

AP Statistics - Least-Squares Regression and Residuals - Worksheet

Common Core AP STAT 2.6, AP STAT 2.7, AP STAT 2.8, AP STAT 2.9, HSS-ID.B.6, HSS-ID.C.7 · maths-pdf.fr/en/

Name: _____ Class: _____ Date: ____/____/____

These 20 problems practice AP Statistics topics 2.6 to 2.9. You will predict with a regression equation, interpret slope, intercept, and r squared, compute residuals, read residual plots, and judge influential points and extrapolation.

AP Statistics · Common Core State Standards: **AP STAT 2.6, AP STAT 2.7, AP STAT 2.8, AP STAT 2.9, HSS-ID.B.6, HSS-ID.C.7**

Exercise 1 - Predict

A regression of a house's price (in \$1,000s) on its size (in square feet) is $\hat{y} = 40 + 0.12x$. Predict the price of a 1,800-square-foot house.

Exercise 2 - Interpret the slope

For the equation $\hat{y} = 40 + 0.12x$ (price in \$1,000s, size in square feet), interpret the slope.

Exercise 3 - Interpret the intercept

The equation $\hat{y} = 3.5 + 2.1x$ predicts minutes of homework (y) from the number of classes (x). Interpret the intercept and say whether it is meaningful.

Exercise 4 - Compute a residual

A model $\hat{y} = 12 + 3x$ predicts the number of points. Sofia studied $x = 5$ hours and scored $y = 30$. Find her residual and interpret it.

Exercise 5 - Find the slope and intercept

Heights (x) have mean 66 in and sd 3 in; weights (y) have mean 150 lb and sd 20 lb. The

correlation is $r = 0.60$. Find the least-squares line.

Exercise 6 - Point of means

For the line $\hat{y} = -114 + 4x$ with $\bar{x} = 66$, verify that the line passes through the point of means.

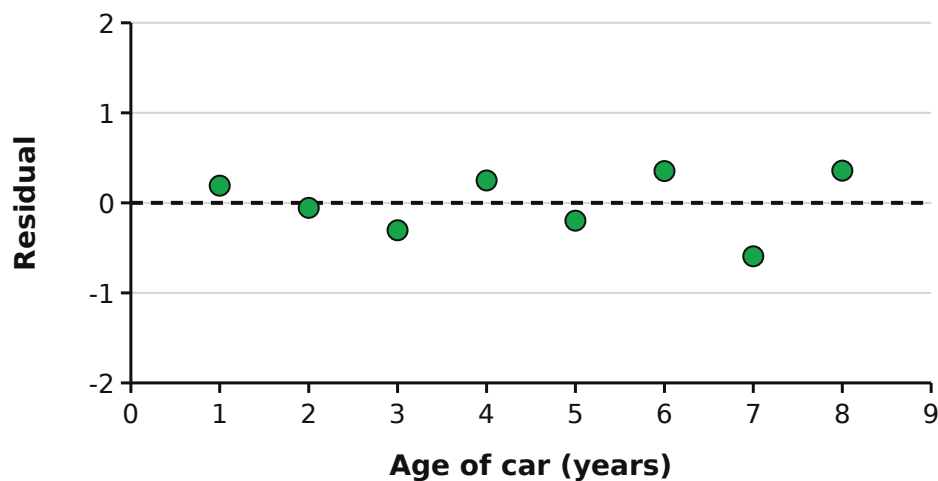
Exercise 7 - Interpret r squared

A regression of battery life on screen brightness has r squared = 0.64. Interpret it.

Exercise 8 - Get r from r squared

A regression of fuel used on distance has r squared = 0.81 and a positive slope. Find r .

