Two problems arising in Population Modeling, Fast Detection of Unobservable Disorder Time and Cohort effect in Population dynamics

Workshop, "Extremes, Copula, and Actuarial Sciences"

N. El Karoui, S. Loisel, Y. Sahli

UPMC-Paris 6/LPMA/ISFA-Lyon 1

with the financial support of ANR LoLitA, and Chair "Risques financiers"

CIRM, 22/26 February 2016

Plan

ANR Program, 2013-2017

Dynamic Population Models

for Human Longevity with Lifestyle Adjustements

Topics

Main Objectives and Tasks

- Population Dynamics
- Long term Care Contracts
- Advanced Simulation methods
- Multi-year Solvency
- Statistical aspects of longevity risk
- Inter-generational Approach and pension issues

Perspectives

- Limit results and inference for population dynamics models, random environment
- Combining longevity, epidemics customer behaviour, joint stress- tests
- Long term care: measuring the impact of preventive actions, innovative care, etc...
- Synthetic internal risk model simulation issues
- Multi-population models
- Incorporating random longevity and interest rates into inter-generational risk models

Conferences

Data

- Requesting access to SNDS Health Database
- Study on AERAS convention thanks to SNDS database societal issues
  - Big data, longevity public health: solidarity risk sharing

Conferences

- LoLitA parallel sessions, EAJ Conference, Lyon, Sept. 2016
- LoLitA international conference, 2017

Part I

Demographic Transition in short

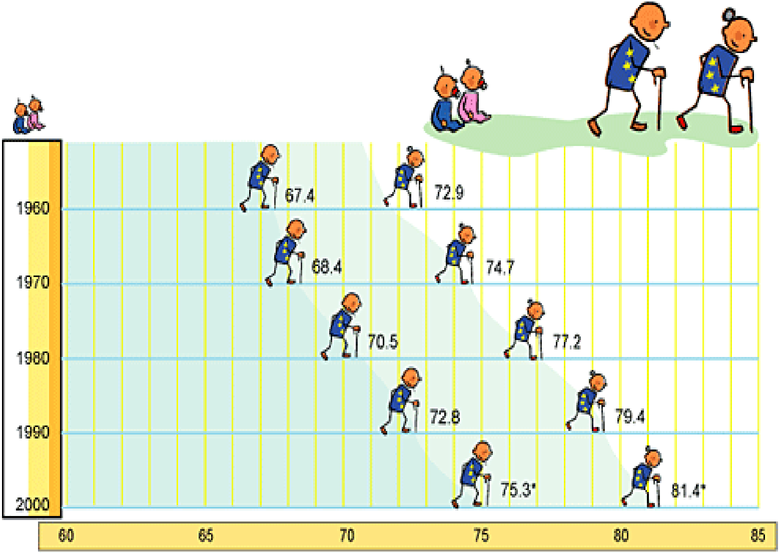


Figure: Longevity Risk

Thank you, Alexandre Boumezoued

+30 years for Life Expectancy (LE) in the last century

The demographic observation

- Substantial decline in mortality rate, in particular in small ages
- followed by reduction in fertility rate
- Heath transition (physical and cognitive development) and compression of morbidity

Economics aspects

- Economic growth, (income by head) and
- increase in social and political policy (education, democratie..)
- Growth in world population, citer Cohen

Health point of view from Cutler,Deaton,alii (2006)

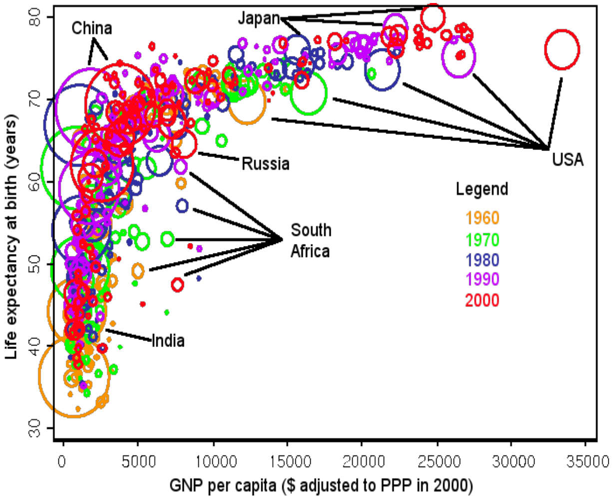
- Decline in infectious disease (60% of deaths in1848, $< 5\%$ in 1971 in UK)
- Nutritional improvement (debate on the importance)
- Progress in medecine, vaccins, ...

Public policies

- Macro public health: big public work projects (water purification explain half of mortality reduction in US ($1900 \sim 1930$)
- Reduction in alcoholism, in smoking
- Public and private health, contribute with complex impacts,
- *large heterogenity* with differences by age, type of sub-population, countries, with reverse or delayed effects.

Economic and Wealth Point of view

- Strong evidence on the links, but only 20% as impact
- Relation non-linear and concave
- Unexplained recent slower pace for LifeExp in US /Europa



Importance of inequalities

Evolutionary theory for aging

Example of Biological views

- Aging is characterized by the decline of physiological capacity
- Explain heterogeneity and randomness in individual patterns
- Nevertheless a robust observation in evolution theory, Gompertz (1825): The log mortality rate between 35-80y is linear in age.
- After 80y, large debate on the rectangularization of the survival curve, the question of "limited human life span"?

Life exptectancy at birth

- Lifetime of an individual: τ
- Life expectancy at birth: $\mathbb{E}[\tau]$, at ten $\mathbb{E}[\tau - 10/\tau > 10]$

Death rate

- Death rate $d(a)$ such that $\mathbb{P}(\tau > a) = e^{-\int_0^a d(s)ds}$
- In practice annual death probability reduction
  $q(a) = \mathbb{P}(\tau < a + 1 \mid \tau \geq a)$
- Mortality plateau (old ages)

Fertility rate

- Complex notion
- With large political connotation (Fertility, Immigration)

Data

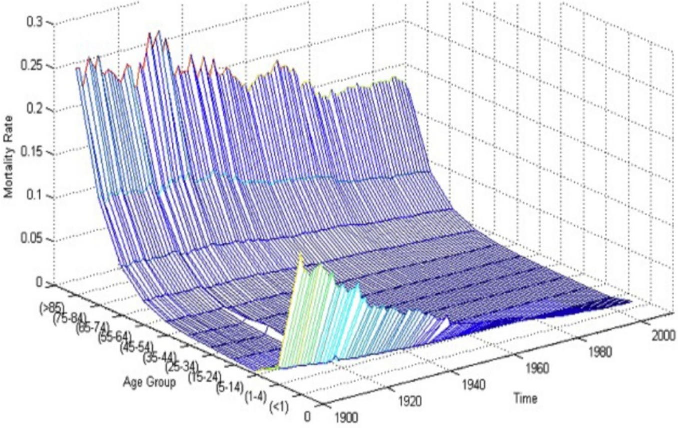
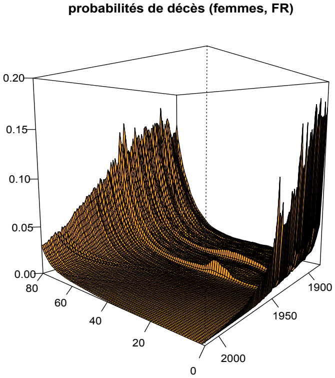
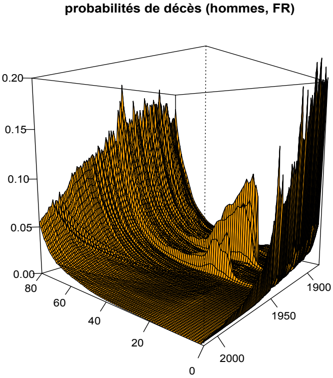


Figure 1. 1900-2004 Mortality Rate

National mortality by gender (France)

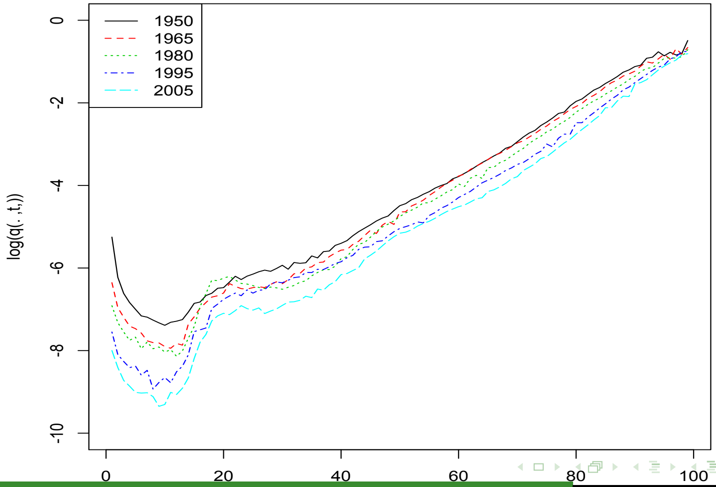




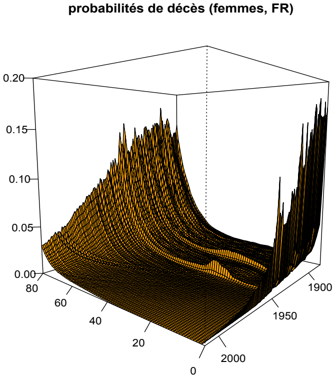
Femmes

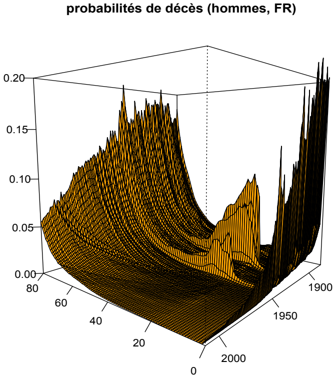
Hommes

- Looking at *log* $q(a,t)$ age a in [0,100]
- for different years t (1950,1965,1980,1995,2005)



