On the Estimation of Poisson Parameter: An Alternative Approach

Khairul Islam[1], Tanweer Shapla[2]

[1,2] Department of Mathematics, Eastern Michigan University, Ypsilanti, Michigan 48197

Abstract
Poisson distribution plays an important role in modeling rare events data. In this paper, a new estimator of the Poisson parameter has been proposed by using moment generating function. Some important characteristics of the estimator are studied. The performance of the new estimator has been compared with the one using the maximum likelihood method, theoretically and empirically. An empirical study and a real-life application suggest that the proposed new estimator is more efficient than usual estimator.

Key Words: Moment generating function, method of moments, Poisson parameter, relative efficiency

1. Introduction

Poisson distribution is well known for modeling rare events data. The estimation of Poisson parameter using maximum likelihood method appears in any standard book of statistics, e.g., see Walpole, R. E. et al. (2012), Hogg, R.V., McKean, J.W. and Craig, A.T. (2013), Casella, G. and Berger, R.L. (2002), etc. We say that a discrete random variable X follows a Poisson distribution with parameter μ if the probability mass function is given by

$$p(x) = P(X = x) = \frac{e^{-\mu}\mu^x}{x!}; x = 0, 1, 2, \cdots; \text{ and } \mu > 0$$

In general, μ is unknown and is estimated using sample data. Let $x_1, x_2, \cdots, x_n$ be a random sample of size n. Then, the maximum likelihood function of $p(x)$ is given by

$$L(\mu) = \frac{e^{-n\mu}\mu^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

Taking logarithm on both sides

$$l(\mu) = logL(\mu) = -n\mu + \sum_{i=1}^{n} x_i \, log(\mu) - \sum_{i=1}^{n} log(x_i!)$$

[1] kislam@emich.edu

Taking derivative of l with respect to μ, and setting equal to zero, a maximum likelihood estimate (MLE) of μ, $\hat{\mu}$ is given by

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

It is easy to see that $\hat{\mu}$ is an unbiased estimate of μ, i.e.,

$$E(\hat{\mu}) = \mu$$

with variance of $\hat{\mu}$ given by

$$V(\hat{\mu}) = \frac{\mu}{n}$$

In the next section, we propose a new estimator of the Poisson parameter μ and study some statistical properties of the proposed estimator.

2. Proposed Estimator

The moment generating function of $X \sim Poisson(\mu)$ is

$$M_X(t) = E(e^{Xt}) = e^{\mu(e^t - 1)}$$

Given a random sample $X_1, X_2, \cdots, X_n$ of size n, the moment generating function of $Y = \sum_{i=1}^{n} X_i$ is given by

$$M_Y(t) = E(e^{Yt}) = E\left(e^{\sum_{i=1}^{n} X_i t}\right) = \prod_{i=1}^{n} e^{\mu(e^t - 1)} = e^{n\mu(e^t - 1)} \text{ ... (1)}$$

Then, by the method of moments, a new estimator of μ, $\tilde{\mu}$ follows from the equation (1) and is given by

$$e^{\sum_{i=1}^{n} X_i t} = e^{n\tilde{\mu}(e^t - 1)}$$

After an algebraic manipulation, we have the following new estimator of μ

$$\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}; t \neq 0$$

3. Properties of New Estimator

In this section we study some properties of the proposed estimator, which we state in terms of the following theorems:

Theorem 3.1 The expected value of $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$ is $E(\tilde{\mu}) = \frac{\mu t}{e^t - 1}$ and if $t \to 0$, then $\tilde{\mu}$ is an unbiased estimate of μ.

Proof: The expected value of $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$ is

$$E(\tilde{\mu}) = \frac{tE(\bar{x})}{e^t - 1} = \frac{\mu t}{e^t - 1}; t \neq 0$$

Taking limit as $t \to 0$ and applying the L' Hospital Rule, we have

$$\lim_{t\to 0} E(\tilde{\mu}) = \lim_{t\to 0} \frac{E(\bar{x})}{e^t} = \mu$$

Theorem 3.2 The bias of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ is $B(\tilde{\mu}) = \frac{\mu(t-e^t+1)}{e^t-1}$ and if $t \to 0$, then bias of $\tilde{\mu}$ is 0.

