Bookmark this page

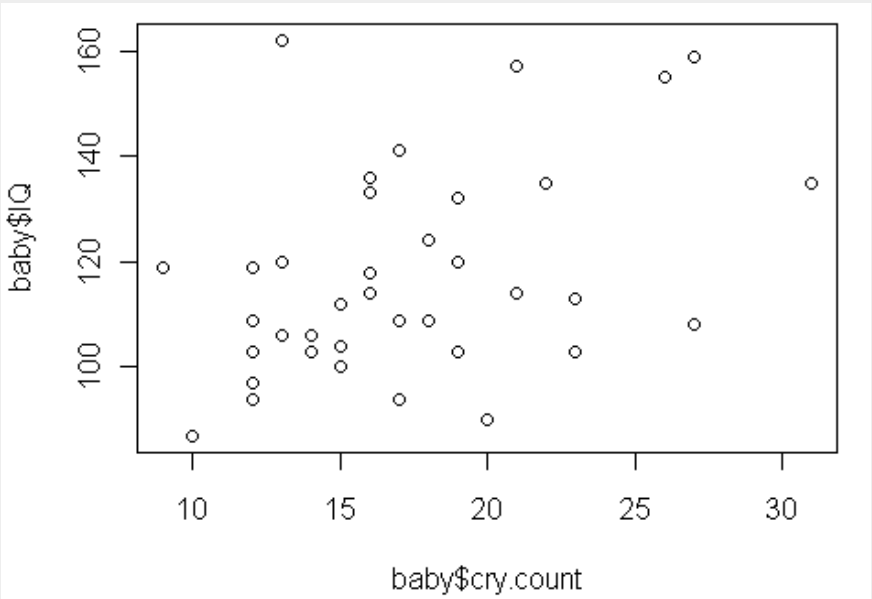
Statistics Package Exercise: Making Predictions with a Least Squares Regression Line

Learning Objective: In the special case of linear relationship, use the least squares regression line as a summary of the overall pattern, and use it to make predictions.

The purpose of this activity is to complete our discussion about our example that examines the relationship between vocalization soon after birth and IQ at age three. So far we explored the data using a scatterplot supplemented with the correlation r:

- R
- StatCrunch
- TI Calculator
- Minitab
- Excel

R



and discovered that the data display a moderately weak positive linear relationship. In addition, when we carried out the t-test for assessing the significance of this linear relationship:

- R StatCrunch TI Calculator Minitab Excel

R

```
> summary(lm(IQ~cry.count,baby))

Call:
lm(formula = IQ ~ cry.count, data = baby)

Residuals:
    Min      1Q  Median      3Q     Max
-31.482 -11.593  -2.591  10.286  51.272

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  90.7550    10.4734   8.665 2.47e-10 ***
cry.count     1.5364     0.5835   2.633   0.0124 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.99 on 36 degrees of freedom
Multiple R-squared: 0.1615,     Adjusted R-squared: 0.1382
F-statistic: 6.932 on 1 and 36 DF,  p-value: 0.01240
```

We concluded (based on the small p-value of 0.012) that the data provide fairly strong evidence of a moderately weak linear relationship between cry count soon after birth and IQ and age 3.

We would now like to consider using the least squares regression line for predicting IQ at age 3 based on cry count soon after birth. Plot the least squares regression line on the scatterplot:

- R StatCrunch TI Calculator Minitab Excel

R Instructions

To open R with the data set preloaded, right-click here and choose "Save Target As" to download the file to your computer. Then find the downloaded file and double-click it to open it in R.

The data have been loaded into the data frame

```
baby
```

. Enter the command

```
baby
```

to see the data. The two variables in the data frame are

```
cry.count
```

and

```
IQ
```

.

Let's rerun the previous code to acquire the plot and equation.

- ```
  plot(baby$cry.count,baby$IQ, xlab="Number of Crying Events", ylab="IQ")
  ```
- ```
  model=lm(baby$IQ~baby$cry.count)
  ```
- ```
  abline(model)
  ```
- ```
  cf=coef(model)
  ```
- ```
  cf
  ```
- ```
  legend(10,160,legend=paste("IQ = ", round(cf[1],1), "+", round(cf[2],2), "(cry.count)"))
  ```

Learn By Doing (1/1 point)

Would you say that the least squares regression line fits the data well?

