A Quasi-Poisson Approach Goals in Football, What contributes to them?

Nayani Jani
Christopher Odoom
Denis Folitse
Animesh Sengupta

University of Massachusetts, Amherst

November 27, 2022

Background of study

Football has become a part of the world and is one of the biggest source of entertainment in the world. It has created job and opportunities to people all over the world. It's important in this modern time cannot be overlooked.There are five major football leagues in the world; English premier league, German BundesLiga, Spanish Laliga, Italian Seria A and France League one. Football supporters of English primier league all over the world always claim that, their league is the difficult one among all the five major leagues, is this claim true?. Football fans all over the world have insufficient answer to this question. This study tries to provide answer to this question.

Problem statement

Argument: If a league has a lot of goals scored in a season it is deemed not competitive because of the quality of opponents (vice versa as well). Goal: Does the factor league determine the amount of goals combined with other variables? If not, what predictors are associated with the response Goals?

Objectives

This study is directed by the specific objectives stated below;

1. Determine if there is a significant difference in goals across the 5 major leagues.

2. Determine if there are other predictors that are associated with the goals other than the different leagues.

Methodology

This study used a data which was collected from https://fbref.com, We used the 2021-2022 season statistics for EPL, La Liga, Bundesliga, Ligue 1 and Serie A in the analysis. The data comprise of 98 observations and 18 variables. Models we considered includes Poisson regression, Quasi - Poisson Regression. Diagnostic Tests: Test for Overdispersion, Residual Vs. Fitted Plot, Normal QQ-Plot

Poisson Regression

This is type a regression used when the response variable is a COUNT variable. It models the expected value of occurrence of an event as a function of some predictor variables.

The mathematical equation for the model is given below:

$$log(E(no.Goals|X)) = \beta_0 + \beta' X$$

Alternatively, we can write it in an exponential form as:

$$E(No.Goals|X) = e^{(\beta_0 + \beta' X)}$$