Lab 7 - Comparing two means and violation of assumptions

Contents

Lab Objectives:

- Compare the means of two groups
- Understand paired vs. two-sample designs
- Explore data transformation
- Detect violations of assumptions and use R for nonparametric alternatives to one-sample, paired and two-sample t-tests

---

Exercise 1: Paired t-test

When comparing two means, if each data point in one group is uniquely paired with a data point in the other group, then the data are said to be paired. A paired t-test examines the mean difference between the members of each pair. Paired t-tests assume that the differences are normally distributed and that the pairs are randomly sampled.

A paired sampling design is usually more powerful than an unpaired design, as it reduces extraneous variation between sampling units. However, this sampling method is not always possible. When data are not paired, we compare the means of two groups using a two-sample t-test. The two-sample t- test assumes that there

1

is an independent random sample from each group, and that the variable is normally distributed in each population with the same variance.

Welch's t-test is similar to a two-sample t-test, but does not require the assumption of equal variances. The data output from a two-sample t-test in R allows you to specify if to assume equal variances, or not.

Biological studies often examine the same patients before and after a particular treatment. These types of studies can be analyzed using a paired t-test. Let's begin with an example of weight gain in anorexia patients as a function of family therapy. We want to see if the difference in weight before and after a family therapy intervention is significantly different from zero.

Anorexia therapy

Open the file anorexia family therapy.csv:

```
d <- read.csv("anorexia family therapy.csv")
str(d)
```

```
## 'data.frame':    17 obs. of  3 variables:
##  $ BEFORE: num  83.8 83.3 86 82.5 86.7 79.6 76.9 94.2 73.4 80.5 ...
##  $ AFTER : num  95.2 94.3 91.5 91.9 100.3 ...
##  $ GAIN  : num  11.4 11 5.5 9.4 13.6 -2.9 -0.1 7.4 21.5 -5.3 ...
```

Examine the correlation between the BEFORE and AFTER data using the **cor.test()** function.

Use the **t.test()** function, and select the paired t-test option by setting paired=TRUE, to examine the difference between the BEFORE and AFTER:

```
cor.test(x = d$BEFORE, y = d$AFTER)
```

```
## 
##  Pearson's product-moment correlation
## 
## data:  d$BEFORE and d$AFTER
## t = 2.4732, df = 15, p-value = 0.02584
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.07764331 0.80945446
## sample estimates:
##      cor
## 0.538203
```

```
t.test(x = d$BEFORE, y = d$AFTER, paired = TRUE, var.equal = TRUE)
```

```
## 
##  Paired t-test
## 
## data:  d$BEFORE and d$AFTER
## t = -4.1849, df = 16, p-value = 0.0007003
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -10.94471  -3.58470
## sample estimates:
## mean of the differences
##               -7.264706
```

The first output displays the correlation statistic. We haven't covered correlation in class yet; however it is fairly easy to interpret. The correlation coefficient is a measure of the strength and direction of a linear relationship between the two variables. The correlation coefficient can range from -1 to +1, with -1 indicating a perfect negative correlation, and +1 indicating a perfect positive correlation, and 0 indicating no correlation at all. This data shows a correlation of 0.538, indicating that the two variables are positively correlated, but not perfectly. 0.538 is a numerical description of how tightly around an imaginary linear line the points lie. Even though it is not a perfect correlation, you can see that the correlation is significant, with a P-value of 0.026 (rounded to 0.03).

Examine the t.test output. This output contains summary statistics, including mean, df, significance (P-value) for a two-tailed test, and confidence intervals for the difference in weight before and after therapy. The mean difference is negative because it was calculated by subtracting the weight after therapy from the weight before. The 95% confidence interval of the difference is indicated, and shows that we are confident that there is a weight gain with therapy, as both the lower and upper limits are negative values.

Remember from class that a paired t-test is identical to a one-sample t-test on the differences. Let's check this using the data in the gain column, using a mu=0, as you are testing the null hypothesis that weight does not change with family therapy:

```
t.test(d$GAIN, mu = 0)
```

```
##
##  One Sample t-test
##
## data:  d$GAIN
## t = 4.1849, df = 16, p-value = 0.0007003
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##    3.58470 10.94471
## sample estimates:
## mean of x
##   7.264706
```

Verify that the two approaches lead to identical conclusions and statistical parameters.

Answer question 1 on Sakai

Exercise 2: Two-sample t-test

Often in Biology, you will want to compare means of two independent samples. In this case each measurement is not directly paired with another measurement. Instead we have several measurements in each group, and we want to compare the mean of one group to the other.

Spinocerebellar ataxia

The data in the file ataxia.sav are from an experiment on a mouse model of the neurodegenerative disease Spinocerebellar ataxia type 1. Researchers examined whether life span differed between mice with no exercise, compared with mice with a mild exercise regime.

Open the file ataxia.csv, peform a normal **t.test()**, but this time we will make use of the formula (~) to indicate the groups you want to compare, and *data* to indicate the data.frame where the data are located:

```
d <- read.csv("ataxia.csv")
t.test(lifespan ~ treatment, data = d)
```

```
##
##  Welch Two Sample t-test
##
## data:  lifespan by treatment
## t = 3.1531, df = 7.4065, p-value = 0.01494
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  11.02150 74.31183
## sample estimates:
##    mean in group Exercise mean in group No Exercise
##                  310.0000                  267.3333
```

In this case the test variable is lifespan and the grouping variable is treatment. This ensures we will be comparing the mean lifespan (test variable) between the two treatment groups.

In this file the treatment groups are assigned numeric values of 0 or 1, and the values are then labelled no exercise or exercise. R treats categorical data quite specifically, but behind the scenes, it converts categories to a form of numeric value for coding. Going forward, if you are unsure about what numeric values are assigned to your groups, consult the **str()** or the **levels()** functions:
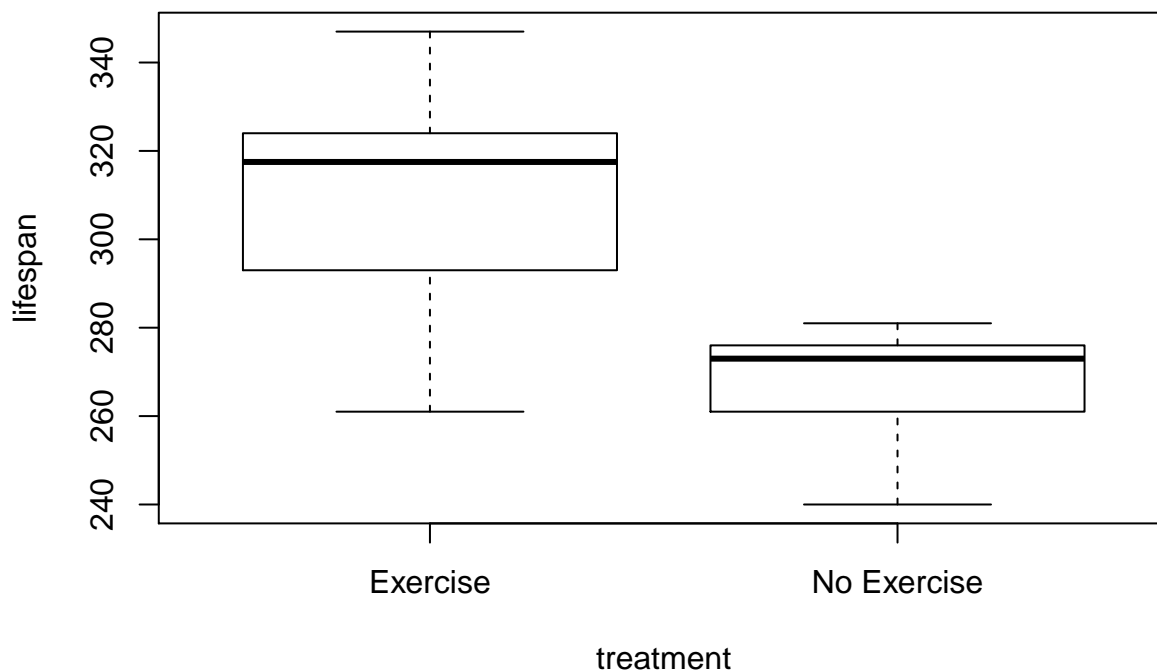
```
str(d)
```

```
## 'data.frame':    12 obs. of  2 variables:
##  $ treatment: Factor w/ 2 levels "Exercise","No Exercise": 2 2 2 2 2 2 1 1 1 1 ...
##  $ lifespan : int  240 261 271 275 276 281 261 293 316 319 ...
```

```
levels(d$treatment)
```

```
## [1] "Exercise"    "No Exercise"
```

```
plot(lifespan ~ treatment, data = d)
```


