Estimating the Parameter n of the Binomial Distribution Using Moments Generating Function Approach

By

Mohammed Khalil Hussein Shakhatreh

B. Sc. (Statistics), 1993

Yarmouk University

Estimating the Parameter n of the Binomial Distribution Using Moments Generating Function Approach

By

Mohammed Khalil Hussein Shakhatreh

B.Sc. (Statistics), 1993

Yarmouk University

Thesis submitted in partial fullfillment of the requirements for the degree of Master of Science (Statistics) at Yarmouk University

Thesis Defense Committee

Dr. Mohammad F. Al-Saleh Chairman:

Prof. Mohammad S. Ahmed Member:

Dr. Hassen A. Muttlak Member:

December 1995

Acknowledgment

I am deeply grateful to my thesis advisor Dr. Mohammad F. Al-Saleh for his guidance, advice and valuable suggestions. I would like to express my appreciation to my committee members, Prof. Mohammed S. Ahmed and Dr. Hassen A. Muttlak. Special thanks are extended to my friend Abdel-Qader Al-Masri for helping me in preparing the computer programs. My respects to all faculty members in the Department of Statistics.

My deepest gratitude to my brothers, sisters and cousins for their continuous encouragement, care and support.

Special thanks are extended to my friends in master program, to them and all other friends I wish successful life and brilliant future.

Special thanks are extended to the director of AL-RISALHA for typing this thesis.

Mohammed

بسم الله الرحمن الرحيم

﴿وقَضَى رَبُّكَ أَنْ لا تَعْبُدُوا إلا إِيَّاهُ وبِالوالِدَين إِحْسَاناً إِمّا يَبْلُغَنَّ عِنْدَكَ الكِبَرَ أَحَدُهُمَا أَوْ كِلاهُمَا فَلا تَقُلْ لَهُمَا أُفٍ ولا تَنْهَرْهُمَا وَقُلْ لَهُمَا قَوْلاً كَرِيماً﴾ (الإسراء، آية ٢٣)

الإهداء

الى نور عُيُونِي وَيُنْبُوعُ حَيَاتِي

أُمِّــي و أَبِــي

محمد خليل حسين شخاترة

Table of Contents

CHAPTER ONE: INTRODUCTION AND LITERATURE REVIEW 1

1.1 Preface 1

1.2 Statement of the problem 4

1.3 Review of the literature 6

CHAPTER TWO: MOMENT GENERATING FUNCTION BASED ESTIMATORS 16

2.1 Introduction 16

2.2 Derivation of the estimator based on moment generatng function (MGF) when p is known 17

2.3 The asymptotic behavior of the MGF based estimator $\hat{n}_{m,t}$ 22

2.4 Derivation of the maximum likelihood estimator $\hat{n}_L$ 28

2.5 Numerical comparisons 33

2.6 Results of numerical comparisons 55

2.7 Derivation of the estimator based on moment generating function when p is unknown 57

2.8 Numerical work 59

CHAPTER THREE: BAYESIAN ESTIMATION OF THE BINOMIAL PARAMETER 65

3.1 Introduction 65

3.2 Bayes Estimators 65

3.3 Numerical comparisons 69

3.4 Results of numerical comparisons 79

Chapter four: CONCLUDING REMARKS AND RECOMMENDATIONS 80

4.1 Concluding remarks 80

4.2 Recommendations 80

References 81

Appendix 83

List of Abreviations

a.s.	almost surely
i.i.d.	independent identically distributed
MGF	Moment Generating Function
$\hat{n}_{B_1}$	Bayes estimator with respect to Poisson prior
$\hat{n}_{B_2}$	Bayes estimator with respect to non-informative prior
$\hat{n}_L$	maximum likelihood estimator (MLE) for n
$\hat{n}_m$	method of moments estimator (MME) for n
$\hat{n}_{m,t}$	moment generating function based estimator for n when p is known
$\hat{n}_{(\hat{p},t)}$	moment generating function based estimator for n when p is unknown
MLE	maximum likelihood estimator
MME	method of moment estimator
$\hat{n}_{m:s}$	the stabilized version of MME for n
$\hat{n}_{L:s}$	the stabilized version of MLE for n
r.v.	random variable
w.p.1	with probability one
w.r.t	with respect to
b(n,p)	binomial with parameters n and p

ABSTRACT

ESTIMATING THE PARAMETER n OF THE BINOMIAL DISTRIBUTION USING MOMENTS GENERATING FUNCTION APPROACH

Suppose $\underline{X} = (X_1, \ldots, X_m)$ is a random sample from $b(n,p)$. Given observations $\underline{x} = (x_1, \ldots, x_m)$, we want to estimate n by using the moment generating function approach. We discuss the two cases, p known and p unknown, separately. For p known, the behavior of the moment generating function based estimator for the parameter n is studied. This estimator, say $\hat{n}_{m,t}$, which depends on the sample size m and on an auxiliary variable t is obtained as a solution of an equation generated by equating the theoretical moment generating function to its empirical counter part. It is shown that for any fixed t, $\hat{n}_{m,t}$ is strongly consistent for n, and for any fixed m, $\hat{n}_{m,t}$ converges to the method of moment estimator for n as $t \longrightarrow 0$ and $\hat{n}_{m,t}$ converges to $X_{(m)} = \max(X_1, \ldots, X_m)$, as $t \longrightarrow \infty$. Moreover, the limiting distribution of $\hat{n}_{m,t}$, when either $m \longrightarrow \infty$ and $t \longrightarrow 0$ or $t \longrightarrow 0$ and $m \longrightarrow \infty$, is shown to coincide with that of the method of moment estimator. We compare the mean square error of $\hat{n}_{m,t}$ with the mean square error of the other estimators such as, maximum likelihood estimator, method of moments and Bayes estimators, for selected values of t. Also, a comparison is made based on the bias. For p unknown, an estimator say $\hat{n}_{(\hat{p},t)}$

for n can be obtained using this approach, but by solving two equations for given t_1, t_2. The stability of this estimator is compared with the stability of other existing estimators such as MLE, MME.

Based on these comparisons and taking into consideration the simple form of $\hat{n}_{m,t}$, we recommend using it.

تقدير المعلمة n لتوزيع ذات الحدين بإستخدام طريقة دالة العزوم المولدة

الملخص

تتناول هذه الرسالة دراسة حول تقدير المعلمة n بإستخدام الدالة المولدة للعزوم عندما تكون المشاهدات مأخوذة من توزيع ذات الحدين ذو معلمتين n,p.

لقد قمنا بدراسة هذا التقدير في حالتين عندما تكون p معلومة وعندما تكون p مجهولة.

الحالة الاولى: عندما تكون p معلومة، قمنا بدراسة سلوك هذا التقدير، ولنفرض أنه $n_{m,t}$، والذي يعتمد على حجم العينة m وعلى المتغير المساعد t، والذي يمكن الحصول عليه عن طريق مساواة الدالة المولدة للعزوم الحقيقية مع الدالة المولدة للعزوم التجريبية.

تم إثبات أنه لأي قيمة ثابتة لـ t، يقترب هذا التقدير من القيمة الحقيقية لـ (n) بإحتمال واحد، أيضا عندما يكون حجم العينة ثابت، فإن هذا التقدير يقترب من تقدير العزوم عندما تكون قيمة t صغيرة جدا $(t \to 0)$، ويقترب الى اكبر مشاهدة من المشاهدات عندما تكون قيمة t كبيرة جدا $(t \to \infty)$.

بالإضافة الى ذلك، قمنا بدراسة توزيع التقدير $n_{m,t}$ عندما يكون حجم العينة كبير $(m \to \infty)$ والمتغير المساعد صغير جدا $(t \to 0)$ وبالعكس، ووجدنا أنهما يتطابقان مع توزيع تقدير العزوم. أيضا قمنا بمقارنة متوسط مربعات الأخطاء لـ $(n_{m,t})$ مع متوسط مربعات الاخطاء لتقديرات اخرى مثل: التقدير الاعظم للدالة الاحتمالية، تقدير العزوم ومقدرات بييز.

الحالة الثانية: عندما تكون p مجهولة، قمنا بإشتقاق التقدير المعتمد على الدالة المولدة للعزوم، ولنفرض انه $n(p,t)$ والذي يمكن الحصول عليه عن طريق حل معادلتين لأي قيم معطاة لـ t_1, t_2 وقد قمنا بمقارنة اتزان هذا التقدير مع اتزان تقديرات أخرى مثل التقدير الأعظم للدالة الإحتمالية وتقدير العزوم.

