A General Definition of Means and Corresponding Inequalities

Pranjal Jain
Indian Institute of Science Education and Research, Pune

pranjal.jain@students.iiserpune.ac.in

Abstract

This paper proves inequalities among generalised f-means and provides formal conditions which a function of several inputs must satisfy in order to be a 'meaningful' mean. The inequalities we prove are generalisations of classical inequalities including the Jensen inequality and the inequality among the Quadratic and Pythagorean means. We also show that it is possible to have meaningful means which do not fall into the general category of f-means.

1 Introduction

Definition 1. Given an invertible and continuous real function f, define the f-mean (denoted by $F(x_1\,,\, x_2\,,\, \ldots\,,\, x_n)$) of n numbers $x_1\,,\, x_2\,,\, \ldots\,,\, x_n$ in the domain of f as follows

$$F(x_1\,,\, x_2\,,\, \ldots\,,\, x_n) = f^{-1}\left(\frac{f(x_1) + f(x_2) + \ldots + f(x_n)}{n}\right)$$

If the choice of f is arbitrary, we will call this a 'functional-mean'.

This definition was first given by Kolmogorov Andrey[1], and we see that it generalises the classical means as follows – the Quadratic Mean (QM) is the special case $f(t) = t^2$, the Arithmetic Mean (AM) is the special case $f(t) = t$, the Geometric Mean (GM) is the special case $f(t) = \ln(t)$ and the Harmonic

Mean (HM) is the special case $f(t) = \frac{1}{t}$.

Definition 2 (Weighted f-mean)**.** Similar to the f-mean, we will define the 'weighted f-mean' (where $f : \mathbb{R} \to \mathbb{R}$ is an invertible function) of n numbers $x_1\,, x_2\,, \ldots\,, x_n \in \mathbb{R}$, given n positive real numbers $a_1\,, a_2\,, \ldots\,, a_n$ such that $a_1 + a_2 + \ldots + a_n = 1$ (denoted by $F_0(x_1\,, x_2\,, \ldots\,, x_n)$) as follows

$$F_0(x_1\,, x_2\,, \ldots\,, x_n) = f^{-1}\left[a_1\, f(x_1) + a_2\, f(x_2) + \ldots + a_n\, f(x_n)\right]$$

Property 1. *The $(pf + q)$-mean is the same as the f-mean, where p and q are some real numbers ($p \neq 0$). This is also true for the weighted f-mean. We may re-frame this by saying that under the transformation T, defined as $T\{f\} = pf + q$, the f-mean is an invariant.*

Furthermore, $\frac{f''}{f'}$ is an important entity in *Theorems 1* to *5* as we will see later, and it too is an invariant under T. This is no coincidence.

Remark. In the context of working with f-means, this property allows us to assume WLG (without loss of generality) that given a monotonic (alternatively, invertible) differentiable function $f : \mathbb{R} \to \mathbb{R}$, $f'(t) \geq 0\ \forall\, t \in \mathbb{R}$. This assumption will be used extensively in this paper.

In *Theorems 1* and *3*, we prove the Jensen inequality[2] (JI for short) for two variables for the non-weighted and weighted cases respectively. These theorems have been proved using a method similar to the proofs of *Theorems 2* and *4* in order to build intuition and generality gradually for the reader.

Theorem 2 shows that given a pair of invertible and twice differentiable functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ with $\frac{f''}{f'} > \frac{g''}{g'}$, the f-mean is greater than or equal to the g-mean for an input of two variables. *Theorem 4* generalises this to the weighted version.

Theorem 5 (from §*3*) generalises *Theorem 4* to an arbitrary number of inputs via induction, and we get the final generalisation of the JI stating that given a pair of invertible and twice differentiable functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ with $\frac{f''}{f'} > \frac{g''}{g'}$, we have

$$f^{-1}\left[a_1\, f(x_1) + a_2\, f(x_2) + \ldots + a_n\, f(x_n)\right] \geq$$

$$g^{-1}\left[a_1\, g(x_1) + a_2\, g(x_2) + \ldots + a_n\, g(x_n)\right]$$

Where $a_1, a_2, \ldots, a_n$ are positive weights adding up to 1. Comparing this to the JI, we see that the JI is the special case in which g is the identity function.
This level of generalisation also brings about the question of exactly what properties we expect of a 'mean' if the definition of the same can be made so general; an attempt is made in §*4* to formalise the intuitive notion of what makes a given function of several inputs a 'meaningful' mean using a check-list of 4 conditions which are as follows. Here, $M(S)$ denotes the mean (under some definition of the concept) of the elements of a finite subset S of $\mathbb{R}$. Note that by definition, we are assuming that the order of inputs of M does not matter.

