Hypothesis Tests

Tests of Variance

DASC 512

Overview

- One-Sample Test of Variance (χ^2 test)
- Two-Sample Test of Variance (F test)

χ^2 Test

One-Sample Test of Variance

Assumptions

- Type of data – Quantitative
- Randomization – Data gathered randomly (iid)
- Population distribution – Assumes population data distributed $N(\mu, \sigma^2)$
- Sample size – May be used for any sample size, but normality is extra important for small samples

Hypotheses

Null hypothesis: The variance is equal to some known value

$$H_0: \sigma^2 = \sigma_0^2$$

Alternative hypothesis: The variance is not equal to the known value in some way (depends on the research hypothesis)

$$H_a: \sigma^2 < \sigma_0^2, \qquad H_a: \sigma^2 > \sigma_0^2, \qquad H_a: \sigma^2 \neq \sigma_0^2$$

χ^2 Distribution Review

Recall that the sum of k squared standard normal random variables is distributed $\chi^2(k)$

$$\sum_{i=1}^{k} Z^2 = X \sim \chi^2(k)$$

χ^2 Distribution Review

Let $Z_1, \dots, Z_k$ be distributed *N*(0,1). Let $X_i = \sigma Z_i$.

$$\sum_{i=1}^{k} (Z_i - \bar{Z})^2 \sim \chi_{k-1}^2$$

$$\frac{\sum_{i=1}^{k} (X_i - \bar{X})^2}{\sigma^2} \sim \chi_{k-1}^2$$

$$\frac{\sum_{i=1}^{k} (X_i - \bar{X})^2}{\sigma^2} \times \frac{n-1}{n-1} = \frac{(n-1)s^2}{\sigma^2} \sim \chi_{k-1}^2$$

