Sur la recherche de ϕ-entropie à maximisante donnée

Jean-François BERCHER[1], Valérie GIRARDIN[2], Justine LEQUESNE[2], Philippe REGNAULT[3], et Steeve ZOZOR[4]

[1]Université Paris-Est, Laboratoire d'Informatique Gaspard Monge (UMR CNRS 8049), ESIEE-Paris, France

[2]Université de Caen Normandie, Laboratoire de Mathématiques Nicolas Oresme (UMR CNRS 6139), Caen, France

[3]Université de Reims Champagne-Ardenne, Laboratoire de Mathématiques de Reims, EA 4535, Reims, France

[4]Univ. Grenoble Alpes, GIPSA-Lab, (UMR CNRS 5216), Grenoble, France

jf.bercher@esiee.fr, {valerie.girardin, justine.lequesne}@unicaen.fr, philippe.regnault@univ-reims.fr, steeve.zozor@gipsa-lab.inpg.fr

Résumé – Nous nous intéressons ici au problème de lois d'entropie maximum, sous contraintes de moments. Contrairement au problème usuel de recherche de maximisante d'une entropie donnée, ou de contraintes pour qu'une loi fixée soit maximisante, nous considérons la recherche de l'entropie elle-même telle qu'une loi donnée en soit sa maximisante. Il s'agit en quelques sorte d'adapter l'entropie à la maximisante. Cette approche trouve potentiellement des applications dans les problèmes de tests d'adéquation basés sur des critères entropiques. Elle permet de sortir du cadre des lois de la famille exponentielle, correspondant aux maximisantes de l'entropie de Shannon, et également de se limiter à des contraintes de moment simples, en pratique estimés à partir de l'échantillon observé. Cette approche nous conduit enfin à définir des fonctionnelles entropiques fonction à la fois de la densité de probabilité et de l'état, permettant de traiter des lois non symétriques ou multimodales.

Abstract – In this paper, we are interested in maximum entropy problems under moment constraints. Contrary to the usual problem of finding the maximizer of a given entropy, or of selecting constraints such that a given distribution is a maximizer, we focus here on the determination of an entropy such that a given distribution is its maximizer. The goal is in some sense to adapt the entropy to its maximizer, with potential application in entropy-based goodness-of-fit tests. It allows us to consider distributions out the exponential family – to which the maximizers of the Shannon entropy belong, and also to consider simple moment constraints, estimated from the observed sample. Finally, this approach also yields entropic functionals that are function of both probability density and state, allowing us to include skew-symmetric or multimodal distributions in the setting.

1 Introduction

Le principe de recherche des lois à maximum d'entropie sous certaines contraintes de moments est largement utilisé dans le domaine de la physique pour la modélisation statistique de système satisfaisant les dites contraintes [1], ou dans divers problèmes de traitement de données comme en communication, clustering, reconnaissance de forme (voir par exemple [1, 2] et références). De même, en statistique, divers problèmes de test d'adéquation à une loi s'appuient sur des critères entropiques, l'entropie retenue étant celle dont la loi testée en est sa maximisante sous contraintes de moments [3]. Le principe en est de comparer l'entropie de la loi testée où les moments théoriques sont remplacés par leur estimés, à une estimation directe de l'entropie à partir des données [4–6].

Parallèlement à ce constat, il est également connu que les lois à maximum d'entropie de Shannon sous contraintes de moments appartiennent à la famille exponentielle [1]. Une loi de cette famille peut être vue à maximum d'entropie sous contraintes particulières ; la loi Gamma par exemple est à maximum d'entropie de Shannon sous contraintes de moyenne et de moment logarithmique.

Dans des problèmes d'adéquation à des lois, on peut vouloir élargir le champ des lois testées hors du cadre de la famille exponentielle, et/ou se focaliser sur des lois contraintes par des moments simples, d'ordre 1 ou 2, d'estimation mieux maîtrisée. On doit alors considérer d'autres entropies que celle de Shannon, et en particulier construire des entropies adaptées aux lois étudiées. Nous en donnons ici quelques idées directrices.

