Random variate generation and connected computational issues for the Poisson–Tweedie distribution

Random variate generation and connected computational issues for the Poisson-Tweedie distribution

A. Baccini[1], L. Barabesi[1], L. Stracqualursi[2]
[1]Department of Economics and Statistics, University of Siena,
P.zza S.Francesco 7, 53100 Siena, Italy
[2]Department of Statistical Sciences "Paolo Fortunati", University of Bologna,
Via Belle Arti 41, 40126 Bologna, Italy

Abstract

After providing a systematic outline of the stochastic genesis of the Poisson-Tweedie distribution, some computational issues are considered. More specifically, we introduce a closed form for the probability function, as well as its corresponding integral representation which may be useful for large argument values. Several algorithms for generating Poisson-Tweedie random variates are also suggested. Finally, count data connected to the citation profiles of two statistical journals are modeled and analyzed by means of the Poisson-Tweedie distribution.

Key words. Mixture Poisson distribution; Compound Poisson distribution; Random variate generation; Journal citation profile data.

1. Introduction. The Poisson-Tweedie integer-valued distribution has been introduced independently by Gerber (1991) - as the generalized Negative Binomial distribution - and Hougaard *et al.* (1997) - as the P-G distribution (where the acronym emphasizes its stochastic representation as a mixture Poisson, see Section 2 for more details). At first, the distribution seems to be named 'Poisson-Tweedie' by Kokonendji *et al.* (2004) - even if Johnson *et al.* (2005, p.480) refer to it as 'Tweedie-Poisson' in their encyclopedia. The denomination is obviously due to the strong connection of this integer-valued law with the absolutely-continuous Tweedie distribution introduced by Hougaard (1986) on the basis of the seminal proposal by Tweedie (1984) - see Section 2 of the present paper. The Poisson-Tweedie is a very flexible model and contains as special cases classical families such as the Poisson and the Negative Binomial, as well as large families such as the Generalized Poisson Inverse Gaussian and the Poisson-Pascal - and even the Discrete Stable family (see *e.g.* El-Shaarawi *et al.*, 2011).

By following the notation adopted by El-Shaarawi *et al.* (2011), the probability generating function (p.g.f.) of the Poisson-Tweedie random variable (r.v.) X_{PT} is given by

$$G_{X_{\mathrm{PT}}}(s) = e^{\frac{b}{a}[(1-c)^a-(1-cs)^a]}\ , \tag{1}$$

where $(a,b,c) \in \{]-\infty,0] \times]0,\infty[\times [0,1[\} \cup \{]0,1] \times]0,\infty[\times [0,1]\}$. It should be remarked that the case $a=0$ may be managed for analytical continuation. In the following, the Poisson-Tweedie r.v. with p.g.f. given by (1) is eventually denoted as $\mathcal{PT}(a,b,c)$ for convenience. The probability function (p.f.) of the r.v. X_{PT} is usually computed by means of the recursive algorithm given by El-Shaarawi *et al.* (2011). Many properties of the

$$G_{X_{\mathrm{SI}}}(s) = 1 - (1-s)^\gamma \ ,$$

where $\gamma \in \,]0,1]$. The Sibuya distribution is a special case of the (shifted) Negative Binomial Beta distribution introduced by Sibuya (1979) with parameters given by 1, γ and $(1-\gamma)$. In the following, the Sibuya r.v. is also denoted by $\mathcal{SI}(\gamma)$. On the basis of the findings by Sibuya (1979), if $\mathcal{B}(\phi, \varphi)$ represents a Beta r.v. with shape parameters ϕ and φ, the Sibuya r.v. has the following stochastic representation

$$\mathcal{SI}(\gamma) \stackrel{\mathcal{L}}{=} 1 + \mathcal{NB}(1, \mathcal{B}(1-\gamma, \gamma)) \stackrel{\mathcal{L}}{=} 1 + \mathcal{P}(\mathcal{G}(1,1)\mathcal{G}(1-\gamma,1)/\mathcal{G}(\gamma,1)) \ , \tag{6}$$

where the Exponential and the two Gamma r.v.'s involved in the previous expression are assumed to be independent. In this case, the geometric down-weighting Sibuya r.v. X_{DSI} displays the p.g.f.

$$G_{X_{\mathrm{DSI}}}(s) = 1 - G_{X_{\mathrm{SI}}}(\beta) + G_{X_{\mathrm{SI}}}(\beta s) = 1 + (1-\beta)^\gamma - (1-\beta s)^\gamma \ ,$$

where $(\gamma, \beta) \in \{]0,1] \times]0,1]\}$ (for more details, see Zhu and Joe, 2009). The geometric down-weighting Sibuya r.v. is also denoted by $\mathcal{DSI}(\gamma, \beta)$. Therefore, if $\mathcal{U}$ represents a Uniform r.v. on $[0,1]$, the following representation holds

$$\mathcal{DSI}(\gamma, \beta) \stackrel{\mathcal{L}}{=} I_{\mathbb{R}^+}(\beta^{\mathcal{SI}(\gamma)} - \mathcal{U})\mathcal{SI}(\gamma) \ , \tag{7}$$

where the r.v. $\mathcal{U}$ is independent of the r.v. $\mathcal{SI}(\gamma)$, while I_B is the usual indicator function of a set B. Finally, by reparametrizing in such a way that $\gamma = a$ and $\beta = c$, from (1) it turns out that

