21 juin 2013
Journées de l'ANR Boole
Notes d'exposés

Sur le problème de la détermination des moments, avec application aux grandes urnes de Pólya

1 Le problème des moments

1.1 La question, premiers exemples

Poser le problème des moments d'une variable aléatoire réelle : existence et unicité. On ne s'occupe ici que de l'unicité.
La situation type (probabilités, statistiques, analyse d'algorithmes, ...) : on montre une convergence (en loi) d'une suite de variables aléatoires réelles en montrant la convergence des moments (cela suffit). La variable limite X a une suite de moments

$$E\left(X^p\right) = \int_{\mathbb{R}} t^p d\mu_X(t), \quad p = 1, 2, 3 \dots$$

Avec un peu de chance, on reconnaît la suite des moments de sa distribution de probabilité préférée, qu'on note B comme Boole (on désigne par une même lettre une variable aléatoire et sa loi). Autrement dit, pour tout $p \in \mathbb{N}$, on a

$$E\left(X^p\right) = E\left(B^p\right).$$

Question : X et B sont-elles distribuées pareil ?
Autrement dit, définissent-elles la même mesure de probabilité sur $\mathbb{R}$? Une autre manière de poser la question : est-il vrai qu'alors, pour tout $x \in \mathbb{R}$,

$$\mathrm{P}\left(X \leq x\right) = \mathrm{P}\left(B \leq x\right) \ ?$$

Lorsque la réponse est oui (pour n'importe quelle loi X), on dit que la loi B *est déterminée par ses moments*. Savoir qu'une loi est déterminée par ses moments est intéressant, par exemple pour identifier une distribution dont on a pu calculer les moments.

Premier cas pour débuter lentement : on suppose que le support de B est fini. Là, la réponse est oui, par la théorie des polynômes symétriques (si $\sum x_k^p = \sum y_k^p$ pour tout p, les polynômes symétriques élémentaires des x_k et des y_k sont les mêmes, donc $\prod(X - x_k) = \prod(X - y_k)$). *Une mesure de probabilité à support fini est toujours déterminée par ses moments.*

Second cas : on suppose que le support de B est compact (problème de Hausdorff). La réponse est encore oui, par Stone Weierstrass (si $\int x^p d\mu = \int x^p d\nu$ pour tout p, alors par densité uniforme, $\int f d\mu = \int f d\nu$ pour toute f continue ; par ailleurs, si

$\int f d\mu = 0$ pour toute f continue, alors $\mu([0,x]) = 0$ pour tout x et les intervalles $[0,x]$ engendrent les boréliens). *Une mesure de probabilité à support compact est toujours déterminée par ses moments.*
Par exemple, les lois Bêta sont déterminées par leurs moments puisque leur support est $[0,1]$.

1.2 Cas des supports non bornés

Si le support de B est inclus dans une demi-droite, c'est le problème de Stieltjes. La question sans contrainte sur le support de B, c'est celui de Hamburger.
Dans le cas d'un support non borné, la réponse est non : il existe des mesures de probabilité qui ne sont pas déterminées par leurs moments. On en sait davantage : il n'est pas très difficile de montrer qu'une mesure de probabilité ayant tous ses moments étant donnée, l'ensemble des mesures de probabilité ayant les mêmes moments est convexe. En particulier, si loi n'est pas déterminée par ses moments, il existe une famille non dénombrable de lois qui ont les mêmes moments.

Exercice de calcul intégral Si $w > 0$, et si $\Re(z) > 0$,

$$\int_0^{+\infty} t^{w-1} e^{-zt} dt = \frac{1}{z^w} \Gamma(w).$$

[Il suffit de le voir pour z réel et de prolonger analytiquement ; changer de variable.]
Du coup, pour tout $p \in \mathbb{N}$,

$$\int_0^{+\infty} x^{3p-1/2} e^{-x} \cos\left(x\sqrt{3} + \frac{\pi}{3}\right) dx = \frac{\Gamma\left(3p - \frac{1}{2}\right)}{2^{3p-1/2}} \Re\left((-1)^p i\right) = 0.$$

