journal of inequalities in pure and applied mathematics

http://jipam.vu.edu.au

issn: 1443-5756

Volume 10 (2009), Issue 2, Article 34, 7 pp.

A CHARACTERIZATION OF THE UNIFORM DISTRIBUTION ON THE CIRCLE BY STAM INEQUALITY

PAOLO GIBILISCO, DANIELE IMPARATO, AND TOMMASO ISOLA

DIPARTIMENTO SEFEMEQ, FACOLTÀ DI ECONOMIA
UNIVERSITÀ DI ROMA "TOR VERGATA"
VIA COLUMBIA 2, 00133 ROME, ITALY.
gibilisco@volterra.uniroma2.it
URL: http://www.economia.uniroma2.it/sefemeq/professori/gibilisco

DIPARTIMENTO DI MATEMATICA
POLITECNICO DI TORINO
CORSO DUCA DEGLI ABRUZZI 24, 10129 TURIN, ITALY.
daniele.imparato@polito.it

DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI ROMA "TOR VERGATA"
VIA DELLA RICERCA SCIENTIFICA, 00133 ROME, ITALY.
isola@mat.uniroma2.it
URL: http://www.mat.uniroma2.it/~isola

Received 08 November, 2008; accepted 20 March, 2009
Communicated by I. Pinelis

ABSTRACT. We prove a version of Stam inequality for random variables taking values on the circle $\mathbb{S}^1$. Furthermore we prove that equality occurs only for the uniform distribution.

Key words and phrases: Fisher information, Stam inequality.

2000 *Mathematics Subject Classification.* 62F11, 62B10.

1. INTRODUCTION

It is well-known that the Gaussian, Poisson, Wigner and (discrete) uniform distributions are maximum entropy distributions in the appropriate context (for example see [18, 6, 7]). On the other hand all the above quoted distributions can be characterized as those distributions giving equality in the Stam inequality. Let us describe what Stam inequality is about.

The Fisher information I_X of a real random variable (with strictly positive differentiable density function f) is defined as

$$I_X := \int (f'(x)/f(x))^2 f(x)dx. \tag{1.1}$$

303-08

For X, Y independent random variables such that I_X, $I_Y < \infty$, Stam was able to prove the inequality

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{1.2}$$

where equality holds iff X, Y are Gaussian (see [16, 1]).

It is difficult to overestimate the importance of the above result because of its links with other important results in analysis, probability, statistics, information theory, statistical mechanics and so on (see [2, 3, 9, 17]). Different proofs and deep generalizations of the theorem appear in the recent literature on the subject (see [19, 13]).

A free analogue of Fisher information has been introduced in free probability. Also in this case one can prove a Stam-like inequality. It is not surprising that the equality case characterizes the Wigner distribution that, in many respects, is the free analogue of the Gaussian distribution (see [18]).

In the discrete setting, one can introduce appropriate versions of Fisher information and prove the Stam inequality. On the integers $\mathbb{Z}$, equality characterizes the Poisson distribution, while on a finite group G equality occurs for the uniform distribution (see [8, 15, 10, 11, 12, 14, 4, 5]).

In this short note we show that also on the circle $\mathbb{S}^1$ one can prove a version of the Stam inequality. This result is obtained by suitable modifications of the standard proofs. Moreover, equality occurs for the maximum entropy distribution, namely for the uniform distribution on the circle.

2. FISHER INFORMATION AND STAM INEQUALITY ON $\mathbb{R}$

Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable, strictly positive density. One may define the f-score $J_f : \mathbb{R} \to \mathbb{R}$ by

$$J_f := \frac{f'}{f}.$$

Note that J_f is f-centered in the sense that $\mathbb{E}_f(J_f) = 0$. In general, if $X : (\Omega, \mathcal{F}, p) \to \mathbb{R}$ is a random variable with density f, we write $J_X = J_f$ and

$$I_X = \text{Var}_f(J_f) = \mathbb{E}_f[J_f^2];$$

namely

$$I_X := \int_{\mathbb{R}} \left(f'(x)/f(x)\right)^2 f(x)dx. \tag{2.1}$$

Let us suppose that I_X, $I_Y < \infty$.

