Simple Linear Regression

Simple Linear Regression

Charlene Garridos

2023-06-21

Contents

1 The Rocket Propellant Data

A rocket motor is manufactured by bonding an igniter propellant and a sustainer propellant together inside a metal housing. The shear strength of the bond between the two types of propellant is an important quality characteristic. It is suspected that shear strength is related to the age in weeks of the batch of sustainer propellant. Twenty observations on shear strength and the age of the corresponding batch of propellant have been collected

```
# Specify the file path relative to the working directory
file_path <- "/Users/User/Downloads/rocket_propellant_data.csv"

# Read the CSV file
rocket_propellant_data <- read.csv(file_path)

# View the data
print(rocket_propellant_data)
```

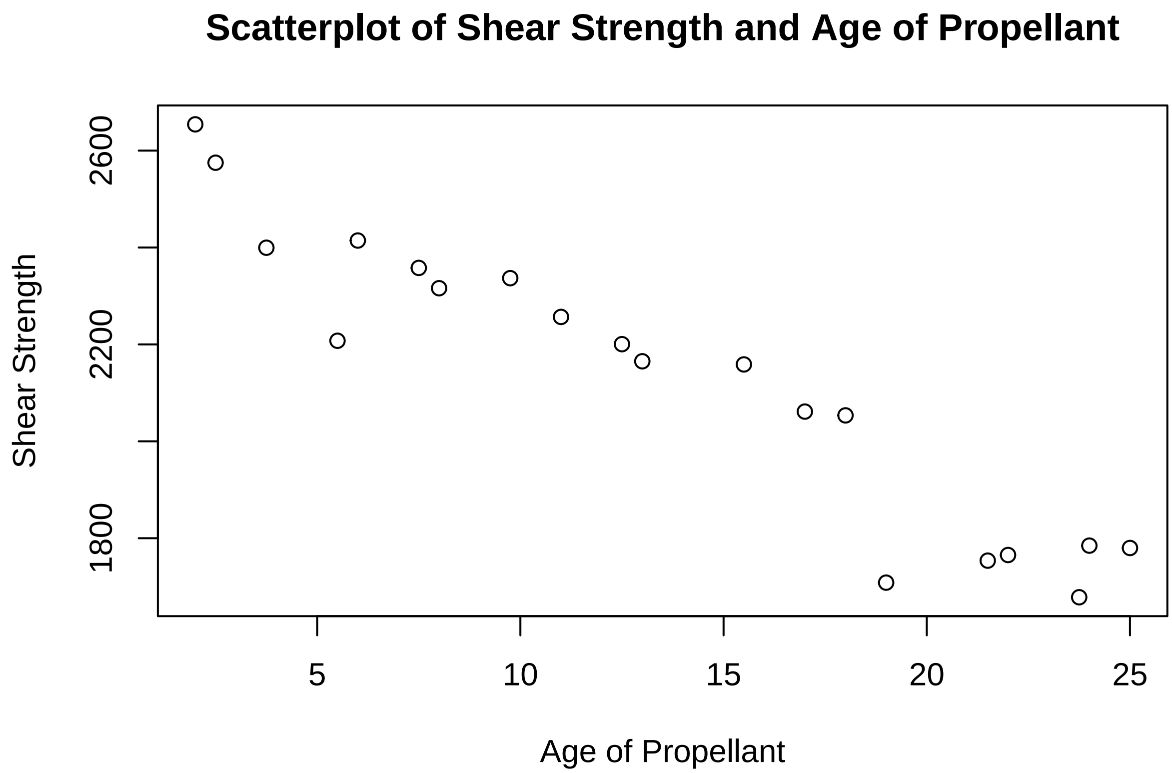
```
##    observation y_shear_strength x_propellant_age
## 1            1          2158.70            15.50
## 2            2          1678.15            23.75
```

```
## 3             3          2316.00              8.00
## 4             4          2061.30             17.00
## 5             5          2207.50              5.50
## 6             6          1708.30             19.00
## 7             7          1784.70             24.00
## 8             8          2575.00              2.50
## 9             9          2357.90              7.50
## 10           10          2256.70             11.00
## 11           11          2165.20             13.00
## 12           12          2399.55              3.75
## 13           13          1779.80             25.00
## 14           14          2336.75              9.75
## 15           15          1765.30             22.00
## 16           16          2053.50             18.00
## 17           17          2414.40              6.00
## 18           18          2200.50             12.50
## 19           19          2654.20              2.00
## 20           20          1753.70             21.50
```

