Chapter 10: Inference on Population Variances

Lecture slides: PDF pages 1–3 (open)

Probability and Statistics for the Life Sciences **Fall 2026**

So far, we have often asked about a population **mean**: its typical value. This chapter asks about its **spread**. Two instruments can have the same average reading but very different precision. Two treatments can have the same average effect but very different consistency. A variance test addresses that second kind of question.

The whole chapter in one sentence

Use a sample variance to learn about a population variance, while accounting for the variation that arises simply because a different random sample would give a different answer.

A route through these notes

Question you want to answer	Tool you need
How do I calculate spread from data or summary totals?	Sample variance and standard deviation
Is one population more or less variable than a claimed value?	One-sample chi-square test
What values of the population spread are plausible?	Chi-square confidence interval
Do two independent populations have different spreads?	Two-sample F test or F confidence interval
How likely is a test to miss a real difference?	Type II error and power under a specified alternative
Do two populations have different means?	A different question: a two-sample t procedure

How to study asynchronously

- Read the explanation before each example. Cover the calculation and predict which quantity comes next.
- Work each arithmetic step yourself, then compare. Keep extra digits until the final answer.
- At each conclusion, say whether you are discussing **variance, standard deviation, or mean**. Those are not interchangeable.
- Use the final homework checklist to decide which worked example is the closest model for each assigned problem. The examples are practice models, not a replacement for your own homework solution.

How the notes and slides work together

Each section heading links to the relevant pages of the **51-page lecture-slide PDF**. These are PDF viewer page numbers, not incremental slide numbers. The footer links to the chapter section index. If a viewer does not jump to the requested page, enter that page number manually. Examples labeled “from the lecture slides” retain the lecture data; additional examples build skills without supplying the assigned problem-set answers.

What are we estimating?

Lecture slides: PDF pages 2–3, 8 (open)

A **population** is the full distribution of measurements we care about. A **sample** is the set of observations we actually have. We use Greek letters for population parameters and ordinary letters for sample summaries.

Population quantity	Sample estimate	Meaning
μ (mu)	$\bar{x}$	Center: a typical value
σ^2 (sigma squared)	s^2	Variance: average squared spread
σ	s	Standard deviation: spread in original units

If readings are measured in millimeters, standard deviation is in millimeters and variance is in square millimeters. A claim about a standard deviation of 2 mm is a claim about a variance of 4 mm^2, **not** a variance of 2.

Why do we square deviations?

Let n be the number of observations, x_i observation i, and $\bar{x}$ their sample mean. The deviations $x_i - \bar{x}$ sum to zero: positive and negative deviations cancel. Squaring makes each contribution nonnegative and gives larger deviations more weight. The symbol $\sum$ means to add over all observations. The sample variance is

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}, \qquad s = \sqrt{s^2}.$$

Why n − 1, rather than n?

Once the sample mean is estimated, the deviations must add to zero. If you know the first n − 1 deviations, the last is forced. This leaves n − 1 **degrees of freedom**. Dividing by n − 1 also makes s^2 unbiased: across repeated samples, its average equals σ^2. This does not mean any one sample variance equals the population variance.

Parameter versus statistic

σ^2 is fixed but unknown. s^2 is calculated from your sample and would change if you collected a new sample. Statistical inference describes what that sample-to-sample variation allows us to conclude about σ^2.

The motivating examples in the slides

An aircraft altimeter can be correct *on average* yet unsafe if individual readings vary by roughly ±1000 feet. Conversely, an exam with almost no score variation may be unable to distinguish levels of preparation. Whether large or small spread is desirable depends on the scientific purpose.

The data-collection design matters

For repeated measurements on the same subject, time-series measurements, or other dependent observations, the independent-sample procedures in this chapter may not apply. Do not silently treat those observations as independent.

A variance calculation, line by line

Lecture slides: PDF pages 3 (open)

Suppose four independent readings are 2, 4, 5 and 9 units. We will calculate the mean, variance and standard deviation.

1. Find n and the sample mean.

There are n = 4 observations. Add them and divide by 4: $\bar{x} = (2 + 4 + 5 + 9)/4 = 20/4 = 5$.

2. Subtract the mean, then square each deviation.

Do not square the observations first and call that the variance: variance measures distance from the mean.

Reading x_i	Deviation $x_i - 5$	Squared deviation
2	-3	9
4	-1	1
5	0	0
9	4	16
Total	0	26

3. Divide the sum of squared deviations by n − 1.

The sum of squared deviations is 26. There are $4 - 1 = 3$ degrees of freedom. Thus $s^2 = 26/3 = 8.6667$ square units.

4. Take a square root if you need standard deviation.

s = $\sqrt{8.6667} = 2.9439$ units. Keep the units straight: a variance of 8.6667 is not a standard deviation of 8.6667.

Shortcut when the problem gives $\sum x_i$ and $\sum x_i^2$

$$s^2 = \frac{\sum x_i^2 - (\sum x_i)^2/n}{n-1}.$$

In this example, $\sum x_i^2 = 2^2 + 4^2 + 5^2 + 9^2 = 4 + 16 + 25 + 81 = 126$. Also, $(\sum x_i)^2/\text{n} = 20^2/4 = 100$. Therefore $s^2 = (126 - 100)/3 = 26/3$, exactly as above.

Two totals that are easy to confuse

$\sum x_i^2$ means square each reading and then add. $(\sum x_i)^2$ means add the readings and then square the total. Here they are 126 and 400, respectively. They are different quantities.

Why a chi-square distribution?

Lecture slides: PDF pages 4–6, 8–12 (open)

A confidence interval or test needs a reference distribution: what values would our statistic take over many repetitions of the sampling process? For variance, the key result is different from the normal or t distributions used for means.

Build the chi-square idea from familiar pieces

1. Start with a standard normal variable.

If Z has a standard normal distribution, its mean is 0 and its standard deviation is 1. Its square Z^2 can never be negative.

2. Add independent squared standard normals.

Let ν (the Greek letter "nu") count the number of independent squared pieces. Then $Z_1^2 + \cdots + Z_\nu^2$ has a chi-square distribution with ν degrees of freedom, written χ^2_ν.

3. Connect this to sample variance.

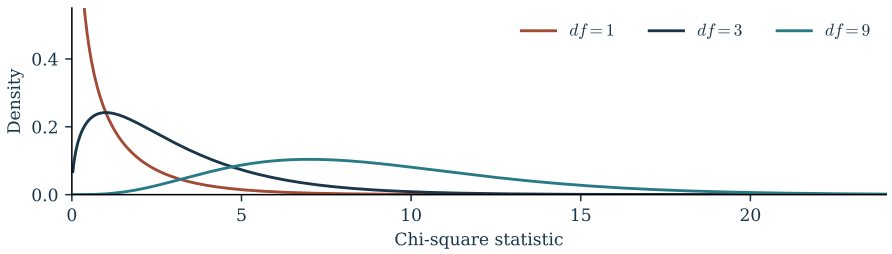
When $X_1, \ldots, X_n$ are independent observations from a normal population, subtracting the sample mean uses one degree of freedom. Dividing by the true population variance removes the units. The result is

$$Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Capital S^2 means the sample variance before we know which sample we will get. Lowercase s^2 is its observed value after we collect the data. For inference, we plug the observed s^2 into the formula.

