Hypergeometric means and their completion

Jairo Bochi
Penn State

Ohio State Math. Dept. Colloquium
March 9, 2023

What is a mean? An ad hoc definition

$\Omega =$ set (perhaps finite)

To "each" function $f : \Omega \to \mathbb{R}_+$, we want to associate a number $\mathfrak{m}(f)$ so that the following properties hold:

- **reflexivity**: $f = c$ (constant) $\Rightarrow \mathfrak{m}(f) = c$
- **monotonicity**: $f \leq g \Rightarrow \mathfrak{m}(f) \leq \mathfrak{m}(g)$
- **homogeneity**: $\forall \lambda \in \mathbb{R}_+, \ \mathfrak{m}(\lambda f) = \lambda \mathfrak{m}(f)$

Examples

- **Arithmetic mean**:

$$\mathfrak{am}(x_1, \ldots, x_n) := \frac{x_1 + \cdots + x_n}{n}$$

Functional version: if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, then

$$\mathfrak{am}(f) = \mathbb{E}(f) := \int f \, d\mathbb{P} \qquad (a.k.a. \textbf{ expectation}).$$

Examples

- **Arithmetic mean**:

$$\mathfrak{am}(x_1, \ldots, x_n) := \frac{x_1 + \cdots + x_n}{n}$$

Functional version: if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, then

$$\mathfrak{am}(f) = \mathbb{E}(f) := \int f \, d\mathbb{P} \qquad (a.k.a. \textbf{ expectation}).$$

- **Geometric mean**:

$$\mathfrak{gm}(x_1, \ldots, x_n) := (x_1 \cdots x_n)^{\frac{1}{n}}$$

Functional version:

$$\mathfrak{gm}(f) := \exp \int \log f \, d\mathbb{P}$$

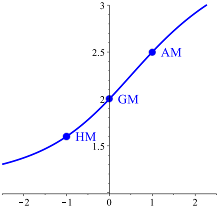
Hölder means (a.k.a. power means)

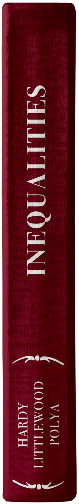
Definition

Given a positive function f on the probability space $(\Omega, \mathbb{P})$, the **Hölder mean** of f with parameter $p \in \mathbb{R}$ is:

$$\mathcal{M}_p(f) := \begin{cases} \left(\int f^p \, d\mathbb{P}\right)^{\frac{1}{p}} & \text{if } p \neq 0\,, \\ \exp \int \log f \, d\mathbb{P} & \text{if } p = 0\,. \end{cases}$$

Continuity and monotonicity wrt parameter:





Characterization of Hölder means

Given a mean $\mathfrak{m}$ and a homeomorphism ϕ (increasing or decreasing), consider a new (not necessarily homogeneous) mean

$$\widetilde{\mathfrak{m}}(f) := \phi^{-1}(\mathfrak{m}(\phi \circ f))$$

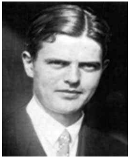
If $\mathfrak{m} = \mathfrak{am}$, then $\widetilde{\mathfrak{m}}$ is called a **quasi-arithmetic mean**. Example: $\widetilde{\mathfrak{m}} = \mathfrak{gm}$, $\phi = \log$.

Andrey Kolmogorov 1903–1987

Mitio Nagumo 1905–1993

Bruno de Finetti 1906–1985

Georg Aumann 1906–1980



Børge Jessen 1907–1993

Theorem (∼ 1930)

Hölder means are the only quasi-arithmetic homogeneous means.

A famous example of "nonlinear averaging"

Given numbers $b \geq a > 0$, define $[a_0, b_0] := [a, b]$ and recursively

$$[a_{n+1}, b_{n+1}] := [\mathfrak{gm}(a_n, b_n), \mathfrak{am}(a_n, b_n)] .$$

These intervals shrink superexponentially fast to a point c, called the **arithmetic geometric mean (AGM)** of a and b.

A famous example of "nonlinear averaging"

Given numbers $b \geq a > 0$, define $[a_0, b_0] := [a, b]$ and recursively

$$[a_{n+1}, b_{n+1}] := [\mathfrak{gm}(a_n, b_n), \mathfrak{am}(a_n, b_n)]\,.$$

These intervals shrink superexponentially fast to a point c, called the **arithmetic geometric mean (AGM)** of a and b.

Theorem (Lagrange 1785, Gauss 1800)

The AGM is related to an elliptic integral:

$$\frac{1}{\mathfrak{agm}(a,b)} = \frac{1}{2\pi}\int_0^{2\pi} \frac{d\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}}$$

Symmetric means

Using the elementary symmetric polynomials, we form the **symmetric means**:

$$\mathfrak{sym}_k(x_1, \ldots, x_n) := \left(\frac{\sum_{i_1 < \cdots < i_k} x_{i_1} \cdots x_{i_k}}{\binom{n}{k}} \right)^{\frac{1}{k}}$$

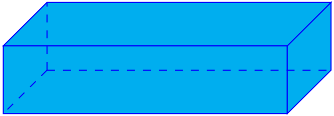
Thus, $\mathfrak{sym}_1 \equiv \mathfrak{am}$ and $\mathfrak{sym}_n \equiv \mathfrak{gm}$.

Symmetric means

Using the elementary symmetric polynomials, we form the **symmetric means**:

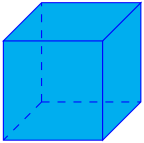
$$\mathfrak{sym}_k(x_1, \ldots, x_n) := \left(\frac{\sum_{i_1 < \cdots < i_k} x_{i_1} \cdots x_{i_k}}{\binom{n}{k}} \right)^{\frac{1}{k}}$$

Thus, $\mathfrak{sym}_1 \equiv \mathfrak{am}$ and $\mathfrak{sym}_n \equiv \mathfrak{gm}$.



box with sides a, b, c

$\rightarrow$



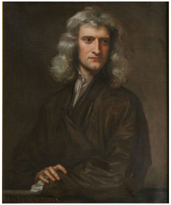
cube of side $\mathfrak{sym}_2(a, b, c) = \sqrt{\frac{ab+ac+bc}{3}}$

equal surface areas

Symmetric means (continued)

Monotonicity wrt parameter (Maclaurin, 1729):

$$\mathfrak{sym}_1 \geq \mathfrak{sym}_2 \geq \cdots \geq \mathfrak{sym}_{n-1} \geq \mathfrak{sym}_n$$



Isaac Newton
1642–1727

Colin Maclaurin
1698–1746

Limit laws for symmetric means?

Let $X_1, X_2, \ldots$ be a sequence of positive i.i.d. random variables (or, more generally, a stationary ergodic process). To avoid discussing integrability issues, let's assume uniform boundedness away from 0 and ∞.

