The course of well-being over the life span - Restricted Likelihood Ratio Testing (RLRT) in the presence of correlated errors

Andrea Wiencierz in cooperation with:
Sonja Greven and Helmut Küchenhoff

Department for Statistics, LMU Munich

June 30, 2010

Contents

"Happiness economics"

Question of interest:

"How does (subjective) well-being change over the life cycle?"

“Happiness economics”

Question of interest:

> *“How does (subjective) well-being change over the life cycle?”*

Previous work:

- Consulting project *“Well-being and Age II”*:
  several parametric analyses of SOEP data and construction of a semiparametric model,
- Discussion paper *“Well-being over the life span: semiparametric evidence from British and German longitudinal data”*:
  application of semiparametric models to BHPS and SOEP data.

"Happiness economics"

Question of interest:

> *"How does (subjective) well-being change over the life cycle?"*

Previous work:

- Consulting project *"Well-being and Age II"*:
  several parametric analyses of SOEP data and construction of a semiparametric model,
- Discussion paper *"Well-being over the life span: semiparametric evidence from British and German longitudinal data"*:
  application of semiparametric models to BHPS and SOEP data.

$\Rightarrow$ Lack of a valid test to confirm the findings.

Linear mixed model (LMM)

Simple linear mixed model for longitudinal data of individuals $i = 1, \ldots, n$ with n_i observations:

$$
\begin{aligned}
\mathbf{y_i} &= \mathbf{X_i}\,\beta + \mathbf{Z_i}\,\mathbf{b_i} + \epsilon_\mathbf{i}, \\
\mathbf{b_i} &\sim \mathrm{N}(\mathbf{0},\,\tilde{\mathbf{D}}), \quad \text{where } \tilde{\mathbf{D}} = \sigma_b\,\mathbf{I_K}, \\
\epsilon_\mathbf{i} &\sim \mathrm{N}(\mathbf{0},\,\mathbf{R_i}), \quad \text{where } \mathbf{R_i} = \sigma_\epsilon\,\mathbf{I_{n_i}}, \\
\mathbf{b_i}, \mathbf{b_j} &\text{ and } \epsilon_\mathbf{i}, \epsilon_\mathbf{j} \quad \text{independent for all } i \neq j, \\
\mathbf{b_i}, \epsilon_\mathbf{j} &\quad \text{independent for all } i, j.
\end{aligned}
$$

Linear mixed model (LMM)

Simple linear mixed model for longitudinal data of individuals $i = 1, \ldots, n$ with n_i observations:

$$\begin{aligned}
\mathbf{y_i} &= \mathbf{X_i}\,\boldsymbol{\beta} + \mathbf{Z_i}\,\mathbf{b_i} + \boldsymbol{\epsilon}_\mathbf{i}, \\
\mathbf{b_i} &\sim \mathrm{N}(\mathbf{0},\, \tilde{\mathbf{D}}), \quad \text{where } \tilde{\mathbf{D}} = \sigma_b\,\mathbf{I_K}, \\
\boldsymbol{\epsilon}_\mathbf{i} &\sim \mathrm{N}(\mathbf{0},\, \mathbf{R_i}), \quad \text{where } \mathbf{R_i} = \sigma_\epsilon\,\mathbf{I_{n_i}}, \\
\mathbf{b_i}, \mathbf{b_j} &\text{ and } \boldsymbol{\epsilon}_\mathbf{i}, \boldsymbol{\epsilon}_\mathbf{j} \quad \text{independent for all } i \neq j, \\
\mathbf{b_i}, \boldsymbol{\epsilon}_\mathbf{j} & \quad \text{independent for all } i, j.
\end{aligned}$$

Corresponding global model:

$$\begin{aligned}
\mathbf{y} &= \mathbf{X}\,\boldsymbol{\beta} + \mathbf{Z}\,\mathbf{b} + \boldsymbol{\epsilon}, \\
\mathbf{b} &\sim \mathrm{N}(\mathbf{0},\, \mathbf{D}), \quad \text{where } \mathbf{D} = diag(\tilde{\mathbf{D}}) = \sigma_b\,\mathbf{I_{n\,K}}, \\
\boldsymbol{\epsilon} &\sim \mathrm{N}(\mathbf{0},\, \mathbf{G}), \quad \text{where } \mathbf{G} = diag(\mathbf{G_i}) = \sigma_\epsilon\,\mathbf{I_N}, \\
\mathbf{b}, \boldsymbol{\epsilon} & \quad \text{independent.}
\end{aligned}$$

