Hypotheses testing about two Population Variances

Session 24

10/05/2021

Hypotheses testing about two Population Variances

- Suppose a manufacturing plant made two batches of an item produced on two different machines, or produced items on two different shifts. It might be of interest to management to compare the variances from two batches or two machines to determine whether there is more variability in one than another.
- In quality control, analysts often examine a measure of variability. Variance is sometimes used as a measure of the risk of a stock in the stock market. The greater the variance, the greater the risk. By using techniques discussed here, a financial researcher could determine whether the variances (or risk) of two stocks are the same.
- In testing hypotheses about two variances, the sample variances are used.

F distribution

- If X and Y are two independent chi-square variates with ν_1 and ν_2 d.f. respectively. Then F statistic is defined by

$$F = \frac{X/\nu_1}{Y/\nu_2}$$

- In other words, F is defined as the ratio of two independent chi-square variates divided by their respective degrees of freedom and it follows Snedecor's F distribution with (ν_1, ν_2) d.f. with probability function given by

$$f(F) = \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2}}{B(\nu_1, \nu_2)} \frac{F^{\frac{\nu_1}{2}-1}}{\left(1+\frac{\nu_1}{\nu_2}F\right)^{(\nu_1+\nu_2)/2}}, 0 \leq F < \infty$$

Remarks

- The sampling distribution of F statistic does not involve any population parameters and depends only on the degrees of freedom ν_1 and ν_2.
- A statistic F following Snedecor's F distribution with (ν1 ,ν2) d.f. will be denoted as $F \sim F(\nu_1, \nu_2)$.

F test for two population variances

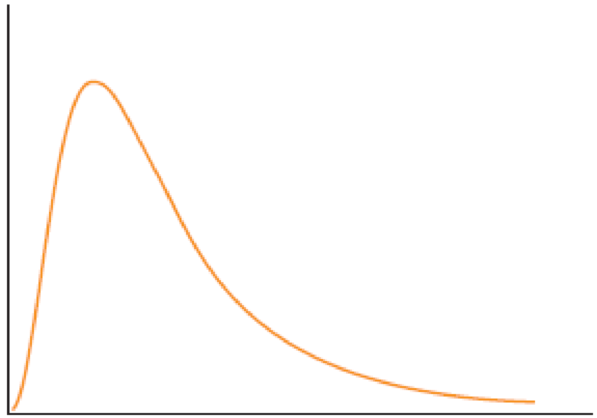
- An assumption underlying F distribution is that the populations from which the samples are drawn are Normally distributed.
- X $\sim N(0,1)$ and Y$\sim N(0,1)$ are independent, $X^2 \sim \chi^2{}_{(1)}$ and $Y^2 \sim \chi^2{}_{(1)}$ are also independent. Hence by definition of F statistic,

$$\frac{X^2/1}{Y^2/1} \sim F_{(1,1)} \Rightarrow \frac{X^2}{Y^2} \sim F_{(1,1)}$$

F distribution

FIGURE 10.14

An F Distribution for $v_1 = 6$, $v_2 = 30$



F test for two population variances

- Test statistic is

$$F = \frac{S_1^2}{S_2^2}$$

$$df_{numerator} = \nu_1 = n_1 - 1$$

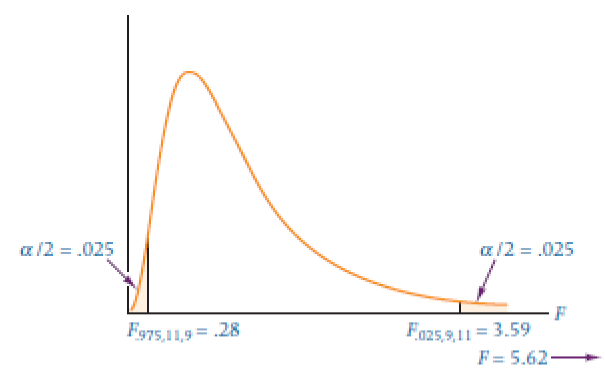
$$df_{denominator} = \nu_2 = n_2 - 1$$

Where ${S_1}^2$ and ${S_2}^2$ are the variances of first and second samples respectively.

