*Have you heard of the "**Theory of Sampling**"?*

I bet you haven't (at least not this one!)

It sounds like standard stats theory...

BUT IT'S NOT!

The "Theory of Sampling" (ToS) - have statisticians missed the boat?

Damian Collins and Anne Harris
NSW DPIRD

Motivation - sampling compost

A researcher was interested in sampling compost for biological and chemical contaminants. He pointed us to the "Theory of Sampling" as a framework.

Compost Piles

Theory of Sampling (ToS)

ToS arose from work by chemical engineer **Pierre Gy** over 50 years ago.

Pierre Gy

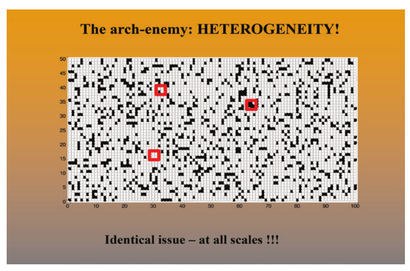
It concerns sampling of **bulk materials** such as soils, composts and ore deposits.

It now enjoys a strong advocacy network particularly amongst chemical engineers and geologists. As well as a formidable body of literature. Even a regular "ToS forum" journal. It is used to justify the recent review of composts and mulches in Australia ([1]).

ToS - sampling recommendations

ToS encourages practitioners to stop and think about the sampling process. ✔

ToS recognises heterogeneity in the sampling target and identify sources of variation in the sampling process. ✔



Heterogeneity *(from [2]*

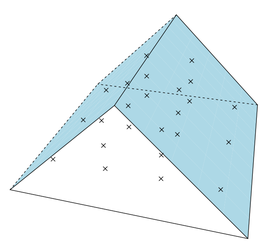
To deal with heterogeneity, rather than just taking a convenient "grab sample" from the surface...

"Grab sampling"

(from [2]

...take subsamples all across the entire sampling target - for instance, a compost pile. ✔

And combine these subsamples into one composite sample for analysis. ?



"Representative sampling"

of a compost pile

A "correct" sample for analysis

ToS encourages the analysis of a single composite sample. In [2], Esbensen concludes

"Applied properly, the TOS allows us to forward only one aliquot to the laboratory for analytical determination. Only one is needed because the entire from-lot-to-analysis process honours the TOS's principles for representativity."

One "correct" representative sample?

Of course, one sample for analysis means ***neither*** sampling ***nor*** analytical error can be estimated (let alone separated).

It's almost like this sample is considered a "population in miniature" as in the Kruskal-Mostellor papers on representative sampling [3].

Model train set - a "population in miniature" *(from [3])*

Have statisticians missed the boat?

Despite the obvious intrinsic statistical nature of the ToS, the theory has developed almost entirely independently of statisticians. Some related literature might be [4], [5] and [6].

Consequently, some statistical ideas appear confused or misconstrued in the ToS literature, as in [2].

This poster hopefully increases your awareness of the inherent problems in ToS. It is a reminder to always critically assess so-called "statistical theories". And it demonstrates the dangers of not involving statisticians in the application of statistics.

Acknowledgments

We would like to acknowledge the input of John Maindonald and Alistair Gray. In particular, their referral to the Kruskal and Mosteller papers on "representative sampling". As well as our NSW DPIRD biometrical colleagues for helpful advice!

NSW GOVERNMENT

References

<- Use the QR code to download

[1] K. Wilkinson, J. Jasonsmith, J. Biala, C. Lee-Steere, and A. Borshall. Review of AS4454 - Composts, Soil Conditioners and Mulches. Online, 2023.

[2] Esbenson, K.H. Sampling: Theory and Practice. *Alchemist*, 85:3–6, 2017.

[3] William Kruskal and Frederick Mosteller. Representative Sampling, I: Non-Scientific Literature. *International Statistical Review*, 47(1):13–24, 1979.

[4] Eurachem and CITAC. Measurement uncertainty arising from sampling: a guide to methods and approaches. Eurachem/CITAC guide, 2019.

[5] AMCTB, Analytical Methods Committee, et al. Sampling theory and sampling uncertainty. *Analytical Methods*, 7(24):10085–10087, 2015.

[6] Charles H Proctor. Statistical considerations in bulk sampling. Technical report, North Carolina State University. Dept. of Statistics, 1990.