University of Technology Sydney
School of Mathematical and Physical Sciences

Mathematical Statistics (37262) – Class 9 Preparation Work

1. Let $X \sim Beta(\alpha,\beta)$ where $\beta > 0$ is known but $\alpha > 0$ is not.

   That is, X has density function $f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)x^{\alpha-1}(1-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$.

   i) Show that X belongs to the exponential family.

   Let $Y \sim Gamma(\alpha,\beta)$ where neither $\alpha > 0$ or $\beta > 0$ is known.

   That is, Y has density function $f(y) = \begin{cases} \dfrac{\beta^{\alpha} y^{\alpha-1} e^{-\beta y}}{\Gamma(\alpha)} & y \in [0,\infty) \\ 0 & \text{otherwise} \end{cases}$.

   ii) Show that Y belongs to the (two parameter) exponential family. That is, show that it can be written in the form $f(y) = a(y)b(\boldsymbol{\Theta})\exp\left(\boldsymbol{c}(\boldsymbol{\Theta}).\boldsymbol{d}(y)\right)$.

2. Let $S \sim \chi^2_v$

   That is, S has density function $f(s) = \begin{cases} \dfrac{1}{\sqrt{2^v}\,\Gamma\left(\frac{v}{2}\right)} e^{-\frac{s}{2}} s^{\frac{v}{2}-1} & s \in [0,\infty) \\ 0 & \text{otherwise} \end{cases}$.

   i) Show that $S \sim \chi^2_v$ belongs to the exponential family.

   ii) Find the natural parameter for this distribution.

   iii) Hence show that $E(\ln(S)) = \ln(2) + \dfrac{\Gamma'\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v}{2}\right)}$.

   iv) Find the variance of $\ln(S)$.

3. A group of children aged between 3 and 18 years old are asked to run 100m. The following information is recorded

y_i : an indicator variable which is 1 is the *i*th child takes more than 30s to run 100m and which is 0 if the *i*th child takes 30s or less to run 100m.

x_{1i} : the height of the *i*th child in cm.

x_{1i} : the mass of the *i*th child in kg.

Consider the interpretation of a binary logistic regression model with the logit link function.

i) If the underlying model is $\hat{y}_i = \alpha + \beta_1 x_{1i}$, explain why you might expect the coefficient β_1 to be negative.

ii) If the underlying model is $\hat{y}_i = \alpha + \beta_2 x_{2i}$, explain why you might expect the coefficient β_2 to be negative.

iii) If the underlying model is $\hat{y}_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2_i}$, if one of the coefficients (β_1 or β_2) was positive and the other was negative, which would you expect to be positive? Justify your answer.