Entropic Means

Maria Barouti

University of Maryland Baltimore County
Department of Mathematics and Statistics
bmaria2@umbc.edu

October 22, 2015

Data: Let $a = (a_1, ..., a_n)$ be strictly positive numbers and let $w = (w_1, ...w_n)$ be given weights, i.e., $\sum_{i=1}^{n} w_i = 1$, $w_i > 0$, $i = 1, 2, ..., n$.

Data: Let $a = (a_1, ..., a_n)$ be strictly positive numbers and let $w = (w_1, ...w_n)$ be given weights, i.e., $\sum_{i=1}^{n} w_i = 1$, $w_i > 0$, $i = 1, 2, ..., n$.

Problem: Generate means as optimal solutions of the minimization problem

$$\min\{\sum_{i=1}^{n} w_i \text{dist}(x, a_i) : x \in \mathbb{R}_+\} \text{ (E)}$$

where w_i : "relative importance" to the error "dist(x, a_i)".

Data: Let $a = (a_1, ..., a_n)$ be strictly positive numbers and let $w = (w_1, ...w_n)$ be given weights, i.e., $\sum_{i=1}^{n} w_i = 1$, $w_i > 0$, $i = 1, 2, ..., n$.

Problem: Generate means as optimal solutions of the minimization problem

$$\min\{\sum_{i=1}^{n} w_i \text{dist}(x, a_i) : x \in \mathbb{R}_+\} \text{ (E)}$$

where w_i : "relative importance" to the error "dist(x, a_i)".

Solution: $x = mean(a_1, ..., a_n)$

Data: Let $a = (a_1, ..., a_n)$ be strictly positive numbers and let $w = (w_1, ...w_n)$ be given weights, i.e., $\sum_{i=1}^{n} w_i = 1$, $w_i > 0$, $i = 1, 2, ..., n$.

Problem: Generate means as optimal solutions of the minimization problem

$$\min\{\sum_{i=1}^{n} w_i \text{dist}(x, a_i) : x \in \mathbb{R}_+\} \text{ (E)}$$

where w_i : "relative importance" to the error "dist(x, a_i)".

Solution: $x = mean(a_1, ..., a_n)$

Idea: Choose "distance" function

Data: Let $a = (a_1, ..., a_n)$ be strictly positive numbers and let $w = (w_1, ...w_n)$ be given weights, i.e., $\sum_{i=1}^{n} w_i = 1$, $w_i > 0$, $i = 1, 2, ..., n$.

Problem: Generate means as optimal solutions of the minimization problem

$$\min\{\sum_{i=1}^{n} w_i \text{dist}(x, a_i) : x \in \mathbb{R}_+\} \text{ (E)}$$

where w_i : "relative importance" to the error "$\text{dist}(x, a_i)$".

Solution: $x = mean(a_1, ..., a_n)$

Idea: Choose "distance" function $\rightarrow$ Entropy-like function

Definition

Let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be a strictly convex differentiable function with $(0, 1] \subset \mathrm{dom}\phi$ such that $\phi(1) = 0$, $\phi'(1) = 0$. Then ϕ is called ϕ-divergence or ϕ-relative entropy.

The"distance" between two discrete probability measures is defined as

$$I_\phi(p, q) := \sum_{j=1}^{n} q_j \phi(\frac{p_j}{q_j}),$$

where $p, q \in D_n = \{x \in \mathbb{R}^n : \sum_{j=1}^{n} x_j = 1, x > 0\}$.

Let $\mathrm{dist}(x, a_i) := d_\phi(x, a_i) := a_i \phi(\frac{x}{a_i})$ for each $i = 1, ..., n$.

Restate the optimization problem (E) as

$$\min\{\sum_{i=1}^n w_i a_i \phi(\frac{x}{a_i}) : x \in \mathbb{R}_+\} \quad (E_{d_\phi})$$

Then the optimal solution of (E_{d_ϕ}) is the entropic mean

$$\bar{x}_\phi(a) := \bar{x}(a_1, ..., a_n).$$

Remarks:

- Entropic mean possesses all the essential properties of a general mean.
- $d_\phi(x, a_i)$ is not symmetric, does not satisfy the triangle inequality but it is homogeneous.

Lemma 1

Let $\phi \in \Phi = \{\phi : \mathbb{R}_+ \to \mathbb{R}$ strictly convex st: $\phi(1) = 0, \phi'(1) = 0\}$.

