Advances and Applications in Statistical Sciences
Volume 7, Issue 3, March-April 2012, Pages 35-46

A DISTRIBUTION DESCRIBING DIFFERENCES IN COUNT DATA CONTAINING COMMON DISPERSION LEVELS

KIMBERLY F. SELLERS

Department of Mathematics and Statistics
Georgetown University
Washington, DC 20057, U. S. A.

Abstract

The Skellam distribution is a discrete probability distribution whose associated random variable is defined as the difference of two independent Poisson random variables with different corresponding expected values, λ_1 and λ_2. This manuscript generalizes the aforementioned framework by instead considering the difference of two independent Conway-Maxwell-Poisson (COM-Poisson) random variables with differing parameters, λ_1 and λ_2, and a common associated dispersion parameter, ν. The resulting distribution, which I name the Conway-Maxwell-Skellam (COM-Skellam) distribution, is then studied, highlighting its associated properties and use as an alternative means to study differences in count data.

1. Introduction

The Skellam distribution (also known as the Poisson difference distribution) is a discrete probability distribution that describes the distribution associated with a difference in counts. Derived from the difference of Poisson random variables, this distribution has broad applications, e.g. image denoising and detection [4-7], studying the effect of a treatment (e.g. Karlis and Ntzoufras [8]), or describing the point spread distribution in certain sports or games where the amount of points obtained per successful score is the same (e.g. soccer; see Karlis and Ntzoufras [9]). We describe this distribution in Section 1.1.

2010 Mathematics Subject Classification: 60E05, 60E10, 62F03, 62F10.

Keywords: count data, dispersion, generalized Poisson regression.

Received February 3, 2012

1.1. The Skellam distribution

Consider two independent Poisson random variables, $Y_i \sim$ Poisson (λ_i), $i = 1, 2$, and let $S = Y_1 - Y_2$. Accordingly, S has the probability mass function (pmf),

$$f(s) = e^{-(\lambda_1+\lambda_2)}\left(\frac{\lambda_1}{\lambda_2}\right)^{s/2} I_{|s|}(2\sqrt{\lambda_1\lambda_2}),$$

where s is an integer, and $I_\alpha(z) \doteq \sum_{m=0}^{\infty} \frac{1}{[\Gamma(m+\alpha+1)m!]}\left(\frac{z}{2}\right)^{2m+\alpha}$ denotes a modified Bessel function of the first kind; see Skellam [12]. This is the pmf of a Skellam distribution with parameters λ_1 and λ_2, and we denote the Skellam random variable accordingly as $S \sim$ Skellam (λ_1, λ_2).

The Skellam distribution holds numerous properties; see Alzaid and Omair [1] for a comprehensive listing. Romani [10] showed that all odd cumulants of a Skellam distribution equal $\lambda_1 - \lambda_2$, while all even cumulants equal $\lambda_1 + \lambda_2$. In particular, $E(S) = \lambda_1 - \lambda_2$ and $\text{Var}(S) = \lambda_1 + \lambda_2$. For large $\lambda_1 + \lambda_2$, the Skellam distribution can be approximated by a normal distribution [8]. Various statistical computations regarding the Skellam distribution can be performed via the `skellam` package in *R*.

1.2. Motivation and outline of the paper

Because the Skellam distribution is derived from the difference of Poisson random variables, the associated inference is gained under the constraining assumption that the underlying count distributions are equi-dispersed. Many references (e.g., Barron [2]) however note that real data oftentimes exhibit some form of under- or over-dispersion. Karlis and Ntzoufras [9] note that, for the case of modeling the difference in football/soccer scoring, while "the Poisson distribution has been widely used as a simple modelling approach for describing the number of goals in football, this assumption can be questionable in certain leagues where over-dispersion (sample variance exceeds the sample mean) has been observed in the number of goals." Thus, we must consider a generalized count distribution that allows for comparing counts in light of associated dispersion. To date, there does not appear to be an extension of the Skellam distribution that allows for over- or under-dispersion in the dataset.

