Probabilités

Pour une expérience aléatoire, la probabilité p de l'événement A se défini comme

$$p(A) = \lim_{n\to\infty} \frac{f_a}{n}.$$

Le nombre de permutations de n éléments distincts est

$$p_e = n! = \prod_{i=0}^{n-1}(n-i) = n \times (n-1) \times (n-2) \times (n-3) \times \cdots \times 3 \times 2.$$

S'il existe des catégories avec des éléments similaires, il faudra diviser le nombre de permutations totale par le nombre de permutations au sein de chaque catégorie :

$$p_e = \frac{n!}{k_1!k_2!\ldots k_k!}.$$

Sans répétition, le nombre d'arrangements de p éléments pris au sein d'une population de n éléments est

$$A_n^p = \prod_{i=0}^{p-1}(n-i) = \frac{n!}{(n-p)!}.$$

Avec répétition ce nombre devient

$$\alpha_n^p = n^p.$$

Une combinaison ne prend pas en compte l'organisation des éléments au sein d'un l'ensemble. Le nombre de combinaisons de p éléments pris au sein d'une population de n éléments est donc

$$C_n^p = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}.$$

Les symboles $\cup$, $\cap$ et $|$ représentent respectivement l'union, l'intersection, et la condition. Si deux événements E et F sont incompatibles (i.e. ils ne peuvent se produire simultanément), la probabilité d'avoir E ou F s'écrit

$$P(E \cup F) = P(E) + P(F).$$

Pour deux événements compatibles, la probabilité d'avoir E ou F est

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

Deux résultats sont indépendants si

$$P(E|F) = P(E) \qquad \text{et} \qquad P(F|E) = P(F).$$

Les résultats d'un tirage avec remise sont indépendants.
Les résultats d'un tirage sans remise sont en général dépendants.

Si E et F sont deux événements indépendants, la probabilité que ces deux se produisent est donnée par

$$P(E \cap F) = P(E) \times P(F).$$

Si ces événements sont dépendants, la probabilité conditionnelle de l'événement E en connaissant (en sachant) F est donnée par:

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

Nous avons trois cartes, une de couleur rouge (recto-verso), une de couleur noire (recto-verso), et la derniere de couleur rouge (recto) et noire (verso). On tire une carte au hasard et une de ces faces est visible. Sachant qu'elle est de couleur noire quelle est la probabilité que l'autre face soit rouge?

Les probabilités conditionnelles sont un outil, souvent très pratique, pour calculer la probabilité d'intersections car

$$\begin{aligned} P(E \cap F) &= P(E|F) \times P(F) \\ &= P(F|E) \times P(E). \end{aligned}$$

Le théorème de Bayes

Soient $B_1, \ldots, B_N$ des événements incompatibles tels que $B_1 \cup \ldots \cup B_N$ engendrent tous les résultats possibles et tels que $P(B_1), \ldots, P(B_N)$ soient connus. Soit E un événement avec $P(E|B_1), \ldots, P(E|B_N)$ connus. Alors

$$
\begin{array}{lll}
1. \ P(E) & = & P(E \cap B_1) + \ldots P(E \cap B_N) \\
& = & P(E|B_1)P(B_1) + \ldots P(E|B_N)P(B_N) \\
2. \ P(B_i|E) & = & \dfrac{P(E|B_i) \times P(B_i)}{P(E|B_1)P(B_1) + \ldots P(E|B_N)P(B_N)}.
\end{array}
$$

Un test pour la grippe aviaire a les propriétés suivantes:

- Le test est positif à 99% lorsque le virus est présent.
- Le test est positif à 0.5% lorsque le virus est non présent.

On pense que la proportion des porteurs du virus dans la population est de 0.1%. Quelle est la probabilité qu'une personne ayant un test positif soit réellement infectée?

Distribution d'une variable aléatoire

Une variable aléatoire est un nombre réel qui dépend du résultat d'une expérience non-déterministe. Toute une s'erie d'expériences détermine les propriétés de la variable aléatoire : fonctions de répartition, densités, expérances, etc ...
La **distribution** (ou **fonction de répartition** ou **loi**) d'une variable aléatoire X est la fonction $F_X(x)$ définie comme

$$F_X(x) = P(X \leq x).$$

Cette fonction est définie pour toutes les valeurs x réelles et prend des valeurs entre 0 et 1. Elle est non-décroissante, elle tend vers 0 si $x \to -\infty$ et vers 1 si $x \to \infty$.
La **fonction de fréquence** (ou **fonction de densité de probabilité** ou **"PDF"**) est la dérivée de la fonction de répartition et s'écrit

