CHAPTER 10

1. When would the mean grade in a class on a final exam be considered a statistic? When would it be considered a parameter?

Solution : A parameter is a value calculated in a population. A statistic is a value computed in a sample estimate a parameter

5. When you construct a 95% confidence interval, what are you 95% confident about?

Solution : If you repeated the sampled process by which the observed data was obtained, the fraction of confidence intervals containing the population parameter would be 95%. In other words, a 95% confidence interval interprets as the interval for which we are 95% confident that the population parameter which we are estimating will lie in this interval.

10. The effectiveness of a blood-pressure drug is being investigated. How might an experimenter demonstrate that, on average, the reduction in systolic blood pressure is 20 or more?

Solution : The experimenter would have to calculate the confidence interval for the mean blood pressure difference and determine if the lower limit of the interval is greater than or equal to 20

15. You take a sample of 22 from a population of test scores, and the mean of your sample is 60. (a) You know the standard deviation of the population is 10. What is the 99% confidence interval on the population mean. (b) Now assume that you do not know the population standard deviation, but the standard deviation in your sample is 10. What is the 99% confidence interval on the mean now?

(a)
We can construct a z-confidence interval as follows
x + or - z * σ/√(n)
Where, x = mean = 60, σ = 10, sample size n = 22
We know that for 99% confidence interval, the critical value is 2.576

Using the above information,
60 +/- 2.576(10/√(22))
60 - 5.4921 = 54.51
60 + 5.4921 = 65.49
Solution : (54.51, 65.49)

(b)
Assuming we do not know the population standard deviation, we can compute the t-confidence interval using the available information
x + or - s * σ/√(n)

Where, x = mean =60, s = 10, sample size n = 22
For confidence interval 99%, with degrees of freedom 21, we have the t critical value = 2.831

60 +/- 2.831(10/√(22))
60 - 6.0357 = 53.96
60 + 6.0357 = 66.04

Solution : (53.96, 66.04)

20. True/false: You have a sample of 9 men and a sample of 8 women. The degrees of freedom for the t value in your confidence interval on the difference between means is 16.

Solution: False.
Degrees of freedom = (Sample size of group 1 – 1) + (Sample size of group 2 -1) = (9-1) + (8-1) = 8+7 = 15

24. Is there a difference in how much males and females use aggressive behaviour to improve an angry mood? For the "Anger-Out" scores, compute a 99% confidence interval on the difference between gender means.

To compute the confidence interval :

$$\text{Lower Limit} = M_1 - M_2 - (t_{CL})(S_{M_1-M_2})$$
$$\text{Upper Limit} = M_1 - M_2 + (t_{CL})(S_{M_1-M_2})$$

```python
# import libraries
import pandas as pd
import numpy as np
import statistics
import math

# read by default 1st sheet of an excel file
angry_moods = pd.read_excel('angry_moods.xls')

angry_moods["Gender_Desc"] = np.where(angry_moods["Gender"].isin([1]), 'Male', 'Female')

mean_men = statistics.mean(angry_moods[angry_moods['Gender_Desc'] == 'Male']["Anger-Out"].tolist())
mean_women = statistics.mean(angry_moods[angry_moods['Gender_Desc'] == 'Female']["Anger-Out"].tolist())

sd_men = statistics.stdev(angry_moods[angry_moods['Gender_Desc'] == 'Male']["Anger-Out"].tolist())
sd_women = statistics.stdev(angry_moods[angry_moods['Gender_Desc'] == 'Female']["Anger-Out"].tolist())

angry_moods.groupby(['Gender_Desc'])['Gender_Desc'].count()
```

```
Gender_Desc
Female    48
Male      30
Name: Gender_Desc, dtype: int64
```

```python
#Standard error of means = √(sd_men_sqrd/#men + sd_women_sqrd/#women)
std_error = math.sqrt((sd_men*sd_men)/30 + (sd_women*sd_women)/48)

deg_of_freedom = (30-1) + (48-1)
#degrees of freedom = 76

#using the t for confidence interval calculator : https://onlinestatbook.com/2/calculators/inverse_t_dist.html
t = 2.642

#difference between means
mean_diff = mean_men - mean_women
lower_limit = mean_diff - (t*std_error)
upper_limit = mean_diff + (t*std_error)
```

```python
print("The lower limit is :", round(lower_limit,4))
print("The upper limit is :", round(upper_limit,4))
```

```
The lower limit is : -1.8346
The upper limit is : 3.4262
```

Solution : The lower and upper threshold of a 99% confidence interval is (-1.8346, 3.4263)