By the Law of Large Numbers (or the Ergodic Theorem),

$$\frac{X_1 + \cdots + X_n}{n} \to \mathbb{E}(X) \quad \text{almost surely as } n \to \infty,$$

where X is a replica of the X_i's.

Limit laws for symmetric means?

Let $X_1, X_2, \ldots$ be a sequence of positive i.i.d. random variables (or, more generally, a stationary ergodic process). To avoid discussing integrability issues, let's assume uniform boundedness away from 0 and ∞.

By the Law of Large Numbers (or the Ergodic Theorem),

$$\frac{X_1 + \cdots + X_n}{n} \to \mathbb{E}(X) \quad \text{almost surely as } n \to \infty,$$

where X is a replica of the X_i's.

Theorem

For any $k \geq 1$,

$$\mathfrak{sym}_k(X_1, \ldots, X_n) \to \mathbb{E}(X) \quad \text{a.s. when } n \to \infty.$$

Boring situation: No new limit. 😴

A more exciting limit law

Theorem (Hálasz–Székely 1976)

Let $k_1, k_2, \ldots$ be a sequence of integers such that $1 \le k_n \le n$ and k_n/n tends to some $c \in [0,1]$ as $n \to \infty$. Then, with probability 1,

$$\lim_{n\to\infty} \mathfrak{sym}_{k_n}(X_1, \ldots, X_n)$$

exists and equals a "computable" number $\mathfrak{HS}_{-c}(X)$ that depends only on the distribution of X and on the parameter c.

A more exciting limit law

Theorem (Hálasz–Székely 1976)

Let $k_1, k_2, \ldots$ be a sequence of integers such that $1 \le k_n \le n$ and k_n/n tends to some $c \in [0,1]$ as $n \to \infty$. Then, with probability 1,

$$\lim_{n\to\infty} \mathfrak{sym}_{k_n}(X_1, \ldots, X_n)$$

exists and equals a "computable" number $\mathfrak{HS}_{-c}(X)$ that depends only on the distribution of X and on the parameter c.

Gábor Hálasz
1941–

Gábor J. Székely
1947–

Note: Székely was personally instigated by Kolmogorov to work on this (or a closely related) problem.

(Extended) Halász-Székely means

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

Let f be a positive function on Ω with $\|f^{\pm 1}\|_\infty < \infty$.

Definition

The **Halász-Székely mean** of f with parameter $\lambda \in \mathbb{R}$ is:

$$\mathfrak{HS}_\lambda(f) := \sup_{g>0} \left(\frac{\exp \int \log g \, d\mathbb{P}}{\int g \, d\mathbb{P}} \right)^{\frac{1}{\lambda}} \frac{\int fg \, d\mathbb{P}}{\int g \, d\mathbb{P}} \quad \text{if } \lambda > 0.$$

(Extended) Halász-Székely means

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

Let f be a positive function on Ω with $\|f^{\pm 1}\|_\infty < \infty$.

Definition

The **Halász-Székely mean** of f with parameter $\lambda \in \mathbb{R}$ is:

$$\mathfrak{HS}_\lambda(f) := \sup_{g>0} \left(\frac{\exp \int \log g \, d\mathbb{P}}{\int g \, d\mathbb{P}} \right)^{\frac{1}{\lambda}} \frac{\int fg \, d\mathbb{P}}{\int g \, d\mathbb{P}} \quad \text{if } \lambda > 0.$$

If $\lambda < 0$, we use the same formula with inf in place of sup.
If $\lambda = 0$, then $\mathfrak{HS}_0(f) := \int f \, d\mathbb{P}$.

(Extended) Halász-Székely means

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

Let f be a positive function on Ω with $\|f^{\pm 1}\|_\infty < \infty$.

Definition

The **Halász-Székely mean** of f with parameter $\lambda \in \mathbb{R}$ is:

$$\mathfrak{HS}_\lambda(f) := \sup_{g>0} \left(\frac{\exp \int \log g \, d\mathbb{P}}{\int g \, d\mathbb{P}} \right)^{\frac{1}{\lambda}} \frac{\int fg \, d\mathbb{P}}{\int g \, d\mathbb{P}} \quad \text{if } \lambda > 0.$$

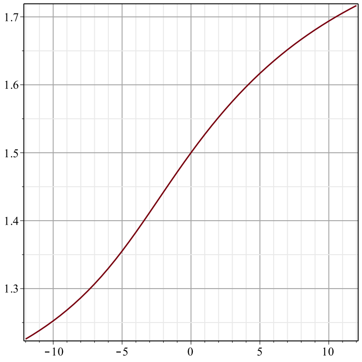
If $\lambda < 0$, we use the same formula with inf in place of sup.
If $\lambda = 0$, then $\mathfrak{HS}_0(f) := \int f \, d\mathbb{P}$.

Note: $\mathfrak{HS}_\lambda(f)$ only depends on the measure $f_* \mathbb{P}$. It is a "nonlinear barycenter" of this measure.

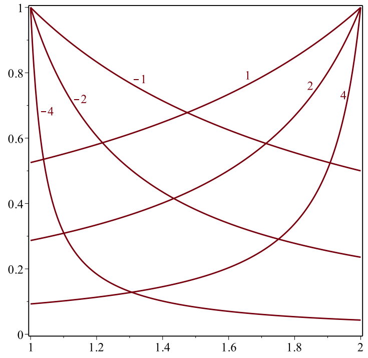
Special case: $\mathfrak{HS}_{-1}(f) = \exp \int \log f \, d\mathbb{P}$ (geometric mean).

Worked-out example

Plot of $\mathfrak{HS}_\lambda(f)$ as a function of λ if $\mathbb{P}$ = Lebesgue measure on $[1, 2]$ and $f(x) \equiv x$:



Functions $g = g_\lambda$ that attain the sup or inf (chosen so that $\max g = 1$):



How did I compute this?

Practical computation

Recall: $$\mathfrak{HS}_\lambda(f) := \begin{cases} \sup_{g>0} \left(\dfrac{\exp \mathbb{E}(\log g)}{\mathbb{E}(g)} \right)^{\frac{1}{\lambda}} \dfrac{\mathbb{E}(gf)}{\mathbb{E}(g)} & \text{if } \lambda > 0, \\ \mathbb{E}(f) & \text{if } \lambda = 0, \\ \inf \text{(same stuff)} & \text{if } \lambda < 0. \end{cases}$$

A positive function g that attains either the sup (if $\lambda > 0$) or the inf (if $\lambda < 0$) above is called an **equilibrium state**.

Proposition

Equilibrium states g are the essentially unique, and are characterized as the positive solutions of the eq.:

$$\frac{1}{g} + \frac{\lambda f}{\mathbb{E}(gf)} = \frac{1+\lambda}{\mathbb{E}(g)}$$

(which reduces to a scalar equation for $\xi := \mathbb{E}(fg)/\mathbb{E}(g)$...)

