The inhouse scientific approach to official statistics at Statistics Sweden

Gustaf Strandell, Statistics Sweden, gustaf.strandell@scb.se

Ingegerd Jansson, Statistics Sweden, ingegerd.jansson@scb.se

Abstract.

As producers of official statistics, we are committed to use scientific criteria in our selection of data, methods and procedures. But what does scientific criteria actually mean, and how should such criteria affect how we as statisticians approach our everyday tasks? Based on science as an idea and generator of utility, in this paper we discuss the role of research and science in official statistics and the need for an inhouse scientific approach at the statistical office. Having access to such an approach is especially important in modernization of official statistics. The discussion converges in four concrete methodological principles; Learn from what others have done, Test, evaluate and draw conclusions, Document and justify choices and Let others take part of the work and be open to a constructive dialogue. At Statistics Sweden these principles guide us to ensure the scientificity of official statistics.

Keywords: Scientific criteria, inhouse scientific approach, methodological principles, research

1. Introduction

The scientific validity of official statistics is laid down in Regulation (EC) No 223/2009 of the European Parliament and of the Council, which states that scientific criteria shall be used in the selection of sources, methods and procedures (European Union 2024). Thus, science and especially statistical science is of fundamental importance for the production of official statistics, in particular the development and application of sound statistical methods.

This is particularly true today with rapid technological developments providing new data sources, machine learning, and AI. These developments have the potential to improve many aspects of official statistics production, but it is not evident how they should be utilized. It is up to us as statisticians, both practitioners and researchers, to find the way forward. From a historic perspective, the scientific path has the best track record.

But what does ‘scientific’ and ‘scientific criteria’ mean for official statistics production, and what implications do the terms have for us as producers of official statistics? In this paper we make a brief overview of the impact of science on official statistics. We present four principles, intended to support scientifically sound decisions in our daily work as methodologists and statisticians and our ability to demonstrate statistical leadership.

2. Science and official statistics production

The roots of science can be traced back many thousand years to the early gathering and sharing of knowledge on agriculture and population. It was not until the 17th century, with for example the formulation of Isaac Newton’s laws of motion, that science established itself as the main method for explaining how the world works. By moving away from past beliefs in authority, religious literalism, and superstition, in favor of rational thinking, objectivity (Gelman and Hennig 2017) and actual observations, science has generated great benefits for humankind such as higher life expectancy, growing prosperity, and the ability to make informed decisions.

The purpose of science is to generate knowledge, and to understand why science is particularly suited to produce useful results we need to ask what is meant by

knowledge. In philosophy, knowledge is often characterized as a true belief that can be justified based on the process in which the belief has arisen. It is not enough to be convinced of something, even if it is true; knowledge also requires acceptable reasons underlying the conviction. Rolling four sixes in a row using a fair dice does not imply great knowledge in the field of dice rolling. Three important factors in the justification of scientific knowledge are the use of scientific methods, scientific theory, and constructive criticism, i.e. a fruitful dialogue with others in the same field. These factors are discussed shortly below.

To work scientifically means to limit oneself to sound scientific methods that promote consistency with actual observations and objective inference. These scientific methods have emerged over time and their suitability varies depending on what is to be investigated. Source criticism has its roots in historical research but is in principle important in all fields of science. Mathematical reasoning is a foundation in several of the natural sciences. In quantitative sciences, methods for conducting experiments or other forms of data collection are combined with statistical inference and hypothesis testing to determine which conclusions can be drawn from the observations.

Scientific knowledge does not usually appear in isolation but as building blocks of a scientific theory, i.e. in a rationally coherent system of definitions, assumptions, and knowledge. A theory makes it possible to understand the scientific results from a broader perspective and is often used as a basis for explaining observed phenomena. Some scientific theories, such as classical mechanics, will supply clear instructions on how to carry out certain actions and what to expect as an outcome. Other theories are less hands-on, such as the theory of evolution or the Big Bang theory. It might happen that new knowledge overturns an existing theory, but it is much more common for agreement with the theory to be seen as a piece of the puzzle in justifying a new result.

Constructive criticism, scrutiny, and discussion were highlighted already by the thinkers of ancient Greece as important elements in the creation of knowledge. An approach that gained early recognition consisted of developing a hypothesis and subjecting it to sharp criticism so that a better hypothesis could emerge from its remains. To ensure new research results, the system of peer review is currently

applied by scientific journals, where researchers critically review each other's work before it is made available to a wider audience.

Since the first half of the 20th century, statistics is an established science with the main purpose to provide theories and methods that can be used to produce knowledge on different topics or in other sciences. Through design-based estimation theory, statistics provide, for example, clear instructions on how a survey should be designed to provide objective statistics of appropriate quality. Statistical theory is often formulated using mathematics, but statistical science also encompasses less mathematical answers, for example on how to conduct experiments or how to formulate objective questions in a questionnaire.

