Stat 543 4-18-05

Recall Using the large sample dsn of "the MLE" to make confidence ellipsoids for a k-dimensional parameter vector ... or a part thereof

Basic Fact: $Y \sim N_k(\mu, \Sigma) \Rightarrow (Y-\mu)'\Sigma^{-1}(Y-\mu) \sim \chi^2_k$

Application(s): General $\{\mu \mid (y-\mu)'\Sigma^{-1}(y-\mu) < c\}$

serves as a confidence set for μ

To MLE's $\theta \in \mathbb{R}^k$, $\hat{\theta}_n$ is a nicely-behaved "MLE"

$\underset{k \times k}{I_1(\theta)}$ the FI in X_1

now apply the general facts to $\hat{\theta}_n$

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathcal{L}_\theta} N_k\left(0, I_1^{-1}(\theta)\right)$$

So

$$\underbrace{\underbrace{\sqrt{n}\left(\hat{\theta}_n - \theta\right)'\left(I_1(\theta)^{-1}\right)^{-1}\sqrt{n}\left(\hat{\theta}_n - \theta\right)}_{\left(\hat{\theta}_n - \theta\right)'\left(\frac{1}{n}I_1(\theta)^{-1}\right)^{-1}\left(\hat{\theta}_n - \theta\right)}}_{\left(\hat{\theta}_n - \theta\right)' nI_1(\theta)\left(\hat{\theta}_n - \theta\right)} \xrightarrow{\mathcal{L}_\theta} \chi^2_k$$

So an (unusable) confidence set for θ is

$$\left\{\theta \;\middle|\; \left(\hat{\theta}_n - \theta\right)'\left(nI_1(\theta)\right)\left(\hat{\theta}_n - \theta\right) < c\right\}$$

upper α pt of χ^2_k

this involves the unknown $I_1(\theta)$ ... there are 2 fixes

1) use of the "expected FI" – replace $I_1(\theta)$ with $I_1(\hat{\theta}_n)$

2) use of the "observed FI" ... think: with

$$\underset{k\times k}{H_n(\theta)} = \left(\frac{\partial^2 l_n(\theta)}{\partial\theta_i \partial\theta_j} \right)$$

the Hessian matrix for the log-likelihood

$H_n(\theta)$ is a sum of iid terms so

$$\frac{1}{n} H_n(\theta) \xrightarrow{P_\theta} -I_1(\theta) \qquad (\text{LLN})$$

and further it is at least plausible that

$$-\frac{1}{n} H_n(\hat{\theta}_n) \xrightarrow{P_\theta} I_1(\theta)$$

So replace $n I_1(\theta)$ in my confidence ellipsoid formula with

$$n\left(-\frac{1}{n} H_n(\hat{\theta}_n)\right) = -H_n(\hat{\theta}_n)$$

to get

$$\left\{ \theta \,\middle|\, (\hat{\theta}_n - \theta)\left(-H_n(\hat{\theta}_n)\right)(\hat{\theta}_n - \theta) < c \right\}$$

as a confidence ellipsoid for θ

Think for $\hat{\theta}_n$

covariance matrix of approximating dsn

$$\| $$

$$I^{-1}(\theta) = \left(n I_1(\theta)\right)^{-1}$$

for iid models

$$\approx \left(n I_1(\hat{\theta}_n) \right)^{-1}$$

$$\approx \left(-H_n(\hat{\theta}_n) \right)^{-1}$$

Need to be careful here about what this says for subvector of $\hat{\theta}_n$

$$\underset{k\times 1}{\theta} = \begin{pmatrix} \underset{l\times 1}{\theta_1} \\ \underset{(k-l)\times 1}{\theta_2} \end{pmatrix}$$ and same partitioning for $\hat{\theta}_n$

covariance matrix for an approximating dsn for $\hat{\theta}_{n1}$

upper left $l \times l$ block of $I^{-1}(\theta) = \left(n I_1(\theta) \right)^{-1}$

upper left $\ell \times \ell$ block of $\left(nI_1(\hat{\theta}_n) \right)^{-1}$

upper left $\ell \times \ell$ block of $\left(-H_n(\hat{\theta}_n) \right)^{-1}$

and these blocks are <u>not</u> just the inverses of the upper left $\ell \times \ell$ blocks of matrices inside $(\cdot)^{-1}$

In multiparameter cases as in single parameter cases folklore says that use of "observed FI" is better than "expected FI" in terms of holding nominal coverage probabilities — something that works even better (according to folk lore) is a method based on large sample dsn of "LRT statistic"

Theorem 9 on handout ($\theta \in \mathbb{R}^1$)

Under appropriate conditions in an iid model if $\{\delta_n(X)\}$ is consistent for θ at θ_0 and with θ_0 probability approaching 1 is a root of the likelihood equation

$$l_n'(\theta) = 0$$

then

$$2\left(l_n(\delta_n(X)) - l_n(\theta_0)\right) \xrightarrow{\mathcal{L}_{\theta_0}} \chi_1^2$$

loglikelihood at "MLE"

loglikelihood evaluated at truth