25. Calculate the 95% confidence interval for the difference between the mean Anger-In score for the athletes and non-athletes. What can you conclude?

```python
angry_moods["Sports_Desc"] = np.where(angry_moods["Sports"].isin([1]), 'Athletes', 'Non-Athletes')

mean_ath = statistics.mean(angry_moods[angry_moods['Sports_Desc'] == 'Athletes']["Anger-In"].tolist())
mean_non_ath = statistics.mean(angry_moods[angry_moods['Sports_Desc'] == 'Non-Athletes']["Anger-In"].tolist())

sd_ath = statistics.stdev(angry_moods[angry_moods['Sports_Desc'] == 'Athletes']["Anger-In"].tolist())
sd_non_ath = statistics.stdev(angry_moods[angry_moods['Sports_Desc'] == 'Non-Athletes']["Anger-In"].tolist())

angry_moods.groupby(['Sports_Desc'])['Sports_Desc'].count()
```

```
Sports_Desc
Athletes        25
Non-Athletes    53
Name: Sports_Desc, dtype: int64
```

```python
#Standard error of means = √(sd_ath_sqrd/#men + sd_non_ath_sqrd/#women)
std_error = math.sqrt((sd_ath*sd_ath)/25 + (sd_non_ath*sd_non_ath)/53)

deg_of_freedom = (25-1) + (53-1)
#degrees of freedom = 76

#using the t for confidence interval calculator : https://onlinestatbook.com/2/calculators/inverse_t_dist.html
t =1.992

#difference between means
mean_diff = mean_non_ath - mean_ath
lower_limit = mean_diff - (t*std_error)
upper_limit = mean_diff + (t*std_error)
```

```python
print("The lower limit is :", round(lower_limit,4))
print("The upper limit is :", round(upper_limit,4))
```

```
The lower limit is : 0.8069
The upper limit is : 4.7765
```

Solution : The 95% confidence interval is (0.8069, 4.7765)

$$0.8069 \leq \mu_{non\_ath} - \mu_{ath} \leq 4.7765$$

where μ_{non_ath} is the population mean for non-athletes and μ_{ath} is the population mean for athletes. This analysis provides evidence that the **mean for non-athletes is higher than the mean for athletes, and that the difference between means in the population is likely to be between 0.8069 and 4.7765.**

26. Find the 95% confidence interval on the population correlation between the Anger-Out and Control-Out scores.

```python
#correlation between anger out and control out scores
r = angry_moods["Anger-Out"].corr(angry_moods["Control-Out"])

#conversion of r to z' using the calculator : https://onlinestatbook.com/2/calculators/r_to_z.html
z = -0.667
N = 78
std_error = 1/math.sqrt(N-3)

#Z for 95% confidence interval using : https://www.mathsisfun.com/data/confidence-interval.html
Z_95 = 1.96

#upper limit = z' + Z.std_error and lower limit = z' - Z.std_error
lower_limit = z - Z_95*std_error
upper_limit = z + Z_95*std_error

print("The lower limit is :", round(lower_limit,4))
print("The upper limit is :", round(upper_limit,4))
```

```
The lower limit is : -0.8933
The upper limit is : -0.4407
```

Solution : The lower and upper limit of a 95% confidence interval on the population correlation between the Anger-Out and Control-Out scores is (-0.8933, -0.4407)

CHAPTER 11

4. State the null hypothesis for:

a. *An experiment testing whether echinacea decreases the length of colds.*

   H0 : Echinacea does not affect the length of colds

b. *A correlational study on the relationship between brain size and intelligence.*

   H0 : There is no correlation between brain size and intelligence

c. *An investigation of whether a self-proclaimed psychic can predict the outcome of a coin flip.*

   H0 : A self-proclaimed psychic cannot predict the outcome of a coin flip

d. *A study comparing a drug with a placebo on the amount of pain relief. (A one-tailed test was used.)*

   H0 : The pain relief from a placebo control group is greater than or equal to pain relief in the test group

8. A significance test is performed and p = .20. Why can't the experimenter claim that the probability that the null hypothesis is true is .20?

Solution: The p value does not indicate the probability of the null hypothesis being true. Rather, a p value indicates the statistical significance of validating the null hypothesis against the observed data.

14. Why is "Ho: "M1 = M2" not a proper null hypothesis?

Solution : This is not a proper null hypothesis because, the null hypothesis has to be defined using the population parameters (μ_1, μ_2) and not the statistics (M1, M2).

18. You choose an alpha level of .01 and then analyze your data. (a) What is the probability that you will make a Type I error given that the null hypothesis is true? (b) What is the probability that you will make a Type I error given that the null hypothesis is false?

