Two-Sample Intervals for the Difference of Means

AP Statistics · Mr. Merrick · February 22, 2026

We compare two population means by estimating the parameter $\mu_1 - \mu_2$.
From two independent random samples (or two randomized groups), we compute:

$$\overline{x}_1,\ s_1,\ n_1 \qquad \text{and} \qquad \overline{x}_2,\ s_2,\ n_2.$$

All intervals follow the same structure:

$$\boxed{(\overline{x}_1 - \overline{x}_2)\ \pm\ \text{(critical value)(standard error)}}.$$

1) Two-Sample z-interval

When to use: σ_1, σ_2 known (rare).
Check conditions:

- Random: each sample is from a random sample or randomized experiment.
- Independence:
  - within groups: $n_1 \le 0.10N_1$ and $n_2 \le 0.10N_2$ (if sampling w/o replacement)
  - between groups: the two samples/groups are independent
- Normal/Large Sample: each population is Normal, or each n is large enough for CLT (≥ 30).

If conditions are satisfied, then

$$(\overline{x}_1 - \overline{x}_2)\ \sim\ N\left(\mu_1 - \mu_2,\ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

Thus the interval is

$$\boxed{(\overline{x}_1 - \overline{x}_2)\ \pm\ z^\star\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

2) Two-Sample t-interval (Welch's)

When to use: σ_1, σ_2 unknown (typical).
Check conditions:

- Random: each sample is random or groups are randomized.
- Independence:
  - within groups: $n_1 \le 0.10N_1$ and $n_2 \le 0.10N_2$ (if sampling w/o replacement)
  - between groups: independent samples (or randomized groups)
- Normal/Large Sample:
  - if n_1 and/or n_2 are small: check each group's sample distribution is roughly symmetric with no outliers
  - if both are large (≥ 30): CLT supports the procedure

Standard error: $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.

Thus the interval is

$$\boxed{(\overline{x}_1 - \overline{x}_2)\ \pm\ t^\star\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $t^\star$ comes from a t distribution with Welch df:

$$df\ \approx\ \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}.$$

$$\boxed{\min\{n_1 - 1, n_2 - 1\} \le df \le n_1 + n_2 - 2}$$

3) Enrichment: Two-Sample t-interval with pooled variance (NOT required for AP)

Big idea: If the population variances are equal ($\sigma_1^2 = \sigma_2^2$), we can combine (pool) information from both groups to estimate the common variance.

Important: This is *not* needed for AP Statistics. The AP standard is Welch's two-sample t interval.

In more advanced settings, equality of variances might be assessed by:

- comparing sample spreads (s_1 vs. s_2), or sample variances (s_1^2 vs. s_2^2),
- using a formal procedure such as Levene's test (beyond AP).

Pooled standard deviation:

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

Pooled standard error:

$$SE_p = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Pooled interval:

$$\boxed{(\overline{x}_1 - \overline{x}_2) \ \pm \ t^{\star}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad df = n_1 + n_2 - 2}$$

For AP Stats: Use Welch's approximation unless the problem explicitly explores this concept (possible on an investigative task).

Example 1

A manufacturer compares the fill amounts of two machines. From historical calibration, the population standard deviations are known:

$$\sigma_1 = 1.2 \text{ mL}, \qquad \sigma_2 = 1.6 \text{ mL}.$$

A random sample of $n_1 = 40$ bottles from Machine 1 has mean $\overline{x}_1 = 502.3$ mL. A random sample of $n_2 = 35$ bottles from Machine 2 has mean $\overline{x}_2 = 500.9$ mL. Construct and interpret a 95% confidence interval for $\mu_1 - \mu_2$ (Machine 1 minus Machine 2).

Solution. Step 1 — State

Let μ_1 be the true mean fill amount (in mL) for all bottles produced by Machine 1, and let μ_2 be the true mean fill amount (in mL) for all bottles produced by Machine 2.