$$f_X(x) = P(X = x) = \frac{dF_X(x)}{dx}.$$

Lois discrètes

La loi binômiale

La loi binômiale est une distribution discontinue qui s'applique à deux événements complémentaires. Elle donne les probabilités de voir apparaître un événement de probabilité p respectivement $0, 1, 2, 3, \ldots, i, \ldots, n$ fois sur n tirages ou expériences. La probabilité elle même s'écrit

$$P(x) = C_n^x (1-p)^{n-x} p^x = \frac{n!}{(n-x)!x!}(1-p)^{n-x}p^x$$

Moyenne ou espérance mathématique	$E = \sum_{i=1}^{n} x_i P(x_i) = np$
Variance	$\sigma^2 = np(1-p)$
Ecart type	$\sigma = \sqrt{np(1-p)}$
Coefficient d'asymétrie	$\alpha_3 = \frac{(1-p)-p}{\sqrt{np(1-p)}}$
Coefficient d'aplatissement	$\alpha_4 = 3 + \frac{1-6p(1-p)}{np(1-p)}$

La loi binômiale se caractèrise par deux paramètres $\{n, p\}$ et se note $\mathcal{B}(n, p)$.

La loi multinômiale

La loi multinômiale s'applique a k événements complémentaires de probabilité $p_1, \ldots, p_k$. Sur n tirages ou expériences, la probabilité que l'on observe x_i fois les événements de probabilité p_i s'écrit

$$P(x_1, x_2, \ldots, x_k) = \frac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}.$$

La loi multinômiale se caractèrise par $k+1$ paramètres $\{n, p_1, p_2, ..., p_k\}$ et se note $\mathcal{M}(n, p_1, p_2, ..., p_k)$.

La loi de Poisson

Lorsqu'un événement est rare, il faut multiplier les tirages afin d'obtenir des résultats fiables. Dans ce cas, l'évaluation des diff'erents termes du binôme de Newton put devenir laborieux. Tout l'interêt de la loi de Poisson est de décrire la limite de la loi binômiale si n est grand et p petit. Elle s'écrit

$$P(x) = \frac{(np)^x}{x!} \exp(-np)$$

ou

$$P(x) = \frac{\mu^x}{x!} \exp(-\mu)$$

en posant $\mu = np$.

Moyenne ou espérance mathématique	$E = \mu$
Variance	$\sigma^2 = \mu$
Ecart type	$\sigma = \sqrt{\mu}$
Coefficient d'asymétrie	$\alpha_3 = \frac{1}{\sqrt{\mu}}$
Coefficient d'aplatissement	$\alpha_4 = 3 + \frac{1}{\mu}$

La loi de Poisson se caractèrise par 1 paramètre μ et se note $\mathcal{P}(\mu)$.

La loi binômiale négative

La loi binômiale négative permet de calculer la probabilité d'être obligé d'effectuer x expériences identiques et indépendantes pour obtenir r fois un éveńement de probabilité p. Elle s'écrit

$$P(x) = \frac{(x-1)!}{(r-1)!(x-r)!} p^r (1-p)^{x-r}.$$

Moyenne ou espérance mathématique	$E = \frac{r}{p}$
Variance	$\sigma^2 = \frac{r(1-p)}{p^2}$
Ecart type	$\sigma = \sqrt{\frac{r(1-p)}{p^2}}$

La loi binômiale négative se caractèrise par 2 paramètres $\{r, p\}$ et se note $\mathcal{BN}(r, p)$.

La loi hypergéométrique

La loi hypergéométrique correspond à la loi binômiale sans remise. Pour une population de N éléments dont on extrait aléatoirement et indépendamment n éléments sur lesquels on observe un caractère avec la probabilité p, elle s'écrit

$$P(x) = \frac{C_x^{Np} C_{n-x}^{N(1-p)}}{C_n^N}.$$

Moyenne ou espérance mathématique	$E = np$
Variance	$\sigma^2 = npq\frac{N-n}{N-1}$
Ecart type	$\sigma = \sqrt{npq\frac{N-n}{N-1}}$

Pour $n > 1$ et N fini,

$$npq\frac{N-n}{N-1} < npq.$$

La loi hypergéométrique se caractèrise par 3 paramètres $\{N, n, p\}$ et se note $\mathcal{H}(r, p)$.

Les lois continues

La distribution normale

Si $n \to \infty$ et p n'est pas trop voisin de 0 ni de 1, la distribution binômiale tend vers la distribution normale d'équation

$$f(x) = \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{(x-np)^2}{2npq}\right)$$

en remplaçant npq par σ^2 et np par $\bar{x}$ on obtient

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right).$$

On obtient la distribution normale centrée réduite

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right),$$

en posant

$$X = x - \bar{x}, \qquad z = \frac{X}{\sigma} \qquad \text{et} \qquad \sigma = 1.$$

Moments de la loi normale centrée réduite.

