Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2023/24, Solutions to Tutorial No. 10

Solution to Problem 10.1: Likelihood-ratio test

In order to calculate the test function

$$\lambda = 2\left[\ln L(\hat{\vec{\beta}}) - \ln L^{\mathrm{r}}(\hat{\vec{\beta}}^{\mathrm{r}})\right] \sim \chi^2(J - J_r)$$

for the different model comparisons, we determine, from the contour plots, the ML parameter estimates and associated log-likelihoods for all four possible specifications of the Logit and Probit models:

	Probit			Logit		
V_{ni} specification	$\hat{\beta}_1$	$\hat{\beta}2$	$\ln L(\hat{\vec{\beta}})$	$\hat{\beta}_1$	$\hat{\beta}2$	$\ln L(\hat{\vec{\beta}})$
Full model M_1	-0.35	-0.08	-16	-0.4	-0.095	-16
AC-only model M_2	-0.5	–	-19.5	-0.55	–	-19.5
Time-onlymodel M_3	–	-0.08	-16.5	–	-0.955	-16.5
Trivial model M_4	–	–	-20.5	–	–	-20.5

(a) Compare the full model $M = M_1$ with the reduced "time-only" model $M_r = M_3$:

$$\lambda_{\text{Probit}} = 1, \quad \lambda_{\text{Logit}} = 1$$

Since the rejection region is given by $\lambda > \chi^2_{1,0.95} = 3.9$ (cross section of the black $\chi^2(1)$ curve with the black $F = 0.95$ line), the null hypothesis H_0: "no ad-hoc preferences" cannot be rejected. Alternatively, the p value can also directly be read off from the black graph of the $\chi^2(1)$ distribution:

$$p = 1 - F_{\chi^2(1)}(1) = 1 - 0.7 = 0.3$$

(b) Here, $M = M_1$ and $M_r = M_2$, so

$$\lambda_{\text{Probit}} = 7 > \chi^2_{1,0.95}, \quad \lambda_{\text{Logit}} = 7 > \chi^2_{1,0.95},$$

or

$$p_{\text{Logit}} = p_{\text{probit}} = 1 - F_{\chi^2(1)}(7) < 0.01$$

For both the Logit and Probit models, Model $M = M_1$ describes the data significantly better than $M_r = M_2$, so the travel time is a significant factor.

(c) Since the $\tilde{L}$ values of the four specifications are essentially the same for the Logit and Probit models, the following applies for both.

(i) $M = M_1$ (full model) vs. $M_r = M_4$ (trivial model):

$$\lambda = 2 * 4.5 = 9 > \chi^2_{2,0.95} \approx 6, \quad \text{rejection, } p = 1 - F_{\chi^2(2)}(9) \approx 0.01$$

(ii) $M = M_2$ (AC-only) vs. $M_r = M_4$ (trivial model):

$$\lambda = 2 * 1 = 2 < \chi^2_{1,0.95} \approx 4, \quad \text{no rejection, } p = 1 - F_{\chi^2(1)}(2) \approx 0.15$$

(iii) $M = M_3$ (time-only) vs. $M_r = M_4$ (trivial model):

$$\lambda = 2 * 4 = 8 > \chi^2_{1,0.95} \approx 4, \quad \text{rejection, } p = 1 - F_{\chi^2(1)}(2) \approx 0.005$$

Discussion

When performing a model selection using the top-down ansatz (starting with M_1 and eliminating the worst factors, one by one) or the bottom-up ansatz (starting with M_4 and adding the best factors, one by one, we arrive at the time-only model.

However, there are theoretical reasons (substantially different modes of transport) to keep the AC. Then, the full model will be selected.