Theorem 2.1 ([16])**.** *If $X, Y : (\Omega, \mathcal{F}, p) \to \mathbb{R}$ are independent random variables then*

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{2.2}$$

with equality if and only if X, Y are Gaussian.

3. STAM INEQUALITY ON $\mathbb{S}^1$

We denote by $\mathbb{S}^1$ the circle group, namely the multiplicative subgroup of $\mathbb{C} \setminus \{0\}$ defined as

$$\mathbb{S}^1 := \{z \in \mathbb{C} : |z| = 1\}.$$

We say that a function $f : \mathbb{S}^1 \to \mathbb{R}$ has a *tangential derivative* in $z \in \mathbb{S}^1$ if the following limit exists and is finite

$$D_T f(z) := \lim_{h \to 0} \frac{1}{h} \left[f(ze^{ih}) - f(z) \right].$$

From now on we consider functions $f : \mathbb{S}^1 \to \mathbb{R}$ that are twice differentiable strictly positive densities.
Then, the f-score is defined as

$$J_f := \frac{D_T f}{f},$$

and is f-centered, in the sense that $\mathbb{E}_f(J_f) = 0$, where $\mathbb{E}_f(g) := \int_{\mathbb{S}^1} gf\, d\mu$, and μ is the normalized Haar measure on $\mathbb{S}^1$.

If $X : (\Omega, \mathcal{F}, p) \to \mathbb{S}^1$ is a random variable with density f, we write $J_X = J_f$ and define the Fisher information as

$$I_X := \mathrm{Var}_f(J_f) = \mathbb{E}_f[J_f^2].$$

The main result of this paper is the proof of the following version of Stam inequality on the circle.

Theorem 3.1. *If $X, Y : (\Omega, \mathcal{F}, p) \to \mathbb{S}^1$ are independent random variables then*

$$\frac{1}{I_{XY}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{3.1}$$

with equality if and only if X or Y are uniform.

4. PROOF OF THE MAIN RESULT

To prove our result we identify $\mathbb{S}^1$ with the interval $[0, 2\pi]$, where 0 and 2π are identified and the sum is modulo 2π. Any function $f : [0, 2\pi] \to \mathbb{R}$, such that $f(0) = f(2\pi)$, can be thought of as a function on $\mathbb{S}^1$. In this representation, the tangential derivative must be substituted by an ordinary derivative.

In this context, a density will be a nonnegative function $f : [0, 2\pi] \to \mathbb{R}$ such that

$$\frac{1}{2\pi} \int_0^{2\pi} f(\theta) \mathrm{d}\theta = 1.$$

The uniform density is the function

$$f(\theta) = 1, \qquad \forall \theta \in [0, 2\pi].$$

From now on, we shall consider f belonging to the class

$$\mathcal{P} := \left\{ f : [0, 2\pi] \to \mathbb{R} \,\middle|\, \int_0^{2\pi} f(\theta) \mathrm{d}\theta = 2\pi,\ f > 0 \ \text{a.e.}, \right.$$
$$\left. f \in \mathcal{C}^2(\mathbb{S}^1),\ f^{(k)}(0) = f^{(k)}(2\pi),\ k = 0, 1, 2 \right\}.$$

Let $f \in \mathcal{P}$; then

$$\int_0^{2\pi} f'(\theta) \mathrm{d}\theta = 0$$

and therefore

$$J_f := \frac{f'}{f}$$

is f-centered. Note that $J_f(0) = J_f(2\pi)$.

If $X : (\Omega, \mathcal{F}, p) \to [0, 2\pi]$ is a random variable with density $f \in \mathcal{P}$, from the score $J_X := J_f$ it is possible to define the Fisher information

$$I_X := \mathrm{Var}_f(J_f) = \mathbb{E}_f[J_f^2].$$

In this additive (modulo 2π) context the main result we want to prove takes the following (more traditional) form.

