Set: $Y_1, \dots, Y_n \overset{iid}{\sim} f_Y(y|\underset{\sim}{\theta})$ where $\underset{\sim}{\theta} \in \Omega \subseteq \mathbb{R}^p$

↳ Could also be multivariate → $f_Y(y|\underset{\sim}{\theta}, \underset{\sim}{\xi})$

Likelihood: $L(\underset{\sim}{\theta} | \underset{\sim}{Y})$

Log likelihood: $\ell(\underset{\sim}{\theta}|\underset{\sim}{Y}) = \log L(\underset{\sim}{\theta}|\underset{\sim}{Y}) = \sum_{i=1}^{n} \log f_Y(Y_i|\underset{\sim}{\theta})$

(Score) likelihood equation (Sample-level)

$$S(\underset{\sim}{\theta}|\underset{\sim}{Y}) = \frac{\partial \ell(\underset{\sim}{\theta}|\underset{\sim}{Y})}{\partial \underset{\sim}{\theta}} = \sum_{i=1}^{n} \frac{\frac{\partial}{\partial \theta} f_Y(Y_i|\underset{\sim}{\theta})}{f_Y(Y_i|\underset{\sim}{\theta})} = 0$$

Information Matrix (Population)

$$I(\underset{\sim}{\theta}) = E_{\underset{\sim}{\theta}}\left[\frac{\partial \log f_Y(Y|\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} \cdot \frac{\partial \log f_Y(Y|\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}^T} \right]$$

$$\frac{1}{n}\sum_{i=1}^{n} -\frac{\partial^2}{\partial\theta\partial\theta^T} \log f_Y(Y_i|\underset{\sim}{\theta})$$

$$= E_{\underset{\sim}{\theta}}\left[-\frac{\partial^2}{\partial\underset{\sim}{\theta}\partial\underset{\sim}{\theta}^T} \log f_Y(Y|\underset{\sim}{\theta}) \right]$$

Assume that all necessary regularity conditions hold.

Let $\underset{\sim}{\theta}_0$ be the true value of $\underset{\sim}{\theta}$ "true" generated the sample.

and $\hat{\underset{\sim}{\theta}}$ be the consistent root to the likelihood equation.

Then: 1. $\sqrt{n}\left(\frac{1}{n} S(\underset{\sim}{\theta}_0|\underset{\sim}{Y})\right) \xrightarrow{d} N(\underset{\sim}{0}, I(\underset{\sim}{\theta}_0))$

$Y_i \sim f_Y(y|\theta_0)$

CLT

$$E_\theta\left[\frac{\frac{\partial}{\partial\theta} f_Y(Y_i|\theta_0)}{f_Y(Y_i|\theta_0)}\right] = \int \frac{\frac{\partial}{\partial\theta} f_Y(y|\theta)}{\cancel{f(y|\theta)}} \times \cancel{f_Y(y|\theta_0)}\, dy$$

$$= \int \frac{\partial}{\partial\theta} f_Y(y|\theta)\, dy$$

$$= \frac{\partial}{\partial\theta} 1 = 0$$

$$Var_\theta\left[\frac{\frac{\partial}{\partial\theta} f_Y(Y_i|\theta_0)}{f_Y(Y_i|\theta_0)}\right] = I(\theta_0)$$

2. $\sqrt{n}\, I(\underset{\sim}{\theta}_0)^{-1} \cdot \frac{1}{n} S(\underset{\sim}{\theta}_0|\underset{\sim}{Y}) = \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}_0) + o_p(1)$

$$0 = S(\hat{\underset{\sim}{\theta}} | \underset{\sim}{Y}) = S(\underset{\sim}{\theta_0} | \underset{\sim}{Y}) + \frac{\partial S(\theta_0 | \underset{\sim}{Y})}{\partial \theta} (\hat{\theta} - \theta_0)$$

$$+ \frac{1}{2} (\hat{\theta} - \theta_0)^T \frac{\partial^2 S(\theta^* | \underset{\sim}{Y})}{\partial \theta \, \partial \theta^T} (\hat{\theta} - \theta_0)$$, where θ^* between $\hat{\theta}$ and θ_0

$$\sqrt{n}(\hat{\theta} - \theta_0) = \left(\frac{\partial S(\theta_0 | \underset{\sim}{Y})}{\partial \theta} \right)^{-1} \left[- \underbrace{\frac{\sqrt{n}}{2} (\hat{\theta} - \theta_0)^T}_{O_p(1)} \frac{\partial^2 S(\theta^* | \underset{\sim}{Y})}{\partial \theta \partial \theta^T} \underbrace{(\hat{\theta} - \theta_0)}_{o_p(1)} - \sqrt{n} S(\theta_0 | \underset{\sim}{Y}) \right] \rightarrow o_p(1)$$

$$= \left(\underbrace{- \frac{1}{n} \frac{\partial S(\theta_0 | Y)}{\partial \theta}}_{} \right)^{-1} \times \frac{1}{n} \times \sqrt{n} \, S(\theta_0 | Y)$$

$$= \left\{ I(\theta_0) + o_p(1) \right\}^{-1} \times \frac{S(\theta_0 | Y)}{\sqrt{n}}$$

3. $$\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta_0}) \xrightarrow{d} N(0, I(\theta_0)^{-1})$$

Now, we consider two hypothesis testing situation.

