

# Estimation of the Poisson Parameter with Moment Generating Method

Khairul Islam<sup>1</sup> & Tanweer Shapla<sup>2</sup>

<sup>1,2</sup>Department of Mathematics and Statistics, Eastern Michigan University, USA.

Correspondence: Khairul Islam, Department of Mathematics and Statistics, Eastern Michigan University, Ypsilanti, MI 48197, USA.

Received: September 2, 2018 Accepted: October 8, 2018 Online Published: October 15, 2018

doi:10.5539/ijsp.v7n6p113

URL: <https://doi.org/10.5539/ijsp.v7n6p113>

## Abstract

A new estimator of the Poisson parameter is proposed using the moment generating function. Some statistical properties of the proposed estimator are studied. The performance of the new estimator is compared with the maximum likelihood estimator (MLE) via examples and simulation in terms of goodness of fit and relative efficiency. Simulation and examples to real-life data suggest that the new estimator has higher relative efficiency compared to the MLE, while both are comparable in goodness of fit. The R program utilized in all computation and simulation is incorporated to facilitate the implementation of the new estimator in computation.

**Keywords:** maximum likelihood estimator, moment generating function, poisson parameter, relative efficiency

## 1. Introduction

Poisson distribution is well known for modeling rare events data. For applications of the Poisson distribution, one could refer to Black (2012), Bortkewitsch (1898), De Veaux, Velleman and Bock (2006), Doane and Seward (2010), Casella and Berger (2002), Jaggia and Kelly (2012), Letkowski (2014) and Rice (2007). The estimation of Poisson parameter using maximum likelihood method appears in any standard book of statistics, e.g., see Walpole et al. (2012), Hogg, McKean and Craig (2013), Casella and Berger (2002), Rice (2007), etc.

A discrete random variable  $X$  is said to follow a Poisson distribution with parameter  $\mu$  if the probability mass function is given by

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

In general,  $\mu$  is unknown and estimated using a sample. Let  $X_1, X_2, \dots, X_n$  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

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

Taking derivative of  $l$  with respect to  $\mu$ , and setting equal to zero, a maximum likelihood estimator (MLE) of  $\mu$ ,  $\hat{\mu}$  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 estimator of  $\mu$ , i.e.,

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

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

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

While the estimator  $\hat{\mu}$  of  $\mu$  by the MLE method is unbiased, there is a possibility that other form of estimator of  $\mu$ , even biased, might have smaller variance. Indeed, in classical statistics, it is well-known that estimators might be biased but have increased accuracy (i.e., smaller variance), which is termed as bias-variance trade-off. For example, one might refer to James-Stein estimator (Stein, 1956; James and Stein, 1961) or LASSO (Tibshirani, 1996, 1997), where the estimator achieves lower mean square error (MSE) than the ML estimator.

In this paper, we proposed a new estimator using the moment generating function, which is biased but provides an increased efficiency compared to the MLE estimator. Method of moments is widely used in different areas of statistics, such as causal inference (e.g., Lu, 2016).

### 2. Proposed Estimator

In this section, we propose a new estimator of the Poisson parameter  $\mu$  using the moment generating function. The moment generating function of  $X \sim \text{Poisson}(\mu)$  is

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

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

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

By the method of moments, the proposed estimator of  $\mu$ ,  $\tilde{\mu}$  follows from the solving the equation

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

After an algebraic manipulation of (1), we have the following new estimator  $\tilde{\mu}$  of  $\mu$

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

A similar method of estimation exists in literature; for example, see Sidhu, Tailor and Singh (2009).

### 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  $t \rightarrow 0$ , then  $\tilde{\mu}$  is an unbiased estimate of  $\mu$ .

**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  $t \rightarrow 0$ , then bias of  $\tilde{\mu}$  is 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  $t \rightarrow 0$ , then variance of  $\tilde{\mu}$  is same as the variance of  $\hat{\mu}$ .

**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  $t \rightarrow 0$ , then MSE of  $\tilde{\mu}$  is the same as the variance of  $\hat{\mu}$ .

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

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

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

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

or,  $t^2 + n\mu(t - e^t + 1)^2 < (e^t - 1)^2$  (2)

We can easily search for values of  $t$ , for selected values of  $\mu$  and  $n$ , satisfying the relation (2) and estimate the percent relative efficiency of the proposed estimator  $\tilde{\mu}$  with respect to  $\hat{\mu}$ .

