Problem 1: (20 pts) A botanist is trying to determine whether the seeds treated with an experimental growth substance resulted in a higher average height of plants than the standard height of 305.3 mm. The botanist treated a random sample of 38 seeds with the extract and subsequently obtained the height data provided in the file PlantHeight.csv. What is his null hypothesis and should it be accepted at 5% significance level?

Solution: Since the sample size of 38 is likely to be large enough to produce approximately normal distribution of the sample mean (furthermore from the problem description it follows that the data is likely to be normally distributed), we can use the classical tests for the sample mean. We have only one set (vector) of data so we use one-sample test, and since the population variance is unknown we use the t-test. Finally we need the upper-tail t-test with the null hypothesis \( \mu = 305.3 \) mm and the alternative hypothesis \( \mu > 305.3 \) mm.

Null hypothesis: \( \mu = 305.3 \) mm

Alternative hypothesis: \( \mu > 305.3 \) mm

*R code:*

```r
d = read.csv("PlantHeight.csv")
t.test(d$Height,mu=305.3,alternative ="greater")  # default confidence level is 0.95
```

*R console output:*

```
One Sample t-test
data:  d$Height
t = 5.0126, df = 37, p-value = 6.784e-06
alternative hypothesis: true mean is greater than 305.3
95 percent confidence interval:
            306.7176      Inf
sample estimates:
  mean of x
307.4368
```

Test conclusion: Since the \( p \)-value is less than 0.05, reject the null hypothesis in favor of the alternative hypothesis \( \mu > 305.3 \). 

Problem 2: (20 pts) The addition of bran to the diet has been reported to benefit patients with diverticulosis. Several different bran preparations are available, and a clinician wants to test the efficacy of two of them on patients, since favorable claims have been made for each. Among the consequences of administering bran that requires testing is the transit time through the alimentary canal. By random allocation the clinician selects two groups of patients aged 40-65 with diverticulosis of comparable severity. SampleA.csv contains 45 patients who are given treatment A, and SampleB.csv contains 62 patients who are given treatment B. The transit times of food through the gut are measured by a standard technique with marked pellets and the results are recorded. Does it differ in the two groups of patients taking these two preparations? The null hypothesis is that the two groups come from the same population. Should it be accepted at 10% significance?

Solution: We have two samples of different sizes so the independent samples (unpaired) two-sample (two-tailed) t-test should be used. The difference in sample sizes is large (62 vs. 45) so separate variances test (and not the pooled variance) should be used. We shall compute the sample variances and their ratio, but even without this calculation the use of separate variances test is more appropriate based on the large difference in sample sizes.

Null hypothesis: Populations A and B have the same population means: \( \mu_A = \mu_B \).

Alternative hypothesis: Populations A and B have different population means: \( \mu_A \neq \mu_B \).

*R code:*

```r
a = read.csv("SampleA.csv")$A  # Note: a is the data vector, not frame
b = read.csv("SampleB.csv")$B  # b is the data vector
sd(a); sd(b); sd(b)/sd(a)  # The ratio is ~ 2.6 > 2
t.test(a,b,paired=FALSE,var.equal=FALSE,conf.level = 0.9)  # Separate variance
```

*R console output:*

```
Welch Two Sample t-test
```

...
data: a and b
\[ t = -1.7187, \text{ df } = 83.603, \text{ p-value } = 0.08937 \]
alternative hypothesis: true difference in means is not equal to 0
90 percent confidence interval:
\[ -2.35206630 \text{ -0.03852151} \]
sample estimates:
mean of x mean of y
9.404222 10.599516

Test conclusion: Since the \( p \)-value is less than the significance level 0.1, reject the null hypothesis in favor of the alternative hypothesis \( \mu_A \neq \mu_B \).

Problem 3: (20 pts) A manufacturer claims that the thickness of the steel plates it produces is 61 mils (thousandth of an inch). Car manufacturer’s quality control officer regularly checks the claim. From a recent shipment he took a random sample of 50 plates and measured their thickness. The data obtained is in the file PlateThickness.csv. In order to verify that the manufacturer’s claim is accurate, what should his null hypothesis be and should it be accepted at 10% significance level?

Solution: Since the sample size of 50 is likely to be large enough to produce approximately normal distribution of the sample mean (and the data itself is likely to be normally distributed), we can use the classical tests for the sample mean. We have only one set (vector) of data so we use one-sample test, and since the population variance is unknown we use the \( t \)-test. Finally this will be a two-sided test since we are only checking the accuracy of the 61 mils thickness claim.

Null hypothesis: \( \mu = 61 \)  
Alternative hypothesis: \( \mu \neq 61 \)

\( R \) code:

```r
pt = read.csv("PlateThickness.csv")$Thickness  # pt is the data vector, not frame
t.test(pt,mu=61,conf.level = 0.9)  # alternative="two.sided" is default
```

\( R \) console output:

One Sample t-test
data: pt
t = -0.76918, df = 49, p-value = 0.4455
alternative hypothesis: true mean is not equal to 61
90 percent confidence interval:
\[ 57.62958 \text{ 62.25042} \]
sample estimates:
mean of x
59.94

Test conclusion: Since the \( p \)-value is greater than 0.1, accept the null hypothesis.

Problem 4: (15 pts) NormalData.csv contains random numbers generated from the normal distribution with variance 1. Test the null hypothesis \( H_0: \mu = 28.41 \).

Solution: Since the data is normally distributed with known population variance we use the (one-sample) \( z \)-test. It is a two-sided test by the problem statement, since no indication is given whether the population mean should be larger or smaller than 28.41. Apparently \( R \) has no standard \( z \)-test function, so we can write a custom one (done below, adopted from Session08+.R) or import a function from some package.

Null hypothesis: \( \mu = 28.41 \)  
Alternative hypothesis: \( \mu \neq 28.41 \)

\( R \) code:

```r
z.test.pvals = function(v,mu0,pop_sd)
# returns p-values for lower-, two-, and upper-tailed z-test
{
zstat = (mean(v)-mu0)/(pop_sd/sqrt(length(v)))
aux = pnorm(zstat)
return(c(aux,2*pnorm(-abs(zstat)),1-aux))
}
```

Typo: should be 100.
\texttt{nd = read.csv("NormalData.csv")$Data}
\texttt{z.test.pvals(nd, mu0=28.41, pop_sd=1)}

\textit{R} console output:
\[
[1] \ 1.000000e+00 \quad 4.789598e-103 \quad 0.000000e+00
\]

Test conclusion: Since the \(p\)-value is less than 0.05, reject the null hypothesis in favor of the alternative hypothesis \(\mu \neq 28.41\).

\textbf{Note 1:} If you use the one-sided two-tailed \(t\)-test (\texttt{t.test(nd, mu=28.41)}) the null hypothesis will be accepted:

\begin{verbatim}
One Sample t-test
data:  nd
t = 1.9736, df = 119, p-value = 0.05075
alternative hypothesis: true mean is not equal to 28.41
95 percent confidence interval:
28.40349 32.35183
sample estimates:
mean of x
30.37766
\end{verbatim}

\textbf{Note 2:} The problem statement has a typo – the data was generated from the normal distribution with variance 100, not 1. Check the sample standard deviation, it is approx. 10.92. Notice that \texttt{z.test.pvals(nd, mu0=28.41, pop_sd=10)} produces the output
\[
[1] \ 0.98443724 \quad 0.03112553 \quad 0.03112553 \quad 0.03112553 \quad 0.01556276
\]
which gives a \(p\)-value of approx. 0.0311 and the null hypothesis is still rejected. That said, the \(p\)-values obtained by the \(z\)-test (0.0311) and the \(t\)-test (0.05075) are not that different.

\textbf{Problem 5:} (20 pts) A company researcher wants to test a new formula for a sports drink that has been designed to improve running performance. To carry out the experiment, the researcher recruited 41 middle distance runners. All of these participants performed two trials (on different days) in which they had to run as far as possible for 2 hours on a treadmill. In one of the trials, all participants drank from a bottle containing the new formula. In the other trial, the same participants drank from a bottle containing no formula. At the end of the two trials, the distance each participant ran (in km) was recorded and stored in \texttt{Distance.csv}. The researcher wants to know whether the formula truly improves performance. What is the null hypothesis and should it be accepted at 5% significance level?

