An extended Poisson distribution

Hassan Bakouch, Maher Kachour, Saralees Nadarajah

Noname manuscript No.
(will be inserted by the editor)

An extended Poisson distribution

Hassan Bakouch · Maher Kachour · Saralees Nadarajah

Received:

Abstract The Poisson distribution is extended over the set of all integers. The motivation comes from the many reflected versions of the gamma distribution, the continuous analog of the Poisson distribution, defined over the entire real line. Various mathematical properties of the extended Poisson distribution are derived. Estimation procedures by the methods of moments and maximum likelihood are also derived with their performance assessed by simulation. Finally, a real data application is illustrated.

2000 Mathematics Subject Classification 62G30, 62E10.

Keywords Estimation · Moments · Poisson distribution · Skellam distribution

1 Introduction

It is well known that the Poisson distribution is the discrete analog of the gamma distribution. The gamma distribution is a popular model for continuous data. Many extensions of the gamma distribution have been proposed in the literature for improved modeling, see Johnson *et al.* (1994, 1995). These extensions include reflected versions of the gamma distribution defined over

Hassan Bakouch
Department of Statistics, Faculty of Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia. E-mail: hbakouch@kau.edu.sa

Maher Kachour
Ecole supérieure de commerce IDRAC, 47, rue Sergent Michel Berthet CP 607, 69258 Lyon Cedex 09, France. E-mail: maher.kachour@idraclyon.com

Saralees Nadarajah
School of Mathematics, University of Manchester, Manchester M13 9PL, UK. E-mail: mbbsssn2@manchester.ac.uk

the entire real line. The first reflected gamma distribution was proposed by Borghi (1965). This distribution was defined by taking a mirror image of the probability density function of the gamma distribution along the y axis. Several other reflected gamma distributions have been proposed since Borghi (1965): Ali and Woo (2006) defined the skew reflected gamma distribution; Ali *et al.* (2008) defined skewed reflected distributions generated by a reflected gamma kernel; Ali *et al.* (2009) defined skewed reflected gamma distributions generated by a Laplace kernel; Ali *et al.* (2010) defined skewed inverse reflected gamma distributions. One should also mention that the Laplace distribution is the reflected version of the exponential distribution.

Each of the reflected gamma distributions including the Laplace distribution has received many real data applications. But data are never continuous. Data are discrete by nature. So, it would be useful to have reflected versions of the Poisson distribution. We are not aware of any reflected version of the Poisson distribution proposed in the literature. Another motivation is that the distribution is under-dispersed, equi-dispersed and over-dispersed, see Section 2.5.

The aim of this paper is to propose the first reflected version of the Poisson distribution, referred to as the extended Poisson (E-Po) distribution. We say that a random variable $X \sim \text{E-Po}(p, \lambda)$ with parameters $\lambda > 0$ and $0 \leq p \leq 1$ if its probability mass function is defined as

$$\mathbb{P}(X = k) = \begin{cases} e^{-\lambda}, & \text{if } k = 0, \\ pe^{-\lambda}\dfrac{\lambda^k}{k!}, & \text{if } k = 1, 2, \ldots, \\ (1-p)e^{-\lambda}\dfrac{\lambda^{|k|}}{|k|!}, & \text{if } k = \ldots, -2, -1. \end{cases} \tag{1}$$

In other words, for all $x \in \mathbb{Z}$, we have

$$\mathbb{P}\left(X = x\right) = p\mathbb{P}\left(Y = x\right)\mathbf{1}_{\{x \geq 0\}} + (1-p)\mathbb{P}\left(Y = |x|\right)\mathbf{1}_{\{x \leq 0\}}, \tag{2}$$

where Y is a Poisson random variable with parameter $\lambda > 0$. Note that

- If $p = 1$, then X has the same distribution than that of Y;
- If $p = 0$, then X has the same distribution than that of $-Y$.

Note also that if $X \sim \text{E-Po}(p, \lambda)$ then $-X \sim \text{E-Po}(1-p, \lambda)$ and $|X|$ is a Poisson random variable with parameter λ.

The contents of this paper are organized as follows. Various mathematical properties of the E-Po distribution are derived in Section 2. These include

the cumulative distribution function (Section 2.1), the quantile function (Section 2.2), the survival function (Section 2.3), the failure rate function (Section 2.3), the reverse failure rate function (Section 2.3), the conditional moments (Section 2.4), the probability generating function (Section 2.5), the moments (Section 2.5), the Shannon entropy (Section 2.6), the mean deviation about the mean (Section 2.7), the mean deviation about the median (Section 2.7), the distribution of sums and differences (Section 2.8). Estimation procedures by the method of moments (Section 3.1) and the method of maximum likelihood (Section 3.2) as well as a simulation study (Section 3.3) to assess their performance are given in Section 3. Finally, Section 4 illustrates practical usefulness of the E-Po distribution.

2 Mathematical properties

2.1 Cumulative distribution function

Let $F_X(x) = \mathbb{P}(X \leq x)$ denote the cumulative distribution function and $\lfloor \cdot \rfloor$ the floor function. Furthermore, for all $x \geq 0$ recall that

$$e^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!} = \frac{\Gamma(\lfloor x+1 \rfloor, \lambda)}{\lfloor x \rfloor !},$$

where $\Gamma(\cdot, \cdot)$ denotes the incomplete gamma function. Suppose now that $x < 0$. Then, we have

$$\begin{aligned} F_X(x) &= \sum_{k=-\infty}^{\lfloor x \rfloor} \mathbb{P}(X = k) = (1-p)e^{-\lambda} \sum_{k=-\infty}^{\lfloor x \rfloor} \frac{\lambda^{|k|}}{|k|!} \\ &= (1-p)e^{-\lambda} \sum_{k=-\lfloor x \rfloor}^{\infty} \frac{\lambda^k}{k!} \\ &= (1-p)\left[1 - \frac{\Gamma\left(-\lfloor x \rfloor, \lambda\right)}{(-\lfloor x \rfloor - 1)!}\right]. \end{aligned}$$

Now if $0 \leq x < 1$, then

$$\begin{aligned} F_X(x) &= \mathbb{P}(X = 0) + F_X(-1) \\ &= e^{-\lambda} + (1-p)\left(1 - \Gamma(1, \lambda)\right) = (1-p) + pe^{-\lambda}. \end{aligned}$$

Finally, for $x \geq 1$, we have

$$\begin{aligned} F_X(x) &= \sum_{k=-\infty}^{\lfloor x \rfloor} \mathbb{P}(X=k) \\ &= F_X(0) + \sum_{k=1}^{\lfloor x \rfloor} \mathbb{P}(X=k) \\ &= (1-p) + pe^{-\lambda} + pe^{-\lambda} \sum_{k=1}^{\lfloor x \rfloor} \frac{\lambda^k}{k!} \\ &= (1-p) + pe^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!} \\ &= (1-p) + p\frac{\Gamma\left(\lfloor x+1 \rfloor, \lambda\right)}{\lfloor x \rfloor !}. \end{aligned}$$

2.2 Quantile function

Letting $Q(a,x) = \Gamma(a,x)/\Gamma(a)$ denote the regularized incomplete gamma function, we can rewrite the cdf of X as

$$F_X(x) = \begin{cases} (1-p)\left[1 - Q\left(-\lfloor x \rfloor, \lambda\right)\right], & \text{if } x < 0, \\ (1-p) + pe^{-\lambda}, & \text{if } 0 \leq x < 1, \\ (1-p) + pQ\left(\lfloor x+1 \rfloor, \lambda\right), & \text{if } x \geq 1. \end{cases}$$

Hence, the quantile function of X can be expressed as

$$F_X^{-1}(q) = \begin{cases} -Q^{-1}\left(1 - \dfrac{q}{1-p}, \lambda\right), & \text{if } q < 1 - p\left[1 - Q(0,\lambda)\right], \\ 0, & \text{if } 1 - p\left[1 - Q(0,\lambda)\right] \\ & \qquad \leq q < (1-p) + pe^{-\lambda}, \\ Q^{-1}\left(\dfrac{q-1+p}{p}, \lambda\right) - 1, & \text{if } q \geq (1-p) + pe^{-\lambda}, \end{cases}$$

where $Q^{-1}(a,\lambda)$ denotes the inverse function of $Q(a,\lambda)$. This quantile function can be used to simulate variates of the EP distribution. The same simulation can be performed perhaps more easily by using (2).

