Mortality Maps Based On Spatial Extrapolation

Joint work with B. Babel, R. Schmidt and V. Schmidt

Stefanie Eckel
Department of Stochastics
Graduate College 1100
University of Ulm

Overview

- Introduction
- Data description
- Construction of mortality maps
- Explorative space-time analysis and its results
- Application in life insurance pricing
- Outlook: Extensions
- References

Introduction
Motivation

- Mortality described by death rate depending on gender, age and year
- Death rates as input factor for contractual calculations of life insurances or pension plans (out of life tables)
- German life insurance contracts do not depend on regional characteristics of mortality

Aim

Analysis of differences and similarities of the spatial demographical structure (in space and time)

Data description

Demographical data set provided by the Federal Statistical Office in Wiesbaden

- Population sizes and numbers of deaths
- 322 administrative districts and 106 non-district towns
- Time range: 1995 to 2003
- Gender and age groups: younger than 50, 50 to 65, 65 to 75, older than 75

Construction of mortality maps
Calculation of death rates

- Death rate

$$m_t = \frac{d_t}{p_{t-1}}, \ t = 1996, ..., 2003$$

with number of deaths d_t and size of population p_t in year t

- Measurement points $u_1, ..., u_n$ in sampling window W
- Time horizons: $T_1 = 1996 - 1998$, $T_2 = 1999 - 2001$, $T_3 = 2002 - 2003$
- Weighted mean death rate

$$m_{T_j}(u_i) = \frac{\sum_{t \in T_j} p_{t-1}(u_i) m_t(u_i)}{\sum_{t \in T_j} p_{t-1}(u_i)}, \ j = 1, 2, 3; i = 1, ..., n$$

Construction of mortality maps

Topology

Locations of measurement points

Construction of mortality maps
Extrapolation method

Inverse distance method

Estimation of $m_{T_j}(u_0)$ at a non-observed location u_0 by the linear convex combination

$$\hat{m}_{T_j}(u_0) = \sum_{i=1}^{n} \lambda_{ij} m_{T_j}(u_i),\ j = 1, 2, 3$$

with some weights λ_{ij} such that

- $\lambda_{ij} \geq 0$
- $\sum_{i=1}^{n} \lambda_{ij} = 1$

Construction of mortality maps
Extrapolation method

Under the assumption that $u_0 \neq u_i$ for all $i = 1, ..., n$

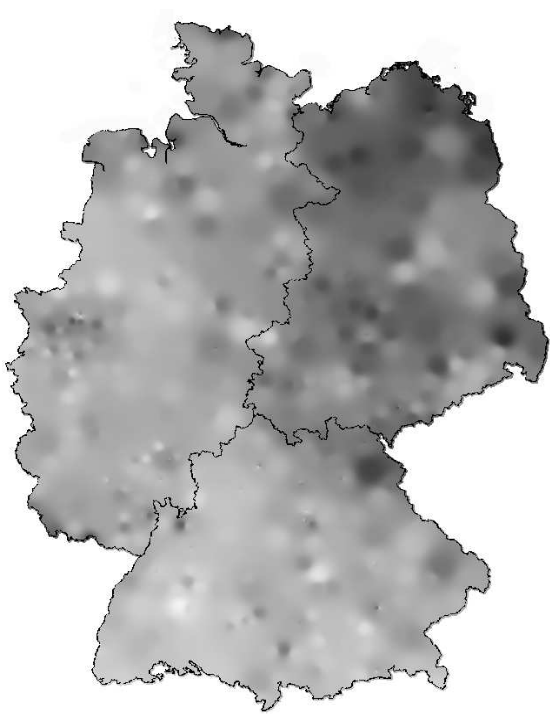
$$\lambda_{ij} = \begin{cases} \frac{p_{T_j}(u_i)}{|u_i - u_0|^3} \left(\sum_{k=1}^{n} \frac{p_{T_j}(u_k)}{|u_k - u_0|^3} \right)^{-1} & \text{if } |u_i - u_0| \leq r \\ 0 & \text{if } |u_i - u_0| > r \end{cases}$$

where $|\cdot|$ denotes the Euclidean distance and
$p_{T_j}(u_i) = \sum_{t \in T_j} p_{t-1}(u_i)$

