Homework #7

Due: March 31, 2017

Vladimir Putin (1952 - present), current Russian President (he really said this?):

> History proves that all dictatorships, all authoritarian forms of government are transient. Only democratic systems are not transient. Whatever the shortcomings, mankind has not devised anything superior.

1. Consider the posterior $p(p|y_1, ..., y_n)$ =Beta$(\alpha^\star, \beta^\star)$ for $\alpha > 1$ and $\beta > 1$. Show that the MAP estimator is $\hat{p}_{MAP} = \frac{\alpha^\star - 1}{\alpha^\star + \beta^\star - 2}$. Be sure to perform either the 1st or 2nd derivative test. EXTRA CREDIT: Show (a) that if either $\alpha < 1$ or $\beta < 1$, then $\frac{\alpha^\star - 1}{\alpha^\star + \beta^\star - 2}$ may actually be a minimum; (b) if both $\alpha < 1$ <u>and</u> $\beta < 1$, then $\frac{\alpha^\star - 1}{\alpha^\star + \beta^\star - 2}$ minimizes the posterior and is not the MAP.

2. Consider a random sample $y_1, ..., y_n \sim$ Geometric(p) distribution as in Exercise 9.97. Assuming a noninformative prior for p, do the following:

   (a) Give the likelihood $p(y_1, ..., y_n|p)$.

   (b) Give the prior $p(p)$.

   (c) Find the posterior $p(p|y_1, ..., y_n)$.

   (d) Find the estimator $\hat{p}_B$, the mean of the posterior.

   (e) Find the MAP estimator $\hat{p}_{MAP}$.

3. On each day, a machine is used produce plastic crappets. The probability of machine failure on any given day is p. You will use your analysis from #2 to complete this problem. Data was collected from 10 randomly chosen crappet machines, and the number of days to failure was recorded for each machine:

   $$\{y_i\} = \{362, 51, 200, 511, 211, 420, 299, 280, 398, 323\}.$$

   Assuming an un-informative prior for p:

   (a) Give the posterior $p(p|y_1, ..., y_{10})$.

   (b) Give the estimator $\hat{p}_B$ for p.

   (c) Give the MAP estimator $\hat{p}_{MAP}$ for p.

   (d) Give the MLE of p.

   (e) Give a 95% credible interval for p. Use R code as given in the examples in the course notes.

   (f) Give a proper conclusion in terms of the problem.

4. Consider the likelihood $p(y_1, ..., y_n|p)$ for a SRS of Bernoulli data, and a prior $p(p)$=Beta(α, β). We showed in class that the posterior is $p(p|y_1, ..., y_n)$ =Beta$(\alpha^\star = \sum y_i + \alpha, \beta^\star = n - \sum y_i + \beta)$.

   (a) Show that the mean of the posterior is $\hat{p}_B = \frac{\sum y_i + \alpha}{n + \alpha + \beta}$.

   (b) Find $E(\hat{p}_B)$.

   (c) Is $\hat{p}_B$ biased?

   (d) Find $Var(\hat{p}_B)$.

(e) Which is larger, $Var(\hat{p}_B)$ or $Var(\hat{p}_{MLE})$, where $\hat{p}_{MLE}$ is the MVUE for p?

(f) Show that $\hat{p}_B$ is consistent for p.

5. Do exercise 8.56 using the Bayesian analysis from #4. Assume that the data are a SRS from a Bernoulli distribution, and use a non-informative prior for p.

   (a) Give $\hat{p}_{MLE}$, the MLE for p (use previous results, you do not need to derive it), and give the Bayesian estimate $\hat{p}_B$

   (b) Give the 98% CI for p.

   (c) Give a 98% credible interval for p. Use R or some other software package.

   (d) Interpret the credible interval in #5c in terms of the problem.

   (e) Do the evidence suggest that either a majority or minority of adults say that movies are getting better?

6. Consider a SRS $y_1, ..., y_n$ from $N(\mu, \sigma^2)$ when σ^2 is known, and assume an uninformative, flat prior for μ.

   (a) Show that $p(\mu|y_1, ..., y_n) = N(\bar{y}, \sigma^2/n)$.

   (b) Give the estimators $\hat{\mu}_B$, $\hat{\mu}_{MAP}$ and $\hat{\mu}_{MLE}$.

7. Let $y_1, ..., y_n$ denote a SRS from a Poisson(λ) distribution (as in Exercise 16.11).

   (a) Show that $p(\lambda)$=Gamma(α, β) is a conjugate prior. That is, show that $p(\lambda|y_1, ..., y_n)$ =Gamma$(\alpha^\star, \beta^\star)$ for some $\alpha^\star$ and $\beta^\star$.

   (b) Give the posterior parameters $\alpha^\star$ and $\beta^\star$.

   (c) Give the estimator $\hat{\lambda}_B$.

   (d) Use previous results to give the estimator $\hat{\lambda}_{MLE}$.

   (e) Find $E(\hat{\lambda}_B)$.

   (f) Is $\hat{\lambda}_B$ biased?

   (g) Find $Var(\hat{\lambda}_B)$.

   (h) Which is larger, $Var(\hat{\lambda}_B)$ or $Var(\hat{\lambda}_{MLE})$?

   (i) Show that $\hat{\lambda}_B$ is consistent for λ. Use the Generalized Weak Law of Large Numbers (similar to Theorem 9.1): If $\hat{\theta}_n$ is an estimator of θ with $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$ and $\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$, then $\hat{\theta}_n$ is consistent for θ, i.e. $\hat{\theta}_n$ converges in probability to θ.