R Tutorial for STAT 350 Lab 6

Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi

t-Procedure for One Population

The same function, `t.test()`, is used for both the confidence interval and the associated hypothesis test. When you are doing the analysis, only one function call should be used for both the confidence interval and hypothesis test for a particular situation. Points will be taken off if there are two `t.test()` calls for one question unless more than one confidence interval or more than one hypothesis test is requested.

Example (DATA SET: DMS.txt) Many food products contain small quantities of substances that would give an undesirable taste or smell if they were present in large amounts. An example is the "off-odors" caused by sulfur compounds in wine. Oenologists (wine experts) have determined the odor threshold, the lowest concentration of a compound that the human nose can detect. For example, the odor threshold for dimethyl sulfide (DMS) is given in the oenology literature as 25 micrograms per liter of wine (μg/L). Untrained noses may be less sensitive, however. Listed below are the DMS odor thresholds for 10 beginning students of oenology:

31 31 43 36 23 34 32 30 20 24

We are interested to see if untrained noses have a higher odor threshold than trained noses.

a) Make a boxplot and histogram to verify that the distribution is roughly symmetric with no outliers. The histogram must include the two curves (i.e., the estimated normal density curve in blue, and the kernel density curve in red) that help to assess normality.
b) Make a Normal probability plot to confirm that there are no systematic departures from Normality.
c) From your observations in parts a) and b), is it appropriate to use inferences using the t-procedure?
d) Calculate and interpret the 95% lower bound for the mean DMS odor threshold among all beginning oenology students.
e) Are you convinced that the mean odor threshold for beginning students is higher than the published threshold, 25 μg/L? Carry out a significance test to justify your answer. Your significance level should be consistent with what was given in part d).
f) Compare the results from parts d) and e). Explain how they lead to equivalent conclusions. Also discuss the notion of practical significance (i.e., whether or not there is a practical difference between the mean odor threshold and the 25 μg/L from the literature).

Solution: Code

To read in the data set: Import Dataset → From CSV → Browse to find file → Delimiter: Tab, Name: "wine" → Import

```
library(ggplot2)
#BOXPLOT
windows()
```

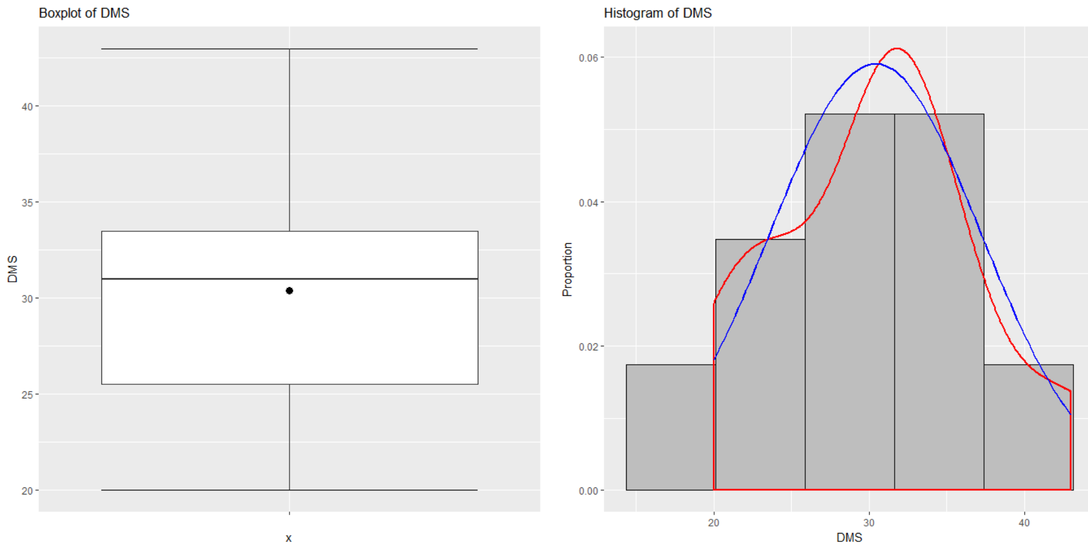
Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi

```
ggplot(wine, aes(x = "", y = DMS)) +
  stat_boxplot(geom = "errorbar") +
  geom_boxplot() +
  ggtitle("Boxplot of DMS") +
  stat_summary(fun.y = mean, col = "black", geom = "point", size = 3)
# HISTOGRAM
xbar <- mean(wine$DMS)
s <- sd(wine$DMS)
windows()
ggplot(wine, aes(DMS)) +
  geom_histogram(aes(y = ..density..),
                 bins = sqrt(nrow(wine))+2,
                 fill = "grey", col = "black") +
  geom_density(col = "red", lwd = 1) +
  stat_function(fun = dnorm, args = list(mean = xbar, sd = s),
                col="blue", lwd = 1) +
  ggtitle("Histogram of DMS") +
  xlab("DMS") +
  ylab("Proportion")
# QQPlot
windows()
ggplot(wine, aes(sample = DMS)) +
  stat_qq() +
  geom_abline(slope = s, intercept = xbar) +
  ggtitle("QQ Plot of DMS")
# Parts d) and e), the same code should be used
# Parameters for t.test():
# - mu: mu_0 of the null hypothesis
# - conf.level: confidence level of the confidence interval,
#     which also indicates the significance level of the hypothesis
#     test, alpha, by "alpha = 1 - confidence level"
# - alternative: form of the alternative hypothesis and confidence
#     interval/bounds, possible options including
#     - "two.sided" (not equal to, confidence interval)
#     - "less" (<, upper confidence bound)
#     - "greater" (>, lower confidence bound)
t.test(wine$DMS, conf.level = 0.95, mu = 25, alternative = "greater")
```

