Biological hypothesis testing using Python

Author: Hansi Thewarapperuma

Date: 31.01.2023

1.

Normal human body temperature, as kids are taught in North America, is 98.6°F. But how well is this supported by data? Researchers obtained body-temperature measurements on randomly chosen healthy people (Shoemaker 1996), which can be found in "Temperature.csv" file.

μ_1 = sample mean body temperature

μ_2 = population mean body temperature

- Null hypothesis (H0): The mean normal human body temperature is not different from 98.6°F /The normal human body temperature is equal to 98.6°F ($\mu_1 - \mu_2 = 0$)
- Alternative hypothesis (HA): The mean normal human body temperature is different from 98.6°F /The normal human body temperature is not equal to 98.6°F ($\mu_1 - \mu_2 \neq 0$)

Assumptions:

- Data are normally distributed
- Random sampling
- Independent measurements

Statistics for the human temperature variable:

Code:

```python
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
from scipy import stats
from statsmodels.graphics.gofplots import qqplot
```

```python
import pandas as pd

temperature = pd.read_csv('Temperature.csv')
# print(temperature)
print(temperature.describe())
```

Output:

```
         temperature
count     25.000000
mean      98.524000
std        0.677791
min       97.400000
25%       98.000000
50%       98.600000
75%       99.000000
max      100.000000
```

Testing the normality assumption for the temperature variable

- Drawing a histogram and Quantile-Quantile plot (QQ plot)
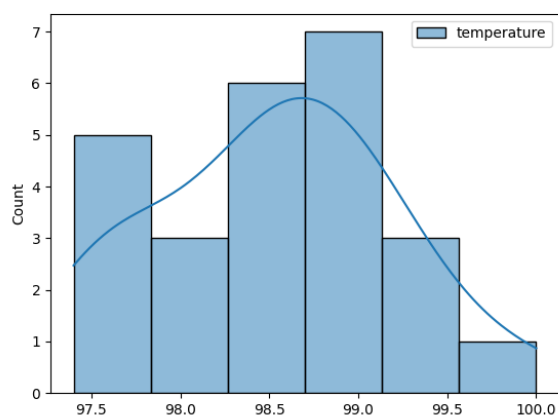- Perform Shapiro-Wilk test for the normality on the variable

Histogram for the variable

Code:

```python
import matplotlib.pyplot as plt
import seaborn as sns

sns.histplot(temperature, kde= True)
plt.show()
plt.savefig('temperature histogram using seaborn')
plt.title('Histogram of temperature')
```
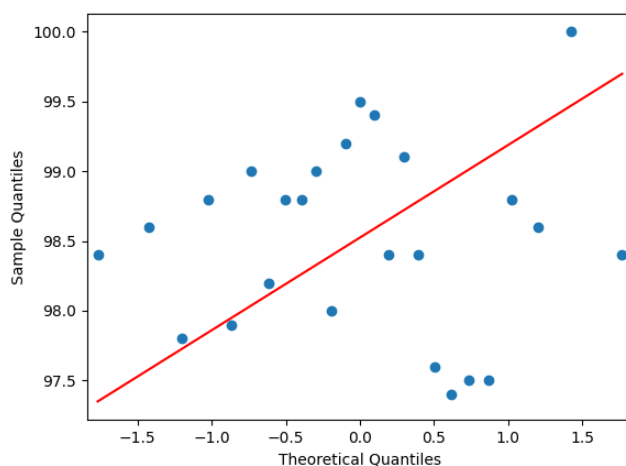
Result:

QQ-plot

Code:

```python
from statsmodels.graphics.gofplots import qqplot
qqplot(temperature, line= 's')
plt.show()
plt.savefig('QQplot for temperature.jpg')
plt.title('QQplot for temperature')
```

Output:

Interpretation: In the histogram the curve is not exactly bell shaped i.e. slightly deviated from the bell shape, with one peak and the curve is asymmetric around the mean hence it is slightly skewed. This depicts a slight deviation from normality

In the QQ plot the data points are highly scattered from the reference line (If the data is normally distributed, the points will fall on the 45-degree reference line) This also depicts a slight deviation from normality

- Null hypothesis (H0) : Data are normally distributed
- Alternative hypothesis (HA) : Data are not normally distributed

Shapiro-Wilk test for normality on the variable

Code:

```
from scipy import stats
stat,p = stats.shapiro(temperature)
print('stat=%.3f p-value=%.3f'%(stat,p))
```