- The larger the distance $|u_i - u_0|$ the smaller its weight w.r.t. this measurement
- Influence only from districts in the immediate neighborhood ($r = 80$ km)

Construction of mortality maps
Grey scale images

- Each pixel has a value in $[0, 255]$
- Light pixels refer to low death rates, dark pixels to high death rates



Grey scale images

Explorative space-time analysis
Regional mortality

Differences and similarities in South, East and West Germany

Threshold method to construct binary images

- Threshold μ_{T_j} as the weighted average of death rates per period T_j

$$\mu_{T_j} = \frac{1}{\sum_{i=1}^{n} p_{T_j}(u_i)} \sum_{i=1}^{n} p_{T_j}(u_i) m_{T_j}(u_i) \ , j = 1, 2, 3$$

- Pixels with death rates below the threshold are white, otherwise they are black

Explorative space-time analysis

Regional mortality - results

Female population aged between 50 and 65 over time horizons 1996-1998 and 2002-2003

- Lower mortality in South Germany in both time periods
- More homogeneous spatial distribution in period T_3

Explorative space-time analysis
Regional mortality - results

Possible reasons for regions with higher mortality

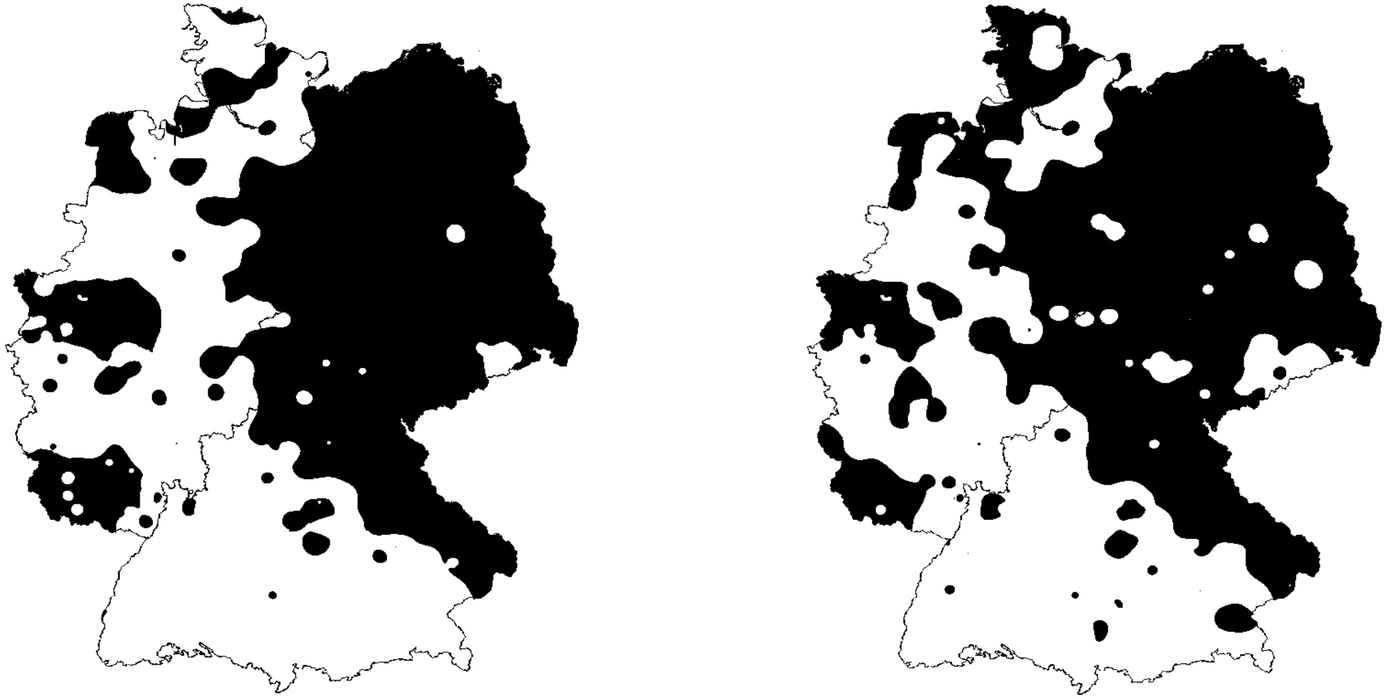
- More frequently occurrence of deadly diseases, more fatal traffic accidents
- Psycho-social stress caused by social, political and economic changes
- Poor economic conditions, unemployment
- Higher environmental pollution
- Increased consumption of alcohol

