A Retrospective Study on Skellam and Related Distributions

Lishamol Tomy
Deva Matha College

Veena G
St.Thomas College

Abstract
This paper reviews works on Skellam distribution, its extensions and areas of applications. Available literature shows that this distribution is flexible for modelling integer data where they appear as count data or paired count data in the field of finance, medicine, sports, and science. Bivariate Skellam distribution, dynamic Skellam model and other extensions are also discussed and additional literature are provided.

Keywords: Skellam distribution, discrete distribution, integer set, Poisson difference distribution, dynamic Skellam model, bivariate Skellam model.

1. Introduction
In various areas of research, researchers have often come across variables that are discrete in nature or practice. For example, in life testing experiments, it is sometimes not possible or inconvenient to measure the life length of a device on a continuous scale; in many applied situations, such as in reliability, data are measured in terms of the number of runs, cycles, or shocks the device endures before failing. In survival analysis, one may note the number of days of survival for patients since the treatment or the time from remission to relapse is noted as the number of days as well. In all the cases, the most suitable model for fitting the data is a discrete random variable.

Before we progress further, we should underline that the Skellam distribution introduced by Skellam (1946), can be used to model any discrete measurement which lie in the set of integers. Some contributions which have elevated interest in the distribution of Skellam to model integer results, can be seen in Karlis and Ntzoufras (2006), Karlis and Ntzoufras (2009) and Barndorff-Nielsen, Pollard, and Shephard (2012). It can be observed as a distribution on positive and negative integers, but can also be made from the differences in pairs of Poisson counts; see Skellam (1946). Paired count observations and their differences appear in many circumstances and research fields.

The article is structured as follows. Section 2 presents the insights to the distribution wherein the basic features of Skellam distribution, bivariate Skellam distribution and Skellam distribution in Time series analysis are discussed. The applications of the distribution are mentioned
in Section 3 and Section 4 concludes the article.

2. Insights to the distribution

2.1. Basic features of Skellam and related distributions

• Irwin (1937) derived the probability mass function (pmf) of the difference between two independent Poisson variates termed as Poisson difference (PD) distribution, given by

\[ f(d; m) = \sum_{x_1=0}^{\infty} \frac{\exp(-2m) m^{(2x_1+d)}}{x_1! (x_1 + d)!} = \exp(-2m)I_d(2m) \]  

where \( x_1 \) is a Poisson random variable with \( m \) being the parameter and \( d \) is the difference of the two Poisson variates which can take all integral values from \(-\infty\) to \(\infty\). \( I_d(2m) \) is Bessel’s modified function of the first kind of order \(d\) and argument \(2m\) as seen in Abramowitz and Stegun (1948). It is also shown that PD has a property of its even semi-invariants being equal and odd ones being zero.

• Infact, Skellam (1946) derived the pmf of the differences between two Poisson variates belonging to different populations with parameters \(m_1\) and \(m_2\) (by using the property, the probability of a variate difference of exactly \(r\) is the coefficient of \(t^r\) in \(G(t) = G_1(t)G_2(-t)\) where \(G_1(t)\) and \(G_2(-t)\) are the moment generating functions of the Poisson distributions with transformations \(m_1 = \frac{z}{\phi}\) and \(m_2 = \frac{z}{\phi}\) and \(\phi\) being a constant) and the pmf is given by

\[ P(r|m_1; m_2) = \exp \{- (m_1 + m_2) \} \left( \frac{m_1}{m_2} \right)^{r/2} I_r(2\sqrt{m_1m_2}) . \]

A special case of the same was derived when \(m_1 = m_2 = m\),

\[ P(r|m_1; m_2) = \exp(-2m) I_r(2m) , \]

which is same as the equation (1) derived by Irwin (1937).

• Castro (1952) also studied the difference of Poisson random variables of mean \(\lambda\) with pmf given by

\[ f(x; \lambda) = \exp(-2\lambda)I_x(2\lambda) \]

Prékopa (1952) discussed the case of PD distribution in terms of unequal parameters. Fisz (1953) proved that the standardized difference of two Poisson variates is asymptotically distributed in a \(N(0, 1)\), a table is presented which advocates that for fairly small Poisson parameters the normal approximation is rather good. Romani (1956) discussed the maximum likelihood estimate (MLE) of the population mean of difference between two Poisson variates, it was also seen that the odd cumulants of the PD distribution are \(\theta_1 - \theta_2\), and the even cumulants are equal to \(\theta_1 + \theta_2\).

• In Section 3 of Katti (1960), moments of the absolute differences for two independent Poisson random variables are derived. Strackee and van der Gon (1962) presented the tables of the cumulative density function of the PD distribution to four decimal places for some combinations of values of the two parameters. These tables also display the differences between the normal approximations seen in Fisz (1953).

