Negative Variance Components in an Underdispersed, Repeated Time-to-Event setting

(Leuven Statistics Days 2016-2017)

Martial Luyts **Geert Molenberghs** **Geert Verbeke**

Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat)

Katholieke Universiteit Leuven, Belgium

martial.luyts@kuleuven.be

www.ibiostat.be

Interuniversity Institute for Biostatistics
and statistical Bioinformatics

Belgium, October 21, 2016

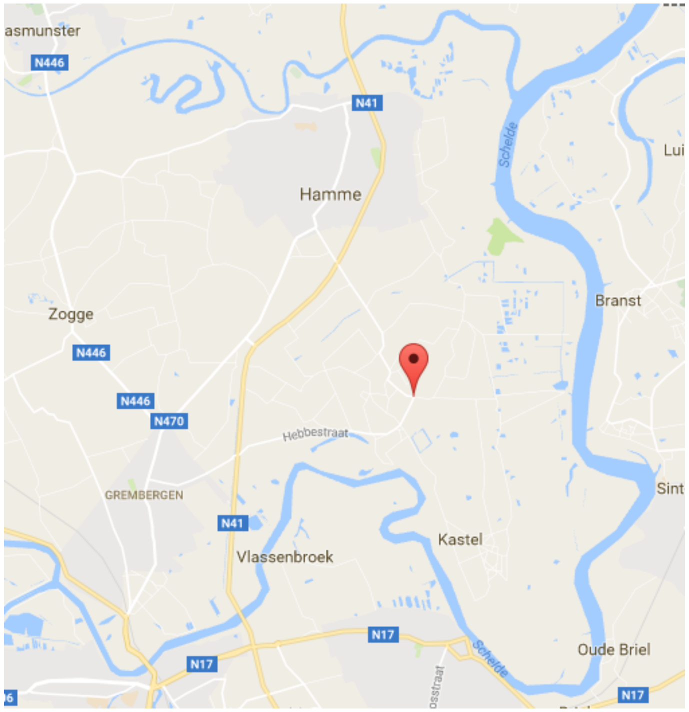
Contents

Part 1:

Introductory material

1.1 Demographic, historical data of Moerzeke

- **Moerzeke** is a small village in the center of Flanders, within the province of East Flanders
  - It is a **geographical isolate**
  - Mainly **populated by farmers** until well into the 20th century
  - More textile industry oriented from the middle of the 19th century onwards
  - **Fertility** was traditionally **high** and dropped at the beginning of the 20th century

