Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Hana Křížovská

Edukometrie - Item Response Theory

Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: Prof. Ing. František Fabian, CSc.

Studijní program: Matematika, studijní obor: Matematická statistika

Na tomto místě bych chtěla poděkovat především vedoucímu diplomové práce panu Prof. Ing. Františku Fabianovi, CSc., který mi poskytl mnoho odborných rad a zároveň mi nechal volnost při sestavování obsahu této práce. Další dík patří panu Prof. RNDr. Stanislavu Komendovi, CSc. z Lékařské fakulty Univerzity Palackého v Olomouci, který se kus svého života věnoval této teorii a troufám si tvrdit, že je v této oblasti nejzasvěcenějším člověkem v České republice. Za poskytnutí dat děkuji panu Doc. Ing. Petru Byčkovskému, CSc. z Pedagogické fakulty Univerzity Karlovy. Poslední ale ne nejmenší dík patří mé mamince a tatínkovi, kteří mě psychicky podporovali (a že to bylo nutné) a mému příteli Davidu Voňkovi, který měl se mnou nejen trpělivost, ale i mi pomohl se zpracováním dat v systému Ox.

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze 1.8.2003

Hana Křížovská

Hana Křížovská

Obsah

Title: Measurement of Education - Item Response Theory
Author: Hana Křížovská
Department: Department of Probability and Mathematical Statistics
Supervisor: Prof. Ing. František Fabian, CSc.
Supervisors' e-mail address: he has not

Abstract
If we want to study abilities of certain pupils or other people in a given area, we can give them a didactial test, which would consist of items testing the abilities in question. The goal is then to determine the qualities of the people (ability parameter) and the relative difficulty of the questions (parameters of the questions) on the basis of the answers. Further we estimate how would a pupil, who have passed just one part of the test (assuming we have the item parameters of the whole test), would perform in the rest of the test. This is the research area of the Item Reponse Theory (IRT). In the introductory section the reasons are given for why we should use the IRT rather than the classical test theory for evaluating didactical test results, althought it's not regulary done in our country. The dichotomous Rasch model, which is one of the main IRT models, is described in detail. First the model is derived and then its main features presented. For the case of Rasch models we also demonstrate the most used methods of estimation of the ability and difficulty parameters. In the conclusion a possible continuation of the work is drafted.

Keywords: Dichotomous Rasch model, Item characteristic curve, difficulty parameter, ability parameter.

Název práce: Edukometrie - Item Reponse Theory
Autor: Hana Křížovská
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: Prof. Ing. František Fabian, CSc.
Email vedoucího: nemá

Abstrakt
Pokud chceme zjistit, jak schopní jsou nějací žáci či jiné osoby, můžeme jim zadat didaktický test, který bude sestaven z otázek testujících danou schopnost. Cílem pak je, na základě odpovědí žáků či jiných osob, odhadnout jejich schopnost (parametr schopnosti) a vlastnosti otázek (parametry otázek) v testu. Dále předpovídáme, jak by žák, kterému jsme zadali jen několik otázek v testu (mějme pro něj zjištěny parametry otázek), odpovídal na zbylé otázky v testu. Touto problematikou se zabývá teorie odpovědí na položky (zkratka TOP, anglický ekvivalent je "Item Response Theory"). V úvodu diplomové práce jsou popsány důvody, proč je k vyhodnocování didaktických testů lepší využívat modely u nás málo používané teorie odpovědí na položky než klasickou teorii testů. Podrobně je popsán Raschův model pro dichtomickou veličinu, což je jeden z hlavních modelů TOP. Nejprve je ukázáno jeho odvození a následně jsou uvedeny jeho základní vlastnosti. Na Raschových modelech pro dichotomickou veličinu jsou ukázány nejpoužívanější metody odhadu parametru obtížnosti otázek v didaktickém testu a parametru schopností zkoušeného, které používá TOP. V závěr je naznačeno, jak by šlo na tuto práci navázat a prozkoumat tak i jiné modely TOP.

Klíčová slova: Raschův model pro dichotomickou veličinu, charakteristická křivka položky, parametr obtížnosti položky, parametr schopností.

Kapitola 1

Úvod

Předpokládejme, že člověk má jisté vlastnosti (např. IQ, aritmetické schopnosti, čtenářské dovednosti atd.), které nejsou přímo měřitelné, navenek se zcela projevující. Tyto vlastnosti nazveme latentními vlastnostmi. Dále předpokládejme, že latentní vlastnosti můžeme odvodit z odpovědí na dobře zvolený soubor otázek, tj. latentní vlastnosti lze odvodit na základě dobře sestavených didaktických testů. (Testem v této práci nebude myšlen statistický test, pokud to nebude explicitně napsáno. Testem bude myšlen didaktický test, či jiný test o k otázkách, který je určen ke zjištění latentní vlastnosti dotazovaných žáků.)

V období mezi roky 1940 a 1980 byla velmi rozšířená tzv. klasická teorie testů (zkratka KTT), která byla založena na myšlence rozkladu testového skóre na pravdivé skóre a chybu. Postupně se však stávala stále rošířenější teorie odpovědí na položku (zkratka TOP), která umožnila zvýšit kvalitu testování.

(V tomto odstavci uvedeme některé myšlenky z [6].) Blahuš P. ve svém článku upozorňuje, že sice mnohé z principů klasické teorie testů jsou platné, ale "běžná psychologická praxe v tuzemských podmínkách pořád setrvává poněkud strnule u jeho některých korelačních a regresních metod vyvinutých téměř před sto lety." V praxi se dle něj u nás málo pracuje s modely teorie odpovědí na položky. Přitom tato teorie a další moderní teorie testů umožňují plnit na vyšší úrovni hlavní praktické a klinické cíle nejen psychometrie: "být teoretickým a interdisciplinárním nástrojem pro vytváření kvalitních diagnostických metod." Často se v praxi charakterizují vlastnosti testu pomocí statistických ukazatelů a pak se mezi nimi zkoumají matematické vztahy. Např. validita testu X vzhledem k cílovému kritériu Y se běžně měří pomocí absolutní hodnoty korelačního koeficientu r_{XY}. Důležité však je "prakticky působit na některé ovlivnitelné vlastnosti a tím změnit jiné, cílové vlastnosti testu." Je proto např. důležité sledovat u otázek jejich obtížnost. Vidíme-li, že test je pro danou populaci příliš obtížný, či naopak příliš snadný, nemá téměř žádnou rozlišovací schopnost, a proto ani validitu k jakémukoli účelu.

Teorie odpovědí na položku explicitně modeluje odpověď subjektu na úrovni položky, zatímco klasické modely testu kladou důraz na odpovědi na úrovni skóre testu. TOP se tedy explicitně zabývá charakteristikou jednotlivých položek a zkoumá např. jejich obtížnost. Málo či příliš obtížné otázky lze pak z analýzy vyřadit a zvýšit tak validitu testu.

Dalším a pravděpodobně tím nejpodstatnějším rozdílem mezi klasickou teorií testů a teorií odpovědí na položku je to, že zatímco v KTT existuje tzv. "circular dependency" - tj. odhad obtížnosti položky závisí na míře znalosti subjektu a obráceně odhady míry znalosti subjektu závisí na obtížnosti položek, v TOP mohou být tyto odhady vzájemně nezávislé - viz podkapitola 2.3.

Abychom však klasickou teorii testů pouze nekritizovali, uveďme její přednosti před teorií odpovědí na položku. KTT se používá s poměrně dobrými výsledky při zpracování u nás běžně aplikovaných testů typu multiple-choice na relativně velmi malých souborech, zatímco TOP se více používá na soubory o tisících testovaných subjektů.

Zaměřme se nyní více na teorii odpovědí na položku. Její hlavní myšlenka spočívá v tom, že pravděpodobnost odpovědi osoby na položku v testu může být (ideálně) popsána funkcí latentních vlastností osoby a několika parametrů charakterizující danou položku. Tato pravděpodobnost je pro každou položku nazvána charakteristická křivka položky (anglický ekvivalent je item characteristic curve, zkratka ICC). Popisu této křivky je věnována podkapitola 2.3.

Jednou z hlavních linií TOP představují Raschovy modely. Tato práce je zaměřena na dichotomické Rachovy modely, v nichž jsou odpovědi na jednotlivé položky v testu skórovány buď 1 (správná odpověď) nebo 0 (nesprávná odpověď). Druhá forma Rachových modelů jsou polytomické Rachovy modely, které jsou zaměřeny na více než dvě kategorie odpovědí. Existují i jiné modely TOP s více obecnějšími předpoklady než mají Raschovy modely (předpoklady Rachových modelů - viz podkapitola 2.1), a proto je lze lépe užít na již existující testy. Na druhou stranu se u nich náročněji odhadují parametry a hůře interpretují výsledky.

V případě dichotomických Raschových modelů je charakteristická křivka položky v testu funkcí parametru schopnosti za pevných parametrů charaktrizujících vlastnosti této položky. Popisuje, s jakou pravděpodobností odpoví testovaní s určitým parametrem schopnosti θ na položku správně. V TOP však existují dva možné pohledy na interpretaci této pravděpodobnosti v dané hodnotě latentní vlastnosti θ při dané obtížnosti položky β:

1. Z hlediska jednoho testovaného subjektu - pravděpodobnost popisuje, s jakou pravděpodobností odpoví daná osoba s latentní vlastností θ na danou otázku s obtížností β správně. Jako příklad vezměme $p = 0,6$. Teoreticky: osobě bychom "vymyli" mozek pokaždé, když odpoví na otázku. Pak by platilo, že v šedesáti procentech odpoví správně a ve čtyřiceti procentech chybně.

2. Z hlediska náhodného výběru testovaných subjektů - pravděpodobnost popisuje počet osob, které správně odpoví na otázku s obtížností β ku počtu všech osob odpovídajících na tuto otázku a majících stejnou latentní úroveň vlastností θ.

Ani autoři v rámci jedné knihy týkající se TOP se neshodují na právě jedné z těchto interpretací.

S čím se musíme vypořádat, když chceme výsledek didaktického testu zobecnit? Pokusme se nějakým testem zjistit, zda jsou chlapci lepší čtenáři než dívky. Nechť výsledky měření ukazují, že dívky jsou lepší čtenáři než chlapci. Tento výsledek můžeme zobecnit, jestliže zůstává platný:

1. pro měření stejných dětí stejným testem v jiný den (lze těžko ověřit, děti si odpovědi pamatují, což zkresluje výsledek);

2. pro jiné chlapce a dívky ze stejné populace;

3. pro jiné otázky měřící čtenářskou schopnost.

Výběr osob a výběr otázek jsou dva zdroje náhodností, se kterými se musíme při zobecňování vyrovnat. Teorie odpovědí na položky nabízí lepší možnosti zobecňování výsledků na jiné osoby ze stejné populace testovaných jedinců než klasická teorie testů, která více závisí na složení skupiny právě testovaných jedinců.

Teorie odpovědí na položku se v kruzích českých pedagogů ani psychologů moc nepěstuje. Důvodem může být, že TOP je pro nematematika příliš obtížná. Právě těmto uživatelům je věnována podkapitola 2.3.

Kapitola 2

Odvození Raschova modelu

Tato kapitola bude věnována odvození dichotomického Raschova modelu (zkratka DRM). Na základě podobné myšlenky se odvozují i polytomické Raschovy modely, které jsou zobecněním dichotomických Raschových modelů.

Nejprve se zamyslíme, jak DRM vysvětlují výskyt dichotomicky skórovaných odpovědí (odpovědi skórované 0 nebo 1) n osob na k otázek zaznamenaných v datové matici $\mathbf{x}$ typu $n \times k$.

Mějme n zkoušených osob (testovaných subjektů), označme je $S_1, S_2, \ldots, S_n$. Dále mějme k otázek (položek) v testu, měřící stejnou latentní vlastnost θ, označme je $I_1, I_2, \ldots, I_k$. Předpokládejme, že každá osoba S_ν má svůj parametr θ_ν označující její úroveň latentní vlastnosti. Dále předpokládejme, že každá otázka I_i má svůj parametr β_i označující její obtížnost. Nechť $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ označuje n-rozměrný vektor latentních vlastností osob $S_1, S_2, \ldots, S_n$ a $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$ označuje k-rozměrný vektor obtížností otázek $I_1, I_2, \ldots, I_k$. Odpověd ν-té osoby na i-tou otázku označme $x_{\nu i}$, přičemž $x_{\nu i} = 1$, pokud osoba odpoví správně a $x_{\nu i} = 0$, pokud osoba odpoví chybně. Všechny odpovědi n osob na k otázek zaznamenejme do matice $\mathbf{x_{n\times k}}$. Pravděpodobnost každého prvku této matice je v Raschových modelech určena následovně ($X_{\nu i}$ označuje náhodnou veličinu, $x_{\nu i}$ pozorovanou hodnotu):

$$P(X_{\nu i} = x_{\nu i}|\theta_\nu, \beta_i) = \frac{exp[x_{\nu i}(\theta_\nu - \beta_i)]}{1 + exp[(\theta_\nu - \beta_i)]}. \tag{2.1}$$

Abychom vyjádřili pravděpodobnost celé matice $\mathbf{X_{n\times k}}$ s prvky $X_{\nu i}$ ($x_{n\times k}$ má prvky $x_{\nu i}$) učiňme následující předpoklady nezávislosti:

1. Vektory odpovědí i-té a j-té osoby jsou nezávislé pro všechna $i \neq j$, $i, j = 1, 2, \ldots n$. V praxi může být tento předpoklad porušen tím, že jeden dotazovaný zná odpovědi jiného dotazovaného. Učiňme předpoklad, že máme nejen náhodný výběr osob, ale i vektorů odpovědí.

2. Vezměme v úvahu odpovědi jedné osoby na k otázek, které všechny měří latentní vlastnost θ. Nechť jsou odpovědi na všechny otázky jedné osoby za podmínky θ na sobě nezávislé. Tj., veškerá závislost mezi odpověďmi jedné osoby lze vysvětlit pouze a jen přítomností θ. Tento předpoklad se obvykle nazývá lokální nezávislost. V jistém smyslu mají být tedy otázky homogenní (měřit stejnou latentní vlastnost) a zároveň heterogenní (nemají nic více společného než latentní vlastnost).

Za těchto předpokladů vyjádřeme pravděpodobnost, že matice $\mathbf{X_{n \times k}}$ je rovna pozorovaným hodnotám $\mathbf{x_{n \times k}}$ následovně:

$$P(\mathbf{X_{n \times k}} = \mathbf{x_{n \times k}} | \theta, \beta) = \prod_{\nu=1}^{n} \prod_{i=1}^{k} \frac{exp[x_{\nu i}(\theta_\nu - \beta_i)]}{1 + exp[(\theta_\nu - \beta_i)]}. \tag{2.2}$$

Nyní popišme výhodu Raschových modelů v rámci modelů teorie odpovědí na položku. Raschovy modely vychází z předpokladu, že počet správných odpovědí osoby S_ν je suficientní statistikou pro její neznámý parametr latentní vlastnosti θ_ν a počet správných odpovědí osob na otázku I_i je suficientní statistikou pro její obtížnost β_i. Toto znamená, že v obou případech získáme hodnoty, které obsahují veškerou statistickou informaci o parametrech. Je to zároveň důležité při odhadování parametrů.

Na základě technického předpokladu lokální nezávislosti a řádkového skóre R_ν (R_ν označuje součet prvků v ν-tém řádků matice $\mathbf{X_{n \times k}}$) jako suficientní statistiky pro θ_ν odvodíme v následující pokapitole 2.1 jednorozměrný (zkoumáme jen jednu latentní vlastnost) dichotomický Raschův model.

2.1 Odvození z předpokladu suficience

Označme si v úvodu této podkapitoly parametr latentních vlastností jako ξ, obtížnost i-té otázky jako δ_i a charakteristickou křivku položky (zkratka ICC) i jako $g_i(\xi)$. Dosud jsme neřešili, na jaké škále (stupnici) jsou ξ a δ_i měřeny. Kdykoli mohou být tyto parametry změněny monotónní transformací (např. je můžeme měřit na logaritmické škále), což vede ke změně tvaru ICC. Nechť $i = 1, 2, \ldots, k$. Model s charakteristickými křivkami položky ve tvaru $g_i(\xi)$ nazveme ekvivalentním s modelem s charakteristickými křivkami $f_i(\theta)$, v němž je parametr latentních vlastností označen θ a obtížnost i-té položky označena β_i, jestliže budou existovat takové ryze monotónní transformace (funkce) $\theta \doteq \phi(\xi)$ a $\beta_i = \psi(\delta_i)$, že $g_i(\xi) = f_i(\theta)$. V následující větě 1 bude ukázáno, že θ a β_i vystupují ve funkci f ve tvaru rozdílu, což bude implikovat, jak ukážeme v lemmatu 2, že θ a β jsou měřeny na intervalových škálách. Stručná charakteristika základních typů škál tak, jak jsou uváděny v literatuře (např. [9]), následuje za lemmatem 2.

