
Lecture 15: One-Sample Hypothesis Tests

STAT 515 — Statistical Methods I

Dr. Yen-Yi Ho

Department of Statistics, University of South Carolina

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

Last time we built the **machinery** of a hypothesis test: H_0 and H_a , Type I and Type II errors, rejection regions, and p -values.

Today

We put that machinery to work on **one sample**. Four parameters, four test statistics — but **one and the same six-step recipe**.

Parameter	Situation	Test statistic
μ	large n ($n > 30$)	z
μ	small n , normal population	t_{n-1}
p	large n	z
σ^2	normal population	χ_{n-1}^2

The Six Elements of a Hypothesis Test

1. **Null hypothesis H_0** — a claim about the value of a parameter; the status quo. By convention we put the *equality* under H_0 .
2. **Alternative hypothesis H_a** — the research hypothesis; the claim we accept only if the data strongly indicate it.
3. **Test statistic** — a sample statistic used to decide whether to reject H_0 , computed *assuming H_0 is true*.
4. **Rejection region** — the values of the test statistic that lead us to reject H_0 in favour of H_a .
5. **Conclusion** — the statistical decision plus a sentence in plain English.
6. **Assumptions** — what we assumed about the population (normality, random sampling, sample size).

General form

$$\text{test statistic} = \frac{\text{estimate} - \text{hypothesized value}}{\text{standard error of the estimate}}$$

- It measures **how many standard errors** the estimate sits away from the value claimed by H_0 .
- Large in absolute value $\implies$ the data are far from what H_0 predicts $\implies$ evidence against H_0 .
- Everything today is this one formula with a different numerator and a different standard error.

Two Equivalent Decision Rules

Rejection-region rule

Fix α *first*. Reject H_0 when the computed statistic falls in the rejection region.

p -value rule

Compute

$$p = P_{H_0}(\text{as extreme or more})$$

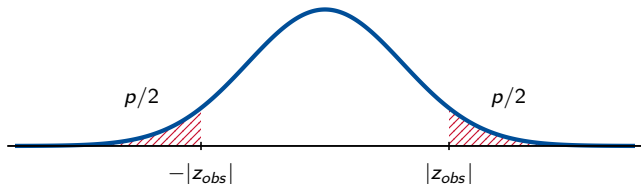
Reject H_0 when $p < \alpha$.

Note

These two rules always give the same answer. The p -value says *how much* evidence there is.

Computing the p -Value

- **One-tailed test:** the p -value is the tail area beyond the computed test statistic, **in the direction of H_a** .
- **Two-tailed test:** the p -value is **twice** the tail area beyond the computed test statistic, in the direction that favours H_a .



1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

Large-Sample Test on the Population Mean

For a **large** random sample we test μ :

Test statistic

$$H_0 : \mu = \mu_0 \qquad Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim N(0, 1) \text{ under } H_0$$

If σ happens to be known, use $\sigma/\sqrt{n}$ in the denominator.

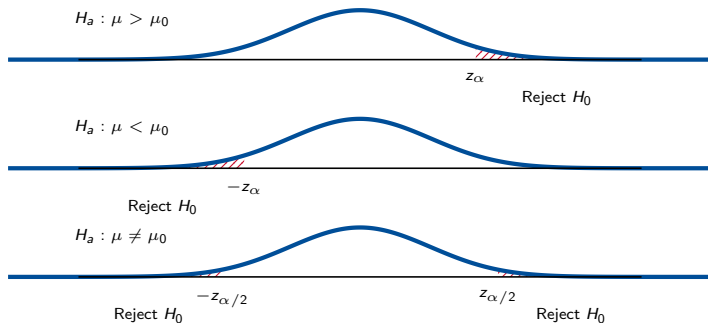
Alternative	Rejection region
-------------	------------------

$H_a : \mu > \mu_0$	$\{ Z > z_\alpha \}$
---------------------	----------------------

$H_a : \mu < \mu_0$	$\{ Z < -z_\alpha \}$
---------------------	-----------------------

$H_a : \mu \neq \mu_0$	$\{ Z > z_{\alpha/2} \}$
------------------------	----------------------------

The Three Rejection Regions, Drawn



A one-tailed alternative puts *all* of α in one tail; the two-tailed alternative splits it into two tails of area $\alpha/2$.

