

AP Statistics · Regression Analysis · Worksheet Answer Key

Common Core AP Statistics Unit 5 (2026-27) · mathovore.fr

This ap statistics two worksheet answer key solves all 22 problems of the regression analysis worksheet step by step.

Every solution shows the formula, the substitution and the final answer in bold. Interpretations are written in context, the way the AP exam scoring guides expect, and several solutions include a graph of the data, a fitted line or the residuals.

AP StatisticsCommon Core: AP Statistics Unit 5 (2026-27)

AP Statistics Two Worksheet Answer Key: problems 1-5

Problem 1: Explanatory and response variables

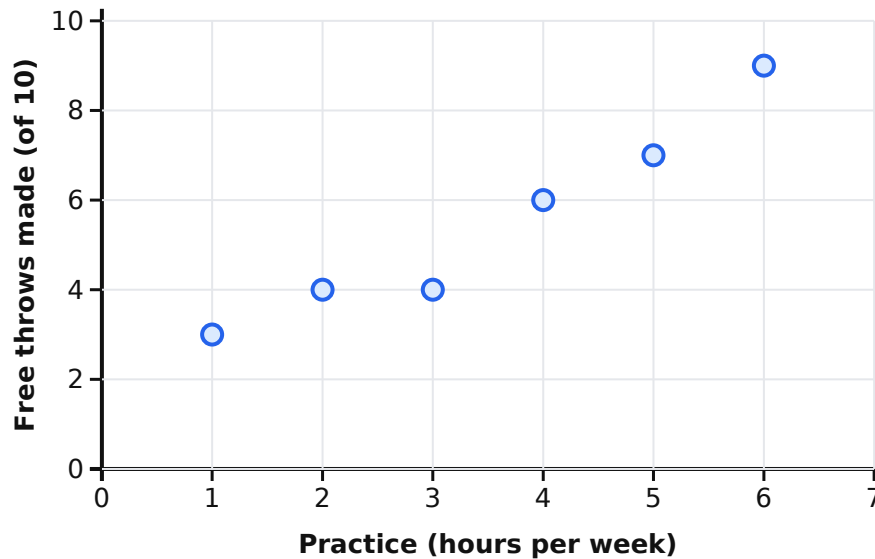
- a. Explanatory: **hours studied**; response: **exam score**.
- b. Explanatory: **age of the car**; response: **selling price**.

Problem 2: Describe a scatterplot

The association is negative, linear and strong: older cars have lower prices, the points follow a straight line closely, and there are no unusual points. Price falls by about 2 thousand dollars for each extra year.

Problem 3: Plot and describe data

Put practice hours on the horizontal axis and free throws made on the vertical axis, then plot the six points.



Free throws made versus weekly practice for six players.

Positive, linear and strong (r is about 0.97): more practice goes with more free throws made. No outliers appear.

Problem 4: Match correlations to graphs

A: $r = -0.92$ (downward trend, fairly tight). **B: $r = 0.996$** (upward and almost exactly on a line). **C: $r = 0.10$** (no clear pattern).

Problem 5: Properties of correlation

- The association is **positive and strong**.
- It does not change.** r has no units and is not affected by a change of units.
- No.** The correlation is always between -1 and 1.

AP Statistics Two Worksheet Answer Key: problems 6-10

Problem 6: Compute the correlation

Compute $z_x = \frac{x-4}{2.74}$ and $z_y = \frac{y-6}{2.83}$ for each point, and multiply.

x	y	z of x	z of y	product
1	3	-1.10	-1.06	1.16
2	4	-0.73	-0.71	0.52
4	5	0.00	-0.35	-0.00
5	9	0.37	1.06	0.39
8	9	1.46	1.06	1.55

The products sum to 3.61, so $r = \frac{3.61}{5-1} \approx 0.90$. **r is about 0.90.**

Problem 7: Slope and intercept from statistics

$$b = r \frac{s_y}{s_x} = 0.6 \cdot \frac{20}{8} = 1.5.$$

$$a = \bar{y} - b\bar{x} = 120 - 1.5(40) = 60.$$

$$\hat{y} = 60 + 1.5x$$

Problem 8: Make a prediction

$$\hat{y} = 12.4 + 3.1(20) = 74.4 \text{ and } \hat{y} = 12.4 + 3.1(45) = 151.9.$$

About 74.4 calories and about 151.9 calories. Both x values are inside the range of the data.