Notice the plot function also takes categorical data as a grouping variable, creating a box and whisker

plot from your data. From this graph, you might suspect that the variances of the two groups are not the same. Unfortunately, R expects you to know this, so you will need to install the package , using the **install.packages("car")** command, in order to perform a Levene Test for variance homogeneity:

```r
# install.packages('car')
library(car)
```

```
## Loading required package: carData
```

```r
leveneTest(d$lifespan, d$treatment)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  1  0.9922 0.3427
##       10
```

Remember that an important assumption of a two-sample t-test is that the two groups have equal variances. Levene's test checks this. If the results of Levene's test are significant (P<0.05), this indicates that the two groups DO NOT have equal variances. If that were the case, we would need to adjust the t.test inputs to reflect this. Compare the results of the t.test with different variance assumptions:

```r
t.equal <- t.test(lifespan ~ treatment, data = d, var.equal = TRUE)
t.equal
```

```
##
##  Two Sample t-test
##
## data:  lifespan by treatment
## t = 3.1531, df = 10, p-value = 0.01028
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  12.51581 72.81752
## sample estimates:
##    mean in group Exercise mean in group No Exercise
##                  310.0000                  267.3333
```

```r
t.unequal <- t.test(lifespan ~ treatment, data = d, var.equal = FALSE)
t.unequal
```

```
##
##  Welch Two Sample t-test
##
## data:  lifespan by treatment
## t = 3.1531, df = 7.4065, p-value = 0.01494
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  11.02150 74.31183
## sample estimates:
##    mean in group Exercise mean in group No Exercise
##                  310.0000                  267.3333
```

```r
t.equal$p.value
```

```
## [1] 0.01027946
```

```r
t.unequal$p.value
```

```
## [1] 0.01494339
```

Welch's approximate t-test does not assume equal variances, and you should see different degrees of freedom and confidence intervals.

If the results of Levene's test are not significant (P>0.05), then we can assume that the variances are equal and can use the normal Two Sample t-test. The P-value can be found in the third row of each output, or derived from a result variable as shown above.

Answer question 2 on Sakai

Exercise 3: Choose the right t-test

Homophobia

In a paper published in 1996, Adams et al examined whether homophobia is associated with homosexual arousal. Men were assigned to either the homophobic or nonhomophobic group based on their scores on a Homophobia Index test. The men then watched sexually explicit erotic homosexual videotapes, and an arousal score was determined. These data are found in the file homophobia.csv. Conduct a statistical test to determine if there is any difference in arousal between the two groups.

```
d <- read.csv("Homophobia.csv")
str(d)
```

```
## 'data.frame':    64 obs. of  2 variables:
##  $ Group  : Factor w/ 2 levels "Homophobic","Nonhomophobic": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Arousal: num  39.1 38 14.9 20.7 19.5 32.2 11 20.7 26.4 35.7 ...
```

Answer question 3 on Sakai.

NZ Noggins

The head sizes of 18 New Zealand army recruits were measured with both cardboard and metal calipers. Researchers want to determine whether there is a significant difference between head size when measured with these two methods. The data is found in NZnoggins.csv. Open NZnoggins.csv and conduct an appropriate test to determine if there is a difference in head size when measured with cardboard and metal calipers.

```
d <- read.csv("NZnoggins.csv")
str(d)
```

```
## 'data.frame':    18 obs. of  3 variables:
##  $ Individual       : Factor w/ 18 levels "A  ","B  ","C  ",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ CardboardCalipers: int  146 151 163 152 151 151 149 166 149 155 ...
##  $ MetalCalipers    : int  145 153 161 151 145 150 150 163 147 154 ...
```

Answer question 4 on Sakai

Beer & mosquitos

A study in West Africa on mosquitos that cause malaria examined whether drinking beer influenced attractiveness to mosquitos. Mosquitos were released next to 25 alcohol-free participants, and the proportion of mosquitos flying toward the participants was measured. The procedure was repeated 15 minutes after each of the same participants had consumed a liter of beer. Open mosquitos beer.csv and conduct an appropriate test to determine if there is a change in attractiveness to mosquitos following drinking beer.

```
d <- read.csv("mosquitos beer.csv")
str(d)
```

```
## 'data.frame':    25 obs. of  3 variables:
##  $ drink     : Factor w/ 1 level "beer ": 1 1 1 1 1 1 1 1 1 1 ...
##  $ beforeDrink: num  0.13 0.13 0.21 0.25 0.25 0.32 0.43 0.44 0.46 0.5 ...
##  $ afterDrink : num  0.49 0.59 0.27 0.43 0.5 0.5 0.37 0.3 0.58 0.89 ...
```

