

SDS315_HW8: https://github.com/Panda-nny/SDS315_HW8

Danny Pan, EID: dp36627

2025-04-13

Problem 1 - Regression Warm Up

A linear regression model was fitted to the given dataset comparing changes in creatinine clearance rates with age. The fitted line's equation is modeled: creatine clearance rate (in mL/minute) = $147.81 - 0.62 * \text{age}$. Thus, there would be an expected 0.62 mL/minute decrease in creatinine clearance rate per year.

Plugging in an age of 55 into this equation results in an expected creatinine clearance rate of about $147.81 - 0.62 * 55$ or 113.72 mL/minute for a 55-year-old.

In order to compare whether two adults are “healthier” for their age, residuals are computed taking the difference between their current rate and the expected rate for their age using the linear regression model. Since higher creatinine clearance rates are considered healthier despite older people naturally having a lower rate, an adult is considered healthier for their age if their residual is higher.

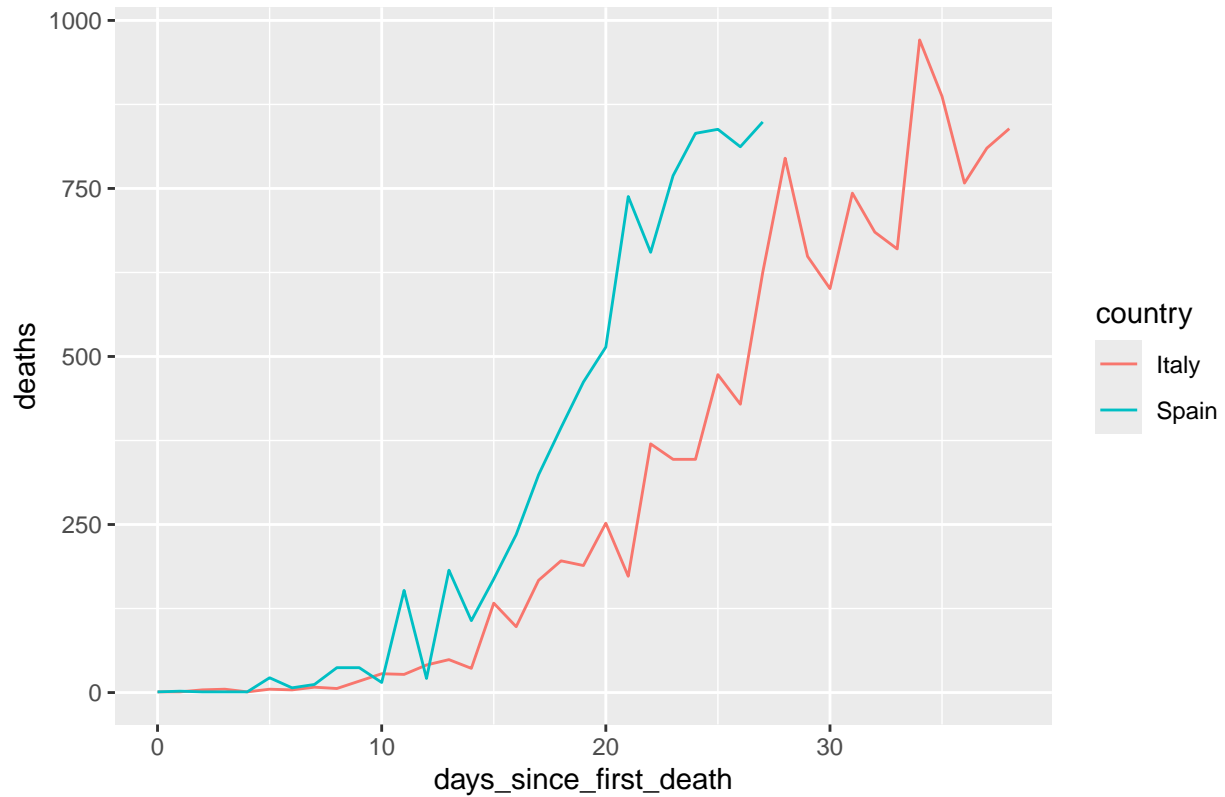
For a 40-y/o with a rate of 135, their residual comes out to be equal to about $135 - (147.81 - 0.62 * 40)$ or 11.98

For a 60-y/o with a rate of 112, their residual comes out to be equal to about $112 - (147.81 - 0.62 * 60)$ or 1.38

Since 11.98 is greater than 1.38, we determine that the 40-y/o likely has a healthier creatinine clearance rate for their age.

