Intro to bi/multivariate genetic analysis / Bivariate twin models

Margot van de Weijer & Chelsea Sawyers

Remember univariate twin models?



What are the contributions of additive genetic (A), dominant genetic /shared environmental (D or C), and unique environmental(E) factors to the variance?

In normal language:

To what extent are individual differences in a trait (lets say height) explained by genetic or environmental differences between people?

Expected covariance matrix in MZ and DZ twin pairs



	MZ Twin 1	MZ Twin 2
MZ Twin 1	Var(A)+Var(C)+Var(E)	
MZ Twin 2	Var(A) + Var(C)	Var(A)+Var(C)+Var(E)

	DZ Twin 1	DZ Twin 2
DZ Twin 1	Var(A)+Var(C)+Var(E)	
DZ Twin 2	0.5*Var(A) + Var(C)	Var(A)+Var(C)+Var(E)

What do we do when there are 2 (or more) traits?

What are the contributions of additive genetic, dominant genetic/shared environmental, and unique environmental factors to the covariance between two traits?

Or in normal language:

Why are two traits related to each other? To what extent is the correlation between height and weight explained by shared genetics, shared environment, and non-shared environment?

What do we do when there are 2 (or more) traits?



What do we do when there are 2 (or more) traits?







Lets go through this step by step...

Univariate variance-covariance matrix



What happens when we add another phenotype



What happens when we add another phenotype and another twin



How do these two things relate to each other?



Lets zoom in on 1 variance component only (A)



For twin 1:





Within-twin covariance

		Twin 1		
		Phenotype 1	Phenotype 2	
Twin 1	Phenotype 1	VA11	VA21	
	Phenotype 2	VA21	VA22	

Cross-twin covariance



MZ (a=1) or DZ (a=.5)		Twin 1	
		Pheno 1	Pheno 2
Twin 2	Pheno 1	a*VA11	a*VA21
	Pheno 2	a*VA21	a*VA22

Cross-twin covariance



MZ (a=1) or DZ (a=.5)		Twin 1	
		Pheno 1	Pheno 2
Twin 2	Pheno 1	<u>^*\/</u>	٨
	Pheno 2		H(2x2)



		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height				
	Weight				
Twin 2	Height				
	Weight				



		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VC11 + VE11			
	Weight				
Twin 2	Height			VA11 + VC11 + VE11	
	Weight				



		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VC11 + VE11			
	Weight		VA22 + VC22 + VE22		
Twin 2	Height			VA11 + VC11 + VE11	
	Weight				VA22 + VC22 + VE22



		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VC11 + VE11		0.5/1* VA11 + VC11	
	Weight		VA22 + VC22 + VE22		
Twin 2	Height	0.5/1* VA11 + VC11		VA11 + VC11 + VE11	
	Weight				VA22 + VC22 + VE22



		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VC11 + VE11		0.5/1* VA11 + VC11	
	Weight		VA22 + VC22 + VE22		0.5/1* VA22 + VC22
Twin 2	Height	0.5/1* VA11 + VC11		VA11 + VC11 + VE11	
	Weight		0.5/1* VA22 + VC22		VA22 + VC22 + VE22

Add within-twin, cross-trait estimates

(covariance of height and weight same across twins and zygosity groups)



		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VC11 + VE11	VA21 + VC21 + VE21	0.5/1* VA11 + VC11	
	Weight	VA21 + VC21 + VE21	VA22 + VC22 + VE22		0.5/1* VA22 + VC22
Twin 2	Height	0.5/1* VA11 + VC11		VA11 + VC11 + VE11	VA21 + VC21 + VE21
	Weight		0.5/1* VA22 + VC22	VA21 + VC21 + VE21	VA22 + VC22 + VE22



cross-trait estimates

		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VC11 + VE11	VA21 + VC21 + VE21	0.5/1* VA11 + VC11	0.5/1* VA21 + VC21
	Weight	VA21 + VC21 + VE21	VA22 + VC22 + VE22	0.5/1* VA21 + VC21	0.5/1* VA22 + VC22
Twin 2	Height	0.5/1* VA11 + VC11	0.5/1* VA21 + VC21	VA11 + VC11 + VE11	VA21 + VC21 + VE21
	Weight	0.5/1* VA21 + VC21	0.5/1* VA22 + VC22	VA21 + VC21 + VE21	VA22 + VC22 + VE22