Answer question 5 on Sakai

More beer & mosquitos

The West African study was expanded by including a second control group of 18 participants who were given water instead of beer. The change in activation of mosquitos was measured in both the group of 25 participants who drank beer, and the 18 participants who drank water. Open mosquitos beer water.csv and conduct a statistical test to examine the difference in the change in mosquito activation between the beer group and the water group.

```
d <- read.csv("mosquitos beer water.csv")
str(d)
```

```
## 'data.frame':    43 obs. of  4 variables:
##  $ drink     : Factor w/ 2 levels "beer","water": 1 1 1 1 1 1 1 1 1 1 ...
##  $ beforeDrink: num  0.13 0.13 0.21 0.25 0.25 0.32 0.43 0.44 0.46 0.5 ...
##  $ afterDrink : num  0.49 0.59 0.27 0.43 0.5 0.5 0.37 0.3 0.58 0.89 ...
##  $ change    : num  0.36 0.46 0.06 0.18 0.25 0.18 -0.06 -0.14 0.12 0.39 ...
```

```
plot(d$change ~ d$drink)
```
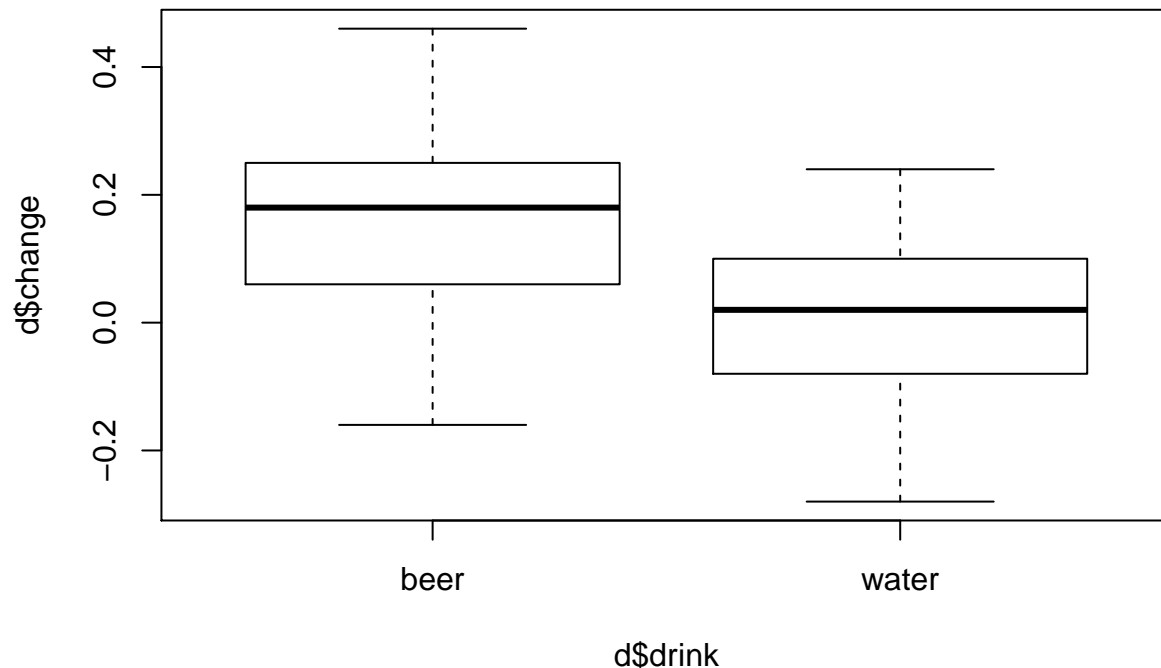
```
t.test(d$change ~ d$drink, data = d)
```

```
##
##   Welch Two Sample t-test
##
## data:  d$change by d$drink
## t = 3.3219, df = 40.663, p-value = 0.001897
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.05746134 0.23578311
## sample estimates:
##  mean in group beer mean in group water
##          0.154400000          0.007777778
```

Answer question 6 on Sakai

Exercise 4: Violations of assumptions and nonparametric tests

Many parametric tests, including t-tests, are robust to their assumptions. This means that the test still works well even when the assumptions are violated. For example, with large sample sizes, the two-sample t-test works acceptably even if the two groups' standard deviations differ by threefold. However, using a parametric test when the assumptions are not met increases the *type I* error rate (probability of rejecting a true null hypothesis). If the assumptions are violated too severely to ignore, data can sometimes be transformed to make new variables that better match the assumptions of the methods. For example, with the log-transformation, the natural logarithm (Ln) of the value of a variable is computed, and is used in parametric tests instead of the variable itself.

If the assumptions are violated and data transformation doesn't produce data that better meet the assumptions, then nonparametric tests can be used. These include the Mann-Whitney U test (or Wilcoxon rank sum test),

which examines the distributions of two independent groups, and the Sign test, which can be used in place of both a one-sample t-test and a paired samples t-test. Nonparametric tests are not as powerful as parametric tests. That is, they are less likely to reject a false null hypothesis (higher type II error). It is therefore better to use parametric tests if the assumptions can be met, and turn to nonparametric methods only after data transformation has failed to meet the assumptions of the parametric methods.

A *Shapiro-Wilk* test is a formal method for testing for departures from normality. A P-value of less than 0.05 indicates that the data are not a good fit for the normal distribution. Formal normality tests should be used with caution, as small samples may not have enough power to reject normality, and large samples may reject normality even when the departure is not severe to give up on methods that assume normality.

Biology & sports

A group of French researchers was interested in whether people who play a lot of sports have more sexual partners than those who do not. They asked two groups of students how many sexual partners they had in the previous year (Faurie et al 2004). One group was composed of 103 physical education majors who regularly participated in sports, and the other group was composed of 100 biology majors who did not regularly participate in sports. The data is found in the file Biology sports.csv.

```r
d <- read.csv("biology sports.csv")
str(d)
```

```
## 'data.frame':    203 obs. of  2 variables:
##  $ major           : Factor w/ 2 levels "biology","sports": 1 1 1 1 1 1 1 1 1 1 ...
##  $ numberSexPartners: int  0 0 0 0 0 0 0 0 0 0 ...
```

