Student name: Wai Yu Amanda, Ng

Student number (last 5 digits): 17350

Course code: ENV399

Instructor: Dr. Brad Bass

Topic: Frequentist approach vs Bayesian approach

Basic background ideas:

Frequentist and Bayesian approaches fundamentally differ in terms of the interpretation of statistical probability. The frequentist approach views the probability of an event as its proportion in a prespecified repeatable random event, equivalating to a “long-run frequency” *(Lecoutre & Poitevineau, 2014)*. In this case, probability should be objective, observable and uniquely defined since it is deducted directly through past outcomes and relies on experimental design. The frequentist approach uses maximum likelihood for parameter estimation which assumes parameters are unknown but fixed values. It then incorporates data from the population and returns the set of parameters that best fits all observed sample data. Bayesian approach views probability as a composition of not only past events but also personal experiences and judgments *(Lecoutre & Poitevineau, 2014)*. Bayesian models involve flexible parameters (not fixed) and require an additional step of stating one's prior expectations about the model parameters. Each parameter is assigned with a probability distribution of possible values using prior knowledge about the parameter values. As new data is observed, this new piece of information will be used to update the distribution and form a new posterior distribution for future subsequent predictions iteratively. In short, Bayesian inferences produce fixed probability intervals for unknown quantities, but frequentist inferences produce random intervals for fixed quantities.

<u>Statement: Frequentist approach is better than Bayesian approach.</u>

Inferences:

1) **The Bayesian approach is too subjective for statistical inferences while the frequentist approach is more objective.**

   The Bayesian approach involves a subjective definition of probabilities that relies on both data and hypothesis (which may be false) while the frequentist approach relies solely on objective past events. In principle, scientists should draw conclusions from objective facts instead of subjective beliefs. However, the Bayesian approach starts from assuming hypotheses, that may be implausible, to be true and derives the ratio of probabilities of them happening. This approach, theoretically speaking, violates the principle stated above since the whole process begins with a subjective stance *(Gelman, 2008)*. Several research projects show that results arising from the Bayesian approach are highly dependent on potentially subjective prior hypotheses *(Flori, 2019)*. In contrast, the frequentist approach is more conservative. It only looks at past events and provides a confidence interval for the unbiased estimates, hence guiding the decision of accepting or rejecting null hypotheses.

Sources/Example:

- *"Objections to Bayesian statistics." (Gelman, 2008)*
- *"News and subjective beliefs: A Bayesian approach to Bitcoin investments. Research in International Business and Finance" (Flori, 2019)* ---> This research demonstrates how

the prior knowledge of Bitcoin market/beliefs put into the Bayesian model can affect the predictive results.

2) **The Bayesian approach requires a full specification of a model to describe the parameter values which may not always be feasible.**

The Bayesian approach involves a full probability model for the data, including a high degree of specification of the prior distribution for unknown parameters. However, developing a good model and deciding which model to use is often challenging *(Efron 2005)*. It seems absurd whether the model chosen can fully reflect the nature of the parameters and hence its credibility. Most models are always simplified for idealization, but they can be disastrously wrong and lead to implausible answers. On top of that, models are vulnerable to subtle misspecification errors that are not easily picked up by model diagnostics *(Little, 2006)*. It is worth noting that model selection through Bayesian evidence can only assess the *relative* performance of the models *(Koo, 2022)*. Therefore, it is unable to rule out a better but still false model.

Sources/Example:

- *"Bayesians, Frequentists, and Scientists." (Efron 2005)* ---> This paper explains the level of difficulty in obtaining a high degree specification of prior distribution in Bayesian models.
- *"Calibrated Bayes: A Bayes/Frequentist Roadmap" (Little, 2006)*
- *"Bayesian vs frequentist: comparing Bayesian model selection with a frequentist approach using the iterative smoothing method" (Koo, 2022)* ---> This research demonstrates model selection through Bayesian evidence cannot rule out better but still false model while Frequentist evidence can.

3) The Bayesian approach produces too many possible solutions.

