Statistical Emulators for Pricing and Hedging Longevity Risk Products

Jimmy Risk

August 6, 2015

What is the problem?

(i) Longevity risk is of growing importance

- Affects pension funds, life insurance companies

(ii) Stochastic mortality models are becoming more popular

- Combining (i) and (ii) creates a difficult problem (pricing, hedging, etc.)
  - Industry utilizes crude extrapolation and approximation methods

What is the problem?

(i) Longevity risk is of growing importance

- Affects pension funds, life insurance companies

(ii) Stochastic mortality models are becoming more popular

- Combining (i) and (ii) creates a difficult problem (pricing, hedging, etc.)
  - Industry utilizes crude extrapolation and approximation methods

What is the problem?

(i) Longevity risk is of growing importance
  - Affects pension funds, life insurance companies

(ii) Stochastic mortality models are becoming more popular

- Combining (i) and (ii) creates a difficult problem (pricing, hedging, etc.)
  - Industry utilizes crude extrapolation and approximation methods

Mathematical background to the problem

- Assume Markov state process $Z(\cdot)$ that captures evolution of mortality
- The time T present value of a $T-$year deferred annuity paying \$1 annually for an individual aged x with remaining lifetime $\tau(x)$ is

$$a(Z(T), T, x) \doteq \sum_{t=1}^{\infty} e^{-rt}\mathbb{E}\left[\mathbb{1}_{\{\tau(x)\geq t\}} \mid Z(T)\right] \qquad (1)$$

- Equation 1 depends on the mortality model.
  - $\mathbb{P}(\tau(x) \geq t \mid Z(T))$ is not available in closed form under any commonly used stochastic mortality model
  - $a(Z(T); T, x)$ needs to be accurately estimated!

Mathematical background to the problem

- Assume Markov state process $Z(\cdot)$ that captures evolution of mortality
- The time T present value of a $T-$year deferred annuity paying \$1 annually for an individual aged x with remaining lifetime $\tau(x)$ is

$$a(Z(T), T, x) \doteq \sum_{t=1}^{\infty} e^{-rt}\mathbb{E}\left[\mathbb{1}_{\{\tau(x)\geq t\}} \mid Z(T)\right] \qquad (1)$$

- Equation 1 depends on the mortality model.
  - $\mathbb{P}(\tau(x) \geq t \mid Z(T))$ is not available in closed form under any commonly used stochastic mortality model
  - $a(Z(T); T, x)$ needs to be accurately estimated!

Mathematical background to the problem

- Assume Markov state process $Z(\cdot)$ that captures evolution of mortality
- The time T present value of a $T-$year deferred annuity paying \$1 annually for an individual aged x with remaining lifetime $\tau(x)$ is

$$a(Z(T), T, x) \doteq \sum_{t=1}^{\infty} e^{-rt}\mathbb{E}\left[\mathbb{1}_{\{\tau(x)\geq t\}} \mid Z(T)\right] \qquad (1)$$

- Equation 1 depends on the mortality model.
  - $\mathbb{P}(\tau(x) \geq t \mid Z(T))$ is not available in closed form under any commonly used stochastic mortality model
  - $a(Z(T); T, x)$ needs to be accurately estimated!

Mathematical background to the problem

- Assume Markov state process $Z(\cdot)$ that captures evolution of mortality
- The time T present value of a $T-$year deferred annuity paying \$1 annually for an individual aged x with remaining lifetime $\tau(x)$ is

$$a(Z(T), T, x) \doteq \sum_{t=1}^{\infty} e^{-rt}\mathbb{E}\left[\mathbb{1}_{\{\tau(x)\geq t\}} \mid Z(T)\right] \qquad (1)$$

- Equation 1 depends on the mortality model.
  - $\mathbb{P}(\tau(x) \geq t \mid Z(T))$ is not available in closed form under any commonly used stochastic mortality model
  - $a(Z(T); T, x)$ needs to be accurately estimated!

Ways to evaluate $\mathbb{E}[a(Z(T), T, x)]$

(i) *Nested Monte Carlo*: simulate trajectories of $Z(T)$ and simulate $a(Z(T), T, x)$ given each realization.