Proof: The bias of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ is

$$B(\tilde{\mu}) = E(\tilde{\mu}) - \mu = \frac{tE(\bar{x})}{e^t - 1} - \mu = \frac{\mu t}{e^t - 1} - \mu = \frac{\mu(t - e^t + 1)}{e^t - 1}; t \neq 0$$

Taking limit as $t \to 0$ and applying the L' Hospital Rule, we have

$$\lim_{t\to 0} B(\tilde{\mu}) = \lim_{t\to 0} \frac{\mu(1 - e^t)}{e^t} = 0$$

Theorem 3.3 The variance of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ is $V(\tilde{\mu}) = \frac{\mu t^2}{n(e^t-1)^2}$ and if $t \to 0$, then variance of $\tilde{\mu}$ is same as the variance of $\hat{\mu}$.

Proof: The variance of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ is

$$V(\tilde{\mu}) = \frac{t^2 V(\bar{x})}{(e^t - 1)^2} = \frac{t^2 \mu}{n(e^t - 1)^2}; t \neq 0$$

Taking limit as $t \to 0$ and applying the L' Hospital Rule, we have

$$\lim_{t\to 0} V(\tilde{\mu}) = \lim_{t\to 0} \frac{2t\mu}{2n\,(e^t-1)(e^t)}$$

$$= \lim_{t\to 0} \frac{2\mu}{2n\{(e^t-1)(e^t)+(e^t)(e^t)\}}$$

$$= \frac{2\mu}{2n\{(1-1)(1)+(1)(1)\}}$$

$$= \frac{\mu}{n}$$

$$= V(\hat{\mu})$$

Theorem 3.4 The mean square error (MSE) of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ is $MSE(\tilde{\mu}) = \frac{\mu t^2 + n\mu^2(t-e^t+1)^2}{n(e^t-1)^2}$ and if $t \to 0$, then MSE of $\tilde{\mu}$ is the same as the variance of $\hat{\mu}$.

Proof: The MSE of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ is

$$MSE(\tilde{\mu}) = V(\,\tilde{\mu}) + [B(\,\tilde{\mu})]^2$$

$$= \frac{t^2\mu}{n(e^t - 1)^2} + \frac{\mu^2(t - e^t + 1)^2}{(e^t - 1)^2}$$

$$= \frac{t^2\mu + n\mu^2(t-e^t+1)^2}{n(e^t-1)^2}$$

Taking limit as $t \to 0$ and applying the L' Hospital Rule, we have

$$\lim_{t\to 0} MSE(\tilde{\mu}) = \lim_{t\to 0} \frac{2t\mu + 2n\mu^2\,(t-e^t+1)(1-e^t)}{2n\,(e^t-1)(e^t)}$$

$$= \lim_{t\to 0} \frac{2\mu + 2n\mu^2\,\{(t-e^t+1)(-e^t)+(1-e^t)(1-e^t)\}}{2n\{(e^t-1)(e^t)+(e^t)(e^t)\}}$$

$$= \frac{2\mu + 2n\mu^2\,\{(0-1+1)(-1)+(1-1)(1-1)\}}{2n\{(1-1)(1)+(1)(1)\}}$$

$$= \frac{\mu}{n}$$

$$= V(\hat{\mu})$$

Theorem 3.5 The relative efficiency (RE) of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ with respect to $\hat{\mu}$ is

$$RE = \frac{(e^t-1)^2}{t^2 + n\mu(t-e^t+1)^2} \times 100\%$$

Proof: The relative efficiency of $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$ with respect to $\hat{\mu}$ is given by

$$RE = \frac{V(\hat{\mu})}{MSE(\tilde{\mu})} \times 100\%$$

$$= \frac{\frac{\mu}{n}}{\frac{t^2\mu + n\mu^2(t-e^t+1)^2}{n(e^t-1)^2}} \times 100$$

$$= \frac{(e^t-1)^2}{t^2 + n\mu(t-e^t+1)^2} \times 100$$

It is easy to see that as $t \to 0$, $\tilde{\mu}$ and $\hat{\mu}$ are the same. If $t \neq 0$, then there may exist a non-zero t such that

$$MSE(\tilde{\mu}) < V(\hat{\mu})$$

$$\text{or, } t^2 + n\mu(t-e^t+1)^2 < (e^t-1)^2 \;\ldots (2)$$

In section 4 below, we search for some values of t for which the relation (2) holds true and hence find the percent relative efficiency of the proposed estimator $\tilde{\mu}$ with respect to $\hat{\mu}$ using R code.

