Statistical Inference

Lecture 07a

ANU - RSFAS

Last Updated: Sun Apr 15 13:50:52 2018

Some Asymptotics (MLE) - Score Function

Lemma: Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x;\theta)$ and let $\hat{\theta}$ be the MLE of θ. Under regularity conditions of $f(x;\theta)$ and thus $L(\theta;\boldsymbol{x})$ (under appropriate smoothness conditions), we can state:

$$W = \frac{1}{\sqrt{n}}\ell'(\theta;\boldsymbol{x}) \xrightarrow{D} \mathrm{normal}(0, i(\theta))$$

Proof:

$$\frac{\ell'(\theta;\boldsymbol{x})}{\sqrt{n}} = \frac{\sum_{i=1}^{n}\ell'(\theta;x_i)}{\sqrt{n}} = \frac{\frac{n}{n}\sum_{i=1}^{n}\ell'(\theta;x_i)}{\sqrt{n}} = \sqrt{n}\ \bar{\ell'}$$

- $\bar{\ell'}$ is the sample average of the first derivative of the log likelihood.

MLEs - Asymptotics

- We can use the Central Limit theorem! We need to know the mean and variance of $\bar{\ell}'$

$$
\begin{aligned}
E[\bar{\ell}'] &= E\left[\frac{1}{n}\sum_{i=1}^{n}\ell'(\theta;x_i)\right] = E[\ell'(\theta;x_i)] \\
&= \int_{-\infty}^{\infty}\ell'(\theta;x_i)f(x_i;\theta)dx_i \\
&= \int_{-\infty}^{\infty}\boxed{\frac{\frac{\partial}{\partial\theta}f(x_i;\theta)}{\cancel{f(x_i;\theta)}}}\cancel{f(x_i;\theta)}dx_i \\
&= \int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}f(x_i;\theta)dx_i \\
&= \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}f(x_i;\theta)dx_i \\
&= \frac{\partial}{\partial\theta}1 = 0
\end{aligned}
$$

MLEs - Asymptotics

$V(X) = E(X^2) - [E(X)]^2$

$E(X) = 0$

$$V[\bar{\ell}'] = \frac{1}{n}V[\ell'(\theta; x_i)] = \frac{1}{n}E[\{\ell'(\theta; x_i)\}^2] = -\frac{1}{n}E[\ell''(\theta; x_i)] = \frac{1}{n}i(\theta)$$

- So let's subtract off the mean and divide by the standard deviation:

$$\frac{(\bar{\ell}' - 0)}{\sqrt{i(\theta)/n}} = \frac{\sqrt{n}(\bar{\ell}' - 0)}{\sqrt{i(\theta)}} = \frac{\frac{\ell'(\theta; \boldsymbol{x})}{\sqrt{n}}}{\sqrt{i(\theta)}} \xrightarrow{D} \text{normal}(0, 1)$$

- So

$$\frac{\ell'(\theta; \boldsymbol{x})}{\sqrt{n}} \xrightarrow{D} \text{normal}(0, i(\theta))$$

MLEs - Asymptotics

Lemma 3.3: Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x;\theta)$. Let $\hat{\theta}$ be the MLE of θ. Under regularity conditions of $f(x;\theta)$ and thus $L(\theta;\boldsymbol{x})$ (under appropriate smoothness conditions), we have:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} \text{normal}(0, i(\theta)^{-1})$$

asymptotically unbiased & CR

& consistent (weakly)

MLEs - Asymptotics

Proof:

- Conduct a Taylor's series expansion of the first derivative of the log likelihood around the true value θ_0:

$$\ell'(\theta; \boldsymbol{x}) = \ell'(\theta_0; \boldsymbol{x}) + (\theta - \theta_0)\ell''(\theta_0; \boldsymbol{x}) + \cdots$$

- Substitute the MLE for θ:

$$\overset{0}{\overset{\shortparallel}{\ell'}}(\hat{\theta}; \boldsymbol{x}) = \ell'(\theta_0; \boldsymbol{x}) + (\hat{\theta} - \theta_0)\ell''(\theta_0; \boldsymbol{x}) + \cdots$$

- Under the regularity conditions we will ignore higher order terms. Also we know $\ell'(\hat{\theta}; \boldsymbol{x}) = 0$:

$$0 = \ell'(\theta_0; \boldsymbol{x}) + (\hat{\theta} - \theta_0)\ell''(\theta_0; \boldsymbol{x})$$

