Homework 3. (due March 25)

(1) Suppose $Y_1, \dots, Y_n$ are i.i.d. with density

$$f_{Y_i}(y;\mu) = \frac{1}{\mu}\exp(-\frac{y}{\mu}), y > 0, \mu > 0.$$

Show that the leading term in the saddlepoint approximation to the density of $\bar{Y} = \hat{\mu}$ reproduces the gamma density, with $\Gamma(n)$ replaced by Stirling's approximation to it. Deduce that the renormalized saddlepoint approximation is exact.

(2) (Calibration) Suppose we have n observations $y_1, \dots, y_n$ from a linear regression with known x's:

$$Y_i = \alpha + \beta x_i + \sigma e_i$$

where $e_1, \dots, e_n$ are i.i.d. $N(0,1)$. We will take a further observation y_0, say, from the model

$$Y_0 = \alpha + \beta\psi + \sigma e_0$$

where e_0 is also $N(0,1)$ and independent of $e_1, \dots, e_n$. The parameter of interest is ψ, the value of x corresponding to the new observation y_0. Derive the likelihood ratio statistic for testing ψ, and show that it is a function of a suitably defined t-statistic.

(3) (Severini, Ch.4). Let $Y_1, \dots, Y_n$ be independent and identically distributed from a density $f(y \mid \theta), \theta \in \mathbb{R}$, and let $\pi(\theta)$ be a probability density function for θ. Define the estimator $\tilde{\theta}$ by

$$\tilde{\theta}(y) = \frac{\int \theta\pi(\theta \mid y)d\theta}{\int \pi(\theta \mid y)d\theta}.$$

Using the Laplace approximation for integrals, show that $\tilde{\theta} - \hat{\theta} = O_p(1/n)$, where $\hat{\theta} = \hat{\theta}(y)$ is the maximum likelihood estimator of θ.

(4) (Wasserman, Ch.11; Berger & Wolpert, 1984). Let X_1, X_2 be independent Bernoulli random variables with $\Pr(X_i = 1) = \Pr(X_i = -1) = 1/2$, and define $Y_i = X_i + \theta, -\infty < \theta < \infty$, and suppose Y_1 and Y_2 are observed.

(a) Show that

$$C = \begin{cases} \{Y_1 - 1\} & \text{if } Y_1 = Y_2, \\ \{(Y_1 + Y_2)/2\} & \text{if } Y_1 \neq Y_2, \end{cases}$$

is a confidence set for θ with confidence 0.75. Suppose we observe $y_1 = 15$ and $y_2 = 17$. What interpretation should we give to the statement that C is a 0.75-confidence set, since we are certain $\theta = 16$?

(b) Suppose θ is an integer, and let $\pi(\theta)$ be a probability mass function that is positive for every θ. Compute the posterior distribution for θ when $y_1 = 15$ and $y_2 = 17$, and contrast the inference statement with the confidence statement.

(5) (A Neyman-Scott problem). A class of problems where maximum likelihood estimators are not consistent are those in which the number of nuisance parameters increases with the sample size. These are often called Neyman-Scott problems.[1] For example, if $Y_{ij}, j = 1, \ldots, m_i; i = 1, \ldots, n$ follow a $N(\mu_i, \sigma^2)$ distribution, the maximum likelihood estimator of σ^2 is inconsistent as $n \to \infty$; in particular if $m_i \equiv 2$, then $\hat{\sigma}^2 \to \sigma^2/2$; see CH Example 9.24.

(a) Suppose that Y_{i1} and Y_{i2} are independent observations from exponential distributions with means $\psi\lambda_i$ and ψ/λ_i, respectively, $i = 1, \ldots, n$. Show that the maximum likelihood estimator of ψ is not consistent, but converges in probability to $(\pi/4)\psi$.

(b) A modification to the profile likelihood to account for estimation of nuisance parameters was proposed in Cox & Reid (1987, JRSS B):

$$\ell_m(\psi) = \ell(\psi, \hat{\lambda}_\psi) - \frac{1}{2}\log|j_{\lambda\lambda}(\psi, \hat{\lambda}_\psi)|,$$

where $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\hat{\lambda}_\psi$ is the constrained maximum likelihood estimator of λ. This is to be computed using a parametrization of the nuisance parameter that is *orthogonal* to the parameter of interest ψ, with respect to expected Fisher information. (The correction term $\frac{1}{2}\log|j_{\lambda\lambda}(\psi, \hat{\lambda}_\psi)|$ is not invariant to reparameterizations,) Show for the exponential case that λ is orthogonal to ψ, and that the value of ψ that solves $\ell_m'(\psi) = 0$, $\hat{\psi}_m$, say, converges to $(\pi/3)\psi$.

[1] Neyman, J. and Scott, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrica* **16**, 1–32.