

Teorie odpovědi na položku (a další teorie měření)

PSYb2590: Základy psychometriky | Přednáška 10

23. 4. 2024 | Hynek Cígler a Petr Palíšek

Opakování: Teorie měření

Teorie měření: postuluje model/y měření.

- Epistemologická východiska a další předpoklady.
- Předpoklady o povaze měřeného atributu.

Model měření: popisuje vztah pozorování a atributu.

- Formalizace teorie měření do konkrétní matematické podoby..
- Způsob konstrukce skóru (resp. odhad „parametru“) z pozorovaných výsledků.
- Způsob odhadu chyby tohoto skóru.

Nikoli všechny psychometrické teorie jsou teoriemi „měření“!

- Atribut nemusí nutně být kvantitativní.

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).

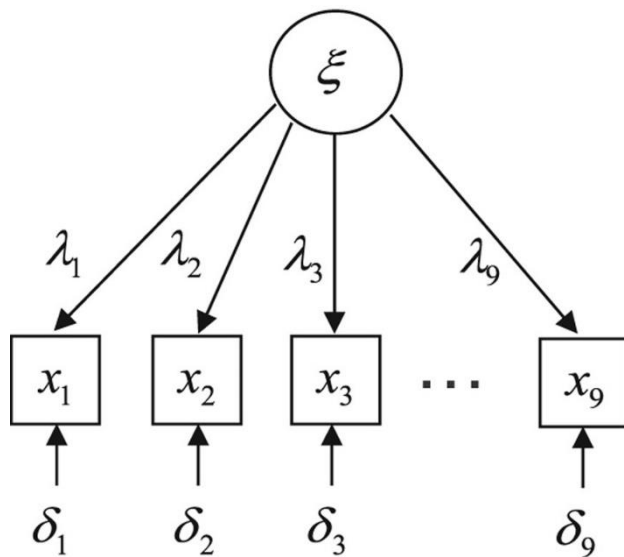
Podle Schmittmanna ([2013](#)):

- Reflektivní model
 - Rys způsobuje chování.
- Formativní model.
 - Rys je konstrukcí na základě chování.
- Dynamické systémy.
 - Jevy spolu souvisejí a rys neexistuje.
- Schmittmann vychází implicitně z realistické epistemologie.

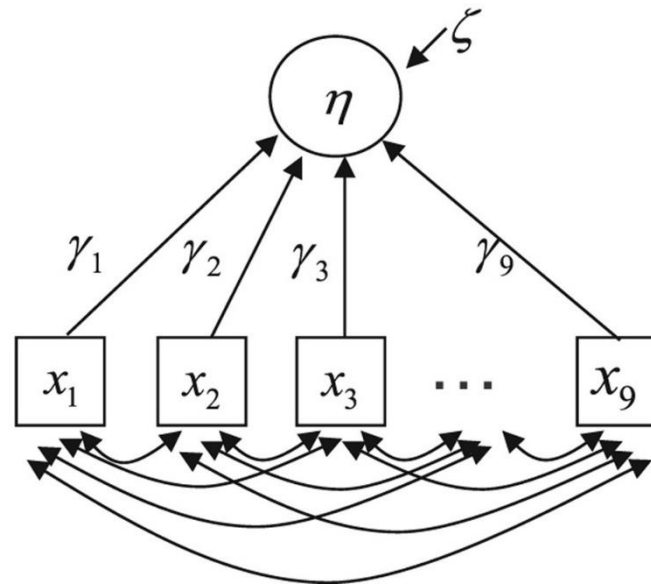
Tady jde spíše o teorie měření (které mohou obsahovat různé konkrétní modely).

Modely měření v psychologii

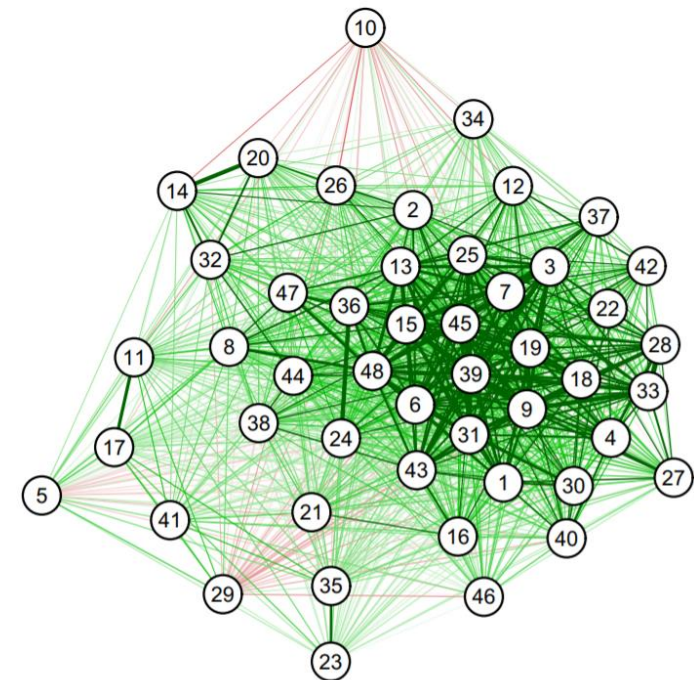
(a) Reflective measure model



(b) Formative measure model



síťový model



Další vybrané modely měření v psychologii

DALŠÍ = DALŠÍ MIMO KLASICKOU TESTOVOU TEORII (TENTO KURZ)
A TEORII ODPOVĚDI NA POLOŽKU (TATO PŘEDNÁŠKA)

VYBRANÉ = PSEUDONÁHODNÝ VÝBĚR ZE STRAŠNĚ VELKÉHO BALÍKU

Teorie zobecnitelnosti

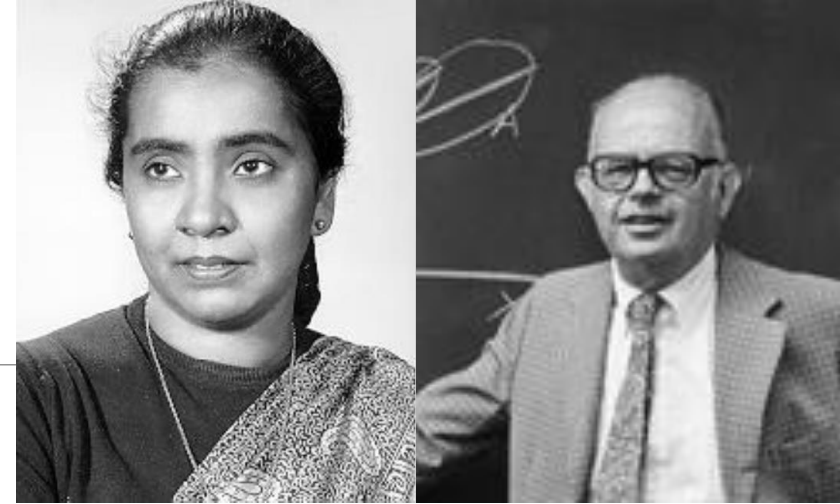
Klasická testová teorie: $X = \tau + e$

Teorie zobecnitelnosti: $X = \tau + e_1 + e_2 + e_3 + \dots + e_n$

- Cronbach, Nageswari & Gleser (1963)
- Zobecnění klasické testové teorie pro případy s více zdroji chybového/systematického rozptylu.
- Příklad: „Osoba P je v rámci assessment centra hodnocena 2 hodnotiteli R na pěti kritériích I při dvou různých scénkách O . Každé kritérium 0–10 bodů.
- Celkový skór je průměrem všech udělených kritérií.
- Jaká je chyba spojená při odhadu *obecného skóru*?

Přínosy a využitelnost:

- Relativní vs. absolutní design měření.
- Vývoj testů (průměrná chyba měření napříč kohortami, podobnost s SB vzorcem).



Reprezentační model Teorie spojitého měření



Alternativní reakce na Fergusonovu komisi.

- Nezávisle na sobě ekonom a laureát Nobelovy ceny Gérard Debreu (1960) a psycholog Duncan Luce s matematikem Johnem Tukeym (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 (1PL IRT) může být jednou ze stochastických specifikací CM.

- Za dodržení několika dalších předpokladů.

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

- Jinými slovy: Se spojitým měřením jsou spojeny hned dvě Nobelovy ceny! 😊

Teorie vědomostních prostorů

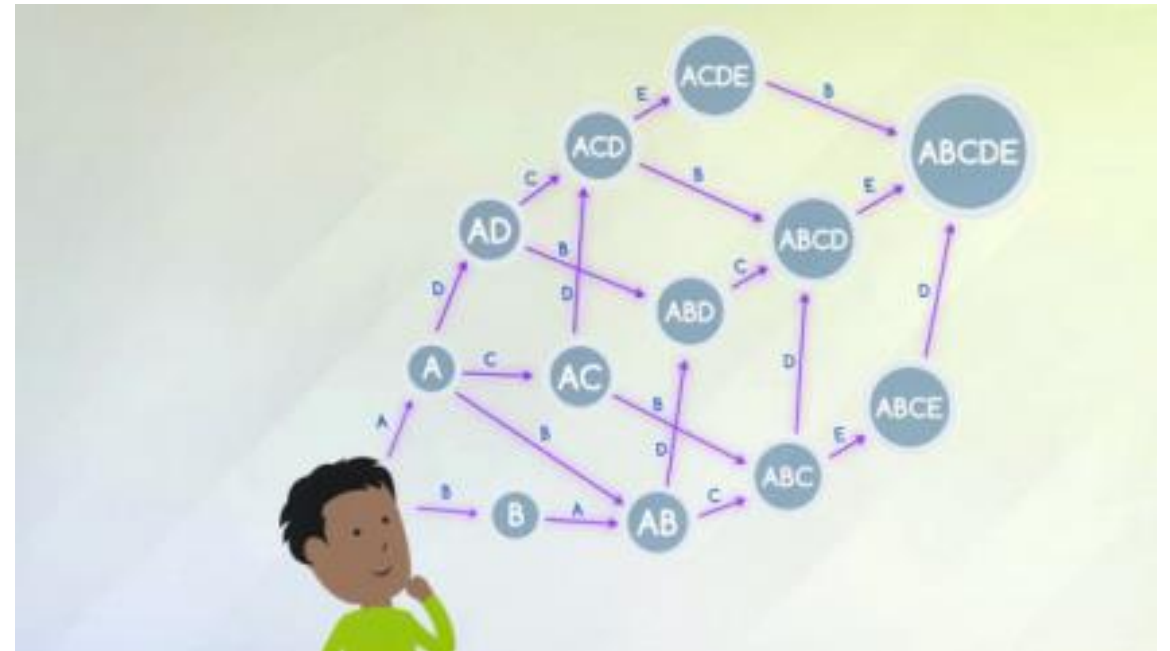
Pochází z didaktického a pedagogického testování.