1. Simple H_0: $H_0: \underset{\sim}{\theta} = \underset{\sim}{\theta}^*$ vs $H_1: \underset{\sim}{\theta} \neq \underset{\sim}{\theta}^*$ (two sided)
2. Composite H_0: $H_0: h(\underset{\sim}{\theta}) = \underset{\sim}{0}$ vs $H_1: h(\underset{\sim}{\theta}) \neq \underset{\sim}{0}$ (two sided)

   where $h: \mathbb{R}^p \rightarrow \mathbb{R}^q$: $q \leq p$ (often $q = 1$)

Focuse on Simple H_0:

1. Wald → 1943
2. Score → 1947
3. LR (Likelihood Ratio) → (1928)

Wald test: under H_0: $\underset{\sim}{\theta}_0 = \underset{\sim}{\theta}^*$ holds

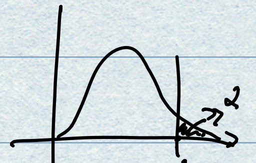
(a) $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I(\theta_0)^{-1})$

So under H_0: $\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*) \xrightarrow{d} N(0, I(\theta_0)^{-1})$

(b) $I(\hat{\underset{\sim}{\theta}}) \xrightarrow{p} I(\underset{\sim}{\theta}^*) \left(= I(\underset{\sim}{\theta}_0) \text{ under } H_0\right)$

Test statistic: $T_W = n(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T I(\hat{\underset{\sim}{\theta}})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)$

under H_0: $T_W \xrightarrow[H_0]{d} \chi^2_p$


$\chi^2_{1-\alpha,p}$

Rejection region: $T_W > \chi^2_{1-\alpha,p}$

$$P(T_W > \chi^2_{1-\alpha,p} \mid H_0) = \alpha$$

Power? $P(T_W > \chi^2_{1-\alpha,p} \mid H_1)$

under H_1: $\underset{\sim}{\theta} \neq \underset{\sim}{\theta}^*$

$$P(T_W > \chi^2_{1-\alpha,p} \mid \underset{\sim}{\theta})$$

under H_1: $T_W = n(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta} + \underset{\sim}{\theta} - \underset{\sim}{\theta}^*)^T I(\hat{\underset{\sim}{\theta}})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta} + \underset{\sim}{\theta} - \underset{\sim}{\theta}^*)$

$$= \underbrace{n(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T I(\hat{\underset{\sim}{\theta}})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})}_{\xrightarrow{d} \chi^2_p} + \underbrace{n(\underset{\sim}{\theta} - \underset{\sim}{\theta}^*)^T I(\hat{\underset{\sim}{\theta}})(\underset{\sim}{\theta} - \underset{\sim}{\theta}^*)}_{\text{increase as } \|\underset{\sim}{\theta} - \underset{\sim}{\theta}^*\| \text{ becomes large}} + \underbrace{2n(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T I(\hat{\underset{\sim}{\theta}})(\underset{\sim}{\theta} - \underset{\sim}{\theta}^*)}$$

$\theta_0 = \underset{\sim}{\theta}$

$$2 \times \sqrt{n} \times \underbrace{\sqrt{n}(\hat{\underset{\sim}{\theta}} - \theta)^T}_{O_p(1)} \cdot \underbrace{I(\hat{\theta})}_{O_p(1)} (\underset{\sim}{\theta} - \underset{\sim}{\theta}^*)$$

$$P(T_W > \chi^2_{1-\alpha,p} \mid \underset{\sim}{\theta}) \nearrow 1 \text{ as } \|\underset{\sim}{\theta} - \underset{\sim}{\theta}^*\| \nearrow$$

why $T_W \xrightarrow[H_0]{d} \chi^2_p$

$$T_W = n(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T I(\hat{\underset{\sim}{\theta}})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)$$

$$= \underbrace{\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T}_{O_p(1)} \times \underbrace{[I(\hat{\underset{\sim}{\theta}}) - I(\theta_0)]}_{o_p(1)} \times \underbrace{\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)}_{O_p(1)}$$
$$+ \underbrace{\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T}_{\downarrow d} I(\theta_0) \times \underbrace{\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)}_{\downarrow d}$$
$$= o_p(1) + N(0, I(\theta_0)^{-1})^T I(\theta_0) \cdot N(0, I(\theta_0)^{-1})$$
$$= o_p(1) + \chi^2_p$$

Calculating $I(\underset{\sim}{\theta})$ is very difficult.