What does this distribution look like?

It is nonnegative and usually right-skewed, not symmetric about zero. Its mean is ν and its variance is 2ν. Small degrees of freedom produce considerable uncertainty; larger samples make the estimated variance relatively more stable.



Different degrees of freedom give different chi-square curves. Squaring produces nonnegative values, and the curves are not symmetric.

Assumptions are part of the answer

The exact chi-square result requires **independent observations from a normal population**. A random sample supports independence. Normality concerns the population of measurements, not the distribution of s^2. Unlike many procedures for means, variance procedures can be quite sensitive to nonnormality and outliers. In a classroom problem, state normality as an assumption when it is not explicitly supplied; do not claim the data prove it.

Why the variance statistic has n minus one degrees of freedom

Lecture slides: PDF pages 8–12 (open)

Define the random quantities before doing the algebra

Let $X_1, \ldots, X_n$ be independent measurements from one normal population with population mean μ and population standard deviation $\sigma > 0$. Their sample mean is $\bar{X} = (X_1 + \cdots + X_n)/n$, and their sample variance is $S^2 = \sum(X_i - \bar{X})^2/(n-1)$. These capital-letter statistics vary from sample to sample.

For each observation, define $Z_i = (X_i - \mu)/\sigma$. Subtracting μ centers the observation and dividing by σ expresses it in population-SD units. Consequently the Z_i are independent standard normal variables. Their average is

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma} = \frac{\bar{X} - \mu}{\sigma}.$$

The mean of $\bar{Z}$ is 0 and its variance is $1/n$, because it is the average of n independent variance-one variables.

Rewrite each centered observation

Insert and subtract the population mean, then divide by σ:

$$\frac{X_i - \bar{X}}{\sigma} = \frac{(X_i - \mu) - (\bar{X} - \mu)}{\sigma} = Z_i - \bar{Z}.$$

Substituting into the sample-variance formula gives

$$\begin{aligned}\frac{(n-1)S^2}{\sigma^2} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \\ &= \sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 = \sum_{i=1}^{n}(Z_i - \bar{Z})^2 \sim \chi^2_{n-1}.\end{aligned}$$

The final step uses the lecture result about centered independent normal observations. The centered residuals themselves are not independent: they sum to zero. That single constraint leaves $n-1$ free directions, which is why the final distribution has $n-1$, not n, degrees of freedom.

What the simulation illustrates

The slides simulate **500 separate samples, each with** $n = 7$, from a standard normal population. In that population $\sigma^2 = 1$, so each sample produces $Q = 6S^2$. The histogram of those 500 statistics approximately follows χ^2_6. The number 500 counts repeated experiments; it is *not* the sample size used for the degrees of freedom. The theorem gives the reference distribution; the histogram illustrates it rather than proving it.

Check your understanding. If the same simulation used 2,000 samples of size 7, would the degrees of freedom change?

Reasoning. No. Each variance still comes from seven observations and one fitted mean, so $df = 6$. More repetitions only make the histogram a more stable picture of the same distribution.

Which tail does the number describe?

Lecture slides: PDF pages 7 (open)

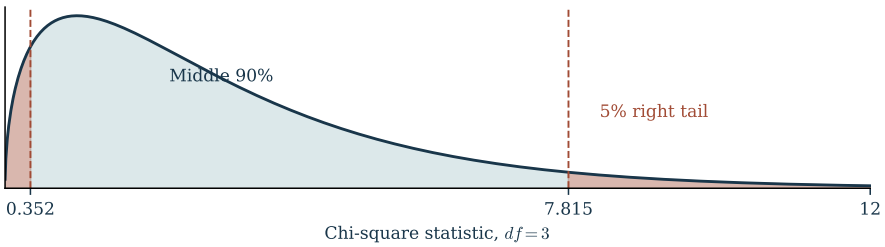
The lecture slides use **upper-tail critical values**. Let Q have a chi-square distribution with ν degrees of freedom, and let a denote a chosen right-tail probability between 0 and 1. We use the same convention throughout:

$$P\left(Q > \chi^2_{\nu,a}\right) = a.$$

The subscript a is the area **to the right** of the critical value, not the area to the left. The degrees of freedom ν identify the distribution.

Critical value	Area to its right	Where it sits
$\chi^2_{\nu,0.05}$	0.05	Farther to the right
$\chi^2_{\nu,0.95}$	0.95	Farther to the left

For example, with $\nu = 3$, $\chi^2_{3,0.05} = 7.8147$ and $\chi^2_{3,0.95} = 0.3518$. The middle 90% of the distribution lies between 0.3518 and 7.8147. There is 5% below the lower value and 5% above the upper value.



The middle 90% lies between two different positive cutoffs. The left and right excluded tails each have probability 5%.

Why a small right-tail probability gives a big number

To leave only 5% of the distribution to the right, move your cutoff far to the right. To leave 95% to the right, move it near the left edge. Always draw or imagine the shaded tail before choosing a table column.

Translate software conventions carefully

Some software asks for the probability to the **left**. If a is a right-tail probability, the corresponding left-tail probability is 1 − a. Therefore the same critical value can be computed either way:

Excel translations

CHISQ.INV.RT(a, ν) uses right-tail area a.
CHISQ.INV(1 − a, ν) uses left-tail area 1 − a.
For example, CHISQ.INV.RT(0.05,3) and CHISQ.INV(0.95,3) both give 7.8147.

The distinction is not just notation: using the wrong tail can reverse a test or produce a wrong interval. In the rest of these notes, every χ^2 or F critical value has the upper-tail meaning used in the lecture slides.

From a middle probability to an interval

Lecture slides: PDF pages 13 (open)

Suppose we want a $100(1-\alpha)\%$ confidence interval for σ^2. For 95% confidence, $\alpha = 0.05$; the total excluded probability is 0.05, with 0.025 in each tail.

1. Capture the middle $1-\alpha$ probability.

For $\nu = \mathrm{n} - 1$, the lower cutoff leaves $1 - \alpha/2$ to its right; the upper cutoff leaves $\alpha/2$ to its right.

$$P\left(\chi^2_{\nu,1-\alpha/2} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{\nu,\alpha/2}\right) = 1-\alpha.$$

2. Solve the inequalities for σ^2.

Let $\mathrm{A} = (\mathrm{n}-1)s^2$, L be the lower cutoff, and U be the upper cutoff. All are positive. From $\mathrm{A}/\sigma^2 \leq \mathrm{U}$, we obtain $\mathrm{A} \leq \mathrm{U}\sigma^2$ and then $\sigma^2 \geq \mathrm{A}/\mathrm{U}$. From $\mathrm{L} \leq \mathrm{A}/\sigma^2$, we obtain $\mathrm{L}\sigma^2 \leq \mathrm{A}$ and then $\sigma^2 \leq \mathrm{A}/\mathrm{L}$.