4. Application: Fitting a Poisson Distribution

Poisson distribution is useful for describing rare, random events such as severe storms. Below in Table 1 are the number of land-falling hurricane in the USA in 98-year period from 1900 to 1997, appeared in Glover, T and Mitchell, K. (2002). We intend to fit model to the dataset using MLE of μ ($\hat{\mu}$) and the proposed estimator $\tilde{\mu}$, and compare the model fitting performance.

Table 1: Number of land-falling hurricane in the USA in 98-year period from 1900 to 1997

Hurricane per year	x_i	0	1	2	3	4	5	6
Frequency	f_i	18	34	24	16	3	1	2

We want to test H_0: Annual number of the USA land-falling hurricane follows a Poisson distribution with an unknown parameter μ.

4.1 Fitting a Poisson Model Using MLE $\hat{\mu}$

In order to fit above data by a Poisson distribution with parameter μ, we first need to estimate μ. The MLE of μ is given by

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{k} f_i\, x_i}{\sum_{i=1}^{k} f_i} = \frac{159}{98} = 1.622\ hurricanes/year$$

Given above estimate, we assume that the number of the USA land-falling hurricane follows a Poisson distribution with parameter $\hat{\mu} = 1.622$, i.e. we have

$$P(x) = \frac{e^{-\hat{\mu}}\hat{\mu}^{x}}{x!} = \frac{e^{-1.622}\,1.622^{\ x}}{x!}; x = 0, 1, 2 \ldots$$

Then, we have the following probabilities:

$$P(0) = \frac{e^{-1.622}\,1.622^{\ 0}}{0!} = 0.198$$

$$P(1) = \frac{e^{-1.622}\,1.622^{\ 1}}{1!} = 0.320$$

$$P(2) = \frac{e^{-1.622}\,1.622^{\ 2}}{2!} = 0.260$$

$$P(3) = \frac{e^{-1.622}\,1.622^{\ 3}}{3!} = 0.140$$

$$P(4) = \frac{e^{-1.622}\,1.622^{\ 4}}{4!} = 0.057$$

$$P(5) = \frac{e^{-1.622}\,1.622^{\ 5}}{5!} = 0.018$$

$$P(\geq 6) = 1 - P(\leq 5) = 1 - 0.994 = 0.006$$

We summarize the observed and expected frequencies in Table 2 below.

Table 2: Observed and expected frequencies of the USA land-falling hurricane based on the MLE estimate

Hurricane per year	x_i	0	1	2	3	4	5	6
Frequency	f_i	18	34	24	16	**3**	**1**	**2**
Expected frequency	e_i	19.404	31.36	25.48	13.72	**5.586**	**1.764**	**0.588**

where $e_i = n \times p_i = 98 \times p_i$

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} = 1.262$$

We computed the chi-square value by amalgamating expected frequencies for four or more hurricanes (5.586+1.764+.588=7.938), since their expected frequencies are less than 5. With this modification, we have a chi-square $df = 5 - 1 - 1(est\ of\ mu) = 3$, and $chi(df = 3, alpha = 0.05) = 7.81$. Then, by comparing the observed value of chi-square (1.262, p-value $= 0.7382$) with 7.81, we may accept the null hypothesis that the US land-falling hurricane follows a Poisson (1.622) distribution.

4.2 Fitting a Poisson Model Using the Proposed Estimate $\tilde{\mu}$

Now, let us assess the goodness of fit using the new estimate $\tilde{\mu}$. By trial and error method, it is easy to find a value of t for which we can estimate μ by $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$. In order to achieve this, we execute a search of t using the following code in R, where the statements after the symbol # are R comment lines.

```
n=98;                               # Sample size of the given example

m=1.622;                            # MLE estimate, the sample mean for given example

t=seq(0.0001,0.0251,0.0001);        # Values of t for the search, in an arbitrary neighborhood
                                          of zero

re=c();                             # Empty storage for relative efficiency

k=length(t);                        # Length of vector in the search of t

for (j in 1:k){                     # Beginning of the loop of the search

a<-t[j]^2+n*m*(t[j]-exp(t[j])+1)^2;      # Left quantity in equation (2)

b<-(exp(t[j])-1)^2;                      # Right quantity in equation (2)

ifelse (a<b,{t[j]=t[j];re[j]=b/a*100},{t[j]=0;re[j]=0})

                                         # Computing relative efficiency satisfying (2)

}                                        # End of the loop of the search for a given t

re                                  # View relative efficiency for all possible values of t

t                                   # View possible values of t in the search
```

