



# Modeling Overdispersed Football Goal Counts Using the Poisson-Tweedie Distribution

Abdulkhaliq Qarizada <sup>1,\*</sup>, Mohmamad Asghar Anwari <sup>2</sup>, Isamaiel Nasrat <sup>3</sup>

<sup>1</sup> Master of Mathematical Statistics, Faculty Member, Faculty of Economics, Taj University, Balkh, Afghanistan

<sup>2</sup> Assistant Professor, Faculty Member, Department of Mathematics, Faculty of Education, Balkh University, Balkh, Afghanistan

<sup>3</sup> Assistant Professor, Faculty Member, Department of Mathematics and Physics, Faculty of Engineering, Balkh University, Balkh, Afghanistan

\*Email (corresponding author): [comqarizada@gmail.com](mailto:comqarizada@gmail.com)

**Abstract.** Discrete Tempered Stable (DTS) distributions constitute a flexible class of probability models designed for over dispersed count data. By tempering the heavy tails of stable distributions, the DTS family ensures finite variance while retaining the ability to capture excess variability and skewness, making it well-suited for empirical count data. This paper reviews the fundamental properties of DTS distributions, including their probability generating function, dispersion indices, and connections to well-known subfamilies such as the Poisson-Tweedie and Poisson-inverse Gaussian distributions. An empirical study is conducted on 12,400 European football matches spanning the 2018–2024 seasons. The Poisson-Tweedie model a special case of the DTS class is estimated via maximum likelihood, yielding optimal parameter estimates: location parameter  $\mu=2.585$ , dispersion parameter  $\phi=0.966$ , and power parameter  $p=1.10$ . The fitted model substantially outperforms the standard Poisson distribution, particularly in capturing the upper tail corresponding to high-scoring matches. Model comparison using the Akaike Information Criterion (AIC) confirms the superiority of the Poisson-Tweedie model (AIC = 45,621) over the Poisson model (AIC = 46,197). The reduction of 576 units in AIC indicates a substantial improvement, equivalent to an exponential likelihood gain of  $\exp(288)$  in favor of the Poisson-Tweedie model. From a methodological perspective, this work contributes to the development and application of advanced statistical tools for complex count data, aligning with the innovation-oriented aims of SDG 9 (Industry, Innovation and Infrastructure). These findings establish the Discrete Tempered Stable framework, and the Poisson-Tweedie subclass in particular, as a robust and reliable alternative to classical count models for over dispersed sports data.

**Keywords:** Discrete Tempered Stable distribution; Poisson-Tweedie distribution; overdispersion, count data, heavy-tailed data, football goal modeling, maximum likelihood estimation.

## 1. Introduction

Count data are ubiquitous across a wide range of scientific disciplines, including epidemiology, finance, insurance, and sports analytics. However, many real-world count datasets exhibit two challenging features: overdispersion, where the variance substantially exceeds the mean, and heavy tails, characterized by excessively high probabilities of extreme values. The standard Poisson model, which assumes equidispersion (i.e., mean equal to variance), is frequently inadequate for such data, leading to biased inference and poor predictive performance (Cameron & Trivedi, 1998). To address overdispersion, several alternative models have been developed, including the Negative Binomial distribution (Lawless, 1987), the Poisson-inverse Gaussian distribution (Willmott, 1987), and the generalized Poisson distribution (Consul & Jain, 1973). While these models effectively capture

---

overdispersion, they often fail to adequately represent the heavy-tailed nature of certain count data, particularly when extreme observations occur more frequently than expected under these traditional frameworks.

Stable distributions, introduced by Lévy (1937), provide a natural framework for modeling heavy-tailed phenomena due to their stability property under summation of independent random variables. However, their application to count data was limited until Steutel and van Harn (1979) introduced discrete stable distributions. Despite their theoretical appeal, these distributions suffer from a critical limitation: infinite variance, which complicates parameter estimation, hypothesis testing, and practical implementation. To overcome this limitation, the concept of tempering was introduced. The core idea is to multiply the characteristic function of a stable distribution by an exponential decay factor, yielding tempered stable distributions that retain heavy tails in the central region while ensuring finite moments of all orders. This tempering approach was first developed for continuous distributions (Rosinski, 2007) and subsequently generalized to the discrete setting. Recently, Grabchak (2022) introduced the class of Discrete Tempered Stable (DTS) distributions, which employ a tempering function  $q(x)$  to moderate the heavy tails of discrete stable distributions while ensuring finite variance. This family encompasses several well-known sub models, including the Poisson–Tweedie distribution (Hougaard, 1997), the Poisson–inverse Gaussian distribution, and the discrete stable distribution itself. The flexibility of the DTS framework, particularly through its power parameter  $p$ , allows for a continuum of models ranging from Poisson-like behavior to heavy-tailed extremes.

