ANALYSIS OF COVARIANCE

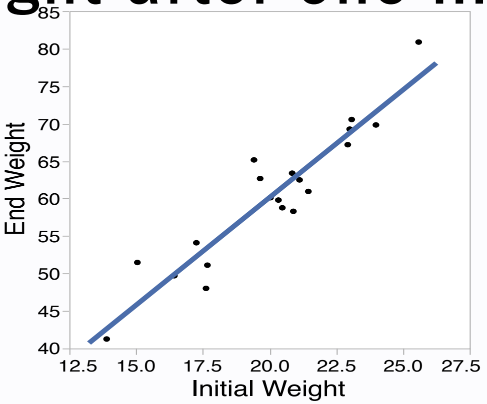
Chapter 9

LEARNING OBJECTIVES

- Explain differences between treatment factors and covariates
- Explain usual assumptions between covariates, treatments, and responses
- Write ANCOVA models for a single covariate
- Perform ANCOVA in R (upcoming Lab)

MOTIVATING EXAMPLE

- Goal: compare effects of several diets on the weights of month-old piglets
- Response: weight after being on diet for one month
- The initial weight (x_{ij}) is significantly correlated with the weight after one month



- Prevent potential confounding and experimental error due to the variation in initial weight

COVARIATES AND NUISANCE FACTORS

- Nuisance factor: potential major source of variation unrelated to treatments
- Control: set the x_{ij} to similar level, no modeling
  - Not always possible
  - Could reduce representativeness of results
- Blocking: group EU's having same/similar levels x_{ij}
  - Need to measure nuisance factor prior to randomization
  - Separate randomizations for each level!
- Analysis of covariance: Measure x_{ij} prior to treatment application and include relationship in the model

NOTATION COMMENTS

- Before x_{ij} denoted the treatment factor values we studied
  - $x_{i1} = x_{i2} = \cdots = x_{ir_i}$

- For ANCOVA, x_{ij} denotes measurement of a nuisance factor for the *j*-th replicate of treatment settings *i*

- Expect the x_{ij} to be different for given *i*

RANDOMIZATION AND COVARIATES

- How does randomization affect covariates?
  - Reduce possibility that treatment effects are confused with covariate effects….but how?

- Let x_k be covariate value for EU *k* (hasn't yet been assigned to a treatment)
- Randomly partitioning the x_k into *t* groups

- Key: distribution of each group's covariates, denoted x_{ij}, similar to distribution of all x_k**
  - **less likely for small runs

- Covariates affect treatment responses similarly

ANCOVA ASSUMPTIONS

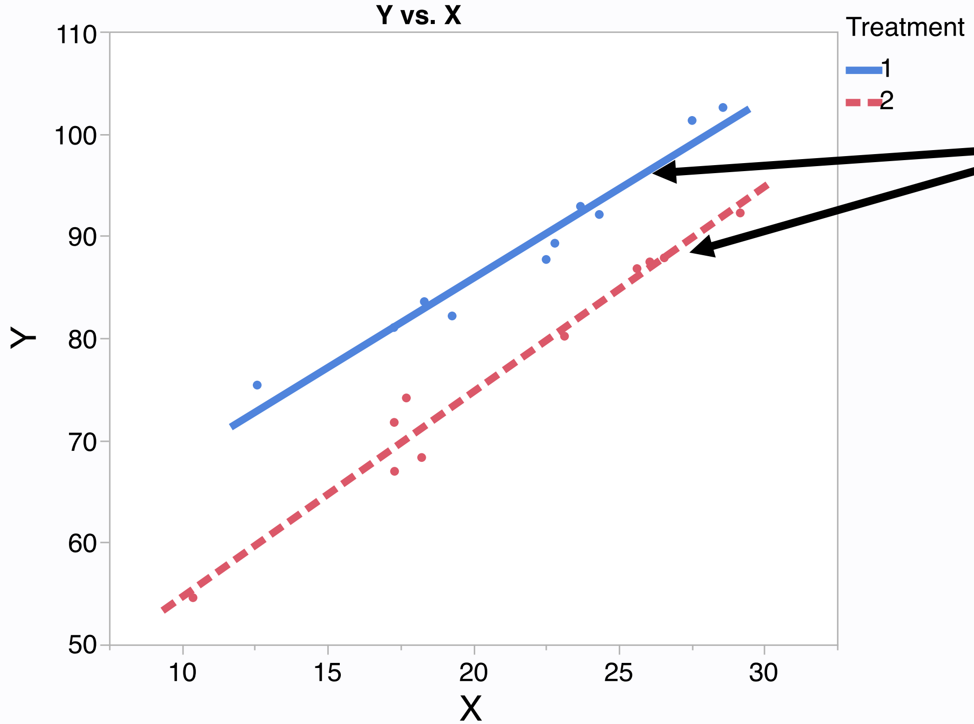
- Covariates explain EU/OU variation
- The x_k are **not influenced** by treatment
  - Measure covariate before treatment application
- If false, then you introduce **potential confounding** between covariate effects and treatment effects
- Won't be able to distinguish which is important

WHY INCLUDE COVARIATES?

