Final, MA 331-B

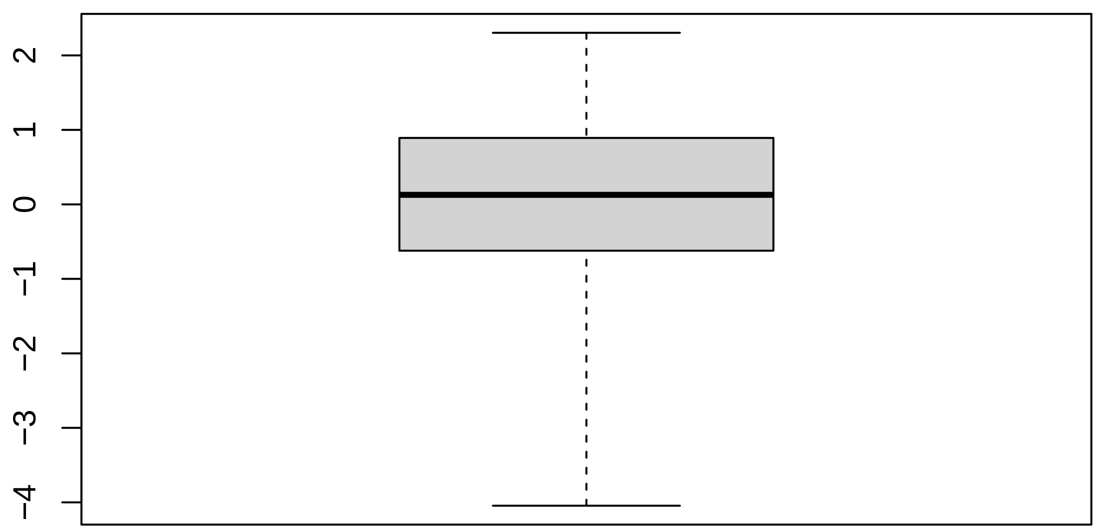
Daniel Detore

May 15, 2025

I pledge my honor that I have abided by the Stevens Honor System.

Problem 1

```
e <- c(-4.045, -0.622, 0.128, 0.892, 2.303)
boxplot(e, range = 15)
```



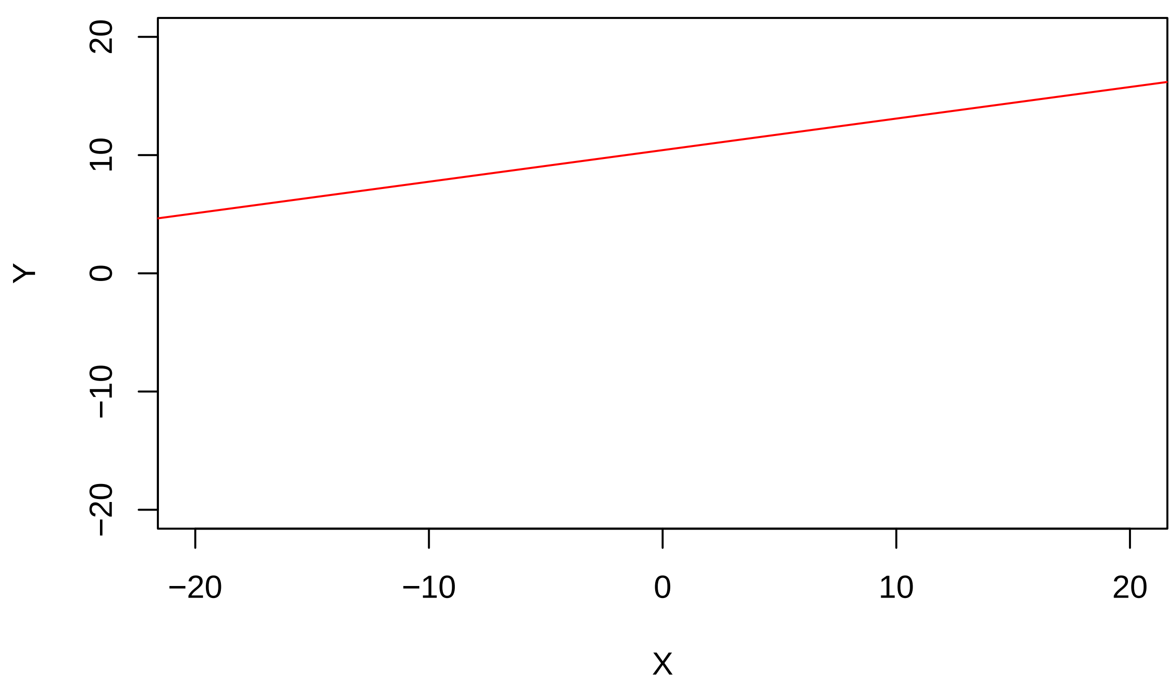
The readout tells us that we use distribution $F_{d_1,d_2} = F_{1,58}$. We find $n_{residuals} = n_{observations} = d_2 + 2 = 60$.

This model seems to be a middling fit to the data. The residuals are somewhat asymmetric about 0, but the R^2 is quite high.