Jednotlivé „znalosti“ mají své „prerekvizity“.

Kompenzační vs. nekompenzační modely.

Deterministické vs. stochastické modely.

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



(Dynamické) síťové modely

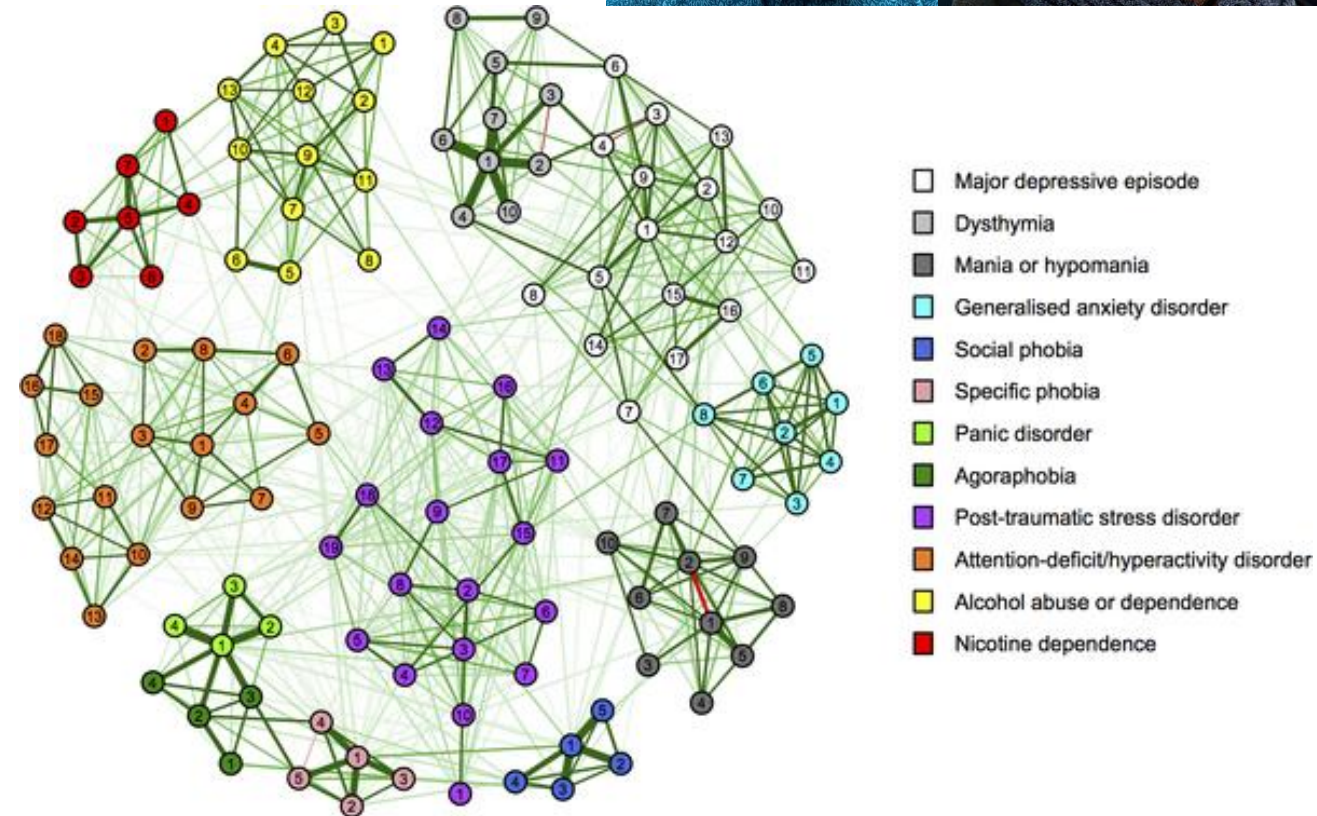


Nejde o teorii měření, ale o *realistický popis struktury a dynamiky psychických fenoménů*.

- Síťové modely ale lze využít při konstrukci měřicích nástrojů (stejně, jako FA lze využít pro vývoj testu skórovaného pomocí CTT).
- Velmi rychle rostoucí oblast psychometrie.

Aplikace:

- Personalizované dynamické modely.
- Popis dynamiky vývoje psychických obtíží a podpora evidence-based terapie.
- Vysvětlení jejich komorbidity.
- Omezené usuzování na kauzalitu (Pearls, DAG).



Ještě další modely/teorie

Signal Detection Theory (SDT).

- Odvozená z Shannonovy informační teorie. Využívá se v kognitivní psychologii pro modelování rozhodování.

Sociometrie (Moreno, Jennings).

- Faktorová analýza korelační matice transponovaných dat.
- Méně schopná sestřička síťových modelů, **eticky problematické při běžném využití.**

Q-sort a Q-metoda obecně (Stephenson, aplikace Rogers).

- O tom skoro nic nevím, ale už rok nám běží projekt TAČR, tak třeba zase za rok 😊

Multidimenzionální škálování.

Modely s nominálními latentními rysy jako specifický případ modelů s latentní proměnnou.

- **Latent Class Analysis (LCA)**, Latent Profile Analysis, **Diagnostic Classification Models (DCM)**

A řada dalších...

Teorie odpovědi na položku

Item Response Theory

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

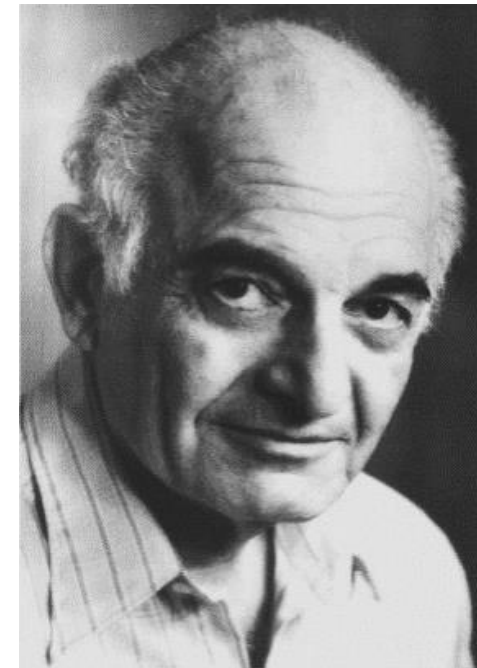
a mnoho dalších



Paul Felix Lazarsfeld (1901–1976)



Georg Rasch (1901-1980)



Louis Guttman (1916–1987)

Teorie odpovědi na položku

Item Response Theory (IRT), dnes někdy též „**item-factor analysis**“.

Motivací vzniku IRT bylo studium **vztahu jedné položky a celkového skóre**.

- CTT a FA předpokládají lineární vztah, což není realistické zejm. u nominálních/ordinálních položek.
 - Lineární vztah je přitom nezbytný předpoklad sčítání položek do celkového skóre.
- V případě spojitě odpovědi může být IRT totožná s faktorovou analýzou.

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

- Stochastická úprava původně deterministického Guttmanova škálování.
- Postupně vzniklo masivní množství různých modelů v rámci IRT paradigmatu.
- Vysoce flexibilní: různé modely pro různé typy položek.

Počátky v 50. a 60. letech 20. stol.

- Původně limitované výpočetní náročností.
- Cca od 70. let první praktické aplikace s rozvojem výpočetní techniky.
- Od 90. let prudký rozvoj.

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 (LCA).

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 typicky 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.

