Chapter 2 LinReg-1 (pg 4)

A screenshot of a computer

Description automatically generated

This output is from a linear regression model fitted using the `lm()` function in R. Let's break down the key components:

### Model Summary:

- \*\*Call:\*\* Specifies the formula used for the linear regression (`y ~ x`), indicating that `y` is the dependent variable and `x` is the independent variable.

- \*\*Residuals:\*\* Provides summary statistics for the residuals (the differences between observed and predicted values):

- \*\*Min:\*\* The minimum residual value (-2.45122).

- \*\*1Q:\*\* The first quartile of the residuals (-0.68319).

- \*\*Median:\*\* The median of the residuals (0.02913).

- \*\*3Q:\*\* The third quartile of the residuals (0.69861).

- \*\*Max:\*\* The maximum residual value (2.88937).

- \*\*Coefficients:\*\* Displays the estimated coefficients for the intercept and the independent variable `x`:

- \*\*Estimate:\*\* The estimated coefficients.

- \*\*Std. Error:\*\* The standard errors associated with the estimates.

- \*\*t value:\*\* The t-statistic, calculated as the estimate divided by its standard error.

- \*\*Pr(>|t|):\*\* The p-value associated with the t-statistic, testing if the coefficients are significantly different from zero.

- \*\*Signif. codes:\*\* Indicates the level of significance.

### Model Performance:

- \*\*Residual standard error:\*\* Represents the standard deviation of the residuals, providing a measure of the model's goodness-of-fit. Here, it's approximately 0.9912.

- \*\*Multiple R-squared:\*\* Measures the proportion of variance in the dependent variable (`y`) explained by the independent variable (`x`). In this case, it's 43.11%.

- \*\*Adjusted R-squared:\*\* Similar to R-squared but adjusted for the number of predictors. It's 42.82%.

- \*\*F-statistic:\*\* Evaluates the overall significance of the model. Here, the F-statistic value is 150 with associated degrees of freedom, indicating that the model is statistically significant in predicting `y` using `x`.

- \*\*p-value:\*\* The p-value associated with the F-statistic, indicating that the overall model is highly significant (p < 2.2e-16).

### Interpretation:

- The intercept is estimated to be approximately -0.40114, but it's not significantly different from zero (p = 0.256).

- The coefficient for `x` is estimated to be approximately 0.43330, indicating that for every one unit increase in `x`, `y` is predicted to increase by 0.43330 units. It's highly significant (p < 2e-16).

- The model explains around 43.11% of the variance in `y`.

- Overall, the model seems to have a good fit, with `x` being a significant predictor of `y`.

A paper with text on it

Description automatically generated