Entropic Limit Theorems

Arnaud Marsiglietti

University of Florida

Conference in honour of Sergey Bobkov
May 31, 2023

Joint with Sergey Bobkov
Jiange Li, James Melbourne

Rényi Entropy and Central Limit Theorem

For independent and identically distributed (i.i.d.) centered random variables with finite second moment $X_1, X_2, \ldots$, consider

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

Rényi entropy of order $r \in [0, +\infty]$:

$$h_r(X) = \frac{1}{1-r} \log \left(\int f^r \right), \qquad X \sim f$$

Rényi Entropy and Central Limit Theorem

For independent and identically distributed (i.i.d.) centered random variables with finite second moment $X_1, X_2, \ldots$, consider

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

Rényi entropy of order $r \in [0, +\infty]$:

$$h_r(X) = \frac{1}{1-r} \log \left(\int f^r \right), \qquad X \sim f$$

Rényi divergence:

$$D_r(X||Y) = \frac{1}{r-1} \log \int \left(\frac{f}{g} \right)^r g, \qquad X \sim f, Y \sim g$$

Rényi Entropy and Central Limit Theorem

For independent and identically distributed (i.i.d.) centered random variables with finite second moment $X_1, X_2, \ldots$, consider

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

Rényi entropy of order $r \in [0, +\infty]$**:**

$$h_r(X) = \frac{1}{1-r} \log \left(\int f^r \right), \qquad X \sim f$$

Rényi divergence:

$$D_r(X||Y) = \frac{1}{r-1} \log \int \left(\frac{f}{g} \right)^r g, \qquad X \sim f, Y \sim g$$

- **Bobkov-Chistyakov-Götze ('18):** When Z Gaussian with same covariance as X_1, established necessary and sufficient conditions for

$$D_r(Z_n||Z) \to 0.$$

Rényi Entropy and Central Limit Theorem

For independent and identically distributed (i.i.d.) centered random variables with finite second moment $X_1, X_2, \dots$, consider

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

Rényi entropy of order $r \in [0, +\infty]$**:**

$$h_r(X) = \frac{1}{1-r} \log \left(\int f^r \right), \qquad X \sim f$$

Rényi divergence:

$$D_r(X||Y) = \frac{1}{r-1} \log \int \left(\frac{f}{g} \right)^r g, \qquad X \sim f, Y \sim g$$

- **Bobkov-Chistyakov-Götze ('18):** When Z Gaussian with same covariance as X_1, established necessary and sufficient conditions for

$$D_r(Z_n||Z) \to 0.$$

Rem: For $r = 1$, $D(Z_n||Z) = h(Z) - h(Z_n)$.

Rényi Entropy and Central Limit Theorem

For independent and identically distributed (i.i.d.) centered random variables with finite second moment $X_1, X_2, \ldots$, consider

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

Rényi entropy of order $r \in [0, +\infty]$:

$$h_r(X) = \frac{1}{1-r} \log \left(\int f^r \right), \qquad X \sim f$$

Rényi divergence:

$$D_r(X||Y) = \frac{1}{r-1} \log \int \left(\frac{f}{g} \right)^r g, \qquad X \sim f, Y \sim g$$

- **Bobkov-Chistyakov-Götze ('18):** When Z Gaussian with same covariance as X_1, established necessary and sufficient conditions for

$$D_r(Z_n||Z) \to 0.$$

Rem: For $r = 1$, $D(Z_n||Z) = h(Z) - h(Z_n)$. For $r \neq 1$,

$$D_r(Z_n||Z) \neq h_r(Z) - h_r(Z_n)$$

Rényi Entropy and Central Limit Theorem

For independent and identically distributed (i.i.d.) centered random variables, consider

$$Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

Rényi entropy of order $r \in [0, +\infty]$:

$$h_r(X) = \frac{1}{1-r} \log \left(\int f^r \right), \qquad X \sim f$$

Rényi divergence:

$$D_r(X||Y) = \frac{1}{r-1} \log \int \left(\frac{f}{g} \right)^r g, \qquad X \sim f, Y \sim g$$

- **Bobkov-Chistyakov-Götze ('18):** When Z Gaussian with same covariance as X_1, established necessary and sufficient conditions for

$$D_r(Z_n||Z) \to 0.$$

Rem: For $r = 1$, $D(Z_n||Z) = h(Z) - h(Z_n)$. For $r \neq 1$,

$$\underbrace{D_r(Z_n||Z)}_{\geq 0} \neq \underbrace{h_r(Z) - h_r(Z_n)}_{\geq 0 \; or \leq 0}$$

Convergence of Z_n in Rényi entropy of order $r > 1$

- $Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$
- Z Gaussian with same covariance as X_1.

Theorem (Bobkov-M. '19)

Let $r > 1$. The following statements are equivalent.

(1) $h_r(Z_n) \to h_r(Z)$.

(2) $h_r(Z_{n_0})$ *is finite for some integer* n_0.

(3) Z_{n_0} *has a bounded density for some integer* n_0.

Convergence of Z_n in Rényi entropy of order $r > 1$

- $Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$
- Z Gaussian with same covariance as X_1.

