The “80-20 rule”

Douglas A. Kerr

Issue 1
January 26, 2017

ABSTRACT

We often hear, for example, that in a certain population, “80% of the wealth is held by 20% of the population”. (Almost always, by the way, with those particular numbers.) We hear the more generic formulation, “In many natural or societal situations, 80% of the overall outcome results from 20% of the causes.”

This “situation” is often called the “80-20 rule”. It is often used to “describe” a certain inequality of income, or wealth holdings, or the like, and is often thought to widely apply to such matters. But there are many misunderstandings about what this means. For example, in the matter of wealth, does this “description” actually completely define a certain “distribution of wealth”? And, if so, is that distribution of wealth “typical” for modern societies? This article looks into many aspects of this matter.

1. DEDICATION AND PROLOGUE

This article is dedicated to the memory of my recently-deceased long-time colleague and good friend, Paul J. P. Gayet, of the Washington D. C. area. Paul was an engineer (mostly of the telecommunication persuasion), scientist, pianist, humanist, philosopher, lover of precise language, and *bona fide* eccentric.

He had a persistent interest in the so-called “80-20 rule”, and over the years we have had episodic discussions of the matter. About two months ago, during one of our sporadic telephone calls to “get caught up”, he raised anew a certain point in that topic. The result is that I decided to refresh, refine, and expand my understanding of the matter. In the course of doing so I developed a number of equations, made a plethora of graphs, and crafted various ways of explaining my outlook on the matter. I sent these, as they arose, to Paul in a series of e-mail messages.

I got back a brief acknowledgment of the first of these missives (sort of, “I’ll look at this when I get a chance”), but nothing beyond. That was of course not too peculiar, as Paul had a wide-ranging attention span. But after a couple of weeks, I decided to call him to make certain he was all right, and to press a little for his reaction to the material I sent.

One of his sons answered the call, and told me that Paul had been stricken by a massive stroke and was in the hospital, in critical condition and with a not-too-optimistic prognosis. I sent my best wishes for recovery, and said I would keep in touch.

About two weeks later I called again, and was advised that Paul had died.

Requiescat in pace, amici mei.

2. INTRODUCTION

2.1 The "80-20 rule"

We often hear, for example, that "In Middleburg, 80% of the overall value of homes comes from 20% of the homes." Or, "In Springvale, 80% of the income earned by the residents is earned by 20% of the residents." Or, "In this software project, 80% of the bugs came from 20% of the code"[1]. Almost always, the description uses those particular numbers.

This situation is often said to be the "80-20 rule". It is often used to "describe" a certain inequality of income, or wealth holdings. It is sometimes said to typically apply to many natural or societal situations.

But there are many ambiguities, and misunderstandings, and downright misstatements in this matter. Let me examine some of these.

2.2 Misunderstandings

2.2.1 *Does the "rule" actually speak precisely of a situation?*

If I go to the mythical town of Middleburg, I can surely identify a set of 20% of all the homes that does not represent 80% of the total home value, seemingly violating the "rule". In fact, to accurately state what is probably meant by most utterances of the "rule", we would need to say (and I will underline the critical added wording):

> In Middleburg, the 20% of the homes <u>that have the highest individual values</u> collectively represent 80% of the total value of all the homes.

Now some critics may say, "Surely, Kerr, it is obvious that this is what is meant."

---

[1] However, this is actually a somewhat different kind of matter, and I won't speak further of that example.

But I always feel most comfortable when “what is meant” is actually said.

2.2.2 *Does this rule actually “describe” a distribution?*

If we again consider the mythical town of Middleburg, where the “80-20 rule” is said to obtain with regard to the distribution of home values, the 20% of the homes with the greatest individual values are collectively responsible for 80% of the value of all the homes. But within those “top 20%” of the homes, their values might be distributed in many ways. In one extreme, they might all have almost the same values. But more likely the “top of the top” might have far greater values than “the rest of the top”.

These two situations of course represent substantially different overall distributions of the home values. And, similarly, there could as well be any number of specific distributions within the “80% of least valuable homes”.

In fact, when we (shortly) will use a plotted curve to completely describe a distribution, such as the distribution of home values in Middleburg, we will find that conformity with the “80-20 rule” merely dictates a single point on the distribution curve. It is in fact merely one “result” of some distribution(s) of interest. And it is not a “result” of many distributions that occur in the real world.

So as not to suggest that the 80-20 “situation” is actually any sort of a “rule”, I will subsequently refer to it as the “80-20 result”.

