

Fakulta informatiky a statistiky

The background of the cover features a blurred cityscape with silhouettes of people walking. The silhouettes are rendered in a gradient of red and blue, creating a sense of movement and urban activity. The city buildings in the background are also blurred, emphasizing the foreground figures.

TEORIE

DEMOGRAFICKÝCH

PROJEKCÍ

Vysoká škola ekonomická v Praze

Filip Hon a kolektiv

Vysoká škola ekonomická v Praze

TEORIE

DEMOGRAFICKÝCH PROJEKČÍ

Filip Hon

Tomáš Fiala

Jitka Langhamrová

2024

VŠE / NAKLADATELSTVÍ
OECONOMICA

Autoři:

Ing. Filip Hon, Ph.D.

RNDr. Tomáš Fiala, CSc.

doc. Ing. Jitka Langhamrová, CSc.

Recenzenti:

doc. Ing. Jan Mertl, Ph.D.

Ing. Ondřej Šimpach, Ph.D.

Obsah

Předmluva	5
1. Základní terminologie demografických projekcí	7
2. Extrapolační projekční metody	13
3. Kohortně komponentní metoda bez migrace	19
4. Přidání migrace do kohortně komponentní metody	23
5. Tvorba scénářů v kohortně komponentní metodě – deterministický přístup	29
6. Tvorba scénářů s využitím modelu Lee-Carter a jeho modifikací – stochastický přístup	37
7. Problematika výpočtu projekce malých územních celků	43
8. Odvozené projekce	49
9. Metodika a hlavní výsledky projekcí významných pracovišť	53
Seznam obrázků	60
Literatura	61
Rejstřík	63

Předmluva

Každá společnost ve svém vlastním zájmu určitým způsobem usměrňuje svůj vývoj žádoucím směrem. Při tom se ale neobejde bez realistických a vědecky podložených představ o budoucím vývoji, tedy bez prognóz.

Základním prvkem každého sociálního systému jsou lidé. Právě početní velikost populace, její pohlavní a věková struktura a jejich budoucí vývoj jsou hlavním předmětem demografických neboli populačních prognóz.

„Prognóza může být definována jako nepodmíněná, na vědeckém poznání založená výpověď o očekávaném a v době jejího vzniku nejpravděpodobnějším budoucím vývoji sledovaného jevu.“ (viz Kučera, 1998)

„Populačními odhady v širokém smyslu můžeme chápat veškeré odhady počtu obyvatel a jeho struktur do budoucna a do minulosti, tj. ty, které nejsou přímým výsledkem statistického šetření.“ (Pavlík a kol., 1986) „Z časového hlediska můžeme populační odhady uvažovat do minulosti, kdy jde obvykle o intercensální odhady pomocí interpolace, nebo do budoucna, kdy počítáme demografické projekce, většinou extrapolací současných populačních trendů. Pod pojmem demografická projekce budeme rozumět souhrn výpočtů, kterými odhadujeme další vývoj populace. Demografické projekce mohou, ale také nemusí sloužit k předpovědím budoucího populačního vývoje. Někdy může být jejich smyslem pouze analýza současného stavu populační reprodukce, jsou určitým modelem ukazujícím např., jak by probíhal budoucí populační vývoj za předpokladu daných úrovní plodnosti, úmrtnosti a migrace, anebo naopak při uvažování různých kombinací jejich změn. Takové projekce se někdy označují jako projekce varovné. Pouze populační projekce, které se snaží poskytnout pokud možno nejspolehlivější předpověď budoucího vývoje, označíme za populační prognózy. Podle výše uvedeného vymezení jsou zvláštním případem populačních projekcí.“ (Pavlík a kol., 1986)

Cílem této odborné knihy je seznámit čtenáře se základem teorie demografických projekcí. Kniha předkládá matematickou definici základních nejpoužívanějších modelů pro účely predikce budoucího počtu a věkové a pohlavní struktury obyvatel. Současně jsou zde obsaženy ukázky potenciálních možných modifikací klasických modelů pro účely modernizace a zpřesnění výsledků odhadů budoucího populačního vývoje.

Kniha je určena čtenářům se zájmem o problematiku demografického prognózování a také se základní znalostí demografického značení a demografických ukazatelů (viz např. Langhamrová, Šimpach, 2013), které zde nebudou definovány od zcela triviálních základů.

Kniha je rozčleněna na několik kapitol, v úvodních kapitolách je věnován prostor koherentně komponentní metodě výpočtu populační projekce. Následně se publikace zabývá možným zpřesněním výsledků pomocí přidání vlivu migrace, zahrnutím specifík malých celků a definováním scénářů různých výchozích předpokladů. Osmá kapitola knihy nastiňuje problematiku takzvaných odvozených projekcí, které rozšiřují praktickou využitelnost populačních odhadů. Poslední kapitola věnuje se metodice a hlavním výsledkům

současných projekcí významných pracovišť má následně za úkol seznámit stručně čtenáře s reálnými výsledky populační prognostiky ze současné praxe pro vytvoření lepší představy o aplikaci uvedené teorie.

Publikace je vhodná nejen pro čtenáře s obecným zájmem o demografické prognózování, ale také pro studenty předmětu 4DM302 Demografické projekce na VŠE, jehož prezentace k přednáškám v některých částech rozvíjí. Autoři při psaní publikace vycházeli ze svých dlouholetých zkušeností. Publikace v několika málo případných tématech obsahuje úryvky z kvalifikačních prací autorů, přičemž tyto práce je současně možné využít k dalšímu podrobnějšímu studiu některých témat. (Langhamrová, 2008; Hon, 2022)

Autoři

Základní terminologie demografických projekcí

Při sestavování populačních prognóz řešíme dva okruhy problémů. Jednak musíme formulovat co nejlépe hypotézu budoucího vývoje reprodukce obyvatelstva a jednak musíme vybrat způsob výpočtu, který použijeme – metodologii. Spolehlivost prognózy je závislá téměř výhradně na tom, jak se nám podaří správně vystihnout předpoklady o budoucím vývoji reprodukce obyvatelstva. Stanovení reálných předpokladů je velice obtížné. K tomu, abychom správně stanovili hypotézu o budoucím vývoji, musíme znát obecnější zákonitosti populačního vývoje, inspirovat se vývojem v zemích, které jsou z demografického pohledu vyspělejší a u nichž předpokládáme, že se sledovaná populace tomuto chování bude přibližovat.

Prognóza má řadu charakteristických rysů, kterými se odlišuje od ostatních obdobných typů předpovědí. Snahou prognózy je maximální přiblížení se budoucímu vývoji. Tím se prognóza liší od projekce. Projekce je chápána jako produkt určité činnosti a představuje nám výpověď o budoucím vývoji, která je založena na libovolně vymezených předpokladech bez nároku na jejich reálnost. (viz Kučera, 1998)

Tvorba populačních prognóz představuje komplexní problém. Těžištěm řešení je ve vytvoření si správných prognostických představ o perspektivním vývoji počtu obyvatel, pohlavní a věkové struktury. Pod pojmem prognózování se rozumí soubor činností, které vedou k vytvoření prognózy. Při tvorbě populačních prognóz si musíme uvědomit, že reálné systémy jsou ve své podstatě systémy velice složité. Proto je třeba odhlédnout od nepodstatných či méně podstatných vlastností a souvislostí a převést reálný systém na vhodný model. Proces prognózování tvoří podle Kučery (1998) sedm dílčích etap. Nejprve si musíme vymezit systém, v jehož rámci budeme vývoj objektu v souvislosti se zadáním prognózovat. Další etapou je pak popis a analýza vymezeného systému a cílem je získat a utřídit informace o dosavadním vývoji prognózovaného objektu. Následuje konstrukce projekčního modelu. Potom model aplikujeme. Nejprve prognózujeme hodnoty parametrů, na základě nich pak provedeme projekční výpočet. Získané výsledky dokumentujeme, následně prezentujeme způsobem vhodným pro uživatele. To, zda se nám podařilo správně odhadnout budoucí vývoj, lze vyhodnotit až porovnáním se skutečností.

Lze přehledně v bodech definovat dva základní způsoby prognózování obyvatelstva:

- Pomocí nějaké analytické funkce – prognózuje pouze celkový počet obyvatel, ne věkovou a pohlavní strukturu.
- Komponentní metodou – na základě předpokládaného vývoje (scénáře) plodnosti, úmrtnosti a migrace (tři komponenty demografického vývoje) prognózuje nejen vývoj celkového počtu obyvatel, ale i jeho složení podle věku a pohlaví.

Při prognózování pomocí analytické funkce jsou často užívané známé základní matematické funkce například exponenciální, logistická nebo lineární. Vzhledem k již řečenému triviálnímu výpočtu, kdy se prognózuje pouze celkový počet obyvatel, ne věková a pohlavní struktura, je potřeba při využití těchto jednoduchých postupů zachovat určitou opatrnost. Přesto je dodnes možné tyto metody využít například v případech odhadu budoucího počtu obyvatel v rámci:

- populace, kde nemáme údaje o jejím věkovém složení,
- populace, kde se nepředpokládá změna plodnosti a úmrtnosti a migrace je zanedbatelná (stabilní populace),
- prognózy na velmi dlouhé období.

Podstatně běžnější je dnes nicméně prognózování komponentní metodou, které je ve většině případů přesnější než prognózování pomocí nějaké funkce. Mezi hlavní důvody zvýšené přesnosti patří skutečnost, že tento model prognózuje nejen celkový počet, ale i věkové a pohlavní složení populace (z čehož lze navíc jako bonus odvodit řadu dalších charakteristik, např. počty žáků, počty pracovních sil, počty důchodců...).

Je dále také potřeba zmínit, že v ekonomicky vyspělých zemích se v dnešní době většinou tolik nemění počet obyvatelstva, jako jeho věkové složení, prognóza pouhého počtu obyvatel by nebyla příliš užitečná.

Jak bude patrné v dalších kapitolách, je také výhodou, že vlastní výpočet komponentní metodou není technicky příliš složitý (je možno jej provést například i v Excelu).

Technické zázemí vytvoření korektní prognózy dobře vypovídající o budoucím vývoji by mělo sestávat z několika kroků, konkrétně z:

- identifikace problému,
- popisu obyvatelstva či populace,
- konstrukce modelu,
- vytvoření scénáře prognózy,
- provedení, publikace a dokumentace,
- monitorování.

Identifikace problému spočívá zejména ve vymezení obyvatelstva, které prognózuujeme (bydlící na určitém území trvale, dlouhodobě...), případně je potřeba rozhodnout, zda je vhodné vyloučit část obyvatelstva. Dále je třeba identifikovat cíl (za jakým účelem prognózu děláme). Zde narazíme na otázky, zda je vhodné prognózovat veškeré obyvatelstvo, nebo pouze obyvatelstvo určitého věku (např. pro prognózu žáků škol – „stačí“ osoby do 50 let (děti a potenciální rodiče), tato prognóza není ovlivněna úmrtností ve vyšším věku; naopak prognóza seniorů – osob nad 65 let – není po dobu 65 let ovlivněna plodností atd.). Identifikovat je dále například potřeba také podrobnost věkové struktury obyvatelstva ve stupních letech, která může být uvedena v jednoletých nebo pětiletých věkových intervalech.

Konstrukce modelu a další související kroky technického zázemí budou následně podrobně probrány v kontextu příslušných vzorců v dalších kapitolách.

Scénář představuje odhad parametrů modelu. V případě komponentní metody je potřeba odhadnout specifické úmrtnosti a plodnosti, případně i velikost a věkovou a pohlavní strukturu migrace. Tyto údaje jsou nezbytné pro každý krok prognózy, tedy pro každý rok. Pokud jde o projekci zaměřenou pouze na celkový počet obyvatel, je třeba zvolit vhodnou funkci, která popisuje budoucí vývoj počtu obyvatel. Scénář obsahuje tzv. věšteckou stránku prognózy, což znamená, že scénář je vlastně sám o sobě prognózou. Samotný výpočet prognózy pak zahrnuje projekci obyvatelstva jako „mechanickou“ aplikaci zvoleného modelu a scénáře.

Teoreticky by měly být publikovány všechny prognózy, avšak v praxi tomu tak často není. Například u prognóz vytvářených na zakázku závisí jejich publikace na rozhodnutí objednatele. Při publikaci je otázkou, zda zveřejnit pouze výsledky, nebo také celý postup včetně předpokladů. Ideální by bylo zveřejnit kompletní metodiku, ale autoři mohou mít obavy z prozrazení svého know-how. Navíc by podrobný popis mohl být příliš dlouhý a plný technických detailů. Pro interní potřeby a pro zájemce by však podrobný popis měl být vždy k dispozici.

Publikovat by se vždy měly výsledky v dostatečně zajímavém detailu a základní informace o scénáři, jakož i o vývoji souhrnných charakteristik úmrtnosti, plodnosti a migrace doplněné nezbytnými údaji o jejich struktuře. Rozsah publikovaných scénářů se však liší. Například Český statistický úřad (ČSÚ) zveřejňuje pouze souhrnné charakteristiky a základní údaje o struktuře plodnosti, úmrtnosti a migrace, zatímco Eurostat poskytuje podrobnější informace, včetně specifických měr plodnosti, specifických měr úmrtnosti a čisté migrace podle jednotek věku pro každý rok.

Porovnání prognózy se skutečností by mělo být pravidelnou činností autorů i zadavatelů prognóz, zejména těch, kteří prognózy vytvářejí opakovaně, jako například ČSÚ. Toto porovnání slouží jako poučení pro tvorbu dalších prognóz. Po vyhodnocení, proč došlo k případným odchylkám, by měla následovat korekce celé prognózy. V praxi se však často nic takového neděje. I když se při tvorbě nové prognózy zohledňuje úspěšnost té předchozí, neexistuje systém, který by dlouhodobě takové poznatky shromažďoval a nabízel návod, jak s nimi pracovat.

Demografické prognózy často nevycházejí a odchylky mohou být značné. Jedním z důvodů je, že je obtížné předvídat výrazné kulturní a technické změny, stejně jako jejich dopad na demografické chování, zejména na migraci. Důkladnější monitorování prognóz by mohlo výrazně zlepšit jejich přesnost. Klíčovou složkou monitorování by mělo být porovnání důvodů, proč byl určitý scénář zvolen, s realitou. Důvody volby konkrétních scénářů však bývají často pragmatické, nikoli teoretické. Například scénář neměnné plodnosti je sice jednoduchý, ale obtížně zdůvodnitelný. Problém tedy není jen v nedostatečném monitorování, ale také v samotné tvorbě scénáře, jeho zdůvodnění a dokumentaci.

Ze základních pojmů demografického prognózování je následně nezbytné znát rozdělení prognóz dle délky. Prognózy a projekce se rozlišují dle takzvaného horizontu, což je časový okamžik, do kdy prognózujeme počty obyvatel. Tento horizont může být vyjádřen buď absolutně (např. do konce roku 2100), nebo relativně (např. na 80 let).

Rozdělení prognóz podle délky se pak často definuje následujícím způsobem:

- krátkodobé prognózy (do 10 let) – předpoklad současných kulturních vzorců a současné technické úrovně,

- střednědobé prognózy (do 25–30 let) – měli bychom odhadnout eventuální společenské i technické změny a jejich důsledky na demografické chování obyvatelstva,
- dlouhodobé prognózy (více než 30 let) – klesá přesnost odhadu počtu narozených – počet potenciálních matek vychází již z (ne vždy přesného) odhadu počtu narozených.

Na tomto místě je vhodné uvést logicky pochopitelný fakt, že s rostoucí délkou prognózy zpravidla klesá její přesnost. Například v roce 2020 lze snadněji odhadnout, jaké bude chování populace v roce 2025 než v roce 2100.

Kromě délky prognózovaného období se demografické prognózy dále rozčleňují také na základě takzvaných variant. Jedná se o scénáře založené na různých předpokladech budoucího vývoje. Zpravidla se definují tři varianty:

- nízká (nízká plodnost, nízká délka života, nízký migrační přírůstek → nízký počet obyvatel),
- střední (střední plodnost, ... → střední počet obyvatel),
- vysoká (vysoká plodnost, ... → vysoký počet obyvatel).

Někdy se tyto varianty nazývají též pesimistická, realistická, optimistická.

Střední varianta je obvykle nejpravděpodobnější a případně slučitelná s prognózou. Nejedná se nicméně o striktní pravidlo, může být například i pět variant, případně krátkodobé prognózy nebo prognózy malých územních celků mohou mít pouze jednu variantu. V případě stochastického typu prognózy mohou být nízká varianta s vysokou variantou vlastně mezemi intervalů spolehlivosti a střední varianta bodovým odhadem.

Stěžejní při tvorbě modelu projekce je schopnost definovat a shromáždit potřebná data o obyvatelstvu a být tedy schopen pracovat s jejich zdroji. Současně je potřeba se orientovat v rozlišnosti datových zdrojů na základě velikosti území, pro ilustraci z hlediska regionů Česka je možné shromáždit údaje o věkové a pohlavní struktuře obyvatelstva Česka do úrovně obcí s rozšířenou působností (ORP) každoročně s dostupností na webových stránkách ČSÚ. Za menší územní celky do úrovně obcí jsou však již údaje pouze na vyžádání. Za části obcí lze dále získat údaje pouze ze Sčítání lidu, domů a bytů. Na malé územní úrovni je následně dostupnost i nižší s ohledem na problém ochrany osobních údajů.

V souvislosti s definováním vstupní věkové struktury je vhodné definovat po horizontu projekce také její práh (začátek). Prahem projekce se rozumí obvykle počátek, resp. konec roku (půlnoc 31. 12. / 1. 1.) může se nicméně jednat i o střed roku, 1. září (pro školské účely) či jiné datum. Závěrem k tomuto tématu je vhodné zmínit, že někdy je nutný přepočítání věkové a pohlavní struktury k požadovanému datu (např. při užití dat ze sčítání lidu).

Výpočet se provádí na základě výchozí pohlavní a věkové struktury (prahová struktura) po krocích. Krok odpovídá posunu v čase. Délka kroku zpravidla odpovídá délce věkových intervalů (jednoleté intervaly – krok 1 rok, pětileté – krok 5 let). Pak platí, že při každém kroku se všichni (kteří přežijí) přesunou do další věkové skupiny. Je tedy vhodné zapamatovat si základní tři pojmy prahu, kroku a horizontu projekce.

Kromě terminologie věnující se demografické prognostice je třeba znát také základní demografické ukazatele a matematické symboly sloužící k popisu obyvatelstva, v této publikaci uvedme alespoň základní metriky (podrobněji např. Langhamrová, Šimpach, 2013). Pro dobrou prognózu je totiž nutná znalost nejen aktuálního stavu, ale i trendů dosavadního vývoje v poslední době. Analýzu je třeba provádět v kontextu ekonomických, společenských a dalších vlivů na demografický vývoj. V demografii konkrétně ve stručnosti rozlišujeme:

- syntetické charakteristiky (střední délka života podle pohlaví a úhrnná plodnost),
- charakteristiky struktury úmrtnosti a plodnosti (specifické míry úmrtnosti a plodnosti).

Dále je možné sledovat např. střední délku života 60-, 70- a 80letých, normální délku života, průměrný, případně modální věk matky při porodu aj. V případě analýzy dosavadního vývoje plodnosti a úmrtnosti Česka platí, že důležité jsou primárně údaje až po roce 1990 a trendy v demografickém vývoji v kontextu změn po změně politického zřízení. Vývoj je nicméně i v posledních desetiletích zejména z hlediska plodnosti poznamenán řadou výkyvů, v ideálním případě je tak vhodné analyzovat například i generační pohled. Například pokles a následný nárůst plodnosti v Česku na přelomu tisíciletí pomáhá pochopit právě kohortní analýza vystihující změny v časování plodnosti jednotlivých generací.

Analýza dosavadního vývoje migrace je následně ještě složitější než plodnost a úmrtnost, migrace obvykle nevykazuje žádné jasné tendence. Například od roku 1990 zde bylo podstatně více výkyvů než u výše řečených demografických procesů, což je způsobeno tím, že migrace je výrazně ovlivněna geopolitickou momentální situací, případně situací na trhu práce a bydlení, které se často mění. Pro poslední roky je pak typický velmi výrazný vliv uprchlických vln (v případě Česka měla bezprecedentní vliv uprchlická vlna z Ukrajiny). Ze zkušenosti vyplývá, že pro základní představu o rozsahu migrace stačí relativně krátká, zhruba 10letá časová řada.

Extrapoláční projekční metody

Před definováním komplexních modelů je vhodné podívat se na možnosti projektování a prognózování budoucího počtu obyvatel pomocí analytické funkce. Jedná se o jednoduchý předpoklad využití známých matematických funkcí v kontextu demografických projekcí. Konkrétně se kapitola více zaměří na funkci lineární, geometrickou, exponenciální a logistickou. Současně je zde ukázáno, jak na základě funkcí odvodit další analyticky zajímavé ukazatele, jako je například doba zdvojnásobení populace.

Nevýhodou takto jednoduchého odhadu je, že se týká pouze odhadu vývoje celkového počtu obyvatel, ne věkové a pohlavní struktury. Nejsou tedy využity všechny dostupné informace a výsledky mohou být častěji nereálné. Přesto jsou extrapoláční metody dodnes používány.

