Univariate/ MonoPhenotype Twin Modeling in OpenMx

Boulder workshop 2021

Hermine H. Maes

with credit to Nick Martin, Elizabeth Prom-Wormley, Lindon Eaves, Tim Bates, Michael Neale & many others

Files & Code

- On https://hermine-maes.squarespace.com
- Saturated Model
  - oneSATc.R
- Direct Variance Estimation of Genetic Models
  - oneACEvc.R
  - oneADEvc.R
- Path Coefficient Estimation of Genetic Models
  - oneACEc.R
  - oneADEc.R
- miFunctions.R
  - miFunctionsDocs.pdf

Questions

- Does trait of interest run in families?
- Can familial resemblance be explained by genetic and/or environmental effects?
- Which sources of variance contribute significantly to the variance of the trait?
- How much of trait variation is accounted for by genetic and environmental factors?

Roadmap for Univariate Analysis

- Use data to test basic assumptions (equal means & variances for members of twin pairs)
  - Saturated Model
- Estimate contributions of genetic/environmental effects on total variance of a phenotype
  - ACE or ADE Models
- Test ACE / ADE submodels to identify and report significant genetic and environmental contributions
  - AE / CE / E Only Model

Practical Example

- Dataset: NH&MRC Twin Register
- 1981 Questionnaire
- BMI (body mass index): weight/height squared
  - kg/m^2, transformed: 7*log(BMI), simulated based on real data
- Young Female Cohort: 18-30 years
- Sample Size:
  - MZf: 534 pairs (zyg=1; zygosity='MZFF' & cohort='younger')
  - DZf: 328 pairs (zyg=3; zygosity='DZFF' & cohort='younger')

Dataset

```
> head(twinData)
  fam age zyg part wt1 wt2    ht1    ht2   htwt1   htwt2    bmi1    bmi2
1   1  21   1    2  58  57 1.7000 1.7000 20.0692 19.7232 20.9943 20.8726
2   2  24   1    2  54  53 1.6299 1.6299 20.3244 19.9481 21.0828 20.9519
3   3  21   1    2  55  50 1.6499 1.6799 20.2020 17.7154 21.0405 20.1210
4   4  21   1    2  66  76 1.5698 1.6499 26.7759 27.9155 23.0125 23.3043
5   5  19   1    2  50  48 1.6099 1.6299 19.2894 18.0662 20.7169 20.2583
6   6  26   1    2  60  60 1.5999 1.5698 23.4375 24.3418 22.0804 22.3454
....
```

My Naming Conventions

```
name of variable(s)              vars     <- 'bmi'
number of variables              nv       <- 1
number of twin variables         ntv      <- nv*2
variables per twin pair          selVars <-c('bmi1','bmi2')
definition variables             covVars
number of factors                nf       <- 2
number of thresholds             nth      <- 3

starting values                  sv
lower bound / upper bound        lb / ub
labels                           lab

built model                      modelNAME
fitted model                     fitNAME
summary of fitted model          sumNAME
```

Classical Twin Study Background

- The Classical Twin Study/Design (CTS/CTD) uses monozygotic (MZ) and dizygotic (DZ) twins reared together
  - MZ twins share 100% of their genes
  - DZ twins share on average 50% of their genes
- Genetic factors are assumed to contribute to a phenotype when MZ twins are more similar for that phenotype than DZ twins

Partitioning Variation & Estimating Heritability

- Partition phenotypic variance (V) in genetic and environmental components
- $V = V_{genetic} + V_{environmental}$
- Assumptions: additivity & independence of effects
- Heritability (h^2): proportion of variance due to genetic influences ($h^2 = V_{genetic} / V$)
- Property of a group (not an individual), thus specific to that group in place & time

Sources of Variance

- Additive genetic factors (VA, **A**, a^2): sum of all average effects of single alleles at individual loci
- Dominance: result of interactions between alleles at same locus (VD, **D**, d^2)
- Common environment: (VC, **C**, c^2): aspects of environment shared by family members, which contribute to similarity between relatives [shared]
- Environmental factors (VE, **E**, e^2): unique to individual, contribute to variation within family [specific, unique or within-family]

Classical Twin Study Assumptions

- Equal Environments of MZ and DZ pairs (EEA)
  - MZ & DZ twins equally correlated in exposure to environmental events of etiologic importance for trait
- Random Mating
- No rGE Correlation
- No G x E Interaction
- No Sex Limitation
- No G x Age Interaction

Classical Twin Study Basic Data Assumptions

- MZ and DZ twins are sampled from the same population, therefore we expect :
  - Equal means/variances in Twin 1 and Twin 2
  - Equal means/variances in MZ and DZ twins
- Further assumptions would need to be tested if we introduce male twins and opposite sex twin pairs

Observed Values

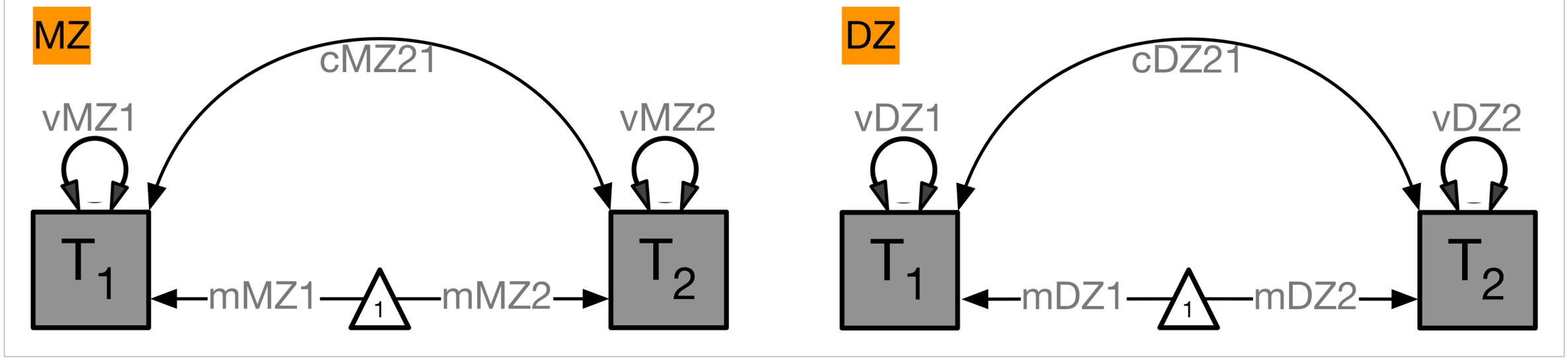
	Descriptive Statistics					
	MZ twins			DZ twins		
		T1	**T2**		**T1**	**T2**
mean	**MZ**	21.34	21.35	**DZ**	21.45	21.46
		T1	**T2**		**T1**	**T2**
covariance	**T1**	0.73	0.59	**T1**	0.77	0.24
	T2	0.59	0.79	**T2**	0.24	0.82

‘Old Fashioned’ Data Checking

Nice, but how can we actually be sure that these means and variances are truly the same?

Saturated Model

SAT model
oneSATc.R



Intuition behind Maximum Likelihood (ML)

- Likelihood: probability that observation (data point) is predicted by specified model
- Maximum Likelihood Estimates (MLE): most likely values of population parameter values (e.g, μ, σ, β) given observed sample values
  - Define model
  - Define probability of observing a given event conditional on a particular set of parameters
  - Choose set of parameters which are most likely to have produced observed results

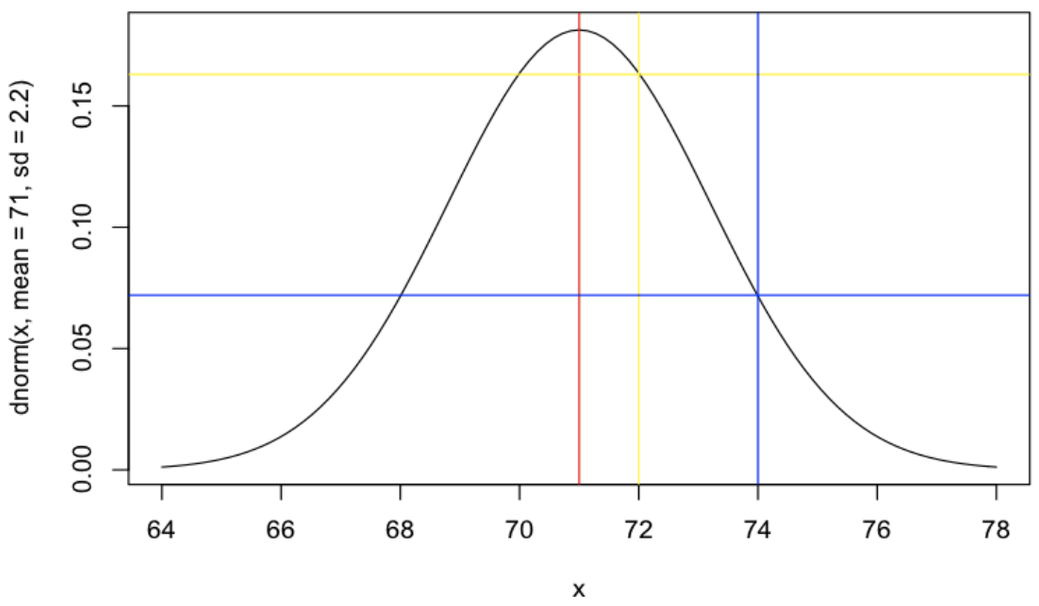
Likelihood Ratio Test

- Likelihood Ratio (LR) test: simple comparison of Log-Likelihoods under 2 separate models:
  - Model Mu: Unconstrained (has more parameters)
  - Model Mc: Constrained (has fewer parameters)
- LR statistic equals:
  - LR (Mc | Mu) = 2ln(L(Mu) - 2ln(L(Mc)
- LR is asymptotically distributed as χ^2 with degrees of freedom (df) equal to number of constraints

Probability Density Function Φ(xi)

- Φ(xi): likelihood of data point x_i for particular mean and variance estimates
- Univariate: height of probability density function

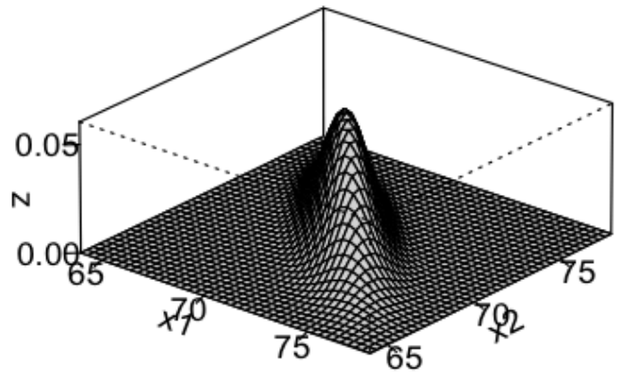
$$\Phi\left(x_i\right) = -\left|2\pi\sigma^2\right|^{-.5}\ e^{-5((x_i-\mu)^2/\sigma^2)}$$



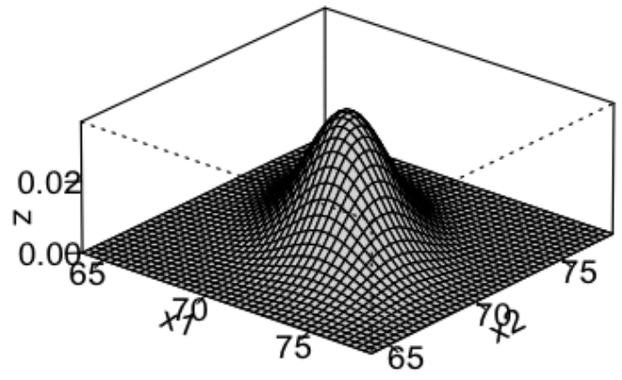
π: pi=3.14; x_i: observed value of variable i; μ: expected mean; σ: expected variance

Multinormal Probability Function

- $\Phi(x_i)$: likelihood of pair of data points x_i and y_i for particular means, variances & correlation estimates
- Multivariate: height of multinormal probability density function



rMZ=.85



rDZ=.49

$$\Phi\left(x_i\right) = -|2\pi\Sigma|^{-n/2}\ e^{-5((x_i-\mu)\Sigma^{-1}(x_i-\mu)')}$$

π= 3.14; x_i: value of variable i; μ: expected mean; ∑: expected covariance matrix

OpenMx Scripts

- oneSATc.R
  - Saturated model estimating means & (co)variances for continuous data in MZ & DZ twins
- oneACEvc.R
  - Univariate/Monophenotype model estimating A, C & E components for continuous data in MZ & DZ twins
- oneADEvc.R
  - Univariate/Monophenotype model estimating A, D & E components for continuous data in MZ & DZ twins

OpenMx Development Team (most active ones)

Steve Boker

https://openmx.ssri.psu.edu

Mike Neale

OpenMx is free and open source software for use with R that allows estimation of a wide variety of advanced multivariate statistical models.

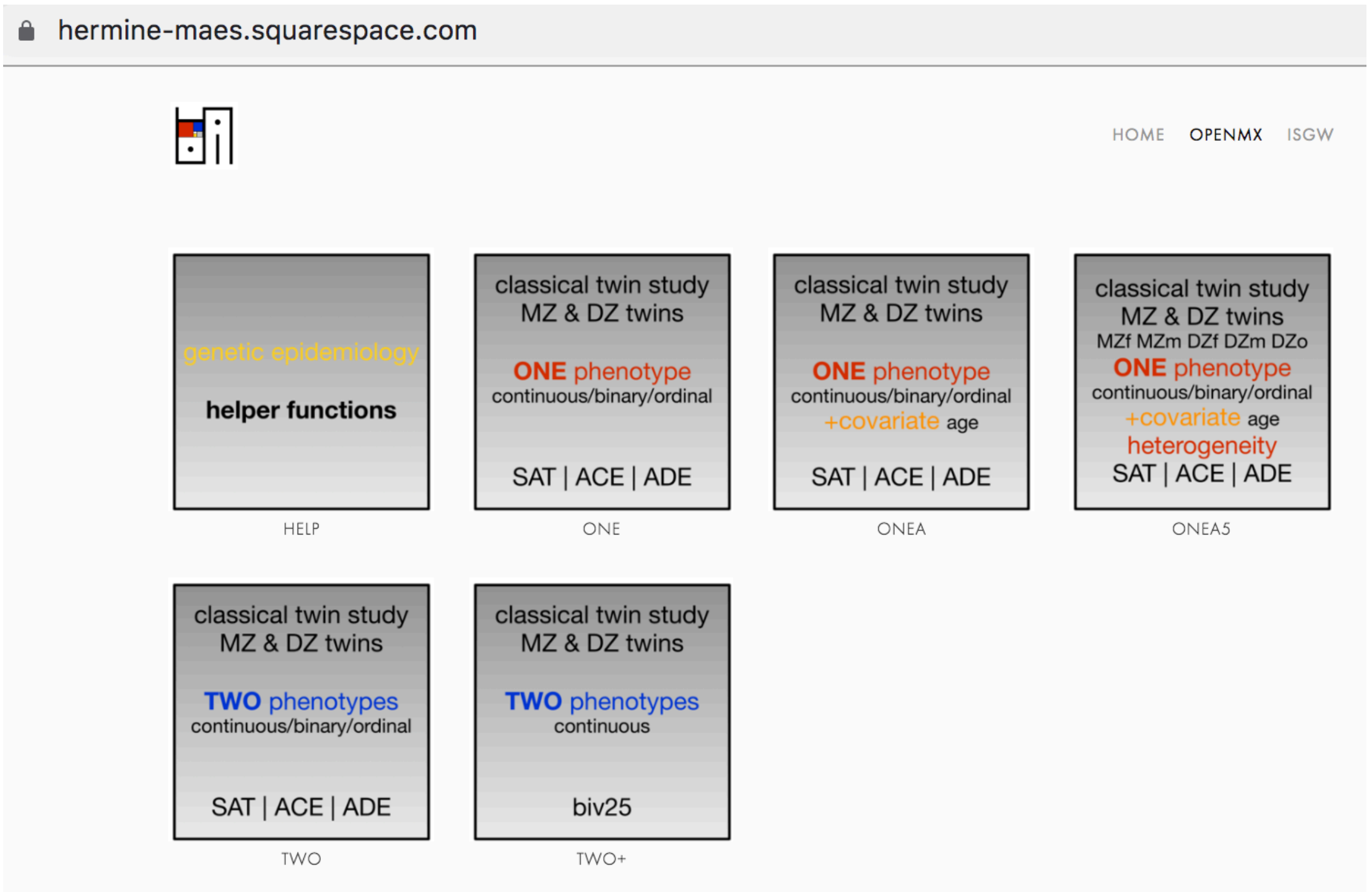
Mike Hunter

Rob Kirkpatrick

Joshua Pritikin

Tim Bates

https://hermine-maes.squarespace.com



helper functions

HOME OPENMX ISGW

Parallel standard OpenMx scripts & matching umx versions

This site offers OpenMx and matching umx scripts to fit standard biometrical models to data collected in MZ and DZ twins. The models can be used to estimate the role of genetic (A: additive genetic factors; D: dominance genetic factors) and environmental factors (C: common/shared environmental factors, E: unique environmental factors) to the variance of phenotypes of interest and covariances between phenotypes of interest.

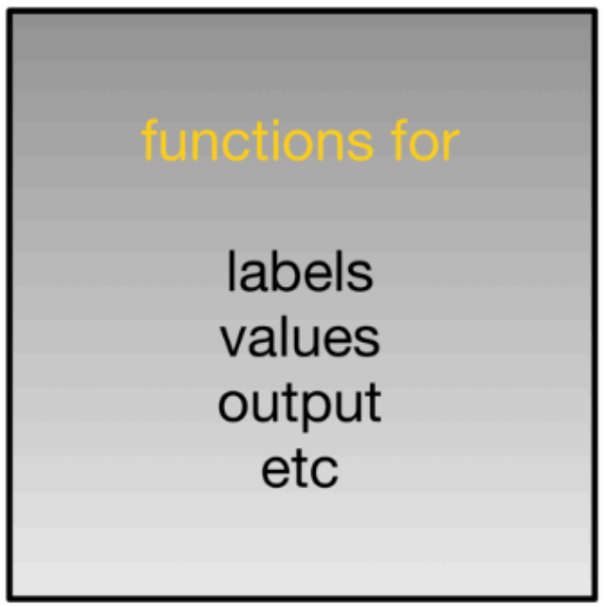
Scripts are organized in pages by number of phenotypes, addition of covariates etc. Within each page, scripts are organized by type of model in rows (Saturated, ACE estimating variance components, ADE estimating variance components, ACE estimating path coefficients, ADE estimating path coefficients) and by type of data in columns (continuous, binary, ordinal (estimating all thresholds), ordinal (fixing two thresholds and estimating means/variances) using standard code, and using umx (for different data types) in the last column. All scripts source the R code attached here that includes a number of functions that automate various aspects of the models such as labels, starting values, output generated etc.

Note that each of the scripts is represented by a path diagram. If you click on the diagram, a PDF of the associated script will be displayed in a separate window. If you click on the filename.R below the diagram, the R script will be downloaded.

Comments, suggestions, corrections welcome!! **Email hmaes@vcu.edu.**

New pages/scripts will be added as they are ready to go! If you have scripts that you'd like to add, send them my way!

Last updated: 02/28/2020



miFunctions.R

miFunctions.R

```
# ---------------------------------------------------------------------------
# Program: miFunctions.R
#  Author: Hermine Maes
#    Date: 02 28 2020
#
# Set of my options & functions used in basic twin methodology scripts
#   Email: hmaes@vcu.edu
# -------|---------|---------|---------|---------|---------|---------|---------|

# Options
mxOption( NULL, "Default optimizer", "NPSOL" )
#mxOption( NULL, "Checkpoint Prefix", filename )
mxOption( NULL, "Checkpoint Units", "iterations" )
mxOption( NULL, "Checkpoint Count", 1 )
options(width=120)
options(digits=8)
mxVersion()


# Functions to assign labels
# ---------------------------------------------------------------------------
labLower  <- function(lab,nv) { paste(lab,rev(nv+1-sequence(1:nv)),rep(1:nv,nv:1),sep="") }
labSdiag  <- function(lab,nv) { paste(lab,rev(nv+1-sequence(1:(nv-1))),rep(1:(nv-1),(nv-1):1),sep="") }
labOdiag  <- function(lab,nv) { paste(lab,c(rev(nv+1-sequence(1:(nv-1))),rep(1:(nv-1),(nv-1):1)),c(rep(1:(nv-1),
 (nv-1):1),rev(nv+1-sequence(1:(nv-1)))),sep="") }
```

….