2.3 Failure functions

Let Y denote a Poisson random variable with parameter λ. Let F_Y, R_Y, k_Y, and τ_Y denote, respectively, the cumulative distribution function, survival function, failure rate function, and the reverse failure function of Y. Then, the survival function, failure rate function, and the reverse failure function of X can be expressed as

$$\begin{aligned} R_X(x) &= \mathbb{P}(X > x) = 1 - F_X(x) \\ &= \begin{cases} pR_Y(x), & \text{if } \ x \geq 0, \\ pR_Y(-x-1) + F_Y(-x-1), & \text{if } x \leq -1, \end{cases} \end{aligned}$$

$$\begin{aligned} k_X(x) &= \mathbb{P}(X = x \mid X \geq x) = \frac{\mathbb{P}(X = x)}{R_X(x-1)} \\ &= \begin{cases} k_Y(x), & \text{if } \ x \geq 1, \\ \dfrac{e^{-\lambda}}{p(1-e^{-\lambda}) + e^{-\lambda}}, & \text{if } \ x = 0, \\ \dfrac{(1-p)k_Y(-x)}{p\frac{R_Y(-x)}{R_Y(-x-1)} + \frac{F_Y(-x)}{R_Y(-x-1)}}, & \text{if } x \leq -1, \end{cases} \end{aligned}$$

and

$$\begin{aligned} \tau_X(x) &= \mathbb{P}(X = x \mid X \leq x) = \frac{\mathbb{P}(X = x)}{F_X(x)} \\ &= \begin{cases} \dfrac{\tau_Y(x)}{1 + \frac{(1-p)}{pF_Y(x)}}, & \text{if } \ x \geq 1, \\ \dfrac{e^{-\lambda}}{pe^{-\lambda} + (1-p)}, & \text{if } \ x = 0, \\ k_Y(-x), & \text{if } x \leq -1, \end{cases} \end{aligned}$$

respectively. Figures and show respectively the behaviour of the failure rate and the reverse failure functions of X, for different values of λ and p

2.4 Conditional moments

For lifetime models, it is of interest to know the conditional measure $E\{(X - x)^k | X > x\}$, $k = 1, 2, ...$, which is important in prediction and can be used to measure the reliability of systems. For all $x \in \mathbb{Z}$, the kth-order conditional moment is

$$c_X^k(x) = \mathbb{E}\left((X-x)^k \mid X \geq x\right) = \frac{\sum_{j=x}^{\infty}(j-x)^k \mathbb{P}(X=j)}{R_X(x-1)}.$$

In particular, for $k = 1, 2$, we have

$$c_X^1(x) = \frac{\sum_{j=x}^{\infty} R_X(j)}{R_X(x-1)}$$

$$= \begin{cases} c_Y^1(x), & \text{if } \ x \geq 1, \\ \dfrac{p(1-e^{-\lambda})}{p(1-e^{-\lambda})+e^{-\lambda}}\left(c_Y^1(1)+1\right), & \text{if } \ x = 0, \\ \dfrac{\sum_{j=x}^{-1} R_X(j) + c_X^1(0)\left(p+(1-p)e^{-\lambda}\right)}{R_X(x-1)}, & \text{if } x \leq -1, \end{cases}$$

and

$$c_X^2(x) = \frac{2\sum_{j=x}^{\infty} jR_X(j)}{R_X(x-1)} - (2x-1)c_X^1(x)$$

$$= \begin{cases} c_Y^2(x), & \text{if } \ x \geq 1, \\ \dfrac{p(1-e^{-\lambda})}{p(1-e^{-\lambda})+e^{-\lambda}}\left(c_Y^2(1)+c_Y^1(1)\right) + c_X^1(0), & \text{if } \ x = 0, \\ Q(x) - (2x-1)c_X^1(x), & \text{if } x \leq -1, \end{cases}$$

where

$$Q(x) = \frac{p\left[C_Y^1(1) + C_Y^2(1)\right] R_Y(0) + 2\sum_{j=x}^{-1} jR_X(j)}{R_X(x-1)}$$

and $c_Y^k(x)$ be the kth-order conditional moment of Y. Note that c_X^1 is called the mean residual life function of X, and the variance residual life function of X is

$$\beta_X(x) = \mathbb{V}\left(X - x \mid X \geq x\right) = c_X^2(x) - \left(c_X^1(x)\right)^2.$$

2.5 Probability generating function and Moments

Several interesting characteristics of a distribution can be studied by its probability generating function and moments, so we get them to the E-Po distribution.

Define

$$G_{X,+}(s) = p\mathbb{P}(X = 0) + \sum_{k=1}^{\infty} s^k \mathbb{P}(X = k) = pG_Y(s),$$

and

$$G_{X,-}(s) = (1-p)\mathbb{P}(X = 0) + \sum_{k=-\infty}^{-1} s^{|k|} \mathbb{P}(X = k) = (1-p)G_Y(1/s),$$

where $G_Y(s) = e^{-\lambda(1-s)}$ is the pgf of Y, a Poisson random variable with parameter λ. Then, the pgf of X is

$$G_X(s) = e^{-\lambda}\left[pe^{s\lambda} + (1-p)e^{\frac{\lambda}{s}}\right]. \tag{3}$$

For$s = -\infty$, the pgf (3) is given by $G_X(-\infty) = (1-p)e^{-\lambda}$. This quantity is known as the survival fraction and it plays an important role in long-term survival analysis. For more details, see Rodrigues et al. (2008).

The characteristic function of X follows from (3) as

$$\varphi_X(t) = \mathbb{E}\left(e^{\mathrm{i}tX}\right) = pe^{-\lambda\left(1-e^{\mathrm{i}t}\right)} + (1-p)e^{-\lambda\left(1-e^{-\mathrm{i}t}\right)},$$

where $\mathrm{i} = \sqrt{-1}$.