- Now, replace $\ell''(\theta_0; \boldsymbol{x})$ with its expectation:

$$
\begin{aligned}
0 &= \ell'(\theta_0; \boldsymbol{x}) + (\hat{\theta} - \theta_0) E[\ell''(\theta_0; \boldsymbol{x})] \\
&= \ell'(\theta_0; \boldsymbol{x}) + (\hat{\theta} - \theta_0) E\left[\sum_{i=1}^{n} \ell''(\theta_0; x_i)\right] \\
&= \ell'(\theta_0; \boldsymbol{x}) + (\hat{\theta} - \theta_0) \sum_{i=1}^{n} E\left[\ell''(\theta_0; x_i)\right] \\
&= \ell'(\theta_0; \boldsymbol{x}) + (\hat{\theta} - \theta_0)[-ni(\theta_0)]
\end{aligned}
$$

$ni(\theta) = I(\theta)$

$$
\Rightarrow (\hat{\theta} - \theta_0) = \frac{-\ell'(\theta_0; \boldsymbol{x})}{-ni(\theta_0)}
$$

- Note: $\frac{1}{n}\ell''(\theta_0; \boldsymbol{x}) \overset{\text{LLN}}{\rightarrow} E[\frac{1}{n}\ell''(\theta_0; \boldsymbol{x})] = -i(\theta)$

MLEs - Asymptotics

- Multiply through by $\sqrt{n}$:

$$\begin{aligned}\sqrt{n}(\hat{\theta} - \theta_0) &= \sqrt{n}\frac{\ell'(\theta_0; \boldsymbol{x})}{ni(\theta_0)} = \sqrt{n}\frac{\ell'(\theta_0; \boldsymbol{x})}{\mathbf{I}(\theta_0)} \\ &= \frac{\frac{1}{\sqrt{n}}\ell'(\theta_0; \boldsymbol{x})}{\frac{1}{n}\mathbf{I}(\theta_0)} = \frac{\frac{1}{\sqrt{n}}l'(\theta_0; \boldsymbol{x})}{i(\theta_0)}\end{aligned}$$

- Now we saw that:

$$W = \frac{1}{\sqrt{n}}\ell'(\theta; \boldsymbol{x}) \xrightarrow{D} \mathrm{normal}(0, i(\theta))$$

- Since a linear transformation of a normal is normal, we just need the mean and variance:

$$E\left[\frac{W}{i(\theta_0)}\right] = \frac{E[W]}{i(\theta)} = \frac{0}{i(\theta)} = 0$$

$$V\left[\frac{W}{i(\theta_0)}\right] = \frac{V[W]}{i(\theta)^2} = \frac{i(\theta)}{i(\theta)^2} = \frac{1}{i(\theta)}$$

- So we have:

C-R LB

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} \mathrm{normal}(0, i(\theta)^{-1})$$

Or

$$\hat{\theta} \,\dot{\sim}\, n\left(\theta, \frac{1}{ni(\theta)}\right) = \mathrm{normal}(\theta, \mathbf{I}(\theta)^{-1})$$

Delta Method

Theorem: Let Y_n be a sequence of random variables such that:

$$\sqrt{n}(Y_n - \theta) \xrightarrow{D} \mathrm{normal}(0, \sigma^2)$$

- For a given function g and a specific value θ, suppose that $g'(\theta)$ exists and is not 0, then

$$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{D} \mathrm{normal}(0, \sigma^2[g'(\theta)]^2)$$

MLEs - Asymptotics

- We can extend the theorem to functions $\tau(\theta)$:

Lemma: Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x;\theta)$. Let $\hat{\theta}$ be the MLE of θ and let *tau*(θ) be a continuous function of θ. Under regularity conditions (i.e. under appropriate smoothness conditions) of $f(x;\theta)$ and thus $L(\theta;\boldsymbol{x})$, we have:

$$\sqrt{n}(\tau(\hat{\theta}) - \tau(\theta)) \overset{D}{\to} \text{normal}(0, \nu(\theta))$$

- Where $\nu(\theta) = \frac{[\tau'(\theta)]^2}{i(\theta)}$ is the Cramer-Rao lower bound for a single data point.

Or

$$\tau(\hat{\theta}) \mathrel{\dot{\sim}} \text{normal}\left(\tau(\theta), \frac{[\tau'(\theta)]^2}{\mathbf{I}(\theta)}\right)$$

- We can get this result from the Delta method!

MLEs - Asymptotics

- So asymptotically, MLEs are:

1. unbiased;
2. achieve the Cramer-Rao lower bound (efficient);
3. asymptotically normally distributed.

- We can also note that MLEs are consistent estimators.
- Because these estimators achieve (1-3) they are *asymptotically efficient! **best asymptotically normal (BAN) estimators**