Moyenne ou espérance mathématique	μ	$= 0$
Variance	σ^2	$= 1$
Ecart type	σ	$= 1$
Coefficient d'asymétrie	α_3	$= 0$
Coefficient d'aplatissement	α_4	$= 3$
Ecart à la moyenne	$\sigma\sqrt{\frac{2}{\pi}}$	$= \sqrt{\frac{2}{\pi}}$

Quelques propriétés de la loi normale centrée réduite

$$f(-z) = f(z).$$

$$f'(z=0) = 0$$

$$f(z=0) = \frac{1}{\sqrt{2\pi}} = 0.399$$

Par conséquent le mode, la médiane et la moyenne sont confondus.

Relation entre la fonction de répartition et la fonction de fréquence pour une loi normale centrée réduite

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz = 1$$

Pour une distribution normale, on obtient les probabilités en intégrant la fonction de densité de probabilité. Ainsi la probabilité d'obtenir une valeur z inférieure à z_1 s'écrit

$$\Phi(z_1) = P(z < z_1) = \int_{-\infty}^{z_1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz.$$

$$P(z > z_1) = 1 - \Phi(z_1)$$

$$P(z_1 < z < z_2) = \Phi(z_2) - \Phi(z_1)$$

$$P(0 < z < z_2) = \Phi(z_2) - 0.5$$

$$P(z_1 < z < 0) = \Phi(-z_1) - 0.5$$

$$\Phi(1) - \Phi(-1) = 0.6827$$

$$\Phi(2) - \Phi(-2) = 0.9545$$

$$\Phi(3) - \Phi(-3) = 0.9973$$

$$\Phi(1.96) = 0.975 \quad \Rightarrow \quad \Phi(1.96) - \Phi(-1.96) = 0.95 \quad \Rightarrow \quad 1 - (\Phi(1.96) - \Phi(-1.96)) = 0.05$$

$$\Phi(2.57) = 0.995 \quad \Rightarrow \quad \Phi(2.57) - \Phi(-2.57) = 0.99 \quad \Rightarrow \quad 1 - (\Phi(2.57) - \Phi(-2.57)) = 0.01$$

La loi exponentielle

Une loi de Poisson de moyenne μ suit une loi exponentielle de paramètre $\Theta = 1/\mu$ et de densité de probabilité

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \Theta \exp(-\Theta x) & \text{si } x > 0, \end{cases}$$

Si la loi de Poisson représente le nombre d'apparitions d'un phénomène aléatoire dans un intervalle de temps fixe, la variable x représente ici l'intervalle de temps séparant deux événements (e.g. pannes, passages d'individus indépendant). Cette loi peut aussi s'appliquer à des distances entre des éléments répartis aléatoirement. Il est possible de translater la loi exponentielle d'un décalage ν et la densité de probabilité devient alors

$$f(x) = \begin{cases} 0 & \text{si } x \leq \nu, \\ \Theta \exp(-\Theta(x-\nu)) & \text{si } x > \nu, \end{cases}$$

On peut démontrer que

$$\mu = \nu + \frac{1}{\Theta},$$

et que

$$\sigma_x^2 = \frac{1}{\Theta^2}.$$

La loi exponentielle s'exprime au travers de deux paramètres $\{\Theta, \nu\}$ et peut s'écrire $\mathcal{E}(\Theta, \nu)$.

La loi uniforme

La densité d'une loi uniforme est

$$f(x) = \begin{cases} \frac{1}{\Theta} & \text{si } x \in [0, \frac{1}{\Theta}], \\ 0 & \text{else.} \end{cases}$$

Il est possible de translater la loi uniforme d'un décalage ν et la densité de probabilité devient alors

$$f(x) = \begin{cases} \frac{1}{\Theta} & \text{si } x \in [\nu, \nu + \frac{1}{\Theta}], \\ 0 & \text{else.} \end{cases}$$

On peut démontrer que

$$\mu = \nu + \frac{\Theta}{2}$$

et que

$$\sigma_x^2 = \frac{\Theta^2}{12}$$

La loi uniforme s'exprime au travers de deux paramètres $\{\Theta, \nu\}$ et peut s'écrire $\mathcal{U}(\Theta, \nu)$.

