STATS 730: Midterm test

Monday 17 Sept 2012

Total marks is 60

Question 1: 10 mks: Let $\boldsymbol{y} = (y_1, y_2, ..., y_n)$ be iid observations from an exponential distribution with density function

$$f(y;\mu) = \frac{e^{-y/\mu}}{\mu} \quad , y > 0$$

where $\mu = E[Y]$.

1. Show that the MLE is $\widehat{\mu} = \overline{y}$.
2. Calculate the curvature of the log-likelihood, and hence provide an estimate of the variance of $\widehat{\mu}$.
3. The exponential distribution is often parameterized using $\lambda = \mu^{-1}$. Use the delta method to obtain the approximate variance of $\widehat{\lambda} = \widehat{\mu}^{-1}$.

Question 2: 7 mks: In a 2-parameter model, the MLE was $\widehat{\boldsymbol{\theta}} = (\widehat{\theta}_1, \widehat{\theta}_2) = (-2, 2)$ and the negative of the Hessian of the log-likelihood evaluated at $\widehat{\boldsymbol{\theta}}$ was

$$-\mathbf{H}(\widehat{\boldsymbol{\theta}}) = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

which has inverse $\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$.

1. Calculate the Wald test statistic for $H_0 : \theta_1 = 0$.
2. Calculate the Wald test statistic for $H_0 : \boldsymbol{\theta} = \mathbf{0}$.

Question 3: 10 mks: Suppose that $y_1 = 1$ is observed from a $\text{Bin}(10, p_1)$ distribution, and $y_2 = 9$ is observed from $\text{Bin}(10, p_2)$ distribution.

Calculate the likelihood ratio test statistic for $H_0 : p_1 = p_2$.

Question 4: 5 mks: It is known that the likelihood ratio and Wald test statistics are equal when the log-likelihood is quadratic in shape. Using this fact, provide an argument to establish that the likelihood ratio test statistic is preferred to the Wald test statistic.

Question 5: 9 mks: Suppose that observations $\boldsymbol{Y} = (Y_1, ..., Y_n)$ are iid from a distribution with true density function $f(y; \boldsymbol{\theta}_0)$ within the statistical model $f(y; \boldsymbol{\theta}), \theta \in \Theta$, and let $f(\boldsymbol{y}; \boldsymbol{\theta})$ denote the (joint) density function of $\boldsymbol{Y}$. For any fixed $\boldsymbol{\theta}$ not equal to $\boldsymbol{\theta}_0$, show that

$$f(\boldsymbol{Y}; \boldsymbol{\theta}_0) > f(\boldsymbol{Y}; \boldsymbol{\theta})$$

with probability that converges to 1 as n goes to infinity. (You may assume all necessary "regularity" conditions.)

[*Hint:* Key ingredients of the proof are Jensen's inequality and the weak law of large numbers, which says that the sample mean of a random variable converges to the population mean (subject to weak regularity conditions).]

Question 6: 13 mks: For these questions, be sure to explain all notation and terminology used.

1. Describe the plug-in approach to prediction and explain why it is not a good method to use. Briefly describe a better method for prediction.
2. Briefly describe how a bootstrap simulation can be used to assess the sampling distribution of a test statistic.
3. In the context of survival analysis, define the hazard function, and show that it can be interpreted as a measure of the instantaneous risk of failure over a small time interval $(t + \Delta t)$.

Question 7: 6 mks: Two groups of patients each were used in a study of the effectiveness of a drug. The first group received the drug and the second group received a placebo. There were 3 patients in the drug group and 4 in the placebo group and the survival times were

Drug group:	4 (5) (12)
Placebo group:	3 7 (8) 10

where values in parentheses denote censored observations.

Write down Cox's partial likelihood function for these data, assuming a proportional hazards model. Be sure to define your notation. (Do not attempt to simplify or maximize this partial likelihood).