The Neyman-Scott Paradox

David A Freedman

Let (X_i, Y_i) be independent $N(\alpha_i, \sigma^2)$ for i = 1, ..., n. The MLE for α_i is $\hat{\alpha}_i = (X_i + Y_i)/2$. The MLE for σ^2 is $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n s_i^2$, where $s_i^2 = [(X_i - \hat{\alpha}_i)^2 + (Y_i - \hat{\alpha}_i)^2]/2 = (X_i - Y_i)^2/4$, because $X_i - \hat{\alpha}_i = (X_i - Y_i)/2$ and $Y_i - \hat{\alpha}_i = (Y_i - X_i)/2$. So $E(s_i^2) = \sigma^2/2$ and the MLE is inconsistent as $n \to \infty$. This is a "fixed-effects" model with two observations on each effect α_i . The effect is estimated by the mean of the relevant observations: with only two observations per parameter, $\hat{\alpha}_i$ remains quite variable as $n \to \infty$. The common variance σ^2 is estimated by the mean of the sample size as the divisor, rather than degrees of freedom. The number of observations relevant to estimating σ^2 grows without bound, but inconsistency follows from the bias in the MLE.

To verify the formulas for the MLE, set $v = 1/\sigma^2$: this makes the calculus a little easier. The log likelihood is

$$2n\log\frac{1}{\sqrt{2\pi}} + n\log v - v\sum_{i=1}^{n}\frac{(X_i - \alpha_i)^2 + (Y_i - \alpha_i)^2}{2}$$

It's "obvious" that $\hat{\alpha}_i$ is as claimed. At these values for α_i , the derivative with respect to v is

$$\frac{n}{v} - \sum_{i=1}^{n} s_i^2$$

so

$$\hat{v} = \frac{n}{\sum_{i=1}^{n} s_i^2}$$

and

$$\hat{\sigma}^2 = \frac{1}{\hat{v}} = \frac{1}{n} \sum_{i=1}^n s_i^2$$

as required.

Notes for Statistics 215 Department of Statistics UC Berkeley, CA 94720-3860 November, 2005