Problem 9: Interpret the slope

For each additional year of age, the predicted selling price decreases by about 1.9 thousand dollars (\$1,900).

Problem 10: Interpret the intercept

The intercept 30.5 is the predicted height, in inches, at age 0. **It is not meaningful:** age 0 is outside the data (3 to 10 years), and the line is not a reliable model there.

AP Statistics Two Worksheet Answer Key: problems 11-15

Problem 11: Interpret r squared

- About 81% of the variation in exam scores is explained by the linear relationship with hours studied.**
- $r = +\sqrt{0.81} = 0.9$: the slope is positive, so **r = 0.9**.
- $(-0.8)^2 = 0.64$, so $r^2 = 0.64$.

Problem 12: Compute a residual

$$\hat{y} = 48.5 + 3.2(7) = 70.9.$$

$$\text{residual} = y - \hat{y} = 78 - 70.9 = 7.1. \text{ The residual is 7.1 points.}$$

Problem 13: Meaning of a residual

$$\hat{y} = 24.8 - 1.9(6) = 13.4, \text{ or } \$13,400.$$

$$y - \hat{y} = 9.2 - 13.4 = -4.2.$$

The residual is -4.2 thousand dollars: the car sold for \$4,200 less than the model predicted (the model overestimated the price).

Problem 14: Read a residual plot

Yes. The residuals are scattered randomly around zero, with no curve or change in vertical spread, so the linear form fits the data.

Problem 15: Computer regression output

- a. $\hat{y} = 12.35 + 4.62x$, where x is hours studied and y is the exam score.
- b. $r = \sqrt{0.724} \approx 0.85$, positive because the slope is positive: **r is about 0.85.**
- c. **The typical distance between the actual exam scores and the scores predicted by the line is about 5.10 points.**

AP Statistics Two Worksheet Answer Key: problems 16-20

Problem 16: Regression line from data

$$\bar{x} = 5.5, s_x \approx 2.74, \bar{y} \approx 9.333, s_y \approx 4.46, r \approx 0.967.$$

$$b = 0.967 \cdot \frac{4.46}{2.74} \approx 1.57 \text{ and } a = 9.333 - 1.57(5.5) \approx 0.68.$$

$$\hat{y} = 0.68 + 1.57x$$

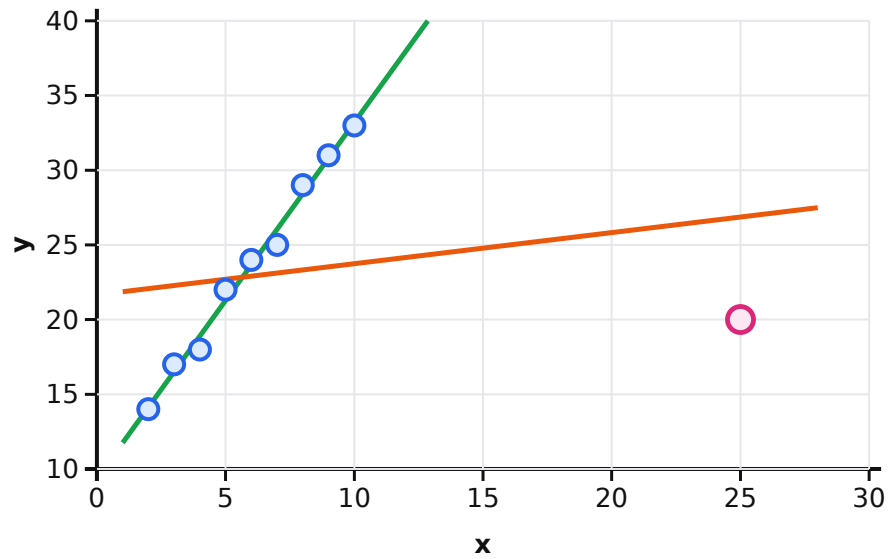
Problem 17: Extrapolation

$$\hat{y} = 2.1 + 0.9(60) = 56.1 \text{ inches.}$$

The prediction is about 56.1 inches, but it is an extrapolation: day 60 is far outside the observed days 1 to 14, and plants stop growing at a steady rate, so there is no reason to expect the linear pattern to continue.