2.2.3 *Does this “rule” actually usually pertain in typical natural or societal distributions?*

I suggested the answer just above: maybe yes, maybe no. I speak of this a little in section 7. .

2.3 The Lorenz curve

A common way to graphically present a distribution of the kind we are discussing is with a *Lorenz curve*, first used in 1905 by Max O. Lorenz. It is a special case of a *cumulative distribution function* curve. Figure 1 shows how it works.

The value on the F axis refers to a fraction of the population, reckoned from the bottom up in order of individual “contribution”. For example, if the matter being described is the distribution of income (as is often the case), an F value of 0.6 would refer to the 60% of the income-earning population with the lowest individual incomes.

The value on the L(F) axis (the “Lorenz function of F”) represents the total “outcome” for that subset of the population. In the matter of an income distribution, that would be the collective income of that subset, as a fraction of the total income for the entire population.

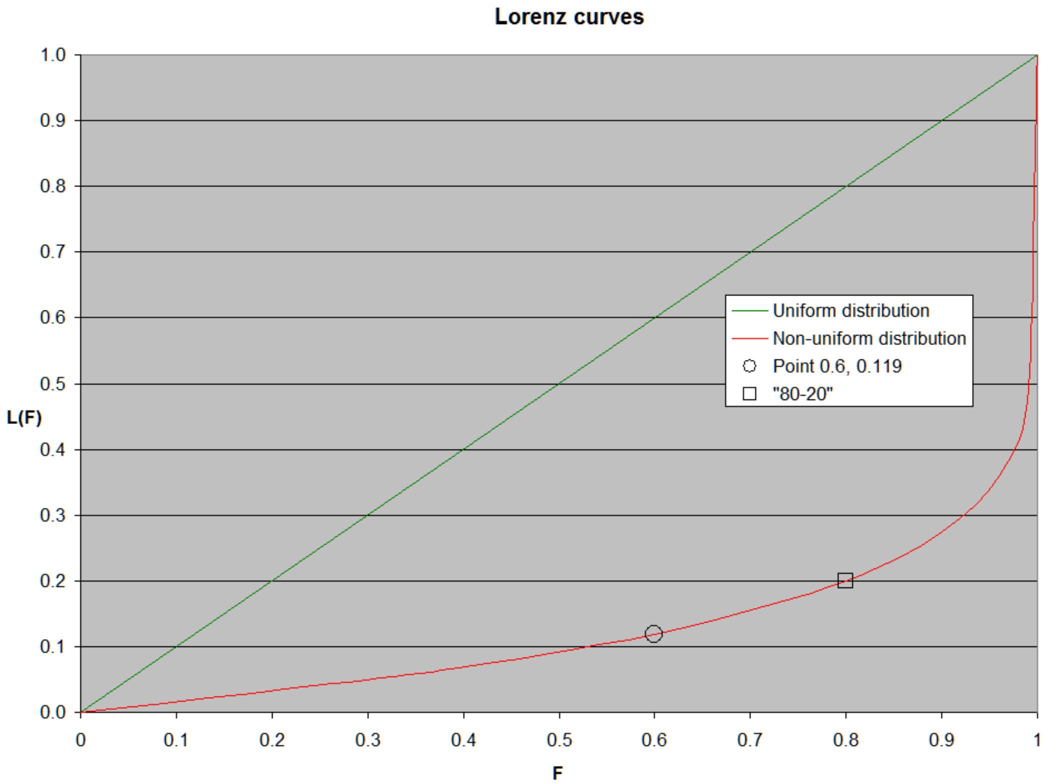

Figure 1. Lorenz curves

Imagine in fact that we are dealing with a distribution of income. The green Lorenz curve represents the "uniform distribution", in which each member of the population has the same individual income. Thus, if that distribution were in effect, inevitably, for example, 70% of the population (we need not here say "with the lowest individual incomes", since all individual incomes are the same) would receive 70% of the total income.

The red curve describes some specific "non-uniform" distribution, one that exhibits the "80-20 result".

To confirm how to interpret the curve, first consider the point on the curve surrounded by the small circle, whose coordinates are 0.60, 0.119 That says, "The 60% of the population with the lowest individual incomes accounts for 11.9% of the total income.

Now, let's consider the "80-20 result", which I said only speaks of one point on the curve that describes a distribution.

Here, the "80-20 result" appears on this distribution curve as the point indicated by the small square. We note that the coordinates of that point are 0.8, 0.2. It is tempting to say, "Oh of course–that is just 80, 20 converted to decimal, just as seen in the name of the 'rule'".