In section 4, we provide an example of a Poisson distributional fit using the two estimators  $\hat{\mu}$  and  $\tilde{\mu}$ . We have utilized an R program to search for values of  $t$  while using  $\tilde{\mu}$  for estimating  $\mu$  and assessing goodness of fit.

In section 5, we perform an empirical study for assessing relative efficiency of the estimator  $\tilde{\mu}$  compared to  $\hat{\mu}$ .

#### 4. Applications: Fitting Poisson Distributions to Real-Life Data

In this section, we fit a Poisson distribution model to the number of land-falling hurricane in the USA in 98-year period from 1900 to 1997, appeared in Glover and Mitchell (2002). The data in Table 1, refers to number of hurricanes per year ( $x_i$ ) and their frequencies ( $f_i$ ).

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

$x_i$	0	1	2	3	4	5	6
$f_i$	18	34	24	16	3	1	2

We wish to test the null hypothesis  $H_0$ : The annual number of the US land-falling hurricane follows a Poisson distribution with an unknown parameter  $\mu$ , using estimators  $\hat{\mu}$  and  $\tilde{\mu}$ .

Note that an MLE of  $\mu$  is given by

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

Then, the number of the US land-falling hurricane is assumed to follow a Poisson distribution with parameter  $\hat{\mu} = 1.622$  given by

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

The expected frequency  $e_i(\hat{\mu})$  under the above model is estimated by

$$e_i(\hat{\mu}) = n \times p_i(x; \hat{\mu}) = 98 \times p_i(x; \hat{\mu})$$

On the other hand, for the proposed new estimator  $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$ , we execute a search of  $t$  satisfying the equation (2) using an R program. By searching, we consider a value of  $t = 0.0125$ . Then, we have

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

The R program that we have used for the search is provided in the Appendix.

Using the estimator  $\tilde{\mu}$ , the Poisson distribution to hurricane data and the expected frequency  $e_i(\tilde{\mu})$ , take the form:

$$p(x; \tilde{\mu}) = \frac{e^{-\tilde{\mu}} \tilde{\mu}^x}{x!} = \frac{e^{-1.612} 1.612^x}{x!}; x = 0, 1, 2 \dots$$

$$e_i(\tilde{\mu}) = n \times p_i(x; \tilde{\mu}) = 98 \times p_i(x; \tilde{\mu})$$

The estimated expected frequency corresponding to the observed frequency  $f_i$  using two estimators  $\hat{\mu}$  and  $\tilde{\mu}$  are provided in Table 2.

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

$x_i$ :	0	1	2	3	4	5	6
$f_i$ :	18	34	24	16	<b>3</b>	<b>1</b>	<b>2</b>
$e_i(\hat{\mu})$ :	19.404	31.36	25.48	13.72	<b>5.586</b>	<b>1.764</b>	<b>0.588</b>
$e_i(\tilde{\mu})$ :	19.502	31.556	25.382	13.622	<b>5.488</b>	<b>1.764</b>	<b>0.588</b>

The value of the Chi-squared test statistic under the MLE is given by

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

We computed the chi-square value by amalgamating expected frequencies for four or more hurricanes per year so that the expected frequency for each cell is at least 5. With this modification, we have a chi-square  $df = 5 - 1 - 1 = 3$ , 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.

The value of the Chi-squared test statistic under the proposed estimator is given by

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

As before, 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 using both the estimators we accept the null hypothesis that the US land-falling hurricane follows a Poisson with mean 1.622 using the MLE estimator and mean 1.612 using the proposed estimator, with comparable  $p$ -value. Therefore, the proposed estimator is better (higher  $p$ -value) or at least as good as the MLE estimator in goodness of fit of Poisson distribution to the given example case.

**5. Relative Efficiency of the Proposed Estimator**

In this section, we consider an empirical study to evaluation relative efficiency ( $RE$ ) of the proposed estimator  $\tilde{\mu}$  compared to the MLE  $\hat{\mu}$  for varying values of sample size  $n$  and  $t$ . We consider four fixed values of estimator  $\hat{\mu}$  at 0.25, 0.50, 0.75 and 1.25, chosen arbitrarily, and sample size ranging between 5 and 50 at  $n=5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45$  and 50. For each combination of  $\hat{\mu}$  and  $n$ , we consider values of  $t$  between  $a$  and  $b$  with an increment of 0.01, denoted by  $t \in [a: b@0.01]$ , where  $a = 0.01$  and values of  $b$  are evaluated using the search so as to satisfy the equation (2). The estimated  $RE$  of  $\tilde{\mu}$  compared to the MLE  $\hat{\mu}$  are reported in Table 3 for varying sample size  $n$  and  $t$ .