2 ϕ-entropies à maximisante donnée

2.1 ϕ-entropies : définition et maximisantes

Définition 1. *Soit $\phi : \Omega \subset \mathbb{R}_+ \mapsto \mathbb{R}$ une fonction définie, différentiable et strictement convexe sur l'ensemble convexe Ω. On appelle ϕ-entropie d'une probabilité sur $\mathcal{X}$ (convexe de $\mathbb{R}$), de densité f par rapport à la mesure de Lebesgue, le réel*

$$H_\phi[f] := -\int_{\mathcal{X}} \phi(f(x))\,\mathrm{d}x. \qquad (1)$$

Par exemple, l'entropie de Shannon est associée à $\phi(x) = x\log x$ et celle de Havrda-Charvàt-Tsallis à $\phi(x) = (x^q - x)/(q-1)$.

De manière générale, on pourrait considérer des h-ϕ entropies de la forme $h(H_\phi[f])$ avec h croissante (ou ϕ concave et h décroissante) [7]. Comme h n'intervient pas dans le problème de maximisation, nous nous limiterons ici à la définition 1.

Etant données $n+1$ fonctions mesurables $T_0 \equiv 1$, $T_1, \ldots, T_n$ linéairement indépendantes de $\mathcal{X}$ dans $\mathbb{R}$ et $(t_0 = 1, t_1, \ldots, t_n) \in \mathbb{R}^{n+1}$, on note

$$\mathcal{F} := \left\{ f : \mathcal{X} \to \mathbb{R}_+ : \int_{\mathcal{X}} T_i(x) f(x) \, \mathrm{d}x = t_i, i = 0, \ldots, n \right\} \quad (2)$$

l'ensemble des densités sur $\mathcal{X}$ vérifiant les contraintes de moments.

Nous considérons ici le problème de maximisation de la ϕ-entropie (1) sous contraintes de moments

$$f^* = \operatorname*{argmax}_{f \in \mathcal{F}} H_\phi[f] \quad (3)$$

Notons qu'en raison de la stricte convexité de ϕ, la fonctionnelle H_ϕ est concave, et le problème de maximisation (3) admet donc une unique solution.

Proposition 1. *Supposons qu'il existe $f_0 \in \mathcal{F}$ telle que*

$$\phi'\big(f_0(x)\big) = \sum_{i=0}^{n} \lambda_i \, T_i(x), \quad \textit{pour p.t. } x \in \mathcal{X}, \quad (4)$$

où $(\lambda_0, \ldots, \lambda_n) \in \mathbb{R}^{n+1}$. Alors f_0 est l'unique solution du problème de maximisation d'entropie (3).

Démonstration. Soit la divergence de Bregman associée à ϕ, $D_\phi(y_1, y_2) = \phi(y_1) - \phi(y_2) - \phi'(y_2)\,(y_1 - y_2)$ définie sur le convexe $\mathcal{X} \times \mathcal{X}$. En raison de la convexité stricte de ϕ, D_ϕ est positive, et nulle ssi $y_1 = y_2$. Pour tout $f \in \mathcal{F}$, en posant $y_2 = f_0(x)$ et $y_1 = f(x)$ puis en intégrant sur $\mathcal{X}$, on obtient

$$0 \le -H_\phi[f] + H_\phi[f_0] - \sum_i \lambda_i \int_{\mathcal{X}} T_i(x) \big(f(x) - f_0(x)\big) \, \mathrm{d}x, \quad (5)$$

avec égalité si $f = f_0$ presque partout. La conclusion en découle immédiatement en remarquant que les intégrales dans (5) s'annulent car f et f_0 partagent les mêmes moments. □

La réciproque n'est pas nécessairement vraie [8]. Elle est cependant exacte lorsque $\mathcal{X}$ est un compact [9] ou pour des fonctions f localement bornées sur $\mathcal{X}$ quelconque [10].