Limit theorems, revisited

Theorem (Halász–Székely 1976)

$$\frac{k_n}{n} \to c \in [0, 1] \quad \Rightarrow \quad \mathfrak{sym}_{k_n}(X_1, \ldots, X_n) \to \mathfrak{HS}_{-c}(X) \ a.s.$$

Limit theorems, revisited

Theorem (Halász–Székely 1976)

$$\frac{k_n}{n} \to c \in [0,1] \quad \Rightarrow \quad \mathfrak{sym}_{k_n}(X_1, \dots, X_n) \to \mathfrak{HS}_{-c}(X) \ a.s.$$

A deterministic version:

Theorem (B.–Iommi–Ponce 2021)

Let X be a discrete random variable assuming distinct values $a_1, \dots, a_n$, each with probability $\frac{1}{n}$. If $1 \le k \le n$, then

$$\mathfrak{HS}_{-k/n}(X) \le \mathfrak{sym}_k(a_1, \dots, a_n) \le (9k)^{\frac{1}{2k}} \, \mathfrak{HS}_{-k/n}(X) \, .$$

Combining these inequalities with continuity properties of $\mathfrak{HS}$ (both as function of the parameter and the distribution), the previous theorem follows.

Mean rates of expansion of a linear operator

Let $A\colon \mathbb{R}^n \to \mathbb{R}^n$ be linear. The **expansion rate** of A is the function

$$v \in S^{n-1} \mapsto \|Av\|$$

where $S^{n-1} := \{v \in \mathbb{R}^n : \|v\| = 1\}$ is the unit sphere.

Mean rates of expansion of a linear operator

Let $A\colon \mathbb{R}^n \to \mathbb{R}^n$ be linear. The **expansion rate** of A is the function

$$v \in S^{n-1} \mapsto \|Av\|$$

where $S^{n-1} := \{v \in \mathbb{R}^n : \|v\| = 1\}$ is the unit sphere. We compute the Hölder means of this function wrt the area measure σ:

$$\mathfrak{em}_p(A) := \left(\int_{S^{n-1}} \|Av\|^p \, d\sigma(v)\right)^{\frac{1}{p}} \quad \text{if } p \neq 0\,,$$

$$\mathfrak{em}_0(A) := \exp \int_{S^{n-1}} \log \|Av\| \, d\sigma(v)\,.$$

Mean rates of expansion of a linear operator

Let $A\colon \mathbb{R}^n \to \mathbb{R}^n$ be linear. The **expansion rate** of A is the function

$$v \in S^{n-1} \mapsto \|Av\|$$

where $S^{n-1} := \{v \in \mathbb{R}^n : \|v\| = 1\}$ is the unit sphere. We compute the Hölder means of this function wrt the area measure σ:

$$\mathfrak{em}_p(A) := \left(\int_{S^{n-1}} \|Av\|^p \, d\sigma(v) \right)^{\frac{1}{p}} \quad \text{if } p \neq 0\,,$$

$$\mathfrak{em}_0(A) := \exp \int_{S^{n-1}} \log \|Av\| \, d\sigma(v)\,.$$

Definition

The **ellipsoidal mean** with parameter p of numbers $a_1, \ldots, a_n \geq 0$ is $\mathfrak{em}_p(a_1, \ldots, a_n) = \mathfrak{em}_p(A)$, where $A = \operatorname{diag}(a_1, \ldots, a_n)$.

Some special cases

Recall:

$$\mathfrak{em}_p(a_1, \ldots, a_n) := \left(\int_{S^{n-1}} \|Av\|^p \, d\sigma(v) \right)^{\frac{1}{p}}, \quad A = \operatorname{diag}(a_1, \ldots, a_n).$$

Exercises:

$$\mathfrak{em}_2(A) = \sqrt{\tfrac{1}{n} \sum a_i^2} = \mathfrak{M}_2(a_1, \ldots, a_n)$$

$$\mathfrak{em}_{-n}(A) = \sqrt[n]{\prod a_i} = \mathfrak{M}_0(a_1, \ldots, a_n)$$

Are ellipsoidal means always Hölder? Definitely no!

Other special ellipsoidal means

- If $n = 2$, $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then

$$
\begin{aligned}
\mathfrak{em}_{-1}(a, b) &= \left(\int_{S^1} \|Av\|^{-1} \, d\sigma(v) \right)^{-1} \\
&= \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} \right)^{-1} \\
&= \mathfrak{agm}(a, b)
\end{aligned}
$$

Other special ellipsoidal means

- If $n = 2$, $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then

$$\mathfrak{em}_{-1}(a, b) = \left(\int_{S^1} \|Av\|^{-1} \, d\sigma(v) \right)^{-1}$$
$$= \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} \right)^{-1}$$
$$= \mathfrak{agm}(a, b)$$

- $\mathfrak{em}_0(a, b) = \frac{a+b}{2}$ (Haruki, 1991).

Other special ellipsoidal means

- If $n = 2$, $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then

$$\mathfrak{em}_{-1}(a, b) = \left(\int_{S^1} \|Av\|^{-1} \, d\sigma(v) \right)^{-1}$$
$$= \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} \right)^{-1}$$
$$= \mathfrak{agm}(a, b)$$

- $\mathfrak{em}_0(a, b) = \frac{a+b}{2}$ (Haruki, 1991).
- $\mathfrak{em}_1(a, b) = \frac{1}{2\pi} \cdot$ perimeter of the ellipse with semiaxes a, b (the quintessential elliptic integral).

Other special ellipsoidal means

- If $n = 2$, $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then

$$\mathfrak{em}_{-1}(a, b) = \left(\int_{S^1} \|Av\|^{-1} \, d\sigma(v) \right)^{-1}$$
$$= \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} \right)^{-1}$$
$$= \mathfrak{agm}(a, b)$$

- $\mathfrak{em}_0(a, b) = \frac{a+b}{2}$ (Haruki, 1991).
- $\mathfrak{em}_1(a, b) = \frac{1}{2\pi} \cdot$ perimeter of the ellipse with semiaxes a, b (the quintessential elliptic integral).
- Generalization: $\text{area}(A(S^{n-1})) = \mathfrak{em}_1(\Lambda^{n-1} A) \cdot \text{area}(S^{n-1})$.

Other special ellipsoidal means

- If $n = 2$, $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, then

$$\mathfrak{em}_{-1}(a, b) = \left(\int_{S^1} \|Av\|^{-1} \, d\sigma(v) \right)^{-1} = \left(\frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}} \right)^{-1} = \mathfrak{agm}(a, b)$$

- $\mathfrak{em}_0(a, b) = \frac{a+b}{2}$ (Haruki, 1991).
- $\mathfrak{em}_1(a, b) = \frac{1}{2\pi} \cdot$ perimeter of the ellipse with semiaxes a, b (the quintessential elliptic integral).
- Generalization: $\text{area}(A(S^{n-1})) = \mathfrak{em}_1(\Lambda^{n-1} A) \cdot \text{area}(S^{n-1})$.
- Also related to: mean width and electrostatic capacities of ellipsoids, spherical functions, Lyapunov exponents, Kullback-Leiber divergence.

