

Univariate ACE Model

Boulder Workshop 2014

Hermine H. Maes, Elizabeth Prom-Wormley

Questions to be Answered

- Does a trait of interest cluster among related individuals?
- Can clustering be explained by genetic or environmental effects?
- What is the best way to explain the degree to which genetic and environmental effects affect a trait?

Practical Example

- Dataset: NH&MRC Twin Register
- 1981 Questionnaire
- BMI (body mass index): $\text{weight}/\text{height squared}$
- Young Female Cohort: 18-30 years
- Sample Size:
 - MZf_young: 534 pairs (zyg=1)
 - DZf_young: 328 pairs (zyg=3)

Dataset

```
> head(twinData)
```

| | fam | age | zyg | part | wt1 | wt2 | ht1 | ht2 | htwt1 | htwt2 | bmi1 | bmi2 |
|-------|-----|-----|-----|------|-----|-----|--------|--------|---------|---------|---------|---------|
| 1 | 115 | 21 | 1 | 2 | 58 | 57 | 1.7000 | 1.7000 | 20.0692 | 19.7232 | 20.9943 | 20.8726 |
| 2 | 121 | 24 | 1 | 2 | 54 | 53 | 1.6299 | 1.6299 | 20.3244 | 19.9481 | 21.0828 | 20.9519 |
| 3 | 158 | 21 | 1 | 2 | 55 | 50 | 1.6499 | 1.6799 | 20.2020 | 17.7154 | 21.0405 | 20.1210 |
| 4 | 172 | 21 | 1 | 2 | 66 | 76 | 1.5698 | 1.6499 | 26.7759 | 27.9155 | 23.0125 | 23.3043 |
| 5 | 182 | 19 | 1 | 2 | 50 | 48 | 1.6099 | 1.6299 | 19.2894 | 18.0662 | 20.7169 | 20.2583 |
| 6 | 199 | 26 | 1 | 2 | 60 | 60 | 1.5999 | 1.5698 | 23.4375 | 24.3418 | 22.0804 | 22.3454 |
| | | | | | | | | | | | | |

Univariate Twin Saturated

twinsatConCov.R [1]

```
#=====
# Program: twinsatConCov.R
# Author: Hermine Maes
# Date: 03 03 2014
#
# Univariate Twin Saturated model to estimate means and (co)variances
# Matrix style model - Raw data - Continuous data
#=+----|-----|-----|-----|-----|-----|-----|-----
# Load Libraries
require(OpenMx)
require(psych)
# PREPARE DATA
# Load Data
data(twinData)
dim(twinData)
describe(twinData, skew=F)
```

Univariate Twin Saturated

twinsatConCov.R [2]

Select Variables for Analysis

```
Vars      <- 'bmi'  
nv        <- 1                # number of variables  
ntv       <- nv*2            # number of total variables  
selVars   <- paste(Vars,c(rep(1,nv),rep(2,nv)),sep="") #c('bmi1','bmi2')
```

Select Covariates for Analysis

```
twinData[, 'age'] <- twinData[, 'age']/100  
twinData <- twinData[-which(is.na(twinData$age)),]  
covVars   <- 'age'
```

Select Data for Analysis

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

Set Starting Values

```
svMe      <- 20      # start value for means  
svVa      <- .8      # start value for variance  
lbVa      <- .0001   # start value for lower bounds  
svVas     <- diag(svVa,ntv,ntv)  
lbVas     <- diag(lbVa,ntv,ntv)  
laMeMZ    <- c("m1MZ","m2MZ") # labels for means for MZ twins  
laMeDZ    <- c("m1DZ","m2DZ") # labels for means for DZ twins  
laVaMZ    <- c("v1MZ","c21MZ","v2MZ") # labels for (co)variances for MZ twins  
laVaDZ    <- c("v1DZ","c21DZ","v2DZ") # labels for (co)variances for DZ twins
```

My Naming Conventions

| | | |
|---------------------------|----------------------|-------------------------------------|
| name of variable(s) | <code>Vars</code> | <code><- 'bmi'</code> |
| number of variables | <code>nv</code> | <code><- 1</code> |
| number of twin variables | <code>ntv</code> | <code><- nv*2</code> |
| variables per twin pair | <code>selVars</code> | <code><-c('bmi1', 'bmi2')</code> |
| number of factors | <code>nf</code> | <code><- 2</code> |
| number of thresholds | <code>nth</code> | <code><- 3</code> |
| MZ data | <code>mzData</code> | |
| DZ data | <code>dzData</code> | |
| starting values | <code>sv</code> | |
| lower bound / upper bound | <code>lb / ub</code> | |
| labels | <code>la</code> | |

Classical Twin Study Background

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

Classical Twin Study Assumptions

- Equal Environments of MZ and DZ pairs
- Random Mating
- No GE 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

'Old Fashioned' Data Checking

| | MZ | | DZ | |
|------------|-------|-------|-------|-------|
| | T1 | T2 | T1 | T2 |
| mean | 21.35 | 21.34 | 21.45 | 21.46 |
| variance | 0.73 | 0.79 | 0.77 | 0.82 |
| covariance | 0.59 | | 0.25 | |