To examine the data more closely, we'll create two new data.frames using the **subset()** function to pull out only those records that correspond to each major, creating a biology and a sports data.frame. This makes plotting histograms of the data easier to do in a basic R tutorial:
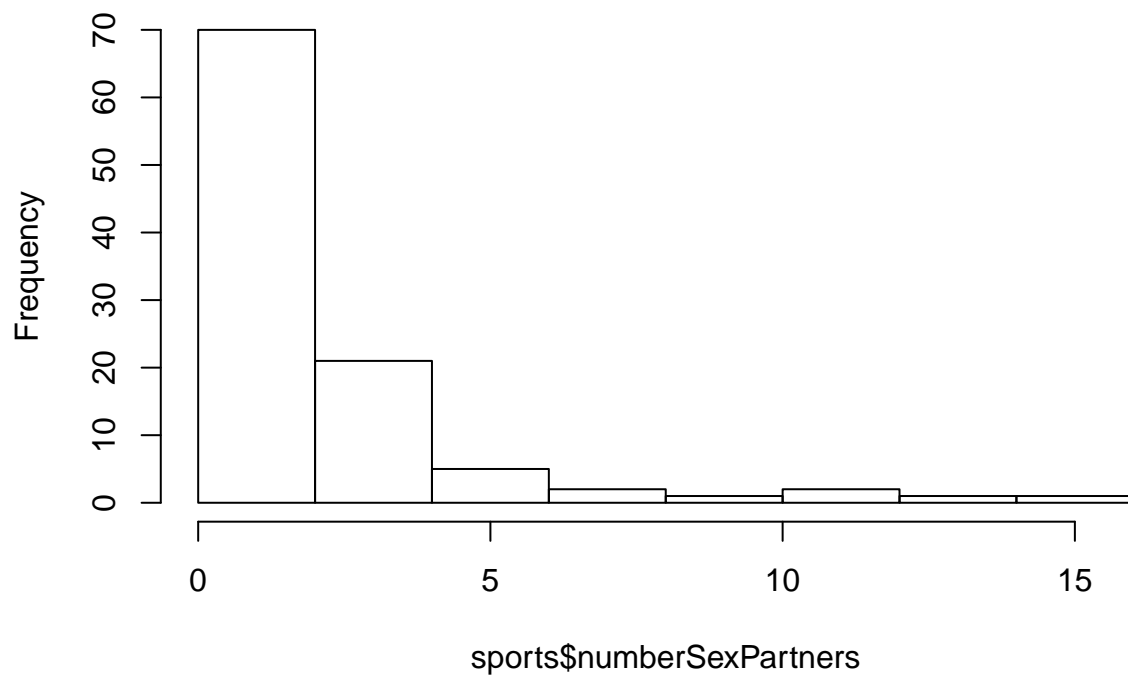
```r
library(car)  # needed for the qqp plots
biology <- subset(d, major == "biology")
sports <- subset(d, major == "sports")
hist(biology$numberSexPartners)
```
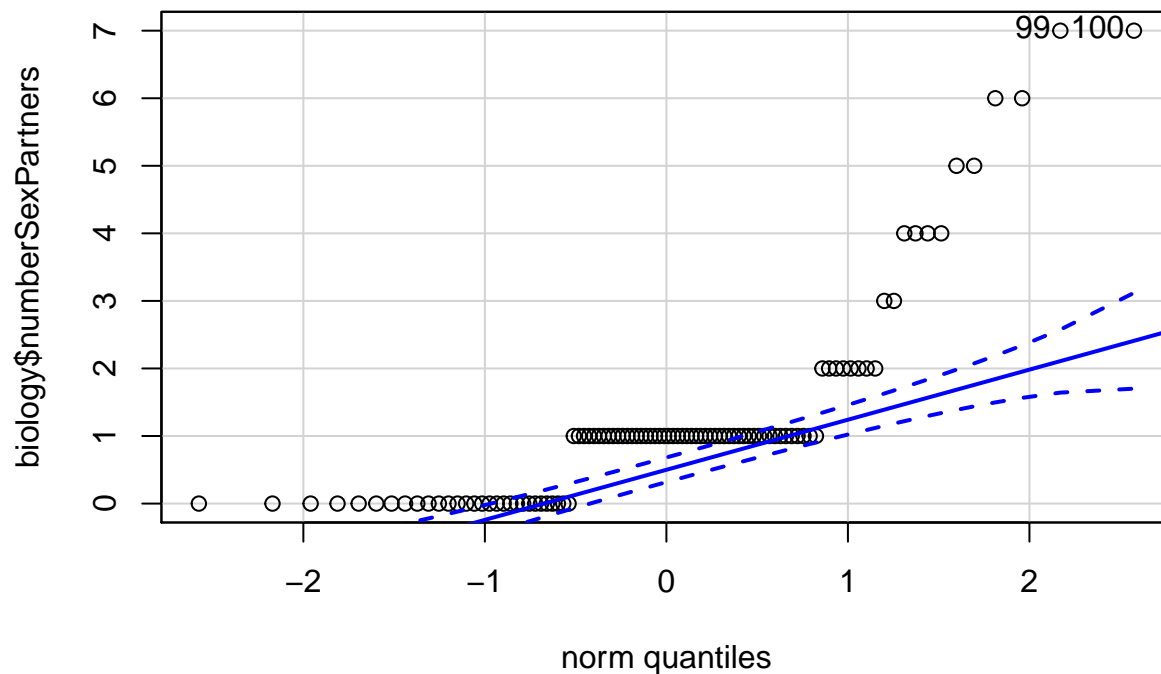
Histogram of biology$numberSexPartners

biology$numberSexPartners
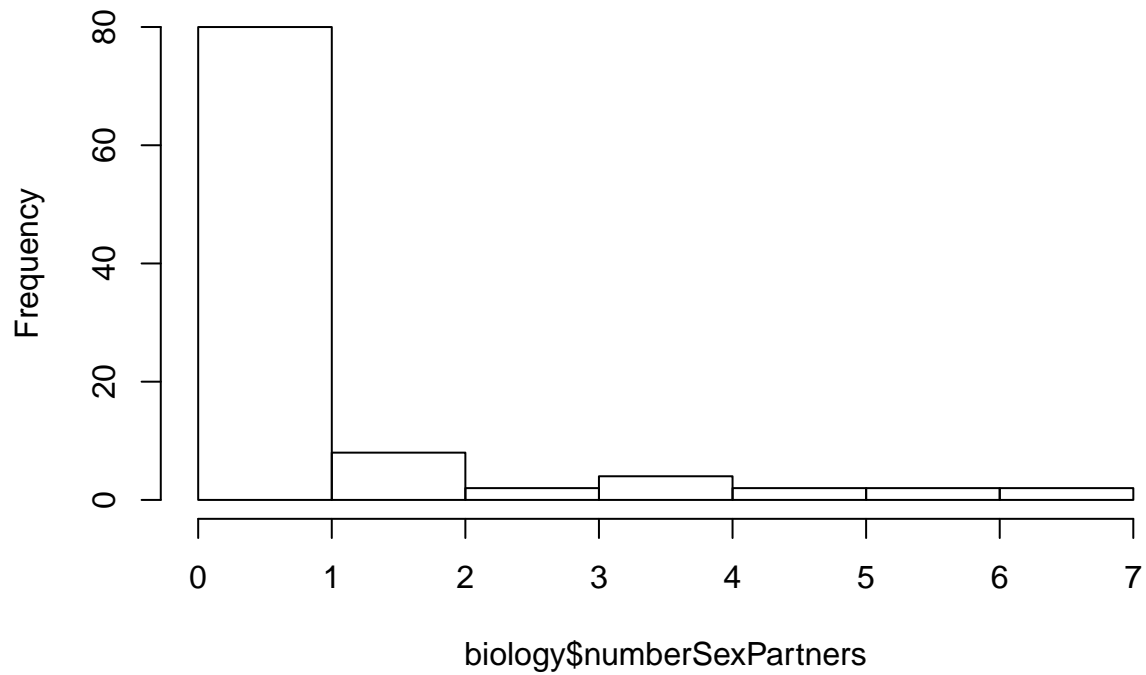
```
hist(sports$numberSexPartners)
```

