

Hypothesis Testing: Tests Involving Variances

Lecture 7.1

1 Outline

- Ratio of Population Variance Tests
- Association vs. Causation

1.1 Hypothesis Test: Ratio of Population Variances

Unlike the comparisons done in this chapter, comparing two population variances are usually based on the ratio of the variances (σ_1^2/σ_2^2).

Note

When $\sigma_1^2/\sigma_2^2 = 1$, the variances are equal. However, this ratio can not be negative, which is not consistent with the distributions we have considered (standard normal distribution and t-distribution)

1.2 F-distribution

The F-distribution is a distribution that has the following characteristics:

- Defined in the range $0 < X < \infty$
- Defined by two degrees of freedom: the numerator degrees of freedom, df_{num} or df_1 , and the denominator degrees of freedom, df_{den} or df_2 .

In R, the following functions can be used to generate the properties of the *central* F-distribution.

- `df(x,df1,df2)`: PDF
- `pf(x,df1,df2)`: CDF

- `qf(p,df1,df2)`: Calculates F value for a given probability
- `rf(n,df1,df2)`: Generates n samples from the F-distribution

1.3 Hypothesis Test: Ratio of Population Variances

If we are interested in testing the null hypothesis $H_0 : \sigma_1/\sigma_2 = 1$, the alternative hypothesis can be one of the following:

- One-sided hypotheses: $H_a : \sigma_1/\sigma_2 < 1$ or $H_a : \sigma_1/\sigma_2 > 1$.
- Two-sided hypotheses: $H_a : \sigma_1/\sigma_2 \neq 1$

The problem dictates which alternative is appropriate.

1.4 Test Statistic

Assuming the following assumptions hold: s_1^2 and s_2^2 are computed from independent samples of size n_1 and n_2 , respectively drawn from two normally distributed populations.

The quantity $\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$ follows the F-distribution with degrees of freedom $(df_1, df_2) = (n_1 - 1, n_2 - 1)$. Under the null hypothesis, $\sigma_1^2/\sigma_2^2 = 1$. Hence, our test statistic is expressed as:

$$F = S_1^2/S_2^2 \sim F(n_1 - 1, n_2 - 1)$$

1.5 Rejection Region

We will use the F-distribution to define the rejection region and the corresponding p-values.

For one-sided tests, the rejection region is defined by the range:

$$F < F_{1-\alpha}(n_1 - 1, n_2 - 1)$$

As much as possible, we set the group with the expected larger variance to be in the numerator to avoid complications in calculating the rejection region.

For two-sided tests, the rejection region is defined by the ranges:

$$F < F_{\alpha/2}(n_1 - 1, n_2 - 1) \text{ or } F > F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$$

i Note

The p-values can be calculated using statistical software like R and follows the same rejection rules as the previous tests (reject H_0 when $p < \alpha$)

1.6 Hypothesis Testing: R

The function `var.test()` provides the test statistic and the corresponding p-values for testing the ratio between population variances. Typically, we are interested in the null hypothesis of equality or $H_0 : \sigma_1^2 / \sigma_2^2 = 1$.

The function can be implemented in two ways. If the data is stored in a data frame `df`,

i Note

`var.test(formula,data=df,ratio=1,alternative)` can be used if the groups are defined by a column in `df` and the response variable is defined by another column in `df`. The formula can be expressed as `response~group`. `alternative` can take on the following values: `"two.sided"`, `"less"`, or `"greater"`

i Note

`var.test(x,y,ratio=1,alternative)` can be used if the responses/measurements from each group are found in different columns in `df`. Suppose the measurements are in the columns `group1` and `group2`. The groups `x` and `y` can be expressed as `x=df$group1,y=df$group2`. `alternative` can take on the following values: `"two.sided"`, `"less"`, or `"greater"`.

i Note

To determine the rejection region for the F statistic, we can use the `qf()` function.

1.7 Example

Consider the sleep health data set from `SleepHealthData.csv`.

```
sleep <- read.csv("SleepHealthData.csv")
```

1.7.1 Question

Perform a hypothesis test to determine if there is evidence of unequal variances between the average sleep duration of males and females. Assume equal variances and a significance level of 0.05.