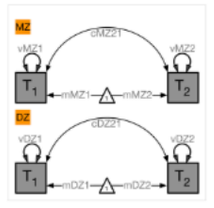
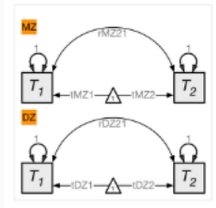
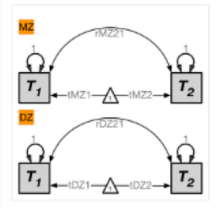
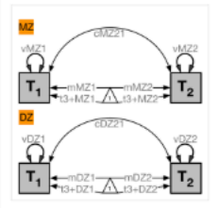
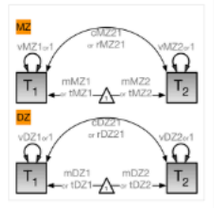
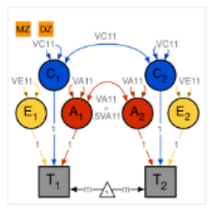
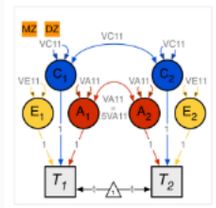
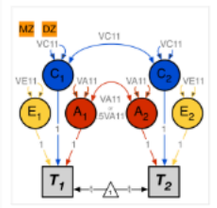
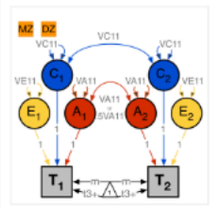
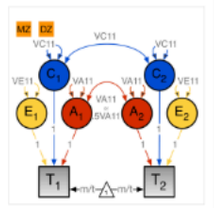
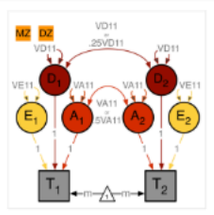
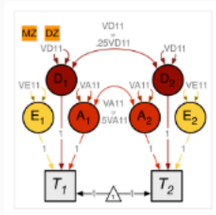
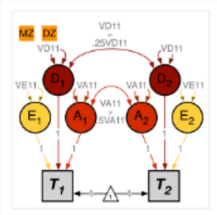
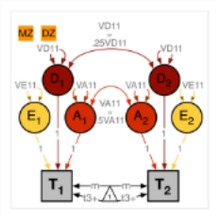
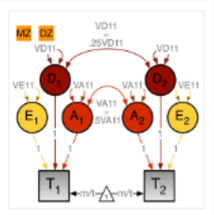
miFunctionsDocs.pdf

one

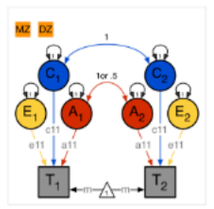
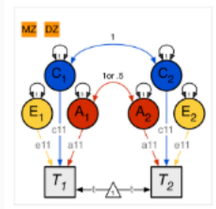
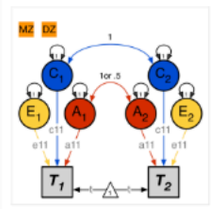
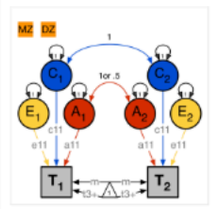
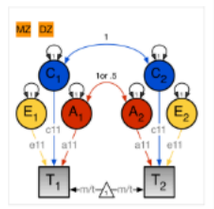
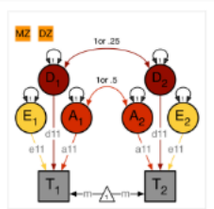
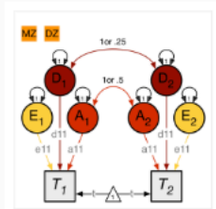
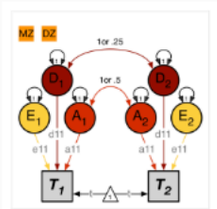
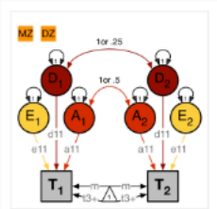
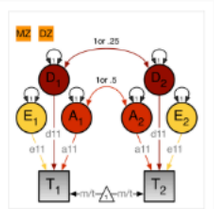
classical twin study
MZ & DZ twins

ONE phenotype
continuous/binary/ordinal

SAT | ACE | ADE

.	One Phenotype **CONTINUOUS** c	One Phenotype **BINARY** b	One Phenotype **ORDINAL** o	One Phenotype **ORDINAL** m *	One Phenotype **UMX** c/b/m
SAT estimating means/thresholds, variances & covariances * > 2 categories					
	oneSATc.R	**oneSATb.R**	**oneSATo.R**	**oneSATm.R**	**oneSATu.R**
ACE estimating variance components					
	oneACEvc.R	**oneACEvb.R**	**oneACEvo.R**	**oneACEvm.R**	**oneACEvu.R**
ADE estimating variance components					
	oneADEvc.R	**oneADEvb.R**	**oneADEvo.R**	**oneADEvm.R**	**oneADEvu.R**

NOTE! Models below estimating path coefficients may provide biased estimates of the parameters. Use of these scripts is discouraged.

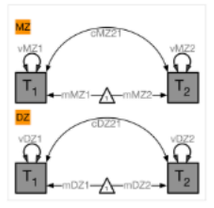
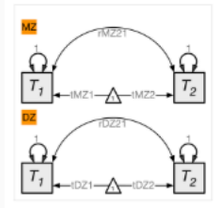
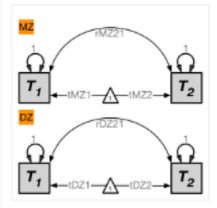
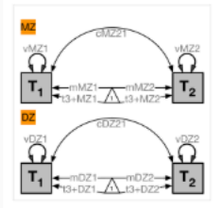
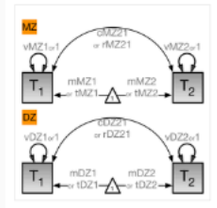
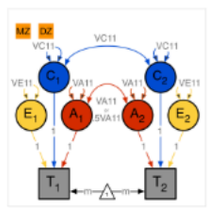
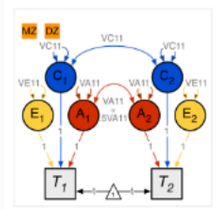
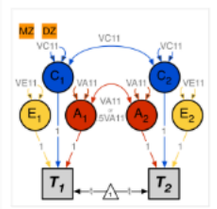
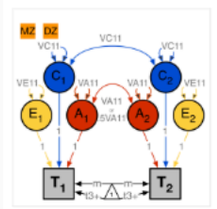
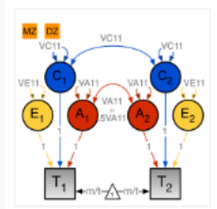
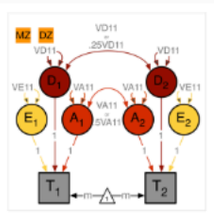
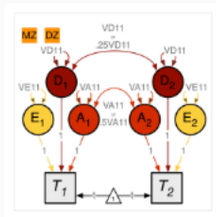
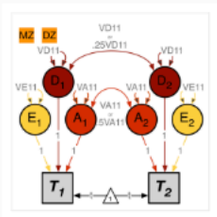
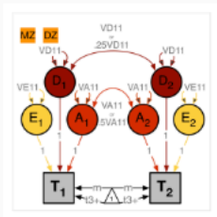
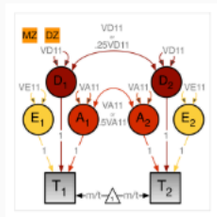
ACE estimating path coefficients					
	oneACEc.R	**oneACEb.R**	**oneACEo.R**	**oneACEm.R**	**oneACEu.R**
ADE estimating path coefficients					
	oneADEc.R	**oneADEb.R**	**oneADEo.R**	**oneADEm.R**	**oneADEu.R**

one

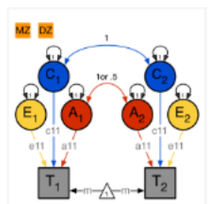
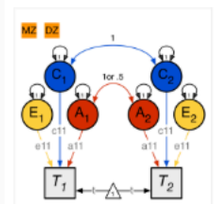
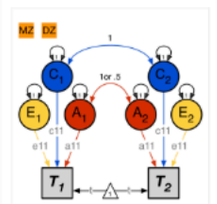
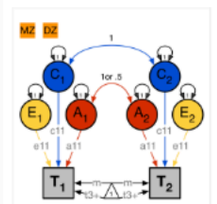
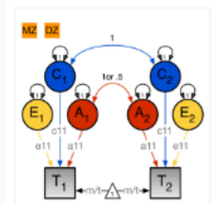
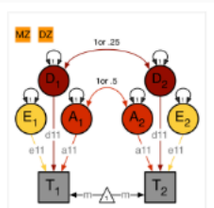
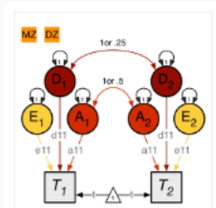
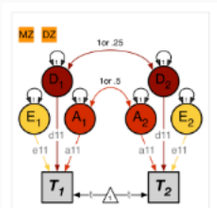
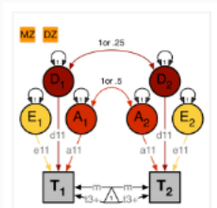
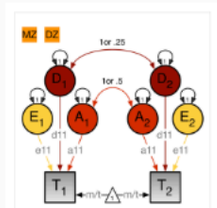
classical twin study
MZ & DZ twins

ONE phenotype
continuous/binary/ordinal

SAT | ACE | ADE

.	One Phenotype **CONTINUOUS** c	One Phenotype **BINARY** b	One Phenotype **ORDINAL** o	One Phenotype **ORDINAL** m *	One Phenotype **UMX** c/b/m
SAT estimating means/thresholds, variances & covariances * > 2 categories	 **oneSATc.R**	 **oneSATb.R**	 **oneSATo.R**	 **oneSATm.R**	 **oneSATu.R**
ACE estimating variance components	 **oneACEvc.R**	 **oneACEvb.R**	 **oneACEvo.R**	 **oneACEvm.R**	 **oneACEvu.R**
ADE estimating variance components	 **oneADEvc.R**	 **oneADEvb.R**	 **oneADEvo.R**	 **oneADEvm.R**	 **oneADEvu.R**

NOTE! Models below estimating path coefficients may provide biased estimates of the parameters. Use of these scripts is discouraged.

ACE estimating path coefficients	 **oneACEc.R**	 **oneACEb.R**	 **oneACEo.R**	 **oneACEm.R**	 **oneACEu.R**
ADE estimating path coefficients	 **oneADEc.R**	 **oneADEb.R**	 **oneADEo.R**	 **oneADEm.R**	 **oneADEu.R**

Layout of my OpenMx Scripts

```
# ----------------------------------------------------------------------------
# Program: oneSATc.R
#  Author: Hermine Maes
#    Date: 10 22 2018
#
# Twin Univariate Saturated model to estimate means and (co)variances across multiple groups
# Matrix style model - Raw data - Continuous data
# -------|---------|---------|---------|---------|---------|---------|---------|
# Load Libraries & Options
# Create Output
# ----------------------------------------------------------------------------
# PREPARE DATA
# Load Data
# Select Variables for Analysis
# Select Data for Analysis
# Generate Descriptive Statistics
# Set Starting Values
# ----------------------------------------------------------------------------
# PREPARE MODEL
# Create Algebra for expected Mean Matrices
# Create Algebra for expected Variance/Covariance Matrices
# Create Data Objects for Multiple Groups
# Create Expectation Objects for Multiple Groups
# Create Model Objects for Multiple Groups
# Create Confidence Interval Objects
# Build Saturated Model with Confidence Intervals
# ----------------------------------------------------------------------------
# RUN MODEL
# Run Saturated Model
# Print Goodness-of-fit Statistics & Parameter Estimates
```

Univariate Saturated Model oneSATc.R

```
# ------------------------------------------------------------------------------
# Program: oneSATc.R
#  Author: Hermine Maes
#    Date: 10 22 2018
#
# Twin Univariate Saturated model to estimate means and (co)variances across multiple groups
# Matrix style model - Raw data - Continuous data
# -------|---------|---------|---------|---------|---------|---------|---------|

# Load Libraries & Options
rm(list=ls())
library(OpenMx)
library(psych); library(polycor)
source("miFunctions.R")

# Create Output
filename    <- "oneSATc"
sink(paste(filename,".Ro",sep=""), append=FALSE, split=TRUE)
```

load OpenMx

my functions which you can edit as you like

creates output file with extension .Ro

Preparing Data

oneSATc.R

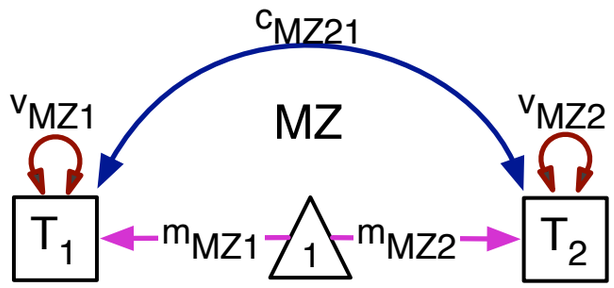
```
# ------------------------------------------------------------------------------
# PREPARE DATA

# Load Data
data(twinData)
dim(twinData)
describe(twinData[,1:12], skew=F)

# Select Variables for Analysis
vars      <- 'bmi'                    # list of variables names
nv        <- 1                        # 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)  # zygosity='MZFF' & cohort='younger'
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")

# Set Starting Values
svMe      <- 20                       # start value for means
svVa      <- .8                       # start value for variance
lbVa      <- .0001                    # lower bound for variance
```

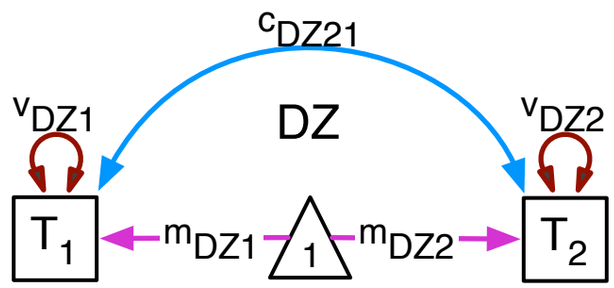
load 'twinData' or read in your own

analyzing c('bmi1','bmi2')

get right codes for zygosity

SAT Deconstructed: *Covariance Matrices & Means*






```
meanMZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv,
 free=TRUE, values=svMe, labels=c("mMZ1","mMZ2"),name="meanMZ" )
meanDZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv,
 free=TRUE, values=svMe, labels=c("mDZ1","mDZ2"),name="meanDZ" )
```

mMZ1	mMZ2

meanMZ *1x2*

mDZ1	mDZ2

meanDZ *1x2*

```
covMZ      <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv,
 free=TRUE, values=valDiag(svVa,ntv), lbound=valDiag(lbVa,ntv),
 labels=c("vMZ1","cMZ21","vMZ2"), name="covMZ" )
covDZ      <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv,
 free=TRUE, values=valDiag(svVa,ntv), lbound=valDiag(lbVa,ntv),
 labels=c("vDZ1","cDZ21","vDZ2"), name="covDZ" )
```

vMZ1	cMZ21
cMZ21	vMZ2

covMZ *2x2*

vDZ1	cDZ21
cDZ21	vDZ2

covDZ *2x2*

OpenMx Commands

```
mxMatrix( type="Full", nrow=1, ncol=ntv,
 free=TRUE, values=svMe, labels=c("mMZ1","mMZ2"),name="meanMZ" )

mxData( observed=mzData, type="raw" )

mxExpectationNormal( covariance="covDZ", means="meanDZ", dimnames=selVars )

mxFitFunctionML()

mxModel( meanMZ, covMZ, dataMZ, expMZ, funML, name="MZ" )

mxFitFunctionMultigroup( c("MZ","DZ") )

mxCI( c('MZ.covMZ','DZ.covDZ') )

mxRun( modelSAT, intervals=F )

mxGetExpected( fitSAT, c("means","covariance") )

omxSetParameters( modelEMVZ, label=c("vMZ","vDZ"), free=TRUE, newlabels='vZ' )

mxCompare( fitSAT, fitEMO )
```

Preparing Model

oneSATc.R

```
# ------------------------------------------------------------------------------
# PREPARE MODEL


# Create Algebra for expected Mean Matrices
meanMZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe, labels=c("mMZ1","mMZ2"), name="meanMZ" )
meanDZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe, labels=c("mDZ1","mDZ2"), name="meanDZ" )

# Create Algebra for expected Variance/Covariance Matrices
covMZ      <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=TRUE, values=valDiag(svVa,ntv), lbound=valDiag(lbVa,ntv),
 labels=c("vMZ1","cMZ21","vMZ2"), name="covMZ" )
covDZ      <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=TRUE, values=valDiag(svVa,ntv), lbound=valDiag(lbVa,ntv),
 labels=c("vDZ1","cDZ21","vDZ2"), name="covDZ" )

# 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="covMZ", means="meanMZ", dimnames=selVars )
expDZ      <- mxExpectationNormal( covariance="covDZ", means="meanDZ", dimnames=selVars )
funML      <- mxFitFunctionML()
```

full matrix for means

symmetric matrix for covariances

fitting to raw data

link to data

using FIML: full information maximum likelihood

Run Model

oneSATc.R

```
# Create Model Objects for Multiple Groups
modelMZ    <- mxModel( meanMZ, covMZ, dataMZ, expMZ, funML, name="MZ" )
modelDZ    <- mxModel( meanDZ, covDZ, dataDZ, expDZ, funML, name="DZ" )
multi      <- mxFitFunctionMultigroup( c("MZ","DZ") )

# Create Confidence Interval Objects
ciCov      <- mxCI( c('MZ.covMZ','DZ.covDZ') )
ciMean     <- mxCI( c('MZ.meanMZ','DZ.meanDZ') )

# Build Saturated Model with Confidence Intervals
modelSAT     <- mxModel( "oneSATc", modelMZ, modelDZ, multi, ciCov, ciMean )
```

model object contains all matrices etc.

evaluating 2 groups simultaneously

built model

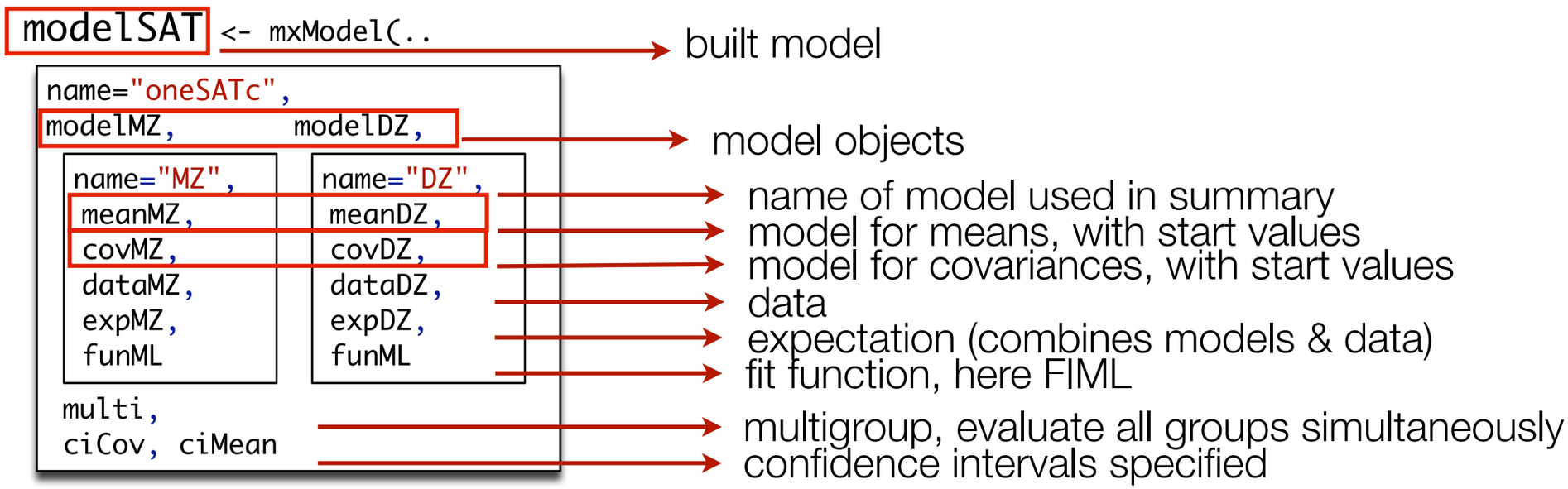
```
# ------------------------------------------------------------------------------
# RUN MODEL

# Run Saturated Model
fitSAT     <- mxRun( modelSAT, intervals=F )
sumSAT     <- summary( fitSAT )

# Print Goodness-of-fit Statistics & Parameter Estimates
fitGofs(fitSAT)
fitEsts(fitSAT)
mxGetExpected( fitSAT, c("means","covariance") )
```

