Tests

Michel Bierlaire

michel.bierlaire@epfl.ch

Transport and Mobility Laboratory

Introduction

- Impossible to determine the most appropriate model specification
- A good fit does not mean a good model
- Formal testing is necessary, but not sufficient
- No clear-cut rules can be given
- Subjective judgments of the analyst
- Good modeling = good judgment + good analysis

Introduction

Hypothesis testing. Two propositions

- H_0 null hypothesis
- H_1 alternative hypothesis

Analogy with a court trial:

- H_0: the defendant
- "Presumed innocent until proved guilty"
- H_0 is accepted, unless the data argue strongly to the contrary
- Benefit of the doubt

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

Errors are always possible:

	Accept H_0	Reject H_0
H_0 is true		Type I error (proba. α)
H_0 is false	Type II error (proba. β)	

- Type I error: send an innocent to jail
- Type II error: free a culprit

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Errors

- For a given sample size N, there is a trade-off between α and β.
- The only way to reduce both Type I and Type II error probabilities is to increase N.
- $\pi = 1 - \beta$ is the *power* of the test, that is the probability of rejecting H_0 when H_0 is false.
- H_1 is usually a composite hypothesis. π can only be determined for a simple hypothesis.
- In general, α is fixed by the analyst, and the power is maximized by the test.

Introduction

- Informal tests
- Asymptotic t-test
- Likelihood ratio tests
  - Specific attributes
  - Taste variations, market segmentation
  - Nonlinear specifications
- Goodness-of-fit measures
- Non nested hypotheses,
- Prediction tests
  - Outlier analysis
  - Market segmentation tests

Informal tests

Wilkinson (1999) “The grammar of graphics”. Springer

> ... some researchers who use statistical methods pay more attention to goodness of fit than to the meaning of the model... Statisticians must think about what the models mean, regardless of fit, or they will promulgate nonsense.

- Is the sign of the coefficient consistent with expectation?
- Are the trade offs meaningful?

Informal tests

Sign of the coefficient

Example: Netherlands Mode Choice Case

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.798	0.275	-2.90	0.00
2	β_{cost}	-0.0499	0.0107	-4.67	0.00
3	β_{time}	-1.33	0.354	-3.75	0.00

Informal tests

Value of trade-offs

- How much are we ready to pay for an improvement of the level-of-service?
- Example: reduction of travel time
- The increase in cost must be exactly compensated by the reduction of travel time

$$\beta_{\text{cost}}(C+\Delta C)+\beta_{\text{time}}(T-\Delta T)+\ldots=\beta_{\text{cost}}C+\beta_{\text{time}}T+\ldots$$

Therefore,

$$\frac{\Delta C}{\Delta T}=\frac{\beta_{\text{time}}}{\beta_{\text{cost}}}$$

Informal tests

Value of trade-offs

In general:

- Trade-off: $\frac{\partial V/\partial x}{\partial V/\partial x_C}$
- Units: $\frac{1/\text{Hour}}{1/\text{Guilder}} = \frac{\text{Guilder}}{\text{Hour}}$

Name	**Value**	Guilders	Euros	CHF	
Cte. car	-0.798	15.97	7.25	11.21	
β_{cost}	-0.0499				
β_{time}	-1.33	26.55	12.05	18.64	(/Hour)

t-test

Is the parameter θ significantly different from a given value θ^*?

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

Under H_0, if $\hat{\theta}$ is normally distributed with known variance σ^2

$$\frac{\hat{\theta} - \theta^*}{\sigma} \sim N(0, 1).$$

Therefore

$$P(-1.96 \leq \frac{\hat{\theta} - \theta^*}{\sigma} \leq 1.96) = 0.95 = 1 - 0.05$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

t-test

$$P\left(-1.96 \leq \frac{\hat{\theta}-\theta^*}{\sigma} \leq 1.96\right) = 0.95 = 1 - 0.05$$

H_0 can be rejected at the 5% level ($\alpha = 0.05$) if

$$\left|\frac{\hat{\theta}-\theta^*}{\sigma}\right| \geq 1.96.$$

- If $\hat{\theta}$ asymptotically normal
- If variance unknown
- A t test should be used with n degrees of freedom.
- When $n \geq 30$, the Student t distribution is well approximated by a $N(0,1)$

Estimator of the asymptotic variance for ML

- Cramer-Rao Bound with the estimated parameters

$$\hat{V}_{CR} = -\nabla^2 \ln L(\hat{\theta})^{-1}$$

- Berndt, Hall, Hall & Haussman (BHHH) estimator

$$\hat{V}_{BHHH} = \left(\sum_{i=1}^{n} \hat{g}_i \hat{g}_i^T \right)^{-1}$$

where

$$\hat{g}_i = \frac{\partial \ln f_X(x_i; \theta)}{\partial \theta}$$

Estimator of the asymptotic variance for ML

Robust estimator:

$$\hat{V}_{CR}\hat{V}_{BHHH}^{-1}\hat{V}_{CR}$$

- The three are asymptotically equivalent
- This one is more robust when the model is misspecified
- Biogeme uses Cramer-Rao and the robust estimators

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

t-test

Example: Netherlands Mode Choice

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.798	0.275	-2.90	0.00
2	β_{cost}	-0.0499	0.0107	-4.67	0.00
3	β_{time}	-1.33	0.354	-3.75	0.00

- $H_0 : \beta_{\text{time}} = 0$: rejected at the 5% level

t-test

Swissmetro: model specification

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	time	time	time
β_{headway}	0	headway	headway

t-test

Swissmetro: coefficient estimates

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.262	0.0615	-4.26	0.00
2	Cte. train	-0.451	0.0932	-4.84	0.00
3	β_{cost}	-0.0108	0.000682	-15.90	0.00
4	β_{headway}	-0.00535	0.000983	-5.45	0.00
5	β_{time}	-0.0128	0.00104	-12.23	0.00

