Mathématiques et sciences humaines
Mathematics and social sciences

158 | Été 2002
Varia

Une définition fonctionnelle de la dispersion en statistique et en calcul des probabilités : les fonctions de concentration de Paul Lévy

A functional definition of the dispersion in statistics and probability theory: Paul Lévy's concentration functions

Marc Barbut

Édition électronique
URL : http://journals.openedition.org/msh/2907
DOI : 10.4000/msh.2907
ISSN : 1950-6821

Éditeur
Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée
Date de publication : 1 mars 2002
ISSN : 0987-6936

Référence électronique
Marc Barbut, « Une définition fonctionnelle de la dispersion en statistique et en calcul des probabilités : les fonctions de concentration de Paul Lévy », *Mathématiques et sciences humaines* [En ligne], 158 | Été 2002, mis en ligne le 10 février 2006, consulté le 19 avril 2019. URL : http://journals.openedition.org/msh/2907 ; DOI : 10.4000/msh.2907

Math. & Sci. hum., (40e année, n° 158, *2002*, p. 31-57)

UNE DÉFINITION FONCTIONNELLE DE LA DISPERSION, EN STATISTIQUE ET EN CALCUL DES PROBABILITÉS□ LES FONCTIONS DE CONCENTRATION DE PAUL LÉVY

Marc BARBUT[1]

RÉSUMÉ – *On étudie certaines propriétés des fonctions de concentration de Paul Lévy et principalement la question de leur inversion. Ces fonctions constituent des résumés fonctionnels de la dispersion dans la répartition de variables numériques en Calcul des Probabilités et en Statistique.*

MOTS-CLÉS – Concentration, Dispersion, Inégalité, Sous-additivité, Densité, Fonction de répartition, Unimodal, Courbes en cloche, Courbes en U.

SUMMARY – A functional Definition of the Dispersion in Statistics and Probability Theory: Paul Lévy's Concentration Functions
Some properties of Paul Lévy's concentration functions are exhibited and mainly the problem of their inversion. These functions summarize in a functional way the dispersion in the distribution of random or statistical numerical variables.

KEYWORDS – Concentration, Dispersion, Inequality, Sub-additivity, Density, Distribution functions, Unimodal, Bell curves, U-curves.

1. INTRODUCTION

Concentration, dispersion, variabilité, inégalités, etc., dans la répartition d'une variable, tous ces mots ont un sens, et souvent une grande importance pour le statisticien.

La pratique la plus courante consiste à leur associer des indicateurs numériques tels que la variance, l'écart type, l'écart moyen absolu, le coefficient de variation, etc.

Mais ceci a deux inconvénients majeurs. D'une part, certains de ces indicateurs ne sont pas toujours définis, notamment si la variable en jeu n'a pas de moments, comme c'est le cas par exemple, pour la distribution de Cauchy, ou certaines de celles de Pareto-Lévy (lois stables pour l'addition des variables aléatoires indépendantes). D'autre part, résumer la concentration par un indicateur numérique se paie au prix d'une considérable «□perte d'information□.

[1] École des hautes études en sciences sociales, Centre d'analyse et de mathématique sociales, 54 bd Raspail 75270 Paris cedex 06, e-mail□cams@ehess.fr.

À titre d'exemple, la distribution Beta symétrique définie sur l'intervalle [0,1] par la densité :

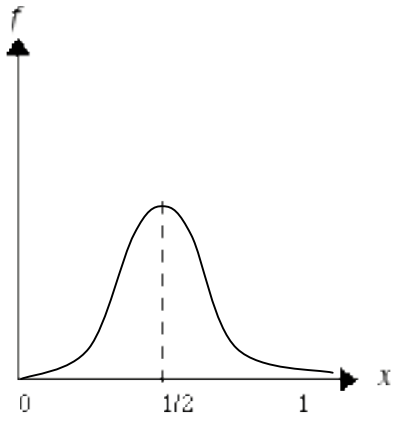
$$f(x) = 140\, x^3\, (1 - x)^3$$

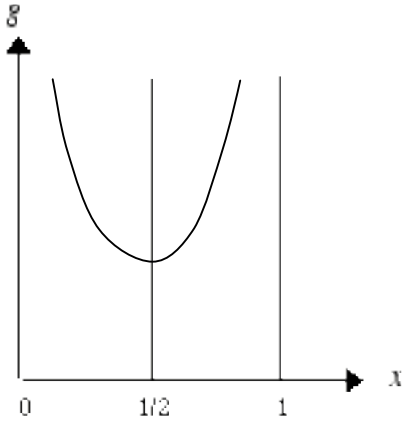
a un écart type égal à $\frac{1}{6}$.

Or, la distribution « arc sinus » définie sur le même intervalle par la densité :

$$g(x) = \frac{2}{\Pi}\ \frac{1}{\sqrt{1 - (2x - 1)^2}}$$

a un écart type égal à $\frac{1}{2\sqrt{3\Pi}}$ qui diffère peu de $\frac{1}{6}$. Or, pour la première, qui est « en cloche », la masse est concentrée autour du milieu $\frac{1}{2}$ de l'intervalle, alors que pour la seconde, qui est « en U », c'est au voisinage des deux extrémités 0 et 1 de l'intervalle que se fait la concentration (cf. Figures 1 et 1bis)





Une façon de pallier ce type d'antinomies, et bien d'autres inconvénients, est de caractériser la dispersion dans la répartition d'une variable par une autre fonction de répartition, qui soit plus simple et se prête à la comparaison lorsque plusieurs variables sont en jeu.

C'est ce que font notamment les fonctions de concentration de Lorenz et Gini, qui ont souvent fait l'objet d'articles dans la revue *Mathématiques et Sciences humaines*[2]. Mais ce procédé n'a de sens que si l'on a affaire à *deux* répartitions : la courbe de concentration de Lorenz et Gini est en effet toujours celle d'une variable Y *par rapport* à une autre variable X définie sur le même ensemble que Y.

[2] Cf. n° 88, 1984 ; 90, 1985 ; 131, 1995 ; 134, 1996 ; 142, 1998.

Ce fait est souvent oublié, pour deux raisons au moins :

1. il arrive que la variable X soit la variable uniforme sur un intervalle ;
2. il arrive aussi, notamment dans les applications à l'étude des inégalités économiques ou sociales (inégalités de revenus, de richesses, de populations, etc.) qu'il y ait une relation fonctionnelle entre les fonctions de répartition des deux variables (en l'occurrence $G(x) = K \int_{-\infty}^{x} td\,F(t)$, ce qui suppose que la moyenne de X existe).

La démarche adoptée par Paul Lévy a été autre : définir une fonction de concentration d'une variable X qui ne dépende que de celle-ci et qui soit toujours définie, que la variable ait des moments ou non.

La question qui se pose dès lors au statisticien est de déterminer les ensembles de répartitions admettant même fonction de concentration.

L'ordre partiel de comparaison uniforme entre leurs fonctions de concentration permet ensuite une analyse beaucoup plus fine (qu'avec un indicateur numérique) de la comparaison, pour deux variables X et Y, de leur comportement en matière de dispersion.

Daniel Dugué, qui pensait à juste titre que la statistique pourrait tirer un grand profit de la mise en œuvre de l'outil proposé par Paul Lévy, s'était posé ce problème de l'inversion de ses fonctions de concentration aux alentours de 1957-1958. À l'époque, les choses en étaient à peu près restées là. Le texte qui suit fournit des réponses pour deux classes étendues de distributions, les plus utilisées en statistique : les répartitions « en cloche » et celles « en U ».

2. DÉFINITIONS ET EXEMPLES

2.1. DÉFINITIONS

Dans ses travaux en Calcul des Probabilités, Paul Lévy définit et utilise, notamment pour démontrer des théorèmes de convergence, la *fonction de dispersion* et la *fonction de concentration* d'une variable aléatoire[3].

La fonction de dispersion ω associée à une variable aléatoire (v.a. dans la suite) X de fonction de répartition (abrégée en f.r. dans la suite) F définie sur $\mathbb{R}$ a pour définition : pour tout α $(1 \geq \alpha \geq 0)$, $\omega(\alpha)$ est la longueur minimum d'un intervalle fermé de probabilité supérieure ou égale à α (ces notations sont celles de P. Lévy).

Elle est la fonction inverse de la *fonction de concentration* (f.c. dans la suite), notée Q_X ou Q_F de la v.a. X : $Q_F(y)$ est, pour tout $y \geq 0$, la *probabilité maximum* d'un intervalle fermé de longueur y.

[3] Cf. par exemple, *Théorie de l'addition des variables aléatoires* (T.A.V.A.), Paris, Gauthier-Villard, 1937.

(1) $\forall y \geq 0,\ Q_F(y) = \max_x \left(\overline{F}(x+y) - \underline{F}(x)\right)$

où $\overline{F}(t)$ et $\underline{F}(t)$ désignent respectivement les limites à droite et à gauche de $F(u)$ lorsque u tend vers t.

Il est clair que $Q_F(y)$ est monotone non décroissante sur la demi-droite $y \geq 0$ et que $Q_F(\infty) = 1$, c'est donc une f.r. $Q_F(0)$ est le plus grand saut de F, de sorte que $Q_F(0) = 0$ ssi F est continue.