2.2 ϕ-entropies contraintes par la maximisante

Nous cherchons à caractériser une densité f_0 maximisant une ϕ-entropie sous des contraintes de moments. Une approche classique consiste à fixer *a priori* la ϕ-entropie – typiquement, l'entropie de Shannon – puis à déterminer les contraintes de moments appropriées, s'il en existe. L'approche alternative développée ici consiste à choisir *a priori* des contraintes simples – typiquement, des contraintes sur les moments d'ordres 1 et 2 – puis à déterminer une fonction ϕ, si elle existe, telle que f_0 réalise le maximum de la ϕ-entropie sous les contraintes données. Précisément, la démarche consiste à

1. Fixer *a priori* $T_0 \equiv 1, T_1, \ldots, T_n$;
2. Exhiber une fonction strictement croissante ϕ' telle que f_0 satisfasse l'équation (4) ;
3. La primitiver pour obtenir $\phi(x) = \displaystyle\int \phi'(y) \, \mathrm{d}y$.

Clairement, une telle fonction ϕ' existe si et seulement si les fonctions $\sum_i \lambda_i T_i$ et f_0 définissent le même ordre sur $\mathcal{X}$, *i.e.*,

$$\left(\sum_i \lambda_i T_i(x) \le \sum_i \lambda_i T_i(y) \right) \Leftrightarrow \big(f_0(x) \le f_0(y)\big), \; x, y \in \mathcal{X}. \quad (6)$$

Si (6) est vérifiée, les fonctions $\sum_i \lambda_i T_i$ et f_0 ont un comportement similaire (mêmes variations et symétries). En particulier, $\sum_i \lambda_i T_i$ est constante sur $f_0^{-1}(\{y\})$ et

$$\phi'(y) := \sum_{i=0}^{n} \lambda_i \, T_i\big(f_0^{-1}(y)\big), \quad y \in f_0(\mathcal{X}) \quad (7)$$

est bien définie. Notons que les paramètres λ_i font partie intégrante de la définition de ϕ' (qui n'est donc pas unique) et pourront être choisis de manière à simplifier son expression.

2.3 Fonctionnelle dépendant de l'état

La condition (6) n'est pas toujours satisfaite, en particulier si l'on souhaite se limiter à des moments simples. Par exemple, la densité de la loi Gamma n'a pas le même comportement qu'un polynôme du second degré : elle ne pourra maximiser aucune ϕ-entropie sous contraintes de moments d'ordre 1 ou 2. Cette restriction peut être contournée pour certaines lois en étendant la classe des ϕ-entropies à des fonctionnelles dont l'intégrande dépend non seulement de la densité, mais aussi de l'état.

Définition 2. *Soit $\phi : \mathcal{X} \times \Omega \mapsto \mathbb{R}$ telle que pour tout $x \in \mathcal{X}$, $\phi(x, \cdot)$ soit strictement convexe et différentiable sur l'ensemble convexe $\Omega \subset \mathbb{R}_+$. On appelle ϕ-entropie état-dépendante d'une probabilité sur $\mathcal{X}$ de densité f le réel*

$$H_\phi[f] := -\int_{\mathcal{X}} \phi(x, f(x)) \, \mathrm{d}x \quad (8)$$

Comme $\phi(x, \cdot)$ est convexe, $H_\phi[f]$ est concave. Le problème de maximisation sous contraintes défini en (3), étendu aux ϕ-entropies état-dépendantes admet donc une unique solution ; le résultat obtenu à la proposition 1 se généralise comme suit.