plot(t,re) *# Plotting relative efficiency versus value of t*

Given the search above, let us consider a value of $t = 0.0125$. Then, we have

$$\tilde{\mu} = \frac{t\bar{x}}{e^t - 1} = \frac{0.0125 * 1.622}{e^{0.0125} - 1} = 1.612$$

$$P(0) = \frac{e^{-1.612}\, 1.612^{\,0}}{0!} = 0.199$$

$$P(1) = \frac{e^{-1.612}\, 1.612^{\,1}}{1!} = 0.322$$

$$P(2) = \frac{e^{-1.612}\, 1.612^{\,2}}{2!} = 0.259$$

$$P(3) = \frac{e^{-1.612}\, 1.612^{\,3}}{3!} = 0.139$$

$$P(4) = \frac{e^{-1.612}\, 1.612^{\,4}}{4!} = 0.056$$

$$P(5) = \frac{e^{-1.612}\, 1.612^{\,5}}{5!} = 0.018$$

$$P(\geq 6) = 1 - P(\leq 5) = 1 - 0.994 = 0.006$$

Table 3: Observed and expected frequencies of the USA land-falling hurricane based on the proposed estimate

Frequency	f_i	18	34	24	16	**3**	**1**	**2**
Expected frequency	e_i	19.502	31.556	25.382	13.622	**5.488**	**1.764**	**0.588**

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} = 1.227$$

As before, we computed the chi-square value by amalgamating expected frequencies for four or more hurricanes (5.488+1.764+.588=7.84), since their expected frequencies are less than 5. With this modification, we have a chi-square $df = 5 - 1 - 1(est\ of\ mu) = 3$, and $chi(df = 3, alpha = 0.05) = 7.81$. Then, by comparing the observed value of chi-square (1.227, p-value $= 0.7465$) with 7.81, we may accept the null hypothesis that the US land-falling hurricane follows a Poisson (1.612) distribution.

Note that with the new estimate, we accept the null hypothesis that the US land-falling hurricane follows a Poisson (1.612) distribution with a little stronger evidence (p-value $= 0.7465$). Therefore, the new proposed estimator provides a better fit of Poisson distribution to the given example case.

5. A simulation Study Investigating the Effect of Sample Size and t on Relative Efficiency

In this section, we consider a simulation study to understand the effect of the sample size and the value of t on the relative efficiency of the proposed estimate as compared to the MLE. We consider two fixed values of parameter μ at 0.25 and 0.5, arbitrarily and sample size ranging between 5 and 30 in an increment of 5, denoted as $n \in [5, 30, @\ 5]$. For each combination of μ and n, we consider values of t between a and b with an increment of 0.01, denoted as $t \in [a, b, @\ 0.01]$, where $a = 0.01$ and values of b are evaluated using the search so as to satisfy (2) and are reported along with the relative efficiency for a given combination of μ and n in the Table 4 below:

Table 4: Relative efficiency of proposed estimate as compared to the maximum likelihood estimate for varying sample size and t

μ	n	$t \in [0.01, b, @0.01]$	re
0.25	5	$t \in [0.01, 3.47, @0.01]$	$100.07 \leq re \leq 180.00$
	10	$t \in [0.01, 1.50, @0.01]$	$100.45 \leq re \leq 140.00$
	15	$t \in [0.01, 1.00, @0.01]$	$100.61 \leq re \leq 126.66$
	20	$t \in [0.01, 0.75, @0.01]$	$100.95 \leq re \leq 119.98$
	25	$t \in [0.01, 0.61\ @0.01]$	$100.36 \leq re \leq 116.00$
	30	$t \in [0.01, 0.50\ @0.01]$	$100.99 \leq re \leq 113.33$
0.50	5	$t \in [0.01,1.50, @0.01]$	$100.45 \leq re \leq 140.00$
	10	$t \in [0.01, 0.76, @0.01]$	$100.20 \leq re \leq 120.00$
	15	$t \in [0.01, 0.51, @0.01]$	$100.37 \leq re \leq 113.33$
	20	$t \in [0.01, 0.38, @0.01]$	$100.75 \leq re \leq 110.00$
	25	$t \in [0.01, 0.31\ @0.01]$	$100.23 \leq re \leq 108.00$
	30	$t \in [0.01, 0.26\ @0.01]$	$100.12 \leq re \leq 106.66$