Věta 1
Nechť jsou splněny následující předpoklady:

1. *charakteristické křivky položek $g_i(\xi)$ jsou spojité ryze rostoucí funkce ξ pro všechny I_i;*
2. *dolní limita $g_i(\xi)$ je $\lim_{\xi\to-\infty} g_i(\xi) = 0$ a horní limita $g_i(\xi)$ je $\lim_{\xi\to\infty} g_i(\xi) = 1$;*
3. *"lokální nezávislost" všech odpovědí, tj. za podmínky testu o k otázkách platí, že*

$$P(X_{\nu 1} = x_{\nu 1}, X_{\nu 2} = x_{\nu 2}, \ldots, X_{\nu k} = x_{\nu k}) = \prod_{i=1}^{k} [g_i(\xi_\nu)]^{x_{\nu i}} [1 - g_i(\xi_\nu)]^{1-x_{\nu i}};$$

4. *za podmínky testu o k otázkách je statistika $R_\nu = \sum_{i=1}^{k} X_{\nu i}$ suficientní statistikou pro ξ_ν,*

pak platí, že model teorie odpovědí na položku je ekvivalentní s Raschovým modelem s charakteristickými křivkami položek:

$$f_i(\theta) = \frac{exp(\theta - \beta_i)}{1 + exp(\theta - \beta_i)}. \tag{2.3}$$

Poznámka: Předpoklad 4 je klíčovým předpokladem.

Důkaz:
Pro přehlednější zápis odstraňme index ν. Náhodný vektor odpovědí označme $\mathbf{X} = (X_1, X_2, \ldots, X_k)$, pozorovaný vektor odpovědí označme $\mathbf{x} = (x_1, x_2, \ldots, x_k)$. (V celém textu nebudeme užívat označování vektoru způsobem: $(x_1, x_2, \ldots, x_k)^T$, tj. nebudeme označovat transponování.)

Vzhledem k tomu, že R je suficientní statistika pro ξ a r je pozorovaná hodnota suficientní statistiky R, pak dle definice suficientní statistiky uvedené v [4] při daném ξ platí, že:

$$\begin{aligned} P(\mathbf{X}=\mathbf{x}|S(\mathbf{X})=r) &= \frac{P(\mathbf{X}=\mathbf{x},S(\mathbf{X})=r)}{P(S(\mathbf{X})=r)} = \frac{p(r,\xi)q(\mathbf{x})}{\sum\limits_{\{\mathbf{x}:S(\mathbf{x})=r\}} P(\mathbf{X}=\mathbf{x})} = \\ &= \frac{p(r,\xi)q(\mathbf{x})}{\sum\limits_{\{\mathbf{x}:S(\mathbf{x})=r\}} p(r,\xi)q(\mathbf{x})} = \frac{q(\mathbf{x})}{\sum\limits_{\{\mathbf{x}:S(\mathbf{x})=r\}} q(\mathbf{x})}, \end{aligned} \tag{2.4}$$

kde $p(r,\xi)$ a $q(\mathbf{x})$ jsou nezáporné funkce. 2.4 platí, je-li $P(S(\mathbf{X})=r)>0$. Dále 2.4 platí, je-li $S(\mathbf{X})=r$, jinak je $P(\mathbf{X}=\mathbf{x}|S(\mathbf{X})=r)$ rovna nule. Podržme tyto dvě podmínky v paměti a ve zbytku práce ji neuvádějme.

Podmíněné rozdělení vektoru $\mathbf{X}$ při známé hodnotě postačující statistiky $S(\mathbf{X})=r$ tedy nezávisí na ξ, je závislé pouze na $\mathbf{x}$. Platí (ξ je daná hodnota):

$$P(\mathbf{X}=\mathbf{x}|S(\mathbf{X})=r) = \frac{P(\mathbf{X}=\mathbf{x}|\xi)}{P(S(\mathbf{X})=r|\xi)} = c(\mathbf{x}). \tag{2.5}$$

Nechť $k \geq 2$ a položka I_i je jakákoli položka s indexem $2 \leq i \leq k$. Definujme:
$\mathbf{x}$ jako vektor odpovědí s řádkovým skórem r, pro nějž platí, že $x_1 = 1$, $x_i = 0$;
$\mathbf{x}^{(1,i)}$ jako vektor $\mathbf{x}$, v němž chybí první a i-tá složka, tj. $(x_2,\ldots,x_{i-1},x_{i+1},\ldots,x_k)$, kde $k>2$;
$\mathbf{y}$ jako vektor odpovědí s řádkovým skórem r, pro nějž platí, že $y_1=0$, $y_i=1$ a v ostatních případech $y_l = x_l$;
$\mathbf{y}^{(1,i)}$ jako vektor $\mathbf{y}$, v němž chybí 1-ní a i-tá složka, tj. $(y_2,\ldots,y_{i-1},y_{i+1},\ldots,y_k)$, kde $k>2$;
Vzhledem k předpokladům platí (místo $\mathbf{X}=\mathbf{x}$ pišme $\mathbf{x}$, místo $S(\mathbf{X})=r$ pišme r):

$$P(\mathbf{x}|\xi) = P(\mathbf{x}^{(1,i)}|\xi)g_1(\xi)[1-g_i(\xi)]. \tag{2.6}$$

Jestliže $k=2$, pak $P(\mathbf{x}^{(1,i)}|\xi) := 1$ ($x := y$ znamená v celém textu, že y vložíme do x). Dosadíme-li 2.6 do 2.5 dostaneme:

$$P(\mathbf{x}|r) = \frac{P(\mathbf{x}^{(1,i)}|\xi)g_1(\xi)[1-g_i(\xi)]}{P(r|\xi)} = c(\mathbf{x}). \tag{2.7}$$

Podobně pro vektor $\mathbf{y}$ platí:

$$P(\mathbf{y}|r) = \frac{P(\mathbf{y}^{(1,i)}|\xi)[1-g_1(\xi)]g_i(\xi)}{P(r|\xi)} = c(\mathbf{y}). \tag{2.8}$$

Dělíme-li pravděpodobnost 2.7 pravděpodobností 2.8 dostaneme:

$$\frac{P(\mathbf{x}^{(1,i)}|\xi)g_1(\xi)[1-g_i(\xi)]}{P(\mathbf{y}^{(1,i)}|\xi)[1-g_1(\xi)]g_i(\xi)} = \frac{c(\mathbf{x})}{c(\mathbf{y})}. \tag{2.9}$$

Vzhledem k definici $\mathbf{x}^{(1,i)} = \mathbf{y}^{(1,i)}$ platí

$$P(\mathbf{x}^{(1,i)}|\xi) = P(\mathbf{y}^{(1,i)}|\xi). \tag{2.10}$$

Dosadíme-li 2.10 do 2.9 dostaneme:

$$\frac{g_1(\xi)[1-g_i(\xi)]}{[1-g_1(\xi)]g_i(\xi)} = \frac{c(\mathbf{x})}{c(\mathbf{y})} =: d_i(\mathbf{x}). \tag{2.11}$$

Prozatím jsme nic nepředpokládali o škále latentní vlastnosti ξ. Pomocí vhodné ryze monotónní transformace $\xi \rightarrow \theta$ vyjádříme prozatím nespecifikovanou charakteristickou křivku položky $g_1(\xi)$ ve tvaru

$$f_1(\theta) = \frac{exp(\theta)}{1+exp(\theta)}. \tag{2.12}$$

Dosaďme do 2.11 za $g_1(\xi)$ a $g_i(\xi)$ postupně $f_1(\theta)$ a $f_i(\theta)$.

$$\frac{\frac{exp(\theta)}{1+exp(\theta)}(1-f_i(\theta))}{(1-\frac{exp(\theta)}{1+exp(\theta)})f_i(\theta)} = \frac{(1-f_i(\theta))exp(\theta)}{f_i(\theta)} = d_i(\mathbf{x}) \tag{2.13}$$

Proveďme transformaci $\beta_i = \ln(d_i(\mathbf{x}))$. Je zřejmé, že β_i závisí pouze a jen na vektoru $\mathbf{x}$. Využijme této transformace a pokračujme v úpravách 2.13:

$$\begin{aligned} f_i(\theta)[exp(\theta)+exp(\beta_i)] &= exp(\theta) \\ f_i(\theta) &= \frac{exp(\theta)}{exp(\theta)+exp(\beta_i)} \\ f_i(\theta) &= \frac{exp(\theta-\beta_i)}{1+exp(\theta-\beta_i)}. \end{aligned} \tag{2.14}$$

Libovolný model splňující předpoklady věty 1 může být transformován do Raschova modelu s parametry otázek $\beta_1 = 0$ a β_i pro $i = 1, \ldots, k$. Tímto máme větu dokázanou.

♣♣♣

Motivace k následujícímu lemmatu 1:
Jestliže $R = \sum_{i=1}^{k} X_i$ je suficientní statistika pro test s otázkami $I_1, I_2, \ldots, I_k$, pak i $R' = \sum_{i=1}^{k-1} X_i$ je suficientní statistika pro test s otázkami $I_1, I_2, \ldots, I_{k-1}$. Tento závěr je intuitivní, protože jediná cesta, jak vstupuje informace o otázkách $I_1, I_2, \ldots, I_{k-1}$ do $R = R' + X_k$ je skrz R'. Formální důkaz suficience R' by byl bez použití věty 1 netriviální.

Lemma 1

Nechť platí předpoklady 1., 2., 3. věty 1. Jestliže je v testu $I_1, I_2, \ldots, I_k$ řádkové skóre $R = \sum_{i=1}^{k} X_i$ suficientní statistikou pro ξ, pak $R_S = \sum_{I_i \in S} X_i$ je suficientní statistikou pro ξ v testu S, kde S označuje neprázdnou podmnožinu otázek $I_1, I_2, \ldots, I_k$.

Důkaz:

(Označme $\mathbf{X}' = (X_1, X_2, \ldots, X_{k-1})$, $\mathbf{x}' = (x_1, x_2, \ldots, x_{k-1})$, $R' = \sum_{i=1}^{k-1} X_i$ a $r' = \sum_{i=1}^{k-1} x_i$.) Budeme chtít dokázat, že $P(\mathbf{X}' = \mathbf{x}' | R' = r')$ je nezávislá na ξ. Pak bude platit, že r' je suficientní statistikou pro ξ v testu $I_1, I_2, \ldots, I_{k-1}$. Po provedení tohoto důkazu bude zjevné, proč je podmíněná pravděpodobnost nezávislá na ξ i v libovolném testu S.

Nechť $\mathbf{x}$ je pozorovaný vektor odpovědí v testu $I_1, I_2, \ldots, I_k$ se skórem r. Nechť $\mathbf{z}$ je libovolný vektor odpovědí se stejným skórem, tj. se skórem r. Protože R je suficientní statistika, platí následující (stejně jako v důkazu věty 1 budeme v celém následujícím textu značit $P(\mathbf{X} = \mathbf{x} | R = r)$ jako $P(\mathbf{x}|r)$ a $P(\mathbf{X}' = \mathbf{x}' | R = r)$ jako $P(\mathbf{x}'|r')$):

$$\begin{aligned} P(\mathbf{x}|r) &= \frac{\prod_{i=1}^{k} g_i^{x_i}(\xi)(1-g_i(\xi))^{1-x_i}}{\sum_{\mathbf{z}} \prod_{i=1}^{k} g_i^{z_i}(\xi)(1-g_i(\xi))^{1-z_i}} = \frac{\prod_{i=1}^{k} (1-g_i(\xi))(\frac{g_i(\xi)}{1-g_i(\xi)})^{x_i}}{\sum_{\mathbf{z}} \prod_{i=1}^{k} (1-g_i(\xi))(\frac{g_i(\xi)}{1-g_i(\xi)})^{z_i}} = \\ &= \frac{\pi_k \prod_{i=1}^{k} h_i^{x_i}(\xi)}{\sum_{\mathbf{z}} \pi_k \prod_{i=1}^{k} h_i^{z_i}(\xi)} = \frac{\prod_{i=1}^{k} h_i^{x_i}(\xi)}{\sum_{\mathbf{z}} \prod_{i=1}^{k} h_i^{z_i}(\xi)} = c(\mathbf{x}), \end{aligned} \tag{2.15}$$

kde $\pi_k = \prod_{i=1}^{k}(1 - g_i(\xi))$ a $h_i(\xi) = \frac{g_i(\xi)}{1-g_i(\xi)}$. Dále označme

$$\gamma_r(h_1(\xi), h_2(\xi), \ldots, h_k(\xi)) = \sum_{\mathbf{z}} \prod_{i=1}^{k} h_i^{z_i}(\xi). \tag{2.16}$$

Tato funkce je nazývána v [7] elementární symetrickou funkcí. Dále se s ní setkáme v podkapitole 3.2 při odvozování odhadu parametrů obtížnosti otázek v testu.

Z 2.15 a 2.16 vyplývá, že

$$\prod_{i=1}^{k} h_i^{x_i}(\xi) = c(\mathbf{x})\gamma_r(h_1(\xi), h_2(\xi), \ldots, h_k(\xi)). \tag{2.17}$$

Vezměme jiný vektor odpovědí $\mathbf{y}$ s řádkovým skórem r v testu $I_1, I_2, \ldots, I_k$. Vztah 2.17 platí zřejmě i pro $\mathbf{y}$, tj.

$$\prod_{i=1}^{k} h_i^{y_i}(\xi) = c(\mathbf{y})\gamma_r(h_1(\xi), h_2(\xi), \ldots, h_k(\xi)). \tag{2.18}$$

Podělme vztah 2.18 vztahem 2.17, pak dostáváme:

$$\frac{\prod_{i=1}^{k} h_i^{y_i}(\xi)}{\prod_{i=1}^{k} h_i^{x_i}(\xi)} = \frac{c(\mathbf{y})}{c(\mathbf{x})} \tag{2.19}$$

Položme $x_1 = y_j = 1$; $x_j = y_1 = 0$ a v ostatních případech $x_l = y_l$. Nechť je řádkové skóre těchto vektorů rovno r. Po dosazení těchto vektorů do vztahu 2.19, který platí pro všechny vektory odpovědí s řádkovým skórem r, dostáváme

$$\begin{aligned} \frac{h_j(\xi)}{h_1(\xi)} &= \frac{c(\mathbf{y})}{c(\mathbf{x})} =: d_j(\mathbf{x}) \\ h_j(\xi) &= d_j(\mathbf{x})h_1(\xi) \qquad j = 1, 2, \ldots, k. \end{aligned} \tag{2.20}$$

Tímto jsme si připravili půdu pro důkaz, že $P(\mathbf{x}'|r')$ je nezávislá na ξ, (nechť $z' = (z_1, z_2, \ldots, z_{k-1})$ je libovolný vektor odpovědí s řádkovým skórem r'):

$$\begin{aligned} P(\mathbf{x}'|r') &= \frac{\prod_{i=1}^{k-1} g_i^{x_i}(\xi)(1-g_i(\xi))^{1-x_i}}{\sum_{\mathbf{z}'}\prod_{i=1}^{k-1} g_i^{x_i}(\xi)(1-g_i(\xi))^{1-x_i}} = \frac{\prod_{i=1}^{k-1} h_i^{x_i}(\xi)}{\sum_{z'}\prod_{i=1}^{k-1} h_i^{z_i'}(\xi)} = \\ &\overset{2.20, d_1=1}{=} \frac{\prod_{i=1}^{k-1} (d_i(\mathbf{x})h_1(\xi))^{x_i}}{\sum_{z'}\prod_{i=1}^{k-1} (d_i(\mathbf{x})h_1(\xi))^{z_i'}} = \frac{h_1^{r'}(\xi)\prod_{i=1}^{k-1} d_i(\mathbf{x})}{h_1^{r'}(\xi)\sum_{\mathbf{z}'}\prod_{i=1}^{k-1} d_i(\mathbf{x})} = \end{aligned}$$

$$= \frac{\prod_{i=1}^{k-1} d_i(\mathbf{x})}{\gamma_{r'}(d_1(\mathbf{x}), d_2(\mathbf{x}), \ldots, d_{k-1}(\mathbf{x}))}, \tag{2.21}$$

což je funkce pouze a jen $\mathbf{x}$.

Je zřejmé, že důkaz platí pro jiné subtesty délky $k-1$ stejně jako pro testy délky $k-2$ atd. Tímto je lemma 1 dokázáno.

♣♣♣

Ve větě 1 jsme došli k závěru, že jakýkoli model splňující předpoklady této věty, může být transformován do Raschova modelu. V charakteristických křivkách položek vystupují právě v Raschově modelu parametry θ a β v podobě rozdílu $(\theta - \beta)$. Předpokládejme, že bychom se hned od počátku rozhodli přidat následující předpoklad.

Předpoklad 5.: Nechť existuje systém otázek s charakteristickými křivkami položek ve tvaru $f(\theta - \beta)$ pro $\forall\theta \in \mathbf{R}$ a $\forall\beta \in \mathbf{R}$, kde β je obtížnost položky. Funkce $f(x)$ je spojitá, rostoucí funkce $f : \mathbf{R} \to (0, 1)$.

Motivace k lemmatu 2: Jaká je pak obecná třída transformací měřítka θ a β, která by vyhovovala tomu, že by po ní vystupovali transformované parametry θ a β v charakteristických křivkách položek v podobě rozdílu?

