Examination in School of Mathematical Sciences
Semester 1, mid-term examination, 2010

011346 Advanced Statistical Inference IV
STATS 4000A, STATS 7004

Official Reading Time: 5 mins
Writing Time: 45 mins
Total Duration: 50 mins

NUMBER OF QUESTIONS: 4 TOTAL MARKS: 30

Instructions

• Attempt all questions.
• You may commence writing when you are ready.
• Begin each answer on a new page.
• Examination materials must not be removed from the examination room.

Materials

• 1 Purple book is provided.
• Calculators are not permitted.

DO NOT COMMENCE WRITING UNTIL INSTRUCTED TO DO SO.
1. The cumulant generating function for the distribution of a random variable $Y$ is defined by $K_Y(t) = \log M_Y(t)$. Using this definition, or otherwise, find the first two cumulants $\kappa_1$ and $\kappa_2$ of the distribution of $Y$ in terms of the moments about the origin. [5 marks]

2. Show that the normal density function

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left\{ -(y - \mu)^2 / 2\sigma^2 \right\}, \quad -\infty < y < \infty$$

is an exponential family model by writing the density in the appropriate form, then defining the natural parameters, the natural statistics, the natural parameter space, and $\kappa(\theta)$. [8 marks]

3. (a) Define what it means to say that a sufficient statistic $T(Y)$ is minimal sufficient.

(b) Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the location exponential distribution with density

$$f(y; \theta) = e^{-(y-\theta)}, \quad \theta < y < \infty, \quad -\infty < \theta < \infty.$$ 

Find a minimal sufficient statistic for $\theta$. [8 marks]

4. Suppose data $Y_1, \ldots, Y_n$ satisfy a normal linear model

$$Y = X\beta + \epsilon$$

where $X$ is the $(n \times p)$ design matrix, $\beta$ is a $(p \times 1)$ vector of regression parameters, and the $e_j$ are i.i.d. $N(0, \sigma^2)$. We know the minimal sufficient statistic for $(\beta, \sigma^2)$ is

$$\hat{\beta} = (X^T X)^{-1} X^T y, \quad S^2 = \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p}$$

(a) Suppose interest lies in $\sigma^2$, and $\beta$ is regarded as a nuisance parameter. Given that $(n - p)S^2 / \sigma^2 \sim \chi^2_{n-p}$, explain why the joint density of $Y$ can be written as follows:

$$f(y; \beta, \sigma^2) = f(y|\hat{\beta}, s^2) f(\hat{\beta}; \beta, \sigma^2) f(s^2; \sigma^2)$$ \hspace{1cm} (1)

(b) Using the last term in the right-hand-side of Equation (1) above, show that the marginal maximum likelihood estimator of $\sigma^2$ is $S^2$. [Hint: the chi-squared density for $n$ degrees of freedom is the gamma density, $f(x; \alpha, \lambda) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} \exp(-\lambda x), \quad x \geq 0$, with $\alpha = n/2$ and $\lambda = 1/2$.] [9 marks]