Two sample test

Kingsuk_Jana

2023-03-25

1.Solution(Problem 35, Page 465)

An experiment was done to measure the effects of ozone, a component of smog. A group of 22 seventy-day-old rats were kept in an environment containing ozone for 7 days, and their weight gains were recorded. Another group of 23 rats of a similar age were kept in an ozone-free environment for a similar time, and their weight gains were recorded. The data (in grams) are given below. Analyze the data to determine the effect of ozone. Write a summary of your conclusions. [This problem is from Doksum and Sievers (1976) who provide an interesting analysis.]

Data set

```
Controls <- c(41,25.9, 13.1, -16.9, 15.4, 22.4,  29.4,26.0,38.4,21.9,27.4, 17.7, 21.4,26.6,24.9,18.3,28
26.0,22.7)
Ozone <- c(10.1,7.3,-9.9,17.9,6.6,39.9,-14.7,-9.0,6.1,14.3,6.8,-12.9,12.1,-15.9,44.1,20.4,15.5,28.2,14,
```

Step1

a. Set up the hypothesis

Null hypothesis: Average weights of rates are same for control and ozone (treated) group. Alternative hypothesis: Average weight of rates different for control and treated group.

b. Condition

I. Check normality: Check whether two sample coming from normal distribution or not. II. Independence: Here observations in control group totally independent on observation in ozone group. III. Check Variance: Check Whether two samples have equal variance or not.

Normality (using Wilks and Sapiro test) test and variance (equality) test.

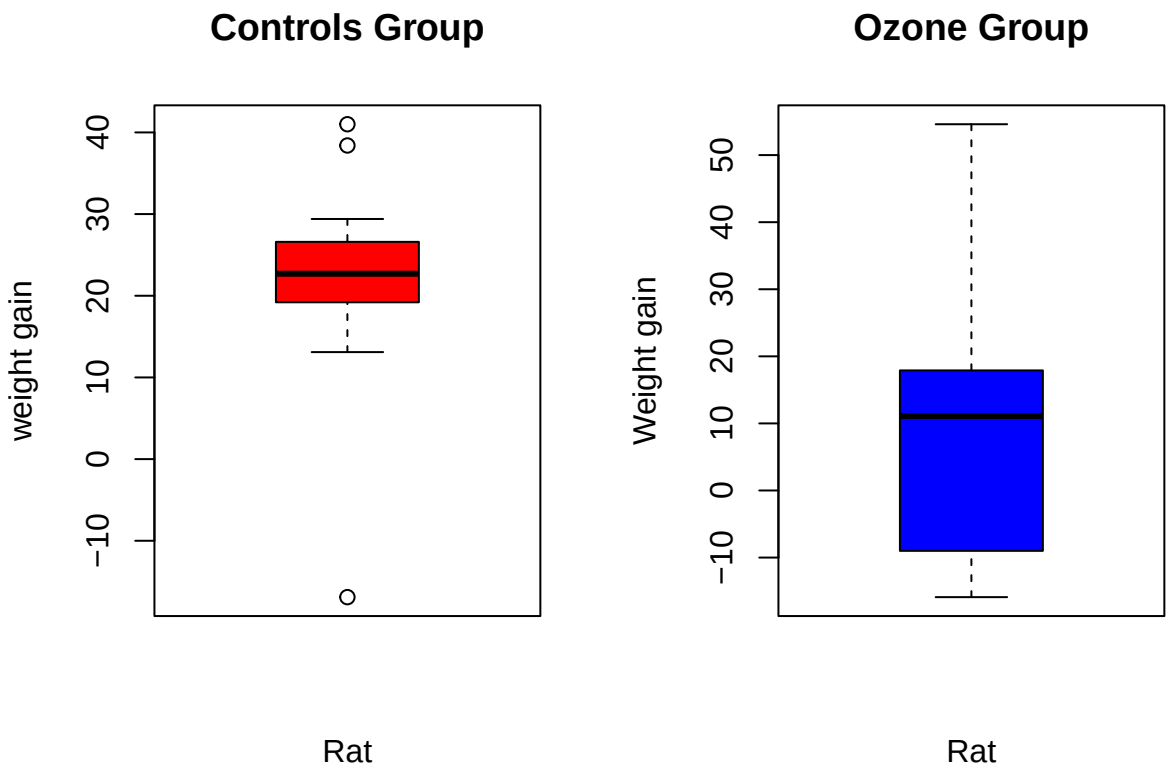
```r
par(mfrow = c(1,2))
shapiro.test(Controls) # Test normality for Controls group.
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Controls
## W = 0.79558, p-value = 0.0005619
```

```r
boxplot(Controls, xlab = "Rat", ylab = "weight gain", main="Controls Group", col = "red")
shapiro.test(Ozone) # Test for normality for Ozone group.
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Ozone
## W = 0.93558, p-value = 0.1603
```

```r
boxplot(Ozone, xlab = "Rat", ylab = "Weight gain", main = "Ozone Group", col = "blue")
```




```r
var.test(Controls,Ozone) # Test for variance.
```

```
##
##  F test to compare two variances
##
## data:  Controls and Ozone
## F = 0.34647, num df = 20, denom df = 21, p-value = 0.02115
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1428902 0.8483043
## sample estimates:
## ratio of variances
##          0.3464709
```

We checked the normality assumption for the Control group and Ozone group using Shapiro-Wilk normality test. We found that the normality assumption is rejected for control group(p-value=0.0005) and accepted for Ozone group (p-value = 0.1603). We used F test to check whether the variance of two group are same or not and we got the p-value = 0.02, which means variance are significantly different fro two groups.

Step2

Decide the significance level, alpha:

This value is used as a probability cutoff for making a decisions about the null hypothesis. The most common alpha **value** is 0.05 or 5%. Other popular choice are 0.01(1%) and 0.1(10%). Here we consider **alpha** as 0.05.

Step3

Calculate the test statistic.

```
# Two independent sample T-test, with unequal variance
T_test <- t.test(Controls, Ozone, var.equal = FALSE)
T_test
```

```
##
##  Welch Two Sample t-test
##
## data:  Controls and Ozone
## t = 2.4135, df = 34.271, p-value = 0.02129
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   1.807547 21.040938
## sample estimates:
## mean of x mean of y
##  22.43333  11.00909
```

From the above result our t statistics value is 2.41 and corresponding p-value is 0.021.

Step_4

Calculate the P-value :

The p-value corresponding to the test statistics is 0.021.

Step_5

Make decesion about the null hypothesis:

Here as we can see the p-value=0.021, which is less than the significance level 0.05. So we can reject the null hypothesis and conclude that average weight of rats in Ozone group significantly different from the average weight of rats in controls group.