Lorsque F est *absolument continue* de densité f (ce qui est le cas le plus usuel dans les applications du calcul des probabilités à la statistique), la définition (1) de Q_F devient :

(2) $\forall y \geq 0,\ Q_F(y) = \max_x \left(F(x+y) - F(x)\right) = \max_x \left(\int_x^{x+y} f(t)\ dt\right)$

Il est d'ailleurs équivalent de noter :

(2') $Q_F(y) = \max_x \left(F(x) - F(x-y)\right)$

Ou plus généralement :

(2") $Q_F(y) = \max_x \left(F(x+py) - (F(x-qy)\right)$

avec, $p,q \geq 0,\ p+q = 1$.

Le maximum intervenant dans ces expressions est atteint pour au moins une valeur bien déterminée $x(y)$ de x.

Lorsque F n'est pas absolument continue, on remplacera dans (2), (2') ou (2") l'opérateur *maximum* par le *supremum*.

D'autre part, il est clair que la f.c. d'une v.a. X est invariante par rapport aux translations (ou changements d'origine) de celle-ci, ainsi que par rapport aux symétries ; un changement d'unité de rapport $\lambda > 0$ sur la v.a. X se traduit par le même changement d'unité sur la v.a. Y positive de f.r. Q_X ; pour toutes les v.a. d'un même type (*i. e.* définies à une transformation affine près), les f.c. sont également toutes de même type.

1.2. Quelques exemples

Pour une distribution de Laplace-Gauss centrée, de densité :

(1.2.1) $g(x) = \dfrac{1}{\sigma\sqrt{2\Pi}}\ e^{-\frac{x^2}{2\sigma^2}}$

et de f.r. $G(x) = \int_{-\infty}^{x} f(t)dt$, on a $x(y) = -\dfrac{y}{2}$

(3) $Q_G(y) = G\left(\dfrac{y}{2}\right) - G\left(-\dfrac{y}{2}\right) = 2G\left(\dfrac{y}{2}\right) - 1$

Q_G est absolument continue, de densité□

(3') $$Q'_G(y) = q(y) = 2G'\left(\frac{y}{2}\right) = g\left(\frac{y}{2}\right)$$

Les relations (3) et (3') sont d'ailleurs vérifiées pour toute v.a. X *absolument continue, unimodale* et *symétrique* (cf. Figure 2) par rapport au mode qui est, dans ce cas, également la médiane et, lorsqu'elle existe, la moyenne de X.

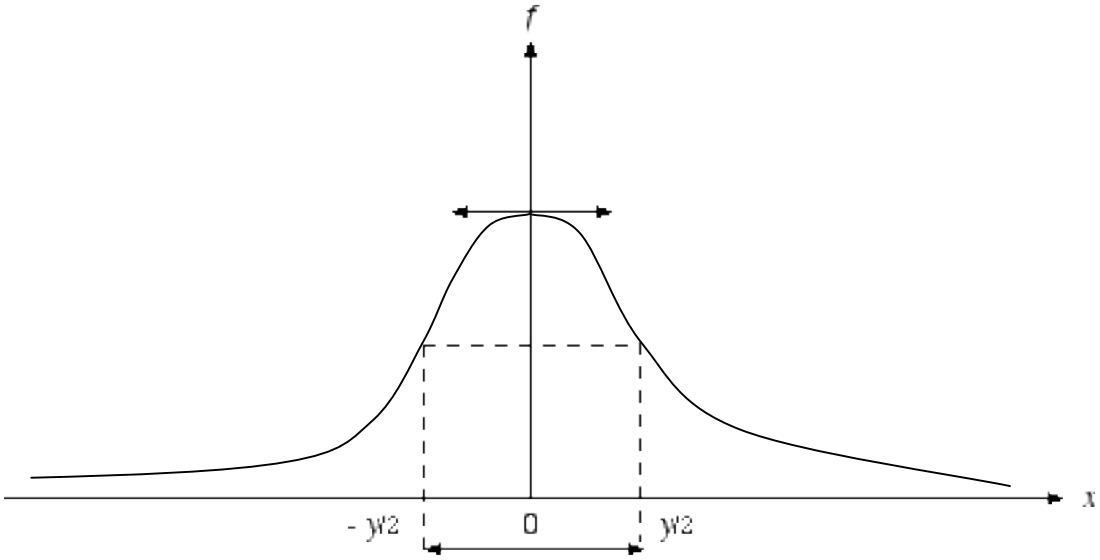


Figure 2

Par exemple, on a□

– pour la distribution de Cauchy, qui n'a pas de moyenne,

(1.2.2) $$F(x) = \frac{1}{2} + \frac{1}{\Pi}\,\text{Arc tg}\, x \qquad f(x) = \frac{1}{\Pi}\cdot\frac{1}{1+x^2}$$

$$G_F(y) = \frac{2}{\Pi}\,\text{Arc tg}\,\frac{y}{2} \qquad q(y) = \frac{1}{\Pi}\,\frac{1}{1+\left(\frac{y}{2}\right)^2}$$

– pour la première distribution de Laplace, de densité□

(1.2.3) $$f(x) = \frac{1}{2\lambda}\, e^{-\frac{|x-x_0|}{\lambda}} \qquad (\lambda > 0)$$

on a□

$$Q_F(y) = 1 - e^{-\frac{y}{2\lambda}} \qquad \text{et} \qquad q(y) = \frac{1}{2\lambda}\, e^{-\frac{y}{2\lambda}}$$

Mais on peut aussi considérer le cas d'une distribution de Laplace *dissymétrique* (cf. Figure 3).

$$(1.2.4)\quad f(x) = \frac{1}{\lambda+\mu}\, e^{-\frac{|x-x_0|}{\lambda}} \qquad \text{si} \qquad x \le x_0$$

$$f(x) = \frac{1}{\lambda+\mu}\, e^{-\frac{|x-x_0|}{\mu}} \qquad \text{si} \qquad x > x_0$$

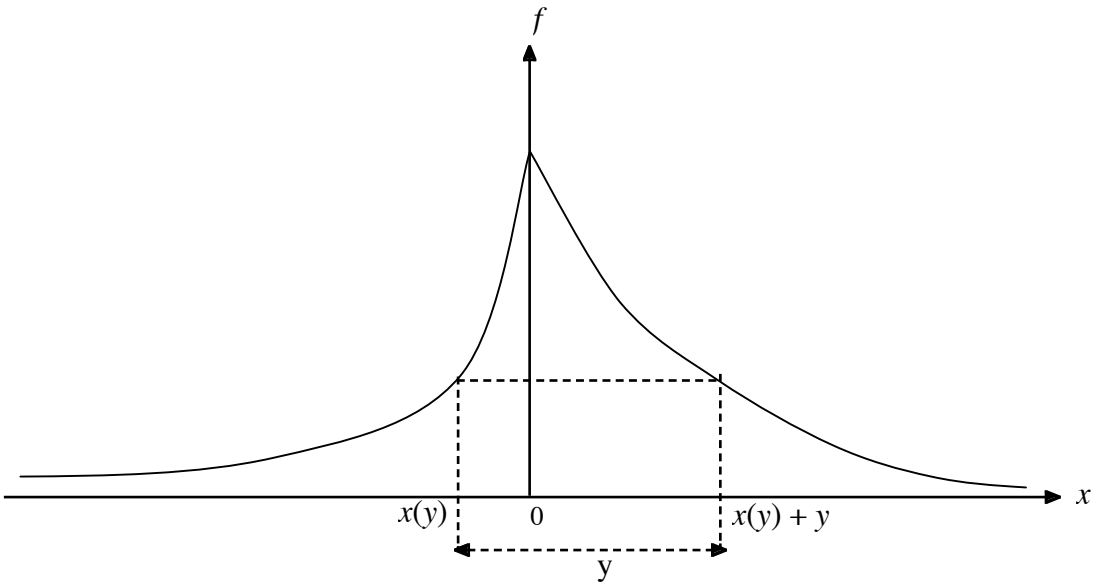


Figure 3

L'origine $x(y)$ de l'intervalle donnant le maximum de $\int_x^{x+y} f(t)\,dt$ est ici la solution unique de :

(4) $\quad f(x) = f(x+y)$

soit : $\quad x(y) = \dfrac{-\lambda y}{\lambda+\mu} \qquad$ et $\qquad x(y) + y = \dfrac{\mu y}{\lambda+\mu}$

d'où : $\quad Q(y) = 1 - e^{-\frac{y}{\lambda+\mu}} \qquad q(y) = \dfrac{1}{\lambda+\mu}\, e^{-\frac{y}{\lambda+\mu}}$

C'est la distribution exponentielle de paramètre $\nu = \lambda + \mu$; elle ne dépend que de ce seul paramètre (cf. Figure 4), et non des deux paramètres λ et μ.

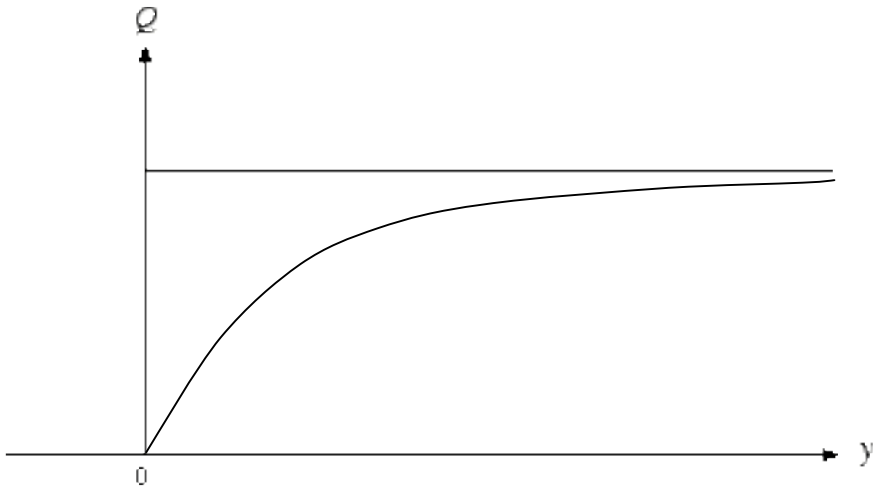


Figure 4

Autres exemples de distributions unimodales absolument continues et dissymétriques□

(1.2.5) $f(x) = \frac{2}{a}x$ pour $0 \le x \le a$

$f(x) = \frac{2}{1-a}(1-x)$ pour $a \le x \le 1$

$f(x) = 0$ partout ailleurs (cf. Figure 5).