3. Put the lower and upper endpoints in order.

The larger chi-square cutoff belongs in the denominator of the **lower** endpoint. The smaller cutoff belongs in the denominator of the **upper** endpoint.

$$\boxed{\frac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{n-1,1-\alpha/2}}}$$

4. If asked about σ instead, take square roots.

Square root is increasing on positive numbers, so take the square root of each endpoint without reversing their order. A confidence interval for variance is not already an interval for standard deviation.

What does 95% confidence mean?

If we repeatedly sampled and constructed intervals this way, 95% of those intervals would contain the fixed true parameter. For one observed interval, we say we are 95% confident that the parameter lies between the computed endpoints. We do not interpret this as 95% of individual observations lying inside the interval.

Expect an asymmetric interval

A chi-square interval is generally not $s^2 \pm$ a margin of error. Its endpoints arise by division by two different positive cutoffs. Small samples can give very wide intervals, especially above s^2.

A 90% interval from four observations

Lecture slides: PDF pages 14–15 (open)

Lecture Example 1: observations 4.1, 5.2, 6.7 and 10.2 from an approximately normal population. Construct a 90% confidence interval for the population variance. Assume the observations are independent.

1. Calculate the mean.

n = 4 and $\bar{x} = (4.1 + 5.2 + 6.7 + 10.2)/4 = 26.2/4 = 6.55$.

2. Calculate each squared deviation and add.

The deviations are -2.45, -1.35, 0.15 and 3.65. Their squares are 6.0025, 1.8225, 0.0225 and 13.3225. The sum is 21.17.

3. Calculate the sample variance and degrees of freedom.

$s^2 = 21.17/(4 - 1) = 7.0567$. The degrees of freedom are $\nu = 3$. Notice that $(\text{n} - 1)s^2 = 21.17$: it is simply the sum of squared deviations.

4. Translate the confidence level into tail areas.

90% confidence means $\alpha = 0.10$. Each excluded tail has area $\alpha/2 = 0.05$. We need the critical values with right-tail areas 0.05 and 0.95.

$$\chi^2_{3,0.05} = 7.8147, \qquad \chi^2_{3,0.95} = 0.35185.$$

5. Divide in the correct order.

For the lower endpoint, divide by the bigger cutoff. For the upper endpoint, divide by the smaller cutoff.

$$\text{Lower} = \frac{21.17}{7.8147} = 2.709, \qquad \text{Upper} = \frac{21.17}{0.35185} = 60.168.$$

6. Interpret the answer.

We are 90% confident that the population variance is between **2.709 and 60.168 square units**. The interval is wide because four observations provide little information about population spread. Extra precision in the critical values may give a slightly different endpoint from a rounded lecture table.

If the question asked for standard deviation

Take square roots: $\sqrt{2.709} = 1.646$ and $\sqrt{60.168} = 7.757$. The corresponding 90% confidence interval for σ is **(1.646, 7.757) units**. Do not square these endpoints again.

Pause and predict

Would a 95% interval from the same data be wider or narrower? **Wider.** To capture the true parameter more often over repeated sampling, the procedure must produce a larger range of plausible values.

Turn the scientific question into a tail

Lecture slides: PDF pages 16–19 (open)

Suppose a claim specifies σ_0^2. The null hypothesis is H_0: $\sigma^2 = \sigma_0^2$. The alternative hypothesis describes the departure you are trying to detect. Write the alternative **before** looking at whether your sample variance is large or small.

$$Q_{\text{obs}} = \frac{(n-1)s^2}{\sigma_0^2}, \qquad \nu = n - 1.$$

Under H_0, the denominator equals the true population variance, so Q has the chi-square reference distribution. A bigger s^2 gives a bigger Q; a smaller s^2 gives a smaller Q.

Alternative	What counts as evidence?	Reject H_0 when
$\sigma^2 > \sigma_0^2$	Unusually large Q	$Q > \chi^2_{\nu,\alpha}$
$\sigma^2 < \sigma_0^2$	Unusually small Q	$Q < \chi^2_{\nu,1-\alpha}$
$\sigma^2 \neq \sigma_0^2$	Unusually small or large Q	$Q < \chi^2_{\nu,1-\alpha/2}$ or $Q > \chi^2_{\nu,\alpha/2}$

A repeatable six-step solution

- **State the parameter and hypotheses.** Translate any claimed standard deviation into a variance by squaring.
- **State assumptions and α.** Independent observations; normal population; chosen significance level.
- **Calculate the statistic.** Show n − 1, s^2, σ_0^2 and the substitution.
- **Find the correct critical value(s).** Use n − 1 degrees of freedom and the appropriate tail area(s).
- **Make the decision.** Is your statistic in the rejection region?
- **Answer the scientific question.** Say what there is or is not sufficient evidence to conclude about population spread.

Do not “accept” the null

Failing to reject H_0 means the data do not provide sufficiently strong evidence against it at the chosen α. It does not establish that the claimed variance is true. A test may miss a real difference, especially with small samples.

A p-value is the null probability of results at least as unfavorable to H_0 in the direction(s) specified by H_a. Reject when p $\leq \alpha$. Critical-value and p-value approaches give the same decision when they implement the same test. The lecture slides emphasize critical-value solutions; software can also calculate tail probabilities.

Is the concrete too variable?

Lecture slides: PDF pages 20–24 (open)

Lecture Example 2: a claimed range of 40 kg/cm^2 is approximated in the problem by 4σ. A normal-population sample has n = 10 and $s^2 = 195$ (kg/cm^2)^2. Test whether variability is too large at $\alpha = 0.05$. Assume independent observations.

1. Translate the stated model into a variance.

Under the problem's range $\approx 4\sigma$ rule, $\sigma_0 = 40/4 = 10$ kg/cm^2. Therefore $\sigma_0^2 = 10^2 = 100$ (kg/cm^2)^2. This "range $\approx 4\sigma$" conversion is a stipulated heuristic, not a universal identity for sample ranges.

2. State what "too variable" means mathematically.

H_0: $\sigma^2 = 100$ versus H_a: $\sigma^2 > 100$. It is an **upper-tail** test because we are looking for excessive spread, not just any difference. If the scientific null is "at most 100," its boundary value 100 supplies the critical calculation.

3. Compute the observed statistic.

The degrees of freedom are $10 - 1 = 9$. Multiply the sample variance by 9 and divide by the null variance.

$$Q_{\text{obs}} = \frac{9(195)}{100} = \frac{1755}{100} = 17.55.$$

4. Find the upper 5% cutoff.

With $\nu = 9$ and right-tail area 0.05, $\chi^2_{9,0.05} = 16.919$. Reject if Q > 16.919.

5. Compare the statistic to the cutoff.

$17.55 > 16.919$, so Q falls in the rejection region. Reject H_0 at the 5% significance level. The upper-tail p-value is approximately 0.0408, also below 0.05.

6. Return to the original context.

There is sufficient evidence that the population variance exceeds 100 (kg/cm^2)^2, equivalently that the population standard deviation exceeds 10 kg/cm^2. Under the stated range approximation, the variability is greater than the claimed acceptable level.

