Lecture 29/Chapter 24

Significance vs. Importance
Undetected Differences

- Review Decision in z Test
- Factors Impacting Decision in z Test
- Importance of Sample Size
- Examples

Hypothesis Test for Means: Details

1. null hypothesis: pop mean = proposed value
   alt hyp: pop mean < or > or ≠ proposed value
2. Find sample mean (and sd) and standardize to z.
3. Find the P-value= prob of z this far from 0
4. If the P-value is small, conclude alt hyp is true.

Final conclusion hinges on size of P-value (is it small?) which hinges on size of z (is it large?).

Standardized Sample Mean

- To test a hypothesis about an unknown population mean, find sample mean (and standard deviation) and standardize to

$$z = \frac{\text{sample mean - pop mean}}{\frac{\text{standard deviation}}{\sqrt{\text{sample size}}}} = \frac{(\text{sample mean - pop mean})\sqrt{\text{sample size}}}{\text{standard deviation}}$$

- z is called the **test statistic**.

Note that "sample mean" is what we've observed, "population mean" is the value proposed in the null hypothesis, and "standard deviation" is from population (preferred) or sample (OK if sample size ≥ 30).

Standardized Sample Mean

To test a hypothesis about an unknown population mean, find sample mean (and standard deviation) and standardize to

$$z = \frac{\text{sample mean - pop mean}}{\frac{\text{standard deviation}}{\sqrt{\text{sample size}}}} = \frac{(\text{sample mean - pop mean})\sqrt{\text{sample size}}}{\text{standard deviation}}$$

What makes z large?

- Large difference between observed sample mean and hypothesized population mean
- Large sample size
- Small standard deviation (recall HW1 Ch. 1 #18(a))

Standardized Sample Mean

To test a hypothesis about an unknown population mean, find sample mean (and standard deviation) and standardize to

$$z = \frac{\text{sample mean - pop mean}}{\dfrac{\text{standard deviation}}{\sqrt{\text{sample size}}}} = \frac{(\text{sample mean - pop mean})\sqrt{\text{sample size}}}{\text{standard deviation}}$$

Conversely, when is z *not* large?

- Observed sample mean close to hypothesized population mean
- Small sample size
- Large standard deviation

Making Decision Based on *P*-value *(Review)*

If the *P*-value in our hypothesis test is small, our sample mean is improbably low/high/different, assuming the null hypothesis to be true. We conclude it is **not** true: we reject the null hypothesis and believe the alternative. A common cut-off for "small" is < 0.05.

If the *P*-value is not small, our sample mean is believable, assuming the null hypothesis to be true. We are willing to believe the null hypothesis.

P-value small ⟶ reject null hypothesis

P-value not small ⟶ don't reject null hypothesis

Example: *Hypothesis Test with Large Sample*

- **Background**: Number of credits taken by a sample of 400 students has mean 15.3, sd 2.
- **Question:** Can we believe population mean is 15?___
- **Response:**

1. Null: _______________ Alt: ________________
2. Sample mean ____, sd=2, z =__________________
3. *P*-value=prob of z this far from 0: ______________
4. Because the *P*-value is very small, we ____________ null hypothesis.

Example: *Hypothesis Test with Large Sample*

- **Background**: Number of credits taken by a sample of 400 students has mean 15.3, sd 2. A test to see if the population mean is 15 has very small *P*-value (0.0026).
- **Question:** Does this mean… (a) The true population mean is very different from 15? Or (b) We have very strong evidence that the true population mean is not 15?
- **Response:** ____ because other things factor into a small *P*-value besides how far what we observed is from what the null hypothesis claims. In fact, 15.3 seems quite close to 15. How close?

Example: *Confidence Interval after Test*

- **Background**: Number of credits taken by a sample of 400 students has mean 15.3, sd 2.
- **Question:** What is a 95% confidence interval for the population mean?
- **Response:** ________________________________
  so the population mean is apparently quite close to 15.

