1. Fixed Design Regression

(a) We consider the distribution of $\underline{Y} = (Y_1, \dots, Y_n)$, with

$\underline{Y} = X\theta + \underline{\varepsilon}$, $\varepsilon \sim N(0, I_n)$. Let $X := \begin{bmatrix} -x_1^T- \\ \vdots \\ -x_n^T- \end{bmatrix}$

$$dP_\theta^n(\underline{Y}) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \langle Y - X\theta, Y - X\theta \rangle \right\}$$

$$dP^n_{\theta + h/\sqrt{n}}(\underline{Y}) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \langle Y - X\left(\theta + \frac{h}{\sqrt{n}}\right), Y - X(\theta + h/\sqrt{n}) \rangle \right\}, \ h \in \mathbb{R}^n$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \Big[\langle Y - X\theta, Y - X\theta \rangle - 2\langle Y - X\theta, X\frac{h}{\sqrt{n}} \rangle + \frac{1}{n} \langle Xh, Xh \rangle \right.$$

$$\log\left[\frac{dP^n_{\theta + \frac{h}{\sqrt{n}}}}{dP^n_\theta} \right] = \langle Y - X\theta, X\frac{h}{\sqrt{n}} \rangle + \frac{1}{2n} \langle Xh, Xh \rangle$$

$$= \frac{1}{\sqrt{n}} \langle \varepsilon, Xh \rangle + \frac{1}{2n} \langle h, X^TXh \rangle$$

$$\frac{1}{\sqrt{n}} \langle \varepsilon, X^T h \rangle = \sqrt{n} \cdot \frac{1}{n} \langle \varepsilon, Xh \rangle \xrightarrow{d} N(0, h^T \Sigma h) \quad \text{by CLT}$$

$$\frac{1}{2n} \langle h, X^TXh \rangle \xrightarrow{d} \frac{1}{2} h^T \Sigma h \quad \text{(given)}$$

$$\Rightarrow \log\left[\frac{dP^n_{\theta + \frac{h}{\sqrt{n}}}}{dP^n_\theta} \right] \xrightarrow{d} N\left(\frac{1}{2} h^T \Sigma h, \ h^T \Sigma h \right)$$

$$\Rightarrow \frac{dP^n_{\theta + \frac{h}{\sqrt{n}}}}{d^n P_\theta} \xrightarrow[\theta]{d} \exp(Z), \quad Z \sim N\left(\frac{1}{2} h^T \Sigma h, h^T \Sigma h \right)$$

Since $\mathbb{P}_\theta[\exp(Z) > 0] = 1$, with $\frac{1}{L_n} := \frac{dP^n_{\theta + \frac{h}{\sqrt{n}}}}{dP_\theta} \to \exp(Z)$, by Le Cam's third Lemma, $P_\theta \triangleleft P_{\theta+\frac{h}{\sqrt{n}}}$

Further, with $\mathbb{E}[\exp(Z)] = \exp\left[-\frac{1}{2}h^T\Sigma h + \frac{1}{2}h^T\Sigma h\right] = 1$,

with $L_n = \frac{dP^n_{\theta+\frac{h}{\sqrt{n}}}}{dP^n_\theta} \xrightarrow[P_\theta]{d} L$, $\mathbb{E}[L] = 1$ so $P_{\theta+\frac{h}{\sqrt{n}}} \triangleleft P_\theta$

Hence $P_{\theta+\frac{1}{\sqrt{n}}h} \triangleleft\triangleright P_\theta$.

(b) We are interested in $\sup_\theta \mathbb{E}_\theta\left[n\|\hat\theta-\theta\|^2\right]$ for

$\hat\theta = (X^TX)^{-1}X^TY$, $X = \begin{bmatrix} - x_1^T - \\ \vdots \\ - x_n^T - \end{bmatrix}$

Under P_θ, $\hat\theta = (X^TX)^{-1}X^T(X\theta + \underline{\varepsilon})$, $\varepsilon \in \mathbb{R}^n$

$= (X^TX)^{-1}(X^TX)\theta + (X^TX)^{-1}X^T\underline{\varepsilon}$

$= \theta + (X^TX)^{-1}X^T\underline{\varepsilon}$

$= \theta + X^+\underline{\varepsilon}$, $X^+ = (X^TX)^{-1}X^T$

$\|\hat\theta-\theta\|^2 = \langle X^+\varepsilon, X^+\underline{\varepsilon}\rangle = \mathrm{tr}\left(X^+\varepsilon(X^+\varepsilon)^T\right) = \mathrm{tr}\left(X^+(\varepsilon\varepsilon^T)X^{+T}\right)$