Limit laws for ellipsoidal means?

Consider a sequence A_n of diagonal matrices of increasing dimensions $n \times n$

$$A_n = \operatorname{diag}(a_{n,1}, \ldots, a_{n,n}), \quad a_{n,i} > 0\,.$$

Assume that there exists a well-defined limit distribution of the diagonal entries:

$$\mu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{a_{n,i}} \xrightarrow{\textit{weakly}} \mu \quad \text{as } n \to \infty\,.$$

Limit laws for ellipsoidal means?

Consider a sequence A_n of diagonal matrices of increasing dimensions $n \times n$

$$A_n = \operatorname{diag}(a_{n,1}, \dots, a_{n,n}), \quad a_{n,i} > 0 .$$

Assume that there exists a well-defined limit distribution of the diagonal entries:

$$\mu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{a_{n,i}} \xrightarrow{weakly} \mu \quad \text{as } n \to \infty .$$

Theorem

If the numbers $a_{n,i}$ are bounded away from 0 and ∞, then, for any $p \in \mathbb{R}$,

$$\mathfrak{em}_p(A_n) \to \mathcal{M}_2(\mu) \quad \textit{as } n \to \infty .$$

Rivin 2004: More precise results (weaker assumption, estimates), but only for $p = 0$ or 1.

Proving that $\mathfrak{em}_p(A_n) \to \mathcal{M}_2(\mu)$

- Consider the expansion rate function

$$f_n \colon S^{n-1} \to \mathbb{R}_+ , \qquad f_n(v) := \|A_n v\| .$$

- As mentioned before, $\mathfrak{em}_2 \equiv \mathcal{M}_2$ (in particular, the theorem is trivial for $p = 2$). This means that f_n^2 has (arithmetic) mean equal to $\mathcal{M}_2(\mu_n)^2$.
- The function f_n^2 is Lipschitz (with bounds).

Proving that $\mathfrak{em}_p(A_n) \to \mathcal{M}_2(\mu)$

- Consider the expansion rate function

$$f_n \colon S^{n-1} \to \mathbb{R}_+ , \qquad f_n(v) := \|A_n v\| .$$

- As mentioned before, $\mathfrak{em}_2 \equiv \mathcal{M}_2$ (in particular, the theorem is trivial for $p = 2$). This means that f_n^2 has (arithmetic) mean equal to $\mathcal{M}_2(\mu_n)^2$.
- The function f_n^2 is Lipschitz (with bounds).
- By **concentration of measure**, if $n \gg 1$, then f_n^2 is very close to its mean on 99% of the sphere.

Proving that $\mathfrak{em}_p(A_n) \to \mathcal{M}_2(\mu)$

- Consider the expansion rate function

$$f_n \colon S^{n-1} \to \mathbb{R}_+ , \qquad f_n(v) := \|A_n v\| .$$

- As mentioned before, $\mathfrak{em}_2 \equiv \mathcal{M}_2$ (in particular, the theorem is trivial for $p = 2$). This means that f_n^2 has (arithmetic) mean equal to $\mathcal{M}_2(\mu_n)^2$.
- The function f_n^2 is Lipschitz (with bounds).
- By **concentration of measure**, if $n \gg 1$, then f_n^2 is very close to its mean on 99% of the sphere.
- Since p is fixed, it follows that $\mathfrak{em}_p(A_n)$, which is the p-Hölder mean of f_n, is very close to $\mathcal{M}_2(\mu_n)$, and therefore to $\mathcal{M}_2(\mu)$. QED

Proving that $\mathfrak{em}_p(A_n) \to \mathcal{M}_2(\mu)$

- Consider the expansion rate function

$$f_n \colon S^{n-1} \to \mathbb{R}_+ , \qquad f_n(v) := \|A_n v\| .$$

- As mentioned before, $\mathfrak{em}_2 \equiv \mathcal{M}_2$ (in particular, the theorem is trivial for $p = 2$). This means that f_n^2 has (arithmetic) mean equal to $\mathcal{M}_2(\mu_n)^2$.
- The function f_n^2 is Lipschitz (with bounds).
- By **concentration of measure**, if $n \gg 1$, then f_n^2 is very close to its mean on 99% of the sphere.
- Since p is fixed, it follows that $\mathfrak{em}_p(A_n)$, which is the p-Hölder mean of f_n, is very close to $\mathcal{M}_2(\mu_n)$, and therefore to $\mathcal{M}_2(\mu)$. QED

This theorem isn't very exciting, since the limit is nothing new. 😴

An unexpected limit law for ellipsoidal means

As before, consider a sequence A_n of matrices of increasing dimensions $n \times n$, whose singular values $a_{n,i}$ are bounded away from 0 and ∞, and have a well-defined limit distribution.

Theorem

Let $p(n)$ be a sequence such that $\frac{p(n)}{n} \to c \in \mathbb{R}$. Then,

$$\mathfrak{em}_{p(n)}(A_n) \to$$

An unexpected limit law for ellipsoidal means

As before, consider a sequence A_n of matrices of increasing dimensions $n \times n$, whose singular values $a_{n,i}$ are bounded away from 0 and ∞, and have a well-defined limit distribution.

Theorem

Let $p(n)$ be a sequence such that $\frac{p(n)}{n} \to c \in \mathbb{R}$. Then,

$$\mathfrak{em}_{p(n)}(A_n) \to \mathfrak{HS}_{c/2}(\mu^{(2)}),$$

where $\mu^{(2)}$ is the limit distribution of $a_{n,i}^2$ (the squares of the singular values).

What's the explanation for this miracle? Is there some relation between symmetric means and ellipsoidal means?

An unexpected limit law for ellipsoidal means

As before, consider a sequence A_n of matrices of increasing dimensions $n \times n$, whose singular values $a_{n,i}$ are bounded away from 0 and ∞, and have a well-defined limit distribution.

Theorem

Let $p(n)$ be a sequence such that $\frac{p(n)}{n} \to c \in \mathbb{R}$. Then,

$$\mathfrak{em}_{p(n)}(A_n) \to \mathfrak{HS}_{c/2}(\mu^{(2)}),$$

where $\mu^{(2)}$ is the limit distribution of $a^2_{n,i}$ (the squares of the singular values).

What's the explanation for this miracle? Is there some relation between symmetric means and ellipsoidal means? Well, there is no direct relation (as far as I could determine). The clue is that these two means are members of a larger family...

