

Modelli stocastici per le proiezioni di longevità

Valeria D'Amato

Dipartimento di Scienze Economiche e Statistiche

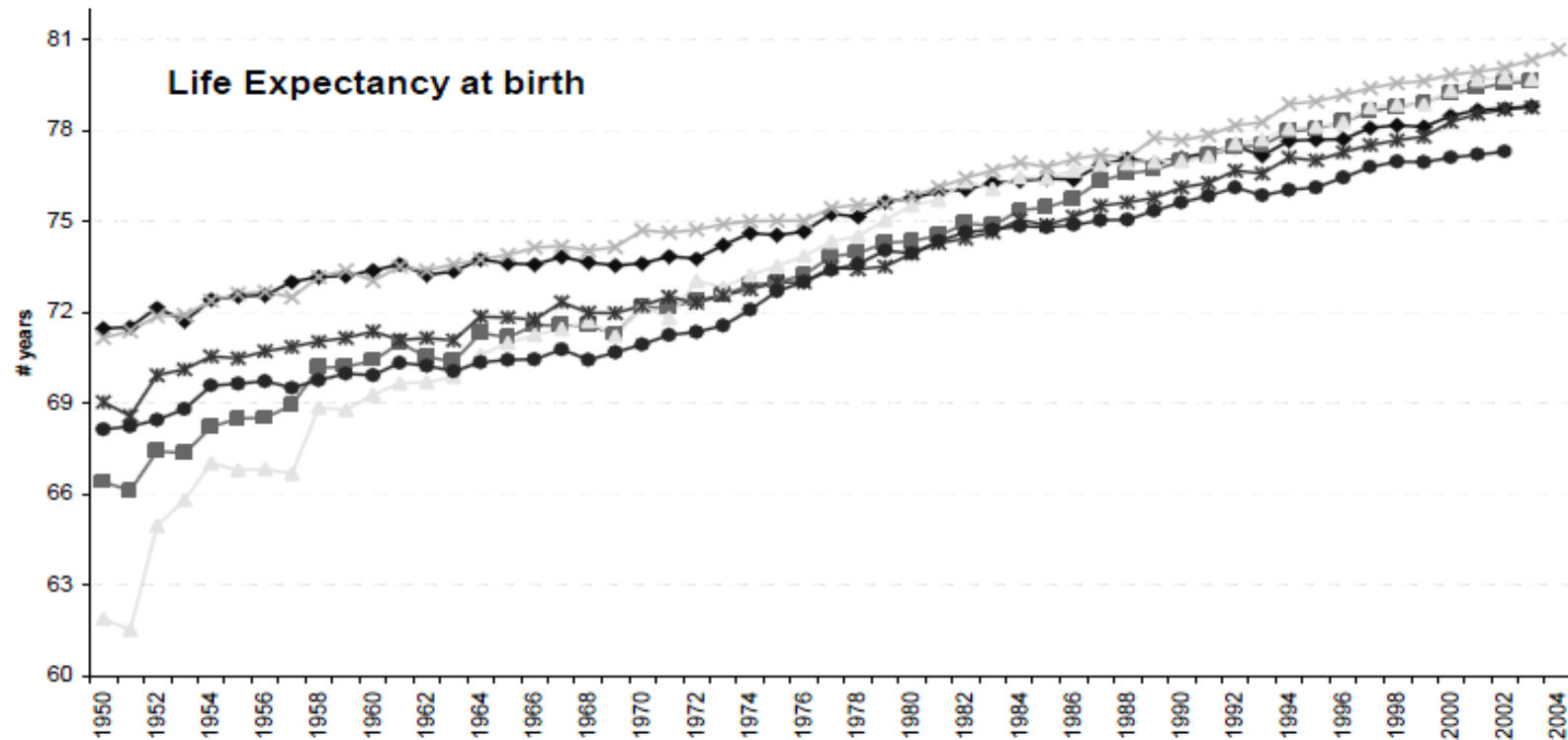
Università degli Studi di Salerno



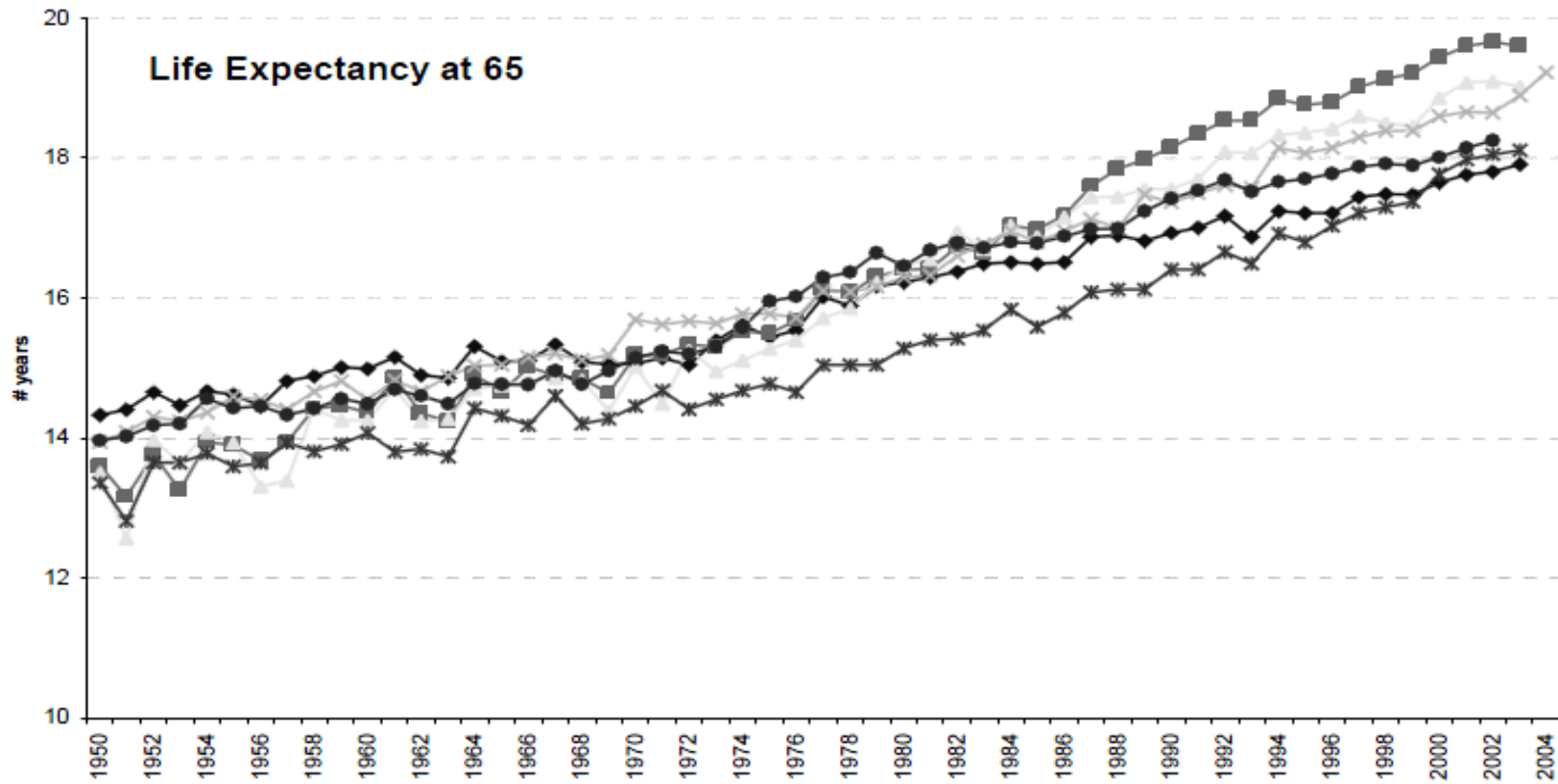
MOTIVATION

- Il secolo scorso è stato caratterizzato da rilevanti progressi in materia di riduzione del rischio di mortalità e di conseguente incremento della sopravvivenza della popolazione in termini di aumento della durata media di vita.
- L'attenzione rivolta verso nuove conoscenze e scoperte mediche, verso una maggiore prevenzione, all'adozione di stili di vita sempre più salutari ha, non solo migliorato il rapporto tra l'uomo e la vita, la morte, la malattia, ma ha anche reso il raggiungimento di età avanzate un traguardo sempre più accessibile.
- Di estrema importanza ai fini di questa ricerca, si rivela il concetto di **speranza di vita**, inteso quale indicatore statistico che esprime il numero medio di anni di vita di un soggetto osservato a partire da una certa età, all'interno della popolazione indicizzata.

