

Décision et Prévision Statistiques

La loi normale

*L'objectif du chapitre est de présenter la loi normale qui est le modèle probabiliste le plus utilisé pour décrire de très nombreux phénomènes observés dans la pratique. Une grande attention devra être accordée aux concepts, essentiels en statistiques, d'**espérance mathématique** et de **variance** et aux opérations qui leurs sont attachées.*

1. Variables aléatoires continues

1.1. Loi de probabilité

Rappelons que si, à un ensemble de possibilités Ω , nous attachons un nombre X prenant les valeurs : $x_1, \dots, x_i, \dots, x_n$, lorsque se produit l'un des événements : $e_1, \dots, e_i, \dots, e_n$, on dit que X est une variable aléatoire. Cette variable est définie lorsqu'on connaît les probabilités :

$$p(x_1), \dots, p(x_i), \dots, p(x_n)$$

correspondant aux valeurs possibles de X . Ces probabilités sont obligatoirement telles que :

$$p(x_1) + \dots + p(x_i) + \dots + p(x_n) = 1,$$

et la correspondance $\{x_i, p(x_i)\}$ est appelée : *loi de probabilité* (ou *distribution de probabilité*) de la variable aléatoire X .

Si les valeurs possibles de X , sont réparties de façon continue sur un intervalle fini ou infini, X est une *variable aléatoire continue*. Une telle variable est définie si l'on connaît la probabilité pour que X prenne une valeur dans tout intervalle $[x, x + h[$. On se donne pour cela la *fonction de répartition* de X :

$$P(x) = \text{Prob}\{X < x\}$$

qui permet de calculer, pour tout intervalle :

$$\text{Prob}\{x \leq X < x + h\} = P(x + h) - P(x).$$

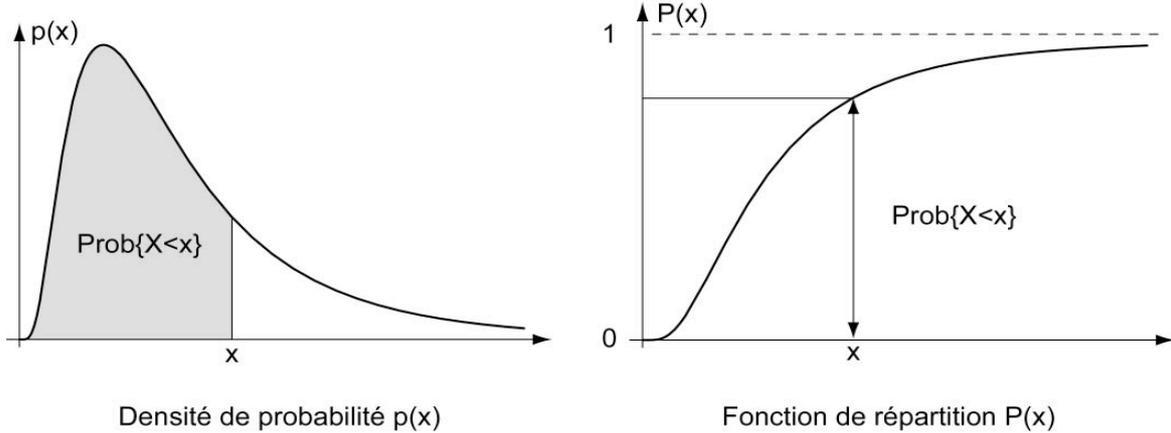
Un cas particulier important, auquel nous nous attacherons exclusivement dans ce qui suit, est celui où la fonction de répartition est continue, et peut être mise sous la forme :

$$P(x) = \int_{-\infty}^x p(u) du$$

où $p(x)$ s'appelle la *densité de probabilité* de X , appellation qui résulte du fait que :

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{\text{Prob}\{x \leq X < x + \Delta x\}}{\Delta x}$$

La densité de probabilité $p(x)$ ou la fonction de répartition $P(x)$ définissent la loi de probabilité d'une variable aléatoire continue X . Elles donnent lieu aux représentations graphiques suivantes :



(2.1)

Il est important de bien noter que, conformément aux axiomes qui définissent les probabilités :

$$\int_{-\infty}^{+\infty} p(x) dx = 1 \quad \text{et} \quad 0 \leq P(x) \leq 1$$

