Math650 Homework 3.1

Yu Huang

October 31, 2023

Abstract

Fish Oil and Blood Pressure

1 Introduction

The question is whether fish oil could reduce the blood pressure.

2 Materials and Methods

Questions from *Statistical Sleuth*, Chap 1, No.12, Chap 2, No.13, No.14.
Data from the CDROM. Using software R.

3 Results

Fish Oil Average	6.571429
Regular Oil Average	-1.142857
Fish Oil standard deviation	5.8554
Regular Oil standard deviation	3.184785
pooled estimate of standard deviation	4.713203
standard error for the difference	2.519313
degree of pooled estimate of standard deviation	12
corresponding t-dist 97.5th percentile	2.179
95% confidence interval of $\mu_2 - \mu_1$	2.224702-13.20387
t-statistic for testing equality	3.062059
one-sided p-value(looking up table)	0.005
below is generated by R	
corresponding t-dist 97.5th percentile	2.178813
95% confidence interval of $\mu_2 - \mu_1$	2.225174-13.20340
one-sided p-value	0.00493062

All R codes are appended(5).

4 Conclusion and Discussion

Regarding question Chap1 No. 12, though the volunteers do not constitute a random sample from any population and we can't do *Inference to Population*, the *Allocation of Units to Groups* is random. It's still valid to do *Causal Inference*.

Based on the 0.005 p-value, it's very unlikely that the reduction of blood pressure after taking fish oil happens by chance.

5 Appendix

```
t_test_func = function(sample_data1, sample_data2)
{
len_f = length(sample_data1)
len_r = length(sample_data2)

mean_f = mean(sample_data1)
cat("mean_f:", mean_f, "\n")
mean_r = mean(sample_data2)
cat("mean_r:", mean_r, "\n")
cat("mean difference:", mean_f-mean_r, "\n")

sd_f = sd(sample_data1)
cat("sd_f:", sd_f, "\n")
sd_r = sd(sample_data2)
cat("sd_r:", sd_r, "\n")

degrees_of_freedom_of_pooled_sd = len_f + len_r -2
cat("degrees_of_freedom_of_pooled_sd:", degrees_of_freedom_of_pooled_sd, "\n")

pooled_sd = sqrt( ( (len_f-1)*sd_f^2 + (len_r-1)*sd_r^2  )/degrees_of_freedom_of_pooled_sd
cat("pooled_sd:", pooled_sd, "\n")

standard_error_for_the_difference = pooled_sd * sqrt(1/len_f + 1/len_r)
cat("standard_error_for_the_difference:", standard_error_for_the_difference, "\n")

#by looking up a table with 12=degrees_of_freedom_of_pooled_sd (only for 1st part of hw3)
percentile_97_5th = 2.179
cat("percentile_97_5th, df=12:", percentile_97_5th, "\n")
conf_interv_of_difference_of_mu_lower = mean_f-mean_r- percentile_97_5th*standard_error_fo
conf_interv_of_difference_of_mu_upper = mean_f-mean_r + percentile_97_5th*standard_error_f
cat("conf_interv_of_difference_of_mu_lower, df=12:", conf_interv_of_difference_of_mu_lower
cat("conf_interv_of_difference_of_mu_upper, df=12:", conf_interv_of_difference_of_mu_upper

t_stat = (mean_f - mean_r - 0)/standard_error_for_the_difference
cat("t_stat:", t_stat, "\n")

#by looking up a table with degrees_of_freedom_of_pooled_sd=12
#p_value = 1-0.995 = 0.005 corresponding to t_stat=3.055, actual p_value should be less th


#below is getting confidence interval and p_value through R

percentile_97_5th = qt(0.025, degrees_of_freedom_of_pooled_sd, lower.tail=FALSE)
cat("percentile_97_5th:", percentile_97_5th, "\n")
```

```
conf_interv_of_difference_of_mu_lower = mean_f-mean_r- percentile_97_5th*standard_error_fo
conf_interv_of_difference_of_mu_upper = mean_f-mean_r + percentile_97_5th*standard_error_f
cat("conf_interv_of_difference_of_mu_lower:", conf_interv_of_difference_of_mu_lower, "\n")
cat("conf_interv_of_difference_of_mu_upper:", conf_interv_of_difference_of_mu_upper, "\n")
p_value = pt(t_stat, degrees_of_freedom_of_pooled_sd, lower.tail=FALSE)
cat("p-value of t_stat", p_value, "\n")

}

#chap2 No13, No14

data = read.csv("/usr/local/doc/statistical_sleuth/ASCII/ex0112.csv")
sample_data1 = data[data$DIET=="fishoil",]$BP
sample_data2 = data[data$DIET=="regularoil",]$BP
t_test_func(sample_data1, sample_data2)
```

Output is this:

```
mean_f: 6.571429
mean_r: -1.142857
mean difference: 7.714286
sd_f: 5.8554
sd_r: 3.184785
degrees_of_freedom_of_pooled_sd: 12
pooled_sd: 4.713203
standard_error_for_the_difference: 2.519313
percentile_97_5th, df=12: 2.179
conf_interv_of_difference_of_mu_lower, df=12: 2.224702
conf_interv_of_difference_of_mu_upper, df=12: 13.20387
t_stat: 3.062059
percentile_97_5th: 2.178813
conf_interv_of_difference_of_mu_lower: 2.225174
conf_interv_of_difference_of_mu_upper: 13.20340
p-value of t_stat 0.00493062
```