Explorative space-time analysis
Regional mortality - results

Female population aged between 65 and 75 over time horizons 1996-1998 and 2002-2003

Homogenization is less pronounced for the older female population

Explorative space-time analysis
Regional mortality - results



Male population aged between 50 and 65 over time horizons 1996-1998 and 2002-2003

- Almost no homogenization of regions with higher mortality
- Lower mortality in South Germany except for the Bavarian forest

Explorative space-time analysis
Regional mortality - results

Possible reason for the absence of homogenization

- Reasons for higher mortality as in the case of female population
- Lower ability of men to cope with these problems

Possible reasons for higher mortality in the Bavarian forest

- Cancer in the respiratory or alimentary system
- Lower economical development

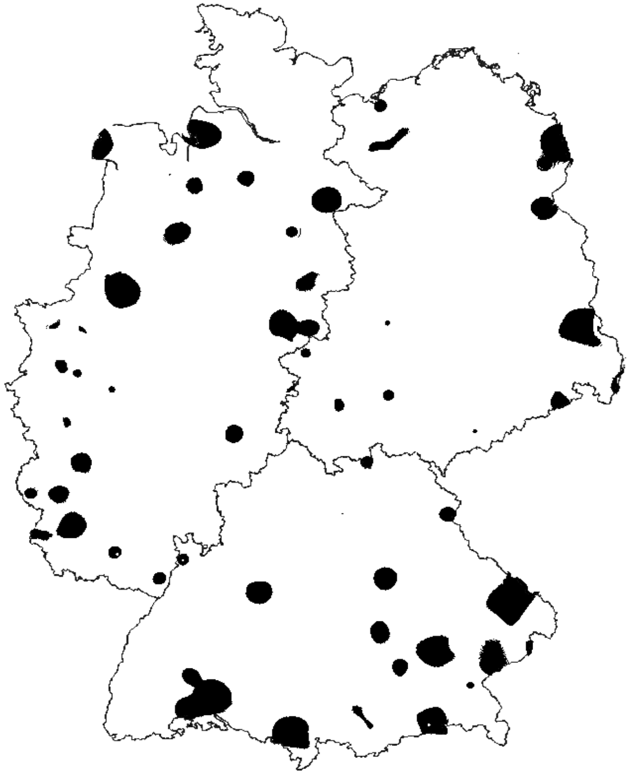
Explorative space-time analysis
Regional increment of mortality

Difference method: alternative method to construct binary images

- Pixel-wise difference of two mortality maps
- $T_1 - T_2$ and $T_2 - T_3$, where $T_1 = 1996 - 1998$, $T_2 = 1999 - 2001$, $T_3 = 2002 - 2003$
- Black pixels refer to a negative sign, i.e. increase of death rate, otherwise mortality improvement

