



Center for Strategic and Budgetary Assessments

ANALYSIS

Statistics in pursuit of social justice

June 6, 2018 | *The Royal Statistical Society*

By: Harrison Schramm, Tracey Pérez Koehlmoos and Scott Nestler

Related Expert: Harrison Schramm

Data skills are in demand. E-commerce, advertising and artificial intelligence companies are willing to pay substantial sums for statistical knowhow. But statisticians can do more than make their mark in business. They can have real impact on people's lives, write Harrison Schramm, Tracey Pérez Koehlmoos and Scott Nestler

"Statistician is the coolest job you've never heard of", says the American Statistical Association (ASA). The ASA's This Is Statistics campaign, which encourages students to pursue careers in statistics (bit.ly/2qRFAbc), explains how "Statisticians contribute to society in many ways, from protecting endangered species and managing the impacts of climate change to making medicines more effective and reducing hunger and disease".

These are all lofty and attractive features of a career in statistics, but they are often underpromoted. Instead, job postings for "data scientists" and careers articles discussing "big data" roles emphasise the statistician's role in "monetisation", the process by which data and statistics are used to develop "insights" about customers in order to maximise profit.

Helping companies to grow and profit is certainly a worthwhile application of statistics and helps advance the science. Statisticians who make a success of this work stand to earn a sizeable salary (bit.ly/2HKvmbM). However, the emphasis on statistics as a business tool risks downplaying one crucial aspect of the discipline: while statistics is generally thought of as the science of data or the science of uncertainty, it is also the science of social justice.

We define social justice as an ideal that seeks to advance society by promoting opportunity and encouraging institutions to work together for the betterment of all humankind. We assume that good policies advance social justice, and that vigorous policy debates should be informed by responsible practitioners and reliable data. Statisticians therefore have a vital role to play. For many issues in public life, statistics – and the rigorous application thereof – is a necessary scientific and disciplined thought approach to understand how actions and policies affect people's lives. Our concern is that many students or job seekers – and even a few already in the profession – now fail to recognise this.

A foundational purpose

The pursuit of social justice has been a feature of the statistics discipline at least since the 1830s, back when the Royal Statistical Society (RSS) was first established. The early basis of the RSS was in “social issues rather than mathematics”, and it lists among its founding aims the “collection and classification of all facts illustrative of the present condition and prospects of Society” (bit.ly/2HqFeEs). Of particular note, the first female member of the RSS was Florence Nightingale, considered by many to be the founder of modern nursing.

The ASA also claims Nightingale as an early member. Its aims are: “to foster statistics and its applications, to promote unity and effectiveness of effort among all concerned with statistical problems, and to increase the contribution of statistics to human welfare”. This commitment has continued, with the ASA issuing statements in the past year supporting the March for Science, opposing the Trump administration's (then) proposed travel ban, and refusing to sign contracts in states with discriminatory laws (bit.ly/2HaWycG).

This tradition of statistics in social justice first acknowledges that the results of statistical work are going to have tangible impact on real people, and this necessitates that statisticians hold their work to high standards, such as those set out in the ASA's Ethical Guidelines for Statistical Practice (bit.ly/2JkhTVE). In our view, the successful application of statistics to the issues of social justice requires our work to be:

1. Accessible to policy-makers and programme managers, who are unlikely to be statisticians or have substantial statistical training.

2. Enduring, in the sense that the recommendations will not be obsolete shortly after implementation or – in the worst case – wrong. We consider the principles of relevance and validity to be necessary preconditions for an enduring solution.

3. Responsible, meaning that practitioners are exceptionally conscientious about data sources, paying particular attention to individuals and populations that may have been inadvertently “missed”

because of sampling methods. This is especially important when using largescale automated data collection, where the practitioner does not have their hands directly on the data.

4. Caring about outliers. In human terms, outliers are those “special” cases that do not fit cleanly into the mould or model. These are wrongly thought of in terms of missed classifications or residuals, and more properly thought of in terms of individual lives.

5. Traceable. Given the explosion in “black box” methods, by which we mean some machine learning or artificial intelligence techniques, traceability and mechanisms for appeal of decisions need to be included. We must be prepared to explain how the inference machinery works.

The work of public policy is often slow and long-lasting. Once a policy becomes law, it gains institutional inertia and can be difficult to change. However, the application of statistics for social good can be very swift. For example, the 1854 Broad Street cholera outbreak in London is a case of an observant practitioner, Dr. John Snow, having data at hand, who then isolated the source of the epidemic to a single contaminated well.¹ He was able to do this because of empirical knowledge from his practice as a physician, as well as his understanding of what the data meant. In this case, his knowledge of both the mechanics of epidemics and the underlying social structure – as well as the set of corrective actions – enabled a timely and life-saving intervention.

Decisions made in public policy over short time periods – including policies influencing healthcare, resources, and the environment – affect people of all ages, but particularly during childhood, and these can have lifelong impacts.² Studies and discussion of how people's lives are affected over long time horizons by rapidly made policy decisions are particularly poignant in this moment when, both in the United States and the United Kingdom, long-standing policies on health, education and the environment are changing at a pace that is difficult to keep up with. These issues are critical to individual and societal well-being and are not independent; changes in one affect the others. Therefore, now is an opportunity for statisticians to bring dispassionate, scientifically based results into the international discourse.

Making an impact

We set about writing this article following an exchange with a student who was trying to decide on an undergraduate major. When we discussed aspects of our practice as statisticians – which, in truth, tend towards the technical – the student's response was to cringe and say, "I want to do some good in the world, not just crunch numbers and do math". The history of statistics shows that such a laudable goal has been shared by generations of statisticians – and for those students who want to pursue the same goal today, we offer the following advice.

