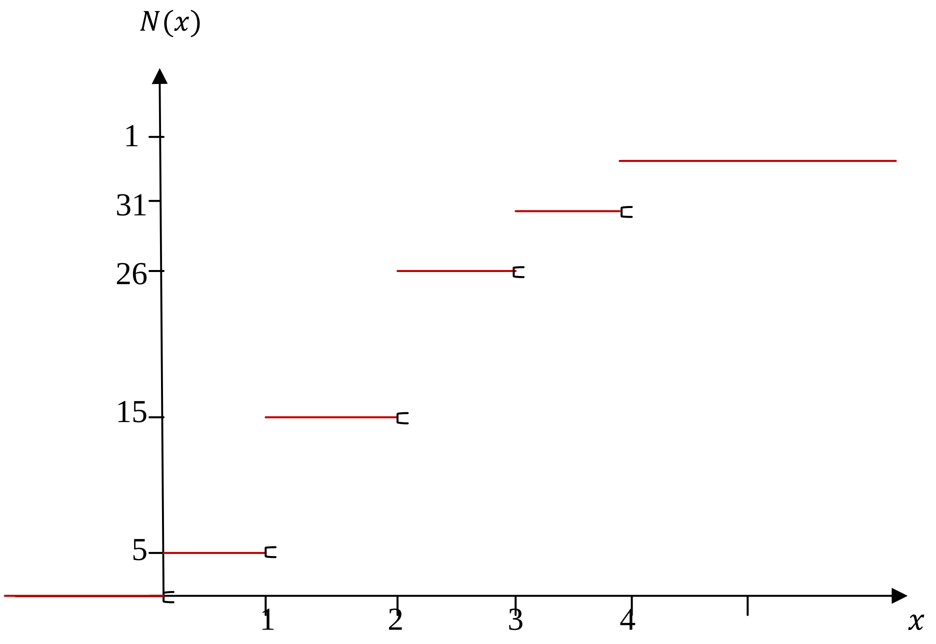


Diagramme en bâtons



Courbe cumulative du nombre d'absences par étudiant

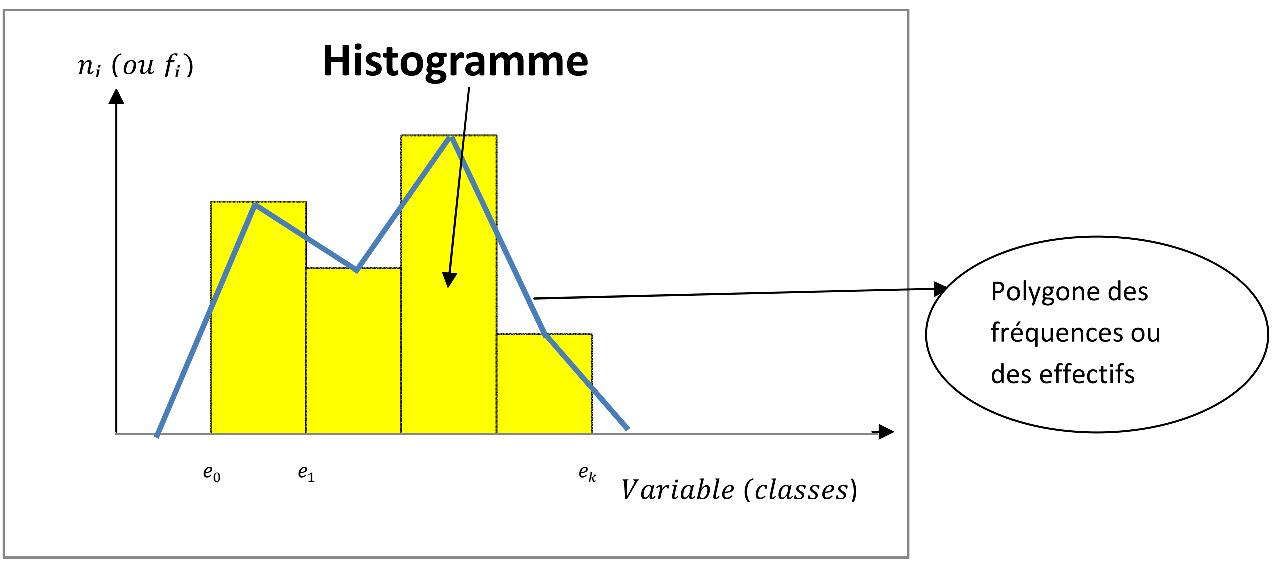
2.2. Caractère quantitatif continu :

On suppose que les données sont réparties en k classes $[e_0, e_1[\ , \ldots,[e_{k-1}, e_k[$ de même amplitude a.

a. Histogramme

On appelle histogramme un diagramme composé d'un ensemble de rectangles contigus d'aire proportionnelle aux effectifs (ou aux fréquences) et de bases déterminées par les extrémités de classes.

