Neyman-Pearson theory

- Understand the optimality of Neyman Pearson test
- Simple illustrations

Neyman-Pearson lemma

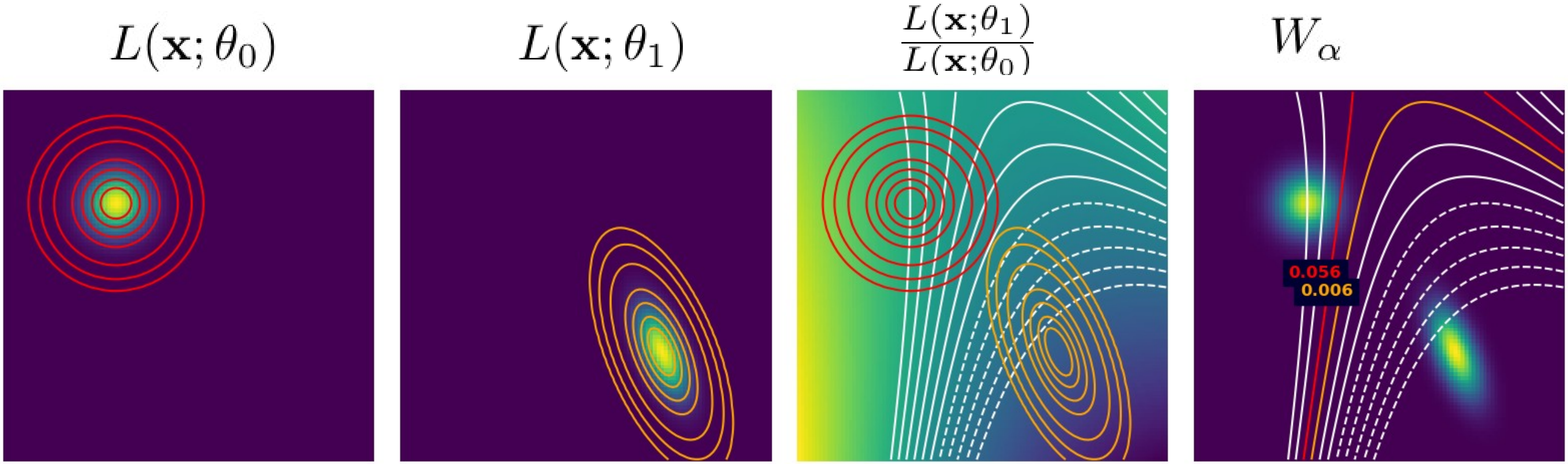
Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.

Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.

Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

Illustration

Neyman-Pearson Lemma: Maximize $1 - \beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.
Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$



Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.
Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

One can assume that $L(\mathbf{x};\theta_1) > 0$ and $L(\mathbf{x};\theta_0) > 0 \, \forall \, \mathbf{x} \in \mathbb{R}^n$

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0:\theta=\theta_0$, $H_1:\theta=\theta_1$.
Then the optimal critical region is defined by: $W_\alpha=\{\mathbf{x}\in\mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)}>k_\alpha\}$

One can assume that $L(\mathbf{x};\theta_1)>0$ and $L(\mathbf{x};\theta_0)>0\ \forall\,\mathbf{x}\in\mathbb{R}^n$

Note that $\mathbb{P}(W|H_0)=\int_W L(\mathbf{x};\theta_0)d\mathbf{x}$

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.
Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

One can assume that $L(\mathbf{x};\theta_1) > 0$ and $L(\mathbf{x};\theta_0) > 0 \; \forall\, \mathbf{x} \in \mathbb{R}^n$
Note that $\mathbb{P}(W|H_0) = \int_W L(\mathbf{x};\theta_0) d\mathbf{x}$

Unicity: $\forall\, \alpha \in]0,1[$, if $\exists\, k_\alpha$ s.t. $W_\alpha = \left\{\mathbf{x} \in \mathbb{R}^n : \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} > k_\alpha\right\}$
realizes $\mathbb{P}(W_\alpha|H_0) = \alpha$
Then $W = W_\alpha$ maximizes $1-\beta = \mathbb{P}(W|H_1)$