Lemma 2

Nechť pravděpodobnosti $P(+|\theta, \beta) = f(\theta - \beta)$ jsou pevně dány pro $\forall\theta \in \mathbf{R}$ a $\forall\beta \in \mathbf{R}$. Pak z předpokladu 5. vyplývá, že všechny transforamce θ a β lze zapsat ve tvaru $a\theta + b_1$ a $a\beta + b_2$, kde $a > 0$, $b_1 \in \mathbf{R}$ a $b_2 \in \mathbf{R}$.(Tj. θ a β jsou měřené na intervalových škálách se společnou jednotkou měření.)

Důkaz:

Předpokládejme model s charakteristickými křivkami ve tvaru $f(\theta - \beta)$. Z toho vyplývá, že $P(+|\theta, \beta)$ je pevně dána pro $\forall\theta \in \mathbf{r}$ a $\forall\beta \in \mathbf{R}$.

Domnívejme se, že existuje jiné vyjádření $P(+|\theta, \beta)$ s novými transformovanými parametry $k(\theta)$ a $l(\beta)$ ve tvaru $m(k(\theta) - l(\beta))$, pro nějž platí

$$m(k(\theta) - l(\beta)) = f(\theta - \beta) \qquad \forall\theta \in \mathbf{R}, \forall\beta \in \mathbf{R}, \tag{2.22}$$

kde $k(\theta)$ a $l(\beta)$ jsou spojité rostoucí transformace měřítka θ a β a $m(x)$ je definována vztahem 2.22. Víme tedy, že $m(x)$ je ryze monotónní, a proto k ní existuje inverzní funkce, čehož využijeme:

$$k(\theta) - l(\beta) = m^{-1} \circ f(\theta - \beta), \tag{2.23}$$

kde symbol $\circ$ znamená skládání funkcí. Položme $\beta \equiv 0$, pak úpravou 2.23 dostáváme:

$$\begin{aligned} k(\theta) &= m^{-1} \circ f(\theta) + l(0) \\ k(\theta) &= h(\theta) + l_0, \end{aligned} \tag{2.24}$$

kde $h = m^{-1} \circ f$ a $l_0 = l(0)$. Podobně položme $\theta \equiv 0$, pak úpravou 2.23 dostáváme:

$$\begin{aligned} l(\beta) &= -m^{-1} \circ f(-\beta) + k(0) \\ l(\beta) &= -h(-\beta) + k_0, \end{aligned} \tag{2.25}$$

kde $k_0 = k(0)$. Položme $\theta \equiv 0$ a zároveň $\beta \equiv 0$, pak úpravou 2.23 dostáváme:

$$k_0 - l_0 = h_0, \tag{2.26}$$

kde $h_0 = h(0)$. Dosaďme 2.24 a 2.25 do 2.23 a zároveň využijme vztahu 2.26, pak dostáváme:

$$\begin{aligned} h(\theta) + l_0 + h(-\beta) - k_0 &= h(\theta - \beta) \\ h(\theta) - h_0 + h(-\beta) - h_0 &= h(\theta - \beta) - h_0. \end{aligned} \tag{2.27}$$

Položme pro zjednodušení $s := \theta$, $t := -\beta$ a $h^{\star}(x) := h(x) - h_0$, dosazením do 2.27 dostáváme:

$$h^{\star}(s) + h^{\star}(t) = h^{\star}(s + t) \tag{2.28}$$

Dostali jsme funkcionální rovnici, která má dle [1] obecné řešení ve tvaru

$$h^\star = ax \qquad a > 0, \tag{2.29}$$

přičemž $a > 0$, protože h je rostoucí funkce (vznikla složením dvou rostoucích funkcí). Pro funkci tedy $h(x)$ platí

$$h(x) = ax + h_0. \tag{2.30}$$

Konečně, transformační funkce $k(\theta)$ a $l(\beta)$ dostaneme dosazením 2.30 do 2.24 a 2.25:

$$k(\theta) = a\theta + h_0 + l_0 = a\theta + b_1 \tag{2.31}$$

$$l(\beta) = a\beta - h_0 + k_0 = a\beta + b_2. \tag{2.32}$$

Vidíme, že všechny přípustné transformace stupnic, na nichž měříme θ a β, jsou ve tvaru $a\theta + b_1$ a $a\beta + b_2$, kde $a > 0$, $h_0 + l_0 = b_1 \in \mathbf{R}$ a $-h_0 + k_0 = b_2 \in \mathbf{R}$. Tímto jsme dokázali lemma 2.

♣♣♣

Zjistili jsme, že pro modely splňující předpoklady 1. až 4. věty 1 platí, že θ a β jsou měřeny na intervalových škálách se společnou měřící jednotkou.

Jaké typy škál ještě známe a jaké z nich jsou nejpřesnější? V knize [9] jsou uvedeny čtyři základní úrovně měření:

1. nominální měření

   Někteří autoři nepovažují nominální měření za měření, ale hovoří spíše o kategorizaci. V tomto případě se data roztřídí do vzájemně se vylučujících kategorií. Každý prvek lze zařadit do jedné kategorie. Příkladem může být roztřídění lidí na muže a ženy, tj. na dvě kategorie.

2. pořadové (ordinální) měření

   V řadě případů nedokážeme v pedagogickém výzkumu změřit přesné hodnoty. Např. se jednoznačně nepodaří změřit píli žáků, jejich snahu atd. Lepším řešením pak je porovnat snahu žáka s ostatními žáky. Žáci se seřadí podle kritéria a na tomto základu je jim přiděleno pořadí. Lze stanovit pořadový vztah, tj. jev B je menší než jev C, ale větší než jev A. Nevíme však, jak velký je interval mezi jevy A,B a B,C.

3. intervalové měření

   Tento typ měření představuje již typ metrického měření, lze stanovit (změřit), o kolik je A lepší než B, lze změřit jejich vzdálenost. Tento způsob měření nemá nulový bod. Nula však může být stanovena výzkumníkem. Např. u schopností lze těžko stanovit nulovou schopnost, každý člověk je nějak schopný. Protože se jedná o metrické měření, lze čísla získaná tímto typem měření mezi sebou sčítat a odčítat, nikoliv však násobit a dělit. Nelze říci, že někdo je "dvakrát schopnější než někdo jiný". Připojme typový příklad intervalového měření - měření teploty. Při měření ve stupnici Celsiově bychom mohli jeden den naměřit 10 stupňů, ve stupnici Fahrenheita 50 stupňů. Další den bychom mohli ve stupnici Celsiově naměřit 20 stupňů, ve stupnici Fahrenheita 68 stupňů. Zatímco teplota na stupnici Celsiově byla druhý den dvakrát větší než první den, ve stupnici Fahrenheita tomu rozhodně tak není.

4. poměrové měření

   Tento typ měření je ze všech uvedených měření nejpřesnější. Naměřené hodnoty jsou ukazatelem množství, míry vlastnosti. Patří sem např. měření výšky, váhy, elektrického proudu atd. Hodnoty, které získáme tímto typem měření lze mezi sebou násobit a dělit.

Poměrové měření je sice nejpřesnější, ale ve společenských vědách ho lze využít jen velmi málo. Parametry našeho modelu lze měřit na intervalových škálách, což odpovídá poměrně velké kvalitě měření.

Použijeme-li předpokladů 1. až 5. získáme rodinu Raschových modelů ve tvaru:

$$P(+|S_\nu, I_i) = \frac{exp[a(\theta_\nu - \beta_i) + b]}{1 + exp[a(\theta_\nu - \beta_i) + b]}, \tag{2.33}$$

kde $a > 0$ a $b := b_1 - b_2$ je prvkem $\mathbf{R}$. Interpretace parametrů je následující (podrobněji v podkapitole 2.3):

a... diskriminační parametr všech otázek, který rozhoduje o tom, jak dobře rozlišují otázky mezi horšími a lepšími studenty;

b... definuje vztah mezi počátky β- a θ- škály. Je zřejmé, že libovolný posun θ- škály může být kompenzován posunem β- škály a naopak, tj. b může být rovno nule. Kdybychom dosadili do 2.33 právě hodnotu $b = 0$, dostali bychom model, který je v literatuře (např. v [5], [7] a [12]) znám jako dvouparametrický logistický model.

Předcházející úvahy shrňme do následující věty 2.

Věta 2

Předpoklady 3. a 4. ve větě 1 a předpoklad 5. implikují, že model je Raschův model a parametry θ a β jsou jedinečné až na kladné lineární transformace se společnou multiplikativní konstantou, tj. mají vlastnosti intervalových škál se společnou jednotkou měření.

Poznámka: Předpoklad 5. zahrnuje předpoklady 1. a 2. ve větě 1.

Důkaz:

Vyplývá z předcházejících úvah.

♣♣♣

Motivace k následujícímu lemmatu 3.: Předpoklad 5. je obecně velmi silný, "potřebuje" systém otázek s charakteristickými křivkami ve tvaru $f(\theta - \beta)$ pro $\forall\theta \in \mathbf{R}$ a $\forall\beta \in \mathbf{R}$. Pokusme se tento předpoklad zeslabit a uvažujme systém, který se skládá z konečného počtu otázek. Výsledek o vlastnostech škál θ a β pak bude nepatrně oslaben, avšak z praktického hlediska nemá téměř žádný význam. Pro úplnost odvození Rachových modelů a jeho vlastností si tento výsledek budeme přesně formulovat a dokazovat.

Předpoklad 6.: Nechť je systém otázek složen z konečného počtu otázek $k \geq 3$, které mají charakteristické křivky položek ve tvaru $f(\theta - \beta_i)$ pro $\forall\theta \in \mathbf{R}$ a $i = 1, 2, \ldots, k$, $f : \mathbf{R} \to (0, 1)$ je spojitá rostoucí funkce.

Lemma 3

Nechť platí předpoklad 6..

1. *Jestliže všechny koeficienty $\frac{\beta_j - \beta_i}{\beta_q - \beta_p}$ ($\beta_i \neq \beta_j$ a $\beta_p \neq \beta_q$) jsou racionální čísla, pak ve všech bodech $Q_{in} = \beta_i + \Delta n$ platí, že θ jsou jedinečné až na kladné lineární transformace. Ve všech bodech mezi Q_{in} jsou θ jedinečné až na spojité, ryze monotónní transformace. Body $Q_{in} = \beta_i + \Delta n$ jsou definovány pro $n = 0, 1, \ldots,$ dále $i, j, p, q = 1, 2, \ldots, k$, $i \neq j \neq p \neq q$. Δ je největší reálné číslo splňující $\beta_j - \beta_i = \omega_{ji}\Delta$, kde $\omega_{ji} \in \mathbf{Z}$.*

2. *Jestliže máme alespoň tři otázky I_i, I_j, I_l takové, že $\frac{\beta_j-\beta_i}{\beta_l-\beta_j}$ ($\beta_i < \beta_j < \beta_l$) jsou iracionální čísla, pak θ-škála je intervalová škála.*

Důkaz:

Funkcionální rovnici 2.22 přepišme pro případ konečného počtu otázek $I_1, I_2, \ldots, k$ s pevnými parametry obtížnosti $\beta_1, \beta_2, \ldots, \beta_k$ do tvaru

$$m(k(\theta) - b_i) = f(\theta - \beta_i) \qquad \forall\theta \in \mathbf{R} \quad i = 1, 2, \ldots, k, \tag{2.34}$$

kde b_i jsou konstanty. Zaměříme se na hledání $k(\theta)$. Nejprve však odvodíme vlastnost funkce $h(x)$, která stejně jako v lemmatu 2 je rovna složení funkcí m^{-1} a f, tj. (upravíme 2.34):

$$\begin{aligned} k(\theta) &= m^{-1} \circ f(\theta - \beta_i) + b_i \qquad \forall\theta \in \mathbf{R} \quad i = 1, 2, \ldots, k \\ k(\theta) &= h(\theta - \beta_i) + b_i \qquad \forall\theta \in \mathbf{R} \quad i = 1, 2, \ldots, k. \end{aligned} \tag{2.35}$$

Pro libovolný pár otázek I_i, I_j, s obtížnostmi $\beta_i \neq \beta_j$ díky 2.35 platí (budeme uvažovat $\forall\theta \in \mathbf{R}$ a $i \neq j$, přičemž $i, j = 1, 2, \ldots, k$):

$$\begin{aligned} h(\theta - \beta_i) + b_i &= h(\theta - \beta_j) + b_j \\ h(\theta - \beta_i) &= h(\theta - \beta_i) + (b_j - b_i) \\ h(x) &= h(x - (\beta_j - \beta_i)) + (b_j - b_i). \end{aligned} \tag{2.36}$$

Pokud budeme h aplikovat samu na sebe v tom smyslu, že $y := x - (\beta_j - \beta_i)$ a dosadíme do 2.36 dostaneme:

$$\begin{aligned} h(y) &= h(y - (\beta_j - \beta_i)) + (b_j - b_i) \\ h(x - (\beta_j - \beta_i)) &= h(x - 2(\beta_j - \beta_i)) + (b_j - b_i) \\ h(x) - (b_j - b_i) &= h(x - 2(\beta_j - \beta_i)) + (b_j - b_i) \\ h(x) &= h(x - 2(\beta_j - \beta_i)) + 2(b_j - b_i) \end{aligned} \tag{2.37}$$

Provedeme-li tento krok n-krát, dostaneme:

$$h(x) = h(x - n(\beta_j - \beta_i)) + n(b_j - b_i) \qquad n = 0, 1, \dots \tag{2.38}$$

Kdybychom vzali $z := x - (\beta_j - \beta_i)$ a aplikovali bychom na něj n-krát h, stejně jako jsme aplikovali h na y, dostaneme:

$$h(x) = h(x + n(\beta_j - \beta_i)) - n(b_j - b_i) \qquad n = 0, 1, \dots. \tag{2.39}$$

Spojením 2.38 a 2.39 dostáváme:

$$h(x) = h(x - n(\beta_j - \beta_i)) + n(b_j - b_i) \qquad n \in \mathbf{Z}. \tag{2.40}$$

Mějme jiný pár otázek I_p, I_q s $\beta_p \neq \beta_q$, vzhledem k 2.36 a 2.40 platí:

$$h(x) = h(x - n(\beta_j - \beta_i) - m(\beta_q - \beta_p)) + n(b_j - b_i) + m(b_q - b_p) \qquad n \in \mathbf{Z} \quad m \in \mathbf{Z}. \tag{2.41}$$

2.41 lze zobecnit na tvar:

$$h(x) = h(x - \sum_{i<j} n_{ji}(\beta_j - \beta_i)) + \sum_{i<j} n_{ji}(b_j - b_i) \qquad n_{ji} \in \mathbf{Z} \tag{2.42}$$

2.42 říká, že pokud se posuneme z bodu x o $\sum_{i<j} n_{ji}(\beta_j - \beta_i)$ zvětší se funkční hodnota h o $\sum_{i<j} n_{ji}(b_j - b_i)$, kde $n_{ji} \in \mathbf{Z}$. Rozlišíme dva případy dle toho, zda je zlomek $\frac{\beta_j - \beta_i}{\beta_q - \beta_p}$ racionální či iracionální číslo a využijeme 2.42.

1. Předpokládejme, že nejprve platí:

$$\frac{\beta_j - \beta_i}{\beta_q - \beta_p} \in \mathbf{Q} \qquad \forall I_i, I_j \quad \forall I_p, I_q. \tag{2.43}$$

Tvrzení 1

Položme $\Delta > 0$ největší reálné číslo splňující $\beta_j - \beta_i = \omega_{ji}\Delta$ pro všechny páry I_i, I_j, kde $\omega_{ji} \in \mathbf{Z}$, pak největší společný dělitel ω_{ji} je 1.

Konstrukce:

Zkonstruujme si pro názornost takové Δ a takové ω_{ji}, o kterých se hovoří ve tvrzení:

Nechť $\Delta_1 = |\beta_q - \beta_p|$ je nejmenší vzdálenost hodnot dvou obtížností. Z předpokladu 2.43 vyplývá, že $\forall i \forall j$ platí, že $\beta_j - \beta_i = r_{jiqp}(\beta_q - \beta_p)$, přičemž $r_{jiji} = 1$. Protože $|\beta_j - \beta_i| > |\beta_q - \beta_p|$ je $|r_{jiqp}| > 1$. Převeďme r_{jiqp} na zlomky se stejným jmenovatelem J. Mají-li čitatelé těchto zlomků společného dělitele většího než jedna, nalezněme toho největšího z nich a označme ho D. Vzhledem k tomu, že D je dělitel i jmenovatele J (jeden ze zlomků je $\frac{J}{J}$), podělíme tímto D čitatele i jmenovatele každého ze zlomků. Pak už mají čitatelé největšího společného dělitele 1. Tito čitatelé nechť jsou ω_{ji} a $\Delta := \frac{\Delta_1}{J/D}$. Kdyby existovalo $\Delta^\star > \Delta$ splňující předpoklad tvrzení, pak $\Delta^\star = k\Delta$, kde $k > 1$, což implikuje, že $\omega^\star_{ji} = \frac{\omega_{ji}}{k}$. Kdyby $k \in \mathbf{R} - \mathbf{Z}$, pak $\omega^\star_{ji} \notin \mathbf{Z}$. Kdyby $k \in \mathbf{Z} - 1$, pak $\omega^\star_{ji} \in \mathbf{Q}$. Kdyby $k = 1$, pak $\omega^\star_{ji} = \omega_{ji}$. Tudíž $\Delta^\star_{ji} = \Delta_{ji}$. Tímto jsou Δ a ω_{ji} jednoznačně zkounstruována.