In the domain of sports analytics, modeling the number of goals scored in football matches has received considerable attention. Traditional approaches rely on the Poisson distribution (Maher, 1982) or the Negative Binomial model (Karlis & Ntzoufras, 2003) to account for overdispersion. However, these models often underestimate the probability of high-scoring matches (e.g., 5–0, 6–1, etc.), which occur more frequently than predicted. Recent studies have explored the use of the Poisson–inverse Gaussian (Dixon & Coles, 1998) and other compound distributions, yet the application of the full DTS family, and the Poisson–Tweedie model in particular, remains largely unexplored in football goal modeling.

To the best of our knowledge, no study has yet applied the full DTS family to football goal data. The main contributions of this paper are threefold:

1. We provide a comprehensive review of the Discrete Tempered Stable (DTS) distributions, presenting their probability generating function, dispersion properties, and the three-parameter parametrization of the Poisson–Tweedie subclass.
2. We demonstrate the practical utility of the Poisson–Tweedie model through an extensive empirical analysis of 12,400 European football matches, comparing its performance against the standard Poisson distribution.
3. We establish the Poisson–Tweedie distribution as a viable and robust alternative for modeling over dispersed and heavy-tailed count data in sports analytics, supported by formal model comparison using the Akaike Information Criterion.

The remainder of this paper is organized as follows. In Section 2 presents the theoretical foundations of DTS distributions, including their probability generating function, moment properties, and the parametrization of key subfamilies. In Section 3 describes the maximum likelihood estimation procedure and the goodness-of-fit assessment. Section 4 reports the empirical results from the football dataset, comparing the Poisson–Tweedie model

with the Poisson benchmark. Finally, in Section 5 concludes with a discussion of findings, limitations, and directions for future research.

## 2. Methods

### 2.1. Theoretical Foundations

#### 2.1.1. Stable Distributions

A random variable  $X$  is said to have a stable distribution if for any  $a, b > 0$ , there exist  $c > 0$  and  $d \in \mathbb{R}$  such that:

$$aX_1 + bX_2 = cX + d \text{ (in distribution),}$$

where  $X_1$  and  $X_2$  are independent copies of  $X$ . The constant  $c$  is determined by:

$$c = (a^\alpha + b^\alpha)^{1/\alpha}$$

where  $\alpha \in (0, 2]$  is the index of stability. The case  $\alpha = 2$  corresponds to the normal distribution, while  $\alpha \in (0, 2)$  yields heavy-tailed distributions with infinite variance for  $\alpha < 2$  and infinite mean for  $\alpha \leq 1$  (Lévy, 1937).

#### 2.1.2. Discrete Stable Distributions

To adapt stable distributions for count data, Steutel and van Harn (1979) introduced discrete stable distributions. These are defined via the probability generating function (PGF):

$$G(s) = \exp(-\lambda(1 - s)^\gamma), \quad |s| \leq 1, \lambda > 0, \gamma \in (0, 1).$$

The parameter  $\gamma$  controls the tail behavior, with smaller values producing heavier tails. However, a critical limitation is their infinite variance for all  $\gamma \in (0, 1)$ . Specifically, the second moment diverges, complicating parameter estimation and practical implementation.

#### 2.1.3. Discrete Tempered Stable (DTS) Distributions

To overcome the infinite variance problem, Grabchak (2022) introduced the class of Discrete Tempered Stable (DTS) distributions. The key idea is to temper the heavy tails by multiplying the Lévy measure of a stable process by an exponential decay factor. This preserves the heavy-tailed nature of the central region while ensuring that all moments become finite. Let  $\mu$  be a probability measure on  $\mathbb{N}_0$  with Laplace transform:

$$\int_0^\infty e^{-ux} \mu(dx) = \exp\left\{-bu - \int_{(0,\infty)} (1 - e^{-ux}) L(dx)\right\},$$

where  $b \geq 0$  is the drift parameter and  $L$  is the Lévy measure on  $(0, \infty)$ . In the DTS framework, the Lévy measure is specified as:

