

## La loi normale

### Quelle est cette loi et à quoi sert-elle en botanique ?

La loi normale est une des principales lois de distributions de probabilité. Elle a été formulée par le mathématicien Abraham de Moivre en 1733. Elle a ensuite été plus amplement étudiée par Carl Friedrich Gauss, mathématicien, physicien et astronome allemand (1777-1855). Depuis, elle est utilisée dans de nombreux tests statistiques.

Quantités de données biologiques se distribuent selon des lois normales de paramètres spécifiques à chaque jeu de données. C'est le cas de la taille ou du poids d'enfants d'une classe d'âge donné, ou encore du poids des grains de blé dans un épi, ou de la taille des spores d'un lichen (Fig.1). Toutes ces lois normales ont en commun certaines propriétés. L'une d'entre elles est que 95% des données se trouvent dans un intervalle qui ne dépend que de la **moyenne** des données ( $\mu$ ) et de la **dispersion** de ces données autour de la moyenne (ou variance  $\sigma^2$ ). Dans le cas ci-dessous (Fig.1), 95% des spores ont une taille qui varie entre 15.6 et 28.2 micromètres. Les données observées à l'extérieur de cet intervalle sont considérées comme non représentatives de l'espèce avec un risque de se tromper de 5%.

Distribution de la longueur de spores chez un lichen *Buellia* sp.

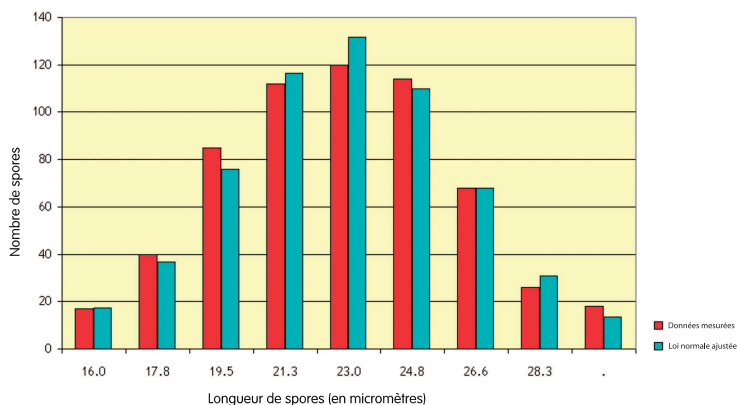


Figure 1. Distribution de la longueur de spores chez un lichen (*Buellia* sp.). Les données observées (en rouge) s'ajustent presque parfaitement à une loi normale (en bleu). Dans ce cas, l'on dit que les données observées suivent une loi normale de moyenne  $m = 21.9\text{mm}$  et de variance  $\sigma^2 = 10.24$ .

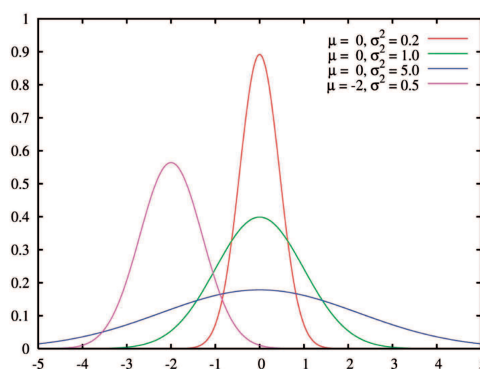


Figure 2. Densité de probabilité d'une loi normale, avec en axe horizontal la valeur prise par la variable aléatoire étudiée et en axe vertical la fréquence de cette valeur. L'une de ces distributions est très largement utilisée en statistiques : la loi normale centrée réduite (de moyenne  $\mu=0$  et de variance  $\sigma^2=1$ ; courbe verte).

La figure 2 donne l'illustration de plusieurs lois normales ayant des caractéristiques de moyenne ( $\mu$ ) et de dispersion ( $\sigma^2$  ou variance) différentes. On parle aussi de courbes de Gauss ou de courbes en cloche. Pour toutes ces lois, 95% des données se trouvent dans l'intervalle délimité par  $(\mu - 1.96\sigma)$  et  $(\mu + 1.96\sigma)$ . C'est cette formule, valable pour toutes les lois normales, qui a été utilisée dans l'exemple précédent.