$\diamondsuit$

V 2.42 nahraďme $(\beta_j - \beta_i)$ výrazem $\omega_{ji}\Delta$, dostaneme:

$$h(x) = h(x - \Delta \sum_{i<j} n_{ji}\omega_{ji}) + \sum_{i<j} n_{ji}(b_j - b_i). \tag{2.44}$$

Dle Scholz [10] má diofantovská rovnice

$$\sum_{i<j} n_{ji}\omega_{ji} = 1 \tag{2.45}$$

celočíselné řešení $n^\star_{ji}$, jestliže ω_{ji} nemají většího společného dělitele než 1. Dle tvrzení "naše" ω_{ji} skutečně nemají většího společného dělitele než jedna. Dosaďme 2.45 do 2.44:

$$h(x) = h(x - \Delta) + d, \tag{2.46}$$

kde $d = \sum_{i<j} n_{ji}(b_j - b_i)$. Úpravou 2.46 zjistíme periodu $h^\star(x) = h(x) - \frac{d}{\Delta}x$:

$$\begin{aligned} h(x) - \frac{d}{\Delta}x &= h(x - \Delta) - \frac{d}{\Delta}(x - \Delta) \\ h^\star(x) &= h^\star(x - \Delta), \end{aligned} \tag{2.47}$$

takže je zřejmé, že perioda $h^\star$ je Δ. Funkce $h(x)$ je součet $h^\star(x)$ a lineární funkce $\frac{d}{\Delta}x$. Funkce "h" vypadá jako "nekonečně mnoho stejných oblouků spojených body (nazveme je Q_{in}), které leží na přímce se směrnicí:

$$c := \frac{d}{\Delta} = \frac{\sum\limits_{i<j} n^{\star}_{ji}(b_j - b_i)}{\sum\limits_{i<j} n^{\star}_{ji}(\beta_j - \beta_i)} = \frac{b_j - b_i}{\beta_j - \beta_i} = \frac{b_q - b_p}{\beta_q - \beta_p} \qquad \forall I_i, I_j, I_p, I_q." \tag{2.48}$$

Vzhledem k tomu, že transformační funkce $k(\theta)$ má tvar 2.35, platí pro ní, že ve všech bodech $Q_{in} = \beta_i + \Delta n$, $n \in \mathbf{Z}$, $i = 1, 2, \ldots, k$ lze provádět pouze kladnou lineární transformaci, mezi body Q_{in} je transformace libovolná (předpokládá se jen její ryze monotónní vlastnost). Jinými slovy, platí závěr uvedený v lemmatu 3 v bodu 1. Tímto je bod 1 lemmatu 3 dokázán.

2. Mějme nejméně tři otázky I_i, I_j, I_l s $\beta_i < \beta_j < \beta_l$ takové, že

$$\frac{\beta_j - \beta_i}{\beta_l - \beta_j} \in \mathbf{R} - \mathbf{Q}. \tag{2.49}$$

Dokážeme, že neexistuje $\Delta > 0$. h je pak ve tvaru $h(x) = cx + h_0$ jako v lemmatu 2. Definujme $\Delta(n, m) = n(\beta_j - \beta_i) - m(\beta_l - \beta_j)$ pro $n \in \mathbf{Z}$ a $m \in \mathbf{Z}$. S použitím 2.46 a 2.48 dostáváme

$$h(x) = h(x - \Delta(n, m)) + c\Delta(n, m). \tag{2.50}$$

$|\Delta(n, m)|$ je perioda $h^{\star}(x)$. Neexistuje však kratší perioda? Pokud zjistíme, že funkce:

$$M_N := min\{|\Delta(n, m)| : \quad 1 \leq |n| \leq N, \quad 1 \leq |m| \leq N\} \tag{2.51}$$

jde k 0 pro $N \to \infty$, máme dokázáno, že transformační funkce je skutečně pouze lineární funkcí.

Je zřejmé, že M_N je nerostoucí funkce a že $M_N \geq 0$. Existuje dolní mez této funkce $\inf M_N$. Dle definice infima platí, že pokud zvolíme libovolné číslo větší než infimum (BÚNO zvolme $\inf M_N + 10^{-1} \inf M_N$), můžeme nalézt hodnotu M_N mezi $\inf M_N$ a $\inf M_N + 10^{-1} \inf M_N$. Zvolme

$$n_0, m_0 : \inf M_N \leq \Delta(n_0, m_0) < \inf M_N + 10^{-1} \inf M_N \tag{2.52}$$

Předpokládejme (pro spor), že $\inf M_N > 0$. Najdeme takovou periodu, která je menší než $\inf M_N$. BÚNO zvolme $n_0 \geq 1$ (s opačnou nerovností analogicky). Definujme

$$\Delta^{\star} = n_0(\beta_j - \beta_i) - |t|\Delta(n_0, m_0), \tag{2.53}$$

kde $t \in \mathbf{Z}$ je největší takové číslo, aby $\Delta^\star > 0$ $(t \geq 1)$. Protože $(\beta_j - \beta_i)$ je perioda a $\Delta(n_0, m_0)$ je perioda je i $\Delta^\star$ perioda, tj. platí

$$h(x) = h(x \mp \Delta^\star) \pm c\Delta^\star. \tag{2.54}$$

Definujme $p : \Delta^\star = p\Delta(n_0, m_0)$, pak pro p platí $0 < p < 1$. Proč?

- $p > 0$, protože $\Delta^\star > 0$;
- Je-li $p > 1$, pak $\Delta^\star - \Delta(n_0, m_0) > 0$, tj. t není největší;
- Zbývají dvě možnosti $p = 1$ a $0 < p < 1$. Kdyby $p = 1$, pak $\Delta^\star = n_0(\beta_j - \beta_i) - t\Delta(n_0, m_0) = \Delta(n_0, m_0)$, přičemž $\Delta(n_0, m_0) = n_0(\beta_j - \beta_i) - m_0(\beta_l - \beta_j)$. Jinými slovy $tn_0(\beta_j - \beta_i) - (t+1)m_0(\beta_l - \beta_j) = 0$, což nelze protože $\frac{\beta_j - \beta_i}{\beta_l - \beta_j}$ je iracionální. Tedy $0 < p < 1$.

Nalezli jsme novou periodu $p\Delta(n_0, m_0)$, tj.

$$h(x) = h(x \mp p\Delta(n_0, m_0)) \pm cp\Delta(n_0, m_0). \tag{2.55}$$

Toto implikuje, že

$$\begin{aligned} h(x) &= h(x - p\Delta(n_0, m_0) + \Delta(n_0, m_0)) + cp\Delta(n_0, m_0) - c\Delta(n_0, m_0) = \\ &= h(x + (1-p)\Delta(n_0, m_0)) - c(1-p)\Delta(n_0, m_0), \end{aligned} \tag{2.56}$$

tedy $(1-p)\Delta(n_0, m_0)$ je též perioda.

Proveďme následující úvahu: odečteme-li od $\Delta(n_0, m_0)$ číslo větší než $10^{-1} \inf M_N$ bude vzhledem k 2.52 platit, že se dostaneme pod číslo $\inf M_N$. Když $p\Delta(n_0, m_0) < 10^{-1} \inf M_N$, pak $(1-p)\Delta(n_0, m_0) > (1 - 10^{-1}) \inf M_N > 10^{-1} \inf M_N$. Když $(1 - p)\Delta(n_0, m_0) < 10^{-1} \inf M_N$, pak $p\Delta(n_0, m_0) > \Delta - 10^{-1} \inf M_N > (1 - 10^{-1}) \inf M_N > 10^{-1} \inf M_N$. Z toho vyplývá, že buď $p\Delta(n_0, m_0) < \inf M_N$ nebo $(1-p)\Delta(n_0, m_0) < \inf M_N$, tj. nalezli jsme periodu menší než infimum, což je spor. Z toho vyplývá, že $\inf M_N = 0$, tj. pokud $N \to \infty$ tak $M_N \to 0$, tj. $h(x) = cx + h_0$. Tímto je dokázán bod 2 v lemmatu 3.

Tímto je lemma 3 dokázáno.

♣♣♣

Tímto jsme odvodili Raschův model pro dichotomické veličiny ze suficience a popsali jeho základní vlastnosti.

Modely latentní struktury byly důležitým krokem vpřed v klasické teorii testování. Jejich základem byla analýza skálogramu testu či Guttmanova škála. Další pokapitola je věnována právě jí, protože obsahuje důležité myšlenky pro vznik "dokonalejší" teorie Raschových modelů. "Dokonalejší" v tom smyslu, že se stanovují nezávisle na sobě parametry schopností subjektu a obtížnosti úloh.

2.2 Guttmanův model - "předchůdce" Raschova modelu

Vznik Guttmanova modelu byl ovlivněn některými měřeními ze sociologických či psychologických dotazníků, kde se sledované znaky nedaly uchopit metricky. Klasická teorie testů vyžaduje uchopování znaků metrickým způsobem.

Guttman vychází z toho, že daná úroveň vlastnosti jedince by se měla, s přihlédnutím k obtížnosti řešené úlohy, manifestovat v chování subjektu. Tedy - na základě reakce subjektu s určitou úrovní schopnosti na různé otázky by se mělo dát zpětně usoudit na jeho úroveň schopnosti. Na základě toho formuloval model vycházející z těchto předpokladů:

1. Vychází se z množiny k dichotomických veličin $X_1, X_2, \ldots, X_k$ (položek v testu) schopných měřit stejnou latentní vlastnost;

2. Každému testovanému přísluší ve vztahu k latentní vlastnosti parametr s_i, který lze interpretovat jako určitý stupeň latentní vlastnosti;

3. Každá veličina (položka v testu) j je doprovázena parametrem o_j, který lze interpretovat jako úroveň obtížnosti úlohy j.

S ohledem na tyto předpoklady formuluje Guttman svůj model: Subjekt i vyřeší úlohu j právě tehdy, když jeho parametr schopností s_i je větší než parametr obtížnosti položky o_j, tj.

$$\begin{aligned} P(X_j = 1|s_i) &= 1 \qquad s_i \geq o_j, \\ &= 0 \qquad s_i < o_j. \end{aligned} \tag{2.57}$$

Tímto položka j rozděluje testované jedince na dvě skupiny - ty, kteří mají $s_i \geq o_j$ a ty, kteří mají $s_i < o_j$. Uspořádává je tak způsobem odpovídajícímu ordinálnímu škálování. Je totiž jasné, že při užití více veličin (úloh, položek) bude latentní dimenze (na ní měříme latentní vlastnost) rozdělena na větší počet částí. Každá část by pak odpovídala určitému typu chování (odpovědi) zkoušeného: ti dokáží řešit úlohy nižšího stupně obtížnosti a nedokáží řešit úlohy vyššího stupně obtížnosti.

Výsledek měření, tj. přiřazení zkoušenému určitou oblast schopnosti, závisí zřejmě jen na jeho chování a ne na chování ostatních zkoušených. V tomto smyslu je výsledek měření nezávislý na souboru zkoušených, což je velice důležitý výsledek v oblasti měření.

Gutmanovo měření je v jistém smyslu deterministické. Odpověď subjektu na jakoukoli položku je zcela určena jeho úrovní schopností. Nás by však zajímalo: zvětší-li se parametr schopnosti o nějakou hodnotu, o kolik pak vzroste pravděpodobnost správné odpovědi na určitou položku. Toto řeší teorie odpovědí na položky.

Guttman podpořil nový směr měření psychických vlastností, který je založen na binárních datech a z kterého čerpá teorie odpovědí na položky. Na základě jeho modelu si můžeme uvědomit souvislosti mezi manifestními pozorováními a latentními vlastnostmi.

2.3 Vlastnosti charakteristické křivky položky

(V této pokapitole budeme čísla v intervalu oddělovat středníkem, desetinnou čárku budeme značit jako ",", ve výčtu čísel budou čísla oddělena středníkem.) V podkapitole 2.1 jsme se věnovali formálnímu odvození Raschových modelů, které by měli využívat naši pedagogové. Je jasné, že pedagog nematematik se v odvozování ztratí úplně. Učitelé matematici to zřejmě také po několika stránkách vzdají.

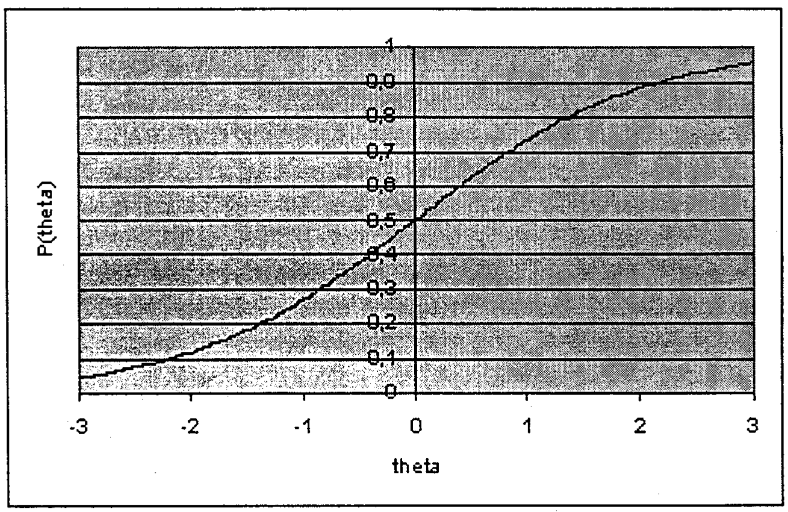
Chtěla bych proto tuto podkapitolu věnovat právě pedagogické veřejnosti a teoretické základy důležité pro pochopení této teorie sepsat co nejjednodušeji to půjde.

Test má měřit určitou schopnost u lidí. Je sestaven z k otázek měřících právě tuto schopnost. Každý testovaný disponuje určitou úrovní této schopnosti. Cílem vždy je, zjistit (odhadnout) tuto schopnost. Označme si v souladu s předchozím textem parametr schopností (úroveň schopností) θ. Pro testovaného s určitou úrovní schopností existuje jistá pravděpodobnost, že odpoví na danou otázku správně. Tuto pravděpodobnost si označme v souladu s předchozím $f(\theta)$. Tato pravděpodobnost má být malá pro testovaného s malou úrovní schopnosti (s malým θ) a velká pro testovaného s vysokou úrovní schopnosti (tj. čím schopnější je testovaný, tím spíše odpoví na danou otázku správně.)

Tvar této křivky je pro každou otázku jiný, což je dáno určitými vlastnostmi otázek, jako je jejich obtížnost, diskriminační vlastnost atd. Právě na základě toho, jaké vlastnosti otázek bereme v úvahu při charakterizování této křivky, rozlišujeme v toerii odpovědí na položky tři základní modely:

1. Jednoparametrický logistický model - při popisu charakteristické křivky položky bere v úvahu její obtížnost. (Známe-li její obtížnost, víme jednoznačně, jak bude tato křivka vypadat.)

   Pro modelování tvaru charakteristické křivky položky se užívá logistická funkce, která připomíná písmeno "S" - viz obrázek 2.1. Poprvé byla v TOP použita roku 1950 věd-



Obrázek 2.1: Typický tvar charakteristické křivky

cem Birnbauem. Logistická funkce nahradila do té doby oblíbenou distribuční funkci normálního rozdělení, protože (jak uvidíme na obrázkách 2.2, 2.3 a 2.4) parametry položky lze graficky dobře interpretovat. Jak je uvedeno v [11], logistická funkce byla poprvé odvozena roku 1844 belgickým matematikem Verhulstem, který se zabýval otázkou růstu populace a formuloval tzv. logistický zákon růstu - růst obyvatelstva lze modelovat logistickou křivkou. Navázal tak na předchozí snahy matematiků a filozofů, kteří se zabývali touto problematikou. Verhulst své práce nikdy nepublikoval. O zveřejnění jeho díla se zasloužil statistik Quételet. Verhulstova teorie však upadla na určitý čas v zapomenutí. Jeho práce byly "znovuobjeveny" až za války roku 1916, když na ně upozornil Francouz Mansion americké biology Lowella, Reeda a Pearla. Ti totiž při studiu biologických pochodů a autokatických zákonů objevili též (nezávisle na Verhulstovi) logistický zákon růstu. Od té doby se Verhulstova teorie logistického růstu stala středem zájmu mnoha statistiků. V současné době je široce užívána v biologických vědách např. k modelování růstu rostlin a živočichů.

Položme si otázku: Proč se charakteristické křivky položky liší na základě různých parametrů obtížnosti?

Situci lze vysvětlit slovně: Čím lehčí bude otázka, tím spíše na ni odpoví testovaný s určitou úrovní schopností správně, tj. pravděpodobnost správné odpovědi se bude s klesající obtížností pro dané θ zvyšovat.

