Rényi entropy and variance comparison for symmetric log-concave random variables

(joint work with Piotr Nayar)

Maciej Białobrzeski

Definition of the Rényi entropy

For a random variable X with density f its Rényi entropy of order $\alpha \in (0, \infty) \setminus \{1\}$ is defined as

$$h_\alpha(X) = h_\alpha(f) = \frac{1}{1-\alpha} \log \left(\int f^\alpha(x) \mathrm{d}x \right),$$

assuming that the integral converges.

Definition of the Rényi entropy

For a random variable X with density f its Rényi entropy of order $\alpha \in (0, \infty) \setminus \{1\}$ is defined as

$$h_\alpha(X) = h_\alpha(f) = \frac{1}{1-\alpha} \log \left(\int f^\alpha(x)\mathrm{d}x \right),$$

assuming that the integral converges.

When $\alpha \to 1$, we recover in the limit the usual Shannon differential entropy $h(f) = h_1(f) = -\int f \log f$. Another limiting cases are

$$h_0(f) = \log |\operatorname{supp} f|,$$

$$h_\infty(f) = -\log ||f||_\infty.$$

Bounding Rényi entropy by variance

It is a well known fact that for any random variable with density f one has

$$h(f) \leq \frac{1}{2} \log \mathrm{Var}(f) + \frac{1}{2} \log(2\pi e)$$

with equality only for Gaussian random variables.

Proof: by inequality $\log x \geq x - 1$ we have $\int f \log(f/g) \geq 0$ for any densities f, g. If g is a gaussian density of the same variance as f, then $\int f \log g = -h(g)$.

Bounding Rényi entropy by variance

It is a well known fact that for any random variable with density f one has

$$h(f) \leq \frac{1}{2} \log \mathrm{Var}(f) + \frac{1}{2} \log(2\pi e)$$

with equality only for Gaussian random variables.

Proof: by inequality $\log x \geq x - 1$ we have $\int f \log(f/g) \geq 0$ for any densities f, g. If g is a gaussian density of the same variance as f, then $\int f \log g = -h(g)$.

For $\alpha \in \left(\frac{1}{3}, \infty\right) \setminus \{1\}$ the maximizer of Rényi entropy under fixed variance is of the form

$$f(x) = c_0(1 + (1-\alpha)(c_1 x)^2)_+^{\frac{1}{\alpha - 1}} \quad \text{(Lutwak, Yang, Zhang, 2005).}$$

For $\alpha \leq \frac{1}{3}$ the supremum of h_α under fixed variance is infinite.

Reverse bounds

For density $f_n(x) = \frac{n}{2}1_{[1,1+n^{-1}]}(|x|)$ variance stays bounded, whereas $h_\alpha(f_n) \to -\infty$ for any $\alpha \in [0, \infty]$.
We will consider the problem of minimizing Rényi entropy under fixed variance in the class of log-concave densities.

A random variable X is log-concave, if it has density of the form e^{-V}, where $V : \mathbb{R} \to (-\infty, \infty]$ is convex.

Main theorem

Theorem 1

Let X be a symmetric log-concave random variable in $\mathbb{R}$ and $\alpha > 0$, $\alpha \neq 1$. Define α^* to be the unique solution to the equation $\frac{1}{\alpha-1}\log\alpha = \frac{1}{2}\log 6$ ($\alpha^* \approx 1.241$). Then

$$h_\alpha(X) \geq \frac{1}{2}\log \mathrm{Var}(X) + \frac{1}{2}\log 12 \qquad \text{for } \alpha \leq \alpha^*$$

and

$$h_\alpha(X) \geq \frac{1}{2}\log \mathrm{Var}(X) + \frac{1}{2}\log 2 + \frac{\log\alpha}{\alpha-1} \qquad \text{for } \alpha \geq \alpha^*.$$

The only cases of equality are uniform random variable on a symmetric interval for $\alpha \leq \alpha^*$ and two-sided exponential distribution for $\alpha \geq \alpha^*$.

Preliminary simplifications

The following well known lemma reduces Theorem 1 to the case $\alpha = \alpha^*$.

Lemma 2 (Fradelizi, Madiman, Wang, 2015)

Let f be a log-concave probability density in $\mathbb{R}^n$. Then for any $p \geq q > 0$ we have

$$0 \leq h_q(f) - h_p(f) \leq n\frac{\log q}{q-1} - n\frac{\log p}{p-1}.$$

Preliminary simplifications

By an approximation argument it is enough to show that for every $\sigma, L > 0$ we have

$$\inf\{h_{\alpha^*}(f) \;:\; f \in A\} \geq \log \sigma + \frac{1}{2}\log 2 + \frac{\log \alpha^*}{\alpha^* - 1},$$

where $A = \{f \;:\; \mathrm{Var}(f) = \sigma^2, \mathrm{supp}(f) \subset [-L, L]\}$.
Furthermore, the infimum is attained on A.

