Statistics Kingdom

<u>Home</u> > <u>Regression</u> > Linear Regression

Linear Regression Calculator

Linear regression calculator and prediction interval calculator with step-by-step solution.

Simple Linear regression Multiple Linear regression Logistic regression Multinomial logistic regression

How to do with R?

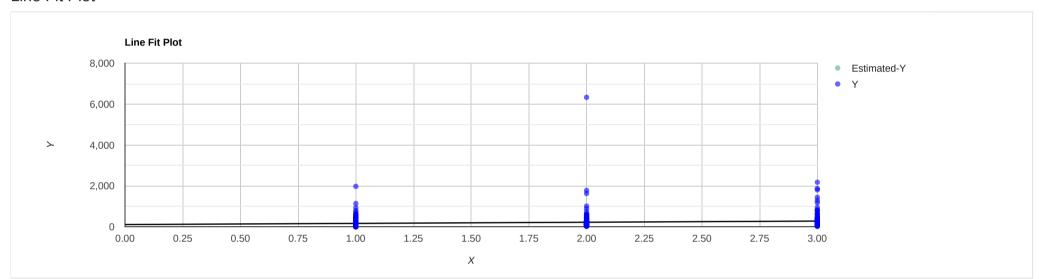
Regression line equation

$\hat{\mathbf{Y}} = \mathbf{105.3922} + \mathbf{56.7701X}$

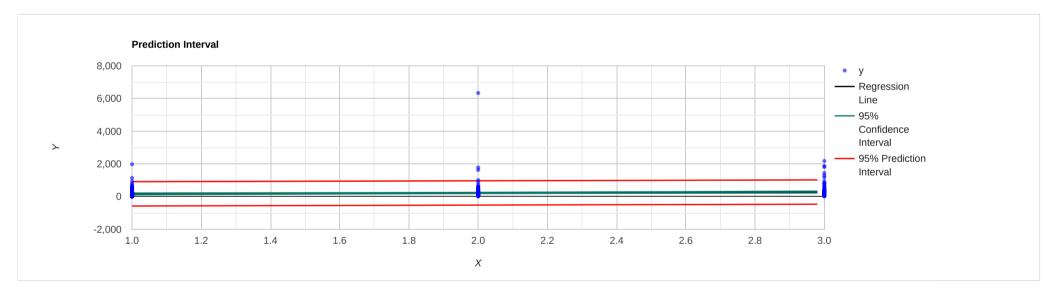
Reporting linear regression in APA style

X predicted Y, R^2 = .016, F(1,498) = 8.2, p = .004. β = 56.77, p = .004, α = 105.39, p = .010.

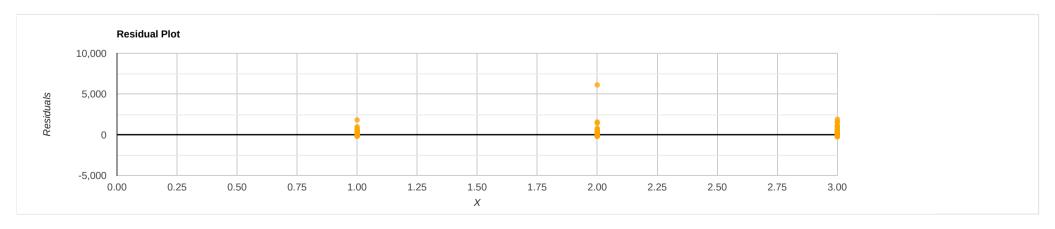
Line Fit Plot



Prediction online



Residual Plot



Prediction

Interpretation of the results



Hover over the cells to see the formulas.

Source	DF	Sum of Square	Mean Square	F Statistic (df ₁ ,df ₂)	P-value
Regression (between $\hat{y_i}$ and \bar{y})	1	1169248.7719	1169248.7719	8.2041 (1,498)	0.004356
	498	70974786.0281	142519.6507		
Total (between y_i and \bar{y})	499	72144034.8	144577.224		

1. Y and X relationship

R-Squared (R²) equals **0.01621.** This means that 1.6% of the variability of Y is explained by X.

Correlation (R) equals 0.1273. This means that there is a very weak direct relationship between X and Y.

The Standard deviation of the residuals (S_{res}) equals **377.5177**.