إعتمادا على هذه المقارنات، وآخذين بعين الإعتبار سهولة هذا التقدير فإننا نوصي بإستخدامه.

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

1.1. Preface

Estimators based on transforms of a distribution function have been extensively discussed and investigated in numerous works in the literature (see, for example Titterington, Smith and Makov, 1985).

There are two main approaches for obtaining such estimators. In one approach, the estimator is chosen to minimize a certain distance between the theoretical and empirical transforms. Suppose that for some auxiliary variable $t \in T$,

$$G(t|\theta) = Eg(t,X) = \int g(t,x)\, dF(x|\theta)$$

provided the integral exists. If the sample space is discrete, we replace the integral by sum and if x is multivariate, then t is also vector valued.

If, also $Eg^2(t,X) < \infty$, for all $t \in T$, $X_1, \ldots, X_m$ represent a random sample from $F(.|\theta)$ and if we define

$\bar{g}_m(t) = m^{-1} \sum_{i=1}^{m} g(t,x_i)$, then by the law of large numbers $\bar{g}_m(t) \longrightarrow G(t|\theta)$. A natural source of estimators for θ, therefore, is the minimization of some distance measure between $\bar{g}_m(.)$ and $G(.|\theta)$ say

$$\delta[\bar{g}_m(.),\ G(.|\theta)],\ \text{say}$$

For example one may use the quadratic distance, that is

$$\delta[\bar{g}_m(.),G(.|\theta)] = \int |G(t|\theta) - \bar{g}_m(t)|^2\ dw(t)$$

where w(.) is a positive weighting measure on T.

Quandt and Ramsey (1978) applied the moment generating function (MGF) method to five parameters normal mixture, using the quadratic distance, that is, if a sample observations $x_1, \ldots, x_m$ is given on a random variable X, where it is known that

$$X \sim N(\mu_1,\sigma_1^2) \quad \text{with probability } \lambda$$

and

$$X \sim N(\mu_2,\sigma_2^2) \quad \text{with probability } 1-\lambda$$

The parameters λ, μ_1, μ_2, σ_1^2, σ_2^2 being unknown. The parameter estimates are obtained by minimizing

$$\sum_{j=1}^{5}\left[\lambda\ e^{\mu_1 t_j+\frac{1}{2}\sigma_1^2 t_j^2} + (1-\lambda)\ e^{\mu_2 t_j+\frac{1}{2}\sigma_2^2 t_j^2} - m^{-1}\sum_{i=1}^{m} e^{t_j x_i}\right]^2$$

In the second approach, the estimator is taken to be the solution of an equation obtained by equating the theoretical transform with its empirical counter part.

Shaul, K. Bar-Lev, N. Barkan and N. A. Langberg (1993) adopt the second approach and consider the problem of estimating the natural parameter of a natural exponential family. The transform that they used for this purpose is the moment generating function (MGF). That is, if X is a r.v distributed according to the probability distribution $F_\theta(dx) = \exp[\theta X - k(\theta)]\, V(dx)$, where $k(\theta) = \ln T(\theta)$, $T(\theta) = \int e^{\theta x}\, V(dx)$ and V be a positive, σ-finite measure on $\mathbb{R}$, with support S containing at least two points. The MGF based estimator for θ, on the basis of the sample $\underline{X} = (X_1, \ldots, X_m)$ is the solution of the equation

$$E_\theta(e^{tX_1}) = m^{-1} \sum_{i=1}^{m} e^{t\chi_i} \qquad (1)$$

or, equivalently, of the equation

$$k(\theta+t) - k(\theta) = \ln\left(m^{-1} \sum_{i=1}^{m} e^{t\chi_i}\right).$$

The solution for (1), if it exists, depends on the sample size m, the sample elements $\chi_1, \ldots, \chi_m$, and the auxiliary variable t. They denoted this solution by $\hat{\theta}_{m,t}$, it exists with probability 1, and is given as the unique solution of (1). Also, they showed that for any fixed t, $\hat{\theta}_{m,t}$ is strongly consistent for θ as $m \longrightarrow \infty$; and for any fixed m, $\hat{\theta}_{m,t}$ converges to $\tilde{\theta}_m$, the MLE for θ, as $t \longrightarrow 0$. Furthermore, they showed that the limiting distribution of $\hat{\theta}_{m,t}$, as either $m \longrightarrow \infty$ and $t \longrightarrow 0$, or as $t \longrightarrow 0$ and $m \longrightarrow \infty$, coincides with that of $\tilde{\theta}_m$.

These asymptotic results suggest, in some situations, the use of $\hat{\theta}_{m,t}$, with large m and small t as an alternative to the MLE.

1.2 Statement of the problem

The problem is as follows: Suppose $X_1, \ldots, X_m$ are a random sample from binomial distribution with parameter n and p b(n,p). Given observations $x_1, \ldots, x_m$, we want to estimate n by using the method of moment generating function. Also we discuss the cases p known and p unknown separately.

Note that the binomial distribution when n is unknown does not belong to the exponential family and hence the problem we are considering is not a special case of the work of Shaul k-Bar-Lev, N. Barkan and N. A. Langberg (1993).

Estimation of the parameter n in the binomial distribution can be useful in practice. Draper and Guttman (1971) gave the following example: "Suppose for example, that the Apex Appliance company wishes to estimate the number of a certain type of appliance in use in a certain service area. Suppose further that the company believes that the weekly total of defective appliances sent in for repair (irrespective of age) arises with a binomial probability p about whose value they have some prior knowledge. Then a count x of the number of defective appliances received during a routine week could be used to cast light on the population size n".

The following is another example which was introduced by Rukhin (1975): "Let us assume n animals are randomly and uniformly distributed in some field. A statistician wants to make inference about the number n on the basis of number of animals captured by successively placed traps. It is supposed that the probability for the i-th trap to capture one animal is known to the statistician and is p_i, $0 < p_i < 1$, i = 1, ..., m, (in the simplest case p_i is the relative area of the i-th trap). If x_i represents the number of animals in the i-th trap (i = 1, ..., m) and the animals are captured independently, then the joint distribution of the random variables $X_1, \ldots, X_m$ has form

$$f(x_1,\ldots,x_m) = \frac{n!}{\prod_{i=1}^{m} x_i!\,(n-\sum_{i=1}^{m} x_i)!}\, p_1^{x_1}(p_2q_1)^{x_2}\cdots(q_1\cdots q_m)^{n-\sum_{i=1}^{m} x_i}$$

$x_i \geq 0$, $\sum_{i=1}^{m} x_i \leq n$, $q_i = 1 - p_i$.

The statistic $\sum_{i=1}^{m} X_i$ is sufficient for parameter n. It has binomial distribution with parameter n and $p_0 = 1 - \prod_{i=1}^{m} (1-p_i)$. Thus $\sum_{i=1}^{m} X_i$ can be used to estimate n.

We give the following example: Suppose that the chief of a police center wishes to estimate the total number of crimes in a certain locality. Suppose further that monthly of total reported crimes received at the center arises with a binomial probability p. Then a count x the number of reported crimes received during a month could be used to cast light on the total number of crimes.

1.3. Review of the Literature.

The standard estimation problem associated with a binomial distribution, with probability function $f(x,p) = \binom{n}{x} p^{x} (1-p)^{n-x}$, $x = 0, 1, \ldots, n$ is that of estimating p. A much harder and less studied problem is that of estimating n. Feldman and Fox (1968) discussed the estimation of the parameter n in the binomial distribution when the other parameter p is known. Based on a random sample, $X_1, \ldots, X_m$ from b(n,p), they derived the Maximum Likelihood Estimator $\hat{n}_L$ and they showed that it is consistent in a relative sense and asymptotically normal. They gave bounds on $\hat{n}_L$ which is:

$$\frac{\max\limits_{1 \le i \le m}(X_i)}{1-q^m} \le \hat{n}_L \le \frac{\sum_{i=1}^{m} X_i}{1-q^m}$$

Also, they defined a new random variable $Y_i = X_i/q$, which is, for large, n approximately normal (i.e., $N(\mu,\mu)$, where $\mu = np/q$). Using Y_i they examined three estimators for μ;

i. Minimum variance unbiased estimator (MVUE) of μ which is equal to $\hat{\mu}_1 = \left(\frac{Z}{m}\right)^{1/2} \frac{I_{m/2}\left(\sqrt{mZ}\right)}{I_{m/2-1}\left(\sqrt{mZ}\right)}$ where $Z = \sum_{i=1}^{m} Y_i^2$,

I_λ is the modified Bessel function of type I.

ii. The maximum likelihood estimator (MLE) which is equal to $\hat{\mu}_2 = \left(\frac{Z}{m} + \frac{1}{4}\right)^{1/2} - \frac{1}{2}$.

iii. The usual estimator of μ, which is equal to $\hat{\mu}_3 = \bar{Y}$.