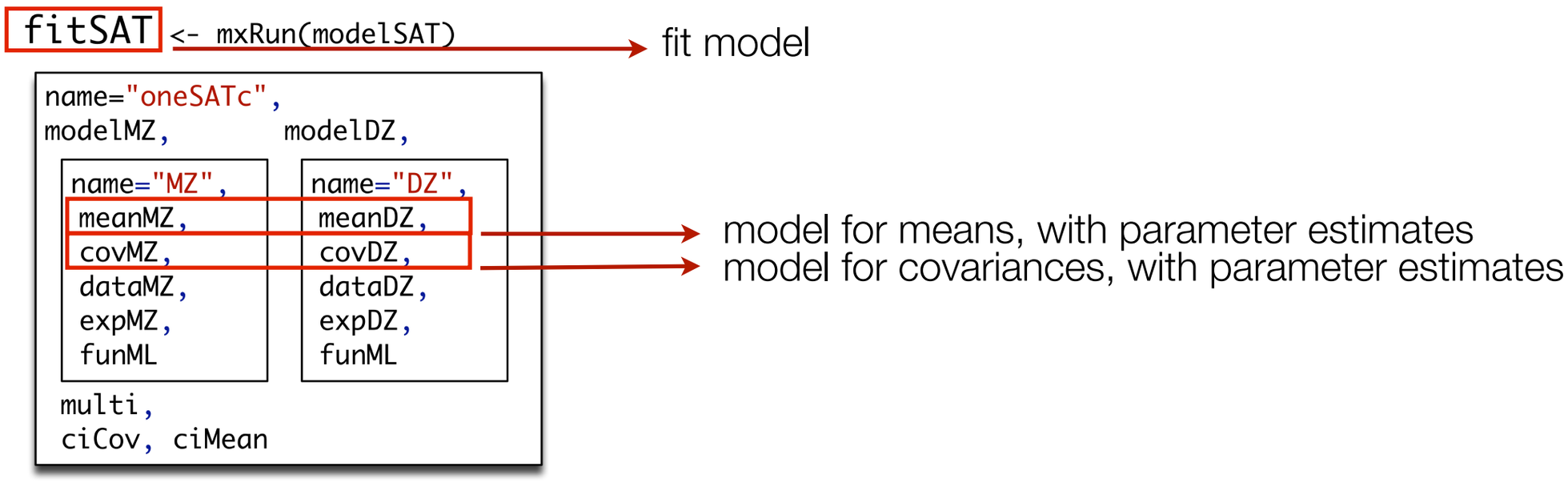
fitted model

standard summary function in OpenMx

my short summary function in miFunctions.R

Model Building - Model Fitting





more of miFunctions.R

```
# Functions to generate output

fitGofs   <- function(fit) {
          summ <- summary(fit)
          cat(paste("Mx:", fit$name,"  os=", summ$ob,"  ns=", summ$nu,"   ep=", summ$es,
                    "   co=", sum(summ$cons),"  df=", summ$de, "  ll=", round(summ$Mi,4),
                    "  cpu=", round(summ$cpu,4),"  opt=", summ$op,"  ver=", summ$mx,
                    "  stc=", fit$output$status$code, "\n",sep=""))
}


fitEsts   <- function(fit) {
  print(round(fit$output$estimate,4))
}

fitEstCis   <- function(fit) {
  print(round(fit$output$estimate,4))
  print(round(fit$output$confidenceIntervals,4))
}
```

Print Goodness-of-Fit Statistics

```
> summary(fitSAT)
Model Statistics:
               |  Parameters  |  Degrees of Freedom  |  Fit (-2lnL units)
       Model:             10                  1767             4055.9346
   Saturated:             NA                    NA                   NA
Independence:             NA                    NA                   NA
Number of observations/statistics: 920/1777

Information Criteria:
      |  df Penalty  |  Parameters Penalty  |  Sample-Size Adjusted
AIC:      521.93461              4075.9346                     NA
BIC:    -8002.73367              4124.1783               4092.4195
CFI: NA
TLI: 1   (also known as NNFI)
RMSEA:  0  [95% CI (NA, NA)]
Prob(RMSEA <= 0.05): NA
To get additional fit indices, see help(mxRefModels)
timestamp: 2020-03-01 16:55:41
Wall clock time: 0.095155001 secs
optimizer:  NPSOL
OpenMx version number: 2.17.2
Need help?  See help(mxSummary)

> fitGofs(fitSAT)
Mx:oneSATc  os=1777  ns=920   ep=10   co=0  df=1767  ll=4055.9346  cpu=0.0952  opt=NPSOL  ver=2.17.2  stc=0
```

Print Estimates

```
> summary(fitSAT)$parameters
free parameters:
   name    matrix  row  col    Estimate   Std.Error A lbound ubound
1   mMZ1 MZ.meanMZ    1 bmi1 21.34437690 0.036061832
2   mMZ2 MZ.meanMZ    1 bmi2 21.34901242 0.037650856
3   vMZ1  MZ.covMZ bmi1 bmi1  0.72766891 0.043658984    1e-04
4  cMZ21  MZ.covMZ bmi1 bmi2  0.59163768 0.040794161        0
5   vMZ2  MZ.covMZ bmi2 bmi2  0.79319915 0.047647906    1e-04
6   mDZ1 DZ.meanDZ    1 bmi1 21.44752035 0.047571928
7   mDZ2 DZ.meanDZ    1 bmi2 21.45784215 0.049233334
8   vDZ1  DZ.covDZ bmi1 bmi1  0.76919130 0.059007266    1e-04
9  cDZ21  DZ.covDZ bmi1 bmi2  0.24004049 0.045201541        0
10  vDZ2  DZ.covDZ bmi2 bmi2  0.82163163 0.063154677    1e-04

> fitEsts(fitSAT)
   mMZ1    mMZ2    vMZ1   cMZ21    vMZ2    mDZ1    mDZ2    vDZ1   cDZ21    vDZ2
21.3444 21.3490  0.7277  0.5916  0.7932 21.4475 21.4578  0.7692  0.2400  0.8216
```

Estimated Values

	Saturated Model					
	MZ twins			DZ twins		
		T1	**T2**		**T1**	**T2**
mean	**MZ**	21.34	21.35	**DZ**	21.45	21.46
		T1	**T2**		**T1**	**T2**
cov	**T1**	0.73		**T1**	0.77	
	T2	0.59	0.79	**T2**	0.24	0.82

10 parameters estimated:

mMZ1, mMZ2, vMZ1, vMZ2, cMZ21

mDZ1, mDZ2, vDZ1, vDZ2, cDZ21

Goodness-of-Fit Statistics

	os	ep	-2ll	df	AIC	diff -2ll	diff df	p
Saturated	1777	10	4055.93	1767	521.93			

os	observed statistics	
ep	estimated parameters	
-2ll	-2 LogLikelihood	
df	degrees of freedom	os - ep
AIC	Akaike's Information Criterion	-2ll -2df

Fitting Nested Models

oneSATc.R

```
# Constrain expected Means to be equal across twin order
modelEMO  <- mxModel(fitSAT, name="oneEMOc" )
modelEMO  <- omxSetParameters( modelEMO, label=c("mMZ1","mMZ2"), free=TRUE, values=svMe, newlabels='mMZ' )
modelEMO  <- omxSetParameters( modelEMO, label=c("mDZ1","mDZ2"), free=TRUE, values=svMe, newlabels='mDZ' )
fitEMO    <- mxRun( modelEMO, intervals=F )
fitGofs(fitEMO); fitEsts(fitEMO)

# Constrain expected Means and Variances to be equal across twin order
modelEMVO <- mxModel(fitEMO, name="oneEMVOc" )
modelEMVO <- omxSetParameters( modelEMVO, label=c("vMZ1","vMZ2"), free=TRUE, values=svVa, newlabels='vMZ' )
modelEMVO <- omxSetParameters( modelEMVO, label=c("vDZ1","vDZ2"), free=TRUE, values=svVa, newlabels='vDZ' )
fitEMVO   <- mxRun( modelEMVO, intervals=F )
fitGofs(fitEMVO); fitEsts(fitEMVO)

# Constrain expected Means and Variances to be equal across twin order and zygosity
modelEMVZ <- mxModel(fitEMVO, name="oneEMVZc" )
modelEMVZ <- omxSetParameters( modelEMVZ, label=c("mMZ","mDZ"), free=TRUE, values=svMe, newlabels='mZ' )
modelEMVZ <- omxSetParameters( modelEMVZ, label=c("vMZ","vDZ"), free=TRUE, values=svVa, newlabels='vZ' )
fitEMVZ   <- mxRun( modelEMVZ, intervals=F )
fitGofs(fitEMVZ); fitEsts(fitEMVZ)

# Print Comparative Fit Statistics
mxCompare( fitSAT, subs <- list(fitEMO, fitEMVO, fitEMVZ) )

# ------------------------------------------------------------------------------
sink()
save.image(paste(filename,".Ri",sep=""))
```

changing parameters

existing parameters

new parameters

generate likelihood ratio test

close .Ro file & save image as file with .Ri extension

Goodness-of-Fit Stats

	os	ep	-2ll	df	AIC	diff -2ll	diff df	p
Saturated	1777	10	4055.93	1767	521.93			
mT1=mT2	1777	8	4056.00	1769	518.00	0.07	2	0.97
mT1=mT2 varT1=varT2	1777	6	4058.94	1771	516.94	3.01	4	0.56
Zyg MZ=DZ	1777	4	4063.45	1773	517.45	7.52	6	0.28

diff -2ll	likelihood ratio Chi-square	
diff df	difference in degrees of freedom	
p	probability of Chi-square	

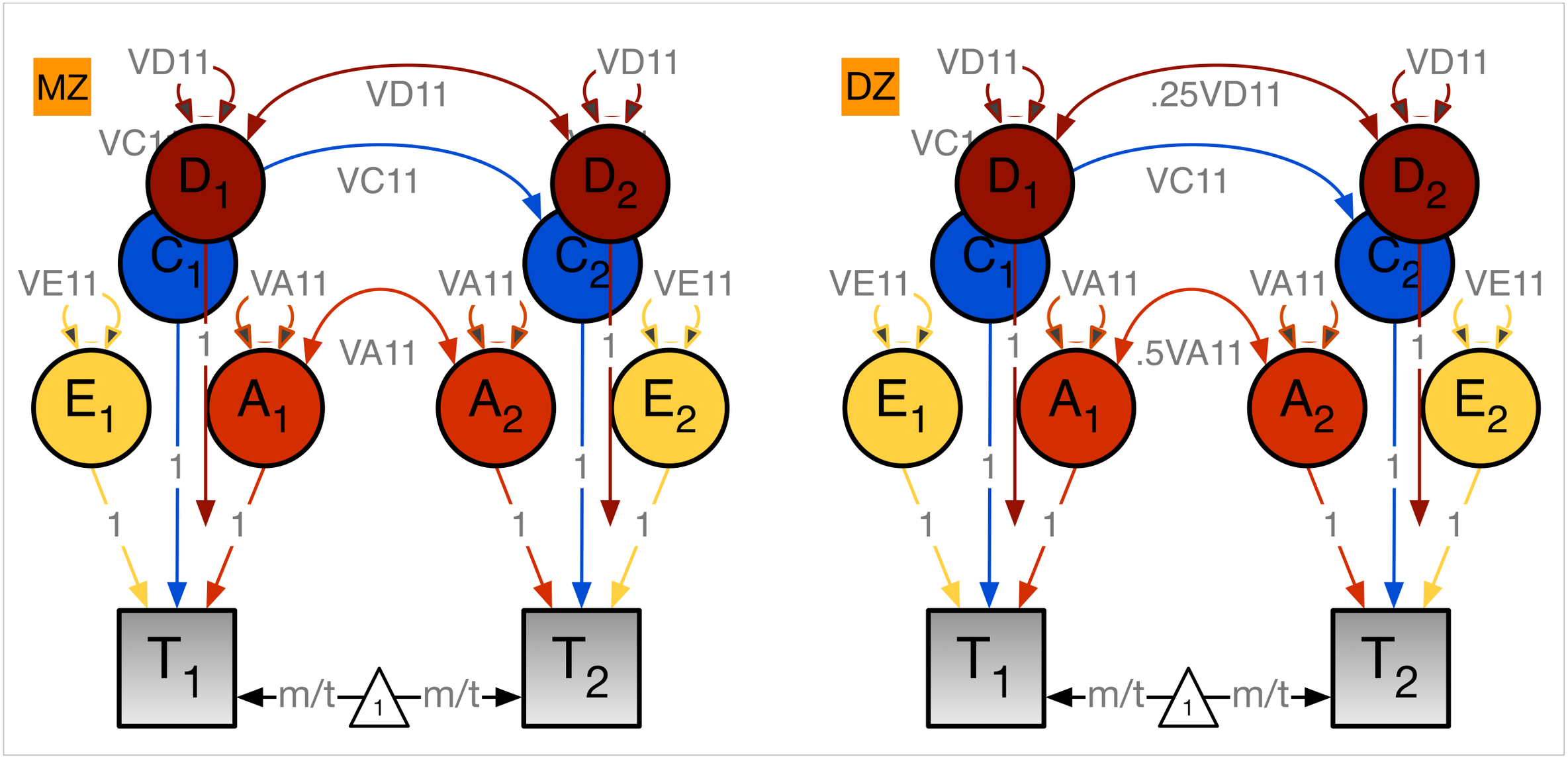
Conclusions so far

- BMI in young OZ females (age 18-30)
  - means of twin 1 and twin 2 not significantly different from one another in MZ & DZ pairs
  - variances of twin 1 and twin 2 not significantly different from one another in MZ & DZ pairs
  - means and variances of MZs and DZs not significantly different from one another
  - basic data assumptions about CTS met

Genetic Model(s)

ACE/ADE model

oneACEvc.R & oneADEvc.R



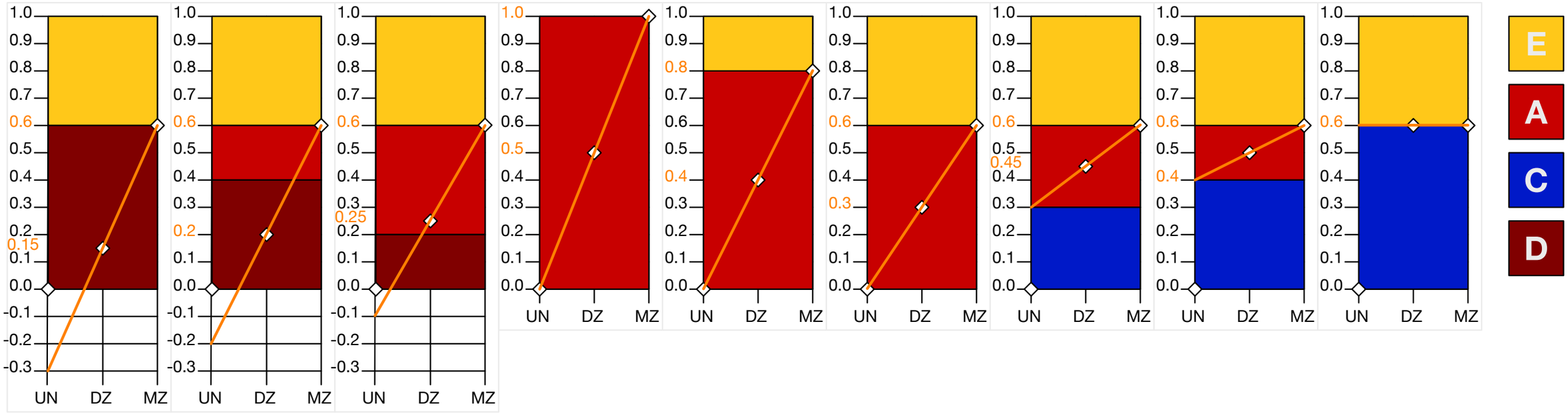
Estimating ACE Model

- Three unique observed statistics
  - V = VA+VC+VE
  - cMZ = VA+VC
  - cDZ = .5VA+VC
- Three unknown parameters
  - V – cMZ = (VA+VC+VE) – (VA+VC) = VE
  - cMZ – cDZ = (VA+VC) – (.5VA+VC) = .5VA
  - Thus VA = 2(cMZ – cDZ)
  - VC = –cMZ +2cDZ, or V – VA – VE = VC
- What if negative VC?
  - VD' = –2VC
  - VA' = cMZ +2VC = VA+3VC

Estimating ADE Model

- Three unique observed statistics
  - V = VA+VD+VE
  - cMZ = VA+VD
  - cDZ = .5VA+.25VD
- Three unknown parameters
  - V – cMZ = (VA+VD+VE) – (VA+VD) = VE
  - cMZ – 4*cDZ = (VA+VD) – (2VA+VD) = -VA
  - Thus VA = –(cMZ – 4*cDZ)
  - VD = 2cMZ – 4*cDZ, or V – VA – VE = VD
- What if negative VD?
  - VC' = –1/2VD
  - VA' = cMZ +1/2VD

From Twin Correlations to Sources of Variance



rDZ < 1/2 rMZ

E + A & D

D= -2*C

rMZ > rDZ

rMZ=2*rDZ

only A

1–rMZ

E + only A

rDZ > 1/2rMZ

rMZ = rDZ

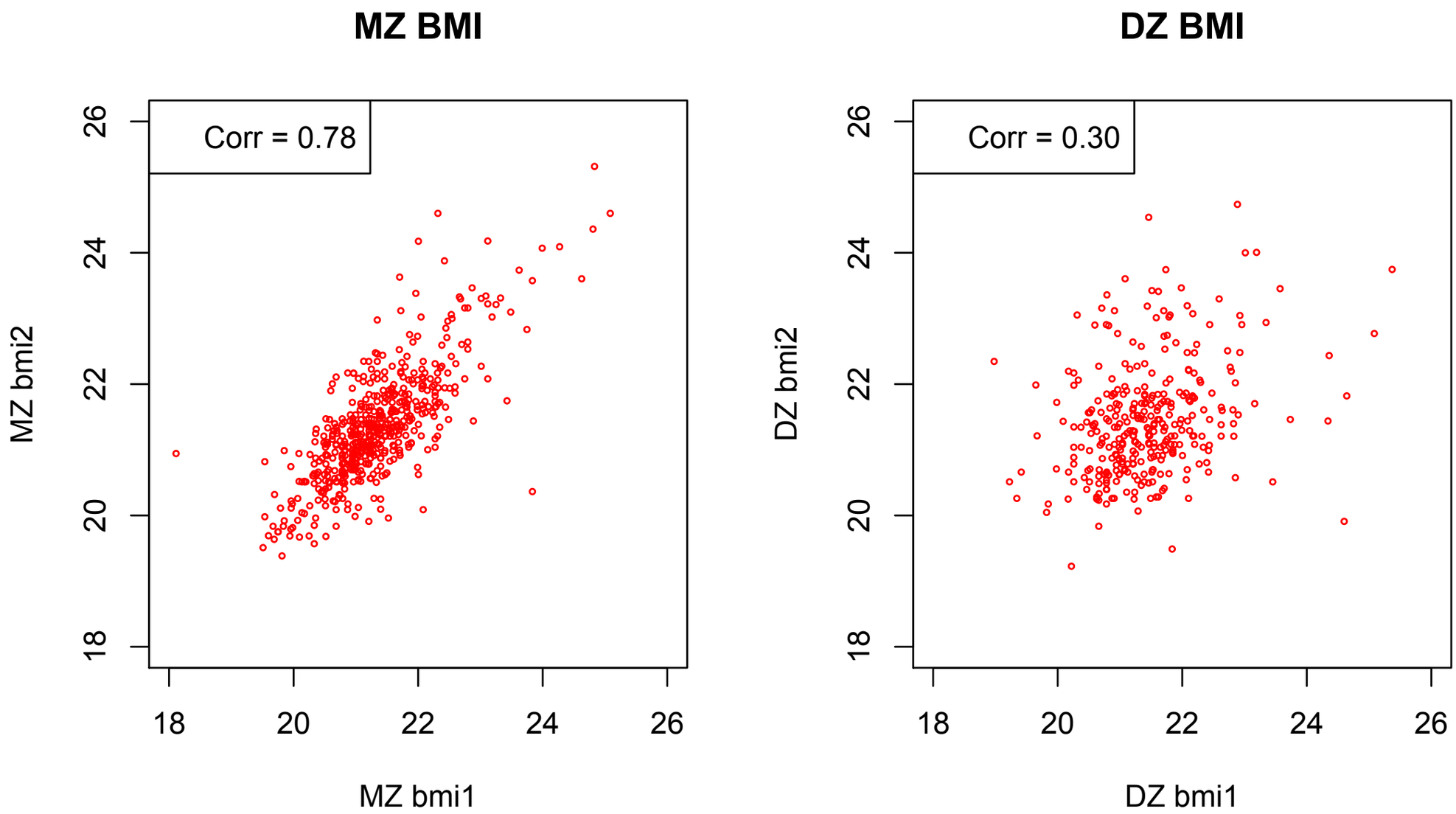
E + only C

E + A & C

C= -1/2*D

rMZ: monozygotic twin correlation; rDZ: dizygotic twin correlation;
A: additive genetic factors; E: unique environment: C: common environment; D: dominance