The slope: b_1 =**56.7701** Cl[17.829, 95.7113] means that when you increase X by 1, the value of Y increases by 56.7701.

The y-intercept: b_0 =105.3922 CI[25.0186, 185.7658] means that when X equals 0, the prediction of Y's value is 105.3922.

The x-intercept equals -1.8565.

2. Goodness of fit

Overall regression: right-tailed, F(1,498) = 8.2041, p-value = 0.004356. Since p-value < α (0.05), we reject H_0 .

The linear regression model, $Y = b_0 + b_1 X + \epsilon$, provides a better fit than the model without the independent variable resulting in $Y = b_0 + \epsilon$.

The slope (b_1) : two-tailed, T(498)=2.8643, p-value = 0.004356. For one predictor it is the same as the p-value for the overall model.

The y-intercept (b_0): two-tailed, T(498) = 2.5763, p-value = 0.01027. Hence, b_0 is significantly different from zero.

3. Residual normality

The linear regression model assumes normality for residual errors. The Shapiro-Wilk p-value equals 0. It is assumed that the data is not normally distributed, But since the sample size is large, it should not adversely affect the regression model.

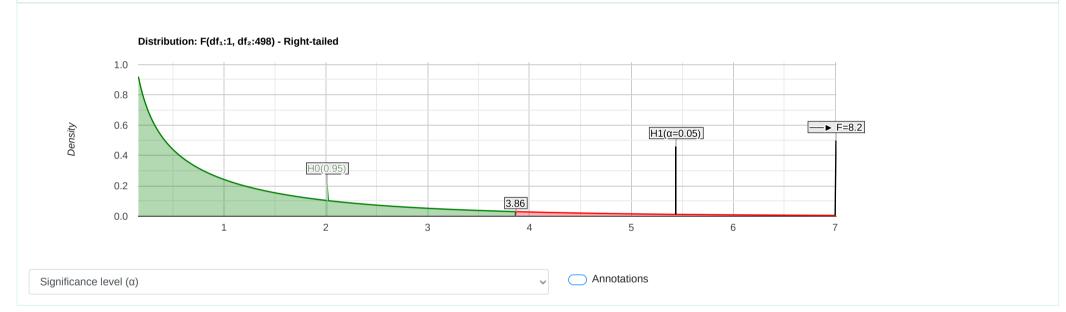
4. Outliers

Outliers may affect the regression line.

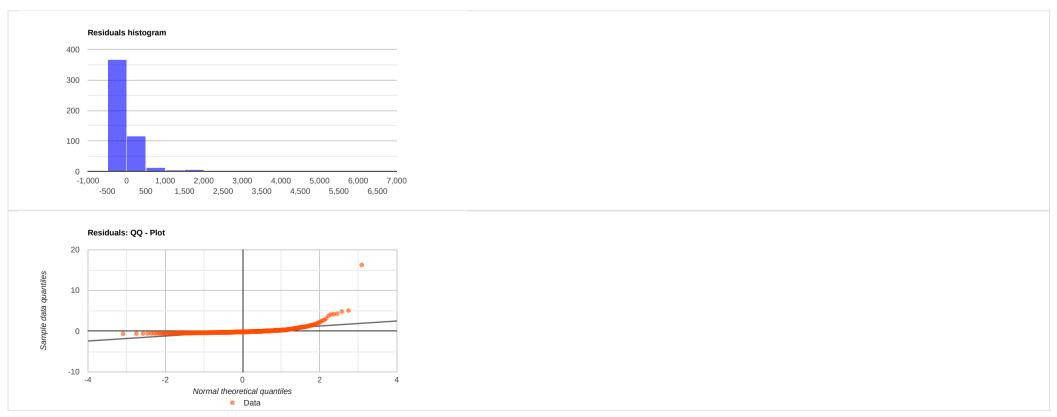
In this case, the distribution of the residuals is normal. Therefore, the probability of detecting 8 valid outliers or more is 0.9999, (outliers:

1163.2975,1407.0676,1558.0676,1530.2975,1597.2975,1811.8377,1901.2975,6115.0676).

You should only remove outliers if you identify them as errors!



Residuals normality



Calculation

 $\bar{x} = 1.88$

Step-by-step solution

$$\hat{Y} = b_0 + b_1 X$$

$$b_1 = \frac{SP_{xy}}{SS_x} = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\Sigma (x_i - \bar{x})^2}$$

$$b_1 = \frac{20596.2}{362.8} = 56.7701$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$\bar{y} = 212.12$$

b₀ = 212.12-56.7701*1.88 = 105.3922

$$R^2 = \frac{SS_{Regression}}{SS_{total}} = \frac{\Sigma (\ \hat{y_i} - \bar{y})^2}{\Sigma (\ y_i - \bar{y})^2} = \frac{1169248.7719}{72144034.8} = 0.01621$$

The standard deviation of the residuals is:

$$MS_{residual} = S_{res}^2 = \frac{\sum (y_i - \hat{y})^2}{n - 2}$$

Residual outliers

 $S_{res} = \sqrt{MSE} = \sqrt{142519.6507} = 377.5177.$

The average of the residuals is always zero.

The thresholds used to calculate the outliers are: $\pm k*S_{res}$.

In this case, the thresholds are $\pm 3*377.5177 = \pm 1132.5532$.

We tagged the outliers with an arrow (\Leftarrow) at the 'Residual' column.

SS_x and SP_{xy}

SOX all a S. xy			
x-x	у-ў	$(x-\bar{x})^2$	(x-x̄)(y-ȳ)

10/23/25, 5:52 PM	Linear regression calcul	ator - calculates the linear regression equation, draws the p	rediction interval, generat
-0.88	-212.12	0.7744	186.6656
-0.88	-212.12	0.7744	186.6656
-0.88	-212.12	0.7744	186.6656
-0.88	-212.12	0.7744	186.6656
-0.88	-212.12	0.7744	186.6656
-0.88	-212.12	0.7744	186.6656
-0.88	-210.12	0.7744	184.9056
-0.88	-210.12	0.7744	184.9056
-0.88	-210.12	0.7744	184.9056
-0.88	-209.12	0.7744	184.0256
-0.88	-209.12	0.7744	184.0256
-0.88	-208.12	0.7744	183.1456
-0.88	-208.12	0.7744	183.1456
-0.88	-207.12	0.7744	182.2656
-0.88	-207.12	0.7744	182.2656
-0.88	-206.12	0.7744	
			181.3856
-0.88	-206.12	0.7744	181.3856
-0.88	-206.12	0.7744	181.3856
-0.88	-205.12	0.7744	180.5056
-0.88	-205.12	0.7744	180.5056
-0.88	-205.12	0.7744	180.5056
-0.88	-205.12	0.7744	180.5056
-0.88	-204.12	0.7744	179.6256
-0.88	-203.12	0.7744	178.7456
-0.88	-203.12	0.7744	178.7456
-0.88	-202.12	0.7744	177.8656
-0.88	-202.12	0.7744	177.8656
-0.88	-202.12	0.7744	177.8656
-0.88	-202.12	0.7744	177.8656
-0.88	-200.12	0.7744	176.1056
-0.88	-198.12	0.7744	174.3456
-0.88	-198.12 107.13	0.7744	174.3456
1.12	-197.12	1.2544	-220.7744
-0.88	-197.12	0.7744	173.4656
-0.88	-195.12	0.7744	171.7056
-0.88	-195.12	0.7744	171.7056
0.12	-195.12	0.0144	-23.4144
-0.88	-195.12	0.7744	171.7056
-0.88	-194.12	0.7744	170.8256
-0.88	-194.12	0.7744	170.8256
-0.88	-194.12	0.7744	170.8256
0.12	-191.12	0.0144	-22.9344
-0.88	-188.12	0.7744	165.5456
-0.88	-187.12	0.7744	164.6656
-0.88	-187.12	0.7744	164.6656
0.12	-184.12	0.0144	-22.0944
-0.88	-184.12	0.7744	162.0256
-0.88	-183.12	0.7744	161.1456
1.12	-182.12	1.2544	-203.9744
-0.88	-180.12	0.7744	158.5056
			158.5056
-0.88	-180.12	0.7744	
-0.88	-178.12	0.7744	156.7456
-0.88	-178.12	0.7744	156.7456
-0.88	-178.12	0.7744	156.7456
-0.88	-177.12	0.7744	155.8656
-0.88	-176.12	0.7744	154.9856
-0.88	-176.12	0.7744	154.9856
-0.88	-174.12	0.7744	153.2256
-0.88	-174.12	0.7744	153.2256
-0.88	-174.12	0.7744	153.2256
0.12	-173.12	0.0144	-20.7744
-0.88	-173.12	0.7744	152.3456
-0.88	-173.12	0.7744	152.3456
-0.88	-173.12	0.7744	152.3456
-0.88	-173.12	0.7744	152.3456
-0.88	-173.12	0.7744	152.3456
-0.88	-172.12	0.7744	151.4656
-0.88	-171.12	0.7744	150.5856
1.12	-170.12	1.2544	-190.5344
-0.88	-170.12	0.7744	149.7056
		0.7744	147.9456
-0.88	-168.12 -168.12		
-0.88	-168.12 -167.12	0.7744	147.9456
-0.88	-167.12	0.7744	147.0656
1.12	-167.12	1.2544	-187.1744
-0.88	-166.12	0.7744	146.1856
-0.88	-164.12	0.7744	144.4256
0.12	-163.12	0.0144	-19.5744
0.12	-163.12	0.0144	-19.5744
0.12	-162.12	0.0144	-19.4544
-0.88	-159.12	0.7744	140.0256
-0.88	-159.12	0.7744	140.0256
-0.88	-159.12	0.7744	140.0256
-0.88	-159.12	0.7744	140.0256
1.12	-158.12	1.2544	-177.0944
-0.88	-158.12	0.7744	139.1456
-0.88	-158.12	0.7744	139.1456
-0.88	-157.12	0.7744	138.2656
1.12	-157.12	1.2544	-175.9744
-0.88	-156.12	0.7744	137.3856
-0.88	-156.12	0.7744	137.3856
0.12	-155.12	0.0144	-10.0144
-0.88	-153.12	0.7744	135.6256
0.12	-154.12 -154.12	0.0144	-18.4944
-0.88	-154.12 -154.12	0.7744	135.6256
		0.1144	133.0230
https://www.statskingdom.com	linear-regression-calculator.html		

