English Premier League 2022-2023

In this dataset we are looking at the results off all the games throughout the 2022-2023 season in the English Premier League. We have the amount of home and away goals at halftime and fulltime. We will be going through the sections that we covered in Probability and Statistics and create problems for anyone to solve using the dataset that we found. We might have to look at the dataset itself to solve the problems. Will also need some knowledge on the schedule of the English Premier League. I provided an answer key at the end of the problems with showing the work for the problems, so it is easy to understand how I got the answers. I hope this helped you learn about the advantage of playing at home compared to playing away.

**Chapter 2**

1. Suppose the first week of games in the season, and we are interested in the games that occurred on 08/06/2022. Let F denote the first day of games, S denote the second day of games, and T denote the third day of games. The set looks like the following:
   1. What is the set when both F and S happen?
   2. What is the set when only S happens?
   3. What is the set on the fourth day?
2. Assuming the second week of games, a sample space of home first half goals, , away first half goals, , home full-time goals, , away full-time goals, , and full-time win to the away team, .
   1. If , , and . Find the probability of .
   2. Now what will happen if we flip the probabilities of and . Find the probability of .
3. Continue looking at the second week of the games, there was a total of ten games in the week. Out of those games there were three games where the home team goals scored more than four goals. Suppose that three games are surveyed to see which games had the home team score four goals.
   1. List the sample space.
   2. Identify the simple events in each of the following events:
      1. At least two games had four goals scored.
      2. Exactly two games had four goals scored.
      3. Exactly one game had less than four goals scored.
   3. Make use of the given interpretation for the median to assign probabilities to simple events and find and
4. An experiment consists of two weeks of games, where each week consists of ten games.
   1. Use the combinatorial theorems to determine the number of sample points in the sample space .
   2. Find the probability that the sum of the numbers of games at 3:00 pm.
5. Suppose that we ask randomly selected teams whether they are in the English Premier League.
   1. Give an expression for the probability that no teams share a game in the league. (Total games are 296)
   2. How many teams do we need to select so that the probability is at least that at least one shares a game in the league?
6. For a certain month of games, the percentage of home teams winning a game was as shown in the accompanying table. That is, of all the games happening, were wins, 20% were loses, and 33% were draws. We can assume that the opposite results are for the away team. A game is to be selected randomly from this population. Let be the event that the game is a win for the home team and let be the event that the away team won.

|  |  |
| --- | --- |
| Outcomes | Home Team |
| Win | 31 |
| Lose | 13 |
| Draw | 22 |
| Total | 66 |

* 1. Are the events and independent?

1. Looking at a random week in the season, assume there are and . Where is home team wins and is away team wins. If a game ends, the probability that the home team wins is ; the away team wins is ; and we know both teams cannot win and the same time.
   1. When a game ends, find the probability that either home or away wins.
   2. Find the probability that neither team will win.
2. A population of results contains 40% home wins and 60% away wins. It is reported that 30% of the home wins and 70% of the away wins easily won their games. A game chosen at random from this population is found to favor the issue in question. Find the conditional probability that the game was won by the away team.

**Chapter 3**

1. You and a friend bet on two games during a random week. If the home team wins both games, you win $1; if the home team loses both games, you win $2; if both the home and away team win a game, you lose $1. Assume there are no draws in the games. Give the probability distribution for your winnings, , on a single play of this game.
2. A fan wants to analyze the games from two weeks where week 1 is , and week 2 is . The fan asks four people to help him with analyzing the games. Each of the people are randomly given three games, two from and other from . Each person is asked to state the stats from the games. Suppose that the games are fair with both teams’ equal odds to win. Let be the number of people stating which week the game was from.
   1. Find the probability function for .
   2. What is the probability that at least three of the four people state that the game was from week 1?
   3. Find the expected value of .
   4. Find the variance of .
3. Suppose that 30% of the results are in favor of the away team. Games are selected at random, and each is independent of each other. Find the probability that the first game is the part of one of five games that went in the favor of the away team.
4. Refer to Problem 11. What is the expected value of games till we find the first result in the favor of the away team.
5. An analyzer got given a list of 20 games, of which ten are home wins, three are away wins, and two are draws. Six games are to be selected from the list, one at a time without replacement. What is the probability that all six games were home wins?
6. Refer to Problem 13. The analyzer wants to know what the mean and variance is with the same values.
7. From a random two consecutive weeks of results. The Poisson distribution at an average of 0.35 the away team wins per game. During a given game, what are the probabilities that:
   1. No more than four away teams win?
   2. At least three away teams win?
   3. Exactly two away teams win?
8. Would you rather take a bet on the away team or a draw? If you have absolutely no knowledge of the results. If you bet on the away team, there is a 29% chance of them winning and a 24% chance of a draw. There are a total of 296 games in a season. Find the answer if you bet on a draw.
   1. What is the expected value of the number games that you bet on that you will win?
   2. Find the standard deviation of .
   3. Calculate the intervals and .