$$G_{X_{\mathrm{PT}}}(s) = e^{\frac{b}{a}[(1+(1-c)^a - (1-cs)^a) - 1]} \ ,$$

i.e. a Poisson compounding of a geometric down-weighting Sibuya r.v. is actually achieved. Hence, the stochastic representation holds

$$\mathcal{PT}(a,b,c) \stackrel{\mathcal{L}}{=} \sum_{i=1}^{\mathcal{P}(b/a)} \mathcal{DSI}_i(a,c) \ , \tag{8}$$

where the $\mathcal{DSI}_i(a)$'s are *i.i.d.* geometric down-weighting Sibuya r.v.'s, which are in turn independent of $\mathcal{P}(b/a)$.

When $a \in \,]-\infty, 0[$, from (1) it promptly follows that

$$G_{X_{\mathrm{PT}}}(s) = e^{-\frac{b(1-c)^a}{a}\left[\left(\frac{1-cs}{1-c}\right)^a - 1\right]} \ ,$$

i.e. a Poisson compounding of a Negative Binomial r.v. with parameters $(-a)$ and c is obtained. Hence, the following representation holds

$$\mathcal{PT}(a,b,c) \stackrel{\mathcal{L}}{=} \sum_{i=1}^{\mathcal{P}(-b(1-c)^a/a)} \mathcal{NB}_i(-a,c) \ ,$$

where the $\mathcal{NB}_i(-a,c)$'s are *i.i.d.* Negative Binomial r.v.'s, which are in turn independent of $\mathcal{P}(-b(1-c)^a/a)$. Owing to the reproductive property of the Negative Binomial, the previous expression reduces to

$$\mathcal{PT}(a,b,c) \stackrel{\mathcal{L}}{=} \mathcal{NB}(-a\mathcal{P}(-b(1-c)^a/a), c) \ ,$$

which is stochastically equivalent to (4) by considering (5). Finally, when $a=0$ the representation (5) is in turn achieved.

3. Easy-computable expressions for the p.f. First, it is worth noting that the p.f. corresponding to the p.g.f. (1) may be obtained as a finite sum. Indeed, from Result 1 in the Appendix, it turns out to be

$$p_{X_{\mathrm{PT}}}(k) = e^{\frac{b}{a}[(1-c)^a-1]}(-c)^k \sum_{m=0}^{k} \frac{(b/a)^m}{m!} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \binom{aj}{k} I_{\mathbb{N}}(k) \, . \qquad (9)$$

In addition, from Result 2 in the Appendix, for $a \in \,]0,1]$ it also follows that

$$p_{X_{\mathrm{PT}}}(k) \leq \frac{b}{a}\left(1+\frac{b}{a}\right) e^{\frac{b}{a}[(1-c)^a+1]} c^k k^{-a-1}.$$

Moreover, by adopting the expression given in Result 2 for $p_{X_{\mathrm{PT}}}$, for a fixed k_* the following simple approximation of $p_{X_{\mathrm{PT}}}$ holds

$$p^*_{X_{\mathrm{PT}}}(k) = e^{\frac{b}{a}(1-c)^a}(-c)^k \sum_{m=0}^{k^*} (-1)^m \binom{am}{k} \frac{(b/a)^m}{m!} \, . \qquad (10)$$

We have numerically assessed that $k_* = 2$ usually suffices for obtaining an adequate approximation for a large k - which may avoid the computational burden involved in the evaluation of (9) in such a case.

Incidentally, it is interesting to remark that for $a \in \,]0,1]$ the Poisson-Tweedie r.v. may be rephrased as an exponentially-tilted Discrete Stable r.v. - *i.e.* the integer-valued counterpart of an exponentially-tilted Stable r.v. Indeed, let us notice that the p.g.f. of the Discrete Stable r.v. X_{DS} of parameters γ and λ is given by

$$G_{X_{\mathrm{DS}}}(s) = e^{-\lambda(1-s)^\gamma} \, ,$$

where $(\gamma,\lambda) \in \{]0,1] \times]0,\infty[\}$ (for more details on this heavy-tailed distribution, see *e.g.* Marcheselli *et al.*, 2008). In the following, this r.v. is also denoted as $\mathcal{DS}(\gamma,\lambda)$. Hence, by reparametrizing in such a way that $\gamma = a$ and $\lambda = b/a$, on the basis of expression (1) it follows that

$$G_{X_{\mathrm{PT}}}(s) = \frac{G_{X_{\mathrm{DS}}}(cs)}{G_{X_{\mathrm{DS}}}(c)} \, .$$

Thus, if $p_{X_{\mathrm{DS}}}$ represents the p.f. of the Discrete Stable r.v. it turns out that

$$p_{X_{\mathrm{PT}}}(k) = e^{\frac{b}{a}(1-c)^a} c^k p_{X_{\mathrm{DS}}}(k) \, , \qquad (11)$$

i.e. an exponentially-tilted Discrete Stable r.v. with tilting parameter c is actually achieved. From expressions (9) and (11), a closed form for the p.f. of the Discrete Stable r.v. $\mathcal{DS}(\gamma,\lambda)$ can be obtained as a by-product, *i.e.*

$$p_{X_{\mathrm{DS}}}(k) = e^{-\frac{b}{a}}(-1)^k \sum_{m=0}^{k} \frac{(b/a)^m}{m!} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \binom{aj}{k} I_{\mathbb{N}}(k) \, .$$

For further discussion of integer-valued exponentially-tilted distributions related to the Discrete Stable distribution see Barabesi and Pratelli (2014a).