Problem 18: Influential point

- a. **The point (25, 20).** Its x value is far from the others, so it pulls the line toward it, even though the other points follow a steep line.
- b. With all ten points $b \approx 0.21$ and $r \approx 0.22$. Without it $b \approx 2.38$ and $r \approx 0.995$. **Removing the point makes the slope much steeper and r much closer to 1.**

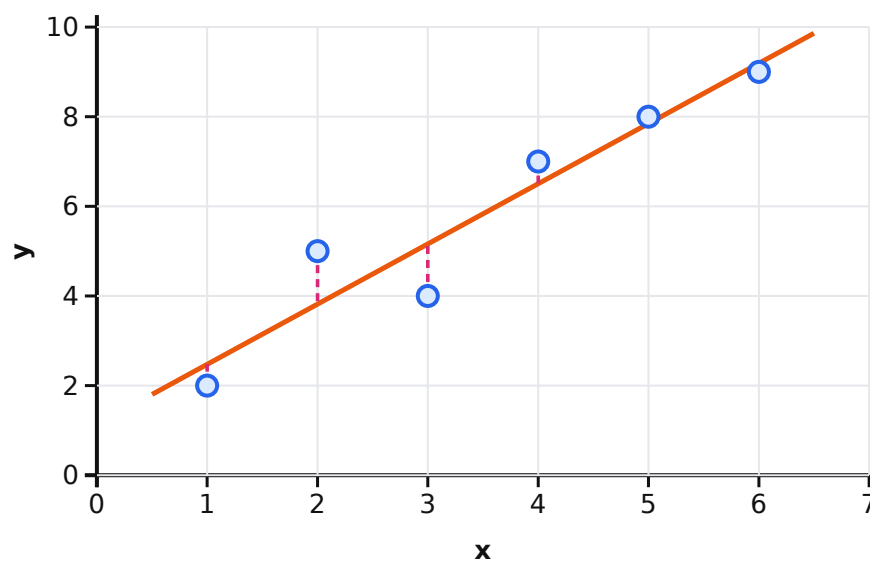


Orange: all ten points (slope 0.21). Green: without (25, 20) (slope 2.38).

Problem 19: Residuals for a small data set

- $x=1: 2-2.47=-0.47$
- $x=2: 5-3.81=1.19$
- $x=3: 4-5.15=-1.15$
- $x=4: 7-6.49=0.51$
- $x=5: 8-7.83=0.17$
- $x=6: 9-9.17=-0.17$

Sum: $-0.47+1.19-1.15+0.51+0.17-0.17=0.08$, close to 0 (rounding). **The largest residual is about 1.19 at the point (2, 5).** With the unrounded line the residual is 1.18.



Dashed segments are residuals; the largest, 1.18, is at $x = 2$.

Problem 20: Complete regression analysis

- a. $b = 0.994 \cdot \frac{11.263}{1.225} \approx 9.14$ and $a = 67 - 9.14(2.25) \approx 46.43$: $\hat{y} = 46.43 + 9.14x$.
- b. **Each additional hour of study is associated with about 9.14 more points on the quiz.**
- c. $r^2 \approx 0.988$: **about 98.8% of the variation in quiz scores is explained by the linear relationship with hours studied.**
- d. $\hat{y} = 46.43 + 9.14(2.5) \approx 69.28$, so the residual is $70 - 69.28 \approx 0.7$. **The residual is about 0.7 points** (0.71 with unrounded values).

AP Statistics Two Worksheet Answer Key: problems 21-22

Problem 21: Curved pattern

No. The scatterplot rises and then falls, and the residual plot shows a clear upside-down U pattern, which means the form is curved, not linear. A line would predict sales poorly. The low r of 0.20 comes from this curvature; it does not mean the variables are unrelated.

Problem 22: Laptop resale value

- a. $b = -0.90 \cdot r_{ac} 2.010 = -1.8$ and $a = 5.6 - (-1.8)(2.8) = 10.64$: $\hat{y} = 10.64 - 1.8x$.
- b. **For each additional year of age, the predicted resale value decreases by about 1.8 hundred dollars (\$180).**
- c. $\hat{y} = 10.64 - 1.8(4) = 3.44$, so about **\$344**.
- d. The value is 4.2 hundred dollars, so the residual is $4.2 - 3.44 = 0.76$. **The laptop is listed about \$76 above the predicted value.**
- e. $r^2 = (-0.90)^2 = 0.81$: **about 81% of the variation in resale value is explained by the linear relationship with age.**

How to use this ap statistics two worksheet answer key

Read this ap statistics two worksheet answer key after trying each problem yourself, and redo any problem you missed.

