Test 4: Z-test for a Proportion (Binomial Distribution) Object

# Load Required Packages:  
library(package = dplyr)  
library(package = psych)  
  
# Load Required Data:  
data = read.csv(file = "data/Data\_Test\_04.csv", header = TRUE) %>%  
 mutate(Khorramabad\_Occr = ifelse(test = Khorramabad >= 0.1, yes = 1, no = 0),  
 Mashhad\_Occr = ifelse(test = Mashhad >= 0.1, yes = 1, no = 0))  
  
# Prepare Data:  
data\_1359\_2017 = data %>% filter(Year < 2018)  
data\_2018 = data %>% filter(Year >= 2018)  
  
# Show Data:  
headTail(x = data, top = 2, bottom = 2)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ## |  | Year | Month | Day | Khorramabad | Mashhad | Khorramabad\_Occr | Mashhad\_Occr |
| ## | 1 | 1959 | 1 | 1 | 0 | 0.3 | 0 | 1 |
| ## | 2 | 1959 | 1 | 2 | 2 | 1 | 1 | 1 |
| ## | ... | ... | ... | ... | ... | ... | ... | ... |
| ## | 21914 | 2018 | 12 | 30 | 0 | 0.3 | 0 | 1 |
| ## | 21915 | 2018 | 12 | 31 | 0 | 0 | 0 | 0 |

# Assumption Checking:  
# 1. The test is approximate and assumes that the number of observations in the sample is  
# sufficiently large (i.e. n>=30) to justify the normal approximation to the binomial.  
nrow(x = data\_2018)

## 365

# Khorramabad:  
# setting initial parameter values:  
x\_Khorramabad = sum(data\_2018$Khorramabad\_Occr, na.rm = TRUE)  
n\_Khorramabad = length(x = data\_2018$Khorramabad\_Occr)  
p0\_Khorramabad = mean(x = data\_1359\_2017$Khorramabad\_Occr, na.rm = TRUE)  
  
# use prop.test function - alternative = "two.sided":  
(result = prop.test(x = x\_Khorramabad, n = n\_Khorramabad, p = p0\_Khorramabad,  
 alternative = "two.sided", conf.level = 0.95, correct = FALSE))

## 1-sample proportions test without continuity correction:  
## data: x\_Khorramabad out of n\_Khorramabad, null probability p0\_Khorramabad  
## X-squared = 4.030, df = 1, p-value = 0.045  
## alternative hypothesis: true p is not equal to 0.189  
## 95 percent confidence interval:  
## 0.190 0.276  
## sample estimates:  
## p = 0.230

# Extract the Z Statistic:  
sqrt(x = result$statistic)

## 2.007

# Mashhad:  
# setting initial parameter values:  
x\_Mashhad = sum(data\_2018$Mashhad\_Occr, na.rm = TRUE)  
n\_Mashhad = length(x = data\_2018$Mashhad\_Occr)  
p0\_Mashhad = mean(x = data\_1359\_2017$Mashhad\_Occr, na.rm = TRUE)  
  
# use prop.test function - alternative = "two.sided":  
(result = prop.test(x = x\_Mashhad, n = n\_Mashhad, p = p0\_Mashhad,  
 alternative = "two.sided", conf.level = 0.95, correct = FALSE))

## 1-sample proportions test without continuity correction:  
## data: x\_Mashhad out of n\_Mashhad, null probability p0\_Mashhad  
## X-squared = 3.655, df = 1, p-value = 0.056  
## alternative hypothesis: true p is not equal to 0.169  
## 95 percent confidence interval:  
## 0.101 0.170  
## sample estimates:  
## p = 0.132

# Extract the Z Statistic:  
sqrt(x = result$statistic)

## 1.912