**Lab session I Tuesday – Chp. 5 – Random Slopes**

-----------------------------------------------------------------------------------------------------

name: <unnamed>

log: J:\Multilevel TA\Chp4\_Random\_Slopes.log

log type: text

opened on: 16 Apr 2012, 22:19:37

. \*\*chap5.do\*\*

. \*\*Examples 5.1, 5.2, 5.4\*\*

. \*\*The Random slopes model\*\*

.

.

. clear all

. set more off

. cd "J:\Multilevel TA"

J:\Multilevel TA

.

. \*Read in the data

.

. \*infile schoolnr pupilNR\_new langpost ses IQ\_verb sex Minority denomina sch\_ses sch\_iqv sch\_min usi

> ng "/Users/Jon/Documents/Snijders/SecondEditionExamples/DataSnijdersSecondEdition/mlbook2\_r.txt" in

> 2/3759, clear

. infile schoolnr pupilNR\_new langpost ses IQ\_verb sex Minority denomina sch\_ses sch\_iqv sch\_min usin

> g mlbook2\_r.txt

'schoolnr' cannot be read as a number for schoolnr[1]

'pupilNR\_new' cannot be read as a number for pupilNR\_new[1]

'langPOST' cannot be read as a number for langpost[1]

'ses' cannot be read as a number for ses[1]

'IQ\_verb' cannot be read as a number for IQ\_verb[1]

'sex' cannot be read as a number for sex[1]

'Minority' cannot be read as a number for Minority[1]

'denomina' cannot be read as a number for denomina[1]

'sch\_ses' cannot be read as a number for sch\_ses[1]

'sch\_iqv' cannot be read as a number for sch\_iqv[1]

'sch\_min' cannot be read as a number for sch\_min[1]

(3759 observations read)

.

.

. egen gmeanIQverb = mean(IQ\_verb), by(schoolnr)

(1 missing value generated)

.

. \*Example 5.1

. \*A random slope for IQ

. \*Observe that the covariance(unstr) option is necessary to get the covariance between the random ef

> fects

.

. xtmixed langpost IQ\_verb gmeanIQverb || schoolnr: IQ\_verb, mle covariance(unstr) var

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = -12433.317

Iteration 1: log likelihood = -12432.785

Iteration 2: log likelihood = -12432.784

Computing standard errors:

Mixed-effects ML regression Number of obs = 3758

Group variable: schoolnr Number of groups = 211

Obs per group: min = 4

avg = 17.8

max = 34

Wald chi2(2) = 1614.71

Log likelihood = -12432.784 Prob > chi2 = 0.0000

------------------------------------------------------------------------------

langpost | Coef. Std. Err. z P>|z| [95% Conf. Interval]

-------------+----------------------------------------------------------------

IQ\_verb | 2.480199 .0643305 38.55 0.000 2.354113 2.606284

gmeanIQverb | 1.002813 .2627389 3.82 0.000 .4878543 1.517772

\_cons | 41.09789 .2341602 175.51 0.000 40.63895 41.55684

------------------------------------------------------------------------------

------------------------------------------------------------------------------

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

-----------------------------+------------------------------------------------

schoolnr: Unstructured |

var(IQ\_verb) | .1953124 .0771219 .0900784 .4234858

var(\_cons) | 8.916067 1.129264 6.956078 11.42831

cov(IQ\_verb,\_cons) | -.8444576 .2343484 -1.303772 -.3851432

-----------------------------+------------------------------------------------

var(Residual) | 39.68833 .9644273 37.84239 41.6243

------------------------------------------------------------------------------

LR test vs. linear regression: chi2(3) = 421.55 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. est store mod1

.

. \*Figure 5.2

. predict BLUP\*, reff

(1 missing value generated)

(1 missing value generated)

.

. gen predRandomSlope= \_b[\_cons] + BLUP2+(\_b[IQ\_verb]+BLUP1)\*IQ\_verb ///

> +\_b[gmeanIQverb]\*gmeanIQverb

(1 missing value generated)

.

. sort schoolnr IQ\_verb

. twoway (line predRandomSlope IQ\_verb, connect(ascending)), ///

> title("Random intercepts and random slopes") ytitle("Test score") ///

> xtitle(IQ) scheme(s1color)

.

.

. \*Might Avg. IQ be an explanation for the different IQ-effects?

. \*Plotting country level residuals (random slopes or random intercepts) against potential explanator

> y variables

. \*is usually a good idea.

.

