STATS 730: Midterm test

Monday 23 Sept 2013

Total marks is 60

Question 1: 10 mks: In a 2-parameter model, the MLE was $\widehat{\boldsymbol{\theta}} = (\widehat{\theta}_1, \widehat{\theta}_2) = (2, 3)$ and the negative of the Hessian of the log-likelihood evaluated at $\widehat{\boldsymbol{\theta}}$ was

$$-\mathbf{H}(\widehat{\boldsymbol{\theta}}) = \begin{pmatrix} 10 & 5 \\ 5 & 5 \end{pmatrix}$$

which has inverse $\begin{pmatrix} 0.2 & -0.2 \\ -0.2 & 0.4 \end{pmatrix}$.

1. Calculate the Wald test statistic for $H_0 : \theta_1 = 1$.
2. Calculate the Wald test statistic for $H_0 : \boldsymbol{\theta} = (1, 1)$.
3. Now, suppose that the parameterization was changed to $\boldsymbol{\zeta} = (\theta_1, \theta_1\theta_2)$. Calculate the approximate variance matrix for the MLE $\widehat{\boldsymbol{\zeta}}$.

Question 2: 12 mks: One hundred iid observations $\boldsymbol{y} = (y_1, ..., y_{100})$ were observed from a $N(\mu, \sigma^2)$ distribution, and the MLEs were calculated to be $\widehat{\mu} = \overline{y} = 2$ and $\widehat{\sigma} = \sqrt{\frac{\sum_{i=1}^{100}(y_i - \overline{y})^2}{100}} = 10$.

1. Calculate the value of the log-likelihood evaluated at the MLE, $l(\widehat{\mu}, \widehat{\sigma})$.
2. Determine the formula for the profile log-likelihood function

$$l^*(\mu) = \max_{\sigma} l(\mu, \sigma) \ .$$

3. Using the profile log-likelihood function from above (or otherwise if you prefer), calculate the likelihood ratio test statistic for the null hypothesis $H_0 : \mu = 0$. You may use the fact that $\sum_{i=1}^{100} y_i^2 = 10400$.

Question 3: 5 mks: Suppose that observations $\boldsymbol{y} = (y_1, ..., y_n)$ are iid from a Uniform$(0, m)$ distribution. We know that the MLE of m is $\widehat{m} = y_{\max}$. Since the usual regularity conditions do not hold, it has been suggested to obtain an approximate 95% confidence interval for m using the percentile method from a parametric bootstrap.

Is this a good idea? Be sure to justify your answer.

Question 4: 8 mks: For these questions, be sure to explain all relevant notation and terminology used.

1. State Jensen's inequality for convex functions.
2. Define AIC.
3. What does it mean for a model to be *identifiable*.
4. Give an example of an MLE that is approximately normal (in the sense that its distribution becomes arbitrarily close to that of a normal distribution when n increases), yet does not possess a mean.

Question 5: 5 mks: Subject to regularity conditions, it has been shown that the likelihood is higher at the true parameter value, θ_0, than at any other fixed parameter value with probability that becomes arbitrarily close to 1 as the sample size increases. Using this fact, show that the maximum likelihood estimator $\widehat{\theta}$ is consistent. (You may assume $\theta \in \mathbb{R}$ and that the likelihood is unimodal.)

Question 6: 8 mks: Suppose that y is observed from a Binomial$(n, 0.4)$ distribution, where n is the integer-valued parameter to be estimated. Find the MLE of n. Is it unique?

Question 7: 12 mks: Suppose that $y_1, ..., y_{100}$ are iid observations from a χ^2_d distribution, and it is desired to make inference about parameter d.

Write an R program to (you may assume that the data are in vector `y`),

1. Compute the MLE, $\widehat{d}$, by numerical optimization of the log-likelihood.
2. Compute the (approximate) 95% Wald confidence interval for d.
3. Compute the likelihood ratio test statistic for $H_0 : d = 10$.
4. Implement a parametric bootstrap simulation to obtain the bootstrap p-value for the above LRT statistic.

Recall that R function `rchisq(n,df)` generates n iid $\chi^2_{\texttt{df}}$ random variables, and `dchisq(y,df)` evaluates the $\chi^2_{\texttt{df}}$ density function at `y`, which may be a vector.