Assumptions

- A **random sample** from the target population is required.
 - The sample size n is **large** (i.e. $n > 30$).
-
- Why large n ? The **Central Limit Theorem** makes $\bar{X}$ approximately normal *whatever* the population looks like, and makes s a reliable stand-in for σ .
 - **No normality assumption on the population is needed here** — that is exactly what the large sample buys us.

Example 1: Truck Tire Mileage

Example 1 (one-sided)

A trucking firm doubts a tire manufacturer's claim that a certain type of truck tire averages *at least* 28,000 miles before a blowout. To check the claim, the firm put 40 of these tires on its trucks and obtained a mean life of 27,563 miles with a standard deviation of 1,324 miles. What conclusion should they reach if they allow a Type I error rate of 1%?

Given: $n = 40$, $\bar{x} = 27,563$, $s = 1,324$, $\mu_0 = 28,000$, $\alpha = .01$.

μ = the true mean mileage before a blowout.

Example 1: Hypotheses and Test Statistic

Hypotheses

$$H_0 : \mu = 28,000 \ (\mu \geq 28,000) \quad \text{vs.} \quad H_a : \mu < 28,000$$

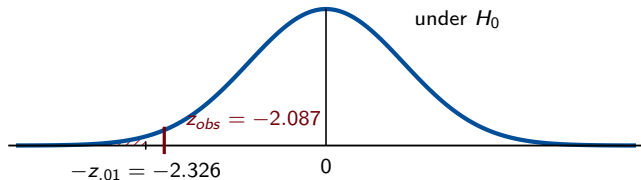
Test statistic

$$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{27,563 - 28,000}{1,324/\sqrt{40}} = \frac{-437}{209.34} = -2.087$$

The firm's doubt goes in *one* direction, so this is a lower-tailed test.

Example 1: Rejection Region and Decision

Rejection region: $\{ Z < -z_\alpha \} = \{ Z < -z_{.01} \} = \{ Z < -2.326 \}$



Since $-2.087 \not< -2.326$, the statistic is **not** in the rejection region: **fail to reject H_0** .

Example 1: p -Value and Conclusion

$$p\text{-value} = P_{H_0}(Z < -2.087) = \mathbf{0.0184} \quad 0.0184 \not< 0.01$$

Same decision, either way

The p -value rule and the rejection-region rule agree: **fail to reject** H_0 at $\alpha = .01$.

Conclusion

At the .01 significance level we **do not** have sufficient evidence to conclude that the true mean mileage before a blowout is less than 28,000 miles.

Careful: with $\alpha = .05$ we *would* have rejected. The significance level is part of the conclusion.

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

What Goes Wrong When n Is Small

- With σ known, $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ exactly, provided the population is normal.
- In practice σ **is never known**, so we substitute s .
- For large n that substitution costs almost nothing. For **small** n , s is itself quite variable, and the ratio has **more spread** than a standard normal.

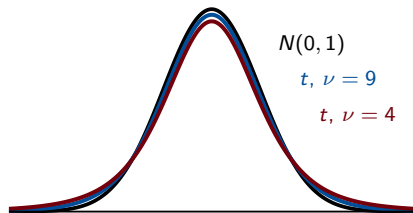
Student's t

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1} \quad (\text{population normal})$$

Properties of the t Distribution

- Mean = median = mode = 0.
- Symmetric about 0.
- Ranges from $-\infty$ to $+\infty$.
- A **family** indexed by the **degrees of freedom** $\nu = n - 1$.
- Heavier tails than $N(0, 1)$; as $\nu \rightarrow \infty$ the t curve approaches the normal curve.

Fewer degrees of freedom $\Rightarrow$ flatter centre, fatter tails $\Rightarrow$ **larger critical values**.



Critical Values: t Versus z

One-sided area	0.025	0.01	0.005
Two-sided area	0.05	0.02	0.01
$\nu = 1$	12.71	31.82	63.66
$\nu = 10$	2.23	2.76	3.17
$\nu = 30$	2.04	2.46	2.75
$\nu = 500$	1.97	2.33	2.59
z (i.e. $\nu = \infty$)	1.96	2.33	2.58

- The t critical values are always **larger** than the matching z values — the price of estimating σ .
- By $\nu = 30$ the gap is already small — which is why “ $n > 30$ ” lets us use z .

Small-Sample Test on the Population Mean

Test statistic

$$H_0 : \mu = \mu_0 \qquad t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \qquad \text{with } df = n - 1$$

Alternative	Rejection region
-------------	------------------

$H_a : \mu > \mu_0$	$\{ t > t_{n-1, \alpha} \}$
---------------------	-----------------------------

$H_a : \mu < \mu_0$	$\{ t < -t_{n-1, \alpha} \}$
---------------------	------------------------------

$H_a : \mu \neq \mu_0$	$\{ t > t_{n-1, \alpha/2} \}$
------------------------	---------------------------------

Identical in shape to the z test — only the reference distribution has changed.

Conditions for a Valid Small-Sample t Test for μ

Assumptions

- A **random sample** from the target population is required.
- The population from which the sample is selected is **approximately normally distributed**.

Note

The t test is **robust** to the normality assumption: moderate departures from normality do not badly distort the test, especially if the departure is roughly symmetric. Strong skewness or outliers with a very small n **are** a problem.

In practice: look at a dotplot, boxplot, or normal QQ plot of the data before trusting a small-sample t test.

Example 2: Typing Speed

Example 2 (one-sided, small sample)

A random sample of 12 students in a high school typing class averaged 73.2 words per minute with a standard deviation of 7.9 on a typing test. What can we conclude, at the .05 level of significance, regarding the claim that students at this school average *less than* 75 words per minute? What are the assumptions for the validity of this test?

Given: $n = 12$, $\bar{x} = 73.2$, $s = 7.9$, $\mu_0 = 75$, $\alpha = .05$, $df = 11$.

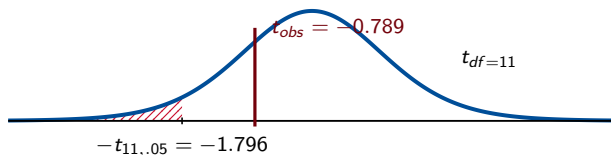
μ = the true average number of typing words per minute.

Example 2: Test Statistic and Rejection Region

$$H_0 : \mu = 75 \quad \text{vs.} \quad H_a : \mu < 75$$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{73.2 - 75}{7.9/\sqrt{12}} = \frac{-1.8}{2.281} = \mathbf{-0.789}$$

$$\text{Rejection region: } \{ t < -t_{11,.05} \} = \{ t < -1.796 \}$$



Example 2: Decision, p -Value, Assumptions

- Since $-0.789 \not< -1.796$, the statistic is **not** in the rejection region: **fail to reject H_0** at the .05 level.
- $p\text{-value} = P_{H_0}(t_{11} < -0.789) = \mathbf{0.223} > .05$ — the same decision.

Conclusion

We do **not** have sufficient evidence to conclude that students at this high school average less than 75 words per minute.

Assumptions

1. The 12 students are a random sample of students at this high school.
2. The number of typing words per minute follows a normal distribution.

Example 3: Weight of a Dog Breed

Example 3 (two-sided, raw data)

Five mature dogs of a certain breed weigh 66, 63, 64, 62, and 65 pounds. A kennel club claims that the average weight for this breed is 60 pounds. Using the .05 level of significance, do we have reason to doubt this claim? What are the assumptions for the validity of this test?

$$\bar{x} = \frac{66 + 63 + 64 + 62 + 65}{5} = 64, \quad s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = \frac{10}{4} = 2.5, \quad s = 1.581$$

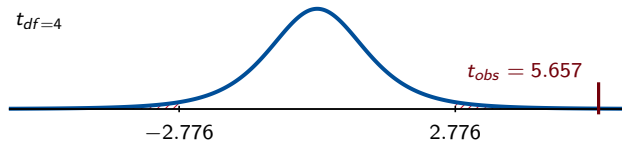
μ = the true mean weight for this breed. $df = n - 1 = 4$.

Example 3: Test Statistic and Rejection Region

$$H_0 : \mu = 60 \quad \text{vs.} \quad H_a : \mu \neq 60$$

$$t = \frac{64 - 60}{1.581/\sqrt{5}} = \frac{4}{0.707} = \mathbf{5.657}$$

$$\text{Rejection region: } \{ |t| > t_{4, .025} \} = \{ |t| > 2.776 \}$$



Example 3: Decision, p -Value, Assumptions

- Since $|5.657| > 2.776$, the statistic **falls in the rejection region: reject H_0** at the .05 level.
- $p\text{-value} = 2 P_{H_0}(t_4 > 5.657) = \mathbf{0.0048} < .05$ — the same decision.