Then

- (a) For any $b_2 > b_1 \geq a > 0$ or $0 < b_2 < b_1 \leq a$,
$$d_\phi(b_2, a) > d_\phi(b_1, a).$$
- (b) For any $a_2 \geq a_1 > b > 0$ or $0 < a_2 < a_1 < b$,
$$d_\phi(b, a_2) > d_\phi(b, a_1).$$

Corollary 1

Let $\phi \in \Phi$ and $a, b > 0$. Then $d_\phi(b, a) \geq 0$ with equality if and only if $a = b$.

Proof

Since $\phi(1) = 0$, then $d_\phi(a, a) = a\phi(\frac{a}{a}) = a\phi(1) = 0$.
Set $b_2 = b$, $b_1 = a$ in Lemma 1(a), then $d_\phi(b, a) > d_\phi(a, a) = 0$.

Theorem 1

Let $\phi \in \Phi$. Then

(i) There is a unique continuous function $\bar{x}_\phi$ which solves (E_{d_ϕ}) such that

$$\min\{a_i\} \leq \bar{x}_\phi(a) \leq \max\{a_i\},$$

for $i = 1, 2, ..., n$ and $a_i > 0$. In particular $\bar{x}_\phi(a, a, ..., a) = a$.

(ii) The mean $\bar{x}_\phi$ is strict, i.e.,

$$\min\{a_i\} < \max\{a_i\} \Rightarrow \min\{a_i\} < \bar{x}_\phi(a) < \max\{a_i\}$$

for $i = 1, 2, ..., n$.

(iii) $\bar{x}_\phi$ is homogeneous (scale invariant), i.e.,

$$\bar{x}_\phi(\lambda a) = \lambda \bar{x}_\phi(a)$$

for $\lambda > 0$, $a_i > 0$.

Theorem 1 (Cont.)

(iv) If $w_i = w$ for all i then $\bar{x}_\phi$ is symmetric; i.e.,

$\bar{x}_\phi(a_1, a_2, ..., a_n)$ is invariant to permutations of the $a_i\ 's > 0$.

(v) $\bar{x}_\phi$ is isotone; i.e., for all i and fixed $\{a_j\}_{j=1}^n > 0$, $j \neq i$

$\bar{x}_\phi(a_1, .., a_{j-1}, \cdot\,, a_{i+1}, ..., a_n)$ is an increasing function.

Remarks:

- ϕ: convex $\Rightarrow \bar{x}_\phi$: unique and $\bar{x}_\phi \in [\min\ a_i, \max\ a_i]$.
- Means not possessing the properties of Theorem 1 cannot derived from the solution of (E_{d_ϕ}) with the entropy type distance $d_\phi(x, a_i) = a_i \phi(\frac{x}{a_i})$.

Examples

Classical means as well as many others are special cases of entropic means, for particular choice of the kernel function ϕ.

For all the cases below we solve the optimality condition equation

$$\sum_{i=1}^{n} w_i \phi'(\frac{x}{a_i}) = 0 \quad \text{(OC)}$$

(1) Arithmetic mean. Choose $\phi(t) = -\log t + t - 1$.

Then by solving the (OC) we get $\bar{x}_\phi(a) = \sum_{i=1}^{n} w_i a_i := A(a)$.

(2) Harmonic mean. Choose $\phi(t) = (t-1)^2$.

Then by solving the (OC) we get
$\bar{x}_\phi(a) = (\sum_{i=1}^{n} (\frac{w_i}{a_i}))^{-1} := H(a)$.

(3) Root mean square. Choose $\phi(t) = 1 - 2\sqrt{t} + t$.

Then by solving the (OC) we get
$\bar{x}_\phi(a) = (\sum_{i=1}^{n} w_i \sqrt{a_i})^2 := R(a)$.

Examples (Cont.)

(4) Geometric mean. Choose $\phi(t) = t \log t - t + 1$.

Then by solving the (OC) we get $\bar{x}_\phi(a) = \Pi_{i=1}^n a_i^{w_i} := G(a)$.

(5) Mean of order p. Choose $\phi_p(t) = (\frac{1}{p-1})(t^{1-p} - p) + t$, $p \neq 1, p > 0$.

Then by solving the (OC) we get $\bar{x}_{\phi_p}(a) = (\sum_{i=1}^n w_i a_i^p)^{\frac{1}{p}}$.

