

President's Message – October 2017

On a recent flight, I read an article in the airline's magazine on employment in workplaces increasingly driven by 'big data' and 'big data analytics'. In commenting on the various 'hybrids' of skills and backgrounds needed, not once did the word statistics appear, but the only workplace person quoted was a statistician, who was also explicitly identified as a statistician! The statistician stressed the need for ability to analyse and communicate as well as technical skills, and it was very clear that the emphasis was on the key skills that statistics professionals and educators have been highlighting for decades, including collaboration, communication, and interpretation of data in context. There are many great relevant quotes from statisticians I use in speaking on this, including two of my favourites:



"we see, tied up together, the role of the statistician as consultant, consultancy as the stimulus for research in statistics, and consultancy as the basis for teaching statistics" (Barnett, 1986).

"In university courses.....The practice [of providing data and a description of context] has led some graduates to be of the opinion that taking part in the collection of data is a waste of the statistician's precious time" (Kenett and Thyregod, 2006).

In discussion at the 9th Australian Conference on Teaching Statistics (http://iase-web.org/documents/anzcots/OZCOTS_2016_Proceedings.pdf), a number of delegates complained about data science programs being set up in their universities with no or negligible statistics. What are we to make of this, plus statements such as "Data scientists are not statisticians"? I also witnessed beliefs of some at the first UN World Data Forum that statistical literacy and data literacy are different, and that the former is 'smaller' than the latter. Such beliefs contradict the long-time and extensive emphasis that producing, sourcing, handling, managing, interpreting and communicating data are core parts of the full statistical investigation process.

A number of colleagues in other disciplines have expressed concern that the data science presented in workshops or seminars they have attended did not seem to critique or question - not just analysis, but also data sources, nature, quality and interpretation. Does data science therefore manage and handle data, and statistics analyse, interpret and question? The answer of 'Of course not' can be seen above. Tackling real problems requires collaboration and combinations of skills. However one wishes to describe statistics and data science, it is clear that their educational challenges relate to their pervasiveness and reach - being fed by, and feeding into, almost all other disciplines.

We do not want over-extended shallow hybrids, but graduates with skills for genuine collaboration and foundations for ongoing learning. It is not possible, at any educational level, to 'cover' such a large and far-reaching discipline as statistics, let alone the 'big tent' of

statistics + data science. No matter how one views their interrelationships, educationally, both need a balanced development of thinking and skills within active learning environments looking to students' current and future needs. The power of statistics lies in its full investigative approach and the transferability of statistical thinking. But foundational learning of this, in other disciplines and in mainstream statistics, requires experiential learning with a considerable variety of real data experiences. The challenges in implementation are many, whether at school or tertiary level, as evidenced by insufficient penetration of authentic practice despite decades of dedicated work. Even with genuine collaboration and the best of intentions, designing curricula and learning experiences for statistics in other disciplines (including mathematics) and in mainstream statistics programs, is impeded by a multitude of habits (not just the dreaded $p=0.05$) and the tendency for over-focus on training for research, whether in the other discipline or in statistics itself.

An interesting addendum is that at the conference to which I was flying, there were some excellent presentations from other disciplines on their statistical challenges of small numbers of observations on multiple dependent variables and repeated measures. Not all current real problems involve big data, but all are complex requiring authentic statistical, technical and contextual thinking, collaboration, analysis, interpretation and communication.

References

Barnett, V. (1986). Statistical consultancy - a basis for teaching and research. In Davidson, R. (ed) *The Proceedings IASE/ISI 2nd International Conference on Teaching Statistics*, Vancouver: ISI, Voorburg, The Netherlands.

<http://iase-web.org/documents/papers/icots2/Barnett.pdf>

Kenett, R. and Thyregod, P. (2006). Aspects of statistical consulting not taught by academia. *Statistica Neerlandica*, 60(3), 396-411.

Helen MacGillivray
ISI President