- $H_0 : \beta_{\text{time}} = 0$: rejected at the 5% level
- $H_0 : \beta_{\text{cost}} = 0$: rejected at the 5% level
- $H_0 : \beta_{\text{headway}} = 0$: rejected at the 5% level

t-test

Comparing two coefficients:
$H_0 : \beta_1 = \beta_2$. The t statistic is given by

$$\frac{\widehat{\beta}_1 - \widehat{\beta}_2}{\sqrt{\mathrm{var}(\widehat{\beta}_1 - \widehat{\beta}_2)}}$$

$$\mathrm{var}(\widehat{\beta}_1 - \widehat{\beta}_2) = \mathrm{var}(\widehat{\beta}_1) + \mathrm{var}(\widehat{\beta}_2) - 2\,\mathrm{cov}(\widehat{\beta}_1, \widehat{\beta}_2)$$

t-test

Example: alternative specific coefficient

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
$\beta_{\text{time car}}$	time	0	0
$\beta_{\text{time train}}$	0	time	0
$\beta_{\text{time Swissmetro}}$	0	0	time
β_{headway}	0	headway	headway

t-test

Coefficient estimates:

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.371	0.120	-3.08	0.00
2	Cte. train	0.0429	0.121	0.36	0.72
3	β_{cost}	-0.0107	0.000669	-16.00	0.00
4	β_{headway}	-0.00532	0.000994	-5.35	0.00
5	$\beta_{\text{time car}}$	-0.0112	0.00109	-10.28	0.00
6	$\beta_{\text{time Swissmetro}}$	-0.0116	0.00182	-6.40	0.00
7	$\beta_{\text{time train}}$	-0.0156	0.00109	-14.29	0.00

t-test

Variance-covariance matrix:

Parameter	Parameter 2	Covariance	Correlation	t-stat
$\beta_{\text{time car}}$	$\beta_{\text{time train}}$	7.57e-07	0.634	4.70
$\beta_{\text{time car}}$	$\beta_{\text{time Swissmetro}}$	1.38e-06	0.696	0.31
$\beta_{\text{time Swissmetro}}$	$\beta_{\text{time train}}$	1.47e-06	0.740	3.19

- $H_0 : \beta_{\text{time car}} = \beta_{\text{time train}}$: reject
- $H_0 : \beta_{\text{time car}} = \beta_{\text{time Swissmetro}}$: cannot reject
- $H_0 : \beta_{\text{time Swissmetro}} = \beta_{\text{time train}}$: reject

Likelihood ratio test

- Used for "nested" hypotheses
- One model is a special case of the other obtained from a set of restrictions on the parameters
- H_0: restrictions are valid

$$-2(\mathcal{L}(\hat{\beta}_R) - \mathcal{L}(\hat{\beta}_U)) \sim \chi^2_{(K_U - K_R)}$$

- $\mathcal{L}(\hat{\beta}_R)$ is the log likelihood of the restricted model
- $\mathcal{L}(\hat{\beta}_U)$ is the log likelihood of the unrestricted model
- K_R is the number of parameters in the restricted model
- K_U is the number of parameters in the unrestricted model

Likelihood ratio test

Example: Netherlands Mode Choice Case.

- Unrestricted model:
  - 3 parameters: β_{time}, β_{cost}, Cte. car.
  - Final log likelihood: -123.133
- Restricted model
  - Restrictions: $\beta_{\text{time}} = \beta_{\text{cost}} = 0$
  - 1 parameter: Cte. car.
  - Final log likelihood: -148.347
- Test: $-2(-148.35 - 123.13) = 50.43$
- χ^2, 2 degrees of freedom, 95% quantile: 5.99
- H_0 is rejected
- The unrestricted model is preferred.

Likelihood ratio test

Test of generic attributes: Swissmetro

- Unrestricted model:

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
$\beta_{\text{time car}}$	time	0	0
$\beta_{\text{time train}}$	0	time	0
$\beta_{\text{time Swissmetro}}$	0	0	time
β_{headway}	0	headway	headway

Likelihood ratio test

Test of generic attributes: Swissmetro

- Restricted model:

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	time	time	time
β_{headway}	0	headway	headway

- Restrictions: $\beta_{\text{time car}} = \beta_{\text{time train}} = \beta_{\text{time Swissmetro}}$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Likelihood ratio test

- Log likelihood of the restricted model: -5315.386
- Number of parameters for the restricted model: 5
- Log likelihood of the unrestricted model: -5297.488
- Number of parameters for the restricted model: 7
- Test: 35.796
- χ^2, 2 degrees of freedom, 95% quantile: 5.99
- Reject the restrictions
- The alternative specific specification is preferred

Likelihood ratio test

Test of taste variations

- Unrestricted model: a different set of parameters for each income group
- 1: [0–50], 2: [50–100], 3:[100–], 4: unknown (KCHF)
- Restricted model: same parameters across income groups
- Socio-economic characteristics: for $i = 1, \ldots, 4$

$$I_i = \begin{cases} 1 & \text{if individual belongs to income group } i \\ 0 & \text{otherwise} \end{cases}$$

Likelihood ratio test: restricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
$\beta_{\text{time car}}$	time	0	0
$\beta_{\text{time train}}$	0	time	0
$\beta_{\text{time Swissmetro}}$	0	0	time
β_{headway}	0	headway	headway

Likelihood ratio test: unrestricted model

	Car	Train	Swissmetro
Cte. car (income 1)	I_1	0	0
Cte. train (income 1)	0	I_1	0
$\beta_{\text{cost},1}$	cost $\cdot I_1$	cost $\cdot I_1$	cost $\cdot I_1$
$\beta_{\text{time car},1}$	time $\cdot I_1$	0	0
$\beta_{\text{time train},1}$	0	time $\cdot I_1$	0
$\beta_{\text{time Swissmetro},1}$	0	0	time $\cdot I_1$
$\beta_{\text{headway},1}$	0	headway $\cdot I_1$	headway $\cdot I_1$
Cte. car (income 2)	I_2	0	0
Cte. train (income 2)	0	I_2	0
$\beta_{\text{cost},1}$	cost $\cdot I_2$	cost $\cdot I_2$	cost $\cdot I_2$
$\beta_{\text{time car},1}$	time $\cdot I_2$	0	0
$\beta_{\text{time train},1}$	0	time $\cdot I_2$	0
$\beta_{\text{time Swissmetro},1}$	0	0	time $\cdot I_2$
$\beta_{\text{headway},1}$	0	headway $\cdot I_2$	headway $\cdot I_2$