MOTIVATION



MOTIVATION



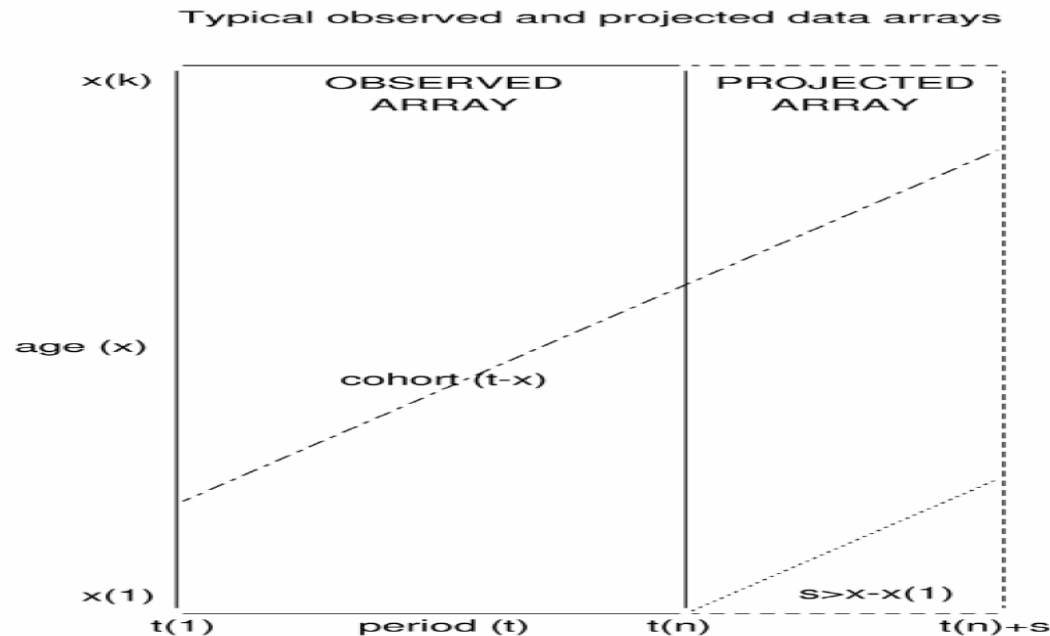
MOTIVATION

- Le basi attuariali per il calcolo delle prestazioni e delle controprestazioni si rinvencono nelle tavole di mortalità di riferimento, che si modificano in modo sostanziale nel corso degli anni a causa delle evoluzioni demografiche.
- Gli scostamenti tra le previsioni ex-ante e le valutazioni ex-post sui fenomeni di rischio di mortalità di una popolazione e di aspettativa di vita a una data età x , possono essere pregiudizievoli per la sopravvivenza di una impresa assicurativa.

MOTIVATION

- Le tavole proiettate esplicitano la loro funzione nell'estrapolazione degli andamenti di mortalità osservati dalle esperienze precedenti, e sulla base di questi prevedere quelli futuri.

Fig 1



MOTIVATION

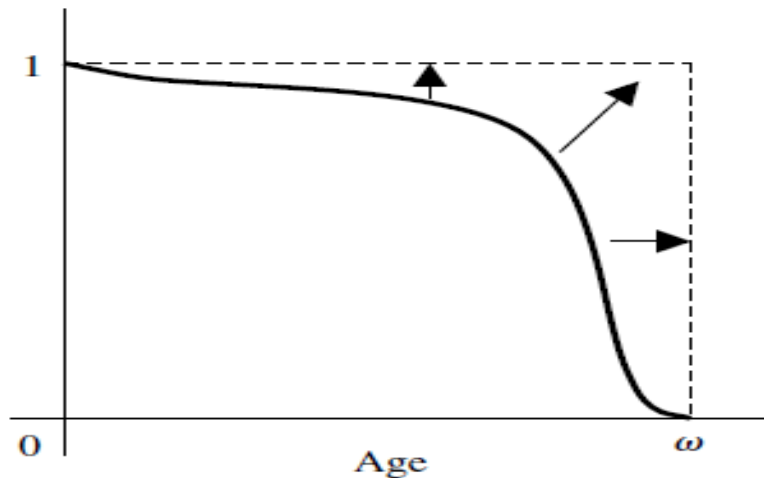
GRANDEZZE D'INTERESSE

- tasso di mortalità annuale è la percentuale di persone di età x che muoiono durante l'anno t ,
- la probabilità di morte è la percentuale di persone all'età x di morire nell'intervallo d'età $x, x+t$

MOTIVATION

- Fenomeni rilevanti, nonché aspetti significativi nell'ambito dell'esperienza di mortalità negli ultimi decenni riguardano la rettangolarizzazione, l'espansione e un livello più alto e una più larga dispersione delle morti accidentali ad età giovani (la cosiddetta “young mortality hump”)

(a) Rectangularization



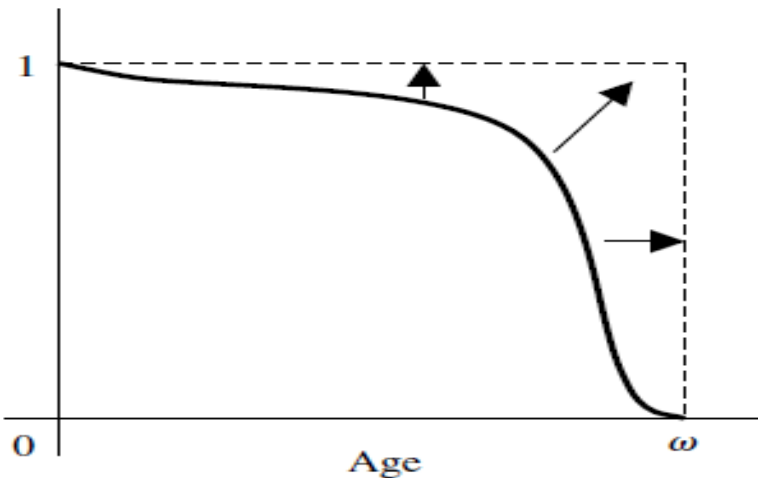
(b) Expansion



MOTIVATION

- la rettangolarizzazione consiste in una maggiore concentrazione della probabilità di decesso attorno ad età avanzate (o meglio al punto di Lexis, il quale tende all'età estrema ω);

