Homework 6: Frequentists vs Bayesians

Instructor: Daniel L. Pimentel-Alarcón Due 03/20/24

DO NOT POLLUTE! AVOID PRINTING, OR PRINT 2-SIDED MULTIPAGE.

In this homework you will compare the *maximum likelihood estimator* (MLE; frequentist) and the *maximum a posteriori* estimator (MAP; Bayesian) to decide once and for all in which cases each is better.

To do so, consider the setup in Example 10.6 in the lecture notes (Topic 10). That is, suppose scientists at a big pharmaceutical company have designed a COVID-19 vaccine, and want to estimate its probability of success $\mathrm{p}^\star$. To this end they will conduct a clinical trial where they will test their treatment on N individuals, and record whether they react favorably. This can be modeled as

$$x_1, \ldots, x_\mathrm{N} \overset{iid}{\sim} \text{Bernoulli}(\mathrm{p}^\star),$$

and the goal is to estimate $\mathrm{p}^\star$. Each subproblem is worth 30 points.

Problem 6.1. Frequentist (MLE). The first option is to estimate $\mathrm{p}^\star$ with a maximum likelihood approach. Derive an expression for the MLE of $\mathrm{p}^\star$.

Problem 6.2. Bayesian (MAP). The second option is to use a Bayesian approach, using *prior* information about this treatment, which has shown to be very effective in vitro and on other animals, like mice, rabbits, and chimpanzees. Based on these related experiments, we expect $\mathrm{p}^\star$ to be closer to 1 than to 0. Hence, we can model the *prior* of $\mathbb{P}(\mathrm{p})$ as $\text{Beta}(\alpha, \beta)$ with parameters $\alpha > \beta > 1$, so that the density is skewed towards 1 (see Figure 10.1 in the lecture notes (Topic 10) for some intuition). Based on this information, derive an expression for the MAP estimator of $\mathrm{p}^\star$.

Problem 6.3. (Correct prior). Suppose $\mathrm{p}^\star = 0.99$. In expectation:

(a) How many samples N do you need for $\hat{\mathrm{p}}_{MLE}$ to be within 0.01 of $\mathrm{p}^\star$?

(b) How many samples N do you need for $\hat{\mathrm{p}}_{MAP}$ to be within 0.01 of $\mathrm{p}^\star$?

(You may answer these questions analytically, numerically, or with a reference.)

Problem 6.4. (Incorrect prior). Suppose $\mathrm{p}^\star = 0.01$. In expectation:

(a) How many samples N do you need for $\hat{\mathrm{p}}_{MLE}$ to be within 0.01 of $\mathrm{p}^\star$?

(b) How many samples N do you need for $\hat{\mathrm{p}}_{MAP}$ to be within 0.01 of $\mathrm{p}^\star$?

(You may answer these questions analytically, numerically, or with a reference.)

Problem 6.5. Based on your results, which do you prefer, $\hat{\mathrm{p}}_{MLE}$, or $\hat{\mathrm{p}}_{MAP}$, and why? In which scenarios would you use one or the other? Discuss advantages and disadvantages of each.