Théorème dit central limite” ou de normalisation par la méthode des cumulants

Ce théorème dit très globalement qu'un phénomène macroscopique aléatoire somme d'un très grand nombre de contributions microscopiques, comme par exemple la pression d'un gaz sur une paroi, résultat des chocs sur cette paroi d'atomes ou molécules, est très souvent décrit par une loi de Gauss, ou loi normale, d'où l'expression de normalisation.

La méthode la plus rapide consiste à utiliser les cumulants d'une variable aléatoire (VA) qu'il convient d'abord de définir. Soit X une VA qu'on suppose continue et décrite par sa densité de probabilité (DDP) $p_X(x)$.

On appelle moment m_n d'ordre n de X, n entier, l'espérance de X^n soit

$$m_n \stackrel{\triangle}{=} \mathsf{E}(X^n) = \int_{-\infty}^{\infty} p_X(x) x^n dx. \tag{1}$$

On admet que cette intégrale est convergente $\forall n$, bien qu'il existe des lois de probabilité où ce n'est pas le cas, qui sortent donc du cadre de la suite. De telles lois apparaissent très rarement dans les applications du calcul des probabilités.

Il convient également d'introduire la fonction caractéristique (FC) $\phi(u)$ de X définie par

$$\phi_X(u) \stackrel{\triangle}{=} \mathsf{E}(\mathrm{e}^{iuX}) = \int_{-\infty}^{\infty} p_X(x) \mathrm{e}^{iux} dx. \tag{2}$$

Il s'agit évidemment d'une fonction en général à valeurs complexes de la variable réelle u. L'inégalité classique sur le module de intégrales indique que cette fonction existe toujours puisque $p_X(x)$ est par définition normée (soit d'intégrale égale à 1). En effet on a

$$|\phi_X(u)| \leq \int_{-\infty}^{\infty} |p_X(x) \mathrm{e}^{iux}| dx = \int_{-\infty}^{\infty} p_X(x) dx = 1. \tag{3}$$

Cette fonction porte le nom de caractéristique car sa connaissance est équivalents à celle de $p(x)$. Il apparaît en effet que (2) est une transformée de Fourier et la théorie de ces transformations introduit ce qu'on appelle la transformation inverse, soit

$$p_X(x) = \left(\frac{1}{2\pi}\right) \int_{-\infty}^{\infty} \phi_X(u) \mathrm{e}^{-iux} du. \tag{4}$$

En raison de cette équivalence, les moments définis par (1) doivent pouvoir se déduire de la FC. Pour ceci il suffit d'utiliser la définition de l'exponentielle par sa série, soit $\exp(a) = \sum a^k/k!$, ce qui donne en utilisant (1)

$$\phi(u) = \sum_{k=0}^{\infty} m_k \frac{(iu)^k}{k!} \ . \tag{5}$$

Il est clair qu'il peut se poser des problèmes de convergence puisque l'on a vu qu'un ou plusieurs moments peuvent ne pas être finis, alors que la FC est toujours définie en raison de (3). Nous admettrons ici que toutes les lois considérées correspondent à des cas où cette série converge et l'on vérifie que c'est effectivement le cas de pratiquement toutes les lois utilisées dans les applications.

La fonction caractéristique joue un rôle fondamental dans tous les problèmes où apparaissent des sommes de VA indépendantes. Pour le voir considérons deux VA indépendantes X et Y et soit $Z = X + Y$ leur somme. En utilisant l'indépendance et la règle du produit de deux exponentielles vous devez pouvoir démonter la relation suivante entre fonctions caractéristiques

$$\phi_Z(u) = \phi_X(u).\phi_Y(u). \tag{6}$$

Comme en algèbre linéaire il convient d'étudier les transformations de moments dans les opérations de multiplication par un scalaire ou d'addition. Soit α un scalaire réel et Y la VA déduite de X par $Y = \alpha X$. Il résulte de la définition des moments que l'on a $m_{Y,k} = \alpha^k m_{X,k}$.

