Understanding
& Interpreting
Regression Analysis

OCTRI BERD Program

28 November 2018

Workshop overview

- Welcome
- What this workshop is not
  - a "first course" in statistics for those who desire a fundamental understanding of what to do, or not do. We assume regression analysis is the appropriate tool for your problems and you've seen it before
  - a detailed review, development or extension of what is typically seen in a standard course on regression analysis
- What this workshop is
  - an adjuvant or corrective therapy for the interpretation of key scientific quantities (estimators) obtained from regression analyses
    - we mean *means* (viewed through the lens of regression coefficients)
  - is narrow in scope; providing the opportunity for much needed insight to clearly communicate research findings

Workshop info: about the instructors

- Kyle Hart

  Biostatistician, OHSU, Department of Obstetrics and Gynecology, Biostatistics & Design Program (BDP)

  14 years experience in biomedical research, 4th year at the OHSU & BDP

  Previously worked as a data manager at the VA Portland Health Care System and at a private medical device company

  Degrees in Biostatistics (MS), English & Technical Writing (BS)

- Broad practitioner across many types of methods, including lots of simple methods, some more exciting stuff, and lots of time mentoring junior investigators

- . . . and I like to play with synthesizers

Workshop info: about the instructors

- David Yanez

  Professor of Biostatistics, OHSU/PSU School of Public Health

  4th year at OHSU, Co-Director of BDP

  Prior to that, was Professor at the Department of Biostatistics, UW

- Collaboratively, worked extensively in CVD research & on projects in anesthesia, emergency medicine, nephrology, nursing & pediatrics

- Statistical interests include: clinical trials, observational studies, longitudinal data analysis, robust methods, measurement error models

- Taught a lot: medical biometry, regression, survival, longitudinal data analysis, mathematical statistics, measurement error models, biostatistical consulting & technical writing

- . . . tries not to be boring

Workshop info: fair warning

Please note:

- Some material may conflict with textbooks, wikipedia, and other easily accessible resources
- Statistical methods (e.g., t-test) can be motivated/interpreted in more than one way
  - Not everyone writing about statistics knows this, or admits it
- In this workshop we will make/use motivations and interpretations that make fewer assumptions
- Why this approach?
  - It should free you to think about what is relevant to your science and not whether potentially relevant statistical assumptions are satisfied or violated

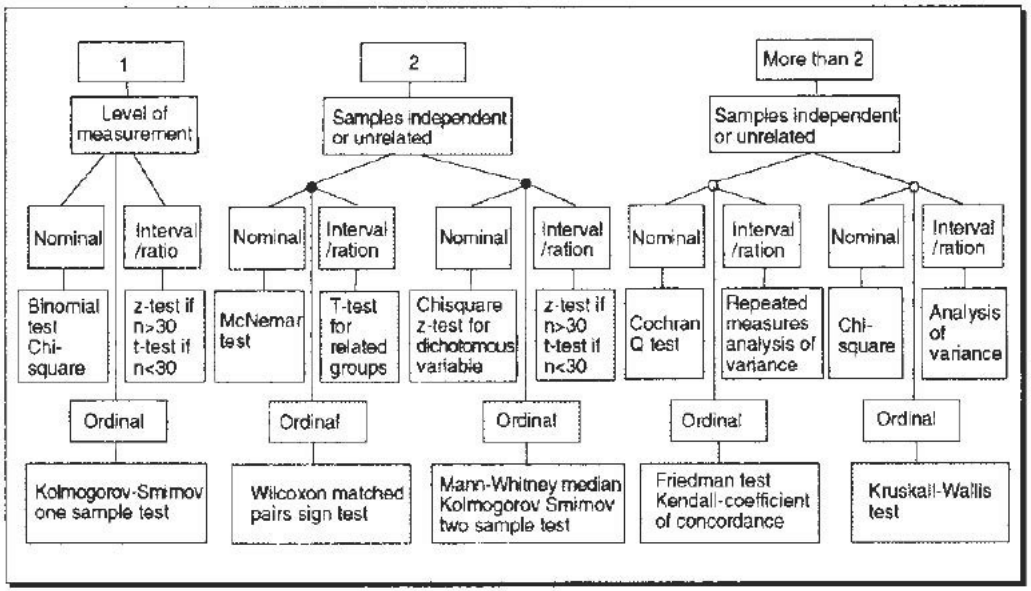
Workshop info: fair warning

Background knowledge tends to vary **a lot**. If issues arise, consider how you might solve them

- *The material was covered too fast*
  - Which parts/topics were confusing?
  - The workshop lectures (slides, scribbles & audio) are being recorded; you can watch them repeatedly
- *I've seen this all before . . .*
  - What was different in these presentations, why present it like that to this audience?
  - Of course, you are welcome/encouraged to ask questions
- *It doesn't feel like a math course*
  - It shouldn't, but that will depend on your background. Statistics uses math but is not math

Workshop info: fair warning

- *There was no "cookbook", flowchart or template of what I should do*
  - Such courses tend to cover a lot of methods but with little depth
    - In some simple scenarios may be okay
    - For more complex studies (involving humans), maybe not
  - We'll keep the scope narrow with more depth
    - Memorizing many formulas (recipes) – tend to forget them
    - Understanding concepts, fewer formulas tend to "stick"



Disclaimer

- Much of these materials have been acquired through courses taught and discussions had over the past 25+ years. The lion's share of the credit goes to Scott Emerson & Kenneth Rice, first-rate statisticians, consummate pedagogues. Through their excellent expositions and discussions on statistical methods, in general, and general regression methods, in particular, these works flow
- Any/all errors contained herein and presented forthwith are the sole responsibility of this presenter
- To begin . . .

Homily

- "Everything is regression."

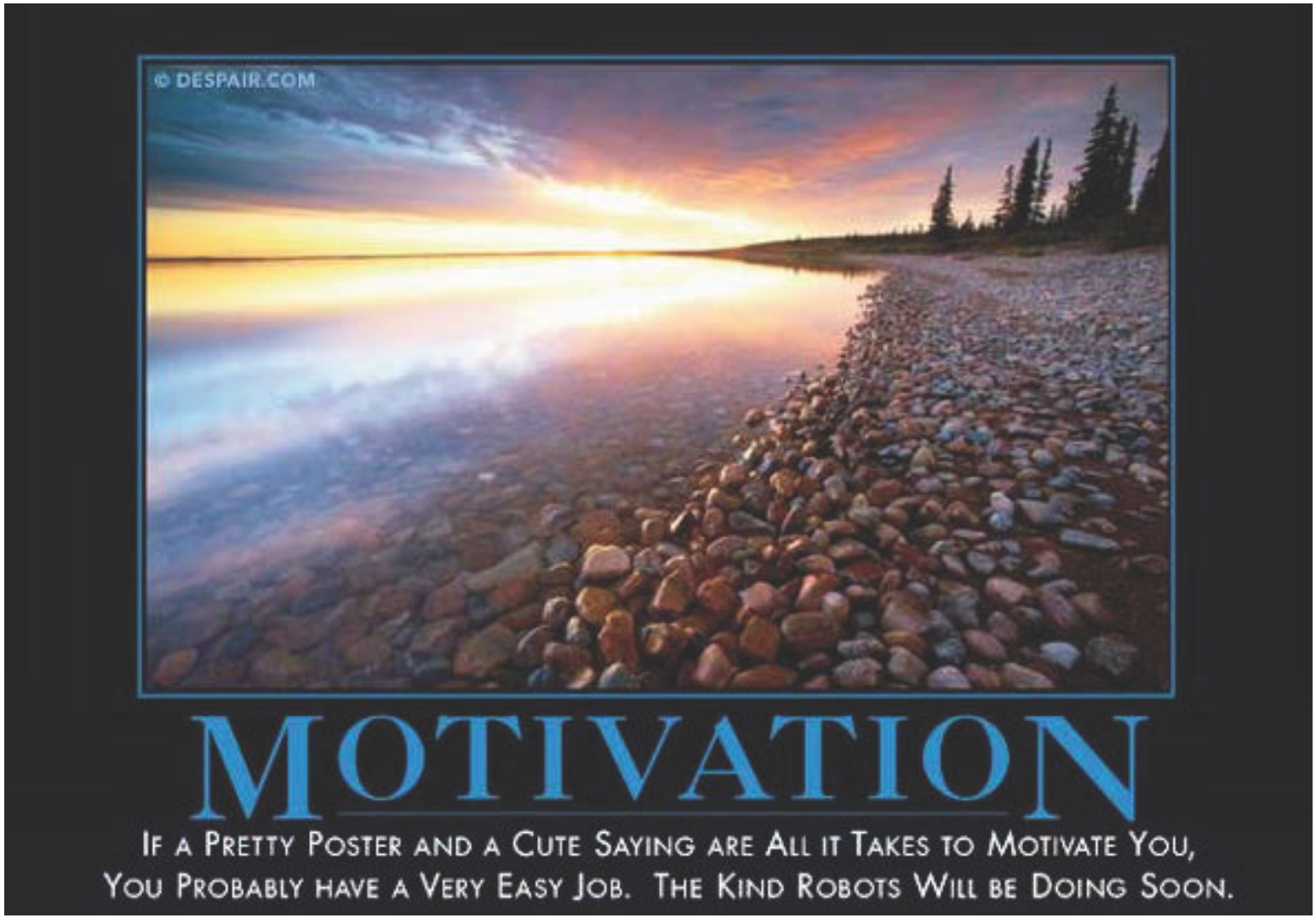
  -Scott Emerson, Emeritus Professor of Biostatistics, UW

- While there are no absolutes, this is *mostly* true
- Statistical methods used to answer scientific questions almost exclusively fall into the following categories:
  1. Prediction of individual observations
  2. Identifying clusters of observations or variables
  3. Quantifying the distribution of some variable
  4. Comparing the distributions of some variable across groups
- We use regression to address items 1, 3, 4
- We will focus solely upon item 4
  - *If you remember one thing, please remember regression is an all-purpose tool for "comparing groups"*

Preliminaries: notation

- For regression, it is common (99 out of 100 statisticians agree. . .) to use the following notation:
  - N or n denote the number of subjects
    - It is also called the sample size
  - Y denotes the outcome (or response) variable (e.g., FEV_1, weight)
  - X denotes the grouping (or predictor, independent) variable (e.g., treatment group, exposure group, age)
  - When appropriate, we will use mnemonic variable names for the characteristics being studied (e.g., FEV_1, AGE, HEIGHT, TRT, etc.)

Preliminaries: motivation



Preliminaries: motivation

- **Q**: Why regression?
- **A**: It is our best all-purpose tool for comparing a response variable across populations defined by some "grouping" variable
- Let's start with the basics. . .

Preliminaries: Anatomy of a Line

- In mathematics, we model a straight line of two variables, X and Y, as

$$Y = a + bX,$$

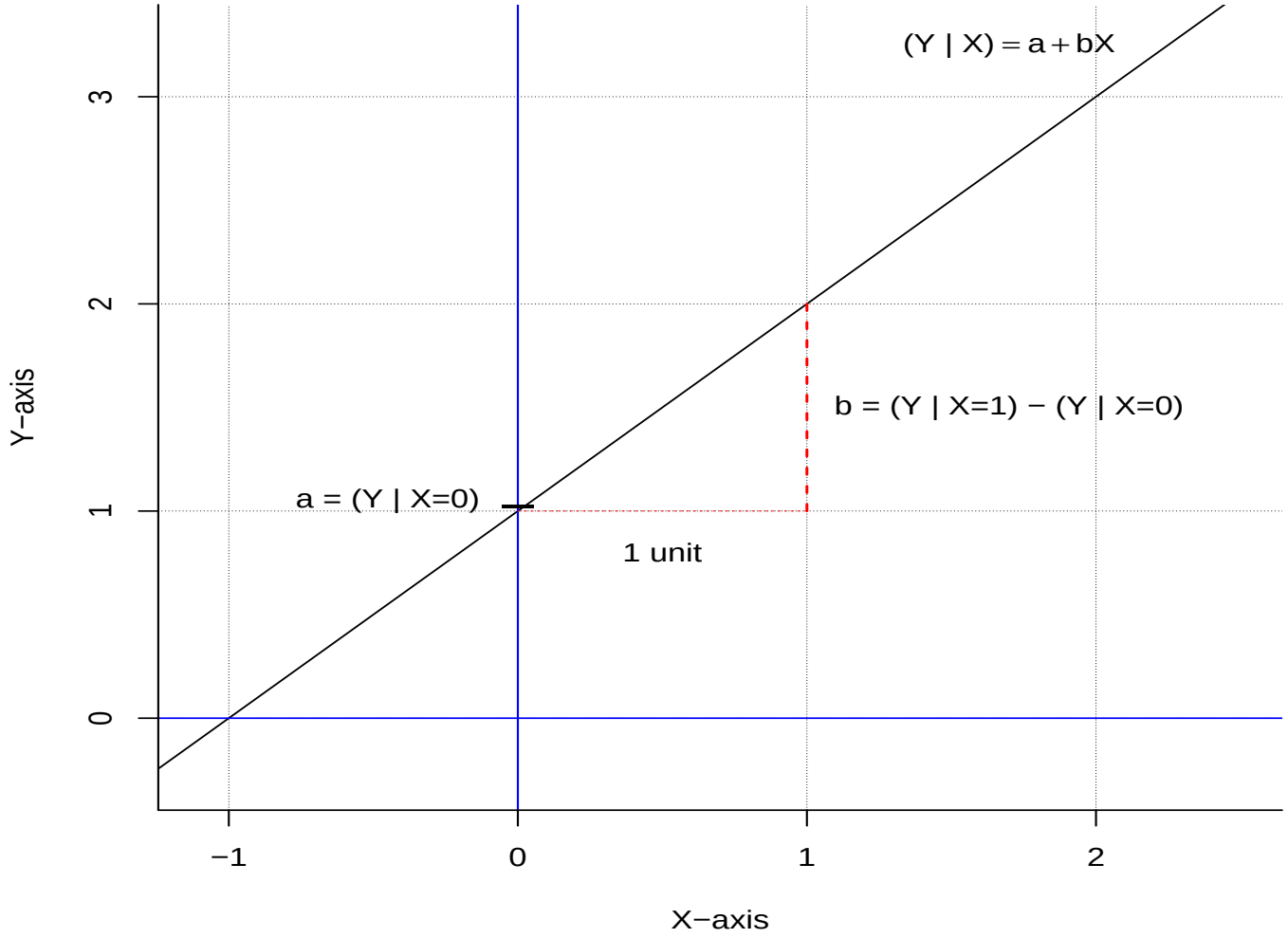
- where
  - Y denotes the *dependent* (aka *output*) variable
  - X denotes the *independent* (aka *input*) variable
- We generally consider Y to be a function of X, i.e., $(Y \mid X) = a + bX$
  - a is the value of Y when $X = 0$

$$\begin{aligned}(Y \mid X = 0) &= a + b(0) \\ &= a.\end{aligned}$$

  - b is the *difference* in Y for a unit difference in X

$$\begin{aligned}(Y \mid X = 1) - (Y \mid X = 0) &= a + b(1) - [a + b(0)] \\ &= a + b - a \\ &= b.\end{aligned}$$

Preliminaries: Anatomy of a Line



Preliminaries: 'Simple'* Linear Regression

- For regression, we model the *average* or *expected value* of Y as

$$\mathbb{E}(Y \mid X) = \beta_0 + \beta_1 X$$

- $\mathbb{E}[\cdot]$ denotes the mean or *expected value*
  - β_0 is the mean value of Y when $X = 0$