Output:

```
stat=0.972 p-value=0.700
```

P value = 0.700

α value = 0.05

Decision: Since p value is greater than α value, null hypothesis fails to be rejected

Conclusion: At 5% of significant level, data are normally distributed

Therefore, parametric test is recommended. – One sample t test

Test statistics: one sample t test

Code:

```
known_temperature = 98.6
t,p = stats.ttest_1samp(temperature, known_temperature)
print('t stat: ', t)
print('p value: ',p)
```

Output:

```
t stat:  [-0.56064519]
p value:  [0.58023628]
```

P value = 0.58023628

Decision: Since p value is greater than α value, null hypothesis fails to be rejected

Conclusion: At 5% of significant level, the mean normal human body temperature is not different from 98.6°F /The normal human body temperature is equal to 98.6°F

2.

The horned lizard *Phrynosoma mcallii* has many unusual features, including the ability to squirt blood from its eyes. The species is named for the fringe of spikes surrounding the head. Herpetologists recently tested the idea that long spikes help protect horned lizards from being eaten, by taking advantage of the gruesome but convenient behavior of one of their main predators—the loggerhead shrike, *Lanius ludovicianus*. The loggerhead shrike is a small predatory bird that skewers its victims on thorns or barbed wire, to save for later eating. The researchers identified the remains of 30 horned lizards that had been killed by shrikes and measured the lengths of their horns (Young et al. 2004). As a comparison group, they measured the same trait on 154 horned lizards that were still alive and well. These data can be found in "HornedLizards.csv" file. The mean horn lengths of the dead lizards were compared with those of the living lizards.

μ_1 = The mean horn length of survived lizards

μ_2 = The mean horn length of dead lizards

- Null hypothesis (H0): The mean horn length of survived lizards are less than or equal to the mean horn length of dead lizards ($\mu_1 - \mu_2 \leq 0$)
- Alternative hypothesis (HA): The mean horn length of survived lizards are greater than the mean horn length of dead lizards ($\mu_1 - \mu_2 > 0$)

Assumptions:

- Data are normally distributed
- Random sampling
- Independent measurements

Code:

```python
import pandas as pd
```

```
# Question 2

horn_length = pd.read_csv('HornedLizards.csv')
# print(horn_length)

# get rid of NA values in csv file
new_data = horn_length.dropna(axis = 0, how ='any')
# print(len(horn_length))
# print(len(new_data))

print(horn_length.describe())
```

Output statistics:

```
        Squamosal horn length
count            184.000000
mean              23.907065
std                2.769862
min               13.100000
25%               22.400000
50%               24.150000
75%               25.725000
max               30.300000
```

Testing the normality assumption for the two independent samples.

- Drawing histograms and Quantile-Quantile plots (QQ plots) for each sample
- Perform Shapiro-Wilk test for the normality on the variable

**Create the two samples (survived lizards/ dead lizards) from the single csv file

```
# ******filter pandas data values by column value using df.loc*****

# extract horn lengths of survived
# step1- extract both values and annotation columns from mixed csv and assign it to 'survived'
survived = new_data.loc[new_data['Survive']=='survived']
# print(survived)
# step2 - extract the values from 'survived'
survived_val = survived.loc[:,"Squamosal horn length"]
# print(survived_val)
# print(len(survived_val))
# check = survived['Squamosal horn length']
# print(len(check))

# extract horn lengths of dead
# step1- extract both values and annotation columns from mixed csv and assign it to 'dead'
dead = new_data.loc[new_data['Survive']=='dead']
# step2 - extract the values from 'dead'
dead_val = dead.loc[:,"Squamosal horn length"]
# print(dead_values)
```