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

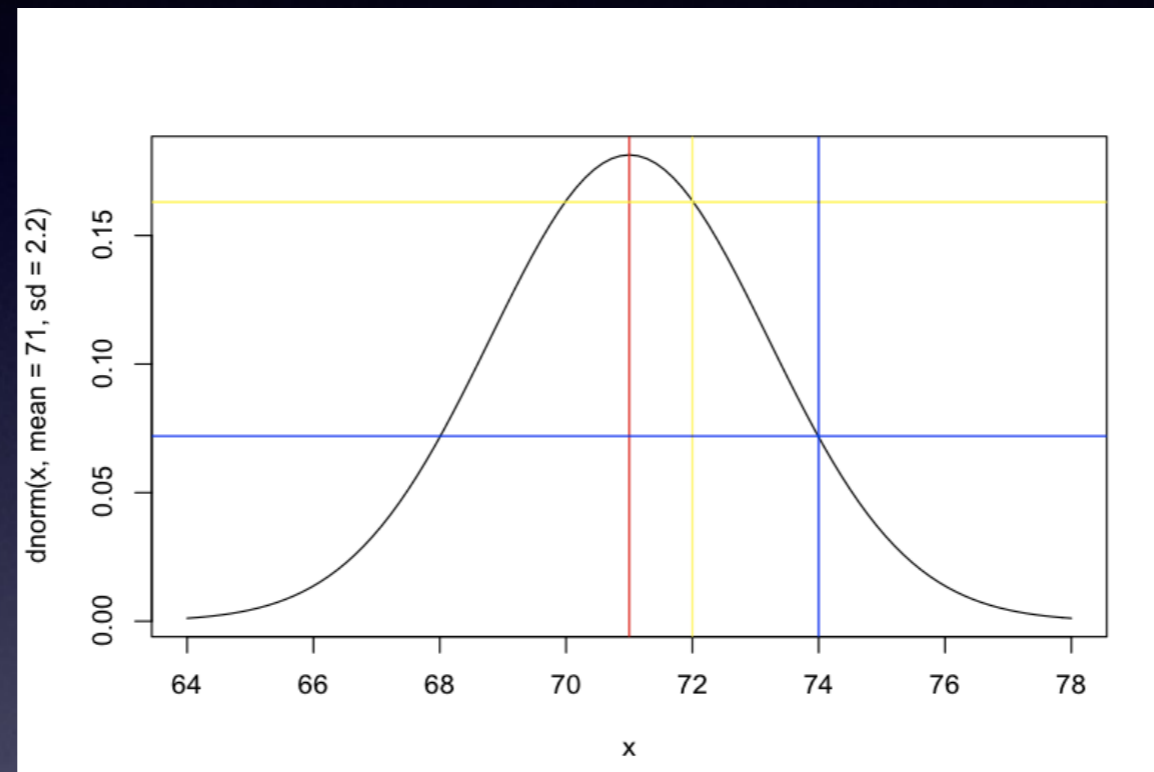
Univariate Analysis

A Roadmap

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

Probability Density Function

$\phi(x_i)$



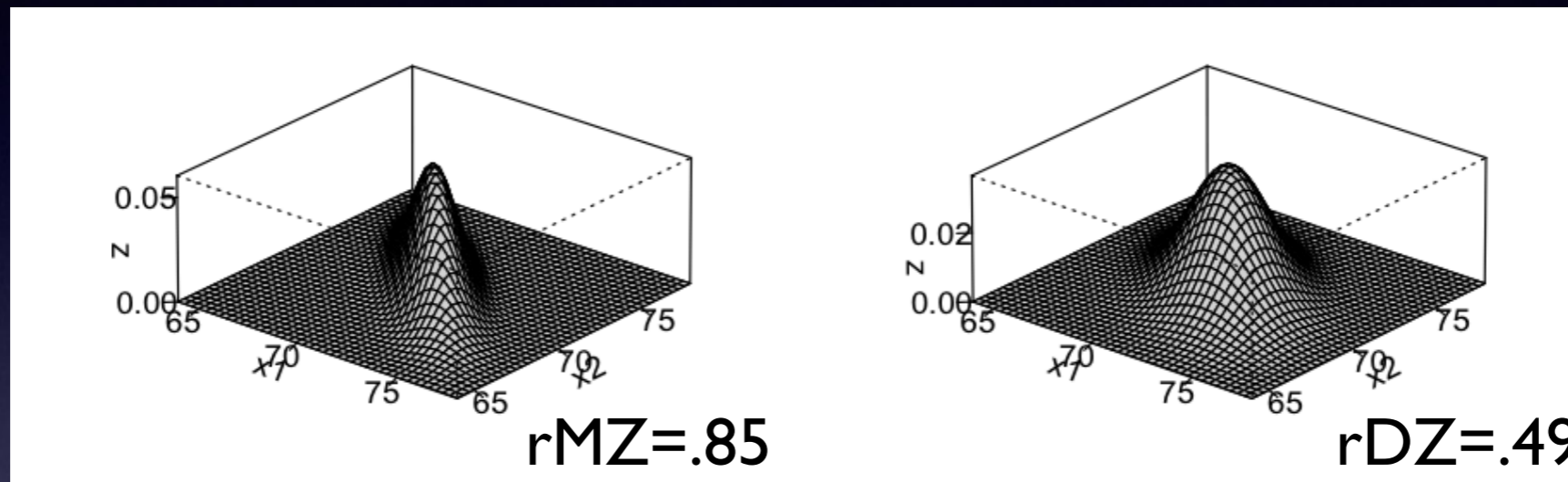
$\phi(x_i)$ is likelihood of data point x_i for particular mean and variance estimates

$$\phi(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-.5((x_i - \mu)^2 / \sigma^2)}$$

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

Univariate: height of probability density function

Multinormal Probability Function



$\Phi(x_i)$ is likelihood of pair of data points x_i and y_i
for particular means, variances and correlation estimates

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

$\pi = 3.14$; x_i : value of variable i ; μ : expected mean; Σ : expected covariance matrix

Multivariate: height of multinormal probability density function

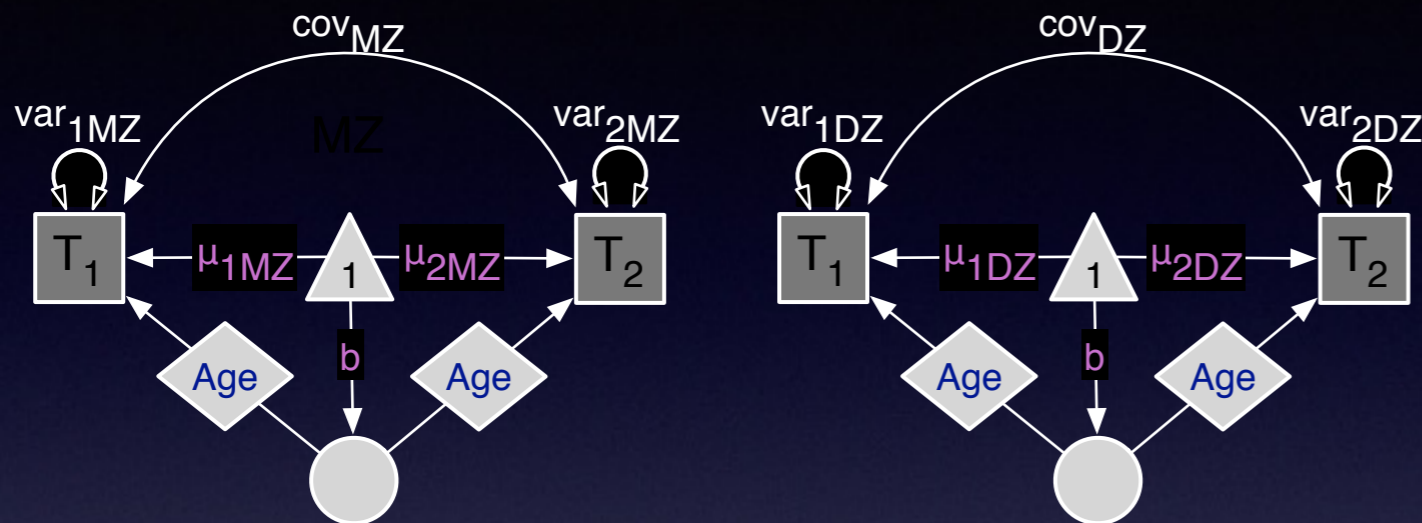
Intuition behind Maximum Likelihood (ML)

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

Likelihood Ratio Test

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

Predicted Means



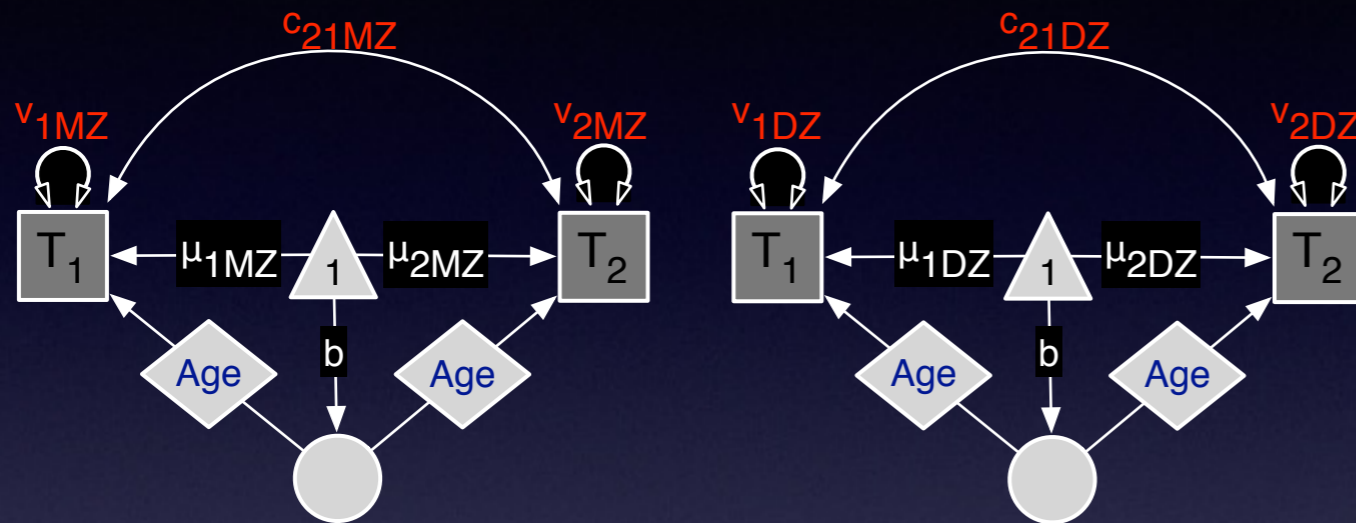
1x2 matrices

| | T1 | T2 |
|--|----------------|----------------|
| | $m1MZ + b*Age$ | $m2MZ + b*Age$ |
| | T1 | T2 |
| | $m1DZ + b*Age$ | $m2DZ + b*Age$ |

```
defAge      <- mxMatrix( type="Full", nrow=1, ncol=1,
  free=FALSE, labels=c("data.age"), name="Age" )
pathB      <- mxMatrix( type="Full", nrow=1, ncol=1,
  free=TRUE, values= .01, label="l11", name="b" )
laMeMZ    <- c("m1MZ", "m2MZ"); laMeDZ    <- c("m1DZ", "m2DZ")
```

```
meanMZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv,
  free=TRUE, values=svMe, labels=laMeMZ, name="MeanMZ")
meanDZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv,
  free=TRUE, values=svMe, labels=laMeDZ, name="MeanDZ")
expMeanMZ  <- mxAlgebra( expression= meanMZ +
  cbind(b%*%Age, b%*%Age), name="expMeanMZ" )
expMeanDZ  <- mxAlgebra( expression= meanDZ +
  cbind(b%*%Age, b%*%Age), name="expMeanDZ" )
```

Predicted Covariances



2x2 matrices

| | T1 | T2 |
|----|-------|-------|
| T1 | v1MZ | c21MZ |
| T2 | c21MZ | v2MZ |

| | T1 | T2 |
|----|-------|-------|
| T1 | v1DZ | c21DZ |
| T2 | c21DZ | v2DZ |

```
laVaMZ <- c("v1MZ", "c21MZ", "v2MZ")
laVaDZ <- c("v1DZ", "c21DZ", "v2DZ")
```

```
covMZ <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv,
  free=TRUE, values=svVas, lbound=lbVas, labels=laVaMZ,
  name="expCovMZ" )
covDZ <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv,
  free=TRUE, values=svVas, lbound=lbVas, labels=laVaDZ,
  name="expCovDZ" )
```

