STAT 520 Homework 5

Yifan Zhu

October 12, 2017

1. They are different. The gls estimator is the minimizer of

$$\sum_{i=1}^{2} \frac{[y_i - g(\boldsymbol{x}_i, \boldsymbol{\beta})]^2}{g_2^2(\boldsymbol{x}_i, \boldsymbol{\beta}, \theta)}$$

while the mle is the maximizer of log likelihood function $\ell(\boldsymbol{\beta})$, where

$$\ell(\boldsymbol{\beta}) = -\frac{1}{2} \sum_{i=1}^{n} \left(\log(2\pi g_2^2(\boldsymbol{x}_i, \boldsymbol{\beta}, \theta)\sigma^2) + \frac{1}{\sigma^2} \frac{[y_i - g(\boldsymbol{x}_i, \boldsymbol{\beta})]^2}{g_2^2(\boldsymbol{x}_i, \boldsymbol{\beta}, \theta)} \right)$$

We can not guarantee that the minimizer is the same as maximizer because of the difference in the term $\log(2\pi g_2^2(\boldsymbol{x}_i, \boldsymbol{\beta}, \theta)\sigma^2)$.

2. We obtain

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^{n} \left[\frac{Y_i - g_1(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}})}{g_2(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}, \theta)} \right]^2$$

Then estimate using `nonlin` as if σ is known. The 95% confidence interval is obtain by

$$\hat{\beta}_k \pm 1.96(\widehat{cov}(\hat{\boldsymbol{\beta}})_{kk})^{1/2}$$

where $\widehat{cov}\hat{\boldsymbol{\beta}}_{kk}$ is the (k, k) element of matrix $\widehat{cov}(\hat{\boldsymbol{\beta}})$ obtained from

$$\widehat{cov}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \left(\sum_{i=1}^{n} \boldsymbol{v}(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}) \boldsymbol{v}^T(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}) / g_2^2(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}, \theta) \right)^{-1}$$

in which $\boldsymbol{v}(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}})$ is a column vector and

$$v_j(\boldsymbol{x}_i, \hat{\boldsymbol{\beta}}) = \frac{\partial}{\partial \beta_j} g_1(\boldsymbol{x}_i, \boldsymbol{\beta}) \bigg|_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}}$$

Results are shown in the table below.

Parameter	Estimate	95% CI
β_1	20.062	(19.364, 20.760)
β_2	3.224	(2.921, 3.527)
β_3	0.539	(0.480, 0.598)
σ^2	0.210	

Next we estimate $u = \beta_2/\beta_3$. The estimated value $\hat{u} = \frac{\hat{\beta}_2}{\hat{\beta}_3} = 5.985$. In order to get the 95% confidence interval, we obtain its variance by Delta Method.

$$\frac{\partial u}{\partial \beta_1} = 0$$
$$\frac{\partial u}{\partial \beta_2} = \frac{1}{\beta_3}$$
$$\frac{\partial u}{\partial \beta_3} = -\frac{\beta_2}{\beta_3^2}$$

Then let

$$D = \begin{bmatrix} 0 & \frac{1}{\beta_3} & -\frac{\beta_2}{\beta_3^2} \end{bmatrix}$$

Then $\hat{D} = D|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}$, and

$$\widehat{Var}(\hat{\beta}_2/\hat{\beta}_3) = \hat{D}\widehat{cov}(\hat{\boldsymbol{\beta}})\hat{D}^T \Rightarrow SE = \sqrt{\widehat{Var}(\hat{\beta}_2/\hat{\beta}_3)} = 0.0957$$

Thus the 95% confidence interval is

$$(5.797, 6.172)$$

3. Using the Wald confidence interval as in **2**. Results are shown in the table below.

Parameter	Estimate	95% CI
β_1	20.064	(19.589, 20.539)
β_2	3.259	(3.082, 3.436)
β_3	0.544	(0.508, 0.579)
σ^2	0.202	