(ii) *Deterministic projection*: Use Taylor series expansion or similar to develop an analytic estimate for $\mathbb{P}(\tau(x) \geq t \mid Z(T))$.

(iii) *Statistical emulator*: Train a model with a design $(z^1, \ldots, z^n)$ by estimating $a(Z(T), T, x)\mid_{Z(T)=z^i}, i = 1, \ldots, n$ through Monte Carlo.

- (ii) and (iii) develop intermediate functionals that estimate

$$\hat{f}(z) \approx \mathbb{E}[a(Z(T), T, x) \mid Z(T) = z]$$

  - Final value $\mathbb{E}[a(Z(T), T, x)]$ is determined through Monte Carlo

Ways to evaluate $\mathbb{E}[a(Z(T), T, x)]$

(i) *Nested Monte Carlo*: simulate trajectories of $Z(T)$ and simulate $a(Z(T), T, x)$ given each realization.

(ii) *Deterministic projection*: Use Taylor series expansion or similar to develop an analytic estimate for $\mathbb{P}(\tau(x) \geq t \mid Z(T))$.

(iii) *Statistical emulator*: Train a model with a design $(z^1, \ldots, z^n)$ by estimating $a(Z(T), T, x) \mid_{Z(T)=z^i}, i = 1, \ldots, n$ through Monte Carlo.

- (ii) and (iii) develop intermediate functionals that estimate

$$\hat{f}(z) \approx \mathbb{E}[a(Z(T), T, x) \mid Z(T) = z]$$

  - Final value $\mathbb{E}[a(Z(T), T, x)]$ is determined through Monte Carlo

What is statistical emulation?

- Statistical emulation deals with a sampler

$$Y(z) = f(z) + \epsilon(z), \tag{2}$$

where f is the unknown response surface and ϵ is the sampling noise.

- Examples of f include:
  - $T-$year deferred annuity: $f(z) = \mathbb{E}[a(Z(T), T, x) \mid Z(T) = z]$.
  - Quantile $q(\alpha, z)$ (Value-at-Risk)
  - Correlation between two functionals, $Corr(F_1(T, Z(\cdot)), F_2(T, Z(\cdot)))$

What is statistical emulation?

- Statistical emulation deals with a sampler

$$Y(z) = f(z) + \epsilon(z), \tag{2}$$

where f is the unknown response surface and ϵ is the sampling noise.

- Examples of f include:
  - $T-$year deferred annuity: $f(z) = \mathbb{E}[a(Z(T), T, x) \mid Z(T) = z]$.
  - Quantile $q(\alpha, z)$ (Value-at-Risk)
  - Correlation between two functionals, $Corr(F_1(T, Z(\cdot)), F_2(T, Z(\cdot)))$

Fitting process for statistical emulation

- Goal:
  - Represent state process $Z(T)$ with a design $\mathcal{D} = \{z^1, \dots, z^N\}$
  - For each z^i, produce realizations $\{y^1, \dots, y^N\}$ of (2)
  - Use pairs $(z^i, y^i)_{i=1}^N$ to construct a fitted response surface $\hat{f}$.
- Possible frameworks:
  - Kernel regressions
  - Splines
  - Kriging (Gaussian processes)

Fitting process for statistical emulation

- Goal:
  - Represent state process $Z(T)$ with a design $\mathcal{D} = \{z^1, \ldots, z^N\}$
  - For each z^i, produce realizations $\{y^1, \ldots, y^N\}$ of (2)
  - Use pairs $(z^i, y^i)_{i=1}^N$ to construct a fitted response surface $\hat{f}$.
- Possible frameworks:
  - Kernel regressions
  - Splines
  - Kriging (Gaussian processes)

Fitting process for statistical emulation

- Goal:
  - Represent state process $Z(T)$ with a design $\mathcal{D} = \{z^1, \ldots, z^N\}$
  - For each z^i, produce realizations $\{y^1, \ldots, y^N\}$ of (2)
  - Use pairs $(z^i, y^i)_{i=1}^N$ to construct a fitted response surface $\hat{f}$.
- Possible frameworks:
  - Kernel regressions
  - Splines
  - Kriging (Gaussian processes)