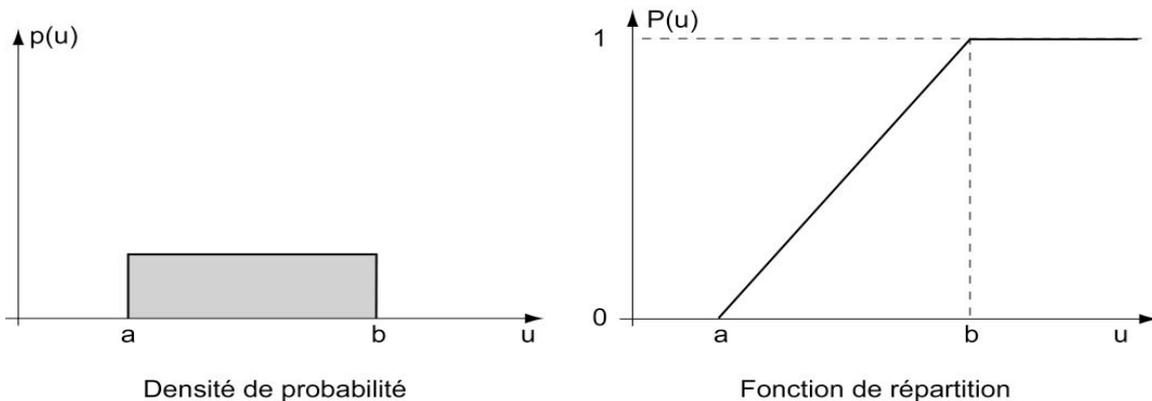
1.2. La loi uniforme

La variable aléatoire U est distribuée *uniformément* sur l'intervalle $[a, b]$ si sa densité de probabilité est constante sur cet intervalle :

$$p(u) = \frac{1}{b-a}$$

et si sa fonction de répartition a, par conséquent, l'équation suivante :

$$P(u) = \frac{u-a}{b-a}$$



(2.2)

1.3. La loi exponentielle

Nous avons montré dans le chapitre précédent que, si la probabilité d'apparition d'un événement pendant un intervalle de temps Δt était égale à $\lambda \Delta t$, la probabilité pour qu'il se produise k fois pendant un intervalle de temps t , était donnée par la loi de Poisson de paramètre λt :

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

Si, maintenant, on considère les intervalles de temps T qui s'écoulent entre les événements successifs d'un processus de Poisson, soit $F(t)$ la fonction de répartition de T . On a :

$$F(t) = 1 - \text{Prob} \{T > t\}.$$

Or cette dernière probabilité est égale à la probabilité pour qu'il ne se produise aucun évènement jusqu'à l'instant t , c'est-à-dire $p_0(t)$, et par conséquent :

$$F(t) = 1 - e^{-\lambda t}.$$

C'est la loi exponentielle qui peut constituer un modèle intéressant pour les durées de vie aléatoires de certains matériels (tubes électroniques) : à chaque instant t de la vie du matériel, la probabilité de défaillance pendant l'intervalle de temps Δt qui suit, est indépendante de t et égale à $\lambda \Delta t$.

1.4. Loi de probabilité à deux dimensions

Si, à un évènement aléatoire, sont attachés deux nombres X et Y , ces deux nombres définissent un *vecteur aléatoire à deux dimensions*. La loi de probabilité d'un tel vecteur peut être définie par la fonction de répartition :

$$P(x, y) = \text{Prob} \{X < x, Y < y\},$$

où la virgule se lit " *et* ". Si cette fonction est dérivable en x et y , on peut définir la densité de probabilité $p(x, y)$ qui est telle que :

$$p(x, y) dx dy = \text{Prob} \{x \leq X < x + dx, y \leq Y < y + dy\}.$$

Géométriquement $p(x, y) dx dy$ peut s'interpréter comme la probabilité pour que l'extrémité du vecteur aléatoire (X, Y) se trouve dans une aire $ds = dx dy$ autour du point (x, y) .

S'intéressant à la variable X , par exemple, indépendamment de la variable Y , on obtient la *distribution marginale* de X en calculant sa densité de probabilité $p_1(x)$:

$$p_1(x) dx = \text{Prob} \{x \leq X < x + dx\} = dx \int_{-\infty}^{+\infty} p(x, y) dy.$$

De même la distribution marginale de Y est définie par la densité de probabilité $p_2(y)$:

$$p_2(y) dy = \text{Prob} \{y \leq Y < y + dy\} = dy \int_{-\infty}^{+\infty} p(x, y) dx.$$

1.5. Indépendance de deux variables aléatoires

Par définition, deux variables aléatoires sont *indépendantes* si la probabilité pour que la valeur de l'une d'elle tombe dans un intervalle donné, ne dépend pas de la valeur prise par l'autre. La probabilité conditionnelle de X , par exemple, est donc indépendante de Y . Or elle s'écrit, avec les notations ci-dessus :

$$\text{Prob} \{x \leq X < x + dx / y \leq Y < y + dy\} = \frac{p(x, y) dx dy}{p_2(y) dy} = \frac{p(x, y) dx}{p_2(y)}.$$

Si les deux variables sont indépendantes, cette expression doit dépendre de x seulement et l'on doit donc pouvoir faire disparaître $p_2(y)$ par simplification. De la même façon, l'expression :

$$\text{Prob} \{y \leq Y < y + dy / x \leq X < x + dx\} = \frac{p(x, y) dy}{p_1(x)}.$$

doit dépendre de y seulement et l'on doit pouvoir la simplifier pour faire disparaître $p_1(y)$. Il en résulte que $p(x, y)$ doit être égal au produit des densités de probabilité marginales de X et de Y :

$$p(x, y) = p_1(x) p_2(y).$$

Cette condition est évidemment *nécessaire et suffisante*. Elle se généralise pour un nombre quelconque de variables aléatoires indépendantes dans leur ensemble.