(a) Rectangularization



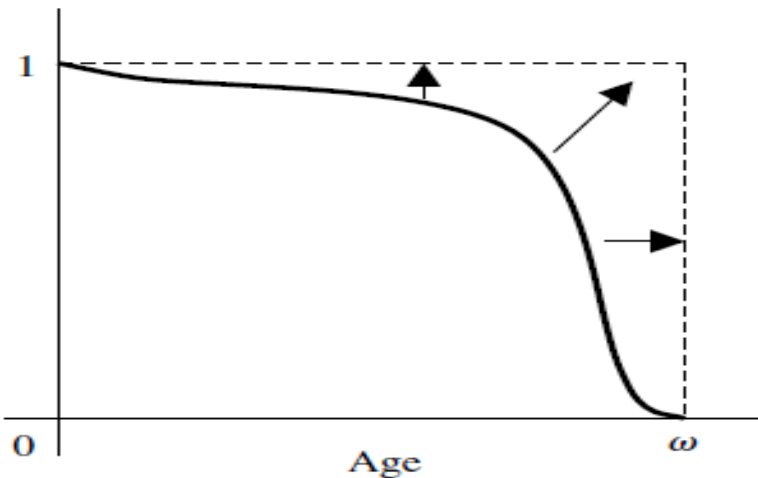
(b) Expansion



MOTIVATION

- l'espansione consiste nello spostamento in avanti, di entità aleatoria, del punto di Lexis, comportando un allungamento della vita, che determina l'incertezza dell'ampiezza della rettangolarizzazione.

(a) Rectangularization



(b) Expansion



MOTIVATION

- un modello dinamico di sopravvivenza può essere rappresentato con una funzione $\Gamma(x, t)$ che può essere formalizzata tramite una matrice le cui righe corrispondono alle età e le colonne agli anni di calendario.

La matrice $\Gamma(x, t) = q_x(t)$ può essere letta secondo:

l'approccio verticale: presenta la tavola di mortalità (*tavola per contemporanei o di periodo*) relativa a soggetti aventi differenti età nell'anno t ;

l'approccio diagonale: configura la tavola di mortalità (*tavola di generazione o di coorte*) relativa a soggetti nati tutti nell'anno t ;

l'approccio orizzontale: rappresenta il *profilo temporale (o trend) della mortalità*, ossia la sua evoluzione per ogni generica età x , riferita a una data popolazione.

MOTIVATION

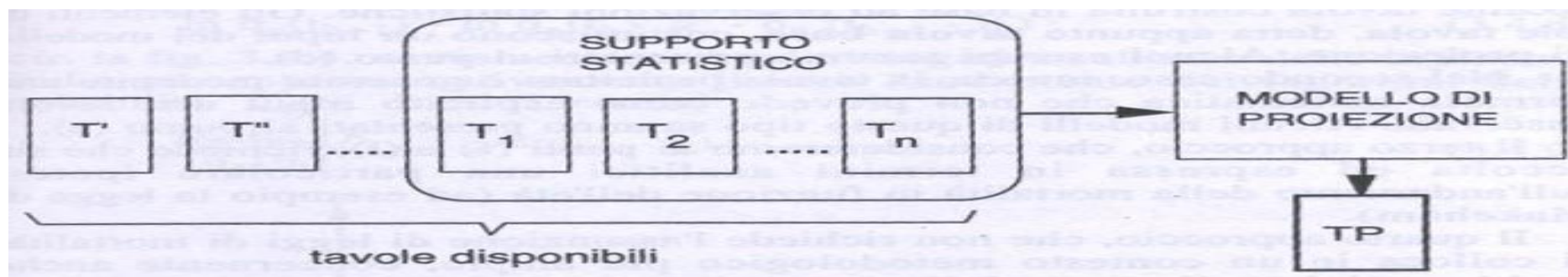
- Per la costruzione di una tavola proiettata si rendono necessarie una serie di scelte:

supporto statistico, ovvero insieme di tavole da impiegare nella procedura di estrapolazione;

modello di proiezione, cioè struttura matematica da utilizzare per la procedura di estrapolazione;

tavola base, alla quale applicare il modello di proiezione.

La tavola proiettata sarà quindi, costruita sulla base di un modello di proiezione, i cui parametri saranno stimati sulla scorta del supporto statistico scelto.