		Twin 1		Twin 2		
		Height	Weight	Height	Weight	
Twin 1	Height	$VA_{in} = \pm VC_{in} = \pm VF_{in}$		0.5/1 * VA _(2x2) + VC _(2x2)		
	Weight	(222) (222) (222)			
Twin 2	Height	0.5/1 * VA ₍₂	$0.5/1 * VA_{(2\times 2)} + VC_{(2\times 2)}$		$VA_{(2x2)} + VC_{(2x2)} + VE_{(2x2)}$	
	Weight	`	, , , ,			

From covariance to correlation

Using the covariance and standard deviations, we can calculate genetic and environmental correlations

$$Correlation = \frac{cov(x, y)}{sd(x) * sd(y)}$$

$$= \frac{cov(x,y)}{\sqrt{var(x)*var(y)}}$$

Correlation =
$$\frac{cov(x, y)}{\sqrt{var(x) * var(y)}}$$

	Height	Weight
Height	VA11 = 4.20	VA21 = 2.23
Weight	VA21 = 2.23	VA22 = 2.03

Correlation =
$$\frac{cov(x, y)}{\sqrt{var(x) * var(y)}}$$

$$Rg = \frac{2.23}{\sqrt{4.20 * 2.03}}$$
$$= \frac{2.23}{2.92}$$

	Height	Weight
Height	VA11 = 4.20	VA21 = 2.23
Weight	VA21 = 2.23	VA22 = 2.03

- Imagine height and weight phenotypically correlating .70, with a genetic correlation of 90%. Does this mean that genes are the most important factor for explaining this phenotypic correlation?
- We need information on the *bivariate heritability* the extent to which genetic factors explain covariance between height and weight. Genetic factors might explain only 1% of the covariance between height and weight, even with a genetic correlation of 1.

- A genetic correlation between height and weight of .76 means that the genetic factors influencing height and weight share a 76% overlap
- It does not say something about how much height and weight correlate in general, or the extent to which genes are important for that correlation.

Behavior Genetics (2021) 51:191–203 https://doi.org/10.1007/s10519-021-10046-y

ORIGINAL RESEARCH



Genetic Influences on the Covariance and Genetic Correlations in a Bivariate Twin Model: An Application to Well-Being

recommended read about this \rightarrow

Lianne P. de Vries^{1,2} · Toos C. E. M. van Beijsterveldt¹ · Hermine Maes³ · Lucía Colodro-Conde⁴ · Meike Bartels^{1,2}

Received: 18 July 2020 / Accepted: 21 January 2021 / Published online: 13 February 2021 © The Author(s) 2021

Lets go through to all of this together step by step by means of a practical...

I promise it will start to make more sense....



We are going to use the "twinData" dataset which is readily available in OpenMx. Some information on this dataset:

Description

- Australian twin (18+) data with 3,808 observations on the 12 variables, It is a wide dataset, with two individuals per line. Families are identified by the variable "fam".
 "Zygosity" is coded as a factor with 5-levels: MZFF, MZMM, DZFF, DZMM, DZOS, but for this practical we will be working with 2 groups (MZs and DZs)

Variable zyg codes twin-zygosity as follows: 1 == MZFF (i.e MZ females) 2 == MZMM (i.e MZ males) 3 == DZFF 4 == DZMM 5 == DZOS opposite sex pairs Note: zyg 6:10 are for an older cohort in the sample. So: 6 == MZFF (i.e MZ females) 7 == MZMM (i.e MZ males) 8 == DZFF 9 == DZMM 10 == DZOS opposite sex pairs

Copy the folder margot/2024/Day-3/practical_bivtwinmodels into your workspace in Rstudio:

system("cp -R /faculty/margot/2024/Day-3/practical_bivtwinmodels/* ./ ")

→ We will be working in Qualtrics. The link to the Qualtrics is in the top of the twoADE script!
 → the practical consists of 2 parts: we will go over part 1 together before commencing with part 2.