Proposition 2. *Supposons qu'il existe $f_0 \in \mathcal{F}$ telle que*

$$\frac{\partial \phi}{\partial y}\big(x, y = f_0(x)\big) = \sum_{i=0}^{n} \lambda_i \, T_i(x), \quad \textit{pour p.t. } x \in \mathcal{X}, \quad (9)$$

où $(\lambda_0, \ldots, \lambda_n) \in \mathbb{R}^{n+1}$. Alors f_0 est l'unique solution du problème de maximisation d'entropie (3), où H_ϕ est donnée par (8).

Démonstration. La preuve est en tous points similaire à celle de la proposition 1. □

La démarche développée dans la partie 2.2 peut encore être étendue à la recherche d'une entropie état-dépendante telle qu'une densité donnée f_0 la maximise sous des contraintes de moments, grâce à une partition de $\mathcal{X}$ telle que f_0 soit monotone sur chaque partie, en considérant les moments partiels sur chacune de ces parties.

Définition 3. *Soit $\phi : \mathcal{X} \times \Omega \mapsto \mathbb{R}$ et une partition $(\mathcal{X}_1, \ldots, \mathcal{X}_m)$ de convexes de $\mathcal{X}$ telle que la fonction ϕ s'écrit*

$$\phi(x, y) = \sum_{k=1}^{m} \phi_k(y) \mathbb{1}_{\mathcal{X}_k}(x), \tag{10}$$

avec $\mathbb{1}_A$ fonction indicatrice de l'ensemble A et chaque ϕ_k étant convexe. La ϕ-entropie état-dépendante associée à ϕ sera dite $(\mathcal{X}_1, \ldots, \mathcal{X}_m)$-dépendante.

Pour le problème (3) où les contraintes sont restreintes sur chacun des ensembles de la partition, la proposition 2 implique immédiatement le résultat suivant.

Proposition 3. *Supposons qu'il existe $f_0 \in \mathcal{F}$ telle que*

$$\sum_{k=1}^{m} \Big(\phi_k'(f_0(x)) - \lambda_0 - \sum_{i_k=1}^{n_k} \lambda_{k,i_k} T_{k,i_k}(x) \Big) \mathbb{1}_{\mathcal{X}_k}(x) = 0,$$

avec des contraintes de moments données par des fonctions

$$T_{k,i_k}(x) \mathbb{1}_{\mathcal{X}_k}(x), \quad k = 1, \ldots, m \quad i_k = 1, \ldots, n_k.$$

Alors f_0 est solution du problème (3)*, où H_ϕ est l'entropie $(\mathcal{X}_1, \ldots, \mathcal{X}_m)$-dépendante associée à ϕ donnée par* (10)*.*

Par suite, l'identification d'une ϕ-entropie état-dépendante telle qu'une distribution f_0 la maximise sous contraintes de moments peut se résoudre selon la démarche suivante :

1. Fixer une partition $(\mathcal{X}_1, \ldots, \mathcal{X}_m)$ de $\mathcal{X}$ telle que la distribution f_0 soit monotone sur chaque partie $\mathcal{X}_k$.
2. Fixer un ensemble de fonctions $T_{k,i_k}(x)$, $k = 1, \ldots, m$, $i_k = 1, \ldots, n_k$ et des paramètres λ_{k,i_k} tels que $\displaystyle\sum_{i_k=1}^{n_k} \lambda_{k,i_k} T_{k,i_k}$ et f_0 définissent le même ordre sur $\mathcal{X}_k$ – voir (6).
3. Considérer ϕ_k' définie par
$$\phi_k'(y) = \sum_{i_k=0}^{n_k} \lambda_{k,i_k} T_{k,i_k} \left(f_{0,k}^{-1}(y) \right)$$
avec $f_{0,k}^{-1}$ fonction inverse de f_0 sur $\mathcal{X}_k$.
4. Primitiver les résultats pour obtenir $\phi_k(y) = \displaystyle\int \phi_k'(y)\, \mathrm{d}y$.

Ici encore, les λ_{k,i_k} font partie intégrante de la définition des ϕ_k et pourront être choisis judicieusement de façon à en simplifier l'expression.

