STATISTICS
TEXAS A&M UNIVERSITY

Handout 09

Mixed Models Analyses

Mixed Models with Covariates

Objectives of Analysis of Covariance with Random Effects

– Perform an analysis of covariance using the MIXED procedure.

– Interpret the parameter estimates from an analysis of covariance.

– Compute adjusted means.

Analysis of Covariance

Continuous Response

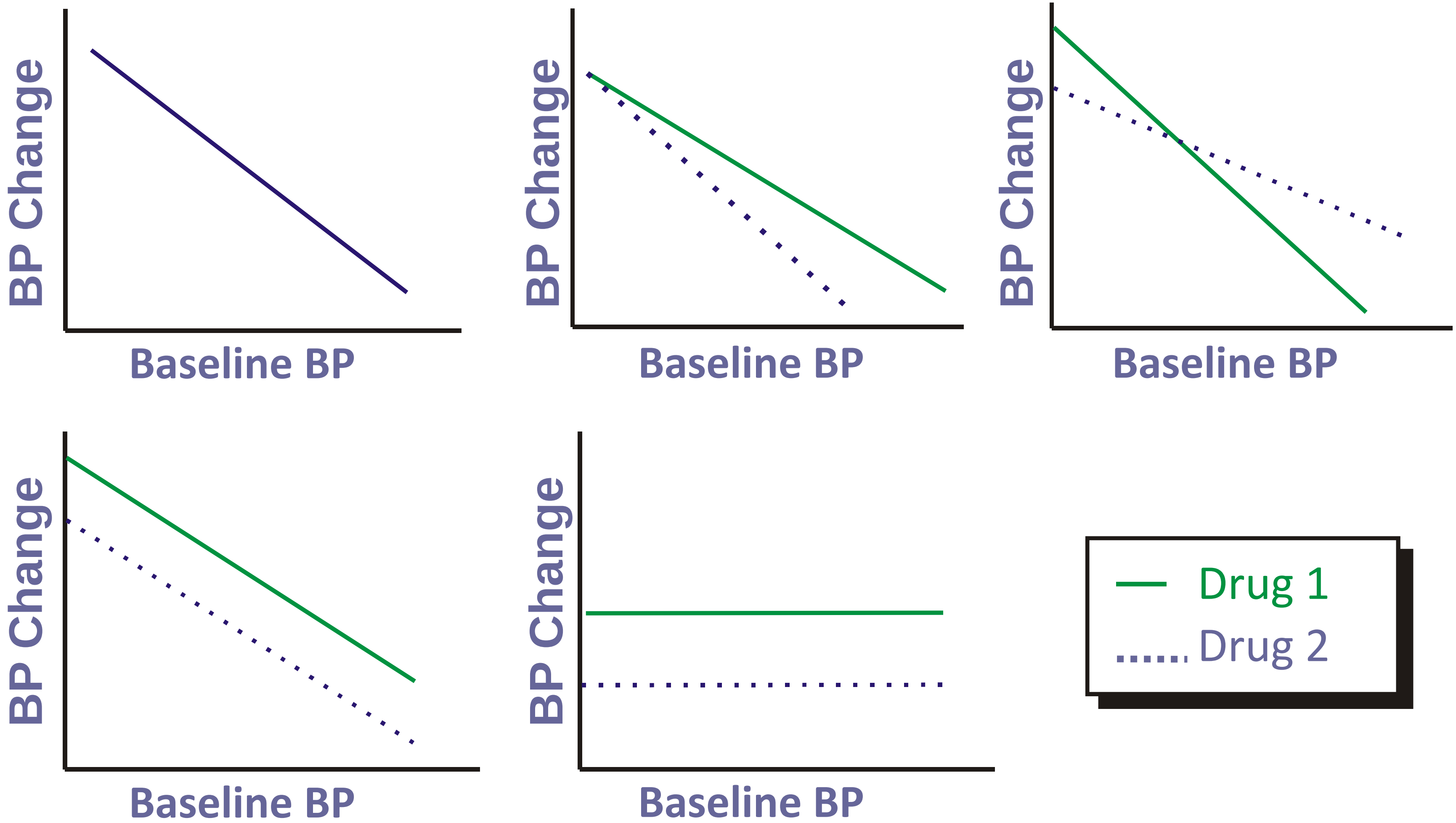
Discrete and Continuous Explanatory

Blood Pressure Change

Drug

Baseline BP

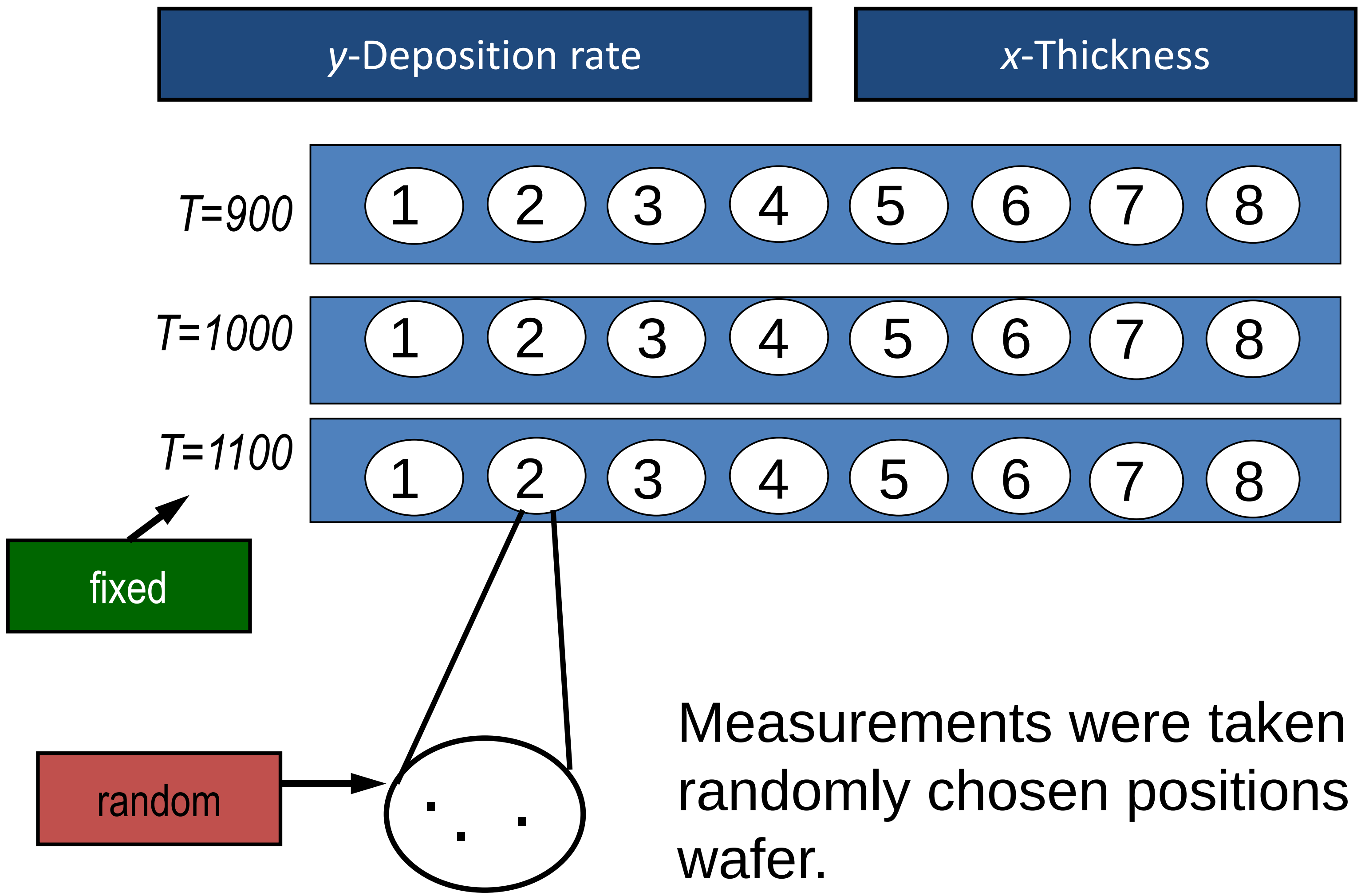
Possible Scenarios



Wafer Example

- The effect of temperature on the deposition rate of a layer of polysilicon in the fabrication of wafers is the interest.
- It was thought that the wafer thickness before the deposition process was applied might have an effect on the deposition rate.
- Therefore, the average thickness of each wafer (thick) was determined and used as a possible covariate.
- A random sample of 24 wafers was collected and used in the experiment.
- Wafers were randomly assigned to one of 3 levels of temperature (990, 1000, 1100). So each level of temperature had 8 wafers assigned.
- The amount of deposited material at 3 randomly chose sites from each wafer was measured.

Wafer Example



Measurements were taken at three randomly chosen positions on each wafer.

The Data

temp	wafer	site	deposit	thick
900	1	1	291	1919
900	1	2	295	1919
900	1	3	294	1919
900	2	1	318	2113
900	2	2	315	2113
900	2	3	315	2113
900	3	1	306	1841
900	3	2	302	1841
900	3	3	305	1841
900	4	1	342	2170
900	4	2	341	2170
900	4	3	336	2170
900	5	1	318	2019
900	5	2	323	2019
900	5	3	323	2019
900	6	1	307	1872
900	6	2	308	1872
900	6	3	308	1872
900	7	1	295	1862
900	7	2	297	1862
...	...	...	...	...

The ANCOVA Model

overall slope

slope effect of **`temp`** *i*

$$y_{ijk} = \beta_0 + \alpha_i + \beta_1 x_{ij} + \delta_i x_{ij} + w_{j(i)} + e_{ijk}$$

overall intercept

intercept effect of **`temp`** *i*

`Wafer` effect, random

i = 1, 2, 3 (temp)

j = 1 to 8 (wafer)

k = 1, 2, 3 (site)

$$w_{j(i)} \sim N(0, \sigma_w^2)$$

$$e_{ijk} \sim N(0, \sigma^2)$$

Wafer4Example.sas

Question

In the example, which of the following is **false** regarding the interaction of **`temp*thick`** in the model?

