

MODELOVÁNÍ RŮSTOVÝCH KŘIVEK

MAREK BRABEC

ABSTRAKT. This paper presents a growth curve model for human growth data. It has a nonlinear regression form, allowing for interindividual variability in growth by introducing random effects describing variability in some of the model parameters. Interpretation of model parameters as well as some other features of biological interest is mentioned. The model is fitted to a longitudinal dataset and quality of the fit is checked from various viewpoints. Estimation of various features of anthropological interest is illustrated.

Резюме: Исследуется проблема оценивания параметров функции роста детей при помощи модели случайных эффектов.

Růstové křivky, zejména ty vzniklé jako výsledek měření velikosti živých subjektů jsou zajímavým zdrojem komplikací pro formulaci statistického modelu a následnou analýzu naměřených dat. Komplikace jsou spojeny

- s výraznou *nelinearitou* (přítomností asymptoty)
- *chybějícími daty* (nestejným rozvrhem měření pro různé subjekty)
- nezanedbatelnou *korelací* mezi měřeními provedenými na stejném subjektu.

Na analýzu růstových křivek tedy můžeme pohlížet jako na jistou směs problémů nelineární regrese a časových řad. Vzhledem k tomu, že počet měření provedených na jednom jedinci je poměrně malý (např. křivka pro tělesnou výšku dosahuje asymptoty kolem 18. roku věku a normálního jedince lze jen stěží nutit k více než 1 až 2 měřením za rok), musí být model z hlediska časových řad či procesů nutně poněkud skromnější. Výhodou je naopak to, že k dispozici je vícero nezávislých realizací (řad měření různých jedinců).

1. RŮSTOVÉ KŘIVKY DĚTÍ

Typická data (laskavě poskytnutá doc. Prokopcem) kterými se budeme dále zabývat, pocházejí z longitudinální studie z níž máme k dispozici měření tělesné výšky 85 chlapců a 86 dívek od 3 měsíců do 22 let (s různým počtem měření na dítě) ukazuje obrázek 1. Co se týče tvaru růstové křivky jako funkce času, existuje celá řada víceméně empirických modelů. Snadno použitelné a proto populární jsou modely polynomického tvaru (se všemi typickými nectnostmi). Mezi antropology patří ke klasice několik modelů nelineárního charakteru a různého stupně sofistikovanosti - od po intervalech formulovaných předpisů (bez požadavku spojitosti v hraničních bodech!) [Karlberg et al., 1987] až po různé spojité modely s asymptotou [Preece a Baines, 1978], [Jolicoeur et al., 1988]. Typicky jsou v praxi takové křivky prokládány pro každého jedince zvlášť a výsledné parametry sumarizovány průměrováním. Jednou z alternativ je vícerozměrná analýza [Rao, 1965], založená na nelineárním modelu a odhadu (nestrukturované) kovarianční matice mezi pozorováními v různých časech. Ta zohledňuje závislost mezi měřeními jednoho individua, nemá však přímou biologickou interpretaci. Odhad pak vyžaduje kompletní a poměrně

2000 *Mathematics Subject Classification.* Primary 62J02; Secondary 62J10, 62P10.

Klíčová slova. Růstové křivky, individuální data, průřezová data.

rozsáhlá data (bez „děr“ způsobených chybějícími měřeními i nedodržením zcela stejných časů měření u všech jedinců).

Chceme-li růst charakterizovat alespoň trochu realisticky, měli bychom se zajímat o model popisující nejen závislost průměrné hodnoty na čase (eventuelně i kovariátách), ale i variabilitu okolo ní, při zohlednění autokorelace mezi měřeními provedenými v různých časech na stejném jedinci. Biologicky interpretovatelný model by měl umožňovat rozklad celkové variability

- na část spojenou s odlišností dynamiky růstu mezi různými jedinci (tedy takříkajíc variabilitu *strukturální*)
- část spojenou s variabilitou uvnitř daného jedince (interpretovatelnou převážně jako *měřicí chyba*).

Rozklad na inter a intra-individuální variabilitu má poměrně dlouhou tradici v kontextu lineárního modelu se smíšenými efekty, tedy i v kontextu klasické analýzy longitudinálních dat - např. [Laird a Ware, 1982], včetně tzv. „hierarchických“ modelů polynomického tvaru [Goldstein, 1995], tedy polynomů s (některými) náhodnými koeficienty. Takové modely jsou však v podstatě čistě popisného charakteru - bez interpretace svého tvaru či hodnot koeficientů. Trpí samozřejmě neduhy známými ze situace bez přítomnosti náhodných efektů a navíc problémy spojenými s tím, že variabilita koeficientů implikuje jistý, dosti striktně vymezený tvar závislosti mezi rozptylem okolo průměrné křivky a časem.

Tvar nelineární regresní funkce, stejně jako model pro náhodné efekty by měl být kromě kvality empirického proložení diktován i věcnou interpretací. V následujícím se o formulaci jednoho takového modelu pokusíme.