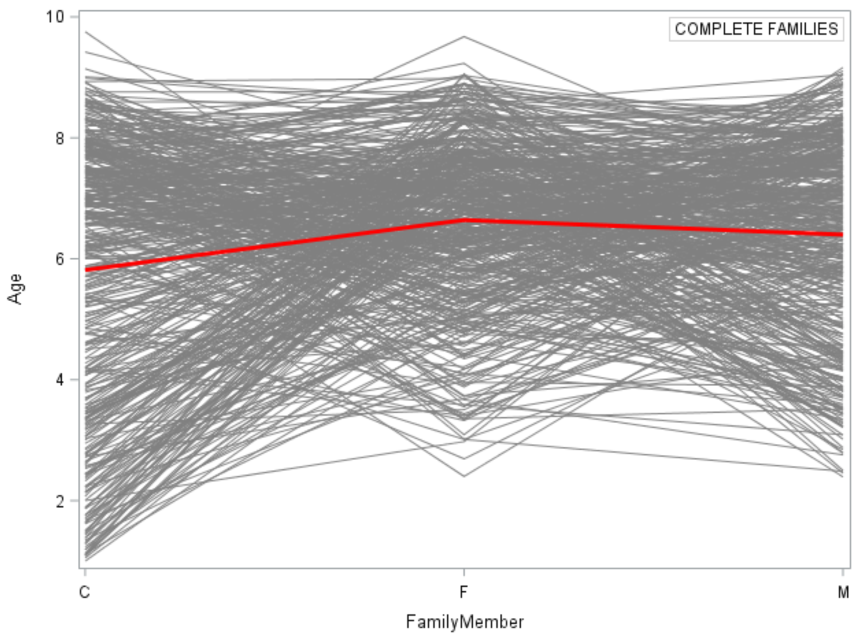


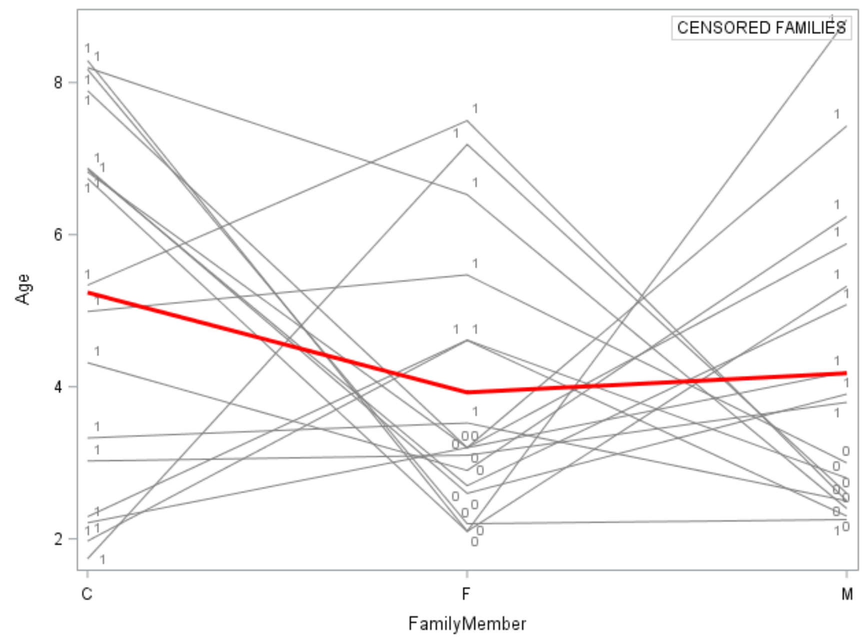
- The information in the database is drawn from church and civil registers

- The database contains information of individuals who were born, married or died in Moerzeke

- Focus is laid on the familial transmission of longevity, i.e., a time-to-event outcome

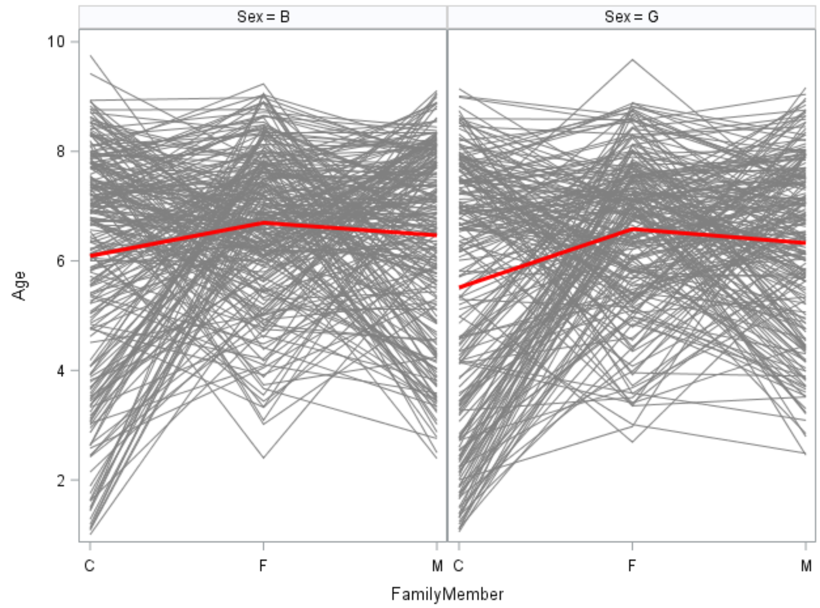
- A sample of 474 families is taken:

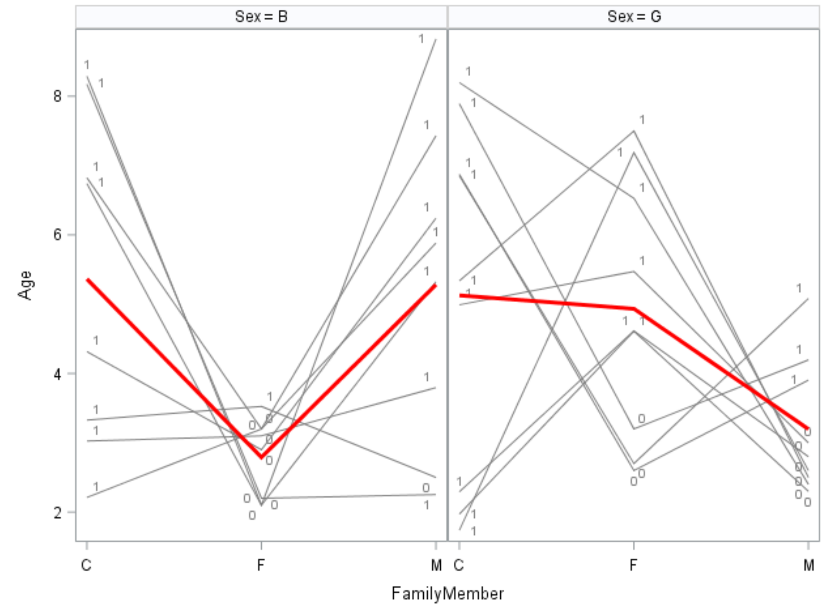




- A total of 457 'complete' families, based on specific criteria
- Recover additional observations ⇒ 17 'censored' families
- Much between- and within-household variability

- And even be categorized in the sex of the first child:





Part 2:

Methodology

2.1 The classical Weibull- and exponential model

- Let T_i be the longevity of mother, father and first-born child, independently of each other $(i = 1, \ldots, 3)$

- Outcome belongs to the family of non-Gaussian outcomes

- The Generalized Linear Model:
  - All T_i have densities $f(t_i|\theta_i, \phi)$ which belong to the exponential family:

$$f(t_i|\theta_i, \phi) = \exp\left\{\phi^{-1}[t_i\theta_i - \psi(\theta_i)] + c(t_i, \phi)\right\}$$

    - *natural parameter* $\longrightarrow$ $\theta_i = \boldsymbol{x_i}'\boldsymbol{\beta}$ $\longleftarrow$ *linear predictor*

    - *Scale parameter (dispersion parameter):* ϕ

    - *Inverse link function:* $\psi'(.)$

- *Mean-variance relationship:* $\mathsf{Var}(T_i) = \phi\psi''\left[\psi'^{-1}(\mathsf{E}(T_i))\right] = \phi v(\mathsf{E}(T_i))$

- Special cases: Exponential- and Weibull model

Element	notation	time to event	
Model		**Exponential**	**Weibull**
Model	$f(t_i)$	$\varphi e^{-\varphi t_i}$	$\varphi\rho t_i^{\rho-1} e^{-\varphi t_i^\rho}$
Nat. param	θ_i	$-\varphi$	
Mean function	$\psi(\theta_i)$	$-\ln(-\theta_i)$	
Norm. constant	$c(t_i, \phi)$	0	
Dispersion	ϕ	1	
Mean	$\mathsf{E}(T_i)$	φ^{-1}	$\varphi^{-1/\rho}\Gamma(\rho^{-1}+1)$
Variance	$\mathsf{Var}(T_i)$	φ^{-2}	$\varphi^{-2/\rho}\left[\Gamma(2\rho^{-1}+1) - \Gamma(\rho^{-1}+1)^2\right]$

