

c) Le coefficient de variation c_v :

$$c_v = s / \bar{x} = 19,16 / 56,28 = 0,34$$

Il mesure la dispersion relative d'une série. C'est un nombre sans dimension.

C. BIBLIOGRAPHIE

- Spiegel, M.R. (1961) : *Statistics*, Schaum Publishing Company, New York.
- Dalrymple, T. (1962) : *Flood Frequency Analysis, Manual of Hydrology : Part 3, Flood - Flow Techniques*, United States Government Printing Office, Washington, D.C.
- Roche M. (1963) : *Hydrologie de Surface*, Gauthier-Villars ed. Paris.
- Parl, B. (1967) : *Basic Statistics*, Doubleday, New York.
- Riggs, H.C., (1968) : *Some Statistical Tools in Hydrology*, United States Government Printing Office, Washington, D.C.
- Pacé, P. et Cluzel R. (1969) : *Statistiques et Probabilités*, Librairie Delagrave, Paris.
- Viallet, F. (1970) : *Statistiques et Recherche Appliquée*, Chotard et Associés éd., Paris.
- Laborde, J.P. (1982) : *Eléments d'Hydrologie de Surface*, Institut Nationale Polytechnique de Lorraine, France.
- Sachs, L. (1884) : *Applied Statistics, a Handbook of Techniques*, Spring-Verlag Inc., New York.
- Baillargeon, G. (1990) : *Méthodes Statistiques de l'Ingénieur*, Les Editions S.M.G., Trois Rivières, Québec, Canada.

IV

LA LOI NORMALE

A - INTRODUCTION

Nous avons vu au chapitre III que l'ensemble des couples (x_i, n_i) ou (x_i, f_i) définit la fonction de distribution d'une variable, c'est-à-dire :

$$f_c(x_i) = n_i / N$$

et que la fonction de répartition est :

$$FND = F_c(x_i) = \sum_{j=1}^i f_c(x_j) \text{ ou } FD = F_1(x_i) = \sum_{j=1}^k f_c(x_j)$$

k est le nombre d'intervalles.

Ces fonctions f_c et F_c (e pour échantillon) sont définies pour un échantillon. Pour une population ces fonctions sont $f(x)$ et $F(x)$ et sont définies comme suit :

$$\begin{aligned} f(x) &= \lim_{N \rightarrow \infty} \left\{ f_c(x) / \Delta x \right\} & \text{et} & & F(x) &= \lim_{N \rightarrow \infty} F_c(x) \\ \Delta x &\rightarrow 0 \end{aligned}$$

$f(x)$ = fonction de densité de probabilité (f.d.p.),
et $F(x)$ = Fonction de probabilité (f.p.).

Pour que $f(x)$ soit une f.d.p., on doit avoir : $\int_{-\infty}^{+\infty} f(x) dx = 1$

Les fonctions $f(x)$ et son intégrale $F(x)$ sont des fonctions théoriques. Les mathématiciens ont étudié et développé un bon nombre de ces fonctions dont chacune peut s'appliquer à un ou plusieurs phénomènes concrets et permettre ainsi d'améliorer considérablement leur compréhension. Nous en aborderons brièvement quelques unes qui sont appliquées le plus fréquemment en hydrologie.

Il s'agit de:

- la loi normale ou loi de Laplace-Gauss,
- la loi log-normale ou loi de Galton,
- la loi exponentielle ou loi de Gumbel.