1.6. Fonctions de variables aléatoires

Etant données une variable aléatoire X , définie par sa densité de probabilité $p(x)$, et une fonction f , la variable aléatoire $f(X)$, *fonction de la variable aléatoire X* , est définie de la façon suivante : $f(X)$ prend la valeur $f(x)$, lorsque X prend la valeur x .

On remarquera qu'une variable $f(X)$ peut prendre la même valeur pour deux valeurs différentes x_i et x_j de X . La probabilité de la valeur $f(x_i)$ ou $f(x_j)$ est alors égale à $p(x_i) + p(x_j)$.

On peut aussi définir des fonctions de plusieurs variables aléatoires. La *somme de deux variables aléatoires*, notamment, se définit de la façon suivante : étant donné deux variables aléatoires X et Y , leur somme est la variable aléatoire $(X + Y)$ qui prend la valeur $(x + y)$ lorsque X prend la valeur x et Y prend la valeur y . Là encore, une même valeur de la somme peut être obtenue pour deux couples différents de valeurs de X et Y . Penser, par exemple à la variable : on jette deux dés et on en fait la somme.

2. Espérance mathématique

2.1. Espérance et moments d'une variable aléatoire

Etant données une variable aléatoire X définie par sa densité de probabilité $p(x)$ et une fonction f , on désigne par le terme d'espérance mathématique de la variable aléatoire $f(X)$, et on la note $E[f(X)]$, l'expression :

$$E[f(X)] = \int_{-\infty}^{+\infty} f(x) p(x) dx$$

C'est donc un *opérateur* qui transforme la variable aléatoire $f(X)$ en un nombre. Appliqué à la variable X elle-même, l'opérateur donne sa moyenne μ :

$$\mu = E(X) = \int_{-\infty}^{+\infty} x p(x) dx.$$

Dans le cas d'une variable aléatoire discrète à valeurs positives ou nulles : $x_0, \dots, x_i, \dots, x_n$, l'expression précédente devient :

$$\mu = E[f(X)] = \sum_{i=0}^n f(x_i) p(x_i).$$

et toutes les propriétés de l'opérateur E que nous démontrerons par la suite, pour une variable continue, s'étendent sans difficulté au cas d'une variable discrète.

Lorsque $f(X)$ est une puissance de X , l'expression $E(X^k)$ est appelée *moment* d'ordre k de la variable aléatoire X . La moyenne $\mu = E(X)$ est ainsi le moment d'ordre 1 de la variable X . Elle s'interprète comme l'abscisse du centre de gravité de la distribution de probabilité et c'est, à ce titre,

une caractéristique de tendance centrale de la distribution : les valeurs d'une variable aléatoire se répartissent autour de sa moyenne.

2.2. Variance et écart-type

Il peut alors s'avérer intéressant de rapporter une variable aléatoire à sa moyenne, autrement dit de la *centrer*. On obtient alors le *moment centré* d'ordre k :

$$E[(X - \mu)^k] = \int_{-\infty}^{+\infty} (x - \mu)^k p(x) dx.$$

Le moment centré d'ordre 2 est appelé *variance* et il revêt une importance toute particulière. Nous le noterons $\sigma^2(X)$, ou très souvent σ^2 :

$$\sigma^2(X) = E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 p(x) dx.$$

Par analogie avec la mécanique, on peut dire que μ est le centre de gravité et que σ^2 est le moment d'inertie de la distribution.

Sa racine carrée σ est appelée l'*écart-type*, qui s'interprète comme une mesure de la dispersion d'une variable aléatoire : plus les valeurs de la variable sont susceptibles de s'éloigner de sa moyenne, plus son écart-type est grand. C'est ce que montre l'inégalité de Bienaymé-Tchebichef.

2.3. Inégalité de Bienaymé-Tchebichef

Soit X une variable aléatoire de moyenne μ et d'écart-type σ , mais à ceci près quelconque. La probabilité pour qu'elle prenne une valeur à l'extérieur d'un intervalle $[\mu - a, \mu + a]$, où a est un nombre positif, est donnée par l'intégrale :

$$\text{Prob} \{ |X - \mu| > a \} = \int_{|X - \mu| > a} p(x) dx.$$

Pour les valeurs x extérieures à l'intervalle $[\mu - a, \mu + a]$, $(x - \mu)^2$ est supérieur à a^2 , et l'on a, par conséquent :

$$\text{Prob} \{ |X - \mu| > a \} < \frac{1}{a^2} \int_{|X - \mu| > a} (x - \mu)^2 p(x) dx.$$

On majore encore l'intégrale en intégrant, pas seulement pour $|x - \mu| > a$, mais de $-\infty$ à $+\infty$. On obtient ainsi :

$$\text{Prob} \{ |X - \mu| > a \} < \frac{1}{a^2} \int_{-\infty}^{+\infty} (x - \mu)^2 p(x) dx,$$

c'est à dire :

$$\text{Prob} \{ |X - \mu| > a \} < \frac{\sigma^2}{a^2}.$$

C'est l'inégalité de Bienaymé-Tchebichef qui s'interprète ainsi : plus σ est petit, plus les grandes valeurs de $|X - \mu|$ sont improbables, donc moins la variable X est dispersée autour de sa moyenne μ .