Condition 1. *$M(S)$ varies continuously if any elements of S are varied continuously.*

Condition 2. $M(S \cup \{x_1\}) > M(S \cup \{x_2\})$ *iff* $x_1 > x_2$.

Condition 3. $M(\{x\,,\, x\,,\, \ldots\,,\, x\}) = x \ \forall\, x \in \mathbb{R}$.

Condition 4. *For some positive integer k, let $x\,,\, y_1\,,\, y_2\,,\, \ldots\,,\, y_k \in \mathbb{R}$ be any real numbers. Let S_n be the set containing n copies of x and all of the y_i's. Then, $\lim\limits_{n\to\infty} M(S_n) = x$.*

Remark. Another important condition is that $\min_{x\in S}\{x\} \leq M(S) \leq \max_{x\in S}\{x\}$. However, this is implied from *Condition 2* and *Condition 3*.

We see that functional-means always satisfy this check-list.

Our next goal is to show that it is possible to define meaningful means which are not functional-means under the said formalisation. Special cases of such means are given below, the general versions of which are derived in §*4*.

For a pair of inputs, the mean M defined as follows is a meaningful mean and also not a functional mean.

$$M(x,y) = \frac{F^2\,(x,y)}{G(x,y)} \text{ for } f(t) = t \text{ and } g(t) = t^3$$

For more than 2 inputs, $M(x_1\,,\, x_2\,,\, \ldots\,,\, x_n)$ defined as follows for $x_1\,,\, x_2\,,\, \ldots\,,\, x_n \in \mathbb{R}^+$ is a meaningful mean as well as not a functional mean.

$$M(x_1\,,\, x_2\,,\, \ldots\,,\, x_n) = \frac{F^2(x_1\,,\, x_2\,,\, \ldots\,,\, x_n)}{G(x_1\,,\, x_2\,,\, \ldots\,,\, x_n)} \text{ for } f(t) = t \text{ and } g(t) = t^{\frac{n+1}{n}}$$

2 Intuition and motivation

Consider the comparison between the f-mean with $f(t) = e^t$ (exp-mean for short) and the AM for the inputs $x = 1$ and $y = 1000$. The exp-mean will be

$$\ln(e + e^{1000}) - \ln(2) \approx 999.5$$

However, the AM will be 500.5. We can say that the bigger number will 'get more attention' if the function grows faster. We can see this with the four standard means QM, AM, GM and HM too, but an exaggerated example like exp-mean versus AM makes it clearer. Likewise, the HM gives much more attention to the smaller number (the HM of 1 and 1000 is ~ 2).

The visual intuition of this phenomenon can be obtained by noticing how sharply the graph of e^t curves up in comparison to that of t (which does not curve at all). This intuition works for the JI as well; the generalisation proved in this paper is based on the intuition that the functional-mean corresponding to a 'less curving' function should be lesser than the functional-mean corresponding to a 'more curving' function.

Armed with this visual intuition and the observation that both the f-mean and the quantity $\frac{f''}{f'}$ are invariants under the transform $T\{f\} = pf + q$ (as seen in *Property 1*), it is but a natural guess that perhaps it is $\frac{f''}{f'}$ which controls the value of the f-mean!

On the other hand, the proven result deviates from the above intuition when comparing the QM and the exp-mean with the dataset lying in $(0, 1)$; one would expect that the exp-mean wins the fight, but the math disagrees. We see a verification of the same in the following graph, which plots the difference between the QM and exp-mean for a given pair of inputs.
Link to graph : `https://www.desmos.com/calculator/wdqk3dlwcq`

In §*4* where we formalise the concept of a 'meaningful' mean, the definition for M used to provide an example of a meaningful mean which is not a functional-mean is motivated by the fact that $\frac{\mathrm{GM}^2}{\mathrm{AM}} = \mathrm{HM}$. In this case we combine a pair of functional-means to get another functional-mean, so the natural next question about whether it is possible to do the same combination – while maintaining the meaningfulness of the result as a mean – and get a result which is not a functional mean was asked.

Inspiration from this relation between the Pythagorean means was also taken in noticing that the smaller among the means in the LHS (i.e. GM) was squared in the numerator and the larger (i.e. AM) was in the denominator. The same is done in the definition of M in *Theorems 6* and *7*, since doing it the other way around violates *Condition 2* in the check-list for meaningful means.

3 Generalising the Jensen inequality

Remark. Theorems 1, *2*, and *3* are special cases of *Theorem 4*. The prior have been proved separately in order to build generality and intuition for the methods used gradually, but the reader may skip them without loss of context.