1.7.2 Answer

The parameter of interest is the ratio of the variances in sleep duration between females and males. The corresponding hypotheses are:

$$H_0 : \sigma_f^2 / \sigma_m^2 = 1; H_a : \sigma_f^2 / \sigma_m^2 \neq 1$$

We can now use the `var.test()` function to solve for the test statistic and p-value. Note that the groups are defined by the `gender` variable and the response is defined by the `sleep_duration` variable.

```
fctest <- var.test(sleep_duration~gender,  
  data=sleep,  
  ratio=1,  
  alternative="two.sided"  
)
```

```
fctest
```

F test to compare two variances

```
data:  sleep_duration by gender  
F = 1.6095, num df = 184, denom df = 188, p-value = 0.001242  
alternative hypothesis: true ratio of variances is not equal to 1  
95 percent confidence interval:  
 1.206633 2.147875  
sample estimates:  
ratio of variances  
 1.60951
```

The test statistic value is 1.6095104 corresponding to a p-value of 0.0012421. The rejection region can be calculated using the `qf()` function.

```
alpha<- 0.05  
lower_bound <- qf(alpha/2,df1=184,df2=188)  
lower_bound
```

```
[1] 0.74935
```

```
upper_bound<- qf(1-alpha/2,df1=184,df2=188)  
upper_bound
```

```
[1] 1.333886
```

The rejection region is defined as: $F < 0.74935$ or $F > 1.3338859$. The test statistic is in this rejection region, hence, we reject the null hypothesis. There is sufficient evidence that the ratio of the two population variances is not equal to one, implying that the population variances are unequal.

i Note

The same decision and conclusion is reached when the p-value is used.

1.8 Exercise

Consider the sleep health data set from `SleepHealthData.csv`.

1.8.1 Question

Recall that you were asked to assume equal variances in testing for a difference between the average heart rate of males and females in the previous lecture. Perform a hypothesis test to determine if there is sufficient evidence to claim that the population variances of the heart rate of males and females are different. Use a significance level of 0.001.

1.8.2 Answer

The parameter of interest is the ratio of the variances between females and males. The corresponding hypotheses are:

$$H_0 : \sigma_f^2 / \sigma_m^2 = 1; H_0 : \sigma_f^2 / \sigma_m^2 \neq 1$$

We can now use the `var.test()` function to solve for the test statistic and p-value. Note that the groups are defined by the `gender` variable and the response is defined by the `heart_rate` variable.

```
fctest <- var.test(heart_rate~gender,
  data=sleep,
  ratio=1,
  alternative="two.sided"
)

fctest
```

F test to compare two variances

```
data: heart_rate by gender
F = 1.4975, num df = 184, denom df = 188, p-value = 0.006088
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 1.122678 1.998430
sample estimates:
ratio of variances
 1.497524
```

The test statistic value is 1.4975238 corresponding to a p-value of 0.0060876. The rejection region can be calculated using the `qf()` function.

```
alpha<- 0.001
lower_bound <- qf(alpha/2,df1=184,df2=188)
lower_bound
```

```
[1] 0.6150367
```

```
upper_bound<- qf(1-alpha/2,df1=184,df2=188)
upper_bound
```

```
[1] 1.624297
```

The rejection region is defined as: $F < 0.6150367$ or $F > 1.6242974$. The test statistic is NOT in this rejection region, hence, we fail to reject the null hypothesis. There is insufficient evidence that the ratio of the two population variances is not equal to one.

Note

The same decision and conclusion is reached when the p-value is used.

2 Association vs. Causation

2.1 Association vs. Causation

When comparing two groups, it is important to be mindful of how to interpret the conclusion. Suppose we are comparing two groups. When we reject the null hypothesis and conclude that there is sufficient evidence supporting the alternative hypothesis, there is an implied *association* between the groups and the response variable.

Warning

However, this does not necessarily mean that the group variable was the reason why there was a significant difference. There might be other variables in play, also known as confounding variables, that could be associated with both groups and response variable and contributed to this significant differences. The hypothesis tests discussed do not account for this, hence causation cannot be claimed.

In experiments, the effects of these confounding variables can be mitigated by performing random assignment of participants to the treatment groups. In observational and retrospective studies, the process to prove causation is much harder. There is a whole field that addresses this problem called *causal inference*, which we will not discuss in this class.