CHE384, From Data to Decisions: Measurement,
Uncertainty, Analysis, and Modeling

Lecture 67
Blocking in Experimental Design

Chris A. Mack

Adjunct Associate Professor

http://www.lithoguru.com/scientist/statistics/

Blocking

- Let X be the results from treatment 1, and Y the results from treatment 2. We wish to measure X – Y (the difference in results)

$$var[X-Y] = v\quad [X] + var[Y] - 2cov[X,Y]$$

- We can reduce the variance of X – Y by increasing the covariance of X and Y
  - An error that is the same for X and Y will cancel
  - We can increasing the covariance with blocking

Blocking vs. Covariate Analysis

- For uncontrolled, measured inputs, use blocking or covariate analysis to remove an effect we are not interested in (nuisance), reducing known variability
  - Different measurement tools, process batches
  - Spatial or temporal variations
- Blocking: grouping the experimental results that you wish to compare into blocks that are similar to one another
- Covariate Analysis: put the uncontrolled but measured inputs into the model. Then ignore them when the modeling is complete

Covariate Analysis Example

- Testing various process inputs, the response is measured using two different tools (A and B) that I can't select but can observe
- Add an indicator variable in the model for tool A vs. tool B
  - If the coefficient of that term is not significant, drop it from the model and refit
  - If the coefficient of that term is significant (there is a measureable tool bias), keep it in the regression but leave it out when using the model

Classic Blocking Example

- I have a new shoe sole that I claim will last longer, but is otherwise identical to the existing sole
- I've recruited 100 volunteers to test the new versus old soles. What experimental design should I use?
  - Randomization: randomly assign half of the participants to wear the new soles, and half to wear the old soles
  - Blocking: every participant is given one shoe with the new sole and one with the old (randomly assigned to left or right foot)

"Block what you can; randomize what you cannot."

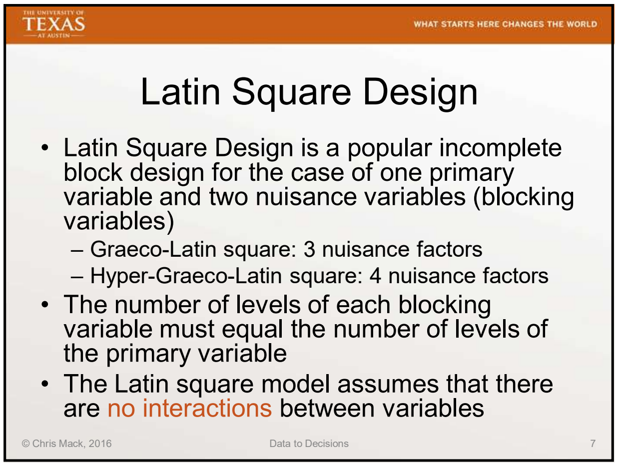
Randomized Complete Block Design

- The complete experiment is performed for each block
  - Each block sees each treatment exactly once
  - Within each block, the testing order is randomized
- Examples of blocks (nuisance variables)
  - Raw material batches
  - People (operators)
  - Process or measurement tools
  - Time

Model:

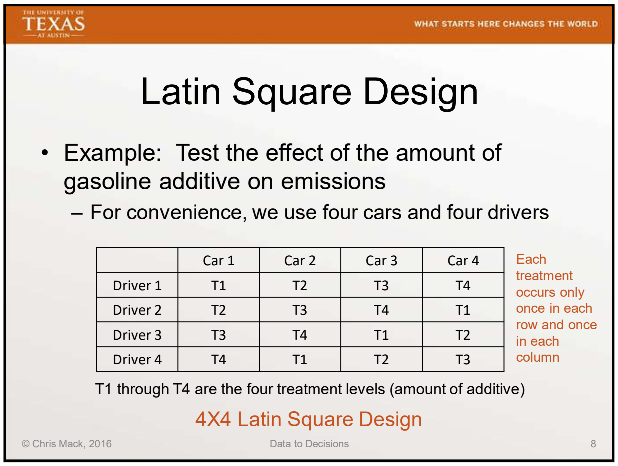
$$y_{ij} = f(\underset{\uparrow\ \text{treatment}}{\tau_i}) + \overset{\text{blocking}\ \downarrow}{\beta_j} + \varepsilon_{ij}$$

