

Q6 review

Stat 301

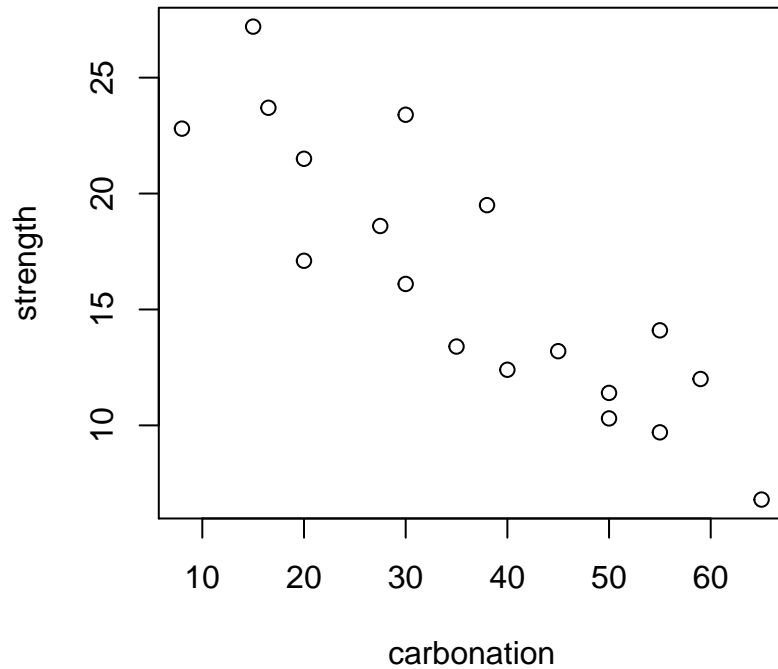
Summer 2019

- (1) *Carbonation*: Corrosion of steel reinforcing bars is the most important durability problem for reinforcing structures. Carbonation of concrete results from a chemical reaction that lowers the pH value by enough to initiate corrosion of the rebar. Data on the carbonation depth (*mm*) and strength (*MPa*) for a sample of core specimens was taken from a particular building, and all the regression output is provided. We are interested in modeling the strength ¹.
- (a) State the regression model and define its components
 - (b) Looking at the raw data scatterplot, does it appear as if there is a linear relationship? Positive or negative slope?
 - (c) State the equation of the regression equation (from output). Use it to estimate the strength when the carbonation depth is 8 *mm* and estimate it again when the depth is 20 *mm*
 - (d) Calculate the residuals for both of your estimates in part c. The observed value for 8 *mm* is 22.8 *MPa* ((8, 22.8)) and for 20 *mm* is 17.1 *MPa* ((20, 17.1))
 - (e) Interpret slope and intercept *in context*. If something does not make sense in context, state it and describe why
 - (f) Do a significance test of the slope. State hypotheses, *t* statistic, *pvalue*, results, and conclusion of the test.
 - (g) State, define, and describe R^2 and r (R^2 is on the output and r will require a calculation from the output).
 - (h) List assumptions of regression. Are the assumptions of regression met? Briefly explain how each assumption is met or not
 - (i) How is the model? Good, bad, etc.? Give specific evidence (use answers from parts *f*, *g*, and *h*)

¹“The Carbonation of Concrete Structures in the Tropical Environment of Singapore” (*Magazine of Concrete Research*, 1996: 293-300)

```
carbonation=c(8,15,16.5,20,20,27.5,30,30,35,38,40,45,50,50,55,55,59,65)
strength=c(22.8,27.2,23.7,17.1,21.5,18.6,16.1,23.4,13.4,19.5,12.4,13.2,11.4,10.3,14.1,9.7,12,6.8)
plot(strength~carbonation,main='Raw data scatterplot')
```