2. FORMULACE MODELU

Chování růstových rychlostí (derivace růstové křivky tělesné výšky, Y'), ilustrované na obrázku 2 pro jedno konkrétní dítě (odhad pomocí smoothing spline) nás motivuje k tomu uvažovat pro výšku daného jedince (v mm) změřenou v čase x_j (let) model

$$(1) \quad k + p(1 - \exp(-a(x_j - 0.25))) + \sum_{l=1}^2 \frac{p_l}{1 + \exp(-t_l(x_j - m_l))} + \epsilon_{ij}$$

kde ϵ_{ij} má v zásadě interpretaci měřicí chyby (pomineme-li jistý systematický nesoulad mezi komplikovaným procesem růstu a hladkým, poměrně jednoduchým modelem). Regresní model tedy sestává z aditivních částí; dvou logistických křivek, exponenciály a konstanty. Pro derivace máme (v souladu s typickým empirickým chováním růstových rychlostí) dva píky logistického tvaru a počáteční exponenciální pokles. Interpretace parametrů je přímočará: p , p_1 , p_2 jsou celkové přírůstky v jednotlivých „fázích“ růstu, m_1 , m_2 odpovídají polohám logistických píků, hodnoty $\frac{1}{t_1}$, $\frac{1}{t_2}$ kontrolují jejich tvar.

Není bez zajímavosti si povšimnout formální podobnosti daného problému s problémem odhadu směsi několika parametrizovaných hustot. Exponenciální a logistické části hrají roli směřovaných měř. Komplikací zde je, že váhy samozřejmě nejsou normovány (naopak, jejich součet souvisí s celkovým přírůstkem od 3 měsíců do dospělosti a jako takový je parametrem značného praktického významu).

Při modelování variability růstové dynamiky mezi subjekty bereme v úvahu empirický fakt, dobře známý antropologům, totiž že podstatná část „strukturální“ variability, je spojena s proměnlivostí posledního píku růstové rychlosti (tzv. píku

pubertálního) – a to jak jeho polohy, tak tvaru a velikosti. To nás vede k následující specifikaci celkového nelineárního modelu se smíšenými efekty. Pro subjekty $i = 1, \dots, N$ a měření uvnitř i -tého subjektu $j = 1, \dots, n_i$:

$$(2) \quad Y_i(x_j) = K_i + P_i(1 - \exp(-a(x_j - 0.25))) + \frac{p_1}{1 + \exp(-t_1(x_j - m_1))} + \frac{P_{2i}}{1 + \exp(-T_{2i}(x_j - M_{2i}))} + \epsilon_{ij}$$

Některé „parametry“ individuálního modelu bereme nyní jako náhodné (normálně rozdělené) veličiny, proměnlivé mezi subjekty. Tedy: $K_i \sim N(k, \sigma_k^2)$, $P_i \sim N(p, \sigma_p^2)$, $M_{2i} \sim N(m_2, \sigma_{m_2}^2)$, $P_{2i} \sim N(p_2, \sigma_{p_2}^2)$, $T_{2i} \sim N(t_2, \sigma_{t_2}^2)$, $\epsilon_{ij} \sim N(0, \sigma^2)$, vše navzájem nezávislé. Kromě náhodného pubertálního píku zmíněného výše tedy ještě dovoluujeme interindividuální variabilitu ve velikosti přírůstku v počáteční (exponenciální) fázi a v konstantě.

Model má 9 parametrů spojených s průměry (tzv. *pevných efektů*), 5 *komponent rozptylu*, spojených s *náhodnými efekty* (tj. prvky „strukturální“ variability) a 1 komponentu rozptylu spojenou s *reziduální variabilitou* („měřicí chybou“). Počet parametrů je samozřejmě mnohem nižší než v praxi často používaném „individuálním“ modelu, který je prokládán pro každý subjekt zvlášť (kde máme počet parametrů individuální křivky krát počet subjektů a náhodné efekty se odhadují zcela bez jakékoli struktury přes subjekty).

3. ODHADY

Odhad parametrů (simultánně pevných efektů i komponent rozptylu) byl proveden s pomocí algoritmu NLME (Non Linear Mixed Effects) [Lindstrom a Bates, 1990], implementovaného v programovém balíku S-plus. Výsledkem je maximálně věrohodný odhad pevných efektů i komponent rozptylu a empiricky Bayesovský odhad náhodných efektů. Je založen na maximalizaci funkce

$$(3) \quad g(\beta, \mathbf{b}, \theta, \sigma^2 | \mathbf{Y}) = - \frac{1}{2\sigma^2} \sum_{i=1}^N (\mathbf{Y}_i - \eta_i(\mathbf{A}\beta + \mathbf{B}\mathbf{b}_i))^T (\mathbf{Y}_i - \eta_i(\mathbf{A}\beta + \mathbf{B}\mathbf{b}_i)) - \frac{1}{2} \sum_{i=1}^N \mathbf{b}_i^T \text{diag}(\theta)^{-1} \mathbf{b}_i$$

$\beta = (k, p, a, p_1, t_1, m_1, p_2, t_2, m_2)^T$ je vektor *pevných efektů*, $\mathbf{b}_i = (K_i - k, P_i - p, P_{2i} - p_2, T_{2i} - t_2, M_{2i} - m_2)^T$ je vektor *náhodných efektů*, $\eta(\cdot)$ je dána specifikací modelu výše, \mathbf{A}, \mathbf{B} jsou příslušné matice jejichž modifikací můžeme snadno modelovat i vliv případných kovariát, $\theta = (\sigma_K^2, \sigma_P^2, \sigma_{P_2}^2, \sigma_{T_2}^2, \sigma_{M_2}^2)^T$ je vektor *komponent rozptylu* náhodných efektů, \mathbf{Y}_i je vektor pozorování na i -tém subjektu. Vidíme tedy, že prakticky jde o minimalizaci reziduálního součtu čtverců, s penalizací variability náhodných efektů. Na druhý člen pravé strany (3) se můžeme dívat i např. jako na člen spojený s (normálním) apriorním rozdělením „náhodných koeficientů“ apod. Maximalizace probíhá iterováním dvou kroků:

- (1) Kroku PD (Pseudo Data), ve kterém jsou aktualizovány hodnoty β a $\mathbf{b}_1, \dots, \mathbf{b}_N$ (při ostatních parametrech fixovaných na hodnotách z předchozího kroku). Výpočet je realizován doplněním pozorovaných dat tzv. *pseudodat* a následným řešením problému na nelineární nejmenší čtverce. Jde o obdobu metody, známé z lineárních modelů se smíšenými efekty, kde odhady

pevných i náhodných efektů jsou řešením soustavy „mixed model equations“. Ekvivalentně tyto rovnice dostaneme doplněním pozorovaných dat tzv. pseudodaty.