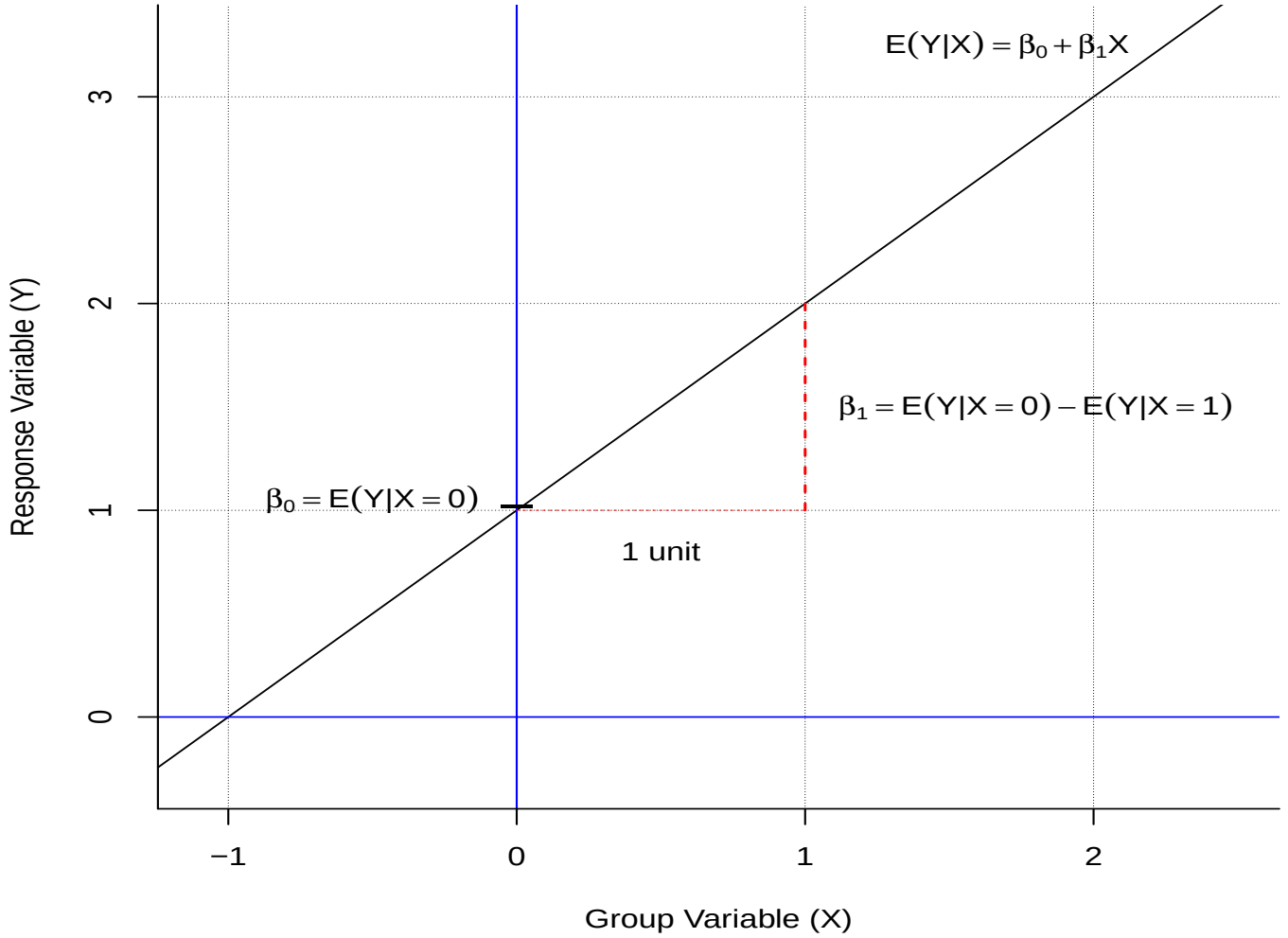
$$\begin{aligned}\mathbb{E}(Y \mid X = 0) &= \beta_0 + \beta_1(0) \\ &= \beta_0.\end{aligned}$$

  - β_1 is the mean *difference* in Y for a unit difference in X

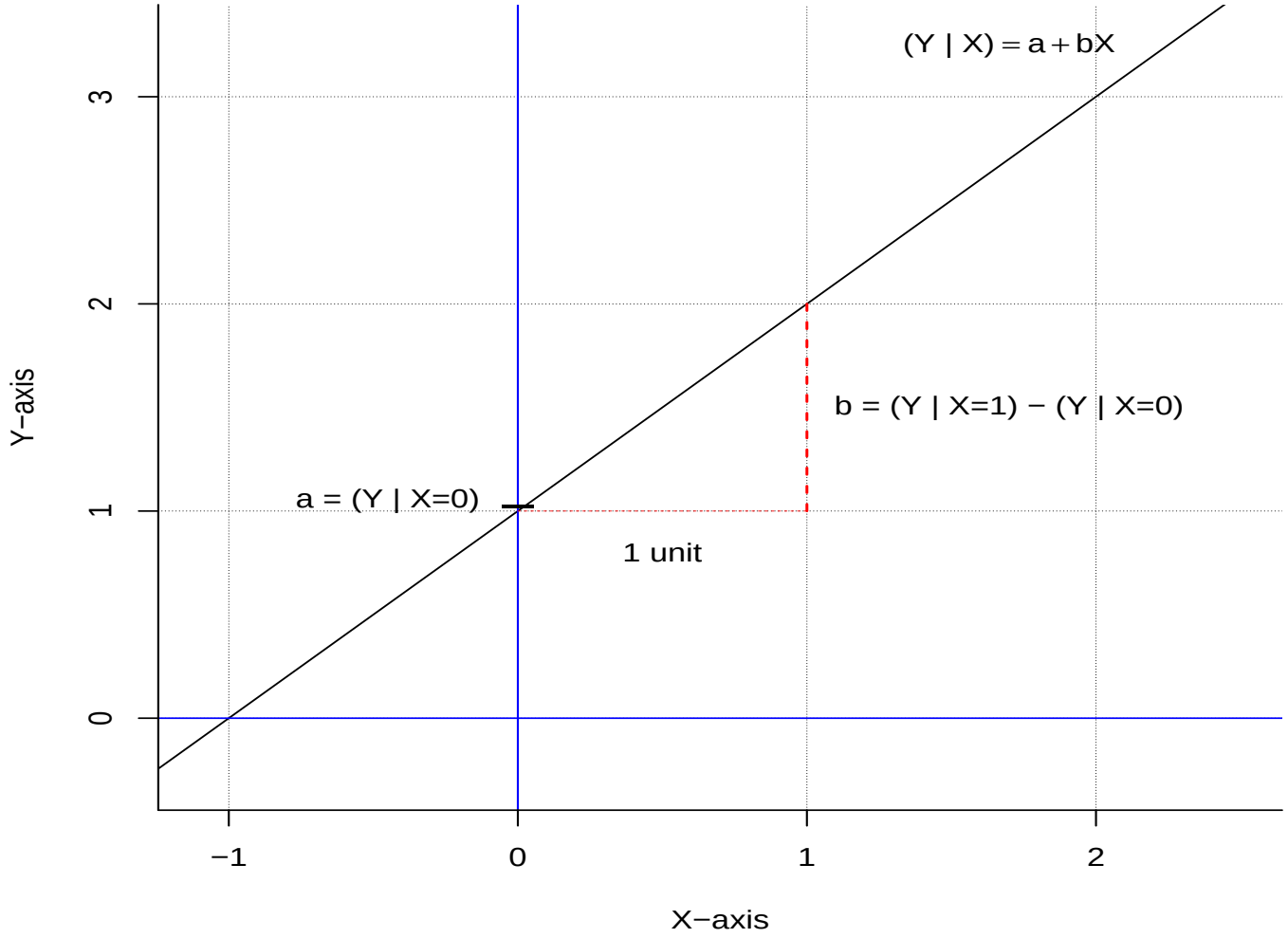
$$\begin{aligned}\mathbb{E}(Y \mid X = 1) - \mathbb{E}(Y \mid X = 0) &= \beta_0 + \beta_1(1) - [\beta_0 + \beta_1(0)] \\ &= \beta_0 + \beta_1 - \beta_0 \\ &= \beta_1.\end{aligned}$$

*'Simple' here means only one independent variable and an intercept

Preliminaries: Simple Linear Regression



Preliminaries: Anatomy of a Line

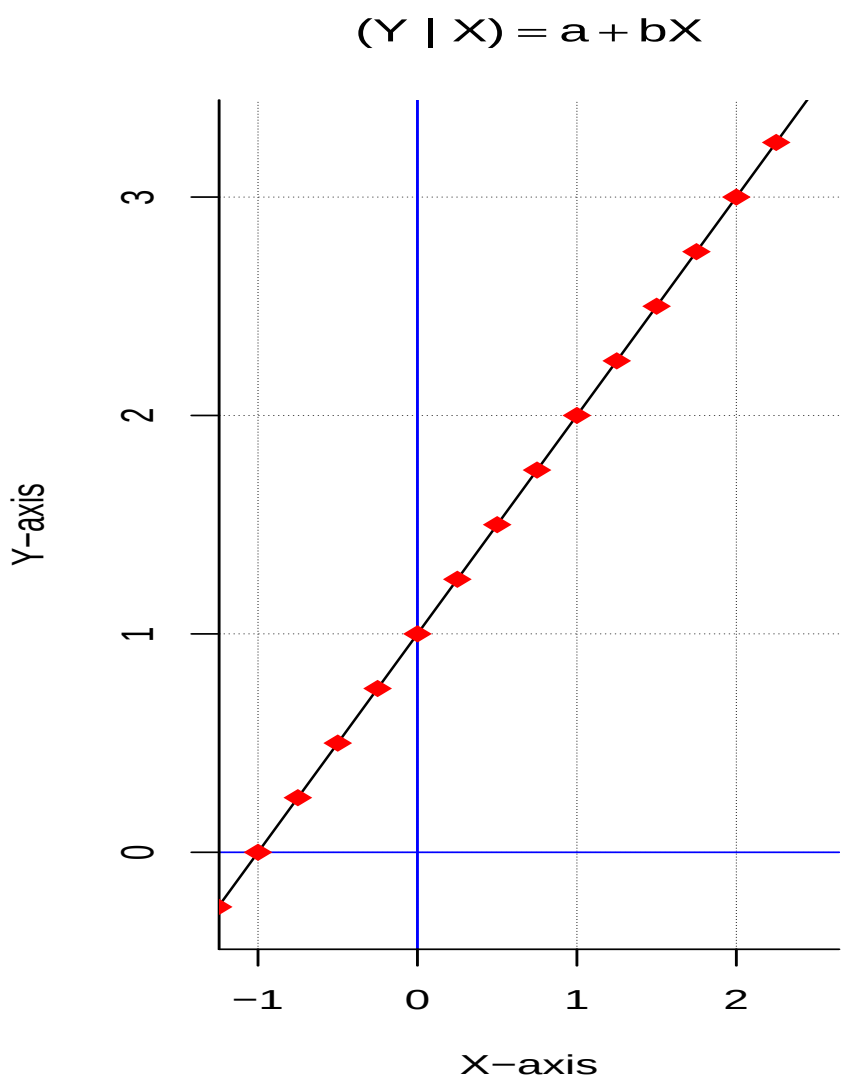


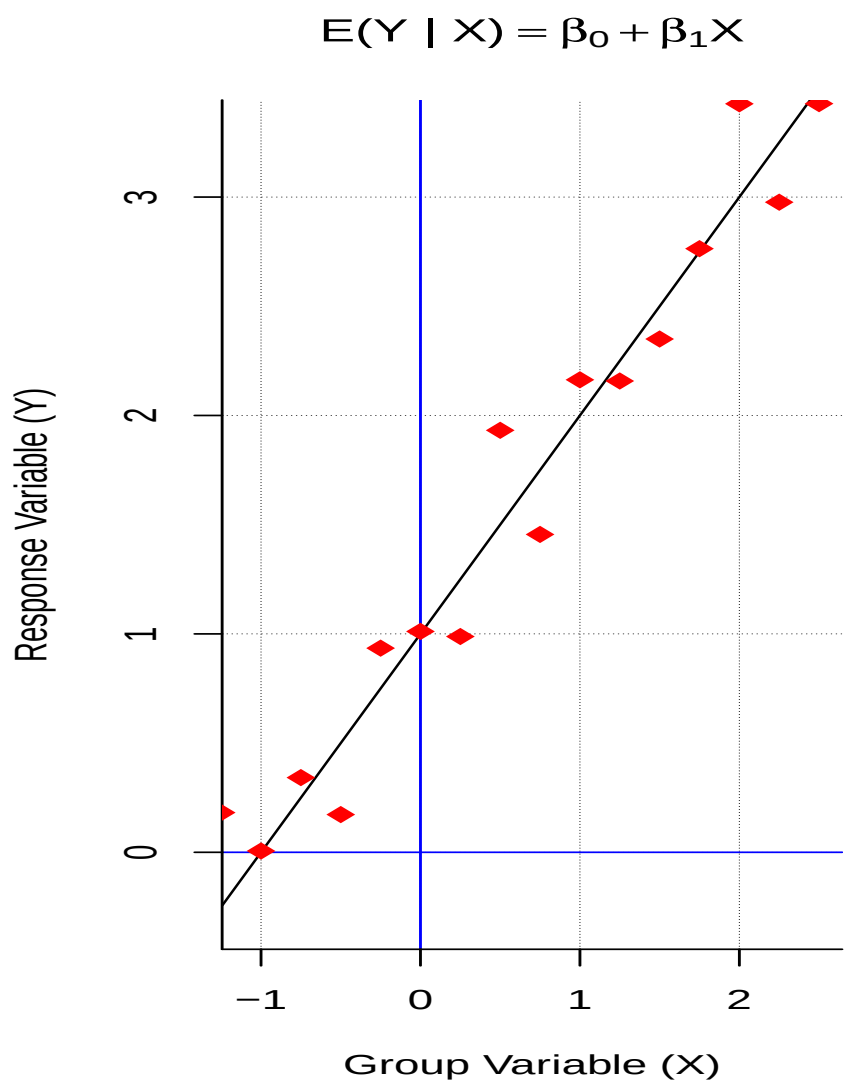
Preliminaries: Differences Between the Linear Equation & Simple Linear Regression

Relationships between Y and X:

- The mathematical linear equation: $(Y \mid X) = a + bX$
  - Deterministic relationship between Y and X
  - We know what Y is *exactly*, given X
  - Typically know a and b
- The simple linear regression model: $\mathbb{E}(Y \mid X) = \beta_0 + \beta_1 X$
  - Non-deterministic (stochastic) relationship between Y and X
  - We know what Y *tends to*, given X
    - the *averaged* Y at a given value of X
  - Typically don't know β_0 and β_1
    - need to estimate them

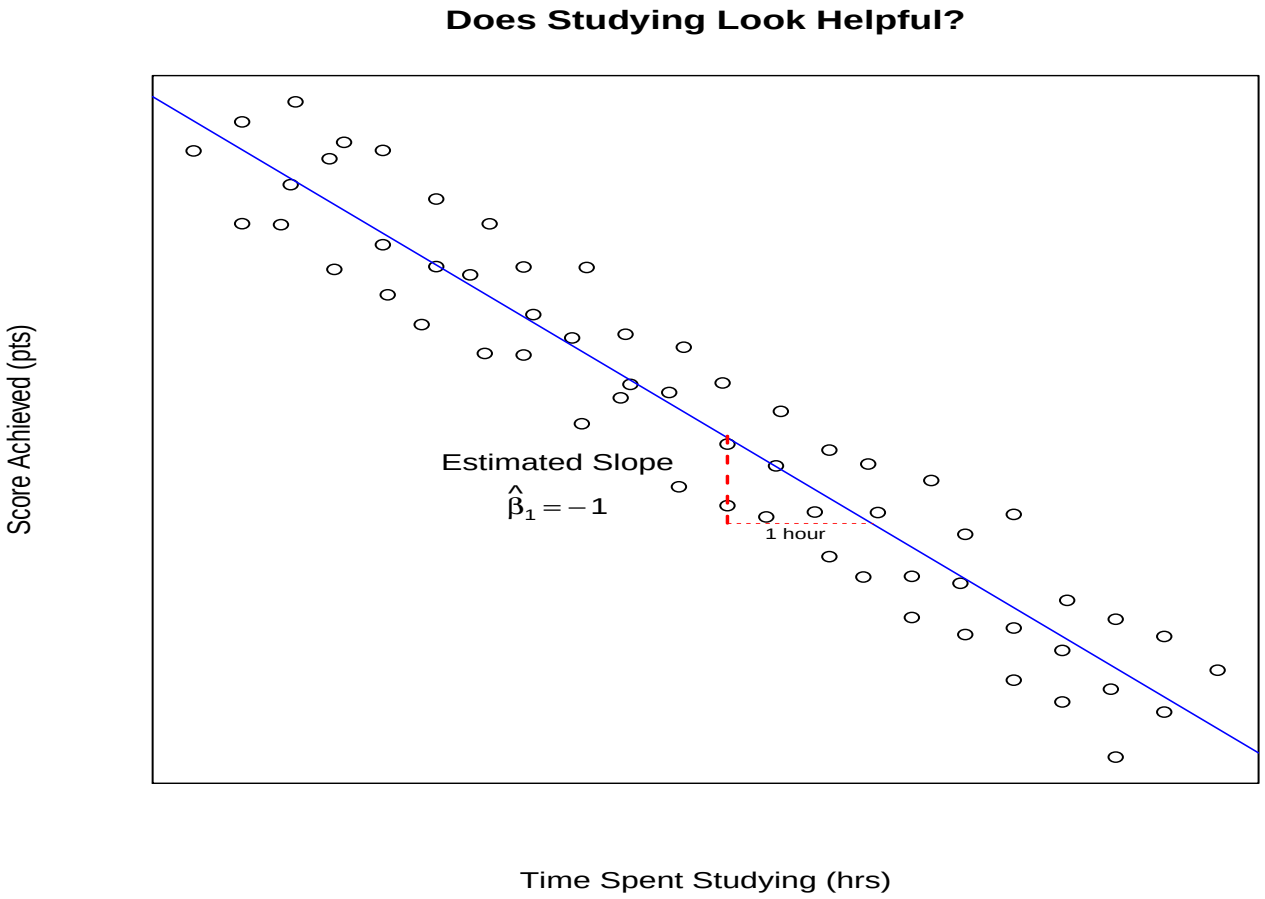
Preliminaries: Differences. . .