Si l'on pose $a = t\sigma$, l'inégalité s'écrit :

$$\text{Prob} \{ |X - \mu| > t\sigma \} < \frac{1}{t^2}.$$

et exprime que la probabilité pour que X s'éloigne de sa moyenne de plus que t fois son écart-type σ , est inférieure à $\frac{1}{t^2}$. Il y a, en particulier, moins de 1 chance sur 4 pour que X

s'éloigne de μ de plus de 2 fois σ , et il est très improbable (probabilité inférieure à $\frac{1}{100}$) que X s'éloigne de μ de plus de 10 fois σ .

Ces résultats sont très généraux puisqu'ils ne nécessitent aucune hypothèse sur la forme de la loi de probabilité de X . On peut, bien sûr, les rendre beaucoup plus précis et les serrer davantage dans chaque cas particulier où la loi de probabilité de X est connue.

2.4. Linéarité de l'opérateur Espérance

On montre très facilement que, a et b étant deux constantes, on a :

$$E(aX + b) = aE(X) + b.$$

Soient maintenant deux variables X et Y définies par leur densité de probabilité $p(x, y)$. L'espérance mathématique de leur somme s'écrit :

$$E(X + Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) p(x, y) dx dy$$

$$E(X + Y) = \int_{-\infty}^{+\infty} x dx \int_{-\infty}^{+\infty} p(x, y) dy + \int_{-\infty}^{+\infty} y dy \int_{-\infty}^{+\infty} p(x, y) dx$$

$$E(X + Y) = \int_{-\infty}^{+\infty} x p_1(x) dx + \int_{-\infty}^{+\infty} y p_2(y) dy$$

$$E(X + Y) = E(X) + E(Y)$$

Cette propriété s'applique, bien entendu, à la somme d'un nombre quelconque de variables.

On notera, d'autre part, que la démonstration n'a requis aucune hypothèse sur l'indépendance des variables aléatoires considérées.

2.5. Calculs de variances de variables aléatoires

Rappelons qu'étant donnée une variable aléatoire de moyenne μ , sa variance s'écrit :

$$\sigma^2(X) = E[(X - \mu)^2].$$

De la forme même de la variance, il résulte que :

- la variance d'une constante est égale à 0,
- dans un changement d'échelle de rapport k , la variance est multipliée par k^2 .

Si, d'autre part, on développe l'expression de la variance, et qu'on utilise les propriétés de linéarité de l'espérance, il vient :

$$\sigma^2(X) = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2$$

et, puisque $E(X) = \mu$:

$$\sigma^2(X) = E(X^2) - \mu^2$$

Cette relation sera très souvent utilisée. On la retiendra facilement en écrivant que :

$$\sigma^2(X) = E(X^2) - E(X)^2$$

et en énonçant que : *la variance est égale à l'espérance du carré moins le carré de l'espérance*. Appliquons la à quelques-unes des variables aléatoires déjà rencontrées.

2.5.1. Loi uniforme

Soit la variable U distribuée uniformément sur l'intervalle $[0, a]$. Elle a pour densité de probabilité $\frac{1}{a}$, donc pour moyenne :

$$E(U) = \int_0^a x \frac{1}{a} dx = \frac{a}{2},$$

pour moment d'ordre 2 :

$$E(U^2) = \int_0^a x^2 \frac{1}{a} dx = \frac{a^2}{3},$$

et, enfin, pour variance :

$$\sigma^2(U) = E(U^2) - E(U)^2 = \frac{a^2}{12}$$

2.5.2. Variable de Bernouilli

Les moments d'ordres 1 et 2 sont égaux :

$$E(X) = E(X^2) = 1 \times \varpi + 0 \times (1 - \varpi) = \varpi.$$

La variance est donc égale à :

$$\sigma^2(X) = E(X^2) - E(X)^2 = \varpi - \varpi^2 = \varpi(1 - \varpi).$$

2.5.3. Loi de Poisson

La loi de Poisson est définie par la densité de probabilité :

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!},$$

et nous avons déjà calculé sa moyenne :

$$E(K) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

Pour trouver sa variance, calculons d'abord :

$$E[K(K-1)] = \sum_{k=0}^{\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2.$$

On a, par conséquent :

$$E[K(K-1)] = E(K^2) - E(K) = \lambda^2,$$

d'où l'on déduit que :

$$E(K^2) = \lambda^2 + \lambda$$

et, enfin, que :

$$\sigma^2(K) = E(K^2) - E(K)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

La loi de Poisson est donc telle que moyenne et variance soient toutes les deux égales à λ .