3/25, 5:52 PM		regression calculator - calculates the linear regressi	
0.88	-153.12	0.7744	134.7
L.12	-153.12	1.2544	-171.
).12	-152.12	0.0144	-18.2
L.12	-152.12	1.2544	-170.
0.88	-151.12	0.7744	132.9
L.12	-150.12	1.2544	-168.
0.88	-149.12	0.7744	131.2
0.88	-149.12	0.7744	131.2
).12	-148.12	0.0144	-17.7
0.88	-147.12	0.7744	129.4
L.12	-147.12	1.2544	-164.
L.12	-146.12	1.2544	-163.
0.88	-146.12	0.7744	128.5
L.12	-146.12	1.2544	-163.
0.12	-145.12	0.0144	-17.4
0.88	-145.12	0.7744	127.7
0.12	-145.12	0.0144	-17.4
12	-145.12	1.2544	-162.
0.88	-145.12	0.7744	127.7
0.88	-144.12	0.7744	126.8
.12	-144.12	0.0144	-17.2
0.12	-144.12	0.0144	-17.2
0.88	-144.12	0.7744	126.8
0.88	-143.12	0.7744	125.9
0.88	-143.12	0.7744	125.9
0.88	-142.12	0.7744	125.0
0.12	-142.12	0.0144	-17.0
).12	-142.12	0.0144	-17.0
.12	-142.12	0.0144	-17.0
0.88	-141.12	0.7744	124.1
.12	-141.12	0.0144	-16.9
12	-141.12	1.2544	-158.
0.12	-141.12	0.0144	-16.9
0.88	-140.12	0.7744	123.3
0.88	-140.12	0.7744	123.3
12	-140.12	1.2544	-156.
12	-140.12	1.2544	-156.
0.88	-139.12	0.7744	122.4
12	-139.12	1.2544	-155.
12	-139.12	1.2544	-155.
0.88	-138.12	0.7744	121.5
12	-137.12	1.2544	-153.
0.88	-137.12	0.7744	120.6
.12	-137.12	0.0144	-16.4
12	-137.12	1.2544	-153.
0.88	-137.12	0.7744	120.6
.12	-137.12	0.0144	-16.4
12	-135.12	1.2544	-151.
12	-133.12	1.2544	-149.
0.12	-133.12	0.0144	-15.9
12	-133.12	1.2544	-149.
12	-132.12	1.2544	-147.
0.12	-131.12	0.0144	-15.7
0.88	-131.12	0.7744	115.3
0.88	-131.12	0.7744	115.3
.12	-130.12	1.2544	-145.
0.88	-129.12	0.7744	113.6
.12	-129.12	1.2544	-144.
.12	-129.12	1.2544	-144.
.12	-129.12	1.2544	-144.
0.88	-128.12	0.7744	112.7
.12	-128.12	1.2544	-143.
	-128.12	1.2544	-143.
.12			
.12	-128.12	1.2544	-143.
0.88	-127.12	0.7744	111.8
.12	-127.12	0.0144	-15.2
.12	-127.12	1.2544	-142.
.12	-126.12	0.0144	-15.1
0.88	-126.12	0.7744	110.9
0.88	-126.12	0.7744	110.9
0.88	-125.12	0.7744	110.1
.12	-125.12	0.0144	-15.0
0.88	-125.12	0.7744	110.1
.12	-124.12	0.0144	-14.8
.12	-123.12	0.0144	-14.6
.12	-123.12	1.2544	-137.
0.88	-122.12	0.7744	107.4
.12	-122.12	0.0144	-14.6
.12	-122.12	0.0144	-14.6
.12	-122.12	1.2544	-136.
.12	-122.12	1.2544	-136.
.12	-121.12	0.0144	-14.5
.12	-121.12	1.2544	-135.
).88	-120.12	0.7744	105.7
0.88	-119.12	0.7744	104.8
.12	-118.12	1.2544	-132.
.12	-118.12	0.0144	-14.1
.12	-118.12	0.0144	-14.1
.12	-117.12	0.0144	-14.0
0.88	-117.12	0.7744	103.0
		1 2544	-130
.12	-116.12	1 23///	
	-116.12	0.7744	102.1
0.88			
	-116.12 -115.12	0.0144 0.7744	-13.9 101.3