Dirichlet distribution

Definition

The **Dirichlet distribution** with parameter $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n_+$ is the probability measure $\theta_{\mathbf{b}}$ on the standard unit simplex

$$\Delta^{n-1} := \{(u_1, \ldots, u_n) \in \mathbb{R}^n_+ : u_1 + \cdots + u_n = 1\}$$

whose density wrt the area measure is proportional to the function

$$u_1^{b_1-1} \cdots u_n^{b_n-1} .$$

Dirichlet distribution

Definition

The **Dirichlet distribution** with parameter $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n_+$. is the probability measure $\theta_{\mathbf{b}}$ on the standard unit simplex

$$\Delta^{n-1} := \{(u_1, \ldots, u_n) \in \mathbb{R}^n_+ : u_1 + \cdots + u_n = 1\}$$

whose density wrt the area measure is proportional to the function

$$u_1^{b_1-1} \cdots u_n^{b_n-1}.$$

Two special cases:

- If $\mathbf{b} = (1, \ldots, 1)$, then $\theta_{\mathbf{b}}$ = normalized area on Δ^{n-1}.

Dirichlet distribution

Definition

The **Dirichlet distribution** with parameter $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n_+$. is the probability measure $\theta_\mathbf{b}$ on the standard unit simplex

$$\Delta^{n-1} := \{(u_1, \ldots, u_n) \in \mathbb{R}^n_+ : u_1 + \cdots + u_n = 1\}$$

whose density wrt the area measure is proportional to the function

$$u_1^{b_1 - 1} \cdots u_n^{b_n - 1}.$$

Two special cases:

- If $\mathbf{b} = (1, \ldots, 1)$, then $\theta_\mathbf{b}$ = normalized area on Δ^{n-1}.
- If $\mathbf{b} = (\frac{1}{2}, \ldots, \frac{1}{2})$, then $\theta_\mathbf{b}$ = the push-forward of normalized area on the sphere under the map

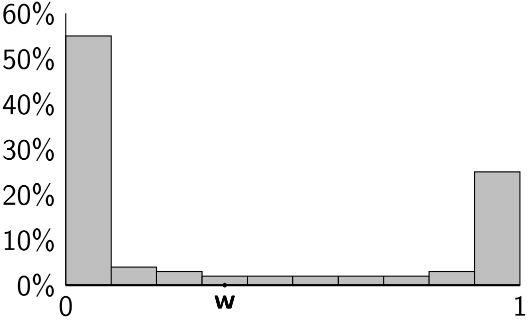
$$(x_1, \ldots, x_n) \in S^{n-1} \mapsto (x_1^2, \ldots, x_n^2) \in \Delta^{n-1}.$$

Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so Δ^{n-1} = an interval:

$s = 1/4$

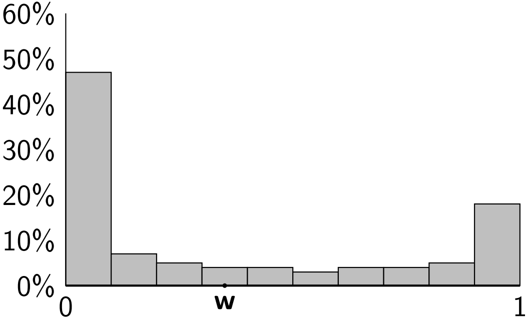


Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so Δ^{n-1} = an interval:

$s = \;\; 1/2$

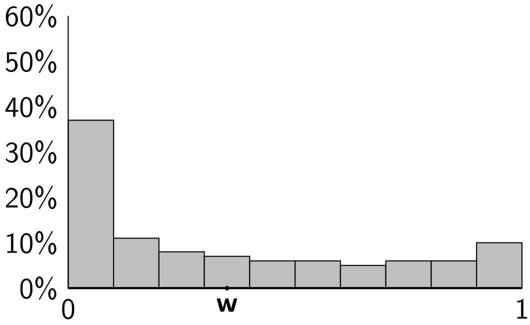


Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so Δ^{n-1} = an interval:

$s = 1$

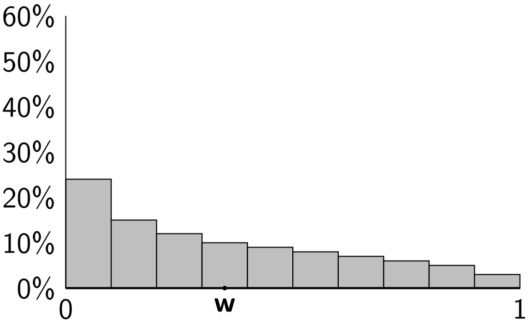


Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so Δ^{n-1} = an interval:

$s = 2$

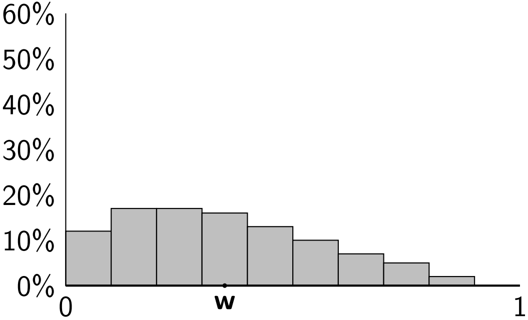


Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so $\Delta^{n-1} =$ an interval:

$s = \; 4$

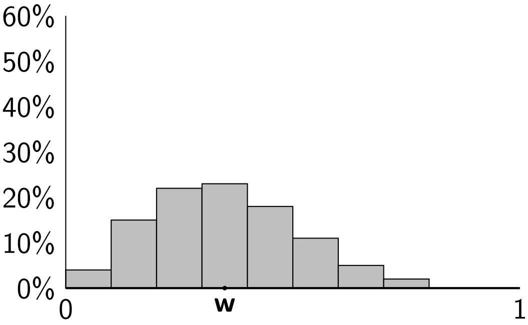


Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so $\Delta^{n-1} =$ an interval:

$s = \;\; 8$

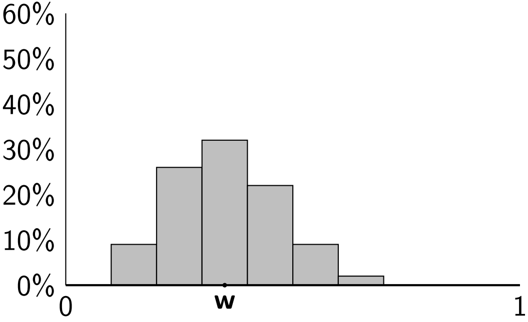


Limit cases

$$\mathbf{w} \in \Delta^{n-1} \quad \Rightarrow \quad \theta_{s\mathbf{w}} \text{ tends to } \begin{cases} \sum_i w_i \delta_{e_i} & \text{as } s \to 0^+ \\ \delta_{\mathbf{w}} & \text{as } s \to \infty \end{cases}$$

Visualization for $n = 2$, so Δ^{n-1} = an interval:

$s = 16$

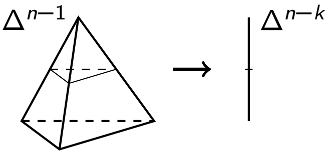


Properties (aggregation and neutrality)

Given $n > k > 1$, consider the map

$$\Delta^{n-1} \to \Delta^{n-k}$$
$$(u_1, \ldots, u_n) \mapsto (u_1 + \cdots + u_k, u_{k+1}, \ldots, u_n)$$

whose fibers are (scaled copies of) Δ^{k-1}.