For any $r \in \mathbb{N} \setminus \{0\}$, the rth-order descending factorial moment is

$$\begin{aligned}\mu_X(r) &= \mathbb{E}\left[X(X-1)\ldots(X-(r-1))\right] \\ &= pG_Y^{(r)}(1) + (-1)^r(1-p)G_{Y'}^{(r)}(1),\end{aligned}$$

where $G_{Y'}(s) = s^{r-1}G_Y(s)$. For any $k \in \mathbb{N} \setminus \{0\}$, the kth-order moment of X is

$$\mathbb{E}\left(X^k\right) = \begin{cases} M_Y^{(k)}(0), & \text{if } k \text{ even}, \\ (2p-1)M_Y^{(k)}(0), & \text{if } k \text{ odd}, \end{cases}$$

where

$$M_Y(t) = \mathbb{E}\left(e^{tY}\right) = G_Y\left(e^t\right) = e^{-\lambda\left(1-e^t\right)}.$$

Thus, the first four moments of X are

$$\mathbb{E}(X) = (2p-1)\lambda, \mathbb{E}\left(X^2\right) = \lambda(1+\lambda),$$
$$\mathbb{E}\left(X^3\right) = (2p-1)\lambda\left(1+3\lambda+\lambda^2\right), \mathbb{E}\left(X^4\right) = \lambda\left(1+7\lambda+6\lambda^2+\lambda^3\right).$$

The corresponding variance, coefficient of variation, skewness and kurtosis are

$$\mathbb{V}(X) = \lambda + 4p(1-p)\lambda^2,$$
$$\mathrm{CV}(X) = \frac{\sigma_X}{\lambda(2p-1)},$$

$$\begin{aligned} \text{Skewness}(X) &= \frac{\mathbb{E}\left(X-\mathbb{E}(X)\right)^3}{\sigma_X^3} \\ &= \frac{\lambda^3(1-3(2p-1)+(2p-1)^3)}{\sigma_X^3} + \frac{6\lambda^2(1-p)}{\sigma_X^3} + \frac{\lambda}{\sigma_X^3} \end{aligned}$$

and

$$\begin{aligned} \text{Kurtosis}(S) &= \frac{\mathbb{E}\left(X-\mathbb{E}(X)\right)^4}{\sigma_X^4} - 3 \\ &= \frac{\lambda^4(1+2(2p-1)^2-3(2p-1)^4)}{\sigma_X^4} + \frac{6\lambda^3(1-(2p-1)^2)}{\sigma_X^4} \\ &\quad + \frac{\lambda^2(7-4(2p-1)^2)}{\sigma_X^4} + \frac{\lambda}{\sigma_X^4} - 3, \end{aligned}$$

where $\sigma_X = \sqrt{\lambda + 4p(1-p)\lambda^2}$. The index of dispersion is defined as the ratio of the variance to the mean, i.e.,

$$ID = \frac{\mathbb{V}(X)}{\mathbb{E}(X)} = \frac{1+4p(1-p)\lambda}{(2p-1)}.$$

Note that $ID = -1$, for $p = 0$; $ID < 0$, for $0 < p < \frac{1}{2}$; $ID = 1$, for $p = 1$ and $ID > 0$, for $\frac{1}{2} < p < 1$. So, the E-Po distribution is equi-dispersed if $p = 0$ or

$p = 1$. On the other hand, one can see that $ID < 1 \Leftrightarrow \lambda < -\frac{1}{2p}$ which is not possible since $\lambda > 0$ and $ID > 1 \Leftrightarrow \lambda > -\frac{1}{2p}$ which is always real for the same reason. Similary, we have that $ID > -1 \Leftrightarrow \lambda > -\frac{1}{2(1-p)}$ which is always real. Thus, one can deduce that E-Po distribution is over-dispersed if $\frac{1}{2} < p < 1$ and under-dispersed if $0 < p < \frac{1}{2}$.

2.6 Entropy

Entropy is used to measure the randomness of systems and it is widely used in various sciences such as physics. In this section, we give the Shannon entropy, which is the most popular entropy, defined by $E[-\log p(X)]$, where $p(k) = \mathbb{P}(X = k)$ denotes the probability mass function. Using (1), we obtain

$$\begin{aligned}
E\left[-\log p(X)\right] &= -\log p(0)\mathbb{P}(X=0) - \sum_{k=1}^{\infty} \log p(k)\mathbb{P}(X=k) \\
&\quad - \sum_{k=-\infty}^{-1} \log p(k)\mathbb{P}(X=k) \\
&= \lambda e^{-\lambda} + p\left[\lambda - \log p\right] e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} - pe^{-\lambda}\log\lambda \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \\
&\quad + pe^{-\lambda} \sum_{k=1}^{\infty} \log k! \frac{\lambda^k}{k!} \\
&\quad + (1-p)\left[\lambda - \log(1-p)\right] e^{-\lambda} \sum_{k=-\infty}^{-1} \frac{\lambda^{-k}}{(-k)!} \\
&\quad - (1-p)e^{-\lambda}\log\lambda \sum_{k=-\infty}^{-1} (-k)\frac{\lambda^{-k}}{(-k)!} \\
&\quad + (1-p)e^{-\lambda} \sum_{k=-\infty}^{-1} \log(-k)! \frac{\lambda^{-k}}{(-k)!} \\
&= \lambda - \left[p\log p + (1-p)\log(1-p)\right]\left(1 - e^{-\lambda}\right) - \lambda\log\lambda \\
&\quad + e^{-\lambda} \sum_{k=1}^{\infty} \log k! \frac{\lambda^k}{k!}. \qquad (4)
\end{aligned}$$

Note that (4) is a closed form expression expect for the single infinite summation.

2.7 Mean deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. These are known as the mean deviation about the mean and the mean deviation about the median – defined by

$$\delta_1(X) = \mathbb{E}\left[|k - \mathbb{E}(X)|\right] = \sum_{k=-\infty}^{\infty} |k - \mathbb{E}(X)|\, \mathbb{P}(X = k) \tag{5}$$

and

$$\delta_2(X) = \mathbb{E}\left[|k - \text{Median}(X)|\right] = \sum_{k=-\infty}^{\infty} |k - \text{Median}(X)|\, \mathbb{P}(X = k), \tag{6}$$

respectively. The measures can be calculated using the general relationship that

$$\begin{aligned}
\delta(X) &= \mathbb{E}\left[|k - m|\right] \\
&= \sum_{k=-\infty}^{\infty} |k - m|\, \mathbb{P}(X = k) \\
&= \sum_{k=\lfloor m \rfloor + 1}^{\infty} (k - m)\mathbb{P}(X = k) + \sum_{k=-\infty}^{\lfloor m \rfloor} (m - k)\mathbb{P}(X = k) \\
&= \sum_{k=\lfloor m \rfloor + 1}^{\infty} k\mathbb{P}(X = k) - \sum_{k=-\infty}^{\lfloor m \rfloor} k\mathbb{P}(X = k) - m + 2m \sum_{k=-\infty}^{\lfloor m \rfloor} \mathbb{P}(X = k) \\
&= 2 \sum_{k=\lfloor m \rfloor + 1}^{\infty} k\mathbb{P}(X = k) - \mathbb{E}(X) - m + 2mF_X(m),
\end{aligned} \tag{7}$$

where $\mathbb{E}(X) = (2p-1)\lambda$ and $F_X(\cdot)$ is given by Section 2.1. Substituting $e^{-\lambda}\sum_{k=0}^{\lfloor x \rfloor}\frac{\lambda^k}{k!}$, the summation term in (7) can be expressed as

$$\begin{aligned}
\sum_{k=\lfloor m \rfloor+1}^{\infty} k\mathbb{P}(X=k) &= pe^{-\lambda}\sum_{k=\lfloor m \rfloor+1}^{\infty}\frac{\lambda^k}{(k-1)!} \\
&= p\lambda e^{-\lambda}\sum_{k=\lfloor m \rfloor}^{\infty}\frac{\lambda^k}{k!} \\
&= \begin{cases} p\lambda + p\lambda e^{-\lambda}\sum_{k=\lfloor m \rfloor}^{-1}\frac{\lambda^k}{k!}, & \text{if } \lfloor m \rfloor \leq -1, \\ p\lambda, & \text{if } \lfloor m \rfloor = 0, \\ p\lambda\left\{1 - \frac{\Gamma(\lfloor m \rfloor, \lambda)}{\lfloor m-1 \rfloor!}\right\}, & \text{if } \lfloor m \rfloor \geq 1. \end{cases}
\end{aligned} \tag{8}$$

Combing (7) and (8), we have an expression for $\delta(X)$. Expressions for the mean deviations in (5) and (6) can be obtained as particular cases.

