Asymptotic approach for testing hypotheses
Numerical comparison.

The setting is the same as in 8x. 51 p. 125 Lehne.

Let we have observations $X_1, X_2, ..., X_n$, iid, according to

$v$-densities $D_0(x)$ or $D_1(x), D_0(x) \neq D_1(x)$. Hypotheses are

$H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$. Approach is MP, i.e.,

$\text{do}(\theta) = 2$.

(1)

What is the MP test? Suppose that $F(x) = \frac{P_1(x)}{P_0(x)}$

is distribution of the likelihood $2 \rightarrow 0$.

The last means that is continuous. The last means that

$P(f(x) = c | H_0) = P(f(x) = c | H_1) = 0$. Then in accordance

$P(f(x) = c | H_0) = P(f(x) = c | H_1) = 0$. Then in accordance

$P(f(x) = c | H_0) = P(f(x) = c | H_1) = 0$. Then in accordance

with our theory the MP test $C$ is

$C = \begin{cases} 0, & \text{if } \sum_{i=1}^{n} f(x_i) \leq C \text{ where } C \text{ is a constant,} \\ 1, & \text{otherwise.} \end{cases}$

The constant $C$ is (may be) a function of $n$

and is determined from equation

$\text{do}(\theta) = P_{\theta_0} \left( \sum_{i=1}^{n} f(x_i) > C \right) = \alpha$.

(2)

The idea is that we can determine $C$ with

no knowledge (i) about $P_0(x)$ and $P_1(x)$ (more precisely

we have to know only some statistical

characteristics).

From (1) it is naturally to study instead of

$\sum_{i=1}^{n} f(x_i)$ the $\ln$ of $\prod_{i=1}^{n} f(x_i)$, i.e.,

$\text{do}(\theta) = P_{\theta_0} \left( \sum_{i=1}^{n} \ln f(x_i) > C \left| P_0 \right) = \alpha \right.$

(3)

Now we will consider the recall that $X_i$ are iid,

therefore $\ln\left( f(x_i) \right) = \eta_i$ and $\eta_i$ are iid as well.
So we can rewrite Equation (3) as

\[(5) \quad r_0(B) = P_0 \left( \frac{1}{n} \sum_{i=1}^{n} \phi(x_i) \right) = \frac{1}{n} \ln c \quad P_0 \{ x \in \mathbb{R}^n : \frac{1}{n} \sum_{i=1}^{n} \phi(x_i) > \frac{1}{n} \}
\]

This we converted our problem into the problem of studying
\[\frac{1}{n} \sum_{i=1}^{n} \phi(x_i), \text{ where } \phi_i \text{ are i.i.d and}
\]

\[(6) \quad \phi_i = \ln c \phi(x_i) = \ln \frac{P_0(x_i)}{P_0(x_i)}.\]

Since we would like to use CLT (central limit theorem). Therefore we have to consider the moments of \(\phi_i\).

**Lemma:** The distance of Kullback-Leibler divergence between \(P_0 \neq P_1\) is defined as

\[(7) \quad \chi_2(P_0, P_1) = \int \phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx
\]

and always

\[(8) \quad \chi_2(P_0, P_1) > 0 \quad (\text{for } P_0 \neq P_1, \text{ a.e.}) \quad \iff P_0 \neq P_1\]

**Proof:**

a) Jensen's inequality \(\Rightarrow -\ln z \geq \phi(x)\)

convex \(\Rightarrow -\ln z \geq \phi(x)\)

\[\int \phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx = 0 \quad \iff P_0(x) = P_1(x)\]

\[\ln(2) \geq \int \phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx \geq 1 - \ln 2, \quad \text{i.e., } P_0 = P_1 \iff \int \phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx = 0\]

So, if \(P_0(x) \neq P_1(x)\) a.e. then

\[(9) \quad \chi_2(P_0, P_1) = \int \phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx = -\phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx = \chi_2(P_1, P_0) = -\chi_2(P_0, P_1) = -\alpha < 0,
\]

\[(10) \quad \chi_2(P_0, P_1) = \int \phi(x) \ln \frac{P_0(x)}{P_1(x)} \, dx = \chi_2(P_1, P_0) = -\chi_2(P_0, P_1) = \alpha > 0
\]

(\text{true if } P_0 \neq P_1 \text{ in a general case})
Well, now we can use it from CLT

\[ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\eta_i + a) \xrightarrow{D} N(0, \text{Var}(\eta_1 P_0)) \]

\[ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\eta_i - b) \xrightarrow{D} N(0, \text{Var}(\eta_1 P_1)) \]

that is, we have a splitting distributions of \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \), because \( a > 0 \) and \( b > 0 \)!

