2-24-09 Math 152

Last time

Large Sample Theory for MLE

When f is smooth ($f(X|\theta)$).

$$\sqrt{nI(\theta_0)}\,(\hat{\theta} - \theta_0) \longrightarrow N(0,1)$$

in distribution.

where $$I(\theta) = E\left(\left[\frac{\partial}{\partial\theta} \log f(X|\theta)\right]^2\right)$$

$$= -E\left[\frac{\partial^2}{\partial\theta^2} \log f(X|\theta)\right].$$

and ($\therefore$)

$$nI(\theta_0) = -E(\ell''(\theta_0)).$$

where ℓ is the log-likelihood.

i.e.

The asymptotic variance of

$$\hat{\theta} \text{ is } \frac{1}{nI(\theta_0)} = -\frac{1}{E(\ell''(\theta_0))}$$

e.g. $\hat{\lambda} = \overline{X}$

for $X_1, \dots, X_n$ i.i.d. Poisson

$$f(x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}.$$

$$I(\lambda) = E\left(\frac{\partial}{\partial\lambda} \log f(X|\lambda)\right)^2$$

$$= E\left(\frac{d}{d\lambda}(X \log \lambda - \log x! - \lambda)\right)^2$$

$$= E\left(\frac{X}{\lambda} - 1\right)^2$$

$$= E\left(\frac{X^2}{\lambda^2} - 2\frac{X}{\lambda} + 1\right)$$

$$= \frac{1}{\lambda^2}(E(X^2)) - \frac{2}{\lambda}E(X) + 1.$$

$$= \frac{1}{\lambda^2}(\lambda^2 + \lambda) - \frac{2}{\lambda}\cdot\lambda + 1$$

$$= 1 + \frac{1}{\lambda} - 2 + 1 = \frac{1}{\lambda}.$$

alternatively

$$\frac{\partial^2}{\partial\lambda^2}(\log f(X|\lambda)) = -\frac{X}{\lambda^2}$$

so $I(\lambda) = -E\left[\frac{\partial^2}{\partial\lambda^2} \log f(X|\lambda)\right] = \frac{E(X)}{\lambda^2} = \frac{1}{\lambda}.$

The asymptotic variance is

$$\frac{1}{n I(\lambda_0)} = \frac{\lambda_0}{n}.$$

We don't know λ_0 so we use $\hat{\lambda} = \bar{X}$ to get a $100(1-\alpha)\%$ confidence interval.

$$\bar{X} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\bar{X}}{n}}$$

(We obtained this earlier by simply approximating large parameter Poisson distribution ~~variables~~ by the Normal).

It can be shown that the result

$$\mathrm{Var}(\hat{\theta}) \approx -\frac{1}{E(l''(\theta_0))} = \frac{1}{E[l'(\theta_0)^2]}$$

remains valid when the parameter is estimated from random multinomial counts. (Even though the X_i are not i.i.d.)

and that the distribution of

$\hat{\theta}$ is asymptotically approx. normal.

e.g. Hardy-Weinberg for

Hong Kong Blood type data.

re call:

$$f(x_1, \dots, x_m / p_1, \dots, p_m) = \frac{n!}{x_1! \dots x_m!} p_1^{x_1} \dots p_m^{x_m}$$

so $l(p_1, \dots, p_m) = \log n! - \sum_{i=1}^{m} \log x_i! + \sum_{i=1}^{m} x_i \log p_i$

and if $p_i = p_i(\theta)$

for some parameter θ

then

$$l(\theta) = \log n! - \sum_{i=1}^{m} \log x_i! + \sum_{i=1}^{m} x_i \log p_i(\theta)$$

In the Hong-Kong Blood type example

	M	MN	N	Total
Frequency	342	500	187	1029

We have

$$p_1(\theta) = (1-\theta)^2$$
$$p_2(\theta) = 2\theta(1-\theta)$$
$$p_3(\theta) = \theta^2$$

by the "Hardy-Weinberg Law".

So

$$\ell(\theta) = \log n! - \sum_{i=1}^{3} \log X_i! + X_1 \log (1-\theta)^2 + X_2 \log 2\theta(1-\theta) + X_3 \log \theta^2.$$

$$\ell'(\theta) = -\frac{2X_1}{1-\theta} + \frac{X_2}{\theta} - \frac{X_2}{1-\theta} + \frac{2X_3}{\theta}.$$

$$= -\frac{2X_1 + X_2}{1-\theta} + \frac{X_2 + 2X_3}{\theta}.$$

and

$$\ell''(\theta) = -\frac{2X_1 + X_2}{(1-\theta)^2} - \frac{2X_3 + X_2}{\theta^2}.$$

So that

$$E(\ell''(\theta)) = -\frac{(2E(X_1) + E(X_2))}{(1-\theta)^2} - \frac{2E(X_3) + E(X_2)}{\theta^2}.$$

Now,

$$X_1 \sim \text{bin}((1-\theta)^2, n)$$

$$X_2 \sim \text{bin}(2\theta(1-\theta), n)$$

$$X_3 \sim \text{bin}(\theta^2, n)$$

So

$$E(X_1) = n(1-\theta)^2$$

$$E(X_2) = 2n\theta(1-\theta)$$

$$E(X_3) = n\theta^2.$$

and

$$E(\ell''(\theta)) = -\frac{(2n(1-\theta)^2 + 2n\theta(1-\theta))}{(1-\theta)^2}$$

$$-\frac{(2n\theta^2 + 2n\theta(1-\theta))}{\theta^2}$$

$$= (\text{algebra}) = -\frac{2n}{\theta(1-\theta)}.$$

$\therefore$

$$Var(\hat{\theta}) \approx \frac{\theta_0(1-\theta_0)}{2n}.$$

Since we don't know θ_0, our approximation to the standard error is $\quad s_{\hat{\theta}} = \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{2n}}$

In the earlier MLE example with this data, we found

$$\hat{\theta} \approx .4247.$$

so $$S_{\hat{\theta}} = \sqrt{\frac{.4247(1-.4247)}{2(1029)}} \approx .011.$$

This is also the value ~~of~~ we obtained in class for the bootstrap approximation to the standard error.

An approximate 95% confidence interval for θ

is $$\hat{\theta} \pm 1.96\, S_{\hat{\theta}} \approx (.403, .447)$$