Fitting process for statistical emulation

- Goal:
  - Represent state process $Z(T)$ with a design $\mathcal{D} = \{z^1, \ldots, z^N\}$
  - For each z^i, produce realizations $\{y^1, \ldots, y^N\}$ of (2)
  - Use pairs $(z^i, y^i)_{i=1}^N$ to construct a fitted response surface $\hat{f}$.
- Possible frameworks:
  - Kernel regressions
  - Splines
  - Kriging (Gaussian processes)

How to deterine the design $\mathcal{D}$

- Design $\mathcal{D}$ should correctly describe $Z(T)$
  - Can be catered to the problem at hand
    - Example: VaR vs expectation
  - Should accurately reflect correlation structure
- Can be determined by
  - Simulation
  - Uniformly spaced grid
  - Pseudo-random grid (e.g. Latin hypercube, Sobol sequence)
  - Weighted grid

How to deterine the design $\mathcal{D}$

- Design $\mathcal{D}$ should correctly describe $Z(T)$
  - Can be catered to the problem at hand
    - Example: VaR vs expectation
  - Should accurately reflect correlation structure
- Can be determined by
  - Simulation
  - Uniformly spaced grid
  - Pseudo-random grid (e.g. Latin hypercube, Sobol sequence)
  - Weighted grid

Smoothing Splines

- Given design $\mathcal{D} = (z^1, \ldots, z^N)$ and paired response $(y^1, \ldots, y^N)$ with $z^i, y^i \in \mathbb{R}$
  - Minimize penalized residual sum of squares

$$\sum_{i=1}^{n} \left(y^i - f(z^i)\right)^2 + \lambda \int \left(f''(u)\right)^2 du \qquad (3)$$

  - Constraint: f', f'' continuous
- $\lambda \geq 0$ is *smoothing parameter*
- Can be extended to $z^i, y^i \in \mathbb{R}^d$
  - Called *Thin Plate Spline*
  - Replace integral in (3) with $\mathbb{R}^d$ penalty function

Smoothing Splines

- Given design $\mathcal{D} = (z^1, \ldots, z^N)$ and paired response $(y^1, \ldots, y^N)$ with $z^i, y^i \in \mathbb{R}$
  - Minimize penalized residual sum of squares

$$\sum_{i=1}^{n} \left(y^i - f(z^i)\right)^2 + \lambda \int \left(f''(u)\right)^2 du \qquad (3)$$

  - Constraint: f', f'' continuous
- $\lambda \geq 0$ is *smoothing parameter*
- Can be extended to $z^i, y^i \in \mathbb{R}^d$
  - Called *Thin Plate Spline*
  - Replace integral in (3) with $\mathbb{R}^d$ penalty function

Mathematical background for kriging

- Consider f as a random field $(f(z))_{z\in\mathbb{R}^d}$
- Given $\mathcal{D} = (z^1, \ldots, z^N)$
  - Access to noisy observations $\mathbf{y} = (y^1, \ldots, y^N)$
  - y^i are draws from the process

$$Y(z) = f(z) + \epsilon(z), \qquad \epsilon(z) \sim N(0, \tau(z))$$

- Goal: Make predictions using $f(z)|Y(\mathcal{D}) = \mathbf{y}$ for new z

Mathematical background for kriging

- Consider f as a random field $(f(z))_{z\in\mathbb{R}^d}$
- Given $\mathcal{D} = (z^1, \dots, z^N)$
  - Access to noisy observations $\mathbf{y} = (y^1, \dots, y^N)$
  - y^i are draws from the process

$$Y(z) = f(z) + \epsilon(z), \qquad \epsilon(z) \sim N(0, \tau(z))$$

- Goal: Make predictions using $f(z)|Y(\mathcal{D}) = \mathbf{y}$ for new z

Kriging model details

- *Kriging* assumes

$$f(z) = \mu(z) + X(z)$$

- μ is a trend function
- X is centered square integrable process
  - X has known covariance kernel C
  - If X is Gaussian,

$$f(z)|Y(\mathcal{D}) = \mathbf{y} \sim N(m_{SK}(z), s^2_{SK}(z))$$

where $m_{SK}(z)$ and $s^2_{SK}(z)$ depend on $\mathcal{D}, \mathbf{y}, \mu, \tau(\mathcal{D})$