National mortality by gender (France)





Femmes

Hommes

New economic and social challenge

Aging populations: new phenomenon, without past historical reference

- viability of shared collective systems, in particular (state or private) pension systems
- new generational equilibrium
- role and place of aging population in the society

Complex phenomenon, multi-causes

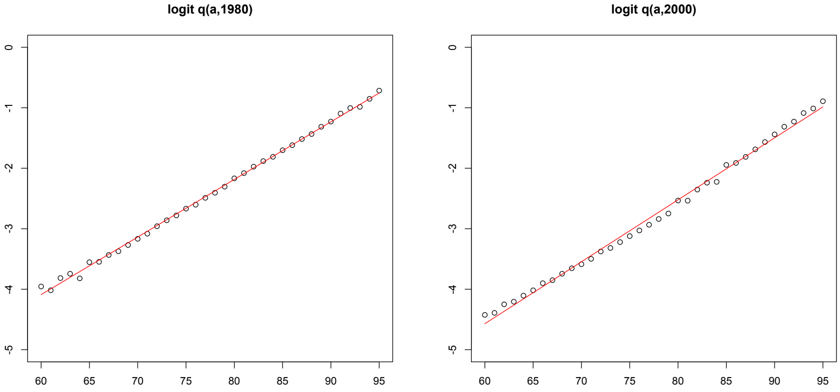
- Difficult to model.
- The role of age
- The heterogenity

Complex Estimation

- Coherence of the data

Cairns-Blake-Dowd model

- Logit of annual death probabilities for years 1980 and 2000 (French males)



Model for high ages (Cairns, Blake, Dowd, 2006):

$$\text{logit}\ (q(a,t)) = Y_1(t) + a.Y_2(t) + \epsilon_{a,t},$$

- $Y_1(t)$: overall reduction in mortality through time, for all ages,
- $Y_2(t)$: specific adjustment at each age,
- $\epsilon_{a,t}$ is the residual noise.

$\Rightarrow$ choice of a particular form of age dependency (Compertz=linear)

$\Rightarrow$ 2 time factors

Estimating parameters: for each year t between 1980 and 2007, we perform the linear regression over ages between 60 and 95, which gives parameters $Y_1(t)$ and $Y_2(t)$ (for men and women separately)

- Compression effect: constraint linking Y_1 and Y_2
  $\Rightarrow$ Mortality improvement transferred from old (~ 95) to younger ages (~ 60)

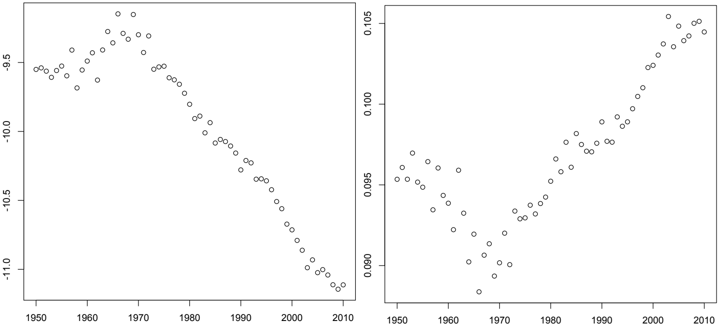


Figure: Processes Y_1 (left) and Y_2 (right) estimated for French males (ages 60-95) between 1950 and 2010

Cairns-Blake-Dowd model, IV

Time series Y_1 and Y_2 can be viewed as a fluctuating environment

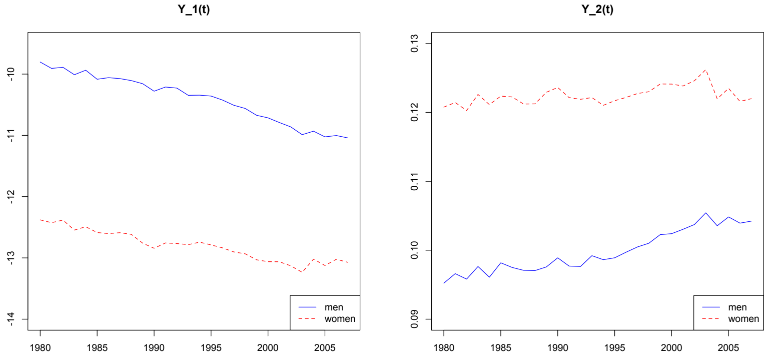


Figure: Estimated environment four factors on French data for ages 60-95 and years 1980-2007.

Plan

Part II

The disorder problem

Longevity Motivation

Motivations

All countries are experiencing a reduction in mortality over time.

- New, without past historical reference
- Societies are facing new different challenges: new generational equilibrium, viability of shared collective systems, in particular pension systems etc...
- Large Heterogeneity

Basis risk for insurer and pension funds, or two populations model

- Difference between national population and insured
- Mainly concerned with mortality evolution which may change over a projection period
- To detect any change any variation in the basis risk

Mortality evolution

- Mainly concerned with mortality evolution which may change over a projection period
- Mortality profile/level/trend : Heterogeneity and other factors
- Interest in experienced mortality : Deaths can be observed sequentially
- Model risk, parameter uncertainty: Long-term projection

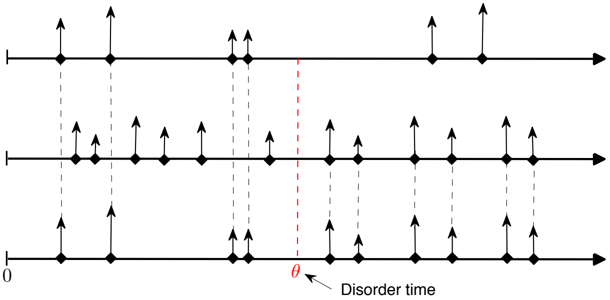
Risk Mangement of Mortality evolution

- Sequential information on death occurrences
- Monitoring and surveillance of mortality dynamics
- Updating mortality assumptions

What is the Disorder Problem?

The Poisson disorder problem is less formally stated as follows

- Observe a trajectory of the Poisson process (N_t) whose intensity changes from λ to $\rho\lambda$ at some (unknown) time θ.
- The problem is to find a rule to detect θ as quickly as possible with a limited number of false alarms



Where do the disorder problems arise

- Insurance compagnies: Recalculate the premiums for the future sales of insurance policies when the risk structure changes ($\rho > 1$)
- Pension funds: Large exposure to change in mortality risk : $\rho < 1$
- Quality control and maintenance: Inspect, recalibrate, or repair as soon a manufacturing process goes out of control
- Fraud and computer intrusion detection: Alert the inspectors for an immediate investigation as soon as abnormal credit card activity, cell phone calls, or computer network traffic are detected.

The observation process

- Let $N = (N_t)_{t \geq 0}$ be a counting process (of claims, or deaths) with stochastic intensity $\lambda = (\lambda_t)_{t \geq 0}$,
- N is said to be doubly stochastic point process, (DSPP)
- A change in the intensity occurs at an unobservable date θ, from λ_t to $\rho\lambda_t, \rho > 0$.
- Proportional hazard models

Change of Point θ=Model-risk

- Random with known prior (Bayesian)
- Deterministic but unknown (Non-Bayesian but Robust)

Change of Probability measures: Known statistics $\mathbb{P}$ and $\widetilde{\mathbb{P}}$

- Under $\mathbb{P}$, no change, (λ_t) holds
- Under $\widetilde{\mathbb{P}}$, immediate change, $(\rho\lambda_t))$ holds

Bayesian setup for random change-point

Brownian framework with abrupt change in the drift

- Based on the conditional distribution of the time of change,
- Formulated as an optimal stopping problem for partially observable process
- Page(1954), Shiryaev(1963), Roberts(1966), Beibel(1988), Moustakides (2004), and Dayanik (2006),....