Example: *Asbestos & Lung Cancer?*

- **Background**: M. Kanarek found a "strong relationship" between the rate of lung cancer among white males and the concentration of asbestos fibers in the drinking water: *P*-value<0.001. An increase of 100 times the asbestos concentration went with an increase of 1.05 per 1000 in the lung cancer rate: 1 more case per year per 20,000 people…In tests of 200+ relationships, the *P*-value for lung cancer in white males was the smallest…They adjusted for age & other demographic variables, but not smoking.
- **Question:** Is there really a strong relationship between asbestos in drinking water and lung cancer in white males?
- **Response:**

Example: *Asbestos & Lung Cancer?*

- **Background**: M. Kanarek found a "strong relationship" between the rate of lung cancer among white males and the concentration of asbestos fibers in the drinking water: *P*-value<0.001. An <u>increase of 100 times the asbestos concentration went with an increase of 1.05 per 1000 in the lung cancer rate: 1 more case per year per 20,000 people…</u>In tests of 200+ relationships, the *P*-value for lung cancer in white males was the smallest…They adjusted for age & other demographic variables, but not smoking.
- **Question:** Is there really a strong relationship between asbestos in drinking water and lung cancer in white males?
- **Response:** (1) The evidence might be_______ (small *P*-value thanks to large samples) but the relationship is ______:1 more case per 20,000 people, when asbestos increases × 100, is minimal.

Example: *Asbestos & Lung Cancer?*

- **Background**: M. Kanarek found a "strong relationship" between the rate of lung cancer among white males and the concentration of asbestos fibers in the drinking water: *P*-value<0.001. An increase of 100 times the asbestos concentration went with an increase of 1.05 per 1000 in the lung cancer rate: 1 more case per year per 20,000 people…<u>In tests of 200+ relationships, the *P*-value for lung cancer in white males was the smallest</u>…They adjusted for age & other demographic variables, but not smoking.
- **Question:** Is there really a strong relationship between asbestos in drinking water and lung cancer in white males?
- **Response:** (2) Beware of ____________! If we reject the null for every P-value < 0.05, then____% of the time, in the long run, we make a Type I Error, rejecting the null even though it's true. For every 100 tests of a true null hyp, about___ incorrectly reject it.

Example: *Asbestos & Lung Cancer?*

- **Background**: M. Kanarek found a "strong relationship" between the rate of lung cancer among white males and the concentration of asbestos fibers in the drinking water: P-value<0.001. An increase of 100 times the asbestos concentration went with an increase of 1.05 per 1000 in the lung cancer rate: 1 more case per year per 20,000 people…In tests of 200+ relationships, the *P*-value for lung cancer in white males was the smallest…They adjusted for age & other demographic variables, but not smoking.
- **Question:** Is there really a strong relationship between asbestos in drinking water and lung cancer in white males?
- **Response:** (3) Principles learned in Part One shouldn't be forgotten: they failed to control for an obvious confounding variable, __________. Perhaps there were more smokers in areas that had higher asbestos concentrations.

Example: *Hypothesis Test with Small Sample*

- **Background**: A manufacturer bragged: "Tests comparing our product to the more expensive competitor's product showed no statistically significant difference in quality.
- **Question:** How impressed should we be?
- **Response:** ___________ If they only sampled a few products, a very small sample size would tend to produce a small *z*, which in turn yields a large *P*-value, failing to show a statistically significant difference.

Example: *Another Test with a Small Sample*

- **Background**: An experiment compared decrease in blood pressure over a 12-wk period for 10 men taking calcium vs. 11 taking placebo. The two-sample *t* was 1.634, with *P*-value=0.06. Using 0.05 as the cut-off, the test has failed to produce statistically significant evidence of the benefits of calcium for blood pressure.
- **Question:** Can we be sure calcium doesn't help b.p.?
- **Response:** __________; a *P*-value of 0.06 is still on the small side. Perhaps larger samples would yield significant results.

Example: *Small vs. Large Samples*

- **Background**: An experiment compared decrease in blood pressure over a 12-wk period for 10 men taking calcium vs. 11 taking placebo. The two-sample t was 1.634, with *P*-value=0.06. Using 0.05 as the cut-off, the test has failed to produce statistically significant evidence of the benefits of calcium for blood pressure.
- **Question:** Why didn't the study use more subjects?
- **Response:**