The Conway-Maxwell-Poisson (COM-Poisson) distribution is a generalization of the Poisson distribution that can better accommodate data dispersion. This distribution is described in greater detail in Section 2. We then introduce (in Section 3) the difference of two COM-Poisson distributions to derive the Conway-Maxwell-Skellam (COM-Skellam) distribution as a means for comparing associated counts from two distributions that exhibit the same level of data dispersion. In particular, Section 3.2 describes how to determine the associated maximum likelihood estimators for this distribution, and Section 3.3 considers two hypothesis tests of interest. We apply the statistical ideas and concepts developed in the previous sections to an example data set in Section 4, and conclude the manuscript with discussion in Section 5.

2. The Conway-Maxwell-Poisson Distribution

In order to introduce the COM-Skellam distribution, we must first provide background regarding its associated foundational distribution: the COM-Poisson distribution. The COM-Poisson distribution (introduced by Conway and Maxwell [3], and revived by Shmueli et al. [11]) is a viable count distribution that generalizes the Poisson distribution in light of associated data dispersion. The COM-Poisson pmf takes the form

$$P(Y = y \mid \lambda, \nu) = \frac{\lambda^y}{(y!)^\nu Z(\lambda, \nu)}, \quad y = 0, 1, 2, \ldots, \tag{1}$$

for a random variable Y, where $\lambda = E(Y^\nu)$, ν is an associated dispersion parameter, and $Z(\lambda, \nu) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}$ is a normalizing constant. The COM-Poisson distribution includes three well-known distributions as special cases: Poisson ($\nu = 1$), geometric ($\nu = 0$, $\lambda < 1$), and Bernoulli ($\nu \to \infty$ with probability $\frac{\lambda}{1+\lambda}$).

In Shmueli et al. [11], the moments are given in the form

$$\mathrm{E}(Y^{r+1}) = \begin{cases} \lambda[\mathrm{E}(Y+1)]^{1-\nu}, & r = 0, \\ \lambda \frac{\partial}{\partial \lambda} \mathrm{E}(Y^r) + \mathrm{E}(Y)\mathrm{E}(Y^r), & r > 0, \end{cases} \tag{2}$$

and the expected value is approximated by

$$\mathrm{E}(Y) = \lambda \frac{\partial \log Z(\lambda, \nu)}{\partial \lambda} \approx \lambda^{1/\nu} - \frac{\nu - 1}{2\nu}, \tag{3}$$

which particularly holds for $\nu \leq 1$ or $\lambda > 10^{\nu}$. Note that the expected value and variance can also be written in the form

$$\mathrm{E}(Y) = \frac{\partial \log Z(\lambda, \nu)}{\partial \log \lambda}, \tag{4}$$

$$\mathrm{Var}\,(Y) = \frac{\partial \mathrm{E}(Y)}{\partial \log \lambda}. \tag{5}$$

Finally, the associated moment generating function of Y is $M_Y(t) = \mathrm{E}(e^{Yt}) = \frac{Z(\lambda e^t, \nu)}{Z(\lambda, \nu)}$, and its probability generating function is $\mathrm{E}(t^Y) = \frac{Z(\lambda t, \nu)}{Z(\lambda, \nu)}$.

3. The Conway-Maxwell-Skellam Probability Distribution and its Statistical Properties

3.1. The Conway-Maxwell-Skellam (COM-Skellam) probability mass function

Suppose that, instead of defining S as in Section 1.1, we generalize its definition such that $S = Y_1 - Y_2$ is the difference of the two independent COM-Poisson random variables, each containing the same level of dispersion ν but allowing for different values of λ_i, $i = 1, 2$. Accordingly, the pmf of the COM-Skellam random variable, $S \sim$ COM-Skellam $(\lambda_1, \lambda_2, \nu)$, is

$$P(S = s) = \frac{1}{Z(\lambda_1, \nu)Z(\lambda_2, \nu)}\left(\frac{\lambda_1}{\lambda_2}\right)^{s/2} I_{|s|}^{(\nu)}(2\sqrt{\lambda_1\lambda_2}), \quad s \in \mathbb{Z}. \tag{6}$$

where $I_{\alpha}^{(\nu)}(z) \doteq \sum_{m=0}^{\infty} \frac{1}{[\Gamma(m + \alpha + 1)m!]^{\nu}}\left(\frac{z}{2}\right)^{2m+\alpha}$ is a generalized form of the modified Bessel function of the first kind; see the supplementary material for the derivation of the COM-Skellam distribution. For the special case where $\nu = 1$, $I_{\alpha}^{(1)}(z) \doteq I_{\alpha}(z)$ is precisely a modified Bessel function of the first kind.

