Lab 2.3 - Squid

Alessandra Ragni

October 11, 2022

$$\text{Testisweight}_{im} = \beta_0 + \beta_1 \cdot \text{DML}_i + \beta_{2m} + \beta_{3m} \cdot \text{DML}_i + \epsilon_{im} \;, \quad \epsilon_{im} \sim \mathcal{N}(0, \sigma^2_{im})$$

for squid i $(i = 1, ..., 768)$ at month m $(m = 1, ..., 12)$
with $\beta_{2m} = \beta_{3m} = 0$ if $m = 1$

```
library(nlme)
lm.form <- Testisweight ~ DML * fMONTH
```

Model M.lm

$\epsilon_{im} \sim \mathcal{N}(0, \sigma^2)$ which means

$$\sigma^2_{i,1} = \sigma^2_{i,2} = ... = \sigma^2_{i,12} \quad \forall i \in \{1, ..., 768\}$$

```
M.lm <- gls(lm.form, data = Squid)
```

Model M.gls1 - λ_is known

$\epsilon_{im} \sim \mathcal{N}(0, \sigma_i^2)$ where $\sigma_i^2 = \sigma^2 \cdot \text{DML}_i$, allowing for a larger residual spread if DML increases, i.e.,

$$\sigma_i^2 = \begin{cases} \sigma_1^2 \\ ... \\ \sigma_{768}^2 \end{cases} = \begin{cases} \sigma^2 \cdot \text{DML}_1 & \text{if } i = 1; \\ ... \\ \sigma^2 \cdot \text{DML}_{768} & \text{if } i = 768. \end{cases}$$

```
vf1Fixed <- varFixed(~DML)
# the variance covariate needs to be continuous!
M.gls1 <- gls(lm.form, weights = vf1Fixed, data = Squid)
```

M.lm is not nested within *varFixed*.

Model M.gls2 - $< \delta >$-group

Month-specific variance: $\epsilon_{im} \sim \mathcal{N}(0, \sigma_m^2)$

$$\sigma_m^2 = \begin{cases} \sigma_1^2 \\ \sigma_2^2 \\ ... \\ \sigma_{12}^2 \end{cases} = \begin{cases} \sigma^2 \cdot \delta_1^2 & \text{if } m = 1; \\ \sigma^2 \cdot 1^2 & \text{if } m = 2; \\ ... \\ \sigma^2 \cdot \delta_{12}^2 & \text{if } m = 12. \end{cases}$$

we get: $\delta_1 = \frac{\sigma_1}{\sigma_2}$; $\delta_{12} = \frac{\sigma_{12}}{\sigma_2}$

```
vf2 <- varIdent(form= ~ 1|fMONTH)
M.gls2 <- gls(lm.form, weights = vf2, data = Squid)
```

M.lm is nested within *varIdent*: we can make use of likelihood ratio test.

Model M.gls3 - $< \delta >$-group

$\epsilon_{im} \sim \mathcal{N}(0, \sigma_{im}^2)$ with $\underline{\delta} = \delta$ (scalar) since we do not include any stratification in the model and

$$\begin{aligned} \sigma_{im}^2 &= \sigma^2 \cdot \lambda_{im}^2 \\ &= \sigma^2 \cdot \lambda^2(\ \delta, \text{DML}_{im}\ ) \quad \text{with } \lambda \text{ modeled as varPower}(\cdot) \\ &= \sigma^2 \cdot |\text{DML}_{im}|^{2\delta} \end{aligned}$$

```
vf3 <- varPower(form =~ DML)
M.gls3 <- gls(lm.form, weights = vf3, data = Squid)
```

M.lm is nested within *varPower*: if $\delta = 0$ we get the "classical model". *varIdent* is nested within the *varPower* structure.

Model M.gls4 - $< \delta >$-group

$\epsilon_{im} \sim \mathcal{N}(0, \sigma_{im}^2)$ with $\underline{\delta} = [\delta_1, ..., \delta_{12}]'$ since we include stratification by month and

$$\begin{aligned} \sigma_{im}^2 &= \sigma^2 \cdot \lambda_{im}^2 \\ &= \sigma^2 \cdot \lambda^2(\ \underline{\delta}, \text{DML}_{im}\ ) \quad \text{with } \lambda \text{ modeled as varPower}(\cdot) \\ &= \sigma^2 \cdot |\text{DML}_{im}|^{2\delta_m} \\ &= \begin{cases} \sigma^2 \cdot |\text{DML}_{i,1}|^{2\delta_1} & \text{if } m = 1 \\ ... \\ \sigma^2 \cdot |\text{DML}_{i,12}|^{2\delta_{12}} & \text{if } m = 12 \end{cases} \end{aligned}$$

```
# we are modeling an increase in spread for larger DML values,
# but only in certain months!
vf4 <- varPower(form=~ DML | fMONTH)
M.gls4 <- gls(lm.form, weights = vf4, data = Squid)
```

If variance covariate has values equal to 0, the variance of the residuals is 0 as well; this causes problems in the numerical estimation process and *varPower()* should not be used. Notice that if $\delta = 0.5$ and variance covariate assumes positive values, we get the same variance structure as *varFixed*($\sim$DML). If $\delta = 0$ we get the "classical model" **M.lm**.