Možné užití extrapoláčních metod zahrnuje několik různých scénářů. V případě malé populace na malém územním celku se může rozlišení obyvatel podle věku a pohlaví ukázat jako problematické. Důvodem jsou příliš malé počty osob, což vede k značně nespolehlivým odhadům dalšího vývoje. Tyto odhady jsou navíc citlivé na náhodné výkyvy, zejména pokud dojde k výrazné migraci.

Naopak u velkých oblastí, jako jsou světadily nebo celý svět, nás často zajímá především počet obyvatel a hustota budoucího zalidnění. Specifickým případem je populace, kde nemáme k dispozici údaje o věkovém a pohlavním složení, nebo populace, kde se nepředpokládá změna plodnosti a úmrtnosti, a migrace je zanedbatelná. Taková populace je označována jako stabilní.

Extrapoláční přístup se dále většinou využívá pro střednědobé a dlouhodobé prognózy, protože u krátkodobých odhadů existuje poměrně velké riziko odchýlení od teoretického vývoje. U velmi dlouhodobých prognóz naopak převládají základní tendence, přičemž se snižuje vliv různých náhodných i nenáhodných faktorů. Dochází ke kompenzaci výkyvů plodnosti, úmrtnosti i migrace, což odráží „přirozený“ vývoj obyvatelstva na daném území.

Začněme patrně nejjednodušší lineární funkcí, kde se jedná pro představu vlastně o proložení konkrétních datových bodů přímkou (interpolace) a prodloužení této přímky do budoucna (extrapolace).

Při modelování vývoje počtu obyvatel lineární funkcí se počet obyvatel řídí rovnicí:

$$S(t + 1) = S(t) + R,$$

případně zobecněnou podobou:

$$S(t) = S(t_0) + R \cdot (t - t_0),$$

kde R – celkový roční přírůstek (úbytek) obyvatelstva – je konstantní.

Relativní roční přírůstek (úbytek) obyvatelstva lze definovat:

$$r_t = \frac{R}{S_t}.$$

Pro $R > 0$ (rostoucí populace) relativní přírůstek v čase klesá k 0.

Pro $R < 0$ (klesající populace) relativní úbytek v čase roste do ∞ .

Hodnota ročního přírůstku se nejčastěji odhaduje pomocí průměrného ročního (absolutního) přírůstku. Mezi obdobími 1 a 2 lze vypočítat jako:

$$R = \frac{S(t_2) - S(t_1)}{t_2 - t_1}.$$

Přesnější způsob odhadu vychází z myšlenky, že pokud známe počet obyvatel ve více časových okamžicích a předpokládáme náhodné odchyly od lineárního trendu, je korektnější odhadnout průměrný roční přírůstek jako parametr směrnice příslušné regresní přímky počtů obyvatel.

Analytické funkce následně umožňují poměrně jednoduché odvození dalších charakteristik, jako je doba D , za kterou se velikost (rostoucí) populace zdvojnásobí:

$$\begin{aligned} S(t + D) &= S(t) + R \cdot D = 2 \cdot S(t), \\ R \cdot D &= S(t), \\ D &= \frac{S(t)}{R}. \end{aligned}$$

Doba zdvojnásobení závisí na velikosti populace. Pro $R > 0$ (rostoucí populace) v čase roste do ∞ .

Ekvivalentně lze dopočítat i dobu D , za kterou se velikost (klesající) populace zmenší na polovinu:

$$\begin{aligned} S(t + D) &= S(t) + R \cdot D = \frac{S(t)}{2}, \\ R \cdot D &= -\frac{S(t)}{2}, \\ D &= \frac{-S(t)}{2 \cdot R} = \frac{S(t)}{2 \cdot |R|}. \end{aligned}$$

Případně také dobu, za kterou klesající populace zcela vymře (pak záporné hodnoty):

$$\begin{aligned} S(t + D) &= S(t) + R \cdot D = 0, \\ R \cdot D &= -S(t), \\ D &= \frac{-S(t)}{R} = \frac{S(t)}{|R|}. \end{aligned}$$

Střední stav obyvatelstva v intervalu $(t_1; t_2)$ je roven aritmetickému průměru počátečního a koncového stavu:

$${}_{t_2-t_1}\bar{S}_{t_1} = S\left(t_1 + \frac{t_2-t_1}{2}\right) = S\left(\frac{t_1+t_2}{2}\right) = S(t_1) + R \cdot \frac{t_2-t_1}{2} = \frac{S(t_1)}{2} + \frac{S(t_1)+R \cdot (t_2-t_1)}{2} = \frac{S(t_1)+S(t_2)}{2},$$

a doba expozice v tomto intervalu je pak rovna střednímu stavu vynásobenému délkou intervalu.

Případně lze doplnit definici, že doba expozice v intervalu $(t_1; t_2)$ je matematicky definována jako:

$$\begin{aligned}
 {}_{t_2-t_1}E_{t_1} &= \int_{t_1}^{t_2} S(t) dt = \int_{t_1}^{t_2} (S(t_1) + R \cdot (t - t_1)) dt = \int_{t_1}^{t_2} (S(t_1) - R \cdot t_1 + R \cdot t) dt = \\
 &= (S(t_1) - R \cdot t_1) \cdot (t_2 - t_1) + R \cdot \int_{t_1}^{t_2} t dt = (S(t_1) - R \cdot t_1) \cdot (t_2 - t_1) + R \cdot \left[\frac{t^2}{2} \right]_{t_1}^{t_2} = \\
 &= (S(t_1) - R \cdot t_1) \cdot (t_2 - t_1) + R \cdot \frac{t_2^2 - t_1^2}{2} = \\
 &= (S(t_1) - R \cdot t_1) \cdot (t_2 - t_1) + R \cdot \frac{(t_2 + t_1) \cdot (t_2 - t_1)}{2} = \\
 &= (S(t_1) - R \cdot t_1 + R \cdot \frac{(t_2 + t_1)}{2}) \cdot (t_2 - t_1) = (S(t_1) + R \cdot \frac{(t_2 - t_1)}{2}) \cdot (t_2 - t_1) = \\
 &= S\left(\frac{t_1 + t_2}{2}\right) \cdot (t_2 - t_1) = \frac{S(t_1) + S(t_2)}{2} \cdot (t_2 - t_1)
 \end{aligned}$$

Výhody lineárního přístupu v modelování počtu obyvatel spočívají v tom, že pracuje pouze se dvěma parametry – nějakým výchozím počtem obyvatel a ročním nárůstem nebo poklesem. Tento model umožňuje snadný odhad, například pomocí lineární regrese nebo metody nejmenších čtverců.

Nevýhody zahrnují skutečnost, že neregulovaná reprodukce (při neměnné úmrtnosti a plodnosti) má spíše exponenciální než lineární charakter. Tento přístup je také diskutabilní pro dlouhodobé prognózy, protože v případě rostoucí populace model předpokládá neomezený nárůst, zatímco u klesající populace hrozí poměrně rychlý pokles, což může vést k nereálným záporným hodnotám. Lineární vývoj počtu obyvatel tedy nemusí často odpovídat skutečnosti.

Příklady použití zahrnují prognózu krátkodobého vývoje, kdy se exponenciální, logistická a jiné funkce v krátkém časovém horizontu příliš neliší od lineární. Tento přístup je rovněž využíván při regulovaném demografickém vývoji, například při stanovení ročních migračních kvót, případně s přihlédnutím k velikosti přirozeného přírůstu či úbytku. V minulosti byla regulace demografického vývoje zajišťována například regulací porodnosti prostřednictvím regulace sňatků. Na malém územním celku se může použít lineární model v případě lineárního nárůstu počtu bytů nebo při plánovaném a regulovaném vylidňování určitého území.

Kromě lineárního vývoje lze samozřejmě využít i jiné analytické funkce, podívejme se na příklad exponenciální funkce. Zde se počet obyvatel řídí rovnicí:

$$\begin{aligned}
 S(t) &= S(t_0) \cdot e^{r(t-t_0)}, \\
 S(t+1) &= S(t) \cdot e^r,
 \end{aligned}$$

kde r je tzv. vnitřní míra růstu; relativní roční přírůstek (úbytek) obyvatelstva je konstantní ($e^r - 1$). Roční kvocient změny počtu obyvatel je $q = e^r$, pak můžeme psát též:

$$\begin{aligned}
 S(t) &= S(t_0) \cdot q^{(t-t_0)}, \\
 S(t+1) &= S(t) \cdot q.
 \end{aligned}$$

Jedná se vlastně o geometrickou řadu, někdy se proto setkáme s pojmem geometrický růst, je to totéž, co exponenciální růst, pouze zapsáno jiným způsobem.

Pro $r > 0$ je $q > 1$, populace roste, pro $r < 0$ je $q < 1$, populace klesá, ale počet osob je stále kladný, konverguje k 0. Průměrnou vnitřní míru růstu lze vypočítat jako:

$$r = \frac{\ln(S(t_2)) - \ln(S(t_1))}{t_2 - t_1}.$$

Pokud známe počet obyvatel ve více časových okamžicích a předpokládáme náhodné odchylky od exponenciálního trendu, je korektnější odhadnout průměrnou vnitřní míru růstu jako parametr směrnice příslušné regresní přímky logaritmů počtů obyvatel. Jedná se tedy vlastně o složitější ekvivalent se stejnou logikou jako u lineární funkce.

Případně lze v modelu geometrického růstu průměrný roční kvocient změny počtu obyvatel vypočítat podle vzorce:

$$q = \sqrt[t_2 - t_1]{\frac{S(t_2)}{S(t_1)}},$$

přitom platí

$$q = \left(\frac{S(t_2)}{S(t_1)}\right)^{\frac{1}{t_2 - t_1}} = e^{\frac{1}{t_2 - t_1} \ln\left(\frac{S(t_2)}{S(t_1)}\right)} = e^{\frac{\ln(S(t_2)) - \ln(S(t_1))}{t_2 - t_1}} = e^r,$$

jedná se tedy jen o jiný způsob zápisu exponenciálního růstu.

Exponenciální vývoj počtu obyvatel s využitím definovaného přírůstku lze již následně vypočítat poměrně jednoduše, roční absolutní přírůstek je dopočten jako:

$$R_t = S(t + 1) - S(t) = S(t) \cdot e^r - S(t) = S(t) \cdot (e^r - 1).$$

V rostoucí populaci ($r > 0$) roční absolutní přírůstky exponenciálně rostou.

V klesající populaci ($r < 0$) roční úbytky exponenciálně klesají.

Analogicky jako v případě lineární funkce lze i u této analytické funkce jednoduše definovat dobu, za kterou se populace zdvojnásobí, resp. zmenší na polovinu. Uvedme stručně vzorcové vyjádření.

Dobu D , za kterou se velikost (rostoucí) populace zdvojnásobí, lze definovat:

$$\begin{aligned} S(t + D) &= S(t) \cdot e^{rD} = 2 \cdot S(t), \\ r \cdot D &= \ln 2, \\ D &= \frac{\ln 2}{r}, \\ N &= \frac{\ln n}{r}. \end{aligned}$$

Doba zdvojnásobení zde nezávisí na velikosti populace, je konstantní.

Doba D , za kterou se velikost (klesající) populace zmenší na polovinu:

$$\begin{aligned} S(t + D) &= S(t) \cdot e^{rD} = \frac{S(t)}{2}, \\ r \cdot D &= \ln\left(\frac{1}{2}\right) = -\ln 2, \\ D &= \frac{-\ln 2}{r} = \frac{\ln 2}{|r|}. \end{aligned}$$

Pro samotný exponenciální vývoj počtu obyvatel platí, že střední stav obyvatelstva v intervalu $(t_1; t_2)$ je roven geometrickému průměru počátečního a koncového stavu:

$$\begin{aligned} t_2 - t_1 \bar{S}_{t_1} &= S\left(t_1 + \frac{t_2 - t_1}{2}\right) = S(t_1) \cdot e^{r \cdot \frac{t_2 - t_1}{2}} = (S(t_1))^{1/2} \cdot (S(t_1))^{1/2} \cdot e^{r \cdot \frac{t_2 - t_1}{2}} = \\ &= (S(t_1))^{1/2} \cdot (S(t_2))^{1/2} = \sqrt{S(t_1) \cdot S(t_2)}, \end{aligned}$$

při výpočtu doby expozice se ale hodnota středního stavu nepoužívá.

Doba expozice v intervalu $(t_1; t_2)$ je totiž definována jako:

$$\begin{aligned} t_2 - t_1 E_{t_1} &= \int_{t_1}^{t_2} S(t) dt = \int_{t_1}^{t_2} (S(t_1) \cdot e^{r \cdot (t-t_1)}) dt = S(t_1) \cdot \int_{t_1}^{t_2} \frac{e^{r \cdot t}}{e^{r \cdot t_1}} dt = \frac{S(t_1)}{e^{r \cdot t_1}} \\ &\quad \cdot \left[\frac{1}{r} \cdot e^{r \cdot t} \right]_{t_1}^{t_2} = \frac{S(t_1)}{r \cdot e^{r \cdot t_1}} \cdot (e^{r \cdot t_2} - e^{r \cdot t_1}) = \\ &= \frac{S(t_1)}{r} \cdot \frac{e^{r \cdot t_2} - e^{r \cdot t_1}}{e^{r \cdot t_1}} = \frac{S(t_1)}{r} \cdot (e^{r \cdot (t_2-t_1)} - 1) = \frac{S(t_1)}{r} \cdot (e^{r \cdot (t_2-t_1)} - 1) = \frac{1}{r} \cdot (S(t_2) - S(t_1)), \end{aligned}$$

přitom $S(t_2) = S(t_1) \cdot e^{r \cdot (t_2-t_1)}$, tedy $\frac{S(t_2)}{S(t_1)} = e^{r \cdot (t_2-t_1)}$,

$\ln(S(t_2)) - \ln(S(t_1)) = r \cdot (t_2 - t_1)$, takže $r = \frac{\ln(S(t_2)) - \ln(S(t_1))}{t_2 - t_1}$,

doba expozice $t_2 - t_1 E_{t_1} = \frac{S(t_2) - S(t_1)}{\ln S(t_2) - \ln S(t_1)} \cdot (t_2 - t_1)$.

Výhody exponenciální extrapolace spočívají v tom, že pracuje pouze se dvěma parametry – nějakým výchozím počtem obyvatel a ročním relativním nárůstem nebo poklesem počtu obyvatel. Snadný odhad je možný díky tomu, že po zlogaritmování dostáváme lineární funkci, což umožňuje využití metod, jako je lineární regrese nebo metoda nejmenších čtverců. Tento model také lépe odpovídá realitě reprodukce při neměnné úmrtnosti a plodnosti.

Nevýhody zahrnují skutečnost, že v poslední době, v důsledku demografické revoluce a obav z přelidnění, dochází ke snižování plodnosti a úmrtnosti, a tím i tempa růstu v řadě populací. Tento model je také diskutabilní pro dlouhodobé prognózy. U rostoucí populace totiž předpokládá velmi rychlý růst, což vede k nerealistickým počtům osob, protože při zachování vysoké plodnosti by dříve či později došlo k nárůstu úmrtnosti. Exponenciální vývoj počtu obyvatel je proto problematický zejména pro predikce na delší období.

Příklady použití zahrnují prognózu krátkodobého či střednědobého vývoje, zejména tam, kde dochází k malým změnám v plodnosti, úmrtnosti a migraci. Dlouhodobou prognózu pak tento model využívá spíše jako varovnou prognózu, například pro zobrazení, jak by se vyvíjel počet obyvatel v Africe, kdyby plodnost žen i úmrtnost zůstaly několik desetiletí na současné úrovni.

Závěrem kapitoly lze alespoň stručně definovat již poměrně složitější logistickou funkci, kde je možné vyjádřit počet obyvatel jako:

$$S(t) = \frac{A}{1 + e^{-c \cdot (t-t_0)}}$$

kde A je limitní velikost populace, c rychlost růstu populace a t_0 inflexní bod (okamžik nejvyššího absolutního přírůstku).

Pro logistickou funkci samozřejmě platí, že roční absolutní přírůstek od počátku do dosažení inflexního bodu roste, pak klesá a relativní roční přírůstek po celou dobu klesá.

Výhody tohoto modelu zahrnují skutečnost, že se jedná o asi nejrealističtější model reprodukce v dlouhodobém časovém horizontu, protože žádná populace nemůže růst do-

nekonečna. Analogicky se například vyvíjí populace rychle se množících banánových mušek v uzavřeném prostoru – pomalý absolutní růst postupně přechází ve stále rychlejší, ale toto zrychlování se v určitém okamžiku zastaví, přírůstky začnou postupně klesat a velikost populace se stabilizuje. Podobně světová populace, která po bouřlivém růstu v 19. století někdy v 90. letech překročila inflexní bod, začala zpomalovat svůj růst.

Nevýhody spočívají v poněkud složitějším odhadu parametrů. Počáteční odhad je třeba provést na základě empirických dat a následně jej zpřesnit pomocí nějaké numerické metody. Podrobnější popis logistické funkce a možných souvisejících numerických metod je již poměrně obsáhlý a nad rámec této publikace shrnující základy demografické prognostiky.

Kohortně komponentní metoda bez migrace

Tato kapitola se věnuje modelu kohortně komponentní metody, který je nejčastějším prostředkem při tvorbě populačních projekcí a prognóz, jak již bylo stručně zmíněno výše.

Metoda se může použít buď bez migrace, nebo s migrací. K prognózování budoucího vývoje bez migrace se přistupuje z již řečeného důvodu velké nestálosti vývoje migračních indikátorů v důsledku silné podmíněnosti migrace sociokulturními faktory. Tato kapitola se věnuje modelu bez migrace, ve kterém se uvažuje pouze o vlivu úmrtnosti a plodnosti, a je tedy z hlediska teoretického ukotvení jednodušší na pochopení. Variantě s migrací bude věnována následující kapitola po seznámení čtenáře s jednodušší variantou.

Popišme podrobněji komponentní metodu s jednoletým krokem, což umožňuje zjednodušený zápis vzorců. Výpočet projekce níže uvedeným postupem se provádí pro každé pohlaví zvlášť.

Pro účely úvodního vysvětlení kohortně komponentní metody jsou vzorce odvozeny pro obecný scénář, který předpokládá změny specifických měr úmrtnosti i plodnosti během období projekce. Způsoby modelování budoucích měr plodnosti a úmrtnosti (tedy hledání potenciálně lepší cesty, než nabízí triviální předpoklad konstantního vývoje), budou popsány v dalších kapitolách.

Známe-li věkové složení populace na počátku roku t , je logické, že na počátku dalšího roku budou všechny osoby, které se tohoto okamžiku dožijí, o rok starší. Odhad věkového složení populace na začátku roku $t + 1$ se provádí tedy podle vzorce:

$$S_{t+1,x+1} = S_{t,x} \cdot P_{t,x}, \text{ pro } x = 0, 1, \omega-2,$$

kde $S_{t,x}$ je počet žijících osob na počátku roku t v dokončeném věku x , $P_{t,x}$ je tzv. projekční koeficient v roce t pro věk x (udává, jaký podíl z osob ve věku x na počátku roku t se dožije začátku roku $t + 1$).

Tento projekční koeficient vypočteme jako poměr počtu žijících z úmrtnostních tabulek (model úmrtnostních tabulek je poměrně komplikovaný a není součástí této publikace, podrobněji viz například Langhamrová, Šimpach, 2013) pro rok t :

$$P_{t,x} = \frac{L_{t,x+1}}{L_{t,x}}.$$

Výše uvedeným postupem pochopitelně nedostaneme počet 0letých na počátku roku $t + 1$, neboť tyto osoby nebyly na počátku roku t ještě naživu, narodily se až během roku t . Počet živě narozených v roce t (chlapců i děvčat dohromady) odhadneme na základě specifických měr plodnosti a středních stavů žen v roce t :

$$N_t = \sum_{x=\alpha}^{\beta-1} \frac{S^{(2)}_{t,x} + S^{(2)}_{t+1,x}}{2} \cdot f_{t,x}.$$

Evidentně je k odhadu počtu živě narozených v roce t nezbytné znát počet potenciálních matek v tomto roce, používáme hodnoty počtu žen v reprodukčním věku z projekce žijících popsané výše.

Rozdělení počtu živě narozených na chlapce a děvčata se provádí v poměru 0,515 : 0,485, tedy:

$$N_t = 0,515 \cdot N_t^{(c)} \text{ pro chlapce, resp. } N_t = 0,485 \cdot N_t^{(c)} \text{ pro děvčata.}$$

Počet 0letých na počátku roku $t + 1$ je pochopitelně menší než počet živě narozených v roce t , neboť ne všechny narozené děti se počátku roku $t + 1$ dožijí. Odhad počtu 0letých na začátku roku $t + 1$ dostaneme podle vzorce:

$$S_{t+1,0} = N_t \cdot P_{t,*},$$

kde $P_{t,*}$ je projekční koeficient v roce t pro novorozence (udává, jaký podíl z dětí živě narozených v roce t se dožije začátku roku $t + 1$). Vypočteme jej opět z úmrtnostních tabulek pro příslušné pohlaví a rok t :

$$P_{t,*} = \frac{l_{t,0}}{l_0}.$$

Projekci na další roky lze následně snadno získat opakováním postupu krok po kroku až k horizontu projekce. Věkové a pohlavní složení na začátku roku $t + 2$ vypočteme na základě (projektované) věkové a pohlavní struktury k počátku roku $t + 1$ a tak dále. Teoreticky můžeme projekci počítat na libovolný počet let, je třeba nicméně opět připomenout, že s rostoucí délkou prognózy logicky klesá v krocích příliš vzdálených od prahu projekce přesnost populačního odhadu.

