Statistical Tests

Matthieu de Lapparent
matthieu.delapparent@epfl.ch

Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne

Introduction

Modeling

- Difficult to determine the most appropriate model specification
- A good fit does not imply a good model
- Formal testing is necessary, but not sufficient
- No clear-cut rules can be given
- Good modeling = good (subjective) judgment + good analysis

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.

Introduction

Hypothesis testing

Four steps

- step 1: State the hypotheses
  - H_0 null hypothesis
  - H_1 alternative hypothesis
- step 2: Set the criteria for a decision
- step 3: Compute a test statistic
- step 4: Make a decision

Step 1: Analogy with a court trial

- H_0: defendant is “presumed innocent until proved guilty”
- H_0 is accepted, unless the data argue strongly to the contrary

Introduction

Step 2: Criterion for a decision

- Court-room: criterion is to show guilt beyond reasonable doubt
- Implies defining the level of significance α

Step 3: Test statistic

- Determine the likelihood of obtaining a sample outcome if the H_0 hypothesis were true
- How far we accept to be from the H_0

Step 4: Decide

- Decide if null is retained or rejected
- Gives the probability value (p-value of obtaining an outcome, given that the H_0 is true)

Introduction

Possible decision outcomes

	Accept H_0	Reject H_0
H_0 is true	Correct (1-α)	Type I error (prob. α)
H_0 is false	Type II error (prob. β)	Correct (1-β)

Relations

- For a given sample size N, there is a trade-off between α and β.
- Only way to reduce both types of error probabilities is to increase N.
- $\pi = 1 - \beta$ is the *power* of the test, that is, the probability of correctly rejecting H_0.
- Researcher directly controls Type I errors by fixing α

Summary of case-studies

Netherlands mode choice

- Intercity travelers
- Choice between train & car
- 228 respondents
- Revealed preference data with self-reported trip characteristics

Swissmetro

- Travelers St. Gallen - Geneva
- Choice between train, car & swissmetro
- 441 respondents
- Stated preference (swissmetro is a non-existing mag-lev train)

Informal tests

Sign of the coefficient

Do the estimated parameters have the right sign?

Example: Netherlands Mode Choice Case

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

Informal tests

Value of trade-offs

Are the trade-offs reasonable?

- How much are we ready to pay for a marginal 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: example with Netherlands data

In general:

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

Parameter	Coeff.	Guilders	Euros	CHF	
ASC 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

Question

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

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

Statistic

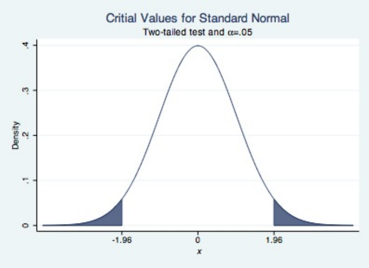
Under H_0, if $\hat{\theta}$ is normally distributed with known variance σ^2

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

Therefore

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

t-test



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

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

Comments

- 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 $N(0, 1)$

t-test

Swissmetro: model specification

	Car	Train	Swissmetro
ASC car	1	0	0
ASC 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	ASC car	-0.262	0.0615	-4.26	0.00
2	ASC 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

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

Comparing two coefficients

Example: alternative specific or generic coefficients? Below alternative specific time

	Car	Train	Swissmetro
ASC car	1	0	0
ASC 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

Swissmetro: coefficient estimates (alternative specific time)

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC car	-0.371	0.120	-3.08	0.00
2	ASC 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 1	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