Theorem 4.1. *If $X, Y : (\Omega, \mathcal{F}, p) \to [0, 2\pi]$ are independent random variables then*

$$\frac{1}{I_{X+Y}} \geq \frac{1}{I_X} + \frac{1}{I_Y}, \tag{4.1}$$

with equality if and only if X or Y are uniform

Note that, since $[0, 2\pi]$ is compact, the condition $I_X < \infty$ always holds. However, we cannot ensure in general that $I_X \neq 0$. In fact, it is easy to characterize this degenerate case.

Proposition 4.2. *The following conditions are equivalent*

(i) *X has uniform distribution;*
(ii) *$I_X = 0$;*
(iii) *$J_X = constant$.*

Proof. $(i) \Longrightarrow (ii)$ Obvious.
$(ii) \Longrightarrow (iii)$ Obvious.
$(iii) \Longrightarrow (i)$ Let $J_X(x) = \beta$ for every x. Then f_X is the solution of the differential equation

$$\frac{f'_X(x)}{f_X(x)} = \beta, \qquad f(0) = f(2\pi).$$

Thus $f_X(x) = ce^{\beta x}$ and the symmetry condition implies $\beta = 0$, so that f_X is the uniform distribution. □

Proposition 4.3. *Let $X, Y : (\Omega, \mathcal{F}, p) \to [0, 2\pi]$ be independent random variables such that their densities belong to $\mathcal{P}$. If X (or Y) has a uniform distribution then*

$$\frac{1}{I_{X+Y}} = \frac{1}{I_X} + \frac{1}{I_Y},$$

in the sense that both sides of equality are equal to infinity.

Proof. Because of independence one has, by the convolution formula, that if X is uniform then so is $X + Y$ and therefore we are done by Proposition 4.2. □

As a result of the above proposition, in what follows we consider random variables with strictly positive Fisher information. Before the proof of the main result, we need the following lemma.

Lemma 4.4. *Let $X, Y : (\Omega, \mathcal{F}, p) \to [0, 2\pi]$ be two independent random variables with densities $f_X, f_Y \in \mathcal{P}$ and let $Z := X + Y$. Then*

$$J_Z(Z) = \mathbb{E}_p[J_X(X)|Z] = \mathbb{E}_p[J_Y(Y)|Z]. \tag{4.2}$$

Proof. Let f_Z be the density of Z; namely,

$$f_Z(z) = \frac{1}{2\pi} \int_0^{2\pi} f_X(z - y) f_Y(y) dy, \qquad z \in [0, 2\pi],$$

with $f_Z \in \mathcal{P}$. Then,

$$\begin{aligned} f'_Z(z) &= \frac{1}{2\pi} \frac{d}{dz} \int_0^{2\pi} f_X(z - y) f_Y(y) dy \\ &= \frac{1}{2\pi} \int_0^{2\pi} f_Y(y) f'_X(z - y) dy \\ &= f'_X * f_Y(z). \end{aligned}$$

Therefore, given $z \in [0, 2\pi]$,

$$\begin{aligned} J_Z(z) &= \frac{f'_Z(z)}{f_Z(z)} \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{f_X(x) f_Y(z-x)}{f_Z(z)} \frac{f'_X(x)}{f_X(x)} dx \\ &= \frac{1}{2\pi} \int_0^{2\pi} J_X(x) f_{X|Z}(x|z) dx \\ &= \mathbb{E}_{f_X}[J_X|Z] \\ &= \mathbb{E}_p[J_X(X)|Z]. \end{aligned}$$

Similarly, by symmetry of the convolution formula one can obtain

$$J_Z(z) = \mathbb{E}_p[J_Y(Y)|Z], \qquad z \in [0, 2\pi],$$

proving Lemma 4.4. □

We are ready to prove the main result.