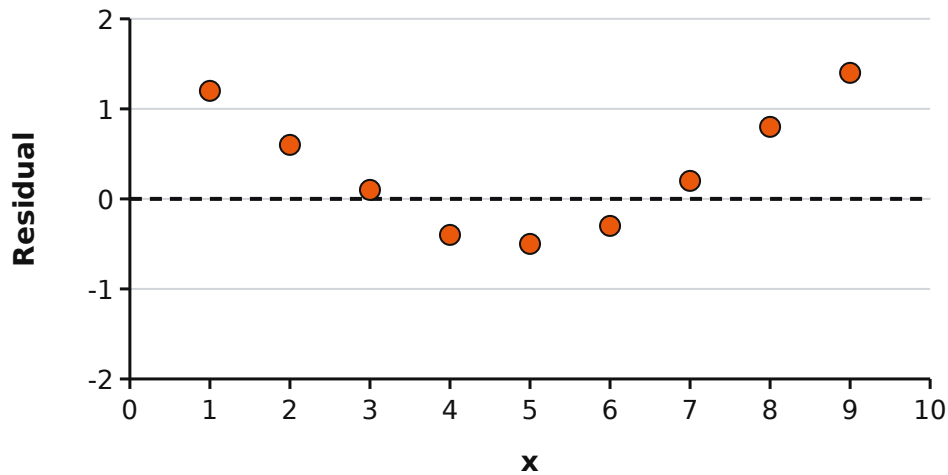
Exercise 9 - Residual plot: good fit



A residual plot with no pattern: a linear model is appropriate.

The residual plot for the car data is shown. Is a linear model appropriate?

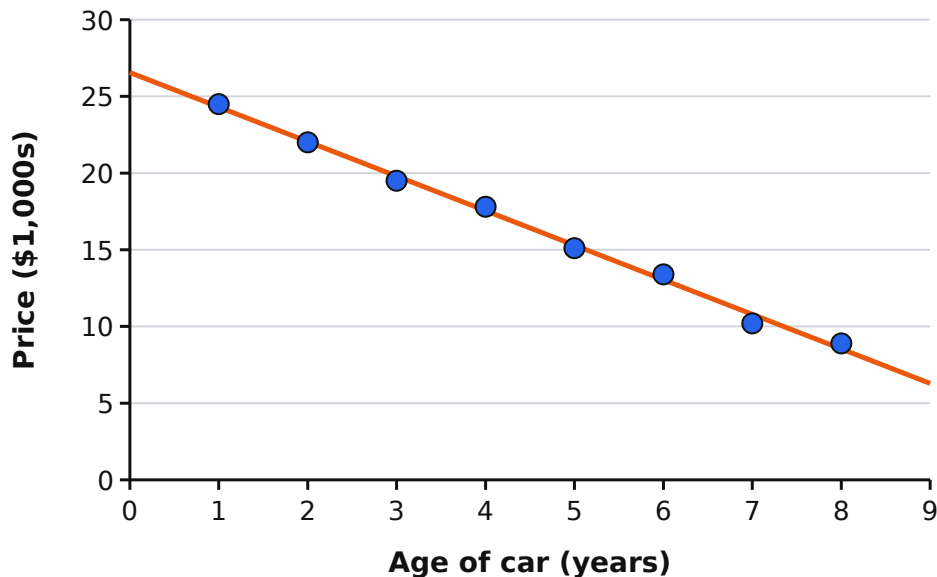
Exercise 10 - Residual plot: curved



A curved residual pattern signals that a straight line is not the right model.

This residual plot comes from a different data set. Is a linear model appropriate? What might help?

Exercise 11 - Extrapolation



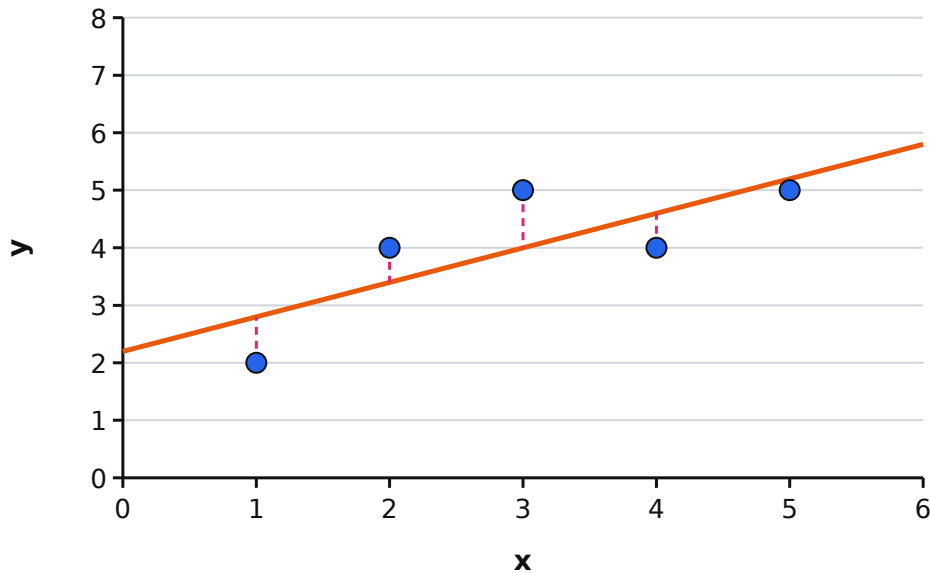
Least-squares line for car price (in \$1,000s) against age: $\text{price-hat} = 26.56 - 2.25(\text{age})$. Dashed segments are residuals.