Likelihood ratio test: unrestricted model (ctd)

	Car	Train	Swissmetro
Cte. car (income 3)	I_3	0	0
Cte. train (income 3)	0	I_3	0
$\beta_{\text{cost},1}$	cost $\cdot I_3$	cost $\cdot I_3$	cost $\cdot I_3$
$\beta_{\text{time car},1}$	time $\cdot I_3$	0	0
$\beta_{\text{time train},1}$	0	time $\cdot I_3$	0
$\beta_{\text{time Swissmetro},1}$	0	0	time $\cdot I_3$
$\beta_{\text{headway},1}$	0	headway $\cdot I_3$	headway $\cdot I_3$
Cte. car (income 4)	I_4	0	0
Cte. train (income 4)	0	I_4	0
$\beta_{\text{cost},1}$	cost $\cdot I_4$	cost$\cdot I_4$	cost $\cdot I_4$
$\beta_{\text{time car},1}$	time $\cdot I_4$	0	0
$\beta_{\text{time train},1}$	0	time $\cdot I_4$	0
$\beta_{\text{time Swissmetro},1}$	0	0	time $\cdot I_4$
$\beta_{\text{headway},1}$	0	headway $\cdot I_4$	headway $\cdot I_4$

Likelihood ratio test: unrestricted model (ctd)

Estimation:

- Divide the sample into 4 subsets, corresponding to the income groups
- Estimate the restricted model on each of the sample separately
- Add up the log likelihood

Group	Log likelihood	Sample size
1	-926.84	1161
2	-1679.53	2133
3	-1946.75	2907
4	-478.4	567
Total	-5031.51	6768

Likelihood ratio test

- Unrestricted model:
  - $7 \times 4 = 28$ parameters
  - Final log likelihood: -5031.51
- Restricted model:
  - 7 parameters
  - Final log likelihood: -5297.488
- Test: 531.956
- χ^2, 21 degrees of freedom, 95% quantile: 32.67
- H_0 is rejected
- There is evidence of taste variation per income group

Nonlinear specifications

- Consider a variable x of the model (travel time, say)
- Unrestricted model: V is a nonlinear function of x
- Restricted model: V is a linear function of x
- We consider the following nonlinear specifications:
  - Piecewise linear
  - Power series
  - Box-Cox transforms
- For each of them, the linear specification is obtained using simple restrictions on the nonlinear specification

Piecewise linear specification

- Partition the range of values of x into M intervals $[a_m, a_{m+1}]$, $m = 1, \ldots, M$
- For example, the partition [0–500], [500–1000], [1000–] corresponds to

$$M = 3, a_1 = 0, a_2 = 500, a_3 = 1000, a_4 = +\infty$$

- The slope of the utility function may vary across intervals
- Therefore, there will be M parameters instead of 1
- The function must be continuous

Piecewise linear specification

- Linear specification:

$$V_i = \beta x_i + \cdots$$

- Piecewise linear specification

$$V_i = \sum_{m=1}^{M} \beta_m x_{im} + \cdots$$

where

$$x_{im} = \max(0, \min(x - a_m, a_{m+1} - a_m))$$

that is

$$x_{im} = \begin{cases} 0 & \text{if } x < a_m \\ x - a_m & \text{if } a_m \leq x < a_{m+1} \\ a_{m+1} - a_m & \text{if } a_{m+1} \leq x \end{cases}$$

Piecewise linear specification

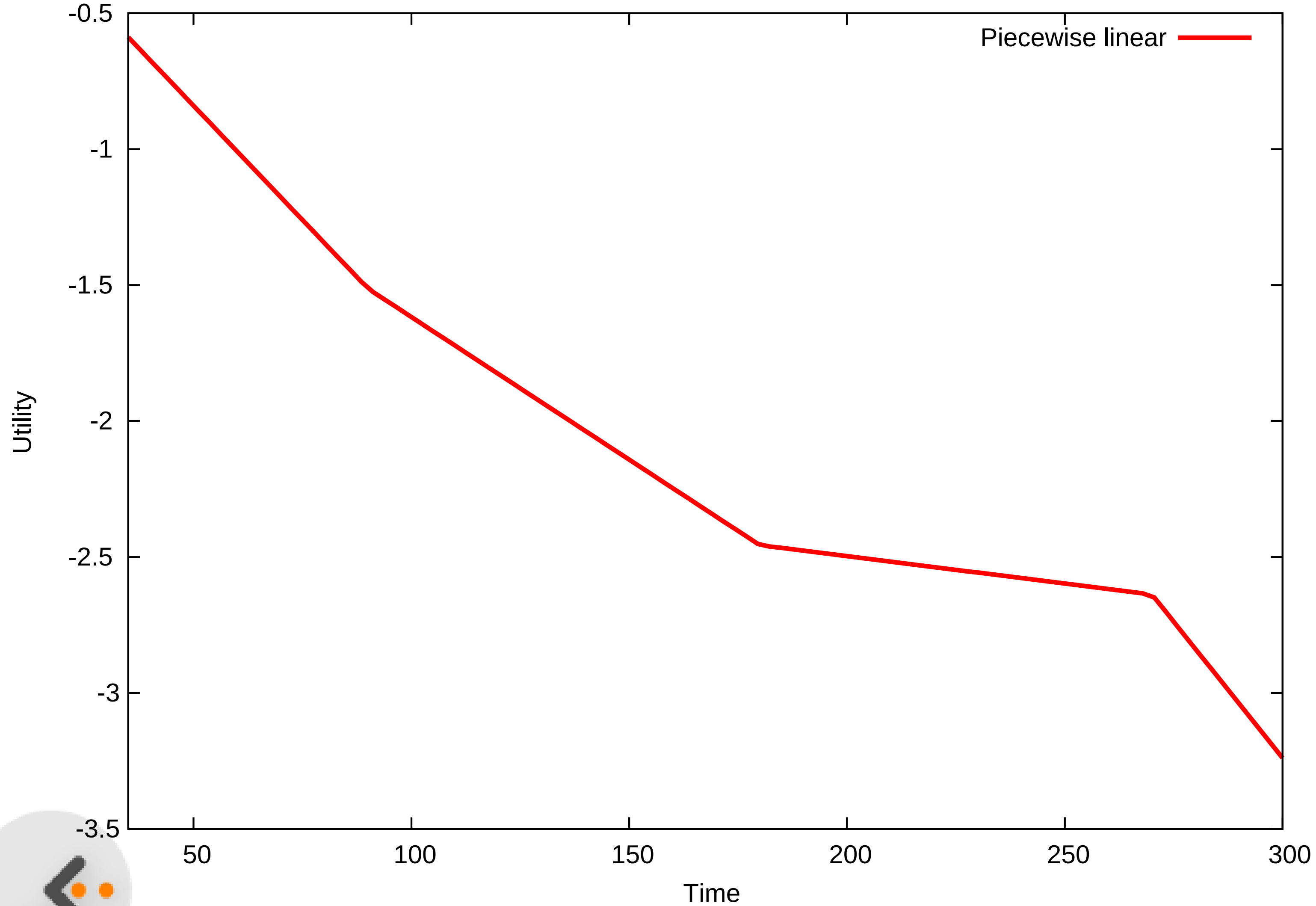
Example: $M = 3, a_1 = 0, a_2 = 500, a_3 = 1000, a_4 = +\infty$

