

Další teorie měření

PSYb2590: Základy psychometriky | Přednáška 6

25. 4. 2022 | Hynek Cígler

Opakování: Teorie měření

Teorie měření postuluje model/y měření.

Model měření: vztah pozorovaného a měřeného.

Protože měření je v psychologii nepřímé, tato inference zahrnuje zejména:

- Předpoklady o podstatě měřeného atributu.
- Způsob tvorby skóru (resp. odhad „parametru“) z pozorovaných výsledků.
- Odhad chyby tohoto odhadu 😊

Slovo „měření“ budeme nadále používat velmi vágně (viz první přednáška a povaha atributu).

Modely měření v psychologii

Podle Borsbooma (2005):

- Model klasické testové teorie (CTT).
 - Operacionalismus.
 - Není měřením v pravém slova smyslu.
 - Kromě CTT i teorie zobecnitelnosti.
- Model latentních proměnných.
 - Realismus.
 - Faktorová analýza, teorie odpovědi na položku.
- Reprezentační model měření.
 - Spojité měření (conjoint measurement).

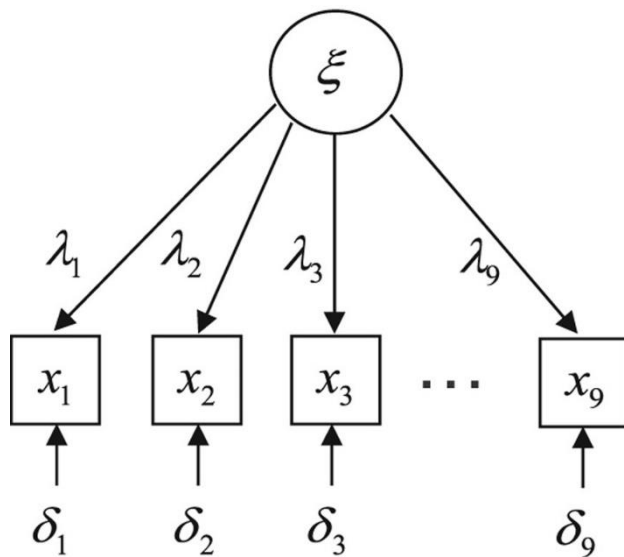
Podle Schmittmanna ([2013](#)):

- Reflektivní model
 - Rys způsobuje chování.
- Formativní model.
 - Rys je konstrukcí na základě chování.
- Dynamické systémy.
 - Jevy spolu souvisejí a rys neexistuje.

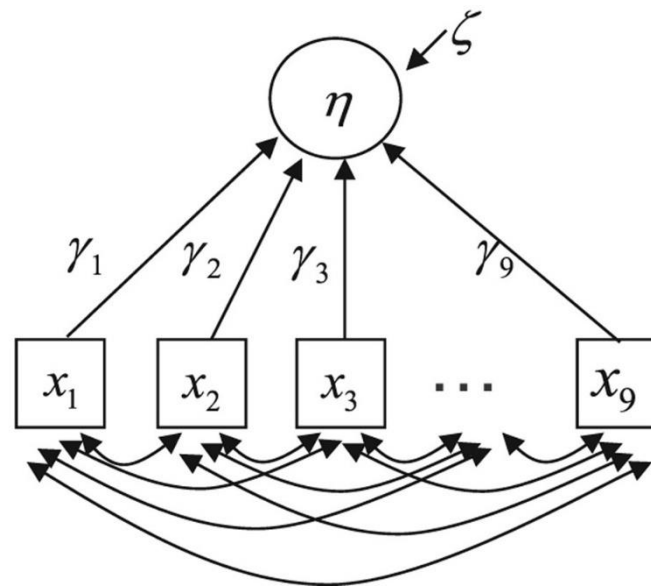
Tady jde spíše o teorie měření (které mohou obsahovat různé konkrétní modely).

Modely měření v psychologii

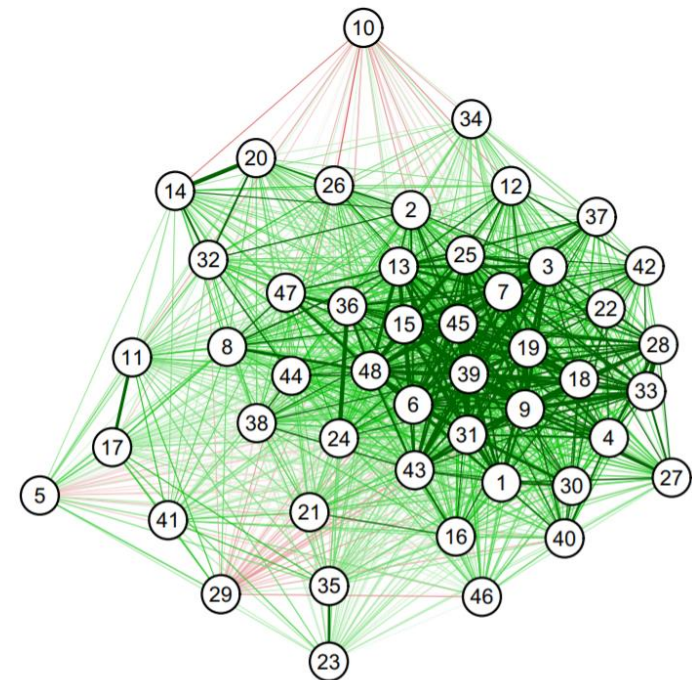
(a) Reflective measure model



(b) Formative measure model



síťový model



Modely měření v psychologii pragmaticky

Klasická testová teorie

- CTT, classical test theory, (classical/weak) true score theory.
- Rozšíření do faktorové analýzy.

Teorie zobecnitelnosti

- GT, generalizability theory; rozšíření CTT.

Teorie odpovědi na položku / Raschův model

- IRT, item response theory, (modern) latent trait theory, strong true score theory.
- Explanační IRT modely.

(Dynamic) Networks Models

Teorie vědomostních prostorů

- KST, Knowledge Space Theory.

Teorie spojitého měření

- CM, Conjoint Measurement.

Signal detection theory

- SDT; někdy též teorie zpracování informace

Další škálovací postupy

- Typicky propojené s formou sběru dat.
- Q-sort, sociometrie...

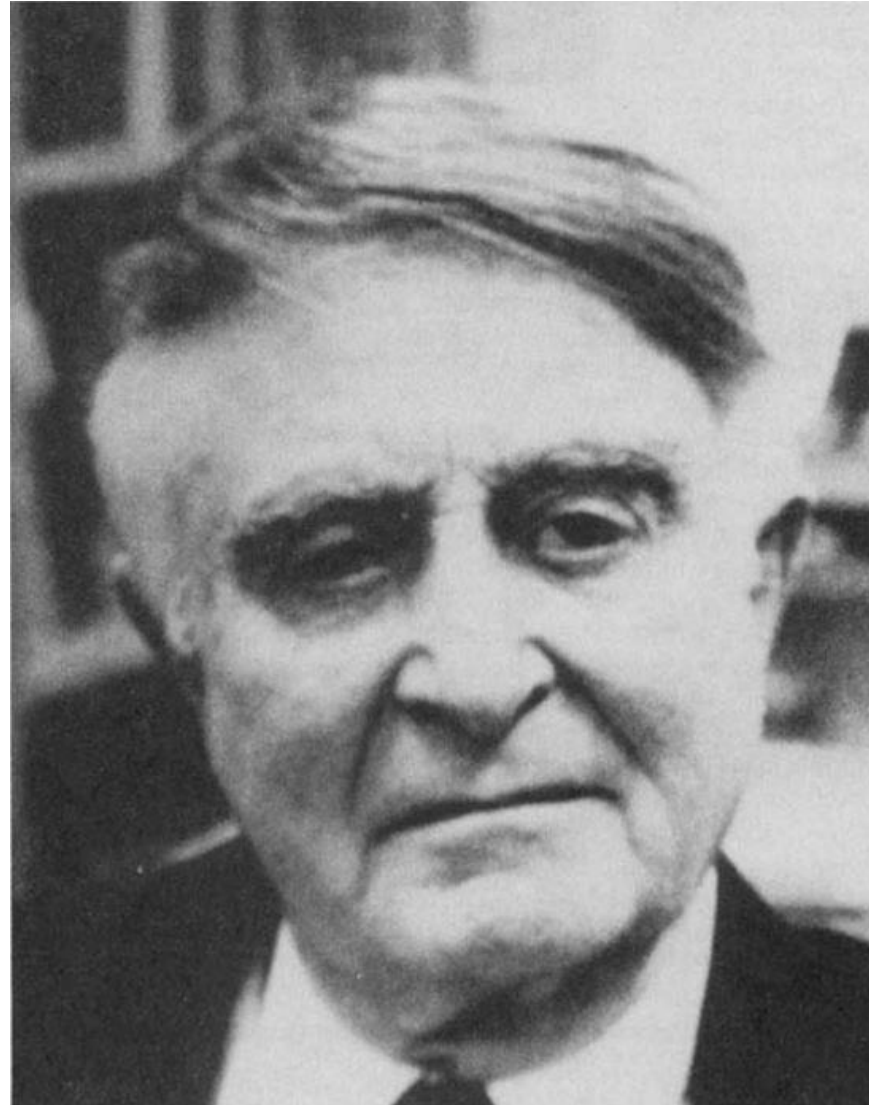
Tady jde spíše o konkrétní modely (které mohou spadat do stejných či různých teorií).

Klasická testová teorie

Spearman (1904)

...

Lord & Novick (1968)



CTT: Opakování

Měřeným atributem je pravý skór.

- Očekávaný skór člověka v daném testu.

Klíčovým konceptem je reliabilita.

- Založená na konceptu paralelních testů.

Operacionalismus, antirealismus.

- Měřený atribut je definován podobou měření.

Pravé skóre je zpravidla definované sumou/průměrem položek.

- Pozorované skóre je proto zpravidla součtem/průměrem pozorovaných odpovědí.

V praxi úzké propojení s faktorovou analýzou.

Nedostatky CTT

CTT založena na **operacionalismu**. **Antirealismus**.

- Nezabývá se existencí měřených atributů, není vhodná pro výzkum.
- V praxi nechceme měřit pravý skór (v daném testu), ale chceme měřit rys (nezávislý na měření).

Nejde o model *měření* v pravém slova smyslu.

- Nelze testovat jeho shodu s daty, lze použít na jakákoli data i neexistující atributy.
- Tohle částečně řeší propojení CTT s faktorovou analýzou.

Předpoklad lineárních, homoskedastických vztahů.

- CTT lze použít k inferenci na úrovni celkového skóre, nikoliv (správně) na úrovni ordinálních položek.

Nezvažuje povahu chyby měření. Chybu měření nelze rozdělit na jednotlivé „zdroje“.

- Např. ověřuji buď test-retest, nebo vnitřní konzistenci.

Sčítání položek vyžaduje paralelní, intervalové položky.

- Zkreslené odhady reliability, suboptimální validita a reliabilita skórů při porušení předpokladu.