Problem 2 - Modeling Disease Growth

Covid Death Rate in Italy and Spain

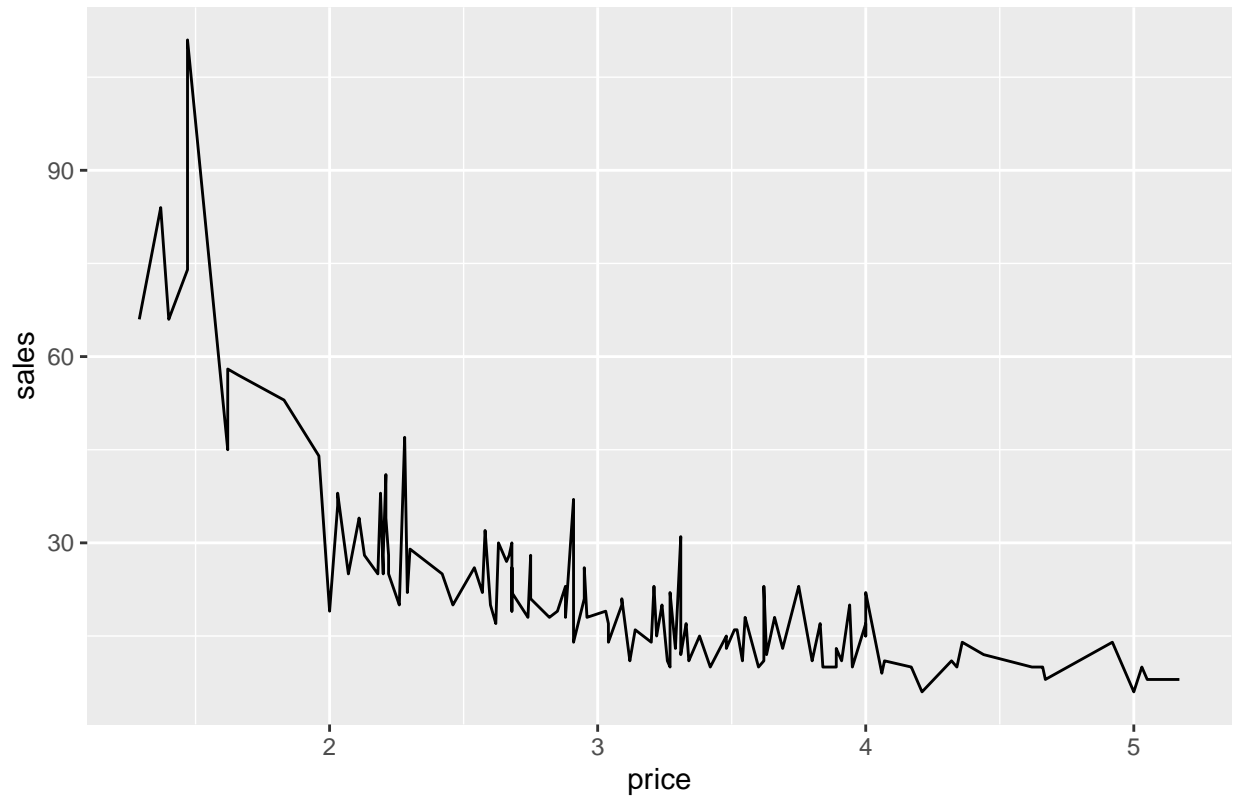


According to the bootstrapped samples, the true growth rate for Covid deaths per date is within the 95% confidence interval of $[0.160, 0.208]$ for Italy and $[0.236, 0.318]$ for Spain. This results in a confidence interval for doubling time of about $[3.4, 4.4]$ days for Italy and $[2.2, 3.0]$ days for Spain, with 95% confidence.

A line graph for death rate vs. days for each country is shown above.

Problem 3 - Price Elasticity of Demand

Sales vs Price of Milk



In order to linearize the provided data, which followed a power-law model, the natural logarithm was taken for both the x and y variables of price and sales, respectively. The slope of the linear regression line for this linearized model can be used to estimate the true price elasticity of demand, which according to the bootstrapped sample, lies within the 95% confidence interval of $[-1.775, -1.451]$. Our estimated value of the price elasticity of demand is -1.618 .