

Milestone2

Unnati

2024-06-26

Question 1: Difference in mean MoCA scores between patients with and without Parkinson's disease

Hypotheses:

H0: $\mu_1 = \mu_2$ (Mean MoCA scores are the same for both groups)

#H1: $\mu_1 \neq \mu_2$ (Mean MoCA scores are different for both groups) # Two-sample t-test

```
park_data <- filter(data, Diagnosis == 1) %>% select(MoCA)
no_park_data <- filter(data, Diagnosis == 0) %>% select(MoCA)
t.test(park_data$MoCA, no_park_data$MoCA, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: park_data$MoCA and no_park_data$MoCA
## t = -8.06, df = 2103, p-value = 1.267e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.830418 -2.331212
## sample estimates:
## mean of x mean of y
## 13.92199 17.00281
```

#Interpretation #t-value: -8.06 #Degrees of freedom (df): 2103 #p-value: 1.267e-15 (which is extremely small) #95% confidence interval for the difference in means: -3.83 to -2.33 #Mean MoCA scores: #Patients with Parkinson's disease (Diagnosis == 1): 13.92 #Patients without Parkinson's disease (Diagnosis == 0): 17.00 #Since the p-value is much less than 0.05, we reject the null hypothesis (H0). This indicates a significant difference in the mean MoCA scores between patients with and without Parkinson's disease. Patients without Parkinson's disease have higher mean MoCA scores.

Question 2: Is the average UPDRS score significantly greater than 50 for patients diagnosed with Parkinson's disease?

Hypotheses:

H0: $\mu = 50$ (Mean UPDRS score is equal to 50)

H1: $\mu > 50$ (Mean UPDRS score is greater than 50)

```
updrs_data <- data %>% filter(Diagnosis == 1) %>% select(UPDRS)
t.test(updrs_data$UPDRS, mu = 50, alternative = "greater")
```

```
##
## One Sample t-test
##
## data: updrs_data$UPDRS
## t = 51.559, df = 1303, p-value < 2.2e-16
## alternative hypothesis: true mean is greater than 50
## 95 percent confidence interval:
## 116.8592 Inf
## sample estimates:
## mean of x
## 119.0641
```

#Interpretation # t-value: 51.559 #Degrees of freedom (df): 1303 #p-value: < 2.2e-16 (which is extremely small) #95% confidence interval for the mean: from 116.86 to infinity #Mean UPDRS score: 119.06 #Since the p-value is much less than 0.05, we reject the null hypothesis (H0). This indicates that the mean UPDRS score for patients with Parkinson's disease is significantly greater than 50.

Question 3: Is there a difference in mean BMI between patients with and without a history of traumatic brain injury?

Hypotheses:

H0: $\mu_1 = \mu_2$ (Mean BMI is the same for both groups)

H1: $\mu_1 \neq \mu_2$ (Mean BMI is different for both groups)

```
tbi_data <- data %>% filter(TraumaticBrainInjury == 1) %>% select(BMI)
no_tbi_data <- data %>% filter(TraumaticBrainInjury == 0) %>% select(BMI)

t.test(tbi_data$BMI, no_tbi_data$BMI, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: tbi_data$BMI and no_tbi_data$BMI
## t = -1.5467, df = 2103, p-value = 0.1221
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.7865639 0.2110559
```

```

## sample estimates:
## mean of x mean of y
## 26.50557 27.29332

#Interpretation: #t-value: -1.5467 #Degrees of freedom (df): 2103 #p-value: 0.1221 #95% confidence
interval for the difference in means: -1.79 to 0.21 #Mean BMI values: # Patients with a history of traumatic
brain injury: 26.51 #Patients without a history of traumatic brain injury: 27.29 #Since the p-value is greater
than 0.05, we fail to reject the null hypothesis (H0). This indicates that there is no significant difference in
the mean BMI between patients with and without a history of traumatic brain injury.

# Output results
print("MoCA scores comparison:")

## [1] "MoCA scores comparison:"
print(t.test(park_data$MoCA, no_park_data$MoCA, var.equal = TRUE))

##
## Two Sample t-test
##
## data: park_data$MoCA and no_park_data$MoCA
## t = -8.06, df = 2103, p-value = 1.267e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.830418 -2.331212
## sample estimates:
## mean of x mean of y
## 13.92199 17.00281

print("UPDRS scores comparison:")

## [1] "UPDRS scores comparison:"
print(t.test(updrs_data$UPDRS, mu = 50, alternative = "greater"))

##
## One Sample t-test
##
## data: updrs_data$UPDRS
## t = 51.559, df = 1303, p-value < 2.2e-16
## alternative hypothesis: true mean is greater than 50
## 95 percent confidence interval:
## 116.8592 Inf
## sample estimates:
## mean of x
## 119.0641

print("BMI scores comparison:")

## [1] "BMI scores comparison:"
print(t.test(tbi_data$BMI, no_tbi_data$BMI, var.equal = TRUE))

##
## Two Sample t-test
##
## data: tbi_data$BMI and no_tbi_data$BMI
## t = -1.5467, df = 2103, p-value = 0.1221
## alternative hypothesis: true difference in means is not equal to 0

```

```
## 95 percent confidence interval:  
## -1.7865639 0.2110559  
## sample estimates:  
## mean of x mean of y  
## 26.50557 27.29332
```