One-Sample z and t Tests for a Population Mean

AP Statistics · Mr. Merrick · February 21, 2026

Assume we have a random sample of size n from a population with an unknown distribution. We want to test a claim about the parameter μ (the population mean).

We compute $\overline{x}$ and s from the sample. Sometimes the population standard deviation σ is known. When σ is unknown (most common), we estimate it using s.

z-test for μ (when σ is known)

Step 1 — State hypotheses

$$H_0 : \mu = \mu_0$$

$$H_a : \mu < \mu_0 \quad \text{or} \quad \mu > \mu_0 \quad \text{or} \quad \mu \neq \mu_0$$

Step 2 — Check conditions

- Random: random sample or randomized experiment.
- Independence: $n \leq 0.10N$ if sampling without replacement.
- Normal/Large Sample: population Normal, or n large enough for CLT.

If conditions are satisfied, then under H_0,

$$Z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1).$$

Step 3 — Compute the p-value

Match to H_a:

- Left-tailed ($\mu < \mu_0$): $p = P(Z \leq z)$
- Right-tailed ($\mu > \mu_0$): $p = P(Z \geq z)$
- Two-sided ($\mu \neq \mu_0$): $p = 2P(Z \geq |z|)$

Step 4 — Conclude

At significance level α:

- If $p \leq \alpha$, reject H_0.
- If $p > \alpha$, fail to reject H_0.

$$\boxed{Z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}}$$

t-test for μ (when σ is unknown)

Step 1 — State hypotheses

$$H_0 : \mu = \mu_0$$

$$H_a : \mu < \mu_0 \quad \text{or} \quad \mu > \mu_0 \quad \text{or} \quad \mu \neq \mu_0$$

Step 2 — Check conditions

- Random: random sample or randomized experiment.
- Independence: $n \leq 0.10N$ if sampling without replacement.
- Normal/Large Sample: if n is small, sample data are roughly symmetric with no outliers.

When σ is unknown, we estimate it with s. Under H_0,

$$T = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}.$$

Step 3 — Compute the p-value

Match to H_a (use $df = n - 1$):

- Left-tailed: $p = P(T \leq t)$
- Right-tailed: $p = P(T \geq t)$
- Two-sided: $p = 2P(T \geq |t|)$

Step 4 — Conclude

At significance level α:

- If $p \leq \alpha$, reject H_0.
- If $p > \alpha$, fail to reject H_0.

$$\boxed{T = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}, \quad df = n - 1}$$

When n is small and the population distribution is unknown

The t-test requires that the population distribution be approximately Normal when n is small. In practice, because we do not see the entire population, we examine the sample data. If the sample distribution is roughly symmetric with no strong skewness or outliers, it is reasonable to proceed with a t-test.

In AP Statistics, when n is small, we examine the sample data directly. If the sample distribution is:

- roughly symmetric,
- free of strong skewness,
- and free of outliers,

then it is reasonable to proceed with a t-test.

If the sample shows strong skewness or clear outliers and n is small, the t-test may not be reliable.

AP Exam Tip
When n is small, and the population isn't stated as normal, describe the sample distribution. If a graph is provided, reference it directly in your explanation. If raw data are given, use your calculator to create a graph (such as a histogram), then describe the shape.

A strong AP response includes words such as:

- roughly symmetric
- no obvious outliers
- not strongly skewed

Do not simply write "the calculator says it is Normal." Always justify using the shape of the distribution.

Example 1

A company that manufactures protein bars claims that the mean protein content is 20 grams per bar. A nutrition inspector suspects the bars may contain less protein than advertised. A random sample of 15 bars is tested. The sample mean protein content is 19.2 grams with a standard deviation of 1.5 grams. Does the sample provide convincing evidence at the $\alpha = 0.05$ level that the true mean protein content is less than 20 grams per bar?

(a) State the appropriate hypotheses and define the parameter.

Solution. Let μ = the true mean protein content (in grams) of all protein bars produced by this company.

$$H_0 : \mu = 20$$
$$H_a : \mu < 20$$

(b) Check the conditions for inference.

Solution. Random: The problem states that a random sample of 15 bars was selected.

Independence: The sample size of 15 bars is less than 10% of all protein bars produced, so the observations can be considered independent.

Normality: Because no graph or raw data are provided, we cannot verify normality; we proceed assuming the population distribution is approximately normal.

Since all conditions are reasonably satisfied, a one-sample t-test is appropriate.

(c) Perform the test at the $\alpha = 0.05$ level.

Solution.

$$T = \frac{19.2 - 20}{1.5/\sqrt{15}} = -2.07$$

$df = 14$

Using a t distribution with 14 degrees of freedom,

$$p = P(T < -2.07) \approx 0.028$$

Since $p = 0.028 < 0.05$, we reject H_0.

(d) Interpret the p-value in context.

Solution. If the true mean protein content were 20 grams, the probability of obtaining a sample mean of 19.2 grams or lower from a random sample of 15 bars is about 2.8%.

(e) Explain what a Type I error would mean in this context.