- (2) Kroku LME (Linear Mixed Effects), ve kterém jsou aktualizovány hodnoty β, θ, σ^2 . Je obdobou výpočtů v lineárním modelu se smíšenými efekty, kde je MLE odhad komponent rozptylu založen na marginálním rozdělení $p(\mathbf{Y}) = \int p(\mathbf{Y} | \mathbf{b})p(\mathbf{b})d\mathbf{b}$. Vzhledem k nelinearitě η je zmíněná integrace obtížná, proto je skutečná věrohodnostní funkce nahrazena aproximací založenou na normalitě a lokální linearizaci η okolo odhadů $\beta, \mathbf{b}_1, \dots, \mathbf{b}_N$ z předchozího kroku.

Počáteční hodnoty pro pevné efekty byly získány jako průměr, resp. medián parametrů individuálně proložených křivek, pro náhodné efekty pak $\mathbf{0}$. Takto získané počáteční hodnoty byly navíc opakovaně perturbovány náhodnými odchylkami (vygenerovanými z odhadnutého rozdělení náhodných efektů) pro posouzení stability konečných odhadů.

4. PROLOŽENÝ MODEL

Proložení modelu naměřenými daty se zdá být vcelku uspokojivé - viz obr. 3 reziduí pro chlapce. Otázkou zůstává do jisté míry kvalita proložení pro věk okolo 2-2.5 let. Dočasná deprese hodnot naměřených v tomto období může souviset s esovitým zakřívováním páteře na počátku „chodícího období“. Dále se můžeme zajímat o kvalitu proložení z hlediska odhadu „marginálních“ charakteristik $\mu(x) := EY(x)$, $\sigma(x) := \sqrt{varY(x)}$. Srovnání odhadů z modelu (získanými jako průměr, resp. směrodatná odchylka, simulací z modelu při odhadnutých parametrech) s odhady empirickými (tj. průměry a směrodatnými odchylkami (s) spočtenými zvláště pro jednotlivé časy) poskytuje obrázek 4. „Vybočující“ empirické odhady směrodatných odchylek v pravé části křivky jsou založeny na velmi nízkém počtu pozorování (do 4). Podobně vypadá srovnání pro kvantily, včetně poměrně extrémních (0.1 a 0.9).

Z přítomnosti individuálních náhodných efektů je zřejmé, že model implikuje jistý stupeň závislosti mezi měřeními provedenými na témže subjektu. Obrázek 5 porovnává odhad korelace mezi výškou ve věku x let a ve 3 měsících, resp. 5 letech. Vidíme, že i z hlediska již poněkud subtilnější charakteristiky jakou je korelační struktura uvnitř opakovaných měření se model nechová špatně. Obrázky také ilustrují poměrně komplikované, nestacionární chování řady měření jednoho subjektu. Není také nezajímavé si povšimnout toho, že empirické korelace naznačují, že $\lim_{\Delta \rightarrow 0} Cor(Y(x + \Delta), Y(x)) \neq 1$, tedy jev známý v kontextu „spatial analysis“ jako tzv. *nugget effect*. I tento aspekt je vcelku věrně reprodukován naším modelem s poměrně velmi jednoduchou strukturou náhodných efektů.

Odhady parametrů (pevných efektů i komponent rozptylu) shrnuje pro chlapce i dívky následující tabulka.

Efekt	Parametr	Chlapci	Dívky
Pevné efekty	m_1 (let)	7.0	5.3
	m_2 (let)	13.4	11.2
	p_1 (mm)	438.8	383.1
	p_2 (mm)	310.7	339.8
	t_1	0.479	0.593
	t_2	1.011	0.801
	a	0.435	0.586
	p (mm)	431.8	353.7
	k (mm)	617.9	595.4
Komponenty rozptylu	σ (mm)	10.62	9.98
	σ_{m_2} (let)	0.81	0.68
	σ_{p_2} (mm)	41.0	42.1
	σ_{t_2}	0.209	0.132
	σ_p (mm)	46.1	32.4
	σ_k (mm)	17.2	18.2

Odtud jasně vidíme některé dobře známé rozdíly v dynamice růstu mezi pohlavími. Např. poslední pík růstové rychlosti („pubertální spurt“) nastupuje u dívek dříve a je poněkud plošší. U dívek je patrný také rychlejší pokles růstové rychlosti v období po narození.

5. ODVOZENÉ CHARAKTERISTIKY

Z modelu lze poměrně snadno získat informaci o poněkud subtilnějších charakteristikách růstových křivek. Derivací odhadu průměrné křivky lze (za předpokladu zaměnitelnosti pořadí derivování a integrování) snadno získat křivku „průměrné růstové rychlosti“ či zrychlení jako explicitní funkce parametrů. Obrázek 6 obsahuje takové odhady $\frac{d}{dx}\widehat{\mu}(x)$ a $\frac{d^2}{dx^2}\widehat{\mu}(x)$.