Poisson framework with abrupt change in intensity

- Based on the conditional distribution of the time of change, with exponential or geometric prior distribution
- More recent studies : Gal (1971), Gapeev (2005), Bayraktar (2005, 2006), Dayanik (2006) for compound Poisson, Peskir, Shyriaev(2009) and others

New methods using particle filters

Andrieu, Legland (2004), Zhang (2005)

Robust Detection Problem, Page(1954)

Robust Detection

- Non-Bayesian framework, mainly motivated by the lack of any priors on the statistical behavior of change-point.
- Concerned with general counting process, in particular inhomogeneous Poisson process

Robust Criterium, Lorden(1971)

- Let τ be a stopping time, candidate for the estimation of θ
- The robust Lorden criterium for Poisson process is

$$C^{\mathrm{Lorden}}(\tau) = \sup_{\theta \in [0,\infty]} \operatorname{esssup}_{\omega} \mathbb{E}_\theta \left[(\tau - \theta)^+ \middle| \mathcal{F}_\theta \right]$$

- Our criterium, well-adapted to our general framework

$$C(T) = \sup_{\theta \in [0,\infty]} \operatorname{esssup}_{\omega} \mathbb{E}_\theta \left[(N_T - N_\theta)^+ \middle| \mathcal{F}_\theta \right]$$

Robust Optimization Problem

- Find T^* such that $C^* = C(T^*) = \min C(T)$, so that $\forall T$,
- $\sup_{\theta \in [0,\infty]} \operatorname{esssup}_\omega \mathbb{E}_\theta\left[(N_T - N_\theta)^+ \middle| \mathcal{F}_\theta\right] \geq \mathbb{E}_\theta\left[(N_{T^*} - N_\theta)^+ \middle| \mathcal{F}_\theta\right]$
- subject to the false alarm constraint, that the error of the second type, ("sound the alarm when there is not change" $(\theta = \infty)$) is small or equivalently if $\mathbb{E}[N_T] \geq \pi$.
- Min-Max Robust problem,
- To be stable by time rescaling

Conditional probability ratio between $\mathbb{P}$ =no change, $\widetilde{\mathbb{P}}$= immediate change.

- The conditional probability ratio process of $\widetilde{\mathbb{P}}$ w.r to $\mathbb{P}$ is
$$d\widetilde{\mathbb{P}}/d\mathbb{P} = \mathcal{E}_t = \exp\Big(\log(\rho)(N_t - (\rho-1)\Lambda_t\Big)$$
- Put $U_t^\rho = N_t - \beta(\rho)\Lambda_t$, so $\rho^{U_t^\rho}$ is a $\mathbb{P}$-martingale.
- Then the CPR of $\mathbb{P}$ w.r to $\widetilde{\mathbb{P}}$ is : $\rho^{-U_t^\rho} = (1/\rho)^{U_t^\rho}$

The β coefficient

- where $\beta(\rho) = \frac{\rho-1}{\log(\rho)}$, with $\beta(\rho) = \int_0^1 \rho^u du.$= Laplace transform of $U[0,1]$.
- Put $\tilde{\beta}(\rho) = \beta(1/\rho) = \beta(\rho)/\rho,\quad$ and $\tilde{\rho} = 1/\rho$.

Sequential conditional probability ratio:

- The $\log_\rho$ ratio of density processes on $\mathcal{F}_t$ of $\mathbb{P}^\theta$ w.r to $\mathbb{P}$ is

$$\log_\rho(d\mathbb{P}^\theta/d\mathbb{P}) = U^\rho_{t\vee\theta} - U^\rho_\theta \quad \text{where } U^\rho_t = N_t - \beta^\rho \Lambda_t$$

The Cumulative Sum rule (cusum)= Max Likelihood in time

- Based on $\max_{s\leq t}(\rho^{U_t - U_s})$, (sign of $\ln(\rho)$)
- Depend on the sign of $\rho - 1$
- Cusum stopping time, $\tau_m = \inf\{t\,;\mathrm{cusum}_t \geq m\}$

Main question: The optimality of τ_m, if the false alarm constraint is achieved, $\mathbb{E}(N_{\tau^\pi_m}) = \pi$.

Connection to insurance theory

Surplus process

- premium rate $(\rho - 1)\lambda_t$ and constant size of claims $\log \rho$
- Surplus by claim is $X_t(z) = z + X_t = z - N_t + \beta(\rho)\Lambda_t$

Ruin problem and Viability condition

- $\beta > 1$ is called the *security loading condition*, and $\mathbb{E}(X_t) = z + (\beta(\rho) - 1)\Lambda_t$ drifts to ∞
- $\mathbb{P}(\sup_t U_t \leq m) = \bar{u}(m)$ is finite and of main interest. Well known for a long time (Feller 1971) as scale function

Exit times For any cadlag process Z

- $\tau_m^Z = \inf\{t : Z_t \geq m\}$ and $\sigma_b^Z = \inf\{t : Z_t \leq b\}$
- $\bar{u}(m) = \mathbb{P}(\tau_m^U = +\infty)$ and $\bar{u}(m - U_t)$ is a martingale

Plan

Running supremum

- Running supremum $\bar{Z}_t = \sup_{s \leq t} Z_s$
- Put $X_t = -U_t = N_t - \beta\Lambda_t$. So, $\bar{X}_t$ is continuous.
- $\bar{U}_t$ is not continuous and increases only at jumps of N, such that $\bar{U}_t = U_t$

Reflected processes and Cusum processes

- Cusum rule based on $\max_{s \leq t} \rho^{U_t - U_s}$
- $\rho > 1$: Reflected process X at its maximum,

$$V_t = \sup_{\theta \leq t}(U_t - U_\theta) = \bar{X}_t - X_t$$

- $\rho < 1$ Reflected process U process at its maximum:

$$Y_t = \sup_{\theta \leq t}(X_t - X_\theta) = \bar{U}_t - U_t,$$

Typical paths for $\rho > 1$

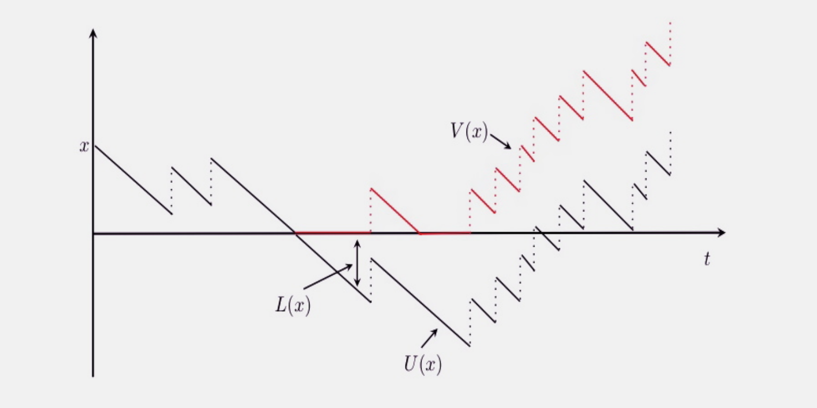
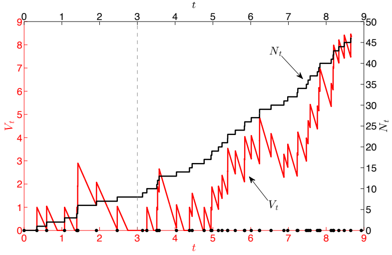
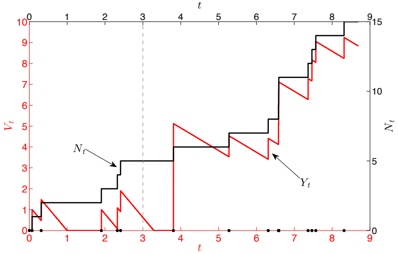


Figure: Sample paths of the processes $U(x)$, $V(x)$ and $L(x)$ when λ is time-homogeneous.

Typical paths with change of regime at date 3



(a) Processes N and V_t^ρ



Definitions with initial conditions

$$\begin{aligned} V_t(Z_0) &= U_t + \sup\{Z_0, \bar{X}_t\} = U_t(Z_0) + (\bar{X}_t - Z_0)^+, \\ Y_t(Z_0) &= X_t + \sup\{Z_0, \bar{U}_t\} = X_t(Z_0) + (\bar{U}_t - Z_0)^+, \\ \bar{X}_t^{ad} &= (\bar{X}_t - Z_0)^+, \quad \bar{U}_t^{ad} = (\bar{U}_t - Z_0)^+. \end{aligned}$$

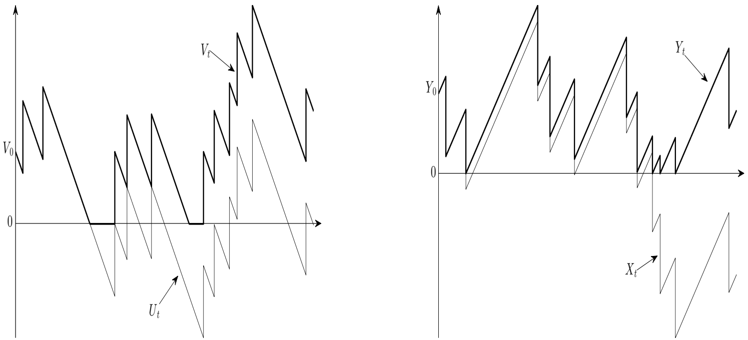
Differential Point of view of reflected processes$(j(y) = y \wedge 1)$

- V is the unique solution of the ODE driven by N,

$$dV_t = dN_t - \beta \,\mathrm{Ind}_{(0,\infty)}(V_t) d\Lambda_t, \;\; d\bar{X}_t^{ad} = \beta \mathrm{Ind}_{\{V_t=0\}} d\Lambda_t$$

- Y is the unique solution of the ODE driven by N

$$dY_t = -j(Y_t)dN_t + \beta d\Lambda_t, \quad d\bar{U}_t^{ad} = \mathrm{Ind}_{\{Y_t=0\}}(1 - Y_{t-})dN_t$$



(a) $\rho > 1$ (b) $\rho < 1$

Figure: Sample paths of the processes V, U, X and Y when λ is time-homogeneous.