But that is in fact just a coincidence, the result of the fact that the two numbers happen to add to 100%. (More on the significance of that in section 6. .)

Suppose that, to generalize the matter, we speak of such a situation as the “P-Q situation”, so that in our case P = 80 and Q = 20. Because of the way the F axis for the Lorenz curve is defined (indicating the fraction of the population reckoned “from the bottom up” in terms of the amounts of the individual incomes), the coordinates of the point for a “P-Q situation” are as follows:

$$F = 1 - \frac{Q}{100} \tag{1}$$

$$L(F) = 1 - \frac{P}{100} \tag{2}$$

We can easily see that, for the case where $P + Q = 100$ (as we have for the “80-20 result”), F will come out to be P (converted from percentage) and L(F) will come out to be Q (converted from percentage). But that is, for all practical purposes, a coincidence. It would not be so if $P \neq Q$.

To venture now momentarily into the realm of social significance, for an income distribution, a Lorenz curve that falls below the curve for a uniform distribution represents an “income inequality” of this “direction”: “Any ‘top down’ subset of the population receives a total income that is greater than proportional to the size of the subset.”

What if we have a distribution whose Lorenz curve falls “above” the uniform distribution curve? That can’t happen. The convention for the Lorenz curve is that the population must be considered to be arrayed (along the F axis) in order of increasing individual income. Once we have done that, it is inevitable that the curve will lie below the uniform distribution curve (except of course for the case of a uniform distribution, in which case the concept of arraying the population in order of increasing individual income has no meaning).

In any case, remember that a specific distribution is not described by giving a single point on its Lorenz curve (which what we do when we speak of an “80-20 result”). In figure 2 we see the curves for three quite different distributions which nevertheless all exhibit the “80-20 result”.

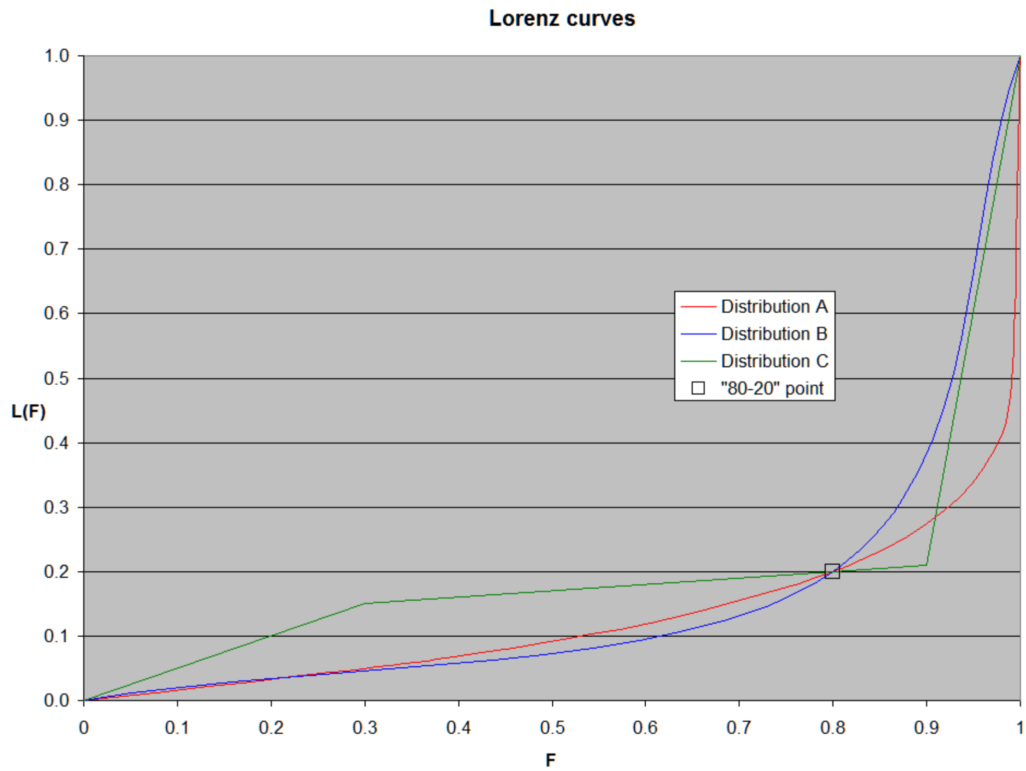

Figure 2. Three “80-20 result” distributions

3. THE PARETO DISTRIBUTION

One type of distribution often found to, at least approximately, describe the distribution of income, wealth, and the like in various societal groups was first characterized by Vilfredo Pareto, an Italian civil engineer, economist, and sociologist, in about 1896. In fact, Pareto characterized various distribution functions, of which the simplest (“type I”) is of most direct interest to us. When I speak of “the Pareto distribution”, I mean that one.