x	x_1	x_2	x_3
40	40	0	0
600	500	100	0
1200	500	500	200

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Piecewise linear specification



Piecewise linear specification: restricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	time	time	time
β_{headway}	0	headway	headway

Piecewise linear specification: unrestricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
$\beta_{\text{time},1}$	time_1	time_1	time_1
$\beta_{\text{time},2}$	time_2	time_2	time_2
$\beta_{\text{time},3}$	time_3	time_3	time_3
β_{headway}	0	headway	headway

Piecewise linear specification

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.145	0.0473	-3.05	0.00
2	Cte. train	-0.265	0.0730	-3.64	0.00
3	β_{cost}	-0.0113	0.000703	-16.04	0.00
4	β_{headway}	-0.00544	0.000996	-5.46	0.00
5	$\beta_{\text{time},1}$	-0.0155	0.000655	-23.58	0.00
6	$\beta_{\text{time},2}$	0.0137	0.00144	9.47	0.00
7	$\beta_{\text{time},3}$	-0.0168	0.00471	-3.56	0.00

Likelihood ratio test

- Unrestricted model:
  - 7 parameters
  - Final log likelihood: -5214.741
- Restricted model:
  - 5 parameters
  - Final log likelihood: -5315.386
- Test: 201.29
- χ^2, 2 degrees of freedom, 95% quantile: 5.99
- H_0 is rejected
- The linear specification is rejected

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Power series

- Idea: if the utility function is nonlinear in x, it can be approximated by a polynomial if degree M
- Linear specification:

$$V_i = \beta x_i + \cdots$$

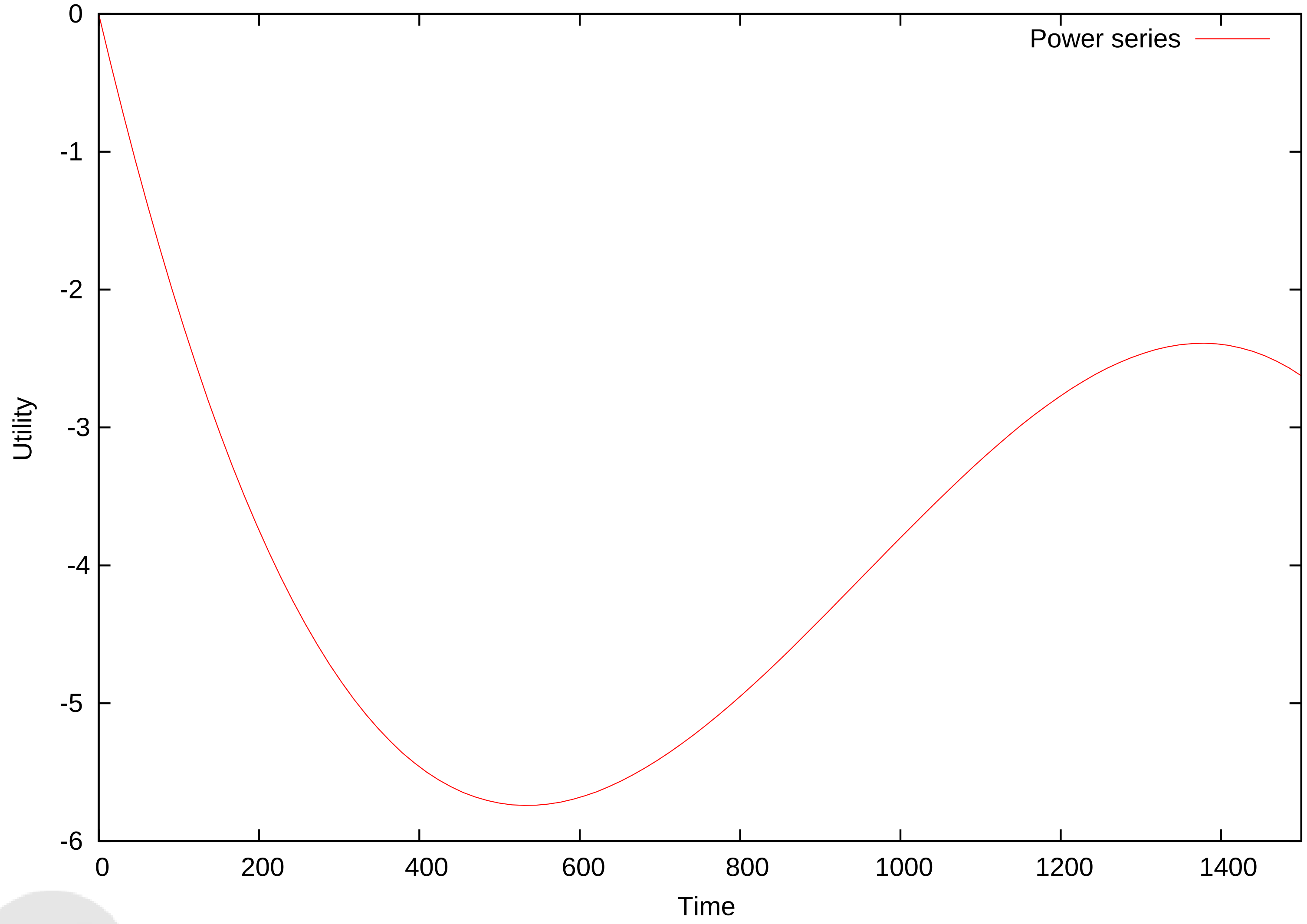
- Power series

$$V_i = \sum_{m=1}^{M} \beta_m x_i^m + \cdots$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Power series: M=3