To extend $\bar{x}_{\phi_p}(a)$ for negative order, choose

$\phi_q(t) = (\frac{t^q - tq}{q-1}) + 1$, $q \neq 1, q > 0$, which yields

$\bar{x}_{\phi_q}(a) = (\sum_{i=1}^n w_i a_i^{1-q})^{\frac{1}{1-q}}$.

Examples (Cont.)

(6) Composition of means. Let $\phi(t) = -\frac{2}{3}\log t + (\frac{t^2}{3}) - \frac{1}{3}$.

Then by solving the (OC) we get

$\bar{x}_\phi(a) = \left(\frac{\sum_{i=1}^n w_i a_i}{\sum_{i=1}^n \frac{w_i}{a_i}}\right)^{\frac{1}{2}} = \sqrt{A(a)H(a)}$, i.e, the geometric mean of A and H.

Note: In two dimensions, $n = 2$ with $w_1 = w_2 = \frac{1}{2}$,
$\bar{x}_\phi(a_1, a_2) = \sqrt{a_1 a_2} = G(a_1, a_2)$.

Hence different choice of ϕ (compare with the geometric mean) may induce the same mean.

Theorem 2

Let ϕ, $\psi \in \Phi$ and denote by $\bar{x}_\phi$, $\bar{x}_\psi$respectively the corresponding entropic means. If there exists a constant $K \neq 0$ such that

$$K\phi'(t) \leq \psi'(t) \;\; \forall t \in \mathbb{R}_+ \backslash \{1\}$$

then

$$\bar{x}_\phi(a) \geq \bar{x}_\psi(a).$$

Theorem 2 Proof

Proof

From the optimality conditions it follows that

$$\sum_{i=1}^{n} w_i \phi'(\frac{\bar{x}_\phi}{a_i}) = 0 \text{ and } \sum_{i=1}^{n} w_i \psi'(\frac{\bar{x}_\psi}{a_i}) = 0.$$

Suppose $\bar{x}_\phi < \bar{x}_\psi$. By assumption we have that

$$K\phi'(t) \leq \psi'(t) \quad \forall t \in \mathbb{R}_+ \backslash \{1\}.$$

Then $K\phi'(\frac{\bar{x}_\phi}{a_i}) \leq \psi'(\frac{\bar{x}_\phi}{a_i}) < \psi'(\frac{\bar{x}_\psi}{a_i})$, since ψ' is strictly increasing.

Multiplying by $w_i > 0$ and summing the above inequalities imply

$$K \sum_{i=1}^{n} w_i \phi'(\frac{\bar{x}_\phi}{a_i}) < \sum_{i=1}^{n} w_i \psi'(\frac{\bar{x}_\psi}{a_i}) \Rightarrow 0 < 0.$$

Contradiction.

Example

By using examples (1) and (4) with $\phi(t) = -\log t + t - 1$ and $\psi(t) = t \log t - t + 1$ we have $\bar{x}_\phi = A(a)$ and $\bar{x}_\psi = G(a)$. The condition of Theorem 2 is satisfied with $K = 1$, i.e, $\psi'(t) \geq \phi'(t)$. Hence $G(a) \leq A(a)$.

Using now examples (2) with $z(t) = (t-1)^2$, where $\bar{x}_z = H(a)$ and (4). The condition of Theorem 2 is satisfied with $K = 2$, i.e, $z'(t) \geq 2\psi'(t)$. Hence $H(a) \leq G(a)$. Thus,

$$H(a) \leq G(a) \leq A(a).$$

Theorem 3

Let ϕ_1, $\phi_2 \in \Phi$ and $\phi_\lambda(t) := \lambda\phi_1(t) + (1-\lambda)\phi_2(t)$.

Then for all $0 \leq \lambda \leq 1$

$$\min\{\bar{x}_{\phi_1}(a), \bar{x}_{\phi_2}(a)\} \leq \bar{x}_{\phi_\lambda}(a) \leq \max\{\bar{x}_{\phi_1}(a), \bar{x}_{\phi_2}(a)\}.$$

Theorem 3 Proof

Proof

Note $0 \leq \lambda \leq 1$ and $\phi_\lambda \in \Phi$.

