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 $\left(\frac{1}{n}\sum_i a_i^r\right)^{1/r}$, 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 <u>arithmetic</u> mean

$$M_{-1}(\vec{a}) = H(\vec{a}) = \frac{1}{\frac{1}{n}\sum \frac{1}{a_i}} = \text{the } \underline{\text{harmonic}} \text{ mean}$$

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

Essentially without loss of generality, we can add positive weights $p_i > 0$ so

$$M_r(\vec{a};\vec{p}) = \left(\frac{\sum p a^r}{\sum p}\right)^{1/r}$$

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

Proposition. Basic properties.

$$M_r(\vec{a}) = \left(U(\vec{a}^r)\right)^{1/r} \quad (\vec{a}^r = (a_1^r, \dots, a_n^r))$$

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

$$M_{-r}(\vec{a}) = \frac{1}{M_r(1/\vec{a})} \quad \left(\frac{1}{\vec{a}} = \dots\right)$$

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

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

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

$$G(K\vec{a}) = K G(\vec{a})$$

$M_r(\vec{a}) \leq M_r(\vec{b})$ if $a_i \leq b_i$ for all i.

We now warm up with an easy theorem.

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

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

$$\mathcal{U} = \sum p_i a_i, \text{ so } \sum p_i \mathcal{U} = \sum p_i a_i$$

and thus

$$\sum p_i (a_i - \mathcal{U}) = 0.$$

Since all the p_i are positive, this means $\text{Max } a_i - \mathcal{U} \geq 0$ and $\text{Min } a_i - \mathcal{U} \leq 0$. This proves the theorem for $M_1 = \mathcal{U}$.

$$(M_r(\vec{a}))^r = \mathcal{U}(\vec{a}^r),$$

so $\text{Min } a_i^r \leq M_r(\vec{a})^r \leq \text{Max } a_i^r$

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

$$(\text{Min } a_i)^r \leq M_r(\vec{a})^r \leq (\text{Max } a_i)^r$$

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

If $r<0$ (and $a>0$, as the other case is trivial), then x^r and $x^{1/r}$ are both <u>decreasing</u> functions, and

$$\text{Min } a_i^r = (\text{Max } a_i)^r$$
$$\text{Max } a_i^r = (\text{Min } a_i)^r$$

and

$$(\mathrm{Max}\ a_i)^r \leq M_r(\vec{a})^r \leq (\mathrm{Min}\ a_i)^r$$

$$\Rightarrow \quad \mathrm{Max}\ a_i \geq M_r(\vec{a}) \geq \mathrm{Min}\ a_i \qquad \square$$

We now show

Proposition. $\lim_{r \to 0} M_r(\vec{a}) = G(\vec{a})$

which justifies our notational convention $G(\vec{a}) = M_0(\vec{a})$ (otherwise meaningless, as "$M_0(\vec{a}) = \left(\frac{\Sigma p a_i^0}{\Sigma p}\right)^{\infty}$" makes no sense!).

Proof. Again, assuming $\Sigma p_i = 1$, we have

$$M_r(\vec{a}) = \left(\Sigma p_i a_i^r\right)^{1/r}$$

$$= e^{\log\left(\Sigma p_i a_i^r\right)^{1/r}}$$

$$= e^{\frac{\log(\sum p_i a_i^r)}{r}}$$

Now since we want to take $\lim_{r \to 0} M_r(\vec{a})$, we first compute

$$\lim_{r \to 0} \frac{\log(\sum p_i a_i^r)}{r} = \lim_{r \to 0} \frac{\frac{d}{dr} \log(\sum p_i a_i^r)}{1}$$

$$= \lim_{r \to 0} \frac{\sum p_i a_i^r \log a_i}{\sum p_i a_i^r} = \sum p_i \log a_i$$

Now we can easily compute

$$\lim_{r \to 0} M_r(\vec{a}) = e^{\sum p_i \log a_i} = a_1^{p_1} \cdots a_n^{p_n}. \quad \square$$

Note that this provides the natural definition of $G(\vec{a}; \vec{p})$, which we didn't state above.

We now understand $\lim\limits_{r \to 0} M_r(\vec{a})$.

What about $\lim\limits_{r \to \infty} M_r(\vec{a})$? $\lim\limits_{r \to -\infty} M_r(\vec{a})$?

Proposition. $\lim\limits_{r \to \infty} M_r(\vec{a}) = \text{Max } a_i$ and $\lim\limits_{r \to -\infty} M_r(\vec{a}) = \text{Min } a_i$.

Proof. We already know $M_r(\vec{a}) \leq \text{Max } a_i$.

Suppose $a_k = \text{Max } a_i$. We know

$$M_r(\vec{a}) = \left(p_1 a_1^r + \dots + p_k a_k^r + \dots + p_n a_n^r\right)^{1/r}$$

$$\geq \left(p_k a_k^r\right)^{1/r} \quad \text{(remember, all } p_i > 0,\ a_i \geq 0\text{)}$$

$$= p_k^{1/r} a_k.$$

As $r \to \infty$, $p_k^{1/r} a_k \to a_k$, so by the squeeze theorem, $\lim\limits_{r \to \infty} M_r(\vec{a}) = \text{Max } a_i$.

Going the other way, we need only use the fact that

$$M_r(\vec{a}) = \frac{1}{M_{-r}\left(\frac{1}{\vec{a}}\right)}. \quad \square$$

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

Theorem of the Means.
If $r<s$, then $M_r(\vec{a}) \le M_s(\vec{a})$ with equality when all a_i are equal or ($s \le 0$, and some $a_i = 0$).

Proposition. (Cauchy's inequality)
If $r>0$, $M_r(\vec{a}) \leq M_{2r}(\vec{a})$ with equality only if all a_i are equal.

Proof. We are going to show $M_r(\vec{a})^{2r} \leq M_{2r}(\vec{a})^{2r}$.
Now

$$M_r(\vec{a})^{2r} = \left(\frac{\sum p_i a_i^r}{\sum p_i}\right)^{\frac{2r}{r}} = \frac{\left(\sum p_i a_i^r\right)^2}{\left(\sum p_i\right)^2}$$

$$M_{2r}(\vec{a})^{2r} = \frac{\sum p_i a_i^{2r}}{\sum p_i}$$

so we must show

$$\frac{\left(\sum p_i a_i^r\right)^2}{\left(\sum p_i\right)^2} \leq \frac{\sum p_i a_i^{2r}}{\sum p_i} \iff \left(\sum p_i a_i^r\right)^2 \leq \left(\sum p_i\right)\left(\sum p_i a_i^{2r}\right)$$

We note that this can be (sneakily!) rewritten, ~~as~~ by letting $u_i = \sqrt{p_i}$, $v_i = a_i^r \sqrt{p_i}$ as

$$\left(\sum u_i v_i\right)^2 \leq \left(\sum u_i^2\right)\left(\sum v_i^2\right).$$

This is the form of Cauchy's inequality you probably know, and the proof is

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

Just to convince myself, the rhs expands to

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

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

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

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