One sample t-test

Cédric HASSEN-KHODJA

1 juillet 2016

Introduction

According to the World Health Organization (WHO), hypertension is the leading cause of death in the world. 51% and 45% of deaths associated with stroke and ischemic heart disease (reduced blood flow to the heart), respectively, can be attributed to hypertension.
Blood pressure is the power or pressure exerted by the blood on the walls of the arteries. It is expressed in two values, the systolic pressure (SBP) and the diastolic pressure (DBP). The first corresponds to the pressure when the heart contracts (systole), pumping blood in the arteries towards the lungs. The second designates the pressure exerted by the flow of blood when the ventricles expand (diastole) in order to collect blood channeled to the atrial chambers by the pulmonary veins and vena cava. Blood pressure is frequently measured in millimeters of mercury (mmHg).

Methods

A man has taken this systolic blood pressure for 24 hour. The average result of the french population for men is equal to 127.2 mmHg. You want to check if the sample mean is significantly similar (when the significance level is 95%) to the average population, assuming that the variance of the population is not known.
Data for the average population is provided by withings

155, 164, 163, 149, 158, 143, 163, 146, 144, 138, 137, 147, 136, 139, 123, 141, 139, 148, 134, 130, 137, 133, 132, 156, 163, 137, 154, 132, 130, 155, 147, 132, 142, 129, 119, 125, 139, 123, 124, 129, 134, 126, 123, 139, 131

Assumptions

Assumption 1

Your dependent variable should be continous.

```
is.numeric(data_clean$systolic)
```

```
## [1] TRUE
```

Assumption 2

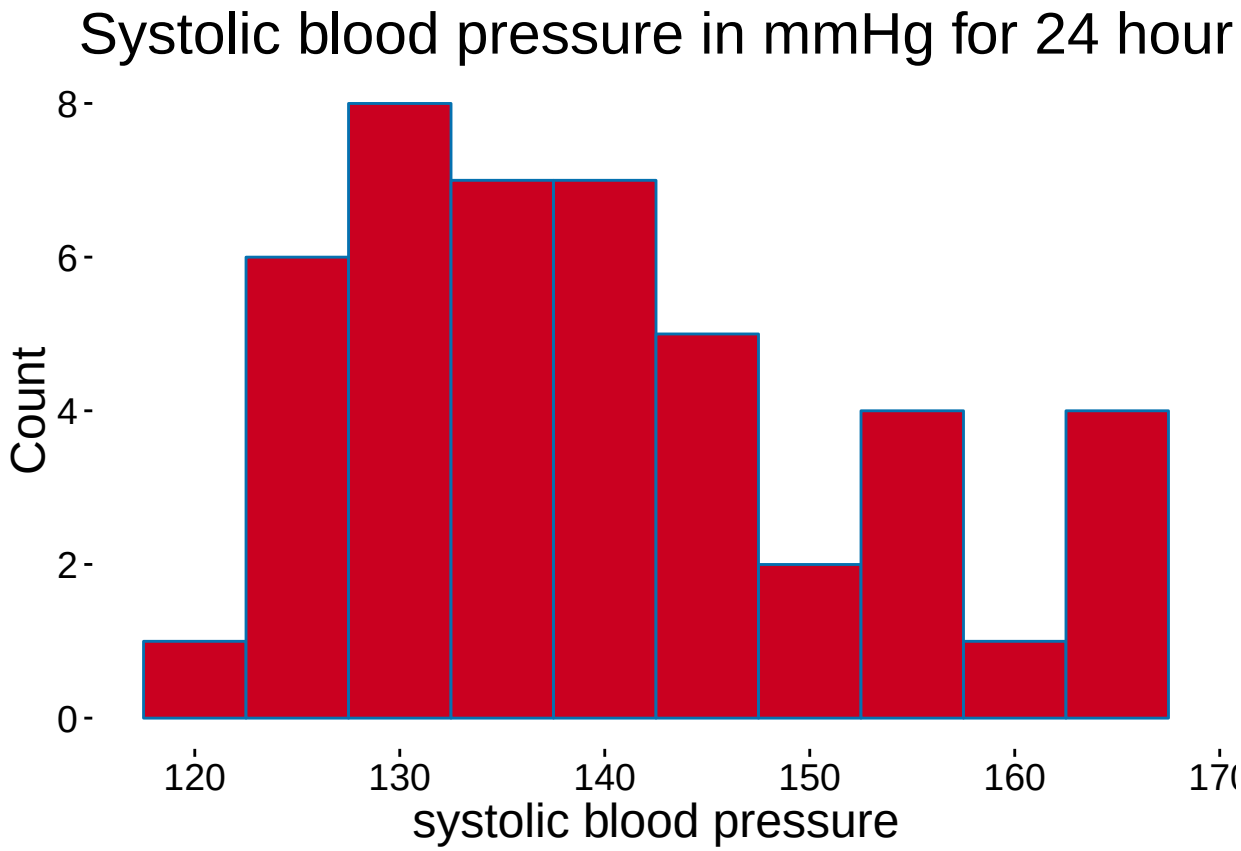
Your data is independent. Measures taken on the same subject every hour for 24 hour.