Now $\bar{x}_{\phi_\lambda}$ is obtained by solving the (OC): $\sum_{i=1}^{n} w_i \phi_\lambda'(\frac{\bar{x}_{\phi_\lambda}}{a_i}) = 0$. Then

$$\sum_{i=1}^{n} w_i \{\lambda \phi_1'(\frac{\bar{x}_{\phi_\lambda}}{a_i}) + (1-\lambda)\phi_2'(\frac{\bar{x}_{\phi_\lambda}}{a_i})\} = 0 \quad (1)$$

Assume $\bar{x}_{\phi_\lambda} < \min(\bar{x}_{\phi_1}, \bar{x}_{\phi_2})$ then since ϕ_1', ϕ_2' are strictly increasing we have with (1)

$$\lambda \sum_{i=1}^{n} w_i \phi_1'(\frac{\bar{x}_{\phi_1}}{a_i}) + (1-\lambda) \sum_{i=1}^{n} w_i \phi_2'(\frac{\bar{x}_{\phi_2}}{a_i}) > 0 \quad (2)$$

Theorem 3 Proof (Cont.)

But from the (OC) conditions we have that

$$\sum_{i=1}^{n} w_i \phi_1'\left(\frac{\bar{x}_{\phi_1}}{a_i}\right) = 0$$

and

$$\sum_{i=1}^{n} w_i \phi_2'\left(\frac{\bar{x}_{\phi_2}}{a_i}\right) = 0$$

Then (2) implies that $0 > 0$.

Contradiction.

Similarly for $\bar{x}_{\phi_\lambda} > \max(\bar{x}_{\phi_1}, \bar{x}_{\phi_2})$.

Example

Let $\phi_1(t) = -\log t + t - 1$ and $\phi_2(t) = (t-1)^2$. Then $\bar{x}_{\phi_1} = A(a)$ $\bar{x}_{\phi_2} = H(a)$.

Consider for $\lambda = \frac{2}{3}$, $\phi_\lambda(t) := \lambda\phi_1(t) + (1-\lambda)\phi_2(t)$ which is the function used in example 6 and so $\bar{x}_{\phi_\lambda} = \sqrt{A(a)H(a)}$.

Since $H(a) \leq A(a)$ then by Theorem 3

$$H(a) \leq \sqrt{A(a)H(a)} \leq A(a).$$

Let A be a nonnegative random variable (r.v.) with distribution F and $\text{supp}A := [a, b]$, $0 \leq a \leq b \leq +\infty$. A natural generalization of

$$\min\{\sum_{i=1}^{n} w_i a_i \phi(\tfrac{x}{a_i}) : x \in \mathbb{R}_+\} \quad (E_{d_\phi})$$

is

$$\min\{E\left(A\phi(\tfrac{x}{A})\right) := \int_a^b t\phi(\tfrac{x}{t})dF(t) : x \in \mathbb{R}_+\},$$

where $E(\cdot)$ denotes the mathematical expectation with respect to the r.v. A with distribution $F(\cdot)$.

Remark:

- (E_{d_ϕ}) corresponds to the discrete r.v. A with $Pr\{A = a_i\} := w_i$

Theorem 4

Assume that all the expectations expressions exist and are finite. Similar to Theorem 1 we have the following result.

Then for any positive random variable A:

(i) There exist a unique $\bar{x}_\phi$ which solves (E_{d_ϕ}) such that

$$\bar{x}_\phi \in \text{supp}A.$$

(ii) If A is a degenerate r.v, i.e., $A = C$ where C is a positive finite constant, $\bar{x}_\phi = C$.

(iii) For all $\lambda > 0$, $\bar{x}_\phi(\lambda A) = \lambda \bar{x}_\phi(A)$.

Examples

We can similarly derive the associated integral means by solving the optimality condition equation

$$\int_a^b \phi'(\tfrac{x}{t})dF(t) = 0 \quad \text{(OC)}$$

For example, choosing ϕ as defined in Examples (1), (2), (4) and (5) one obtains respectively:

(1) The Expectation $\bar{x}_\phi = E(A) = \int t dF(t)$.

(2) The Harmonic Expectation $\bar{x}_\phi = \frac{1}{E(1/A)} = \frac{1}{\int \frac{dF(t)}{t}}$.

(3) The Geometric integral mean $\bar{x}_\phi = e^{E \log A} = e^{\int \log t dF(t)}$.

