STA 365: Homework 1

Professor Boris Babic

Due Date: Friday, February 4, 2022 (11:59pm via Quercus)

Instructions. You are encouraged to type out your answers in LaTeX or a word processor. If you need to handwrite your responses, make sure that they are clear and legible, and that you scan a high quality image (any ambiguity may be resolved against you).

While you are permitted to work on these problems in a study group, you must ultimately complete the problems on your own and submit the assignment individually. Do not copy your colleagues' work.

Where a problem requires the use of R, produce your associated R code. While you may use other software, solutions will be provided only in R.

Problem 1. (15 points)

(a) Provide an example of a sequence of random variables, $X_1, ..., X_n$, which are exchangeable but not iid.

(b) Let $X_1, ...X_n$ be a sequence of continuous exchangeable random variables. Prove that $X_1, ...X_n$ are identically distributed.

Problem 2. (20 points)

You are tasked with predicting the occurrence of a binary event E. That is, $\Omega = E \cup \overline{E}$, and $E \cap \overline{E} = \emptyset$. Let p and $1 - p$ denote your true probabilities for E and $\overline{E}$, respectively. Let r and $1 - r$ denote your reported probabilities for E and $\overline{E}$, respectively. And suppose that your payoff is given by

$$s(r) = \begin{cases} -\log(r), & E \\ -\log(1-r), & \overline{E} \end{cases}$$

(a) Recall that $s(r)$ is strictly proper if $\mathrm{E}_p s(p) > \mathrm{E}_p s(r)\ \forall r \neq p$. Prove that $s(r)$ is (or is not) strictly proper.

(b) Let $r = p$ and plot $\mathrm{E}_p[s(p)]$ (using a base 2 log) in R. This function is also known as the Shannon information entropy corresponding to the Bernoulli random variable X where E corresponds to $X = 1$ and $\overline{E}$ corresponds to $X = 0$.

(c) If you want to pick a prior probability for E, i.e., $p(E)$, that maximizes Shannon information entropy, which value of p would you select? Hint: find $\arg\max_{p \in [0,1]} \mathrm{E}_p[s(p)]$.

(d) Argue in one paragraph why (or why not) using the measure of Shannon information entropy to select a prior probability is a good/bad idea.

Problem 3. (15 points)

Consider two coins C_1 and C_2, with the following characteristics: $\Pr(\text{heads}|C_1) = 0.7$ and $\Pr(\text{heads}|C_2) = 0.3$. Choose one of the coins at random and imagine spinning it repeatedly. Given that the first two spins from the chosen coin are tails, what is the expectation of the number of additional spins until a head shows up?

Problem 4. (20 points)

Suppose we are going to sample 100 individuals from a county (of size much larger than 100) and ask each sampled person whether they support policy Z or not. Let $Y_i = 1$ if person i in the sample supports the policy, and $Y_i = 0$ otherwise.

(a) Assume $Y_1, ...Y_{100}|\theta \overset{\text{iid}}{\sim} \text{Bernoulli}(\theta)$. Write down the joint distribution of $\Pr(Y_1 = y_1, ...Y_{100} = y_{100}|\theta)$ in a compact form.

(b) Let $X = \sum_i Y_i$. Write down the distribution of X in a compact form.

(c) Suppose that the results of the survey are $X = 66$. Suppose that $\theta \in \{0.0, 0.1, ...0.9, 1.0\}$. Compute $\Pr(X = 66|\theta)$ (the likelihood) for each of these 11 values of θ and plot these probabilities as a function of θ.

(d) Now suppose $\theta \in [0, 1]$. Using the uniform prior density for θ so that $\pi(\theta) = 1$, write down the posterior density $\pi(\theta|X = 66)$.

Problem 5. (30 points)

Estimate the probability θ of Covid-19 reinfections based on a study in which there were $n = 43$ previously infected individuals and $y = 15$ re-infected individuals within 36 months.

(a) Using a beta(2, 8) prior for θ, write down $\pi(\theta)$, and $\pi(y|\theta)$, and derive $\pi(\theta|y)$.

(b) Find the posterior mean, mode and standard deviation for θ.

(c) Find a 95% credible interval for θ.

(d) Plot $\pi(\theta)$, $\pi(y|\theta)$, and $\pi(\theta|y)$ as functions of θ.

(e) Repeat a-d using a beta(8, 2) prior for θ.

(f) Consider the following prior distribution for θ:

$$\pi(\theta) = \frac{1}{4}\frac{\Gamma(10)}{\Gamma(2)\Gamma(8)}[3\theta(1-\theta)^7 + \theta^7(1-\theta)],$$

which is a 75-25% mixture of a beta(2, 8) and a beta(8, 2) prior distribution.

(i) Given this prior, write down $\pi(\theta)\pi(y|\theta)$ and simplify as much as possible.

(ii) The posterior distribution is a mixture of two distributions. Identify these distributions.