Explorative space-time analysis

Regional increment of mortality - results

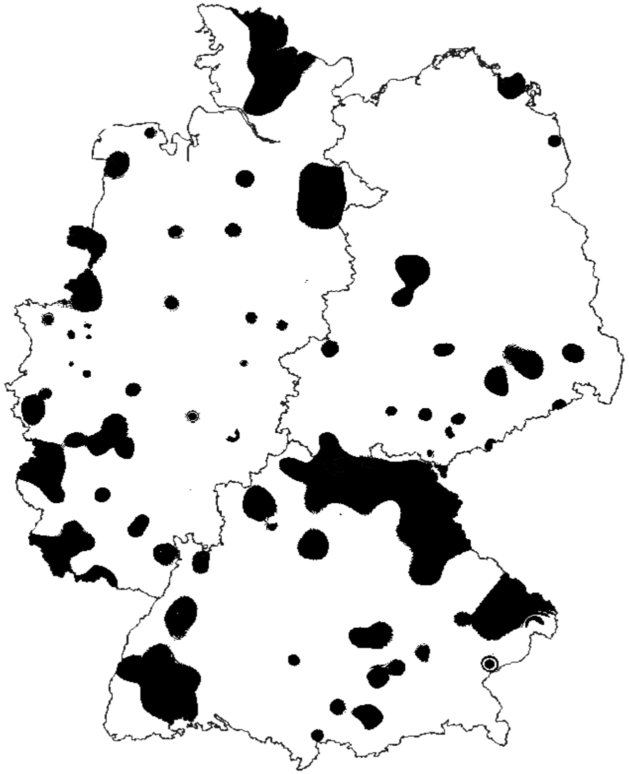


German population aged between 65 and 75 with respect to $T_1 - T_2$ *and* $T_2 - T_3$

Improvement of mortality: innovations in medicine and health care, improved welfare system

Explorative space-time analysis

Regional increment of mortality - results



German population aged 50 and younger with respect to $T_1 - T_2$ *and* $T_2 - T_3$

No further improvement of the mortality within the younger population

Application in life insurance pricing
Premium formula

- Initial cohort of size l_{x_0,I_j} all aged x_0 in region I_j, $j = 1, 2, 3$ (East, West, South)
- Term insurance which pays an amount P at death occurring up to a maximum age of $x_0 + N$
- Weighted mean death rate $\bar{m}_{x,I_j}$ in region I_j at age x, where

$$\bar{m}_{x,I_j} = \frac{1}{\sum_{u_i \in I_j} p_{x,T_j}(u_i)} \sum_{u_i \in I_j} p_{x,T_j}(u_i) m_{x,T_j}(u_i)$$

with size of population $p_{x,T_j}(u_i)$ and death rate $m_{x,T_j}(u_i)$ at measurement point u_i with age x in time horizon T_j

Application in life insurance pricing
Premium formula

- Expected number of deaths $\bar{d}_{x,I_j}$ at age x in region I_j

$$\bar{d}_{x,I_j} = l_{x,I_j} \cdot \bar{m}_{x,I_j},$$

where l_{x,I_j} is given by the recursion

$$l_{x,I_j} = l_{x-1,I_j} \cdot (1 - \bar{m}_{x-1,I_j}) \qquad \forall x = x_0 + 1, ..., x_0 + N$$

- Premium A_{x_0,I_j} calculated via the equivalence principle

$$A_{x_0,I_j} = P \cdot \sum_{x=x_0}^{x_0+N} (1+\rho)^{x_0-x-1} \cdot \frac{\bar{d}_{x,I_j}}{l_{x_0,I_j}}$$

where ρ denotes the risk free rate

Application in life insurance pricing
Numerical example

Computation of term–insurance premium A_{x_0,I_j} for various time horizons and regions

- Death rates are assumed to be constant for the next N years
- Initial cohorts all aged $x_0 = 30$ and $x_0 = 40$
- Maximum age $x_0 + 35$ years
- Payoff $P = 30000$
- Risk free rate $\rho = 0.0325$

Application in life insurance pricing

Numerical Example

	men			women		
	1996–1998	1999–2001	2002–2003	1996–1998	1999–2001	2002–2003
Initial cohort all aged $x_0 = 30$						
East	3168	2799	2651	1456	1255	1189
West	2626	2402	2318	1364	1298	1252
South	2413	2203	2068	1222	1143	1091
Initial cohort all aged $x_0 = 40$						
East	6206	6051	5764	3797	3166	3157
West	5657	5543	5320	3439	3108	3107
South	5243	5122	4830	3130	2793	2760