Theorem 4.5. *Let $X, Y : (\Omega, \mathcal{F}, p) \to [0, 2\pi]$ be two independent random variables such that $I_X, I_Y > 0$. Then*

$$\frac{1}{I_{X+Y}} > \frac{1}{I_X} + \frac{1}{I_Y}. \tag{4.3}$$

Proof. Let $a, b \in \mathbb{R}$ and let $Z := X + Y$; then, by Lemma 4.4

$$\begin{aligned} \mathbb{E}_p[aJ_X(X) + bJ_Y(Y)|Z] &= a\mathbb{E}_p[J_X(X)|Z] + b\mathbb{E}_p[J_Y(Y)|Z] \\ &= (a+b) J_Z(Z). \end{aligned} \tag{4.4}$$

Hence, applying Jensen's inequality, we obtain

$$\begin{aligned} \mathbb{E}_p[(aJ_X(X) + bJ_Y(Y))^2] &= \mathbb{E}_p[\mathbb{E}_p[(aJ_X(X) + bJ_Y(Y))^2|Z]] \\ &\geq \mathbb{E}_p[\mathbb{E}_p[aJ_X(X) + bJ_Y(Y)|Z]^2] \\ &= \mathbb{E}_p[(a+b)^2 J_Z(Z)^2] \\ &= (a+b)^2 I_Z, \end{aligned} \tag{4.5}$$

and thus

$$\begin{aligned} (a+b)^2 I_Z &\leq \mathbb{E}_p[(aJ_X(X) + bJ_Y(Y))^2] \\ &= a^2 \mathbb{E}_p[J_X(X)^2] + 2ab\mathbb{E}_p[J_X(X)J_Y(Y)] + b^2 \mathbb{E}_p[J_Y(Y)^2] \\ &= a^2 I_X + b^2 I_Y + 2ab\mathbb{E}_p[J_X(X)J_Y(Y)] \\ &= a^2 I_X + b^2 I_Y, \end{aligned}$$

where the last equality follows from independence and since the score is a centered random variable.

Now, take $a := 1/I_X$ and $b := 1/I_Y$; then we obtain

$$\left(\frac{1}{I_X} + \frac{1}{I_Y}\right)^2 I_Z \leq \frac{1}{I_X} + \frac{1}{I_Y}. \tag{4.6}$$

It remains to be proved that equality cannot hold in (4.6). Define $c := a + b$, where, again, $a = 1/I_X$ and $b = 1/I_Y$; then equality holds in (4.6) if and only if

$$c^2 I_Z = a^2 I_X + b^2 I_Y. \tag{4.7}$$

Let us prove that (4.7) is equivalent to

$$aJ_X(X) + bJ_Y(Y) = cJ_Z(X+Y) \qquad \text{a.e.} \tag{4.8}$$

Indeed, let $H := aJ_X(X) + bJ_Y(Y)$; then equality occurs in (4.5) if and only if

$$\mathbb{E}_p[H^2|Z] = (\mathbb{E}_p[H|Z])^2, \qquad \text{a.e.}$$

i.e.

$$\mathbb{E}_p[(H - \mathbb{E}_p[H|Z])^2|Z] = 0, \qquad \text{a.e.}$$

Therefore, $H = \mathbb{E}_p[H|Z]$ a.e., so that, by (4.4),

$$cJ_Z(Z) = \mathbb{E}_p[aJ_X(X) + bJ_Y(Y)|Z] = aJ_X(X) + bJ_Y(Y) \qquad \text{a.e.},$$

i.e. (4.8) is true. Conversely, if (4.8) holds, then by applying the squared power and taking the expectations we obtain (4.7).

Let $x, y \in [0, 2\pi]$; because of independence

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \neq 0.$$

Thus, it makes sense to write equality (4.8) for $x, y \in [0, 2\pi]$

$$aJ_X(x) + bJ_Y(y) = cJ_Z(x+y). \tag{4.9}$$

By deriving (4.9) with respect to both x and y and subtracting such relations one obtains

$$aJ'_X(x) = bJ'_Y(y), \qquad \forall x, y \in [0, 2\pi],$$

which implies $J'_X(x) = \alpha = constant$, i.e.

$$J_X(x) = \beta + \alpha x, \qquad x \in [0, 2\pi].$$

In particular, by symmetry conditions one obtains

$$\beta = J_X(0) = J_X(2\pi) = \beta + 2\pi\alpha.$$

This implies that $\alpha = 0$, that is, $J_X = constant$. By Proposition 4.2 one has $I_X = 0$. This fact contradicts the hypotheses and ends the proof. □

REFERENCES

[1] N.M. BLACHMAN, The convolution inequality for entropy powers, *IEEE Trans. Inform. Theory*, **11** (1965), 267–271.