We also provide graphs in Figures 1(a) - 1(f) and 2(a) - 2(f), of estimated efficiency resulting from the simulation so as to understand the effect of sample size (n) and t on the efficiency of the proposed estimate compared to the ML estimate. From these graphs, it is evident that there is a unique value of t for a given interval of t for which the relative efficiency is maximum given the sample size n and the fixed parameter μ.

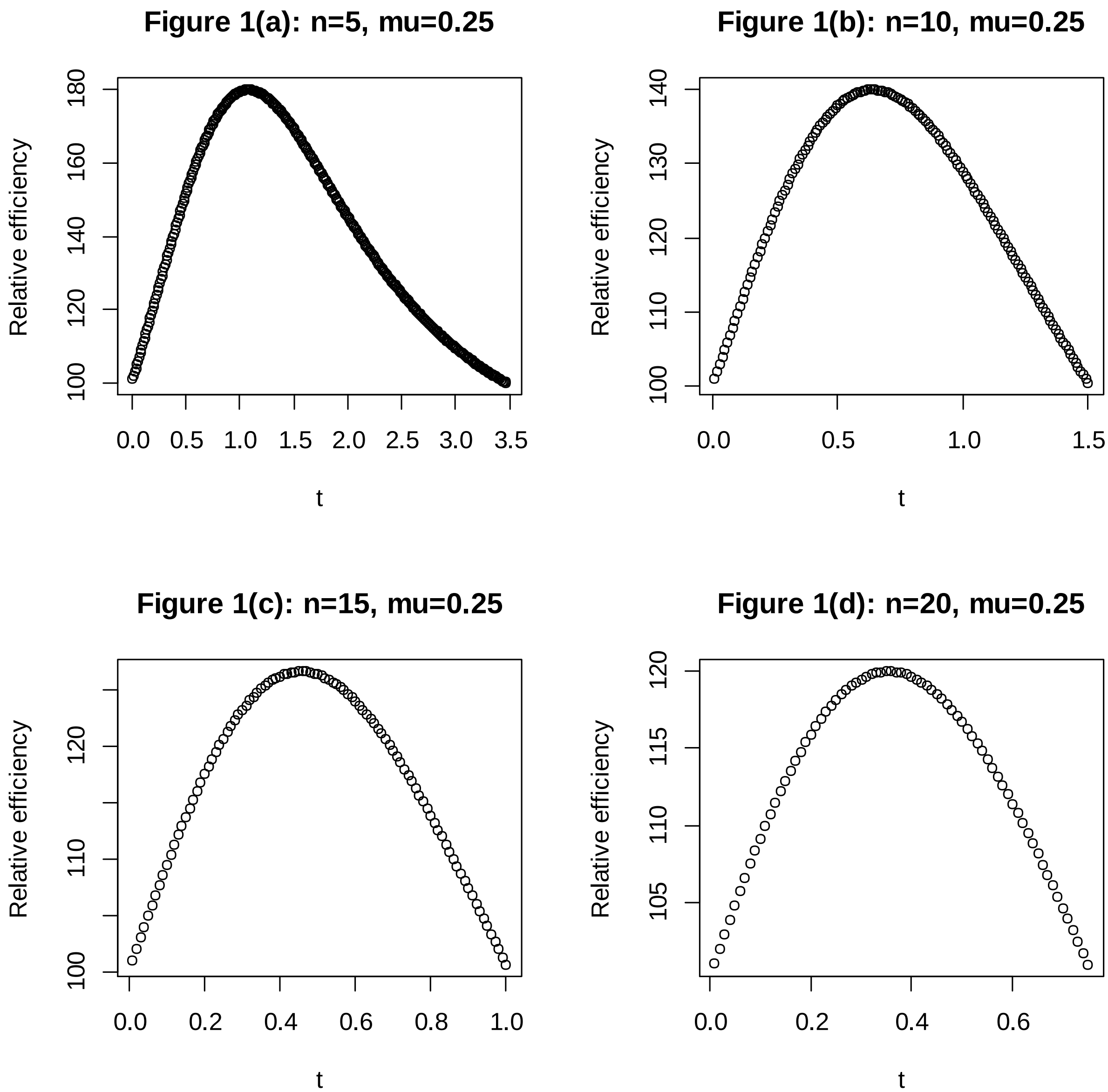


Figure 1(a): n=5, mu=0.25
Relative efficiency
100
120
140
160
180
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
t
Figure 1(b): n=10, mu=0.25
