Registration Number: 24MDT0184
Name: Tufan Kundu
Slot: L23+L24
Course Code: PMDS503P
Course Title: Statistical Inference Lab
DA 2

Q1: Medical researchers have developed a new artificial heart constructed primarily of titanium and plastic. The heart will last and operate almost indefinitely once it is implanted in the patient's body,but the battery pack needs to be recharged about every four hours. A random sample of 50 battery packs is selected and subjected to a life test. The average life of these batteries is 4.05 hours. Assume that battery life is normally distributed with a standard deviation of 0.2 hour.Is there evidence to support the claim that the mean battery life exceeds 4 hours? Use $\alpha = 0.01$ and $\alpha = 0.05$.

```
#Given information:
# H0: mu <=4
# HA: mu > 4
# n = 50
# sigma = 0.2
# xbar = 4.05


# Loading the BSDA package
library("BSDA")

## Loading required package:  lattice
##
## Attaching package:  'BSDA'
## The following object is masked from 'package:datasets':
##
##     Orange

# Given Data
x_bar <- 4.05   # Sample mean
mu_0  <- 4
sigma <- 0.2    # Population standard deviation
n     <- 50     # Sample size
alpha_1 <- 0.05 # Significance level 1
alpha_2 <- 0.01 # Significance level 2

# Performing one-sample Z-test (right-tailed test)
z_test <- zsum.test(mean.x = x_bar, sigma.x = sigma, n.x = n, mu =
                      mu_0, alternative = "greater")

# Extracting the results
p_value <- z_test$p.value
z_score <- z_test$statistic

# Printing the results
cat("Z-score:", z_score, "\n")

## Z-score: 1.767767

cat("P-value:", p_value, "\n")

## P-value: 0.03854994
```

```r
if (p_value < alpha_1) {
  cat("At alpha = 0.05, we reject H0.
There is enough evidence that the mean battery life exceeds 4 hours.\n")
} else {
  cat("At alpha = 0.05, we fail to reject H0.
There is not enough evidence to conclude that the mean battery life exceeds 4 hours.\n")
}

## At alpha = 0.05, we reject H0.
## There is enough evidence that the mean battery life exceeds 4 hours.

if (p_value < alpha_2) {
  cat("At alpha = 0.01, we reject H0.
There is evidence that the mean battery life exceeds 4 hours.\n")
} else {
  cat("At alpha = 0.01, we fail to reject H0.
There is not enough evidence to conclude that the mean battery life exceeds 4 hours.\n")
}

## At alpha = 0.01, we fail to reject H0.
## There is not enough evidence to conclude that the mean battery life exceeds 4 hours.
```

Conclusion:

- At $\alpha = 0.01$, we fail to reject the null hypothesis.
  There is not enough evidence to conclude that the mean battery life exceeds 4 hours.
- At $\alpha = 0.05$, we reject the null hypothesis.
  There is enough evidence to conclude that the mean battery life exceeds 4 hours.

Q2.Two different formulations of an oxygenated motor fuel are being tested to study their road octane numbers. The variance of road octane number for formulation-1 is 1.5, and for formulation-2 it is 1.2.Two random samples of size n1 = 35 and n2 = 40 are tested, and the mean road octane numbers observed are 89.6 and 92.5. Assume normality,
(a) Construct a 95% two-sided confidence interval on the difference in mean road octane number.
(b) If formulation 2 produces a higher road octane number than formulation 1, the manufacturer would like to detect it. Formulate and test an appropriate hypothesis, using $\alpha = 0.10$ and $\alpha = 0.05$.

```r
# loading the BSDA package
library(BSDA)

# Given data
mu1 <- 89.6         # Mean of Formulation 1
mu2 <- 92.5         # Mean of Formulation 2
sigma1 <- sqrt(1.5)  # Standard deviation of Formulation 1
sigma2 <- sqrt(1.2)  # Standard deviation of Formulation 2
n1 <- 35            # Sample size of Formulation 1
n2 <- 40            # Sample size of Formulation 2
alpha_1 <- 0.10     # Significance level 1
alpha_2 <- 0.05     # Significance level 2

# (a) 95% Confidence Interval for the Difference in Means
ci_result <- zsum.test(mean.x = mu1, sigma.x = sigma1, n.x = n1,

                       mean.y = mu2, sigma.y = sigma2, n.y = n2,
```

```r
                        alternative = "two.sided", conf.level = 0.95)

cat("95% Confidence Interval for difference in means (mu1 - mu2): (",
    ci_result$conf.int[1], ",", ci_result$conf.int[2], ")\n")

## 95% Confidence Interval for difference in means (mu1 - mu2): ( -3.429035 , -2.370965 )

# (b)
# Null hypothesis H0: mu1>=mu2
# Alternate hypothesis H1: mu1<mu2 (i.e. formulation 2 is better)
test_result <- zsum.test(mean.x = mu1, sigma.x = sigma1, n.x = n1,
                         mean.y = mu2, sigma.y = sigma2, n.y = n2,
                         alternative = "less")

# Extracting Z-score and p-value
z_score <- test_result$statistic
p_value <- test_result$p.value

# Print results
cat("Z-score:", z_score, "\n")

## Z-score: -10.7439

cat("P-value:", p_value, "\n")

## P-value: 3.165628e-27

# Decision at alpha = 0.10
if (p_value < alpha_1) {
  cat("At alpha = 0.10, reject H0.
Formulation 2 has a significantly higher road octane number.\n")
} else {
  cat("At alpha = 0.10, fail to reject H0.
Not enough evidence to conclude Formulation 2 is better.\n")
}

## At alpha = 0.10, reject H0.
## Formulation 2 has a significantly higher road octane number.

# Decision at alpha = 0.05
if (p_value < alpha_2) {
  cat("At alpha = 0.05, reject H0.
Formulation 2 has a significantly higher road octane number.\n")
} else {
  cat("At alpha = 0.05, fail to reject H0.
Not enough evidence to conclude Formulation 2 is better.\n")
}

## At alpha = 0.05, reject H0.
## Formulation 2 has a significantly higher road octane number.
```

