Inferences About Slope

DASC 512

AF INSTITUTE OF TECHNOLOGY

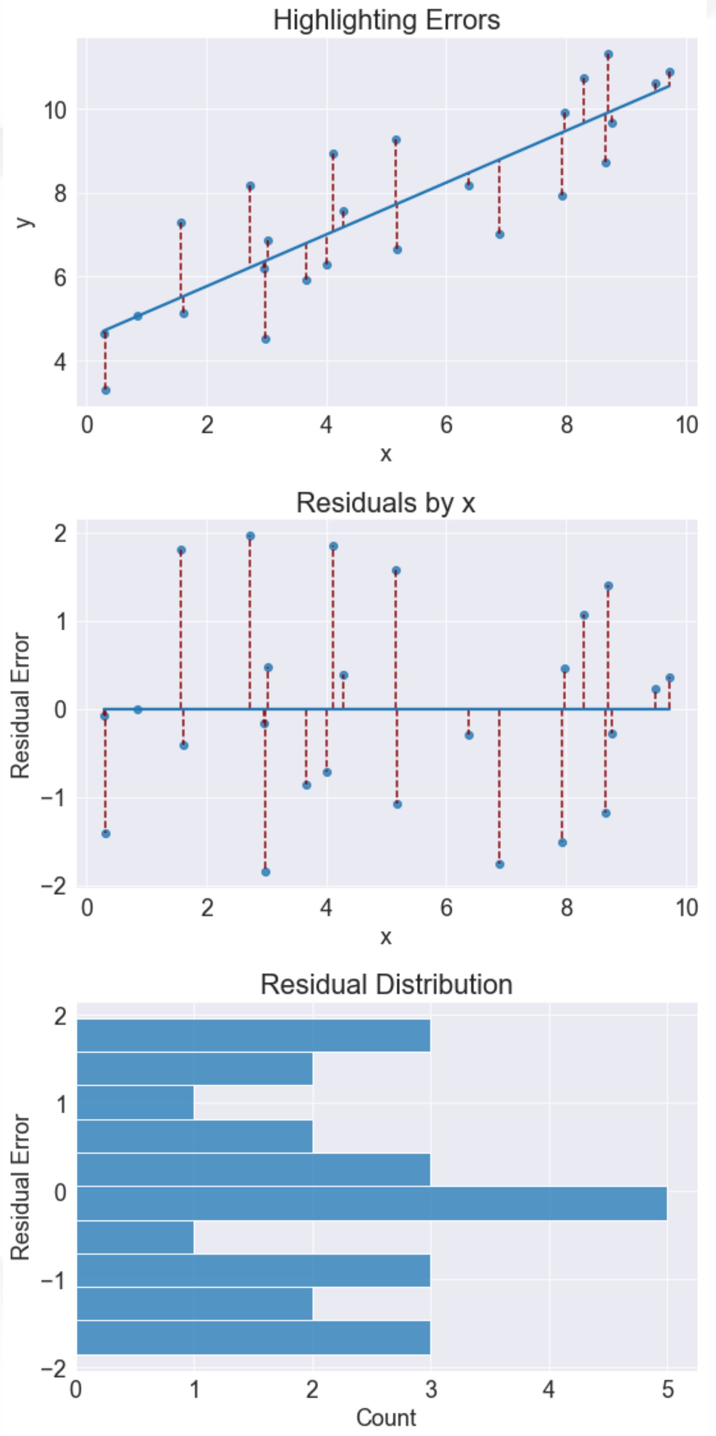
Inferences about Slope

- "All models are wrong, but some are useful." – George Box

- As with the ANOVA, we may be interested in whether our independent variable's effect on the response variable is real
  - If it is real, then the model provides useful information
- If the slope $\beta_1 = 0$, then there is no effect – $y = \beta_0 + \epsilon = \mu + \epsilon$
  - The OLS model provides no information
  - The variables are not related (aka not correlated) according to this model

Assumptions

1. The mean value of ϵ is zero
2. The variance of ϵ (some σ^2), is constant for all values of x
3. ϵ is normally distributed
4. Each ϵ_i is iid (independent and identically distributed)



Hypotheses

Null hypothesis: The slope is zero.

- $\beta_1 = 0$

Alternative hypothesis: The slope is not zero in some specified way.

- $\beta_1 < 0,\ \beta_1 > 0,\ \beta_1 \neq 0$

Test Statistic

If we assume that the residuals are distributed normally, we have the standard error of the sampling distribution of $\hat{\beta}_1$ is

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{SS_{xx}}}$$

This gives us a test statistic for a t-distribution with $n - 2$ degrees of freedom

$$t = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}} = \frac{\hat{\beta}_1}{s/\sqrt{SS_{xx}}}$$

Conclusion

If $H_a: \beta_1 < 0$,

$$t < t_{PPF}(\alpha, \nu = n - 2)$$

If $H_a: \beta_1 > 0$,

$$t > t_{ISF}(\alpha, \nu = n - 2)$$

If $H_a: \beta_1 \neq 0$,

$$|t| > t_{ISF}\left(\frac{\alpha}{2}, \nu = n - 2\right)$$

Also, a $100(1 - \alpha)\%$ confidence interval for β_1:

$$\hat{\beta}_1 \pm t_{ISF}\left(\frac{\alpha}{2}, \nu = n - 2\right) s_{\hat{\beta}_1}$$

P-value

If $H_a\!: \beta_1 < 0$,

$$p = t_{CDF}(t, \nu = n - 2)$$

If $H_a\!: \beta_1 > 0$,

$$p = t_{SF}(t, \nu = n - 2)$$

If $H_a\!: \beta_1 \neq 0$,

$$p = 2 \times t_{SF}(|t|, \nu = n - 2)$$

Example: Rocket Propellant

After removing the outliers due to faulty instrumentation, we had

$$\hat{y} = 2659 - 37.7x$$

Let's put a 95% confidence interval on the estimate for slope, β_1

Example: Rocket Propellant

After removing the outliers due to faulty instrumentation, we had

$$\hat{y} = 2659 - 37.7x$$

We can say with 95% confidence that propellant loses between 33.5 and 41.9 psi of sheer strength per week **on average** between 2 and 25 weeks.

We can also reject the null hypothesis with an associated p-value

$$p < 0.0001$$

Equations Recap (1 of 2)

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$s^2 = MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n-2}$$

$$= \frac{SS_{yy} - \hat{\beta}_1 SS_{xy}}{n-2}$$

$$SS_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$SS_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

$$SS_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Equations Recap (2 of 2)

$$s_{\hat{\beta}_1} = \frac{s}{\sqrt{SS_{xx}}}$$

$$t = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$$

Next time…

Coefficient of Correlation (r)

Coefficient of Determination (r^2)