Mezi antropology se těší značné popularitě i charakteristiky komplikovanější, jako je např. poloha maxima rychlosti v pubertálním spurtu, spolu s hodnotou výšky a růstové rychlosti v tomto bodě. Zájem je přitom nejen o odhad populačního průměru ale i o odhad variability takových veličin mezi individui. Pro pubertální spurt dostáváme z proloženého modelu následující hodnoty:

Veličina	Pohlaví	Střední hodnota	Směrodatná odchylka
Poloha maxima, x_0 (v letech)	Chlapci	13.2	0.8
	Dívky	10.9	0.7
Výška v maximu rychlosti, $Y(x_0)$ (v mm)	Chlapci	1609	52
	Dívky	1471	39
Maximální rychlost, $\dot{Y}(x_0)$ (v mm/rok)	Chlapci	88	13
	Dívky	76	7.2

6. PODMÍNĚNÉ STANDARDY

Jedním z hlavních praktických důvodů pro měření růstu dětí je potřeba tzv. „růstových standardů“ pro děti různého věku jako prostředku pro posuzování toho zda daný jedinec roste méně, stejně nebo více než je pro danou populaci typické. Standardy jsou tedy dány pomocí několika kvantilů rozdělení výšky. Jsou stanovovány

různým způsobem, v praxi však nejčastěji s referencí k normalitě, z průměru a směrodatné odchylky pro děti daného věku. V každém případě však po jednotlivých hodnotách věku (bodově). Takovéto standardy lze samozřejmě konstruovat i z tzv. *průřezových studií*, kdy jsou změřeny ve stejné době děti různého věku, každé jen jednou. Naopak longitudinální data spolu s naším modelem umožňují zkonstruovat kromě výše zmíněných standardů „marginálních“ i standardy „podmíněné“ - tedy standardy typu podmíněných kvantilů (kvantil výšky ve věku x_2 pro jedince, jehož výška ve věku $x_1 < x_2$ je dána). Takové podmíněné standardy mohou být prakticky užitečné např. pokud má být posouzeno opakované měření jednoho jedince, jež byl při prvním měření dle „marginálního“ standardu posouzen jako „malý“ s cílem rozeznat zda se mezi prvním a druhým měřením jeho původní zaostávání v růstu ještě dále prohlubuje nebo zda je jeho růst již vpořádku byt je jeho celková výška nižší díky nižšímu tempu růstu před prvním měřením. V praxi není neobvyklé, že bodově konstruovaný standard je nesprávně použit jako standard simultánní pro obě měření. Obrázek 7 ilustruje jak se rozchází (aproximace pro) podmíněný standard a standard marginální. Zobrazena je střední hodnota (pro chlapce), spolu s 20. a 80. percentilem. Plnou čarou marginální, tečkovaně podmíněné hodnoty, spočtené pro každý čas zvlášť (za podmínky, že výška 5 letých je na úrovni marginálního 20. percentilu).

Podmíněný standard lze snadno získat z našeho modelu s pomocí jednoduché aproximace dvojrozměrného rozdělení výšek v časech x_1, x_2 normálním rozdělením jehož momenty jsou získány Taylorovým rozvojem. S ohledem na nelinearitu regresní funkce vzhledem k náhodným efektům, nedostaneme střední hodnotu $\mu(x)$ pouhým vyhodnocením „náhodných parametrů“ v jejich střední hodnotě. Takovéto vyhodnocení, $\mu_1(x)$ je jen první aproximací. Druhé přiblížení pak dostáváme jako $\mu_2(x) := \mu_1(x) + (\exp(-t_2(x - m_2)) - 1) \frac{p_2 \exp(-t_2(x - m_2))}{2(1 + \exp(-t_2(x - m_2)))^3} (\sigma_{m_2}^2 t_2^2 + \sigma_{t_2}^2 (x - m_2)^2)$. Z provedených srovnání je patrné, že i takto jednoduchá aproximace dává překvapivě dobré výsledky. Podobně dostáváme přibližnou marginální směrodatnou odchylku $\sigma_1^2(x)$ i kovarianci $c_1(x_1, x_2)$ mezi měřeními v časech x_1 a x_2 . Přibližné podmíněné normální rozdělení $Y(x_2)$ dáno $Y(x_1) = y_1$ má pak momenty $\mu_2(x_2) + \frac{c_1(x_1, x_2)}{\sigma_1^2(x_1)}(y_1 - \mu_2(x_1))$ a $\sigma_1^2(x_2) - \frac{c_1^2(x_1, x_2)}{\sigma_1^2(x_1)}$. Odtud pak již snadno dostáváme i kvantily.

