Systematic effects versus nuisance parameters

Richard Lockhart

Simon Fraser University

Nov 3, 2021

Why am I talking?

- No new suggested analysis techniques.
- Just mention some extra topics on which I have done no work.
- Some discussion of language.

Why am I talking?

- No new suggested analysis techniques.
- Just mention some extra topics on which I have done no work.
- Some discussion of language.
- Warning: almost every bullet point I wrote needs caveats.

Nuisance parameters – rambling

- The statisticians speaking so far have not been unhappy about the use of the term 'nuisance parameter'.
- But usually statistics sounds like: there is an iid data set.
- Or at least: there is a single data set.
- Then we have 'interest parameter' and 'nuisance parameters'.
- Our papers essentially never have $125 \pm 2 \pm 3$.
- And I hear the ± 3 is $\sqrt{s_1^2 + \cdots + s_k^2}$.
- I want to make sure I ask: what is the ideal number of $\pm$ terms?
- Why is it ever different from 1?
- Or should you usually frame your uncertainty some other way?

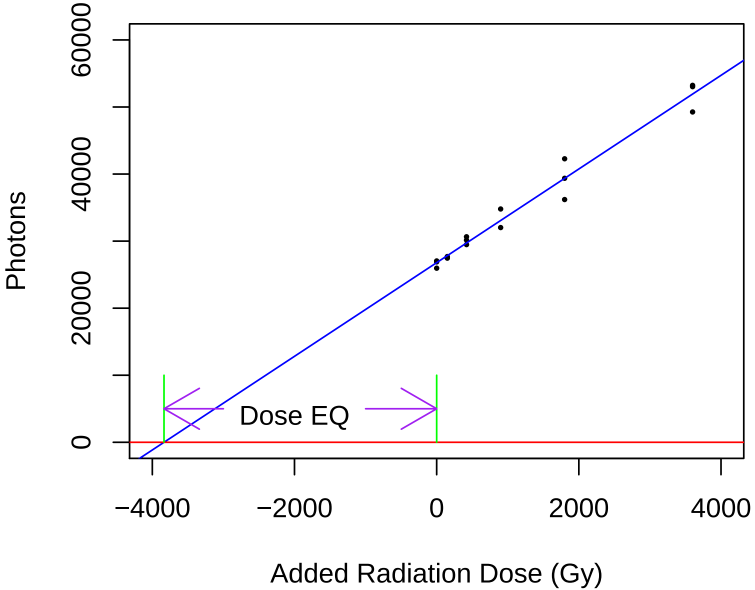
My subjects

- An old toy geophysics problem.
- Identifiability, ‘constrain’, ‘chi-square’;
- Single Experiment Analysis.
- Inference after model selection.
- Statistical vs non-statistical systematics.
- Theory via survey sampling.
- A list of ideas I have seen.

Old example

- Dating sand dunes.
- Drill core, cut into small samples.
- Varying doses of added radiation.
- Heated in oven; glows blue; count photons as sample heats.
- Extrapolate added dose back to 0 photon count.
- Convert dose to years – involves ‘nuisance parameter’, historic dose per year in ground.
- Plot photon count vs age.

Simplest Case



Discussion

- Parameter of interest is age of samples (time buried).
- Does not appear explicitly in model for data.
- Could, though, by writing Dose $=$ Age $\times$ Rate.
- This model is *unidentifiable*; extra information required.
- I wasn't involved in this issue.
- Common: collaborators produce some values for some things you need to know.
- Problem is sometimes abstracted by others before statistician thinks about it.
- Very challenging trade-off for statistician.

Identifiability – Statisticians' Attitudes

- We used to avoid unidentifiable models.
- Sand dune problem is unidentifiable.
- Background + signal is unidentifiable.
- Extra data (and maybe extra theory) needed.
- They 'constrain' the fit.

Inference after model selection

- Many modern problems require unidentifiable models: over-parametrized regressions, expansion truncation, etc.
- We use penalty terms or risk estimation or ... to choose between competing fits.
- This is model selection.
- Background modelling appears to me to involve model selection.
- Inference after model selection is a mildly hot topic.
- Mostly in regression modelling – but cf Sasha Glazov.
- POSI group at Penn State; Zurich: van de Geer, Bhlmann, Meinshausen; Stanford: Tibshirani, Taylor, Tibshirini.

Major criticisms of the new methods

- Uncheckable assumptions.
- Redefined targets of inference.
- Uncomputable constants.
- Stringent distributional assumptions. Outlier prone?
- Very wide confidence intervals.
- Big problems estimating variability – deliberate overestimation.
- Murder mystery problem: one of them did it but no idea which.

Issues we went by in first two days

- Model order selection; impact on inference.
- Likelihoods with only Gaussian terms and no acknowledgement of uncertainty in variability.
- Profile-log-likelihood in Gaussian model is not chi-squared but log-chi-squared.
- So when we see profile log-likelihoods which are just Gaussian they probably don't handle the tails properly.

Theory uncertainties: buyer beware

- One kind arises from expansions involving Feynman diagrams.
- Each order increase involves (a sum over I think) many more Feynman diagrams.
- Lots of work to compute all the contributions.
- One kind of statistician, not represented here, does survey sampling.
- Tentative proposal
  - draw a possibly stratified random sample of diagrams at a given order,
  - compute the contributions,
  - estimate total and attach sampling uncertainty.
- Suggestion may not be compatible with Frank Tackmann's "Theory Nuisance Parameters".

Hodge Podge

- Uncertainty Quantification (UQ): (e.g. SIAM / ASA journal). Numerical and statistical uncertainties
- Gaussian Processes in UQ: estimating non-random objects.
- My colleague Derek Bingham. Large scale slow computational code, GP emulator, data assimilation.
- Neyman-Scott problems: lots of parameters profiled over gives narrower log-likelihoods.
- Statisticians should be able to tell you when it is safe not to worry.
- We have seen many genuinely different problems so far and many analyses which have many of these at the same time.
- So I worry that useful advice from a statistician will likely require an embedding in a small group.

References I

- Berk, Brown, Buja, Zhang, & Zhao (2013). Valid post-selection inference. *Ann Statist*, **41**, 802–837.
- Bühlmann, Kalisch, and Meier (1014). High-dimensional statistics with a view toward applications in biology. *Ann Rev Stat Appl*, **1**, 255 - 278.
- Dezeure, Bühlmann, Meier, and Meinshausen (2015). High-Dimensional Inference: Confidence Intervals, *p*-Values, and `R`-Software `hdi`. *Stat Sci*, **30**, 533–558.
- Fithian, Sun and Taylor (2014). Optimal inference after model selection. *arXiv: 1410.2597*.
- Fithian, Taylor, Tibshirani, & Tibshirani (2015). Selective sequential model selection. *arXiv: 1512.02565*.
- Jewell, Fearnhead, and Witten (2019). Testing for a Change in Mean After Changepoint Detection, *arXiv: 1910.04291*.

References II

- Liu, Markovic, and Tibshirani (2018). More powerful post-selection inference, with application to the Lasso, *arXiv: 1801:09037*.
- Markovic, Xia, and Taylor (2017). Unifying approach to selective inference with applications to cross-validation, *arXiv: 1703:06559*.
- Markovic, Taylor, and Taylor (2019). Inference after black box selection, *arXiv: 1910.09973*.
- Meinshausen & Bühlmann (2010). Stability selection. *JRSS*-B, **72**, 417–473
- Wasserman & Roeder. (2009) High-dimensional variable selection. *Ann Statist*, **37**, 2178–2201.
- van de Geer, Bühlmann, Ritov, & Dezeure (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann Statist*, **42**, 1166-1201.

A useful post-selection blog entry