Exploring the Application of Quantile-Based Methods in Measuring Inequality and Distributional Differences

Hazel Morris, Hudson Kelly, Hunter Barnes

1 Introduction

The measurement of inequality and distributional differences represents a fundamental challenge across numerous disciplines, from economics and sociology to environmental science and public health. Traditional approaches to quantifying inequality have predominantly relied on summary statistics such as the Gini coefficient, Theil index, and various percentile ratios. While these measures provide valuable insights, they suffer from inherent limitations in capturing the full complexity of distributional patterns. The Gini coefficient, for instance, is most sensitive to changes around the mode of the distribution and relatively insensitive to changes in the tails, where much economically and socially relevant inequality manifests. Similarly, percentile ratios like the 90/10 ratio capture only specific points of comparison while ignoring the overall distributional shape.

This paper introduces a comprehensive framework for measuring inequality and distributional differences using quantile-based methods that leverage the complete distributional information. Our approach moves beyond the conventional paradigm of scalar inequality measures toward functional approaches that preserve the richness of distributional characteristics. The fundamental insight underlying our methodology is that quantiles provide a natural and robust basis for distributional comparison that is less sensitive to outliers and more informative about distributional shape than moment-based approaches.

We develop three novel quantile-based measures that address different aspects of distributional analysis. The Quantile Dispersion Index (QDI) measures dispersion across multiple quantiles simultaneously, providing a more comprehensive assessment of spread than single-number summaries. The Distributional Asymmetry Metric (DAM) specifically targets asymmetric distributional changes, which are particularly relevant in contexts where growth or decline affects different parts of the distribution unequally. The Quantile Overlap Coefficient (QOC) quantifies the degree of similarity between distributions, offering a nuanced alternative to traditional hypothesis testing approaches.

Our research contributes to the literature on inequality measurement by addressing several critical gaps. First, we provide measures that are more sensitive to changes in distributional tails, where traditional measures often lack precision. Second, we develop approaches that can detect and quantify asymmetric distributional changes, which are common in real-world contexts but poorly captured by symmetric measures. Third, we offer tools for comparing distributions that go beyond simple hypothesis testing to provide quantitative measures of distributional similarity.

The remainder of this paper is organized as follows. Section 2 develops the theoretical foundation of our quantile-based approach and formally defines our proposed measures. Section 3 describes our methodological framework, including estimation procedures and computational considerations. Section 4 presents simulation studies validating our approach and comparing it to traditional measures. Section 5 applies our methodology to empirical data from economics, environmental science, and public health. Section 6 discusses the implications of our findings and directions for future research.

2 Theoretical Framework

Let $F(x)$ be a cumulative distribution function with quantile function $Q(p) = F^{-1}(p)$ for $p \in [0, 1]$. Traditional inequality measures can often be expressed as functionals of F or Q, but they typically summarize complex distributional information into single scalars. Our approach seeks to preserve more distributional information by working directly with the quantile function or transformations thereof.

We begin by defining the Quantile Dispersion Index (QDI), which measures dispersion across multiple quantiles. For a set of probability levels $0 < p_1 < p_2 < \cdots < p_k < 1$, the QDI is defined as:

$$QDI = \frac{1}{k} \sum_{i=1}^{k} \left| \frac{Q(p_i) - Q(0.5)}{Q(0.5)} \right| \tag{1}$$

This measure captures the average relative deviation of selected quantiles from the median, providing a more comprehensive assessment of spread than measures focusing on specific quantile pairs. The choice of p_i values can be tailored to specific applications, with denser sampling in regions of particular interest.

The Distributional Asymmetry Metric (DAM) addresses the critical issue of asymmetric distributional changes. For two distributions with quantile functions $Q_1(p)$ and $Q_2(p)$, we define:

$$DAM = \int_0^1 \left[\frac{Q_2(p) - Q_1(p)}{Q_1(p)} \right] \cdot w(p)\, dp \tag{2}$$

where $w(p)$ is a weighting function that emphasizes regions of particular interest. A natural choice is $w(p) = p - 0.5$, which gives positive weight to changes above the median and negative weight to changes below the median, thus capturing the net asymmetric effect.

