2020

36th IARIW General Conference

Paper Prepared for the 36th IARIW General Conference, Oslo, Norway, August 24-28, 2020

Inequality Measurement For Bounded Variables

Suman Seth

Inaki Permanyer

Gaston Yalonetzky

In his seminal contribution, Atkinson (1970) set the foundations of inequality measurement as we know it. After five decades, the contributions to this burgeoning field of research have expanded in multiple directions, and 'inequality' can arguably be considered one of the most hotly debated topics in an increasingly globalised world, as witnessed by the popularity of several recent books on the subject (e.g. Piketty, 2015; Bourguignon, 2017; Atkinson, 2018; Milanovic, 2018). Moreover, the interest in inequality has gone well beyond the study of monetary distributions. Nowadays scholars and policy-makers alike are particularly interested in studying the distribution of the non-pecuniary dimensions of well-being (e.g. health and education outcomes, or the post-2015 Sustainable Development Goals [SDG] agenda).

Bounded variables abound. Unlike monetary variables such as income or consumption expenditure, the majority of non-pecuniary aspects of well-being are gauged by variables that cannot take indefinitely large values. This seemingly unimportant and technical point has key implications as regards the way in which we measure and interpret 'inequality' in the corresponding distributions. The measurement of inequality for bounded variables poses specific challenges that, while acknowledged in the literature, have not been jointly addressed in a satisfactory manner thus far. In this paper, we propose an approach to inequality measurement with bounded variables that addresses several key problems simultaneously.

The first limitation that researchers encounter when studying inequality for bounded variables is the 'consistency problem'. When a variable is bounded it is a priori possible to focus either on the distribution of achievements or the corresponding distribution of shortfalls with respect to the upper bound. For instance, improvements in the coverage of public health plans could be assessed via the percentage of vaccinated children (an achievement indicator) or through the percentage of unvaccinated children (a shortfall indicator). As highlighted by Micklewright and Stewart (1999), Clarke et al. (2002), Erreygers (2009), Lambert and Zheng (2011), Lasso de la Vega and Aristondo (2012) and many others after them, traditional relative inequality measures fail to rank distributions consistently when measurement is switched from attainments to shortfalls and poses several practical challenges to the study of inequality for bounded variables.

The second limitation is the 'boundary' problem (in some contexts also known as 'floor-' and 'ceiling-effect' problems). Whenever a variable is bounded, one can observe a clustering of the distribution as its mean converges towards any of its bounds. In these situations, the corresponding inequality levels mechanically go to zero, simply because there is no room left for further variation. This problem persists even when the consistency problem is solved. As the mean of the distribution increases from the lower bound to the upper bound of the distribution, the level of maximum feasible inequality first increases and then decreases—making the maximum feasible inequality a parabolic function of the mean. In these circumstances, it is not clear whether studying inequality with a bounded variable can provide new insights above and beyond what we already know from studying the values of the mean alone. This point was already highlighted by Wagstaff (2005) in his study of the concentration index and discussed by many others after him (e.g., Erreygers, 2009; Erreygers and Van Ourti, 2011; Wagstaff, 2009, 2011). In this paper we extrapolate some of these ideas to the context of inequality measurement for bounded variables.

Third, it is not uncommon to find situations where the bounds of the variable we are working with differ across the groups one might want to compare. For purely biological reasons, it might well be the case that the upper bounds for certain variables vary across groups (e.g. life expectancy for women and men is a case in point). When this happens, one could argue that the potentially different supports that such variables have for the different groups we are comparing should be allowed to play a non-trivial role when assessing the corresponding inequality levels. In addition, there might be cases in which the variables we are dealing with have uncertain bounds—a common circumstance when studying health related issues. Sometimes, the uncertainty around the bounds of such variables has led researchers to implicitly treat that variable as unbounded. While such assumption is a practical way of circumventing the uncomfortable decision of fixing the bounds, it is not realistic. Instead, we suggest categorising them as members of a new class of variables: the set of uncertainly bounded variables. In this paper, we discuss the methodological and substantive implications that such categorisation has for our understanding of the variability of bounded variables.

To address these issues, in this paper we contribute in several directions. First, we introduce a new class of inequality indices, the so-called class of 'normalised inequality measures.' This class is obtained by combining a seemingly weak normalisation axiom with the consistency requirement and other minimally desirable properties. As a result, we obtain new inequality measures quantifying observed inequality levels as a proportion of the maximal inequality levels that could be attained with the same index evaluated at a hypothetical distribution with the same mean as the observed distribution. The new measures satisfy the basic requirements of inequality measurement and are neither affected by the boundary effects nor by the (in)consistency problems. Second, we develop techniques to assess whether inequality rankings are robust to the choice of alternative upper bounds in contexts where their values (i) differ across the groups we are comparing, and/or (ii) are highly uncertain.

After defining our normalised inequality measures, we illustrate how they perform empirically and compare them vis-à-vis their absolute counterparts. For that purpose, we study the evolution of international inequality in life expectancy across world countries from 1950 to 2050 (from 1950 to 2015 we use observed data and from 2020 to 2050 the medium variant of the United Nations' World Population Prospects data). Our results show that (i) the trends in international health inequality strongly depend on the choice of the upper bound for life expectancy, (ii) our new inequality indices are significantly less sensitive to the values of the mean than currently existing measures.