Par ailleurs soit X et Y deux VA indépendantes et Z leur somme. En appliquant la formule du binôme exprimant Z^n en fonction des puissances de X et Y, on obtient

$$m_{Z,n} = \sum_{p=0}^{n} C_n^p \, m_{X,p} \, m_{Y,n-p}, \tag{7}$$

où C_n^p est le coefficient du binôme égal à $n!/[p!(n-p)!]$. Si les VA X et Y ne sont plus indépendantes, le résultat est beaucoup plus compliqué car dans ce cas $\mathsf{E}(X^k Y^l)$ ne vaut plus nécessairement $m_{X,k} \, m_{Y,l}$.

Seconde fonction caractéristique et cumulants

La seconde fonction caractéristique $\psi(u)$ est le logarithme népérien de $\phi(u)$. Il s'agit donc en général du logarithme d'un nombre complexe. Mais en dehors du caractère multiforme de cette fonction que l'on corrige avec une coupure, toutes les relations algébriques connues dans le cas réel sont applicables dans le cas complexe. Son développement en série analogue à (5) introduit des coefficients dénommés cumulants c_k. On a donc

$$\psi(u) = \sum_{k=1}^{\infty} c_k \frac{(iu)^k}{k!} \ . \tag{8}$$

Le fait que cette série débute à $k = 1$ provient de la relation fondamentale $\phi(0) = 1$ indiquée ci-dessus et qui entraîne son équivalent $\psi(0) = 0$. De plus, si $Y = \alpha X$ il en résulte par définition que $\phi_Y(u) = \phi_X(\alpha u)$, et on en déduit la même relation pour $\psi(.)$. En remplaçant dans (8) on obtient que, comme pour les moments, les cumulants de X et Y sont reliés par $c_{Y,n} = \alpha^n \, c_{X,n}$.

La relation sur la somme est plus intéressante. En effet si X et Y sont indépendantes, la relation (6) devient

$$\psi_Z(u) = \psi_X(u) + \psi_Y(u), \tag{9}$$

de sorte que les cumulants correspondants satisfont

$$c_{Z,n} = c_{X,n} + c_{Y,n} \; . \tag{10}$$

Ainsi le cumulant d'ordre quelconque de la somme de deux VA indépendantes est égal à la somme de leurs cumulants du même ordre. Cette relation va jouer un rôle essentiel dans la suite.

Il est évident qu'il y a une relation entre les moments et les cumulants puisqu'ils sont tous déterminés par la même fonction caractéristique de départ. Le calcul est un peu plus laborieux comme on va le voir pour les moments et cumulants d'ordre un à trois. Nous donnons le principe de la méthode laissant au lecteur le soin d'effectuer les calculs complets. On part du développement (5) qu'on limite à l'ordre trois en puissances de u. Ceci peut alors s'écrire

$$\phi(u) = 1 + A(u) = 1 + m_1 iu - (1/2) m_2 u^2 - (1/6) m_3 iu^3. \tag{11}$$

Mais on sait aussi que pour $|A|$ suffisamment petit on peut développer le logarithme sous la forme limitée au troisième ordre

$$\ln(1 + A) = A - (1/2)A^2 + (1/3)A^3 \tag{12}$$

Cette formule peut se démontrer en dérivant les deux termes de l'équation et en utilisant les premiers termes de la série géométrique. Il suffit alors de calculer les trois puissances de A et d'ordonner leurs sommes pondérées selon les puissances de u. En comparant avec le développement (8) limité au troisième ordre on obtient une expression de cumulants en fonction des moments. Après un calcul algébrique simple mais nécessitant ordre et attention on trouve le résultat suivant : $c_1 = m_1$, $c_2 = m_2 - m_1^2$ qui est en fait la variance σ^2 de la VA dont $\phi(u)$ est la fonction caractéristique et

$$c_3 = m_3 - 3m_1 m_2 + 2m_1^3. \tag{13}$$

Il est évidemment aisé de faire l'opération inverse donnant les moments en fonction des cumulants, soit

$$m_3 = c_3 + 3c_1c_2 + c_1^3. \tag{14}$$

Il existe des formules complètes valables pour tous les ordres mais dont nous n'avons pas l'usage dans toute la suite. Notons tout de suite que si la VA de départ est centrée, soit si $m_1 = 0$, les trois premiers cumulants sont égaux aux trois premiers moments.

