Winter 2019/20 TERM 1 EPSE 581C
Due: Friday, Dec. 20th

- Please make sure you write your answers to these questions in your own words. Even if you work with a group to formulate your responses, do not just copy someone else's sentences/words.
- All problem data are available online in .csv format; R code is also available.

Question 1: (Bayesianism as good science). These questions can be answered succinctly either in words or mathematically: your choice!

(a) Explain why the adage *"Extraordinary claims demand extraordinary evidence"* reflects a distinctly Bayesian view of inference, and also why the frequentist (i.e. p-value based) approach might contradict this adage.

(b) Explain why a significant p-value (say, < 0.05) is *not* sufficient to conclude that the probability that the null hypothesis is true is small.

(c) Classically, we require an observed effect to be *replicable* if it is to be considered an accurate reflection of reality. What effect does replication of an observed effect (e.g. one observes nearly the same average outcome 10 samples in a row) have on the shape of a posterior distribution that is functionally related to the observed effect?

Question 2: (Statistical elicitation). Read through the *JASA* paper "Statistical methods for eliciting probability distributions" (2005) by Garthwaite, Kadane, & O'Hagan. Answer the following questions in one or two *sentences* each.

(a) What are the four main steps of the elicitation process?

(b) What is the major problem with the *judgement-by-representativeness* heuristic to elicit probability statements? What can be done to (partially) combat this?

(c) Contrast the elicitation phenonema of *conservatism* with the *"law of small numbers"*.

(d) Think about the mathematical definitions of *mean* and *median*. In general, which quantity is easier to elicit from individuals and why?

(e) In general, is it easier to elicit probability statements about a *joint distribution* of parameters, or about their *conditional and marginal distributions*? Why?

(f) What is one advantage and one disadvantage to using elicitation techniques to fit a *parametric* distribution (as opposed to a *nonparametric* distribution) to a parameter's prior?

(g) In general, what is/are the most difficult (hyper)-parameters of a regression model for which to elicit prior probability distributions? Why?

(h) What are the most common tools for assessing the adequacy of elicitation process?

(i) What are some major flaws with *opinion pooling*? What are some alternatives? Which of these sythensis methods do you think would be most feasible for *your own* particular area of research and why?

Question 3: (Regression modelling). For this exercise, make sure you have the 'lattice' and 'rethinking' packages loaded into your R installation.

Load the 'seabirdData' dataset from the webpage. This dataset contains information on 6 variables for 120 nesting female seabirds that were banded off the coast of southern California: **age** of the seabird (from date of banding), **cortizol** level of seabird at sampling, an estimate of the total mass of **fish** in the bird's diet (estimated by pellet contents), an estimate of the total mass of **plastic** in the bird's diet, an estimate of **reproductive success** for the bird from previous breeding seasons (0 = no success; 3 = total success), and a standardized measure of androgen vs. **estrogen** expression of the seabird at sampling. We are interested in building a model to understand the relationship between **estrogen** response and the other variables.

(a) Many people start a model building problem like this by first examining the model consisting of the sum of all first-order explanatory variables:

$$Est = \beta_0 + \beta_{rep}Rep + \beta_{plas}Plastic + \beta_{cort}Cortizol + \beta_{age}Age + \beta_{fish}Fish + \varepsilon$$

Use the R code on the webpage to fit this model using Bayesian estimation, and record its AIC and BIC. Also examine the marginal posteriors and indicate which variables (if any) do *not* appear to explain much variation in the response. [Note: if you get an error message about "non-finite finite-difference value" or "start values for parameters may be too far from MAP," try rerunning the code. The way the 'rethinking' package is encoded makes the MCMC approximation to the posterior distributions finicky.]

(b) It is then common practice to examine simpler models where one or more of the explanatory variables have been removed. Pick 2 or 3 such models, fit them, and compute their AIC and BIC values (adapt the code from above). According to the AIC and BIC values, which model is "best" between the original model and the ones you have just considered?

(c) Examine a plot of the residuals vs. fitted values for the full first-order model of part (a). If there is evidence of model misfit, find which variable seems to be causing it by plotting the residuals vs. the individual predictors (code is on the webpage).

(d) Propose a new model to correct for the misfit you found in part (c). FIt this model by adapting the code from part (a). Record the new model's AIC and BIC, note if these values have *decreased* from the first model (indicating better fit), and note which predictors (if any) now appear to *not* explain much variation in the response.

(e) Examine a plot of your new model's residuals vs. fitted values. Is there still evidence of misfit, and if so, is the misfit as bad as before? Is more variance explained by this new model? [An easy way to check this is to compare the magnitudes of the residuals in this model to the previous one: less residual error means more variation explained by the model.]

(f) Investigate the residuals vs. individual predictors plots for your new model. Do you find any suggestions as to how to improve your model fit? Recall that we look for patterns in the residuals, notably curves, "pinching," uneven vertical spread, etc. Propose a refined model, fit it, compute its AIC and BIC, and examine its residuals. In your opinion, is this newest model superior to the model you proposed in part (d)? Why or why not?