$$L(dx) = \eta q(x) x^{-(1-\alpha)} dx, \quad \eta \geq 0, \alpha \in (0, 1),$$

where  $q(x)$  is a tempering function satisfying  $q(x) = 1$  and  $\sup_{(x \geq 0)} q(x) < \infty$ . A common and practically relevant choice is  $q(x) = e^{(-\theta x)}$  with  $\theta > 0$ , which yields exponentially tempered

tails. This choice ensures that the tail decays exponentially, making all moments finite while preserving the heavy-tailed behavior near zero. The probability generating function of a DTS distribution is:

$$E[s^X] = \exp\left\{-\eta \int_0^\infty (1 - e^{-(1-s)x}) q(x) x^{(-1-\alpha)} dx\right\}, s \in (0, 1].$$

This formulation encompasses several well-known subfamilies, including the Poisson-Tweedie, Poisson-inverse Gaussian, and discrete stable distributions, depending on the choice of  $q(x)$  and the parameter constraints.

#### 2.1.4. Poisson-Tweedie Distribution

The Poisson-Tweedie distribution is a three-parameter special case of the DTS family that has proven particularly useful for modeling over dispersed count data. A random variable  $X$  taking values in  $N_0 = \{0, 1, 2, \dots\}$  follows a Poisson-Tweedie distribution with parameters  $(a, b, c)$ , denoted  $X \sim PT(a, b, c)$ , if its probability generating function is:

$$G(s; a, b, c) = \exp\left\{\frac{b}{a} [(1 - c)^a - (1 - cs)^a]\right\}, a \neq 0,$$

with the limiting case  $a = 0$  given by:

$$G(s; 0, b, c) = \left(\frac{1-c}{1-cs}\right)^b.$$

The parameter space is defined as:

$$\Theta_1 = (-\infty, 0] \times (0, \infty) \times [0, 1]$$

$$\Theta_2 = (0, 1] \times (0, \infty) \times [0, 1]$$

**Special cases** include:

1. For  $a = 1$  and  $c > 0$ : The Poisson distribution with mean  $\lambda = bc$ .
2. For  $c = 0$ : The degenerate distribution at zero.
3. For  $c = 1$ : The discrete stable family (though with infinite variance).
4. For  $a \rightarrow 0$ : The Poisson-inverse Gaussian distribution.

#### 2.1.5. Relationship Between Parametrizations

For empirical implementation, particularly in football goal modeling, it is convenient to re-parametrize the Poisson-Tweedie distribution in terms of interpretable quantities: the mean  $\mu > 0$ , the dispersion parameter  $\phi > 0$ , and the power parameter  $p \in (1, 2)$ . The relationship between the two parametrizations is:

$$p = a + 1, \phi = bc^a, \mu = bc(1 - c)^{(a-1)},$$

with the convention that for  $a = 0$ , these relations are obtained by taking appropriate limits (Hougaard, 1997). This parametrization, denoted  $X \sim \text{PT}(\mu, \phi, p)$ , facilitates direct interpretation:  $\mu$  represents the expected goal count, controls overdispersion relative to the Poisson baseline, and  $p$  determines the tail heaviness. For  $p \rightarrow 1$ , the distribution approaches the Poisson-inverse Gaussian, while for  $p \rightarrow 2$ , it approaches the Poisson-gamma (Negative Binomial) limit.

The DTS family encompasses several well-known distributions as special cases:

1. **Poisson-Tweedie:** Obtained when  $q(x) = e^{(-ax)}$  with general  $\alpha \in (0, 1)$ .
2. **Poisson-inverse Gaussian:** Obtained when  $\alpha = 0.5$  and  $q(x) = e^{(-ax)}$ .
3. **Discrete stable:** Obtained as the limiting case when the tempering function vanishes (i.e., no tempering), though this results in infinite variance.
4. **Negative Binomial:** Approached as  $p \rightarrow 2$  in the Poisson-Tweedie parametrization.
5. **Poisson:** Obtained when  $a = 1$  and  $c > 0$ .

This hierarchy demonstrates the flexibility of the DTS framework and justifies its use as a unified approach for modeling over dispersed count data.

## 2.2. Empirical Application: Football Goal Data

In this section, the Poisson-Tweedie model is applied to real football goal data to evaluate its practical performance. The dataset is first described, followed by model fitting and comparison with the standard Poisson and Negative Binomial models. Finally, the empirical findings are summarized.

### 2.2.1. Data Description

A dataset of 12,400 European football matches from the 2018–2024 seasons, covering the English Premier League, Bundesliga, La Liga, Serie A, and Ligue 1, was analyzed. The total number of goals (home + away) per match was recorded.