Dans une optique d'application à des tests d'adéquation à une loi de densité f_0, l'estimation de moment restreints à chaque sous domaine $\mathcal{X}_k$ de $\mathcal{X}$ est envisageable facilement par seuillage des données à tester (avec éventuellement la nécessité d'estimer les seuils définissant la partition).

3 Quelques exemples

3.1 Gaussienne, Student et moment d'ordre 2

Dans le cas d'école de la distribution gaussienne, $f_0(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)$, et cette loi est maximisante d'une ϕ-entropie sous contrainte de moment $T_1(x) = x^2$ sur $\mathcal{X} = \mathbb{R}$ si ϕ vérifie l'équation (7), soit $\phi'(y) = (\lambda_0 - \sigma^2 \log(2\pi\sigma^2)\lambda_1) - 2\sigma^2 \lambda_1 \log y$. Le choix judicieux $\lambda_0 = 1 - \log(2\pi\sigma^2)$ et $\lambda_1 = -\frac{1}{2\sigma^2}$ conduit immédiatement à $\phi(y) = y \log y$: on retrouve l'entropie de Shannon dont la gaussienne est maximisante [1].

Le second cas d'école consiste en la distribution q-gaussienne (distribution Student, ou de Tsallis) [11], soit $f_0(x) = C_q \left(1 - (q-1)\beta x^2\right)_+^{\frac{1}{q-1}}$ avec $x_+ = \max(x, 0)$ et $q > 0$. Sous contrainte de moment d'ordre 2, ϕ doit vérifier $\phi'(y) = \left(\lambda_0 + \frac{\lambda_1}{(q-1)\beta}\right) - \frac{\lambda_1 \, y^{q-1}}{C_q^{q-1}(q-1)\beta}$. Le choix judicieux $\lambda_0 = \frac{q \, C_q^{q-1} - 1}{q-1}$ et $\lambda_1 = -q \, C_q^{q-1} \beta$ conduit à $\phi(y) = \frac{y^q - y}{q-1}$, et par suite à l'entropie de Havrdat-Charvàat ou de Tsallis, dont la q-gaussienne est la maximisante [11].

3.2 Loi arcsinus et moment d'ordre 2

Pour cette loi, cas spécial de distribution beta, $f_0(x) = \frac{1}{\pi\sqrt{\alpha^2 - x^2}}$ sur $\mathcal{X} = (-\alpha\,;\,\alpha)$. Lorsque le moment d'ordre 2 est contraint, cette distribution est maximisante d'une ϕ-entropie si $\phi'(y) = (\lambda_0 + \lambda_1 \alpha^2) - \frac{\lambda_1}{\pi^2 y^2}$. Le choix $\lambda_0 = -\frac{\alpha^2}{\pi^2}$ et $\lambda_1 = \pi^2$ conduit à

$$\phi(y) = \frac{\mathbb{1}_{\left[\frac{1}{\pi\alpha}\,;\,+\infty\right)}(y)}{y}, \qquad H_\phi[f] = -\int_{-\alpha}^{\alpha} \frac{\mathrm{d}x}{f(x)}.$$

3.3 Loi Gamma : un exemple asymétrique

Les lois gamma, soit $f_0(x) = \frac{\beta^\alpha x^{\alpha-1} \exp(-\beta x)}{\Gamma(\alpha)}$ sur $\mathcal{X} = \mathbb{R}_+$, sont non monotones et unimodales pour $\alpha > 1$. Précisément, une telle loi est monotone sur chaque sous-domaine

$$\mathcal{X}_0 = \left[0\,;\, \frac{\alpha - 1}{\beta}\right) \quad \text{et} \quad \mathcal{X}_{-1} = \left[\frac{\alpha-1}{\beta}\,;\, +\infty\right).$$