Table 3. The estimated  $RE$  of the proposed estimator  $\tilde{\mu}$  compared to the MLE  $\hat{\mu}$  for varying sample size  $n$  and  $t$ , along with an optimum value of  $t$ ,  $t_{opt}$

$\hat{\mu}$	$n$	$t \in [0.01, b, @0.01]$	$t_{opt}$	$RE$
0.25	5	$t \in [0.01, 3.47@0.01]$	1.08	$100.07 \leq RE \leq 180.00$
	6	$t \in [0.01, 2.66@0.01]$	0.95	$100.01 \leq RE \leq 166.67$
	7	$t \in [0.01, 2.20@0.01]$	0.84	$100.28 \leq RE \leq 157.14$
	8	$t \in [0.01, 1.90@0.01]$	0.76	$100.17 \leq RE \leq 150.00$
	9	$t \in [0.01, 1.68@0.01]$	0.70	$100.04 \leq RE \leq 144.44$
	10	$t \in [0.01, 1.50@0.01]$	0.64	$100.45 \leq RE \leq 140.00$
	15	$t \in [0.01, 1.00@0.01]$	0.46	$100.61 \leq RE \leq 126.66$
	20	$t \in [0.01, 0.75 @0.01]$	0.35	$100.95 \leq RE \leq 119.98$
	25	$t \in [0.01, 0.61 @0.01]$	0.29	$100.36 \leq RE \leq 116.00$
	30	$t \in [0.01, 0.50 @0.01]$	0.25	$100.99 \leq RE \leq 113.33$
	35	$t \in [0.01, 0.44 @0.01]$	0.21	$100.99 \leq RE \leq 113.33$
	40	$t \in [0.01, 0.38 @0.01]$	0.19	$100.99 \leq RE \leq 113.33$
	45	$t \in [0.01, 0.34 @0.01]$	0.17	$100.99 \leq RE \leq 113.33$
	50	$t \in [0.01, 0.31 @0.01]$	0.15	$100.99 \leq RE \leq 113.33$
0.50	5	$t \in [0.01, 1.50@0.01]$	0.64	$100.45 \leq RE \leq 140.00$
	6	$t \in [0.01, 1.25@0.01]$	0.55	$100.39 \leq RE \leq 133.33$
	7	$t \in [0.01, 1.07@0.01]$	0.48	$100.62 \leq RE \leq 128.57$
	8	$t \in [0.01, 0.94@0.01]$	0.43	$100.51 \leq RE \leq 125.00$
	9	$t \in [0.01, 0.84, @0.01]$	0.39	$100.35 \leq RE \leq 122.22$
	10	$t \in [0.01, 0.76, @0.01]$	0.35	$100.20 \leq RE \leq 120.00$
	15	$t \in [0.01, 0.51, @0.01]$	0.25	$100.37 \leq RE \leq 113.33$
	20	$t \in [0.01, 0.38, @0.01]$	0.19	$100.75 \leq RE \leq 110.00$
	25	$t \in [0.01, 0.31 @0.01]$	0.15	$100.23 \leq RE \leq 108.00$
	30	$t \in [0.01, 0.26 @0.01]$	0.13	$100.12 \leq RE \leq 106.66$
	35	$t \in [0.01, 0.22 @0.01]$	0.11	$100.42 \leq RE \leq 105.71$
	40	$t \in [0.01, 0.19 @0.01]$	0.10	$100.63 \leq RE \leq 104.99$
	45	$t \in [0.01, 0.17 @0.01]$	0.09	$100.49 \leq RE \leq 104.44$
	50	$t \in [0.01, 0.15 @0.01]$	0.08	$100.72 \leq RE \leq 103.99$

