La loi normale et la loi normale centrée réduite

1 Introduction

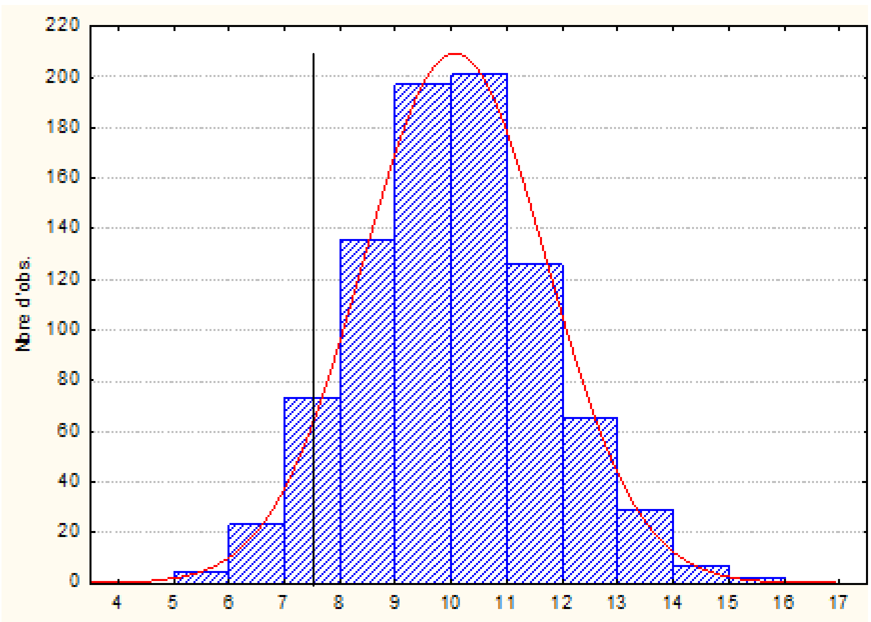
En statistique descriptive, nous avons vu comment construire un histogramme pour représenter la distribution d'une variable quantitative continue. La technique de construction d'un histogramme, repose sur le respect de la proportionnalité entre la surface des rectangles et la fréquence relative de la classe correspondante. Grace à cette proportionnalité, on peut calculer le pourcentage de données dans une classe à l'aide d'un rapport d'aires

$$\% \text{ de données d'une classe} = \frac{aire\ du\ réctangle\ de\ la\ classe}{aire\ totale\ de\ l'histogramme} \times 100$$

2 La courbe normale ou la courbe de Gauss

Chaque fois que l'on prend des mesures analogues (variable quantitative continue), on obtient une distribution ayant la forme d'une cloche. On donne le nom de la courbe Normale ; ou la courbe Gauss à ce type de graphique oű le pourcentage de données est élevé autour de la moyenne, et de plus en plus faible à mesure que l'on s'éloigne de celle-ci, dans un sens ou dans l'autre.

L'adjectif « Normale », s'explique par le fait que cette loi décrit et modélise des situations statistiques aléatoires, concrètes et naturelles.



On désigne une courbe normale par la notation $N(\mu, \sigma^2)$, oű μ représente la moyenne de la distribution, et σ l'écart-type de la distribution, l'equation de la courbe normale est comme suite :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Bien qu'il existe un très grand nombre de courbes normales, elles présentent toutes les caractéristiques suivantes :μ

1- La courbe sous forme de cloche parfaitement symétrique par rapport à la moyenne : le mode, la médiane et la moyenne ont la même valeur. Théoriquement, la courbe s'étend indéfiniment de chaque coté de la moyenne ;
2- L'aire totale comprise entre la courbe et l'axe des abscisses est égale à 1 ;
3- La surface entre la courbe et l'axe des x se répartit comme suit :
   - Environ 68% de la surface sous la courbe est comprise entre la moyenne moins 1 écart-type, et moyenne plus 1 écart-type soit entre les bornes de l'intervalle $[\mu - \sigma, \mu + \sigma[$, les données appartenant à cet intervalle ont une cote z comprise entre-1 et +1 ;
   - Environ 99.7% de la surface sous la courbe est comprise entre la moyenne moins 3 écart-type, et moyenne plus 3 écart-type soit entre les bornes de l'intervalle $\left[\mu - 3\sigma, \mu + 3\sigma\right[$, les données appartenant à cet intervalle ont une cote z comprise entre-3 et +3 ;

Par exemple pour une courbe normale $\mathrm{N}(50,100)$, qui possède les caractéristiques suivantes : une moyenne $\mu = 50$ et un écart-type $\sigma = 10$, on aura :

- 68.3% de la surface de la courbe est comprise entre 40 et 60
- 99.7% de la surface de la courbe est comprise entre 20 et 80

A l'aide de la loi normale, on peut déterminer des pourcentages qui n'apparaissent pas dans le tableau de distribution.

<u>Exemple</u>

Soit la distribution ci-dessous, qui représente la répartition de 300 feuille selon la teneur en métabolites

Teneur en métabolites	Nombre de feuilles	Surface de chaque rectangle
0.65-0.75	4	2S
0.75-0.85	18	9 S
0.85-0.95	42	21S
0.95-1.05	63	31.5S
1.05-1.15	75	37.5S
1.15-1.25	54	27S
1.25-1.35	30	15S
1.35-1.45	12	6S
1.45-1.55	2	1S
Total	300	150 S

Quel est le pourcentage de données dont la teneur en métabolites est supérieure ou égale à 1.35 ?

C'est-à-dire P(X≥1.35)

Puisque le polygone de fréquence de cette distribution possède la forme d'une cloche (symétrique), on utilise la loi normale comme modèle mathématique de cette distribution, les paramètres de cette distribution sont : La moyenne $\mu = 1.08$ µg et un écart-type $\sigma = 0.16$ µg

On va répondre à cette question par 2 méthodes différentes :

- 1ére methode par la surface du rectangle
- 2éme méthode par l'utilisation de la table de la loi normale centrée réduite

La méthode surface du rectangle

$$P(X\geq 1.35)=\frac{Aire\ du\ rectangle\ \ 1.35-1.55}{Aire\ totale\ de\ l'histogramme}$$

On prend comme unité de mesure la surface de la classe 1.45-1.55= 1S

$$P(X\geq 1.35)=\frac{6S+1S}{150S}=\frac{7S}{150S}=0.046=4.6\%$$

La méthode de la table de la loi normale centrée réduite

P(X≥1.35), cette expression signifie qu'on doit calculer l'aire sous la courbe de la loi normale centrée réduite supérieur à la cote z $>$ 1.35 :

P(X≥1.35)= P(z $\geq \frac{1.35-1.08}{0.16}$)

P(X≥1.35)= P(z ≥ 1.68)=0.0465 (lire dans la table Z, l'intersection entre 1.6 en colonne et 0.08 en ligne)