Solution. A Type I error would occur if we conclude that the true mean protein content is less than 20 grams when, in reality, the true mean protein content is exactly 20 grams.

(f) Explain what a Type II error would mean in this context.

Solution. A Type II error would occur if we fail to conclude that the true mean protein content is less than 20 grams when, in fact, the true mean protein content is less than 20 grams.

(g) Describe one way to increase the power of this test.

Solution. One way to increase the power of this test is to increase the sample size. A larger sample size reduces variability in the sampling distribution, making it easier to detect a true difference from the null value.

Example 2

A bottling plant fills soda cans with 12 ounces of soda. The known process standard deviation is $\sigma = 0.25$ ounces. A random sample of 36 cans has a mean of 12.09 ounces. Does the sample provide convincing evidence at the $\alpha = 0.01$ level that the machine is overfilling soda cans?

(a) State hypotheses.

Solution. Let μ = the true mean fill amount (in ounces) of soda cans produced by the machine.

$$H_0 : \mu = 12$$
$$H_a : \mu > 12$$

(b) Check conditions.

Solution. Random: The problem states that a random sample of 36 cans was selected.

Independence: The sample size of 36 cans is less than 10% of all cans produced, so independence is reasonable.

Normal/Large Sample: The sample size is $n = 36$, which is sufficiently large for the Central Limit Theorem to apply.

Since σ is known and the conditions are satisfied, a one-sample z-test is appropriate.

(c) Perform the test at $\alpha = 0.01$.

Solution.

$$Z = \frac{12.09 - 12}{0.25/\sqrt{36}} = 2.16$$

$$p = P(Z > 2.16) = 0.0154$$

Since $p = 0.0154 > 0.01$, we fail to reject H_0.

(d) Interpret the conclusion in context.

Solution. At the 0.01 significance level, there is insufficient evidence to conclude that the machine is overfilling soda cans.

(e) Define a Type I error in context.

Solution. A Type I error would occur if we conclude that the machine is overfilling cans when, in reality, the true mean fill amount is exactly 12 ounces.

(f) Suppose the true mean is 12.15 ounces. Explain what power represents.

Solution. In this context, power is the probability that the test correctly rejects H_0 and concludes that the machine is overfilling when the true mean fill amount is actually 12.15 ounces.

Example 3

A pharmaceutical company claims that a new medication results in an average recovery time of 8 days. A hospital study of 20 patients finds a mean recovery time of 9.1 days with standard deviation 2.4 days. Does the sample provide convincing evidence at the $\alpha = 0.05$ level that the true mean recovery time differs from 8 days?

(a) State hypotheses.

Solution. Let μ = the true mean recovery time (in days) for patients using the medication.

$$H_0 : \mu = 8$$
$$H_a : \mu \neq 8$$

(b) Perform the test at $\alpha = 0.05$.

Solution.

$$T = \frac{9.1 - 8}{2.4/\sqrt{20}} = 2.05$$

$df = 19$

Two-sided $p \approx 0.054$

Since $p = 0.054 > 0.05$, we fail to reject H_0.

(c) State the conclusion in context.

Solution. At the 0.05 significance level, there is insufficient evidence to conclude that the true mean recovery time differs from 8 days.

(d) Describe a Type II error in context.

Solution. A Type II error would occur if we fail to conclude that the true mean recovery time differs from 8 days when, in fact, the true mean recovery time is not equal to 8 days.

(e) If the true mean were 9.5 days instead of 8.5 days, would power be higher or lower? Explain.

Solution. Power would be higher if the true mean were 9.5 days because 9.5 is farther from the null value of 8. When the true parameter is farther from the null value, it is easier for the test statistic to fall into the rejection region.

Example 4

A school district claims the average SAT math score is 520. A random sample of 25 students has mean 508 and standard deviation 40. Does the sample provide convincing evidence at the $\alpha = 0.05$ level that the true mean SAT math score is below 520?

(a) State hypotheses.

Solution. Let μ = the true mean SAT math score for students in the district.

$$H_0 : \mu = 520$$
$$H_a : \mu < 520$$

(b) Conduct the test at $\alpha = 0.05$.

Solution.

$$T = \frac{508 - 520}{40/\sqrt{25}} = -1.50$$

$df = 24$

$$p = P(T < -1.50) \approx 0.073$$

Since $p = 0.073 > 0.05$, we fail to reject H_0.

(c) Interpret the conclusion in context.

Solution. At the 0.05 significance level, there is insufficient evidence to conclude that the true mean SAT math score in the district is below 520.

(d) What is the probability of committing a Type I error?

Solution. The probability of committing a Type I error is the significance level, $\alpha = 0.05$.

(e) Suppose the true mean is 500. Explain what power represents.

Solution. Power is the probability that this test correctly rejects H_0 and concludes that the mean SAT math score is below 520 when the true mean is actually 500.

(f) Describe two ways to increase the power of this test.

Solution. Two ways to increase power are:

1. Increase the sample size. A larger sample reduces standard error and makes it easier to detect a true difference.

2. Increase the significance level α, which enlarges the rejection region.