CHAPTER 6

A PROBABILITY GENERATING FUNCTION BASED MINIMUM HELLINGER TYRE DISTANCE ESTIMATION

6.0 Introduction

The use of transforms (characteristic function, moment generating function) in statistical inference has been investigated by many researchers; see, for example, Press (1972), Heathcote (1972), Feuerverger and Mureika (1977), Feuerverger and McDunnough (1984), Meintanis and Swanepoel (2007) and the references therein. The main motivation for using transforms is that these transforms are usually much simpler than their corresponding probability density functions, and this leads to simpler inference procedures. In particular, for count variables, the probability generating function (pgf) has been proposed for testing goodness-of-fit and estimation.

Kocherlakota and Kocherlakota (1986) developed an empirical generating function based goodness-of-fit test for discrete distributions and this is further studied by Marques and Pérez-Abreu (1989). Baringhaus and Henze (1992) showed that the test is not consistent against each alternative distribution with finite first moment. Stute *et al.* (1993) showed that the parametric bootstrap method employed by Baringhaus and Henze (1992) provides an alternative approximation in situations when the true quantiles of test statistics are not tabulated. Rueda *et al*. (1991, 1999) extended the work of Kocherlakota and Kocherlakota (1986) for the goodness-of-fit.

The empirical probability generating function has been proposed for parameter estimation besides goodness-of-fit tests. Kemp and Kemp (1988) have suggested methods to estimate the unknown parameters by using the epgf. The basic idea is to solve the equations obtained by equating functionals of the pgf and epgf at selected values of the variable t. Dowling and Nakamura (1997) provided the asymptotic theory for these epgf based methods and showed that this theory can be used as a guide for the selection of t. The estimates obtained depend on the choice of the t values leading to arbitrariness of the estimates.

In this chapter, we propose a pgf-based divergence statistic for parameter estimation which avoids the arbitrariness due to the choice of values for the variable t. A particular case of the proposed statistic is the quadratic statistic of Rueda *et al*. (1991) which has been considered for testing goodness-of-fit. It seems there is no study on the performance and robustness of this quadratic statistic, and our proposed statistic, for parameter estimation especially in comparison with the maximum likelihood estimation (MLE) and minimum Hellinger distance estimation (MHDE). It is well-known that the MHDE and related procedures enjoy the property of robustness and asymptotic efficiency of the first order (see, for example, Basu *et al.*, 1997). In contrast, the popular MLE is not robust to outliers.

Kullback-Leibler (KL) divergence is not considered because the pdf (pmf) based KL divergence could not be compared to pgf based KL divergence since the probability could be zero. The HD does not have such a problem. In this chapter, we give the statistics and provide the basic asymptotic theory for the proposed pgf-based HD type estimation. The performance of the proposed statistics in estimation is examined with respect to the mean square error (MSE). Since the MHD estimator is robust we also examine the robustness of the proposed pgf-based statistics to outliers. Results of the Monte Carlo simulation studies are reported. Finally, the numerical examples are reported in the last section.

6.1 Generalized Minimum Hellinger Divergence Statistic based on Probability Generating Function

A general form of a statistic for goodness-of-fit (Meintanis and Swanepoel, 2007, page 1006) for discrete variables is given by

$$T_{N,\beta}(\theta) = N\int_0^\infty D_N^2(t)\beta(t)\,dt$$

where $D_N(t)$ is a distance between the pgf $G(t;\theta)$ and the epgf $G_N(t)$ and $\beta(t)$ is a weight function. Rueda *et al.* (1991, 1999) considered $D_N(t) = G_N(t) - G(t;\theta)$ and $\beta(t) \equiv 1$. Gürtler and Henze (2000) generalized this by taking $\beta(t) = t^b,\ b \ge 0$.

We consider a parametric family of models $\{g_\theta\}$ with pgf $\{G(t;\theta)\}$ and let $\Im$ be the family of all distributions with pgf $F(t)$. For count data, Basu *et al.* (1997) have proposed a generalized Hellinger divergence involving densities. Along similar lines but employing pgf, a pgf-based MHD type divergence for estimation will be proposed. We first define the divergence between two pgf's $F(t)$ and $G(t;\theta)$ as follows:

$$T(\theta;\gamma) = \int_0^1 (F(t)^\gamma - G(t;\theta)^\gamma)^2 \beta(t)\,dt, \qquad 0 < \gamma \le 1.$$

Observe that $T(\theta;\gamma) = 0$ if, and only if, $F(t) = G(t;\theta)$. Furthermore, $T(\theta;\gamma) \ge 0$.

If a random sample $X_1, X_2, ..., X_N$ is obtained from a distribution with pgf $F(t)$, the epgf is $F_N(t) = \frac{1}{N}\sum_{i=1}^{N} t^{x_i}$. We propose the pgf-based divergence statistic

$$T(\theta;\gamma,N) = \int_0^1 (F_N(t)^\gamma - G(t;\theta)^\gamma)^2 \beta(t)\,dt, \qquad 0 < \gamma \le 1. \tag{6.1}$$

when $\gamma = 1/2$, a Hellinger type statistic is obtained

$$T(\theta;1/2,N)=\int_0^1(\sqrt{F_N(t)}-\sqrt{G(t;\theta)})^2\beta(t)\,dt \tag{6.2}$$

Note that $G(t;\theta)$($F_N(t)$) is nonnegative for $t\in[0,1]$. When $\gamma=1$, we get an L_2 type statistic:

$$T(\theta;1,N)=\int_0^1(F_N(t)-G(t;\theta))^2\beta(t)\,dt \tag{6.3}$$

For the weight function $\beta(t)$ we consider (a) $\beta(t)=t^b,\ b\geq 0$ and (b) $\beta(t)=1/\omega(t)$ with $\omega(t)$ being the beta pdf

$$\omega(t)=\frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)},\quad B(\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)},\ t\in[0,1].$$

Note that $\beta(t)=t^b/(b-1), b=\alpha-1>1$ is a special case of $\omega(t)$ when $\beta=1$.

6.2 Asymptotic Results

In this section we discuss the consistency of the proposed estimator and its asymptotic distribution. The estimator obtained from minimizing $T(\theta;\gamma,N)$ is given by

$$\hat{\theta}_N=\arg\min_{\theta\in\Theta}T(\theta;\gamma,N)$$

We wish to show that if

$$T(\theta;\gamma,N)\underset{a.s.}{\rightarrow}T(\theta;\gamma),N\rightarrow\infty,$$

where $T(\theta;\gamma)=\int_0^1(F(t)^\gamma-G(t;\theta)^\gamma)^2\beta(t)\,dt$, then $\widehat{\theta}_N\underset{a.s.}{\rightarrow}\theta$. The notation a.s. denotes almost sure convergence. In other words, the estimator $\widehat{\theta}_N$ obtained from minimizing $T(\theta;\gamma,N)$ is a consistent estimator of θ.