Theorem (Bobkov-M. '19)

Let $r > 1$. The following statements are equivalent.

(1) $h_r(Z_n) \to h_r(Z)$.

(2) $h_r(Z_{n_0})$ *is finite for some integer* n_0.

(3) Z_{n_0} *has a bounded density for some integer* n_0.

Rem:

- The statement fails when $r \in (0, 1]$.
- For $r = 1$, Barron noted that (3) implies (1) but (1) does **not** imply (3).

Convergence of Z_n in Rényi entropy of order $r > 1$

- $Z_n = \frac{X_1 + \cdots + X_n}{\sqrt{n}}$
- Z Gaussian with same covariance as X_1.

Theorem (Bobkov-M. '19)

Let $r > 1$. The following statements are equivalent.

(1) $h_r(Z_n) \to h_r(Z)$.

(2) $h_r(Z_{n_0})$ *is finite for some integer* n_0.

(3) Z_{n_0} *has a bounded density for some integer* n_0.

Rem:

- The statement fails when $r \in (0, 1]$.
- For $r = 1$, Barron noted that (3) implies (1) but (1) does **not** imply (3).
- For $r \in (0, 1)$, one can see that (3) does **not** imply (1).

Convergence of Z_n in Rényi entropy of order $r \in (0,1)$

Motivated by Rényi EPIs of order $r \in (0,1)$ we established convergence for special classes of distribution:

Convergence of Z_n in Rényi entropy of order $r \in (0,1)$

Motivated by Rényi EPIs of order $r \in (0,1)$ we established convergence for special classes of distribution:

Def: $X \sim f$ log-concave if $f = e^{-V}$, with V convex

Theorem (Li-M.-Melbourne '20)

For $X_1, \ldots, X_n$ i.i.d. **log-concave** *random variables in $\mathbb{R}^d$,*

$$\lim_{n \to +\infty} h_r(Z_n) = h_r(Z)$$

Convergence of Z_n in Rényi entropy of order $r \in (0,1)$

Motivated by Rényi EPIs of order $r \in (0,1)$ we established convergence for special classes of distribution:

Def: $X \sim f$ log-concave if $f = e^{-V}$, with V convex

Theorem (Li-M.-Melbourne '20)

For $X_1, \ldots, X_n$ i.i.d. **log-concave** *random variables in $\mathbb{R}^d$,*

$$\lim_{n \to +\infty} h_r(Z_n) = h_r(Z)$$

Def: f spherically symmetric if $f(x) = F(\|x\|)$, with some function $F\colon [0,+\infty) \to [0,+\infty)$, $\|\cdot\|$ Euclidean norm

Theorem (Li-M.-Melbourne '20)

For $X_1, \ldots, X_n$ i.i.d. random variables in $\mathbb{R}^d$ with spherically symmetric unimodal density with compact support,

$$\lim_{n \to +\infty} h_r(Z_n) = h_r(Z)$$

Application to Rényi EPI

- $N_r(X) = e^{\frac{2}{d} h_r(X)}$

Shannon EPI:
For independent random variables $X_1, \ldots, X_n$ in $\mathbb{R}^d$,

$$N\left(X_1 + \cdots + X_n\right) \geq \sum_{i=1}^{n} N\left(X_i\right)$$

Application to Rényi EPI

- $N_r(X) = e^{\frac{2}{d} h_r(X)}$

Shannon EPI:
For independent random variables $X_1, \ldots, X_n$ in $\mathbb{R}^d$,

$$N_r(X_1 + \cdots + X_n) \geq c \sum_{i=1}^{n} N_r(X_i) \qquad ?$$

Application to Rényi EPI

- $N_r(X) = e^{\frac{2}{d} h_r(X)}$

Shannon EPI:
For independent random variables $X_1, \ldots, X_n$ in $\mathbb{R}^d$,

$$N_r(X_1 + \cdots + X_n) \geq c \sum_{i=1}^{n} N_r(X_i) \qquad ?$$

When $r > 1$: Bobkov-Chistyakov '15, Ram-Sason '16.

Theorem (Li-M.-Melbourne '20)

For any $r \in (0,1)$, and $\varepsilon > 0$, there exist independent random variables $X_1, \ldots, X_n$ in $\mathbb{R}^d$, for some $d \geq 1$ and $n \geq 2$, such that

$$N_r(X_1 + \cdots + X_n) < \varepsilon \sum_{i=1}^{n} N_r(X_i).$$

Smoothed CLT

Let $(X_n)_{n\geq 1}$ be independent, identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ with an isotropic distribution:

$$\mathrm{E}[X_1] = 0, \quad \mathrm{Cov}(X_1) = Id.$$

Let X be a random variable in $\mathbb{R}^d$ with finite second moment, independent of all X_n's.

Smoothed CLT

Let $(X_n)_{n\geq 1}$ be independent, identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ with an isotropic distribution:

$$\mathrm{E}[X_1] = 0, \quad \mathrm{Cov}(X_1) = Id.$$

Let X be a random variable in $\mathbb{R}^d$ with finite second moment, independent of all X_n's. By the central limit theorem (CLT), the normalized sums

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in distribution (as $n \to \infty$) to the standard normal Z.