0.75	5	$t \in [0.01, 1.00@0.01]$	0.46	$100.61 \leq RE \leq 126.66$
	6	$t \in [0.01, 0.84@0.01]$	0.39	$100.35 \leq RE \leq 122.22$
	7	$t \in [0.01, 0.72@0.01]$	0.34	$100.56 \leq RE \leq 119.05$
	8	$t \in [0.01, 0.63@0.01]$	0.30	$100.71 \leq RE \leq 116.67$
	9	$t \in [0.01, 0.56@0.01]$	0.27	$100.81 \leq RE \leq 114.81$
	10	$t \in [0.01, 0.51@0.01]$	0.25	$100.37 \leq RE \leq 113.33$
	15	$t \in [0.01, 0.34@0.01]$	0.17	$100.57 \leq RE \leq 108.89$
	20	$t \in [0.01, 0.26 @0.01]$	0.13	$100.12 \leq RE \leq 106.66$
	25	$t \in [0.01, 0.20 @0.01]$	0.10	$100.89 \leq RE \leq 105.33$
	30	$t \in [0.01, 0.17 @0.01]$	0.09	$100.49 \leq RE \leq 104.44$
	35	$t \in [0.01, 0.15 @0.01]$	0.07	$100.05 \leq RE \leq 103.80$
	40	$t \in [0.01, 0.13 @0.01]$	0.07	$100.18 \leq RE \leq 103.31$
	45	$t \in [0.01, 0.11 @0.01]$	0.06	$100.67 \leq RE \leq 102.96$
	50	$t \in [0.01, 0.10 @0.01]$	0.05	$100.53 \leq RE \leq 102.66$
1.25	5	$t \in [0.01, 0.60, @0.01]$	0.29	$100.99 \leq RE \leq 116.00$
	6	$t \in [0.01, 0.51, @0.01]$	0.25	$100.37 \leq RE \leq 113.33$
	7	$t \in [0.01, 0.44, @0.01]$	0.21	$100.24 \leq RE \leq 111.43$
	8	$t \in [0.01, 0.38, @0.01]$	0.19	$100.75 \leq RE \leq 110.00$
	9	$t \in [0.01, 0.34, @0.01]$	0.17	$100.57 \leq RE \leq 108.89$
	10	$t \in [0.01, 0.31, @0.01]$	0.15	$100.23 \leq RE \leq 108.00$
	15	$t \in [0.01, 0.20, @0.01]$	0.10	$100.89 \leq RE \leq 105.33$
	20	$t \in [0.01, 0.15, @0.01]$	0.08	$100.73 \leq RE \leq 104.00$
	25	$t \in [0.01, 0.12 @0.01]$	0.06	$100.61 \leq RE \leq 103.19$
	30	$t \in [0.01, 0.10 @0.01]$	0.05	$100.53 \leq RE \leq 102.66$
	35	$t \in [0.01, 0.09@0.01]$	0.05	$100.07 \leq RE \leq 102.26$
	40	$t \in [0.01, 0.07 @0.01]$	0.04	$100.82 \leq RE \leq 102.00$
	45	$t \in [0.01, 0.07 @0.01]$	0.04	$100.07 \leq RE \leq 101.74$
	50	$t \in [0.01, 0.06 @0.01]$	0.03	$100.34 \leq RE \leq 101.60$

As appears in Table 3, the estimated  $RE$  of the proposed estimator  $\tilde{\mu}$  as compared to the MLE  $\hat{\mu}$  is sensitive to sample size. For example, when  $n = 5$ , the  $RE$  ranges from 100.07 to 180, when  $t$  ranges from 0.01 to 3.47, with an optimum relative efficiency of 180 observed at optimum  $t_{opt} = 1.08$ . However, when  $n = 10$ , the  $RE$  ranges from 100.07 to 140, when  $t$  ranges from 0.01 to 1.50, with an optimum relative efficiency of 140 observed at optimum  $t_{opt} = 0.64$ . Overall, as  $\hat{\mu}$  increases the  $RE$  decreases.

The estimated  $RE$  of the proposed estimator  $\tilde{\mu}$  as compared to the MLE  $\hat{\mu}$  for selected sample sizes between 5 and 50, and  $t$  satisfying equation (2) are presented in Figures 1-4.