Result 6.1 Consider $T\left(\theta;\gamma,N\right)=\int_0^1 (F_N(t)^{\gamma}-G(t;\theta)^{\gamma})^2\beta\left(t\right)dt, \qquad 0<\gamma\leq 1$. Then

$$T\left(\theta;\gamma,N\right)\underset{a.s.}{\rightarrow}T\left(\theta;\gamma\right), N\rightarrow\infty, \tag{6.4}$$

where $T\left(\theta;\gamma\right)=\int_0^1 (F(t)^{\gamma}-G(t;\theta)^{\gamma})^2\beta\left(t\right)dt$.

Proof. The Strong Law of Large Numbers implies that $F_N(t)-F(t)\underset{a.s.}{\rightarrow}0,\ N\rightarrow\infty$, that is, $\Pr\left\{\lim_{N\rightarrow\infty}F_N(t)=F(t)\right\}=1$. Application of the continuous mapping theorem shows that

$$\left(F_N(t)^{\gamma}-G(t;\theta)^{\gamma}\right)^2\underset{a.s.}{\rightarrow}\left(F(t)^{\gamma}-G(t;\theta)^{\gamma}\right)^2.$$

A further application gives the result (6.4).

The following result given in Amemiya (1973) (see also Newey and McFadden, 1994, page 2121) is required.

Lemma 6.1 (Amemiya, 1973, Lemma 3):

Let $Q_N\left(\omega,\theta\right)$ be a measurable function on a measurable space Ω and for each ω in Ω a continuous function for θ in a compact set Θ. Then there exists a measurable function $\hat{\theta}_N\left(\omega\right)$ such that

$$Q_N\left(\omega,\widehat{\theta}_N\left(\omega\right)\right)=\sup_{\theta\in\Theta}Q_N\left(\omega,\theta\right) \text{ for all } \omega \text{ in } \Omega.$$

If $Q_N\left(\omega,\theta\right)$ converges to $Q\left(\theta\right)$ *a.e.* (almost everywhere) uniformly for all θ in Θ, and if $Q\left(\theta\right)$ has a unique maximum at $\theta_0\in\Theta$, then $\widehat{\theta}_N$ converges to θ_0 *a.e.*

Remark. The condition of convergence *a.e.* uniformly in the above Lemma may be replaced by convergence *a.e.* as explained in Newey and McFadden (1994, page 2122).

Result 6.2 Suppose $T(\theta;\gamma,N)$ is a measurable function on the given measurable space and has a unique minimum at $\theta_0 \in \Theta$ where Θ is a compact set. Let $\hat{\theta}_N$ be the estimator obtained from minimizing $T(\theta;\gamma,N)$, that is,

$$\hat{\theta}_N = \arg\min_{\theta\in\Theta} T(\theta;\gamma,N).$$

Then $\widehat{\theta}_N \underset{a.s.}{\rightarrow} \theta$, that is , $\widehat{\theta}_N$ is a strongly consistent estimator of θ .

Proof. Let $Q_N(\omega,\theta) = -T(\theta,\gamma,N)$ in the Lemma and apply Result 3.4.

6.3 Monte Carlo Simulation Design

In this section, an empirical study has been designed to investigate the accuracy and robustness to outliers for the proposed estimation methods. To illustrate the methods, we consider the well-known negative binomial (NB) distribution which may be formulated as a mixed or compound Poisson distribution. Since there are various parameterizations of the NB (see Johnson, Kemp and Kotz, 2005), the following orthogonal parameterization for the pmf is used:

$$\Pr(X = x) = \binom{r+x-1}{r-1}\left(\frac{\mu}{\mu+r}\right)^x\left(1-\frac{\mu}{\mu+r}\right)^r \quad x = 0,\ 1,\ 2,\ ...,\ r > 0,\ \mu > 0.$$

Based upon (3.1) we consider the following six estimators

$$T_1(\theta;1/2,N) = \int_0^1 \left(\sqrt{F_N(t)} - \sqrt{G(t;\theta)}\right)^2 dt$$

$$T_2(\theta;1,N) = \int_0^1 \left(F_N(t) - G(t;\theta)\right)^2 dt$$

$$T_3(\theta;1/2,N)=\int_0^1\left(\sqrt{F_N(t)}-\sqrt{G(t;\theta)}\right)^2\times t^b dt, \quad b>0$$

$$T_4(\theta;1,N)=\int_0^1\left(F_N(t)-G(t;\theta)\right)^2\times t^b dt$$

$$T_5(\theta;1/2,N)=\int_0^1\left(\sqrt{F_N(t)}-\sqrt{G(t;\theta)}\right)^2\times\frac{1}{\omega(t)}dt$$

$$T_6(\theta;1,N)=\int_0^1\left(F_N(t)-G(t;\theta)\right)^2\times\frac{1}{\omega(t)}dt$$

where $\omega(t)=\dfrac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}, \quad B(\alpha,\beta)=\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$

The performance of the proposed estimators is compared with the MLE and the MHDE with respect to the mean square error (MSE) and bias.

In our simulation design, the sample sizes considered are N =100, 200, 500 and 1000 to represent small to large samples. For the feasible number of simulation runs, we performed 1000, 2000, 5000 and 10,000 simulation runs with selected set of parameters. From Figure 6.1, it is found that 1000 simulation runs give results of sufficient accuracy.

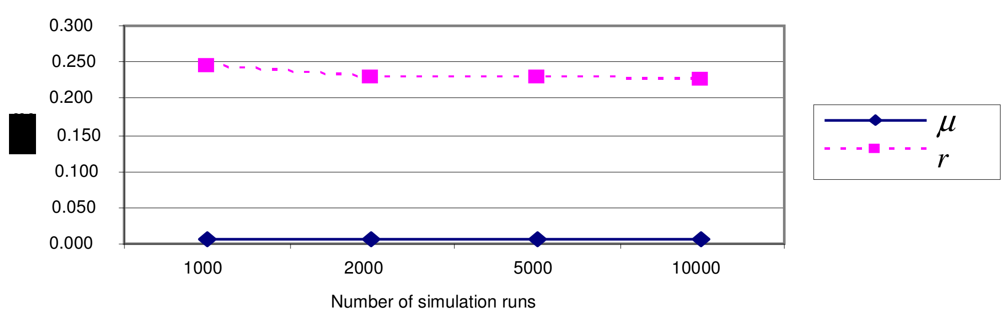

Figure 6.1: MSE of parameter estimates for NB distribution (μ =1.5; r = 2.0)

Outliers, amounting to (1-ω) percent of the sample size, are generated from the Poisson distribution with parameters $\lambda=30$ and $\lambda=65$ to represent moderate to extreme outliers. Figure 6.2 and 6.3 show the mixture of NB distribution with Poisson (POI) distribution. The mixed pmf is

$$\Pr(X = x) = \omega \binom{r+x-1}{r-1} \left(\frac{\mu}{\mu+r}\right)^x \left(1-\frac{\mu}{\mu+r}\right)^r + (1\text{-}\omega)\frac{e^{-\lambda}\lambda^x}{x!}, \quad \text{for } x = 0, 1, 2, \ldots \quad (6.5)$$