Smoothed CLT

Let $(X_n)_{n\geq 1}$ be independent, identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ with an isotropic distribution:

$$\mathrm{E}[X_1] = 0, \quad \mathrm{Cov}(X_1) = Id.$$

Let X be a random variable in $\mathbb{R}^d$ with finite second moment, independent of all X_n's. By the central limit theorem (CLT), the normalized sums

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in distribution (as $n \to \infty$) to the standard normal Z.

Smoothed CLT: Suppose that X has an absolutely continuous distribution, so that Z_n has some density p_n.

Smoothed CLT

Let $(X_n)_{n\geq 1}$ be independent, identically distributed (i.i.d.) random variables in $\mathbb{R}^d$ with an isotropic distribution:

$$\mathrm{E}[X_1] = 0, \quad \mathrm{Cov}(X_1) = Id.$$

Let X be a random variable in $\mathbb{R}^d$ with finite second moment, independent of all X_n's. By the central limit theorem (CLT), the normalized sums

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

converges in distribution (as $n \to \infty$) to the standard normal Z.

Smoothed CLT: Suppose that X has an absolutely continuous distribution, so that Z_n has some density p_n.

Question: Can the weak CLT be strengthened to the convergence of entropies:

$$h(Z_n) \to h(Z)$$

to the entropy of the Gaussian limit Z?

Entropic CLT

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

Entropic CLT

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

Shannon entropy: If X has density p, then

$$h(X) = -\int_{\mathbb{R}^d} p(x) \log p(x)\, dx$$

Entropic CLT

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

Shannon entropy: If X has density p, then

$$h(X) = -\int_{\mathbb{R}^d} p(x) \log p(x)\, dx$$

The usual entropic CLT corresponds to $X = 0$:

Theorem: (Barron '86)
The entropic CLT $h(Z_n) \to h(Z)$ holds if and only if Z_n have densities p_n with finite $h(Z_n)$ for some n large enough.

Entropic CLT

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

Shannon entropy: If X has density p, then

$$h(X) = -\int_{\mathbb{R}^d} p(x) \log p(x) \, dx$$

The usual entropic CLT corresponds to $X = 0$:

Theorem: (Barron '86)
The entropic CLT $h(Z_n) \to h(Z)$ holds if and only if Z_n have densities p_n with finite $h(Z_n)$ for some n large enough.

Theorem: (Artstein-Ball-Barthe-Naor '04, Madiman-Barron '07)
Monotonicity in the entropic CLT: $h(Z_n) \leq h(Z_{n+1})$.

Entropic CLT

$$Z_n = \frac{X + X_1 + \cdots + X_n}{\sqrt{n}}$$

Shannon entropy: If X has density p, then

$$h(X) = -\int_{\mathbb{R}^d} p(x) \log p(x)\, dx$$

The usual entropic CLT corresponds to $X = 0$:

Theorem: (Barron '86)
The entropic CLT $h(Z_n) \to h(Z)$ holds if and only if Z_n have densities p_n with finite $h(Z_n)$ for some n large enough.

Theorem: (Artstein-Ball-Barthe-Naor '04, Madiman-Barron '07)
Monotonicity in the entropic CLT: $h(Z_n) \leq h(Z_{n+1})$.

Rate of convergence in the entropic CLT:
Miclo '03, Bobkov-Chistyakov-Götze '13, Eldan-Mikulincer-Zhai '18

Results - Part I

Introduce the **characteristic function** of X:

$$f(t) = \mathrm{E}[e^{i\langle t, X\rangle}], \quad t \in \mathbb{R}^d.$$

Results - Part I

Introduce the **characteristic function** of X:

$$f(t) = \mathrm{E}[e^{i\langle t,X\rangle}], \quad t \in \mathbb{R}^d.$$

Theorem (Bobkov, M. '20)

If f is compactly supported, and X_1 has a non-lattice distribution, then

$$h(Z_n) \to h(Z) \quad \text{as } n \to \infty.$$

Results - Part I

Introduce the **characteristic function** of X:

$$f(t) = \mathrm{E}[e^{i\langle t, X\rangle}], \quad t \in \mathbb{R}^d.$$

Theorem (Bobkov, M. '20)

If f is compactly supported, and X_1 has a non-lattice distribution, then

$$h(Z_n) \to h(Z) \quad \text{as } n \to \infty.$$

Convergence in entropy also holds for lattice distributions, if f is supported on the ball $|t| \leq T$ for some $T > 0$ depending on the distribution of X_1.

Results - Part I

Introduce the **characteristic function** of X:

$$f(t) = \mathrm{E}[e^{i\langle t, X\rangle}], \quad t \in \mathbb{R}^d.$$

Theorem (Bobkov, M. '20)

If f is compactly supported, and X_1 has a non-lattice distribution, then

$$h(Z_n) \to h(Z) \quad \text{as } n \to \infty.$$

Convergence in entropy also holds for lattice distributions, if f is supported on the ball $|t| \leq T$ for some $T > 0$ depending on the distribution of X_1.