Conclusion

We have sufficient evidence at the .05 level that the true mean weight for this breed is **significantly different from 60 pounds**.

Assumptions and a caution

Random sample of mature dogs of this breed; weight is normally distributed.
If in fact $\mu = 60$, we have just committed a **Type I error** — which happens 5% of the time by design.

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

The Birth Weight Setting

Setting

Birth weights are normally distributed. Birth weights are obtained on a sample of 10 infants; the sample mean is 2500 grams.

We ask the *same data* four different questions:

Question	H_a	Known?
1 different from 3000g?	$\mu \neq 3000$	$\sigma = 1000$ known
2 less than 3000g?	$\mu < 3000$	$\sigma = 1000$ known
3 greater than 3000g?	$\mu > 3000$	$\sigma = 1000$ known
4 different from 3000g?	$\mu \neq 3000$	σ unknown, $s = 900$

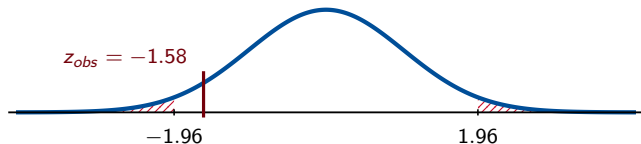
Only the **direction** of H_a and **which distribution** we compare against change.

Scenario 1: Two-Sided, σ Known

$$H_0 : \mu = 3000 \quad \text{vs.} \quad H_a : \mu \neq 3000, \quad \alpha = .05$$

$$z_{obs} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{2500 - 3000}{1000/\sqrt{10}} = \frac{-500}{316.2} = -1.58$$

Decision rule: reject H_0 if $z_{obs} \geq 1.96$ or $z_{obs} \leq -1.96$.



$z_{obs} = -1.58$ is **not** in the rejection region $\Rightarrow$ **fail to reject H_0** .

Scenario 1: p -Value and Interval

$$p\text{-value} = P(Z < -1.58) + P(Z > 1.58) = 2(0.0571) = \mathbf{0.1142}$$

What $p = 0.1142$ means

In about 11 out of 100 samples we would get, **due to chance alone**, a sample mean as extreme or more extreme (in either direction) than the one we observed — **when H_0 is true**. That is not rare enough to doubt H_0 .

A 95% confidence interval:

$$\bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 2500 \pm 1.96 \cdot \frac{1000}{\sqrt{10}} = 2500 \pm 620 = (\mathbf{1880}, \mathbf{3120})$$

The interval **covers** 3000 — the same inference once more.

Scenarios 2 and 3: One-Sided Tests, Same Data

Scenario 2: is $\mu < 3000$?

$H_0 : \mu \geq 3000$ vs. $H_a : \mu < 3000$

Reject if $z_{obs} \leq -1.645$

$z_{obs} = -1.58$: **fail to reject**

$p = P(Z < -1.58) = \mathbf{0.0571}$

Scenario 3: is $\mu > 3000$?

$H_0 : \mu \leq 3000$ vs. $H_a : \mu > 3000$

Reject if $z_{obs} \geq 1.645$

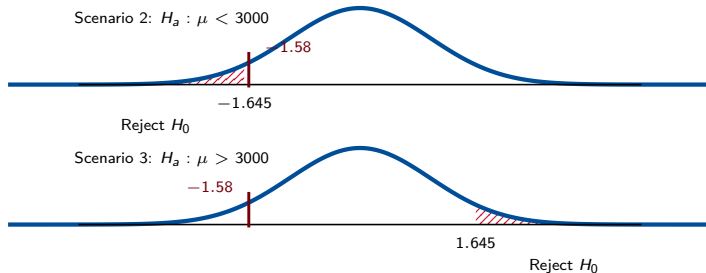
$z_{obs} = -1.58$: **fail to reject**

$p = P(Z > -1.58) = \mathbf{0.9429}$

Read the two p -values carefully

The *same* z_{obs} gives $p = .0571$ against one alternative and $p = .9429$ against the other, and $.0571 + .9429 = 1$. **A p -value is only meaningful once H_a has been fixed** — and H_a must be chosen *before* seeing the data.