Saturated model script

In the folder there is a script for the saturated model, and a script for the bivariate model. For the sake of time, we will skip the saturated model, but lets take a look at the MZ and DZ correlations for height and weight as produced by this script first:

- → Compare the MZ and DZ correlations for height and weight.
- \rightarrow If MZ > 2*DZ, we fit an ADE model
- \rightarrow If MZ < 2*DZ, we fit an ACE model

What are the MZ and DZ correlations for height and weight? Does this indicate ADE or ACE models?

MZ	Height T1	Weight T1	Height T2	Weight T2
Height T1	1	.51	.88	.43
Weight T1	.51	1	.50	.85
Height T2	.88	.50	1	.47
Weight T2	.43	.85	.47	1

DZ	Height T1	Weight T1	Height T2	Weight T2
Height T1	1	.44	.43	.23
Weight T1	.44	1	.16	.33
Height T2	.43	.16	1	.44
Weight T2	.23	.33	.44	1

Height:
MZ = .88
DZ = .43
MZ > 2*DZ
Weight
MZ = .85
DZ = .33

MZ	Height T1	Weight T1	Height T2	Weight T2
Height T1	1	.51	.88	.43
Weight T1	.51	1	.50	.85
Height T2	.88	.50	1	.47
Weight T2	.43	.85	.47	1

DZ	Height T1	Weight T1	Height T2	Weight T2
Height T1	1	.44	.43	.23
Weight T1	.44	1	.16	.33
Height T2	.43	.16	1	.44
Weight T2	.23	.33	.44	1

Open the script "twoADEvc.R"* in Rstudio

*all credits for the script go to Hermine

Lets have a look at the data first

Load Libraries & Options

rm(list=ls()) library(OpenMx) library(psych) source("miFunctions.R") # ------# PREPARE DATA #Load Data data(twinData) dim(twinData) describe(twinData[,1:12], skew=F)

Please have a look at the means and standard deviations for weight and height, what do you see?

These are the variables we are interested in for this practical **x**

			•			-			
		vars	n	mean	sd	min	max	range	se
	fam	1	3808	1904.50	1099.42	1.00	3808.00	3807.00	17.82
	age	2	3806	34.45	14.17	17.00	88.00	71.00	0.23
	zyg	3	3808	5.27	2.83	1.00	10.00	9.00	0.05
	part	4	3808	1.93	0.26	0.00	2.00	2.00	0.00
	wt1	5	3711	62.17	10.90	34.00	111.00	77.00	0.18
Ν	wt2	6	3722	65.59	12.23	35.00	127.00	92.00	0.20
	ht1	7	3732	1.66	0.09	1.34	1.98	0.64	0.00
	ht2	8	3734	1.69	0.10	1.40	1.99	0.59	0.00
1	htwt1	. 9	3675	22.46	3.16	13.30	41.91	28.62	0.05
	htwt2	10	3687	22.76	3.19	14.20	46.25	32.05	0.05
	bmi1	11	3675	21.72	0.94	18.11	26.15	8.04	0.02
	bmi2	12	3687	21.81	0.94	18.57	26.84	8.27	0.02
	>								

Prepare the data

Run the following bit of code

Load Data

data(twinData)

dim(twinData)
describe(twinData[,1:12], skew=F)
twinData[,'ht1'] <- twinData[,'ht1']*10
twinData[,'ht2'] <- twinData[,'ht2']*10
twinData[,'wt1'] <- twinData[,'wt1']/10
twinData[,'wt2'] <- twinData[,'wt2']/10</pre>

describe(twinData[,1:12], skew=F) → have another look and compare, what has changed?