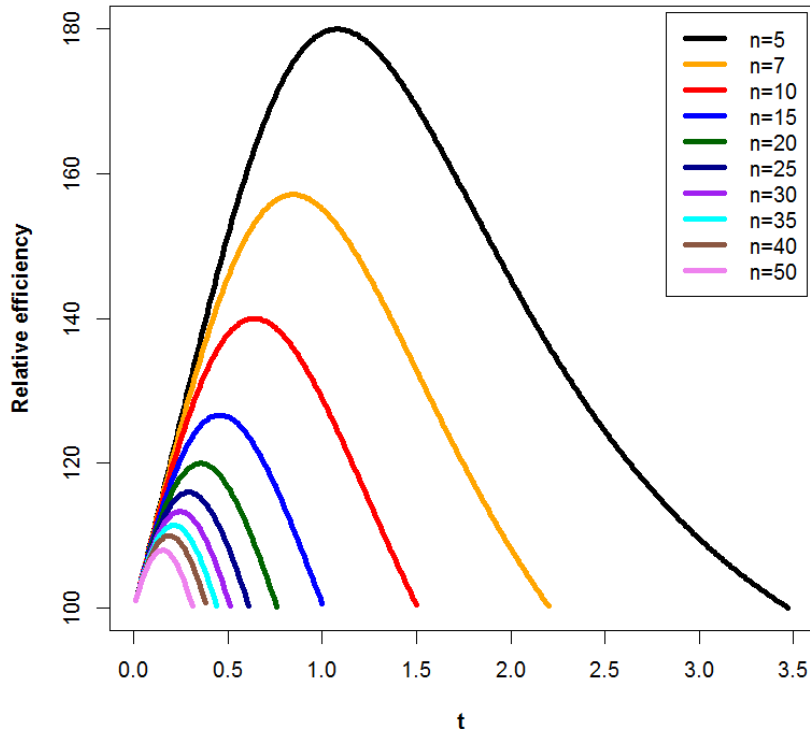


Figure 1. Relative efficiency of  $\tilde{\mu}$  compared to  $\hat{\mu}=0.25$  for varying n and t

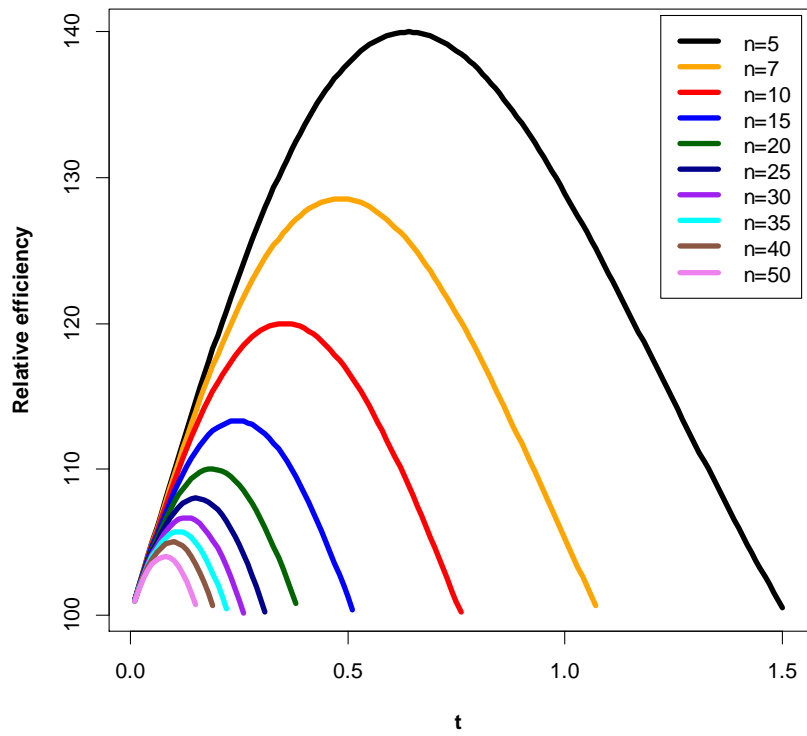


Figure 2. Relative efficiency of  $\tilde{\mu}$  compared to  $\hat{\mu}=0.50$  for varying n and t

