Anorexia Data Lab

Rory Quinlan

2024-08-31

```
# Load data
Anorexia <- read.table("C:\\Users\\roryq\\Downloads\\Anorexia.txt", header=TRUE)
```

Context

The primary diagnosis for anorexia is a weight that is less than 85% of what is considered normal for that person's height and age. In this population of patients, researchers consider a patient anorexic if his or her weight is lower than 90 lbs. The columns Before and After are the weights of the patients before and after the treatments, respectively. The last column Y is the difference of weights: After - Before.

Check if there evidence that the patients are anorexic after treatment? Let α=0.05

- reject null at 95% confidence level and conclude that there is evidence anorexia is present after treatment

```
# Remove control to test treatment
w <- Anorexia$After[Anorexia$Therapy != "Control"]

# T test  to see of weight increases to above 90 (diagnostic for anorexia) with one tail, becaus
e we are checking if it is less

t.test(w,mu=90,alternative= "less")
```

```
##
##  One Sample t-test
##
## data:  w
## t = -1.8692, df = 44, p-value = 0.03413
## alternative hypothesis: true mean is less than 90
## 95 percent confidence interval:
##      -Inf 89.75508
## sample estimates:
## mean of x
##  87.57778
```

Was the average weight gain in the Cog/Behav group large enough to provide evidence that the Cog/Behav treatment worked even if anorexia is still present? Let α=0.05.

- Sufficient evidence to reject null at 95% confidence level and conclude that the gains from Cog/Behav therapy significantly differ from zero (greater than zero)

```
# filter out only cog/behav treatment and do one sided t test
Gains.cog <- Anorexia$Y[ Anorexia$Therapy == "Cog/Behav" ]
t.test(Gains.cog, mu=0, alternative= "greater")
```

```
##
##  One Sample t-test
##
## data:  Gains.cog
## t = 2.1535, df = 27, p-value = 0.02018
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  0.6331638      Inf
## sample estimates:
## mean of x
##  3.028571
```

Test to see if the cog/behavior treatment had greater weight gain than control treatment

- Fail to reject, not sufficient evidence to say that the true mean of cog/behav is significantly greater than control

```
# Filter data to seperate treatment types
Gains.cog_b <- Anorexia$Y[ Anorexia$Therapy == "Cog/Behav" ]
Gains.Cont <- Anorexia$Y[ Anorexia$Therapy == "Control" ]
Gains.Fam <- Anorexia$Y[ Anorexia$Therapy == "Family" ]


# two sample t and one sided t test
t.test(Gains.cog_b, Gains.Cont, alternative="greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  Gains.cog_b and Gains.Cont
## t = 1.6523, df = 50.913, p-value = 0.05232
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.04855955      Inf
## sample estimates:
## mean of x mean of y
##  3.028571 -0.450000
```

Test same hypothesis above with a permutation test and α= 0.01

- Fail to reject null at 99% confidence level, and conclude that there is insufficient evidence to say that the true mean weight gain is greater on cog/behav than in the control

```
# Create funtion to simulate distribution of differences between means
perm_fun <- function(x, nA, nB)
{
  n <- nA + nB
  idx_b <- sample(1:n, nB)
  idx_a <- setdiff(1:n, idx_b)
  mean_diff <- mean(x[idx_b]) - mean(x[idx_a])
  return(mean_diff)
}
```

```
set.seed(99)

# Create data for test
Y1 = Anorexia$Y[ Anorexia$Therapy != "Family"]
Gains.cog_bb <- Anorexia$Y[ Anorexia$Therapy == "Cog/Behav" ]
Gains.Cont <- Anorexia$Y[ Anorexia$Therapy == "Control" ]
nA = length(Gains.cog_bb)
nB = length(Gains.Cont)

# Call the perm_fun 1000 times to get 1000 mean differences
perm_diffs <- rep(0, 1000)
for (i in 1:1000) {
  perm_diffs[i] = perm_fun(Y1, nA, nB)
}

# Calculate the "p-value"
mean_diff = mean(Gains.cog_bb)-mean(Gains.Cont)
mean(perm_diffs > mean_diff)
```

```
## [1] 0.05
```

