

AP Statistics - Inference for the Difference of Two Means - Answer Key

Common Core AP STAT 7.6, AP STAT 7.7, AP STAT 7.8, AP STAT 7.9, HSS-IC.B.5 · maths-pdf.fr/en/

Each solution gives the conditions, the standard error, the degrees of freedom from a calculator, the interval or t statistic, and a conclusion in context. Probabilities and critical values use the exact (non-integer) degrees of freedom.

AP Statistics · Common Core State Standards: **AP STAT 7.6, AP STAT 7.7, AP STAT 7.8, AP STAT 7.9, HSS-IC.B.5**

Exercise 1 - Parameters and hypotheses

μ_A = true mean score of all students taught with Method A; μ_B likewise for Method B.

$$H_0: \mu_A - \mu_B = 0, H_a: \mu_A - \mu_B > 0.$$

Exercise 2 - Difference and standard error

Difference: $78.4 - 72.1 = 6.3$.

$$SE = \sqrt{\frac{9.2^2}{35} + \frac{11.5^2}{40}} = \sqrt{2.4183 + 3.3062} = 2.393.$$

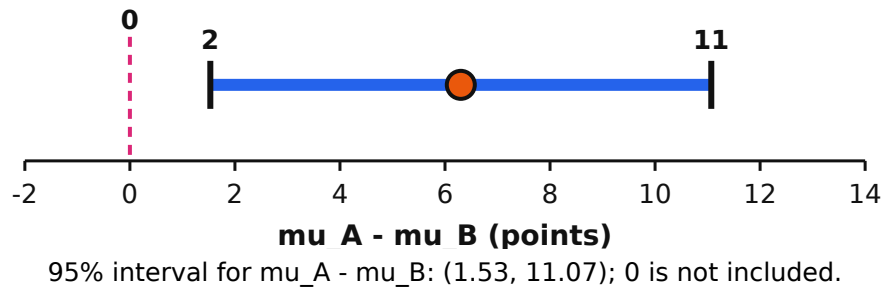
6.3 points; SE = 2.393.

Exercise 3 - Degrees of freedom

Conservative: $\min(35-1, 40-1) = 34$. Calculator: $df \approx 72.5$.

34 compared with 72.5; either gives a similar critical value.

Exercise 4 - Confidence interval



$t^* = 1.993$, $ME = 4.769$. Interval: 6.3 ± 4.769 .

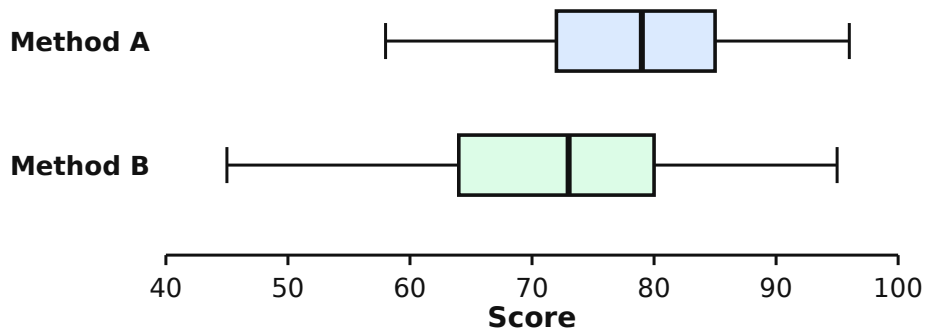
(1.531, 11.069) points.

Exercise 5 - Interpret the interval

State the parameter, the groups, and the direction.

We are 95% confident that the true mean score with Method A is between 1.531 and 11.069 points higher than with Method B. Because 0 is not in the interval, there is convincing evidence of a difference.

Exercise 6 - Check conditions



Random and independent: given. 10%: each sample is far less than 10% of its population.
Normal/Large: both sample sizes are at least 30, and the boxplots show no outliers.

