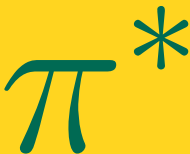


$\pi^*$



Juraj Medzihorsky



12. december 2013

# Outline

Model fit

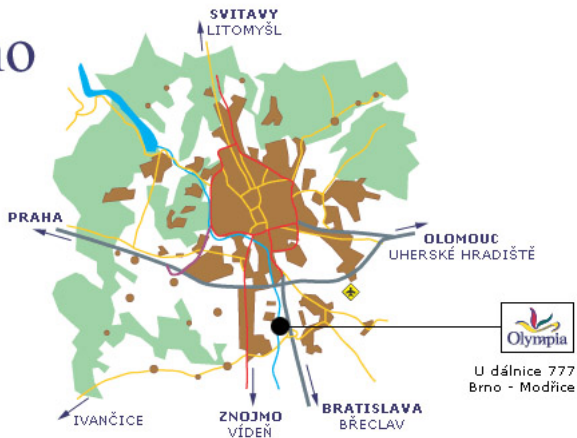
Čo je  $\pi^*$

Balík pistar

Použitia

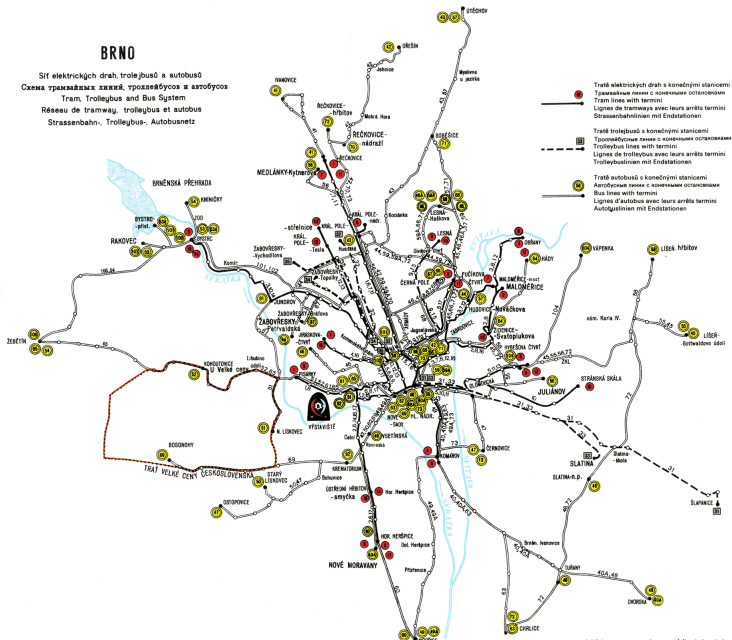


# Brno



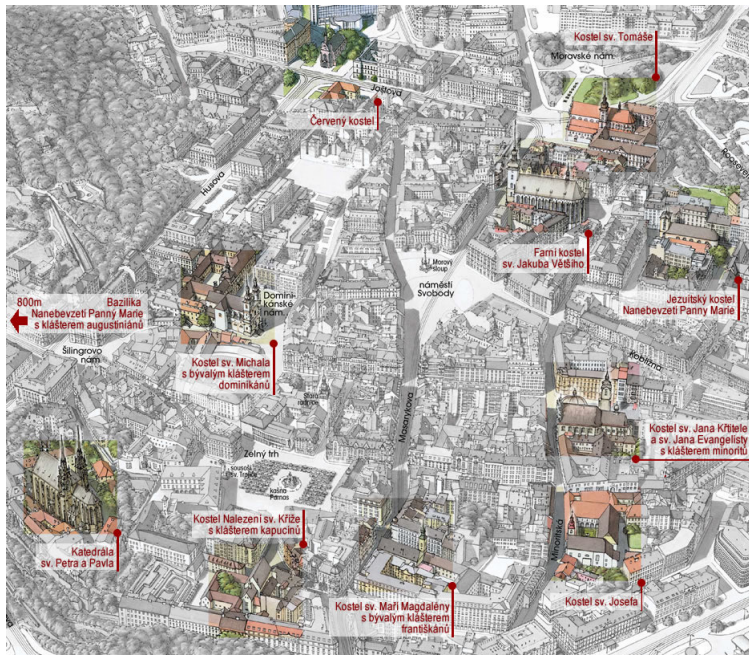
# BRNO

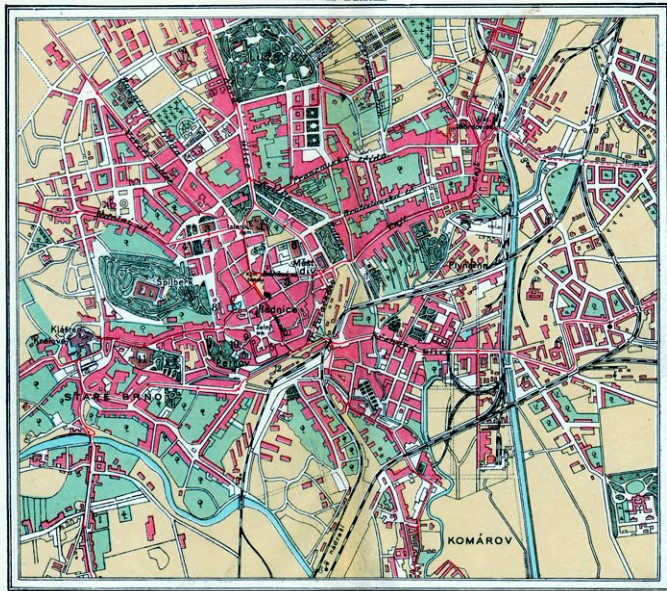
Síť elektrických drah, trolejbusů a autobusů  
 Схема трамвайных линий, троллейбусов и автобусов  
 Tram, Trolleybus and Bus System  
 Réseau de tramway, trolleybus et autobus  
 Strassenbahn, Trolleybus, Autobusnetz



stav: 1974

<http://bmhd.vhd.cz>





Měřítko 1 : 25,000

- 1 Zem. správa politická.
- 2 Národní divadlo.
- 3 Starý zemský dům.

- 4 Radnice
- 5 Chrám sv. Jakuba.
- 6 Nejvyš. s. dvůr.

*Význačné budovy:*

- 7 Ředitelství pošt a telegrafů.
- 8 Průmyslové museum.
- 9 Zemské museum.
- 10 Dům.
- 11 Zem. nemocnice.
- 12 Státní nádraží.

Reprodukce Voj. zeměp. ústavu v Praze.



# Vyhodnocovanie modelov.

Rudas-Clogg-Lindsay-ov zmesový index vhodnosti.

Miera vhodnosti modelov.

Populácia ako zmes dvoch tried:

1. trieda popísaná modelom;
2. trieda nepopísaná modelom.

$\pi^*$

Hodnota indexu je veľkosť časti populácie nepopísanej modelom.

Nadobúda hodnoty od 0 do 1.

$\pi^*$ 

$$\pi^*(O, \mathcal{M}) = \inf \left\{ \pi : \begin{array}{l} O = (1 - \pi)M + \pi U, \\ M \in \mathcal{M}, \\ U \text{ unspecified} \end{array} \right\}$$

# $\pi^*$ : výpočet

Hľadáme také hodnoty parametrov modelu, ktoré dávajú najmenšiu možnú časť populácie mimo modelu.

- ▶ Predpoklady vždy splnené.
- ▶ Priamočiary výklad.
- ▶ Nezávisí od veľkosti vzorky.
- ▶ Neistotu zhŕňa interval spoľahlivosti.
- ▶ Nevyžaduje stochastické vzorky.
- ▶ Umožňuje porovnávať rozdielne modely.

$\pi^*$

Analýza modelom nepopísanej časti populácie  
ako analýza zbytkov (reziduálov).