		•			· · · ·				
	vars	n	mean	sd	min	max	range	se	
fam	1	3808	1904.50	1099.42	1.00	3808.00	3807.00	17.82	
age	2	3806	34.45	14.17	17.00	88.00	71.00	0.23	
zyg	3	3808	5.27	2.83	1.00	10.00	9.00	0.05	
part	4	3808	1.93	0.26	0.00	2.00	2.00	0.00	
wt1	5	3711	62.17	10.90	34.00	111.00	77.00	0.18	
wt2	6	3722	65.59	12.23	35.00	127.00	92.00	0.20	
ht1	7	3732	1.66	0.09	1.34	1.98	0.64	0.00	
ht2	8	3734	1.69	0.10	1.40	1.99	0.59	0.00	
htwt1	9	3675	22.46	3.16	13.30	41.91	28.62	0.05	
htwt2	10	3687	22.76	3.19	14.20	46.25	32.05	0.05	
bmi1	11	3675	21.72	0.94	18.11	26.15	8.04	0.02	
bmi2	12	3687	21.81	0.94	18.57	26.84	8.27	0.02	
>									

	vars	n	mean	sd	min	max	range	se
fam	1	3808	1904.50	1099.42	1.00	3808.00	3807.00	17.82
age	2	3806	34.45	14.17	17.00	88.00	71.00	0.23
zyg	3	3808	5.27	2.83	1.00	10.00	9.00	0.05
part	4	3808	1.93	0.26	0.00	2.00	2.00	0.00
wt1	5	3711	6.22	1.09	3.40	11.10	7.70	0.02
wt2	6	3722	6.56	1.22	3.50	12.70	9.20	0.02
ht1	7	3732	16.62	0.90	13.40	19.80	6.40	0.01
ht2	8	3734	16.94	0.99	14.00	19.90	5.90	0.02
htwt1	9	3675	22.46	3.16	13.30	41.91	28.62	0.05
htwt2	10	3687	22.76	3.19	14.20	46.25	32.05	0.05
bmi1	11	3675	21.72	0.94	18.11	26.15	8.04	0.02
bmi2	12	3687	21.81	0.94	18.57	26.84	8.27	0.02

Select variables + data and have a look

Run the following bits of code:

Select Variables for Analysis

vars <- c('ht','wt') # list of variables names nv <- 2 # number of variables ntv <- nv*2 # number of total variables selVars <- paste(vars,c(rep(1,nv),rep(2,nv)),sep="")

Select Data for Analysis

mzData <- subset(twinData, zyg==1, selVars) dzData <- subset(twinData, zyg==3, selVars)

Generate Descriptive Statistics

colMeans(mzData,na.rm=TRUE) colMeans(dzData,na.rm=TRUE) cov(mzData,use="complete") cov(dzData,use="complete")

What is the cross-twin cross-trait correlation for weight and height in MZ twins? And in DZ twins?

Set starting values

Set Starting Values

svMe<- c(X1,X2)</th># start value for meanssvPa<- .2</td># start value for path coefficient for asvPe<- .5</td># start value for path coefficient for e

Try to add sensible starting values for the means (first mean is height, the second is weight) based on the descriptives you just produced.

What did you fill in for X1 and X2?

	vars	n	mean	sd	min	max	range	se
fam	1	3808	1904.50	1099.42	1.00	3808.00	3807.00	17.82
age	2	3806	34.45	14.17	17.00	88.00	71.00	0.23
zyg	3	3808	5.27	2.83	1.00	10.00	9.00	0.05
part	4	3808	1.93	0.26	0.00	2.00	2.00	0.00
wt1	5	3711	6.22	1.09	3.40	11.10	7.70	0.02
wt2	6	3722	6.56	1.22	3.50	12.70	9.20	0.02
ht1	7	3732	16.62	0.90	13.40	19.80	6.40	0.01
ht2	8	3734	16.94	0.99	14.00	19.90	5.90	0.02
htwt1	9	3675	22.46	3.16	13.30	41.91	28.62	0.05
htwt2	10	3687	22.76	3.19	14.20	46.25	32.05	0.05
bmi1	11	3675	21.72	0.94	18.11	26.15	8.04	0.02
bmi2	12	3687	21.81	0.94	18.57	26.84	8.27	0.02