Dans un système d'axes orthogonaux, les valeurs de la variable sont portées sur l'axe des abscisses et les effectifs (ou fréquences) sont reportés sur l'axe des ordonnées. L'histogramme est la surface définie par l'ensemble des rectangles.

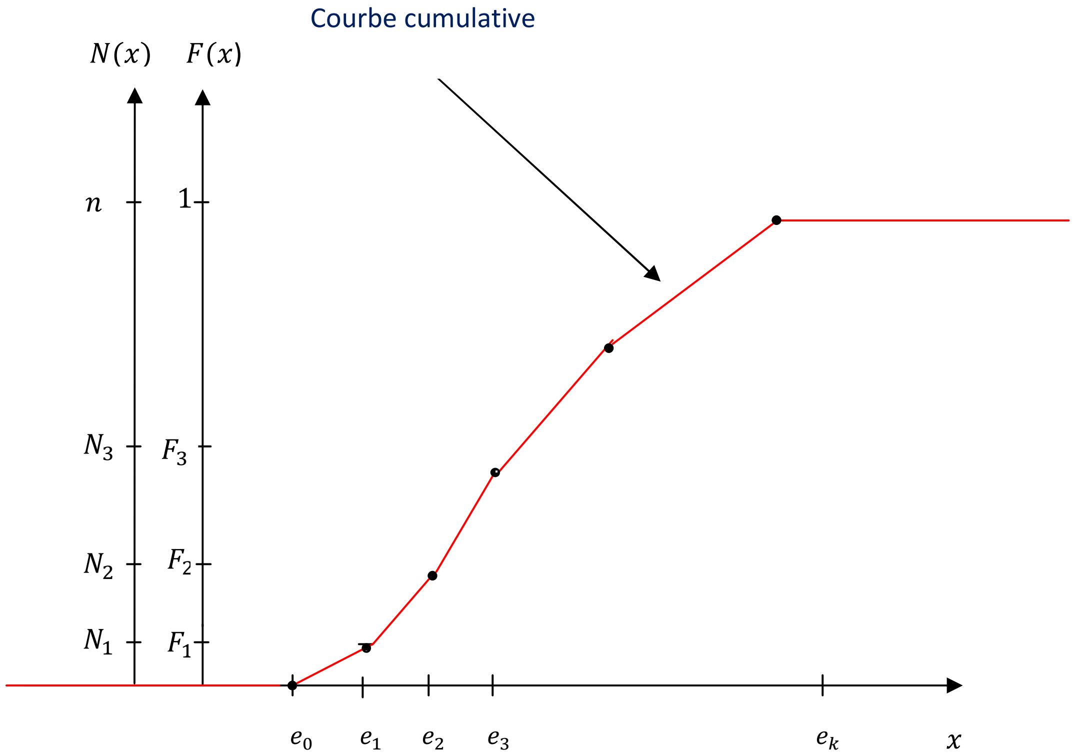


b. Polygone des effectifs (ou des fréquences)

On appelle polygone des fréquences (ou des effectifs) un polygone dont l'aire est égale à celle de l'histogramme et dont les côtés sont obtenus en joignant les milieux des segments supérieurs de chaque rectangle de l'histogramme. Pour tracer, on ajoute deux fausses classes (deux classes fictives) aux extrémités d'effectifs 0.

c. Courbe cumulative

La courbe des fréquences cumulées croissantes (ou des effectifs cumulés croissants) est obtenue en joignant, dans un système d'axes orthogonaux, les points de coordonnées (e_i, F_i) (ou les points de coordonnées (e_i, N_i)), $i = 0, \dots, k$, avec $F_0 = 0, F_k = 1$ $(N_0 = 0, N_k = n)$. On prolonge par une horizontale de niveau 1 pour les points d'abscisses supérieure à e_k(horizontale de niveau n) et par une horizontale de niveau 0 pour les points d'abscisses inférieure à e_0. On obtient ce qu'on appelle polygone des fréquences cumulées (ou effectifs cumulées