Example Twin Correlations



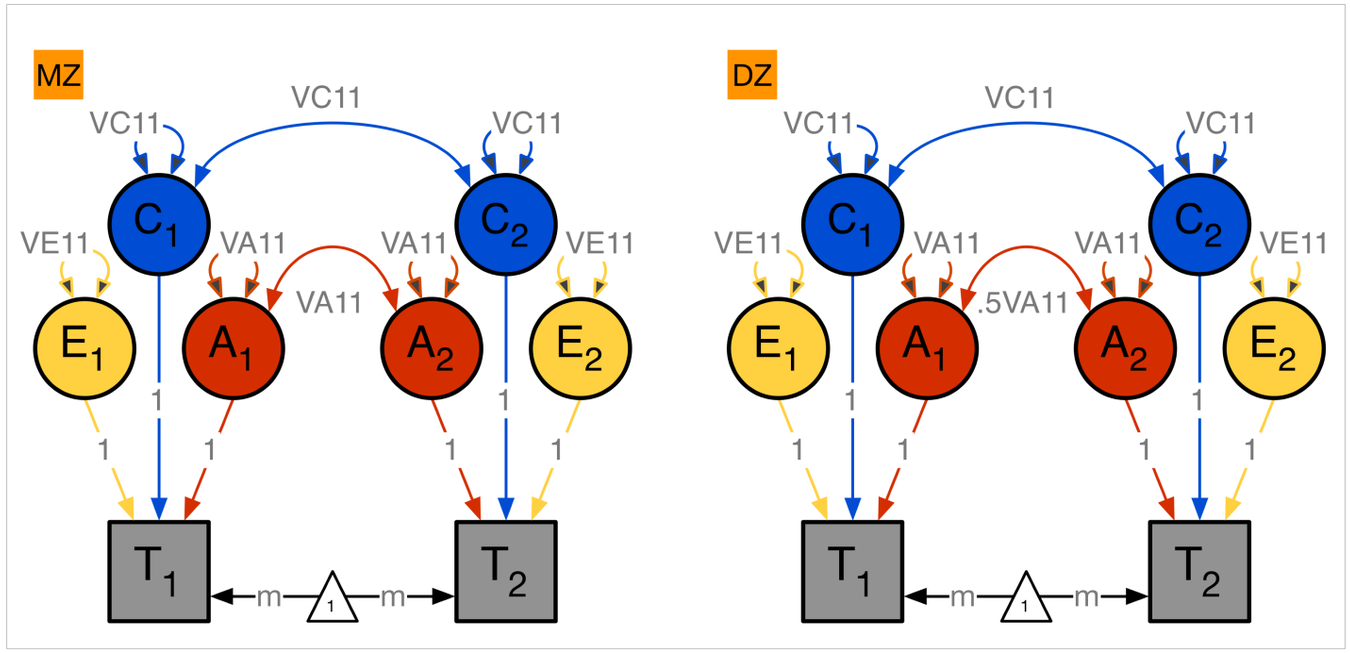
Roadmap for Univariate Analysis

- Use data to test basic assumptions (equal means & variances for twin 1/twin 2 and MZ/DZ pairs)
  - Saturated Model
- Estimate contributions of genetic/environmental effects on total variance of a phenotype
  - ACE or ADE Models
- Test ACE / ADE submodels to identify and report significant genetic and environmental contributions
  - AE / CE / E Only Models

Univariate ACE / ADE Model **variance estimation**

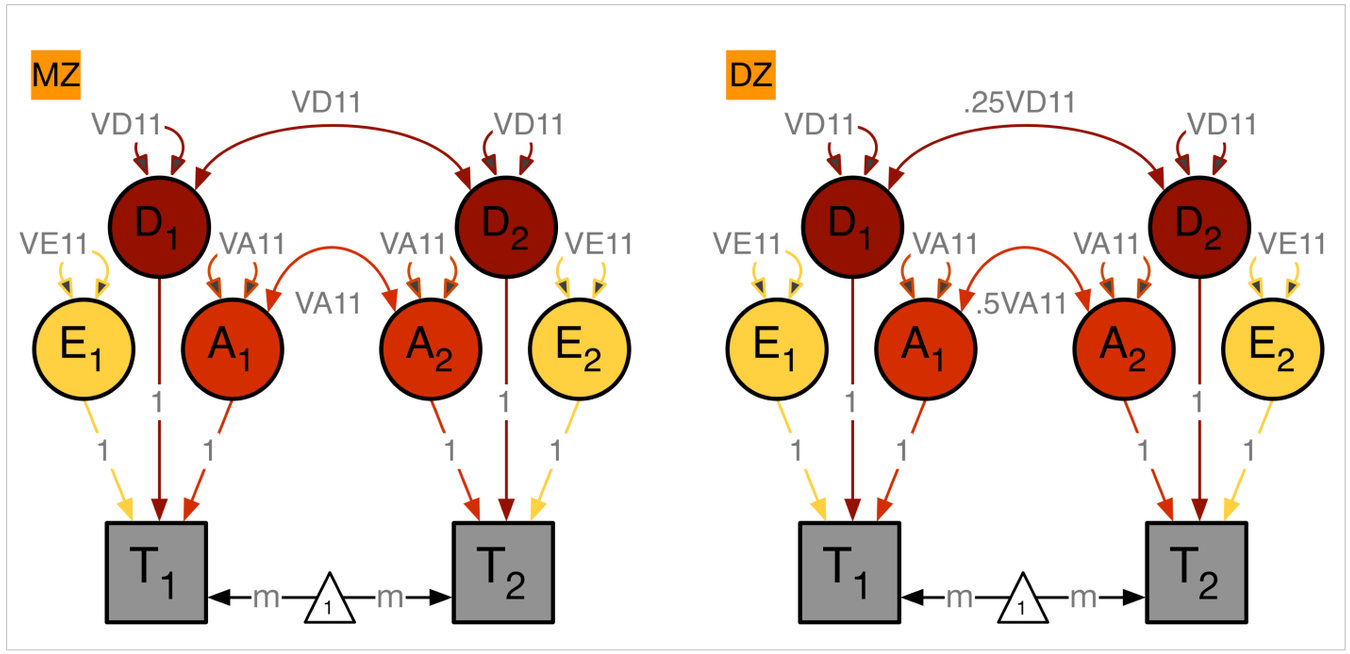
ACE model

oneACEvc.R

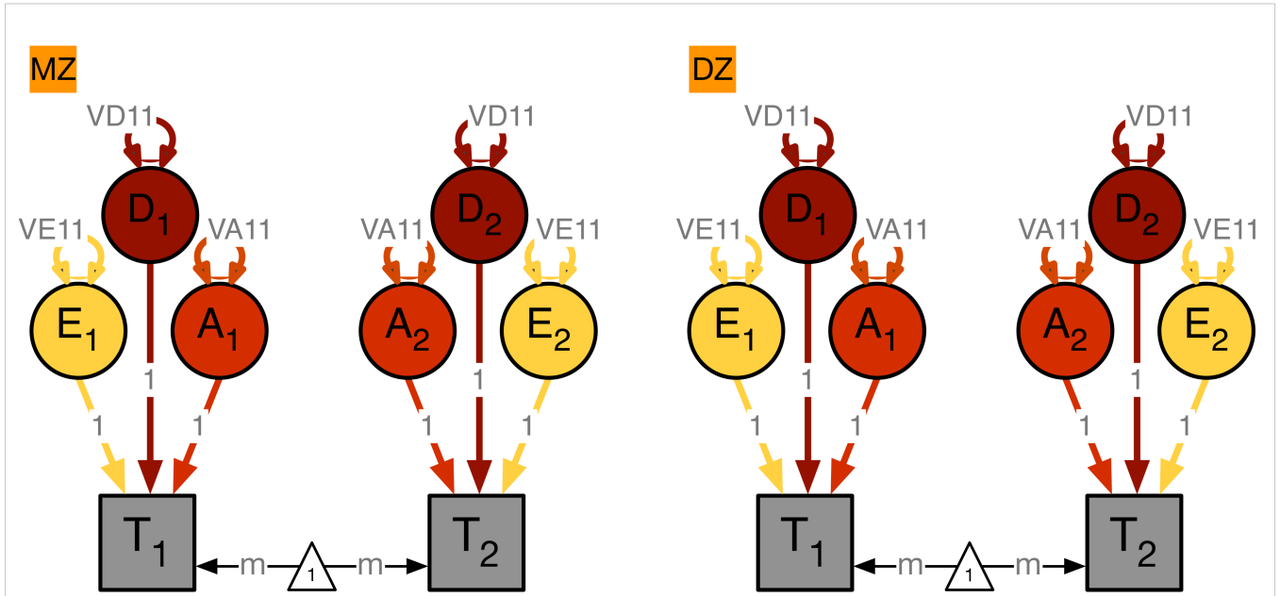


ADE model

oneADEvc.R



ADE Deconstructed: *Variance Components*




```
covA      <- mxMatrix( type="Symm", nrow=nv, ncol=nv,
 free=TRUE, values=svPa, label="VA11", name="VA" )

covD      <- mxMatrix( type="Symm", nrow=nv, ncol=nv,
 free=TRUE, values=svPa, label="VD11", name="VD" )

covE      <- mxMatrix( type="Symm", nrow=nv, ncol=nv,
 free=TRUE, values=svPe, label="VE11", name="VE" )
```

VA_{11}

VA *1x1*

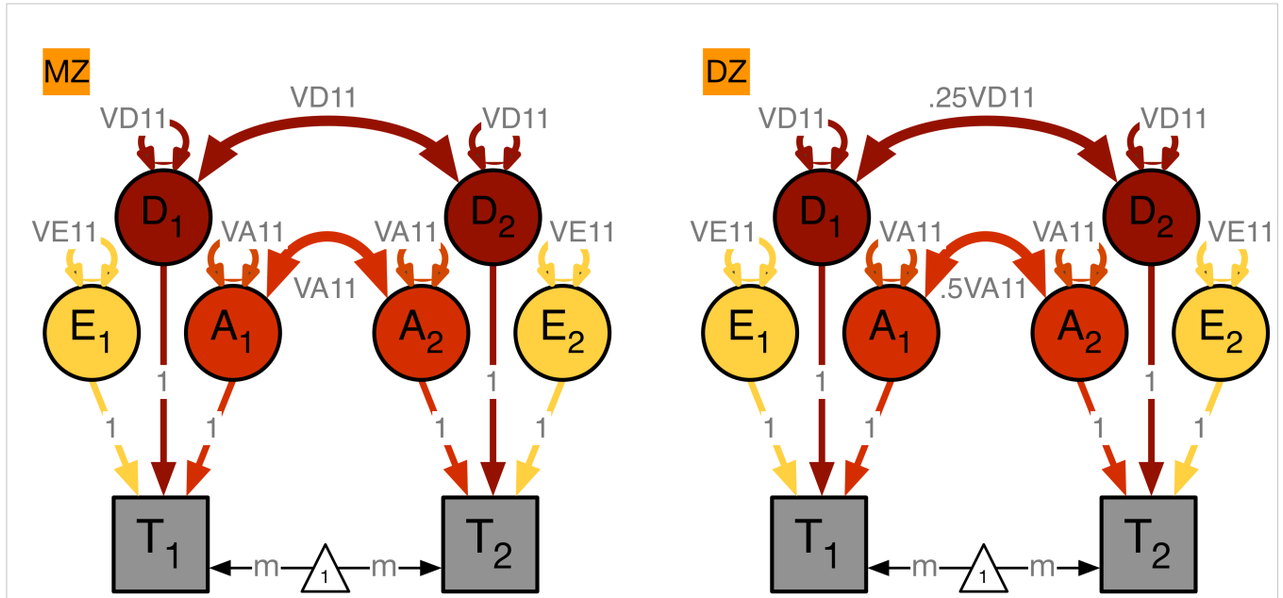
VD_{11}

VD *1x1*

VE_{11}

VE *1x1*

ADE Deconstructed: *Variances + Covariances*




```
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" )
```

V	VA+VD+VE

V *1x1*

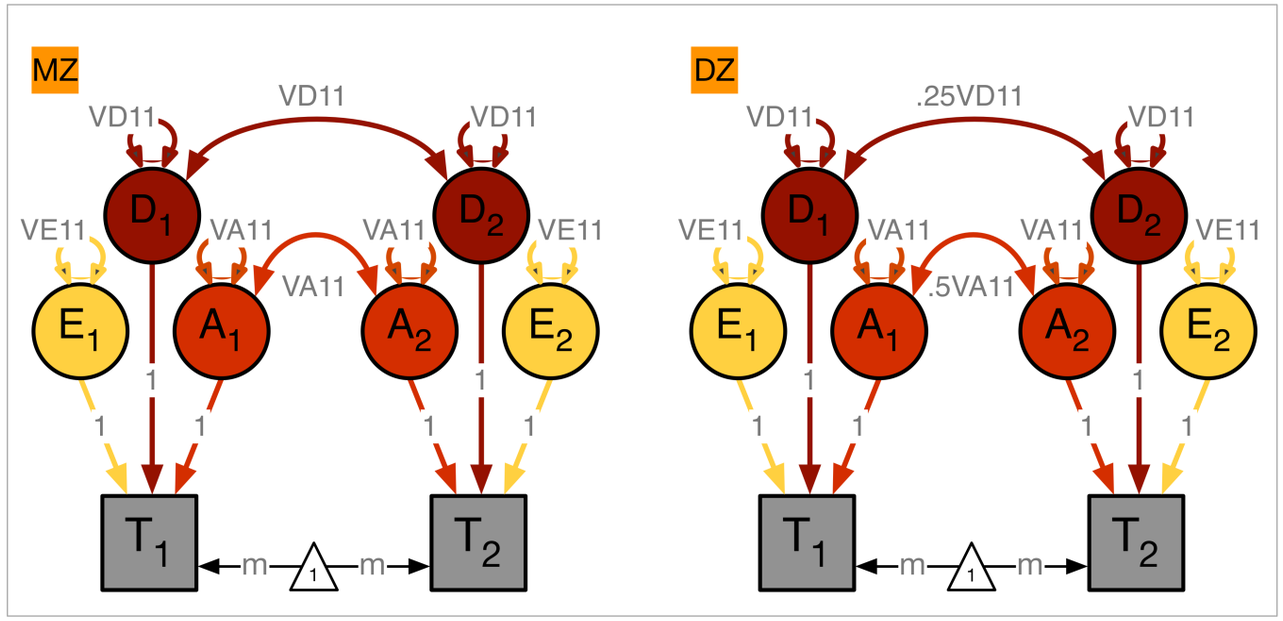
cMZ	VA+VD

cMZ *1x1*

cDZ	.5VA+.25VD

cDZ *1x1*

ADE Deconstructed: *Covariance Matrices & Means*




```
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" )



meanG     <- mxMatrix( type="Full", nrow=1, ncol=ntv,
 free=TRUE, values=svMe, labels=labVars("mean",vars),
 name="meanG" )
```

V	cMZ
cMZ	V

expCovMZ *2x2*

V	cDZ
cDZ	V

expCovDZ *2x2*

x_1	x_1

meanG *1x2*

Model Specification

oneADEvc.R

```
# ------------------------------------------------------------------------------
# PREPARE MODEL

# 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=svPa, label="VA11", name="VA" )
covD      <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svPa, label="VD11", name="VD" )
covE      <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svPe, label="VE11", name="VE" )

# 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" )
```

variance components: VA, VD & VE

Model Specification 2

oneADEvc.R

```
# 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 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
ciADE     <- mxCI( "US[1,1:3]" )

# Build Model with Confidence Intervals
modelADE  <- mxModel( "oneADEvc", pars, modelMZ, modelDZ, multi, estUS, ciADE )
```

list of common elements

calculate standardized variance components

list of matrix elements to calculate confidence intervals (CI)

ADE model object

Run Model

oneADEvc.R

```
# ---------------------------------------------------------------------------
# RUN MODEL

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

# Compare with Saturated Model
#if saturated model fitted in same session
mxCompare( fit, fitADE )
#if saturated model prior to genetic model
#lrtSAT( fitADE,4055.9346,1767)
```

estimate CI's

function in miFunctions.R to provide -2ll & df of previously fit model

```
# Print Goodness-of-fit Statistics & Parameter Estimates
fitGofs(fitADE)
fitEstCis(fitADE)

round(fitADE$US$result,4)
```

print estimates of variance components

summary(fitADE)

```
free parameters:
     name matrix row col    Estimate   Std.Error A
1 meanbmi  meanG   1   1 21.39464927 0.025973494
2    VA11     VA   1   1  0.32092995 0.150909584
3    VD11     VD   1   1  0.28942518 0.147886812
4    VE11     VE   1   1  0.16935016 0.010363413

confidence intervals:
                       lbound   estimate     ubound note
oneADEvc.US[1,1] 0.016290870 0.32092995 0.61208265
oneADEvc.US[1,2] 0.011924028 0.28942518 0.59556124
oneADEvc.US[1,3] 0.150553156 0.16935016 0.19139089

Model Statistics:
               |  Parameters  |  Degrees of Freedom  |  Fit (-2lnL units)
       Model:             4                  1773             4063.4496
   Saturated:            NA                    NA                   NA
Independence:            NA                    NA                   NA
Number of observations/statistics: 920/1777

Information Criteria:
      |  df Penalty  |  Parameters Penalty  |  Sample-Size Adjusted
AIC:       517.44962             4071.4496                     NA
BIC:    -8036.16490             4090.7471              4078.0436
```

miFunctions: fitGofs & fitEsts

```
> fitGofs(fitADE)
Mx:oneADEvc  os=1777  ns=920   ep=4   co=0  df=1773  ll=4063.4496  cpu=0.1513  opt=NPSOL  ver=2.17.2  stc=0
>
> fitEstCis(fitADE)
meanbmi    VA11    VD11    VE11
21.3946  0.3209  0.2894  0.1694
                lbound estimate ubound
oneADEvc.US[1,1] 0.0163   0.3209 0.6121
oneADEvc.US[1,2] 0.0119   0.2894 0.5956
oneADEvc.US[1,3] 0.1506   0.1694 0.1914
```

```
> round(fitADE$US$result,4)

       VA     VD     VE     SA     SD     SE
US 0.3209 0.2894 0.1694 0.4116 0.3712 0.2172
```

Goodness-of-Fit Stats & Estimates

	os	ep	-2ll	df	AIC	diff -2ll	diff df	p
Saturated	1777	10	4055.93	1767	521.93			
ADE	1777	4	4063.45	1773	517.45	7.51	6	0.27

	unstandardized variance components			standardized variance components		
	VA	VD	VE	SA	SD	SE
ADE	0.32	0.29	0.17	0.41	0.37	0.22

Roadmap for Univariate Analysis

- Use data to test basic assumptions (equal means & variances for twin 1/twin 2 and MZ/DZ pairs)
  - Saturated Model
- Estimate contributions of genetic/environmental effects on total variance of a phenotype
  - ACE or ADE Models
- Test ACE / ADE submodels to identify and report significant genetic and environmental contributions
  - AE / CE / E Only Models