V případě projekce obyvatelstva podle širších věkových skupin je postup obdobný. Zpravidla se s ohledem na dostupná data v oficiálních zdrojích (5leté věkové skupiny) často volí 5letý krok (během 5 let se všichni, kteří přežijí, přesunou do následující věkové skupiny). Projekční koeficienty udávají pravděpodobnost přežití 5 let. Projekce 0–4letých se počítá na základě projekce živě narozených během 5letého období a projekčního koeficientu udávajícího pravděpodobnost, že tyto děti přežijí do konce 5letého období. Analogickým způsobem by se počítala projekce pro věkové intervaly jiné délky a odpovídající krok (např. 10 let).

Krok vlastně odpovídá posunu v čase, je to interval mezi dvěma okamžiky, pro které počítáme projekci věkové struktury. Délka kroku zpravidla odpovídá délce věkových intervalů (1leté intervaly – krok 1 rok, 5leté intervaly – krok 5 let). Potom platí, že během období jednoho kroku se všechny osoby, které toto období přežijí, přesunou do další věkové skupiny, jejich věk je tedy vždy jednoznačně určen.

Rozšířme úvahu o skutečnost, že na reálných datech je často problém s projekcí poslední věkové skupiny, kdy charakteristiky úmrtnosti i věkové složení osob bývají často zakončeny otevřeným věkovým intervalem.

Věkové složení podle jednotek věku bývá zpravidla zakončeno otevřeným intervalem $x_{max} +$ (osoby v dokončeném věku x_{max} a vyšším – není to 1letý, ale širší interval). Podobně je tomu často u úplných úmrtnostních tabulek. Například úmrtnostní tabulky ČSÚ končí ve věku 105 let, přitom střední délka života v tomto věku je vyšší než 1. Předpokládá se tedy, že se lidé mohou dožít 106 i více let, ale takovýto věk v tabulkách již není. Projekce

poslední věkové skupiny lze v ideálním případě odhadnout jednoduše, pokud je nejvyšší věk x_{max} tak vysoký, že můžeme pro jednoduchost předpokládat, že lidé v tomto věku již další rok nepřežijí (a nedopustíme se tím příliš velké chyby). To nastane speciálně v případě, kdy $x_{max} = \omega - 1$ (pak to plyne přímo z definice věku ω jako věku, kterého se již nikdo nedožije). Můžeme následně doložit:

$$P_{t,x_{max}} = 0,$$

počítáme tedy projekci obyvatelstva s vyloučením osob $(x_{max} + 1)$ letých a starších.

Pokud nelze předpokládat, že se nikdo nedožije věku $x_{max} + 1$ (např. pokud $x_{max} = 85$, fakticky i pokud $x_{max} = 100$, záleží na účelu projekce), existují další možnosti výpočtu projekce:

- Zvýšení maximálního věku na hodnotu blízkou nebo rovnou $\omega - 1$.
- Prodloužení úmrtnostních tabulek (přijmeme nějaký předpoklad o dalším vývoji měř úmrtnosti či pravděpodobnosti úmrtí a dopočítáme ostatní biometrické míry).
- Rozdělení nejstarších osob $(x_{max} + 1)$ letých do 1letých intervalů ve stejném poměru jako počty žijících v úmrtnostních tabulkách:

$$S_{t,x} = S_{t,x_{max}} \cdot \frac{L_{t,x}}{T_{t,x_{max}}}.$$

Pokud nelze předpokládat, že osoby v nejstarší věkové skupině nepřežijí další rok, je důležité se dobře orientovat v posunech mezi jednotlivými kroky projekce osob v nejvyšším věku. Osoby, které byly (na počátku roku t) x_{max} leté a starší, budou na počátku roku $t + 1$ (pokud budou naživu) x_{max} leté a starší. Na počátku roku $t + 1$ budou tedy ve skupině $(x_{max} + 1)$ letých osoby, kterým bylo na počátku roku t $x_{max} - 1$ let a přežily rok t , a dále osoby, kterým bylo na počátku roku t $x_{max} + 1$ let a přežily rok t .

Projekce nejstarší věkové skupiny se tedy vypočte podle vzorce:

$$S_{t+1,x_{max}+} = S_{t,x_{max}-1} \cdot P_{t,x_{max}-1} + S_{t,x_{max}+} \cdot P_{t,x_{max}+}.$$

Pro projekční koeficient v tomto případě platí:

$$P_{t,x_{max}+} = \frac{T_{t,x_{max}+}}{T_{t,x_{max}}},$$

(pokud je v úmrtnostních tabulkách věk $x_{max} + 1$).

Pokud úmrtnostní tabulky končí stejně jako výchozí věková struktura ve věku x_{max} , počítá se projekce poslední věkové skupiny podle vzorců:

$$S_{t+1,x_{max}+} = (S_{t,x_{max}-1} + S_{t,x_{max}+}) \cdot P_{t,x_{max}+},$$

$$P_{t,x_{max}+} = \frac{T_{t,x_{max}+}}{T_{t,x_{max}-1}}.$$

Je asi výrazně patrné, že vzorce pro výpočet projekce posledního otevřeného věkového intervalu jsou složitější než pro předchozí intervaly. Proto je jednodušší prodloužit úmrtnostní tabulky a ve věkovém složení rozdělit nejstarší osoby do jednoletých intervalů tak, aby nejvyšší věk byl blízký nebo rovný věku $\omega - 1$. Pak můžeme předpokládat, že žádná z osob v nejvyšší věkové skupině nepřežije další rok, výpočet se zjednoduší. Nedopustíme se velké chyby v projekci, protože se odhaduje úmrtnost a věkové složení pouze u nejstarších osob, kterých je málo a brzy zemřou.

Projekce s otevřeným posledním věkovým intervalem byla do publikace zahrnuta i pro ilustraci, jak se v závislosti na dostupnosti dat může poměrně logicky triviální výpočet základního modelu kohortně komponentní metody zkomplikovat.

Přidání migrace do kohortně komponentní metody

Existují dva přístupy k práci s migrací v rámci kohortně komponentní metody. Jedním z nich je práce pouze s čistou migrací (migrační saldo), což je jednodušší, ale méně realistické. Druhou možností je rozlišení imigrace a emigrace, které poskytuje podrobnější a přesnější obraz migračních procesů.

Pro zahrnutí migrace do prognóz existují dále ještě dva způsoby. Prvním je samostatná projekce imigrantů, kteří se na závěr přičtou k projekci populace bez migrace. Druhý, jednodušší přístup, spočívá v přičítání imigrantů (eventuálně odečítání emigrantů) v každém kroku projekce. Tento přístup však předpokládá, že imigranti mají stejnou plodnost, úmrtnost i tendenci emigrovat jako domácí populace, což je sice nerealistické, ale obvykle přijatelné. Imigrantů obvykle nebývá tolik, aby významně ovlivnili demografické charakteristiky, rozlišovat demografické chování imigrantů podle země původu by velmi komplikovalo výpočet a navíc by bylo často obtížné získat potřebná data. Kromě toho se demografické chování imigrantů často přizpůsobuje domácí populaci (tzv. demografická asimilace). Proto je tento druhý přístup častější.

Shrňme a rozveďme pro přehlednost informace ohledně řečených variant zahrnutí migrace v bodech. Dvojí pojetí migrace:

- Pouze čistá migrace (migrační saldo) zpravidla pro zemi, kde převažuje imigrace nad emigrací; pokud převažuje emigrace, má počet imigrantů záporné znaménko. Jednodušší, dostupnější data, méně realistické (neexistují čistí imigranti, pouze imigranti a emigranti).
- Rozlišení imigrace a emigrace. Realističtější, počet a pohlavní a věkové složení vystěhovalých může záviset na velikosti a složení prognózované populace výrazně více než počet a věkové složení přistěhovalých. Trendy imigrace a emigrace se mohou lišit (např. emigrace mladších, imigrace starších či naopak).

Zahrnutí přistěhovalých / vyloučení vystěhovalých do/z projektované populace v okamžiku přistěhování/vystěhování:

- Výpočetně mnohem jednodušší.

- U přistěhovalých nemusím počítat projekci narozených,
 - narození před přistěhováním – zahrnutí mezi přistěhovalé,
 - narození po přistěhování – zahrnutí mezi narozené v projektované populaci.
- U vystěhovalých nemusím brát v úvahu, zda přežijí do konce roku,
 - v době vystěhování byli pochopitelně naživu,
 - zemřeli po vystěhování již neovlivní velikost projektované populace.
- Předpoklad stejné plodnosti, úmrtnosti i tendenci emigrovat u imigrantů jako u domácí populace – nerealistické.
- Imigrantů obvykle není tolik ve srovnání s domácí populací.
- Často se demografické chování imigrantů přizpůsobuje domácí populaci (demografická asimilace). Mají (teoreticky) stejnou lékařskou péči, stejné sociální dávky, stejné pracovní podmínky.
- Proto je tento postup obvyklejší, budeme jej vždy používat.

Samostatná projekce imigrantů a emigrantů (teprve na závěr se eventuálně k projekci bez migrace přičtou imigranti):

- Teoreticky mnohem přesnější, mohu uvažovat u imigrantů jiné demografické chování než u domácí populace.
- Neexistuje „univerzální imigrant“.
- Demografické chování imigrantů může záviset a často závisí na tom, z jaké země přicházejí, jaké mají občanství, a na zemi původu.
- Tedy bychom museli prognózovat imigranty z každé země zvlášť, velmi komplikovaný scénář: nejen celkový počet imigrantů každý rok a rozdělení podle pohlaví a věku, ale i rozdělení imigrantů podle zemí, odkud přicházejí, a podle jejich občanství (zemi původu statistiky obvykle neuvádějí).
- Je zde otázka, jak prognózovat jejich plodnost, úmrtnost a migraci (např. žena přistěhovalá z Německa s britským občanstvím pocházející z Pákistánu).
- Demografické chování imigrantů se může měnit, zpravidla postupně přizpůsobovat demografickému chování domácí populace.
- Projekce s rozlišením imigrantů podle zemí by se prováděla spíše metodou vícetavové projekce, resp. prognózy.

Pro výpočty je nutné stanovit předpoklady nejen o celkovém počtu imigrantů, ale také o jejich pohlavní a věkové struktuře pro jednotlivé roky. Zvláštní specifikum zahrnutí migrace spočívá v práci s věkovým složením imigrantů. Pokud známe věk imigrantů v okamžiku přistěhování, není jednoznačně určen jejich věk na začátku následujícího roku. Ti, kteří v roce přistěhování slavili narozeniny ještě ve své domovské zemi, budou na začátku dalšího roku stejně staří jako při přistěhování. Ti, kteří oslavili narozeniny v roce migrace

teprve po přistěhování, budou mít na začátku dalšího roku věk o 1 vyšší než v okamžiku migrace.

Lepší alternativou je používat věkové složení imigrantů podle tzv. dosaženého věku (age reached) v daném roce. Tento věk je určen jako rozdíl mezi rokem imigrace a ročníkem narození. Tento způsob publikace dat o imigraci a emigraci využívá například aktuálně u většiny zemí i Eurostat, což umožňuje jednoznačné začlenění imigrantů do projektované populace.

Při zohlednění třídění imigrantů dle věku je třeba mít na paměti, že k výpočtu projekcí kohortně komponentní metodou je nezbytné stanovit nejen celkový počet imigrantů, ale také jejich věkové a pohlavní složení v jednotlivých letech. Věkové složení migrantů by mělo být dostupné ke stejnému okamžiku jako věkové složení prognózované populace, ideálně k počátku nebo konci roku. To umožňuje zařadit všechny migranty jednoznačně do příslušné věkové skupiny.

U věkového složení k počátku, resp. ke konci roku existuje jednoznačná korespondence mezi dokončeným věkem a ročníkem narození. Osoba v dokončeném věku x na počátku roku t (tj. na konci roku $t - 1$) se narodila v roce $t - x - 1$.

Pokud jsou imigranti v daném roce tříděni podle dokončeného věku v okamžiku přistěhování, není možné jednoznačně určit jejich ročník narození a dokončený věk na začátku následujícího roku $t + 1$. To závisí na okamžiku narození:

- Kdo slavil v roce přistěhování narozeniny ještě ve své domovské zemi, (tedy se přistěhoval po narozeninách) bude mít na začátku roku $t + 1$ stejný dokončený věk jako v okamžiku přistěhování.
- Kdo bude slavít v roce přistěhování své narozeniny až v zemi, kam se přistěhoval (tedy se přistěhoval před narozeninami), bude mít na začátku roku $t + 1$ dokončený věk o 1 vyšší než v okamžiku přistěhování.

Lepší je proto mít věkové složení imigrantů tříděné podle tzv. dosaženého věku v daném roce (age reached). Dosažený věk osoby v roce t je definován jako rozdíl daného roku t a ročníku narození dané osoby. Je to věk, kterého daná osoba dosáhne v daném roce, až bude mít narozeniny.

Zatímco dokončený věk se zvyšuje o 1 v den narozenin, dosažený věk se zvyšuje o 1 již 1. ledna (tedy dříve). Na začátku roku je dosažený věk všech osob o 1 vyšší než jejich dokončený věk, v den narozenin se dokončený věk osoby zvýší na hodnotu dosaženého věku, na konci roku je dokončený věk všech osob rovný dosaženému věku. Eurostat publikuje data o imigrantech a emigrantech většinou podle dosaženého věku.

Při využití počtů imigrantů v roce t tříděných podle dosaženého věku je zde výhoda, že na začátku roku $t + 1$ bude jejich dokončený věk stejný jako dosažený věk v roce imigrace.

Naopak pokud máme počty imigrantů v roce t tříděné podle dokončeného věku (v okamžiku imigrace), na začátku roku $t + 1$ bude jejich dokončený věk buď stejný (pokud se přistěhovali po narozeninách), nebo o 1 vyšší (pokud se přistěhovali před narozeninami). Často se v tomto případě stanoví předpoklad rovnoměrného rozdělení porodů a okamžiku

imigrace během roku (nerealistické, ale používá se), kde předpokládáme, že polovina imigrantů přišla před narozeninami a polovina po narozeninách. Potom polovina imigrantů v každé jednotce dokončeného věku bude mít na počátku roku $t + 1$ stejný dokončený věk a druhá polovina bude mít dokončený věk o 1 vyšší.

Pro zahrnutí imigrantů do populační projekce je nutné zohlednit pouze ty, kteří přežijí do konce roku t , kdy se přistěhovali. Je třeba zahrnout logiku shrnutou v několika následujících tezí:

- Jedná se o přežití kratší doby než celý rok, pouze období od okamžiku imigrace do konce kalendářního roku.
- Opět předpokládáme rovnoměrné rozdělení okamžiku imigrace během roku.
- Počty imigrantů v roce t tříděné podle dosaženého věku, předpokládáme, že do konce roku každý musí prožít v průměru $1/2$ roku.
- Počty imigrantů v roce t tříděné podle dokončeného věku (v okamžiku imigrace), průměrná doba prožitá do konce roku závisí na tom, zda se přistěhovali před narozeninami, nebo po narozeninách.
- Přistěhovalí před narozeninami musí přežít v průměru $2/3$ roku, přistěhovalí po narozeninách v průměru pouze $1/3$ roku.

V případě varianty s čistou migrací tříděnou podle dosaženého věku se projekce s migrací počítá podle následujících vzorců:

$$S_{t+1,x+1} = S_{t,x} \cdot P_{t,x} + I_{t,x+1} \cdot P_{t,x}^{1/2},$$

$$S_{t+1,0} = N_t \cdot P_{t,*} + I_{t,0} \cdot P_{t,*}^{1/2},$$

kde $I_{t,x}$ je migrační saldo osob v dosaženém věku x v roce t (tj. osob narozených v roce $t - x$).

Je třeba uvažovat, že na počátku roku t těmto migrantům bylo $x - 1$ let a od okamžiku přistěhování do konce roku „musí“ přežít v průměru $1/2$ roku. Na začátku roku $t + 1$ (tj. na konci roku t) bude dokončený věk všech imigrantů (pokud budou naživu) stejný jako jejich dosažený věk v roce t . K projekci počtu x -letých osob domácí populace na počátku roku t přičtu projekci imigrantů v roce t v dosaženém věku $x + 1$ (na počátku roku t jim bylo také x let). K projekci narozených v roce t přičtu projekci imigrantů v dosaženém věku 0, tedy imigrantů narozených v roce t .

Dosažený věk u imigrantů ve vzorci musí být o 1 vyšší než dokončený věk u domácí populace, jedná se totiž o stejný ročník narození. Věk v projekčním koeficientu je stejný (roven x) a pro prognózovanou populaci i imigranty odpovídá věku osob na počátku roku t . Žijící v dokončeném věku x na počátku roku t ($S_{t,x}$) měli x -té narozeniny v roce $t - 1$, tedy se narodili v roce $t - x - 1$. Migranti v roce t v dosaženém věku $x + 1$ měli/budou mít $(x + 1)$ -ní narozeniny v roce t , tedy se narodili rovněž v roce $t - x - 1$, na počátku roku t jim bylo také x let. Na počátku roku $t + 1$ (tj. na konci roku t) bude dokončený věk všech těchto osob (domácí populace i imigrantů) stejný, $x + 1$.

Pokud je čistá migrace tříděna podle dokončeného věku, vzorcové vyjádření se výrazně liší:

$$S_{t+1,x+1} = S_{t,x} \cdot P_{t,x} + \frac{I_{t,x} \cdot P_{t,x}^{2/3} + I_{t,x+1} \cdot P_{t,x}^{1/3}}{2},$$

$$S_{t+1,0} = N_t \cdot P_{t,*} + \frac{I_{t,0} \cdot P_{t,*}^{1/3}}{2},$$

kde $I_{t,x}$ je migrační saldo osob v dokončeném věku x v roce t (věk při imigraci). Používá se předpoklad rovnoměrného rozdělení porodů a migrace během roku. Předpokládáme, že polovina imigrantů v každé skupině dokončeného věku se přistěhovala před narozeninami, tedy do konce roku musí přežít v průměru $2/3$ roku a na konci roku jim bude $x + 1$ let. Polovina se přistěhovala po narozeninách, do konce roku musí přežít v průměru $1/3$ roku a na konci roku jim bude x let. K projekci počtu x -letých osob domácí populace na počátku roku t přičtu projekci poloviny imigrantů v roce t ve věku x a poloviny imigrantů ve věku $x + 1$. K projekci narozených v roce t přičtu projekci poloviny imigrantů ve věku 0.

Přejděme nyní ve výkladu na komponentní metodu zahrnující migraci s rozlišením imigrace a emigrace. Nově tedy uvažují zahrnutí imigrantů do projekce:

- U emigrantů se nebere v úvahu, zda přežijí do konce roku. V době emigrace byli všichni naživu, tedy všechny odečtu; eventuální úmrtí po emigraci už nemá vliv na velikost prognózované populace v dalším roce.
- Korektní postup výpočtu projekce s rozlišením imigrace a emigrace:
 - od složení domácí populace na začátku roku t a od projekce počtu narozených v roce t nejprve odečíst emigranty z roku t ,
 - teprve pak provést projekci přežití „zbyvajících“ osob na další rok s přičtením imigrantů.

Lehce diferentní je také logika týkající se určení věku emigrantů na začátku roku emigrace. Pokud jsou data o vystěhovaných tříděna podle jejich dokončeného věku v okamžiku vystěhování, nelze u osob vystěhovaných v roce t jednoznačně určit dokončený věk na začátku roku t . Kdo se v roce t vystěhoval před svými narozeninami, měl na začátku roku t stejný dokončený věk jako v okamžiku vystěhování.

Kdo se v roce t vystěhoval až po svých narozeninách, měl na začátku roku t dokončený věk o 1 rok nižší, než v okamžiku vystěhování. Lepší je věkové složení emigrantů podle dosaženého věku; dokončený věk emigranta na začátku roku, kdy emigroval, je vždy o 1 nižší než jeho dosažený věk v roce emigrace. Pokud uvažujeme počty emigrantů v roce t tříděné podle dosaženého věku, tam platí, že na začátku roku t byl jejich dokončený věk o 1 nižší než dosažený věk v roce imigrace. Naopak u počtů emigrantů v roce t tříděných podle dokončeného věku (v okamžiku emigrace), byl na začátku roku t jejich dokončený věk buď stejný (pokud emigrovali před narozeninami), nebo o 1 nižší (pokud imigrovali po narozeninách). Opět se zde používá předpoklad rovnoměrného rozdělení porodů a okamžiku emigrace během roku, polovina emigrantů odešla před narozeninami a polovina po narozeninách. Následně by polovina emigrantů v každé jednotce dokončeného věku měla na počátku roku t stejný dokončený věk a druhá polovina dokončený věk o 1 nižší.