$H_0 : \beta_{\text{time car}} = \beta_{\text{time train}}$

$$\begin{aligned}
\text{var}(\widehat{\beta}_{t.car} - \widehat{\beta}_{t.train}) &= \text{var}(\widehat{\beta}_{t.car}) + \text{var}(\widehat{\beta}_{t.train}) - 2\,\text{cov}(\widehat{\beta}_{t.car}, \widehat{\beta}_{t.train}) \\
&= 1.188 \times 10^{-6} + 3.312 \times 10^{-06} - 2 \times 7.570 \times 10^{-07} \\
&= 8.622 \times 10^{-07}
\end{aligned}$$

t-test

$H_0 : \beta_{\text{time car}} = \beta_{\text{time train}}$

$$\frac{\widehat{\beta}_{t.car} - \widehat{\beta}_{t.train}}{\sqrt{\text{var}(\widehat{\beta}_{t.car} - \widehat{\beta}_{t.train})}} = \frac{-0.0112 - (-0.0156)}{\sqrt{8.622 \times 10^{-07}}} = 4.739$$

We can reject the H_0 of parameter equality

What about $\beta_{\text{time car}} = \beta_{\text{time metro}}$ and $\beta_{\text{time metro}} = \beta_{\text{time train}}$?

Homework to calculate the t-ratios for these parameter differences!

Likelihood ratio test

Comparing two models

- Used for "nested" hypotheses
- One model is a special case of another obtained from a set of linear restrictions on the parameters
- H_0: the restricted model is the true model

Statistic under H_0

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

- $\mathcal{L}(\hat{\beta}_R)$ 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

Test of parameters being equal to zero: Netherlands

- Unrestricted model:
  - 3 parameters: β_{time}, β_{cost}, ASC car.
  - Final log likelihood: -123.133
- Restricted model
  - Restrictions: $\beta_{\text{time}} = \beta_{\text{cost}} = 0$
  - 1 parameter: ASC car.
  - Final log likelihood: -148.347

Statistic

- 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

- Restricted model:

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

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

- Unrestricted model:

	Car	Train	Swissmetro
ASC car	1	0	0
ASC 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:
  - Final log likelihood: -5315.386
  - 5 parameters
- Unrestricted model:
  - Final log likelihood: -5297.488
  - 7 parameters

Statistic

- -2(-5315.386 - -5297.488) = 35.796
- χ^2, 2 degrees of freedom, 95% quantile: 5.99
- Reject the restrictions (H_0)
- The alternative specific specification is preferred

Test of taste variations

Segmentation

- Classify the data into G groups. Size of group g: N_g.
- The same specification is considered for each group.
- A different set of parameters is estimated for each group.
- Restrictions:

$$\beta^1 = \beta^2 = \ldots = \beta^G$$

  where β^g is the vector of coefficients of market segment g.
- Statistic:

$$-2\left[\mathcal{L}_N(\widehat{\beta}) - \sum_{g=1}^{G} \mathcal{L}_{N_g}(\widehat{\beta}^g)\right]$$

- χ^2 with $\sum_{g=1}^{G} K_g - K$ degrees of freedom.
- In general, $\sum_{g=1}^{G} K_g - K = (G-1)K$.

Test of taste variations

Segmentation according to income: Swissmetro

- 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

Hypothesis

H_0 the true parameters are the same across income classes

Estimation results by income groups

Estimation procedure

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

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

Different taste across income groups?

Test of taste variations

- Restricted model:
  - 7 parameters
  - Final log likelihood: -5297.488
- Unrestricted model:
  - $7\times\ 4 = 28$ parameters
  - Final log likelihood: -5031.51

Statistic

- Likelihood ratio test gives: 531.956
- χ^2, 21 degrees of freedom, 95% quantile: 32.67
- $531.956 > 32.67$ hence 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 case, the linear specification is obtained using simple restrictions on the nonlinear specification

