STAT 210
Applied Statistics and Data Analysis
Problem List 3
(due on week 4)

Joaquin Ortega

Fall 2022

Exercises

The Quality Control Engineer at a soft drinks plant wants to test a bottling machine that fills bottles with 500 ml of soda. From previous experience, the engineer knows that the normal distribution is a good approximation to the distribution of the content of the bottles.

The engineer draws a sample of size 10 from the machine's production and obtains the values

```
bottles <- c(494.04, 499.07, 497.03, 502.79, 495.57, 498.09, 500.40,
             491.80, 494.34, 498.65)
```

Part 1

1) Using this sample, estimate mean, variance, and standard deviation.

For this we use the functions `mean`, `var`, and `sd`.

```
(mean.btt <- mean(bottles))
```

```
## [1] 497.178
```

```
var(bottles)
```

```
## [1] 10.8762
```

```
(sd.btt <- sd(bottles))
```

```
## [1] 3.297908
```

2) Which is the correct sampling distribution for the sample mean in this situation?

Since we have to estimate the variance from the sample, the correct distribution is the t-distribution with $n-1=9$ degrees of freedom. Recall that

$$\frac{\hat{\mu}_n - \mu}{s_n/\sqrt{n}} \sim t_{n-1} \tag{1}$$

This is equation (5) in V15-IntervalEstimation.

3) Using this distribution, find a two-sided confidence interval at the 98% level for the mean.

We have from (1) that

$$P(t_{9,0.01} < \frac{\hat{\mu}_n - \mu}{s_n/\sqrt{n}} < t_{9,0.99}) = 0.98$$

where $t_{9,\alpha}$ is the α quantile for the t_9 distribution. Multiplying by $s_n/\sqrt{n}$ inside the probability, we get

$$P\Big(\frac{s_n t_{9,0.01}}{\sqrt{n}} < \hat{\mu}_n - \mu < \frac{s_n t_{9,0.99}}{\sqrt{n}}\Big) = 0.98$$

And from this, we get that

$$P\Big(\hat{\mu}_n - \frac{s_n t_{9,0.99}}{\sqrt{n}} < \mu < \hat{\mu}_n - \frac{s_n t_{9,0.01}}{\sqrt{n}}\Big) = 0.98$$

By symmetry, $t_{9,0.99} = -t_{9,0.01}$ and we finally get that

$$P\Big(\hat{\mu}_n - \frac{s_n t_{9,0.99}}{\sqrt{n}} < \mu < \hat{\mu}_n + \frac{s_n t_{9,0.99}}{\sqrt{n}}\Big) = 0.98$$

You can find this formula in slide 22 of V15-IntervalEstimation. We now proceed to evaluate the terms in the formula

```
n <- 10
error <- qt(0.99,df=n-1)*sd.btt/sqrt(n)
c(mean.btt-error,mean.btt+error)
```

```
## [1] 494.2356 500.1204
```

4) Using this sample, the engineer wants to test whether the machine fills the bottles correctly. What hypothesis test should he carry out? What is the test statistic? Calculate the value for this statistic using the sample above.

To test whether the machine is working properly, she should test

$$H_0 : \mu = 500 \qquad \text{vs.} \qquad \mu \neq 500$$

The test statistic is

$$\hat{t} = \frac{\hat{\mu}_n - \mu}{s_n/\sqrt{n}}$$

which by equation (1) has a t_9 distribution. The value for this statistic is

```
(t.hat <- (mean.btt - 500)/(sd.btt/sqrt(10)))
```

```
## [1] -2.705942
```

Part 2

5) The engineer always sets a level of confidence of 98% for these tests. Find a rejection region for the test at this level. Is the value of the test statistic inside or outside the rejection region? What is your conclusion?

Since this is a two-tailed test, the acceptance region is the interval $[t_{9,0.01}, t_{9,0.9}]$ since, under the null hypothesis,

$$P(\hat{t} \in [t_{9,0.01}, t_{9,0.99}]) = P(\hat{t} \in [t_{9,0.01}, -t_{9,0.01}]) = 0.98$$

The acceptance region is given by

```
c(qt(0.01,9), -qt(0.01,9))
```

```
## [1] -2.821438  2.821438
```

and we see that $\hat{t} = -2.705942$ falls inside this region, so we do not have enough evidence from this sample to reject the null hypothesis that the average filling level for the machine is 500 ml.

6) Carry this test out using a command in R and look at the p-value that you obtain. What is your conclusion?

