



# *Анализа на демографски животни табели*

Статистичко моделирање

Изработила: Симона Ристовска

Ментор: Проф. Билјана Тојтовска Рибарски



# Вовед

Животните табели претставуваат статистички инструменти кои ни овозможуваат да ја прикажеме смртноста и преживувањето во една популација преку сите возрасти. Тие служат како основа за формирање на различни демографски индикатори. За подобро да го објасниме ова ќе земеме една популација од 100 луѓе кои сите умираат во рок од 6 години (од 0-та година до крајот на 5-та година):

	A	B	C	D
1	age	qx	lx	dx
2	0	0.280	100	28
3	1	0.125	72	9
4	2	0.111	63	7
5	3	0.143	56	8
6	4	0.313	48	15
7	5	1.000	33	33

Променливата на возраста  $x$  го претставува интервалот на возраста од  $x$  до  $x+1$

$lx$  го претставува бројот на луѓе кои преживеале до  $x$ -годишна возраст.

$dx$  бројот на луѓе кои починале помеѓу  $x$  и  $x+1$  година

А  $qx$  е стапката на смртност, т.е. процентот на луѓе кои починале помеѓу  $x$  и  $x+1$  година,  $qx=dx/lx$

# Продолжение - Очекуван животен век

**Очекуван животен век (Life expectancy):** Просечниот број на дополнителни години кои се очекува да живее лице на дадена возраст, врз основа на досегашните податоци за смртноста.

	A	B	C	D	E	F	
1	age	qx	lx	dx	Lx	Tx	ex
2	0	0.280	100	28	86	322	3.2
3	1	0.125	72	9	67.5	236	3.3
4	2	0.111	63	7	59.5	168.5	2.7
5	3	0.143	56	8	52	109	1.9
6	4	0.313	48	15	40.5	57	1.2
7	5	1.000	33	33	16.5	16.5	0.5

Како можеме да го пресметаме животниот век од тие бројки? Постојат два еквивалентни начини за толкување на очекуваниот животен век:

- **просечниот (просечен) број на години живеени во дадена популација** или
- **просечната (средна) возраст на смртта.**

**Lx**, вкупниот број на живеени години над возраста  $x$  или вкупниот број на години живеени во популацијата за луѓе на возраст меѓу  $x$  и  $x+1$ .

Од **Lx** можеме да го изведеме **Tx**, вкупниот број на живеени години (**Tx**) од возраста  $x$  до крајот, па едноставно збирот на сите **Lx** од возраста  $x$  до 5-тата година во нашата табела (последната опција).

Последниот чекор е да се земе **Tx** и да се подели со бројот на луѓе на почетокот на секој период, **lx**. Ние го правиме тоа на секоја возраст, за да го добиеме **очекуваниот животен век на возраст  $x$**  (означен како **ex**).



# Продолжение - $m_x$



- **возрасно-специфична стапка на смртност или централна стапка на смртност.**
- **колку често се случуваат смртни случаи за таа популација-година.**

Замислете популација од 100 000 новородени. На секој роденден  $x$ , дел од нив ќе починат во текот на таа година, а останатите се живи. Вредноста  $m_x$  ни кажува колку брзо луѓето умираат на возраст  $x$ .

Во годината од возраст  $x$  до  $x+1$ , сите луѓе придонесуваат „години“ во одреден збир. Тој збир е  $L_x$ . Бројот на смртни случаи во истата година е  $d_x$ . Тогаш:  $m_x = d_x / L_x$ , односно број на смртни случаи поделен со вкупните години живеење во тој возрастен интервал.

Тоа не е „веројатноста“ за умирање (тоа би било  $q_x$ ), но многу се приближува кон  $q_x$  кога стапките се ниски. „Ако  $m_{80} = 0.05$ , тоа значи во просек 5 смртни случаи на 100 лица меѓу 80-годишници.“



