Problem Set - Generalized Linear Model

October 1, 2021

1 Exponential Family

The exponential family has PDF function

$$p(y, \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right)$$

when

$$\begin{aligned} b(\theta) &= e^{\theta}, \\ a(\phi) &= 1, \\ c(y, \phi) &= -\log(y!) \end{aligned}$$

Prove that $p(y, \theta, \phi)$ is a Poisson distribution. Hint: re-parameterize $\theta = \log(\lambda)$

2 GLM

Given a data set of n data points $[y_i, \vec{x}_i]$, $i = 1, 2, 3, ..., n$. Assume that ϕ is constant and θ is dependent on $\vec{x}_i$, the log-likelihood function have a format $\ell = \sum_{i=1}^{n}[\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)]$. To maximize the ℓ is equivalent to maximize $\ell = \sum_{i=1}^{n}[y\theta_i - b(\theta_i)]$. In GLM, we model θ_i as a function of $\vec{x_i} \cdot \vec{\beta}$ i.e. $b'(\theta_i) = g(\vec{x_i} \cdot \vec{\beta})$

1. Prove that the jth component of the gradient $\nabla \ell_j = \sum_{i=1}^{n}[y_i - b'(\theta_i)]\frac{\partial \theta_i}{\partial \beta_j}$

2. Prove that $\frac{\partial \theta_i}{\partial \beta_j} = \frac{g'(\vec{x}_i \cdot \beta)}{b''(\theta_i)} x_{ij}$, here x_{ij} is the jth component of vector $\vec{x}_i$. Hint: calculate the partial derivative of the equation $b'(\theta_i) = g(\vec{x_i} \cdot \vec{\beta})$ against β_j on both side.