Histograms

Code:

```python
# histogram for the survived sample
fig, axs = plt.subplots(1,2,figsize=(12,4))
# ******mistake- replace survived by x
sns.histplot(survived_val, kde= True, ax=axs[0])
axs[0].set_title('Histogram of survived lizards')

# histogram for the dead sample
sns.histplot(dead_val, kde=True, ax=axs[1])
axs[1].set_title('Histogram of dead lizards')
plt.show()
plt.savefig('Histogram of 2 samples.jpg')
```
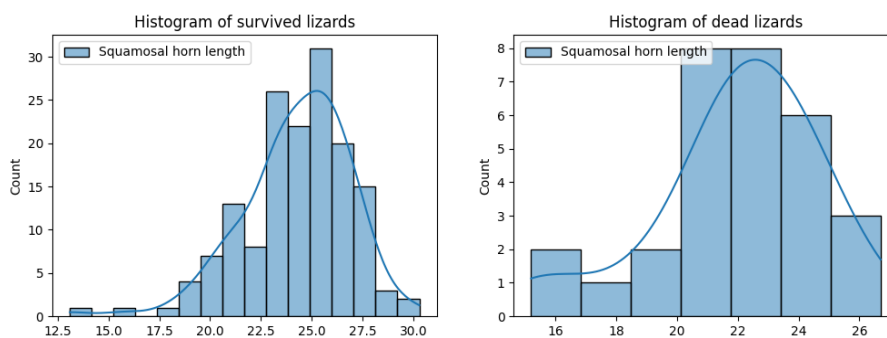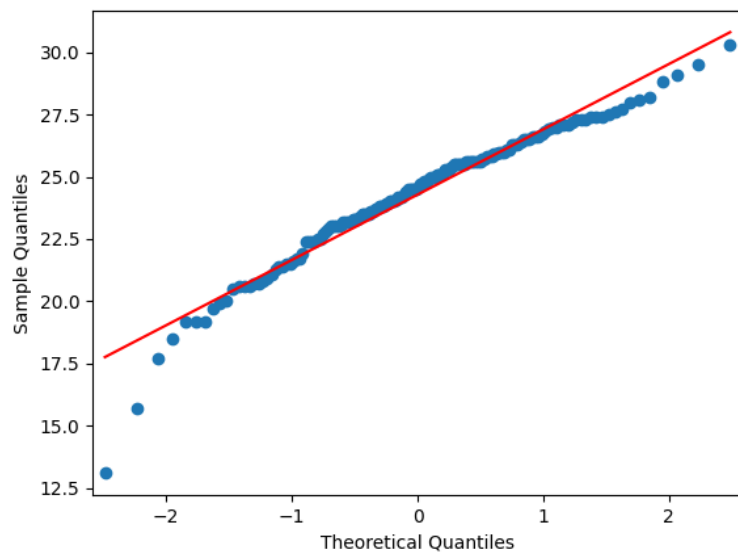
Output:

QQ plot

Survived sample

Code:

```python
# QQplot for survived sample
# from statsmodels.graphics.gofplots import qqplot
qqplot(survived_val, line= 's')
plt.show()
plt.title('QQplot for survived sample')
plt.savefig('QQplot for survived sample.jpg')
```
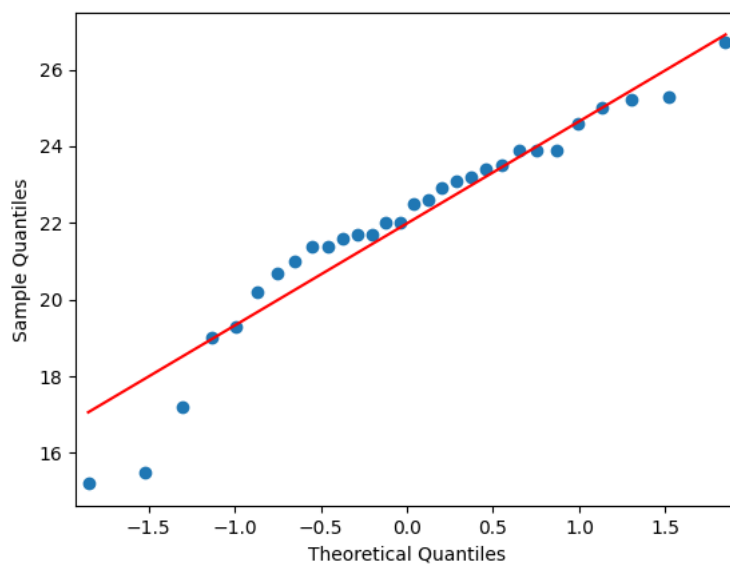
Output:

Dead sample

Code:

```
# QQplot for dead sample
qqplot(dead_val, line= 's')
plt.show()
plt.title('QQplot for dead sample')
plt.savefig('QQplot for dead sample.jpg')
```

Output:

Interpretation: In the histogram both curves sre not exactly bell shaped i.e. slightly deviated from the bell shape, with one peak and the curves are asymmetric around the mean hence they are skewed. This depicts a slight deviation from normality

In the QQ plots the data points are scattered from the reference line (If the data is normally distributed, the points will fall on the 45-degree reference line) This also depicts a slight deviation from normality However in the survived sample qq plot data points are comparatively closer to the best fit line

- Null hypothesis (H0) : Data are normally distributed
- Alternative hypothesis (HA) : Data are not normally distributed