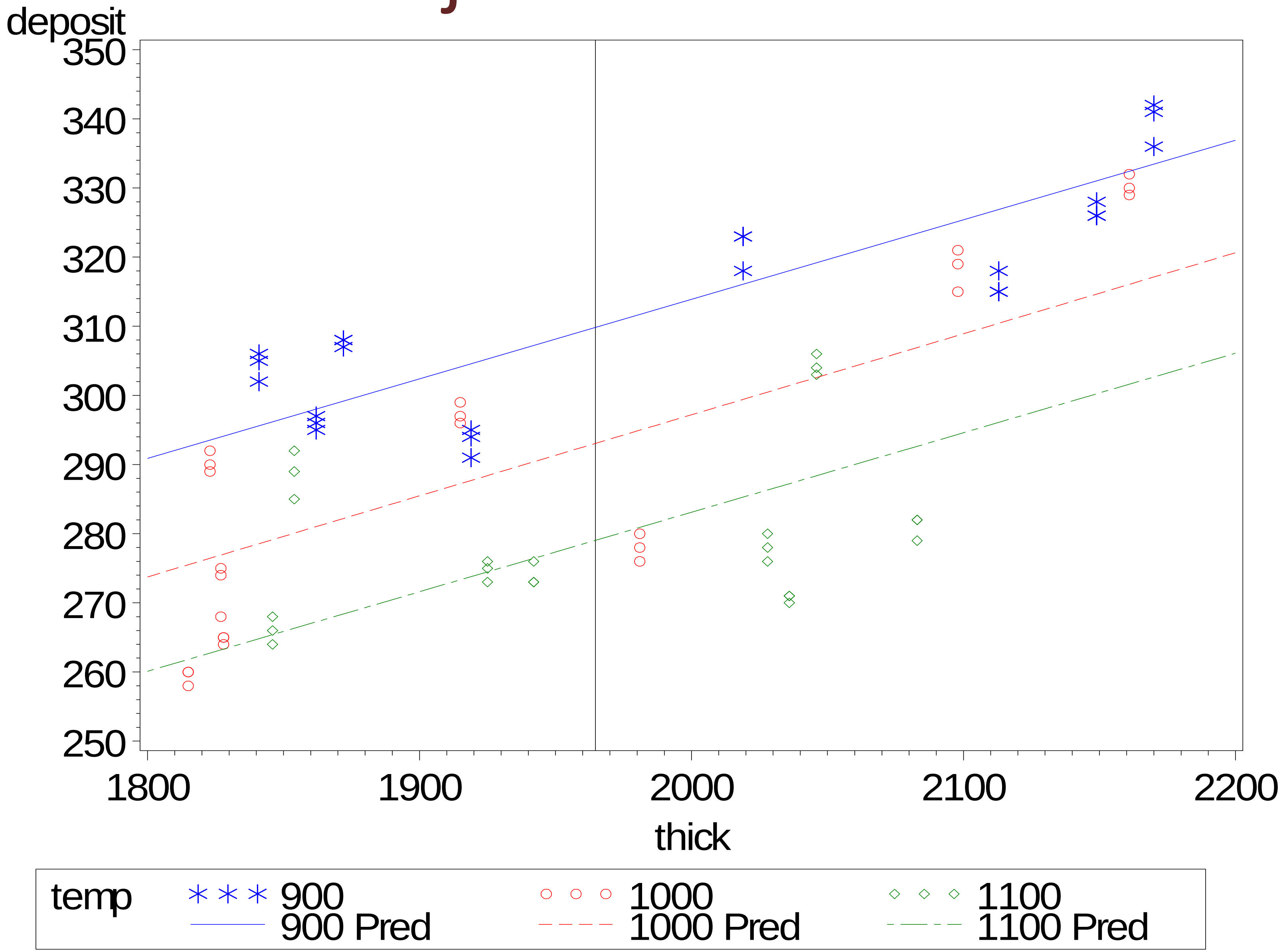
a. The **`temp*thick`** term fits an unequal slope model.

b. The **`temp*thick`** term tests whether the slopes are equal to each other.

c. The **`temp*thick`** term tests whether the slopes are equal to zero.

d. When the **`temp*thick`** term is not significant, you can remove the term and fit a common slope model.

LSMEANS Adjusted for the Covariate



Computing the Least Squares Means

Which of the following is true?

a. You cannot use the LSMEANS statement for treatment effects when your model includes covariates.

b. In ANCOVA models, the least squares means are computed accounting for the covariates.

c. Least squares means are the same as the arithmetic means.

This demonstration illustrates the concepts discussed previously.

Wafer4Example.sas

An Alternative Formulation

slope for **`temp`** *i*

$$y_{ijk} = \beta_{0i} + \beta_{1i} x_{ij} + w_{j(i)} + e_{ijk}$$

intercept for **`temp`** *i*

`wafer` effect, random

$$\beta_{0i} = \beta_0 + \alpha_i, \quad \beta_{1i} = \beta_1 + \delta_i$$

i = 1, 2, 3 (temp)

j = 1 to 8 (wafer)

k = 1, 2, 3 (site)