Premium A_{x_0,I_j} based on death rates for various time horizons and regions

Application in life insurance pricing

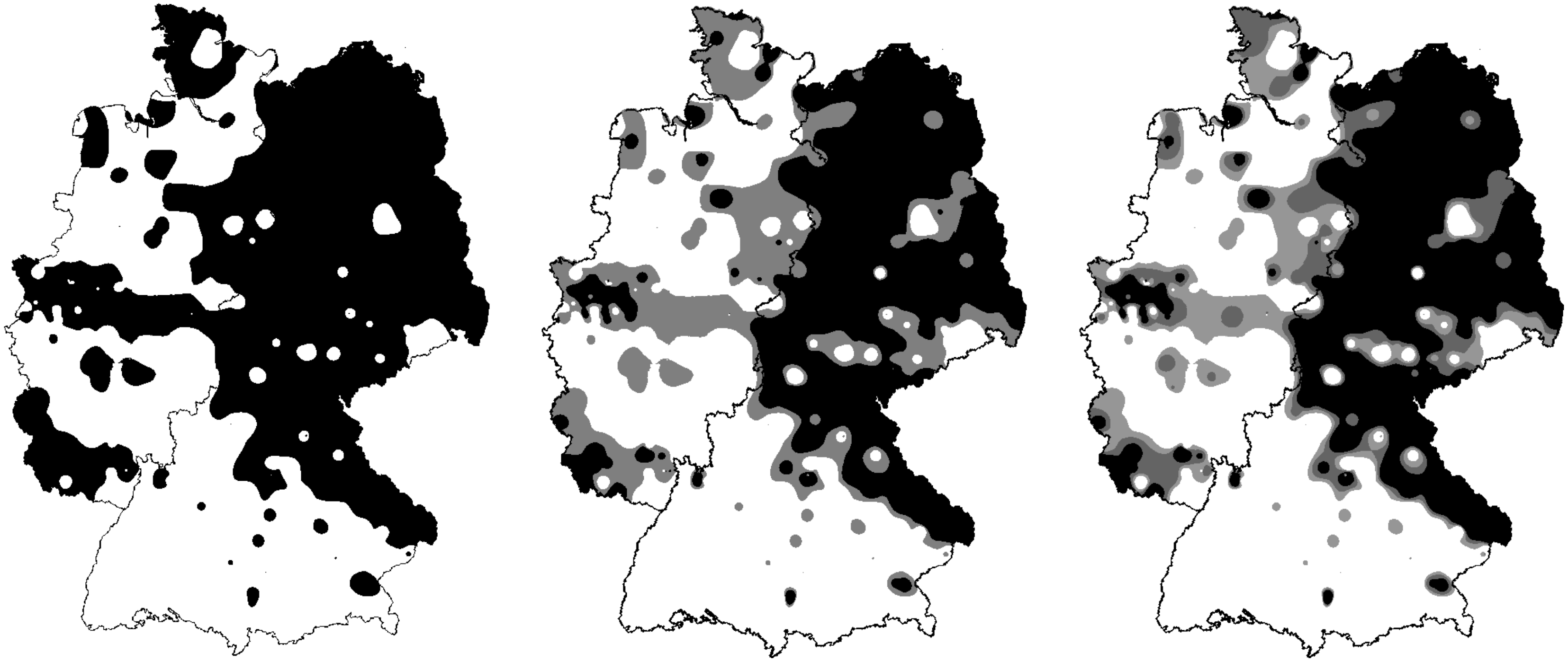
Conclusions

- Premiums for the male population are greater than the corresponding premiums for the female population.
- All premiums for a specific combination of age and region decrease from T_1 to T_3.
- Premiums for the male population in East Germany exceed the corresponding premiums in West and South Germany.
- Premiums for the female population in East Germany fall below the corresponding premiums in West Germany in T_2 and T_3 for initial cohort aged $x_0 = 30$.