(4) The Integral mean of order p: $\bar{x}_\phi = \{\int t^p dF(t)\}^{\frac{1}{p}}, p > 0$.

This example illustrates the derivation of an important "average" concept arising in statistics which is obtained by choosing a nondifferential kernel ϕ.

Example (θth quantile)

$$\phi(\xi) = \begin{cases} (1-\theta)(\xi - 1) & \text{if } \xi > 1 \\ \theta(1-\xi) & \text{if } 0 < \xi \leq 1 \end{cases}$$

$0 < \theta < 1$.

Remarks:

- ϕ is not differentiable at $\xi = 1$
- the subdifferential of ϕ is $\vartheta\phi(1) = [-\theta, 1-\theta]$, so $0 \in \vartheta\phi(1)$
- $d_\phi(\cdot\,,\cdot\,)$ satisfies Lemma 1 and Corollary 1

(Cont.)

The objective function of problem (E_{d_ϕ}) is

$$h(x) := E(A\phi(\tfrac{x}{A})) = \int_0^\infty t\phi(\tfrac{x}{t})dF(t) =$$

$$(1-\theta)\int_0^x (x-t)dF(t) + \theta\int_x^\infty (t-x)dF(t).$$

Since $F(\cdot)$ is the distribution function of the positive random variable A then $\int_0^\infty dF(t) = 1$. Hence,

$$h(x) = xF(x) - \theta x - \int_0^x tdF(t) + \theta E(A).$$

It follows that $h'(\bar{x}_\phi) = 0$ is simply $F(\bar{x}_\phi) = \theta$.

Note: $\bar{x}_\phi$ is the θth quantile of the continuous r.v. A.
In particular, for $\theta = \frac{1}{2}$, $\bar{x}_\phi$ is the median.

Consider the generalized mean of Hardy, Littlewood, and Polya (HLP)

$$M_\psi(a, w) = \psi^{-1}\{\sum_{i=1}^{n} w_i \psi(a_i)\},$$

where ψ is strictly monotone function with inverse ψ^{-1}.

Note: This mean is not scale invariant.

Take for example the Logarithmic mean defined in two variables $a, b > 0$, $a \neq b$, by

$$L(a, b) := \left(\frac{\log b - \log a}{b - a}\right)^{-1} = \psi^{-1}\left(\frac{\int_a^b \psi(x)dx}{b - a}\right)$$

where $\psi(x) := \frac{1}{x}$.

Goal: Consider a new entropy-like distance function $\Rightarrow$ derive the generalized mean of HLP.

Let $h : \mathbb{R}_+ \to \mathbb{R}$ be a strictly convex differentiable function and let $\alpha, \beta \in \mathbb{R}_+$. We define

$$D_h(\alpha, \beta) := h(\alpha) - h(\beta) - (\alpha - \beta)h'(\beta).$$

Since h is convex $\Rightarrow$ gradient inequality holds for $h(\cdot)$.

Then it follows that $D_h(\alpha, \beta) = 0$ if $\alpha = \beta$ and $D_h(\alpha, \beta) > 0$ if $\alpha \neq \beta$.

Remark:

- $D_h(\alpha, \beta)$ is not symmetric (except for $h(t)$ quadratic), does not satisfy the triangle inequality and is not homogeneous (as $d_\phi(\cdot\,,\cdot)$ was).

Problem: Adapting the "distance " D_h as

$$D_h(x, a_i) := h(x) - h(a_i) - (x - a_i)h'(a_i)$$

problem (E_{D_h}) is now

$$\min\{\sum_{i=1}^n w_i D_h(x, a_i) : x \in \mathbb{R}_+\} \; (E_{D_h}).$$

Solution: By substituting $D_h(x, a_i)$ in the objective function of (E_{D_h}) we have to solve the convex minimization problem

$$\min\{h(x) - x \sum_{i=1}^n w_i h'(a_i) : x \in \mathbb{R}_+\},$$

since a_i are given numbers and $\sum_{i=1}^n w_i = 1$.

Then $\bar{x}_h := \bar{x}_h(a)$ solves (E_{D_h}) if and only if

$$h'(\bar{x}_h) = \sum_{i=1}^{n} w_i h'(a_i) \ (3)$$

- h' is strictly convex differentiable function $\Rightarrow$ h' is continuous and strictly increasing in $\mathbb{R}_+$

By Theorem 82 [1, p.65] there is a unique $\bar{x}_h$ solving (3) and

$$\min\, a_i \leq \bar{x}_h \leq \max\, a_i.$$

Thus

$$\bar{x}_h(a) = (h')^{-1}\{\sum w_i h'(a_i)\}.$$

[1] G.H. Hardy, J.E Littlewood, and G. Polya, "Inequalities," 2nd edition, Cambridge Univ. Press, London, 1959.