Shapiro-Wilk test for normality for survived sample

Code:

```
# from scipy import stats
stat,p = stats.shapiro(survived_val)
print('stat=%.3f p-value=%.3f'%(stat,p))
```

Output:

```
stat=0.961 p-value=0.000
```

P value = 0.000

α value = 0.05

Decision: Since p value is less than α value, null hypothesis is rejected

Conclusion: At 5% of significant level, data are not normally distributed

Shapiro-Wilk test for normality for dead sample

Code:

```
# from scipy import stats
stat,p = stats.shapiro(survived_val)
print('stat=%.3f p-value=%.3f'%(stat,p))
```

Output:

```
stat=0.935 p-value=0.065
```

P value = 0.065

α value = 0.05

Decision: Since p value is greater than α value, null hypothesis fails to be rejected

Conclusion: At 5% of significant level, data are normally distributed

Hence non parametric test is recommended. – MannWhitney

**Comparison of means. To further visualize the two samples, draw boxplots and a violin plots to compare the distribution of the two samples
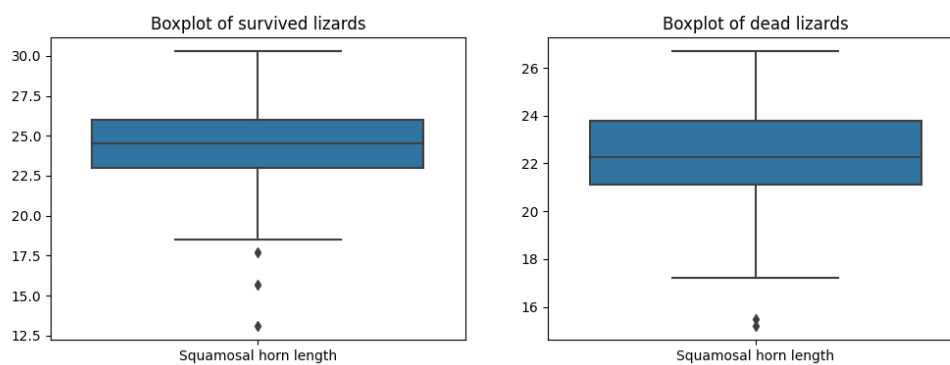
Box plot

Code:

```
# comparison of means using boxplot and violin plot
# boxplot for survived sample
fig, axs = plt.subplots(1,2,figsize=(12,4))
sns.boxplot(survived, ax=axs[0])
axs[0].set_title('Boxplot of survived lizards')

# boxplot for the dead sample
sns.boxplot(dead, ax=axs[1])
axs[1].set_title('Boxplot of dead lizards')
plt.show()
plt.savefig('Histogram of 2 samples.jpg')
```
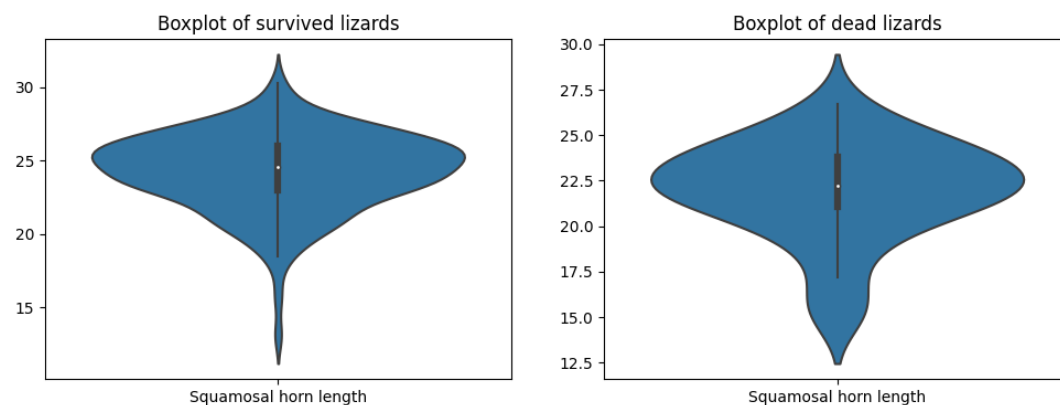
Output:

Violin plot

Code:

```
# violinplot for survived sample
fig, axs = plt.subplots(1,2,figsize=(12,4))
sns.violinplot(survived, ax=axs[0])
axs[0].set_title('Violinplot of survived lizards')

# violinplot for the dead sample
sns.violinplot(dead, ax=axs[1])
axs[1].set_title('Violinplot of dead lizards')
plt.show()
plt.savefig('Violinplots of 2 samples.jpg')
```