Histogram of sports$numberSexPartners
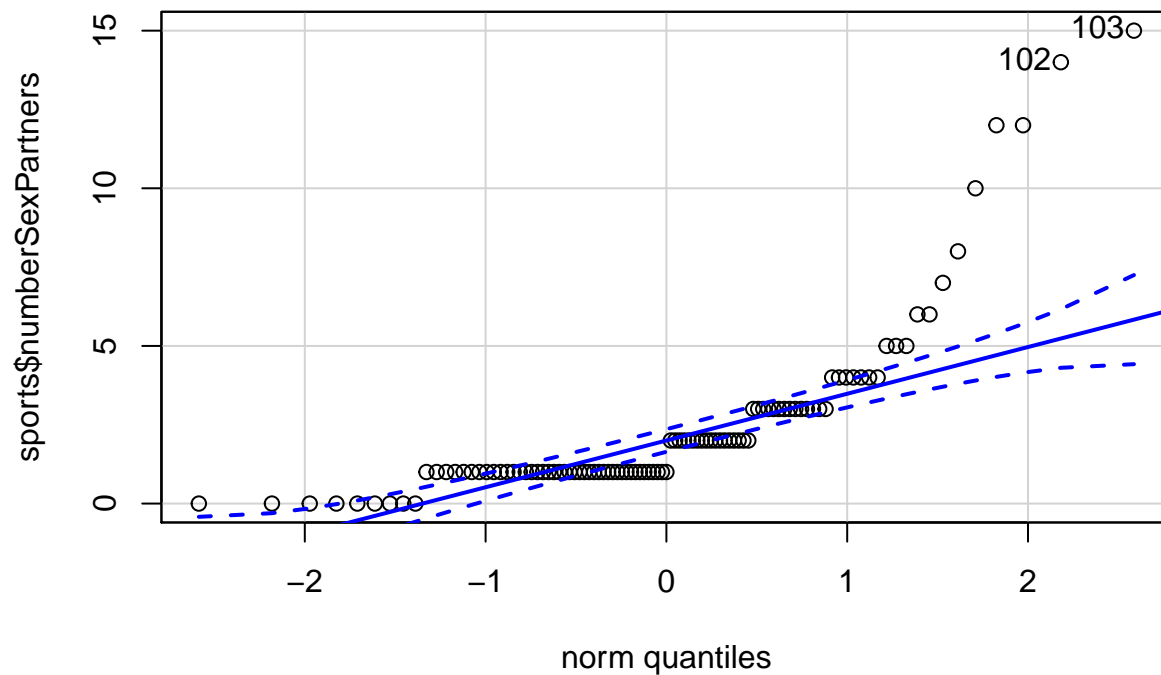

sports$numberSexPartners

```
qqp(biology$numberSexPartners)
```

```
## [1]  99 100
```

```
qqp(sports$numberSexPartners)
```



```
## [1] 103 102
```

Examine the two histograms of the data. You should be able to see that the data are not normally distributed. The important consideration for statistical tests is whether the variable is normally distributed in the population. However, random samples that are very skewed are often reflective of a non-normal population distribution.

Next, take a look at the two Normal Q-Q plots. These plots are another method for detecting deviations

from normality. If the data are normally distributed, the data points should fall roughly along a straight line. The plots include a straight line as a guide.

We are also interested in the results of the Shapiro-Wilk test. If the Sig. value of this test is less than 0.05, then the data are not a good fit for the normal distribution.

```
shapiro.test(biology$numberSexPartners)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  biology$numberSexPartners
## W = 0.68733, p-value = 2.76e-13
```

```
shapiro.test(sports$numberSexPartners)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sports$numberSexPartners
## W = 0.67166, p-value = 7.507e-14
```

Now that we know the data are not normal, let's try a data transformation

Since the data include zeros, we will try log transforming the variable + 1. Simply add a +1 in the numeric expression as follows:

```
biology$lnnumber <- log(biology$numberSexPartners + 1)
sports$lnnumber <- log(sports$numberSexPartners + 1)
```

Another common transformation for data that are skewed right is a reciprocal transformation.

```
biology$recipnumber <- 1/(biology$numberSexPartners + 1)
sports$recipnumber <- 1/(sports$numberSexPartners + 1)
```

Finally, let's try the square root transformation:

```
biology$sqrtnumber <- 1/(biology$numberSexPartners + 0.5)
sports$sqrtnumber <- 1/(sports$numberSexPartners + 0.5)
```

Now let's check to see if any of these transformations improved the normality of our data.