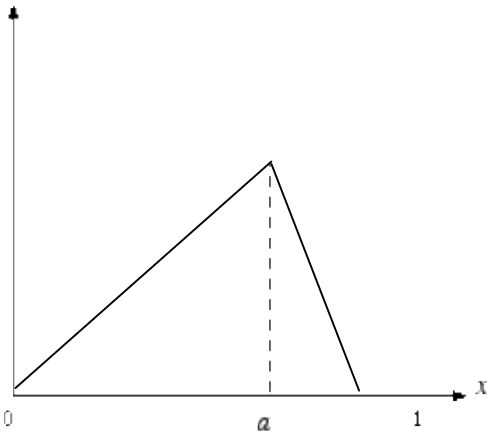


Figure 5

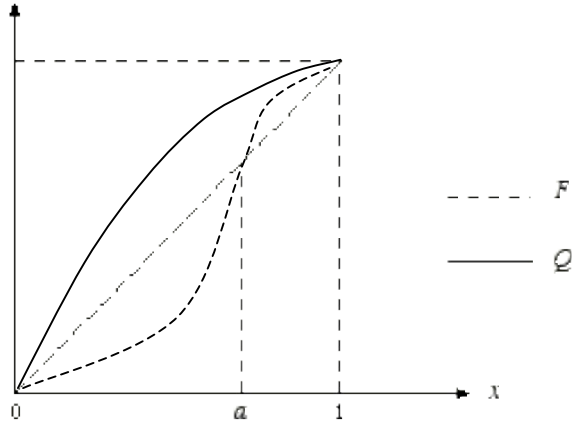


Figure 5 bis

Sa f.r. est□

$F(x) = \frac{x^2}{a}$ pour $0 \le x \le a$ et

$$F(x) = 1 - \frac{(1-x)^2}{1-a} \qquad \text{pour} \qquad a \le x \le 1$$ (cf. Figure 5bis).

Par un calcul simple, on obtient, pour $0 \le y \le 1$,

$$x(y) = a(1-y)$$

et $$Q(y) = y(2-y)$$

Sa courbe représentative (cf. Figure 5bis) est un arc de parabole, elle ne dépend pas du paramètre a.

Autre exemple laissé en exercice au lecteur.

(1.2.6) $$f(x) = \frac{2}{(c+d)} \frac{x}{a} \qquad \text{pour} \qquad 0 \le x \le a$$

$$f(x) = \frac{2}{c+d} \qquad \text{pour} \qquad a \le x \le b \quad (\text{avec} \quad b = a + d)$$

$$f(x) = \frac{2}{c+d} \frac{c-x}{c-b} \qquad \text{pour} \qquad b \le x \le d$$

$$f(x) = 0 \qquad \text{partout ailleurs (cf. Figure 6).}$$

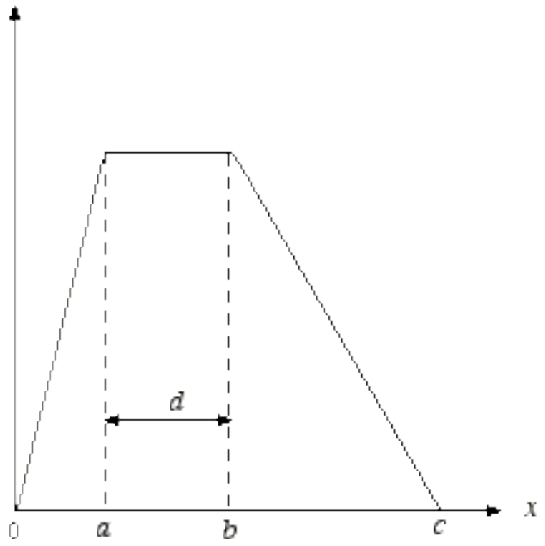


Figure 6

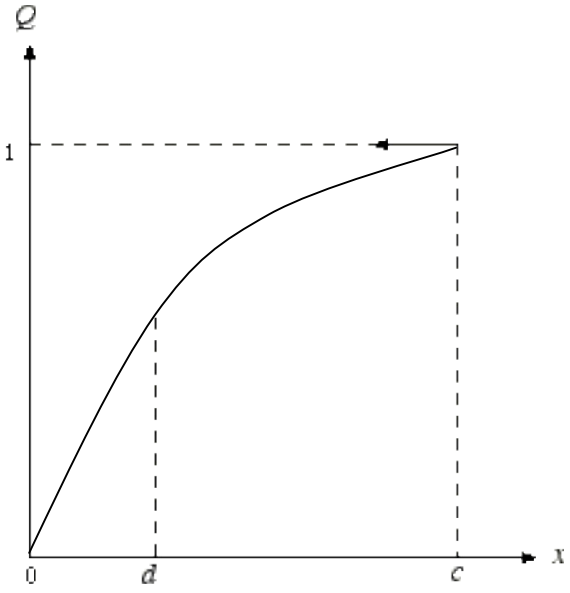


Figure 6 bis

La f.c. est (cf. Figure 6bis).

$$Q(y) = \frac{2y}{c+d} \qquad \text{pour} \qquad 0 \le y \le d$$

$$Q(y) = 1 - \frac{(c-y)^2}{c^2 - d^2} \qquad \text{pour} \qquad d \le y \le c$$

Elle ne dépend que des seuls paramètres c et d.

Voici maintenant trois derniers exemples de nature assez différente des précédents, qui tous concernaient des distributions unimodales.

(1.2.7) La loi «arc sinus», définie ici sur l'intervalle (-1, +1) par

$$f(x) = \frac{1}{\Pi} \cdot \frac{1}{\sqrt{1 - x^2}} \qquad F(x) = \frac{1}{2} + \frac{1}{\Pi} \arcsin x$$

C'est un premier cas déjà évoqué plus haut de «distribution en U», c'est-à-dire celles pour lesquelles la concentration se fait non autour d'une «valeur centrale», mais aux deux extrémités de l'intervalle de définition (cf. Figures 7 et 7 bis).

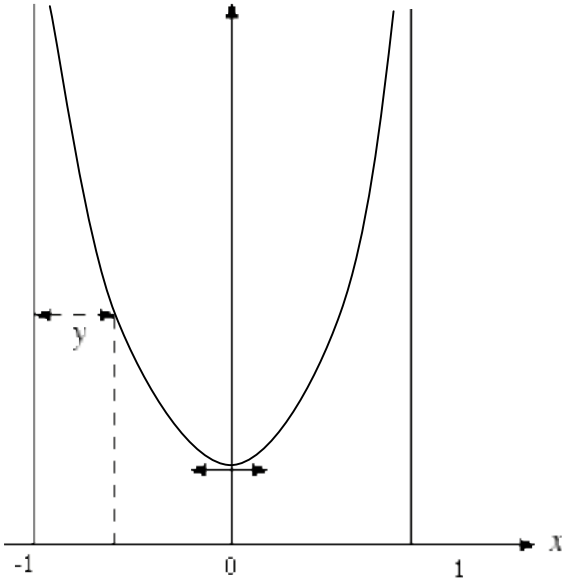


Figure 7

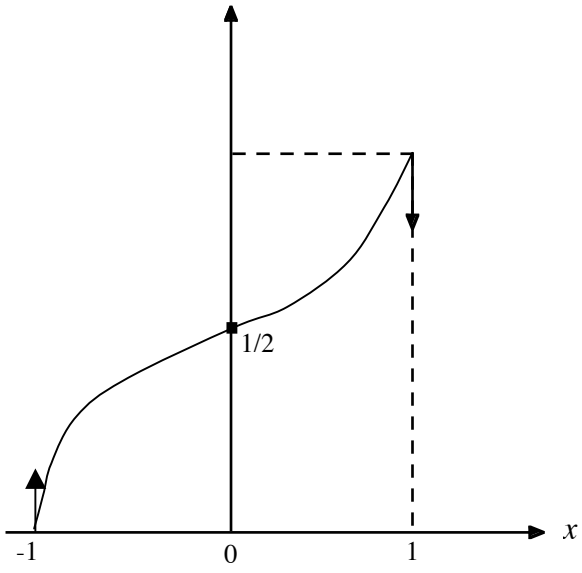


Figure 7 bis

Ici, en raison de la symétrie, il est clair que

$$H(x,y) = \int_x^{x+y} f(t)\ dt$$

est maximum pour $x = -1$. Par conséquent

$$Q_F(y) = F(y-1) = \frac{1}{2} + \frac{1}{\Pi} \arcsin (y-1) \qquad \text{avec} \qquad 0 \le y \le 2.$$

À une translation près, F est *sa propre fonction de concentration*, dont on remarquera en outre qu'elle n'est pas concave, contrairement à tous les exemples précédents.