As previously remarked, expression (9) is not obviously convenient for large values of k, even if its computation solely requires a finite summation. The recursive expression for $p_{X_{\rm PT}}$ provided by El-Shaarawi *et al.* (2011) actually involves the same drawback. In such a case, the p.f. $p_{X_{\rm PT}}$ may be alternatively computed by adopting a generalization of the Inversion Theorem. Indeed, by following Barabesi and Pratelli (2014b), if X is an integer-valued r.v. with p.f. p_X and g is a measurable function defined on $\mathbb{Z}$ such that $\mathrm{E}[|g(X)|] < \infty$, then it holds

$$p_X(k) = \frac{1}{2\pi g(k)} \int_{-\pi}^{\pi} e^{-\mathbf{i}tk} \mathrm{E}[g(X) e^{\mathbf{i}tX}] \, \mathrm{d}t$$

for $k \in \mathbb{Z}$ and $g(k) \neq 0$ and where $\mathbf{i}$ represents the imaginary unit. In the case of the Poisson-Tweedie distribution, by selecting $g(k) = q^k$ for a given $q \in]0, 1/c[$, the previous expression gives rise to

$$\begin{aligned} p_{X_{\rm PT}}(k) &= \frac{e^{\frac{b}{a}(1-c)^a}}{2\pi} q^{-k} \int_{-\pi}^{\pi} e^{-\mathbf{i}tk - \frac{b}{a}(1-cqe^{\mathbf{i}t})^a} \, \mathrm{d}t \\ &= \frac{e^{\frac{b}{a}(1-c)^a}}{\pi} q^{-k} \int_0^{\pi} e^{-b\rho_{a,cq}(t)\cos[\psi_{a,cq}(t)]} \cos[tk - b\rho_{a,cq}(t)\sin(\psi_{a,cq}(t))] \, \mathrm{d}t \, , \end{aligned} \qquad (12)$$

where

$$\rho_{a,d}(t) = \frac{1}{a} (1 + d^2 - 2d\cos t)^{a/2} \, ,$$

and

$$\psi_{a,d}(t) = a \arctan \frac{d \sin t}{1 - d \cos t} \, .$$

It is at once apparent that expression (12) reduces to the usual Inversion Theorem for $q = 1$, while a different and suitable choice of q may lead to a faster and more accurate evaluation of $p_{X_{\rm PT}}$. Indeed, for given a and b, the parameter q may be chosen in such a way that cq is fixed at a convenient value. In addition, the remarks provided by Dunn and Smyth (2008) for the numerical integration with oscillating integrands may be helpful in this setting. Finally, it should be remarked that (12) implies that the r.v. $\mathcal{PT}(a, b, cq)$ may be achieved by the exponentially-tilting of the r.v. $\mathcal{PT}(a, b, c)$ with tilting parameter q - when $q \in]0, 1[$.

4. Random variate generation. As to the computer generation from the Poisson-Tweedie distribution, in principle the mixture Poisson representation (3) should be the cornerstone. Actually, this equivalence in law leads to the following simple algorithm:

Algorithm 1

```
input a, b, c
generate Y absolutely-continuous Tweedie T(a, bc^a, (1 − c)/c)
generate X Poisson P(Y)
return X
```

Algorithm 1 is easy-to-implement when $a \in]-\infty, 0]$ owing to the further representation (4), since Poisson and Gamma variates are commonly available. However, the algorithm may be not convenient when $a \in]0,1]$, since in this case the absolutely-continuous Tweedie r.v. is cumbersome to generate and simple and efficient algorithms are not at disposal (see Devroye, 2009, and Hofert, 2011). Hence, the main focus of the present section is devoted to Poisson-Tweedie variate generation in this parameter range.

When $a \in]0,1]$, a second procedure may be achieved by means of the compound Poisson representation (8) - and by suitably considering expressions (6) and (7) - which actually leads to the following algorithm:

Algorithm 2

```
input a, b, c
generate N Poisson P(b/a)
for i = 1,...,N
   generate W_i geometric down-weighting Sibuya DSI(a,c)
continue
set X = sum_{i=1}^N W_i
return X
```

Unfortunately, the geometric down-weighting Sibuya distribution does not possess a reproductive property and hence the cycles in Algorithm 2 cannot be avoided. In addition, it should be remarked that the average number of cycles is given by (b/a) and hence Algorithm 2 is not suitable as $a \downarrow 0$ or $b \to \infty$. Moreover, since a geometric down-weighting Sibuya variate is obtained on the basis of representations (6) and (7), each cycle actually requires a Geometric variate, a Beta variate and a Uniform variate - alternatively and less conveniently, in turn on the basis of representation (6), each cycle involves a Poisson variate, an Exponential variate, two Gamma variates and a Uniform variate.