Výhody CTT

Lze měřit cokoli.

- Z čistě operacionalistické pozice to nemusí vadit.
- CTT lze použít pro parametrizaci „měření“, i když předpokládáme jinou povahu měřeného atributu (například IRT).

Sčítání položek je jednoduché.

- Součet položek může sloužit jako aproximativní odhad skóre v jiné teorii měření (např. IRT).

Velmi jednoduché → snadno použitelné.

V řadě případů nám CTT parametrizace stačí.

Nedostatky CTT podruhé

Lokální nezávislost subjektů i položek.

Chyba měření je plně náhodná.

- Co když mají žáci v jedné třídě ale stejné „specifické znalosti“?
- Co když při test-retest odpovídám v osobnostním dotazníku stále stejně?

CTT pracuje vždy s jediným zdrojem chyby.

Teorie zobecnitelnosti

Generalizability Theory.

Cronbach, L.J., Nageswari, R.,
& Gleser, G.C. (1963).



Teorie zobecnitelnosti

„Zobecnění“ CTT pro všechny případy, kdy se měření skládá z dílčích paralelních testů.

Řeší problém „mnoho chyb měření, mnoho odhadů reliability“.

Univerzální teorie, která umožňuje parcelovat chybu měření na dílčí zdroje chyb.

Vychází z CTT: operacionalismus.

- Jde o její rozšíření se všemi neduhy a nevýhodami.

GT příklad: 2fasetový design

N osob je měřeno 3položkovým testem ve 2 situacích.

2fasetový design $p \times i \times o$:

- $n_i=3$ položky a $n_o=2$ administrace. Skóre je průměrem všech pozorování.

Odpověď na každou jednu položku se skládá z:

- $X_{pio} = T_p + e_i + e_o + e_{p \times i} + e_{p \times o} + e_{i \times o} + e_{p \times i \times o}$

Celkový rozptyl v datech:

- Rozptyl těchto jednotlivých pozorování X_{pio} .
- $\sigma_{X_{pio}}^2 = \sigma_p^2 + \sigma_i^2 + \sigma_o^2 + \sigma_{pi}^2 + \sigma_{po}^2 + \sigma_{io}^2 + \sigma_{pio,e}^2$
 - Pro připomenutí: sčítání náhodně rozdělených proměnných. Jednotlivé složky jsou nezávislé.

TABLE 36-1
Crossed Person \times Item \times Occasion G Study of Self-Concept Scores

	<i>Occasion</i>					
	<i>I</i>			<i>II</i>		
	<i>Item 1</i>	<i>Item 2</i>	<i>Item 3</i>	<i>Item 1</i>	<i>Item 2</i>	<i>Item 3</i>
<i>Person</i>						
1	4	2	5	4	3	4
2	3	1	4	4	2	3
3	2	3	3	3	2	4
...						
<i>p</i>	4	5	4	3	4	2
...						
<i>N</i>	3	4	4	3	3	3

GT příklad: 2fasetový design

N osob je měřeno 3položkovým testem ve 2 situacích.

2fasetový design $p \times i \times o$:

- $n_i=3$ položky a $n_o=2$ administrace. Skóre je průměrem všech pozorování.

Jinými slovy – chybami v tomto kontextu jsou:

- Každá **situace** je jinak obtížná.
- Každá **položka** je jinak obtížná (to většinou nezvažujeme).
- Výkon **respondentů** není stabilní (liší se v situacích).
- **Obtížnost** položek variuje napříč **situacemi**.
- Každá **položka** je jinak obtížná pro každého **respondenta**, systematicky napříč situacemi (korelace není perfektní).
- Zbytek (pro každého **respondenta** je každá **položka** v různých **situacích** různě obtížná.

TABLE 36-1
Crossed Person \times Item \times Occasion G Study of Self-Concept Scores

<i>Person</i>	<i>Occasion</i>					
	<i>I</i>			<i>II</i>		
	<i>Item 1</i>	<i>Item 2</i>	<i>Item 3</i>	<i>Item 1</i>	<i>Item 2</i>	<i>Item 3</i>
1	4	2	5	4	3	4
2	3	1	4	4	2	3
3	2	3	3	3	2	4
...						
<i>p</i>	4	5	4	3	4	2
...						
<i>N</i>	3	4	4	3	3	3

Teorie zobecnitelnosti

1. krok: G-studie (generalizability study)

- Rozdělení pozorovaného rozptylu na jednotlivé složky.
- Jak velkou část rozptylu měření tvoří rozptyl měřeného atributu, jak velkou část jednotlivé chyby?
- Zobecňuje z dat na prostor měření.

2. krok: D-studie (decision study)

- Odhaduje velikost chybového rozptylu pro konkrétní design měření.
- Případně hledá optimální design s nejlepším poměrem náročnosti měření a chyby měření (reliability).
- Specifikuje z prostoru měření na testovou situaci.

G-studie: Příklad výsledků

Table 36–2
Estimated Variance Components in the Example $p \times i \times o$ design

<i>Source</i>	<i>Variance Component</i>	<i>Estimate</i>	<i>Percent of Total Variability</i>
Person (p)	σ_p^2	1.108	30
Item (i)	σ_i^2	0.102	03
Occasion (o)	σ_o^2	0.030	01
$p \times i$	σ_{pi}^2	0.810	22
$p \times o$	σ_{po}^2	0.230	06
$i \times o$	σ_{io}^2	0.001	00
$p \times i \times o, e$	$\sigma_{pio,e}^2$	1.413	38

D-studie

Definuje tzv. „prostor zobecnění“ (počtem pozorování, mírou zobecnění atp.), na který budeme zobecňovat pozorování daného respondenta.

- V rámci tohoto prostoru má každý respondent tzv. U-skór (universe score; obecný skór).

Dva typy zobecnění:

- **Relativní** (norm-referenced) – pořadí osob v rámci stejné kombinace faset.
 - Reliabilita odhadována pomocí koeficient zobecnitelnosti (generalizability).
 - Přímo srovnatelný s různými druhy CTT reliability.
- **Absolutní** (kriteriální) – dosažení určité kriteriální hodnoty universe skóru.
 - Reliabilita odhadována pomocí koeficientu spolehlivosti (dependability)
 - Shodný počet, avšak jiných faset (např. 2 pozorovatelé, ale pro každého jiní).
 - Zajímá nás absolutní, kriteriální výsledek (např. u znalostních testů), nechceme jen srovnávat s normou.

D-studie: Příklad použití

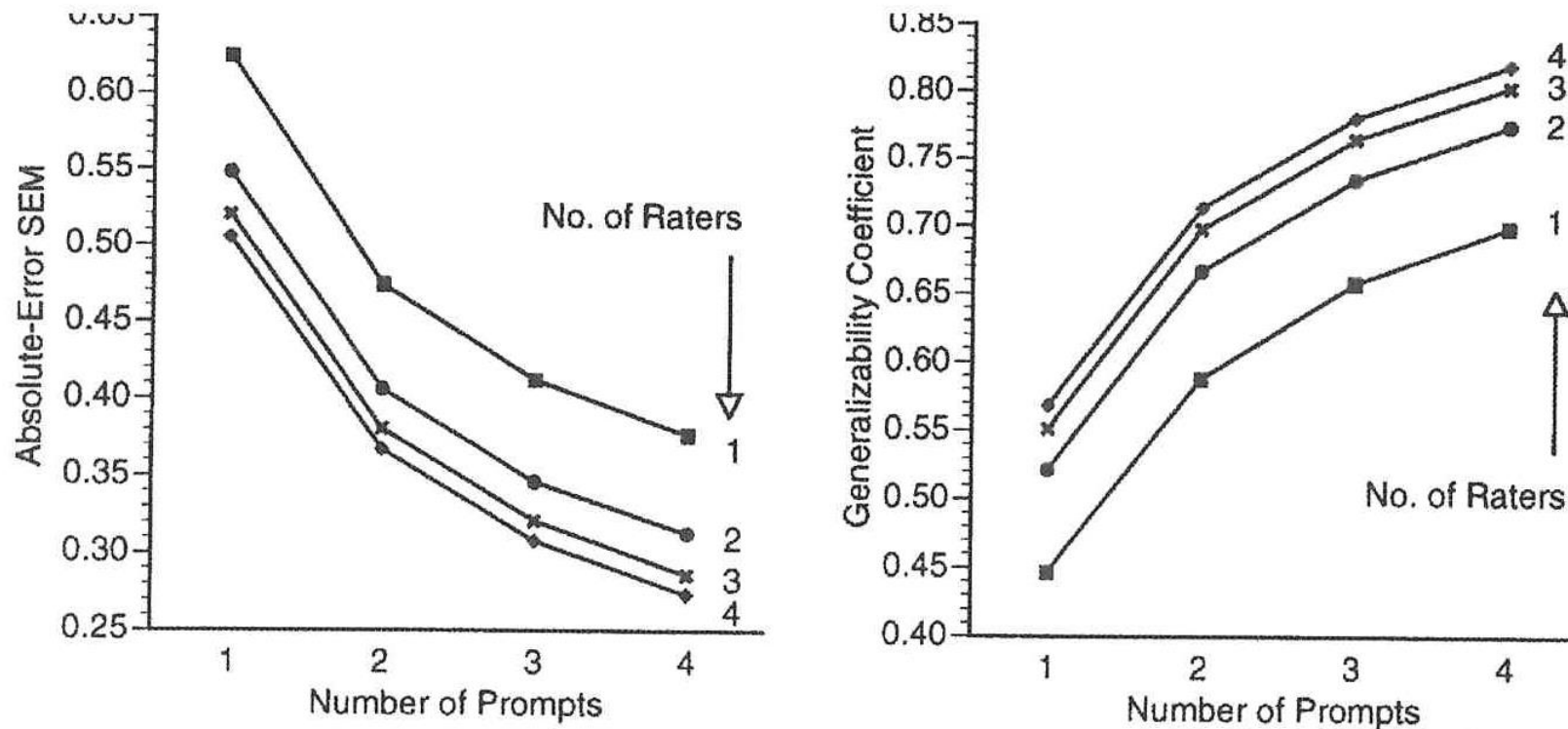


FIGURE 1.2. $\hat{\sigma}(\Delta)$ and $E\hat{\rho}^2$ for scenario with $p \times T \times R$ design.

Software a příklad

Lze použít ANOVA nebo smíšené lineární modely v běžném SW (SPSS, R, JASP...)

R: balíčky hamp, gtheory, lme4

- Balíček hamp není na CRANu (jde o balíček ke knize *Psychometrics in R*). Lze nainstalovat pomocí.

Příklad použití GT v češtině včetně R skriptu:

- Cígler, H., Ježek, S., Širůček, J., & Lacinová, L. (in press). Hodnocení bakalářských prací jako přijímací kritérium do navazujícího magisterského studia: Psychometrická kazuistika. *Studia Paedagogica*.
- preprint: <https://doi.org/10.31234/osf.io/dt7nr>

Teorie zobecnitelnosti: Shrnutí

Teorie zobecnitelnosti: „*Ultimátní*“ teorie „*měření*“ 😊

- Rozšíření a překonání klasické testové teorie.
- Využití mimo psychologii.
- Agregace většího množství chyb za předpokladu paralelního měření.