Situaci lze vysvětlit matematicky: Použijme logistickou funkci, pomocí níž budeme modelovat charakteristickou křivku položky v jednoparametrickém modelu

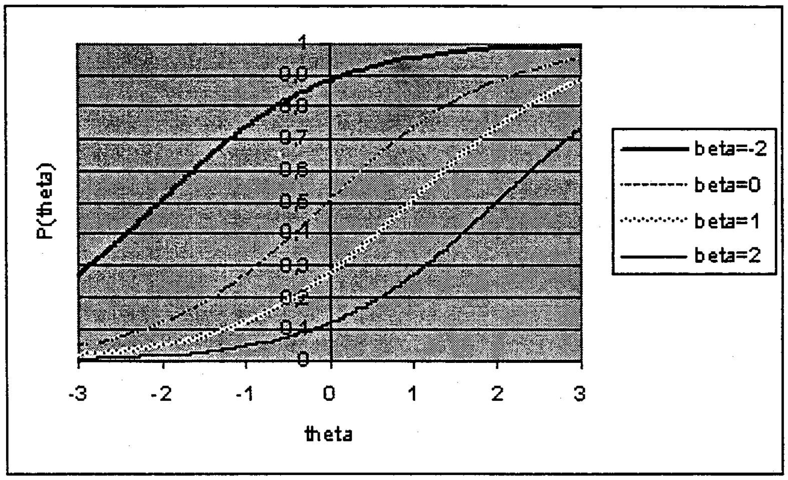
$$f(\theta) = \frac{1}{1 + exp[-(\theta - \beta)]}, \tag{2.58}$$

kde θ je parametr schopností, β je parametr obtížnosti, $(\theta - \beta)$ se nazývá logit. Je-li $\beta_1 < \beta_2$, pak $\frac{1}{1+exp[-(\theta-\beta_1)]} > \frac{1}{1+exp[-(\theta-\beta_2)]}$.

Situaci lze vysvětlit graficky: Na obrázku 2.2, vidíme čtyři charakteristické křivky s parametrem obtížnosti $\beta = -2; 0; 1; 2$. Křivka pro $\beta = -2$ "leží" nad křivkou pro $\beta = 0$ atd. Právě na základě tohoto obrázku je dobře vidět, že definice parametru obtížnosti zní: Parametr obtížnosti je taková hodnota na škále schopností, v níž je pravděpodobnost správné odpovědi rovna 0,5. Matematicky lze toto ověřit snadným dosazením $\theta = \beta$ do 2.58: $\frac{1}{1+exp[-(\theta-\theta)]} = \frac{1}{2}$. Slovně lze situaci vysvětlit tak, že pokud obtížnost otázky je tak velká jako jsou naše schopnosti, tak nemůžeme říci, jestli na ni odpovíme spíše správně nebo spíše chybně - "je to tak půl na půl".

Poznamenejme, že teoreticky může parametr β nabývat hodnot z intervalu $(-\infty; \infty)$, v praxi jsou běžné hodnoty z intervalu $(-3; 3)$.

2. Dvouparametrický logistický model - při popisu charakteristické křivky položky bere



Obrázek 2.2: Jednoparametrický model – charakteristická křivka obtížnosti s parametry obtížnosti $\beta = -2; 0; 1; 2$.

v úvahu její obtížnost a diskriminační schopnost. Druhou jmenovanou vlastnost položky jsme zaváděli již v podkapitole 2.1. Popišme ji ještě podrobněji.

Charakteristické křivky položky se budou lišit pro různé diskriminační parametry. Opět bychom se mohli zeptat proč.

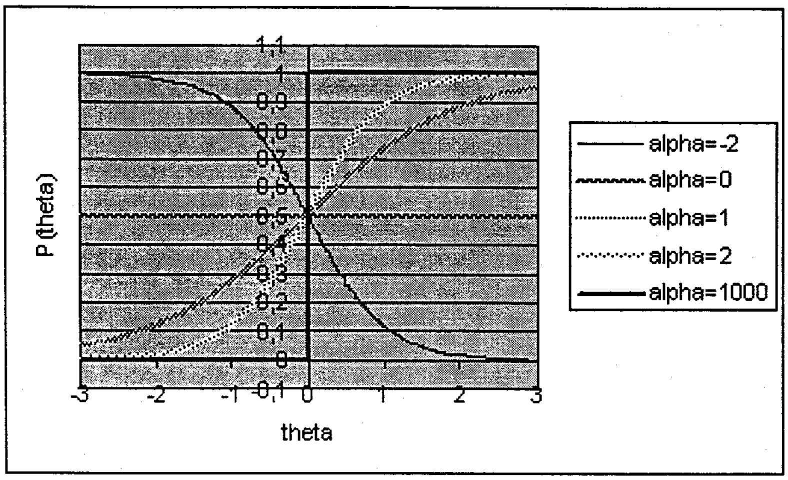
Situaci lze vyvětlit slovně: Intuitivně by mělo platit, že pokud se zvyšuje diskriminační schopnost otázky, tj. jak dobře otázka rozlišuje mezi dobrými a špatnými žáky, tím "zakřivenější" by funkce měla být.

Situaci lze vysvětlit matematicky: Použijme logistickou funkci, pomocí níž budeme modelovat charakteristickou křivku položky v dvouparametrickém modelu:

$$f(\theta) = \frac{1}{1 + exp[-\alpha(\theta - \beta)]}, \tag{2.59}$$

kde θ a β jsou definovány u jednoparametrického modelu a α je diskriminační parametr. Jak jsme si vysvětlili u jednoparametrických modelů, je vždy v bodě $\theta = \beta$ funkční hodnota rovna $1/2$. Derivace křivky se liší pro různá α. Pro jedno dané α nabývá derivace 2.59 maxima v bodě $\beta = \theta$, hodnota této derivace je $\alpha/4$. Definice tohoto parametru α tedy zní, že $\alpha/4$ je rovno derivaci 2.59 v bodě $\theta = \beta$.

Situaci lze vysvětlit graficky. Na obrázku 2.3 vidíme pět charakteristických křivek, s diskriminačním parametrem $\alpha = -2; 0; 1; 2; 1000$. U všech křivek byla zvolena



Obrázek 2.3: Dvouparametrický model – charakteristická křivka obtížnosti s diskriminačním parametrem $\alpha = -2; 0; 1; 2; 1000$ (parametr $\beta = 0$.)

obtížnost rovna nule. Právě v této hodnotě se všechny křivky protínají. $\alpha = -2$ je příkladem otázky se záporným diskriminačním parametrem, křivka je klesající, tj. s větší pravděpodobností řeší otázky správně méně schopnější testovaní. Je zřejmé, že takové otázky jsou v testu nežádoucí. $\alpha = 0$ je příkladem otázky s nulovým dikriminačním parametrem, tj. s nulovou schopností rozlišovat mezi dobrými a špatnými žáky. Z tohoto důvodu je funkce konstantní. Případy $\alpha = 1; 2$ je příkladem otázky s v praxi běžným kladným diskriminačním parametrem, křivka je rostoucí, tj. s větší pravděpodobností řeší otázky správně více schopnější testovaní. $\alpha = 1000$ je příkladem otázky, která téměř dokonale rozlišuje mezi testovanými, zda jsou jejich schopnosti větší či menší než nula, ale nerozlišuje téměř nijak mezi testovanými se schopnostmi pod nulou a zároveň nijak nerozlišuje mezi testovanými se schopnostmi nad nulou. Vidíme, že zakřivení roste, pokud se $|\alpha|$ zvyšuje.

Poznamenejme, že teoreticky může α nabývat hodnot z intervalu $(-\infty; \infty)$, v praxi jsou běžné hodnoty z intervalu $(-3; 3)$.

3. Tříparametrický logistický model - při popisu charakteristické křivky položky bere v úvahu její obtížnost, diskriminační schopnost a tzv. parametr uhodnutí správné

odpovědi.

Charakteristické křivky položky se budou lišit pro různé parametry uhodnutí správné odpovědi. I zde bychom se mohli zeptat proč tomu tak je.

Situaci lze vysvětlit slovně: Parametr uhodnutí správné odpovědi λ je definován jako pravděpodobnost získání správné odpovědi pouhým hádáním. λ není funkcí parametru schopností, tj. zkoušení s malou i velkou hodnotou parametru schopností mají stejnou pravděpodobnost získání správné odpovědi pouhým hádáním. Zároveň je jasné, že nejnižší možná pravděpodobnost správné odpovědi bude λ, tj. charakteristické křivky budou zdola omezeny tímto parametrem.

Situaci lze přesněji vysvětlit matematicky: Použijme funkci, pomocí níž budeme modelovat charakteristickou křivku položky v tříparametrickém modelu

$$f(\theta) = \lambda + (1-\lambda)\frac{1}{1 + exp[-\alpha(\theta - \beta)]}, \tag{2.60}$$

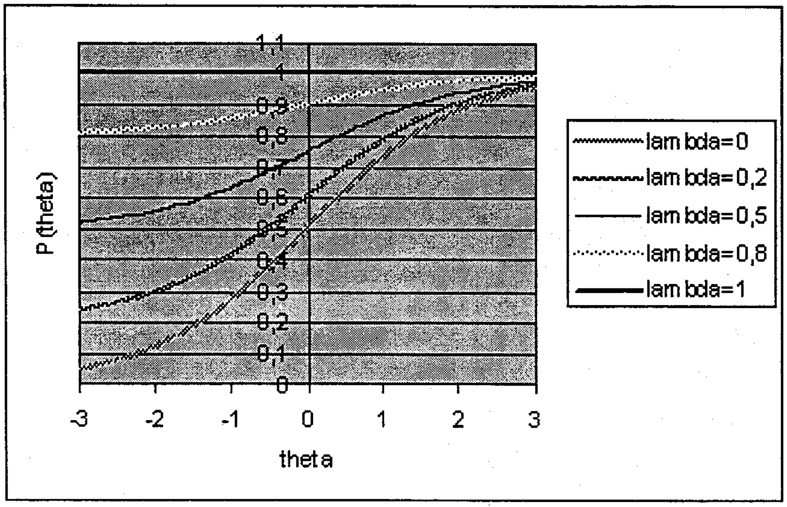
kde parametry α, β a θ jsou postupně diskriminační parametr, parametr obtížnosti a parametr schopnosti; λ je parametr získání správné odpovědi pouhým hádáním. Zkoušený tedy buď zná správnou odpověď s pravděpodobností 2.59 nebo hádá s pravděpodobností úspěchu λ. Je-li $\alpha > 0$, $f(\theta)$ je rostoucí funkce a pro $\theta \to -\infty$ platí, že $f(\theta) \to \lambda$. Je-li $\alpha < 0$, $f(\theta)$ je klesající funkce a pro $\theta \to \infty$ platí, že $f(\theta) \to \lambda$. Je-li $\alpha = 0$ je $f(\theta) = (1+\lambda)/2$, tj. $f(\theta)$ je konstantní funkce. Pokud α není rovno nule, je dolní hranicí $f(\theta)$ právě parametr λ, čímž opravujeme předchozí slovní vysvětlení.

U předchozích jedno- a dvouparametrických modelů byl parametr obtížnosti β definován jako taková hodnota na škále schopností, v níž je pravděpodobnost správné odpovědi rovna $0,5$. Vzhledem k tomu, že dolní hranice charakteristické křivky položky není 0, ale λ (opět $\alpha \neq 0$), změní se i interpretace parametru obtížnosti. Dosaďme do 2.60 $\theta = \beta$, dostaneme $f(\theta) = \lambda + (1-\lambda)\frac{1}{2} = \frac{1+\lambda}{2}$. Tedy parametr obtížnosti je definován jako hodnota na škále schopností, v níž je pravděpodobnost správné odpovědi rovna $\frac{1+\lambda}{2}$, což je v polovině mezi hodnotami λ a 1.

U dvouparametrického modelu je diskriminační parametr α úměrný zakřivení charakteristické křivky v bodě $\theta = \beta$. Toto platí i u tříparametrických modelů, změna je pouze v hodnotě derivace v bodě $\theta = \beta$, je rovna $\frac{\alpha(1-\lambda)}{4}$.

Přestože 2.60 nedefinuje logistickou funkci, je tento model v literatuře (např. [5]) nazýván jako tříparametrický logistický model.

Situaci lze vyvětlit graficky: Na obrázku 2.4 vidíme pět charakteristických křivek s parametrem $\lambda = 0; 0,2; 0,5; 0,8; 1$. U všech křivek byla zvolena obtížnost rovna



Obrázek 2.4: Tříparametrický model – charakteristická křivka obtížnosti s parametrem uhodnutí $\lambda = 0; 0,2; 0,5; 0,8; 1$ (parametry $\beta = 0$ a $\alpha = 1$)

nule a diskriminační parametr roven jedné. U všech křivek je dolní mez rovna právě hodnotě λ. Otázka s parametrem $\lambda = 1$ je extrémní, všichni by na ni odpověděli správně.

Poznamenejme, že teoreticky může λ nabývat hodnot z intervalu $[0; 1]$, v praxi jsou běžné hodnoty z intervalu $[0; 0,35]$.

Důležité je, z pozorovaných dat odhadnout parametry otázek a hlavně parametry schopností jednotlivých lidí. Popisu jednotlivých metod odhadů, jsou věnovány kapitoly 3 a 4. Pro nematematika je postačující vědět, že existují procedury na odhadování těchto parametrů a že vyžadují práci s počítačem (musíme totiž řešit soustavy nelineárních rovnic).

Víme, co je charakteristická křivka položky, a že její parametry lze odhadovat.

Na závěr této podkapitoly bychom mohli na příkladu naznačit vysvětlení nezávislosti odhadů parametru schopnosti a parametru obtížnosti otázky jako jedné z jejích vlastností. Představme si situaci, kdy parametry schopnosti testovaných v jedné skupině jsou z intervalu $(-3; -1)$ a v druhé skupině z $(1; 3)$. Oběma skupinám je předložena stejná otázka. Po odhadu obtížnosti otázky bychom měli dojít k závěru, že je v obou skupinách stejný. Když si představíme možný tvar charakteristické křivky, viz obrázek 2.1, tak v první skupině by měl odhad charakteristické křivky odpovídat její dolní části (pro $\theta \in (-3; -1)$) a v druhé skupině by měl odhad charakteristické křivky odpovídat její horní části (pro $\theta \in (1; 3)$) charakteristické křivky. Pozorované podíly správných odpovědí bychom v první skupině nalezli kolem dolní části a v druhé skupině kolem horní části charakteristické křivky. Pořád však mluvíme o jedné charakteristické křivce s jedním parametrem obtížnosti.

Kapitola 3

Odhad parametru obtížnosti položky

Cílem této kapitoly bude prezentovat nejčastější metody odhadu parametru obtížnosti položky. Tyto metody jsou užívány nejen k odhadům dalších parametrů položky (případy dvou- a tříparametrických modelů), ale i k odhadům parametrů otázek v jiných modelech, které užívá teorie odpovědí na položky. Abychom si situaci algebraicky zjednodušili, budeme prezentovat jednotlivé metody odhadu pro případ jednoparametrického modelu, tj. zaměříme se na odhad parametru obtížnosti.

K odhadům parametru obtížnosti užijeme metodu maximální věrohodnosti, přičemž rozlišíme:

- metodu společné maximální věrohodnosti - parametry schopnosti jsou odhadovány společně s parametrem obtížnosti, podrobněji viz podkapitola 3.1;
- metodu podmíněné maximální věrohodnosti - parametry schopnosti jsou eliminovány podmíněním, podrobněji viz podkapitola 3.2;
- metodu marginální maximální věrohodnosti - parametry schopnosti jsou "vyintegrovány", podrobněji viz podkapitola 3.3.

3.1 Metoda společné maximální věrohodnosti

Abychom získali jednoznačné řešení, musíme učinit následující omezení:

$$\sum_i \beta_i = 0. \tag{3.1}$$

Parametr b v 2.33 je nyní roven nule, tj. počátky β-škály a θ-škály jsou totožné. Abychom upevnili počátek β-škály zvolili jsme omezení 3.1.