En changeant de variable, si on pose

$$g(x) = \frac{2}{3\sqrt{\pi}} |x|^{-2/3} \exp\left(-|x|^{2/3}\right) \cos\left(\frac{\pi}{3} + \sqrt{3}|x|^{2/3}\right),$$

alors pour tout $p \in \mathbb{N}$,

$$\int_{-\infty}^{+\infty} x^p g(x) dx = 0$$

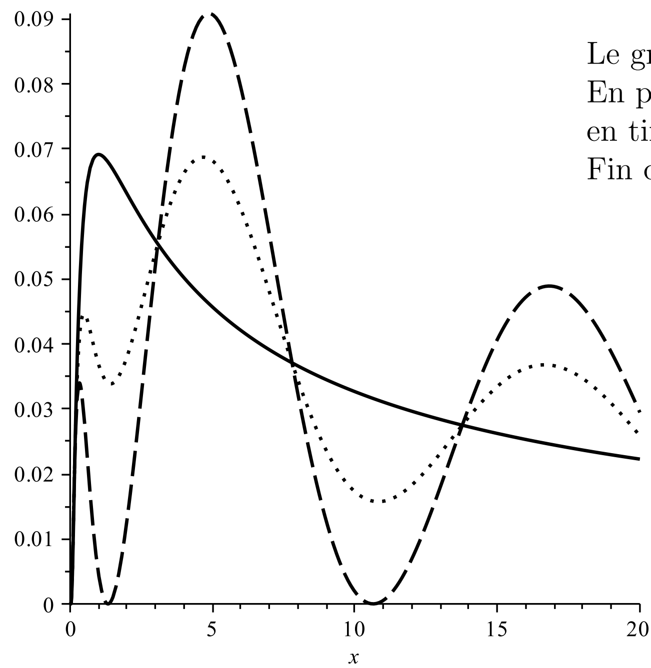
(c'est trivial si p est impair).
Ca y est, nous y sommes : soit f la fonction

$$f(x) = \frac{1}{3\sqrt{\pi}} |x|^{-2/3} \exp\left(-|x|^{2/3}\right).$$

C'est une densité de probabilité sur $\mathbb{R}$. Par ailleurs, si $\rho \in [0, 1/2]$, la fonction $f + \rho g$ est positive sur $\mathbb{R}$, et pour tout $p \in \mathbb{N}$,

$$\int_{\mathbb{R}} x^p f(x) dx = \int_{\mathbb{R}} x^p (f + \rho g)(x) dx \ :$$

les densités de probabilité f et $f + \rho g$ ont les mêmes moments.



Le graphe de f est en traits pleins.
En pointillés, celui de g pour $\rho = 1/4$;
en tirets, celui de g pour $\rho = 1/2$.
Fin de l'exercice.

Si l'on regarde bien, f est la densité du cube d'une normale $\mathcal{N}\left(0, 1/\sqrt{2}\right)$. Ainsi, la loi du cube d'une variable aléatoire normale n'est pas déterminée par ses moments (elle ne se reconnaît pas à la seule lecture de ses moments). On peut montrer par le même genre de calcul que les puissances impaires supérieures ou égales à 3 d'une distribution normale ne sont pas déterminées par leurs moments.

D'autres exemples célèbres :
(i) l'exponentielle d'une loi normale, dite *distribution log-normale*. Dans le cas d'une normale $\mathcal{N}\left(0, 1/\sqrt{2}\right)$, sa densité sur $\mathbb{R}_+$ est $\frac{1}{y\sqrt{\pi}} e^{-\log^2 y} = \frac{1}{\sqrt{\pi}} y^{-1-\log y}$ et son $p^{\text{ième}}$ moment vaut $e^{p^2/4}$.
(ii) Le cube d'une loi exponentielle de paramètre 1. Sa densité est $\frac{1}{24}\exp\left(-\sqrt[4]{x}\right)$ sur $\mathbb{R}_+$, son $p^{\text{ième}}$ moment est $\frac{1}{6}\Gamma\left(4p+4\right)$;
- *etc.*
Bref, elles ne sont pas si exotiques, les lois qui ne sont pas déterminées par leurs moments !

On a des conditions suffisantes pour qu'une loi soit déterminée par ses moments. La première est la suivante, qui porte sur le rayon de la série génératrice exponentielle des moments.

Proposition 1 *Soit μ une mesure de probabilité sur $\mathbb{R}$ ayant des moments (entiers) de tous ordres $m_p = \int_{\mathbb{R}} x^p d\mu$, $p \geq 1$. Si la série (entière) de Laplace*

$$\sum_n \frac{m_p}{p!} z^p$$

a un rayon non nul, alors μ est déterminée par ses moments.

