Notes

Xingyu Chen[1]

November 07, 2025

Contents

1 Tpyst symbols

Tpyst symbols.

2 Random effects model for election polls

Andrew Gelman's recent blog post, Polls & Betting Odds & Nonsampling Errors & Win Probabilities & Vote Margins, discusses how to estimate winning probabilities based on polling data. A ChatGPT summary is available here: chatgpt-history.

He presents three examples: the New Jersey governor race, the Virginia governor race, and the New York City mayoral race. Individual polls are often unreliable, so systematic differences between polls must be considered. A simple random-effects model can capture this phenomenon:

$$Y_{i,t} = \theta + \eta_{i,t} + \varepsilon_i$$

Here, $Y_{i,t}$ is the t-th poll for the i-th media, θ is the true underlying support level, ε_i represents the systematic error for poll i, and $\eta_{i,t}$ represents the sampling error of the poll.

When focusing on a single media's polls (fixing $i = i_0$), Gelman uses historical information to estimate ε_t, assuming $\varepsilon_t \sim \mathcal{N}(0, 2\%)$. He then uses polling data of media i_0 to estimate the sample variance $\sigma_{i_0}^2$ of η_{i,i_0}.

Let $\overline{Y_{i_0}} = \frac{1}{n}\sum_{i=1}^{n} Y_{i,i_0}$. Under his Bayesian viewpoint, the posterior distribution of θ is approximated by:

$$\theta \mid Y \sim \mathcal{N}\left(\overline{Y_{i_0}}, \sqrt{\sigma_{i_0}^2 + 2\%}\right)$$

The winning probability can then be computed as:

$$\Pr(\theta > 0.5 \mid Y).$$

3 Good title: The unreasonable effectiveness of XX in/of YY

Good title form The unreasonable effectiveness of mathematics in the natural sciences by Eugene Wigner in 1960 [1].

Richard Hamming in computer science, "The Unreasonable Effectiveness of Mathematics".

Arthur Lesk in molecular biology, "The Unreasonable Effectiveness of Mathematics in Molecular Biology"

Peter Norvig in artificial intelligence, "The Unreasonable Effectiveness of Data"

Vela Velupillai in economics, "The Unreasonable Ineffectiveness of Mathematics in Economics"

Terrence Joseph Sejnowski in Artificial Intelligence: The Unreasonable Effectiveness of Deep Learning in Artificial Intelligence".

4 Good title: Rebel With a Cause

Form the famous movie Rebel Without a Cause.

The special issue Volume 8, Issue 2, 2022 Issue of *Observational Studies* titleed Rebel With a Cause

5 On the undistinguishable or identification of statistical models

Based on talk with Ruiqi Zhang, Lin Liu and Chatgpt.

The undistinguishable means you can not distinguish two models form data. Which will corresponds to 3 cases.

One and two are both in the modeling stage. When you consider one model P_θ, this corresponds to non-identifiable model. When you consider two models, this corresponds to two models sharing some common distributions.

The third case is in the hypothesis testing stage, two models can not be distinguished by data since they are too close.

In a word, two models $\mathcal{P}_1$ and $\mathcal{P}_2$ as two distribution calsses are undistinguishable if $\mathcal{P}_1 \cap \mathcal{P}_2 \neq \emptyset$ or they are too close under some metric.

5.1 On latent structure models

The potential outcome model is an example of latent structure model. The observed random variable is determined by some unobservable/latent variable is in this calss.

Definition 5.1. *The observed random variable X is determined by a high dimensional latent variable Z by a map $X = f(Z)$.*

***Example* 5.2.** The observed random variable is (A, Y) determined by three latent variable $(A, Y(0), Y(1)), A \in \{0, 1\}, Y = AY(1) + (1 - A)Y(0)$, consider two submodels:
- $\mathcal{P}_1$: $Y(1) - Y(0) = 0, Y(1) \perp A, Y(0) \perp A$
- $\mathcal{P}_2$: $Y(1) - Y(0) = Z \neq 0$, but $Y(1) \overset{d}{=} Y(0), Y(1) \perp A, Y(0) \perp A$.

The model 2 is not empty, take

$$Y(0), \varepsilon \sim \mathcal{N}(0, 1), Y(0) \perp \varepsilon, Z = -\frac{1}{2}Y(0) + \frac{\sqrt{3}}{2}\varepsilon$$

then $Y(1) \sim \mathcal{N}(0, 1)$ and $Y(1) \overset{d}{=} Y(0)$.

On the observed data level, we can not distinguish these two models since they both have the same conditional distribution of $Y|A$, therefore they are undistinguishable in modeling stage.

This example is the reason why the sharp null hypothesis can not be tested in randomized experiment, and also the joint distribution of $(Y(0), Y(1))$ is not identifiable.

[2] talk about the identification of joint distribution of potential outcomes under some assumptions.

6 On the cluster analysis

7 Numerical calcaulation of influence function

[3]

Bibliography

[1] E. P. Wigner and others, “The unreasonable effectiveness of mathematics in the natural sciences,” *Mathematics and science*, vol. 13, pp. 1–14, 1990.

[2] P. Wu and X. Mao, “The Promises of Multiple Experiments: Identifying Joint Distribution of Potential Outcomes,” *arXiv preprint arXiv:2504.20470*, 2025.

[3] Y. Mukhin, “Kernel von Mises Formula of the Influence Function,” in *The Thirty-ninth Annual Conference on Neural Information Processing Systems*,