$\mathbb{E}\left[n\|\hat\theta-\theta\|^2\right] = n\left(\mathrm{tr}(X^+\mathbb{E}[\varepsilon\varepsilon^T]X^{+T})\right) = \mathrm{tr}(X^+X^{+T})$ using $\mathbb{E}[\varepsilon\varepsilon^T] = I_n$

Since $\sqrt{n}\,P_n\varepsilon_i \xrightarrow{d} N(0,1)$

$X^+X^{+T} = (X^TX)^{-1}X^T\left(X(X^TX)^{-T}\right) = (X^TX)^{-1\,T} = (X^TX)^{-1}$ by symmetry

So $\sup_\theta \mathbb{E}\left[n\|\hat\theta-\theta\|^2\right] = \mathrm{tr}\left[\left(\frac{X^TX}{n}\right)^{-1}\right] \to \mathrm{tr}(\Sigma^{-1})$

by continuity of $x \mapsto \mathrm{tr}\,x$.

(c) Consider $\theta = 0$. Since $P_0^n [\tilde{\theta} \neq 0] \to 0$, by $P_\theta^n \triangleleft\triangleright P^n_{\frac{h}{\sqrt{n}}}$
$\mathbb{P}^n_{\frac{h}{\sqrt{n}}} [\tilde{\theta} \neq 0] \to 0$. Hence $\mathbb{P}^n_{\frac{h}{\sqrt{n}}} [\tilde{\theta} = 0] \to 1$

$$\mathbb{E}_\theta [n \| \tilde{\theta} - \theta \|^2] \geq \mathbb{E}_{h/\sqrt{n}} [n \| \tilde{\theta} - \frac{h}{\sqrt{n}} \|^2]$$

$$\geq \mathbb{E}_{h/\sqrt{n}} [n \| \frac{-h}{\sqrt{n}} \|^2] \, \mathbb{P}(\tilde{\theta} = 0)$$

$$= \mathbb{E}_{h/\sqrt{n}} [\| h \|^2] \, \mathbb{P}^n(\tilde{\theta} = 0) = \| h \|^2 \, \mathbb{P}^n_{h/\sqrt{n}}(\tilde{\theta} = 0)$$

$$\Rightarrow \limsup_{n \to \infty} \sup_\theta \mathbb{E}_\theta [n \| \tilde{\theta} - \theta \|^2] \geq \lim_{n \to \infty} \| h \|^2 \, \mathbb{P}^n_{h/\sqrt{n}} (\tilde{\theta} = 0) = \| h \|^2$$

Take $\| h \| \to \infty$, we get the desired result.

2. Testing Lower Bounds in QMD families via Hellinger distance

(a) In HWH, we showed

$$H^2\left(P^{\otimes n}, Q^{\otimes n}\right) = 1 - \left(1 - H^2(P, Q)\right)^n$$

Define
$$g_\theta := \nabla_\theta \sqrt{p_\theta} = \frac{1}{2} \dot{\ell}_\theta \cdot \sqrt{p_\theta}$$

$$R_\theta(h) := \sqrt{p_{\theta+h}} - \sqrt{p_\theta} + h^T g_\theta$$

Consider
$$H^2\left(P^n_{\theta_0}, P^n_{\theta_0 + \frac{1}{\sqrt{n}}h}\right) = \frac{1}{2} \int \left(\sqrt{p_{\theta_0}} - \sqrt{p_{\theta_0 + \frac{1}{\sqrt{n}}h}}\right)^2 d\mu$$

$$= \frac{1}{2} \int \left(-R_{\theta_0}(h/\sqrt{n}) + \frac{1}{\sqrt{n}} h^T g_\theta\right)^2 d\mu$$

QMD implies $\int R^2_\theta\left(\frac{1}{\sqrt{n}}h\right) d\mu = o\left(\frac{\|h\|^2}{n}\right)$,

also
$$\int \left(\frac{1}{\sqrt{n}} h^T g_{\theta_0}\right)^2 d\mu = \frac{1}{4n} \int \left(h^T \dot{\ell}_\theta \sqrt{p_{\theta_0}}\right)^T \left(h^T \dot{\ell}_\theta \sqrt{p_{\theta_0}}\right) d\mu$$

$$= \frac{1}{4n} \int \left(h^T \dot{\ell}_{\theta_0} \sqrt{p_{\theta_0}}\right)\left(h^T \dot{\ell}_{\theta_0} \sqrt{p_{\theta_0}}\right)^T d\mu$$

$$= \frac{1}{4n} \int h^T \left(p_{\theta_0} \dot{\ell}_{\theta_0} \dot{\ell}^T_{\theta_0}\right) h \, d\mu$$

$$= \frac{1}{4n} h^T \int p_{\theta_0} \dot{\ell}_{\theta_0} \dot{\ell}^T_{\theta_0} \, d\mu \, h$$

