1. Reduction between Testing & Estimation

(a) Consider the sequence of tests

$$T_n = \begin{cases} 1 & \text{if } r_n d(\Theta_0, \hat{\theta}_n) > \sqrt{r_n} \\ 0 & \text{otherwise} \end{cases}$$

By definition and with Θ_0 closed,

$$d(\Theta_0, \hat{\theta}_n) = \inf_{\theta_0 \in \Theta_0} \| \theta_0 - \hat{\theta}_n \| = \min_{\theta_0 \in \Theta_0} \| \theta_0 - \hat{\theta}_n \|$$

· pointwise asymptotic level 0:

$$r_n (\theta - \hat{\theta}_n) = O_{P_\theta}(1) \rightarrow \forall \varepsilon > 0, \exists M^\theta(\varepsilon) \text{ s.t.}$$

$$\mathbb{P}_\theta \left[\| r_n (\theta - \hat{\theta}_n) \| > M^\theta(\varepsilon) \right] \leq \varepsilon$$

Hence if $\theta_0 \in \Theta_0$,

$$\mathbb{P}_{\theta_0} [T_n = 1] = \mathbb{P}_{\theta_0} \left[r_n d(\Theta_0, \hat{\theta}_n) > \sqrt{r_n} \right]$$

$$\leq \mathbb{P}_{\theta_0} \left[r_n \| \theta - \hat{\theta}_n \| > \sqrt{r_n} \right]$$

For $n(M^{\theta_0}(\varepsilon))$ large enough, $\forall\ n > n(M^{\theta_0}(\varepsilon))$, we have

$$\sqrt{r_n} \geq M^{\theta_0}(\varepsilon) \quad \text{and}$$

$$\mathbb{P}_{\theta_0}\left[T_n = 1\right] \leq \mathbb{P}_{\theta_0}\left[r_n \| \theta - \hat{\theta}_n \| > M^{\theta_0}(\varepsilon)\right] \leq \varepsilon$$

$$\Rightarrow \limsup_{n\to\infty} \mathbb{P}_{\theta_0}\left[T_n = 1\right] \leq \limsup_{n\to\infty} \mathbb{P}_{\theta_0}\left[r_n \| \theta - \hat{\theta}_n \| > M^{\theta_0}(\varepsilon)\right] \leq \varepsilon.$$

Since ε arbitrary, for $\varepsilon_n \downarrow 0$,

$$\limsup_{n\to\infty} \mathbb{P}\left[T_n = 1\right] \leq \limsup_{n\to\infty} \mathbb{P}_{\theta}\left[r_n \| \theta - \hat{\theta}_n \| > M^{\theta_0}(\varepsilon_n)\right] = 0.$$

This holds $\forall\, \theta_0 \in \Theta_0$, and Θ_0 is closed. taking $\sup_{\theta_0 \in \Theta_0}$,

$$\sup_{\theta_0 \in \Theta_0} \limsup_{n\to\infty} \mathbb{P}\left[T_n = 1\right] \leq 0$$, sup is attained for some $\theta_0 \in \Theta_0$, so

Hence $$\sup_{\theta_0 \in \Theta_0} \limsup_{n\to\infty} \mathbb{P}\left[T_n = 1\right] = 0.$$

<u>Pointwise asymptotic power 1</u>

Suppose $\theta = \theta_1 \in \Theta_1$, Θ_0 closed $\Rightarrow d(\Theta_0, \theta_1) > 0$.

Take $\eta := d(\Theta_0, \theta) > 0$

$$\mathbb{P}_{\theta_1}[T_n = 1] = \mathbb{P}_{\theta_1}\left[r_n d(\theta_0, \hat{\theta}_n) > \sqrt{r_n}\right]$$

$$\geq \mathbb{P}_{\theta_1}\left[r_n\left(d(\theta_0, \theta_1) - d(\theta_1, \hat{\theta}_n)\right) > \sqrt{r_n}\right]$$

$$\geq \mathbb{P}_{\theta_1}\left[r_n(\eta - d(\theta_1, \hat{\theta}_n)) > \sqrt{r_n}\right]$$

$$= \mathbb{P}_{\theta_1}\left[r_n\eta - r_n d(\theta_1, \hat{\theta}_n) > \sqrt{r_n}\right]$$

$$= \mathbb{P}_{\theta_1}\left[\sqrt{r_n}(\sqrt{r_n}\eta - 1) > r_n d(\theta_1, \hat{\theta}_n)\right]$$

$$\geq \mathbb{P}_{\theta_1}\left[\sqrt{r_n}(\sqrt{r_n}\eta - 1) > r_n \|\theta_1 - \hat{\theta}_n\|\right]$$

$$= 1 - \mathbb{P}_{\theta_1}\left[\sqrt{r_n}(\sqrt{r_n}\eta - 1) \leq r_n \|\theta_1 - \hat{\theta}_n\|\right]$$

Under $\theta = \theta_1$, using $r_n(\hat{\theta}_n - \theta_1) = O_p(1)$, $\forall \varepsilon > 0$ $\exists$ $M^{\theta_1}(\varepsilon)$ s.t.

$$\lim_{n\to\infty} \mathbb{P}_{\theta_1}\left[r_n \|\theta_1 - \hat{\theta}_n\| > M^{\theta_1}(\varepsilon)\right] \leq \varepsilon.$$

$\Rightarrow$ $\limsup_{n\to\infty} \mathbb{P}_{\theta_1}\left[\|r_n(\theta_1 - \hat{\theta}_n)\| > M^{\theta_1}(\varepsilon)\right] = 0$ as ε arbitrary.

Take n large enough so that $M^{\theta_1}(\varepsilon) < \sqrt{r_n}(\sqrt{r_n}\eta - 1)$.

Then $\limsup_{n\to\infty} \mathbb{P}_{\theta_1}\left[\sqrt{r_n}(\sqrt{r_n}\eta - 1) \leq r_n \|\theta_1 - \hat{\theta}_n\|\right] = 0$

$$\Rightarrow \liminf_{n\to\infty} 1 - \mathbb{P}_{\theta_1}\left[\sqrt{r_n}\,\left[\sqrt{r_n}\,\eta - 1\right] \le v_n \|\theta_1 - \hat{\theta}_n\|\right] = 1$$

$$\Rightarrow \liminf_{n\to\infty} \mathbb{P}_{\theta_1}\left[T_n = 1\right] = 1$$

Since this is true $\forall\, \theta_1 \in \Theta_1$,

$$\inf_{\theta_1 \in \Theta_1} \liminf_{\theta_1 \in \Theta_1} \mathbb{P}\left[T_n = 1\right] = 1$$

(b) Consider the following inductive construction of $\hat{\delta}_n$

We will follow a procedure at each step $\{1, \dots n, \dots N\}$ which requires splitting our parameter space into smaller subsets, performing test T_n^{subsets} on each and discard the rejected sets. We continue splitting sets with $T_n^{\text{subset}} = 0$, creating a subsequence of parameter spaces where $T_n = 0$ on each when evaluated. This is possible because $|\Theta| < \infty$.