For example, one may take $T = 1/\beta_3$, where

$$\beta_3 = \sup_{|\theta|=1} \mathrm{E}\,|\,\langle X_1, \theta\rangle\,|^3 < +\infty.$$

Results - Part I

Introduce the **characteristic function** of X:

$$f(t) = \mathrm{E}[e^{i\langle t,X\rangle}], \quad t \in \mathbb{R}^d.$$

Theorem (Bobkov, M. '20)

If f is compactly supported, and X_1 has a non-lattice distribution, then

$$h(Z_n) \to h(Z) \quad \text{as } n \to \infty.$$

Convergence in entropy also holds for lattice distributions, if f is supported on the ball $|t| \leq T$ for some $T > 0$ depending on the distribution of X_1.

For example, one may take $T = 1/\beta_3$, where

$$\beta_3 = \sup_{|\theta|=1} \mathrm{E}\,|\,\langle X_1, \theta\rangle\,|^3 < +\infty.$$

Remark: The assumption of compactness on the support of the characteristic function of X requires its density p to be the restriction to $\mathbb{R}^d$ of an entire function on $\mathbb{C}^d$ of exponential type (Paley-Wiener theorems).

Motivations - Part I

The usual entropic CLT doesn't handle the discrete case.

Motivations - Part I

The usual entropic CLT doesn't handle the discrete case.

Distances:

• The (quadratic) Kantorovich distance between $X \sim \mu$ and $Y \sim \nu$ is defined as

$$W_2(X,Y) = \inf_{\widetilde{X}\sim\mu,\widetilde{Y}\sim\nu} \mathrm{E}[|\widetilde{X}-\widetilde{Y}|^2]^{\frac{1}{2}}.$$

Motivations - Part I

The usual entropic CLT doesn't handle the discrete case.

Distances:

- The (quadratic) Kantorovich distance between $X \sim \mu$ and $Y \sim \nu$ is defined as

$$W_2(X, Y) = \inf_{\widetilde{X} \sim \mu, \widetilde{Y} \sim \nu} \mathrm{E}[|\widetilde{X} - \widetilde{Y}|^2]^{\frac{1}{2}}.$$

- The Kullback-Leibler distance between $X \sim p$ and $Y \sim q$ is defined as

$$D(X||Y) = \int_{\mathbb{R}^d} p(x) \log \frac{p(x)}{q(x)} \, dx.$$

Motivations - Part I

The usual entropic CLT doesn't handle the discrete case.

Distances:

- The (quadratic) Kantorovich distance between $X \sim \mu$ and $Y \sim \nu$ is defined as

$$W_2(X,Y) = \inf_{\widetilde{X}\sim\mu,\widetilde{Y}\sim\nu} \mathrm{E}[|\widetilde{X}-\widetilde{Y}|^2]^{\frac{1}{2}}.$$

- The Kullback-Leibler distance between $X \sim p$ and $Y \sim q$ is defined as

$$D(X||Y) = \int_{\mathbb{R}^d} p(x) \log \frac{p(x)}{q(x)}\, dx.$$

When Y is the standard Gaussian measure Z on $\mathbb{R}^d$, the relationship of W_2 with relative entropy was emphasized by Talagrand (1996) who showed that

$$W_2^2(X,Z) \leq 2\, D(X||Z).$$

Motivations - Part I

Returning to the setting of the above Theorem, define the normalized sums

$$Z_n' = Z_n - \frac{1}{\sqrt{n}} X = \frac{X_1 + \cdots + X_n}{\sqrt{n}}.$$

Motivations - Part I

Returning to the setting of the above Theorem, define the normalized sums

$$Z_n' = Z_n - \frac{1}{\sqrt{n}} X = \frac{X_1 + \cdots + X_n}{\sqrt{n}}.$$

Choose for f a characteristic function supported on a suitable small ball $|t| \leq T$, so that $D(Z_n||Z) \to 0$ by our theorem.

Motivations - Part I

Returning to the setting of the above Theorem, define the normalized sums

$$Z_n' = Z_n - \frac{1}{\sqrt{n}} X = \frac{X_1 + \cdots + X_n}{\sqrt{n}}.$$

Choose for f a characteristic function supported on a suitable small ball $|t| \leq T$, so that $D(Z_n||Z) \to 0$ by our theorem. Applying the Talagrand transport-entropy inequality, we get

$$W_2^2(Z_n', Z) \leq 2W_2^2(Z_n, Z) + \frac{2}{n} \mathrm{E}\,|X|^2 \leq 4\,D(Z_n||Z) + \frac{2}{n} \mathrm{E}\,|X|^2 \ \to \ 0.$$

Motivations - Part I

Returning to the setting of the above Theorem, define the normalized sums

$$Z_n' = Z_n - \frac{1}{\sqrt{n}} X = \frac{X_1 + \cdots + X_n}{\sqrt{n}}.$$

Choose for f a characteristic function supported on a suitable small ball $|t| \leq T$, so that $D(Z_n||Z) \to 0$ by our theorem. Applying the Talagrand transport-entropy inequality, we get

$$W_2^2(Z_n', Z) \leq 2W_2^2(Z_n, Z) + \frac{2}{n} \mathrm{E}\,|X|^2 \leq 4\,D(Z_n||Z) + \frac{2}{n} \mathrm{E}\,|X|^2 \;\to\; 0.$$

$\Longrightarrow$ As the result, Z_n' converges to Z in Wasserstein distance. The X_i's can be discrete.