2.8 Distribution of sums and differences

It is well known that the sum of independent Poisson random variables follows a Poisson distribution. It is also well known that the difference of two independent Poisson random variables follows a Skellam distribution (Skellam, 1946). But these properties do not hold for E-Po random variables, see Section 2.5. Here, we derive the distributions of the sum and the difference of two independent E-Po random variables and get their moments.

Suppose $X_1 \sim \text{E-Po}(p_1, \lambda_1)$ and $X_2 \sim \text{E-Po}(p_2, \lambda_2)$ are independent random variables. Since $X \sim \text{E-Po}(p, \lambda)$ implies $-X \sim \text{E-Po}(1-p, \lambda)$, we only

consider the distribution of the sum $X_1 + X_2 = S$ say. For $k > 0$, we have

$$\begin{aligned}
\mathbb{P}(S=k) &= \mathbb{P}(X=0)\mathbb{P}(Y=k) + \mathbb{P}(X=k)\mathbb{P}(Y=0) \\
&\quad + \sum_{j=1}^{k-1} \mathbb{P}(X=j)\mathbb{P}(Y=k-j) + \sum_{j=-\infty}^{-1} \mathbb{P}(X=j)\mathbb{P}(Y=k-j) \\
&\quad + \sum_{j=-\infty}^{-1} \mathbb{P}(Y=j)\mathbb{P}(X=k-j) \\
&= p_2 e^{-\lambda_1-\lambda_2} \frac{\lambda_2^k}{k!} + p_1 e^{-\lambda_1-\lambda_2} \frac{\lambda_1^k}{k!} + p_1 p_2 e^{-\lambda_1-\lambda_2} \sum_{j=1}^{k-1} \frac{\lambda_1^j \lambda_2^{k-j}}{j!(k-j)!} \\
&\quad + (1-p_1)\, p_2 e^{-\lambda_1-\lambda_2} \sum_{j=-\infty}^{-1} \frac{\lambda_1^{-j} \lambda_2^{k-j}}{(-j)!(k-j)!} \\
&\quad + p_1\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \sum_{j=-\infty}^{-1} \frac{\lambda_1^{k-j} \lambda_2^{-j}}{(k-j)!(-j)!} \\
&= p_2 e^{-\lambda_1-\lambda_2} \frac{\lambda_2^k}{k!} + p_1 e^{-\lambda_1-\lambda_2} \frac{\lambda_1^k}{k!} \\
&\quad + p_1 p_2 e^{-\lambda_1-\lambda_2} \sum_{j=1}^{k-1} \frac{\lambda_1^j \lambda_2^{k-j}}{j!(k-j)!} \\
&\quad + (1-p_1)\, p_2 e^{-\lambda_1-\lambda_2} \frac{\lambda_2^k}{k!} \left[{}_0F_1\left(;k+1;\lambda_1\lambda_2\right) - 1\right] \\
&\quad + p_1\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_1^k}{k!} \left[{}_0F_1\left(;k+1;\lambda_1\lambda_2\right) - 1\right], \qquad (9)
\end{aligned}$$

where ${}_0F_1(;a;x)$ denotes a hypergeometric function defined by

$${}_0F_1\left(;a;x\right) = \sum_{k=0}^{\infty} \frac{1}{(a)_k} \frac{x^k}{k!},$$

where $(f)_k = f(f+1)\cdots(f+k-1)$ denotes the ascending factorial. Similar calculations for $k = 0$ show that

$$\begin{aligned}
\mathbb{P}(S=0) &= e^{-\lambda_1-\lambda_2} + (1-p_1)\, p_2 e^{-\lambda_1-\lambda_2} \left[{}_0F_1\left(;1;\lambda_1\lambda_2\right) - 1\right] \\
&\quad + p_1\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \left[{}_0F_1\left(;1;\lambda_1\lambda_2\right) - 1\right]. \qquad (10)
\end{aligned}$$

For $k < 0$, we have

$$\begin{aligned}
\mathbb{P}(S=k) &= (1-p_2)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_2^{-k}}{(-k)!} + (1-p_1)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_1^{-k}}{(-k)!} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \sum_{j=1}^{k-1} \frac{\lambda_1^{-j}\lambda_2^{j-k}}{(-j)!(j-k)!} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \sum_{j=-\infty}^{-1} \frac{\lambda_1^{-j}\lambda_2^{j-k}}{(-j)!(j-k)!} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \sum_{j=-\infty}^{-1} \frac{\lambda_2^{-j}\lambda_1^{j-k}}{(-j)!(j-k)!} \\
&= (1-p_2)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_2^{-k}}{(-k)!} + (1-p_1)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_1^{-k}}{(-k)!} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \sum_{j=1}^{k-1} \frac{\lambda_1^{-j}\lambda_2^{j-k}}{(-j)!(j-k)!} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_2^{-k}}{(-k)!} \left(1+\frac{\lambda_1}{\lambda_2}\right)^{-k} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_1^{-k}}{(-k)!} \left(1+\frac{\lambda_2}{\lambda_1}\right)^{-k} \\
&= (1-p_2)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_2^{-k}}{(-k)!} + (1-p_1)\, e^{-\lambda_1-\lambda_2} \frac{\lambda_1^{-k}}{(-k)!} \\
&\quad + (1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2} \sum_{j=1}^{k-1} \frac{\lambda_1^{-j}\lambda_2^{j-k}}{(-j)!(j-k)!} \\
&\quad + \frac{2\,(1-p_1)\,(1-p_2)\, e^{-\lambda_1-\lambda_2}}{(-k)!\,(\lambda_1+\lambda_2)^k}. \qquad (11)
\end{aligned}$$

To find the distribution of the difference $X_1 - X_2$, just replace p_2 by $1-p_2$ in (9), (10) and (11).

The moments of $X_1 + X_2$ and $X_1 - X_2$ can be calculated using the facts that

$$\mathbb{E}\left[(X_1+X_2)^k\right] = \sum_{m=0}^{k} \binom{k}{m} \mathbb{E}\left[X_1^m\right] \mathbb{E}\left[X_2^{k-m}\right]$$

and

$$\mathbb{E}\left[(X_1-X_2)^k\right] = \sum_{m=0}^{k} \binom{k}{m} (-1)^{k-m} \mathbb{E}\left[X_1^m\right] \mathbb{E}\left[X_2^{k-m}\right],$$

respectively.

3 Estimation and inference with simulation

Let $X_1, \ldots, X_n$ denote a sample of i.i.d. observations from the EP distribution with parameters p and λ. Here, we consider estimation of the parameters by the method of moments and the method of maximum likelihood with inference of the distribution parameters. Also, simulation results on the behavior of estimators are presented.