2.2 Adjust for extra dispersion

- Mean-variance relationship available

- Different approaches to account for extra dispersion

  - **Approach 1:** $\phi \neq 1 \Longrightarrow \mathsf{Var}(T_i) = \phi \cdot \mathsf{E}(T_i)^2$

  - **Approach 2:** Two-stage approach

$$f(t_i \mid \theta_i) \;=\; \exp\{\phi^{-1} \cdot [t_i \cdot h(\theta_i) - g(\theta_i)] + c(t_i, \phi)\},$$

$$f(\theta_i) \;=\; \exp\{\gamma \cdot [\psi \cdot h(\theta_i) - g(\theta_i)] + c^*(t_i, \psi)\},$$

- Special cases: Exponential-gamma and Weibull-gamma model

Element	notation	time to event	
Model		**Exponential-gamma**	**Weibull-gamma**
Hier. model	$f(t_i \mid \theta_i)$	$\varphi\theta_i e^{-\varphi\theta_i t_i}$	$\varphi\theta_i\rho t_i^{\rho-1}e^{-\varphi\theta_i t_i^\rho}$
RE model	$f(\theta_i)$	$\frac{\theta_i^{\alpha-1}e^{-\theta_i/\beta}}{\beta^\alpha\Gamma(\alpha)}$	$\frac{\theta_i^{\alpha-1}e^{-\theta_i/\beta}}{\beta^\alpha\Gamma(\alpha)}$
Marg. model	$f(t_i)$	$\frac{\varphi\alpha\beta}{(1+\varphi\beta t_i)^{\alpha+1}}$	$\frac{\varphi\rho t_i^{\rho-1}\alpha\beta}{(1+\varphi\beta t_i^\rho)^{\alpha+1}}$
	$h(\theta_i)$	$-\theta_i$	$-\theta_i$
	$g(\theta_i)$	$-\ln(\theta_i)/\varphi$	$-\ln(\theta_i)/\varphi$
	ϕ	$1/\varphi$	$1/\varphi$
	γ	$\varphi(\alpha-1)$	$\varphi(\alpha-1)$
	ψ	$[\beta\varphi(\alpha-1)]^{-1}$	$[\beta\varphi(\alpha-1)]^{-1}$
	$c(t_i,\phi)$	$\ln(\varphi)$	$\ln\left(\varphi\rho t_i^{\rho-1}\right)$
	$c^*(\gamma,\psi)$	$\frac{\gamma+\varphi}{\varphi}\ln(\gamma\psi)-\ln\Gamma\left(\frac{\gamma+\varphi}{\varphi}\right)$	$\frac{\gamma+\varphi}{\varphi}\ln(\gamma\psi)-\ln\Gamma\left(\frac{\gamma+\varphi}{\varphi}\right)$
Mean	$\mathrm{E}(Y)$	$[\varphi(\alpha-1)\beta]^{-1}$	$\frac{\Gamma(\alpha-\rho^{-1})\Gamma(\rho^{-1}+1)}{(\varphi\beta)^{1/\rho}\Gamma(\alpha)}$
Variance	$\mathrm{Var}(Y)$	$\alpha[\varphi^2(\alpha-1)^2(\alpha-2)\beta^2]^{-1}$	$\frac{1}{\rho(\varphi\beta)^{2/\rho}\Gamma(\alpha)}\left[2\Gamma(\alpha-2\rho^{-1})\Gamma(2\rho^{-1}) - \frac{\Gamma(\alpha-\rho^{-1})^2\Gamma(\rho^{-1})^2}{\rho\Gamma(\alpha)}\right]$

2.3 Adjust for hierarchical structures

- Family members are allocated within a household
  $\Longrightarrow$ **Hierarchical structure** is present!