As a further option, since in Section 3 it is emphasized that a Poisson-Tweedie r.v. may be seen as an exponentially-tilted Discrete Stable r.v. when $a \in]0,1]$, a naive algorithm is initially introduced. Let us remind that for the Discrete Stable r.v., Devroye (1993) proved that

$$\mathcal{DS}(\gamma,\lambda) \stackrel{\mathcal{L}}{=} \mathcal{P}(\mathcal{PS}(\gamma,\lambda))\ . \tag{13}$$

Moreover, from the classical Kanter's (1975) representation it turns out that

$$\mathcal{PS}(\gamma,\lambda) \stackrel{\mathcal{L}}{=} \left(\frac{\sin((1-\gamma)\pi\mathcal{U})}{\mathcal{G}(1,1)\sin(\gamma\pi\mathcal{U})}\right)^{(1-\gamma)/\gamma}\left(\frac{\lambda\sin(\gamma\pi\mathcal{U})}{\sin(\pi\mathcal{U})}\right)^{1/\gamma}\ , \tag{14}$$

where the r.v.'s $\mathcal{G}(1,1)$ and $\mathcal{U}$ are independently distributed. Hence, since from (11) it promptly follows that

$$p_{X_{\mathrm{PT}}}(k) \le e^{\frac{b}{a}(1-c)^a} c^k\ ,$$

and by considering (13) and (14), an acceptance-rejection algorithm for the generation of an exponentially-tilted Discrete Stable variate is given by:

Algorithm 3

```
input a, b, c
repeat
   generate Z Discrete Stable DS(a, b/a)
   generate U Uniform on ]0, 1[
until U ≤ c^Z
set X = Z
return X
```

Unfortunately, Algorithm 3 may display a poor performance since the corresponding rejection constant, say A_{N}, is given by

$$A_{\text{N}} = e^{\frac{b}{a}(1-c)^a} \,.$$

As usual for an acceptance-rejection algorithm, the rejection constant represents the expected number of iterations to obtain a random variate. Obviously, the best performance is achieved for $c = 1$, while the worst performance is obtained when $a \downarrow 0$ or $b \to \infty$. In any case, the algorithm is not practically acceptable since $A_{\text{N}} = O(\exp(b/a))$. In addition, on the basis of representation (13), the algorithm requires an average of A_{N} Poisson variates, $2A_{\text{N}}$ Uniform variates and A_{N} Exponential variates.

An improved version of the Algorithm 3 may be achieved. Indeed, it is worth remarking that the sum of m *i.i.d.* Poisson-Tweedie r.v.'s $\mathcal{PT}(a, b/m, c)$ is a Poisson-Tweedie r.v. $\mathcal{PT}(a, b, c)$ - *i.e.* the distribution is actually infinitely divisible with respect to the parameter b. Hence, the random generation of m such r.v.'s by means of Algorithm 3 implies a rejection constant given by

$$A_{\text{IN}}(m) = m e^{\frac{b}{ma}(1-c)^a} \,,$$

which is minimized in $\mathbb{N}$ when $m = m^* = \max(1, [\![(b/a)(1-c)^a]\!])$ and where $[\![\,\cdot\,]\!]$ represents the rounding function. Hence, the following improved algorithm may be considered:

Algorithm 4

```
input a, b, c
set m = max(1, [[(b/a)(1 − c)^a]])
for i = 1, ..., m
   repeat
      generate Z_i Discrete Stable DS(a, b/(ma))
      generate U Uniform on ]0, 1[
   until U ≤ c^(Z_i)
continue
set X = ∑_{i=1}^m Z_i
return X
```

It should be remarked that $A_{\text{IN}}(m^*) = O(b/a)$, while $A_{\text{IN}}(m^*) \leq A_{\text{IN}}(1) = A_{\text{N}}$. Therefore, even if Algorithm 4 always improves over Algorithm 3, in turn its performance deteriorates when $a \downarrow 0$ or $b \to \infty$. Moreover, on the basis of the considerations carried out for Algorithm 3, it should be remarked that Algorithm 4 involves an average of $A_{\text{IN}}(m^*)$ Poisson variates, $2A_{\text{IN}}(m^*)$ Uniform variates and $A_{\text{IN}}(m^*)$ Exponential variates.

A further algorithm could be implemented by considering a different acceptance-rejection technique. Barabesi and Pratelli (2014b, 2015) provides a universal algorithm which is likely to conjugate efficiency and simplicity if applied to the Poisson-Tweedie distribution for $a \in]0,1]$ and $c \neq 1$. In this case, by following Barabesi and Pratelli (2015), let us consider the function

$$\alpha(q) = \frac{e^{\frac{b}{a}(1-c)^a}}{\pi} \int_0^\pi |\mathrm{E}[q^X e^{\mathrm{i}tX}]\,|\,\mathrm{d}t = \frac{e^{\frac{b}{a}(1-c)^a}}{\pi} \int_0^\pi \exp[-b\rho_{a,cq}(t)\cos(\psi_{a,cq}(t))]\,\mathrm{d}t\ ,$$

where $\rho_{a,d}$ and $\psi_{a,d}$ are introduced in Section 3. It should be also remarked that $\alpha(q)$ is defined for $q \in]0,1/c]$. Moreover, by assuming that $\alpha_1 = \alpha(q_1)$, $\alpha_2 = \alpha(q_2^{-1})$ and $\nu = \alpha(1)$ for the sake of simplicity, let us denote by