Why not compare 195 with 16.919?

195 is a variance with physical units. 16.919 is a cutoff for the dimensionless statistic Q. First transform the sample variance into Q; then compare Q with its reference cutoff.

When “too little spread” is the concern

Chapter 10 lecture slides (open) Additional practice and connections

Additional practice example: an assessment should have a population standard deviation of at least 8 points to distinguish performance levels. A random normal-population sample of n = 13 scores has s = 6 points. Is the spread below the target at $\alpha = 0.05$?

1. Convert standard deviations to variances.

The target variance is $\sigma_0^2 = 8^2 = 64$ points^2. The observed sample variance is $s^2 = 6^2 = 36$ points^2. Both must be squared before entering the chi-square formula.

2. Choose the alternative from the scientific concern.

We want to detect insufficient spread: H_0: $\sigma^2 = 64$ versus H_a: $\sigma^2 < 64$. This is a lower-tail test. The boundary $\sigma^2 = 64$ is used if the scientific null is $\sigma^2 \geq 64$.

3. Calculate the statistic and degrees of freedom.

$\nu = 13 - 1 = 12$. Then

$$Q_{\text{obs}} = \frac{12(36)}{64} = \frac{432}{64} = 6.75.$$

4. Find a cutoff with 5% to its left.

The lecture table labels the **right** tail, so the right-tail area must be 0.95. The cutoff is $\chi^2_{12,0.95} = 5.226$. Reject if Q < 5.226.

5. Make the comparison and interpret it.

6.75 is greater than 5.226, so it is **not** in the rejection region. Fail to reject H_0. There is insufficient evidence at the 5% level that the population standard deviation is below 8 points. The lower-tail p-value is about 0.1263.

A smaller sample spread is not automatically significant

The sample standard deviation, 6, is indeed below the target, 8. The test asks whether the discrepancy is large enough relative to the sampling uncertainty. With only 13 observations, this particular discrepancy is not enough to reject at $\alpha = 0.05$.

Check your direction before computing

“More consistent,” “more precise” or “below the minimum variability” usually points toward a smaller variance. “Less consistent” or “too variable” points toward a larger variance. “Different variability” is two-sided. Read the actual claim rather than inferring the tail from the observed data.

Why an F distribution?

Lecture slides: PDF pages 25–26, 32 (open)

Now compare two population variances, σ_1^2 and σ_2^2. Define the parameter of interest as $\rho = \sigma_1^2/\sigma_2^2$. The Greek letter ρ ("rho") denotes the ratio. A value of 1 means equal variances; a value below 1 means population 1 is less variable.

1. Write a chi-square quantity for each sample.

For independent normal samples, U $= (n_1 - 1)S_1^2/\sigma_1^2$ has $\nu_1 = n_1 - 1$ degrees of freedom. V $= (n_2 - 1)S_2^2/\sigma_2^2$ has $\nu_2 = n_2 - 1$. Independence between the samples makes U and V independent.

2. Divide each chi-square quantity by its degrees of freedom.

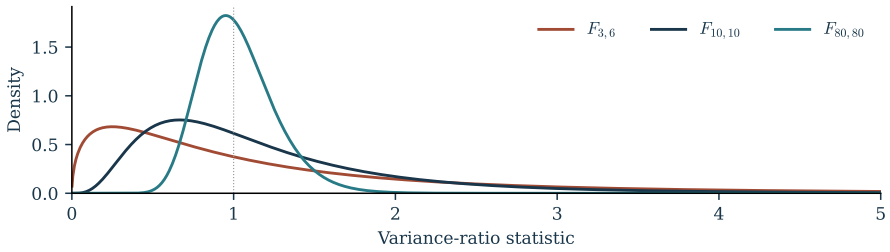
$U/\nu_1 = S_1^2/\sigma_1^2$ and $V/\nu_2 = S_2^2/\sigma_2^2$. Each is a sample variance expressed relative to its true population variance.

3. Take their ratio.

The ratio of two independent chi-square quantities, each divided by its degrees of freedom, has an F distribution:

$$F = \frac{U/\nu_1}{V/\nu_2} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2/S_2^2}{\rho} \sim F_{\nu_1,\nu_2}.$$

This distribution is positive and asymmetric. It needs **two** degrees of freedom: numerator first, denominator second. When H_0 says $\rho = 1$, the statistic reduces to the observed ratio R $= s_1^2/s_2^2$.



The numerator and denominator degrees of freedom both affect the curve. When both are large, the ratio becomes concentrated near 1.

Always write the orientation

If you use s_1^2/s_2^2, use degrees of freedom $(n_1 - 1, n_2 - 1)$. If you invert the ratio to s_2^2/s_1^2, **swap the degrees of freedom too**. For unequal sample sizes, forgetting the swap changes your cutoff.

The lecture slides again use upper-tail notation: P(F $> F_{\nu_1,\nu_2,a}$) = a. The lower-tail cutoff is obtained through a reciprocal:

$$F_{\nu_1,\nu_2,1-a} = \frac{1}{F_{\nu_2,\nu_1,a}}.$$

Assumptions: independent observations within each group, independent groups, and normal populations. The F variance test is sensitive to departures from normality. A paired design is not the independent-samples setup used here.

How sample size changes an F curve

Lecture slides: PDF pages 27–30 (open)

An F statistic contains randomness from *both* its numerator and its denominator. This section explains the limiting statements in the slides rather than introducing another test to memorize.

Why a large denominator degrees of freedom stabilizes the denominator

Let $V \sim \chi^2_{\nu_2}$. Here ν_2 is the denominator degrees of freedom. Write E for expectation (mean) and Var for variance. Recall that $\mathrm{E}(V) = \nu_2$ and $\mathrm{Var}(V) = 2\nu_2$. For a constant c, $\mathrm{E}(cV) = c\,\mathrm{E}(V)$ and $\mathrm{Var}(cV) = c^2\,\mathrm{Var}(V)$. With $c = 1/\nu_2$,

$$\mathrm{E}\left(\frac{V}{\nu_2}\right) = 1, \qquad \mathrm{Var}\left(\frac{V}{\nu_2}\right) = \frac{2\nu_2}{\nu_2^2} = \frac{2}{\nu_2}.$$

As ν_2 grows, the mean stays at 1 and the variance approaches 0. Thus V/ν_2 becomes concentrated near 1. If $U \sim \chi^2_{\nu_1}$ is independent of V, then

$$F = \frac{U/\nu_1}{V/\nu_2} \;\approx\; \frac{U}{\nu_1} \quad \text{when } \nu_2 \text{ is large.}$$

Only the denominator is becoming nearly constant in this argument; the numerator may still fluctuate considerably.

Translate that statement into a cumulative probability

A cumulative distribution function (CDF) at C gives the probability that a random variable is at most C. Here C is a fixed positive cutoff. Multiplying the inequality by the positive number ν_1 gives

$$P(F \le C) \approx P\left(\frac{U}{\nu_1} \le C\right) = P(U \le \nu_1 C).$$

This explains the comparison of an F CDF with a chi-square CDF in the slides. The chi-square cutoff must be multiplied by ν_1; the two curves do not have identical numerical inputs.