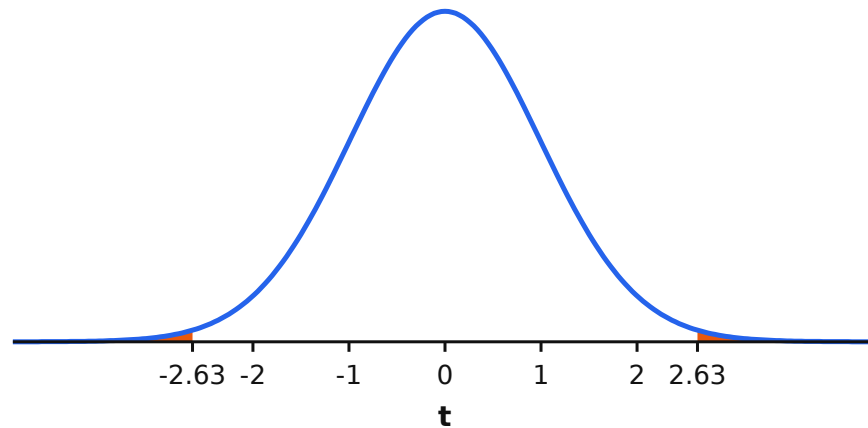
All conditions are met.

Exercise 7 - Compute t

$$t = \frac{6.3}{2.393} = 2.63.$$

t = 2.63.

Exercise 8 - p-value



Two-sided p-value for $t = 2.63$: 0.0104.

$$p = 2P(t_{72.5} > 2.63) = 0.0104.$$

0.0104.

Exercise 9 - Conclusion

$$0.0104 < 0.05.$$

Reject H_0 . There is convincing evidence of a difference in the true mean scores of students taught with Method A and Method B (Method A higher).

Exercise 10 - Experiment with caffeine

$H_0: \mu_c - \mu_p = 0$, $H_a: \mu_c - \mu_p > 0$. Random assignment; small samples, assume approximately normal.

$$SE = 1.963, df \approx 23.8, t = 1.68, p\text{-value} = 0.0530.$$

Fail to reject H_0 : not convincing evidence that caffeine increases the mean score (the p-value 0.0530 exceeds 0.05).

Exercise 11 - From raw data

Fertilizer 1: mean 59.000, s 4.472. Fertilizer 2: mean 52.000, s 5.071.

$$SE = 2.390, df \approx 13.8, t^* = 2.148, ME = 5.135.$$

(1.865, 12.135) cm.

Exercise 12 - Paired or independent?

(a) The same runners are measured twice: paired. (b) Different people in two groups: two-sample. (c) Twin pairs are matched: paired.

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

Exercise 13 - Wrong procedure

The scores are dependent because they come from the same students.

Use a matched-pairs t procedure on the differences (after minus before) instead.

Exercise 14 - Interval and test

0 is inside the interval.

Fail to reject H_0 : there is not convincing evidence that the means differ.

Exercise 15 - Homework hours

$$SE = 0.260, df \approx 83.4, t = 1.92, \text{p-value} = 0.0583.$$

Fail to reject H_0 : p-value 0.0583 exceeds 0.05, so there is not convincing evidence that the mean weekly homework time differs between boys and girls.

Exercise 16 - Effect of sample size

$$n = 25: \sqrt{\frac{100}{25} + \frac{100}{25}} = \sqrt{8} = 2.828. \quad n = 100: \sqrt{\frac{100}{100} + \frac{100}{100}} = \sqrt{2} = 1.414.$$

The larger samples halve the standard error, giving a narrower interval and a larger $|t|$ for the same difference.

Exercise 17 - Outlier problem

The outlier inflates the mean and standard deviation of group 2, and the sample is small.

The Normal condition fails (small n with an outlier), so the t interval may be misleading; investigate the outlier first.

Exercise 18 - 99 percent interval

$SE = 1.072$, $df \approx 113.0$, $t^* = 2.620$, $ME = 2.810$.

(0.690, 6.310) minutes.

Exercise 19 - Multi-step

$t = 2.33$, $df \approx 55.8$, $p\text{-value} = 0.0117$.

90% interval: (0.652, 3.948).

Reject H_0 ($p < 0.05$); the 90% interval is entirely positive, which agrees.

Exercise 20 - Describe the difference

0 is in the interval.

We are 95% confident that the true mean reaction time for men is between 8 ms less and 21 ms more than for women. Since 0 is plausible, there is not convincing evidence of a difference.