(1.2.8) C'est, également le cas, par exemple, de F définie sur l'intervalle [0,1] par (Cf. Figure 8) :

$$F(x) = \frac{3}{2}\, x \qquad \text{pour} \qquad 0 \le x \le \frac{1}{3},$$

$$F(x) = \frac{1}{2} \qquad \text{pour} \qquad \frac{1}{3} \le x \le \frac{2}{3},$$

$$F(x) = \frac{3}{2}\, (x - \frac{1}{3}) \qquad \text{pour} \qquad \frac{2}{3} \le x \le 1$$

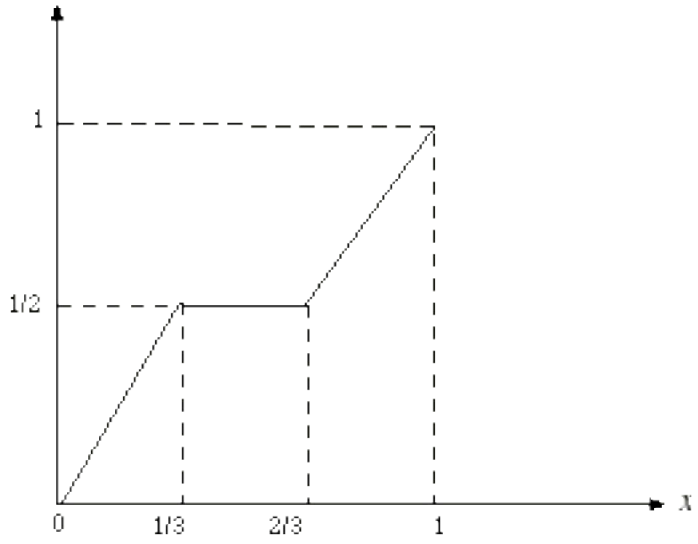


Figure 8

Elle est sa propre f.c., et n'est évidemment pas concave.

(1.2.9) On peut d'ailleurs itérer ce mode de définition d'une f.r. non concave qui est sa propre f.c., et construire (cf. Figure 9) F_2, F_3, …, F_n, … qui ont pour limite uniforme la fameuse f.r. continue, ayant pour support l'ensemble ternaire de Cantor (de mesure nulle) et non dérivable sur son support.

On verra ultérieurement que lorsqu'une suite F_n converge uniformément vers une f.r. limite F, la f.c. Q_{F_n} de F_n converge aussi uniformément vers celle Q_F de F ; la f.r. C « de Cantor » est donc une fonction de concentration.

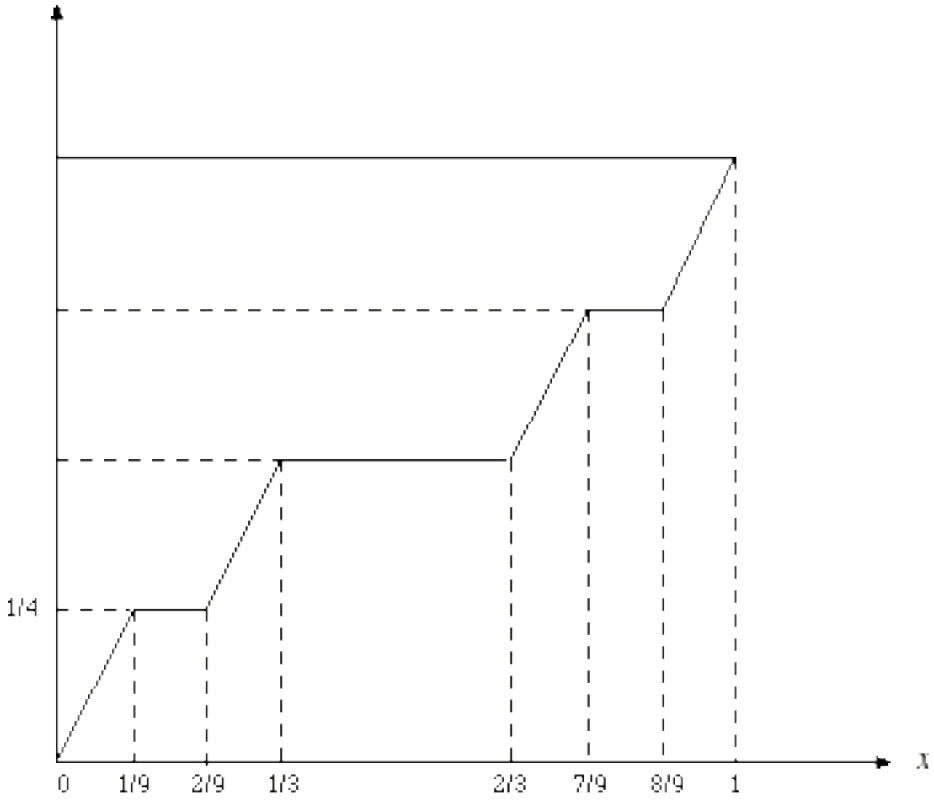


Figure 9

1.3 CONCENTRATION DE LÉVY ET CONCENTRATION DE LORENZ ET GINI

Il ne faut pas confondre la f.c. de P. Lévy avec la courbe de concentration de M.O. Lorenz et C. Gini, utilisée en statistique, dans l'étude des inégalités économiques ou sociales notamment.

En effet, la courbe de concentration de Lorenz et Gini est définie comme celle d'une f.r. F *par rapport* à une autre f.r. G, si on la note $Q_{F/G}$, $Q_{F/G}(y)$ est pour chaque y de l'intervalle fermé [0,1], le supremum de la mesure selon F des ensembles de mesure y selon G.

La concentration de Lorenz et Gini par rapport à la mesure uniforme sur un intervalle coïnciderait avec celle de Lévy à condition de remplacer dans la définition de celle-ci les intervalles de longueur y par les ensembles de mesure de Lebesgue égale à y. P. Lévy l'a du reste parfois fait (cf. T.A.V.A., ch. IV, par. 29).

D'ailleurs la courbe de concentration de Lorenz et Gini est toujours concave, alors qu'il n'en est pas du tout de même, on l'a vu, pour celle de P. Lévy.

Un exemple très simple illustre bien la différence.

Soit sur l'intervalle $[0, n]$ de $\mathbb{Z}$, ensemble des entiers, une distribution discrète $(p_0, p_1, \ldots, p_n)$ avec

$$p_0 > p_1 > p_2 > \ldots > p_n > 0$$

La fonction de concentration associée est :

$$Q_0 = p_0,\ Q_1 = p_0 + p_1,\ \ldots,\ Q_n = p_0 + p_1 + \ldots + p_n = 1$$

Moyennant un changement d'unité de rapport $\frac{1}{n}$, elle est f.c. de Lorenz et Gini (par rapport à la mesure uniforme) pour les $(n+1)!$ distributions obtenues en permutant les p_i de toutes les façons possibles.

Par contre, elle n'est f.c. de Lévy que pour les 2^n de ces distributions qui sont *unimodales*.

2. PRINCIPALES PROPRIÉTÉS DES f.c. DE PAUL LÉVY

2.1. CONCENTRATION D'UNE SOMME DE VARIABLES ALÉATOIRES INDÉPENDANTES (v.a.i.)

Soient X_1 et X_2 deux v.a.i. de f.r. F_1 et F_2 respectivement. Leur somme :

$$Z = X_1 + \mathrm{X}_2$$

a pour f.r. G le produit de convolution de F_1 et F_2 :

$$\forall x, \quad G(x) = \int_{-\infty}^{+\infty} F_1(x - t)\, d\, F_2(t)$$

d'où :

$$\leq \sup_x \left[F_1(x + y) - F_1(x)\right] \int_{-\infty}^{+\infty} d\, F(t) = \sup_x \left[F_1(x + y) - F_1(x)\right] = Q_{F_1}(y)$$

$$\forall x, \forall y \geq 0,\ G(x + y) - G(x) = \int_{-\infty}^{+\infty} \left[F_1(x + y - t) - F_1(x - t)\right] d\, F_2(t)$$

En faisant varier x dans le premier membre, il vient :

$$\forall y \geq 0,\ \sup_x \left(G(x + y) - G(x)\right) = Q_G(y) \leq Q_{F_1}(y)$$

On a de même : $Q_G(y) \leq Q_{F_2}(y)$.

D'où pour tout couple X_1, X_2 de v.a.i.

(5) $$Q_{X_1 + X_2} \leq \min\left(Q_{X_1}, Q_{X_2}\right)$$

Cette inégalité montre que, comme il se doit, la concentration n'augmente pas (elle diminue en général), par addition de v.a.i.

En outre, comme $Q_X(0) = 0$ ssi X est continue (cf. § 1.1 *supra*), (5) montre qu'il suffit qu'une seule v.a. X_i d'une somme ΣX_i soit continue pour que la somme le soit.

2.2. LIMITES DE f.c.

Soit F_n une suite de f.r. qui convergent *uniformément* vers F.