Wald test is easy to apply. $\sqrt{n}(\hat{\theta}-\theta_0) \xrightarrow{d} N(0, \Sigma(\theta_0))$

Score-test: motivated by the fact that under H_0

$$\sqrt{n}\left(\frac{1}{n} S(\underset{\sim}{\theta}^* | \underset{\sim}{Y})\right) \overset{H_0}{=} \underline{\sqrt{n}\left(\frac{1}{n} S(\underset{\sim}{\theta}_0 | \underset{\sim}{Y})\right) \xrightarrow{d} N(0, I(\theta_0))}$$

Test Statistic:

$$T_S = S(\underset{\sim}{\theta}^* | \underset{\sim}{Y}) [n I(\theta^*)]^{-1} S(\underset{\sim}{\theta}^* | \underset{\sim}{Y}) \xrightarrow[H_0]{d} \chi_p^2$$

Rejection Region: $T_S > \chi^2_{1-\alpha, p}$

LR-test:

$$T_{LR} = -2\log\left[\frac{\sup_{H_0} L(\underset{\sim}{\theta} | \underset{\sim}{Y})}{\sup_{\Omega} L(\underset{\sim}{\theta} | \underset{\sim}{Y})}\right]$$

under H_0: $\underset{\sim}{\theta} = \underset{\sim}{\theta}^*$

$$T_{LR} = -2\log\left[\frac{L(\underset{\sim}{\theta}^* | \underset{\sim}{Y})}{L(\hat{\underset{\sim}{\theta}} | \underset{\sim}{Y})}\right]$$ where $\hat{\theta}$ is the global MLE

we have

$$T_{LR} \xrightarrow[H_0]{d} \chi_p^2 \quad \text{as } n \to +\infty$$

Rejection Region: $T_{LR} > \chi^2_{1-\alpha, p}$

illustration: $\frac{1}{2} T_{LR} = \log L(\hat{\underset{\sim}{\theta}}) - \log L(\underset{\sim}{\theta}^*)$

$$\log L(\hat{\underset{\sim}{\theta}}) = \log L(\underset{\sim}{\theta}^*) + \sqrt{n}(\hat{\theta}-\theta^*)^T \frac{1}{\sqrt{n}} \frac{\partial}{\partial\theta} \log L(\underset{\sim}{\theta}^*)$$
$$+ \frac{1}{2}\sqrt{n}(\hat{\underset{\sim}{\theta}}-\underset{\sim}{\theta}^*)^T \left\{\frac{1}{n} \frac{\partial^2}{\partial\theta\partial\theta^T} \log L(\tilde{\underset{\sim}{\theta}})\right\} \sqrt{n}(\hat{\underset{\sim}{\theta}}-\underset{\sim}{\theta}^*)$$

where $\tilde{\underset{\sim}{\theta}}$ is between $\hat{\underset{\sim}{\theta}}$ and $\underset{\sim}{\theta}^*$.

$$T_{LR} = 2\sqrt{n}(\hat{\theta}-\theta^*)^T \frac{1}{\sqrt{n}} \frac{\partial}{\partial\theta} \log L(\underset{\sim}{\theta}^*)$$

$$+ \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T \left\{ \frac{1}{n} \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) \right\} \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)$$

$$\frac{\partial}{\partial \theta} \log L(\underset{\sim}{\theta}^*) = \frac{\partial}{\partial \theta} \log L(\hat{\theta}) + \frac{\partial^2}{\partial \theta \partial \theta^T} \log L(\tilde{\underset{\sim}{\theta}}) (\underset{\sim}{\theta}^* - \hat{\underset{\sim}{\theta}})$$

where $\tilde{\underset{\sim}{\theta}}$ is between $\underset{\sim}{\theta}^*$ and $\hat{\underset{\sim}{\theta}}$

$\hat{\underset{\sim}{\theta}}$ solves $\frac{\partial}{\partial \theta} \log L(\theta) = 0$

meaning $\frac{\partial}{\partial \theta} \log L(\hat{\theta}) = 0$

$$\frac{\partial}{\partial \theta} \log L(\theta^*) = - \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) (\hat{\theta} - \theta^*)$$

$$T_{LR} = -2\sqrt{n}(\hat{\theta} - \theta^*)^T \cdot \frac{1}{n} \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) \sqrt{n}(\hat{\theta} - \theta^*)$$

$$+ \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T \frac{1}{n} \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) \sqrt{n}(\hat{\theta} - \theta^*)$$

$$= \sqrt{n}(\hat{\theta} - \underset{\sim}{\theta}^*)^T \cdot \left\{ -\frac{2}{n} \cdot \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) + \frac{1}{n} \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) \right\} \sqrt{n}(\hat{\theta} - \theta^*)$$