- step 0: $\Theta^0 = \Theta$

Apply T^{Θ^0}. If $T^{\Theta^0} = 1$, there are no non-rejected regions so we pick any arbitrary $\theta_n^1 \in \Theta$

Otherwise, let $\mathcal{C}_0 := \{\Theta^0\}$, $R_0 = \emptyset$

- $\underline{\text{step } n}$: Split each $\Theta^{n-1}_{(i)} \in \mathcal{C}^{n-1}$ into $\Theta^{n(1)}_{(i)}, \Theta^{n(2)}_{(i)}$

both of size $\frac{|M|}{2^n}$ (This is possible due to $\|\Theta\| < \infty$

Hence $\exists M > 0$ s.t. $\|\Theta\| < \infty$). Then

$$\mathcal{C}_n := \left\{ \Theta^{n(j)}_{(i)} \,\middle|\, \Theta^{n(j)}_{(i)} \subseteq \Theta^{n-1}_{(i)} \text{ for some } \Theta^{n-1}_{(i)} \in \mathcal{C}_{n-1}, \right.$$
$$\left. \text{and } T_n^{\Theta^{n(j)}_{(i)}} = 0 \right\}$$

$$R_n = \left\{ \Theta^{n(j)}_{(i)} \,\middle|\, \Theta^{n(j)}_{(i)} \subseteq \Theta^{n-1}_{(i)} \text{ for some } \Theta^{n-1}_{(i)} \in \mathcal{C}_{n-1} \right.$$
$$\left. \text{and } T_n^{\Theta^{n(j)}_{(i)}} = 1 \right\}$$

If $\mathcal{C}_n = \phi$, then Pick $\hat{\delta}_n \in \Theta^{n(j)}_{(i)}$ for some $\Theta^{n(j)}_{(i)} \in \mathcal{C}_{n-1}$

otherwise, continue to step $n+1$.

After n iterations, Θ is split into a rejection region of T_n and regions of size $\frac{M^d}{2^n}$ and we have performed at most $\sum_{i=1}^{n} 2^{i-1} = 2^{n-1} - 1$ tests.

Consistency:

But since T_n has asymptotic power 1, so $\forall \varepsilon > 0$, any sequence that keeps a distance ε away from θ will be rejected w/ probability 1 as $n \to \infty$.

Indeed, under $\mathbb{P}_{\theta^*}$, $\forall \delta > 0$, $\varepsilon > 0$

$\forall \{\hat{\theta}_n\}_{n \geq 1} \in \Theta$ s.t. $d(\hat{\theta}_n, \theta^*) > \delta$, $\forall n$,

Then $$\liminf_{n \to \infty} \inf_{\theta \in (\theta^* - \frac{M^d}{2^n}, \theta^* + \frac{M^d}{2^n})} \|\theta - \theta^*\| > \delta$$

$$\Rightarrow \mathbb{P}_{\theta^*}\left[T_n^{\Theta_{(i)}^{(n-1)(j)}} = 0\right] \leq \varepsilon$$

Since by construction we have $\hat{\theta}_n \in \Theta_{(i)}^{(n-1)(j)}$

Further, since we have asymptotic level 0, $\forall$ subset of volume $\frac{M^d}{2^{n-1}}$ containing θ^* will be rejected with asymptotic probability 0.

We can find n s.t. $\forall \varepsilon > 0$, $\frac{M^d}{2^{n-1}} < \varepsilon$, and hence

$$\lim_{n \to \infty} \mathbb{P}_\theta\left[|\hat{\theta}_n - \theta^*| > \varepsilon\right] = 0$$

Hence $\hat{\theta}_n$ is consistent

2. Uniform vs Pointwise Testing

(a) WTS: $\inf_{T: \mathcal{X} \to \{0,1\}} \{ P_0[T=1] + P_1[T=0] \} = 1 - \| P_0 - P_1 \|_{TV}$

① WTS $\inf_{T: \mathcal{X} \to \{0,1\}} \{ P_0[T=1] + P_1[T=0] \} \leq 1 - \| P_0 - P_1 \|_{TV}$

Consider a sequence of events $\{A_n\}_{n \geq 1}$ which converges to some A, the event which achieves $|P_1[A] - P_0[A]| = \sup_{B \subseteq \mathcal{X}} |P_1[B] - P_0[B]|$

I.e. consider $\{A_n\}_{n \geq 1} \xrightarrow{P} A$ s.t.

$$|P_1[A] - P_0[A]| = \sup_{A \subseteq \mathcal{X}} |P_1(A) - P_0(A)|$$

It will be enough to show that $\exists\, T: \mathcal{X} \to \{0,1\}$ with
$P_0[T=1] + P_1[T=0] + \sup_{A \subseteq \mathcal{X}} |P_1(A) - P_0(A)| = 1$

If $P_1(A) \geq P_0(A)$: $T := \mathbb{1}_A$, and

$$\begin{aligned} &P_0[T=1] + P_1[T=0] - \sup_{A \subseteq \mathcal{X}} |P_1[A] - P_0[A]| \\ = \; &P_0[A] + P_1[A^c] + P_1[A] - P_0[A] \\ = \; &P_1[\mathcal{X}] = 1 \end{aligned}$$

If $P_1[A] < P_0[A]$: $T = \mathbb{1}_{A^c}$ and

$$P_0[T=1] + P_1[T=0] + \sup_{A \subset \mathcal{X}} |P_1[A] - P_0[A]|$$
$$= P_0[A^c] + P_1[A] + P_0[A] - P_1[A]$$
$$= P_0[A^c] + P_0[A]$$
$$= P_0[\mathcal{X}] = 1.$$

Hence $\exists T$ s.t. $f_{P_0,P_1}(T) := P_0[T=1] + P_1[T=0] + \|P_0 - P_1\|_{TV} = 1$

$$\Rightarrow \inf_{T:\mathcal{X}\to\{0,1\}} f_{P_0,P_1}(T) \leq 1.$$

(2) WTS $\inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + P_1[T=0] \geq 1 - \|P_0 - P_1\|_{TV}.$

Suppose WLOG $P_1[T=1] \geq P_0[T=1]$.

indeed $\inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + P_1[T=0]$

$$= \inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + (1 - P_1[T=1])$$

so if $P_1[T=1] \leq P_0[T=1]$ we can take $\hat{T} := \mathbb{1}_{\{T=0\}}$ as the new test and the test statistic would give

$$\mathbb{P}_1[\hat{T}=1] = \mathbb{P}_1[T=0] \geq \mathbb{P}_0[T=0] = \mathbb{P}_0[\hat{T}=1]$$

Consider the sequence $\{T_n\}_{n\geq 1}$ which converges to a limit statistic $T^*: \mathcal{X} \to \{0,1\}$ s.t.

$$\inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + P_1[T=0] = P_0[T^*=1] + P_1[T^*=0].$$

Define the corresponding sequence of r.vs. $A_n := \mathbb{1}_{\{T_n=1\}}$. $A_n \nearrow A$, with $A = \mathbb{1}_{\{T^*=1\}}$

Together, this gives $\inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + P_1[T=0] = P_0[A] + P_1[A^c]$.

