Nov 29, 2021

ML Estimation (Ch 4)

Suppose we have $E(Y_i|X_i) = X_i\beta$ and $Cov(Y_i|X_i) = \Sigma_i = \overbrace{\Sigma_i(\theta)}$

Σ_i is a function of θ

θ = vector of covariance parameters (e.g., σ_{ij})

Using ML to estimate β, we get $\hat{\beta}_{ML} = \left[\sum_{i=1}^{N}(X_i' \underbrace{\hat{\Sigma}_i^{-1}} X_i)\right]^{-1} \sum_{i=1}^{N} X_i' \underbrace{\hat{\Sigma}_i^{-1}} y_i$

Σ_i is unknown so we must estimate it (meaning we must estimate all the σ_{ij})

It can be shown that in very $\overbrace{\text{large samples}}^{\text{large \# of clusters}}$ $E(\hat{\beta}_{ML}) = \beta$ and $Cov(\hat{\beta}) = \left[\sum_{i=1}^{N} X_i' \Sigma_i^{-1} X_i\right]^{-1}$ and $\hat{\beta}_{ML} \sim MVN\left(\beta, Cov(\hat{\beta})\right)$

We could use this large sample approximation to conduct inference (e.g., for $H_0: \beta_1 = 0$, we might compute $Z = \frac{\hat{\beta}_1 - 0}{\sqrt{Cov(\hat{\beta}_1)}}$)

But this falls apart in several ways:

(1) We don't know Σ_i so must estimate it with $\hat{\Sigma}_i$.

(2) Since we don't know Σ_i, we can't compute $Cov(\hat{\beta})$ and thus must estimate it with $\hat{Cov}(\hat{\beta}) = \left[\sum_{i=1}^{N} X_i' \hat{\Sigma}_i^{-1} X_i\right]^{-1}$

(3) The results are based on large sample approximations and do not hold for finite (smaller) samples

Specifically, $\hat{\Sigma}_i$ underestimates Σ_i in finite samples (i.e., $\hat{\Sigma}_i$ is too small $\Rightarrow$ $\hat{Cov}(\hat{\beta})$ is also too small)

So test statistics (e.g. $t = \frac{\hat{\beta} - \beta_0}{SE(\hat{\beta})}$) are too large and CIs (e.g. $\hat{\beta} \pm t^*[SE(\hat{\beta})]$) are too narrow

and we are therefore more likely to falsely reject H_0 ("anti-conservative" inference)

Satterthwaite and KR adjustments

For a simple test of $H_0: \beta_1 = 0$:

Under H_0, we assume $T = \frac{\hat{\beta}_1 - \beta_0}{SE(\hat{\beta})} \sim t_{df_r}$

but $SE(\hat{\beta})$ is too small

Satterthwaite fix: inflate df of H_0 distribution

K-R fix: inflate df of H_0 distr'n and SE($\hat{\beta}$)