Paired Samples

MAT150 - Stats for Life Sciences

2023-05-01

R Commands covered in this lesson

`t.test` , `pt` , `if` , `print` , `mean` , `sd` , `sqrt` , `as.numeric`

Paired Samples for Comparing Two Populations or Treatement means

In many cases we are interested in comparing means before and after a treatment or between dependent individuals. For example,

- Pre and Post-treatment
- Twins
- Husband and wife

Summary

1. Null Hypotheis, $H_0 : \mu_d = D$
2. Test statistics

$$t = \frac{\bar{x}_d - D}{s/\sqrt{n}}$$

3. Calulate p-value using $df = n - 1$.

```
p = pt(t,df,lower.tail = F)   # Right tail test
p = pt(t,df,lower.tail = T)   # Left tail test
p = 2*pt(t,df,lower.tail = F) # Two tail, but t > 0
p = 2*pt(t,df,lower.tail = T) # Two tail, but t < 0
```

4. Reject H_0 if $p < \alpha$.

Example

```
memory = read.table('https://raw.githubusercontent.com/jamesquinlan/Intro-Stats-MAT150/main/data/examples/ex11_06.txt', header = T)
```

```
pre = memory$memory_test[memory$test == "pre"]
post = memory$memory_test[memory$test == "post"]
results = t.test(pre, post,
                 alternative = "less",
                 paired = T,
                 conf.level = 0.95)
n = length(pre)
tcrit = as.numeric(results$statistic["t"])
```

Confidence Interval

```
a = mean(pre - post) -  tcrit * sd(pre - post)/sqrt(n)
b = mean(pre - post) +  tcrit * sd(pre - post)/sqrt(n)
c(a,b)
```

```
## [1]    0.0000 -289.1667
```

Example

Ultrasound is often used as treatment for soft tissue injury. A range of motion study conducted on knee extension. The data is:

```
pre  = c(31,53,45,57,50,43,32)
post = c(32,59,46,64,49,45,40)
```

Given 90% confidence level, is treatment effective?

```
t.test(pre, post,
       paired = T,
       alternative = "less",
       conf.level = 0.9)
```

```
##
##  Paired t-test
##
## data:  pre and post
## t = -2.588, df = 6, p-value = 0.02066
## alternative hypothesis: true mean difference is less than 0
## 90 percent confidence interval:
##       -Inf -1.521179
## sample estimates:
## mean difference
##       -3.428571
```

Exercise

Do men overstate their height in online dating profiles? We have $\bar{x}_d = 0.57$ (profile - actual), $s_d = 0.81$, and $n = 40$.

The null hypothesis is: $H_0 : \mu_d = 0$ and $H_1 : \mu_d > 0$.

```
# Aggregate Data
xd = 0.57
s  = 0.81
n = 40

# Calculate test statistic
t = xd/(s/sqrt(n))

# Calculate p-value
p = pt(t,n-1, lower.tail = F)

# Make Decision
if(p < 0.05){
  print("Reject the Null")
  }else{
  print("Fail to reject the Null")
}
```

```
## [1] "Reject the Null"
```

References

Peck, R., & Devore, J. L. (2011). *Statistics: The Exploration & Analysis of Data*. Cengage Learning.