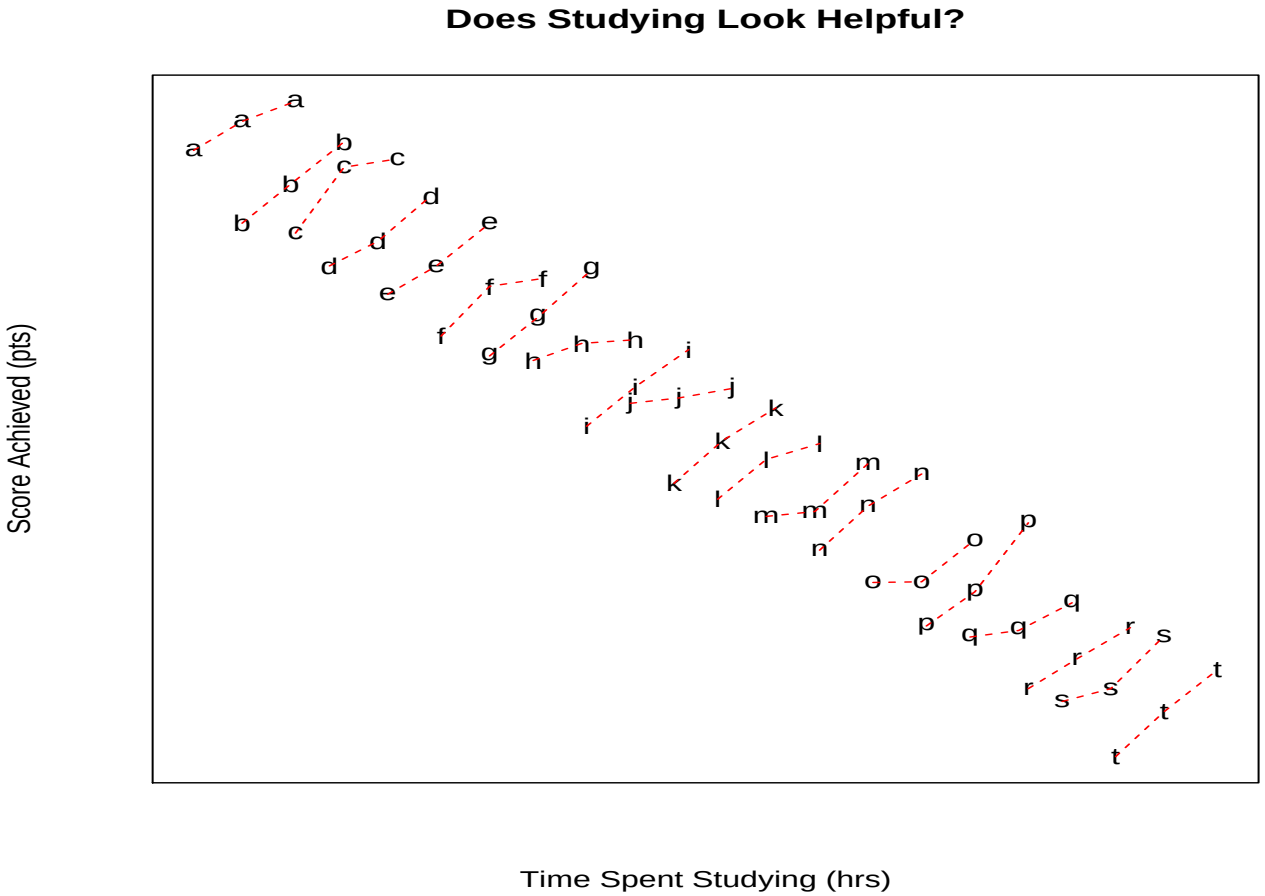
Example 1: Is Studying Helpful?

- Please interpret the regression coefficient, $\widehat{\beta}_1 = -1$, below



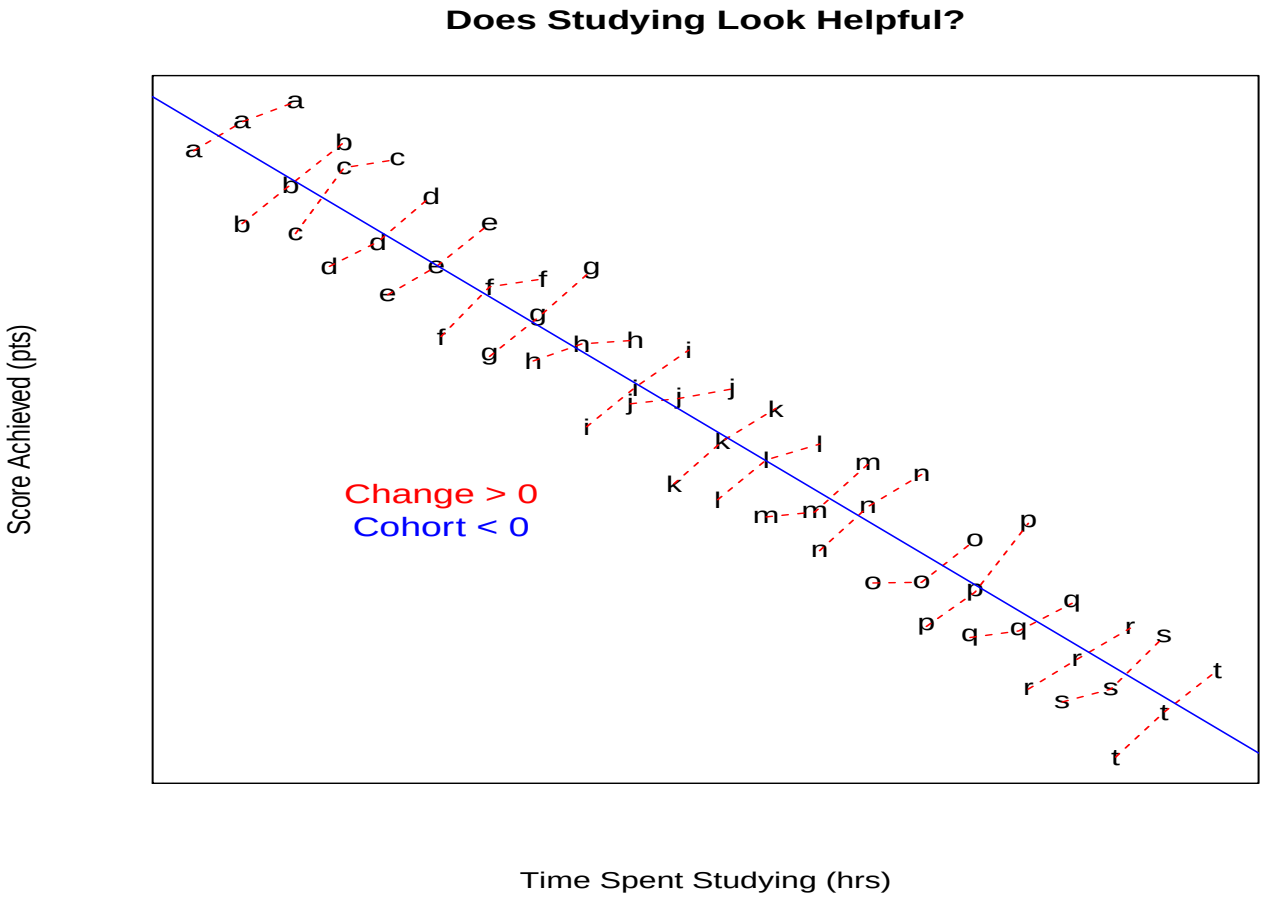
Example 1: Is Studying Helpful?

- Letters denote measurements for 20 different subjects



Example 1: Is Studying Helpful?

- Does your interpretation change now?



Example 1: Summary

Going back to our definition

$$\beta_1 = \mathbb{E}(Y \mid X = 1) - \mathbb{E}(Y \mid X = 0)$$

- β_1 is
  - the *difference* in means of the outcome, Y, comparing two groups that *differ* by one unit in X
  - not the *change* in the mean (of Y) obtained by *increasing* X by one unit
- As this example illustrates, the effects of change may differ from the observe association averaged over the population

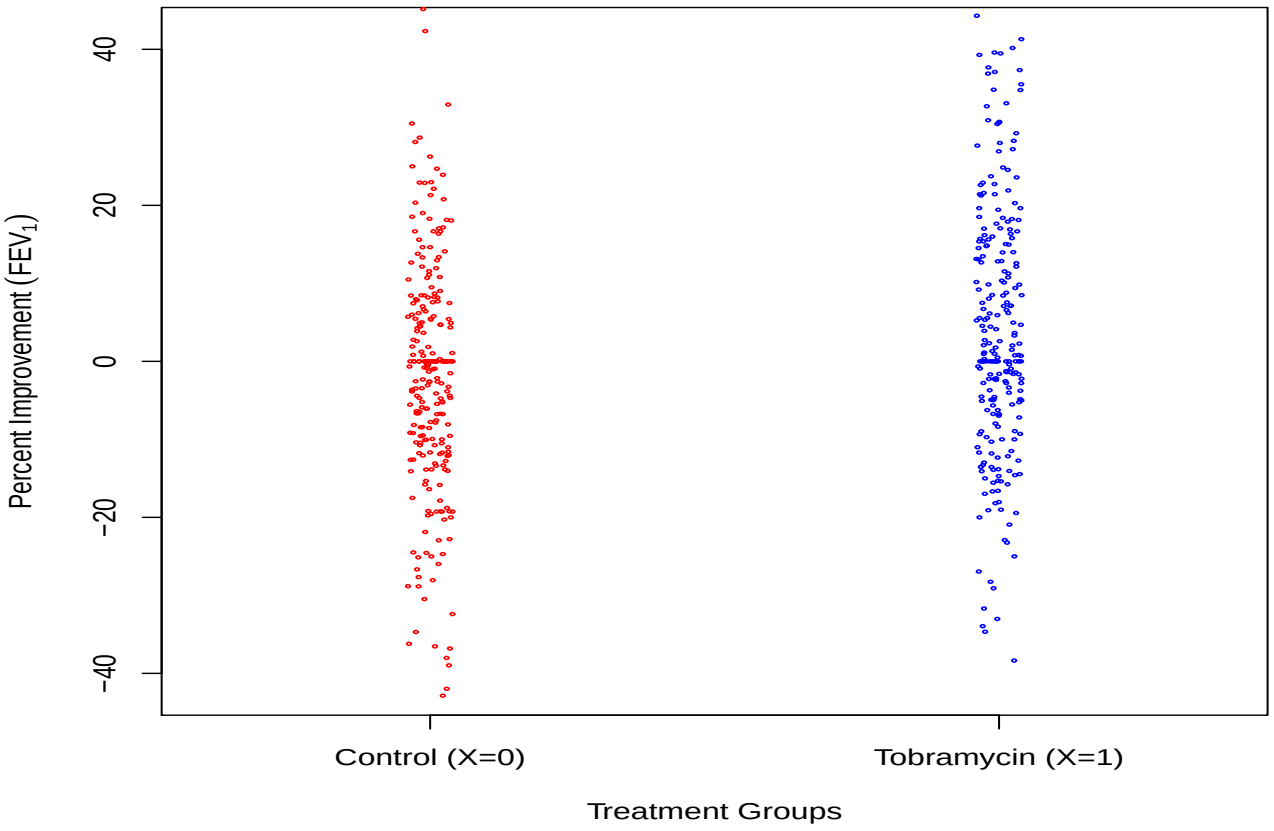
Example 2: CF clinical trial

Cystic fibrosis (CF) is a common serious genetic disorder, frequently complicated by recurrent pulmonary infection caused by *pseudomonas aruginosa*

- Study to determine if the aerosolized antibiotic, tobramycin, is efficacious in treating infection in patients with CF (Ramsey et al., 1999)
- N=520 CF patients (10-60 years) were randomized to receive tobramycin or placebo in a double-blind controlled trial
- Primary endpoint: percent improvement in FEV_1 (pulmonary fct. test)
- FEV_1 measurements were collected pre-randomization, and again at the end of the 24-week study
  - The outcome was percent improvement in FEV_1:

$$\%FEV_1 = 100 \times \frac{FEV_1(24) - FEV_1(0)}{FEV_1(0)}$$

Example 2: The Data



Example 2: CF clinical trial

We have a quantitative outcome, percent improvement in FEV_1 (i.e., $\%FEV_1$), and a binary (0/1) "grouping" or predictor variable, X

- **Q**: What is a reasonable approach to evaluating whether mean $\%FEV_1$ differs between the two treatment groups?
- **A**:

Example 2: CF clinical trial

We have a quantitative outcome, percent improvement in FEV_1 (i.e., $\%FEV_1$), and a binary (0/1) "grouping" or predictor variable, X

- **Q**: What is a reasonable approach to evaluating whether mean $\%FEV_1$ differs between the two treatment groups?
- **A**: A t-test!