Extending from the issue of choosing the correct models (Inference 2), the next challenge of the Bayesian approach is picking one among them. It is revealed in several critics and research projects that the step of "feeding prior distributions" greatly affects the conclusions drawn in some applications *(Little, 2006; Smid, 2020)*. When the "input" prior distribution changes, the "output" parameter values may change drastically *(Smid, 2020)*. Additionally, the iterative nature of the Bayesian approach increases the differences in results between different models over the process. Hence, given multiple possible distributions that may work well with the data, different sets of parameter values generated create an overwhelming number of possible solutions *(Little, 2006)*. On the other hand, although the frequentist paradigm does not give exact answers to hypotheses, it indicates the extent to which the hypotheses may hold and are useful in ruling out hypotheses (such as rejecting null hypotheses).

Sources/Example:

- *"Calibrated Bayes: A Bayes/Frequentist Roadmap" (Little, 2006)*

  ---> This paper discusses in general terms the problem with the surfeit of posterior distributions.

- *"Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review" (Smid, 2020)*

---> This paper tests several Bayesian models on the same dataset, and each gives different results due to the highly dependent property on the strength of prior distributions.

4) **Frequentist approach is better calibrated in survey sampling setting than Bayesian approach.**

Approaches to finite population inferences are mainly classified into three streams: design-based, model-assisted and model-dependent. Since frequentist's definition of probability carries repeated sampling properties, such design-based inference automatically takes account of survey features like survey weighting to reduce sample selection bias *(Rao, 2011)*. In the frequentist model-assisted approach, there are analytical methods such as mean squared error, which is essentially the basis of the repeated sampling approach, to ensure the results' reasonability *(Rao, 2011).* Lastly, the frequentist model-dependent approach focuses on point estimation, variance estimation and associated normal theory confidence intervals, rather than distributional assumptions, which is the same case as the design-based inference mentioned above *(Rao, 2011)*. However, Bayesian approaches are fundamentally model-based and deny the role of randomization for design, their inclusion of distributional assumptions (see Inferences 2) and random effects further put them into a disadvantaged position.

Sources/Example:

- *"Impact of Frequentist and Bayesian Methods on Survey Sampling Practice: A Selective Appraisal" (Rao, 2011)* ---> This paper discusses in detail about different approaches to finite population inferences in mathematical terms and how survey features are incorporated to frequentist approach.

<u>Statement: Bayesian approach is better than Frequentist approach.</u>

Inferences:

1) **Frequentist methods are vulnerable to producing unreliable results while Bayesian methods take advantage of priors to incorporate all data and beliefs to constrain results within reasonable bounds.**

   Due to inadequate understanding and stipulated recipe-like implementation of statistical modelling, frequentist testing often contributes significantly to untrustworthy results and is argued to be indirectly related to replication crisis *(Spanos, 2022).* On the contrary, several literature and research projects have pointed out the importance of prior knowledge in drawing reasonable conclusions and facilitating learning of causal relationships between variables *(Hatswell, 2019)*. The additional prior distributions required in Bayesian methods are explicitly stated and their influences can also be studied through sensitivity analyses separately *(Chan, 2012)*.

Sources/Example:

- *"Frequentist Model-based Statistical Induction and the Replication Crisis." (Spanos, 2022)*
- *"Frequentist and Bayesian meta-regression of health state utilities for multiple myeloma incorporating systematic review and analysis of individual patient data" (Hatswell, 2019)*
- *"Sensitivity analysis in Bayesian networks: From single to multiple parameters." (Chan, 2012)*

2) **Frequentist approach does not provide sufficient answers.**

Frequentist absorbs information solely from the data and gives answers with exact frequentist properties like 95% confidence interval covering the true value 95% of the time in a repeated sampling setting. Therefore, in exact finite-sample settings, frequentist solutions are very limited *(Little, 2006)*. A famous example is the Behrens-Fisher problem: two independent normally distributed populations with same means but distinct variances. The frequentist approach fails to produce an efficient confidence interval to account for the difference in means with exact confidence coverage (some approximations are available, but they are not optimal) *(Kim, 1998)*. The lack of component in measuring uncertainty for variance parameters is a root problem of Frequentist approach. Fortunately, Bayesian procedures give posterior distributions to all parameters (including variance) which solves this problem *(Little, 2006; Kim, 1998)*.

Sources/Example:

- *"Calibrated Bayes: A Bayes/Frequentist Roadmap" (Little, 2006)* ---> It provides a brief view of the limited ability of frequentist methods.
- *"On the Behrens-Fisher Problem: A Review" (Kim, 1998)* ---> It summarizes solutions from both approaches to the Behrens-Fisher problem and highlights the effectiveness of Bayesian approach towards finding a solution.