under H_0, θ^* is the true value. $\hat{\theta} \xrightarrow{P} \theta^*$

$$\sqrt{n}(\hat{\theta} - \theta^*) \xrightarrow{d} N(0, I^{-1}(\theta^*))$$

$$-\frac{1}{n} \frac{\partial^2}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \log L(\tilde{\underset{\sim}{\theta}}) \xrightarrow{P} I(\theta^*)$$

$$-\frac{1}{n} \frac{\partial^2}{\partial \theta \partial \theta^T} \log L(\tilde{\underset{\sim}{\theta}}) \xrightarrow{P} I(\theta^*)$$

$$T_{LR} \xrightarrow[H_0]{d} N(0, I^{-1}(\theta^*))^T \cdot I(\theta^*) \; N(0, I^{-1}(\theta^*))$$

$$= \chi^2_p$$

Summary. $H_0: \underset{\sim}{\theta} = \underset{\sim}{\theta}^*$ vs $H_1: \underset{\sim}{\theta} \neq \underset{\sim}{\theta}^*$

$$\left.\begin{matrix} T_W \\ T_S \\ T_{LR} \end{matrix}\right\} \xrightarrow[H_0]{d} \chi^2_p \quad \text{as } n \to +\infty \qquad RR: \begin{matrix} T_W \\ T_S \\ T_{LR} \end{matrix} > \chi^2_{1-\alpha, p}$$

Eg: $Y_1, \dots, Y_n \overset{iid}{\sim} \text{Bernoulli}(\theta)$ $\theta \in \mathbb{R}$

$H_0: \theta = \theta^*$, $H_1: \theta \neq \theta^*$

$$L(\theta|\underset{\sim}{Y}) = \prod_{i=1}^{n} \theta^{Y_i}(1-\theta)^{1-Y_i}$$

$$\ell(\theta|Y) = \left(\sum_{i=1}^{n} Y_i\right)\log\theta + \left(n - \sum_{i=1}^{n} Y_i\right)\log(1-\theta)$$

$$S(\theta|Y) = \frac{\partial}{\partial\theta}\ell(\theta|Y) = \frac{\sum Y_i}{\theta} - \frac{n-\sum Y_i}{1-\theta} = \frac{\sum Y_i - n\theta}{\theta(1-\theta)} = 0$$

$$\hat{\theta} = \bar{Y}_n$$

Find $I(\theta)$

$$\frac{\partial^2}{\partial\theta^2}\log f_Y(y|\theta) = \frac{\partial^2}{\partial\theta^2}\left(y\log\theta + (1-y)\log(1-\theta)\right)$$

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

$$I(\theta) = E\left[-\frac{\partial^2}{\partial\theta^2}\log f_Y(Y|\theta)\right]$$

est $\hat{I}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left\{-\frac{\partial^2}{\partial\theta^2}\log f_Y(Y_i|\theta)\right\}$

$$= E\left[-\left(-\frac{Y}{\theta^2} - \frac{1-Y}{(1-\theta)^2}\right)\right]$$

$$= \frac{E(Y)}{\theta^2} + \frac{E(1-Y)}{(1-\theta)^2} = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}$$

$$T_W = n(\hat{\theta} - \theta^*) \cdot I(\hat{\theta})(\hat{\theta} - \theta^*)$$

$$= \frac{n(\hat{\theta} - \theta^*)^2}{\hat{\theta}(1-\hat{\theta})} = \frac{n(\bar{Y}_n - \theta^*)^2}{\bar{Y}_n(1-\bar{Y}_n)}$$

$$T_S = S(\theta^*)[nI(\theta^*)]^{-1}S(\theta^*)$$

$$= \frac{\sum Y_i - n\theta^*}{\theta^*(1-\theta^*)} \frac{\theta^*(1-\theta^*)}{n} \frac{\sum Y_i - n\theta^*}{\theta^*(1-\theta^*)} = \frac{n(\bar{Y}_n - \theta^*)^2}{\theta^*(1-\theta^*)}$$

$$T_{LR} = -2\left[\log L(\theta^*) - \log L(\hat{\theta})\right]$$

$$= -2\left[\left(\sum_{i=1}^{n} Y_i\right)\log\left(\frac{\theta^*}{\bar{Y}_n}\right) + \left(n-\sum_{i=1}^{n} Y_i\right)\cdot\log\left(\frac{1-\theta^*}{1-\bar{Y}_n}\right)\right]$$

RR $\quad \begin{matrix} T_W \\ T_S \\ T_{LR} \end{matrix} > \chi^2_{1-\alpha,\, p=1}$