• Karlis and Ntzoufras (2000) studied the properties and behaviour of the distribution of the difference of two Poisson (dependent or not) distributions as well as the distribution of the absolute Poisson difference (APD). The MLE of the PD is derived by the method
of Expectation Maximization (EM) algorithm. In Section 3 of Karlis and Ntzoufras (2000), the pmf of \( W \), (where \( W = |X - Y| \), \( X \) and \( Y \) being poisson variates with parameters \( \theta_1 \) and \( \theta_2 \)) following APD is derived, given by

\[
P(W = w) = \exp\left(-\left(\frac{\theta_1}{\theta_2} - 1\right)w\right) \left(\frac{\theta_1}{\theta_2}\right)^w \left(\frac{\theta_2}{\theta_1}\right)^w.
\]

- In Lemma 1 of Karlis and Ntzoufras (2006), it is proved that the pmf of the random variable \( Z = X - Y \), where \( (X, Y) \) jointly follows the bivariate Poisson distribution is given by

\[
f_{PD}(Z = z|\theta_1; \theta_2) = P(Z = z|\theta_1; \theta_2) = \exp\left(-\left(\frac{\theta_1}{\theta_2} - 1\right)z\right) \left(\frac{\theta_1}{\theta_2}\right)^z \left(\frac{\theta_2}{\theta_1}\right)^z,
\]

for all \( z \in \mathbb{Z} \) where \( I_r(x) \) is the modified Bessel’s function in \( r \), also formulated the bayesian estimates and used the bayesian approach to test the equality of the parameters of the PD distribution. Also an extended model named, the zero inflated Poisson difference (ZPD) distribution with the parameters being \( \theta_1 \) and \( \theta_2 \) is introduced in Section 4 and is given by,

\[
f_{ZPD}(Z = 0|p; \theta_1; \theta_2) = p + (1 - p) f_{PD}(Z = 0|\theta_1; \theta_2),
\]

\[
f_{ZPD}(Z = z|p; \theta_1; \theta_2) = (1 - p) f_{PD}(Z = z|\theta_1; \theta_2),
\]

for \( z \in \mathbb{Z} \); where \( p \in (0; 1) \) and \( f_{PD}(z|\theta_1; \theta_2) \) is given by equation (2). It has been studied to account for the excess of zero counts. Bayesian estimation and hypothesis testing of the model are also developed.

- Alzaid and Omair (2010) obtained inference results for the parameters of the PD distribution, maximum likelihood estimates of \( \theta_1 \) and \( \theta_2 \) were obtained by maximizing the likelihood function using the properties of modified Bessels function of first kind. For the purpose of simplification, alternative formulae for the PD distribution were formulated for which poisson distribution and negative of poisson distribution can be shown by direct replacement to be special cases of the PD distribution.

- Koopman, Lit, and Lucas (2014) presented a dynamic Skellam model with stochastically time varying intensities, where they replaced \( Y, y, \lambda_1 \), and \( \lambda_2 \) in

\[
P(Y = y|\theta_1; \theta_2) = \exp\left(-\left(\frac{\theta_1}{\theta_2} - 1\right)y\right) \left(\frac{\theta_1}{\theta_2}\right)^y \left(\frac{\theta_2}{\theta_1}\right)^y,
\]

by their time-varying counterparts \( Y_t, y_t, \lambda_{1t}, \) and \( \lambda_{2t} \), respectively and defined the dynamic model as

\[Y_t \sim Skellam(\lambda_{1t}, \lambda_{2t}); \ t = 1, \ldots n.\]

where \( n \) is the length of the time series with an assumption that the serial correlation in \( Y_t \) is accounted by the time-variation in the intensities \( \lambda_{1t} \) and \( \lambda_{2t} \), which is formulated in terms of a nonlinear non gaussian state space process and its basic properties are studied. They showed that the dynamic Skellam model in its state space formulation can handle larger sections of missing data. It is also observed that how the model can be extended to include regression effects, heterogeneous dynamics in the panel, and extensions of the Skellam distribution that assign different probability mass to a small number of discrete outcomes, an example of the latter is the dynamic zero inflated Skellam model.

- The zero modified Skellam distribution was introduced and its applications were studied by Conceição, Suzuki, and Andrade (2020).
2.2. Bivariate Skellam distribution

This section includes the works of researchers where bivariate Skellam distribution and its extensions are introduced.

- Genest and Mesfioui (2014) introduced two models of bivariate Skellam distribution which can be easily interpreted and simulated, which were based on a probabilistic construction involving a common shock, their basic features were studied and moment based estimators of their parameters were derived.

