STAT 740, Fall 2017: Homework 1

1. **Poisson regression**. The Ache hunting data set has $n = 47$ observations (x_i, d_i, a_i) where x_i is the number of monkeys killed over d_i days for the ith hunter, aged a_i years. It is of interest to estimate and quantify the *monkey kill rate* as a function of hunter's age. Hunting prowess confers elevated status among the group, so a natural question is whether hunting ability improves with age, and at which age hunting ability is best. The following R code imports the data, makes the design matrix & log-likelihood function, and fits the model using the `glm` function to obtain the MLE $\hat{\boldsymbol{\theta}}$ and covariance matrix. This covariance matrix is based on the expected Fisher information, not observed; they are asymptotically equivalent.

```
d=read.table("http://people.stat.sc.edu/hansont/stat740/ache.txt",header=T)
attach(d) # makes 'age', 'days', and 'monkeys' available to R
n=length(age)
X=cbind(rep(1,n),age,age^2)
ll=function(theta){
 sum(dpois(monkeys,exp(log(days)+X%*%theta),log=T))
}
f=glm(monkeys~age+I(age^2),family="poisson",offset=log(days))
f$coef # MLE
vcov(f) # covariance matrix from Fisher scoring
```

   (a) Compute two different sets of crude starting values.

      i. For the "empirical log- rates" $r_i = \log(\frac{x_i}{d_i} + 0.1)$ use least-squares to obtain $\boldsymbol{\theta}_0 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{r}$, and

      ii. A first-order Taylors approximation $e^u \approx 1 + u$ leads to use of least-squares on $r_i = x_i - 1 - \log d_i$ to obtain $\boldsymbol{\theta}_0 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{r}$.

      Which starting values are closer to the true MLE?

   (b) As shown in class, hand-code Newton-Raphson in R to fit the Poisson regression model

   $$x_i \overset{ind.}{\sim} \text{Pois}\{\exp(\log d_i + \theta_1 + \theta_2 a_i + \theta_3 a_i^2)\}.$$

      Feel free to use `jacobian` and `hessian` in the `numDeriv` R package.

   (c) Dr. McMillan was interested in the age at which Ache hunters reached "maximum effectiveness" in hunting, e.g. the age where the mean kill rate reached a maximum. Note that the maximum of $\exp\{f(u)\}$ occurs at the same $\hat{u}$ as the maximum of $f(x)$ due to monitonicity. Find the function $\tau = g(\boldsymbol{\theta})$ that is the age where the mean kill-rate (per day) reaches a maximum. Obtain an estimate and 95% CI using the multivariate delta method.

2. For the O-ring failure data and logistic regression model use large-sample normal approximations and the multivariate delta method to get an estimate and 95% CI for the relative risk of failure at 36 degrees (the temperature at launch time) versus 70 degrees, using MLE theory and the Bayesian approach with the normal prior in the R examples. Specifically let the log relative risk be

$$\tau = g(\boldsymbol{\theta}) = (\theta_1 + 36\theta_2) - \log[1 + \exp(\theta_1 + 36\theta_2)] - (\theta_1 + 70\theta_2) + \log[1 + \exp(\theta_1 + 70\theta_2)].$$

   (a) Use Newton-Raphson as in the example to obtain the MLEs $\hat{\theta}$ and $\hat{\tau} = g(\hat{\theta}_1, \hat{\theta}_2)$; use the multivariate delta method to obtain $se(\hat{\tau})$. Next, obtain and interpret the MLE of the relative risk $e^{\hat{\tau}}$ and approximate 95% CI $(e^{\hat{\tau}-1.96se(\hat{\tau})}, e^{\hat{\tau}+1.96se(\hat{\tau})})$.

   (b) Repeat part (a) but instead using the posterior mode and approximate posterior covariance matrix from a Bayesian approach using the bivariate normal prior on $\boldsymbol{\theta}$ in the R examples. How does the estimate and CI change from the frequentist to the Bayesian approach?

3. Use accept-reject to sample from this bimodal density:

$$f(x) \propto 3\exp\{-0.5(x+2)^2\} + 7\exp\{-0.5(x-2)^2\}.$$

   The normalizing constant is 25.066. For your proposal $g(\cdot)$, use a $N(0, 2^2)$ distribution. Verify that your method works via a plot of the true normalized density, the proposal density, and a histogram of the generated values.

4. Use Metropolis-Hastings with an independence $N(0, 2^2)$ proposal $g(\cdot)$ to sample from the $f(\cdot)$ in problem 3. As in problem 3, show that your method works with a plot.