Test Statistic

Assuming the population is distributed $N(\mu, \sigma^2)$, the test statistic

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(k = n-1)$$

Conclusion

If $H_a\!: \sigma^2 < \sigma_0^2$,

$$\chi^2 < \chi^2_{PPF}(\alpha, n-1)$$

If $H_a\!: \sigma^2 > \sigma_0^2$,

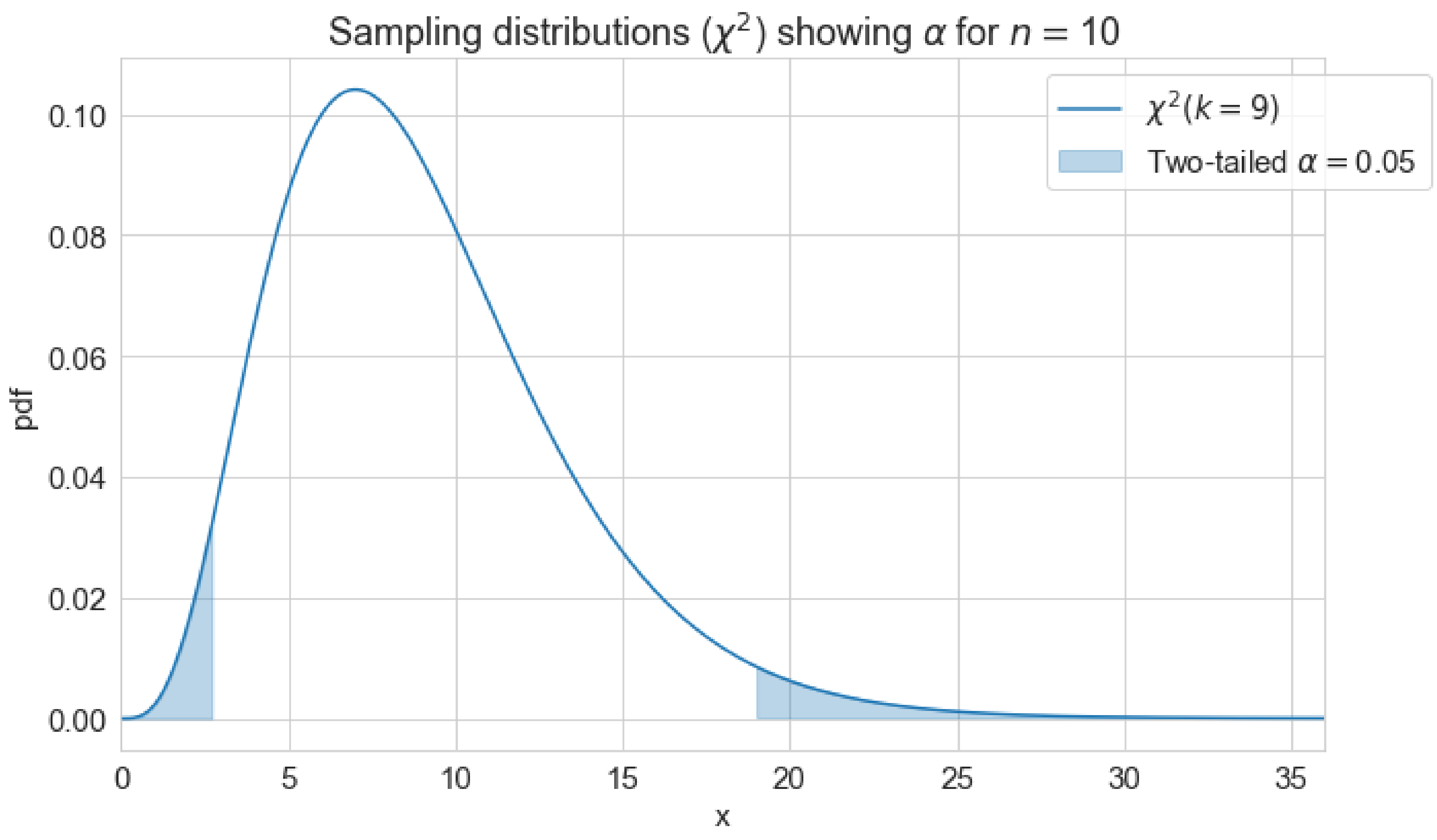
$$\chi^2 > \chi^2_{ISF}(\alpha, n-1)$$

If $H_a\!: \sigma^2 \neq \sigma_0^2$,

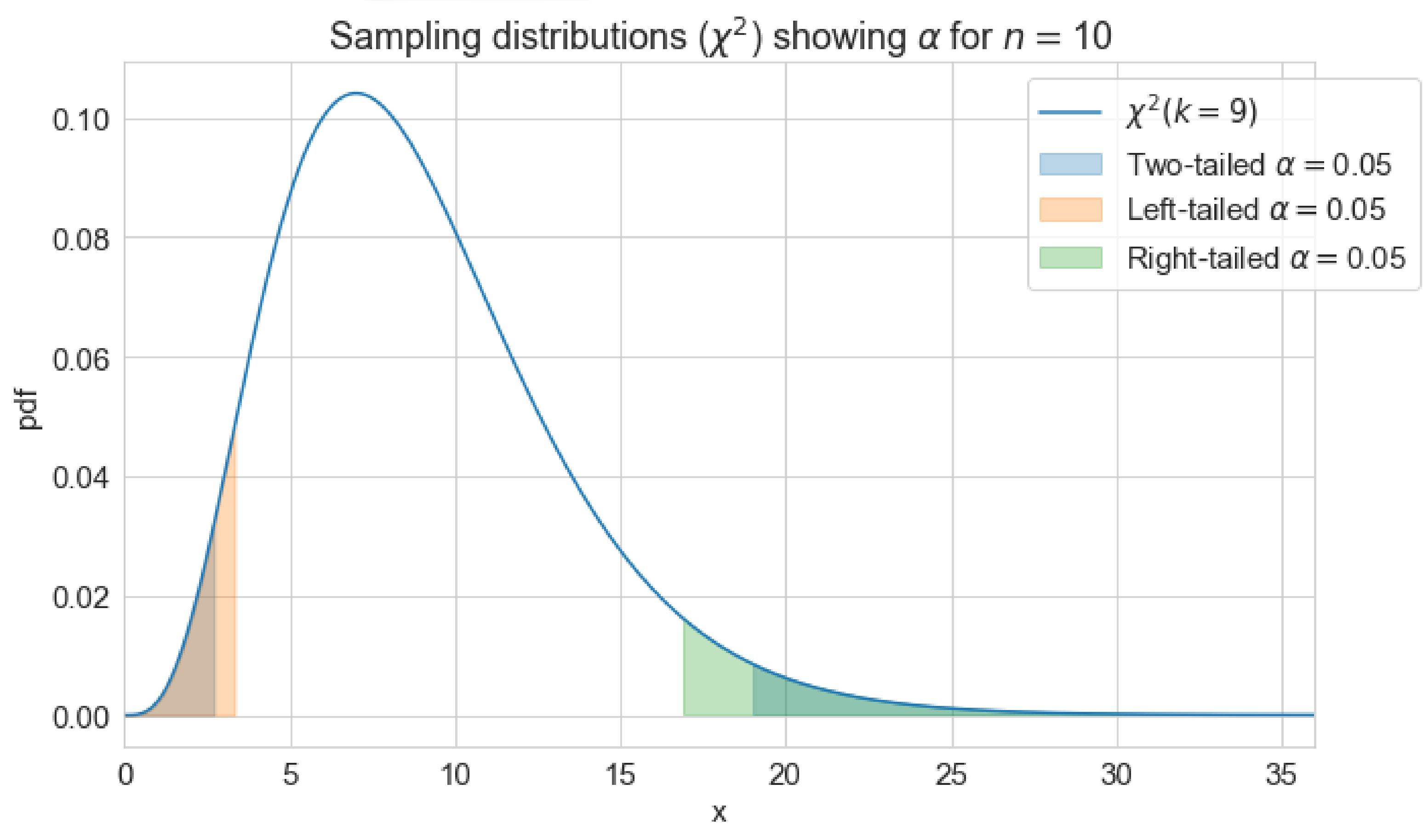
$$\chi^2 > \chi^2_{ISF}\left(\frac{\alpha}{2}, n-1\right), \qquad \text{or}$$

