Homework 8

Pranav Belligundu(psb898) - SDS 315 UT Austin

https://github.com/pranav-B21/SDS-315

Problem 1: regression warm up

A. Load and examine the data

```
##
## Call:
## lm(formula = creatclear ~ age, data = creatinine)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.2249  -4.6175   0.2221   4.7212  15.8221
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 147.81292    1.37965  107.14   <2e-16 ***
## age          -0.61982    0.03475  -17.84   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.911 on 155 degrees of freedom
## Multiple R-squared:  0.6724, Adjusted R-squared:  0.6703
## F-statistic: 318.2 on 1 and 155 DF,  p-value: < 2.2e-16
```

```
##       1
## 113.723
```

```
##        age
## -0.6198159
```

```
##        1
## 11.97972
```

```
##        1
## 1.376035
```

A

The expected value for a 55-year old is 133.72(mL/minute) clearance rate and I calcualted this through the lm function which gives me the intercept and the slope –> clearance rate = -0.6198159(age) + 147.81292.

B

For each increase in age by one, the creatine clearance rate decreases by -0.62 ml/minute per year with age. I got this value from the slope of the lin reg model.

C

The difference in creatine rate for the 40 year old with a rate of 135 is 11.97 and the difference in creatine rate for the 60 year old with a rate of 112 is 1.38. Because the differnce of the 40 year old is higher, they are more healthier for their age. I determined this by using the model to predict it based on their ages and subtracted the actual rate to find the difference.

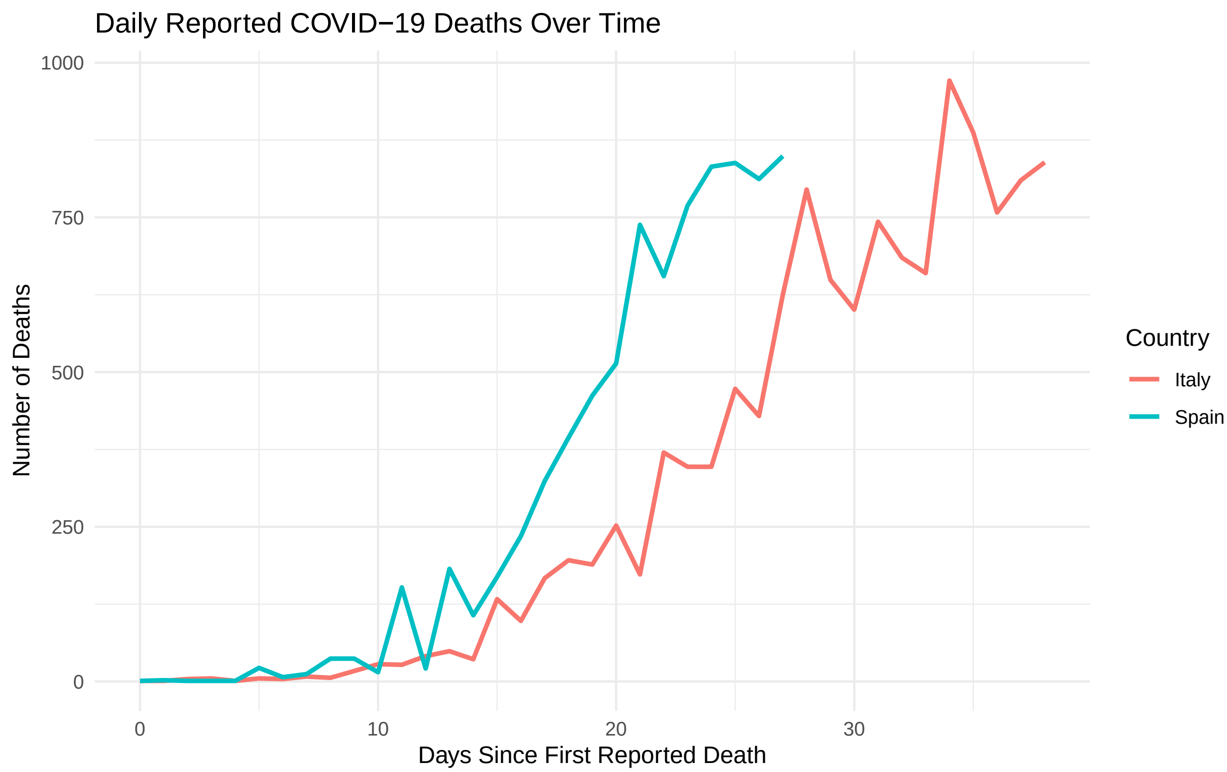
Problem 2: Modeling disease growth

```
##                     name       lower       upper level     method    estimate
## 1              Intercept   0.5437208   1.6044025  0.95 percentile   1.0186023
## 2 days_since_first_death   0.1595732   0.2082654  0.95 percentile   0.1832180
## 3                  sigma   0.5584201   0.8345044  0.95 percentile   0.7248213
## 4              r.squared   0.8543197   0.9328082  0.95 percentile   0.8950791
## 5                      F 216.9807351 513.6625229  0.95 percentile 315.6466194
```

Italy: The growth rate with a 95% CI is between 0.159 and 0.208. The doubling time with a 95% CI is between 3.3 and 4.4 days.

```
##                     name       lower       upper level     method    estimate
## 1              Intercept  -0.1595855   1.2463876  0.95 percentile   0.4652173
## 2 days_since_first_death   0.2351316   0.3182956  0.95 percentile   0.2762447
## 3                  sigma   0.5930165   0.9501547  0.95 percentile   0.8168767
## 4              r.squared   0.8317826   0.9398591  0.95 percentile   0.8893316
## 5                      F 128.5619321 406.3180157  0.95 percentile 208.9360824
```

Spain: The growth rate with a 95% CI is between 0.234 and 0.318. The doubling time with a 95% CI is between 2.2 and 3.0 days.



 # Problem 3: price elasticity of demand

```
##        name       lower       upper level     method    estimate
## 1 Intercept   4.5363627   4.8939185  0.95 percentile   4.7206042
## 2 log_price  -1.7749846  -1.4517843  0.95 percentile  -1.6185778
## 3     sigma   0.2314283   0.3001677  0.95 percentile   0.2687036
## 4 r.squared   0.6875275   0.8447800  0.95 percentile   0.7772187
## 5         F 250.8321060 620.4414497  0.95 percentile 397.7126271
```

The estimated price elasticity of demand for milk is -1.62, with a 95% bootstrapped confidence interval between -1.77 and -1.45. To estimate this, I log-transformed both price and sales to linearize the power-law demand model Q=KP^ , then used linear regression to estimate as the slope of log_sales on log_price. I computed the confidence interval using 10,000 bootstrap resamples.