We will construct a 95% confidence interval for the difference in population means, $\mu_1 - \mu_2$ (Machine 1 minus Machine 2).

Step 2 — Justify

We will use a two-sample z interval for $\mu_1 - \mu_2$ because the population standard deviations are known ($\sigma_1 = 1.2$ mL and $\sigma_2 = 1.6$ mL).

- **Random:** The problem states that a random sample of $n_1 = 40$ bottles was selected from Machine 1 and a random sample of $n_2 = 35$ bottles was selected from Machine 2.
- **Independent:**
  - **Within groups:** Because the samples are taken from production runs that are much larger than 40 and 35 bottles, it is reasonable to assume $n_1 = 40 \le 0.10N_1$ and $n_2 = 35 \le 0.10N_2$ (10% condition), so observations within each sample are approximately independent.
  - **Between groups:** Bottles from Machine 1 are produced separately from bottles from Machine 2, and the samples were taken independently, so the two samples are independent of each other.
- **Normal/Large Sample:** Both sample sizes are large ($n_1 = 40 \ge 30$ and $n_2 = 35 \ge 30$), so by the Central Limit Theorem the sampling distribution of $\overline{x}_1 - \overline{x}_2$ is approximately Normal.

Since the conditions are met, a two-sample z interval is appropriate.

Step 3 — Carry Out

$$(\overline{x}_1 - \overline{x}_2) \pm z^\star \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\overline{x}_1 - \overline{x}_2 = 502.3 - 500.9 = 1.4 \qquad z^\star = 1.96$$

$$SE = \sqrt{\frac{1.2^2}{40} + \frac{1.6^2}{35}} = \sqrt{\frac{1.44}{40} + \frac{2.56}{35}} = \sqrt{0.0360 + 0.0731} = \sqrt{0.1091} \approx 0.330$$

$$ME = 1.96(0.330) \approx 0.647$$

$$\text{CI: } 1.4 \pm 0.647 = (0.753,\ 2.047)$$

Step 4 — Interpret

We are 95% confident that the true mean fill amount for Machine 1 is between 0.753 mL and 2.047 mL higher than the true mean fill amount for Machine 2.

Example 2

A school compares weekly study time for students in two different programs. Two independent random samples are taken.

Group	n	$\overline{x}$ (hours)	s (hours)
Program A	18	6.8	1.9
Program B	14	5.4	2.3

Construct and interpret a 90% confidence interval for $\mu_A - \mu_B$.

Solution. Step 1 — State

Let μ_A be the true mean weekly study time (in hours) for all students in Program A, and let μ_B be the true mean weekly study time (in hours) for all students in Program B. We will construct a 90% confidence interval for the difference in population means, $\mu_A - \mu_B$ (Program A minus Program B).

Step 2 — Justify

We will use a two-sample t interval for $\mu_A - \mu_B$ using Welch's approximation because the population standard deviations are not known and we are using s_A and s_B as estimates.

- **Random:** The problem states that two independent random samples of students were taken from the two programs.
- **Independent:**
  - **Within groups:** Because the programs each contain many more than 18 and 14 students, it is reasonable to assume $n_A = 18 \leq 0.10N_A$ and $n_B = 14 \leq 0.10N_B$ (10% condition), so observations within each sample are approximately independent.
  - **Between groups:** The samples are from two different programs and were selected independently, so the two samples are independent of each other.
- **Normal/Large Sample:** The sample sizes are below 30 ($n_A = 18$ and $n_B = 14$), so we cannot rely on the CLT alone. We would need to check that each group's sample distribution is roughly symmetric with no outliers (using a histogram/boxplot or a graph provided). Assuming those graphs show no strong skewness and no outliers, the Normality condition is reasonable.

Since the conditions are met, a two-sample t interval with Welch's approximation is appropriate.