2.5.4. Loi exponentielle

La densité de probabilité est :

$$p(t) = \lambda e^{-\lambda t}.$$

En intégrant par parties, on trouve pour la moyenne :

$$E(T) = \int_0^{\infty} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda},$$

puis, pour le moment d'ordre 2 :

$$E(T^2) = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = \frac{2}{\lambda^2},$$

ce qui montre que l'écart-type est égal à la moyenne :

$$\sigma^2(T) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

2.6. Indépendance et covariance de deux variables aléatoires

Soient X et Y deux variables aléatoires définies par leur densité de probabilité $p(x, y)$. L'espérance mathématique du produit de ces deux variables est, par définition :

$$E(X Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x y p(x, y) dx dy.$$

Si les deux variables aléatoires sont *indépendantes* :

$$p(x, y) = p_1(x) p_2(y)$$

et, par suite :

$$E(X Y) = \int_{-\infty}^{+\infty} x p_1(x) dx \int_{-\infty}^{+\infty} y p_2(y) dy = E(X) E(Y).$$

Il s'agit là d'un théorème très important qui peut s'énoncer de la façon suivante : *si deux variables aléatoires sont indépendantes, l'espérance de leur produit est égale au produit de leurs espérances*.

Ce théorème justifie la définition d'une quantité appelée la *covariance* de X et Y :

$$\sigma(X, Y) = E(X Y) - E(X) E(Y)$$

qui a la propriété suivante : *la covariance de deux variables aléatoires indépendantes est nulle*.

Il faut faire attention au fait que la réciproque n'est pas vraie, en général.

2.7. Variance d'une somme (ou d'une différence) de variables aléatoires indépendantes

Soient X et Y deux variables aléatoires. La variance de leur somme (ou de leur différence) peut s'écrire (espérance du carré moins carré de l'espérance) :

$$\sigma^2(X \pm Y) = E[(X \pm Y)^2] - E(X \pm Y)^2.$$

Le premier terme du second membre se développe en :

$$E[(X \pm Y)^2] = E(X^2) \pm 2 E(X Y) + E(Y^2)$$

et le second terme en :

$$E(X \pm Y)^2 = E(X)^2 \pm 2 E(X) E(Y) + E(Y)^2.$$

On a donc, en réorganisant les termes :

$$\sigma^2(X \pm Y) = (E(X^2) - E(X)^2) \pm 2(E(XY) - E(X)E(Y)) + (E(Y^2) - E(Y)^2),$$

où l'on reconnaît :

$$\sigma^2(X \pm Y) = \sigma^2(X) \pm 2\sigma(X, Y) + \sigma^2(Y).$$

Si, maintenant, les deux variables X et Y sont *indépendantes*, leur covariance est nulle, et :

$$\sigma^2(X \pm Y) = \sigma^2(X) + \sigma^2(Y),$$

propriété qui s'énonce ainsi : la variance de la somme ou de la différence de deux variables aléatoires indépendantes est égale à la *somme* de leurs variances.

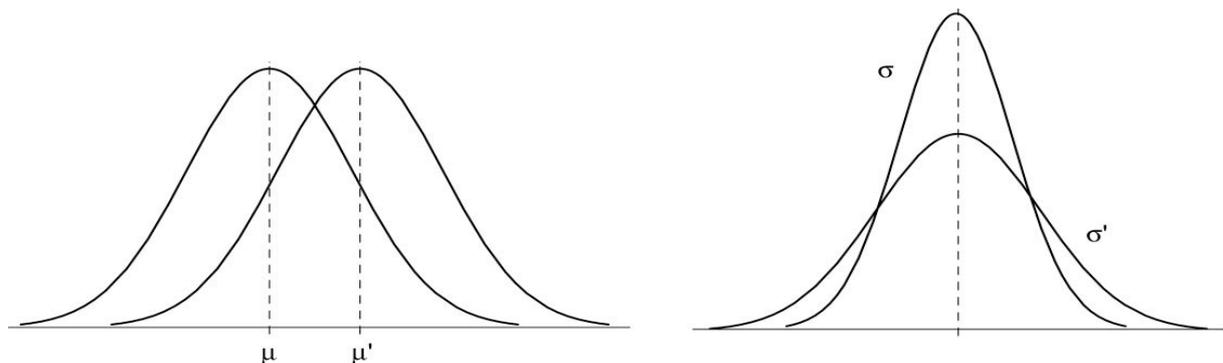
3. Loi normale

3.1. Définition et propriétés

La loi de probabilité la plus utilisée en statistique est la loi normale, encore appelée loi de Gauss, ou de Laplace-Gauss. Une variable aléatoire X suit une loi normale si sa densité de probabilité a pour équation :

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

dont le graphe est la fameuse "courbe en cloche", et qui dépend de deux paramètres μ et σ qui ne sont autres, comme on le montrera au paragraphe suivant, que la *moyenne* et l'*écart-type* de X .