$$w_{j(i)} \sim N(0, \sigma_w^2)$$
$$e_{ijk} \sim N(0, \sigma^2)$$

Wafer4Example.sas

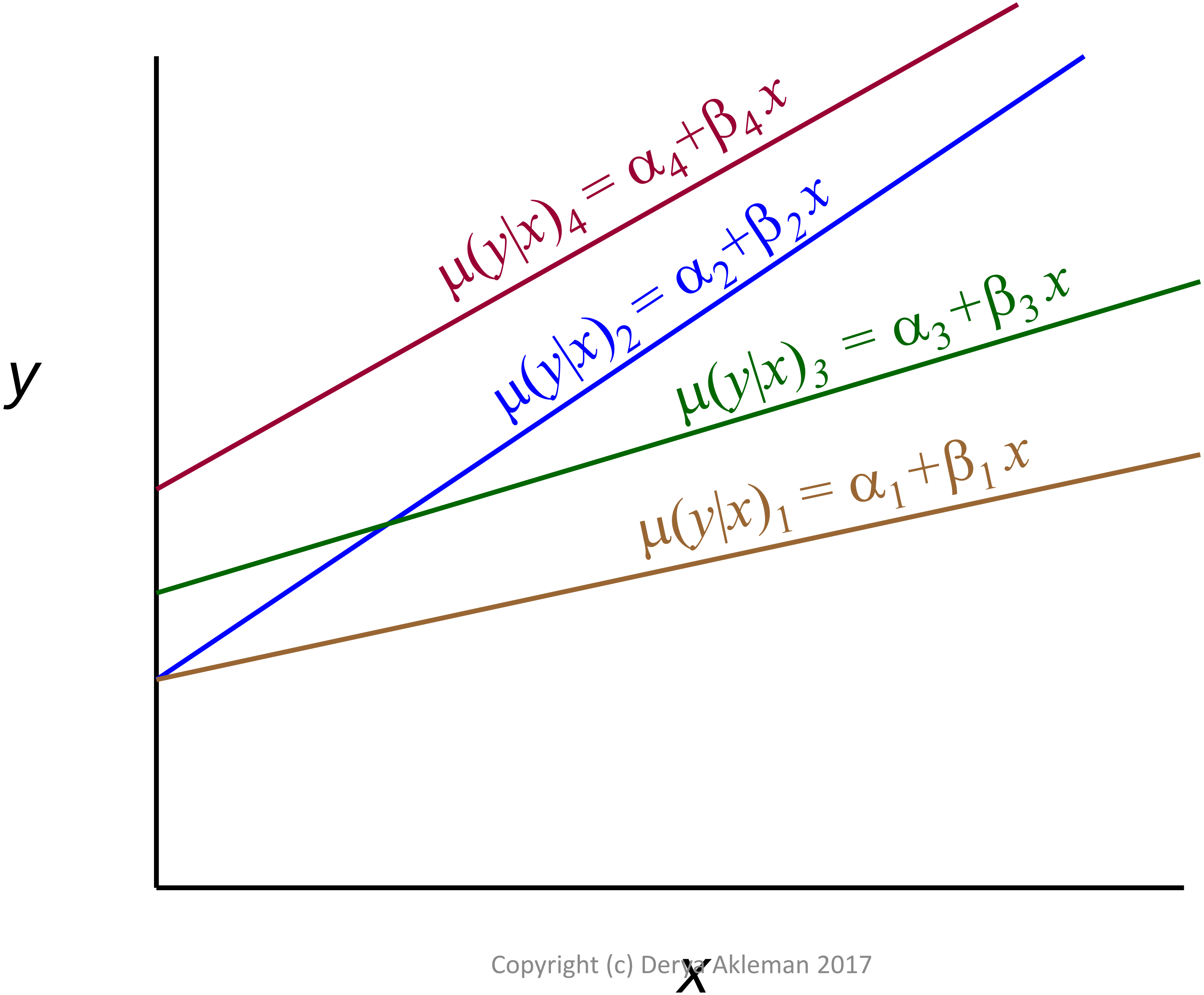
Random Coefficient Models

- In the analysis of covariance, the regression coefficients for the covariates are assumed to be fixed effects, that is, unknown fixed parameters estimated from data.

- In the random coefficient model, the regression coefficients for one or more covariates are assumed to be a random sample from some population of possible coefficients.