Power series: restricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	time	time	time
β_{headway}	0	headway	headway

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Power series: unrestricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
$\beta_{\text{time},1}$	time	time	time
$\beta_{\text{time},2}$	$\text{time}^2/10^5$	$\text{time}^2/10^5$	$\text{time}^2/10^5$
$\beta_{\text{time},3}$	$\text{time}^3/10^5$	$\text{time}^3/10^5$	$\text{time}^3/10^5$
β_{headway}	0	headway	headway

Power series: unrestricted model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.0556	0.0493	-1.13	0.26
2	Cte. train	-0.148	0.0752	-1.96	0.05
3	β_{cost}	-0.0111	0.000693	-15.98	0.00
4	β_{headway}	-0.00536	0.000991	-5.41	0.00
5	$\beta_{\text{time},1}$	-0.0247	0.00123	-20.04	0.00
6	$\beta_{\text{time},2}$	3.21	0.322	9.98	0.00
7	$\beta_{\text{time},3}$	-0.00112	0.000181	-6.18	0.00

Likelihood ratio test

- Unrestricted model:
  - 7 parameters
  - Final log likelihood: -5223.233
- Restricted model:
  - 5 parameters
  - Final log likelihood: -5315.386
- Test: 184.306
- χ^2, 2 degrees of freedom, 95% quantile: 5.99
- H_0 is rejected
- The linear specification is rejected

Box-Cox transform

- Let $x > 0$ be a positive variable
- Its Box-Cox transform is defined as

$$B(x, \lambda) = \frac{x^\lambda - 1}{\lambda},$$

- Special cases:

$$B(x, 1) = x - 1, \ \lim_{\lambda \to 0} B(x, \lambda) = \ln x.$$

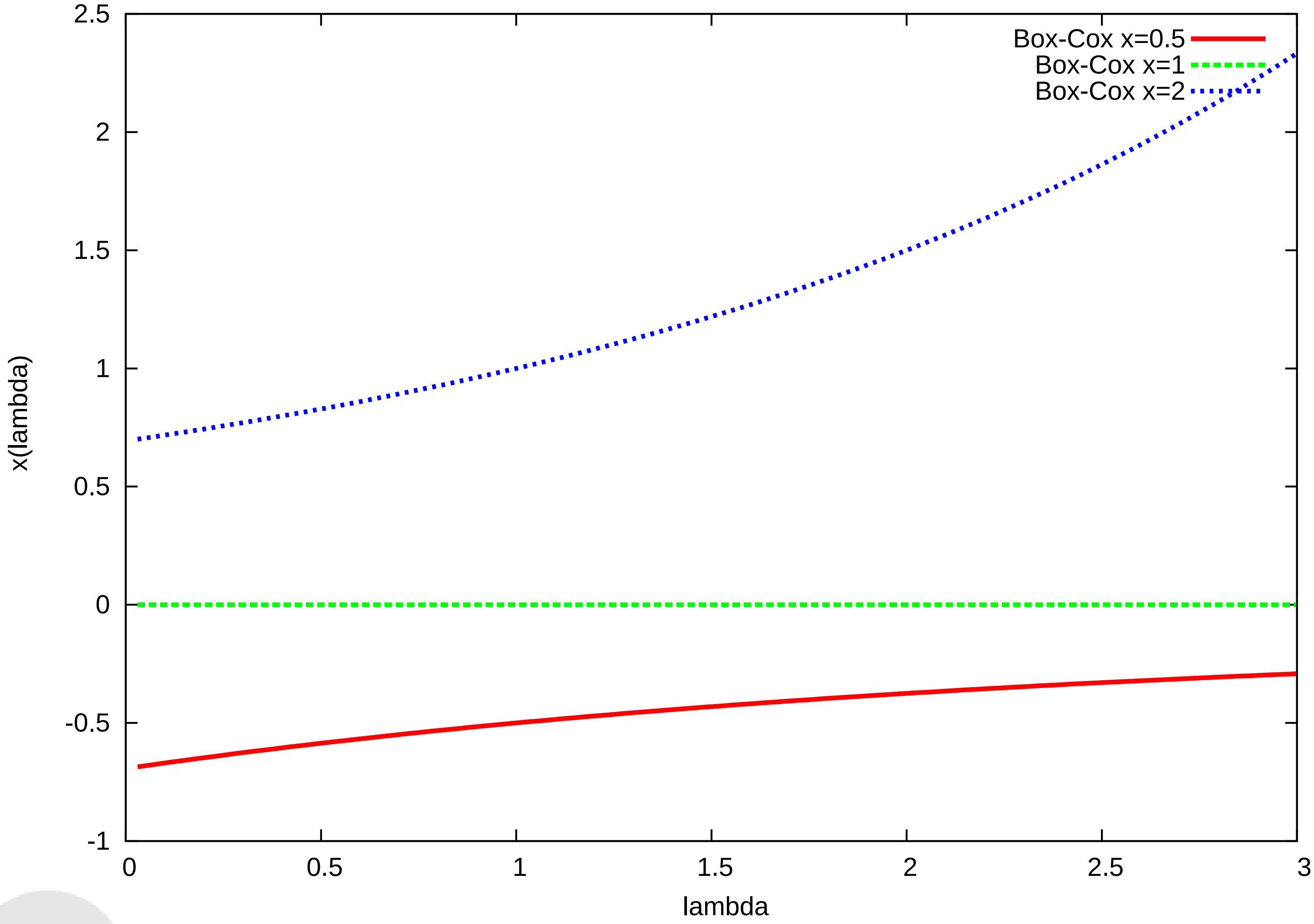
- Linear specification:

$$V_i = \beta x_i + \cdots$$

- Box-Cox specification

$$V_i = \beta B(x, \lambda) + \cdots = \beta \frac{x^\lambda - 1}{\lambda} + \cdots$$