Two useful consequences

- If *both* ν_1 and ν_2 are large, both scaled chi-square quantities are close to 1, so F is concentrated near 1. This is a concentration statement, not a claim that every finite-*df* F distribution has mean exactly 1.
- For a normal sample, $S^2/\sigma^2 = Q/(n-1)$ with $Q \sim \chi^2_{n-1}$. Its mean is 1 and its variance is $2/(n-1)$. Larger samples therefore make S^2 more precise as an estimate of σ^2.

Check your understanding. Why does a small sample give a wide variance interval?

Reasoning. Its sample variance changes substantially across possible samples. More data reduce that sampling variability; they do not necessarily reduce the spread of the individual measurements in the population.

F tables, reciprocals and the three alternatives

Lecture slides: PDF pages 31–38, 47 (open)

Let $\rho = \sigma_1^2/\sigma_2^2$ denote the population variance ratio and $R = s_1^2/s_2^2$ its observed sample estimate. Fix that order before choosing a test. Let $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ be the corresponding degrees of freedom.

How to read the table

The notation $F_{\nu_1,\nu_2,a}$ means the cutoff with right-tail probability a:

$$P(F > F_{\nu_1,\nu_2,a}) = a.$$

Choose the table for the required tail area, then the numerator and denominator degrees of freedom. For example, $F_{3,6,.05} = 4.7571$: an $F_{3,6}$ variable exceeds 4.7571 with probability .05. The order $(3, 6)$ cannot be swapped without changing the quantity.

Why reciprocals swap the degrees of freedom

If $F = (U/\nu_1)/(V/\nu_2)$, then

$$\frac{1}{F} = \frac{V/\nu_2}{U/\nu_1} \sim F_{\nu_2,\nu_1}.$$

The numerator and denominator literally change places. Because taking a reciprocal reverses order for positive numbers, a small value of F corresponds to a large value of $1/F$. Thus

$$F_{\nu_1,\nu_2,1-a} = \frac{1}{F_{\nu_2,\nu_1,a}}.$$

This lets us obtain lower-tail cutoffs from tables that list only small upper-tail probabilities.

Rejection rules for H-zero: equal population variances

Under $H_0 : \rho = 1$, R has the F_{ν_1,ν_2} null distribution. At significance level α:

Alternative	Reject when
$H_a : \rho > 1$	$R > F_{\nu_1,\nu_2,\alpha}$.
$H_a : \rho < 1$	$R < 1/F_{\nu_2,\nu_1,\alpha}$, equivalently $1/R > F_{\nu_2,\nu_1,\alpha}$.
$H_a : \rho \neq 1$	$R < 1/F_{\nu_2,\nu_1,\alpha/2}$ or $R > F_{\nu_1,\nu_2,\alpha/2}$.

The slides note that the larger of R and $1/R$ is the convenient one to compare to an upper-tail cutoff in the usual two-sided tests. You must still use the degrees of freedom appropriate to that ratio and split α between two tails. The fixed-order, two-boundary rule above is the safest general way to keep the bookkeeping correct.

A clarification about the null value. Equal variances correspond to $\rho_0 = 1$, which is the null used in these slides. A test of another specified positive ratio is also possible: use R/ρ_0 as the null F statistic. The unscaled ratio R has the stated F distribution only under equality.

Does production line 1 have less variation?

Lecture slides: PDF pages 39–42 (open)

Lecture Example 3: independent normal-population samples from two production lines have $n_1 = 20$, $n_2 = 25$, $s_1^2 = 0.51$ and $s_2^2 = 1.10$. Test whether line 1 has smaller impurity variance at $\alpha = 0.05$.

1. State the hypotheses in a fixed order.

Let $\rho = \sigma_1^2/\sigma_2^2$. H_0: $\rho = 1$ versus H_a: $\rho < 1$. This is not a claim about the mean impurity; it is about variability.

2. Choose a convenient ratio for a one-sided upper-tail table.

Because the alternative says line 1 is less variable, s_2^2/s_1^2 should be unusually large if the alternative is true. We can therefore invert the ratio and use an upper-tail test.

$$F_{\text{obs}} = \frac{s_2^2}{s_1^2} = \frac{1.10}{0.51} = 2.1569.$$

3. Assign degrees of freedom in that same order.

The numerator is sample 2, so $\nu_{\text{numerator}} = 25 - 1 = 24$. The denominator is sample 1, so $\nu_{\text{denominator}} = 20 - 1 = 19$. The reference distribution is $F_{24,19}$, not $F_{19,24}$.

4. Find the upper 5% critical value.

$F_{24,19,0.05} = 2.1141$. Under H_0, only 5% of ratios in this orientation exceed this cutoff.

5. Compare and conclude.

$2.1569 > 2.1141$, so reject H_0. There is sufficient evidence at the 5% level that line 1 has smaller population variance than line 2. The one-sided p-value is approximately 0.0457.

Check the same test without inverting

Keep R $= s_1^2/s_2^2 = 0.51/1.10 = 0.4636$. Its degrees of freedom are (19,24). The lower 5% cutoff is $1/F_{24,19,0.05} = 1/2.1141 = 0.4730$. Since $0.4636 < 0.4730$, we again reject H_0.

One procedure, two equivalent presentations

Either orientation is valid if you keep the alternative, rejection tail and degrees of freedom consistent. Do not select whichever orientation gives a preferred conclusion: the two correct calculations agree.

Are two response variances different?

Lecture slides: PDF pages 43–47 (open)

Lecture Example 4: $n_1 = 11$, $n_2 = 9$, $s_1^2 = 14.14$ and $s_2^2 = 3.21$ for independent normal-population samples. Test equality of population variances at $\alpha = 0.05$.

1. State a two-sided alternative.

H_0: $\sigma_1^2/\sigma_2^2 = 1$ versus H_a: $\sigma_1^2/\sigma_2^2 \neq 1$. Either population could have the larger variance. Divide α between the tails: $\alpha/2 = 0.025$.

2. Calculate the ratio and keep its labels.

$\mathrm{R} = s_1^2/s_2^2 = 14.14/3.21 = 4.4050$. Its numerator degrees of freedom are $11 - 1 = 10$; denominator degrees of freedom are $9 - 1 = 8$.

3. Find the upper and lower rejection cutoffs.

The upper cutoff is $F_{10,8,0.025} = 4.2951$. The lower cutoff uses swapped degrees of freedom before taking a reciprocal:

$$L = \frac{1}{F_{8,10,0.025}} = \frac{1}{3.8549} = 0.2594, \qquad U = 4.2951.$$

Thus the full rejection region is $\mathrm{R} < 0.2594$ **or** $\mathrm{R} > 4.2951$. The middle interval [0.2594, 4.2951] is the nonrejection region.