Output:

Interpretation:

Boxplot: The median value for survived lizards is greater than of its dead lizards. The minimum horn length of survived lizards is greater than the minimum horn length of dead lizards. Thus survived ones are more likely to have longer horns

Violin plot: in survived lizards higher probability of data points are around 25 while in dead lizards its around 22.5 . Thus majority oof survived lizards are more likey to have longer horns

Performing MannWhitney test

Code:

```
# perform MannWhitney test
t,p = stats.mannwhitneyu(survived_val, dead_val, alternative= 'greater')
print('t stat: ',t)
print('p value: ', p)
```

Output:

```
t stat:  3438.5
p value:  1.1827706967305409e-05
```

P value = 1.1827706967305409e-05

Decision: Since p value is less than α value, null hypothesis is rejected

Conclusion: At 5% of significant level, the mean horn length of survived lizards are greater than the mean horn length of dead lizards

3.

In many species, males are more likely to attract females if the males have high testosterone levels. Are males with high testosterone paying a cost for this extra mating success in other ways? One hypothesis is that males with high testosterone might be less able to fight off disease—that is, their high levels of testosterone might reduce their immunocompetent. To test this idea, Hasselquist et al. (1999) experimentally increased the testosterone levels of 13 male red-winged blackbirds by surgically implanting a small permeable tube filled with testosterone. They measured immunocompetence as the rate of antibody production in response to a nonpathogenic antigen in each bird's blood serum both before and after the implant. The antibody production rates were measured optically, in units of log 10−3 optical density per minute (ln[mOD/min]). The data is available in "BlackbirdTestosterone.csv"

$\mu 1$ = The mean rate of antibody production in response to a nonpathogenic antigen in bird's blood serum after implant

$\mu 2$ = The mean rate of antibody production in response to a nonpathogenic antigen in bird's blood serum before implant

- Null hypothesis (H0):): The mean antibody production after implanting the testosterone tube is greater than or equal to antibody production before implanting ($\mu1-\mu2>=0$)
- Alternative hypothesis (HA): The mean antibody production after implanting the testosterone tube is less than of its before implanting ($\mu1- \mu2<0$)

Assumptions:

- Data are normally distributed
- Random sampling
- Independent measurements

Code:

```
# create dataframe
antibody_production = pd.read_csv('BlackbirdTestosterone.csv')

# select a subset of a dataframe
log_before = antibody_production['log before']
log_after = antibody_production['log after']
log_difference = antibody_production['dif in logs']

# log_before.head()

print(log_before.describe())
print(log_after.describe())
print(log_difference.describe())
```

Output:

```
count    13.000000
mean      4.733846
std       0.279837
min       3.910000
25%       4.700000
50%       4.800000
75%       4.880000
max       4.980000
Name: log before, dtype: float64
count    13.000000
mean      4.790000
std       0.261598
min       4.300000
25%       4.600000
50%       4.960000
75%       5.000000
max       5.020000
Name: log after, dtype: float64
count    13.000000
mean     -0.056154
std       0.159245
min      -0.390000
25%      -0.130000
50%      -0.070000
75%       0.050000
max       0.210000
```

Log before:

mean   4.733846

std   0.279837

Log after:

mean   4.790000

std   0.261598

log difference:

mean   -0.056154

std   0.159245

Testing the normality assumption for the log difference variable.

- Drawing histograms and Quantile-Quantile plots (QQ plots) for each sample
- Perform Shapiro-Wilk test for the normality on the variable