Relative efficiency
100
110
120
130
140
0.0
0.5
1.0
1.5
t
Figure 1(c): n=15, mu=0.25
Relative efficiency
100
110
120
0.0
0.2
0.4
0.6
0.8
1.0
t
Figure 1(d): n=20, mu=0.25
Relative efficiency
105
110
115
120
0.0
0.2
0.4
0.6
t

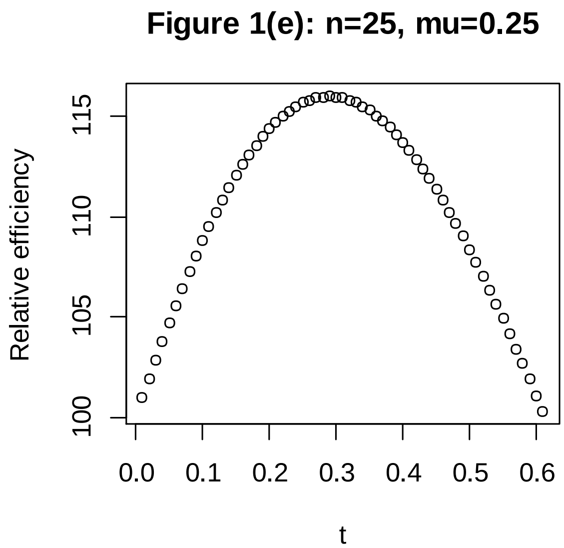


Figure 1(e): n=25, mu=0.25
Relative efficiency
100
105
110
115
0.0
0.1
0.2
0.3
0.4
0.5
0.6
t

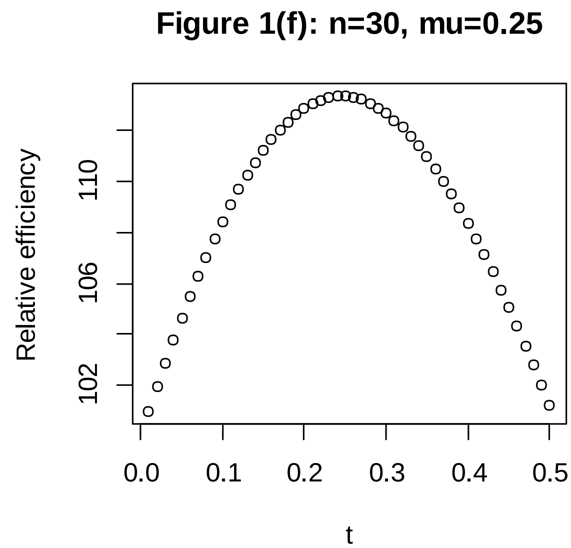


Figure 1(f): n=30, mu=0.25
Relative efficiency
102
106
110
0.0
0.1
0.2
0.3
0.4
0.5
t

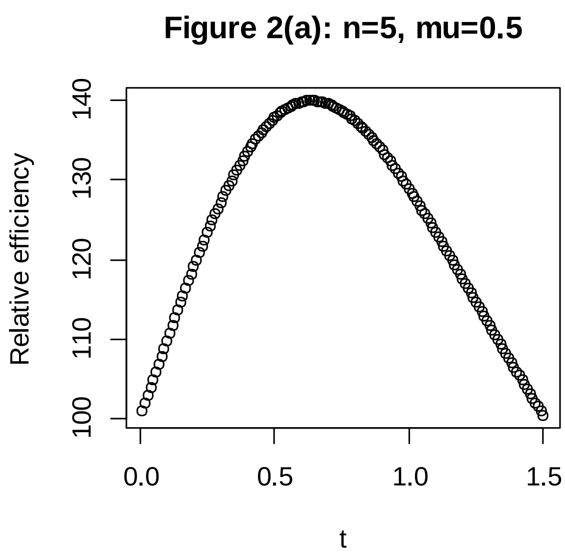


Figure 2(a): n=5, mu=0.5
Relative efficiency
100
110
120
130
140
0.0
0.5
1.0
1.5
t