V kapitole 2 jsme vyjádřili za předpokladů nezávislosti odpovědí i-té a j-té osoby a lokální nezávislosti pravděpodobnost matice $X_{n\times k}$ jako

$$P(\mathbf{X_{n\times k}} = \mathbf{x_{n\times k}}|\theta,\beta) = \prod_{\nu=1}^{n}\prod_{i=1}^{k}\frac{exp[x_{\nu i}(\theta_\nu - \beta_i)]}{1 + exp[(\theta_\nu - \beta_i)]}.$$

Cílem bude odvodit odhady obou vektorů parametrů $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ a $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$ pomocí věrohodnostní funkce $L(\theta, \beta|\mathbf{x_{n\times k}})$, která je rovna $P(\mathbf{X_{n\times k}} = \mathbf{x_{n\times k}}|\theta,\beta)$. Pro získání odhadů položíme derivaci (dle neznámých parametrů) logaritmu věrohodnostní funkce rovnu nule:

$$\begin{aligned}\ln L(\theta, \beta|\mathbf{x_{n\times k}}) &= \sum_{\nu=1}^{n}\sum_{i=1}^{k} x_{\nu i}(\theta_\nu - \beta_i) - \sum_{\nu=1}^{n}\sum_{i=1}^{k}\ln\left(1 + \exp\left(\theta_\nu - \beta_i\right)\right) = \\ &= \sum_{\nu=1}^{n} x_{\nu\cdot}\theta_\nu - \sum_{i=1}^{k} x_{\cdot i}\beta_i - \sum_{\nu=1}^{n}\sum_{i=1}^{k}\ln\left(1 + \exp\left(\theta_\nu - \beta_i\right)\right), \qquad (3.2)\end{aligned}$$

kde $x_{\nu\cdot} = \sum_{i=1}^{k} x_{\nu i}$, což je suficientní statistika pro θ_ν a $x_{\cdot i} = \sum_{\nu=1}^{n} x_{\nu i}$, což je suficientní statistika pro β_i. Derivujme 3.2 dle θ_ν a β_i:

$$\begin{aligned}\frac{\partial \ln L(\theta,\beta)}{\partial\theta_\nu} &= x_{\nu\cdot} - \sum_{i=1}^{k}\frac{\exp\left(\theta_\nu - \beta_i\right)}{1 + \exp\left(\theta_\nu - \beta_i\right)} \qquad \nu = 1, 2, \ldots, n \\ \frac{\partial \ln L(\theta,\beta)}{\partial\beta_i} &= -x_{\cdot i} + \sum_{\nu=1}^{n}\frac{\exp\left(\theta_\nu - \beta_i\right)}{1 + \exp\left(\theta_\nu - \beta_i\right)} \qquad i = 1, 2, \ldots, k. \qquad (3.3)\end{aligned}$$

Položíme-li derivace rovny nule získáme následující rovnice:

$$\begin{aligned}x_{\nu\cdot} &= \sum_{i=1}^{k}\frac{\exp\left(\theta_\nu - \beta_i\right)}{1 + \exp\left(\theta_\nu - \beta_i\right)} \qquad \nu = 1, 2, \ldots, n \qquad (3.4) \\ x_{\cdot i} &= \sum_{\nu=1}^{n}\frac{\exp\left(\theta_\nu - \beta_i\right)}{1 + \exp\left(\theta_\nu - \beta_i\right)} \qquad \imath = 1, 2, \ldots, k. \qquad (3.5)\end{aligned}$$

Dostáváme speciální případ obecného výsledku pro exponenciální rodiny (viz [7]): Věrohodnostní rovnice pro odhady parametrů $\theta_1, \theta_2, \ldots, \theta_n$, a $\beta_1, \beta_2, \ldots, \beta_k$ získáme tak, že pozorované hodnoty suficientních statistik pro tyto parametry (vlevo v rovnicích 3.4 a 3.5) položíme rovné jejich očekávaným hodnotám (vpravo v rovnicích 3.4 a 3.5), tj.

$$x_{\nu\cdot} = E[X_{\nu\cdot}] \quad \nu = 1, 2, \ldots, n \tag{3.6}$$

$$x_{\cdot i} = E[X_{\cdot i}] \quad \imath = 1, 2, \ldots, k. \tag{3.7}$$

Celkově můžeme z 3.6 a 3.7 dostat nejvýše $2k-2$ nezávislých rovnic. $x_{\nu\cdot}$ může totiž nabývat hodnot $0, 1, \ldots, k$, ale případy $0, k$ z analýzy vyloučíme, což bude zdůvodněno v následujícím odstavci. Dále máme k otázek, ale jedna rovnice je nadbytečná vzhledm k omezení 3.1. Tedy dostáváme skutečně nejvýše $(k-1)+(k-1)$ nezávislých rovnic.

Speciální případ je $x_{\cdot i} = 0$, formální řešení 3.7 diverguje k $\beta_i = \infty$. Mohli bychom předpokládat, že všichni budoucí testovaní tuto položku nezodpoví správně, je obtížnější než ostatní otázky, ale daný výběr zkoušených neposkytuje dostatečnou informaci k odhadnutí jejího umístnění na β-škále. Podobně je tomu s případem $x_{\cdot i} = n$. Formální řešení diverguje k $\beta_i = -\infty$, z čehož bychom mohli odvodit, že je otázka méně obtížnější než ostatní, ale opět nemáme dostatečnou informaci pro její umístnění na β-škále. Podobné situace nastávají pro případy $x_{\nu\cdot} = 0$ a $x_{\nu\cdot} = k$, formální řešení 3.6 konverguje postupně k $\theta_\nu = -\infty$ a k $\theta_\nu = \infty$. Toto naznačuje, že zkoušení s $\theta_\nu = -\infty$ v budoucnu nikdy neodpoví správně na otázky testující schopnost θ a zkoušení s $\theta_\nu = \infty$ naopak vždy odpoví na tento typ otázek správně. Tyto čtyři případy (nulové a maximální možné součty v řádcích a sloupcích matice dat) odstraňme z datové matice. Pak se může stát, že opět bude nějaký součet v řádku či sloupci datové matice nulový či maximální, tento řádek či sloupec opět vypustíme. Toto odstraňování řádků a sloupců končí většinou naštěstí po druhém či třetím kole. Zbytek dat vstupuje do iteračního procesu.

Bylo zjištěno (Andersen [2]), že odhady metodou společné maximální věrohodnosti jsou nekonzistentní pro k pevné a $n \to \infty$, což je jejich značná nevýhoda. Většinou máme velký výběr testovaných a pevně daný počet otázek. Z toho důvodu se dává přednost metodám uvedených v pokapitolách 3.2 a 3.3, které "odstraňují" rušivé parametry.

3.2 Metoda podmíněné maximální věrohodnosti

Jako druhou metodu na odhadování parametrů obtížnosti si představíme metodu podmíněné maximální věrohodnosti. Jak je uvedeno v [7], na základě obecné vlastnosti exponenciálních rodin víme, že rozdělení podmíněné suficientní statistikou pro rušivý parametr (v našem případě je rušivý parametr θ_ν) nezávisí na tomto parametru. Maximalizace věrohodnostní funkce podmíněné řádkovými skóry $x_{\nu.}$ (suficientní statistika pro θ_ν) vede ke zlepšení odhadů parametrů obtížnosti (pro k pevné a $n \to \infty$ jsou odhady β_i konzistentní).

V této podkapitole si označme $\xi_\nu = \exp\theta_\nu$ a $\epsilon_i = \exp(-\beta_i)$, pak pravděpodobnost prvku v matici $\mathbf{X_{n\times k}}$ za podmínky ξ_ν a ϵ_i má tvar

$$P(X_{\nu i} = x_{\nu i}|\xi_\nu, \epsilon_i) = \frac{(\xi_\nu)^{x_{\nu i}}(\epsilon_i)^{x_{\nu i}}}{(1+\xi_\nu\epsilon_i)}. \tag{3.8}$$

Odvoďme pravděpodobnost matice $\mathbf{X_{n\times k}}$ za podmínky znalosti řádkových skórů $\mathbf{r} = (r_1, r_2, \ldots, r_n)$, kde $r_\nu = x_{\nu.}$. Náhodné veličiny $R_\nu = X_{\nu.}$ zaznamenejme do vektoru $\mathbf{R} = (R_1, R_2, \ldots, R_n)$. Libovolný vektor odpovědí $\mathbf{y} = (y_1, y_2, \ldots, y_k)$, pro který platí $\sum_{i=1}^{k} y_i = r_\nu$, označme $(\mathbf{y}|r_\nu)$:

$$\begin{aligned}
P(\mathbf{X_{n\times k}} = \mathbf{x_{n\times k}}|\mathbf{R}=\mathbf{r}) &= \frac{P(\mathbf{X_{n\times k}} = \mathbf{x_{n\times k}})}{P(\mathbf{R}=\mathbf{r})} = \\
&= \frac{\prod_{\nu=1}^{n}\prod_{i=1}^{k}\frac{(\xi_\nu)^{x_{\nu i}}(\epsilon_i)^{x_{\nu i}}}{(1+\xi_\nu\epsilon_i)}}{\prod_{\nu=1}^{n}\sum_{(\mathbf{y}|r_\nu)}\prod_{i=1}^{k}\frac{(\xi_\nu)^{y_{\nu i}}(\epsilon_i)^{y_{\nu i}}}{(1+\xi_\nu\epsilon_i)}} = \\
&= \prod_{\nu=1}^{n}\frac{(\prod_{i=1}^{k}\frac{1}{1+\xi_\nu\epsilon_i})(\xi_\nu)^{r_\nu}\prod_{i=1}^{k}(\epsilon_i)^{x_{\nu i}}}{(\prod_{i=1}^{k}\frac{1}{1+\xi_\nu\epsilon_i})(\xi_\nu)^{r_\nu}\sum_{(\mathbf{y}|r_\nu)}\prod_{i=1}^{k}(\epsilon_i)^{y_{\nu i}}} = \\
&= \prod_{\nu=1}^{n}\frac{\prod_{i=1}^{k}(\epsilon_i)^{x_{\nu i}}}{\sum_{(\mathbf{y}|r_\nu)}\prod_{i=1}^{k}(\epsilon_i)^{y_{\nu i}}} = \\
&= \frac{\prod_{i=1}^{k}(\epsilon_i)^{x_{.i}}}{\prod_{\nu=1}^{n}\gamma_{r_\nu}},
\end{aligned} \tag{3.9}$$

kde

$$\gamma_{r_\nu}(\epsilon_1, \epsilon_2, \ldots, \epsilon_k) = \sum_{(\mathbf{y}|r_\nu)}\prod_{i=1}^{k}(\epsilon_i)^{y_i} \tag{3.10}$$

označuje symetrickou funkci, o které jsme se zmiňovali v důkazu lemmatu 1 (místo funkcí $h_i(\xi)$ v 2.16 máme parametry ϵ_i). γ_{r_ν} má pro $r_\nu = 0, 1, \ldots, k$ tvar

$$\gamma_0 = 1,$$

$$\begin{aligned}\gamma_1 &= \epsilon_1 + \epsilon_2 + \ldots + \epsilon_k,\\ \gamma_2 &= \epsilon_1\epsilon_2 + \epsilon_1\epsilon_3 + \ldots + \epsilon_{k-1}\epsilon_k,\\ &\vdots\\ \gamma_k &= \epsilon_1\epsilon_2\ldots\epsilon_k.\end{aligned} \tag{3.11}$$

Vidíme, že $P(\mathbf{X_{n\times k}} = \mathbf{x_{n\times k}}|\mathbf{R} = \mathbf{r})$ skutečně není závislá na parametrech schopnosti, podmíněné rozdělení je závislé jen na parametrech obtížnosti. Označme příslušnou věrohodnostní funkci jako $L(\epsilon|\mathbf{r})$, kde $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_k)$. Rovnice pro odhady těchto parametrů nalezneme derivováním dle ϵ_j logaritmu věrohodnostní funkce, kterou položíme rovnu nule:

$$\begin{aligned}\ln L(\epsilon|\mathbf{r}) &= \ln \frac{\prod_{i=1}^{k}(\epsilon_i)^{x_{\cdot i}}}{\prod_{\nu=1}^{n}\gamma_{r_\nu}} = \sum_{i=1}^{k} x_{\cdot i}\ln\epsilon_i - \sum_{\nu=1}^{n}\ln\gamma_{r_\nu},\\ \frac{\partial \ln L(\epsilon|\mathbf{r})}{\partial\epsilon_j} &= \frac{x_{\cdot j}}{\epsilon_j} - \sum_{\nu=1}^{n}\frac{1}{\gamma_{r_\nu}}\frac{\partial\gamma_{r_\nu}}{\partial\epsilon_j} \qquad j = 1, 2, \ldots, k,\end{aligned} \tag{3.12}$$

kde

$$\begin{aligned}\frac{\partial\gamma_{r_\nu}}{\partial\epsilon_j} &= \frac{\partial \sum_{(\mathbf{y}|r_\nu)}\prod_{i=1}^{k}(\epsilon_i)^{y_{\nu i}}}{\partial\epsilon_j} =\\ &= \sum_{(\mathbf{y}|r_\nu)} y_{\nu j}\prod_{i\neq j=1}^{k}\epsilon_i^{y_{\nu i}} =\\ &= \sum_{(\mathbf{y}|r_\nu-1)}\sum_{i\neq j=1}^{k}\epsilon_i^{y_{\nu i}} =\\ &= \gamma^{(j)}_{(r_\nu-1)} \qquad j = 1, 2, \ldots, k,\end{aligned} \tag{3.13}$$

kde $\gamma^{(j)}_{(r_\nu-1)}$ je symetrická funkce proměnných ϵ_i, přičemž $i = 1, 2, \ldots, k$ a $i \neq j$. Dosadíme 3.13 do 3.12 a pak 3.12 položíme rovno nule, dostaneme

$$\begin{aligned}\frac{x_{\cdot j}}{\epsilon_j} - \sum_{\nu=1}^{n}\frac{\gamma^{(j)}_{r_\nu-1}}{\gamma_{r_\nu}} &= 0\\ x_{\cdot j} &= \sum_{\nu=1}^{n}\frac{\epsilon_j\gamma^{(j)}_{r_\nu-1}}{\gamma_{r_\nu}} \qquad j = 1, 2, \ldots, k.\end{aligned} \tag{3.14}$$

Rozdělení matice $\mathbf{X_{n\times k}}$ za podmínky znalosti pozorovaných hodnot minimálních suficientních statistik $r_1, r_2, \ldots, r_n$ pro $\theta_1, \theta_2, \ldots, \theta_n$ patří do exponenciální rodiny. $x_{\cdot j}$ je pozorovaná hodnota suficientní statistiky pro β_j. Dostáváme opět speciální případ pro exponenciální rodiny: věrohodnostní rovnice pro odhad β_j získáme tak, že položíme pozorovanou hodnotu suficientní statistiky pro β_j rovnu její očekávané hodnotě za podmínky znalosti pozorovaných hodnot suficientních statistik pro $\theta_1, \theta_2, \ldots, \theta_n$, tj.

$$x_{\cdot j} = E[X_{\cdot j}|r_1, r_2, \ldots, r_k] \qquad j = 1, 2, \ldots, k. \tag{3.15}$$

Vzhledem k tomu, že platí

$$\begin{aligned} E[\mathbf{X}_{\cdot\mathbf{j}}|R_1 = r_1, R_2 = r_2, \ldots, R_n = r_n] &= E[\sum_{\nu=1}^{n} X_{\nu j}|R_1 = r_1, R_2 = r_2, \ldots, R_n = r_n] = \\ &= \sum_{\nu=1}^{n} E[X_{\nu j}|R_\nu = r_\nu] = \\ &= \sum_{\nu=1}^{n} 1P(X_{\nu j} = 1|R_\nu = r_\nu) = \\ &= \sum_{\nu=1}^{n} \frac{P(X_{\nu j} = 1, R_\nu = r_\nu)}{P(R_\nu = r_\nu)} = \\ &= \sum_{\nu=1}^{n} \frac{\epsilon_j \gamma^{(j)}_{r_\nu - 1}}{\gamma_{r_\nu}} \qquad j = 1, 2, \ldots, k, \end{aligned} \tag{3.16}$$

v 3.14 jsme dospěli skutečně k závěru v 3.15.

Stejně jako v metodě společné maximální věrohodnosti vyloučíme z analýzy nulové a maximální součty v řádcích a sloupcích datové matice $\mathbf{x_{n \times k}}$. Vzhledm k tomu, že držíme omezení 3.1, vyloučíme jednu rovnici v 3.15. Do iteračního procesu vstupuje tedy $k-1$ rovnic z 3.15 a omezující podmínka 3.1. Po jejich vyřešení dostáváme odhady parametrů $\beta_1, \beta_2, \ldots, \beta_k$.

3.3 Metoda marginální maximální věrohodnosti

Jestliže chceme poukázat na rozdělení schopností buď v celé skupině testovaných nebo v její velké podskupině, je výhodnější použít metody marginální maximální věrohodnosti oproti metodám společné a podmíněné maximální věrohodnosti. Na druhou stranu, pokud je rozdělení nesprávně odhadnuto nebo nesprávně předem určeno, je metoda marginální maximální věrohodnosti nejméně výhodná.

Nechť rozdělení schopností v populaci má distribuční funkci $G(\theta)$. Předpokládejme, že parametry schopnosti pozorovaných osob (testovaných) jsou náhodným výběrem z této distribuční funkce. Existují různé přístupy k volbě $G(\theta)$. Prvním přístupem je předpokádat, že $G(\theta)$ je distribuční funkce dané parametrické rodiny rozdělení s omezeným počtem neznámých parametrů, které musí být odhadovány společně s $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$. Často se předpokládá, že $G(\theta)$ je distribuční funkce normálního rozdělení s neznámou střední hodnotou a neznámým rozptylem. Druhý přístup je odhadnout $G(\theta)$ z pozorovaných dat.

V následujícím odvození rovnic, jejichž numerickým vyřešením dostaneme odhad β, budeme předpokládat, že $G(\theta)$ je distribuční funkce normálního rozdělení se střední hodnotou μ a rozptylem σ^2.

