Stat 543 4-8-05

Example Stat 101 "Fact"

For $\hat{p}_n$ = a sample fraction of "type A" items from a large dicotomous population

$$Z = \frac{\hat{p}_n - p}{\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}}$$ is "approximately std normal"

So approximate confidence limits for p are

$$\hat{p}_n \pm z\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}$$

How / Why?

Define

$$Y_i = \begin{cases} 1 & \text{if the } i\text{th sampled object is of type A} \\ 0 & \text{otherwise} \end{cases}$$

$Y_i \sim$ Bernoulli (p) population fraction of type A and "if the population is large in comparison to sample then these are 'approximately independent' "

$$X_n = \sum_{i=1}^{n} Y_i \qquad \left(\text{This is } Bi(n,p) \right)$$

$$\hat{p}_n = \frac{X_n}{n} = \overline{Y}_n$$

CLT says for $p \in (0,1)$

$$\sqrt{n}\left(\overline{Y}_n - p\right) = \sqrt{n}\left(\hat{p}_n - p\right) \xrightarrow{\mathcal{L}_p} N(0, \underbrace{\operatorname{Var} Y_1}_{p(1-p)})$$

Since $\hat{p}_n = \bar{Y}_n \xrightarrow{P_p} EY_1 = p$ (weak law of large numbers)

I'll be able to "replace $p(1-p)$ with $\hat{p}_n(1-\hat{p}_n)$" . . .

$$\begin{pmatrix} \sqrt{n}(\hat{p}_n - p) \\ \hat{p}_n \end{pmatrix} \xrightarrow{\mathcal{L}_p} \begin{pmatrix} N(0, p(1-p)) \\ p \end{pmatrix}$$

$g(u,v) = \dfrac{u}{\sqrt{v(1-v)}}$ is cont$\underline{^s}$ except where

$v = 0$ or $v = 1$ so as long as $p \in (0,1)$

$$g\left(\sqrt{n}(\hat{p}_n - p), \hat{p}_n\right) \xrightarrow{\mathcal{L}_p} g\left(N(0, p(1-p)), p\right)$$

That is

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \xrightarrow{\mathcal{L}_p} N(0,1)$$

The most important large sample Thy for inference at this level concerns ML and the behavior of LRT statistics — regarding the 1st

"Often" MLE's are

Thm 5.2.2 → 1) consistent

Thm 5.3.5 → 2) approximately normal with variance of the approximate dsn $I^{-1}(\theta)$

for exponential families

<u>Def</u> A sequence of estimators $\{\delta_n\}$ of θ is consistent (weakly consistent / consistent in probability) at $\theta_0 \in \Theta$ provided

$$\delta_n \xrightarrow{P_{\theta_0}} \theta_0$$

<u>Aside</u>:
Recall an earlier comment about MoM estimation ... in an iid model

Provided $E|X_1|^r < \infty$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^s = \hat{\mu}_{sn} \xrightarrow{P_\theta} \mu_s(\theta) = E_\theta X_1^s \quad \forall\, s \le r$$

(by WLLN)

Then if $h(\cdot, \ldots, \cdot)$ is cont's at

$$(\mu_1(\theta), \mu_2(\theta), \ldots, \mu_r(\theta))$$

and
$$\delta_n = h(\hat{\mu}_{1n}, \hat{\mu}_{2n}, \ldots, \hat{\mu}_{rn})$$
is consistent for $\gamma(\theta) = h(\mu_1(\theta), \ldots, \mu_r(\theta))$

i.e. MoM estimators are typically consistent

Thy for "MLE's" — In class IID Theory (mostly for $\Theta \subset \mathbb{R}^1$) — handout is somewhat more general + B+D results are for multi-dimensional η exponential families (iid)

Suppose $X_1, X_2, \ldots$ are iid with pdf or pmf $f(x|\theta)$

Let

$$\ell_n(\theta) = \sum_{i=1}^{n} \ln f(X_i|\theta)$$

be the nth loglikelihood — what can be said in general ?

<u>Thm 1</u> (handout) If $\Theta \subset \mathbb{R}^1$ under appropriate regularity conditions if $\epsilon > 0$

$$P_{\theta_0}\left[\begin{array}{l}\text{The equation } \ell_n'(\theta) = 0 \\ \text{has a root within } \epsilon \\ \text{of } \theta_0\end{array}\right] \longrightarrow 1$$

(The θ_0 probability that the score function has a 0 close θ_0 goes to 1)

Corollary 2 (on handout)