Lecture 4a: Practice Problem Solutions

John McGready
Johns Hopkins University
1. Which of the following examples involve the comparison of paired data?
   - If so, on what are we pairing the data?
a) In Baltimore, a real estate practice known as “flipping” has elicited concern from local and Federal government officials

- “Flipping” occurs when a real estate investor buys a property for a low price, makes little or no improvement to the property, and then resells it quickly at a higher price
- This practice has raised concern because the properties involved in “flipping” are generally in disrepair and the victims generally have low incomes
- Fair housing advocates are launching a lawsuit against three real estate corporations accused of this practice
Practice: Paired Data?

a) In Baltimore, a real estate practice known as “flipping” has elicited concern from local and Federal government officials.
   - As part of the suit, these advocates have collected data on all houses purchased by these three corporations which were sold in less than one year after they were purchased.
   - Data were collected on the purchase price and the resale price for each of these properties.
   - The data were collected to investigate whether the resale prices were, on average, higher than initial purchase price.
   - A confidence interval was constructed for the average profit in these quick turnover sales.

   - This is a paired data scenario: each before price is connected to a specific after price, the unit of pairing is the house.
b) Researchers are testing a new blood pressure-reducing drug; participants in this study are randomized to either a drug group or a placebo group
- Baseline blood pressure measurements are taken on both groups and another measurement is taken three months after the administration of the drug/placebo
- Researchers are curious as to whether the drug is more effective in lowering blood pressure than the placebo

- This is an unpaired comparison ultimately. Paired difference are computed (after-before BP) for each subject in each group (drug or placebo) but the comparison of interest is the comparison of the differences between the two independent groups: drug and placebo.
Researchers are interested in the impact of a vegan diet on risk factors for coronary heart disease (CHD) in subjects with a family history of such CHD. Researchers randomly select 100 such families with more than one child and randomize two siblings from each family to either a vegan diet or an omnivorous diet (one sibling on each diet). These diets, prescribed by a nutritionist, are to last for six weeks.

- Baseline CHD risk factor measurements are taken (BP, cholesterol level, percent body fat) on each participant.
- Follow-up CHD risk factor measurements are taken at the end of the 6-week diet period.
- Changes in risk-fact levels are to be compared between those on the vegan diet and those on the omnivorous diet.
Researchers are interested in the impact of a vegan diet on risk factors for coronary heart disease (CHD) in subjects with a family history of such CHD. Researchers randomly select 100 such families with more than one child and randomize two siblings from each family to either a vegan diet or an omnivorous diet (one sibling on each diet). These diets, prescribed by a nutritionist, are to last for six weeks.

- This is a paired comparison (of paired differences)
- The change in cholesterol is computed for each subject in the study and these changes are compared between the two diet groups
- But each subject in the vegan group is matched with his/her sibling in the omnivorous group: so the diet group comparison is paired by sibling