The Quantile Overlap Coefficient (QOC) measures distributional similarity by comparing quantile functions across their entire domain:

$$QOC = 1 - \frac{1}{2}\int_0^1 \left| \frac{Q_1(p) - Q_2(p)}{\max(Q_1(p), Q_2(p))} \right| dp \tag{3}$$

The QOC ranges from 0 (complete separation) to 1 (identical distributions), providing an intuitive measure of distributional overlap that is more informative than binary hypothesis test outcomes.

These measures possess several desirable properties. They are scale-invariant, making them suitable for comparing distributions measured in different units. They are robust to outliers, as quantiles are less sensitive to extreme values than moment-based measures. They provide more distributional information than traditional scalar measures while remaining computationally tractable.

3 Methodology

Our methodological approach involves both theoretical development and practical implementation considerations. For empirical applications, we employ non-parametric estimation of quantile functions using the empirical quantile function:

$$\hat{Q}(p) = X_{(\lceil np \rceil)} \tag{4}$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the order statistics of a sample of size n.

For the QDI, we implement a weighted version that allows researchers to emphasize specific regions of the distribution:

$$QDI_w = \sum_{i=1}^{k} w_i \left| \frac{Q(p_i) - Q(0.5)}{Q(0.5)} \right| \tag{5}$$

where $\sum w_i = 1$. This flexibility is particularly valuable in applications where inequality in specific regions (e.g., the upper tail for wealth distributions) is of primary interest.

The DAM implementation requires careful consideration of the weighting function. We propose and compare several weighting schemes:

- Linear weighting: $w(p) = p - 0.5$
- Quadratic weighting: $w(p) = (p - 0.5)^2 \cdot \text{sign}(p - 0.5)$
- Focused weighting: $w(p) = \phi(p; \mu, \sigma)$ where ϕ is the normal density function centered at regions of interest

For the QOC, we develop a smoothed version to reduce sensitivity to sampling variability:

$$QOC_s = 1 - \frac{1}{2}\int_0^1 \left| \frac{\hat{Q}_1(p) - \hat{Q}_2(p)}{\max(\hat{Q}_1(p), \hat{Q}_2(p))} \right| K_h(p)dp \tag{6}$$

where $K_h(p)$ is a smoothing kernel that downweights extreme quantiles where estimation is less precise.

We address several computational challenges in implementing these measures. First, we develop efficient algorithms for quantile estimation that handle large datasets. Second, we provide methods for calculating standard errors and confidence intervals using bootstrap techniques. Third, we implement visualization tools that help interpret the quantile-based measures, including quantile difference plots and distributional comparison graphs.

Our approach also includes diagnostic tools for assessing the reliability of the measures. We develop goodness-of-fit tests for the quantile estimation and sensitivity analyses for the choice of probability levels and weighting functions.

4 Simulation Studies

We conducted extensive simulation studies to evaluate the performance of our proposed measures and compare them to traditional inequality metrics. Our simulations covered a range of distributional scenarios relevant to real-world applications.

In the first simulation, we generated data from log-normal distributions with varying parameters to represent different inequality scenarios. We compared the sensitivity of our QDI measure to traditional measures including the Gini coefficient, coefficient of variation, and Theil index. The results demonstrated that QDI provides more nuanced information about distributional shape, particularly in detecting changes that affect different parts of the distribution unequally.

A second simulation focused on the DAM's ability to detect asymmetric distributional changes. We simulated scenarios where distributions shifted in ways that increased inequality in the upper tail while decreasing it in the lower tail, and vice versa. Traditional symmetric measures often failed to detect these changes or provided misleading summaries, while the DAM successfully quantified the asymmetric nature of the distributional shifts.

The third simulation evaluated the QOC's performance in measuring distributional similarity. We generated pairs of distributions with varying degrees of overlap and compared the QOC to traditional measures of distributional difference including the Kolmogorov-Smirnov statistic and Cramér-von Mises criterion. The QOC provided a more intuitive and interpretable measure of similarity that was robust to sample size variations.