Results - Part II

Necessary and sufficient condition for the uniform on discrete cube:

Theorem (Bobkov, M. '20)

Suppose that X_1 has a uniform distribution on the discrete cube $\{-1,1\}^d$, that is, with independent Bernoulli coordinates. Assume the characteristic function f of X satisfies

$$\int_{\mathbb{R}^d} |f(t)|\, dt < \infty, \quad \int_{\mathbb{R}^d} \frac{|f'(t)|}{\|t\|^{d-1}}\, dt < \infty,$$

where $\|t\|$ denotes the distance from the point t to the lattice $\pi\mathbb{Z}^d$. Then, the entropic CLT holds true, if and only if

$$f(\pi k) = 0 \quad \text{for all } \ k \in \mathbb{Z}^d,\ k \neq 0.$$

Results - Part II

Necessary and sufficient condition for the uniform on discrete cube:

Theorem (Bobkov, M. '20)

Suppose that X_1 has a uniform distribution on the discrete cube $\{-1,1\}^d$, that is, with independent Bernoulli coordinates. Assume the characteristic function f of X satisfies

$$\int_{\mathbb{R}^d} |f(t)|\, dt < \infty, \quad \int_{\mathbb{R}^d} \frac{|f'(t)|}{\|t\|^{d-1}}\, dt < \infty,$$

where $\|t\|$ denotes the distance from the point t to the lattice $\pi\mathbb{Z}^d$. Then, the entropic CLT holds true, if and only if

$$f(\pi k) = 0 \quad \text{for all } k \in \mathbb{Z}^d,\ k \neq 0.$$

Remark: The second moment assumption on X guarantees that f has a bounded continuous derivative $f'(t) = \nabla f(t)$ with its Euclidean length $|f'(t)|$. The assumption of integrability is fulfilled, for example, under decay assumptions (say, $\frac{1}{|t|^{1+\varepsilon}}$).

Motivations - Part II

Theorem (Bobkov, M. '20)

Let $(X_n)_{n\geq 1}$ be a sequence of independent, integer valued random variables, whose components have variance one. Then

$$\limsup_{n\to\infty} h(Z_n) \leq h(X) + h(Z).$$

As a consequence, if $h(Z_n) \to h(Z)$ as $n \to \infty$, then necessarily $h(X) \geq 0$.

Motivations - Part II

Theorem (Bobkov, M. '20)

Let $(X_n)_{n\geq 1}$ be a sequence of independent, integer valued random variables, whose components have variance one. Then

$$\limsup_{n\to\infty} h(Z_n) \leq h(X) + h(Z).$$

As a consequence, if $h(Z_n) \to h(Z)$ as $n \to \infty$, then necessarily $h(X) \geq 0$.

Hyperplane Conjecture: (Bourgain '86)
For any convex body K in $\mathbb{R}^d$ there is a hyperplane H such that the $(d-1)$-dimensional volume of the slice $H \cap K$ is bounded away from zero by a universal positive constant.

Motivations - Part II

Theorem (Bobkov, M. '20)

Let $(X_n)_{n\geq 1}$ be a sequence of independent, integer valued random variables, whose components have variance one. Then

$$\limsup_{n\to\infty} h(Z_n) \leq h(X) + h(Z).$$

As a consequence, if $h(Z_n) \to h(Z)$ as $n \to \infty$, then necessarily $h(X) \geq 0$.

Hyperplane Conjecture: (Bourgain '86)
For any convex body K in $\mathbb{R}^d$ there is a hyperplane H such that the $(d-1)$-dimensional volume of the slice $H \cap K$ is bounded away from zero by a universal positive constant.

Bobkov-Madiman ('11): The hyperplane conjecture is equivalent to the following statement: If X is a random variable in $\mathbb{R}^d$ with an isotropic log-concave distribution then

$$h(X) \geq -cd.$$

with some universal constant $c > 0$.

Proof of Theorems

Let $(X_n)_{n\geq 1}$ be a sequence of integer valued random variables in $\mathbb{R}^d$, and let X be a continuous random variable in $\mathbb{R}^d$ with finite second moment, independent of this sequence. As before, we define the normalized sums

$$Z_n = \frac{1}{\sqrt{n}}\,(X + X_1 + \cdots + X_n).$$

Proof of Theorems

Let $(X_n)_{n\geq 1}$ be a sequence of integer valued random variables in $\mathbb{R}^d$, and let X be a continuous random variable in $\mathbb{R}^d$ with finite second moment, independent of this sequence. As before, we define the normalized sums

$$Z_n = \frac{1}{\sqrt{n}}\,(X + X_1 + \cdots + X_n).$$

Well known: When the second moment $\mathrm{E}\,|U|^2$ of a continuous random variable U in $\mathbb{R}^d$ is fixed, its entropy is maximized on the normal distribution with the same second moment.

