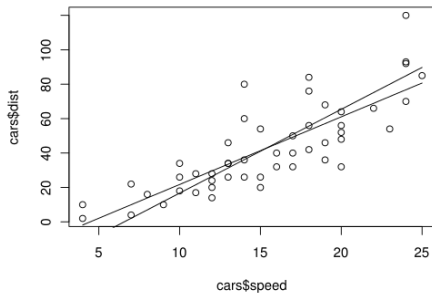


Math 314: Statistics

Chapter 10: Regression



Dr. Ralph Wojtowicz

CME Department

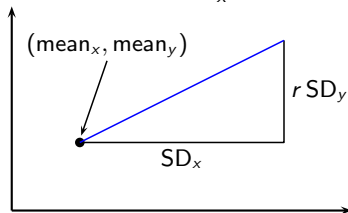
Shepherd
UNIVERSITY

- 1 Regression Line
 - The Regression Line
 - Using R to Find a Regression Line
 - Plot the SD and Regression Lines
 - Cars Data

The Regression Line

- The regression line for y on x estimates the average value for y corresponding to each value of x .
- Associated with each increase of one SD in x , there is an increase of only r SDs in y , on the average.
- To see why r is the right factor, consider the cases $r = 0$, $r = 1$ and $r = -1$.
- The regression line is a smoothed version of the graph of averages.
- The equation for the regression line is (where r = the correlation coefficient):

$$(y - \text{mean}_y) = r \frac{\text{SD}_y}{\text{SD}_x} (x - \text{mean}_x)$$



Using R to Find a Regression Line

- Example from pages 132–133.

```
> x <- c(1, 3, 4, 5, 7)
> y <- c(5, 9, 7, 1, 13)
> source("http://www.adjoint-functors.net/SDline.R")
> SDline(x, y)
    $meanX=4, $meanY=7, $slope=2, $correlation=0.4

> linearModel <- lm(y~x)
> summary(linearModel)
```

Coefficients:

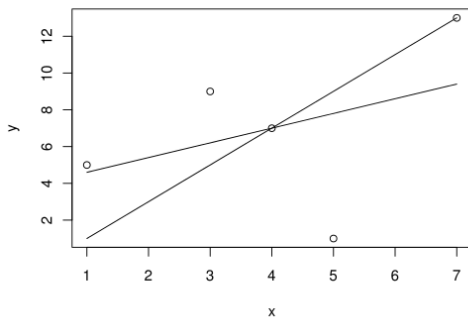
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.800	4.733	0.803	0.481
x	0.800	1.058	0.756	0.505

Residual standard error: 4.733 on 3 degrees of freedom
 Multile R-squared: 0.16, Adjusted R-squared: -0.12
 F-statistic: 0.5714 on 1 and 3 DF, p-value: 0.5046

Plot the SD and Regression Lines

- Example from pages 132–133

```
> x <- c(1, 3, 4, 5, 7)
> y <- c(5, 9, 7, 1, 13)
> plot(x, y)
> lines(x, 2*x - 1, type="l")
> lines(x, 0.8*x + 3.8, type="l")
```



Cars Data

- Use the following to get the equation for the SD line of the cars data
- `SDline(cars$speed, cars$dist)`
- `plot(cars$speed, cars$dist)`
- Add the SD line to your plot.
- Use the following to get the regression line
- `linearModel <- lm(cars$dist~cars$speed)`
- Add the regression line to your plot.