Problem 2

The LSE of the two regression parameters are $\beta_1 = 0.2670$ and $\beta_0 = 10.4183$ which gives the regression equation $Y = 0.2670X + 10.4183$.

```
plot(x=NA, type="n", xlim=c(-20, 20), ylim=c(-20, 20),
    xlab="X", ylab="Y")
abline(a = 10.4183, b =0.2670, col = "red")
```



Problem 3

The coefficient of determination $R^2 = 0.854$.

We know $RSE = S = \sqrt{\frac{SSE}{n-2=df}} \implies SSE = RSE^2 \times df =$

```
SSE <- (1.18^2) * 58; SSE
```

```
## [1] 80.7592
```

We also know $SST = SSM + SSE$ and $R^2 = \frac{SSM}{SST} \implies SSM = R^2 \times SST$ thus $SST = R^2 \times SST + SSE \implies SST - R^2 \times SST = SSE \implies SST(1 - R^2) = SSE \implies SST = \frac{SSE}{1-R^2} =$

```
SST <- SSE/(1-0.854); SST
```

```
## [1] 553.1452
```

Problem 4

We have $H_0 : \beta_1 = 0$ and $H_a : \beta_0 \neq 0$ The readout gives $\hat{\beta}_1 = 0.2670$ and $SE_{\hat{\beta}_1} = 0.4251$. This makes our observed testing statistic $t = \frac{\hat{\beta}_1}{SE_{\hat{\beta}_1}} =$

```
t <- 0.2670/0.4251; t
```

```
## [1] 0.6280875
```

```
# which gives us p-value:
2 * (1 - pt(t, 58))
```

```
## [1] 0.5324121
```

This p-value is too high for any reasonable confidence level. We must accept H_0.

Problem 5

The CI of the intercept parameter is $\hat{\beta}_0 \pm t_{1-\alpha/2}(n-2) \times SE_{\hat{\beta}_0}$.

```
b0 <- 10.4183
SEb0 <- 0.4251
cl <- 0.95
sprintf('Our confidence interval is: [%f, %f]',
    b0 - pt(1 - (cl / 2), 58) * SEb0,
    b0 + pt(1 - (cl / 2), 58) * SEb0)
```

```
## [1] "Our confidence interval is: [10.121067, 10.715533]"
```

Problem 6

The estimate of the variance of random error is $RSE^2 =$

```
1.18^2
```

```
## [1] 1.3924
```

Problem 7

We have $H_0 : \beta_1 = 0$ and $H_a : \beta_1 \neq 0$. Our testing statistic $f = \frac{SSM}{SSE/(n-2)} = \frac{SST-SSE}{SSE/(df)} =$

```
f <- (SST-SSE) / (SSE/58); f
```

```
## [1] 339.2603
```

```
# which gives us p-value:
1 - pf(f, 1, 58)
```

```
## [1] 0
```

$> 1 - \alpha = 0.01$. Thus we reject H_0 and assume H_a.

Problem 8

We can estimate Y^* as $y^* = \beta_0 + \beta_1 x^*$ (we must ignore ε^* because its value is experimental). When $x^* = 49$, $y^* =$

```r
10.4183 + 49 * 0.2670
```

```
## [1] 23.5013
```

Problem 9

We get the prediction interval by $\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{1-\alpha/2}(n-2) \times SE_{\bar{Y}^*}$. Without the data, we cannot calculate $SE_{\bar{Y}^*}$. Thus our best prediction is still 23.5013.

Problem 10

$$Y = \tilde{X}\alpha + \varepsilon \implies \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 - \bar{x} \\ 1 & x_2 - \bar{x} \\ \vdots & \vdots \\ 1 & x_n - \bar{x} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Problem 11

We know

$$(\tilde{X}'\tilde{X}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ (x_1 - \bar{x}) & (x_2 - \bar{x}) & \dots & (x_n - \bar{x}) \end{bmatrix} * \begin{bmatrix} 1 & x_1 - \bar{x} \\ 1 & x_2 - \bar{x} \\ \vdots & \vdots \\ 1 & x_n - \bar{x} \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n (x_i - \bar{x}) \\ \sum_{i=1}^n (x_i - \bar{x}) & \sum_{i=1}^n (x_i - \bar{x})^2 \end{bmatrix}$$

If we know that $\sum_{i=1}^n (x_i - \bar{x}) = 0$, this becomes

$$\begin{bmatrix} n & 0 \\ 0 & \sum_{i=1}^n (x_i - \bar{x})^2 \end{bmatrix}$$

and if we know the property of the inverse of a diagonally dominant matrix, we can find $(\tilde{X}'\tilde{X})^{-1} =$

$$\begin{bmatrix} n & 0 \\ 0 & \sum_{i=1}^n (x_i - \bar{x})^2 \end{bmatrix}.$$

We can find

$$\tilde{X}'Y = \begin{bmatrix} 1 & 1 & \dots & 1 \\ (x_1 - \bar{x}) & (x_2 - \bar{x}) & \dots & (x_n - \bar{x}) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i (x_i - \bar{x}) \end{bmatrix}.$$

Using these two equations, we can find $\hat{\alpha} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'Y =$

$$\begin{bmatrix} n & 0 \\ 0 & \sum_{i=1}^n (x_i - \bar{x})^2 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i (x_i - \bar{x}) \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^n y_i}{n} \\ \frac{\sum_{i=1}^n y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{bmatrix}$$