Model comparison

Key concepts & study plan

Experimental design

Data collection & processing

Model specification & estimation

Interpretation & application

Model comparison

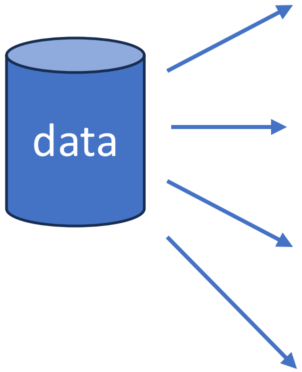
Overview

- Different models might result in different conclusions
- An infinite number of possible models exists
- Need approach to select a good model

Model comparison

Important

- It only makes sense to compare models estimated on the same data set

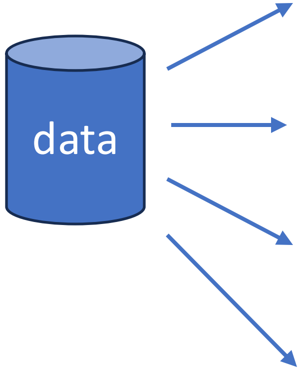


1. MNL model without gender effects
2. MNL model with gender effects
3. MNL model with only alternative-specific constants
4. Mixed MNL model with random preference heterogeneity

Model comparison

Typical question

- Which model fits the data best?

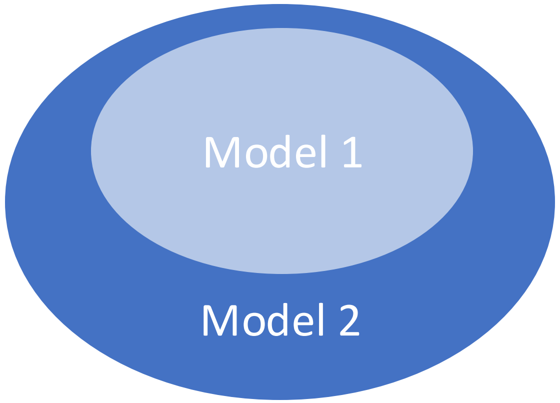


1. MNL model without gender effects
2. MNL model with gender effects
3. MNL model with only alternative-specific constants
4. Mixed MNL model with random preference heterogeneity

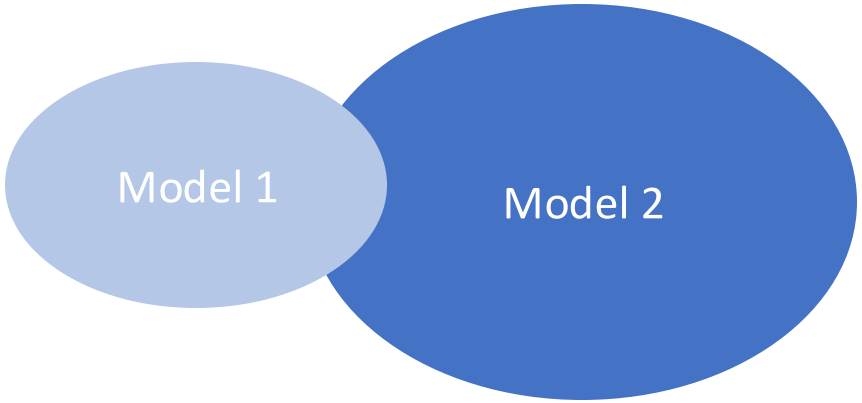
Model comparison

Two types of model comparisons

- Comparing nested models



- Comparing non-nested models



Models are nested if one is a restricted version of the other
(e.g., by setting one or more parameters to zero)

Model comparison

Two types of model comparisons

- Comparing nested models

1. $V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j$
2. $V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j$

Model 1 is a restricted version of Model 2

- Comparing non-nested models

1. $V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j$
2. $V_j = \beta_{time} * TT_j + \beta_{log-cost} * \log(TC_j)$

Model 1 is <u>not</u> a restricted version of Model 2

Model comparison

Should we just look at log-likelihood value?