Kriging model details

- *Kriging* assumes

$$f(z) = \mu(z) + X(z)$$

- μ is a trend function
- X is centered square integrable process
  - X has known covariance kernel C
  - If X is Gaussian,

$$f(z)|Y(\mathcal{D}) = \mathbf{y} \sim N(m_{SK}(z), s_{SK}^2(z))$$

where $m_{SK}(z)$ and $s_{SK}^2(z)$ depend on $\mathcal{D}, \mathbf{y}, \mu, \tau(\mathcal{D})$

Why we should consider kriging

- Nonparametric regression tool
- Combines trend and flexible residual modeling
- Trend function can be pre-specified ("Simple Kriging") or estimated ("Universal Kriging")
- Widely used in simulation literature
- Easy to implement (R package `DiceKriging`)
- Bayesian framework provides posterior credible intervals to understand model accuracy

Why we should consider kriging

- Nonparametric regression tool
- Combines trend and flexible residual modeling
- Trend function can be pre-specified ("Simple Kriging") or estimated ("Universal Kriging")
- Widely used in simulation literature
- Easy to implement (`R` package `DiceKriging`)
- Bayesian framework provides posterior credible intervals to understand model accuracy

Kriging Illustration

$\mathcal{D} = \{-1, -0.5, 0, 0.5, 1\}$ $\mathbf{y} = \{-9, -5, -1, 9, 11\}$ $\sigma(\mathcal{D}) = \{0.1, 0.5, 2, 4, 8\}$

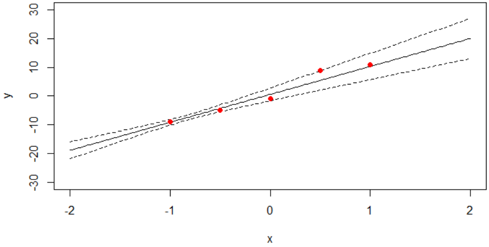

Figure: Bayesian credibility bands under the above setup. Fit assuming first order linear trend. Red dots are training points.

Case Studies

Case Studies

Analysis Overview

Case studies:

- 10-year deferred annuity hedge portfolio analysis under a two-population Lee-Carter model
- 20-year deferred annuity evaluation using the CBD framework

10-Year Deferred Annuity Hedge Portfolio Problem

Two population hedge portfolio

- Insured population dynamics should be different from the general population
- If a tradeble mortality index were available, how effective could a hedge be?
- Goal: predict hedge portfolio values

10-Year Deferred Annuity Hedge Portfolio Problem

Two population hedge portfolio

- Insured population dynamics should be different from the general population
- If a tradeble mortality index were available, how effective could a hedge be?
- Goal: predict hedge portfolio values

10-Year Deferred Annuity Hedge Portfolio Problem

Two population hedge portfolio

- Insured population dynamics should be different from the general population
- If a tradeble mortality index were available, how effective could a hedge be?
- Goal: predict hedge portfolio values

Case study data and model

Following Cairns et al. [2014]

- Ages 50–89, Years 1961–2005
- "General Population" data is represented by England & Wales male mortality data
- "Insured Population" data is represented by Continuous Mortality Investigation (CMI) male mortality data
  - CMI produces a life table with data supplied by private UK life insurance companies and actuarial consultancies
- Case study uses cointegrated two-population Lee Carter model from Cairns et al. [2011]

Case study data and model

Following Cairns et al. [2014]

- Ages 50–89, Years 1961–2005
- “General Population” data is represented by England & Wales male mortality data
- “Insured Population” data is represented by Continuous Mortality Investigation (CMI) male mortality data
  - CMI produces a life table with data supplied by private UK life insurance companies and actuarial consultancies
- Case study uses cointegrated two-population Lee Carter model from Cairns et al. [2011]

Cointegrated Two-Population Model

Following Cairns et al. [2014]

- Both the general population (index 1) and insured subpopulation (index 2) follow Lee-Carter with cohort effect

$$\log m_i(t,x) = \beta_i^{(1)}(x) + \beta_i^{(2)}(x)\kappa_i^{(2)}(t) + \beta_i^{(3)}(x)\gamma_i^{(3)}(t-x), i = 1,2$$