$$= \frac{1}{4n} h^T I_\theta h$$

Also, $\int \left(R_0(h/\sqrt{n})\right)\left(\frac{1}{\sqrt{n}} h^T g_{\theta_0}\right) d\mu$

$$\leq \| R_0(h/\sqrt{n}) \|_{L^2} \cdot \| \frac{1}{\sqrt{n}} h^T g_{\theta_0} \|_{L_2}$$

$$\int \left(R_0(h/\sqrt{n})\right)^2 d\mu = o\left(\frac{\|h\|^2}{n}\right)$$

$$\| \frac{1}{n} h^T g_{\theta_0} \|_{L_2}^2 = \int \left(\frac{1}{n} h^T g_{\theta_0}\right)^2 d\mu$$

$$= \int \frac{1}{n} h^T \frac{1}{2} \nabla \sqrt{p_{\theta_0}}\, d\mu$$

$$= \frac{1}{n} h^T \frac{1}{2} \sqrt{p_{\theta_0}} = O(1)$$

Hence $\int R(h/\sqrt{n}) \left(\frac{1}{\sqrt{n}} h^T g_{\theta_0}\right) d\mu = O(1) \cdot o(1) = o(1)$

Thus $H^2\left(P_{\theta_0}, P_{\theta_0 + \frac{1}{\sqrt{n}} h}\right) = \frac{1}{8n} h^T I_{\theta_0} h + o\left(\frac{\|h\|^2}{n}\right)$

$$H^2\left(P^n_{\theta_0}, P^n_{\theta_0 + \frac{1}{\sqrt{n}} h}\right) = 1 - \left(1 - H^2\left(P_{\theta_0}, P_{\theta_0 + \frac{1}{\sqrt{n}} h}\right)\right)^n$$

$$= 1 - \left[1 - \frac{1}{n}\left(\frac{h^T I_{\theta_0} h}{8} + o(\|h\|^2)\right)\right]^n$$

$$\xrightarrow[n \to \infty]{} 1 - \exp\left(-\frac{1}{8} h^T I_{\theta} h\right)$$

(b) By the previous HW,

$$\liminf_{n\to\infty} \left(P_{\theta_0}[T_n \neq 0] + P_{\theta_0 + \frac{1}{\sqrt{n}}h}[T_n = 1] \right) \geq 1 - \| P_{\theta_0}^n - P_{\theta_0 + h/\sqrt{n}}^n \|_{TV}$$

Enough to show $\| P_{\theta_0}^n - P_{\theta_0 + \frac{h}{\sqrt{n}}}^n \| < 1$.

Recalling, with $H_n := H^2(P_{\theta_0}, P_{\theta_0 + \frac{1}{\sqrt{n}}})$ that

$$\| P_{\theta_0}^n - P_{\theta_0 + \frac{h}{\sqrt{n}}}^n \|_{TV} < \sqrt{2H_n(1 - H_n/2)},$$

$$\begin{aligned}
\sqrt{2H_n(1-H_n/2)} < 1 &\iff 2H_n(1 - H_n/2) < 1 \\
&\iff H_n(2 - H_n) < 1 \\
&\iff H_n^2 - 2H_n + 1 > 0 \\
&\iff (H_n - 1)^2 > 0 \\
&\iff |H_n - 1| > 0 \\
&\iff H_n \neq 1
\end{aligned}$$

$H_n = 1 - \left[1 - \frac{1}{n}\left(\frac{1}{8} h^T I_\theta h + o(\|h\|^2)\right)\right]^n$ decreasing in n, and attaining

Since $H_n \to 1 - \exp\left(-\frac{1}{8} h^T I_0 h\right)$ hence $H_n \neq 1 \ \forall n \in \mathbb{N}$.

Hence

$$\liminf_{n\to\infty} \left(P_{\theta_0}[T_n = 0] + P_{\theta_0 + \frac{1}{\sqrt{n}}h}[T_n = 1] \right) > 0.$$

3. QMD implies LAN

(a) $$\log \frac{dP^n_{\theta+h_n}}{dP^n_\theta} = \sum_{i=1}^n \log\left(\frac{p_{\theta+h_n}}{p_\theta}(X)\right) = \sum_{i=1}^n \log\left((1+B_i^{(n)})^2\right) = \sum_{i=1}^n 2\log(1+B_i^{(n)})$$

(b) Using $B_i = \frac{1}{\sqrt{p_\theta}}\left(\sqrt{p_{\theta+h_n}} - \sqrt{p_\theta}\right) = \frac{1}{\sqrt{p_\theta}}\left(\frac{1}{2}h^T \dot{\ell}_\theta \sqrt{p_\theta} + R(h_n)\right)$

