1. Find the least squares regression line for using payroll to predict number of wins. Interpret the slope and the intercept in the context of the problem.



The slope of the line is 0.1056 wins for every million dollars spent in this regression equation. The intercept represents the amount of wins if payroll were zero. This is not a realistic number in the context of this problem as you must pay players a salary for them to play. No free lunch.

1. Is the slope of the regression line significantly different from zero? Carry out the appropriate test and interpret the results.

Hypothesis Test:



We are testing for the null hypothesis assumption that the slope is equal to zero. This will allow us to determine if our slope value is statistically significant and not due to chance or random variation.



Based on a p value of 0.0465, we reject the null hypothesis that the slope is equal to zero.

There is evidence to show that payroll and number of wins have a positive relationship, with 0.1056 wins per million dollars of payroll.

1. Calculate a confidence interval for the slope and interpret this interval.



Based on an error of 0.0507, the confidence interval of the slope is 0.0017 to 0.2095. We can assume that the procedure used will capture the mean in its interval 95% of the time. Further, we see that 0 is not in the confidence interval, confirming that the slope is indeed not 0.

I followed Dr. McGee’s methodology for this question. However, the book states that we should look at tests for the model without an intercept.

1. Give a 95% CI for the expected number of wins for a team with $100 million payroll. Give a 95% PI for the number of wins for a team with $100 million payroll. Explain the difference between these two intervals.

Only supposed to do 1H of question:

Confidence interval for expected mean number of wins for teams with $100m payroll:

