1. The research department of a company that produces industrial string proposes a new technique for strengthening the string and making it able to withstand stronger forces. To find out whether the new technique really does strengthen the string, five lengths of string are produced by the standard techniques and five by the new technique and samples are compared for their respective breaking points, in pounds of force. The breaking points are as follows:

<table>
<thead>
<tr>
<th>Standard technique</th>
<th>144</th>
<th>131</th>
<th>155</th>
<th>126</th>
<th>134</th>
</tr>
</thead>
<tbody>
<tr>
<td>New technique</td>
<td>139</td>
<td>154</td>
<td>132</td>
<td>143</td>
<td>147</td>
</tr>
</tbody>
</table>

Does the data support the contention of the research department that the new technique produces stronger string?

(a) Use the two-sample $t$-test.

(b) Use the Wilcoxon rank-sum test.

(State formally the hypothesis to be tested and the assumptions required in both testing procedures)

Solution:

(a) Let $X$ be the breaking point of the string in the standard technique and $Y$ be the breaking point of the string in the new technique. Because the sample size is small, to perform a two-sample $t$-test, we have to assume

$$X \sim N(\mu_X, \sigma^2) \quad \text{and} \quad Y \sim N(\mu_Y, \sigma^2).$$

Note that $\mu_X$ and $\mu_Y$ are the breaking points of the string in the standard technique and new technique respectively.

The hypothesis to be tested is

$$H : \mu_X = \mu_Y \quad \text{vs} \quad H_A : \mu_X < \mu_Y.$$ 

Based on the information contained in Question 1, we have

- $\bar{x} = 138, \quad \bar{y} = 143$;
- $s^2_x = 133.5, \quad s^2_y = 68.5$;
- $s^2 = (4 \times s^2_x + 4 \times s^2_y)/8 = 101, \quad t_{m,n} = (\bar{x} - \bar{y})/\left(s\sqrt{1/5 + 1/5}\right) = -0.786646$. 

1
The \( p \)-value is given by
\[
p = P(t_8 < -0.786646) = 0.2270794.
\]

**Conclusion:** The data is consistent with the null hypothesis, that is, there is no evidence to support the contention of the research department.

(b) The hypothesis to be tested is
\[
H : \mu_X = \mu_Y \quad \text{vs} \quad H_A : \mu_X < \mu_Y,
\]
where \( \mu_X \) and \( \mu_Y \) denotes the average breaking points of the string corresponding to the standard technique and the new technique respectively.

The ranks corresponding to the standard technique and the new technique are given as follows:

<table>
<thead>
<tr>
<th>Standard: 144 131 155 126 134</th>
<th>New: 139 154 132 143 147</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranks: 7 2 10 1 4</td>
<td>5 9 3 6 8</td>
</tr>
</tbody>
</table>

By using the Wilcoxon rank-sum-test, the \( p \)-value for testing the hypothesis is given by
\[
p = P(W \leq w),
\]
where by using the information contained in the sample,
\[
w = 7 + 2 + 10 + 1 + 4 = 24.
\]

With \( n_1 = n_2 = 5 \), it follows from table 8 that
\[
p = P(W \leq 24) = 0.2738.
\]

The data is consistent with the null hypothesis \( H \).

2. Let \( X_1, X_2, ..., X_m \) be iid \( N(\mu_X, \sigma^2) \), \( Y_1, Y_2, ..., Y_n \) be iid \( N(\mu_Y, \sigma^2) \) and the \( X_j \)'s are independent of the \( Y_j \)'s. We want to test the hypothesis:
\[
H : \mu_X = \mu_Y \quad \text{vs} \quad H_A : \mu_X \neq \mu_Y.
\]

(a) Find the critical value at level \( \alpha = 0.05 \) for the two-sample \( t \)-test.

(b) Find the rejection region based on \( \bar{X} - \bar{Y} \) at level \( \alpha = 0.05 \). For the data in Question 1, is the two-sample \( t \)-test significant at level \( \alpha = 0.05 \)?

(c) Find a 95\% confidence interval for the \( \mu_X - \mu_Y \).

**Solution:** It follows from Lect.9 that the test statistic (two sample \( t \)-statistic) is given by
\[
T_{m,n} = \frac{\bar{X} - \bar{Y}}{S \sqrt{\frac{1}{m} + \frac{1}{n}}},
\]
where $\bar{X}, \bar{Y}, S^2_X$ and $S^2_Y$ are defined as before, 

$$S^2 = \frac{(m - 1)S^2_X + (n - 1)S^2_Y}{m + n - 2},$$

and under $H: \mu_X = \mu_Y$, 

$$T_{m,n} \sim t_{n+m-2}.
$$

(a) According to Lect.4, a critical value $c_\alpha$ (for two-sided) at level $\alpha$ is defined by 