Fitting Nested Models

oneADEvc.R

```
# ---------------------------------------------------------------------------
# RUN SUBMODELS

# Run AE model
modelAE   <- mxModel( fitADE, name="oneAEc" )
modelAE   <- omxSetParameters( modelAE, labels="VD11", free=FALSE, values=0 )
fitAE     <- mxRun( modelAE, intervals=T )
fitGofs(fitAE); fitEsts(fitAE)

# Run E model
modelE    <- mxModel( fitAE, name="oneEc" )
modelE    <- omxSetParameters( modelE, labels="VA11", free=FALSE, values=0 )
fitE      <- mxRun( modelE, intervals=T )
fitGofs(fitE); fitEsts(fitE)

# Print Comparative Fit Statistics
mxCompare( fitADE, nested <- list(fitAE, fitE) )
round(rbind(fitADE$US$result,fitAE$US$result,fitE$US$resulT ),4)
```

dropping parameters

Nested Models

- ‘Full’ ADE Model
- Nested Models
  - AE Model vs ADE Model: test significance of **D**
  - E Model vs AE Model: test significance of **A**
  - E Model vs ADE Model: test significance simultaneously of both **A & D**

Goodness-of-Fit Statistics **variance estimation**

	os	ep	-2ll	df	AIC	diff -2ll	diff df	p
Saturated	1777	10	4055.93	1767	521.93			
ADE	1777	4	4063.45	1773	517.45	7.51	6	0.27
AE	1777	3	4067.66	1774	519.66	4.21	1	0.04
E	1777	2	4591.79	1775	1041.79	528.34	2	0.00

Under the null hypothesis, test is distributed as a chi-square with 1df

Estimated Values **variance estimation**

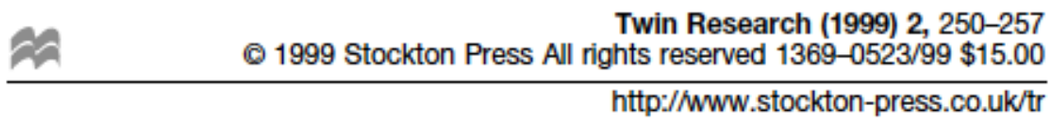
	unstandardized variance components			standardized variance components		
	VA	VD	VE	SA	SD	SE
ADE	0.32 0.02-0.61	0.29 0.01-0.60	0.17 0.15-0.19	0.41	0.37	0.22
AE	0.62 0.56-0.68	-	0.17 0.17-0.19	0.78	-	0.22
E	-	-	0.78 0.73-0.83	-	-	1.00

Conclusions

- BMI in young OZ females (age 18-30)
  - **additive** genetic factors: highly significant
  - **dominance**: borderline significant
  - **specific environment**al factors: significant
  - **shared environment**al factors: not

Publications

- Eaves LJ: Inferring the causes of human variation. J. R. Stat. Soc. Ser. A 140, 324–355,1977.
- Neale MC, Cardon LR: Methodology for Genetic Studies of Twins and Families (NATO ASI Series), Dordrecht, The Netherlands: Kluwer Academic Publishers, 496p, 1992.
- Posthuma P, Beem AL, de Geus EJC, van Baal GCM, von Hjelmborg JB, Iachine I, Boomsma DI: Theory and Practice in Quantitative Genetics. Twin Research 6:361-376, 2003.
- Eaves LJ, Chen S, Neale M, Maes HH, Silberg J: Questions, Models and Methods in Psychiatric Genetics, in Psychiatric Genetics (Review of Psychiatry Vol 24), Kendler KS & Eaves LJ (Eds). Washington, DC: American Psychiatric Publishing, Inc., 2005.
- Maes HH: The ACE model, in Encyclopedia for Behavioral Statistics (Wiley Series in Probability and Statistics), Purcell S (Volume Editor). John Wiley & Sons, Inc., 2005.
- Neale MC: Biometrical Models in Behavioral Genetics, in Handbook of Behavior Genetics, Yong-Kyu, K. (Volume Editor). Springer, 2009.
- Evans DM, Frazer IH, Martin NG: Genetic and environmental causes of variation in basal levels of blood cells, Twin Research 2: 250-257, 1999.



Twin Research (1999) 2, 250–257

http://www.stockton-press.co.uk/tr

Genetic and environmental causes of variation in basal levels of blood cells

David M Evans[1], Ian H Frazer[2] and Nicholas G Martin[1]

Calculate Correlations

- Add calculations to OpenMx scripts
  - as part of script (calculated with every iteration)

```
# Create Algebra for Maximum Likelihood Estimates of Twin Correlations

corMZ      <- mxAlgebra( cov2cor(covMZ), name="corMZ" )
corDZ      <- mxAlgebra( cov2cor(covDZ), name="corDZ" )
```

→ function to calculate correlations from covariances

  - after script has been run

```
# Create Algebra for Maximum Likelihood Estimates of Twin Correlations

corMZ      <- mxEval( cov2cor(covMZ), fitSAT$MZ )
corDZ      <- mxEval( cov2cor(covDZ), fitSAT$DZ )
```

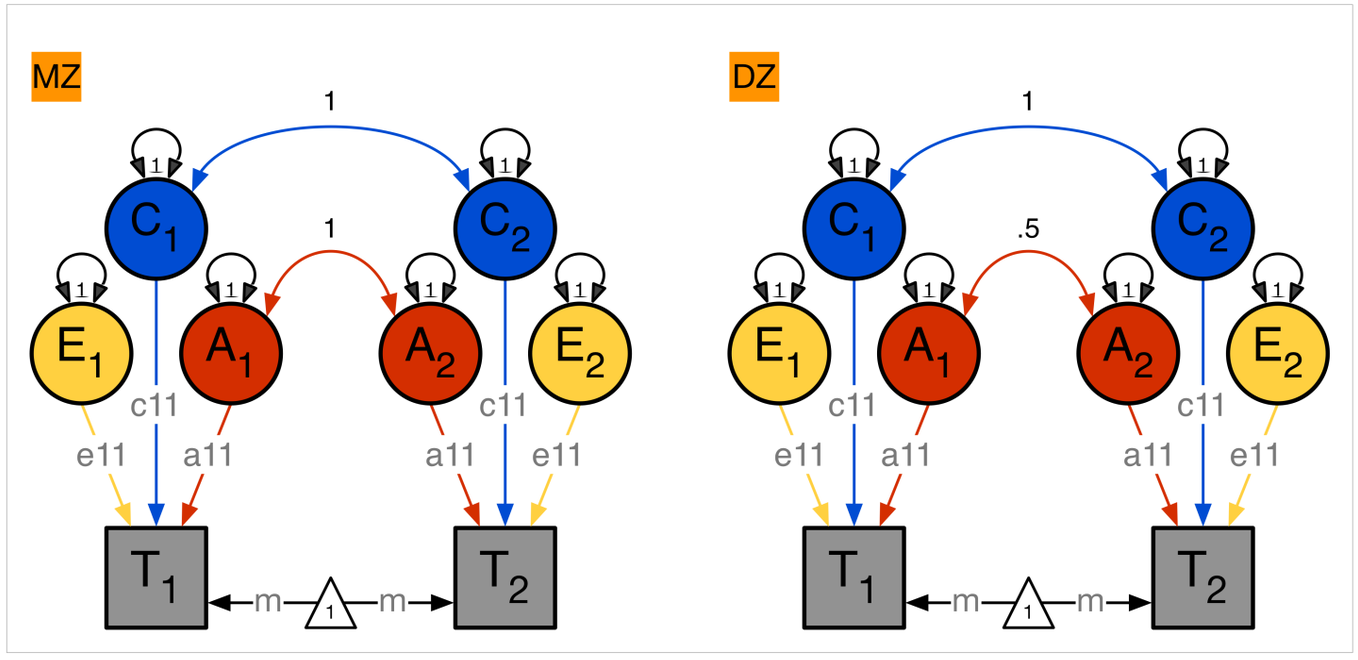
Alternative Approaches to Fitting Genetic Models

- Direct Variance Estimation
  - Allows variances to be estimated as negative
  - Makes correct inferences when comparing alternative models
- Path Estimation
  - Bounds variances to be positive
  - May make incorrect inferences when comparing alternative models

Univariate ACE / ADE Model - **path estimation**

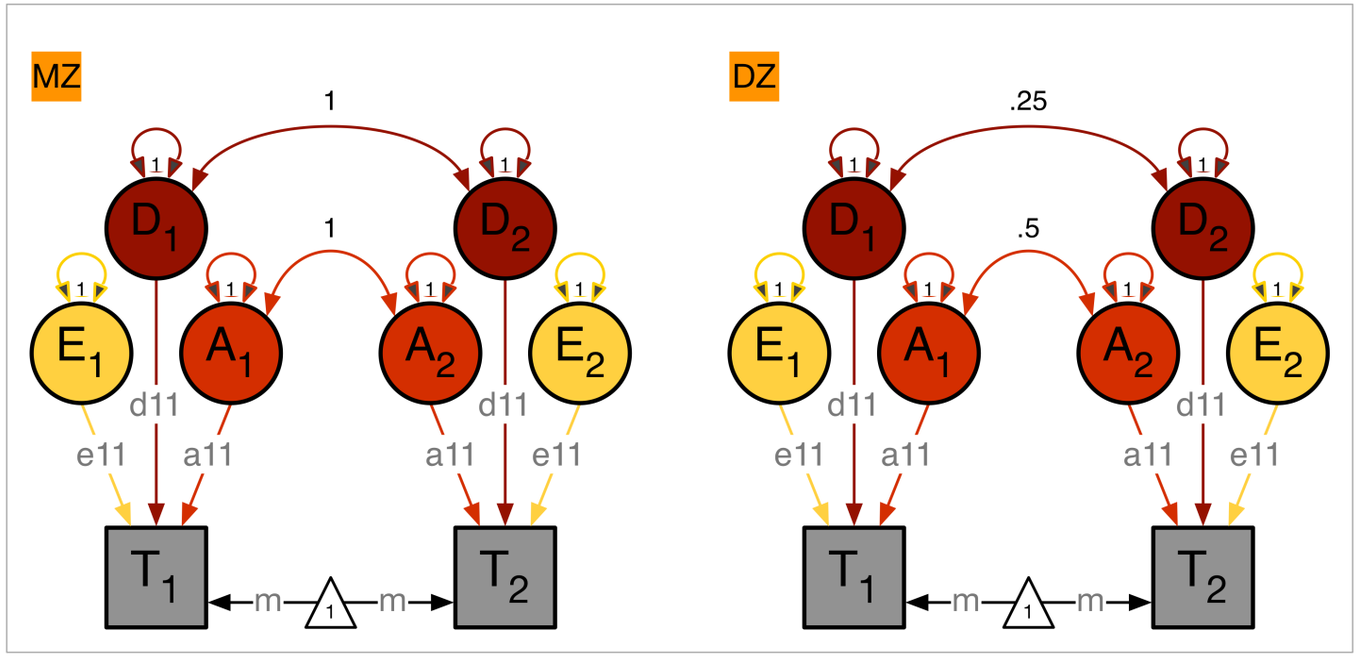
ACE model

oneACEc.R

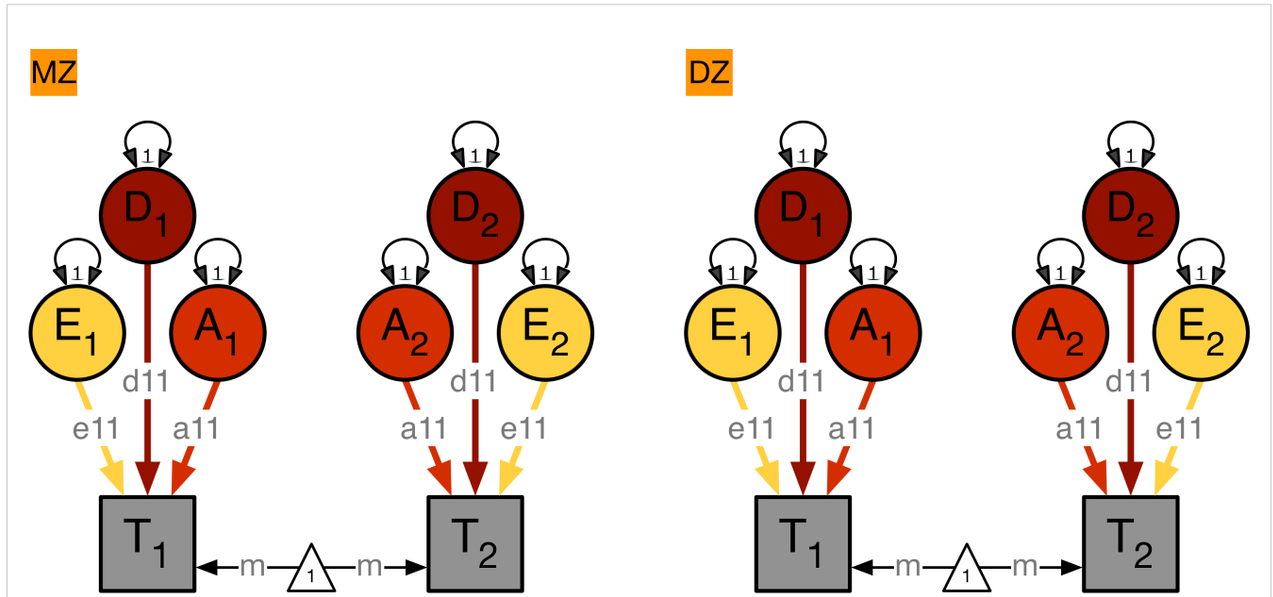


ADE model

oneADEc.R



ADE Deconstructed: *Path Coefficients*




```
pathA     <- mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE,
values=svPa, label="a11", lbound=lbPa, name="a" )

pathD     <- mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE,
values=svPa, label="d11", lbound=lbPa, name="d" )

pathE     <- mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE,
values=svPe, label="e11", lbound=lbPa, name="e" )
```

a_{11}

a *1x1*

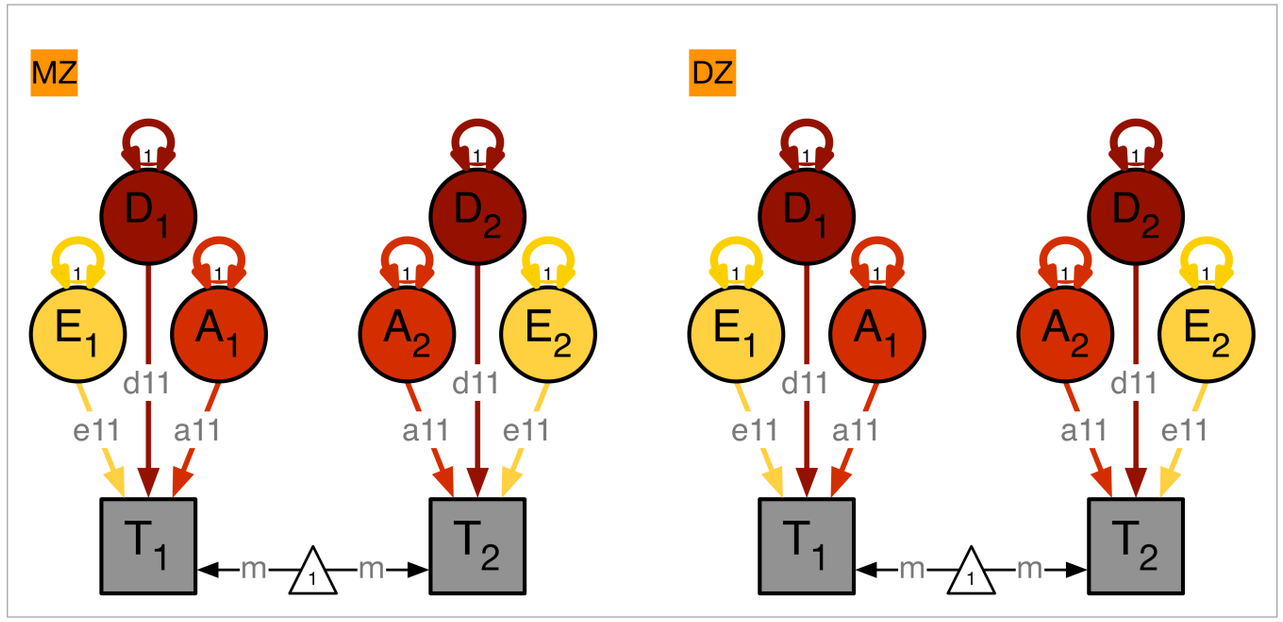
d_{11}

d *1x1*

e_{11}

e *1x1*

ADE Deconstructed: *Variance Components*




```
covA       <- mxAlgebra( expression=a %*% t(a),
 name="A" )

covC       <- mxAlgebra( expression=d %*% t(d),
 name="D" )

covE       <- mxAlgebra( expression=e %*% t(e),
 name="E" )
```

$a_{11} * t(a_{11})$

A *1x1*

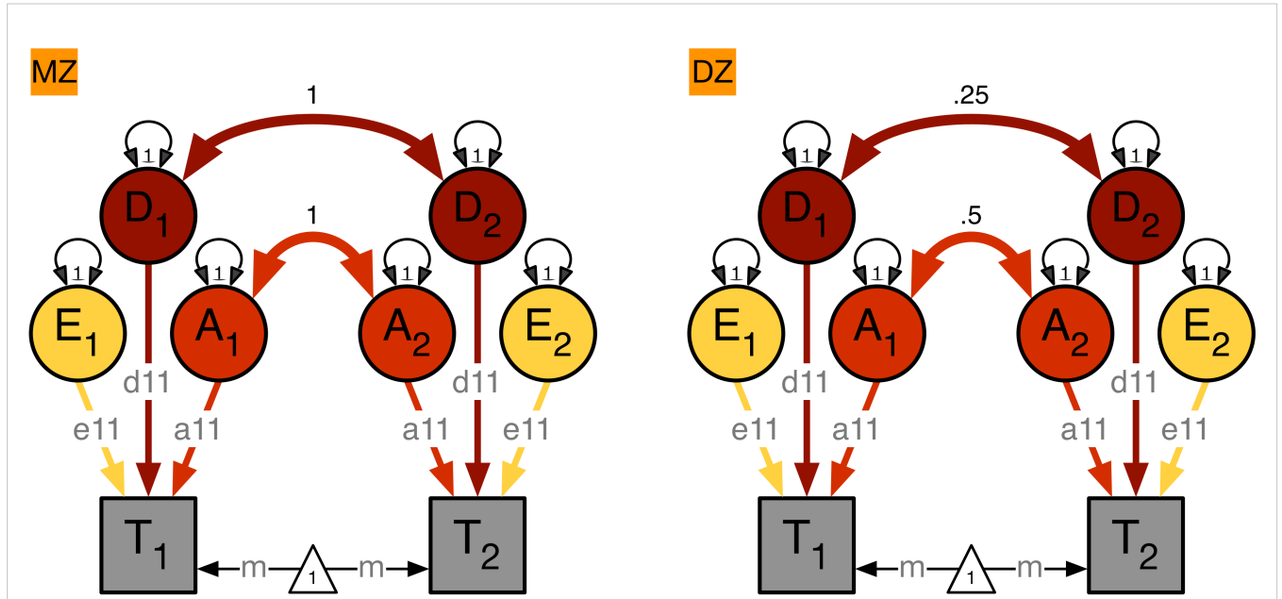
$d_{11} * t(d_{11})$

D *1x1*

$e_{11} * t(e_{11})$

E *1x1*

ADE Deconstructed: *Variances + Covariances*




```
covP       <- mxAlgebra( expression= A+D+E,
 name="V" )

covMZ      <- mxAlgebra( expression= A+D,
name="cMZ" )

covDZ      <- mxAlgebra( expression= 0.5%x%A+ 0.25%x%D,
name="cDZ" )
```