Projekční vzorce projekce s migrací s rozlišením imigrace a emigrace lze při třídění migrantů podle dosaženého věku zapsat ve tvaru:

$$S_{t+1,x+1} = (S_{t,x} - E_{t,x+1}) \cdot P_{t,x} + I_{t,x+1} \cdot P_{t,x}^{1/2},$$

$$S_{t+1,0} = (N_t - E_{t,0}) \cdot P_{t,*} + I_{t,0} \cdot P_{t,*}^{1/2},$$

kde $I_{t,x}$ je počet imigrantů v dosaženém věku x v roce t a $E_{t,x}$ je počet emigrantů v dosaženém věku x v roce t .

Jedná se zde tedy o osoby narozené v roce $t - x$, kterým na počátku roku t bylo $x - 1$ let. Od počtu x -letých osob projektované populace na počátku roku t odečtu vystěhovalé v roce t v dosaženém věku $x + 1$, provedu projekci „zbývajících“ osob a přičtu projekci přistěhovalých v roce t v dosaženém věku $x + 1$. Pro druhý vzorec od počtu narozených v roce t odečtu vystěhovalé v dosaženém věku 0, provedu projekci „zbývajících“ osob, přičtu projekci přistěhovalých v dosaženém věku 0.

Při třídění podle dokončeného věku je potřeba upravit vzorcový zápis:

$$S_{t+1,x+1} = (S_{t,x} - \frac{E_{t,x} + E_{t,x+1}}{2}) \cdot P_{t,x} + \frac{I_{t,x} \cdot P_{t,x}^{2/3} + I_{t,x+1} \cdot P_{t,x}^{1/3}}{2},$$

$$S_{t+1,0} = (N_t - \frac{E_{t,0}}{2}) \cdot P_{t,*} + \frac{I_{t,0} \cdot P_{t,*}^{1/3}}{2},$$

kde $I_{t,x}$ je počet imigrantů v dokončeném věku x v roce t a $E_{t,x}$ je počet emigrantů v dokončeném věku x v roce t .

Zde se opět vychází z předpokladu rovnoměrného rozdělení narození a okamžiku migrace. Polovina migrovala před narozeninami a na začátku roku t jim bylo x let. Druhá polovina migrovala po narozeninách a na začátku roku t jim tedy bylo $x - 1$ let. Od počtu x -letých osob projektované populace na počátku roku t odečtu polovinu emigrantů v roce t v dokončeném věku x a polovinu emigrantů v dokončeném věku $x + 1$, provedu projekci „zbývajících“ osob, přičtu projekci poloviny imigrantů v roce t ve věku x a poloviny imigrantů ve věku $x + 1$. V případě druhého vzorce od počtu živě narozených v roce t odečtu polovinu emigrantů v dokončeném věku 0, provedu projekci „zbývajících“ osob a přičtu projekci poloviny imigrantů v dokončeném věku 0.

Závěrem kapitoly zmiňme, že je samozřejmě případně možné odhadem převádět věk dosažený a dokončený, ilustrujme na poměrně jednoduché metodě, kde:

$$I_{t,x(\text{dosaž.})} = \frac{I_{t,x(\text{dokonč.})} + I_{t,x-1(\text{dokonč.})}}{2},$$

$$I_{t,0(\text{dosaž.})} = \frac{I_{t,0(\text{dokonč.})}}{2},$$

kde $x(\text{dosaž.})$ je dosažený věk, $x(\text{dokonč.})$ dokončený věk.

Je samozřejmě nezbytné pro funkci těchto vzorců vycházet z předpokladu rovnoměrného rozdělení narozených i okamžiku migrace během roku. Polovina migruje před narozeninami a je tedy u nich dosažený věk o 1 vyšší než dokončený věk. Polovina migruje po narozeninách a u této skupiny je dosažený věk stejný jako dokončený věk. Vzorce pro emigranty by byly analogické.

Tvorba scénářů v kohortně komponentní metodě – deterministický přístup

Proces tvorby prognóz obyvatelstva zahrnuje několik klíčových fází, o kterých byla řeč výše, počínaje vytvořením scénáře prognózy. Tento scénář slouží jako odhad parametrů modelu, který je základem dalšího výpočtu. V rámci kohortně komponentní metody prognózy se zaměřujeme na odhad specifických úmrtností a plodností, případně na odhad velikosti a věkové a pohlavní struktury migrace. Tyto údaje jsou nezbytné pro každý krok prognózy, obvykle pro každý rok projekce. Zároveň je třeba mít na paměti, že scénář v sobě obsahuje „věštecký“ aspekt prognózy – představuje totiž samotnou podstatu odhadu budoucího vývoje.

Jakmile je scénář sestaven, přechází proces do fáze výpočtu prognózy. Tento krok, označovaný jako projekce obyvatelstva, spočívá v mechanické aplikaci zvoleného modelu a scénáře. Tím se získávají odhady budoucího počtu a struktury populace na základě předpokládaných trendů.

Pro detailní projekci je třeba sestavit podrobné scénáře zaměřené na tři klíčové oblasti – úmrtnost, plodnost a migraci. V každé z těchto oblastí existují specifické požadavky na data a metody odhadu.

Dodejme, že v případě triviálního modelu, kde by se ponechávaly charakteristiky úmrtnosti, plodnosti a migrace na úrovni posledních dostupných dat (nebo na úrovni průměru hodnot z několika posledních let), se vlastně také jednalo o scénář. Tento scénář měl specifickou hypotézu založenou na neměnnosti demografických procesů do budoucna. Jedná se pochopitelně o nejjednodušší předpoklad, který je vhodný při seznámení s modelem v rámci této publikace, nicméně v řadě případů je konstantní scénář využíván i v praxi v případě nestálosti demografických procesů nebo z důvodu komparace a podobně. Tyto scénáře se používají i v některých variantách projekcí OSN (zpravidla bez migrace), nazývají se „no change“ nebo „benchmark“. Cílem takovýchto projekcí je ukázat, nakolik je aktuální demografické chování udržitelné, tj. jaký by byl populační vývoj, kdyby se několik desítek let neměnila aktuální plodnost ani úmrtnost a nedocházelo k migraci.

Tato kapitola se věnuje takzvaným deterministickým scénářům. Tyto scénáře jsou předem jednoznačně určeny na základě určité funkce či parametru. Neurčitost v těchto modelech vzniká pouze díky různým scénářům použitých parametrů (návaznost na nízkou atd. variantu řešenou v předchozích kapitolách). Přestože jednotlivé varianty nebo scénáře nemají přiřazené konkrétní pravděpodobnosti, bývají obvykle vytvořeny na základě expertního, a tedy subjektivního odhadu. Přiřazování pravděpodobnosti je následně subjektem stochastického modelování, které bude probráno v další kapitole.

Úmrtnost se zpravidla modeluje odděleně pro muže a ženy, přičemž jako základní ukazatel bývá často uváděna střední délka života novorozence. Tento ukazatel však sám o sobě nestačí, protože charakterizuje pouze celkovou úroveň úmrtnosti, nikoliv její věkovou strukturu. Pro podrobný výpočet jsou nezbytné tabulkové počty žijících $L_{t,x}$ pro každý rok projekce. Ty lze odvodit z jiných charakteristik úmrtnosti, jako jsou specifické míry úmrtnosti $m_{t,x}$, tabulkové pravděpodobnosti úmrtí $q_{t,x}$ nebo tabulkové počty dožívajících $l_{t,x}$.

Pro dlouhodobější projekce se v souladu s aktuálními trendy vývoje často předpokládá pokračující pokles úmrtnosti, tj. růst střední délky života. Jedním z používaných modelů je exponenciální pokles pravděpodobností úmrtí. Tento model může být transformován do lineární podoby pomocí logaritmování, což umožňuje odhad parametrů lineární regrese metodou nejmenších čtverců na základě historických dat. Pro odhad budoucího trendu se obvykle používají data za několik posledních desetiletí, přičemž se mohou vyloučit roky s extrémní úmrtností způsobenou například epidemiemi nebo katastrofami.

Z hlediska matematického zápisu by bylo možné exponenciální pokles definovat vzorcem:

$$q_{t+1,x} = k_x \cdot q_{t,x} \text{ pro } 0 < k_x < 1,$$

případně u víceletého období jako:

$$q_{t,x} = k_x^{t-t_0} \cdot q_{t_0,x},$$

logaritmováním možno zápis převést na lineární závislost:

$$\ln q_{t,x} = (t - t_0) \cdot \ln k_x + \ln q_{t_0,x},$$

a po transformaci časové osy (položíme $t' = t - t_0$) dostáváme model lineární regrese. Odhady parametrů $\ln k_x$ a $\ln q_{0,x}$ pro každý věk x získáme metodou nejmenších čtverců na základě empirických hodnot pravděpodobností úmrtí v nějakém období považovaném za vhodné pro prognózu budoucího trendu vývoje.

Za předpokladu pokračování dosavadních trendů pro daný rok t' je odvození teoreticky jednoduché podle rovnice regresní přímky:

$$\ln q_{t',x} = t' \cdot \ln k_x + \ln q_{0,x}.$$

Jedním z hlavních problémů je otázka zachování tzv. koherence. V každém roce prognózy by měla být pravděpodobnost úmrtí závislá na věku monotónně. Typicky dochází ke snižování úmrtnosti v raném dětství, dosažení minima mezi 5. a 10. rokem věku a následnému růstu pravděpodobností úmrtí až do nejvyššího věku. Výjimkou může být tzv. „nehodový hrb“ – krátkodobý pokles úmrtnosti po 20. roce věku, zejména u mužů. Další podmínkou je, aby v každém věku byla pravděpodobnost úmrtí vyšší u mužů než u žen, s výjimkou nejvyšších věkových kategorií, kde se rozdíly mohou stírat.

K dosažení požadované monotonie výchozích pravděpodobností úmrtí lze využít logaritmickou transformaci, která zachovává monotónní závislosti. Následně se hodnoty $\ln q_{0,x}$ (namísto $q_{0,x}$) vyrovnávají, například pomocí klouzavých průměrů. Tento proces může být opakován několikrát, dokud se nedosáhne hladké a monotónní závislosti na věku. Současně se porovnávají hodnoty pro muže a ženy, aby byla zajištěna vyšší úmrtnost mužů. Po dosažení požadované monotonie lze přistoupit k výpočtu parametrů regresní

přímky, přičemž směrnice regresní přímky se odhaduje metodou lineární regrese bez absolutního členu. Matematicky:

$$\ln q_{t',x} - \ln q_{0,x} = t' \cdot \ln k_x.$$

S odhadem jediného $\ln k_x$ parametru:

$$\ln \hat{k}_x = \frac{\sum_r (\ln q_{r,x} - \ln q_{0,x}) \cdot r}{\sum_r r^2}.$$

Zejména při dlouhodobých projekcích se však může stát, že dojde k porušení monotonie pravděpodobností úmrtí ve vyšším věku, například pokud klesá pravděpodobnost úmrtí u 70letých rychleji než u 60letých, což by potenciálně vedlo k nerealistickým výsledkům, kdy by od určitého roku projekce mohla být pravděpodobnost úmrtí 70letých nižší než 60letých. Proto je nezbytné zajistit monotonii nejen na začátku, ale i v posledním roce prognózy, tím je zajištěna i v každém kroku.

K dosažení monotonie může pomoci vyrovnání hodnot parametrů $\ln k_x$, např. klouzavými průměry (často nutné mnohonásobně opakované vyrovnání) až do dosažení monotonie hodnot $\ln q_{t,x}$ v posledním roce projekce. Poté je třeba porovnání hodnot $\ln q_{t,x}$ v posledním roce pro muže a ženy, eventuální korekce hodnot $\ln k_x$, pro muže a pro ženy, aby úmrtnost mužů byla vyšší. Následovala by opakovaná kontrola monotonie hodnot $\ln q_{t,x}$ v posledním roce a tak dále.

Výchozí pravděpodobnosti úmrtí $q_{0,x}$ lze samozřejmě stanovit i jinými způsoby. Například jako průměr hodnot pravděpodobností úmrtí za několik posledních let. Následně je vhodné provést ekvivalentní vyrovnání klouzavými průměry.

Obecně zde je možné dosáhnout podstatně jednodušších výpočtů za cenu méně realistického modelu, pokud se nastaví předpoklad, že tempo poklesu úmrtnosti nezávisí na věku. Následně lze jednoduše zapsat:

$$q_{t+1,x} = k \cdot q_{t,x}, \text{ pro } 0 < k < 1.$$

Pokud je závislost na věku u výchozích pravděpodobností úmrtí monotónní, je monotonie zajištěna po celé období prognózy. Výchozí hodnotu indexu poklesu úmrtnosti k zvolíme zkusmo (např. 0,99) a později ji upravíme podle nějakého předpokladu o střední délce života při narození v posledním roce projekce (hodnota pro muže a pro ženy se může lišit).

Scénáře úmrtnosti zahrnují věkovou škálu až do maximálního věku $\omega - 1$, kde lze předpokládat, že nikdo se nedožije dalších narozenin. Pro vyspělé země se tento věk stanovuje obvykle na 110 až 120 let. Ve vysokých věkových kategoriích se pravděpodobnosti úmrtí modelují individuálně pro každý rok věku. Místo sestavování úmrtnostních tabulek pro každý rok se často využívají rekurentní výpočty na základě tabulkových počtů dožívajících. Stručně nastiňme definici výpočtů:

$$l_{t+1,0} = 100\,000,$$

$$l_{t+1,x+1} = l_{t+1,x} \cdot (1 - q_{t+1,x}) = l_{t+1,x} \cdot (1 - k_x \cdot q_{t,x}) = l_{t+1,x} \cdot (1 - k_x \cdot \frac{d_{t,x}}{l_{t,x}}) = l_{t+1,x} \cdot (1 - k_x \cdot \frac{l_{t,x} - l_{t,x+1}}{l_{t,x}}),$$

$$T_x = \sum_{u=x}^{110} L_u = \frac{l_u + l_{u+1}}{2} + \frac{l_{u+1} + l_{u+2}}{2} + \dots + \frac{l_{109} + l_{110}}{2} + \frac{l_{110} + l_{111}}{2} = \sum_{u=x}^{110} l_u - \frac{l_x}{2},$$

$$e_x^\circ = \frac{T_x}{l_x} = \frac{\sum_{u=x}^{110} l_u - \frac{l_x}{2}}{l_x} = \frac{\sum_{u=x}^{110} l_u}{l_x} - \frac{1}{2}.$$

Po provedení výpočtů je důležité ověřit, zda je výsledná prognóza realistická, – například porovnáním střední délky života při narození se scénáři jiných projekcí. Pokud jsou zjištěny odchylky, lze hodnoty koeficientů poklesu pomocí předpisu:

$$k_{x'} = K \cdot k_x$$

upravit. Například nižší hodnota $K < 1$ znamená rychlejší pokles úmrtnosti a vyšší délku života, zatímco vyšší hodnota $K > 1$ ukazuje na pomalejší pokles a nižší délku života. Najdeme takové K , aby střední délka života v posledním roce dosahovala požadované hodnoty.

Jednoduché modely mají své limity, zejména pokud nepředpokládají změnu tempa poklesu úmrtnosti v průběhu prognózovaného období. Pro pokročilejší analýzy lze použít sofistikovanější metody, například Lee-Carterův model, který umožňuje přesnější modelování trendů demografických procesů plodnosti a úmrtnosti a jejich závislosti na čase i věku. Tento model patří mezi stochastické způsoby modelování a bude více rozebrán v následující kapitole.

V případě plodnosti se scénář obvykle vytváří pouze pro ženy, bez ohledu na pohlaví dítěte. Základním ukazatelem bývá úhrnná plodnost, která však podobně jako střední délka života u úmrtnosti poskytuje jen celkový obraz a nezohledňuje věkovou strukturu plodnosti. Pro projekce počtu živě narozených jsou proto nezbytné specifické míry plodnosti žen $f_{i,x}$ pro každý rok projekce.

Scénář plodnosti představuje klíčový aspekt populačních prognóz a lze jej formulovat relativně jednoduše. Základem scénáře jsou specifické míry plodnosti pro jednotlivé roky projekce, které umožňují popsat jak úroveň, tak věkovou strukturu plodnosti v populaci. V oblastech s vysokou plodností je často realistické předpokládat její pokles, zatímco v regionech s nízkou plodností může být naopak možné očekávat růst. Výchozí hodnoty specifických měr plodnosti se obvykle určují na základě dat z posledního roku nebo z průměru několika posledních let. Tyto hodnoty pak slouží jako základ pro prognózu, v níž může být zohledněna jak změna celkové úrovně plodnosti, tak její věková struktura.

Změny v plodnosti mezi vybranými lety se často předpokládají lineární. Například pokud úhrnná plodnost v prahovém roce dosahuje určité hodnoty a má se do určitého roku zvýšit na danou hodnotu, jsou specifické míry plodnosti v mezidobí odvozeny lineární interpolací mezi oběma body. Obdobný přístup se uplatňuje při dalším prodloužení prognózy, kdy úhrnná plodnost pokračuje v růstu nebo se stabilizuje na jiné cílové hodnotě, a zároveň dochází k úpravám věkové struktury podle vzorové plodnosti. Takto definovaný jednoduchý scénář plodnosti zajišťuje plynulý vývoj bez výrazných výkyvů, což je pro účely projekcí žádoucí.

Při matematickém zápisu vlastně vyjdeme ze specifických měr plodnosti v prahovém roce t_1 a předpokládáme, že do roku t_2 se:

- úhrnná plodnost postupně zvýší na hodnotu $úf_{t_2}$,
- struktura plodnosti postupně přiblíží struktuře nějaké vzorové plodnosti se specifickými mírami $f_x^{(l)}$,
- v jednotlivých letech specifické míry plodnosti mění lineárně.

Specifické míry plodnosti v roce t_2 vypočteme podle vzorce:

$$f_{t_2,x} = f_x^{(V)} \cdot \frac{\dot{u}f_{t_2}}{\dot{u}f^{(V)}} = \dot{u}f_{t_2} \cdot \frac{f_x^{(V)}}{\dot{u}f^{(V)}}.$$

Specifické míry plodnosti v roce t z intervalu $(t_1; t_2)$ vypočteme podle vzorce pro výpočet lineární interpolace:

$$f_{t,x} = f_{t_1,x} + \frac{f_{t_2,x} - f_{t_1,x}}{t_2 - t_1} \cdot (t - t_1).$$

Stejným způsobem by se počítala další část scénáře vycházející z předpokladu, že do roku t_3 se úhrnná plodnost postupně zvýší na hodnotu $\dot{u}f_{t_3}$ a struktura plodnosti se postupně přiblíží struktuře nějaké vzorové plodnosti se specifickými mírami $f_x^{(W)}$.

Sofistikovanější scénáře plodnosti využívají pokročilé metody modelování, například analogie Lee-Carterova modelu, který se používá v případě úmrtnosti, quadratic spline model a další matematické přístupy. Tyto metody umožňují podrobnější a flexibilnější popis trendů plodnosti a její věkové struktury, což je užitečné zejména při komplexnějších analýzách. Pro ověření realističnosti scénáře se provádějí kontrolní výpočty, například porovnání specifických měr plodnosti v jednotlivých letech, analýza průměrného věku matek či zhodnocení hodnot modálního věku plodnosti a její výše v tomto věku. Také se sleduje, zda vývoj plodnosti zůstává plynulý, a to jak v úrovni, tak ve struktuře.

Migrace je další klíčovou komponentou prognózy, přičemž i zde se scénáře obvykle sestavují odděleně pro muže a pro ženy. Celkové migrační saldo nebo celkový počet imigrantů a emigrantů neposkytuje dostatečné informace, protože neodráží věkovou strukturu migrace. Pro použití komponentní metody projekce je proto třeba odhadnout předpokládané migrační saldo podle věku (ideálně dosaženého, jak již bylo rozvedeno v předchozích kapitolách), případně počty přistěhovalých a vystěhovalých podle věku $I_{t,x}$ a $E_{t,x}$ pro každý rok projekce.

Scénář migrace je nejobtížněji predikovatelnou součástí populačních prognóz. Migrace často nevykazuje jasné dlouhodobé tendence, protože je významně ovlivňována aktuálními podmínkami, jako jsou situace na trhu práce a bydlení, vládní politiky, politická a ekonomická situace, války, přírodní katastrofy či klimatické změny. V důsledku těchto faktorů je obtížné stanovit přesné scénáře.