Exemples de cumulants

Il existe deux lois qui sont les plus importantes pour les applications : la loi normale pour les VA continues et la loi de Poisson pour les variables discrètes.

Loi normale

La DDP d'un VA normale de moyenne m et de variance σ^2 est donnée par l'expression

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp[-(x-m)^2/2\sigma^2]. \tag{15}$$

Sa FC se calcule dans les cours d'analyse mathématique et vaut

$$\phi(u) = \exp[imu - (1/2)\,\sigma^2 u^2]. \tag{16}$$

On mémorise ce résultat fondamental en disant que la transformée de Fourier d'une forme gaussienne est un forme gaussienne, c'est-à-dire, à un coefficient de normalisation près, l'exponentielle d'une forme quadratique négative. La normalisation assure que $\phi(0) = 1$, comme indiqué ci-dessus. En calculant le logarithme on obtient la seconde fonction caractéristique qui vaut donc

$$\psi(u) = \ln[\phi(u)] = imu - (1/2)\,\sigma^2 u^2, \tag{17}$$

et en comparant à (8) on obtient le résultat fondamental suivant : **tous les cumulants d'une loi normale sont nuls, sauf les deux premiers**. Si de plus la loi est centrée ($m = 0$) elle ne possède qu'un seul cumulant non nul qui est sa variance. Bien entendu ce résultat a une réciproque : une VA dont la variance σ^2 est le seul cumulant non nul est nécessairement une VA de type $\mathrm{N}(0, \sigma^2)$, ce qui veut dire normale, centrée et de variance σ^2. Pour le montrer il suffit de remonter tous les calculs à l'envers et de tomber sur (17) et par transformation de Fourier inverse sur (15)

Loi de Poisson

Un VA discrète X de prenant que des valeurs entières non-négatives est dite de Poisson si

$$p_n \triangleq \Pr(X = n) = \exp(-m)\frac{m^n}{n!}. \tag{18}$$

Un calcul que je ne fais pas mais que je vous recommande de faire indique que la moyenne et la variance sont égales et valent m. Le calcul de la FC est encore plus simple (faites le) et le résultat est

$$\phi(u) = \exp[m(\mathrm{e}^{iu} - 1)] \tag{19}$$

dont le logarithme s'écrit

$$\psi(u) = m(\mathrm{e}^{iu} - 1). \tag{20}$$

En comparant à (8) on en conclut que tous les cumulants sont égaux à m, ce qui est également une caractéristique très importante de la loi de Poisson.

Le théorème central limite

Soit une suite de n VA centrées IID X_i de cumulants c_k et S_n leur somme. D'après ce qui a été vu précédemment le cumulant d'ordre k de S_n vaut nc_k. Soit maintenant Y_n la VA $(1/\sqrt{n})S_n$. D'après la règle sur le produit d'une VA par un scalaire indiquée ci-dessus le cumulant d'ordre k de Y_n vaut $c_{Y,k} = (n)^{1-(k/2)}c_k$. Pour $k = 2$ cela donne $c_{Y,2} = c_2$, et comme le cumulant d'ordre 2 est la variance, cela signifie la somme pondérée $(1/\sqrt{n})S_n$ des X_i a la même variance que chacune des VA X_i. Par contre pour $k > 2$ on voit que tous les cumulants de Y_n tendent vers 0, ce qui caractérise une loi normale $\mathrm{N}(0, \sigma^2)$. C'est le théorème central limite ou de normalisation disant qu'un somme correctement pondérée de variables aléatoires IID a une distribution qui tend vers une distribution normale. Il convient de noter l'importance de la pondération qui assure la constance de la variance quand le nombre n de termes de la somme croît. Avec toute autre pondération la variance de la somme pondérée tendrait soit vers 0 soit vers l'infini quand n croit indéfiniment.

B. P.

11/10/19