Homework Solutions

Royou Kiku

December 20th, 2021

1.

(1)

The binomial distribution is given by:

$$f(x) = \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}$$

The likelihood function is

$$L(\beta) = \prod_{i=1}^{T} \left[\frac{n!}{y_i!\,(n-y_i)!} (F(X_i\beta))^{y_i} (1-F(X_i\beta))^{n-y_i} \right]$$

(2)

$$logL(\beta) = \sum_{i=1}^{T} \left(log\frac{n!}{y_i!\,(n-y_i)!} + y_i logF(X_i\beta) + (n-y_i)log(1-F(X_i\beta)) \right)$$

The first order condition is

$$\frac{\partial logL(\beta)}{\partial \beta} = \sum_{i=1}^{T} \left(\frac{y_i X'_i f(X_i\beta)}{F(X_i\beta)} - \frac{(n-y_i)X'_i f(X_i\beta)}{1-F(X_i\beta)} \right) = 0$$

2.

(3)

$$E[y_i] = 1 \times P(y_i = 1) + 0 \times P(y_i = 0) = P(y_i = 1)$$

(4)

If in linear regression model (y_i is a continuous type), when $E[u_i] = 0$, we can get $E[y_i] = X_i\beta$. And $X_i\beta$ takes the value from $-\infty$ to ∞.

But in this case, $E[y_i]$ indicates the ratio of people who answer Yes out of all the people, and $E[y_i]$ has to be between zero to one. Therefore, $E[y_i]$ is not approximate as $X_i\beta$.

(5)

$$\begin{aligned} f(y_i) &= (P(y_i = 1))^{y_i}(1 - P(y_i = 1))^{1-y_i}, \qquad y_i = 0,1 \\ &= (F(X_i\beta))^{y_i}(1-F(X_i\beta))^{1-y_i} \end{aligned}$$

$$L(\beta) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} (F(X_i\beta))^{y_i}(1-F(X_i\beta))^{1-y_i}$$

$$logL(\beta)=\sum_{i=1}^{n}\{y_i logF(X_i\beta)+(1-y_i)\log(1-F(X_i\beta)\}$$

$$\frac{\partial logL(\beta)}{\partial\beta}=\sum_{i=1}^{n}\left(\frac{y_iX'_if(X_i\beta)}{F(X_i\beta)}-\frac{(1-y_i)X'_if(X_i\beta)}{1-F(X_i\beta)}\right)=\sum_{i=1}^{n}\frac{X'_if(X_i\beta)(y_i-F(X_i\beta))}{F(X_i\beta)(1-F(X_i\beta))}=0$$

$$\frac{\partial^2 logL(\beta)}{\partial\beta\partial\beta'}=\sum_{i=1}^{n}\frac{X'_iX_if'(X_i\beta)\big(y_i-F(X_i\beta)\big)}{F(X_i\beta)\big(1-F(X_i\beta)\big)}-\sum_{i=1}^{n}\frac{X'_iX_if^2(X_i\beta)}{F(X_i\beta)\big(1-F(X_i\beta)\big)}$$

$$+\sum_{i=1}^{n}X'_i f(X_i\beta)(y_i-F(X_i\beta))\frac{X_if(X_i\beta)(1-2F(X_i\beta))}{\big(F(X_i\beta)(1-F(X_i\beta))\big)^2}$$

For maximization, the method of scoring is given by:

$$\beta^{(j+1)}=\beta^{(j)}+\left(-E\left(\frac{\partial^2 logL(\beta^{(j)})}{\partial\beta\partial\beta'}\right)\right)^{-1}\frac{\partial logL(\beta^{(j)})}{\partial\beta}$$

$$\beta^{(j+1)}=\beta^{(j)}+\left(\sum_{i=1}^{n}\frac{X'_iX_i\big(f(X_i\beta)^{(j)}\big)^2}{F(X_i\beta)^{(j)}(1-F(X_i\beta)^{(j)})}\right)^{-1}\sum_{i=1}^{n}\frac{X'_if(X_i\beta)^{(j)}(y_i-F(X_i\beta)^{(j)})}{F(X_i\beta)^{(j)}(1-F(X_i\beta)^{(j)})}$$

3.