V	A+D+E

V *1x1*

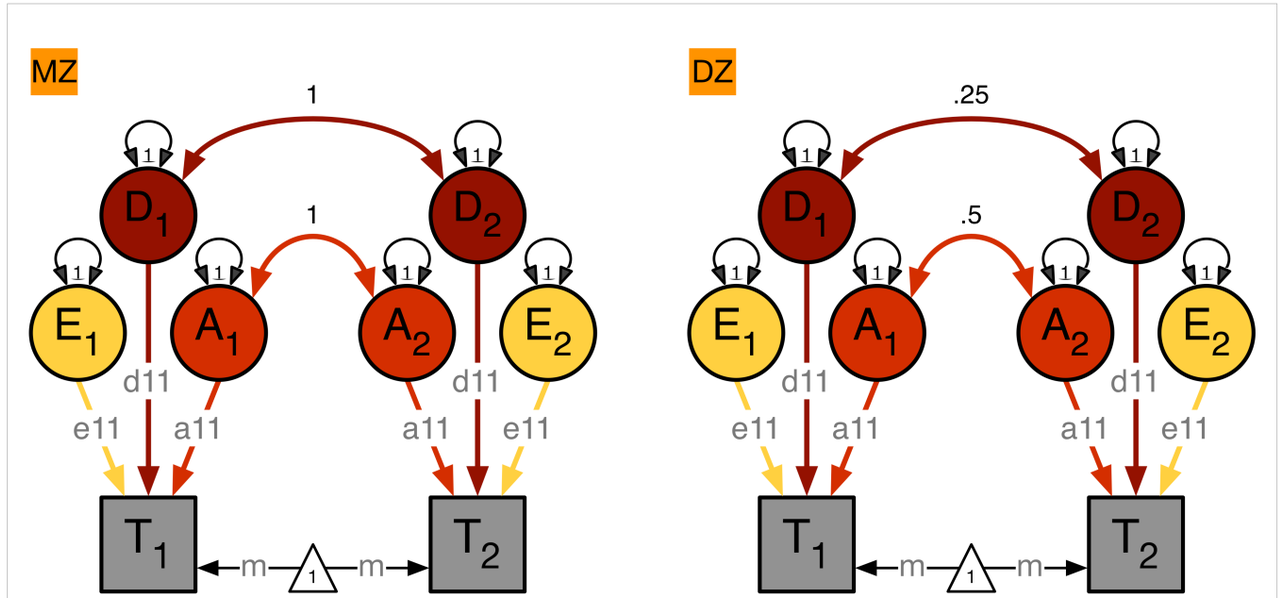
cMZ	A+D

cMZ *1x1*

cDZ	.5A+.25D

cDZ *1x1*

ADE Deconstructed: *Covariance Matrices & Means*




```
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" )



meanG     <- mxMatrix( type="Full", nrow=1, ncol=ntv,
 free=TRUE, values=svMe, labels=labVars("mean",vars),
 name="meanG" )
```

V	cMZ
cMZ	V

expCovMZ *2x2*

V	cDZ
cDZ	V

expCovDZ *2x2*

x_1	x_1

meanG *1x2*

Model Specification

oneADEc.R

```
# ------------------------------------------------------------------------------
# PREPARE MODEL

# ADE Model
# Create Algebra for expected Mean Matrices
meanG     <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe, labels="x1", name="meanG" )

# Create Matrices for Path Coefficients
pathA     <- mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE, values=svPa, label="a11", lbound=lbPa, name="a" )
pathD     <- mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE, values=svPa, label="d11", lbound=lbPa, name="d" )
pathE     <- mxMatrix( type="Lower", nrow=nv, ncol=nv, free=TRUE, values=svPe, label="e11", lbound=lbPa, name="e" )

# Create Algebra for Variance Components
covA      <- mxAlgebra( expression=a %*% t(a), name="A" )
covD      <- mxAlgebra( expression=d %*% t(d), name="D" )
covE      <- mxAlgebra( expression=e %*% t(e), name="E" )

# Create Algebra for expected Variance/Covariance Matrices in MZ & DZ twins
covP      <- mxAlgebra( expression= A+D+E, name="V" )
covMZ     <- mxAlgebra( expression= A+D, name="cMZ" )
covDZ     <- mxAlgebra( expression= 0.5%x%A+ 0.25%x%D, 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" )
```

path matrices: a, d & e

variance components: a^2, d^2 & e^2

Model Specification 2

oneADEc.R

```
# 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, pathA, pathD, pathE, 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 Variance Components
rowUS     <- rep('US',nv)
colUS     <- rep(c('A','D','E','SA','SD','SE'),each=nv)
estUS     <- mxAlgebra( expression=cbind(A,D,E,A/V,D/V,E/V), name="US", dimnames=list(rowUS,colUS))

# Create Confidence Interval Objects
ciADE     <- mxCI( "US[1,1:3]" )

# Build Model with Confidence Intervals
modelADE  <- mxModel( "oneADEc", pars, modelMZ, modelDZ, multi, estUS, ciADE )
```

list of common elements

ADE variance components

Fitting Nested Models

oneADEc.R

```
# ------------------------------------------------------------------------------
# RUN SUBMODELS

# Run AE model
modelAE   <- mxModel( fitADE, name="oneAEc" )
modelAE   <- omxSetParameters( modelAE, labels="d11", free=FALSE, values=0 )
fitAE     <- mxRun( modelAE, intervals=T )
mxCompare( fitADE, fitAE )
fitGofs(fitAE)
fitEsts(fitAE)

# Run E model
modelE    <- mxModel( fitAE, name="oneEc" )
modelE    <- omxSetParameters( modelE, labels="a11", free=FALSE, values=0 )
fitE      <- mxRun( modelE, intervals=T )
mxCompare( fitAE, fitE )
fitGofs(fitE)
fitEsts(fitE)

# Print Comparative Fit Statistics
mxCompare( fitADE, nested <- list(fitAE, fitE) )
round(rbind(fitADE$US$result,fitAE$US$result,fitE$US$resulT ),4)

# ------------------------------------------------------------------------------
sink()
save.image(paste(filename,".Ri",sep=""))
```

dropping path parameters

Goodness-of-Fit Stats **ADE** **variance estimation**

	os	ep	-2ll	df	AIC	diff -2ll	diff df	p	p/2
Saturated	1777	10	4055.93	1767	521.93				
ADE	1777	4	4063.45	1773	517.45	7.51	6	0.27	
AE	1777	3	4067.66	1774	519.66	4.21	1	0.04	
E	1777	2	4591.79	1775	1041.79	528.34	2	0.00	
path ADE	1777	4	4063.45	1773	517.45	7.51	6	0.27	0.13
path AE	1777	3	4067.66	1774	519.66	4.21	1	0.04	0.02
path E	1777	2	4591.79	1775	1041.79	528.34	2	0.00	0.00

path estimation

Should be divided by 2, as ADE parameters are bounded to be positive
Under the null hypothesis, test is distributed 50:50 as mixture of 0 and a chi-square with 1df

Estimated Values **ADE** **variance estimation**

	path coefficients			unstandardized variance components			standardized variance components		
	a	d	e	VA	VD	VE	SA	SD	SE
ADE				0.32 .02-.61	0.29 .01-.60	0.17 .15-.19	0.41	0.37	0.22
AE				0.62	-	0.17	0.78	-	0.22
E				-	-	0.78	-	-	1.00
path ADE	0.57	0.44	0.41	0.32 .02-.61	0.29 .01-.60	0.17 .15-.19	0.41	0.37	0.22
path AE	0.77	-	0.41	0.62	-	0.17	0.78	-	0.22
path E	-	-	0.87	-	-	0.79	-	-	1.00

path estimation

Goodness-of-Fit Stats **ACE** **variance estimation**

	os	ep	-2ll	df	AIC	diff -2ll	diff df	p	p/2
Saturated	1777	10	4055.93	1767	521.93				
ACE	1777	4	4063.45	1773	517.45	7.51	6	0.27	
AE	1777	3	4067.66	1774	519.66	4.21	1	0.04	
CE	1777	3	4220.31	1774	672.31	156.86	1	0.00	
E	1777	2	4591.79	1775	1041.79	528.34	2	0.00	
path ACE	1777	4	4067.66	1773	519.66	4.21	6	0.27	0.13
path AE	1777	3	4067.66	1774	519.66	0	1	0.04	0.02
path CE	1777	3	4220.31	1774	672.31	152.65	1	0.00	0.00
path E	1777	2	4591.79	1775	1041.79	524.13	2	0.00	0.00

path estimation

Estimated Values **ACE** **variance estimation**

	path coefficients			unstandardized variance components			standardized variance components		
	a	c	e	VA	VC	VE	SA	SC	SE
ACE				0.75	-0.14	0.17	0.97	-0.19	0.22
AE				0.62	-	0.17	0.78	-	0.22
CE				-	0.46	0.32	-	0.59	0.41
E				-	-	0.78	-	-	1.00
path ACE	0.79	0.00	0.41	0.62	0.00	0.17	0.41	0.37	0.22
path AE	0.77	-	0.41	0.62	-	0.17	0.78	-	0.22
path CE	-	0.68	0.56	-	0.46	0.32	-	0.59	0.41
path E	-	-	0.87	-	-	0.79	-	-	1.00

path estimation

Estimated Values **ADE|ACE** **variance estimation**

	path coefficients			unstandardized variance components			standardized variance components		
	a	d	e	VA	VD	VE	SA	SD	SE
ADE				0.32 .02-.61	0.29 .01-.60	0.17 .15-.19	0.41	0.37	0.22
AE				0.62	-	0.17	0.78	-	0.22
E				-	-	0.78	-	-	1.00
	a	c	e	VA	VC	VE	SA	SC	SE
ACE				0.75 .60-.92	-0.14 -.3 -.01	0.17 .15-.19	0.97	-0.19	0.22
AE				0.62	-	0.17	0.78	-	0.22
CE				-	0.46	0.32	-	0.59	0.41
E				-	-	0.78	-	-	1.00

A'= A’+3C'=.33

D'=-2C'=.28

Publications 2

- Hao Wu, Michael C Neale: On the Likelihood Ratio Tests in Bivariate ACDE Models. Psychometrika 78 (3), 441-63 Jul 2013.
- Brad Verhulst, Elizabeth Prom-Wormley, Matthew Keller, Sarah Medland, Michael C Neale: Type I Error Rates and Parameter Bias in Multivariate Behavioral Genetic Models. Behav Genet 49 (1), 99-111 Jan 2019.

Thank you !

- Functions to run saturated / ADE / ACE models
- **umx**
- Tim Bates

Univariate Twin Modeling 7 ways

Hermine Maes, Elizabeth Prom-Wormley

Sarah Medland, Lucía Colodro Conde, Jose Morosoli Garcia

2021

Previous Videos (Michael Neale, Conor Dolan)

- Intro to Structural Equation Modeling (27 min)
  - Measuring variation
  - Structural equation modeling
  - Path analysis, RAM algebra
  - OpenMx & other SEM software
- ACE Model & Likelihood (29 min)
  - Derive ACE model expectations from twin data
  - Twin correlations -> Variance components
  - Likelihood & Maximum Likelihood model fitting
- Genetically Informative Designs (? min)
  - Linear regression -> Covariance structure
  - Genetic covariance structure analysis with twin data
  - Extension to multivariate case

Rationale

- Understanding causes of variation in phenotype of interest & partitioning variation in genetic and environmental variance components
  - Genetically informative designs -> twin heritability
    - infer genetic and environmental contributions to variance from phenotypic covariances (correlations) among family members, often twins, using expected relatedness
  - Genomically informative designs -> SNP heritability
    - infer genetic and environmental contributions to variance from genetic relationship matrices (correlations) based on measured genotypes of individuals, using actual relatedness

Outline

- Start from basic Twin Model for monozygotic MZ & dizygotic DZ twins
  - See videos on Univariate Twin Modeling in OpenMx
    - 1: intro to Classical Twin Study Design (18 min)
    - 2: saturated model: estimating means/covariances (22 min)
    - 3: twin model: estimating variance components (34 min)
    - 4: path versus variance estimation (11min)
- Model extensions
- Alternate parameterizations
- Focus on expected **variance/covariance matrices** defining the model!

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

OpenMx details

www.openmx.ssri.psu.edu

- OpenMx has a very fluid and flexible structure
- Code snippets saved as objects
- Object names often reused across scripts
  - Very few “reserved" names
  - Naming a matrix "mean" does not make it a mean
- Projects also contain data so files can be large

Matrices as Building Blocks

- covA <- mxMatrix(type="Lower", nrow=nv, ncol=nv, free=TRUE, values=.6, labels="a11", name="a")
  - Many types eg. type="Lower","Full","Symm", etc.
  - Size eg. nrow=nv, ncol=nv
- mxMatrix name eg. name="a", used in mxAlgebra
- Object name, eg. covA, used in mxModel
- Estimated parameters must be placed in matrices

Simulated Dataset

- 2000 pairs of twins with extra full sibling
- Variables:
  - Phenotypes: **Twin1 Twin2 Sib**
  - Ages of siblings: age1 age2 age3
  - Sex of siblings (0=male, 1=female): sex1 sex2 sex3
  - Assigned Zygosity (1=MZ, 2=DZ): zyg
  - Expected relatedness: relT (twins) relS (siblings)
  - Genomic relatedness: rel12 rel13 rel23

Actual Data: top 1000: MZ; bottom 1000: DZ

MZ

```
> round(head(twsDataS),2)
  Twin1 Twin2   Sib rel12 rel13 rel23  age1  age2  age3 sex1 sex2 sex3 relT relS zyg
1  1.64  0.57 -0.59     1  0.49  0.52 35.74 23.02 18.08    1    1    0    1  0.5   1
2  1.30  1.40  0.19     1  0.51  0.48 38.44 30.50 21.67    0    1    1    1  0.5   1
3  0.65  1.53  0.69     1  0.48  0.52 26.05 23.91 17.53    1    1    0    1  0.5   1
4 -0.66 -0.18 -0.93     1  0.55  0.52 39.33 27.62 24.83    0    0    0    1  0.5   1
5 -0.19 -0.09 -0.97     1  0.47  0.49 30.20 27.01 18.89    0    0    1    1  0.5   1
6  1.24  1.37  1.52     1  0.47  0.59 23.23 20.30 34.77    1    1    1    1  0.5   1
```

DZ

```
> round(head(twsDataS[twsDataS$zygosity==2,]),2)
     Twin1 Twin2   Sib rel12 rel13 rel23  age1  age2  age3 sex1 sex2 sex3 relT relS zyg
1001 -0.66  1.46 -0.74  0.47  0.51  0.44 20.32 27.64 29.78    0    0    0  0.5  0.5   2
1002  1.62  0.69  0.15  0.51  0.53  0.54 26.29 25.78 30.43    0    0    0  0.5  0.5   2
1003  0.71 -1.36  0.71  0.55  0.47  0.48 23.68 30.47 16.09    1    1    0  0.5  0.5   2
1004  0.67 -0.73  0.14  0.47  0.50  0.53 23.82 33.37 22.59    1    1    1  0.5  0.5   2
1005 -0.13 -0.60 -0.14  0.50  0.48  0.54 31.19 33.84 19.37    1    1    0  0.5  0.5   2
1006 -0.75 -0.67  0.59  0.50  0.52  0.45 36.44 32.91 28.37    1    1    1  0.5  0.5   2
```

Notation of Variance Components

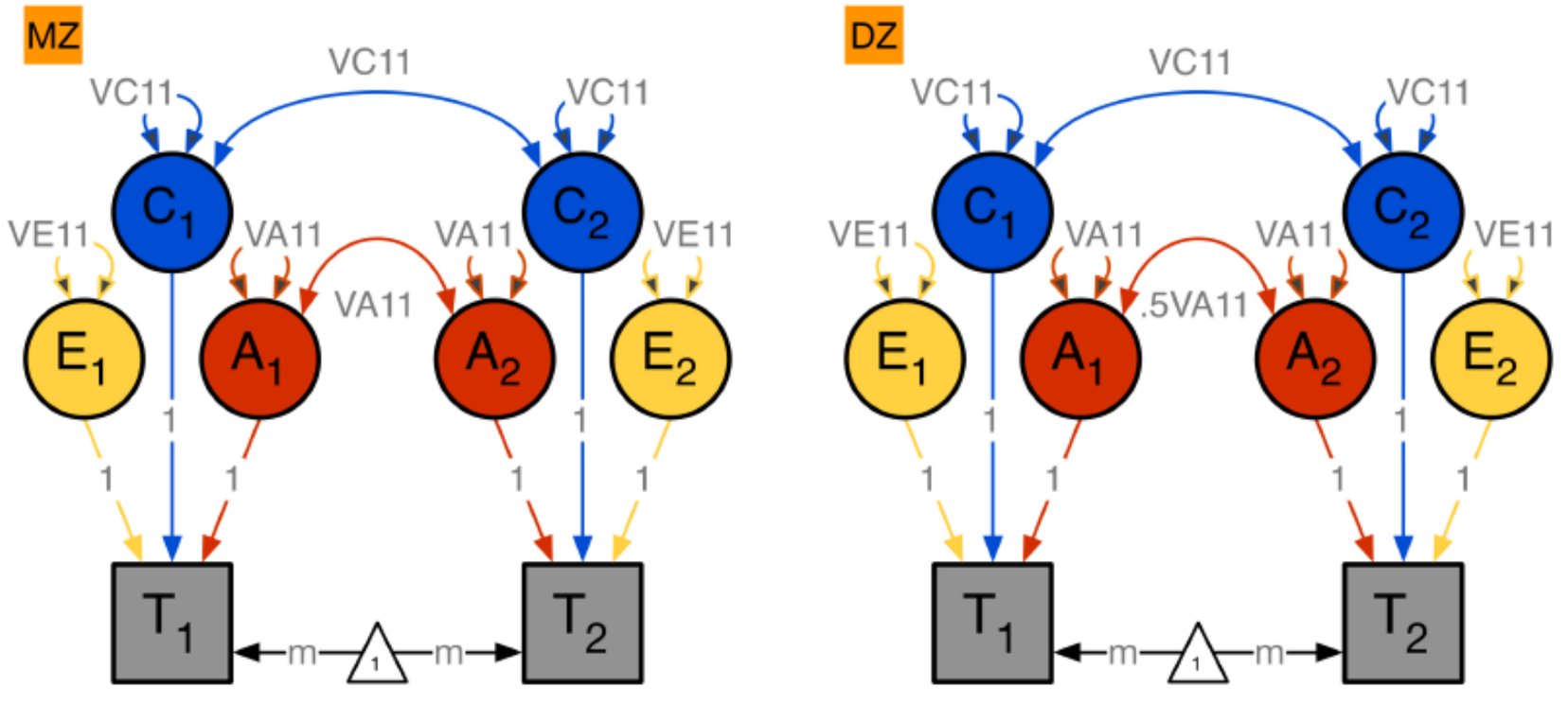
- **V**: Phenotypic variance
  - VA, a^2 or **A**: Additive genetic variance
  - VC, c^2 or **C**: Common/shared environmental variance
  - VE, e^2 or **E**: Unique environmental variance
  - VD, d^2 or **D**: Dominance genetic variance
- **h^2**: Heritability
  - h^2_B = **A**+**D** / **V**: Broad heritability
  - h^2_N = **A** / **V**: Narrow heritability

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

Basic ACE model: Expected Covariance Matrices

MZ and DZ pairs – estimating A, C and E



	T1	T2
T1	**A+C+E**	**A+C**
T2	**A+C**	**A+C+E**

	T1	T2
T1	**A+C+E**	**.5⊗A+C**
T2	**.5⊗A+C**	**A+C+E**

OpenMx Script I: Data

oneACEvc_**1cov**.R

```
# ----------------------------------------------------------------------------------------------------------------------
#  Script: oneACEvc_1cov.R
#  Author: Sarah Medland - Hermine Maes
#    Date: 29 02 2020 - 05 05 2021
#
# Twin Univariate ACE model to estimate causes of variation across multiple groups
# Matrix style model - Raw data - Continuous data
# -------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|

# Load Libraries & Options
rm(list=ls())
library(OpenMx)
library(psych)
source("miFunctions.R")
mxOption(NULL,"Default optimizer","NPSOL")

# Create Output
filename  <- "oneACEvc_1cov"
sink(paste(filename,".Ro",sep=""), append=FALSE, split=TRUE)

# ----------------------------------------------------------------------------------------------------------------------
# PREPARE DATA

# Load Data
tsDataS  <- read.table("tsDataS2.txt", header=T)
dim(tsDataS)
describe(tsDataS, skew=F)

# Select Variables for Analysis
nv        <- 1                                # number of variables
ntv       <- nv*2                             # number of total variables
selVars   <- c('Twin1','Twin2')               # list of variables names of observed variables
covVars   <- c('age1','age2','sex1','sex2')   # list of variables names of covariates

# Select Data for Analysis
mzData    <- subset(tsDataS, zyg==1, c(selVars,covVars))
dzData    <- subset(tsDataS, zyg==2, c(selVars,covVars))
cov(mzData[,selVars],use="complete")
cov(dzData[,selVars],use="complete")

# Set Starting Values
svBe      <- .01                              # start value for regressions
svMu      <-  0                               # start value for means
svVa      <- .2                               # start value for path coefficient
svVe      <- .5                               # start value for path coefficient for e
```

Prepare Data

OpenMx Script II: Model

oneACEvc_**1cov**.R

Means

```
# Create Matrices for Covariates and linear Regression Coefficients
betaS     <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe, labels="betaS", name="bS" )
betaA     <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe, labels="betaA", name="bA" )
defSex    <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE, labels=c("data.sex1","data.sex2"), name="Sex" )
defAge    <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE, labels=c("data.age1","data.age2"), name="Age" )

# Create Algebra for expected Mean Matrices
intercept <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMu, labels="interC", name="intercept" )
expMean   <- mxAlgebra( expression = intercept + bS*Sex + bA*Age, name="expMean" )
```