Just as the COM-Poisson distribution reduces to the special cases of the Poisson, geometric, and Bernoulli distributions, we can likewise consider analogous cases associated with the COM-Skellam distribution. For the special case where $\nu = 1$, the COM-Skellam distribution reduces to the Skellam distribution as described in Section 1.1. When $\nu = 0$ and $\lambda_i < 1$ for $i = 1, 2$ (i.e. where $S = Y_1 - Y_2$ is the difference of two independent geometric distributions),

$$P(S = s) = \frac{\lambda_1^s (1 - \lambda_1)(1 - \lambda_2)}{1 - \lambda_1 \lambda_2}, \quad s \in \mathbb{Z};$$

for the special case where $\lambda_1 = \lambda_2 = \lambda$, we have $P(S = s) = \dfrac{\lambda^s (1 - \lambda)}{1 + \lambda}$, $s \in \mathbb{Z}$. Finally, as $\nu \to \infty$, the COM-Skellam distribution approaches a distribution whose probability mass function is

$$P(S = s; \lambda_1, \lambda_2) = \begin{cases} \dfrac{\lambda_2}{(1 + \lambda_1)(1 + \lambda_2)} = (1 - p_1)p_2, & \text{if } s = -1, \\ \dfrac{1 + \lambda_1 \lambda_2}{(1 + \lambda_1)(1 + \lambda_2)} = (1 - p_1)(1 - p_2) + p_1 p_2, & \text{if } s = 0, \\ \dfrac{\lambda_1}{(1 + \lambda_1)(1 + \lambda_2)} = p_1(1 - p_2), & \text{if } s = 1, \end{cases}$$

which is the difference of two independent Bernoulli $\left(p_i = \dfrac{\lambda_i}{1 + \lambda_i}\right)$ distributions.

Properties of the COM-Skellam distribution include an interesting pseudo-symmetrical result, $P(S = s; \lambda_1, \lambda_2) = P(S = -s; \lambda_2, \lambda_1)$. The mean of the COM-Skellam distribution can be determined via

$E(S) = E(Y_1) - E(Y_2) \approx \lambda_1^{1/\nu} - \lambda_2^{1/\nu}$ by Equation (3) for $\nu \le 1$ or $\lambda > 10^\nu$, while the variance is $\mathrm{Var}(S) = \mathrm{Var}(Y_1) + \mathrm{Var}(Y_2)$ by independence of Y_i, $i = 1, 2$, where the variance can be computed via Equation (5). For the case where $\nu = 1$, the respective results for the expectation and variance simplify to that for a Skellam distribution. The moment generating function (mgf) of S is

$$M_S(t) = M_{Y_1}(t) M_{Y_2}(-t) \text{ by independence of } Y_1, Y_2$$

$$= \frac{Z(\lambda_1 e^t, \nu)}{Z(\lambda_1, \nu)} \cdot \frac{Z(\lambda_2 e^{-t}, \nu)}{Z(\lambda_2, \nu)}$$

and the probability generating function is

$$E(t^S) = E(t^{Y_1})E(t^{-Y_2}) \text{ by independence of } Y_1, Y_2$$
$$= \frac{Z(\lambda_1 t, \nu)}{Z(\lambda_1, \nu)} \cdot \frac{Z(\lambda_2 / t, \nu)}{Z(\lambda_2, \nu)}.$$

3.2. Estimating the Conway-Maxwell-Skellam parameters

Romani [10] considered the properties of the maximum likelihood estimator of the Skellam expected value, $E(S) = \lambda_1 - \lambda_2$. Here, we consider the analogous question, determining the maximum likelihood estimators for the COM-Skellam parameters. Given the pmf as defined in Equation (6), we can use maximum likelihood estimation to determine the associated estimates for λ_1, λ_2 and ν, respectively. Accordingly, we write the log-likelihood as

$$\log L(\lambda_1, \lambda_2, \nu \mid s) = -n[\log Z(\lambda_1, \nu) + \log Z(\lambda_2, \nu)] + \frac{\sum_{i=1}^{n} s_i}{2}(\log \lambda_1 - \log \lambda_2)$$
$$+ \sum_{i=1}^{n} \log I_{|s_i|}^{\nu}(2\sqrt{\lambda_1 \lambda_2}),$$

and solve it numerically through a computational software tool (in this case, R via the commands, `nlminb` or `optim`), thus determining the maximum likelihood estimates. Further, we can determine the associated standard errors via the corresponding Information matrix.