Examine these results to determine if any of the transformations improved the fit of the data to a normal distribution.
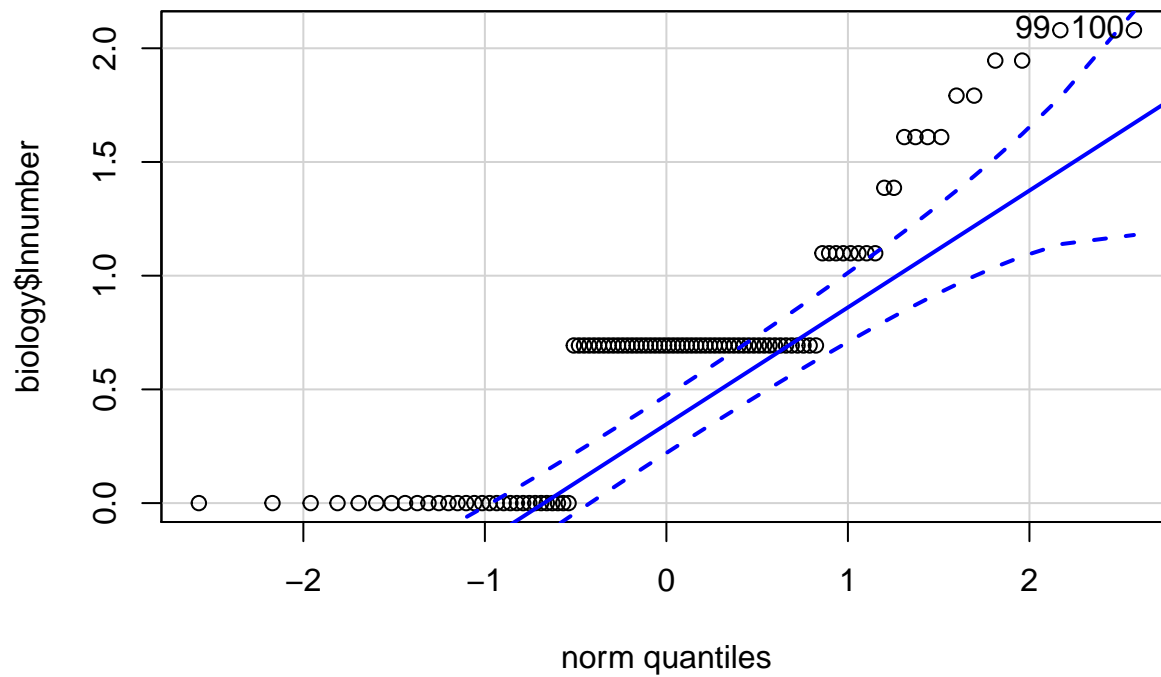
```
shapiro.test(biology$lnnumber)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  biology$lnnumber
## W = 0.83192, p-value = 2.691e-09
```
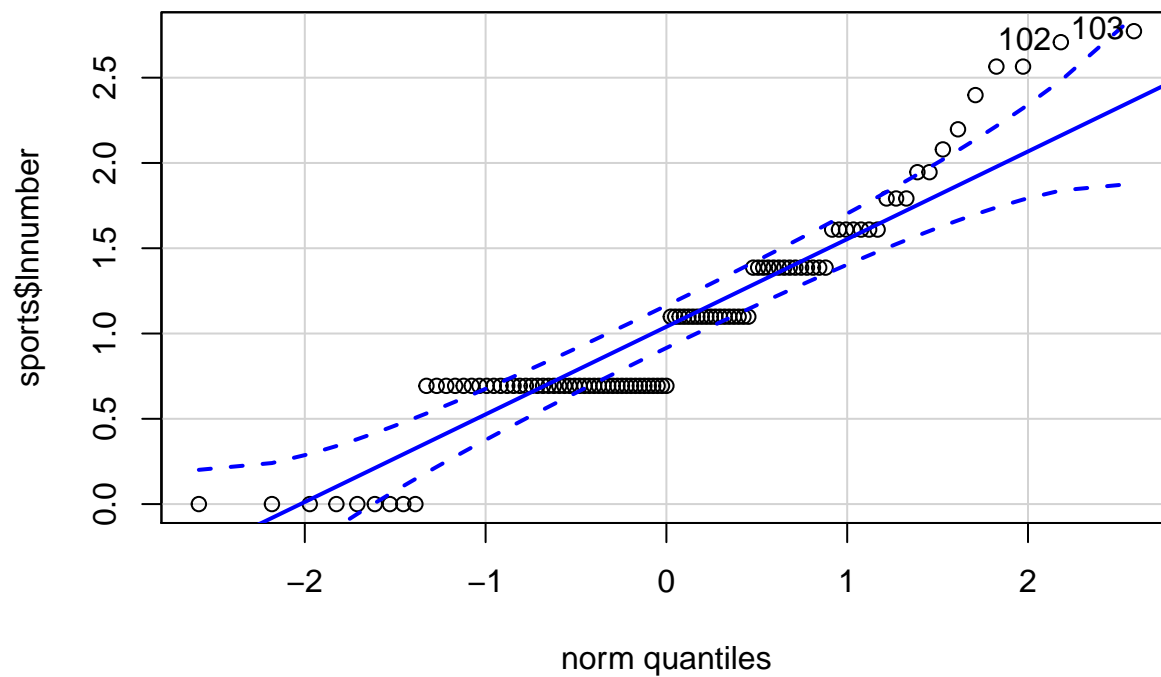
```
shapiro.test(sports$lnnumber)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sports$lnnumber
## W = 0.90135, p-value = 1.207e-06
```

```r
qqp(biology$lnnumber)
```



```
## [1]  99 100
```

```r
qqp(sports$lnnumber)
```



```
## [1] 103 102
```
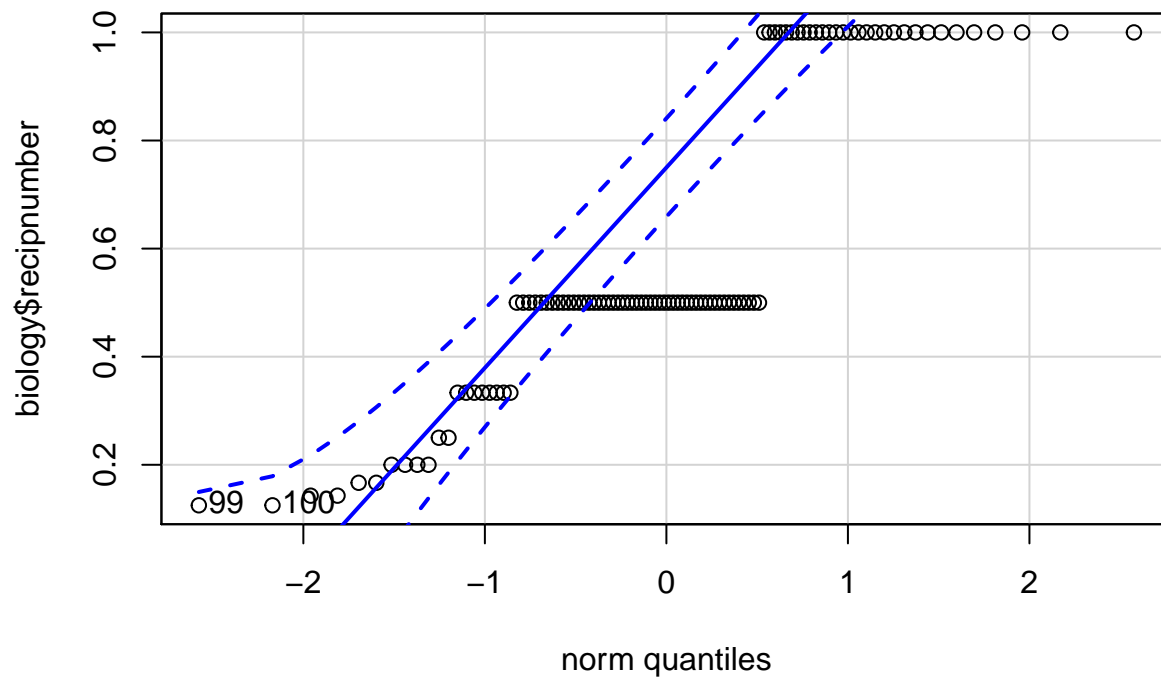
```r
shapiro.test(biology$recipnumber)
```

```
## 
##  Shapiro-Wilk normality test
```

```
##
## data:  biology$recipnumber
## W = 0.79914, p-value = 2.33e-10
```