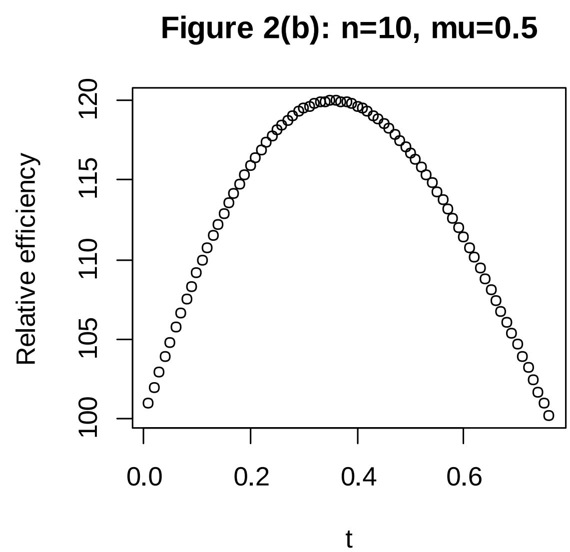


Figure 2(b): n=10, mu=0.5
Relative efficiency
100
105
110
115
120
0.0
0.2
0.4
0.6
t

Figure 2(c): n=15, mu=0.5

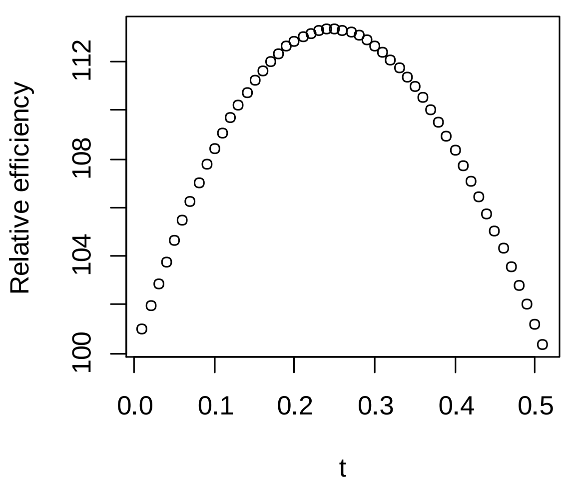


Figure 2(d): n=20, mu=0.5

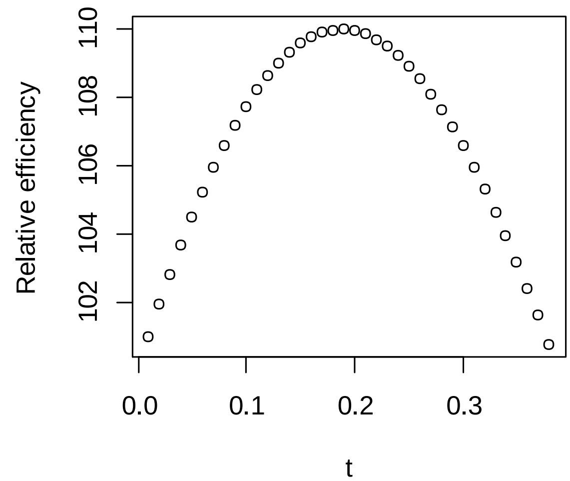


Figure 2(e): n=25, mu=0.5

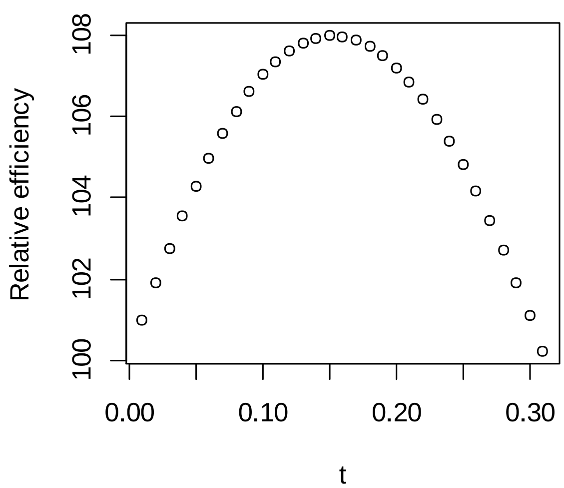
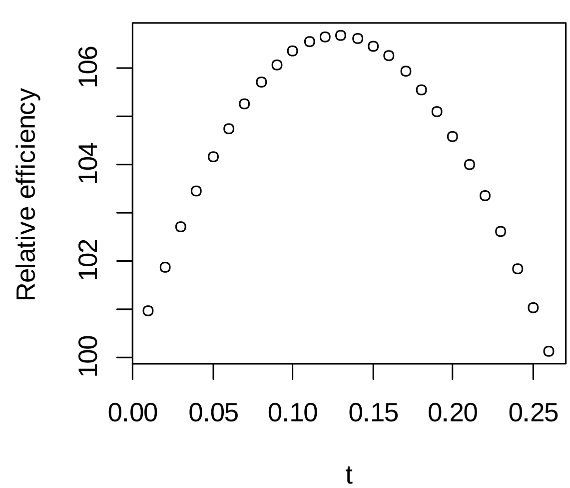


Figure 2(f): n=30, mu=0.5



6. Results and Discussion

We write program in R to search for values of t and the relative efficiency of the proposed estimator of Poisson parameter as compared to the MLE estimator. It appears that the values of t for the example data model remain positive for relative efficiency to be more than 100% for the proposed estimator compared to the MLE estimator. In simulation study, we only need to search positive values of t nearing to 0 for relative efficiency more than 100% for the proposed estimator. Theoretically, since the proposed estimate is unbiased as $t \to 0$, we wish to achieve efficiency as well as nearing unbiased estimate by choosing values of t nearing 0. For example, when $\mu = 0.25$ and the sample size $n = 5$, the relative efficiency of the proposed estimate ranges from 100.07 to 180 when t ranges from 0.01 to 1.50 with an increment of 0.01. However, when $\mu = 0.25$ and the sample size $n = 10$, the relative efficiency ranges from 100.45 to 140 when t ranges from 0.01 to 0.76 with an increment of 0.01. From the reported results, it appears that for a fixed parameter, lower sample size provides better efficiency for the proposed estimate, which makes sense because as sample size gets larger, the values of $MSE(\tilde{\mu})$ and $V(\hat{\mu})$ both get smaller so as to lead to the equally efficient estimates $\tilde{\mu}$ and $\hat{\mu}$. It also follows that relative efficiency of the proposed estimate is better when the value of μ is fixed at a lower value 0.25 than 0.50. Therefore, the proposed estimate is efficiently applicable to the rare events.