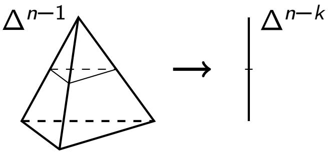


Properties (aggregation and neutrality)

Given $n > k > 1$, consider the map

$$\Delta^{n-1} \to \Delta^{n-k}$$
$$(u_1, \ldots, u_n) \mapsto (u_1 + \cdots + u_k, u_{k+1}, \ldots, u_n)$$

whose fibers are (scaled copies of) Δ^{k-1}.



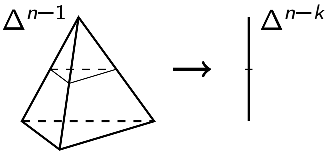
- **Aggregation property:** Dirichlet on Δ^{n-1} projects to Dirichlet on Δ^{n-k};

Properties (aggregation and neutrality)

Given $n > k > 1$, consider the map

$$\Delta^{n-1} \to \Delta^{n-k}$$
$$(u_1, \dots, u_n) \mapsto (u_1 + \cdots + u_k, u_{k+1}, \dots, u_n)$$

whose fibers are (scaled copies of) Δ^{k-1}.



- **Aggregation property:** Dirichlet on Δ^{n-1} projects to Dirichlet on Δ^{n-k};
- **Neutrality property:** the conditional measures are Dirichlets.

Carlson's hypergeometric function R (1963)

Definition

Given $\mathbf{b} = (b_1, \ldots, b_n), \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_+$, $t \in \mathbb{R}$, define:

$$R_t(\mathbf{b}; \mathbf{x}) := \int_{\Delta^{n-1}} \langle \mathbf{u}, \mathbf{x} \rangle^t \, d\theta_{\mathbf{b}}(\mathbf{u})$$

Carlson's hypergeometric function R (1963)

Definition

Given $\mathbf{b} = (b_1, \ldots, b_n), \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_+$, $t \in \mathbb{R}$, define:

$$R_t(\mathbf{b}; \mathbf{x}) := \int_{\Delta^{n-1}} \langle \mathbf{u}, \mathbf{x} \rangle^t \, d\theta_{\mathbf{b}}(\mathbf{u})$$

A special case:

$$R_{\frac{p}{2}}\left(\tfrac{1}{2}, \cdots, \tfrac{1}{2}; a_1^2, \ldots, a_n^2\right) = \int_{S^{n-1}} \|Av\|^p \, d\sigma(v)$$

where $A = \operatorname{diag}(a_1, \ldots, a_n)$.

This follows from the change of variables $u_i = v_i^2$, which (as mentioned before) sends the area measure σ on S^{n-1} to the Dirichlet measure $\theta_{\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)}$ on Δ^{n-1}.

Properties of the function R

- homogeneity of degree t:

$$R_t(\mathbf{b}; \lambda \mathbf{x}) = \lambda^t R_t(\mathbf{b}; \mathbf{x})$$

- symmetry under permutations of indices;
- aggregation property:

$$x_1 = x_2 \;\Rightarrow\; R_t(\mathbf{b}, \mathbf{x}) = R_t(b_1 + b_2, b_3, \ldots, b_n; x_1, x_3, \ldots, x_n)$$

- and LOTS of others

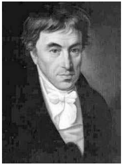
Classical hypergeometric function $F = {}_2F_1$ (Euler 1769)

$$x(1-x)\frac{d^2y}{dx^2} + [c-(a+b+1)x]\frac{dy}{dx} - aby = 0$$

$$F(a,b;c;x) = 1 + \frac{ab}{c}x + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{x^2}{2!} + \cdots$$

$$= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 u^{b-1}(1-u)^{c-b-1}(1-ux)^{-a}\,du\,.$$


Euler
1707–1783


Pfaff
1765–1825


Gauss
1777–1855


Kummer
1810–1893


Riemann
1826–1866

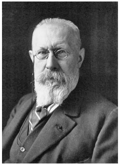
…

R versus F's

Carlson's R is related to Euler's $F = {}_2F_1$ if $n = 2$, to Appell's F_1 (1880) if $n = 3$, and to Lauricella's F_D (1893) for any n.

$$F(a, b; c; x) = R_{-a}(b, c - b; 1 - x, 1)$$
$$R_t(b_1, b_2; x_1, x_2) = x_2^t F(-t, b_1; b_1 + b_2; 1 - \frac{x_1}{x_2})$$

"The symmetry of R entails the cost of an extra variable resulting from homogeneous coordinates. To use $ax^2 + bxy + cy^2$ instead of $ax^2 + bx + c$ would be analogous."



Paul Émile Appell
1855–1930

Giuseppe Lauricella
1867–1913

Bille Chandler Carlson
1924–2013

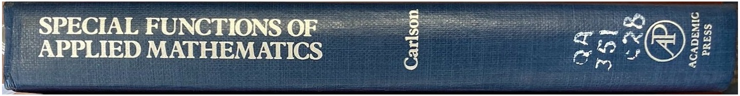
Advantages of symmetry

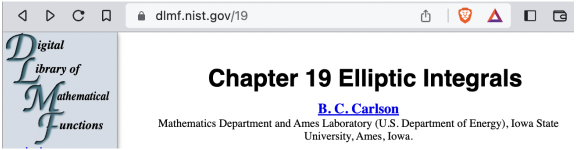
One example (among many): Pfaff's reflection law

$$\frac{1}{(1-z)^a} F\left(a, b; c; -\frac{z}{1-z}\right) = F(a, c-b; c; z) \qquad (1)$$

becomes

$$R_t(b_1, b_2; z_1, z_2) = R_t(b_2, b_1; z_2, z_1)\,. \qquad (2)$$





Carlson's hypergeometric means (1964)

Given an tuple of positive numbers $\mathbf{x}$, a weight vector $\mathbf{w} \in \Delta^{n-1}$, and parameters $t \in \mathbb{R} \setminus \{0\}$ (the **exponent**), $s > 0$ (the **concentration**), the **hypergeometric mean** is

$$\mathfrak{hygm}_{t,s}(\mathbf{x}, \mathbf{w}) := [R_t(s\mathbf{w}; \mathbf{x})]^{\frac{1}{t}} = \left(\int_{\Delta^{n-1}} \langle \mathbf{u}, \mathbf{x} \rangle^t \, d\theta_{\mathbf{b}}(\mathbf{u}) \right)^{\frac{1}{t}}$$

Carlson's hypergeometric means (1964)