Theorem 1. *Given a monotonic and twice differentiable function $f : \mathbb{R} \to \mathbb{R}$, we have the following inequality.*

$$\frac{f''(t)}{f'(t)} > 0 \;\forall\, t \in \mathbb{R} \implies F(x,y) > \frac{x+y}{2} \;\;\forall\; y > x \in \mathbb{R}$$

Proof. For simplicity, we will write $F(x,y)$ as F henceforth. Also, note that our assumption (using *Property 1*) that $f'(t) > 0 \;\forall\, t \in \mathbb{R}$ implies that $f''(t) > 0 \;\forall\, t \in \mathbb{R}$.

Now, we shall manipulate $f(F)$ in such a way that it becomes straightforward to compare it to $f\left(\frac{x+y}{2}\right)$.

$$\begin{aligned} f(F) = \frac{f(x)+f(y)}{2} &= f(y) - \frac{1}{2}\,[\,f(y) - f(x)] \\ &= f(y) - \frac{1}{2}\int_x^y f'(t)\,dt \\ &= f(y) - \frac{1}{2}\int_x^{\frac{x+y}{2}} f'(t)\,dt - \frac{1}{2}\int_{\frac{x+y}{2}}^y f'(t)\,dt \end{aligned}$$

$$\implies f(F) = f(y) - \frac{1}{2}\int_0^{\frac{y-x}{2}} f'(x+t)\,dt - \frac{1}{2}\int_{\frac{x+y}{2}}^{y} f'(t)\,dt \tag{1}$$

Likewise, we may also manipulate $f\left(\frac{x+y}{2}\right)$ in a similar fashion.

$$\begin{aligned} f\left(\frac{x+y}{2}\right) &= f(y) - \left[f(y) - f\left(\frac{x+y}{2}\right)\right] \\ &= f(y) - \int_{\frac{x+y}{2}}^{y} f'(t)\,dt \\ &= f(y) - \frac{1}{2}\int_{\frac{x+y}{2}}^{y} f'(t)\,dt - \frac{1}{2}\int_{\frac{x+y}{2}}^{y} f'(t)\,dt \end{aligned}$$

$$\implies f\left(\frac{x+y}{2}\right) = f(y) - \frac{1}{2}\int_0^{\frac{y-x}{2}} f'\left(\frac{x+y}{2}+t\right)dt - \frac{1}{2}\int_{\frac{x+y}{2}}^{y} f'(t)\,dt \tag{2}$$

Hence, subtracting *(2)* from *(1)* yields

$$f(F) - f\left(\frac{x+y}{2}\right) = \frac{1}{2}\int_0^{\frac{y-x}{2}} \left[f'\left(\frac{x+y}{2}+t\right) - f'(x+t)\right] dt$$

Since $f''(t) > 0 \;\forall\, t \in \mathbb{R}$, this means that

$$f(F) - f\left(\frac{x+y}{2}\right) > 0$$

Now, since $f'(t) > 0 \;\forall\, t \in \mathbb{R}$, we have

$$F > \frac{x+y}{2} \quad \square$$

Theorem 2. *Given monotonic and twice differentiable functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, we have the following inequality.*

$$\frac{f''(t)}{f'(t)} > \frac{g''(t)}{g'(t)} \;\forall\, t \in \mathbb{R} \implies F(x,y) > G(x,y) \;\;\forall\; y > x \in \mathbb{R}$$

Proof. As before, we will write $F(x,y)$ and $G(x,y)$ as F and G respectively for simplicity.

Now, we shall manipulate $f(F)$ to make it easier to compare to $f(G)$ as follows

$$
\begin{aligned}
f(F) = \frac{f(x)+f(y)}{2} &= f(y) - \frac{1}{2}\,[\,f(y)-f(x)]\\
&= f(y) - \frac{1}{2}\int_x^y f'(t)\,dt\\
\implies f(F) &= f(y) - \frac{1}{2}\int_x^G f'(t)\,dt - \frac{1}{2}\int_G^y f'(t)\,dt
\end{aligned}
$$

Substituting $g(t)-g(x)=u$ into the left integral, we have

$$
f(F) = f(y) - \frac{1}{2}\int_0^{\frac{1}{2}\left(g(y)-g(x)\right)} h\left[\,g^{-1}(u+g(x))\,\right]\,du - \frac{1}{2}\int_G^y f'(t)\,dt \tag{3}
$$

Where $h(t) = \frac{f'(t)}{g'(t)}$.

