The above analysis is valid only if the model is correctly specified: The true density is truly described by a parametric family $f(X|\theta), \theta \in \Theta$ for one specific θ_0 which we are trying to determine from our sampling

Otherwise the model is mis-specified:

$\nexists$ a true θ_0 so that the $f(X|\theta_0)$ to be the true density.

However, it can be seen as a good approximation of the true model. Then the pseudo-true parameter $\theta_0 \in \Theta$ is selected so that the approximate $f(X|\theta_0)$ to diverge as little as possible from the true one: $f(X)$

Kullback-Leibler divergence

$$KLIC(f, f_\theta) \equiv \int f(x) \cdot \log \frac{f(x)}{f(X|\theta)} dx = \int f(x) \cdot \log f(x) dx - \underbrace{\int f(x) \log f(X|\theta) dx}_{l(\theta)}$$

- $\to KLIC(f,f) = 0$
- $\to KLIC(f,g) \geq 0$
- $\to \operatorname{argmin} KLIC(f,g) \equiv f$

pseudo-true parameter:

$$\boxed{\theta_0 \equiv \underset{\theta \in \Theta}{\operatorname{argmin}}\, KLIC(f, f_\theta) \equiv \underset{\theta \in \Theta}{\operatorname{argmax}}\, l(\theta)} \Rightarrow$$

for a mis-specified model the MLE-$\hat{\theta}$ is the sample analog of the pseudo-true parameter θ_0

So the whole analysis remains untouched except for the difference that for mis-specified model:

$$\boxed{\mathcal{F}_\theta \neq \mathcal{H}_\theta} \quad \text{thus} \quad \boxed{\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N\left(0, \frac{1}{\mathcal{H}_\theta} \cdot \mathcal{F}_\theta \cdot \frac{1}{\mathcal{H}_\theta}\right)}$$

Practical comments

take n-sample

- $\to$ to estimate the (pseudo) true θ_0: calculate the MLE-$\hat{\theta} \equiv \operatorname{argmax} \bar{l}_n(\theta)$
- $\to$ to estimate the variance of $\hat{\theta}$, assume that n is big enough for the normal distribution to be valid. Thus $var[\hat{\theta}] \equiv \mathcal{H}_\theta^{-1} \cdot \mathcal{F}_\theta \cdot \mathcal{H}_\theta^{-1}$ or $var[\hat{\theta}] \equiv \mathcal{F}_\theta^{-1}$ for the pseudo-true or true case.
- $\to$ the previous exact expression must be replaced by appropriate estimates of $\hat{\mathcal{H}}_\theta, \hat{\mathcal{F}}_\theta$.
- $\to$ After calculating the estimate $\hat{V}$, our experimental estimation of θ_0 is MLE-$\hat{\theta}$ with standard error $s(\hat{\theta}) = \sqrt{\hat{V}/n}$.

??? $\hat{V}$ from:

- $\to \mathcal{F}_\theta^{-1} = \mathcal{H}_\theta^{-1} \Rightarrow \oplus$ stricter conditions BUT $var[\hat{V}]$ small $\oplus$
- OR
- $\to \mathcal{H}_\theta^{-1} \mathcal{F}_\theta \mathcal{H}_\theta^{-1} \Rightarrow \ominus$ $var[\hat{V}]$ bigger BUT broader conditions $\oplus$