- Notation of T_i now extends to T_{ij}, which presents the longevity of mother, father and first child ($j = 1, 2, 3$) in household i ($i = 1, \ldots, 474$)

- The Generalized Linear Mixed Model:

$$\begin{aligned} f_i(t_{ij}|\boldsymbol{b_i}, \boldsymbol{\beta}, \phi) &= \exp\left\{\phi^{-1}[t_{ij}\theta_{ij} - \psi(\theta_{ij})] + c(t_{ij}, \phi)\right\} \\ \eta(\mu_{ij}) &= \eta[E(T_{ij}|\boldsymbol{b_i})] = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b_i} \\ \boldsymbol{b_i} &\sim N(\mathbf{0}, D) \end{aligned}$$

- Special case: The Weibull-normal model

$$f(\mathbf{t}_i \mid \mathbf{b}_i) = \prod_{j=1}^{3} \lambda \cdot \rho \cdot t_{ij}^{\rho-1} \cdot e^{\mathbf{x}_{ij}^{'} \cdot \beta + \mathbf{z}_{ij}^{'} \cdot \mathbf{b}_i} \cdot e^{-\lambda \cdot t_{ij}^{\rho} \cdot e^{\mathbf{x}_{ij}^{'} \cdot \beta + \mathbf{z}_{ij}^{'} \cdot \mathbf{b}_i}},$$

$$f(\mathbf{b}_i) = \frac{1}{(2 \cdot \pi)^{q/2} \cdot \mid D \mid^{1/2}} \cdot e^{-\frac{1}{2} \cdot \mathbf{b}_i^{'} \cdot D^{-1} \cdot \mathbf{b}_i}.$$

2.4 The combined model

- Until now, both complexities were treated separately
- Is there a way to combine both strategies simultaneously? **YES!!!**
- The Combined Model:

$$
\begin{aligned}
f_i(t_{ij}|\boldsymbol{b_i}, \boldsymbol{\beta}) &= \exp\{\phi^{-1}[t_{ij}\lambda_{ij} - \psi(\lambda_{ij})] + c(t_{ij}, \phi)\} \\
E(T_{ij}|\theta_{ij}, \boldsymbol{b_i}) &= \mu_{ij}^c = \theta_{ij}\kappa_{ij} \\
\kappa_{ij} &= g(\boldsymbol{x}_{ij}'\boldsymbol{\xi} + \boldsymbol{z}_{ij}'\boldsymbol{b_i}) \\
\theta_{ij} &\sim \mathcal{G}_{ij}(\beta_{ij}, \sigma_{ij}^2) \\
\boldsymbol{b_i} &\sim N(\mathbf{0}, D) \\
\eta_{ij} &= \boldsymbol{x}_{ij}'\boldsymbol{\xi} + \boldsymbol{z}_{ij}'\boldsymbol{b_i}
\end{aligned}
$$

- Special case: Weibull-gamma-normal model

$$f(\mathbf{t}_i \mid \theta_i, \mathbf{b}_i) = \prod_{j=1}^{3} \lambda \cdot \rho \cdot \theta_{ij} \cdot t_{ij}^{\rho-1} \cdot e^{\mathbf{x}_{ij}^{'} \cdot \beta + \mathbf{z}_{ij}^{'} \cdot \mathbf{b}_i} \cdot e^{-\lambda \cdot t_{ij}^{\rho} \cdot \theta_{ij} \cdot e^{\mathbf{x}_{ij}^{'} \cdot \beta + \mathbf{z}_{ij}^{'} \cdot \mathbf{b}_i}},$$

$$f(\theta_i) = \prod_{j=1}^{3} \frac{1}{\beta_j^{\alpha_j} \cdot \Gamma(\alpha_j)} \cdot \theta_{ij}^{\alpha_j - 1} \cdot e^{-\theta_{ij}/\beta_j},$$

$$f(\mathbf{b}_i) = \frac{1}{(2 \cdot \pi)^{q/2} \cdot \mid D \mid^{1/2}} \cdot e^{-\frac{1}{2} \cdot \mathbf{b}_i^{'} \cdot D^{-1} \cdot \mathbf{b}_i}.$$

- Such complex models can have some **drawbacks**:
  - Attendance of analytically closed-form expressions? If not, **approximation methods** need to be used (e.g., Taylor-series expansion based methods; Laplace approximations; numeric integration)
  - Weibull-gamma-normal model $\Rightarrow$ Analytical closed-form expressions exist!!