Power consideration. (simple H_0) $\quad H_0: \theta = \theta^*$

$H_1: \theta \neq \theta^*$

Suppose that $\underset{\sim}{\Delta} \in \mathbb{R}^p$ is fixed and consider

$$H_1: \underset{\sim}{\theta} = \underset{\sim}{\theta}^* + \underset{\sim}{\Delta}$$

assume H_1 is true : $\theta_0 = \underset{\sim}{\theta}^* + \underset{\sim}{\Delta}$

$$N(0, I^{-1}(\theta_0)) \leftarrow \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}_0) = \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^* - \underset{\sim}{\Delta})$$

$$= \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*) - \sqrt{n}\underset{\sim}{\Delta}$$

$$\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*) = \underbrace{\sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}_0)}_{\downarrow d \; N(0, I^{-1}(\theta_0))} + \underbrace{\sqrt{n}\underset{\sim}{\Delta}}_{\downarrow\;\infty}$$

$$\left.\begin{matrix} T_W \\ T_S \\ T_R \end{matrix}\right\} \xrightarrow{d} +\infty \quad \text{under } H_1 \quad \text{as } n \to +\infty$$

$$P(\text{rejecting } H_0) = P(T_W / T_S / T_R > \chi^2_{1-\alpha, p})$$

$$\to 1 \quad \text{as } n \to +\infty$$

proves the consistency of the three tests!

Now. consider local Pitman-type alternatives

$$H_{1:n} : \underset{\sim}{\theta}_n = \underset{\sim}{\theta}^* + \frac{1}{\sqrt{n}} \underset{\sim}{\Delta}$$

Under $H_{1:n}$:

$$T_W, T_S, T_R \xrightarrow{d} \chi^2_{p,(\lambda = \underset{\sim}{\Delta}^T I(\underset{\sim}{\theta}^*) \underset{\sim}{\Delta})}$$

Proof. (Intro: $\chi^2_p : \sum_{i=1}^{p} Z_i^2 \quad Z_i \overset{iid}{\sim} N(0,1)$

non center $\chi^2_{p,\lambda} : \sum_{i=1}^{p} X_i^2 , \quad X_i \overset{ind}{\sim} N(\mu_i, 1)$

$\lambda = \sum_{i=1}^{p} \mu_i^2$)

under $H_{1:n}$. $\quad \underset{\sim}{\theta}_0 = \underset{\sim}{\theta}^* + \frac{1}{\sqrt{n}} \underset{\sim}{\Delta}$

$$\sqrt{n}(\hat{\theta} - \theta_0) = \sqrt{n}(\hat{\theta} - \underset{\sim}{\theta}^*) - \underset{\sim}{\Delta}$$

$$\underbrace{\sqrt{n}(\hat{\theta} - \theta^*)}_{\text{Test statistic}} = \sqrt{n}(\hat{\theta} - \theta_0) + \underset{\sim}{\Delta} \rightarrow N(\underset{\sim}{\Delta}, I^{-1}(\theta_0))$$

$$T_W = \sqrt{n}(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}^*)^T I(\hat{\theta}) \sqrt{n}(\hat{\theta} - \theta^*)$$

$$\downarrow d \qquad \downarrow P \qquad \downarrow$$

$$\underline{N(\underset{\sim}{\Delta}, I^{-1}(\theta_0))\, I(\theta_0)\, N(\underset{\sim}{\Delta}, I^{-1}(\theta_0))} = \chi^2_{p,\lambda = \underset{\sim}{\Delta}^T I(\theta_0) \underset{\sim}{\Delta}}$$

$$\text{Power}(\underset{\sim}{\Delta}) = P(\underbrace{\chi^2_{p,\lambda = \underset{\sim}{\Delta}^T I(\theta_0) \underset{\sim}{\Delta}}}_{\text{no } n} > \underline{\chi^2_{p,1-\alpha}})$$

is a fixed probability

Same notation. $H_0: h(\underset{\sim}{\theta}) = \underset{\sim}{0}$ where $h: \mathbb{R}^p \to \mathbb{R}^q$.

Example $\underset{\sim}{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$. $H_0: \theta_1 - \theta_2 = 0$

$h(\underset{\sim}{\theta}) = h(\theta_1, \theta_2, \theta_3) = \theta_1 - \theta_2$

$h: R^3 \to R$

assume h is a continuous differentiable function

Define: $H(\underset{\sim}{\theta}) = \dfrac{\partial h(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}}$ gradient of h.

If $h: \mathbb{R}^p \to \mathbb{R}^q$. Then H is a $p \times q$ matrix.

Assume all pq partial derivative in $H(\underset{\sim}{\theta})$ are continuous and the rank of $H(\underset{\sim}{\theta})$ is q

$\{(\theta_1, \theta_2, \theta_3)\ \theta_1^2 = \theta_2, \theta_2^2 = \theta_3\}$

Example. $\underset{\sim}{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$ $H_0: \begin{matrix} \theta_1^2 - \theta_2 = 0 \\ \theta_2^2 - \theta_3 = 0 \end{matrix}$ vs. H_0^c

$$h(\underset{\sim}{\theta}) = \begin{pmatrix} h_1(\underset{\sim}{\theta}) \\ h_2(\underset{\sim}{\theta}) \end{pmatrix} = \begin{pmatrix} \theta_1^2 - \theta_2 \\ \theta_2^2 - \theta_3 \end{pmatrix}$$ $: \mathbb{R}^3 \to \mathbb{R}^2$