Joindre les points de coordonnées (e_i, F_i) par des segments revient à faire l'hypothèse d'une répartition uniforme des individus à l'intérieur des classes (on fait une interpolation linéaire). La fonction de répartition F est alors donnée par

$$F(x) = \begin{cases} 0 & si \;\; x < e_0 \\ F_{i-1} + \dfrac{x - e_{i-1}}{e_i - e_{i-1}} f_i & si \;\; e_{i-1} \leq x < e_i, \;\; i = 1, \dots, k \\ \vdots & \\ 1 & si \; x \geq e_k \end{cases}$$

En effet, sur $[e_{i-1},\; e_i[$, $F(x) = \alpha_i x + \beta_i$ avec $F(e_{i-1}) = F_{i-1}$ et $F(e_i) = F_i$. On obtient le système d'équations

$$\begin{cases} F_{i-1} = \alpha_i e_{i-1} + \beta_i, \\ F_i = \alpha_i e_i + \beta_i. \end{cases}$$

Ce qui donne

$$\alpha_i = \frac{F_i - F_{i-1}}{e_i - e_{i-1}} = \frac{f_i}{e_i - e_{i-1}},$$

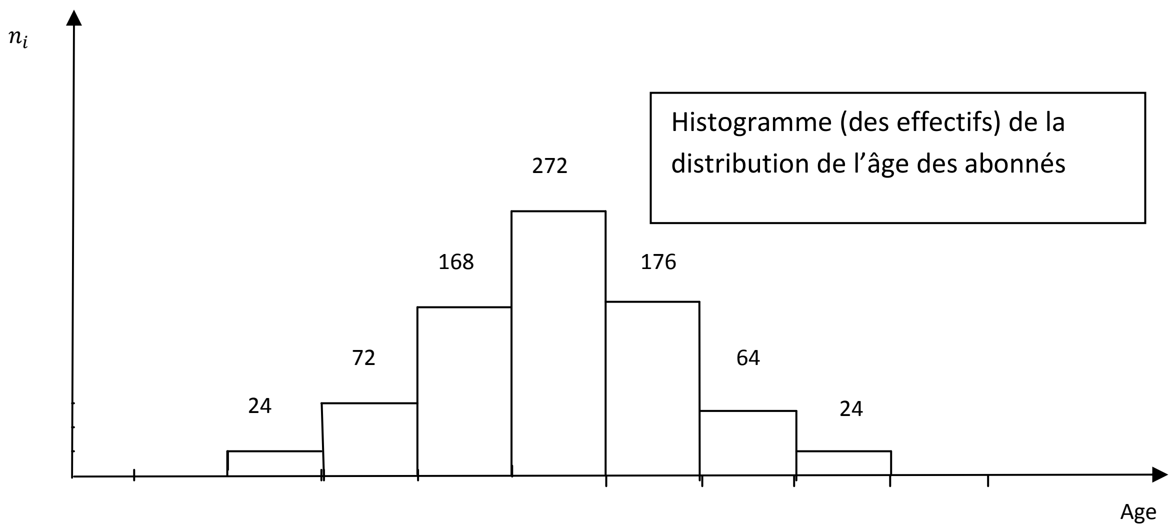
$$\beta_i = F_{i-1} - \alpha_i e_{i-1} = F_{i-1} - \frac{f_i}{e_i - e_{i-1}} e_{i-1}.$$