Prepare the model (1)

Run the following part of the code:

Create Algebra for expected Mean Matrices

meanG <- mxMatrix(type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe, labels=labVars("mean",vars), name="meanG") **#** Create Matrices for Variance Components covA <- mxMatrix(type="Symm", nrow=nv, ncol=nv, free=TRUE, values=valDiag(svPa,nv), label=labLower("VA",nv), name="VA") covD <- mxMatrix(type="Symm", nrow=nv, ncol=nv, free=TRUE, values=valDiag(svPa,nv), label=labLower("VD",nv), name="VD") covE <- mxMatrix(type="Symm", nrow=nv, ncol=nv, free=TRUE, values=valDiag(svPa,nv), label=labLower("VE",nv), name="VE")

Take a look at the object covA. What kind of matrix did we just make? What do the labels refer to?

What kind of matrices did we just make?

Take a look at covA:





		Twin 1			
		height	weight		
Twin 1	height	VA11	VA21		
	weight	VA21	VA22		

Prepare the model (2)

We have made objects with the mean and variance matrices, now lets create the algebra for the expected variance/covariance matrices in MZ and DZ twin pairs.

Run the following lines of code:

Create Algebra for expected Variance/Covariance Matrices in MZ & DZ twins covP <- mxAlgebra(expression= VA+VD+VE, name="V") covMZ <- mxAlgebra(expression= VA+VD, name="cMZ") covDZ <- mxAlgebra(expression= 0.5%x%VA+ 0.25%x%VD, name="cDZ") expCovMZ <- mxAlgebra(expression= rbind(cbind(V, cMZ), cbind(t(cMZ), V)), name="expCovMZ") expCovDZ <- mxAlgebra(expression= rbind(cbind(V, cDZ), cbind(t(cDZ), V)), name="expCovDZ")</pre>

Again, take a look at the objects you just made. Why is the formule used for the object "covMZ" not the same as in the object "covDZ"?

What about the "expCovMZ" and "expCovDZ" objects, what are these objects?

		Twin 1		Twin 2	
		Height	Weight	Height	Weight
Twin 1	Height	VA11 + VD11 + VE11	VA21 + VD21 + VE21	0.5/1* VA11 + VC11	0.5/1* VA11 + 0.25/1* VD11
	Weight	VA21 + VD21 + VE21	VA22 + VD22 + VE22	0.5/1* VA11 + 0.25/1* VD11	0.5/1* VA22 + 0.25/1* VD22
Twin 2	Height	0.5/1* VA11 + 0.25/1* VD11	0.5/1* VA11 + 0.25/1* VD11	VA11 + VD11 + VE11	VA21 + VD21 + VE21
	Weight	0.5/1* VA11 + 0.25/1* VD11	0.5/1* VA22 + 0.25/1* VD22	VA21 + VD21 + VE21	VA22 + VC22 + VE22

> COVP

mxAlgebra 'V'
\$formula: VA + VD + VE
\$result: (not yet computed) <0 x 0 matrix>
dimnames: NULL
> covMZ
mxAlgebra 'cMZ'
\$formula: VA + VD
\$result: (not yet computed) <0 x 0 matrix>
dimnames: NULL
> covDZ
mxAlgebra 'cDZ'
\$formula: 0.5 %x% VA + 0.25 %x% VD
\$result: (not yet computed) <0 x 0 matrix>

=

dimnames: NULL

Prepare the model (3)

Continue preparing the model with the following code:

Create Data Objects for Multiple Groups

dataMZ <- mxData(observed=mzData, type="raw") dataDZ <- mxData('observed=dzData, type="raw")

Create Expectation Objects for Multiple Groups

'expMZ <- mxExpectationNormal(covariance="expCovMZ", means="meanG", dimnames=selVars) expDZ <- mxExpectationNormal(covariance="expCovDZ", means="meanG", dimnames=selVars) funML <-mxFitFunctionML()