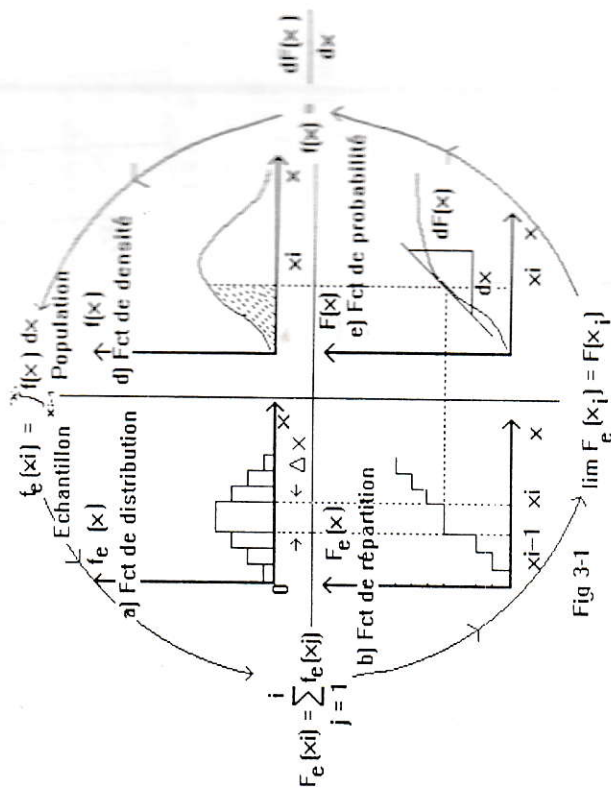


Figure IV-1 Relations entre les fonctions de distribution, de répartition, de densité et de probabilité d'une variable aléatoire

B - DEFINITION DE LA LOI NORMALE

OU LOI DE LAPLACE - GAUSS

L'expression de la fonction de densité de probabilité de la loi normale est:

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

Celle de la fonction de probabilité de la loi normale est:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

où μ est la moyenne et σ l'écart-type de la population étudiée.

En remplaçant x par $z = (x - \mu) / \sigma$ qui est la variable normale réduite, ou encore appelée variable centrée réduite, on a:

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

$$\text{et } F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{z^2}{2}} dz$$

Les valeurs de $F(x)$ sont fournies par les tables de l'intégrale de Gauss en fonction de la variable réduite z (annexe 1).

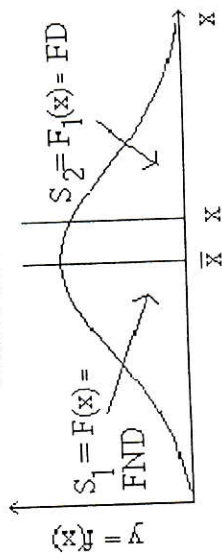
Supposons qu'un échantillon de hauteurs de pluies annuelles s'ajuste bien à une loi de Gauss. Pour chaque valeur de x_i de notre échantillon, nous pouvons calculer sa variable réduite z_i . La table de Gauss nous donne la probabilité de non-dépassement (PND ou FND) de z_i , c'est-à-dire la probabilité de ne pas dépasser ou d'égaliser z_i , qui est la même que la probabilité de non-dépassement de x_i . Notez que $F(x)$ est égale à la surface sous la courbe $f(x)$ qui va de $-\infty$ à x ; elle est donc égale à la probabilité d'avoir x_i inférieur ou égal à x , c'est-à-dire la probabilité d'égaliser ou de ne pas dépasser x . $F(x)$ est aussi appelée la fréquence au non-dépassement (FND) de x .

Notez qu'il existe des tables de l'intégrale de Gauss qui donnent la probabilité au dépassement $FD(x)$:

$$FD(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{z^2}{2}} dz$$

Donc $FD(x)$ est égale à la surface sous la courbe $f(x)$ qui va de x à $+\infty$; elle est donc égale à la probabilité d'avoir x_i supérieur ou égal à x , c'est-à-dire la probabilité d'égaliser ou de dépasser x .

LOI NORMALE



$F(x)$ = fréquence au non-dépassement = F.N.D.
 = probabilité au non-dépassement = P.N.D.
 $F_1(x)$ = fréquence au dépassement = F.D.
 = probabilité au dépassement = P.D.

Figure IV-2 Définition de la FD et de la FND

Noter que $FND(x) + FD(x) = 1$. Avant d'utiliser une table de Gauss, il y a lieu donc de s'assurer quelle surface elle donne.

