Stat 543 4-27-05

Recall

"Result" In an iid model where a prior for $\theta \in R^1$ has a density that is cont$^{\underline{s}}$ and positive at θ_0 and regularity conditions hold, if $\delta_n(X)$ is the MLE of θ, under the θ_0 dsn for X the posterior density of

$$\sqrt{-l_n''(\delta_n(X))}\,\left(\theta - \delta_n(X)\right)$$

"converges" to the std normal density $\left(\frac{1}{\sqrt{2\pi}} \exp - \frac{\theta^2}{2}\right)$

i.e. a posterior density starts to look like a normal density with

$$\text{mean } \delta_n(X) \quad \text{and} \quad \text{variance } \frac{1}{-\ell_n''(\delta_n(X))}$$

observed FI in X about θ

(This really comes from the fact that $L_n(\theta)$ tends to be peaked near θ_0 and take on the shape of a normal pdf ...)

<u>Example</u> $X_1, X_2, \ldots, X_n$ iid Bernoulli(p)

and prior $\quad p \sim \text{Beta}(\alpha, \beta)$

posterior for $p \mid X_1, \ldots, X_n$ is

$$\text{Beta}\left(\alpha + \sum X_i,\ \beta + \left(n - \sum X_i\right)\right)$$

a random dsn with a therefore random pdf

What happens to this posterior as $n \to \infty$ for a $p \in (0,1)$?

Note $\quad \delta_n(X) = \overline{X}_n \qquad -\ell_n''(\overline{X}_n) = \dfrac{n}{\overline{X}_n(1-\overline{X}_n)}$

The claim is that for $X_1, \ldots, X_n$ iid Bernoulli $p_0 \in (0,1)$ and

$$r \sim \text{Beta}\left(\alpha + \sum X_i,\ \beta + \left(n - \sum X_i\right)\right)$$

the pdf for

$$\sqrt{\frac{n}{\bar{X}_n(1-\bar{X}_n)}}\left(r-\bar{X}_n\right)$$

tends to look like a std normal density — i.e. with p_0 probability 1 I get an observed sequence of $X_1, X_2, \ldots, X_n$ for which the Beta$(\alpha+\sum X,$
$\beta+(n-\sum X))$

density is approximately normal with mean $\bar{X}_n$ and variance $\frac{\bar{X}_n(1-\bar{X}_n)}{n}$

?!? Note that α, β wash out of the above statement — i.e. every Beta prior will tend to have approximately normal posteriors ... and there is nothing in this development that prescribes a different approximation for different (α, β) pairs --- ☹

There are multivariate versions of this — for $\theta \in \mathbb{R}^k$ the posterior density of

$$\left(-H_n(\hat{\theta}_n)\right)^{1/2}\left(\theta - \hat{\theta}_n\right)$$

a matrix square root of $-H_n$ (Hessian of $-$ loglikelihood)

MLE

tends to $N_k\left(0, \underbrace{I_k}_{k\times k \text{ identity}}\right)$

i.e. the posterior density of θ is approximately k-dimensional normal with

mean vector $\hat{\theta}_n$ and covariance matrix $\left(-H_n(\hat{\theta}_n)\right)^{-1}$

inverse of observed FI

That gives a way to do approximate inference — but this boils down to inference based on large sample normal dsn of $\hat{\theta}_n$ — ☹

A slightly more Bayesian-looking version of this is to replace

$\hat{\theta}_n$ by $\tilde{\theta}_n =$ maximizer of $g(\theta|X)$
$(g(\theta) L_n(\theta))$

$H_n(\cdot)$ by $\tilde{H}_n(\cdot) =$ matrix of 2nd partials of log posterior density

There are more refined approximations available —

A bit of insight as to where the normal limit for posteriors comes from (no tight argument, just some rough motivation)

Lemma (not really enough to what I'm claiming -- but perhaps enough to motivate it)

Under regularity conditions in an iid model where $\theta \in \mathbb{R}^1$ and $\hat{\theta}_n$ is a consistent MLE of θ, for any Δ

$$Q_n(\Delta) = \ell_n(\hat{\theta}_n) - \ell_n\left(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}\right) \xrightarrow{P_\theta} \frac{1}{2}\Delta^2 I_1(\theta)$$

Picture

while $\ell_n(\hat{\theta}_n) - \ell_n(\hat{\theta} + \Delta)$ can expected to go to ∞, $Q_n(\Delta)$ tells me about local shape of $\ell_n(\theta)$ near $\hat{\theta}_n$