Předpoklady

- tau-ekvivalence/paralelnost položek.
- jednodimenzionalita (ale existují i multivariační aplikace).

Využití v psychologii:

- Vývoj testů (a odhad počtu položek/hodnotitelů...). Aka power analýza při vývoji metod.
- Vždy, když nemáme data výhradně jen pro vnitřní konzistenci.
- Princip parcelování rozptylu využitelný i v jiných modelech (FA, IRT).

Teorie odpovědi na položku

Item Response Theory

G. Rasch
F. M. Lord
P. F. Lazarsfeld

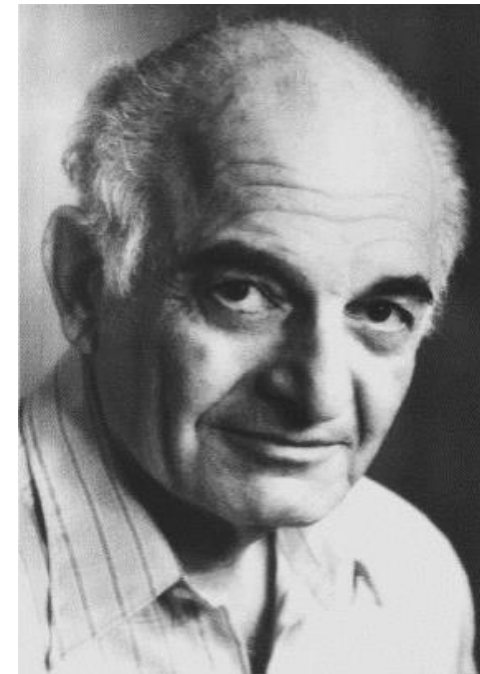
a mnoho dalších



Paul Felix Lazarsfeld (1901–19076)



Georg Rasch (1901-1980)



Louis Guttman
(1916–1987)

Teorie odpovědi na položku

Teorie odpovědi na položku (IRT; Item Response Theory) pomocí různých modelů popisuje vztah mezi latentním rysem (latent trait) a odpovědí na položky.

Odpovědi na položky jsou realisticky považovány za nominální či ordinální.

- V případě intervalové odpovědi může být IRT totožná s faktorovou analýzou.

IRT modely: popisují pravděpodobnost správné (nějaké) odpovědi respondenta s určitou mírou latentního rysu na položku o určité obtížnosti (a dalších parametrech).

Vztah odpovědí a latentního rysu není lineární jako v případě FA.

- Protože samy položky (většinou) nejsou považovány za intervalové.
- Zpravidla je logistický (IRT je vlastně série logistických regresí).

Faktorová analýza vs. IRT

Technicky vzato je faktorová analýza a IRT velmi podobný model měření.

- Viz Borsboom (2005) a modely s latentními proměnnými.
- Dalším modelem s latentními proměnnými je například analýza latentních tříd (LTA).

Faktorová analýza: model pro spojitá intervalová data.

- Vztah faktoru a odpovědi na položku je lineární.
- Odhad faktorového skóru lze provést jako vážený průměr položek.

$$X = \lambda\theta + \tau + \epsilon$$

IRT: model pro libovolná, typicky kategorická (nominální, ordinální) data.

- Vztah faktoru a odpovědi na položku je logistický (příp. probitový).

$$\ln \frac{P(x = 1)}{P(x = 0)} = a\theta + b$$

Někdy též „**item-factor analysis**“ – kombinace IRT přístupu a *ordinální faktorové analýzy*.

- Tzv. „ordinální/kategorická“ CFA je zcela shodná s IRT Graded Response Modelem s probitovým linkem.

Protože data mohou být velmi různá, IRT je vysoce flexibilní nástroj.

IRT modely

Vznik v 50. a 60. letech, další rozvoj v 80. letech (počítače).

Nezávisle na sobě G. Rasch (matematik), F. M. Lord (psycholog) a P. F. Lazarsfeld (sociolog).

Jde o stochastickou úpravu původně deterministického Guttmanova modelu.

Základní typy modelů podle počtu parametrů:

- 1 – parametr obtížnosti položky
- 2 – diskriminační parametr položky
- 3 – parametr (pseudo)uhádnutelnosti

Raschovy modely:

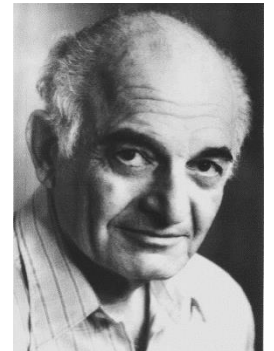
- Řadí se mezi jednoparametrové, ale mají odlišnou logiku.
- (Složitější RM navíc svým způsobem zahrnují i jiné parametry.)



Paul Felix Lazarsfeld
(1901–1976)



Georg Rasch (1901–1980)



Louis Guttman
(1916–1987)

Charakteristická funkce položky Item Characteristic Curve (ICC)

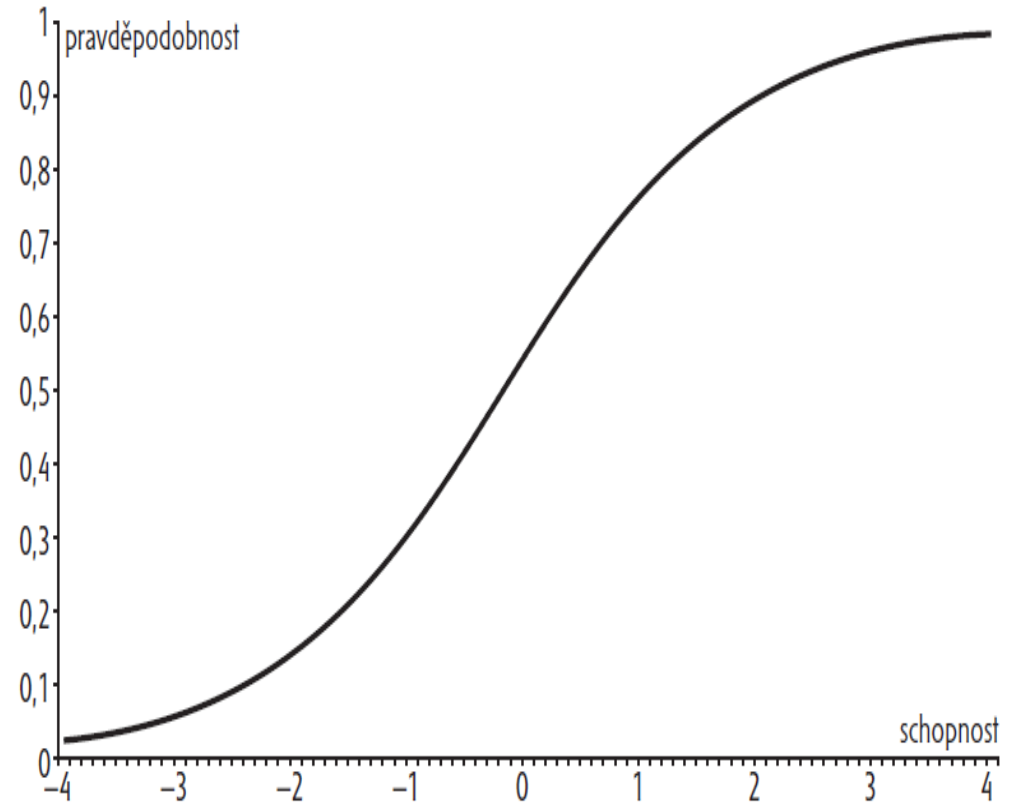
Operacionalizace daného IRT modelu.

- Jednotlivé IRT modely se liší právě svou ICC.

Obtížnost položek a schopnost respondenta jsou na stejné škále.

Charakteristická funkce položky:

- Má zpravidla přibližně tvar kumulativního normálního rozložení.
- Popisuje vztah mezi schopností probanda a jeho očekávaným výkonem v dané položce.
- Predikuje pravděpodobnost správné/špatné odpovědi v závislosti na obtížnosti položky a probandově schopnosti.



Raschův (1PL) model

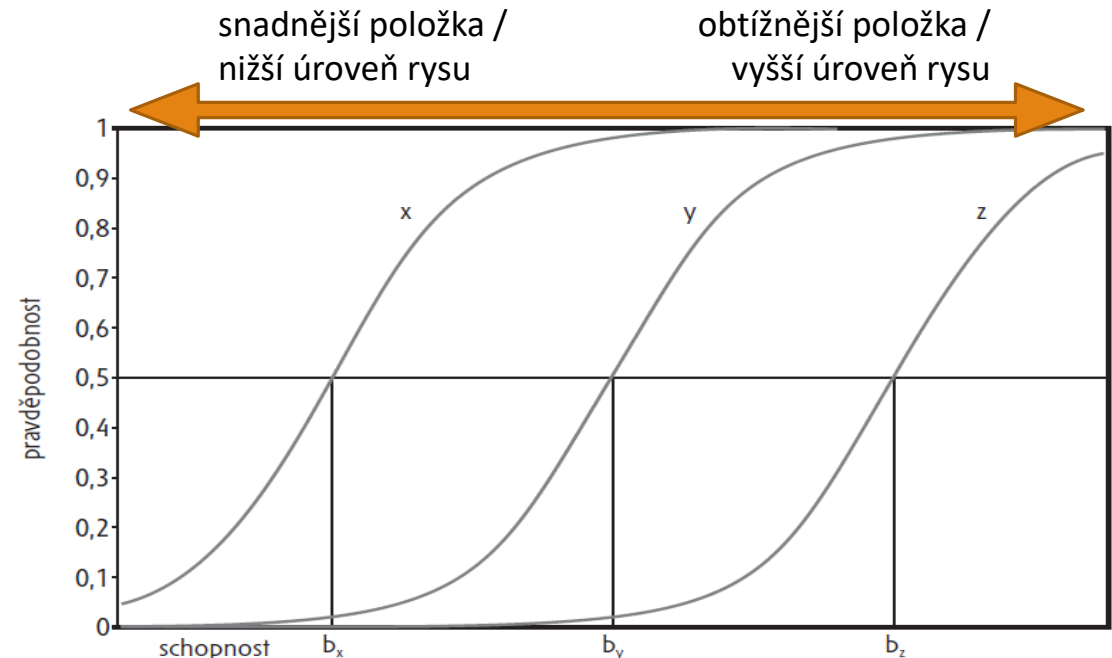
Charakteristická funkce:

$$P_i(\theta) = \frac{e^{\theta - b_i}}{1 + e^{\theta - b_i}}$$

Přímé rozšíření guttmanovské škály o pravděpodobností usuzování.

Parametr obtížnosti b_i položky i je bod na škále schopnosti, v němž je pravděpodobnost správné odpovědi respondenta j se stejnou mírou schopnosti ($\theta_j = b_i$) na danou položku $P_i(\theta_j) = 0,5$.