By QMD, $\int (R(h_n))^2 \, d\mu \to 0$, hence

$$E_{p_\theta}\left[(B_i^{(n)})^2\right] = \int \left(\sqrt{\frac{p_{\theta+h_n}}{p_\theta}} - 1\right)^2 p_\theta \, d\mu$$

$$= \int \left(\left(\sqrt{p_{\theta+h_n}} - \sqrt{p_\theta}\right)\frac{1}{\sqrt{p_\theta}}\right)^2 p_\theta \, d\mu \qquad = 2H^2(P_{\theta+h_n}, P_\theta)$$

$$= \int \left(R_\theta(h_n) + h^T \tfrac{1}{2}\dot{\ell}_\theta \sqrt{p_\theta}\right)^2 \frac{1}{p_\theta} \cdot p_\theta \, d\mu$$

$$= E_\theta\left[\left(R_\theta(h_n) + h^T\tfrac{1}{2}\dot{\ell}_\theta\sqrt{p_\theta}\right)^2 \frac{1}{p_\theta}\right]$$

$$= E_\theta\left[\frac{R_\theta^2(h_n)}{p_\theta}\right] + \frac{1}{4}E_\theta\left[h_n^T \dot{\ell}_\theta \dot{\ell}^T h\right] + E_\theta\left[\frac{R_\theta(h_n)\cdot h^T \frac{1}{2}\dot{\ell}_\theta}{\sqrt{p_\theta}}\right]$$

$\int R_\theta^2(h_n)\, d\mu \to 0$ by QMD $\Rightarrow E_\theta\left[\frac{R_\theta^2(h_n)}{p_\theta}\right] \to 0$

$$\int \frac{h^T \dot{\ell} R_\theta(h_n)}{\sqrt{p_\theta}} p_\theta \, d\mu \le \left(\int h^T \dot{\ell}_\theta \dot{\ell}^T h \, d\mu \cdot \underbrace{\int R_\theta^2(h_n)\, d\mu}_{=o(\|h_n\|^2)}\right)^{1/2} \longrightarrow 0$$

Hence $\mathbb{E}_\theta\left[(B_i^{(n)})^2\right] = \frac{1}{4} h_n^T I_\theta h_n + o\left(\frac{1}{n}\right)$

$$\mathbb{E}_{p_\theta}\left[B_i^{(n)}\right] = \int p_\theta\left(\sqrt{\frac{p_{\theta+h_n}}{p_\theta}} - 1\right) d\mu$$

$$= \int \left(\sqrt{p_\theta\, p_{\theta+h_n}} - p_\theta\right) d\mu$$

$$= \int \sqrt{p_\theta\, p_{\theta+h_n}} - \frac{1}{2} p_\theta - \frac{1}{2} p_{\theta+h_n}\, d\mu$$

$$= \frac{-1}{2} \int \left(\sqrt{p_{\theta+h_n}} - \sqrt{p_\theta}\right)^2 d\mu$$

$$= -H^2\left(P_{\theta+h_n}, P_\theta\right)$$

$$= -\frac{1}{8} h_n^T I_\theta h_n + o(\|h_n\|^2) \quad \text{by} \quad 2(a)$$

(c) $\mathbb{P}\left[\max_{i\in[n]} |B_i^{(n)}| \geq \varepsilon\right]$

$$\leq n\, \mathbb{P}\left[|B_i^{(n)}| \geq \varepsilon\right]$$

$$\leq n\, \mathbb{P}\left[|B_i^{(n)} - \tfrac{1}{2} h_n^T \dot{\ell}_\theta(X_i)| \geq \varepsilon/2\right] + \mathbb{P}\left[|\tfrac{1}{2} h_n^T \dot{\ell}_\theta(X_i)| \geq \tfrac{\varepsilon}{2}\right]$$

Let $C_i^{(n)} := B_i^{(n)} - \frac{1}{2} h_n^T \dot{\ell}_\theta(X_i)$

$$\leq \frac{4n}{\varepsilon^2} \mathbb{E}\left[(C_i^{(n)})^2\right] + \mathbb{P}\left[|h^T \dot{\ell}_\theta(X_i)| \geq \varepsilon\sqrt{n}\right]$$

$$B_i^{(n)} - \frac{1}{2} h_n^T \dot{\ell}_\theta(X_i) = \frac{1}{\sqrt{p_\theta}}\left(\sqrt{p_{\theta+h_n}(X_i)} - \sqrt{p_\theta(X_i)} - \frac{1}{2}\langle h, \dot{\ell}_\theta(X_i)\rangle \sqrt{p_\theta(X_i)}\right)$$

$$\Rightarrow \mathbb{E}_\theta\left[(C_i^{(n)})^2\right] = \int R^2(h_n)\, d\mu = o(\|h_n\|^2) = o(\tfrac{1}{n}) \text{ by QMD}$$