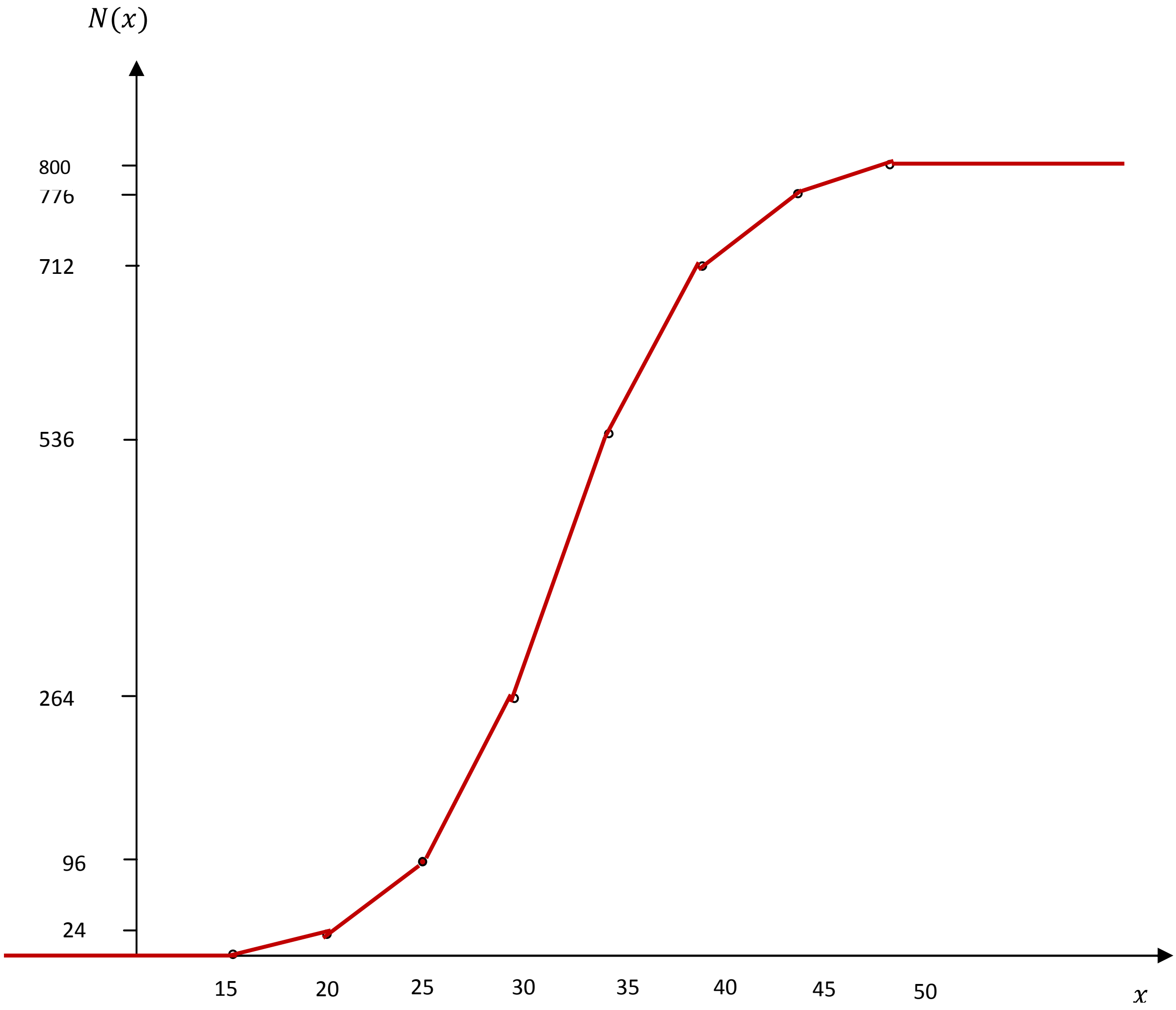
Par suite

$$F(x) = \frac{f_i}{e_i - e_{i-1}} x + F_{i-1} - \frac{f_i}{e_i - e_{i-1}} e_{i-1} = F_{i-1} + \frac{x - e_{i-1}}{e_i - e_{i-1}} f_i.$$

Exemple : Une entreprise d'édition a relevé l'âge X de ses abonnés aux différentes revues et a obtenu (après regroupement en classes) les données du tableau suivant :

Classe	[15,20[	[20,25[	[25,30[	[30,35[	[35,40[	[40,45[	[45,50[
n_i	24	72	168	272	176	64	24
f_i	0.03	0.09	0.21	0.34	0.22	0.08	0.03
N_i	24	96	264	536	712	776	800
F_i	0.03	0.12	0.33	0.67	0.89	0.97	1





Courbe cumulative des effectifs

Remarque : On peut également tracer (dans les cas discret et continu) la courbe des fréquences cumulées décroissantes (ou effectifs cumulés décroissants). Il suffit de remplacer F_i par F_i' (N_i par N_i').