MOTIVATION

- Si hanno a disposizione diverse osservazioni di tipo cross-section, cioè rilevate tutte nello stesso istante temporale, per una data popolazione (ad esempio i pensionati iscritti ad un piano pensionistico).

Ogni osservazione è un vettore che rappresenta lo schema di mortalità per un dato insieme di età x , ove $x = \{x_{min}, x_{min} + 1, \dots, x_{max}\}$. Per un generico anno di calendario t si ha:

$$\{q_x(t)\}_{x \in \mathcal{X}} = \{q_{x_{min}}(t), q_{x_{min}+1}(t), \dots, q_{x_{max}}(t)\}$$

Fissando un limitato numero di anni $\mathcal{T} = \{t_1, t_2, \dots, t_n\}$, il data base di dati a disposizione è rappresentabile mediante una matrice:

$$\{q_x(t)\}_{x \in \mathcal{X}, t \in \mathcal{T}} = \{q_x(t_1), q_x(t_2), \dots, q_x(t_n)\}_{x \in \mathcal{X}}$$

MOTIVATION

- Se valgono le ipotesi che i trend osservati negli anni passati (ovvero negli anni compresi in \mathcal{T}) possano essere interpolati attraverso un'opportuna funzione e che il loro andamento continui anche negli anni a venire, allora la mortalità futura può essere stimata estrapolando il trend stesso.

L'estrapolazione delle probabilità q_x rappresenta un caso particolare dell'approccio orizzontale il quale può anche essere applicato ad altre quantità come ad esempio il mortality odds $\frac{q_x}{p_x}$.

- L'odds è definito come il rapporto tra la probabilità q che si verifichi un certo evento e la probabilità $p = 1 - q$ che lo stesso non si verifichi.

MOTIVATION

- Si può esprimere la relazione tra la probabilità di morte per l'età x nell'anno t' (per esempio $t = t_n$) e quella in un generico anno $t > t'$ come segue:

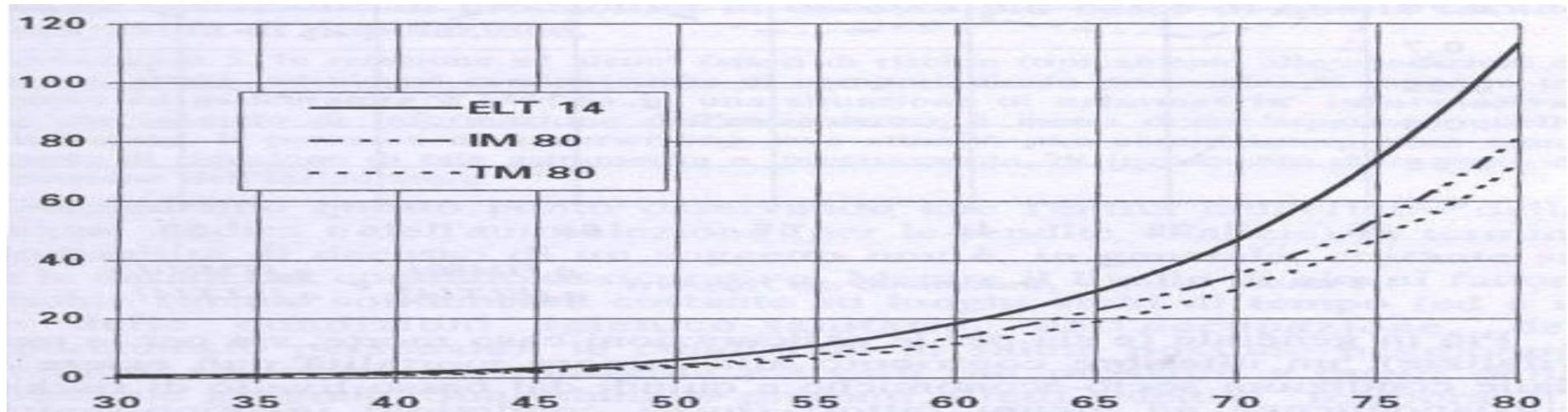
$$q_x(t) = q_x(t')R_x(t - t')$$

La quantità $R_x(t - t')$ è detta “*fattore di riduzione*” all'età x per l'intervallo (t', t) e ci si aspetta che il suo valore sia inferiore all'unità, viste le dinamiche decrescenti delle probabilità di morte.

Una semplificazione può essere ottenuta ipotizzando che il fattore di riduzione non dipenda dall'età $R_x(t - t') = R(t - t')$.