- Bulla, Chesneau, and Kachour (2015) defined the bivariate Skellam ($BSkellam(\lambda_0, \lambda_1, \lambda_2)$) distribution where $(X_1, X_2)$ are bivariate random variables with $X_1 \sim Skellam(\lambda_1, \lambda_0)$ and $X_2 \sim Skellam(\lambda_2, \lambda_0)$, the pmf of $BSkellam(\lambda_0, \lambda_1, \lambda_2)$ distribution is given as

$$P(X_1 = x_1, X_2 = x_2) = \exp(-\lambda_1 + \lambda_2 + \lambda_0)\lambda_1^{x_1}\lambda_2^{x_2} \sum_{i=max(0,-x_1,-x_2)}^{\infty} \frac{(\lambda_1\lambda_2\lambda_0)^i}{(x_1 + i)! (x_2 + i)!i!}$$

for all $(x_1, x_2) \in \mathbb{Z}^2$, $\lambda_0 \geq 0, \lambda_1 > 0, \lambda_2 > 0$ and its basic properties with methods of estimation are studied. Also, the extensions of the bivariate Skellam distribution, namely shifted $BSkellam$ distribution and mixture of $BSkellam$ distribution are studied.

- Aissaoui, Genest, and Mesfioui (2017) derived the limiting distribution of the moment estimators for the parameters of the models introduced by Genest and Mesfioui (2014) and their asymptotic efficiency is compared with the MLE of the parameter.

2.3. Skellam distribution in time series

In previous studies, the Skellam distribution is used with a static viewpoint. When we analyze the time series of differences in counts, we often obtain substantial improvements in model fit and in forecasting the performance when the parameters of the Skellam distribution are allowed to vary over time. Time variation in the parameters of the Skellam distribution may capture the developments of relative team strengths over elongated periods of time in sports, trends in health and demography in medical applications, or market environments and risk behaviours in economic and finance applications. Some of the notable contributions of Skellam distribution in time series analysis is given.

- Freeland (2010), Andersson and Karlis (2014) introduced a new integer-valued autoregressive (INAR) model and studied the case where the innovation has Skellam (PD) distribution, which allows negative values and negative correlations.

- Baroud (2011) developed a discrete-valued time series model, Pegram’s operator was used to build the autoregressive model of order $p$ (AR($p$)). The marginal distribution of the AR($p$), considered here is the Skellam distribution which is a symmetric distribution, due to which it can have both negative and positive correlations.

- Kerss, Leonenko, and Sikorskii (2014) studied the fractional Skellam processes, via the time changes in Poisson and Skellam processes by the inverse of a standard stable subordinator and the process is defined in Section 3. Alzaid and Omair (2014) introduced the Poisson difference integer valued AR(1) model to model integer valued time series with possible negative value and either positive or negative correlations.

- Alomani, Alzaid, and Omair (2018a) studied modeling nonstationary integer valued time series with conditional heteroskedasticity using Skellam distribution and the proposed model is applied to real time series. Alomani, Alzaid, and Omair (2018b) introduced an INGARCH(1,1) model based on the symmetric Skellam distribution, but they have only considered the simplest case, when the conditional expectation of process is equal to zero.

• Doukhan, Khan, and Neumann (2020) introduced the integer-valued generalized autoregressive conditional heteroskedasticity (GARCH) processes, where the count variable conditioned on past values of the count and state variables follow the Skellam distribution, using arguments for contractive Markov chains they proved that the process has a unique stationary regime. Gupta, Kumar, and Leonenko (2020) introduced the space-fractional Skellam process in Section 7.1 and tempered space-fractional Skellam Process in Section 7.2. Cui, Li, and Zhu (2020) extended the symmetric Skellam integrated generalized autoregressive conditional heteroskedasticity (INGARCH) model to an asymmetric version, which can deal with non-zero-mean processes, they also proposed a modified Skellam model which adopted a careful treatment on integers 0 and ±1 to satisfy a special feature of the data.

3. Applications of the Skellam distribution

The Skellam distribution has been applied in various fields, of which a few contributions are mentioned:

3.1. Finance

In financial markets, change in prices from one trade to the next can be modeled by the Skellam distribution, as prices move on a grid of integer multiples of a fixed ‘tick size’. The Skellam distribution is characterized by two ‘intensity’ parameters. Alzaid and Omair (2010) applied the PD distribution in the data from Saudi stock exchange, as they focussed only on studying the price formation during ordinary trading. Baroud (2011) used the Skellam AR model to analyse financial data. Koopman et al. (2014) analyzed long univariate high-frequency time series of U.S. stock price changes using the dynamic Skellam model, which evolved as discrete multiples of a fixed tick size of one dollar cent. Alzaid and Omair (2014) applied Poisson difference integer valued AR(1) to the data from Saudi stock exchange. Kerss et al. (2014) applied the fractional Skellam process in an empirical investigation of waiting times in high frequency financial data. Bulla et al. (2015) applied the bivariate Skellam distribution and its extensions to analyse Euro-Bund and Schatz future. Koopman, Lit, and Lucas (2017) studied the intraday stochastic volatility for four liquid stocks traded on the New York Stock Exchange using a new dynamic Skellam model for high-frequency tick-by-tick discrete price changes. Catania et al. (2019) applied the hierarchical HMM model for multivariate count data to the prices of stocks traded on NYSE and observed at very high frequencies (15 seconds). Lee (2019) considered the Skellam process to model financial high frequency data. Cui et al. (2020) applied the proposed models to a real time series data, the daily difference between the close and open prices of the Saudi Telecom in 2012.