Call $\psi := h'$. Then ψ is strictly monotone and

$$\bar{x}_\psi(a) = \psi^{-1}\{\sum w_i \psi(a_i)\} \text{ (HLP)}$$

Remarks:

- $\bar{x}_\psi$ is not necessarily homogeneous (while $\bar{x}_\phi$ was)
- the only homogeneous means $\bar{x}_\psi$ are the means $\bar{x}_p$ obtained with $\psi(t) = t^p$ (see Theorem 84 [1, p.68])
- $\psi(t) = t^p$: special case of entropic means

[1] G.H. Hardy, J.E Littlewood, and G. Polya, "Inequalities," 2nd edition, Cambridge Univ. Press, London, 1959

Generalization to the case of random variables can be obtained by solving

$$\min\{\int_\alpha^\beta D_h(x,t)dF(t) : x \in \mathbb{R}_+\},$$

where

$$D_h(x,t) := h(x) - h(t) - (x-t)h'(t).$$

Equivalently,

$$\min\{h(x) - x\int_\alpha^\beta h'(t)dF(t) : x \in \mathbb{R}_+\}.$$

The solution is

$$\bar{x}_h(A) = (h')^{-1}\{E[h'(A)]\}.$$

Note: $\text{supp}A = [\alpha, \beta]$

Call $\psi := h'$ then

$$\bar{x}_\psi(A) = (\psi)^{-1}\{E[\psi(A)]\} = \psi^{-1}\{\int_\alpha^\beta \psi(t)dF(t)\}.$$

Example

Let A be a r.v. uniformly distributed to the interval $[a, b]$, i.e, with density

$$f(t) = \begin{cases} \frac{1}{b-a} & , a \leq t \leq b \\ 0 & , \text{else} \end{cases}$$

Then

$$\bar{x}_\psi(A) = \bar{x}_\psi(a, b) = \psi^{-1}\{\int_\alpha^\beta \psi(t) dF(t) = \psi^{-1}\{\int_a^b \frac{\psi(t)}{b-a} dt\}.$$

Note: $dF(t) = f(t)dt = \frac{1}{b-a} dt$ when $a \leq t \leq b$.

Example (Cont.)

Choose now $\psi(t) = \frac{1}{t}$. Then $\int_a^b \frac{\psi(t)}{b-a} dt = \int_a^b \frac{1}{t(b-a)} dt = \frac{\ln(\frac{b}{a})}{b-a}$.

The solution will be

$$\bar{x}_\psi(A) = \bar{x}_\psi(a,b) = \psi^{-1}\{\int_a^b \frac{\psi(t)}{b-a} dt\} = L(a,b).$$

More generally, with $\psi(t) = t^{p-1}$, $(p \neq 1)$, the Stolarsky's power mean is recovered

$$\bar{x}_\psi(A) = S_p(a,b) = \{\frac{b^p - a^p}{p(b-a)}\}^{\frac{1}{(p-1)}} \ ,$$

since $\int_\alpha^\beta \psi(t) dF(t) = \int_a^b \frac{t^{p-1}}{b-a} = \frac{b^p - a^p}{p(b-a)}$ and $\psi^{-1}(t) = t^{\frac{1}{(p-1)}}$.

In [2] Hoehn and Niven discovered that the mean of order p

$$x_p(a_1, a_2, ..., a_n) = \left(\sum_{i=1}^n w_i a_i^p\right)^{\frac{1}{p}}$$

has the asymptotic property

$$x_p(a_1 + \xi, a_2 + \xi, ..., a_n + \xi) - \xi \to \sum_{i=1}^n w_i a_i, \quad (4)$$

$\xi \to +\infty$.

Remarks:

- (4) is proved for some specials values of p
- Extended to unweighted power means of all orders by [3] , and for non-homogeneous means by [4]
- Extension to a great variety of homogeneous means, including weighted means is given in [5]

[2] L. Hoehn and I. Niven, Averages on the move, Math. Mag. 58 (1985), 151-156

[3] J. L. Brenner, Limits of means for large values of the variables, Pi Mu Epsilon J. 8(1985), 160-163.