4. Compare the observed statistic with both boundaries.

$\mathrm{R} = 4.4050$ exceeds the upper boundary 4.2951. Therefore reject H_0 at $\alpha = 0.05$. The reciprocal check in the slides gives $1/R = 3.21/14.14 = 0.2270$. Its degrees of freedom are $(8, 10)$, so compare it with $F_{8,10,0.025} = 3.8549$. Although $0.2270 < 3.8549$, the original ratio already exceeds its cutoff: rejection requires *either* upper-tail comparison, not both.

5. State the scientific conclusion without changing the question.

There is sufficient evidence that the population variances differ. The observed direction is that population 1 has greater variability. This does not by itself show that population 1 has a greater mean response.

Why do we use 0.025 instead of 0.05?

Using a 5% rejection tail at each end would spend 10% total Type I error. For the equal-tailed two-sided test with total $\alpha = 0.05$, allocate 2.5% to each tail.

Optional p-value check

For this equal-tailed test, calculate both F tail probabilities and double the smaller one (capped at 1). Here the upper-tail probability is about 0.02323, so the two-sided p-value is about 0.04646. It is below 0.05, agreeing with the critical-value decision.

An interval for the variance ratio

Lecture slides: PDF pages 48–51 (open)

Let $\mathrm{R} = s_1^2/s_2^2$ and $\rho = \sigma_1^2/\sigma_2^2$. Since R/ρ follows F_{ν_1,ν_2}, the same inversion idea used for a one-population variance gives an interval for the ratio.

1. Start with the central F probability.

Let $\mathrm{L} = F_{\nu_1,\nu_2,1-\alpha/2}$ and $\mathrm{U} = F_{\nu_1,\nu_2,\alpha/2}$. Then $\mathrm{L} \le \mathrm{R}/\rho \le \mathrm{U}$ contains the middle probability $1-\alpha$.

2. Isolate ρ.

From $\mathrm{R}/\rho \le \mathrm{U}$, obtain $\rho \ge \mathrm{R}/\mathrm{U}$. From $\mathrm{L} \le \mathrm{R}/\rho$, obtain $\rho \le \mathrm{R}/\mathrm{L}$. Replace the small lower-tail cutoff L by its reciprocal expression to use upper-tail tables.

$$\boxed{\frac{s_1^2/s_2^2}{F_{n_1-1,n_2-1,\alpha/2}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} F_{n_2-1,n_1-1,\alpha/2}}$$

3. Check the meaning of the endpoints.

A ratio interval is dimensionless when both populations use the same measurement units. If it includes 1, equal variances remain compatible with the interval. If it is entirely above 1, population 1 has greater variance; entirely below 1 indicates smaller variance.

Lecture Example 5: prey sizes of two spider groups

The slides compare the prey sizes, in millimeters, of adult *Deinopis* and *Menneus* spiders in eastern Australia. Their reported sample means are 10.26 mm and 9.02 mm; those means are not needed to estimate the *variance ratio*. The observed standard deviations are the relevant measures of spread.

Use group 1 for *Deinopis* and group 2 for *Menneus*. The lecture slides give $n_1 = n_2 = 10$ and sample **standard deviations** $s_1 = 2.51$, $s_2 = 1.90$. Assume independent normal samples. We want a 95% interval for σ_1^2/σ_2^2.

- **Square first:** $s_1^2 = 6.3001$ and $s_2^2 = 3.6100$. $\mathrm{R} = 6.3001/3.6100 = 1.74518$. Do not use $2.51/1.90$ as a variance ratio.
- **Degrees of freedom:** $\nu_1 = \nu_2 = 9$. For 95% confidence, $\alpha = 1 - 0.95 = 0.05$, so $\alpha/2 = 0.025$. $F_{9,9,0.025} = 4.02599$.
- **Lower endpoint:** $1.74518/4.02599 = 0.43348$.
- **Upper endpoint:** $1.74518 \times 4.02599 = 7.02608$.

The 95% confidence interval for σ_1^2/σ_2^2 is **(0.4335, 7.0261)**. It includes 1 and is wide, so these small samples do not estimate the relative spread very precisely. For σ_1/σ_2 instead, take square roots to obtain approximately **(0.6584, 2.6507)**.

When you are given sums instead of raw data

Chapter 10 lecture slides (open) Additional practice and connections

Additional practice example: two independent normal samples summarize measurement variation. Group A has $n_A = 5$, $\bar{x}_A = 12$, and $\sum x_{A,i}^2 = 760$. Group B has $n_B = 5$, $\bar{x}_B = 16$, and $\sum x_{B,i}^2 = 1320$. Calculate the variances and outline a variance comparison.

1. Use a form of the sample-variance shortcut that matches the input.

Because $\bar{x} = (\sum x_i)/\mathrm{n}$, we have $(\sum x_i)^2/\mathrm{n} = \mathrm{n}\bar{x}^2$. Therefore

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}.$$

2. Calculate group A's centered sum of squares.

$n_A\bar{x}_A^2 = 5(12^2) = 5(144) = 720$. Subtract this from the sum of squared raw observations: $760 - 720 = 40$. Then $s_A^2 = 40/(5-1) = 10$.

3. Repeat for group B, not mixing the groups.

$n_B\bar{x}_B^2 = 5(16^2) = 5(256) = 1280$. The centered sum of squares is $1320 - 1280 = 40$. Thus $s_B^2 = 40/4 = 10$.

4. Form the correctly labeled variance ratio.

$\mathrm{R} = s_A^2/s_B^2 = 10/10 = 1$, with degrees of freedom (4,4). A two-sided variance-equality test at a conventional level does not reject when the ratio is exactly 1.

5. Do not confuse this with a test about means.

The sample means, 12 and 16, differ, while the sample variances are equal. A variance comparison does not answer whether the population means differ. The next pages show how to answer that separate question.

Arithmetic checks before doing inference

A sample variance cannot be negative. If $\sum x_i^2 - \mathrm{n}\bar{x}^2$ is negative, check the given units, your transcription, parentheses and rounding. Keep sufficient precision: subtracting two large nearly equal totals can magnify early rounding errors.

For a confidence interval, insert the computed variances into the ratio formula. For a test, use the α specified for that test. A 95% two-sided interval corresponds to a 5% two-sided test, **not** a 10% test; if the task gives different levels, do the two calculations separately.

Variance equality is not mean equality

Chapter 10 lecture slides (open) Additional practice and connections

A problem may first ask whether two populations have equal variances, then ask whether they have equal means. These are **two different hypotheses**, even though both calculations use the same samples.

Question	Null hypothesis	Typical statistic
Same variability?	$\sigma_1^2 = \sigma_2^2$	F variance ratio
Same average?	$\mu_1 = \mu_2$	Two-sample t statistic

Pooled two-sample t test

If the samples are independent, the populations are normal and a common population variance is a justified assumption, combine the two estimates of variance. Each sample contributes its centered sum of squares, $(n_i - 1)s_i^2$. Add those sums and divide by their total degrees of freedom:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

The subscript p means "pooled." The sampling variance of the difference between two independent sample means is $\sigma^2/n_1 + \sigma^2/n_2$. Replacing the common unknown σ^2 by s_p^2 gives the estimated standard error:

$$\mathrm{SE}(\bar{X}_1 - \bar{X}_2) = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}.$$

Let Δ_0 denote the hypothesized difference between the two population means. For H_0: $\mu_1 - \mu_2 = \Delta_0$, subtract the hypothesized difference from the observed difference and divide by that standard error:

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{s_p^2(1/n_1 + 1/n_2)}}, \qquad df = n_1 + n_2 - 2.$$

Usually $\Delta_0 = 0$. For a two-sided test, reject if |T| exceeds the upper-tail t cutoff with area $\alpha/2$. Use the t distribution for this mean comparison, not the F cutoff from the preceding variance test.

