Homework7

2024-04-18

Question 1

Step 1:
Model for data: X ~ Normal(μ = 16.00, σ)

Step 2:
Null Hypothesis:
H0 = Coca Cola cans have 16.00 ounces of Cola.

Alternative Hypothesis:
Ha = Coca Cola cans have < 16.00 ounces of Cola.

Step 3:
We will pick the t statistic as our test statistic.

Step 4:

```
sample <- c(15.997, 16.005, 15.981, 15.954, 15.986, 16.021, 15.985, 16.001, 16.018, 16.056)

s_mean <- mean(sample)
s_sd <- sd(sample)

s_mean
```

```
## [1] 16.0004
```

```
s_sd
```

```
## [1] 0.02755278
```

Sample mean, $\bar{x}$ = 16.0004, sample standard deviation s = 0.02755278, Hypothesized mean μ = 16.00, Sample Size = 10.

```
Observed_tvalue <- (16.0004 - 16.00)/(0.02755278/sqrt(10))
Observed_tvalue
```

```
## [1] 0.04590865
```

t = 0.04590865

Step 5:
Computing p-value under H0,
Degrees of freedom = 9

```r
df <- 9
p_val <- pt(Observed_tvalue, df)
p_val
```

```
## [1] 0.5178072
```

p-value under H0 = 0.5178072

Step 6:
Significance level $\alpha = 0.05$.

Step 7:
p_value = 0.5178072 $> \alpha$

Therefore, we cannot reject the hypothesis that the mean fill of Coca Cola cans is 16.00 oz. The sample data does not provide enough evidence to prove that it is filled with less than 16.00 oz of Cola.

Question 2

Step 1:
Model for data: X ~ Binomial(n, p = 0.7)

Step 2:
Null Hypothesis:
H0 = 70% of college students in the US are stressed.

Alternative Hypothesis:
Ha = The percentage of college students in the US that are stressed is less than 70%.

Step 3:
We will pick the z statistic as our test statistic.

Step 4:
Sample proportion $\hat{p}$ = 130/200 = 0.65.
Hypothesized proportion p0 = 0.7.
Sample size n = 200

Observed z value = ($\hat{p}$ - p0)/$\sqrt{p0(1-p0)}/n$

```r
z <- (0.65 - 0.7)/sqrt((0.7 * 0.3)/200)
z
```

```
## [1] -1.543033
```

Step 5:
Computing p-value under H0,

```r
p <- 1 - pnorm(-1.543033, mean = 0, sd = 1)
p
```

```
## [1] 0.9385886
```

p-value under H0 = 0.9385886

Step 6:
Significance level $\alpha = 0.05$.

Step 7:
p_value = 0.9385886 $> \alpha$

Therefore, we cannot reject the hypothesis that 70% of college students in the US are stressed. The sample data does not provide enough evidence to prove that the number of stressed college students is different than 70%.

Question 3

Step 1:
Model for data: X ~ Normal($\mu = 2$, $\sigma = 0.812$) for ages 18 - 50.
Y ~ Normal($\mu = 1.85$, $\sigma = 0.837$) for ages 50 and above.

Step 2:
Null Hypothesis:
H0 = The mean cell phone data usage among people above 50 years of age is greater than or equal to that in people of ages 18 - 50.

Alternative Hypothesis:
Ha = The mean cell phone data usage among people above 50 years of age is less than that in people of ages 18 - 50.

Step 3:
We will pick the t statistic as our test statistic.

Step 4:
Sample size of customers from 18-50 (n1) = 350
Sample size of customers > 50 (n2) = '150
Sample mean of n1 ($\hat{X}$) = 2
Sample mean of n2 ($\hat{Y}$) = 1.85

Let $\hat{w}$ = (n1($\hat{X}$) + n2($\hat{y}$)) / (n1 + n2)

```
w <- ((350*2) + (150*1.85))/700
w
```

```
## [1] 1.396429
```

t-value = ($\hat{X}$ - $\hat{Y}$)/ $\sqrt{\hat{w}(1 - \frac{n1}{n1+n2})(\sigma1^2/n1 + \sigma2^2/n2)}$

```
t <- (2 - 1.85) / sqrt(w* (1 - (350/(350+150))) * (0.812/350) + (0.837/150))
t
```

```
## [1] 1.853135
```

Step 5:
Computing p-value under H0,

```
p_val <- pt(t, df = 350 + 150 - 2)
p_val
```

```
## [1] 0.9677729
```

p-value under H0 = 0.9677729

Step 6:
Significance level $\alpha = 0.05$.