Charakteristická funkce položky Item Characteristic Curve (ICC)

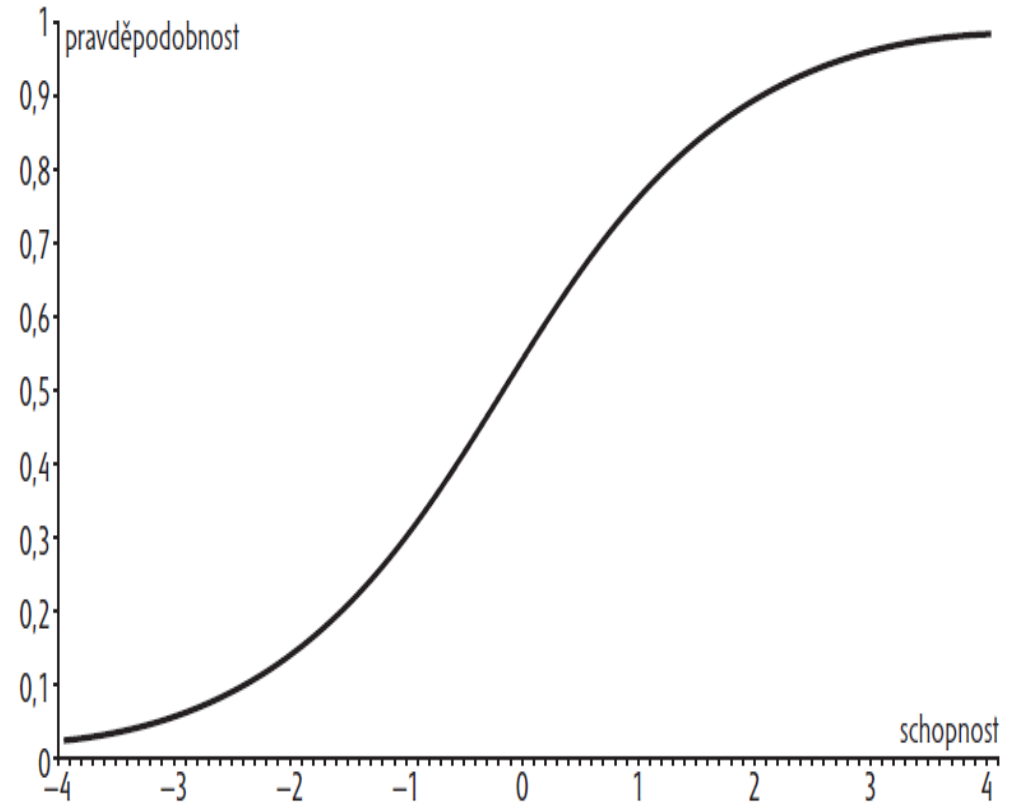
Operacionalizace daného IRT modelu.

- Jednotlivé IRT modely se liší právě svou ICC.

Obtížnost položek a schopnost respondenta jsou na stejné škále.

Charakteristická funkce položky:

- Má 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) logistický model

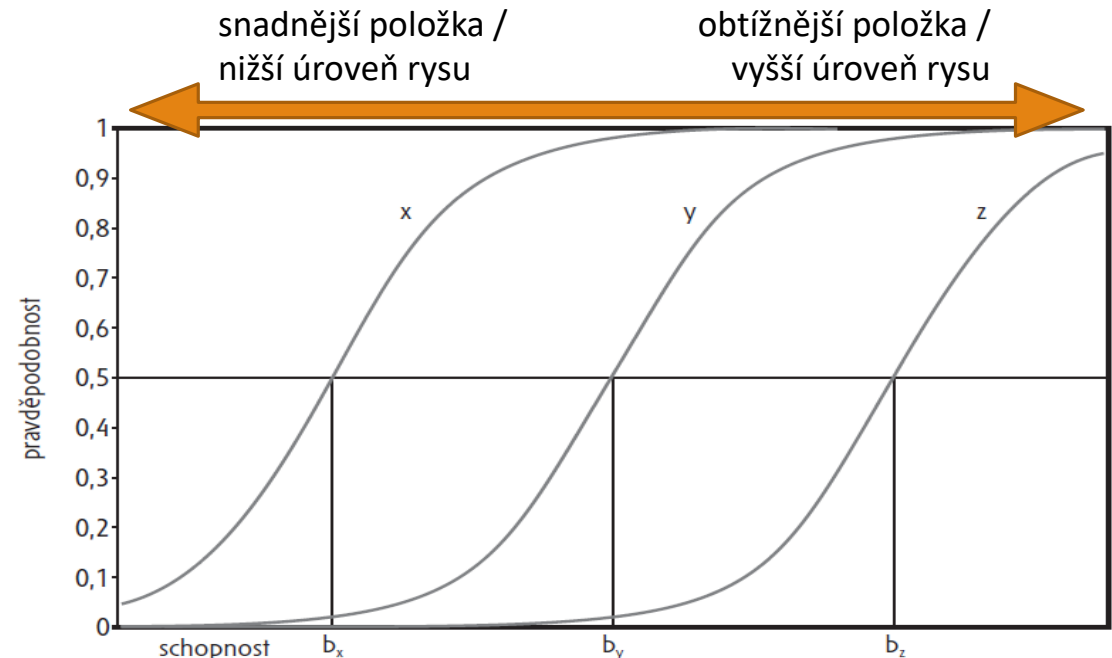
Charakteristická funkce:

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

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 se stejnou mírou schopnosti ($\theta - b_i$) na danou položku $P_i(\theta) = 0,5$.

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



Základní logistické IRT modely pro binární položky

1PL – technicky shodný s Raschovým modelem.

- 1PL = 1 Parameter Logistic (test model)
- Pouze parametr obtížnosti b_i .

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

2PL – jak dobře model diskriminuje?

- Přidává strmost ICC prostřednictvím diskriminačního parametru a_i .

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

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

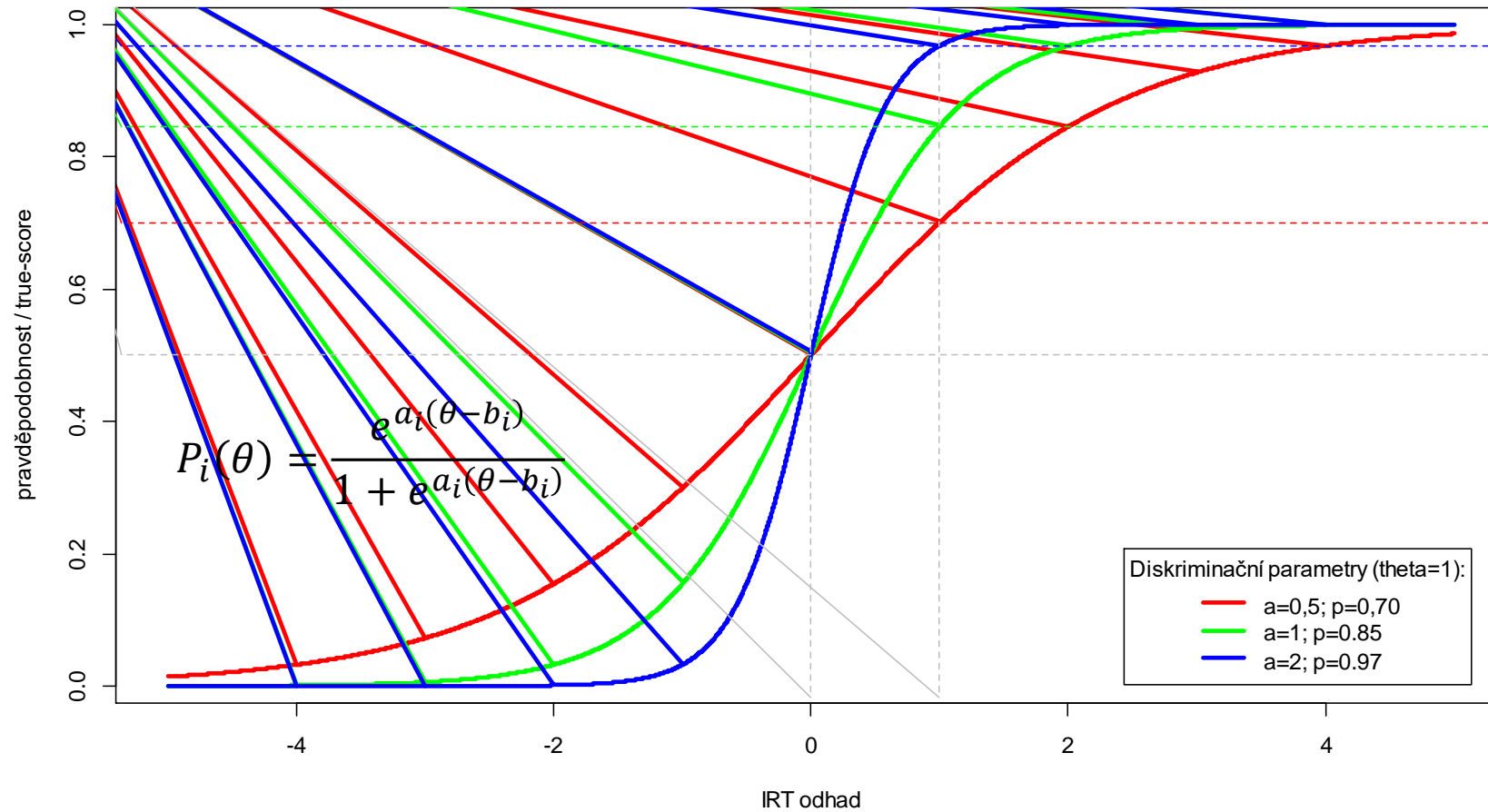
- Přidává spodní limitu ICC pomocí parametru pseudouhádnutelnosti c_i .
- Modeluje tipování u multiple-choice testů.