# pistar

- ▶ R balík pre výpočet  $\pi^*$ .
- ▶ Dostupný na [GitHub-e](#).



$\pi^*$

# Použitia

- ▶ Jednorozmerné modely.
- ▶ Viacrozmerné normálne rozdelenie.
- ▶ Modely kontingenčných tabuliek.
- ▶ Lineárna a logistická regresia.

# Nový vývoj

- ▶ Zovšeobecnenie pre chýbajúce dáta.  
(Rudas 2005, Rudas & Verdes 2012)
- ▶ Bayesovská verzia.

$\pi^*$

# References

- ▶ Rudas, T., Clogg. C., Lindsay, B. (1994) 'A New Index of Fit Based on Mixture Methods for the Analysis of Contingency Tables.' *JRSS(b)*. 56(4): 623–639.
- ▶ Rudas, T. Zwick, R. (1997) 'Estimating the Importance of Differential Item Functioning.' *Journal of Educational and Behavioral Statistics*. 22(1): 31–45.
- ▶ Rudas, T. (1999) 'The mixture index of fit and minimax regression.' *Metrika*. 50: 163–172.
- ▶ Rudas, T. (2005) 'Mixture Models of Missing Data.' *Quality & Quantity*. 39: 19–36.
- ▶ Verdes, E., Rudas, T. (2003), 'The  $\pi^*$  Index as a New Alternative for Assessing Goodness of Fit of Logistic Regression'.

# References

- ▶ Revuelta, J. (2008) 'Estimating the  $\pi^*$  goodness of fit index for finite mixtures of item response models.' *BJMSP*. 61: 93–113.
- ▶ Dayton, C.M. (2003) 'Applications and computational strategies for the two-point mixture index of fit.' *BJMSP*. 56: 1–13.
- ▶ Dayton, C.M. (2007) 'Applications and extensions of the two-point mixture index of fit.' in: *Advances in Latent Variable Mixture Models*.
- ▶ Knott, M. (2005) 'A measure of independence for a multivariate normal distribution and some connections with factor analysis'. *Journal of Multivariate Analysis*. 96(2): 374–383.
- ▶ Langeheine, R. (2012) 'Konsequenzen des Ignorierens von globalen Modelltests in Studien wie TIMSS, PISA und PIRLS'. in *Item-Response-Modelle in der sozialwissenschaftlichen Forschung*.