ProblemSet3

Aaron Graybill

320/2021

Problem 2.6

a.

$$y_i - \hat{y}_i = y_i - \hat{\beta}_0 + \hat{\beta}_1 x_i = y_i - \overline{y} - \hat{\beta}_1 \overline{x} + \hat{\beta}_1 x_i = (y_i - \overline{y}) - \hat{\beta}_1 (x_i - \overline{x})$$

b.

$$\hat{y}_i - \overline{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i - \overline{y} = \overline{y} - \hat{\beta}_1 \overline{x} + \hat{\beta}_1 x_i - \overline{y} = \hat{\beta}_1 (x_i - \overline{x})$$

c.

$$\begin{aligned}
&\sum_{i=1}^{n} (\hat{y}_i - \overline{y})(y_i - \hat{y}_i) \\
&\sum_{i=1}^{n} \Big((y_i - \overline{y}) - \hat{\beta}_1 (x_i - \overline{x})\Big)\Big(\hat{\beta}_1 (x_i - \overline{x})\Big) \\
&\sum_{i=1}^{n} (y_i - \overline{y})\hat{\beta}_1 (x_i - \overline{x}) - \hat{\beta}_1^2 (x_i - \overline{x})^2 \\
&\hat{\beta}_1 SXY - \hat{\beta}_1^2 SXX \\
&\frac{SXY}{SXX} SXY - \left(\frac{SXY}{SXX}\right)^2 SXX \\
&\frac{SXY^2 - SXY^2}{SXX} = 0
\end{aligned}$$

Simultaneous Question 1.

a.

```
## here() starts at /Users/aarongraybill/Documents/Haverford Stuff/Math/Math286
```

```
##
## Call:
## lm(formula = AS95 ~ AS94, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1391 -2.0531 -0.5961  1.1873 27.0122
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.40333    0.75339   7.172 8.27e-12 ***
```

```
## AS94         0.93471    0.01306  71.552  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.851 on 251 degrees of freedom
## Multiple R-squared:  0.9533, Adjusted R-squared:  0.9531
## F-statistic:  5120 on 1 and 251 DF,  p-value: < 2.2e-16
```

b.

Using the Bonferroni method, I find two individual intervals at confidence .95 and they together for a joint confidence interval. We have:

```
##                 2.5 %    97.5 %
## (Intercept) 3.9195711 6.8870968
## AS94        0.9089871 0.9604428
```

So we are (more than) 90% sure that neither coefficient is zero.

c.

We need 95% independent levels for a joint prediction interval when there are two predictions, we have:

```
##        fit      lwr      upr
## 1 52.13908 44.53918 59.73899
## 2 54.94323 47.34413 62.54232
```

d.

Since we have 3 intervals, we need a confidence level of $1 - \frac{.1}{3} = 96.66667\%$

```
##        fit      lwr      upr
## 1 50.26965 42.01126 58.52804
## 2 54.94323 46.68678 63.19967
## 3 58.68209 50.42549 66.93868
```

Simultaneous Problem 2.

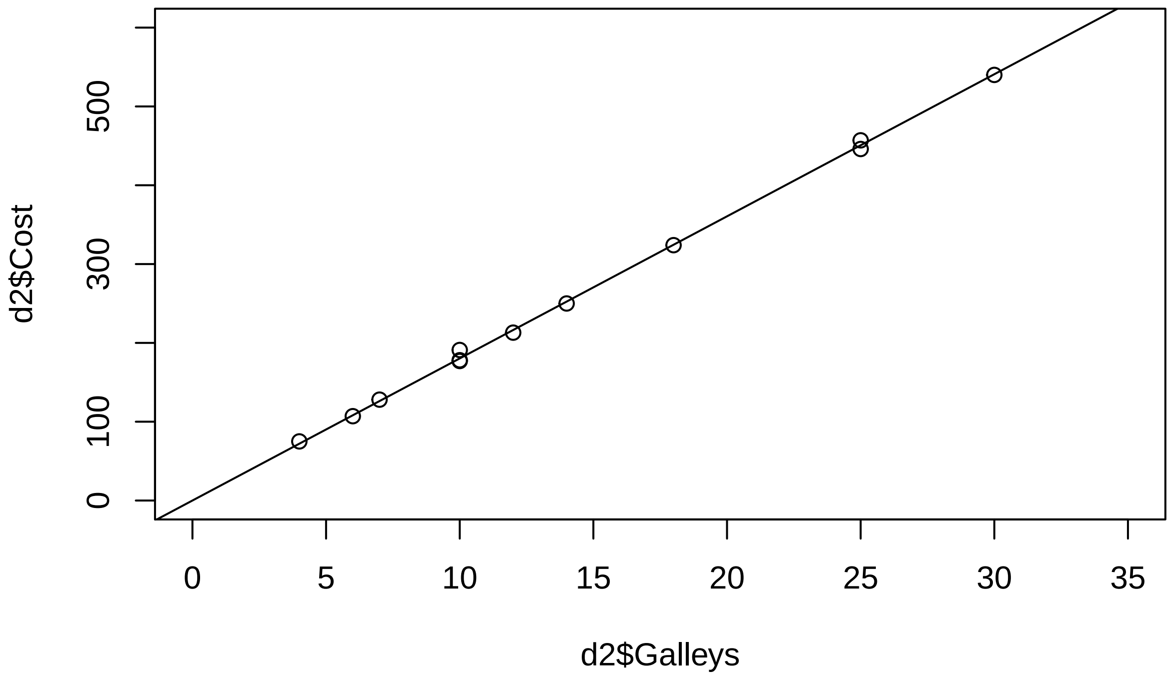
a.

Here's a regression through the origin:

```
##
## Call:
## lm(formula = Cost ~ Galleys - 1, data = d2)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.708 -2.618 -1.010  2.073 10.717
##
## Coefficients:
##         Estimate Std. Error t value Pr(>|t|)
## Galleys 18.02830    0.07948   226.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.507 on 11 degrees of freedom
## Multiple R-squared:  0.9998, Adjusted R-squared:  0.9998
```

```
## F-statistic: 5.145e+04 on 1 and 11 DF,  p-value: < 2.2e-16
```

b.



The relationship looks appropriate, so if we have a theoretical reason for regressing through the origin, this seems like an appropriate model.

c.

Management is essentially asserting that $\beta = 17.5$ and we can run a hypothesis to see whether or not the observed data suggests a value significantly different than that. We have:

$$H_0 : \beta = 17.5 H_a : \beta \neq 17.5$$

In the last problem set we showed that for a e_i with variance σ^2, a through-the-origin has $\hat{\beta} \sim \left(\beta, \frac{\sigma^2}{\sum x_i^2}\right)$. So we need to normalize this to t-distribution it. We know that:

$$\frac{\hat{\beta} - \beta}{\frac{S}{\sqrt{\sum x_i^2}}} = \frac{\hat{\beta} - \beta}{se(\hat{\beta})} \sim t_{n-2}$$

And doing that I now realize that this is just the standard hypothesis testing. Okay now we have all that we need to run hypothesis testing on β I compute the test statistic:

$$\frac{18.02830 - 17.5}{0.07948} = 6.646955$$

```
## [1] 4.703294e-05
```

Computing the p-value gives: 9.4065879×10^{-5} which is well below the $\alpha = .02$ level, which indicates that we reject the null hypothesis that $\beta = 17.5$ and accept that it is not equal to $\beta = 17.5$. The managers assumption is faulty and likely should be revised.

d.

```
##       fit      lwr     upr
## 1 180.283 167.8441 192.722
```

We estimate the cost of the Job wopuld be between 167 and 192 money units.