Hence $\inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + P_1[T=0] + \sup_{B\subseteq\mathcal{X}} |P_1[B] - P_0[B]|$

$$= P_0[A] + P_1[A^c] + \sup_{\subseteq\mathcal{X}} |P_1[B] - P_0[B]|$$

$$\geq P_0[A] + P_1[A^c] + P_1[A] - P_0[A] = P_1[\mathcal{X}] = 1$$

This is what we wanted to show.

<u>Conclusion</u>: By ① & ②, $\inf_{T:\mathcal{X}\to\{0,1\}} P_0[T=1] + P_1[T=0] + \|P_1 - P_0\|_{TV} = 1$

(b) $H^2(P_0,P_1) = \frac{1}{2}\int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu$

WTS: $H^2(P_0,P_1) \underset{(1)}{\leq} \| P_0 - P_1 \|_{TV} \underset{(2)}{\leq} \sqrt{2H^2(P_0,P_1)\left(1-\frac{1}{2}\circ H^2(P_0,P_1)\right)}$

(1) WLOG, suppose $p_0(x) \geq p_1(x)$ $\forall x \in \mathcal{X}$,

$$|p_0(x)-p_1(x)| = p_0(x)-p_1(x) = p_0(x)+p_1(x) - 2\sqrt{p_1(x)}\sqrt{p_1(x)}$$
$$\geq p_0(x)+p_1(x) - 2\sqrt{p_1(x)}\sqrt{p_0(x)}$$

Hence $H^2(P_0,P_1) = \frac{1}{2}\int_{\mathcal{X}} p_0(x)+p_1(x)-2\sqrt{p_1(x)}\sqrt{p_0(x)} \, d\mu(x)$

$$\leq \frac{1}{2}\int_{\mathcal{X}} |p_0(x)-p_1(x)| \, d\mu(x)$$

$$= \| P_0 - P_1 \|_{TV}$$

(2) WTS

$$\left(\frac{1}{2}\int_{\mathcal{X}} |p_0-p_1| \, d\mu\right) \leq \left[2\cdot\frac{1}{2}\int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu\right]^{1/2} \left[1-\frac{1}{2}\cdot\frac{1}{2}\int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu\right]^{1/2}$$

$$\Leftrightarrow \left(\frac{1}{2}\int_{\mathcal{X}} |p_0-p_1| \, d\mu\right) \leq \left(\int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu\right)^{1/2} \left[1-\frac{1}{4}\int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu\right]^{1/2}$$

$$\Leftrightarrow \left(\frac{1}{2}\int_{\mathcal{X}} \sqrt{(p_0-p_1)^2} \, d\mu\right)^2 \leq \int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu \cdot \left(1-\frac{1}{4}\int_{\mathcal{X}} (\sqrt{p_0}-\sqrt{p_1})^2 \, d\mu\right)$$

By Cauchy-Schwartz, we have

$$\left(\int_{\mathcal{X}} \sqrt{(p_0 - p_1)^2}\, d\mu \right)^2 \leq \int_{\mathcal{X}} (\sqrt{p_0} - \sqrt{p_1})^2\, d\mu \underbrace{\int_{\mathcal{X}} (\sqrt{p_0} + \sqrt{p_1})^2\, d\mu}_{(*)}$$

Considering the $(*)$ term,

$$\frac{1}{4} \int_{\mathcal{X}} (\sqrt{p_0} + \sqrt{p_1})^2\, d\mu \leq 1 - \frac{1}{4} \int_{\mathcal{X}} (\sqrt{p_0} - \sqrt{p_1})^2\, d\mu$$

$$\Leftrightarrow \frac{1}{4} \int_{\mathcal{X}} 2(p_0 + p_1)\, d\mu \leq 1$$

$$\Leftrightarrow \int_{\mathcal{X}} (p_0 + p_1)\, d\mu \leq 4 \quad \Leftrightarrow \quad 2 \leq 4$$

Hence

$$\sqrt{2H^2(P_0, P_1)(1 - \tfrac{1}{2} H^2(P_0, P_1))} \geq \| P_0 - P_1 \|_{TV}$$

(c) WTS: $H^2(P_0^n, P_1^n) = 1 - (1 - H^2(P_0, P_1))^n$

$$\frac{1}{2}\int_{\mathcal{X}} (\sqrt{p_0} - \sqrt{p_1})^2 \, d\mu = \frac{1}{2}\int_{\mathcal{X}} p_0 + p_1 - 2\sqrt{p_0}\sqrt{p_1} \, d\mu$$

$$= \frac{1}{2}\left(2 - 2\int_{\mathcal{X}} \sqrt{p_0}\sqrt{p_1} \, d\mu\right)$$

$$= 1 - \int_{\mathcal{X}} \sqrt{p_0}\sqrt{p_1} \, d\mu .$$

$$\Rightarrow H^2(P_1^n, P_2^n) = 1 - \int_{\mathcal{X}} \prod_{i=1}^{n} \sqrt{p_0(x_i)}\sqrt{p_1(x_i)} \, d\mu^n(x)$$

$$= 1 - \int_{\mathcal{X}} \left(\prod_{i=1}^{n} p_0(x_i) p_1(x_i)\right)^{1/2} d\mu^n(x)$$

$$= 1 - \prod_{i=1}^{n} \int_{\mathcal{X}} (p_0(x_i) p_1(x_i))^{1/2} \, d\mu(x_i) \quad \text{independence}$$

$$= 1 - \left(\int_{\mathcal{X}} (p_0(x_i) p_1(x_i))^{1/2} \, d\mu(x_i)\right)^n \quad \text{iid}$$

$$= 1 - (1 - H^2(P_0, P_1))^n .$$

(d) i) Consider $\hat{\theta}_n := (\bar{X}_n, \hat{\sigma}_n)$, with $\hat{\sigma}_n := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$.