General form linear mixed model

General formulation of a mixed model for $i = 1, \dots, N$ data points:

$$\begin{aligned}
\mathbf{y} &= \mathbf{X}\,\beta + \mathbf{Z_1}\,\mathbf{b_1} + \ldots + \mathbf{Z_S}\,\mathbf{b_S} + \epsilon, \\
\mathbf{b} &\sim \mathrm{N}(\mathbf{0},\, \sigma_s\,\mathbf{\Sigma_s}), \quad s = 1, \dots, S, \\
\epsilon &\sim \mathrm{N}(\mathbf{0},\, \sigma_\epsilon\,\mathbf{I_N}), \\
\mathbf{b_1}, \dots, \mathbf{b_S}, \epsilon & \quad \text{independent.}
\end{aligned}$$

General form linear mixed model

General formulation of a mixed model for $i = 1, \ldots, N$ data points:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\,\beta + \mathbf{Z_1}\,\mathbf{b_1} + \ldots + \mathbf{Z_S}\,\mathbf{b_S} + \epsilon, \\ \mathbf{b} &\sim \mathrm{N}(\mathbf{0},\, \sigma_s\,\mathbf{\Sigma_s}), \quad s = 1, \ldots, S, \\ \epsilon &\sim \mathrm{N}(\mathbf{0},\, \sigma_\epsilon\,\mathbf{I_N}), \\ \mathbf{b_1}, \ldots, \mathbf{b_S}, \epsilon & \quad \text{independent.} \end{aligned}$$

Nonparametric model (with cubic TP basis) as a mixed model:

$$y_i = f(x_i) + \epsilon_i = \sum_{j=1}^{K} \gamma_j\, B_j(x_i) + \epsilon_i$$

General form linear mixed model

General formulation of a mixed model for $i = 1, \dots, N$ data points:

$$
\begin{aligned}
\mathbf{y} &= \mathbf{X}\,\beta + \mathbf{Z_1}\,\mathbf{b_1} + \ldots + \mathbf{Z_S}\,\mathbf{b_S} + \epsilon,\\
\mathbf{b} &\sim \mathrm{N}(\mathbf{0},\, \sigma_s\,\mathbf{\Sigma_s}), \quad s = 1, \dots, S,\\
\epsilon &\sim \mathrm{N}(\mathbf{0},\, \sigma_\epsilon\,\mathbf{I_N}),\\
\mathbf{b_1}, \dots, \mathbf{b_S}, \epsilon &\quad \text{independent.}
\end{aligned}
$$

Nonparametric model (with cubic TP basis) as a mixed model:

$$
\begin{aligned}
y_i = f(x_i) + \epsilon_i &= \sum_{j=1}^{K} \gamma_j\, B_j(x_i) + \epsilon_i\\
&= \underbrace{\gamma_0 + \gamma_1\, x_i + \gamma_2\, x_i^2 + \gamma_3\, x_i^3}_{=:\ \mathbf{x_i}\,\beta} + \underbrace{\sum_{j=1}^{K} \gamma_{j+3}\,(x_i - \kappa_j)_+^3}_{=:\ \mathbf{z_i}\,\mathbf{b}} + \epsilon_i
\end{aligned}
$$

Testing for zero variance components

We want to test:

$$H_0 : b_1 = \ldots = b_K = 0 \quad \text{vs.} \quad H_1 : b_j \neq 0 \text{ for one } j.$$

This is equivalent to testing:

$$H_0 : \sigma_b = 0 \quad \text{vs.} \quad H_1 : \sigma_b > 0.$$

Testing for zero variance components

We want to test:

$$H_0 : b_1 = \ldots = b_K = 0 \quad \text{vs.} \quad H_1 : b_j \neq 0 \text{ for one } j.$$

This is equivalent to testing:

$$H_0 : \sigma_b = 0 \quad \text{vs.} \quad H_1 : \sigma_b > 0.$$

Violation of assumptions of usual asymptotics for Likelihood Ratio tests:

- the tested parameter is on the boundary of the parameter space,
- observations are not i.i.d.

Testing for zero variance components

We want to test:

$$H_0 : b_1 = \ldots = b_K = 0 \quad \text{vs.} \quad H_1 : b_j \neq 0 \text{ for one } j.$$

This is equivalent to testing:

$$H_0 : \sigma_b = 0 \quad \text{vs.} \quad H_1 : \sigma_b > 0.$$

Violation of assumptions of usual asymptotics for Likelihood Ratio tests:

- the tested parameter is on the boundary of the parameter space,
- observations are not i.i.d.

$\Rightarrow$ Test distribution?

RLRT for zero variance components

Exact finite sample null distribution of the RLRT (Crainiceanu and Ruppert, 2004):

$$RLRT_n \stackrel{d}{=} \sup_{\lambda \geq 0} \left\{ (n-p) \log \left[1 + \frac{N_n(\lambda)}{D_n(\lambda)} \right] - \sum_{l=1}^{K} \log(1 + \lambda\, \mu_{l,n}) \right\},$$

where

$$N_n(\lambda) = \sum_{l=1}^{K} \frac{\lambda\, \mu_{l,n}}{1 + \lambda\, \mu_{l,n}}\, \omega_l^2, \qquad D_n(\lambda) = \sum_{l=1}^{K} \frac{\omega_l^2}{1 + \lambda\, \mu_{l,n}} + \sum_{l=K+1}^{n-p} \omega_l^2,$$

$\mu_{l,n}$ eigenvalues of $\boldsymbol{\Sigma}^{\frac{1}{2}}\, \mathbf{Z}'\, (\mathbf{I_n} - \mathbf{X}\, (\mathbf{X}'\, \mathbf{X})^{-1}\, \mathbf{X}')\, \mathbf{Z}\, \boldsymbol{\Sigma}^{\frac{1}{2}}$ and $\omega_l \stackrel{iid}{\sim} \mathrm{N}(0,\, 1)$.

RLRT for zero variance components

Exact finite sample null distribution of the RLRT (Crainiceanu and Ruppert, 2004):

$$RLRT_n \stackrel{d}{=} \sup_{\lambda \geq 0} \left\{ (n-p) \log \left[1 + \frac{N_n(\lambda)}{D_n(\lambda)} \right] - \sum_{l=1}^{K} \log(1 + \lambda\, \mu_{l,n}) \right\},$$

where

$$N_n(\lambda) = \sum_{l=1}^{K} \frac{\lambda\, \mu_{l,n}}{1 + \lambda\, \mu_{l,n}}\, \omega_l^2, \qquad D_n(\lambda) = \sum_{l=1}^{K} \frac{\omega_l^2}{1 + \lambda\, \mu_{l,n}} + \sum_{l=K+1}^{n-p} \omega_l^2,$$

$\mu_{l,n}$ eigenvalues of $\boldsymbol{\Sigma}^{\frac{1}{2}}\, \mathbf{Z}' \,(\mathbf{I_n} - \mathbf{X}\,(\mathbf{X}'\,\mathbf{X})^{-1}\,\mathbf{X}')\, \mathbf{Z}\, \boldsymbol{\Sigma}^{\frac{1}{2}}$ and $\omega_l \stackrel{iid}{\sim} \mathrm{N}(0,\, 1)$.

$\Rightarrow$ This distribution can be simulated in R with the RLRsim package (Scheipl, 2010).

The idea of GLS transformation

Consider the simple linear model:

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \text{where } \boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{V}) \text{ and } \mathbf{V}\text{is known.}$$

Then the generalized least squares estimator is efficient:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{X})^{-1}\,\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{y}$$

The idea of GLS transformation

Consider the simple linear model:

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \text{where } \boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{V}) \text{ and } \mathbf{V}\text{is known.}$$

Then the generalized least squares estimator is efficient:

$$\hat{\beta} = (\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{X})^{-1}\,\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{y} = (\tilde{\mathbf{X}}'\,\tilde{\mathbf{X}})^{-1}\,\tilde{\mathbf{X}}'\,\tilde{\mathbf{y}},$$

where $\tilde{\mathbf{X}} := \mathbf{V}^{-\frac{1}{2}}\,\mathbf{X}$ and $\tilde{\mathbf{y}} := \mathbf{V}^{-\frac{1}{2}}\,\mathbf{y}$.
The idea is to estimate $\mathbf{V}$, then transform the data and estimate $\hat{\beta}$ from the transformed data.

The idea of GLS transformation

Consider the simple linear model:

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \text{where } \boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{V}) \text{ and } \mathbf{V}\text{is known.}$$

Then the generalized least squares estimator is efficient:

$$\hat{\beta} = (\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{X})^{-1}\,\mathbf{X}'\,\mathbf{V}^{-1}\,\mathbf{y} = (\tilde{\mathbf{X}}'\,\tilde{\mathbf{X}})^{-1}\,\tilde{\mathbf{X}}'\,\tilde{\mathbf{y}},$$

where $\tilde{\mathbf{X}} := \mathbf{V}^{-\frac{1}{2}}\,\mathbf{X}$ and $\tilde{\mathbf{y}} := \mathbf{V}^{-\frac{1}{2}}\,\mathbf{y}$.
The idea is to estimate $\mathbf{V}$, then transform the data and estimate $\hat{\beta}$ from the transformed data.