- Which model would you use?

	Model specification	Loglikelihood value
1	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j$	LL = -538.3
2	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j$	LL = -534.9
3	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j + \beta_{cost*age} * TC_j * age + \beta_{seat} Seat_j + \beta_{time*age} TT_j * age$	LL = -533.5

Model comparison

Assessing model fit

Assessment methods

	Nested models	Non-nested models
❑ Likelihood ratio (LR) test	■	
❑ Akaike Information Criterion (AIC)	■	■
❑ Bayesian Information Criterion (BIC)	■	■
❑ Adjusted ρ^2	■	■
❑ Ben-Akiva and Swait test		■

Likelihood ratio test

Key concepts & study plan

Experimental design

Data collection & processing

Model specification & estimation

Interpretation & application

Likelihood ratio test

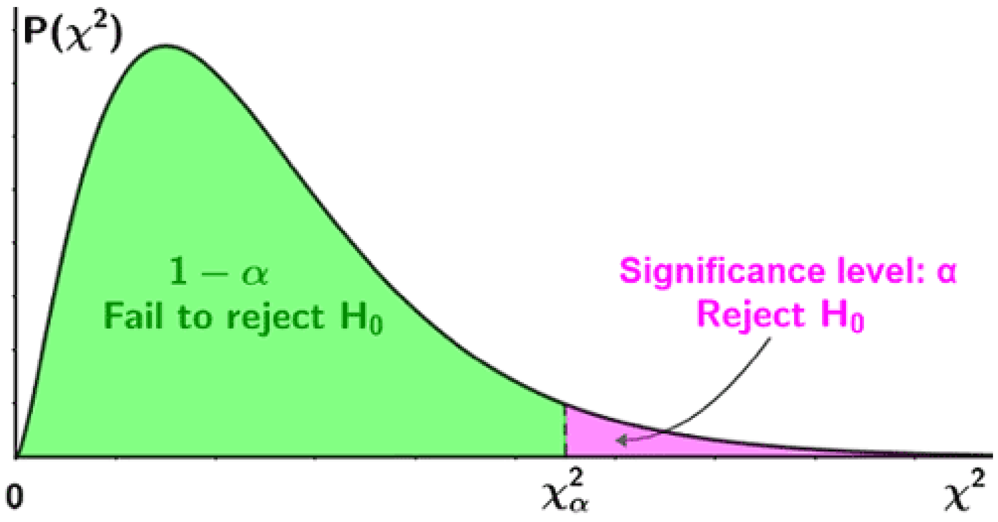
Statistical significance test

- Compare nested models
  - Restricted model with K parameters
  - General model with K+p parameters
- Hypothesis testing
  - H_0: The p additional parameters do NOT improve the model
  - H_1: The p additional parameters do improve the model

- H_0 is rejected if $-2log\left(\frac{L_{restricted}}{L_{general}}\right) = -2\left(LL_{restricted} - LL_{general}\right) > \chi^2_{p,1-\alpha}$

Likelihood ratio test

The χ_p^2 distribution



(source: math.net)

$\chi^2_{p,1-\alpha}$	1-α				
v	0.90	0.95	0.975	0.99	0.999
1	2.706	3.841	5.024	6.635	10.828
2	4.605	5.991	7.378	9.210	13.816
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.833	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322
8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588
11	17.275	19.675	21.920	24.725	31.264
12	18.549	21.026	23.337	26.217	32.91

Likelihood ratio test

Example 1

- Assume $a = 0.05$
- Test statistic: $-2(LL_{restricted} - LL_{general}) = -2(-538.3 - (-534.9)) = 6.8$
- Critical value: $\chi^2_{1,0.95}$=3.841
- Model 2 has a significantly better model fit since 6.8 > 3.841

	Model specification	Loglikelihood value
1	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j$	LL = -538.3
2	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j$	LL = -534.9