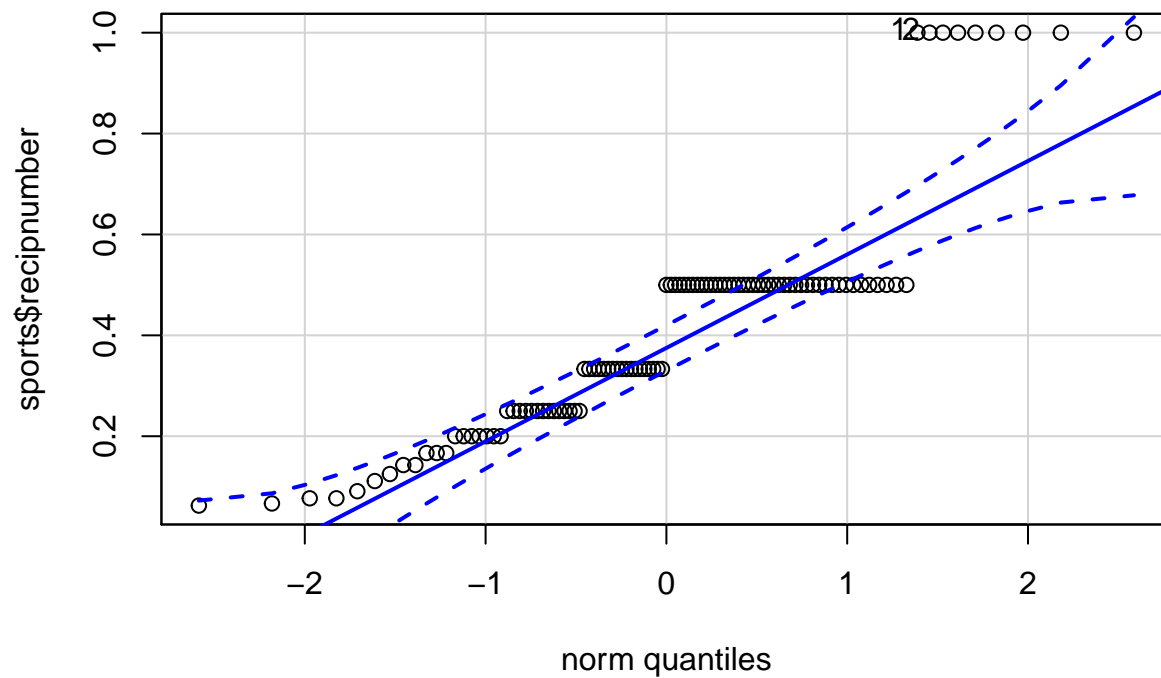
```
shapiro.test(sports$recipnumber)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sports$recipnumber
## W = 0.83343, p-value = 2.062e-09
```

```
qqp(biology$recipnumber)
```



```
## [1]  99 100
```

```
qqp(sports$recipnumber)
```
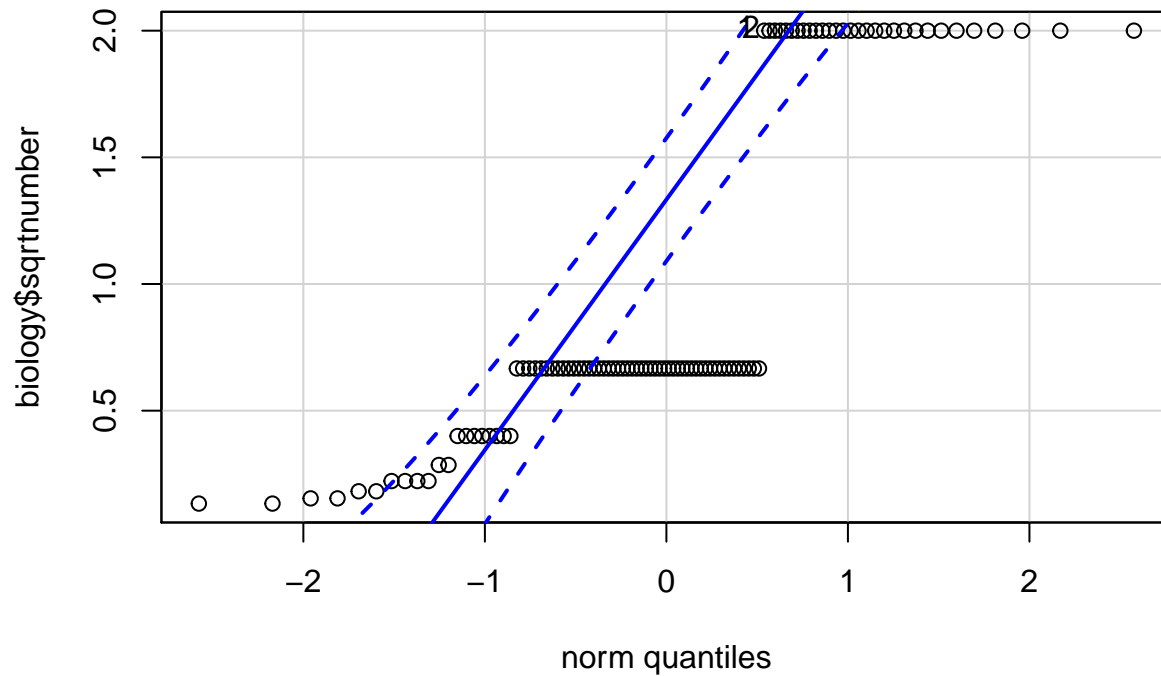
```
## [1] 1 2
```

```
shapiro.test(biology$sqrtnumber)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  biology$sqrtnumber
## W = 0.7301, p-value = 2.879e-12
```
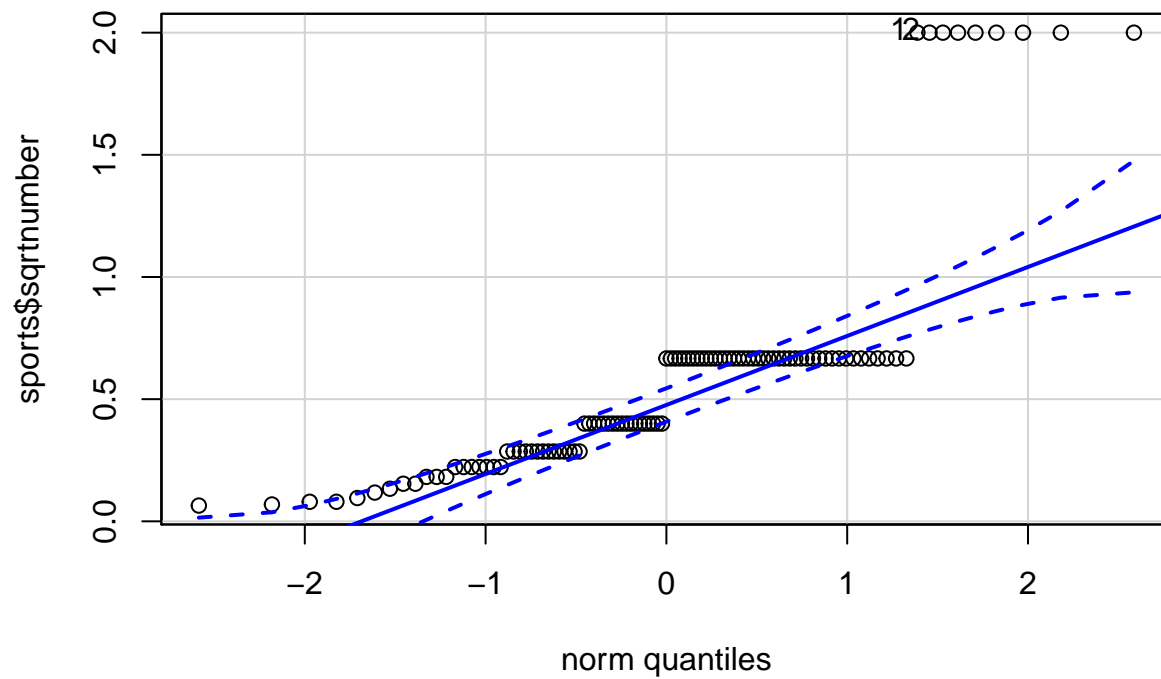
```
shapiro.test(sports$sqrtnumber)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sports$sqrtnumber
## W = 0.68653, p-value = 1.62e-13
```

```
qqp(biology$sqrtnumber)
```

```
## [1] 1 2
```

```
qqp(sports$sqrtnumber)
```



```
## [1] 1 2
```