. twoway (scatter BLUP1 gmeanIQverb) (lfit BLUP1 gmeanIQverb), xtitle("Avg. IQ") ytitle("IQ country l

> evel residuals") scheme(s1mono)

.

. \* Example 5.2

. \*Cross level interaction between IQ and group-mean IQ

.

. \*gen IQxMeanIQ= IQ\_verb\*gmeanIQverb

.

.

. xtmixed langpost c.IQ\_verb##c.gmeanIQverb || schoolnr: IQ\_verb, mle covariance(un)

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = -12429.786

Iteration 1: log likelihood = -12428.963

Iteration 2: log likelihood = -12428.962

Computing standard errors:

Mixed-effects ML regression Number of obs = 3758

Group variable: schoolnr Number of groups = 211

Obs per group: min = 4

avg = 17.8

max = 34

Wald chi2(3) = 1693.93

Log likelihood = -12428.962 Prob > chi2 = 0.0000

-----------------------------------------------------------------------------------------

langpost | Coef. Std. Err. z P>|z| [95% Conf. Interval]

------------------------+----------------------------------------------------------------

IQ\_verb | 2.468985 .0629405 39.23 0.000 2.345624 2.592346

gmeanIQverb | 1.094822 .2619394 4.18 0.000 .5814298 1.608214

|

c.IQ\_verb#c.gmeanIQverb | -.1811133 .0641514 -2.82 0.005 -.3068479 -.0553788

|

\_cons | 41.21933 .2355939 174.96 0.000 40.75757 41.68108

-----------------------------------------------------------------------------------------

------------------------------------------------------------------------------

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

-----------------------------+------------------------------------------------

schoolnr: Unstructured |

sd(IQ\_verb) | .4047489 .090912 .2606121 .6286036

sd(\_cons) | 2.942895 .1871715 2.59799 3.33359

corr(IQ\_verb,\_cons) | -.7086838 .1694809 -.914093 -.2138496

-----------------------------+------------------------------------------------

sd(Residual) | 6.305471 .0766981 6.156924 6.457603

------------------------------------------------------------------------------

LR test vs. linear regression: chi2(3) = 404.57 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. est store mod2

.

. esttab mod1 mod2, b(%9.2f) se(%9.2f)

--------------------------------------------

(1) (2)

langpost langpost

--------------------------------------------

langpost

IQ\_verb 2.48\*\*\* 2.47\*\*\*

(0.06) (0.06)

gmeanIQverb 1.00\*\*\* 1.09\*\*\*

(0.26) (0.26)

c.IQ\_verb#~b -0.18\*\*

(0.06)

\_cons 41.10\*\*\* 41.22\*\*\*

(0.23) (0.24)

--------------------------------------------

lns1\_1\_1

\_cons -0.82\*\*\* -0.90\*\*\*

(0.20) (0.22)

--------------------------------------------

lns1\_1\_2

\_cons 1.09\*\*\* 1.08\*\*\*

(0.06) (0.06)

--------------------------------------------

atr1\_1\_1\_2

\_cons -0.76\*\* -0.88\*\*

(0.26) (0.34)

--------------------------------------------

lnsig\_e

\_cons 1.84\*\*\* 1.84\*\*\*

(0.01) (0.01)

--------------------------------------------

N 3758 3758

--------------------------------------------

Standard errors in parentheses

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

.

. /\* Unfortunately, Stata is a bit unpractical sometimes.

> This is how to convert the logged variance components properly:

>

> mat list e(b)

> estadd scalar v1 = exp(2\*[lns1\_1\_1]\_b[\_cons])

> estadd scalar v2 = exp(2\*[lns1\_1\_2]\_b[\_cons])

> estadd scalar cov = tanh([atr1\_1\_1\_2]\_b[\_cons])\* exp([lns1\_1\_1]\_b[\_cons])\* exp([lns1\_1\_2]\_b[\_cons])

> estadd scalar v\_e = exp(2\*[lnsig\_e]\_b[\_cons])

>

> esttab mod1 mod2, b(%9.2f) se(%9.2f) stats(v1 v2 cov v\_e)

> \*/

.

. \*Plotting the cross-level interaction

. sum gmeanIQverb

Variable | Obs Mean Std. Dev. Min Max

-------------+--------------------------------------------------------

gmeanIQverb | 3758 .0441831 .8205808 -5.12 2.47375

.

