quiz w4

2024-01-30

```r
# Step 1: Simulate the Data
set.seed(0)
num_friends <- 20
true_heights <- runif(num_friends, 150, 200) # True heights in cm

# Measurement errors
error_edward <- rnorm(num_friends, 1, 2)
error_hugo <- rnorm(num_friends, 0, 3)
error_lucy <- rnorm(num_friends, -1, 2)

heights_edward <- true_heights + error_edward
heights_hugo <- true_heights + error_hugo
heights_lucy <- true_heights + error_lucy

# Combine data into a dataframe
data <- data.frame(
  friend = rep(1:num_friends, 3),
  height = c(heights_edward, heights_hugo, heights_lucy),
  measurer = rep(c("Edward", "Hugo", "Lucy"), each=num_friends)
)

# Step 2: Conduct Three Tests

# Test 1: ANOVA
anova_result <- aov(height ~ measurer, data = data)
summary(anova_result)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## measurer     2     44   21.86   0.101  0.904
## Residuals   57  12385  217.28
```

```r
# Test 2: Paired t-test between Edward and Hugo
t_test_result <- t.test(heights_edward, heights_hugo, paired = TRUE)
t_test_result
```

```
##
##  Paired t-test
##
## data:  heights_edward and heights_hugo
## t = 0.80963, df = 19, p-value = 0.4282
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -0.9926709  2.2451202
## sample estimates:
## mean difference
##       0.6262247
```

```r
# Test 3: Regression analysis between Hugo and Lucy
reg_result <- lm(heights_lucy ~ heights_hugo)
summary(reg_result)
```

```
##
## Call:
## lm(formula = heights_lucy ~ heights_hugo)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.5076 -1.4344  0.0232  1.9613  6.2947
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.55988    8.80944   0.404    0.691
## heights_hugo  0.97208    0.04929  19.721 1.23e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.188 on 18 degrees of freedom
## Multiple R-squared:  0.9558, Adjusted R-squared:  0.9533
## F-statistic: 388.9 on 1 and 18 DF,  p-value: 1.227e-13
```

```r
# Load the ggplot2 library

install.packages("ggplot2")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```r
library("ggplot2")

# Simulate the data
set.seed(0)
num_friends <- 20
true_heights <- runif(num_friends, 150, 200) # True heights in cm

# Measurement errors
error_edward <- rnorm(num_friends, 1, 2)
error_hugo <- rnorm(num_friends, 0, 3)
error_lucy <- rnorm(num_friends, -1, 2)

heights_edward <- true_heights + error_edward
heights_hugo <- true_heights + error_hugo
heights_lucy <- true_heights + error_lucy

# Combine data into a dataframe
data <- data.frame(
  friend = rep(1:num_friends, 3),
  height = c(heights_edward, heights_hugo, heights_lucy),
  measurer = rep(c("Edward", "Hugo", "Lucy"), each=num_friends)
)

# Create the plot using ggplot2
ggplot(data, aes(x = friend, y = height, color = measurer)) +
```

```
geom_point() +
theme_minimal() +
labs(title = "Comparison of Height Measurements by Edward, Hugo, and Lucy",
     x = "Individual Friends",
     y = "Measured Height (cm)",
     color = "Measurer")
```