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

Postupnou fixací parametrů 3PL modelu lze dosáhnout jednodušších modelů:

- $c_i = 0 \rightarrow$ 2PL model
- $a_i = 1 \rightarrow$ 1PL model

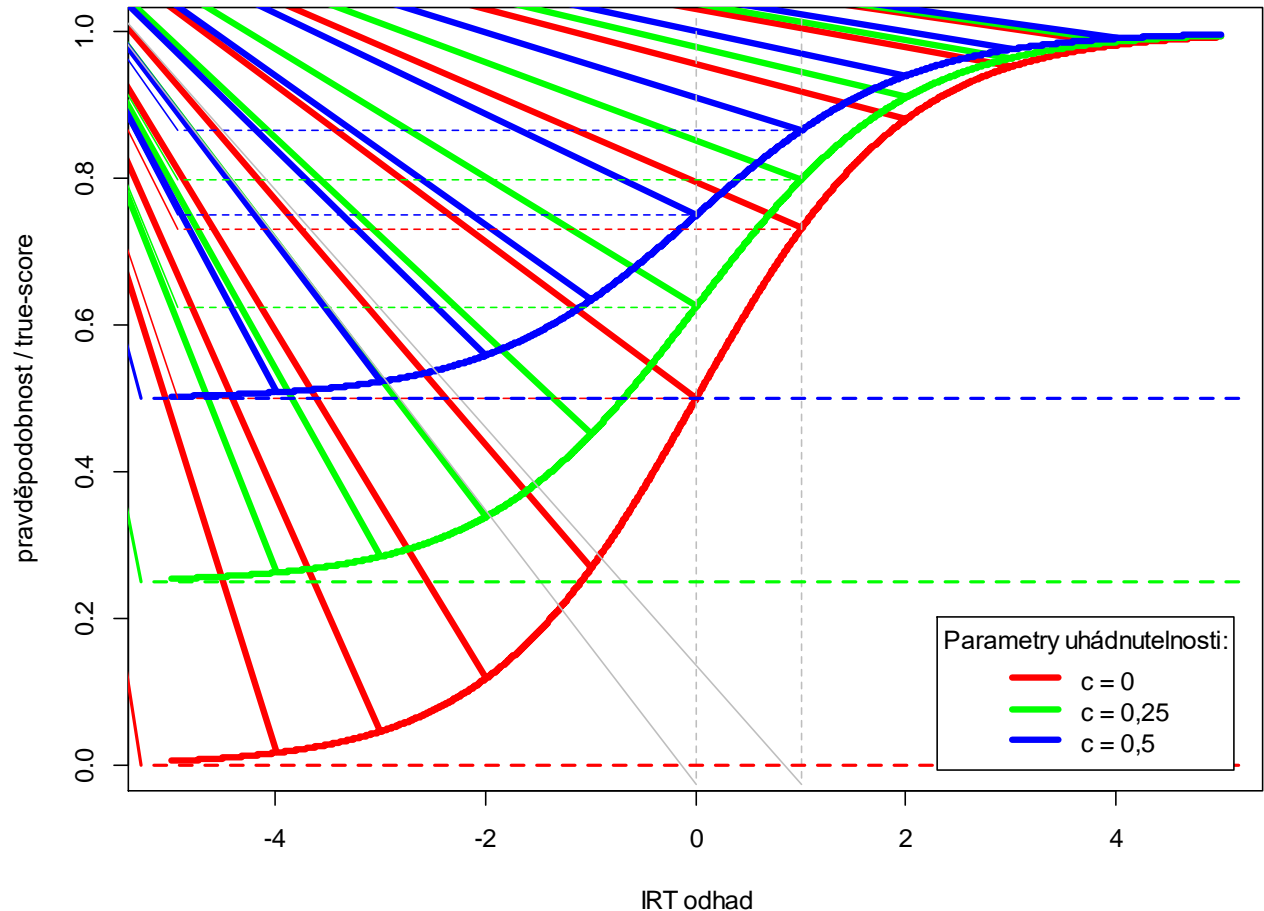
Charakteristické křivky položek 2PL modelu



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 strmosti ICC.
- b_i – parametr obtížnosti
 - Umístění ICC na škále latentního rysu.
- c_i – parametr pseudouhánutelnosti
 - Spodní limita ICC.



Estimace IRT modelů

Full-information odhad nad plnou datovou maticí.

- Na rozdíl od „limited-information“ odhadem nad kovarianční maticí jako u FA.

Díky tomu:

- Jen minimálně ovlivněné (náhodně, tedy MAR/MCAR) chybějícími daty.
- Relativní nezávislost parametrů položek na vzorku.
 - Platí u Raschova modelu při použití JML estimátoru, jinak je to komplikovanější.
- Výhodné při designech s plánovaně chybějícími daty.
- Lze „uložit“ kalibrované parametry položek a použít je jinde.
 - Např. adaptivní testování nebo naopak vyvažování testů (test, resp. anchoring/equating/linking)

Další typy modelů

Polytomní IRT modely určené pro položky s více odpověďmi (Likertova škála aj.).

- Partial Credit model, Nominal Response model, Graded Response Model a řada dalších.
- Graded response model je logistickou, full-information variantou ordinální faktorové analýzy (probitový model s limited-information estimací).

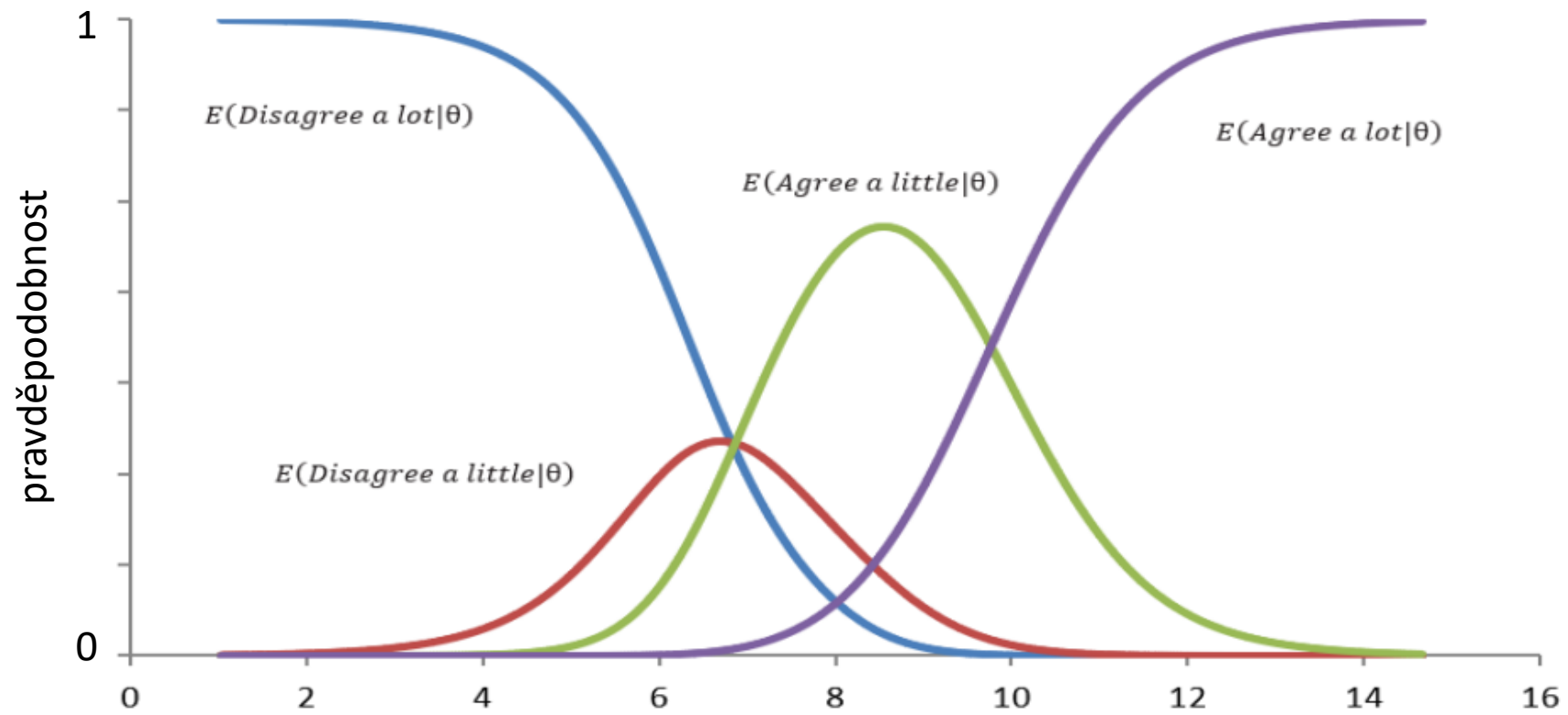
Multidimenzionální IRT modely; např.:
$$P_i(\theta) = \frac{e^{a_{i1}\theta_1 + a_{i2}\theta_2 + \dots + a_{in}\theta_n + b_i}}{1 + e^{a_{i1}\theta_1 + a_{i2}\theta_2 + \dots + a_{in}\theta_n + b_i}}$$

- Explorační i konfirmační.
- Kompenzatorní, parciálně kompenzatorní či non-kompenzatorní IRT modely.