. gen predlowIQ=\_b[\_cons]+\_b[IQ\_verb]\*IQ\_verb+\_b[gmeanIQverb]\*r(min) ///

> +\_b[IQ\_verb#c.gmeanIQverb]\*IQ\_verb\*r(min)

(1 missing value generated)

.

. gen predhiIQ=\_b[\_cons]+\_b[IQ\_verb]\*IQ\_verb+\_b[gmeanIQverb]\*r(max) ///

> +\_b[IQ\_verb#c.gmeanIQverb]\*IQ\_verb\*r(max)

(1 missing value generated)

.

. sort IQ\_verb

. twoway (line predlowIQ IQ\_verb, legend(label(1 "Low IQ school"))) ///

> (line predhiIQ IQ\_verb, lpattern(dash) ///

> legend(label(2 "High IQ school"))), scheme(s1mono) ///

> xtitle("Individual IQ") ytitle("Predicted language score")

.

.

. \*Example 5.3

. \*A model with many fixed effects

. \*Differs from the specification in the book, because xtmixed does not allow the covariance restrict

> ions.

. \*See Addendum!

.

. bysort schoolnr: egen gmeanSES=mean(ses)

(1 missing value generated)

.

. xtmixed langpost IQ\_verb ses c.IQ\_verb#c.ses gmeanIQverb gmeanSES c.gmeanIQverb#c.gmeanSES ///

> c.IQ\_verb#c.gmeanIQverb c.IQ\_verb#c.gmeanSES c.ses#c.gmeanIQverb c.ses#c.gmean

> SES || schoolnr: IQ\_verb, cov(unstr) var

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = -12314.27

Iteration 1: log likelihood = -12313.253

Iteration 2: log likelihood = -12313.249

Iteration 3: log likelihood = -12313.249

Computing standard errors:

Mixed-effects ML regression Number of obs = 3758

Group variable: schoolnr Number of groups = 211

Obs per group: min = 4

avg = 17.8

max = 34

Wald chi2(10) = 2010.73

Log likelihood = -12313.249 Prob > chi2 = 0.0000

------------------------------------------------------------------------------------------

langpost | Coef. Std. Err. z P>|z| [95% Conf. Interval]

-------------------------+----------------------------------------------------------------

IQ\_verb | 2.233239 .0634882 35.18 0.000 2.108804 2.357674

ses | .1717136 .0118958 14.43 0.000 .1483983 .195029

|

c.IQ\_verb#c.ses | -.019102 .0062636 -3.05 0.002 -.0313785 -.0068255

|

gmeanIQverb | .7625186 .3122987 2.44 0.015 .1504245 1.374613

gmeanSES | -.0854514 .0446985 -1.91 0.056 -.1730589 .0021561

|

c.gmeanIQverb#c.gmeanSES | -.1286327 .0381265 -3.37 0.001 -.2033593 -.0539061

|

c.IQ\_verb#c.gmeanIQverb | -.0772729 .0816733 -0.95 0.344 -.2373497 .0828039

|

c.IQ\_verb#c.gmeanSES | .004814 .0134621 0.36 0.721 -.0215713 .0311993

|

c.ses#c.gmeanIQverb | .0206205 .0181859 1.13 0.257 -.0150233 .0562643

|

c.ses#c.gmeanSES | .0003298 .0022241 0.15 0.882 -.0040294 .0046891

|

\_cons | 41.6081 .2583287 161.07 0.000 41.10178 42.11441

------------------------------------------------------------------------------------------

------------------------------------------------------------------------------

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

-----------------------------+------------------------------------------------

schoolnr: Unstructured |

var(IQ\_verb) | .1668484 .0717865 .0717948 .3877494

var(\_cons) | 8.428271 1.061685 6.584384 10.78852

cov(IQ\_verb,\_cons) | -.9477128 .2176479 -1.374295 -.5211307

-----------------------------+------------------------------------------------

var(Residual) | 37.36361 .9089777 35.62385 39.18833

------------------------------------------------------------------------------

LR test vs. linear regression: chi2(3) = 425.29 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

.

.

. \*Example 5.4

. \*A parsimonius model with several variables

. \*All cross-level interactions have been found to be nonsignificant

. \*Also no random slope for ses

.

.