1.1 Scatter Diagram

a. Create a scatter diagram for the data.

```r
# scatter plot of propellant_age and shear_strength
plot(x = rocket_propellant_data$x_propellant_age, y = rocket_propellant_data$y_shear_strength,
     xlab = "Age of Propellant", ylab = "Shear Strength",
     main = "Scatterplot of Shear Strength and Age of Propellant")
```



2 Least-Squares Estimation of the Parameters

Use the lm() function to calculate the linear model based on the data set.

```
# calculate model
model <- lm(data = rocket_propellant_data,
formula = y_shear_strength ~ x_propellant_age)
```

The model object is a list of a number of different pieces of information, which can be seen by looking at the names of the objects in the list.

```
# view the names of the objects in the model
names(model)
```

```
##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "xlevels"       "call"          "terms"         "model"
```

```
model$coefficients
```

```
##      (Intercept) x_propellant_age
##       2627.82236        -37.15359
```

2.1 The Least-Squares Fit

The summary() function is a useful way to gather critical information in your model.

```
model_summary <- summary(model)
model_summary$sigma
```

```
## [1] 96.10609
```

2.2 The Estimate

The estimate of σ^2 is

```
sigma_squared<- (model_summary$sigma)^2
sigma_squared
```

```
## [1] 9236.381
```

2.3 Alternative Method of Calculating Least-Squares Estimation of the Parameters

Estimate model parameters Sxx and Sxy

```
S_xx <- sum((rocket_propellant_data$x_propellant_age)^2) -
  ((sum(rocket_propellant_data$x_propellant_age)^2)/20)

S_xy <- sum((rocket_propellant_data$x_propellant_age)*
              (rocket_propellant_data$y_shear_strength))-
  ((sum(rocket_propellant_data$x_propellant_age))*
     (sum(rocket_propellant_data$y_shear_strength))/20)

S_xx
```

```
## [1] 1106.559
```

```
S_xy
```

```
## [1] -41112.65
```

```
B_1 <- (S_xy/S_xx)
B_1
```

```
## [1] -37.15359
```

```
B_0 <- mean(rocket_propellant_data$y_shear_strength)-
  B_1*mean(rocket_propellant_data$x_propellant_age)

B_0
```

```
## [1] 2627.822
```

```r
y_sqrd <- sum((rocket_propellant_data$y_shear_strength)^2)
y_sqrd
```

```
## [1] 92547433
```

```r
y_sum <- sum(rocket_propellant_data$y_shear_strength)
y_sum
```

```
## [1] 42627.15
```

```r
SS_t <- (y_sqrd-((y_sum)^2/20))
SS_t
```

```
## [1] 1693738
```

Residual sum of squares

```r
SS_Res <- SS_t - ((B_1)*(S_xy))
SS_Res
```

```
## [1] 166254.9
```

Therefore, the estimate of σ^2 is

```r
n=20
sigma_sqrd <- SS_Res/(n-2)
sigma_sqrd
```

```
## [1] 9236.381
```

3 Hypothesis Testing on the Slope and Intercept

```r
model_summary
```

```
##
## Call:
## lm(formula = y_shear_strength ~ x_propellant_age, data = rocket_propellant_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -215.98  -50.68   28.74   66.61  106.76
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2627.822     44.184   59.48  < 2e-16 ***
## x_propellant_age  -37.154      2.889  -12.86 1.64e-10 ***
## ---
```