**Chapter 4**

1. For a team to gain up to three points from a result. A win gives a team three points, a draw gives them one point, and a loss gives them zero points. Let three points equal one in the density function. The total amount of points gained from the result is a random variable with a probability density function given by
   1. Graph .
   2. Find and graph it.
   3. Find the probability that the team will gain between 0 to 1 point in the result.
   4. Find the probability that the team will gain more than 1 point in the result.
2. Upon studying the number of goals scored by the away team throughout the entire season. The number of goals is uniformly distributed between 0 to 5. Find the probability that the goals scored next season:
   1. Are below 3 goals.
   2. Are above 4 goals.
3. Refer to problem 18. Find the expected values of goals of the type described there.
4. If, as in Problem 17, has density function,

Find the mean and variance of .

**Chapter 5**

1. Results from two random games and they are randomly assigned by a win, , a draw, , or a loss, . Let denote the number of wins and denote the number of draws. Recall that each result can receive 0, 1, or 3 points.
   1. Find the joint probability function for and .
   2. Find .
2. Let and denote the proportion of games (out of their schedule) during which team I and II, respectively, play their games. The joint frequency behavior of and is modeled by the density function.
   1. Find
   2. Find .
3. In problem 21, we determined that the joint distribution of , the number of wins, and , the number of draws, is given by the entries in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| \ | 0 | 1 | 3 |
| **0** |  |  |  |
| **1** |  |  | 0 |
| **3** |  | 0 | 0 |

* 1. Find the marginal probability distribution of .
  2. According to results in Chapter 4, has a binomial distribution with and . Is there any conflict between this result and the answer you provided in part(a)?

1. In problem 21, we determined that the joint distribution of , the number of wins, and , the number of draws, is given by the entries in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| \ | 0 | 1 | 3 |
| **0** |  |  |  |
| **1** |  |  | 0 |
| **3** |  | 0 | 0 |

The marginal probability function of was derived in problem 23 to be binomial with and . Are and independent? Why?

**Answer Key**

Chapter 2

1. The probabilities are:
2. The answers are:
   1. The sample space is: is less than four goals and is four or more goals.
   2. Events are:
      1. through
      2. through }
      3. through
3. Sample space :
   1. games
   2. sum of the games at 3:00pm and
4. Share a game:
   1. and thus
5. and
   1. it is not independent.
6. Let be the home team, be the away team, and is the easy win.

Chapter 3

1. Possible outcomes: Let be a win, be a lost, and be a draw.

win $1, win $-1, win $-1, win $2.

Each outcome has a probability.

1. People = 4 # of games per person = 3

1. The mean is:

The variance is:

2. Answers are:
   1. Binomial distribution 0.24
   2. The intervals are:

Chapter 4

1. The answers are:
   1. The graph of is:

Chart

Description automatically generated

* 1. is:

Chart, line chart

Description automatically generated

1. The answers are:
2. The mean is:

Chapter 5

* 1. The joint probability function looks like:

|  |  |  |  |
| --- | --- | --- | --- |
| \ | 0 | 1 | 3 |
| **0** |  |  |  |
| **1** |  |  | 0 |
| **3** |  | 0 | 0 |

* 1. or

1. The answers are:
2. The answers are:

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0 | 1 | 2 |
|  |  |  |  |

* 1. **No**

1. The marginal density of is:

The marginal density of is:

If you consider . Therefore and are not independent.

Work-Cited

https://www.kaggle.com/datasets/evangower/premier-league-2022-2023?select=epl\_results\_2022-23.csv