14 March Assignment

May 18, 2023

[ ]: Q1. Write a Python function that takes `in` two arrays of data `and` calculates the↳
     ↳F-value `for` a variance ratio
     test. The function should `return` the F-value `and` the corresponding p-value `for`↳
     ↳the test.

ANS -

[ ]: To calculate the F-value `for` a variance ratio test `in` Python, you can use the↳
     ↳f_oneway() function `from` `the` scipy.stats module.
     This function takes `in` two `or` more arrays of data `and` returns the F-value `and`↳
     ↳the corresponding p-value `for` the test.

```python
from scipy.stats import f_oneway

data1 = [1, 2, 3, 4, 5]
data2 = [6, 7, 8, 9, 10]

F_value, p_value = f_oneway(data1, data2)

print("F-value:", F_value)
print("p-value:", p_value)
```

F-value: 25.0
p-value: 0.0010528257933665396

[ ]:

[ ]:

[ ]: Q2. Given a significance level of `0.05` `and` the degrees of freedom `for` the↳
     ↳numerator `and` denominator of an
     F-distribution, write a Python function that returns the critical F-value `for` a↳
     ↳two-tailed test.

ANS -

[ ]: Python function to find the critical F-value `for` a two-tailed test given a↳
     ↳significance level of `0.05` `and` the degrees of freedom `for`

```
the numerator and denominator of an F-distribution:
```

```
[2]: from scipy.stats import f

     def critical_f_value(significance_level, dfn, dfd):
         return f.ppf(1 - significance_level / 2, dfn, dfd)
```

```
[ ]:
```

```
[ ]:
```

```
[ ]: Q3. Write a Python program that generates random samples from two normal␣
     ↪distributions with known

     variances and uses an F-test to determine if the variances are equal. The␣
     ↪program should output the F-
     value, degrees of freedom, and p-value for the test.
```

ANS -

```
[ ]: Python program that generates random samples from two normal distributions with␣
     ↪known variances and uses an F-test to determine if
     the variances are equal. The program outputs the F-value, degrees of freedom,␣
     ↪and p-value for the test.
```

```
[3]: import numpy as np
     from scipy.stats import f

     def f_test(var1, var2):
         f_value = np.var(var1) / np.var(var2)
         df1 = len(var1) - 1
         df2 = len(var2) - 1
         p_value = 1 - f.cdf(f_value, df1, df2)
         return f_value, df1, df2, p_value

     variance_1 = 5
     variance_2 = 10

     sample_1 = np.random.normal(0, np.sqrt(variance_1), size=100)
     sample_2 = np.random.normal(0, np.sqrt(variance_2), size=100)

     f_value, df1, df2, p_value = f_test(sample_1, sample_2)

     print(f"F-value: {f_value:.3f}")
     print(f"Degrees of freedom: {df1}, {df2}")
     print(f"P-value: {p_value:.3f}")
```

```
F-value: 0.424
Degrees of freedom: 99, 99
P-value: 1.000
```

[ ]:

[ ]:

[ ]: Q4.The variances of two populations are known to be 10 and 15. A sample of 12␣
    ↪observations is taken from
    each population. Conduct an F-test at the 5% significance level to determine if␣
    ↪the variances are
    significantly different.

ANS -

[ ]: To test if the variances of two populations are significantly different, we can␣
    ↪use an F-test. The null hypothesis is that the
    variances are equal. The alternative hypothesis is that the variances are not␣
    ↪equal.

    In this case, we have two populations with known variances of 10 and 15. We␣
    ↪take a sample of 12 observations from each population.
    To conduct an F-test at the 5% significance level to determine if the variances␣
    ↪are significantly different, we can use the following
    formula:

[ ]: F = s1^2 / s2^2

[ ]: where s1^2 is the sample variance of the first population and s2^2 is the␣
    ↪sample variance of the second population

[ ]: F = s1^2 / s2^2 = 10 / 15 = 0.6667

[ ]: The degrees of freedom for this test are (n1 - 1) and (n2 - 1), where n1 and n2␣
    ↪are the sample sizes for each population. In this
    case the degrees of freedom are 11 and 11.

    Using an F-distribution table or calculator with 11 and 11 degrees of freedom␣
    ↪at a significance level of 5%, we find that the
    critical value is approximately 3.11.

    Since our calculated F-value (0.6667) is less than the critical value (3.11),␣
    ↪we fail to reject the null hypothesis that the
    variances are equal at a significance level of 5%. Therefore, we can conclude␣
    ↪that there is not enough evidence to suggest that the
    variances are significantly different at a significance level of 5%.

[ ]:

[ ]:

Q5. A manufacturer claims that the variance of the diameter of a certain product is 0.005. A sample of 25 products is taken, and the sample variance is found to be 0.006. Conduct an F-test at the 1% significance level to determine if the claim is justified.

ANS -

```
[ ]: The F-test is used to test the null hypothesis that two normal populations have␣
     ↪the same variance. In this case, the manufacturer
     claims that the variance of the diameter of a certain product is 0.005. A␣
     ↪sample of 25 products is taken, and the sample variance is
     found to be 0.006. The null hypothesis is that the variance of the diameter of␣
     ↪a certain product is 0.005. The alternative hypothesis
     is that the variance of the diameter of a certain product is greater than 0.005.

     The test statistic for this F-test is calculated as follows:
```

```
[ ]: F = (sample variance) / (hypothesized variance)
```

```
[6]: F = 0.006 / 0.005
```

```
[6]: 1.2
```

```
[ ]: The degrees of freedom for this F-test are (n1 - 1) and (n2 - 1), where n1 and␣
     ↪n2 are the sample sizes of the two populations being
     compared. In this case, n1 = n2 = 25, so the degrees of freedom are 24 and 24.

     At the 1% significance level, the critical value for an F-distribution with␣
     ↪degrees of freedom (24,24) is 2.491. Since our
     calculated F-value (1.2) is less than the critical value (2.49), we fail to␣
     ↪reject the null hypothesis that the variance of the
     diameter of a certain product is 0.005 at the 1% significance level.

     Therefore, we can conclude that there is not enough evidence to suggest that␣
     ↪the manufacturer's claim is not justified.
```