3.3. Hypothesis testing

Two interesting questions (and, therefore, hypothesis tests) arise from consideration of a COM-Skellam distribution. The first question asks if the data dispersion existing in the dataset is signficant enough that one must use the COM-Skellam distribution, as opposed to the Skellam distribution, to model count differences. The second question compares the resulting means of the two COM-Poisson distributions, i.e. assessing whether or not the mean of the COM-Skellam distribution is approximately zero or otherwise.

To address the first question, we perform the hypothesis test about the dispersion parameter,

$$H_0 : \nu = 1 \text{ versus } H_1 : \nu \neq 1.$$

This hypothesis test determines whether the amount of data dispersion warrants the need to use the more general COM-Skellam distribution to compare differences in the count data, or if the Skellam distribution is sufficient for such data comparisons. Accordingly, for this hypothesis test, we consider the likelihood ratio statistic,

$$\Lambda_\nu = \frac{L(\hat{\lambda}_{1,H_0};\ \hat{\lambda}_{2,H_0};\ \nu_{H_0} = 1)}{L(\hat{\lambda}_1,\ \hat{\lambda}_2,\ \hat{\nu})},$$

where $-2\log\Lambda_\nu \to \chi_1^2$. Note that this two-sided test does not directly identify the direction of dispersion (i.e., over or under); however, the test result, along with the maximum likelihood estimate for ν, will provide sufficient information about the dispersion type.

The COM-Skellam distribution can be used as a tool to compare two count datasets containing data dispersion. In particular, one can consider a hypothesis test that compares the respective means from the two groups modeled by the COM-Poisson distribution. The expected value of the COM-Poisson distribution is difficult to describe, given its nonlinear form. We can, however, compare the expected values of two independent COM-Poisson distributions in the following manner. We can consider the hypothesis test in reference to the associated COM-Skellam distribution, namely

$$H_0 : \mu_S = \mathrm{E}(S) = 0 \text{ versus } H_1 : \mu_S = \mathrm{E}(S) \neq 0.$$

Because we assume a common dispersion level (ν) for this framework, this hypothesis test is equivalent to $H_0 : \lambda_1 = \lambda_2$ versus $H_1 : \lambda_1 \neq \lambda_2$. Thus, for this hypothesis test, we consider the associated likelihood ratio statistic,

$$\Lambda_{\mu_S} = \frac{L(\hat{\lambda}_{1,H_0} = \hat{\lambda}_{2,H_0} = \hat{\lambda};\ \hat{\nu}_{H_0})}{L(\hat{\lambda}_1,\ \hat{\lambda}_2,\ \hat{\nu})},$$

where $-2\log\Lambda_{\mu_S} \to \chi_1^2$. While this test is described in a two-sided form, we can analogously consider one-sided tests to assess directional performance of the associated data.

4. Example: Soccer Scoring

Karlis and Ntzoufras [9] used Bayesian modeling of the Skellam distribution to model the difference in the number of goals in soccer matches. In this work, they use data from the English Premiership's 2006-2007 soccer season; Table 4 provides the team names and their observed points scored and corresponding goal difference for each of the 20 teams.

The maximum likelihood estimates assuming a Skellam model are $\hat{\lambda}_1 = \hat{\lambda}_2 = 268.3215$, while the COM-Skellam maximum likelihood estimates are $\hat{\lambda}_1 = 1.9480$, $\hat{\lambda}_2 = 1.9213$, and $\hat{\nu} = 1.5794$. While we cannot directly compare the estimates associated with the two distributions, we can draw some interesting inferences from this information. The resulting COM-Skellam dispersion estimate (i.e. $\hat{\nu} = 1.5794$) implies possible data under-dispersion in this dataset. Performing the associated hypothesis test to determine if this level of dispersion is significant, we obtain $-2 \log \Lambda = 850.7189$ (p-value $= 5.0732 \times 10^{-187}$); thus we see that the amount of dispersion in this dataset is statistically significantly under-dispersed. Thus, using a Skellam distribution to model the data is inappropriate here, because of the constraining equidispersion assumption of the Poisson distribution. Meanwhile, if we compare the COM-Skellam parameters, λ_1 and λ_2, the associated hypothesis test concludes that the parameter estimates are not statistically significantly different $-2 \log \Lambda = 0.2119$ with associated p-value $= 0.6453$), implying that the teams are reasonably matched together as a league.