Step 3 — Carry Out

$$(\overline{x}_A - \overline{x}_B) \pm t^{\star}\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

$$\overline{x}_A - \overline{x}_B = 6.8 - 5.4 = 1.4$$

$$SE = \sqrt{\frac{1.9^2}{18} + \frac{2.3^2}{14}} = \sqrt{\frac{3.61}{18} + \frac{5.29}{14}} = \sqrt{0.2006 + 0.3779} = \sqrt{0.5785} \approx 0.7606$$

Welch df:

$$df \approx \frac{(0.2006 + 0.3779)^2}{\frac{0.2006^2}{17} + \frac{0.3779^2}{13}} = \frac{0.5785^2}{\frac{0.0402}{17} + \frac{0.1428}{13}} = \frac{0.3347}{0.00237 + 0.01099} \approx \frac{0.3347}{0.01336} \approx 25.1$$

So use $df \approx 25$.

For a 90% confidence interval, $\alpha = 0.10$, so we use $t^{\star} = t_{0.95,\ 25} \approx 1.708$.

$$ME = 1.708(0.7606) \approx 1.299$$

$$\text{CI: } 1.4 \pm 1.299 = (0.101,\ 2.699)$$

Step 4 — Interpret

We are 90% confident that the true mean weekly study time for Program A students is between 0.101 and 2.699 hours higher than the true mean weekly study time for Program B students.

Example 3

A nutritionist compares sodium content (mg) for two brands of soup. Independent samples are taken.

Brand	n	$\overline{x}$ (mg)	s (mg)
Brand 1	10	710	48
Brand 2	9	742	55

Construct and interpret a 95% confidence interval for $\mu_1 - \mu_2$ (Brand 1 minus Brand 2).

Solution. Step 1 — State

Let μ_1 be the true mean sodium content (in mg) for all cans of Brand 1 soup, and let μ_2 be the true mean sodium content (in mg) for all cans of Brand 2 soup.

We will construct a 95% confidence interval for the difference in population means, $\mu_1 - \mu_2$ (Brand 1 minus Brand 2).

Step 2 — Justify

We will use a two-sample t interval for $\mu_1 - \mu_2$ using Welch's approximation because the population standard deviations are not known and we are using s_1 and s_2 as estimates.

- **Random:** The problem does not explicitly state random sampling or random assignment. To proceed with inference, we must assume each sample is a random sample from its population (or that subjects were randomly assigned in an experiment). Without randomization, conclusions may not generalize.
- **Independent:**
  - **Within groups:** It is reasonable to assume each brand's production is much larger than the sample sizes, so $n_1 = 10 \leq 0.10N_1$ and $n_2 = 9 \leq 0.10N_2$ (10% condition). Thus observations within each sample are approximately independent.
  - **Between groups:** The samples come from two different brands and were taken independently, so the two samples are independent of each other.
- **Normal/Large Sample:** Both sample sizes are small ($n_1 = 10$ and $n_2 = 9$), so we cannot rely on the CLT. We would need to examine the sample distributions for each brand (histograms/boxplots or any graph provided) to verify they are roughly symmetric with no outliers and not strongly skewed. Assuming the graphs show this, the Normality condition is reasonable.

Since the conditions are met, a Welch two-sample t interval is appropriate.

Step 3 — Carry Out

$$\overline{x}_1 - \overline{x}_2 = 710 - 742 = -32$$

$$SE = \sqrt{\frac{48^2}{10} + \frac{55^2}{9}} = \sqrt{\frac{2304}{10} + \frac{3025}{9}} = \sqrt{230.4 + 336.1} = \sqrt{566.5} \approx 23.80$$

Welch df:

$$df \approx \frac{(230.4 + 336.1)^2}{\frac{230.4^2}{9} + \frac{336.1^2}{8}} = \frac{566.5^2}{\frac{53084}{9} + \frac{112963}{8}} \approx \frac{321{,}000}{5898 + 14120} \approx \frac{321{,}000}{20018} \approx 16.0$$

Use $df \approx 16$. For 95% CI, $t^\star = t_{0.975,16} \approx 2.120$.

$$ME = 2.120(23.80) \approx 50.5$$

$$\text{CI: } -32 \pm 50.5 = (-82.5,\ 18.5)$$

Step 4 — Interpret

We are 95% confident that the true mean sodium content for Brand 1 soup is between 82.5 mg lower and 18.5 mg higher than the true mean sodium content for Brand 2 soup.