(2.3)

La distribution est symétrique par rapport à μ qui caractérise donc la tendance centrale. Quant à σ , il caractérise la dispersion de la distribution. Plus il est grand, plus la distribution est étalée de part et d'autre de μ . Les points d'inflexion se trouvent à $(\mu - \sigma)$ et $(\mu + \sigma)$, avec :

$$\text{Prob}\{|X - \mu| < \sigma\} = 68,26\%.$$

Il sera bon de conserver en mémoire deux autres valeurs intéressantes :

$$\text{Prob}\{|X - \mu| > 1,96\sigma\} = 5\%,$$

$$\text{Prob}\{|X - \mu| > 2,58\sigma\} = 1\%,$$

que l'on pourra aussi comparer à celles que donnait l'approximation de Bienaymé-Tchebichef pour une loi de probabilité quelconque.

3.2. Calcul de la moyenne et de la variance

Nous allons montrer que la moyenne d'une variable qui suit une loi normale est égale à μ , et que sa variance est égale à σ^2 . La démonstration est uniquement calculatoire et pourra être omise.

La moyenne est, par définition, égale à :

$$E(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

Pour calculer cette intégrale, faisons le changement de variable, classique pour les calculs sur la loi normale :

$$u = \frac{x-\mu}{\sigma}.$$

Il vient :

$$E(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\mu + \sigma u) e^{-\frac{u^2}{2}} du,$$

$$E(X) = \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u e^{-\frac{u^2}{2}} du.$$

La seconde intégrale est nulle. Quant à la première, elle est égale à $\sqrt{2\pi}$, ce qui se montre en posant :

$$I = \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du,$$

d'où :

$$I^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy,$$

que l'on intègre facilement en passant en coordonnées polaires. D'où finalement :

$$E(X) = \mu.$$

De la même façon :

$$E[(X - \mu)^2] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x - \mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx,$$

qui devient, après changement de variable :

$$E[(X - \mu)^2] = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}} du.$$

L'intégrale se calcule par parties, et il vient :

$$E[(X - \mu)^2] = \sigma^2.$$

3.3. La loi normale réduite

Etant donnée une variable suivant une loi normale de moyenne μ et d'écart-type σ , sa densité de probabilité est telle que :

$$p(x) dx = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

Faisant le changement de variable : $U = \frac{X-\mu}{\sigma}$, on peut écrire : $du = \frac{dx}{\sigma}$, d'où :

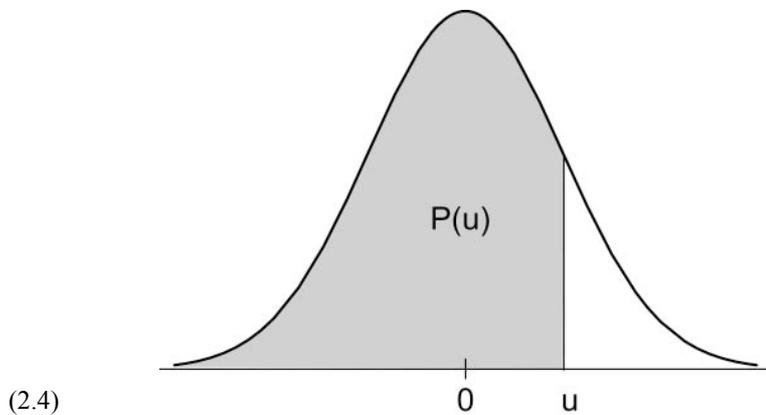
$$p(u) du = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$$

La nouvelle variable suit donc une loi normale qui est dite *centrée* ($E(U) = 0$) et *réduite* ($\sigma^2(U) = 1$), et qui ne fait intervenir aucun paramètre.

Cette loi, appelée *loi normale réduite*, est tabulée. Les tables donnent généralement, pour différentes valeurs de u , les probabilités :

$$\text{Prob} \{-\infty < U < u\} = P(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{t^2}{2}} dt,$$

qui, compte tenu de la symétrie de la loi et de l'additivité des probabilités pour des intervalles disjoints, permettent de connaître la probabilité attachée à n'importe quel intervalle.