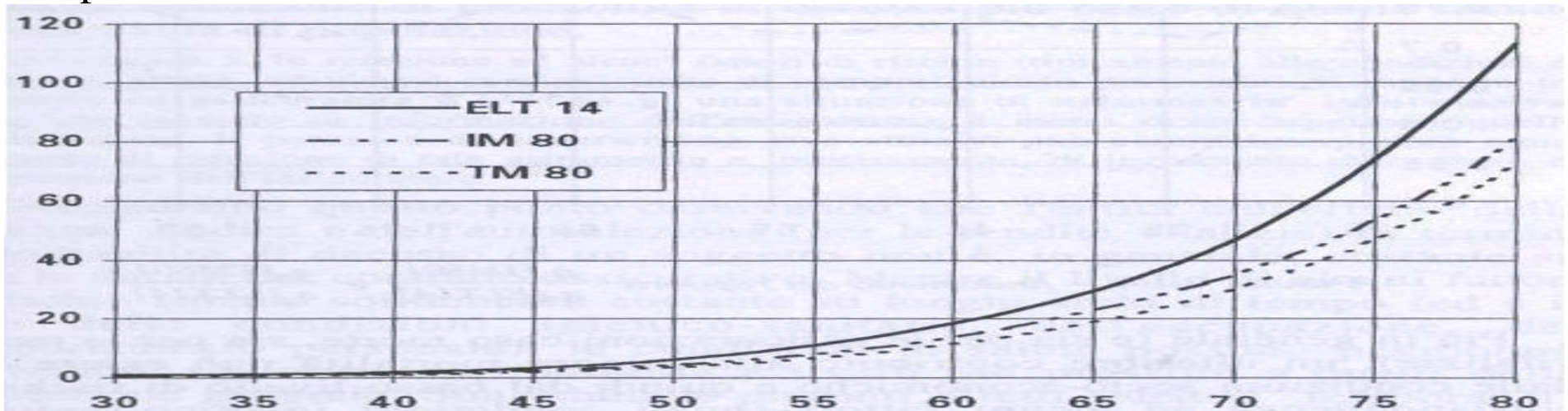
MOTIVATION

- Confronto la mortalità generale di una popolazione con quella degli assicurati (caso vita e caso morte) all'interno della stessa, si ottengono una serie di evidenze.



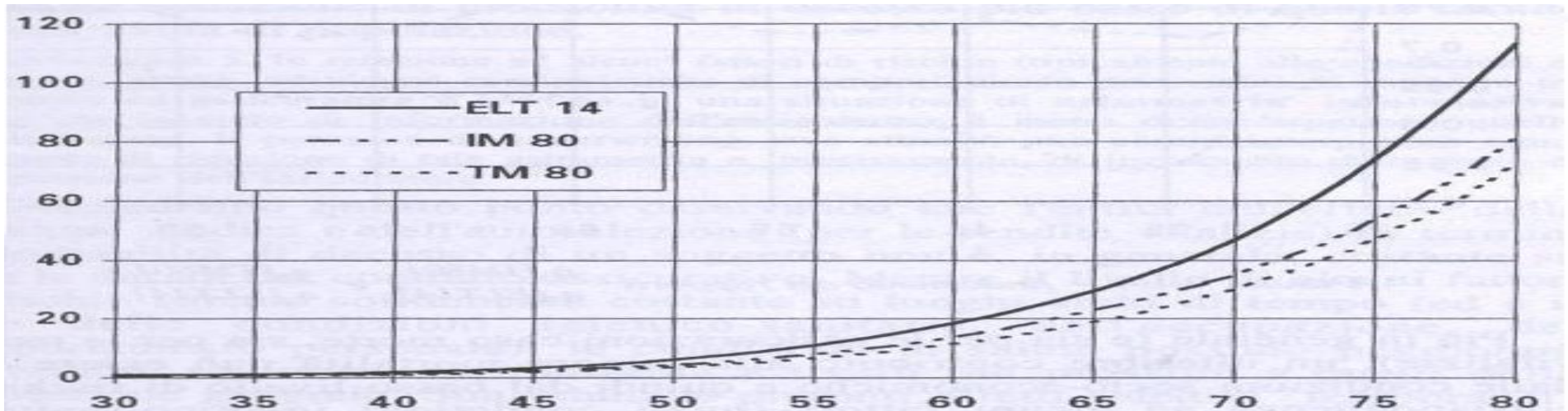
MOTIVATION

- La più marcata differenza si riscontra tra la funzione q_x relativa alla mortalità della popolazione in generale e quelle degli assicurati nelle due diverse configurazioni; ciò si spiega in quanto la popolazione degli assicurati risulta selezionata rispetto alla collettività di appartenenza, si *autoseleziona*. Il fenomeno di *autoselezione*, è riferibile agli assicurati caso vita, i quali contano su di una probabilità di sopravvivenza elevata per essere portati a stipulare un contratto assicurativo.