0.13	Linea	r regression calculator - calculates the linear regression	requation, araws are prediction
0.12	19.88	0.0144	2.3856
0.12	19.88	0.0144	2.3856
-0.88	20.88	0.7744	-18.374
-0.88	21.88	0.7744	-19.254
-0.88	21.88	0.7744	-19.254
0.12	22.88	0.0144	2.7456
-0.88	23.88	0.7744	-21.014
-0.88			-21.894
	24.88	0.7744	
0.12	25.88	0.0144	3.1056
1.12	25.88	1.2544	28.9856
0.12	25.88	0.0144	3.1056
-0.88	29.88	0.7744	-26.294
1.12	29.88	1.2544	33.4656
0.12	30.88	0.0144	3.7056
1.12	30.88	1.2544	34.5856
-0.88	32.88	0.7744	-28.934
-0.88	36.88	0.7744	-32.454
1.12	38.88	1.2544	43.5456
0.12	39.88	0.0144	4.7856
-0.88	39.88	0.7744	-35.094
0.12	39.88	0.0144	4.7856
1.12	41.88	1.2544	46.9056
-0.88	42.88	0.7744	-37.734
1.12	44.88	1.2544	50.2656
0.12	49.88	0.0144	5.9856
0.12	50.88	0.0144	6.1056
1.12	51.88	1.2544	58.1056
1.12	52.88	1.2544	59.2256
-0.88	53.88	0.7744	-47.414
0.12	56.88	0.0144	6.8256
-0.88	56.88	0.7744	-50.054
0.12	58.88	0.0144	7.0656
1.12	58.88	1.2544	65.9456
1.12	58.88	1.2544	65.9456
-0.88	58.88	0.7744	-51.814
1.12	60.88		
		1.2544	68.1856
1.12	61.88	1.2544	69.3056
-0.88	63.88	0.7744	-56.214
1.12	66.88	1.2544	74.9056
1.12	67.88	1.2544	76.0256
0.12	69.88	0.0144	8.3856
1.12	69.88	1.2544	78.2656
0.12	72.88	0.0144	8.7456
1.12	74.88	1.2544	83.8656
0.12	75.88	0.0144	9.1056
-0.88	77.88	0.7744	-68.534
1.12	81.88	1.2544	91.7056
1.12	82.88	1.2544	92.8256
0.12	86.88	0.0144	10.4256
0.12	87.88	0.0144	10.5456
-0.88	92.88	0.7744	-81.734
0.12	96.88	0.0144	11.6256
1.12	98.88	1.2544	110.745
1.12	98.88	1.2544	110.745
1.12	100.88	1.2544	112.985
-0.88	100.88	0.7744	-88.774
1.12	103.88	1.2544	116.345
0.12	117.88	0.0144	14.1456
1.12	122.88	1.2544	137.625
-0.88	124.88	0.7744	-109.89
1.12	126.88	1.2544	142.105
1.12	129.88	1.2544	145.465
1.12	142.88	1.2544	160.025
1.12	150.88	1.2544	168.985
-0.88	154.88	0.7744	-136.29
-0.88	163.88	0.7744	-144.21
0.12	165.88	0.0144	19.9056
1.12	165.88	1.2544	185.785
0.12		0.0144	20.9856
	174.88		
-0.88	177.88	0.7744	-156.53
0.12	180.88	0.0144	21.7056
0.12	191.88	0.0144	23.0256
1.12	196.88	1.2544	220.505
-0.88	196.88	0.7744	-173.25
1.12	197.88	1.2544	221.625
-0.88	207.88	0.7744	-182.93
-0.88	208.88	0.7744	-183.81
1.12	216.88	1.2544	242.905
-0.88	217.88	0.7744	-191.73
1.12	228.88	1.2544	256.345
1.12	234.88	1.2544	263.065
0.12	241.88	0.0144	29.0256
0.12	243.88	0.0144	29.2656
-0.88	245.88	0.7744	-216.37
1.12	257.88	1.2544	288.825
	259.88	0.7744	-228.69
-0.88	277.88	0.7744	-244.53
-0.88	281.88	0.0144	33.825
-0.88 -0.88	201.00		-257.73
-0.88 -0.88 0.12	202.00	111 ((((((((((((((((((- 26 / 72
-0.88 -0.88 0.12 -0.88	292.88	0.7744	
-0.88 -0.88 0.12 -0.88 0.12	302.88	0.0144	36.3456
-0.88 -0.88 0.12 -0.88			36.3456
-0.88 -0.88 0.12 -0.88 0.12	302.88	0.0144	36.3456 -277 07 -297.33
-0.88 -0.88 0.12 -0.88 0.12 -0.88	302.88 315.88	0.0144	36.3456 -277 07