On en déduit facilement la probabilité pour qu'une variable suivant une loi normale quelconque $X(\mu, \sigma)$ de moyenne μ et d'écart-type σ , soit comprise dans un intervalle donné $[x_1, x_2]$:

$$\text{Prob} \{x_1 < X < x_2\} = \text{Prob} \left\{ \frac{x_1 - \mu}{\sigma} < U < \frac{x_2 - \mu}{\sigma} \right\}.$$

3.4. Fonctions linéaires de variables normales

Etant données deux variables X_1 et X_2 indépendantes et suivant toutes les deux des lois normales, et deux constantes a_1 et a_2 , on peut montrer que la fonction linéaire $(a_1 X_1 + a_2 X_2)$ suit elle-même une *loi normale*. Pour spécifier cette loi, il faut connaître sa moyenne et sa variance. Les propriétés générales (c'est-à-dire indépendantes des lois de probabilité en jeu) des opérateurs espérance et variance permettent d'établir que :

$$E(a_1 X_1 + a_2 X_2) = a_1 E(X_1) + a_2 E(X_2),$$

et que :

$$\sigma^2(a_1 X_1 + a_2 X_2) = a_1^2 \sigma^2(X_1) + a_2^2 \sigma^2(X_2).$$

Cette propriété s'étend évidemment à une fonction linéaire d'un nombre *quelconque* de variables indépendantes suivant des lois normales.

3.5. Théorème central limite

Ce théorème établit une propriété encore plus générale que la précédente, et qui va justifier l'importance considérable de la loi normale, à la fois comme *modèle* pour décrire des situations pratiques, mais aussi comme *outil* théorique. Il s'énonce ainsi :

Soit $X_1, \dots, X_i, \dots, X_n$, une suite de n variables aléatoires *indépendantes*, de moyennes $\mu_1, \dots, \mu_i, \dots, \mu_n$, et de variances $\sigma_1^2, \dots, \sigma_i^2, \dots, \sigma_n^2$, et de *lois de probabilité quelconques*, leur somme suit une loi qui, lorsque n augmente, tend vers une *loi normale* de moyenne $\mu = \sum_{i=1}^n \mu_i$ et de variance $\sigma^2 = \sum_{i=1}^n \sigma_i^2$. Il y a une seule condition restrictive, c'est que les variances soient finies et qu'aucune ne soit prépondérante devant les autres.

3.5.1. La loi normale comme modèle

Prenons l'exemple du fonctionnement d'un tour. Le réglage du tour a pour but d'obtenir des pièces présentant une cote bien définie ; mais on sait que de multiples causes perturbatrices agissent au cours de l'usinage d'une pièce : vibrations, usures, variations de courant ... Or si les causes perturbatrices sont nombreuses, si leurs effets interviennent de façon *additive*, enfin si la dispersion provoquée par chacune d'elles reste faible par rapport à la dispersion totale, alors le théorème central limite signifie qu'on doit observer une fluctuation globale très voisine de la loi normale. Et, comme ce mécanisme d'intervention de causes perturbatrices est très répandu dans la nature, il en résulte que la loi normale occupe en statistique une place privilégiée.

3.5.2. Limite de la loi binomiale

On a défini une variable de Bernoulli comme une variable qui prend la valeur 1 avec la probabilité ϖ , et la valeur 0 avec la probabilité $(1 - \varpi)$, et montré que sa moyenne est égale à ϖ et sa variance à $\varpi(1 - \varpi)$. Or on peut considérer une variable binomiale comme la somme K_n de n variables de Bernoulli. Il résulte du théorème central limite que, si n est suffisamment grand (en pratique à partir de $n = 50$), la loi binomiale peut être approximée par une *loi normale* de moyenne $n\varpi$ et de variance $n\varpi(1 - \varpi)$. C'est pourquoi les tables de la loi binomiale s'arrêtent généralement à $n = 50$.

3.6. La loi log-normale

On appelle *loi log-normale* une loi qu'une transformation du type $\log(X - x_0)$ ramène à une loi normale. Cette loi est très répandue, notamment dans le domaine des sciences naturelles (géologie, biologie, ...) et des sciences humaines (psychologie, économie, ...). Chaque fois que les causes perturbatrices à l'origine des fluctuations observées répondent aux conditions du théorème central limite mais sont *multiplicatives*, on enregistre des distributions modélisables par des lois log-normales.

Exercices du chapitre 2

Exercice 1

Le profilé laminé à partir d'un lingot est découpé en billettes de 8 mètres de longueur. L'extrémité du profilé correspondant au laminage de la tête du lingot présente un défaut sur une certaine longueur X qui suit une loi normale de moyenne égale à 15 mètres et d'écart-type égal à 5 mètres.

- Pour tenter d'éliminer la longueur atteinte, on décline systématiquement les trois billettes de tête. Quel est le risque pour que la quatrième billette présente encore un défaut ?
- Calculer le nombre de billettes à déclasser pour que la première billette retenue soit propre avec un risque inférieur à 1 %.

Exercice 2

On usine des axes sur un tour automatique. L'intervalle de tolérance sur le diamètre de ces axes est (355, 365) en centièmes. Les diamètres des axes usinés sur la machine sont distribués suivant une loi normale de moyenne correspondant au réglage de la machine et d'écart-type 2 centièmes. Montrer que le meilleur réglage est 360. Quelle est alors la proportion de déchets ? Quelle est la proportion de déchets si la machine se dérègle et que la moyenne devient égale à 358 ?