Pravděpodobnost odpovědi $x_{\nu i}$ zkoušeného S_ν na položku I_i je dána

$$P(X_{\nu i} = x_{\nu i}) = \int_{-\infty}^{\infty} \frac{\exp\left[x_{\nu i}(\theta - \beta_i)\right]}{(1 + \exp(\theta - \beta_i))} \varphi(\theta) d\theta, \tag{3.17}$$

kde $\varphi(\theta)$ je hustota normálního rozdělení $N(\mu, \sigma^2)$, tj.

$$\varphi(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\theta - \mu)^2}{2\sigma^2}\right). \tag{3.18}$$

Vzhledem k lokální nezávislosti, tj. nezávislosti odpovědí na úrovni θ_ν dostáváme, že pravděpodobnost vektoru odpovědí zkoušeného S_ν je vzhledem k 3.17

$$P(X_{\nu 1} = x_{\nu 1}, X_{\nu 2} = x_{\nu 2}, \ldots, X_{\nu k} = x_{\nu k}) = \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp\left[x_{\nu i}(\theta - \beta_i)\right]}{(1 + \exp(\theta - \beta_i))} \varphi(\theta) d\theta \tag{3.19}$$

Označme

$$\begin{aligned}
\tau_{\mathbf{z}} &= P(Z_1 = z_1, Z_2 = z_2, \ldots, Z_k = z_k) = \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp\left[z_i(\theta - \beta_i)\right]}{1 + \exp(\theta - \beta_i)} \varphi(\theta) d\theta = \\
&= \exp\left(-\sum_{i=1}^{k} z_i \beta_i\right) C_r(\beta),
\end{aligned} \tag{3.20}$$

kde $\mathbf{z} = (z_1, z_2, \ldots, z_k)$ označuje libovolný vektor odpovědí, těchto vektorů je tedy 2^k; $r = \sum_{i=1}^{k} z_i$ a kde

$$C_r(\beta) = \int_{-\infty}^{\infty} \frac{\exp(r\theta)}{\prod_{i=1}^{k} (1 + \exp(\theta - \beta_i))} \varphi(\theta) d\theta. \tag{3.21}$$

Dále nechť $n_{\mathbf{z}}$ označuje počet zkoušených, kteří mají vektor odpovědí $\mathbf{z}$ a n_r označuje počet zkoušených se skórem r.

Vzhledem k tomu, že budeme odhadovat i parametry normálního rozdělení společně s vektorem $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$ označme si hledané parametry vektorem α, tj.

$$\alpha = (\beta_1, \beta_2, \ldots, \beta_k, \mu, \sigma). \tag{3.22}$$

Věrohodnostní funkce L bude vzhledem 3.20 k rovna

$$L = \prod_{\mathbf{z}} {\tau_{\mathbf{z}}}^{n_{\mathbf{z}}}. \tag{3.23}$$

Logaritmus věrohodnostní funkce L zderivujme dle jednotlivých proměnných v α. Pak položme všechny derivace nule, čímž získáme rovnice pro odhadování parametrů v α.

$$\ln L = \ln \prod_{\mathbf{z}} {\tau_{\mathbf{z}}}^{n_{\mathbf{z}}} = \sum_{\mathbf{z}} \ln {\tau_{\mathbf{z}}}^{n_{\mathbf{z}}} = \sum_{\mathbf{z}} n_{\mathbf{z}} \ln \tau_{\mathbf{z}}, \tag{3.24}$$

$$\frac{\partial \ln L}{\partial \alpha_j} = \frac{\partial \sum_{\mathbf{z}} n_{\mathbf{z}} \ln \tau_{\mathbf{z}}}{\partial \alpha_j} = \sum_{\mathbf{z}} \frac{n_{\mathbf{z}}}{\tau_{\mathbf{z}}} \frac{\partial \tau_{\mathbf{z}}}{\partial \alpha_j} = 0 \qquad j = 1, 2, \ldots, k, k+1, k+2, \tag{3.25}$$

kde pro $j = 1, 2, \ldots, k$

$$\begin{aligned}
\frac{\partial \tau_{(\mathbf{z})}}{\partial \alpha_j} &= \frac{\partial \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp[z_i(\theta-\beta_i)]}{1+\exp(\theta-\beta_i)} \varphi(\theta) d\theta}{\partial \alpha_j} = \\
&= \int_{-\infty}^{\infty} \Big(\prod_{i=1}^{k} \frac{\exp[z_i(\theta-\beta_i)]}{1+\exp(\theta-\beta_i)}(-z_j) + \prod_{i=1}^{k} \frac{\exp[z_i(\theta-\beta_i)]}{1+\exp(\theta-\beta_i)} \frac{\exp(\theta-\beta_j)}{(1+\exp(\theta-\beta_j))}\Big)\varphi(\theta)d\theta = \\
&= (-z_j) \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp[z_i(\theta-\beta_i)]}{1+\exp(\theta-\beta_i)} \varphi(\theta)d\theta + \\
&\quad + \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp[z_i(\theta-\beta_i)]}{1+\exp(\theta-\beta_i)} \frac{\exp(\theta-\beta_j)}{(1+\exp\theta-\beta_j)} \varphi(\theta)d\theta = \\
&= (-z_j) \exp(-\sum_{i=1}^{k} z_i\beta_i) C_r(\beta) + \\
&\quad + \exp(-\sum_{i=1}^{k} z_i\beta_i) \int_{-\infty}^{\infty} \frac{\exp(r\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \frac{\exp(\theta-\beta_j)}{(1+\exp(\theta-\beta_j))} \varphi(\theta)d\theta,
\end{aligned} \tag{3.26}$$

pro $j = k+1$

$$\begin{aligned}
\frac{\partial \tau_{\mathbf{z}}}{\partial \alpha_{k+1}} &= \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp\left[z_i(\theta-\beta_i)\right]}{1+\exp\left(\theta-\beta_i\right)} \left(\frac{\mu-\theta}{\sigma^2}\right) \varphi(\theta) d\theta = \\
&= \frac{\mu \exp\left(-\sum_{i=1}^{k} z_i\beta_i\right)}{\sigma^2} C_r(\beta) - \\
&\quad - \frac{\exp\left(-\sum_{i=1}^{k} z_i\beta_i\right)}{\sigma^2} \int_{-\infty}^{\infty} \theta \frac{\exp\left(r\theta\right)}{\prod_{i=1}^{k}\left(1+\exp\left(\theta-\beta_i\right)\right)} \varphi(\theta) d\theta
\end{aligned} \tag{3.27}$$

a pro $j = k+2$

$$\begin{aligned}
\frac{\partial \tau_{\mathbf{z}}}{\partial \alpha_{k+1}} &= \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp\left[z_i(\theta-\beta_i)\right]}{1+\exp\left(\theta-\beta_i\right)} \left(-\frac{1}{\sigma} + \frac{(\theta-\mu)^2}{\sigma^3}\right) \varphi(\theta) d\theta = \\
&= \int_{-\infty}^{\infty} \prod_{i=1}^{k} \frac{\exp\left[z_i(\theta-\beta_i)\right]}{1+\exp\left(\theta-\beta_i\right)} \left(\frac{\theta^2}{\sigma^3} - \frac{2\theta\mu}{\sigma^3} + \frac{(\mu^2-\sigma^2)}{\sigma^3}\right) \varphi(\theta) d\theta = \\
&= \frac{(\mu^2-\sigma^2) \exp\left(-\sum_{i=1}^{k} z_i\beta_i\right)}{\sigma^3} C_r(\beta) - \\
&\quad - \frac{2\mu \exp\left(-\sum_{i=1}^{k} z_i\beta_i\right)}{\sigma^3} \int_{-\infty}^{\infty} \theta \frac{\exp\left(r\theta\right)}{\prod_{i=1}^{k}\left(1+\exp\left(\theta-\beta_i\right)\right)} \varphi(\theta) d\theta + \\
&\quad + \frac{\exp\left(-\sum_{i=1}^{k} z_i\beta_i\right)}{\sigma^3} \int_{-\infty}^{\infty} \theta^2 \frac{\exp\left(r\theta\right)}{\prod_{i=1}^{k}\left(1+\exp\left(\theta-\beta_i\right)\right)} \varphi(\theta) d\theta.
\end{aligned} \tag{3.28}$$

Dosadíme-li do 3.25 postupně 3.26, 3.27 a 3.28 dostaneme pro $j = 1, 2, \ldots, k$

$$\begin{aligned}
\sum_{\mathbf{z}} n_{\mathbf{z}} z_j &= \sum_{\mathbf{z}} \frac{n_{\mathbf{z}}}{C_r(\beta)} \int_{-\infty}^{\infty} \frac{\exp\left(r\theta\right)}{\prod_{i=1}^{k}\left(1+\exp\left(\theta-\beta_i\right)\right)} \frac{\exp\left(\theta-\beta_j\right)}{\left(1+\exp\left(\theta-\beta_j\right)\right)} \varphi(\theta) d\theta \\
\sum_{(\mathbf{z}:z_j=1)} n_{\mathbf{z}} &= \sum_{r=0}^{k} \frac{n_r}{C_r(\beta)} \int_{-\infty}^{\infty} \frac{\exp\left(r\theta\right)}{\prod_{i=1}^{k}\left(1+\exp\left(\theta-\beta_i\right)\right)} \frac{\exp\left(\theta-\beta_j\right)}{\left(1+\exp\left(\theta-\beta_j\right)\right)} \varphi(\theta) d\theta,
\end{aligned} \tag{3.29}$$

pro $j = k + 1$

$$\begin{aligned}
\sum_{\mathbf{z}} n_{\mathbf{z}}\mu &= \sum_{r=0}^{k} \frac{n_r}{C_r(\beta)} \int_{-\infty}^{\infty} \theta \frac{\exp(r\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \varphi(\theta) d\theta \\
n\mu &= \sum_{r=0}^{k} \frac{n_r}{C_r(\beta)} \int_{-\infty}^{\infty} \theta \frac{\exp(r\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \varphi(\theta) d\theta, \qquad (3.30)
\end{aligned}$$

použijeme-li 3.30, pak pro $j = k + 2$ dostaneme

$$\begin{aligned}
2\mu \overbrace{\sum_{r=0}^{k} \frac{n_r}{C_r(\beta)} \int_{-\infty}^{\infty} \theta \frac{\exp(r\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \varphi(\theta) d\theta}^{n\mu} - n(\mu^2-\sigma^2) &= \\
= \sum_{r=0}^{k} \frac{n_r}{C_r(\beta)} \int_{-\infty}^{\infty} \theta^2 \frac{\exp(r\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \varphi(\theta) d\theta & \\
n(\mu^2+\sigma^2) &= \sum_{r=0}^{k} \frac{n_r}{C_r(\beta)} \int_{-\infty}^{\infty} \theta^2 \frac{\exp(r\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \varphi(\theta) d\theta. \qquad (3.31)
\end{aligned}$$

Platí následující vztahy:

- $$\sum_{\{\mathbf{z}:z_j=1\}} n_{\mathbf{z}} = x_{\cdot j}; \qquad (3.32)$$

- pro $j = 1, 2, \ldots, k$

$$\begin{aligned}
E[X_{\cdot j}|\theta] &= E[\sum_{\nu=1}^{n} X_{\nu j}|\theta] = \sum_{\nu=1}^{n} E[X_{\nu j}|\theta] = \\
&= \sum_{\nu=1}^{n} \frac{\exp(\theta-\beta_j)}{(1+\exp(\theta-\beta_j))} = n\frac{\exp(\theta-\beta_j)}{(1+\exp(\theta-\beta_j))}; \qquad (3.33)
\end{aligned}$$

- z rozdělení vektoru odpovědí $\mathbf{x}$ za podmínky θ, tj. z

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{k} \frac{\exp(x_i(\theta-\beta_i))}{(1+\exp(\theta-\beta_i))} = \frac{\exp(r\theta)\exp(-\sum_{i=1}^{k}\beta_i x_i)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \qquad (3.34)$$

a z 3.21 vyplývá, že

$$
\begin{aligned}
f(\theta|\mathbf{x}) &= \frac{f(\mathbf{x}|\theta)\varphi(\theta)}{\int\limits_{-\infty}^{\infty} f(\mathbf{x}|\theta)\varphi(\theta)} = \\
&= \frac{\exp(r\theta)\varphi(\theta)}{C_r(\beta)\prod\limits_{i=1}^{k}(1+\exp(\theta-\beta_i))} = f(\theta|r). \qquad (3.35)
\end{aligned}
$$

Rovnice 3.29, 3.30 a 3.31 upravíme pomocí 3.32, 3.33, 3.34 a 3.35 do následujícího konečného tvaru:

$$
\begin{aligned}
x_{.j} &= \frac{1}{n}\sum_{r=0}^{k} n_r E[E[X_{.j}|\theta]|r] \qquad j = 1,2,\ldots,k \qquad (3.36) \\
n\mu &= \sum_{r=0}^{k} n_r E[\theta|r] \qquad (3.37) \\
n(\mu^2+\sigma^2) &= \sum_{r=0}^{k} n_r E[\theta^2|r] \qquad (3.38)
\end{aligned}
$$

Rovnice 3.36, 3.37 a 3.38 vstupují do iteračního procesu, jejich vyřešením dostáváme odhady parametrů $\beta_1, \beta_2, \ldots, \beta_k, \mu, \sigma$. Poznamenejme, že z analýzy nevylučujeme řádky či sloupce matice $\mathbf{x_{n\times k}}$, které mají nulový či maximální možný součet oproti předchozím metodám. I zde počítáme s omezením 3.1.

Kapitola 4

Odhad parametru schopnosti zkoušeného

Prvotním cílem při vyhodnocování didtaktických testů nejsou odhady parametrů otázek, ale zjištění odhadu parametrů schopnosti jednotlivých testovaných. Abychom vyřešili tento problém, použijeme:

1. metodu maximální věrohodnosti - v případě, kdy nic nepředpokládáme o tvaru rozdělení schopností;
2. bayesovské metody - v případě, kdy předpokládáme nebo máme odhadnutý tvar rozdělení schopností.

4.1 Metoda maximální věrohodnosti

Odhad θ je získán z maximalizace logaritmu pravděpodobnosti řádkového skóre r za podmínky θ. Tato pravděpodobnost je nepřímo uvedená v 3.9, když jsme odvozovali pravděpodobnost matice $\mathbf{X_{n\times k}}$ za podmínky znalosti jejích řádkových skórů, tj.

$$f(r) = (\prod_{i=1}^{k} \frac{1}{1+\xi\epsilon_i})(\xi)^r \gamma_r(\epsilon). \tag{4.1}$$

Zlogaritmováním 4.1 dostaneme (připomeňme, že $\xi = \exp(\theta)$ a $\epsilon_i = \exp(-\beta_i)$):

$$\ln f(r) = -\sum_{i=1}^{k} \ln(1+\exp(\theta-\beta_i)) + r\theta + \ln\gamma_r(\exp\beta_1, \exp\beta_2, \ldots, \exp\beta_k). \tag{4.2}$$

Derivováním 4.2 a položením nule dostaneme rovnici

$$r = \sum_{i=1}^{k} \frac{\exp(\theta - \beta_i)}{(1 + exp(\theta - \beta_i))}, \tag{4.3}$$

kam za β_i dosadíme jejich odhady získané metodou podmíněné maximální věrohodnosti (viz podkapitola 3.2). Odhady neexistují pro nulové a maximální skóre.

4.2 Užití bayesovských metod

Jestliže máme odhadnuto rozdělení schopností (nechť má hustotu $g(\theta)$), můžeme odvodit bayesovský modální odhad tím, že budeme maximalizovat vzhledem k θ apostriorní hustotu $f(\theta|r)$, kterou jsme odvodili v 3.35, tj. maximalizujeme vzhledem k θ výraz

$$\frac{\exp(r\theta)g(\theta)}{C_r(\beta)\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} \propto \frac{\exp(r\theta)g(\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))}, \tag{4.4}$$

kam za β_i $i = 1, 2, \ldots, k$ dosadíme jejich odhady, které jsme získali metodou marginální maximální věrohodnosti (viz podkapitola 3.3), přičemž předpokládáme, že hustota $g(\theta)$ je hustota normálního rozdělení $\varphi(\theta)$.

Další možností využití bayesovských metod pro odhad θ je očekávaný aposteriorní odhad θ, tj.

$$\widehat{\theta} = E[\theta|data]. \tag{4.5}$$

Vzhledem k tomu, že r je suficientní pro θ stačí odvodit podobně jako u bayesovského modálního odhadu očekávanou hodnotu θ pouze za podmínky r, tj.

$$E[\theta|r] = \frac{1}{C_r(\beta)} \int_{-\infty}^{\infty} \theta \frac{\exp(r\theta)g(\theta)}{\prod_{i=1}^{k}(1+\exp(\theta-\beta_i))} d\theta, \tag{4.6}$$

kam dosadíme (jako u bayesovského modálního odhadu) za β_i jejich odhady, které jsme získali metodou marginální maximální věrohodnosti (rozdělení schopností je předpokládáno normální, $g(\theta)$ je opět i zde $\varphi(\theta)$).

Příjemnou vlastností těchto dvou bayesovských odhadů je, že existují pro nulové i maximální skóre r oproti odhadu metodou maximální věrohodnosti.

Kapitola 5

Příklad odhadů parametrů na reálných datech

Abychom nezůstali jen u teoretického odvozování, budeme demonstrovat odhad parametru obtížnosti a parametru schopnosti na konkrétních datech datech. Parametr obtížnosti odhadneme pomocí metody podmíněné maximální věrohodnosti (viz podkapitola 3.2) a parametr schopnosti pomocí metody maximální věrohodnosti (viz podkapitola 4.1).