Univariate Twin Saturated

twinSatCon.R [3]

```
# Matrices for Covariates and linear Regression Coefficients
```

```
defAge      <- mxMatrix( type="Full", nrow=1, ncol=1, free=FALSE,
  labels=c("data.age"), name="Age" )
pathB      <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE,
  values=.01, label="l11", name="b" )
```

```
# Algebra for expected Mean Matrices in MZ & DZ twins
```

```
meanMZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe,
  labels=laMeMZ, name="meanMZ" )
meanDZ     <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe,
  labels=laMeDZ, name="meanDZ" )
expMeanMZ  <- mxAlgebra( expression= meanMZ + cbind(b%*%Age,b%*%Age), name="expMeanMZ" )
expMeanDZ  <- mxAlgebra( expression= meanDZ + cbind(b%*%Age,b%*%Age), name="expMeanDZ" )
```

```
# Algebra for expected Variance/Covariance Matrices in MZ & DZ twins
```

```
expCovMZ   <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=TRUE,
  values=svVas, lbound=lbVas, labels=laVaMZ, name="expCovMZ" )
expCovDZ   <- mxMatrix( type="Symm", nrow=ntv, ncol=ntv, free=TRUE,
  values=svVas, lbound=lbVas, labels=laVaDZ, name="expCovDZ" )
```

Univariate Twin Saturated

twinSatCon.R [4]

```
# Data objects for Multiple Groups
```

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

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

```
# Objective objects for Multiple Groups
```

```
objMZ <- mxFIMLObjective( covariance="expCovMZ", means="expMeanMZ",  
  dimnames=selVars )
```

```
objDZ <- mxFIMLObjective( covariance="expCovDZ", means="expMeanDZ",  
  dimnames=selVars )
```

```
# Combine Groups
```

```
modelMZ <- mxModel( "MZ", defAge, pathB, meanMZ, expMeanMZ, expCovMZ, dataMZ,  
  objMZ )
```

```
modelDZ <- mxModel( "DZ", defAge, pathB, meanDZ, expMeanDZ, expCovDZ, dataDZ,  
  objDZ )
```

```
minus2ll <- mxAlgebra( MZ.objective+ DZ.objective, name="minus2sumloglikelihood" )
```

```
obj <- mxFitFunctionAlgebra( "minus2sumloglikelihood" )
```

```
ciCov <- mxCI( c( 'MZ.expCovMZ', 'DZ.expCovDZ' ) )
```

```
ciMean <- mxCI( c( 'MZ.expMeanMZ', 'DZ.expMeanDZ' ) )
```

```
twinSatModel <- mxModel( "twinSat", modelMZ, modelDZ, minus2ll, obj,  
  ciCov, ciMean )
```

Univariate Twin Saturated

twinsatcon.R [5]

```
#-----  
# RUN MODEL  
  
# Run Saturated Model  
twinsatfit      <- mxRun( twinsatmodel, intervals=F )  
twinsatsum      <- summary( twinsatfit )  
twinsatsum  
  
# Generate Saturated Model Output  
twinsatfit$MZ.expMeanMZ@result  
twinsatfit$DZ.expMeanDZ@result  
twinsatfit$MZ.expCovMZ@values  
twinsatfit$DZ.expCovDZ@values  
  
twinsatsum$observedStatistics  
length(twinsatsum$parameters[[1]])  
twinsatfit@output$Minus2LogLikelihood  
twinsatsum$degreesOfFreedom  
twinsatsum$AIC  
round(twinsatfit@output$estimate,4)
```

summary(MxModel)

free parameters:

| | name | matrix | row | col | Estimate | Std.Error | lbound |
|----|-------|-------------|------|------|------------|------------|--------|
| 1 | l11 | MZ.b | 1 | 1 | 2.7536831 | 0.69736577 | |
| 2 | mMZ1 | MZ.meanMZ | 1 | 1 | 20.6886948 | 0.16989840 | |
| 3 | mMZ2 | MZ.meanMZ | 1 | 2 | 20.6934721 | 0.17017616 | |
| 4 | vMZ1 | MZ.expCovMZ | bmi1 | bmi1 | 0.7213571 | 0.04326293 | 1e-04 |
| 5 | cMZ21 | MZ.expCovMZ | bmi1 | bmi2 | 0.5840681 | 0.04034558 | 0 |
| 6 | vMZ2 | MZ.expCovMZ | bmi2 | bmi2 | 0.7842634 | 0.04713262 | 1e-04 |
| 7 | mDZ1 | DZ.meanDZ | 1 | 1 | 20.7833537 | 0.17217692 | |
| 8 | mDZ2 | DZ.meanDZ | 1 | 2 | 20.8080578 | 0.17278599 | |
| 9 | vDZ1 | DZ.expCovDZ | bmi1 | bmi1 | 0.7281361 | 0.05594823 | 1e-04 |
| 10 | cDZ21 | DZ.expCovDZ | bmi1 | bmi2 | 0.2414624 | 0.04378434 | 0 |
| 11 | vDZ2 | DZ.expCovDZ | bmi2 | bmi2 | 0.8030416 | 0.06183128 | 1e-04 |

observed statistics: 1775
estimated parameters: 11
degrees of freedom: 1764
-2 log likelihood: 4015.118
number of observations: 919
Information Criteria AIC: 1837.118

Estimated Values

| | | T1 | T2 | | T1 | T2 |
|-----------------|----|-------|-------|----|-------|-------|
| Saturated Model | | | | | | |
| mean | MZ | 20.68 | 20.69 | DZ | 20.78 | 20.80 |
| cov | T1 | 0.72 | | T1 | 0.73 | |
| | T2 | 0.58 | 0.78 | T2 | 0.24 | 0.80 |

|| parameters Estimated: || |
 m1MZ, m2MZ, v1MZ, v2MZ, c21MZ
 m1DZ, m2DZ, v1DZ, v2DZ, c21DZ

Goodness-of-Fit Statistics

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|-----|----|---------|------|--------|--------------|------------|---|
| Sat | 11 | 4015.12 | 1764 | 487.12 | | | |

| | | | |
|-----------|----------------------------------|--|-----------|
| os | observed statistics | | |
| ep | estimated parameters | | |
| -2ll | -2 LogLikelihood | | |
| df | degrees of freedom | | os - ep |
| AIC | Akaike's Information Criterion | | -2ll -2df |
| diff -2ll | likelihood ratio Chi-square | | |
| diff df | difference in degrees of freedom | | |

Fitting Submodels

```
# Test significance of Covariate

# Copy model, provide new name
testCovModel <- mxModel(twinSatFit, name="testCov")

# Change parameter by changing attributes for label
testCovModel <- omxSetParameters( testCovModel,
  label="l11", free=FALSE, values=0 )

# Fit Nested Model
testCovFit <- mxRun(testCovModel)

# Compare Nested Model with 'Full' Model
mxCompare(twinSatFit, testCovFit)
```

Goodness-of-Fit Stats

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|-----------|----|---------|------|--------|--------------|------------|---|
| Saturated | 11 | 4015.12 | 1764 | 487.12 | | | |
| drop beta | 10 | 4030.57 | 1765 | 500.57 | 15.45 | 1 | 0 |

Fitting Nested Models

- **Saturated Model**
 - likelihood of data without any constraints
 - fitting as many means and (co)variances as possible
- **Equality of means & variances by twin order**
 - test if mean of twin 1 = mean of twin 2
 - test if variance of twin 1 = variance of twin 2
- **Equality of means & variances by zygosity**
 - test if mean of MZ = mean of DZ
 - test if variance of MZ = variance of DZ

Equate Means across twin order

twinSatCon.R [7]

```
# -----  
# RUN SUBMODELS  
  
# Constrain expected Means to be equal across twin order  
eqMeansTwinModel <- mxModel(twinSatFit, name="eqMeansTwin" )  
eqMeansTwinModel <- omxSetParameters( eqMeansTwinModel,  
  label=c("m1MZ", "m2MZ"), free=TRUE, values=svMe, newlabels='mMZ' )  
eqMeansTwinModel <- omxSetParameters( eqMeansTwinModel,  
  label=c("m1DZ", "m2DZ"), free=TRUE, values=svMe, newlabels='mDZ' )  
  
eqMeansTwinFit <- mxRun( eqMeansTwinModel, intervals=F )  
eqMeansTwinSum <- summary( eqMeansTwinFit )  
eqMeansTwinLLL <- eqMeansTwinFit@output$Minus2LogLikelihood  
  
twinSatLLL <- twinSatFit@output$Minus2LogLikelihood  
chi2Sat_eqM <- eqMeansTwinLLL-twinSatLLL  
pSat_eqM <- pchisq( chi2Sat_eqM, lower.tail=F, 2)  
chi2Sat_eqM; pSat_eqM  
mxCompare(twinSatFit, eqMeansTwinFit)
```

