Advanced Topics in Statistics (ST414): Asymptotic Statistics Exercise sheet

Ioannis Kosmidis

May 28, 2010

1. Consider a random variable X distributed according to a Poisson distribution with parameter λ ($P(\lambda)$), that is
$$P(X = x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}\,, \quad x = 0, 1, \dots, \lambda > 0\,.$$
   (a) Show that the distribution of X is an exponential family and find the natural statistic $s(x)$, the natural parameter $\theta \equiv \theta(\lambda)$ and the cumulant generating function $K_X(t)$.
   (b) If $X_1, \dots, X_n$ are independent random variables with $X_i \sim P(\lambda_i)$, show that $Y = \sum_{i=1}^n X_i \sim P(\sum_{i=1}^n \lambda_i)$.
   (c) According to the Central Limit Theorem, if $X_1, \dots, X_n$ are independent and identically distributed random variables with $X_i \sim P(\lambda)$ ($i = 1, \dots, n$), then
$$Z = \frac{\sum_{i=1}^n X_i - n\lambda}{\sqrt{n\lambda}} \xrightarrow{d} N(0,1)\,,$$
   where $N(0,1)$ is used to denote the standard normal distribution. Prove this result by examining the asymptotic behaviour of an expansion of the cumulant generating function of Z.

2. Suppose that $X_1, \dots, X_n$ are independent and identically distributed random variables with $X_i \sim P(\lambda)$. Using the results of Exercise 1:
   (a) Write down the log-likelihood $l(\lambda)$ for λ, obtain the score function $U(\lambda)$, the maximum likelihood estimator $\hat\lambda$ and the expected information $i(\lambda)$.
   (b) Consider the new parameterization $\theta \equiv \theta(\lambda) = \log(\lambda)$. Using the results in (a) or otherwise find the maximum likelihood estimator $\hat\theta$ of θ and write down the score function $\tilde U(\theta)$ and the expected information $\tilde i(\theta)$ on θ.
   (c) In general parametric problems with scalar parameter θ and under the usual regularity conditions, the $O(n^{-1})$ term in the asymptotic expansion of the bias of maximum likelihood estimator $\hat\theta$ has the form
$$b(\theta) = \left\{ E_\theta\left(\frac{\mathrm{d}^3\tilde l(\theta)}{\mathrm{d}\theta^3}\right) + 2E_\theta\left(\frac{\mathrm{d}\tilde l(\theta)}{\mathrm{d}\theta}\frac{\mathrm{d}^2\tilde l(\theta)}{\mathrm{d}\theta^2}\right)\right\} \Big/ \left[2\left\{\tilde i(\theta)\right\}^2\right]\,,$$
   where $\tilde l(\theta)$ is the log-likelihood on θ and $\tilde i(\theta)$ is the expected information on θ.
   For the current setting, show that $b(\theta) = -\{2n\exp(\theta)\}^{-1}$ and hence based on $\hat\theta$ construct a bias-corrected estimator $\hat\theta_{BC}$ of θ.

3. Consider a random variable X distributed according to the Inverse Gaussian distribution with mean parameter μ and shape parameter λ ($IG(\lambda, \mu)$), that is
$$f_X(x; \mu, \lambda) = \frac{\sqrt{\lambda}}{\sqrt{2\pi}} x^{-3/2} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}\,, \quad x > 0\,, \lambda > 0\,, \mu > 0\,.$$

(a) Show that the distribution of X is an exponential family and find the natural statistics $s_1(x)$ and $s_2(x)$, the natural parameters $\theta_1 \equiv \theta_1(\lambda, \mu)$ and $\theta_2 \equiv \theta_2(\lambda, \mu)$, and the cumulant transform $k(\theta_1, \theta_2)$.

(b) Calculate the covariance of X and $1/X$ in terms of the parameters λ and μ.

4. Suppose that Y is distributed according to a density of the form

$$f_Y(y; \psi, \lambda) = \exp\{s_1(y)\theta_1(\psi) + s_2(y)\theta_2(\lambda) - k(\theta_1(\psi), \theta_2(\lambda)) + c(y)\} ,$$

and suppose that ψ is the parameter of interest. Furthermore, let ξ be the *complementary mean parameter* given by

$$\xi \equiv \xi(\psi, \lambda) = E\{s_2(Y)\} .$$

(a) Show that ψ and ξ are orthogonal.

(b) Let Y have Gamma distribution with scale parameter ϕ and shape parameter n ($G(n, \phi)$), that is

$$f_Y(y; \phi, n) = \frac{y^{n-1}\exp(-y/\phi)}{\phi^n \Gamma(n)}, \quad y > 0\,, \phi > 0\,, n > 0\,.$$

Show that $n\phi$ is orthogonal to n.

5. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with $X_i \sim N(\mu, \sigma^2)$ (Normal distribution with mean μ and variance σ^2) and suppose that the parameter of interest is μ.

(a) Write down the profile log-likelihood $l_p(\mu)$ on μ.

(b) Show that the $100(1-\alpha)\%$ approximate confidence interval for μ based on $l_p(\mu)$ has the form

$$\left\{\hat{\mu} - \hat{\sigma}\sqrt{\exp\left(\frac{\chi^2_{1-\alpha}}{n}\right) - 1}\,,\ \hat{\mu} + \hat{\sigma}\sqrt{\exp\left(\frac{\chi^2_{1-\alpha}}{n}\right) - 1}\right\} ,$$

where $\chi^2_{1-\alpha}$ is the $(1-\alpha)$th quantile of the χ^2 distribution with 1 degree of freedom, and $\hat{\mu} = \sum_{i=1}^n X_i/n$ and $\hat{\sigma}^2 = \sum_{i=1}^n (X_i - \hat{\mu})^2/n$ are the maximum likelihood estimates of μ and σ^2, respectively.

6. Let $X_1, \ldots, X_n$ ($n > 5$) be independent and identically distributed random variables with $X_i \sim N(\mu, \sigma^2)$ and consider testing the hypothesis $H_0 : \mu = \mu_0$. Show that the likelihood ratio statistic for testing H_0 may be expressed as

$$w(\mu_0) = n \log\left\{1 + \frac{t^2}{n-1}\right\} ,$$

where t is the Student's t statistic, i.e.

$$t = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sqrt{\sum_{i=1}^n (X_i - \bar{Y})^2/(n-1)}} ,$$

where $\bar{X} = \sum_{i=1}^n X_i/n$. Show that $E_{\mu_0}\{w(\mu_0)\} = 1 + 3/(2n) + O\left(n^{-2}\right)$ and hence construct a Bartlett corrected version of $w(\mu_0)$.

HINT: You may use without justification the following results:

- R1: $E(t^2) = \dfrac{n-1}{n-3}$ and $E(t^2) = \dfrac{3(n-1)^2}{(n-3)(n-5)}$.
- R2: If $u = O(n^{-1})$ then $\dfrac{1}{1-u} = 1 + u + O\left(n^{-2}\right)$.