Your Answer:

No because it looks bad

Our Answer:

RStatCrunch TI CalculatorMinitabExcel R The least squares regression line is IQ = 90.8 + 1.54(cry count), and this graph displays it on the scatterplot: While the line captures the general linear trend of the data, it does not fit the data very well due to the moderately weak linearity in the data. Visually, we see that the data points do not lay close to the line, resulting in a relatively poor fit of the line to the data. StatCrunch The least squares regression line is IQ = 90.8 + 1.54(cry count), and this graph displays it on the scatterplot. While the line captures the general linear trend of the data, it does not fit the data very well due to the moderately weak linearity in the data. Visually, we see that the data points do not lay close to the line, resulting in a relatively poor fit of the line to the data. TI Calculator The least squares regression line is IQ=90.8+1.54(cry count), and this graph displays it on the scatterplot: While the line captures the general linear trend of the data, it does not fit the data very well due to the moderately weak linearity in the data. Visually, we see that the data points do not lay close to the line, resulting in a relatively poor fit of the line to the data. Minitab The least squares regression line is IQ = 90.8 + 1.54(cry count), and this graph displays it on the scatterplot: While the line captures the general linear trend of the data, it does not fit the data very well due to the moderately weak linearity in the data. Visually, we see that the data points do not lay close to the line, resulting in a relatively poor fit of the line to the data. Excel The least squares regression line is IQ = 90.8 + 1.54(cry count), and this graph displays it on the scatterplot: While the line captures the general linear trend of the data, it does not fit the data very well due to the moderately weak linearity in the data. Visually, we see that the data points do not lay close to the line resulting in a relatively poor fit of the line to the data.

Resubmit Reset

Learn By Doing (1/1 point)

Use the least squares regression line to predict the IQ at age 3 of a newborn whose cry count is 25. Based on our discussion in question 1, comment on the accuracy of your prediction.

Your Answer:

plugging in the value, we got 129.3. Accuracy will probably not be that great because the linear relationship within the data is already not that strong to begin with.

Our Answer:

RStatCrunch TI CalculatorMinitabExcel R The predicted IQ is obtained by plugging in a cry count of 25 into the regression line. We therefore obtain: Predicted IQ = 90.8 + 1.54 * 25 = 129.3. Here is the graph:. It should be noted, however, that due to the moderately weak linearity, this prediction is not very accurate in the sense that the actual IQ level can be either much lower or much higher than this predicted value (in the same way that many of the data points are well below or above the line). In fact, given how far the data points are from the line, the IQ could be roughly anywhere between 90 and 167 as this graph displays: Our prediction of the IQ being 129 is therefore not very accurate. StatCrunch The predicted IQ is obtained by plugging in a cry count of 25 into the regression line. We therefore obtain: Predicted IQ = 90.8 + 1.54 * 25 = 129.3. Here is the graph: It should be noted, however, that due to the moderately weak linearity, this prediction is not very accurate in the sense that the actual IQ level can be either much lower or much higher than this predicted value (in the same way that many of the data points are well below or above the line). In fact, given how far the data points are from the line, the IQ

could be roughly anywhere between 90 and 167 as this graph displays: Our prediction of the IQ being 129 is therefore not very accurate. TI Calculator The predicted IQ is obtained by plugging in a cry count of 25 into the regression line. We therefore obtain: Predicted IQ = 90.8 + 1.54 * 25 = 129.3. Here is the graph: It should be noted, however, that due to the moderately weak linearity, this prediction is not very accurate in the sense that the actual IQ level can be either much lower or much higher than this predicted value (in the same way that many of the data points are well below or above the line). In fact, given how far the data points are from the line, the IQ could be roughly anywhere between 98 and 167 as this graph displays: Our prediction of the IQ being 129 is therefore not very accurate. Minitab The predicted IQ is obtained by plugging in a cry count of 25 into the regression line. We therefore obtain: Predicted IQ = 90.8 + 1.54 * 25 = 129.3. Here is the graph: It should be noted, however, that due to the moderately weak linearity, this prediction is not very accurate in the sense that the actual IQ level can be either much lower or much higher than this predicted value (in the same way that many of the data points are well below or above the line). In fact, given how far the data points are from the line, the IQ could be roughly anywhere between 90 and 167 as this graph displays: Our prediction of the IQ being 129 is therefore not very accurate. Excel The predicted IQ is obtained by plugging in a cry count of 25 into the regression line. We therefore obtain: Predicted IQ = 90.8 + 1.54 * 25 = 129.3. Here is the graph: It should be noted, however, that due to the moderately weak linearity, this prediction is not very accurate in the sense that the actual IQ level can be either much lower or much higher than this predicted value (in the same way that many of the data points are well below or above the line). In fact, given how far the data points are from the line, the IQ could be roughly anywhere between 98 and 167 as this graph displays: Our prediction of the IQ being 129 is therefore not very accurate.

Resubmit Reset