When a deviation from normality is too strong to ignore, and transformation does not meet the assumptions of a two-sample t-test (normal distribution and equal standard deviation), the next option is to try a non-parametric test.

The Wilcoxon rank-sum test (available in R, as **wilcox.test()**) is equivalent to the Mann-Whitney U-test and used to compare samples that do not follow normal distributions.

Note from the **wilcox.test()** function's help: The literature is not unanimous about the definitions of the Wilcoxon rank sum and Mann-Whitney tests. The two most common definitions correspond to the sum of the ranks of the first sample with the minimum value subtracted or not: R subtracts and S-PLUS does not, giving a value which is larger by m(m+1)/2 for a first sample of size m. (It seems Wilcoxon's original paper used the unadjusted sum of the ranks but subsequent tables subtracted the minimum.)

See here for an explanation of how the tests are basically the same: https://stats.stackexchange.com/questions/79843/is-the-w-statistic-output-by-wilcox-test-in-r-the-same-as-the-u-statistic.

The output of the test statistic in R gives you the statistic called Wilcoxon W and the P-value (Asymp. Sig. 2-tailed).

If the significance value is less than 0.05, then we can conclude that the distribution of the variable (number of sex partners) is not the same between the two majors. If both distributions have the same shape (in this case they do, both are right-skewed) then we can also conclude that the two groups do not have the same median.

```
wilcox.test(biology$lnnumber, sports$lnnumber)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  biology$lnnumber and sports$lnnumber
## W = 3220, p-value = 1.078e-06
## alternative hypothesis: true location shift is not equal to 0
```

Answer question 7 on Sakai

Health expenditure

Health expenditure per person from a random sample of 20 countries is found in the file health expenditure.csv. You suspect that this data is not normally distributed. Use R to conduct a natural log transformation of this data, and conduct a Shapiro-Wilk test on both the untransformed and transformed data to determine if this improves the distribution.

```
d <- read.csv("health expenditure.csv")
str(d)
```

```
## 'data.frame':    20 obs. of  2 variables:
##  $ country    : Factor w/ 20 levels "Bahrain                              ",..: 1 2 3 4
##  $ expenditure: int  864 320 239 882 472 72 60 607 3984 4668 ...
```

Answer question 8 on Sakai

Killer lions

When an intruding male lions take over a pride of females, they often kill most or all of the infants in the group. Following takeover, the stability of the pride is unpredictable, and scientists hypothesized that females may delay ovulation until the uncertainty has passed. To examine this, a study of lions in the Serengeti measured the time to reproduction of female lions after losing cubs to infanticide, compared to the time to reproduction after losing cubs to accidental death. The data are not normally distributed, even following

transformation. Conduct an appropriate statistical test on the data in the file Lions.csv to determine if there is a significant difference in reproduction time.

```
d <- read.csv("Lions.csv")
str(d)
```

```
## 'data.frame':    14 obs. of  2 variables:
##  $ death      : Factor w/ 2 levels "Accidental","Infanticide": 1 1 1 1 1 1 1 1 1 2 ...
##  $ daysToNextCub: int  110 117 133 135 140 168 171 238 255 211 ...
```

Answer question 9 on Sakai

Parasitic ducks

Female goldeneye ducks lay eggs in other females' nests, in addition to the eggs they produce and raise in their own nests. Females lay just one egg per day, so the total time span for all eggs in one nest to hatch can be affected by whether the eggs were laid early or late. To determine whether a female tends to lay her earliest eggs in her own nest, Andersson and Ahlund (2012) measured eggs produced by 14 female goldeneyes and assigned each a parasitism index. The index is negative if she tends to lay her last eggs in another female's nest, and positive if she tends to lay her last eggs in her own nest. An index of zero indicates that she alternates laying eggs in her nest and the other female's nest. This data is not normally distributed, and transformation does not make it normal.

To determine whether the parasitism index is significantly different from zero, we could conduct a one-sample sign test. R has a solution, but you'll have to install a new package .

For the one-sample sign-test, the null hypothesis is that the median (md) of the population from which x is drawn is md (we can set to md=0).

```
d <- read.csv("Parasitic ducks.csv")
str(d)
```

```
## 'data.frame':    14 obs. of  2 variables:
##  $ femaleName         : Factor w/ 14 levels "female 1              ",..: 1 5 2 3 4 6 7 8 9 10
##  $ parasitismFirstIndex: num  -2.3 8 10.5 4.6 5.5 6 0 5 4.5 6.5 ...
# the number of scores > 0:
table(d$parasitismFirstIndex > 0)
```

```
##
## FALSE   TRUE
##     2     12
```

```
# install.packages('BSDA')
library(BSDA)
```

```
## Loading required package: lattice
```

```
##
## Attaching package: 'BSDA'
```

```
## The following objects are masked from 'package:carData':
##
##     Vocab, Wool
```

```
## The following object is masked from 'package:datasets':
##
```

```
##      Orange
```

```
SIGN.test(d$parasitismFirstIndex, md = 0)
```

```
##
##  One-sample Sign-Test
##
## data:  d$parasitismFirstIndex
## s = 12, p-value = 0.003418
## alternative hypothesis: true median is not equal to 0
## 95 percent confidence interval:
##  4.383846 6.082967
## sample estimates:
## median of x
##          5
##
## Achieved and Interpolated Confidence Intervals:
##
##                  Conf.Level L.E.pt U.E.pt
## Lower Achieved CI    0.9426 4.5000  6.000
## Interpolated CI      0.9500 4.3838  6.083
## Upper Achieved CI    0.9871 3.8000  6.500
```

The first line performs a simply boolean test of how many of the index scores are > 0 and we return this as a table of FALSE and TRUE values.

In this case, 12 ducks had an index greater than zero. Is this enough to reject the null hypothesis? Examine the test-statistic output, which provides the P-value. If this is less than 0.05, we reject the null hypothesis and conclude that the parasitism index is significantly different from zero.

Answer question 10 on Sakai

Incompatible mates

There are two genetically distinct types of Gouldian finches, one having a red face and the other having a black face. Previous experiments have shown that females have a strong preference for mating with males with the same colour face as themselves, and that the offspring are more likely to survive when both parents are the same type. To examine stress caused by having an incompatible mate, researchers measured corticosterone concentration in 43 females following breeding with a compatible male (same colour face), and in the same 43 females following breeding with an incompatible male (different colour face). The data are not normally distributed. We will therefore conduct a paired-samples sign test.

19

```
d <- read.csv("incompatible mates.csv")
str(d)
```

```
## 'data.frame':    43 obs. of  2 variables:
##  $ Compatible  : int  10 10 39 30 6 21 0 4 16 10 ...
##  $ Incompatible: int  130 105 91 82 77 65 64 60 56 51 ...
```

To run the sign test in R and incorporate the paired nature of the experimental design, you will have to set paired=TRUE:

```
library(BSDA)
SIGN.test(x = d$Compatible, y = d$Incompatible, paired = TRUE)
```

```
##
##  Dependent-samples Sign-Test
##
## data:  d$Compatible and d$Incompatible
## S = 1, p-value = 1e-11
## alternative hypothesis: true median difference is not equal to 0
## 95 percent confidence interval:
##  -30.46421 -22.00000
## sample estimates:
## median of x-y
##           -26
##
## Achieved and Interpolated Confidence Intervals:
##
##                  Conf.Level   L.E.pt U.E.pt
## Lower Achieved CI    0.9340 -30.0000    -22
## Interpolated CI      0.9500 -30.4642    -22
## Upper Achieved CI    0.9685 -31.0000    -22
```

Consult the output to determine if there is a significant difference in the amount of corticosterone produced in females with incompatible vs compatible mates.

Answer question 11 on Sakai

Exercise 5: Additional Practice

Contraception prevalence

The data in the file contraceptive prevalence provides the percentage of married women ages $15 - 49$ who are practicing, or whose sexual partners are practicing, any form of contraception. Data have been compiled from World Bank for 57 countries for the year 2002 and the year 2012. You are interested in analyzing the data to determine if there is a difference in contraception between these two years.

First, analyze the data to determine whether it violates the assumptions of the t-test that would apply to this data.

Then, analyze the data using the appropriate statistical test to determine whether there is a change in contraception prevalence between 2002 and 2012.

Answer questions 12 & 13 on Sakai