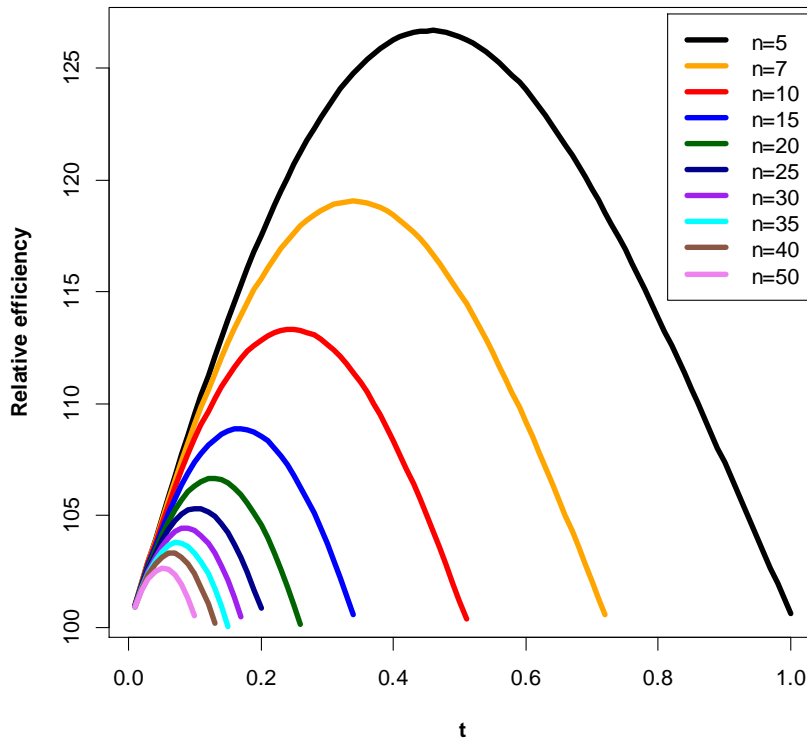


Figure 3. Relative efficiency of compared to  $\lambda=0.75$  for varying  $n$  and  $t$

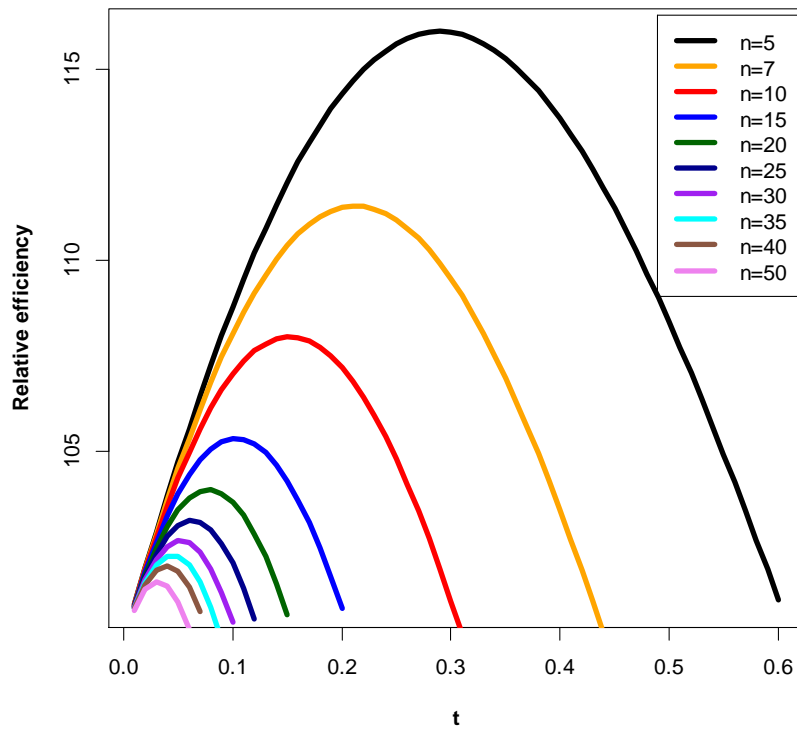


Figure 4. Relative efficiency of compared to  $\lambda=1.25$  for varying  $n$  and  $t$

From Figures 1-4, it follows that we need to search  $t$  in an interval of the positive neighborhood of 0 for higher value of the relative efficiency of the proposed estimator compared to the MLE. Also, the estimated relative efficiency decreases as the sample size increases. It is also evident that there is an optimum value of  $t$ ,  $t_{opt}$ , at which the  $RE$  attains its maximum value, for a given sample size  $n$  and the MLE estimator  $\hat{\mu}$ .

## 6. Results and Discussion

In order to search for  $t$  satisfying equation (2), we have utilized a program written in R, which is incorporated in the Appendix. Once values of  $t$  are obtained, we have evaluated the relative efficiency of the proposed estimator as compared to the MLE. 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 empirical study, we have restricted our search for  $t$  nearing 0 at positive values using trial and error method, by choosing values of  $t$  in some interval. Theoretically, since the proposed estimate is unbiased as  $t \rightarrow 0$ , we wish to achieve efficiency as well as nearing unbiased estimate by choosing values of  $t$  nearing 0. For example, when  $\hat{\mu} = 0.50$  and the sample size  $n = 5$ , the relative efficiency of the proposed estimator ranges from 100.45 to 140 as  $t$  ranges from 0.01 to 1.50, with a maximum relative efficiency of 140 observed at  $t = 0.64$ . However, when the sample size increases to  $n = 10$  ( $\hat{\mu} = 0.50$ ), the relative efficiency ranges from 100.20 to 120, when  $t$  ranges from 0.01 to 0.76, with a maximum  $RE$  of 120 occurring at 0.35. From the reported results, it appears that for a fixed estimator of the Poisson parameter, the lower sample size provides better efficiency for the proposed estimator. It makes sense because as sample size gets larger, the values of  $MSE(\tilde{\mu})$  and  $V(\tilde{\mu})$  both get smaller to lead to equally efficient estimator  $\tilde{\mu}$  and  $\hat{\mu}$ . It also follows that relative efficiency of the proposed estimate is better when the value of MLE estimator  $\hat{\mu}$  is fixed at a lower value, a rare event rate. For example, the maximum value of the estimated  $RE$  of the proposed estimator  $\tilde{\mu}$  is 180 when  $\hat{\mu}$  is 0.25 and  $n = 5$ , whereas the maximum estimated  $RE$  decreases to 116 when the value of  $\hat{\mu}$  decreases to 1.25 and  $n = 5$ . Therefore, the proposed estimate is efficiently applicable to the rare events data.