Application in life insurance pricing
Conclusions

- Converging effect of East and West German mortality is more pronounced for the female population.
- Effects of mortality improvement are less pronounced with increasing age of the initial cohort for the female population in East Germany.
- Differences between South Germany and West Germany are increasing.

Construction of images with two and more grey levels by using several thresholds



Binary image *2 grey levels* *3 grey levels*

Outlook

Extensions

Model-based statistical analysis

Development and application of significance tests for intrinsic volumes of random closed sets

- Mortality differs significantly in two regions
- Mortality is significantly improving between two time periods
- Influence of parameters, e.g. extrapolation radius and cubical decay
- Extension to two and more grey level images

Outlook
Random closed sets

Construction of asymptotic Gaussian significance tests

- Black part of the image is a realization of a stationary random set Ξ in $\mathbb{R}^2$ observed in a sampling window W
- Area fraction p_Ξ of Ξ (expected area of Ξ per unit area)

$$p_\Xi = \mathbb{E}|\Xi \cap [0,1]^2|$$

- Unbiased estimator $\hat{p}_\Xi(W)$ for p_Ξ

$$\hat{p}_\Xi(W) = \frac{|\Xi \cap W|}{|W|}$$

is asymptotically normal distributed (if $|W| \to \infty$)

- Asymptotic variance $\sigma^2 = \lim_{|W|\to\infty} \sqrt{|W|}\mathrm{Var}\,\hat{p}_\Xi(W)$
- Estimator $\hat{\sigma}^2_\Xi(W)$ for σ^2

$$\hat{\sigma}^2_\Xi(W) = \int_{\mathbb{R}^2} \widehat{\mathrm{Cov}}_{\Xi,W}(x)\gamma_W(x)dx$$

  - Weighting function $\gamma_W(x) \geq 0$
  - Consistent estimator $\widehat{\mathrm{Cov}}_{\Xi,W}(x)$ for the covariance $\mathrm{Cov}(\mathrm{Y}_\Xi(\mathrm{o}), \mathrm{Y}_\Xi(\mathrm{x}))$ of the stationary random field $\{Y_\Xi(x), x \in \mathbb{R}^2\}$, where

$$Y_\Xi(x) = \begin{cases} 1 & \text{if } x \in \Xi \\ 0 & \text{else} \end{cases}$$

References
Demographical analysis

- Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit (2004) Bericht: Regionale Unterschiede der Sterblichkeit in Bayern. München.
- C. Gräb (1994) Sterbefälle 1993 nach Todesursachen. *Wirtschaft und Statistik* **12**, 1033-1041.
- B. Häussler, E. Hempel & P. Reschke (1995) Die Entwicklung von Lebenserwartung und Sterblichkeit in Ostdeutschland nach der Wende (1989-1992). *Öffentliches Gesundheitswesen* **57**, 365-372.
- C. Höhn & J. Pollard (1991) Mortality in the two Germanies in 1986 and trends 1976-1986. *European Journal of Population* **7**, 1-28.
- R.T. Riphahn & K.F. Zimmermann (1998) The mortality crisis in East Germany. In.: G.A. Cornia, R. Paniccia (eds), *The mortality crisis in transitional economics*. Oxford Press, Oxford, 227-252.

References
Statistical Analysis

- S. Böhm, L. Heinrich & V. Schmidt (2004) Asymptotic properties of estimators for the volume fractions of jointly stationary random sets. *Statistica Neerlandica* **58**, 388-406.
- P.A. Burrough (1986) *Principles of geographical information systems for land resource assessment*. Monographs on Soil and Resources Survey **12**, Clarendon Press, Oxford.
- T.A. Jones, D.E. Hamilton & C.R. Johnson (1986) *Contouring geologic surfaces with the computer*. Van Nostrand Reinhold, London.
- D. Stoyan, W.S. Kendall & J. Mecke (1995) *Stochastic geometry and its applications* (2nd ed.). J. Wiley & Sons, Chichester.

Thank you for your attention!