Plan

Performance functions of the V-reflected process

- The performance of the cusum stopping is based on

$$\Gamma_t^m(x) = \widetilde{\mathbb{E}}_x \left[\mathrm{Ind}_{\tau_m^V \geq t}(N_{\tau_m^V} - N_t)\,|\mathcal{F}_\theta\right] = \tilde{h}_m(V_t(x)) \quad \widetilde{\mathbb{P}}\ a.s.$$

- $\widetilde{H}_t(x,m) = \tilde{h}_m(V_t) + N_t$ is a $\widetilde{\mathbb{P}}_x$-local martingale on $[0, \tau_m^V)$,
- Similar definition under $\mathbb{P}_x$, with $h_m(x) = \mathbb{E}_x(N_{\tau_m^V})$ and $H_t(x,m) = h_m(V_t) + N_t$ is a $\mathbb{P}_x$-local martingale on $[0, \tau_m)$

Performance functions of the Y-cusum rule

- Same criterium with $g_m(x) = \mathbb{E}_x(N_{\tau_m^Y})$ and $\tilde{g}_m(x) = \mathbb{E}_x(N_{\tau_m^Y})$
- $\widetilde{G}_t(x,m) = \tilde{g}_m(Y_t) + N_t$ is a $\widetilde{\mathbb{P}}_x$-local martingale on $[0, \tau_m^Y)$

Plan

Toward optimality: Two problems:

- Extension of the martingale property for any T
- Computation of the functions $h_n, \tilde{h}_m, g_m, \tilde{g}_m$

Differential finite variation calculus

- $\phi(z) = \phi(0) + \int_0^z \phi'(u)du + \sum_{\alpha \leq z} \phi(\alpha) - \phi(\alpha-)$. for ϕ with finite variation, a.e differentiable, with finite jumps
- Extension without problem for the g_m function since g_m is continuous $G_t(x, m) = \tilde{g}_m(Y_t) + N_t^{m,Y}$ is a $\widetilde{\mathbb{P}}_x$-local martingale.

Discontinuous finite variation function

Differential calculus for discontinuous functions

- The function h_m is discontinuous at m, and $h_m(V_t)$ has negative jumps when $V_t = V_{t-} = m$
- $J_t^{d,Z}$= number of down-crossings of m continuously
- H_t^m is the martingale $h_m(V_t) + N_t^{m,V} - h_m(m-)J_t^{d,V}$

Delayed Differential Equation

- $\beta u'(x) = u(x) - u(x-1), \quad \beta > 0.$
- $\beta h_m'(x) = h_m(x+1) - h_m(x) + 1, \quad x \in (0,m), \quad h_m'(0) = 0,$
- $\beta g_m'(x) = g_m(x) - g_m(x-1) - 1, \quad x \in (0,m).$
- $u(x),\ g_m(x) = 0$ for $x < 0$, and $h_m(x) = 0$ for $x > m$ with jump at $m = -h_m(m-)$

The basic DDE defined on $(0,\infty)$ (0 for $x<0$) with only one jump at 0.

$$\beta u'(x) = u(x) - u(x-1), \quad \beta > 0.$$

Delayed equation properties

- If $\beta = \beta(\rho)$, then $\rho^x u(x)$ is sol of DDE with $\tilde{\beta}(\rho) = \beta(1/\rho)$
- $\hat{u}(x) = \int_0^x u(z)dz$ is sol of DDE
$$\beta \hat{u}'(x) = \hat{u}(x) - \hat{u}(x-1) + \beta u(0) \quad \beta > 0, \quad u(0) = \hat{u}'(0).$$
- The solution W^ρ such that $W^\rho(0) = 1/\beta$ is the scale function in the Levy framework, characterized by its Laplace transform.
- the function $(x) = \rho^x u(x)$ is solution of the tilded DDE, and $\widetilde{W}^{\tilde{\rho}}(x) = \rho W^\rho(x)$

Old result, Feller 1971, $\beta > 1$

- The derivative is a solution (null for $x < 0$) of the convolution equation,

$$u'(x) = (1/\beta)\,\mathrm{Ind}_{[0,1)}(x)u(0) + (1/\beta)\int_0^1 u'(x-z)dz$$

- When $\beta > 1$, let S_n be sum of i.i.d. unif on $[0,1]$, S_n, and ν an indep. geometric r.v. with , $\mathbb{P}(\nu = j) = (1-1/\beta)\beta^{-j}$.

$$u'(x) = u(0)/(\beta-1)\mathbb{P}(S_\nu \in [x-1, x)),$$

- and $\bar{u}(x) = \mathbb{P}(\bar{U}_\infty \le x)$ is equal to $\mathbb{P}(S_\nu \le x)$
- The Laplace transform of $\frac{1}{\lambda(\beta-1)}\bar{u}(x)$ is equal to the inverse of rhe Laplace exponent of U^ρ, $\psi^\rho(\alpha) = \alpha\lambda(\beta(\rho) - \beta(e^{-\alpha}))$.
- $\frac{1}{\lambda(\beta-1)}\bar{u}(x)$ is a scale function (Bertoin(1996)).

Scale functions

Scale functions for increase intensity ($\rho > 1$)

- $W(x) = \frac{1}{(\beta-1)}\mathbb{P}(\bar{U}^{\rho}_{\infty} \leq x) = \frac{1}{(\beta-1)}\mathbb{P}(S_{\nu} \leq x)$
- $\widetilde{W}(x) = \rho^{x}\, W(x)$

Scale functions for decrease intensity ($\rho < 1$)

- $W(x) = \rho^{-x}\widetilde{W}(x)$
- $\widetilde{W}(x) = \frac{1}{\rho(\tilde{\beta}-1)}\, \widetilde{\mathbb{P}}(\bar{U}^{\rho}_{\infty} \leq x).$

Performance functions

Y-and V performance functions

Y performance

- $g_m(y) = \int_y^m W(z)dz, \quad \tilde{g}_m(y) = \int_y^m \rho \widetilde{W}(z)dz, \quad y \in [0, m].$

V performance

- $h_m(m-) = W(0)\frac{W(m)}{W'(m)}, \tilde{h}_m(m-) = \rho W(0)\frac{\widetilde{W}(m)}{\widetilde{W}'(m)}.$
- $h_m(x) = W(m-x)\frac{W(m)}{W'(m)} - \int_0^{m-x} W(y)dy$
- $\tilde{h}_m(x) = \rho\Big(\widetilde{W}(m-x)\frac{\widetilde{W}(m)}{\widetilde{W}'(m)} - \int_0^{m-x} \widetilde{W}(y)dy\Big)$

Comparison functions

- $\psi(y) = \rho^{-(m-y)} g_m(y) - \tilde{g}_m(y)/\rho$ is positive if $\rho < 1$.
- $\phi_m(m-z) = \frac{\tilde{h}_m(m-)}{h_m(m-)}\,\rho^{m-z}\, h_m(z) - \tilde{h}_m(z)$ is positive for $\rho > 1$ and $\tilde{h}'_m(x) = \rho^{m-x}\frac{\tilde{h}_m(m-)}{h_m(m-)} h'_m(x)$

Integration by parts

- $\Gamma_t^T := \widetilde{\mathbb{E}}_x(\int_t^T dN_s \big| \mathcal{F}_t)$.
- Let $(\bar{Z}_t)$ be a monotonic process as $\bar{X}_t$ or $-\bar{U}_t$.
- By integration de Γ_t^T with respect to $\rho^{\bar{Z}_t}$
  $\widetilde{\mathbb{E}}_x\big[\int_t^T \Gamma_\alpha^T d\rho^{\bar{Z}_\alpha} \big| \mathcal{F}_t\big] = \widetilde{\mathbb{E}}_x\big[\int_t^T (\rho^{\bar{Z}_{s-}} - \rho^{\bar{Z}_t}) dN_s \big| \mathcal{F}_t\big]$.

