Median vs. Mean

Eric Johannesson

October 25, 2021

Let $n > 0$ be a natural number, and consider a finite population $I = \{i_1, ..., i_n\}$ with a distribution of values $w : I \to \mathbb{R}_{\geq 0}$. The values may represent anything, such as the wealth or income of the individuals in question. The **mean value** $\bar{w}$ of the population is defined as

$$\bar{w} = \frac{w(i_1) + ... + w(i_n)}{n}$$

A **median value** of the population is any real number r such that the number of individuals with a value smaller than r is at most $n/2$, and the number of individuals with a value greater than r is at most $n/2$. There is always at least one individual $i \in I$ such that $w(i)$ is a median value. [1] If there is more than one such individual, it is common to define **the median value** of the population as the mean value of all such individuals. [2]

Let $0 \leq a < b \leq 1$, and define two distributions $w_a : I \to \mathbb{R}_{\geq 0}$ and $w_a : I \to \mathbb{R}_{\geq 0}$ such that, for each $i \in I$,

$$w_a(i) = (1-a)w(i) + a\frac{w(i_1) + ... + w(i_n)}{n}$$

and

$$w_b(i) = (1-b)w(i) + b\frac{w(i_1) + ... + w(i_n)}{n}$$

[1] We can assume without loss that $w(i_1) \leq ... \leq w(i_n)$. If $n = 1$, then $w(i_1)$ is clearly a median value. Let $m > 0$ be a natural number. If $n = 2m$, we have

$$w(i_1) \leq ... \leq w(i_m) \leq w(i_{m+1}) \leq ... \leq w(i_{2m})$$

The number of individuals with a value smaller than $w(i_{m+1})$ is at most m, and the number of individuals with a value greater than $w(i_m)$ is also at most m. Since $m \leq n/2$, it follows that both $w(i_m)$ and $w(i_{m+1})$ are median values. If $n = 2m + 1$, we have

$$w(i_1) \leq ... \leq w(i_{m+1}) \leq ... \leq w(i_{2m+1})$$

The number of individuals with a value smaller than $w(i_{m+1})$ is at most m, and the number of individuals with a value greater than $w(i_{m+1})$ is also at most m. Since $m \leq n/2$, it follows that $w(i_{m+1})$ is a median value.

[2] There are never more than two real numbers $p \neq q$ such that, for any individual $i \in I$ with a median value, $w(i) = p$ or $w(i) = q$. To see why, assume towards contradiction that there are $i, j, k \in I$ with median values $w(i) < w(j) < w(k)$. Let a be the number of individuals with values less than $w(i)$, and b the number of individuals with values greater than $w(k)$. It follows that $a + b + 3 = n$. By assumption, we also have $a + 2, b + 2 \leq n/2$, which means that $a + b + 4 \leq n$, yielding a contradiction.

Thus, w_a and w_b can be seen as the results of collecting taxes a and b, respectively, and then redistributing the sum of the collected values equally among the individuals in the population. For each $i \in I$, we have

$$w_a(i) < w_b(i)$$

just in case

$$(1-a)w(i) + a\frac{w(i_1) + ... + w(i_n)}{n} < (1-b)w(i) + b\frac{w(i_1) + ... + w(i_n)}{n}$$

just in case

$$(b-a)w(i) < (b-a)\frac{w(i_1) + ... + w(i_n)}{n}$$

just in case

$$w(i) < \frac{w(i_1) + ... + w(i_n)}{n}$$

Thus,

(1) For any taxes $0 \leq a < b \leq 1$ and individual $i \in I$, we have $w_a(i) < w_b(i)$ just in case $w(i) < \bar{w}$.

Now, assume that the mean value is larger than any median value. It then follows that more than half of the population has a value smaller than the mean value. Hence, by (1), more than half of the population benefits (value-wise) from larger taxes. Vice versa, assuming that more than half of the population benefits from larger taxes, it likewise follows that the mean value is larger than any median value. In conclusion,

(2) A majority benefits from higher taxes just in case the mean value is larger than any median value.

In most countries, the mean income is significantly larger than the median income (and, we can safely assume, significantly larger than *any* median income). A few examples:[3]

Country	Median Income	Mean Income
Sweden	\$17,625	\$20,193
Denmark	\$17,432	\$20,304
Netherlands	\$17,154	\$19,690
Australia	\$17,076	\$21,329

Thus, a clear majority of taxpayers in these countries would benefit if the state increased the income tax and redistributed the money equally among all taxpayers. As far as I am aware, no political party has ever proposed such a thing, at least not explicitly. Implicitly, however, one might say they have. Suppose that everyone currently pays p per unit of their income to the state, not assuming anything about how the state spends the money. Let $\bar{w}$ be the mean income, and let q be real number such that $0 < q \leq 1 - p$. With the

[3]Source: https://worldpopulationreview.com/country-rankings/median-income-by-country. Note that the amounts are in current "international dollars", which is a theoretical dollar often used in country-to-country comparisons.

new proposal, for any income x, one would effectively pay $(p+q)x - q\bar{w}$ in income tax, yielding a new (progressive) tax rate of $p + q(1 - \bar{w}/x)$ per unit. In one form or other, progressive tax rates have been adopted by most democratic countries, as one would expect.

Caveat: perhaps not all voters are included in the income statistics. But including voters who are not (elderly people, essentially) would probably only increase the difference between mean and median income.