[2] E. CARLEN, Superadditivity of Fisher's information and logarithmic Sobolev inequalities, *J. Funct. Anal.*, **101**(1) (1991), 194–211.

[3] A. DEMBO, T. COVER AND J. THOMAS, Information theoretic inequalities, *IEEE Trans. Inform. Theory*, **37**(6) (1991), 1501–1518.

[4] P. GIBILISCO, D. IMPARATO AND T.ISOLA, Stam inequality on $\mathbb{Z}_n$, *Statis. Probab. Lett.*, **78**(13) (2008), 1851–1856.

[5] P. GIBILISCO AND T. ISOLA Fisher information and Stam inequality on a finite group, *Bull. Lond. Math. Soc.*, **40**(5) (2008,) 855–862.

[6] P. HARREMOES, Binomial and Poisson distribution as maximum entropy distributions, *IEEE Trans. Inform. Theory*, **47**(5) (2001), 2039–2041.

[7] O.T. JOHNSON, Log-concavity and the maximum entropy property of the Poisson distribution, *Stoch. Proc. Appl.*, **117**(6) (2007), 791–802.

[8] I.M. JOHNSTONE AND B. MACGIBBON, Une mesure d'information caractérisant la loi de Poisson, in *Séminaire de Probabilités, XXI*, vol. 1247 of Lecture Notes in Math., 563–573, Springer, Berlin, 1987.

[9] A. KAGAN AND Z. LANDSMAN Statistical meaning of Carlen's superadditivity of the Fisher information, *Statis. Probab. Lett.*, **32** (1997), 175–179.

[10] A. KAGAN, A discrete version of Stam inequality and a characterization of the Poisson distribution, *J. Statist. Plann. Inference*, **92**(1-2) (2001), 7–12.

[11] A. KAGAN, Letter to the editor: "A discrete version of Stam inequality and a characterization of the Poisson distribution" [*J. Statist. Plann. Inference*, **92**(1-2), (2001), 7–12], *J. Statist. Plann. Inference*, **99**(1) (2001), 1.

[12] I. KONTOYANNIS, P. HARREMOËS AND O. JOHNSON, Entropy and the law of small numbers, *IEEE Trans. Inform. Theory*, **51**(2) (2005), 466–472.

[13] M. MADIMAN AND R.A. BARRON, Generalized entropy power inequalities and monotonicity properties of information, *IEEE Trans. Inform. Theory*, **53**(7) (2007), 2317–2329.

[14] M. MADIMAN, O. JOHNSON AND I. KONTOYANNIS, Fisher information, compound Poisson approximation and the Poisson channel, *Proc. IEEE Intl. Symp. Inform. Theory, Nice, France*, 2007.

[15] V. PAPATHANASIOU, Some characteristic properties of the Fisher information matrix via Cacoullo-type inequalities, *J. Multivariate Anal.*, **44**(2) (1993), 256–265.

[16] A.J. STAM, Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, **2** (1959), 101–112.

[17] C. VILLANI, Cercignani's conjecture is sometimes true and always almost true. *Comm. Math. Phys.*, **234**(3)(2003), 455–490.

[18] D. VOICULESCU, The analogues of entropy and of Fisher's information measure in free probability theory. V. Noncommutative Hilbert transforms, *Invent. Math.*, **132**(1) (1998), 189–227.

[19] R. ZAMIR, A proof of the Fisher information inequality via a data processing argument, *IEEE Trans. Inform. Theory*, **44**(3) (1998), 1246–1250.