Given an tuple of positive numbers $\mathbf{x}$, a weight vector $\mathbf{w} \in \Delta^{n-1}$, and parameters $t \in \mathbb{R} \setminus \{0\}$ (the **exponent**), $s > 0$ (the **concentration**), the **hypergeometric mean** is

$$\mathfrak{hygm}_{t,s}(\mathbf{x}, \mathbf{w}) := [R_t(s\mathbf{w}; \mathbf{x})]^{\frac{1}{t}} = \left(\int_{\Delta^{n-1}} \langle \mathbf{u}, \mathbf{x} \rangle^t \, d\theta_{\mathbf{b}}(\mathbf{u}) \right)^{\frac{1}{t}}$$

Note that this is a t-Hölder wrt the Dirichlet measure $\theta_{s\mathbf{w}}$. So it makes sense to extend it to $t = 0$ as the corresponding geometric mean:

$$\mathfrak{hygm}_{0,s}(\mathbf{x}, \mathbf{w}) := \exp \int_{\Delta^{n-1}} \log \langle \mathbf{u}, \mathbf{z} \rangle \, d\theta_{\mathbf{b}}(\mathbf{u})$$

(related to another hypergeometric function $L_0 := \left.\frac{\partial R_t}{\partial t}\right|_{t=0}$).

Hypergeometric means extend ellipsoidal means

$$\mathfrak{em}_p(a_1,\ldots,a_n) = \sqrt{\mathfrak{hygm}_{\frac{p}{2},\frac{n}{2}}\left(a_1^2,\ldots,a_n^2;\frac{1}{n},\cdots,\frac{1}{n}\right)}.$$

Hypergeometric means extend ellipsoidal means

$$\mathfrak{em}_p(a_1, \ldots, a_n) = \sqrt{\mathfrak{hygm}_{\frac{p}{2}, \frac{n}{2}}\left(a_1^2, \ldots, a_n^2; \frac{1}{n}, \cdots, \frac{1}{n}\right)}.$$

The hypergeometric mean is more flexible: s doesn't need to be a half-integer, and it allows for weights.

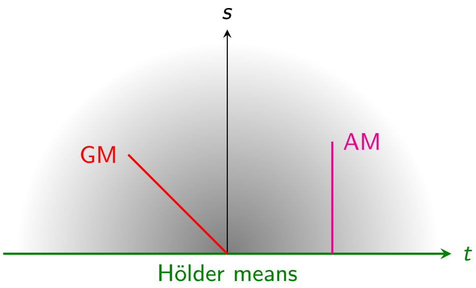
Note that weights work as they should (by the aggregation property of R).

Special cases of the hypergeometric mean

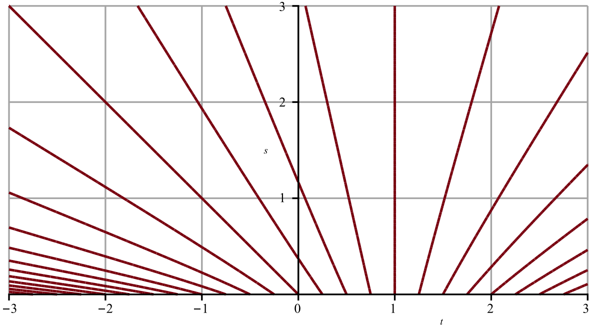
$\mathfrak{hygm}_{t,s}$ can be defined on the closed half-plane $s \geq 0$.

Special cases:

- Concentration $s = 0 \Rightarrow$ Hölder mean with exponent t
- Exponent $t = 1 \Rightarrow$ arithmetic mean
- $t = -s \Rightarrow$ geometric mean
- and many other particular means...



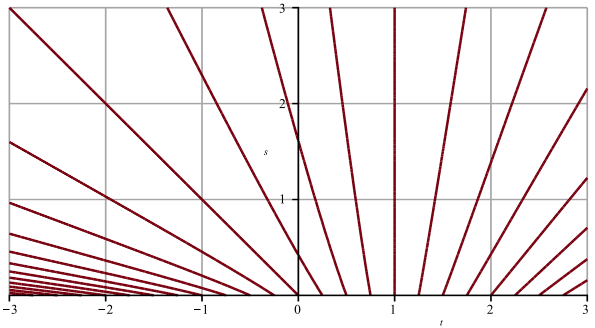
Examples: level sets



(Levels correspond to Hölder means with spacing $\frac{1}{4}$)

The levels sets are **not** straight lines, except for $s = -t$ (GM), $t = 1$ (AM) (and $t = \frac{1}{2} - \frac{s}{2}$ if $n = 2$)).

Examples: level sets



(Levels correspond to Hölder means with spacing $\frac{1}{4}$)

The levels sets are **not** straight lines, except for $s = -t$ (GM), $t = 1$ (AM) $\left(\text{and } t = \frac{1}{2} - \frac{s}{2} \text{ if } n = 2\right)$.

Boring limit theorems

Theorem (Carlson 1964)

For any fixed $t \in \mathbb{R}$,

$$\lim_{s \to +\infty} \mathfrak{hygm}_{t,s}(\mathbf{x}, \mathbf{w}) = \mathfrak{am}(\mathbf{x}, \mathbf{w}) \,.$$

Boring limit theorems

Theorem (Carlson 1964)

For any fixed $t \in \mathbb{R}$,

$$\lim_{s \to +\infty} \mathfrak{hygm}_{t,s}(\mathbf{x}, \mathbf{w}) = \mathfrak{am}(\mathbf{x}, \mathbf{w})\,.$$

Theorem (Brenner–Carlson 1987)

Suppose μ_n *is a sequence of discrete probability measures supported on a common compact subinterval of* $\mathbb{R}_+$, *and converging weakly to some probability* μ. *Let* $t \in \mathbb{R}$ *be fixed and* $s = s(n) \to +\infty$. *Then,*

$$\lim_{n \to \infty} \mathfrak{hygm}_{t,s(n)}(\mu_n) = \mathfrak{am}(\mu)\,.$$

No new limits 😴

Exciting limit theorem

Theorem

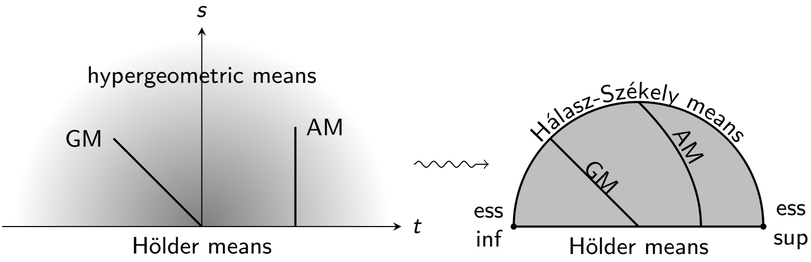
Suppose μ_n is a sequence of discrete probability measures supported on a common compact subinterval of $\mathbb{R}_+$, and converging weakly to some probability μ. If $t = t(n)$ and $s = s(n) \to +\infty$ are such that $t/s \to \lambda \in \mathbb{R}$, then

$$\lim_{n\to+\infty} \mathfrak{hygm}_{t(n),s(n)}(\mu_n) = \mathfrak{HS}_\lambda(\mu).$$

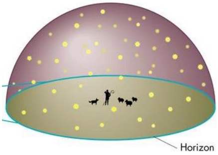
Disclaimer: $\exists$ related results on approximation of hypergeometric means using saddle point method: Jiang–Kadane–Dickey 1991, Butler–Wood 2015

A compactification of the space of parameters

Projective compactification of the space of parameters (s, t): from a half-plane to a half-disk.