$$\beta_1 = \min(\lfloor \log_{q_1 q_2}(\alpha_1/\alpha_2) \rfloor, \lfloor \log_{q_1}(\alpha_1/\nu) \rfloor)$$

and

$$\beta_2 = \max(\lfloor \log_{q_1 q_2}(\alpha_1/\alpha_2) \rfloor + 1, \lceil \log_{q_2}(\nu/\alpha_2) \rceil)\ ,$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ represent the floor function and the ceiling function, respectively. Moreover, let us consider the quantities

$$\omega_1 = \frac{\alpha_1 q_1^{-\beta_1}}{A(1-q_1)}\ , \omega_2 = \frac{\alpha_2 q_2^{\beta_2}}{A(1-q_2)}\ , \omega_3 = \frac{\nu(\beta_2 - \beta_1 - 1)}{A}\ ,$$

and

$$A_{\text{BP}}(q_1, q_2) = \frac{\alpha_1 q_1^{-\beta_1}}{1-q_1} + \frac{\alpha_2 q_2^{\beta_2}}{1-q_2} + \nu(\beta_2 - \beta_1 - 1)\ .$$

Finally, let us choose q_1 and q_2 as $(q_1^*, q_2^*) = \arg\min A(q_1, q_2)$ - where minimization is carried out under the constrain on the domain of $\alpha(q)$. It is worth noting that in such a case $A_{\text{BP}}(q_1^*, q_2^*)$ represents the rejection constant (see Barabesi and Pratelli, 2015). Thus, for the Poisson-Tweedie distribution the universal algorithm specializes to:

Algorithm 5

```
input a, b, c
input q1, q2
compute ν, α1, α2, β1, β2
compute ω1, ω2, ω3
repeat
   generate U1, U2, U3 uniformly on ]0,1[
   if U1 > ω1 + ω2 set X := β1 + ⌊(β2 − β1 − 1)U2 + 1⌋
   else
      if U1 ≤ ω1 set X := β1 − ⌊log_{q1} U2⌋
      else
         set X := ⌊log_{q2} U2⌋ + β2
until p_{X_PT}(X) < min(α1 q1^{−X}, α2 q2^{X}, ν)U3
return X
```

Table IV. Quality assessment of the generated variates for the considered algorithms. Empirical indexes are computed on the basis of $1,000(10,000)100,000$ replicates.

a	b	c	Algorithm					
0.304	0.463	0.902		$\mu = 2.10$	$\sigma = 3.95$	$\alpha_3 = 4.11$	$\alpha_4 = 30.11$	χ^2
			1D	2.22(2.18)2.12	3.95(4.08)3.94	3.49(4.07)4.01	19.74(28.51)28.18	12.65(18.18)23.44
			1H	2.15(2.09)2.10	4.17(3.93)3.96	3.98(4.24)4.17	24.92(31.57)30.74	16.48(18.15)12.41
			2	2.14(2.11)2.10	3.87(3.93)3.95	3.74(3.70)4.09	25.11(21.74)29.20	23.38(18.37)16.98
			4	2.26(2.12)2.10	4.02(3.89)3.93	3.78(4.02)4.09	24.35(31.15)30.03	22.13(23.94)18.66
			5	2.02(2.10)2.10	3.82(4.00)3.95	4.38(4.13)4.11	32.50(28.20)30.22	22.66(18.31)23.07
0.263	0.513	0.909		$\mu = 2.73$	$\sigma = 4.78$	$\alpha_3 = 3.78$	$\alpha_4 = 25.60$	χ^2
			1D	2.68(2.73)2.73	4.27(4.67)4.76	2.78(3.43)3.70	12.84(19.81)23.94	16.36(20.91)21.56
			1H	2.55(2.75)2.74	4.45(4.90)4.82	3.89(3.81)3.73	26.17(25.27)24.70	20.56(27.20)18.65
			2	2.69(2.73)2.72	4.60(4.81)4.77	3.26(3.73)3.80	16.88(24.32)25.13	10.98(16.95)23.67
			4	2.78(2.84)2.73	4.97(5.01)4.78	4.87(3.93)3.65	42.48(27.04)23.30	22.30(18.99)22.46
			5	2.56(2.73)2.73	4.36(4.86)4.81	3.77(3.92)3.79	25.33(27.15)25.27	15.34(23.81)21.30

Acknowledgements

The authors would like to thank the two anonymous reviewers for their valuable comments and suggestions which have improved the early version of the paper.

References

Aalen, O.O. (1992) Modelling heterogeneity in survival analysis by the compound Poisson distribution, *Annals of Applied Probability* **2**, 951-972.

Ahrens, J. H. and Dieter, U. (1982) Computer generation of Poisson deviates from modified Normal distributions, *ACM Transactions on Mathematical Software* **8**, 163-179.

Baccini, A., Barabesi, L., Cioni, M. and Pisani, C. (2014) Crossing the hurdle: the determinants of individual scientific performance, *Scientometrics* **101**, 2035-2062.

Barabesi, L. and Pratelli, L. (2014a) Discussion of "On simulation and properties of the Stable law" by L. Devroye and L. James, *Statistical Methods and Applications* **23**, 345-351.