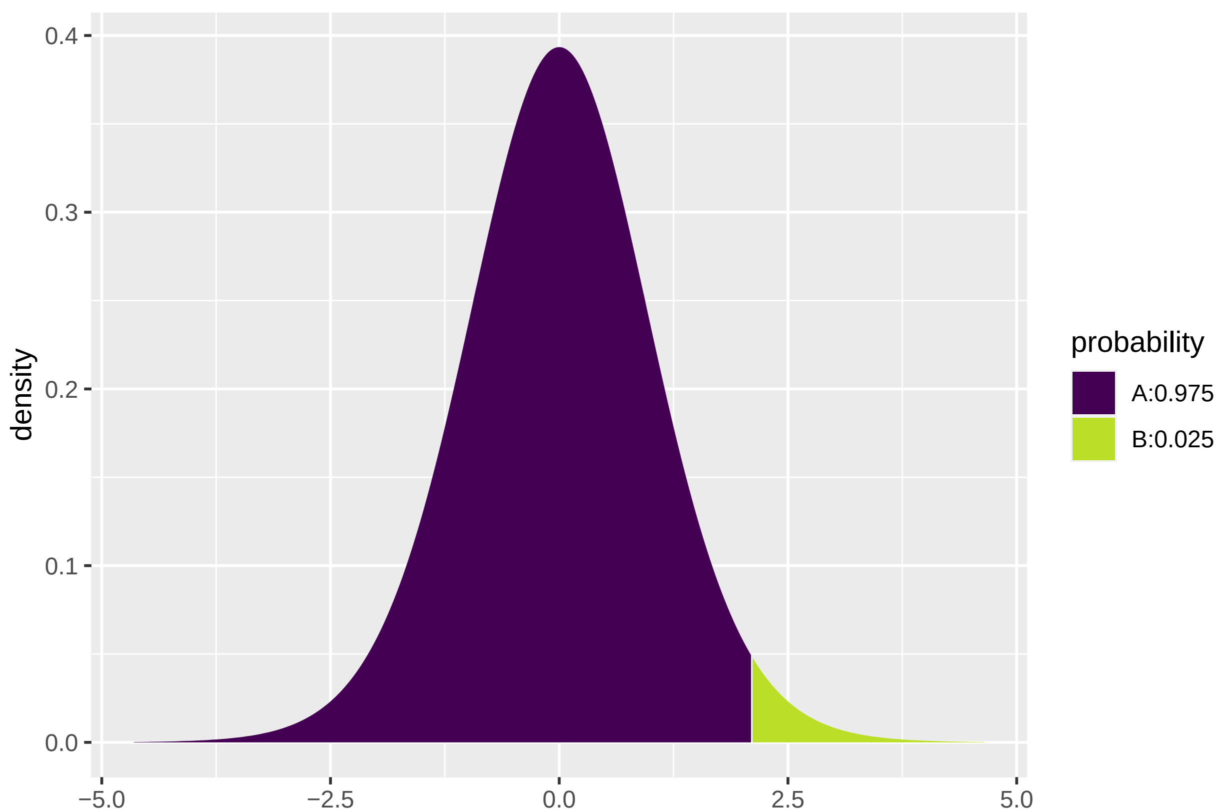
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 96.11 on 18 degrees of freedom
## Multiple R-squared:  0.9018, Adjusted R-squared:  0.8964
## F-statistic: 165.4 on 1 and 18 DF,  p-value: 1.643e-10
```

```
model_summary$coefficients["x_propellant_age",]
```

```
##      Estimate    Std. Error       t value      Pr(>|t|)
## -3.715359e+01  2.889107e+00 -1.285989e+01  1.643344e-10
```

```
mosaic::xqt(0.975, 18)
```

```
## Registered S3 method overwritten by 'mosaic':
##   method                           from
##   fortify.SpatialPolygonsDataFrame ggplot2
```