In a similar fashion, we shall manipulate $f(G)$ as follows

$$
\begin{aligned}
f(G) &= f(y) - [f(y)-f(G)]\\
&= f(y) - \int_G^y f'(t)\,dt\\
\implies f(G) &= f(y) - \frac{1}{2}\int_G^y f'(t)\,dt - \frac{1}{2}\int_G^y f'(t)\,dt
\end{aligned}
$$

Substituting $g(t)-g(G)=u$ into the left integral, we have

$$
f(G) = f(y) - \frac{1}{2}\int_0^{\frac{1}{2}\left(g(y)-g(x)\right)} h\left[g^{-1}\left(u+\frac{g(x)+g(y)}{2}\right)\right]du - \frac{1}{2}\int_G^y f'(t)\,dt \tag{4}
$$

Hence, subtracting *4* from *3* yields

$$f(F)-f(G) = \frac{1}{2}\int_0^{\frac{1}{2}\left(g(y)-g(x)\right)} \left\{ h\left[g^{-1}\left(u + \frac{g(x)+g(y)}{2}\right)\right] - h\left[g^{-1}(u+g(x))\right]\right\} du$$

Given that $\frac{f''(t)}{f'(t)} > \frac{g''(t)}{g'(t)}$ and $f'(t), g'(t) > 0 \; \forall \, t \in \mathbb{R}$, it's easy to see that $h'(t) > 0 \; \forall \, t \in \mathbb{R}$. Hence, we have

$$f(F) - f(G) > 0$$

Now, since $f'(t) > 0 \; \forall \, t \in \mathbb{R}$, we have

$$F > G \quad \square$$

Theorem 3. *Given a monotonic and twice differentiable function* $f : \mathbb{R} \to \mathbb{R}$, *with* $F_0(x, y) = f^{-1}[af(x) + bf(y)]$ *where* a *and* b *are positive real numbers with* $a + b = 1$, *we have the following inequality.*

$$\frac{f''(t)}{f'(t)} > 0 \; \forall \, t \in \mathbb{R} \implies F_0(x,y) > ax + by \;\; \forall \; y > x \in \mathbb{R}$$

Proof. As before, we will write $F_0(x, y)$ as F_0 for simplicity.

In the same vein as the proof for *Theorem 1*, we shall manipulate $f(F_0)$ as follows.

$$f(F_0) = af(x) + bf(y) = f(y) - a[f(y) - f(x)]$$

$$= f(y) - a\int_x^y f'(t)\, dt$$

$$= f(y) - a\int_x^{ax+by} f'(t)\, dt - a\int_{ax+by}^y f'(t)\, dt$$

$$= f(y) - a\int_0^{b(y-x)} f'(x+t)\, dt - a\int_{ax+by}^y f'(t)\, dt$$

$$\implies f(F_0) = f(y) - \int_0^{ab(y-x)} f'\left(x + \frac{t}{a}\right) dt - a\int_{ax+by}^y f'(t)\, dt \qquad (5)$$

Likewise, we may also manipulate $f\left(ax+by\right)$ in a similar fashion.

$$f\left(ax+by\right) = f(y) - \left[f(y) - f\left(ax+by\right)\right]$$

$$= f(y) - \int_{ax+by}^{y} f'(t)\,dt$$

$$= f(y) - b\int_{ax+by}^{y} f'(t)\,dt - a\int_{ax+by}^{y} f'(t)\,dt$$

$$= f(y) - b\int_{0}^{a(y-x)} f'\left(ax+by+t\right)\,dt - a\int_{ax+by}^{y} f'(t)\,dt$$

$$\implies f(ax+by) = f(y) - \int_{0}^{ab(y-x)} f'\left(ax+by+\frac{t}{b}\right)dt \;-\; a\int_{ax+by}^{y} f'(t)\,dt \tag{6}$$

Hence, subtracting *(6)* from *(5)* yields

$$f(F_0) - f\left(ax+by\right) = \int_{0}^{ab(y-x)} \left[f'\left(ax+by+\frac{t}{b}\right) - f'\left(x+\frac{t}{a}\right)\right] dt$$

Since $f''(t) > 0\ \forall\, t \in \mathbb{R}$, this means that

$$f(F_0) - f\left(ax+by\right) > 0$$

Now, since $f'(t) > 0\ \forall\, t \in \mathbb{R}$, we have

$$F_0 > ax+by \quad \square$$

Theorem 4. *Given monotonic and twice differentiable functions $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$, with $F_0(x,y) = f^{-1}[af(x)+bf(y)]$ and $G_0(x,y) = g^{-1}[ag(x)+bg(y)]$ where a and b are positive real numbers with $a+b=1$, we have the following inequality.*

$$\frac{f''(t)}{f'(t)} > \frac{g''(t)}{g'(t)}\ \forall\, t\in\mathbb{R} \implies F_0(x,y) > G_0(x,y)\ \ \forall\ y > x \in \mathbb{R}$$

Proof. As before, we will write $F_0(x,y)$ and $G_0(x,y)$ as F_0 and G_0 respectively.

