MATH 423/533 – SIMPLE LINEAR REGRESSION: INFERENCE AND TESTING

For the simple linear regression model

$$\mathbb{E}[Y_i|x_{i1}] = \beta_0 + \beta_1 x_{i1} \qquad i = 1, 2, \ldots, n$$

- x_{i1} is Age, the single continuous predictor;
- Y_i is the Shear Strength outcome random variable;
- y_i is the observed version of Y_i.

A summary of inference and test results for the model is obtained in `R` using the `summary` function:

Simple Linear Regression: Inference

```
1  > x<-RocketProp$Age
2  > y<-RocketProp$Strength
3  > fit.RP<-lm(y ~ x)
4  > #Summary of inference
5  > summary(fit.RP)
6
7  Call:
8  lm(formula = y ~ x)
9
10 Residuals:
11     Min      1Q  Median      3Q     Max
12 -215.98  -50.68   28.74   66.61  106.76
13
14 Coefficients:
15             Estimate Std. Error t value Pr(>|t|)
16 (Intercept) 2627.822     44.184   59.48  < 2e-16 ***
17 x            -37.154      2.889  -12.86 1.64e-10 ***
18 ---
19 Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1     1
20
21 Residual standard error: 96.11 on 18 degrees of freedom
22 Multiple R-squared:  0.9018,    Adjusted R-squared:  0.8964
23 F-statistic: 165.4 on 1 and 18 DF,  p-value: 1.643e-10
```

- line 21: the quantity `Residual standard error`, taking the value 96.11, is $\widehat{\sigma}$, where

$$\widehat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$$

- lines 14-19: contains the regression coefficient estimates (`Estimate` column) and estimated standard errors(e.s.e, in the `Std. Error` column). The `t value` column contains the test statistics

$$\frac{\widehat{\beta}_j}{\text{e.s.e.}(\widehat{\beta}_j)} \qquad j = 0, 1$$

in tests of the null hypothesis

$$\mathrm{H}_0 \ : \ \beta_j = 0.$$

The column `Pr(>|t|)` contains the p-value in each test: here both p-values are below significance level $\alpha = 0.05$, so both null hypotheses are rejected. The asterisks $\star$ indicate the strength of rejection (line 19).