hw01

Samuel Swain

2023-01-20

Problem 1

$$Joint\ PDF \Rightarrow\ f(y;\mu,\sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} e^{\frac{-1}{2\sigma^2}\Sigma_{i=1}^{n}(y_i-\mu)^2}$$

$$Likelihood\ function \Rightarrow\ ln(1) - (\frac{n}{2}ln(2\pi) + nln(\sigma)) - \frac{1}{2\sigma^2}\Sigma_{i=1}^{n}(y_i-\mu)^2$$

$$\frac{\partial(f(y;\mu,\sigma))}{\partial\sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^3}\Sigma_{i=1}^{n}(y_i-\mu)^2 \Rightarrow\ Set = 0$$

$$\Rightarrow\ \frac{\Sigma_{i=1}^{n}(y_i-\mu)^2}{\sigma^3} = \frac{n}{\sigma}$$

$$\Rightarrow from\ class:\ \mu = \hat{y} \Rightarrow \sigma = \sqrt{\frac{\Sigma_{i=1}^{n}(y_i-\bar{y})^2}{n}}$$

Problem 2

2(a)

```
## Linear Model
##  ---
##   Function:
##    - y = 5.42 + 0.49 * x
##   Coefs:
##    - gamma_0: 0.1844778
##    - gamma_1: 0.09032769
```

2(b)

```
## NLM:
##  - Estimates: 28.13688 12.57428
```

```
## NLS:
##  - Estimates: 28.13702 12.57442
```

Problem 3

3(a)

```
## NLM:
##  ---
##   Fisher Info Matrix:
##   15.37455 -13.73842
##   -13.73842 13.92197
##   Standard Errors:
##    - gamma_0: 0.7418084
##    - gamma_1: 0.7795475
```

3(b)

```
## NLS:
##  ---
##   Covariance Matrix:
##   0.5299511 0.52028
##   0.52028 0.5822471
##   Standard Errors:
##    - gamma_0: 0.7279774
##    - gamma_1: 0.7630512
```

The results agree with part a as they are very close.

3(c)