Applications to functions

- $\rho^x(\tilde{h}_m(x) - \tilde{h}_m(0)) = \rho\,\mathbb{E}_x\big(\int_0^{\tau_m} \rho^{V_{s-}} dN_s\big) - \tilde{h}_m(0)\,\mathbb{E}_x(\rho^{V_{\tau_m}})$,
- $\rho^{-y}(\tilde{g}_m(y) - \tilde{g}_m(0)) = \rho\,\mathbb{E}_y\big(\int_0^{\tau_m} \rho^{-Y_{s-}} dN_s\big) - \tilde{g}_m(0)\,\mathbb{E}_y(\rho^{-Y_{\tau_m}})$.

Applications to Lower bounds

- for $\rho > 1, \rho\,\mathbb{E}\big[\int_t^T \rho^{V_{s-}} dN_s \big| \mathcal{F}_t\big] \leq C(T)\,\widetilde{\mathbb{E}}(\rho^{V_T} \big| \mathcal{F}_t)$
- for $\rho < 1$ $\rho\,\mathbb{E}\big[\int_t^T \rho^{-Y_{s-}} dN_s \big| \mathcal{F}_t\big] \leq C(T)\,\mathbb{E}(\rho^{-Y_T} \big| \mathcal{F}_t)$
- $\tilde{h}_m(0)$ and $\tilde{g}_m(0)$ are respectively the cusum bounds of the stopping times τ_m^V $(\rho > 1)$ and τ_m^Y $(\rho < 1)$.

Plan

Optimality results

Optimality in decrease in intensity

- Let T be a stopping times with finite cusum bound, such that $\mathbb{E}(N_T) = \mathbb{E}(N_{\tau^Y_m}) = g_m(0).$, then
- $\mathbb{E}\big(\int_0^T \rho^{1-Y_{s-}} dN_s \big) \geq \tilde{h}_m(0)\, \mathbb{E}\big(\rho^{-Y_T}\big),$

Optimality in increase in intensity

- Let T be a stopping times with finite cusum bound,such that $\mathbb{E}(N_T) = \mathbb{E}(N_{\tau^V_m}) = h_m(0).$, then
- $\mathbb{E}\big(\int_0^T \rho^{V_s} dN_s \big) \geq \tilde{h}_m(0)\, \mathbb{E}\big(\rho^{V_T}\big),$

Argument for the proofs

Recall Comparison functions

- $\psi(y) = \rho^{-(m-y)} g_m(y) - \tilde{g}_m(y)/\rho$ is positive if $\rho < 1$.
- $\phi_m(m-z) = \frac{\tilde{h}_m(m-)}{h_m(m-)} \rho^{m-z} h_m(z) - \tilde{h}_m(z)$ is positive for $\rho > 1$ and $\tilde{h}'_m(x) = \rho^{m-x} \frac{\tilde{h}_m(m-)}{h_m(m-)} h'_m(x)$
- Generalized martingales properties
- Use a multiple of the constraint to compare the integral for cusum processes above m
- Eliminate the discontinuous local time ($\rho > 1$) and use the comparison functions
- The case $\rho > 1$ is difficult and its was not clear that cusum strategy are optimal.
- In the Brownian case, the proofs are easy : thanks to continuity

Detection Procedure – Algorithm

Step 1: Fix the input parameters: The post-change intensity through the specification of ρ and the false alarm constraint π.

Step 2: Determine the threshold m as the solution of the equation $\mathbb{E}_\infty[N_{\tau_m}] = \pi$.

Step 3: For each new observation at time t compute the value of the CUSUM process V given by the iterative relation $V_{t+1} = (V_{t-1} + U_t)^+$.

Step 4: Compare the current value of V to the threshold m and stop the procedure once $V_t \geq m$ and sound an alarm. Hence $\tau_m(0) = t$.

Numerical instability of $\bar{u}_m$, and h_m

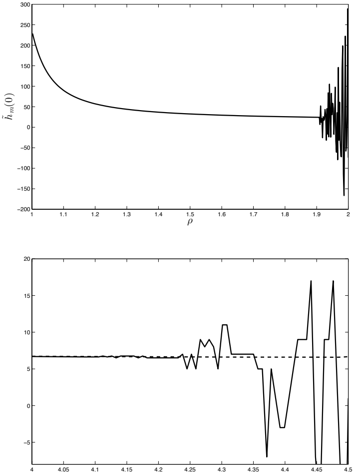


Figure: Scale function $W(x)$ for different values of ρ.

Evolution of the global population

$$\lambda_t = a\big(1 + exp(-(t-b)/c)\big)^{-1},\ t \geq 0,$$

where a, b and c are some constant parameters given in

a	b	c	σ
13.80	11.85	26.40	0.0907

- 10000 simulations for $\rho = 1.1, 1.5, 2$

	ρ	1.1	1.5	2
$m=5$	$h_0(0)$	25.01	17.77 ()	14.15
	$h_\infty(0)$	32.92 ()	58.52 ()	116.82
$m=10$	$h_0(0)$	81.48 ()	44.49 ()	31.97
	$h_\infty(0)$	147.73 ()	643.29 ()	4174.49

Detection Procedure – Real World

We consider the Continuous Mortality Investigation assured lives dataset and England & Wales national population.
We split data into two periods:

- We consider the period 1947-1969 as a training period.
- The Cox model is estimated over this period using the MLE.
- It is not easy to estimate the post change coefficient ρ

Hence we monitor sequentially the dataset over the period 1970-2005 and look for changes on the mortality of assured lives.

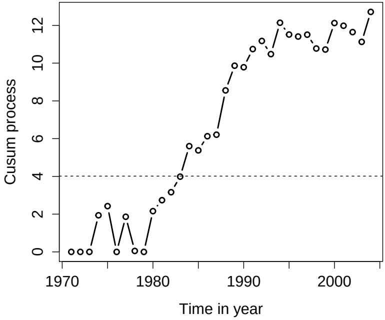


Figure: Detection scheme for age groups $50-59$ (right) and $80-89$ (left). The post-change is set to $\rho = 1.15\%$ and the false alarm constraint to $\pi = 100 \times \bar{\lambda}$.

Detection Procedure – Real World

	τ_m	
Age	$\rho = 1.50$	$\rho = 1.15$
$50 - 59$	1984	1978
$60 - 69$	1991	1985
$70 - 79$	1988	1984
$80 - 89$	1983	1978

Table: Detection of mortality change with a post-change ratio of $\rho = 1.15$ and an average run length (false alarm) constraint of 100. The right column reports the detected change-point using an off-line procedure.

- When ρ is larger the detection delay is longer
- Difficile to disctinct effective change or false alarm
- Other results mention that a breakpoint on the data on the

- The CUSUM detection rule is optimal in the case of non-homogeneous Poisson process with a modified Lorden criterion
- Very easy to implement
- Non asymptotic criterium
- Based on fine properties of scale functions, easy to extend to Levy processes
- The proof provide lower bound for some conditional ratio (Basseville)
- Applications in non-life insurance
- Further research on possible extensions

Part III

Individual Based Model In Longevity

Demographic rates: an individual of traits $x_t \in \mathcal{X} \subset \mathbb{R}^d$ and age $a_t \in [0, \bar{a}]$ at time t, (born at time 0)

- Dies at rate $d\,(x_t, a_t, t, Y)$
- Gives birth at rate $b\,(x_t, a_t, t, Y)$
  and the new individual has traits $x' \sim K^b(x_t, a_t, dx')$
- Evolves during life at rate $e\,(x_t, a_t, t, Y)$
  from traits x_t to $x' \sim K^e(x_t, a_t, dx')$

Environmental factors

- Demographic rates depend on characteristics, age, time and on the stochastic environment Y
- Conditionally on the environment Y, the events for a given individual are jumps of a counting process

The thinning construction can be used to define a wide variety of processes as solution to stochastic equations.

Intensity for spatial Birth Process

- naissance +"mutation"individual = rate becomes $b(x')k_b(x', x)m(dx')$
- aggregated rate of birth mutation=

$$\beta(\xi, x) = \sum_{x' \in \xi} b(x')k_b(x', x)$$

- Equation $Z(dt, dx) = \int_{\mathbb{R}_+} \mathbf{1}_{\theta \leq \beta(Z_{t-}, x)} Q_b(dt, dx, d\theta)$

Birth with age

- First define the new kernel with the age
- Applied the previous relation to process $d\langle Z_t, f\rangle = \langle I^b_{Z_{t-}} f(., 0), Q_b\rangle(dt) + \langle Z_t, \partial_a f\rangle dt$
- Existence result similar to the linear case

Fondamental asymmetry

- since the newborn is from outside,
- then the death remove an individual in the population

How to select an individual by its characteristics

- the counting measure on E is not a "Radon" σ-finite measure on E
- Necessity to give a measurable and adapted process to select individual in a given population

Envelop process of population path without accumulation

- The age desagregated population $\tilde{\xi}_s(dx) = \xi_s(dx, \mathbb{R}_+)$
- The non decreasing envelop $\bar{\xi}_t = \bigcup_{s\in[0,t]} \tilde{\xi}_s$ process
- $\bar{\xi}_t$ has only finite number of jumps on $[0, t]$, denoted by (S_k)
- (S_k) are also times of jumps for the path ξ_t
- The sequence $(X_k, A_k(.))_{k\geq N_0+1}$, $X_k = \bar{\xi}_{S_{k-N_0}} \setminus \bar{\xi}_{S^-_{k-N_0}}$, and $A_k(t) = t - S_k$

Spatial death process

- A Poisson point measure $Q_d(ds, di, d\theta)$ on $\mathbb{R}_+ \times \mathbb{N}^* \times \mathbb{R}_+$
- with intensity measure $q_d(ds, di, d\theta) = ds\, n(di)\, d\theta$
- $I^d(Z_{t-}, i, \theta) = \mathbf{1}_{X_i\in Z_{t-}} \mathbf{1}_{\theta\leq d(X_i)}$
- Using the previous numbering, we see that $Z(dt, dx) = -\int_{i\in\mathbb{N}^*} \int_{\theta\in\mathbb{R}_+} I^d(Z_{t-}, i, \theta)\delta_{X^i}(dx)Q_d(dt, di, d\theta),$

Theorem Bezborodov (2014), Garcia 1999, ...