Covariances

```
# Create Matrices for Variance Components
covA      <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVa, label="VA11", name="VA" )
covC      <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVa, label="VC11", name="VC" )
covE      <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVe, label="VE11", name="VE" )

# Create Algebra for expected Variance/Covariance Matrices in MZ & DZ twins
covP      <- mxAlgebra( expression= VA+VC+VE, name="V" )
covMZ     <- mxAlgebra( expression= VA+VC, name="cMZ" )
covDZ     <- mxAlgebra( expression= 0.5%x%VA+ VC, 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" )
```

```
# 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="expMean", dimnames=selVars )
expDZ     <- mxExpectationNormal( covariance="expCovDZ", means="expMean", dimnames=selVars )
funML     <- mxFitFunctionML()
```

Models for MZ&DZ

```
# Create Model Objects for Multiple Groups
defs      <- list( defAge, defSex)
pars      <- list( intercept, betaS, betaA, covA, covC, covE, covP )
modelMZ   <- mxModel( pars, defs, expMean, covMZ, expCovMZ, dataMZ, expMZ, funML, name="MZ" )
modelDZ   <- mxModel( pars, defs, expMean, covDZ, expCovDZ, dataDZ, expDZ, funML, name="DZ" )
multi     <- mxFitFunctionMultigroup( c("MZ","DZ") )
```

```
# Create Algebra for Unstandardized & Standardized Variance Components
rowUS     <- rep('US',nv)
colUS     <- rep(c('VA','VC','VE','SA','SC','SE'),each=nv)
estUS     <- mxAlgebra( expression=cbind(VA,VC,VE,VA/V,VC/V,VE/V), name="US", dimnames=list(rowUS,colUS) )

# Create Confidence Interval Objects
ciACE     <- mxCI( "US[1,1:6]" )
```

Overall Model

```
# Build Model with Confidence Intervals
modelACE  <- mxModel( "oneACEvc_1cov", pars,  modelMZ, modelDZ, multi, estUS, ciACE )
```

OpenMx Script III: Run

oneACEvc_**1cov**.R

```
# ------------------------------------------------------------------------------------------------------------------------
# RUN MODEL

# Run ACE Model
fitACE    <- mxRun( modelACE, intervals=T )
sumACE    <- summary( fitACE )

# Print Goodness-of-fit Statistics & Parameter Estimates
fitGofs(fitACE)
fitEstCis(fitACE)

# ------------------------------------------------------------------------------------------------------------------------
sink()
```

Run Model to Data

Output

Descriptive statistics

```
> colMeans(mzData[,selVars],na.rm=TRUE)
     Twin1      Twin2
0.15939615 0.19972023
> colMeans(dzData[,selVars],na.rm=TRUE)
     Twin1      Twin2
0.17799030 0.15720902

> cov(mzData[,selVars],use="complete")
           Twin1      Twin2
Twin1 0.89973996 0.61810038
Twin2 0.61810038 0.89006913
> cov(dzData[,selVars],use="complete")
           Twin1      Twin2
Twin1 0.83616769 0.39190014
Twin2 0.39190014 0.91150869
```

Goodness-of-fit statistics

```
> fitGofs(fitACE)
Mx:oneACEvc_1cov  os=4000  ns=2000   ep=6   co=0  df=3994  ll=9975.7666  cpu=32.9968  opt=NPSOL  ver=2.17.3.145  stc=0
```

Parameter Estimates & Confidence Intervals

```
> fitEstCIs(fitACE,colUS)
interC  betaS  betaA   VA11   VC11   VE11
0.0746 0.0982 0.0016 0.4128 0.1907 0.2763

             VA     VC     VE     SA     SC     SE
lbound   0.3210 0.1019 0.2537 0.3656 0.1169 0.2858
estimate 0.4128 0.1907 0.2763 0.4692 0.2168 0.3140
ubound   0.5100 0.2764 0.3016 0.5774 0.3096 0.3452
```

Variance Components

oneACEvc_**1cov**.R

```
nv      <- 1
covA    <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVa,
             label="VA11", name="VA" )
covC    <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVa,
             label="VC11", name="VC" )
covE    <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVe,
             label="VE11", name="VE" )
covP    <- mxAlgebra( expression= VA +VC +VE, name="V" )
covMZ   <- mxAlgebra( expression= VA +VC, name="cMZ" )
covDZ   <- mxAlgebra( expression= 0.5%x%VA +VC, name="cDZ" )
```

Object	covA	covC	covE	covP	covMZ	covDZ
Matrix	VA	VC	VE	V	cMZ	cDZ
	A	+ C	+ E	= A+C+E		
	A	+ C		=	A+C	
.5 ⊗	A	+ C		=		.5⊗A+C

Variance Components

oneACEvc_**1cov**.R

```
nv       <- 1
covA     <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVa,
              label="VA11", name="VA" )
covC     <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVa,
              label="VC11", name="VC" )
covE     <- mxMatrix( type="Symm", nrow=nv, ncol=nv, free=TRUE, values=svVe,
              label="VE11", name="VE" )
covP     <- mxAlgebra( expression= VA +VC +VE, name="V" )
covMZ    <- mxAlgebra( expression= VA +VC, name="cMZ" )
covDZ    <- mxAlgebra( expression= 0.5%x%VA +VC, name="cDZ" )
```

Object	covA	covC	covE	covP	covMZ	covDZ
Matrix	VA	VC	VE	V	cMZ	cDZ
	A	+ C	+ E	= A+C+E		
	A	+ C		=	A+C	
.5 ⊗	A	+ C		=		.5⊗A+C

Expected Covariance Matrices oneACEvc_**1cov**.R

```
covP      <- mxAlgebra( expression= VA +VC +VE, name="V" )
covMZ     <- mxAlgebra( expression= VA +VC, name="cMZ" )
covDZ     <- mxAlgebra( expression= 0.5%x%VA +VC, 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" )
```

MZ	T1	T2
T1	**A+C+E**	**A+C**
T2	**A+C**	**A+C+E**

=

	T1	T2
T1	**V**	**cMZ**
T2	**cMZ'**	**V**

DZ	T1	T2
T1	**A+C+E**	**.5⊗A+C**
T2	**.5⊗A+C**	**A+C+E**

=

	T1	T2
T1	**V**	**cDZ**
T2	**cDZ'**	**V**

Definition Variables

- "data.varName" in label argument of mxMatrix indicates definition variable
- Matrix element updated dynamically with value from dataset when row likelihood is calculated

```
> round(head(twsDataS),2)
  Twin1 Twin2   Sib rel12 rel13 rel23  age1  age2  age3 sex1 sex2 sex3 relT relS zyg
1  1.64  0.57 -0.59     1  0.49  0.52 35.74 23.02 18.08    1    1    0    1  0.5   1
2  1.30  1.40  0.19     1  0.51  0.48 38.44 30.50 21.67    0    1    1    1  0.5   1

betaS     <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe,
             labels="betaS", name="bS" )
defSex    <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE,
             labels=c("data.sex1","data.sex2"), name="Sex" )
intercept <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMu,
             labels="interC", name="intercept" )
expMean   <- mxAlgebra( expression= intercept +bS*Sex, name="expMean" )
```

Means: Intercept + Covariates oneACEvc_**1cov**.R

```
ntv      <- nv*2
betaS    <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe,
              labels="betaS", name="bS" )
betaA    <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe,
              labels="betaA", name="bA" )
defSex   <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE,
              labels=c("data.sex1","data.sex2"), name="Sex" )
defAge   <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE,
              labels=c("data.age1","data.age2"), name="Age" )

intercept <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMu,
              labels="interC", name="intercept" )
expMean   <- mxAlgebra( expression= intercept +bS*Sex +bA*Age, name="expMean" )
```

	T1	T2
Mu	**icept + bS*Sex + bA*Age**	**icept + bS*Sex + bA*Age**

Models for MZ, DZ, combined oneACEvc_**1cov**.R

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

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

defs       <- list( defAge, defSex)
pars       <- list( intercept, betaS, betaA, covA, covC, covE, covP )
modelMZ    <- mxModel( pars, defs, expMean, covMZ, expCovMZ, dataMZ, expMZ,
              funML, name="MZ" )
modelDZ    <- mxModel( pars, defs, expMean, covDZ, expCovDZ, dataDZ, expDZ,
              funML, name="DZ" )
multi      <- mxFitFunctionMultigroup( c("MZ","DZ") )

modelACE   <- mxModel( "oneACEvc_1cov", pars,  modelMZ, modelDZ, multi, estUS,
              ciACE )
```

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

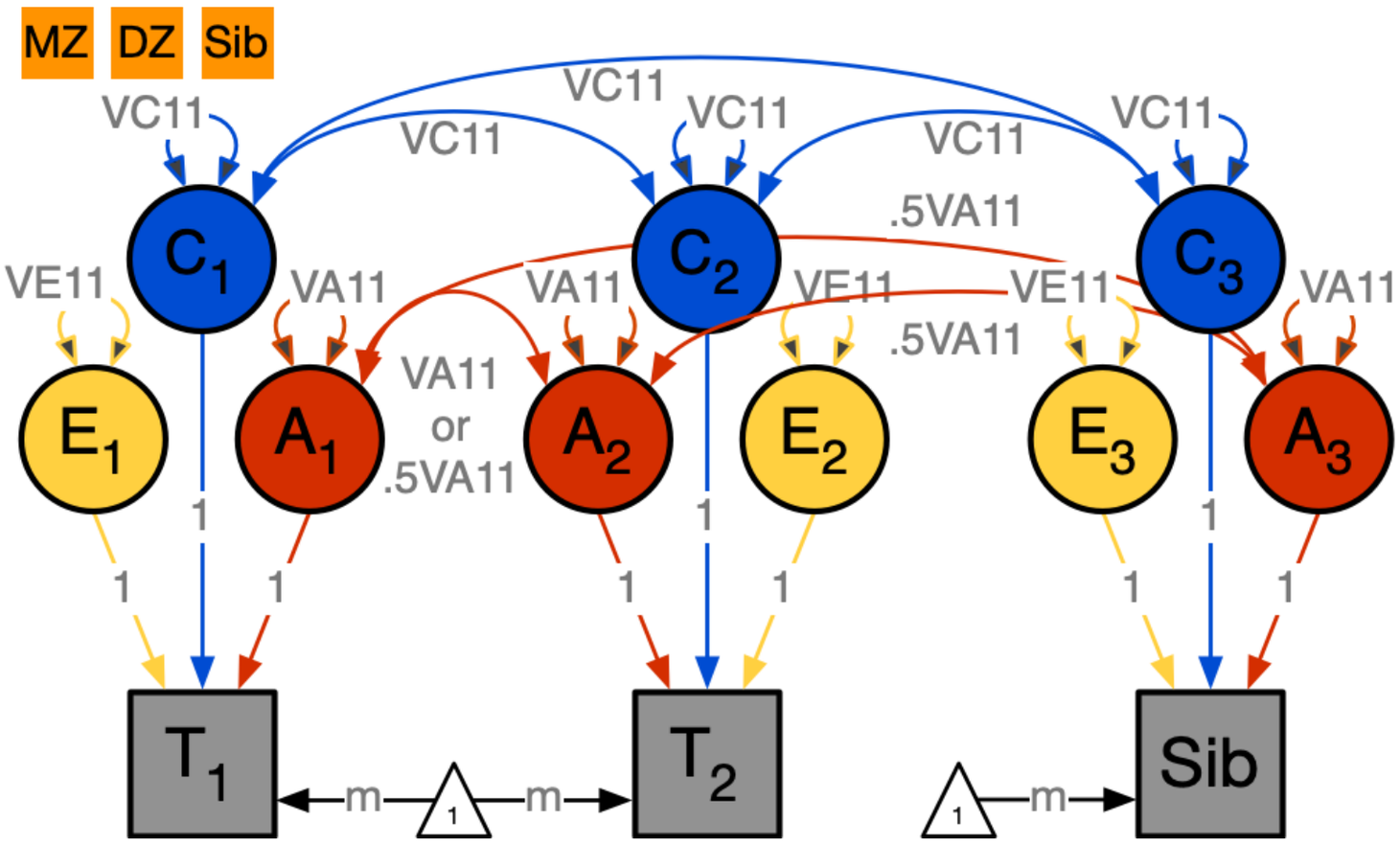
Additional Sibling

- 1 extra non-twin full sibling per twin pair
- Variance of sibling in ACE model?
- Covariance between sibling and twin1 / twin 2?
- Same for MZ and DZ families?

	T1	T2	Sib
T1	**A+C+E**	**(.5⊗)A+C**	**?**
T2	**(.5⊗)A+C**	**A+C+E**	**?**
Sib	**?**	**?**	**?**

Extra Sibling

- MZ and DZ pairs + extra sibling – still estimating A, C and E



Means: Intercept + Covariates oneACEvc_**2sib**.R

```
ntv       <- nv*3
betaS     <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe,
               labels="betaS", name="bS" )
betaA     <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE, values=svBe,
               labels="betaA", name="bA" )
defSex    <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE,
               labels=c("data.sex1","data.sex2","data.sex3"), name="Sex" )
defAge    <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=FALSE,
               labels=c("data.age1","data.age2","data.age3"), name="Age" )

intercept <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMu,
               labels="interC", name="intercept" )
expMean   <- mxAlgebra( expression= intercept +bS*Sex +bA*Age, name="expMean" )
```

	Twin1	Twin2	Sib
Mu	**icept + bS*Sex + bA*Age**	**icept + bS*Sex + bA*Age**	**icept + bS*Sex + bA*Age**

ACE model: Add Sibling

oneACEvc_**2sib**.R

```
expCovMZ  <- mxAlgebra( rbind( cbind(V,       cMZ,    cDZ),
                               cbind(t(cMZ), V,       cDZ),
                               cbind(t(cDZ), t(cDZ), V)), name="expCovMZ" )
expCovDZ  <- mxAlgebra( rbind( cbind(V,       cDZ,    cDZ),
                               cbind(t(cDZ), V,       cDZ),
                               cbind(t(cDZ), t(cDZ), V)), name="expCovDZ" )
```

MZ	T1	T2	Sib
T1	**A+C+E**	**A+C**	**.5⊗A+C**
T2	**A+C**	**A+C+E**	**.5⊗A+C**
Sib	**.5⊗A+C**	**.5⊗A+C**	**A+C+E**

	T1	T2	Sib
T1	**V**	**cMZ**	**cDZ**
T2	**cMZ'**	**V**	**cDZ**
Sib	**cDZ'**	**cDZ'**	**V**

DZ	T1	T2	Sib
T1	**A+C+E**	**.5⊗A+C**	**.5⊗A+C**
T2	**.5⊗A+C**	**A+C+E**	**.5⊗A+C**
Sib	**.5⊗A+C**	**.5⊗A+C**	**A+C+E**

	T1	T2	Sib
T1	**V**	**cDZ**	**cDZ**
T2	**cDZ'**	**V**	**cDZ**
Sib	**cDZ'**	**cDZ'**	**V**

Multiple Siblings

- Some families have siblings, others don't
- Full information maximum likelihood (FIML) methods, assumes missing at random
- Model biggest family size
- Missing phenotypes for non-existent sibs BUT non-missing 'dummy' covariates

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

Variation on a Theme (of Paganini*)

- Writing out full covariance matrix quickly unwieldy
- Imagine doing this if largest family = 10 sibs…
- Example with just 3 extra siblings:

```
expCovMZ  <- mxAlgebra( rbind( cbind(V,      cMZ,    cDZ,    cDZ,    cDZ),
                               cbind(t(cMZ), V,      cDZ,    cDZ,    cDZ),
                               cbind(t(cDZ), t(cDZ), V,      cDZ,    cDZ),
                               cbind(t(cDZ), t(cDZ), t(cDZ), V,      cDZ),
                               cbind(t(cDZ), t(cDZ), t(cDZ), t(cDZ), V  )),
                               name="expCovMZ" )
```

* https://www.allmusic.com/performance/variations-on-a-theme-by-paganini-mq0002271781

Alternate Parameterization A

oneACEvc_**3alt**.R

```
relAmz    <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
             values=c(1,1,.5,1,.5,1), name="rAmz" )
relAdz    <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
             values=c(1,.5,.5,1,.5,1), name="rAdz" )
```

MZ	T1	T2	Sib
T1	**A+C+E**	**A+C**	**.5⊗A+C**
T2	**A+C**	**A+C+E**	**.5⊗A+C**
Sib	**.5⊗A+C**	**.5⊗A+C**	**A+C+E**

relAmz	T1	T2	Sib
T1	**1**	**1**	**.5**
T2	**1**	**1**	**.5**
Sib	**.5**	**.5**	**1**

⊗A

DZ	T1	T2	Sib
T1	**A+C+E**	**.5⊗A+C**	**.5⊗A+C**
T2	**.5⊗A+C**	**A+C+E**	**.5⊗A+C**
Sib	**.5⊗A+C**	**.5⊗A+C**	**A+C+E**

relAdz	T1	T2	Sib
T1	**1**	**.5**	**.5**
T2	**.5**	**1**	**.5**
Sib	**.5**	**.5**	**1**

⊗A

Alternate Parameterization CE oneACEvc_**3alt**.R

```
relC      <- mxMatrix( type="Unit", nrow=ntv, ncol=ntv, free=FALSE, name="rC" )
relE      <- mxMatrix( type="Iden", nrow=ntv, ncol=ntv, free=FALSE, name="rE" )
```

	T1	T2	Sib
T1	**A+C+E**	**A+C**	**.5⊗A+C**
T2	**A+C**	**A+C+E**	**.5⊗A+C**
Sib	**.5⊗A+C**	**.5⊗A+C**	**A+C+E**

relC	T1	T2	Sib
T1	**1**	**1**	**1**
T2	**1**	**1**	**1**
Sib	**1**	**1**	**1**

⊗C

	T1	T2	Sib
T1	**A+C+E**	**.5⊗A+C**	**.5⊗A+C**
T2	**.5⊗A+C**	**A+C+E**	**.5⊗A+C**
Sib	**.5⊗A+C**	**.5⊗A+C**	**A+C+E**

relE	T1	T2	Sib
T1	**1**	**0**	**0**
T2	**0**	**1**	**0**
Sib	**0**	**0**	**1**

⊗E

Alternate Parameterization ACE oneACEvc_**3alt**.R

```
relAmz    <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
              values=c(1,1,.5,1,.5,1), name="rAmz" )
relAdz    <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
              values=c(1,.5,.5,1,.5,1), name="rAdz" )
relC      <- mxMatrix( type="Unit", nrow=ntv, ncol=ntv, free=FALSE, name="rC" )
relE      <- mxMatrix( type="Iden", nrow=ntv, ncol=ntv, free=FALSE, name="rE" )
expCovMZ  <- mxAlgebra( VA%x%rAmz +VC%x%rC +VE%x%rE, name="expCovMZ" )
expCovDZ  <- mxAlgebra( VA%x%rAdz +VC%x%rC +VE%x%rE, name="expCovDZ" )
```

rAmz	T1	T2	S
T1	**1**	1	.5
T2	**1**	**1**	.5
Sib	**.5**	**.5**	**1**

rAdz	T1	T2	S
T1	**1**	.5	.5
T2	**.5**	**1**	.5
Sib	**.5**	**.5**	**1**

⊗**A**

rC	T1	T2	S
T1	**1**	**1**	**1**
T2	**1**	**1**	**1**
Sib	**1**	**1**	**1**

⊗**C**

rE	T1	T2	S
T1	**1**	0	0
T2	0	**1**	0
Sib	0	0	**1**

⊗**E**

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

More Efficient Approach?

Difference between MZ & DZ groups

```
relAmz     <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
              values=c(1,1,.5,1,.5,1), name="rAmz" )
relAdz     <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
              values=c(1,.5,.5,1,.5,1), name="rAdz" )
```

rAmz	T1	T2	S
T1	**1**	1	.5
T2	**1**	**1**	.5
Sib	**.5**	**.5**	**1**

rAdz	T1	T2	S
T1	**1**	.5	.5
T2	**.5**	**1**	.5
Sib	**.5**	**.5**	**1**

⊗**A**

rC	T1	T2	S
T1	**1**	**1**	**1**
T2	**1**	**1**	**1**
Sib	**1**	**1**	**1**

⊗**C**

rE	T1	T2	S
T1	**1**	0	0
T2	0	**1**	0
Sib	0	0	**1**

⊗**E**

Definition Variables

oneACEvc_**4def**.R

- Read relationship coefficient for zygosity from data
- Only one group with definition variable