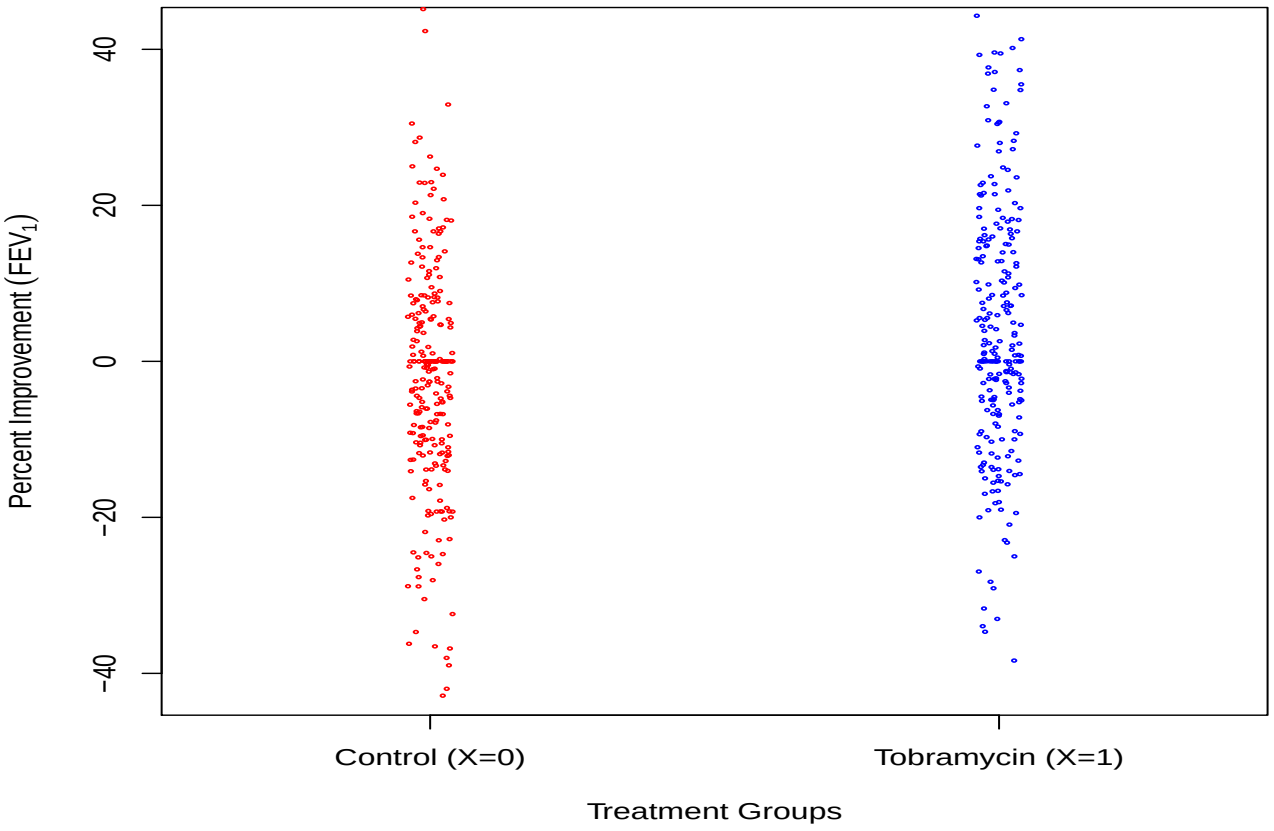
Example 2: CF clinical trial

The two-sample t-test:

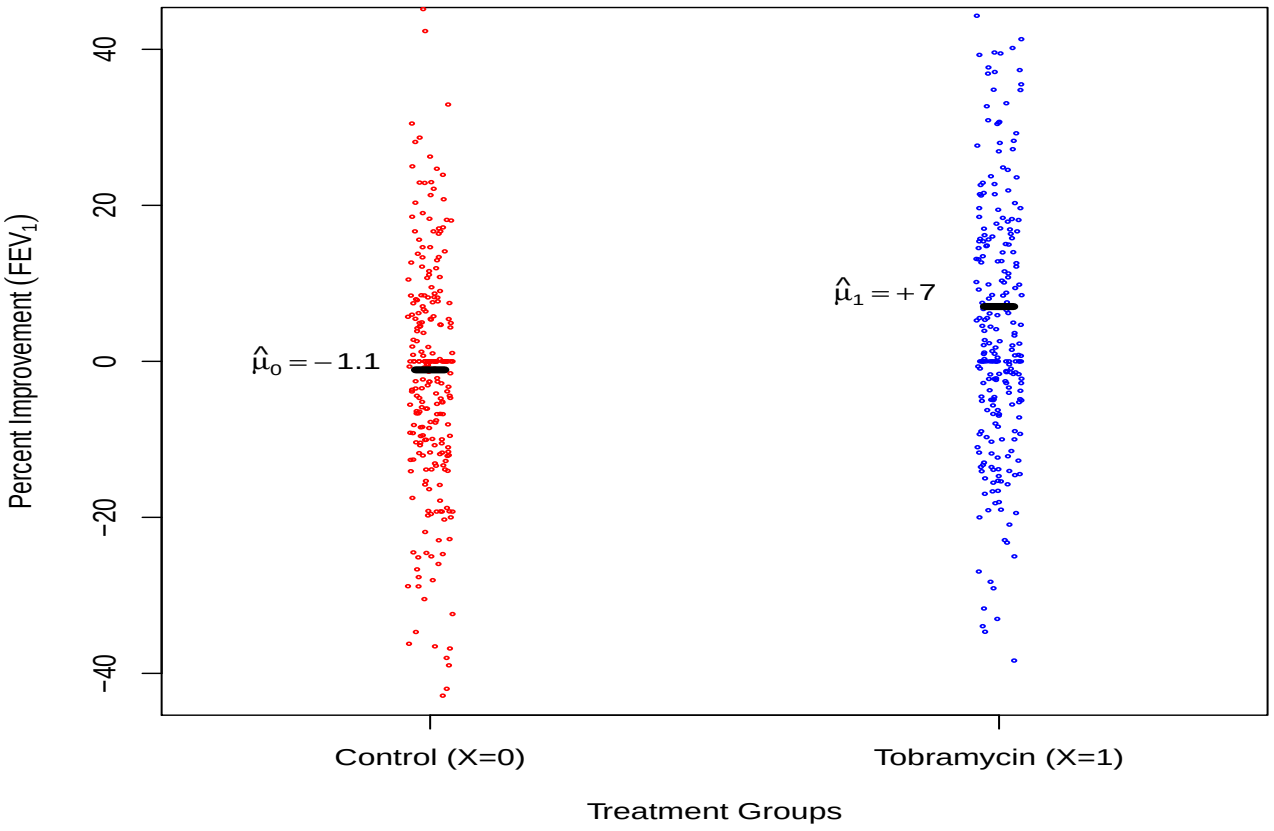
- We have 262 %FEV$_1$ values for the controls ($X = 0$), and 258 %FEV$_1$ values for the treatment group ($X = 1$)
- **Goal**: To evaluate whether the mean percent improvement in FEV$_1$ for patients treated with tobramycin differs from the mean percent improvement for the control patients
- Define
  - $\mu_0 = \mathbb{E}[\%\text{FEV}_1 \mid X = 0]$ = mean %FEV$_1$ for placebo patients
  - $\mu_1 = \mathbb{E}[\%\text{FEV}_1 \mid X = 1]$ = mean %FEV$_1$ for tobramycin patients

    * μ is the lowercase Greek letter (pronounced "mū")
    * *Often use Greek letters to denote unknown model parameters*
    * vertical bar " $\mid$ " means "conditioned on" or "given"
    * circumflex " $\widehat{\ }$ " used to denote sample estimates (e.g., $\widehat{\mu}_0, \widehat{\beta}_1$)

Example 2: The Data



Example 2: Model for t-test



Example 2: Model for t-test

Results:

- $\widehat{\mu}_0 = \widehat{\mathbb{E}}(\%\mathsf{FEV}_1 \mid X = 0) = -1.1\%$ for placebo patients
- $\widehat{\mu}_1 = \widehat{\mathbb{E}}(\%\mathsf{FEV}_1 \mid X = 1) = +7.0\%$ for tobramycin patients

- We estimate the difference in the two estimated means as

$$\begin{aligned}\widehat{\mu}_1 - \widehat{\mu}_0 &= \widehat{\mathbb{E}}(\%\mathsf{FEV}_1 \mid X = 1) - \widehat{\mathbb{E}}(\%\mathsf{FEV}_1 \mid X = 0) \\ &= 7.0\% - (-1.1)\% = 8.1\%\end{aligned}$$

- and estimate the precision of the mean difference (e.g., standard error) to formally test whether the difference is statistically different from zero

- **Q**: Can we evaluate this problem using regression?

Example 2: Simple Linear Regression approach

Recall the simple linear regression model:

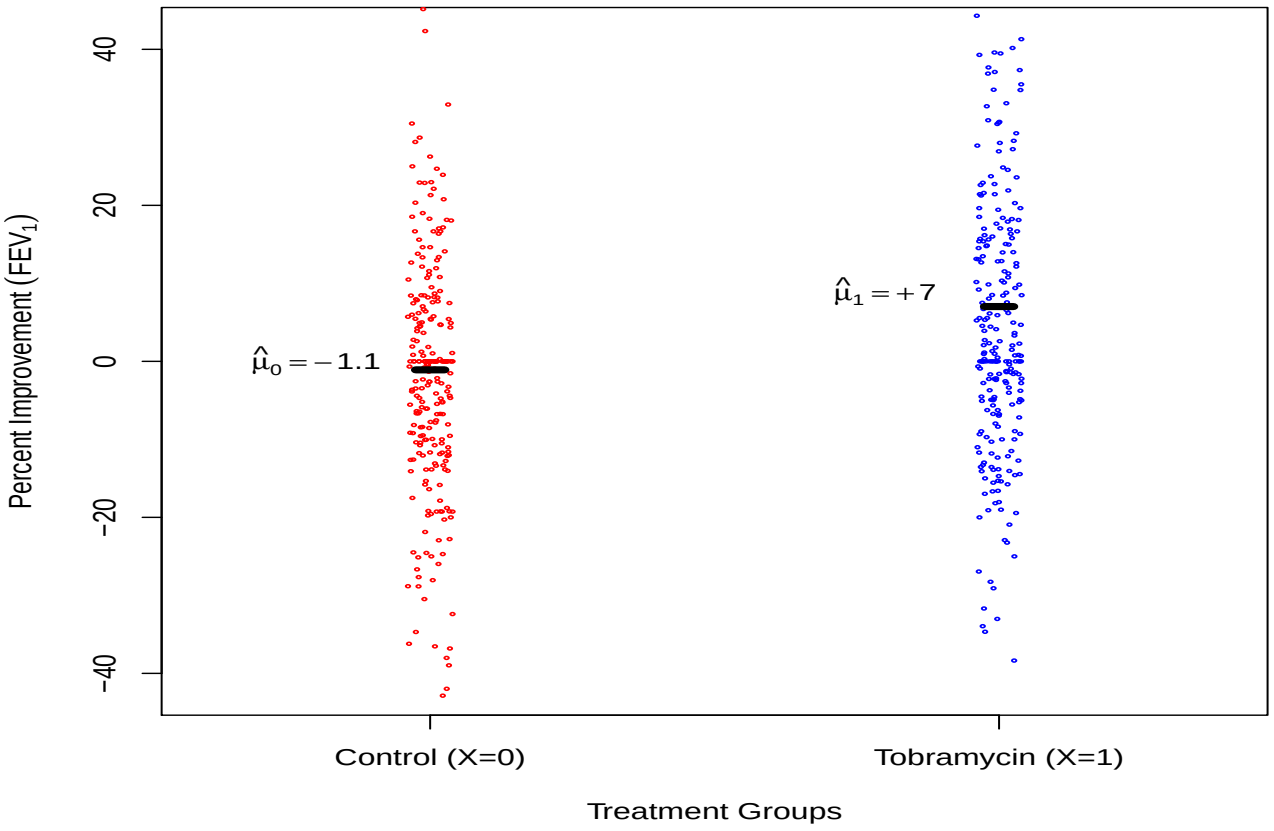
- $\mathbb{E}(\%\mathsf{FEV}_1 \mid X) = \beta_0 + \beta_1 X$
  - Controls: $\mathbb{E}(\%\mathsf{FEV}_1 \mid X = 0) = \beta_0$
  - Treatment: $\mathbb{E}(\%\mathsf{FEV}_1 \mid X = 1) = \beta_0 + \beta_1$
- This should look familiar
  - We model means for groups defined by the grouping variable, X

Example 2: Simple Linear Regression approach

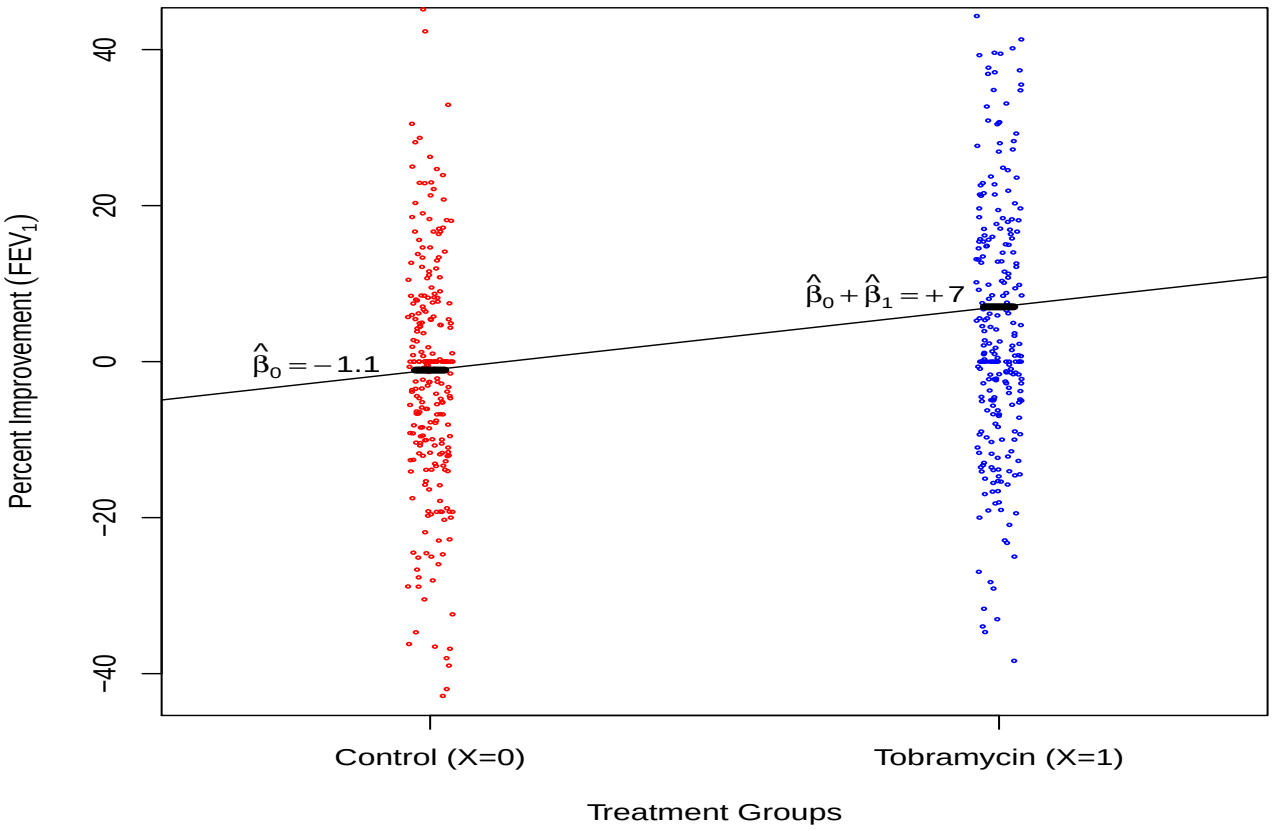
Now recall the model for the t-test:

- $\mathbb{E}(\%\mathsf{FEV}_1 \mid X) = \beta_0 + \beta_1 X$
  - Controls: $\mathbb{E}(\%\mathsf{FEV}_1 \mid X = 0) = \beta_0 = \mu_0$
  - Treatment: $\mathbb{E}(\%\mathsf{FEV}_1 \mid X = 1) = \beta_0 + \beta_1 = \mu_1$
- We have
  - $\beta_1 = \mathbb{E}(\%\mathsf{FEV}_1 \mid X = 1) - \mathbb{E}(\%\mathsf{FEV}_1 \mid X = 0) = \mu_1 - \mu_0$