Likelihood ratio test

Example 2

- Assume $a = 0.05$
- Test statistic: $-2(LL_{restricted} - LL_{general}) = -2(-534.9 - (-533.5)) = 2.8$
- Critical value: $\chi^2_{3,0.95}$=7.815
- Model 3 does not have a significantly better model fit since 2.8 < 7.815

	Model specification	Loglikelihood value
2	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j$	LL = -534.9
3	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j + \beta_{cost*age} * TC_j * age + \beta_{seat} Seat_j + \beta_{time*age} TT_j * age$	LL = -533.5

AIC and BIC

Key concepts & study plan | Experimental design | Data collection & processing | **Model specification & estimation** | Interpretation & application

AIC and BIC

Indicators not based on statistical significance test

- Compare any two models (nested or non-nested)
- Information criteria
  - Akaike Information Criterion (AIC) $= 2 * K - 2 * LL$
  - Bayesian Information Criterion (BIC) $= \log(N) * K - 2 * LL$

LL = loglikelihood value

$K =$ number of model parameters

$N =$ number of choice observations

- A model has a better model fit if it has a smaller AIC or BIC
- If you prefer a more parsimonious model (with less parameters) then use BIC

AIC and BIC

Example

- Model 1
  - $AIC = 2 * K - 2 * LL = 2 * 2 - 2 * (-538.3) = 1080.6$
  - $BIC = \log(N) * K - 2 * LL$=2*log(1200)-2*(-538.3) = 1090.78

Model 1 has better fit according to BIC

- Model 2
  - $AIC = 2 * K - 2 * LL = 2 * 3 - 2 * (-534.9) = 1075.8$
  - $BIC = \log(N) * K - 2 * LL$=3*log(1200)-2*(-534.9) = 1091.07

Model 2 has better fit according to AIC

	Model specification	Loglikelihood value	Observations
1	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j$	LL = -538.3	N = 1200
2	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j$	LL = -534.9	N = 1200

Adjusted ρ^2

Key concepts
& study plan

Experimental
design

Data collection
& processing

**Model specification
& estimation**

Interpretation
& application

Adjusted ρ^2

Indicators not based on statistical significance test

- Compare any two models (nested or non-nested)
- Criterion resembles R^2 in linear regression
  - $\bar{\rho}^2 = \dfrac{LL - K}{LL(0)}$ $LL(0)$=N*log(1/J)

LL = loglikelihood value

K = number of model parameters

N = number of choice observations

J = number of alternatives

- A model has a better model fit if it has a larger adjusted $\bar{\rho}^2$
- Model selection is the same as with AIC

Adjusted ρ^2

Example

- Model 1
  - $\bar{\rho}^2 = 1 - \frac{LL - K}{LL(0)} = \frac{-538.3 - 2}{1200 * \log(0.5)} = 0.35$
- Model 2
  - $\bar{\rho}^2 = 1 - \frac{LL - K}{LL(0)} = \frac{-534.9 - 3}{1200 * \log(0.5)} = 0.353$

Model 2 has better fit

	Model specification	Loglikelihood value	Observations
1	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j$	LL = -538.3	N = 1200
2	$V_j = \beta_{time} * TT_j + \beta_{cost} * TC_j + \beta_{time*seating} * TT_j * Seat_j$	LL = -534.9	N = 1200

Adjusted ρ^2

Ben-Akiva and Swait (1986) test

- Tests for statistical significance of difference in $\bar{\rho}^2$
- Mainly for non-nested models (for nested models the Likelihood Ratio test is preferred)

Model comparison guidance

Key concepts & study plan

Experimental design

Data collection & processing

Model specification & estimation

Interpretation & application

Model comparison guidance

So which method should you use?

Comparing nested models

- Likelihood Ratio test

Comparing non-nested models

- AIC and Ben-Akiva & Swait test give similar outcome
- BIC if a more parsimonious model is preferred (with fewer parameters)