3.2. Medicine

In medical research, experiments for measuring the effect of treatments and drug intake lead to paired counts. This is important since discrete distributions for such measurements are rare while such data usually arise in medicine as differences of discrete outcomes. The reason to which, Skellam distribution can be applied in medical research.
Karlis and Ntzoufras (2006) studied the effectiveness of the zero inflated Skellam distribution by applying them to a data taken from dental epidemiology. The application considered in the study by Alzaid and Omair (2010) used PD distribution to analyze the change in number of beds during 24 hours in NICU. Skellam distribution was used by Das Gupta, Srinivasa, Madhukara, and Antony (2015) in generating an erythema grading classification for erythema regions in psoriasis. Conceição et al. (2020) applied the zero modified Skellam distribution in an experiment in which 59 epileptic patients suffering from simple or complex partial seizures were randomized to receive either the antiepileptic drug (progabide) or a placebo, they also applied the zero modified Skellam to the difference in number of reported AIDS cases in the municipalities of Bahia 2015-2016.

3.3. Sports

In sports such as ice-hockey, handball and soccer, the score difference between the teams can be viewed as the difference between two Poisson counts and thus be modeled by a Skellam distributed random variable.

Karlis and Ntzoufras (2000) examined two distributions namely, the PD and the absolute PD distribution for modelling soccer data, these models can also be used for games where the number of goals scored is larger. Karlis and Ntzoufras (2003) modified the PD to model the goal difference of two opposing teams. Karlis and Ntzoufras (2009) applied the PD model in a data from the English premiership for the season 2006-2007. Koopman and Lit (2019) developed a multivariate model to analyse a high dimensional panel of football match results, wherein the score driven methodology is applied to three classes of models where one of them assumes the difference between the number of goals or the margin of victory of a team which is generated by Skellam distribution. Koopman et al. (2014) analyzed the score differences between rival soccer teams using a large, unbalanced panel of seven seasons of weekly matches in the German Bundesliga. Bulla et al. (2015) has applied the bivariate Skellam distribution and its extensions to the analysis of soccer results and also to analyze the score differences between rival soccer teams using a large, unbalanced panel of seven seasons of weekly matches in the German Bundesliga. Ntzoufras, Palaskas, and Drikos (2019) introduced a model by truncating the Skellam distribution to model the difference of sets in volleyball.

3.4. Science and technology

“In a steady state the number of light quanta, emitted or absorbed in a definite time, is distributed according to a Poisson distribution. In view thereof, the physical limit of perceptible contrast in vision can be studied in terms of the difference between two independent variates each following a Poisson distribution”. The distribution of differences may also be relevant when a physical effect is estimated as the difference between two counts, one when a “cause” is acting, and the other a “control” to estimate the “background effect”, as stated by Strackee and van der Gon (1962), Skellam distribution has been applied in the field of science in many ways:

Hwang, Kim, and Kweon (2007) introduced the Skellam distribution as a sensor noise model for CCD or CMS cameras, they showed that the distribution can be used to measure the intensity difference of pixels in spatial domain as well as in temporal domain. Hirakawa, Baqai, and Wolfe (2009) presented that certain wavelet and filter bank transform coefficients corresponding to vector valued measurements of this type are distributed as sums and differences of Poisson counts, taking the Skellam distribution. Alzaid and Omair (2014) defined the PD INAR model based on the extended binomial thinning operator. Chen, Zhang, and Zhao (2020) proposed a novel market model built upon the integer-valued Skellam distribution named Skellam market model. Skellam distribution was used by Choi, Park, and Kyung (2020) to estimate grey pixels.
4. Conclusion

In this review, we have discussed the Skellam distribution and provided additional literature. As it is the difference of two Poisson random variables and is defined on integers, it is more flexible to be applied to model real situation in a better manner than the basic distributions. These are also easily tractable and effective models for integer set data. Also, Skellam distributions provide applications in various areas like science, medicine, sports, time series modeling etc. We aspire this review may help to identify the gaps so that new distributions related to Skellam can be constructed.

References