Scénář migrace lze ilustrovat na příkladu jednoduchého scénáře imigrace (čisté migrace). Nejjednodušší přístup spočívá v prognóze celkového počtu přistěhovalých na základě předpokladu konstantních ročních počtů imigrantů v delším časovém horizontu. V některých případech je možné uvažovat postupné snižování, nebo naopak zvyšování migračních toků. Pohlavní a věková struktura imigrantů bývá považována za konstantní v čase a určuje se jako podíl počtu přistěhovalých daného pohlaví a věku z celkového počtu přistěhovalých na základě dat za několik posledních let. Pro podíl x -letých pohlaví p jako:

$$i_x^{(p)} = \frac{\sum_r I_{r,x}^{(p)}}{\sum_r (I_r^{(M)} + I_r^{(Z)})},$$

kde I je počet přistěhovalých, horní index označuje pohlaví, dolní indexy pak rok a věk. Pro vyrovnání hodnot se často využívají klouzavé průměry.

Počty imigrantů v jednotlivých letech podle pohlaví a věku určíme rozdělením celkového počtu podle předpokládané struktury:

$$I_{t,x}^{(p)} = I_t \cdot i_x^{(p)}.$$

Scénář je možné modifikovat například změnou pohlavní a věkové struktury imigrantů v čase (např. předpoklad nárůstu podílu žen a dětí – slučování rodin).

Podobný přístup je aplikován i na scénáře emigrace. Tendence k emigraci závisí na pohlaví a věku jednotlivců v populaci a je možné ji kvantifikovat specifickými kvocienty emigrace. Tyto kvocienty vyjadřují poměr mezi počtem vystěhovalých a celkovým počtem osob daného pohlaví a věku (resp. počtu živě narozených). Předpokládá se, že kvocienty emigrace zůstávají v čase neměnné, což umožňuje projekci budoucích migračních toků. Při projekci počtu vystěhovalých podle věku a pohlaví se obdobně využívají data za poslední roky.

Doplňme, že použití specifických měř emigrace (namísto výše uvedených kvocientů) by komplikovalo výpočet – střední stavy závisí mimo jiné na počtu vystěhovalých v daném roce –, došlo by k zacyklení výpočtu. Opět je vhodnější použít počty vystěhovalých tříděné podle dosaženého věku (který je vždy o 1 vyšší než dokončený věk osoby na začátku roku vystěhování).

Výpočet kvocientů emigrace (podle dosaženého věku) provedeme (podobně jako v případě struktury přistěhovalých) na základě dat za několik posledních let:

$$e_0 = \frac{\sum_r E_{r,0}}{\sum_r N_r},$$

$$e_x = \frac{\sum_r E_{r,x}}{\sum_r S_{r,x-1}} \quad \text{pro } x = 1, 2, \dots,$$

i zde je obvykle nutné provádět vyrovnání hodnot klouzavými průměry, aby bylo dosaženo hladkého vývoje. Projekce počtů vystěhovalých v roce t je následně definována jako:

$$E_{t,0} = N_t \cdot e_0,$$

$$E_{t,x} = S_{t,x-1} \cdot e_x \quad \text{pro } x = 1, 2, \dots$$

Prognózy migračních toků mohou být dále modifikovány podle dalších specifických předpokladů, nejen zvýšením podílu žen a dětí mezi imigranty v důsledku slučování rodin. Stejně tak může být scénář přizpůsoben, pokud se očekávají výrazné změny v ekonomické či politické situaci, které by mohly ovlivnit charakter migrace. Tyto přístupy jsou klíčové pro zajištění komplexnosti prognóz v oblasti migrace a toho, aby tyto prognózy byly realistické. Migrace totiž stále zůstává jednou z nejvíce variabilních složek demografického vývoje.

Příklad závislosti kvocientů emigrace na věku ilustruje, jak se migrace různých věkových skupin liší v závislosti na jejich specifických charakteristikách. U osob mladších 15 let lze obecně předpokládat, že jejich migrace probíhá spolu s rodiči. Rozhodnutí o migraci tedy obvykle nezávisí na pohlaví dítěte, což znamená, že poměr mezi migrujícími chlapci a dívkami by měl být zhruba stejný jako poměr pohlaví při narození. Pro tyto věkové skupiny je proto vhodné nejprve počítat strukturu migrantů podle věku bez ohledu na pohlaví. Poté lze celkové počty rozdělit na chlapce a dívky podle podílu děvčat při

narození, označovaného symbolem δ . Chlapci tedy tvoří podíl $(1 - \delta)$ z migrujících dětí, zatímco děvčata tvoří podíl δ . Stejný přístup lze aplikovat i na výpočet kvocientů emigrace pro děti do 15 let, kde opět není z výše uvedených důvodů vhodné provádět rozlišení hodnot kvocientů podle pohlaví.

V rámci scénářů emigrace může být nutné počítat s tím, že kvocienty emigrace se mohou v průběhu času měnit. Tyto změny mohou být směrem nahoru, například v důsledku ekonomických krizí, válek nebo politické nestability, nebo naopak dolů, například v období ekonomické prosperity a stabilní politické situace. Nejjednodušším způsobem, jak změny kvocientů modelovat, je předpoklad, že v konkrétním roce t jsou všechny kvocienty násobkem výchozích hodnot, zápis:

$$e_{t,x} = k_t \cdot e_x.$$

Tento způsob modelování je přehledný a efektivní, což jej činí vhodným pro základní projekce.

Tento postup ukazuje, jak lze relativně jednoduchými metodami zachytit některé dynamiky spojené s migrací a jejich věkovou strukturou, a to i za předpokladu, že jednotlivé aspekty migrace nejsou vždy snadno kvantifikovatelné.

Tvorba scénářů s využitím modelu Lee-Carter a jeho modifikací – stochastický přístup

Stochastické scénáře jsou na rozdíl od deterministických založeny na principu náhodnosti výskytu určitého jevu v populaci. Míry intenzity v těchto modelech jsou chápány jako realizace náhodné veličiny, která se může týkat jak minulosti, tak budoucnosti. Tyto modely umožňují provádět bodové i intervalové odhady, přičemž obsahují neurčitost vyplývající z pravděpodobnostního rozdělení sledovaného jevu. Obecně lze v jednoduchosti shrnout, že stochastické modely jsou komplikovanější z hlediska teoretického odvození a poměrně výrazně navázány na statistickou problematiku časových řad a vícerozměrné statistiky. Nepochybně nejznámějším modelem z této rodiny tvorby demografických scénářů je Lee-Carterův model, který bude podrobněji řešen v rámci této kapitoly.

Dále je zde model rozveden na takzvaný funkcionální demografický model, který je vlastně rozšířením Lee-Carterova modelu o modernější statistickou terminologii umožňující zjednodušeně řečeno vyčerpávat více variability v původních datech a odstranit i některé další nedostatky původního modelu plynoucí například z jeho malého počtu parametrů a tedy omezenější flexibility.

Jinak právě v kontextu prolínání se statistickou problematikou existují doslova stovky stochastických modelů, kdy se často jedná v podstatě o realizaci modelu odvozeného třeba i pro jiné účely na demografických datech. Vybrané modely v této kapitole jsou v praxi nejvíce používané, a navíc i dobře dostupné z hlediska výpočetního prostředí, proto byly zvoleny pro ilustraci této široké problematiky v rámci publikace seznamující se základy demografických projekcí.

Některé pojmy užití v kapitole vyžadují znalost statistických metod a terminologie z poměrně širokého spektra záběru. Protože nelze tyto pojmy jednoduše definovat, odkazujeme čtenáře v těchto příkladech v případě zájmu na doporučenou literaturu uvedenou v textu.

Původní Lee-Carterův model byl odvozen pro projekci úmrtnosti, jeho použití na příkladu plodnosti je nicméně ekvivalentní. V kapitole jsou odvozeny vzorce pro příklad plodnosti, kde je pro čtenáře možné si výpočty v celkovém kontextu lépe představit. U plodnosti je analyzováno méně věků, rozdíly v distribuci jsou patrné i bez logaritmického měřítka a dosažení do kohortně komponentní metody triviální (bez potřeby dopočtu projekčních koeficientů přes úmrtnostní tabulky z modelovaných měr úmrtnosti). U úmrt-

nosti se navíc musí řešit již v minulé kapitole zmíněná koherence (zde se například koherence někdy zajišťuje pomocí využití modelu na geometrický mezíprůměr s přičítáním odchylek). Stochastické modelování migrace je sice možné, nicméně ještě podstatně komplikovanější (a vzhledem k nestabilitě migrace obecně i v odborné komunitě místy považováno za diskutabilní z hlediska přínosu) a v publikaci nebude řešeno.

Je vhodné zmínit, že oba modely zde uvedené jsou založeny na redukci počtu dimenzí prostřednictvím singulárního rozkladu matice a metody hlavních komponent. Existují i rodiny modelů založené například na regresních splinech a dalších metodách.

Lee-Carterův model byl původně navržen v příspěvku „*Modeling and forecasting U. S. mortality*“ (Lee a Carter, 1992), nicméně se mu autor následně hojně věnoval i v dalších letech své publikační činnosti, jak dokládá například podstatně novější příspěvek „*Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method*“ (Li a Lee, 2005). Model je široce používán a modifikován i dalšími autory, jak již bylo nicméně zmíněno.

Před definováním samotného modelu je třeba upozornit na fakt, že vstupní data jsou z důvodu stabilizace jejich variability před samotnou analýzou prostřednictvím Lee-Carterova modelu podrobena logaritmické transformaci – možné zapsat ve formě:

$$f_{t,x}^{\text{LN}} = \ln(f_{t,x}),$$

kde $f_{t,x}$ označuje míru plodnosti v běžné podobě a $f_{t,x}^{\text{LN}}$ značí míru plodnosti po úvodní transformaci, která vstupuje do modelu (Lee a Carter, 1992).

Model vychází z myšlenky, že zahrnutím metod k redukci dimenzí do analýzy lze vývoj křivek měř plodnosti dle věku vyjádřit pomocí několika parametrů. Důležitý z hlediska využitelnosti modelu pro účely projekcí je fakt, že jeden z parametrů charakterizuje změnu probíhající v čase. Další parametry se pak zaměřují za rozdíly mezi jednotlivými roky věku. Konkrétně vychází z následující rovnice:

$$f_{t,x}^{\text{LN}} = \mu_x + \beta_t \cdot \varphi_x + e_{t,x},$$

kde věky reprodukčního období jsou klasicky nejčastěji $x = 15, 16, \dots, 49$ a čas se značí $t = 1, 2, \dots, T$. Úroveň věkově specifické plodnosti v daném věku a období je zachycena parametrem μ_x , dále model obsahuje věkově specifickou konstantu φ_x , která představuje kolísání plodnosti v konkrétním věku oproti β_t , přičemž tento parametr zachycuje změnu v čase t a v české terminologii se nazývá celkovou úrovní plodnosti (Šimpach, 2016). Kromě toho rovnice obsahuje $e_{t,x}$, které je možné chápat jako běžný residuální člen s vlastnostmi bílého šumu. Kvantifikace parametrů φ_x a β_t je provedena metodou singulárního rozkladu matice, která extrahuje první hlavní komponentu (Lee a Carter, 1992).

Úroveň věkově specifické plodnosti je následně zachycena prostým aritmetickým průměrem:

$$\mu_x = \frac{\sum_{t=1}^T f_{t,x}^{\text{LN}}}{T},$$

tím je model stručně definován, některé skutečnosti ještě budou nicméně rozepsány z důvodu přehlednosti srovnání obou modelů redukce dimenzí dále.

V případě funkcionálního demografického modelu se jedná, jak sami autoři uvádějí, o přirozené statistické rozšíření Lee-Carterova modelu. Statistickým rozšířením se myslí například skutečnost, že model využívá k vysvětlení úrovně plodnosti více než jen první hlavní komponentu. Kromě toho autoři přistupují k datům z funkcionálního pohledu, mění způsob stanovení procesu časové řady pro účely projekcí a nabízejí ještě i další možné úpravy. Za účelem tohoto statistického rozšíření využívají buď pokroky v teorii statistiky od dalších autorů, nebo případně navrhují vlastní postup.

Konkrétně lze uvedené nalézt v příspěvku „*Robust forecasting of mortality and fertility rates: A functional data approach*“ (Hyndman a Ullah, 2007). Užívání modelu a případným modifikacím se hlavní autor věnuje i v dalších publikacích, jako je například „*Stochastic population forecasts using functional data models for mortality, fertility and migration*“ (Hyndman a Booth, 2008). Významným počinem je pak sestavení balíčku „*demography*“ do programu RStudio (Hyndman a kol., 2023).

Každopádně tyto změny s sebou přináší úpravu ve vzorcovém vyjádření příslušných výpočtů, která je rozepsána v následujících odstavcích. Rovnice modelu je vzhledem k uvedenému pochopitelně o něco komplikovanější než v případě Lee-Carterova modelu.

Před uvedením rovnice modelu je nicméně třeba zmínit, že se opět realizuje transformace vstupních dat. Jako u Lee-Carterova modelu, také v případě funkcionálního demografického modelu byla nejprve navržena logaritmizace (Hyndman a Ullah, 2007), nicméně poměrně záhy začal autor v reakci na další výzkum využívat Box-Coxovu transformaci. Tu nakonec využívá i ve výpočetním prostředí demografického balíčku, který následně využívá většina autorů. Konkrétně se jedná o transformaci, která má pro účely měř plodnosti podle věku definovanou speciální formu:

$$f_{t,x}^{B-C} = \frac{1}{0,4} \cdot ((f_{t,x})^{0,4} - 1),$$

kde $f_{t,x}$ značí pozorovanou míru plodnosti ve věku x a čase t , $f_{t,x}^{B-C}$ značí míru v tom samém věku i čase, nicméně transformovanou a vstupující do modelu. Za vhodný transformační parametr je v případě plodnosti považována často úroveň čtyř desetin (viz např. Hyndman a Booth, 2008; Shang, 2015). Pro případ nulového transformačního parametru, který se používá například u úmrtnosti, by se vzorcový zápis lišil a jednalo by se v podstatě zjednodušeně opět o prostou logaritmizaci.

Rozklad transformovaných dat se provádí pomocí metody hlavních komponent, tentokrát lze vývoj plodnosti charakterizovat více než jen první hlavní komponentou a logicky tak vysvětlit i více variability. Funkci tohoto procesu je možné zapsat následujícím způsobem:

$$f_t^{B-C}(x) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \cdot \varphi_k(x) + e_t(x),$$

kde úroveň věkově specifické plodnosti v daném věku je charakterizována prostým aritmetickým průměrem a residuální člen lze chápat analogicky, jako ve vzorci u předchozího modelu. Co se týká dalších parametrů, $\beta_{t,k}$ charakterizuje změnu v čase $t = 1, 2, \dots, T$, pro komponenty $k = 1, 2, \dots, K$ a $\varphi_k(x)$ je sada ortonormálních funkcí věkově specifických konstant pro věky $x = 15, 16, \dots, 49$ a komponenty $k = 1, 2, \dots, K$.

Parametry je možné odhadnout metodou hlavních komponent, bylo však navrženo několik přístupů k ní, jak bude řešeno dále. Pro účely projekce budoucích hodnot měr plodnosti dle věku je obdobně jako u minulého modelu třeba odhadnout parametr $\beta_{t,k}$, který vlastně vyjadřuje komponentní skóre jednotlivých komponent. V případě, kde počet hlavních komponent přesahuje hodnotu jedné, tedy $K > 1$, je tak evidentně nezbytné provést projekci většího počtu časových řad než u původní varianty modelu.

Na vzorci je zřejmá patrné, že se jedná vlastně pouze o rozšířenou variantu vzorce uvedeného u Lee-Carterova modelu. Za předpokladu jiné varianty některého z modelů, ve které by se nelišila mezi modely vstupní transformace dat a přístup k datům, by se v případě $K = 1$ vlastně jednalo o stejný vzorec. Je tak dokázána teze týkající se úvah o funkcionálním demografickém modelu, jako o statistickém rozšíření modelu předchozího.

V této chvíli je možné považovat stručné shrnutí hlavních tezí, ze kterých modely vychází, a jejich metodologický popis za odpovídající pro potřeby dalšího textu. Vyčerpávající definice modelů včetně naznačení možných diferenčních přístupů k výpočtu je značně rozsáhlá, pro podrobnou metodologii je vhodné doplnit pohled o rešerši zmíněných příspěvků, ve kterých jsou modely definovány.

Z hlediska korektnosti je aktuálně stěžejní přejít spíše na specifika modelů, která se v případě těchto dvou modelů mohou lišit v každé konkrétní realizaci na základě preferencí autorů. Jak již bylo rozepsáno v rešerši, modely byly řadou autorů modifikovány z různých hledisek a za různým účelem.

Na rozdíl od předchozích kapitol, kde bylo naprostou většinou výpočtů možné vyřešit prostým dosazením do vzorců například v programu Excel, u stochastického modelování podpořeného odhadem parametrů metodou hlavních komponent je nepochybně pro alespoň minimální efektivitu výpočtu potřeba využít primárně statisticky zaměřený software. Věnujme tedy výpočetní realizovatelnosti modelů alespoň několik odstavců, které v předchozích kapitolách nebyly nezbytné. Následující odstavce navíc čtenáři dobře ilustrují komplikovanost výpočtů a volby postupu v případě volby stochastického modelování. Jak bude patrné, ani například informace, že odhad parametrů je proveden prostřednictvím metody hlavních komponent, nemusí být dostatečná, protože zde existují metody jako RAPCA (v případě zájmu viz Hubert a kol., 2002) a další, mezi kterými je třeba vybrat.

Patrně nejčastější, respektive určitě nejjednodušší, je předpoklad využití defaultních nastavení modelů v podobě, v jaké jsou dostupné v balíčku „*demography*“ pro software RStudio v jeho aktuální verzi 2.0 (Hyndman a kol., 2023), konkrétně jde o funkce *lca*, *lca.forecast* a *fdm*, *fdm.forecast*, které jsou určené k modelování demografických dat včetně měr plodnosti a úmrtnosti dle věku.

Přepisování kódového nastavení modelu, natož jeho automatizace, samozřejmě v řadě případů vyžaduje podstatně větší schopnosti a pro řadu autorů demografických projekcí tak při volbě varianty modelu nemusí být akceptovatelná.

Jsou zde například pro zvětšení uživatelské přívětivosti dle typu vložených dat automatizovány některé úkony, jako je například výše řešený transformační parametr definovaný pro účely plodnosti u Box-Coxovy transformace. Výpočetní prostředí vyšlo více než deset let od původní publikace s definovanou teorií a v určitých ohledech se liší od původního příspěvku. Opět lze jako příklad uvést právě transformaci vstupních dat. V případě

Lee-Carterova modelu se v uvedeném výpočetním prostředí jedná o variantu vycházející výhradně z původního příspěvku (Lee a Carter, 1992).

Je třeba rozhodnout, který přístup k redukci dimenzí model využije. V případě Lee-Carterova modelu se jednoduše jedná o singulární rozklad matice (Lee a Carter, 1992), který umožňuje nalézt první hlavní komponentu s minimalizovaným součtem čtverců (např. Good, 1969). V případě funkcionálního demografického modelu však přibýlo požadavků na metodu hlavních komponent a bylo navrženo více způsobů, jak tohoto cíle dosáhnout. Například lze využít metodu RAPCA (Hubert a kol., 2002), která sestaví odhady robustní z hlediska odlehlých pozorování. To je možné chápat jako jedno z potenciálních rozšíření, které může být zajímavé například u úmrtnosti v případě válek a podobně (Hyndman a Ullah, 2007). Nicméně jako doporučená defaultní metoda z hlediska výpočetního prostředí je v tomto případě stanovena „klasická“ metoda funkcionálních hlavních komponent (např. Ramsay a Dalzell, 1991).

Pro projekci časových řad parametrů komponentních skór se následně u obou modelů používají procesy vycházející z Box-Jenkinsovy metodologie (viz Box a Jenkins, 1970). V rámci modelů byly definovány i konkrétní způsoby automatizace volby procesu časové řady. V případě Lee-Carterova modelu se v jeho klasickém pojetí v publikacích i ve výpočetním prostředí jedná vždy o proces náhodné procházky s konstantou, jak bylo navrženo již při původní definici modelu (Lee a Carter, 1992). Funkcionální demografický model následně vybírá vhodný řád autoregresního integrovaného procesu klouzavých průměrů (známý více pod zkratkou ARIMA, z anglického autoregressive integrated moving average) u parametrů všech pořadí pomocí Akaikeho korigovaného informačního kritéria (např. Burnham a Anderson, 2002). Lze říci, že v postupu výběru procesu modelování časové řady autor vlastně vychází z funkce „*auto.arima*“ v balíčku „*forecast*“ (Hyndman a kol., 2024). Budoucí hodnoty měř plodnosti modely odhadnuté jsou samozřejmě na závěr upraveny inverzní transformací dat v podobě dle konkrétního modelu.

Závěrem kapitoly tedy zdůrazníme stěžejní fakt, který z uvedených skutečností plyne. Deterministické modely jsou sice definovány často přímo odborným odhadem, u stochastického modelování však je také na autorovi volba způsobu modelování časové řady, volba informačního kritéria a další rozhodnutí. Stejně jako nelze najít jediné kritérium vhodné pro všechny procesy z hlediska statistické teorie (v této se následně například sleduje porovnání úspěšnosti Akaikeho versus Schwarzova kritéria na různých datech apod.), nelze ani z demografického hlediska ponechat scénář (založený třeba na složitých statistických kritériích a procesech modelování), který je logicky nepravděpodobný bez ohledu na vysokou složitost jeho výpočtu. Populační, a tedy lidské chování se nemusí vyvíjet podle žádné křivky, bez ohledu na komplikovanost jejího matematické odvození, i u stochastických projekcí tak zůstává odborný názor autora stěžejní z hlediska finálních výsledků projekce.