Scenarios 2 and 3, Drawn



In Scenario 3 the observed value sits on the *opposite* side from the rejection region — the data point away from H_a entirely.

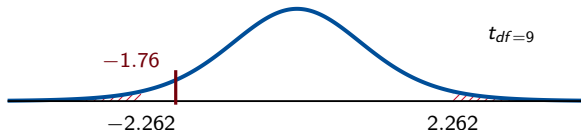
Scenario 4: Two-Sided, σ Unknown

Now σ is unknown; the sample standard deviation is $s = 900$ grams.

$$H_0 : \mu = 3000 \quad \text{vs.} \quad H_a : \mu \neq 3000$$

$$t_{obs} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{2500 - 3000}{900/\sqrt{10}} = \frac{-500}{284.6} = \mathbf{-1.76}, \quad df = 9$$

Decision rule: reject H_0 if $t_{obs} \leq -2.262$ or $t_{obs} \geq 2.262$.



Scenario 4: Decision, p -Value, Interval

- $t_{obs} = -1.76$ is **not** in the rejection region $\Rightarrow$ **fail to reject H_0** at $\alpha = .05$ with $df = 9$.
- $p\text{-value} = P(t_9 < -1.76) + P(t_9 > 1.76) = \mathbf{0.113} > .05$.
- 95% CI: $2500 \pm 2.262(284.6) = 2500 \pm 644 = (\mathbf{1856}, \mathbf{3144})$, which **covers** 3000.

Conclusion

Fail to reject H_0 : this sample could well have come from a population whose mean birth weight is 3000 grams.

Note the t interval is *wider* than the z interval of Scenario 1 even though $s < \sigma$ here — the critical value grew from 1.96 to 2.262.

All Four Scenarios at a Glance

	H_a	Statistic	Critical value	p -value
1	$\mu \neq 3000$ (σ known)	$z = -1.58$	± 1.96	.1142
2	$\mu < 3000$ (σ known)	$z = -1.58$	-1.645	.0571
3	$\mu > 3000$ (σ known)	$z = -1.58$	$+1.645$	.9429
4	$\mu \neq 3000$ ($s = 900$)	$t_9 = -1.76$	± 2.262	.113

Key point

Every scenario fails to reject H_0 — but for different reasons and with very different amounts of evidence. Scenario 2 comes close ($p = .057$); Scenario 3 is not even in the running.

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

Large-Sample Test on a Population Proportion

For a **large** random sample we test p :

Test statistic

$$H_0 : p = p_0 \qquad Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}, \qquad \hat{p} = \frac{x}{n}, \quad q_0 = 1 - p_0$$

Alternative	Rejection region
$H_a : p > p_0$	$\{ Z > z_\alpha \}$
$H_a : p < p_0$	$\{ Z < -z_\alpha \}$
$H_a : p \neq p_0$	$\{ Z > z_{\alpha/2} \}$

Watch the standard error: it uses p_0 , the *hypothesized* value — not $\hat{p}$, because everything is computed assuming H_0 is true.

Assumptions

- A **random sample** from the target population is required.
 - The sample size n is **large**: both $np_0 \geq 15$ and $nq_0 \geq 15$.
-
- The condition guarantees the binomial count is well approximated by a normal curve, so that $\hat{p}$ is approximately normal.
 - For the **confidence interval** the analogous check uses $\hat{p}$: $n\hat{p} \geq 15$ and $n\hat{q} \geq 15$, because there is no hypothesized p_0 to lean on.

Example 4: Concern About Drug Trafficking

Example 4 (one-sided)

A city council member claims that 45% of her constituency is “very concerned” about drug trafficking. To see whether this claim is too low, a sample of 265 citizens was taken and 135 of them indicated that they were “very concerned” about drug trafficking. Test at the .05 level of significance.

p = the true proportion of her constituency who are “very concerned” about drug trafficking.

$$n = 265, \quad x = 135, \quad \hat{p} = \frac{135}{265} = 0.509, \quad p_0 = 0.45, \quad \alpha = .05$$

Check: $np_0 = 119.25 \geq 15$ and $nq_0 = 145.75 \geq 15$. ✓

Example 4: Test, Decision, Conclusion

$$H_0 : p = 0.45 \quad \text{vs.} \quad H_a : p > 0.45$$
$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{0.509 - 0.45}{\sqrt{\frac{(0.45)(0.55)}{265}}} = \frac{0.059}{0.0306} = \mathbf{1.931}$$

Rejection region: $\{ Z > z_{.05} \} = \{ Z > 1.645 \}$

Since $1.931 > 1.645$ the statistic **falls in the rejection region**, so we **reject** H_0 . $p\text{-value} = P(Z > 1.931) = \mathbf{0.027} < .05$. ✓

Conclusion

At the .05 level we have sufficient evidence that the true proportion of citizens who are “very concerned” is **significantly larger than 45%**.