## 7. Concluding Remarks

We have proposed a new estimator,  $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$ ,  $t \neq 0$ , for estimating the unknown Poisson parameter  $\mu$  using the moment generating function. It appears that the new estimator is a constant multiple of the MLE of  $\mu$ . Some properties of the new estimator such as *Expected value*, *Bias*, *MSE*, *Variance* and *RE* have been studied. As  $t \rightarrow 0$ , 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 estimator as compared to the MLE  $\hat{\mu} = \bar{x}$ . The new estimator has been justified using an example, where the new estimator provides a better fit or at least as good as a fit similar to MLE. In empirical study, it appears that the proposed estimator has much higher relative efficiency as compared to the MLE for smaller sample size and lower value of the MLE. We write program in R to search for the range of  $t$  and range of relative efficiency (RE), along with an optimum value of  $t$  at which relative efficiency is maximum, which will provide a guide to implement the new method. Given the facts of the empirical 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 the usual MLE estimator. Therefore, we recommend the new method of estimation for fitting Poisson model to rare events data and the estimation of Poisson parameter.

## References

- Black, K. (2012). *Business Statistics for Contemporary Decision Making*. New York, NY: John Wiley and Sons, Inc.
- Bortkewitsch, L. (1898). *Das Gesetz der Kleinen Zahlen*. Leipzig, Germany: Teubner.
- Casella, G., & Berger, R. L. (2002). *Statistical Inference*. Second Edition, Duxbury.
- De Veaux, R. D., Velleman, P. F., & Bock, D. E. (2006). *Intro to Stats*. Boston, MA: Addison Wesley, Pearson Education, Inc.
- Doane, D., & Seward, L. (2010). *Applied Statistics in Business and Economics*, 3rd Edition, McGraw-Hill.
- Glover, T., & Mitchell, K. (2002). *An Introduction to Biostatistics*. McGraw-Hill Higher Education, First Edition, ISBN: 0-07-241841-9.
- Hogg, R. V., McKean, J. W., & Craig, A. T. (2013). *Introduction to Mathematical Statistics*. Seventh Edition, Prentice Hall.
- Jaggia, S., & Kelly, A. (2012). *Business Statistics - Communicating with Numbers*. New York, NY: McGraw-Hill Irwin.
- James, W., & Stein, C. (1961). Estimation with quadratic loss, *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, 1, 361–379.
- Letkowsky, J. (2014). Developing Poisson probability distribution applications in a cloud. *Journal of Case Research in*



*Business and Economics*, 5.

Lu, J. (2016). Covariate adjustment in randomization-based causal inference for 2k factorial designs. *Statistics & Probability Letters*, 119, 11-20.

Rice, J. A. (2007). *Mathematical Statistics and Data Analysis*, Cengage Learning.

Sidhu, S. S., Tailor, R., & Singh, S. (2009). On the Estimation of Population Proportion. *Applied Mathematical Sciences*, 3(35), 1739-1744.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate distribution, *Proc. Third Berkeley Symp. Math. Statist. Prob., I*, 197–206.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the lasso. *Journal of the Royal Statistical Society. Series B (methodological)*. Wiley, 58(1), 267–88.

Tibshirani, R. (1997). The lasso Method for Variable Selection in the Cox Model. *Statistics in Medicine*, 16, 385–395.