Production of official statistics is not research, but it is a scientific activity. For Statistics Sweden, science has been involved from the very beginning. The world's first continuous production of statistics, Tabellverket[1], published population statistics for Sweden and Finland. Together with the work of Linnaeus and Celsius, it formed the basis for Sweden's scientific awakening during the 18th century's so called freedom period. Tabellverket has recently been selected as one of 100 items for the Swedish culture canon.

Official statistics production includes areas that do not have a comprehensive theoretical basis and that are not much dealt with at universities. Although theories and established methods are important foundations, methods for official statistics production also include elements where statistical science does not produce direct answers. The knowledge of how these areas should be handled comes largely from applications in the production of statistics. This is where practice and evidence-based experience come into play, i.e. there is evidence that something works through knowledge that has grown over time in the profession. This is not contrary to scientific or innovative thinking (Box 1976), even though there could be obstacles (Dillman 1996). Through a scientific approach, the soundness of these practices can be guaranteed.

[1] Startsida | Tabellverket

3. Four principles

From the above discussion on how scientific knowledge is justified we have derived four principles intended to support a scientifically sound way of working. They do not define science, but rather a scientific process. The principles have been subjected to criticism within Statistics Sweden, and they have been communicated to all methodologists.

As methodologists at Statistics Sweden, we apply a scientific approach to production of statistics and development of the production process. The scientific approach includes elements that are well established in scientific contexts.

Learn from what others have done

Learning from others, both within Statistics Sweden and elsewhere, is an investment that saves time, gives better results, and helps to avoid mistakes. By keeping up to date, we benchmark our processes and results to the current state of sound practices in our field.

Test, evaluate and draw conclusions

It is not uncommon for intuition to lead us astray. Testing on real or simulated data is therefore necessary, and evidence always weighs heavily. A methodological challenge is to be able to propose a design or a method, while at the same time being prepared to reconsider the choices made based on new knowledge from evaluations or tests.

Document and justify choices

The choices made in production of the statistics are reported and documented. It is just as important to explain and justify why we have done what we have done, and why other alternatives have been rejected. It benefits us, our successors, and all users of the statistics.

Let others take part of the work and be open to a constructive dialogue

Within and outside Statistics Sweden, there are many opportunities to conduct an open dialogue by sharing and receiving views on our work. Some examples are talking to colleagues, asking colleagues to read a text, give a presentation for colleagues, Statistics Sweden's Scientific Advisory Board, or at a conference, or write

a methodology report or a scientific article. It creates the opportunity to benefit from the skills and experiences of others and thereby contribute to a learning organization.

4. Discussion

The scientific integrity of the official statistics producer is closely connected to the trust and confidence that stakeholders and users of the statistics have in the producer and the statistics. With that in mind, scientific principles should not only be applied in-house but also communicated to the outside world. Transparency regarding methods, data sources, and technologies should be part of the scientific 'infrastructure' applied by the producer. This becomes even more important as new approaches are deployed, such as machine learning methods for classification or methods for estimation using non-probability data sources. The importance of ethics, trust and transparency in official statistics are topics that deserves a deeper discussion than we can provide in this paper, see for example Eltinge (2025) or Fellegi (2004).

Science is old. So old in fact that we might take the scientific process for granted, and here lays a great danger. This is particularly true given today's political tendences towards a value relativism where opinions held by some politicians are viewed just as 'true' as evidence and scientific knowledge. Alleged "facts" become a weapon against political opponents. But as Groshen (2021) puts it, "Indeed, evidence is especially important for building joint solutions in fraught and partisan contexts". In a world where science as well as official statistics lose their value, benefits such as increasing mean life expectancy are lost too.

With this paper we like to remind ourselves as producers of official statistics to always be aware of some core principles in the scientific process. Science is better.

5. References

Box, G. E. P. 1976. Science and Statistics. *Journal of the American Statistical Association*, 71:356. https://doi.org/10.2307/2286841

Dillman, D. A. 1996. Why Innovation is Difficult in Governments Surveys. *Journal of Official Statistics*, 12:2. Tiff2PDF

Eltinge, J. L. 2025. Some Dimensions of Statistical Ethics and Scientific Integrity That Warrant Exploration Through Empirical Studies of Stakeholder Information Needs. *Journal of Official Statistics*, 41:3. https://doi.org/10.1177/0282423X251331327

European Union. 2024. REGULATION (EC) No 223/2009 of the European parliament and of the council. eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009R0223

Fellegi, I.P. 2004. Official Statistics—Pressures and Challenges ISI President's Invited Lecture, 2003. *International Statistical Review*, 72:1. https://doi.org/10.1111/j.1751-5823.2004.tb00229.x

Gelman, A. and Hennig, C. 2017. Beyond Subjective and Ojective in Statistics. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* 180:4. https://www.jstor.org/stable/44682661

Groshen, E. L. 2021. The Future of Official Statistics. *Harvard Data Science Review*, 3:4. https://doi.org/10.1162/99608f92.591917c6