Neyman-Pearson lemma

Unicity: $\forall\, \alpha \in]0,1[$, if $\exists\, k_\alpha$ s.t. $W_\alpha = \left\{ \mathbf{x} \in \mathbb{R}^n : \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} > k_\alpha \right\}$
realizes $\mathbb{P}(W_\alpha | H_0) = \alpha$
Then $W = W_\alpha$ maximizes $1 - \beta = \mathbb{P}(W | H_1)$

Neyman-Pearson lemma

Unicity: $\forall\, \alpha \in]0,1[$, if $\exists\, k_\alpha$ s.t. $W_\alpha = \left\{ \mathbf{x} \in \mathbb{R}^n : \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} > k_\alpha \right\}$
realizes $\mathbb{P}(W_\alpha | H_0) = \alpha$
Then $W = W_\alpha$ maximizes $1 - \beta = \mathbb{P}(W | H_1)$

Indeed let W' $s.t.$ $\mathbb{P}(W' | H_0) = \alpha$

Neyman-Pearson lemma

Unicity: $\forall\ \alpha\ \in]0,1[,$ if $\exists\ k_\alpha$ s.t. $W_\alpha = \left\{ \mathbf{x} \in \mathbb{R}^n : \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} > k_\alpha \right\}$
realizes $\mathbb{P}(W_\alpha|H_0) = \alpha$
Then $W = W_\alpha$ maximizes $1 - \beta = \mathbb{P}(W|H_1)$

Indeed let W' $s.t.$ $\mathbb{P}(W'|H_0) = \alpha$

$$\mathbb{P}(W - W'|H_0) = \mathbb{P}(W' - W|H_0)$$

Neyman-Pearson lemma

Unicity: $\forall\, \alpha \;\in]0,1[$, if $\exists\, k_\alpha$ s.t. $W_\alpha = \left\{ \mathbf{x} \in \mathbb{R}^n : \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} > k_\alpha \right\}$
realizes $\mathbb{P}(W_\alpha | H_0) = \alpha$
Then $W = W_\alpha$ maximizes $1 - \beta = \mathbb{P}(W|H_1)$

Indeed let $W' \; s.t.\, \mathbb{P}(W'|H_0) = \alpha$

$\mathbb{P}(W - W'|H_0) = \mathbb{P}(W' - W|H_0)$

Then $\int_{W-W'} \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} L(\mathbf{x};\theta_0) d\mathbf{x} \geq \int_{W'-W} \cdots$

Neyman-Pearson lemma

Unicity: $\forall\, \alpha\ \in]0,1[$, if $\exists\, k_\alpha$ s.t. $W_\alpha = \left\{ \mathbf{x} \in \mathbb{R}^n : \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} > k_\alpha \right\}$
realizes $\mathbb{P}(W_\alpha | H_0) = \alpha$
Then $W = W_\alpha$ maximizes $1 - \beta = \mathbb{P}(W|H_1)$

Indeed let W' $s.t.$ $\mathbb{P}(W'|H_0) = \alpha$

$\mathbb{P}(W - W'|H_0) = \mathbb{P}(W' - W|H_0)$

Then $\int_{W-W'} \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} L(\mathbf{x};\theta_0) d\mathbf{x} \geq \int_{W'-W} \cdots$

ie $\int_W \frac{L(\mathbf{x};\theta_1)}{L(\mathbf{x};\theta_0)} L(\mathbf{x};\theta_0) d\mathbf{x} \geq \int_{W'} \ldots$ i.e. $\mathbb{P}(W|H_1) \geq \mathbb{P}(W'|H_1)$

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.

Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0: \theta = \theta_0$, $H_1: \theta = \theta_1$.

Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0:\theta=\theta_0$, $H_1:\theta=\theta_1$.
Then the optimal critical region is defined by: $W_\alpha=\{\mathbf{x}\in\mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)}>k_\alpha\}$

Existence: Let $W(k)=\{\mathbf{x}\in\mathbb{R}^n : L(\mathbf{x};\theta_1)>kL(\mathbf{x};\theta_0)\}$

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.
Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

Existence: Let $W(k) = \{\mathbf{x} \in \mathbb{R}^n : L(\mathbf{x};\theta_1) > kL(\mathbf{x};\theta_0)\}$
Then $\mathbb{P}(W(k)|H_0) \searrow k$ and is continuous

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.
Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

Existence: Let $W(k) = \{\mathbf{x} \in \mathbb{R}^n : L(\mathbf{x};\theta_1) > kL(\mathbf{x};\theta_0)\}$
Then $\mathbb{P}(W(k)|H_0) \searrow k$ and is continuous
$\mathbb{P}(W(0)|H_0) = 1 \; as \; L(\mathbf{x};\theta_1) > 0$

Neyman-Pearson lemma

Neyman-Pearson Lemma: Maximize $1-\beta$ for a given α.
Assume $\mathbf{x}$ has density $L(\mathbf{x};\theta)$, $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$.
Then the optimal critical region is defined by: $W_\alpha = \{\mathbf{x} \in \mathbb{R}^n s.t. \frac{L(x;\theta_1)}{L(x;\theta_0)} > k_\alpha\}$

Existence: Let $W(k) = \{\mathbf{x} \in \mathbb{R}^n : L(\mathbf{x};\theta_1) > kL(\mathbf{x};\theta_0)\}$
Then $\mathbb{P}(W(k)|H_0) \searrow k$ and is continuous
$\mathbb{P}(W(0)|H_0) = 1 \; as \; L(\mathbf{x};\theta_1) > 0$
$\lim_{k\to\infty} \mathbb{P}(W(k)|H_0) = 0$

Key properties of NP tests

$$\boxed{\text{proposition: } 1-\beta \geq \alpha, \text{ ie } \mathbb{P}(W|H_1) \geq \mathbb{P}(W|H_0)}$$

If $k_\alpha > 1$, this results from the definition of W

Key properties of NP tests

$\boxed{\text{proposition: } 1-\beta \geq \alpha, \text{ ie } \mathbb{P}(W|H_1) \geq \mathbb{P}(W|H_0)}$

If $k_\alpha > 1$, this results from the definition of W

If $k_\alpha \leq 1$, then we show that $\beta < 1-\alpha$, ie $\mathbb{P}(W^c|H_1) \leq \mathbb{P}(W^c|H_0)$

Key properties of NP tests

$\boxed{\text{proposition: } 1-\beta \geq \alpha, \text{ ie } \mathbb{P}(W|H_1) \geq \mathbb{P}(W|H_0)}$

If $k_\alpha > 1$, this results from the definition of W

If $k_\alpha \leq 1$, then we show that $\beta < 1-\alpha$, ie $\mathbb{P}(W^c|H_1) \leq \mathbb{P}(W^c|H_0)$

Indeed, by definition W^c is such that $\forall \mathbf{x} \in W^c, L(\mathbf{x};\theta_1) \leq k_\alpha L(\mathbf{x};\theta_0)$

Key properties of NP tests

$$\boxed{\text{proposition: } 1-\beta \geq \alpha, \text{ ie } \mathbb{P}(W|H_1) \geq \mathbb{P}(W|H_0)}$$

If $k_\alpha > 1$, this results from the definition of W

If $k_\alpha \leq 1$, then we show that $\beta < 1-\alpha$, ie $\mathbb{P}(W^c|H_1) \leq \mathbb{P}(W^c|H_0)$

Indeed, by definition W^c is such that $\forall \mathbf{x} \in W^c, L(\mathbf{x};\theta_1) \leq k_\alpha L(\mathbf{x};\theta_0)$

Then $\int_{W^c} L(\mathbf{x};\theta_1)d\mathbf{x} \leq \int_{W^c} L(\mathbf{x};\theta_0)d\mathbf{x}$

Key properties of NP tests

proposition: $1-\beta \geq \alpha$, ie $\mathbb{P}(W|H_1) \geq \mathbb{P}(W|H_0)$