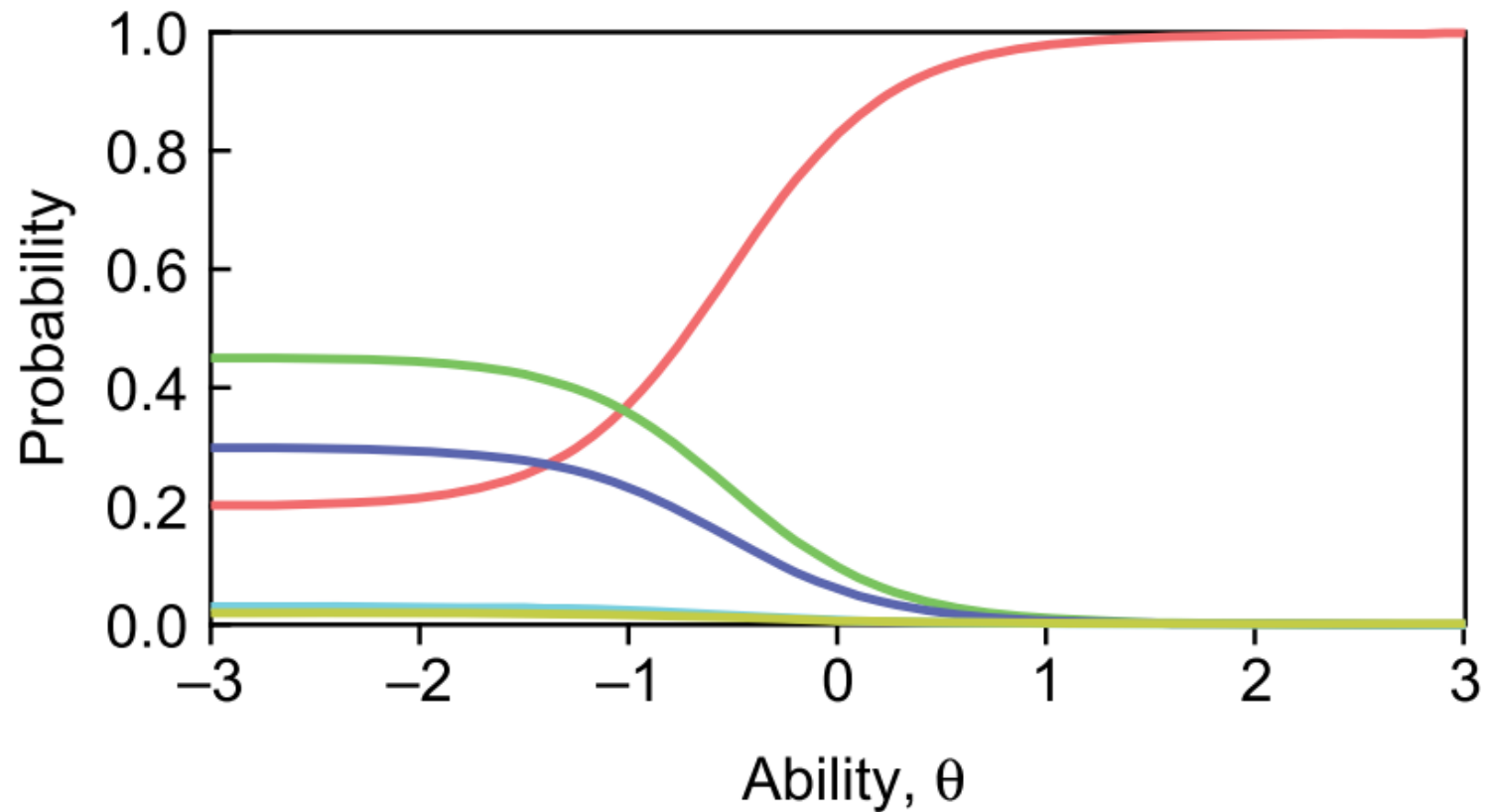
Další specifické modely.

- Například modely pro práci s reakčními časy, explanační IRT modely a podobně.
 - Mimochodem, právě model pro práci s rychlostí čtení byl prvním Raschovým (1960) IRT modely (samotný RM publikoval až 1966).

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/>

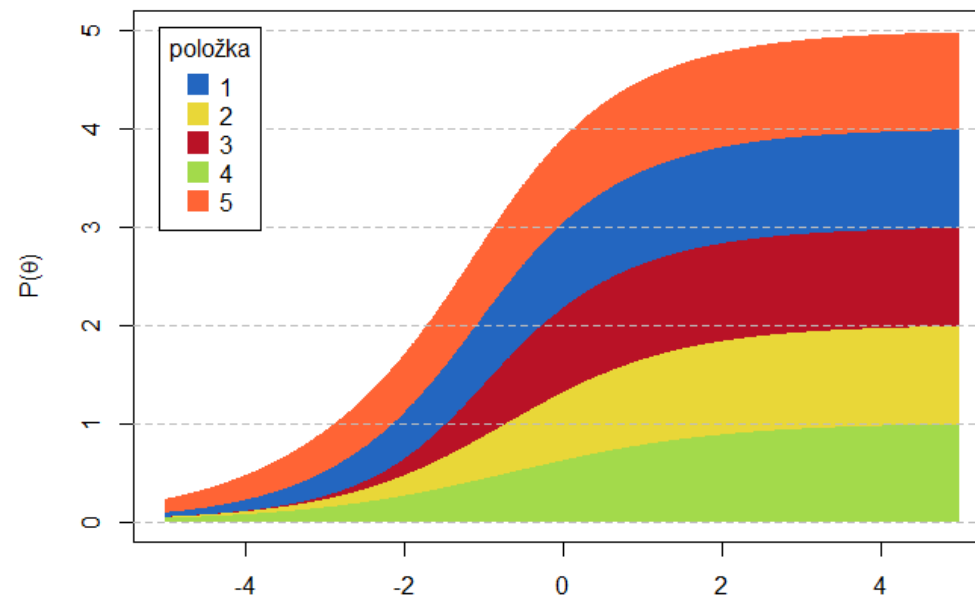
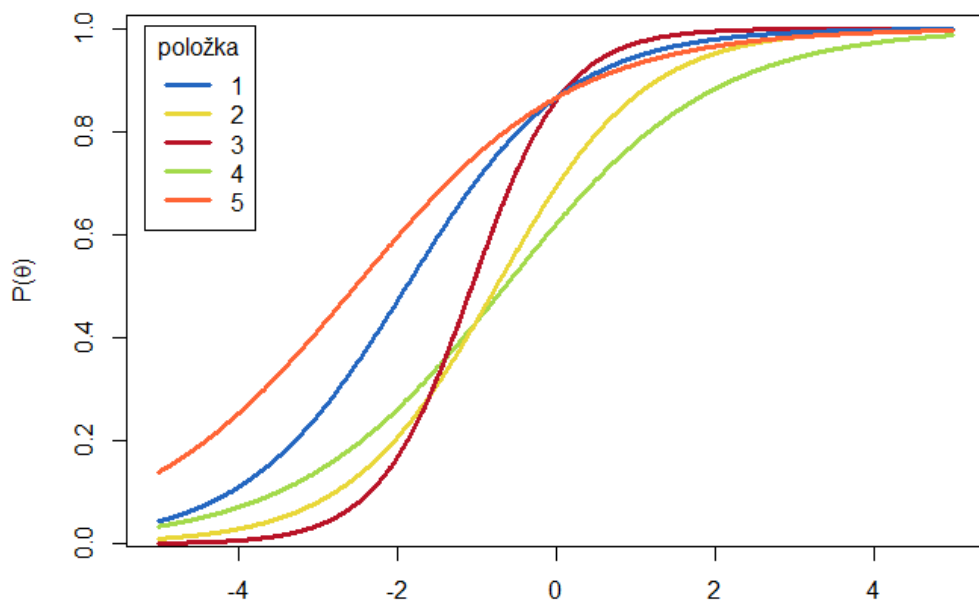
- 4PL binární model: záložka IRT models → Training → Dichotomous model
- Polytomické modely (GRM, GPCM, NRM): záložka ... → Polytomous model

Charakteristická funkce testu (TCF)

Popisuje vztah celkového hrubého (součtového) skóru a latentního rysu.

- Prostý součet jednotlivých ICC:

$$TCC(\theta) = \sum_{i=1}^n P_i(\theta) = E(T|\theta)$$



Chyba měření v IRT

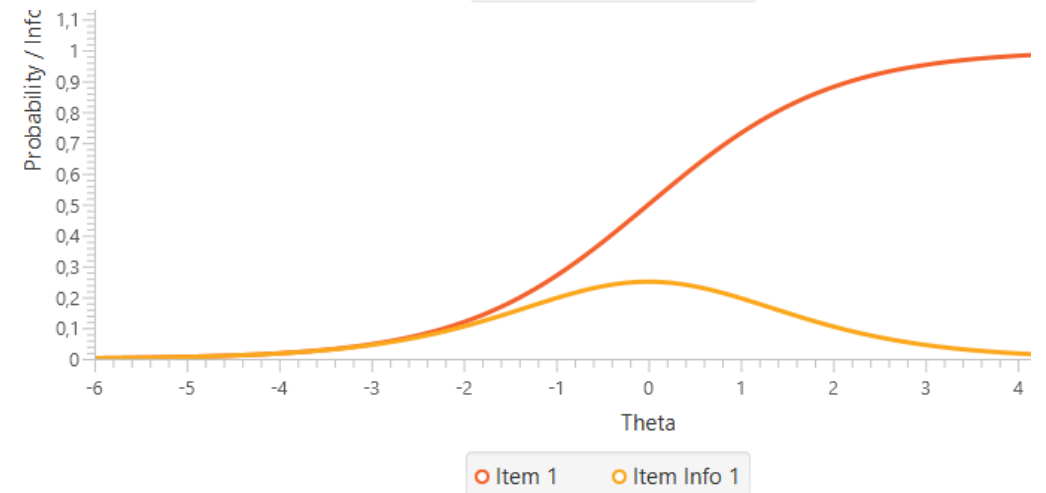
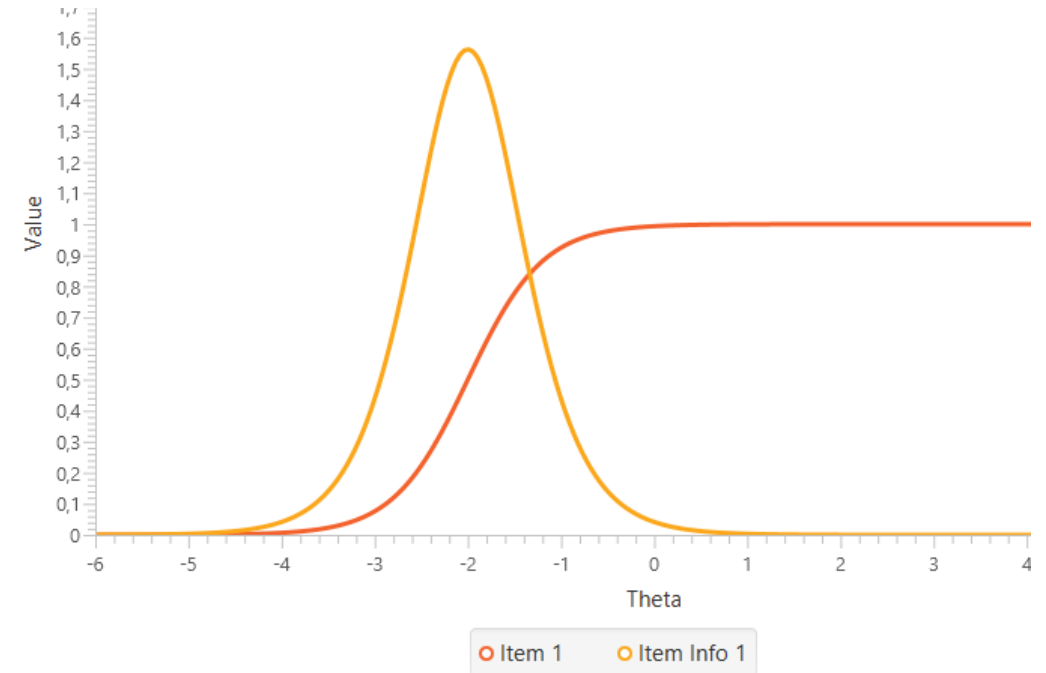
Každá položka přináší jiné množství informace o každé úrovni latentního rysu.