Step 7:
p_value = 0.9677729 > α

Therefore, we cannot reject the hypothesis that the mean cell phone usage of those above the age of 50 is greater than or equal to those between ages 18 and 50. The statistical data does not provide enough evidence to prove that the phone usage of those above 50 is less than those from 18 - 50.

Question 4

Step 1:
Model for data: X ~ Binomial(n = 150, p = 14/150) for Apple iPhone.
Y ~ Binomial(n = 125, p = 15/150) for Samsung Galaxy.

Step 2:
Null Hypothesis:
The proportion of returned iPhones is $\geq$ the proportion of returned Samsung Galaxy phones.

Alternative Hypothesis:
Ha = The proportion of returned iPhones is < the proportion of returned Samsung Galaxy phones.

Step 3:
We will pick the z statistic as our test statistic.

Step 4:
Sample size of Apple iPhones (n1) = 150
Sample size of Samsung Galaxy (n2) = 125
Sample Proportion of returned iPhones (p1) = 14/150 = 0.093
Sample Proportion of returned Samsung Galaxy (p2) = 15/125 = 0.12

Let $\hat{w}$ = (p1 + p2) / (n1 + n2)

```
w <- (0.093 + 0.12)/275
w
```

```
## [1] 0.0007745455
```

Observed z value = (p1 - p2)/$\sqrt{\hat{w}(1-\hat{w})}/(1/n1 + 1/n2)$

```r
z <- -0.027/sqrt(0.000773945579268 * 0.0146666666667)
z
```

```
## [1] -8.01388
```

Step 5:
Computing p-value under H0,

```r
p <- pnorm(-8.01388, mean = 0, sd = 1)
p
```

```
## [1] 5.557258e-16
```

p-value under H0 = 5.557258e-16

Step 6:
Significance level $\alpha = 0.05$.

Step 7:
p_value = 5.557258e-16 $< \alpha$

Thus, we can reject the hypothesis that the proportion of returned iPhones is greater than or equal to the proportion of returned Samsung Galaxy phones. The sample data provides enough statistical evidence to prove that Apple iPhones have a smaller chance of being returned compared to Samsung Galaxy phones.

Question 5

```r
library(UsingR)
```

```
## Loading required package: MASS

## Loading required package: HistData

## Loading required package: Hmisc

##
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```r
data(babies)
m_age <- babies$age
d_age <- babies$dage
```

Step 1:
Model for data: X ~ Normal(μ, σ) for Mother's age.
Y ~ Normal(μ, σ) for Father's age.

Step 2:
Null Hypothesis:
H0 = The mean ages of mothers and fathers are equal.

Alternative Hypothesis:
Ha = The mean ages of mothers and fathers are different.

Step 3:
We will pick the t statistic as our test statistic.

Step 4:

```
X = mean(m_age)
Y = mean(d_age)
s1_square = var(m_age)
s2_square = var(d_age)
n1 = length(m_age)
n2 = length(d_age)

t <- (X - Y) / sqrt((s1_square / n1) + (s2_square / n2))
t
```

```
## [1] -11.0671
```

t = -11.0671

Step 5:
Computing p-value under H0,

```
p <- pnorm(-11.0671, mean = 0, sd = 1)
p
```

```
## [1] 9.05844e-29
```

p-value under H0 = 9.05844e-29

Step 6:
Significance level $\alpha = 0.05$.

Step 7:
p_value = 9.05844e-29 $< \alpha$

Thus, we can reject the hypothesis that the mean ages of mothers and fathers are equal. The sample data provides enough statistical evidence to prove that the mean ages of mothers and fathers are different.