$$\chi^2 < \chi^2_{PPF}\left(\frac{\alpha}{2}, n-1\right)$$

Conclusion


Sampling distributions (χ^2) showing α for $n = 10$
$\chi^2(k = 9)$
Two-tailed $\alpha = 0.05$
pdf
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30
35
x

Conclusion


Sampling distributions (χ^2) showing α for $n = 10$
$\chi^2(k = 9)$
Two-tailed $\alpha = 0.05$
Left-tailed $\alpha = 0.05$
Right-tailed $\alpha = 0.05$
pdf
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
25
30
35
x

P-value

Presuming that H_0 is true, the probability of observing a test statistic (χ^2_{n-1}) more extreme than the one observed (χ^2) is:

If $H_a: \sigma^2 < \sigma_0^2$,

$$p = P(\chi^2_{n-1} < \chi^2) = \chi^2_{CDF,n-1}(\chi^2)$$

If $H_a: \sigma^2 > \sigma_0^2$,

$$p = P(\chi^2_{n-1} > \chi^2) = \chi^2_{SF,n-1}(\chi^2)$$

If $H_a: \sigma^2 \neq \sigma_0^2$,

$$p = 2 \times \min\left(P(\chi^2_{n-1} > \chi^2), P(\chi^2_{n-1} < \chi^2)\right) = 2 \times \min\left(\chi^2_{CDF,n-1}(\chi^2), \chi^2_{SF,n-1}(\chi^2)\right)$$

Python Functions

Bad news: The χ^2 test is not included in Python.

Good news: I’ve coded it up for you – see the notebook for this lesson.

- chi2_1samp_var_stats if you have summary statistics
- chi2_1samp_var_data if you have raw data

Example

A specific filling machine should fill containers with a standard deviation of no more than 1.3 mL. It is assumed that the filling of containers follows a normal distribution. Every hour, a random sample of containers is selected and the variability in filling is tested.

Over the last hour, 31 containers were sampled. The sample standard deviation was 1.4 mL.

Is there enough evidence to say that the variability in container filling is outside of spec? Assume significance $\alpha = 0.05$.

Example

- Hypotheses: $H_0\colon \sigma^2 = 1.3^2,\ \ H_a\colon \sigma^2 > 1.3^2$
- Test Statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(30)\left(1.4^2\right)}{1.3^2} = 34.8$
- Rejection Region: $\chi^2 > \chi^2_{PPF}(1-\alpha, n-1) = 43.8$
- P-value: $p = 0.25$

We fail to reject the null hypothesis. There is insufficient evidence to conclude that the filler is performing out of spec.

Critical observation: $s > \sqrt{\frac{\chi^2_{PPF}(1-\alpha, n-1)\sigma_0^2}{n-1}} = \sqrt{\frac{(43.8)(1.3)^2}{30}} = 1.57$

F Test

Two-Sample Test of Variance

Assumptions

- Type of data – Quantitative
- Randomization – Data gathered randomly (iid)
- Population distribution – Assumes population data distributed normally
- Sample size – May be used for any sample size, but normality is extra important for small samples

Hypotheses

Null hypothesis: The variances of the samples are equal

$$H_0: \sigma_1^2 = \sigma_2^2$$

Alternative hypothesis: The variances of the samples differ in some way (depends on the research hypothesis)

$$H_a: \sigma_1^2 < \sigma_2^2, \qquad H_a: \sigma_1^2 > \sigma_2^2, \qquad H_a: \sigma_1^2 \neq \sigma_2^2$$