Box-Cox transform



Box-Cox: restricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	time	time	time
β_{headway}	0	headway	headway

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Box-Cox: unrestricted model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	B(time,λ)	B(time,λ)	B(time,λ)
β_{headway}	0	headway	headway
λ			

Note: specification tables are not designed for nonlinear specifications.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

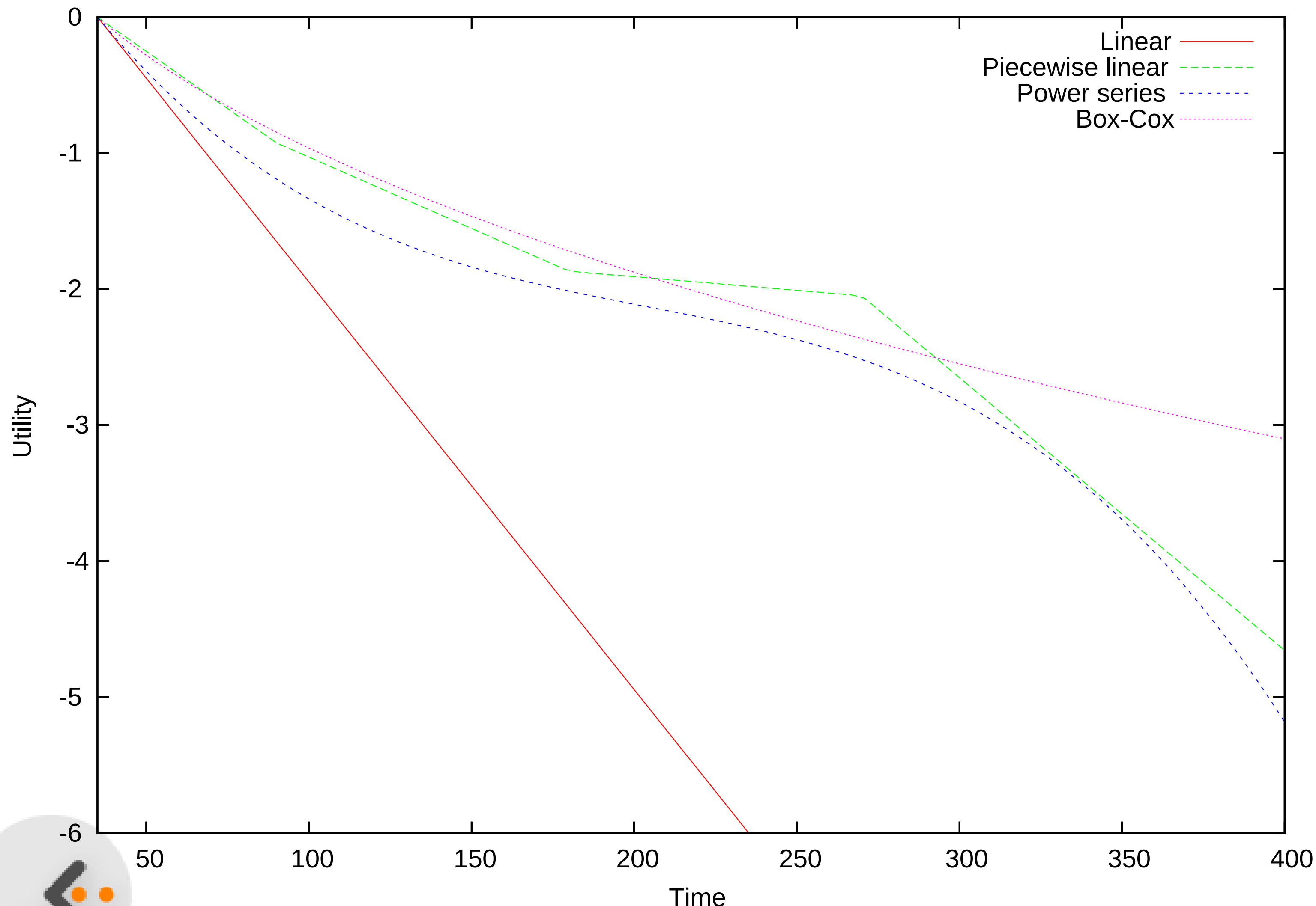
Box-Cox: unrestricted model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.112	0.0517	-2.16	0.03
2	Cte. train	-0.236	0.0781	-3.02	0.00
3	β_{cost}	-0.0108	0.000680	-15.87	0.00
4	β_{headway}	-0.00533	0.000985	-5.41	0.00
5	β_{time}	-0.160	0.0568	-2.82	0.00
6	λ	0.510	0.0776	6.57	0.00

Likelihood ratio test

- Unrestricted model:
  - 6 parameters
  - Final log likelihood: -5276.353
- Restricted model:
  - 5 parameters
  - Final log likelihood: -5315.386
- Test: 78.066
- χ^2, 1 degree of freedom, 95% quantile: 3.84
- H_0 is rejected
- The linear specification is rejected

Comparison



0
-1
-2
-3
-4
-5
-6
Utility
50
100
150
200
250
300
350
400
Time
Linear
Piecewise linear
Power series
Box-Cox

Non-nested hypotheses

- Need to compare two different models
- If none of the models is a restricted version of the other, we talk about non-nested models
- The likelihood ratio test cannot be used
- Three possible tests:
  - Cox test
  - Davidson-McKinnon J-test
  - Horowitz test $\bar{\rho}^2$

Cox test

- We want to test model 1 against model 2
- We generate a composite model C such that both models 1 and 2 are restricted cases of model C.
- We test 1 against C using the likelihood ratio test
- We test 2 against C using the likelihood ratio test
- Possible outcomes:
  - Only one of the two models is rejected. Keep the other.
  - Both models are rejected. Better models should be developed.
  - Both models are accepted. Use another test.