Preuve. On s'appuie sur le fait qu'une mesure de probabilité μ est caractérisée par sa transformée de Fourier $x \mapsto \int_{\mathbb{R}} e^{itx} d\mu(t)$. Soit ν une mesure de probabilité dont les moments existent et égalent ceux de μ. Il suffit de montrer que la transformée de Fourier de ν est analytique en 0 pour conclure (les deux transformées de Fourier seront analytiques en 0 et auront le même développement de Taylor). On y va. Soient $z \in \mathbb{C}$ et $r > 0$ tels que $\Im(z) \in [-r, r]$. Alors, pour tout $t \in \mathbb{R}$, $|e^{itz}| = e^{-t\Im z} \leq 2\cosh(tr)$ et

$$\int_{\mathbb{R}} \cosh(tr) d\nu(t) = \sum_{p\geq 0} \frac{r^{2p}}{(2p)!} \int_{\mathbb{R}} t^{2p} d\nu(t) = \sum_{p\geq 0} \frac{m_p}{p!} r^p.$$

Si R est le rayon de la série de Laplace de μ, cette somme positive est finie dès que $r < R$. Cela montre que la série de Laplace de ν a aussi un rayon non nul. ■

Exemples (i) La série de Laplace d'une loi normale centrée et réduite est $\exp\left(z^2/2\right)$ dont le rayon est infini : *les lois normales sont déterminées par leurs moments.*
(ii) La loi Gamma de paramètre $s > 0$ est la densité $x^{s-1}e^{-x}/\Gamma(s)$ sur $\mathbb{R}_+$. Sa série de Laplace est $(1-t)^{-s}$: *les lois Gamma sont déterminées par leurs moments.*

En revanche, les moments des contre-exemples ci-haut (cube de normale et compagnie) sont trop gros (du e^{p^2} et du $\Gamma(4p)$) : les séries de Laplace ont des rayons nuls.

Développer davantage la question de l'unicité en termes modernes amène naturellement à la théorie des opérateurs autoadjoints sur des espaces de Hilbert (étude de leur spectre). Je passe. Une autre condition suffisante de détermination, célébrissime, assez fine : le critère de Carleman.

Théorème 1 (Critère de Carleman)
(i) (Hamburger) Soit μ une mesure de probabilité sur $\mathbb{R}$ admettant des moments (entiers) de tous ordres $m_p = \int_{\mathbb{R}} x^p d\mu(x)$, $p \geq 1$. Alors, il suffit que

$$\sum_{p\geq 0} \left(m_{2p}\right)^{-1/2p} = +\infty$$

pour que μ soit déterminée par ses moments.
(ii) (Stieltjes) Soit μ une mesure de probabilité sur $\mathbb{R}_+$ admettant des moments (entiers) de tous ordres $m_p = \int_{\mathbb{R}} x^p d\mu(x)$, $p \geq 1$. Alors, il suffit que

$$\sum_{p\geq 0} \left(m_p\right)^{-1/2p} = +\infty$$

pour que μ soit déterminée par ses moments.

Preuve. On admet. Se prouve par des considération sur le domaine d'analyticité de la transformé de Stieltjes. Voir par exemple *The problem of moments* de Shohat et Tamarkin (AMS). ■

Là, les moments des contre-exemples à la détermination (cube de normale et compagnie) sont encore trop gros et font converger les séries de Carleman. Un corollaire du critère de Carleman : dès qu'une mesure de probabilité admet des moments exponentiels, elle est déterminée par ses moments. C'est une condition sur la légéreté de la queue (de la distribution, honni soit qui mal y pense).

Corollaire 1 *(i) Une mesure de probabilité admettant une densité φ sur $\mathbb{R}$ est déterminée par ses moments s'il existe $q \geq 1$ et $\delta > 0$ tels que*

$$\int_{\mathbb{R}} \varphi^q(t) e^{\delta|t|} dt < +\infty.$$

(ii) Une mesure de probabilité admettant une densité φ sur $\mathbb{R}_+$ est déterminée par ses moments s'il existe $q \geq 1$ et $\delta > 0$ tels que

$$\int_0^{+\infty} \varphi^q(t) e^{\delta\sqrt{|t|}} dt < +\infty.$$

Preuve. Par Hölder, on peut supposer que $q = 1$. Pour tout p, on écrit le maximum de $e^{-\delta|t|}t^{2p}$ et on applique Carleman. ■

Pour finir, mention d'un critère d'indétermination, dû à M.G. Krein.

Théorème 2 (Condition de Krein) *Soit μ une mesure de probabilité définie par une densité strictement positive f sur $\mathbb{R}_+$.*
(i) Si

$$\int_0^{+\infty} \frac{-\log f(t)}{1+t^2} dt < +\infty,$$

alors μ n'est pas *déterminée par ses moments.*
(ii) Si

$$\int_0^{+\infty} \frac{-\log f(t)}{1+t^2} dt = +\infty,$$

et si f vérifie la condition de Lin, alors μ est *déterminée par ses moments.*

La condition de Lin sur f est la suivante : f est dérivable et, asymptotiquement, $\frac{-xf'}{f}$ tend en croissant vers $+\infty$.