- Pravděpodobnost správné odpovědi závisí na rozdílu schopnosti respondenta a obtížnosti položky $\theta - b_i$.



Další modely

1PL – technicky shodný s Raschovým modelem.

- (To „PL“ značí počet parametrů jedné položky).

2PL – jak dobře model diskriminuje?

- Strmost ICC prostřednictvím diskriminačního parametru.

3PL – jakou pravděpodobnost uhádnutí má respondent s nekonečně malou úrovní daného rysu?

- Spodní limita funkce pomocí parametru pseudouhádnutelosti.

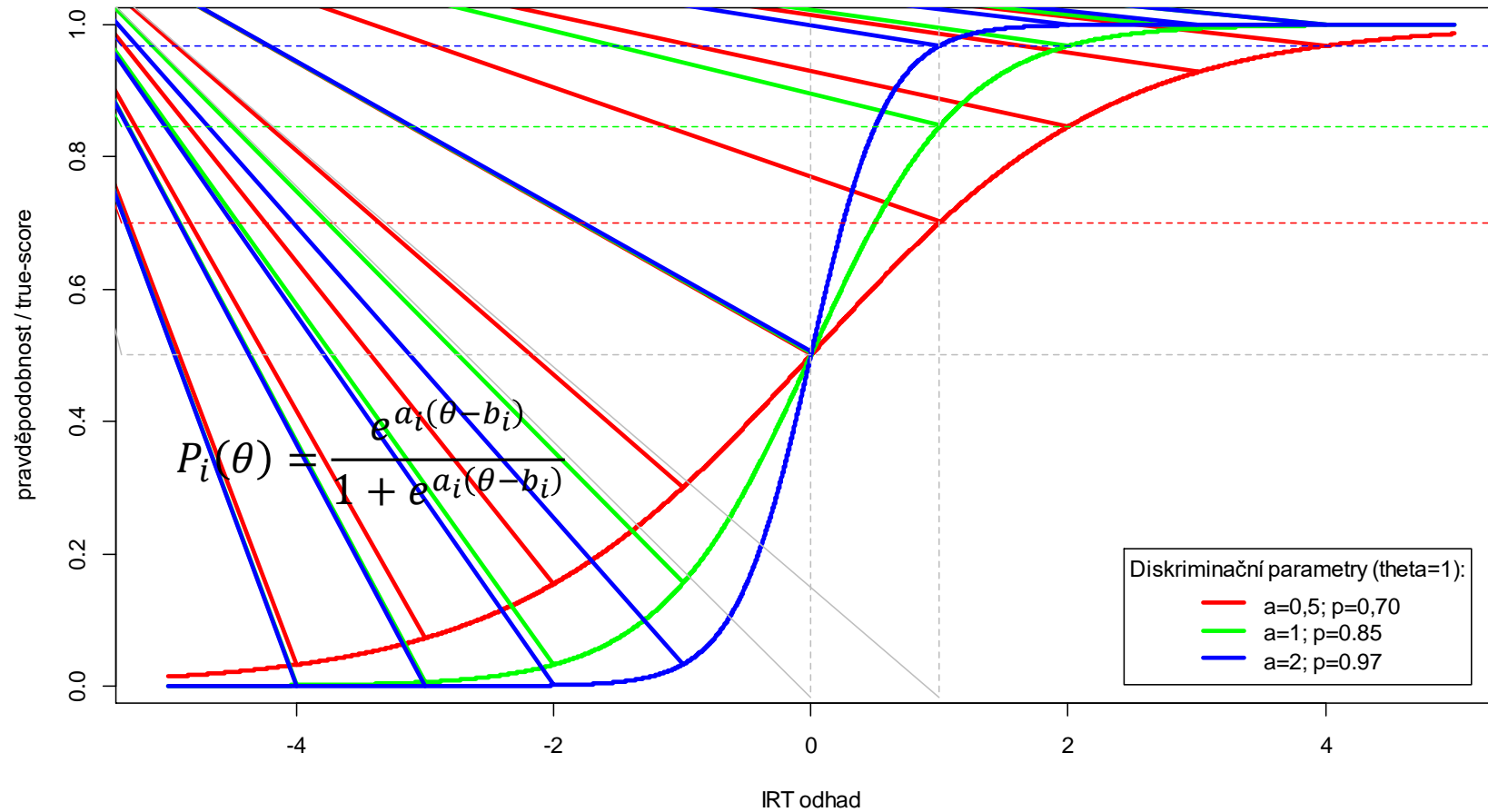
Různé polytomické modely (Likertovy škály, multiple-choice, nucená volba...).

- Partial Credit model, Nominal Response model, Graded Response Model a řada dalších.

Multidimenzionální modely.

- Pro případy, kdy odpověď závisí na více latentních rysech.

Charakteristická křivka položky 2PL modelu



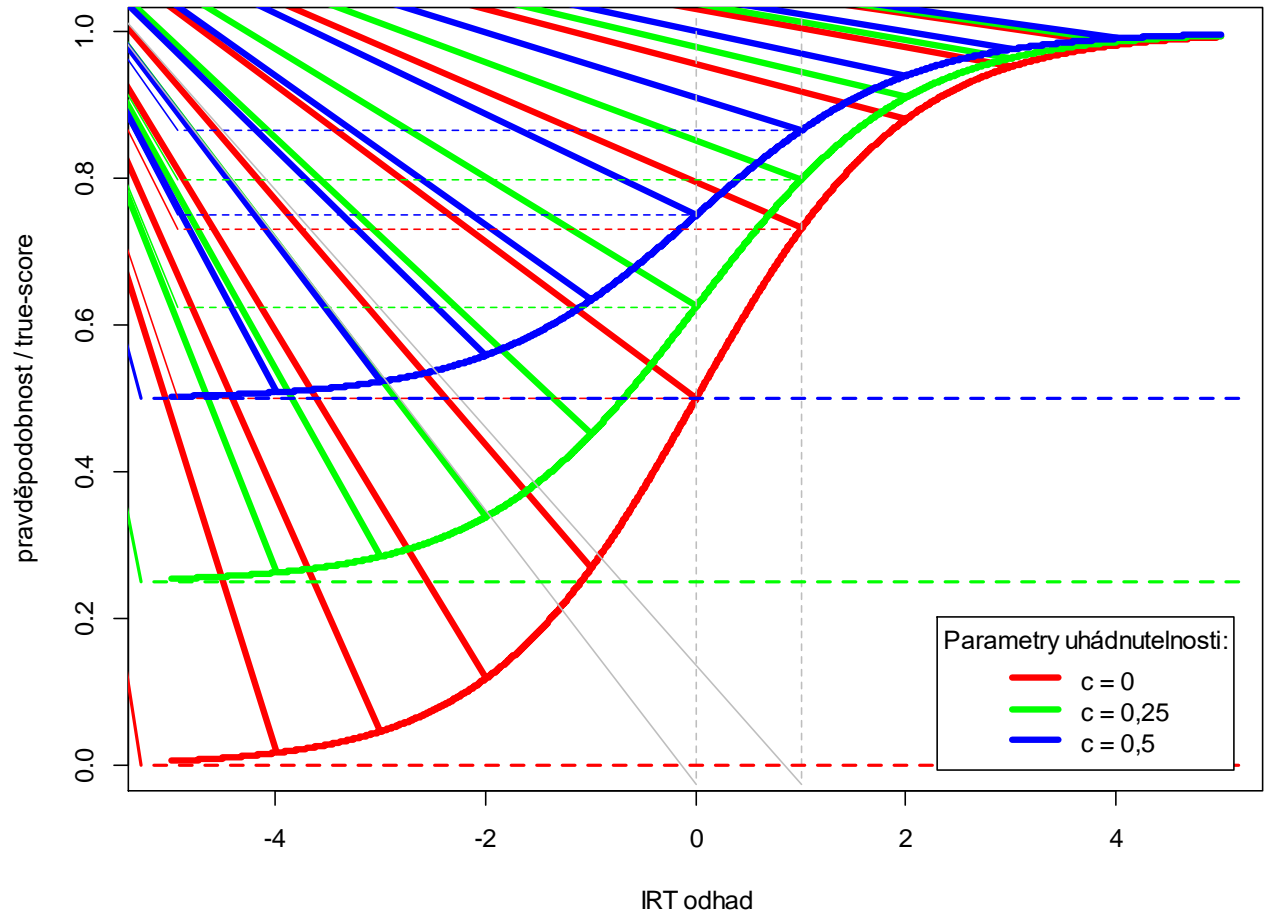
Charakteristická křivka položky 3PL modelu

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}}$$

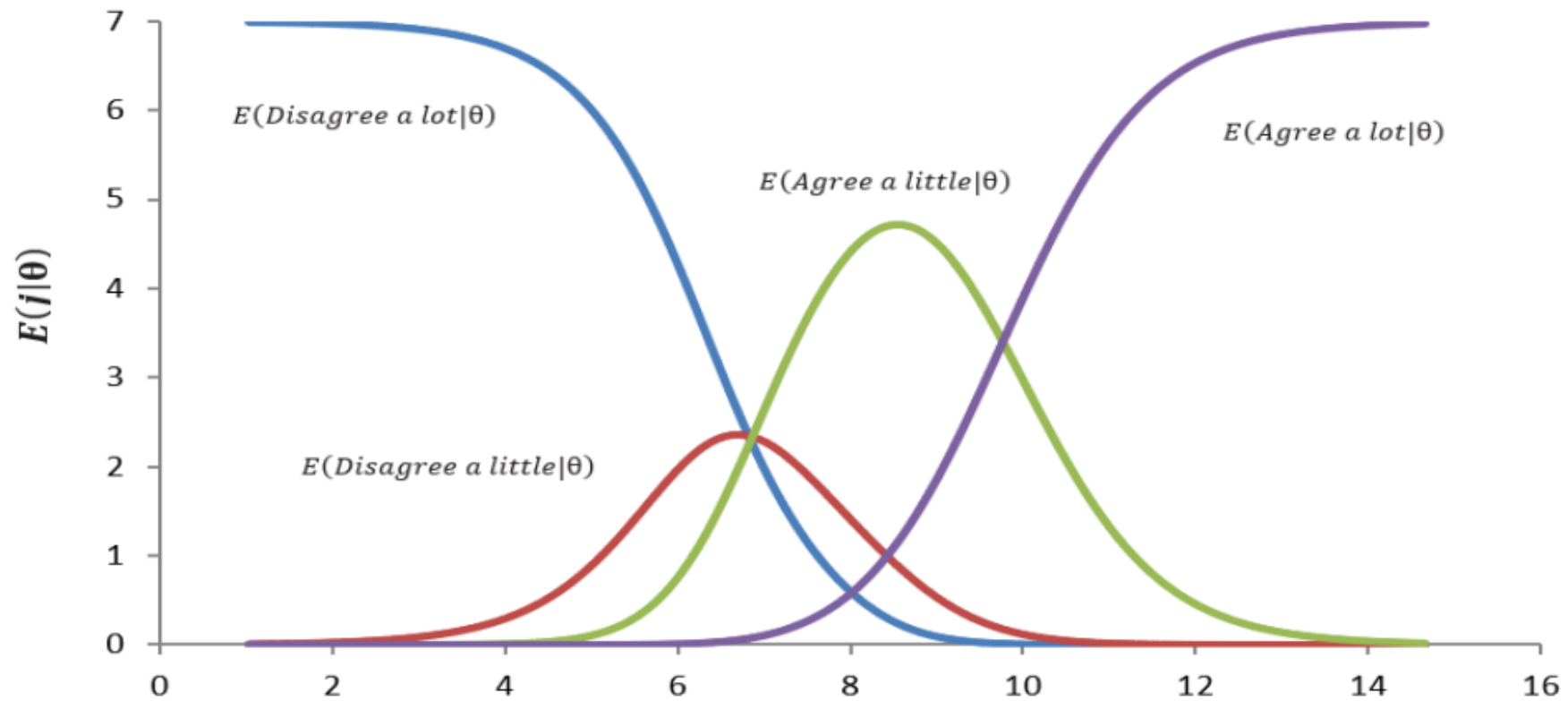
- a_i – diskriminační parametr
 - Míra náklonu ICC.
- c_i – parametr pseudouhádnutelnosti
 - Spodní limita ICC.