$$\mathbb{P}\left[|h^T \dot{\ell}_\theta(X_i)| \geq \varepsilon\sqrt{n}\right] \leq \frac{1}{n\varepsilon^2} \mathbb{E}_\theta\left[(h^T \dot{\ell}_\theta(X_i))^2\right] = \frac{1}{n\varepsilon^2} \mathbb{E}_\theta\left[h^T \dot{\ell}_\theta \dot{\ell}_\theta^T h\right]$$

by $P_\theta \dot{\ell}_\theta \dot{\ell}_\theta^T < \infty$ (given), $\frac{1}{n\varepsilon^2} h^T (P_\theta \dot{\ell}_\theta \dot{\ell}_\theta^T) h = o(1) \cdot O(1) = o(1)$

So $\mathbb{P}\left[\max_{i \in [n]} |B_i^{(n)}| \geq \varepsilon\right] \to 0$

(d) Recalling from above

$$B_i^{(n)} = C_i^{(n)} + \frac{1}{2\sqrt{n}} h^T \dot{\ell}_\theta(X_i), \qquad (1)$$

$$\mathbb{E}_\theta\left[(C_i^{(n)})^2\right] = \int (R(h_n))^2\, d\mu = o(\tfrac{1}{n})$$

By (a), and using the Taylor Expansion of $x \mapsto \log(1+x)$,

$$\log\left(\frac{dP^n_{\theta+h_n}}{dP^n_\theta}\right) = \sum_{i=1}^n 2\log(1+B_i^{(n)})$$

$$= \sum_{i=1}^n 2B_i^{(n)} - \sum_{i=1}^n (B_i^{(n)})^2 + o\left(2\sum_{i=1}^n (B_i^{(n)})^2\right)$$

$$= \sum_{i=1}^n \frac{1}{\sqrt{n}} h^T \dot{\ell}_\theta(X_i) + 2\sum_{i=1}^n C_i^{(n)} - \sum_{i=1}^n (B_i^{(n)})^2 + o\left(2\sum_{i=1}^n (B_i^{(n)})^2\right) \quad (1)$$

$$= \sum_{i=1}^n \frac{1}{\sqrt{n}} h^T \dot{\ell}_\theta(X_i) + 2\sum_{i=1}^n C_i^{(n)} - \sum_{i=1}^n (B_i^{(n)})^2 + o\left(2\sum_{i=1}^n (B_i^{(n)})^2\right)$$

$$= \sum_{i=1}^n \frac{1}{\sqrt{n}} h^T \dot{\ell}_\theta(X_i) + 2\sum_{i=1}^n C_i^{(n)} - \frac{1}{4n}\sum_{i=1}^n h^T \dot{\ell}_\theta \dot{\ell}_\theta^T h$$

$$- \frac{1}{2}\sum_{i=1}^n \frac{1}{2\sqrt{n}} h^T \dot{\ell}_\theta \cdot C_i - \sum_{i=1}^n C_i^2$$

$$+ o\left(2\sum_{i=1}^n (B_i^{(n)})^2\right)$$

- $\frac{1}{n}\sum_{i=1}^n (C_i^{(n)})^2 \xrightarrow{P} \mathbb{E}\left[(C_i^{(n)})^2\right] = o\left(\frac{1}{n}\right)$ by QMD $\rightarrow n P_n (C_i^{(n)})^2 = o(1)$
- $o\left(2\sum_{i=1}^n (B_i^{(n)})^2\right) = 2\,o\left(\sum_{i=1}^n -\frac{1}{8} h_n^T I_\theta h_n + o\left(\frac{\|h\|}{n}\right)\right)$

$$= 2n\,o\left(\frac{1}{n}\left(\frac{1}{8} h^T I_\theta h + o(\|h\|)\right)\right) = o\left(\frac{1}{8} h^T I h + o(\|h\|)\right)$$

- $P_n \dot{\ell}_\theta \rightarrow P_\theta \dot{\ell}_\theta = 0$, with $C_i^{(n)} = o_p\left(\frac{1}{\sqrt{n}}\right)$,

$$-\frac{1}{2}\frac{1}{\sqrt{n}}\sum_{i=1}^n h^T \dot{\ell}_\theta\, o_p\left(\frac{1}{\sqrt{n}}\right) \xrightarrow{P} 0$$