Equate Means & Variances across twin order & zygosity

twinSatCon.R [8]

```
# Constrain expected Means and Variances to be equal across twin order
eqMVarsTwinModel    <- mxModel(eqMeansTwinFit, name="eqMVarsTwin" )
eqMVarsTwinModel    <- omxSetParameters( eqMVarsTwinModel,
  label=c("v1MZ","v2MZ"), free=TRUE, values=svMe, newlabels='vMZ' )
eqMVarsTwinModel    <- omxSetParameters( eqMVarsTwinModel,
  label=c("v1DZ","v2DZ"), free=TRUE, values=svMe, newlabels='vDZ' )
eqMVarsTwinFit      <- mxRun( eqMVarsTwinModel, intervals=F )
subs <- list(eqMeansTwinFit, eqMVarsTwinFit)
mxCompare(twinSatFit, subs)

# Constrain expected Means and Variances to be equal across order and zygosity
eqMVarsZygModel     <- mxModel(eqMVarsTwinModel, name="eqMVarsZyg" )
eqMVarsZygModel     <- omxSetParameters( eqMVarsZygModel,
  label=c("mMZ","mDZ"), free=TRUE, values=svMe, newlabels='mZ' )
eqMVarsZygModel     <- omxSetParameters( eqMVarsZygModel,
  label=c("vMZ","vDZ"), free=TRUE, values=svMe, newlabels='vZ' )
eqMVarsZygFit       <- mxRun( eqMVarsZygModel, intervals=F )
mxCompare(eqMVarsTwinFit, eqMVarsZygFit)
```

Estimated Values

| | | T1 | T2 | | T1 | T2 |
|---|----|----|----|----|----|----|
| Equate Means & Variances across Twin Order | | | | | | |
| mean | MZ | | | DZ | | |
| cov | T1 | | | T1 | | |
| | T2 | | | T2 | | |
| Equate Means Variances across Twin Order & Zygosity | | | | | | |
| mean | MZ | | | DZ | | |
| cov | T1 | | | T1 | | |
| | T2 | | | T2 | | |

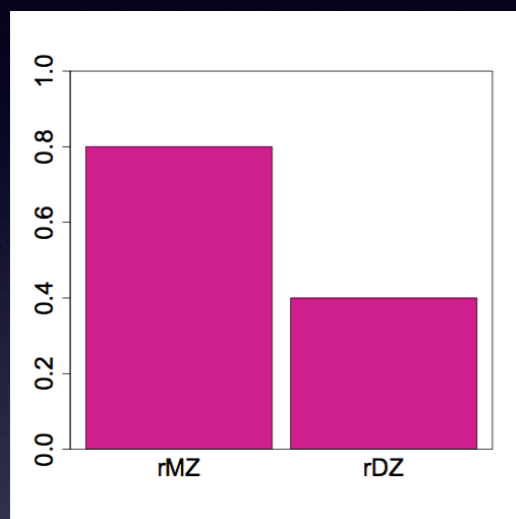
Goodness-of-Fit Stats

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|----------------------------|----|------|----|-----|--------------|------------|---|
| Saturated | | | | | | | |
| $mT1=mT2$ | | | | | | | |
| $mT1=mT2$ $varT1=varT2$ | | | | | | | |
| Zyg MZ=DZ | | | | | | | |

Goodness-of-Fit Stats

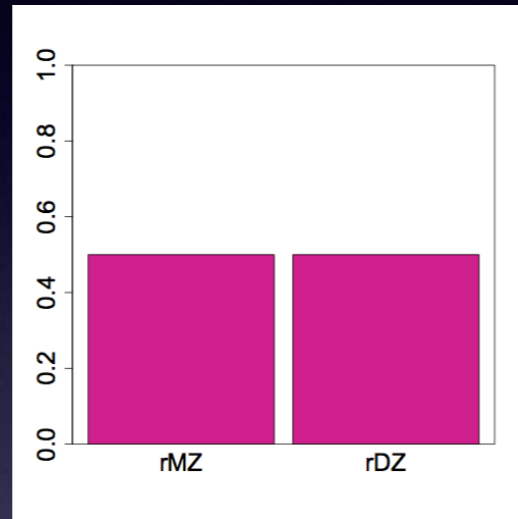
| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|----------------------------|----|---------|------|--------|--------------|------------|------|
| Saturated | 11 | 4015.12 | 1764 | 487.12 | | | |
| $mT1=mT2$ | 9 | 4015.35 | 1766 | 483.35 | 0.23 | 2 | 0.89 |
| $mT1=mT2$ $varT1=varT2$ | 7 | 4018.61 | 1768 | 482.61 | 3.49 | 4 | 0.48 |
| Zyg $MZ=DZ$ | 5 | 4022.79 | 1779 | 482.78 | 7.67 | 6 | 0.26 |