Postupnou fixací parametrů 3PL modelu lze dosáhnout jednodušších modelů:

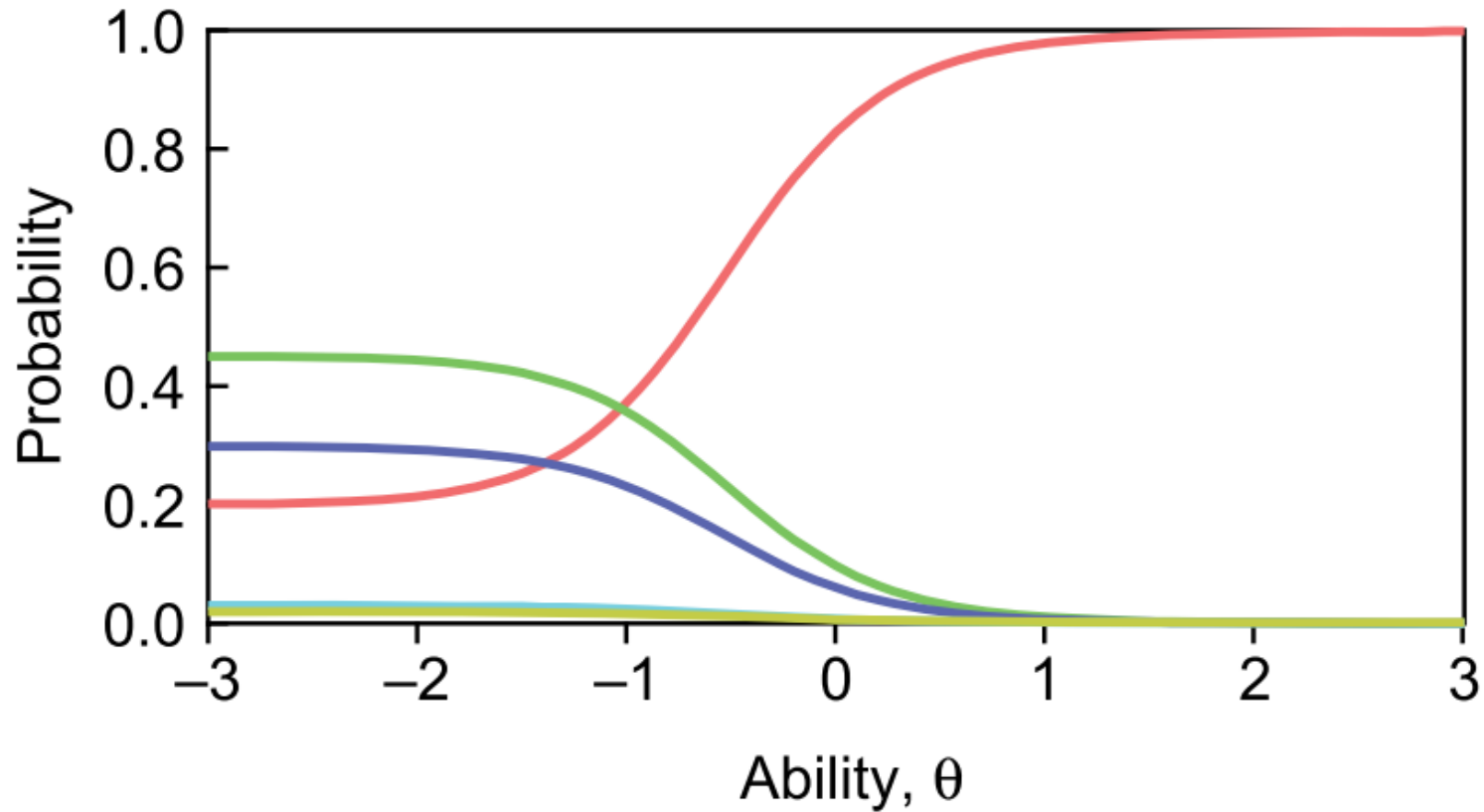
- $c_i = 0$ – 2PL model
- $a_i = 1$ – 1PL model



Polytomické modely: př. Graded Response Model



Polytomické modely: Multiple-choice Model



Praktické příklady

<http://fssvm6.fss.muni.cz/ICC/>

- 4PL binární model

<https://shiny.cs.cas.cz/ShinyItemAnalysis/>

- 4PL binární model: záložka IRT models → Training → Dichotomous model
- Polytomické modely (GRM, GPCM, NRM): záložka ... → Polytomous model

Typické použití IRT

1. Použití namísto faktorové analýzy (EFA i CFA) pro ověření faktorové struktury.
 - Odhad parametrů položek a shody modelu s daty.
 - V případě krátké odpověďové škály (zejm. binární) realističtější výsledky.
 - *IRT jako nástroj pro validizaci.*
2. Odhad latentních rysů (faktorových skóre) respondentů ve stávajícím vzorku.
 - Namísto součtu položek – stejná data, jako v bodě 1.
 - *IRT jako nástroj pro vytvoření skóre.*
3. Odhad latentních rysů nových respondentů.
 - Předem odhadnuté parametry položek (v bodě 1) lze použít pro odhad latentního rysu libovolného nového respondenta.
 - *IRT jako nástroj pro skórování testu v praxi.*

Výhody IRT a srovnání s CTT

Realistické modelování odpověďového procesu.

Měření je (velmi) nezávislé na konkrétním testu a výzkumném souboru.

Chyba měření je specifická pro různé úrovně latentního rysu.

Explorační i konfirmační podoba modelů.

- Jako u faktorové analýzy; tzv. item-factor analysis.

Chybějící data nejsou problém (full-information)

- IRT se zaměřuje na položku, nikoliv celý test jako CTT.

Poskytuje více informací o fungování testu ve srovnání s FA/CTT.

Občas náročnější na velikost vzorku a na výpočetní sílu (dimenzionalita).

Typické aplikace IRT do praxe

Počítačové adaptivní testování: Vybírá položky podle průběhu testování.

- Zvýšení reliability, resp. zkrácení testu.

Vyvažování paralelních verzí testu (large-scale assessment), kalibrace.

- Anchor tests/items, předkalibrované sety položek atd.

Různé výhody při **konstrukci psychodiagnostických testů.**

- Nižší závislost skóru na obtížnosti testu, speciální typy skóru, plánovaně chybějící data.
- V ČR např. WJ-IV, KIT či některé části Vídeňského testového systému.

Další

- Např. explanační IRT modely ve výzkumu pro výzkum odpověďových procesů.

Dostupný software

R

- Zejm. balíček **mirt**, případně **lavaan**, TAM, irtoys, eRm, psychometrics aj.

Winsteps

- Pouze Raschovy modely, zaměřený na praktickou konstrukci testů (WJ-IV aj.).
- Grafické GUI; HC může poskytnout licenci.

Mplus

- Obecný SW pro modelování s latentními proměnnými.
- Katedra má malý počet licencí.

Klasika a tradice: BILOG-MG, flexMIRT, jMetrik, PARSCALE, Xcalibre aj.

[ShinyItemAnalysis](#): [on-line aplikace](#), nebo [package](#) do R s vlastním GUI.

Reprezentační model měření: Spojité měření

Měření v sociálních vědách může být měřením.

Od ordinality ke kardinalitě.

Luce a Tukey, Debreu.



Vlevo John Tukey, vpravo Gérard Debreu.

Spojité měření

Nezávisle na sobě vyvinuli francouzský ekonom Gérard Debreu (1960) a psycholog Duncan Luce s matematikem Johnem Tukey (1964).

Conjoint measurement theory (CM) je postup pro inferenci intervalové proměnné s aditivní strukturou z ordinálních proměnných.

- A tedy vyvrací závěry Fergusonovy komise.
- Raschův model může být jednou ze stochastických specifikací CM.

V současnosti určitý rozvoj v oblasti dalších stochastických aplikací pro různé účely, celkově ale není příliš využíván kvůli značné matematické náročnosti.

Zajímavost: Tversky (Nobelova cena Kahneman a Tversky, 1979) se zaměřoval právě na CM (např. [1967](#)) a jejich *prospektivá teorie* je na CM přímo založena.

CM: Klíčové pojmy a koncepty

Původně byla teorie plně deterministická, stochastické úpravy jsou komplikované.

- Je tedy problém s aplikací v sociálních vědách, kde hraje velkou roli chyba měření.

Typicky máme dvě proměnné, u kterých nevíme, zda jsou nominální či kardinální.

- V typickém případě v psychologii položky a osoby.

Dále máme množinu všech možných seřazení těchto proměnných.

- Může být seřazená (ordinální), nebo číselná (intervalová).

CM podle několika pravidel odvodí, zda jsou proměnné intervalové.

- A my pozorujeme jen jejich konkrétní „instance“.
- Jednoduché a dvojité vykrácení, řešitelnost, archimédovská podmínka, posloupnost vykrácení (tranzitivita rozdílů).

Spojité měření



Představte si, že na zemi přiletí mimozemšťani, kteří posuzují pokročilost civilizací podle kvality měřicích nástrojů.

Ukradnou proto několik předmětů denní potřeby a v obchodě si koupí nástroje pro jejich „měření“.

Mimozemšťani nevidí a jsou celkově divní, takže neví, zda měřidla měří intervalovově.

Jak to ověří?

Spojité měření

Délka.

- Položky: $m > cm > mm$
- Respondenti: $stůl > kniha > tužka$

Jsou tužka-kniha-stůl a mm-cm-m pravítko kvantitami?