On peut donc choisir n_ε assez grand pour que :

$$\forall n > n_\varepsilon\ ,\ \forall t,\quad \left|F_n(t) - F(t)\right| < \frac{\varepsilon}{2}$$

On a d'autre part :

$$\forall n, \forall x, \forall y \geq 0,\ F_n(x+y) - F_n(x) = F_n(x+y) - F(x+y) + F(x+y) - F(x) + F(x) - F_n(x)$$

$$\leq \left|F_n(x+y) - F(x+y)\right| + F(x+y) - F(x) + \left|F(x) - F_n(x)\right|$$

Donc, pour $n > n_\varepsilon$:

$$\forall x, \forall y \geq 0,\ F_n(x+y) - F_n(x) \leq \frac{\varepsilon}{2} + F(x+y) - F(x) + \frac{\varepsilon}{2}$$

En prenant le supremum par rapport à x de chacun des deux membres, on obtient :

$$\forall \varepsilon > 0,\ \forall y \geq 0,\ n > n_\varepsilon \Rightarrow Q_n(y) \leq Q(y) + \varepsilon$$

Mais comme on a aussi :

$$F(x+y) - F(x) = F(x+y) - F_n(x+y) + F_n(x+y) - F_n(x) + F_n(x) - F(x)$$

On obtient également :

$$\forall \varepsilon > 0,\ \forall y \geq 0,\ n \leq n_\varepsilon \Rightarrow Q(y) \leq Q_n(y) + \varepsilon$$

D'où finalement :

(6) $$\forall \varepsilon > 0,\ \forall y \geq 0\ ,\ \forall n > n_\varepsilon,\quad \left|Q_n(y) - Q(y)\right| < \varepsilon$$

Q_n converge donc uniformément vers Q. La « fonctionnelle » $F \to Q_F$ est continue pour la métrique L_∞.

Ceci démontre en particulier le résultat indiqué *supra* § 1.2, exemple 1.2.9, concernant la f.r. « de Cantor », qui est sa propre f.c.

2.3. SOUS-ADDITIVITÉ DES f.c.

Soient y et z positifs ou nuls. On a :

$$\forall x\ ,\ F(y+z+x) - F(x) = F(y+z+x) - F(z+x) + F(z+x) - F(x)$$

En prenant le supremum de chacun des termes du second membre, il vient⊡

$$\forall x\ ,\ F(y+z+x)-F(x)\le Q_F(y)+Q_F(z)$$

D'où la propriété de *sous-additivité* des f.c.⊡

(7) $$Q_F(y+z)\le Q_F(y)+Q_F(z),\quad \forall y\ge 0,\ \ \forall z\ge 0$$

Cette relation (7) est condition *nécessaire* (on vient de le voir) mais aussi *suffisante* pour qu'une f.r. définie sur la demi-droite $x\ge 0$ soit une f.c.

En effet, soit F une telle f.r. (cf. Figure 10).

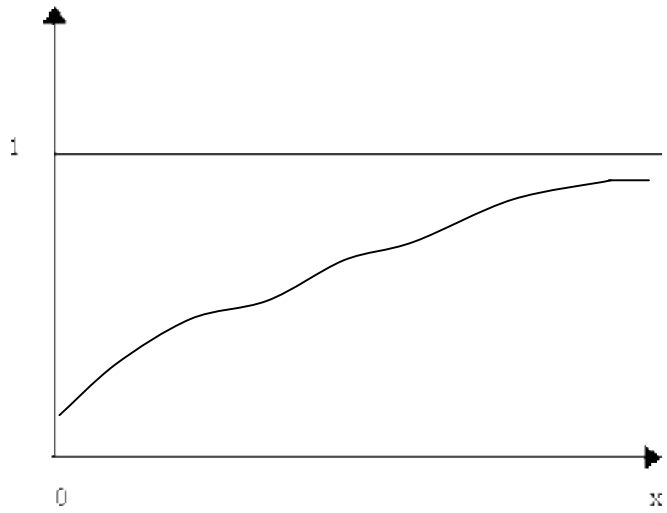


Figure 10

En reprenant la définition 1, (§ 1.1) de P. Lévy, il vient⊡

$$\forall\ y\ge 0,\ \ Q_F(y)=\max_x\left[\overline{F}(x+y)-\underline{F}(x)\right]$$

Comme $\underline{F}(0)=0$, il vient⊡

$$\forall y\ge 0,\quad Q_F(y)\ge \overline{F}(y)$$

Si d'autre part F satisfait à la condition (7) de sous-additivité, on a⊡

$$\forall x\ge 0,\ \forall y\ge 0,\quad F(x+y)-F(x)\le F(y)$$

D'où, en prenant le supremum du premier membre⊡ $\forall y\ \ge 0,\quad Q_F(y)\le F(y)$

Comme, par définition⊡ $F(y)\le \overline{F}(y)$

On a finalement⊡

$$Q_F(y)=F(y)=\overline{F}(y)$$

La «fonctionnelle» $F \to Q_F$ est donc *idempotente*, et la *sous-additivité* (7) est la *propriété caractéristique* des f.c. (parmi les f.r. à support ≥ 0).

Il résulte en outre de (7) qu'il suffit, pour une f.c., d'être continue à l'origine pour l'être partout.

De plus, cette inégalité montre que si Q est dérivable à l'origine, en tout autre point où elle est dérivable, sa dérivée à droite est au plus égales à $Q'(0)$.

Bien entendu, toute f.r. à support ≥ 0 et concave satisfait à (7) et est donc une f.c.

Mais, comme l'ont montré les exemples (1.2.7), (1.2.8) et (1.2.9) *supra*, la réciproque n'est pas vraie.

2.4 CAS DES v.a. À SUPPORT BORNÉ

Nous supposons maintenant que le support de la v.a. X et de sa f.r. F est borné. Nous supposerons en outre que F est définie sur l'intervalle fermé $[0,1]$ de $\mathbb{R}$ (on peut se ramener à ce cas par une transformation affine sur X).

La propriété (7) de sous-additivité a quelques conséquences supplémentaires.

D'abord, pour tout entier naturel n et tout $y > 0$, il vient :

$$Q_F(ny) \leq n\, Q_F(y)$$

D'où, pour $y = \frac{1}{n}$:

$$1 = Q_F(1) \leq n\, Q_F\left(\frac{1}{n}\right)$$

Soit :

(8) $$Q_F\left(\frac{1}{n}\right) \geq \frac{1}{n}\ , \quad \forall n \in \mathbb{N}$$

En particulier : $Q_F\left(\frac{1}{2}\right) \geq \frac{1}{2}$.

La médiane μ de Q est inférieure ou égale à $\frac{1}{2}$, car :

$$Q_F(\mu) = \frac{1}{2} \leq Q_F\left(\frac{1}{2}\right)$$

D'après (8), la même propriété vaut pour tous les premiers quantiles.

D'autre part, comme on l'a vu : $Q_F(y) \geq F(y)$.

Posons : $x = 1 - y$ dans l'expression $(F(x + y) - F(x))$.

Il vient : $F(x + y) - F(x) = F(1) - F(1 - y) = 1 - F(1 - y)$

d'où $Q_F(y) \geq 1 - F(1-y)$

et

(9) $\forall y, \quad 1 \geq y \geq 0, \quad Q_F(y) \geq \max(F(y), 1 - F(1-y))$

$1 - F(1-y) = F^*(y)$ est la f.r. dont la courbe représentative est symétrique de celle de F par rapport au point $\left(\frac{1}{2}, \frac{1}{2}\right)$ (cf. Figure 11).

Mais de (9) il résulte en outre

$$\forall y, \quad 0 \leq y \leq 1, \qquad Q_F(y) + Q_F(1-y) \geq Q_F(y) + F(1-y) \geq 1$$
$$Q_F(y) + Q_F(1-y) \geq F(y) + Q_F(1-y) \geq 1$$

D'où

(10) $\forall y, 0 \leq y \leq 1, \qquad Q_F(y) \geq F(y) \geq 1 - Q_F(1-y) = Q_F^*(y)$

La même relation vaut pour F^*.

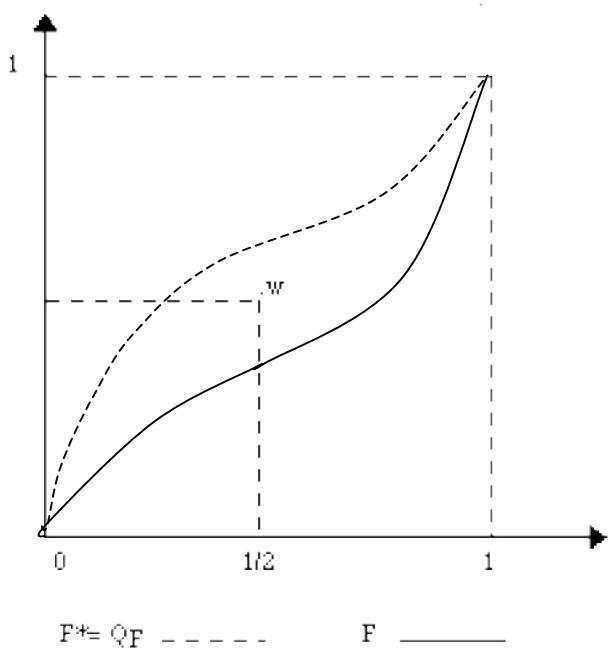


Figure 11

En particulier, si Q est une f.c. symétrique par rapport au point ω de coordonnées $\left(\frac{1}{2}, \frac{1}{2}\right)$, comme dans les exemples (1.2.7), (1.2.8) et (1.2.9) *supra*, elle est sa propre f.c. et aucune autre f.r. ne l'admet comme f.c. ; elle est un « point fixe » pour la fonctionnelle $F \to Q_F$.

$$Q_F = \mathrm{Q} = \mathrm{Q}^* \Rightarrow F = Q$$

C'est notamment le cas pour toute distribution uniforme sur un intervalle.

3. DISTRIBUTIONS UNIMODALES ABSOLUMENT CONTINUES

3.1 Nous supposons la densité f de F continue et strictement croissante avant le mode x_0, strictement décroissante au-delà. Ce cas est celui qui est le plus fréquemment utilisé en statistique. On appellera «□famille C□» (C pour «□en cloche□») cette famille de distributions.