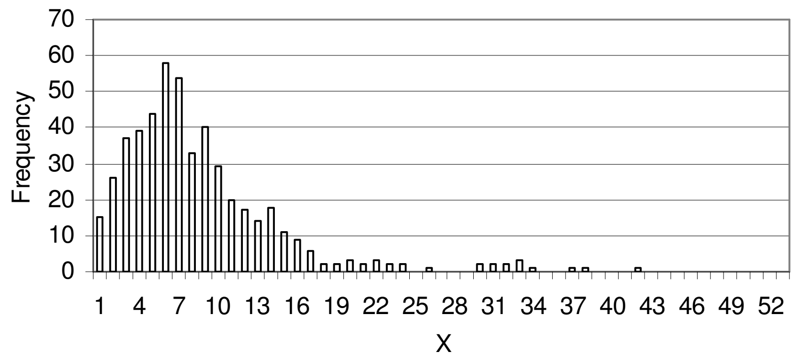

Figure 6.2: 97.5% NB (μ =6.53; r = 4.0) + 2.5% Poi (λ =30) with sample size of 500

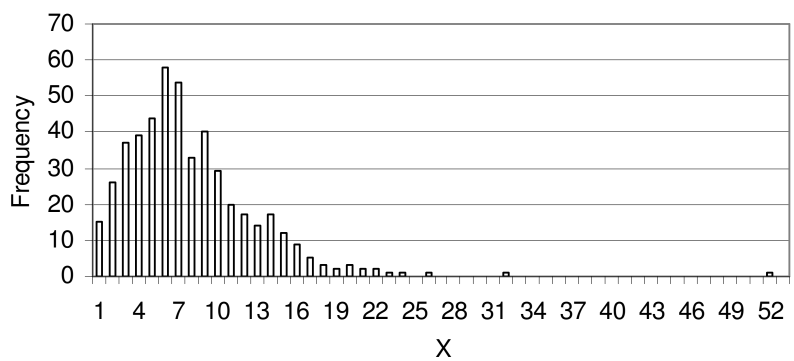

Figure 6.3: 97.5% NB (μ =6.53; r = 4.0) + 2.5% Poi (λ =65) with sample size of 500

The design of simulation study is shown in Figure 6.4 where we consider data with contamination (existence of outliers) by mixing the NB distribution ($\mu = 1.5$ and $r = 2.0$ or $\mu = 6.53$ and $r = 4.0$) with POI distribution ($\lambda = 30$ and $\lambda = 65$). Note that NB ($\mu = 6.53$ and $r = 4.0$) has longer tail than NB ($\mu = 1.5$ and $r = 2.0$).

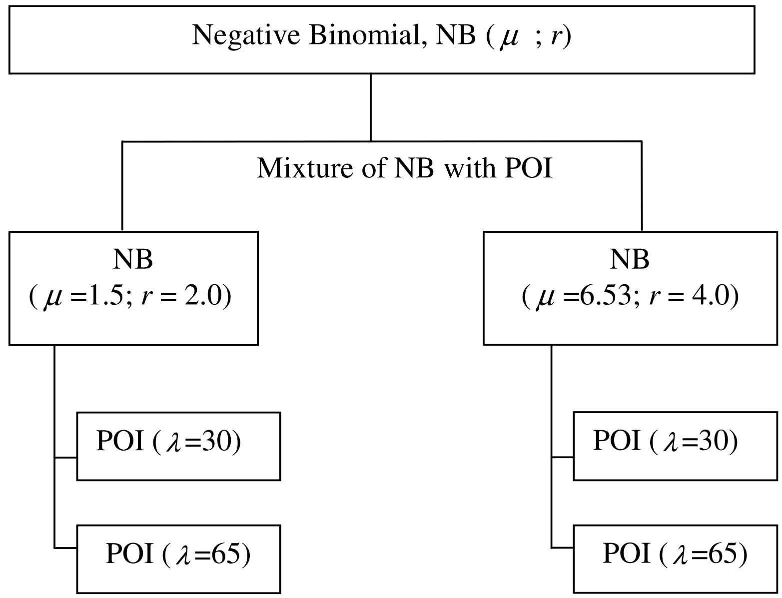

Figure 6.4: Design of simulation study

The estimation is done through numerical minimization by simulated annealing (Metropolis *et al.*, 1953). A Gaussian quadrature method is used to evaluate the integrals. IMSL FORTRAN routine GQRUL (or DQRUL) has been used.

Basically, a higher number of quadrature points produce a higher accuracy of the integral approximation. However, higher accuracy requires much computation time. Hence, the following empirical observation is made to choose the number of quadrature points. We consider six-points, twelve-points, twenty four-points Gaussian quadrature and the result is displayed in Figure 6.5. Clearly, from the figure, a six-point Gaussian quadrature method is sufficient for a good approximation of the integral which takes the range from 0 to 1.

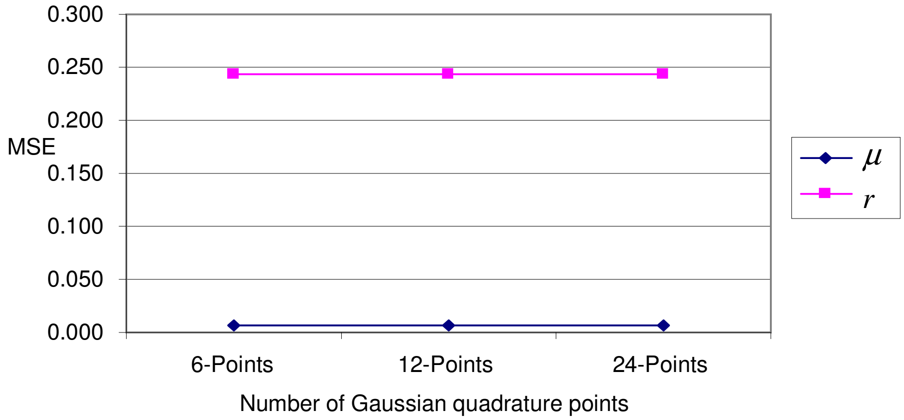

Figure 6.5: MSE of parameter estimates for NB distribution (μ =1.5; r = 2.0) with different number of Gaussian quadrature points

Next, we examined the range for parameter b (>0) under T_3 and T_4. From Gürtler and Henze (2000), choosing a large value of b means to put more weight near to the endpoint t =1 under the integration. Here, we fix the parameter b with values in the range 0 to 4. We examined the performance of T_3 and T_4 for data with and without outliers. Tables 6.1 and 6.2 exhibit the MSE and bias for the estimates of the parameters μ and r (NB) for data with and without outlier respectively.

Table 6.1: MSE and bias (in bracket) for proposed estimators T_3 and T_4 (with 5% outliers)

Sample Size, N	T_3 μ	$b=\frac{1}{4}$ r	T_4 μ	$b=\frac{1}{4}$ r	T_3 μ	$b=\frac{1}{2}$ r	T_4 μ	$b=\frac{1}{2}$ r	T_3 μ	$b=\frac{3}{4}$ r	T_4 μ	$b=\frac{3}{4}$ r	T_3 μ	$b=4$ r	T_4 μ	$b=4$ r
200	0.54	5.30	0.90	1.13	0.54	3.18	0.96	1.18	0.56	2.04	1.03	1.26	1.02	1.36	1.02	1.36
	(0.52)	(0.07)	(0.82	(-0.75)	(0.55	(-0.09)	(0.85)	(-0.84)	(0.58)	(-0.21)	(0.89)	(-0.91)	(0.89)	(-0.98)	(0.89)	(-0.98)
500	0.40	0.65	0.76	0.93	0.42	0.61	0.82	1.04	0.44	0.60	0.88	1.15	0.88	1.28	1.78	2.78
	(0.55)	(-0.37)	(0.82	(-0.87)	(0.57)	(-0.44)	(0.85)	(-0.94)	(0.60)	(-0.49)	(0.89)	(-1.01)	(0.89)	(-1.07)	(1.29)	(-1.65)
1000	0.35	0.44	0.70	0.89	0.37	0.45	0.76	1.00	0.40	0.47	0.82	1.12	0.82	1.25	1.68	2.78
	(0.55)	(-0.44)	(0.81)	(-0.89)	(0.57)	(-0.49)	(0.85)	(-0.96)	(0.59)	(-0.54)	(0.88)	(-1.02)	(0.88)	(-1.09)	(1.27)	(-1.66)

Table 6.2: MSE and bias (in bracket) for proposed estimators T_3 and T_4 (without outliers)

Sample size ,N	T_3 μ	$b=\frac{1}{4}$ r	T_4 μ	$b=\frac{1}{4}$ r	T_3 μ	$b=\frac{1}{2}$ r	T_4 μ	$b=\frac{1}{2}$ r	T_3 μ	$b=\frac{3}{4}$ r	T_4 μ	$b=\frac{3}{4}$ r	T_3 μ	$b=4$ r	T_4 μ	$b=4$ r
200	0.18	10.99	0.11	1.60	0.03	0.33	0.02	0.19	0.14	5.49	0.10	1.27	0.10	1.04	0.09	0.72
	(-0.02)	(0.79)	(0.00)	(0.29)	(0.00)	(0.05)	(0.00)	(0.02)	(-0.01)	(0.53)	(0.00)	(0.24)	(0.00)	(0.20)	(0.00)	(0.14)
500	0.07	0.82	0.04	0.40	0.06	0.69	0.04	0.37	0.05	0.61	0.04	0.34	0.04	0.29	0.03	0.23
	(0.00)	(0.16)	(0.01)	(0.07)	(0.00)	(0.13)	(0.01)	(0.06)	(0.00)	(0.12)	(0.01)	(0.06)	(0.01)	(0.05)	(0.01)	(0.03)
1000	0.03	0.38	0.02	0.21	0.03	0.33	0.02	0.19	0.03	0.30	0.02	0.18	0.02	0.15	0.02	0.12
	(0.00)	(0.06)	(0.00)	(0.03)	(0.00)	(0.05)	(0.00)	(0.02)	(0.00)	(0.05)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.01)

In Table 6.1, the disparity of MSE and bias achieved under $b = 1/4$ and $b = 4$ shows that a large value of b is preferred when the sample size is smaller than 500. In contrary, a smaller value of b is favoured for $N \geq 500$. Table 6.2 displays the MSE and bias for data without outlier and it is clear from the table that the proposed estimators T_3 and T_4 tend to favour large values of b, regardless of the sample size. Since the parameter b varies according to sample size and the existence of outlier, we shall treat the parameter b as a real number in the range 0 to 4.

We investigated the range for parameters α, β under the proposed estimator T_5 and T_6. Table 6.3 gives the estimated parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{r}$ and $\hat{\mu}$ for the mixed distribution of 97.5%ONB (μ=6.53 and r = 4.0) and 2.5% POI (λ=65) under the proposed estimator T_5 with sample size of 500.

Table 6.3: Estimated parameters for α and β for different range

Range for α and β = (0-4)				Range for α and β = (0-12)			
$\hat{\alpha}$	$\hat{\beta}$	$\hat{r}$	$\hat{\mu}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{r}$	$\hat{\mu}$
1.2404	1.0010	7.4642	2.6391	1.2417	1.0010	7.4651	2.6383
1.4616	1.0010	6.6547	3.3733	1.4638	1.0011	6.6557	3.3725
1.4195	1.0010	6.7159	4.0547	1.4240	1.0010	6.7150	4.0565
1.0010	1.0010	6.0454	4.9093	1.0010	1.0010	6.0448	4.9106
1.0010	1.0010	6.5308	4.5477	1.0010	1.0010	6.5299	4.5492
3.0371	1.0011	7.1687	2.4953	3.0513	1.0011	7.1680	2.4956
1.0745	1.0010	7.1419	4.0954	1.0746	1.0010	7.1419	4.0956
1.0010	1.0010	6.4316	5.1252	1.0011	1.0010	6.4309	5.1258
1.1910	1.0010	6.7659	3.4826	1.1942	1.0010	6.7660	3.4826
1.0010	1.0010	6.5504	4.6510	1.0010	1.0010	6.5505	4.6511

It is obvious in Table 6.3 that the range 0 to 4 provides results of sufficient accuracy. Therefore, we only need to consider α, β and b in the range 0 to 4.

6.3 Discussion of the Monte Carlo simulation results

Tables 6.4 to 6.7 show the MSE and bias for the simulated NB data (μ =1.5, r = 2.0) with outliers coming from the Poisson distributions ($\lambda = 30$ and 65). Meanwhile, the result for data without outlier is displayed in Table 6.8. Tables 6.9 to 6.13 also show the MSE and bias for NB data but with different parameters which are μ =6.53 and r =4.0. The same percentages of outliers are applied (2.5% and 5%). Finally, Table 6.14 exhibits the data without outlier under the NB data, μ =6.53 and r =4.0.

Firstly, we studied the result in Table 6.4 for the mixed distribution under NB data (μ =1.5, r = 2.0) with 2.5% outliers. It is shown that for small sample size (N=100), on the basis of the MSE, T_3 and T_4 are to be preferred. With increasing sample sizes (N=200 to 1000) all of the proposed estimation methods achieve low MSE which are better or comparable to MHDE. As expected MLE does not fare that well for data with outliers. Similar result is achieved when the percentage of outliers is increased to 5%, as displayed in Table 6.5.

In Tables 6.6 and 6.7, with the existence of extreme outliers, T_3 and T_4 appear to perform well with small MSE for small samples (N=100). All of the proposed estimation methods are rather good with small MSE for larger sample sizes (N=200 to 1000). It is seen that T_1, T_2, T_5 and T_6 performed consistently well and comparable to MHDE. MLE (high values of MSE and bias) is clearly not a desirable estimator for data with extreme outliers.

For data without outliers, Table 6.8 shows comparable performance of the proposed estimators with MLE and MHDE for large sample sizes but are poor for smaller sample sizes (N=100).

Next, we look into the result under NB data with μ =6.53 and r =2.0 which provides a distribution with a longer tail than μ =1.5 and r = 4.0. Tables 6.9 and 6.10 with contamination of 2.5% and 5% (POI, $\lambda = 30$), respectively indicate that for small sample sizes (*N*=100 and 200), T_3 and T_4 are favoured based on the small MSE values. Except for MLE, all of the proposed estimation methods achieve small and comparable MSE compare to MHDE when the sample size is large (*N*=500 and *N*=1000).

In Tables 6.11 and 6.12, we present the data with extreme outlier where we have the mixture of NB distribution (μ =6.53 and r =2.0) with POI distribution ($\lambda = 65$). For small sample size (*N*=100), T_3 and T_4 achieve the lowest MSE values among all of the estimators. When *N*=200, T_2 , T_3 , T_4 and T_6 are comparable to MHDE. For large sample sizes (*N*=500 and *N*=1000), all of the proposed estimation methods perform well and are comparable to MHDE. MLE performs badly with the existence of outliers.

For the data without contamination under the NB (μ =6.53 and r =2.0) distribution, MLE, MHDE and the proposed estimators attained small MSE values when the sample size is large (*N*=200 to *N*=1000). Large MSE values are achieved for small sample size (*N*=100) for all of the estimators.

Besides MSE and bias, we also look into the Qualitative Robustness Index (QRI) which was introduced by Alam *et al.* (2010), in order to measure the effect of contamination on different estimators at different contamination models. The QRI is defined as

$$QRI = \frac{1}{\sum \left| \xi^s_{100\alpha} - \xi^c_{100\alpha} \right|}$$

where $\xi^s_{100\alpha}$ = 100^{th}_{α} percentile of the simulated sampling distribution of the estimators for the standard model.

$\xi^{c}_{100\alpha}$ = 100^{th}_{α} percentile of the simulated sampling distribution of the estimators for the contaminated model.

We examine the QRI by using the same α as considered by Alam *et al.* (2010) which are 0.005, 0.01, 0.025, 0.05, 0.1, 0.5, 0.9, 0.95, 0.975, 0.99, and 0.995. Note that the larger value of QRI, the more stable is the model.

Our study shows that the value of QRI is violated if the estimation for the standard model (data without contamination) fails. This happened especially when the sample size is not large enough, as QRI is calculated according to how large is the disparity of the estimator from the data with contamination and without contamination. The result is not shown here.

Besides the mixture of NB distribution and POI distribution, we study the mixture of (i) the $\mathrm{GIT}_{3,1}$, (ii) the Neyman type A distribution and (iii) the Poisson inverse Gaussian (PIG) distribution with POI distribution for the performance of T_1 and T_2 only. Appendix A displays the tables of MSE and bias for these distributions.

Table 6.4: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 2.5% outliers)

$\mu = 1.5,\ r = 2.0,\ \lambda = 30$

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	0.76 (0.72)	1.64 (-1.17)	0.03 (-0.10)	31.07 (1.87)	0.07 (0.17)	10.46 (0.14)	0.08 (0.18)	4.95 (-0.05)	0.15 (0.28)	0.84 (-0.64)	0.17 (0.30)	0.88 (-0.71)	0.06 (0.15)	13.20 -0.28	0.07 (0.16)	5.74 -0.10
200	0.64 (0.72)	1.69 (-1.28)	0.02 (0.12)	0.21 (-0.23)	0.05 (0.16)	0.45 (-0.27)	0.06 (0.18)	0.40 (-0.34)	0.11 (0.28)	0.69 (-0.76)	0.13 (0.29)	0.78 (-0.82)	0.04 (0.14)	0.49 (-0.16)	0.04 (0.15)	0.42 (-0.21)
500	0.57 (0.72)	1.75 (-1.32)	0.01 (-0.03)	0.25 (0.25)	0.03 (0.16)	0.23 (-0.36)	0.04 (0.18)	0.26 (-0.42)	0.09 (0.27)	0.69 (-0.81)	0.10 (0.29)	0.79 (-0.87)	0.03 (0.13)	0.18 (-0.25)	0.03 (0.14)	0.19 (-0.28)
1000	0.53 (0.71)	1.77 (-1.33)	0.00 (-0.02)	0.09 (0.13)	0.03 (0.16)	0.19 (-0.38)	0.03 (0.17)	0.23 (-0.43)	0.08 (0.27)	0.69 (-0.82)	0.08 (0.27)	0.69 (-0.82)	0.02 (0.13)	0.13 (-0.26)	0.02 (0.13)	0.13 (-0.29)

Table 6.5: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 5% outliers)

$\mu = 1.5,\ r = 2.0,\ \lambda = 30$

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	2.47 (1.42)	2.12 (-1.44)	0.03 (1.41)	35.07 (-1.42)	0.19 (0.35)	3.48 (-0.37)	0.23 (0.39)	3.15 (-0.50)	0.54 (0.62)	1.27 (-1.06)	0.64 (0.68)	1.40 (-1.13)	0.15 (0.30)	4.21 (-0.18)	0.16 (0.32)	3.36 (-0.29)
200	2.24 (1.42)	2.19 (-1.48)	0.02 (-0.06)	1.82 (0.62)	0.15 (0.35)	0.53 (-0.58)	0.18 (0.38)	0.58 (-0.67)	0.44 (0.60)	1.31 (-1.12)	0.51 (0.66)	1.45 (-1.19)	0.11 (0.28)	0.43 (-0.41)	0.12 (0.30)	0.43 (-0.46)
500	2.14 (1.43)	2.22 (-1.49)	0.01 (-0.03)	0.25 (0.23)	0.13 (0.34)	0.46 (-0.64)	0.16 (0.38)	0.55 (-0.71)	0.38 (0.60)	1.35 (-1.15)	0.45 (0.65)	1.50 (-1.22)	0.08 (0.27)	0.28 (-0.46)	0.09 (0.29)	0.32 (-0.50)
1000	2.07 (1.43)	2.23 (-1.49)	0.00 (-0.02)	0.09 (0.12)	0.12 (0.34)	0.45 (-0.65)	0.15 (0.37)	0.55 (-0.73)	0.36 (0.59)	1.35 (-1.16)	0.43 (0.64)	1.51 (-1.22)	0.07 (0.26)	0.25 (-0.46)	0.08 (0.28)	0.29 (-0.51)

Table 6.6: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 2.5% outliers)

μ =1.5, r = 2.0, λ =65

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	3.65 (1.60)	5.06 (-1.32)	0.03 (-0.10)	31.08 (1.87)	0.08 (0.17)	10.18 (0.11)	0.09 (0.19)	4.7 (-0.08)	0.2 (0.33)	0.94 (-0.74)	0.25 (0.37)	1.03 (-0.83)	0.06 (0.14)	13.68 (0.34)	0.07 (0.15)	8.2 (0.20)
200	3.14 (1.61)	2.37 (-1.53)	0.02 (-0.06)	1.50 (0.66)	0.05 (0.17)	0.45 (-0.29)	0.06 (0.19)	0.41 (-0.36)	0.14 (0.32)	0.84 (-0.85)	0.18 (0.35)	0.98 (-0.93)	0.04 (0.13)	0.56 (-0.12)	0.04 (0.14)	0.48 (-0.15)
500	2.78 (1.60)	2.44 (-1.56)	0.01 (-0.03)	0.25 (0.27)	0.04 (0.17)	0.24 (-0.37)	0.04 (0.18)	0.27 (-0.44)	0.11 (0.31)	0.85 (-0.90)	0.14 (0.34)	1.00 (-0.98)	0.02 (0.12)	0.18 (-0.21)	0.02 (0.13)	0.18 (-0.23)
1000	2.63 (1.59)	2.46 (-1.57)	0 (-0.02)	0.1 (0.15)	0.03 (0.16)	0.2 (-0.39)	0.04 (0.18)	0.25 (-0.46)	0.1 (0.30)	0.85 (-0.91)	0.12 (0.34)	1.01 (-0.99)	0.02 (0.12)	0.12 (-0.22)	0.02 (0.12)	0.12 (-0.24)

Table 6.7: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 5% outliers)

μ =1.5, r = 2.0, λ =65

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	12.13 (3.15)	2.69 (-1.61)	0.03 (-0.09)	35.08 (1.91)	0.21 (0.37)	3.44 (-0.40)	0.26 (0.41)	3.16 (-0.53)	0.90 (0.78)	1.55 (-1.16)	1.25 (0.91)	1.76 (-1.25)	0.13 (0.28)	6.56 (-0.04)	0.14 (0.29)	4.10 (-0.16)
200	11.01 (3.16)	2.74 (-1.66)	0.02 (-0.06)	1.80 (0.63)	0.17 (0.36)	0.55 (-0.60)	0.20 (0.40)	0.63 (-0.70)	0.66 (0.74)	1.57 (-1.23)	0.89 (0.85)	1.79 (-1.32)	0.09 (0.26)	0.45 (-0.33)	0.10 (0.27)	0.43 (-0.37)
500	10.58 (3.19)	2.77 (-1.66)	0.01 (-0.03)	0.25 (0.26)	0.14 (0.35)	0.49 (-0.66)	0.17 (0.40)	0.60 (-0.75)	0.56 (0.72)	1.62 (-1.27)	0.74 (0.83)	1.85 (-1.35)	0.07 (0.24)	0.25 (-0.39)	0.08 (0.26)	0.26 (-0.42)
1000	10.27 (3.17)	2.78 (-1.67)	0.00 (-0.02)	0.10 (0.15)	0.13 (0.35)	0.48 (-0.67)	0.16 (0.39)	0.60 (-0.76)	0.52 (0.71)	1.63 (-1.27)	0.68 (0.81)	1.86 (-1.36)	0.06 (0.24)	0.20 (-0.40)	0.07 (0.25)	0.23 (-0.43)

Table 6.8: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (without outliers)

μ =1.5; r = 2.0

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	0.03	1.59	0.04	23.79	0.03	26.62	0.03	22.09	0.03	8.29	0.03	7.47	0.03	26.19	0.03	21.41
	(0.00)	(0.38)	(-0.10)	(1.64)	(0.00)	(1.10)	(0.00)	(0.96)	(0.00)	(0.53)	(0.00)	(0.50)	(0.01)	(1. 07)	(0.00)	(0.94)
200	0.01	0.46	0.02	2.14	0.02	3.27	0.02	3.15	0.01	0.53	0.01	0.51	0.02	3.25	0.02	3.14
	(0.00)	(0.17)	(-0.06)	(0.62)	(0.01)	(0.28)	(0.00)	(0.27)	(0.00)	(0.16)	(0.00)	(0.16)	(0.01)	(0.28)	(0.00)	(0.26)
500	0.01	0.13	0.01	0.24	0.01	0.21	0.01	0.19	0.01	0.14	0.01	0.14	0.01	0.21	0.01	0.19
	(0.00)	(0.06)	(-0.03)	(0.24)	(0.00)	(0.08)	(0.00)	(0.07)	(0.00)	(0.05)	(0.00)	(0.05)	(0.00)	(0.07)	(0.00)	(0.07)
1000	0.00	0.06	0.00	0.09	0.00	0.10	0.00	0.10	0.00	0.06	0.00	0.06	0.00	0.10	0.00	0.09
	(0.00)	(0.03)	(-0.02)	(0.13)	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.04)	(0.00)	(0.04)

Table 6.9: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 2.5% outliers)

μ =6.53, r = 4.0, λ =30

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	0.66	1.70	0.28	6.90	0.49	103.99	0.37	5.22	0.35	2.85	0.43	1.50	0.45	75.05	0.35	5.25
	(0.58)	(-0.98)	(-0.30)	(1.32)	(0.16)	(3.26)	(0.28)	(0.34)	(0.31)	(0.01)	(0.40)	(-0.36)	(0.17)	(2.54)	(0.27)	(0.37)
200	0.51	1.51	0.12	1.04	0.28	12.72	0.24	1.18	0.23	0.75	0.30	0.71	0.25	8.92	0.22	1.21
	(0.59)	(-1.13)	(-0.12)	(0.38)	(0.20)	(0.66)	(0.30)	(-0.09)	(0.32)	(-0.28)	(0.40)	(-0.55)	(0.20)	(0.56)	(0.28)	(-0.04)
500	0.41	1.52	0.05	0.30	0.14	0.82	0.14	0.40	0.16	0.37	0.21	0.55	0.13	0.78	0.14	0.41
	(0.59)	(-1.20)	(0.08)	(-0.25)	(0.23)	(-0.04)	(0.30	(-0.28)	(0.33)	(-0.40)	(0.40)	(-0.63)	(0.22)	(-0.03)	(0.28)	(-0.23)
1000	0.37	1.51	0.06	0.37	0.10	0.40	0.14	0.40	0.13	0.29	0.18	0.49	0.09	0.39	0.11	0.27
	(0.59)	(-1.21)	(0.19)	(-0.52)	(0.23)	(-0.13)	(0.30)	(-0.28)	(0.32)	(-0.43)	(0.40)	(-0.65)	(0.22)	(-0.12)	(0.28)	(-0.27)

Table 6.10: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 5% outliers)

μ =6.53, r = 4.0, λ =30

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	1.83	2.84	0.27	6.56	0.69	1.29	0.71	2.82	0.76	1.84	1.06	1.59	0.64	65.06	0.67	3.00
	(1.17)	(-1.57)	(-0.20)	(0.95)	(0.39)	(-1.12)	(0.60)	(-0.09)	(0.66)	(-0.44)	(0.83)	(-0.90)	(0.39)	(2.20)	(0.57)	(0.00)
200	1.61	2.92	0.16	0.92	0.47	9.98	0.55	1.01	0.61	0.92	0.89	1.34	0.44	7.28	0.51	1.04
	(1.17)	(-1.67)	(0.06)	(-0.12)	(0.44)	(0.39)	(0.61)	(-0.42)	(0.67)	(-0.66)	(0.83)	(-1.04)	(0.42)	(0.34)	(0.57)	(-0.33)
500	1.48	3.01	0.21	0.90	0.32	0.74	0.45	0.57	0.52	0.73	0.77	1.31	0.31	0.74	0.40	0.53
	(1.18)	(-1.72)	(0.36)	(-0.83)	(0.47)	(-0.24)	(0.61	(-0.57)	(0.67)	(-0.76)	(0.84)	(-1.10)	(0.45)	(-0.21)	(0.57)	(-0.47)
1000	1.17	-1.73	0.52	-1.09	0.47	-0.32	0.61	-0.60	0.67	-0.78	0.83	-1.12	0.45	-0.29	0.55	-0.48
	(1.42)	(3.02)	(0.31)	(1.26)	(0.27)	(0.42)	(0.41)	(0.49)	(0.48)	(0.69)	(0.72)	(1.29)	(0.26)	(0.42)	(0.34)	(0.40)

Table 6.11: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 2.5% outliers)

μ =6.53, r = 4.0, λ =65

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	3.23	4.69	0.31	8.35	0.52	100.64	0.46	4.21	0.47	2.46	0.76	1.69	0.47	81.48	0.42	4.79
	(1.46)	(-1.98)	(-0.37)	(1.72)	(0.18)	(3.16)	(0.36)	(0.15)	(0.41)	(-0.21)	(0.60)	(-0.77)	(0.18)	(2.70)	(0.32)	(0.29)
200	2.72	4.88	0.15	1.78	0.30	12.07	0.30	1.09	0.33	0.82	0.56	1.23	0.28	10.33	0.27	1.23
	(1.47)	(-2.17)	(-0.24)	(0.89)	(0.23)	(0.61)	(0.37)	(-0.23)	(0.42)	(-0.48)	(0.60)	(-0.94)	(0.21)	(0.58)	(0.33)	(-0.11)
500	2.38	5.09	0.05	0.43	0.15	0.80	0.20	0.46	0.23	0.54	0.43	1.16	0.15	0.82	0.17	0.46
	(1.47)	(-2.24)	(-0.13)	(0.41)	(0.25)	(-0.08)	(0.37)	(-0.41)	(0.42)	(-0.59)	(0.60)	(-1.01)	(0.24)	(-0.04)	(0.32)	(-0.27)
1000	0.02	0.17	0.02	0.17	0.11	0.40	0.17	0.35	0.20	0.47	0.39	1.12	0.11	0.41	0.13	0.31
	(-0.08)	(0.23)	(-0.08)	(0.23)	(0.25)	(-0.17)	(0.37)	(-0.44)	(0.42)	(-0.61)	(0.59)	(-1.03)	(0.24)	(-0.13)	(0.31)	(-0.29)

Table 6.12: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (with 5% outliers)

μ =6.53, r = 4.0, λ =65

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	10.41	6.90	0.31	9.54	0.77	75.12	1.05	2.29	1.24	1.92	2.38	2.73	0.73	68.08	22.52	8.71
	(2.90)	(-2.58)	(-0.37)	(1.76)	(0.44)	(2.43)	(0.77)	(-0.39)	(0.89)	(-0.79)	(1.30)	(-1.47)	(0.42)	(2.30)	(-4.74)	(-2.15)
200	9.43	7.19	0.14	1.80	0.54	9.38	0.84	1.11	1.02	1.36	2.00	2.71	0.51	8.01	0.66	1.18
	(2.91)	(-2.67)	(-0.23)	(0.85)	(0.49)	(0.30)	(0.78)	(-0.66)	(0.89)	(-0.98)	(1.29)	(-1.59)	(0.46)	(0.34)	(0.65)	(-0.37)
500	8.99	7.37	0.05	0.42	0.38	0.74	0.70	0.84	0.88	1.28	1.78	2.78	0.35	0.78	0.48	0.60
	(2.94)	(-2.71)	(-0.13)	(0.40)	(0.52)	(-0.30)	(0.78)	(-0.80)	(0.89)	(-1.07)	(1.29)	(-1.65)	(0.48)	(-0.22)	(0.62)	(-0.46)
1000	8.71	7.40	0.02	0.17	0.33	0.45	0.64	0.78	0.82	1.25	1.68	2.78	0.28	0.45	0.39	0.44
	(2.92)	(-2.72)	(-0.08)	(0.23)	(0.52)	(-0.39)	(0.77)	(-0.82)	(0.88)	(-1.09)	(1.27)	(-1.66)	(0.46)	(-0.28)	(0.58)	(-0.45)

Table 6.13: MSE and bias (in bracket) for proposed estimators $T_1 - T_6$ (without outliers)

μ =6.53, r = 4.0

Sample size	MLE		MHDE		T_1		T_2		T_3		T_4		T_5		T_6	
	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r	μ	r
100	0.18	1.62	0.31	8.44	0.41	135.86	0.23	8.70	0.19	5.66	0.18	3.45	0.36	78.67	0.23	8.18
	(0.00)	(0.32)	(-0.36)	(1.71)	(-0.07)	(4.08)	(-0.01)	(0.89)	(-0.01)	(0.59)	(0.00)	(0.42)	(-0.03)	(2.79)	(0.00)	(0.82)
200	0.09	0.51	0.14	1.62	0.21	18.28	0.12	1.89	0.10	1.04	0.09	0.72	0.19	10.98	0.11	1.76
	(0.00)	(0.13)	(-0.23)	(0.84)	(-0.03)	(1.03)	(0.00)	(0.32)	(0.00)	(0.20)	(0.00)	(0.14)	(-0.01)	(0.78)	(0.01)	(0.30)
500	0.03	0.19	0.05	0.40	0.08	1.01	0.04	0.45	0.04	0.29	0.03	0.23	0.07	0.87	0.04	0.43
	(0.00)	(0.04)	(-0.12)	(0.39)	(0.00)	(0.19)	(0.01)	(0.08)	(0.01)	(0.05)	(0.01)	(0.03)	(0.01)	(0.16)	(0.01)	(0.07)
1000	0.02	0.09	0.02	0.16	0.04	0.45	0.02	0.23	0.02	0.15	0.02	0.12	0.04	0.41	0.02	0.22
	(0.00)	(0.02)	(-0.07)	(0.22)	(0.00)	(0.08)	(0.00)	(0.03)	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)	(0.07)	(0.00)	(0.03)

6.5 Examples of Data Fitting

Four well-known data sets (a) European red mite on apple leaves. (Garman, 1951) and (b) yeast cells counts in a haemocytometer (Student, 1907) (c) Soil bacteria in microscopic counts per field (Jones and Mollison, 1948) (d) Soil bacteria in microscopic counts per colony (Jones and Mollison, 1948) have been fitted with the proposed estimation methods. The fits are given in the tables below.