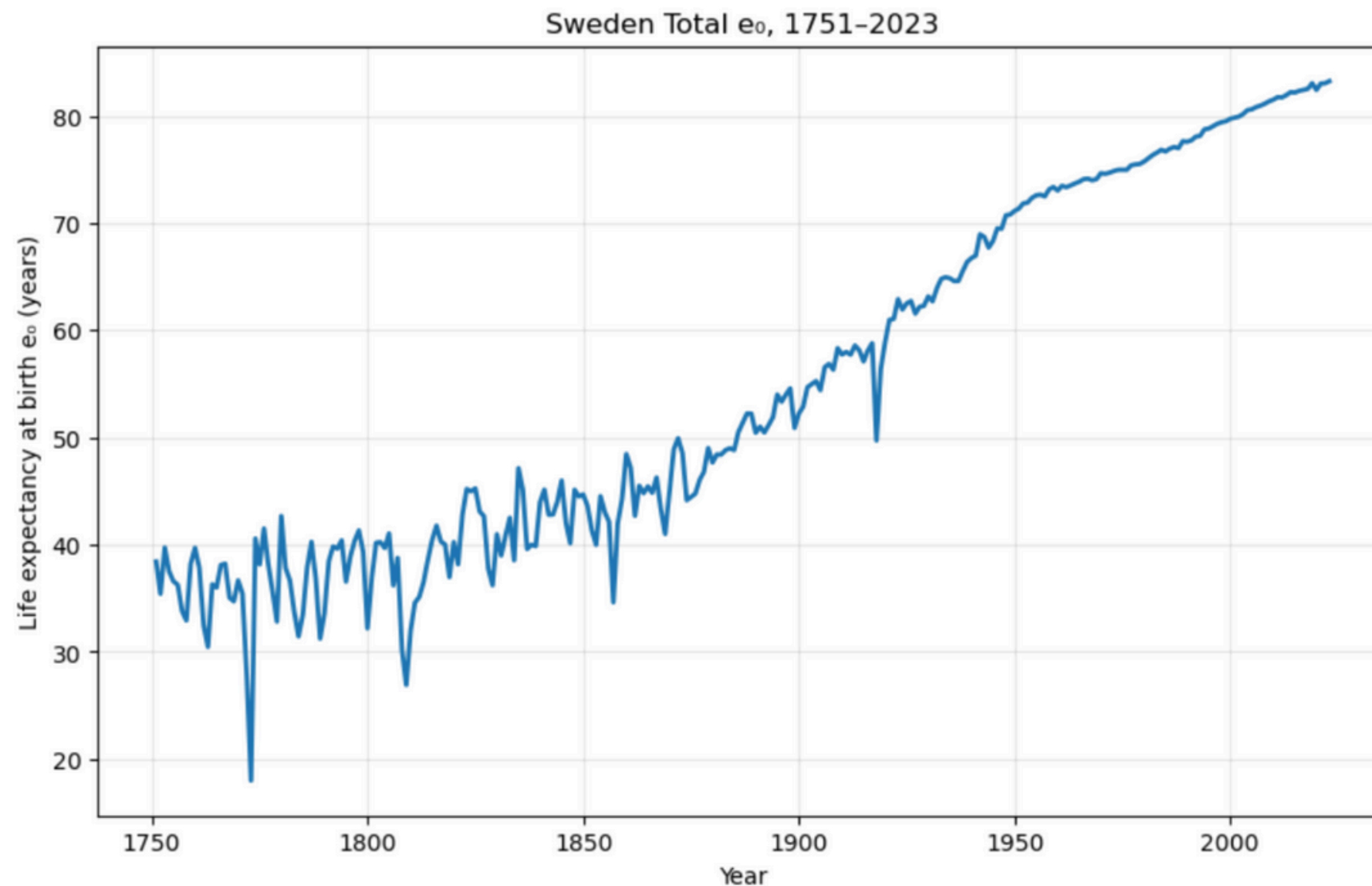
# Избор на држава



Целта на овој проект е да се направат најразлични визуелизации кои ќе ни дадат некаква претстава за демографијата на одредена популација. Како држава за која сметам дека најдобро би можеле да се направат било какви анализи ја избрав **Шведска**, причината за тоа беше бидејќи за неа имаше **најмногу достапни податоци** кои датираат уште од 1751, па се до 2023 година.



# Прва и наједонставна визуелизација



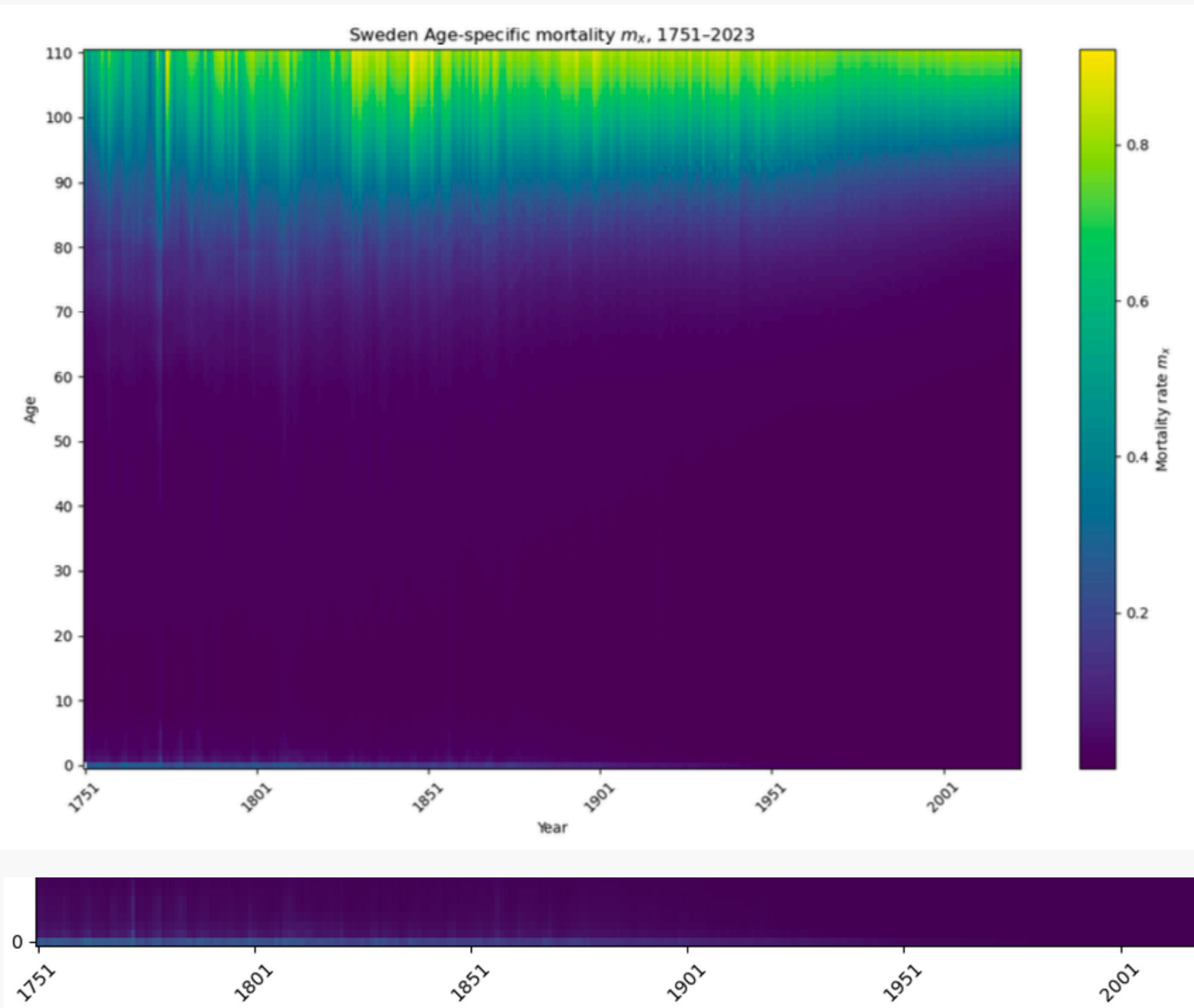
**Очекувана должина на животот на едно новороденче ( $e_0$ ) во Шведска, 1751-2023.**

На хоризонталната оска е годината (**Year**), а по вертикалната е  **$e_x$**  (очекуваниот број на години живот во моментот на раѓање).