From these estimators; $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\mu}_3$, they obtained respectively three estimators of n in terms of the original random variables,

$$\hat{n}_1 = \frac{q}{p}\left(\sum_{i=1}^{m} X_i^2/m\right)^{1/2} \frac{I_{m/2}\left(\frac{1}{q}\sqrt{m\sum_{i=1}^{m} X_i^2}\right)}{I_{m/2-1}\left(\frac{1}{q}\sqrt{m\sum_{i=1}^{m} X_i^2}\right)}$$

$$\hat{n}_2 = \frac{q}{p}\left(\sum_{i=1}^{m} X_i^2/mq + \frac{1}{4}\right)^{1/2} - \frac{1}{2}\,\frac{q}{p}$$

and $\hat{n}_3 = \frac{\sum_{i=1}^{m} X_i^2}{mq}$. These estimators, $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\mu}_3$, were shown to be asymptotically equivalent.

Draper and Guttman (1971), gave a Bayesian treatment of the problem both when p is known and when p is unknown. For known p, they proposed a uniform prior of n on the set {1, 2, ..., N}, ($p_0(n) = 1/N$, $n = 1, \ldots, N$) and they derived the posterior distribution for n which is given by $\pi(n|\underline{x},p) \propto (1-p)^{mn}\, p_0(n) \prod_{i=1}^{m} \frac{n!}{(n-x_i)!}$ $x_{(m)} \le n \le N$ where $X_1, \ldots, X_m$ is a random sample from b(n,p) and $x_{(m)}$ is the order statistics of $x_1, \ldots, x_m$. They used the mode of posterior distribution as an estimate of n also, they claimed that the posterior distribution can be examined to cast light on the precision of the estimate. For p unknown, they assumed that n and p are independent and they proposed the following prior: $g_0(n,p) = p_0(n)\, h_0(p)$, where $h_0(p) \propto p^{v_1-1}(1-p)^{v_2-1}$, $0 < p < 1$, $v_1, v_2 > 0$. They derived the joint posterior of p and n and they integrate p out from

$$\pi(n,p|\underline{x}) \propto p^{t-v_1-1}(1-p)^{mn-t+v_2-1}\, p_0(n) \prod_{i=1}^{m} \frac{n!}{(n-x_i)!} \quad \text{where}$$

$t = \sum_{i=1}^{m} x_i$, to get the marginal distribution for n and as in the previous case, they used the mode as an estimate of n. They provided two examples for the two cases. For p known they take $p = 0.8$, $N = 15$, $\alpha = 10$ they observed that the posterior distribution of n is essentially unchanged to three decimal places for $N \geq 21$ and the mode is at $n = 12$. For unknown p they assume that p has uniform prior distribution and obtain results that are considerably different. As N increases the mode of their posterior remains the same at $n = 10$.

Ghosh and Meeden (1975), showed that the estimator $T(X) = X/p$, where $X \sim b(n,p)$ with known $p \in (0,1)$ and n is an unknown parameter contained in the set $\mathbb{N} = \{0,1,\ldots\}$, is admissible under quadratic loss function. Rukhin (1975), made some statistical inference about the parameter n of the binomial distribution with known p. He showed that the estimator $T(X) = X/p$ is (i) a variant of the maximum likelihood estimator, (ii) the only unbiased estimator of n, (iii) minimax with respect to the weighted squared error loss

$$L(\delta(x),n) = \begin{cases} \dfrac{(\delta(x)-n)^2}{n} & \text{for } n \geq 1 \\ A\delta^2(x) & n = 0 \end{cases}, \quad A > 0$$

Blumenthal and Dahiya (1981), extended the results of Feldman and Fox (1968) to cover some additional cases and show how they may be applied to a goodness-of-fit test, also they considered the problem of zero-truncated observations.

Also, they compared the following estimators: $\hat{n}_L$, $\hat{n}_m = \bar{X}/p$, $\hat{n}_2 = \frac{1}{p}\left(\sum_{i=1}^{m} X_i^2/m\right)^{1/2}$ (the estimator that minimizes the χ^2-statistic), $\hat{n}_3 = \frac{q}{p}\left(\frac{1}{m}\sum_{i=1}^{m} X_i^2 + \frac{1}{4}\right)^{1/2} - \frac{q}{2p}$ (proposed by Feldman and Fox (1968)) for small m, n, in terms of their efficiencies. They observed that, in general, the MLE is preferred.

Olikn, Petkau and Zidek (1981), discussed the estimation of the parameter n in the binomial distribution when the other parameter p is also unknown. They considered the MME ($\hat{n}_m$) and the MLE ($\hat{n}_L$) of n and they showed when p is small, while n is large, both estimators become highly unstable (in the sense that a small change in one of the observation yields a large change in the estimators). They formulated the stabilized versions of the MME and MLE. The stabilized version of MME is $\hat{n}_{m:s} = \max\left\{\frac{\hat{\sigma}^2\phi^2}{\phi-1}, S_{max}\right\}$, where

$$\phi = \begin{cases} \dfrac{\hat{\mu}}{\hat{\sigma}^2} & \text{if } \dfrac{\hat{\mu}}{\hat{\sigma}^2} \geq \left(1+\dfrac{1}{\sqrt{2}}\right) \\[2ex] \max\left(\dfrac{S_{max}-\hat{\mu}}{\hat{\sigma}^2}, 1+\sqrt{2}\right) & \text{if } \dfrac{\hat{\mu}}{\hat{\sigma}^2} < \left(1+\dfrac{1}{\sqrt{2}}\right) \end{cases}$$

S_{max} = maximum $(X_1, \ldots, X_m)$, $\hat{\mu} = \sum_{i=1}^{m} X_i/m$, $\hat{\sigma}^2 = \sum_{i=1}^{m} (X_i-\hat{\mu})^2/m$

if $\frac{\hat{\mu}}{\hat{\sigma}^2} > 1 + 1/\sqrt{2}$ the case is called stable, otherwise unstable. The stabilized version of MLE is

$$\hat{n}_{L:s} = \begin{cases} \hat{n}_L & \text{if } \frac{\hat{\mu}}{\hat{\sigma}^2} \geq \left(1+\frac{1}{\sqrt{2}}\right) \\ J_m & \text{if } \frac{\hat{\mu}}{\hat{\sigma}^2} < \left(1+\frac{1}{\sqrt{2}}\right) \end{cases}$$

where $J_m = S_{max} + \left(\frac{m-1}{m}\right)\left(S_{max} - S_{(m-1)}\right)$ is the Jakknife estimator of n and $S_{(m-1)}$ is the (m-1)-th order statistics of $(X_1, \ldots, X_m)$. The simulation experiments has shown that $\hat{n}_{m:s}$ and $\hat{n}_{L:s}$ are not as sensitive to small perturbations in the x_i's as $\hat{n}_m$ and $\hat{n}_L$.

Carroll and Lombard (1985), considered the problem of estimating the parameter n based on independent random sample $(X_1, \ldots, X_m)$ from a binomial distribution with unknown parameters n and p. They took a beta prior distribution for p with parameters α, $\beta > 0$ and they integrated the product of the likelihood function $L(n,p|\underline{x}) = \prod_{i=1}^{m} \left(\binom{n}{x_i} p^{x_i}(1-p)^{n-x_i}\right)$, $n \geq x_{(m)}$ and the prior probability density function of p, over p to obtain the beta-binomial likelihood for n.

$$L(n|\underline{x}) = \left(\prod_{i=1}^{m} \binom{n}{x_i}\right) \times \left((mn+a+b+1)\binom{mn+a+b}{a+\sum_{i=1}^{m} x_i}\right)^{-1}$$

The Carroll-Lombard estimator [CLE(α,β)] is obtained by maximizing this beta-binomial likelihood for n. A numerical work shows that the CLE(α,β) is reasonably stable for the choices (a,b) = (0,0), (1,1).