Important assumption, not a fact established by a test

Failure to reject equal variances does not prove the variances are equal. A pooled t calculation is valid under its common-variance model. In a course exercise, explicitly state that assumption. If equal variances are not justified, Welch's two-sample t procedure is generally used instead; do not mechanically pool just because a small-sample F test was nonsignificant.

Two groups, two different questions

Chapter 10 lecture slides (open) Additional practice and connections

Additional practice example: measurements from independent normal populations are A: 14, 16, 18, 20, 22 and B: 19, 21, 23, 25. Test a difference in means at $\alpha = 0.05$ under the common-variance model.

1. Find each sample mean.

$n_A = 5$; $\bar{x}_A = 90/5 = 18$. $n_B = 4$; $\bar{x}_B = 88/4 = 22$. The observed difference A − B is $18 - 22 = -4$.

2. Calculate each sample variance.

For A, deviations are −4, −2, 0, 2, 4; their squares add to 40. So $s_A^2 = 40/4 = 10$. For B, deviations are −3, −1, 1, 3; their squares add to 20. So $s_B^2 = 20/3 = 6.6667$.

3. Distinguish the preceding variance comparison.

The F ratio is $10/6.6667 = 1.5$, with degrees of freedom (4,3). A two-sided 5% F test has cutoffs $1/F_{3,4,0.025} = 0.1002$ and $F_{4,3,0.025} = 15.1010$. Since 1.5 lies between them, it does not reject equal variances. We proceed here under the explicitly assumed common-variance model, not because nonsignificance proves it.

4. State the mean hypotheses and pool the variances.

H_0: $\mu_A - \mu_B = 0$ versus H_a: $\mu_A - \mu_B \neq 0$. The total degrees of freedom are $5 + 4 - 2 = 7$.

$$s_p^2 = \frac{4(10) + 3(6.6667)}{7} = \frac{40 + 20}{7} = 8.5714.$$

5. Compute the standard error, then the t statistic.

$1/5 + 1/4 = 0.20 + 0.25 = 0.45$. Multiply by the pooled variance and take the square root.

$$\text{SE} = \sqrt{8.5714(0.45)} = 1.9640, \qquad T = \frac{-4 - 0}{1.9640} = -2.0367.$$

6. Compare with the t cutoff and interpret.

The two-sided 5% cutoff is $t_{7,0.025} = 2.3646$. Since $|-2.0367| = 2.0367 < 2.3646$, fail to reject H_0. Under the assumptions, there is insufficient evidence at the 5% level that the population means differ (p ≈ 0.0811). The negative sign describes the observed direction, not a negative probability.

A test can miss a real difference

Chapter 10 lecture slides (open) Additional practice and connections

A test's decision is based on a random sample. Even a real population difference may produce a sample statistic inside the nonrejection region.

Reality	Reject H_0	Do not reject H_0
H_0 is true	Type I error; probability α	Correct nonrejection
A specified alternative is true	Detection; probability $1-\beta$	Type II error; probability β

The **significance level** α is chosen using the null distribution. The **Type II error probability** β depends on a particular true alternative, as well as sample sizes and α. The **power** is $1-\beta$. There is not one β for the whole statement "the variances differ."

The key distinction for an F variance test

Write $\mathrm{R} = S_1^2/S_2^2$ and let the true variance ratio be ρ. The general sampling result is $\mathrm{R}/\rho \sim F_{\nu_1,\nu_2}$. Under equal variances, $\rho = 1$ and R itself has this F distribution. Under unequal variances, **R itself is scaled**.

$$\frac{R}{\rho} = W \sim F_{\nu_1,\nu_2}, \qquad R = \rho W.$$

Keep the rejection rule fixed; change the probability model

For an equal-tailed, two-sided test of H_0: $\rho = 1$ at level α, define the null nonrejection bounds

$$L = \frac{1}{F_{\nu_2,\nu_1,\alpha/2}}, \qquad U = F_{\nu_1,\nu_2,\alpha/2}.$$

The test does not reject when $\mathrm{L} \le \mathrm{R} \le \mathrm{U}$. If the true ratio is ρ, calculate the probability of that same event under the alternative:

$$\beta = P_\rho(L \le R \le U) = P\left(\frac{L}{\rho} \le W \le \frac{U}{\rho}\right).$$

Why divide both bounds by ρ?

Substitute $\mathrm{R} = \rho\mathrm{W}$ into $\mathrm{L} \le \mathrm{R} \le \mathrm{U}$, giving $\mathrm{L} \le \rho\mathrm{W} \le \mathrm{U}$. Because ρ is positive, dividing by it preserves the inequality directions. This rescaling is the essential step; using the null F distribution directly for R under unequal variances would be wrong.

Calculate β with an F cumulative distribution

Chapter 10 lecture slides (open) Additional practice and connections

Additional practice example: independent normal samples have $n_1 = n_2 = 8$. We test equal variances against unequal variances at $\alpha = 0.05$. What is β if the true variances are $\sigma_1^2 = 1$ and $\sigma_2^2 = 3$?

1. Construct the test using the null hypothesis.

Both degrees of freedom are $8 - 1 = 7$. Each rejection tail has probability 0.025 under H_0. The upper cutoff is $\mathrm{U} = F_{7,7,0.025} = 4.99491$. The lower cutoff is $\mathrm{L} = 1/F_{7,7,0.025} = 0.200204$.

2. Write the event that counts as a Type II error.

Under the specified unequal variances, H_0 is false. A Type II error occurs if the test nevertheless does not reject: $0.200204 \leq \mathrm{R} \leq 4.99491$.

3. Compute the true variance ratio and rescale.

$\rho = 1/3$. Since $\mathrm{W} = \mathrm{R}/\rho = 3\mathrm{R}$ follows $F_{7,7}$, divide each bound by 1/3, which is the same as multiplying by 3:

$$\beta = P(0.600612 \leq W \leq 14.98473), \qquad W \sim F_{7,7}.$$

4. Turn the interval probability into a CDF difference.

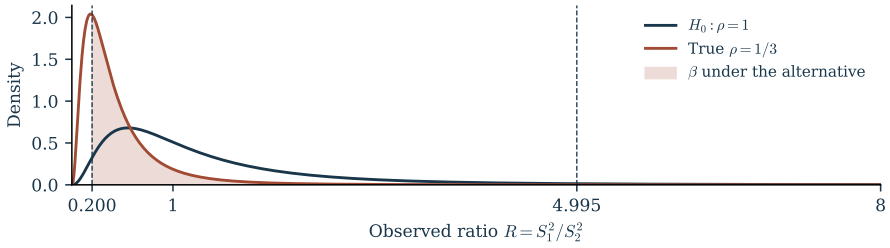
A cumulative distribution function, CDF, gives the probability to the **left** of a value. Subtract the probability below the lower bound from the probability below the upper bound:

$$\beta = P(W \leq 14.98473) - P(W \leq 0.600612).$$

In Excel, use **F.DIST(14.98473,7,7,TRUE) − F.DIST(0.600612,7,7,TRUE)**. TRUE asks for cumulative probability, not density. The two CDF values are approximately 0.998993 and 0.258659.