Cox test

Model 1

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
$\beta_{\text{cost car}}$	cost	0	0
$\beta_{\text{cost Swissmetro}}$	0	0	cost
$\beta_{\text{cost train}}$	0	cost	0
$\beta_{\text{gen. abo.}}$	0	GA	GA
β_{headway}	0	headway	headway
$\beta_{\text{time car}}$	time	0	0
$\beta_{\text{time Swissmetro}}$	0	0	time
$\beta_{\text{time train}}$	0	time	0

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Cox test: estimates for model 1

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.403	0.116	-3.48	0.00
2	Cte. train	0.126	0.116	1.08	0.28
3	$\beta_{\text{cost car}}$	-0.00776	0.00150	-5.18	0.00
4	$\beta_{\text{cost Swissmetro}}$	-0.0108	0.000828	-12.99	0.00
5	$\beta_{\text{cost train}}$	-0.0300	0.00200	-14.97	0.00
6	$\beta_{\text{gen. abo.}}$	0.513	0.194	2.65	0.01
7	β_{headway}	-0.00535	0.00101	-5.31	0.00
8	$\beta_{\text{time car}}$	-0.0129	0.00162	-7.94	0.00
9	$\beta_{\text{time Swissmetro}}$	-0.0111	0.00179	-6.19	0.00
10	$\beta_{\text{time train}}$	-0.00866	0.00120	-7.22	0.00

Cox test

Model 2: cost of car appears as a log

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
$\beta_{\text{log cost car}}$	log(cost)	0	0
$\beta_{\text{cost Swissmetro}}$	0	0	cost
$\beta_{\text{cost train}}$	0	cost	0
$\beta_{\text{gen. abo.}}$	0	GA	GA
β_{headway}	0	headway	headway
$\beta_{\text{time car}}$	time	0	0
$\beta_{\text{time Swissmetro}}$	0	0	time
$\beta_{\text{time train}}$	0	time	0

Cox test: estimates for model 2

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	1.39	0.437	3.18	0.00
2	Cte. train	0.138	0.117	1.18	0.24
3	$\beta_{\text{log cost car}}$	-0.547	0.135	-4.04	0.00
4	$\beta_{\text{cost Swissmetro}}$	-0.0105	0.000812	-12.96	0.00
5	$\beta_{\text{cost train}}$	-0.0297	0.00199	-14.93	0.00
6	$\beta_{\text{gen. abo.}}$	0.560	0.193	2.90	0.00
7	β_{headway}	-0.00531	0.00101	-5.28	0.00
8	$\beta_{\text{time car}}$	-0.0133	0.00170	-7.83	0.00
9	$\beta_{\text{time Swissmetro}}$	-0.0110	0.00179	-6.16	0.00
10	$\beta_{\text{time train}}$	-0.00868	0.00120	-7.23	0.00

Cox test

	Log likelihood	# parameters
Model 1 (linear car cost)	-5047.205	10
Model 2 (log car cost)	-5056.262	10

- The fit of model 1 is better
- But we cannot apply a likelihood ratio test
- We estimate a composite model

Cox test

Composite model

	Car	Train	Swissmetro
Cte. car	1	0	0
Cte. train	0	1	0
$\beta_{\text{cost car}}$	cost	0	0
$\beta_{\text{log cost car}}$	log(cost)	0	0
$\beta_{\text{cost Swissmetro}}$	0	0	cost
$\beta_{\text{cost train}}$	0	cost	0
$\beta_{\text{gen. abo.}}$	0	GA	GA
β_{headway}	0	headway	headway
$\beta_{\text{time car}}$	time	0	0
$\beta_{\text{time Swissmetro}}$	0	0	time
$\beta_{\text{time train}}$	0	time	0

Cox test: estimates of the composite model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-1.26	0.865	-1.46	0.14
2	Cte. train	0.118	0.116	1.02	0.31
3	$\beta_{\text{cost car}}$	-0.0105	0.00279	-3.76	0.00
4	$\beta_{\text{log cost car}}$	0.258	0.267	0.97	0.33
5	$\beta_{\text{cost Swissmetro}}$	-0.0108	0.000827	-13.00	0.00
6	$\beta_{\text{cost train}}$	-0.0299	0.00200	-14.96	0.00
7	$\beta_{\text{gen. abo.}}$	0.501	0.193	2.59	0.01
8	β_{headway}	-0.00535	0.00101	-5.31	0.00
9	$\beta_{\text{time car}}$	-0.0130	0.00170	-7.65	0.00
10	$\beta_{\text{time Swissmetro}}$	-0.0110	0.00179	-6.16	0.00
11	$\beta_{\text{time train}}$	-0.00858	0.00120	-7.18	0.00

Cox test

- Test 1: model 1 vs. composite
  - Unrestricted model:
    - 11 parameters
    - Final log likelihood: -5046.418
  - Restricted model:
    - 10 parameters
    - Final log likelihood: -5047.205
  - Test: 1.58
  - χ^2, 1 degree of freedom, 95% quantile: 3.84
  - H_0 cannot be rejected
  - Model 1 cannot be rejected

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Cox test

- Test 2: model 2 vs. composite
  - Unrestricted model:
    - 11 parameters
    - Final log likelihood: -5046.418
  - Restricted model:
    - 10 parameters
    - Final log likelihood: -5056.262
  - Test: 18.104
  - χ^2, 1 degree of freedom, 95% quantile: 3.84
  - H_0 can be rejected
  - Model 2 can be rejected

Conclusion: model 1 (linear car cost) is preferred over model 2 (log car cost).