Casella (1986), proposed a method for assessing the sensitivity for the MLE. Suppose that $X_1, \ldots, X_m$ be a random sample from b(n,p), where both n and p are unknown. The Ln-likelihood function is

$$L(n,p|\underline{x}) = \sum_{i=1}^{m} \ln \binom{n}{x_i} + \bar{x} \ln p + k(n-\bar{x}) \ln(1-p)$$

they approximated the $\ln$-likelihood function by

$$\ell_\alpha(n,p|\underline{x}) = mh_\alpha(n) - \sum_{i=1}^{m} h_{1-\alpha}(n-x_i) - \sum_{i=1}^{m} \ln x_i! + m\bar{x} \ln p$$
$$+ m(n-\bar{x}) \ln(1-p)$$

where $h_\alpha(y) = (1-\alpha)y \ln y + \alpha(y+1) \ln(y+1) - y$, $0 \le \alpha \le 1$ for α near 1/2, $\ell_\alpha(n,p)$ is very close to $\ell(n,p)$. They treat $\ell_\alpha(n,p|\underline{x})$ as a likelihood function and they obtained $\hat{n}_\alpha$, $\hat{p}_\alpha$ the maximum likelihood estimators based on $\ell_\alpha(n,p|\underline{x})$, where $\hat{p}_\alpha = \frac{\bar{x}}{\hat{n}_\alpha}$ and $\hat{n}_\alpha$ is the solution of the following equation:

$$\ln \left(\frac{n^{(1-\alpha)m} \, (n+1)^{\alpha m} \left[\sum_{i=1}^{m}(n-x_i)/mn\right]^m}{\prod_{i=1}^{m} (n-x_i)^{\alpha} \prod_{i=1}^{m} (n-x_i+1)^{1-\alpha}} \right) = 0$$

They considered the problem to be stable if $\hat{n}_\alpha$ was not overly sensitive to changes in α, for values of α near 1/2.

Sadoghi-Alvandi (1986), proved that if $X \sim b(n,p)$ with p known, $p \in (0,1)$ and unknown, $n \in \{1, 2, \ldots\}$, the estimator $T(0) = -\frac{(1-p)}{p \ln p}$, $T(X) = \frac{X}{p}$, $X = 1, 2, \ldots$ is admissible under quadratic loss, and the only admissible estimator for $p \geq 1/2$. Also he proved that the natural estimator $T(0) = 1$, $T(X) = \frac{X}{p}$, $X = 1, 2, \ldots$ is inadmissible estimator under quadratic loss.

Kahn (1987), showed that, if n is large, then the prior distribution for n alone determines which moments of the posterior distribution exist, that is, if $X_1, \ldots, X_m$ is a random sample from b(n,p) where both n and p are unknown and consider priors on n and p that are factorable and can be written as f(n) g(p), further, let g(p) be a beta density with parameters a and b. The posterior density on n given $\underline{x}$ after integrate out p is

$$\pi(n|\underline{x}) \propto \frac{\Gamma(mn-t+b)}{\Gamma(mn+a+b)} \prod_{i=1}^{m} \frac{\Gamma(n+1)}{\Gamma(n-x_i+1)}, \text{ for } n \geq x_{(m)}$$

where $t = \sum_{i=1}^{m} x_i$. Then for some positive constant C and all f such that $f(n) > 0$ and for all sufficiently large n,

$$\lim_{n\to\infty} \frac{\pi(n|\underline{x})}{f(n)/n^a} = C.$$

Hamedani and Walter (1988), considered the Bayesian approach for estimating the parameter n when p is known and when p is unknown. For p known, they took a Poisson prior of n and they obtained the posterior distribution function of n and they used the mean of the posterior distribution as an estimate of n. For p unknown, they suggest a beta prior for p and they obtained the posterior distribution function and as in the previous case, they used the mean of the posterior distribution as an estimate of n. They used the examples that introduced by Draper and Guttman (1971) and they showed that the assumption of improper priors in both p and n leads to implausible results.

Gunel and Chilko (1989), considered the problem of estimating the parameter n based on a random sample of size m from a binomial distribution with unknown parameters n, p, using a Bayesian approach for estimating n. They took a continuous prior for n (i.e., $n \sim G(\alpha+\beta,\delta)$) and a beta prior distribution as an estimate of n. They observed that the mean of the posterior distribution proposed a stable estimator and dominates $\hat{n}_m$: $\hat{n}_L$: S and CLE(α,β) in terms of the mean squared error for $(\alpha,\beta) = (1,1), (2,2)$.

Sadoghi-Alvandi (1992), showed that if $X \sim b(n,p)$ with known p and unknown, $n \in \{0, 1, \ldots\}$ the linear estimator $\hat{n} = CX+d$ relative to the linex loss function ($L(\Delta) = b[e^{a\Delta} - a\Delta - 1]$, where Δ is the estimation error and $b > 0$, $a \neq 0$, are the parameters of the loss function) is inadmissible for $C < 1$ and $d \geq 0$ and the estimator X+d is admissible for $d \geq 0$. Also, he showed that the admissibility of the usual estimator $\hat{n} = x/p$ depends on the sign of the shape parameter a, if $a < 0$, then x/p is admissible, otherwise, it is inadmissible.

CHAPTER TWO

MOMENT GENERATING FUNCTION BASED ESTIMATORS

2.1. Introduction

In this chapter, we study the behavior of the moment generating function based estimator for the parameter n of the binomial distribution b(n,p). We use a random sample of size m, and we discuss the cases p known and p unknown separately. For p known we derive the MGF based estimator $\hat{n}_{m,t}$ and we discuss some properties of this estimator. Also, we study its asymptotic behavior, as either $m \longrightarrow \infty$ or $t \longrightarrow 0$. Also, we make comparisons among the various estimators; the estimator based on MGF $\hat{n}_{m,t}$, the MLE $\hat{n}_L$ and MME $\hat{n}_m$ for small m, n.

For p unknown, our main concern is the stability of the estimator of n, since the MME $\hat{n}_m$ and the MLE $\hat{n}_L$ of n are unstable in the sense that they are highly sensitive to small perturbations, that is, an increase or decrease in an observed success count by one can result in a drastic change. We provide the estimator $\hat{n}_{(\hat{p},t)}$ which is based on MGF and we analyze the examples listed in Table (2) of Olkin, Petkau and Zidek (1981) who computed $\hat{n}_m$, $\hat{n}_{m:s}$, $\hat{n}_L$ and $\hat{n}_{L:s}$ for some particular cases.

2.2. Derivation of the estimator based on moment generating function MGF when p is known

Let $\underline{X} = (X_1, \ldots, X_m)$ be a random sample taken from $x \sim b(n,p)$, where n is the parameter of interest, $n \in \{1, 2, \ldots\}$ and p is known, $0 < p < 1$.

The MGF based estiamtor for n, on the basis of the sample X, is the solution of the equation:

$$E_n\left(e^{X_1 t}\right) = \frac{\sum_{i=1}^{m} e^{x_i t}}{m}$$

$$(q+pe^t)^n = \frac{\sum_{i=1}^{m} e^{x_i t}}{m} \quad ; \; q = 1-p$$

$$\hat{n}_{m,t} = \frac{\ln\left(m^{-1} \sum_{i=1}^{m} e^{X_i t}\right)}{\ln(q+pe^t)}$$

This estimator may not be an integer. In this case we define another estimator and we call it the modified MGF based estimator, say $\hat{n}^*_{m,t}$, where

$$\hat{n}^*_{m,t} = \begin{cases} 1 & \text{if } \hat{n}_{m,t} \le 1 \\ [(\hat{n}_{m,t}+.5)] & \text{if } \hat{n}_{m,t} > 1 \end{cases}$$

where [a] is the largest integer in a.