This simplest distribution, which may be expressed in different forms, has only a single parameter, usually represented by α (lower-case Greek *alpha*), which controls the exact “shape” of the distribution.

If we cast the defining equation in terms of a Lorenz curve, we find that for approximately $\alpha = 1.161$, the curve will pass through the point 0.8, 0.2 (and the distribution will thus exhibit the “80-20 result”).

And in fact the red curve in figure 2 is just that Pareto distribution.

Of course, Pareto distributions with other values of α are often of interest. In figure 3, we see the Lorenz curve for the aforementioned Pareto distribution for $\alpha = 1.161$ (which exhibits the “80-20 result”), along with the curves for Pareto distributions with $\alpha = 1.333$, $\alpha = 1.500$, and $\alpha = 2.000$.

As we can see, the greater the value of α, the less “inequality” is represented by the distribution. When $\alpha = \infty$, the curve represents “complete equality”; that is, it is the “uniform distribution” curve we have seen before.

When $\alpha = 0$, the curve represents “complete inequality” (where one person, for example, earns all the income in a population). That curve is flat (L(F) = 0) except for a spike to L(F) = 1 at the right end (where F = 1)

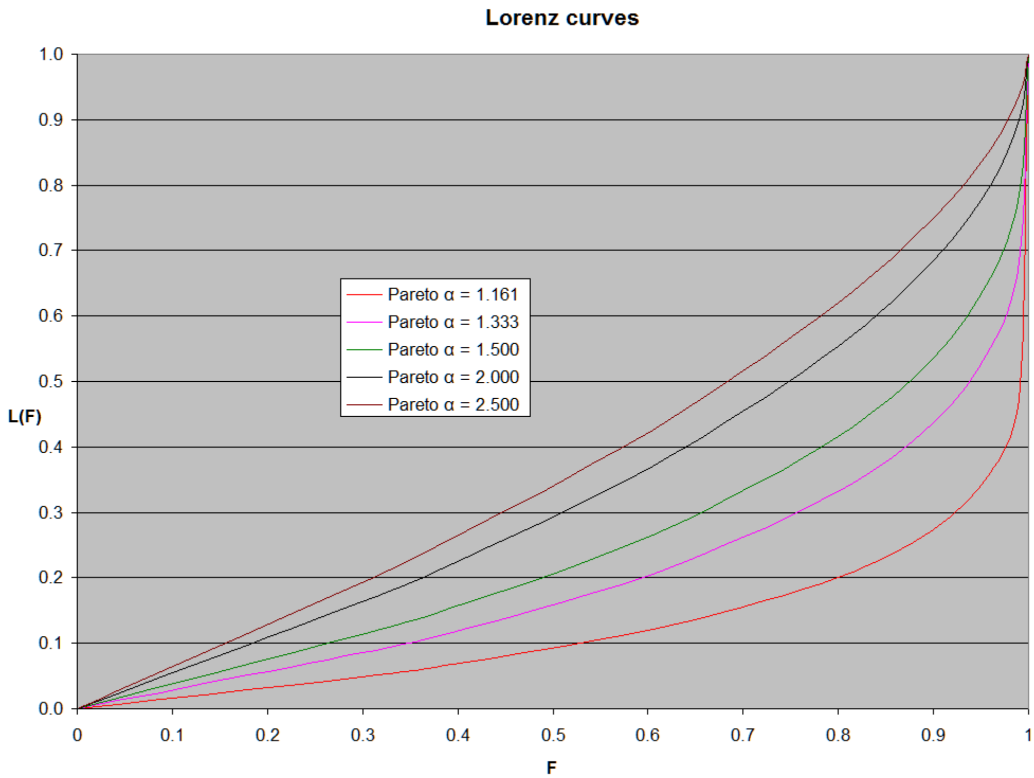

Figure 3. Lorenz curves for several Pareto distributions

4. THE GINI COEFFICIENT

The Gini coefficient is a single-valued metric that indicates, in a certain way, the degree of inequality of a distribution of wealth, income, or such. It was devised by the Italian statistician and sociologist Corrado Gini and published in 1912.