$p = 3.\quad q = 2.$

$$H(\underset{\sim}{\theta}) = \frac{\partial h(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = \begin{pmatrix} \frac{\partial h_1(\underset{\sim}{\theta})}{\partial \theta_1} & \frac{\partial h_2(\underset{\sim}{\theta})}{\partial \theta_1} \\ \frac{\partial h_1(\underset{\sim}{\theta})}{\partial \theta_2} & \frac{\partial h_2(\underset{\sim}{\theta})}{\partial \theta_2} \\ \frac{\partial h_1(\underset{\sim}{\theta})}{\partial \theta_3} & \frac{\partial h(\underset{\sim}{\theta})}{\partial \theta_3} \end{pmatrix} = \begin{pmatrix} 2\theta_1 & 0 \\ -1 & 2\theta_2 \\ 0 & -1 \end{pmatrix}$$

Three approaches.
1. Wald
2. Score
3. Likelihood ratio

Wald test: $T_W = n\, h(\hat{\theta}) [H(\hat{\theta}) I(\hat{\theta})^{-1} H(\hat{\theta})]^{-1} h(\hat{\theta}) \xrightarrow[H_0]{d} \chi_q^2$

Score test: $T_S = S(\tilde{\theta})^T [n I(\tilde{\theta})]^{-1} S(\tilde{\theta}) \xrightarrow[H_0]{d} \chi_q^2$

where $\tilde{\theta}$ is the restricted MLE under H_0

$\tilde{\theta} = \arg\sup_{\underset{\sim}{\theta} \in H_0} \log L(\underset{\sim}{\theta} | \underset{\sim}{X})$

Likelihood ratio: $T_{LR} = -2\log\left[\dfrac{\sup_{\theta\in H_0} L(\underset{\sim}{\theta}|\underset{\sim}{Y})}{\sup_{\theta\in\Omega} L(\underset{\sim}{\theta}|\underset{\sim}{Y})}\right] \xrightarrow[H_0]{d} \chi^2_q$

RR: $\begin{matrix} T_W \\ T_S \\ T_{LR} \end{matrix} > \chi^2_{q,1-\alpha}$

show $T_W \xrightarrow[H_0]{d} \chi^2_q$

Motivation. under H_0. $h(\underset{\sim}{\theta})=0$ is true meaning $h(\underset{\sim}{\theta}_0)=0$

we have: (a) $\sqrt{n}(\hat{\underset{\sim}{\theta}}-\underset{\sim}{\theta}_0) \xrightarrow{d} N_p(0, I(\theta_0)^{-1})$

(b) $\sqrt{n}(h(\hat{\underset{\sim}{\theta}})-h(\underset{\sim}{\theta}_0)) \xrightarrow{d} N_q(0, H(\theta_0)^T I(\theta_0)^{-1} H(\theta_0))$

(c) $I(\hat{\underset{\sim}{\theta}}) \xrightarrow{P} I(\theta_0)$, $H(\hat{\underset{\sim}{\theta}}) \xrightarrow{P} H(\theta_0)$

Proof. under H_0. $h(\underset{\sim}{\theta}_0)=0$

$$T_W = n\, h(\hat\theta)^T [H(\hat\theta) I(\hat\theta)^{-1} H(\hat\theta)]^{-1} h(\hat\theta)$$

$$= \underbrace{\sqrt{n}(h(\hat\theta)-h(\underset{\sim}{\theta}_0))^T}_{\downarrow d} \underbrace{[H(\hat\theta) I(\hat\theta)^{-1} H(\hat\theta)]^{-1}}_{\downarrow} \underbrace{\sqrt{n}(h(\hat\theta)-h(\theta_0))}_{\downarrow d}$$

$$N_q(0, H(\theta_0)^T I(\theta_0)^{-1} H(\theta_0))^T [H(\theta_0) I(\theta_0)^{-1} H(\theta_0)]^{-1} N_q(0, H(\theta_0)^T I(\theta_0)^{-1} H(\theta_0))$$

$$= \chi^2_q$$

Example Test equality of 2 exponential means

Suppose $\underset{\sim}{Y}_1, \cdots, \underset{\sim}{Y}_n$ iid from

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}|\underset{\sim}{\theta}) = \begin{cases} \frac{1}{\theta_1}\frac{1}{\theta_2}\exp\left\{-\frac{y_1}{\theta_1}-\frac{y_2}{\theta_2}\right\}, & y_1, y_2 > 0 \\ 0, & \text{o.w.} \end{cases}$$

where $\underset{\sim}{Y}_i = \begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix}$ $\underset{\sim}{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$

$$H_0: \frac{\theta_1}{\theta_2} = 1 \quad \text{vs} \quad H_1: \frac{\theta_1}{\theta_2} \neq 1$$

i.e. $h(\underset{\sim}{\theta}) = h(\theta_1, \theta_2) = \frac{\theta_1}{\theta_2} - 1$