Theorem 2.2.1. Let $\underline{X} = (X_1, \ldots, X_m)$ be a random sample taken from binomial distribution with parameter n and known $p \in (0,1)$, then

i. For fixed t, $p\left(\lim_{m\to\infty} \hat{n}_{m,t} = n\right) = 1$. i.e., $\hat{n}_{m,t}$ is strongly consistent estimator of n.

ii. For fixed m, $\lim_{t\to 0} \hat{n}_{m,t} = \bar{X}/p$.

iii. The estimator $\hat{n}_{m,t}$ underestimates n for $t > 0$ and overestimates n for $t < 0$.

iv. For $t > 0$, $\dfrac{t\bar{X}}{\ell n(q+pe^t)} \le \hat{n}_{m,t} \le \dfrac{t\max(X_1,\ldots,X_m)}{\ell n(q+pe^t)}$

and for $t < 0$, $\hat{n}_{m,t} \le \dfrac{t\bar{X}}{\ell n(q+pe^t)}$.

v. For fixed m, $\lim_{t\to\infty} \hat{n}_{m,t} = \max(X_1, \ldots, X_m)$

Proof: (i) Fix $|t| < h$, let $Y_i = e^{tX_i}$, $i = 1, 2, \ldots, m$, then $\{Y_i\}$ is a sequence of iid r.v.s by the strong law of large number

$$\frac{\sum_{i=1}^{m} Y_i}{m} \xrightarrow{a.s} E(Y_1)$$

This implies that

$$\frac{\sum_{i=1}^{m} Y_i}{m} \xrightarrow{a.s.} (q+pe^t)^n$$

Therefore,

$$\ln\left(m^{-1}\sum_{i=1}^{m} e^{tX_i}\right) \xrightarrow{a.s} n\,\ln(q+pe^t)$$

and this implies that

$$\frac{\ln\left(m^{-1}\sum_{i=1}^{m} e^{X_i t}\right)}{\ln(q+pe^t)} \xrightarrow{a.s} n$$

i.e., $$p\left(\lim_{m\to\infty} \hat{n}_{m,t} = n\right) = 1.$$

(ii) Fix m, then

$$\lim_{t\to 0} \hat{n}_{m,t} = \lim_{t\to 0} \frac{\ln\left(m^{-1}\sum_{i=1}^{m} e^{X_i t}\right)}{\ln(q+pe^t)}$$

using L'Hopitial's rule we see that

$$\lim_{t\to 0} \frac{\ln\left(m^{-1}\sum_{i=1}^{m} e^{X_i t}\right)}{\ln(q+pe^t)} = \lim_{t\to 0} \frac{(q+pe^t)\sum_{i=1}^{m} X_i e^{X_i t}}{pe^t \sum_{i=1}^{m} e^{X_i t}} = \frac{\overline{X}}{p} = \hat{n}_m$$

Note: $\hat{n}_m = \frac{\overline{X}}{p}$ is actually (MME).

(iii) Since $g(x) = \ln x$ is concave for $x > 0$ then by Jensen's Inequality we have

$$E\, g(X) \le g(EX)$$

This implies that

$$E\left\{ \ln\left(m^{-1} \sum_{i=1}^{m} e^{tX_i} \right)\right\} \le n\, \ln(q+pe^t)$$

For $t > 0$, since $\ln(q+pe^t) > 0$, we have

$$E(\hat{n}_{m,t}) = E\left(\frac{\ln\left(m^{-1} \sum_{i=1}^{m} e^{X_i t} \right)}{\ln(q+pe^t)} \right) \le n$$

in a similar way we can prove that

$$E(\hat{n}_{m,t}) \ge n \quad \text{for } t < 0$$

(iv) For $t > 0$, since $x_i \le \max(x_1, \ldots, x_m)$, $i = 1, \ldots, m$, then $\sum_{i=1}^{m} e^{tx_i} \le m\, e^{t \max(x_1,\ldots,x_m)}$ or equivalently, $\ln\left(m^{-1} \sum_{i=1}^{m} e^{tx_i} \right) \le t \max(x_1, \ldots, x_m)$. Since $\ln(q+pe^t) > 0$, we have $\hat{n}_{m,t} \le \dfrac{t \max(x_1,\ldots,x_m)}{\ln(q+pe^t)}$ which is the upper bound. The lower bound comes from convexity of e^{tX_i}. Hence $t\bar{x} \le \ln\left(m^{-1} \sum_{i=1}^{m} e^{tx_i} \right)$ or equivalently $\dfrac{t\bar{X}}{\ln(q+pe^t)} \le \hat{n}_{m,t}$. Similarly, we can prove that,

$$\hat{n}_{m,t} \le \frac{t\bar{X}}{\ln(q+pe^t)}, \quad \text{for } t < 0.$$

(v) Fix m, let $X_{(1)}, \ldots, X_{(m)}$ be the order statistics of $X_1, \ldots, X_m$. Assume there are no ties. Then

$$\lim_{t\to\infty} \hat{n}_{m,t} = \lim_{t\to\infty} \frac{\ell n\left(m^{-1} \sum_{i=1}^{m} e^{X_i t}\right)}{\ell n(q+pe^t)}$$

$$= \lim_{t\to\infty} \frac{\ell n\left(m^{-1} \sum_{i=1}^{m} e^{tX_{(i)}}\right)}{\ell n(q+pe^t)}$$

using L'Hopitial rule we see that

$$\lim_{t\to\infty} \hat{n}_{m,t} = \lim_{t\to\infty} \left(\frac{(q+pe^t) \sum_{i=1}^{m} X_{(i)} e^{tX_{(i)}}}{pe^t \sum_{i=1}^{m} e^{tX_{(i)}}} \right)$$

$$= \lim_{t\to\infty} \left(\frac{\sum_{i=1}^{m-1} X_{(i)} e^{t(X_{(i)} - X_{(m)})} + X_{(m)}}{\sum_{i=1}^{m-1} e^{t(X_{(i)} - X_{(m)})} + 1} \right)$$

Since $X_{(i)} \le X_{(m)}$, for $i = 1, \ldots, m-1$ we have

$\lim_{t\to\infty} \hat{n}_{m,t} = X_{(m)}$ which is the desired result.

2.3. The asymptotic behavior of the MGF based estimator ($\hat{n}_{m,t}$)

In this section, we study the asymptotic behavior of $\hat{n}_{m,t}$, as either $m \longrightarrow \infty$ or $t \longrightarrow 0$, and we compare it with that of $\hat{n}_m$, the estimator obtained by the MME.

It can be easily shown that the limiting distribution of $\sqrt{m}\left(\hat{n}_m - n\right)$, as $m \longrightarrow \infty$ is normal with mean 0 and variance nq/p.

Before presenting our next theorem, we recall two results:

a. Let X be a r.v with binomial distribution b(n,p). Then, for $|t| < h$,

$$\text{Var}(e^{tX}) = (q+pe^{2t})^n - (q+pe^t)^{2n} = \sigma_t^2(n)$$

b. Let $\{Y_j\}_{j=1}^{\infty}$ be a sequence of iid r.v.s with $E(Y_1) = 0$ and $\text{Var}(Y_1) = \sigma^2$, then by Berry-Esseen's theorem, (S. K. Bar-Lev, N. Barken and N. A. Langberg, 1993)

$$\left| P\left(m^{-1/2}\sigma^{-1}\sum_{j=1}^{m} Y_j \le x\right) - \Phi(x)\right| \le Cm^{-1/2}\sigma^{-3}E|Y_1|^3, \ \forall\ x \in \mathbb{R} \qquad (2.3.1)$$

where Φ is the cummulative standard normal distribution and C is some constant.

Theorem 2.3.1. Let $\underline{X} = (X_1, \ldots, X_m)$ be a random sample taken from binomial distribution with parameter n and known

$p \in (0,1)$ and denote $L_{m,t} = \sqrt{m}\,(\hat{n}_{m,t} - n)$. Then

(i) For fixed t, $\lim_{n\to\infty} p(L_{m,t} \le \chi) = p[L_t \le \chi],\ \forall\, \chi \in \mathbb{R}$, where L_t is normal random variable with mean 0 and vairance

$$\eta_t^2(n) = \frac{(q+pe^{2t})^n - (q+pe^t)^{2n}}{(q+pe^t)^{2n}\,(\ell n(q+pe^t))^2}$$

(ii) $\lim_{t\to 0}\ \lim_{m\to\infty}\ p(L_{m,t} \le x) = p[L \le x],\ \forall\, x \in \mathbb{R}$, where L is a normal random variable with mean zero and variance nq/p.

(iii) $\lim_{m\to\infty}\ \lim_{t\to 0}\ p(L_{m,t} \le \chi) = p[L \le \chi],\ \forall\, \chi \in \mathbb{R}$.