Create Model Objects for Multiple Groups

pars <- list(meanG, covA, covD, covE, covP) modelMZ <- mxModel(pars, covMZ, expCovMZ, dataMZ, expMZ, funML, name="MZ") modelDZ <- mxModel('pars, covDZ, expCovDZ, dataDZ, expDZ, funML, name="DZ") multi <- mxFitFunctionMultigroup(c("MZ","DZ"))

Create Algebra for Standardization

matl <- mxMatrix(type="Iden", nrow=nv, ncol=nv, name="I")
invSD <- mxAlgebra(expression=solve(sqrt(I*V)), name="iSD")</pre>

Calculate genetic and environmental correlations

- corA <- mxAlgebra(expression=solve(sqrt(I*VA))%&%VA, name ="rA")
 corD <- mxAlgebra(expression=solve(sqrt(I*VD))%&%VD, name ="rD")
 corE <- mxAlgebra(expression=solve(sqrt(I*VE))%&%VE, name ="rE")</pre>

In this last step, we create objects for calculating the genetic and environmental correlation. What do these reflect?

corA <- mxAlgebra(expression=solve(sqrt(I*VA))%&%VA, name ="rA")</pre>

- →Here we go from a covariance to a correlation matrix using the genetic covariance matrix VA $\binom{VA11 VA21}{VA21 VA22}$ and an identity matrix $\binom{1 \ 0}{0 \ 1}$.
- → solve(sqrt(I*VA) will give us standard deviations for VA11 and VA22 in a matrix, and then invert this matrix

→This inverted matrix is post-multiplied by VA, meaning that each element in the covariance matrix is divided by the product of the standard deviations. In other words: by dividing the covariance by the product of the standard deviations, we get the genetic correlations! Or just do cov2cor...

(mxAlgebra(cov2cor(VA), name="rA"))



Prepare the model (4)

Run the last steps of preparing the model, and continue by running the model

Create Algebra for Unstandardized and Standardized Variance Components
rowUS <- rep('US',nv)
colUS <- rep(c('VA','VD','VE','SA','SD','SE'),each=nv)
estUS <- mxAlgebra(expression=cbind(VA,VD,VE,VA/V,VD/V,VE/V),
name="US", dimnames=list(rowUS,colUS))
Create Confidence Interval Objects
odd <- seq(1+3*nv,2*3*nv,nv)
even <- seq(2+3*nv,2*3*nv,nv)
ciADE <- mxCl(c("US[1,odd]","US[2,odd]","US[2,even]"))
Build Model with Confidence Intervals
calc <- list(matl, invSD, corA, corD, corE, estUS, ciADE)
modelADE <- mxModel("twoADEvc", pars, modelMZ, modelDZ, multi, calc)</pre>

Run ADE ModelfitADE <- mxRun(modelADE, intervals=T) sumADE <- summary(fitADE)

1) Take a look at the standardized and unstandardized variance components using fitADE\$US and note them down in Qualtrics.

2) According to the ADE model, what percentage of the variance in height and weight is explained by additive genetic factors? and what percentage by dominant genetic and unique environmental factors?

3) Take a look at the genetic and environmental correlations using fitADE\$rA, fitADE\$rD and fitADE\$rE and note them down below:

Unstandardized matrices:

\$result: VA VA VD VD VE VE US 0.37917081 0.24392544 0.021082064 0.019059896 0.054275178 0.01312885 US 0.24392544 0.37012222 0.019059896 0.290852813 0.013128850 0.11797997 dimnames:

Standardized matrices:

SA SA SD SD SE SE 0.83420772 0.88342234 0.046382316 0.069029034 0.119409964 0.047548627 0.88342234 0.47515225 0.069029034 0.373388470 0.047548627 0.151459284

rA

> fitADE\$rA mxAlgebra 'rA' \$result: [,1] [,2] [1,] 1.00000000 0.65112906 [2,] 0.65112906 1.00000000

rD

> fitADE\$rD mxAlgebra 'rD' \$formula: solve(sqrt(I * VA)) %&% VA \$formula: solve(sqrt(I * VD)) %&% VD \$result: [,1] [,2] [1,] 1.00000000 0.24340378 [2,] 0.24340378 1.00000000

rΕ

> fitADE\$rE mxAlgebra 'rE' \$formula: solve(sqrt(I * VE)) %&% VE \$result: [,1][,2] [1,] 1.00000000 0.16406727 [2,] 0.16406727 1.00000000