Jiným, mezi antropology živě diskutovaným tématem je i spor o „křížení růstových kanálů“, tedy spor o to nakolik je či není obvyklé, že jedinec nacházející se svou výškou v intervalu $I_1 := (F^{-1}(p_{11}, x_1), F^{-1}(p_{12}, x_1))$ daném „marginálními“ kvantily v čase x_1 je v čase x_2 mimo interval $I_2 := (F^{-1}(p_{21}, x_2), F^{-1}(p_{22}, x_2))$ s $p_{21} = p_{11}, p_{22} = p_{12}$. Pomocí výše popsané normální aproximace snadno spočteme pravděpodobnost $P\{Y(x_2) \in I_2 | Y(x_1) \in I_1\}$ pro libovolné I_1, I_2 . Obrázek 8 ukazuje tuto pravděpodobnost pro chlapce s $p_{11} = 0.1, p_{12} = 0.2$, tři různá x_1 a o 2.5 roku pozdější x_2 : 9.2; 11.7 (plnou čarou), 4.2; 6.7 (tečkovanou čarou), 0.25; 2.75 (čerchovanou čarou). Čísla 1; ...; 10 na horizontální ose odpovídají volbě I_2 s $p_{21} = 0, p_{22} = 0.1; \dots; p_{21} = 0.9, p_{22} = 1$. Vidíme tedy, že tendence zůstat ve stejném tzv. „růstovém kanálu“ (intervalu vymezeném stejnými kvantily) silně závisí na uvažovaném počátečním věku - nejvyšší je mezi obdobím počátečního prudkého růstu a obdobím variabilního posledního píku růstové rychlosti. Variabilita tohoto píku (tvaru i polohy) ke „křížení růstových kanálů“ výrazně přispívá.

7. PRŮŘEZOVÉ STUDIE

Model, o kterém jsme mluvili byl navržen pro longitudinální data. Zajímavé může být, jak se chová v situaci na kterou nebyl cvičen, tedy při použití na data ze studie průřezové. Je třeba poznamenat, že rozdíl mezi těmito typy studií nespočívá jen ve formátu dat (longitudinální s opakovanými měřeními téhož jedince, průřezové jen s jedním měřením), ale částečně i ve struktuře studované populace. Jde o to, že v longitudinální studii jsou typicky sledováni jedinci s podobným rokem narození, zatímco v průřezové studii se kalendářní rok narození samozřejmě značně liší. To může hrát podstatnou roli v souvislosti s tzv. sekulárním trendem (tak antropologové říkají odlišnostem v dynamice růstu různých ročníků narození) např. pozdější ročníky mají tendenci nabývat dospělé výšky dříve.

Průřezová data (ve formě průměrů a směrodatných odchylek výšek, spočtených po jednotlivých letech věku) byla proložena aproximacemi zmíněnými výše.

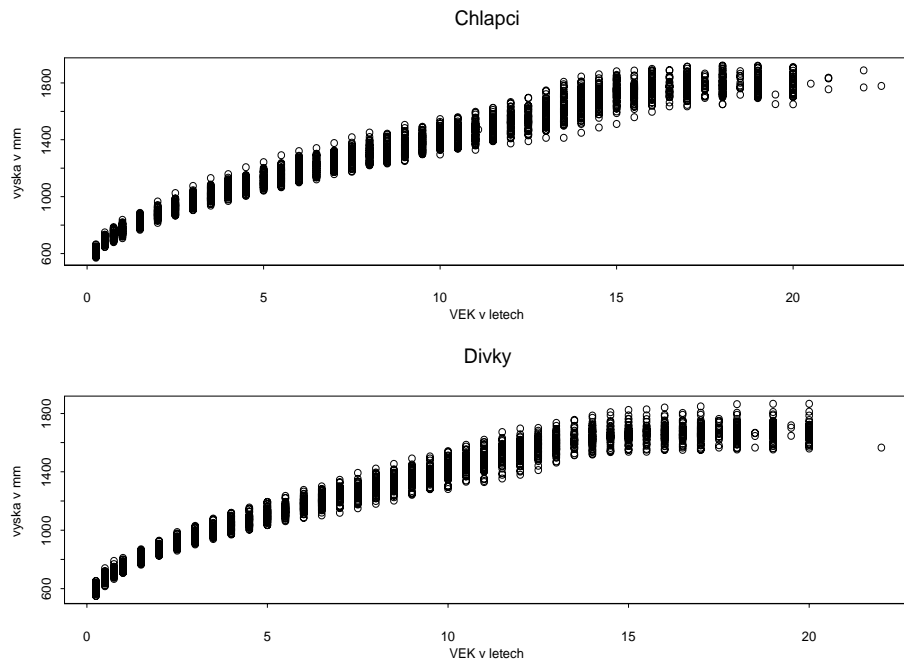
- Nejprve průměry aproximací střední hodnoty dané modelem,
- potom směrodatné odchylky aproximací $\sigma_1(x)$ (přičemž byly odhadnuty komponenty rozptylu, pevné efekty byly fixovány na hodnotách z prvního kroku).

Obrázek 9 srovnává data pro chlapce ze studie [Lhotská et al., 1993] (průměry resp. směrodatné odchylky jako kroužky) a odhady z (aproximace) modelu (plná čára). Vidíme, že i v této situaci není proložení špatné, včetně vystižení poměrně komplikovaného tvaru závislosti marginální variability na věku. Mírné podhodnocení marginální směrodatné odchylky v předpubertálním období může být způsobeno přítomností sekulárního trendu, který vnaší do dat další zdroj variability (spojený s rozdílným *tvarem* růstové křivky pro různé ročníky narození zastoupené ve studii) s nímž longitudinální model nepočítá. Toto vysvětlení je podpořeno i tím, že proložení je lepší pro pozdější průřezové studie, v nichž je dle antropologů sekulární trend mezi sledovanými ročníky slabší.