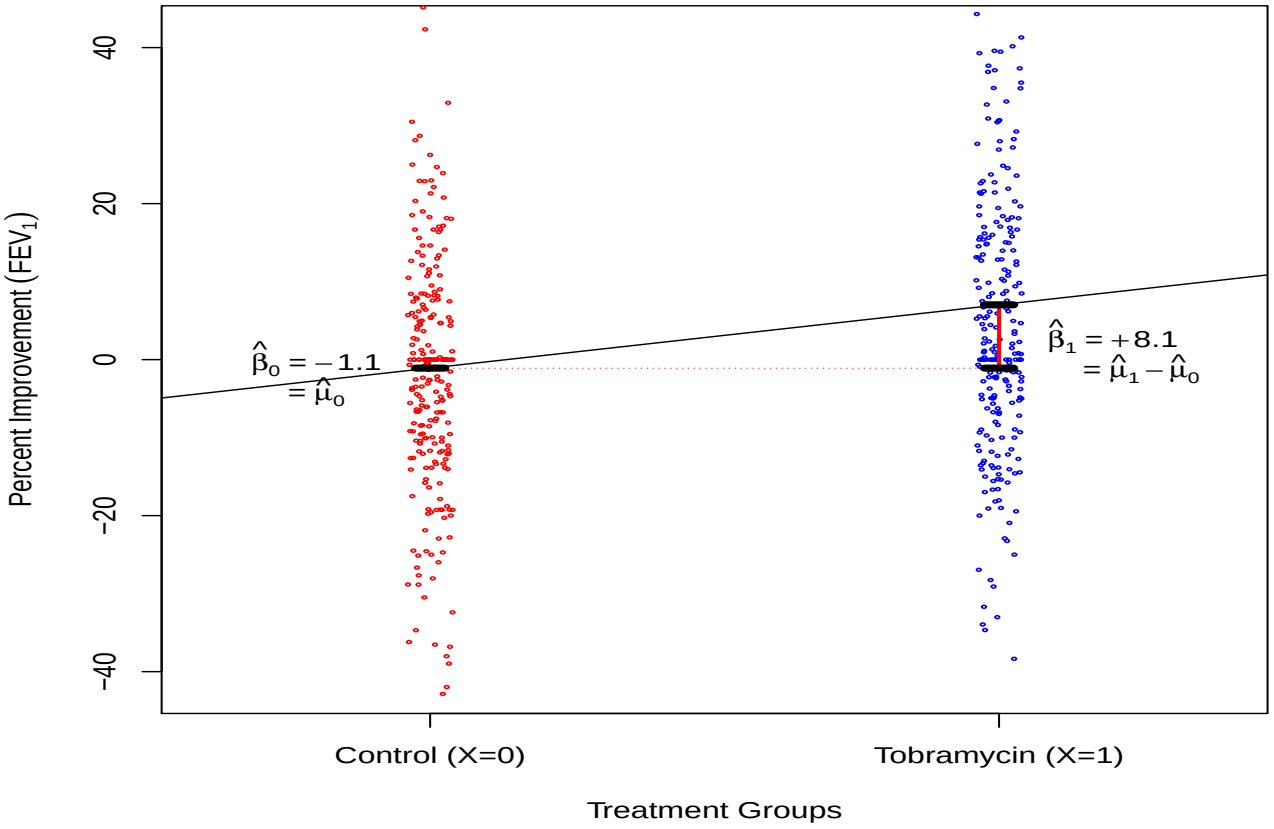
Example 2: Model for t-test



Example 2: Model for Simple Linear Regression



Example 2: Model for Simple Linear Regression



Example 2: Summary

The t-test and simple linear regression

- Yield exactly the same results
  - The t-test estimates the two group means directly
  - The regression estimates the mean of
    - a "referent" group (here, the X=0 group mean), and
    - the difference in the two group means
- Scientific interest is often evaluating a difference in the group means

- **Q**: How can we compare the means of more than two groups?
- **A**:

Example 2: Summary

The t-test and simple linear regression

- Yield exactly the same results
  - The t-test estimates the two group means directly
  - The regression estimates
    - a "referent" group (here, the X=0 group mean), and
    - the difference in the two group means
- Scientific interest is often evaluating a difference in the group means

- **Q**: How can we compare the means of more than two groups?
- **A**: ANOVA

Example 2: Summary

The t-test and simple linear regression

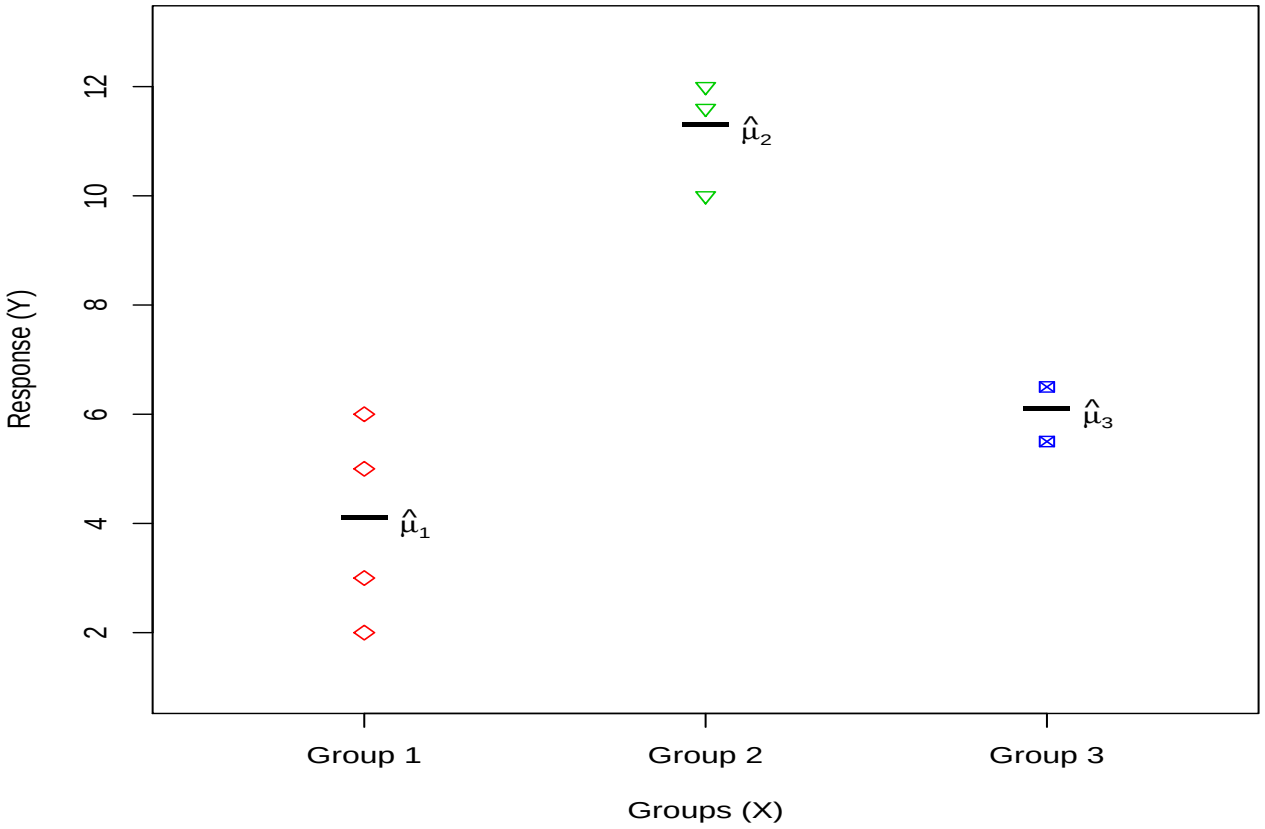
- Yield exactly the same results
  - The t-test estimates the two group means directly
  - The regression estimates
    - a "referent" group (here, the X=0 group mean), and
    - the difference in the two group means
- Scientific interest is often evaluating a difference in the group means

- **Q**: How can we compare the means of more than two groups?
- **A**: ANOVA and/or regression

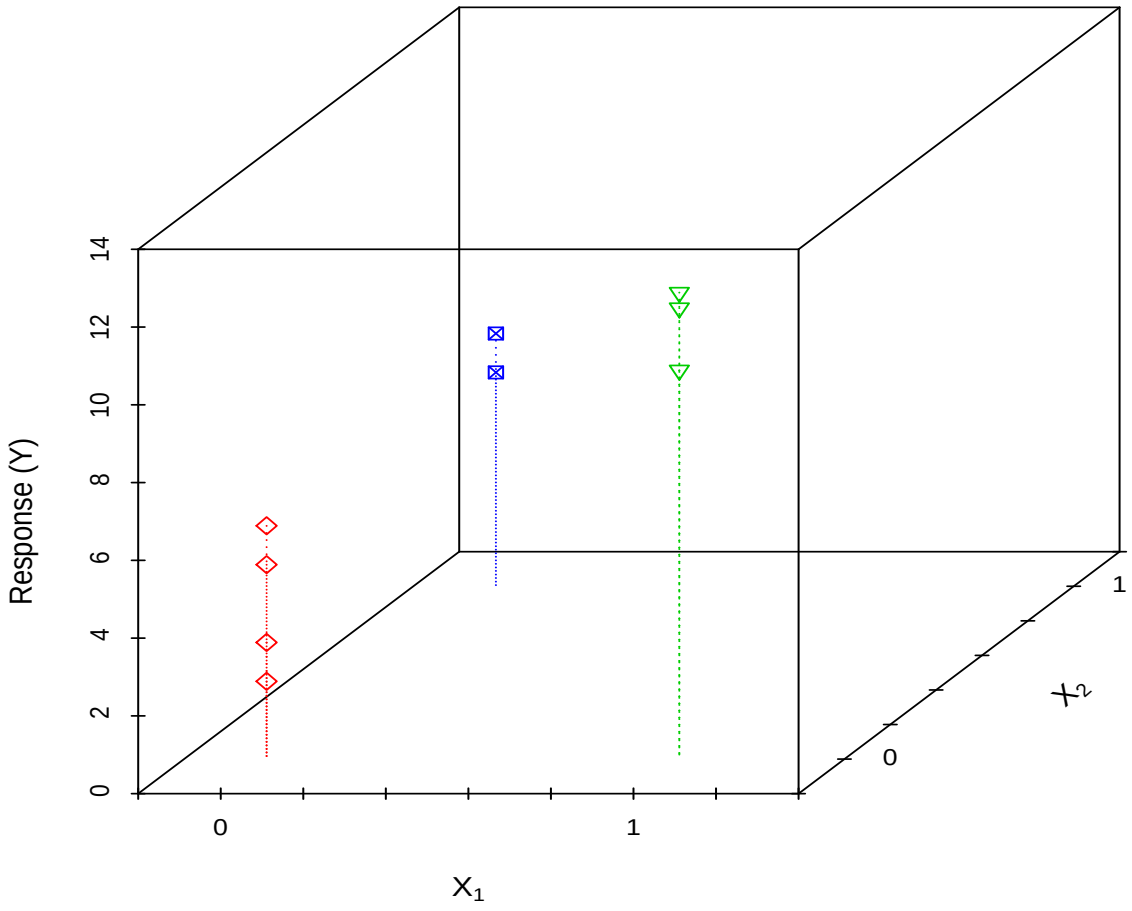
T-test, ANOVA & Regression

- The (model for the) t-test allows us to compare the means of two groups
- ANOVA is a generalization of the t-test for more than two groups
  - ANalysis Of VAriance
    - Formally developed by RA Fisher for designed experiments
    - Used to compare means of groups (2 or more)
    - Used as a form of hypothesis testing, like the t-test
    - ANOVA "table" group-level summaries in terms of variance estimates
- ANOVA & Regression Models (for 3 groups):
  - ANOVA: $\mathbb{E}(Y \mid \text{Group}_j) = \mu_j \qquad j = 1, 2, 3$
  - Regression: $\mathbb{E}(Y \mid X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$
    - X_1, X_2 are indicators that uniquely define the three groups

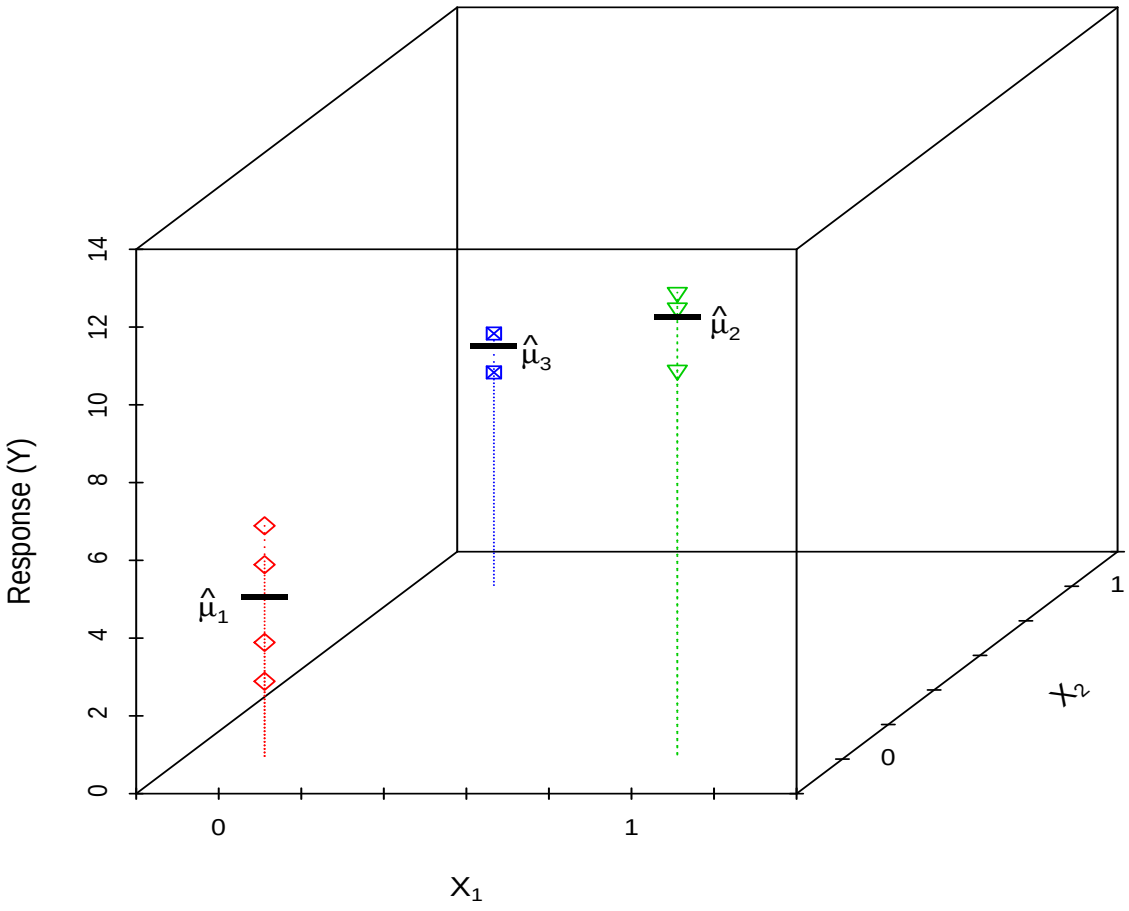
ANOVA (with 3 groups)