Patterns of Twin Correlations

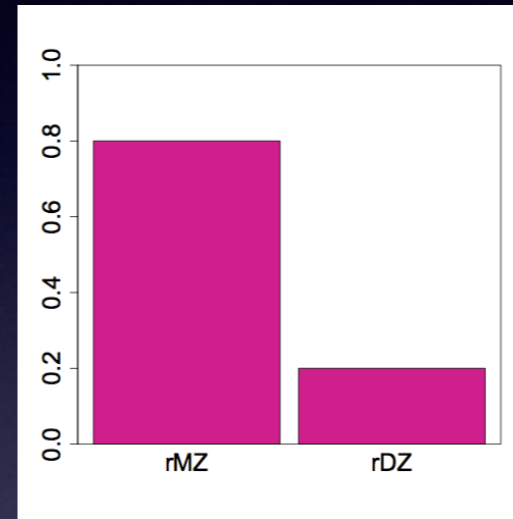


$r_{MZ} = 2r_{DZ}$
Additive

*DZ twins
on average
share 50% of
additive
effects*

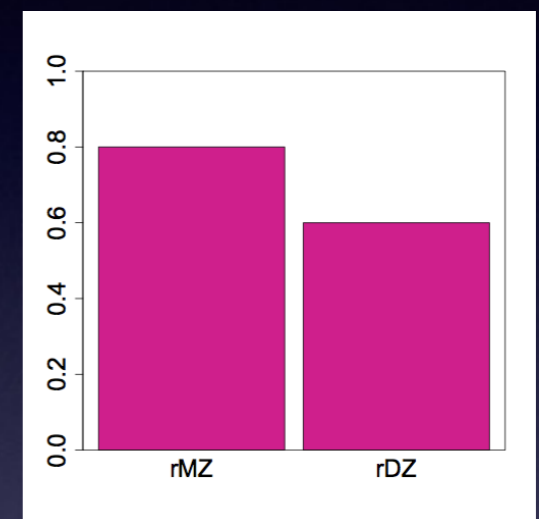


$r_{MZ} = r_{DZ}$
Shared
Environment



$r_{MZ} > 2r_{DZ}$
Additive &
Dominance

*DZ twins
on average share
25% of
dominance
effects*



$r_{DZ} > \frac{1}{2} r_{MZ}$
Additive &
Shared
Environment

$A = 2(r_{MZ} - r_{DZ})$
 $C = 2r_{DZ} - r_{MZ}$
 $E = 1 - r_{MZ}$

Twin Correlations ~ Sources of Variance

$1 - r_{MZ}$

E

$r_{MZ} > r_{DZ}$

A

$r_{MZ} = 2 r_{DZ}$

only A

$r_{MZ} = r_{DZ}$

only C

$r_{MZ} < 1/2 r_{DZ}$

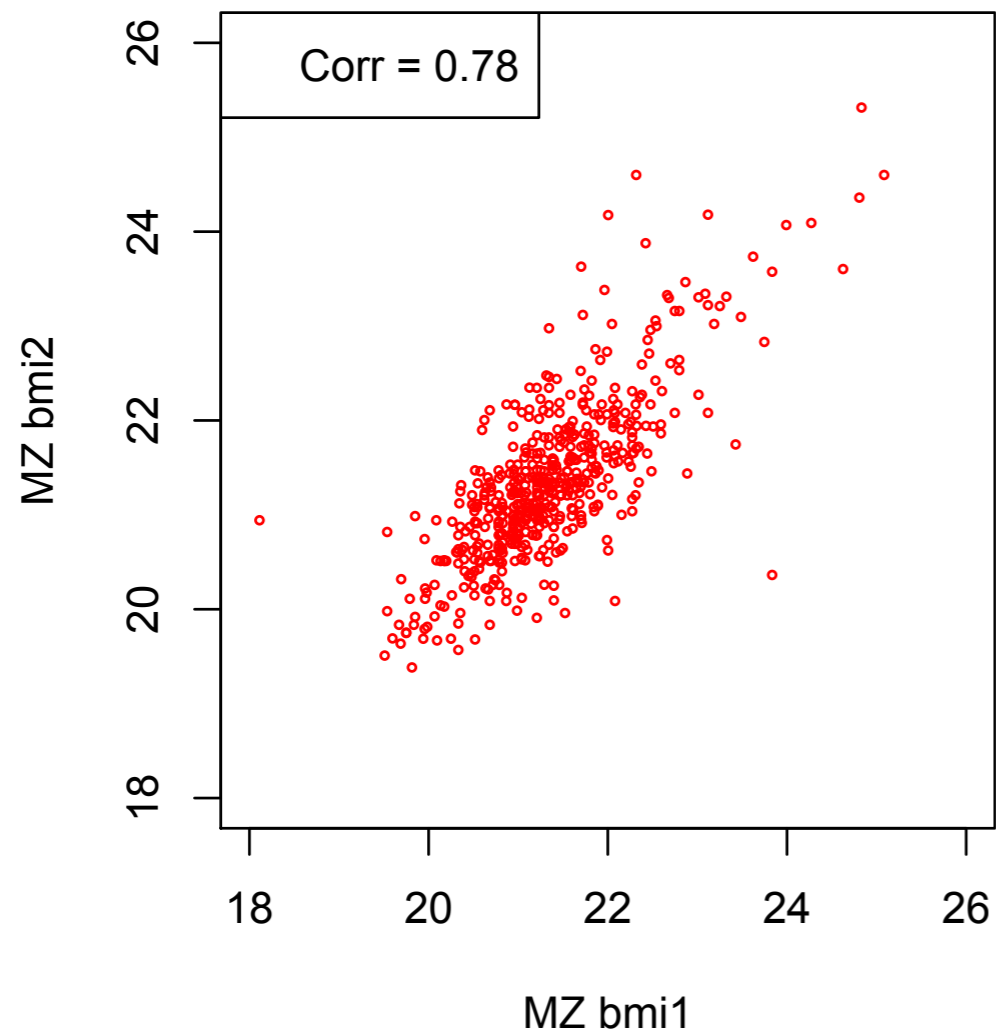
A & C

$r_{MZ} > 1/2 r_{DZ}$

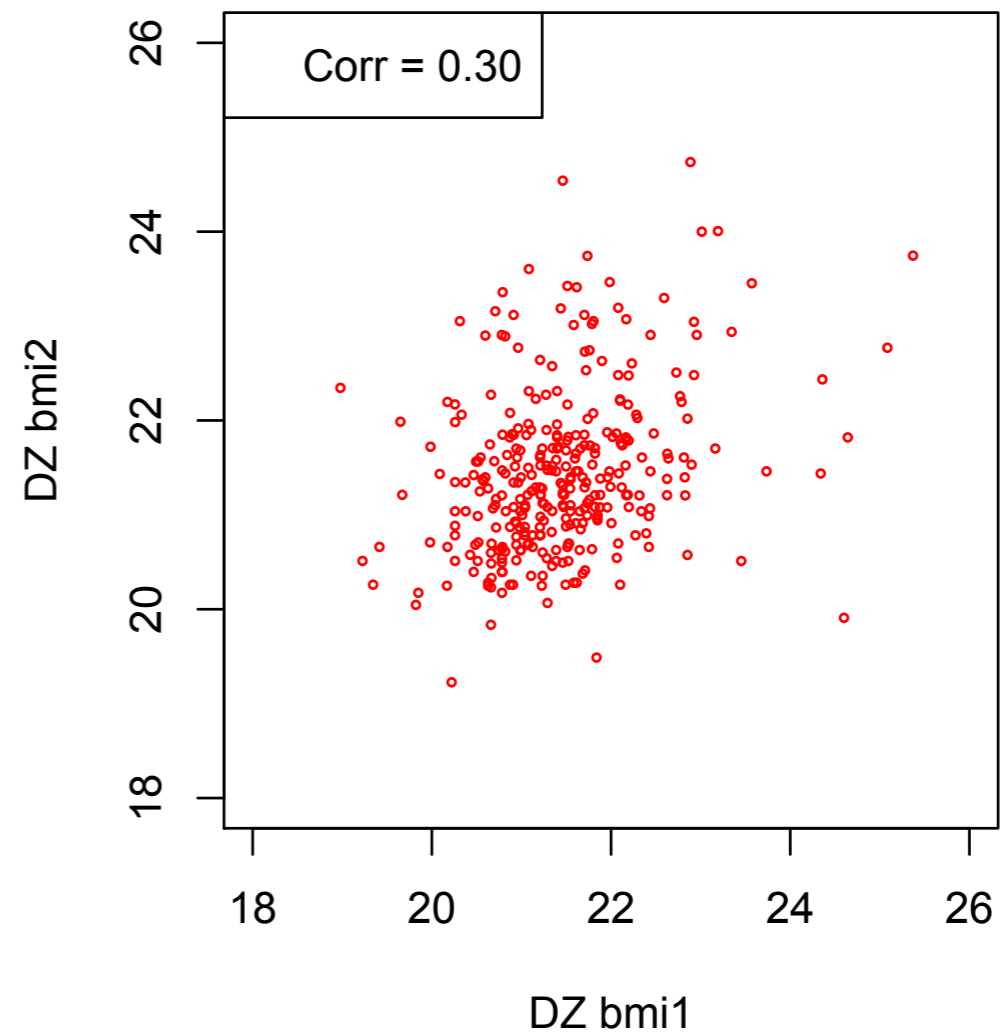
A & D

Twin Correlations

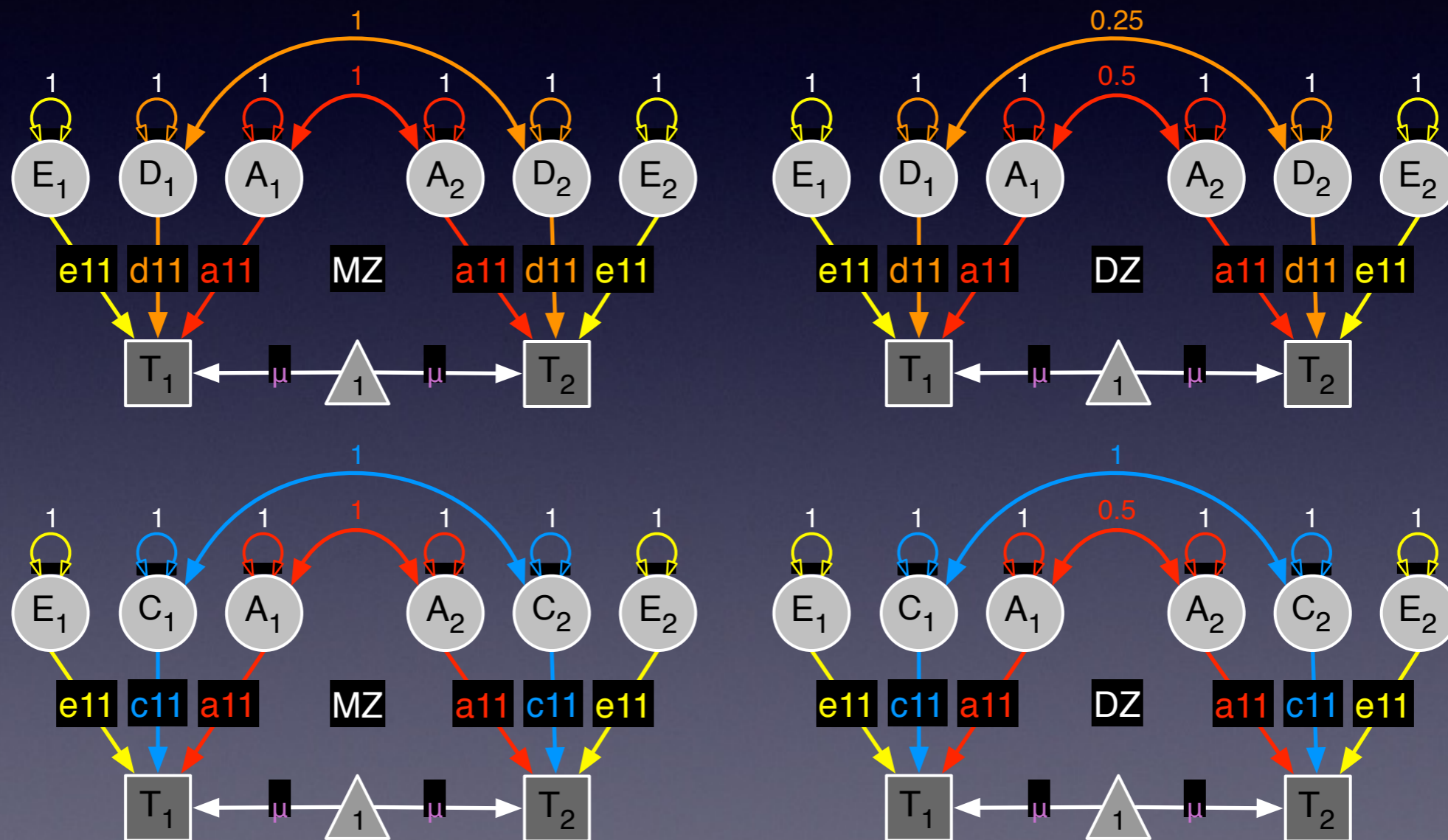
MZ BMI



DZ BMI



Univariate ADE / ACE Model



Univariate Analysis

A Roadmap

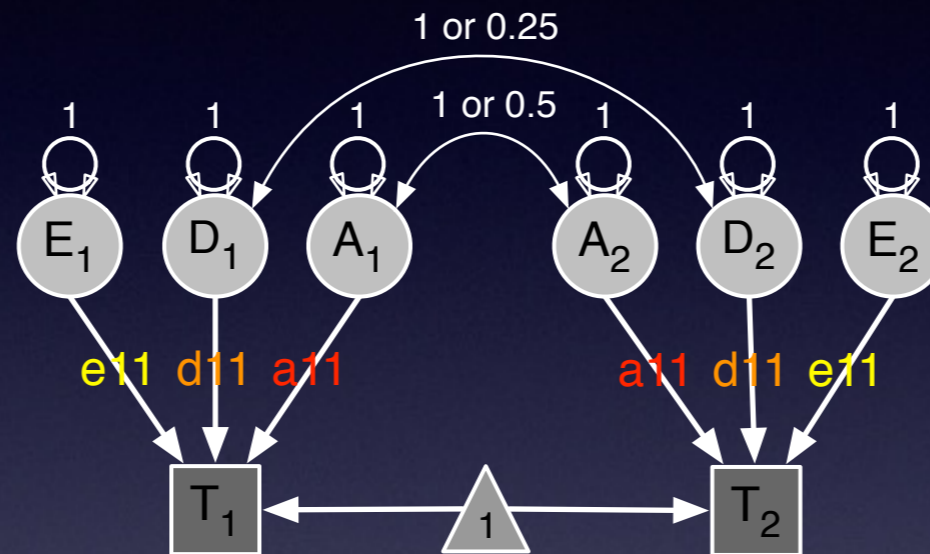
1. Use data to test basic assumptions (equal means & variances for twin 1/twin 2 and MZ/DZ pairs)
 - Saturated Model
2. Estimate contributions of genetic/environmental effects on total variance of a phenotype
 - ACE or ADE Models
3. Test ACE (ADE) submodels to identify and report significant genetic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