A group squarely at the intersection of public policy and statistics are public health professionals. Master's and PhD programmes in public health offer concentrations or specialisation in biostatistics, and this is a direct line to improving the understanding of the needs of populations and evaluating interventions that improve the lives of the disadvantaged. Further, for those already in possession of a statistical degree (even if that degree's name does not explicitly include "statistics"), there are ample opportunities – and a pressing need – for savvy individuals who possess both mathematical and domain expertise to have a direct, positive impact in the areas of social justice, international development, health and healthcare, education, and behaviour through various levels of research service – whether primary, secondary or synthesis.

There is demand for more rigorous primary research for those working closest to the population, specifically in the areas of public policy, social sciences, and international development in resource-poor settings, in which real life is often improved by programmes built rapidly but without robust analytical plans. Organisations like the International Initiative for Impact Evaluation (3ieimpact.org) seek to create opportunities for this evaluation to happen.

For those students more inclined to secondary analysis, in these days of "big data" and the Data Liberation Movement, there are ample opportunities to use statistical skills to combine existing data sets – one example might be to link Demographic and Health Surveys data collected in low- and middleincome countries (dhsprogram.com/data/) with data sets looking at quality of services, or geographic information systems, to better understand disease burden (bit.ly/2Jhk5wS).

However, there are also opportunities to engage at a higher level of analysis, with an even more robust means of addressing the real impact of interventions through research synthesis. Much of the best-known work in this area is by the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre), and the Cochrane and Campbell systematic review organisations.

Systematic review is more than a statistical technique for combining effect sizes from primary research; it is a replicable approach to summarising the results from existing studies, offering complete transparency in the process of question design, methodology, and search strategy. Systematic reviews apply established tools of critical appraisal to the assessment of primary studies included in the review – whether randomised controlled trials (RCTs) or otherwise – so that, in the end, the result is more than “this works”, but rather “this works and with such-and-such a degree of confidence”.

While the work done by Cochrane and Campbell enjoys a reputation for being conducted by armies of volunteers with a burning desire to answer pressing questions, build an evidence base, and inform policy and decision-making, increasingly each organisation and other organisations offer funding opportunities to move into and work in the field. Of particular note is the Campbell and Cochrane Equity Methods Group (bit.ly/2Jjx3hf), which aims to encourage systematic reviewers to incorporate subgroup analysis focusing on potentially disadvantaged populations (the poor, women, rural versus urban, racial groups, etc.) in order to present a more robust picture of how an intervention or approach is affecting the lives of the entire population. Meanwhile, programs like the new What Works centres in the UK and data clearinghouses in the USA, as well as government-sponsored centres in Denmark, Sweden, and Norway, attempt to produce synthesised research and create accessible evidence-informed policy documents on a variety of topics.

At first glance, systematic review might seem like simple meta-analysis. But there is often a need to sort through and make sense of sometimes messy and incomparable primary studies, and to expand beyond RCTs. These trials are wonderful when well-planned and implemented, but are often not reflective of the methods used on the ground – especially in international development, where different methods are needed. Here, a well-trained and engaged statistician can make the immeasurable suddenly tangible. In fields such as education, health, development and justice, there may be plenty of subject matter experts, but no one who is fluent in mathematics and practical statistics. An analysis or full review is never the end product; the true end is the creation of policy briefs and other knowledge translation materials that can be easily accessed and understood by decision-makers and program planners. The practitioner who can make good science happen can change the world and how we live in it.

A particular irony is that organisations which advocate for disenfranchised groups and individuals lack the resources to perform analyses – statistical or otherwise – that could be of great value to those they seek to assist. Recognising this, several professional societies have set up programmes to fill the gap. For example, the ASA founded the Statistics Without Borders outreach group in 2008, along the lines of the well-known Doctors Without Borders, and in 2014 the RSS launched Statisticians for Society. More recently, in 2016, the Institute for Operations Research and the Management Sciences (INFORMS) created Pro Bono Analytics, which uses statistics alongside other applied quantitative techniques, such as simulation and optimisation. Organised efforts like these are starting to have an impact, but the demand for such services will likely exceed the supply, especially as their availability becomes more widely known.

How do you want to change the world?

For the student hoping to “do some good in the world”, it should be clear that the study and professional practice of statistics can have a real impact on the policies that shape people’s lives, and – as set out above – there are opportunities for practitioners of all levels to make a difference.

Today, as ever, quality analytics and data sources are of the utmost importance. But to be an agent of change, our work must be accessible, enduring, responsible, caring and traceable. Job adverts might highlight the important role statisticians have to play in supporting customer segmentation, targeted advertising, and revenue optimisation. But let us not forget that statistics is also the science of social justice.

Acknowledgements

HS and SN would like to thank many of their colleagues, particularly Professor Ron Fricker, for early readings of this article. TPK would like to thank Howard White for comments on an early version.

Disclaimer

The contents in this article are the sole responsibility of the authors and do not necessarily reflect the views, assertions, opinions or policies of the Uniformed Services University of the Health Sciences, the Department of Defense, or the Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organisations does not imply endorsement by the US government.

References

1. Snow, J. (1855) On the Mode of Communication of Cholera. London: John Churchill.
[http://www.ph.ucla \(http://www.ph.ucla\).edu/epi/snow/snowbook.html](http://www.ph.ucla.edu/epi/snow/snowbook.html)