$\hat{\theta}_n$ is an efficient estimator for θ. (MLE for normal, which is "nice")

Let: $\left\{ \begin{array}{l} \mathcal{C}_n^{\theta} := \{ \mu \in \mathbb{R} \mid \mu^2 \hat{\sigma}_n^{-2} \le \frac{\theta}{n} \} \\ u_\alpha \text{ the } 1-\alpha \text{ quantile of } \chi_1^2 \text{ distribution} \\ T_n(\alpha) = \begin{cases} 1 & \text{if } \bar{X}_n \notin \mathcal{C}_n^{u_\alpha} \\ 0 & \text{otherwise} \end{cases} \end{array} \right.$

Asymptotic level α:

$$\mathbb{P}_{\theta_0}[T_n = 1] = \mathbb{P}_{\theta_0}[n \bar{X}_n^2 \hat{\sigma}^{-2} \le u_\alpha] = \mathbb{P}_{\theta_0}[(\sqrt{n}\bar{X}_n \hat{\sigma}^{-1})^2 \le u_\alpha] = \alpha$$

Asymptotic power 1:

If $\theta_1 \notin \Theta_0$,

$$\mathbb{P}_{\theta_1}[T_n = 1] = \mathbb{P}_{\theta_1}[n \bar{X}_n^2 \hat{\sigma}_n^{-2} > u_\alpha]$$

$$= \mathbb{P}_{\theta_1}[n(\bar{X}_n - \theta_1)^2 \hat{\sigma}_n^{-2} + 2n\bar{X}_n \theta_1 \hat{\sigma}_n^{-2} - n\theta_1^2 \hat{\sigma}_n^{-2} > u_\alpha]$$

$$= \mathbb{P}_{\theta_1}[n(\bar{X}_n - \theta_1)^2 \hat{\sigma}_n^{-2} + \underbrace{n\theta_1 \hat{\sigma}_n^{-2}(2\bar{X}_n - \theta_1)}_{A} > u_\alpha]$$

$\sqrt{n}(\bar{X}_n - \theta_1)\,\hat{\sigma}_n^{-1} \xrightarrow{d} N(0,1)$ by CLT ($\theta_1 < \infty$, $\sigma^2 < \infty$)

$\Rightarrow n(\bar{X}_n - \theta_1)^2 \hat{\sigma}_n^{-2} \xrightarrow{d} \chi^2$ by CMT.

and $2\bar{X}_n - \theta_1 \to \theta_1$ by WLLN, $\hat{\sigma}_n^{-2} \xrightarrow{P} \sigma^{-2}$ by WLLN & CMT,
Hence A grows linearly with n.

$\Rightarrow \forall \varepsilon, \alpha \ \exists N(\alpha)$ s.t. $\forall n > N(\alpha)$,

$$\mathbb{P}_{\theta_1}\left[n\theta_1 \hat{\sigma}_n^{-2} (2\bar{X}_n - \theta_1) > u_{1-\alpha} \right] \geq 1-\varepsilon.$$

Take $\varepsilon \downarrow 0$, we obtain

$$\mathbb{P}_{\theta_1}\left[n(\bar{X} - \theta_1)^2 \hat{\sigma}_n^{-2} + n\theta_1 \hat{\sigma}_n^{-2} (2\bar{X}_n - \theta_1) > u_{1-\alpha} \right] = 1 - o_{n\to\infty}(1)$$

This holds $\forall\, \theta_1 \in \Theta_1$, so $\inf\limits_{\theta_1 \in \Theta_1} \lim\limits_{n\to\infty} \mathbb{P}_{\theta_1}[T_n = 1] = 1$

This is what we wanted to show.

(iii) WTS $\forall n \in \mathbb{N}, \varepsilon > 0,$

$$\sup_{\theta_0 \in \Theta_0} \mathbb{P}_{\theta_0}[T=1] = \alpha \Rightarrow \forall \varepsilon > 0, \inf_{\theta_1 \in \Theta_1} \mathbb{P}_{\theta_1}[T=1] \leq \alpha + \varepsilon$$

(1) Let P_θ^n denotes the distribution of n variables sampled from P_θ.

From part (b), we know $H^2(P_{\theta_0}^n, P_{\theta_1}^n) \leq \| P_{\theta_0}^n - P_{\theta_1}^n \|_{TV}$

By the continuity of $P_{\theta_0}, P_{\theta_1}$ over $\mathcal{X}$, we take $\theta_1 \in \Theta \setminus \Theta_0$ s.t. the normal with mean zero is "close" enough to that with mean θ_1 to achieve total variation smaller than $1-(1-\varepsilon^2)^{n/2}$ i.e.

Pick $\theta_1 \in \Theta \setminus \Theta_0$ s.t. $\| P_{\theta_0}^n - P_{\theta_1}^n \|_{TV} < 1-(1-\varepsilon^2)^{n/2}$.

Hence $1-(1-\varepsilon^2)^{n/2} \geq \| P_{\theta_0}^n - P_1^n \|_{TV} \geq H^2(P_0^n, P_1^n)$ by (b)

By part (c), $H^2(P_{\theta_0}^n, P_{\theta_1}^n) = 1-(1-H^2(P_{\theta_0}, P_{\theta_1}))^n$

Hence $(H^2(P_{\theta_0}, P_{\theta_1}))^n \leq 1-\sqrt{1-\varepsilon^2}$

Let $H := H^2(P_{\theta_0}, P_{\theta_1})$.

Note $\sqrt{2H^n(1-\frac{1}{2}H^n)} \leq \varepsilon \Leftrightarrow 2H^n(1-\frac{1}{2}H^n) \leq \varepsilon^2$

$\Leftrightarrow -H^{2n} + 2H^n - \varepsilon^2 \leq 0$

$$\Leftrightarrow \underbrace{H^{2n} - 2H^n + \varepsilon^2}_{(*)} \geq 0$$

For equality to zero, we need $H^n = 1 \pm \dfrac{\sqrt{4-4\varepsilon^2}}{2} = 1 \pm \sqrt{1-\varepsilon^2}$

Since $(*)$ is a quadratic with H^{2n} positive coeff.

The above inequality in $(*)$ holds iff.

$$H^n \leq 1 - \sqrt{1-\varepsilon^2} \quad \text{or} \quad H^n \geq 1 + \sqrt{1-\varepsilon^2} \quad \text{since } H^n \leq 1.$$

In particular, for $H^n \leq 1-\sqrt{1-\varepsilon^2}$, $\sqrt{2H^n(1-H^n/2)} \leq \varepsilon$.

By part (b) again, we know $\sqrt{2H^n(1-H^n/2)} \geq \| P_{\theta_0}^n - P_{\theta_1}^n \|_{TV}$.

Hence $H^n \leq 1 - \sqrt{1-\varepsilon^2} \Rightarrow 1 - \| P_{\theta_0} - \mathbb{P}_{\theta_1} \| \geq 1 - \varepsilon$

We recall our goal: WTS $\forall$ n, $\varepsilon > 0$, and for some

T_n s.t. $\sup_{\theta_0 \in \Theta_0} P_{\theta_0}^n [T_n = 1] = P_{\theta_0}^n [T_n = 1] = \alpha$,

Then $\inf_{\theta_1 \in \Theta_1} P_{\theta_1}^n [T_n = 1] \leq \alpha + \varepsilon$.

Fix T_n s.t. $\sup_{\theta_0 \in \Theta_0} P_{\theta_0}^n [T_n = 1] = P_{\theta_0}^n [T_n = 1] = \alpha$.