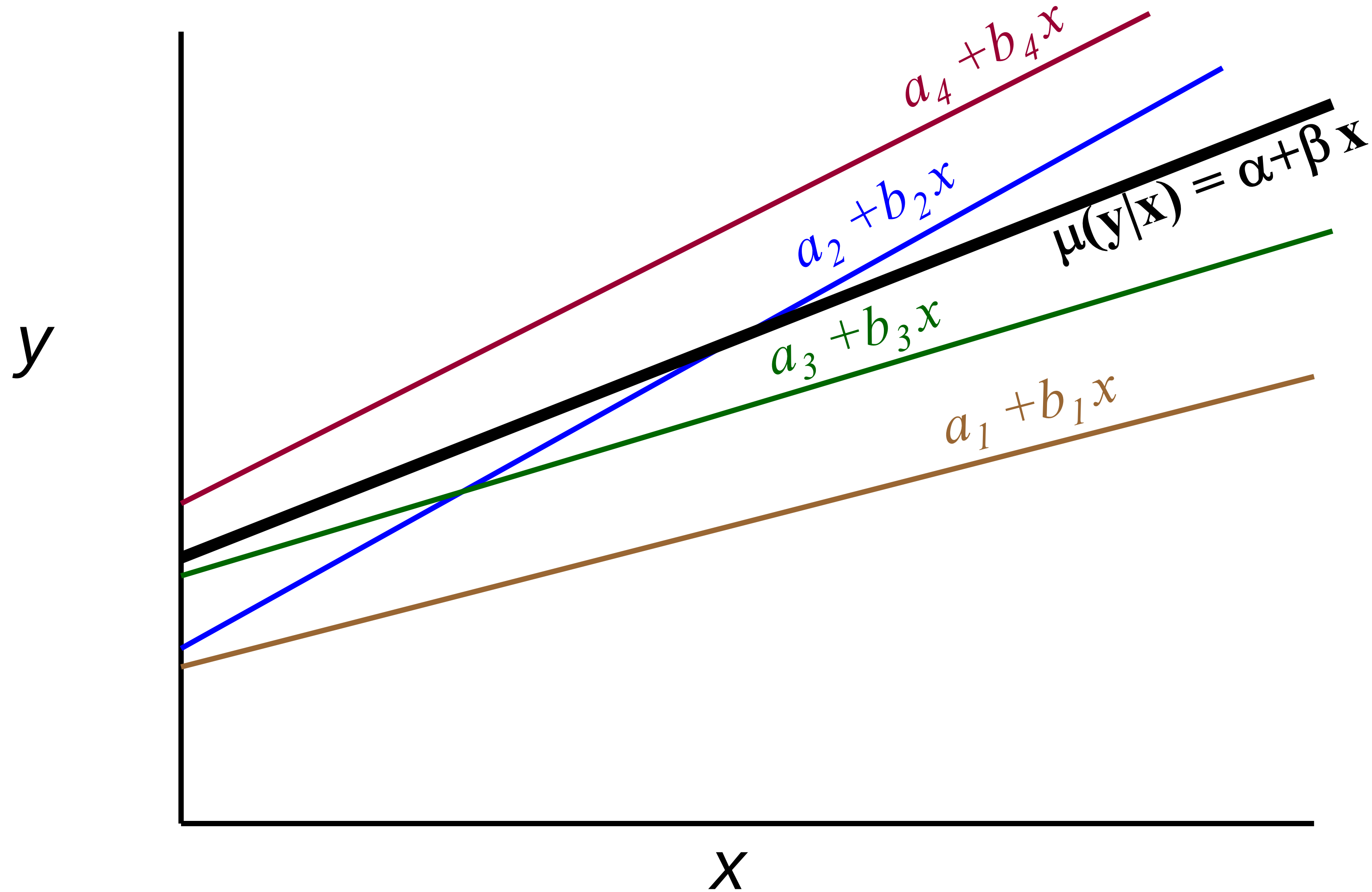
STATISTICS
TEXAS A&M UNIVERSITY

A Graphical Representation of ANCOVA

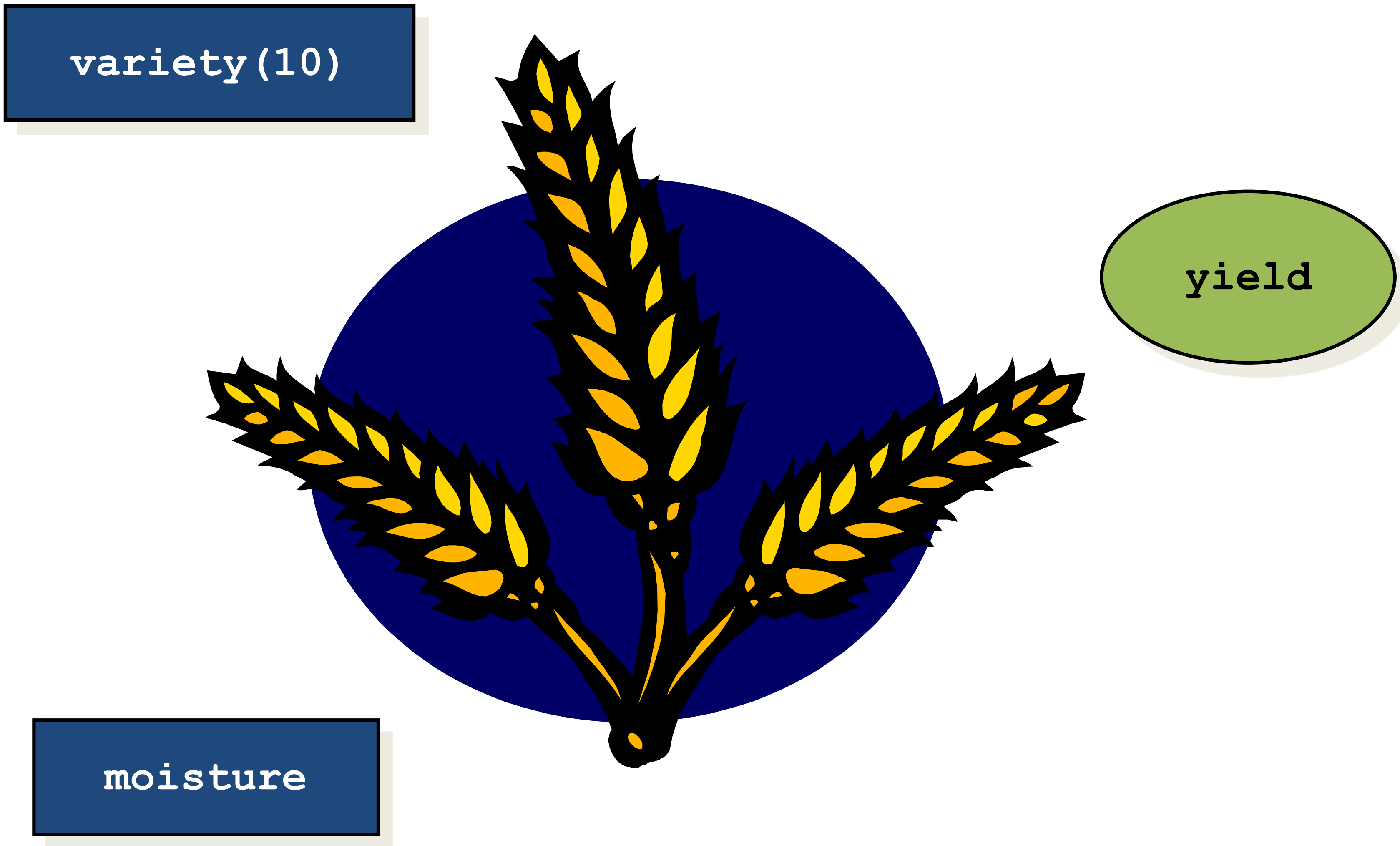


STATISTICS
TEXAS A&M UNIVERSITY

A Graphical Representation of a Random Coefficient Model



Wheat Example



Wheat Example

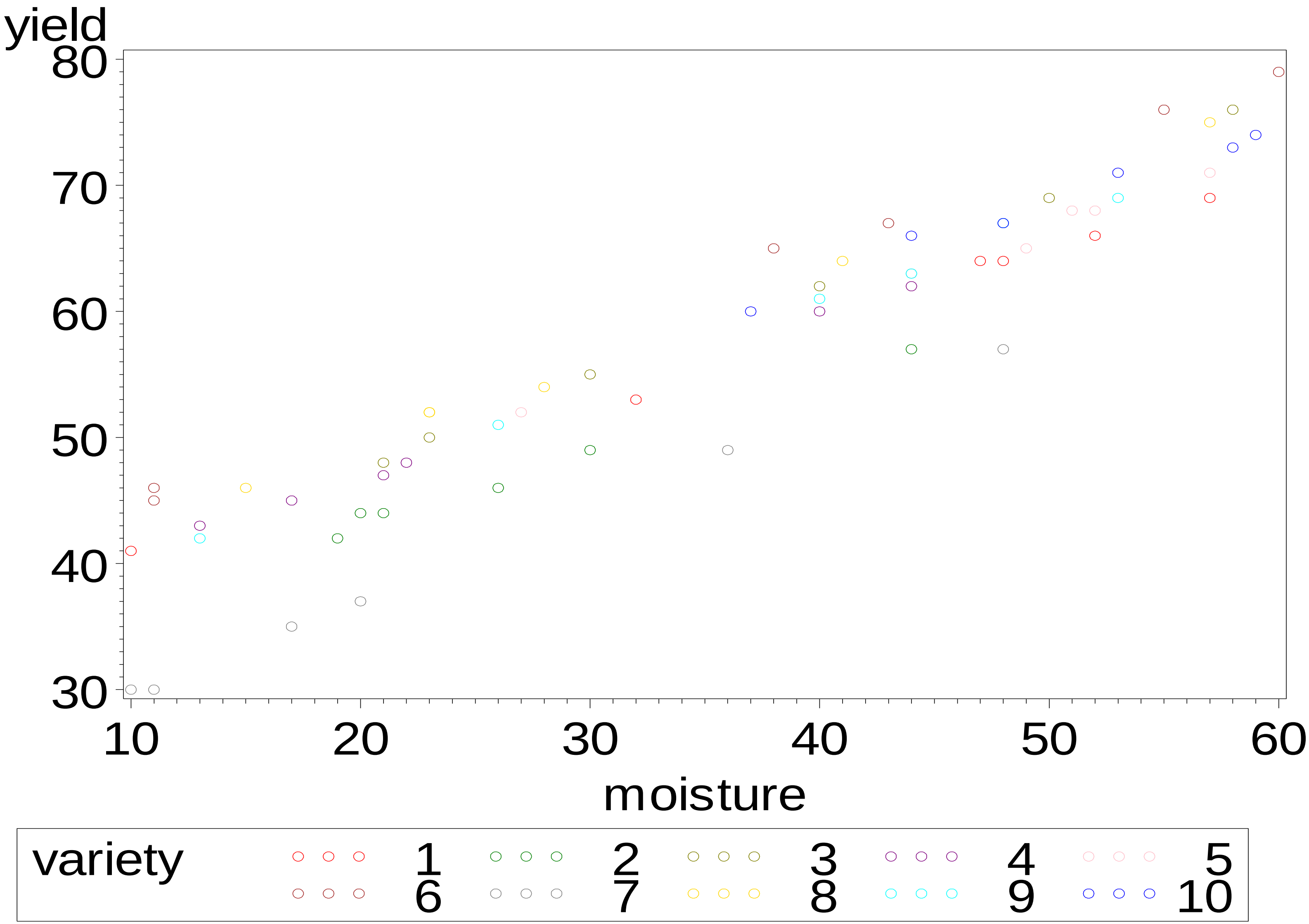
- 10 varieties of wheat are randomly selected from the population of varieties of hard red winter wheat adapted to dry climate conditions.
- Each variety was randomly assigned to 6 one-acre plots of land; thus the EUs are one-acre plots of land in 60-acre field.
- It was thought that the pre-plant moisture content of the plots could have an influence on the germination rate and hence the eventual yield of the plots.
- Therefore, the amount of pre-planting moisture in the top 36 inches of the soil was determined for each plot.
- The response variable is the yield in bushels per acre (**yield**), and the covariate is the measured amount of moisture (**moist**).

STATISTICS
TEXAS A&M UNIVERSITY

The Data

id	variety	moist	yield
1	1	10	41
2	1	57	69
3	1	32	53
4	1	52	66
5	1	47	64
6	1	48	64
7	2	30	49
8	2	21	44
9	2	20	44
10	2	26	46
11	2	44	57
12	2	19	42
13	3	50	69
14	3	40	62
15	3	23	50
16	3	58	76
17	3	21	48
18	3	30	55
