HW Question Week 7

STA2101F 2021

Due November 4 2021 11.59 pm

Homework to be submitted through Quercus

A parametric model is a member of the *exponential family* of distributions if its density takes the form

$$f(y;\theta) = \exp\{s(y)^T\theta - c(\theta) - h(y)\}, \tag{1}$$

where $\theta = (\theta_1, \ldots, \theta_p)$, and $s(y)$ is a p-dimensional sufficient statistic. For example, the normal distribution has density

$$f(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2\sigma^2}(y-\mu)^2\} = \exp\{y\frac{\mu}{\sigma^2} + y^2\frac{-1}{2\sigma^2} - \{\frac{\mu^2}{2\sigma^2} + \frac{1}{2}\log(\sigma^2)\} - \log(\sqrt{2\pi})\},$$

so $s(y) = (y, y^2)$, $\theta = \{\mu/\sigma^2, -1/(2\sigma^2)\}$. The parameter θ is called the canonical parameter of the exponential family.

(a) Show that if y is distributed as $Binom(n,p)$ that this is a member of the exponential family, and give expressions for θ (in terms of p) and $c(\theta)$. Show that if y is distributed as $Poisson(\lambda)$ that this is a member of the exponential family and give expressions for θ (in terms of λ) and $c(\theta)$.

(b) Show that if $y = (y_1, \ldots, y_n)$ are independent, and identically distributed, according to (1), that the joint distribution of y is of the same form, with $s(y) = \sum_{i=1}^n s(y_i)$, $c(\theta)$ replaced by $nc(\theta)$.

(c) Use this result to argue that the marginal density of $s = (s_1, \ldots, s_p)$ has again the same form, i.e.

$$f(s;\theta) = \exp\{s^T\theta - nc(\theta) - k(s)\},$$

where $k(s)$ is defined from an appropriate integral.

(d) Suppose $y_1, \ldots, y_n$ are independent, with $y_i \sim Binom(n_i, p_i)$, where $p_i = \exp(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i})/\{1 + \exp(\theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i})\}$. A slight extension of the argument in (b) shows that the joint density of $y = (y_1, \ldots, y_n)$ is a member of the exponential family (you do not need to derive this). What are the sufficient statistics s.

(e) Show that the maximum likelihood estimate of θ in (d) is defined by

$$\begin{aligned}\sum_{i=1}^n y_i &= \sum_{i=1}^n n_i p_i(\hat{\theta}),\\ \sum_{i=1}^n x_{1i} y_i &= \sum_{i=1}^n x_{1i} n_i p_i(\hat{\theta}),\\ \sum_{i=1}^n x_{2i} y_i &= \sum_{i=1}^n x_{2i} n_i p_i(\hat{\theta}).\end{aligned}$$

(f) *Bonus/PhD SM, Problem 10.9*: At each of the doses $x_1 < x_2 < \cdots < x_n$ of a drug, m animals are treated. At dose x_i, r_i animals have response 1 and the remaining $m - r_i$ have response 0. Derive the maximum likelihood equation when a *probit* link function is used, and the linear predictor takes the form $\eta = \beta x$; i.e. $r_i \sim Binom(m, p_i)$ and $p_i = \Phi(x_i \beta)$. Show that if $x_i \equiv x_0$, i.e. only one dose is used, that

$$\hat{\beta} = \frac{1}{x_0}\Phi^{-1}(r/m), \quad \text{var}(\hat{\beta}) \doteq \frac{\Phi(\beta x_0)\{1 - \Phi(\beta x_0)\}}{m x_0^2 \{\phi(\beta x_0)\}^2},$$

where ϕ and Φ are the standard normal density and distribution functions. Plot the function $\Phi(\eta)\{1 - \Phi(\eta)\}/\phi^2(\eta)$ for η in the range $(-3, 3)$ and comment on the implication sfor the choice of x_0 if there is some prior knowledge of the likely value of β.