36-463/663: Hierarchical Linear Models
Fall 2016
HW08 – Due Thu 03 Nov 2016

Announcements

- Homework due as a pdf on Blackboard as usual.
- Reading:
  - Please make sure you are familiar with the material in G&H, Chapters 12 & 13.
  - Chapters 2 and 3 of Lynch provide a gentle review/introduction to mathematical statistics, maximum likelihood and Bayesian inference. has a quick, gentle introduction to Bayesian inference.
  - G&H, Ch 18 gives a more brisk introduction, oriented toward hierarchical and multilevel modeling.

Exercises

1. *MLE for Poisson data.* In this problem, I want you to derive the maximum likelihood estimator (MLE) and its standard error, using ideas like those presented in class. Suppose the data $x_1, x_2, \ldots, x_n$ is iid Poisson with parameter λ; then

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{x_1!x_2!\cdots x_n!} = \frac{e^{-n\lambda}\lambda^{n\bar{x}}}{x_1!x_2!\cdots x_n!} \qquad (*)$$

   (a) Maximize the log-likelihood to find the MLE $\hat{\lambda}$. *Hints: Take the logarithm of $L(\lambda)$ in $(*)$, simplify, differentiate, set equal to zero, solve for λ.*

   (b) Now find $SE(\hat{\lambda})$, as one over the square root of the Fisher information $I(\lambda)$ (we used observed information in class; if you prefer to use expected information, that is fine, just say that that is what you are doing in your solutions).

2. Find and plot the posterior distribution for a binomial likelihood with $x = 5$ successes out of $n = 10$ trials, using at least three different beta prior distributions. Make appropriate graphs of the posterior distributions (densities) you find. Does the prior make a large difference in the outcome? When? *Hints:*

   - The *likelihood* is going to be $L(p) = \text{Binom}(x|n, p) \propto p^x(1-p)^{n-x}$, with n and x as specified in the problem. The *prior distribution* for p is going to be $f(p) = \text{Beta}(p|\alpha,\beta) \propto p^{\alpha-1}(1-p)^{\beta-1}$. The *posterior distribution* for p will also be $\text{Beta}(p|\alpha^*,\beta^*)$. Figure out what the new parameters α^* and β^* are, as functions of α, β, n and k.
   - To answer the question, first try some values for α and β that move the mean of the (prior) Beta distribution away from $k/n = 5/10 = 1/2$. Then try to keep the mean of the prior distribution constant but increase $\alpha + \beta$ (this will reduce the (prior) variance).
   - For graphing, note that there are a couple examples of graphing beta densities using the `dbeta()` function, in the R handout to go along with lectures in class.

3. *Bayes for Poisson data.* Recall from problem #1 that the pmf for the Poisson distribution is

$$f(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

   for $x = 0, 1, 2, \ldots$.

(a) Show that the Gamma distribution

$$f(\lambda|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}$$

is the conjugate prior distribution for the Poisson distribution above, and give the parameters of the posterior distribution, in terms of α, β, and x.

(b) Now suppose we have n independent counts $x_1, x_2, \ldots, x_n$ from a Poisson distribution with parameter λ. (i) Write the likelihood for this data; (ii) show that the Gamma distribution is still the conjugate prior distribution; and (iii) give the parameters of the posterior distribution, in terms of α, β, the x_i's and n. *Hint:* You worked with the likelihood function for problem #1 above.

(c) There is a certain intersection in town that has many accidents, even though the city installed a traffic light. Some students watched activity at that intersection during the noon hour on each of 16 successive days. They found that the number of cars that ran a red light at the intersection on each day was:

$$3, 7, 1, 7, 6, 1, 5, 4, 6, 2, 11, 3, 1, 6, 4, 4$$

Note that $\lambda = E[X]$, the expected number of cars to run a red light during each noon hour at that intersection. Find a point estimate an approximate 95% CI for λ, four different ways:

i. By calculating the sample average and sd from the data above, and applying the Central Limit Theorem to make a confidence interval for the mean.

ii. By calculating the MLE and SE for the MLE, as derived in problem #1 above.

iii. By fitting the model `glm(x ~ 1, family=poisson)` and making an appropriate transformation for a confidence interval for the intercept.

iv. Using a Gamma prior on λ with parameters $\alpha = 1$ and $\beta = 1$, find the posterior distribution and compute a suitable point estimate and 95% credible interval from the posterior.

(d) Comment on similarities and differences between the four methods above.