Then WTS $\sup_{\theta_1 \in \Theta_1} P_{\theta_1}^n [T_n = 0] \geq 1 - \gamma - \varepsilon$.

$$\iff \sup_{\theta_1 \in \Theta_1} P_{\theta_1}^n [T_n = 0] + P_{\theta_0}^n [T_n = 0] \geq 1 - \varepsilon$$

By part (a),

$$P_{\theta_1}^n [T_n = 0] + P_{\theta_0}^n [T_n = 0]$$

$$\geq \inf_{T : \mathcal{X} \to \{0,1\}} P_{\theta_1}^n [T_n = 0] + P_{\theta_0}^n [T_n = 1]$$

$$= \| P_{\theta_0} - P_{\theta_1} \|_{TV}$$

And we showed above that $\| P_{\theta_0} - P_{\theta_1} \|_{TV} \geq 1 - \varepsilon$

The result follows.

3. Power

(a) ① Consider K_n^{sgn}

$Y_i := sgn(X_i)$ are iid, $Y_i = 1$ w/ probability $\frac{1}{2}$ and $Y_i = -1$ w/ probability $\frac{1}{2}$ under $\theta_0 = 0$

$\Rightarrow E_0[sgn(X_i)] = 0$ and $Var_0[sgn(X_i)] = 1$

By CLT, $\sqrt{n} S_n \xrightarrow{d} N(0,1)$.

Let $Z \sim N(0,1)$, and $z_{1-\alpha}$ is $1-\alpha$ quantile

Then $\mathbb{P}_0\left[S_n > \frac{z_{1-\alpha}}{\sqrt{n}}\right] = \mathbb{P}_0\left[\sqrt{n} S_n > z_{1-\alpha}\right]$

$$\xrightarrow[n\to\infty]{} \mathbb{P}\left[Z > z_{1-\alpha}\right] = \alpha.$$

Hence $K_n^{sgn} = \left(\frac{z_{1-\alpha}}{\sqrt{n}}, +\infty\right)$

② Consider K_n^{mean}, $\sqrt{n}\bar{X}_n \xrightarrow{d} N(0,1)$ under $\theta_0 = 0$,

Hence $\mathbb{P}_0\left[\bar{X}_n > \frac{z_{1-\alpha}}{\sqrt{n}}\right] = \alpha$.

Hence $K_n^{mean} = \left(\frac{z_{1-\alpha}}{\sqrt{n}}, +\infty\right)$ also.

(b) ① Consider $\lim_{n \to \infty} \Pi_n^{sgn}(\theta)$. Under P_θ, $\theta > 0$

$sgn(X_1) = -1$ w/ probability $\Phi(-\theta) =: z_{-\theta}$
$sgn(X_i) = 1$ w/ probability $1 - \Phi(-\theta) =: 1 - z_{-\theta}$

Hence $\mathbb{E}[sgn(X_i)] = 1 - 2z_{-\theta} > 0$
$\mathbb{V}[sgn(X_i)] = 4z_{-\theta}(1 - z_{-\theta})$

Under P_θ, $\theta > 0$, $\sqrt{n}\left(S_n - (1 - 2z_{-\theta})\right) \xrightarrow{d} N(0, 4z_{-\theta}(1 - z_\theta))$

Hence

$$\lim_{n \to \infty} P_\theta\left[S_n \in K_n^{sgn}\right] = \lim_{n \to \infty} \mathbb{P}_\theta\left[\sqrt{n}\left(S_n - (1 - 2z_{-\theta})\right) + \sqrt{n}\,(1 - 2z_{-\theta}) > z_{1-\alpha}\right]$$

Since $\sqrt{n}\left(S_n - (1 - 2z_{-\theta})\right) \xrightarrow[n \to \infty]{d} N(0, 4z_{-\theta}(1 - z_{-\theta}))$, and

Since $z_{-\theta} < \frac{1}{2}$ with $\theta > 0$, $\sqrt{n}\,(1 - 2z_{-\theta}) \xrightarrow[n \to \infty]{} +\infty$

Hence $\lim_{n \to \infty} P_\theta\left[S_n \in K_n^{sgn}\right] = 1$

② Consider $\lim_{n \to \infty} \Pi_n^{mean}(\theta)$. Under P_θ, $\theta > 0$,

$\sqrt{n}(\bar{X}_n - \theta) \xrightarrow{d} N(0,1)$.

Hence $\lim_{n\to\infty} \mathbb{P}_\theta \left[\bar{X}_n \in K_n^{mean} \right]$

$= \lim_{n\to\infty} \mathbb{P}_\theta \left[\sqrt{n}\left(\bar{X}_n - \theta\right) + \sqrt{n}\,\theta > z_{1-\alpha} \right]$

$\sqrt{n}\left(\bar{X}_n - \theta\right) \xrightarrow{d} \mathcal{N}(0,1)$, $\sqrt{n}\,\theta \to \infty$ as $n \to \infty$, hence

$\lim_{n\to\infty} \mathbb{P}_\theta \left[\bar{X}_n \in K_n^{mean} \right] = 1$

Hence the two estimators have equal asymptotic pointwise level and power, so they are equally good asymptotically.

c) Taking $\inf\limits_{\theta_1 \in \Theta_1}$ on both expressions obtained in b,

we can find a sequence $\{\theta_n\}_{n \geq 1} \subseteq \Theta_1$ with a subsequence converging to $\theta = 0$.

$\theta = 0$ minimizes both expressions for $\overline{\Theta}_1 = [0, +\infty)$,

as $n \to \infty$. Hence we shall take $\inf\limits_{\theta_n \to 0} \pi^{sgn / mean}(\theta_n)$

to achieve $\inf\limits_{\theta_1 \in \Theta_1} \Pi_n^{sgn/mean}(\theta)$

By (a), $\lim\limits_{n\to\infty} \Pi(\theta) = \alpha$ for $\theta = 0$.

Hence $\limsup\limits_{n\to\infty} \inf\limits_{\theta_1 \in \Theta_1} \Pi_n^{sgn}(\theta) = \limsup\limits_{n\to\infty} \inf\limits_{\theta_1 \in \Theta_1} \Pi_n^{mean}(\theta) = \alpha$

(d) Under P_{μ,σ^2}, $sgn(X_i) = -1$ w/ probability $z_{-\mu/\sigma^2}$

Again, $\sqrt{n}\left(S_n - (1 - 2q_{-\mu/\sigma^2})\right) \xrightarrow{d} N\left(0, 4 z_{\mu/\sigma^2}(1 - z_{-\mu/\sigma^2})\right)$

and $\mathbb{P}_{(0,\sigma^2)}\left[S_n \in K_n^{sgn}\right]$

$$= \mathbb{P}_{(0,1^2)}\Big[\underbrace{\sqrt{n}\left(S_n - (1 - 2z_{-0})\right)}_{\xrightarrow{d} N(0,1)} + \sqrt{n}\left(1 - 2z_0\right) \geq z_{1-\alpha}\Big]$$

$$1 - 2z_0 \overset{\shortparallel}{0}$$

Hence $\mathbb{P}_{(0,\sigma^2)}\left[S_n \in K_n^{sgn}\right] \to \alpha$ as $n\to\infty$ regardless of (μ, σ^2).