The Gini coefficient is formally defined in terms of the Lorenz curve of the distribution of interest. It is the ratio of:

a. the area between the Lorenz curve for the uniform distribution and the Lorenz curve for the distribution of interest, to

b. the entire area under the uniform distribution curve.

Because the entire area under the uniform distribution curve (b, above) is unavoidably 0.5, we can restate the Gini coefficient as twice the area mentioned as (a) above.

A Gini coefficient of 0 represents no inequality at all (that is, a uniform distribution of wealth, income, or such). A Gini coefficient of 1 (100%) represents the worst possible inequality: where a single person holds all the wealth, earns all the income, or such in the population.

It can be shown that the Gini coefficient, G, for a Pareto distribution with a certain value of α is given by:

$$G = \frac{1}{2\alpha - 1} \tag{3}$$

Thus, for the "80-20" Pareto distribution, for which $\alpha = 1.161$ (approximately), the value of the Gini coefficient is approximately 75.6%.

It can also be shown that for **any** distribution that exhibits the "80-20" result, the Gini coefficient will be at least (approximately) 60%. A Pareto distribution with that Gini coefficient would have α = 1.333 (approximately). We in fact see the Lorenz curve for that Pareto distribution on figure 3. For that distribution of income, the top 20% of the population would earn 66.9% of the total income.

How "severe" is the income inequality indicated by a Gini coefficient of 60%? A chart in the Wikipedia article on the Gini coefficient, showing the Gini coefficient with regard to income for various "major" nations, over the years, shows that only the nation that is for most years "the worst", Brazil, reaches a Gini coefficient of 60%. And the maximum value for Brazil, since 1960, has never exceeded about 62%

Thus, the Gini coefficient for the "80-20" Pareto distribution, approximately 75.6%, represents far greater inequality than the worst attained by any of the world's major nations.

Note however that when the Gini coefficient for the income distribution of some society is cited, that does not imply that the distribution is the corresponding Pareto distribution. The distribution may not be of the Pareto form at all.

As in all matters in this area, it should be recognized that the Gini coefficient can hardly tell the entire story of the distribution of wealth, income, or such for some population.

5. GAYET’S FORMULATION OF AN “80-20” DISTRIBUTION

Paul Gayet had on several occasions expressed the view that the distribution that exhibits the “80-20 rule” (and I must remind us that the “rule”–actually a “result”–does not in fact describe a unique distribution) has these properties, which I will describe in the context of a distribution of wealth; the first one of these properties, included here for continuity, is by definition:

- The top 20% of the population, in decreasing order of individual wealth, collectively holds 80% of the overall wealth.
- Of the remaining population, the top 20% of those holds 80% of the remaining wealth.
- Of the remaining population, the top 20% of those hold 80% of the remaining wealth.
- And so forth.

I describe this formulation of a distribution as a “recursive downward” formulation.