3) **The Bayesian approach can work well in small sample size setting while Frequentist approach may yield to severe bias without sufficient data.**

The Bayesian approach can perform well regardless of the sample size since it takes in both the data and additional prior distributions. In small sample size scenarios, it can proceed not to rely heavily on the sample, but on prior specification instead (Lee, 2004). This is particularly useful in clinical trial research of which the sample sizes are usually small since the number of patients/volunteers willing to participate in the research are very limited. However, researchers may already have a reliable sense or expectation about how the results will perform before performing the analysis. In contrast, the frequentist approach emphasizes the sample data. When the sample size is small, skewed and biased data plays a significant role in frequentist methods, consequently producing severely biased results (Lee, 2004).

Sources/Example:

- *"Evaluation of the Bayesian and Maximum Likelihood Approaches in Analyzing Structural Equation Models with Small Sample Sizes" (Lee, 2004)*
- *"Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors" (Rans, 2015)*

---> This research paper adopts both the Frequentist approach and the Bayesian approach to analyze a small sample in a longitudinal study. Maximum likelihood estimation (a frequentist method) showed insufficient coverage (I.e., biased parameter values) as well as power with very small samples. However, in conjunction with informative priors, Bayesian analysis statistical power increased to acceptable levels.

A note on the sources referenced in this assignment

All the references listed are peer-reviewed research articles. They can be found on the University of Toronto Libraries search engine. Some key strings I used during my search include: “Frequentist vs Bayesian”, “Bayesian subjective”, “Frequentist Bayesian sample size”, “Frequentist Bayesian strength weakness”. After the initial search into papers, I referred to the reference section of the papers to look deeper into specific topics and were able to find more grounded evidence/examples than a general description.

Reflection

I came across the concepts of Bayesian statistics and Frequentist statistics in my first STA355 (Theory of Statistical Practice) lecture. However, the lecturer only stated that we would focus more on the frequentist approach going forward in the course but skipped the explanation of the differences between the two main streams. Although Bayesian statistical methods have been widely used in multidisciplinary research recently, the course's decision to emphasize frequentist theories, which sounds like the course sided with frequentists, provokes thoughts on the reasons behind them. Since then, I have grown interested in learning more about this controversial topic.

Surprisingly, this topic is more complicated than I initially thought. It was a bit of a struggle for me to come up with strong and definite inferences (perhaps because of the theoretical/philosophical nature of the topic). To avoid creating confirmation bias while developing the inferences, I spent a significant amount of time doing relatively deep background research into these two approaches and understanding their fundamental differences from theoretical and mechanical perspectives. Additionally, I was cautious not to immerse myself in the works that have shown stances in either one of the approaches. Combined with the statistical knowledge I possess; I proposed the inferences that are essentially my doubts towards the approaches' principles (such as sample size requirements, violation of important research principles, and difficulties in prior specification).

After establishing my inferences, I searched for peer-reviewed articles as evidence to support my inferences. To suggest well-rounded arguments, I put in extra effort to widen the scope of evidence, ranging from theoretical arguments to practical research projects that incorporated the Frequentist/Bayesian methods which have reflected their relative performance. While reading the articles, I have developed critical thinking skills in determining the strengths of evidence. Due to the theoretical nature of the topic, many literatures involve logical deductions. To prevent referencing implausible sources, I held a relatively conservative attitude while reading them and always questioned myself to make sure the logical flows in the papers are valid. One example is the controversial Likelihood Principle (LP) argument. In the beginning, I was skeptical that the Frequentist approach violates LP using my (in fact shallow) understanding of the theorem from STA261 (Probabilities and Statistics II). As I read more literature that supports or defends the argument, I realized that there are different interpretations towards the extent of violation which involve more complicated concepts (unfortunately, they are beyond my level of understanding due to my limited mathematical background). Therefore, I decided not to recklessly put the LP argument as one of the inferences.

This assignment creates an opportunity for me to explore a statement from both sides objectively. I did not have an initial stance before working on this assignment and I remain against the idea of “picking a side absolutely” after all. Although the inferences I proposed are supported with evidence from credible scientific articles, I recognized that arguments on one side may as well be effective counterarguments against the other side. As shown in many sources’ conclusions, the choice of approach ultimately really depends on the data available and the desired focus of the results.