- Snadná položka → diferencuje respondenty s nízkou úrovní rysu.
- Obtížná položka → diferencuje respondenty s vysokou úrovní rysu.

Informační funkce položky.

- Item information function (IIF).
- Například pro 2PL model:
$$I_i(\theta) = a_i^2 P_i(\theta)[1 - P_i(\theta)]$$
- $P_i(\theta)$ – ICF, a_i – diskriminační parametr.

Umožňuje respondentovi vybírat položky z položkové banky „na míru“ (počítačové adaptivní testování).



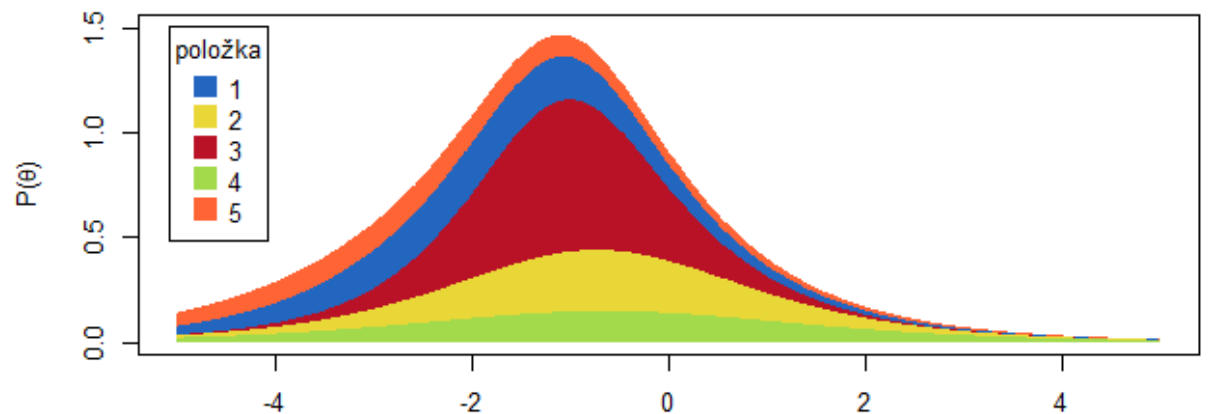
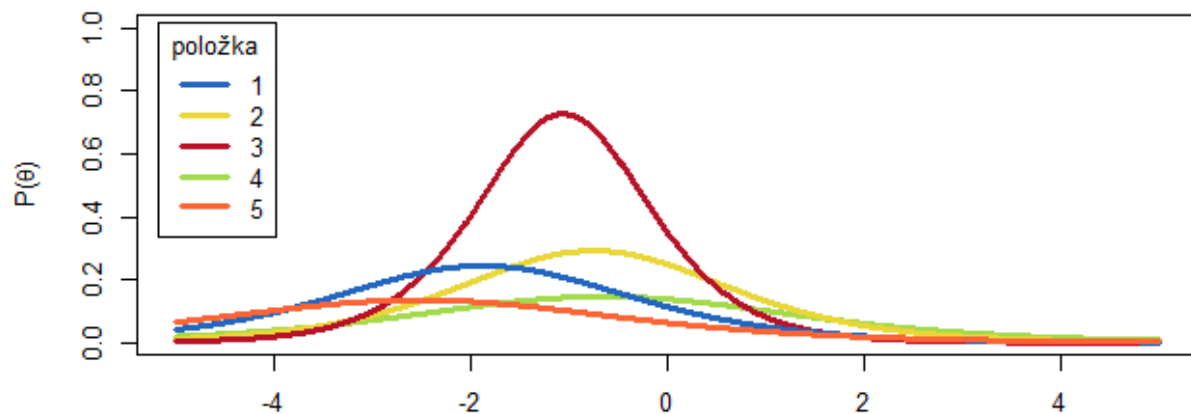
Informační funkce testu (TIF)

Celková míra informace, kterou nám přináší test o osobě s určitou mírou rysu:

- Prostý součet informačních funkcí položek:

$$I(\theta) = \sum_{i=1}^n I_i(\theta)$$

Lze ji využít pro odhad chyby měření (chybu odhadu latentního rysu): $SE(\hat{\theta}) = \frac{1}{\sqrt{I(\theta)}}$



Další ukazatele v rámci IRT

Shoda modelu s daty (podobně jako u CFA).

Shoda položky s daty (item-fit) – řada indikátorů.

- Další možný ukazatel při IRT položkové analýze.

Shoda respondentů s daty (person-fit) – řada indikátorů.

- Indikátor aberantního odpovídání a podobně.

Analýza diferenciálního fungování položek (DIF).

- Nefunguje jedna položka rozdílně pro různé skupiny osob?
- Ukazatel férovosti položky testu.

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 skóru) respondentů ve stávajícím vzorku.
 - Namísto součtu položek – stejná data, jako v bodě 1.
 - **IRT jako nástroj pro vytvoření skóru** při standardizaci psychodiagnostických metod.
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** (namísto součtových skóru).

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.
- Kvalitnější konstrukce skóru (tzv. W-skóry).
- 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.

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

Person	Occasion					
	I			II		
	Item 1	Item 2	Item 3	Item 1	Item 2	Item 3
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 3 položkovým 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).
 - Přímě 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

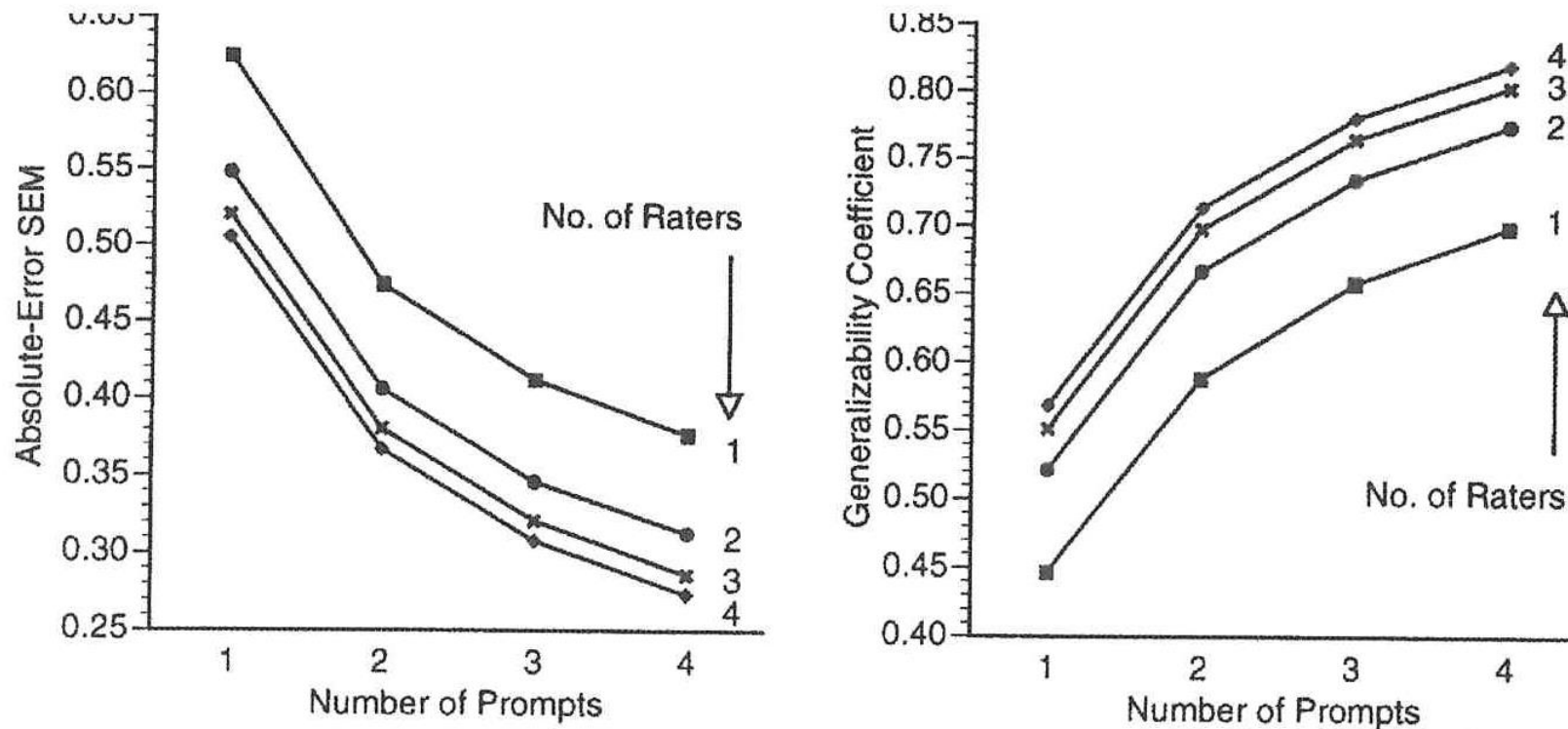


FIGURE 1.2. $\hat{\sigma}(\Delta)$ and $E\hat{\rho}^2$ for scenario with $p \times T \times R$ design.

