**Exploring Logistic Regression Through Cricket**

Cricket is a game watched and played by billions of people across the world. Second in global popularity only to football (soccer), it is extremely popular in South Asia, Australia, Africa, and Europe.

**The Rules of Cricket**

Cricket is played on a rectangular pitch inside an oval boundary. In the pitch, on either side, there is a vertical **wicket**, made up of three vertical wooden "stumps" with two wooden bails (small blocks of wood) resting atop the stumps. There are two teams, each with 11 players (a batting team and a bowling team).

**The bowling team:** The bowler bounces a small leather ball at the batter. The goal of the bowler is to get the batter out and ultimately knock the bails off of the stumps.

**The batting team:** The goal of the batting team is to score as many runs as possible and not let the bowler hit the bail off of the stumps. There are two batters at a time for the batting team. After the batter hits the ball with the wooden, bat like paddle, the two batters can attempt to switch sides. Each batter will go until they get out, when at that time, they will switch with another teammate.

**Scoring Runs:** Each time the two batters switch sides without getting out counts as a run scored. For example, if you hit the ball and are able to run to the other wicket and back without getting out, that would score two runs for your team. If the batter hits the ball on a bounce to the oval boundary, it counts as four runs scored. If the batter hits the ball over the oval boundary on a fly, it counts as six runs.

**Outs:** Below are a few common ways in which a batter can get out

* If the batter hits the ball and it is caught in the air (caught out)
* If a player on the bowling team throws the ball and it hits the bails off the stumps before the batter can cross the line while trying to score a run (run out)
* If the bowler throws the ball past the batter and it knocks a bail off the wicket
* If the batter gets hit in the legs by a ball that would have hit the wicket (Note: if the batter gets hit and the ball wasn't ruled as one that would have hit the wicket, the batter is not out)

A screenshot of a sports game

Description automatically generatedThe duration of cricket matches can range from hours to days, depending on what format is being played, and scores are often high as each individual batter usually scores many runs on multiple pitches before getting out.

**Data**

The asia\_cup data set includes data from each cricket match played in all Asia Cup Tournaments from 1984 (the first one) to 2022. The Asia Cup is a tournament that now takes place every two years, alternating host cities in different countries throughout Asia. To the right is a description of the variables in the data set you will be using.

1. Provide the R code that when executed, would run a logistic regression that predicts the result of the match using runs scored, sixes, highest score, and extras given.
2. Given the R output below that runs a multiple logistic regression, interpret all of the coefficients in context.

**A screenshot of a computer

Description automatically generated**

1. Does the number of extras a team gives up during the match seem to matter? Why or why not
2. Write the logit form and the probability form of the multiple logistic regression model above.
3. Find the estimated log odds of a team that starts batting wins the match given they score 265 runs, 8 sixes, and give up 5 extras to the other team. Interpret the result in context.
4. Find the estimated probability of that same team winning the cricket match and interpret it in context.
5. In this model, does the exact number of runs scored by a team or the number of sixes scored by a team seem to matter more when predicting the outcome of the cricket match? Explain.