

ST440/540 – Mid-term exam 1 - Due February 26

THIS IS AN EXAM - DO NOT DISCUSS THE PROBLEM WITH ANYONE (INCLUDING OTHER STUDENTS OR THE TA)! If you have questions, please email me.

Developing a COVID-19 vaccine is one of the most important scientific endeavors our of lifetime. In this exam, you will conduct a Bayesian analysis of the data produced by the clinical trial of the Moderna vaccine¹. The data² are:

	Placebo group		Vaccine group	
	Infected	Participants	Infected	Participants
All patients	185	14,073	11	14,134
White	144	8,916	10	9,023
Communities of color	41	5,132	1	5,088

Let $\theta_g \in (0, 1)$ be the infection probability in group $g \in \{\text{placebo, vaccine}\}$. That is, a randomly-selected person given the placebo has probability θ_{placebo} of contracting the virus, and a randomly-selected person given the vaccine has probability θ_{vaccine} of contracting the virus. The efficacy of the vaccine is often defined as

$$E = 1 - \frac{\theta_{\text{vaccine}}}{\theta_{\text{placebo}}}.$$

Using these data, answer the following questions:

1. **Placebo analysis:** Using all 14,073 patients in the placebo group, state a reasonable likelihood for the data and an uninformative conjugate prior distribution for θ_{placebo} . What are the main assumptions in your model and do you think they are valid?
2. **Derivation:** Using the model from the previous question, derive (i.e., a 3-4 line mathematical derivation) the posterior distribution of θ_{placebo} . Your final answer should be a distribution, e.g., $\theta_{\text{placebo}}|Y \sim \text{Poisson}(Y + 1)$, and you must show your work.
3. **Efficacy analysis:** Using all 14,073+14,134 patients, perform a Bayesian analysis of the efficacy of the vaccine, E . Are the results sensitive to the prior (i.e., do they change a lot when you change the prior)? Be sure to give all details such as likelihoods, priors and a clear description of the results you report. Summarize the results in a clearly-labeled plot and a table.
4. **Hypothesis test:** According to the protocol² (Page 14), the stated objective was to test the hypothesis that the vaccine has efficacy at least 0.70. Carry out a Bayesian test of this hypothesis. Are the results sensitive to your the prior?

¹<https://www.modernatx.com/sites/default/files/mRNA-1273-P301-Protocol.pdf>

²Taken from Figure 4 of <https://www.nejm.org/doi/full/10.1056/NEJMoa2035389>

5. **Subgroup analysis:** Carry out the Bayesian analysis separately for white subjects and subjects from communities of color, and test whether there is a difference in the efficacy across these groups. Summarize the results in a clearly-labeled plot and a table.

Your paper should be written as a professional document with full sentences and paragraphs, clearly labeled and numbered figures and tables, and few spelling/grammar errors. Organize your report with five sections, one for each question. You should include enough detail that another student in class could reproduce your results. Summarize your analysis in a PDF document that is no more than four pages long (excluding code). Append your code to the end of this document and submit a single document on moodle.

HAVE FUN!