

13.E: Linear Regression (Exercises)

1. How are ANOVA and linear regression similar? How are they different?

Answer:

ANOVA and simple linear regression both take the total observed variance and partition it into pieces that we can explain and cannot explain and use the ratio of those pieces to test for significant relations. They are different in that ANOVA uses a categorical variable as a predictor whereas linear regression uses a continuous variable.

2. What is a residual?

3. How are correlation and regression similar? How are they different?

Answer:

Correlation and regression both involve taking two continuous variables and finding a linear relation between them. Correlations find a standardized value describing the direction and magnitude of the relation whereas regression finds the line of best fit and uses it to partition and explain variance.

4. What are the two parameters of the line of best fit, and what do they represent?

5. What is our criteria for finding the line of best fit?

Answer:

Least Squares Error Solution; the line that minimizes the total amount of residual error in the dataset.

6. Fill out the rest of the ANOVA tables below for simple linear regressions:

a.

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Model	34.21			
Error				
Total	66.12	54		

b.

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Model			6.03	
Error		16		
Total	19.98			

7. In chapter 12, we found a statistically significant correlation between overall performance in class and how much time someone studied. Use the summary statistics calculated in that problem (provided here) to compute a line of best fit predicting success from study times: $\bar{X} = 1.61$, $s_X = 1.12$, $\bar{Y} = 2.95$, $s_Y = 0.99$, $r = 0.65$.

Answer:

$$b = r(s_y/s_x) = 0.65 * (0.99/1.12) = 0.72; a = \bar{Y} - b\bar{X} = 2.95 - (0.72 * 1.61) = 1.79; \hat{Y} = 1.79 + 0.72X$$

8. Using the line of best fit equation created in problem 7, predict the scores for how successful people will be based on how much they study:

- $X = 1.20$
- $X = 3.33$
- $X = 0.71$
- $X = 4.00$

9. You have become suspicious that the draft rankings of your fantasy football league have no predictive value for how teams place at the end of the season. You go back to historical league data and find rankings of teams after the draft and at the end of the season (below) to test for a statistically significant predictive relation. Assume $SSM = 2.65$ and $SSE = 337.35$

Draft Projection	Final Rankings
1	14
2	6
3	8
4	13
5	2
6	15
7	4
8	10
9	11
10	16
11	9
12	7
13	14
14	12
15	1
16	5

Answer:

Step 1: $H_0 : \beta = 0$ "There is no predictive relation between draft rankings and final rankings in fantasy football," $H_A : \beta \neq 0$, "There is a predictive relation between draft rankings and final rankings in fantasy football."

Step 2: Our model will have 1 (based on the number of predictors) and 14 (based on how many observations we have) degrees of freedom, giving us a critical value of $F^* = 4.60$.

Step 3: Using the sum of products table, we find : $\bar{X} = 8.50$, $\bar{Y} = 8.50$, $SSX = 339.86$, $SP = 29.99$ giving us a line of best fit of: $b = 29.99/339.86 = 0.09$; $a = 8.50 - 0.09 \cdot 8.50 = 7.74$; $\hat{Y} = 7.74 + 0.09X$. Our given SS values and our df from step 2 allow us to fill in the ANOVA table:

Source	SS	df	MS	F
Model	2.65	1	2.65	0.11
Error	337.35	14	24.10	
Total	339.86	15		

Step 4: Our obtained value was smaller than our critical value, so we fail to reject the null hypothesis. There is no evidence to suggest that draft rankings have any predictive value for final fantasy football rankings, $F(1, 14) = 0.11, p > .05$

10. You have summary data for two variables: how extroverted some is (X) and how often someone volunteers (Y). Using these values, calculate the line of best fit predicting volunteering from extroversion then test for a statistically significant relation

using the hypothesis testing procedure:

$$\bar{X} = 12.58, s_X = 4.65, \bar{Y} = 7.44, s_Y = 2.12, r = 0.34, N = 67, SSM = 19.79, SSE = 215.77.$$

This page titled [13.E: Linear Regression \(Exercises\)](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Foster et al. \(University of Missouri's Affordable and Open Access Educational Resources Initiative\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.