Degrees of freedom

The degree of freedom (Fradelizi, Guédon, 2000) of a log-concave function $g : \mathbb{R} \to [0, \infty)$ is the largest integer k such that there exist $\delta > 0$ and linearly independent continuous functions $h_1, \ldots, h_k$ defined on $\{x \in \mathbb{R}, g(x) > 0\}$ such that for every $(\varepsilon_1, \ldots, \varepsilon_k) \in [-\delta, \delta]^k$ the function $g + \sum_{i=1}^{k} \varepsilon_i h_i$ is log-concave.

Degrees of freedom

The degree of freedom (Fradelizi, Guédon, 2000) of a log-concave function $g : \mathbb{R} \to [0, \infty)$ is the largest integer k such that there exist $\delta > 0$ and linearly independent continuous functions $h_1, \ldots, h_k$ defined on $\{x \in \mathbb{R}, g(x) > 0\}$ such that for every $(\varepsilon_1, \ldots, \varepsilon_k) \in [-\delta, \delta]^k$ the function $g + \sum_{i=1}^k \varepsilon_i h_i$ is log-concave.

Fact: if $V : (a, b) \to \mathbb{R}$ is a convex function, then the degree of freedom of e^{-V} is $k+1$ if and only if there exist k (and not less than k) affine functions $\phi_1, \ldots, \phi_k$ such that

$$V = \max_{1 \le i \le k} \phi_i.$$

Minimizer has low degree of freedom

Suppose that $f \in A$ has more than two degrees of freedom, so $f + \varepsilon_1 h_1 + \varepsilon_2 h_2 + \varepsilon_3 h_3$ is log-concave for some linearly independent h_1, h_2, h_3 and every $(\varepsilon_1, \varepsilon_2, \varepsilon_3) \in [-\delta, \delta]^3$. Then the system of equations

$$\varepsilon_1 \int h_1 + \varepsilon_2 \int h_2 + \varepsilon_3 \int h_3 = 0$$

$$\varepsilon_1 \int x^2 h_1 + \varepsilon_2 \int x^2 h_2 + \varepsilon_3 \int x^2 h_3 = 0$$

has space of solutions of dimension at least 1, hence there are $(\eta_1, \eta_2, \eta_3) \in [-\delta, \delta]^3$ such that $f_+ = f + \eta_1 h_1 + \eta_2 h_2 + \eta_3 h_3$ and $f_- = f - \eta_1 h_1 - \eta_2 h_2 - \eta_3 h_3$ are both members of A and $f_+ \neq f_-$.

Minimizer has low degree of freedom

Suppose that $f \in A$ has more than two degrees of freedom, so $f + \varepsilon_1 h_1 + \varepsilon_2 h_2 + \varepsilon_3 h_3$ is log-concave for some linearly independent h_1, h_2, h_3 and every $(\varepsilon_1, \varepsilon_2, \varepsilon_3) \in [-\delta, \delta]^3$. Then the system of equations

$$\varepsilon_1 \int h_1 + \varepsilon_2 \int h_2 + \varepsilon_3 \int h_3 = 0$$

$$\varepsilon_1 \int x^2 h_1 + \varepsilon_2 \int x^2 h_2 + \varepsilon_3 \int x^2 h_3 = 0$$

has space of solutions of dimension at least 1, hence there are $(\eta_1, \eta_2, \eta_3) \in [-\delta, \delta]^3$ such that $f_+ = f + \eta_1 h_1 + \eta_2 h_2 + \eta_3 h_3$ and $f_- = f - \eta_1 h_1 - \eta_2 h_2 - \eta_3 h_3$ are both members of A and $f_+ \neq f_-$. Then, by strict concavity of $x \to x^{\alpha^*}$,

$$\int f^{\alpha^*} = \int \left(\frac{1}{2} f_+ + \frac{1}{2} f_- \right)^{\alpha^*} < \frac{1}{2} \int f_+^{\alpha^*} + \frac{1}{2} \int f_-^{\alpha^*},$$

so f cannot be the minimizer.

Two-point form of the inequality

It follows that the minimizer of Rényi entropy in A must have at most two degrees of freedom, so it is of the form

$$f(x) = c\mathbf{1}_{[0,a]}(|x|) + ce^{-\gamma(|x|-a)}\mathbf{1}_{[a,a+b]}(|x|),$$

where $a + b \leq L$, $c > 0$, $a, b, \gamma \geq 0$ and also $\int f = 1$. We can restrict ourselves to such densities.

Two-point form of the inequality

It follows that the minimizer of Rényi entropy in A must have at most two degrees of freedom, so it is of the form

$$f(x) = c\mathbf{1}_{[0,a]}(|x|) + ce^{-\gamma(|x|-a)}\mathbf{1}_{[a,a+b]}(|x|),$$

where $a + b \leq L$, $c > 0$, $a, b, \gamma \geq 0$ and also $\int f = 1$. We can restrict ourselves to such densities.

The inequality we are left to prove is

$$G(a,b) = (2a\alpha^* + 1 - e^{-\alpha^* b})^{\frac{2}{1-\alpha^*}}(a + 1 - e^{-b})^{\frac{1-3\alpha^*}{1-\alpha^*}} - \left(\frac{a^3}{3} + \int_0^b (x+a)^2 e^{-x} \mathrm{d}x\right) \geq 0$$

Fixed sign of fourth derivative

$$G(a,b) = (2a\alpha^* + 1 - e^{-\alpha^* b})^{\frac{2}{1-\alpha^*}}(a + 1 - e^{-b})^{\frac{1-3\alpha^*}{1-\alpha^*}} - \left(\frac{a^3}{3} + \int_0^b (x+a)^2 e^{-x}\mathrm{d}x\right) \geq 0$$

Note that the second term is polynomial of degree 3 in a. It turns out that $\frac{\partial^4}{\partial a^4} G(a,b) \geq 0$ for every $a, b \geq 0$. It follows from the following observation.

Fixed sign of fourth derivative

$$G(a,b) = (2a\alpha^* + 1 - e^{-\alpha^* b})^{\frac{2}{1-\alpha^*}}(a + 1 - e^{-b})^{\frac{1-3\alpha^*}{1-\alpha^*}}$$
$$-\left(\frac{a^3}{3} + \int_0^b (x+a)^2 e^{-x}\mathrm{d}x\right) \geq 0$$

Note that the second term is polynomial of degree 3 in a. It turns out that $\frac{\partial^4}{\partial a^4} G(a,b) \geq 0$ for every $a, b \geq 0$. It follows from the following observation.

Lemma 3

Let $f(x) = (x+a)^\gamma (x+b)^\delta$, where $a, b \geq 0$ and $\gamma + \delta = m \in \mathbb{N}$. Then $(m+1)$-th derivative of f has fixed sign on $[0, \infty)$. The sign of f is equal to $\operatorname{sgn}(\gamma(\gamma-1)\cdots(\gamma-m))\operatorname{sgn}((b-a)^{m+1})$.

Finishing the proof

Now it suffices to prove the following:
(a) $\lim_{a\to\infty} \frac{\partial^3}{\partial a^3} G(a,b) = 0$ for every $b \geq 0$,
(b) $\lim_{a\to\infty} \frac{\partial^2}{\partial a^2} G(a,b) \geq 0$ for every $b \geq 0$,
(c) $\frac{\partial}{\partial a} G(a,b) \mid_{a=0} \geq 0$ for every $b \geq 0$,
(d) $G(0,b) \geq 0$ for every $b \geq 0$.

Finishing the proof

Now it suffices to prove the following:
(a) $\lim_{a\to\infty} \frac{\partial^3}{\partial a^3} G(a,b) = 0$ for every $b \geq 0$,
(b) $\lim_{a\to\infty} \frac{\partial^2}{\partial a^2} G(a,b) \geq 0$ for every $b \geq 0$,
(c) $\frac{\partial}{\partial a} G(a,b) \mid_{a=0} \geq 0$ for every $b \geq 0$,
(d) $G(0,b) \geq 0$ for every $b \geq 0$.

With those claims, combined with $\frac{\partial^4}{\partial a^4} G(a,b) \geq 0$, we get one by one

$$\frac{\partial^3 G}{\partial a^3} \leq 0, \quad \frac{\partial^2 G}{\partial a^2} \geq 0, \quad \frac{\partial G}{\partial a} \geq 0, \quad G \geq 0.$$

Extension to the non-symmetric case

We can partially extend our inequality to general, not necessarily even, log-concave densities. To that end, we use the following result.

Theorem 4 (Melbourne, Tkocz, 2021)

Let X and Y be iid log-concave random variables in $\mathbb{R}$. If $\alpha \in [2, \infty]$, then

$$h_\alpha(X - Y) \leq h_\alpha(X) + \log 2$$

with equality when X has exponential distribution $\mathbf{1}_{(0,\infty)}(x)e^{-x}$.

Extension to the non-symmetric case

We can partially extend our inequality to general, not necessarily even, log-concave densities. To that end, we use the following result.

Theorem 4 (Melbourne, Tkocz, 2021)

Let X and Y be iid log-concave random variables in $\mathbb{R}$. If $\alpha \in [2, \infty]$, then

$$h_\alpha(X - Y) \leq h_\alpha(X) + \log 2$$

with equality when X has exponential distribution $\mathbf{1}_{(0,\infty)}(x)e^{-x}$.

Corollary 5

Let X be a log-concave random variable in $\mathbb{R}$ and let $\alpha \geq 2$. Then

$$h_\alpha(X) \geq \frac{1}{2} \log \mathrm{Var}(X) + \frac{\log \alpha}{\alpha - 1}$$

with equality for one-sided exponential variable.