Math 361

Randomization tests for Experimental Studies-
Matched Pairs - Inv. 4.8

Ch. 4 – Comparing two groups (binary EV) on single quantitative response (RV)

- What are appropriate **graphs** to look at?
- What are appropriate **statistics** for summarizing the data numerically?
- How to **test** H_0?
  - Randomization test (simulation)
  - t-test (approximate)
- How **estimate** a difference in population or treatment means?
- Scope of conclusions based on **study design**

Study designs

We'll see how to perform simulation-based tests and approximate tests for data for **three** types of study designs.

1. Observational study of samples from two populations   (Inv. 4.2 in lab on 11/15)

Experimental study with random assignment into

2. two independent groups (Inv. 4.4), or
3. matched pairs (Inv. 4.8)

Investigation 4.8 (p. 289)

Research Question:

Compare melting times of **dark** and **milk** chocolate chips.

Parameter of interest

The difference in mean melting times between dark and milk chocolate chips, $\mu_D - \mu_M$

Hypotheses:

$$H_0: \mu_D - \mu_M = 0 \text{ vs. } H_a: \mu_D - \mu_M \neq 0$$

Study design?

<u>Considerations:</u>

We have access to

- Students,
- Timers on their cellphones,
- A limited number of milk and peanut chips

We want to be able to conclude that any observed difference in ***melting times*** was caused by a difference in ***chip type*** (milk or peanut)…

…so we should use random assignment to eliminate the possibility of confounding variables.

An additional consideration: power

Power is the probability that our test will detect a difference in melting times of milk and peanut chips if a difference truly exists.

We'd like to design our experiment to have as high a power as possible.

Given a fixed sample size (number of students/chips available) and a fixed level of significance (0.05), we should seek out the most powerful **design** for our experiment.

Which design do you think is more **powerful?**

Each student tosses a coin to **randomly pick either a peanut or milk chip**.

Each student then

1. places their chip on their tongue,
2. starts a timer and holds the chip against the roof of their mouths until the chip is melted, and
3. the time until the chip is completely melted, without any “encouragement” by the student, is recorded in seconds.

Each student tosses a coin to **randomly pick to melt a peanut or milk chip first**.

Each student then

1. places the first chip on their tongue,
2. starts a timer and holds the chip against the roof of their mouths until the chip is melted, and
3. the time until the chip is completely melted, without any “encouragement” by the student, is recorded in seconds.

The steps 1-3 are repeated for the **second chip by each student**.

A Matched Pair Design

The melting times of peanut and milk chocolate chips are “paired” by student.

This design results in a more powerful experiment because the variation in melting times due to **differences in students** is eliminated.

Each student tosses a coin to **randomly pick to melt a peanut or milk chip first**.

Each student then

1. places the first chip on their tongue,
2. starts a timer and holds the chip against the roof of their mouths until the chip is melted, and
3. the time until the chip is completely melted, without any “encouragement” by the student, is recorded in seconds.

The steps 1-3 are repeated for the **second chip by each student**.

Collect Data using the “matched pair” design

Instructions:

1. Randomly decide which chip to melt first (toss a coin)
2. Melt each chip:
    - Put the chip on your tongue
    - Press to the roof of your mouth
    - Record time (in seconds)
3. Record both melting times by chip type on the sheet.

Answer as much of ***parts (h)-(j)*** *as you can while the data is being recorded.*

Results of Data Collection

What test of significance should we use?

Test: H_0: $\mu_D - \mu_M = 0$ vs. H_a: $\mu_D - \mu_M \neq 0$

A two-sample t-test or the Randomization Test from last time assume that the samples of milk and dark chocolate chips melting times are

independent

that is, the melting times of a milk chip are unrelated to the melting of any of the dark chips.

This assumption is not satisfied for our data because the melting times are related through students:

A student with a "cool" mouth will produce longer melting times for both milk and dark chips, while "hot" mouthed students will produce longer

Tests of significance for "Matched Pairs"

Test: H_0: $\mu_D - \mu_M = 0$ vs. H_a: $\mu_D - \mu_M \neq 0$

It is appropriate to use a one-sample t-test, where the "one" sample consists of the "differences in melting times", if the number of students is more than 30.

If less than 30 students were in the sample, then the matched pairs randomization test can be used. Here, the random choice of which chip to melt first is mimicked by randomly choosing the sign of the difference in melting times to be + or -.

Applet: Matched Pairs Randomization