- If $\xi_0^1 \subset \xi_0^2$,
- $\beta_1(x, \eta^1) \leq \beta_2(x, \eta^2) \quad \eta^1 \subset \eta^2$
- $d_1(x, \eta^1) \geq d_2(x, \eta^2) \quad \eta^2 \subset \eta^1, x \in \eta^1$

The comparison theorem

There exists a cadlag process (η_t) *such that* $\eta_t \subset \xi_t^2$ *having the same law that* (ξ_t^1)

Sketch of the proof without age, and swap

$$\eta(dt, B) = \int_{B\times\mathbb{R}^+} \mathbf{1}_{[0, b_1(x,\eta_{s-})]}(\theta) dQ^b(dt, dx, d\theta)$$

$$-\int_{\mathbb{N}\times\mathbb{R}^+} \mathbf{1}_{\{x_i^2 \in \eta_{s-}\cap B\}} \mathbf{1}_{[0, d_1(x_i^2, \eta_{s-})]}(\theta) dQ^d(dt, di, d\theta)$$

Applications of comparison theorem

In progess

Study classical properties of population processes

- Agregation by traits and convexity
- Localisation and explosion
- Monotonic convergence

Stochastic order on the space of configuration

- Starting from the result of Preston (1975) on the stochastic order for the Point random field
- Property of the stochastic order on the distributions of the population processes Z_t in terms of demographic characteristics

The Poisson measures driving the equation

- $Q_b,\ Q_d,\ Q_e$
- $I^b(Z_{t-},t,x,\theta),\ I^d(Z_{t-},t,i,x,\theta),\ I^e(Z_{t-},t,i,x,\theta)$

The BSD Population equation

$$d\langle Z_t,f\rangle = \langle I^b_{Z_{t-},t}f(.,0),Q_b\rangle(dt) - \langle I^d_{Z_{t-},t}f(X_.,A_.(t)),Q_d\rangle(dt)$$
$$+\langle I^e_{Z_{t-},t}[f(.,A_.(t)) - f(X_.,A_.(t))],Q_e\rangle(dt) + \langle Z_t,\partial_a f\rangle dt. \tag{1}$$

Hypotheses, $E=\mathbb{R}^d, m(dx)=l(dx)$

- $\int b(x,\eta)dx \le c_1|\eta| + c_2$
- $\sup_x \sup_{\{|\eta|\le m\}} d(x,\eta) < \infty$

Then, existence and strong uniqueness

Plan

Cohort Effect

Cohort effect

An example of numerical experiment to explain an observed phenomenon

Cohort effect

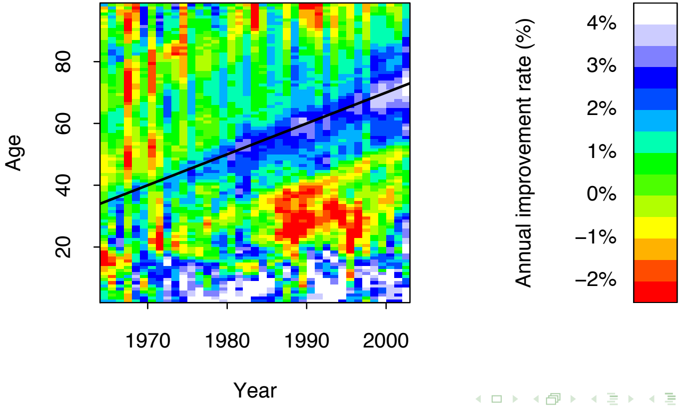
Birth cohort for the period $[t_1, t_2]$: group of individuals born between t_1 and t_2.

- Individuals of the same birth cohort share similar demographic characteristics ("cohort effect")
- Age, Period, Cohort analysis put a lot of problems in practice, in different domains, medecine, sociology,...due to the lag in data,..insurance...
- Huge literature on APC problems

Golden cohort: generations born between 1925 and 1940

Cairns et al. (2009) $$r_{a,t} = (q_{a,t-1} - q_{a,t})/q_{a,t}$$

The Golden cohort has experienced more rapid improvements than earlier and later generations.



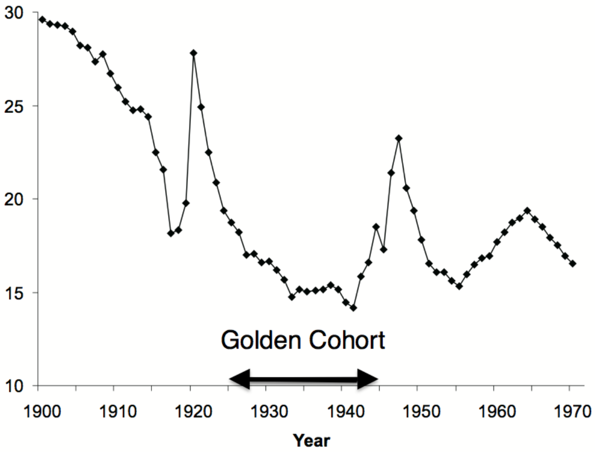
Some possible explanations:

- Impact of World War II on previous generations,
- Changes on smoking prevalence: tobacco consumption in next generations,
- Impact of diet in early life,
- Post World War II welfare state,
- Patterns of birth rates

"One possible consequence of rapidly changing birth rates is that the 'average' child is likely to be different in periods where birth rates are very different. For instance, if trends in fertility vary by socio-economic class, the class mix of a population will change."

The Cohort Effect: Insights And Explanations, 2004, R. C. Willets

Cohort effect and Fertility



Data source: www.mortality.org

Figure 6. Crude birth rate per 1,000 population, England and Wales, 1900 to 1970

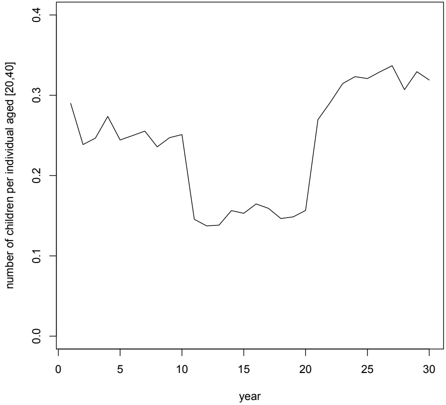
The different rates

- Reference death rate $\bar{d}(a) = A\exp(Ba)$
- Parameters $A \sim 0.0004$ and $B \sim 0.073$ estimated on French national data for year 1925 to capture a proper order of magnitude
- "Upper class": time independent death rate $d^1(a) = \bar{d}(a)$ and birth rate $b^1(a) = c\mathbf{1}_{[20,40]}(a)$ (c=0.1)
- "Lower class": time independent death rate $d^2(a) = 2\bar{d}(a)$ but birth rate $b^2(a,t) = 4c\mathbf{1}_{[20,40]}(a)\mathbf{1}_{[0,t_1]\cup[t_3,\infty)}(t) + 2c\mathbf{1}_{[20,40]}(a)\mathbf{1}_{[t_2,t_3]}(t)$

Comment Constant death rates but reduction in overall fertility between times t_1 (=10) and t_2 (=20).

