Introduction: The Theory of Means

We begin by considering vectors of non-negative numbers:

$$\vec{a} = (a_1, \ldots, a_n), \quad a_i \geq 0.$$

Definition. The r-th power mean $M_r(\vec{a})$ is given by $\left( \frac{1}{n} \sum a_i^r \right)^{1/r}$, unless $r = 0$ or $(r < 0$ and some $a_i = 0$, in which case we take $M_r(\vec{a}) = 0$.)

Note: We will eventually define $M_0(\vec{a})$, but for now it's enough to say that it's not obvious.
We have:

\[ M_1(\vec{a}) = U(\vec{a}) = \text{the average of the } a_i; \]

or arithmetic mean

\[ M_{-1}(\vec{a}) = H(\vec{a}) = \frac{1}{\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_n}} = \text{the harmonic mean} \]

\[ G(\vec{a}) = (a_1 \cdot \ldots \cdot a_n)^{1/n} = \text{the geometric mean} \]

Essentially without loss of generality, we can add positive weights \( p_i > 0 \) so

\[ M_r(\vec{a}; \vec{p}) = \left( \frac{\sum p_i a_i^r}{\sum p_i} \right)^{1/r} \]

Since we can scale the weights without changing \( M_r(\vec{a}; \vec{p}) \) we may as well assume \( \sum p_i = 1 \). All our theorems work with weights, assuming
of course that they are the same weights when we compare means with different $r$ or $\hat{a}$.

Proposition. Basic properties.

$$M_r(\hat{a}) = (U(\hat{a}))^{1/r} \quad (\hat{a} = (a_1, \ldots, a_n))$$

$$G(\hat{a}) = e^{U(\log \hat{a})} \quad (\log \hat{a} = (\log a_1, \ldots, \log a_n))$$

$$M_{-r}(\hat{a}) = \frac{1}{M_r(\frac{1}{\hat{a}})} \quad (\frac{1}{\hat{a}} = \ldots \ldots)$$

$$U(\hat{a} + \hat{b}) = U(\hat{a}) + U(\hat{b})$$

$$G(\hat{a} \hat{b}) = G(\hat{a}) G(\hat{b}) \quad (\hat{a} \hat{b} = (a_1 b_1, \ldots, a_n b_n))$$

$$M_r(K \hat{a}) = K M_r(\hat{a})$$

$$G(K \hat{a}) = KG(\hat{a})$$

$$M_r(\hat{a}) \leq M_r(\hat{b}) \quad \text{if} \quad a_i \leq b_i \quad \text{for all} \ i.$$
We now warm up with an easy theorem.

Proposition. Min $a_i \leq Mr(\bar{a}) \leq Max a_i$ with equality when all $a_i$ are equal and strict inequalities unless (all $a_i$ are equal or $Mr(\bar{a})=0$).

Proof. Suppose $\Sigma p_i = 1$. Then

$$U = \Sigma p_i a_i$$

so $\Sigma p_i U = \Sigma p_i a_i$

and thus

$$\Sigma p_i(a_i - U) = 0.$$

Since all the $p_i$ are positive, this means

* $Max a_i - U \geq 0$ and $Min a_i - U \leq 0$.

This proves the theorem for $M_1 = U$.

$$\left( M_r(\bar{a}) \right)^c = U(\bar{a}^c),$$

so $Min a^c \leq M_r(\bar{a}) \leq Max a_i$.
Now if \( r > 0 \), \( x^r \) is an increasing function on \( \mathbb{R}^+ \), so
\[
\text{Min } a_i^r = (\text{Min } a_i)^r \\
\text{Max } a_i^r = (\text{Max } a_i)^r
\]

Further, \( x^{1/r} \) is an increasing function so
\[
(\text{Min } a_i)^r \leq M_r(\bar{a}) \leq (\text{Max } a_i)^r
\]

\( \Rightarrow \) \( \text{Min } a_i \leq M_r(\bar{a}) \leq \text{Max } a_i \).

If \( r < 0 \) (and \( a > 0 \), as the other case is trivial), then \( x^r \) and \( x^{1/r} \) are both decreasing functions, and
\[
\text{Min } a_i^r = (\text{Max } a_i)^r \\
\text{Max } a_i^r = (\text{Min } a_i)^r
\]
and

\[(\text{Max } a_i)^r \leq M_r(\hat{a})^r \leq (\text{Min } a_i)^r\]

\[\Rightarrow \text{Max } a_i \geq M_r(\hat{a}) \geq \text{Min } a_i \]

We now show

Proposition. \[\lim_{r \to 0} M_r(\hat{a}) = G(\hat{a})\]

which justifies our notational convention \[G(\hat{a}) = M_0(\hat{a})\] (otherwise meaningless, as \[M_0(\hat{a}) = (\frac{\sum p a_i}{\sum p})^\infty\] makes no sense!).