First find $I(\theta)$

$$\log f_{\underset{\sim}{Y}}(\underset{\sim}{y} \mid \underset{\sim}{\theta}) = -\log\theta_1 - \log\theta_2 - \frac{y_1}{\theta_1} - \frac{y_2}{\theta_2}$$

$$\frac{\partial \log f_Y(\underset{\sim}{y} \mid \underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = \begin{pmatrix} -\frac{1}{\theta_1} + \frac{y_1}{\theta_1^2} \\ -\frac{1}{\theta_2} + \frac{y_2}{\theta_2^2} \end{pmatrix}$$

$$\frac{\partial^2 \log f_Y(\underset{\sim}{y} \mid \underset{\sim}{\theta})}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} = \begin{pmatrix} \frac{1}{\theta_1^2} & \\ & \frac{1}{\theta_2^2} \end{pmatrix}$$

$$I(\theta) = E\left[- \frac{\partial^2 \log f_Y(\underset{\sim}{Y} \mid \underset{\sim}{\theta})}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T} \right] = \begin{pmatrix} \frac{1}{\theta_1^2} & \\ & \frac{1}{\theta_2^2} \end{pmatrix}$$

$\uparrow$ est

$$\hat{I}(\theta) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f_Y(\underset{\sim}{Y}_i \mid \underset{\sim}{\theta})}{\partial \underset{\sim}{\theta} \partial \underset{\sim}{\theta}^T}$$

MLE: $\hat{\theta} = \begin{pmatrix} \bar{Y}_{1n} \\ \bar{Y}_{2n} \end{pmatrix}$ where $\bar{Y}_{jn} = \frac{1}{n} \sum_{i=1}^{n} Y_{ji}$ for $j = 1, 2$.

Now T_W, $h(\underset{\sim}{\theta}) = \frac{\theta_1}{\theta_2} - 1$ $\quad H(\underset{\sim}{\theta}) = \begin{pmatrix} \frac{1}{\theta_2} \\ -\frac{\theta_1}{\theta_2^2} \end{pmatrix}$

$$T_W = n\, h(\hat{\theta})^T \left[H(\hat{\theta})^T I(\hat{\theta})^{-1} H(\hat{\theta}) \right]^{-1} h(\hat{\theta})$$

$$= n \left(\frac{\bar{Y}_{1n}}{\bar{Y}_{2n}} - 1 \right)^2 \cdot \left[\left(\frac{1}{\bar{Y}_{2n}}, -\frac{\bar{Y}_{1n}}{\bar{Y}_{2n}^2} \right) \cdot \begin{pmatrix} \bar{Y}_{1n}^2 & \\ & \bar{Y}_{2n}^2 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\bar{Y}_{2n}} \\ -\frac{\bar{Y}_{1n}}{\bar{Y}_{2n}^2} \end{pmatrix} \right]^{-1}$$

$$= \frac{n}{2} \cdot \frac{(\bar{Y}_{1n} - \bar{Y}_{2n})^2}{\bar{Y}_{1n}^2} \xrightarrow[H_0]{d} \chi_1^2$$

Score test: $\tilde{\theta} = \underset{H_0: \theta_1 = \theta_2}{\arg\sup} \; L(\underset{\sim}{\theta} | \underset{\sim}{Y})$

Let $\theta = \theta_1 = \theta_2$ (under H_0)

$$L(\underset{\sim}{\theta} | \underset{\sim}{Y}) \overset{H_0}{=} L(\theta | \underset{\sim}{Y}) = \prod_{i=1}^{n} \frac{1}{\theta^2} \exp\left\{ - \frac{Y_{i1} + Y_{i2}}{\theta} \right\}$$

$$\Rightarrow \tilde{\theta} = \begin{pmatrix} \frac{\bar{Y}_{1n} + \bar{Y}_{2n}}{2} \\ \frac{\bar{Y}_{1n} + \bar{Y}_{2n}}{2} \end{pmatrix} = \begin{pmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{pmatrix}$$

$$S(\hat{\underset{\sim}{\theta}} | \underset{\sim}{Y}) = \begin{bmatrix} -\frac{n}{\theta_1} + \frac{n \bar{Y}_{1n}}{\theta_1^2} \\ -\frac{n}{\theta_2} + \frac{n \bar{Y}_{2n}}{\theta_2^2} \end{bmatrix}$$

$$\Rightarrow T_S = S(\tilde{\underset{\sim}{\theta}} | \underset{\sim}{Y})^T \cdot \left[n \, I(\tilde{\underset{\sim}{\theta}}) \right]^{-1} S(\tilde{\underset{\sim}{\theta}} | \underset{\sim}{Y})$$

$$= n \left(\frac{\bar{Y}_{1n}^2 + \bar{Y}_{2n}^2}{\left(\frac{\bar{Y}_{1n} + \bar{Y}_{2n}}{2} \right)^2} - 2 \right) \xrightarrow[H_0]{d} \chi_1^2$$