Proof:

(i) $$p[\sqrt{m}\,(\hat{n}_{m,t} - n) \le x] = p\left[\sqrt{m}\left(\frac{\ell n\left(m^{-1}\sum_{i=1}^{m} e^{X_i t}\right)}{\ell n(q+pe^t)} - n\right) \le x\right]$$

$$= p\left[\sqrt{m}\left(\ell n\left(m^{-1}\sum_{i=1}^{m} e^{X_i t}\right) - \ell n(q+pe^t)^n\right) \le x\ \ell n(q+pe^t)\right]$$

$$= p\left[\ell n\left(m^{-1}\,\frac{\sum_{i=1}^{m} e^{X_i t}}{(q+pe^t)^n}\right) \le \frac{x}{\sqrt{m}}\ \ell n(q+pe^t)\right]$$

$$= p\left[m^{-1}\,\frac{\sum_{i=1}^{m} e^{X_i t}}{(q+pe^t)^n} \le (q+pe^t)^{(x/m^{1/2})}\right]$$

$$= p\left[m^{-1}\sum_{i=1}^{m} e^{X_i t} \le (q+pe^t)^{(x/m^{1/2}+n)}\right]$$

$$= p\left[m^{-1}\sum_{i=1}^{m}\left(e^{X_i t}-(q+pe^t)^n\right)\le(q+pe^t)^n\left[(q+pe^t)^{(x/m^{1/2})}-1\right]\right]$$

$$=p\left[m^{-1/2}\frac{\sum_{i=1}^{m}\left(e^{X_i t}-(q+pe^t)^n\right)}{\sigma_t(n)} \le \frac{m^{1/2}(q+pe^t)^n\left[(q+pe^t)^{(x/m^{1/2})}-1\right]}{\sigma_t(n)}\right]$$

Denote

$$H_{m,t}(x) = \frac{m^{1/2}(q+pe^t)^n\left[(q+pe^t)^{(x/m^{1/2})}-1\right]}{\sigma_t(n)} \qquad (2.3.2)$$

Apply (2.3.1) with $Y_j = e^{tx_j} - (q+pe^t)^n$, $j = 1, 2, \ldots$ leads to

$$\left|p\left(m^{-1/2}\frac{\sum_{i=1}^{m}\left(e^{X_i t}-(q+pe^t)^n\right)}{\sigma_t(n)} \le H_{m,t}(\chi)\right) - \Phi(H_{m,t}(x)\right|$$

$$\le C\, m^{-1/2}\, \sigma_t^{-3}(n)\, E\left|e^{X_1 t} - (q+pe^t)^n\right|^3 \quad \forall\, \chi \in \mathbb{R} \qquad (2.3.3)$$

The term on the right-hand side of (2.3.3) converges to 0 as $m \longrightarrow \infty$. This implies that,

$$\lim_{m\to\infty} p\left(\sqrt{m}\left(\hat{n}_{m,t}-n\right) \le x\right) = \lim_{m\to\infty} \Phi\left(H_{m,t}(\chi)\right).$$

Now

$$\lim_{m\to\infty} H_{m,t}(\chi) = \lim_{m\to\infty} \sqrt{m}\ \frac{(q+pe^{t})^{n}\left[(q+pe^{t})^{x/m^{1/2}} - 1\right]}{\sigma_t(n)}$$

Using L'Hopitial rule we see that

$$\lim_{m\to\infty} H_{m,t}(x) = \chi\ \frac{\ln(q+pe^{t})(q+pe^{t})^{n}}{\sigma_t(n)} \tag{2.3.4}$$

Hence

$$\lim_{m\to\infty} p\left[\sqrt{m}\ (\hat{n}_n-n) \le x\right] = \Phi\left(\frac{(q+pe^{t})^{n}\ \ln(q+pe^{t})x}{\sqrt{(q+pe^{2t})^{n} - (q+pe^{t})^{2n}}}\right)$$

which ends the proof of part (i).

(ii) Taking limits, as $t \longrightarrow 0$, in (2.3.4), yields

$$\lim_{t\to 0}\ \lim_{m\to\infty} H_{m,t}(x) = x \lim_{t\to 0} \left\{\frac{(\ln(q+pe^{t}))^{2}(q+pe^{t})^{2n}}{(q+pe^{2t})^{n}-(q+pe^{t})^{2n}}\right\}^{\frac{1}{2}}$$

by using L'Hopital rule twice we see that

$$\lim_{t\to 0}\ \lim_{m\to\infty} H_{m,t}(x) = x \left\{\frac{p}{nq}\right\}^{\frac{1}{2}}$$

Taking limits in (2.3.3) as $m \longrightarrow \infty$ and then as $t \longrightarrow 0$ and using (2.3.4), implies

$$\lim_{t\to 0}\lim_{m\to\infty} p\left(\sqrt{m}\left(\hat{n}_{m,t}-n\right) \le x\right) = \lim_{t\to 0}\lim_{m\to\infty} \Phi\left(H_{m,t}(x)\right)$$

$$= \Phi\left(x\sqrt{p/nq}\right) = p[L \le x].$$

(iii) Let $A_{m,t}$ denote the upper bound in (2.3.3). We will show that

$$\lim_{m\to\infty}\lim_{t\to 0} A_{m,t} = 0 \qquad (2.3.5)$$

and

$$\lim_{m\to\infty}\lim_{t\to 0} H_{m,t}(x) = x\left\{\frac{p}{nq}\right\}^{\frac{1}{2}} \qquad (2.3.6)$$

The relations (2.3.5) and (2.3.6) would imply

$$\lim_{m\to\infty}\lim_{t\to 0} p\left(\sqrt{m}\left(\hat{n}_{m,t}-n\right) \le x\right) = \lim_{m\to\infty}\lim_{t\to 0} \Phi\left(H_{m,t}(x)\right)$$

$$= \Phi\left(\chi\sqrt{p/nq}\right) = p[L \le x].$$

Now to show (2.3.5). By an application of Liapounov's inequality $\left((E|Y_1|^r)^{1/r} \le (E|Y_1|^s)^{1/s},\ 0 < r < s < \infty,\right.$ for the r.v $Y_1 = e^{X_i t} - (q+pe^t)^n$ with $r = 3$ and $s = 4\left.\right)$ we have

$$A_{m,t} \le \frac{c\left[E\left(e^{X_i t}-(q+pe^t)^n\right)^4\right]^{3/4}}{m^{1/2}\left[(q+pe^{2t})^n-(q+pe^t)^{2n}\right]^{3/2}} = \frac{B_{m,t}}{C_{m,t}}\ ;\ \text{say}$$

Expand $(e^{x_i t} - (q+pe^t)^n)^4$ and taking the expectation then by using L'Hopital's rule twice, we find that $\lim_{t\to 0} \left(\frac{B_{m,t}}{C_{m,t}}\right)^{4/3} = 0$ the desired result.

To show (2.3.6), taking limits as $t \longrightarrow 0$, in (2.3.2) and using L'Hopital rules twice results in

$$\lim_{t\to 0} H_{m,t}(\alpha) = \alpha \left\{\frac{p}{nq}\right\}^{\frac{1}{2}}$$

and hence (2.3.6) follows, which completes the proof.

The relative asymptotic efficiency, defined by $eff_t(n) = \frac{nq/p}{\eta_t^2(n)}$, of the $\hat{n}_{m,t}$ with respect to $\hat{n}_m$ is summarized in Table 2.A for t = 0.05, n = 3, 6, 9, 12, 15 and p = 0.25, 0.5, 0.75. It can be seen that the relative asymptotic efficiency is always greater than one.