(a)
Given that the null hypothesis is true, a type 1 error (rejecting the null hypothesis by mistake) is a false positive with probability = α. i.e., probability of making a type 1 error = 0.01

(b)
The probability is 0 as the Type 1 error can only be made while the null hypothesis H_0 is true. Here it is given that null hypothesis is false, so the probability is 0

20. True/false: It is easier to reject the null hypothesis if the researcher uses a smaller alpha (α) level.
Solution : False. Lower the α level, harder it gets to reject the null hypothesis and we are more likely to face a type 2 error due to the low α levels.

21. True/false: You are more likely to make a Type I error when using a small sample than when using a large sample.
Solution : False. The type 1 error occurs when we reject the null hypothesis when it is actually true for a population (i.e., it is a false positive). Clearly type 1 error is independent of the sample size. Hence the statement is False.

22. True/false: You accept the alternative hypothesis when you reject the null hypothesis.
Solution : True. If the null hypothesis is rejected we accept the alternative hypothesis.

23. True/false: You do not accept the null hypothesis when you fail to reject it.

Solution : True. Failing to reject the null hypothesis does not mean that we accept the null hypothesis because the truth about the null hypothesis is unknown.

24. True/false: A researcher risks making a Type I error any time the null hypothesis is rejected.

Solution : False. Only when we reject a null hypothesis that is true, do we risk a Type 1 error.

CHAPTER 12

8. Participants threw darts at a target. In one condition, they used their preferred hand; in the other condition, they used their other hand. All subjects performed in both conditions (the order of conditions was counterbalanced). Their scores are shown below.

Preferred	Non-preferred
12	7
7	9
11	8
13	10
10	9

a. Which kind of t-test should be used?
b. Calculate the two-tailed t and p values using this t test.
c. Calculate the one-tailed t and p values using this t test.

Solution :

a. We use **t-test for correlated pairs or paired t test** as all subjects performed in both conditions

b.

```python
# Importing library
import scipy.stats as stats


preferred = [12, 7, 11, 13, 10]

non_preferred = [7, 9, 8, 10, 9]

# Performing the paired sample t-test
stats.ttest_rel(preferred, non_preferred)
```

```
Ttest_relResult(statistic=1.6903085094570331, pvalue=0.1662327518581254)
```

Two-tailed t = 1.69,   p-value = 0.1662

c. For a one tailed t test, the test value t remains the same but the p value is divided by 2

One-tailed t = 1.69,   p-value = 0.0831

9. Assume the data in the previous problem were collected using two different groups of subjects: One group used their preferred hand and the other group used their non-preferred hand. Analyse the data and compare the results to those for the previous problem

Solution :

```python
# Importing library
import scipy.stats as stats


preferred = [12, 7, 11, 13, 10]

non_preferred = [7, 9, 8, 10, 9]

# Performing the two tailed t-test
stats.ttest_ind(a=preferred, b=non_preferred, equal_var=True)
```

```
Ttest_indResult(statistic=1.7407765595569784, pvalue=0.1199027736779093)
```

Two-tailed t = 1.7408,   p-value = 0.12

One-tailed t = 1.7408,   p-value = 0.6

We can see that the independent t tests are not very significant but their p-values are closer together than the t-test for correlated pairs

11. In an experiment, participants were divided into 4 groups. There were 20 participants in each group, so the degrees of freedom (error) for this study was 80 - 4 = 76. Tukey's HSD test was performed on the data. (a) Calculate the p value for each pair based on the Q value given below. You will want to use the Studentized Range Calculator. (b) Which differences are significant at the .05 level?

Comparison of Groups	Q
A - B	3.4
A - C	3.8
A - D	4.3
B - C	1.7
B - D	3.9
C - D	3.7

Given :
Number of means = Number of groups = 4
Total number of observations in the study = 20*4 = 80
Degrees of freedom = 76
(a)
Using the Studentized Range Calculator
(https://onlinestatbook.com/2/calculators/studentized_range_dist.html)

$$Q = \frac{M_i - M_j}{\sqrt{MSE/n}}$$

Comparison of Groups	Q	p
A - B	3.4	**0.0849**
A - C	3.8	**0.0430**
A - D	4.3	**0.0167**
B - C	1.7	**0.6274**
B - D	3.9	**0.0359**
C - D	3.7	**0.0513**

(b)
The test is significant if p value is less than or equal to 0.05. From the given table we can see that, pairs **(A-C), (A-D), and (B-D) are significant** as their p value is less than 0.05 (i.e., our level of significance)

