Exercise Sheet #6

due on Tuesday, 15 December 2020, 12 noon sharp

1. **EXAMple** Of all US states, Arizona saw the closest outcome in the recent US presidential elections: Joseph R. Biden Jr. received a total of 1 672 143 votes, while Donald J. Trump received 1 661 686 votes.[1] The two numbers are so close that their difference may seem insignificant. Is it? Consider the null hypothesis that the votes were drawn iid. with probability $f = 0.5$ for either option.

   (a) Under the null hypothesis, what is the probability for exactly this outcome?

   (b) Write down a formula for a hypothesis test to construct a p-value for this outcome under the null hypothesis. Using a computer or pocket calculator and the result of (a), can you find an upper bound for it?

2. **Theory Question:** Consider data $\boldsymbol{x} := [x_1, \dots, x_n]$ drawn iid. from $p(\boldsymbol{x} \mid \theta^*) = \prod_{i=1}^n p(x_i \mid \theta^*)$ with an unknown $\theta^* \in \mathbb{R}$. Lecture 5 introduced the notion, for general $\theta \in \mathbb{R}$, of the score function

   $$s(\boldsymbol{x};\theta) := \frac{\partial \log p(\boldsymbol{x} \mid \theta)}{\partial \theta}$$

   and the *Fisher information*

   $$I(\theta) = \operatorname{var}_{p(\boldsymbol{x}\mid\theta)}(s(\boldsymbol{x};\theta)).$$

   Show the following properties (introduced without proof in the lecture), which establish that the Fisher information is not just the variance of the score function, but also the expected curvature of the log-likelihood at θ. (These statements hold for all values of θ where the necessary quantities are defined. Note that all the expectations are over $p(\boldsymbol{x} \mid \theta)$).

   (a) $\mathbb{E}_{p(\boldsymbol{x}\mid\theta)}(s(\boldsymbol{x};\theta)) = 0$ (Thus we also have $\operatorname{var}_{p(\boldsymbol{x}\mid\theta)}(s(\boldsymbol{x},\theta)) = \mathbb{E}_{p(\boldsymbol{x}\mid\theta)}(s^2(\boldsymbol{x};\theta))$)

   (b) $I(\theta) = -\mathbb{E}_{p(\boldsymbol{x}\mid\theta)}\left(\frac{\partial^2 \log p(\boldsymbol{x}\mid\theta)}{\partial \theta^2}\right)$

   Hints for (a) and (b): You can assume that you are allowed to *differentiate under the integral*, i.e.

   $$\int \frac{\partial p(\boldsymbol{x} \mid \theta)}{\partial \theta} \mathrm{d}\boldsymbol{x} = \frac{\partial}{\partial \theta} \int p(\boldsymbol{x} \mid \theta)\, \mathrm{d}\boldsymbol{x} \quad \text{and}$$

   $$\int \frac{\partial^2 p(\boldsymbol{x} \mid \theta)}{\partial \theta^2} \mathrm{d}\boldsymbol{x} = \frac{\partial^2}{\partial \theta^2} \int p(\boldsymbol{x} \mid \theta)\, \mathrm{d}\boldsymbol{x}.$$

3. **Practical Question:** You can find this week's practical exercise in `Exercise_06.ipynb`

[1] preliminiary results as reported by the NY Times. For the purposes of this question, we will ignore the 53 497 votes cast for other candidates, noting in passing that this number is indeed larger than the difference between the votes for the two major party candidates.