F Distribution Review

Recall that the sum of k squared standard normal random variables is distributed $\chi^2(k)$

$$\sum_{i=1}^{k} Z^2 = X \sim \chi^2(k)$$

Recall that the ratio of two χ^2 variables divided by their degrees of freedom is distributed $F(r_1, r_2)$

$$\left(\frac{\chi^2_{r_1}}{r_1}\right) \Big/ \left(\frac{\chi^2_{r_2}}{r_2}\right) = F(r_1, r_2)$$

F Distribution Review

Assuming the null hypothesis $(\sigma_1 = \sigma_2 = \sigma)$

$$F = \frac{\chi_1^2}{\chi_2^2} = \frac{\dfrac{\left(\dfrac{(n_1 - 1)s_1^2}{\sigma^2}\right)}{n_1 - 1}}{\dfrac{\left(\dfrac{(n_2 - 1)s_2^2}{\sigma^2}\right)}{n_2 - 1}}$$

F Distribution Review

Assuming the null hypothesis $(\sigma_1 = \sigma_2 = \sigma)$

$$F = \frac{\chi_1^2}{\chi_2^2} = \frac{\dfrac{\left(\dfrac{\cancel{(n_1 - 1)} s_1^2}{\cancel{\sigma^2}}\right)}{\cancel{n_1 - 1}}}{\dfrac{\left(\dfrac{\cancel{(n_2 - 1)} s_2^2}{\cancel{\sigma^2}}\right)}{\cancel{n_2 - 1}}} = \frac{s_1^2}{s_2^2}$$

Test Statistic

Assuming the population is distributed $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$,

$$F = \frac{s_1^2}{s_2^2} \sim F(n_1 - 1, n_2 - 1)$$

Conclusion

If $H_a: \sigma_1^2 < \sigma_2^2$,

$$F < F_{PPF}(\alpha, r_1 = n_1 - 1, r_2 = n_2 - 1)$$

If $H_a: \sigma_1^2 > \sigma_2^2$,

$$F > F_{ISF}(\alpha, r_1 = n_1 - 1, r_2 = n_2 - 1)$$

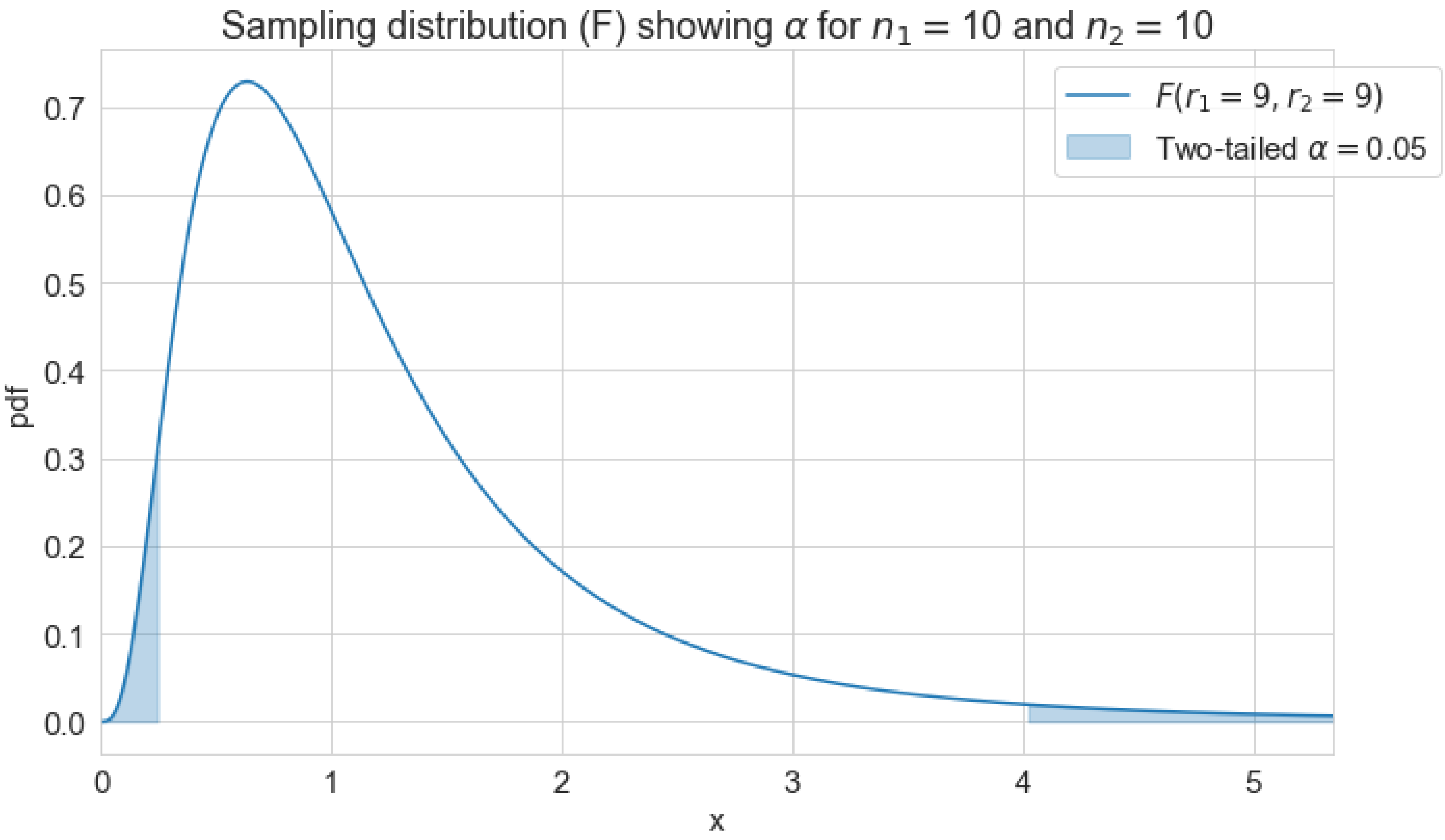
If $H_a: \sigma_1^2 \neq \sigma_2^2$ and $s_1 > s_2$,