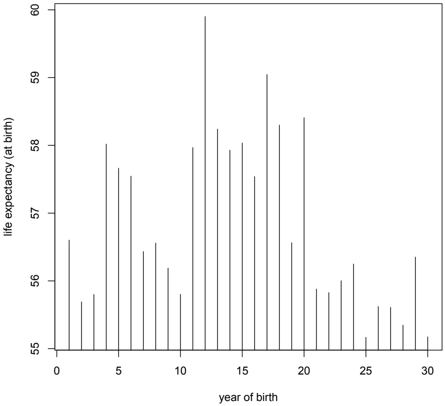
- Aim: Test the cohort effect by computing standard

- One trajectory with 20000 individuals (randomly) splitted between groups. Estimation of aggregate fertility



Life expectancy by year of birth

- "Cohort effect" for aggregate life expectancy



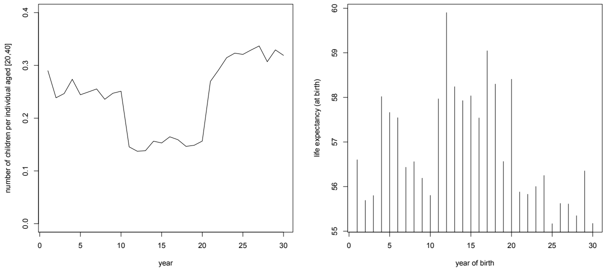


Figure: Observed fertility (left) and estimated life expectancy by year of birth (right)

- Death rates by specific group remain the same
- But reduction in fertility for "lower class" during 10-20 modifies the generations composition
  $\Rightarrow$ "upper class" is more represented among those born between 10 and 20

Plan

Longevity patterns and longevity improvements are very different for different countries, and different geographic area.

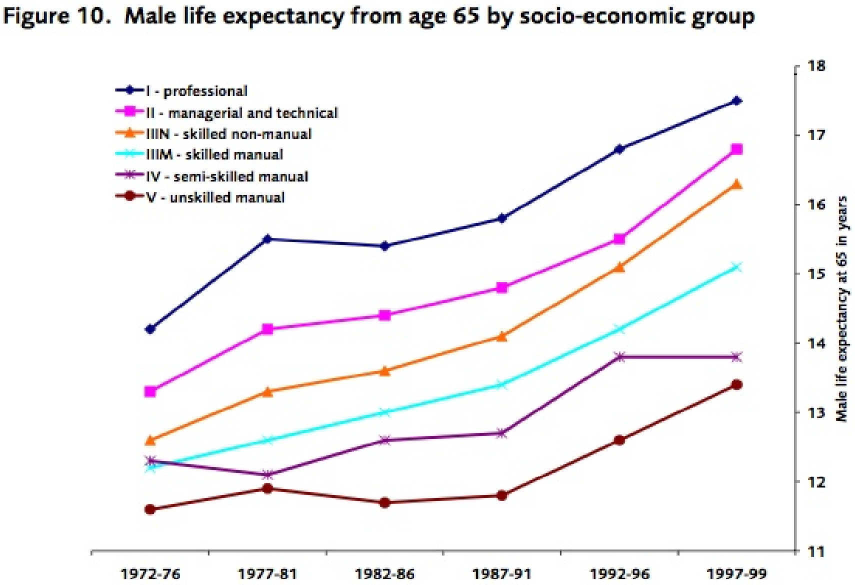
Factor affecting mortality

- socio-economic level (occupation, income, education, wealth...)
- gender
- marital status
- living environment (pollution, nutritional standards, hygienic...)
- Take them into account in a stochastic mortality model

Conditional calibration

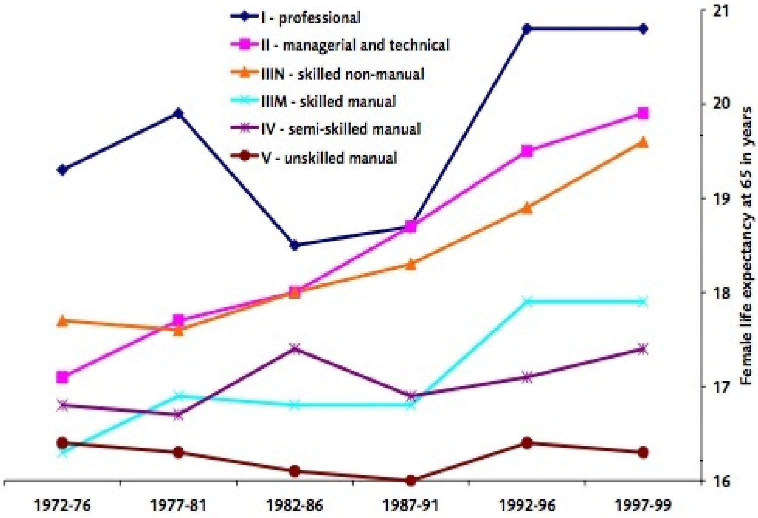
- On national mortality data and on specific data (with information on individual characteristics)

Male Life Expectancy from age 65



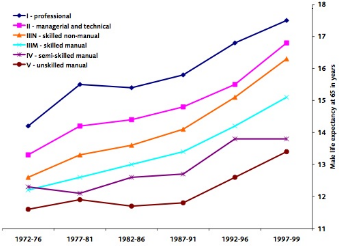
Female Life Expectancy from age 65

Figure 11. Female life expectancy from age 65 by socio-economic group



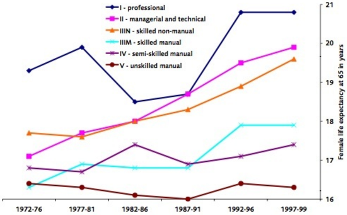
Social Heterogenity of Life Expectancy

Figure 10. Male life expectancy from age 65 by socio-economic group



Source: ONS Longitudinal Survey.

Figure 11. Female life expectancy from age 65 by socio-economic group



Source: ONS Longitudinal Survey.

Basis risk I

Difference : national mortality versus that of specific group

- Insurance companies can use national reliable mortality estimates on large samples
- but the final goal is to model mortality rates specific to subpopulations with owns traits
  - population of a small country or region,
  - individuals with a specific disease,
  - insurance portfolio,
  - annuitants of sectorial pension funds.
- But also how take into account other informations
  - They know the exact ages at death and not only the year of death (time continuous data)

Basis risk II

- Cause of death are specified
- Characteristics of the policyholders : socio economic level, living conditions ...
- selection bias

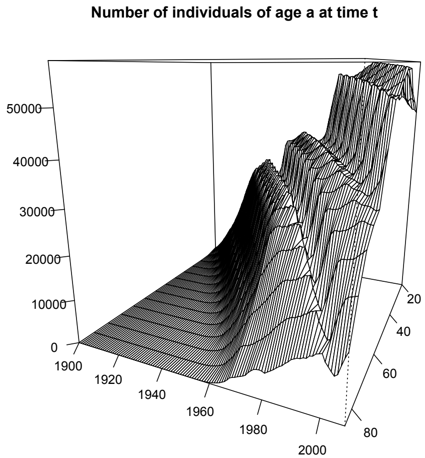
▸ BUT

- limited size of their portfolios (in comparison to national populations : 700 000 individuals from 19 different insurance companies)
- small range of the observation period

This heterogeneity is very important for longevity risk transfer based on national indices: for too important basis risk, the hedge would be too imperfect

Permanent Demographic Sample

- Number of individuals at each year by 5 years age groups in the sample



Heterogeneity of mortality

Some tests on three characteristics of interest

- Education level
  - Group 1: Diploma $\leq$ Baccalauréat (high school diploma)
  - Group 2: Diploma $>$ Baccalauréat (high school diploma)
- Socio-professional category
  - Groupe 1: Employees and workers
  - Groupe 2: Executives and higher intellectual professions, intermediaries professional categories
- Marital status
  - Group 1: Single or divorced
  - Group 2: Married or widowed

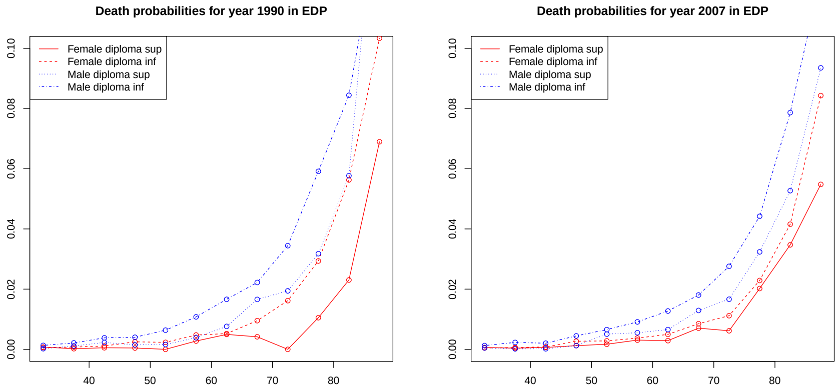


Figure: Death probabilities by education level: years 1990 and 2007

Mortality heterogeneity: socio-professional cat.