$$P_H(|T_{m,n}| \geq c_\alpha) = \alpha.$$ 

Hence the critical value at level $\alpha = 0.05$ for the two-ample $t$-test is 

$$c_\alpha = t_{0.025},$$

where $t_{0.025}$ is given by 

$$P(t_{m+n-2} \geq t_{0.025}) = 0.025.$$

(b) $|T_{m,n}| \geq c_\alpha$ is called the rejection region at level $\alpha$. Hence, the rejection region based on $\bar{X} - \bar{Y}$ at level $\alpha = 0.05$ is given by 

$$|\bar{X} - \bar{Y}| \geq t_{0.025}S\sqrt{\frac{1}{m} + \frac{1}{n}}$$

For the data in Question 1, with $m = n = 5$, 

$$c_\alpha = t_{0.025} = 2.306,$$

and the observed value of the two-sample $t$-test statistic 

$$|t_{m,n}| = 0.786646 < 2.306.$$ 

Hence, for the data in Question 1, the two-sample $t$-test is not significant at level $\alpha = 0.05$.

(c) Note that 

$$\frac{(\bar{X} - \mu_X) - (\bar{Y} - \mu_Y)}{S\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t_{m+n-2}.$$ 

By using a similar method in Lect.6, we can show that a 95% confidence interval for $\mu_X - \mu_Y$ is 

$$\left[(\bar{X} - \bar{Y}) - t_{0.025}S\sqrt{\frac{1}{m} + \frac{1}{n}}, \ (\bar{X} - \bar{Y}) + t_{0.025}S\sqrt{\frac{1}{m} + \frac{1}{n}}\right],$$

where $t_{0.025}$ is given by 

$$P(t_{m+n-2} \geq t_{0.025}) = 0.025.$$
1. The datasets `ratcon` and `ratoz` are weight gains for two groups of rats kept in an ozone free environment and in an environment containing ozone.

   (a) Set up the screen to take a $1 \times 2$ array of plots with `par`.

   (b) Obtain joint boxplots of the two samples and a normal `qq`-plot of the combined residuals from the sample means.

   (c) Comment on your plots in (b) (symmetry, equality of variance, approximate normality, etc) and explain if a t-test is appropriate to test for ozone effects.

   (d) Use `t.test` to obtain a t-test and a 95% confidence interval assuming equal variances and without assuming equal variances. Comment on their similarity.

   (e) Use the Wilcoxon rank-sum test to test for ozone effects.

   Hint: Consult Lect.11 for the idea of the test and Example 11.2 for calculation of the `p`-value in Splus (Note that the alternative hypothesis in this exercise may be two-sided).

   (f) Use `wilcox.test` to confirm your findings in (e).

   (g) Comment on the `p`-values obtained by using `t.test` and `wilcox.test`.

**Solution:**

```r
# (a)
par(mfrow=c(1,2))

#(b)
> boxplot(ratcon, ratoz)
> qqnorm(c(ratcon, ratoz)-mean(c(ratcon, ratoz)))
> qqline(c(ratcon, ratoz)-mean(c(ratcon, ratoz)))

#(c)

# The boxplots indicate that both samples are not symmetric,
# the `{tt ratcon}` is slightly long tailed the variances
# of two samples might not be same, and the qq-plot of the
# combined residuals indicates the sample might not be from
# a normal distribution. However the sample sizes $m$ (23) and
# $n$ (22) are considered to be large. We may use the two sample
# t-test to test for ozone effects. However we will recognize
# that some errors in inference will occur due to these deviations
# from the assumptions.

#(d)
```

4
> t.test<-t.test(ratcon, ratoz, alternative="two.sided", 
  mu=0, var.equal=T, conf.level=0.95)

> t.test$p.value
[1] 0.01664109

> t.test$conf
[1] 2.177186 20.656806
attr(, "conf.level"): [1] 0.95

> t.test1<-t.test(ratcon, ratoz, alternative="two.sided", 
  mu=0, var.equal=F, conf.level=0.95)

> t.test1$p.value
[1] 0.01918275

> t.test1$conf
[1] 1.985043 20.848949
attr(, "conf.level"): [1] 0.95

# Two confidence intervals are nearly same. The difference is 
# because we have used different test statistics.

# (e)
rank<-rank(c(ratcon, ratoz))
m<-length(ratcon)
rankA<-rank[1:m]
w<-sum(rankA)

n<-length(ratoz)
EW<-m*(m+n+1)/2
sumranksq<-sum(rank^2)
g<-(m+n)*((m+n+1)^2)/4
> varW<-m*n*(sumranksq-g)/((m+n)*(m+n-1))
> p<-2*(1- pnorm(abs(w-EW)/sqrt(varW)))
> p
[1] 0.002723284

# (f)

> wilcox.test(ratcon, ratoz, alternative="two.sided",
       mu=0, exact=F, correct=F)

Wilcoxon rank-sum test

data:  ratcon and ratoz
rank-sum normal statistic without correction Z = 2.9974,
p-value = 0.0027
alternative hypothesis: true mu is not 0

# (g)

The $p$-values obtained by using \texttt{t.test} is a little bit larger than the one obtained by using \texttt{wilcox.test}. Both tests indicate that there is strong evidence that weight gains for two groups of rats were affected by ozone.