Homework 6: I Made You a Likelihood Function, But I Ate It

36-350, Statistical Computing

Due at 11:59 pm on Tuesday, 9 October 2012 (note day)

Agenda: Using functions as arguments *and* as return values; maximum likelihood estimation; standard errors of estimates.

This homework follows on from this week's lab. You can use the solution code from the lab.

1. (30) Using the version of gradient descent from lecture 10 and the log-likelihood function from lab, maximize the likelihood of the gamma distribution on the cats' hearts. Start the optimization at the estimate you get from the method of moments.
   (a) (10) What command do you use to maximize the log-likelihood?
   (b) (5) What is the estimate?
   (c) (5) What is the log-likelihood there? The gradient?
   (d) (5) Do the location and the value of the maximum match what you would expect from the plot from the lab? Explain.
2. (25) We need standard errors for the estimated parameters. If the model is right, we can get standard errors for maximum likelihood estimates from the *second* derivatives of the log-likelihood. (The second derivatives tell us how sharp the maximum of the likelihood is.) Specifically,

   $$\mathrm{Var}\left[\hat{\theta}\right] \approx -H^{-1}(\hat{\theta})$$

   where H is the **Hessian** of the log-likelihood, its matrix of second partial derivatives. (This is sometimes called the **observed information matrix**.)
   (a) (10) Install the package `numDeriv`. Its function `hessian` will calculate a numerical approximation to the Hessian of a given function at a given point. What is the Hessian matrix of the log-likelihood at the MLE?
   (b) (5) What is the inverse of the Hessian matrix?

(c) (10) What standard errors does this imply for the shape and scale estimate? *Hint:* Remember how the standard error of an estimate relates to the variance of the estimator.

3. (40) An alternative to using the Hessian is to use the jack-knife. This does *not* assume that the model is correct, but does mean we need to be able to re-compute the MLE after deleting points from the data set.

   (a) (15) Write a function, `make.gamma.loglike`, which takes in a data vector `x` and returns a log-likelihood function. Check that if `x` is `cats$Hwt`, then `make.gamma.loglike` returns a function which matches your old `gamma.loglike` at multiple parameter values.

   (b) (15) Write a function to calculate jack-knife standard errors for the MLE of the gamma distribution on an arbitrary data vector. The only argument should be the data vector. It should return a vector of length two, giving the standard errors for the shape and the scale parameters. Your code should use `make.gamma.loglike` and `gradient.descent`, and may use a `for` loop, or `sapply`. *Hint:* look at the solutions to lab 4.

   (c) (5) What are the jackknife standard errors for the MLE? (If you do not have two, one for the shape and one for the scale parameters, something is wrong.)

   (d) (5) Do the jack-knife errors for the MLE match the jack-knife errors for the method of moments? Do the jack-knife errors for the MLE match those obtained from the Hessian? Should *any* of these match?

4. (5) Are the maximum-likelihood and the method-of-moments estimates compatible with each other? Explain.

5. (5) What is the likelihood (not the log-likelihood) at the maximum? Should you be worried about using the model because the maximum likelihood is so small? Explain.