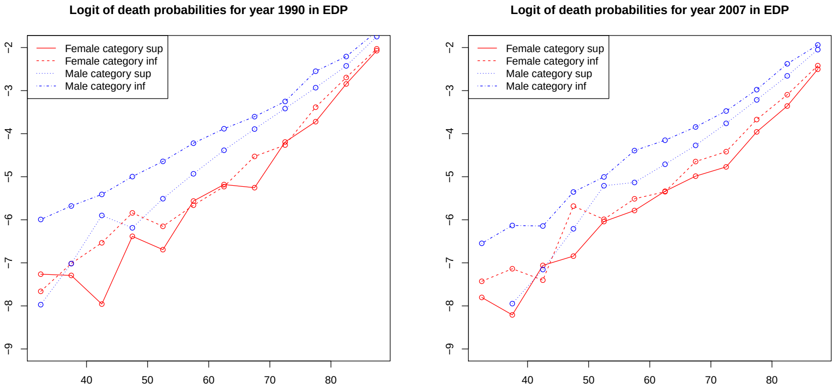


Figure: Logit of death probabilities by socio-professional category: years 1990 and 2007

Mortality heterogeneity: marital status

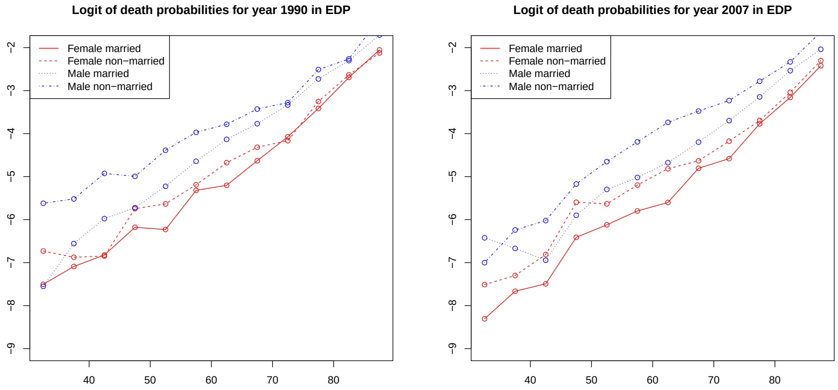
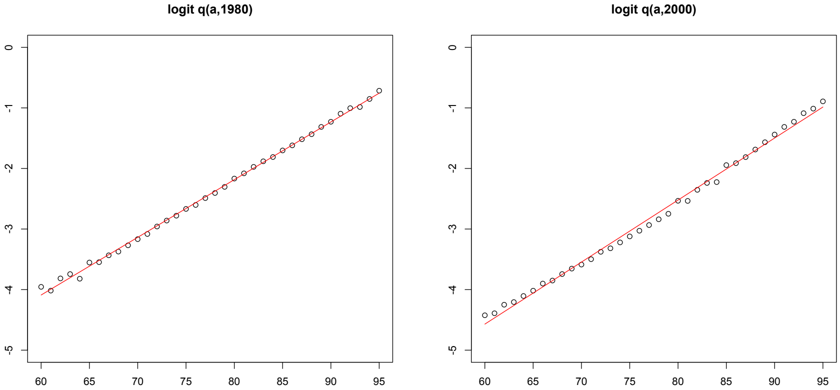


Figure: Logit of death probabilities by marital status: years 1990 and 2007

Cairns-Blake-Dowd model

- Logit of annual death probabilities for years 1980 and 2000 (French males)



Model for high ages (Cairns, Blake, Dowd, 2006):

$$\text{logit}\ (q(a,t)) = Y_1(t) + a.Y_2(t) + \epsilon_{a,t},$$

- $Y_1(t)$: overall reduction in mortality through time, for all ages,
- $Y_2(t)$: specific adjustment at each age,
- $\epsilon_{a,t}$ is the residual noise.

$\Rightarrow$ choice of a particular form of age dependency (Compertz=linear)

$\Rightarrow$ 2 time factors

Estimating parameters: for each year t between 1980 and 2007, we perform the linear regression over ages between 60 and 95, which gives parameters $Y_1(t)$ and $Y_2(t)$ (for men and women separately)

- Compression effect: constraint linking Y_1 and Y_2
  $\Rightarrow$ Mortality improvement transferred from old (~ 95) to younger ages (~ 60)

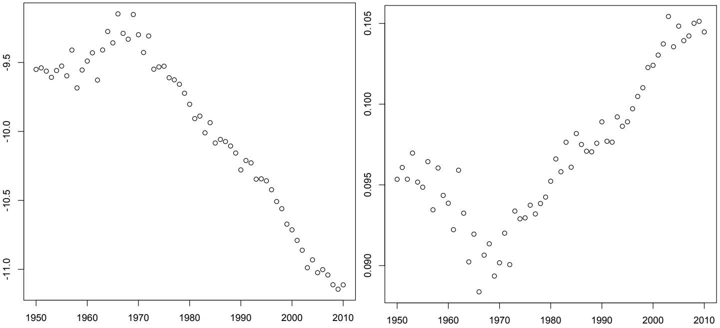


Figure: Processes Y_1 (left) and Y_2 (right) estimated for French males (ages 60-95) between 1950 and 2010

Cairns-Blake-Dowd model, IV

Time series Y_1 and Y_2 can be viewed as a fluctuating environment

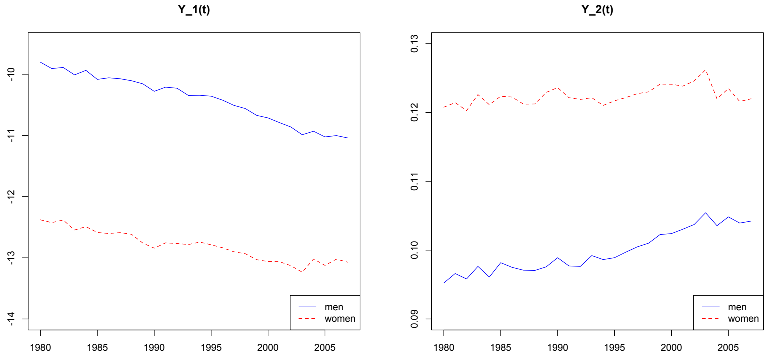
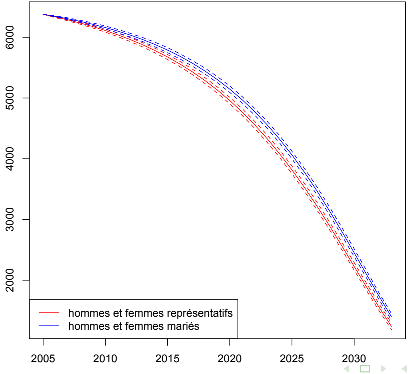


Figure: Estimated environment four factors on French data for ages 60-95 and years 1980-2007.

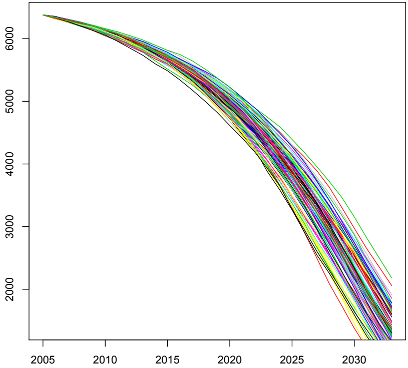
Numerical example: French sub-population

- Cohort of French (males and females) aged 61 at the beginning of year 2005 in the Permanent Demographic Sample
- Confidence intervals at 90% for the number of individuals without environment noise



Numerical example: French sub-population

- The model allows to simulate the evolution of the population, subject to various death rates dues to different environment scenarios



- Application to an insurance portfolio: initial age distribution

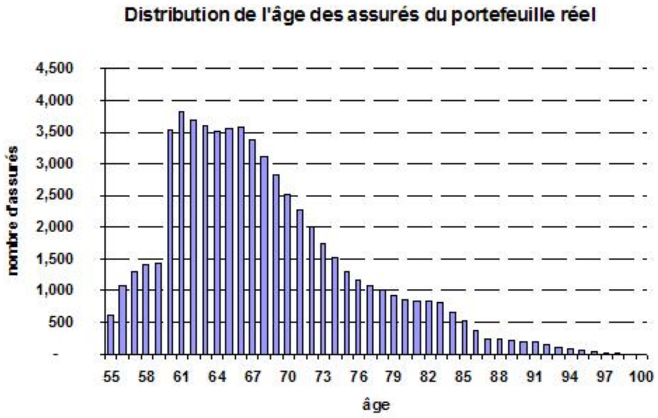
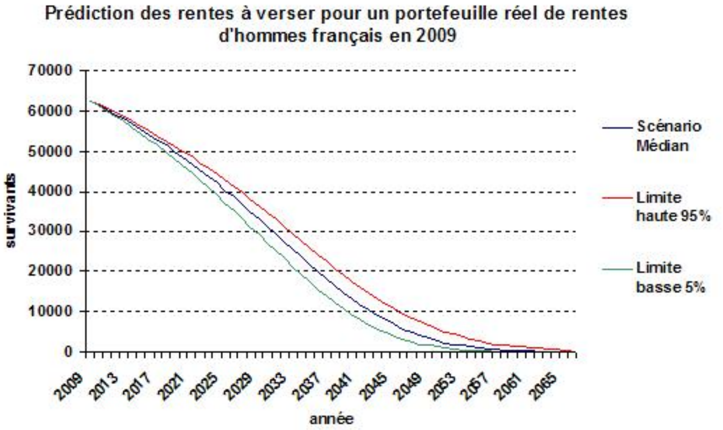


Figure: Confidence interval at 90% on the size of the insured population

- Application to an insurance portfolio: pension amount



With the indulgence of Sau, Gold of health and longevity

Thank you