If $k_\alpha > 1$, this results from the definition of W

If $k_\alpha \leq 1$, then we show that $\beta < 1-\alpha$, ie $\mathbb{P}(W^c|H_1) \leq \mathbb{P}(W^c|H_0)$

Indeed, by definition W^c is such that $\forall \mathbf{x} \in W^c, L(\mathbf{x};\theta_1) \leq k_\alpha L(\mathbf{x};\theta_0)$

Then $\int_{W^c} L(\mathbf{x};\theta_1)d\mathbf{x} \leq \int_{W^c} L(\mathbf{x};\theta_0)d\mathbf{x}$

Other basic properties: $\mathbb{P}(W(k)) \searrow k$, $k_\alpha \searrow \alpha$, $1-\beta \nearrow \alpha$

Exhaustive statistics

Exhaustive statistic: $L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})$

Exhaustive statistics

$$\boxed{\text{Exhaustive statistic: } L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})}$$

Then the critical regions becomes $\frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha$

Exhaustive statistics

$\boxed{\text{Exhaustive statistic: } L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})}$

Then the critical regions becomes $\frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha$

typical cases: Student, Fisher

Exhaustive statistics

$\boxed{\text{Exhaustive statistic: } L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})}$

Then the critical regions becomes $\frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha$

typical cases: Student, Fisher

$\boxed{\text{Example: Gaussian law with known } \sigma,\ H_0 : \mathcal{N}(m_0, \sigma^2),\ H_1 : \mathcal{N}(m_1, \sigma^2)}$

Exhaustive statistics

$\boxed{\text{Exhaustive statistic: } L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})}$

Then the critical regions becomes $\frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha$

typical cases: Student, Fisher

$\boxed{\text{Example: Gaussian law with known } \sigma,\ H_0 : \mathcal{N}(m_0, \sigma^2),\ H_1 : \mathcal{N}(m_1, \sigma^2)}$

$$g(\bar{x}, m) = \frac{1}{\sigma\sqrt{\frac{2\pi}{n}}} \exp\left(-\frac{1}{2}\left(\frac{\bar{x}-m}{\sigma\sqrt{n}}\right)^2\right)$$

Exhaustive statistics

Exhaustive statistic: $L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})$

Then the critical regions becomes $\frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha$

typical cases: Student, Fisher

Example: Gaussian law with known σ, $H_0 : \mathcal{N}(m_0, \sigma^2)$, $H_1 : \mathcal{N}(m_1, \sigma^2)$

$$g(\bar{x}, m) = \frac{1}{\sigma\sqrt{\frac{2\pi}{n}}} \exp\left(-\frac{1}{2}\left(\frac{\bar{x}-m}{\sigma\sqrt{n}}\right)^2\right)$$

$$\log\frac{g(\bar{x},m_1)}{g(\bar{x},m_0)} = -\frac{n}{2\sigma^2}\left((\bar{x}-m_1)^2 - (\bar{x}-m_0)^2\right)$$

Exhaustive statistics

Exhaustive statistic: $L(\mathbf{x};\theta) = g(t,\theta)h(\mathbf{x})$

Then the critical regions becomes $\frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha$

typical cases: Student, Fisher

Example: Gaussian law with known σ, $H_0 : \mathcal{N}(m_0, \sigma^2)$, $H_1 : \mathcal{N}(m_1, \sigma^2)$

$$g(\bar{x}, m) = \frac{1}{\sigma\sqrt{\frac{2\pi}{n}}} \exp\left(-\frac{1}{2}\left(\frac{\bar{x}-m}{\sigma\sqrt{n}}\right)^2\right)$$

$$\log\frac{g(\bar{x},m_1)}{g(\bar{x},m_0)} = -\frac{n}{2\sigma^2}\left((\bar{x}-m_1)^2 - (\bar{x}-m_0)^2\right)$$