Paths & Variance Components

twiAdeCon.R [3]

```
# -----  
# PREPARE MODEL  
  
# Set Starting Values  
svMe      <- 20          # start value for means  
svPa      <- .6          # start value for path coefficients  
            (sqrt(variance/#ofpaths))  
  
# ADE Model  
# Matrices declared to store a, d, and e Path Coefficients  
pathA     <- mxMatrix( type="Full", nrow=nv, ncol=nv, free=TRUE, values=svPa,  
            label="a11", name="a" )  
pathD     <- mxMatrix( type="Full", nrow=nv, ncol=nv, free=TRUE, values=svPa,  
            label="d11", name="d" )  
pathE     <- mxMatrix( type="Full", nrow=nv, ncol=nv, free=TRUE, values=svPa,  
            label="e11", name="e" )  
  
# Matrices generated to hold A, D, and E computed 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" )
```

ADE Deconstructed Path Coefficients



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

all

$a \ 1 \times 1$

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

d11

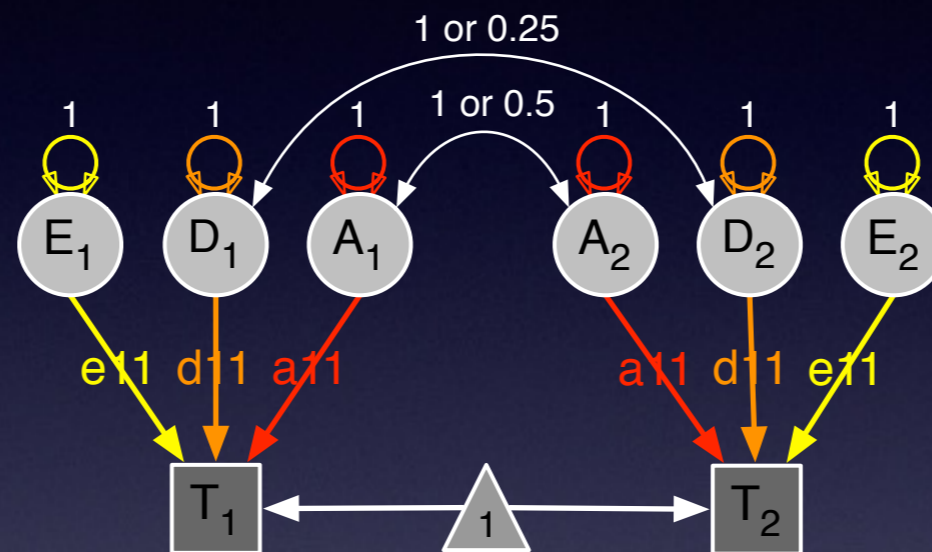
$d \ 1 \times 1$

```
pathE    <- mxMatrix( type="Full", nrow=nv, ncol=nv,
  free=TRUE, values=svPa, label="e11", name="e" )
```

e11

$e \ 1 \times 1$

ADE Deconstructed Variance Components



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

$$\begin{matrix} \boxed{a_{11}} & * & \boxed{a_{11}^T} \\ A & | & x & | \end{matrix}$$

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

$$\begin{matrix} \boxed{d_{11}} & * & \boxed{d_{11}^T} \\ D & | & x & | \end{matrix}$$

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

$$\begin{matrix} \boxed{e_{11}} & * & \boxed{e_{11}^T} \\ E & | & x & | \end{matrix}$$

Expected Means & (Co)Variances

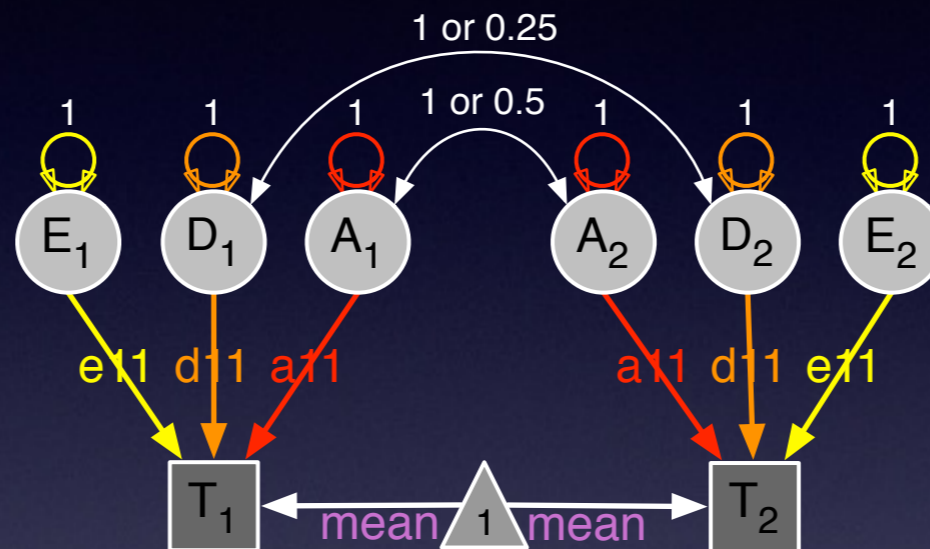
twinAdeCon.R [4]

```
# Matrices for covariates and linear regression coefficients
defAge      <- mxMatrix( type="Full", nrow=1, ncol=1, free=FALSE,
  labels=c("data.age"), name="Age" )
pathB       <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE,
  values= .01, label="l11", name="b" )

# Algebra for expected Mean Matrices in MZ & DZ twins
meanG       <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE, values=svMe,
  labels="xbmi", name="mean" )
expMean     <- mxAlgebra( expression= mean + cbind(b%*%Age,b%*%Age),
  name="expMean" )

# Algebra for expected Variance/Covariance Matrices in MZ & DZ twins
covP        <- mxAlgebra( expression= A+D+E, name="V" )
expCovMZ    <- mxAlgebra( expression= rbind( cbind(V, A+D), cbind(A+D, V)),
  name="expCovMZ" )
expCovDZ    <- mxAlgebra( expression= rbind( cbind(V, 0.5*x%A+ 0.25*x%D),
  cbind(0.5*x%A+ 0.25*x%D , V)), name="expCovDZ" )
```

ADE Deconstructed Means & Variances



| | |
|------------------------|------|
| mean | mean |
| $expMean_{1 \times 2}$ | |

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

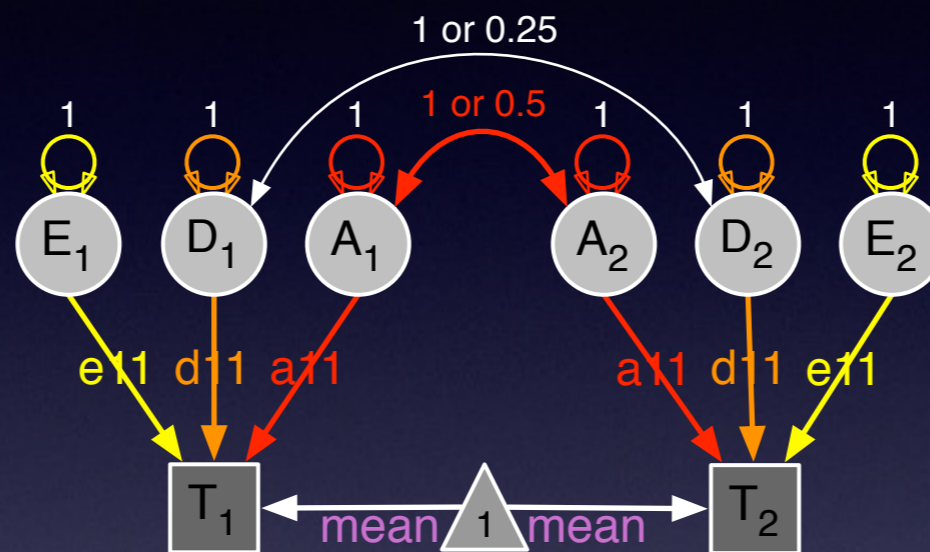
```
covMZ <- mxAlgebra( expression= rbind(
cbind(V, A+D), cbind(A+D, V)), name="expCovMZ")
```

```
covDZ <- mxAlgebra( expression= rbind(
cbind(V, 0.5*x%A+ 0.25*x%D),
cbind(0.5*x%A+ 0.25*x%D, V)), name="expCovDZ")
```

| | |
|-------------------------|-----|
| V | |
| | V |
| $expCovMZ_{2 \times 2}$ | |