Jednoduché vykrácení

- $(tužka, mm) < (kniha, mm)$
- $(tužka, mm) > (tužka, cm)$

Dvojité vykrácení

- $(tužka, cm) < (kniha, mm)$
 - $(kniha, mm)/(tužka, cm) = 300/15 = 20$
- $(kniha, m) < (stůl, cm)$
 - $(stůl, cm)/(kniha, m) = 150/0,3 = 500$

interakce

položky

respondenti

	mm	cm	m
tužka	150	15	0,15
kniha	300	30	0,3
stůl	1500	150	1,5

- $(tužka, m) < (stůl, mm)$
- $(stůl, mm)/(tužka, m) = 1500/0,15 = 10000$

→ **tranzitivita: tužka+cm+kniha+m < kniha+mm+stůl+cm**

- $tužka+m < stůl+mm$
- $20 \cdot 500 = 10000$

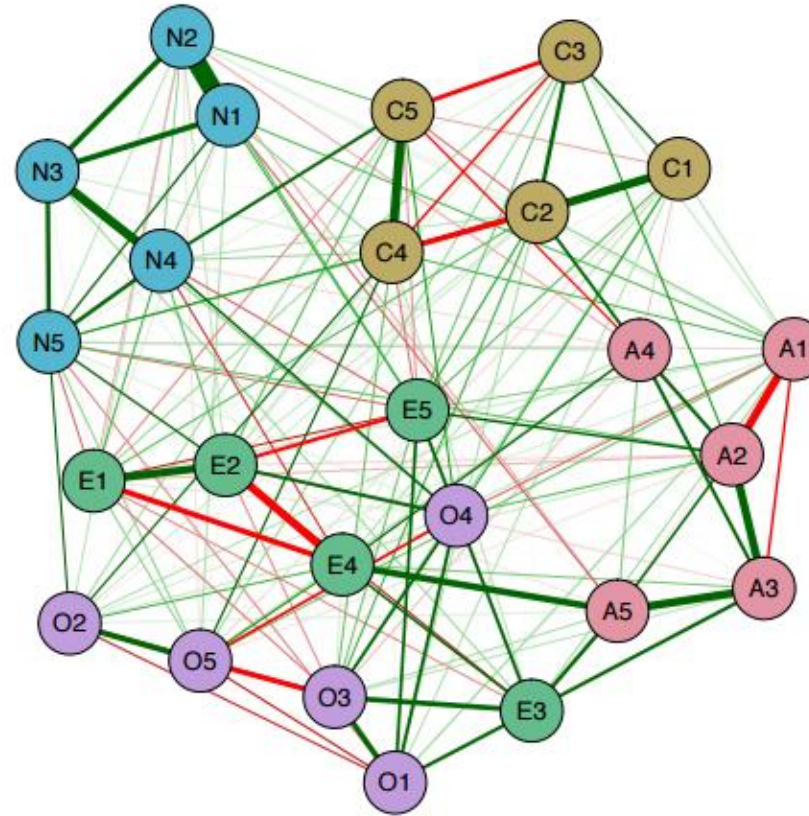
Fundamentální škála: tužka=1, kniha=2, stůl=10
m=1, cm=100, mm=1000

Dynamické síťové modely

Dynamic network models

Van Der Maas et al., 2006; Borsboom,
2008; Cramer, Waldorp, van der Maas,
& Borsboom, 2010;

Borsboom, Cramer, Schmittmann,
Epskamp, & Waldorp, 2011



Agreeableness

- A1: Am indifferent to the feelings of others.
- A2: Inquire about others' well-being.
- A3: Know how to comfort others.
- A4: Love children.
- A5: Make people feel at ease.

Conscientiousness

- C1: Am exacting in my work.
- C2: Continue until everything is perfect.
- C3: Do things according to a plan.
- C4: Do things in a half-way manner.
- C5: Waste my time.

Extraversion

- E1: Don't talk a lot.
- E2: Find it difficult to approach others.
- E3: Know how to captivate people.
- E4: Make friends easily.
- E5: Take charge.

Neuroticism

- N1: Get angry easily.
- N2: Get irritated easily.
- N3: Have frequent mood swings.
- N4: Often feel blue.
- N5: Panic easily.

Openness

- O1: Am full of ideas.
- O2: Avoid difficult reading material.
- O3: Carry the conversation to a higher level.
- O4: Spend time reflecting on things.
- O5: Will not probe deeply into a subject.

Dynamic network model

„We propose that the variables that are typically taken to be indicators of latent variables should be taken to be autonomous causal entities in a network of dynamical systems. Instead of positing a latent variable, one assumes a network of directly related causal entities as a result of which one avoids the three problems discussed above.“

- Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 31(1). doi: [10.1016/j.newideapsych.2011.02.007](https://doi.org/10.1016/j.newideapsych.2011.02.007)

Zejména v psychopatologii či psychologii osobnosti.

V současnosti extrémně rychle se rozvíjející oblast psychometriky; jen pár let staré.

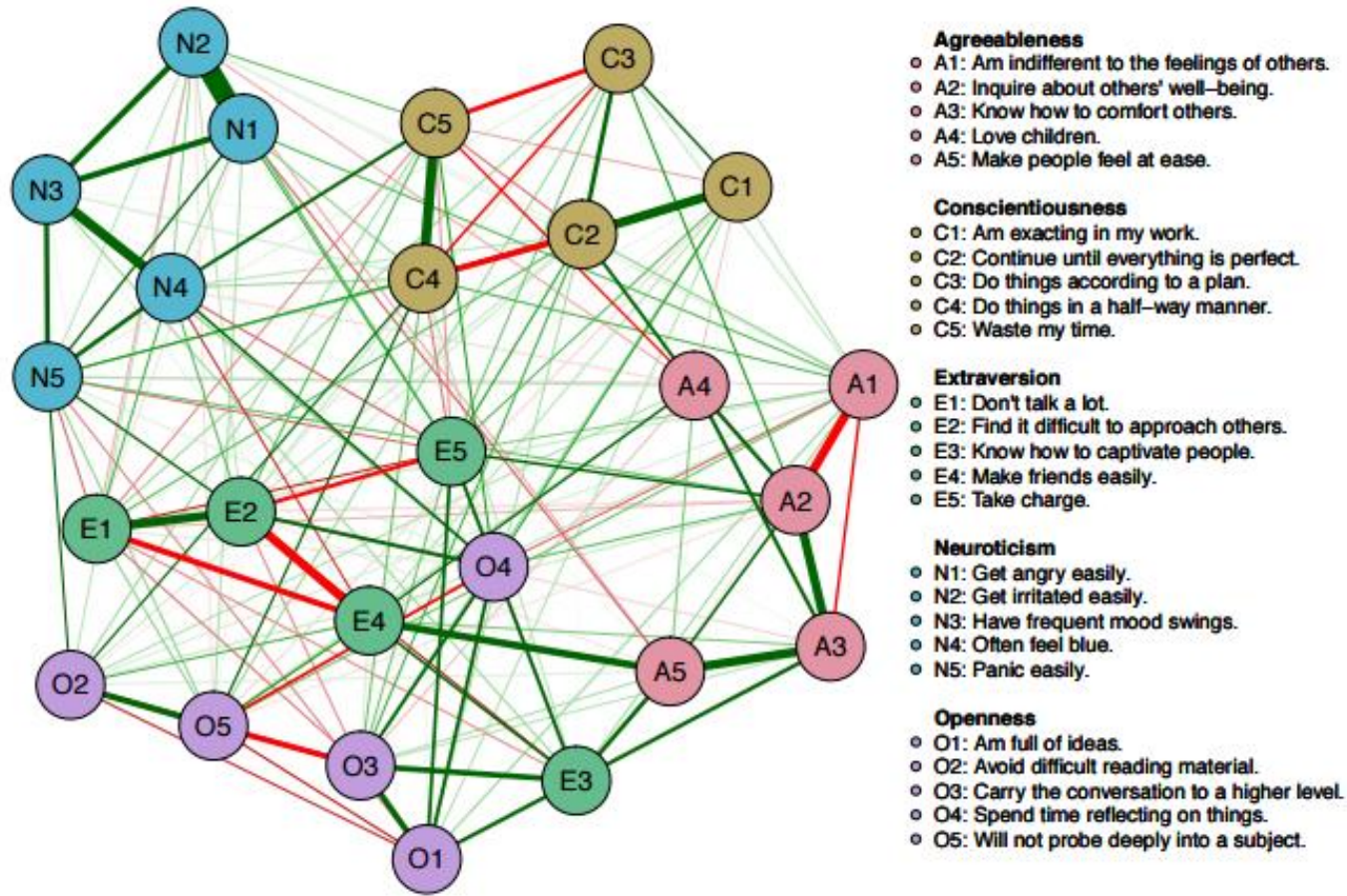


Figure 1.1: Example of a network model estimated the BFI personality dataset from the psych package in R. Nodes represent variables (in this case, personality inventory items) and links between the nodes (also termed *edges*) represent partial correlation coefficients. Green edges indicate positive partial correlations, red edges indicate negative partial correlations, and the width and saturation of an edge corresponds to the absolute value of the partial correlation. Estimation technique as outlined in Chapter 2 was used.

[Epskamp, S. \(2016\). Regularized Gaussian Psychological Networks: Brief Report on the Performance of Extended BIC Model Selection. arXiv preprint, arXiv:1606.05771.](#)