5. Discussion

The Skellam distribution is a handy distribution to study differences in count datasets, yet is restricted to the underlying assumption that the respective datasets contain equal mean and variance. The development of the Conway-Maxwell-Skellam (COM-Skellam) distribution relaxes this assumption for broader comparison of count distributions when a common dispersion level is present among the two datasets.

Table 1. Original data from English Premiership's 2006-2007 soccer season, reproduced from Karlis and Ntzoufras [9].

Team Name	Points	Goal Difference
Man Utd	89	56
Chelsea	83	40
Liverpool	68	30
Arsenal	68	28
Tottenham	60	3
Everton	58	16
Bolton	56	−5
Reading	55	5
Portsmouth	54	3
Blackburn	52	−2
Aston Villa	50	2
Middlesbrough	46	−5
Newcastle	43	−9
Man City	42	−15
West Ham	41	−24
Fulham	39	−22
Wigan	38	−22
Sheff Utd	38	−23
Charlton	34	−26
Watford	28	−30

Future work will further generalize these ideas to consider comparing datasets containing differing dispersion levels and this assumption's impact on the development of a generalized form of the COM-Skellam distribution.

Acknowledgments

The author thanks David Caraballo and Howard Seltman for insightful discussion.

References

[1] A. A. Alzaid and M. A. Omair, On the Poisson difference distribution inference and applications, Bull. Malaysian Math. Sci. Soc. 33(1) (2010), 17-45.

[2] D. N. Barron, The analysis of count data: over-dispersion and autocorrelation, Sociol. Method. 22 (1992), 179-220.

[3] R. W. Conway and W. L. Maxwell, A queuing model with state dependent service rates, J. Ind. Eng. 12 (1962), 132-136.

[4] K. Hirakawa, F. Baqai and P. J. Wolfe, Wavelet-based Poisson Rate Estimation using the Skellam Distribution, in C. A. Bouman, E. L. Miller and I. Pollak, Editors, Proceedings of SPIE 7246 (2009), 72460R-72460R-12.

[5] K. Hirakawa and P. J. Wolfe, Skellam shrinkage: Wavelet-based intensity estimation for inhomogeneous Poisson data, (2009), arXiv:0905.217.

[6] Y. Hwang, H. Joo, J. S. Kim and I. S. Kweon, Statistical Background Subtraction based on the Exact Per-Pixel Distributions, in MVA2007 IAPR Conference on Machine Vision Applications Tokyo, Japan, May 16-18, 2007a, pp. 540-543.

[7] Y. Hwang, I. S. Kweon and J. S. Kim, Color Edge Detection using the Skellam Distribution as a Sensor Noise Model, in SICE Annual Conference 2007, Kagawa University, Japan, 2007b, 1972-1979.

[8] D. Karlis and I. Ntzoufras, Bayesian analysis of the differences of count data, Stat. Med. 25 (2006), 1885-1905.

[9] D. Karlis and I. Ntzoufras, Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference, IMA J. Manag. Math. 20(2) (2008), 133-145.

[10] J. Romani, Distribution of the algebraic sum of Poisson variables, Trabajos de Estadistica 7 (1956), 175-181.

[11] G. Shmueli, T. P. Minka, J. B. Kadane, S. Borle and P. Boatwright, A useful distribution for fitting discrete data: revival of the Conway-Maxwell-Poisson distribution, App. Stat. 54 (2005), 127-142.

[12] J. G. Skellam, The frequency distribution of the difference between two Poisson variates belonging to different populations, J. Royal Stat. Soc. 109(3) (1946), 296.