[4] R. P. Boas and J. L. Brenner, Asymptotic behavior of inhomogeneous means, J. Math. Anal. Appl. 123 (1987), 262-264.

[5] J. L. Brenner and B. C. Carlson, Homogeneous mean values: Weightsand asymptotics, J. Math. Anal. Appl. 123 (1987), 265-280.

Theorem 5

Next Theorem shows that the entropic means have the Hoehn-Niven property. This result is direct application of [5, Theorem 1].

Let $\phi \in \Phi$, $\phi \in C^3$ in the neighborhood of $t = 1$. If $a_1, a_2, ..., a_n$ are fixed and $\xi \to \infty$ then

$$x_\phi(a_1 + \xi, ..., a_n + \xi) = \xi + \sum_{i=1}^{n} w_i a_i + O(\tfrac{1}{\xi}).$$

[5] J. L. Brenner and B. C. Carlson, Homogeneous mean values: Weightsand asymptotics, J. Math. Anal. Appl. 123 (1987), 265-280.

Theorem 3 Proof

Proof

By Theorem 1, $x_\phi(a)$ is a homogeneous mean satisfying $x_\phi(1) := x_\phi(1,1,...,1) = 1$.

Next we can see that x_ϕ is a weighted mean in the sense of Brenner and Carlson [5] , i.e.,

$$\frac{\partial x_\phi}{\partial a_j}(1,1,...,1) = w_j \ , \ j = 1,2,...,n.$$

[5] J. L. Brenner and B. C. Carlson, Homogeneous mean values: Weightsand asymptotics, J. Math. Anal. Appl. 123 (1987), 265-280.

Theorem 3 Proof (Cont.)

Indeed the entropic mean $x_\phi(a)$ is the optimal solution of the problem

$$\min\{\sum_{i=1}^{n} w_i a_i \phi(\tfrac{x}{a_i}) : x \in \mathbb{R}_+\} \quad (E_{d_\phi})$$

Thus it satisfies the optimality condition (OC)

$$\sum_{i=1}^{n} w_i \phi'(\tfrac{x}{a_i}) = 0.$$

Differentiating the identity with respect to a_j we obtain

$$\frac{\partial x_\phi(a)}{\partial a_j} \sum_{i=1}^{n} \frac{w_i}{a_i} \phi''\left(\frac{x_\phi(a)}{a_i}\right) = \frac{w_j}{a_j^2} \phi''\left(\frac{x_\phi(a)}{a_j}\right) x_\phi(a)\ , \ j = 1, 2, ..., n. \ (5)$$

Theorem 3 Proof (Cont.)

Setting $a_i = 1$ for all $i = 1, 2, ..., n$ and using $x_\phi(1) = 1$, $\phi''(1) > 0$ and $\sum_{i=1}^n w_i = 1$, it follows from (5) that

$$\frac{\partial x_\phi}{\partial a_j}(1, 1, ..., 1) = w_j \ , \ j = 1, 2, ..., n.$$

Also, the differentiability assumption of ϕ implies that $x_\phi(\cdot)$ is twice continuously differentiable in the neighborhood of $(1, 1, ..., 1)$.

Thus, invoking [5, Theorem 1], the asymptotic result

$$x_\phi(a_1 + \xi, ..., a_n + \xi) = \xi + \sum_{i=1}^n w_i a_i + O(\tfrac{1}{\xi})$$

follows.

[5] J. L. Brenner and B. C. Carlson, Homogeneous mean values: Weightsand asymptotics, J. Math. Anal. Appl. 123 (1987), 265-280.

Summary

- Generate means as optimal solutions of minimization problem E, where the distance function is the Entropy-like function and the resulting mean is called Entropic mean.
- Entropic mean satisfies the basic properties of a general mean (see proof of Theorem 1).
- All classical means as well as many others are special cases of entropic means.
- Comparison Thm$_s$ used to derive inequalities between various means.
- Derive entropic mean for random variables. Also, show how classical "measures of centrality" (Expectation, Quantiles,etc) are special cases of Entropic means.
- Use new entropy-like function to derive the generalized mean of HLP.
- Entopic means are weighted homogeneous means and have an interesting asymptotic property.

QUESTIONS?