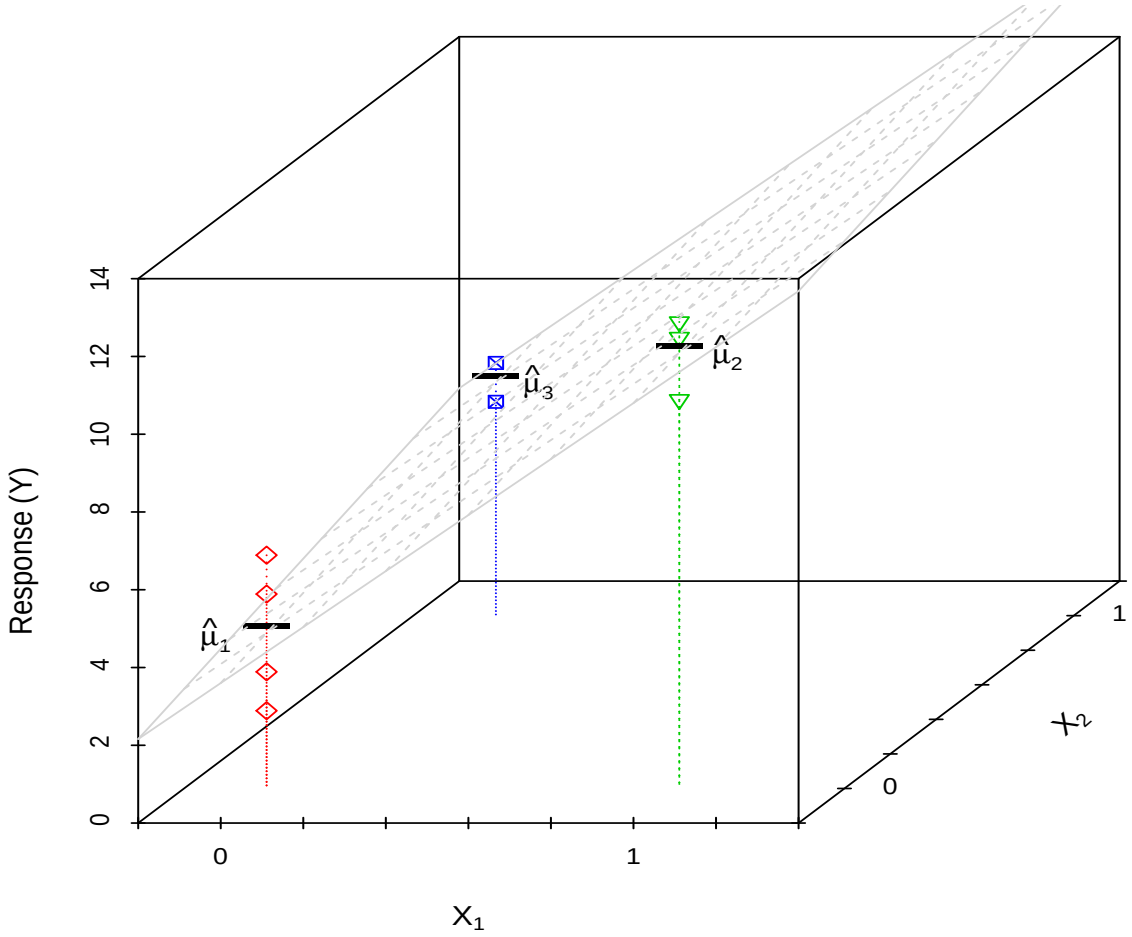
Data in 3D



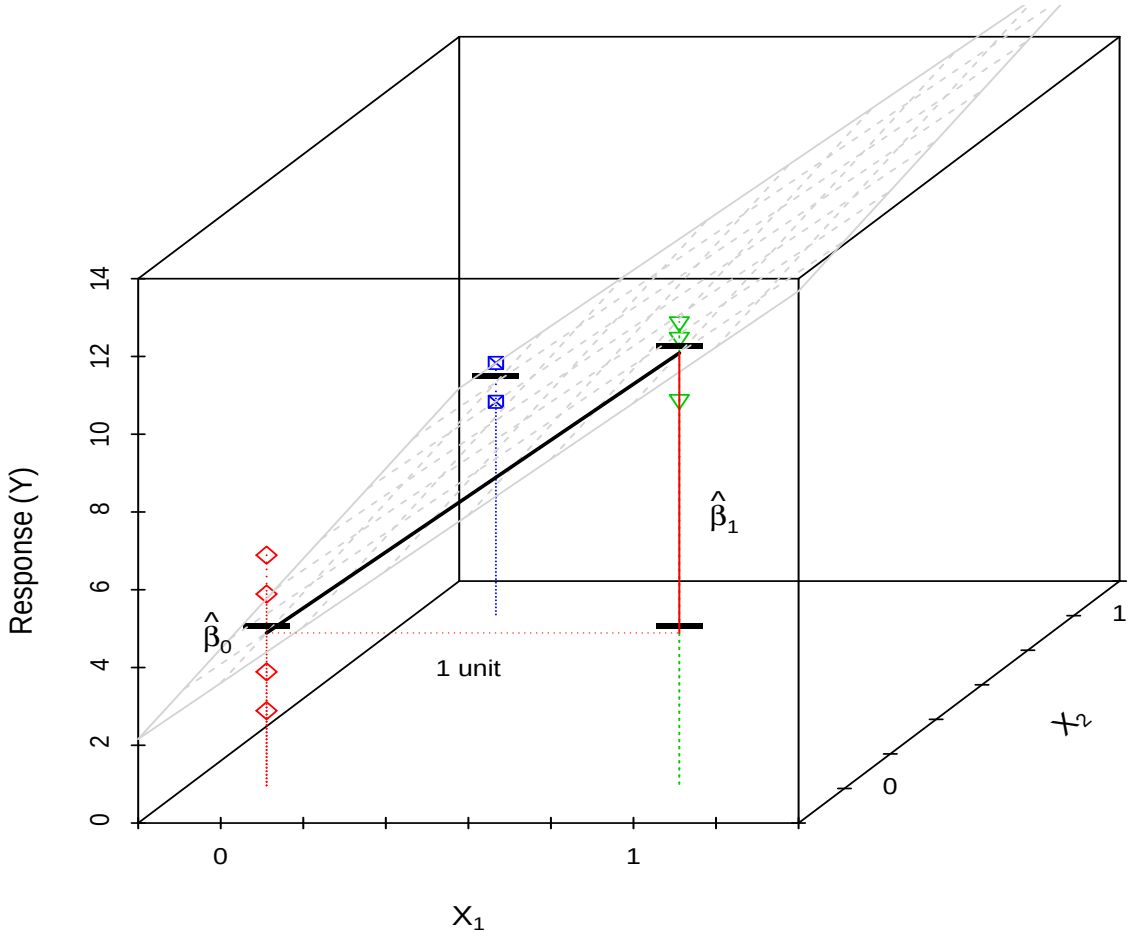
ANOVA in 3D



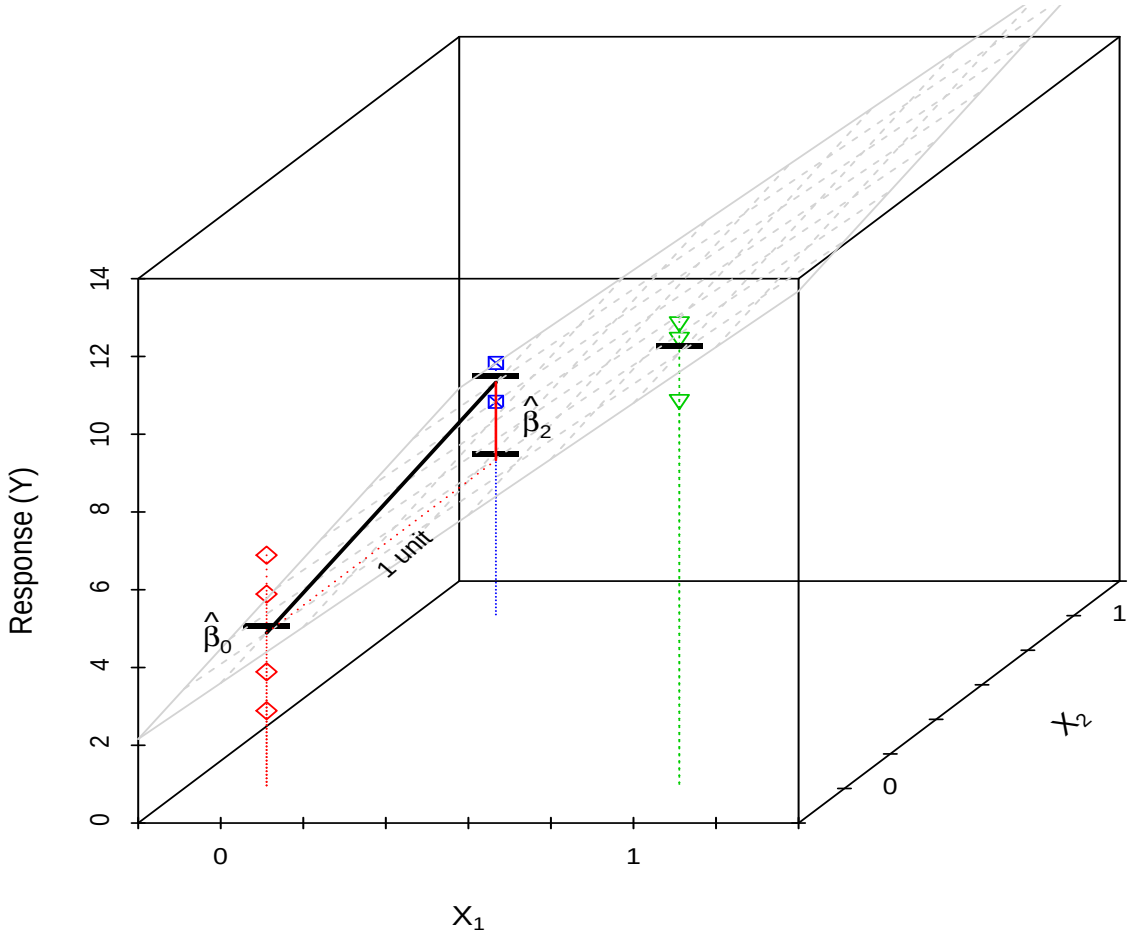
ANOVA with regression plane



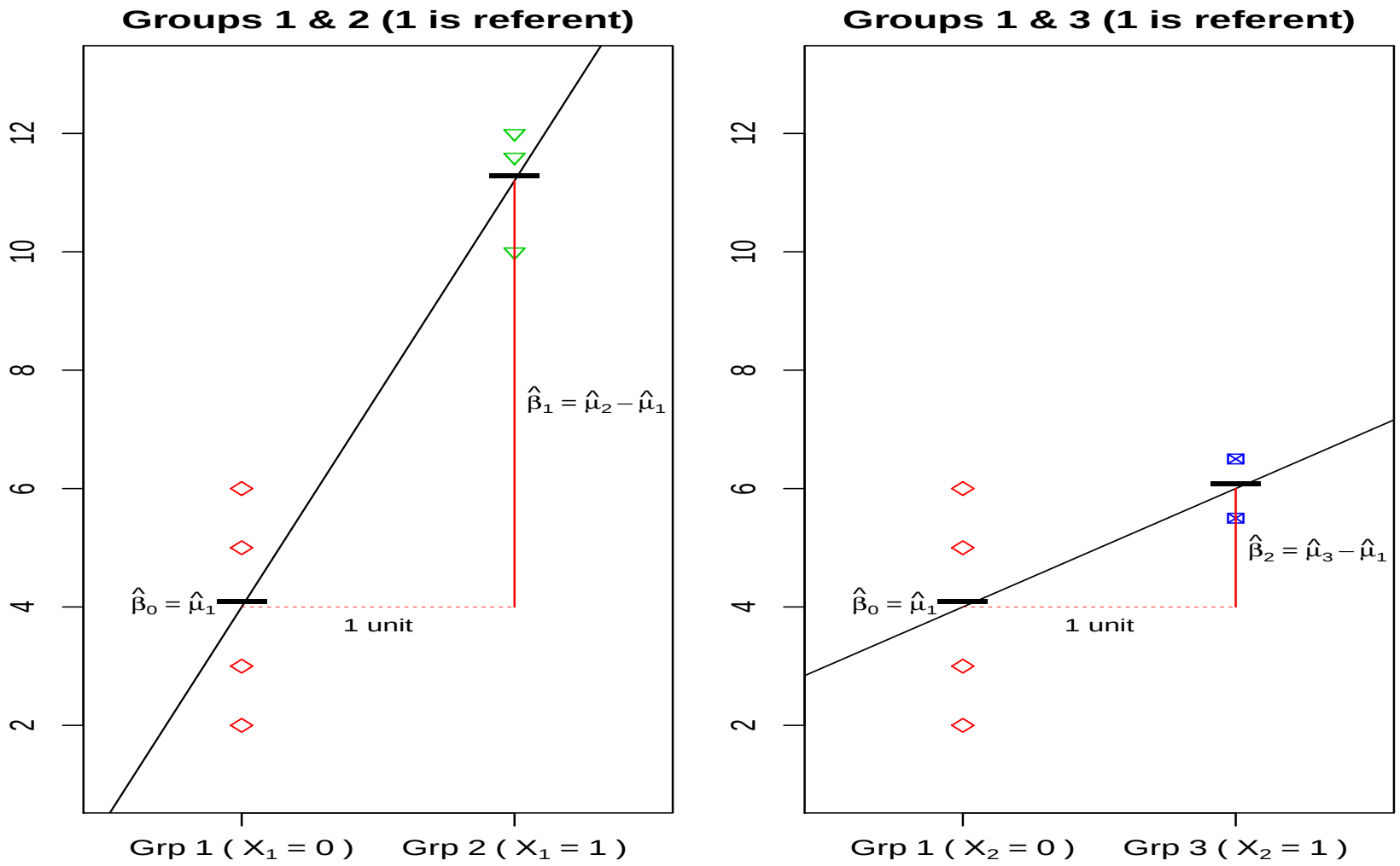
Regression: Groups 1 & 2



Regression: Groups 1 & 3



What regression does



T-test, ANOVA & Regression

Facts:

- T-test & ANOVA are special cases of regression
- T-test & ANOVA construct group means
- Regression constructs pairwise differences of group means

Number of Groups	T-test	ANOVA	Regression
2	μ_1 μ_2		$\beta_0 = \mu_1$ $\beta_1 = \mu_2 - \mu_1$
3		μ_1 μ_2 μ_3	$\beta_0 = \mu_1$ $\beta_1 = \mu_2 - \mu_1$ $\beta_2 = \mu_3 - \mu_1$
etc.			

The Multiple Regression Model

The Multiple Regression Model

- We often model the mean response across groups defined by multiple predictors
  - Simple linear regression: 1 predictor
    - E.g., compare the distribution of weight across groups defined by treatment status
  - Multiple regression: 2 or more predictors
    - E.g., compare the distribution of weight across groups defined by treatment status, age, gender

Interpretation of Regression Parameters

Difference in interpretation of the "slope" parameters

- Unadjusted model:

$$\mathbb{E}(WGT \mid TRT) = \beta_0 + \beta_1 TRT$$

  - β_1 is the difference in mean WGT for groups differing by 1 unit in TRT

$$\begin{aligned}\beta_1 &= \mathbb{E}(WGT \mid TRT = 1) - \mathbb{E}(WGT \mid TRT = 0) \\ &= \mathbb{E}(WGT \mid \text{treatment}) - \mathbb{E}(WGT \mid \text{controls})\end{aligned}$$

- Adjusted model:

$$\mathbb{E}(WGT \mid TRT, FEM) = \gamma_0 + \gamma_1 TRT + \gamma_2 FEM$$

  - γ_1 is the difference in mean WGT for groups differing by 1 unit in TRT **and** they have the same gender

$$\begin{aligned}\gamma_1 &= \mathbb{E}(WGT \mid TRT = 1, FEM) - \mathbb{E}(WGT \mid TRT = 0, FEM) \\ &= \mathbb{E}(WGT \mid \text{treatment, females}) - \mathbb{E}(WGT \mid \text{controls, females})\end{aligned}$$

Impacts of Covariate Adjustment

- The focus of why we adjust for covariates (aka predictors) is thus on
  - The scientific interpretation of the slope parameter estimates
  - Bias of these estimates relative to the scientific parameter of interest
  - The precision of the estimates

Reasons for Adjusting for Covariates

Adjusting for Covariates

- In order to assess whether we adjust for covariates, we must consider our beliefs about the causal relationships among the measured variables
  - We will not be able to assess causal relationships in our statistical analysis
  - Inference of causation comes only from study design
- However, consideration of hypothesized causal relationships helps us decide which statistical question to answer

Causation versus Association

- Statistical analysis can only detect associations reflecting causation in either direction
- Only experimental design and understanding of the variables allows us to infer cause and effect
- Statistical analysis will identify “causation” in either direction
  - We regard that causes of events must be in the correct temporal sequence

Causation versus Association

- Sometimes we can isolate particular pathways of scientific interest by including a third variable into an analysis
  - "Adjusting" for an effect of a third variable
    - Strata are defined based on the value of the third variable
    - Comparisons of the response distribution across groups defined by the predictor of interest are made within strata
    - The effects within strata are then averaged in some way to obtain the adjusted association

Causation versus Association

- Clearly, such adjustment makes most sense only when the association between response and predictor of interest is the same in each stratum
  - If there are different effects across strata, modeling an interaction would be indicated
    - The question should essentially be answered in each stratum separately

Adjustment for Covariates

We include predictors in a regression model for a variety of reasons

- In order of importance
  - Scientific question
    - Predictor(s) of interest
    - Effect modifiers
  - Adjustment for confounding
  - Gain precision
- Adjustment for covariates changes the question being answered by the statistical analysis
  - Adjustment can be used to isolate associations that are of particular interest

Scientific Question

Many times the scientific question dictates inclusion of particular predictors

- Predictor(s) of interest
  - The scientific factor being investigated can be modeled by multiple predictors
    - E.g., indicator (dummy) variables for multiple treatment groups
- Effect modifiers
  - The scientific question may relate to detection of effect modification
- Confounders
  - The scientific question may have been stated in terms of adjusting for known (or suspected) confounders