Првичните години сведочат за високата ранлива фаза во која секое епидемско или воено нарушување веднаш се рефлектира во поголеми падови, додека современите периоди покажуваат стабилно и мирно зголемување благодарение на напредокот во санитарната инфраструктура, родителската здравствена нега и општата социјална заштита.



# Смртност по возраст $m_x$ од 1751 до 2023



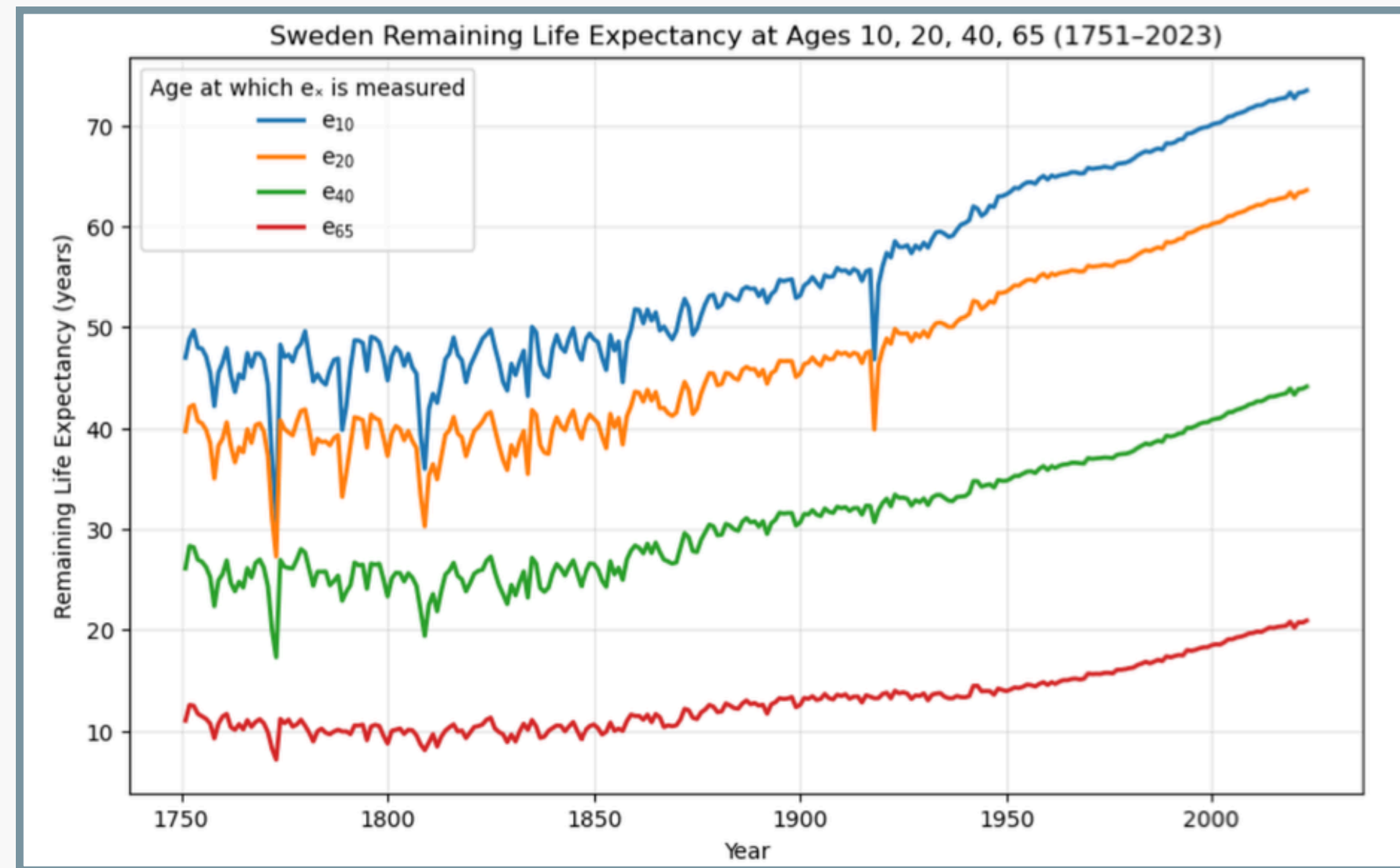
Во оваа визуализација (т.н. heatmap) прикажуваме вредности на централната стапка на смртност  $m_x$  за секоја година од 1751 до 2023 (хоризонтална оска) и за секоја возраст од 0 до 110 (вертикална оска). Секоја селектирана вредност  $m_x$  (веројатна стапка на смртност за одредена возраст) е прикажана со бои: **потемни** (виолетови) нијанси означуваат ниски стапки на смртност, додека **потоплите** (жолто-зелени) нијанси покажуваат повисоки стапки.

- Инфантилна смртност (Age = 0)
- Детска смртност (Age 1-10)
- Младинска и возрасна смртност (Age 10-50)
- Средновечна до постара смртност (Age 50-80)
- Постари возрасти (Age 80-110)



# Очекуван животен век за 10, 20, 40 и 65 години (1751–2023)

- **1751–1850:** Силни годишни колебања поради епидемии и војни. Поголеми падови се јасни кај помладите ( $e_{10}$ ,  $e_{20}$ ), но влијае и кај  $e_{40}$  и  $e_{65}$ .
- **1850–1918:** Постепен пораст на сите  $e_x$ . Неколку кратки падови (најизразен околу 1918, „Spanish flu“).
- **1950–1980:** По Втората светска војна следува стабилен раст.  $e_{10}$  се искачува од 55 на 65 години,  $e_{20}$  од 47 на 55,  $e_{40}$  од 30 на 32–33 и  $e_{65}$  од 13 на 14.
- **1980–2023:** Континуиран раст со намалени годишни осцилации. До 2023,  $e_{10} \approx 75$ ,  $e_{20} \approx 64$ ,  $e_{40} \approx 44$ ,  $e_{65} \approx 21$ .