La loi de Gauss offre une répartition symétrique de part et d'autre de la moyenne, qui est en même temps la médiane et le mode. Certaines tables ne donnent que la moitié de la distribution. On prend avantage de la symétrie pour avoir les valeurs de l'autre moitié.

C - LA DROITE DE HENRY:

L'équation de la variable réduite $z = (x - \mu) / \sigma$ peut s'écrire :

$$x = \mu + z\sigma.$$

Noter que l'on désigne respectivement par μ et σ la moyenne et l'écart-type de la population et par $\bar{x}$ et s la moyenne et l'écart-type de l'échantillon.

Cette équation est l'équation de Henry ; elle représente la courbe de Gauss sur le papier à probabilité normale (Figure IV-4).

Il existe deux types de papier de probabilité normale qui ne diffèrent que par l'échelle des abscisses:

- l'un a une échelle arithmétique sur laquelle on porte les valeurs des variables réduites z_i correspondant chacune à la variable naturelle x_i ,
- l'autre a une échelle gaussienne (ou de probabilité) sur laquelle on porte les fréquences expérimentales $f_i = (n_i - 0,5) / N$

(où n_i = rang de la variable, et N = nombre de valeurs dans la série) et les fréquences théoriques calculées ou lues sur la table de Gauss.

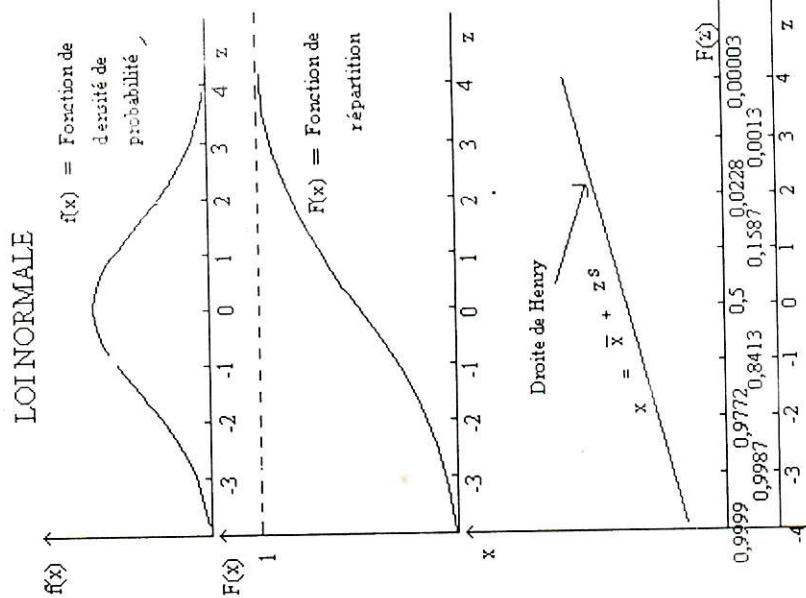


Figure IV-3 Relations entre $f(x)$, $F(x)$ et la droite de Henry

Evidemment, l'on porte sur l'axe des ordonnées les valeurs de x_i expérimentales ou calculées (théoriques).

A partir de l'équation de la droite de Henry l'on peut facilement calculer les valeurs que prendrait la variable analysée (hauteur de pluie annuelle, par exemple) pour une probabilité voulue.

On voudrait, par exemple, connaître la hauteur de la pluie décennale, c'est-à-dire la hauteur de pluie égale ou dépassée au moins une fois pendant dix ans. Cette période de dix ans est appelée période de retour T.

On calcule $\bar{x}$ et s à partir de l'échantillon de pluies annuelles observées en un lieu donné et l'on tire $x_{0.10}$ pour cet endroit, grâce à la formule ci-dessus.

La première table de Gauss donnée dans l'annexe la indique directement la FND, c'est-à-dire la surface sous la courbe de Gauss comprise entre $-\infty$ et z . La seconde table de Gauss de l'annexe lb donne la FD, c'est à dire la surface sous la courbe de Gauss comprise entre z et $+\infty$.