(6)

$$\int L(\theta;x)dx=1$$

$$\frac{\partial}{\partial\theta}\int L(\theta;x)dx=\int\frac{\partial L(\theta;x)}{\partial\theta}dx=0$$

$$\int\frac{\partial L(\theta;x)}{\partial\theta}dx=\int\frac{\partial logL(\theta;x)}{\partial\theta}L(\theta;x)dx=0$$

Suppose $\int g(x)L(\theta;x)dx=E[g(X)]$, we can get

$$\int\frac{\partial logL(\theta;x)}{\partial\theta}L(\theta;x)dx=E\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)=0$$

Here, we prove the first equation.

Then, we differentiate the first equation by θ.

$$\frac{\int\frac{\partial logL(\theta;x)}{\partial\theta}L(\theta;x)dx}{\partial\theta}=\int\frac{\partial^2 logL(\theta;x)}{\partial\theta\partial\theta'}L(\theta;x)dx+\int\frac{\partial logL(\theta;x)}{\partial\theta}\frac{\partial L(\theta;x)}{\partial\theta'}dx=0$$

$$\int\frac{\partial^2 logL(\theta;x)}{\partial\theta\partial\theta'}L(\theta;x)dx+\int\frac{\partial logL(\theta;x)}{\partial\theta}\frac{\partial logL(\theta;x)}{\partial\theta'}L(\theta;x)dx=0$$

$$E\left(\frac{\partial^2 logL(\theta;X)}{\partial\theta\partial\theta'}\right)+E\left(\frac{\partial logL(\theta;X)}{\partial\theta}\frac{\partial logL(\theta;X)}{\partial\theta'}\right)=0$$

$$E\left(\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)^2\right)=-E\left(\frac{\partial^2 logL(\theta;X)}{\partial\theta^2}\right)$$

Because $E\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)=0,$ we can get

$$-E\left(\frac{\partial^2 logL(\theta;X)}{\partial\theta^2}\right)=E\left(\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)^2\right)=V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

Finally, we prove the second equation.

(7)

$$E\big(s(X)\big)=\int s(x)L(\theta;x)dx$$

$$\frac{\partial E(s(X))}{\partial\theta}=\int s(x)\frac{\partial L(\theta;x)}{\partial\theta}dx$$

$$=\int s(x)\frac{\partial Log(\theta;x)}{\partial\theta}L(\theta;x)dx$$

$$=E\left(s(X)\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

$$=COV\left(s(X),\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

Because $E\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)=0,$ we can get $E\left(s(X)\frac{\partial logL(\theta;X)}{\partial\theta}\right)=COV\left(s(X),\frac{\partial logL(\theta;X)}{\partial\theta}\right).$

$$\left(\frac{\partial E(s(X))}{\partial\theta}\right)^2=\left(COV\left(s(X),\frac{\partial logL(\theta;X)}{\partial\theta}\right)\right)^2$$

$$=\frac{\left(COV\left(s(X),\frac{\partial logL(\theta;X)}{\partial\theta}\right)\right)^2}{\left(\sqrt{V(s(X))}\sqrt{V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)}\right)^2}V\big(s(X)\big)V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

$$=\rho^2V\big(s(X)\big)V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

Since ρ denotes the correlation coefficient, we know that $|\rho|\le 1.$

$$\rho^2\le 1$$

$$\rho^2V\big(s(X)\big)V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)\le V\big(s(X)\big)V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

$$\left(\frac{\partial E(s(X))}{\partial\theta}\right)^2\le V\big(s(X)\big)V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)$$

$$V\big(s(X)\big)\ge\frac{\left(\frac{\partial E(s(X))}{\partial\theta}\right)^2}{V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)}$$

When $E\big(s(X)\big)=\theta,$ we can get

$$V\big(s(X)\big)\ge\frac{1}{V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)}=\frac{1}{-E\left(\frac{\partial^2 logL(\theta;X)}{\partial\theta^2}\right)}=I(\theta)^{-1}$$

$$V(s(X))\ge I(\theta)^{-1}$$

(8)

$$\max \quad logL(\theta; x)$$

$$\frac{\partial logL(\theta; x)}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial logf(x_i; \theta)}{\partial\theta} = 0$$