- $\kappa_1^{(2)}$ is random walk with drift
- Define $S(t) \doteq \kappa_1^{(2)}(t) - \kappa_2^{(2)}(t)$. Then $\kappa_2^{(2)}$ is determined through the AR process

$$S(t) = \mu_2 + \phi(S(t-1) - \mu_2) + \sigma_2\epsilon_2(t-1) + c\epsilon_1(t-1)$$

- $\epsilon_2(\cdot) \overset{iid}{\sim} N(0,1)$ independent of $\epsilon_1(\cdot)$
- $\epsilon_1(\cdot) \overset{iid}{\sim} N(0,1)$ is the noise term in $\kappa_1^{(2)}(t)$

Cointegrated Two-Population Model

Following Cairns et al. [2014]

- Both the general population (index 1) and insured subpopulation (index 2) follow Lee-Carter with cohort effect

$$\log m_i(t,x) = \beta_i^{(1)}(x) + \beta_i^{(2)}(x)\kappa_i^{(2)}(t) + \beta_i^{(3)}(x)\gamma_i^{(3)}(t-x), i = 1,2$$

- $\kappa_1^{(2)}$ is random walk with drift
- Define $S(t) \doteq \kappa_1^{(2)}(t) - \kappa_2^{(2)}(t)$. Then $\kappa_2^{(2)}$ is determined through the AR process

$$S(t) = \mu_2 + \phi(S(t-1) - \mu_2) + \sigma_2\epsilon_2(t-1) + c\epsilon_1(t-1)$$

- $\epsilon_2(\cdot) \overset{iid}{\sim} N(0,1)$ independent of $\epsilon_1(\cdot)$
- $\epsilon_1(\cdot) \overset{iid}{\sim} N(0,1)$ is the noise term in $\kappa_1^{(2)}(t)$

Details for evaluating hedge portfolio values

- Deferral period $T = 10$ years
- Begin receiving payments at age $x = 65$
- Models refit at time T to reflect "parameter partial certain" case [Cairns et al. [2014]]
  - State process $Z(T)$ is four dimensional including period effects and (significant) refit parameters

$$Z(T) = \{\kappa_1^{(2)}(T), \kappa_2^{(2)}(T), \mu_2, \phi\}$$

Details for evaluating hedge portfolio values

- Deferral period $T = 10$ years
- Begin receiving payments at age $x = 65$
- Models refit at time T to reflect "parameter partial certain" case [Cairns et al. [2014]]
  - State process $Z(T)$ is four dimensional including period effects and (significant) refit parameters

$$Z(T) = \{\kappa_1^{(2)}(T), \kappa_2^{(2)}(T), \mu_2, \phi\}$$

Methods compared through the case study

- Estimation methods:
  - Analytic Estimate
  - Thin Plate Spline
  - 1st order linear Universal Kriging
  - Simple Kriging
    - Uses analytic estimate as drift
- Training set size (N_{tr}) effect
  - $N_{tr} = 1000$
  - $N_{tr} = 8000$

Methods compared through the case study

- Estimation methods:
  - Analytic Estimate
  - Thin Plate Spline
  - 1st order linear Universal Kriging
  - Simple Kriging
    - Uses analytic estimate as drift
- Training set size (N_{tr}) effect
  - $N_{tr} = 1000$
  - $N_{tr} = 8000$

Methods compared through the case study

- Estimation methods:
  - Analytic Estimate
  - Thin Plate Spline
  - 1st order linear Universal Kriging
  - Simple Kriging
    - Uses analytic estimate as drift
- Training set size (N_{tr}) effect
  - $N_{tr} = 1000$
  - $N_{tr} = 8000$

Details behind the analytic estimate ("Industry Standard")

- Based on Cairns et al. [2014]
- Find $\mathbb{E}[m(T+t,x) \mid Z(T)], i = 1,2$ as a function of $Z(T)$ and t
- The one year survival probability for a person aged x in year t is

$$\mathbb{E}[\exp(-m(t,x))] \approx \exp(-\mathbb{E}[m(t,x)])$$

- We model $\log m(t,x)$, so an additional level of approximating via exponentiation is required

Details behind the analytic estimate ("Industry Standard")

- Based on Cairns et al. [2014]
- Find $\mathbb{E}[m(T+t,x) \mid Z(T)], i = 1, 2$ as a function of $Z(T)$ and t
- The one year survival probability for a person aged x in year t is

$$\mathbb{E}[\exp(-m(t,x))] \approx \exp(-\mathbb{E}[m(t,x)])$$

- We model $\log m(t,x)$, so an additional level of approximating via exponentiation is required

Results of two-population hedge case study

	$N_{tr} = 1000$		$N_{tr} = 8000$	
Type	Bias	MSE	Bias	MSE
Analytic	4.480e-03	2.831e-05	4.480e-03	2.831e-05
Thin Plate Spline	2.577e-03	1.701e-04	5.803e-04	2.596e-05
Universal Kriging	4.363e-04	3.446e-04	1.857e-03	1.662e-04
Simple Kriging	-1.334e-03	1.076e-05	9.390e-04	9.262e-06

Table: Monte Carlo averages based on 1000 simulations of $Z(T)$

- Simple Kriging performs best
- Training set size effect is apparent

Results of two-population hedge case study

	$N_{tr} = 1000$		$N_{tr} = 8000$	
Type	Bias	MSE	Bias	MSE
Analytic	4.480e-03	2.831e-05	4.480e-03	2.831e-05
Thin Plate Spline	2.577e-03	1.701e-04	5.803e-04	2.596e-05
Universal Kriging	4.363e-04	3.446e-04	1.857e-03	1.662e-04
Simple Kriging	-1.334e-03	1.076e-05	9.390e-04	9.262e-06

Table: Monte Carlo averages based on 1000 simulations of $Z(T)$

- Simple Kriging performs best
- Training set size effect is apparent

Results of two-population hedge case study

	$N_{tr} = 1000$		$N_{tr} = 8000$	
Type	Bias	MSE	Bias	MSE
Analytic	4.480e-03	2.831e-05	4.480e-03	2.831e-05
Thin Plate Spline	2.577e-03	1.701e-04	5.803e-04	2.596e-05
Universal Kriging	4.363e-04	3.446e-04	1.857e-03	1.662e-04
Simple Kriging	-1.334e-03	1.076e-05	9.390e-04	9.262e-06

Table: Monte Carlo averages based on 1000 simulations of $Z(T)$

- Simple Kriging performs best
- Training set size effect is apparent

Comments on two-population hedge case study

- Portfolio values are large in practice
  - A portfolio of $1,000,000 would yield an error of $4,480 in using the analytic estimate
- No way to recognize apriori the performance of the analytic estimate
  - Bias may have been subtracted in differencing process

Comments on two-population hedge case study

- Portfolio values are large in practice
  - A portfolio of $1,000,000 would yield an error of $4,480 in using the analytic estimate
- No way to recognize apriori the performance of the analytic estimate
  - Bias may have been subtracted in differencing process

Comments on two-population hedge case study

- Portfolio values are large in practice
  - A portfolio of $1,000,000 would yield an error of $4,480 in using the analytic estimate
- No way to recognize apriori the performance of the analytic estimate
  - Bias may have been subtracted in differencing process

Comments on two-population hedge case study

- Portfolio values are large in practice
  - A portfolio of $1,000,000 would yield an error of $4,480 in using the analytic estimate
- No way to recognize apriori the performance of the analytic estimate
  - Bias may have been subtracted in differencing process

Model for CBD annuity valuation model

- Fit CMI data to the CBD model [Cairns et al., 2006]

$$\operatorname{logit} q(t,x) = \kappa^{(1)}(t) + (x - \bar{x})\kappa^{(2)}(t)$$

- Following Cairns et al. [2009]
  - $\kappa^{(1)}(t)$ and $\kappa^{(2)}(t)$ are period effects (time series fit using `auto.arima` in `R`)
- Auto-regressive time series yields

$$Z(T) = \{\kappa^{(1)}(T), \kappa^{(2)}(T-1), \kappa^{(2)}(T)\}$$

- Key difference from previous case study: we model survival probabilities and not mortality rates

Model for CBD annuity valuation model

- Fit CMI data to the CBD model [Cairns et al., 2006]

$$\operatorname{logit} q(t,x) = \kappa^{(1)}(t) + (x - \bar{x})\kappa^{(2)}(t)$$

- Following Cairns et al. [2009]
  - $\kappa^{(1)}(t)$ and $\kappa^{(2)}(t)$ are period effects (time series fit using `auto.arima` in `R`)
- Auto-regressive time series yields

$$Z(T) = \{\kappa^{(1)}(T), \kappa^{(2)}(T-1), \kappa^{(2)}(T)\}$$

- Key difference from previous case study: we model survival probabilities and not mortality rates

Outline of CBD annuity case study

- Value 20-year deferred annuities beginning payments at age 65.
- Analytic estimator is derived similarly as in the two-population study
- Surrogate models are
  - Thin plate spline
  - Ordinary kriging
  - 1st-order universal kriging

Outline of CBD annuity case study

- Value 20-year deferred annuities beginning payments at age 65.
- Analytic estimator is derived similarly as in the two-population study
- Surrogate models are
  - Thin plate spline
  - Ordinary kriging
  - 1st-order universal kriging

Outline of CBD annuity case study

- Value 20-year deferred annuities beginning payments at age 65.
- Analytic estimator is derived similarly as in the two-population study
- Surrogate models are
  - Thin plate spline
  - Ordinary kriging
  - 1st-order universal kriging

Results of CBD annuity case study

	$N_{tr} = 1000$		$N_{tr} = 8000$	
Type	Bias	MSE	Bias	MSE
Analytic	-4.560e-01	2.764e-01	-4.560e-01	2.764e-01
TPS	-2.358e-02	4.515e-03	4.195e-03	2.955e-03
OK	3.669e-03	9.575e-03	9.734e-03	5.996e-03
1st-Order UK	-1.785e-03	3.415e-03	5.635e-03	1.897e-03

- Longer deferral period reduces effectiveness of analytic estimate
- Training set size effect is slower to converge

Concluding Remarks

Concluding Remarks

Concluding Remarks

- Attacked problem of pricing deferred life annuities
  - Used real data
  - Utilized commonly used mortality models
  - Easy to implement method
  - Outperformed "industry standard"
    - Case studies used drastically different mortality models

Concluding Remarks

- Attacked problem of pricing deferred life annuities
  - Used real data
  - Utilized commonly used mortality models
  - Easy to implement method
  - Outperformed “industry standard”
    - Case studies used drastically different mortality models

Concluding Remarks

- Attacked problem of pricing deferred life annuities
  - Used real data
  - Utilized commonly used mortality models
  - Easy to implement method
  - Outperformed "industry standard"
    - Case studies used drastically different mortality models

Concluding Remarks

- Attacked problem of pricing deferred life annuities
  - Used real data
  - Utilized commonly used mortality models
  - Easy to implement method
  - Outperformed "industry standard"
    - Case studies used drastically different mortality models

Concluding Remarks

- Attacked problem of pricing deferred life annuities
  - Used real data
  - Utilized commonly used mortality models
  - Easy to implement method
  - Outperformed "industry standard"
    - Case studies used drastically different mortality models

Further Work

- Extend to more general problem where input $(Z(T))$ includes
  - Age
  - Deferral period (in the case of annuity)
  - Time 0 parameters
  - Interest rate
- Different mortality assumptions

Further Work

- Extend to more general problem where input $(Z(T))$ includes
  - Age
  - Deferral period (in the case of annuity)
  - Time 0 parameters
  - Interest rate
- Different mortality assumptions

Bibliography

Andrew JG Cairns, David Blake, and Kevin Dowd. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718, 2006.

Andrew JG Cairns, David Blake, Kevin Dowd, Guy D Coughlan, David Epstein, Alen Ong, and Igor Balevich. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1):1–35, 2009.

Andrew JG Cairns, David Blake, Kevin Dowd, Guy D Coughlan, and Marwa Khalaf-Allah. Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin*, 41(01):29–59, 2011.

Andrew JG Cairns, Kevin Dowd, David Blake, and Guy D Coughlan. Longevity hedge effectiveness: A decomposition. *Quantitative Finance*, 14(2):217–235, 2014.

Questions?

- Paper Available

Statistical Emulators for Pricing and Hedging Longevity Risk Products
James Risk, Michael Ludkovski
http://arxiv.org/abs/1508.00310