Raw data scatterplot



```
soda=lm(strength~carbonation); summary(soda)
```

Call:

```
lm(formula = strength ~ carbonation)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.1317	-2.0043	-0.7488	2.1366	5.1439

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	27.18294	1.65135	16.461	1.88e-11 ***
carbonation	-0.29756	0.04116	-7.229	2.01e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

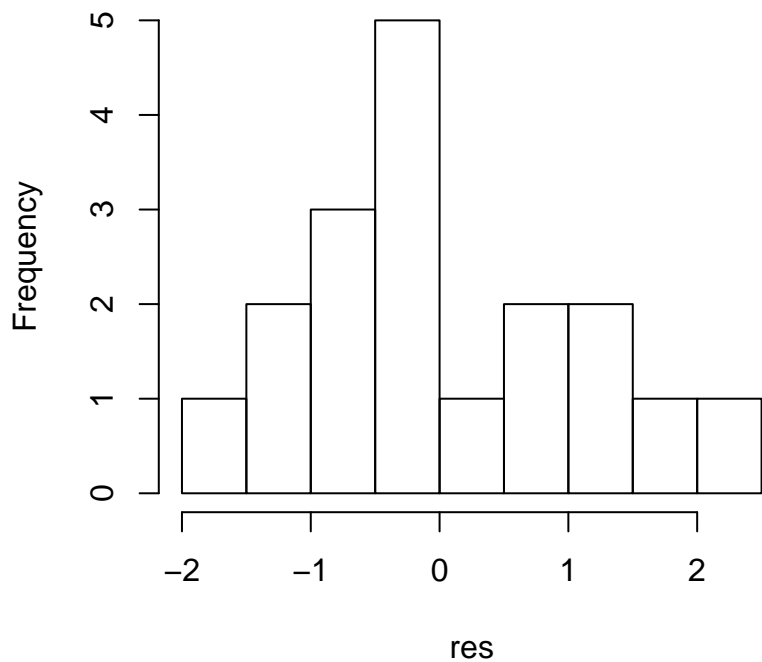
Residual standard error: 2.864 on 16 degrees of freedom

Multiple R-squared: 0.7656, Adjusted R-squared: 0.7509

F-statistic: 52.25 on 1 and 16 DF, p-value: 2.013e-06

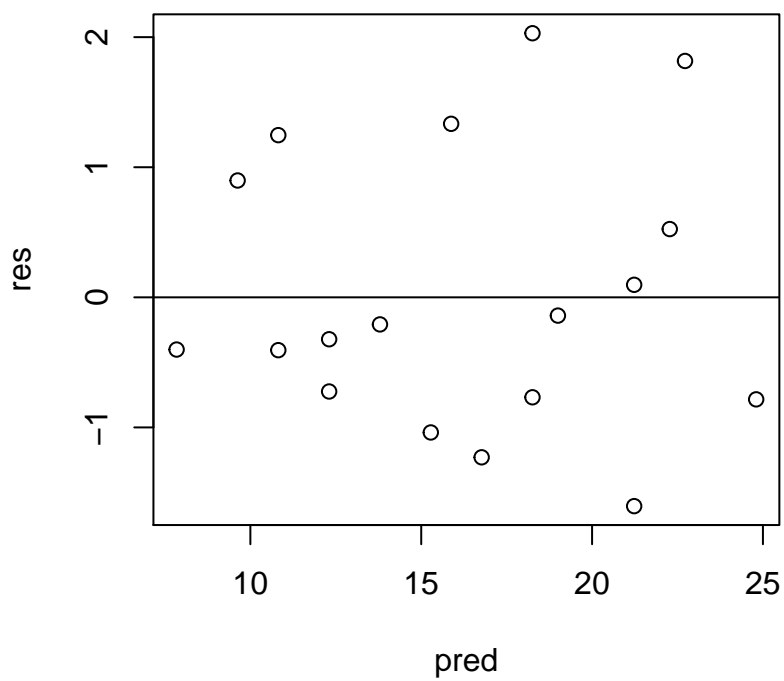
```
res=rstudent(soda); pred=fitted(soda)
hist(res,main='Residuals')
```

Residuals



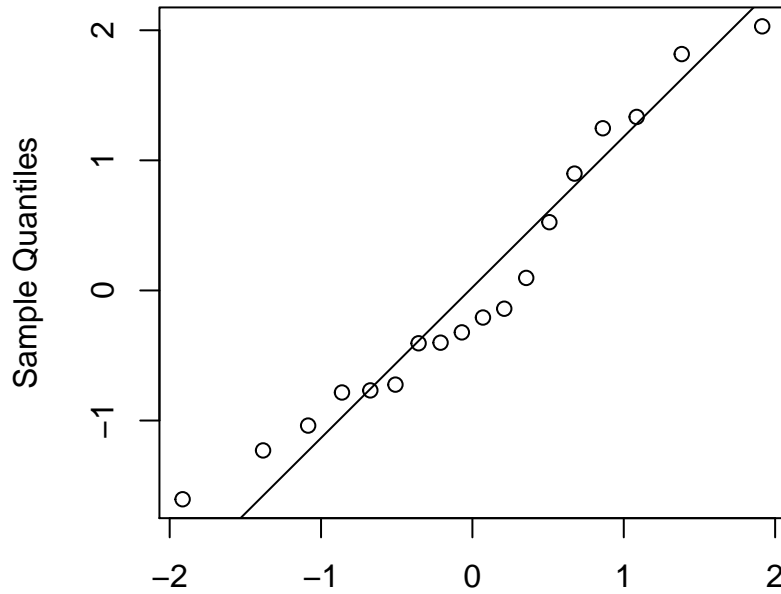
```
plot(pred,res,main='Residuals vs. Predicted'); abline(0,0)
```

Residuals vs. Predicted



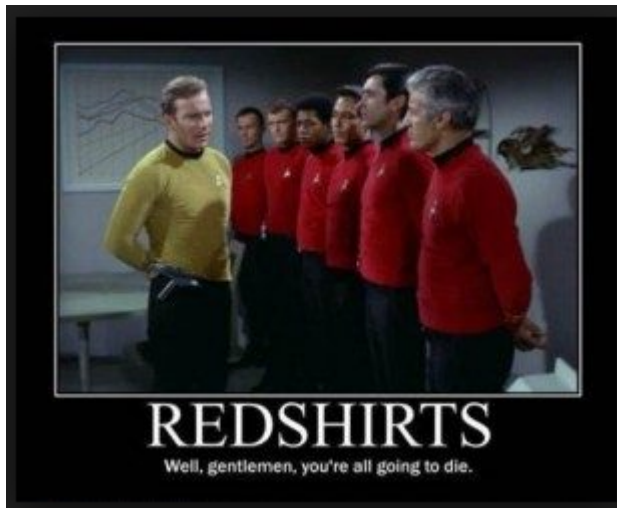
```
qqnorm(res); qqline(res)
```

Normal Q-Q Plot



Theoretical Quantiles

- (2) *Red Shirt of Doooooom*: In Star Trek fandom, there is a running joke that characters on the show who wear a red shirt are doomed^[2]; just another statistic. Shirt colors can be only blue, gold, or red; fatalities can be only dead or alive.
- Is there sufficient evidence determine whether there is an association between shirt color and deaths?
 - State the kind of error that could have been made. *Describe in context*



	Shirt.Colour			
Survival	Blue	Red	Gold	Total
Alive	129	215	46	390
Dead	7	24	9	40
Total	136	239	55	430

Star Trek survival by shirt colour