Histogram

Code:

```
# plotting histogram for log difference using seaborn
sns.histplot(log_difference, kde= True)
plt.show()
plt.savefig('histogram for log difference using seaborn.jpg')
plt.title('Histogram of log difference')
```
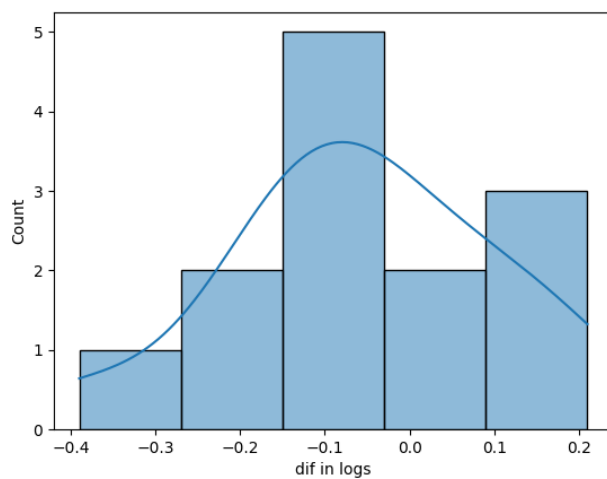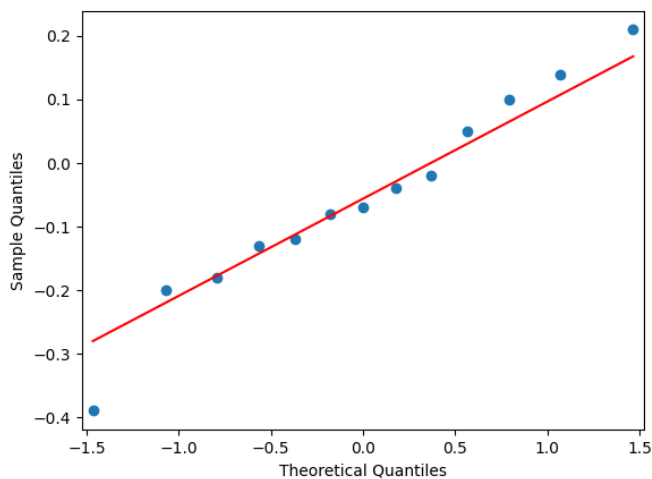

QQ plot

Code:

```
# plotting qqplot for the log difference using statmodels
qqplot(log_difference, line= 's')
plt.show()
plt.title('QQplot for log difference ')
plt.savefig('QQplot for log difference.jpg')
```

Output:

Interpretation: Histogram: the curve is skewed, not exactly a bell shape. Hence it is likely to deviate from normality

QQ plot: the datapoints are much closer to best fit line. Therefore normality can be assumed

- Null hypothesis (H0) : Data are normally distributed
- Alternative hypothesis (HA) : Data are not normally distributed

Shapiro-Wilk test for normality on the variable

Code:

```
stat,p = stats.shapiro(log_difference)
print('stat=%.3f p-value=%.3f'%(stat,p))
```

Output:

```
stat=0.979 p-value=0.977
```

P value = 0.977   α value = 0.05

Decision : Since p value is greater than α value, null hypothesis fails to be rejected

Conclusion : At 5% of significant level, data are normally distributed

Hence, parametric test is recommended – paired sample t test

Comparison of means. To further visualize the two samples, draw boxplots and violin plots to compare the distribution of the two before and after samples
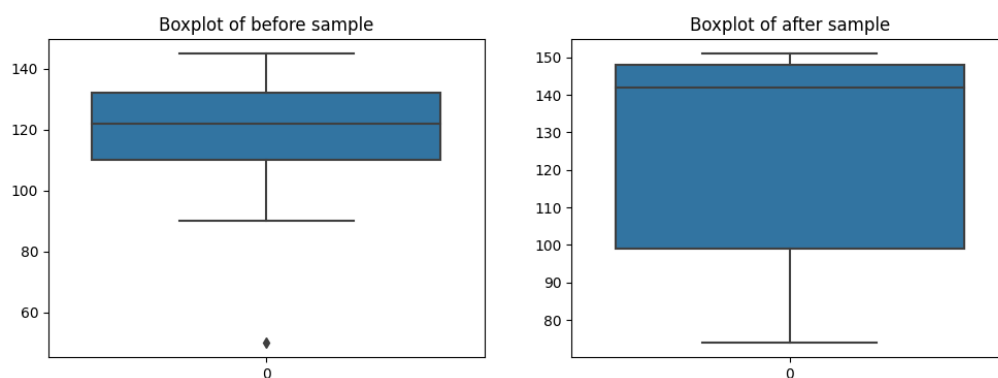
Boxplot

Code:

```
# comparison of means using boxplot and violin plot
# boxplot for before sample
fig, axs = plt.subplots(1,2,figsize=(12,4))
sns.boxplot(before_sample, ax=axs[0])
axs[0].set_title('Boxplot of before sample')

# boxplot for the after sample
sns.boxplot(after_sample, ax=axs[1])
axs[1].set_title('Boxplot of after sample')
plt.show()
plt.savefig('Histogram of 2 before and after samples.jpg')
```

Output:

Violin plot

Code:

```
# violinplot for before sample
fig, axs = plt.subplots(1,2,figsize=(12,4))
sns.violinplot(before_sample, ax=axs[0])
axs[0].set_title('Violinplot of before sample')

# violinplot for the after sample
sns.violinplot(after_sample, ax=axs[1])
axs[1].set_title('Violinplot of after sample')
plt.show()
plt.savefig('Violinplots of 2 samples.jpg')
```

Output:

| Violinplot of before sample | Violinplot of after sample |
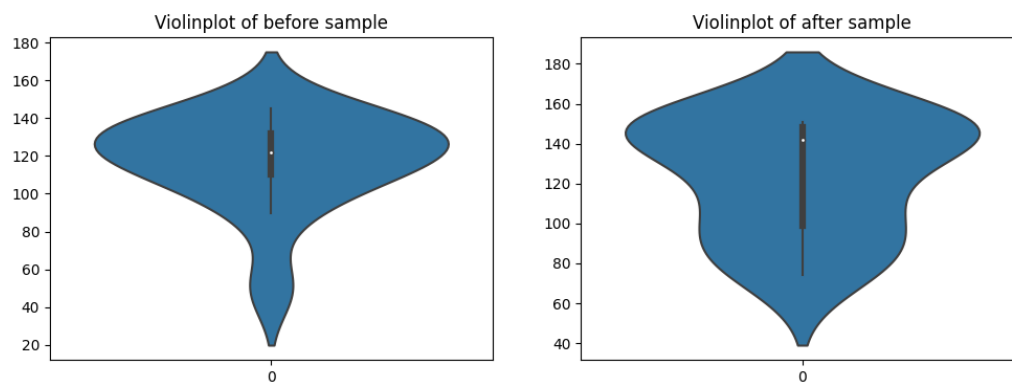
Interpretation: boxplot- median value of before sample is less than of its after sample The minimum value for datapoints in before sample is greater than of its after sample

Violinplot: In the before sample higher probability of datapoints are around the value of 130 while in after sample higher probability of datapoints are around the value of 150(wider section of the plot) Thus after implanting sample would more likely to have higher antibody production than before implanting sample

**Performing paired sample T test

Code:

```
# ********order of log_after anf log_before
t,p = stats.ttest_rel(log_before, log_after, alternative='less')
print('t stat: ',t)
print('p value: ', p)
```

Output:

```
t stat:  -1.27140091765582776
p value:  0.1138369363627144
```

P value = 0.1138

Decision: Since p value is greater than α value, null hypothesis fails to be rejected

Conclusion: At 5% of significant level, The mean antibody production after implanting the testosterone tube is greater than or equal to antibody production before implanting

4.

Many parasites have more than one species of host, so the individual parasite must get from

one host to another to complete its life cycle. Trematodes of the species *Euhaplorchis*

californiensis use three hosts during their life cycle. Worms mature in birds and lay eggs that pass out of the bird in its feces. The horn snail *Cerithidea californica* eats these eggs, which hatch and grow to another life stage in the snail, sterilizing the snail in the process. When an infected snail is eaten by the California killifish *Fundulus parvipinnis*, the parasite develops to the next life stage and encysts in the fish's braincase. Finally, when the killifish is eaten by a bird, the worm becomes a mature adult and starts the cycle again.

Researchers have observed that infected fish spend excessive time near the water surface,

where they may be more vulnerable to bird predation. This would certainly be to the worm's advantage, as it would increase its chances of being ingested by a bird, its next host. Lafferty and Morris (1996) tested the hypothesis that infection influences risk of predation by birds. A large outdoor tank was stocked with three kinds of killifish: unparasitized, lightly infected, and heavily infected. This tank was left open to foraging by birds, especially great egrets, great blue herons, and snowy egrets.

Observed frequencies of fish eaten or not eaten by birds according to trematode infection level is given below.

	Uninfected	Lightly infected	Highly infected	Row total
Eaten by birds	1	10	37	48
Not eaten by birds	49	35	9	93
Column total	50	45	46	141

It is essential to test whether the probability of being eaten by birds differs according to infection status.

- Null hypothesis (H0): The probability of being eaten by birds and infection status are independent
- Alternative hypothesis (HA): The probability of being eaten by birds and infection status are not independent/ related

Assumptions:

- Both variables are categorical
- Independent observations
- Expected value of cells should be 5 or greater in at least 80% of cells
- Cells in the contingency table are mutually exclusive.