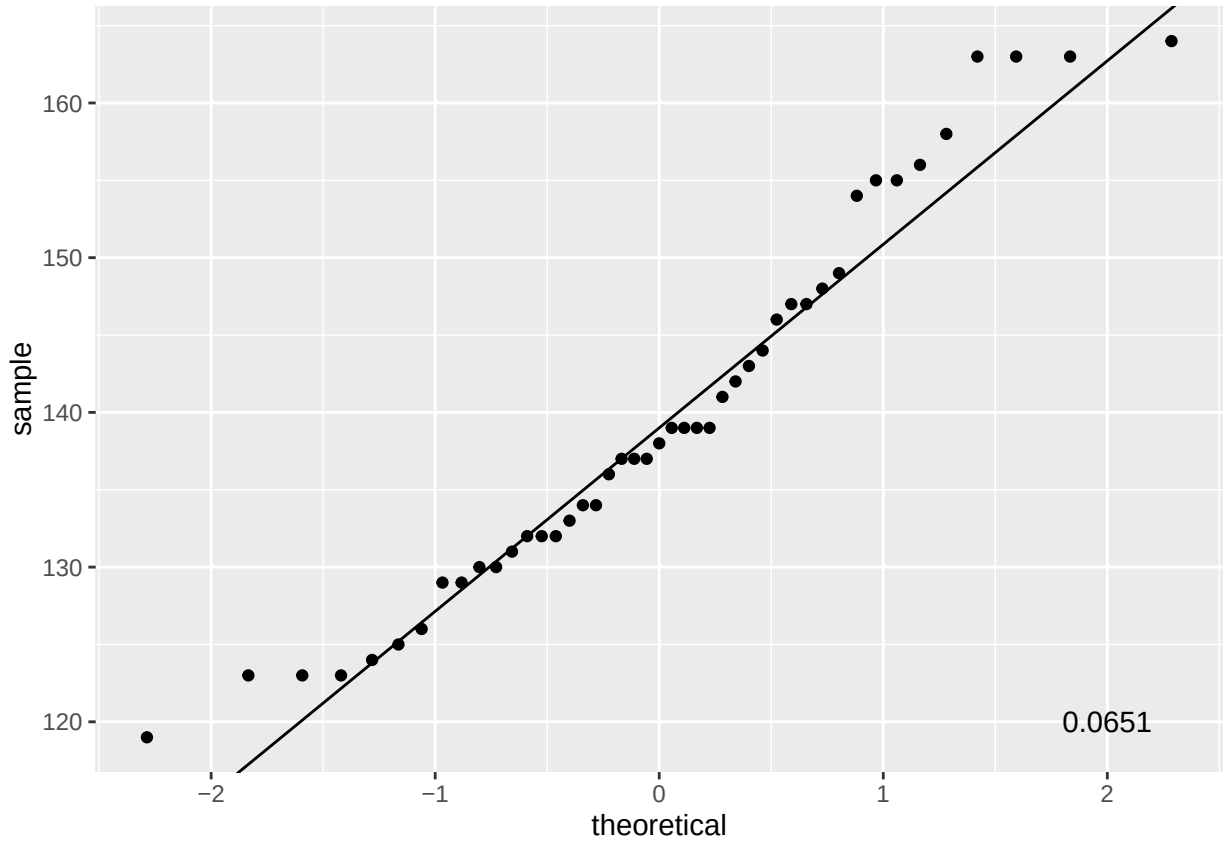
Assumption 3

There should be no significant outliers.
No values are higher than 164.1665447 or less than 115.300122 thresholds. No outliers are present in the data.

Assumption 4

Your dependent variable should be approximately normally distributed.





Results

The Student's t-test for a single sample have a pre-set function in R we can apply immediately.
t.test (a, mu)

```
##
##  One Sample t-test
##
## data:  data_stats$systolic
## t = 6.8821, df = 44, p-value = 1.702e-08
## alternative hypothesis: true mean is not equal to 127.2
## 95 percent confidence interval:
##  136.0631 143.4036
## sample estimates:
## mean of x
##  139.7333
```

The function **t.test** on one sample provides in output the value of t calculated; also gives us degrees of freedom, the confidence interval and the average (mean of x).

Conclusion

In order to take your statistic decision, you can proceed in two ways. We can compare the value of t with the value of the tabulated student t with 9 degrees of freedom. If we do not have tables, we can calculate the

value t-tabulated in the following way: **qt(0.975,44)**

```
## [1] 2.015368
```

The function **qt(p,df)** returns the value of t-tabulated considering the significance level (we chose a significance level equal to 95%, which means that each tail is the 2.5% which corresponds to the value of p = 1 - 0.025), and the degrees of freedom. By comparing the value of t-tabulated with t-computed, t-computed appears higher, which means that we don't accept the null hypothesis of equality of the averages: **our sample mean is significantly different to the mean of the population.**
Alternatively we could consider the p-value. With a significance level of 95%, remember this rule: If p-value is greater than 0.05 then we accept the null hypothesis H0; if p-value is less than 0.05 then we reject the null hypothesis H0 in favor of the alternative hypothesis H1. **Here 1.702066e-08 is less than 0.05 then our systolic blood pressure mean is significantly different to the mean of the systolic blood pressure population.**