D - AJUSTEMENT D'UNE LOI NORMALE À UN ÉCHANTILLON

On se propose d'ajuster une loi de Gauss à un échantillon donné de pluie annuelles. Les étapes à suivre sont les suivantes:

1 - Calcul des caractéristiques empiriques:

- moyenne $\bar{x} = \frac{\sum x_i}{N}$ où x_i = valeurs de l'échantillon
et N = longueur de l'échantillon,

- variance $s^2 = \frac{\{\sum x_i^2 - N\bar{x}^2\}}{N-1}$;

- écart-type: $s = \sqrt{s^2}$

- coefficient de variation: $c_v = \frac{s}{\bar{x}}$

2 - Classement des valeurs :

On classe les valeurs de l'échantillon par ordre croissant ou décroissant en attribuant à chacune des valeurs son numéro d'ordre n_i compté à partir de 1.

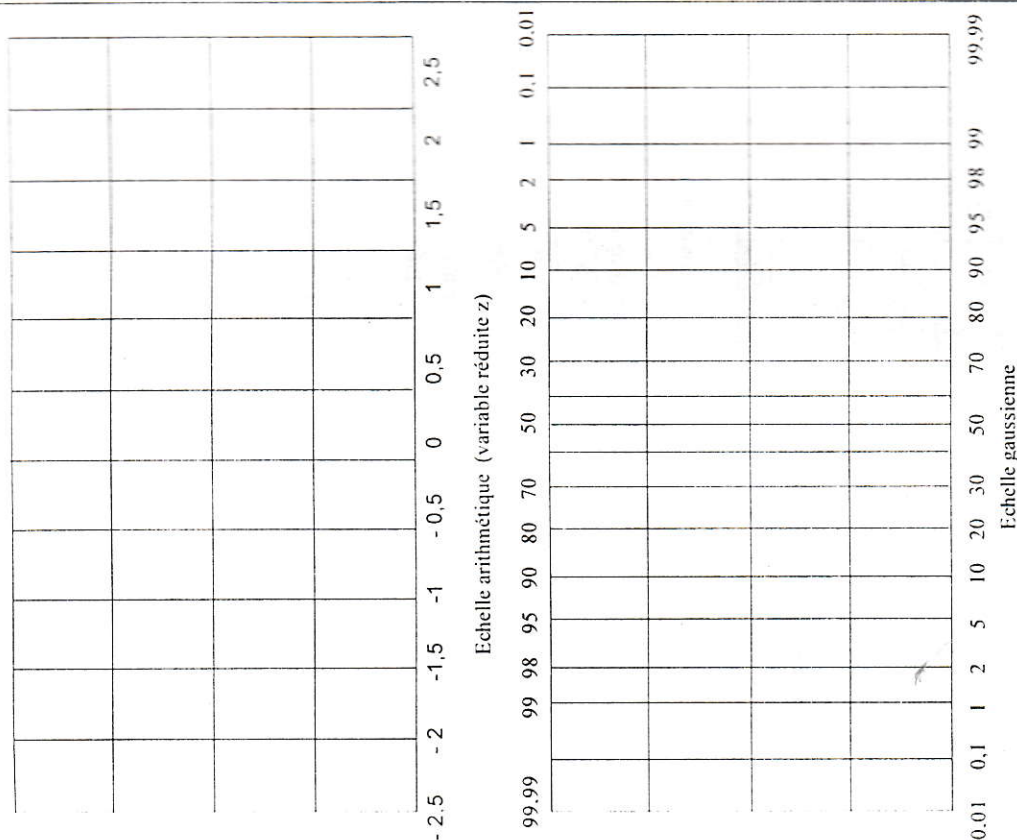


Figure IV - 4 Les deux types de papier de probabilité normale

On commence par chercher la fréquence au dépassement FD. Par définition, $FD = 1 / T = 1 / 10 = 0.1$; la table de Gauss donne, pour $FD = 0.1$, $z = 1.28$ (la table donne exactement $FD = 0.1003$). On en déduit que la pluie décennale est égale à :

$$x_{0.10} = \bar{x} + 1.28 s$$