- **Marginal density:**

$$f(\boldsymbol{t}_i) = \sum_{(m_1,\ldots,m_3)} \prod_{j=1}^{3} \frac{(-1)^{m_j}}{m_j!} \frac{\Gamma(\alpha_j + m_j + 1)\beta_j^{m_j+1}}{\Gamma(\alpha_j)} \lambda^{m_j+1} \rho t_{ij}^{(m_j+1)\rho-1}$$
$$\times \exp\left\{(m_j+1)\left[\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}(m_j+1) \cdot \boldsymbol{z}'_{ij} D \boldsymbol{z}_{ij}\right]\right\}$$

- **Marginal moments:**

$$E(T_{ij}^k) = \frac{\alpha_j B(\alpha_j - k/\rho, k/\rho + 1)}{\lambda^{k/\rho}\beta_j^{k/\rho}} \exp\left(-\frac{k}{\rho}\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{k^2}{2\rho^2}\boldsymbol{z}'_{ij} D \boldsymbol{z}_{ij}\right)$$

$$E(T_{ij}) = \frac{\alpha_j B(\alpha_j - 1/\rho, 1/\rho + 1)}{\lambda^{1/\rho}\beta_j^{1/\rho}} \exp\left(-\frac{1}{\rho}\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2\rho^2}\boldsymbol{z}'_{ij} D \boldsymbol{z}_{ij}\right)$$

$$
\begin{aligned}
\mathsf{Var}(T_{ij}) &= \frac{\alpha_j}{\lambda^{2/\rho}\beta_j^{2\rho}} \exp\left(-\frac{2}{\rho}\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{\rho^2}\boldsymbol{z}'_{ij}D\boldsymbol{z}_{ij}\right) \\
&\times \left[B\left(\alpha_j - 2/\rho, 2/\rho + 1\right)\exp\left(\frac{1}{\rho^2}\boldsymbol{z}'_{ij}D\boldsymbol{z}_{ij}\right) - \alpha_j B\left(\alpha_j - \frac{1}{\rho}, \frac{1}{\rho} + 1\right)^2\right] \\
\mathsf{Cov}(T_{ij}, T_{ik}) &= \frac{\alpha_j\alpha_k}{\lambda^{2/\rho}\beta_j^{1/\rho}\beta_k^{1/\rho}} \exp\left[-\frac{1}{\rho}(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{x}'_{ik}\boldsymbol{\beta})\right] \\
&\times B\left(\alpha_j - \frac{1}{\rho}, \frac{1}{\rho} + 1\right) B\left(\alpha_k - \frac{1}{\rho}, \frac{1}{\rho} + 1\right) \\
&\times \exp\left[\frac{1}{2\rho^2}(\boldsymbol{z}'_{ij}D\boldsymbol{z}_{ij} + \boldsymbol{z}'_{ik}D\boldsymbol{z}_{ik})\right]\left[\exp\left(\frac{1}{\rho^2}\boldsymbol{z}'_{ij}D\boldsymbol{z}_{ik}\right) - 1\right]
\end{aligned}
$$

Part 3:

Analyzing the Moerzeke data

3.1 Findings with the multivariate Plackett Dale model

Main conclusions:

- The estimated **association parameter** between **mother** and **child** is **1.349** (95% CI $=$ [1.002;1.696]), indicating a positive association between them;
- However, for **father-child**, the value seems to be lower (**0.983**; not statistically significant).