The command is `t.test`:

```
t.test(bottles, mu = 500, conf.level = .98)
```

```
##
##  One Sample t-test
##
## data:  bottles
## t = -2.7059, df = 9, p-value = 0.02416
## alternative hypothesis: true mean is not equal to 500
## 98 percent confidence interval:
##  494.2356 500.1204
## sample estimates:
## mean of x
##   497.178
```

The p-value is 0.02416, which is slightly bigger than the level set by the engineer. Therefore, we do not have enough evidence to reject the null hypothesis at this level.

Looking at the confidence interval in the result for the test, we see that it coincides with the interval calculated in (3).

7) Since the p-value is close to the confidence level that she set, the engineer decides to take a new sample of size 20 and obtains the following values

```
bottles2 <- c(497.24, 497.43, 500.64, 490.59, 496.24, 497.44, 501.69,
              489.98, 493.83, 495.60, 504.33, 495.11, 497.94, 495.03,
              490.75, 498.16, 491.87, 492.94, 494.12, 500.08)
```

Using this new sample, repeat the test you carried out in (6) and comment on the results you obtain.

```
t.test(bottles2, mu = 500, conf.level = .98)
```

```
##
##  One Sample t-test
##
## data:  bottles2
## t = -4.5712, df = 19, p-value = 0.0002086
## alternative hypothesis: true mean is not equal to 500
## 98 percent confidence interval:
##  493.8564 498.2446
## sample estimates:
## mean of x
##  496.0505
```

The p-value now is 0.00158, which is small enough to reject the null hypothesis. The conclusion is that the machine is not working as expected.

Part 3

8) Suppose now that the company is not worried about underfilling the bottles since management argues that consumers do not notice a small difference in the amount of soda in the bottle, but it is concerned about overfilling the bottles, as this would mean less profit. What would be a reasonable test of hypothesis in this context? Carry out this test and comment on the results.

In this case the engineer should carry out a one-sided test:

$$H_0 : \mu = 500 \qquad \text{vs.} \qquad \mu > 500.$$

In `R`, we carry out this test with the command

```r
t.test(bottles2, mu = 500, conf.level = .98, alternative = 'greater')
```

```
##
##  One Sample t-test
##
## data:  bottles2
## t = -4.5712, df = 19, p-value = 0.9999
## alternative hypothesis: true mean is greater than 500
## 98 percent confidence interval:
##  494.1457      Inf
## sample estimates:
## mean of x
##  496.0505
```

The p-value is now 0.9992, and we cannot reject the null hypothesis in this test. The machine is not overfilling the bottles.

By the way, observe the confidence interval that we get in the one-sided case: $(496.0842, \infty)$.

9) In the lectures, we worked out an example about the power of a hypothesis test for the normal distribution. Assuming that the variance for the population was known, we calculated and plotted the power function. However, when the variance of the population is not known, calculating the power is not so easy since the sampling distribution under the alternative distribution changes. Fortunately, there is a function in `R` for doing this. It is `power.t.test`. Look at the help for this function to get familiar with the required arguments.

   Using `power.t.test`, calculate the power of the test

$$H_0 : \mu = 500 \qquad \text{vs.} \qquad \mu = 502$$

   using a confidence level of 98% for the two sample sizes that we have considered before.

```r
power.t.test(n = 10, delta = 2, sd = sd(bottles),
             type = 'one.sample', sig.level = 0.02,
             alternative = 'one.sided')
```

```
##
##      One-sample t test power calculation
##
##               n = 10
##           delta = 2
##              sd = 3.297908
##       sig.level = 0.02
##           power = 0.3600212
##     alternative = one.sided
```

```r
power.t.test(n = 20, delta = 2, sd = sd(bottles2),
             type = 'one.sample', sig.level = 0.02,
             alternative = 'one.sided')
```

```
##
##      One-sample t test power calculation
##
##               n = 20
##           delta = 2
##              sd = 3.863872
##       sig.level = 0.02
```

```
##           power = 0.5524943
##     alternative = one.sided
```

We can also combine the two samples since they are independent and come from the same experiment.

```r
power.t.test(n = 30, delta = 2, sd = sd(c(bottles,bottles2)),
             type = 'one.sample', sig.level = 0.02,
             alternative = 'one.sided')
```

```
##
##      One-sample t test power calculation
##
##               n = 30
##           delta = 2
##              sd = 3.66729
##       sig.level = 0.02
##           power = 0.7948682
##     alternative = one.sided
```

We see how the power increases from 0.36 with a sample of size 10 to 0.795 with a sample of size 30.

Part 4

10) We can use the `power.t.test` function to determine the sample size to obtain a given power for a fixed alternative and significance level for the test. Suppose the engineer wants to detect when the machine is overfilling the bottles by 2.5 ml with a probability of at least 0.7. Calculate the sample size.