Now, we shall manipulate $f(F_0)$ to make it easier to compare to $f(G_0)$ as follows

$$f(F_0) = af(x) + bf(y) = f(y) - a\ [\,f(y) - f(x)]$$

$$= f(y) - a\int_x^y f'(t)\,dt$$

$$\implies f(F_0) = f(y) - a\int_x^{G_0} f'(t)\,dt - a\int_{G_0}^y f'(t)\,dt$$

Substituting $a[g(t) - g(x)] = u$ into the left integral, we have

$$f(F_0) = f(y) - \int_0^{ab(g(y)-g(x))} h\left[g^{-1}\left(\frac{u}{a} + g(x)\right)\right]\,du - a\int_{G_0}^y f'(t)\,dt \tag{7}$$

Where $h(t) = \frac{f'(t)}{g'(t)}$.

In a similar fashion, we shall manipulate $f(G_0)$ as follows

$$f(G_0) = f(y) - [f(y) - f(G_0)]$$

$$= f(y) - \int_{G_0}^y f'(t)\,dt$$

$$\implies f(G_0) = f(y) - b\int_{G_0}^y f'(t)\,dt - a\int_{G_0}^y f'(t)\,dt$$

Substituting $b[g(t) - g(G_0)] = u$ into the left integral, we have

$$f(G_0) = f(y) - \int_0^{ab(g(y)-g(x))} h\left[g^{-1}\left(\frac{u}{b} + ag(x) + bg(y)\right)\right]\,du - a\int_{G_0}^y f'(t)\,dt \tag{8}$$

Hence, subtracting *(8)* from *(7)* yields

$$f(F_0)-f(G_0) = \int_0^{ab(g(y)-g(x))} \left\{ h\left[g^{-1}\left(\frac{u}{b}+ag(x)+bg(y)\right)\right] - h\left[g^{-1}\left(\frac{u}{a}+g(x)\right)\right]\right\} du$$

Given that $\frac{f''(t)}{f'(t)} > \frac{g''(t)}{g'(t)}$ and $f'(t)\,, g'(t) > 0 \;\forall\, t \in \mathbb{R}$, it's easy to see that $h'(t) > 0 \;\forall\, t \in \mathbb{R}$. Hence, we have

$$f(F_0) - f(G_0) > 0$$

Now, since $f'(t) > 0 \;\forall\, t \in \mathbb{R}$, we have

$$F_0 > G_0 \quad \square$$

Up until this point, we have proved the final result for the case when the data set has only two entries (x and y). Now, we will show that *Theorem 4* implies the final result, which is as follows.

Theorem 5. *Given n numbers $x_1\,, x_2\,, \ldots\,, x_n \in \mathbb{R}$ and n positive real numbers $a_1\,, a_2\,, \ldots\,, a_n$ (with $a_1 + a_2 + \ldots + a_n = 1$), we have the following*

$$F_0(x_1\,, x_2\,, \ldots\,, x_n) = f^{-1}\left[a_1\, f(x_1) + a_2\, f(x_2) + \ldots + a_n\, f(x_n)\right] \;>$$
$$G_0(x_1\,, x_2\,, \ldots\,, x_n) = g^{-1}\left[a_1\, g(x_1) + a_2\, g(x_2) + \ldots + a_n\, g(x_n)\right]$$

Where $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are invertible and twice differentiable functions, with the condition that

$$\frac{f''(t)}{f'(t)} > \frac{g''(t)}{g'(t)}$$

Proof. We will perform induction on n, with base case $n = 2$.

The claim has been proven for $n = 2$ as *Theorem 4*. Hence, we will now assume that it is true for $n-1$ given some $n \geq 3$ and prove that it's true for n.

Let $\mathbb{F} = f^{-1}\left[\frac{a_1}{1-a_n}\, f(x_1) + \frac{a_2}{1-a_n}\, f(x_2) + \ldots + \frac{a_{n-1}}{1-a_n}\, f(x_n)\right]$, and define $\mathbb{G}$ analogously. Now, we have

$$F_0(x_1\,, x_2\,, \ldots\,, x_n) = f^{-1}\left[\, a_1\, f(x_1) + a_2\, f(x_2) + \ldots + a_n\, f(x_n)\,\right]$$

$$\implies F_0(x_1\,,\,x_2\,,\,\dots\,,\,x_n) = f^{-1}\left[\,(1-a_n)\;f(\mathbb{F}) + a_n\;f(x_n)\,\right] \tag{9}$$