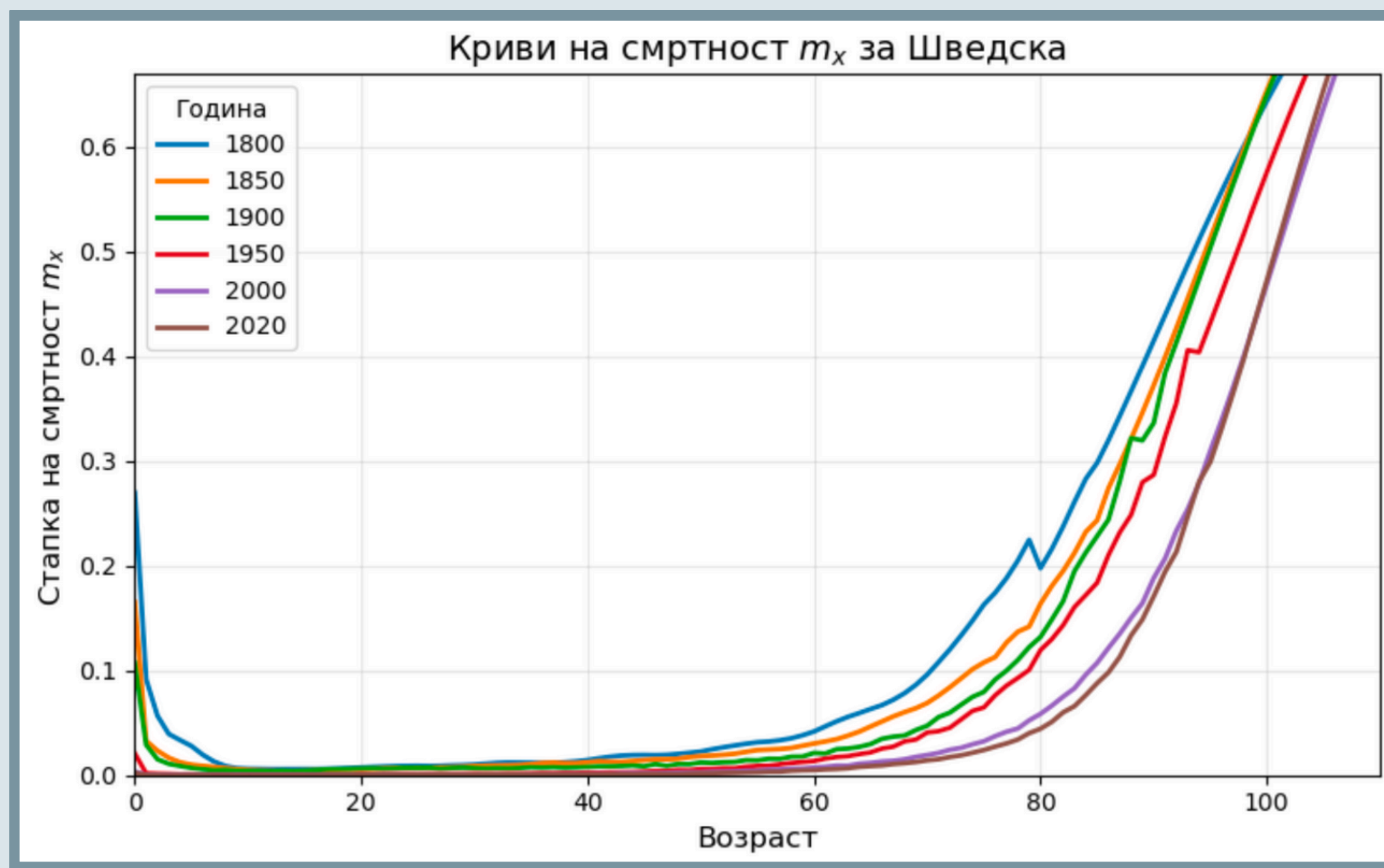




# Криви на смртност $m_x$

На овој линиски дијаграм се прикажани возраст-специфични стапки на смртност  $m_x$  (од 0 до 110 години) во шест фундаментални години (1800, 1850, 1900, 1950, 2000, 2020).