Example 5: The Answering Machine

Example 5 (one-sided, $\alpha = .10$)

I believe that callers to my home fail to leave a message on my answering machine *at least* 40% of the time. In order to test this, I refuse to answer my phone on 50 random calls and observe whether or not the caller leaves a message. If 15 callers do not leave a message, is this reason to doubt my claim at the .10 level of significance?

p = the true proportion of callers who fail to leave a message.

$$n = 50, \quad x = 15, \quad \hat{p} = \frac{15}{50} = 0.3, \quad p_0 = 0.4, \quad \alpha = .10$$

Check: $np_0 = 20 \geq 15$ and $nq_0 = 30 \geq 15$. ✓

Example 5: Test, Decision, Conclusion

$$H_0 : p = 0.4 \text{ (} p \geq 0.4 \text{)} \quad \text{vs.} \quad H_a : p < 0.4$$

$$Z = \frac{0.3 - 0.4}{\sqrt{\frac{(0.4)(0.6)}{50}}} = \frac{-0.1}{0.0693} = \mathbf{-1.443}$$

$$\text{Rejection region: } \{ Z < -z_{.10} \} = \{ Z < -1.282 \}$$

Since $-1.443 < -1.282$, the statistic **falls in the rejection region: reject H_0** .
 $p\text{-value} = P(Z < -1.443) = \mathbf{0.074} < .10. \checkmark$

Conclusion

At the .10 level we have sufficient evidence to **doubt the claim** that at least 40% of callers fail to leave a message.

At $\alpha = .05$ the cutoff would be -1.645 and we would *not* reject — the chosen α matters.

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

Test on a Population Variance

The key fact

For a random sample from a **normal** population,

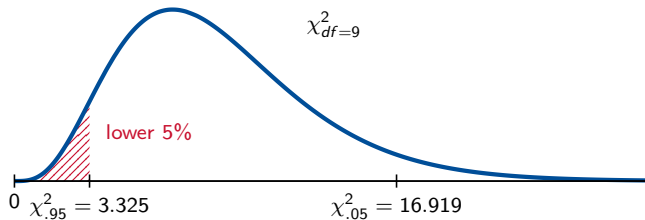
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

Test statistic

$$H_0 : \sigma^2 = \sigma_0^2 \qquad \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \qquad df = n - 1$$

The χ^2 distribution is skewed right and lives on $[0, \infty)$ — so, unlike z and t , its critical values are *not* symmetric: look up each tail separately.

The Chi-Square Distribution ($df = 9$)



Alternative	Rejection region
$H_a : \sigma^2 > \sigma_0^2$	$\{ \chi^2 > \chi^2_{n-1, \alpha} \}$
$H_a : \sigma^2 < \sigma_0^2$	$\{ \chi^2 < \chi^2_{n-1, 1-\alpha} \}$
$H_a : \sigma^2 \neq \sigma_0^2$	$\{ \chi^2 < \chi^2_{n-1, 1-\alpha/2} \text{ or } \chi^2 > \chi^2_{n-1, \alpha/2} \}$

Assumptions

- A **random sample** from the target population is required.
- The sampled population has a distribution that is **approximately normal**.

A serious warning

The chi-square test for the variance is **not robust** to violation of the normality assumption — unlike the t test for the mean. A heavy-tailed or skewed population can make this test badly wrong even for large n . **Check normality before you use it.**

Example 6: Fill Variation in a Cannery

Example 6

A quality control supervisor in a cannery knows that the exact amount each can contains will vary. Regulatory agencies specify that the standard deviation of the amount of fill in 16-ounce cans should be less than .1 ounce. To determine whether the process is meeting this specification, the supervisor randomly selects 10 cans and weighs the contents of each. If the sample standard deviation is .0412, do we have sufficient evidence to conclude that the true standard deviation σ of the fill is less than .1 ounce?