\textbf{Solution:} This is a paired sample since each runner is measured twice, resulting in pairs of observations. The sample size of 40 is large enough to produce approximately normal distribution of the sample mean of the differences. Since the population variance of the differences is unknown we use the \textit{paired two-sample} \textit{t-test}. Finally this will be a one-sided test and we can use either the \textit{upper-tail} \(t\)-test or the \textit{lower-tail} \(t\)-test, with analogous reasoning as used in Problem 1. In this case the null-hypothesis is accepted at 5% significance no matter which tailed-test you choose. Alternatively, you can use the \textit{one-sample} \textit{t-test} on the vector of differences – the paired two-sample test is ‘converted’ into this test anyway.

\textbf{Null hypothesis:} The population means of ‘formula’ and ‘no-formula’ populations are equal: \(\mu_f = \mu_{nf}\)

\textbf{Alternative hypothesis:} The population mean of ‘formula’ population is greater than the population mean of the ‘no-formula’ population: \(\mu_f > \mu_{nf}\)

\textbf{R code:}
\begin{verbatim}
data = read.csv("Distance.csv")
head(data); f = data$Formula; nf = data$NoFormula
t.test(f,nf,paired = TRUE,alternative ="greater")  # paired two-sample t-test
# or
t.test(f-nf,mu=0,alternative ="greater")            # one-sample t-test for differences
\end{verbatim}

\textbf{R console output:}
\begin{verbatim}
Paired t-test
data:  f and nf
t = 0.95554, df = 40, p-value = 0.1725
alternative hypothesis: true difference in means is greater than 0
\end{verbatim}
95 percent confidence interval:
-0.2109974   Inf
sample estimates:
mean of the differences
0.2768293

# or
One Sample t-test
data:  f - nf
t = 0.95554, df = 40, p-value = 0.1725
alternative hypothesis: true mean is greater than 0
95 percent confidence interval:
-0.2109974   Inf
sample estimates:
mean of x
0.2768293

Test conclusion: Since the p-value is greater than the significance level 0.05, accept the null hypothesis. The researcher cannot claim that the formula improves performance.

Problem 6: (5 pts) A random sample is drawn from a distribution and the resulting population mean confidence interval for 90% confidence level is (32.5, 38.1). The null hypothesis for the population mean is \( H_0: \mu = 33.3 \). Should it be accepted at 5% significance level? Justify your answer.

Solution: We have shown in class (slide 2.31) that \( \mu_0 \) belongs to a confidence interval generated by confidence level \( C \) if and only if the (two-tailed test) null hypothesis \( \mu = \mu_0 \) is accepted at significance level \( \alpha = 1 - C \).

Hence the null hypothesis \( H_0: \mu = 33.3 \) is accepted at the significance level \( 1 - 0.9 = 0.1 = 10\% \). That means that the two-tailed test produced a p-value larger than 0.1, so this null hypothesis will certainly be accepted at any significance level smaller than 10%.

Extra Credit: (10 pts) Coin is tossed 11 times and 2 Heads and 9 Tails are recorded. The null hypothesis is that the coin is fair. Do you accept it at 5% significance level?

Hint: Take the number of Heads as your statistic of interest. Can you compute the p value for this statistic? Explain your reasoning.

Solution: As hinted, let \( t \) be the number of Heads in 11 coin tosses. Then \( t \) has a binomial \( B(11, p) \) distribution, with \( p \) the probability of a Head in a single coin toss. Set the null hypothesis \( p = 0.5 \) and calculate the p-value as the probability that a fair coin produces 0, 1, 2, 9, 10, or 11 heads in 11 tosses; these are all the outcomes as likely or less likely than 2 Heads. That probability is calculated as

\[
pbinom(2,11,0.5) + pbinom(8,11,0.5,\text{lower.tail}=\text{FALSE})
\]

and it is approx. 0.0654. Since it is larger than the significance level 0.05, the null hypothesis that the coin is fair is accepted.