Davidson and McKinnon J-test

- Model 1: $U_n^{(1)} = V_n^{(1)}(x_n^{(1)}|\beta) + \varepsilon_n^{(1)}$
- Model 2: $U_n^{(2)} = V_n^{(2)}(x_n^{(2)}|\gamma) + \varepsilon_n^{(2)}$
- Hypothesis H_0: model 1 is correct.
- Procedure:
  1. Estimate model 2 and obtain $\widehat{\gamma}$.
  2. Consider the composite model

$$U_n^{(1)} = (1-\alpha)V_n^{(1)}(x_n^{(1)}|\beta) + \alpha V_n^{(2)}(x_n^{(2)}|\widehat{\gamma}) + \varepsilon_n.$$

  3. Estimate β and α.
  4. Under H_0, we have $\alpha = 0$.
  5. It can be tested with a t-test.

Davidson and McKinnon J-test: test model 1

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	-0.535	0.158	-3.39	0.00
2	Cte. Swissmetro	-0.126	0.117	-1.08	0.28
3	$\beta_{\text{cost car}}$	-0.00584	0.00231	-2.53	0.01
4	$\beta_{\text{cost Swissmetro}}$	-0.0108	0.000831	-12.94	0.00
5	$\beta_{\text{cost train}}$	-0.0300	0.00201	-14.91	0.00
6	$\beta_{\text{gen. abo.}}$	0.513	0.194	2.64	0.01
7	β_{headway}	-0.00535	0.00101	-5.29	0.00
8	$\beta_{\text{time car}}$	-0.0129	0.00163	-7.92	0.00
9	$\beta_{\text{time Swissmetro}}$	-0.0111	0.00180	-6.17	0.00
10	$\beta_{\text{time train}}$	-0.00866	0.00120	-7.19	0.00
11	α	0.00355	0.00277	1.28	0.20

Davidson and McKinnon J-test: test model 1

- The hypothesis that $\alpha = 0$ cannot be rejected.
- Model 1 cannot be rejected.

Davidson and McKinnon J-test: test model 2

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	Cte. car	5.16	13.5	0.38	0.70
2	Cte. Swissmetro	-0.164	0.376	-0.44	0.66
3	$\beta_{\text{cost Swissmetro}}$	-0.0100	0.00221	-4.53	0.00
4	$\beta_{\text{cost train}}$	-0.0292	0.00613	-4.77	0.00
5	$\beta_{\text{gen. abo.}}$	0.665	0.695	0.96	0.34
6	β_{headway}	-0.00524	0.00324	-1.61	0.11
7	$\beta_{\text{log cost car}}$	-1.73	4.09	-0.42	0.67
8	$\beta_{\text{time car}}$	-0.0143	0.00862	-1.66	0.10
9	$\beta_{\text{time Swissmetro}}$	-0.0108	0.00529	-2.05	0.04
10	$\beta_{\text{time train}}$	-0.00873	0.00393	-2.22	0.03
11	α	0.687	0.775	0.89	0.38

Davidson and McKinnon J-test: test model 2

- The hypothesis that $\alpha = 0$ cannot be rejected.
- Model 2 cannot be rejected.
- The test is not conclusive.

Goodness-of-fit

$$\rho^2 = 1 - \frac{\mathcal{L}(\hat{\beta})}{\mathcal{L}(0)}$$

- $\rho^2 = 0$: trivial model, equal probabilities
- $\rho^2 = 1$: perfect fit.

Warning: $\mathcal{L}(\hat{\beta})$ is a biased estimator of the expectation over all samples. Use $\mathcal{L}(\hat{\beta}) - K$ instead.

$$\bar{\rho}^2 = 1 - \frac{\mathcal{L}(\hat{\beta}) - K}{\mathcal{L}(0)}$$

$\bar{\rho}^2$ test (Horowitz)

Compare model 0 and model 1.

- We expect that the best model corresponds to the best fit.
- We will be wrong if M_0 is the true model and M_1 produces a better fit.
- What is the probability that this happens?
- If this probability is low, M_0 can be rejected.

$$P(\bar{\rho_1}^2 - \bar{\rho_0}^2 > z | M_0) \leq \Phi\left(-\sqrt{-2z\mathcal{L}(0) + (K_1 - K_0)}\right)$$

where

- $\bar{\rho_\ell}^2$ is the adjusted likelihood ratio index of model $\ell = 0, 1$
- K_ℓ is the number of parameters of model ℓ
- Φ is the standard normal CDF.

$\bar{\rho}^2$ test (Horowitz)

Back to the example:

	$\bar{\rho}^2$	# parameters
Model 0 (log car cost)	0.272	10
Model 1 (linear car cost)	0.273	10

$$P({\bar{\rho}_1}^2 - {\bar{\rho}_0}^2 > z | M_0) \leq \Phi\left(-\sqrt{-2z\mathcal{L}(0) + (K_1 - K_0)}\right)$$

$$P({\bar{\rho}_1}^2 - {\bar{\rho}_0}^2 > 0.001 | M_0) \leq \Phi\left(-\sqrt{-2z(-6958.425) + (10 - 10)}\right)$$

$$P({\bar{\rho}_1}^2 - {\bar{\rho}_0}^2 > 0.001 | M_0) \leq \Phi\left(-3.73\right) \approx 0$$

Therefore, M_0 can be rejected, and the linear specification is preferred.

$\bar{\rho}^2$ test (Horowitz)

In practice,

- if the sample is large enough (i.e. more than 250 observations),
- if the values of the $\bar{\rho}^2$ differ by 0.01 or more,
- the model with the lower $\bar{\rho}^2$ is almost certainly incorrect.

Outlier analysis

- Apply the model on the sample
- Examine observations where the predicted probability is the smallest for the observed choice
- Test model sensitivity to outliers, as a small probability has a significant impact on the log likelihood
- Potential causes of low probability:
  - Coding or measurement error in the data
  - Model misspecification
  - Unexplainable variation in choice behavior

Outlier analysis

- Coding or measurement error in the data
  - Look for signs of data errors
  - Correct or remove the observation
- Model misspecification
  - Seek clues of missing variables from the observation
  - Keep the observation and improve the model
- Unexplainable variation in choice behavior
  - Keep the observation
  - Avoid over fitting of the model to the data

Market segments

- Compare predicted vs. observed shares per segment
- Let N_g be the set of samples individuals in segment g
- Observed share for alt. i and segment g

$$S_g(i) = \sum_{n \in N_g} y_{in}/N_g$$

- Predicted share for alt. i and segment g

$$\hat{S}_g(i) = \sum_{n \in N_g} P_n(i)/N_g$$

Market segments

Note:

- With a full set of constants for segment g:

$$\sum_{n \in N_g} y_{in} = \sum_{n \in N_g} P_n(i)$$

- Do not saturate the model with constants

Conclusions

- Tests are designed to check meaningful hypotheses
- Do not test hypotheses that do not make sense
- Do not apply the tests blindly
- Always use your judgment.