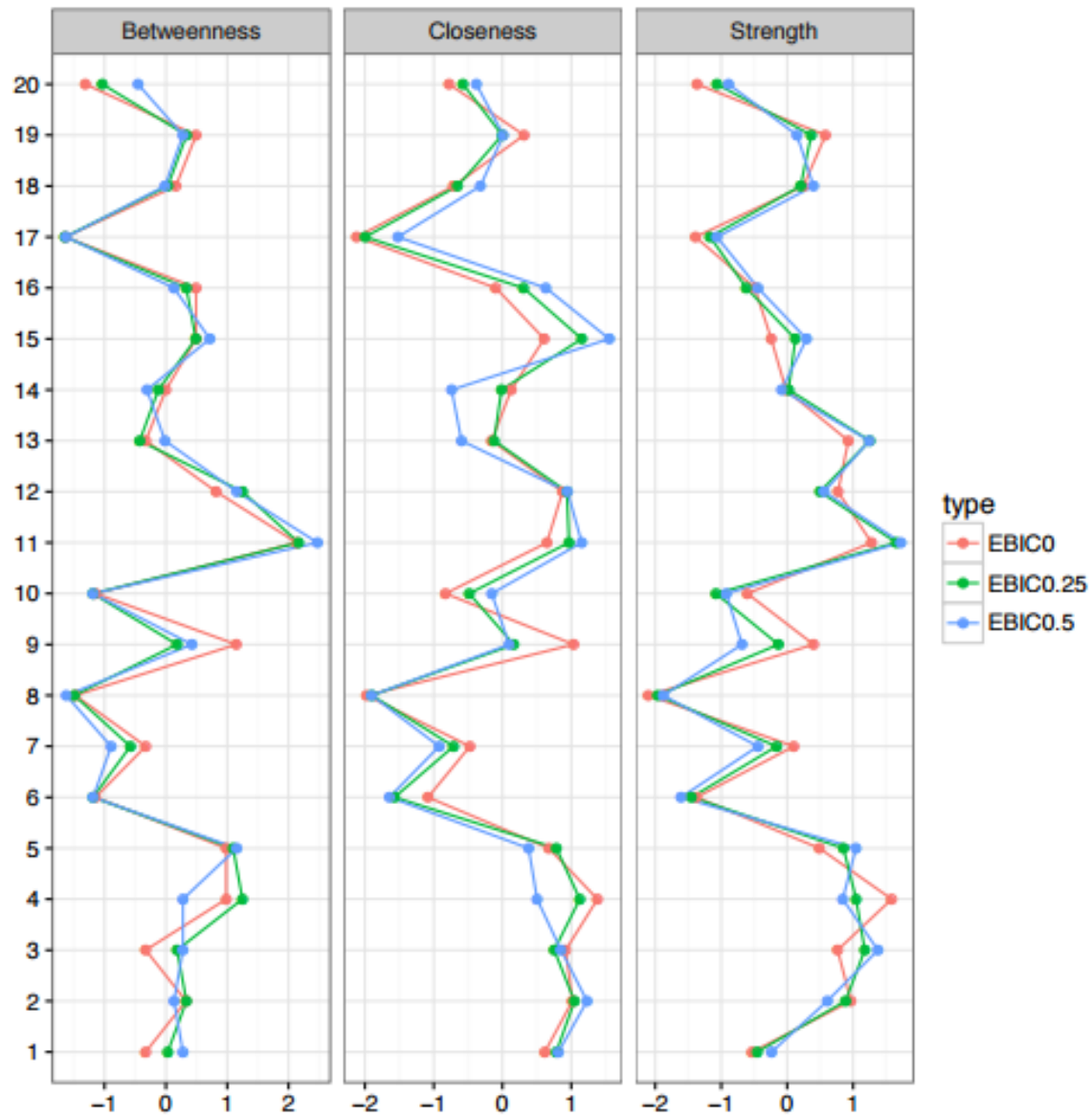


Figure 2.4: Closeness, betweenness, and degree centrality of the three networks described in Figure 2.3 with increasing levels of the LASSO hyperparameter γ . Values are standardized to z-scores.

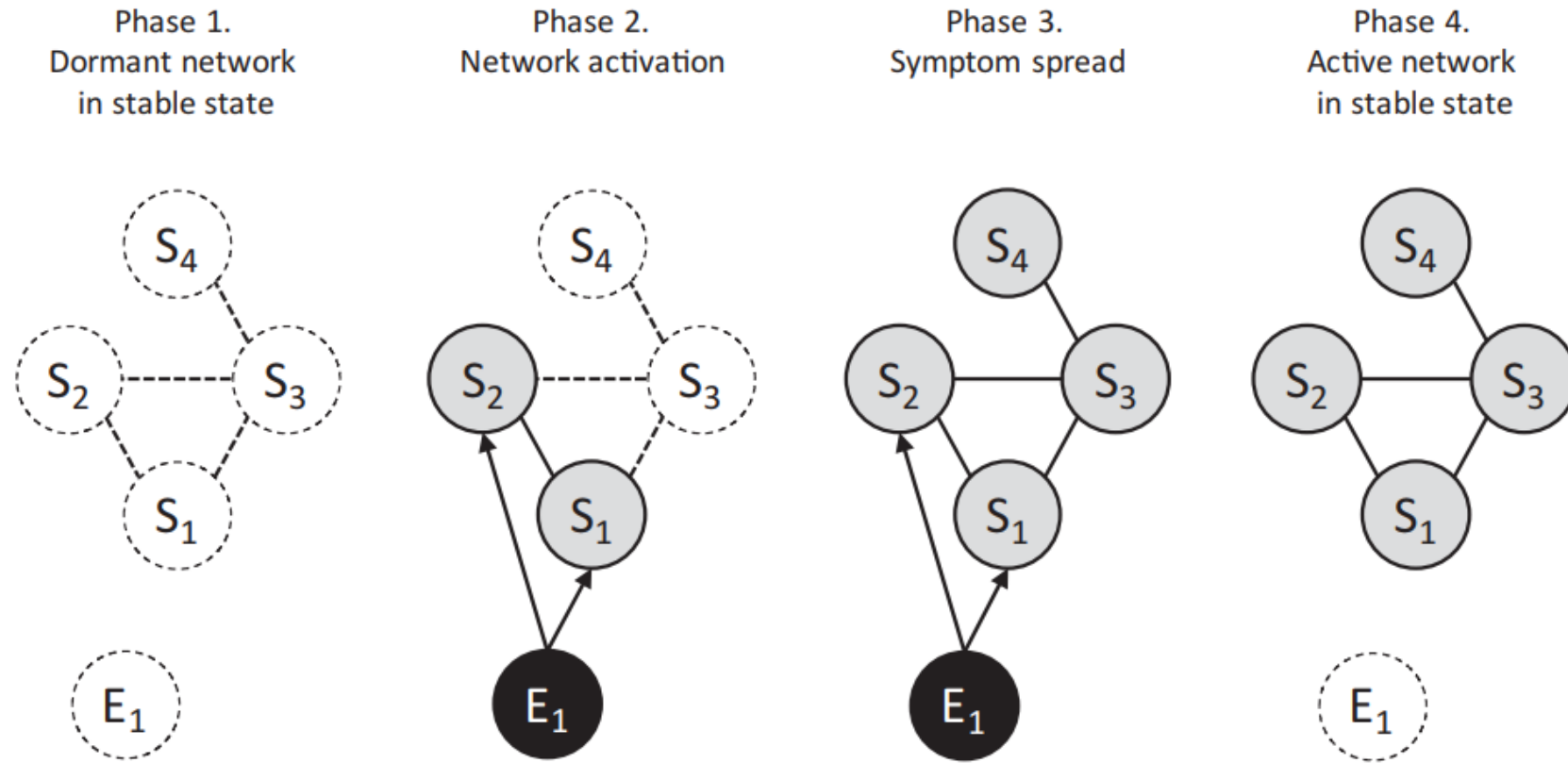
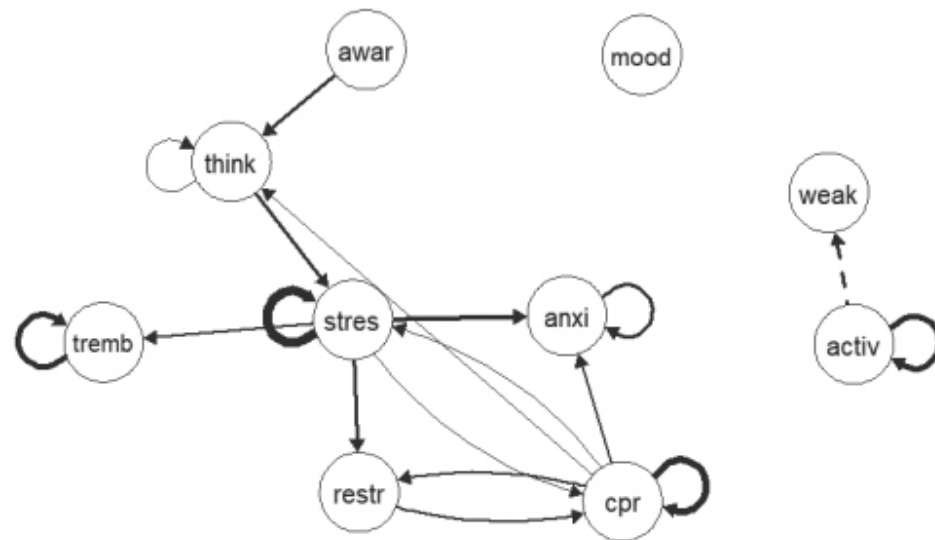


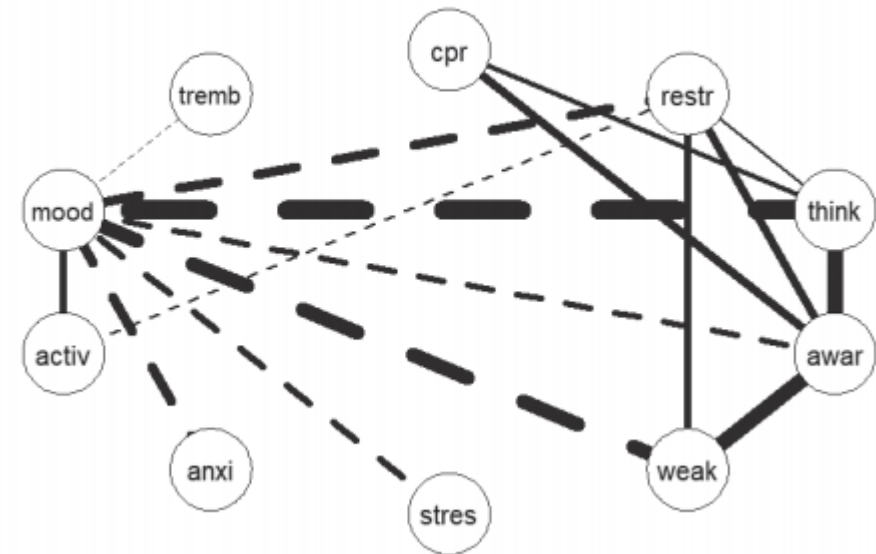
Figure 3 Phases in the development of mental disorders according to the network theory. After an asymptomatic phase, in which the network is dormant (Phase 1), an external event (E₁) activates some of the symptoms (Phase 2), which in turn activate connected symptoms (Phase 3). If the network is strongly connected, removal of the external event does not lead to recovery: the network is self-sustaining and is stuck in its active state (Phase 4).

Deníčková studie (medicínsky nevysvětlitelné symptomy)

(a) Temporální síť



(b) Simultánní síť



P1 byl muž ve věku 42 let s následujícími symptomy: úzkost, nervozita, bušení srdce, ztuhlost, „výboje“ na šíji, rozostřené vidění, náhlá ztráta energie a bolesti nohou, otoky kloubů (diagnostikována revmatoidní artritida, nyní v remisi). Za hlavní spouštěč potíží byla v anamnéze označena pracovní zátěž (stresující a špatně organizovaná práce v jeho tehdejší zaměstnání).

Nováček, T., Řiháček, T., & Cígler, H. (2020). Personalizovaný síťový model pro medicínsky nevysvětlené tělesné symptomy. *Československá psychologie*, 64(3), s. 257-271.

Plná čára vyjadřuje pozitivní vztah, přesušovaná čára negativní vztah; *activ* = právě se věnuji aktivitě, která mě naplňuje, *anxi* = mám z něčeho obavy, *awar* = právě v tomto okamžiku si všímám, kde v těle co cítím, *cpr* = mám pocit tlaku v oblasti hrudníku, *mood* = celková nálada, *restr* = moje tělesné obtíže mi zabraňují dělat to, co chci, *stres* = jsem v napětí, *think* = aktuálně myslím na své tělesné obtíže, *tremb* = třesou se mi ruce/nohy, *weak* = cítím se zesláblý.