Piecewise linear specification

Model procedure

- 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

Piecewise linear specification: restricted model

Test of piecewise specification

- Restricted model

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

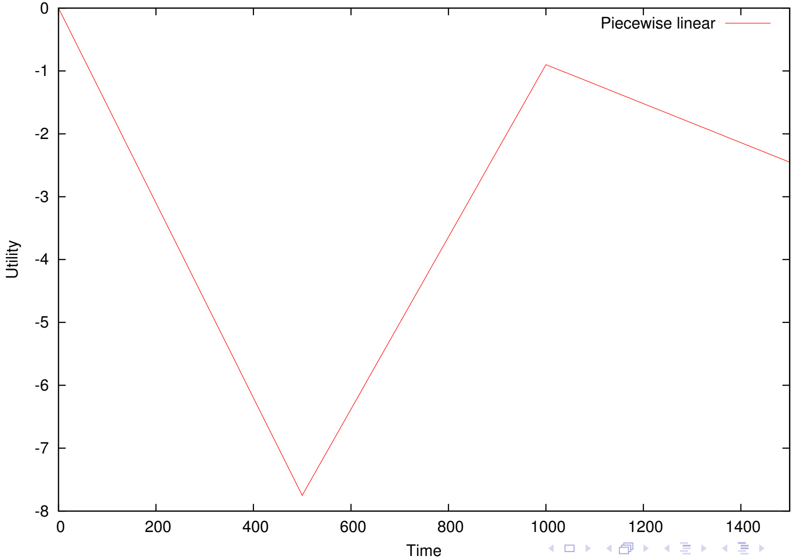
- Unrestricted model

	Car	Train	Swissmetro
ASC car	1	0	0
ASC 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	ASC car	-0.145	0.0473	-3.05	0.00
2	ASC 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

Piecewise linear specification



Likelihood ratio test

Test of piecewise linear specification for time

- Restricted model:
  - 5 parameters
  - Final log likelihood: -5315.386
- Unrestricted model:
  - 7 parameters
  - Final log likelihood: -5214.741

Statistic

- LR Test: 201.29
- χ^2, 2 degrees of freedom, 95% quantile: 5.99
- H_0 is rejected
- The linear specification is rejected

Power series

Idea

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

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

- Power series

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

Power series: restricted model

Test of power series specification for time

- Restricted model

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

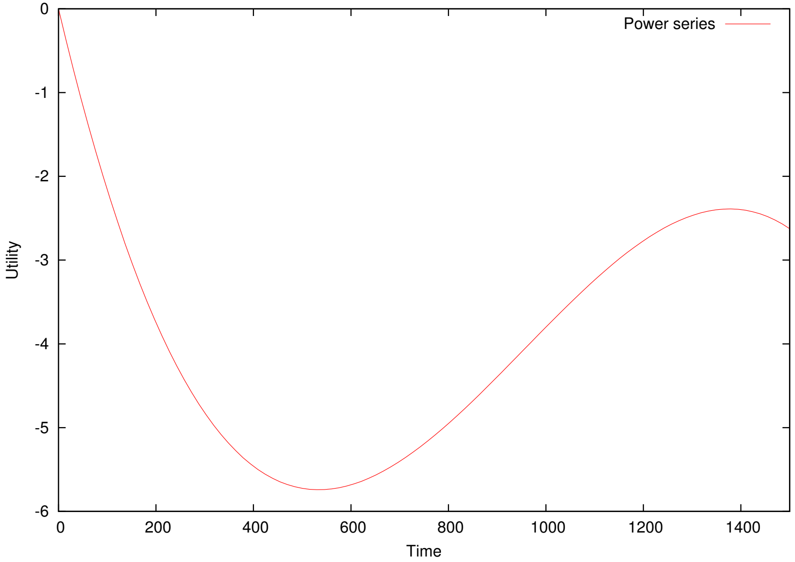
- Unrestricted model

	Car	Train	Swissmetro
ASC car	1	0	0
ASC 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	ASC car	-0.0556	0.0493	-1.13	0.26
2	ASC 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

Power series: M=3



Likelihood ratio test

Test of power series specification for time

- Restricted model:
  - 5 parameters
  - Final log likelihood: -5315.386
- Unrestricted model:
  - 7 parameters
  - Final log likelihood: -5223.233

Statistic

- LR 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

Definition

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

$$B(x,\lambda) = \begin{cases} \dfrac{x^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln x & \text{if } \lambda = 0. \end{cases}$$

where $\lambda \in \mathbb{R}$ is a parameter.