Moreover, several modern Bayesian-Frequentist mixed methods have been proposed in the papers that make use of the advantages of the two approaches to counteract their weaknesses. Therefore, instead of being persuaded to stand with a particular side, I am inspired to consider staying in a neutral position and embrace both approaches to develop merged methods that can maximize the effectiveness and efficiency of the statistical methods branched from either side of the approaches.

References

1. Chan, H., & Darwiche, A. (2012). *Sensitivity Analysis in Bayesian Networks: From Single to Multiple Parameters*. https://doi.org/10.48550/arxiv.1207.4124
2. Efron, B. (2005). Bayesians, Frequentists, and Scientists. *Journal of the American Statistical Association*, *100*(469), 1–5. https://doi.org/10.1198/016214505000000033
3. Flori, A. (2019). News and subjective beliefs: A Bayesian approach to Bitcoin investments. *Research in International Business and Finance*, *50*, 336–356. https://doi.org/10.1016/j.ribaf.2019.05.007
4. Gelman, A. (2008). Objections to Bayesian statistics. *Bayesian Anal*, *3*(3), 445–449. https://doi.org/10.1214/08-BA318
5. Hatswell, A. J., Burns, D., Baio, G., & Wadelin, F. (2019). Frequentist and Bayesian meta-regression of health state utilities for multiple myeloma incorporating systematic review and analysis of individual patient data. *Health Economics*, *28*(5), 653–665. https://doi.org/10.1002/hec.3871
6. Kim, S.-H., & Cohen, A. S. (1998). On the Behrens-Fisher Problem: A Review. *Journal of Educational and Behavioral Statistics*, *23*(4), 356–377. https://doi.org/10.2307/1165281
7. Koo, H., Keeley, R. E., Shafieloo, A., & L'Huillier, B. (2022). Bayesian vs frequentist: Comparing Bayesian model selection with a frequentist approach using the iterative smoothing method. *Journal of Cosmology and Astroparticle Physics*, *2022*(3), 47-. https://doi.org/10.1088/1475-7516/2022/03/047
8. Lecoutre, B., & Poitevineau, J. (2014). Preamble—Frequentist and Bayesian Inference. In *The Significance Test Controversy Revisited* (pp. 9–20). Springer Berlin / Heidelberg. https://doi.org/10.1007/978-3-662-44046-9_2
9. Lee, S.-Y., & Song, X.-Y. (2004). Evaluation of the Bayesian and Maximum Likelihood Approaches in Analyzing Structural Equation Models with Small Sample Sizes. *Multivariate Behavioral Research*, *39*(4), 653–686. https://doi.org/10.1207/s15327906mbr3904_4
10. Little, R. J. (2006). Calibrated Bayes: A Bayes/Frequentist Roadmap. *The American Statistician*, *60*(3), 213–223. https://doi.org/10.1198/000313006X117837
11. Rao, J. N. K., Meeden, G., Sedransk, J., & Slud, E. (2011). Impact of Frequentist and Bayesian Methods on Survey Sampling Practice: A Selective Appraisal/Discussion of "Impact of Frequentist and Bayesian Methods on Survey Sampling Practice: A Selective Appraisal" by J. N. K. Rao/Discussion of "Impact of Frequentist and Bayesian Methods on Survey Sampling Practice: A Selective Appraisal" by J. N. K. Rao/Discussion of "Impact of Frequentist and Bayesian Methods on Survey Sampling Practice: A Selective Appraisal" by J. N. K. Rao/Rejoinder. *Statistical Science*, *26*(2), 240-.
12. Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020). Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review. *Structural Equation Modeling*, *27*(1), 131–161. https://doi.org/10.1080/10705511.2019.1577140
13. Spanos, A. (2022). Frequentist Model-based Statistical Induction and the Replication Crisis. *Journal of Quantitative Economics : Journal of the Indian Econometric Society*, *20*(Suppl 1), 133–159. https://doi.org/10.1007/s40953-022-00312-z
14. Rans van de Schoot, Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., & van Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: The case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal of Psychotraumatology*, *6*(1), 25216–25216. https://doi.org/10.3402/ejpt.v6.25216