Hence

$$\log\left(\frac{d^n P_{\theta+h_n}}{d^n P_\theta}\right) = \sum_{i=1}^n \frac{1}{\sqrt{n}} h^T \dot{\ell}_\theta(X_i) + 2\left(\sum_{i=1}^n B_i + \frac{1}{2\sqrt{n}} h^T \dot{\ell}_\theta(X_i)\right)$$

$$- \frac{1}{4n}\sum_{i=1}^n h^T \dot{\ell}_\theta \dot{\ell}_\theta^T h + o_{P_\theta}(1)$$

$$= \sum_{i=1}^n \frac{1}{\sqrt{n}} h^T \dot{\ell}_\theta(X_i) + 2n\left(\mathbb{P}_n\left(B_i + \frac{1}{2\sqrt{n}} h^T \dot{\ell}_\theta(X_i)\right)\right) + o_{P_\theta}(1)$$

$$- \frac{1}{4} h^T \mathbb{P}_n \dot{\ell}_\theta \dot{\ell}_\theta^T h + o_{P_\theta}(1)$$

$$(2) \quad = \sum_{i=1}^n \frac{1}{\sqrt{n}} h^T \dot{\ell}_\theta(X_i) - \frac{1}{4} h^T I_\theta h - \frac{1}{4} h^T I_\theta h + o(1) + o_{P_\theta}(1)$$

$$= \left\langle h, \frac{1}{\sqrt{n}} \sum_{i=1}^n \dot{\ell}_\theta(X_i) \right\rangle - \frac{1}{2} h^T I_\theta h + o_{P_\theta}(1)$$

(2) Using $\mathbb{P}_n B_i \to E_\theta B_i = -\frac{1}{8} h_n^T I_\theta h_n + o\left(\frac{1}{n}\right)$

This is what we wanted to show.

4 AMP Tests ?

(a) Since $h \in \mathbb{R}$, we are in 1 dimension.
Under necessary assumptions the asymptotic normality of the MLE holds.
(Ed post says, we can assume "nice" family
We can construct a Wald Test: Under $\theta = 0$,

$$\sqrt{n}\, \hat{\theta}_n \xrightarrow{d} N(0, I_0^{-1}). \Rightarrow (n I_0)\hat{\theta}_n^2 \xrightarrow{d} \chi_1^2$$

Consider $R_\alpha := \left\{ \hat{\theta}_n : |\hat{\theta}| > \sqrt{(n I_\theta)^{-1} F^{-1}_{\chi_1^2}(1-\alpha)} \right\}$

$$(1\text{-dim}) = \left\{ \hat{\theta}_n : \left(\sqrt{n}\, I_\theta^{-1/2} \hat{\theta}_n\right)^T \left(\sqrt{n}\, I_0^{-1/2} \hat{\theta}_n\right) > F^{-1}_{\chi_1^2}(1-\alpha) \right\}$$

Then $\mathbb{P}_0 \left[n I_0 \hat{\theta}_n^2 > F^{-1}_{\chi_1^2}(1-\alpha) \right] \to 1 - F_{\chi_1^2}\left[F^{-1}_{\chi_1^2}(1-\alpha) \right] = \alpha$

with $F^{-1}_{\chi_1^2}$ the $1-\alpha$ quantile of χ_1^2.

(b) As showed in HW4 Q4e,
using $\{P_\theta\}_{\theta \in \Theta}$ a "nice" family and hence $P_0^{\otimes n} \triangleleft \triangleright P^{\otimes n}_{\frac{1}{\sqrt{n}}h}$

$$\log L_n = \sqrt{n}\, \mathbb{P}_n\, h^T \dot{\ell}_0 - \frac{1}{2} h^T I_0 h + o_{P_\theta}(1)$$

$$\xrightarrow[\theta=0]{d} N\left(-\frac{1}{2} h^T I_0 h,\ h^T I_0 h\right) \quad \text{(CLT)}$$

$$\sqrt{n}\,\hat{\theta}_n = \left(P_0 \nabla^2 \ell_0 + o_p(1) \right)^{-1} \left(-\sqrt{n}\, P_n \nabla \ell_0 \right) \xrightarrow{d} N(0, I_0^{-1})$$

under consistency assumption of MLE

Since $g : x \mapsto \left(I_0^{-1} x,\ h^T x - \frac{1}{2} h^T I_0 h + o_p(1) \right) \in \mathcal{L}(\mathbb{R}^n)$, by Slutsky / CMT with g applied at $x = \sqrt{n}\, P_n \nabla \ell_0$, we obtain and with $g' = (I_0^{-1}, h)$

$$\begin{bmatrix} \sqrt{n}\,\hat{\theta}_n \\ \log L_n \end{bmatrix} \xrightarrow[P_0]{d} N\left(\begin{bmatrix} 0 \\ -\frac{1}{2} h^T I_0 h \end{bmatrix}, \begin{bmatrix} I_0^{-1} & \tau \\ \tau & h^T I_0 h \end{bmatrix} \right)$$

$$\tau = \operatorname{cov}\left(\sqrt{n}\, I_0^{-1} P_n \nabla \ell_0,\ \sqrt{n}\, h^T P_n \nabla \ell_0 - \frac{1}{2} h^T I_0 h \right) = h$$