3.2 Extra findings with the combined modeling framework

- The proposed exponential-gamma-normal model is formulate as

$$
\begin{aligned}
f_{T_i}(\mathbf{t}_i \mid \boldsymbol{\theta}_i, \mathbf{b}_i) &= \prod_{j=1}^{3} \theta_{ij} \cdot e^{\Delta_{ij}} \cdot e^{-t_{ij} \cdot \theta_{ij} \cdot e^{\Delta_{ij}}}, \\
\Delta_{ij} &= \xi_0 + \xi_{\text{G}} \cdot S_i + \xi_{\text{YB}} \cdot Y_{ij} + \xi_{\text{IN2}} \cdot F_{ij} + \xi_{\text{IN1}} \cdot M_{ij} + b_i, \\
f(\boldsymbol{\theta}_i) &= \prod_{j=1}^{3} \frac{1}{\alpha^{-\alpha} \cdot \Gamma(\alpha)} \cdot \theta_{ij}^{\alpha-1} \cdot e^{-\alpha \cdot \theta_{ij}}, \\
f(b_i) &= \frac{1}{(2 \cdot \pi \cdot d)^{1/2}} \cdot e^{-d/2},
\end{aligned}
$$

where

- $S_i = 0$ if the sex of the first child in household i is female and 1 if it is male;
- $F_{ij} = 1$ if person j in household i is the father and 0 if it is not the father;
- $M_{ij} = 1$ if person j in household i is the mother and 0 if it is not the mother;

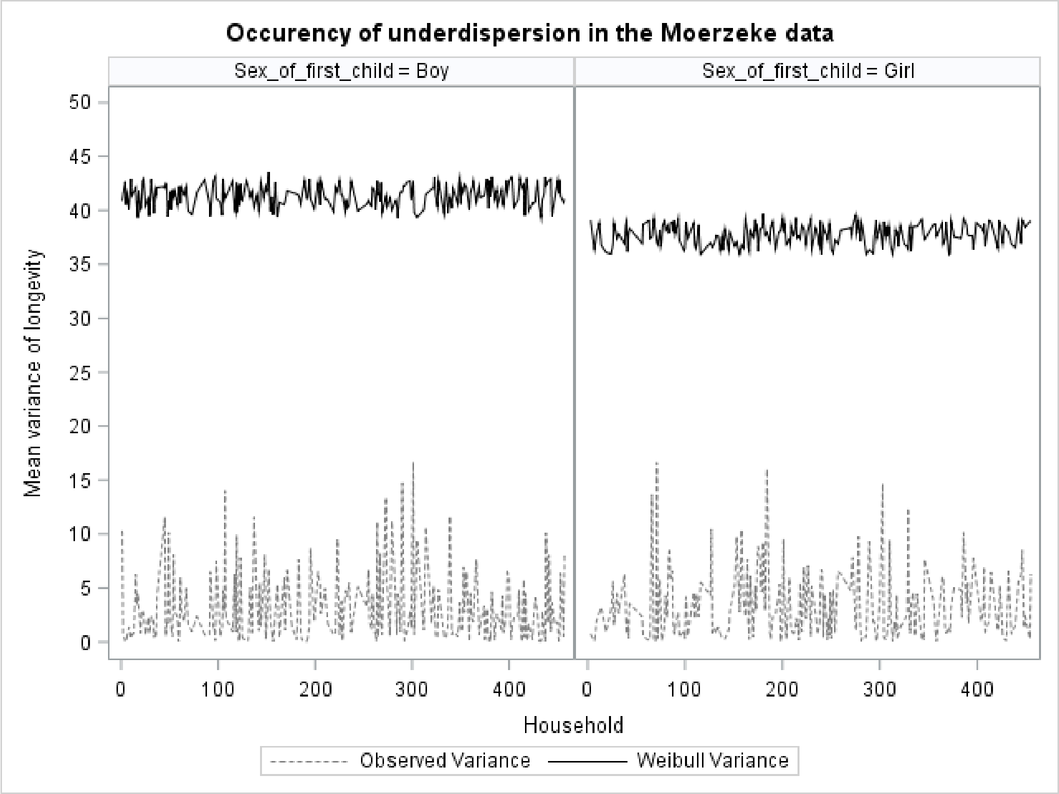
- The year of birth Y_{ij}, which is subject-specific