Proof of Theorems

Let $(X_n)_{n\geq 1}$ be a sequence of integer valued random variables in $\mathbb{R}^d$, and let X be a continuous random variable in $\mathbb{R}^d$ with finite second moment, independent of this sequence. As before, we define the normalized sums

$$Z_n = \frac{1}{\sqrt{n}}(X + X_1 + \cdots + X_n).$$

Well known: When the second moment $\mathrm{E}\,|U|^2$ of a continuous random variable U in $\mathbb{R}^d$ is fixed, its entropy is maximized on the normal distribution with the same second moment. In the case of independent and isotropic X_n's, we have

$$\mathrm{E}[|Z_n|^2] = \frac{1}{n}\mathrm{E}[|X|^2] + d \to_{n\to+\infty} d.$$

Hence

$$\limsup_{n\to\infty} h(Z_n) \leq h(Z), \quad Z \text{ standard Gaussian in } \mathbb{R}^d.$$

Proof of Theorems

Let $(X_n)_{n\geq 1}$ be a sequence of integer valued random variables in $\mathbb{R}^d$, and let X be a continuous random variable in $\mathbb{R}^d$ with finite second moment, independent of this sequence. As before, we define the normalized sums

$$Z_n = \frac{1}{\sqrt{n}}(X + X_1 + \cdots + X_n).$$

Well known: When the second moment $\mathrm{E}\,|U|^2$ of a continuous random variable U in $\mathbb{R}^d$ is fixed, its entropy is maximized on the normal distribution with the same second moment. In the case of independent and isotropic X_n's, we have

$$\mathrm{E}[|Z_n|^2] = \frac{1}{n}\mathrm{E}[|X|^2] + d \to_{n\to+\infty} d.$$

Hence

$$\limsup_{n\to\infty} h(Z_n) \leq h(Z), \quad Z \text{ standard Gaussian in } \mathbb{R}^d.$$

We proved: (When X_n's are integer valued)

$$\limsup_{n\to\infty} h(Z_n) \leq h(Z) + h(X).$$

Proof of Theorems

Let $(X_n)_{n\geq 1}$ be a sequence of integer valued random variables in $\mathbb{R}^d$, and let X be a continuous random variable in $\mathbb{R}^d$ with finite second moment, independent of this sequence. As before, we define the normalized sums

$$Z_n = \frac{1}{\sqrt{n}}(X + X_1 + \cdots + X_n).$$

Well known: When the second moment $\mathrm{E}\,|U|^2$ of a continuous random variable U in $\mathbb{R}^d$ is fixed, its entropy is maximized on the normal distribution with the same second moment. In the case of independent and isotropic X_n's, we have

$$\mathrm{E}[|Z_n|^2] = \frac{1}{n}\mathrm{E}[|X|^2] + d \to_{n\to+\infty} d.$$

Hence

$$\limsup_{n\to\infty} h(Z_n) \leq h(Z), \quad Z \text{ standard Gaussian in } \mathbb{R}^d.$$

We proved: (When X_n's are integer valued)

$$\limsup_{n\to\infty} h(Z_n) \leq h(Z) + h(X).$$

- We expect $h(X)$ to be very negative (otherwise, it would satisfy the hyperplane conjecture).

Proof of Theorems

We proved: (When X_n's are integer valued)

$$\limsup_{n\to\infty} h(Z_n) \leq h(Z) + h(X).$$

Proof of Theorems

We proved: (When X_n's are integer valued)

$$\limsup_{n\to\infty} h(Z_n) \le h(Z) + h(X).$$

Based on two elementary lemmas, which involve the discrete Shannon entropy

$$H(Y) = -\sum_k p_k \log p_k.$$

Here, Y is a discrete random variable taking at most countably many values, say y_k, with probabilities p_k respectively.

Proof of Theorems

Lemma

Let X be a continuous random variable, and let Y be a discrete random variable independent of X. Then,

$$h(X+Y) \leq h(X) + H(Y).$$

Rem: False if Y is continuous.

Proof of Theorems

Proof: Denote by p the density of X and let $p_k = P\{Y = y_k\}$. Since X and Y are independent, $X + Y$ has density

$$q(z) = \sum_k p_k p(z - y_k).$$

Proof of Theorems

Proof: Denote by p the density of X and let $p_k = P\{Y = y_k\}$. Since X and Y are independent, $X + Y$ has density

$$q(z) = \sum_k p_k p(z - y_k).$$

We use the convention $u \log(u) = 0$ if $u = 0$. Note that, if $p(z - y_k) = 0$, then

$$p_k p(z - y_k) \log \sum_i p_i p(z - y_i) = 0 = p_k p(z - y_k) \log(p_k p(z - y_k)),$$