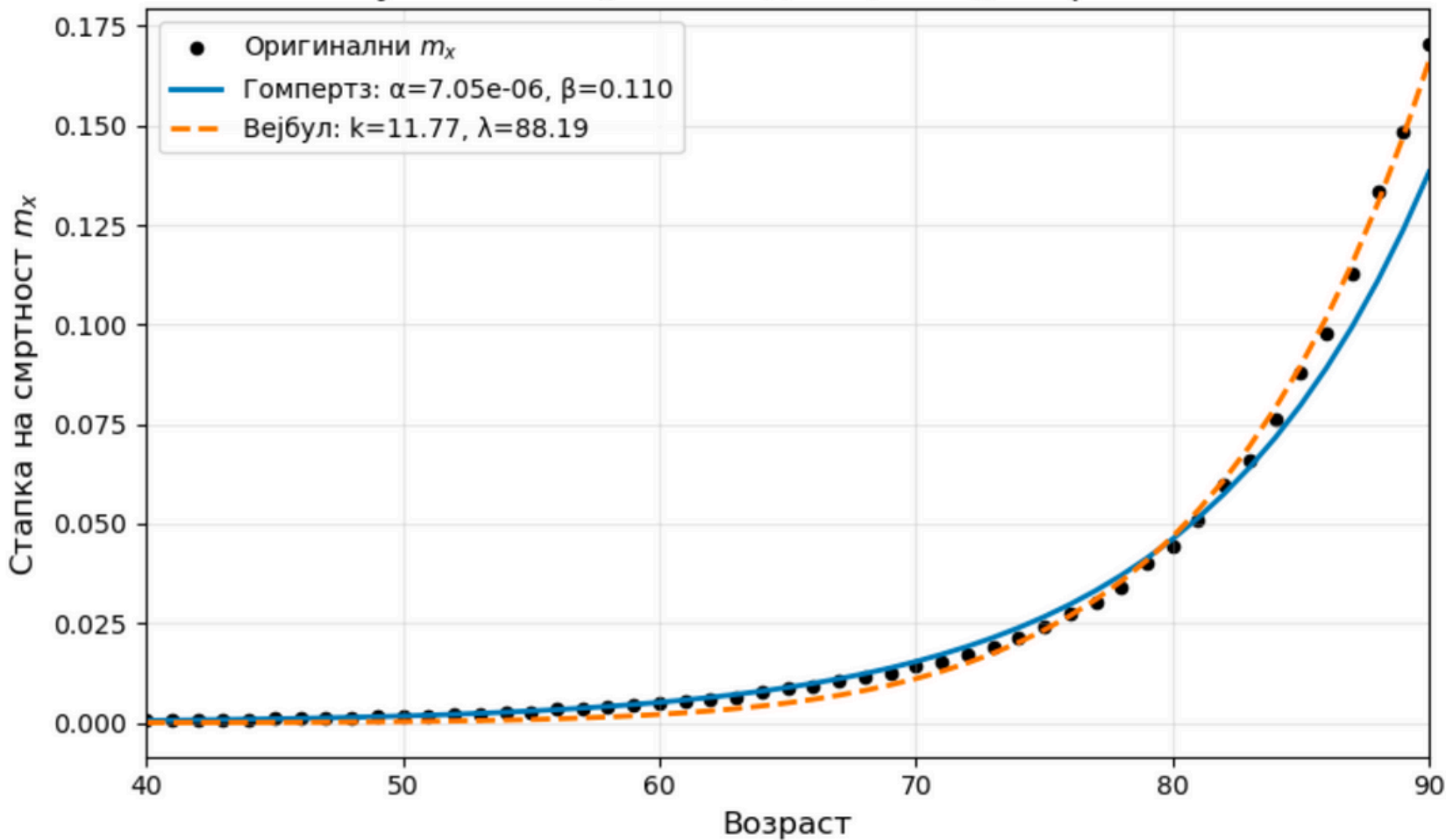
- 1800 и 1850
- 1900 и 1950
- 2000 и 2020



Во текот на 220 години, Шведска поминува од многу високи смртни стапки на новороднчиња и младински смртни стапки кон масивно намалување на смртноста на сите возрасти до средна возраст. Од средината на XX век и особено до 2020, едвај има смртност под 50 години, додека релативната тежина на старите возрасти се преместува кон 80–100 и понатака.

# Распределби

Фитнување на  $m_x$  за Шведска (2020), возраста 40-90



Црните точки се оригиналните вредности на  $m_x$  за возрасниот интервал 40-90 во 2020.

Сината линија е Гомпертз-фит ( $m_x = \alpha e^{\beta x}$ ) со параметри  $\alpha = 7.05 \times 10^{-6}$  и  $\beta = 0.110$ .

Портокаловата непрекината линија е Вејбул-фит  $m_x = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1}$   $k = 11.77$  и  $\lambda = 88.19$ .

- Гомперц претпоставува дека смртноста расте експоненцијално со возраста, а ние ја прилагодуваме со користење  $\ln m_x$  на возраст.
- Веибул презема форма на power-law, и користи директно намалување на квадратните грешки.
- И двата модели добро ја следат смртноста од 40 до 90 - со единечна стапка на Гомперц  $\beta = 0.110$  што го доловува удвојувањето на ризикот приближно на секои 6-7 години, и параметарот на обликот на Веибул  $k > 1$ , што го потврдува забрзувачкиот наклон во напредната возраст.