Next we estimate $u = \beta_2/\beta_3$ like what we did in **1**. The estimated value $\hat{u} = \frac{\hat{\beta}_2}{\hat{\beta}_3} = 5.995$. The 95% confidence interval by Delta Method is

$$(5.868, 6.123)$$

4. Let

$$\mu_i = \beta_1 \exp[-\exp(\beta_2 - \beta_3 x_i)], v_i = \sigma^2\{\mu_i\}^{2\theta}$$

Then the density of Y_i is

$$f_i(y_i|\mu_i, v_i) = \frac{1}{(2\pi v_i)^{1/2}} \exp\left[-\frac{1}{2v_i}(y_i - \mu_i)^2\right]$$

and the log likelihood function is

$$\ell_i = -\frac{1}{2}\log(2\pi v_i) - \frac{1}{2v_i}(y_i - \mu_i)^2$$

Because Y_i's are independent, then

$$\ell(\boldsymbol{\beta}, \sigma^2) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\beta}, \sigma^2)$$

By chain rule, the first derivatives are

$$\frac{\partial \ell_i}{\partial \beta_j} = \frac{\partial \ell_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j} + \frac{\partial \ell_i}{\partial v_i}\frac{\partial v_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}$$

and

$$\frac{\partial \ell_i}{\partial \sigma^2} = \frac{\partial \ell_i}{\partial v_i}\frac{\partial v_i}{\partial \sigma^2}$$

Then the second derivatives,

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \beta_j \partial \beta_k} &= \frac{\partial^2 \ell_i}{\partial \mu_i^2}\frac{\partial \mu_i}{\partial \beta_j}\frac{\partial \mu_i}{\partial \beta_k} + \frac{\partial^2 \ell}{\partial \mu_i \partial v_i}\frac{\partial v_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\frac{\partial \mu_i}{\partial \beta_k} + \frac{\partial \ell_i}{\partial \mu_i}\frac{\partial^2 \mu_i}{\partial \beta_j \partial \beta_k} \\
&+ \frac{\partial^2 \ell_i}{\partial \mu_i \partial v_i}\frac{\partial v_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\frac{\partial \mu_i}{\partial \beta_k} + \frac{\partial^2 \ell_i}{\partial v_i^2}\left(\frac{\partial v_i}{\partial \mu_i}\right)^2 \frac{\partial \mu_i}{\partial \beta_j}\frac{\partial \mu_i}{\partial \beta_k} + \frac{\partial \ell_i}{\partial v_i}\frac{\partial^2 v_i}{\partial \mu_i^2}\frac{\partial \mu_i}{\partial \beta_j}\frac{\partial \mu_i}{\partial \beta_k} \\
&+ \frac{\partial \ell_i}{\partial \mu_i}\frac{\partial v_i}{\partial \mu_i}\frac{\partial^2 \mu_i}{\partial \beta_j \partial \beta_k}
\end{aligned}$$

$$\frac{\partial^2 \ell_i}{\partial (\sigma^2)^2} = \frac{\partial^2 \ell_i}{\partial v_i^2}\left(\frac{\partial v_i}{\partial \sigma^2}\right)^2 + \frac{\partial \ell_i}{\partial v_i}\frac{\partial^2 v_i}{\partial (\sigma^2)^2}$$

$$\frac{\partial^2 \ell_i}{\partial \sigma^2 \partial \beta_j} = \frac{\partial^2 \ell}{\partial \mu_i \partial v_i}\frac{\partial \mu_i}{\partial \beta_j}\frac{\partial v_i}{\partial \sigma^2} + \frac{\partial^2 \ell}{\partial v_i^2}\frac{\partial v_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\frac{\partial v_i}{\partial \sigma^2} + \frac{\partial \ell_i}{\partial v_i}\frac{\partial^2 v_i}{\partial \sigma^2 \partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}$$

