Tutorial Business Analytics

Homework 2

Exercise 2.4

32 randomly selected men and women participated in a clinical trial. The purpose of the study was to compare vegetarian diets to non-vegetarian diets. The hypothesis to be tested is: "On average, vegetarians eat fewer calories than non-vegetarians". The sample mean for the 12 vegetarians is $\bar{x}_1 = 1780$ calories per day, while the sample mean for non-vegetarians amounts to $\bar{x}_2 = 1900$ calories per day. Moreover, the sample standard deviations are: $s_1 = 230$ and $s_2 = 250$.

a) Calculate a 95% confidence interval for the average daily intake of each group.

b) How do you assess above hypothesis considering the confidence intervals from question a)?

c) Which test is suitable for testing above hypothesis? Briefly explain your choice and perform the test with significance level $\alpha = 0.05$ and 25 degrees of freedom.

Exercise 2.5

a) Assess whether the following sample could possibly have been taken from a population with mean equal to 0. $(\alpha = 0.05)$

2 3 2 4 2 4 5 2 1 4 3 0 3 2 4 5 3 3 0 1

Solve this question manually (pen & paper) and then a second time using R (use the function "t.test()")

b) Briefly explain the term p-Value.

Confidence Intervals

Find confidence intervals for μ_x, which—under H_0—contain the true value μ_x with a probability of at least $1-\alpha$ (confidence level). We differentiate two cases:

- σ_x known:

  confidence interval: $$[I_u(x),\ I_o(x)] = \left[\bar{x} - z^c_{1-\frac{\alpha}{2}}\frac{\sigma_x}{\sqrt{n}},\ \bar{x} + z^c_{1-\alpha/2}\frac{\sigma_x}{\sqrt{n}}\right]$$

- σ_x unknown, use s_x as estimate instead:

  confidence interval: $$[I_u(x),\ I_o(x)] = \left[\bar{x} - t^c_{1-\frac{\alpha}{2};\,n-1}\frac{s_x}{\sqrt{n}},\ \bar{x} + t^c_{1-\frac{\alpha}{2};\,n-1}\frac{s_x}{\sqrt{n}}\right]$$

- Values of μ_0 within the confidence interval cannot be rejected regarding a significance level of α
  → Reject H_0 if μ_0 is not in the confidence interval

a) veget. [1554, ~~1846~~] 1634, 1926

$x = $ ~~1700~~ 1780

s 230 230

$[1700 \pm \underbrace{2.201 \times \frac{230}{\sqrt{12}}}_{146}$

non-veg.

1900

250

n = 20

$t_{0.975;19} = 2.093$

[1783, 2017]

2c) two samples, independent

$H_1: \mu_x < \mu_0$, $H_0\ \mu_x \geq \mu_0$

$$s^2_{x-w} = \frac{230^2}{12} + \frac{250^2}{20} = 7555$$

$$s_{x-c} = 86,8$$

$$\frac{1700 - 1900}{86,8} = -1,38 = t_0$$

$$t_{0.95;175} = 1.720$$

$$-1.38 < -1.720$$