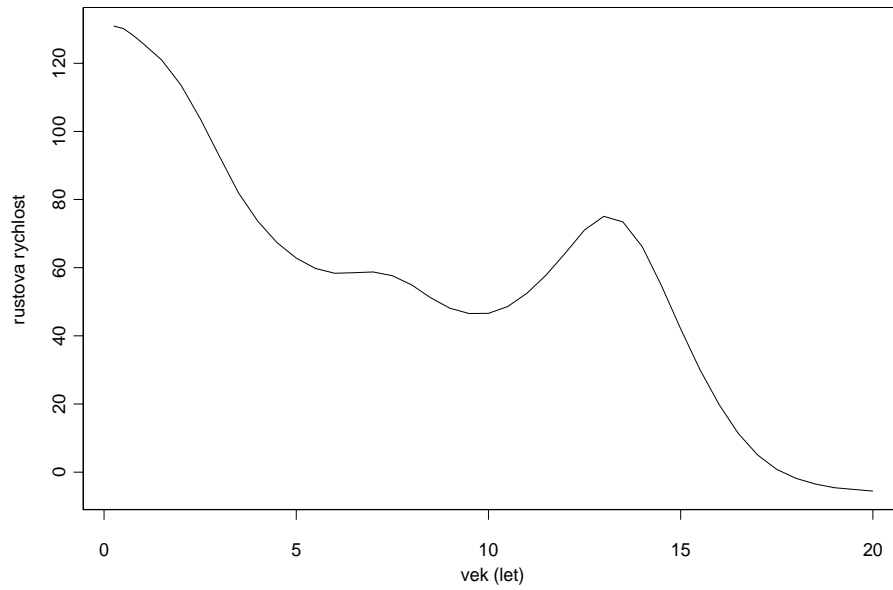
LITERATURA

- [Goldstein, 1995] Goldstein H. (1995): Multilevel statistical models. (2. vyd.) Edward Arnold. London.
- [Jolicoeur et al., 1988] Jolicoeur P., Pontier J., Pernin M.O., Sempe M. (1988): A lifetime asymptotic growth curve for human height. *Biometrics*, 44, 995-1003
- [Laird a Ware, 1982] Laird N.M., Ware J.H. (1982): Random-effects models for longitudinal data. *Biometrics*, 38, 963-974
- [Lhotská et al., 1993] Lhotská L., Bláha P., Vignerová J., Roth Z., Prokopec M. (1993): 5. celostátní antropologický výzkum dětí a mládeže, 1991 (české země). Antropologické charakteristiky. SZÚ Praha.
- [Lindstrom a Bates, 1990] Lindstrom M.J., Bates D.M. (1990): Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46, 673-687
- [Karlberg et al., 1987] Karlberg J., Fryer J.G., Engstrom I., Karlberg P. (1987): Analysis of linear growth using a mathematical model. II. From 3 to 21 years of age. *Acta Paediatr. Scand. Suppl.* 337, 12-29
- [Preece a Baines, 1978] Preece M.A., Baines M.J. (1978): A new family of mathematical models describing the human growth curve. *Annals of Human Biology* 5, 1-24
- [Rao, 1965] Rao C.R. (1965): The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, 52, 447-458

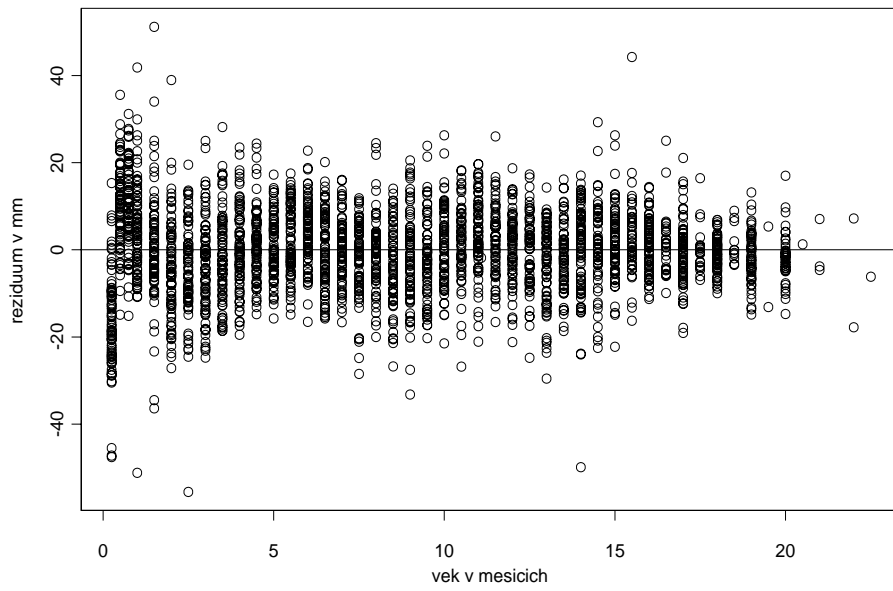
SZÚ MSP, ŠROBÁROVA 48, 100 42 PRAHA 10
E-MAIL: marek.brabec@szu.cz