Using Le Cam's 3rd lemma,

$$\sqrt{n}\,\hat{\theta}_n \xrightarrow[P^{\otimes n}_{\frac{h}{\sqrt{n}}}]{d} N(h, I_0^{-1}).$$

c) Recalling def: asymptotic power $:= \liminf_{\theta_1 \in \Theta_1} \mathbb{P}_{\theta_1}[T_n = 1]$

Because $\{P_\theta\}_{\theta \in \Theta}$ is LAN, it is AMP iff
we can show that $P_0[T_n = 1] \to \alpha$, which we checked in (a),
and that

$$\forall h \in \mathbb{R}^d \quad \lim_{n \to \infty} P_{h/\sqrt{n}}[T_n = 1] = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{h^T I_0 h}\right)$$

We have $P_{\frac{h}{\sqrt{n}}}[T_n = 1] = P_{\frac{h}{\sqrt{n}}}\left[\left|I_{\hat{\theta}_n}^{1/2} \sqrt{n}\, \hat{\theta}_n\right| > \sqrt{F^{-1}_{\chi^2_1}(1-\alpha)}\right]$

$$\to P_{Z \sim N(0,1)}\left[Z + I_0^{1/2} h > \sqrt{F^{-1}_{\chi^2_1}(1-\alpha)}\right]$$

Remark: $\sqrt{F^{-1}_{\chi^2_1}(1-\alpha)} = z_{1-\alpha/2}$. So

$$P_{h/\sqrt{n}}[T_n = 1] \longrightarrow 1 - \Phi\left(\Phi^{-1}(1-\tfrac{\alpha}{2}) - I_0^{1/2} h\right) + \Phi\left(-\Phi^{-1}(1-\tfrac{\alpha}{2}) - I_0^{1/2} h\right)$$

However, by testing for real values, this quantity is different from $1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sqrt{h^T I_0 h}\right)$, so the wald test is not AMP.

5. A Generalization of Anderson's Lemma

(a) Let $x, y \in \mathbb{R}^n$, $\lambda \in [0,1]$.

- if $x \notin C$ or $y \notin C$

  Then $\lambda \log(g(x)) + (1-\lambda)\log(g(y)) = -\infty$

  So $\lambda \log(g(x)) + (1-\lambda)\log(g(y)) \leq \log(g(\lambda x + (1-\lambda)y))$

- if $x \in C$ and $y \in C$.

  Then C is convex so $\lambda x + (1-\lambda)y \in C$. So

  $\lambda \log(g(x)) + (1-\lambda)\log(g(y)) = \lambda + 1 - \lambda = 0 = \log(g(\lambda x + (1-\lambda)y))$

  Hence g is log-concave

(b) WTS: f, g log concave $\Rightarrow$ $\log\left(\int f(x-y)g(y)\,dy\right)$ concave

i.e. $\lambda \log((f*g)(x_1)) + (1-\lambda)\log((f*g)(x_2))$
$\leq \log((f*g)(\lambda x_1 + (1-\lambda)x_2))$

$\Leftrightarrow \log\left((f*g)(x_1)^{\lambda} \cdot (f*g)(x_2)^{1-\lambda}\right) \leq \log((f*g)(\lambda x_1 + (1-\lambda)x_2))$

$\Leftrightarrow (f*g)(x_1)^{\lambda} \cdot (f*g)(x_2)^{1-\lambda} \leq (f*g)(\lambda x_1 + (1-\lambda)x_2)$

$$\Leftrightarrow \left(\int f(x_1-y)\,g(y)\,dy\right)^{\lambda} \left(\int f(x_2-y)\,g(y)\,dy\right)^{1-\lambda}$$

$$\leq \int f\big((\lambda x_1+(1-\lambda)x_2)-y\big)\,g(y)\,dy$$

This will be implied by Prékopa-Leindler inequality, if we can show

$$f((\lambda x_1+(1-\lambda)x_2)-y)\,g(y) \geq f(x_1-y)g(y)^{\lambda} \cdot f(x_2-y)g(y)^{1-\lambda}$$

$$= f(x_1-y)^{\lambda} f(x_2-y)^{1-\lambda} g(y).$$

$$f(\lambda x_1+(1-\lambda)x_2) \leq f(x_1-y)^{\lambda} f(x_2-y)^{1-\lambda}$$

and $g(y) \leq g(y)^{\lambda} g(y)^{1-\lambda}$

because f, g log-concave (take log on both sides and result is immediate.)

