Likelihood Theory and Extensions (BIOS:7110)
Breheny

Assignment 7
Due: Monday, October 18

1. *Exponential dispersion.* Show that for the exponential dispersion family,

$$\mathbb{E}(\mathbf{s}) = \nabla\psi(\boldsymbol{\theta})$$
$$\mathbb{V}(\mathbf{s}) = \phi\nabla^2\psi(\boldsymbol{\theta}).$$

   Hint: Use the fact that $\psi(\boldsymbol{\theta})/\phi$ is the normalizing constant, then take first and second derivatives. For the purposes of this problem, you are not allowed to use any sort of result concerning cumulant generating functions (these are the results we're proving). You are allowed to take derivatives inside the integral sign (this is always justified for exponential families, but you do not need to show it).

2. *Completing the square.*

   (a) Show that if $\mathbf{A}$ is a symmetric, positive definite matrix, then any equation of the form

$$\mathbf{x}^\top\mathbf{A}\mathbf{x} + \mathbf{b}^\top\mathbf{x} + C = 0$$

   can be rewritten as

$$(\mathbf{x} - \mathbf{m})^\top\mathbf{A}(\mathbf{x} - \mathbf{m}) = D;$$

   provide expressions for $\mathbf{m}$ and D in terms of $\mathbf{A}$, $\mathbf{b}$, and C.

   (b) Suppose we have the Taylor series expansion

$$\tfrac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top\mathbf{H}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) + \mathbf{u}^\top(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) + C = 0.$$

   Using your result from part (a), show that this is equivalent (equal plus a constant not depending on $\boldsymbol{\theta}$) to the log-likelihood from the model $\mathbf{y} \sim \mathrm{N}(\boldsymbol{\theta}, \mathbf{V})$; find $\mathbf{y}$ and $\mathbf{V}$.

3. *Observed and Fisher information in the exponential family.* Suppose that $x_1, \ldots, x_n \overset{\text{iid}}{\sim} F$, where F is a distribution in the exponential family. Show that for all $\boldsymbol{\theta}$,

$$n\boldsymbol{\mathscr{I}}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}(\boldsymbol{\theta})$$

   where $\boldsymbol{\theta}$ is the natural parameter.

4. *Quadratic approximation for the Gamma distribution.* The pdf for the Gamma distribution with shape parameter α and rate parameter β is given by

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

   Generate a random sample of size 50 from this distribution using the following code:

```
set.seed(1)
x <- rgamma(50, shape=2, rate = 1)
```

To do this problem, please take note of the following information. The derivative of the log of the gamma function is known as the *digamma function*, usually denoted $\psi_0(\alpha)$:

$$\psi_0(\alpha) = \tfrac{d}{d\alpha} \log \Gamma(\alpha).$$

The second derivative of the log of the gamma function is known as the *trigamma function*, and usually denoted $\psi_1(\alpha)$:

$$\psi_1(\alpha) = \tfrac{d^2}{d\alpha^2} \log \Gamma(\alpha).$$

These functions are available in R and called `digamma()` and `trigamma()`, respectively.

(a) Derive the score vector for (α, β).

(b) Solve for the MLE. This is only available in partially-closed form. To find the MLE, solve for $\widehat{\beta}$ in terms of $\hat{\alpha}$ and use this to derive a one-dimensional score function for α, then use the `uniroot()` function to obtain a numerical answer for $\hat{\alpha}$.

(c) Derive the information matrix.

(d) Plot the two-dimensional 95% confidence region based on the quadratic/normal approximation at the MLE. On the plot, label the true value of (α, β). You can do this however you'd like, but my suggestion is to use the `ellipse` package.

(e) Overlay three gamma densities: the true gamma density for this problem, some other gamma density for parameter values inside the confidence interval, and a gamma density for a value outside the confidence interval. Comment briefly on the results.