Problem Set: One Sample Z-test

Pablo E. Gutiérrez-Fonseca

2023-08-09

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. Some of these questions can be answered by hand, in excel or in JMP. Be sure that you can navigate each of these approaches to running a One-Sample Z-test.

Practicing one-sample z-tests the old fashioned way by hand.

According to the European medical Association, 11 year old boys in this region are known to have a mean weight of $\boldsymbol{\mu = 85}$ **pounds** with population standard deviation $\boldsymbol{\mu = 11.6}$ **pounds**. A complaint is made that the boys living in a municipal children's home are underfed. As one bit of evidence, $\mathbf{n = 25}$ **boys** (all age 11) are weighed and found to have a mean weight of 80.94 pounds.
Assume the data is normally distributed and do this analysis by hand (excel) 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. What is the obtained value for this test?
2. What is the critical value of this test assuming an alpha threshold of 0.05?
3. 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?
4. What is the Actual p-value associated with this obtained test statistic value? (you can use excel functions for this).
5. 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.

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.jmp 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 $\boldsymbol{\mu = 1644}$, $\boldsymbol{\mu = 497}$).

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

R practice.

```
df_water <- read.csv('C:/Users/Guti/OneDrive - University of Vermont/Curriculum/07_ Cursos/NR-2400/Prob
head(df_water)
```

```
##       region  value
## 1 North West 1782.4
## 2 North East 1904.2
## 3   Midlands 3070.1
## 4    Anglian 1773.2
## 5     Thames 1528.3
## 6   Southern 1317.0
```

```r
# Population mean (null hypothesis)
population_mean <- 1644
population_sd <- 497
```

```r
# Calculate sample mean
sample_mean <- mean(df_water$value, na.rm=TRUE)
```

```r
# Sample size
n <- length(df_water$value)
```

```r
# Calculate the test statistic (z-score)
z_score <- (sample_mean - population_mean) / (population_sd / sqrt(n))
```

```r
# Calculate p-value for two-tailed test
p_value_two_tail <- 2 * (1 - pnorm(abs(z_score)))
# Calculate the p-value for the right-tailed test
p_value_right_tailed <- 1 - pnorm(z_score)
# Calculate the p-value for the left-tailed test
p_value_left_tailed <- pnorm(z_score)
```

```r
# Print results
cat("Z-Score:", z_score, "\n")
```

```
## Z-Score: 1.577321
```

```r
cat("Prob >|z|:", p_value_two_tail, "\n")
```

```
## Prob >|z|: 0.1147216
```

```r
cat("Prob > z:", p_value_right_tailed, "\n")
```

```
## Prob > z: 0.05736082
```

```r
cat("Prob < z:", p_value_left_tailed, "\n")
```

```
## Prob < z: 0.9426392
```