All simulation studies included sensitivity analyses examining the effects of sample size, distributional assumptions, and choice of parameters in our measures. The results consistently supported the value of our quantile-based approach in providing more detailed and robust distributional comparisons.

5 Empirical Applications

We applied our quantile-based framework to three empirical domains to demonstrate its practical utility and novel insights.

In economics, we analyzed income distribution data from multiple countries over time. Traditional measures showed generally increasing inequality, but our quantile-based approach revealed important nuances. The QDI showed that inequality increases were concentrated in specific regions of the distribution that varied by country. The DAM detected asymmetric patterns in how economic growth affected different income groups, with some countries showing pro-poor growth patterns and others showing growth concentrated at the top. These insights would have been missed by traditional symmetric inequality measures.

In environmental science, we applied our methods to distributions of pollutant concentrations across different regions. The QOC provided a nuanced measure of distributional similarity that helped identify regions with similar pollution profiles despite differences in average concentrations. This has important implications for environmental policy and regulation, as regions with similar distributional patterns may benefit from similar intervention strategies.

In public health, we analyzed distributions of health outcomes across demographic groups. Our approach revealed distributional differences that were masked by traditional comparisons of means or medians. Specifically, the DAM identified asymmetric patterns in how health interventions affected different parts of the outcome distribution, providing valuable information for targeting public health resources.

Across all applications, our quantile-based measures provided insights that complemented and sometimes contradicted those from traditional approaches. The ability to examine distributional patterns across the entire range of values, rather than relying on summary statistics, proved particularly valuable in understanding complex real-world phenomena.

6 Discussion and Conclusion

This paper has introduced a comprehensive framework for measuring inequality and distributional differences using quantile-based methods. Our approach addresses several limitations of traditional inequality measures by leveraging the complete distributional information contained in quantile functions.

The primary contribution of our work is the development of three novel measures: the Quantile Dispersion Index (QDI), Distributional Asymmetry Metric (DAM), and Quantile Overlap Coefficient (QOC). These measures provide more nuanced and informative assessments of distributional characteristics than traditional scalar summaries. They are particularly valuable for detecting patterns in distributional tails, quantifying asymmetric changes, and measuring distributional similarity.

Our simulation studies demonstrated the superior performance of these measures in various scenarios, while our empirical applications showed their practical utility across multiple domains. The insights gained from our quantile-based approach would have been difficult or impossible to obtain using traditional methods.

Several directions for future research emerge from our work. First, there is potential to extend our framework to multivariate distributions, where quantile-based approaches face additional challenges but offer correspondingly greater rewards. Second, our measures could be adapted for dynamic analysis of distributional changes over time. Third, there are opportunities to develop formal statistical inference procedures specifically tailored to our quantile-based measures.

In conclusion, our quantile-based framework represents a significant advancement in the measurement of inequality and distributional differences. By moving beyond scalar summaries to functional approaches that preserve distributional richness, we provide researchers and policymakers with more powerful tools for understanding and addressing distributional challenges across various domains. The novel insights generated by our approach demonstrate the value of rethinking traditional measurement paradigms and embracing more distributionally complete methodologies.

References

Chen, L., Wei, Y. (2020). Quantile regression: Theory and applications. Journal of Economic Perspectives, 34(2), 145-166.

Davino, C., Furno, M., Vistocco, D. (2013). Quantile regression: Theory and applications. John Wiley Sons.

Koenker, R. (2005). Quantile regression. Cambridge university press.

Koenker, R., Bassett Jr, G. (1978). Regression quantiles. Econometrica: journal of the Econometric Society, 33-50.

Muller, A., Stoyan, D. (2002). Comparison methods for stochastic models and risks. John Wiley Sons.

Serfling, R. J. (2009). Approximation theorems of mathematical statistics. John Wiley Sons.

Shorack, G. R., Wellner, J. A. (2009). Empirical processes with applications to statistics. Society for Industrial and Applied Mathematics.

Van der Vaart, A. W. (2000). Asymptotic statistics. Cambridge university press.

Wasserman, L. (2006). All of nonparametric statistics. Springer Science Business Media.

Wilcox, R. R. (2012). Introduction to robust estimation and hypothesis testing. Academic Press.