5. Subtract, then calculate power.

$\beta = 0.998993 - 0.258659 = \mathbf{0.740334}$. Power $= 1 - \beta = \mathbf{0.259666}$. Under this particular alternative, the test misses the difference about 74% of the time and detects it about 26% of the time.



Keep the test boundaries fixed. The shaded area is the probability of falling between them under the unequal-variance alternative, which is the Type II error probability.

Why is power low despite a real difference?

Each group has only eight observations, so sample variances fluctuate substantially. A stringent two-sided 5% rule has a broad nonrejection interval. A real difference is not a guarantee of detection. β is not the probability that H_0 is true and is not a p-value.

How α, confidence and decisions fit together

Chapter 10 lecture slides (open) Additional practice and connections

1. A larger α makes rejection easier

For the same data, null and alternative, suppose a test rejects at $\alpha = 0.05$. Its p-value is no larger than 0.05. If α increases to 0.10, the p-value is still no larger than α, so the test still rejects. The rejection region expands as α increases. The price is a higher allowed Type I error rate.

Reason through the inequality

If $p \leq 0.05$ and $\alpha > 0.05$, then $p \leq \alpha$. This establishes the conclusion without recomputing a statistic. Moving to a **smaller** α can change rejection to nonrejection.

2. More confidence means a wider interval

For the same data and interval method, reducing α increases confidence $1 - \alpha$ and widens the interval. A value excluded from a 95% interval might be included in a wider 99% interval. Exclusion from one interval does not guarantee exclusion from all wider intervals.

For example, a hypothetical parameter value 12 is outside (4,10) but inside (2,15). This is just a logical illustration of nested intervals, not the numerical interval for a particular sample.

3. Match interval and test levels

The equal-tailed two-sided test at significance α rejects H_0: $\theta = \theta_0$ when θ_0 lies outside the corresponding $100(1 - \alpha)\%$ confidence interval. For a one-population variance, θ_0 is σ_0^2. For a variance ratio under equality, θ_0 is 1.

Two-sided test level	Corresponding confidence interval
$\alpha = 0.10$	90%
$\alpha = 0.05$	95%
$\alpha = 0.01$	99%

A 95% interval can directly resolve the matching 5% two-sided test, not a 10% test and not automatically a one-sided test. If a problem asks for a 95% interval and a test at $\alpha = 0.10$, calculate the test with its own cutoffs.

4. An observed difference is not the same as evidence of a population difference

A sample variance ratio unequal to 1 is common even when the true ratio is 1. A test asks whether the observed ratio is sufficiently extreme relative to its sampling distribution. A nonsignificant result is compatible both with equality and with differences the test lacked power to detect.

A problem-solving checklist

Chapter 10 lecture slides (open) Additional practice and connections

These notes cover the methods in the Chapter 10 lecture slides and the accompanying Chapter 10 problem-set topics. Always use the exact data, α and confidence level on your current assigned version.

If your task asks you to...	Follow this route
Explain a true/false claim about α or confidence	Page 23: compare rejection regions or nested intervals; give a reason, not only T/F.
Test precision from raw measurements	Pages 3, 9–11: calculate s^2, square the claimed σ, select the tail.
Find a confidence interval for σ or σ^2	Pages 7–8: chi-square endpoints; square-root only if asked for σ.
Compare variance using sums of squared observations	Page 18, then pages 12–17: recover s^2 correctly and keep numerator order.
Compare both variances and means	Pages 16, 19–20: separate hypotheses and procedures; state the pooled-t assumptions.
Find the probability of a Type II error for an F test	Pages 21–22: null nonrejection region; true ρ; rescale both bounds; subtract CDFs.

Before you turn in a solution

- Have I identified the population parameter, not merely a sample summary?
- Did I square a standard deviation before using it as a variance?
- Did I state independence and normality assumptions, and equal variances if using pooled t?
- Did I keep the sample order and the degrees of freedom in the same order?
- Did I use $\alpha/2$ only for an equal-tailed two-sided test or interval?
- Does my conclusion name variance, standard deviation or mean accurately?
- For β, did I evaluate under the specified true alternative rather than under H_0?

A useful habit

Write one sentence after each formula: "This number represents ..." If you cannot explain what the statistic, cutoff, endpoint or probability means, return to the definition before doing more arithmetic.

Formulas and software, with conventions

Chapter 10 lecture slides (open) Additional practice and connections

All χ^2 and F critical values below use right-tail areas, as in the lecture slides. Let $\nu = \mathrm{n} - 1$, $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$, and $\mathrm{R} = s_1^2/s_2^2$.

One normal population

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^{\,2}}{n-1}, \qquad Q = \frac{(n-1)s^2}{\sigma_0^2}.$$

$$\mathrm{CI}(\sigma^2) = \left[\frac{(n-1)s^2}{\chi^2_{\nu,\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{\nu,1-\alpha/2}}\right].$$

Two independent normal populations

$$\mathrm{CI}(\rho) = \left[\frac{R}{F_{\nu_1,\nu_2,\alpha/2}}, RF_{\nu_2,\nu_1,\alpha/2}\right], \qquad \rho = \frac{\sigma_1^2}{\sigma_2^2}.$$

Let G_{ν_1,ν_2} denote the F cumulative distribution function. Let L and U be the lower and upper nonrejection bounds chosen under H_0.

$$\beta = G_{\nu_1,\nu_2}(U/\rho) - G_{\nu_1,\nu_2}(L/\rho),$$

Power is $1 - \beta$. Square-root the variance-ratio interval endpoints if you want σ_1/σ_2 instead of σ_1^2/σ_2^2.

Excel functions: do not mix critical values and probabilities

Need	Excel expression
Chi-square upper-tail cutoff	CHISQ.INV.RT(a, df)
Chi-square lower-tail probability at q	CHISQ.DIST(q, df, TRUE)
Chi-square upper-tail probability at q	CHISQ.DIST.RT(q, df)
F upper-tail cutoff	F.INV.RT(a, df1, df2)
F lower-tail probability at x	F.DIST(x, df1, df2, TRUE)
F upper-tail probability at x	F.DIST.RT(x, df1, df2)
Two-sided t cutoff for total α	T.INV.2T(alpha, df)

Use the full stored precision for intermediate arithmetic. Differences in the last digits from lecture tables are expected when using more accurate software quantiles. A probability must lie between 0 and 1; a variance and an F ratio cannot be negative.