Hypothesis Tests for Paired Samples

October 4, 2019

T-Test Example

The following data is on red blood cell counts (in 10^6 cells per microliter) for 9 people:

$$5.4, \quad 5.3, \quad 5.3, \quad 5.2, \quad 5.4, \quad 4.9, \quad 5.0, \quad 5.2, \quad 5.4$$

Test at the 5% level of significance if the average cell count is 5.

Aside: One-Sided Tests

$$
\begin{aligned}
H_0: &\quad \mu \leq \mu_0 \\
H_A: &\quad \mu > \mu_0
\end{aligned}
$$

or

$$
\begin{aligned}
H_0: &\quad \mu \geq \mu_0 \\
H_A: &\quad \mu < \mu_0
\end{aligned}
$$

One-Sided Hypothesis Testing

There is only one major difference in one-sided hypothesis testing.

- For the test statistic approach
  - We use the critical value z_α instead of $z_{\alpha/2}$.
- For the p-value approach
  - We no longer multiply by two: p-value $= P(Z < -|ts|)$

One-Sided Hypothesis Testing

In both cases,

- We are no longer interested in seeing observations as extreme as $\bar{x}$.
- Now we are actively interested in a particular direction corresponding to the direction of the alternative hypothesis.
- This means that we are interested in a particular Z-score or a single tail area.

One-Sided Hypothesis Testing

This is useful for several reasons

1. We don't have to find $z_{\alpha/2}$ or double the p-value, so the level of evidence required to reject H_0 goes down.
2. Sometimes we are really only interested in one direction.

On the flip side, we lose the ability to detect any interesting findings in the opposite direction.

Example

Suppose some doctors are interested in determining whether stents will help people who have a high risk of stroke.

- The researchers believe the stents would help.
- ...but the data suggests the opposite, that stents are actively harmful.

One-Sided Hypotheses

- A one-sided test could have checked whether the stents were helpful.
- But a two-sided test allowed the researchers to see that there was harm being done.
- Using one-sided hypotheses runs the risk of overlooking data supporting an opposite conclusion.

One-Sided Hypothesis Tests

So when should you use a one-sided test? Rarely!

Before using a one-sided test, consider:

- What would we conclude if the data happens to go clearly in the opposite direction?
- Is there any value in learning about the data doing in the opposite direction?

Why Can't We Look at the Data First?

We should always set up our hypotheses and analysis plan ***before taking any data***.

- This is part of doing good science!
- If we pick hypotheses after seeing the data we double our probability of Type I error.

Paired Data

What if we want to compare two populations?
Why might we want to do this?

Paired Data

A medical researcher is interested in testing a new blood pressure medication.

Patient	Before (Week 0)	After (Week 2)
1	141	125
2	135	118
⋮	⋮	⋮
128	138	121

Paired Observations

- Each patient has two corresponding observations.
- Each observation has an explicit connection to exactly one other observation.
- It is natural to pair these observations.

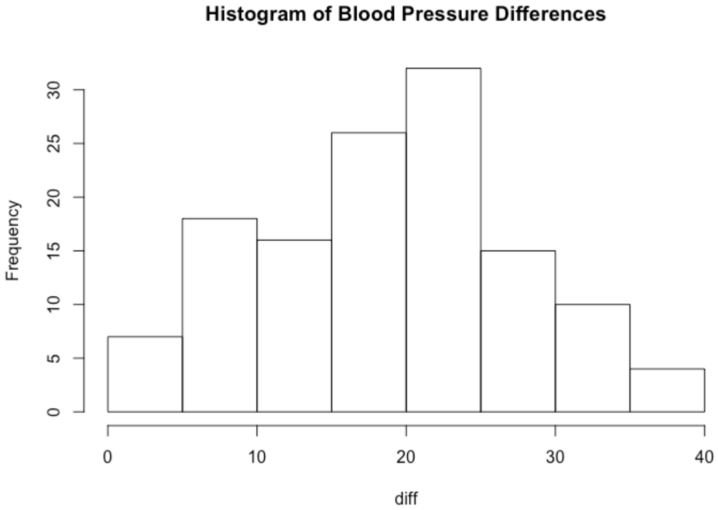
Paired Observations

We often analyze paired data by looking at the differences.

Patient	Before (W0)	After (W2)	Difference
1	141	125	16
2	135	118	17
⋮	⋮	⋮	⋮
128	138	117	21

Note: We want to be consistent with the subtraction order!
Here, we always take Week0 - Week2.

Paired Data


Histogram of Blood Pressure Differences
Frequency
0
5
10
15
20
25
30
0
10
20
30
40
diff

Inference for Paired Data

Consider some sample statistics for these differences:

n_{diff}	$\bar{x}_{\text{diff}}$	s_{diff}
128	18.83	8.45

Taking the differences is going to make our lives easy!

Example

Let's run the hypothesis test for our researcher looking at blood pressure medication. We'll test at the $\alpha = 0.01$ level.

n	$\bar{x}_{\text{diff}}$	s_{diff}
128	18.83	8.45

Example: Climate

We have a set of temperatures taken at 197 locations in 1948 and in 2018. We want to know if there were more days exceeding 90^o in 2018 or in 1948.

Is there a relationship between the observations in 1948 and 2018? Or are they independent?

Example: Climate

The difference in number of days exceeding 90^o was calculated for each location (days in 2018 - days in 1948). The sample statistics for these differences are:

n	$\bar{x}_{\text{diff}}$	s_{diff}
197	2.9	17.2

Test whether there were more days exceeding 90^o in 2018 or in 1948.