Midterm 2

Luke Letizia

April 8, 2018

Question 1

A.

$$\mathrm{L} = \prod_{i=1}^{n} e^{\frac{y_i\theta_i - \mathrm{b}(\theta_i)}{\phi_i} + \mathrm{S}(y_i,\phi_i)}$$

-I started the maximum likelihood procedure by taking the product of the function.

$$\ln(\mathrm{L}) = \sum_{i=1}^{n} \frac{y_i\theta_i - \mathrm{b}(\theta_i)}{\phi_i} + \mathrm{S}(y_i\phi_i)$$

-I then took the natural log of the product.

$$\frac{\delta \mathrm{L}}{\delta \underline{\beta}} = \underline{0} = \sum_{i=1}^{n} \frac{(y_i - \mathrm{b}'(\theta_i)\frac{\delta\theta}{\delta\underline{\beta}})}{\phi_i}$$

-Here I took the derivative of the function.

$$= \sum_{i=1}^{n} \frac{y_i - \mu_i}{\mathrm{Var[Y]}} * \mathrm{b}''(\theta) * \frac{\delta\theta}{\delta\underline{\beta}}$$

-Here I took the derivative of the function and then substituted the following values; $\phi = \frac{\mathrm{Var[Y]}}{\mathrm{b}''(\theta)}$ and $\mathrm{b}'(\theta) = \mu$

$$= \sum_{i=1}^{n} \frac{y_i - \mu_i}{\mathrm{Var[Y]}} (\frac{\delta\mu}{\delta\underline{\beta}})(\frac{\delta\underline{\beta}}{\delta\theta}) * (\frac{\delta\theta}{\delta\underline{\beta}})$$

-I also substituted the following value to solve for the simplest form;

$$\frac{\delta\mu}{\delta\underline{\beta}} = \frac{\delta\mathrm{b}'(\theta)}{\delta\underline{\beta}} = \mathrm{b}''(\theta)\frac{\delta\theta}{\delta\underline{\beta}} <=> \mathrm{b}''(\theta) = \frac{\delta\mu}{\delta\underline{\beta}} * \frac{\delta\underline{\beta}}{\delta\theta}$$

$$= \sum_{i=1}^{n} \frac{y_i - \mu_i}{\mathrm{Var[Y]}} * \frac{\delta\mu}{\delta\underline{\beta}} = \underline{0}$$

B.

$$\mathrm{L} = \prod_{i=1}^{n} e^{\frac{y_i\theta_i - \mathrm{b}(\theta_i)}{\phi_i} + \mathrm{S}(y_i,\phi_i)}$$

-I started the maximum likelihood procedure by taking the product of the function.

$$\ln(\mathrm{L}) = \sum_{i=1}^{n} \frac{y_i\theta_i - \mathrm{b}(\theta_i)}{\phi_i} + \mathrm{S}(y_i\phi_i)$$

-I then took the natural log of the product.

$$\frac{\delta \mathrm{L}}{\delta \underline{\beta}} = 0 = \frac{\delta}{\delta\underline{\beta}_{\mathrm{L}}} \sum_{i=1}^{n} \frac{y_i x_i^{\mathrm{T}}\underline{\beta} - \mathrm{b}'(x_i^{\mathrm{T}}\underline{\beta})}{\phi_i} + S(y_i,\phi_i)$$

-Here I substituted θ_i values for $x_i^{\mathrm{T}}\underline{\beta}$

$$= \sum_{i=1}^{n} \frac{y_i x_{i\mathrm{L}} - b'(x_i^{\mathrm{T}}\underline{\beta} x_{i\mathrm{L}})}{\phi_i}$$

-Here, I took the derivative of the log function.

$= \frac{1}{\phi} \sum_{i=1}^{n} (y_i - b'(x_i^{\mathrm{T}} \underline{\beta})) x_{i\mathrm{L}} w_i$

-I then substituted the following value and simplified; $\phi_i = \frac{\phi}{w_i}$

$= \sum_{i=1}^{n} (y_i - b'(\underline{\beta})) * w_i$

-Setting $x_i = 1$ for $i = 1, 2, \ldots, n$, I got the following result.

$= \sum_{i=1}^{n} (y_i - \mu)) * w_i = 0$

-I finally substituted $\mathrm{b}'(\underline{\beta}) = \mu$ and solve for $\hat{\mu}$

$\sum_{i=1}^{n} y_i w_i = \sum_{i=1}^{n} \mu * w_i <=> \hat{\mu} = \frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i}$

C.

$g(\mu) = \theta$ & $\theta = \mu$

$\hat{\mu} = \hat{\mu} = \frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i}$

-Here, I substituted $\hat{\mu}$, which we found in part b with the value of $\hat{\lambda}$ which I found from a series of equivalences.

$= \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$

-By setting $w_i = 1$ we could solve for $\hat{\mu} = \bar{y}$

D.

$g(\mu) = \theta$ & $\theta = \log\lambda$

$\mu = \mathrm{b}'(\theta) = e^{\theta} = e^{\log\lambda} = \lambda$

$\hat{\mu} = \hat{\lambda} = \frac{\sum_{i=1}^{n} y_i w_i}{\sum_{i=1}^{n} w_i}$

-Here, I substituted $\hat{\mu}$, which we found in part b with the value of $\hat{\mu}$ which I found from a series of equivalences.

$= \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$

-By setting $w_i = 1$ we could solve for $\hat{\lambda} = \bar{y}$

.

.

.

.

.

.

.

.

Question 2

Poisson Distribution

Introduction

The car insurance company has already implemented a Poisson regression model to describe the claim counts distribution. I have been hired to test whether the Poisson model should be replaced by a homogeneous negative binomial model. I have concluded that a poisson regression model better fits the data and should be used to describe the claim counts distribution.

Methods

Linders Consulting is implementing a Poisson regression model to describe the claim counts distribution. Because λ is the parameter of a Poisson distribution and is the value for both the mean and variance, it may be inaccurate to assume all samples follow the same λ variable. When the mean and variance are not equal in a counts distribution, there is an under or overdispersion, depending on if the variance is smaller or larger than the mean. The issue could be resolved by implementing a homogeneous negative binomial regression model. The variance is $\lambda(1 + \lambda/\gamma)$ and mean of λ.

Results

The homogeneous negative binomial has a θ of 1.4464912. The estimated values of dispersion for the Poisson and homogeneous negative binomial model are 0.996757 and 0.9727552, respectiveyl. They are both close to 1 and can be described as the residual being normal, however the Poisson dispersion is more normal than the negative binomial dispersion making it a better fit.

Conclusion

The Poisson distribution will be the better predicttion model.