Solution to Problem 10.2: Likelihood-ratio test for regression models: $\lambda = T^2$

(a) The likelihood-ratio test is based on log-likelihoods. Therefore, it is only applicable to models where the ML estimation can be applied. This is only possible if there are random elements with known distributions.[1]

(b) The likelihood and log-likelihood functions to data $\{(x_i, y_i)\}$, $i = 1, ..., n$ for a random term $\epsilon_i \sim i.i.d. N(0, \sigma^2)$ with known variance σ^2 is given by

$$\begin{aligned} L(\beta_0) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\beta_0)^2}{2\sigma^2}}, \\ \tilde{L}(\beta_0) = \ln L(\beta_0) &= \sum_{i=1}^{n} \left[\frac{-1}{2} \ln(2\pi\sigma^2) - \frac{(y_i - \beta_0)^2}{2\sigma^2} \right]. \end{aligned}$$

Maximizing it:

$$\frac{\partial l(\beta_0)}{\partial \beta_0} = \sum_{i=1}^{n} \left[\frac{y_i - \beta_0}{\sigma^2} \right] \overset{!}{0} \quad \Rightarrow \quad \hat{\beta}_0 = \bar{y}.$$

[1] It does not necessarily need to be Gaussian, neither to be uncorrelated; however, if $\epsilon \sim i.i.d. N(0, \sigma^2)$, the ML estimation is identical to the standard OLS calibration.

(c) The log-likelihood of the restrained model $y = \mu_0 + \epsilon$ is given by

$$\tilde{L}^{\mathrm{r}} = \sum_{i=1}^{n} \left[\frac{-1}{2} \ln(2\pi\sigma^2) - \frac{(y_i - \mu_0)^2}{2\sigma^2} \right]$$

When calculating the LR test statistics λ, the constant terms $-1/2 \ln(2\pi\sigma^2)$ of both log-likelihoods cancel out and we obtain

$$\begin{aligned}
\lambda &= 2 \left[\ln L(\bar{y}) - \ln L^{\mathrm{r}} \right] \\
&= 2 \sum_{i=1}^{n} \left[-\frac{(y_i - \bar{y})^2}{2\sigma^2} + \frac{(y_i - \mu_0)^2}{2\sigma^2} \right],
\end{aligned}$$

or, after further manipulations,

$$\begin{aligned}
\lambda\sigma^2 &= \sum_{i=1}^{n} \left[-(y_i - \bar{y})^2 + (y_i - \mu_0)^2 \right] \\
&= \sum_{i=1}^{n} \left[-y_i^2 + 2y_i\bar{y} - \bar{y}^2 + y_i^2 - 2y_i\mu_0 + \mu_0^2 \right] \\
&= \sum_{i=1}^{n} \left[2y_i(\bar{y} - \mu_0) + \mu_0^2 - \bar{y}^2 \right] \\
&= n \left[\bar{y}^2 + \mu_0^2 - 2\bar{y}\mu_0 \right] \\
&= n \left(\bar{y} - \mu_0 \right)^2
\end{aligned}$$

If H_0: "both models are equivalent" applies, i.e. $E(y) = \mu_0$, we have because of the i.i.d Gaussian random terms,

$$\bar{y} \sim N\left(\mu_0, \frac{\sigma^2}{n} \right),$$

or

$$\sqrt{\lambda} = \sqrt{n} \frac{\bar{y} - \mu_0}{\sigma} := Z \sim N(0, 1).$$

i.e.,

$$\lambda = Z^2 \sim \chi^2(1).$$

The last identity is valid since, by definition, a squared Gaussian random variable is $\chi^2(1)$ distributed (since a sum of m i.i.d squares of standardnormal distributed random variables is $\chi^2(m)$ distributed).

(d) For a known variance, the test statistics of the t-test is a standard Gaussian:

$$T = \frac{\hat{\beta}_0 - \mu_0}{\sqrt{V_{00}}} = \frac{\bar{y} - \mu_0}{\sigma} \sqrt{n} \sim N(0, 1)$$

A comparison with the results of (c) shows that $T = \sqrt{\lambda}$, hence

$$\lambda = T^2$$

Discussion

The LR test of two models with only one parameter difference gives a rejection of the null hypothesis H_0: "Both models are equivalent" if, and only if, the t-test for known error variance rejects the significance of the additional parameter of the full model at the same level. The condition of a known variance follows from the fact that the LR test is only exact in the asympotic limit $n \to \infty$ (but also gives useful results for normal-sized samples).

More generally, the LR test of two models with one or more factors difference rejects the nullhypothesis if, and only if, the F-test rejects the simultaneous null hypotheses: "all additional parameters of the full model are zero or fixed" in the asymptotic limit.

Finally, let us remind the result obtained earlier for regression models: For the general case of unknown variance but only one factor difference, the F and the T tests are equivalent