3.1 Method of moments

Since $\mathbb{E}(|X|) = \lambda$, a method of moments estimator of λ is

$$\overline{\lambda}_n = \frac{1}{n}\sum_{i=1}^{n} |X_i|. \tag{12}$$

Since $\mathbb{P}(X = 0) = e^{-\lambda}$, another method of moments estimator of λ is

$$\widetilde{\lambda}_n = -\log\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i=0\}}\right). \tag{13}$$

The method of moments estimator for p can be obtained using moments of X given in Section 2.5: the ones corresponding (12) and (13) are

$$\overline{p}_n = \frac{1}{2}\left(\frac{\sum_{i=1}^{n} X_i}{n\overline{\lambda}_n} + 1\right) \tag{14}$$

and

$$\widetilde{p}_n = \frac{1}{2}\left(\frac{\sum_{i=1}^{n} X_i}{n\widetilde{\lambda}_n} + 1\right), \tag{15}$$

respectively. The exact sampling distributions of $\overline{\lambda}_n$ and $\widetilde{\lambda}_n$ follow by noting that

$$\sum_{i=1}^{n} |X_i| \sim \text{Poisson}(n\lambda) \tag{16}$$

and

$$\sum_{i=1}^{n} \mathbf{1}_{\{X_i=0\}} \sim \text{Binomial}\left(n, e^{-\lambda}\right), \tag{17}$$

respectively. The exact sampling distributions of $\overline{p}_n$ and $\widetilde{p}_n$ are difficult to find because the sum of independent E-Po random variables does not have a closed form distribution function, see Section 2.8.

3.2 Method of maximum likelihood

Set $x = (x_1, \ldots, x_n)$ and $\theta = (\lambda, p)$. The likelihood function is

$$L(x, \theta) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n |x_i|}}{\prod_{i=1}^n |x_i|!} \prod_{i=1}^n \left[p \mathbf{1}_{\{x_i \geq 0\}} + (1-p) \mathbf{1}_{\{x_i \leq 0\}} \right].$$

The log likelihood function is

$$\begin{aligned} \log L(x, \theta) = -n\lambda + \left(\sum_{i=1}^n |x_i| \right) \log \lambda - \sum_{i=1}^n \log |x_i|! \\ + \sum_{i=1}^n \log \left[p \mathbf{1}_{\{x_i \geq 0\}} + (1-p) \mathbf{1}_{\{x_i \leq 0\}} \right]. \end{aligned}$$

The maximum likelihood estimator $\widehat{\theta}_n = \left(\widehat{\lambda}_n, \widehat{p}_n \right)$ is defined as $\widehat{\theta}_n = \text{argmax}_\theta \log L(x, \theta)$ .Thus, the maximum likelihood estimator of λ is

$$\widehat{\lambda}_n = \frac{1}{n} \sum_{i=1}^n |x_i|. \tag{18}$$

That of p is

$$\widehat{p}_n = \frac{\sum_{i=1}^n \mathbf{1}_{\{x_i > 0\}}}{\sum_{i=1}^n \mathbf{1}_{\{x_i > 0\}} + \sum_{i=1}^n \mathbf{1}_{\{x_i < 0\}}}. \tag{19}$$

The exact sampling distribution of $\widehat{\lambda}_n$ follows from (16). The exact sampling distribution of $\widehat{p}_n$ is difficult to obtain. It can be approximated by asymptotic normality.

An exact $100(1-\alpha)$ percent confidence interval for λ based on (16) is

$$\frac{1}{2n} \chi^2_{2 \sum_{i=1}^n |x_i|, \alpha/2} < \lambda < \frac{1}{2n} \chi^2_{2 \sum_{i=1}^n |x_i|, 1-\alpha/2}, \tag{20}$$

where $\chi^2_{\nu,a}$ is the $100a$ percentile of a chisquare random variable with ν degrees of freedom. An exact $100(1-\alpha)$ percent confidence interval for λ based on (17) is

$$-\log I^{-1}_{1-\alpha/2} \left(1 + \sum_{i=1}^n \mathbf{1}_{\{X_i = 0\}}, n - \sum_{i=1}^n \mathbf{1}_{\{X_i = 0\}} \right) \tag{21}$$

$$< \lambda < -\log I^{-1}_{\alpha/2} \left(\sum_{i=1}^n \mathbf{1}_{\{X_i = 0\}}, n + 1 - \sum_{i=1}^n \mathbf{1}_{\{X_i = 0\}} \right), \tag{22}$$

where $I_x^{-1}(a,b)$ is the inverse function of the incomplete beta function ratio defined by $I_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt/B(a,b)$, where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ denotes the beta function.

Approximate confidence intervals can be based on asymptotic normality of the maximum likelihood estimators. As $n \to \infty$, $\sqrt{n}\left(\widehat{\lambda}_n - \lambda, \widehat{p}_n - p\right)$ approaches a bivariate normal random vector with zero means and variance-covariance matrix V say. Standard calculations show that

$$\mathbb{E}\left(-\frac{\partial^2 \log L}{\partial \lambda^2}\right) = \frac{1}{\lambda^2}\mathbb{E}\left(\sum_{i=1}^{n}|x_i|\right) = \frac{n}{\lambda},$$
$$\mathbb{E}\left(-\frac{\partial^2 \log L}{\partial \lambda \partial p}\right) = 0, \mathbb{E}\left(-\frac{\partial^2 \log L}{\partial p^2}\right) = \frac{n\left(1-e^{-\lambda}\right)}{p(1-p)}.$$

Hence,

$$\mathbb{V}\left(\widehat{\lambda}_n\right) \approx \frac{\lambda_n}{n}, \mathrm{Cov}\left(\widehat{\lambda}_n, \widehat{p}_n\right) \approx 0, \ \mathbb{V}\left(\widehat{p}_n\right) \approx \frac{p(1-p)}{n\left(1-e^{-\lambda}\right)}.$$

Note that $\widehat{\lambda}_n$ and $\widehat{p}_n$ are asymptotically independent.

So, an approximate $100(1-\alpha)$ percent confidence interval for λ is

$$\widehat{\lambda}_n - z_{\alpha/2}\frac{\sqrt{\widehat{\lambda}_n}}{\sqrt{n}} < \lambda < \widehat{\lambda}_n + z_{\alpha/2}\frac{\sqrt{\widehat{\lambda}_n}}{\sqrt{n}}, \tag{23}$$

where z_a is the $100a$ percentile of a standard normal random variable. An approximate $100(1-\alpha)$ percent confidence interval for p is

$$\widehat{p}_n - z_{\alpha/2}\frac{1}{\sqrt{n}}\sqrt{\frac{\widehat{p}_n\left(1-\widehat{p}_n\right)}{1-e^{-\widehat{\lambda}_n}}} < p < \widehat{p}_n + z_{\alpha/2}\frac{1}{\sqrt{n}}\sqrt{\frac{\widehat{p}_n\left(1-\widehat{p}_n\right)}{1-e^{-\widehat{\lambda}_n}}}. \tag{24}$$

3.3 A simulation study

Here, we assess the finite sample performance of the point and interval estimators in Sections 3.1 and 3.2. The assessment of the performance of the estimators is based on a simulation study:

1. generate ten thousand samples of size n from the E-Po distribution. The representation (2) was used to generate samples.
2. compute (13), (14), (15), (18) and (19) for the ten thousand samples, say $\left(\widetilde{\lambda}_{n,i}, \overline{p}_{n,i}, \widetilde{p}_{n,i}, \widehat{\lambda}_{n,i}, \widehat{p}_{n,i}\right)$ for $i = 1, 2, \ldots, 10000$.