19	4	22	48
...	...	...	...

Yield versus Moisture



Question

Which of the following is **false** for the **`wheat`** data?

a. The variable **`variety`** can be considered a random effect.
b. Both **`yield`** and **`moisture`** are continuous variables.
c. You can fit a random coefficient model to this data to model the variety-to-variety variations though the intercepts and slopes.
d. You should fit an analysis of covariance model to this data.

Model for the Wheat Example

slope for **`variety`** *i*, random

$$y_{ij} = a_i + b_i x_{ij} + e_{ij}$$

intercept for **`variety`** *i*, random

i = 1 to 10 (variety)

j = 1 to 6 (plot)

$$a_i \sim N(\alpha, \sigma_a^2)$$

$$b_i \sim N(\beta, \sigma_b^2)$$

$$Cov(a_i, b_i) = \sigma_{ab}$$

$$e_{ijk} \sim N(0, \sigma^2)$$

In Terms of a Mixed Model

$$\left.\begin{aligned} y_{ij} &= a_i + b_i x_{ij} + e_{ij} \\ a_i &= \alpha + a_i^* \\ b_i &= \beta + b_i^* \end{aligned}\right\}$$

$$a_i^* \sim N(0, \sigma_a^2)$$
$$b_i^* \sim N(0, \sigma_b^2)$$
$$Cov(a_i^*, b_i^*) = \sigma_{ab}$$

$$y_{ij} = \alpha + \beta x_{ij} + a_i^* + b_i^* x_{ij} + e_{ij}$$

population intercept

population slope

intercept deviation

slope deviation

i = 1 to 10 (variety)

j = 1 to 6 (plot)