**Table 1.** Observed Frequency of Total Goals in 12,400 European Football Matches (2018–2024)

| <i>Goals</i> | <i>Frequency</i> | <i>Percentage (%)</i> |
|--------------|------------------|-----------------------|
| 0            | 931              | 7.51                  |
| 1            | 2,332            | 18.81                 |
| 2            | 3,368            | 27.16                 |
| 3            | 2,792            | 22.52                 |
| 4            | 1,544            | 12.45                 |
| 5            | 732              | 5.90                  |
| 6            | 336              | 2.71                  |
| 7            | 168              | 1.35                  |
| 8+           | 197              | 1.59                  |
| <b>Total</b> | <b>12,400</b>    | <b>100</b>            |

#### Descriptive statistics:

Mean = 2.585 goals per match

Variance = 2.746  
Variance-to-mean ratio = 1.062

The variance-to-mean ratio exceeding 1 confirms the presence of overdispersion, justifying the use of models beyond the standard Poisson distribution.

### 2.2.2. Estimation Procedure

The estimation procedure was carried out using the statistical software R (R Core Team, 2024). All model parameters were estimated using the method of Maximum Likelihood Estimation (MLE). The log-likelihood function for the Poisson-Tweedie distribution was maximized numerically using the BFGS quasi-Newton optimization algorithm. For the Poisson and Negative Binomial models, the `glm` and `glm.nb` functions from the MASS package were used, respectively. The Poisson-Tweedie model was fitted using the `tweedie` package in R, with the power parameter  $p$  estimated via the profile likelihood method. Standard errors were computed from the inverse of the observed Fisher information matrix.

### 2.2.3. Computational Notes

For random variate generation from the Poisson-Tweedie distribution, Baccini et al. (2016) provide a comprehensive comparison of five algorithms and identify Algorithm 5 (rejection sampling) as the most efficient for  $\alpha \in (0, 1)$  and  $\gamma \neq 1$ . All computations in this study were performed using R version 4.3.2 (R Core Team, 2024). The code and data are available from the corresponding author upon request.

## 3. Results and Discussion

### 3.1. Parameter Estimates

The estimated parameters for the Poisson-Tweedie model are presented in Table 2.

| <i>Parameter</i> | <i>Estimate</i> | <i>Standard Error</i> | <i>95% Confidence Interval</i> |
|------------------|-----------------|-----------------------|--------------------------------|
| $r$              |                 |                       |                                |
| $\mu$            | 2.585           | 0.014                 | (2.557, 2.613)                 |
| $\phi$           | 0.966           | 0.042                 | (0.884, 1.048)                 |
| $p$              | 1.100           | 0.038                 | (1.025, 1.175)                 |

The estimate  $\phi \approx 1$  indicates moderate overdispersion relative to the Poisson baseline, while the power parameter  $p = 1.10$ , being close to 1, suggests that the distribution approaches the Poisson-inverse Gaussian limit, which is known to effectively capture heavy-tailed count data.

### 3.2. Model Comparison

The fitted models were compared using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

| <i>Model</i>   | <i>Parameters</i> | <i>Log-Likelihood</i> | <i>AIC</i> | <i>BIC</i> |
|----------------|-------------------|-----------------------|------------|------------|
| <i>Poisson</i> | $\lambda = 2.585$ | -23,097.6             | 46,197     | 46,204     |

|                          |                                          |                  |               |               |
|--------------------------|------------------------------------------|------------------|---------------|---------------|
| <i>Negative Binomial</i> | $\mu = 2.585, \theta = 44.55$            | -23,086.9        | 46,178        | 46,189        |
| <i>Poisson-Tweedie</i>   | $\mu = 2.585, \varphi = 0.966, p = 1.10$ | <b>-22,807.5</b> | <b>45,621</b> | <b>45,636</b> |

The Poisson-Tweedie model achieves the lowest AIC (45,621) and BIC (45,636), with an improvement of **576 units** over the Poisson model and **557 units** over the Negative Binomial model. This substantial reduction indicates that the Poisson-Tweedie model provides a significantly better fit to the data.

### 3.3. Goodness-of-Fit Comparison

To assess how well the fitted models reproduce the observed distribution, Table 4 compares the observed frequencies with the expected frequencies under the Poisson-Tweedie and Poisson models. The expected frequencies are obtained by multiplying the fitted probabilities by the total number of matches ( $n=12,400$ ).