# Мерки за оценка на дистрибуциите

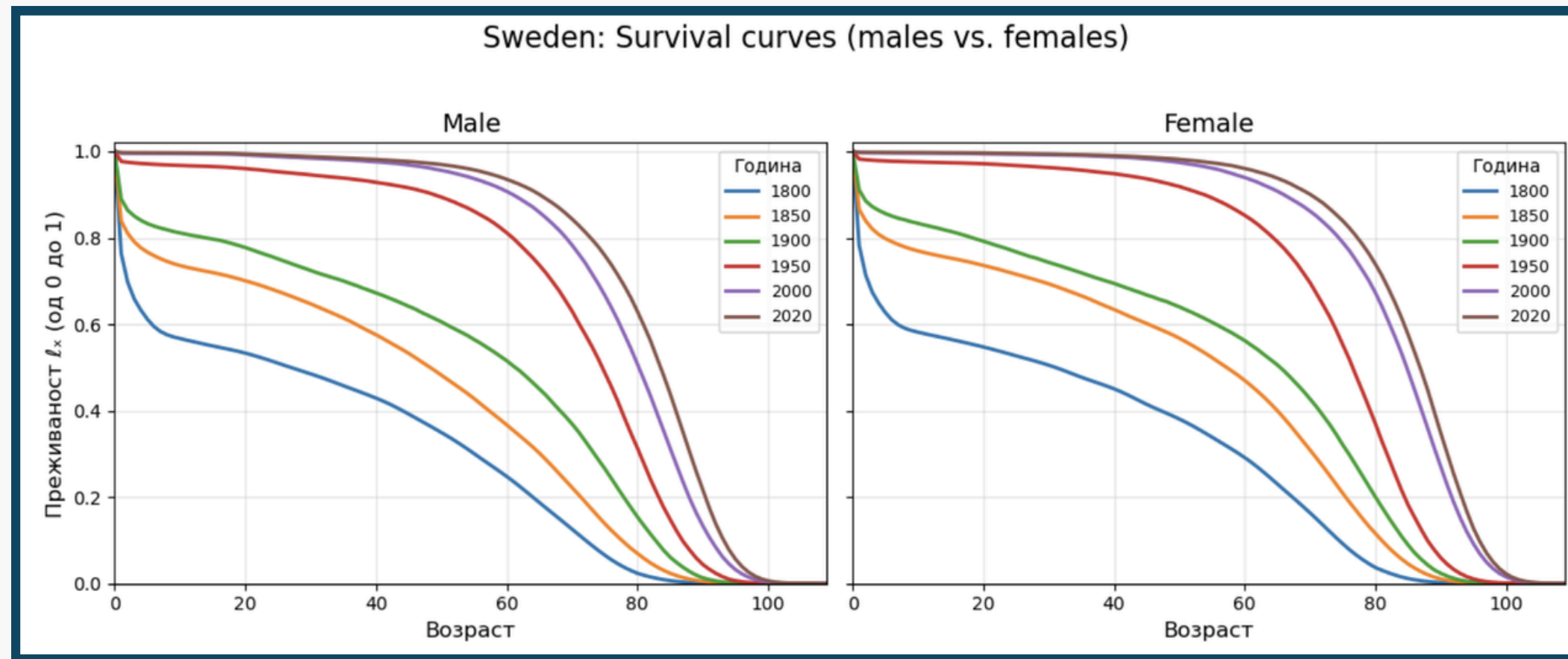


Model	Parameters	RMSE	R2	AIC
Gompertz	2	0.00709735	0.969988	-500.699
Weibull	2	0.00233717	0.996746	-613.999

- **RMSE (Root Mean Squared Error):** колку просечно се разликува предвидената стапка од набљудуваната. Помала вредност = подобар фит.
- **R<sup>2</sup> (Coefficient of Determination):** колкав дел од варијацијата на податоците моделот успева да ја објасни. Поголема вредност (блиску до 1) значи дека моделот добро го следи трендот.
- **AIC (Akaike Information Criterion):** балансира точност и комплексност. Помала вредност = подобар компромис.



# Криви на преживување за мажи и жени

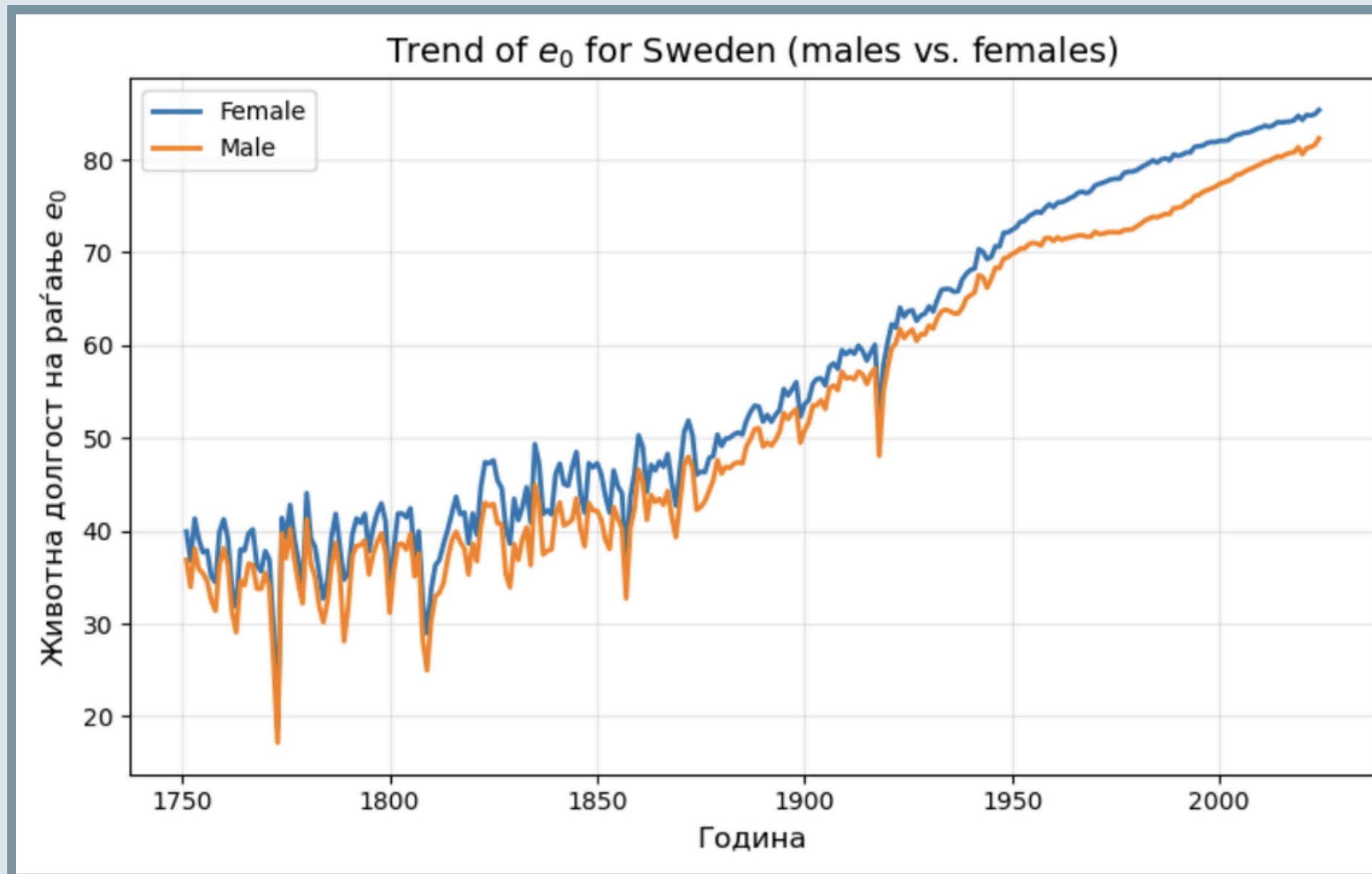


- **Поместување надесно:** Со текот на времето, секоја крива се поместува надесно и издигнува нагоре — тоа значи дека луѓето почесто достигнуваат повисоки години.
- **Правоаголна тенденција (“rectangularization”):** За современите години (2000–2020), кривите се речиси правоаголни: релативно висок  $\ell_x$  до 80–85, па брзо паѓање кон нула. Тоа е индикатор за малку варијабилна возраст на смрт и подобро здравје во староста.

**Животен век:** Жените (десен панел) секогаш имаат повисоко  $\ell_x$  отколку мажите (лев панел) во сите историски точки. Во 2020, на 80-г.  $\ell_x \approx 0.7$  за жени vs.  $\ell_x \approx 0.6$  за мажи.