MOTIVATION

- Il fenomeno speculare dell'*antiselezione* si applica agli assicurati caso morte che presentano delle condizioni salutari inferiori alla media e proprio a causa del loro stato di salute, decidono di assicurarsi



MOTIVATION

- Le principali tipologie di rischio che insistono su di un portafoglio assicurativo e che costituiscono oggetto di studio degli esperti attuariali sono riconducibili al *rischio d'investimento* (o rischio finanziario) ed al *rischio demografico*.
- *rischio d'investimento*: deriva dall'aleatorietà dei tassi di rendimento degli investimenti effettuati dalla compagnia assicurativa rispetto al tasso tecnico garantito agli assicurati e presenta natura sistematica impattando su tutte le polizze nello stesso verso.

MOTIVATION

- Le principali tipologie di rischio che insistono su di un portafoglio assicurativo e che costituiscono oggetto di studio degli esperti attuariali sono riconducibili al *rischio d'investimento* (o rischio finanziario) ed al *rischio demografico*
- *rischio demografico*: va analizzato distinguendone la componente assicurativa e quella di longevità. La componente assicurativa risulta connessa alle oscillazioni del numero effettivo di decessi rispetto al numero previsto.

Il rischio demografico prende in considerazione il *rischio di scarti accidentali* del numero di decessi dal valore atteso, dovuto a fatti non prevedibili e al normale andamento della mortalità in un portafoglio assicurativo, nonché il *rischio di scarti sistematici* (all'interno del quale si inserisce il *longevity risk*), facente riferimento a una mortalità strutturalmente diversa rispetto a quella attesa.

MOTIVATION

- La scelta del modello di proiezione della mortalità deve essere effettuata in base a criteri il più possibile oggettivi, sia di tipo qualitativo che quantitativo. Cairns et al. (2008) suggeriscono una lista di criteri utili per la scelta dei modelli di proiezione:
 - il modello deve risultare coerente con i dati storici
 - le dinamiche future a lungo termine del modello devono essere biologicamente ragionevoli
 - le stime dei parametri del modello e le previsioni a cui conduce devono risultare robuste rispetto sia al periodo temporale che alle età a cui il dato storico si riferisce
 - la previsione del livello di incertezza e le traiettorie centrali prodotte dal modello devono essere plausibili e coerenti con le tendenze storiche e la variabilità dei dati sulla mortalità

MOTIVATION

- La scelta del modello di proiezione della mortalità deve essere effettuata in base a criteri il più possibile oggettivi, sia di tipo qualitativo che quantitativo. Cairns et al. (2008) suggeriscono una lista di criteri utili per la scelta dei modelli di proiezione:
 - il modello deve essere semplice da applicare mediante metodi analitici o efficienti algoritmi numerici
 - il modello deve risultare relativamente parsimonioso
 - il modello deve consentire di integrare l'incertezza del parametro o dei parametri temporali nelle simulazioni e consentire il calcolo di intervalli di confidenza per la previsione.

MOTIVATION

- La letteratura attuariale ha sviluppato numerosi approcci in ordine all'obiettivo di effettuare proiezioni sui tassi di mortalità (si veda in particolare CMI, 2004, 2005a; Wong-Fupuy and Haberman, 2004), che potrebbero essere così classificati (Antolin 2007):

Process-based methods;

Expert-based methods;

Structural Modelling;

Decomposition, Disaggregation modelling;

Trend Modelling

MOTIVATION

- ❑ I metodi process-based fanno riferimento a dati biomedici della popolazione su cui vengono effettuate le proiezioni.
- ❑ I metodi degli esperti presentano una opinione che tende a risultare conservativa, nella misura in cui è legata alla scelta della popolazione target, all'orizzonte temporale considerato ecc.
- ❑ I metodi strutturali esplicitano relazioni econometriche presenti nel trend della sopravvivenza.
- ❑ I modelli di decomposizione si basano sulla distinzione per genere, per causa di morte.
- ❑ I metodi estrapolativi fondano le proiezioni sui trend storici della mortalità. La caratteristica precipua di tali metodi è che la mortalità ad una età x è funzione dell'anno di calendario t sulla base dei dati passati. Nell'ambito dei metodi estrapolativi si distinguono quelli di tipo deterministico, che semplicemente estendono l'andamento della sopravvivenza passata al futuro e quelli di tipo stocastico, che operano previsioni sulla base di date distribuzioni di probabilità.

MOTIVATION

- Il modello di Lee e Carter è stato ampiamente adottato. Esso infatti, rappresenta uno dei metodi utilizzati per l'estrapolazione della mortalità e la costruzione di stime per il futuro.

Attualmente infatti, la gran parte dei metodi usati nella pratica attuariale si basano su procedure interpolative-estrapolative.

- Il modello LC è funzione dell'età x e del tempo t , fondato su caratteristiche stocastiche (aleatorie) di mortalità. Una caratteristica è detta aleatoria quando è possibile ex-ante effettuare valutazioni attendibili nel suo verificarsi e sulle relative conseguenze.