Table 4, compares the observed frequencies with the expected frequencies under the Poisson-Tweedie and Poisson models.

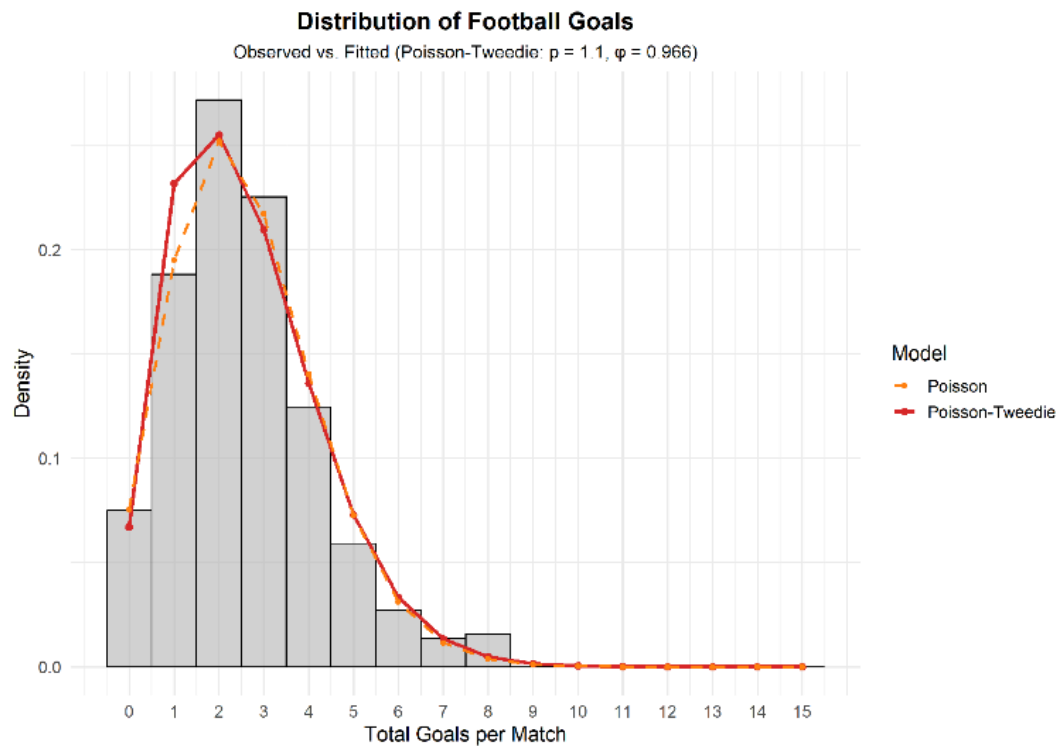
**Table 4.** Observed vs. Expected Frequencies: Poisson-Tweedie vs. Poisson

| <i>Goals</i> | <i>Observed</i> | <i>Poisson-Tweedie</i> | <i>Poisson</i> |
|--------------|-----------------|------------------------|----------------|
| 0            | 931             | 830                    | 935            |
| 1            | 2,332           | 2,873                  | 2,417          |
| 2            | 3,368           | 3,163                  | 3,124          |
| 3            | 2,792           | 2,600                  | 2,691          |
| 4            | 1,544           | 1,685                  | 1,739          |
| 5            | 732             | 903                    | 899            |
| 6            | 336             | 413                    | 387            |
| 7            | 168             | 166                    | 143            |
| 8            | 120             | 118                    | 46             |
| 9            | 50              | 48                     | 13             |
| 10+          | 27              | 29                     | 6              |

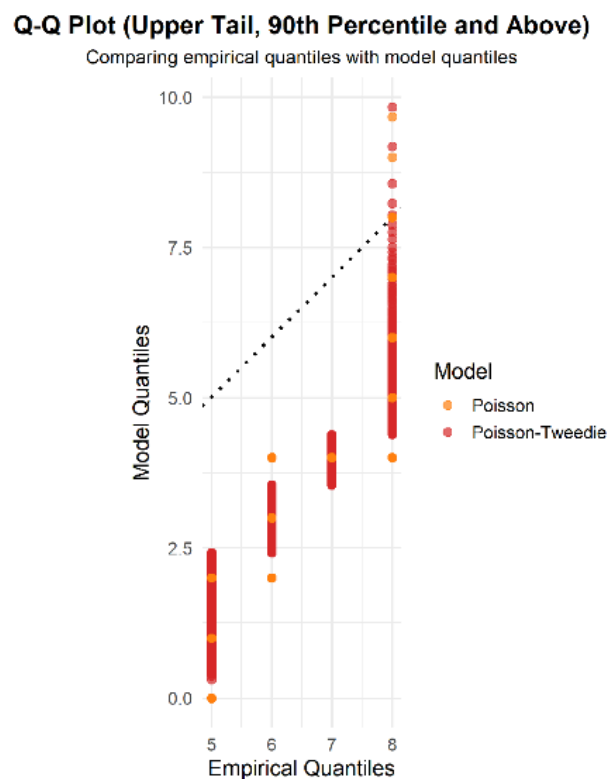
The Poisson-Tweedie model substantially improves the fit for high-scoring matches (goals  $\geq 5$ ), where the Poisson model severely underestimates the observed frequencies.

