Eric Slud May 7, 2014

Stat 701 Take-Home Test, Due 5/12/14

Instructions. Do all of the following problems, to be handed in at the last class on Monday, May 12, 2014. This test must be done **individually** with no help from anyone else except for possible hints from me. In each problem, justify your assertions in terms of results and Theorems from the course.

(1). A large sample of 500 data-values X_i is assumed to come from a Gamma(α, β) distribution with unknown parameter vector $\vartheta = (\alpha, \beta)$. An estimator $\tilde{\vartheta}$ (the generalized method of moments estimator for some smooth 2-vector function $g(x)$, but we are not told g because it is used only in generating $\tilde{\vartheta}$ as a preliminary estimator) is reported as $\tilde{\vartheta} = (1.703, 2.300)$. Moreover, the gradient and Hessian of the Gamma-density log-Likelihood at this point $\tilde{\vartheta}$ are reported as

$$\nabla_{\vartheta} \,\mathrm{logL}(\tilde{\vartheta}) \;=\; \begin{pmatrix} -13.239 \\ 0.171 \end{pmatrix} \quad , \quad \nabla^{\otimes 2}_{\vartheta} \,\mathrm{logL}(\tilde{\vartheta}) \;=\; \begin{pmatrix} -395.71 & 217.44 \\ 217.44 & -161.03 \end{pmatrix}$$

(a). Approximate the MLE $\hat{\vartheta}$ as well as possible from this information, and explain in terms of asymptotic statistical theory how good your answer is.

(b). Give an asymptotically valid two-sided 95% Confidence Interval for α/β based on the reported statistics and your answer in (a).

(2). Suppose that variables Z_i, $1 \le i \le 100$, are *iid* Gamma(β, β^2) variables, with unknown parameter $\beta > 0$. This distribution has probability density function on $[0, 1]$ given by $f(z) = c(\beta)\, z^{\beta-1}\, e^{-z\beta^2}$, with mean $= 1/\beta$ and variance $= 1/\beta^3$.

(a). Find a 'variance-stabilizing' transformation $g(z)$ which gives $g(\bar{Z})$ an approximate large-sample normal distribution with mean $g(1/\beta)$ and variance which does not (to first order) depend on β.

(b). Use the transformation g defined in (a) to find a two-sided approximate 95% large-sample confidence interval for β depending only on $\bar{Z}$ which is symmetric for $g(1/\beta)$ but not for β .

(3) Suppose that a sample of *iid* pairs $\{(Z_i,\, Y_i),\;\; i = 1, \ldots, n\}$ follows the probability model

$$Z_i \sim \text{Expon}(\lambda)\,, \quad E(Y_i|Z_i) \,=\, \alpha + \beta Z_i\,, \quad \text{Var}(Y_i|Z_i) \,=\, \sigma^2\, Z_i \qquad (*)$$

(a) Not knowing the true model (*), we analyze the data by least-squares estimates, i.e. by estimates of (α, β) minimizing $\sum_{i=1}^n (Y_i - \alpha - \beta\, Z_i)^2$. Show that this estimator is $\sqrt{n}$ consistent, and find an expression for the variance-covariance matrix V of the asymptotic distribution of $\sqrt{n}(\hat{\alpha} - \alpha,\, \hat{\beta} - \beta)$. **Note:** this variance will involve the nuisance parameters λ and σ^2.

(b) How can V be estimated consistently, if you know only that $Z_i \sim \text{Expon}(\lambda)$ and that (Z_i, Y_i) are *iid* with $E(Y_i - \alpha - \beta\, Z_i \,|\, Z_i) = 0$ (or that $E(Y_i - \alpha - \beta\, Z_i) = E(Z_i\,(Y_i - \alpha - \beta\, Z_i)) = 0$) ?

(4) Suppose that we observe X_i, $i = 1, \ldots, n$, *iid* scalar-valued measurements, and we wish to test the hypothesis $\mathbf{H}_0$ that they are distributed according to the Logistic(μ, σ) distribution with d.f. $F(x, \vartheta) = (1 + e^{-(x-\mu)/\sigma})^{-1}$ for some $\mu \in \mathbf{R}$, $\sigma > 0$. The parameters $\vartheta = (\mu, \sigma)$ are estimated in two distinct ways:

(i) first, $\hat{\vartheta}$ denotes the MLE's, and

(ii) second, denote by $\tilde{\vartheta}$ the MLE's for the same parameters based on the reduced data (N_1, N_2, N_3, N_4) where N_j is the number $\sum_{i=1}^n I_{[X_i \in (a_{j-1}, a_j]]}$ of observations falling between a_{j-1} and a_j, where the a_j's are fixed and respectively

$$a_0 = -\infty\,, \quad a_1 = -1, \quad a_2 = 0\,, \quad a_3 = 1, \quad a_4 = \infty$$

(a) Find the asymptotic variance matrices of these two estimators.

(b) Find the asymptotic null-hypothesis distribution of the test statistic

$$\sum_{j=1}^{4} \frac{(N_j - n(F(a_j, \tilde{\vartheta}) - F(a_{j-1}, \tilde{\vartheta})))^2}{n(F(a_j, \tilde{\vartheta}) - F(a_{j-1}, \tilde{\vartheta}))}$$

(c) Find the asymptotic null-hypothesis distribution of the random variable $(N_1 - nF(-1, \hat{\vartheta}))/\sqrt{n}$.