Proof of Theorems

Proof: Denote by p the density of X and let $p_k = P\{Y = y_k\}$. Since X and Y are independent, $X + Y$ has density

$$q(z) = \sum_k p_k p(z - y_k).$$

We use the convention $u\log(u) = 0$ if $u = 0$. Note that, if $p(z - y_k) = 0$, then

$$p_k p(z - y_k)\log\sum_i p_i p(z - y_i) = 0 = p_k p(z - y_k)\log(p_k p(z - y_k)),$$

while in the case $p(z - y_k) > 0$, we have

$$p_k p(z-y_k)\log\sum_i p_i p(z-y_i) = p_k p(z-y_k)\log\left(p_k p(z-y_k) + \sum_{i\neq k} p_i p(z-y_i)\right)$$

$$= p_k p(z - y_k)\left[\log(p_k p(z - y_k)) + \log\left(1 + \frac{\sum_{i\neq k} p_i p(z - y_i)}{p_k p(z - y_k)}\right)\right]$$

$$\geq p_k p(z - y_k)\log(p_k p(z - y_k)).$$

Proof of Theorems

Hence, for all z,

$$p_k p(z - y_k) \log \sum_i p_i p(z - y_i) \geq p_k p(z - y_k) \log(p_k p(z - y_k)).$$

Proof of Theorems

Hence, for all z,

$$p_k p(z-y_k)\,\log\sum_i p_i p(z-y_i) \ge p_k p(z-y_k)\,\log(p_k p(z-y_k)).$$

We may therefore conclude that

$$
\begin{aligned}
h(X+Y) &= -\int_{\mathbb{R}^d} q(z)\log q(z)\,dz \\
&= -\sum_k \int_{\mathbb{R}^d} p_k p(z-y_k)\,\log\sum_i p_i p(z-y_i)\,dz \\
&\le -\sum_k \int_{\mathbb{R}^d} p_k p(z-y_k)\log(p_k p(z-y_k))\,dz \\
&= -\sum_k p_k\left(\int_{\mathbb{R}^d} p(z-y_k)\log p_k\,dz + \int_{\mathbb{R}^d} p(z-y_k)\log p(z-y_k)\,dz\right) \\
&= h(X)+H(Y).
\end{aligned}
$$

Proof of Theorems

Lemma

Let X be a continuous random variable, and let Y be a discrete random variable independent of X. Then,

$$h(X+Y) \leq h(X) + H(Y).$$

Proof of Theorems

Lemma

Let X be a continuous random variable, and let Y be a discrete random variable independent of X. Then,

$$h(X+Y) \leq h(X) + H(Y).$$

Sharpening: (Melbourne-Madiman-Salapaka, 2019)

$$h(X+Y) \leq h(X|Y) + TH(Y),$$

where $h(X|Y)$ is the conditional entropy, reducing to $h(X)$ on independence, and T is the supremum of the total variation of the conditional densities from their “mixture complements”, necessarily $T \leq 1$.

Lemma

For any integer valued random variable Y with finite second moment,

$$H(Y) \leq \frac{1}{2}\log\left(2\pi e\left(\mathrm{Var}(Y) + \frac{1}{12}\right)\right).$$

The proof also combines both discrete and differential entropy:

Lemma

For any integer valued random variable Y with finite second moment,

$$H(Y) \le \frac{1}{2}\log\left(2\pi e\left(\mathrm{Var}(Y)+\frac{1}{12}\right)\right).$$

The proof also combines both discrete and differential entropy:

Proof:
Put $p_k = \mathbb{P}\{Y = k\}$, $k \in \mathbb{Z}$. Consider a continuous random variable $\widetilde{Y}$ with density q defined to be

$$q(x) = p_k \quad \text{if } x \in (k - \tfrac{1}{2}, k + \tfrac{1}{2}).$$

In other words, $\quad q(x) = \sum_k p_k 1_{(k-\frac{1}{2},k+\frac{1}{2})}(x), \quad x \in \mathbb{R}.$
Note that

$$\mathrm{E}[\widetilde{Y}] = \sum_k p_k \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} x\,dx = \sum_k \frac{p_k}{2}\left(\left(k+\frac{1}{2}\right)^2 - \left(k-\frac{1}{2}\right)^2\right) = \sum_k kp_k = \mathrm{E}[Y]$$

and similarly

$$\mathrm{E}[\widetilde{Y}^2] = \sum_k p_k \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} x^2\,dx = \mathrm{E}[Y^2] + \frac{1}{12}.$$

Proof of Theorems

Hence $\mathrm{Var}(\widetilde{Y}) = \mathrm{Var}(Y) + \frac{1}{12}$. Also,

$$h(\widetilde{Y}) = -\int_{-\infty}^{\infty} \sum_k p_k 1_{(k-\frac{1}{2},k+\frac{1}{2})}(x) \log \sum_j p_j 1_{(j-\frac{1}{2},j+\frac{1}{2})}(x)\, dx$$

$$= -\sum_k p_k \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \log p_k \, dx \;=\; H(Y).$$

Proof of Theorems

Hence $\mathrm{Var}(\widetilde{Y}) = \mathrm{Var}(Y) + \frac{1}{12}$. Also,

$$h(\widetilde{Y}) = -\int_{-\infty}^{\infty} \sum_k p_k 1_{(k-\frac{1}{2},k+\frac{1}{2})}(x) \log \sum_j p_j 1_{(j-\frac{1}{2},j+\frac{1}{2})}(x)\, dx$$

$$= -\sum_k p_k \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \log p_k \, dx \; = \; H(Y).$$