$\ldots$ ratio $>$ cst iff $(m_1 - m_0)(\bar{x} - \frac{m_0+m1}{2}) > \kappa_\alpha$, ie $(m_1 > m_0)$ $\bar{x} > x_\alpha$

solutions are half lines

Composite hypothesis

Composite Hypothesis. e.g. $H_0 : \theta = \theta_0; H_1 : \theta > \theta_0$ or $\theta \neq \theta_0$

Composite hypothesis

$\boxed{\textbf{Composite Hypothesis}\text{. e.g. } H_0 : \theta = \theta_0; H_1 : \theta > \theta_0 \text{ or } \theta \neq \theta_0}$

Power function $1 - \beta(\theta)$

Uniformly Most Powerful test if $\forall\theta, 1 - \beta(\theta)$ greater than for any other test.

Composite hypothesis

$\boxed{\textbf{Composite Hypothesis.}\ \text{e.g. } H_0 : \theta = \theta_0; H_1 : \theta > \theta_0 \text{ or } \theta \neq \theta_0}$

Power function $1 - \beta(\theta)$
Uniformly Most Powerful test if $\forall \theta, 1 - \beta(\theta)$ greater than for any other test.

E.g. in Gaussian case, exists for $H_1 : \theta > \theta_0$, not for $H_1 : \theta \neq \theta_0$

Composite hypothesis

Composite Hypothesis. e.g. $H_0 : \theta = \theta_0 ; H_1 : \theta > \theta_0$ or $\theta \neq \theta_0$

Power function $1 - \beta(\theta)$

Uniformly Most Powerful test if $\forall \theta, 1 - \beta(\theta)$ greater than for any other test.

E.g. in Gaussian case, exists for $H_1 : \theta > \theta_0$, not for $H_1 : \theta \neq \theta_0$

Maximal likelihood ratio test for $H_0 : \theta = \theta_0 ; H_1 : \theta \neq \theta_0$

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{\sup_\theta L(\mathbf{x};\theta)} = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})}$$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\to} \chi^2_p, \text{ where } p = dim(\theta)$$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\rightarrow} \chi^2_p, \text{ where } p = dim(\theta)$$

Proof (scalar case to simplify notations):

$$\log L(\mathbf{x};\theta_0) - \log L(\mathbf{x};\hat{\theta}) = (\theta_0 - \hat{\theta})\frac{\partial}{\partial\theta}\log L(\mathbf{x};\hat{\theta}) + \frac{1}{2}(\theta_0 - \hat{\theta})^2\frac{\partial^2}{\partial\theta^2}\log L(\mathbf{x};\theta^\star),$$

where $\theta^\star \in [\theta_0, \hat{\theta}]$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\rightarrow} \chi^2_p, \text{ where } p = dim(\theta)$$

Proof (scalar case to simplify notations):

$$\log L(\mathbf{x};\theta_0) - \log L(\mathbf{x};\hat{\theta}) = (\theta_0 - \hat{\theta})\frac{\partial}{\partial\theta}\log L(\mathbf{x};\hat{\theta}) + \frac{1}{2}(\theta_0 - \hat{\theta})^2\frac{\partial^2}{\partial\theta^2}\log L(\mathbf{x};\theta^\star),$$

where $\theta^\star \in [\theta_0, \hat{\theta}]$

but $\frac{\partial}{\partial\theta}\log L(\mathbf{x};\hat{\theta}) = 0$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\to} \chi^2_p, \text{ where } p = dim(\theta)$$

Proof (scalar case to simplify notations):

$$\log L(\mathbf{x};\theta_0) - \log L(\mathbf{x};\hat{\theta}) = (\theta_0 - \hat{\theta})\frac{\partial}{\partial\theta}\log L(\mathbf{x};\hat{\theta}) + \frac{1}{2}(\theta_0 - \hat{\theta})^2\frac{\partial^2}{\partial\theta^2}\log L(\mathbf{x};\theta^\star),$$

where $\theta^\star \in [\theta_0, \hat{\theta}]$

but $\frac{\partial}{\partial\theta}\log L(\mathbf{x};\hat{\theta}) = 0$

$$-2\log\lambda = -(\theta_0 - \hat{\theta})^2\frac{\partial^2}{\partial\theta^2}\log(\mathbf{x};\theta^\star)$$