Conclusion:

- The 95% Confidence Interval for the difference in means $(\mu_1 - \mu_2)$ is $\mathbf{(-3.429, -2.371)}$.
- At both significance levels $\alpha = 0.10$ and $\alpha = 0.05$, we reject the null hypothesis H_0.
- This provides strong evidence that Formulation 2 has a significantly higher road octane number than Formulation 1.

Q3. A random sample of 500 registered voters in Bangalore is asked if they favor the use of oxygenated fuels year-round to reduce air pollution. If more than 400 voters respond positively, we will conclude that at least 60% of the voters favor the use of these fuels.Find the probability of type I error if exactly 60% of the voters favor the use of these fuels.

```
# null hypothesis H0: p = 0.6
# alternate hypothesis: p > 0.6

# Given data
p0 <- 0.60  # Null hypothesis proportion
n <- 500    # Sample size
p_bar <- 400 / 500  # Observed sample proportion

# Standard Error (SE)
SE <- sqrt((p0 * (1 - p0)) / n)

# Compute Z-score
Z_score <- (p_bar - p0) / SE

# type 1 error
alpha <- 1 - pnorm(Z_score)

# Print results
cat("Z-score:", Z_score, "\n")

## Z-score: 9.128709

cat("Probability of Type I error (alpha):", alpha, "\n")

## Probability of Type I error (alpha): 0
```

Conclusion:
Probability of type 1 error is zero if exactly 60% of the voters favor the use of these fuels.

Q4. A random sample of 500 adult residents of Maricopa County ound that 385 were in favor of increasing the highway speed limit to 75 mph, while another sample of 400 adult residents of Pima Country found that 267 were in favor of the increased speed limit. Do these data indicate that there is a difference in the support for increasing the speed limit between the residents of the two countries? Use alpha =0.05. What is the P-value for this test?

```
# Null hypothesis H0: p1=p2
# Alternative Hypothesis H1: p1!=p2

# Given data
x <- c(385, 267)   # Favoring in Maricopa and Pima
n <- c(500, 400)   # Sample sizes

# Perform a two-proportion test (Chi-squared approximation)
test_result <- prop.test(x, n, alternative = "two.sided", correct = FALSE)

# Print results
cat("Chi-squared Test Statistic:", test_result$statistic, "\n")

## Chi-squared Test Statistic: 11.69556

cat("P-value:", test_result$p.value, "\n")

## P-value: 0.0006264947
```

```r
# Decision at alpha = 0.05
if (test_result$p.value < 0.05) {
  cat("At alpha = 0.05, we reject H0. There is a significant
difference in support for the speed limit increase.\n")
} else {
  cat("At alpha = 0.05, we fail to reject H0. No significant
difference in support between the counties.\n")
}

## At alpha = 0.05, we reject H0. There is a significant
## difference in support for the speed limit increase.
```

Conclusion:

- The computed p-value is **0.0006264947**.
- Since the p-value is less than the significance level $\alpha = 0.05$, we **reject the null hypothesis** H_0.
- This indicates that there is a **significant difference** in support for the speed limit increase between the two counties.