<https://goo.gl/6cJs6a>

Wichers, M., Groot, P. C. (2016). Critical Slowing Down as a Personalized Early Warning Signal for Depression. *Psychotherapy and Psychosomatics* 85, 114-116.
doi: [10.1159/000441458](https://doi.org/10.1159/000441458)

Dynamic network models: Využití

Aktuálně je propojován přístup strukturního modelování a dynamických systémů – část modelu je parametrizována jako síť, část jako strukturní model.

Time-series modely.

- Zároveň zjišťují „souvýskyt“ symptomů v jednom časovém okně (contemporaneous network) a sledují, jak na sebe kauzálně symptomy působí v čase (temporal network).

Personalizovaný mixed/random effect network model.

- V případě kauzálního time-series modelu: existují „průměrné“ hodnoty kauzálních vztahů mezi symptomy, které ale pro různé lidi různě variují (každý má vztahy trochu jiné).
- Přímé využití v terapii, např. při identifikaci fokálních symptomů.

Network modely podporují systemický přístup k diagnózám, umožňují vhodnějším způsobem studovat patologii (aplikace zejm. v oblasti deprese, úzkosti apod.).

Blízké sociometrii, analogické k analýze sítí v sociologii.

Software

JASP

R. Balíčky:

- psychometrics
- bootnet
- qgraph

Další zdroje:

- <http://psychosystems.org/>
- <http://psychometrics.org/>

Další přístupy

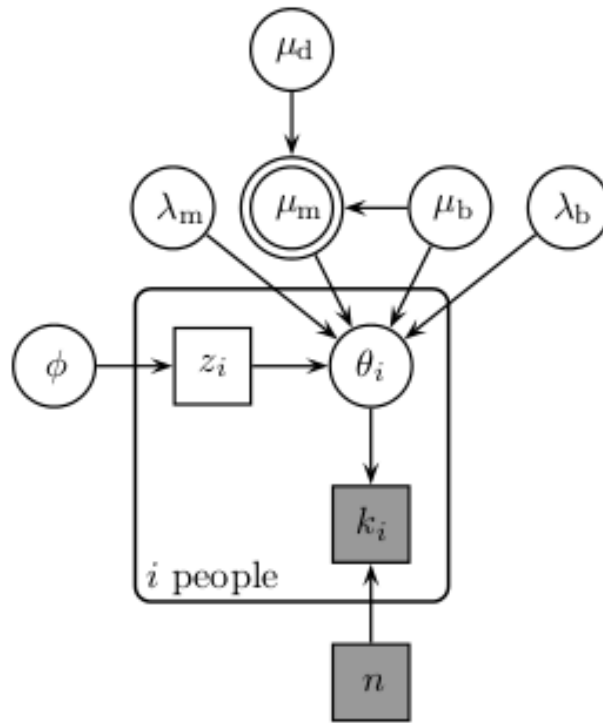
Teorie vědomostních prostorů

Spojité měření

IRT modelování

Bayesovské kognitivní modelování

Spíše ve výzkumu, neslouží primárně pro škálování a individuální diagnostiku.



$$\mu_b \sim \text{Beta}(1, 1)$$

$$\mu_d \sim \text{Gaussian}(0, 0.5)_{\mathcal{I}(0, \infty)}$$

$$\lambda_b \sim \text{Uniform}(40, 800)$$

$$\lambda_m \sim \text{Uniform}(4, 100)$$

$$z_i \sim \text{Bernoulli}(\phi)$$

$$\theta_i \sim \begin{cases} \text{Beta}(\mu_b \lambda_b, (1 - \mu_b) \lambda_b) & \text{if } z_i = 0 \\ \text{Beta}(\mu_m \lambda_m, (1 - \mu_m) \lambda_m) & \text{if } z_i = 1 \end{cases}$$

$$k_i \sim \text{Binomial}(\theta_i, n)$$

$$\text{logit} \mu_m \leftarrow \text{logit} \mu_b - \mu_d$$

$$\phi \sim \text{Beta}(5, 5)$$

Graphical model for inferring membership of two latent groups, consisting of malingers and *bona fide* participants.

Teorie vědomostních prostorů

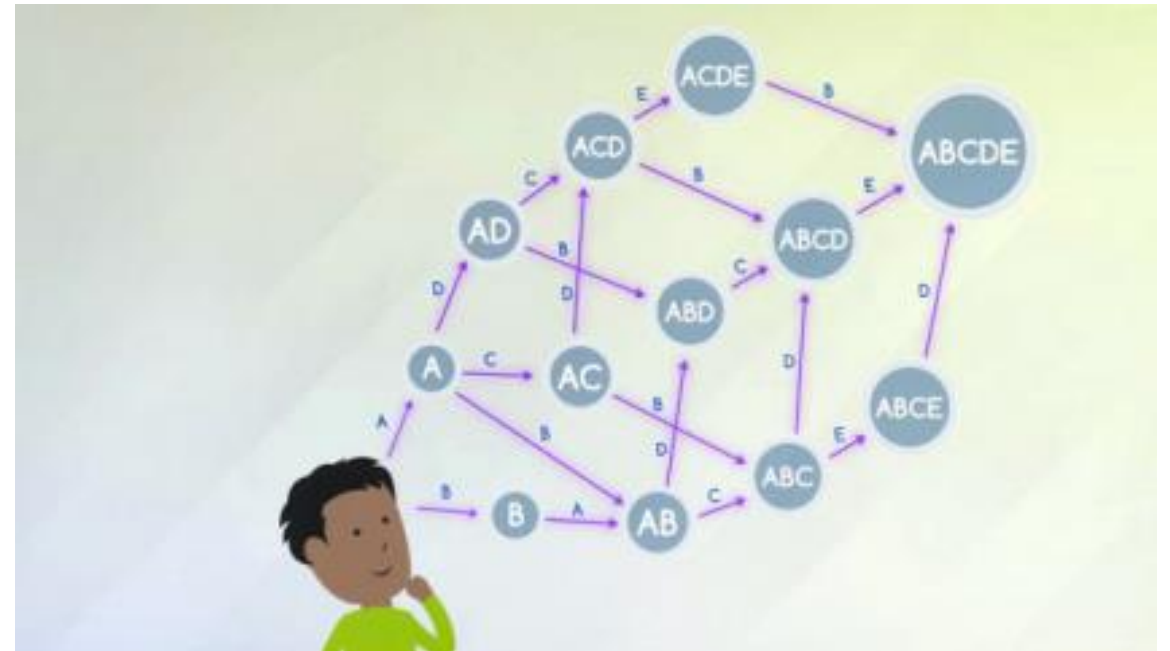
Pochází z didaktického a pedagogického testování.

Jednotlivé „znalosti“ mají své „prerekvizity“.

Kompenzační vs. nekompenzační modely.

Deterministické vs. stochastické modely.

- KST jsou omezený případ tzv. latent-class modelů.



Teorie vědomostních prostorů

Využívá se při plánování kurikula, výuky.

Různé tutorové výukové systémy.

Dva způsoby konstrukce/uvážování o vědomostních prostorech:

- Expertní návrhy.
- Empirický (item-tree analysis, cluster analysis aj.).

IRT modelování

IRT modeling, item response latent modeling...

IRT není v tomto kontextu chápáné jako statická teorie, ale jen jako konkrétní specifikace nějakého generalizovaného lineárního modelu (existují ovšem i nelineární aplikace).

Tzv. „general latent variable modeling“.

- Muthén, Muthén & Asparouhov v rámci SW Mplus.

Univerzální framework, v rámci něhož výzkumník definuje model, popisující jeho představu o tom, jaké složky se podílejí na vzniku pozorovaných dat... a model na těchto datech ověří.

IRTree

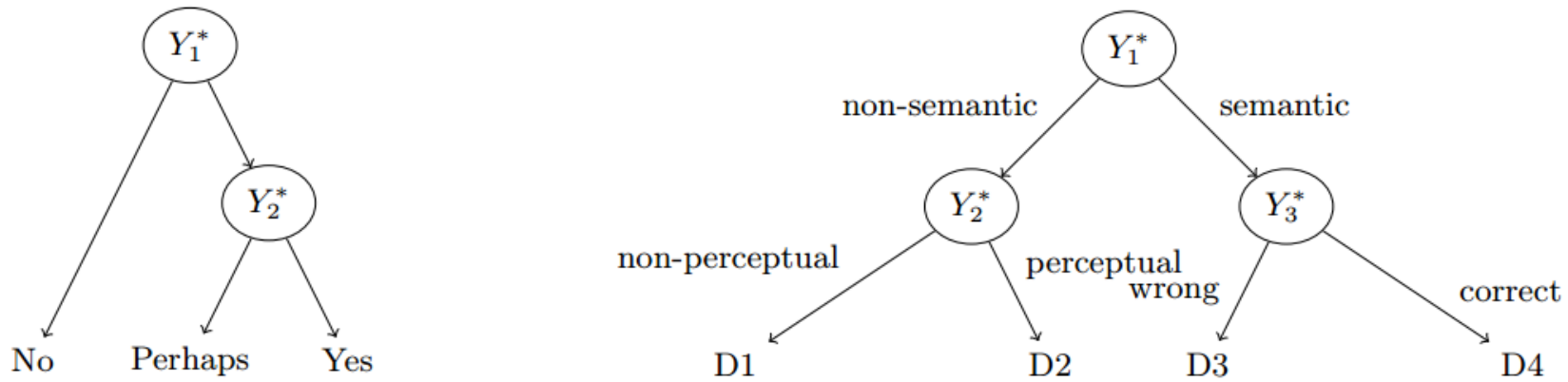


Figure 1: Examples of a linear response tree (left) and a nested response tree (right).

De Boeck, P., & Partchev, I. (2012). IRTrees: Tree-Based Item Response Models of the GLMM family. *Journal of Statistical Software*, 48(1), 1–18.

Explanační IRT modely

IRT modely, které kromě parametrů na straně položek a osob (obtížnosti, diskriminace, schopnosti) predikují odpověď na položku s pomocí dalších externích informací.

- Např. příslušnost respondenta ke skupině, aspekty na straně položky, o kterých se domníváme, že „způsobují“ obtížnost...

Umožňují nám lépe a s větší silou testu ověřovat různé hypotézy o respondentech či položkách.

Bayesovské kognitivní modelování

Velmi moderní přístup, není jasné, kterým směrem se bude ubírat vývoj.

- Umožněno extrémní rozvojem počítačové kapacity – i dnes ale analýzy trvají mnoho dní.

Snaží se „modelovat“ chování lidí v určité situaci.

- Časté použití v psychiatrickém kontextu u neuropsychologických testů.

Jiné zobecnění teorie odpovědi na položku.

Na základě teorie je specifikován konkrétní model pro daný test/situaci. Pomocí bayesovského modelování se odhadují parametry pro konkrétního respondenta, výzkumnou skupinu atp.

Hlavní použití je zatím ve výzkumu.