**Кривина:** Кривата за жени е неколку проценти повисока до подоцнежна возраст, што го рефлектира нивниот повисок животен век.

# Животната должина при раѓање $e_0$ , според пол

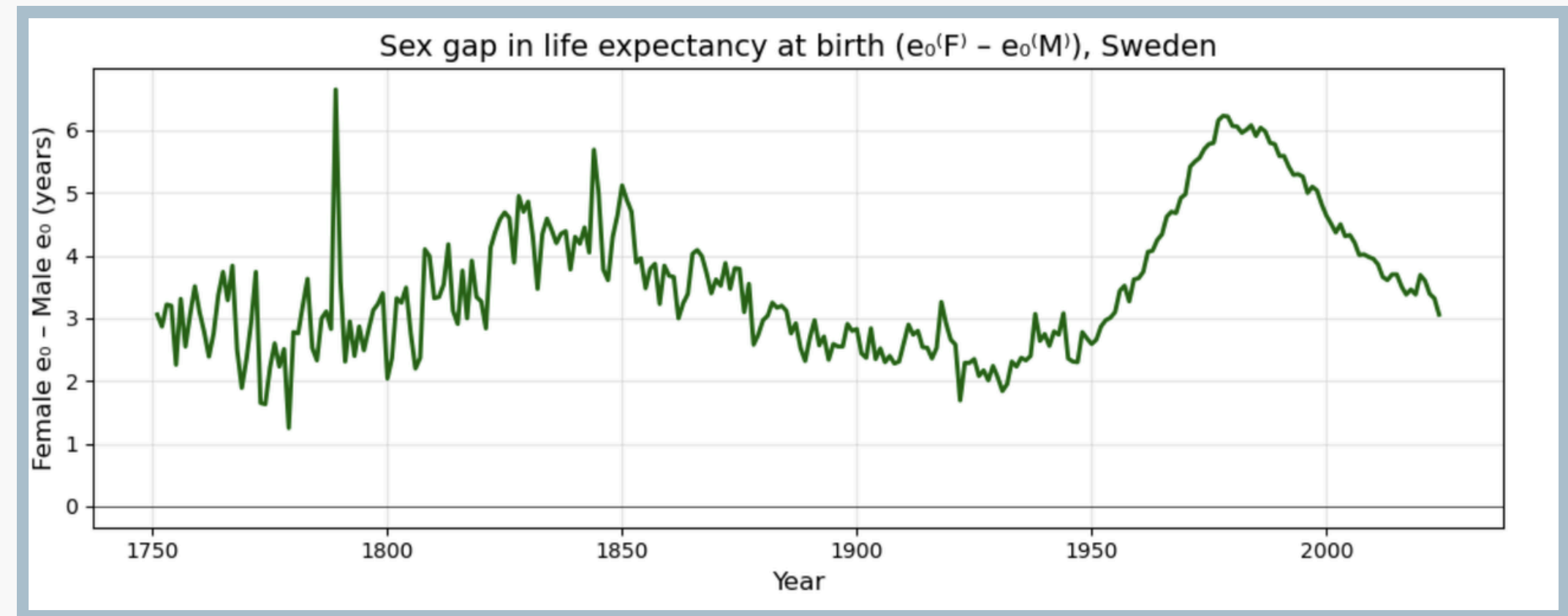


- Продолжување на животниот век: Од 40 г (18 век) до 85 г (денес).
- Жените водат: По секоја ера, жените живеат барем неколку години повеќе.
- Историски „удари“: Војни и епидемии го „дестабилизираа“  $e_0$ , но со напредокот на јавното здравје и технологијата тие стануваат помалку изразени.



# Разликата во животна должина при раѓање помеѓу жени и мажи ( $e_o^F - e_o^M$ )

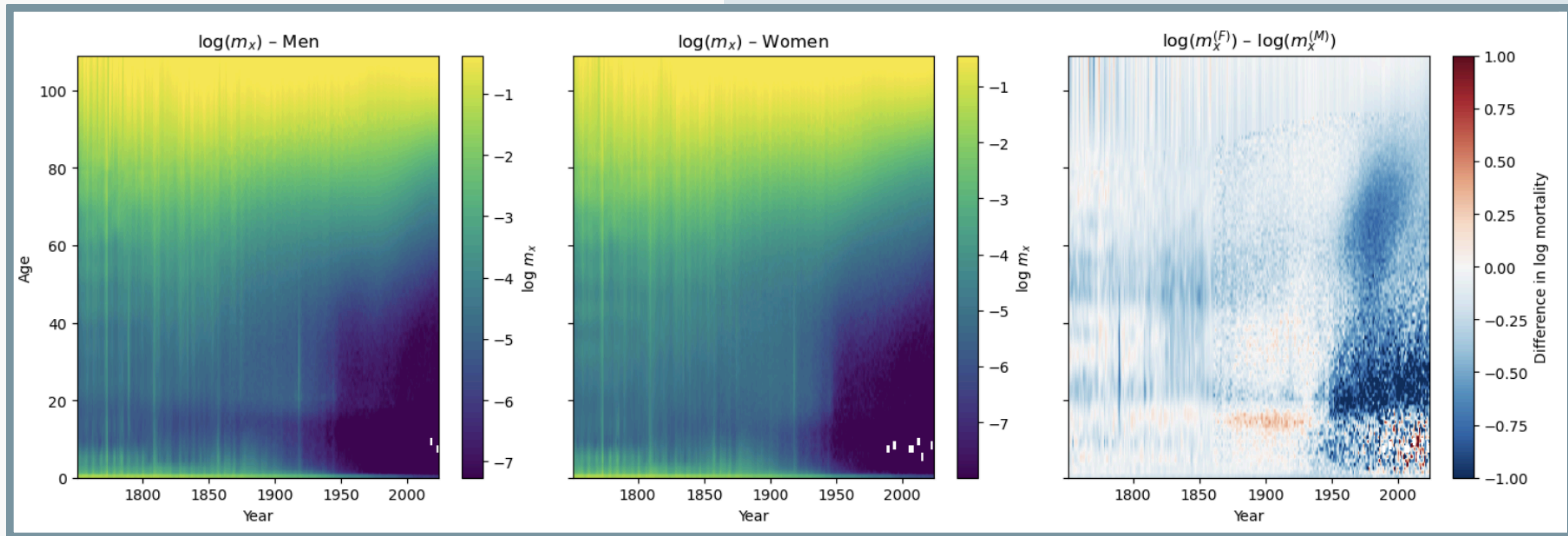
- **Вредности > 0:** Жените живеат во просек повеќе години од мажите.
- **Вредности = 0:** Животниот век е ист. (Никогаш не се пресече)
- **Вредности < 0:** (Не се јавува во податоците) мажите би живееле повеќе.