The above contingency table was created in a Pandas DataFrame

Code:

```
# question 3
# create manual dataframe
df = pd.DataFrame([[1,10,37],[49,35,9]],index=['eaten','not_eaten'],columns=['uninfected','lightly_infected','highly_infected'])
print(df)
```

Output:

```
          uninfected  lightly_infected  highly_infected
eaten              1                10               37
not_eaten         49                35                9
```
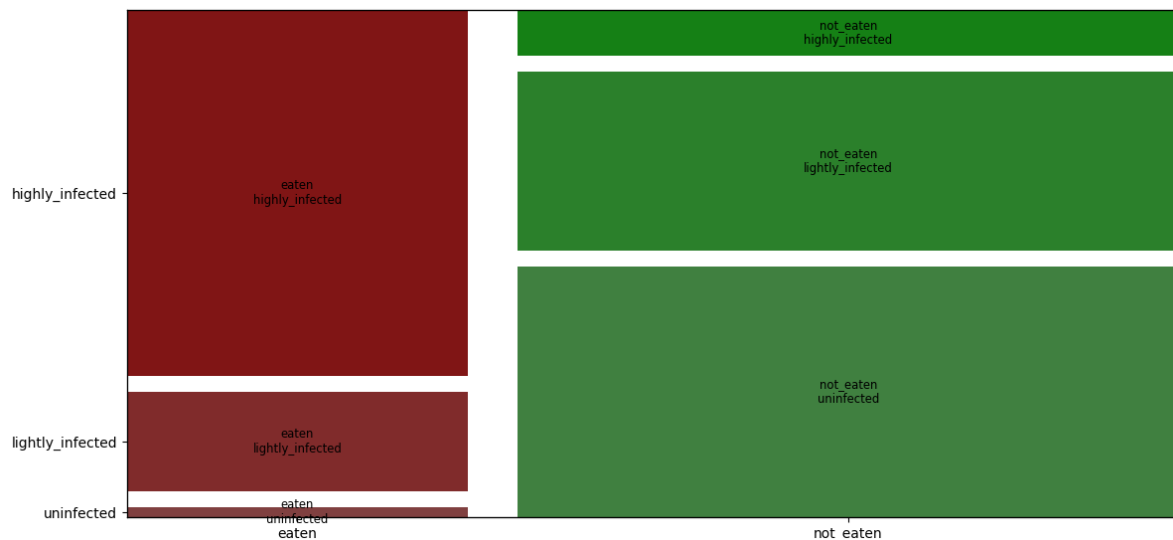
Mosaic plot was drawn for the above table

Code:

```
df2 = df.stack()
print(df2)

df3 = df2.to_dict()
print(df3)

# plotting the mosaic plot
from statsmodels.graphics.mosaicplot import mosaic
mosaic(df3, gap= 0.05)
plt.show()
plt.title('mosaic plot')
plt.savefig('mosaic plot.jpg')
```

Output:

Chi-square contingency test on above data (Chi-square statistic, the P-value and the degree of freedom value)

Code:

```python
t = stats.chi2_contingency(df)
print('t stat: ',t)
```

Output:

```
chi square statistic:  Chi2ContingencyResult(statistic=69.75570515817361, pvalue=7.12428171268388e-16, dof=2, expected_freq=array([[17.0212766 , 15.3191489
    [32.9787234 , 29.68085106, 30.34042553]]))
```

Chi square statistic: 69.75570515817361

P value: 7.12428171268388e-16

Dof = 2

Output the expected value table. A new DataFrame was created to store the expected values.

Code:

```python
expected_values_table = pd.DataFrame(chi.expected_freq, index=['eaten','not_eaten'], columns=['uninfected','lightly_infected','h
print(expected_values_table)
```

Output:

	uninfected	lightly_infected	highly_infected
eaten	17.021277	15.319149	15.659574
not_eaten	32.978723	29.680851	30.340426

Decision: Since p value is less than alpha value (0.05), null hypothesis is rejected

Conclusion: At 5% of significance level, the frequency of being eaten by birds and infection status are not independent/ related

References

•   Whitlock, Michael C., and Dolphcoaut Schluter. The analysis of biological data. No. 574.015195 W5. 2009.

•   Young, Kevin V., Edmund D. Brodie Jr, and Edmund D. Brodie III. "How the horned lizard got its horns." Science 304.5667 (2004): 65-65.

•   Hasselquist, Dennis, et al. "Is avian humoral immunocompetence suppressed by testosterone?." Behavioral Ecology and Sociobiology 45.3-4 (1999): 167-175.

•   Lafferty, Kevin D., and A. Kimo Morris. "Altered behavior of parasitized killifish increases susceptibility to predation by bird final hosts." Ecology 77.5 (1996): 1390-1397.