

DISTRIBUTION-FREE CONFIDENCE INTERVALS FOR FUNCTIONS OF QUANTILES

CHANDIMA N. P. G. ARACHCHIGE 

(Received 15 July 2020; first published online 17 September 2020)

2020 *Mathematics subject classification*: primary 62G08; secondary 62G15.

Keywords and phrases: confidence interval, quantile, estimator, variation, skewness.

The objective of this thesis was to provide new distribution-free point and interval estimators for measures of spread, relative spread and skewness involving quantiles. The main advantages of these new quantile-based measures are that they are comparatively efficient to compute with minimal distributional assumptions. In the two-samples case, the location and scale of two independent samples can be compared using ratios of linear combinations of quantiles. In the single-sample case, distribution-free inferences can be made using quantile versions of the coefficient of variation and measures of skewness. The new estimators can be directly applied to many areas, including economics, medical statistics, biostatistics and social sciences.

The thesis consists of four papers [1–4], which had been either published or submitted at the time of thesis submission, and also includes some introductory material.

In [1], the distribution-free point and interval estimators were introduced for ratios of independent quantiles to compare the locations of two independent samples. The interquantile range, which is the most natural quantile-based estimator of scale, was used to compare the scales of two independent samples. The distribution-free point and interval estimators were applied to the squared ratios of interquantile ranges. The optimal probability to achieve the minimum asymptotic variance for the squared ratio of interquantile ranges was determined. Robustness properties of the estimators were investigated using partial influence functions. The simulation results reveal that all the intervals provide excellent coverage probabilities, even for small sample sizes and for a wide range of distributions. An R Shiny web application was developed and is publicly available to readers to run the simulations as they desire. Some

Thesis submitted to La Trobe University in December 2019; degree approved on 21 April 2020; supervisors Luke Prendergast and Bob Staudte.

© 2020 Australian Mathematical Publishing Association Inc.

real-world data examples were considered; the results suggest that the new estimators perform very well compared with classical parametric tests such as the t -test and F -test.

In [2], the median absolute deviation, which is the most robust estimator of scale with respect to the breakdown point, was considered. The distribution-free point and interval estimators were applied to the median absolute deviation and the difference and squared ratio of median absolute deviations, to make inferences about the spread of a single sample and to compare the spreads of two populations, respectively. Robustness properties of the new estimators were investigated using an influence function and partial influence functions. Simulations were conducted to check the performance of the new estimators; the results suggest that the coverage probabilities are very close to the nominal coverage, even with small sample sizes and for a variety of distributions. A real-world data example was considered; the results suggest that the new estimators are more robust to outliers than the F -test.

In [4], two robust versions of the coefficient of variation based on linear combinations of quantiles were used to make distribution-free inferences of relative dispersion for a single sample. The first measure was the interquartile range divided by the median; the second measure was the median absolute deviation divided by the median. The distribution-free point and interval estimators were constructed for the two robust versions of the coefficient of variation, and the robustness properties were investigated using influence functions. The performance of the new estimators was compared with several existing estimators via simulations; the results suggest that the new methods perform well for a wide variety of distributions. An R Shiny web application was developed to compare the performance of the new estimators of spread with some existing estimators; this application is publicly available. The interval estimators were extended to the ratios of the robust coefficients of variation to compare the relative dispersion between two independent samples. The examples reveal that different conclusions can be made based on robust and nonrobust versions of the coefficient of variation.

In [3], some integrated versions were constructed for existing measures of skewness based on ratios of linear combinations of quantiles. These existing skewness measures are generalisations of Bowley's well-known skewness coefficient. The properties that any measures of skewness should satisfy were validated for the new measures of skewness. The distribution-free point and interval estimators were introduced for the new measures of skewness to make inferences about the skewness of a single sample. A simulation study was conducted to compare the performance of the new estimators with existing estimators; the results suggest that the new measures perform well for a wide range of distributions. An R shiny web application was developed to compare the performances of the new skewness estimators and is publicly available. The interval estimators were applied to the differences between the measures of skewness to compare the skewness between two independent samples. Some real-world examples were used and different conclusions were drawn based on different methods.

References

- [1] C. N. P. G. Arachchige, M. Cairns and L. A. Prendergast, 'Interval estimators for ratios of independent quantiles and interquantile ranges', *Comm. Statist. Simulation Comput.*, Published online (5 July 2019), 17 pages.
- [2] C. N. P. G. Arachchige and L. A. Prendergast, 'Confidence intervals for median absolute deviations', Preprint, [arXiv:1910.00229](https://arxiv.org/abs/1910.00229).
- [3] C. N. P. G. Arachchige and L. A. Prendergast, 'Mean skewness measures', Preprint, [arxiv:1912.06996](https://arxiv.org/abs/1912.06996).
- [4] C. N. P. G. Arachchige, L. A. Prendergast and R. G. Staudte, 'Robust analogues to the coefficient of variation', *J. App. Stat.*, Published online (20 August 2020), 23 pages.

CHANDIMA N. P. G. ARACHCHIGE, Department of Mathematics and Statistics,
La Trobe University, Bundoora, Victoria 3086, Australia
e-mail: chandimanilanthi@gmail.com