7. Concluding Remarks

We proposed a new estimate, $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$, $t \neq 0$, for estimating the unknown Poisson parameter μ using mgf. It appears that the new estimator is a constant multiple of the MLE of μ . Some properties of the new estimator such as Expected value, Bias, MSE, Variance and RE have been studied. As $t \to 0$, the new estimator is unbiased, and MSE and Variance are identical to the variance of the MLE. By searching values of t nearing 0, we can have the higher relative efficiency of the proposed estimate as compared to the ML estimate, $\hat{\mu} = \bar{x}$. The new estimator has been justified using an example, where the new estimate provides a better fit to the data compared to the MLE estimate. In simulation study, it appears that the proposed estimator has much higher relative efficiency as compared to the MLE for smaller sample size. We write program in R to search for the range of t and range of relative efficiency (RE) of the proposed estimate as compared to MLE, which will provide a guide to implement the new method. Given the facts of empirical simulation study and a real-life application to the land-falling hurricane in the USA, we could conclude that the proposed new estimate is more efficient than usual ML estimate for values of t nearing 0, and therefore, we recommend the new method of estimation for fitting Poisson model to rare events data and the estimation of Poisson parameter.

References

Walpole, R.E., Myers, R.H., Myers, S.L. and Ye, K. (2012). Probability & Statistics for Engineers and Scientists. Ninth Edition, Prentice Hall.

Hogg, R.V., McKean, J.W. and Craig, A.T. (2013). Introduction to Mathematical Statistics. Seventh Edition, Prentice Hall.

Casella, G. and Berger, R.L. (2002). Statistical Inference. Second Edition, Duxbury.

Glover, T and Mitchell, K. (2002). An Introduction to Biostatistics. McGraw-Hill Higher Education, First Edition, ISBN: 0-07-241841-9.