Because 0 is contained in the interval, the data do not provide strong evidence of a difference in the true mean sodium contents at the 95% confidence level.

Example 4 (Enrichment: pooled t interval)

(Not needed for AP.) A researcher believes two populations have equal variances. Independent samples produce:

Group	n	$\overline{x}$	s
Group 1	22	15.2	3.1
Group 2	20	12.9	2.9

Construct a 95% pooled t interval for $\mu_1 - \mu_2$.

Solution. Step 1 — State

Let μ_1 be the true mean response for the population represented by Group 1 and let μ_2 be the true mean response for the population represented by Group 2.

We will construct a 95% confidence interval for the difference in population means, $\mu_1 - \mu_2$ (Group 1 minus Group 2), using the pooled two-sample t method.

Step 2 — Justify

We will use a pooled two-sample t interval because the researcher believes the population variances are equal, so it is reasonable (for enrichment purposes) to treat $\sigma_1^2 = \sigma_2^2$ and pool the sample variances.

- **Random:** The problem does not explicitly state random sampling or random assignment. To proceed with inference, we must assume each sample is a random sample from its population (or that subjects were randomly assigned in an experiment). Without randomization, conclusions may not generalize.
- **Independent:**
  - **Within groups:** It is reasonable to assume each population is much larger than the sample sizes, so $n_1 = 22 \leq 0.10N_1$ and $n_2 = 20 \leq 0.10N_2$ (10% condition). Thus observations within each sample are approximately independent.
  - **Between groups:** The samples were taken independently from two different groups, so the two samples are independent of each other.
- **Normal/Large Sample:** The sample sizes are below 30 ($n_1 = 22$ and $n_2 = 20$), so we would check that each group's sample distribution is roughly symmetric with no outliers (using graphs if provided). Assuming that condition holds, using a t procedure is reasonable.
- **Equal variances (pooling condition):** For a pooled procedure, we additionally need it to be reasonable that $\sigma_1^2 = \sigma_2^2$. The sample standard deviations are close ($s_1 = 3.1$ and $s_2 = 2.9$), which supports the assumption of similar variances (in practice one might also use Levene's test, beyond AP).

Since the conditions are met (including the equal-variances assumption for enrichment), a pooled two-sample t interval is appropriate.

Step 3 — Carry Out

$$\overline{x}_1 - \overline{x}_2 = 15.2 - 12.9 = 2.3$$

$$s_p^2 = \frac{(22-1)3.1^2 + (20-1)2.9^2}{22+20-2} = \frac{21(9.61) + 19(8.41)}{40} = \frac{201.81 + 159.79}{40} = \frac{361.60}{40} = 9.04$$

$$s_p = \sqrt{9.04} \approx 3.007$$

$$SE_p = s_p\sqrt{\frac{1}{22}+\frac{1}{20}} = 3.007\sqrt{0.04545+0.05000} = 3.007\sqrt{0.09545} \approx 3.007(0.3090) \approx 0.929$$

Degrees of freedom:

$$df = n_1 + n_2 - 2 = 40$$

For 95% CI, $t^\star = t_{0.975,40} \approx 2.021$.

$$ME = 2.021(0.929) \approx 1.878$$

$$\text{CI: } 2.3 \pm 1.878 = (0.422,\ 4.178)$$

Step 4 — Interpret

We are 95% confident that the true mean for Group 1 is between 0.422 and 4.178 units higher than the true mean for Group 2.