- Randomization only reduces bias (potential confounding) due to covariate effect on average
- msE will be inflated estimate of σ^2 if we only fit the effects model → low power
- Need to include covariate effect in model
- What are some common models for covariates?
  - Add effects to the existing models we have already seen

ANCOVA WITH QUANTITATIVE COVARIATES SIMPLE LINEAR REGRESSION

- Assume that covariate is quantitative and relationship with response is same for all treatments
- SLR ANCOVA model visualized:



Different SLR model for each treatment

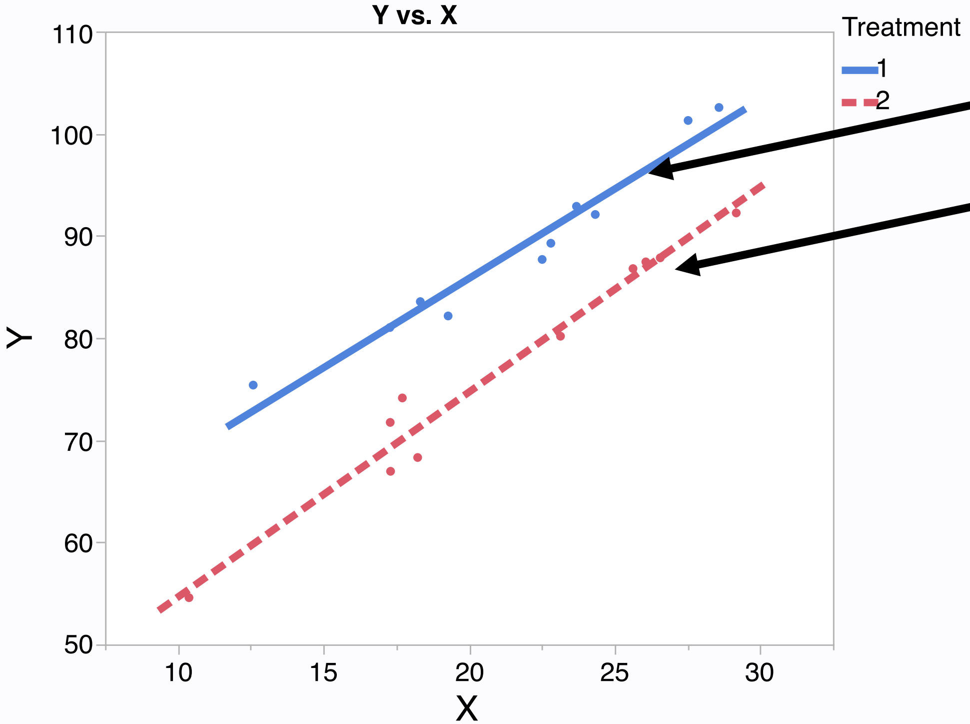
1) Same slope parameter
2) Different intercepts

Comparing treatments for fixed x_{ij} value gives same difference $\tau_i - \tau_{i'}$

ANCOVA WITH QUANTITATIVE COVARIATES SIMPLE LINEAR REGRESSION

- **Write model mathematically as:**

$$Y_{ij} = \mu + \tau_i + \color{red}{\boldsymbol{\beta x_{ij}}} + E_{ij}$$



$\mu + \tau_1 + \beta x_{1j}$

$\mu + \tau_2 + \beta x_{2j}$

Calculate $E(Y_{ij} - Y_{i'j'})$

Calculate $E(\bar{Y}_{i.} - \bar{Y}_{i'.})$

ANCOVA WITH QUANTITATIVE COVARIATES POLYNOMIAL MODELS

- Polynomial ANCOVA model

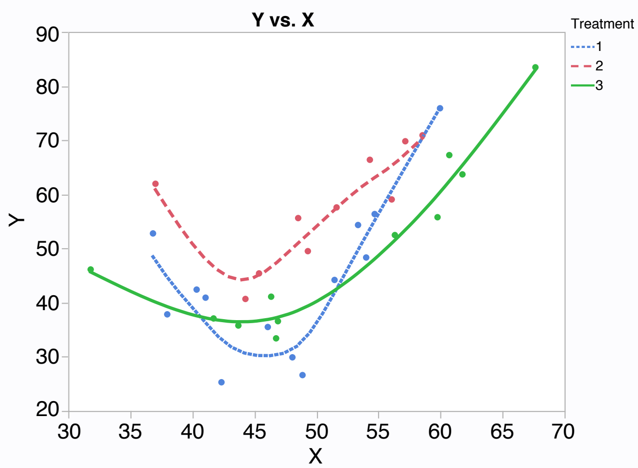
$$Y_{ij} = \mu + \tau_i + \boldsymbol{\beta_1 x_{ij}} + \boldsymbol{\beta_2 x_{ij}^2} + \cdots + \boldsymbol{\beta_p x_{ij}^p} + E_{ij}$$