0	0	362.8 (SS _x)	20596.2 (SP _{xy})
0.12	6121.88	0.0144	734.6256
1.12	1964.88	1.2544	2200.6656
-0.88	1761.88	0.7744	-1550.4544
1.12	1660.88	1.2544	1860.1856
1.12	1593.88	1.2544	1785.1456
0.12	1564.88	0.0144	187.7856
0.12	1413.88	0.0144	169.6656
1.12	1226.88	1.2544	1374.1056
1.12	1077.88	1.2544	1207.2256
1.12	961.88	1.2544	1077.3056
-0.88	923.88	0.7744	-813.0144
0.12	796.88	0.0144	95.6256
1.12	735.88	1.2544	824.1856
-0.88	718.88	0.7744	-632.6144
0.12	668.88	0.0144	80.2656
1.12	610.88	1.2544	684.1856
1.12	605.88	1.2544	678.5856
-0.88	574.88	0.7744	-505.8944
1.12	558.88	1.2544	625.9456
0.12	517.88	0.0144	62.1456
-0.88	500.88	0.7744	-440.7744
1.12	478.88	1.2544	536.3456
-0.88	420.88	0.7744	-370.3744
1.12	411.88	1.2544	461.3056
0.12	409.88	0.0144	49.1856
-0.88	392.88	0.7744	-345.7344
0.12	386.88	0.0144	46.4256
1.12	383.88	1.2544	429.9456
-0.88 0.12	374.88 374.88	0.7744 0.0144	-329.8944 44.9856

Linear regression calculator

The linear regression calculator generates the linear regression equation. It also draws: a linear regression line, a histogram, a residuals QQ-plot, a residuals x-plot, and a distribution chart. It calculates the R-squared, the R, and the outliers, then testing the fit of the linear model to the data and checking the residuals' normality assumption and the priori power.