Using the car data (ages 1 to 8 years), the line predicts price = $26.56 - 2.25(\text{age})$. Predict the price at age 15. Is this prediction trustworthy?

Exercise 12 - LSRL for five points

For the data (1, 2), (2, 4), (3, 5), (4, 4), (5, 5), the means are 3 and 4, $s_x = 1.581$, $s_y = 1.095$, $r = 0.775$. Find the regression line.

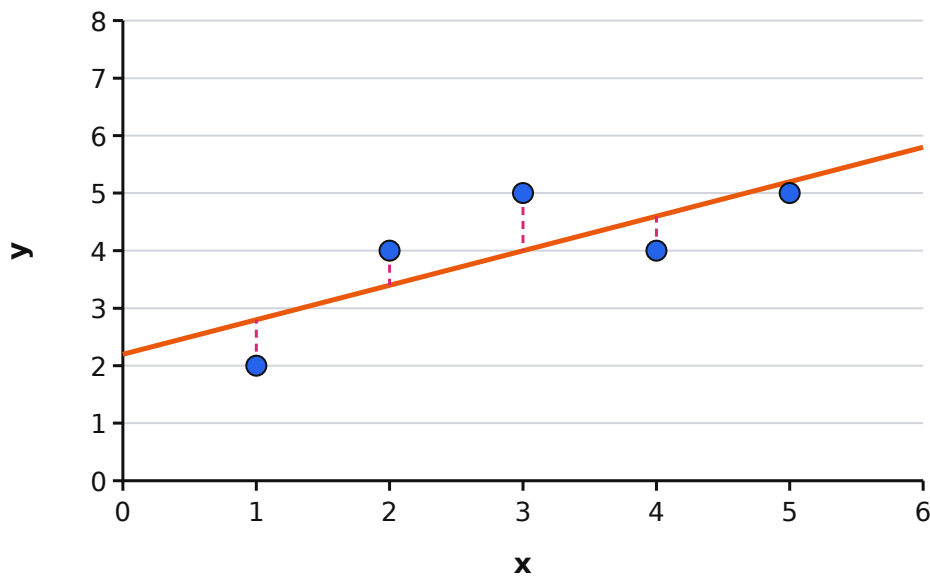
Exercise 13 - Residuals add to zero



Least-squares line through five points.

For $\hat{y} = 2.2 + 0.6x$ and the five points above, compute all five residuals and their sum.

Exercise 14 - Standard deviation of residuals



Least-squares line through five points.

Using the residuals -0.8, 0.6, 1.0, -0.6, -0.2, find s and interpret it.

Exercise 15 - Read computer output

Predictor	Coef
Constant	25.6
Hours	-3.2
R-Sq	78%

Output for sleep data.

A computer gives this output for predicting a student's test score from hours of screen time. Write the regression equation, interpret the slope, and give r .

Exercise 16 - Influential point

A regression of 12 points has slope 2.1. Removing one point far to the right of the other points changes the slope to 0.9. What kind of point is it?

Exercise 17 - Compare two predictions

The line $\hat{y} = 18 + 2.5x$ predicts commute cost (y) from distance (x). A 10-mile commute costs \$39 and a 12-mile commute costs \$52. Find both residuals. Which is the better prediction?

Exercise 18 - Why least squares?

Explain what 'least squares' means for the regression line.

Exercise 19 - Describe a fit

For a regression of mileage on speed, r squared = 0.92, s = 1.4 mpg, and the residual plot shows no pattern. Write a conclusion about the model.

Exercise 20 - Two predictions

The LSRL for a bakery is $\hat{y} = 15 + 0.8x$, where x is the number of cupcakes ordered and y is the

preparation time in minutes. Predict the time for 40 cupcakes. If the actual time was 52 minutes, find the residual.