Problematika výpočtu projekce malých územních celků

Projekce malých územních celků je proces, který zahrnuje odhad budoucího vývoje populace a subpopulací na různých úrovních územní struktury, například na úrovni státu a subpopulací krajů nebo okresů. Důležitým aspektem tohoto procesu je, že projekce nadřazené populace se nemusí vždy rovnat prostému součtu projekcí dílčích populací. K takovým rozdílům může dojít z několika příčin, mezi nimiž jsou nejvýznamnější rozdílné scénáře vývoje, které se aplikují na jednotlivé úrovně, a zaokrouhlovací chyby, jež vznikají při výpočtech.

Existují dva základní přístupy k výpočtu a korekci projekcí označované jako „zdola nahoru“ (from the bottom to the top) a „shora dolů“ (from the top to the bottom). Každý z těchto přístupů má své výhody i nevýhody a jejich aplikace závisí na konkrétních podmínkách a dostupnosti dat.

Přístup „zdola nahoru“ zahrnuje výpočet projekcí nejprve pro dílčí populace, přičemž nadřazená populace je určena jako součet těchto dílčích hodnot. Například projekce jednotlivých okresů se sečtou do projekcí krajů a součet projekcí krajů pak tvoří projekci celé republiky. Tento přístup je vhodný zejména tehdy, když jsou dostupná kvalitní data na nižší úrovni územní struktury. Jedním z hlavních úskalí této metody je, že projekce menších územních celků mohou být méně přesné, protože scénáře vývoje pro tyto oblasti často vycházejí z omezenějších dat za malé celky, která mohou generovat větší statistickou chybu. Navíc se statistické chyby těchto dílčích projekcí mohou kumulovat, což může vést k nepřesnostem na úrovni nadřazených územních celků.

Naopak přístup „shora dolů“ se zaměřuje na výpočet projekce pro nadřazenou populaci, přičemž subpopulace jsou průběžně korigovány tak, aby jejich součet odpovídal výchozí projekci nadřazené populace. Tento proces začíná například výpočtem projekce celé republiky, na jehož základě se stanoví projekce pro kraje. Během tohoto procesu se provádí korekce, aby součet počtu osob v každé věkové a pohlavní skupině za všechny kraje odpovídal projektovanému počtu osob pro celou republiku.

Lze uvést příklad jednoduché korekce počtu osob v roce t ve věku x v kraji k :

$$S_{t,x}^{(k,kor)} = S_{t,x}^{(k)} \cdot \frac{S_{t,x}^{(rep)}}{\sum_i S_{t,x}^{(i)}}$$

Druhý krok zahrnuje výpočet krajských projekcí na základě již korigovaných hodnot z předchozího kroku a postupně se pokračuje obdobně směrem k nižším úrovním, jako

jsou v komparaci například s kraji následně okresy. Tento přístup umožňuje přesnější kontrolu nad součtem, ale může být náročnější na výpočetní kapacity.

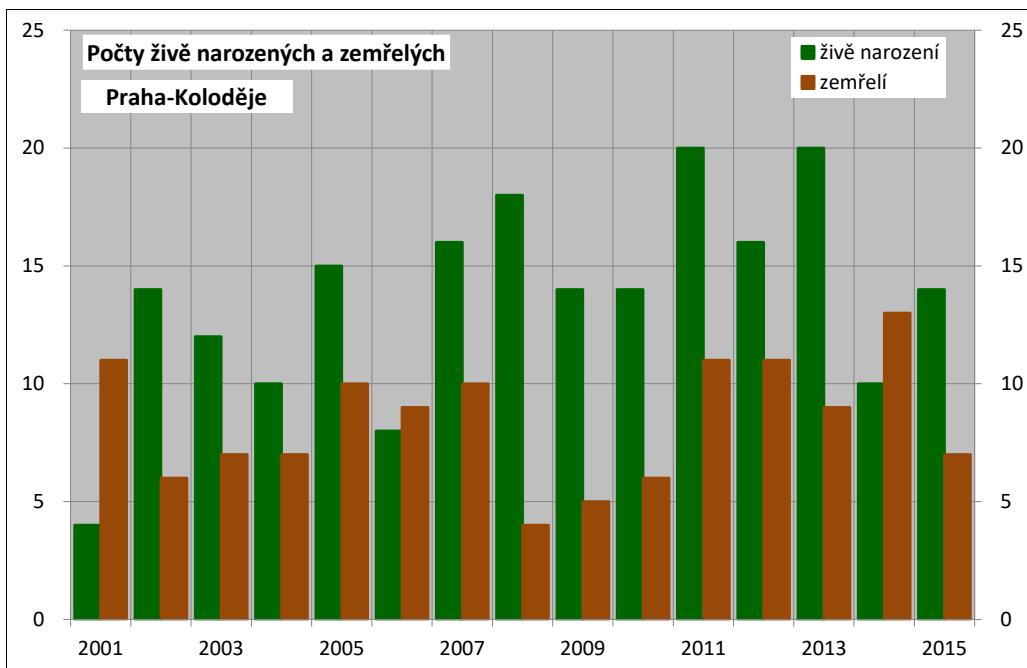
Při regionálních projekcích se základním problémem často ukazuje dostupnost dat. K provedení těchto projekcí je nutné mít aktuální informace o počtu obyvatel podle pohlaví a jednotek věku, počtu zemřelých, živě narozených, přistěhovalých a vystěhovalých podle pohlaví a věku. Tyto údaje jsou nezbytné pro výpočet specifických měr plodnosti, úmrtnosti a migrace, stejně jako pro tvorbu úmrtnostních tabulek a projekčních koeficientů. Data jsou běžně dostupná na úrovni okresů a vyšších územních celků, ale jejich detailnější úroveň může být problémem.

Počty obyvatel podle pohlaví a jednotek věku jsou běžně dostupné pouze za okresy a vyšší územní celky. Počty živě narozených, přistěhovalých, vystěhovalých podle jednotek věku jsou zpravidla zveřejněny pouze za kraje a vyšší územní celky. Český statistický úřad může na požádání poskytnout podrobnější údaje, například na úrovni obcí, avšak tyto údaje mohou být zatíženy poplatky a podléhat ochraně osobních údajů. Samotné poskytnutí dat záleží také na žadateli a konkrétním účelu.

Při práci s těmito daty je klíčové určit, zda postačí data za poslední rok, nebo zda je nutné mít delší časové řady. Pro počty obyvatel podle pohlaví a věku obvykle stačí aktuální rok, ale u jiných údajů, jako jsou počty narozených, zemřelých nebo migrační pohyby, je nutné mít data za několik posledních let. To je důležité zejména kvůli meziročním náhodným výkyvům a možnosti standardizace a odhadu hypotetických hodnot.

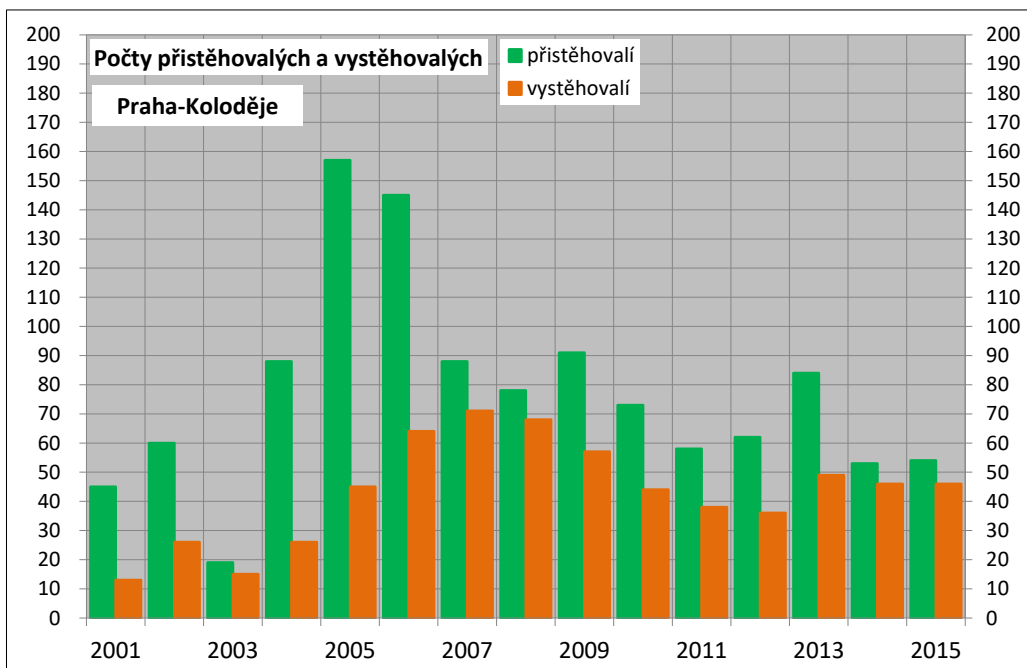
Výkyvy lze ilustrovat na získaných datech velmi malého územního celku Praha-Koloděje pomocí grafů převzatých z jiné aktivity autorů publikace (Fiala a Langhamrová, 2017).

Obrázek 1 | Výkyvy v počtu narozených a zemřelých



Zdroj: Fiala a Langhamrová (2017).

Obrázek 2 | Výkyvy v počtu migrantů



Zdroj: Fiala a Langhamrová (2017).

Scénář úmrtnosti v projekcích zahrnuje výpočet úmrtnostních tabulek, které vycházejí ze specifických měr úmrtnosti. Tyto míry mohou být u malých populací problematické, protože zde nízké počty zemřelých často vedou k nepoužitelným hodnotám. Počet zemřelých může být u malých celků jen v desítkách (např. 30), většina specifických měr úmrtnosti pak vyjde nulových a jsou tedy logicky nepoužitelné pro modelování úmrtnosti. Potenciálně se nabízí jako řešení použít počty zemřelých za více let. Například při 30 zemřelých ročně by nicméně evidentně ani 10 let nestačilo, možná údaje za 100 let, které jsou nicméně nedostupné a neaktuální.

Řešením je využití úmrtnostních tabulek větších populací stanovených za standard a jejich následná korekce pomocí srovnávacího indexu. Podobný přístup se používá i u scénářů plodnosti, kde se míry plodnosti menších populací odvozují od plodnosti větších (zpravidla nadřazených) populací.

Konkrétně v takovém případě výpočtu srovnávací index úmrtnosti lokální populace vzhledem ke standardní populaci (za období několika let) (metoda nepřímé standardizace, viz Langhamrová, Šimpach, 2013). Lokální úmrtnostní tabulky pak vytvořím z tabulek standardních vynásobením pravděpodobností biometrických měr. Na příkladu Prahy-Kolodějí bychom následně za standard použili úmrtnostní tabulky Prahy. Vypočetli bychom srovnávací index úmrtnosti v Kolodějích vzhledem k Praze jako standardu. Zde vyjde např. 1,05. Úmrtnostní tabulky pro Koloděje dostaneme z pražských tabulek vynásobením pravděpodobností úmrtí indexem 1,05 a přepočtem ostatních biometrických měr.

Vzniká zde vždy při konkrétní realizaci otázka optimální délky období. Příliš krátké období nese riziko velkých náhodných odchylek a příliš dlouhé období potřebuje více dat, potenciálně neaktuálních.

V případě plodnosti bychom postupovali obdobně. Použili bychom míry plodnosti nějaké větší populace (okres, kraj) jako standard. Vypočetli bychom srovnávací index plodnosti lokální populace vzhledem ke standardní populaci (za období několika let metodou nepřímé standardizace). Jako lokální míry plodnosti pro projekci bychom pak použili specifické míry plodnosti standardní populace vynásobením srovnávacím indexem.

Migrace je dalším klíčovým faktorem, který ovlivňuje regionální projekce. Scénáře imigrace často zohledňují počet dokončených bytů, velikost bytů a počet vystěhovalých a zemřelých v předchozích letech. Počet přistěhovalých do malé lokality závisí především na možnosti získání bydlení.

Lze proto přijmout předpoklad, že na počet přistěhovalých v daném roce má vliv:

- počet dokončených bytů v daném roce a velikost těchto bytů,
- počet vystěhovalých a zemřelých v předchozím roce (hodnoty z daného roku nelze použít – cyklické výpočty).

Možný jednoduchý vzorec pro scénář celkového počtu přistěhovalých v daném roce může být definován pomocí vzorce:

$$I_t = \delta \cdot (ob \cdot B_t + E_{t-1} + M_{t-1}),$$

kde B_t je předpokládaný počet dokončených bytů v roce t , ob je předpokládaný průměrný počet obyvatel nového bytu, δ je parametr intenzity imigrace.

Hodnotu parametru δ odhadneme na základě hodnot počtů dokončených bytů, počtů přistěhovaných, počtů vystěhovaných a počtů zemřelých v několika minulých letech jako řešení rovnice:

$$\delta = \frac{\sum_{t=t_1}^{t_2} I_t}{ob \cdot \sum_{t=t_1}^{t_2} B_t + \sum_{t=t_1-1}^{t_2-1} E_t + \sum_{t=t_1-1}^{t_2-1} M_t}.$$

Imigranty je následně potřeba rozdělit do počtů specifických dle pohlaví a věku. Vypočteme strukturu přistěhovaných podle pohlaví p a věku x ze známých hodnot za víceleté období:

$$i_x^{(p)} = \frac{\sum_{t=t_1}^{t_2} I_{t,x}^{(p)}}{\sum_{t=t_1}^{t_2} I_t},$$

eventuálně provedeme vyrovnání těchto hodnot klouzavými průměry. Musí samozřejmě platit:

$$\sum_{x,p} i_x^{(p)} = 1.$$

V letech projekce pak rozdělíme přistěhované podle pohlaví a věku podle této struktury:

$$I_{t,x}^{(p)} = I_t \cdot i_x^{(p)}.$$

V případě scénáře emigrace malých územních celků předpokládáme, že v každém věku má určitý podíl lidí tendenci se vystěhovat, intenzitu tohoto jevu charakterizují specifické kvocienty emigrace.

Předpokládáme tedy, že počet a pohlavní a věkové složení vystěhovaných v daném roce závisí na pohlavním a věkovém složení populace na počátku roku a na specifických kvocientech emigrace (pro každé pohlaví zvlášť), vyjádřeno jako:

$$E_{t,0} = N_t \cdot e_0,$$

$$E_{t,x} = S_{t,x-1} \cdot e_x \text{ pro } x = 1, 2, \dots$$

Specifické kvocienty emigrace opět odhadneme na základě počtu vystěhovaných podle pohlaví a věku a věkového složení obyvatelstva, resp. počtu živě narozených v několika posledních letech:

$$e_0 = \frac{\sum_{t=t_1}^{t_2} E_{t,0}}{\sum_{t=t_1}^{t_2} N_t},$$

$$e_x = \frac{\sum_{t=t_1}^{t_2} E_{t,x}}{\sum_{t=t_1}^{t_2} S_{t,x-1}},$$

eventuálně provedeme vyrovnání těchto hodnot klouzavými průměry. Věkem ve vzorcích u emigrantů rozumíme dosažený věk.

Specifické téma představuje migrace osob do 15 let. Tyto osoby migrují zpravidla s rodiči. Lze předpokládat, že rozhodnutí rodičů o migraci nezávisí na pohlaví jejich dětí, že tedy poměr mezi migrujícími chlapci a dívkami je zhruba stejný jako poměr pohlaví při

narození. Strukturu imigrantů podle věku do 15 let počítám bez rozlišení pohlaví, pak rozdělím na chlapce a děvčata v poměru $(1 - \delta)$, resp. δ , kde δ označuje podíl děvčat při narození. Podobně kvocienty emigrace pro děti do 15 let počítám bez rozlišení pohlaví.

Specifickým způsobem výpočtu projekce malého územního celku je postup získání projekčních koeficientů přímo z věkové struktury. Tento postup má smysl, pokud nejsou k dispozici podrobná data za migraci. Použije se zde věková struktura za více let k odhadu projekčních koeficientů bez potřeby znalosti měr úmrtnosti jako:

$$SP_x^{(p)} = \frac{\sum_{t=t_1}^{t_2} S_{t+1,x+1}^{(p)}}{\sum_{t=t_1}^{t_2} S_{t,x}^{(p)}}.$$

Uvedené projekční koeficienty v sobě obsahují současně vliv úmrtnosti i migrace, mohou být vyšší než 1. Ve vzorci pro výpočet projekce odpadá migrační člen, počítá se jako projekce bez migrace. Tato metoda výpočtu je použitelná pro projekci, která předpokládá úmrtnost a migraci po celé období projekce stejnou jako v intervalu $t_1 - t_2$. Pokud se má předpokládat v období projekce měnící se úroveň migrace a úmrtnosti, je problematické, jak upravit uvedené koeficienty. Výpočet je obecně dost zjednodušující a riskantní z hlediska validity výsledků projekce. Přistupuje se k němu spíše v případě, kdy nelze smysluplně realizovat komplikovanější postup.

Celý proces projekce zahrnuje nejen výpočet, ale i prezentaci a interpretaci výsledků. Zatímco po provedení výše definovaných úprav je možné provést výpočet kohortně komponentní metodou podle stejných vzorců jako projekce větších územních celků, existuje rozdíl ve způsobu vizualizace a prezentace výsledků. Tyto výsledky je vhodné prezentovat graficky například v pětiletých věkových skupinách, což je přehlednější než detailní tabulky. Interpretace projekcí je zaměřena na identifikaci trendů a středních hodnot vývoje, přičemž je důležité brát v úvahu, že skutečné hodnoty budou oscilovat kolem projektovaných hodnot.

Odvozené projekce mají zásadní význam v demografii, ekonomii a dalších společenských vědách, neboť umožňují detailnější vzhled do budoucího vývoje populací a jejich potřeb. Základním cílem těchto projekcí je překročit rámec tradičního rozdělení populace pouze podle věku a pohlaví, které nabízí například výpočet kohortně komponentní metodou, a zaměřit se na další charakteristiky, jež mají často klíčový vliv na plánování veřejných politik či ekonomických strategií. Jednoduché projekce založené na těchto odvozených charakteristikách využívají specifických měr, jako jsou míry vzdělanosti, ekonomické aktivity nebo potřeby specifické péče, a jejich aplikace na předpokládané počty obyvatel v jednotlivých věkových a pohlavních skupinách.

Podívejme se nejdříve na principy a příklady aplikace odvozených projekcí. Základní myšlenkou jednoduché odvozené projekce je odhadnout podíl osob s určitou charakteristickou vlastností v každé věkové a pohlavní skupině pro každý rok projekce. Tato metoda umožňuje například odhadnout počet žáků různých typů škol, zájemců o studium, ekonomicky aktivních osob, nebo osob vyžadujících specifickou péči. Výpočet je v nejjednodušší variantě odvozené projekce relativně přímočarý, projektované počty osob se vynásobí specifickými mírami, čímž vzniká odhad počtu osob s danou charakteristikou. Pokud není třeba rozlišovat mezi pohlavími, lze výpočty zjednodušit a aplikovat je na celkový součet populací mužů a žen.

Příkladem může být projekce počtu ekonomicky aktivních osob, kde je nutné znát nejen počet obyvatel ve specifických věkových skupinách, ale také míry ekonomické aktivity pro každé pohlaví. Tyto míry se obvykle liší, a proto je často vhodné provádět výpočty odděleně pro muže a ženy. Projekce zahrnují i scénáře, které zohledňují možný vývoj sledovaných charakteristik. Například změny v mírách ekonomické aktivity mohou být ovlivněny posuny v důchodovém věku nebo jinými faktory pracovního trhu.

Matematický zápis podobné úvahy by následně bylo v případě ekonomické aktivity možné zapsat jako:

$$S_{t,x}^{(EA)} = S_{t,x} \cdot a_{t,x},$$

kde $S_{t,x}^{(EA)}$ je počet ekonomicky aktivních osob v roce t ve věku x , $S_{t,x}$ je počet obyvatel v roce t ve věku x , $a_{t,x}$ je míra ekonomické aktivity v roce t ve věku x .

Rozlišení pohlaví ve vzorci pro jednoduchost uvedeno není, výpočet se nicméně provádí zpravidla pro každé pohlaví zvlášť, jak bylo již zmíněno výše.

Klíčovým krokem při tvorbě odvozených projekcí je stanovení scénáře, který definuje vývoj sledovaných specifických měr. Tyto míry nemusí vždy odrážet zájem o danou aktivitu, ale například dostupnost služeb, což lze ilustrovat na příkladu předškolního vzdělávání. Pokud 60 % dětí v určité věkové kategorii navštěvuje mateřskou školu, nemusí to nutně znamenat, že rodiče zbývajících 40 % dětí nemají zájem o tyto služby – důvodem

může být nedostatečná kapacita. V takových případech je vhodné vytvářet různé varianty projekcí, například optimální scénář (odrážející dostupnost) a maximální scénář (založený na zájmu).