Použijeme výsledky přijímacího řízení na PedF UK v roce 2002 na obor speciální pedagogika. Test, který byl při přijímacích zkouškách na speciální pedagogiku zadáván, měl zjišťovat předpoklady uchazečů k vysokoškolskému studiu oboru speciální pedagogika a zároveň předpoklady k budoucímu vykonávání profese speciálního pedagoga.

Test byl složen z dvaceti otázek a řešilo ho celkem 144 uchazečů. V tabulce 5.1 jsou shrnuty údaje vyplývající z dat, které použijeme při odhadování parametrů obtížnosti a parametrů schopnosti. (Data nalezneme na disketě přiložené k diplomové práci.) Tabulka obsahuje číslo otázky (je označeno jako "otázka j"); počet testovaných, kteří správně odpověděli na danou otázku (v souladu s předchozím textem je tento počet v tabulce označen "$x_{\cdot j}$"); skóre, kterého mohli testovaní dosáhnout, tj. máme-li 20 otázek, dosažená skóre mohou být $0, 1, \ldots, 20$ (označeno jako "r"); počet testovaných se skórem r (označeno jako "n_r").

Nejprve odhadneme parametry obtížnosti pomocí již zmíněné metody podmíněné maximální věrohodnosti, tj. využijeme rovnic 3.15 a 3.1. Aby metoda konvergovala je nutné vyloučit nulové a maximální součty v řádcích a sloupcích datové matice, která má v našem případě rozměry 144×20. V našem případě bylo nutné vyloučit jeden řádek v datové matici

otázka j	1	2	3	4	5	6	7	8	9	10
$x_{.j}$	109	65	50	67	70	100	89	115	99	123
otázka j	11	12	13	14	15	16	17	18	19	20
$x_{.j}$	133	12	127	89	109	134	112	122	124	73

r	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_r	1	0	0	0	0	0	2	1	7	5	9	12	20	15	18	14	14	16	4	5	1

Tabulka 5.1: Shrnutí vstupních dat

- jeden z testovaných dosáhl maximálního skóre.

Samotné numerické řešení rovnic 3.15 je problematické a to zejména s ohledem na vyčíslení funkcí γ_i. Pro výpočet γ_i je nutné vygenerovat všechny vektory o délce $k = 20$ obsahující nuly a jedničky a to tak, aby součet jedniček byl roven i. Pro $k = 20$ trvá výpočet γ_i a následně γ_i^j pro $i = 1, 2, \dots, k$ a $j = 1, 2, \dots, k$ několik minut. Odvoďme proto následující vztahy:

Tvrzení 2

Pro definované veličiny γ v 3.10 platí následující rekurentní vztahy (pro $i = 1, 2, \dots, k$ a $j = 1, 2, \dots, k$):

$$\gamma_i = \frac{\gamma_{(i-1)}\gamma_1 - \sum_{j=1}^{k} \epsilon_j^2 \gamma_{(i-2)}^{(j)}}{i}, \tag{5.1}$$

$$\gamma_i^{(j)} = \gamma_i - \epsilon_j \gamma_{(i-1)}^{(j)}, \tag{5.2}$$

přičemž $\gamma_i = 1$ pro $i = 0$ a $\gamma_i = 0$ pro $i < 0$. Stejné vztahy platí pro $\gamma_i^{(j)}$, tj. pro libovolné j je $\gamma_i^{(j)} = 1$, je-li $i = 0$ a $\gamma_i^j = 0$, je-li $i < 0$.

Konstrukce:

Z definice 3.10 vyplývají následující vztahy

$$\gamma_i = \sum_{(\mathbf{y}|i)} \prod_{j=1}^{k} \epsilon_j^{y_j}, \tag{5.3}$$

$$\gamma_{i-1} = \sum_{(\mathbf{y}|i-1)} \prod_{j=1}^{k} \epsilon_j^{y_j}, \tag{5.4}$$

$$\gamma_1 = \sum_{j=1}^{k} \epsilon_j. \tag{5.5}$$

Vynásobíme-li γ_{i-1} a γ_1, dostáváme vztah

$$\gamma_{i-1}\gamma_1 = \sum_{i=1}^{k} \sum_{(\mathbf{y}|i-1)} \epsilon_i \prod_{j=1}^{k} \epsilon_j^{y_j} \tag{5.6}$$

Jak se liší 5.3 od 5.6? Výraz 5.6 obsahuje všechny i–tice epsilonů, které jsou i v 5.3, protože všechny $(i-1)$–tice pronásobuje všemi epsilony. Každá i–tice však může vzniknout z $(i-1)$–tic i způsoby. Např. pro $k = 7$ trojice $\epsilon_1 \cdot \epsilon_3 \cdot \epsilon_7$ může vzniknout přinásobením ϵ_1 ke dvojici $\epsilon_3 \cdot \epsilon_7$ nebo přinásobením ϵ_3 ke dvojici $\epsilon_1 \cdot \epsilon_7$ a nebo přinásobením ϵ_7 ke dvojici $\epsilon_1 \cdot \epsilon_3$. Tedy třema způsoby. Proto je každá i-tice v $\gamma_{i-1}\gamma_1$ reprezentována i krát. Podělme proto $\gamma_{i-1}\gamma_1$ hodnotou i.

V 5.6 oproti 5.3 jsou i výrazy, kde se přinásobené ϵ_i sešlo s $(i-1)$–ticí, která již ϵ_i obsahovala. Tyto typy výrazů jsou v $\gamma_{i-1}\gamma_1$ jednou. Po podělení $\gamma_{i-1}\gamma_1$ hodnotou i je třeba odečíst jejich i–tinu. Uvědomme si, že jde o výrazy, kde je jeden prvek z $\epsilon_1, \epsilon_2, \ldots, \epsilon_k$ reprezentován právě dvakrát a zbylých je $i-2$. Pro všechna ϵ_i odpovídají tyto výrazy výrazu $\sum_j \epsilon_j^2 \gamma_{(i-2)}^{(j)}$. Tímto jsme zkonstruovali vztah 5.1. Podobným způsobem lze ukázat správnost vztahu 5.2.

◇

Použijeme-li pro vyčíslení funkce γ rovnice 5.1 a 5.2, trvá jejich numerické vyčíslení jen několik sekund, čímž jsme se vyrovnali s problémem rychlosti konvergence. K řešení samotných rovnic 3.15 a 3.1 byla využita metoda Newtonova typu programovaná v systému *Ox 3.30*.[1]

Odhady parametrů obtížnosti jsou uvedeny v tabulce 5.2, v níž je uvedeno číslo otázky (označeno jako "otázka j") a odhad parametru β_j vypočtený pomocí metody podmíněné maximální věrohodnosti (označen jako "$\widehat{\beta_j}$").

Dále se zaměříme na odhad parametru schopnosti jednotlivých testovaných, k čemuž využijeme rovnici 4.3. Připomeňme, že používáme metodu maximální věrohodnosti. V tabulce 5.3 je uvedeno skóre, kterého mohl testovaný dosáhnout (označeno jako "r") a odhad parametru schopnosti vypočtený pomocí metody maximální věrohodnosti (označeno jako "$\widehat{\theta}$"). Pro skóre $r = 0$ a $r = 20$ odhady parametru θ neexistují - políčka v tabulce jsou proškrtnuta.

[1]Ox je registrovanou značkou společnosti Timberlake consultants. Pro akademické účely je k dispozici zdarma.

otázka j	1	2	3	4	5	6	7	8	9	10
$\widehat{\beta}_j$	-0,943	1,088	1,589	1,023	0,928	-0,064	0,316	-0,948	-0,028	-0,963
otázka j	11	12	13	14	15	16	17	18	19	20
$\widehat{\beta}_j$	-0,969	3,578	-0,966	0,316	-0,938	-0,970	-0,952	-0,963	-0,964	0,832

Tabulka 5.2: Odhad parametru obtížnosti pomocí metody podmíněné maximální věrohodnosti

r	0	1	2	3	4	5	6	7	8	9	10
$\widehat{\theta}$	-	-3,354	-2,579	-2,086	-1,703	-1,379	-1,086	-0,813	-0,549	-0,289	-0,029
r	11	12	13	14	15	16	17	18	19	20	-
$\widehat{\theta}$	0,245	0,529	0,833	1,170	1,553	2,012	2,602	3,424	4,666	-	-

Tabulka 5.3: Odhad parametru schopnosti pomocí metody maximální věrohodnosti

Představme si nyní, že zadáme nějaké osobě pouze výběr otázek, které byly v námi analyzovaném testu. Jinými slovy jí zadáme nějaký subtest našeho testu. Na základě jejích odpovědí můžeme odhadnout její parametr schopnosti (použijeme rovnici 4.3, v níž bude suma přes indexy otázek v subtestu). Můžeme předpovídat, s jakou pravděpodobností odpoví správně na otázky, které nebyly v subtestu (ale byly v testu). Využijeme k tomu charakteristickou křivku položky jednoparametrického modelu, tj. funkci $f(\theta) = \frac{1}{1+\exp(-(\theta-\beta))}$, kam za β dosadíme parametr obtížnosti otázky, která v subtestu nebyla (ale byla v testu) a za θ dosadíme parametr schopnosti diskutované osoby. Takto lze výsledky odhadů obtížnosti položek dále využít.

U otázky číslo 12 nabyl parametr obtížnosti velké hodnoty. Lze proto doporučit, aby byla tato otázka při dalším používání přijímacího testu vyřazena.

Kapitola 6

Slovo závěrem

Komenda ve své knize [8] píše: "Schopnost měřit je pro proces vzdělávání otázkou života a smrti." S tímto heslem musí snad každý pedagog souhlasit. Vyjímky by mohly tvořit pedagogové vyznávající různé alternativní směry, mezi něž by šla řadit waldorfská či montessoriovská pedagogika. Měření znalostí slouží nejen hodnotiteli - pedagogovi, aby si utvořil představu o znalostech testovaných - žáků, ale ještě více ho potřebují právě testovaní - žáci, kteří se dozvídají, jak pokročili ve svých znalostech (někdy to může být špatné zjištění). Pro žáky je to zároveň motivační činitel. Rodiče dětí získávají přestavu, jak na tom jejich děti se znalostmi jsou.

Měření znalostí, schopností a dovedností je důležité nejen ve školním prostředí, ale i při přijímání nových zaměstnanců atd. Je proto důležité rozvíjet různé metody na zkvalitňování měření znalostí, schopností, či jiných latentních vlastností.

Jednou z cest je právě teorie odpovědí na položky, která obsahuje modely pro vyhodnocování testů měřících nějakou latentní vlastnost testovaných. V této práci byly popsány dichotomické Raschovy modely, jejichž rozšířením jsou polytomické Raschovy modely.

Slovo autora této práce: "Viděla bych určitý úspěch této práce, kdyby se někdo začal podrobně v České republice věnovat právě těmto modelům a navázal tak na teorii sepsanou v této práci.

Je však důležité, aby se zpracovaná teorie dostávala k uživatelům v praxi. Aby ti věděli, že tyto metody existují a aby znali alespoň jejich základní myšlenky. Matematici by měli tuto teorii prvotně zpracovat, protože je právě z tohoto hlediska náročnější, ale neměl by to být jejich cíl. Cílem by mělo být její co nejširší zpřístupnění lidem, kteří ji potřebují a v náročných matematických vzorcích se ztrácejí."

Literatura

[1] Aczél, J.(1996): Lectures on functional equation and their applications, Academic Press, New York.

[2] Andersen, E.B. (1973): Conditional inference and models for measuring, Mentalhygiejnisk Forlag, Copenhagen.

[3] Andersen, E.B. (1980): Discrete statistical models with social science applications, Amsterdam: North-Holland.

[4] Anděl J. (1978): Matematická statistika, SNTL, Praha.

[5] Baker F. B. (2001): The Basics of Item Response Theory, ERIC Clearinghouse on Assesment and Evaluation, University of Wisconsin.

[6] Blahuš P.(rok v článku neuveden): Elementy z psychologie a teorie testů, internetová adresa: www.psychodiagnostika.cz/index.php?akce=blahus2.

[7] Fischer G.H., Molenaar I. W. (1995): Rasch Models Foundation, Recent Developments, and Application, Springer-Verlag, New York.

[8] Komenda S., Zapletalová J. (1996) :Analýza didaktického testu a její počítačová podpora, Lékařská fakulta Univerzity Palackého, Olomouc.

[9] Pelikán, J. (1998): Základy empirického výzkumu pedagogických jevů, Karolinum, Praha.

[10] Scholz, A. a Schoenberg, B. (1995): Einfürung in die Zahlentheorie (Sammlung Goeschen, Band 1131), Walter de Gruyter, Berlin.

[11] Talacko J. (1939): Příspěvek k matematické teorii růstu populace, Rozpravy jednoty pro vědy pojistné, Praha.

[12] Wim J. van der Linden, Hambleton Ronald K. (1997): Handbook of Item Response Theory, Springer Verlag, New York.

Příloha A

Program ke zpracování dat

```
#include <oxstd.h>
#include <oxprob.h>
#import <maximize>
#import <solvenle>
#import <database>

decl K, n, x, rv,gamma,vyradit,r,beta;


//
// funkce gamma pocitana rekurentnim vzorcem ...
//

gmafnc(const eps)
{
decl i,j,ifact;

gamma = zeros(K+1,K+1);

// Nasledujou ty pocatecni hodnoty

// pro i =0
for (j=0;j<=K;j++) gamma[0][j] = 1;

// pro i = 1
gamma[1][0] = 0;
for (j=1;j<=K;j++) gamma[1][0] += eps[j];

for (j=1;j<=K;j++) gamma[1][j] =gamma[1][0] - eps[j];
```

```
// pro i = 2

gamma[2][0] = (sumc(eps)^2 - sumc(eps.^2))/2;

for (j=1;j<=K;j++) gamma[2][j] = ((sumc(eps)-eps[j])^2 - sumc(eps.^2)+eps[j]^2)/2;


// a pro zbytek budou platit nasledujici rekurzivn vztahy ...
ifact = 2;

for (i=3; i<=K; i++){
     ifact *= i;
     gamma[i][0] = gamma[i-1][0]*gamma[1][0]/i;
     for (j=1;j<=K;j++) gamma[i][0] -= eps[j]*eps[j]*gamma[i-2][j]/i;

     for (j=1;j<=K;j++){
        gamma[i][j] = gamma[i][0] - eps[j]*gamma[i-1][j];
     }
}

}


rovnice(const epsilon, const j, const x, const rv,const n, const K)
{
decl v,vys=0;

for (v=1;v<=n;v++){
  vys += epsilon[j]*gamma[rv[v]-1][j]/gamma[rv[v]][0];
}

vys=vys-x[j];
return vys;
}

systemrovnic(const aF, const epsilon)
{
decl j,j1,pocet = 0, soucin = 1;

aF[0] = 1;
for (j=1;j<=20;j++) aF[0] *= epsilon[j];

aF[0] -= 1;
```

```
gmafnc(epsilon);


for (j=1;j<=K;j++){
   if (j!=1)
     aF[0]=aF[0] | rovnice(epsilon,j,x,rv,n,K);
     }

return 1;
}

poslednirovnice(const aF, const theta)
{
decl hlp,i;

hlp = r;

for (i=1;i<=K;i++) hlp -= exp(theta-beta[i])/(1+exp(theta-beta[i]));

aF[0] = hlp;
return 1;
}


main()
{
decl AF=<0>,epsilon,dbase,x1,x2,x3,x4,rv1,rv2,rv3,rv4,KK=<0,121,71,142,65>,v,theta,X,j,razeni,rhelp;
decl i;

X=loadmat("c:/Documents and Settings/vonkad/dokumenty/ostatni/dh/Varianta2.xls",&X);

X = 0 |X;
X = 0 ~ X;

K=columns(X)-1;
n=rows(X)-1;
x=sumc(X);rv=sumr(X);
epsilon=ones(1,K+1)*0.1;


print(K,"\n",n,"\n",x,"\n",rv);


SolveNLE(systemrovnic,&epsilon,1,1,1e-3,5e-4,10000,20,5000);
```

```
for (v=1;v<=K;v++) print(rovnice(epsilon,v,x,rv,n,K),"    ");

print(" \n A ted ty thety: \n");

beta = sumc(-log(epsilon[1:]))/K;

epsilon = -log(epsilon)-beta;
beta = 0 | epsilon;


for (r=1;r<=K-1;r++){
theta = 1;

SolveNLE(poslednirovnice,&theta);

print(theta,"\n");
}

rv=<0>;
for (i=1;i<=K;i++) rv = rv | i;

razeni =  rv ~ x' ~ epsilon;

// print(razeni,"\n");

for (i=1;i<=K-1;i++)
  for(j=2;j<=K+1-i;j++)
     if (razeni[j][1]<razeni[j-1][1]){
               rhelp = razeni[j][0:2];
               razeni[j][0:2] = razeni[j-1][0:2];
               razeni[j-1][0:2]=rhelp;
     }

print(razeni);

AF=<0>;
}
```