- If it is not possible to run a complete experiment in each block, use a balanced incomplete block design



Latin Square Design

- Latin Square Design is a popular incomplete block design for the case of one primary variable and two nuisance variables (blocking variables)
  - Graeco-Latin square: 3 nuisance factors
  - Hyper-Graeco-Latin square: 4 nuisance factors
- The number of levels of each blocking variable must equal the number of levels of the primary variable
- The Latin square model assumes that there are no interactions between variables



Latin Square Design

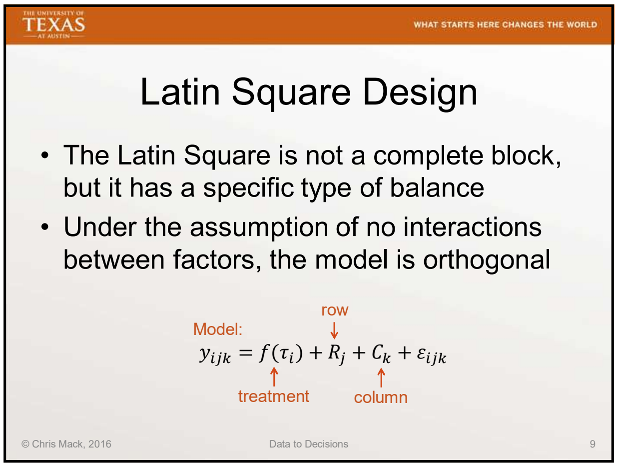
- Example: Test the effect of the amount of gasoline additive on emissions
  - For convenience, we use four cars and four drivers

	Car 1	Car 2	Car 3	Car 4
Driver 1	T1	T2	T3	T4
Driver 2	T2	T3	T4	T1
Driver 3	T3	T4	T1	T2
Driver 4	T4	T1	T2	T3

Each treatment occurs only once in each row and once in each column

T1 through T4 are the four treatment levels (amount of additive)

4X4 Latin Square Design

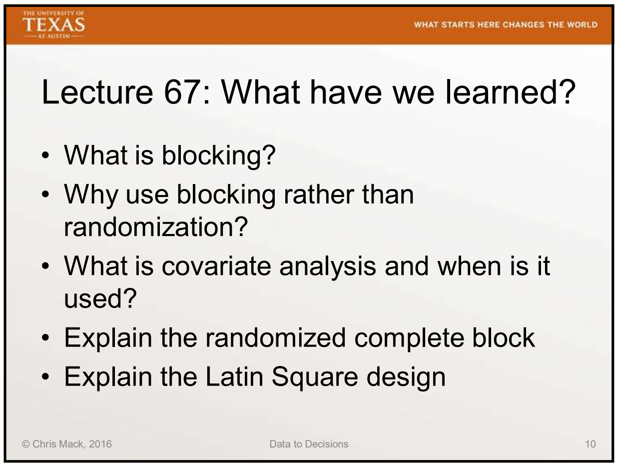


Latin Square Design

- The Latin Square is not a complete block, but it has a specific type of balance
- Under the assumption of no interactions between factors, the model is orthogonal

Model:

$$y_{ijk} = \underset{\text{treatment}}{\underset{\uparrow}{f(\tau_i)}} + \overset{\text{row}}{\overset{\downarrow}{R_j}} + \underset{\text{column}}{\underset{\uparrow}{C_k}} + \varepsilon_{ijk}$$



Lecture 67: What have we learned?

- What is blocking?
- Why use blocking rather than randomization?
- What is covariate analysis and when is it used?
- Explain the randomized complete block
- Explain the Latin Square design