. xtmixed langpost IQ\_verb ses c.IQ\_verb#c.ses gmeanIQverb gmeanSES || schoolnr: IQ\_verb , mle cova

> riance(un) var

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = -12320.95

Iteration 1: log likelihood = -12320.081

Iteration 2: log likelihood = -12320.08

Computing standard errors:

Mixed-effects ML regression Number of obs = 3758

Group variable: schoolnr Number of groups = 211

Obs per group: min = 4

avg = 17.8

max = 34

Wald chi2(5) = 2013.07

Log likelihood = -12320.08 Prob > chi2 = 0.0000

---------------------------------------------------------------------------------

langpost | Coef. Std. Err. z P>|z| [95% Conf. Interval]

----------------+----------------------------------------------------------------

IQ\_verb | 2.2302 .0631551 35.31 0.000 2.106418 2.353982

ses | .1751759 .0116656 15.02 0.000 .1523118 .19804

|

c.IQ\_verb#c.ses | -.0194105 .0048602 -3.99 0.000 -.0289363 -.0098847

|

gmeanIQverb | .9517005 .301435 3.16 0.002 .3608988 1.542502

gmeanSES | -.0906051 .0437531 -2.07 0.038 -.1763597 -.0048505

\_cons | 41.2504 .234182 176.15 0.000 40.79141 41.70939

---------------------------------------------------------------------------------

------------------------------------------------------------------------------

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

-----------------------------+------------------------------------------------

schoolnr: Unstructured |

var(IQ\_verb) | .1563774 .0704422 .0646754 .3781018

var(\_cons) | 8.784658 1.097335 6.876968 11.22155

cov(IQ\_verb,\_cons) | -.8626369 .2188356 -1.291547 -.4337269

-----------------------------+------------------------------------------------

var(Residual) | 37.41262 .9099561 35.67098 39.23929

------------------------------------------------------------------------------

LR test vs. linear regression: chi2(3) = 443.05 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

.

.

.

.

.

. \*\*\*\*\*\*\*\*\*\*\*\*\*\*

. \*\* Addendum \*\*

. \*\*\*\*\*\*\*\*\*\*\*\*\*\*

.

.

. \* Example 5.3

.

. \* No easy way to estimate this particular covariance structure of the random effects in STATA.

.

. \* However, this covariance structure was chosen only because the full covariance structure could no

> t be estimated in MLwiN.

.

. \* Possible solutions not covered here is to estimate the model as a structural equation model using

> the sem package in STATA or to

. \* run the model in MLwin.

. \* Note that, besides the possibility of MLwiN, in the LME4 package in R this covariance structure i

> s one of the defaults.

.

.

.

.

.

.

. \*Example 5.5

.

.

. \*The data source used in the book is presently unavailable.

.

. \*There is however an example (with data) of three level models in the STATA help file for xtmixed.

.

.

.

. help xtmixed

.

.

. webuse productivity, clear

(Public Capital Productivity)

. xtmixed gsp private emp hwy water other unemp || region: || state:, mle

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = 1430.5017

Iteration 1: log likelihood = 1430.5017

Computing standard errors:

Mixed-effects ML regression Number of obs = 816

-----------------------------------------------------------

| No. of Observations per Group

Group Variable | Groups Minimum Average Maximum

----------------+------------------------------------------

region | 9 51 90.7 136

state | 48 17 17.0 17

-----------------------------------------------------------

Wald chi2(6) = 18829.06

Log likelihood = 1430.5017 Prob > chi2 = 0.0000

------------------------------------------------------------------------------

gsp | Coef. Std. Err. z P>|z| [95% Conf. Interval]

-------------+----------------------------------------------------------------

private | .2671484 .0212591 12.57 0.000 .2254814 .3088154

emp | .754072 .0261868 28.80 0.000 .7027468 .8053973

hwy | .0709767 .023041 3.08 0.002 .0258172 .1161363

water | .0761187 .0139248 5.47 0.000 .0488266 .1034109

other | -.0999955 .0169366 -5.90 0.000 -.1331906 -.0668004

unemp | -.0058983 .0009031 -6.53 0.000 -.0076684 -.0041282

\_cons | 2.128823 .1543854 13.79 0.000 1.826233 2.431413

------------------------------------------------------------------------------

------------------------------------------------------------------------------

Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]

-----------------------------+------------------------------------------------

region: Identity |

sd(\_cons) | .038087 .0170591 .0158316 .091628

-----------------------------+------------------------------------------------

state: Identity |

sd(\_cons) | .0792193 .0093861 .0628027 .0999273

-----------------------------+------------------------------------------------

sd(Residual) | .0366893 .000939 .0348944 .0385766

------------------------------------------------------------------------------

LR test vs. linear regression: chi2(2) = 1154.73 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

.

. capture log close