

Inference for Means with t Procedures

Grade 12 – AP Statistics math · Worksheet Answer Key · Standards: AP Statistics Unit 4 (2026-27)

The answer key checks conditions, shows the standard error, the t statistic, the degrees of freedom and the p-value, and ends with a conclusion in context. P-values come from the t distribution with the stated degrees of freedom.

Standards: AP Statistics Unit 4 (2026-27) · **Level:** Grade 12 AP Statistics

Problem 1 Standard Error of a Mean

$$SE = s/\sqrt{n} = 30/\sqrt{25} = 30/5 = 6.0.$$

6.0 minutes

Problem 2 Finding a Critical Value

(a) $df = 11$, $t^* = 2.201$. (b) $df = 19$, $t^* = 2.861$. (c) $df = 7$, $t^* = 1.895$. (calculator: $\text{invT}(0.975, 11)$, $\text{invT}(0.995, 19)$, $\text{invT}(0.95, 7)$).

(a) 2.201; (b) 2.861; (c) 1.895

Problem 3 A One-Sample t Interval

$df = 15$, $t^* = 2.131$. $SE = 0.85/\sqrt{16} = 0.2125$. $ME = 2.131 \times 0.2125 = 0.4529$. Interval: 4.62 ± 0.453 .

(4.17, 5.07) pounds

Problem 4 Checking Conditions

The sample is small ($n = 10$), so the t procedure needs the data to show no strong skew or outliers. The dotplot has one value far above the rest (15 minutes), which is an outlier, so the normal/large sample condition is not met.

Not appropriate: strong outlier with $n = 10$

Problem 5 Interpreting an Interval

We are 95% confident that the mean weekly screen time of all students at the school is between 18.47 and 24.33 hours. The confidence level refers to the method: about 95% of intervals constructed this way from random samples capture the true mean.

95% confident the true mean is in (18.47, 24.33)

Problem 6 Writing Hypotheses for Means

(a) $H_0: \mu = 450$, $H_a: \mu < 450$ (μ is the true mean loaf weight). (b) $H_0: \mu = 8$, $H_a: \mu \neq 8$ (μ is the true mean sleep time on school nights). (c) $H_0: \mu = 12$, $H_a: \mu > 12$ (μ is the true mean battery life in hours).

(a) $\mu < 450$; (b) $\mu \neq 8$; (c) $\mu > 12$

Problem 7 Test Statistic and p-Value

$SE = 4.1/\sqrt{22} = 0.8741$. $t = (448.2 - 450)/0.8741 = -2.059$, $df = 21$. $p\text{-value} = \text{tcdf}(-\infty, -2.059, 21) = 0.0260$.

$t \approx -2.06$, $p\text{-value} \approx 0.0260$

Problem 8 Two-Sided t Test

$SE = 1.3/\sqrt{30} = 0.2373$. $t = (7.4 - 8)/0.2373 = -2.528$, $df = 29$. Two-sided $p\text{-value} = 0.0172$. Since $0.0172 < 0.05$, reject H_0 : there is convincing evidence that the mean sleep time differs from 8 hours.

$t \approx -2.53$, $p\text{-value} \approx 0.0172$; reject H_0

Problem 9 A Claim from an Interval

$df = 23$, $t^* = 2.069$, $ME = 0.633$, interval (5.47, 6.73). Since 7 is outside the interval, a two-sided test at the 5% level rejects $\mu = 7$: there is convincing evidence that the mean is not 7 hours.

Yes; 7 is outside (5.47, 6.73)

Problem 10 Sample Size and Interval Width

$n = 9$: $t^* = 2.306$, $ME = 2.306 \times 12/3 = 9.22$. $n = 36$: $t^* = 2.030$, $ME = 2.030 \times 12/6 = 4.06$. The small sample is wider because the standard error is larger and the critical value is larger.

$ME \approx 9.22$ vs 4.06

Problem 11 Paired or Two-Sample

(a) The same runners are measured twice: paired. (b) Two unrelated groups: two-sample. (c) Twins form natural pairs and treatments are randomized within each pair: paired.

(a) paired; (b) two-sample; (c) paired

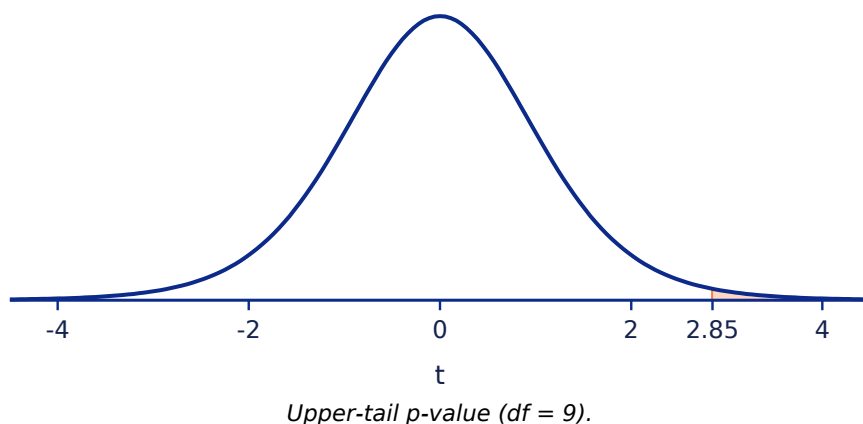
Problem 12 Paired t Interval

The dotplot is roughly symmetric with no outliers, so the conditions are reasonable. $\bar{d} = 3.750$, $s_d = 2.493$. $df = 7$, $t^* = 2.365$, $ME = 2.365 \times 2.493/\sqrt{8} = 2.084$.

(1.67, 5.83) points

Problem 13 Paired t Test

$H_0: \mu_d = 0$; $H_a: \mu_d > 0$. $\bar{d} = 10.0$, $s_d = 11.106$. $t = 10.0/(11.106/\sqrt{10}) = 2.847$, $df = 9$, $p\text{-value} = 0.0096$. Since $0.0096 < 0.05$, reject H_0 : there is convincing evidence that texting increases the mean braking reaction time.



$t \approx 2.85$, $p\text{-value} \approx 0.0096$; reject H_0

Problem 14 Two-Sample t Interval

$SE = \sqrt{(5.2^2/28 + 6.1^2/31)} = \sqrt{(0.9657 + 1.2003)} = 1.4717$. Technology gives $df \approx 56.8$, $t^* = 2.003$. $ME = 2.947$. Difference = 3.5.

(0.55, 6.45) grams

Problem 15 Two-Sample t Test

The boxplots are roughly symmetric with no outliers, and both samples are around 30. $H_0: \mu_1 - \mu_2 = 0$; $H_a: \mu_1 - \mu_2 \neq 0$. $SE = 2.7481$, $t = 5.3/2.7481 = 1.929$, $df \approx 53.4$, $p\text{-value} = 0.0591$. Since $0.0591 > 0.05$, fail to reject H_0 : not convincing evidence that the means differ.

$t \approx 1.93$, $p\text{-value} \approx 0.059$; fail to reject H_0

Problem 16 Interval and Test Agree

$t^* = 2.005$, $ME = 5.511$. Interval: 5.3 ± 5.51 gives $(-0.21, 10.81)$. The interval contains 0, which agrees with failing to reject H_0 at $\alpha = 0.05$.

$(-0.21, 10.81)$; contains 0

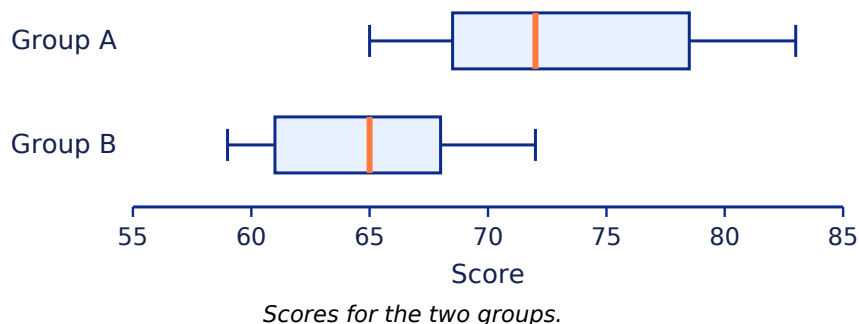
Problem 17 Selecting a Procedure

(a) Two-sample z interval or test for a difference of proportions. (b) One-sample t interval for a mean. (c) Paired t procedure (one-sample t on the differences). (d) Chi-square test for homogeneity.

(a) two-proportion z; (b) one-sample t; (c) paired t; (d) chi-square homogeneity

Problem 18 A Full Two-Sample Test

State: $H_0: \mu_A = \mu_B$; $H_a: \mu_A \neq \mu_B$. **Plan:** two-sample t test; random independent samples; the boxplots below show no outliers and no strong skew. **Do:** $\bar{x}_A = 73.11$, $s_A = 5.925$; $\bar{x}_B = 65.00$, $s_B = 4.346$; $SE = 2.406$; $t = 3.371$; $df \approx 14.58$; $p\text{-value} = 0.0043$. **Conclude:** $0.0043 < 0.05$, so there is convincing evidence that the mean scores differ.



$t \approx 3.37$, $df \approx 14.6$, $p\text{-value} \approx 0.0043$; reject H_0

Problem 19 Confidence Level and Two-Sided Tests

(a) 50 is outside the interval, so reject H_0 at $\alpha = 0.01$. (b) 45 is inside, so fail to reject: 45 is a

plausible value. (c) A lower confidence level uses a smaller t^* , so the interval is narrower.

(a) reject; (b) fail to reject; (c) narrower

Problem 20 Why Pairing Matters

(a) Differences (before minus after): 4, 6, 1, 8, 8, 3, 9, 5; $\bar{d} = 5.500$, $s_d = 2.777$; $t = 5.601$, $df = 7$, $p\text{-value} = 0.00082$. (b) Treating the groups as independent: $t = 1.564$, $p\text{-value} = 0.1429$. (c) The paired analysis removes the person-to-person variation in pulse and gives a much stronger result. The independent analysis is wrong here because the two measurements are linked.

Paired: $t \approx 5.60$, $p \approx 0.0008$; independent: $t \approx 1.56$, $p \approx 0.14$