Likewise we have a similar result for the weighted g-mean, which is as follows

$$G_0(x_1\,,\,x_2\,,\,\dots\,,\,x_n) = g^{-1}\left[\,(1-a_n)\;g(\mathbb{G}) + a_n\;g(x_n)\,\right] \tag{10}$$

We have $\mathbb{F} > \mathbb{G}$ using the induction hypothesis. Hence, we have

$$f^{-1}\left[\,(1-a_n)\;f(\mathbb{F}) + a_n\;f(x_n)\,\right] > f^{-1}\left[\,(1-a_n)\;f(\mathbb{G}) + a_n\;f(x_n)\,\right]$$

Theorem 4 yields

$$f^{-1}\left[\,(1-a_n)\;f(\mathbb{G}) + a_n\;f(x_n)\,\right] > g^{-1}\left[\,(1-a_n)\;g(\mathbb{G}) + a_n\;g(x_n)\,\right]$$

Hence, combining these two inequalities yields the desired result. $\square$

4 What exactly is a 'mean'?

In this section, we explore the following question : Is it possible to define a 'meaningful' mean which is not just a special case of the f-mean for some choice of f? The answer is affirmative if we rigorously define what we mean by 'meaningful' using the check-list discussed previously; the same has been reproduced here for convenience ($M(S)$ denotes the mean (under some definition of the concept) of the elements of a finite subset S of $\mathbb{R}$. Note that by definition, we are assuming that the order of inputs of M does not matter).

Condition 1. *$M(S)$ varies continuously if any elements of S are varied continuously.*

Condition 2. *$M(S \cup \{x_1\}) > M(S \cup \{x_2\})$ iff $x_1 > x_2$.*

Condition 3. *$M(\{x\,,\,x\,,\,\dots\,,\,x\}) = x \;\; \forall\; x \in \mathbb{R}$.*

Condition 4. *For some positive integer k, let $x\,,\,y_1\,,\,y_2\,,\,\dots\,,\,y_k \in \mathbb{R}$ be any real numbers. Let S_n be the set containing n copies of x and all of the y_i's. Then, $\lim\limits_{n\to\infty} M(S_n) = x$.*

We can see that given any monotonic and continuous function $f : \mathbb{R} \to \mathbb{R}$, the f-mean satisfies all of these conditions; the continuity of f guarantees *Conditions 1* and *4*, the monotonic nature of f guarantees *Condition 2*, and *Condition 3* is obvious.
We will now show that

(a) it is possible to define a mean on 2 variables satisfying all the above conditions (with *Condition 4* not being applicable and *Condition 2* being true iff $|S| < 2$), while also not being a functional-mean.

(b) the proposed solution for *(a)* generalises to an arbitrary number of variables if all inputs are positive under some additional conditions.

Consider a pair of monotonic and differentiable functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, and let $x_1 , x_2 , \ldots , x_n \in \mathbb{R}$ be arbitrary numbers. Define M as follows.

$$M(x_1 , x_2 , \ldots , x_n) = \frac{F^2(x_1 , x_2 , \ldots , x_n)}{G(x_1 , x_2 , \ldots , x_n)}$$

Note that regardless of the choice of f and g, *Conditions 1*, *3* and *4* hold under this definition. Hence for any given f and g, showing that *Condition 2* holds for the corresponding M suffices to show that it is a meaningful mean.

Let F' denote the partial derivative of F with respect to x_1, and likewise for G' and M'. We want $M' \geq 0$ always. Note that the choice of x_1 comes without loss of generality, since switching the values of any inputs does not matter.

We will show that if $f(t) = \text{sgn}(t)\,|t|^p$ and $g(t) = \text{sgn}(t)\,|t|^q$ (where $\text{sgn}(t) = 1$ iff $t \geq 0$, and $\text{sgn}(t) = -1$ otherwise) for any real p and q with $0 < p < q$, then M as defined is a meaningful mean as well as not a functional-mean if $n = 2$. We will also show that under certain conditions on p and q, this can also be generalised to a mean of $n > 2$ positive real numbers $x_1 , x_2 , \ldots , x_n$.

This choice of f and g is motivated by the fact that F and G must share all roots in order for *Conditions 1* and *2* to hold. Clearly if we allow $n > 2$ and let signs of $x_1 , x_2 , \ldots , x_n$ be arbitrary simultaneously, F and G would not share all roots and hence M would not be a meaningful mean.