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).

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: $m > \text{cm} > \text{mm}$
- Respondenti: $\text{stůl} > \text{kniha} > \text{tužka}$

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

Jednoduché vykrácení

- $(\text{tužka}, \text{mm}) < (\text{kniha}, \text{mm})$
- $(\text{tužka}, \text{mm}) > (\text{tužka}, \text{cm})$

Dvojité vykrácení

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

respondenti

interakce

položky

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

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

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

- $\text{tužka} + \text{m} < \text{stůl} + \text{mm}$
- $20 * 500 = 10000$

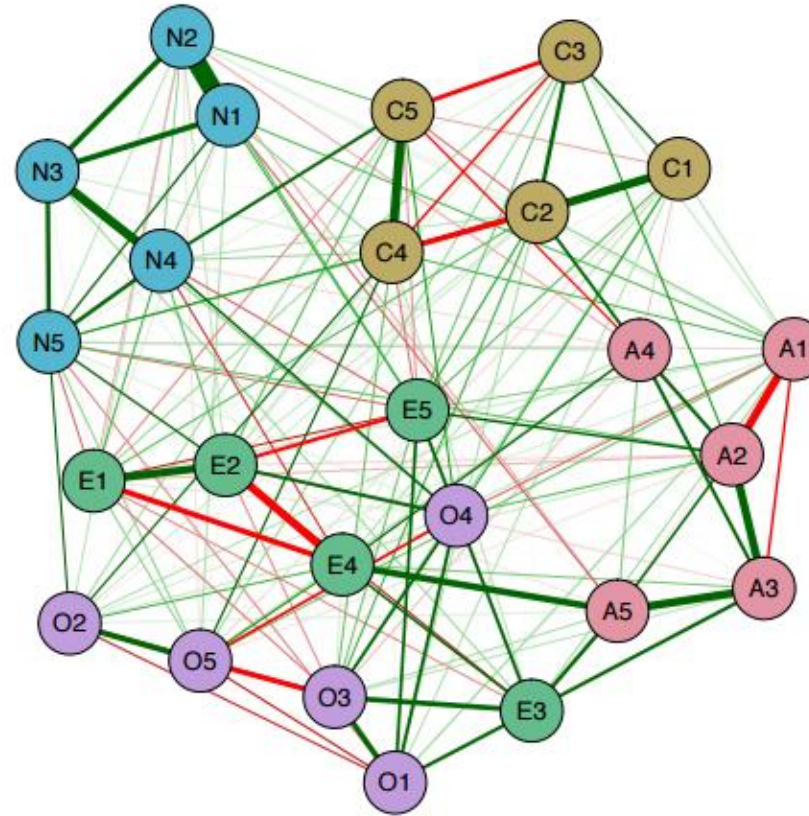
Fundamentální škála: $\text{tužka}=1, \text{kniha}=2, \text{stůl}=10$
 $\text{m}=1, \text{cm}=100, \text{mm}=1000$

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.

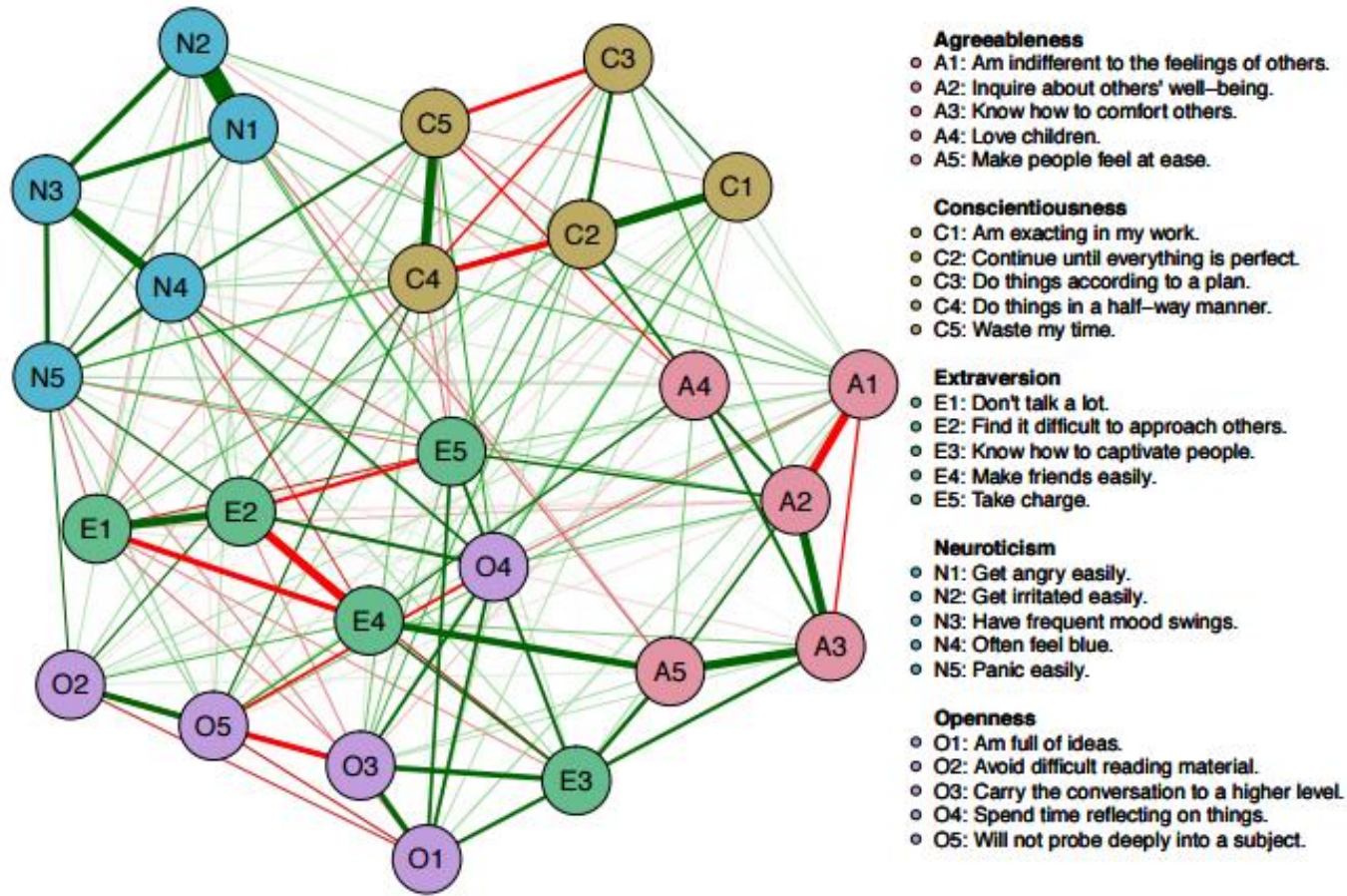


Figure 1.1: Example of a network model estimated the BFI personality dataset from the psych package in R. Nodes represent variables (in this case, personality inventory items) and links between the nodes (also termed *edges*) represent partial correlation coefficients. Green edges indicate positive partial 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.](#)

