Exercise Sheet #4

due on Monday, 22 November 2021, 10am sharp

1. **EXAMple** On 23 November 2020, the British pharma company AstraZeneca published a press release about their vaccine AZD1222 (which has since been licensed under the trademark Vaxzervia). It contains the quote

   > One dosing regimen ($n = 2,741$) showed vaccine efficacy of 90% when AZD1222 was given as a half dose, followed by a full dose at least one month apart, and another dosing regimen ($n = 8,895$) showed 62% efficacy when given as two full doses at least one month apart. The combined analysis from both dosing regimens ($n = 11,636$) resulted in an average efficacy of 70%. All results were statistically significant ($p \leq 0.0001$).

   It may seem puzzling that the *higher* dosage seemingly had a *weaker* effect (the PI, Andrew Pollard, spun this as a good thing: "Excitingly, weve found that one of our dosing regimens may be around 90% effective and if this dosing regime is used, more people could be vaccinated with planned vaccine supply"). Your task in this exercise is to read the press release closely (click on link above) and come up with a hypothesis for how, despite this being a randomized controlled trial, such a result *could* have arisen by selection bias, in a way that is *not* related to the dosage. This gives an example for how subtle such effects can be, and why designing a reliable randomized controlled trial is hard. If the link should stop working, you can also find the press release as a pdf on Ilias.

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 sheet on Ilias as `Exercise_04.ipynb`