La loi Gamma

La loi de distribution gamma repose sur une fonction Γ portant le même nom. Cette fonction est une intégrale n'admettant pas de primitive sous la forme d'une fonction élémentaire. Elle sécrit

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} \exp(-y) dy \qquad \text{pour} \qquad \alpha > 0.$$

Cette intégrale converge seulement si $\alpha > 0$ et on peut vérifier que

$$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha) = \alpha!\Gamma(1) = \alpha!,$$

et que

$$\Gamma(\frac{1}{2}) = \pi.$$

Une variable aléatoire continue x obéit à une loi gamma $G(\alpha, \beta)$, si sa fonction de densité de probabilité s'écrit

$$f(x) = \frac{x^{\alpha-1} \exp(-\frac{x}{\beta})}{\beta^\alpha \Gamma(\alpha)}$$

avec x, α et β supérieurs à zéro.
La loi gamma $G(r, 1/\Theta)$ est une généralisation de la fonction exponentielle. En effet, elle permet de déterminer le temps qui s'écoule entre la i^{eme} et la $(i+r)^{eme}$ apparition de l'événement (r est un réel). Ainsi, $G(1, 1/\Theta)$ est une fonction exponentielle $Exp(\Theta, 0)$.
La loi Gamma s'exprime au travers de deux paramètres $\{\alpha, \beta\}$ et peut s'écrire $\mathcal{G}(\alpha, \beta)$.
On peut démontrer que les moments de $G(\alpha, \beta)$ sont

$$\mu = \alpha\beta,$$

et

$$\sigma_x^2 = \alpha\beta^2.$$

La loi gamma est un bon modèle de probabilité pour prévoir la durée de vie des appareils qui subissent une usure (e.g. véhicules, apparails ménagers, ordinateurs). Cette loi est parfois utile en météorologie (nombre de pluie durant un interval de temps T) est peut utilisée en biostatistique.

La loi du χ^2

Si $Z_1, Z_2, ..., Z_\nu$ sont des variables aléatoires normales centrées réduites et indépendantes entre elles, la somme des carrés de ces variables aléatoires obéit à une loi du χ^2 à ν degré de liberté

$$\chi^2 = Z_1^2 + Z_2^2, ... + Z_\nu^2.$$

$G(\nu/2, 2)$ est une loi du χ^2 avec ν le nombre de composantes indépendantes impliquées impliquées dans le calcul de la valeur du χ^2. La densité de probabilité de χ^2_ν s'écrit donc

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)}\ x^{\frac{\nu}{2}-1}\exp(-x/2) & \text{si } x > 0, \end{cases}$$

Si X_i sont des variables indépendantes respectant des lois normales, de moyennes μ_i et de variance σ_i, on peut montrer que

$$\chi^2_{(n)} = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2.$$

Si χ^2_i sont des variables indépendantes respectant des lois du χ^2 avec ν_i degrés de libertés, on peut montrer que

$$\chi^2 = \chi^2_1 + \chi^2_2 + ... + \chi^2_n$$

est une loi du χ^2 avec

$$\nu = \nu_1 + \nu_2 + ...\nu_n$$

degrés de libertés. La réciproque de cette propriété peut servir à démontrer l'indépendance des variables X_i.
Les moments de la loi du χ^2 sont

$$\mu_{\chi^2} = \nu$$

et

$$\sigma^2_{\chi^2} = 2\nu$$

Lorsque le nombre de variables croît, la loi du χ^2 tend vers la loi normale.

La loi de Fisher-Snedecor

Soit χ^2_1 et χ^2_2 un couple de variables aléatoires indépendantes suivant respectivement des lois du χ^2 à ν_1 et ν_2 degrés de libertés. Alors

$$F = \frac{\chi^2_1/\nu_1}{\chi^2_2/\nu_2}$$

La variable F a une fonction de densité de probabilité de la forme

$$f(F) = \begin{cases} 0 & \text{si } F \leq 0, \\ \dfrac{\dfrac{\nu_1}{\nu_2}^{\nu_1/2}\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \dfrac{F^{\dfrac{\nu_1}{2}} - 1}{(1+\dfrac{\nu_1}{\nu_2}F)^{\dfrac{\nu_1+\nu_2}{2}}} & \text{si } F > 0, \end{cases}$$

On peut montrer que

$$\mu_F = \frac{\nu_2}{\nu_2 - 2} \quad \text{si} \quad \nu_2 > 2$$

et que

$$\sigma^2_F = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \quad \text{si} \quad \nu_2 > 4.$$

La loi de Student

Soit une variable obéissant à une loi du χ^2 à ν degré de liberté et une variable aléatoire Z respectant une loi normale centrée. La variable

$$t = \frac{Z}{\sqrt{\chi^2/\nu}}$$

suit une loi de Student à ν degré de liberté.