Now, let us return to (5). We have to find \( c = c_n \).

Assume that \( \sigma_i^2 = \text{Var}(\eta_1 P_i) \). Then

\[ \chi^2 \left( \frac{n}{\sigma_i^2} \right) = \frac{n}{\sigma_i^2} \sum_{i=1}^{n} \eta_i^2 > \frac{n}{\sigma_i^2} \sigma_i^2 \chi^2_0 = \chi^2_0 \left( \frac{\sum_{i=1}^{n} (\eta_i + a)}{0.5 \sqrt{n}} \right) \]

But \( \frac{1}{0.5 \sqrt{n}} \sum_{i=1}^{n} (\eta_i + a) \xrightarrow{D} \chi^2_0 \sim N(0, 1) \). So let

\( c = c_n \) is any sequence of numbers such that

\[ \frac{\chi^2 \left( c_n \right) + a n}{\sigma_0 \sqrt{n}} \xrightarrow{n \to \infty} Z_\alpha \]

where \( Z_\alpha \) is a \( (1-\alpha) \)-percentile

Then

\[ \chi^2 \left( c_n \right) \xrightarrow{d} \chi^2_0 \text{ as } n \to \infty \]

The Fisher's Test of Variance: Test

\[ S = \left\{ \begin{array}{ll}
0, & \text{if } \prod_{i=1}^{n} \frac{P_i(x_i)}{P_0(x_i)} \leq c_n \\
1, & \text{otherwise}
\end{array} \right. \]

where \( c_n \) satisfies (14), is asymptotically the best \( \chi^2 \) test, i.e.

\[ \chi^2 \left( c_n \right) \xrightarrow{d} \chi^2_0 \text{ as } n \to \infty \]

Conclusion: We do not know properties of \( P_0(x) \) and \( P_0(x) \), except of: \( \alpha = \text{Var}(P_0, P_1) \) and \( \sigma_0^2 = \text{Var}(\eta_1 P_i) \).

And this is enough for looking by ML testing! (i.e., for finding \( c_n \)...)
(16) \[ \ln(C_n) = -a_n + \frac{a}{2} \ln(n) \]

Evidently \( C_n \to 0 \) as \( n \to \infty \).

What can we say about the second type error for alternative hypothesis \( H_i: p_i(x) \)?

It is a more difficult problem. Write,

(17) \[ \lambda_2(D) = \max \mathbb{P}_1 \left( \sum_{i=1}^{n} \xi_i < b_n \right) = \mathbb{P}_1 \left( \frac{1}{\delta_n \sqrt{n}} \sum_{i=1}^{n} (\xi_i - b) < (-a - b) \ln \delta_n^{-1} + \frac{a}{2} \ln \delta_n^{-1} \right). \]

Now \( \frac{1}{\delta_n \sqrt{n}} \sum_{i=1}^{n} (\xi_i - b) \xrightarrow{P} \delta_0 \sim N(0, 1) \),

i.e., we have to find \((8)\) is more common expression than \( \mathbb{P}_1 \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i < x \right) \)

(18) \[ \mathbb{P}_1 \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i > x \right) = ? \] as \( x \to \infty, n \to \infty \)

This is the problem of a large deviation!

The central point is that we have to assume that

(19) \[ \psi(\lambda) = \mathbb{E}_1 \left[ e^{z \xi_i} \right] \leq \infty \quad \text{for } |z| < A \text{ for some } A. \]

In our case \( \psi(z) = \frac{1}{\delta_n} \mathbb{E}_1 (\xi_i - b), \quad \mathbb{E}_1 = \mathbb{E}_1 \frac{p_i(x)}{p_0(x)} \).

