1. (Problem 19, p 313)
(a) Researchers counted the number of electric fish species above and below the entrance points of the same tributaries. Thus, the samples are paired.

<table>
<thead>
<tr>
<th>Upstream number of species</th>
<th>Downstream number of species</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>16</td>
<td>-3</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>25</td>
<td>21</td>
<td>-4</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

95% CI for \( \mu_{d} \) (\( \mu_{down} - \mu_{up} \))

\[
\bar{d} = 1.83
\]

\[
s_{d} = \sqrt{\frac{1}{11} \sum_{i=1}^{12} (d_{i} - 1.83)^{2}} = 3.325749
\]

\[
SE = \frac{s_{d}}{\sqrt{n}} = \frac{3.325749}{\sqrt{12}} = 0.96
\]

\[
t_{11,0.05/2} = 2.20
\]

95% CI for \( \mu_{d} \)

\[
\bar{d} \pm t_{11,0.05/2} SE
\]

\[
\Rightarrow 1.83 \pm 2.20 \times 0.96
\]

\[
\therefore -0.28 < \mu_{d} < 3.96
\]
b) We'll do a paired t-test.

\[ H_0 : \mu_D = 0 \quad \text{vs} \quad H_\alpha : \mu_D \neq 0 \]

\[ T = \frac{\bar{D} - 0}{S_D / \sqrt{12}} \sim t - \text{distribution with d.f. = 11} \]

\[ t = \frac{\bar{d} - 0}{SE} = \frac{1.83}{0.96} = 1.91 \]

\[ p - \text{value} = 2 \cdot P(T \geq 1.91) = 0.083 \]

Since the p-value is bigger than 0.05, we cannot reject Ho. It means that there is no evidence that tributaries have any effect on electric fish species diversity.

c) We had to assume that the species count differences were normally distributed and random sample.

Check the normality.

**R code:**

```r
mat1=c(14, 11, 8, 5, 10, 5, 23, 29, 19, 16, 25, 10)
mat2=c(19, 18, 8, 7, 16, 6, 24, 30, 16, 20, 21, 12)

qqnorm(mat2-mat1)
```

According to the normal quantile plot, we can say this distribution follows a normal distribution. So, normality assumption is satisfied.
2. (Problem 23, p 314)

a)

<table>
<thead>
<tr>
<th>Oxygen consumption after non-feeding dive</th>
<th>Oxygen consumption after feeding dive</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>42.2</td>
<td>71</td>
<td>28.8</td>
</tr>
<tr>
<td>51.7</td>
<td>77.3</td>
<td>25.6</td>
</tr>
<tr>
<td>59.8</td>
<td>82.6</td>
<td>22.8</td>
</tr>
<tr>
<td>66.5</td>
<td>96.1</td>
<td>29.6</td>
</tr>
<tr>
<td>81.9</td>
<td>106.6</td>
<td>24.7</td>
</tr>
<tr>
<td>82</td>
<td>112.8</td>
<td>30.8</td>
</tr>
<tr>
<td>81.3</td>
<td>121.2</td>
<td>39.9</td>
</tr>
<tr>
<td>81.3</td>
<td>126.4</td>
<td>45.1</td>
</tr>
<tr>
<td>96</td>
<td>127.5</td>
<td>31.5</td>
</tr>
<tr>
<td>104.1</td>
<td>143.1</td>
<td>39</td>
</tr>
</tbody>
</table>

\[ \bar{d} = 31.78 \]

The mean change in oxygen consumption was 31.78 ml O2 kg\(^{-1}\).

b.

Researchers measured the metabolic cost after non-feeding dive and after feeding dive using the same seals, respectively. Thus, the samples are paired.

99% CI for \( \mu_D = \mu_{\text{feed}} - \mu_{\text{non-feed}} \)

\[ \bar{d} = 31.78 \]

\[ s_d = \sqrt{\frac{1}{9} \sum_{i=1}^{10} (d_i - 31.78)^2} = 7.29624 \]

SE = \[ \frac{s_d}{\sqrt{10}} = 2.307274 \]

\[ t_{9,0.005} = 3.25 \]

99% CI for \( \mu_0 \)

\[ 31.78 \pm 3.25 \cdot 2.3 \]

\[ 24.3 < \mu_0 < 39.3 \]
c)

Since the samples are paired, we’ll do a paired t-test.

\[ H_0 : \mu_D = 0 \quad \text{vs} \quad H_a : \mu_D > 0 \]

\[ T = \frac{D - 0}{S_D / \sqrt{n}} \sim t \text{ distribution with d.f.} = 9 \]

\[ t = \frac{31.78}{2.307274} = 13.78 \]

\[ p\text{-value} = P(T \geq 13.78) = 1.175706e-07 \]

Since the p-value is smaller than 0.05, we reject Ho. Thus, feeding increases the metabolic costs of a dive.