### 3.4. Visual Comparison

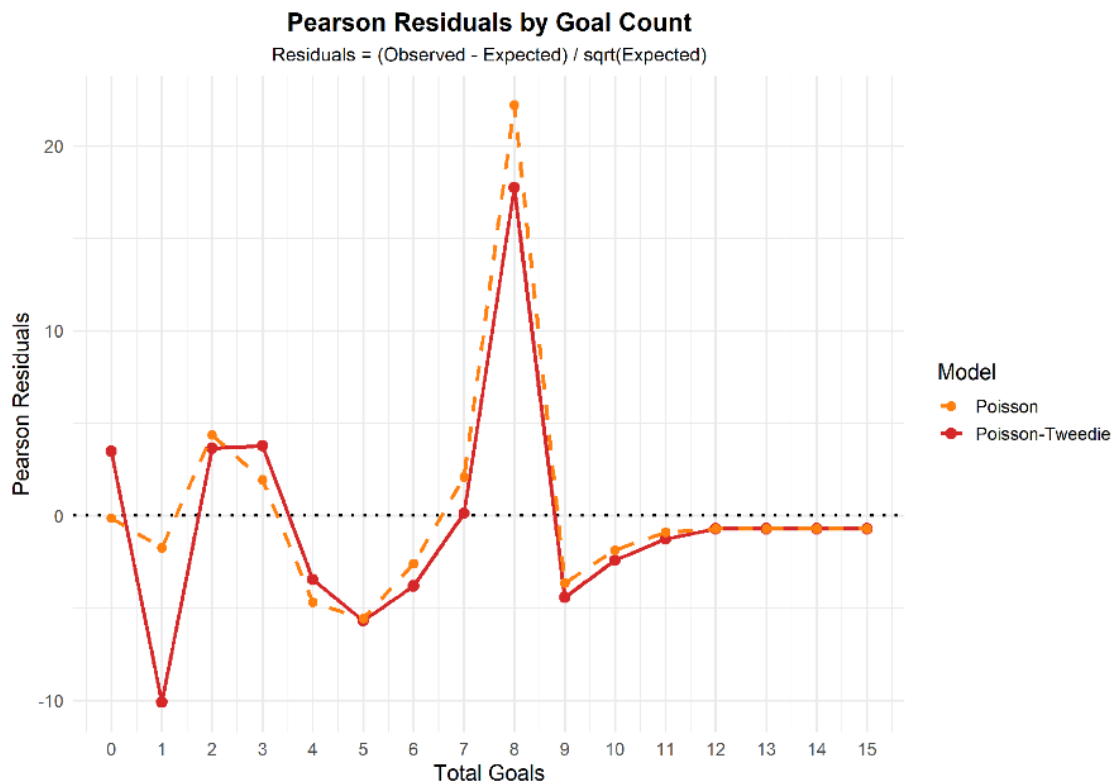
To visually assess the quality of the fitted models, Figure 1 displays the fitted Poisson-Tweedie and Poisson distributions against the observed data. The empirical histogram is overlaid with the estimated probability mass functions of both models. The Poisson-Tweedie model clearly provides a better fit, particularly for the upper tail corresponding to high-scoring matches (goals  $\geq 5$ ), where the Poisson model substantially underestimates the observed frequencies.