What is linear regression?

The linear regression is the linear equation that best fits the points.

There is no one way to choose the best fit ting line, the most common one is the ordinary least squares (OLS). The linear regression describes the relationship between the dependent variable (Y) and the independent variables (X).

The linear regression model calculates the dependent variable (DV) based on the independent variables (IV, predictors).

What is "ordinary least squares"?

The ordinary least squares method chooses the line parameters that minimize the sum of squares of the differences between the observed dependent variables (Y) and the estimated value by the linear regression (Ŷ).

Why do you need linear regression?

We may use linear regression when we want to do one of the following

- Predict the dependent variable (\hat{Y}) .
- Estimate the effect of each independent variable (X) on the dependent variable (Y).
- Calculate the correlation between the dependent variable and the independent variables.
- Test the linear model significance level.

How to calculate linear regression?

Following the linear regression formula:

$$\hat{\mathbf{Y}} = \mathbf{b}_0 + \mathbf{b}_1 \mathbf{x}$$

 $\ensuremath{b_0}$ - the y-intercept, where the line crosses the y-axis.

b₁ - the slope, describes the line's direction and incline.

$$b_1 = \frac{SP_{xy}}{SS_x} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2}$$

$$b_0=\bar{y}-b_1\bar{x}$$

linear regression prediction

The prediction calculator uses the linear regrssion to predict the depdendent variable based on the independent value. The calculator also creates the confidence interval, and the prediction interval.

Confidence interval of the prediction

The prediction interval for the **mean value** of the dependent variable.

This is the interval for the equation line, the true value equation will be in this interval. If we would know the true equation then the width of this interval would be zero.

If you would calculate the confidence interval over an infinite number of regressions with the same sample size, 95% (confidence level) of the calculated confidence intervals will contain the mean's true value. Since this interval is for the mean, the standard error is smaller and the the range is narrower than the range of the prediction interval.

$$\begin{aligned} MS_{residual} &= S^2_{residual} = \frac{\Sigma (y_i - \hat{y})^2}{n - 2} \\ S.E^2_{ci} &= S^2_{residual} (\frac{1}{n} + \frac{(x_0 - \hat{x})^2}{SS_X}) \end{aligned}$$

$$\hat{Y} \pm T_{1-\alpha/2}(n-2)*S.E_{ci}$$

Prediction Interval

The prediction interval for a ${\bf particular\ observation}$ of the dependent variable.

This is the interval for any single value.

The prediction inteval takes into consideration the fact that you don't know the true equatio, and the fact the the liner regression explaned only part of the variance (the part is R-squared). Even if we would know the true equation then the width of this interval would be greater than zero.

Since this interval is for a single observation, the standard error is larger and the range is wider than the range of the confidence interval

$$S.E_{prediction}^2 = S_{residual}^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_x} \right)$$

 $\hat{Y} \pm T_{1\text{-}\alpha/2} (n\text{-}2) * S. E_{prediction}$

How to calculate R squares?

R squares is the percentage of the variance explain by the regression (SS_{Regression}) from the overall variance (SS_{Total}).

$$R^2 = \frac{SS_{Regression}}{SS_{Total}}$$

Linear regression in calculator

This online calculator supports all the basic functionality and more.

The right-tailed F test checks if the entire regression model is statistically significant. Why only right tail?

For Multiple regression calculator with the stepwise method and assumptions validations: multiple regression calculator

The following statistic checks if the linear regression model supports better results than the average of Y.

Hypotheses

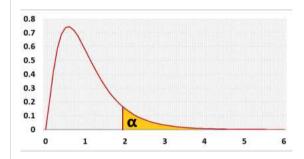
 H_0 : Y = b_0

 H_1 : $Y = b_0 + b_1 X$

Test statistic

 $F = \frac{MS(regression)}{1}$ MS (residual)

F distribution



R Code

The following R code should produce similar results

rm(list = ls())

if(!"car" %in% installed.packages()){install.packages("car")}

library(car)

x10 <-

x11 <-

x1 <- c(x10,x11)

y10 <-

y1 <- c(y10,y11)

model1 = Im(y1~x1)

summary(model1)

What is linear regression?

Tutorial

Calculators

Correlation

Regression sample size