Fitting a Random Coefficient Model

A RANDOM Statement with the SUBJECT= Option

```
model yield=moist / ddfm=kr;
random int moist / type=un subject=variety;
```

$$
\mathbf{G} = \text{variety}\;
\begin{array}{cc}
 & \text{variety} \\
 & \begin{array}{cccccc} \multicolumn{2}{c}{1} & \multicolumn{2}{c}{2} & \ldots & 10 \\ a_1^* & b_1^* & a_2^* & b_2^* & \ldots\ a_{10}^* & b_{10}^* \end{array} \\
\begin{array}{cc} \multirow{2}{*}{1} & a_1^* \\ & b_1^* \\ \multirow{2}{*}{2} & a_2^* \\ & b_2^* \\ \vdots & a_{10}^* \\ 10 & b_{10}^* \end{array} &
\left[\begin{array}{cc|cc|cc}
\sigma_a^2 & \sigma_{ab} & & & & \\
\sigma_{ab} & \sigma_b^2 & & & & 0 \\
\hline
 & & \sigma_a^2 & \sigma_{ab} & & \\
 & & \sigma_{ab} & \sigma_b^2 & & \\
\hline
 & 0 & & & \sigma_a^2 & \sigma_{ab} \\
 & & & & \sigma_{ab} & \sigma_b^2
\end{array}\right]
\end{array}
$$

WheatExample.sas

Questions

Which of the following is **false**?

a. Random coefficient models are the same as the analysis of covariance models.
b. You use the RANDOM statement with the SUBJECT= and TYPE= options in PROC MIXED
to fit a random coefficient model.
c. The RANDOM statement in PROC MIXED defines the **G** matrix.
d. The SUBJECT= option in the RANDOM statement defines a block-diagonal **G** matrix.