Table 2.A

Values of $eff_t(n)$ where t = 0.05

	n				
p	3	6	9	12	15
.25	1.011959	1.011341	1.010596	1.009874	1.009119
.5	1.025066	1.024135	1.023188	1.022227	1.021268
.75	1.237803	1.037101	1.036423	1.035702	1.034984

2.4 Derivation of the maximum likelihood estimator ($\hat{n}_L$)

Let $\underline{X} = (X_1, \ldots, X_m)$ be a random sample from b(n,p), where p is known. Let L(n) be the likelihood of n given $X_i = x_i$, (i = 1, ..., m) and let $\lambda(n) = L(n)/L(n-1)$. The MLE $\hat{n}_L$ is an integer solution of

$$\lambda(n) \geq 1 \text{ and } \lambda(n+1) < 1 \text{ for } n \geq \max(x_1, \ldots, x_m). \qquad (2.4.1)$$

But, $\lambda(n) = (nq)^m \prod_{i=1}^{m} (n-x_i)^{-1}$, for $n \geq \max(x_1,\ldots,x_m)$ where $q = 1-p$. Thus, (2.4.1) becomes

$$\prod_{i=1}^{m} (n-x_i)^{-1} \leq (nq)^m \text{ and } \prod_{i=1}^{m} (n+1-x_i) > \left((n+1)q\right)^m \qquad (2.4.2)$$

Now ignore the integer character of $\hat{n}_L$ and consider the equation obtained by replacing the inequality by equality in the first part of (2.4.1), set $z = 1/n$, we have

$$\prod_{i=1}^{m} (1-x_i z) = q^m \qquad (2.4.3)$$

Let p(z) be the left side of (2.4.3), then p(0) = 1 and $p\left(\frac{1}{\max(x_1, \ldots, x_m)}\right) = 0$, and p is strictly decreasing in z and convex on $\left[0, 1/\max(x_1,\ldots,x_m)\right]$. Hence there is a unique root $\hat{z}$ of (2.4.3) in this interval. So that there is a unique root of (2.4.2). This root is $\hat{n} = [1/\hat{z}]$, where [a] is the largest integer in a. Also, from convexity of p on $\left[0, 1/\max(x_1, \ldots, x_m)\right]$ bounds on $\hat{z}$ are obtainable. These are

$$\frac{1-q^m}{\sum_{i=1}^{m} X_i} \le \hat{z} \le \frac{1-q^m}{\max(X_1, \ldots, X_m)}$$

or equivalently,

$$\frac{\max(X_1, \ldots, X_m)}{1-q^m} \le \hat{n}_L \le \frac{\sum_{i=1}^{m} X_i}{1-q^m}$$

This description of the MLE of n was found by Feldman and Fox (1968). It should be noted that the MLE is not unique. For example, suppose n = 2, p = 1/2 and let (X_1, X_2) be a random sample of size 2, then the sample space is

$$S = \left\{(0,0),(0,1),(0,2),(1,2),(2,1),(1,1),(2,2),(1,0),(2,0)\right\}$$

$$L(n \mid X_1 = x_1, X_2 = x_2) = \binom{n}{x_1}\binom{n}{x_2}\left(\frac{1}{2}\right)^{2n}, \quad n \ge \max(x_1, x_2).$$

For the sample points

(0,0) the MLE is $\hat{n} = 1$

(0,1) the MLE is $\hat{n} = 1$

(0,2) the MLE is $\hat{n} = 2$

(1,1) $\Longrightarrow [4/3] \le \hat{n} \le [8/3] \Longrightarrow \hat{n} = 1, 2$

because $L(1 \mid 1,1) = \binom{1}{1}\binom{1}{1}\left(\frac{1}{2}\right)^2 = 1/4$

$$L(2 \mid 1,1) = \binom{2}{1}\binom{2}{1}\left(\frac{1}{2}\right)^4 = 1/4$$

so we have two MLE.

$(2,2) \Longrightarrow [8/3] \le \hat{n} \le [16/3],\ \hat{n} = 2, 3, 4, 5$

$$L(2|2,2) = \binom{2}{2}\binom{2}{2}\left(\frac{1}{2}\right)^{4} = 1/16$$

$$L(3|2,2) = \binom{3}{2}\binom{3}{2}\left(\frac{1}{2}\right)^{6} = 9/64$$

$$L(4|2,2) = \binom{4}{2}\binom{4}{2}\left(\frac{1}{2}\right)^{8} = 9/64$$

$$L(5|2,2) = \binom{5}{2}\binom{5}{2}\left(\frac{1}{2}\right)^{10} = 6.25/64$$

So, the MLE is at $\hat{n} = 3$ or 4.

In numerical simulation for the MLE if we have more then one we take the largest one. To calculate the exact distribution of the MLE it is very difficult and take long time. So, we used simulation method to approximate the mean square error and bias for it. It observed that numerical simulation is very close to the exact. (see Table 2.1 for exact) and (Table 2.2 for simulation).

We describe the procedure that we have used to calculate the MSE and Bias of $\hat{n}_L$ and $\hat{n}_{m,t}$.

1. Fix n, p, t, m.
2. Generate m observations from b(n,p), $(x_1, \ldots, x_m)$, say
3. Substitute the observations obtained in step 2 in the following equation

$$\hat{n}_{m,t} = \frac{\ln \sum_{i=1}^{m} e^{x_i t}/m}{\ln(q+pe^{t})}$$

4. If all observations in step 2 are zero, we put $\hat{n}_L = 1$.

5. If at least one observation in step (2) not equal zero we substitue these observations in the following bounds

$$\left[\frac{\max\limits_{1\le i\le n}(x_i)}{1-q^m}\right] \le \hat{n}_L \le \left[\frac{\sum_{i=1}^{m} x_i}{1-q^m}\right]$$

6. Order the values obtained in (5) and substitute these values in the likelihood function and the maximum likelihood estimator is the value say $\hat{n}_L$ with large probability.

7. If we have more than one value that maximize likelihood we take the largest one.

8. To approximate the mean square error and the bias of $\hat{n}_{m,t}$ and $\hat{n}_L$ we repeat steps (2-8) 10000 times.

$$MSE(\hat{n}_{m,t}) = \sum_{i=1}^{10000} \left(\hat{n}_{m,t}(i)-n\right)^2 \Big/ 10000,$$

$$Bias(\hat{n}_{m,t}) = \left(\sum_{i=1}^{10000} \hat{n}_{m,t}(i) \Big/ 10000\right) - n$$

$$MSE(\hat{n}_L) = \sum_{i=1}^{10000} \left(\hat{n}_L(i)-n\right)^2 \Big/ 10000$$

$$\text{Bias}(\hat{n}_L) = \left(\sum_{i=1}^{10000} \hat{n}_L(i) \Big/ 10000 \right) - n$$

Table 2.1

Exact MSE and bias of $\hat{n}_L$, $\hat{n}_{m,t}$

$t = 0.05$, $p = 1/2$, $n = 3$

M	MSE($\hat{n}_L$)	MSE($\hat{n}_{m,t}$)	bias($\hat{n}_{m,t}$)	bias($\hat{n}_L$)	$eff_t(n)$
3	.9804687	.9729873	-.012353	-.167735	1.0045917
6	.5167691	.4882278	-.006164	-.130829	1.058459
9	.3152752	.327507	-.004848	-.058835	0.962626517

Table 2.2

Simulated MSE and bias of $\hat{n}_L$, $\hat{n}_{m,t}$

$t = 0.05$, $p = 1/2$, $n = 3$

M	MSE($\hat{n}_L$)	MSE($\hat{n}_{m,t}$)	bias($\hat{n}_{m,t}$)	bias($\hat{n}_L$)	$eff_t(n)$
3	.9757376	966165	-.014293	-.173263	1.009908
6	.5148777	.4843701	-.005648	-.131881	1.062984
9	.3221941	.3252513	-.005568	-.054012	0.9906005

2.5 Numerical comparisons

In this section, we make comparisons among the various estimators; the estimator based on the MGF $\hat{n}_{m,t}$, MLE $\hat{n}_L$ and the MME $\hat{n}_m$ for small m,n. 10000 b(n,p) samples of size m were simulated for m,n = 3, 9, 15, t = -1, 0.05, 1 and p = 0.05, 0.25, 0.5, 0.75, 0.95. The efficiency of the estimator $\hat{\theta}_1$ with respect to the estimator $\hat{\theta}_2$ is defined by eff = $MSE(\hat{\theta}_2)/MSE(\hat{\theta}_1)$, also the absolute ratio of the bias of $\hat{\theta}_1$ with respect to $\hat{\theta}_2$ is defined by $R = |bias(\hat{\theta}_2)/bias(\hat{\theta}_1)|$. We report the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$, in Tables (2.3, 2.5, 2.7, 2.9). Tables (2.4, 2.6, 2.8, 2.10) present the absolute ratio of bias of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$. Tables (2.11-2.14) present the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_m$. Tables (2.15, 2.17, 2.19) present the efficiency of $\hat{n}^*_{m,t}$ with respect to $\hat{n}_L$. Tables (2.16, 2.18, 2.20) present the abolute ratio of bias of $\hat{n}^*_{m,t}$ with respect to $\hat{n}_L$. Tables (2.21-2.23) present the efficiency of $\hat{n}^*_{m,t}$ with respect to the $\hat{n}_L$.