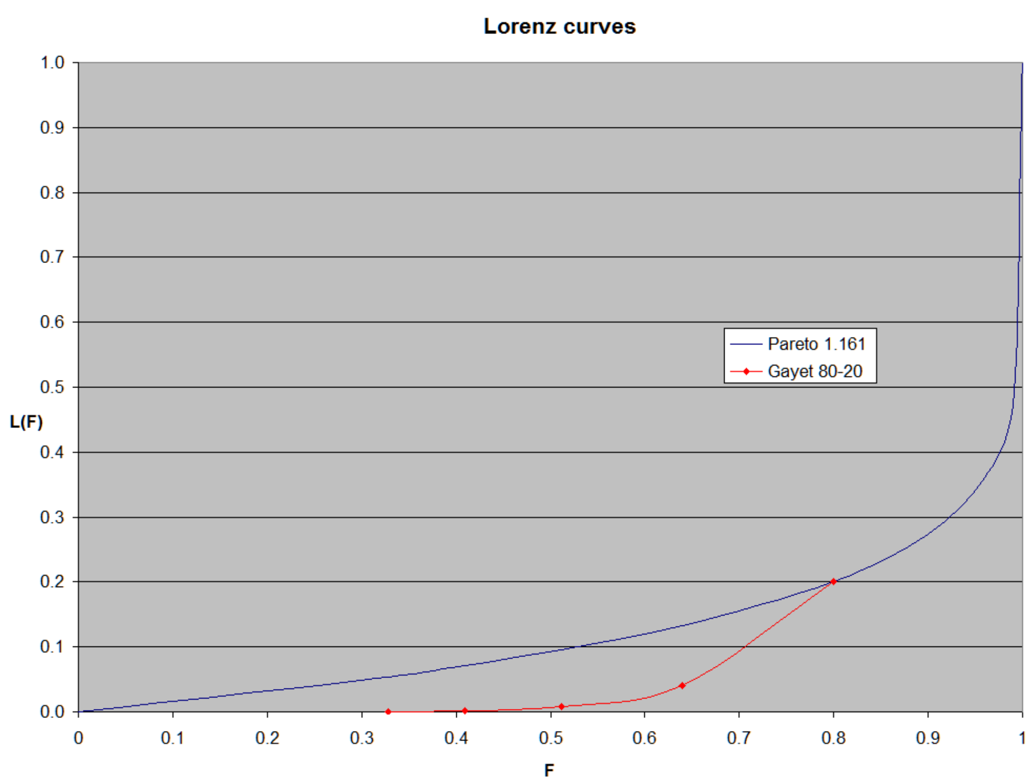

Figure 4. Pareto “80-20” and “Gayet” distributions

Now, does that formulation indeed hold for a distribution exhibiting the “80-20 result”? Well, given that the “80-20 result” does not define a specific distribution, this formulation cannot be determined to “hold” or “not hold” for “that distribution”.

But in fact there is a specific distribution for which the Gayet formulation does hold: the distribution that is defined by it. I refer to that here as the “Gayet distribution”.

Paul Gayet himself took issue with that designation. He said that he had not in any way intended to define a (perhaps previously undiscussed) distribution. Rather, he had only intended to point out an interesting property of the “80-20” distribution.

But, as we have seen above, the “80-20” result does not define a specific distribution, but rather an infinity of distributions, all having the sole common property that their Lorenz curves pass through the point 0.8, 0.2.

So, Paul, like it or not, your “recursive downward” formulation has defined a specific (perhaps previously undiscussed) distribution.

If we consider a widely-cited distribution exhibiting the “80-20 result”, the Pareto distribution with $\alpha = 1.116$, then does the “Gayet formulation” hold for that specific distribution? No. We see that difference in figure 4. (The points shown for the “Gayet” distribution are those developed from Gayet’s “recursive downward” formulation, for the original premise and four “steps” below that.)

Can we work Gayet’s formulation of a distribution backwards to define the distribution for values **above** the “defining” starting point (*e.g.*, “80-20”)? Yes, but the next point we get is at 1,1 on the Lorenz graph—the “end of the road” and the point every distribution must have. (It means that 100% of the income is earned by 100% of the population. It has the P-Q form “0-0”).

However, the portion Gayet’s formulation does define can be characterized exactly by an equation, which (for the “80-20” starting point) is approximately:

$$L(F) = F^{7.21} \qquad (4)$$

and that is defined for the entire range of values of F. So we perhaps reasonably use that equation to extend Gayet’s vision above the starting point. We see the curve of this “extrapolated” Gayet distribution in figure 5.