Now, since Gaussian distributions maximize the differential entropy for a fixed variance, we conclude that

$$H(Y) = h(\widetilde{Y}) \leq \frac{1}{2} \log \left(2\pi e \,\mathrm{Var}(\widetilde{Y})\right) = \frac{1}{2} \log \left(2\pi e \left(\mathrm{Var}(Y) + \frac{1}{12}\right)\right).$$

Proof of Theorems

Theorem (Bobkov, M. '20)

Given a sequence $X_n = (X_{n,1}, \dots, X_{n,d})$ *of random variables with values in* $\mathbb{Z}^d$*, independent of* X*, assume that for each* $k \le d$*, the components* $X_{n,k}$*,* $n \ge 1$*, are uncorrelated and have variance one. Then,*

$$\limsup_{n \to \infty} h(Z_n) \le h(X) + h(Z).$$

Proof of Theorems

Theorem (Bobkov, M. '20)

Given a sequence $X_n = (X_{n,1}, \dots, X_{n,d})$ of random variables with values in $\mathbb{Z}^d$, independent of X, assume that for each $k \le d$, the components $X_{n,k}$, $n \ge 1$, are uncorrelated and have variance one. Then,

$$\limsup_{n\to\infty} h(Z_n) \le h(X) + h(Z).$$

Proof: Putting $S_n = X_1 + \dots + X_n$ and applying Lemma 1 above, we get

$$\begin{aligned} h(Z_n) = h\Big(\frac{X+S_n}{\sqrt{n}}\Big) &= h(X+S_n) - \frac{d}{2}\log n \\ &\le h(X) + H(S_n) - \frac{d}{2}\log n. \end{aligned}$$

Note that

$$S_n = (S_{n,1}, \dots, S_{n,d}), \qquad S_{n,k} = X_{1,k} + \dots + X_{n,k} \quad (1 \le k \le d).$$

Proof of Theorems

By the well-known subadditivity of entropy along components of a random variable (an abstract property on product spaces which is irrelevant to the independence assumption), we have

$$H(S_n) \leq H(S_{n,1}) + \cdots + H(S_{n,d}).$$

Here, the entropy functional on the left is applied to the *d*-dimensional random variable, while on the right-hand side we deal with one-dimensional entropies.

Proof of Theorems

By the well-known subadditivity of entropy along components of a random variable (an abstract property on product spaces which is irrelevant to the independence assumption), we have

$$H(S_n) \leq H(S_{n,1}) + \cdots + H(S_{n,d}).$$

Here, the entropy functional on the left is applied to the d-dimensional random variable, while on the right-hand side we deal with one-dimensional entropies. For each $k \leq d$, the k-th component $S_{n,k}$ of the random variable S_n represents the sum of n uncorrelated integer valued random variables with variance one, so that $\mathrm{Var}(S_{n,k}) = n$.

Proof of Theorems

By the well-known subadditivity of entropy along components of a random variable (an abstract property on product spaces which is irrelevant to the independence assumption), we have

$$H(S_n) \leq H(S_{n,1}) + \cdots + H(S_{n,d}).$$

Here, the entropy functional on the left is applied to the d-dimensional random variable, while on the right-hand side we deal with one-dimensional entropies. For each $k \leq d$, the k-th component $S_{n,k}$ of the random variable S_n represents the sum of n uncorrelated integer valued random variables with variance one, so that $\mathrm{Var}(S_{n,k}) = n$. Hence, by Lemma 2 applied to $Y = S_{n,k}$, we have

$$H(S_{n,k}) \leq \frac{1}{2} \log\left(2\pi e\left(n + \frac{1}{12}\right)\right) = \frac{1}{2} \log(2\pi e n) + O(1/n),$$

Proof of Theorems

By the well-known subadditivity of entropy along components of a random variable (an abstract property on product spaces which is irrelevant to the independence assumption), we have

$$H(S_n) \le H(S_{n,1}) + \cdots + H(S_{n,d}).$$

Here, the entropy functional on the left is applied to the d-dimensional random variable, while on the right-hand side we deal with one-dimensional entropies. For each $k \le d$, the k-th component $S_{n,k}$ of the random variable S_n represents the sum of n uncorrelated integer valued random variables with variance one, so that $\mathrm{Var}(S_{n,k}) = n$. Hence, by Lemma 2 applied to $Y = S_{n,k}$, we have

$$H(S_{n,k}) \le \frac{1}{2} \log\left(2\pi e\left(n + \frac{1}{12}\right)\right) = \frac{1}{2}\log(2\pi e n) + O(1/n),$$

and therefore

$$H(S_n) \le \frac{d}{2}\log(2\pi e n) + O(1/n).$$

We conclude that

$$\limsup_{n\to\infty} h(Z_n) \le h(X) + \frac{d}{2}\log(2\pi e) = h(X) + h(Z).$$

Thank you!