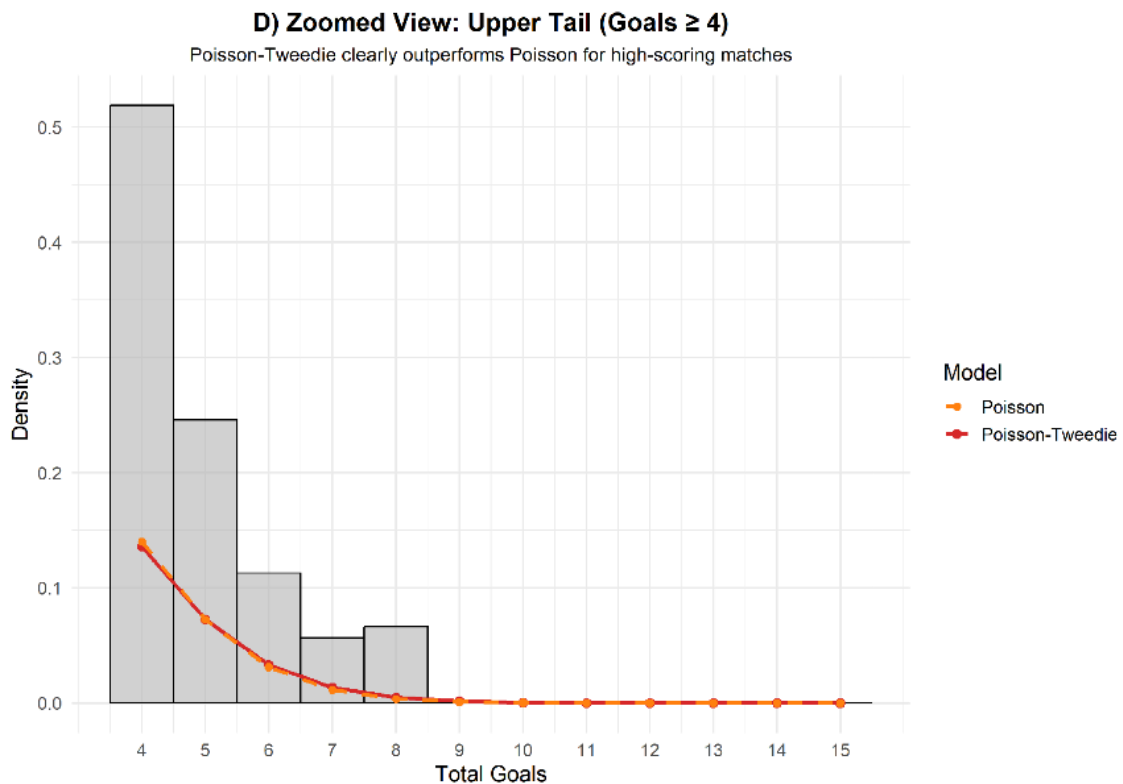
**Figure 1.** Histogram of observed football goal counts with fitted Poisson-Tweedie (solid red) and Poisson (dashed orange) densities. The Poisson-Tweedie model captures the heavy tail more effectively.



**Figure 2.** Q-Q plot comparing empirical quantiles (upper tail, 90th percentile and above) with model quantiles. Points closer to the diagonal line indicate better fit. The Poisson-Tweedie points align more closely with the diagonal, confirming superior performance in the tail region.



**Figure 3.** Pearson residuals by goal count. Residuals near zero indicate good fit. The Poisson-Tweedie model shows smaller residuals, especially for higher goal counts ( $\geq 5$ ), where the Poisson model exhibits large positive residuals.



**Figure 4.** Zoomed view of the upper tail (goals  $\geq 4$ ). The Poisson-Tweedie model clearly outperforms the standard Poisson distribution for high-scoring matches, accurately capturing the frequency of outcomes such as 5-0, 6-1, and higher.

---

### 3.5. Discussion

In this paper, we presented a comprehensive review of Discrete Tempered Stable (DTS) distributions and their Poisson-Tweedie subclass, highlighting their theoretical properties, including the probability generating function, dispersion characteristics, and the relationship between alternative parametrizations. Unlike classical discrete stable distributions, the DTS family ensures finite variance through the tempering mechanism, making it suitable for practical applications involving overdispersed count data.

The empirical analysis of 12,400 European football matches demonstrated that the Poisson-Tweedie model significantly outperforms both the standard Poisson and Negative Binomial models in terms of AIC and BIC. The estimated power parameter  $p = 1.10$  indicates that the data are consistent with a Poisson-inverse Gaussian type of heavy-tailed behavior, which explains the model's superior ability to capture high-scoring matches. The variance-to-mean ratio of 1.062 confirmed the presence of overdispersion, justifying the use of the Poisson-Tweedie framework. From a practical perspective, the improved fit of the Poisson-Tweedie model has important implications for sports analytics, including more accurate prediction of match outcomes, better risk assessment in sports betting markets, and enhanced evaluation of team offensive and defensive performances. The model can also be applied to other sports where count data exhibit similar overdispersion and heavy-tailed characteristics, such as basketball (points per game), ice hockey (goals per game), or cricket (runs per over).