$$F > F_{ISF}\left(\frac{\alpha}{2}, r_1 = n_1 - 1, r_2 = n_2 - 1\right)$$

If $H_a: \sigma_1^2 \neq \sigma_2^2$ and $s_2 > s_1$,

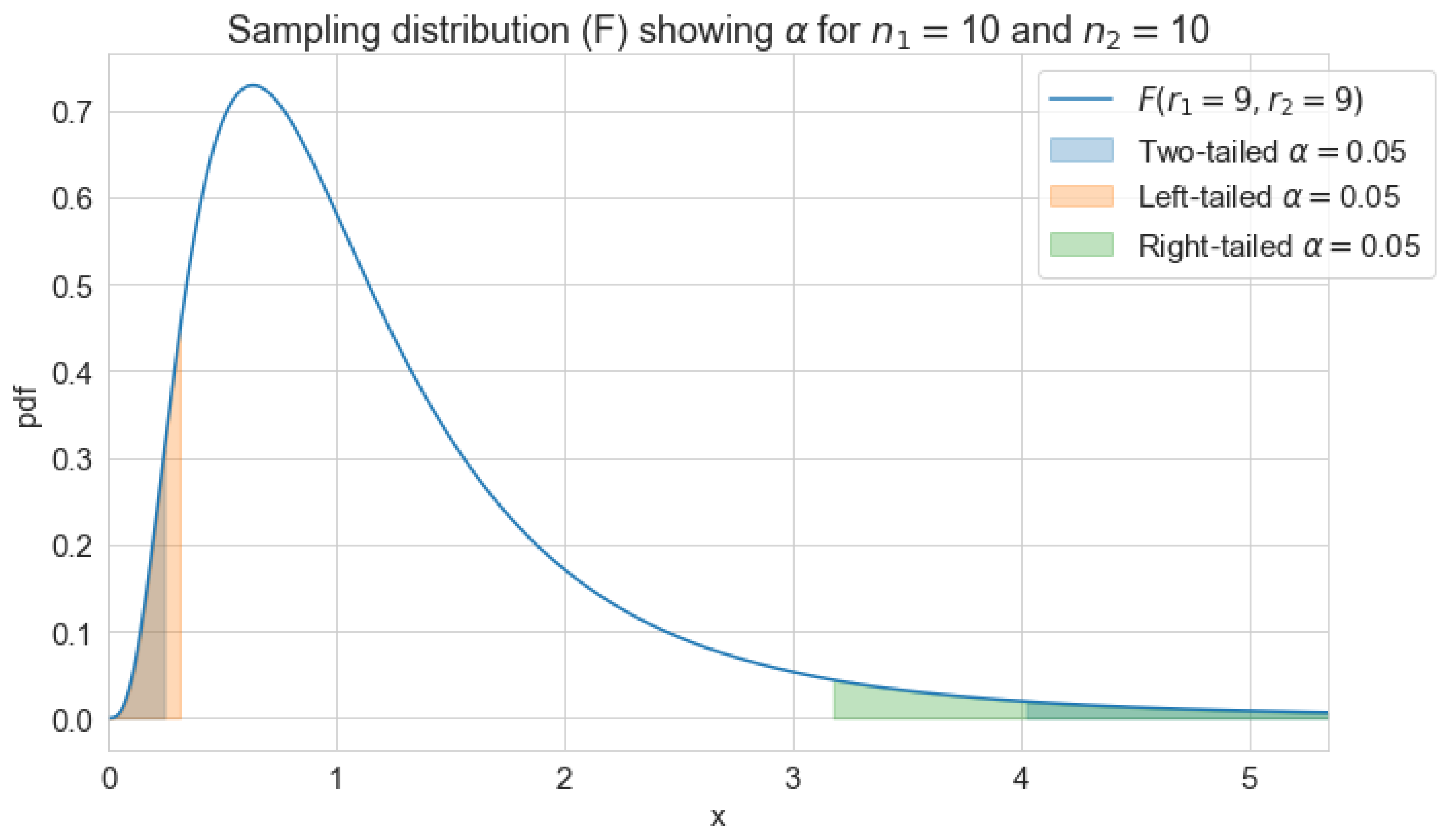
$$F < F_{PPF}\left(\frac{\alpha}{2}, r_1 = n_1 - 1, r_2 = n_2 - 1\right)$$

