The MLE and MME Methods of Estimation of the Parameters of a Zero Inflated Poisson Distribution: A Comparative Study

Name: Sambit Das

Roll Number: 455

Supervisor: Prof. Debjit Sengupta

I affirm that I have identified all my sources and that no part of my dissertation paper uses unacknowledged materials.

Signature: Sambit Das

Date: 10/05/2021

Acknowledgement

I wish to offer my heartiest gratitude to my supervisor, Professor Debjit Sengupta, without whose invaluable expertise, mentorship and guidance this project would not have been a success. I shall forever be indebted to him for his constant support throughout the project in every possible way.

I am also grateful to St. Xaviers College (Autonomous), Kolkata for providing me with an opportunity to explore and analyse a very important and relevant topic which I had chosen for my project.

I further extend my gratitude to the other professors of my department for their support, my parents for their encouragement and my friends for their succor in times when I was woebegone and distressed regarding my dissertation.

Contents

Section 1

Introduction and objective

In statistics, count data is a type of data in which the observations can take only the non-negative integer values {0, 1, 2, 3, ...}, and where these integers arise from counting. An individual piece of count data is often termed a count variable. When such a variable is treated as a random variable, the Poisson, binomial and negative binomial distributions provide a good representation of the data.

In real life there exists many count data where the probability of the zero observations are higher than that of the non-zero ones. In such cases distributions like Poisson, Binomial or Negative binomial fail to provide a good representation to the data as it is highly positively skewed. To obtain a good representation of such a zero inflated data, we require zero inflated distributions.

In this project we consider the Zero Inflated Poisson (Z.I.P) distribution which belongs to the zero inflated family of distributions. A zero inflated poisson distribution is denoted by ZIP (π, λ) where π and λ are the two parameters of a ZIP distribution.

Our main objective is to compare mainly two methods of estimation of the parameters of a ZIP (π, λ) distribution, namely the method of maximum likelihood estimation (MLE) and the method of moments estimation (MME). We compare the two estimators in terms of their bias and mean square errors (MSE).

Section 2

The ZIP Distribution

2.1 PMF of a ZIP distribution with some real life examples

The ZIP distribution is a modified version of a regular Poisson distribution which is used to model count data which has an excessive number of zero observations.

It has two parts:

i. One part generates only the value zero.

ii. The other part follows a Poisson distribution with parameter λ.

Let us see how this distribution arises in real life by taking the following example:

Let us consider a data collected on the number of cigarettes smoked by the students of a college in a day.

Let Y be a random variable defined as:

$$Y=\begin{cases} 1\ ,\ if\ the\ student\ does\ not\ smoke\ a\ cigarette\ with\ probability\ \pi \\ 0\ ,\ otherwise\ with\ probability\ 1-\pi \end{cases}$$

Let X be another random variable denoting the number of cigarettes smoked in a day by a smoker.

Clearly, X ~ Poi (λ) where λ is the expected number of cigarettes smoked by a smoker.

This data is zero inflated by the non-smokers as they only contribute to the zero observations. Some of the smokers might also contribute to the zero observations. Thus this data is zero inflated.

Now, let Z be the reported number of cigarettes smoked by a college student in a day.

Then by the theorem of total probability we can say,

$$P(Z = 0) = P(Z = 0|Y = 1) * P(Y = 1) + P(Z = 0|Y = 0) * P(Y = 0)$$

$$= 1 * \pi + P(X = 0) * (1 - \pi)$$

$$= \pi + (1 - \pi)\, e^{-\lambda}$$

For any z>0,

$$P(Z = z) = P(Z = z|Y = 1) * P(Y = 1) + P(Z = z|Y = 0) * P(Y = 0)$$

$$= 0 + P(X = z) * (1 - \pi)$$

$$= (1 - \pi) e^{-\lambda} \frac{\lambda^z}{z!}$$

Thus the PMF of Z can be written as:

$$P(Z = z) = \begin{cases} \pi + (1 - \pi) e^{-\lambda}, & z = 0 \\ (1 - \pi) e^{-\lambda} \dfrac{\lambda^z}{z!}, & z = 1,2,3, \ldots\ldots \end{cases}$$

We can say that Z ~ ZIP(π, λ) where $0 \leq \pi \leq 1$ and $\lambda \geq 0$

Some more real life examples:

1. The number of billionaires living in every single city in the world follows a zero inflated poisson distribution.

   Here we can divide the class of all cities into poor cities and rich cities. The class of poor cities generates only the value zero, while the number of billionaires in the rich cities follows a poisson distribution. The data will be zero inflated by the poor cities as they only contribute to zero observations.

2. <u>The number of exoplanets discovered each year follows a zero inflated poisson distribution.</u>

   We can divide the class of entire exoplanets into the ones which lie within the telescopic reach and the ones which lie beyond the telescopic reach. The data will be zero inflated by the exoplanets lying beyond telescopic reach as they contribute to only zero observations; while the number of exoplanets discovered among the ones lying within telescopic reach follow a poisson distribution.

3. <u>The number of times a machine fails in a month follows a zero inflated poisson distribution.</u>

   There can be machines among the group of machines which are defective. This data will be zero inflated by the non-defective machines which contributes to only zero observations (they will not fail to work). The number of times the defective machines fail in a month will follow a poisson distribution.

2.2 Comparison of ZIP and Poisson distributions

ZIP distribution is a modified version of a regular Poisson distribution used to model zero inflated count data.

The PMF of a ZIP(π, λ) distribution is given by;

$$P(X = x) = \begin{cases} \pi+(1-\pi)e^{-\lambda}, & x = 0 \\ (1-\pi)e^{-\lambda}\ \dfrac{\lambda^x}{x!}, & x = 1,2,3,\dots\dots \end{cases}$$

Where X~ ZIP(π, λ) and $0 \leq \pi \leq 1$ and $\lambda \geq 0$.

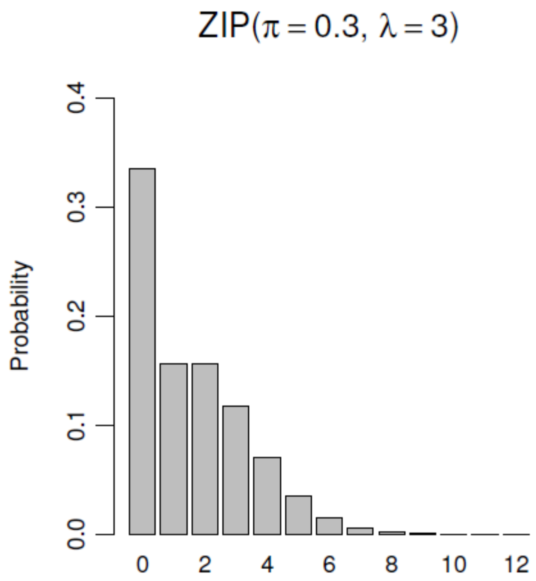
The PMF of a Poisson(λ) distribution is given by;

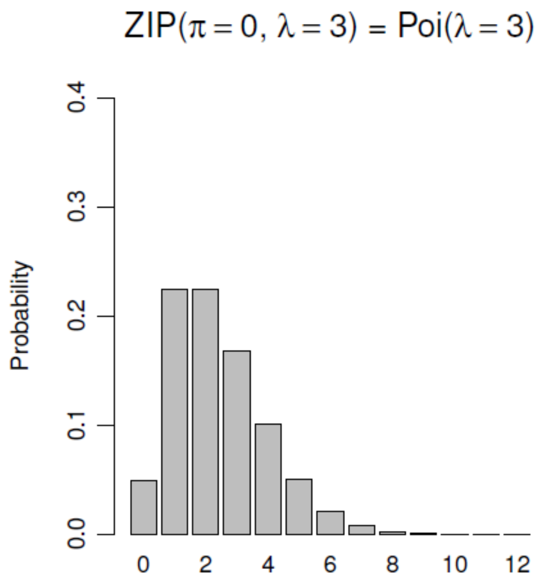
$$P(Y = y) = e^{-\lambda}\ \frac{\lambda^y}{y!}\ , \qquad y = 0,1,2,3,\dots\dots$$

Where Y~Poisson(λ) and λ ≥ 0.

Thus it can be seen that when π=0, the ZIP distribution boils down to a regular Poisson distribution. The function of π is to provide an additional probability boost at the value 0, making the ZIP distribution more positively skewed than the standard poisson distribution.

The following graphs demonstrates this fact:





2.3 Moments of a ZIP distribution

Let X be a random variable where X~ ZIP(π, λ).

The PMF of X is given by,

$$P(X = x) = \begin{cases} \pi+(1-\pi)e^{-\lambda}, & x = 0 \\ (1-\pi)e^{-\lambda}\,\frac{\lambda^x}{x!}, & x = 1,2,3,\ldots\ldots \end{cases} \qquad 0 \leq \pi \leq 1 \text{ and } \lambda \geq 0.$$

We know the expectation of X is given by,

$$E(X) = \sum_{x=0}^{\infty} x * P(X = x)$$

$\Rightarrow E(X) = 0 * (\pi + (1 - \pi)e^{-\lambda}) + \sum_{x=1}^{\infty} x * (1 - \pi)e^{-\lambda} \ \frac{\lambda^x}{x!}$

$\Rightarrow E(X) = (1 - \pi) \sum_{x=1}^{\infty} x * e^{-\lambda} \ \frac{\lambda^x}{x!}$

Let us take Z=X-1

$\Rightarrow E(X) = (1 - \pi) \sum_{z=0}^{\infty} (z + 1) * e^{-\lambda} \ \frac{\lambda^{z+1}}{(z+1)!}$

$\Rightarrow E(X) = (1 - \pi) \ \lambda e^{-\lambda} \sum_{z=0}^{\infty} \ \frac{\lambda^z}{z!}$

$\Rightarrow E(X) = (1 - \pi) \ \lambda e^{-\lambda} e^{\lambda}$

$\Rightarrow \boldsymbol{E}(\boldsymbol{X}) = (\mathbf{1} - \boldsymbol{\pi}) \ \boldsymbol{\lambda}$

<u>Now,</u>

$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) * P(X = x)$

$\Rightarrow E[X(X-1)] = 0 * (\pi + (1 - \pi)e^{-\lambda}) + \sum_{x=1}^{\infty} x(x-1) * (1 - \pi)e^{-\lambda} \ \frac{\lambda^x}{x!}$

$\Rightarrow E[X(X-1)] = (1 - \pi) \sum_{x=1}^{\infty} e^{-\lambda} \ \frac{\lambda^x}{(x-2)!}$

Let us take Z=X-2

$\Rightarrow E[X(X-1)] = (1 - \pi) \sum_{z=0}^{\infty} e^{-\lambda} \ \frac{\lambda^{z+2}}{z!}$

$\Rightarrow E[X(X-1)] = (1 - \pi)\lambda^2 \sum_{z=0}^{\infty} e^{-\lambda} \ \frac{\lambda^z}{z!}$

$\Rightarrow E(X^2) - E(X) = (1-\pi)\lambda^2$ [Since $e^{-\lambda}\ \frac{\lambda^z}{z!}$ is the PMF of a Poisson(λ) distribution. So summing it over it's range will give $\sum_{z=0}^{\infty} e^{-\lambda}\ \frac{\lambda^z}{z!}$ =1]

$\Rightarrow E(X^2) = (1-\pi)\lambda^2 + (1-\pi)\ \lambda$

Thus we have the variance of the distribution as,

$V(X) = E(X^2) - [E(X)]^2$

$\Rightarrow V(X) = \lambda^2 - \lambda^2\pi + \lambda - \lambda\pi - \lambda^2 - \pi^2\lambda^2 + 2\pi\lambda^2$

$\Rightarrow V(X) = \lambda^2\pi + \lambda - \lambda\pi - \pi^2\lambda^2$

$\Rightarrow V(X) = \lambda(1-\pi) + \lambda^2\pi(1-\pi)$

$\Rightarrow V(X) = (\lambda + \lambda^2\pi)(1-\pi)$

$\Rightarrow \boldsymbol{V(X) = \lambda(1 + \lambda\pi)(1-\pi)}$

Thus we have the moments of a ZIP(π, λ) distribution.

The expectation and variance of the distribution are given by;

$\boxed{\boldsymbol{E(X) = (1-\pi)\lambda}}$

$\boxed{\boldsymbol{V(X) = \lambda(1 + \lambda\pi)(1-\pi)}}$

Section 3

MLE and MME methods of estimation

3.1 Maximum Likelihood Estimation (MLE)

Maximum likelihood estimation (MLE) is a method of estimating the parameters of a probability distribution by maximizing a likelihood function, so that under the assumed statistical model the observed data is most probable.

Consider a sample of n observations from a probability distribution with parameter Θ.

Let $f(x_1, x_2, \ldots, x_n \mid \Theta)$ be the joint probability density function or probability mass function of the sample observations. When Θ is fixed, it may be looked upon as a function of the sample observations and then it gives their pdf or pmf. But when the sample observations $x_1, x_2, \ldots, x_n$ are fixed, it may be looked upon as a function of Θ which is called the **likelihood function** of Θ and is denoted by L(Θ).

In many cases it is convenient to work with *ln* L(Θ) rather than L(Θ). Also *ln* L(Θ) attains its highest value for the same value of Θ as L(Θ) does.

The principle of maximum likelihood estimation is that we take that value as the estimate of Θ, for which L(Θ) or *ln* L(Θ) is maximum.

We can obtain that by taking the derivative of the likelihood function with respect to the parameter Θ. Then equating it to 0 and solving for Θ, we will get our maximum likelihood estimator.

Solving the equation $\frac{dL(\Theta)}{d\Theta} = 0$ we get the MLE of Θ.

3.1.1 Variance of MLE estimator

Let X be a random variable following a distribution with a single parameter Θ.

Let f(x) denote the distribution function of X.

Let T be the maximum likelihood estimator of Θ.

The variance of T is then given by the inverse of the information function I(Θ), where I(Θ) is given by,

$$I(\theta) = E(\frac{\partial L(\theta)}{\partial \theta})^2$$

For ease of computational purposes, we write

$$I(\theta) = -E(\frac{\partial^2 L(\theta)}{\partial \theta^2})$$

Thus, $Var(T) = [I(\theta)]^{-1}$

Now,

Let us consider the case where the population distribution has n (n>1) parameters, say$(\theta_1, \theta_2, \theta_3, \dots, \theta_n)$.

The likelihood function is given by,

$$L(\underline{\theta}) = f(x_1, x_2, \dots .., x_n | \theta_1, \theta_2, \theta_3, \dots, \theta_n)$$

Let $T_1, T_2, \dots ., T_n$ be the maximum likelihood estimators of $\theta_1, \theta_2, \theta_3, \dots, \theta_n$ respectively.

Then the variance covariance matrix of the maximum likelihood estimators is given by the inverse of the information matrix $I(\theta)$.

Here $I(\theta)$ is defined as,

$$[I(\theta)]_{ij} = \begin{cases} -E\left(\frac{\partial^2 L(\theta)}{\partial \theta_i^2}\right) \ldots\ldots\ldots, when\ i = j \\ -E\left(\frac{\partial^2 L(\theta)}{\partial \theta_i \partial \theta_j}\right) \ldots\ldots\ldots, when\ i \neq j \end{cases} \qquad i = 1(1)n; j = 1(1)n$$

Thus the variance-covariance matrix of the estimators is given by $[I(\theta)]^{-1}$.

Thus $Var(T_i) = [I(\theta)]_{ii}^{-1}$; $i = 1(1)n$

3.1.2 MLE estimators of Z.I.P distribution and their variances

Let X be a random variable following a ZIP distribution with parameters π and λ.

We know the PMF of X is given by:

$$f(x) = P(X = x) = \begin{cases} \pi + (1-\pi)e^{-\lambda}, & x = 0 \\ (1-\pi)e^{-\lambda}\ \frac{\lambda^x}{x!}, & x = 1,2,3,\ldots\ldots \end{cases}$$

We consider a random sample $X_1, X_2, \ldots\ldots, X_n$ of size n from the ZIP(π, λ) distribution.

We define another variable *Y* as

$$Y = \sum_{i=1}^{n} Y_i \quad ; i = 1(1)n$$

where Y_i is defined as

$$Y_i = \begin{cases} 1, & X_i = 0 \\ 0, & o.w \end{cases} \quad ; i = 1(1)n$$

Thus *Y* follows a binomial distribution with parameter

$$p = P(X_i = 0) = \pi + (1-\pi)e^{-\lambda}.$$

Thus the joint PMF of $X_1, X_2, \ldots\ldots, X_n$ is given by

$$f(x_1, x_2, \dots, x_n) = (\pi + (1\text{-}\pi)e^{-\lambda})^Y + \prod_{\substack{i=1 \\ x_i \neq 0}}^{n} (1\text{-}\pi)e^{-\lambda}\ \frac{\lambda^{x_i}}{x_i!}$$

We now define the likelihood function of π, λ as

$L(\underline{\theta}) = \ln \{f(x_1, x_2, \dots, x_n | \pi, \lambda)\}$; where $\underline{\theta} = (\pi, \lambda)$

$\Rightarrow L(\underline{\theta}) = Y ln\left(\pi + (1\text{-}\pi)e^{-\lambda}\right) + \ln \left(\prod_{\substack{i=1 \\ x_i \neq 0}}^{n} (1\text{-}\pi)e^{-\lambda}\ \frac{\lambda^{x_i}}{x_i!}\right)$

$\Rightarrow L(\underline{\theta}) = Y ln\left(\pi + (1\text{-}\pi)e^{-\lambda}\right) + \sum_{\substack{i=1 \\ x_i \neq 0}}^{n} \ln \left((1\text{-}\pi)e^{-\lambda}\ \frac{\lambda^{x_i}}{x_i!}\right)$

$\Rightarrow L(\underline{\theta}) = Y ln\left(\pi + (1\text{-}\pi)e^{-\lambda}\right) + [(n - Y) ln(1\text{-}\pi) - (n - Y)\lambda +$
$\sum_{\substack{i=1 \\ x_i \neq 0}}^{n} [x_i ln\lambda - \ln(x_i!)]$

$\Rightarrow L(\underline{\theta}) = Y ln\left(\pi + (1\text{-}\pi)e^{-\lambda}\right) + (n - Y) ln(1\text{-}\pi) - (n - Y)\lambda + n\bar{x} ln\lambda$ -
$\sum_{i=1}^{n} \ln (x_i!)$ ……..**(1.1)**

We thus obtain the likelihood function from equation **(1.1)**

Now we differentiate the likelihood function partially with respect to π and λ and equate them to 0. Solving the equations we get their MLE estimators.

Differentiating $L(\underline{\theta})$ w.r.t π and equating it to 0 we get,

$$\frac{\partial L(\underline{\theta})}{\partial \pi} = \frac{Y(1 - e^{-\lambda})}{\pi + (1-\pi)e^{-\lambda}} - \frac{n - Y}{1 - \pi} = 0$$

$\Rightarrow \frac{Y(1 - e^{-\lambda})}{\pi + (1-\pi)e^{-\lambda}} = \frac{n - Y}{1 - \pi}$ ………**(1.2)**

Differentiating $L(\underline{\theta})$ w.r.t λ and equating it to 0 we get,

$$\frac{\partial L(\theta)}{\partial\lambda} = -\frac{Y(1-\pi)e^{-\lambda}}{\pi+(1-\pi)e^{-\lambda}} - (n-Y) + \frac{n\overline{X}}{\lambda} = 0$$

$$\Rightarrow \frac{Y(1-\pi)e^{-\lambda}}{\pi+(1-\pi)e^{-\lambda}} + (n-Y) = \frac{n\overline{X}}{\lambda} \qquad \text{………}\textbf{(1.3)}$$

Solving equations **(1.1)** and **(1.2)**, we can obtain the maximum likelihood estimators of π and λ.

Now,

To obtain the variances of the estimators, we need to find out the information matrix $I(\theta)$. For that we first compute the second order derivatives of $L(\underline{\theta})$ with respect to π and λ.

Second order derivative of $L(\underline{\theta})$ with respect to π

$$\frac{\partial^2 L(\underline{\theta})}{\partial\pi^2} = \frac{\partial}{\partial\pi}\left(\frac{\partial L(\underline{\theta})}{\partial\pi}\right)$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial\pi^2} = -\frac{Y\left(1-e^{-\lambda}\right)}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2}\left(1-e^{-\lambda}\right) - \frac{(n-Y)}{(1-\pi)^2}$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial\pi^2} = -\frac{Y\left(1-e^{-\lambda}\right)^2}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} - \frac{(n-Y)}{(1-\pi)^2} \qquad \text{……..}\textbf{(1.4)}$$

Second order derivative of $L(\underline{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\lambda}$

$$\frac{\partial^2 L(\underline{\theta})}{\partial \lambda^2} = \frac{\partial}{\partial \lambda}\left(\frac{\partial L(\underline{\theta})}{\partial \lambda}\right)$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial \lambda^2} = -\frac{-\left(\pi+(1-\pi)e^{-\lambda}\right)\left(Y(1-\pi)e^{-\lambda}\right)+\left(Y(1-\pi)e^{-\lambda}\right)(1-\pi)e^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} - \frac{n\overline{X}}{\lambda^2}$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial \lambda^2} = \frac{Y\pi(1-\pi)e^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} - \frac{n\overline{X}}{\lambda^2} \qquad \text{.........}\mathbf{(1.5)}$$

Mixed second order derivatives of $L(\underline{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\pi}$ and $\boldsymbol{\lambda}$

- $$\frac{\partial^2 L(\underline{\theta})}{\partial \pi \partial \lambda} = \frac{\partial}{\partial \lambda}\left(\frac{\partial L(\underline{\theta})}{\partial \pi}\right)$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial \pi \partial \lambda} = \frac{\left(\pi+(1-\pi)e^{-\lambda}\right)Ye^{-\lambda}+Y\left(1-e^{-\lambda}\right)(1-\pi)e^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2}$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial \pi \partial \lambda} = \frac{\pi Ye^{-\lambda}+Y(1-\pi)e^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2}$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial \pi \partial \lambda} = \frac{Ye^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} \qquad \text{.........}\mathbf{(1.6)}$$

- $$\frac{\partial^2 L(\underline{\theta})}{\partial \lambda \partial \pi} = \frac{\partial}{\partial \pi}\left(\frac{\partial L(\underline{\theta})}{\partial \lambda}\right)$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial \lambda \partial \pi} = \frac{\left(\pi+(1-\pi)e^{-\lambda}\right)Ye^{-\lambda}+Y\left(1-e^{-\lambda}\right)(1-\pi)e^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2}$$

$$\Rightarrow \frac{\partial^2 L(\underline{\theta})}{\partial\lambda\partial\pi} = \frac{Ye^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} \qquad \text{.........}\mathbf{(1.7)}$$

Now,

$$E(\overline{X}) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)$$

$$\Rightarrow E(\overline{X}) = \frac{\sum_{i=1}^{n} E(X_i)}{n}$$

$$\Rightarrow E(\overline{X}) = \frac{n*(1-\pi)\,\lambda}{n}$$

$$\Rightarrow E(\overline{X}) = (1-\pi)\,\lambda \qquad \text{.........}\mathbf{(2.1)}$$

Also,

Since Y follows $Bin(n, \pi$+(1- $\pi)e^{-\lambda})$

Thus, (1-Y) follows $Bin(n, 1-\pi$+(1- $\pi)e^{-\lambda})$

Therefore $E(Y) = n(\pi$+(1- $\pi)e^{-\lambda})$ ....**(2.2)**

and $E(n-Y) = n$(1- $\pi)$(1-$e^{-\lambda})$ ....**(2.3)**

We now find out the information matrix $I(\theta)$ by taking expectation of the negative of the second order derivatives of $L(\underline{\theta})$.

- $$E\left[-\frac{\partial^2 L(\underline{\theta})}{\partial\lambda^2}\right] = -\frac{n\left(\pi+(1-\pi)e^{-\lambda}\right)\left(\pi(1-\pi)e^{-\lambda}\right)}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} + \frac{n(1-\pi)\,\lambda}{\lambda^2}$$

$$\Rightarrow E\left[-\frac{\partial^2 L(\underline{\theta})}{\partial\lambda^2}\right] = -\frac{n\left(\pi(1-\pi)e^{-\lambda}\right)}{\left(\pi+(1-\pi)e^{-\lambda}\right)} + \frac{n(1-\pi)}{\lambda} \qquad \textbf{.........(3.1)}$$

- $$E\left[-\frac{\partial^2 L(\underline{\theta})}{\partial\pi^2}\right] = \frac{n\left(\pi+(1-\pi)e^{-\lambda}\right)\left(1-e^{-\lambda}\right)^2}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} + \frac{n(1-\pi)(1-e^{-\lambda})}{(1-\pi)^2}$$

$$\Rightarrow E\left[-\frac{\partial^2 L(\underline{\theta})}{\partial\pi^2}\right] = \frac{n\left(1-e^{-\lambda}\right)^2}{\left(\pi+(1-\pi)e^{-\lambda}\right)^2} + \frac{n(1-e^{-\lambda})}{(1-\pi)} \qquad \textbf{........(3.2)}$$

- $$E\left[-\frac{\partial^2 L(\underline{\theta})}{\partial\pi\partial\lambda}\right] = E\left[-\frac{\partial^2 L(\underline{\theta})}{\partial\lambda\partial\pi}\right] = -\frac{ne^{-\lambda}}{\left(\pi+(1-\pi)e^{-\lambda}\right)} \qquad \textbf{.......(3.3)}$$

We can find the inverse of the information matrix using R software. Further discussion has been done in section 4.

3.2 Method of Moments Estimation (MME):

Method of moments estimation is another method of estimating the parameters of a population distribution. In this method the population (theoretical) moments are equated to the corresponding sample moments. The number of equations are same as the number of parameters to be estimated. The equations are then solved to get the estimates of those parameters. The resulting estimators are called the method of moment estimators of the parameters.

For example, if we want to estimate k unknown parameters $\theta_1, \theta_2, \theta_3, \dots, \theta_k$ of the distribution of a random variable X. Suppose the first k moments of the distribution can be expressed as functions of the θs.

$\mu_1 \equiv E(X) = g_1(\theta_1, \theta_2, \theta_3, \dots, \theta_k)$

$\mu_2 \equiv E(X^2) = g_2(\theta_1, \theta_2, \theta_3, \dots, \theta_k)$

.

.

.

$\mu_k \equiv E(X^k) = g_k(\theta_1, \theta_2, \theta_3, \dots, \theta_k)$

Let a sample $x_1, x_2, \dots, x_n$ of size n be drawn. We equate the population j-th moment with the sample j-th moment; j=1(1)k.

$$\hat{\mu}_j = \frac{\sum_{i=1}^{n} X_i^k}{n}$$

Solving the j equations simultaneously, we can find the estimates of the k estimators $\theta_1, \theta_2, \theta_3, \dots, \theta_k$.

MME estimators of a ZIP distribution

Let X be a random variable following a ZIP distribution with parameters π and λ. We take a random sample $X_1, X_2, \dots, X_n$ of size n from the population.

To obtain the MME estimators of π and λ we equate the sample mean with the population mean and the sample variance with the population variance.

We define the sample mean as $\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$ and

the sample variance as $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

We know that $E(X) = (1-\pi)\lambda$ and $V(X) = \lambda(1+\lambda\pi)(1-\pi)$

Thus equating $\bar{X}$ with $E(X)$ and s^2 with $V(X)$ we get,

$\bar{X} = (1-\pi)\lambda$ ....(**4.1**)

$$s^2 = \lambda(1+\lambda\pi)(1-\pi) \quad \text{…………}(\mathbf{4.2})$$

Dividing equation (**4.2**) by (**4.1**) we get,

$$\frac{s^2}{\bar{X}} = 1+\pi\lambda \quad \text{…………}(\mathbf{4.3})$$

$$\Rightarrow \frac{s^2}{\bar{X}} + \bar{X} - 1 = \lambda$$

$$\Rightarrow \hat{\lambda}_{MME} = \frac{s^2}{\bar{X}} + \bar{X} - 1 \quad \text{…………}(\mathbf{4.4})$$

From equation (**4.3**) we get,

$$\hat{\pi}_{MME} = \frac{1}{\hat{\lambda}_{MME}}\left(\frac{s^2}{\bar{X}} - 1\right)$$

$$\Rightarrow \hat{\pi}_{MME} = \frac{s^2-\bar{X}}{\bar{X}^2+s^2-\bar{X}} \quad \text{…………}(\mathbf{4.5})$$

Equations (**4.4**) and (**4.5**) gives the method of moments estimators of λ and π respectively.

We observe that when $\bar{X} > s^2$ the values of $\hat{\pi}_{MME}$ and $\hat{\lambda}_{MME}$ can become negative, whereas the actual parameters π and λ are always positive (0≤π≤1; λ≥0).

So we truncate $\hat{\pi}_{MME}$ at 0 when $\bar{X} > s^2$ and $\hat{\lambda}_{MME}$ at $\bar{X}$ when $\bar{X} > s^2$.

So our modified MME becomes:

$$\hat{\pi}_{MME}{}^{c} = \begin{cases} 0 & if\ \bar{X} \geq s^2 \\ \hat{\pi}_{MME} & o.w \end{cases}$$

$$\hat{\lambda}_{MME}{}^{c} = \begin{cases} \bar{X} & if\ \bar{X} \geq s^2 \\ \hat{\lambda}_{MME} & o.w \end{cases}$$

The modified MMEs makes more sense since under a ZIP model V(X) is always greater than E(X). Thus it is expected to have $\bar{X} < s^2$.

3.2.1 Variance of MME estimator

To find out the variance of the MME estimator, we use the delta method.

Delta method

Let $T_1, T_2, \ldots., T_k$ be a k-dimensional statistic such that the asymptotic distribution of $T_1, T_2, \ldots., T_k$ is k-variate normal with mean $\underline{\theta} = \theta_1, \theta_2, \ldots\ldots, \theta_k$ and dispersion matrix $\Sigma = (\sigma_{ij})_{k\times k}$. That is,

$$T_1, T_2, \ldots., T_k \xrightarrow{D} N_k((\theta_1, \theta_2, \ldots\ldots, \theta_k), \Sigma_{k\times k})$$

Further let $f_1, f_2, \ldots., f_q$ be q functions of k variables and each $f_i, i = 1(1)q$ is totally differentiable. Then the asymptotic distribution of $f_1, f_2, \ldots., f_q$ is q-variate normal with means $f_i(\theta_1, \theta_2, \ldots\ldots, \theta_k), i = 1(1)q$ and dispersion matrix $G'\Sigma G$ where

$$G = \frac{\partial f_i}{\partial \theta_j}, i = 1(1)q; j = 1(1)k$$

We now apply the delta method to find out the approximate variances of $\hat{\pi}_{MME}$ and $\hat{\lambda}_{MME}$.

Let us consider two statistics $T_1 = \bar{X}$ and $T_2 = s^2$.

The asymptotic distribution of T_1, T_2 is bivariate normal with mean $\underline{\theta} = \theta_1, \theta_2$ and dispersion matrix $\Sigma = (\sigma_{ij})_{2\times 2}$. That is,

$$T_1, T_2 \xrightarrow{D} N_2((\theta_1, \theta_2), \Sigma_{2\times 2})$$

Here $V(T_1) = \sigma_{11}$

$V(T_2) = \sigma_{22}$

$Cov(T_1, T_2) = \sigma_{12} = \sigma_{21}$

Now,

$V(T_1) = V(\bar{X}) = V(\frac{1}{n}\sum_{i=1}^{n} X_i)$

$= \frac{1}{n^2}\sum_{i=1}^{n} V(X_i)$

$= \frac{1}{n^2}\sum_{i=1}^{n} \lambda(1 + \lambda\pi)(1 - \pi)$

$= \frac{\lambda(1+\lambda\pi)(1-\pi)}{n}$

$\Rightarrow V(T_1) = V(\bar{X}) = \frac{\lambda(1+\lambda\pi)(1-\pi)}{n}$

It can be shown that the approximate variance of T_2 is given by

$V(T_2) = \frac{\mu_4}{n} - \frac{(n-3)\sigma^4}{n(n-1)}$

where $\mu_4 = E[(X - E(X))^4]$ is the fourth order central moment of X

and $\sigma = \lambda(1 + \lambda\pi)(1 - \pi)$.

Also,

$\theta_1 = E(T_1) = E(\bar{X}) = (1 - \pi)\lambda$ (From equation (**2.1**))

$\theta_2 = E(T_2) = E(s^2)$

$$E(s^2) = E[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}]$$

$$= E[\frac{\sum_{i=1}^{n} X_i{}^2}{n-1} - \frac{n}{n-1}\bar{X}^2]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} E(X_i{}^2) - \frac{n}{n-1}E(\bar{X}^2)$$

$$= \frac{1}{n-1}[\sum_{i=1}^{n}(\lambda(1+\lambda\pi)(1-\pi) + \lambda^2(1-\pi)^2)] -$$

$$\frac{n}{n-1}[\frac{\lambda(1+\lambda\pi)(1-\pi)}{n} + \lambda^2(1-\pi)^2]$$

[Since $E\left(X_i{}^2\right) = V(X_i) + E^2(X_i)$

And $E(\bar{X}^2) = V(\bar{X}) + E^2(\bar{X})$]

$$= \frac{n}{n-1}[\lambda(1+\lambda\pi)(1-\pi) + \lambda^2(1-\pi)^2] - \frac{n}{n-1}[\frac{\lambda(1+\lambda\pi)(1-\pi)}{n} +$$

$$\lambda^2(1-\pi)^2]$$

$$= \frac{n}{n-1}[\frac{(n-1)}{n}\lambda(1+\lambda\pi)(1-\pi)]$$

$$= \lambda(1+\lambda\pi)(1-\pi)$$

$$\Rightarrow \theta_2 = \lambda(1+\lambda\pi)(1-\pi)$$

$Cov(T_1, T_2)$

$$= E[(T_1 - E(T_1))(T_2 - E(T_2))]$$

$$= E(\bar{X} - \theta_1)(s^2 - \theta_2)$$

$= E[\bar{X} * s^2 - \theta_2\bar{X} - \theta_1 s^2 + \theta_1\theta_2]$

$= E(\bar{X} * s^2) - \theta_1\theta_2$ [Since $E(\bar{X}) = \theta_1$ and $E(s^2) = \theta_2$]

Now,

$E(\bar{X} * s^2) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i * \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2\right]$

Let $Y_i = X_i - \theta$ [$\theta = E(X_i) = (1-\pi)\lambda$]

Thus, $E(Y_i) = 0$ and $V(Y_i) = V(X_i)$

$\Rightarrow E(\bar{X} * s^2) = E\left[\frac{1}{n(n-1)}\sum_{i=1}^{n}(Y_i + \theta)\sum_{j=1}^{n} {Y_j}^2\right]$

$= \frac{1}{n(n-1)} E\left[\sum_{i=1}^{n}(Y_i + \theta){Y_i}^2 + \sum_{i \neq j}\sum(Y_i + \theta){Y_j}^2\right]$

$= \frac{1}{n(n-1)}\left[\sum_{i=1}^{n}\left[E\left({Y_i}^3\right) + \theta E\left({Y_i}^2\right)\right] + \sum_{i \neq j}\sum\left[E(Y_i)E\left({Y_j}^2\right) + \theta E\left({Y_j}^2\right)\right]\right]$

$= \frac{1}{n(n-1)}[n(\mu_3 + \theta\mu_2) + n(n-1)(\mu_1\mu_2 + \theta\mu_2)]$

[μ_1, μ_2 and μ_3 are the 1st, 2nd and 3rd order central moments of X]

$= \frac{1}{(n-1)}(\mu_3 + \theta\mu_2) + \theta\mu_2$

Thus,

$Cov(T_1, T_2) = \frac{1}{(n-1)}(\mu_3 + \theta\mu_2) + \theta\mu_2 - \theta_1\theta_2$

To compute the variance-covariance matrix of T_1, T_2 we first need to compute the central moments of X. For that we first calculate the raw moments of X.

1st order raw moment

$$\boxed{\mu_1' = E(X) = \lambda(1-\pi)}$$

2nd order raw moment

$$\mu_2' = E(X^2)$$

$$E\big(X(X-1)\big) = \sum_{x=0}^{\infty} x(x-1) * P(X=x)$$

$$= 0 * \pi(1-\pi)e^{-\lambda} + \sum_{x=1}^{\infty} x(x-1) * (1-\pi)\, e^{-\lambda}\frac{\lambda^x}{x!}$$

$$= (1-\pi)\sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^x}{(x-2)!}$$

Let $(x-2) = z$

$$\Rightarrow E(X^2 - X) = (1-\pi)\sum_{z=0}^{\infty} e^{-\lambda}\frac{\lambda^z\lambda^2}{z!}$$

$$\Rightarrow E(X^2) - E(X) = (1-\pi)\lambda^2$$

[Since $e^{-\lambda}\,\frac{\lambda^z}{z!}$ is the PMF of a Poisson(λ) distribution. So summing it over it's range will give $\sum_{z=0}^{\infty} e^{-\lambda}\,\frac{\lambda^z}{z!}$ =1]

$$\Rightarrow E(X^2) = (1-\pi)\lambda^2 + E(X)$$

$\Rightarrow E(X^2) = (1-\pi)\lambda^2 + \lambda(1-\pi)$

$\Rightarrow \mu_2' = E(X^2) = \lambda(1+\lambda)(1-\pi)$

$\boxed{\mu_2' = \lambda(1+\lambda)(1-\pi)}$

<u>3rd order raw moment</u>

$\mu_3' = E(X^3)$

$$E(X(X-1)(X-2)) = \sum_{x=0}^{\infty} x(x-1)(x-2) * P(X=x)$$

$$= 0 * \pi(1-\pi)e^{-\lambda} + \sum_{x=1}^{\infty} x(x-1)(x-2) * (1-\pi)\, e^{-\lambda} \frac{\lambda^x}{x!}$$

$$= (1-\pi) \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^x}{(x-3)!}$$

Let $(x-3) = z'$

$\Rightarrow E(X^3 - 3X^2 + 2X) = (1-\pi) \sum_{z=0}^{\infty} e^{-\lambda} \frac{\lambda^{z'} \lambda^3}{z'!}$

$\Rightarrow E(X^3) - E(3X^2) + E(2X) = (1-\pi)\lambda^3$

[Since $e^{-\lambda} \frac{\lambda^{z'}}{z'!}$ is the PMF of a Poisson(λ)

distribution. So summing it over

it's range will give $\sum_{z=0}^{\infty} e^{-\lambda} \frac{\lambda^{z'}}{z'!}$ =1]

$$\Rightarrow E(X^3) = (1-\pi)\lambda^3 + 3E(X^2) - 2E(X)$$

$$= \lambda^3 - \pi\lambda^3 + 3\lambda(1+\lambda)(1-\pi) - 2\lambda(1-\pi)$$

$$= \lambda^3 - \pi\lambda^3 + 3\lambda^2 - 3\pi\lambda^2 + 3\lambda - 3\pi\lambda - 2\lambda + 2\pi\lambda$$

$$= (1-\pi)\lambda^3 + 3\lambda^2(1-\pi) + \lambda(1-\pi)$$

$$= (1-\pi)(\lambda^3 + 3\lambda^2 + \lambda)$$

$$\boxed{\mu_3' = (1-\pi)(\lambda^3 + 3\lambda^2 + \lambda)}$$

Similarly we can find out the fourth order raw moment of X. It comes out to be

$$\boxed{\mu_4' = (1-\pi)(\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda)}$$

We thus obtain the central moments as:

1st order central moment

$$\boxed{\mu_1 = E\big(X - E(X)\big) = 0}$$

2nd order central moment

$$\mu_2 = \mu_2' - {\mu_1'}^2$$

$$= \lambda(1+\lambda)(1-\pi) - \lambda^2(1-\pi)^2$$

$$= \pi\lambda^2 + \lambda - \pi\lambda - \pi^2\lambda^2$$

$$\boxed{\mu_2 = \pi\lambda^2 + \lambda - \pi\lambda - \pi^2\lambda^2}$$

3rd order central moment

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2{\mu_1'}^3$$

$$= (1-\pi)(\lambda^3 + 3\lambda^2 + \lambda) - 3\lambda(1-\pi)\lambda(1+\lambda)(1-\pi) + 2\lambda^3(1-\pi)^3$$

$$= \lambda - \pi\lambda^3 + 9\pi\lambda^2 - \pi\lambda + 3\pi^2\lambda^3 - 3\pi^2\lambda^2 - 2\pi^3\lambda^3$$

$$\boxed{\mu_3 = \lambda - \pi\lambda^3 + 9\pi\lambda^2 - \pi\lambda + 3\pi^2\lambda^3 - 3\pi^2\lambda^2 - 2\pi^3\lambda^3}$$

4th order central moment

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'{\mu_1'}^2 - 3{\mu_1'}^4$$

$$= (1-\pi)(\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda) - 4(1-\pi)(\lambda^3 + 3\lambda^2 + \lambda)\lambda(1-\pi) +$$
$$6\lambda(1+\lambda)(1-\pi)\lambda^2(1-\pi)^2 - 3\lambda^4(1-\pi)^4$$

$$= \pi\lambda^4 + 4\pi^2\lambda^4 + 6\pi^3\lambda^4 + 3\pi^4\lambda^4 + 30\pi^2\lambda^3 - 6\pi^3\lambda^3 + 3\lambda^2 + \lambda + \pi\lambda^2 -$$
$$\pi\lambda + 4\pi^2\lambda^2$$

$$\boxed{\begin{gathered}\mu_4 = \pi\lambda^4 + 4\pi^2\lambda^4 + 6\pi^3\lambda^4 + 3\pi^4\lambda^4 + 30\pi^2\lambda^3 - 6\pi^3\lambda^3 + 3\lambda^2 \\ +\lambda + \pi\lambda^2 - \pi\lambda + 4\pi^2\lambda^2\end{gathered}}$$

Having obtained the central moments of X, we can get the variance-covariance matrix of T_1 and T_2. We now proceed to obtain the variances of the MME estimators of π and λ.

- **Variance of $\hat{\lambda}_{MME}$**

Consider two statistics $T_1 = \bar{X}$ and $T_2 = s^2$

The asymptotic distribution of T_1, T_2 is bivariate normal with mean $\underline{\theta} = \theta_1, \theta_2$ and dispersion matrix $\Sigma = (\sigma_{ij})_{2\times 2}$. That is,

$$T_1, T_2 \xrightarrow{D} N_2((\theta_1, \theta_2), \Sigma_{2\times 2})$$

The MME of λ is given by

$$\hat{\lambda}_{MME} = \frac{s^2}{\bar{X}} + \bar{X} - 1$$

Let $f_1(T_1, T_2) = \frac{s^2}{\bar{X}} + \bar{X} - 1 = \hat{\lambda}_{MME}$

The asymptotic distribution of $f_1(T_1, T_2)$ is univariate normal with mean $f_1(\theta_1, \theta_2)$ and dispersion matrix ${G_1}'\Sigma G_1$ where

$$G_1 = \frac{\partial f_1}{\partial \theta_i}\ ; i = 1,2$$

$$f_1(T_1, T_2) \xrightarrow{D} N(f_1(\theta_1, \theta_2), {G_1}'\Sigma G_1)$$

$$f_1(\theta_1, \theta_2) = \frac{\theta_2}{\theta_1} + \theta_1 - 1$$

Therefore

$$\frac{\partial f_1}{\partial \theta_1} = -\frac{\theta_2}{{\theta_1}^2} + 1$$

$$\frac{\partial f_1}{\partial \theta_2} = \frac{1}{\theta_1}$$

Thus $G_1 = \left[-\frac{\theta_2}{{\theta_1}^2} + 1 \quad \frac{1}{\theta_1} \right]$

Let $\boxed{\mathbf{a_1} = -\frac{\boldsymbol{\theta_2}}{\boldsymbol{\theta_1}^2} + \mathbf{1}}$ and $\boxed{\mathbf{a_2} = \frac{\mathbf{1}}{\boldsymbol{\theta_1}}}$

Therefore

$$G_1{}'\Sigma G_1 = [a_1 \quad a_2] \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$= \boxed{{a_1}^2\sigma_{11} + 2a_1a_2\sigma_{12} + {a_2}^2\sigma_{22}} \qquad \text{……….(}\mathbf{5.1}\text{)}$$

- **Variance of $\hat{\boldsymbol{\pi}}_{\boldsymbol{MME}}$**

The MME of π is given by,

$$\hat{\pi}_{MME} = \frac{s^2 - \bar{X}}{\bar{X}^2 + s^2 - \bar{X}}$$

Let $f_2(T_1, T_2) = \frac{s^2 - \bar{X}}{\bar{X}^2 + s^2 - \bar{X}} = \hat{\pi}_{MME}$

The asymptotic distribution of $f_2(T_1, T_2)$ is univariate normal with mean $f_2(\theta_1, \theta_2)$ and dispersion matrix $G_2{}'\Sigma G_2$ where

$$G_2 = \frac{\partial f_2}{\partial \theta_i} \,; i = 1,2$$

$$f_2(T_1, T_2) \xrightarrow{D} N(f_2(\theta_1, \theta_2), G_2{}'\Sigma G_2)$$

$$f_2(\theta_1, \theta_2) = \frac{\theta_2 - \theta_1}{{\theta_1}^2 + \theta_2 - \theta_1}$$

Therefore

$$\frac{\partial f_2}{\partial \theta_1} = \frac{-\left({\theta_1}^2+\theta_2-\theta_1\right)-(\theta_2-\theta_1)(2\theta_1-1)}{({\theta_1}^2+\theta_2-\theta_1)^2}$$

$$= \frac{-{\theta_1}^2-\theta_2+\theta_1-2\theta_1\theta_2+\theta_2+2{\theta_1}^2-\theta_1}{({\theta_1}^2+\theta_2-\theta_1)^2}$$

$$= \frac{{\theta_1}^2-2\theta_1\theta_2}{({\theta_1}^2+\theta_2-\theta_1)^2}$$

$$\frac{\partial f_2}{\partial \theta_2} = \frac{{\theta_1}^2+\theta_2-\theta_1-(\theta_2-\theta_1)}{({\theta_1}^2+\theta_2-\theta_1)^2}$$

$$= \frac{{\theta_1}^2}{({\theta_1}^2+\theta_2-\theta_1)^2}$$

Thus $G_2 = \left[\frac{{\theta_1}^2-2\theta_1\theta_2}{\left({\theta_1}^2+\theta_2-\theta_1\right)^2} \quad \frac{{\theta_1}^2}{\left({\theta_1}^2+\theta_2-\theta_1\right)^2}\right]$

Let $\boxed{\mathbf{b_1} = \frac{\boldsymbol{{\theta_1}^2-2\theta_1\theta_2}}{\left(\boldsymbol{{\theta_1}^2+\theta_2-\theta_1}\right)^2}}$ and $\boxed{\mathbf{b_2} = \frac{\boldsymbol{{\theta_1}^2}}{(\boldsymbol{{\theta_1}^2+\theta_2-\theta_1})^2}}$

Therefore

$$G_2{}'\Sigma G_2 = [b_1 \quad b_2]\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}\begin{bmatrix}b_1\\ b_2\end{bmatrix}$$

$= \boxed{{b_1}^2\sigma_{11} + 2b_1b_2\sigma_{12} + {b_2}^2\sigma_{22}}$ ……….(**5.2**)

Defining the terms a_1, a_2, b_1, b_2 and equations (**5.1**) and (**5.2**) in R, we can compute the theoretical variances of the MME of π and λ.

Section 4

Simulation

Algorithm

- Step 1: Draw a random sample from a ZIP (π, λ) distribution of size n=100 for λ=2 and π=0.1. Repeat this 1000 times.

```
n=100; l=2; p=0.1
for (j in 1:1000)
{ y=array(dim=n)
for(i in 1:n)
{ u=runif(1,0,1)
  if(u<p)
  y[i]=0
  else
  y[i]=rpois(1,l)
}
y
```

- Step 2 : Count the number of 0s in each sample drawn and find the sample mean

```
for (j in 1:1000)
{ Z=0
for(i in 1:n)
{ if(y[i]==0)
Z=Z+1
}
Z
```

```
ybar=mean(y)
ybar
```

- Step 3 : Define the score functions for estimating the MLE of π, λ

```
score=function(x)
{
 S=numeric(2)
 S[1]=(Z*(1-exp(-x[2]))/(x[1]+(1-x[1])*exp(-x[2])))-((n-Z)/(1-x[1]))
 S[2]=((Z*(1-x[1])*exp(-x[2]))/(x[1]+(1-x[1])*exp(-x[2])))+(n-Z)-((n*ybar)/x[2])
 S
}
```

- Step 4 : Solve the score equations using nleqslv function to get the j^{th} (j=1(1)1000) MLE estimates of π and λ and calculate their bias and mean square errors.

```
library(nleqslv)
p_est=l_est=array(dim=1)
MSE_p=MSE_l=array(dim=1)
bias_mle_l=bias_mle_p=array(dim=1)
for(j in 1:1000)
{ par_ini=c(p,l)
par_hat=nleqslv(par_ini, score, jac=NULL, method="Newton")$x
p_est[j]=par_hat[1]
l_est[j]=par_hat[2]
bias_mle_p[j]=(p_mle[j]-p)
bias_mle_l[j]=(l_mle[j]-l)
MSE_p[j]=(p_est[j]-p)^2
MSE_l[j]=(l_est[j]-l)^2
}
```

- Step 5 : Define MMEs of π and λ and calculate their bias and MSEs for every sample drawn

```
p_mme=l_mme=array(dim=1)
MSE_p1=MSE_l1=array(dim=1)
bias_mme_l=bias_mme_p=array(dim=1)
for(j in 1:1000)
{ s2=(sum((y-ybar)^2))/n-1
if(ybar>=s2)
{ l_mme[j]=ybar
p_mme[j]=0
}
else
{ l_mme[j]=(s2/ybar)+ybar-1
p_mme[j]=(s2-ybar)/((ybar^2)+s2-ybar)
}
MSE_l1[j]=(l_mme[j]-l)^2
bias_mme_l[j]=(l_mme[j]-l)
MSE_p1[j]=(p_mme[j]-p)^2
bias_mme_p[j]=(p_mme[j]-p)
}
```

- Step 6 : Obtain an estimate of the MLEs of π and λ along with their simulated bias and MSEs by taking the mean of the 1000 simulated values of the estimates, bias and MSEs

```
pi_est=mean(p_mle)
pi_est                    #Estimate of MLE of π
lambda_est=mean(l_mle)
lambda_est                #Estimate of MLE of λ
```

```
bias_lambda_mle=mean(bias_mle_l)
bias_lambda_mle                    #simulated bias of MLE of λ
bias_pi_mle=mean(bias_mle_p)
bias_pi_mle                        #simulated bias of MLE of π
MSE_pi=mean(MSE_p)
MSE_pi                             #simulated MSE of MLE of π
MSE_lambda=mean(MSE_l)
MSE_lambda                         #simulated MSE of MLE of λ
```

- Step 7 : Obtain an estimate of the MMEs of π and λ along with their simulated bias and MSEs by taking the mean of the 1000 simulated values of the estimates, bias and MSEs

```
MME_pi=mean(p_mme)
MME_pi                             #Estimate of MME of π
MME_lambda=mean(l_mme)
MME_lambda                         #Estimate of MME of λ
MSE_lambda_mme=mean(MSE_l1)
MSE_lambda_mme                     #simulated MSE of MME of λ
bias_lambda_mme=mean(bias_mme_l)
bias_lambda_mme                    #simulated bias of MME of λ
MSE_pi_mme=mean(MSE_p1)
MSE_pi_mme                         #simulated MSE of MME of π
bias_pi_mme=mean(bias_mme_p)
bias_pi_mme                        #simulated bias of MME of π
```

- Step 8 : Define the information matrix and obtain the theoretical variances of π and λ by using the inv function (inverse function)

```
library(matlib)
t1=-(n*(p*(1-p)*exp(-l))/(p+(1-p)*exp(-l)))+(n*(1-p)/l)
```

```
t4=(n*((1-exp(-l))^2)/((p+(1-p)*exp(-l))^2))+(n*(1-exp(-l))/(1-p))

t2=t3=-n*exp(-l)/(p+(1-p)*exp(-l))

M=matrix(c(t1,t2,t3,t4),nrow=2,byrow=TRUE)

M1=inv(M)

M1[1,1]                         #Theoritical MSE of MLE of λ

M1[2,2]                         #Theoretical MSE of MLE of π
```

- Step 9 : Calculate the theoretical variance of the MMEs of π and λ by defining the expressions for the variance of the estimators using delta method in R

- Step 10 : Repeat steps 1 to 9 for a different combination of n (n=100,200), π (π=0.1, 0.3, 0.5) and λ (λ=2, 5, 10).

Section 5

Observations and inference

We denote the MLE estimators of π and λ by $\hat{\pi}_{MLE}$ and $\hat{\lambda}_{MLE}$ respectively and the MME estimators of π and λ by $\hat{\pi}_{MME}$ and $\hat{\lambda}_{MME}$ respectively. For a fixed value of n (100, 200) we compare the estimates, bias and MSEs of the MLE and MME estimators of π and λ for varying values of λ (2, 5, 10) and π (0.1, 0.3, 0.5). The observations are presented in a tabulated form below:

$\boxed{\mathbf{n = 100}}$

Table 1 Comparison of MLE and MME estimators of λ for sample size 100

Parameters		**Estimate**	**Estimate**	**S-MSE**	**S-MSE**	**S-Bias**	**S-Bias**	**Theoretical Variance**	**Theoretical Variance**
$\boldsymbol{\pi}$	**λ**	$(\hat{\boldsymbol{\lambda}}_{MLE})$	$(\hat{\boldsymbol{\lambda}}_{MME})$	$(\hat{\boldsymbol{\lambda}}_{MLE})$	$(\hat{\boldsymbol{\lambda}}_{MME})$	$(\hat{\boldsymbol{\lambda}}_{MLE})$	$(\hat{\boldsymbol{\lambda}}_{MME})$	$(\hat{\boldsymbol{\lambda}}_{MLE})$	$(\hat{\boldsymbol{\lambda}}_{MME})$
0.1	2	1.99210	1.79768	0.0331	0.0629	-0.0078	-0.2023	0.026878	0.053048
0.3	2	1.98355	1.42241	0.0429	0.3671	-0.0164	-0.5776	0.038682	0.397956
0.5	2	2.00104	1.10174	0.0573	0.8668	0.0010	-0.8982	0.056701	0.852198
0.1	5	5.01315	4.78577	0.0579	0.1289	0.0132	-0.2142	0.05739	0.12180
0.3	5	4.98799	4.69349	0.0726	0.2042	-0.0120	-0.3065	0.073901	0.21660
0.5	5	4.98663	4.56399	0.1013	0.3503	-0.0134	-0.4360	0.10349	0.33487
0.1	10	10.0020	9.87662	0.1136	0.1508	0.0021	-0.1234	0.11116	0.15802
0.3	10	9.98901	9.82823	0.1375	0.1966	-0.0109	-0.1718	0.14292	0.18989
0.5	10	9.98859	9.79839	0.2096	0.2950	-0.0114	-0.2016	0.20009	0.29244

Source: Author’s calculation from simulated data

Table 2 Comparison of MLE and MME estimators of π for sample size 100

Parameters		**Estimate**	**Estimate**	**S-MSE**	**S-MSE**	**S-Bias**	**S-Bias**	**Theoretical Variance**	**Theoretical Variance**
$\boldsymbol{\pi}$	**λ**	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$
0.1	2	0.09893	0.00063	0.0031	0.0099	-0.0011	-0.0994	0.00357	0.00744
0.3	2	0.29456	0.01991	0.0036	0.0806	-0.0054	-0.2801	0.002782	0.06766
0.5	2	0.49592	0.06546	0.0036	0.1986	-0.0041	-0.4345	0.00266	0.18752
0.1	5	0.10023	0.05679	0.0010	0.0034	0.0002	-0.0432	0.00113	0.00296
0.3	5	0.29876	0.25229	0.0022	0.0052	-0.0012	-0.0477	0.00198	0.004902
0.5	5	0.49869	0.45123	0.0025	0.0055	-0.0013	-0.0488	0.00270	0.005701
0.1	10	0.09838	0.08672	0.0009	0.0013	-0.0016	-0.0133	0.00098	0.00206
0.3	10	0.29955	0.28791	0.0022	0.0025	-0.0004	-0.0121	0.00197	0.002875
0.5	10	0.50072	0.48898	0.0023	0.0025	0.0007	-0.0110	0.001667	0.003197

Source: Author’s calculation from simulated data

$\boxed{\mathbf{n = 200}}$

Table 3 Comparison of MLE and MME estimators of λ for sample size 200

Parameters		**Estimate** $(\hat{\lambda}_{MLE})$	**Estimate** $(\hat{\lambda}_{MME})$	**S-MSE** $(\hat{\lambda}_{MLE})$	**S-MSE** $(\hat{\lambda}_{MME})$	**S-Bias** $(\hat{\lambda}_{MLE})$	**S-Bias** $(\hat{\lambda}_{MME})$	**Theoretical Variance** $(\hat{\lambda}_{MLE})$	**Theoretical Variance** $(\hat{\lambda}_{MME})$
π	**λ**								
0.1	2	2.00107	1.79999	0.0152	0.0509	0.0011	-0.2000	0.0134392	0.0506251
0.3	2	1.99626	1.41678	0.0225	0.3564	-0.0037	-0.5832	0.01934	0.33359
0.5	2	1.99579	1.07103	0.0290	0.8946	-0.0042	-0.9289	0.02835	0.91487
0.1	5	4.99671	4.77175	0.0283	0.0913	-0.0033	-0.2282	0.028696	0.095455
0.3	5	4.98796	4.69246	0.0327	0.1442	-0.0120	-0.3075	0.03695	0.18743
0.5	5	4.995021	4.587642	0.0493	0.2485	-0.0049	-0.4123	0.051747	0.25384
0.1	10	9.99804	9.87935	0.0555	0.0829	-0.0019	-0.1206	0.05558	0.08424
0.3	10	9.99420	9.84628	0.0671	0.1042	-0.0058	-0.1537	0.071461	0.09323
0.5	10	9.99615	9.78272	0.1022	0.1673	-0.0038	-0.2173	0.100045	0.16122

Source: Author’s calculation from simulated data

Table 4 Comparison of MLE and MME estimators of π for sample size 200

Parameters		Estimate	Estimate	S-MSE	S-MSE	S-Bias	S-Bias	Theoretical Variance	Theoretical Variance
π	λ	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$	$(\hat{\pi}_{MLE})$	$(\hat{\pi}_{MME})$
0.1	2	0.09947	0.00004	0.0015	0.0099	-0.0006	-0.0999	0.000328	0.013678
0.3	2	0.29877	0.01219	0.0019	0.0839	-0.0012	-0.2878	0.000891	0.057363
0.5	2	0.49906	0.05709	0.0017	0.2029	-0.0009	-0.4429	0.00133	0.21318
0.1	5	0.09976	0.05690	0.0005	0.0027	-0.0002	-0.0431	0.00056	0.00344
0.3	5	0.30046	0.25593	0.0010	0.0033	0.0005	-0.0441	0.00117	0.003925
0.5	5	0.49864	0.45319	0.0013	0.0038	-0.0014	-0.0468	0.00085	0.004746
0.1	10	0.09986	0.08894	0.0004	0.0007	-0.0001	-0.0111	0.000494	0.0006906
0.3	10	0.30091	0.29034	0.0010	0.0012	0.0009	-0.0096	0.000899	0.001297
0.5	10	0.50084	0.48990	0.0012	0.0014	0.0008	-0.0101	0.0009334	0.0019896

Source: Author's calculation from simulated data

Note: [Index: S-MSE : Simulated MSE ; S-Bias : Simulated Bias]

From Tables 1-4 we make the following observations:

- The MLE estimators of π and λ gives better estimates of the parameters than the MME estimators.
- The simulated MSEs of the MLE estimators are smaller than the MME estimators of π and λ for every combination of n, π and λ.
- The magnitude of the bias is more for the MME estimators than their MLE counterparts for all n, π and λ.
- The theoretical variance of the MLE estimators are lesser than the MME estimators for every combination of n, π and λ.

Thus from our simulation study we can say that the MLE estimators are superior than the MME estimators of the parameters. The theoretical variances of the MLE and MME estimators also suggests the same. Thus our findings corroborate with the theoretical result.

References

1. A.M. Gun, M.K. Gupta and B. Dasgupta (2002): *Fundamentals of Statistics*, Vol. 1, 7th edition, The World Press Private Limited
2. C. Radhakrishna Rao: Linear Statistical Inference and its Applications, 2nd edition, Wiley Eastern Limited
3. Beckett, S., J. Jee, T. Ncube, S. Pompilus, Q. Washington, A. Singh and N. Pal (2014), "Zero-inflated Poisson (ZIP) distribution: parameter estimation and applications to model data from natural calamities", *involve: a journal of mathematics*, 7(6): 751-767
4. D. Lambert, "Zero-inflated Poisson regression with an application to defects in manufacturing", Technometrics 34:1 (1992), 1–14.
5. D. Böhning, "Zero-inflated Poisson models and C.A.MAN: a tutorial collection ofevidence", Biometric Model 40:7 (1998), 833–843.