- Стабилна „предност“ на жените: Во цел период +2...+6 г.
- Историски причини: Епидемии/војни → силни флукутации; модерна медицина → намалување на осцилациите.
- Современ тренд: По 1980, „просторот“ се стеснува, што покажува сè почеста еднаквост во начинот на однесување и пристапот до здравствени технологии.

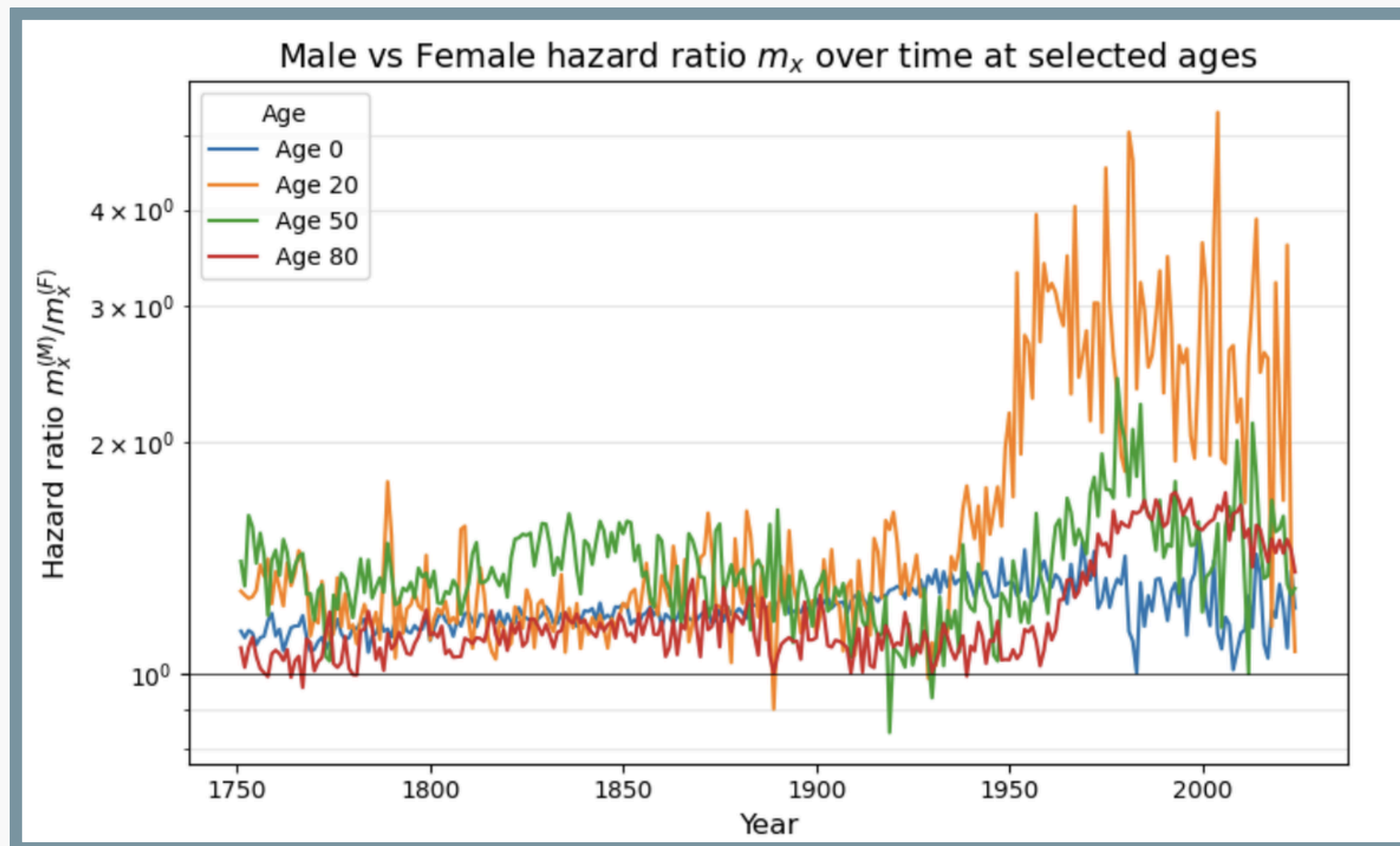


# Смртност по возраст $m_x$ според пол



- Мажите традиционално имаат повисока смртност во сите возрасти, најизразено во зрела возраст (35–65 г).
- Жените го предводат падот на смртноста уште од почетокот на 19 век, па затоа денес ја имаат и поголемата животна должина.
- Главните историски „удари“ (епидемии, војни) се видливи како коридори преку цела возрасна дистрибуција, често со полови разлики во отпорноста.

# Односот на хазард-стапките ( $m_x$ ) мажи/жени низ времето за четири возрасни категории



- Мажите секогаш имале повисока смртност, но пред 20. век разликата била скромна ( $\approx 10$ – $50$  % повисока).
- „Пушачката ера“ и поствоените децении (1950–1980) драматски го зголемиле половиот јаз, особено кај младите и средношколци – ризични навики и професионални опасности го зголемиле односот кон 2–4 пати.
- Денес: јавното здравје, контрола на пушењето и подобрена кардиоваскуларна заштита го намалија ова растојание, па денес односот кај млади и средна возраст е помеѓу 1.5 и 2.0, а кај бебиња и постари лица блиску до 1



*Ви благодарам на  
вниманието!*

*Прашања?*