3. compute the biases and mean squared errors given by

$$\text{bias}_1(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widetilde{\lambda}_{n,i} - \lambda\right), \text{bias}_2(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\overline{p}_{n,i} - p\right),$$

$$\text{bias}_3(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widetilde{p}_{n,i} - p\right), \text{ bias}_4(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widehat{\lambda}_{n,i} - \lambda\right),$$

$$\text{bias}_5(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widehat{p}_{n,i} - p\right), \text{MSE}_1(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widetilde{\lambda}_{n,i} - \lambda\right)^2$$

$$\text{MSE}_2(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\overline{p}_{n,i} - p\right)^2, \text{MSE}_3(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widetilde{p}_{n,i} - p\right)^2,$$

$$\text{MSE}_4(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widehat{\lambda}_{n,i} - \lambda\right)^2, \text{MSE}_5(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widehat{p}_{n,i} - p\right)^2.$$

4. compute the coverage probabilities and coverage lengths given by

$$\text{CP}_1(n) = \frac{1}{10000}\sum_{i=1}^{10000} I\left\{\ell_{1,i} < \lambda < u_{1,i}\right\}, \text{CP}_2(n) = \frac{1}{10000}\sum_{i=1}^{10000} I\left\{\ell_{2,i} < \lambda < u_{2,i}\right\},$$

$$\text{CP}_3(n) = \frac{1}{10000}\sum_{i=1}^{10000} I\left\{\ell_{3,i} < \lambda < u_{3,i}\right\}, \text{CP}_4(n) = \frac{1}{10000}\sum_{i=1}^{10000} I\left\{\ell_{4,i} < p < u_{4,i}\right\},$$

$$\text{CL}_1(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(u_{1,i} - \ell_{1,i}\right), \text{CL}_2(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(u_{2,i} - \ell_{2,i}\right),$$

$$\text{CL}_3(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(u_{3,i} - \ell_{3,i}\right), \text{CL}_4(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(u_{4,i} - \ell_{4,i}\right),$$

where $I\{\cdot\}$ denotes the indicator function, $(\ell_{1,i}, u_{1,i})$ are the confidence limits of (20) for the i th simulated sample, $(\ell_{2,i}, u_{2,i})$ are the confidence limits of (21) for the ith simulated sample, $(\ell_{3,i}, u_{3,i})$ are the confidence limits of (23) for the i th simulated sample and $(\ell_{4,i}, u_{4,i})$ are the confidence limits of (24) for the ith simulated sample.

We repeated these steps for $n = 10, 11, \ldots, 100$ with $\lambda = 1$ and $p = 1/2$, so computing the five biases, the five mean squared errors, the four coverage probabilities and the four coverage lengths for $n = 10, 11, \ldots, 100$.

Figure 3 shows how the five biases vary with respect to n. Figure 4 shows how the five mean squared errors vary with respect to n. Figure 5 shows how the four coverage probabilities vary with respect to n. Figure 6 shows how the four coverage lengths vary with respect to n. The broken line in Figure 3 corresponds to the biases being zero. The broken line in Figure 5 corresponds to the nominal coverage probability of 0.95.

The following observations can be drawn from the Figures (3, 4, 5, and 6): the biases appear largest for the estimators, (13) and (18); the biases appear smallest for the estimators, (14) and (15); the biases for each estimator either decrease or increase to zero as $n \to \infty$; the mean squared errors appear largest for the estimator, (13); the mean squared errors appear smallest for the estimators, (19) and (14); the mean squared errors for each estimator decrease to zero as $n \to \infty$; the coverage probabilities for the confidence intervals, (20), (21) and (23), appear reasonably close to the nominal level for all n; the coverage probabilities for the confidence interval (24) approach the nominal level with increasing n and they appear reasonably close to the nominal level for all $n \geq 50$; the coverage lengths appear largest for the confidence interval, (21); the coverage lengths appear smallest for the confidence interval, (24); the coverage lengths for each confidence interval decrease to zero as $n \to \infty$. These observations are for only one choice for (λ, p), namely that $(\lambda, p) = (1, 1/2)$. But the results were similar for other choices.

4 An application to real data

The students number of a specific (test) group from the Bachelor program (first year) at IDRAC International Management school (Lyon, France) who followed 60 sessions of courses in marketing covering the period, from 1/9/2012 to 1/4/2013 are time dependent. But after taking difference of every 2 consecutive sessions the resulting data set is a random sample of size 59. Thus, this data set represents the change in number of students between two sessions of course. Figure 7 illustrate the plots of the students number of the specific group and the difference of every two consecutive sessions, respectively. Descriptive statistics for the original data and the difference are presented in following table.

Number of students and the difference

Variable	Sample size	Mean	Median	Standard deviation	Minimum	Maximum
No of students	60	25.35	26	5.80	8	37
Difference	59	0.3898	0	2.97	-5	7

The runs test was applied to both the original data set and the differenced one. The results of the test show that the number of students of the specific group is not a random sample (p-value $\simeq 0$), while after differencing the resulting data set is a random sample (p-value $= 0.7077$). Now, in order to fit the data, we propose our extended Poisson distribution. First, we estimated the distribution parameters by using (18) and (19). It follows that $\hat{\lambda}_n = 2.457627$ and $\hat{p}_n = 0.509434$. The Pearson Chi-square test is performed to test the fitting. The null hypothesis is that the sample comes from the extended Poisson distribution and the alternative hypothesis is that sample does not come from extended Poisson distribution. The results of Pearson Chi-square test are given in the following table.

Pearson chi-square test for extended Poisson distribution

Modalities	Observed	Expected	$(O-E)^2/E$
≤ -4	5	6.245263	0.24829698
-3	3	6.132144	1.59981959
-2	11	7.485444	1.65014952
-1	7	6.091603	0.13546275
0	6	5.678383	0.01821597
1	6	6.325895	0.01678936
2	4	7.773346	1.83166164
3	8	6.367995	0.41825400
≥ 4	8	6.899926	0.63918202
Total	59	59	6.557832

$\sum \frac{(O-E)^2}{E} = 6.557832 < \chi^2_{6,0.95} = 12.5916$ indicating that the extended Poisson distribution fits the data well. Moreover, in order to assess the goodness of fit we followed also a computational approach. Explicitly, we seek to compare the quality adjustment for the observed data of Skellam (S) distribution (Skellam, 1946), Extended Poisson (E-Po), Extended Binomial (EB) distribution (Alzaid and Omair, 2012), and discrete analogue of the Laplace (DAL) distribution (Inusah and Kozubowski, 2006). Indeed, using the data, we estimate parameter values associated to each distiribution, we simulated 1000 series of length 59 from each distribution and we kept the expected relative frequencies for each series. The reported frequencies are the mean over the 1000 series. Results are represented in the following table (see also Figure 7)

Relative frequency of students number and fitted distributions.

Modalities	Observed frequency	E-Po frequency	S frequency	DAL frequency	EB frequency
≤ -4	8.60	8.54	6.78	6.68	1.77
-3	5.05	4.96	6.22	8.42	5
-2	18.50	18.44	17.94	20.33	16.9
-1	11.80	11.76	10.56	10.75	17.86
0	10.15	10.16	11.3	9.45	12.4
1	10.15	10.16	11.24	10.50	11.16
2	6.80	6.77	8.4	5.68	5.34
3	13.60	13.95	13.56	12	14.41
≥ 4	15.35	15.26	14	16.19	15.16
Total	100	100	100	100	100

One can see that the extended Poisson distribution is the more appropriate to fit the data, compared to the other distributions.

