Test 5: Z-test for the Equality of Two Proportions (Binomial Distribution))

# Load Required Packages:  
library(package = dplyr)  
library(package = psych)  
  
# Load Required Data:  
data = read.csv(file = "data/Data\_Test\_05.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\_2018 = data %>% filter(Year >= 2018)  
  
# Show Data:  
headTail(x = data\_2018, top = 2, bottom = 2)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ## |  | Year | Month | Day | Khorramabad | Mashhad | Khorramabad\_Occr | Mashhad\_Occr |
| ## | 1 | 2018 | 1 | 1 | 0 | 0 | 0 | 0 |
| ## | 2 | 2018 | 1 | 2 | 4.9 | 0 | 1 | 0 |
| ## | … | … | … | … | … | … | … | … |
| ## | 364 | 2018 | 12 | 30 | 0 | 0.3 | 0 | 1 |
| ## | 365 | 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

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

## 2-sample test for equality of proportions without continuity correction:  
## data: c(x\_Khorramabad, x\_Mashhad) out of c(n\_Khorramabad, n\_Mashhad)  
## X-squared = 11.985, df = 1, p-value = 0.0005  
## alternative hypothesis: two.sided  
## 95 percent confidence interval:  
## 0.0433 0.154  
## sample estimates:  
## prop 1 prop 2   
## 0.230 0.132

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

## 3.462