This is like a "celestial sphere":



The theory is still incomplete...

Note that the Hálasz-Székely means are "functional" means (or "barycenters").

What about the hypergeometric means – do they admit a functional version?

The answer is yes!

Short explanation: Any homogeneous mean coherently defined for weighted finite lists of arbitrary lengths can be extended to a functional mean (i.e. barycenter).

To find this extension concretely, the first step is to extend the Dirichlet measures...

Ferguson-Dirichlet process

The simplex Δ^{n-1} is the space of probability measures on the finite set $F = \{1, \ldots, n\}$. Therefore, each Dirichlet $\theta_{\mathbf{b}}$ is a probability distribution on set of probabilities on F (think "a bag of loaded dice").

Ferguson-Dirichlet process

The simplex Δ^{n-1} is the space of probability measures on the finite set $F = \{1, \ldots, n\}$. Therefore, each Dirichlet $\theta_{\mathbf{b}}$ is a probability distribution on set of probabilities on F (think "a bag of loaded dice").

Theorem (Freedman 1963, Fabius 1964, Ferguson 1973)

Let β be a positive finite measure on $(\Omega, \mathcal{F})$. Then there exist a ***random probability measure*** *μ on Ω such that, for any finite measurable partition $A_1 \sqcup \cdots \sqcup A_n = X$, the distribution of the random vector $(\mu(A_1), \ldots, \mu(A_n)) \in \Delta^{n-1}$ is Dirichlet with parameter $(\beta(A_1), \ldots, \beta(A_n))$.*

This result produces a **probability measure on the space of probability measures on** Ω, called **Ferguson-Dirichlet measure** and denoted Θ_β.

Key: the aggregation property of the (finite-dimensional) Dirichlet.

Infinite-dimensional hypergeometric stuff

β = a finite measure on $\mathbb{R}_+$

$$R_t(\beta) := \int_{\mathcal{P}(\mathbb{R}_+)} \left(\int_{\mathbb{R}_+} x \, d\mu(x) \right)^t d\Theta_\beta(\mu)$$

If β is a discrete measure $\sum b_i \delta_{x_i}$, we get the previous $R_t(\mathbf{b}, \mathbf{x})$.

Disclaimer: Many people considered averages wrt Ferguson–Dirichlet. It was already known that hypergeometric functions have an infinite dimensional generalization: Lijoi–Regazzini 2004

Infinite-dimensional hypergeometric stuff

$\beta =$ a finite measure on $\mathbb{R}_+$

$$R_t(\beta) := \int_{\mathcal{P}(\mathbb{R}_+)} \left(\int_{\mathbb{R}_+} x \, d\mu(x) \right)^t d\Theta_\beta(\mu)$$

If β is a discrete measure $\sum b_i \delta_{x_i}$, we get the previous $R_t(\mathbf{b}, \mathbf{x})$.

Disclaimer: Many people considered averages wrt Ferguson–Dirichlet. It was already known that hypergeometric functions have an infinite dimensional generalization: Lijoi–Regazzini 2004

$\mu =$ a probability measure on $\mathbb{R}_+$

$$\mathfrak{hygm}_{t,s}(\mu) := [R_t(s\mu)]^{\frac{1}{t}}$$

Completion of hypergeometric means

Theorem (Main theorem)

The hypergeometric means can be extended to a ***continuous*** *map*

$$\mathfrak{hygm} : H \times \mathcal{M}_I \to \mathbb{R}_+ \ ,$$

where:

- *H is the closed half-disk (projective compactification of the half-plane);*
- *$\mathcal{M}_I$ is the space of probability measures supported on an interval $I = [a, b] \subset \mathbb{R}_+$*

What about symmetric means?

Fact: $\mathfrak{sym}_k(x_1, \ldots, x_n) = \mathfrak{hygm}_{k,-n}\left(x_1, \ldots, x_n; \frac{1}{n}, \cdots, \frac{1}{n}\right)$

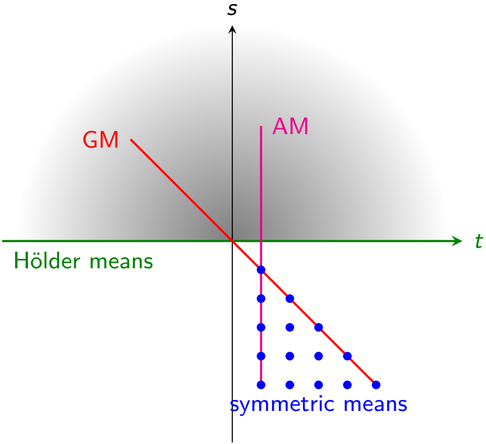
Despite $s < 0$, the RHS makes sense, because:

- R_k is a polynomial in the x_i variables, and
- $R_k(-1, \ldots, -1; x_1, \ldots, x_n) > 0$ if all $x_i > 0$.

What about symmetric means?

Fact: $\mathfrak{sym}_k(x_1, \ldots, x_n) = \mathfrak{hygm}_{k,-n}\left(x_1, \ldots, x_n; \frac{1}{n}, \cdots, \frac{1}{n}\right)$

So the limit formula also works in this case ("below the horizon").



Matrix arguments?

B.–Iommi–Ponce 2016 obtained the $\mathfrak{sym} \to \mathfrak{HS}$ law as a particular case of a "**law of large permanents**":

$$\left(\frac{1}{n!}\,\mathrm{per}(A_n)\right)^{\frac{1}{n}} \to \text{"scaling mean"}$$

(later improved by Balogh–Nguyen 2017) – that's a topic for another talk.

Question

Is this all of this part of something even bigger?

Matrix arguments?

B.–Iommi–Ponce 2016 obtained the $\mathfrak{sym} \to \mathfrak{HS}$ law as a particular case of a "**law of large permanents**":

$$\left(\frac{1}{n!}\operatorname{per}(A_n)\right)^{\frac{1}{n}} \to \text{"scaling mean"}$$

(later improved by Balogh–Nguyen 2017) – that's a topic for another talk.

Question

Is this all of this part of something even bigger?

In an effort to symmetrize other classical hypergeometric functions, Carlson (1971) introduced a hypergeometric function $\mathcal{R}$ of matrix argument. **It turns out that the permanent is a particular case of the $\mathcal{R}$ function.**