a) Make a boxplot and histogram to verify that the distribution is roughly symmetric with no outliers. The histogram must include the two curves (i.e., the estimated normal density curve in blue, and the kernel density curve in red) that help to assess normality.

Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi

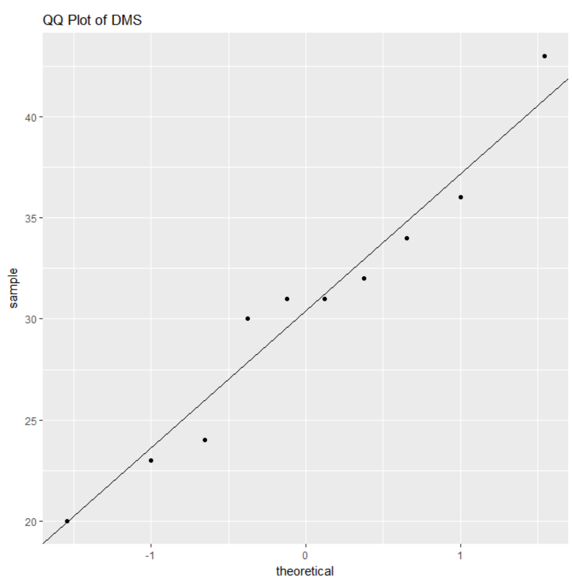
Solution:



The median is approximately equal to the mean suggesting that the distribution is approximately symmetric. The histogram indicates that it is close to normal and also symmetric. I can determine the symmetry and normality better on the histogram than the boxplot because I can relate the blue and red curves easier. I see no outliers in the boxplot or the histogram. Remember to always create a modified boxplot.

b) Make a Normal probability plot to confirm that there are no systematic departures from Normality.

Solution:



The points on the probability plot roughly follow a straight line. This indicates that the distribution is approximately normal.

Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi

c) From your observations in parts a) and b), is it appropriate to use the t-procedure?

Solution:

Assuming that the sample is SRS, the only other assumption that is necessary is that the distribution is normal. Since the sample size is 10, we cannot use CLT. However, from the normal probability plot and the histogram, we can see that the distribution is approximately normal. Therefore, this assumption is met.

d) Calculate and interpret the 95% lower bound for the mean DMS odor threshold among all beginning oenology students.