$expCovDZ_{2 \times 2}$

ADE Deconstructed A Covariances



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

```
covMZ <- mxAlgebra( expression= rbind(
  cbind(V, A+D), cbind(A+D, V)), name="expCovMZ")
```

```
covDZ <- mxAlgebra( expression= rbind(
  cbind(V, 0.5*x%A+ 0.25*x%D),
  cbind(0.5*x%A+ 0.25*x%D, V)), name="expCovDZ")
```

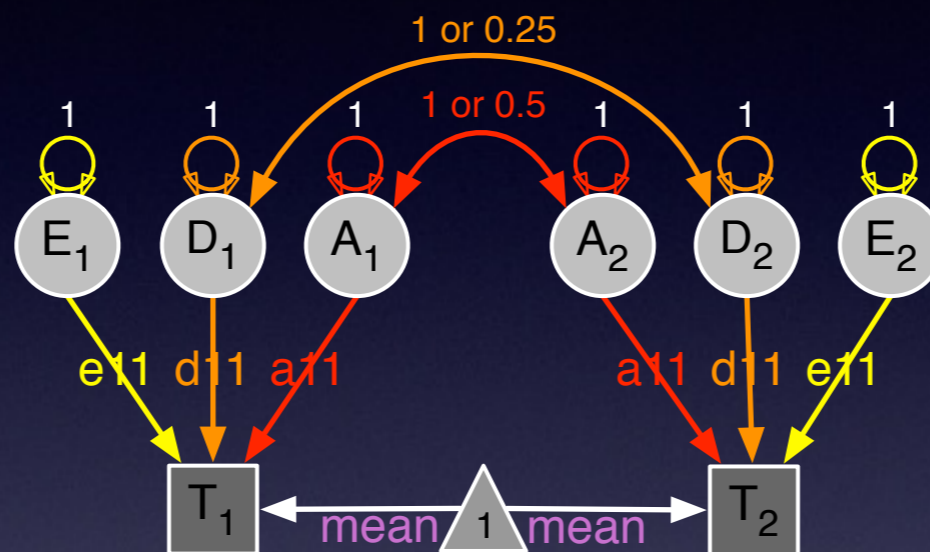
| | |
|-----|-----|
| V | A+D |
| A+D | V |

expCovMZ 2×2

| | |
|----------|----------|
| V | .5A+.25D |
| .5A+.25D | V |

expCovDZ 2×2

ADE Deconstructed D Covariances



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

```
covMZ <- mxAlgebra( expression= rbind(
  cbind(V, A+D), cbind(A+D, V)), name="expCovMZ")
```

```
covDZ <- mxAlgebra( expression= rbind(
  cbind(V, 0.5*x%A+ 0.25*x%D),
  cbind(0.5*x%A+ 0.25*x%D, V)), name="expCovDZ")
```

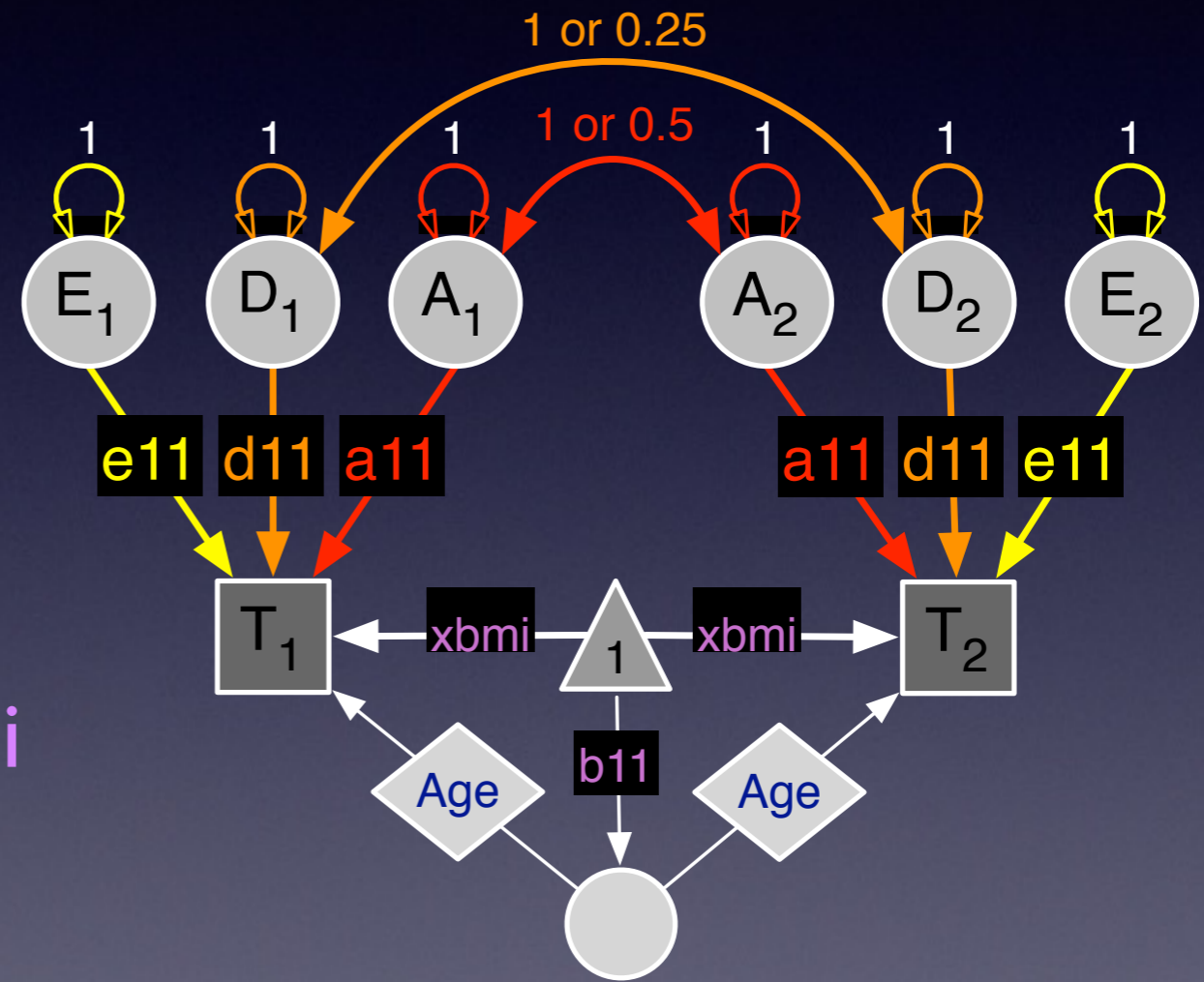
| | |
|-----|-----|
| V | A+D |
| A+D | V |

$expCovMZ_{2 \times 2}$

| | |
|----------|----------|
| V | .5A+.25D |
| .5A+.25D | V |

$expCovDZ_{2 \times 2}$

ADE Deconstructed Parameters



- 5 Parameters Estimated:
- Mean $xbmi$
 - Regression on Age $b11$
 - Variance due to A $a11$
 - Variance due to D $d11$
 - Variance due to E $e11$

Data & Objectives

twinAdeCon.R [5]

Data objects for Multiple Groups

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

Objective objects for Multiple Groups

```
objMZ     <- mxFIMLObjective( covariance="expCovMZ", means="expMean",
  dimnames=selVars )
objDZ     <- mxFIMLObjective( covariance="expCovDZ", means="expMean",
  dimnames=selVars )
```

Combine Groups

```
pars      <- list( pathA, pathD, pathE, covA, covD, covE, covP, pathB )
modelMZ   <- mxModel( pars, defAge, meanG, expMean, expCovMZ, dataMZ, objMZ,
  name="MZ" )
modelDZ   <- mxModel( pars, defAge, meanG, expMean, expCovDZ, dataDZ, objDZ,
  name="DZ" )
minus2ll  <- mxAlgebra( expression=MZ.objective + DZ.objective, name="m2LL" )
obj       <- mxAlgebraObjective( "m2LL" )
AdeModel  <- mxModel( "ADE", pars, modelMZ, modelDZ, minus2ll, obj )
```

Model Fitting

twinAdeCon.R [6]

```
# RUN MODEL
```

```
# Run ADE model
```

```
AdeFit <- mxRun(AdeModel, intervals=T)
```

```
AdeSumm <- summary(AdeFit)
```

```
AdeSumm
```

```
mxCompare(twinSatFit,AdeFit)
```

```
round(AdeFit@output$estimate,4)
```

```
round(AdeFit$Vars@result,4)
```

```
# Generate Table of Parameter Estimates using mxEval
```

```
pathEstimatesADE <- print(round(mxEval(cbind(a,d,e), AdeFit),4))
```

```
varComponentsADE <- print(round(mxEval(cbind(A/V,D/V,E/V), AdeFit),4))
```

```
rownames(pathEstimatesADE) <- 'pathEstimates'
```

```
colnames(pathEstimatesADE) <- c('a','d','e')
```

```
rownames(varComponentsADE) <- 'varComponents'
```

```
colnames(varComponentsADE) <- c('a^2','d^2','e^2')
```

```
pathEstimatesADE
```

```
varComponentsADE
```

