

M6: Problem Set One Sample Z-test for TAs

Pablo E. Gutiérrez-Fonseca

2024-02-27 13:23:44

R practice.

```
library(BSDA)
```

```
## Warning: package 'BSDA' was built under R version 4.3.2
```

```
## Loading required package: lattice
```

```
##
```

```
## Attaching package: 'BSDA'
```

```
## The following object is masked from 'package:datasets':
```

```
##
```

```
##      Orange
```

Remember: We can also calculate the p-value using $2 * \text{pnorm}(\text{abs}(z), \text{lower.tail}=\text{FALSE})$ for a two-sided test and $\text{pnorm}(z)$ or $\text{pnorm}(z, \text{lower.tail}=\text{FALSE})$ for a one-sided test where z is the z-statistic.

```
# Given data
```

```
population_mean <- 85
```

```
population_sd <- 11.6
```

```
sample_mean <- 80.94
```

```
n <- 25
```

```
# Perform one-sample z-test
```

```
# Calculate the Z-test statistic
```

```
z <- (sample_mean - population_mean) / (population_sd / sqrt(n))
```

```
z
```

```
## [1] -1.75
```

```
# Two-tailed test, so multiply p-value by 2
```

```
p_value <- pnorm(z, lower.tail = T)
```

```
p_value
```

```
## [1] 0.04005916
```

Load the water pollution data into R.

```
##           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
## 7      South West 2278.5
## 8      EA Wales  1585.9
## 9      Scotland  2101.7
## 10 Northern Ireland 1577.7
```

```
z.test(x= df_water$value,
alternative = "greater",
mu = 1644,
sigma.x = 497)
```

```
##
## One-sample z-Test
##
## data: df_water$value
## z = 1.5773, p-value = 0.05736
## alternative hypothesis: true mean is greater than 1644
## 95 percent confidence interval:
##  1633.386      NA
## sample estimates:
## mean of x
##    1891.9
```