σ = the true standard deviation of the amount of fill. $n = 10$, $s = .0412$,
 $\sigma_0 = .1$, $\alpha = .05$, $df = 9$.

Example 6: Hypotheses and Test Statistic

$$H_0 : \sigma = .1 \iff \sigma^2 = (.1)^2 = .01 \quad \text{vs.} \quad H_a : \sigma < .1 \iff \sigma^2 < .01$$

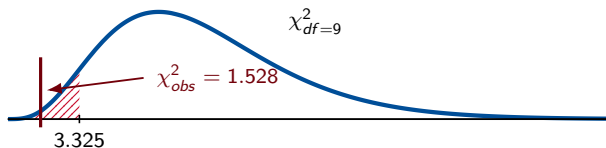
Test statistic

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(10-1)(.0412)^2}{(.1)^2} = \frac{9(.0016974)}{.01} = \mathbf{1.528}$$

$$\text{Rejection region: } \{ \chi^2 < \chi_{9, 1-\alpha}^2 \} = \{ \chi^2 < \chi_{9, .95}^2 \} = \{ \chi^2 < 3.325 \}$$

A one-sided test about a *small* variance puts the whole rejection region in the *lower* tail.

Example 6: Decision and Conclusion



- Since $1.528 < 3.325$, the statistic **falls in the rejection region: reject H_0** at $\alpha = .05$.
- $p\text{-value} = P_{H_0}(\chi_9^2 \leq 1.528) = \mathbf{0.003} < .05$. ✓

Conclusion

We have sufficient evidence that the true standard deviation of the fill is **significantly less than .1 ounce** — the process meets the specification.

1. Review: the elements of a hypothesis test
2. Large-sample z test for a population mean μ
3. Small-sample t test for a population mean μ
4. Birth weight example: all the scenarios side by side
5. Large-sample z test for a population proportion p
6. Test for a population variance σ^2
7. Summary: which test, when

Choosing the Right One-Sample Test

Parameter	Conditions	Test statistic	Reference
μ	normal pop., σ known	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$N(0, 1)$
μ	$n > 30$ (any population)	$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$N(0, 1)$
μ	small n , normal pop., σ unknown	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$	t_{n-1}
p	$np_0 \geq 15, nq_0 \geq 15$	$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0/n}}$	$N(0, 1)$
σ^2	normal population	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	χ_{n-1}^2

The recipe never changes

State H_0 and $H_a \rightarrow$ check the conditions $\rightarrow$ compute the statistic $\rightarrow$ compare with the rejection region (or the p -value with α) $\rightarrow$ **write a conclusion in words.**

Three Views, One Inference

- Does the observed test statistic fall within the **rejection region** of the sampling distribution?
- Is the associated **p-value** less than the significance level (usually $\alpha = 0.05$)?
- Does the $100(1 - \alpha)\%$ **confidence interval** cover the hypothesized value?

Key point

All three procedures result in the same inference — provided the pieces match: a two-sided test at level α pairs with a two-sided $100(1 - \alpha)\%$ interval.

Common Mistakes to Avoid

- **Choosing H_a after seeing the data.** The direction of the alternative must come from the research question.
- **Using $\hat{p}$ in the standard error** of the test for p . Under H_0 the standard error uses p_0 .
- **Saying “accept H_0 .”** We **fail to reject**: absence of evidence is not evidence of absence.
- **Forgetting $df = n - 1$** for the t and χ^2 tests.
- **Confusing statistical with practical significance.** A huge n can make a trivial difference “significant.”
- **Skipping the conditions.** Especially normality for the χ^2 test for σ^2 , which is **not** robust.

- The **large-sample z test for μ** needs only a random sample and $n > 30$; the CLT does the rest.
- The **small-sample t test for μ** replaces σ by s and the normal curve by t_{n-1} ; it requires an approximately normal population but is robust to mild departures.
- The **large-sample z test for p** uses the hypothesized p_0 in the standard error and needs $np_0 \geq 15$, $nq_0 \geq 15$.
- The **χ^2 test for σ^2** uses $(n-1)s^2/\sigma_0^2$ with asymmetric critical values and is **not** robust to non-normality.
- One-sided versus two-sided changes the rejection region and the p -value — never the data.

Next: comparing two populations.