Proof. Again, assuming \[\sum p_i = 1\], we have

\[M_r(\hat{a}) = (\sum p_i a_i^r)^{\frac{1}{r}}\]

\[= e^{\log(\sum p_i a_i^r)^{\frac{1}{r}}}\]
\[
\log \left( \sum_{i=1}^{r} \pi_{ia} \right)
= e
\]

Now since we want to take
\[ \lim_{r \to 0} \mu_r(\hat{a}), \] we first compute
\[
\lim_{r \to 0} \frac{\log \left( \sum_{i=1}^{r} \pi_{ia} \right)}{r} = \lim_{r \to 0} \frac{d}{dr} \log \left( \sum_{i=1}^{r} \pi_{ia} \right)
= \lim_{r \to 0} \frac{\sum_{i=1}^{r} \pi_{ia} \log a_i}{\sum_{i=1}^{r} \pi_{ia}} = \sum_{i=1}^{n} \pi_i \log a_i
\]

Now we can easily compute
\[
\lim_{r \to 0} \mu_r(\hat{a}) = \sum_{i=1}^{n} \pi_i e^{\frac{1}{\lambda_i}} a_i^{\lambda_i} = a_1^{\lambda_1} \ldots a_n^{\lambda_n}
\]

Note that this provides the natural definition of \( G(\hat{a}; \hat{p}) \), which we didn't state above.
We now understand \( \lim_{r \to 0} M_r(\hat{a}) \). What about \( \lim_{r \to \infty} M_r(\hat{a}) \)? \( \lim_{r \to -\infty} M_r(\hat{a}) \)?

Proposition. \( \lim_{r \to \infty} M_r(\hat{a}) = \max a_i \) and \( \lim_{r \to -\infty} M_r(\hat{a}) = \min a_i \).

Proof. We already know \( M_r(\hat{a}) \leq \max a_i \). Suppose \( a_k \to_{\infty} = \max a_i \). We know

\[
M_r(\hat{a}) = \left( p_1 a_1^r + \ldots + p_k a_k^r + \ldots + p_n a_n^r \right)^{\frac{1}{r}} \\
\geq (p_k a_k^r)^{\frac{1}{r}} \quad \text{(remember, all} \ p_i > 0, a_i > 0) \\
= p_k^{\frac{1}{r}} a_k.
\]

As \( r \to \infty \), \( p_k^{\frac{1}{r}} a_k \to a_k \), so by the squeeze theorem, \( \lim_{r \to \infty} M_r(\hat{a}) = \max a_i \).
Going the other way, we need only use the fact that

\[ M_r(\hat{a}) = \frac{1}{M_{-r}(\frac{1}{\hat{a}})} . \]

We can now start our quest: we want to compare \( M_r(\hat{a}) \) and \( M_{-r}(\hat{a}) \). We will pick off a couple of special cases, but it seems unreasonable to avoid stating:

**Theorem of the Means.**

If \( r < s \), then \( M_r(\hat{a}) \leq M_s(\hat{a}) \) with equality when all \( a_i \) are equal or \( s \leq 0 \), and some \( a_i = 0 \).
Proposition. (Cauchy’s inequality)
If \( r > 0 \), \( M_r(\hat{a}) \leq M_{2r}(\hat{a}) \) with equality only if all \( a_i \) are equal.

Proof. We are going to show \( M_r(\hat{a})^{2r} \leq M_{2r}(\hat{a})^{2r} \).
Now
\[
M_r(\hat{a})^{2r} = \left( \frac{\sum p_i a_i}{\sum p_i} \right)^{2r} = \left( \frac{\sum p_i a_i}{\sum p_i} \right)^2
\]
\[
M_{2r}(\hat{a})^{2r} = \frac{\sum p_i a_i^{2r}}{\sum p_i}
\]
so we must show
\[
\left( \frac{\sum p_i a_i^{2r}}{\sum p_i} \right)^2 \leq \frac{\sum p_i a_i^{2r}}{\sum p_i} \left( \sum p_i \right) \leq \left( \sum p_i a_i \right)^2 \left( \sum p_i \right)\]
\[
(\sum p_i a_i)^2 \leq (\sum p_i)(\sum p_i a_i^{2r})\]
(\sum u_i v_i)^2 \leq (\sum u_i^2)(\sum v_i^2).

We note that this can be (sneakily!) rewritten as by letting \( u_i = \sqrt{p_i} \), \( v_i = a_i \sqrt{p_i} \) as
This is the form of Cauchy's inequality you probably know, and the proof is

$$\sum u_i^2 \sum v_i^2 - (\sum u_i v_i)^2 = \frac{1}{2} \sum (u_i v_j - u_j v_i)^2$$

Just to convince myself, the rhs expands to

$$\frac{1}{2} \sum_{i,j} u_i^2 v_j^2 - 2 u_i v_i v_i v_j + u_j^2 v_i^2$$

$$= \frac{1}{2} \sum_{i,j} u_i^2 v_j^2 - 2 u_i v_i u_j v_j + u_j^2 v_i^2$$

$$= \left(\frac{1}{2} \cdot 2\right) \sum_{i,j} u_i^2 v_j^2 - \left(\frac{1}{2} \cdot 2\right) \sum_{i,j} u_i v_i u_j v_j$$

$$= \sum_i u_i^2 \sum_j v_j^2 - (\sum_i u_i v_i)^2.$$