OBRÁZEK 1. Data

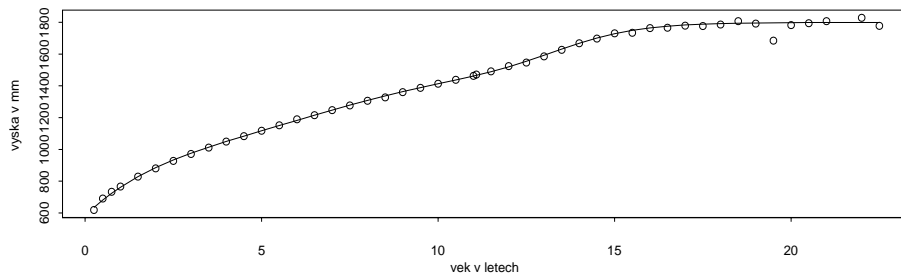


OBRÁZEK 2. Odhad růstové rychlosti jednoho chlapce

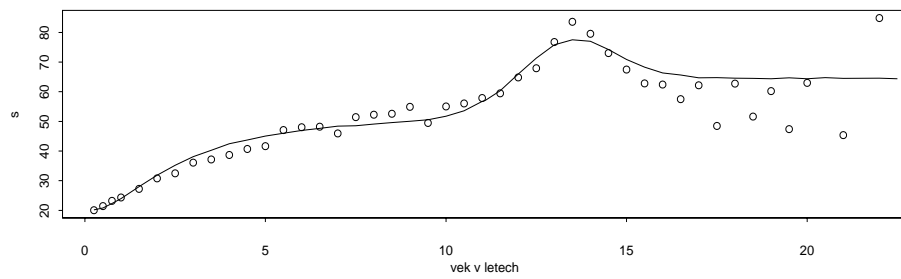


OBRÁZEK 3. Rezidua pro chlapce

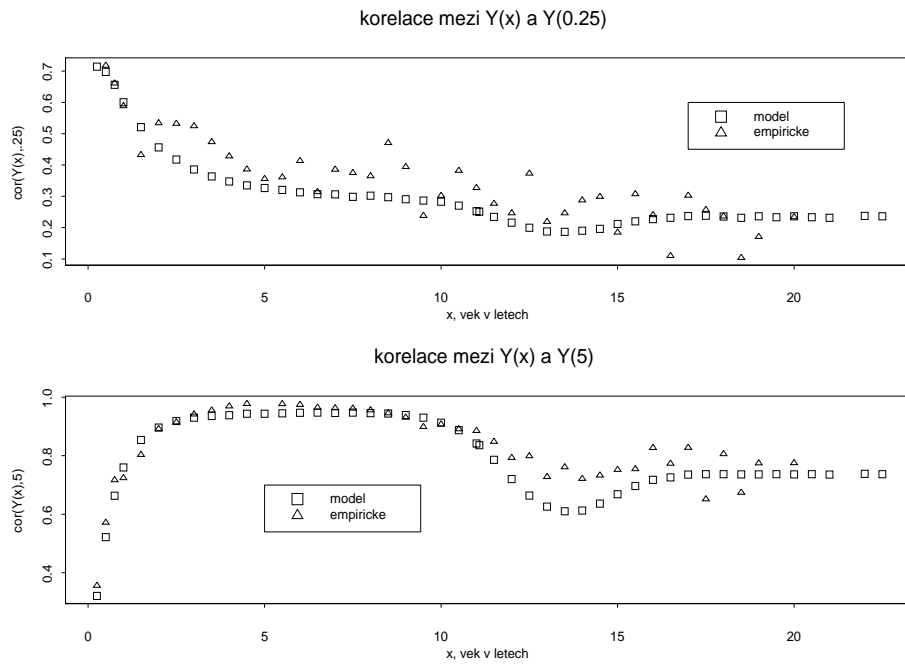
Chlapci, model vs. empirická růstová křivka