Model M.gls5 - $< \delta >$-group

$\epsilon_{im} \sim \mathcal{N}(0, \sigma^2_{im})$ with $\underline{\delta} = \delta$ (scalar) since we do not include any stratification in the model and

$$\begin{aligned} \sigma^2_{im} &= \sigma^2 \cdot \lambda^2_{im} \\ &= \sigma^2 \cdot \lambda^2(\ \delta, \mathrm{DML}_{im}\ ) \quad \text{with } \lambda \text{ modeled as varExp}(\cdot) \\ &= \sigma^2 \cdot e^{2\delta \cdot |\mathrm{DML}_{im}|} \end{aligned}$$

```
vf5 <- varExp(form =~ DML)
M.gls5 <- gls(lm.form, weights = vf5, data = Squid)
# as before, we could allow for different delta per month
# by specifying varExp(form = ~ DML | fMONTH) )
```

In this case, there are no restrictions on δs or *DML*; moreover, this structure allows also a decrease of spread for *DML* values if *delta* is negative. If $\delta = 0$ we get the "classical model" **M.lm**.

Model M.gls6 - $< \delta >$-group

$\epsilon_{im} \sim \mathcal{N}(0, \sigma^2_{im})$ with $\underline{\delta_1} = \delta_1$ (scalar) and $\underline{\delta_2} = \delta_2$ (scalar), since we do not include any stratification in the model and

$$\begin{aligned} \sigma^2_{im} &= \sigma^2 \cdot \lambda^2_{im} \\ &= \sigma^2 \cdot \lambda^2(\ \delta, \mathrm{DML}_{im}\ ) \quad \text{with } \lambda \text{ modeled as varConstPower}(\cdot) \\ &= \sigma^2 \cdot (\delta_1 + |\mathrm{DML}_{im}|^{\delta_2})^2 \end{aligned}$$

```
vf6 <- varConstPower(form =~ DML)
M.gls6 <- gls(lm.form, weights = vf6, data = Squid)
```

This variance structure should work better than the *varExp()* if the variance covariate has values close to zero.

Model M.gls7 - $< \delta >$-group

$\epsilon_{im} \sim \mathcal{N}(0, \sigma^2_{im})$ with $\underline{\delta_1} = [\delta_{1,1}, ..., \delta_{1,12}]'$ and $\underline{\delta_2} = [\delta_{2,1}, ..., \delta_{2,12}]'$, since we include stratification by month in the model and

$$\begin{aligned}
\sigma^2_{im} &= \sigma^2 \cdot \lambda^2_{im} \\
&= \sigma^2 \cdot \lambda^2(\ \delta,\, \text{DML}_{im}\ ) \quad \text{with } \lambda \text{ modeled as varConstPower}(\cdot) \\
&= \sigma^2 \cdot (\delta_{1m} + |\text{DML}_{im}|^{\delta_{2m}})^2
\end{aligned}$$

```r
vf7 <- varConstPower(form =~ DML | fMONTH)
M.gls7 <- gls(lm.form, weights = vf7, data = Squid)
```

Model M.gls8 - $< \delta >$-group

We can allow for both an increase in residual spread for larger DML values as well as a different spread per month:
$\epsilon_{im} \sim \mathcal{N}(0, \sigma^2_{im})$ with $\underline{\delta} = \delta$ (scalar), since we do not include any stratification in the model and

$$\begin{aligned}
\sigma^2_{im} &= \sigma^2_m \cdot \lambda^2_{im} \\
&= \sigma^2_m \cdot \lambda^2(\ \delta,\, \text{DML}_{im}\ ) \quad \text{with } \lambda \text{ modeled as varComb}(\cdot) \\
&= \sigma^2_m \cdot e^{2\delta \cdot |\text{DML}_{im}|}
\end{aligned}$$

```r
vf8 <- varComb(varIdent(form =~ 1 | fMONTH), varExp(form =~ DML) )
M.gls8 <- gls(lm.form, weights = vf8, data = Squid)
```