Replacing x_i by the underlying random variable X_i, and applying Central Limit Theorem II, we can get

$$\frac{\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i; \theta)}{\partial\theta} - E\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i; \theta)}{\partial\theta}\right)}{\sqrt{V\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i; \theta)}{\partial\theta}\right)}}$$

$$= \frac{\frac{1}{n}\frac{\partial logL(\theta; X)}{\partial\theta} - E\left(\frac{1}{n}\frac{\partial logL(\theta; X)}{\partial\theta}\right)}{\sqrt{V\left(\frac{1}{n}\frac{\partial logL(\theta; X)}{\partial\theta}\right)}} \longrightarrow N(0,1)$$

Note that $E\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i;\theta)}{\partial\theta}\right) = E\left(\frac{1}{n}\frac{\partial logL(\theta;X)}{\partial\theta}\right) = 0, V\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i;\theta)}{\partial\theta}\right) = V\left(\frac{1}{n}\frac{\partial logL(\theta;X)}{\partial\theta}\right) = \frac{1}{n^2} I(\theta)$.

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i; \theta)}{\partial\theta} - E\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\partial logf(X_i; \theta)}{\partial\theta}\right)\right)$$

$$= \sqrt{n}\left(\frac{1}{n}\frac{\partial logL(\theta; X)}{\partial\theta} - E\left(\frac{1}{n}\frac{\partial logL(\theta; X)}{\partial\theta}\right)\right)$$

$$= \frac{1}{\sqrt{n}}\frac{\partial logL(\theta; X)}{\partial\theta} \longrightarrow N(0, \frac{1}{n} I(\theta))$$

Note that $E\left(\frac{1}{\sqrt{n}}\frac{\partial logL(\theta;X)}{\partial\theta}\right) = 0, V\left(\frac{1}{\sqrt{n}}\frac{\partial logL(\theta;X)}{\partial\theta}\right) = \frac{1}{n} I(\theta)$.

Now replacing θ by $\tilde{\theta}$, the asymptotic distribution of $\frac{1}{\sqrt{n}}\frac{\partial logL(\tilde{\theta};X)}{\partial\theta}$, which is expanded around $\tilde{\theta} = \theta$ as follows:

$$0 = \frac{1}{\sqrt{n}}\frac{\partial logL(\tilde{\theta}; X)}{\partial\theta} \approx \frac{1}{\sqrt{n}}\frac{\partial logL(\theta; X)}{\partial\theta} + \frac{1}{\sqrt{n}}\frac{\partial^2 logL(\theta; X)}{\partial\theta\partial\theta'}(\tilde{\theta} - \theta)$$

$$-\frac{1}{\sqrt{n}}\frac{\partial^2 logL(\theta; X)}{\partial\theta\partial\theta'}(\tilde{\theta} - \theta) \approx \frac{1}{\sqrt{n}}\frac{\partial logL(\theta; X)}{\partial\theta}$$

$$\sqrt{n}\left(-\frac{1}{n}\frac{\partial^2 logL(\theta; X)}{\partial\theta\partial\theta'}\right)(\tilde{\theta} - \theta) \approx \frac{1}{\sqrt{n}}\frac{\partial logL(\theta; X)}{\partial\theta}$$

$$\sqrt{n}(\tilde{\theta} - \theta) \approx \left(-\frac{1}{n}\frac{\partial^2 logL(\theta; X)}{\partial\theta\partial\theta'}\right)^{-1}\left(\frac{1}{\sqrt{n}}\frac{\partial logL(\theta; X)}{\partial\theta}\right)$$

Assuming $\sigma^2 = \lim_{n\to\infty}\left(\frac{I(\theta)}{n}\right)^{-1}$ and using the law of large number, we can get

$$-\frac{1}{n}\frac{\partial^2 logL(\theta;X)}{\partial\theta\partial\theta'} \longrightarrow \lim_{n\to\infty}\frac{1}{n}\left(-E\left(\frac{\partial^2 logL(\theta;X)}{\partial\theta\partial\theta'}\right)\right)$$

$$= \lim_{n\to\infty}\frac{1}{n}\left(V\left(\frac{\partial logL(\theta;X)}{\partial\theta}\right)\right) = \lim_{n\to\infty}\frac{1}{n}I(\theta) = (\sigma^2)^{-1}$$

Finally, we can get

$$\sqrt{n}(\tilde{\theta}-\theta) \longrightarrow N(0,\sigma^2(\sigma^2)^{-1}\sigma^2) = N(0,\sigma^2) = N(0,\lim_{n\to\infty}\left(\frac{I(\theta)}{n}\right)^{-1})$$