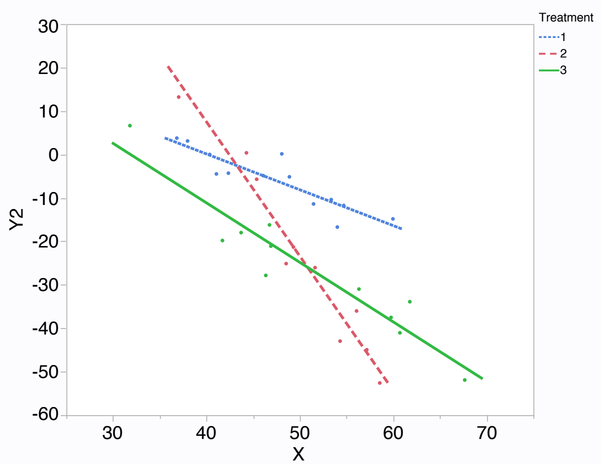
- Treatment comparisons still assumed to be the same throughout the covariate values

- Polynomial coefficients are estimable
- Recommend backwards elimination
- Lack-of-fit unlikely due to no control over x_{ij} so unlikely to have replicates

COVARIATE-TREATMENT INTERACTION

- May find that the **relationship between covariate and response changes with treatment**





- If treatment had no effect on the covariate value, then we need to **assign a unique slope** to each treatment

- **Treatment comparisons depend on covariate value**

MODEL INCLUDING COVARIATE-TREATMENT INTERACTION

- One way to write the model is

$$Y_{ij} = \mu + \tau_i + \beta x_{ij} + \beta_i x_{ij} + E_{ij}$$

- Notes about this model:
  - The β_i are adjustments to β
  - Treatment comparisons depend on x_{ij} value

ANCOVA WITH CATEGORICAL COVARIATES

- Suppose we don't have piglet initial weight values but rather have 3 initial weight groups
  - Small (1), Medium (2), Large (3)

- Distribution of categories:

Small	Medium	Large
11	9	10

- How many small/medium/large assigned to each treatment group?
  - Random assignment means we EXPECT equal distribution
  - For a given experiment, the distribution will not be equal

ANCOVA WITH CATEGORICAL COVARIATES

- Perform randomization and get following assignment

		Small	Medium	Large	Total:
Treatment	1	3	3	4	10
	2	2	4	4	10
	3	6	2	2	10
	Total:	11	9	10	

- Let $\theta_1, \theta_2, \theta_3$ be effects for covariate group effect
  - Like treatment effects but we can't randomize to EUs

- What's a useful model for this data?

CATEGORICAL COVARIATE MODEL 1: TWO-FACTOR MAIN EFFECT MODEL

- **Want to combine two models together**
  - **Treatment model:** $Y_{ij} = \mu + \tau_i + E_{ij};\ \ \mathrm{i} = 1,2,3;\quad \mathrm{j} = 1, \dots, 10$
  - **Initial group model:** $Y_{kl} = \mu + \theta_k + E_{kl};\quad k = 1,2,3;\ \ j = 1, \dots, r_k$

    $r_1 = 11,\ r_2 = 9;\ \ r_3 = 10$
- **Index individual observations using 3 subscripts:**
  - ***i*** **= treatment assigned**
  - ***j*** **= covariate group**
  - ***k*** **= replicate for paired (*i*,*j*)**
- **Final model:**

$$\boxed{\begin{array}{c} Y_{ijk} = \mu + \tau_i + \boldsymbol{\theta_j} + E_{ijk} \\ i = 1,2,3;\quad j = 1,2,3;\quad k = 1, \dots, r_{ij};\quad \mathrm{E_{ijk}} \sim^{iid} N(0, \sigma^2) \\ r_{11} = 3,\ r_{12} = 3,\ r_{13} = 4, \dots, r_{33} = 2 \end{array}}$$

CATEGORICAL COVARIATE MODEL 2: DESIGN-DEPENDENT SUBSCRIPTS

- Keep subscripts from one-way effects model

$$Y_{ij} = \mu + \tau_i + \boldsymbol{\theta}_{?} + E_{ij}$$

- ? subscript should:
  - Equal 1, 2, or 3
  - Correspond to covariate group for *j*th replicate of treatment *I*
- Let $d[i,j] \in \{1,2,3\} =$ initial group for replicate j of treatment i
  - The "*d*" means it depends on the randomized design
- Final model:

$$\boxed{\begin{array}{c} Y_{ij} = \mu + \tau_i + \boldsymbol{\theta}_{\boldsymbol{d[i,j]}} + E_{ij} \\ \mathrm{i} = 1,2,3; \quad \mathrm{j} = 1, \ldots, 10; \quad \mathrm{d}[i,j] \in \{1,2,3\}; \quad E_{ij} \sim^{iid} N(0,\sigma^2) \end{array}}$$

LEARNING OBJECTIVES REVIEW

- Explain differences between treatment factors and covariates
- Explain usual assumptions between covariates, treatments, and responses
- Write ANCOVA models for a single covariate
- Perform ANCOVA in R (upcoming Lab)