LR: $T_{LR} = -2 \left[\log L(\tilde{\theta}) - \log L(\hat{\theta}) \right]$

$$\log L(\theta) = \sum_{i=1}^{n} \left\{ -\log \theta_1 - \log \theta_2 - \frac{Y_{i1}}{\theta_1} - \frac{Y_{i2}}{\theta_2} \right\}$$

$$\log L(\tilde{\theta}) = -2n \log \frac{\bar{Y}_{1n} + \bar{Y}_{2n}}{2} - \frac{n(\bar{Y}_{1n} + \bar{Y}_{2n})}{\left(\frac{\bar{Y}_{1n} + \bar{Y}_{2n}}{2} \right)}$$

$$= -2n \log \frac{\bar{Y}_{1n} + \bar{Y}_{2n}}{2} - 2n$$

$$\log L(\hat{\theta}) = -n \log \bar{Y}_{1n} - n \log \bar{Y}_{2n} - \frac{n \bar{Y}_{1n}}{\bar{Y}_{1n}} - \frac{n \bar{Y}_{2n}}{\bar{Y}_{2n}}$$

$$= -n\left(\log \bar{Y}_{1n} + \log \bar{Y}_{2n}\right) - 2n$$

$$T_{LR} = -2\left[-2n\log\frac{\bar{Y}_{1n}+\bar{Y}_{2n}}{2} + n\log\bar{Y}_{1n} + n\log\bar{Y}_{2n}\right] \xrightarrow[H_0]{d} \chi^2_1$$

$$\underset{\sim}{\theta} = \begin{pmatrix}\theta_1\\ \vdots\\ \theta_p\end{pmatrix}$$

$H_0: \quad \theta_1 = \cdots = \theta_r = 0 \qquad\qquad T_{LR} \xrightarrow{d} \chi^2_r$

$H_0: \quad \theta_1 = \cdots = \theta_r \qquad\qquad T_{LR} \xrightarrow{d} \chi^2_{r-1}$

$$h(\underset{\sim}{\theta}) = \begin{pmatrix}\theta_1-\theta_2\\ \theta_2-\theta_3\\ \vdots\\ \theta_{r-1}-\theta_r\end{pmatrix} \quad : R^p \to R^{r-1}$$

Example. $\quad X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \qquad \underset{\sim}{\theta} = \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix} = \begin{pmatrix}\mu\\ \sigma\end{pmatrix}$

$H_0: \quad \mu = 0,\ \sigma = 1 \iff H_0: \underset{\sim}{\theta} = \begin{pmatrix}0\\ 1\end{pmatrix} = \underset{\sim}{\theta}^*$

$$L_n(\mu, \sigma) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_i (X_i-\mu)^2\right\}$$

$$T_{LR} = -2\log\left[\frac{L_n(0,1)}{L_n(\bar{X}_n, S_n)}\right] \quad \text{where } S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X}_n)^2$$

$$= -n\log S_n^2 + \sum_{i=1}^n X_i^2 - n \xrightarrow[H_0]{d} \chi^2_2$$

$$RR: \quad T_{LR} > \chi^2_{2,1-\alpha}$$

Example. Let $X_1, \cdots, X_c$ have a multinomial distribution based on n trials, each resulting in one of c outcomes (cells) with respective probabilities $p_1, \cdots, p_c$ where $p_i > 0$ for all i. $\sum_{i=1}^{c} p_i = 1$.

$$L_n(p_1, \cdots, p_c) = \begin{cases} \underbrace{\binom{n}{x_1 \cdots x_n}}_{\frac{n!}{x_1! \cdots x_n!}} \prod_{i=1}^{c} p_i^{x_i} & \text{if } \sum x_i = n \\ 0 & \text{o.w.} \end{cases}$$

Test: $H_0: p_1 = \cdots = p_c$

$\sum p_i = 1$

under $H_0 \Rightarrow p_i = \frac{1}{c}$

$$T_{LR} \xrightarrow[H_0]{d} \chi^2_{c-1}$$

$$T_{LR} = -2\log\left(\frac{\binom{n}{x_1 \cdots x_n} \cdot \prod_{i=1}^{n} \left(\frac{1}{c}\right)^{x_i}}{\binom{n}{x_1 \cdots x_n} \prod_{i=1}^{n} \left(\frac{x_i}{n}\right)^{x_i}}\right)$$

$$= -2\log \prod_{i=1}^{n} \left(\frac{n}{x_i c}\right)^{x_i}$$

$$= 2\sum_{i=1}^{c} x_i \log\left(\frac{c x_i}{n}\right) \xrightarrow[H_0]{d} \chi^2_{c-1}$$