F table and formula for determining the critical value for the lower tail of F

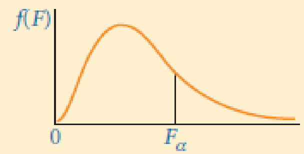
- F distribution is non symmetric which can be a problem for when we are conducting a two tailed test and want to determine the critical value for the lower tail.
- Table of F distribution contains only F values for the upper tail.
- This dilemma can be solved by using the following formula

$$F_{1-\propto,\nu_2,\nu_1} = \frac{1}{F_{\alpha,\nu_1,\nu_2}}$$



F distribution table[P(F>$F_{\alpha,v1,v2}$)=α]

Percentage Points of the *F* Distribution



$\alpha = 0.025$

Numerator Degrees of Freedom (v_1); *ominator Degrees of Freedom* (v_2)

v_2 \ v_1	1	2	3	4	5	6	7	8	9
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84

$F_{.025,9,11}$

Test-9:Testing Hypotheses about two population Variances (Populations are normally distributed)

	Left tail test	Right tail Test	Two tail Test
Hypotheses	$H_0:\sigma_1^2 = \sigma_2^2$ $H_1:\sigma_1^2 < \sigma_2^2$	$H_0:\sigma_1^2 = \sigma_2^2$ $H_1:\sigma_1^2 > \sigma_2^2$	$H_0:\sigma_1^2 = \sigma_2^2$ $H_1:\sigma_1^2$ ‡ σ_2^2
Test Statistic	$F = \frac{S_1^2}{S_2^2}$	$df_{numerator} = \nu_1 = n_1 - 1$ $df_{denominator} = \nu_2 = n_2 - 1$	
Rejection Rule	Reject H_0 if $F \leq F_{1-\propto,\nu_2,\nu_1}$	Reject H_0 if $F \geq F_{\propto,\nu_1,\nu_2}$	Reject H_0 if $F \leq F_{(1-\propto/2),\nu_2,\nu_1}$ or $F \geq F_{\propto/2,\nu_1,\nu_2}$

Example

- An investment advisor believes that the variances of stock prices of manufacturing companies and information technology companies are the same. To verify this claim, variances of stock prices from these two sectors are collected and are given below.

Group	Sample size	Variance estimated from sample
Manufacturing firms	80	42
IT firms	52	36

- Conduct an appropriate test at $\alpha=0.1$, to check whether the variances of stock prices in two industry sectors are equal or not?

Solution

- Consider Population I: Manufacturing firms and population II as IT firms.
- Given S_1^2=42 , S_2^2=36, n_1=80 and n_2=52.
- The null and alternative hypothesis are $H_0: \sigma_1^2 = \sigma_2^2$ against $H_1: \sigma_1^2 \neq \sigma_2^2$ where σ_1^2 and σ_2^2 are the population variances in stock prices from different firms.
- Given α=0.1
- The test statistic is $F = \frac{S_1^2}{S_2^2}$

with ν_1, ν_2 d.f where $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$.

- The calculated $F = \frac{S_1^2}{S_2^2} = \frac{42}{36} = 1.167$
- Rejection rule is reject H_0 if $F \leq F_{(1-\propto/2),\nu_2,\nu_1}$

  or $F \geq F_{\propto/2,\nu_1,\nu_2}$

- Now $F_{\alpha/2,\nu_1,\nu_2}$=$F_{0.05,79,51}$ $=$1.53 (approximately) and

  $F_{(1-\propto/2),\nu_2,\nu_1}$= $F_{(0.05),\nu_2,\nu_1}$ =$\frac{1}{1.53}$ $=$0.654
- The same can be obtained in Excel by FINV(0.05, 79,51) and $\frac{1}{\text{FINV(0.05, 79,51)}}$.

- Here 0.654 < 1.167<1.53. The test statistic value lies in the non-rejection region and hence do not reject H_0.
- Conclusion : There is enough statistical evidence to conclude that the variances of stock prices in two industry sectors are equal.

Example

- An output obtained in Excel after conducting an F test for equality of population variances are given below. Comment the output by framing both one tail and two tail hypothesis.

EXCEL Output

F- Test Two-Sample for Variances

	Machine 1	*Machine 2*
Mean	22.04	21.975
Variance	0.11378	0.02023
Observations	10	12
df	9	11
F	5.62	
P (F<=f) one-tail	0.0047	
F Critical one-tail	3.59	