In the case of the general form LMM:

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \mathbf{Z_1}\,\mathbf{b_1} + \ldots + \mathbf{Z_S}\,\mathbf{b_S} + \boldsymbol{\epsilon}, \quad \text{where } \boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \sigma_\epsilon\,\mathbf{R})$$

the transformation leads to the model:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\,\beta + \tilde{\mathbf{Z}}_\mathbf{1}\,\mathbf{b_1} + \ldots + \tilde{\mathbf{Z}}_\mathbf{S}\,\mathbf{b_S} + \tilde{\epsilon}, \quad \text{where } \tilde{\epsilon} \sim \mathrm{N}(\mathbf{0}, \sigma_\epsilon\,\mathbf{I_N})$$

Scenarios

1. **kub_testRI**
   H_1 model: simple cubic LMM with one random intercept and autocorrelated errors,
   test: RLRT for zero variance of the random intercepts
2. **kubglatt**
   H_1 model: smoothing LMM (TP basis 3rd degree) with autocorrelated errors,
   test: RLRT for zero variance of the penalized smoothing coefficients
3. **kubglatt_RI**
   H_1 model: smoothing LMM (TP basis 3rd degree) with a nuisance random intercept and autocorrelated errors,
   test: RLRT for zero variance of the penalized smoothing coefficients

All scenarios simulated for $I = 20, 100, 500$ individuals with $J = 4, 10, 20$ observations and for $\rho = 0, 0.4, 0.8$.

Scenarios

1. **kub_testRI**
   H_1 model: simple cubic LMM with one random intercept and autocorrelated errors,
   test: RLRT for zero variance of the random intercepts
2. **kubglatt**
   H_1 model: smoothing LMM (TP basis 3rd degree) with autocorrelated errors,
   test: RLRT for zero variance of the penalized smoothing coefficients
3. **kubglatt_RI**
   H_1 model: smoothing LMM (TP basis 3rd degree) with a nuisance random intercept and autocorrelated errors,
   test: RLRT for zero variance of the penalized smoothing coefficients

All scenarios simulated for $I = 20, 100, 500$ individuals with $J = 4, 10, 20$ observations and for $\rho = 0, 0.4, 0.8$.
$\Rightarrow$ All type 1 error rates in $[0.045, 0.06]$, no systematic variation observed.

SOEP data:

- 253 044 observations of 33 451 individuals
- observations from 20 waves between 1986 to 2007
- well-being measured by self-reported overall life satisfaction on a discrete scale from 0 to 10

SOEP data:

- 253 044 observations of 33 451 individuals
- observations from 20 waves between 1986 to 2007
- well-being measured by self-reported overall life satisfaction on a discrete scale from 0 to 10

Considered test:

- H_0: LMM with a cubic polynomial in *age* with nuisance random intercept and autocorrelated errors vs.
- H_1: Semiparametric LMM with nuisance random intercept and autocorrelated errors.

SOEP data:

- 253 044 observations of 33 451 individuals
- observations from 20 waves between 1986 to 2007
- well-being measured by self-reported overall life satisfaction on a discrete scale from 0 to 10

Considered test:

- H_0: LMM with a cubic polynomial in *age* with nuisance random intercept and autocorrelated errors vs.
- H_1: Semiparametric LMM with nuisance random intercept and autocorrelated errors.

Result:

- observed RLRT: 557.0264
- 95% quantile of the approximate distribution: 1.612736

$\Rightarrow H_0$ is rejected.

Open Issues

- Investigate different strategies/ methods to estimate the correlation structure and their potentials/ problems.
- Quantify the approximation bias of GLS transformation.
- Exact finite sample or asymptotic distribution of RLRT in case of more than one random effects variance, instead of well-proven approximation?
- Theoretical derivation of the distribution of RLRT without GLS transformation?
- ...

References

Crainiceanu, C. and D. Ruppert (2004). Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 66*(1), 165–185.

Greven, S. (2007). *Non-Standard Problems in Inference for Additive and Linear Mixed Models*. Ph. D. thesis, LMU Munich.

Scheipl, F. (2010). *RLRsim: Exact (Restricted) Likelihood Ratio tests for mixed and additive models.* R package version 2.0-5.

Wunder, C., A. Wiencierz, J. Schwarze, H. Küchenhoff, S. Kleyer, and P. Bleninger (2009). Well-being over the life span: Semiparametric evidence from british and german longitudinal data. Technical Report 179, DIW Berlin.