Barabesi, L. and Pratelli, L. (2014b) A note on a universal random variate generator for integer-valued random variables, *Statistics and Computing* **24**, 589-596.

Barabesi, L. and Pratelli, L. (2015) Universal methods for generating random variables with a given characteristic function, *Journal of Statistical Computation and Simulation* **85**, 1679-1691.

Burrell, Q.L. (2014) The individual author's publication-citation process: theory and practice, *Scientometrics* **98**, 725-742.

Burrell, Q.L. and Fenton, M.R. (1993) Yes, the GIGP really does work - and is workable!, *Journal of the American Society for Information Science* **44**, 61-69.

Devroye, L. (1986) *Non-uniform random variate generation*, Springer, New York.

Devroye, L. (1993) A triptych of discrete distribution related to the stable law, *Statistics and Probability Letters* **18**, 349-351.

Devroye, L. (2009) Random variate generation for exponentially and polynomially tilted stable distributions, *ACM Transactions on Modeling and Computer Simulation* **19**, Article 18.

Dunn, P.K. and Smyth, G.K. (2008) Evaluation of Tweedie exponential dispersion model densities by Fourier inversion, *Statistics and Computing* **18**, 73-86.

Egghe, L. (2005) *Power laws in the information production process: Lotkaian informetrics*, Elsevier, Oxford.

El-Shaarawi, A.H., Zhu, R. and Joe, H. (2011) Modelling species abundance using the Poisson-Tweedie family, *Environmetrics* **22**, 152-164.

Esnaola, M., Puig, P., Gonzalez, D., Castelo, R. and Gonzalez, J.R. (2013) A flexible count data model to fit the wide diversity of expression profiles arising from extensively replicated RNA-seq experiments, *BMC Bioinformatics* **14**, 254.

Gerber, H.U. (1991) From the generalized gamma to the generalized negative binomial distribution, *Insurance: Mathematics and Economics* **10**, 303–309.

Hao-Chun Chuang, H. and Oliva, R. (2014) Estimating retail demand with Poisson mixtures and out-of-sample likelihood, *Applied Stochastic Models in Business and Industry* **30**, 455-463.

Hofert, M. (2011) Sampling exponentially tilted stable distributions, *ACM Transactions on Modeling and Computer Simulation* **22**, Article 3.

Hougaard, P. (1986) Survival models for heterogeneous populations derived from stable distributions, *Biometrika* **73**, 387-396.

Hougaard, P., Lee M.T. and Whitmore, G.A. (1997) Analysis of overdispersed count data by mixtures of Poisson variables and Poisson processes, *Biometrics* **53**, 1225-1238.

Johnson, N.L., Kemp, A.W. and Kotz, S. (2005) *Univariate discrete distributions*, 3rd edn., Wiley, New York.

Jöhnk, M.D. (1975) Erzeugung von betaverteilten und gammaverteilten Zufallszahlen, Metrika **8**, 5-15.

Kanter, M. (1975) Stable densities under change of scale and total variation inequalities, *Annals of Probability* **3**, 697-707.

Kokonendji, C.C., Dossou-Gbété S. and Demétrio, C.G.B. (2004) Some discrete exponential dispersion models: Poisson-Tweedie and Hinde-Demétrio classes, *SORT* **28**, 201-214.

Lotka, A.J. (1926) The frequency distribution of scientific productivity, *Journal of the Washington Academy of Sciences* **16**, 317-323.

Marcheselli, M., Baccini, A. and Barabesi, L. (2008) Parameter estimation for the discrete stable family, *Communications in Statistics - Theory and Methods* **37**, 815-830.

Rao, I.K.R. (1980) The distribution of scientific productivity and social change, *Journal of the American Society for Information Science* **31**, 111-122.

Schubert, A. and Glänzel, W. (1984) A dynamic look at a class of skew distributions. A model with scientometric applications, *Scientometrics* **6**, 149-167.

Sibuya, M. (1979) Generalized hypergeometric, digamma and trigamma distributions, *Annals of the Institute of Statistical Mathematics* **31**, 373-390.

Sichel, H.S. (1985) A bibliometric distribution which really works, *Journal of the American Society for Information Science* **36**, 314-321.

Tweedie, M.C.K. (1984) An index which distinguishes between some important exponential families, in *Statistics: Applications and New Directions*, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (J.K. Ghosh and J. Roy, eds.), Indian Statistical Institute, Calcutta, pp. 579-604.

Wilson, C.S. (1999) Informetrics, *Annual Review of Information Science and Technology* **34**, 107-247.

Wolfram Research, Inc. (2008) Mathematica, Version 7.0, Champaign, Illinois.

