M6: Problem Set One Sample Z-test

Pablo E. Gutiérrez-Fonseca

2024-02-22 13:16:19

Welcome to our first foray into inferential statistics. The goal is to make sure that you understand the basics of what one-sample z-tests are doing and the steps in conducting any inferential test.

Be sure that you can navigate each of these approaches to running a One-Sample Z-test.

The Water Pollution Concentration data that will be used to answer Question #5 can be downloaded from this link.

Practicing one sample z-test when parameters are provided.

According to the European medical Association, 11 year old boys in this region are known to have a mean weight of $\mu = 85$ **pounds** with population standard deviation $\sigma = 11.6$ **pounds**. A complaint is made that the boys living in a municipal children's home are underfed. As one bit of evidence, **n = 25 boys** (all age 11) are weighed and found to have a **mean weight of 80.94 pounds**.
Assume the data is normally distributed and test to see **if the boys at the municipal home are significantly smaller than the general population of 11 year old boys in the region**.

1.1 What is the obtained value for this test?
1.2 What is the critical value of this test assuming an alpha threshold of 0.05?
2. Do you conclude that there is a significant difference between the weight of the boys in the municipal home and the general population? Yes or no?
3. What is the Actual p-value associated with this obtained test statistic value? (you can use excel functions for this).
4. Based on the available data, can you conclude that the boys in the municipal home are underweight compared to the general population? Write a brief 1 paragraph summary for this analysis.

Running a one-sample z-test in R using *z.test()*.

An environmental law firm is suing the European Union for additional resources to clean its waterways. They claim that the mean concentration (ppm) of contaminants in their water samples collected over a 10-year period are significantly higher than the larger "population" of water quality samples collected across the EU. Citing significantly higher need would warrant the allocation of additional resources to their region. Download the data Water Pollution Concentration (as csv file) and test to see if the concentration (mean ppm) of pollutants from the 10 observation sample in our data set are **significantly higher than** the larger EU "population" of water quality results (of $\mu = 1644$, $\sigma = 497$).

Write a brief 1 paragraph summary for this analysis. Keep in mind this should be a succinct and clear summary. No lab report style of telling me step by step what you did.

This should include: a brief statement of your original hypothesis. A statement of the results that includes the short hand stats summary with z test statistic and p-value. Make sure this summarizes how the groups are different (which is bigger) and then interpret what you think these results mean (i.e. is this meaningful, is your original hypothesis correct, why might this be?)

- Use the code **z.test()**

```
z.test(x = , #values for the sample
alternative = "two.sided" or "greater" or "less", #string value indicating type of null hypothesis
mu = ,  #population mean for comparison
sigma.x = ) #population standard deviation
```

```
## Error: <text>:2:27: unexpected symbol
## 1: z.test(x = , #values for the sample
## 2: alternative = "two.sided" or
##                           ^
```

5. Write a brief 1 paragraph summary for this analysis.