Figures

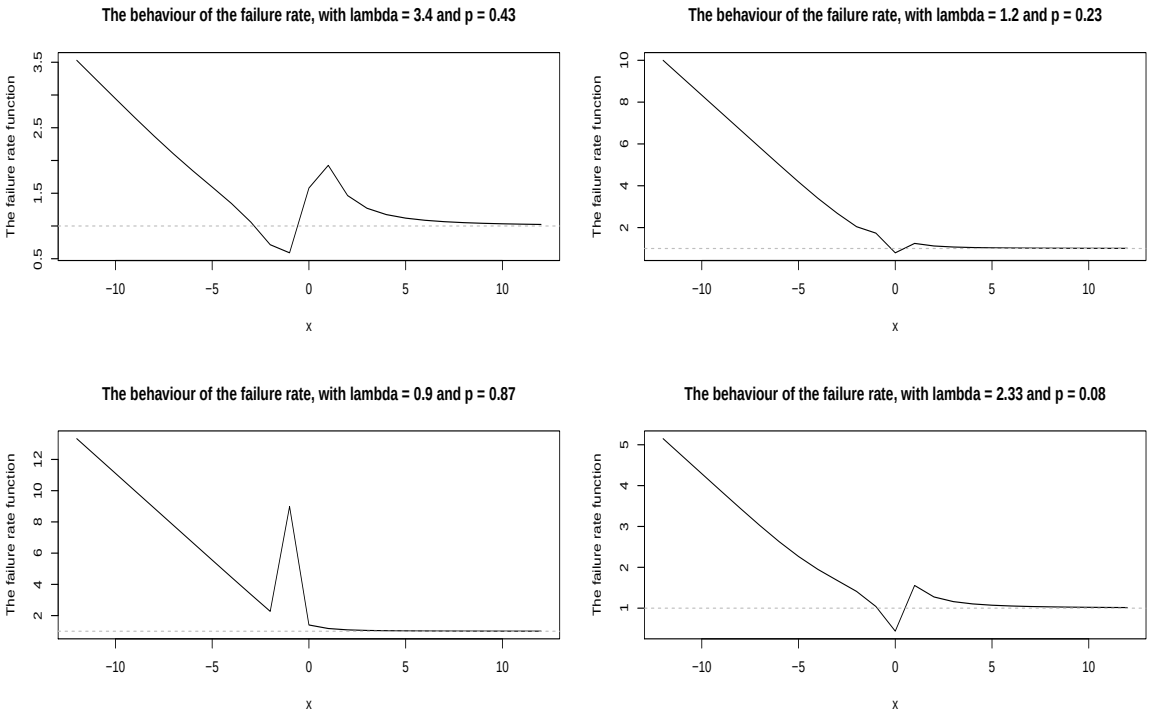

Fig. 1 The behaviour of the failure rate function with different value of λ and p.

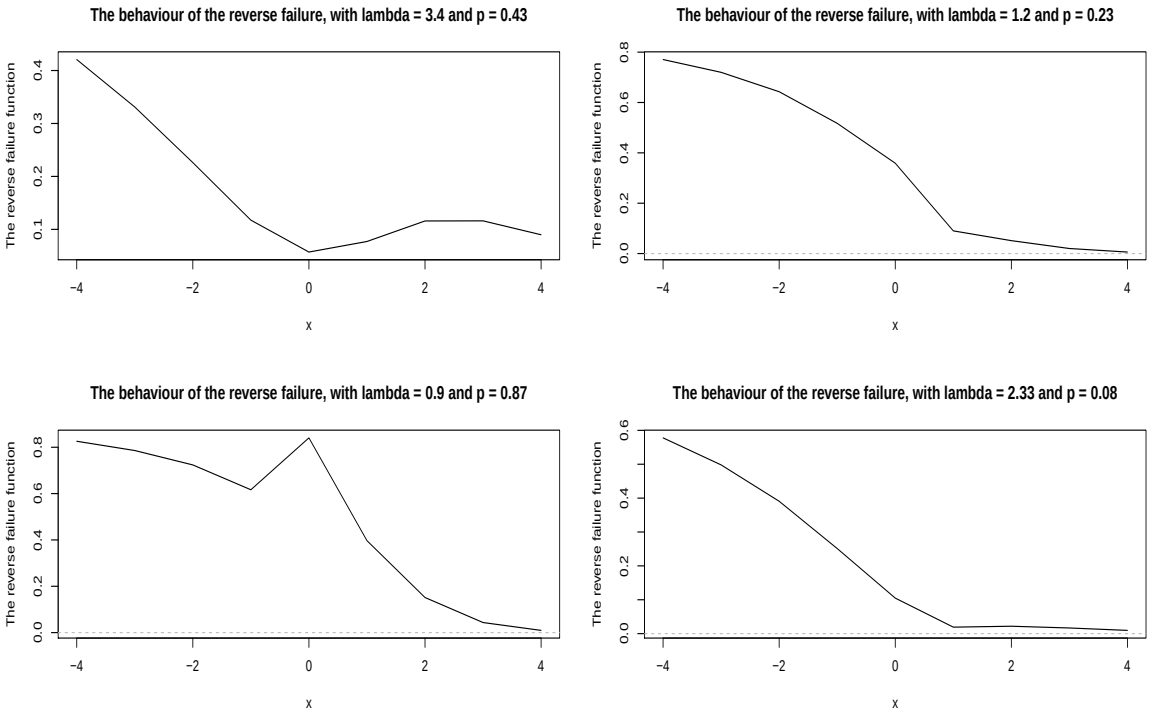

Fig. 2 The behaviour of the reverse failure rate function with different value of λ and p.

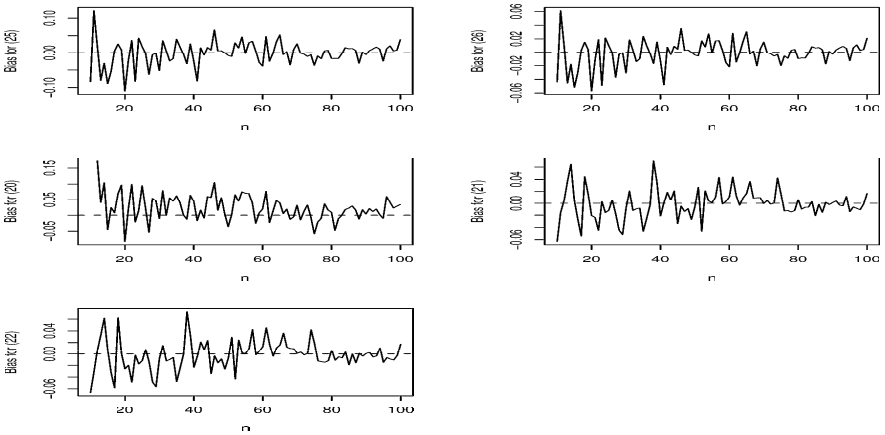

Fig. 3 From top to bottom and from left to right : Biases of (18), (19), (13), (16) and (17) versus $n = 10, 11, \ldots, 100$.

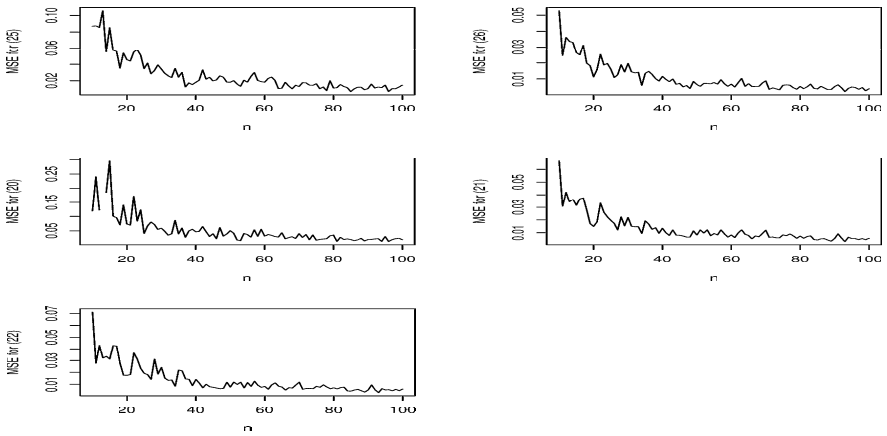

Fig. 4 From top to bottom and from left to right : Mean squared errors of (18), (19), (13), (16) and (17) versus $n = 10, 11, \ldots, 100$.