Exercice 3

La durée de vie d'un certain type de lampes à incandescence suit une loi normale de moyenne égale à 160 heures. Les spécifications impliquent que 80 % au moins de la production tombe entre 120 et 200 heures. Quel est le plus grand écart-type acceptable pour que la fabrication réponde aux spécifications ?

Exercice 4

Il existe deux méthodes pour doser le phosphore dans l'acier d'une coulée :

- méthode C : dosage par voie chimique,
- méthode S : dosage par méthode spectrographique.

Pour toutes les coulées effectuées pendant une certaine période, on a réalisé les deux mesures et calculé la variance $\sigma^2(X)$ des mesures obtenues par la méthode C, la variance $\sigma^2(Y)$ des mesures obtenues par la méthode S, et la variance $\sigma^2(X - Y)$ des différences entre les mesures relatives à une même coulée. On a trouvé : $\sigma^2(X) = 20$, $\sigma^2(Y) = 35$ et $\sigma^2(X - Y) = 50$.

Quelle est la variance de la vraie teneur en phosphore pendant la période considérée ? Quelles sont les variances caractérisant les erreurs de mesure commises par les méthodes C et S ? Quelle est la covariance entre les deux mesures ?

On admettra que le résultat d'une mesure est la somme de la vraie teneur en phosphore et d'une erreur indépendante de cette teneur (modèle de l'erreur "absolue").

Exercice 5

A une heure déterminée, soit Z le nombre de voyageurs arrivant à une station dans une voiture de métro. On compte le nombre X de voyageurs qui montent et le nombre Y de voyageurs qui descendent. La voiture repart avec N personnes. On admettra que les variables X , Y , et Z sont indépendantes et qu'elles suivent des lois normales de moyennes $E(X) = 40$, $E(Y) = 30$, $E(Z) = 100$ et d'écart-type : $\sigma(X) = 9$, $\sigma(Y) = 12$ et $\sigma(Z) = 20$.

Calculer la capacité n de la voiture pour que $\text{Prob}\{N > n\} = 5\%$.

Exercice 6

Une étude a montré que, pour un haut-fourneau donné, lorsqu'on vise un poids de coulée μ_Q , on obtient un poids de fonte distribué suivant une loi normale de moyenne justement égale au poids visé μ_Q et d'écart-type σ_Q tel que $\sigma_Q = \gamma \mu_Q$ où γ est le coefficient de variation du H.F. étudié. Les moyens de transport dont on dispose entre le H.F. et l'aciérie consistent en n poches identiques. Lorsqu'on cherche à remplir au maximum une poche, le poids de fonte contenue suit une loi normale de moyenne μ_P et d'écart-type σ_P .

Calculer le poids de coulée qu'il faut viser pour qu'au risque α près, on ne doive pas reboucher le H.F. non vide. Application numérique : $\gamma = 0.01$, $n = 2$, $\alpha = 10\%$, $\mu_P = 30 T$ et $\sigma_P = 2 T$.

Exercice 7

La production d'une taille (chantier de production du charbon), pendant le cycle d'abattage, est distribuée suivant une loi normale de moyenne $500 T$ et d'écart-type $50 T$. Elle est desservie par un train de n berlines. Lorsqu'on cherche à remplir une berline, le poids chargé est distribué suivant une loi normale de moyenne $10 T$ et d'écart-type $1 T$. Calculer n pour que le risque de "manque à vide" soit égal à 2% seulement.

Exercice 8

Un problème grave est posé en Lorraine par les effondrements d'anciennes exploitations souterraines du minerai de fer, qui menacent des villages construits à leur aplomb. Une des méthodes d'exploitation les plus courantes consistait à laisser en place, dans la couche exploitée, des piliers destinés à soutenir les terrains sus-jacents. Dans l'une de ces anciennes exploitations, les mesures ont montré que les contraintes dans les piliers suivaient une loi normale de moyenne égale à 10 bars et d'écart-type égal à 3 bars, que la résistance à la compression du minerai suivait une loi normale de moyenne égale à 50 bars et d'écart-type égal à 15 bars. On peut par ailleurs raisonnablement accepter l'indépendance entre contrainte et résistance. Quel est la probabilité de ruine d'un pilier ?

Exercice 9

Des pièces de monnaie ont un poids distribué suivant une loi normale de moyenne 10 g et d'écart-type 0.15 g. Pour les compter, on les pèse par lots d'environ un kilo. Quel est le risque de faire une erreur d'une unité ?

Exercice 10

Dans une tréfilerie, l'atelier pour l'émaillage du fil de cuivre comporte 4 machines. Chaque machine est alimentée à partir d'un stock reconstitué chaque jour. Sachant que la production journalière d'une machine suit une loi normale de moyenne 10 tonnes et d'écart-type 2 tonnes, calculer le stock pour que le risque de pénurie soit égal à 1 % seulement.

Que devient le stock total qui serait nécessaire si l'on décidait d'organiser l'atelier de telle sorte que les 4 machines puissent être alimentées à partir d'un stock commun ?