(c) f log-concave probability density

Let C be a convex set. Let $y = x+v$

$$h(v) = \mathbb{P}[X+v \in C] = \int_{\mathbb{R}^n} f(x)\,\mathbb{1}_{\{x+v\in C\}} = \int_{\mathbb{R}^n} f(y-v)\,g(y)\,dy \quad (*)$$

This is different from the earlier convolution, but the same argument holds with signs flipped. The functions in the Prékopa-Leindler inequality will be $h'(y) = f(y-\lambda x_1+(1-\lambda)x_2)\,g(y)$

$f'(y) = f(y-x_1)\,g(y)$, $g' = f(y-x_2)\,g(y)$

By our previous argument with f is by assumption log-concave,
So by (b), the convolution $(*)$ is log-concave.

(d) WTS: f log-concave & symmetric $\forall \underline{x} \in \mathbb{R}^d$.

If $L : \mathbb{R}^k \to \mathbb{R}_+$ is quasiconvex and symmetric,
then for any matrix $A \in \mathbb{R}^{k \times d}$,

$$\inf_{v} \mathbb{E}\left[L(AX-v)\right] = \mathbb{E}\left[L(AX)\right]$$

We know for $v=0$, equality achieved. So WTS

$$\mathbb{E}\left[L(AX-v)\right] \geq \mathbb{E}\left[L(AX)\right] \quad \forall \ Ax-v \in \operatorname{dom} L.$$

Note: $\mathbb{E}\left[L(AX-v)\right] = \int_0^\infty y_v f_Y(y_v)\, dy_v, \quad y_v = L(Ax-v)$

$$= \int_0^\infty \left(\int_0^y ds\right) f_{Y_v}(y_v)\, dy_v$$

$$= \int_0^\infty \int_0^y f_{Y_v}(y_v)\, ds\, dy_v$$

$$= \int_0^\infty \int_s^\infty f_{Y_v}(y_v)\, dy_v\, ds$$

$$= \int_0^\infty \left(\int_s^\infty f_{Y_v}(y_v)\, dy_v \right) ds$$

$$= \int_0^\infty \mathbb{P}[Y_v > s]\, ds$$

$$= \int_0^\infty 1 - \mathbb{P}[Y_v \leq s]\, ds$$

$$= \int_0^\infty 1 - \mathbb{P}[Ax - v \in C_s]\, ds$$

So if we show $\mathbb{P}[Ax - v \in C_s] \overset{\forall v}{\leq} \mathbb{P}[Ax \in C_s]$, we are done.

Define $h(v) = \mathbb{P}[AX - v \in C_s]$, $g(x) := \mathbb{1}_{\{x \in C_s\}}$

Let us show h is log-concave.

$h(v) = \int f(x)\, g(Ax - v)\, dx$, f is log concave by

and $g : x \mapsto \mathbb{1}_{\{x \in C_s\}}$ is log. concave by (b).

Let us define $\phi(x, v) := f(x)\, g(Ax - v)$

Take $(x_1, v_1) \in \mathbb{R}^d \times \mathbb{R}^k$, $(x_2, v_2) \in \mathbb{R}^d \times \mathbb{R}^k$, $\lambda \in (0,1]$

Note that $(x,v) \mapsto g(Ax - v)$ is is log-concave by pre-composition with affine function.

$$\log(\phi(\lambda(x_1, v_1) + (1-\lambda)(x_2, v_2))$$

$$= \log(f(\lambda x_1 + (1-\lambda)x_2) + \log g(A(\lambda x_1 + (1-\lambda)x_2 - \lambda v_1 - (1-\lambda)v_2))$$

$$\geq \lambda \log(\lambda x_1) + (1-\lambda)\log(x_2) + \lambda \log(g(Ax_1 - v_1) + (1-\lambda)\log(g(Ax_2 - v_2))$$

$$= \lambda(\log(f(x_1)) + \log(g(Ax_1 - v_1))) + (1-\lambda)(\log(x_2) + \log(g(Ax_2 - v_2)))$$

$$= \lambda \log \phi(x_1, v_1) + (1-\lambda)\log \phi(x_2, v_2)$$

Hence ϕ is log-concave.
By Prékopa-Leinder inequality, so is h.

hence $h(0) \geq h(-v)^{1/2} h(v)^{1/2}$ i.e.

$$\mathbb{P}[Ax \in C_s] \geq \mathbb{P}[Ax - v \in C_s]^{1/2} \mathbb{P}[Ax + v \in C_s]^{1/2}$$

But f, L symmetric so $\mathbb{P}[Ax - v \in C_s]$
$= \mathbb{P}[-Ax + v \in C_s]$ L symmetric
$= \mathbb{P}[Ax + v \in C_s]$ f symmetric

& $\mathbb{P}[Ax \in C_S] \leq \mathbb{P}[Ax - v \in C_S] \quad \forall v \in \mathrm{dom}$

$\Rightarrow \mathbb{E}[L(AX - v)] \geq \mathbb{E}[L(AX)]$

This is what we wanted to show.