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

**Appendix**

**PROOF OF THEOREM 3.1.** 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  $t \rightarrow 0$  and applying the L' Hospital Rule, we have

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

**PROOF OF THEOREM 3.2.** 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  $t \rightarrow 0$  and applying the L' Hospital Rule, we have

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

**PROOF OF THEOREM 3.3.** The variance of  $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$  is

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

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

$$\begin{aligned} \lim_{t \rightarrow 0} V(\tilde{\mu}) &= \lim_{t \rightarrow 0} \frac{2t\mu}{2n(e^t - 1)(e^t)} \\ &= \lim_{t \rightarrow 0} \frac{2\mu}{2n\{(e^t-1)(e^t)+(e^t)(e^t)\}} \\ &= \frac{2\mu}{2n\{(1-1)(1)+(1)(1)\}} \\ &= \frac{\mu}{n} \\ &= V(\hat{\mu}) \end{aligned}$$

**PROOF OF THEOREM 3.4.** The MSE of  $\tilde{\mu} = \frac{t\bar{x}}{e^t-1}$  is

$$\begin{aligned} MSE(\tilde{\mu}) &= V(\tilde{\mu}) + [B(\tilde{\mu})]^2 \\ &= \frac{t^2\mu}{n(e^t - 1)^2} + \frac{\mu^2(t - e^t + 1)^2}{(e^t - 1)^2} \\ &= \frac{t^2\mu + n\mu^2(t - e^t + 1)^2}{n(e^t - 1)^2} \end{aligned}$$

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

$$\begin{aligned}\lim_{t \rightarrow 0} \text{MSE}(\hat{\mu}) &= \lim_{t \rightarrow 0} \frac{2t\mu + 2n\mu^2 (t - e^t + 1)(1 - e^t)}{2n(e^t - 1)(e^t)} \\ &= \lim_{t \rightarrow 0} \frac{2\mu + 2n\mu^2 \{(t - e^t + 1)(-e^t) + (1 - e^t)(1 - e^t)\}}{2n\{(e^t - 1)(e^t) + (e^t)(e^t)\}} \\ &= \frac{2\mu + 2n\mu^2 \{(0 - 1 + 1)(-1) + (1 - 1)(1 - 1)\}}{2n\{(1 - 1)(1) + (1)(1)\}} \\ &= \frac{\mu}{n} \\ &= V(\hat{\mu})\end{aligned}$$

**PROOF OF THEOREM 3.5.** The relative efficiency of  $\tilde{\mu} = \frac{t\bar{x}}{e^t - 1}$  with respect to  $\hat{\mu}$  is given by

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

and the proof follows immediately after substituting values of  $V(\hat{\mu})$  and  $\text{MSE}(\tilde{\mu})$ .

**R Code to search for  $t$  in the sample data for the proposed estimator  $\tilde{\mu}$**

```
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 in the search, satisfying equation (2)
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)
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
```

**Sample R Program for Evaluating Relative Efficiency**

```
m=0.25; # MLE estimate
n=5;#Sample size
t=seq(0.01,3.47,0.01);#Sequence of t used in the search
k=length(t);
re=c(); #Empty storage for relative efficiency estimates
for (j in 1:k){ a<-t[j]^2+n*m*(t[j]-exp(t[j])+1)^2;
b<-(exp(t[j])-1)^2;
ifelse (a<b,{t[j]=t[j];re[j]=b/a*100},{t[j]=0;re[j]=0})
}
plot(t,re, type="l", col="black", lwd=2, ylab="Relative efficiency",main=expression("Figure 1: Relative efficiency of
"*tilde(mu)*" compared to "*hat(mu)*"=0.25 for varying n and t"))
indx=which(re==max(re))
c(min(re),max(re), t[indx])
rel<-function(n,mu,t, color){
```

```
k=length(t)
re=c();
for (j in 1:k){
a<-t[j]^2+n*m*(t[j]-exp(t[j])+1)^2;
b<-(exp(t[j])-1)^2;
ifelse (a<b,{t[j]=t[j];re[j]=b/a*100},{t[j]=0;re[j]=0})}
lines(t,re,col=color,lwd=2)
indx=which(re==max(re))
c(min(re),max(re), t[indx])}
rel(n=10,mu=0.25, t=seq(0.01,1.50,0.01),col="red");
rel(n=15,mu=0.25, t=seq(0.01,1.00,0.01),col="blue");
rel(n=20,mu=0.25, t=seq(0.01,0.75,0.01),col="darkgreen");
rel(n=25,mu=0.25, t=seq(0.01,0.61,0.01),col="darkblue");
rel(n=30,mu=0.25, t=seq(0.01,0.50,0.01),col="purple");
colors<-c("black","red","blue","darkgreen","darkblue","purple")
labels<-c(" n=5", " n=10"," n=15"," n=20"," n=25"," n=30")
legend("topright", inset=.05,
labels, lwd=2, lty=c(1, 1, 1, 1, 1, 1), col=colors)
```

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).