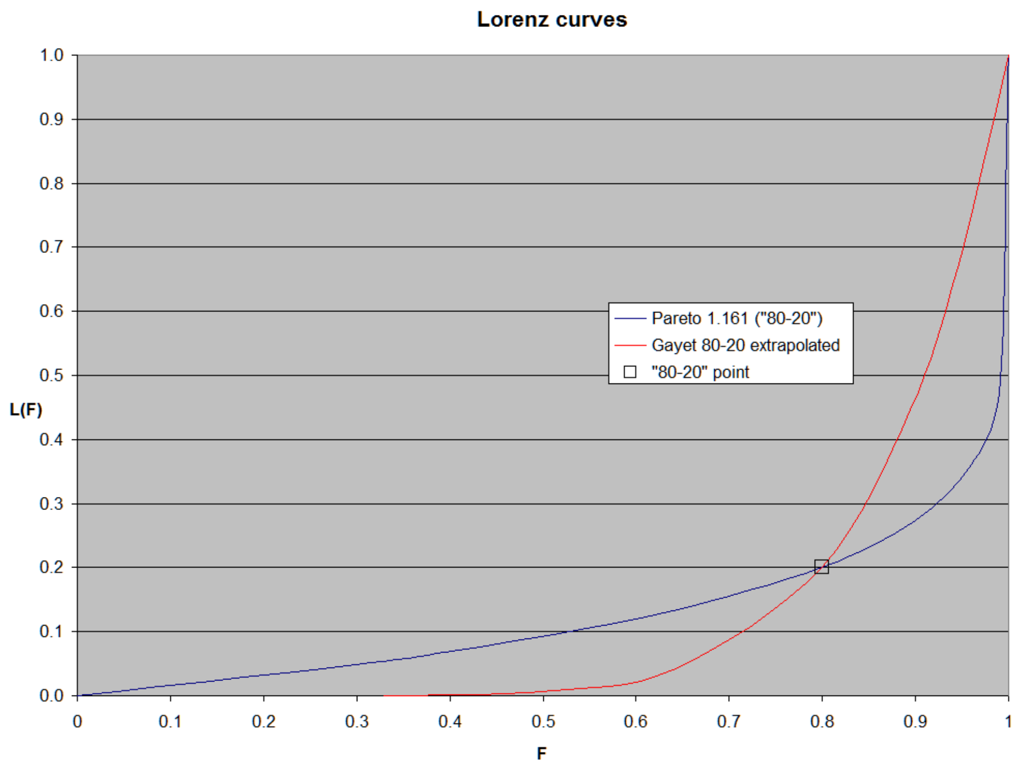

Figure 5. Pareto “80-20” and extrapolated Gayet distributions

But there **is** a “recursive upward” formulation for the part of a Pareto distribution **above** the “reference point”. It works this way for the one seen above, which exhibits the “80-20 result” (again, the first item is “by definition” for that distribution):

- The top 20% of the population, in order of individual wealth, collectively holds 80% of the overall wealth.
- Of that top 20% of the population, the top 20% of that holds 80% of that 80% of the overall wealth).
- Of that group of the population, the top 20% holds 80% of that portion of the overall wealth.
- And so forth.

These relationships match precisely the analytical definition of the distribution. They can perhaps be seen as the “mirror image” of the Gayet formulation.

But there is no corresponding “recursive downward” formulation (such as that described by Gayet) for the Pareto distribution below the “reference point”.

6. THE SIGNIFICANCE OF $P+Q=100$

The most common form of the "P-Q result" (recall that these do not define unique distributions) is the "80-20 result". Others are less-frequently seen, as for example a "70-30 result". Note that for both of these, $P+Q=100$. And when we see others, they too often have that property.

Does that property confer some unique attribute on the distributions exhibiting a given "P-Q result"? No. Then why the preoccupation with P-Q results for which $P+Q=100$?

So far as I can tell, only that these pairs of numbers seem to have a nice symmetrical ring to them. We could equally validly comment that, for some population under the same Pareto distribution we have been considering, 52.75% of the wealth is held by the top 1% of the population, an expression for which $P\neq Q$. And of course that result doesn't define the entire distribution either.

In any case, for any specific distribution (of whatever type), there is always one, and only one, point for which $P+Q=100$.

7. SOME ACTUAL INCOME DISTRIBUTIONS

It can be of interest to take some data from studies done by the World Bank on income distributions in various nations, for recent years, and present it in terms of Lorenz curves.

Before I proceed, let me say that what I am about to discuss is solely for the purpose of illustrating the working of income distribution curves. There are many tricky wrinkles to the collection and analysis of income statistics (for example, how do we reckon the impact of various kinds of taxes, or social benefits). This matter is well beyond the scope of this article, as is the matter of interpretation of the societal implications of various distributions.

That having to be said, we see the Lorenz curves for some national income distributions on figure 6.

The three solid curves represent the income distributions (per the World Bank data) for three nations, Belgium, USA, and Zimbabwe.

For reference, we see also (as dashed lines) the "uniform distribution" curve and the by-now familiar curve for the Pareto distribution with α = 1.161. We also see the curve for the Pareto distribution with α = 2.500, shown because it is broadly comparable to the three "national" distribution curves.