Maximum Likelihood ratio test

$$\boxed{\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\to} \chi^2_p, \text{ where } p = dim(\theta)}$$

$$-2\log\lambda = -(\theta_0 - \hat{\theta})^2 \frac{\partial^2}{\partial\theta^2} \log(\mathbf{x};\theta^\star)$$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\rightarrow} \chi^2_p, \text{ where } p = dim(\theta)$$

$$-2\log\lambda = -(\theta_0 - \hat{\theta})^2 \frac{\partial^2}{\partial\theta^2}\log(\mathbf{x};\theta^\star)$$

$$\text{Under } H_0\ \theta^\star \underset{n\to\infty}{\rightarrow} \theta_0 \text{ thus } \frac{\partial^2}{\partial\theta^2}\log L(\mathbf{x};\theta^\star) \rightarrow nI(\theta_0), \text{ where } I(\theta) = \mathbb{E}\frac{\partial^2 \log L}{\partial\theta^2}$$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\to} \chi^2_p, \text{ where } p = dim(\theta)$$

$$-2\log\lambda = -(\theta_0 - \hat{\theta})^2 \frac{\partial^2}{\partial\theta^2}\log(\mathbf{x};\theta^\star)$$

Under H_0 $\theta^\star \underset{n\to\infty}{\to} \theta_0$ thus $\frac{\partial^2}{\partial\theta^2}\log L(\mathbf{x};\theta^\star) \to nI(\theta_0)$, where $I(\theta) = \mathbb{E}\frac{\partial^2 \log L}{\partial\theta^2}$

But $\frac{\theta_0 - \hat{\theta}}{\sqrt{1/I(\theta_0)}} \to \mathcal{N}(0,1)$

Maximum Likelihood ratio test

$$\boxed{\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\to} \chi^2_p, \text{ where } p = dim(\theta)}$$

$$-2\log\lambda = -(\theta_0 - \hat{\theta})^2 \frac{\partial^2}{\partial\theta^2} \log(\mathbf{x};\theta^\star)$$

$$\text{Under } H_0\ \theta^\star \underset{n\to\infty}{\to} \theta_0 \text{ thus } \frac{\partial^2}{\partial\theta^2} \log L(\mathbf{x};\theta^\star) \to nI(\theta_0), \text{ where } I(\theta) = \mathbb{E}\frac{\partial^2 \log L}{\partial\theta^2}$$

$$\text{But } \frac{\theta_0 - \hat{\theta}}{\sqrt{1/I(\theta_0)}} \to \mathcal{N}(0,1)$$

$$\text{Thus } -2\log\lambda \sim (\theta_0 - \hat{\theta})^2 I(\theta_0) \to \chi^2_1$$

Maximum Likelihood ratio test

$$\lambda = \frac{L(\mathbf{x};\theta_0)}{L(\mathbf{x};\hat{\theta})} \quad \text{Theorem: Under } H_0,\ -2\log\lambda \underset{n\to\infty}{\to} \chi^2_p, \text{ where } p = dim(\theta)$$

$$-2\log\lambda = -(\theta_0 - \hat{\theta})^2 \tfrac{\partial^2}{\partial\theta^2} \log(\mathbf{x};\theta^\star)$$

$$\text{Under } H_0\ \theta^\star \underset{n\to\infty}{\to} \theta_0 \text{ thus } \tfrac{\partial^2}{\partial\theta^2} \log L(\mathbf{x};\theta^\star) \to nI(\theta_0), \text{ where } I(\theta) = \mathbb{E}\frac{\partial^2 \log L}{\partial\theta^2}$$

$$\text{But } \frac{\theta_0 - \hat{\theta}}{\sqrt{1/I(\theta_0)}} \to \mathcal{N}(0,1)$$

$$\text{Thus } -2\log\lambda \sim (\theta_0 - \hat{\theta})^2 I(\theta_0) \to \chi^2_1$$

$$\boxed{\text{Note that when } n \to \infty, 1 - \beta \to 1}$$