[ ]:

[ ]:

```
[ ]: Q6. Write a Python function that takes in the degrees of freedom for the␣
     ↪numerator and denominator of an
     F-distribution and calculates the mean and variance of the distribution. The␣
     ↪function should return the
```

```
mean and variance as a tuple.
```

ANS -

```python
[7]: import numpy as np

     def f_distribution(dfn, dfd):
         mean = dfd / (dfd - 2)
         variance = (2 * dfd ** 2 * (dfn + dfd - 2)) / (dfn * (dfd - 2) ** 2 * (dfd↵
     ↪- 4))
         return mean, variance
```

```python
[8]: mean, variance = f_distribution(3, 5)
     print("Mean:", mean)
     print("Variance:", variance)
```

```
Mean: 1.6666666666666667
Variance: 11.11111111111111
```

[ ]:

[ ]:

```
[ ]: Q7. A random sample of 10 measurements is taken from a normal population with↵
     ↪unknown variance. The
     sample variance is found to be 25. Another random sample of 15 measurements is↵
     ↪taken from another
     normal population with unknown variance, and the sample variance is found to be↵
     ↪20. Conduct an F-test
     at the 10% significance level to determine if the variances are significantly↵
     ↪different.
```

ANS -

```
[ ]: An F-test is used to test the null hypothesis that two normal populations have↵
     ↪the same variance. The F-test is used when the
     difference between two variances needs to be significantly assessed, i.e., when↵
     ↪determining whether or not two samples can be taken
     as representative of the normal population with the same variance1.

     In your case, you have two samples with sample variances 25 and 20 respectively.
     ↪ You want to conduct an F-test at the 10%
     significance level to determine if the variances are significantly different.↵
     ↪The null hypothesis is that the variances are equal
     and the alternative hypothesis is that they are not equal.

     The test statistic for this test is given by:
```

```
F = s1^2 / s2^2

where s1^2 and s2^2 are the sample variances of the two samples.

The degrees of freedom for this test are given by:

df1 = n1 - 1 df2 = n2 - 1

where n1 and n2 are the sample sizes of the two samples.

Using these values, we can calculate the p-value for this test. If the p-value␣
  ↪is less than or equal to 0.10, we reject the null
hypothesis and conclude that the variances are significantly different
```

[ ]:

[ ]:

[ ]:
```
Q8. The following data represent the waiting times in minutes at two different␣
  ↪restaurants on a Saturday
night: Restaurant A: 24, 25, 28, 23, 22, 20, 27; Restaurant B: 31, 33, 35, 30,␣
  ↪32, 36. Conduct an F-test at the 5%
significance level to determine if the variances are significantly different.
```

ANS -

[ ]:
```
To determine if the variances are significantly different, we can conduct an␣
  ↪F-test at the 5% significance level. The formula for the
test statistic is F = s12/s22 where s1^2 is the variance of the first sample␣
  ↪and s2^2 is the variance of the second sample. The null
hypothesis is that the variances are equal and the alternative hypothesis is␣
  ↪that they are not equal. If the calculated F-value is
greater than the critical value from an F-distribution table at a 5%␣
  ↪significance level, then we reject the null hypothesis and
conclude that the variances are significantly different12.

In this case, we have two samples: Restaurant A and Restaurant B. The variance␣
  ↪of Restaurant A is 8.2857 and the variance of
Restaurant B is 4.8. The calculated F-value is 1.7266 which is less than the␣
  ↪critical value of 5.1434 from an F-distribution table
at a 5% significance level1. Therefore, we fail to reject the null hypothesis␣
  ↪and conclude that there is not enough evidence to
suggest that the variances are significantly different.
```

[ ]:

[ ]:

[ ]: Q9. The following data represent the test scores of two groups of students:␣
↪Group A: 80, 85, 90, 92, 87, 83;
Group B: 75, 78, 82, 79, 81, 84. Conduct an F-test at the 1% significance level␣
↪to determine if the variances
are significantly different.

ANS -

[ ]: To conduct an F-test at the 1% significance level to determine if the variances␣
↪are significantly different, we can use the formula
for the test statistic which is F = s12/s22 where s1^2 and s2^2 are the sample␣
↪variances of group A and group B respectively.
The null hypothesis is that the variances are equal and the alternative␣
↪hypothesis is that they are not equal.

The degrees of freedom for the numerator is n1-1 and for the denominator is␣
↪n2-1 where n1 and n2 are the sample sizes of group A
and group B respectively.

At 1% significance level, we can use F0.01 with degrees of freedom (n1-1) and␣
↪(n2-1) to determine if the null hypothesis should be
rejected or not. If F > F0.01, we reject the null hypothesis and conclude that␣
↪the variances are significantly different. Otherwise,
we fail to reject the null hypothesis.

Using this formula and data provided in your question, we can calculate F as␣
↪follows:

s1^2 = 30.5 s2^2 = 6.8 n1 = 6 n2 = 6

F = s12/s22 = 30.5/6.8 = 4.48

Using F0.01 with degrees of freedom (5) and (5), we get F0.01 = 6.61.

Since F < F0.01, we fail to reject the null hypothesis and conclude that there␣
↪is not enough evidence to suggest that the variances
are significantly different at 1% significance level.

[ ]: