

Good proposal patterns

Several interesting things happened during the discussion in the proposal panel. Besides I unfortunately forgot to prepare the presentation, I learned some patterns that people would generally think they are good proposals:

1. **Observation proposal but includes simulations as a backup story:** like the proposal of Gunn-Peterson trough. It's a typical kind of proposal that people like. Since it includes observation and simulation, it's plausible both observers and simulators would vote for it.
2. **Simulation proposal but includes achievable observations:** e.g., neutrino mass proposal and lensing with the ray-tracing technique proposal. People are generally less interested in those simulations which are untestable in the near future, e.g., simulations from different dark matter candidates. People expect simulations would improve our understanding of the observations. So it's crucial that simulators should think about those observers when they are trying to write proposals. Especially, if your simulations are not going to benefit observers' understandings to their data, they will not give you money.
3. **You already have something:** proposals which already include some results are always preferred in our panel. I feel this is a little bit controversial because how is that possible you already have results before you have funds to do the research you are going to do. I also heard some people would use their past published results as their proposed research, and use the new funding to do whatever they want to do.

Things need to be avoided

1. **Avoid to use figures which are hard to explain**
2. **Don't assume everyone knows what you are doing**
3. **Write as easy as possible**

Non-representative panel

One essential thing people noticed during the discussion is that the reviewers (we) are not representative enough to reflect the actual proposal panel. For example, we have too many simulators, some earth scientists, and few observers. The composition of our panel is heterogeneous and non-representative, so the voting result would certainly be biased.

My first thought is “social science is hard”. The psychology survey data or political science voting forecast are always non-representative, I think. Calibrating the non-representative data and making on them is social scientists’ bread and butter.

I’ve read this paper¹ before written by statistician Prof. Gelman. By using hierarchical Bayesian modeling doing logistic regression, you can turn the non-representative survey to make robust predictions. He called this method “multilevel regression and poststratification (MRP)”. He used X-box to do the survey (, which for sure is non-representative because most of the players are male,) and made an accurate forecast for the 2012 presidential election.

The concept of MRP is to combine hierarchical structure in Bayesian modeling with the poststratification. The idea of poststratification is simple: we divide people into several cells based on their outcomes, e.g., genders, (theorists or observers or ...), ages, research interests (galaxy, AGN, exoplanet, IGM, ...), and so on. And to get an unbiased cell-level estimate, we simply do the average to get the mean value of the votings. The poststratification part is to scale the value of voting to an existing representative group. For example, we have five observers in our class, but, in reality, maybe the real panel for an observation proposal composited by 9/10 observers and 1/10 theorists. So the poststratification procedure (if we assume we only have one outcome and one group in consideration) is to first take the average to the voting results within the cell (observer or theorist):

$$a_{\text{theorist}} = (2 + 3 + 2 + 2 + 3 + 2)/6 = 2.33$$

$$a_{\text{observer}} = (1 + 2 + 2 + 1 + 2)/5 = 1.6$$

And then weighting the cell-level estimate to the representative group,

$$a = 1/10a_{\text{theorist}} + 9/10a_{\text{observer}} = 1.67$$

Note that if we are not doing the poststratification, the result of voting we get is 2.0. The assumption for taking average within the cell is that we assumed the sample is drawn randomly from the population.

The real life is not as simple as this example. To get a robust voting prediction, we have to consider all possible outcomes of the voters (us). Moreover, there may have some hierarchical structure between outcomes. For example, the same outcomes (gender, (simulator or observer), research interests) may have different voting results based on which subgroup (astrophysics or earth science) the voter belongs to. For instance, a researcher doing exoplanet in the astrophysics department may have a different voting result than an exoplaneter in the earth science department.

That’s the reason why multilevel modeling would be helpful for poststratification. We could build a Bayes net (multilevel modeling) to do the parameter estimation on different levels.

¹<https://www.microsoft.com/en-us/research/wp-content/uploads/2016/04/forecasting-with-nonrepresentative-polls.pdf>

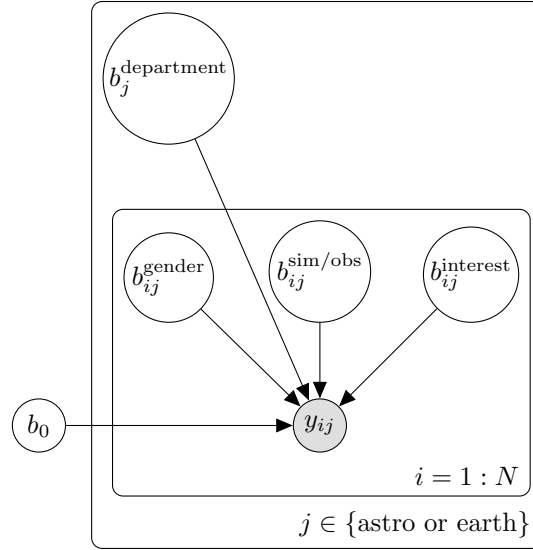


Figure 1: A Bayes net designed for our proposal panel

The final step is just weighing the estimation based on the cell-level (or departmental-level) to an appropriate representative composition of the panel.

See Fig 1 for the Bayes net for my proposed model of voting in our panel. The calculation of the probability would be,

$$\Pr(y_{ij} = \text{yes}) = \text{logit}^{-1}(b_0 + b_j^{\text{department}} + b_{ij}^{\text{gender}} + b_{ij}^{\text{sim/obs}} + b_{ij}^{\text{interest}}). \quad (1)$$

Assume b_{ij}^{outcomes} are randomly sampled from a prior distribution, and then fit the logistic regression model to the data.

Finally, poststratificate the $\Pr(y_{ij} = \text{yes})$ probabilities using the weighting estimated from the representative population. I would be interested in testing the model if I have the voting result and the composition of a real proposal panel.