The discussion henceforth will be that of a case when $F \neq 0 \neq G$. Using the quotient rule, we get the following expression for M'.

$$M' = \frac{2FF'}{G} - \frac{F^2 G'}{G^2}$$

Condition 2 yields that $M' \geq 0$ for all inputs, so the following must be true for all inputs.

$$2\frac{FF'}{G} \geq \frac{F^2}{G^2} G'$$

$$\implies 2F' \geq \frac{F}{G} G' \left(\text{‘cancelling’ } \frac{F}{G} \text{ from both sides}\right) \tag{11}$$

Note that the cancellation of $\frac{F}{G}$ from both sides is allowed since F and G are of the same sign in all cases that we are concerned about (namely the case when $n = 2$ and F and G share all roots, and that when $x_1, x_2, \ldots, x_n > 0$).

Theorem 6. *(11) Holds for $f(t) = sgn(t)\,|t|^p$, $g(t) = sgn(t)\,|t|^q$ and $n = 2$ for all real p and q such that $q > p > 0$.*

Proof. We will use F as short for $F(x_1\,,\,x_2)$, and likewise for F', G and G'. Also, let f be short for $f(x_1)$, and likewise for f', g and g'. Let $x = x_1$ and $y = x_2$. For now, we will deal with the case $x_1 \neq -x_2$ (and so $F \neq 0 \neq G$) unless specified otherwise. We see that

$$F' = \frac{f'}{n\,f'(F)} \tag{12}$$

Likewise,

$$G' = \frac{g'}{n\,g'(G)} \tag{13}$$

We also have

$$f' = p\,|x|^{\,p-1} \tag{14}$$

Likewise,

$$g' = q\,|x|^{\,q-1} \tag{15}$$

Hence, combining *(11)* to *(15)* yields that we wish to prove the following (note that $f', g', F', G' > 0$).

$$
\begin{aligned}
2\,\frac{F'}{G'} &\geq \frac{F}{G} \\
\implies 2\,\frac{f'}{g'} &\geq \frac{f'(F)\,F}{g'(G)\,G} \\
\implies 2\,|x|^{\,p-q} &\geq \frac{|F|^{\,p}}{|G|^{\,q}} \\
\implies 2\,|G|^{\,q} &\geq |x|^{\,q-p}\,|F|^{\,p} \qquad (16) \\
\implies 2\,|\,|x|^{\,q}\,sgn(x) + |y|^{\,q}\,sgn(y)\,| &\geq |x|^{\,q-p}\,|\,|x|^{\,p}\,sgn(x) + |y|^{\,p}\,sgn(y)\,|
\end{aligned}
$$

Since at least one of x and y is non-zero, we may assume that $x \neq 0$. Let $X = \frac{y}{x}$. Hence, wish to show that

$$
2|\,1 + |X|^{\,q}\,sgn(X)\,| \geq |\,1 + |X|^{\,p}\,sgn(X)\,|
$$

If $|X| \leq 1$ or $X > 1$, then this inequality is clearly true (note that this also covers the $x = -y$ case). If $X < -1$, then we wish to prove the following.

$$
2\,|X|^{\,q} - |X|^{\,p} \geq 1
$$

Which is clearly true because $|X|^{\,q} > |X|^{\,p} > 1$. Hence, we have proved the desired result. $\square$

Theorem 7. *(11) Holds for $f(t) = t^p$, $g(t) = t^q$ and $x_1\,,\, x_2\,,\,\ldots\,,\, x_n > 0$ for all real p and q such that $q > p > 0$ and $n - 1 \leq \dfrac{p}{q-p}\left(\dfrac{q}{p}\right)^{\frac{p}{q-p}}$.*

Proof. We will use the same notational short-cuts as before. Also, define X_r (for all $r > 0$) as follows.

$$
X_r = \frac{x_2^r + x_3^r + \ldots + x_n^r}{x^{\,r}}
$$

Hence, *(16)* yields that we wish to prove the following.

$$
\begin{aligned}
2\,(1 + X_q) &\geq 1 + X_p \\
\implies X_q + (X_q - X_p) &\geq -1
\end{aligned}
$$

If $X_p \leq X_q$, then this is clearly true. Hence, assume that $X_p > X_q$. We will now show the following.

$$n - 1 \leq \frac{p}{q-p}\left(\frac{q}{p}\right)^{\frac{p}{q-p}} \implies X_q - X_p \geq -1$$

Using *Theorem 5*, we know that the following is always true.

$$\left(\frac{X_q}{n-1}\right)^{\frac{1}{q}} \geq \left(\frac{X_p}{n-1}\right)^{\frac{1}{p}}$$

Hence, we have that

$$X_q \geq X_p^{\frac{q}{p}}(n-1)^{1-\frac{q}{p}}$$

Since we are interested in the quantity $X_q - X_p$, we subtract X_p from both sides to get

$$X_q - X_p \geq X_p\left(X_p^{\frac{q}{p}-1}(n-1)^{1-\frac{q}{p}} - 1\right) = h(X_p) \quad \text{(let)} \tag{17}$$