(3) *Here be dragons*: An analysis of dragon reserve accident data was made to determine if there is a relationship between the type of accident (fatal or non-fatal) and the location of the dragon reserve (Romania, Canada, Australia). Is there sufficient evidence that more fatal injuries happen at one or two specific reserves? The data for 346 accidents are shown in the accompanying table.

- Is there sufficient evidence determine if injury type is the same at all dragon reserve locations?
- State the kind of error that could have been made. *Describe in context*



	Location			
Survival	Romania	Canada	Australia	Total
Fatal	128	63	46	237
Non Fatal	67	26	16	109
Total	195	89	62	346

Dragon Reserve Injuries by Location

- (4) *Book Mediums:* (not the psychic kind of medium) A professor of an introductory college class uses an open-source textbook for the class. Of interest is the proportions of students that will either purchase a hard copy, print the book online, or just use the downloaded PDF format to read on a device. From earlier semesters, 60% bought a hard copy of the book, 25% printed it online, and 15% used a downloaded PDF format on their devices. At the end of the semester, the professor asks the students to complete a survey and indicate what format of the book they used. Of the 126 students, 71 bought a hard copy, 30 printed it, and 25 downloaded PDF to use.
- (a) Is there evidence that the students used similar mediums for the book?
 - (b) State the kind of error that could have been made. *Describe in context*

Data	Type		
	Hard copy	Printed	PDF
Observed	71.00	30.00	25.00
Probabilities	0.60	0.25	0.15