The Likelihood Ratio (LR) Test

Consider the Log-Likelihood function \( L\{y_1, \ldots, y_n; \theta\} \), where \( \theta = [\theta_1, \theta_2, \ldots, \theta_p]' \). Test the null hypothesis of \( m \) independent restrictions:

\[
H_0: c(\theta) - q = 0,
\]

where \( q_{[m \times 1]} \) is a vector of constants.

- if the restrictions are linear:
  \( c_{[m \times p]} \) is a fixed matrix and:
  \[
  H_0: c \theta - q = 0
  \]

- if the restrictions are non-linear:
  \( c_{[m \times 1]} \) is a vector with elements \( c_i = c_i(\theta), \)
  \[
  i = 1, 2, \ldots, m.
  \]
Consider:

(i) the unrestricted MLE, $\hat{\theta}$, i.e.:
$$\max_{\theta} L\{y_1, \ldots, y_n; \theta\} \implies \theta_{\text{max},L} = \hat{\theta}$$

(ii) the restricted MLE, $\hat{\theta}_0$, i.e.:
$$\max_{\theta} L\{y_1, \ldots, y_n; \theta\} \quad \text{subject to} \quad c(\theta) - q = 0 \implies \theta_{\text{max},L} = \hat{\theta}_0$$

Then, the Likelihood Ratio Test statistic, LR, is:
$$LR = 2\left(L\{y_1, \ldots, y_n; \hat{\theta}\} - L\{y_1, \ldots, y_n; \hat{\theta}_0\}\right) = 2 \ln \frac{l\{y_1, \ldots, y_n; \hat{\theta}\}}{l\{y_1, \ldots, y_n; \theta_0\}}$$

- Under the null hypothesis, the asymptotic distribution of the LR is $\chi^2_m$ (see the next two pages for more detail).
- Reject the null if LR is too large.
Consider the Taylor series expansion of the restricted Log-likelihood function, \( L\{y_1, ..., y_n; \hat{\theta}_0\} \), in a neighborhood of the unrestricted estimate \( \hat{\theta} \):

\[
L\{y_1, ..., y_n; \hat{\theta}_0\} = L\{y_1, ..., y_n; \hat{\theta}\} + \frac{(\hat{\theta}_0 - \hat{\theta})'}{1!} \left\{ \frac{\partial L\{y_1, ..., y_n; \theta\}}{\partial \theta} \right\} |_{\theta = \hat{\theta}} \\
+ \frac{1}{2!}(\hat{\theta}_0 - \hat{\theta})' \left\{ \frac{\partial^2 L\{y_1, ..., y_n; \theta\}}{\partial \theta \partial \theta} \right\} |_{\theta = \hat{\theta}} (\hat{\theta}_0 - \hat{\theta}) + R_2,
\]

\[
\downarrow
\]

\[
L\{y_1, ..., y_n; \hat{\theta}_0\} = L\{y_1, ..., y_n; \hat{\theta}\} + \frac{(\hat{\theta}_0 - \hat{\theta})'}{1} s(y_1, ..., y_n; \hat{\theta}) \\
+ \frac{1}{2}(\hat{\theta}_0 - \hat{\theta})' \left\{ \frac{\partial s\{y_1, ..., y_n; \theta\}}{\partial \theta} \right\} |_{\theta = \hat{\theta}} (\hat{\theta}_0 - \hat{\theta}) + R_2,
\]

where \( R_2 \) is a remainder of order 3.

However, \( s(y_1, ..., y_n; \hat{\theta}) = 0 \)

\[
\implies L\{y_1, ..., y_n; \hat{\theta}_0\} \approx \approx L\{y_1, ..., y_n; \hat{\theta}\} + \frac{1}{2}(\hat{\theta}_0 - \hat{\theta})' \left\{ \frac{\partial s\{y_1, ..., y_n; \theta\}}{\partial \theta} \right\} |_{\theta = \hat{\theta}} (\hat{\theta}_0 - \hat{\theta})
Thus
\[
LR = 2(L\{y_1, ..., y_n; \hat{\theta}\} - L\{y_1, ..., y_n; \hat{\theta}_0\}) =
\]
\[
(\hat{\theta}_0 - \hat{\theta})' \left\{ -\frac{\partial s\{y_1, ..., y_n; \theta\}}{\partial \theta} \right\} \bigg|_{\theta = \hat{\theta}} (\hat{\theta}_0 - \hat{\theta})
\]
\[
\uparrow
\]
\[
LR = \sqrt{n}(\hat{\theta}_0 - \hat{\theta})' \left\{ -\frac{1}{n} \frac{\partial s\{y_1, ..., y_n; \theta\}}{\partial \theta} \right\} \bigg|_{\theta = \hat{\theta}} \sqrt{n}(\hat{\theta}_0 - \hat{\theta})
\]

It follows from Assumption 1 and Assumption 2 that:

(see your last week notes or Lecture notes p.15-16):

\[
p \lim \frac{1}{n} \left\{ -\frac{\partial s}{\partial \theta} \right\} \bigg|_{\theta = \hat{\theta}} = \lim_{n \to \infty} \frac{1}{n} I(\theta)
\]
\[
\uparrow
\]
\[
\left\{ -\frac{1}{n} \frac{\partial s\{y_1, ..., y_n; \theta\}}{\partial \theta} \right\} \bigg|_{\theta = \hat{\theta}} \approx \frac{1}{n} I(\theta)
\]

\[
\implies LR \approx \sqrt{n}(\hat{\theta}_0 - \hat{\theta})' \frac{1}{n} I(\theta) \sqrt{n}(\hat{\theta}_0 - \hat{\theta}) \text{ for } n \to \infty
\]

- Let \( p = m \) and the restrictions are linear.

\[
c \hat{\theta}_0 - q = 0 \implies c \theta - q = c(\theta - \hat{\theta}_0)
\]

Then, under the null that the restrictions are true \( \theta = \hat{\theta}_0 \)

\[
\implies \sqrt{n}(\hat{\theta} - \hat{\theta}_0) = \sqrt{n}(\hat{\theta} - \theta) \approx N(0, nI(\theta)^{-1})
\]

\[
\implies LR \approx \chi^2_p
\]
Two properties of the LR test

- At a fixed significance level, $\alpha$, the LR test is (usually) a consistent test, i.e.:

$$\lim_{n \to \infty} \Pr\{\text{reject null} \mid \text{null false}\} = 1,$$

i.e. the power of the test converges to 1 as $n \to \infty$.

- In some situations the LR test is the most powerful possible test at any level of significance

Two practical problems with the LR test

- the $\chi^2$ limit distribution may be a poor approximation in finite samples, giving inaccurate critical values which lead to tests which have the wrong level of significance,

- to conduct the test we must maximize the Log likelihood both with and without the restrictions imposed, one of which may be inconvenient to calculate.