Continuity

$$\lim_{\lambda \to 0} \frac{x^\lambda - 1}{\lambda} = \ln x.$$

Box-Cox transform

Linear specification

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

Box-Cox specification

$$V_i = \beta B(x, \lambda) + \cdots$$

Properties

- Convex if $\lambda > 1$
- Linear if $\lambda = 1$
- Concave if $\lambda < 1$

Box-Cox specification

Test of Box-Cox transformation on time

- Restricted model

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

- Unrestricted model

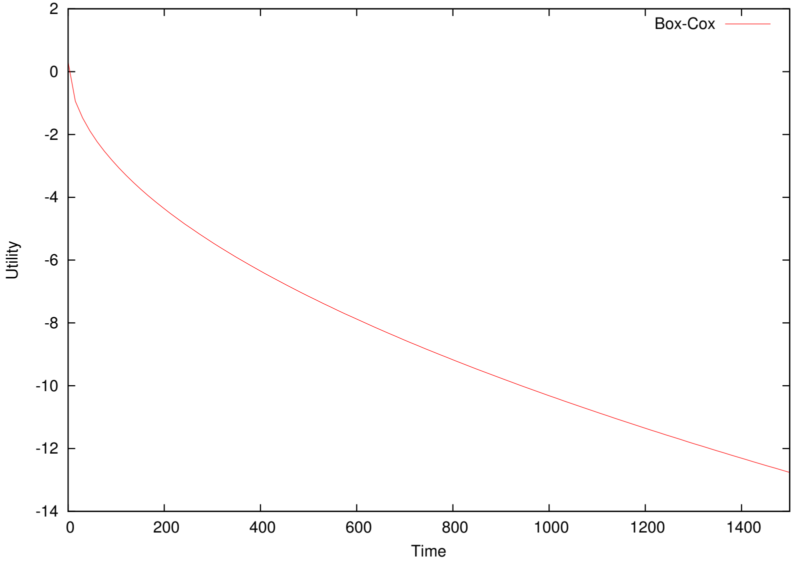
	Car	Train	Swissmetro
ASC car	1	0	0
ASC train	0	1	0
β_{cost}	cost	cost	cost
β_{time}	B(time,λ)	B(time,λ)	B(time,λ)
β_{headway}	0	headway	headway
λ			

Note: specification tables are not designed for nonlinear specifications.

Box-Cox: unrestricted model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC car	-0.112	0.0517	-2.16	0.03
2	ASC 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

Box-Cox transform



Likelihood ratio test

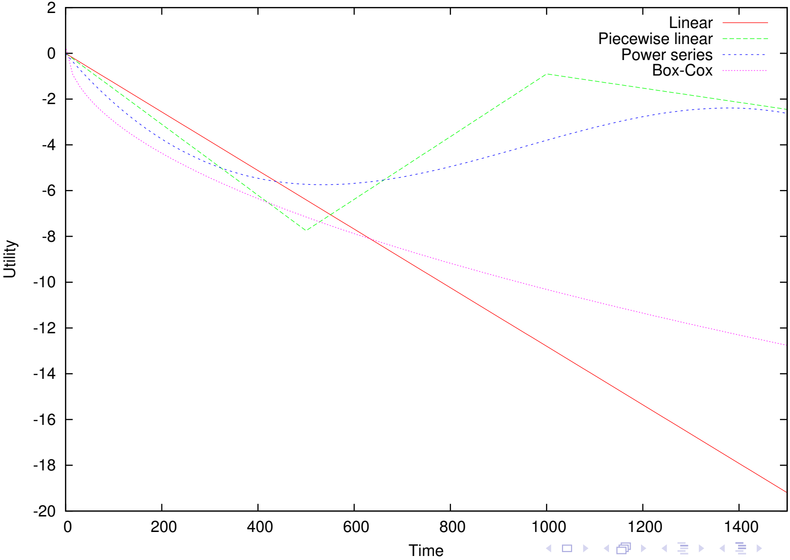
Test of Box-Cox specification for time

- Restricted model:
  - 5 parameters
  - Final log likelihood: -5315.386
- Unrestricted model:
  - 6 parameters
  - Final log likelihood: -5276.353