Then

(20) \[ \psi(\lambda) = \mathbb{E}_1 \left[ e^{\frac{2}{\delta_n} (\xi_i - b)^2} \right] = e^{\frac{-2}{\delta_n} b} \mathbb{E}_1 \left[ e^{\frac{2}{\delta_n} \xi_i} \right] \]

and

(21) \[ \mathbb{E}_1 \left[ e^{\frac{2}{\delta_n} \xi_i} \right] = \int p_i(x) e^{\frac{2}{\delta_n} \xi_i} \frac{\delta_n}{p_i(x)} dx = \int p_i(x) \left( \frac{\delta_n}{p_i(x)} \right) e^{\frac{2}{\delta_n} \xi_i} dx. \]
Now I omit some steps, but the cases when in (21) the right-hand side is undefined + 1/2

\[ \sigma^2_{\chi^2}(c_n) = \frac{\chi^2}{\theta, \nu} \sum_{i=1}^{\nu} (\bar{x}_i - \bar{x})^2 \frac{\theta - 6}{\theta - 6 + \frac{2}{\theta} \sigma^2} \]

\[ \sim \frac{1}{\theta, \nu / 2 \Gamma(\nu)} \exp \{- \frac{1}{\theta} \beta_p (\theta, \nu) + \frac{2}{\nu} \sigma^2 / \nu \}
\]

where

\[ \beta_p (\theta, \nu) = \int \theta \ln \frac{\theta}{\theta} \frac{0 (\theta)}{0 (\theta)} d\theta. \]

**Example.** Let \( x \in \Gamma, \nu, \) i.e.

\( x \) has a distribution density

\[ p(x) = \theta, e^{-\theta x}, x > 0 \]

let we test hypotheses \( H_0 : \theta = 1, \) \( H_1 : \theta = 0.5 \)

Let \( \alpha = 0.025 \) = \( 1 - F(2) \) and \( \delta = 2 \)

\( \eta_i = \ln \frac{\theta_i}{\theta} = \ln (\theta_2) - (\theta - 1) x \)

From (16) we obtain that \( \chi^2 = 2 \)

\[ \chi^2_{\theta, \nu} = \frac{\theta}{\theta} \int \theta \ln \theta \frac{0 (\theta)}{0 (\theta)} d\theta = \frac{\theta}{\theta} \]

\[ \sigma^2_{\chi^2} = E_{\theta, \nu} \left[ \left( \frac{\theta_i}{\theta_i} + \bar{x} \right)^2 \right] = (\theta - 1)^2 \theta \ln \theta, \frac{\theta^2}{\theta} \]

\[ \sigma_{\theta}^2 = \theta \left( \frac{\theta}{\theta} + \frac{2}{\theta} \right)^2 = (\theta - 1)^2 \theta \ln \theta \]
and therefore substituting this into the right-hand side of (24) we get

\[ L_n(n) = n \left[ \ln \frac{\theta_1}{\theta_2} + (\theta_2 - 1) \right] + 2 (\theta_2 - 1)^2 \frac{1}{n} \]

This is our LMP test with \( \Delta = 0.023 \)

But (25) gives us an approximation.

Look at numerical calculations under our hypotheses:

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>0.5</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta )</td>
<td>0.028</td>
<td>0.033</td>
<td>0.029</td>
</tr>
<tr>
<td>( L_n(n) )</td>
<td>(5)10^{-7}</td>
<td>(5)10^{-7}</td>
<td>(5)10^{-7}</td>
</tr>
<tr>
<td>( L_2(n) )</td>
<td>0.35</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Values of \( L_2(n) = P\left( S > \theta_1 \mid \theta_2 \right) \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>30</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_2(n) )</td>
<td>0.031</td>
<td>0.028</td>
<td>0.026</td>
<td>0.024</td>
</tr>
<tr>
<td>( L_2(n) )</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
</tr>
<tr>
<td>( L_2(n) )</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Now the second type error (upper) (against null \( \theta = 1 \)) lower = using (22)

<table>
<thead>
<tr>
<th>( n )</th>
<th>30</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_2 )</td>
<td>0.5</td>
<td>0.8</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>( L_2(n) )</td>
<td>0.028</td>
<td>0.033</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>( L_2(n) )</td>
<td>(5)10^{-7}</td>
<td>(5)10^{-7}</td>
<td>(5)10^{-7}</td>
<td></td>
</tr>
<tr>
<td>( L_2(n) )</td>
<td>0.35</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

we have a different situation!

Conclusion: \( \theta_1 = 1 \) and \( \theta_2 = 0.9 \) are close hypotheses. Even for \( n = 1000 \) they are not indistinguishable. This is a special topic in modern hypothesis testing.