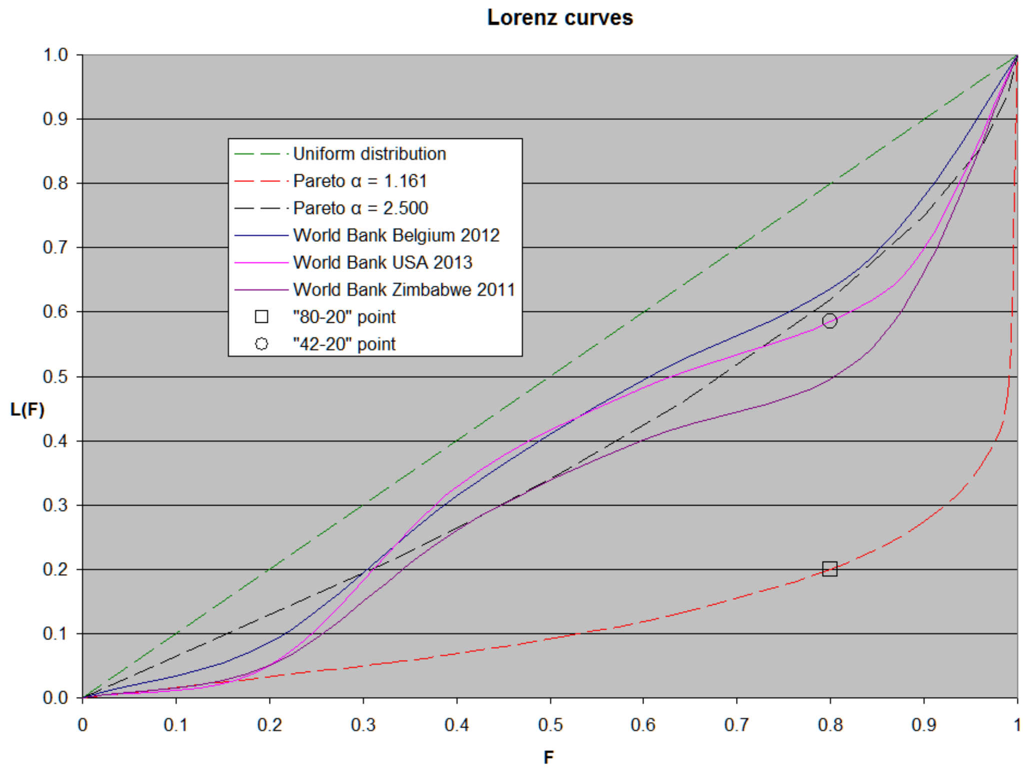

Figure 6. Income distributions

Again for reference we see the infamous "80-20" point, which lies on the Pareto α = 1.161 curve (the value of α for that curve was of course chosen to make that so).

And we see the point "42-20", which lies on the USA income distribution curve. That means, of course, that, under that distribution, 42% of the total income is earned by 20% of the population.

Of course, I have no idea if that is actually so–it is just an implication of the curve derived from the World Bank data.

For an actual Pareto distribution with α = 2.500, the Gini coefficient would be 25%. Is that consistent with the Gini coefficients reported (as part of other analyses) for those nations? Well, not really. One report places the Gini coefficient for the USA, as of 2002 (the last year covered by that report) at about 44%. Going at it from the other direction, if we had that Gini coefficient for a Pareto distribution, its α would have to be about 1.64.

But, as I said, it is not my point here to probe the complications and mysteries of this issue.

8. ANOTHER TYPE OF "RESULT"

Another type of "result" of a distribution is as given in this example: "The top 1% of the population earns as much as the bottom 45% of the population". This is of course different than the "P-Q result" we have been discussing here.

It is not uncommon for a reader to not pay close attention to what is said and to think that what is meant is, "The top 1% of the population earns 45% of the total earnings". That of course is a "P-Q result", namely "45-1".

But, for the situation as I actually stated it first above, if we know that the distribution involved in this kind of result is in fact a Pareto distribution, then from the two numbers involved (1% and 45%) we can analytically determine that the value of α for that distribution is approximately 1.557. The earnings of the top 1%, and the bottom 45%, of the population are then each approximately 19.25% of the total earnings of the population. The calculation is beyond the scope of this article.

9. SUMMARY

The so-called "80-20 rule" describes one "result" of certain distributions, as for example distributions of wealth or income, in which the distribution is substantially "unequal". It can be useful in grasping the impact of the distribution. But this "result" does not define a specific distribution. Many distributions can have the "80-20 result".

We occasionally hear, in this same form, a different "result" of some distribution, such as the "70-30 result". Almost always, these expressions have two numbers whose sum is 100. But there is nothing "magical" about this relationship. It has merely come to be characteristic of a "preferred" property of any given distribution.

It is often suggested that the "80-20 result" itself is typical of actual distributions of income or wealth in various populations. That is not necessarily so.

10. ACKNOWLEDGMENTS

Thanks to my late good friend and colleague Paul J. P. Gayet, one of the last accomplishments of whose very productive life was to direct my attention anew to the fascinating topic of "the 80-20 rule".

Thanks to my wife, Carla Kerr, for her insightful copy editing of this tedious manuscript.

#