Which of the following is **false**?

a. Random coefficient models fit subject-specific models to your data.
b. Random coefficient models accounts for subject variability by modeling the variations among the regression coefficients across subjects.
c. In random coefficient models the type of covariance structure for the random coefficients is difficult to determine.

An Alternative Covariance Structure

```
 model yield=moist / ddfm=kr;
 random int moist / type=FA0(2)
subject=variety;
```

variety

$\mathbf{G} =$ **variety**

variety		1: a_1*	1: b_1*	2: a_2*	2: b_2*	… 10: a_{10}*	10: b_{10}*
1	a_1*	λ_{11}^2	$\lambda_{11}\lambda_{21}$			0	
	b_1*	$\lambda_{21}\lambda_{11}$	$\lambda_{21}^2+\lambda_{22}^2$				
2	a_2*			λ_{11}^2	$\lambda_{11}\lambda_{21}$		
	b_2*		0	$\lambda_{21}\lambda_{11}$	$\lambda_{21}^2+\lambda_{22}^2$		
⋮ 10	a_{10}*					λ_{11}^2	$\lambda_{11}\lambda_{21}$
	b_{10}*					$\lambda_{21}\lambda_{11}$	$\lambda_{21}^2+\lambda_{22}^2$

Alternative covariance structure

Another useful covariance structure in random coefficient regression analysis is the no-diagonal factor-analytic structure with two factors, that is, FA0(2). The $(i, j)^{\text{th}}$ element of the covariance matrix of FA0(q) is given by

$$\sum_{k=1}^{\min(i,j,q)} \lambda_{ik}\lambda_{jk} \text{, where}$$

i is the row position

j is the column position

q is the number of factors

λ_{ij} is the $(i, j)^{\text{th}}$ element of the Cholesky root of the unstructured covariance matrix.

You can use this structure FA0(q) for approximating an unstructured **G** matrix in the RANDOM statement. You can also use this structure to constrain the estimate of **G** to be nonnegative definite.

Details

If **A** is positive definite of size $p \times p$, you can find an upper triangular matrix **U** such that **A**=**U′U**, so that **U** is a sort of square root of **A**. The elements in matrix **U** are referred to as square-root or Cholesky root of matrix **A**. For example,

$$\mathrm{U}' = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & \lambda_{22} \end{bmatrix}, \mathrm{U} = \begin{bmatrix} \lambda_{11} & \lambda_{21} \\ 0 & \lambda_{22} \end{bmatrix}, \quad \mathrm{A} = \mathrm{U'U} = \begin{bmatrix} \lambda_{11}^2 & \lambda_{11}\lambda_{21} \\ \lambda_{21}\lambda_{11} & \lambda_{21}^2 + \lambda_{22}^2 \end{bmatrix}$$

TYPE=FA0(q) versus TYPE=UN

- The two covariance structures are generally interchangeable.

- The covariance parameter estimates are different because they use different parameterizations.

- FA0(q) constrains the **G** matrix to be nonnegative definite, whereas TYPE=UN can occasionally result in an indefinite **G** matrix.

- Using FA0(q) can improve the convergence and stability in the model fitting process.

One Issue in Using DDFM=KR

The DDFM=KR option in the MODEL statement

- adjusts the variance-covariance matrix of fixed and random effects
- might have undesirable consequences for the adjustment when covariance matrices have nonzero second derivatives
  - Adjustment can lead to shrinkage of standard errors
  - Adjusted covariance matrix may not be positive definite
  - Results are not invariant under parameterization.

➢ The DDFM=KENWARDROGER option performs the df calculations detailed by Kenward and Roger (1997).

➢ This approximation involves inflating the estimated variance-covariance matrix of the fixed and random effects by the method proposed by Prasad and Rao (1990) and Harville and Jeske (1992).

➢ Satterthwaite-type df are then computed based on this adjustment.

➢ By default, the observed information matrix of the covariance parameter estimates is used in the calculations.

➢ For covariance structures that have nonzero second derivatives with respect to the covariance parameters, the KR covariance matrix adjustment includes a second-order term. This term can result in standard error shrinkage. Also, the resulting adjusted covariance matrix can then be indefinite and is not invariant under reparameterization.

Affected Covariance Matrices

ANTE(1) , AR(1), ARH(1), ARMA(1,1), CSH, FA0(q), TOEPH, UNR, all SP().

The DDFM=KR(FIRSTORDER) Option

The FIRSTORDER suboption

- eliminates the second derivatives from the calculation of the covariance matrix adjustment
- might be preferred for covariance structures that have nonzero second derivatives.

Fitting a Random Coefficient Model Using TYPE=FA0(2)

This demonstration illustrates the concepts discussed previously.