Další významnou otázkou je, zda lze specifické míry považovat za neměnné v čase. U některých charakteristik, jako je dosažené vzdělání, může být změna velmi omezená, protože dosažená úroveň vzdělání zpravidla přetrvává po zbytek života. U jiných charakteristik, jako je ekonomická aktivita, může být vývoj měř značně ovlivněn vnějšími faktory, například politikami zaměstnanosti nebo demografickým stárnutím, případně zvyšováním důchodového věku.

V jiných typech projekcí je dosažení určité charakteristiky nevratné a přetrvává po zbytek života (například nejvyšší úroveň vzdělání). Při tvorbě scénáře je tedy nutný generační pohled.

Například lze uvést skutečnosti očekávané v rámci současné společnosti předpokládající, že podíl osob s nejvyšším vzděláním by měl v každé generaci růst nebo stagnovat $e_{t+1,x+1}^{(VŠ)} \geq e_{t,x}^{(VŠ)}$.

Naopak podíl osob s nejnižším vzděláním by měl v každé generaci klesat nebo stagnovat $e_{t+1,x+1}^{(ZŠ)} \leq e_{t,x}^{(ZŠ)}$.

U středních stupňů vzdělání pak je nejednoznačný trend. Osoby, které dokončí SŠ, zvyšují podíl osob se středním vzděláním a osoby, které absolvují VŠ, tento podíl snižují. Pro tvorbu racionální odvozené projekce je tak nezbytné mít určitý socioekonomický náhled o zkoumané charakteristice.

V jiných typech projekcí přetrvává dosažení určité charakteristiky po zbytek života s poměrně vysokou pravděpodobností. Například absolventi studia určitých oborů budou s vysokou pravděpodobností celý život či většinu života pracovat v tomto oboru (lékaři, IT-odborníci a další).

Podíl osob s určitou vlastností je však ovlivněn též demografickým vývojem, úmrtností a především migrací. Podíl osob s VŠ vzděláním se může zvyšovat také díky nižší úmrtnosti těchto osob v porovnání s osobami s nižším vzděláním. Větší emigrace osob s VŠ vzděláním naopak způsobí pokles podílu těchto osob. Sofistikovanější způsob projekce podle vzdělání pak poskytuje tzv. vícestavová projekce.

V případech, kdy je potřeba zohlednit více stavů nebo charakteristik, například dosažené úrovně vzdělání nebo pracovní specializace, se využívají vícestavové projekce. Tyto modely berou v úvahu nejen demografické faktory, ale také přechody mezi stavy (například dosažení vyššího vzdělání či změnu zaměstnání). Jejich problematika je nicméně výrazně složitá a přesahuje cíl této publikace seznámit čtenáře se základy demografické prognostiky.

Odvozené projekce mohou být aplikovány na širokou škálu charakteristik, jako je vývoj počtu osob s určitým hendikepem nebo projekce příjmů a výdajů zdravotního či důchodového systému. U zdravotního systému je klíčové zohlednit nejen zaměstnanost a mzdy, které ovlivňují příjmy z pojištění, ale také výdaje na zdravotní péči, které mohou být ovlivněny například snižující se úmrtností. Snižování úmrtnosti může mít dvojí dopad – na jedné straně zvyšuje náklady na zdravotní péči díky delšímu dožití, na druhé straně může prevence a zdravý životní styl náklady snížit.

Podobně v oblasti důchodového systému mohou odvozené projekce pomoci nalézt rovnováhu mezi příjmy a výdaji prostřednictvím úprav důchodového věku, pojistného nebo valorizace důchodů. Důležité je také zohlednit demografické faktory, jako je migrace, která může ovlivnit například podíl osob s vysokoškolským vzděláním. Větší emigrace této skupiny obyvatel může podíl snižovat, zatímco nižší úmrtnost oproti méně vzdělaným skupinám jej může naopak zvyšovat.

Lze uvést matematický zápis ještě některých konkrétních případů pro lepší ilustraci, například projekce příjmů zdravotního systému by se dala kvantifikovat:

$$ZzP_t = \sum_{x=15}^{\omega-1} \frac{S_{t,x} + S_{t+1,x}}{2} \cdot z_{t,x} \cdot w_{t,x} \cdot zp,$$

kde $z_{t,x}$ je míra zaměstnanosti v roce t ve věku x , $w_{t,x}$ je průměrná roční mzda zaměstnané osoby v roce t ve věku x , zp je sazba zdravotního pojištění stanovená procentem ze mzdy.

Měli bychom mít na paměti, že příjmy jsou zde odvozy zdravotního pojištění, vzorec tedy vyjadřuje úhrn těchto odvodů od zaměstnaných osob v roce t . Výpočet by zde opět bylo vhodné provádět zvlášť pro muže a zvlášť pro ženy. Analogicky by se prováděl výpočet úhrnu pojistného odvedeného OSVČ, odvedeného státem za státní pojištěnce, odvedeného samoplátci atd. U posledních dvou skupin by byl vzorec jednodušší – výše odvedeného pojistného zpravidla nezávisí na mzdě těchto osob, tedy ani na jejich věku.

Obdobně by se dala vypočítat projekce výdajů zdravotního systému. Výdaji jsou zde vlastně platby za zdravotní péči a jejich úhrn v roce t lze tedy vyjádřit jako:

$$ZV_t = \sum_{x=0}^{\omega-1} \frac{S_{t,x} + S_{t+1,x}}{2} \cdot zv_{t,x},$$

kde $zv_{t,x}$ jsou průměrné výdaje na zdravotní péči v roce t na osobu ve věku x .

Do výpočtu průměru se musí zahrnout i osoby, které žádnou zdravotní péči nečerpaly. Výpočet je opět vhodné provádět ideálně zvlášť pro muže a zvlášť pro ženy.

Výstupy těchto projekcí mohou být využity pro plánování veřejných politik, například v oblasti školství, zdravotnictví nebo trhu práce. Odvozené projekce tedy představují klíčový nástroj pro pochopení a plánování budoucího vývoje v různých oblastech života společnosti. Díky jejich flexibilitě a schopnosti zohlednit různé scénáře nabízejí možnost informovaného rozhodování založeného na podrobných analýzách dat a trendů.

Metodika a hlavní výsledky projekcí významných pracovišť

Skutečnost, že demografické projekce mají zásadní význam pro plánování v mnoha oblastech veřejného a soukromého života, si samozřejmě uvědomují oficiální statistické instituce. V České republice je jedním z nejdůležitějších subjektů v oblasti demografických projekcí Český statistický úřad, který pravidelně zpracovává projekce obyvatelstva v pětiletých cyklech. Tyto projekce se staly klíčovým nástrojem například při stanovování důchodového věku po roce 2030 a poskytují podklady pro dlouhodobé strategické plánování v dalších oblastech, jako je zdravotnictví, školství nebo bydlení.

Cílem kapitoly je seznámit čtenáře s vybranými aktuálními praktickými aplikacemi demografického prognózování. Budou zde shrnuty hlavní teze konstrukce a výsledky aktuálně nejnovějších projekcí Českého statistického úřadu, Eurostatu a Organizace spojených národů. Projekci Českého statistického úřadu bude věnován největší prostor a nadnárodní projekce následně doplní čtenáři pohled o další možné přístupy. Kromě statistických institucí je samozřejmě řada demografických projekcí zejména na úrovni malých územních celků plněna formou veřejných zakázek ze strany měst, asociací apod. Těchto projekcí je nicméně poměrně velké množství a nelze tedy všechny aktuální postihnout v rámci této publikace, zájemci jistě zvládnou ty z nich, které jsou zveřejněné, dohledat na příslušných webových stránkách zadavatelů.

Poslední projekce obyvatelstva České republiky Českého statistického úřadu byla zveřejněna 30. listopadu 2023. Zpracoval ji tým odborníků z oddělení demografické statistiky Českého statistického úřadu za podpory externích expertů. Projekce je založena na kohortně komponentní metodě, která umožnila zpracovatelům postupovat po jednotkách věku a v jednoletých krocích. Klíčovými proměnnými v této projekci jsou plodnost, úmrtnost a migrace, jejichž vývoj je modelován podle věku a pohlaví.

Základním prahovým stavem byl počet obyvatel podle pohlaví a věku k 1. lednu 2023. Tento údaj vychází z výsledků sčítání lidu, domů a bytů v roce 2021, doplněných o demografické události z let 2021 a 2022. Projekce se zaměřuje na období do okamžiku 1. ledna 2101, kdy poskytuje nejen předpokládané stavy populace, ale také relativní a analytické ukazatele plodnosti, úmrtnosti a migrace až do roku 2100.

Projekce zahrnuje tři varianty – střední, nízkou a vysokou, které zohledňují různé scénáře vývoje demografických procesů.

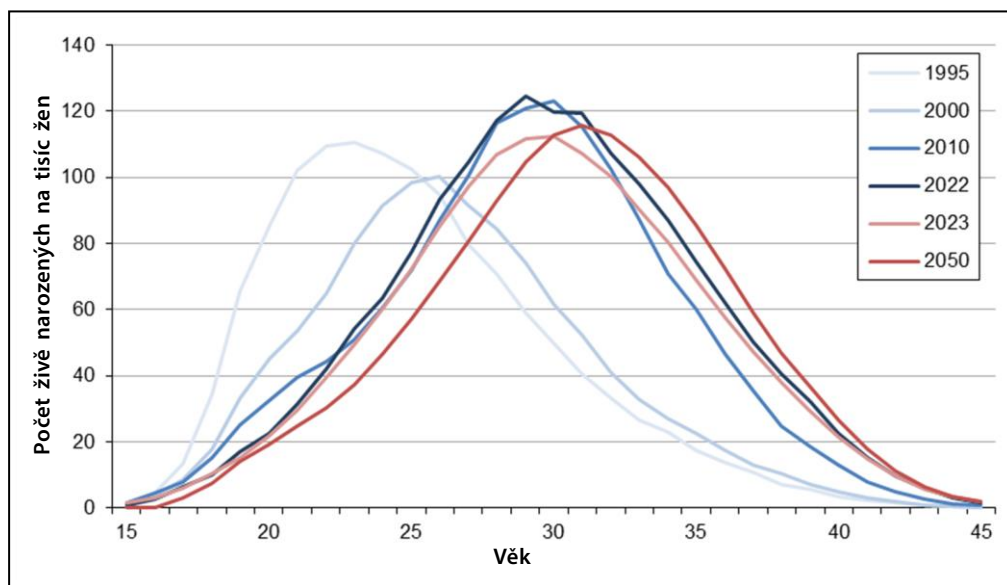
Model střední varianty plodnosti vychází z analýzy trendů od roku 1990, zejména z posledních let před projekcí. Jedním z klíčových rysů české demografie v 90. letech 20. století byl dramatický pokles plodnosti spojený s odkladem rodičovství do vyššího

věku. Od počátku 21. století do roku 2018 byla téměř ve všech letech pozorována kompenzace tohoto poklesu, přičemž úhrnná plodnost se v době posledních dat při tvorbě projekce držela okolo 1,71 dítěte na ženu. Výrazným nárůstem byla hodnota 1,83 v roce 2021, avšak následovala ji opět klesající tendence na 1,62 v roce 2022, zčásti ovlivněná přícho- dem ukrajinských žen s nízkou plodností.

Vzhledem k výrazným výkyvům v letech 2020–2022 autoři projekce přijali konzer- vativní přístup, předpokládající stabilní úroveň plodnosti na hodnotě 1,50 dítěte na ženu. Dalším trendem je pokračující odkládání těžiště reprodukce do vyššího věku, což se pro- jeví růstem průměrného věku matek z 30,4 let v roce 2022 na 31,3 let v roce 2050.

Úroveň úhrnné plodnosti se podle středního scénáře stabilizuje dle odhadu roku 2023 na úrovni 1,50 s podíly podle pořadí na úrovni 0,72 u prvního dítěte, 0,56 u druhého a 0,21 u třetího či dalšího dítěte. Tyto hodnoty reflektují evropské trendy, včetně věkové struktury plodnosti podobné té, která byla v Německu v roce 2021.

Obrázek 3 | Míry plodnosti dle věku v projekci ČSÚ

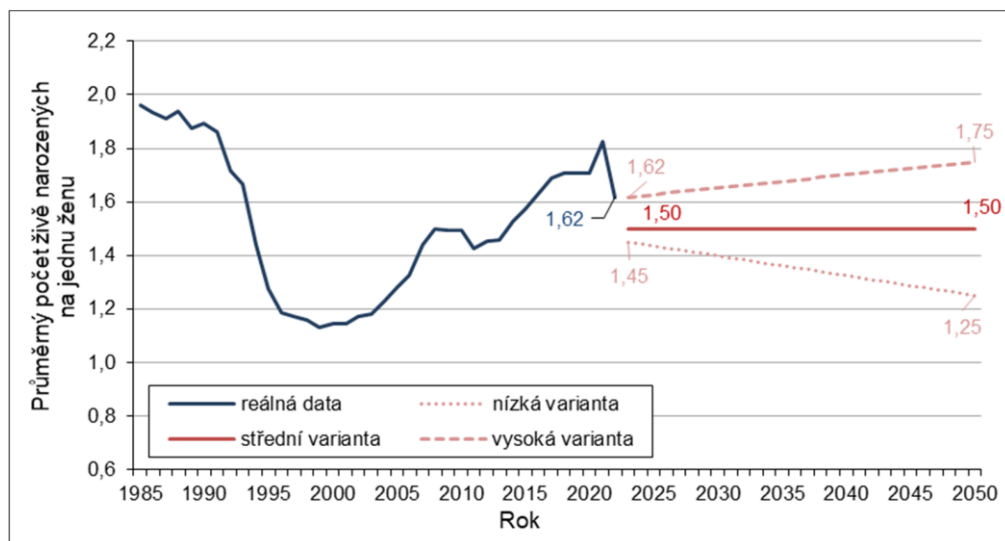


Zdroj: Projekce obyvatelstva České republiky – 2023–2100 (ČSÚ, 2023).

Nízká varianta plodnosti předpokládá zastavení trendu posouvání narození dětí do pozděj- šího věku, přičemž věkový profil plodnosti se zafixuje na úrovni roku 2022. Dále se oč- kává pokračování stávajícího mírného poklesu plodnosti, přičemž úhrnná plodnost bude lineárně klesat na hodnotu 1,25 dítěte na ženu v roce 2050. Vysoká varianta se zaměřuje na pokračující odkládání plodnosti, přičemž vzorem pro věkově specifický průběh plod- nosti bude Švýcarsko (z roku 2021), kde je výrazně vyšší míra plodnosti ve věku 30–40 let a nad 40 let ve srovnání s Českou republikou. Průměrný věk matek při narození dítěte zde činí 32,3 let, což je hodnota projektovaná pro Česko v roce 2050. V této variantě se oč- kává lineární růst úhrnné plodnosti na hodnotu 1,75 dítěte na ženu v roce 2050. Stejně jako

u střední varianty se i v krajních variantách předpokládá zafixování projektované výše a profilu plodnosti na úrovni roku 2050 pro druhou polovinu století.

Obrázek 4 | Úhrnná plodnost v projekci ČSÚ



Zdroj: Projekce obyvatelstva České republiky – 2023–2100 (ČSÚ, 2023).

Scénář střední varianty úmrtnosti vycházel ze skutečnosti, že Česká republika dlouhodobě zaostává v naději dožití za nejvyspělejšími evropskými státy o 2–4 roky, přičemž aktuální hodnoty dosahují úrovně v západní Evropě zhruba před 15 lety. Přesto je hlavním rysem růst naděje dožití, což je klíčovým aspektem i pro tuto projekci.

Projekce střední varianty předpokládá, že Česko bude postupně sledovat trajektorii vývoje úmrtnosti, která byla pozorována v zemích s nejvyšší nadějí dožití. Základem pro výpočet byl vytvořen tzv. evropský úmrtnostní potenciál, který reflektuje nejnižší specifické míry úmrtnosti v jednotlivých věkových skupinách v zemích jako Švýcarsko, Švédsko nebo Nizozemsko a dalších.

Vstupní data pro výpočet specifických měr úmrtnosti, tedy počty zemřelých a počty osob středního stavu, byly pro každou z těchto tří populací sečteny podle věku, pohlaví a roku, čímž vznikla populace uvažovaného potenciálu. Následně byly pro tuto populaci spočítány specifické míry úmrtnosti podle věku, které byly použity v dalších krocích výpočtu. Aby byly eliminovány náhodné fluktuace měr úmrtnosti a pro potřeby extrapolace do věku 110 let, byly specifické míry úmrtnosti potenciálu v letech 1980–2019 vyrovnány. Toto vyrovnání probíhalo stejnou metodou, jaká se používá při výpočtu standardní úmrtnostní tabulky ČSÚ.

Odhadovaný vývoj měr úmrtnosti potenciálu podle jednotek věku a pohlaví je analyzován ve vzájemném vztahu specifických měr úmrtnosti Česka a potenciálu v období posledních 15 let (do roku 2019). Relativní rozdíly, vyjádřené indexy nadúmrtnosti, jsou počítány podle jednotek věku a pohlaví pro každý kalendářní rok. Pro analýzu byl použit nástroj joint-point regrese, která navrhl funkci umožňující modelování průběhu indexů v jednotlivých věkových skupinách a v kalendářních letech. Indexy nadúmrtnosti Česka

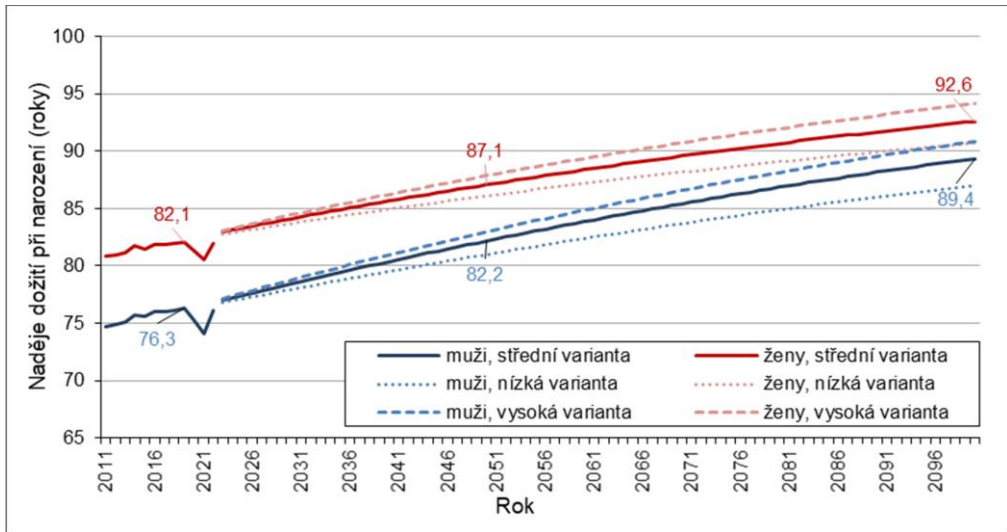
proti potenciálu jsou nejvyšší v dětských věkových skupinách (kde úmrtnost přesahuje dvojnásobek úmrtnosti potenciálu). Tento vliv na střední délku života je však nepatrný a může být ovlivněn nízkými počty úmrtí v obou populacích. U mužů je podobně vysoká nadúmrtnost pozorována u padesátníků, u žen pak nejvyšší index nadúmrtnosti oproti potenciálu nastává ve vyšším věku, přičemž maximální hodnota je kolem 75 let. Tyto rozdíly v dospělém věku odrážejí dosavadní vývoj úmrtnosti a určité rezervy, které si česká populace v porovnání s nejvyspělejšími státy Evropy nese.

Budoucí vývoj úmrtnosti potenciálu je projektován pomocí koherentní varianty modelu Lee-Carter. V rámci střední varianty se očekává zachování úrovně nadúmrtnostních indexů na úrovni průměru z období 2015–2019. S poklesem úmrtnosti potenciálu dojde k poklesu i očekávané úmrtnosti obyvatel Česka, což povede ke zmenšení rozdílu v naději dožití při narození mezi Českem a potenciálem. Nízká varianta předpokládá konstantní rozdíl mezi střední délkou života v Česku a v potenciálu, jaký byl pozorován na počátku projekce. Vysoká varianta počítá s větším přibližováním specifických měr úmrtnosti Česka a potenciálu než ve střední variantě, což znamená dosažení přibližně o 1,5 roku vyšší naděje dožití při narození v roce 2100. V nízké i vysoké variantě je modelován vývoj indexů nadúmrtnosti v čase. Odhad naděje dožití a struktura úmrtnosti podle pohlaví a věku pro rok 2023 vychází i z posledních dostupných dat v době tvorby projekce o počtech zemřelých za část roku 2023. V nízké variantě byla naděje dožití odhadována o 0,2 roku nižší než ve variantě střední, zatímco ve vysoké variantě byla odhadována o 0,1 vyšší.

Naděje dožití při narození má do roku 2050 vzrůst na 82,2 let u mužů a 87,1 let u žen, přičemž do konce století dosáhne 89,4 let u mužů a 92,6 let u žen. Zároveň se očekává postupné snižování rozdílu mezi pohlavími, a to ze současných 5,9 let na 3,2 let.