En raison de l'asymétrie de la loi, elle pourra par exemple être vue comme maximisante sous contraintes de moments d'ordre $p > 0$ partiels (typiquement $p = 1$ ou $p = 2$), $T_{k,1}(x) = x^p \mathbb{1}_{\mathcal{X}_k}(x)$. Sur chaque domaine $\mathcal{X}_k$, f_0 est inversible, sous la forme $f_{0,k}^{-1}(y) = \frac{1-\alpha}{\beta} \; \mathrm{W}_k\left(-\tau \, y^{\frac{1}{\alpha-1}}\right)$ avec

$$y \in \left[0\,;\, \frac{1}{(e\tau)^{\alpha-1}}\right], \qquad \tau = \frac{\Gamma(\alpha)}{\beta\,(\alpha-1)^{\alpha-1}}$$

et W fonction de Lambert multiforme, solution de $z = \mathrm{W}(z)\exp(\mathrm{W}(z))$, où k désigne la branche de cette fonction, principale pour $k = 0$, secondaire pour $k = -1$. Par suite, f_0 est maximisante de la ϕ-entropie sur chaque branche qui satisfait à l'équation $\phi_k'(y) = \lambda_0 + \lambda_{k,1}\left[-\mathrm{W}_k\left(-\tau \, y^{\frac{1}{\alpha-1}}\right)\right]^p$, où le facteur $\frac{1-\alpha}{\beta}$ est absorbé dans les $\lambda_{k,1}$. Nécessairement

$(-1)^k \lambda_{k,1} \geq 0$ pour que ϕ_k' soit croissante. On peut montrer que ϕ_0 et ϕ_{-1} s'écrivent

$$\phi_k(y) = c_k \,+\, \lambda_0\, y \,+\, \lambda_{k,1}\, y \left[-\,\mathrm{W}_k\Big(-\tau y^{\frac{1}{\alpha-1}}\Big)\right]^p \times \left[1 - \frac{p\;{}_1F_1\Big(1\,;\,p+\alpha-1\,;\,(1-\alpha)\,\mathrm{W}_k\Big(-\tau y^{\frac{1}{\alpha-1}}\Big)\Big)}{p+\alpha-1}\right]$$

avec ${}_1F_1$ fonction hypergéométrique confluente (ou de Kummer). On pourra choisir les constantes d'intégration c_k de manière à garantir $\phi_k(0) = 0$, en particulier pour $k = -1$, le domaine $\mathcal{X}_{-1}$ étant non borné.

Les deux composantes ϕ_0 et ϕ_{-1} de ϕ correspondant au cas d'une contrainte d'ordre $p = 2$ sont représentées figure 1 pour $\beta = 3$, $\alpha = 2$ et $\alpha = 5$, avec les choix (pour des raisons d'illustration) $\lambda_0 = 0, \lambda_{0,1} = 1, \lambda_{-1,1} = -0.1$.

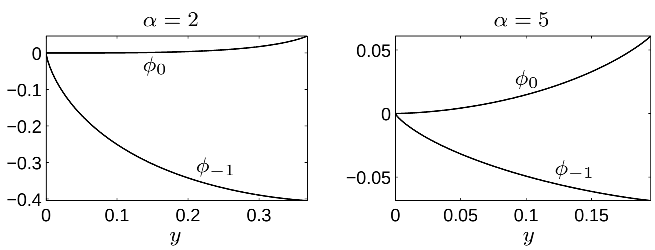

FIGURE 1 – Composantes ϕ_0 et ϕ_{-1} de la fonctionnelle entropique $(\mathcal{X}_0, \mathcal{X}_{-1})$-dépendante telle que la distribution gamma en soit la maximisante sous contraintes de moments partiels $T_{k,1}(x) = x^2 \mathbb{1}_{\mathcal{X}_k}(x)$.