```
## [1] 2.100922
```

3.1 Test for Significance of Regression in the Rocket Propellant Regression Model.

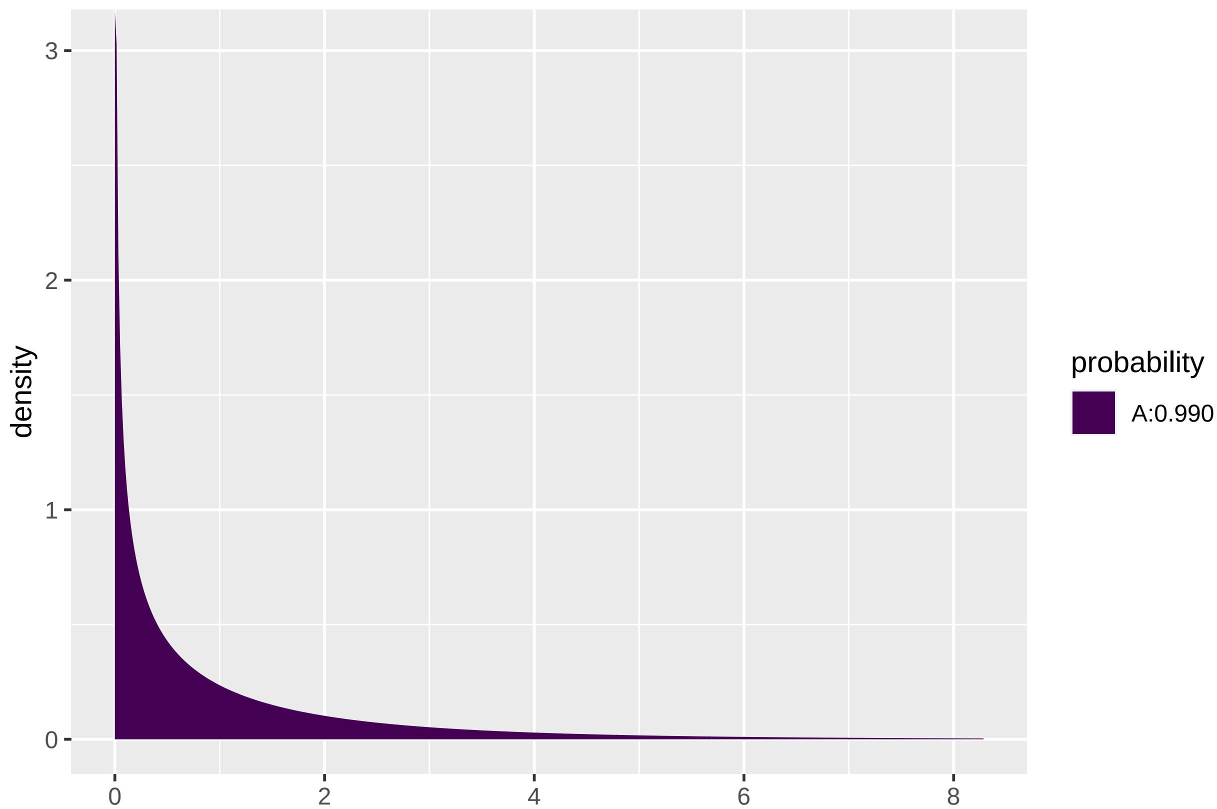
Based on the statistical analysis conducted, the t-statistic for the coefficient of x_propellant_age, with the 97.5% confidence interval, is -12.86 ($p < 0.001$), indicating a significant relationship between propellant age and shear strength. This negative t-value suggests that as the age of the propellant increases, the shear strength tends to decrease. Therefore, we can infer that propellant age has a substantial impact on the shear strength of the rocket propellant.

Thus, the statistically significant t-value provides strong evidence to reject $H_0 : \beta_1 = 0$ and support the conclusion that propellant age is an influential factor in determining the shear strength, highlighting the importance of considering the age of the propellant in optimizing the performance and reliability of rocket systems.

```
model_summary$fstatistic
```

```
##    value    numdf    dendf
## 165.3768   1.0000  18.0000
```

```
mosaic::xqf(0.99,1,18)
```




```
## [1] 8.28542
```

3.2 An analysis-of-variance approach to test significance of regression.

At 99% confidence interval, the F-statistic of 165.4 ($p < 0.001$) obtained from the regression analysis indicates that the overall model, including the intercept and the predictor variable x_propellant_age, is statistically significant. It suggests the inclusion of propellant age as a predictor in the model significantly improves the ability to predict the shear strength compared to a model without this variable. Along with the low p-value, it indicates that there is a low probability of obtaining such a strong relationship between the predictors and the response variable by chance alone. This strengthens our confidence in the conclusion that propellant age has a substantial impact on the shear strength of the rocket propellant. Therefore, the F-statistic provides strong evidence to reject $H_0 : \beta_1 = 0$.