Check assumption of equal variances

```
var.test(Gains.Fam, Gains.Cont,
         alternative = "two.sided",
         conf.level = 0.95)
```

```
##
##  F test to compare two variances
##
## data:  Gains.Fam and Gains.Cont
## F = 0.80271, num df = 16, denom df = 25, p-value = 0.6587
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.3367083 2.0981634
## sample estimates:
## ratio of variances
##           0.8027132
```

- Fail to reject null and conclude we dont have suffiecient evidence to say that family and control group have different variance
- Assumption validated

Asses normality in the family treatment group and control

- They are both approximately normal, you can see when you compare observations to the line and there are no significant deviation from the linear pattern

```
qqnorm(Gains.Cont, pch = 19, frame = FALSE, main="Normal Q-Q Plot (Control")
qqline(Gains.Cont, col = "steelblue", lwd = 2)
```

Normal Q-Q Plot (Control
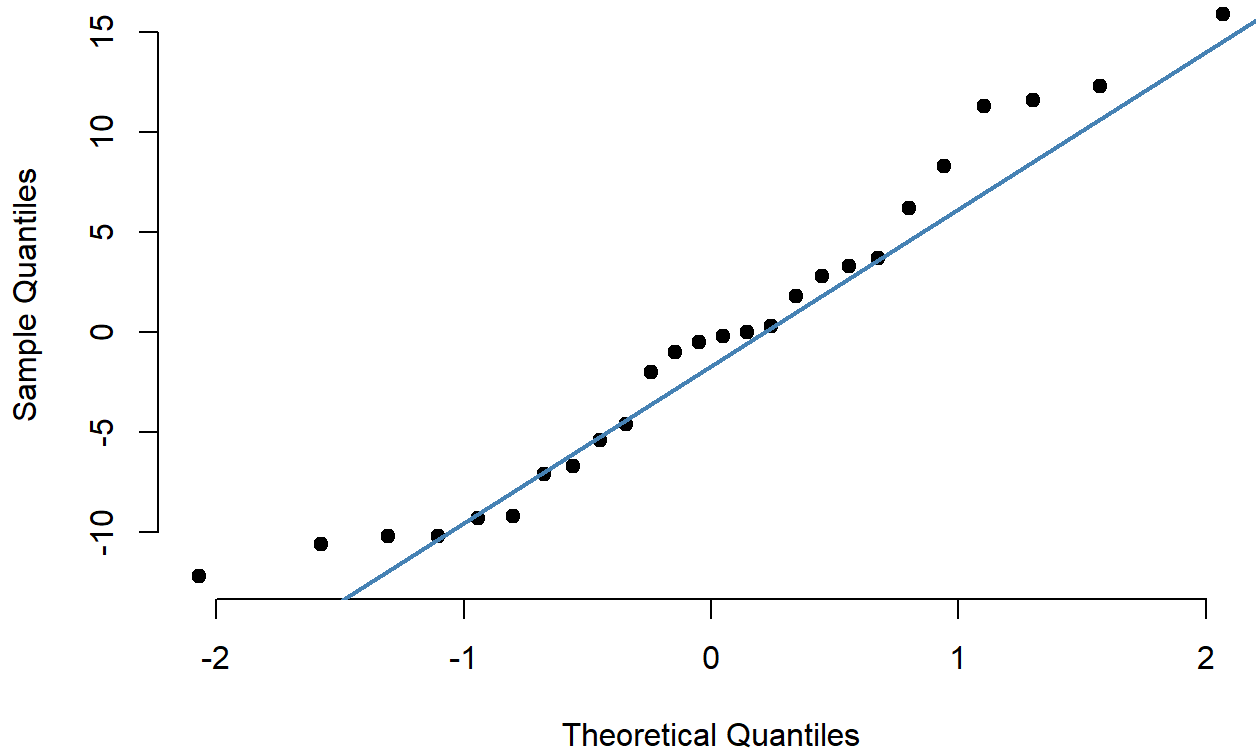

```
Gains.Fam <- Anorexia$Y[ Anorexia$Therapy == "Family" ]
qqnorm(Gains.Fam, pch = 19, frame = FALSE, main="Normal Q-Q Plot (Family")
qqline(Gains.Fam, col = "steelblue", lwd = 2)
```

Normal Q-Q Plot (Family