Sets 31: Hypothesis Testing with Two Paired Samples

Stat 260: July 31, 2024

In our final section, we'll look once more at the difference of two means. This time, however, our means no longer come from independent populations.

Example 1: A new medication is thought to have a side effect of increased heart rate. The resting heart rate in beats per minute (bpm) of 8 patients is measured before taking the medication, and then one hour after taking the medication.

Patient ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Before Medication (B)	73	94	87	67	101	77	75	86
After Medication (A)	82	94	94	78	117	96	78	84

*Since the two measurements are taken from the same patient, they are no longer independent; instead of treating the Before-data and the After-data as separate samples, we treat their Difference-data as a single sample, with a single **sample mean** $\overline{x}_D$, and a single **sample standard deviation** s_D.*

*Now that we have a single sample of **Difference-data**, we proceed exactly as with our previous work on inferences of a single mean: with hypothesis tests as in Sets 25-27, and with confidence intervals as in Sets 22-23.*

We typically denote each parameter and statistic with a subscript D to demonstrate that it comes from a Difference-data set: μ_D, σ_D, $\overline{x}_D$, s_D, n_D, etc.

- **Case 1: (Z) Large Samples ($n_D \geq 30$) or σ_D is known with a normal population**
  - (i) **Requirements:**
    - ▷ Random samples
    - ▷ Paired (not independent) data.
    - ▷ Either $n_D < 30$, or σ_D is known and population is normally distributed.
  - (ii) **Confidence Interval for μ_D:**

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma_D}{\sqrt{n_D}} \qquad \text{or} \qquad \bar{x} \pm z_{\alpha/2} \frac{s_D}{\sqrt{n_D}}$$

  - (iii) **Test Statistic for hypothesis tests:**

$$z_{obs} = \frac{\bar{x}_D - \mu_D}{\sigma_D/\sqrt{n_D}} \qquad \text{or} \qquad z_{obs} = \frac{\bar{x}_D - \mu_D}{s_D/\sqrt{n_D}}$$

$\ast$ *Use σ_D if you know it, use s_D if you don't.*

- **Case 2: (T) Small samples $(n_D < 30)$ and σ_D is unknown**

  (i) **Requirements:**

  - ▷ Random samples
  - ▷ Paired (not independent) data.
  - ▷ Near-normally distributed population.

  (ii) **Confidence Interval for μ_D:**

$$\bar{x} \pm t_{\nu,\alpha/2}\,\frac{s_D}{\sqrt{n_D}} \qquad \text{where } \nu = n_D - 1.$$

  (iii) **Test Statistic for hypothesis tests:**

$$t_{obs} = \frac{\bar{x}_D - \mu_D}{s_D/\sqrt{n_D}}$$

  Test statistic has a T_ν-distribution with $\nu = n_D - 1$ degrees of freedom when H_0 is true.

Example 1 Continued...

(a) Calculate a 90% confidence interval for the mean difference (after-before) of resting heart rates. Assume the paired differences are normally distributed.

(b) Is there evidence to suggest that the mean resting heart rates difference is more than 5 bpm after taking the medication?

(i) Define the parameter(s) to be tested.

(ii) Specify H_0 and H_1.

(iii) Specify the test-statistic to be used, and identify its distribution (assuming H_0 is true).

(iv) Compute the observed value of the test-statistics.

(v) Compute the p-value.

(vi) Report the strength of evidence against H_0.

(vii) Using significance level $\alpha = 0.10$, state your conclusion.

(viii) Report the estimated value of the parameter and (estimated) standard error (ese or se).

Extra Example 1: A wildlife rehabilitation centre hopes that by providing proper nutrition and safe habitat to kiwi birds (*Apteryx mantelli*) in their captive breeding program, that the average weight of a (fully grown) kiwi born into the program will be greater than that of their mother. They record the weights of 50 adult female kiwi birds, and compare them to the weights of their firstborn female offspring (once fully grown). The mean weight difference (Mother-Offspring) is -30g, with a standard deviation of 150g. Is there evidence to suggest that the mean difference weight of the kiwis (Mother-Offspring) is less than 0g?

Textbook Readings: Swartz 7.4 [EPS middle of p. 260 – end of 6.5]
Practice problems: EPS: 5.33, 5.35, 6.43, 6.45