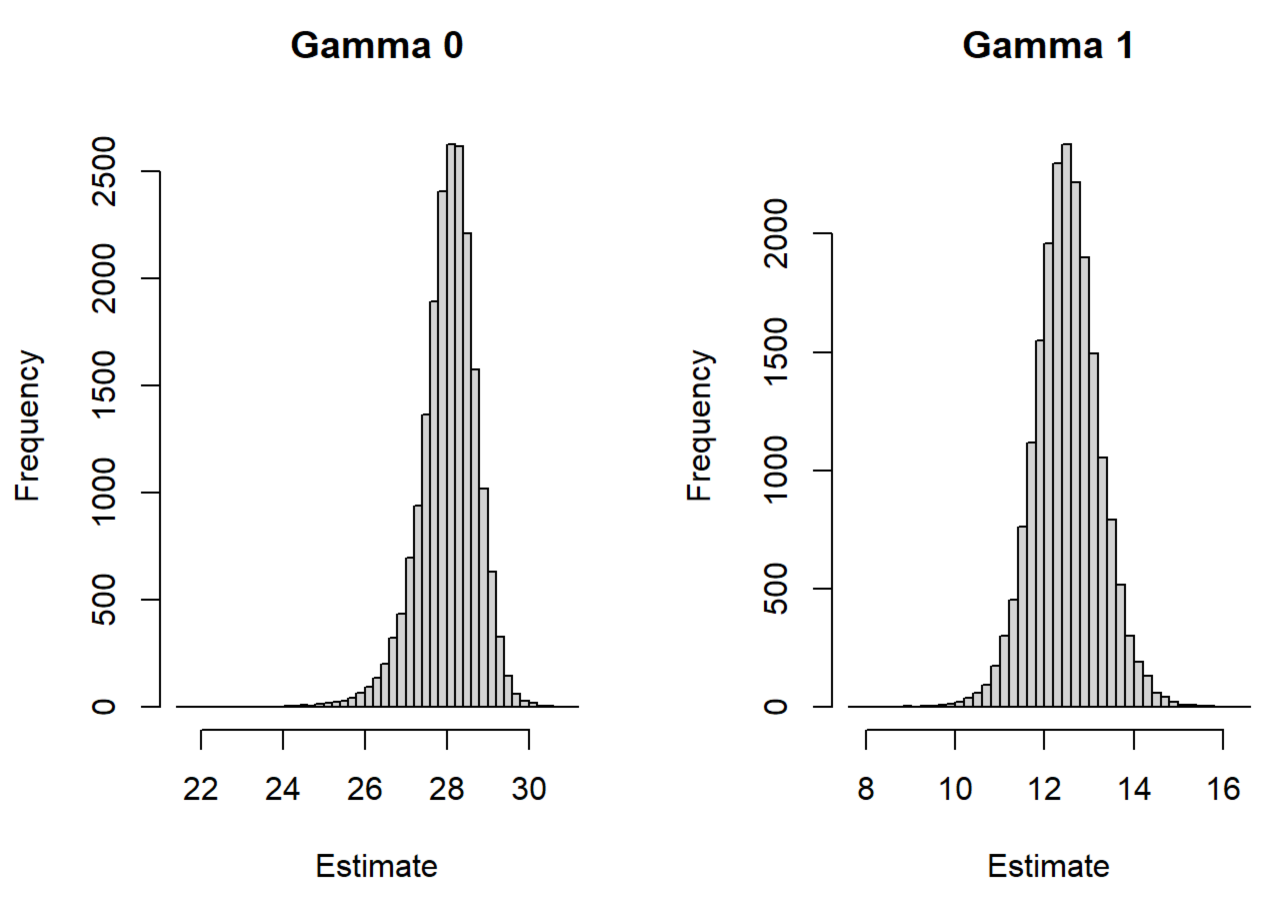
```
## NLM:
##  - gamma_0 CI: 26.68294 29.59083
##  - gamma_1 CI: 11.04637 14.10219
```

```
## NLS:
##  - gamma_0 CI: 26.71021 29.56383
##  - gamma_1 CI: 11.07887 14.06997
```

The two confidence intervals are almost the same.

Problem 4

4(a)




```
## Bootstrapped Standard Errors:
##  - gamma_0: 0.7103538
##  - gamma_1: 0.7359762
```

4(b)

```
## Crude 95% CI's:
##  - gamma_0: 26.64418 29.42877
##  - gamma_1: 11.05972 13.94475
```

4(c)

```
## Two-sided 95% CIs:
##  - gamma_0: 27.02695 29.82897
##  - gamma_1: 11.16272 14.09511
```

4(d)

They do agree. The histograms in part a suggest the gamma distributions are normally distributed. Because of this the "crude" CI calculated in part b will be roughly correct. This is why it is close to the CI calculated using the boot.ci function in part c.

Question 5

```
## Future Response:
##  - L: 18.13702
##  - U: 20.34532
```

```
## Predictable CI:
##  - L: 18.90606
##  - U: 19.7199
```

I think the prediction interval better represents an interval that would contain future events. The prediction interval contains the confidence interval. Because of this it will always be larger to account for the extra noise. Because the interval is larger, it will better represent the interval we can be 95% sure will contain the prediction.

Question 6

```
## AIC:
##  - NLS Model: 1.739982
##  - Square Root Model: 2.947259
```

AIC suggests the old NSL model is better.