Hence $\limsup_{n\to\infty} \sup_{(\mu,\sigma^2)\in \Theta_0} \Pi^{sign}(\mu,\sigma^2) = \alpha$ (independent of μ, σ^2)

(2) Consider now the mean.

Under P_{μ,σ^2}, $\dfrac{\sqrt{n}(X_n - \mu)}{\sigma} \sim N(0,1)$ by CLT

Hence $\forall \sigma^2 > 0$,

$$\Pi_n^{mean}(0,\sigma^2) = \mathbb{P}_{(0,\sigma^2)}\left[\sqrt{n}\,\bar{X}_n > z_{1-\alpha}\right]$$

$$= \mathbb{P}_{(0,\sigma^2)}\left[\frac{\sqrt{n}\,\bar{X}_n}{\sigma} > \frac{z_{1-\alpha}}{\sigma}\right]$$

$$= \mathbb{P}_{(0,\sigma^2)}\left[Z > \frac{z_{1-\alpha}}{\sigma}\right], \quad Z \sim N(0,1).$$

Hence as $n \to \infty$, $\Pi_n^{mean}(0,\sigma^2)$ is maximized for $\sigma \to \infty$ in $\overline{\Theta}_0$.

Hence $\limsup_{n\to\infty} \sup_{(\mu,\sigma^2)\in\Theta_0} \Pi_n^{mean}(\mu,\sigma^2) = \mathbb{P}[Z > 0] = \frac{1}{2}$,

$Z \sim N(0,1)$.

(e) Again, for $\theta_1 = (\mu, \sigma^2) \in \Theta_1$,

$$\lim_{n\to\infty} \mathbb{P}_{\theta_1}\left[S_n \in K_n^{sgn} \right]$$

$$= \lim_{n\to\infty} \mathbb{P}_{\theta_1}\left[\sqrt{n}\left(S_n - (1-2z_{-\mu/\sigma^2})\right) + \sqrt{n}\left(1-2z_{-\mu/\sigma^2}\right) > z_{1-\alpha} \right]$$

$$\sqrt{n}\left(S_n - (1-2z_{-\mu/\sigma^2})\right) \xrightarrow{d} N\left(0, 4z_{-\mu/\sigma^2}(1-z_{-\mu/\sigma^2})\right)$$

$\sqrt{n}\left(1-2z_{-\mu/\sigma^2}\right) \to \infty$ as $n \to \infty$ since $q_{-\mu/\sigma^2} < \frac{1}{2}$ for $\mu/\sigma^2 > 0$.

Hence $\mathbb{P}_{\theta_1}\left[S_n \in K_n^{sgn} \right] \xrightarrow{n\to\infty} 1$.

and

$$\lim_{n\to\infty} \mathbb{P}_{\theta_1}\left[\bar{X}_n \in K_n^{mean} \right] = \lim_{n\to\infty} \mathbb{P}_{\theta_1}\left[\underbrace{\sqrt{n}(\bar{X}_n - \mu)}_{\xrightarrow{d} N(0,\sigma^2)} + \underbrace{\sqrt{n}\mu}_{\to\infty \text{ as } n\to\infty \text{ for } \mu>0} > z_{1-\alpha} \right] = 1$$

Hence they have equal asymptotic pointwise power (equal to 1).

But from (d) we saw that with sgn estimator we can control the level of the test, but for the mean it is constant $\frac{1}{2}$ regardless. Hence we prefer the sign test.

4. Comparing Tests

(a) ① Taylor Expansion of $\mathbb{P}_n \ell_{\hat\theta_n}$ around $\mathbb{P}_n \ell_{\theta_0}$

$$\mathbb{P}_n \ell_{\hat\theta_n} = \mathbb{P}_n \ell_{\theta_0} (\hat\theta_n - \theta_0) + \mathbb{P}_n \nabla \ell_{\theta_0} (\hat\theta_n - \theta_0) + \frac{1}{2} (\hat\theta_n - \theta_0)^T \mathbb{P}_n \nabla^2 \ell_{\theta_0} (\hat\theta_n - \theta_0)$$
$$+ (\hat\theta_n - \theta_0)^T \mathbb{P}_n R (\hat\theta_n - \theta_0)$$

with $\|\mathbb{P}_n \tilde{R}\|_{op} = 1$ as seen recurrently in the lectures.

Then $\Delta_n = n \left(\mathbb{P}_n \ell_{\hat\theta_n} - \mathbb{P}_n \ell_{\theta_0} \right)$

$$(i) \quad = n \left[\mathbb{P}_n \nabla \ell_{\theta_0} (\hat\theta_n - \theta_0) + \frac{1}{2} (\hat\theta_n - \theta_0)^T \mathbb{P}_n \nabla^2 \ell_{\theta_0} (\hat\theta_n - \theta_0) \right] + \sqrt{n} (\hat\theta_n - \theta_0) \mathbb{P}_n R (\hat\theta_n - \theta_0)$$

② Taylor expansion of $\mathbb{P}_n \nabla \ell_{\hat\theta_n}$ around $\mathbb{P}_n \nabla \ell_{\theta_0}$

$$(ii) \quad 0 = \nabla \mathbb{P}_n \ell_{\hat\theta_n} = \mathbb{P}_n \nabla \ell_{\theta_0} + \mathbb{P}_n \nabla^2 \ell_{\theta_0} (\hat\theta_n - \theta_0) + \mathbb{P}_n \tilde{R} (\hat\theta_n - \theta_0)$$

with $\|\mathbb{P}_n \tilde{R}\| = o_p(1)$ again, for a possibly different $\tilde{R}$ than before

With $(i) - (\hat\theta_n - \theta_0)^T n (ii)$

$$\Delta_n = -\frac{1}{2} \sqrt{n} (\hat\theta_n - \theta_0)^T \mathbb{P}_n \nabla^2 \ell_{\theta_0} \sqrt{n} (\hat\theta_n - \theta_0) + \sqrt{n} (\hat\theta_n - \theta_0)^T \mathbb{P}_n (R - \tilde{R}) \sqrt{n} (\hat\theta_n - \theta_0)$$

Note $\|\mathbb{P}_n (R - \tilde{R})\| = o_p(1)$, $\sqrt{n}(\hat\theta_n - \theta_0) = O_p(1)$ by asymptotic normality of the MLE, so

$$\sqrt{n} (\hat\theta_n - \theta_0)^T \mathbb{P}_n (R - \tilde{R}) \sqrt{n} (\hat\theta_n - \theta_0) = o_p(1).$$