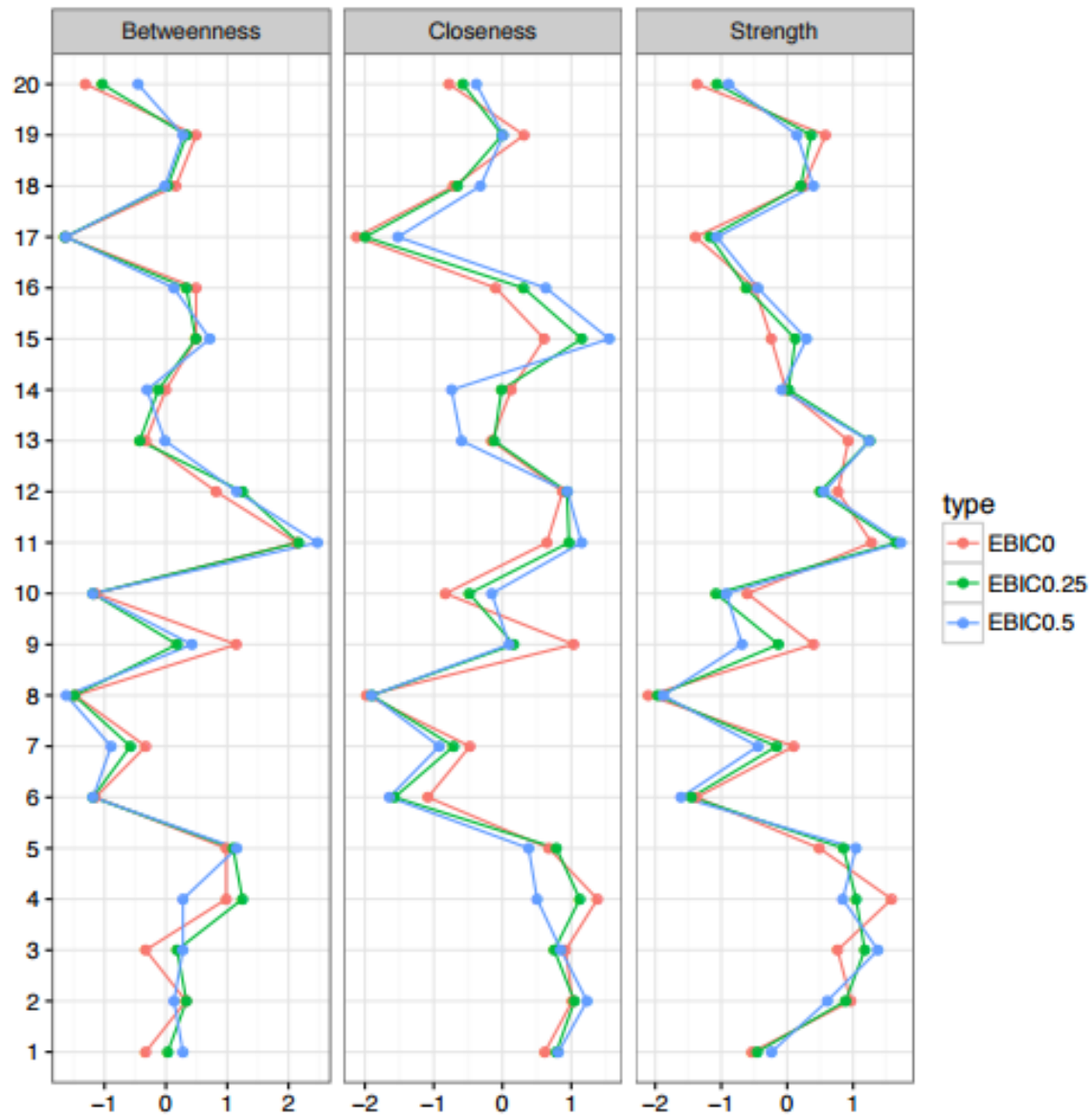


Figure 2.4: Closeness, betweenness, and degree centrality of the three networks described in Figure 2.3 with increasing levels of the LASSO hyperparameter γ . Values are standardized to z-scores.

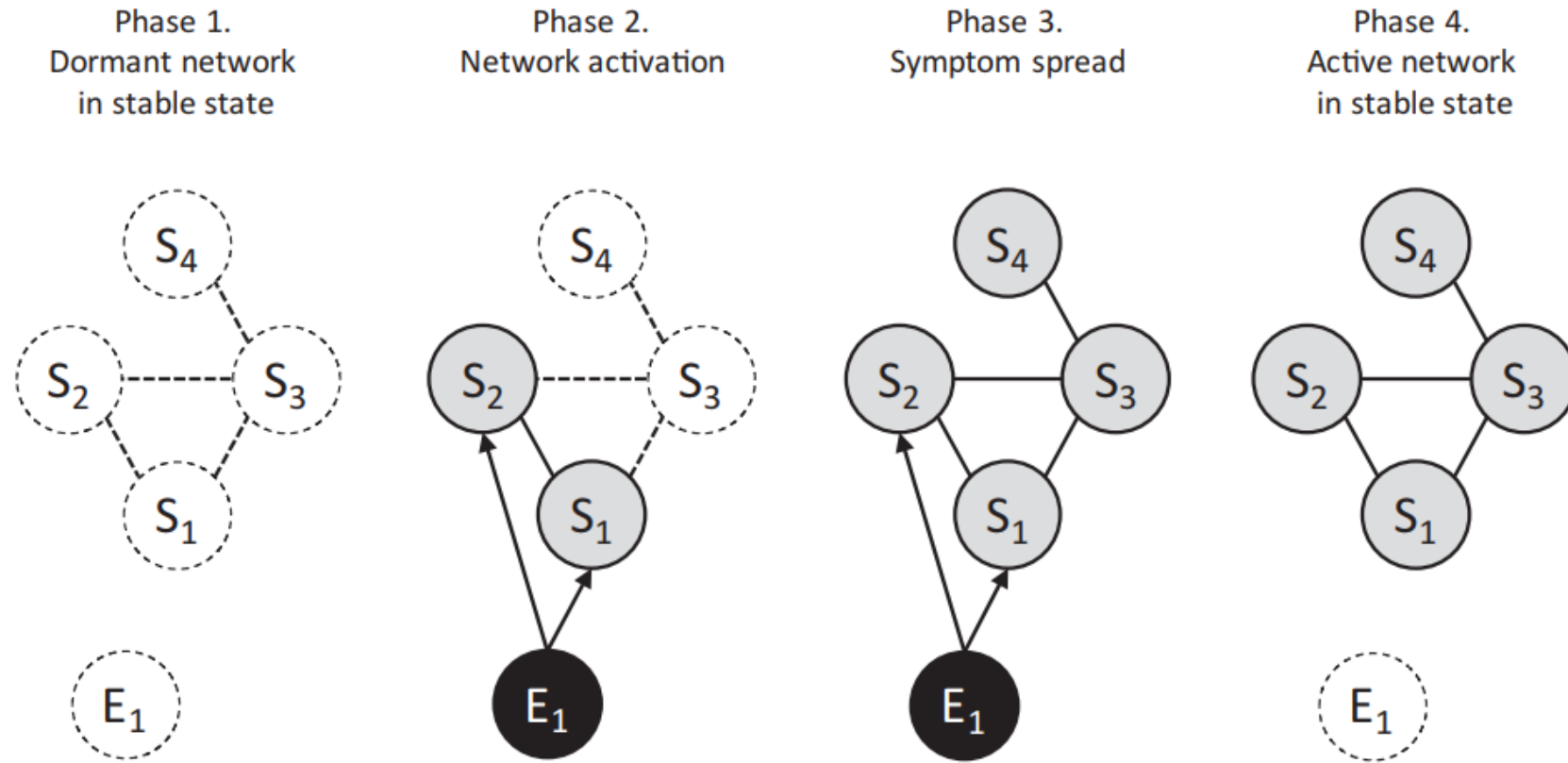
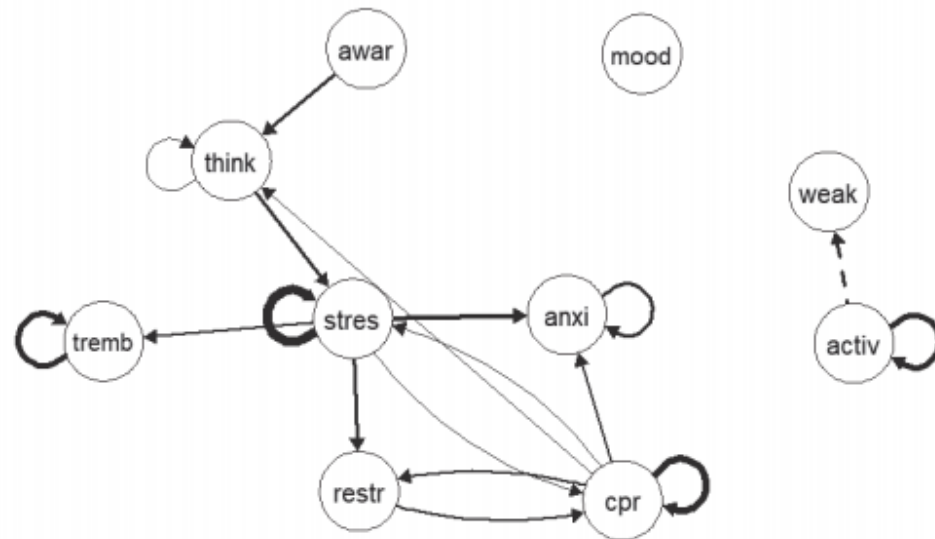


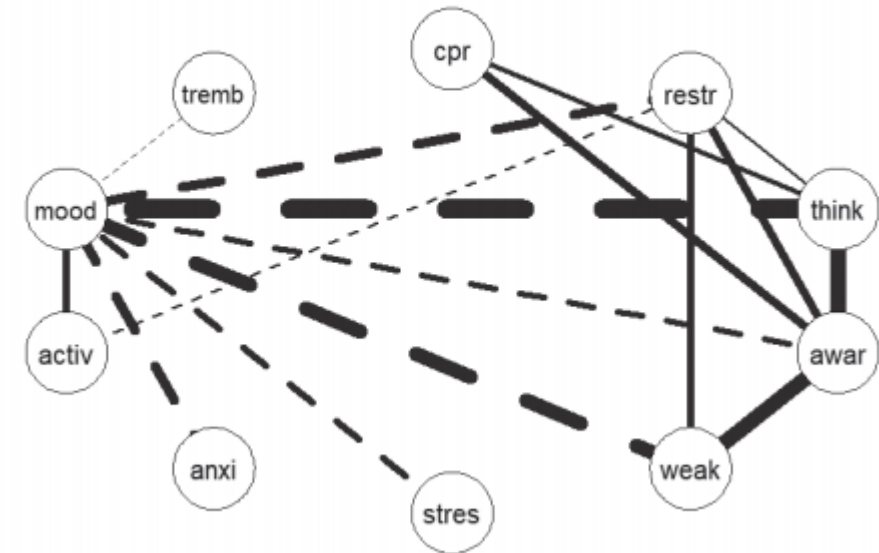
Figure 3 Phases in the development of mental disorders according to the network theory. After an asymptomatic phase, in which the network is dormant (Phase 1), an external event (E₁) activates some of the symptoms (Phase 2), which in turn activate connected symptoms (Phase 3). If the network is strongly 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/>