2 Les lois des grandes urnes

La chanson : on prend une urne à deux couleurs, grande. On la prend à temps continu. L'asymptotique du vecteur composition : quand t tend vers l'infini,

$$X(t) = e^{St}\xi\,(1 + o(1))\, v_1 + e^{mt}W^{CT}\,(1 + o(1))\, v_2$$

presque sûrement et dans tous les L^p. En particulier, la variable aléatoire réelle W admet des moments (entiers) de tous ordres. Sa loi dépend des conditions initiales

de l'urne. C'est une densité dont le support est $\mathbb{R}$ tout entier [ChaPouSah] (la loi de ξ est une Gamma).
Equation de dislocation : si on nomme X et Y les lois W^{DT} en partant d'une seule boule, noire ou rouge respectivement, ces variables vérifient le système en loi

$$\left\{\begin{array}{l} X = U^m\left([a+1]X + [b]Y\right) \\ \\ Y = U^m\left([c]X + [d+1]Y\right) \end{array}\right.$$

où, dans chaque équation, U est une variable uniforme sur $[0,1]$ indépendante de tous les exemplaires de X ou Y qui interviennent dans les membres de droite. En termes de moments, cela s'écrit

$$\left\{\begin{array}{l} (1+mp)\, EX^p = E\left([a+1]X + [b]Y\right)^p \\ \\ (1+mp)\, EY^p = E\left([c]X + [d+1]Y\right)^p. \end{array}\right.$$

On développe, on arrange. On obtient, en notant $x_p = EX^p$ et $y_p = EY^p$,

$$\left\{\begin{array}{l} (mp-a)\, x_p - by_p = \displaystyle\sum_{\substack{p_1+\cdots+p_{S+1}=p \\ p_j \leq p-1}} \frac{p!}{p_1!\ldots p_{S+1}!}\, x_{p_1}\ldots x_{p_{a+1}} y_{p_{a+2}}\ldots y_{p_{S+1}} \\ \\ -dx_p + (mp-d)\, y_p = \displaystyle\sum_{\substack{p_1+\cdots+p_{S+1}=p \\ p_j \leq p-1}} \frac{p!}{p_1!\ldots p_{S+1}!}\, x_{p_1}\ldots x_{p_c} y_{p_{c+1}}\ldots y_{p_{S+1}}. \end{array}\right.$$

Comme l'urne est grande, la matrice $mpI_2 - R$ est inversible pour $p \geq 2$ et ce système détermine récursivement les x_p et les y_p dès qu'on fixe x_1 et y_1 (les moyennes de X et Y).
On montre dans [ChaPouSah] que la série génératrice exponentielle de ces moments diverge (rayon nul). Mais alors, la loi de W est-elle déterminée par ses moments ?
On montre par récurrence [ChaMaiPou] le résultat suivant.

Proposition 2 *Il existe $C > 0$ tel que pour tout $p \geq 2$,*

$$\left(\frac{|x_p|}{p!}\right)^{1/p} \leq C \log p.$$

Du coup, on peut appliquer le critère de Carleman, puisque, grâce à la formule de Stirling,

$$x_{2p} \geq \frac{C'}{p \log p}$$

qui est le terme général d'une série de Bertrand divergente (ouf !, tout juste). Du coup, on obtient le résultat [ChaMaiPou] suivant.

Théorème 3 *Les lois des grandes urnes sont déterminées par leurs moments.*

Remarque. On a utilisé le critère de Carleman pour montrer que les W^{CT} sont déterminées par leurs moments. Les variables aléatoires W^{CT} et son homologue W^{DT} en temps discret ront reliées par la *martingale connexion*

$$W^{CT} = \xi^{\sigma} W^{DT}$$

où $\sigma \in]\frac{1}{2}, 1[$ et ξ une loi Gamma indépendante de W^{DT}, dont le paramètre dépend des conditions initiales de l'urne. Du coup, les moments des X sont reliés par l'égalité

$$E\left(X^{CT}\right)^p = \frac{\Gamma\left(\frac{1}{S} + \sigma p\right)}{\Gamma\left(\frac{1}{S}\right)} \times E\left(X^{DT}\right)^p$$

qui montre, grâce à la Proposition 2 et à la formule de Stirling, que les séries de Laplace des W^{DT} ont toutes un rayon infini.