ECON5529

Problem Set #1

Due Wednesday September 27th

Please type up your answers and email them along with your R code as a pdf file to scharfenakere@umkc.edu.

Analytical Exercises

1. Consider an urn containing N balls each labeled with a unique number $1, 2, ..., N$. M of these balls are colored red and the remaining $N - M$ are white, $0 \leq M \leq N$. Understand that these facts represents our state of knowledge, or information I. Let $R_i \equiv$ red ball on the i^{th} draw and $W_i \equiv$ white ball on the i^{th} draw. According to our knowledge of the composition of the urn, only red or white can be drawn, thus, for the i^{th} draw it must be that $p[R_i|I] + p[W_i|I] = 1$.
   a) Find the probability of drawing and red or white ball on the first draw, i.e. $p[R_1|I]$? and $P[W_1|I]$
   b) Find the probability of drawing a red ball on the first two draws, assuming you are sampling without replacement, i.e. find $p[R_1 R_2|I]$
   c) Following this logic find the probability for drawing a red ball on the first $m \leq M$ consecutive draws. Show your logic.
   d) Find the probability for drawing exactly $m \leq M$ red balls in $n \leq N$ draws, regardless of order.

2. A *sample* is a set of n numbers $x = x_1, x_2, ..., x_n$. The *sample mean* is the average of the sample, $m[x] = \frac{x_1 + ... + x_n}{n}$, the *sample variance* is the average squared deviation of the sample values from the sample mean $s[x]^2 = \frac{(x_1 - m[x])^2 + ... + (x_n - m[x])^2}{n}$, and the *sample standard deviation* is the square root of the sample variance.

a) Suppose we model the sample as a constant, μ. In general the likelihood that all the numbers in the sample will be to equal μ is zero. In order to allow for deviations, suppose that we assume that likelihood of the sample is proportional to $exp[-\frac{(x_1-\mu)^2+...+(x_n-\mu)^2}{n\sigma}]$, where σ is a second model parameter. Derive the expression for the posterior probability $p[\mu,\sigma|x]$, given a prior $p[\mu,\sigma]$.

b) *Jeffreys' prior* is $p[\mu,\sigma] = d\mu\frac{d\sigma}{\sigma}$. Letting $d\mu = d\sigma = 1$ write the posterior probability with Jeffreys' prior.

c) Find the maximum posterior probability model, $(\hat{\mu},\hat{\sigma})$ given Jeffreys' prior. (Hint: Use the first order conditions. You should get familiar looking results)

3. Sometimes Bayesian results are given as *posterior odds ratios*, which for two possible alternative hypothesis is expressed as: $odds[\theta_1,\theta_2] = \frac{p[\theta_1|D]}{p[\theta_2|D]}$. If the prior probabilities for θ_1 and θ_2 are the same, how can this be re-expressed using Bayes' Law?

4. Prove the gamma distribution $p[\mu|\alpha,\beta] = \frac{\beta^\alpha}{\Gamma[\alpha]}\mu^{\alpha-1}e^{-\beta\mu}, \quad \mu,\alpha,\beta > 0$ is the conjugate prior for μ in a Poisson likelihood function,

$$p[x|\mu] = \left(\prod_{i=1}^{n} x_i!\right)^{-1} e^{log[\mu]\sum_{i=1}^{n} x_i} e^{-n\mu} \tag{0.1}$$

that is, calculate the posterior distribution of μ and show that it is also gamma distributed. (Hint: work with the kernel of the distribution)

5. The multinomial model

$$p[x|\theta] = \frac{n!}{x_1!\cdots x_k!}\theta_1^{x_1}\cdots\theta_k^{x_k} = \frac{n!}{\prod_i^k x_i!}\prod_i^k \theta_i^{x_i} \tag{0.2}$$

is a generalization of the binomial model where the number of discrete outcomes is greater than two. Prove that the conjugate prior for the generalized binomial is the generalized beta distribution, which is also called the Dirichlet distribution and is of the form:

$$p[\theta|\alpha] = \frac{\Gamma[\sum_i \alpha_i]}{\prod_i \Gamma[\alpha_i]}\prod_{i=1}^k \theta_i^{\alpha_i-1} \tag{0.3}$$

6. Prove that the entropy of the Normal distribution is equal to $\frac{1}{2}Log[2\pi e\sigma^2]$.

R Exercises

1. A bank has made 100 mortgages of a new type (say it's 2005 and they are subprime mortgages), and all have been outstanding 5 years. Of these 100, 5 of them have defaulted. The bank would like to estimate the probability θ of default in the first five years for this type of mortgage, and get some idea of how much uncertainty there is about the probability, given the observed data. These being a new type of mortgage, the bank assigns a uniform prior over θ.

a) What is the likelihood $p[x|\theta]$?

b) Plot the likelihood in R and indicate on the plot (e.g. use the abline() function) the location of the maximum likelihood value of θ as well as the expected value of θ.

c) Using the quantile function $qbeta()$ calculate and indicate a symmetric 95% confidence interval (cut off 2.5% of the left and right tail). Does this look like a reasonable confidence interval?

d) The shortest interval with 95% probability will have the likelihood the same height at each end. Using the package "TeachingDemos" use the HPD function to find the shortest interval.

e) Compare graphically the 95% confidence interval and the 95% HPD interval. Is the HPD the shortest interval?

2. Imagine an urn filled with 100 balls of unknown composition. You are told that the contents consist of red and green balls of any possible combination. Consider two priors over the distribution of red and green balls which we denoted $p[\{r,g\}]$:

- A) a uniform prior over all possible combinations of balls, which assumes that each pair of numbers of red and green balls, $\{0,100\},\{1,99\},\{2,98\}...,\{100,0\}$ is equally likely to occur in the string.
- B) a prior that each complete data string is equally likely. That is, we assign an equal probability to each possible combination of a complete draw of all balls from the urn. If we denote a draw of a red ball 1 and a draw of a green ball 0 there are 2^{100} possibilities where each data string is of the form:
$\{1,0,0,...,0\},\{1,1,0,...,0\},...,\{0,0,0,...,1\}$

Note: Since $100 - r = g$, where g is the number of green balls, it suffices to determine the probabilities p_r, as it follows that $p_g = 1 - p_r$.

a) What probability does prior A assign to each composition of the urn? Plot this probability mass function in R.

b) What prior probability does prior B assign to each of the possible data strings? Plot this probability mass function in R. (Hint: use the Binomial coefficient and either a *for loop* or the function $sapply()$).

c) Is either of these priors informative? Would you favor either one of them? Is either or both exchangeable?

d) Say we draw 10 balls from the urn with replacement and observe 7 red balls and 3 green balls. If $0 \le \theta \le 1$ is the unknown parameter representing the composition of the urn (as in the Binomial model) what is the likelihood function for this data? Plot the likelihood in R.

e) Plot the normalized posterior distribution in R for each prior. What can you say about posterior inference for the composition of the urn for each prior?

f) Calculate the max posterior from having seen 7 red balls and 3 green balls for each prior.