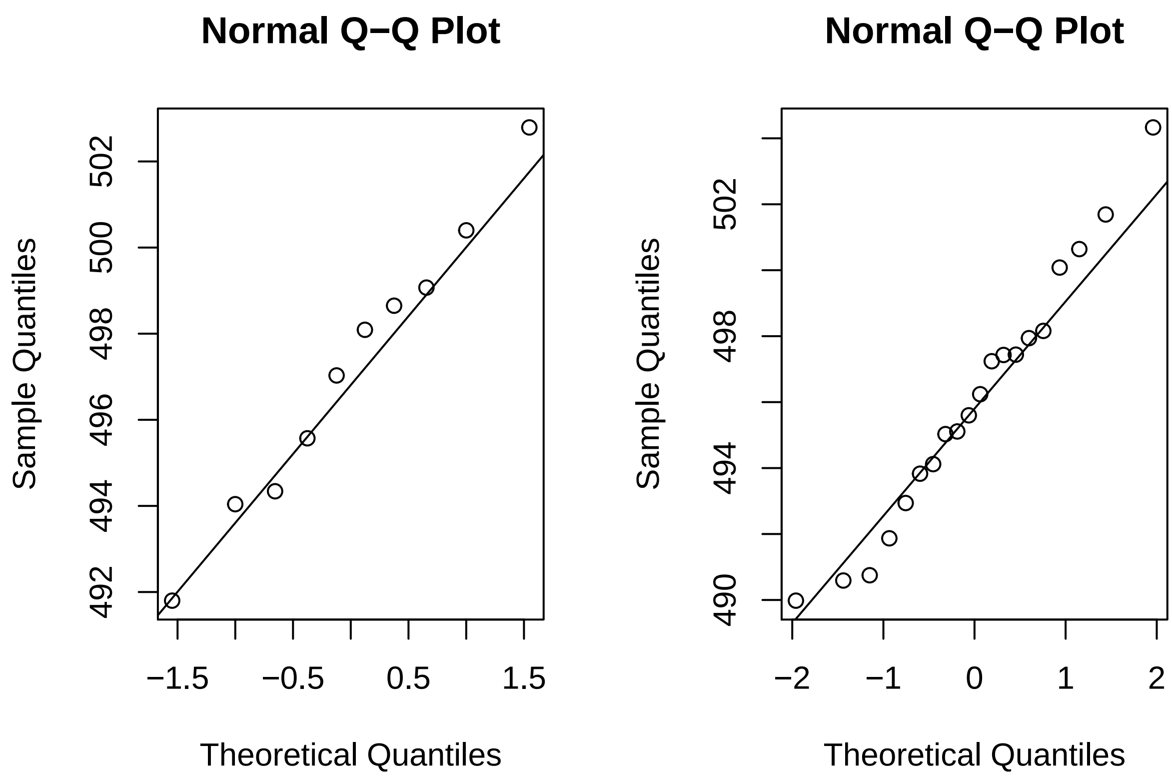
```r
power.t.test(delta = 2.5, sd = sd(bottles2), power = 0.7, sig.level = 0.02,
             type = 'one.sample', alternative = 'one.sided')
```

```
##
##      One-sample t test power calculation
##
##               n = 18.06907
##           delta = 2.5
##              sd = 3.863872
##       sig.level = 0.02
##           power = 0.7
##     alternative = one.sided
```

Therefore, a sample of size 19 would be enough.

11) All the calculations we have made rely on the fact that the engineer 'knows' by experience that the production of the machine follows a normal distribution. How would you verify this assumption for the two samples considered?

```r
par(mfrow=c(1,2))
qqnorm(bottles); qqline(bottles)
qqnorm(bottles2); qqline(bottles2)
```




```r
shapiro.test(bottles)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  bottles
## W = 0.98755, p-value = 0.9928
```

```r
shapiro.test(bottles2)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  bottles2
## W = 0.97604, p-value = 0.8736
```

12) Use a non-parametric test as an alternative to the tests carried out in (6) and (7) and compare your results.

```r
wilcox.test(bottles, mu = 500)
```

```
##
##  Wilcoxon signed rank test
##
## data:  bottles
## V = 6, p-value = 0.02734
## alternative hypothesis: true location is not equal to 500
```

```r
wilcox.test(bottles2, mu = 500)
```

```
##
##  Wilcoxon signed rank test
```

```
##
## data:  bottles2
## V = 16, p-value = 0.0003223
## alternative hypothesis: true location is not equal to 500
```

Although the p-values are different, we would reach the same conclusions for both samples.