- Results of fitting several exponential models:

		E—	EG–	E–N	EGN
Effect	Par.	Estimate (s.e.)	Estimate (s.e.)	Estimate (s.e.)	Estimate (s.e.)
Intercept	ξ_0	-0.7357 (2.0636)	-0.7354 (2.0657)	-0.7357 (2.0636)	-0.7355 (2.0659)
Sex of first child	ξ_{G}	-0.0454 (0.0541)	-0.0454 (0.0542)	-0.0454 (0.0541)	-0.0454 (0.0542)
Year of Birth	ξ_{YB}	-0.5471 (1.1290)	-0.5471 (1.1302)	-0.5471 (0.9600)	-0.5471 (1.1303)
Indicator of Father	ξ_{IN2}	-0.1524 (0.0463)	-0.1526 (0.0765)	-0.1524 (0.0764)	-0.1526 (0.0765)
Indicator of Mother	ξ_{IN1}	-0.1134 (0.0744)	-0.1135 (0.0745)	-0.1134 (0.0744)	-0.1135 (0.0745)
Std. dev. random effect	$\sqrt{d}$	$--$	$--$	$-6.06E-8$ (0.0284)	$5.471E-7$ (0.0284)
Gamma parameter	α	$--$	545.01 (359.54)	$--$	500.01 (315.96)
-2 log-likelihood		7777.0	7779.3	7777.0	7779.6

$\Rightarrow$ **Possible indication of negative variance components!**

- **Digging a little deeper:**



- **Indication that underdispersion is present**

- **Note: Normal random effects induce both correlation and over-underdispersion**

Part 4:

Negative variance components

4.1 Linear mixed model

- Start simple with the random intercept approach
  - **Conditional model**:

$$\begin{aligned} \mathbf{Y}_i|b_i &\sim N(X_i \cdot \beta_i + b_i, \sigma^2), \\ b_i &\sim N(0, d^2). \end{aligned}$$

    - **Constraint:** σ^2 and d^2 require to be POSITIVE!

  - **Marginal model**:

$$\mathbf{Y}_i \sim N(X_i \cdot \beta, d^2 \cdot J + \sigma^2 \cdot I).$$

    - **Constraint:** $d^2 \cdot J + \sigma^2 \cdot I$ require to be POSITIVE DEFINITE! (satisfied if $\rho = d^2/(d^2 + \sigma^2) \geq -(n_i - 1)^{-1}$)

$\Rightarrow$ **NEGATIVE VALUES for d^2 are perfectly acceptable!**

- In the **marginal model**, d^2 is interpreted as a **VARIANCE COMPONENT** (NOT as a variance!);
- Negative values for d^2 are perfectly possible in practice, e.g., in a **COMPETITIVE SETTING**:

- The **asymptotic null distribution** is well known to be χ^2_1 for the marginal model, while this is $\frac{1}{2}\chi^2_1 + \frac{1}{2}\chi^2_0$ for the conditional model

4.2 Generalized linear mixed & combined models

- Investigating **hierarchical and marginal views** for **negative variance components and/or underdispersion** become **more complex** in these frameworks

- Classical software procedures like **NLMIXED** encompasses numerical optimization algorithms, which often follows a **hierarchical viewpoint**

  - **Limitation:** No allowance for negative estimates of the variance components


danke
謝謝
ngiyabonga
teşekkür ederim
спасибо
dank je
thank you
gracias
tapadh leat
bedankt
hvala
mauruuru
mochchakkeram
dziękuję
go raibh maith agat
sagolun
sukriya
kop khun krap
obrigado
grazie
arigatô
takk
dakujem
мерси
terima kasih
감사합니다
ευχαριστώ
merci