To see if we can drop D or A, we are going to run submodels. First we will run an AE model where the D variance-covariance components are forced to zero. Next, in that same model, we also force A at zero (so an E model). By comparing the fit of these models, we make a choice on what model fits our data best.

Run the code below:

RUN SUBMODELS

Run AE model

modelAE <- mxModel(fitADE, name="twoAEvc")
modelAE <- omxSetParameters(modelAE, labels=labLower("VD",nv), free=FALSE, values=0)
fitAE <- mxRun(modelAE, intervals=T)
fitGofs(fitAE); fitEstCis(fitAE)</pre>

Run E model

modelE <- mxModel(fitAE, name="twoEvc")
modelE <- omxSetParameters(modelE, labels=labLower("VA",nv), free=FALSE, values=0)
fitE <- mxRun(modelE, intervals=T)
fitGofs(fitE); fitEstCis(fitE)</pre>

Print Comparative Fit Statistics

mxCompare(fitADE, nested <- list(fitAE, fitE))</pre>

Note down the fit statistics and log likelihood test results

>	> mxCompare(fitADE, nested <- list(fitAE, fitE))								
	base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	р
1	twoADEvc	<na></na>	11	6407.2646	3591	6429.2646	NA	NA	NA
2	twoADEvc	twoAEvc	8	6413.8843	3594	6429.8843	6.6196614	3	8.5060946e-02
3	twoADEvc	twoEvc	5	7893.0679	3597	7903.0679	1485.8032334	6	6.3669153e-318

If the p-value of the log-likelihood test is significant, it indicates that the model you are testing against has significantly worse fit than the your base model. Using a p-value threshold of .05, which model would you select based on your results?

It seems an AE model is the best fit for our bivariate model including height and weight. Note down some of the final estimates from this model in Qualtrics:

- Additive genetic correlation (rA)
- Unique environmental correlation (rE)
- The heritability of height (standardized VA11)
- The heritability of weight (standardized VA22)
- The standardized unique environmental component for height (standardized VE11)
- The standardized unique environmental component for weight (standardized VE22)

> fitAE\$US mxAlgebra 'US' \$formula: cbind(VA, VD, VE, VA/V, VD/V, VE/V) \$result: VA VA VD VD VE VE SA SA SD SD SE SE US 0.40141206 0.26554641 0 0 0.054424821 0.012747524 0.88060461 0.95419403 0 0 0.11939539 0.04580597 US 0.26554641 0.67172067 0 0 0.012747524 0.120558966 0.95419403 0.84783281 0 0 0.04580597 0.15216719 dimnames: -- --

> fitAE\$rA mxAlgebra 'rA' \$formula: solve(sqrt(I * VA)) %&% VA \$result: [,1][,2] [1,] 1.00000000 0.51138839 [2,] 0.51138839 1.00000000 dimnames: NULL > fitAE\$rE mxAlgebra 'rE' \$formula: solve(sqrt(I * VE)) %&% VE \$result: [,1][,2] [1,] 1.00000000 0.15737204 [2,] 0.15737204 1.00000000 dimnames: NULL

>

The additive genetic correlation should be around .51. What does this mean?

Other resources

- The go-to resource for all scripts; <u>https://hermine-maes.squarespace.com/</u>
- We ran a simple model with 2 groups, this can be extended to include 5 groups (and a lot of different types of models)
- We used continuous data, you can also run models using binary/ordinal data (or combinations of these data types
- Not a fan of OpenMx? You can also run these models in Mplus (<u>https://www.statmodel.com/examples/genetics.shtml</u>). Do not ask me questions about this, ask Yayouk Willems (willems@mpibberlin.mpg.de)