Confounding

Definition of confounding

- The association between a predictor of interest and the response variable is confounded by a third variable if
  - The third variable is associated with the predictor of interest in the sample, AND
  - The third variable is associated with the response
    - causally (in truth)
    - in groups that are homogeneous with respect to the predictor of interest, and
    - not in the causal pathway of interest

Confounding

Symptoms of confounding

- Estimates of association from unadjusted analysis are markedly different from estimates of association from adjusted analysis
  - Association within each stratum is similar to each other, but different from the association in the combined data
- In linear regression, these symptoms are diagnostic of confounding
  - Effect modification would show differences between adjusted analysis and unadjusted analysis, but would also show different associations in the different strata
- Note that confounding produces a difference between unadjusted and adjusted analyses, but those symptoms are not proof of confounding

Confounding

Effect of confounding

- A confounder can make the observed association between the predictor of interest and the response variable look
  - stronger than the true association,
  - weaker than the true association, or
  - even the reverse of the true association

Confounding

Some times the scientific question of greatest interest is confounded by unexpected associations in the data

- Confounders
  - Variables (causally) predictive of outcome, but not in the causal pathway of interest
    - (Often assessed in the control group)
  - Variables associated with the predictor of interest in the sample
    - Note that statistical significance is not relevant, because that tells us about associations in the population
  - Detecting confounders must ultimately rely on our best knowledge about possible mechanisms

Precision

- Sometimes we choose the exact scientific question to be answered on the basis of which question can be answered most precisely
  - In general, questions can be answered more precisely if the within group distribution is less variable
    - Comparing groups that are similar with respect to other important risk factors decreases variability

Precision

- Two special cases to consider when attempting to gain precision in a model
  - If stratified randomization or matched sampling was used in order to address possible confounding and / or precision issues, the added precision will NOT be realized UNLESS the stratification or matching variables are adjusted for in the analysis
  - If baseline measurements are available, it is more precise to adjust for those variables as a covariate than to analyze the change

Adjusting for Covariates: Confounding, Precision and Effect Modification

Discriminating between confounding, precision, and effect modifying variables

- Is the estimate of association between response and the predictor of interest the same in all strata?
  - Effect modifier: NO; Confounder, precision: YES
- Is the third variable causally associated with the response after adjusting for the predictor of interest?
  - Confounder, precision: YES
- Is the third variable associated with the predictor of interest?
  - Confounder: YES; Precision: NO

Summary: Adjustment for Covariates

- When I consult with a scientist, it is often very difficult to decide whether the interest in additional covariates is due to confounding, precision, or effect modification
  - We illustrate the difference between precision variables, confounders, and effect modifiers in the following hypothetical example

Weight Loss Study

Scientific Question

- Is there an association between weight and an experimental weight loss therapy in adults?

Causal Pathway of Interest

- We are interested in knowing if an experimental therapy will cause a decrease in weight for adults, as measured by body weight (in pounds)

Causation versus Association

- Statistical analyses, however, can only detect associations between treatment exposure and weight

  - In a randomized trial, we could infer from the design that any association must be causal
  - In an observational study, we must try to isolate causal pathways of interest by adjusting for covariates

Study Design

Observational Study

- Measurements on n=500 adults, ages 20 to 45 years
- Obtain 250 treatment volunteers & sampled 250 control subjects
  - Predictor of interest (POI): experimental treatment/control (TRT = 1/0)
  - Response: Self-reported weight at 6 months (in pounds)
  - Additional covariates: subject (female) gender & age (in years)
    - Effect modifiers
    - Potential confounders
    - Precision variables

Additional Covariates: Effect Modifiers

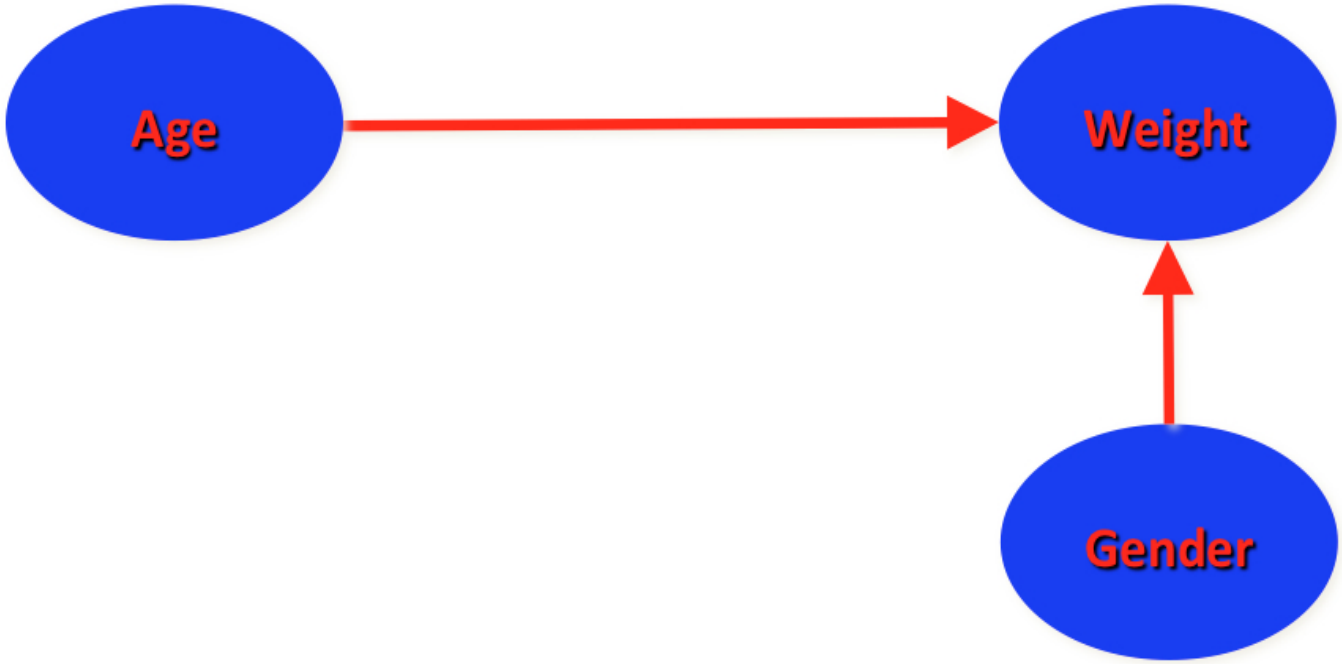
- There are no covariates of major scientific interest for their potential for effect modification
  - Perhaps secondary interest in differential weight loss due to treatment by gender and age groups
- First things first
  - Not generally advisable to go looking for different effects of treatment in subgroups before we have established that an effect exists overall
    - (We may sometimes delay discovery of important facts, but most times this seems the logical strategy)

Additional Covariates: Confounders

- **Think** about potential confounders
  - Necessary requirements for confounders
    - Associated causally with response
    - Associated with predictor of interest in the sample
    - Does not lie in the causal path
  - Prior to looking at data, we cannot be sure of the second criterion
    - But, clearly, any strong predictor of the response has the potential to be a confounder
      - So first consider known predictors of the response
    - Furthermore, in an observational study, known associations in the population will likely also be in the sample

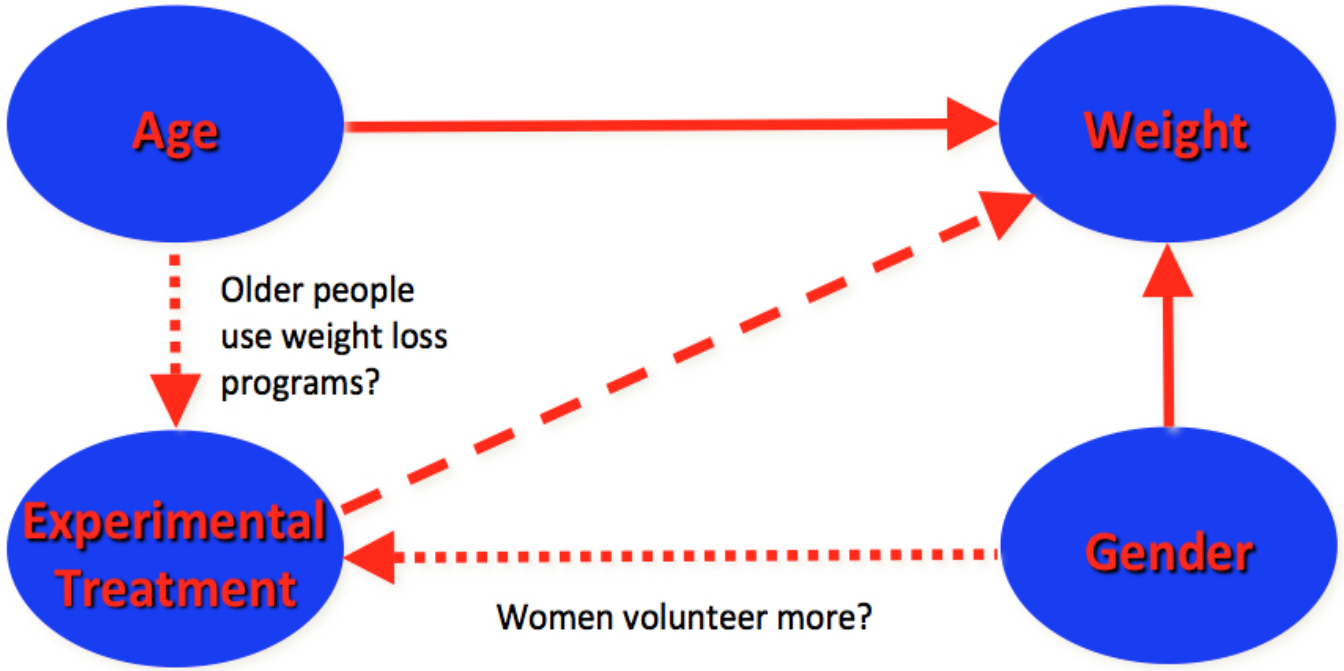
Predictors of Weight

- "Known" predictors of weight



Associations with Weight Loss Treatments

- "Known" associations with weight loss treatments in the population



Adjusting for Potential Confounders

- Investigating the effect of the experimental treatment on weight in adults
  - We are scientifically interested in the possibility that the treatment might cause a decrease in weight
  - We are not scientifically interested in showing that a person's weight might influence treatment propensity
    - (Of course, this is one possible explanation of an observed association, and so we must try to rule this out)

Additional Covariates: Precision Variables

- **Think** about major predictors of the response
  - In an observational study, all predictors of response should be considered potential confounders
  - However, even if strong predictors of response are not confounding (i.e., not associated with POI in sample), we might want to consider adjusting for them in the analysis to gain precision
    - In the weight loss study, age might be a strong predictor of weight
      - It is less clear that age might influence treatment propensity

Planned Analysis: Covariate Adjustment

Based on these issues, a priori we might plan an analysis adjusting for gender and age

- If that had not been specified a priori, I would perform the unadjusted analysis and then report the observed confounding from exploratory analyses
  - **Data driven analyses always provide less confidence than pre-specified analyses**
- In order to illustrate the effects of adjusting for confounders and precision variables, we will explore several analyses
  - Variable for experimental treatment:
    - (TRT) coded 0= control group, 1= treatment group

Planned Analysis: Summary Measures

- Based on the scientific relationship between weight and its strong predictors (gender and age), we will compare mean weights
  - had we collected subjects' weights at the beginning of the study, we might have considered evaluating change in weight or include baseline weight as an adjustment variable
    - In a randomized study, baseline weights are likely to be similar across exposure groups; modeling change or adjusting for baseline will likely improve precision (Frison & Pocock, *Stats in Med*, 1992) and the "treatment effect" will be similar from model to model
    - In a non-randomized study, baseline weights are likely to differ by exposure groups, and the interpretations of the "treatment effect" will differ from model to model (Fitzmaurice, *Nutrition*, 2001)
- Such an analysis is easily performed and interpreted

Unadjusted Analysis: Stata Output

```
. regress wgt i.TRT, robust

Number of obs      =        500
Root MSE           =     30.236

        |          Robust
    wgt |  Coef.    St Err       t     P>|t|      [95% CI]
--------+-----------------------------------------------------
    TRT |
   Yes  |  -8.75   2.70    -3.24   0.001   -14.07   -3.44
  _cons | 191.19   1.99    96.06   0.000   187.28  195.10
--------------------------------------------------------------
```

Unadjusted Analysis: Interpretation

- Treatment effect
  - The estimated mean weight of adults on the experimental treatment is 8 3/4 pounds lower than adults not on the treatment (95% CI: -14.1 to -3.4 pounds)
    - These results are atypical of what we might expect with no true difference between the groups (P = 0.001)
  - (Because TRT is a binary (0-1) variable, this analysis is nearly identical to a two-sample t test allowing for unequal variances)

Unadjusted Analysis: Interpretation

- Intercept
  - The estimated mean weight of adults not on the experimental treatment is 191.2 pounds (95% CI: 187.3 to 195.1 pounds)
    - The scientific relevance is questionable here because we do not know the population our sample represents
      - Comparing the treated to the untreated subjects is more useful than looking at either group by itself
    - (The P value is of no importance; it is testing that the group's mean weight is zero. Why would we care?)
  - (Because TRT is a binary variable, the estimate corresponds to the "referent" group)

Age Adjusted Analysis: Stata Output

```
. regress wgt i.TRT AGE, robust

Number of obs      =        500
Root MSE           =     28.882

        |          Robust
    wgt |  Coef.   St Err      t     P>|t|      [95% CI]
--------+----------------------------------------------------
    TRT |
   Yes  |  -9.32   2.58    -3.61   0.000   -14.39    -4.24
        |
    AGE |   1.27   0.19     6.79   0.000     0.90     1.64
  _cons | 149.56   6.43    23.26   0.000   136.93   162.20
-------------------------------------------------------------
```

Age Adjusted Analysis: Interpretation

- Treatment effect
  - The estimated mean weight of adults on the experimental treatment is 9.3 pounds lower than adults not on the treatment of the same age (95% CI: -14.4 to -4.2 pounds)
    - These results are highly atypical of what we might expect with no true difference in mean weight between the treatment groups ($P < 0.001$)
- Age effect
  - The estimated mean weight of adults is 1.3 pounds higher for each year difference in age between two groups with the same treatment status (95% CI: 0.9 to 1.6 pounds higher for each year difference in age)
    - These results are highly atypical of what we might expect with no true difference in the mean weight between age groups having the same treatment status ($P < 0.001$)

