Math 107

Inference for Two Means
(Sections 6.10 – 6.13)

Comparing means

Central Limit Theorem

- Many studies aim to compare two independent groups
- A natural way to compare groups is to compare sample means
- The sampling distribution of the difference in sample means is

$$\overline{\mathbf{x}}_{\mathbf{1}} - \overline{\mathbf{x}}_{\mathbf{2}} \sim \mathcal{N}\left(\mu_{\mathbf{1}} - \mu_{\mathbf{2}},\ \sqrt{\frac{\sigma_{\mathbf{1}}^{\mathbf{2}}}{\mathbf{n_2}} + \frac{\sigma_{\mathbf{2}}^{\mathbf{2}}}{\mathbf{n_2}}}\right)$$

- This is usually a good approximation if the sample sizes are both sufficiently large (≥30).

Standard Error

- Just as in the one-sample situation, if the population standard deviations are unknown, then we estimate them with the sample standard deviations

$$\mathbf{SE} = \sqrt{\frac{\mathbf{s_1^2}}{\mathbf{n_2}} + \frac{\mathbf{s_2^2}}{\mathbf{n_2}}}$$

t-distribution

- Due to the added uncertainty, the standardized test statistic now has a t-distribution rather than the normal distribution

$$\frac{(\overline{\mathbf{x}}_\mathbf{1} - \overline{\mathbf{x}}_\mathbf{2}) - (\mu_\mathbf{1} - \mu_\mathbf{2})}{\sqrt{\frac{\mathbf{s_1^2}}{\mathbf{n_2}} + \frac{\mathbf{s_2^2}}{\mathbf{n_2}}}} \sim \mathbf{t_{df}}$$

- df = smaller of $\mathbf{n_1 - 1}$ and $\mathbf{n_2 - 1}$

t-distribution

- In order to use the t-distribution,
  - we need random/representative samples within each group
  - the groups must be independent
  - the populations must be approximately normal, or we must have large sample sizes

Confidence intervals

Confidence Interval Formula

General formula:

$$\text{sample statistic} \pm \mathbf{t^*} \times \text{SE}$$

Formula for difference in means:

$$(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \pm \mathbf{t}^*_{\mathbf{df}} \sqrt{\frac{\mathbf{s}_1^2}{\mathbf{n}_2} + \frac{\mathbf{s}_2^2}{\mathbf{n}_2}}$$

Video Games and GPA

- 210 first-year college students were randomly assigned roommates
- For the 78 students assigned to roommates who brought a video game to college: average GPA after the first semester was 2.84, with a SD of 0.669.
- For the 132 students assigned to roommates who did not bring a video game to college, average GPA after the first semester was 3.105, with a SD of 0.625.
- Question: How much does getting assigned a roommate who brings a video game to college affect your first semester GPA? Use a 90% CI to answer this question.

Hypothesis tests

t-Test for a Difference in Means

1. State the hypotheses
2. Check the conditions necessary to use the normal distribution
3. Compute the test statistic
4. Compute the p-value
5. Make a decision and state its implications in the context of the problem

The Pygmalion Effect

- Pygmalion Effect: the greater the expectation placed upon people, the better they perform
- Teachers were told that certain children (chosen randomly) were expected to be "growth spurters," based on the Harvard Test of Inflected Acquisition (a test that didn't actually exist).
- The response variable is the change in IQ over the course of one year.

Source: Rosenthal, R. and Jacobsen, L. (1968). "Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development." Holt, Rinehart and Winston, Inc.

The Pygmalion Effect

	n	Mean	SD
Control Students	255	8.42	12.0
"Growth Spurters"	65	12.22	13.3

Does this provide evidence that the Pygmalion Effect exists? (that merely expecting a child to do better actually *causes* the child to do better?)

*s_1 and s_2 were not given, so I set them to give the correct p-value

The Pygmalion Effect

From the paper:

"The difference in gains could be ascribed to chance about 2 in 100 times"

INTELLECTUAL GROWTH

Expectancy Advantage by Grades

The bottom row of Table 1 gives the over-all results for Oak School. In the year of the experiment, the undesignated control-group children gained over eight IQ points while the experimental-group children, the special children, gained over twelve. The difference in gains could be ascribed to chance about 2 in 100 times ($F = 6.35$).

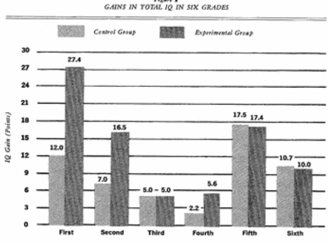
The rest of Table 1 and Figure 1 show the gains by children of the two groups separately for each grade. We find increasing expectancy advantage as we go from the sixth to the first grade; the correlation between grade level and magnitude of expectancy advantage ($r = -.86$) was significant at the .03 level. The interaction effect, or likelihood that at different grades there were significantly greater expectancy advantages, was significant at the .07 level ($F = 2.13$). (Interactions, however, are not sensitive to the ordering of differences unless one makes them so with further statistical efforts; that is, the p of .07 is conservative.)

In the first and second grades the effects of teachers' prophecies were dramatic. Table 1 shows that, and so does Table 2 and Figure 2. There we find the percentage of experimental- and control-group children of the first two grades who achieved various amounts of gain. In these grades about every fifth control-group child gained twenty IQ points or more, but of the special children, nearly every second child gained that much.

So far we have told only of the effects of favorable expectancies on total IQ, but Flanagan's TOGA yields separate IQs for the verbal and reasoning spheres of intellectual functioning. These are sufficiently different from each other so it will not be redundant to give the results of each. In the case of verbal IQ the control-group children of the entire school gained just less than eight points, and the special children gained just less than ten, a difference that could easily have arisen by chance. The interaction term was not very significant ($p < .15$) so that we can not conclude greater ex-

Figure 1

GAINS IN TOTAL IQ IN SIX GRADES



Paired data

Paired Data

- **Paired Data** consist of two measurements taken on related (non-independent) units
- Common paired data examples:
  - Two measurements on each case (compare each case to themselves under different treatments)
  - Twin studies
  - Each case is matched with a similar case, and one case in each pair is given each treatment

Analyzing Paired Data

- For a matched pairs situation, we look at the **difference** between responses **for each unit (pair)**, rather than just the average difference between treatment groups
- Create a new variable of the differences, and do inference for the difference as you would for a single mean

Inference for Paired Data