Statistic

- LR Test: 78.066
- χ^2, 1 degree of freedom, 95% quantile: 3.84
- H_0 is rejected
- The linear specification is rejected
- Also possible to employ t-test to compare Box-Cox to linear

Comparison of nonlinear time specifications



Non nested hypotheses

Nested hypotheses

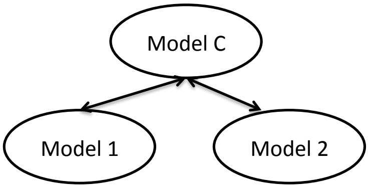
- Restricted and unrestricted models
- Linear restrictions
- H_0: restricted model is correct
- Test: likelihood ratio test

Non nested hypotheses

- Need to compare two models
- None of them is a restriction of the other
- Likelihood ratio test cannot be used
- Two possible tests:
  - Cox composite model
  - 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 model 1 against C using the likelihood ratio test
- We test model 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

Models

- $M_1 : U_{in} = \cdots + \beta x_{in} + \cdots + \varepsilon_{in}^{(1)}$
- $M_2 : U_{in} = \cdots + \theta log(x)_{in} + \cdots + \varepsilon_{in}^{(2)}$
- $M_C : U_{in} = \cdots + \beta x_{in} + \theta log(x)_{in} + \cdots + \varepsilon_{in}.$

Testing M_1 against M_C

Restrictions: $\theta = 0$

Testing M_2 against M_C

Restrictions: $\beta = 0$

Non nested models: estimates for model 1

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC car	-0.403	0.116	-3.48	0.00
2	ASC 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

Non nested models: estimates for model 2

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC car	1.39	0.437	3.18	0.00
2	ASC 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

Non nested models

	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

Non nested models: estimates of the composite model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC car	-1.26	0.865	-1.46	0.14
2	ASC 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

Non nested models

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

Non nested models

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

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

Adjusted likelihood ratio index

Likelihood ratio index

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

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

Adjusted likelihood ratio index

- ρ^2 is increasing with the number of parameters.
- A higher fit (that is a higher ρ^2) does not mean a better model.
- An adjustment is needed.

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

$\bar{\rho}^2$ 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

Procedure

- 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

Procedure

- Compare predicted vs. observed shares per segment
- Let N_g be the set of sampled 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.

90%, 95% and 99% of the χ^2 distribution with K degrees of freedom

K	90%	95%	99%	K	90%	95%	99%
1	2.706	3.841	6.635	21	29.615	32.671	38.932
2	4.605	5.991	9.210	22	30.813	33.924	40.289
3	6.251	7.815	11.345	23	32.007	35.172	41.638
4	7.779	9.488	13.277	24	33.196	36.415	42.980
5	9.236	11.070	15.086	25	34.382	37.652	44.314
6	10.645	12.592	16.812	26	35.563	38.885	45.642
7	12.017	14.067	18.475	27	36.741	40.113	46.963
8	13.362	15.507	20.090	28	37.916	41.337	48.278
9	14.684	16.919	21.666	29	39.087	42.557	49.588
10	15.987	18.307	23.209	30	40.256	43.773	50.892
11	17.275	19.675	24.725	31	41.422	44.985	52.191
12	18.549	21.026	26.217	32	42.585	46.194	53.486
13	19.812	22.362	27.688	33	43.745	47.400	54.776
14	21.064	23.685	29.141	34	44.903	48.602	56.061
15	22.307	24.996	30.578	35	46.059	49.802	57.342
16	23.542	26.296	32.000	36	47.212	50.998	58.619
17	24.769	27.587	33.409	37	48.363	52.192	59.893
18	25.989	28.869	34.805	38	49.513	53.384	61.162
19	27.204	30.144	36.191	39	50.660	54.572	62.428
20	28.412	31.410	37.566	40	51.805	55.758	63.691