Table 6.14: Fitting NB to counts of red mites on apple leaves with MLE, MHD and proposed estimation methods

No. per unit x	Observed frequency	Expected frequency							
		MLE	MHD	T_1	T_2	T_3	T_4	T_5	T_6
0	70	69.48	70.03	70.10	70.10	69.88	69.84	70.10	70.10
1	38	37.60	37.14	36.72	36.72	36.99	37.04	36.72	36.73
2	17	20.10	19.85	19.72	19.72	19.85	19.87	19.72	19.72
3	10	10.71	10.64	10.68	10.68	10.69	10.70	10.68	10.68
4	9	5.69	5.71	5.81	5.80	5.78	5.77	5.81	5.80
5	3	3.02	3.07	3.16	3.16	3.12	3.12	3.16	3.16
6	2	1.60	1.65	1.73	1.73	1.69	1.68	1.73	1.73
7	1	0.85	0.89	0.94	0.94	0.92	0.91	0.94	0.94
8+	0	0.96	1.03	1.14	1.14	1.08	1.07	1.14	1.14
Total	150	150	150	150	150	150	150	150	150
χ^2		3.54	3.49	3.42	3.42	3.43	3.44	3.42	3.42
P-value		0.74	0.75	0.75	0.75	0.75	0.75	0.75	0.75
Estimated parameter	$\hat{\mu}$	1.147	1.150	1.165	1.165	1.165	1.165	1.165	1.165
	$\hat{r}$	1.025	0.985	0.952	0.952	0.974	0.977	0.952	0.952

$\chi^2_{5,0.05} = 11.070$

Table 6.15: Fitting NB to yeast cells per square in a haemocytometer with MLE, MHD and proposed estimation methods

No. per unit x	Observed frequency	Expected frequency							
		MLE	MHD	T_1	T_2	T_3	T_4	T_5	T_6
0	213	214.06	214.95	213.32	213.35	213.99	214.01	213.37	213.40
1	128	122.81	121.74	124.51	124.40	123.07	123.03	124.33	124.22
2	37	45.06	44.87	45.03	45.04	45.04	45.03	45.05	45.06
3	18	13.42	13.58	12.95	12.99	13.34	13.35	13.01	13.05
4	3	3.54	3.66	3.25	3.27	3.49	3.49	3.28	3.30
5	1	0.86	0.92	0.74	0.75	0.84	0.84	0.76	0.76
6+	0	0.25	0.28	0.20	0.20	0.24	0.24	0.20	0.21
Total	400	400	400	400	400	400	400	400	400
χ^2		3.58	3.60	3.36	3.36	3.56	3.54	3.29	3.33
P-value		0.47	0.46	0.50	0.50	0.47	0.47	0.51	0.50
Estimated parameter	$\hat{\mu}$	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68
	$\hat{r}$	3.59	3.32	4.18	4.13	3.67	3.66	4.10	4.06

$\chi^2_{3,0.05} = 7.815$

Table 6.16: Fitting NB to soil bacteria per field with MLE, MHD and proposed estimation methods

Bacteria per field	Observed frequency	Expected frequency							
		MLE	MHDE	T_1	T_2	T_3	T_4	T_5	T_6
0	11	8.89	10.74	10.73	10.65	10.51	10.23	10.73	10.68
1	17	18.52	17.61	19.77	19.76	19.68	19.51	19.77	19.78
2	31	25.13	21.19	25.00	25.07	25.10	25.15	25.00	25.06
3	24	28.04	22.40	26.77	26.90	27.00	27.22	26.77	26.86
4	29	27.89	22.04	26.09	26.23	26.36	26.65	26.09	26.18
5	18	25.73	20.72	23.91	24.03	24.16	24.44	23.91	23.99
6	19	22.49	18.88	20.99	21.08	21.19	21.41	20.99	21.05
7	16	18.89	16.81	17.85	17.90	17.97	18.13	17.85	17.88
8	13	15.37	14.70	14.81	14.82	14.86	14.95	14.81	14.81
9	17	12.20	12.68	12.05	12.03	12.05	12.06	12.05	12.03
10	6	9.48	10.82	9.64	9.61	9.60	9.57	9.64	9.62
11	8	7.25	9.14	7.62	7.58	7.55	7.49	7.62	7.59
12+	31	20.12	42.26	24.77	24.34	23.97	23.18	24.78	24.47
Total	240	240	240	240	240	240	240	240	240
χ^2		15.43	14.24	9.51	9.80	10.10	10.83	9.52	9.70
P-value		0.12	0.16	0.48	0.46	0.43	0.37	0.48	0.47
Estimated parameter	$\hat{\mu}$	5.67	6.99	5.85	5.82	5.81	5.78	5.85	5.83
	$\hat{r}$	3.29	2.14	2.69	2.72	2.76	2.85	2.69	2.71

$\chi^2_{10,0.05} = 18.307$

Table 6.17: Fitting NB to soil bacteria per colony with

MLE, MHD and proposed estimation methods

Bacteria per colony	Observed frequency	Expected frequency							
		MLE	MHD	T_1	T_2	T_3	T_4	T_5	T_6
1	359	356.65	346.76	359.67	359.66	358.29	358.05	359.68	359.65
2	146	143.20	121.29	138.27	138.29	139.85	140.12	138.29	138.30
3	57	73.99	67.66	72.04	72.05	72.74	72.86	72.05	72.05
4	41	41.07	42.44	40.81	40.81	40.99	41.01	40.81	40.82
5	26	23.58	28.09	24.05	24.05	23.98	23.97	24.05	24.05
6	17	13.82	19.17	14.50	14.50	14.34	14.31	14.50	14.50
7+	27	20.69	47.60	23.65	23.64	22.81	22.67	23.63	23.63
Total	673	673	673	673	673	673	673	673	673
χ^2		6.88	16.51	4.64	4.64	5.11	5.21	4.65	4.65
P-value		0.23	0.01	0.46	0.46	0.40	0.39	0.46	0.46
Estimated parameter	$\hat{\mu}$	1.09	1.49	1.12	1.12	1.12	1.11	1.12	1.12
	$\hat{r}$	0.64	0.46	0.58	0.58	0.60	0.60	0.58	0.58

$\chi^2_{4,0.05} = 9.488$

Table 6.14 to 6.17 exhibit the chi-square goodness of fit under MLE, MHDE and the proposed methods. In Table 6.14, it is seen that all the proposed methods have comparable values of χ^2 values with MLE and MHDE. Meanwhile in Table 6.15, a data set with larger sample size than the data set in Table 6.14, all of the proposed methods achieved slightly better χ^2 values than MLE and MHDE. For the counts of bacteria data as shown in Tables 6.16 and 6.17, all of the proposed methods perform better than MLE and MHDE as their χ^2 values are lower. Obviously, T_1 and T_2 are to be favoured due to their simplicity.