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

Appendix

Result 1. By expanding (1) in exponential and binomial series, it follows that

$$\begin{aligned} G_{X_{\mathrm{PT}}}(s) &= e^{\frac{b}{a}(1-c)^a}\sum_{m=0}^{\infty}(-1)^m\frac{(b/a)^m}{m!}(1-cs)^{am} \\ &= e^{\frac{b}{a}(1-c)^a}\sum_{m=0}^{\infty}(-1)^m\frac{(b/a)^m}{m!}\sum_{k=0}^{\infty}\binom{am}{k}(-cs)^k \\ &= e^{\frac{b}{a}(1-c)^a}\sum_{k=0}^{\infty}(-cs)^k\sum_{m=0}^{\infty}(-1)^m\binom{am}{k}\frac{(b/a)^m}{m!} \end{aligned}$$

and hence

$$p_{X_{\mathrm{PT}}}(k) = e^{\frac{b}{a}(1-c)^a}(-c)^k\sum_{m=0}^{\infty}(-1)^m\binom{am}{k}\frac{(b/a)^m}{m!}I_{\mathbb{N}}(k)\,.$$

Moreover, by using the straightforward identity $\binom{am}{k} = \frac{1}{k!}\frac{d^k x^{am}}{dx^k}\Big|_{x=1}$, it also holds that

$$\begin{aligned} \sum_{m=0}^{\infty}(-1)^m\binom{am}{k}\frac{(b/a)^m}{m!} &= \frac{1}{k!}\frac{d^k e^{-\frac{b}{a}x^a}}{dx^k}\Bigg|_{x=1} = \frac{e^{-\frac{b}{a}}}{k!}\frac{d^k e^{\frac{b}{a}(1-x^a)}}{dx^k}\Bigg|_{x=1} \\ &= \frac{e^{-\frac{b}{a}}}{k!}\sum_{m=0}^{k}\frac{(b/a)^m}{m!}\frac{d^k(1-x^a)^m}{dx^k}\Bigg|_{x=1} \\ &= e^{-\frac{b}{a}}\sum_{m=0}^{k}\frac{(b/a)^m}{m!}\sum_{j=0}^{m}(-1)^j\binom{m}{j}\binom{aj}{k}\,, \end{aligned}$$

since $\frac{d^k(1-x^a)^m}{dx^k}\Big|_{x=1} = 0$ when $k > m$. Thus, expression (9) promptly follows.

Result 2. On the basis of the expression of $p_{X_{\mathrm{PT}}}$ given in Result 1, it follows

$$G_{X_{\mathrm{PT}}}(s) = e^{\frac{b}{a}[(1-c)^a+1]}\sum_{k=0}^{\infty}(-cs)^k\,\mathrm{E}\left[(-1)^M\binom{aM}{k}\right],$$

where M represents a $\mathcal{P}(b/a)$ r.v. Hence, it also holds that

$$p_{X_{\mathrm{PT}}}(k) = e^{\frac{b}{a}[(1-c)^a+1]}(-c)^k\,\mathrm{E}\left[(-1)^M\binom{aM}{k}\right].$$

Therefore, for $a \in\,]0,1]$, from the previous expression we obtain

$$\begin{aligned} p_{X_{\mathrm{PT}}}(k) &\le e^{\frac{b}{a}[(1-c)^a+1]}c^k\,\mathrm{E}\left[\left|\binom{aM}{k}\right|\right] \\ &\le \frac{b}{a}\left(1+\frac{b}{a}\right)e^{\frac{b}{a}[(1-c)^a+1]}c^k k^{-a-1}\,. \end{aligned}$$

Reply to Associate Editor

Dear Associate Editor,

we have modified the first draft of the manuscript according to the comments of the two referees. As you have suggested, in addition to the minor changes proposed by the first referee, we have especially considered the major requirements of the second referee. We hope we have addressed the referees' suggestions satisfactorily.

Best regards
Alberto Baccini, Lucio Barabesi e Luisa Stracqualursi

Reply to Referee 1

Comment 1. *In the derivation of equation 9, the authors subtracted 1 and added 1 then expanded their generating function. This is really not needed at all if binomial expansion is done directly and then changing the order of summation so the results can be directly obtained and there are no needs for Results I in the appendix. This way no needs for the many*
lines of equations given with the addition of better clarification of the results.
Reply. Thank you for your suggestion on this subtle issue. According to your comment, we have modified the initial part of Section 3 by also avoiding the use of the quantities $w_{a,m}(k)$. In order to achieve a clearer (and shorter) exposition in the main text of Section 3, we have decided to postpone the expansion in exponential and binomial series in Result 1 of the Appendix, which - even if is kept - turns out to be much shorter in the new version. Obviously, in Result 1, we also needed a passage in order to show that the summations in the expression of the p.f. $p_{X_{\mathrm{PT}}}(k)$ are finite. Incidentally, on the basis of the expansion you suggested, we have also realized that the inequality in Result 2 could be improved and a simple and accurate approximation of $p_{X_{\mathrm{PT}}}(k)$ for large k could be obtained. Thus, Result 2 is in turn modified and shortened.

Comment 2. *The second is with applications one does not just go ahead to apply such a model without considering the factors that led to the observed frequency. It seems to me that the large number of zeros is related to the number of years since the publication appeared. One can infer that a recently published paper will belong to zero class. I suggest adding another column with average number of years since publication as a covariate would likely improve the fit of the model and even may not need this heavy computation.*
Reply.

We hope we have addressed your suggestions satisfactorily.
Thank you for your revision.

