Post-Live Session Homework: Unit 9

Assignments

We will use SAS for all questions during the live session. For homework, you will be asked to verify the SAS results for questions 1-3 using hand calculations.

These are the same data from last week's homework. Now, we are going to use them for simple linear regression.

Team	Payroll	Wins	Team	Payroll	Wins	Team	Payroll	Wins
NYY	206	95	LAD	95	80	KC	71	67
BOS	162	89	HOU	92	76	TOR	62	85
CHC	146	75	SEA	86	61	ARZ	61	65
PHI	142	97	STL	86	86	CLE	61	69
NYM	134	79	ATL	84	91	WAS	61	69
DET	123	81	COL	84	83	FA	57	80
CHW	106	88	BAL	82	66	TEX	55	90
LAA	105	80	MIL	81	77	OAK	52	81
SF	99	92	TB	72	96	SD	38	90
MIN	98	94	CIN	71	91	PIT	35	57

Here are some summary statistics for these data to make doing this by hand a little easier:

$$\sum_{i=1}^{30} x_i = 2707 \qquad \sum_{i=1}^{30} x_i^2 = 286509 \qquad \sum_{i=1}^{30} x_i y_i = 223728 \qquad \sum_{i=1}^{30} (x_i - \bar{x})^2 = 42247.37$$

$$\sum_{i=1}^{30} y_i = 2430 \qquad \sum_{i=1}^{30} y_i^2 = 200342 \qquad \sum_{i=1}^{30} (y_i - \bar{y})^2 = 3512 \qquad \sum_{i=1}^{30} (x_i - \bar{x})(y_i - \bar{y}) = 4461$$

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.

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

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

4. 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.