Each term in the chain rule is as below,

$$\begin{aligned}
\frac{\partial \ell_i}{\partial \mu_i} &= \frac{1}{v_i}(y_i - \mu_i) \\
\frac{\partial \ell_i}{\partial v_i} &= \frac{1}{2v_i}\left[\frac{1}{v_i}(y_i - \mu_i)^2 - 1\right] \\
\frac{\partial^2 \ell}{\partial \mu_i^2} &= -\frac{1}{v_i} \\
\frac{\partial^2 \ell}{\partial v_i^2} &= \frac{1}{v_i^2}\left[\frac{1}{2} - \frac{1}{v_i}(y_i - \mu_i)^2\right] \\
\frac{\partial^2 \ell}{\partial \mu_i \partial v_i} &= -\frac{1}{v_i^2}(y_i - \mu_i) \\
\frac{\partial v_i}{\partial \mu_i} &= 2\theta\sigma^2\mu_i^{2\theta-1} \\
\frac{\partial v_i}{\partial \sigma^2} &= \mu_i^{2\theta} \\
\frac{\partial^2 v_i}{\partial \mu_i^2} &= 2\theta\sigma^2(2\theta-1)\mu_i^{2\theta-2} \\
\frac{\partial^2 v_i}{\partial (\sigma^2)^2} &= 0 \\
\frac{\partial^2 v_i}{\partial \mu_i \partial \sigma^2} &= 2\theta\mu_i^{2\theta-1}
\end{aligned}$$

For derivatives of μ_i, for the simplicity of the expression, denote

$$\begin{aligned}
T_1 &= \exp[-\exp(\beta_2 - \beta_3 x_i)] \\
T_2 &= \exp(\beta_2 - \beta_3 x_i)
\end{aligned}$$

then

$$\frac{\partial \mu_i}{\partial \beta_1} = T_1$$
$$\frac{\partial \mu_i}{\partial \beta_2} = -\beta_1 T_1 T_2$$
$$\frac{\partial \mu_i}{\partial \beta_3} = \beta_1 x_i T_1 T_2$$
$$\frac{\partial^2 \mu_i}{\partial \beta_1^2} = 0$$
$$\frac{\partial^2 \mu_i}{\partial \beta_1 \partial \beta_2} = -T_1 T_2$$
$$\frac{\partial^2 \mu_i}{\partial \beta_1 \partial \beta_3} = x_i T_1 T_2$$
$$\frac{\partial^2 \mu_i}{\partial \beta_2^2} = \beta_1 T_1 T_2 (T_2 - 1)$$
$$\frac{\partial^2 \mu_i}{\partial \beta_2 \partial \beta_3} = \beta_1 x_i T_1 T_2 (1 - T_2)$$
$$\frac{\partial^2 \mu_i}{\partial \beta_3^2} = \beta_1 x_i^2 T_1 T_2 (T_2 - 1)$$

5. Evaluate $\nabla \ell = (\frac{\partial \ell}{\partial \beta_1}, \frac{\partial \ell}{\partial \beta_2}, \frac{\partial \ell}{\partial \beta_3}, \frac{\partial \ell}{\partial \sigma^2})$ by chain rule and formulas we obtained in **4** at gls estimates, we have

$$\nabla \ell = (-0.07476674, 8.76568528, -34.43197118, 149.15362379)$$

Evaluate $\nabla \ell$ at mle, we have

$$\nabla \ell = (-1.324440 \times 10^{-6}, 1.726760 \times 10^{-6}, -3.365130 \times 10^{-5}, -6.110668 \times 10^{-6})$$

6. Incorporate θ as a parameter and do mle again, we obtain the estimates and 95% confidence for $\boldsymbol{\beta}$ as shown in the table below.

Parameter	Estimate	95% CI
β_1	20.147	(19.640, 20.654)
β_2	3.141	(2.978, 3.304)
β_3	0.523	(0.489, 0.557)
σ^2	0.107	
θ	0.624	

7. Let

$$\ell_p(\theta) = \max_{\boldsymbol{\beta}, \sigma^2} \ell(\boldsymbol{\beta}, \sigma^2, \theta)$$

Then the 95% confidence interval is obtained from

$$\{\theta : -2(\ell_p(\theta) - \ell_p(\hat{\theta})) \le \chi^2_{1,0.95}\}$$

Pick the endpoints when the equality holds, then the confidence interval for θ is

$$(0.497, 0.730)$$