Differentiating h, we have

$$h'(a) = \left(\frac{q}{p}\right) a^{\frac{q}{p}-1}(n-1)^{1-\frac{q}{p}} - 1$$

It is easy to see that for $a > 0$, $h(a)$ first decreases and then increases. Hence, the positive value of a for which $h'(a) = 0$ is the value for which h has a minima in $\mathbb{R}^+$. Solving for this value of a, we have

$$a = (n-1)\left(\frac{p}{q}\right)^{\frac{p}{q-p}}$$

Hence, the minimum value of h over $\mathbb{R}^+$ is

$$h\left[(n-1)\left(\frac{p}{q}\right)^{\frac{p}{q-p}}\right] = -\frac{q-p}{p}\left(\frac{p}{q}\right)^{\frac{p}{q-p}}(n-1)$$

Hence, *(17)* yields

$$X_q - X_p \geq -\frac{q-p}{p}\left(\frac{p}{q}\right)^{\frac{p}{q-p}}(n-1)$$

Hence, $X_q - X_p \geq -1$ if

$$\frac{q-p}{p}\left(\frac{p}{q}\right)^{\frac{p}{q-p}}(n-1) \leq 1$$

Yielding the desired result. $\square$

Remark. If we let $z = \frac{p}{q-p}$, then the condition on p and q in *Theorem 7* can be stated as $n - 1 \leq \dfrac{(z+1)^z}{z^{z-1}}$. Since z is unbounded and the RHS of the prior inequality grows as $O(z)$, it is guaranteed that for every n, there are values of p and q for which *Theorem 7* holds (one such example of $p = 1$ and $q = \frac{n+1}{n}$ is given in the introduction).

Theorem 8. *For $f(t) = sgn(t)\,|t|^p$ and $g(t) = sgn(t)\,|t|^q$, $M = \frac{F^2}{G}$ is not a functional-mean for all real p and q with $q > p > 0$.*

Proof. For a proof by contradiction, assume that there is some monotonic and continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $M = H$.

Clearly, h must be differentiable since M is differentiable and h is continuous. Also, we may assume that $h'(t) \geq 0\ \forall\, t \in \mathbb{R}$ and $h(0) = 0$ using *Property 2.3.* We will also be using the same notational short-cuts as before.

In the assumption $H = M$, we will plug $x_2 = x_3 = \ldots = x_n = 0$ and let $x_1 = x$ be arbitrary. Hence, we have

$$h^{-1}\left[\frac{h(x)}{n}\right] = n^{\frac{1}{q}-\frac{2}{p}}\, x$$

$$\implies h(x) = n\, h\left(n^{\frac{1}{q}-\frac{2}{p}} x\right) \tag{18}$$

It is intuitively clear that the only solution for this equation which also satisfies $h(0) = 0$ is $h(t) = t^k$ where $k = \frac{pq}{2q-p}$ (note that k satisfies $1 + k\left(\frac{1}{q} - \frac{2}{p}\right) = 0$). To prove this rigorously, we simply divide both sides of *(18)* by x^k, differentiate both sides and plug *(18)* to form a linear differential equation for h. The detailed calculation for this is left out.

Hence, we have shown that $h(t) = t^k$, but it is also easy to see that in this case $H = M$ is false (leading to a contradiction). Hence, we have shown that our initial assumption (that there is some invertible and continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $M = H$) must be false. $\square$

In conclusion,

(a) Theorems 6 and *8* together show that it is possible to define a meaningful mean on 2 variables which is also not a functional-mean.

(b) Theorems 7 and *8* together show that it is possible to define a meaningful mean on an arbitrary (but fixed) number of positive variables which is also not a functional-mean.

Summary. Bringing together the definition of the 'f-mean' provided by Kolmogorov and the Jensen inequality, we first show that a more general version of the Jensen inequality is true; we explore the intuition behind this generalisation and how parts of it seem to arise naturally from basic properties of the f-mean. Next, we discuss how one can formally define the concept of a 'mean', how all f-means are 'meaningful' means and how it is possible to define 'meaningful' means which are not functional means.

Acknowledgements

I would like to thank Dr. Sinha for her extensive guidance and invaluable feedback for improving this article. I would also like to thank my parents who have always supported me in all my endeavours including this one; helping me when they can and providing important life lessons and advice.

References

[1] Kolmogorov Andrey (1930): "On the Notion of Mean", in "Mathematics and Mechanics" (Kluwer 1991), 144 – 146.

[2] "Jensen inequality", Encyclopedia of Mathematics, EMS Press, 2001 [1994]