Generating Output

twinAdeCon.R [7]

```
# Generate ADE Model Output
estMean    <- mxEval(expMean, AdeFit$MZ)    # expected mean
estCovMZ   <- mxEval(expCovMZ, AdeFit$MZ)   # expected covariance matrix for MZ's
estCovDZ   <- mxEval(expCovDZ, AdeFit$DZ)   # expected covariance matrix for DZ's
estVA      <- mxEval(a*a, AdeFit)          # additive genetic variance, a^2
estVD      <- mxEval(d*d, AdeFit)          # dominance variance, d^2
estVE      <- mxEval(e*e, AdeFit)          # unique environmental variance, e^2
estVP      <- (estVA+estVD+estVE)          # total variance
estPropVA  <- estVA/estVP                  # standardized additive genetic variance
estPropVD  <- estVD/estVP                  # standardized dominance variance
estPropVE  <- estVE/estVP                  # standardized unique environmental var
estADE     <- rbind(cbind(estVA,estVD,estVE), # table of estimates
                   cbind(estPropVA,estPropVD,estPropVE))
LL_ADE    <- mxEval(objective, AdeFit)      # likelihood of ADE model
```


summary(mxModel)

free parameters:

| | name | matrix | row | col | Estimate | Std.Error | lbound | ubound |
|---|------|---------|-----|-----|------------|-----------|--------|--------|
| 1 | a11 | a | 1 | 1 | 0.6060900 | NaN | | |
| 2 | d11 | d | 1 | 1 | 0.4743898 | NaN | | |
| 3 | e11 | e | 1 | 1 | 0.4111268 | NaN | | |
| 4 | l11 | b | 1 | 1 | 2.7677606 | NaN | | |
| 5 | xbmi | MZ.mean | 1 | 1 | 20.7346094 | NaN | | |

confidence intervals:

| | | lbound | estimate | ubound |
|---------------|--------------|-----------|-----------|--------|
| ADE.Vars[1,1] | 6.811043e-02 | 0.3673451 | 0.6333987 | |
| ADE.Vars[1,2] | 5.966440e-21 | 0.2250457 | 0.5243742 | |
| ADE.Vars[1,3] | 1.502990e-01 | 0.1690253 | 0.1909726 | |
| ADE.Vars[1,4] | 8.988222e-02 | 0.4824499 | 0.7934488 | |
| ADE.Vars[1,5] | 1.415784e-13 | 0.2955620 | 0.6894092 | |
| ADE.Vars[1,6] | 1.937827e-01 | 0.2219881 | 0.2545650 | |

observed statistics: 1775

estimated parameters: 5

degrees of freedom: 1770

-2 log likelihood: 4022.789

number of observations: 919

Information Criteria AIC: 1832.789

Goodness-of-Fit Stats

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|-----------|----|---------|------|--------|--------------|------------|------|
| Saturated | 11 | 4015.12 | 1764 | 487.12 | | | |
| ADE | 5 | 4022.79 | 1770 | 482.79 | 7.67 | 6 | 0.26 |

Table of Estimates

```
> # Generate Table of Parameter Estimates using mxEval
> pathEstimatesADE
              a          d          e
pathEstimates 0.6061 0.4744 0.4111
> varComponentsADE
              a^2          d^2          e^2
varComponents 0.4824 0.2956 0.222
```

Univariate Analysis

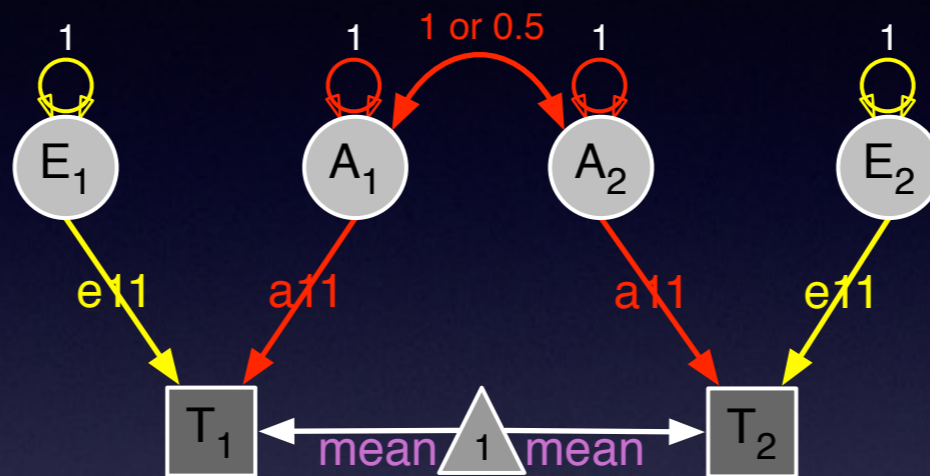
A Roadmap

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

Nested Models

- 'Full' ADE Model
- Nested Models
 - AE Model
 - test significance of D
 - E Model vs AE Model
 - test significance of A
 - E Model vs ADE Model
 - test combined significance of A & D

AE Deconstructed *Parameters*



```
# Test significance of D
```

```
# Copy model, provide new name
```

```
AeModel <- mxModel(AdeFit, name="AE")
```

```
# Change parameter by changing attributes for label
```

```
AeModel <- omxSetParameters( AeModel,  
  label="d11", free=FALSE, values=0 )
```

```
# Fit Nested Model
```

```
AeFit <- mxRun(AeModel)
```

```
# Compare Nested Model with 'Full' Model
```

```
mxCompare(AdeFit, AeFit)
```

AE Model

twinAdeCon.R [8]

```
# -----  
# FIT SUBMODELS  
  
# Run AE model  
AeModel <- mxModel( AdeFit, name="AE" )  
AeModel <- omxSetParameters( AeModel, labels="d11", free=FALSE, values=0 )  
AeFit <- mxRun(AeModel)  
mxCompare(AdeFit, AeFit)  
round(AeFit@output$estimate,4)  
round(AeFit$Vars@result,4)  
  
# Run E model  
eModel <- mxModel( AeFit, name="E" )  
eModel <- omxSetParameters( eModel, labels="a11", free=FALSE, values=0 )  
eFit <- mxRun(eModel)  
mxCompare(AeFit, eFit)  
round(eFit@output$estimate,4)  
round(eFit$Vars@result,4)  
  
# Print Comparative Fit Statistics  
AdeNested <- list(AeFit, eFit)  
mxCompare(AdeFit,AdeNested)  
round(rbind(AdeFit$Vars@result,AeFit$Vars@result,eFit$Vars@result),4)
```

Goodness-of-Fit Statistics

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|-----|----|------|----|-----|--------------|------------|---|
| ADE | | | | | | | |
| AE | | | | | | | |
| E | | | | | | | |

Goodness-of-Fit Statistics

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|-----|----|---------|------|---------|--------------|------------|------|
| ADE | 5 | 4022.79 | 1770 | 482.79 | | | |
| AE | 4 | 4025.41 | 1771 | 483.41 | 2.62 | 1 | 0.10 |
| E | 3 | 4549.61 | 1772 | 1005.61 | 526.8 | 2 | 0 |

Estimated Values

| | path coefficients | | | unstandardized variance components | | | standardized variance components | | |
|-----|-------------------|------|------|------------------------------------|----------------|----------------|----------------------------------|----------------|----------------|
| | a | d | e | a ² | d ² | e ² | a ² | d ² | e ² |
| ADE | 0.61 | 0.47 | 0.41 | 0.37 | 0.25 | 0.17 | 0.48 | 0.30 | 0.22 |
| AE | 0.77 | - | 0.41 | 0.60 | - | 0.17 | 0.77 | - | 0.22 |
| E | - | - | 0.87 | - | - | 0.76 | - | - | 1.00 |

What about C?

- 'Full' ACE Model
- Nested Models
 - AE Model
 - test significance of C
 - CE Model
 - test significance of A
 - E Model vs AE Model
 - test significance of A
 - E Model vs ACE Model
 - test combined significance of A & C

Goodness-of-Fit Statistics

| | ep | -2ll | df | AIC | diff -2ll | diff df | p |
|-----|----|------|----|-----|--------------|------------|---|
| ADE | | | | | | | |
| AE | | | | | | | |
| ACE | | | | | | | |
| CE | | | | | | | |
| E | | | | | | | |

Estimated Values

| | a | d | e | c | a^2 | d^2 | e^2 | c^2 |
|-----|---|---|---|---|-------|-------|-------|-------|
| ADE | | | | | | | | |
| AE | | | | | | | | |
| ACE | | | | | | | | |
| AE | | | | | | | | |
| E | | | | | | | | |

Conclusions

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