Age Adjusted Analysis: Interpretation

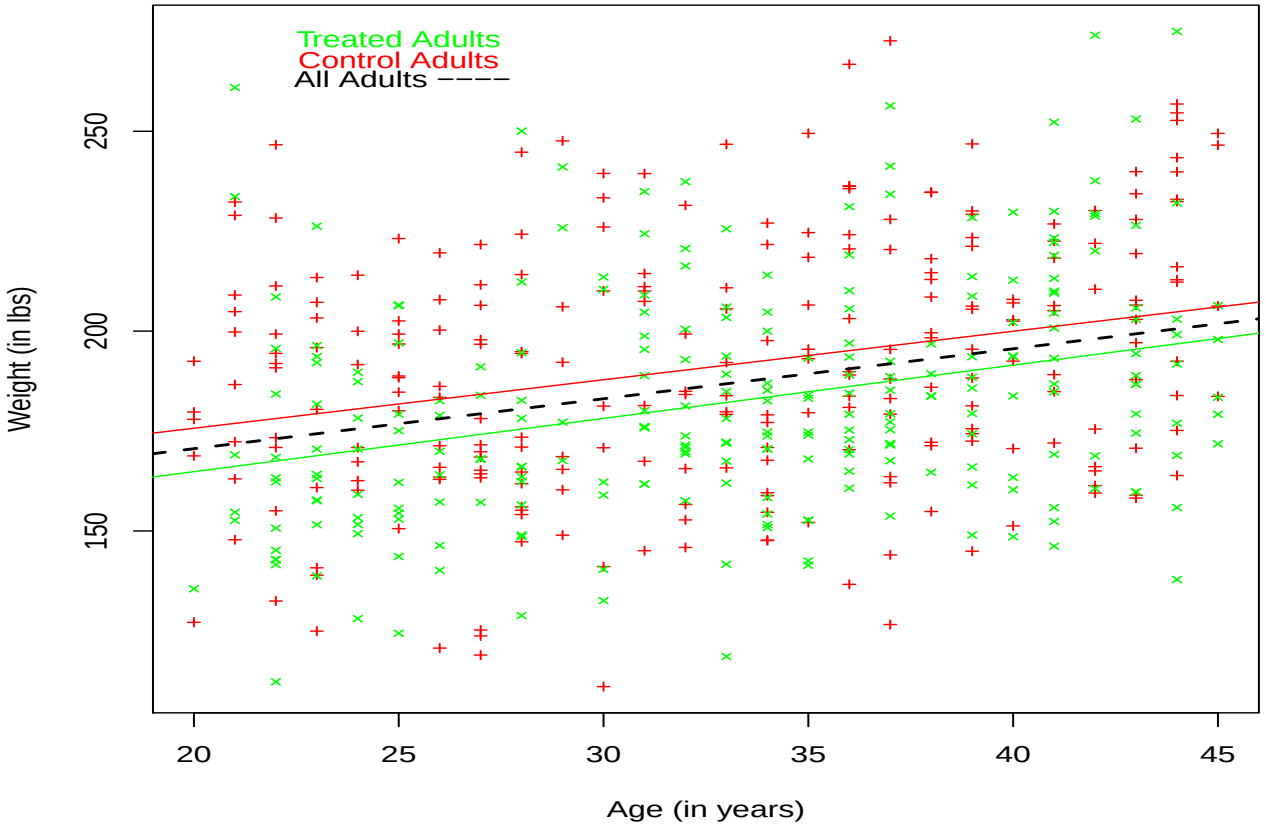
- Intercept
  - The estimated mean weight of newborn adults not on the experimental treatment is 149.6 pounds (95% CI: 136.9 to 162.2 pounds)
    - The intercept corresponds to the mean weight of adults not on the treatment and 0 years of age
    - There is no scientific relevance here because we are extrapolating outside of the range of our data (i.e., 20 to 45 year olds)

Age Adjusted Analysis: Comments

- Comparing unadjusted and age adjusted analyses
  - Little difference in effect of treatment suggests that there was little confounding by age
    - Age is a relatively strong predictor of weight
    - Age is not associated with treatment in the sample
      - Mean (SD) of age in analyzed adults on treatment: 33.2 (6.9) years
      - Mean (SD) of age in analyzed adults not on treatment: 32.8 (7.3) years
  - Effect of age adjustment on precision
    - Lower Root MSE (28.9 vs 30.2) would tend to increase the precision of estimate of the treatment effect
    - Lack of an association between treatment and age tends to not impact precision
    - Net effect: More precision ($\widehat{\mathrm{SE}}(\hat{\beta}_{\mathrm{TRT}})$: 2.58 vs 2.70)

Age Adjusted Analysis: Comments

- Strong association with weight
  - Little difference in associations with weight across treatment groups



Gender Adjusted Analysis: Stata Output

```
. regress wgt i.TRT i.FEM, robust

Number of obs     =         500
Root MSE          =     24.029

             |               Robust
         wgt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         TRT |
         Yes |       1.42   2.300000     0.62   0.538    -3.090000    5.930000
             |
         FEM |
         Yes |     -39.11   2.470000   -15.85   0.000    -43.96000   -34.26000
       _cons |     210.12   2.020000   103.99   0.000     206.1500    214.0900
------------------------------------------------------------------------------
```

Gender Adjusted Analysis: Interpretation

- Treatment effect
  - The estimated mean weight of adults on the experimental treatment is 1.4 pounds higher than adults not on the treatment of the same gender (95% CI: -3.1 to 5.9 pounds)
    - These results are typical of what we might expect with no true difference in mean weight between the treatment groups (P = 0.538)
- Gender effect
  - The estimated mean weight of adult females is 39.1 pounds lower than adult males with the same treatment status (95% CI: -44.0 to -34.3 pounds lower)
    - These results are highly atypical of what we might expect with no true difference in the mean weight between women and men having the same treatment status (P $<$ 0.001)

Gender Adjusted Analysis: Interpretation

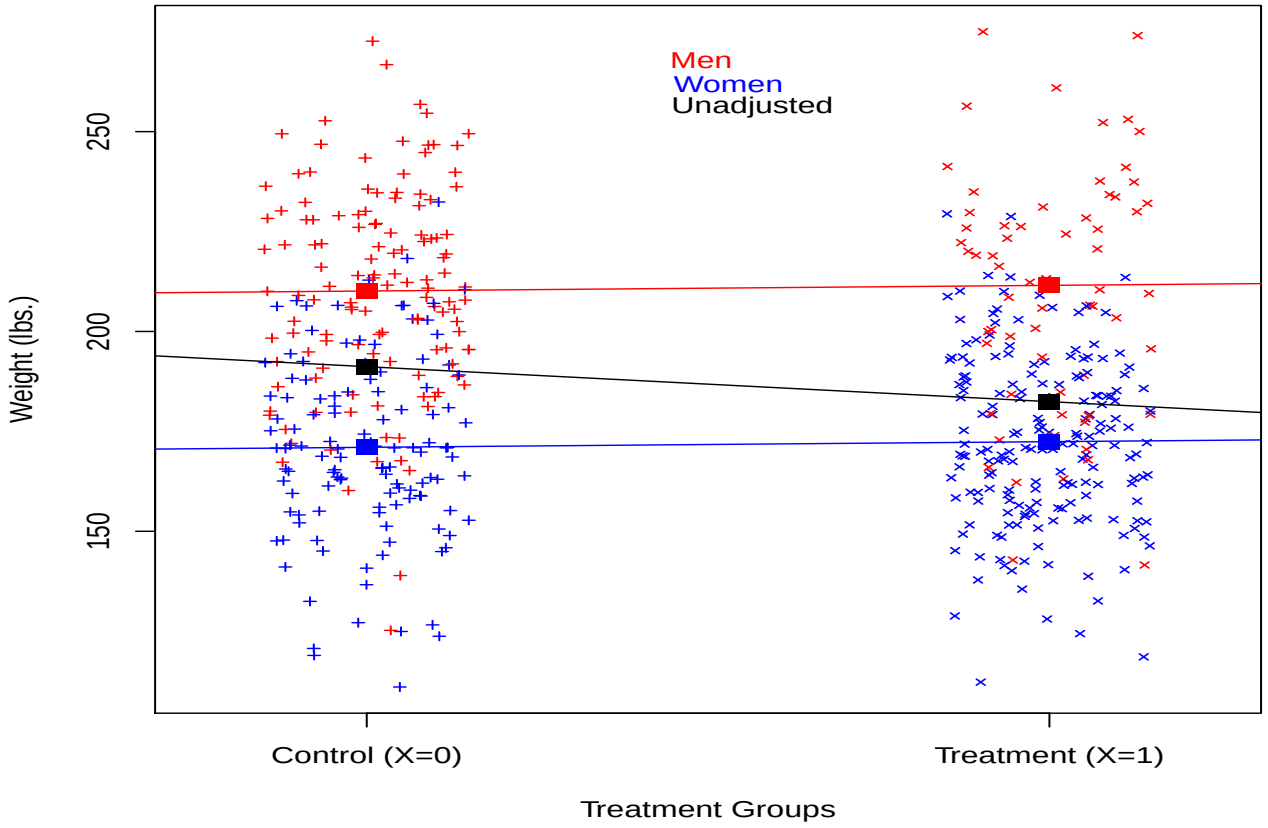
- Intercept
  - The estimated mean weight of adults males not on the experimental treatment is 210.1 pounds (95% CI: 206.2 to 214.1 pounds)
    - The intercept corresponds to the mean weight for the group defined where both predictors (i.e., TRT & FEM) are equal to zero
    - While this is a well-defined group, it is of little importance given we do not know the population our sample represents

Gender Adjusted Analysis: Comments

- Comparing unadjusted and gender adjusted analyses
  - Marked difference in effect of treatment suggests that there was noteworthy confounding by gender
    - Gender is a strong predictor of weight
      - US Women average weight approximately 30 lbs lower than men (CDC)
    - Gender is associated with treatment in the sample
      - Percent women on treatment: 74.4%
      - Percent women on not on treatment: 48.4%
  - This “imbalance” could skew results

Gender Adjusted Analysis: Comments

- Marked difference between unadjusted association between weight and treatment compared to adjusted association (stratified by gender)



Gender Adjusted Analysis: Comments

- Comparing unadjusted and gender adjusted analyses
  - Marked lower Root MSE (24.0 vs 30.2) would tend to increase precision of estimate of treatment effect
  - Association between treatment and gender would tend to lower precision
  - Net effect: More precision ($\widehat{\mathrm{SE}}(\hat{\beta}_{\mathrm{TRT}})$: 2.30 vs 2.70)

Age & Gender Adjusted Analysis: Stata Output

```
. regress wgt i.TRT i.FEM AGE, robust

Number of obs      =        500
Root MSE           =     22.416

        |          Robust
    wgt |  Coef.    St Err       t     P>|t|      [95% CI]
--------+-------------------------------------------------------
    TRT |
   Yes  |   0.78    2.15      0.36    0.716    -3.45     5.02
        |
    FEM |
   Yes  | -38.77    2.32    -16.69    0.000   -43.33   -34.21
        |
    AGE |   1.22    0.15      8.38    0.000     0.94     1.51
  _cons | 169.86    5.26     32.27    0.000   159.52   180.20
----------------------------------------------------------------
```

Age & Gender Adjusted Analysis: Interpretation

- Treatment effect
  - The estimated mean weight of adults on the experimental treatment is 0.8 pounds higher than adults not on the treatment of the same age and gender (95% CI: -3.5 to 5.0 pounds)
    - These results are typical of what we might expect with no true difference in mean weight between the treatment groups (P = 0.716)
- Age effect
  - The estimated mean weight of adults is 1.2 pounds higher for each year difference in age between two age groups with the same treatment status and gender (95% CI: 0.9 to 1.5 pounds higher for each year difference in age)
    - These results are highly atypical of what we might expect with no true difference in the mean weight between age groups having the same treatment status and gender ($P < 0.001$)

Age & Gender Adjusted Analysis: Interpretation

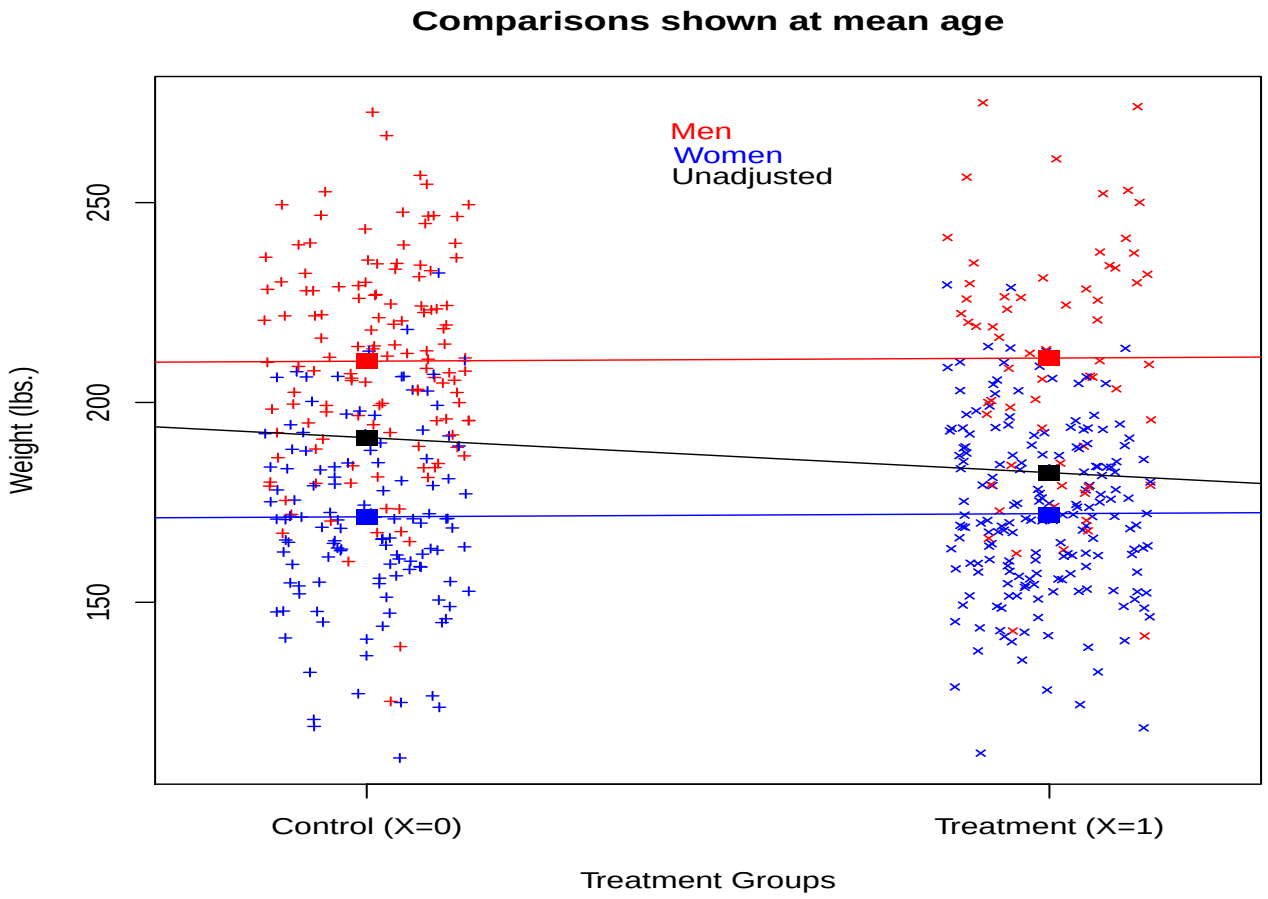
- Gender effect
  - The estimated mean weight of adult females is 38.8 pounds lower than adult males with the same treatment status and age (95% CI: -43.3 to -34.2 pounds)
    - These results are highly atypical of what we might expect with no true difference in the mean weight between women and men having the same treatment status and age ($P < 0.001$)
- Intercept
  - The estimated mean weight of newborn adults males not on the experimental treatment is 169.9 pounds (95% CI: 159.5.2 to 180.2 pounds)
    - The intercept corresponds to the mean weight for the group defined where all predictors are equal to zero
      - non experimental treatment (TRT=0)
      - newborns (AGE=0)
      - men (FEM=0)
    - There is no scientific relevance because there are no such people in our sample OR population

Age & Gender Adjusted Analysis: Comments

- Comparing the gender adjusted analysis to the age and gender adjusted analyses
  - No appreciable difference in the effect of treatment suggesting no further confounding of treatment after we had adjusted for gender
  - Lower Root MSE (22.4 vs 24.0) would tend to increase precision of the estimate of the treatment effect
  - Gender (or treatment) do not appear to confound the effect of age on weight
    - Age is acting as a precision variable
  - Effect of age and gender on precision
    - Association between treatment and gender would tend to lower precision
    - Lack of an association between treatment and age tends not to impact precision
    - Net effect: More precision ($\widehat{\mathrm{SE}}(\hat{\beta}_{\mathrm{TRT}})$: 2.15 vs 2.70)

Age &Gender Adjusted Analysis: Comments

- Marked difference between unadjusted association between weight and treatment compared to adjusted association (by gender and age)



Final Comments

Choosing the model for analysis

- Confirmatory vs Exploratory analyses
  - Every statistical model answers a different question
  - Data driven choice of analyses requires later confirmatory analyses
  - Best strategy
    - Choose appropriate primary analysis based on scientific question identified a priori
      - Provide most robust statistical inference regarding this question
  - Further explore your data to generate new hypotheses and speculate on mechanisms
    - Regard these statistics as descriptive

Final Disclaimer

- In presenting 4 different analyses for the weight loss data, I did not mean to suggest that I would choose from among these
  - Instead, I wanted to show how regression could be used to address confounding and provide greater precision
  - I would have chosen the analysis based on age and gender adjustment a priori, and reported those results as my primary analysis