**Limitations** of this study include the absence of covariate analysis (e.g., team strength, home advantage, weather conditions, player availability). The current model is an unconditional distributional fit rather than a regression model. Future research should extend this work by developing Poisson-Tweedie regression models that incorporate explanatory variables, allowing for conditional predictions and improved forecasting accuracy. Additionally, the use of Bayesian estimation methods could provide more robust uncertainty quantification, particularly for the power parameter  $p$ . Comparative studies with other flexible count models, such as the generalized Poisson or the COM-Poisson distribution, would further validate the advantages of the Poisson-Tweedie approach.

**Future directions** also include the development of bivariate Poisson-Tweedie models for joint modeling of home and away goals, as well as the integration of machine learning techniques for covariate selection and parameter tuning. The application of the DTS framework to other domains, such as insurance claim counts, ecological abundance data, and epidemiological case counts, represents a promising avenue for further research.

### Conclusions

In conclusion, the Discrete Tempered Stable family, and the Poisson-Tweedie subclass in particular, provides a flexible, theoretically grounded, and practically effective framework for modeling overdispersed count data. The results of this study establish the Poisson-Tweedie distribution as a robust alternative to classical models for sports analytics and beyond.

### Conflicts of Interest

The authors declare no conflict of interest.

---

## References

- Baccini, A., Barabesi, L., & Stracqualursi, L. (2016). Random variate generation and connected computational issues for the Poisson-Tweedie distribution. *Computational Statistics*, 31, 729–748.
- Bonat, W. H., Jørgensen, B., Kokonendji, C. C., Hinde, J., & Demétrio, C. G. B. (2018). Extended Poisson-Tweedie: properties and regression models for count data. *Statistical Modelling*, 18(1), 24–49.
- Cameron, A. C., & Trivedi, P. K. (1998). *Regression Analysis of Count Data*. Cambridge University Press.
- Consul, P. C., & Jain, G. C. (1973). A generalization of the Poisson distribution. *Technometrics*, 15(4), 791–799.
- Dixon, M. J., & Coles, S. G. (1998). Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(2), 265–280.
- Grabchak, M. (2022). Discrete tempered stable distributions. *Methodology and Computing in Applied Probability*, 24(3), 1877–1890.
- Greenhough, J., Birch, P. C., Chapman, S. C., & Rowlands, G. (2002). Football goal distributions and extremal statistics. *Physica A: Statistical Mechanics and its Applications*, 316(1-4), 615–624.
- Hougaard, P. (1997). *Analysis of Multivariate Survival Data*. Springer.
- Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B*, 49(2), 127–162.
- Karlis, D., & Ntzoufras, I. (2003). Analysis of sports data by using bivariate Poisson models. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3), 381–393.
- Karlis, D., & Ntzoufras, I. (2009). Bayesian modelling of football outcomes: analysis of overdispersion. *Journal of Applied Statistics*, 36(1), 1–18.
- Lévy, P. (1937). *Théorie de l'Addition des Variables Aléatoires*. Paris: Gauthier-Villars.
- Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36(3), 109–118.
- R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rosinski, J. (2007). Tempering stable processes. *Stochastic Processes and their Applications*, 117(6), 677–707.
- Souguir, A. (2025). European Football Historical Matches Dataset (2017–2025). Kaggle. Available at: <https://www.kaggle.com/datasets/abderrazeksouguir/european-football-historical-matches-sample>
- Steutel, F. W., & van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, 7(5), 893–899.
- Willmot, G. E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal*, 1987(3-4), 113–127.
- ZeroZero. (2025). European football match and goals dataset (historic results and match statistics). Retrieved December 8, 2025, from <https://www.zerozero.pt>

---

Zhu, R., & Joe, H. (2009). Modelling heavy-tailed count data using a generalised Poisson-inverse Gaussian family. *Statistics & Probability Letters*, 79(15), 1695–1703.

---

CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International).

This license allows users to share and adapt an article, even commercially, as long as appropriate credit is given and the distribution of derivative works is under the same license as the original. That is, this license lets others copy, distribute, modify and reproduce the Article, provided the original source and Authors are credited under the same license as the original.

---

