Suppose $\sup_{\theta} |g(W,\theta)| < +\infty \Rightarrow E[\sup_{\theta} |g(W,\theta)|]$ is a uniform upper bound.

Suppose $\theta_0 = \arg\max_{\theta \in \Theta} \log f(Y_i, X_i; \theta)$

WTS $\log f(Y_i, X_i; \theta) - \log f(Y_i, X_i; \theta_0) < 0$

$$\Rightarrow E_0\left[\log \frac{f(Y_i, X_i; \theta)}{f(Y_i, X_i; \theta_0)}\right] \leq \log E_0\left[\frac{f(Y_i, X_i; \theta)}{f(Y_i, X_i; \theta_0)}\right]$$

$$= \log \iint \frac{f(Y_i, X_i; \theta)}{f(Y_i, X_i; \theta_0)} f(Y_i, X_i; \theta_0)\,dx\,dy$$

$$= \log \underbrace{\iint f(Y_i, X_i; \theta)\,dx\,dy}_{=1}$$

$$= \log 1$$

$$= 0$$

$\Rightarrow$ Thus, θ_0 is a maximizer.

In the Logit model,

$$E_0[\log f(Y_i, X_i; \theta_0)] = E_{X_i}[E[\ln f(Y_i, X_i; \theta_0) | X_i]]$$

$$= E_{X_i}\left[\frac{\exp\{X_i'\theta_0\}}{1+\exp\{X_i'\theta_0\}} \log\left(\frac{\exp\{X_i'\theta_0\}}{1+\exp\{X_i'\theta_0\}}\right) + \frac{1}{1+\exp\{X_i'\theta_0\}} \log\left(\frac{1}{1+\exp\{X_i'\theta_0\}}\right)\right]$$

$$+ E_{X_i}[\log g(X_i)]$$

If $F(y) = \frac{e^y}{1-e^y}$, then $f(y) = \frac{dF}{dy} = \frac{(1+e^y)e^y - e^y(+e^y)}{(1+e^y)^2}$

$$= \frac{e^y - e^{2y} + e^{2y}}{(1+e^y)^2}$$

$$= \frac{e^y}{(1+e^y)^2} = F(y)[1-F(y)]$$

Let θ be scalar. Then

$$F(\theta X) = \frac{e^{\theta x}}{1+e^{\theta x}}$$

$$\Rightarrow E_0[\ln f(Y_i, X_i; \theta_0)] = F(\theta X_i)\log F(\theta X_i) + (1-F(\theta X_i))\log(1-F(\theta X_i)]$$

$$\Rightarrow \frac{d E_0[\ln f(Y_i, X_i; \theta_0)]}{d\theta} = -F(\theta X_i)(1-F(\theta X_i))X_i^2 < 0$$

when $0 < \underline{F} < F(\theta X_i) < \bar{F} < 1$

Thus, we have that in the Logit case, $\hat{\theta}_0$ is the unique minimizer.

In general, we cannot check uniqueness. Thus,

Assumption uniqueness: We assume that $\forall \varepsilon > 0$,

$$\sup_{|\theta-\theta_0| \geq \varepsilon} E_0[\ln f(Y_i, X_i; \theta)] < E_0[\ln f(Y_i, X_i; \theta_0)]$$

If Θ is compact and $f(Y_i, X_i; \theta)$ is continuous for a.a. (Y_i, X_i) then a sufficient condition is that $\exists A$ s.t. $f(Y_i, X_i; \theta) \neq f(Y_i, X_i; \theta_0)$ $\forall (Y_i, X_i)$ and $\int_A f(Y_i, X_i; \theta_0)\,dx\,dy > 0$.

Thm: If for some Θ, with $\theta_0 \in \Theta$, the normalized log likelihood converges uniformly to the expected log likelihood and θ_0 is is a unique maximizer, then $\hat{\theta}_n \xrightarrow{p} \theta_0$.

Let $M_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^{n} \log f(W_i,\theta)$

$M(\theta) = E_0[\log f(W_i,\theta_0)]$

By definition of $\hat{\theta}_n$, $M_n(\hat{\theta}_n) \geq M_n(\theta_0)$

We know that $M_n(\theta_0) \longrightarrow M(\theta_0)$. Then $\exists$ sequence of nonnegative $\{Z_n\}$ s.t. $|M_n(\theta_0) - M(\theta_0)| \leq Z_n \; \forall n$. and $Z_n \xrightarrow{P} 0$.

$\Rightarrow -Z_n \leq M(\theta_0) - M_n(\theta_0) \leq Z_n$

$\Rightarrow 0 \leq M(\theta_0) - M(\hat{\theta}_n) \leq M_n(\theta_0) - M(\hat{\theta}_n) + Z_n$

$\Rightarrow 0 \leq M(\theta_0) - M(\hat{\theta}_n) \leq \underbrace{2\sup_{\theta\in\Theta} |M_n(\theta) - M(\theta)|}_{\xrightarrow{P} 0} + \underbrace{Z_n}_{\xrightarrow{P} 0}$

by unif conv. in prob.

$\Rightarrow |M(\theta_0) - M(\hat{\theta}_n)| \xrightarrow{P} 0$

$\Rightarrow M(\hat{\theta}_n) \xrightarrow{P} M(\theta_0)$.

Thus, $\hat{\theta}_n \xrightarrow{P} \theta_0$ since we assumed that θ_0 was the unique maximizer. $\square$

What is the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta_0)$?

Assume that $\frac{\partial \ln f}{\partial \theta}$ and $\frac{\partial^2 \ln f}{\partial\theta\partial\theta'}$ exist and are continuous in θ.

By the FOCs for acquiring $\hat{\theta}_n$ and by the MVT,

$$0 = \frac{1}{n}\frac{\partial \log L_n(\hat{\theta}_n)}{\partial \theta} = \frac{1}{n}\frac{\partial \log L_n(\theta_0)}{\partial \theta} + \frac{1}{n}\frac{\partial^2 \log L_n(\tilde{\theta})}{\partial \theta \partial \theta'} \cdot (\hat{\theta}_n - \theta_0)$$

for some $\tilde{\theta} \in |\hat{\theta}_n, \theta_0|$

Since $\hat{\theta}_n \xrightarrow{P} \theta_0$, $\tilde{\theta} \xrightarrow{P} \theta_0$, assuming $\left[\frac{1}{n}\frac{\partial^2 \log L_n(\tilde{\theta})}{\partial \theta \partial \theta'}\right]$ is nonsingular,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \left[-\frac{1}{n}\frac{\partial^2 \log L_n(\tilde{\theta})}{\partial \theta \partial \theta'}\right]^{-1} \sqrt{n}\left[\frac{1}{n}\frac{\partial \log L_n(\theta_0)}{\partial \theta}\right]$$

$$\xrightarrow{d} N\left(0, V\left(\frac{\partial \log f(w_i, \theta)}{\partial \theta}\right)\right)$$