Tale modello esplicita efficacemente i miglioramenti nel trend della mortalità e riesce a catturare i cambiamenti nel trend della mortalità. I dati sono facilmente implementabili e correggono la descrizione del fenomeno e conseguentemente evidenziano i cambiamenti in tale trend.

Il metodo proposto è puramente estrapolativo; non considera quindi, dati sociali e medici, ma l'attenzione è rivolta all'evoluzione del fenomeno nel tempo.

MOTIVATION

- Per spiegare la mortalità a una età specifica, il modello LC utilizza un tasso di morte centrale con età x e tempo t , $m_x(t)$:

$$\ln m_{x,t} = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

in cui:

$m_{x,t}$ il tasso centrale di mortalità calcolato per un individuo di età x al tempo t ;

α_x rappresenta il comportamento della mortalità alle varie età;

β_x è un parametro di sensitività, descrive la deviazione dal comportamento medio al variare di k_t ;

k_t è un indice univariato di mortalità globale nel tempo;

$\varepsilon_{x,t}$ indica il termine errore con $\mu = 0$ e varianza σ^2 ; esso si presenta come una variabile causale indipendente e identicamente distribuita. Questo termine riflette la variabilità non catturata dal modello.

MOTIVATION

- Per spiegare la mortalità a una età specifica, il modello LC utilizza un tasso di morte centrale con età x e tempo t , $m_x(t)$:

$$\ln m_{x,t} = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

Da questa equazione, parte l'analisi del modello Lee-Carter.

$m_{x,t}$ è regolato come un processo stocastico, da cui si ricavano i valori attesi dei tassi di mortalità.

Lee e Carter adottano un metodo molto semplice per descrivere le dinamiche della mortalità; esso inoltre è non parametrico in quanto è tale la dipendenza dall'età.

Nel modello viene utilizzata una unica storica, la quale indica una perfetta correlazione nei tassi di mortalità a differenti età da un anno all'altro.

MOTIVATION

- Il modello espresso dalla relazione

$$\ln m_x = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

non può essere assimilato ad una semplice regressione, in quanto non vi sono variabili osservabili direttamente al secondo membro dell'equazione.

La stima dei parametri α_x 's, β_x 's e k_t 's sulla base di una serie di dati osservati per n anni t e k anni x è condotta attraverso una serie di vincoli.

Una soluzione al metodo dei minimi quadrati può essere trovata utilizzando il SVD per il primo parametro α_x , basato su una matrice di tassi di mortalità disponibili.

Lee e Carter propongono la normalizzazione $\sum_x \beta_x = 1$, e $\sum_t k_t = 0$, per la quale ogni α_x sia la media dei tassi di mortalità del logaritmo centrale negli anni. I vincoli ai quali sono soggetti i parametri del modello, presentano la finalità di assicurare l'identificabilità del modello stesso.

MOTIVATION

■ Il modello espresso dalla relazione

Una volta stimati i parametri α_x 's, β_x 's e k_t 's, le previsioni ricavano dal modello i valori di k_t come una serie temporale.

Dal modello LC al modello Poisson log-bilineare

Il modello LC assume implicitamente che gli errori casuali siano omoschedastici.

Questa assunzione, che deriva dal metodo dei minimi quadrati ordinari usati per la stima, sembra essere poco realistica, in quanto il logaritmo dell'intensità di mortalità osservata è più variabile alle età anziane rispetto a quelle giovani, a causa del minor numero di morti osservato nelle età avanzate.

Brounhns et al.(2002a, 2002b) hanno studiato un possibile miglioramento del modello LC, usando una variazione casuale del numero di decessi di tipo Poisson invece che aggiungere il termine d'errore additivo $\epsilon_{x,t}$ al logaritmo del tasso di mortalità, risultando tale ipotesi molto più realistica ad età elevate.

MOTIVATION

- In termini di intensità di mortalità $\mu_x(t)$, l'assunzione poissoniana implica che il numero dei decessi all'età x nell'anno di calendario t è dato da:

$$D_x(t) \sim \text{Poisson}(L_x(t)\mu_x(t))$$

ove $L_x(t)$ esprime il numero medio di esposti al rischio.

In tale modello si assume inoltre che il logaritmo dell'intensità di mortalità abbia un andamento bilineare:

$$\ln(L_x(t)\mu_x(t)) = \alpha_x + \beta_x k_t$$

MOTIVATION

- La funzione, che esprime il logaritmo naturale dei tassi centrali di mortalità presenta la stessa struttura del modello Lee-Carter $\ln m_x = \alpha_x + \beta_x k_t + \epsilon_{x,t}$ eccetto che per il termine d'errore, mentre i decessi $D_x(t)$ sono distribuiti come una Poisson.

Anche in questo caso è necessario fissare dei vincoli sui parametri β_x e k_t ai fini dell'identificabilità dei parametri stessi.