Conclusion



Sampling distribution (F) showing α for $n_1 = 10$ and $n_2 = 10$
$F(r_1 = 9, r_2 = 9)$
Two-tailed $\alpha = 0.05$
pdf
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
5
x

Conclusion



Sampling distribution (F) showing α for $n_1 = 10$ and $n_2 = 10$
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
pdf
0
1
2
3
4
5
x
$F(r_1 = 9, r_2 = 9)$
Two-tailed $\alpha = 0.05$
Left-tailed $\alpha = 0.05$
Right-tailed $\alpha = 0.05$

P-Value

If $H_a: \sigma_1^2 < \sigma_2^2$,

$$p = P(F(n_1 - 1, n_2 - 1) < F) = F_{CDF}(F, n_1 - 1, n_2 - 1)$$

If $H_a: \sigma_1^2 > \sigma_2^2$,

$$p = P(F(n_1 - 1, n_2 - 1) > F) = F_{SF}(F, n_1 - 1, n_2 - 1)$$

If $H_a: \sigma_1^2 \neq \sigma_2^2$,

$$p = 2 \times \min\big(F_{CDF}(F, n_1 - 1, n_2 - 1), F_{SF}(F, n_1 - 1, n_2 - 1)\big)$$

Python Functions

Bad news: The F test is not included in Python either.

Good news: I’ve coded it up for you – see the notebook for this lesson.

- F_2samp_var_stats if you have summary statistics
- F_2samp_var_data if you have raw data

Example

A two-sample t-test is planned to determine if there are significant differences in mean time to failure (MTTF) of an electronic component constructed using an old versus new methodology.

- A random sample of 15 components constructed with the old method resulted in a MTTF of 65 hours with sample variance of 5 hours.
- A random sample of 18 components constructed with the new method resulted in a MTTF of 75 hours with a sample variance of 8 hours.

Is there enough evidence to conclude that the variance differs? Assume $\alpha = 0.05$.

Example

Hypotheses:

- $H_0: \sigma^2_{\text{old}} = \sigma^2_{\text{new}}$, $H_a: \sigma^2_{\text{old}} \neq \sigma^2_{\text{new}}$

Test Statistic:

- $F = \frac{s^2_{\text{old}}}{s^2_{\text{new}}} = \frac{5}{8} = 0.625$

Rejection Region: $F < F_{PPF}\left(\frac{\alpha}{2}, r_1 = 14, r_2 = 17\right) = 0.3448$

P-value: $p = 2 \times F_{CDF}(0.625, \mathrm{r}_1 = 14, \mathrm{r}_2 = 17) = 0.3795$

We fail to reject the null hypothesis. There is insufficient evidence to conclude that the variance has changed.

Recap

- One-Sample Test of Variance (χ^2 test)
- Two-Sample Test of Variance (F test)