Si l'on pose□

2 $G(x, y) = F(x) - F(x - y)$

Le maximum $Q_F(y)$ de G quand x parcourt $\mathbb{R}$ est atteint, comme on l'a vu dans plusieurs exemples (1.2.1 à 1.2.6. *supra*) pour une valeur $x(y)$ bien déterminée de x, solution unique de l'équation $\frac{\partial G}{\partial x} = 0$, soit (cf. Figure 12)□

2 $f(x) = f(x - y)$

et l'on a□

$$Q_F(y) = G(x(y), y)$$

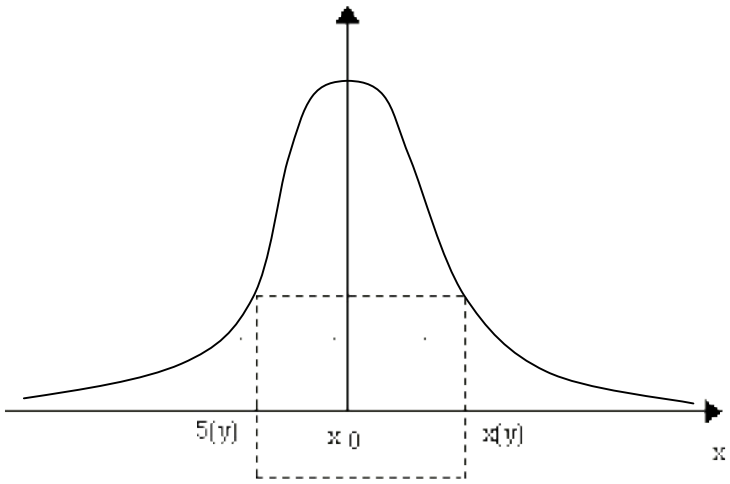


Figure 12

La fonction $x(y)$ est continue, monotone strictement croissante, et la fonction $x(y) - y$ est monotone décroissante□on a en outre $x(0) = x_0$.

Des inégalités de monotonie□

$$\forall h > 0,\ x(y+h) > x(y)$$

$$x(y+h) - (y+h) < x(y) - y$$

résulte immédiatement la double inégalité qui résume la monotonie et la continuité de $x(y)$.

(13) $\quad \forall y \geq 0, \quad \forall h > 0, \qquad 0 < x(y+h) - x(y) < h.$

D'autre part, $G(x, y)$ est dérivable en x et en y, d'après sa définition (11). Et l'on a.

$$\frac{\partial G}{\partial x} = F'(x) - F'(x-y) = f(x) - f(x-y)$$

$$\frac{\partial G}{\partial y} = -F'(x-y) = f(x-y)$$

d'où. $dG = (f(x) - f(x-y))\,dx + f(x-y)\,dy.$

Sous la condition (12), $G = Q_F$ il reste.

$$dQ_F(y) = f(x(y) - y)\,dy = f(x(y))\,dy$$

Donc Q_F est absolument continue, de densité.

(14) $\quad Q_F'(y) = q_F(y) = f(x(y))$

La densité de Q_F est donc monotone décroissante, et Q_F est concave (cf. Figure 13).

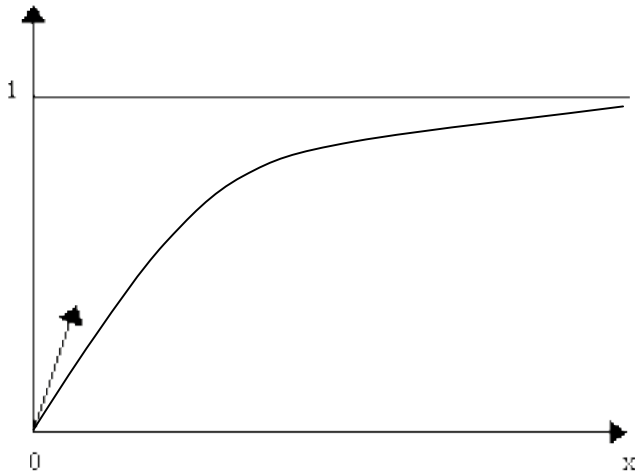


Figure 13

La pente de Q à l'origine est celle de F à son point d'inflexion.

La f.c. Q_F résume de façon simple l'essentiel de ce qu'il y a à dire sur la concentration et l'inégalité de la distribution de la f.r. F, et elle en dit beaucoup plus

qu'une simple mesure de la dispersion (par la variance, ou tel autre indicateur numérique).

3.2 INVERSION DE LA F.C.

Les remarques résumées par les relations (13) et (14) ci-dessus permettent, inversement, de déterminer toutes les distributions de la famille C qui admettent une fonction Q *donnée*, concave, absolument continue sur la demi-droite $y \geq 0$ (cf. Figure 13) comme fonction de concentration.

Soit donc $g(y)$ une fonction de $\mathbb{R}^+$ dans $\mathbb{R}$ satisfaisant à $g(0) = 0$ et à la double inégalité⊡

(14) $\forall y \geq 0, \quad \forall h > 0, \quad 0 < g(y + h) - g(y) < h$

Dans ces conditions, g est strictement croissante (première inégalité) et continue (deuxième inégalité).

On a en outre, $0 < g(y) - g(0) < y$

Soit⊡ $\forall y, \quad g(y) < y.$

Posons⊡ $\gamma(y) = g(y) - y.$

Il résulte de (14) que $\gamma(y)$ est monotone, strictement décroissante et continue, puisque⊡

$$\forall h > 0 \qquad -h < \gamma(y + h) - \gamma(y) = g(y + h) - g(y) - h < 0$$

et $$\forall y \geq 0 \qquad -y < \gamma(y) < 0.$$

La fonction g n'est soumise qu'aux conditions explicitées ci-dessus, et est donc pour le reste totalement arbitraire (cf. Figure 14).

Désignons par g^{-1} et γ^{-1} les fonctions inverses de g et γ respectivement, et posons, par *définition de la densité* f⊡

$$\forall x \geq 0, \; f(x) = q(g^{-1}(x))$$

où q est la densité de Q *donnée* et pour⊡ $x < 0, \; f(x) = q(\gamma^{-1}(x))$.

Lorsque le support de Q est borné, de maximum m, on a⊡ $q(y) = 0$ pour $y > m$.

Ainsi définie, f est monotone strictement croissante pour $x < 0$, et strictement décroissante pour $x > 0$, puisque q est une fonction strictement décroissante de son argument. Si Q est à support borné, f s'annule à l'extérieur d'un intervalle fini.

On a d'ailleurs⊡ $f(0) = q(0) > q(y), \; \forall y \geq 0.$

Donc f a son mode en 0

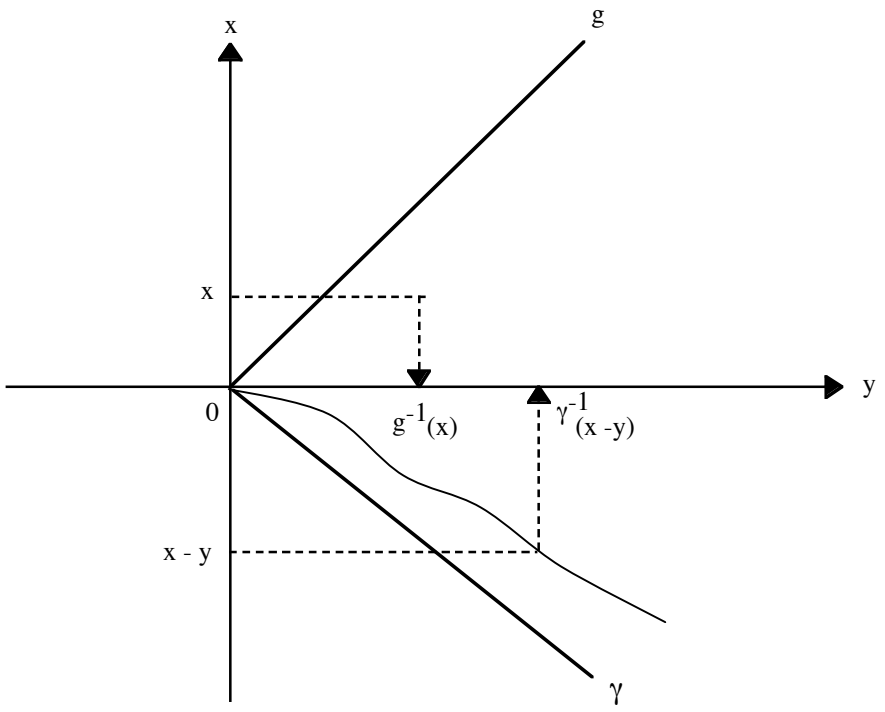


Figure 14

Montrons enfin que l'équation⊡

(15) $f(x) = f(x - y)$

a pour solution unique⊡ $x = g(y)$.

En effet (15) ne saurait être satisfaite si x et $x - y$ sont tous deux du même signe, à cause de la monotonie stricte de f.

Lorsque $x > 0 > x - y$, on a⊡

$$f(x) = q(g^{-1}(x)) \qquad f(x - y) = q(\gamma^{-1}(x - y))$$

Donc (15) est vérifiée ssi⊡

$$g^{-1}(x) = \gamma^{-1}(x - y) \quad \text{avec} \quad (0 < x < y)$$

Or, lorsque x, à y constant, croît de 0 à y, $g^{-1}(x)$ monotone strictement et continue, croît de 0 à $g^{-1}(y)$.

De même, $\gamma^{-1}(x - y)$ décroît de $\gamma^{-1}(-y)$ à 0. Il existe donc un point unique (x, u) de rencontre de leurs courbes représentatives, avec⊡

$$u = g^{-1}(x) = \gamma^{-1}(x - y)$$

c'est-à-dire⊡

(16) $g(u) = x$ et $\gamma(u) = x - y$, d'où⊡

(17) $\gamma(u) = g(u) - y$

Mais, par définition de γ⊡

(18) $\gamma(u) = g(u) - u$

En rapprochant (17) de (18), il vient⊡ $u = y$

et par (16)⊡ $g(y) = x$

On vient de voir que dans le cas étudié il y a une infinité de solutions, dépendant d'une fonction arbitraire, au problème de l'inversion de la f.c.. Ceci pouvait se prévoir intuitivement par le raisonnement suivant (cf. § 1.3. *supra*).

Donnons-nous sur l'ensemble $\mathbb{N}_0$ des entiers naturels, 0 compris, une distribution de poids q_n tels que⊡

$$q_0 > q_1 > q_2 > \ldots > q_n > \ldots, \qquad \sum_0^\infty q_n = 1$$

Posons⊡ $Q_n = \sum_0^n q_i$

Q_n est f.c. de toutes les distributions unimodales sur $\mathbb{Z}$ obtenues par le procédé suivant⊡

poser $p_0 = q_0$

choisir une suite n_k monotone croissante d'entiers naturels, par ailleurs quelconques

désigner par $m_1, m_2, \ldots m_h, \ldots$ la suite monotone des entiers naturels qui n'ont pas été choisis à l'étape précédente

poser, pour tout $k > 0, p_k = q_{n_k}$ et $p_{-k} = q_{m_k}$.

Remplaçons maintenant $\mathbb{Z}$ par une division de $\mathbb{R}$ en intervalles I_n de même longueur ε, et chaque p_n par une répartition uniforme sur I_n, tout en conservant l'unimodalité de la distribution en escalier ainsi obtenue.

Mutatis mutandis, le procédé ci-dessus s'appliquera.

En faisant tendre ε vers zéro, on comprend intuitivement que, dans le cas des f.r. absolument continues unimodales, il y en aura une infinité dépendant d'une *fonction monotone arbitraire* qui admettront la même f.c.

4. DISTRIBUTIONS EN U ABSOLUMENT CONTINUES

Le cas des f.r. de densité «en cloche», où la répartition est concentrée autour d'une seule valeur (valeur centrale, ou éventuellement l'une des extrémités de l'intervalle de définition) n'est pas le seul intéressant pour la statistique.

Sont également très intéressantes, et pourraient être plus souvent utilisées, les distributions «en U» où la répartition se concentre au contraire au voisinage des deux extrémités de l'intervalle de définition.

On appelle ici par commodité «famille U» l'ensemble des distributions absolument continues ayant pour support un intervalle, (que l'on peut ramener à l'intervalle [0,1] par une transformation affine) de densité f continue, strictement décroissantes de 0 à un minimum atteint en x_0 (avec $0 < x_0 < 1$) strictement croissantes ensuite (cf. Figures 15 et 15bis)

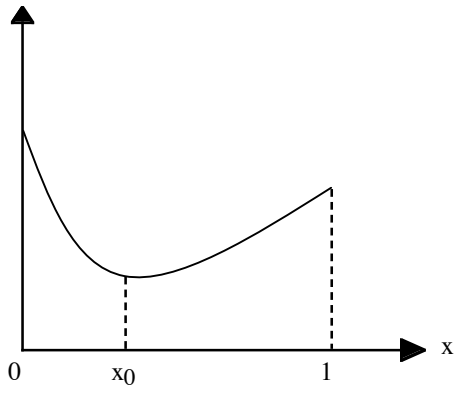


Figure 15

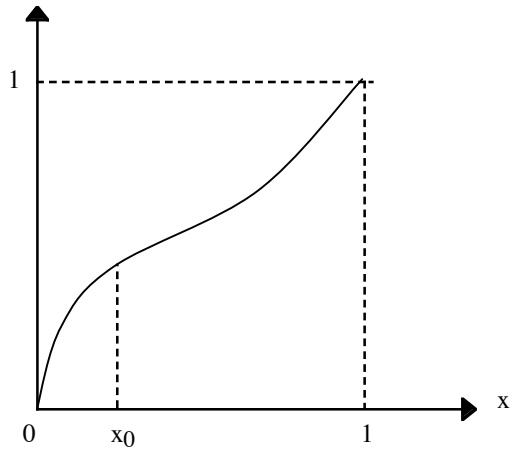


Figure 15bis

N.B. Dans le cas particulier où x_0 aurait été égal à 0 ou 1 (cf. Figure 16), on serait dans le cas précédent. F serait concave, serait sa propre f.c., et aurait une infinité d'inverses «en cloche» construites comme on l'a vu en 3.2 *supra*.

Dans le cas de la famille U, l'équation $f(x + y) = f(x)$ admet au plus une solution; celle-ci existe pour tout y $(0 \le y \le 1)$ dans le cas particulier où $f(0) = f(1)$. Dans le cas contraire, et si, pour fixer les idées, $f(0) > f(1)$, la solution $x(y)$ n'existe que pour

$$0 \le y \le 1 - u$$

où u est la solution de $f(u) = f(1)$ $(0 < u < x_0)$.

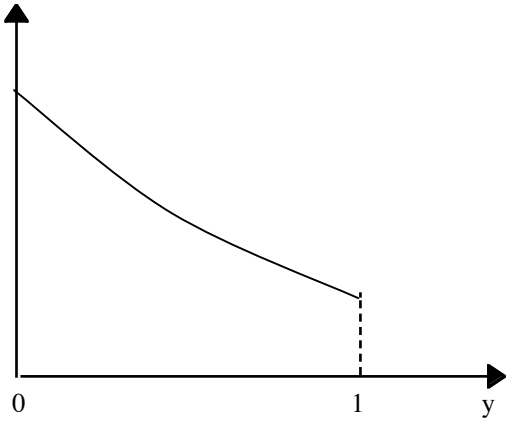


Figure 16

En posant :

$$G(x, y) = F(x + y) - F(x)$$

on a, à y constant :

$$dG = (f(x + y) - f(x))\, dx$$

G est donc décroissante tant que $f(x + y) < f(x)$; elle passe éventuellement par un minimum pour la valeur $x = x(y)$ si celle-ci existe, et est ensuite croissante. On a donc :

$$\forall y \geq 0, \quad Q_F(y) = \max\{G(0, y), G(1 - y, 1)\}$$

c'est-à-dire :

$$(19) \qquad Q_F(y) = \max(F(y), F^*(y))$$

Où $F^*(y) = 1 - F(1 - y)$ est la symétrique de F par rapport au point $\omega\left(\frac{1}{2}, \frac{1}{2}\right)$, et $f^*(y) = f(1 - y)$ celle de f par rapport à l'axe $x = \frac{1}{2}$.

Dès lors, deux cas sont possibles :

1. Ou bien : $\forall x, \quad F(x) \geq F^*(x)$ (ou, dualement $F(x) \leq F^*(x)$)
on a alors :

$$Q_F = F \quad (\text{resp. } Q_F = F^*)$$

Cela signifie que la courbe représentative de Γ de F et celle de Γ^* de F^*, qui sont symétriques l'une de l'autre par rapport au point ω, ne se rencontrent qu'à l'origine, au point I(1,1) et éventuellement en ω (cf. Figure 11, p. 46, *supra*).

Les distributions vérifiant ces conditions constituent une première sous-famille de U, appelons la U_1.

2. Ou bien les deux courbes Γ et Γ^* se rencontrent en d'autres points.

Il en est ainsi, par exemple, pour F définie sur le segment $[0,1]$ par

$$F(x) = \frac{x^4}{4} - \frac{4x^2}{9} + \frac{43x}{36} \qquad \text{(cf. Figure 17)}$$

de densité $f(x) = x^3 - \frac{8x}{9} + \frac{43}{36}$ (cf. Figure 17bis)

Celle-ci a son minimum en $x_0 = \frac{2}{3}\sqrt{\frac{2}{3}}$, compris entre $\frac{1}{2}$ et $\frac{2}{3}$, la courbe représentative de f coupe celle de f^* en $x_1 = \frac{1}{2} - \frac{\sqrt{5}}{6}$ et $1 - x_1 = \frac{1}{2} + \frac{\sqrt{5}}{6}$ (et évidemment en $x = \frac{1}{2}$). Quant à F, sa courbe représentative Γ coupe celle Γ^* de F^* aux points d'abscisse $\frac{1}{3}$ et $\frac{2}{3}$ (cf. Figures 17 et 17bis).

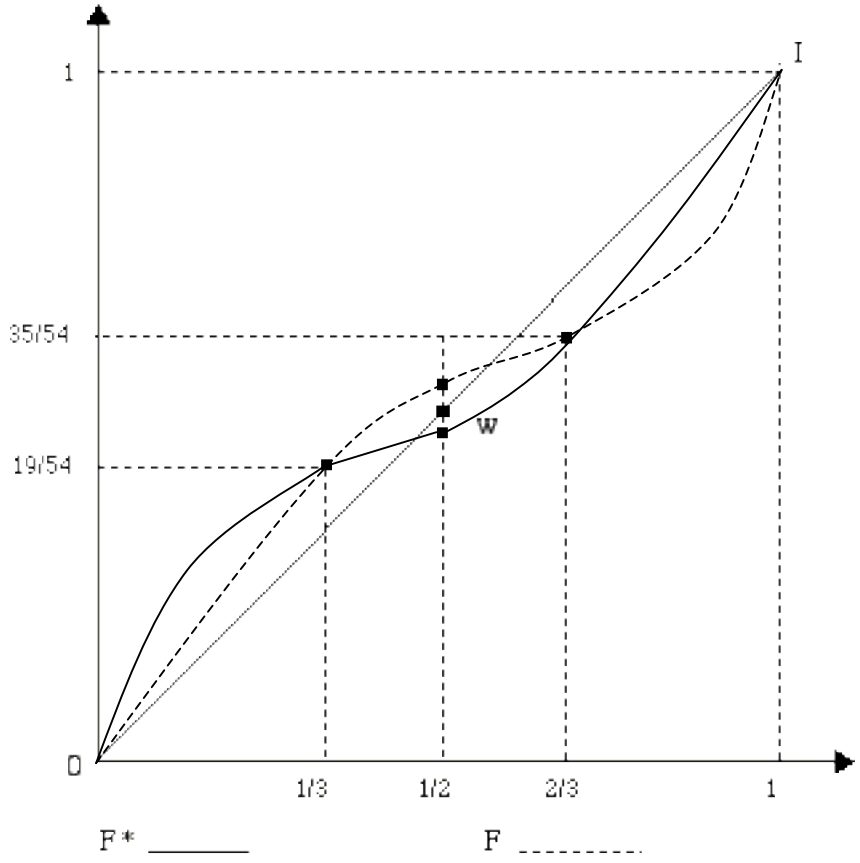


Figure 17

Les f.r. en U pour lesquelles il y a ainsi plusieurs points d'intersection, en nombre pair, autres que 0, I et éventuellement ω, entre Γ et Γ^*, seront dites constituer la famille U_2.

La fonction de concentration associée Q_F est alternativement égale soit à F, soit à F^*, selon celle de ces deux valeurs qui est la plus petite en chaque point x de Γ et de Γ^*.

Sa densité a un saut en chacun des points d'intersection (cf. Figure 17bis), de sorte qu'elle-même n'appartient pas à la famille U.

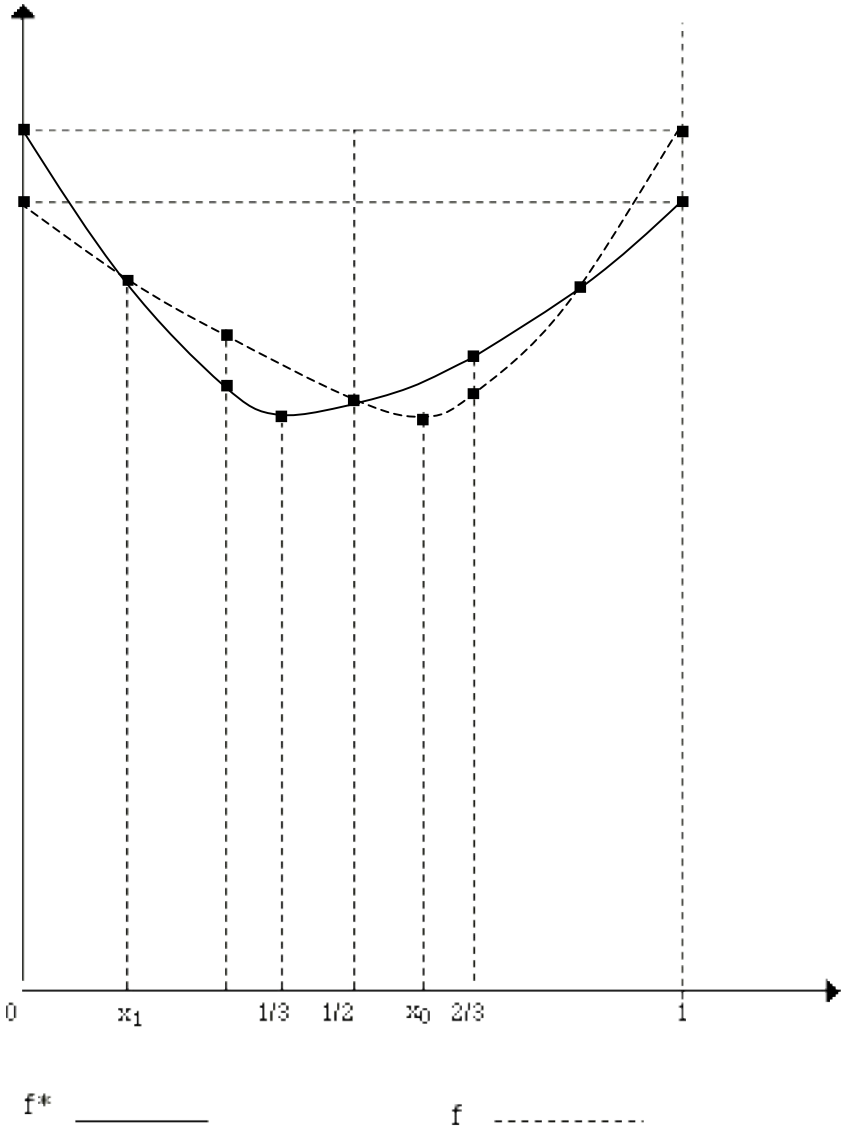


Figure 17bis

Dès lors, la question de l'inversion des f.c. de f.r. de la famille U est immédiatement résolue.

Si Q est de la famille U_1 elle est f.c. d'elle-même et de sa symétrique Q^*, et de celles-ci seulement ; puisque sa densité est continue, elle ne peut être f.c. d'une f.r. de U_2 ; et toute autre f.r. de U_1 est sa propre f.c., donc celle-ci diffère de Q.

Si Q est f.c. d'une f.r. de U_2, c'est par définition qu'il existe une f.r. H telle que

$$Q = \max (H, H^*)$$

les courbes représentatives de H et H^* se coupant en plusieurs points, en nombre pair, autres que 0 et I.

Soit F une inverse de Q distincte de Q, Q^*, H et H^*. Si F appartient à la famille U_1, elle (ou sa symétrique F^*) est sa propre f.c., et ne peut donc admettre Q comme f.c.

Donc F appartient à la famille U_2, et sa f.c. est

$$Q_F = \max (F, F^*) = Q = \max (H, H^*)$$

Soit x un point du segment [0,1]. En ce point, on a donc nécessairement $F(x) = H(x)$ (et $F^*(x) = H^*(x)$) ou $F(x) = H^*(x)$ (et $F^*(x) = H(x)$).

Supposons, pour fixer les idées, que $F(x) = H(x)$.

En tout point d'intersection, les quatre valeurs de F, F^*, H, H^* coïncident.

Si x n'est pas un point d'intersection, il est situé entre deux tels points u et v. Dans l'intervalle $u < x < v$ le signe de $H(x) - H^*(x)$ est constant. Donc, F étant continue, on aura $F(x) = H(x)$ dans tout cet intervalle.

Quand x dépasse un point d'intersection v, $Q(x)$ devient égale à $H^*(x)$. Mais si $F(x)$ subissait le même changement, sa densité aurait une discontinuité en v, ce qui est contraire à la définition de la famille U (cf. Figures 17 et 17bis).

Donc, Q n'a, dans ce cas, que quatre inverses : Q elle-même, sa symétrique Q^*, H et H^*. Seules les deux dernières appartiennent à la famille U.

En résumé, la f.c. d'une f.r. F de la famille U ne peut avoir pour inverses dans cette famille que F et sa symétrique F^* ; lorsqu'en particulier F est symétrique, elle est sa propre f.c., et la seule inverse de celle-ci (cas des exemples 1.2.7, 1.2.8 et 1.2.9 du § I.2).

Les résultats des deux paragraphes 3 et 4 ci-dessus montrent que la f.c. permet de séparer nettement le comportement de deux grandes familles de variables quant à leur dispersion.

D'autre part, on comparera la dispersion de deux variables X_1 et X_2 au moyen de l'ordre de majoration uniforme :

$$(20) \qquad Q_{X_1} \le Q_{X_2} \iff \forall y \ge 0,\ Q_{X_1}(y) \le Q_{X_2}(y)$$

Comme cet ordre est partiel, il arrive que les courbes représentatives des deux f.c. se coupent. C'est le cas, par exemple, pour les deux distributions Beta et Arc sinus

évoquées en introduction (cf. p. 32)□ pour y petit, c'est l'Arc sinus qui est le plus concentré, alors qu'au-delà de l'abscisse du point d'intersection de leurs f.c., et notamment pour $y > \frac{1}{2}$, c'est la distribution Beta.

Mais cette circonstance, justement, fournit une information très pertinente.

L'étude qui précède ne prétend pas épuiser la question de l'inversion des f.c.□il conviendrait notamment d'étudier plus avant le cas des distributions discrètes unimodales, pour les applications à la statistique notamment, ou celui des distributions bi-modales (en «□dos de chameau□»).

D'autre part, l'extension de la notion de f.c. aux variables du plan $\mathbb{R}^2$, ou a fortiori d'un espace $\mathbb{R}^n$ quelconque, présente quelques difficultés□plusieurs définitions sont possibles, qui conduisent à des résultats différents ou ne sont adaptées qu'à un type déterminé de f.r. (par exemple les «□ellipses de concentration□» pour les distributions de Laplace-Gauss)□; des mathématiciens de l'école roumaine (Radu Theodorescu, notamment[4]) sont ceux qui, à ma connaissance, sont allés le plus loin dans cette voie.

[4] Cf. W. Hentgartner, R. Theodorescu, *Concentration functions*, Academis Press, New York, 1973.