t a une fonction de densité de probabilité de la forme

$$f(t) = \frac{\Gamma(\dfrac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\nu/2)}(1+t^2/\nu)^{-(\nu+1)/2}$$

On peut montrer que

$$\mu_t = 0 \quad \text{si} \quad \nu > 1$$

et que

$$\sigma_t^2 = \frac{\nu}{\nu - 2} \quad \text{si} \quad \nu > 2.$$

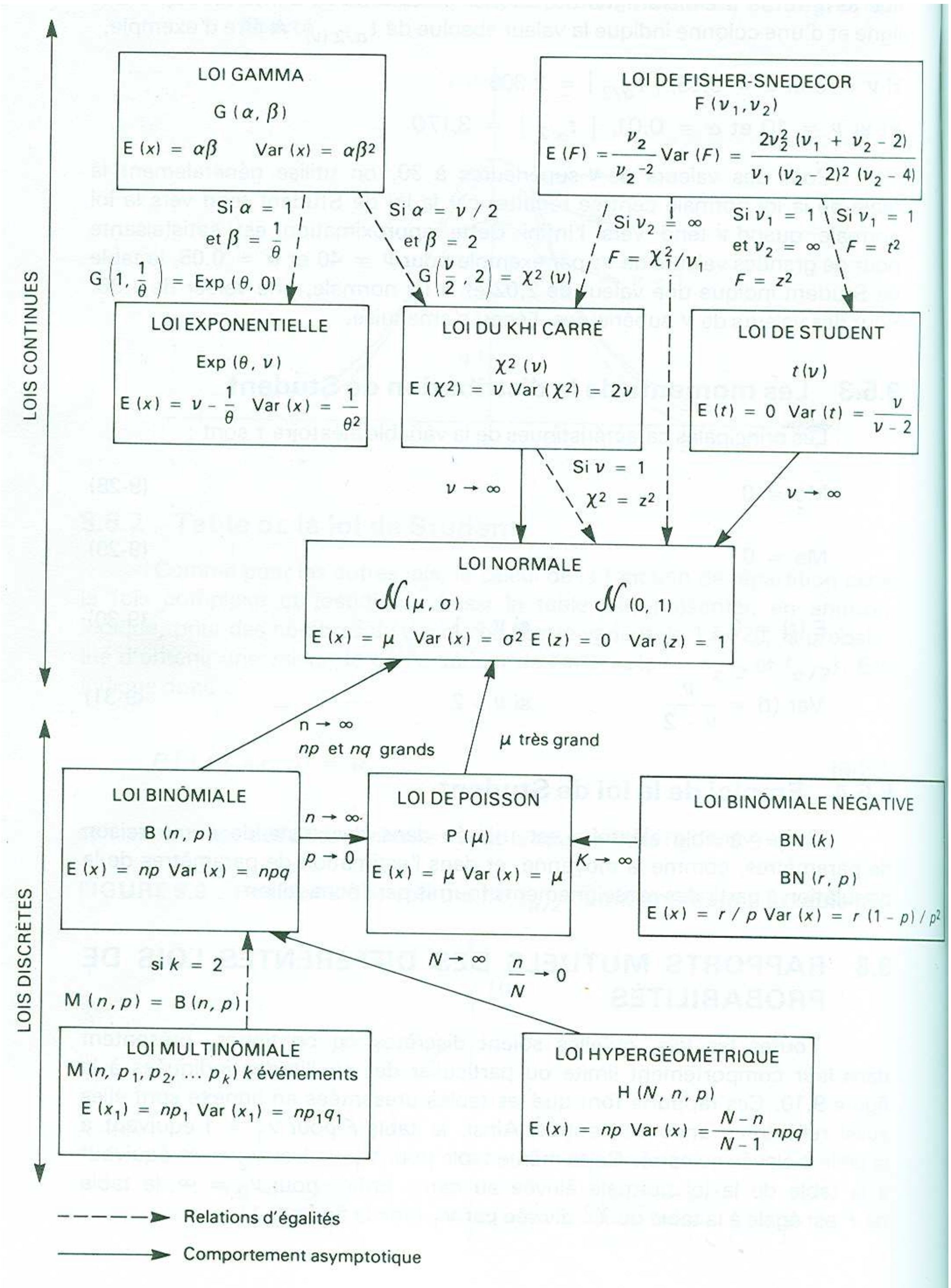

Figure 1: Relations entre les différentes lois de probabilités (extrait de Scherrer, *Biostatistiques*, Gaëtan Morin (2004)).