Hence $2\Delta_n = -\sqrt{n}(\hat{\theta}_n-\theta_0)^T P_n \nabla^2 \ell_{\theta_0} \sqrt{n}(\hat{\theta}_n-\theta_0) + o_p(1)$

By WLLN, $P_n \nabla^2 \ell_{\theta_0} \to E[\nabla^2 \ell_{\theta_0}] = -I_{\theta_0}$

$\Rightarrow P_n \nabla^2 \ell_{\theta_0} = -I_{\theta_0} + o_p(1)$

$\Rightarrow 2\Delta_n = \sqrt{n}(\hat{\theta}_n-\theta_0)^T I_{\theta_0} \sqrt{n}(\hat{\theta}_n-\theta_0) - \sqrt{n}(\hat{\theta}_n-\theta_0)^T o_p(1) \sqrt{n}(\hat{\theta}_n-\theta_0)^T + o_p(1)$

$\sqrt{n}(\hat{\theta}_n-\theta_0)^T \xrightarrow{d} N(0, I)$ hence $O_p(1)$ so $\sqrt{n}(\hat{\theta}_n-\theta_0)^T o_p(1) \sqrt{n}(\hat{\theta}_n-\theta_0)^T = o_p(1)$

$\rightarrow 2\Delta_n = \sqrt{n}(\hat{\theta}_n-\theta_0)^T I_{\theta_0} \sqrt{n}(\hat{\theta}_n-\theta_0) + o_p(1)$.

② Asymptotic Wald Ellipse

By consistency of the MLE, continuity of I_θ, and the CMT, we have that $I_{\hat{\theta}_n} \xrightarrow{P} I_{\theta_0}$, i.e.

$$W_n = \sqrt{n}(\hat{\theta}_n-\theta_0)^T I_{\hat{\theta}_n} \sqrt{n}(\hat{\theta}_n-\theta_0) = \sqrt{n}(\hat{\theta}_n-\theta_0)^T I_{\theta_0} \sqrt{n}(\hat{\theta}_n-\theta_0) + \sqrt{n}(\hat{\theta}_n-\theta_0)^T o_p(1) \sqrt{n}(\hat{\theta}_n-\theta_0)^T$$

By same argument as above, $\sqrt{n}(\hat{\theta}_n-\theta_0)^T o_p(1) \sqrt{n}(\hat{\theta}_n-\theta_0)^T = o_p(1)$

Hence $2\Delta_n = \sqrt{n}(\hat{\theta}_n-\theta_0) I_{\hat{\theta}_n} \sqrt{n}(\hat{\theta}_n-\theta_0) + o_p(1)$

$\Rightarrow 2\Delta_n - W_n = o_p(1)$ which is what we wanted to prove

(b) Let us consider X_i's with Bernoulli density.

$$p_\theta(x) = \begin{cases} \theta & \text{if } X=1 \\ 1-\theta & \text{if } X=0 \end{cases}, \quad \ell_\theta(x) = \begin{cases} \log\theta & \text{if } X=1 \\ \log(1-\theta) & \text{if } X=0 \end{cases}$$

$$\nabla \ell_\theta(x) = \begin{cases} 1/\theta & \text{if } X=1 \\ -1/(1-\theta) & \text{if } X=0 \end{cases}, \quad \nabla^2 \ell_\theta = \begin{cases} -1/\theta^2 & \text{if } X=1 \\ -1/(1-\theta)^2 & \text{if } X=0 \end{cases}$$

Hence $\mathbb{E}[\nabla^2 \ell_\theta] = \dfrac{-\theta}{\theta^2} + \dfrac{1-\theta}{(1-\theta)^2} = \dfrac{-(1-\theta)-\theta}{\theta(1-\theta)} = \dfrac{-1}{\theta(1-\theta)}$

So $I_\theta = -\mathbb{E}[\nabla^2 \ell_{\theta_0}] = \dfrac{1}{\theta(1-\theta)}$.

$$P_n \ell_\theta(x) = \frac{1}{n}\sum_{i=1}^n \left(X_i \log(\theta) + (1-X_i)\log(1-\theta)\right) = \bar{X}_n \log(\theta) + (1-\bar{X}_n)\log(1-\theta)$$

$$P_n \nabla \ell_\theta(x) = \frac{1}{n}\sum_{i=1}^n X_i \frac{1}{\theta} + (1-X_i)\left(\frac{-1}{1-\theta}\right) = \frac{\bar{X}_n}{\theta} - \frac{1-\bar{X}_n}{1-\theta}$$

Since $\hat{\theta}_n$ is the solution to $P_n \nabla \ell_\theta(x) = 0 \Leftrightarrow \dfrac{\bar{X}_n}{\theta} = \dfrac{1-\bar{X}}{1-\theta}$, then $\hat{\theta}_n = \bar{X}_n$, and the obtained tests are as follows

GLR: $2\Lambda_n = n\left(P_n \ell_{\hat{\theta}_n} - P_n \ell_\theta\right)$

$$= n\left[\bar{X}_n \log(\bar{X}_n/\theta) + (1-\bar{X}_n)\log\left(\frac{1-\bar{X}_n}{1-\theta}\right)\right]$$

The Wald Ellipse is $W_n = \dfrac{n(\bar{X}_n-\theta)^2}{\bar{X}_n(1-\bar{X}_n)}$

Consider a sample with $n=10$, $\theta = \frac{1}{2}$, $\bar{X}_n = \frac{4}{10}$

$$2\Delta_n = 4\log(4/5) + 6\log(6/5) = 4\log 4 + 6\log 6 - 10\log 5 \simeq 0.2$$

$$W_n = 10(1/100)/(24/100) = 10/24 = 5/12 \simeq 0.4.$$

(c) By consistency of the MLE and CMT, under θ^* we have

$$(\hat{\theta}_n - \theta_0)^T I_{\hat{\theta}_n} (\hat{\theta}_n - \theta_0) \xrightarrow{P} (\theta^* - \theta_0)^T I_{\theta^*} (\theta^* - \theta_0)$$

Since $I_{\theta^*} > 0$ (given), then $(\theta^* - \theta_0)^T I_{\theta^*} (\theta^* - \theta_0) \geq 0$

Hence $\frac{W_n}{n} \xrightarrow{P} c$, $c > 0$.