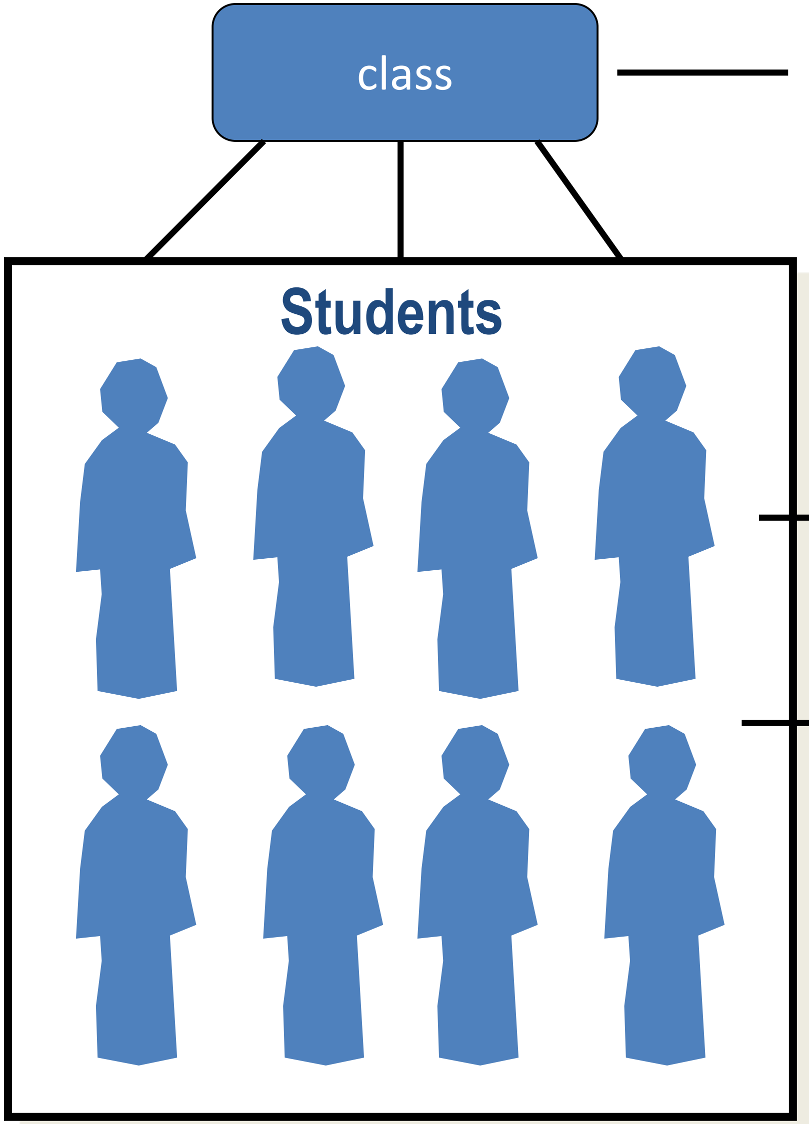
WheatExample.sas

Objectives of Hierarchical Linear Models

– Define hierarchical linear modeling.

– Analyze data from the educational field using the MIXED procedure.

Test Score Gains Data



Percent of class mastering previous curricula (**`Tmastry`**)

Pretest Total (**`PreTotal`**)

- Test Score Gains (**`Gain`**)

Test Score Gains Data (MathscoreExample_sas)

Data was collected for 3111 eight-grade students. The students' test score gains (**Gain**) on one of the mathematics achievement tests were recorded. In addition, the sum of some pretest core items (**PreTotal**) on the same students was also recorded. These students were grouped into 159 classes. A variable measuring the percent of the class with a sufficient degree of mastery of previous curricula (**Tmastry**) was recorded for each class. The data is stored in the SAS dataset which includes the following variables:

Gain: the test score gains on a mathematics achievement test for each student

PreTotal: the sum of some pretest core items for each student

Class: the class each student belongs to

Tmastry: the percent of class mastering previous curricula

The Data

Obs	Class	Tmastry	Gain	Pre Total
1	1	50	2	20
2	1	50	-3	18
3	1	50	5	12
4	1	50	1	9
5	1	50	-3	11
6	1	50	3	12
7	1	50	-8	12
8	1	50	-6	18
9	1	50	-4	13
10	1	50	4	8
11	1	50	1	21
...	...	...	...	...
30	2	80	-2	7
31	2	80	1	13
32	2	80	-10	16
33	2	80	-4	11
34	2	80	-1	9
35	2	80	6	11
...	...	...	...	...

Hierarchical Linear Modeling

- Data has a nested structure, or hierarchy of random effects.

- There might be fixed effects that are measured at each level of the experimental unit, or random effects.

- You perform multilevel analysis:
  - Fit a random coefficient model for level 1 (the smallest sized units).
  - Model the coefficients to be a function of level 2 variables.
  - Continue this pattern if you have more levels.
  - Combine models from all the levels.

A Model at the Student Level

`Gain` for student *i* in class *j*

`PreTotal` for student *i* in class *j*

$$y_{ij} = a_j + b_j x_{ij} + e_{ij}$$

Random error

Random intercept for class *j*

Random slope for class *j*

$$a_j \sim N(\alpha_0, \sigma_a^2) \qquad a_j = \alpha_0 + a_j^* \qquad a_j^* \sim N(0, \sigma_a^2)$$

$$b_j \sim N(\beta_0, \sigma_b^2) \qquad b_j = \beta_0 + b_j^* \qquad b_j^* \sim N(0, \sigma_b^2)$$

$$\mathrm{Cov}(a_j, b_j) = \sigma_{ab} \qquad \mathrm{Cov}(a_j^*,\ b_j^*) = \sigma_{ab}$$

A Model at the Class Level

– Percent of class mastering previous curricula (**`Tmastry`**) is measured at the **`Class`** level.

– You can incorporate this effect to model the intercept and slope for each class:

$$a_j^* \sim N(0, \sigma_a^2)$$
$$b_j^* \sim N(0, \sigma_b^2)$$
$$\mathrm{Cov}(a_j^*, b_j^*) = \sigma_{ab}$$

`Tmastry` for Class *j*

$$a_j = \alpha_0 + \alpha_1 z_j + a_j^*$$
$$b_j = \beta_0 + \beta_1 z_j + b_j^*$$

Fixed, Intercepts

Random Component

Fixed, Slopes

A Model of Multilevel Effects

$$y_{ij} = a_j + b_j x_{ij} + e_{ij}$$
$$a_j = \alpha_0 + \alpha_1 z_j + a_j^*$$
$$b_j = \beta_0 + \beta_1 z_j + b_j^*$$

$$y_{ij} = \alpha_0 + \alpha_1 z_j + \beta_0 x_{ij} + \beta_1 z_j x_{ij} + a_j^* + b_j^* x_{ij} + e_{ij}$$

Do you agree that hierarchical linear models are closely related to random coefficient models?

Fitting a Hierarchical Linear Model

This demonstration illustrates the concepts discussed previously.

MathscoreExample.sas