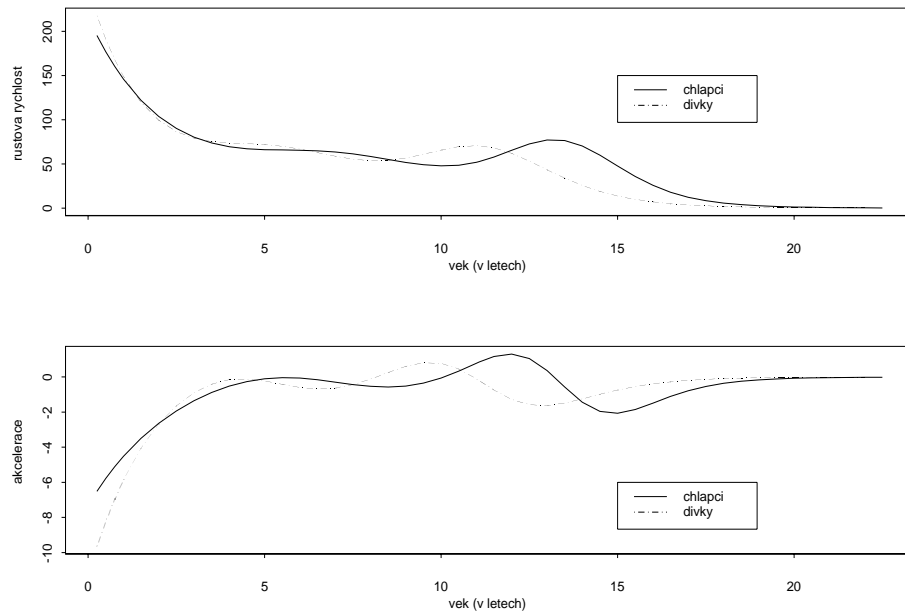
Chlapci, s, model vs. empirická

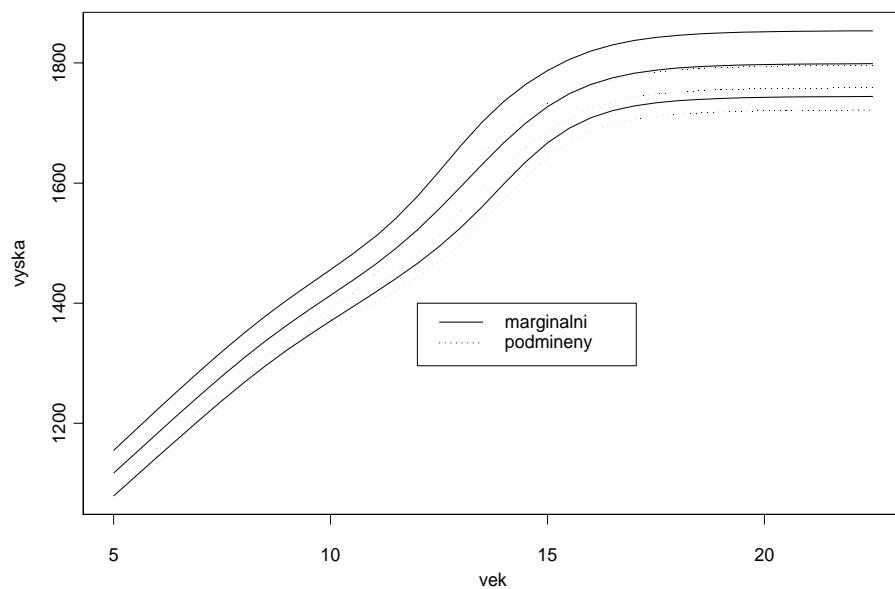


OBRÁZEK 4. Marginální charakteristiky pro chlapce

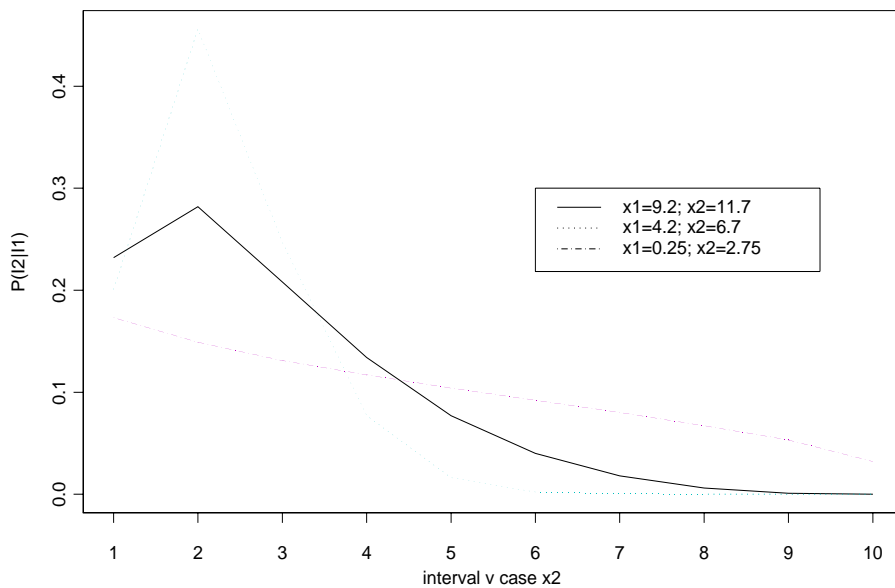


OBRÁZEK 5. Korelace mezi výškou ve dvou vybraných časech, chlapi

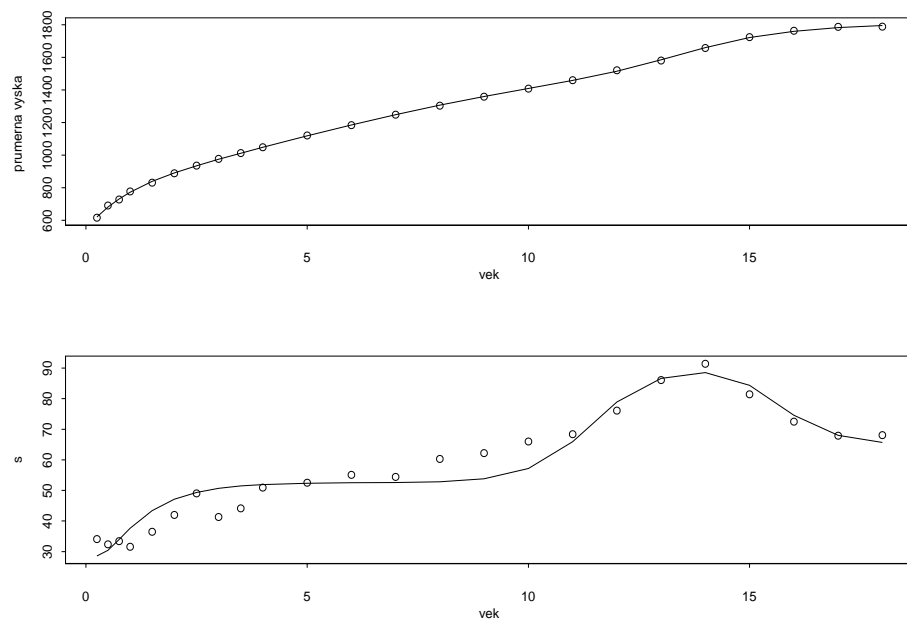
OBRÁZEK 6. $\frac{d}{dx}\widehat{\mu}(x)$ a $\frac{d^2}{dx^2}\widehat{\mu}(x)$



OBRÁZEK 7. „Marginální“ a „podmíněný standard“, chlapci



OBRÁZEK 8. Pravděpodobnost přechodu mezi „růstovými kanály, chlapci“



OBRÁZEK 9. Proložení průřezových dat pro chlapce