Alberto Baccini, Lucio Barabesi e Luisa Stracqualursi

Reply to Referee 2

General Comment. *The authors introduce the stochastic genesis of Poisson-Tweedie distribution and suggest several algorithms for generating the Poisson-Tweedie variates. The authors also give two scientometric data, Metron and SMAP, and state, by the excellent agreement of the estimated frequencies and observed frequencies, that that the two data do follow the Poisson-Tweedie distributions. However, the materials present in the manuscript looks not complete, since the last section (5. An analysis of scientometric data) of the manuscript seems not related to the main topic of the manuscript - the five algorithms that generate the Poisson-Tweedie variates. The readers who read the paper would expect to see stuff like, based on the collected data, quality of variates generated by the five algorithms so that the performance of the algorithms can be further compared. One way to achieve this goal is to assume the scientometric data do follow the Poisson-Tweedie distributions, and generates the Poisson-Tweedie variates using the MLEs of a, b, and c.*

Reply. On the basis of your comments, we have realized that Section 5 is not homogeneous with the other parts of the manuscript. Obviously, as you surely grasp, this Section was introduced in the manuscript since our interest in the Poisson-Tweedie law originated from the need of a flexible model able to fit data which may be eventually zero-inflated or heavy-tailed - indeed, one of our aim consist in persuading practitioners to adopt this model which may be very suitable in scientometrics for this reasons. Thus, on the basis of your suggestion, we have largely modified Section 5 (see the final part of this section). In primis, we have emphasized that the results given in Section 3 - dealing with the expressions of the p.f. $p_{X_{\mathrm{PT}}}(k)$ - may provide a suitable computation of the maximum likelihood estimates of the parameters (in this way, the link between Section 3 and Section 5 is more apparent). Subsequently, as you recommend, we have generated the Poisson-Tweedie variates by using the maximum likelihood estimates as the values for the parameters a, b and c. The quality of the variates (generated according the considered algorithms) has been assessed on the basis of some empirical indexes which are compared with the corresponding true model indexes (*i.e.* the mean, the variance and the skewness and kurtosis coefficients). In addition, the χ^2 statistic was also computed for the same sets of test variates. Hence, the connection between Section 4 and Section 5 should be clearer.

Specific comment 1. *The authors apply rejection constants as the criterion in evaluating the performance of the Algorithms 1, 4, and 5, but it is not so obvious what the rejection constants are. The authors need to have a clearer definition. Also, for Algorithm 1, we have hard time in figuring out how rejection constant plays a role in the performance evaluation. A little bit detail seems necessary. Besides rejection constants, the time elapse in generating the Poisson-Tweedie variates and the quality of the generated data should be evaluated, too.*

Reply. We agree with the referee. Indeed, we attempted to present in a unique table some performance benchmarks for algorithms which are too different in their own genesis. As a matter of fact, Algorithm 1 is actually based on a stochastic representation - *i.e.* expression (3) of our paper - involving the generation of a Poisson variate and a Tweedie variate. Regrettably, the rejection constant reported in Table I of the previous version of the paper is solely referred to the complex algorithm proposed by Devroye (2009) adopted for the generation of the Tweedie variate - in addition, this algorithm stems from the double rejection method for which is even difficult to define the rejection constant in comparison with the usual acceptance-rejection method (indeed, we computed this constant by simulation). Furthermore, Algorithm 2 is in turn based on a rather complex stochastic representation - *i.e.* expression (8) of our paper - involving a (Poisson) stochastic sum of functions of Geometric and Beta random variables. Hence, the adopted performance benchmark - *i.e.* the expected number of cycles in the stochastic sum - does not adequately inform on the complexity of the algorithm. In contrast, Algorithm 4 and Algorithm 5 are actually based on the acceptance-rejection method and - more correctly - they may be judged on the basis of the rejection constants. In such a case, we opted to compare solely Algorithm 4 and Algorithm 5 on the basis of such constants and accordingly we modified Table I. By following your suggestion, we decided to compare the algorithms on the basis of the time elapse in generating the Poisson-Tweedie variates. With this aim, we tried to implement the algorithms as more efficiently as possible (we adopted the Mathematica software in so doing) and we reported the results in Table II of the new version of the manuscript. Hence, as you can see, also Section 4 was quite radically modified. Finally, by following your comment, we also evaluated the quality of the generated variates in Section 5 of the new version of the manuscript (see our reply to the general comment).

Specific comment 2. *The expected number of cycles are calculated for Algorithm 2. Why the expected number, as compared to rejection constants, is a reasonable choice?*
Reply. As remarked in our reply to the specific comment 1, we decided to avoid the comparison on the basis of this benchmark since it was not suitable.

Specific comment 3. *For performance comparison in Table 1, people might expect to see the value of N, and the number of Poisson-Tweedie variates generated.*
Reply. As remarked in our reply to the specific comment 1, in the new version of the manuscript, Table I solely contains the values of the rejection constant of Algorithm 4 and Algorithm 5 - which are computed on the basis of their closed expressions, *i.e.* by means of the values of $A_{\mathrm{IN}}(m^*)$ and $A_{\mathrm{BP}}(q_1^*, q_2^*)$.

Specific comment 4. *From page 11, the authors state that Algorithm 5 is usually the best, except few cases for b=1 and a or c are equal to 0.9. With c values for Metron and SMAP being 0.902 and 0.909, respectively, do the authors have any comments on these cases?*
Reply. Actually, we put too emphasis on the performance of Algorithm 5. In the new version of the manuscript we have modified our comments in Section 4.

We hope we have addressed your suggestions satisfactorily.
Thank you for your revision.

Alberto Baccini, Lucio Barabesi e Luisa Stracqualursi