Solution:

The same code (and output) should be used for both parts d) and e).

```
        One Sample t-test

data:  wine$DMS
t = 2.5288, df = 9, p-value = 0.01615
alternative hypothesis: true mean is greater than 25
95 percent confidence interval:
 26.48554      Inf
sample estimates:
mean of x
     30.4
```

The part highlighted in yellow is used for part d)

The 95% lower bound is 26.48554. R says that the 95% confidence interval is (26.48554, ∞).

We are 95% confident that the mean odor threshold for beginning students is greater than 26.48554 μg/L.

e) Are you convinced that the mean odor threshold for beginning students is higher than the published threshold, 25 μg/L? Carry out a significance test to justify your answer. Your significance level should be consistent with what was given in part d).

Solution:

The output for this part is highlighted in green above.

Step 1: Definition of the terms

μ is the population mean odor threshold for beginning students.

Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi

Step 2: State the hypotheses

H_0: $\mu = 25$
H_a: $\mu > 25$

Step 3: Find the *test statistic, report DF, find the p-value*

$t_{ts} = 2.5288$
DF = 9
P-value = 0.01615

Step 4: Conclusion:

Since 0.01615 ≤ 0.05 (0.05 = 1 – 0.95), we should reject H_0

The data provides evidence (p = 0.01615) to the claim that the mean odor threshold for beginning students is higher than the published threshold, 25 µg/L.

f) Compare the results from parts d) and e). Explain how they lead to equivalent conclusions. Also discuss the notion of practical significance (i.e., whether or not there is a practical difference between the mean odor threshold and the 25 µg/L from the literature).

Solution:

From the confidence bound, we know that we are 95% confident that the true value (i.e., the population mean of beginning students) is greater than 26.4855. Since 26.4855 is greater than 25, we are also confident that the true value is greater than 25. Therefore, the conclusion from the hypothesis test would be to reject H_0.

However, we are always interested in the practical significance, and for that we compare two differences between 1) the sample average and the theoretical value, and 2) the confidence bound and the theoretical value. Here, the “theoretical value” refers to the population mean odor threshold under the null hypothesis.

1) When we are comparing the sample average to the theoretical value, we should consider what our maximum tolerance is in practice. If the difference is greater than that tolerance, then the sample average and the theoretical value are considered practically different. For example, if the sample mean odor threshold is 25.1 µg/L and the tolerance of difference is 0.5 µg/L, people would regard it practically the same as 25 µg/L. However, if the standard error (standard deviation of the estimator) is also very small (for example, when sample size is very large), then it is possible that 25.1 µg/L, though practically the same as 25 µg/L, would be considered statistically different.

In this case, I would say that 30.4 µg/L is practically different from 25 µg/L as their difference is about 20% of the theoretical value: $\frac{30.4-25}{25} = 21.6\%$.

Author: Leonore Findsen, Chunyan Sun, Sarah H. Sellke, Jeremy Troisi

Note that you should check for the practical difference BEFORE you do the inference. If you do not believe that the sample average is different from the theoretical value according to practicality, it doesn't matter what the inference says. On the other hand, if you think the sample average could be different from the theoretical value practically, you should proceed to see whether there is statistical evidence based on the inference. However, in this class, the practicality question always comes last so you should always perform the inference.

2) We also need to compare the confidence bounds to the theoretical value. For instance, if the lower confidence bound was 25.1 μg/L, the null would be rejected. Again, is 25.1 a reasonably likely value of the true parameter (do not forget the correct interpretation of confidence intervals)? We would hardly consider 25.1 μg/L a substantially different value from 25 μg/L. Hence, in this case we might conclude that the mean is statistically significantly different from 25 but the difference is not significant in practice.

Meanwhile, for example, if the lower bound were to be 35, we might conclude that the difference between the true population mean and the theoretical value is practically significant.

In the case of our problem, the lower bound is 26.4855. I consider a different of more than 1 μg/L practically different so this answer is practically different as well as statistically different.

In general, if we fail to reject the null, then we have neither evidence for a statistically significant difference nor a practical significant difference though if the sample size is very small, further studies might be warranted to confirm the results.

A full answer to the original question should consider both the statistical significance and the practical significance.