3. (Problem 19, p 353)

Histogram:

![Histograms](image)

R code:

```r
dat=read.csv("13q19Vines.csv")
diff=(dat$climbing-dat$nonclimbing)
par(mfrow=c(1,3))
hist(dat$climbing,breaks=20,main="Climbing", xlab="number of species")
hist(dat$nonclimbing,breaks=20,main="Non-Climbing", xlab="number of species")
hist(diff,breaks=20,main="Difference between Climbing and Non-Climbing", xlab="number of species")
```
Normal quantile plot:

R code:
```r
par(mfrow=c(1,3))
qqnorm(dat$climbing, main="Climbing")
qqnorm(dat$nonclimbing,main="Non-Climbing")
qqnorm(diff, main="Difference between Climbing and Non-Climbing")
```

The distribution of the number of species in clades that have climbing vines is right skewed and the distribution of the number of species in clades that have non-climbing vines is right skewed. The distribution of difference between climbing and non-climbing has a heavy left tail. Thus, it’s hard to say that these distributions are a normal distribution.

Since this distribution doesn’t follow a normal distribution, we will do a sign test.

$H_0$ : the difference $D$ in the number of species between climbing and non-climbing has a median of 0.

$H_a$ : median of $D$ is not 0

$Y_+ \sim B(48,0.5)$ under $H_0$

$y = # \text{ of observed } ^{+}\text{ differences in the } # \text{ of species} = 38$

p-value $= 2 \cdot P(Y_+ \geq 38)$

$= 2 \cdot \left( \sum_{i=38}^{48} \binom{48}{i} 0.5^i \cdot 0.5^{48-i} \right)$

$= 2 \cdot (1 - \text{pbinom}(37,48,0.5)) \quad (R \text{ code})$

$= 6.169641e-05$
Using the normal approximation.
\[ p \cdot n = 0.5 \cdot 48 = 24 > 5 \]
\[ (1 - p)n = 0.5 \cdot 48 = 24 > 5 \]
where \( p \) comes from \( H_0 \)
\[ Y_+ \sim N(24, 3.464102) \]
\[ p \text{-value} = 2 \cdot P(Y_+ \geq 38) \]
\[ = 2 \cdot P\left(\frac{Y_+ - 24}{3.464102} \geq \frac{38 - 24}{3.464102}\right) \]
\[ = 2 \cdot P(Z \geq 4.041452), \quad Z \sim N(0,1) \]
\[ < 0.00004 \]

Since the p-value is smaller than 0.05, we reject \( H_0 \). It means that clades with climbing vines have more species on average than species with non-climbing vines.

4. (Problem 23, p 355)

Normal quantile plot and Histogram:

![Normal quantile plot and Histogram](image)

R code:

Biomass=read.csv("13q23Clearcuts.csv")
par(mfrow=c(1,2))
qqnorm(Biomass,main="normal quantile plot of biomass different")
hist(Biomass,breaks=10,main="Histogram of Biomass difference",xlab="Biomass change")
This distribution is left skewed. Thus, the normality assumption is not satisfied.

Researchers compiled data on the biomass before and after clear-cutting at the same areas. Thus, the samples are paired. But, since the distribution is not a normal distribution, we'll use a sign test.

\[ H_0 : \text{difference in biomass between before and after clear-cutting has a median of 0} \]
\[ H_a : \text{median of D is not 0} \]
\[ Y_+ \sim B(36, 0.5) \quad \text{under } H_0 \]
\[ y = \# \text{ of observed } "+" \text{ signs} \]
\[ = 21 \]
\[ p - value = 2 \cdot P(Y_+ \geq 21) \]
\[ = 2 \cdot \left( \sum_{i=21}^{36} \binom{36}{i} 0.5^i \cdot 0.5^{36-i} \right) \]
\[ = 2 \cdot (1 - \text{pbinom}(20, 36, 0.5)), \quad (R \text{ code}) \]
\[ = 0.4 \]

Using the normal approximation.

Check the assumption
\[ np = 36 \cdot 0.5 = 18 > 0.5 \]
\[ n(1 - p) = 36 \cdot 0.5 = 18 > 0.5 \]
\[ Y_+ \sim N(18, 3) \]
\[ p - value = 2 \cdot P(Y_+ \geq 21) \]
\[ = 2 \cdot P \left( \frac{Y_+ - 18}{3} \geq \frac{21 - 18}{3} \right) \]
\[ = 2 \cdot P(Z \geq 1), \quad \text{where } Z \sim N(0, 1) \]
\[ = 0.32 \]

Since the p-value is bigger than 0.05, we cannot reject \( H_0 \). It means that there is no evidence that there is an average change in biomass between before and after clear-cutting.