Hence W_n diverges (in probability) to $+\infty$, hence we have asymptotic power of 1 for W_n, i.e.

$$\forall \gamma < \infty, \quad \lim_{n\to\infty} \mathbb{P}_{\theta^*}[W_n > \gamma] = 1$$

For GLR,

$$\begin{aligned} \frac{\Delta_n}{n} &= P_n \ell_{\hat{\theta}_n} - P_n \ell_{\theta_0} \\ &= P_n \ell_{\hat{\theta}_n} - P_n \ell_{\theta^*} + P_n \ell_{\theta^*} - P_n \ell_{\theta_0} \\ &= o_p(1) + \mathbb{E}_{\theta^*}[\ell_{\theta^*} - \ell_{\theta_0}] \\ &= o_p(1) + D_{KL}(p_{\theta^*} \| p_{\theta_0}) \end{aligned}$$

Note: $\{p_\theta\}_{\theta \in \Theta}$ is identifiable by consistency $\forall \theta \in \Theta$, so

$D_{KL}(p_{\theta^*} \| p_\theta) > 0$ for $\theta^* \neq \theta$. Hence

$\lim_{n\to\infty} \mathbb{P}_{\theta^*}[\Delta_n > \gamma] = 1$ and we have asymptotic power of 1 for Δ_n.

(d) Under P_{θ}, using results from lecture 7,

$$\log L_n = n(\mathbb{P}_n \ell_{\theta_n} - \mathbb{P}_n \ell_{\theta_0}) = \sqrt{n}\, h^T \mathbb{P}_n \nabla \ell_{\theta} + \frac{1}{2} h^T I_{\theta_0} h + o_p(\|h\|)$$

$$\Rightarrow \log L_n \xrightarrow[P_{\theta_0}]{d} N\left(-\frac{1}{2} h^T I_{\theta_0} h,\ h^T I_{\theta_0} h\right)$$

From (a), $2\Delta_n - W_n \xrightarrow[P_{\theta_0}]{P} 0$.

By Slutsky's lemma,

$$(2\Delta_n - W_n,\ \log L_n) \xrightarrow[P_{\theta_0}]{d} N\left(\begin{bmatrix} 0 \\ h^T I_{\theta_0} h \end{bmatrix},\ \begin{bmatrix} 0 & 0 \\ 0 & (h^T I_{\theta_0} h)^2 \end{bmatrix}\right)$$

By Corollary 4.5 in revised notes,

$$2\Delta_n - W_n \xrightarrow[P_{\theta_n}]{d} 0 \quad \rightarrow \quad 2\Delta_n - W_n \xrightarrow[P_{\theta_n}]{P} 0$$

By corollary to Slutky's lemma.

(e) Because we have a "nice" family, $\{P_\theta\}_{\theta\in\Theta}$, sequences

$$(P_{\theta_0})^{\otimes n} \triangleleft\triangleright \left(P_{\theta_0+\frac{1}{\sqrt{n}}h}\right)^{\otimes n}$$, and

$$\log L_n = \sqrt{n}\,\mathbb{P}_n h^T \nabla \ell_{\theta_0} - \frac{1}{2} h^T I_{\theta_0} h + o_p(1)$$ (as shown in the lectures)

Since $\sqrt{n}\,\mathbb{P}_n \nabla \ell_\theta \xrightarrow[P_{\theta_0}]{d} N(0, I_\theta)$, by delta method

$$\log L_n \xrightarrow[P_{\theta_0}]{d} N\left(-\frac{1}{2} h^T I_{\theta_0} h,\ h^T I_{\theta_0} h\right)$$

In addition, we take $\sqrt{n}(\hat{\theta}_n - \theta_0) = \sqrt{n}\, I_{\theta_0}^{-1} \mathbb{P}_n \nabla \ell_{\theta_0} + o_p(1)$
(cf lecture 6)

Hence $\left(\sqrt{n}(\hat{\theta}_n - \theta_0),\ \log L_n\right) \xrightarrow[P_{\theta_0}]{d} N\left(\begin{bmatrix} 0 \\ -\frac{1}{2} h^T I_{\theta_0} h \end{bmatrix}, \begin{bmatrix} I_{\theta_0}^{-1} & h^T \\ h & h^T I_{\theta_0} h \end{bmatrix}\right)$

Using Le Cam's third Lemma, $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow[P_{\theta_0+\frac{1}{\sqrt{n}}h}]{d} N(h, I_{\theta_0}^{-1})$

By CMT, $P_{\theta_n}[W_n > \gamma] \to P_{Z\sim N(0,1)}\left[\| Z + I_\theta^{1/2} h \|^2 > \gamma \|\right]$

i.e. $\lim_{n\to\infty} P_{\theta_n}[W_n > \gamma] = P_{Z\sim N(0,1)}\left[\| Z + I_{\theta_0}^{1/2} h \|^2 > \gamma\right]$

(f) The set of null hypotheses that we reject when $W_n \geq \gamma$

is all points which are not in an ellipsoid defined by

$$\mathcal{E}_n^\gamma := \{ \theta \in \mathbb{R}^d / \sqrt{n}(\theta - \hat{\theta}_n)^T I_{\hat{\theta}_n} \sqrt{n}(\theta - \hat{\theta}_n) \leq \gamma \} \text{ , ie.}$$

centered around the MLE and with intervals that shrink

with rate relative to the covariance matrix of the MLE.

The GLR rejection region are not necessarily centered around the MLE.

It reflects the curvature of the likelihood function and may be

less interpretable. But asymptotically, the likelihood surface

becomes sharper and the rejection region tighter around $\hat{\theta}_n$.

For $P_\theta = Poi(\theta)$, $p_\theta(k) = \dfrac{\theta^k e^{-\theta}}{k!}$

$$l_\theta(k) = k \log\theta - \theta - \log(k!)$$

$$\nabla l_\theta(k) = \frac{k}{\theta} - 1 \quad , \quad P_n l_\theta(k) = \frac{1}{n}\sum_i X_i \log\theta - \theta - \frac{1}{n}\sum \log(X_i!)$$

Hence $\mathbb{P}_n D\ell_\theta(\omega) = \frac{\sum_i X_i}{n\theta} - 1 \Rightarrow \hat{\theta}_n$ s.t. $\hat{\theta}_n = \frac{\sum_i X_i}{n} \Rightarrow \hat{\theta}_n = \bar{X}_n$

$I_\theta = -\mathbb{E}\left[\nabla^2 \ell_\theta\right] = -\frac{-\theta}{\theta^2} = \frac{1}{\theta}$

For the GLR, $2\Delta_n > \gamma$

$$\Leftrightarrow 2\sum_{i=1}^n X_i \log\left(\frac{\hat{\theta}_n}{\theta_0}\right) - n(\hat{\theta}_n - \theta_0) > \gamma$$

For the Wald Ellipse,

$$W_n > \gamma$$

$$\Leftrightarrow n(\hat{\theta}_n - \theta_0)\frac{1}{\hat{\theta}_n}(\hat{\theta}_n - \theta_0) > \gamma$$

$$\Leftrightarrow n(\bar{X}_n - \theta_0)^2 / \bar{X}_n > \gamma$$

$$\Leftrightarrow \theta_0 \in \left]\bar{X}_n - \sqrt{\frac{\gamma \bar{X}_n}{n}}, \bar{X}_n + \sqrt{\frac{\gamma \bar{X}_n}{n}}\right[.$$