21. Do athletes or non-athletes calm down more when angry? Conduct a t test to see if the difference between groups in Control-In scores is statistically significant.

```
import statistics
import scipy.stats as stats

# read by default 1st sheet of an excel file
angry_moods = pd.read_excel('angry_moods.xls')

angry_moods["Sports_Desc"] = np.where(angry_moods["Sports"].isin([1]), 'Athletes', 'Non-Athletes')

ath = angry_moods[angry_moods['Sports_Desc'] == 'Athletes']["Control-In"].tolist()
non_ath = angry_moods[angry_moods['Sports_Desc'] == 'Non-Athletes']["Control-In"].tolist()

ath_var = statistics.variance(ath)
non_ath_var = statistics.variance(non_ath)
#after comparing the variances of athletes and non-athletes Control-In scores (20.476666666666667, 22.822931785195937)

stats.ttest_ind(a=ath, b=non_ath, equal_var=True)
```

```
Ttest_indResult(statistic=3.044310234609416, pvalue=0.003203009981395095)
```

Solution : p value = 0.00320 and it is less than the significance level. Hence the difference between the two groups is statistically significant.

22. Do people in general have a higher Anger-Out or Anger-In score? Conduct a t test on the difference between means of these two scores. Are these two means independent or dependent?

The two means are dependent as we are comparing means for two sets of scores that are computed on the same sample.

H_0 : The means of Anger Out and Anger In scores are equal.

H_1 : The means of the two scores are not equal

```
import statistics
import scipy.stats as stats

# read by default 1st sheet of an excel file
angry_moods = pd.read_excel('angry_moods.xls')

stats.ttest_rel(angry_moods["Anger-In"].tolist(), angry_moods["Anger-Out"].tolist())
```

```
Ttest_relResult(statistic=3.525860744424124, pvalue=0.0007145896341213851)
```

Since p = 0.00071 which is very much less than 0.5, we can conclude that the result is significant and people generally have a higher Anger-In score.

Solution : People generally have a higher Anger-In score. The two means are dependant

CHAPTER 13

5. Alan, while snooping around his grandmother's basement stumbled upon a shiny object protruding from under a stack of boxes . When he reached for the object a genie miraculously materialized and stated: "You have found my magic coin. If you flip this coin an infinite number of times you will notice that heads will show 60% of the time." Soon after the genie's declaration he vanished, never to be seen again. Alan, excited about his new magical discovery, approached his friend Ken and told him about what he had found. Ken was skeptical of his friend's story, however, he told Alan to flip the coin 100 times and to record how many flips resulted with heads.

(a) What is the probability that Alan will be able convince Ken that his coin has special powers by finding a p value below 0.05 (one tailed).
Use the Binomial Calculator (and some trial and error)

(b) If Ken told Alan to flip the coin only 20 times, what is the probability that Alan will not be able to convince Ken (by failing to reject the null hypothesis at the 0.05 level)?

(a)
H_0 = The coin does not have any special powers
H_1 = The coin has special powers

Probability of getting a head when the coin has no special powers = 0.5
Probability that head shows up 60% of the time. p = 0.6
Number of times the coin is tossed = 100
Given p value = 0.05,
$P(X > k \mid p = 0.5) < 0.05$
To convince Ken, Alan has to observe k heads that meet the above criteria

Using the Binomial calculator
(https://onlinestatbook.com/2/calculators/binomial_dist.html)

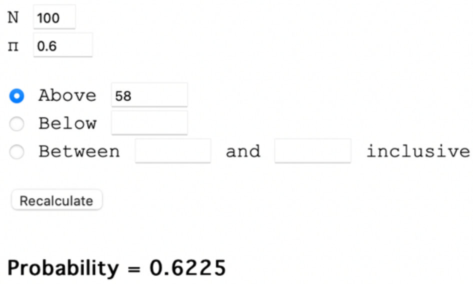
Since the probability of a fair coin is 0.5,
When $P(X > k \mid p = 0.5) < 0.05$
We can see that using the binomial calculator,

X	56	57	58	59
P(X > x)	0.0967	0.0667	0.0443	0.02844397

When k = 58 we observe value < 0.05

Probability that Alan will be able to convince Ken = P(X > 58 | p = 0.6)

Again, we can use the binomial calculator to determine the value of k.



k = 0.6225

Solution : Probability that Alan will be able to convince Ken = **62.25%**

(b)
Repeat the same steps from (a) but using number of coin toss n = 20

When P(X > k | p = 0.5) < 0.05
We can see that using the binomial calculator,

Using Binomial calculator, we get,

X	11	12	13	14	15
P(X > x)	0.2517	0.1316	0.05766	0.02069	0.0059

From the table we can see that,

Alan will not able to convince Ken, when at most 14 heads are observed in 20 tosses.

When the coin has special powers, p = 0.6

Probability that Alan will not be able to convince Ken = P(X $\leq$ 14 | p = 0.6)

Using Binomial calculator, we get, P(X $\leq$ 14 | p = 0.6) = 0.874401

Solution : Probability that Alan will not be able to convince Ken = **87.44%**