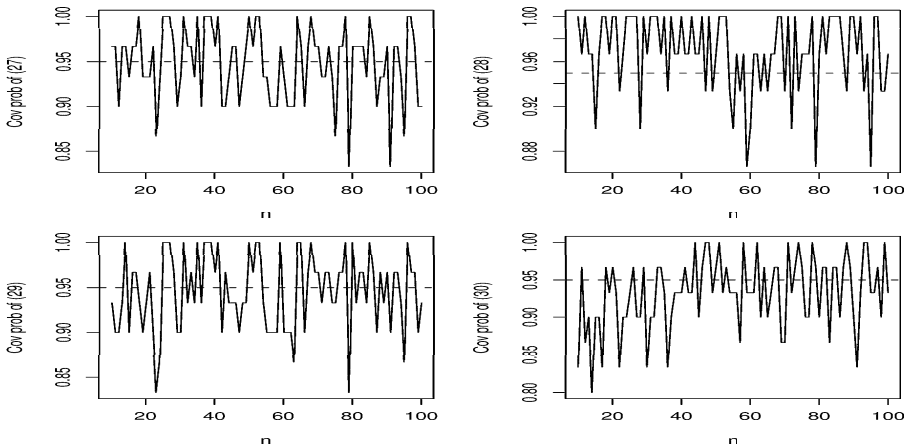

Fig. 5 From top to bottom and from left to right : Coverage probabilities of (20), (22), (23) and (24) versus $n = 10, 11, \ldots, 100$.

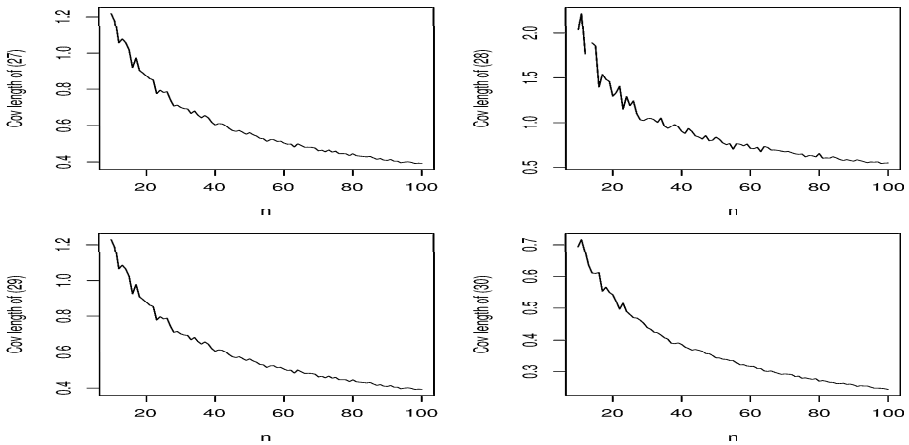

Fig. 6 From top to bottom and from left to right : Coverage lengths of (20), (22), (23) and (24) versus $n = 10, 11, \ldots, 100$.

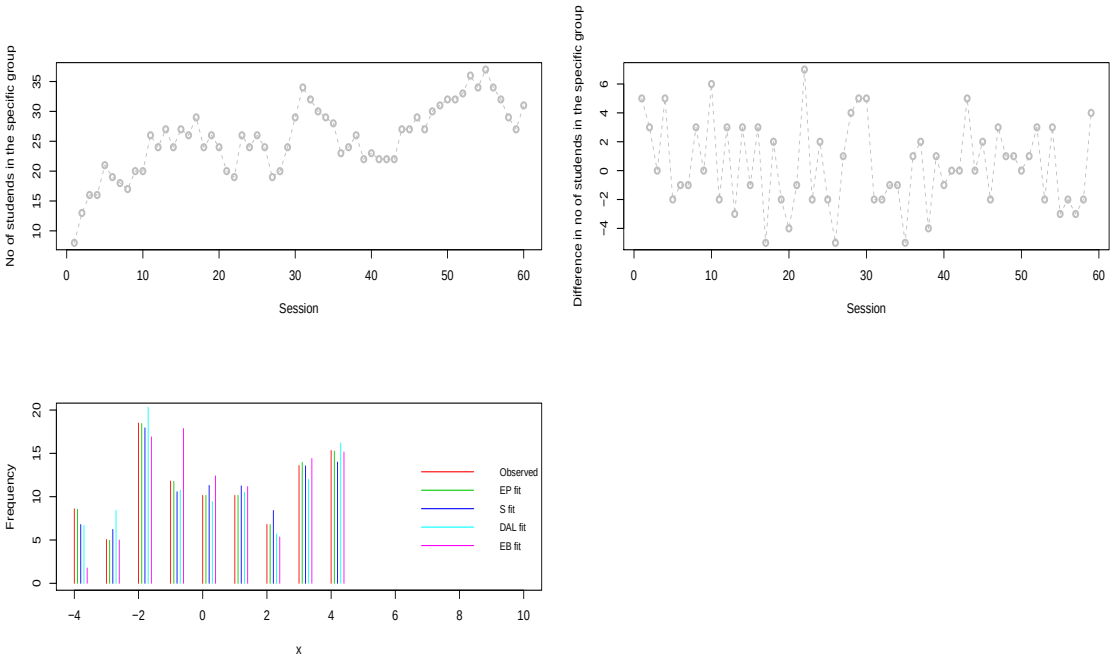

Fig. 7 From top to bottom and from left to right : Number of students of a specific (test) group from the Bachelor program (first year) at IDRAC International Management school (Lyon, France) who followed 60 sessions of courses in marketing from 1/9/2012 to 1/4/2013 ; Difference of number of Number of students of a specific group ; Relative frequency of students number and fitted distributions.

References

1. Ali, M. M., Pal, M., and Woo, J. (2008). Skewed reflected distributions generated by reflected gamma kernel. *Pakistan Journal of Statistics*, **24**, 77-86.
2. Ali, M. M., Pal, M., and Woo, J. (2009). Skewed reflected distributions generated by the Laplace kernel. *Austrian Journal of Statistics*, **38**, 45-58.
3. Ali, M. M. and Woo, J. (2006). Skew-symmetric reflected distributions. *Soochow Journal of Mathematics*, **32**, 233-240.
4. Ali, M. M., Woo, J. and Nadarajah, S. (2010). Some skew symmetric inverse reflected distributions. *Brazilian Journal of Probability and Statistics*, **24**, 1-23.
5. Alzaid, A. A. and Omair, M. A. (2012). An Extended Binomial Distribution with Applications. Communications in Statistics: Theory and Methods, **41**, 3511-3527.
6. Borghi, O. (1965). Sobre una distribución de frecuencias. *Trabajos de Estadistica*, **16**, 171-192.
7. Inusah, S. and Kozubowski, T. J. (2006). A discrete analogue of the Laplace distribution. Journal of Statistical Planning and Inference, **136**, 1090 – 1102
8. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, volume 1, second edition. John Wiley and Sons, New York.
9. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2, second edition. John Wiley and Sons, New York.

10. Rodrigues, J., Cancho, V. G., de Castro, M. and Louzada-Neto, F. (2008). On the unification of the long-term survival models. Statistics and Probability Letters, **39**, 753-759.
11. Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society*, A, **109**, 296.