4 Discussion

Dans cet article, nous avons revisité le problème d'entropie maximale sous contraintes. Dans l'optique de s'appuyer sur des critères entropiques pour des tests d'adéquation à une loi, le problème consiste à voir la loi cible comme à entropie maximale sous contraintes. Nous avons proposé ici de chercher non pas les contraintes adéquates permettant de voir une distribution donnée comme maximisante de l'entropie de Shannon, mais au contraire d'adapter la fonctionnelle entropique à la loi cible, en se limitant à des contraintes simples et en se plaçant dans la classe des ϕ-entropies. L'approche proposée permet en outre potentiellement de sortir des lois de la famille exponentielle. De plus, en permettant à la fonctionnelle entropique de dépendre à la fois de l'état et de la loi, les cas de lois asymétriques ou encore multimodales peuvent être considérés.

Dans un cadre applicatif (tests d'adéquation, problèmes de clustering revisité, ...), la problématique de l'estimation de l'entropie se pose alors. Les approches de type plus proches voisins [12], s'appliquent naturellement à toute ϕ-entropie.

Pour être complet, un problème important à traiter serait celui de la recherche d'inégalités de type Cramér-Rao, reliant moment et information de Fisher généralisée à définir de sorte que l'inégalité soit saturée pour une loi donnée. Il s'agirait typiquement de relier une distribution f_0 donnée, un moment simple de celle-ci, la ϕ-entropie pour laquelle f_0 est sa distribution d'entropie maximale, et une ϕ-information de Fisher permettant de minorer une erreur moyenne sur l'estimation de ce moment avec saturation pour la loi f_0 (voir par exemple [13]). Une telle étude, en cours, s'avère néanmoins délicate, en particulier dans le cadre étendu d'entropies dépendantes de l'état.

Références

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Hoboken, New Jersey, 2nd edition, 2006.

[2] A. O. Hero III, B. Ma, O. J. J. Michel, and J. Gorman. Application of entropic spanning graphs. *IEEE Sig. Proc. Mag.*, 19(5) :85–95, Sept. 2002.

[3] J. Lequesne. *Tests statistiques basés sur la théorie de l'information. Applications en biologie et en démographie*. PhD thesis, Université de Caen Normandie, mai 2015.

[4] O. Vasicek. A test for normality based on sample entropy. *J. of the Royal Stat. Soc. B*, 38(1) :54–59, 1976.

[5] D.V. Gokhale. On entropy-based goodness-of-fit tests. *Comp. Stat. & and Data Anal.*, 1 :157–165, march 1983.

[6] J. Lequesne. A goodness-of-fit test of Student distributions based on Rényi entropy. In *AIP conference proceedings*, volume 1641, pages 487–494, 2014.

[7] I. Csiszàr. Information-type measures of difference of probability distributions and indirect observations. *Studia Scient. Math. Hungarica*, 2 :299–318, 1967.

[8] J. M. Borwein and A. S. Lewis. Partially-finite programming in L_1 and the existence of maximum entropy estimates. *SIAM Journal on Optimization*, 3(2) :248–267, May 1993.

[9] V. Girardin. Méthodes de réalisation de produit scalaire et de problème de moments avec maximisation d'entropie. *Studia Mathematica*, 124(3) :199–213, 1997.

[10] V. Girardin. Relative entropy and spectral constraints : some invariance properties of the arma class. *Journal of Time Series Analysis*, 28 :844–866, 2007.

[11] J. A. Costa, A. O. Hero III, and C. Vignat. On solutions to multivariate maximum α-entropy problems. *Lecture Notes in Computer Sciences*, 2683 :211–226, 2003.

[12] N. Leonenko, L. Pronzato, and V. Savani. A class of Rényi information estimators for multidimensional densities. *Annals of Stat.*, 36(5) :2153–2182, Oct. 2008.

[13] J.-F. Bercher. On generalized Cramér-Rao inequalities, generalized Fisher information and characterizations of generalized q-Gaussian distributions. *J. of Phys. A*, 45(25) :255303, June 2012.