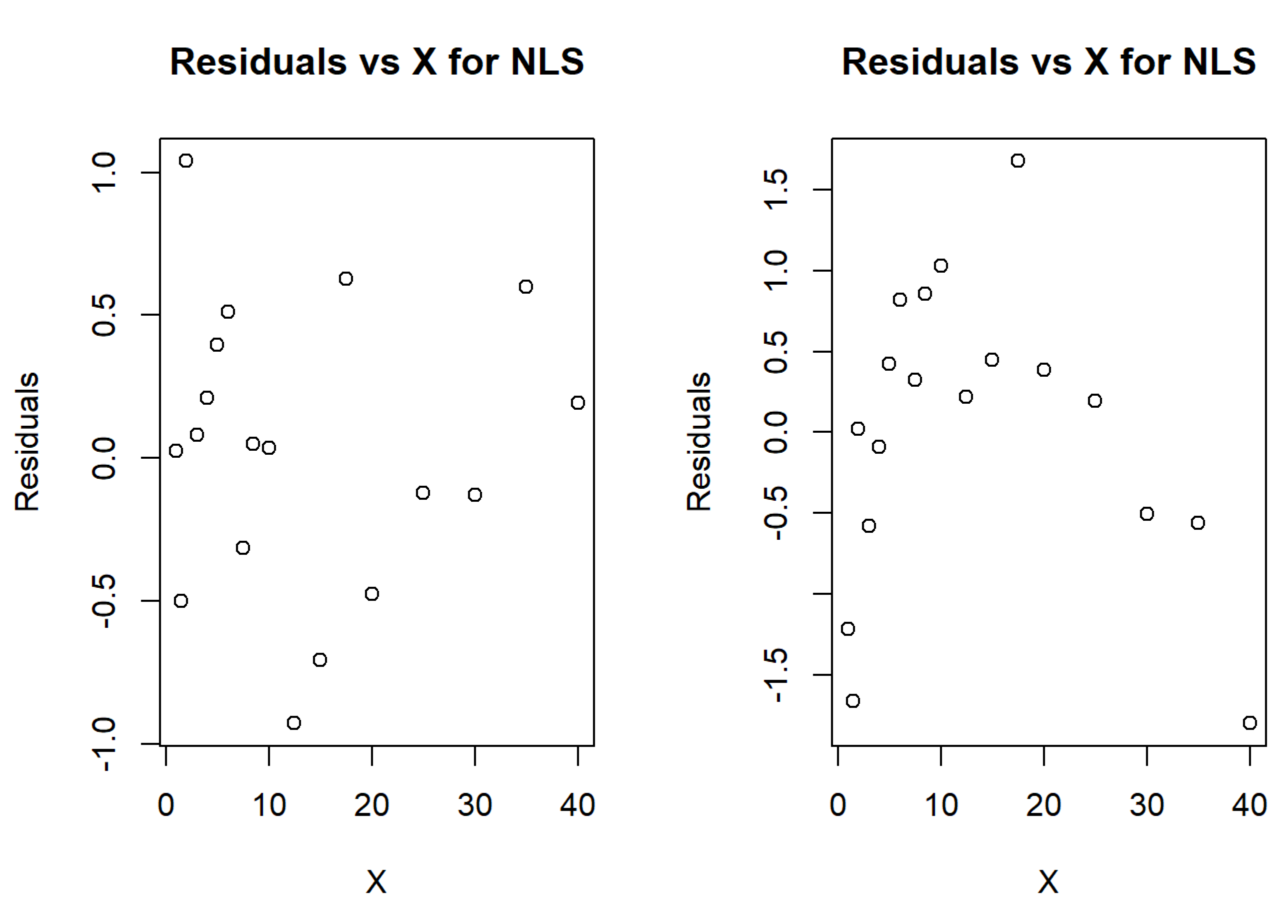
Question 7

```
## NLS Model
##  - Average MSE: 0.3299946
##  - Standard Deviation of MSE: 0.05794992
```

```
## Square Root Model
##  - Average MSE: 1.19611
##  - Standard Deviation of MSE: 0.3717434
```

Cross validation suggests the NLS model performs better than the square root model.

Question 8



The residual plots for the NLS model seems more appropriate than the square root model. This is because the residuals for the NLS model show no pattern whereas the square root residual plot looks like a quadratic function. This agrees with out conclusions from questions 6 and 7 above.