Vzhledem k velmi nízkým úrovním úmrtnosti v kojeneckém a dětském věku je jejich vliv na vývoj střední délky života téměř zanedbatelný, i když dochází k poklesu úmrtnosti v těchto skupinách. Těžiště růstu naděje dožití se přesouvá do věkových skupin 70–74 let u mužů a 80–84 let u žen. V posledních dvaceti letech projekce se pak toto těžiště posouvá do věkových skupin 80–84 let u mužů a 85–89 let u žen. Počty zemřelých se stále více koncentrují kolem modálního věku (nejčastějšího věku) úmrtí, což odpovídá probíhajícímu procesu komprese úmrtnosti. Tento trend je patrný v dlouhodobém vývoji a je běžný v rozvinutých státech po celém světě.

Obrázek 5 | Naděje dožití v projekci ČSÚ



Zdroj: Projekce obyvatelstva České republiky – 2023–2100 (ČSÚ, 2023).

Připomeňme, že migrace je považována za nejvíce nejistou složku demografické projekce, neboť závisí na řadě vnějších faktorů, jako jsou ekonomický vývoj, legislativní opatření či geopolitické události. Scénář střední varianty migrace pro Českou republiku vychází z několika východisek, která se zaměřují na vývoj migračního salda a jeho dopad na demografické složení obyvatelstva. Česko je zemí s pozitivním saldem zahraničního stěhování, přičemž od vstupu do EU v roce 2004 došlo pouze k jednomu zápornému saldu v roce 2013, které bylo způsobeno administrativními ukončeními pobytu řady cizinců. Od roku 2015 migrační saldo vykazuje rostoucí trend, přičemž v roce 2015 činilo necelých 16 tisíc osob, v roce 2018 to bylo již téměř 39 tisíc osob a v roce 2021 dokonce bezmála 50 tisíc osob. Věkově-pohlavní skladba tohoto salda je poměrně stabilní, přičemž nejvyšší hodnoty se nacházejí ve věkové skupině 25–29 let, což představuje pětinu až čtvrtinu celkového salda, a dále ve skupině 30–34 let. Mírně převažovali muži, přičemž za období 2018–2021 tvořili 58,6 % migrujících.

Migrační statistika v ČR však naráží na problémy, jako je částečné neodhlašování při dlouhodobém pobytu v zahraničí či administrativní zásahy do evidencí. Přesto migraci nelze ignorovat, neboť významně ovlivňuje změny počtu a složení obyvatelstva. V zemích s plodností pod hranici prosté reprodukce může migrace být jediným faktorem, který zajistí růst či stabilizaci počtu obyvatel a zmírní populační stárnutí.

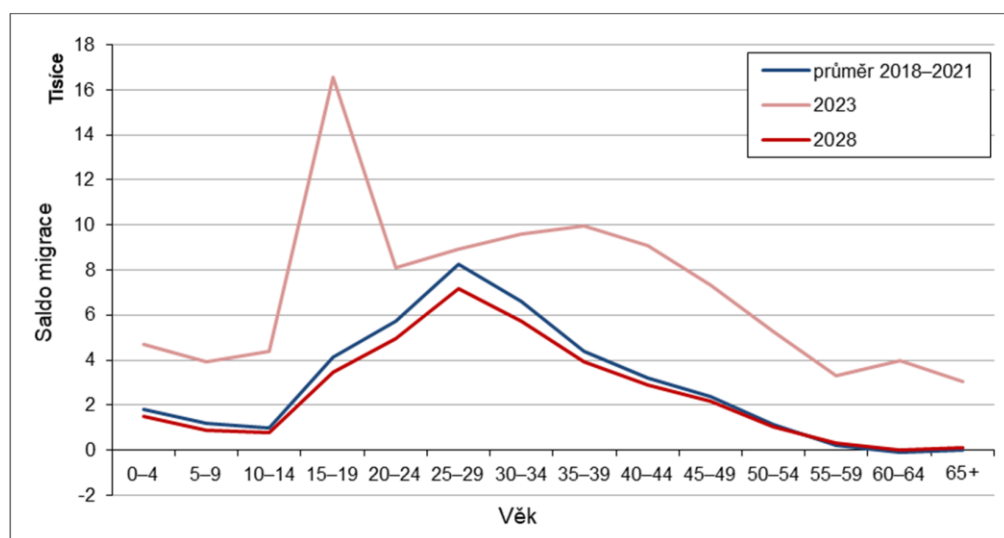
Významná událost, která ovlivnila migrační saldo, byla válka na Ukrajině, kdy Česká republika přijala až několik set tisíc uprchlíků. V roce 2022 bylo saldo zahraniční migrace téměř 330 tisíc, což bylo nejvíce ze všech zemí na počet obyvatel. Pohlavní a věková struktura nově přistěhovaných byla odlišná, s převahou žen, a největší přírůstek byl ve věkových skupinách 15–19 a 35–39 let. V roce 2023 pokračoval příliv migrantů, přičemž i když byl nižší než v roce 2022, stále jej tvořili převážně uprchlíci.

V prvních letech projekce byly formulovány předpoklady zvláště pro „běžné“ migranty a zvláště pro uprchlíky. Scénář střední varianty migrace předpokládá pokračující

migrační atraktivitu ČR, což znamená kladné saldo stěhování. Pro rok 2023 se očekává pokles salda běžné migrace o 60 %, a od roku 2025 se projekce stabilizuje na roční saldo běžné migrace 35 tisíc osob (s variantami: nízkou – 25 tisíc osob a vysokou – až 40 tisíc osob). Do roku 2050 by měl podíl žen mezi migranty vzrůst na 47 %, což souvisí s rostoucími požadavky na pracovní sílu v pečovatelských a zdravotních službách.

V případě uprchlíků z Ukrajiny se odhaduje, že se v roce 2023 přistěhuje 108 tisíc uprchlíků, přičemž tento počet v roce 2024 klesne na polovinu. V následujících letech, od roku 2025, již projekce nepočítá se žádnou specifickou skupinou uprchlíků, pouze s běžnou migrací. Podle střední varianty se očekává, že 60 % přistěhovaných uprchlíků opustí ČR nejpozději do roku 2027, s možností návratu k vyšší míře odchodů v závislosti na vývoji situace na Ukrajině. Vysoká varianta počítá s odchodem 55 % uprchlíků a nízká se 70 % uprchlíků.

Obrázek 6 | Saldo migrace v projekci ČSÚ



Zdroj: Projekce obyvatelstva České republiky – 2023–2100 (ČSÚ, 2023).

Eurostat, statistický úřad Evropské unie, zpracovává pravidelné projekce populace pro členské státy EU a několik dalších evropských zemí (Island, Lichtenštejnsko, Norsko a Švýcarsko).

Projekce Eurostatu využívají také kohortně komponentní metodu. Projekce jsou počítány v jednotkách věku po jednotlivých letech, tedy s jednoletým časovým krokem. Poslední projekce byly zahájeny prahem k 1. lednu 2023 a jejich horizont sahá až k 1. lednu 2100 (o rok později byla publikována krátkodobá aktuální projekce s prahem k 1. lednu 2024 a horizontem k 1. lednu 2050). Předpokládá se, že v dlouhodobém horizontu dojde k sblížení demografických charakteristik mezi jednotlivými zeměmi, což je součástí konvergenčního scénáře. Eurostat vedle základní varianty testuje citlivost výsledků na změny jednotlivých demografických parametrů. Mezi nejčastější scénáře patří snížená plodnost, která předpokládá, že míry plodnosti budou každý rok o 20 % nižší než v základní variantě, snížená úmrtnost, jež předpokládá rychlejší pokles úmrtnosti a zvýšení

střední délky života o přibližně 2 roky do roku 2070, snížená migrace, kdy je migrační saldo každý rok o 33 % nižší než v základní variantě, a zvýšená migrace, která počítá s migračním saldem o 33 % vyšším než v základní variantě. Dále je simulován hypotetický scénář nulové migrace. Tyto varianty poskytují vládám a dalším institucím flexibilní rámec pro plánování politických strategií. Krátká projekce s prahem 1. 1. 2024 je vypracována pouze ve dvou variantách: v základní variantě a ve variantě bez migrace.

Organizace spojených národů poskytuje demografické projekce s celosvětovým záběrem, zahrnující jednotlivé státy, kontinenty, ekonomické skupiny zemí podle hrubého domácího produktu a další regionální úrovně. Organizace spojených národů nabízí projekce založené především na různých variantách vývoje plodnosti a její variabilitě. Mezi hlavní varianty patří základní varianta, která představuje střední odhad budoucího vývoje populace, a simulace nízké a vysoké plodnosti, které zohledňují možný pokles nebo zvýšení plodnosti. OSN také používá stochastickou metodiku projekce, která nabízí intervaly spolehlivosti, jež reflektují pravděpodobnou variabilitu výsledků a jejich možné rozpětí.

Obě organizace, Eurostat a OSN, využívají sofistikované metodiky k poskytování přesných a použitelných nástrojů pro analýzu demografických trendů. Zatímco Eurostat se zaměřuje na detailní analýzu evropských zemí, OSN nabízí globální perspektivu. Projekce OSN zahrnují více variant a často kladou důraz na pravděpodobnostní přístupy, zatímco Eurostat se soustředí na scénářové analýzy konkrétních změn v jednotlivých demografických parametrech.

Závěrem kapitoly je možné shrnout, že je vidět poměrně odlišná metodika tvorby projekcí u různých statistických institucí. Existují nicméně styčné body společné pro všechny uvedené projekce, jako je využívání kohortně komponentní metody, či tvorba rozličných scénářů. Odlišnosti v metodice a výsledcích jsou tvořeny z velké části právě v rozličnosti volby scénářů. Hlavní rámec tvorby demografických projekcí tedy i u těchto významných institucí následuje skutečnosti, se kterými čtenáře seznámila tato publikace. Tvorba scénářů je následně vždy závislá na konkrétní situaci a rovněž na tvůrčím myšlení autora konkrétní projekce, přičemž toto se dá očekávat, že bude dále platné i do budoucna bez ohledu na další úroveň rozvoje výpočetního softwaru. Populační chování se nikdy nemusí nezbytně vyvíjet podle matematické křivky, která vychází nejlépe například dle kritérií z problematiky časových řad.

Seznam obrázků

Obrázek 1 Výkyvy v počtu narozených a zemřelých	45
Obrázek 2 Výkyvy v počtu migrantů	45
Obrázek 3 Míry plodnosti dle věku v projekci ČSÚ	54
Obrázek 4 Úhrnná plodnost v projekci ČSÚ	55
Obrázek 5 Naděje dožití v projekci ČSÚ	57
Obrázek 6 Saldo migrace v projekci ČSÚ	58

Literatura

- BOX, G. E. P. a G. M. JENKINS. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden Day, 1970.
- BURNHAM, K. a D. ANDERSON. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2. vyd. New York: Springer, 2002. ISBN 978-1-4419-2973-0.
- ČSÚ. *Projekce obyvatelstva České republiky – 2023–2100* [online]. 2023. [cit. 2024-12-07]. URL: <https://csu.gov.cz/produkty/projekce-obyvatelstva-ceske-republiky-2023-2100>
- EUROSTAT. *Population projections, EuroPOP 2023* [online]. 2023. [cit. 2024-12-08]. URL: https://ec.europa.eu/eurostat/cache/metadata/en/proj_23n_esms.htm
- FIALA, T. a J. LANGHAMROVÁ. *Demografická projekce počtu žáků mateřských a základních škol pro malé územní celky* [online]. Konference České demografické společnosti, 2017. [cit. 2024-12-12]. URL: <https://www.czechdemography.cz/res/archive/003/000386.pdf?seek=1495222375>
- GOOD, I. J. Some applications of the singular decomposition of a matrix. *Technometrics*. 1969, roč. 11, č. 4, s. 823–831.
- HON, F. *Statistické metody modelování a odhadu budoucího vývoje plodnosti*. Praha, 2022. Disertační práce. Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky.
- HUBERT, M., ROUSSEEUW, P. J. a S. VERBOVEN. A fast robust method for principal components with applications to chemometrics. *Chemometrics and Intelligent Laboratory Systems*. 2002, roč. 60, č. 1–2, s. 101–111.
- HYNDMAN, R. J. a H. BOOTH. Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*. 2008, roč. 24, č. 3, s. 323–342.
- HYNDMAN, R. J., ATHANASOPOULOS, G., BERGMEIR, C., CACERES, G., CHHAY, L., KUROPTEV, K., O'HARA-WILD, M., PETROPOULOS, F., RAZBASH, S., WANG, E., YASMEEN, F., GARZA, F., GIROLIMETTO, D., R CORE TEAM, IHAKA, R., REID, D., SHAUB, D., TANG, Y., WANG, X. a Z. ZHOU. *Forecast: Forecasting functions for time series and linear models* [online]. 2024. [cit. 2024-12-02]. R balíček verze 8.16. URL: <https://cran.r-project.org/web/packages/forecast/forecast.pdf>

-
- HYNDMAN, R. J., BOOTH, H., TICKLE, L., MAINDONALD, J., WOOD, S. a R CORE TEAM. *Forecasting mortality, fertility, migration and population data* [online]. 2023. [cit. 2024-12-02]. R balíček verze 2.0. URL: <https://cran.r-project.org/web/packages/demography/demography.pdf>
- HYNDMAN, R. J. a M. S. ULLAH. Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*. 2007, roč. 51, č. 10, s. 4942–4956.
- KUČERA, T. *Regionální populační prognózy: teorie a praxe prognózování vývoje lidských zdrojů v území*. Praha, 1998. Disertační práce. Univerzita Karlova, Přírodovědecká fakulta.
- LANGHAMROVÁ, J. *Změny ve věkové struktuře obyvatelstva a jejich možné důsledky*. Praha, 2008. Habilitační práce. Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky.
- LANGHAMROVÁ, J. a O. ŠIMPACH. *Základy demografie: (materiály ke cvičením)*. Praha: Oeconomica, 2013. ISBN 978-80-2451-956-2.
- LEE, R. D. a L. R. CARTER. Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association*. 1992, roč. 87, č. 419, s. 659–671.
- LI, N. a R. LEE. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*. 2005, roč. 42, č. 3, s. 575–594.
- PAVLÍK, Z., RYCHTAŘÍKOVÁ, J. a A. ŠUBRTOVÁ. *Základy demografie*. Praha: Academia, 1986.
- RAMSAY, J. O. a C. J. DALZELL. Some tools for functional data analysis (with discussion). *Journal of the Royal Statistical Society*. 1991, roč. 53, č. 3, s. 539–572.
- SHANG, H. L. Selection of the optimal Box-Cox transformation parameter for modelling and forecasting age-specific fertility. *Journal of Population Research*. 2015, roč. 32, č. 1, s. 69–79.
- ŠIMPACH, O. *Statistické metody v demografickém prognózování*. Praha, 2016. Disertační práce. Vysoká škola ekonomická v Praze, Fakulta informatiky a statistiky.
- UNITED NATIONS. *World Population Prospects* [online]. 2024. [cit. 2024-12-08]. URL: <https://population.un.org/wpp/>

A

absolutní přírůstek, 16, 17
analytické funkce, 7, 8, 13–16

Č

Český statistický úřad (ČSÚ), 9, 10, 20, 44, 53
čistá migrace (migrační saldo), 23, 26, 27,
33, 57–59

D

deterministické projekce, 29, 37, 41
deterministický model, 40
deterministický přístup, 29
dlouhodobá projekce (prognóza), 5, 7–10,
13, 15, 17, 19, 20, 23–25, 27, 30–33, 36,
55
doba expozice, 14, 15, 17
doba zdvojnásobení, 13, 14
doba zmenšení na polovinu, 14, 16, 58
dokončený věk, 25–28, 34
dosažený věk, 25–28, 47

E

emigrace, 23, 27, 28, 35, 47, 48, 50, 51
emigrant, 23, 24, 25, 27, 28, 33, 47
Eurostat, 9, 25, 53, 58, 59
exponenciální funkce, 13, 15–16, 30
extrapolace, 5, 13, 17, 55

H

hlavní komponenty, 7, 39
horizont projekce, 10

I

imigrace, 23, 25, 26, 27, 28, 33, 46
imigrant, 23, 24, 25, 26, 28, 33, 34, 48
interpolace, 5, 13, 32

K

koherence, 30, 38

komponentní metoda, 5, 8, 9, 19, 22, 23, 25,
27, 29, 33, 37, 48, 49, 53, 58, 59

koncový stav, 14, 17

krátkodobá projekce (prognóza), 9, 15, 17,
58

krok projekce, 9, 49

kvocienty emigrace, 34, 35, 47, 48

L

Lee-Carterův model, 32, 37, 38

lineární funkce, 16, 17

logistická funkce, 8, 13, 15, 17

M

migrace, 5, 8, 9, 11, 13, 23–25, 27, 29, 33, 34,
38, 44, 46, 47, 48, 51, 53, 57–59

migrační saldo (čistá migrace), 23, 26, 27,
33, 57–59

migrant, 23, 26, 28, 34, 57, 58

N

naděje dožití (střední délka života), 11, 20,
30, 32, 55, 56, 57

nízká varianta, 10, 54, 56

O

odvozená projekce, 5, 49–51

Organizace spojených národů (OSN), 59

P

počáteční stav, 14, 16, 18

počet dokončených bytů, 46, 47

počet dožívajících, 30, 31

počet žijících, 19–21, 30

práh projekce, 10, 20

pravděpodobnost úmrtí, 30, 31

prognóza (dlouhodobá projekce), 5, 7–10,
13, 15, 17, 19, 20, 23–25, 27, 29–32, 53

projekce, 7, 9, 10, 13, 19–22

projekční koeficient, 19–21, 48

R

relativní přírůstek, 14, 15–17, 53

S

scénář, 7–10, 13, 19, 24, 29, 31–35, 37, 41, 43, 49–54

scénář migrace, 35, 47, 57, 59

scénář plodnosti, 32

scénář úmrtnosti, 46

sčítání lidu, 10, 53

singulární rozklad matice, 38, 41

specifické míry plodnosti, 11, 32, 46, 50

specifické míry úmrtnosti, 11, 30, 55

rovnávací index, 46

stochastický model, 32, 37–41

stochastický přístup, 37

střednědobá projekce (prognóza), 10, 13, 18

střední délka života (naděje dožití), 11, 20, 30, 32, 55, 56, 57

střední stav, 14, 16, 34

střední varianta, 55, 56, 57

Ú

úhrnná plodnost, 11, 32, 33, 54

úmrtnostní tabulky, 20, 21, 37, 46, 55

V

varianty projekce, 50

věková struktura, 5, 7, 21, 32, 33, 48, 57

věkový interval, 8, 10, 15, 16, 20, 22

vícetavová projekce, 24, 50

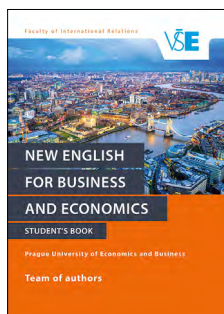
vnitřní míra růstu, 15

vysoká varianta, 54, 56, 58



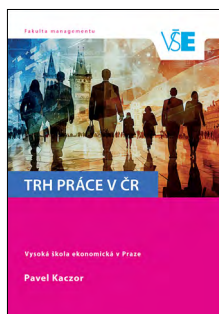
Z produkce Nakladatelství Oeconomica

více informací na <https://oeconomica.vse.cz/>



**Team of authors:
New English for
Business and
Economics.
Student's Book**

ISBN 978-80-245-2427-6,
VŠE, třetí dotisk
1. vydání, 244 stran,
274 Kč



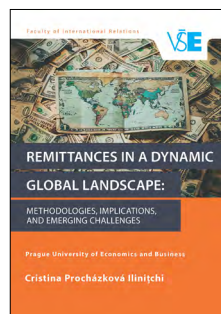
**Pavel Kaczor:
Trh práce v ČR**

ISBN 978-80-245-2513-6,
VŠE, 1. vydání, 216 stran,
401 Kč



**Hana Řezanková:
Statistické metody
se zaměřením na
kategoriální data**

ISBN 978-80-245-2521-1,
VŠE, 1. vydání
v elektronické podobě,
210 stran, zdarma ke
stažení (e-kniha)



**Cristina Procházková
Ilinițchi: Remittances
in a Dynamic Global
Landscape**

ISBN 978-80-245-2511-2,
VŠE, 1. vydání, 238 stran,
496 Kč

Název	Teorie demografických projekcí
Autoři	Ing. Filip Hon, Ph.D. RNDr. Tomáš Fiala, CSc. doc. Ing. Jitka Langhamrová, CSc.
Vydavatel	Vysoká škola ekonomická v Praze Nakladatelství Oeconomica
Vydání	1. vydání v elektronické podobě
Jazyková a redakční úprava	Mgr. Ludmila Doudová
Grafický návrh	Daniel Hamerník, DiS.
Počet stran	65
DTP	Vysoká škola ekonomická v Praze Nakladatelství Oeconomica
Doporučená cena	Zdarma ke stažení

ISBN 978-80-245-2552-5