A. Derivation of COM-Skellam Probability Mass Function

The probability mass function for the COM-Skellam random variable S, is derived as follows. For $s \in \mathbb{Z}$,

$$P(S = s) = \sum_{y_2} P(S = s \mid Y_2 = y_2)P(Y_2 = y_2)$$

$$= \sum_{y_2 = \max(0, -s)}^{\infty} P(Y_1 = s + y_2)P(Y_2 = y_2)$$

$$= \sum_{y_2=\max(0,-s)}^{\infty} \left(\frac{\lambda_1^{s+y_2}}{[(s+y_2)!]^{\nu} Z(\lambda_1, \nu)} \right) \left(\frac{\lambda_2^{y_2}}{(y_2!)^{\nu} Z(\lambda_2, \nu)} \right)$$

$$= \frac{\lambda_1^s}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \sum_{y_2=\max(0,-s)}^{\infty} \frac{(\lambda_1 \lambda_2)^{y_2}}{[(s+y_2)! \, y_2!]^{\nu}}$$

$$= \frac{1}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \frac{1}{(\sqrt{\lambda_1 \lambda_2})^s} \sum_{y_2=\max(0,-s)}^{\infty} \frac{\left(\frac{2\sqrt{\lambda_1 \lambda_2}}{2} \right)^{2y_2+s}}{[(s+y_2)! y_2!]^{\nu}}$$

$$= \frac{1}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \left(\frac{\lambda_1}{\lambda_2} \right)^{s/2} \sum_{y_2=\max(0,-s)}^{\infty} \frac{\left(\frac{2\sqrt{\lambda_1 \lambda_2}}{2} \right)^{2y_2+s}}{[(s+y_2)! y_2!]^{\nu}}. \quad (7)$$

For the case where $s \geq 0$, Equation (7) becomes

$$P(S = s) = \frac{1}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \left(\frac{\lambda_1}{\lambda_2} \right)^{s/2} \sum_{y_2=0}^{\infty} \frac{\left(\frac{2\sqrt{\lambda_1 \lambda_2}}{2} \right)^{2y_2+s}}{[\Gamma(s+y_2+1) y_2!]^{\nu}}$$

$$= \frac{1}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \left(\frac{\lambda_1}{\lambda_2} \right)^{s/2} I_s^{(\nu)}(2\sqrt{\lambda_1 \lambda_2}),$$

where we define $I_{\alpha}^{(\nu)}(z) \doteq \sum_{m=0}^{\infty} \frac{1}{[\Gamma(m+\alpha+1) m!]^{\nu}} \left(\frac{z}{2} \right)^{2m+\alpha}$ as a generalized form of the modified Bessel function of the first kind. Meanwhile, for the case where $s < 0$, Equation (7) becomes

$$P(S = s) = \frac{1}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \left(\frac{\lambda_1}{\lambda_2} \right)^{s/2} \sum_{y_2=-s}^{\infty} \frac{\left(\frac{2\sqrt{\lambda_1 \lambda_2}}{2} \right)^{2y_2+s}}{[(s+y_2)! y_2!]^{\nu}},$$

where shifting the index in the summation (e.g. $k = y + s$) produces

$$P(S = s) = \frac{1}{Z(\lambda_1, \nu) Z(\lambda_2, \nu)} \left(\frac{\lambda_1}{\lambda_2} \right)^{s/2} \sum_{k=0}^{\infty} \frac{\left(\frac{2\sqrt{\lambda_1 \lambda_2}}{2} \right)^{2k-s}}{[k!(k-s)!]^{\nu}}$$

$$= \frac{1}{Z(\lambda_1, \nu)Z(\lambda_2, \nu)}\left(\frac{\lambda_1}{\lambda_2}\right)^{s/2} \sum_{k=0}^{\infty} \frac{\left(\frac{2\sqrt{\lambda_1\lambda_2}}{2}\right)^{2k-s}}{[\Gamma(k-s+1)k!]^{\nu}}$$

$$= \frac{1}{Z(\lambda_1, \nu)Z(\lambda_2, \nu)}\left(\frac{\lambda_1}{\lambda_2}\right)^{s/2} I_{-s}^{(\nu)}(2\sqrt{\lambda_1\lambda_2}),$$

because $-s \geq 0$. Thus, the probability mass function of the COM-Skellam random variable S, is

$$P(S = s) = \frac{1}{Z(\lambda_1, \nu)Z(\lambda_2, \nu)}\left(\frac{\lambda_1}{\lambda_2}\right)^{s/2} I_{|s|}^{(\nu)}(2\sqrt{\lambda_1\lambda_2}), \quad s \in \mathbb{Z}. \tag{8}$$