```
relA      <- mxMatrix( type="Stand", nrow=ntv, ncol=ntv, free=FALSE,
 labels=c("data.relT","data.relS","data.relS"), name="rA" )
```

rA	T1	T2	Sib
T1	**1**	relT	relS
T2	**relT**	**1**	relS
Sib	**relS**	**relS**	**1**

relT = 1 for MZs
relT = .5 for DZs

relS =.5 for siblings

relT,relS: Expected Relatedness oneACEvc_**4def**.R

```
relA       <- mxMatrix( type="Stand", nrow=ntv, ncol=ntv, free=FALSE,
                values=c("data.relT","data.relS","data.relS"), name="rA" )
relC       <- mxMatrix( type="Unit", nrow=ntv, ncol=ntv, free=FALSE, name="rC" )
relE       <- mxMatrix( type="Iden", nrow=ntv, ncol=ntv, free=FALSE, name="rE" )
expCovTW   <- mxAlgebra( VA%x%rA +VC%x%rC +VE%x%rE, name="expCovTW" )
dataTW     <- mxData( observed=twsDataS, type="raw" )
expTW      <- mxExpectationNormal( covariance="expCovTW", means="expMean",
                dimnames=selVars )
modelACE   <- mxModel( "ACEvc_4def", pars, defs, expMean, expCovTW, dataTW,
                expTW, funML, estUS, ciACE )
```

rA	T1	T2	S
T1	**1**	relT	relS
T2	**relT**	**1**	relS
Sib	**relS**	**relS**	**1**

⊗**A**

rC	T1	T2	S
T1	**1**	**1**	**1**
T2	**1**	**1**	**1**
Sib	**1**	**1**	**1**

⊗**C**

rE	T1	T2	S
T1	**1**	0	0
T2	0	**1**	0
Sib	0	0	**1**

⊗**E**

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

Actual Data: top 1000: MZ; bottom 1000: DZ

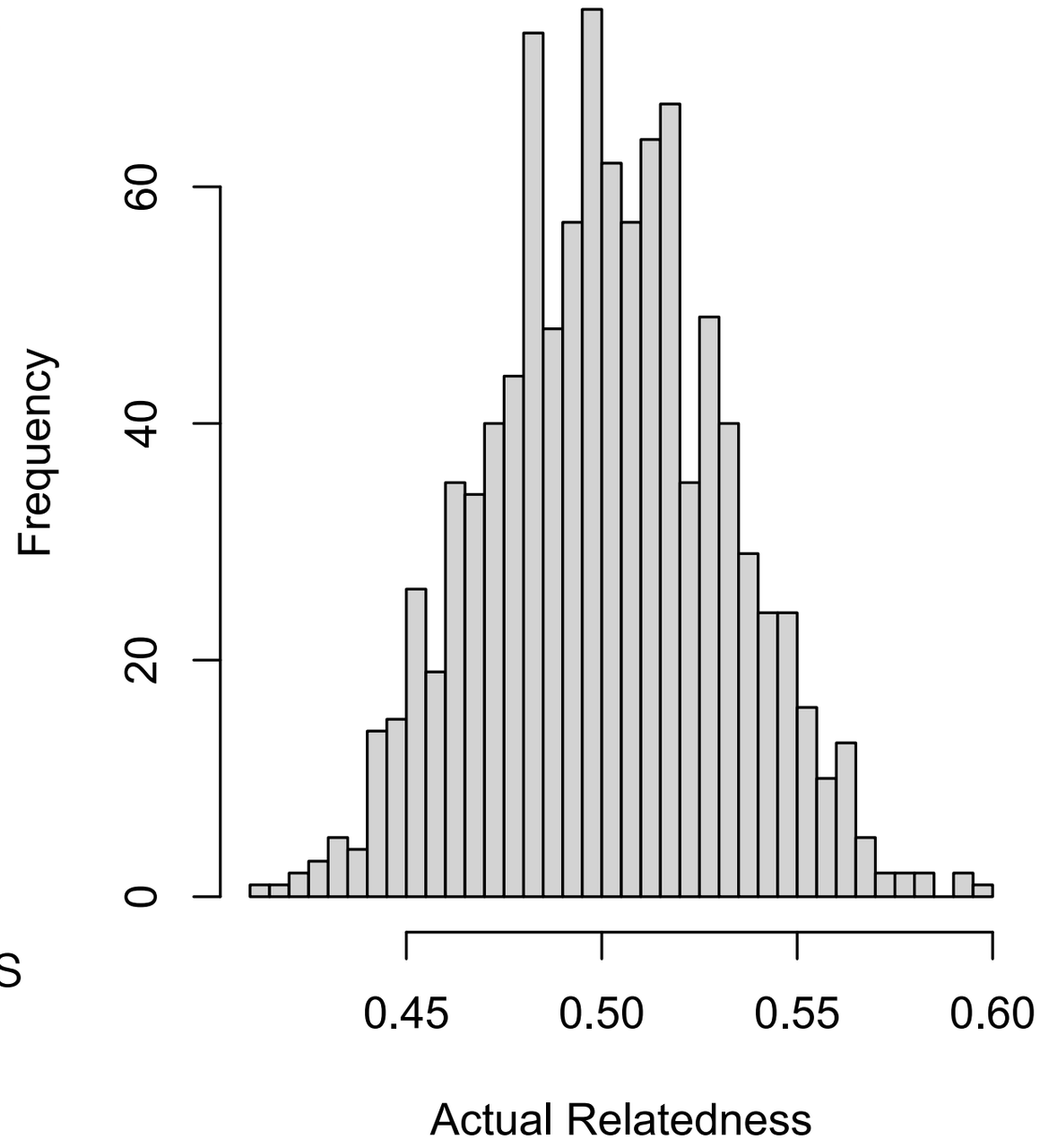
```
> round(head(twsDataS),2)
  Twin1 Twin2   Sib rel12 rel13 rel23  ag
1  1.64  0.57 -0.59     1  0.49  0.52 35.
2  1.30  1.40  0.19     1  0.51  0.48 38.
3  0.65  1.53  0.69     1  0.48  0.52 26.
4 -0.66 -0.18 -0.93     1  0.55  0.52 39.
5 -0.19 -0.09 -0.97     1  0.47  0.49 30.
6  1.24  1.37  1.52     1  0.47  0.59 23.
```

MZ

```
> round(head(twsDataS[twsDataS$zygosity==
     Twin1 Twin2   Sib rel12 rel13 rel23
1001 -0.66  1.46 -0.74  0.47  0.51  0.44
1002  1.62  0.69  0.15  0.51  0.53  0.54
1003  0.71 -1.36  0.71  0.55  0.47  0.48
1004  0.67 -0.73  0.14  0.47  0.50  0.53
1005 -0.13 -0.60 -0.14  0.50  0.48  0.54
1006 -0.75 -0.67  0.59  0.50  0.52  0.45
```

DZ

Actual genomic relatedness for DZs



Actual Relatedness

oneACEvc_**5rel**.R

- Actual genetic relatedness instead of .5 or 1
- Estimate genetic relatedness by computing a genetic relationship matrix GRM in PLINK or GCTA

```
relA      <- mxMatrix( type="Stand", nrow=ntv, ncol=ntv, free=FALSE,
              labels=c("data.rel12","data.rel13","data.rel23"), name="rA" )
```

rA	T1	T2	S
T1	**1**	rel12	rel13
T2	**rel12**	**1**	rel23
Sib	**rel13**	**rel23**	**1**

⊗**A**

rC	T1	T2	S
T1	**1**	**1**	**1**
T2	**1**	**1**	**1**
Sib	**1**	**1**	**1**

⊗**C**

rE	T1	T2	S
T1	**1**	0	0
T2	0	**1**	0
Sib	0	0	**1**

⊗**E**

rel12 etc.: Actual Relatedness oneACEvc_**5rel**.R

```
relA      <- mxMatrix( type="Stand", nrow=ntv, ncol=ntv, free=FALSE,
               values=c("data.rel12","data.rel13","data.rel23"), name="rA" )
relC      <- mxMatrix( type="Unit", nrow=ntv, ncol=ntv, free=FALSE, name="rC" )
relE      <- mxMatrix( type="Iden", nrow=ntv, ncol=ntv, free=FALSE, name="rE" )
expCovTW  <- mxAlgebra( VA%x%rA +VC%x%rC +VE%x%rE, name="expCovTW" )
dataTW    <- mxData( observed=twsDataS, type="raw" )
expTW     <- mxExpectationNormal( covariance="expCovTW", means="expMean",
               dimnames=selVars )
modelACE  <- mxModel( "ACEvc_5rel", pars, defs, expMean, expCovTW, dataTW,
               expTW, funML, estUS, ciACE )
```

rA	T1	T2	S
T1	**1**	rel12	rel13
T2	**rel12**	**1**	rel23
Sib	**rel13**	**rel23**	**1**

⊗**A**

rC	T1	T2	S
T1	**1**	1	1
T2	**1**	**1**	1
Sib	**1**	**1**	**1**

⊗**C**

rE	T1	T2	S
T1	**1**	0	0
T2	0	**1**	0
Sib	0	0	**1**

⊗**E**

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

From Classical Twin Study to Siblings/DZs only

- With measured relationships, MZs technically not needed for model identification

PLOS GENETICS

advanced search

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings

Peter M Visscher, Sarah E Medland, Manuel A. R Ferreira, Katherine I Morley, Gu Zhu, Belinda K Cornes, Grant W Montgomery, Nicholas G Martin

Published: March 24, 2006 • https://doi.org/10.1371/journal.pgen.0020041

Synopsis

Quantitative geneticists attempt to understand variation between individuals within a population for traits such as height in humans and the number of bristles in fruit flies. This has been traditionally done by partitioning the variation in underlying sources due to genetic and environmental factors, using the observed amount of variation between and within families. A problem with this approach is that one can never be sure that the estimates are correct, because nature and nurture can be confounded without one knowing it. The authors got around this problem by comparing the similarity between relatives as a function of the exact proportion of genes that they have in common, looking only within families. Using this approach, the authors estimated the amount of total variation for height in humans that is due to genetic factors from 3,375 sibling pairs. For each pair, the authors estimated the proportion of genes that they share from DNA markers. It was found that about 80% of the total variation can be explained by genetic factors, close to results that are obtained from classical studies. This study provides the first validation of an estimate of genetic variation by using a source of information that is free from nature–nurture assumptions.

- Use if
  - equal environments assumption EEA problematic
  - only data available for sibling pairs

Actual Relatedness DZs only

oneACEvc_**6dzs**.R

```
dzData    <- subset(tsDataS, zyg==2, c(selVars,covVars))
relA      <- mxMatrix( type="Stand", nrow=ntv, ncol=ntv, free=FALSE,
               values=c("data.rel12","data.rel13","data.rel23"), name="rA" )
relC      <- mxMatrix( type="Unit", nrow=ntv, ncol=ntv, free=FALSE, name="rC" )
relE      <- mxMatrix( type="Iden", nrow=ntv, ncol=ntv, free=FALSE, name="rE" )
expCovTW  <- mxAlgebra( VA%x%rA +VC%x%rC +VE%x%rE, name="expCovTW" )
dataDZ    <- mxData( observed=tsDataS[tsDataS$zyg==2,], type="raw" )
expTW     <- mxExpectationNormal( covariance="expCovTW", means="expMean",
               dimnames=selVars )
modelACE  <- mxModel( "ACEvc_6dzs", pars, defs, expMean, expCovTW, dataDZ,
               expTW, funML, estUS, ciACE )
```

rA	T1	T2	S
T1	**1**	rel12	rel13
T2	**rel12**	**1**	rel23
Sib	**rel13**	**rel23**	**1**

⊗**A**

rC	T1	T2	S
T1	**1**	**1**	**1**
T2	**1**	**1**	**1**
Sib	**1**	**1**	**1**

⊗**C**

rE	T1	T2	S
T1	**1**	0	0
T2	0	**1**	0
Sib	0	0	**1**

⊗**E**

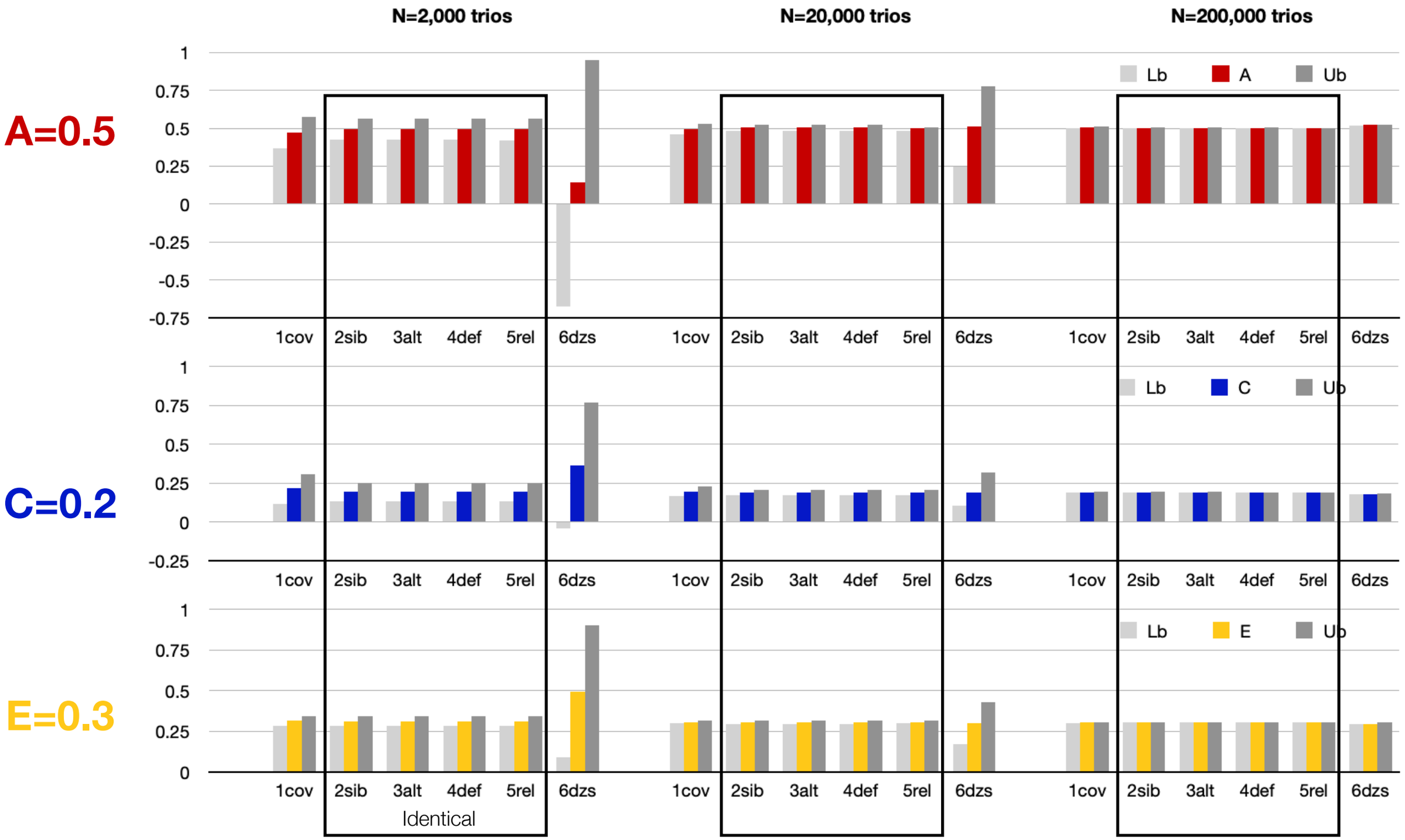
Summary of Variations

oneACEvc_.....R

- 1cov: Original ACE/ADE
  - rMZ=VA+VC, rDZ=.5VA + VC
- 2sib: Additional sibling
  - rMZ=VA+VC, rDZ/rSib=.5VA + VC
- 3alt: Alternate parameterization
  - VA*relAmz/relAdz + VC*relC + VE*relE
- 4def: Definition variables
  - VA*relA[relT|relS] + VC*relC + VE*relE
- 5rel: Actual relatedness
  - VA*relA[rel12 etc] + VC*relC + VE*relE
- 6dzs: Actual relatedness DZ's only

MZ	DZ		2x2
MZ	DZ	Sib	3x3
MZ	DZ	Sib	3x3
MZ	DZ	Sib	3x3
MZ	DZ	Sib	3x3
	DZ	Sib	3x3

Parameter Estimates & CIs across simulations|models



Difference in Efficiency? In seconds, with CIs

		1,000 MZ 1,000 DZ	10,000 MZ 10,000 DZ	100,000 MZ 100,000 DZ
	observed statistics			
oneACEvc_1cov	4000[0[0]]	33.00	400.49	2905.35
oneACEvc_2sib	6000[0[0]]	45.80	548.75	2061.60
oneACEvc_3alt	6000[0[0]]	39.47	644.97	2485.97
oneACEvc_4def	6000[0[0]]	44.24	680.89	2177.31
oneACEvc_5rel	6000[0[0]]	55.36	1010.28	3043.85
oneACEvc_6dzs	3000[0[0]]	17.07	338.14	1849.06

Files /faculty/hmaes/2021/ oneACEvc_.....R

- 1cov: Original ACE/ADE model
- 2sib: Additional sibling
- 3alt: Alternate parameterization
- 4def: Definition variables
- 5rel: Actual relatedness
- 6dzs: Actual relatedness DZ's only
- 7unr: Unrelated individuals

'Unrelated' individuals

oneAEvc_**7unr**.R

```
relA      <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=FALSE,
               values=c("data.rel12","data.rel13","data.rel23",..), name="rA" )
relE      <- mxMatrix( type="Iden", nrow=ntv, ncol=ntv, free=FALSE, name="rE" )
expCovUR  <- mxAlgebra( VA%x%rA +VE%x%rE, name="expCov" )
dataUR    <- mxData( observed=unrDataS, type="raw" )
expUR     <- mxExpectationNormal( "expCov", "expMean", dimnames=selVars )
modelAE   <- mxModel( "AEvc_7unr", pars, expMean, expCov, dataUR, expUR, funML )
```

rA	P1	P2	P3	P4	P5	P6	P7	P8	P9	..	PN
P1	**1**	rel12	rel13	rel14	rel15	rel16	rel17	rel18	rel19		rel1N
P2	**rel12**	**1**	rel23	rel24	rel25	rel26	rel27	rel28	rel29		rel2N
P3	**rel13**	**rel23**	**1**	rel34	rel35	rel36	rel37	rel38	rel39		rel3N
P4	**rel14**	**rel24**	**rel34**	**1**	rel45	rel46	rel47	rel48	rel49		rel4N
P5	**rel15**	**rel25**	**rel35**	**rel45**	**1**	rel56	rel57	rel58	rel59		rel5N
P6	**rel16**	**rel26**	**rel36**	**rel46**	**rel56**	**1**	rel67	rel68	rel69		rel6N
P7	**rel17**	**rel27**	**rel37**	**rel47**	**rel57**	**rel67**	**1**	rel78	rel79		rel7N
P8	**rel18**	**rel28**	**rel38**	**rel48**	**rel58**	**rel68**	**rel78**	**1**	rel89		rel8N
P9	**rel19**	**rel29**	**rel39**	**rel49**	**rel59**	**rel69**	**rel79**	**rel89**	**1**		rel9N
..										**1**	rel..
PN	**rel1N**	**rel2N**	**rel3N**	**rel4N**	**rel5N**	**rel6N**	**rel7N**	**rel8N**	**rel9N**	**rel..**	**1**

rE	P1	P2	P3	..	PN
P1	**1**	0	0		0
P2	0	**1**	0		0
P3	0	0	**1**		0
P4	0	0	0		0
P5	0	0	0	..	0
P6	0	0	0		0
P7	0	0	0		0
P8	0	0	0		0
P9	0	0	0		0
				1	0
PN	0	0	0	0	**1**

Final variation….

- If we can use actual genetic relatedness of DZs, how about using actual genetic relatedness of unrelated individuals

Behavior Genetics
https://doi.org/10.1007/s10519-020-10037-5

ORIGINAL RESEARCH

Combining Structural-Equation Modeling with Genomic-Relatedness-Matrix Restricted Maximum Likelihood in OpenMx

Robert M. Kirkpatrick[1,3] · Joshua N. Pritikin[1] · Michael D. Hunter[2] · Michael C. Neale[1]

All models are wrong

- "Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful"

- George E P Box and Norman R Draper. 1986. Empirical Model-Building and Response Surface. John Wiley & Sons, Inc., New York, NY, USA.