Table 2.3

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$ [t = 2].

m	p	n = 3	9	15
	.05	1.989871	1.595142	1.264394
	.025	1.421788	0.723858	0.469198
3	.5	1.272793	0.660346	0.416110
	.75	1.324184	0.895239	0.613240
	.95	0.373053	0.592724	1.036242
	.05	1.281042	1.067459	0.769975
	.25	1.119503	0.552734	0.294994
9	.5	1.450276	0.591168	0.310664
	.75	0.150738	1.059167	0.636965
	.95	0.000031	0.000921	0.139397
	.05	1.279114	0.851484	0.577984
	.25	1.227665	0.423303	0.214024
15	.5	1.357525	0.535878	0.287554
	.75	0.004980	0.018975	0.714511
	.95	0.000017	0.000820	0.00459

Table 2.4

The absolute ratio of bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_L$ [t = 2]

m	p	n = 3	n = 9	n = 15
3	.05	0.395413	0.109536	0.103693
	.25	0.197019	0.004028	0.001996
	.5	0.484630	0.084990	0.039105
	.85	0.636210	0.465930	0.176206
	.95	0.755923	0.972863	1.470802
9	.05	0.250907	0.035184	0.022835
	.25	0.118323	0.002296	0.001221
	.5	0.052229	0.082650	0.007888
	.75	0.170706	0.316367	0.109144
	.95	0.013039	0.015384	0.158771
15	.05	0.340416	0.005114	0.004049
	.25	0.097128	0.002021	0.036253
	.5	0.209021	0.054236	0.0176200
	.75	0.009951	0.048423	0.017230
	.95	0.140873	0.024221	0.013961

Table 2.5

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$ [t = 1].

m	p	n = 3	n = 9	n = 15
	.05	1.269864	1.370430	0.842082
	.25	1.299266	1.04053	0.818082
3	.5	1.279126	0.923317	0.740842
	.75	1.228654	1.09683	0.909233
	.95	0.100076	0.537870	0.625634
	.05	0.943609	0.850185	0.659413
	.25	1.139656	0.838218	0.655401
9	.5	1.293633	0.882374	0.649167
	.75	0.135088	1.265701	0.951527
	.95	0.00039	0.00699	0.020652
	.05	1.022702	0.906629	0.758253
	.25	1.235479	0.774747	0.567122
15	.5	1.176499	0.896688	0.611030
	.75	0.016361	1.203924	1.044713
	.95	0.000021	0.00043	0.003890

Table 2.6

The absolute ratio of the bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_L$ [t = 1]

m	p	n = 3	n = 9	n = 15
	.05	0.825983	0.386636	0.1127580
	.25	0.401885	0.0486201	0.062161
3	.5	0.786681	0.159047	0.096813
	.75	1.449943	0.768646	0.283272
	.95	0.229683	0.770135	0.852212
	.05	0.547785	0.113836	0.109489
	.25	0.234830	0.048091	0.037488
9	.5	0.889333	0.121183	0.06960
	.75	0.288125	0.754149	0.244279
	.95	0.020526	0.003987	0.023076
	.05	0.245892	0.141942	0.1621061
	.25	0.169008	0.056021	0.047979
15	.5	0.185700	0.127661	0.071247
	.75	0.042788	0.0660421	0.263962
	.95	0.004432	0.045653	0.010257

Table 2.7

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$ [t = 0.05].

m	p	n = 3	n = 9	n = 15
	.05	0.9851801	0.982903	0.9331924
	.25	1.012898	1.006888	1.013526
3	.5	1.009908	1.005548	1.025901
	.75	0.876528	1.023675	1.052734
	.95	0.060164	0.338697	0.681608
	.05	0.961623	1.013964	1.247751
	.25	0.995214	1.02137	1.024149
9	.5	0.990601	1.038449	1.038569
	.75	0.067519	1.068099	1.079486
	.95	0.008132	0.00988	0.095657
	.05	0.983558	1.002982	1.021993
	.25	1.107644	1.046914	1.025015
15	.5	0.857825	1.118694	1.075365
	.75	0.006197	0.983293	1.135589
	.95	0.000739	0.00098	0.001465

Table 2.8

The absolute ratio of the bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_L$ [t = 0.05]

		n		
m	p	3	9	15
3	.05	5.472289	12.48657	1.040615
	.25	20.5223	1.590919	1.379380
	.5	12.121913	3.79622	3.325263
	.75	11.683667	24.108563	10.291247
	.95	0.564672	5.894948	10.49824
9	.05	1.95555	1.258899	1.261764
	.25	23.496808	2.287466	1.88495
	.5	9.700848	3.418789	2.759835
	.75	1.03789	19.445216	3.877796
	.95	0.429964	1.191996	4.192691
15	.05	3.388691	1.494286	1.150074
	.25	2.464699	2.20457	16.010413
	.5	5.574175	2.42572	1.758977
	.75	0.97979	3.339062	5.00366
	.95	0.280992	1.494671	3.982758

Table 2.9

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$ [t = -1].

M	p	n = 3	9	15
3	.05	0.657081	0.604859	0.570910
	.25	0.558287	0.418631	0.343407
	.5	0.470749	0.353148	0.278980
	.75	0.392552	0.364635	0.309866
	.95	0.010267	0.1173909	0.3277608
9	.05	0.752504	0.747433	0.681547
	.25	0.657893	0.471071	0.359412
	.5	0.499291	0.331505	0.249526
	.75	0.027179	0.298749	0.238082
	.95	0.000017	0.000316	0.011584
15	.05	0.85737	0.787521	0.765073
	.25	0.761280	0.511921	0.363899
	.5	0.4252617	0.3341568	0.233327
	.75	0.001217	0.245976	0.235467
	.95	0.000017	0.000038	0.000417

Table 2.10

The absolute ratio of bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_L$ [t = -1]

m	p	n = 3	n = 9	n = 15
	.05	0.526609	0.398023	0.069602
	.25	0.393462	0.0854634	0.0518936
3	.5	0.480341	0.0199609	0.063781
	.75	0.568036	0.295881	0.117089
	.95	0.014651	0.208186	0.639248
	.05	0.343795	0.505996	0.581531
	.25	0.36009	0.0466817	0.0423459
9	.5	0.510737	0.0122360	0.039792
	.75	0.087981	0.201624	0.048263
	.95	0.005380	0.011577	0.168258
	.05	0.985566	0.252599	0.209373
	.25	0.160373	0.101376	0.070041
15	.5	0.140424	0.0036401	0.021807
	.75	0.005454	0.184924	0.074518
	.95	0.010048	0.023857	0.012412

Table 2.11

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_m$ [t = 2].

m	p	n = 3	n = 9	n = 15
	.05	1.35274	1.178737	0.87737
	.25	1.431075	0.746819	0.454183
3	.5	1.266861	0.636938	0.396850
	.75	1.694005	0.935667	0.624440
	.95	2.801843	1.279732	0.597921
	.05	1.550167	1.087864	0.741317
	.25	1.106577	0.523834	0.291600
9	.5	1.518651	0.543441	0.321664
	.75	1.925708	0.970797	0.673763
	.95	2.503138	1.579317	0.53433
	.05	1.338856	0.786323	0.618060
	.25	1.026745	0.447554	0.238818
15	.5	1.693533	0.532610	0.284204
	.75	1.219808	0.185311	0.596574
	.95	2.31463	1.81828	0.724325

Table 2.12

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_m$ [t = 1].

m	p	n = 3	n = 9	n = 15
	.05	1.368778	1.481109	0.300324
	.25	1.340084	1.038133	0.823879
3	.5	1.272209	0.935661	0.751457
	.75	1.432502	1.077087	0.880930
	.95	1.751932	1.631943	0.472134
	.05	1.73744	1.073567	0.709085
	.25	1.161101	0.813796	0.641130
9	.5	1.366343	0.842337	0.645624
	.75	1.705657	1.208568	0.897665
	.95	1.347099	1.023588	0.775731
	.05	1.070953	0.8675864	0.8569815
	.25	1.159534	0.745183	0.558807
15	.5	1.405595	0.839729	0.580976
	.75	1.817863	1.278712	0.928415
	.95	1.260658	0.998493	0.75652

Table 2.13

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_m$ [t = 0.05].

m	p	n = 3	n = 9	n = 15
3	.05	1.026052	1.041879	1.014246
	.25	1.05418	1.03875	1.01078
	.5	1.03502	1.025117	1.016897
	.75	1.024822	1.019362	1.031354
	.95	1.061593	1.025318	1.028297
9	.05	1.046351	1.006781	1.000981
	.25	1.023957	1.005703	1.006831
	.5	1.024849	1.019455	1.017359
	.75	1.013798	1.043497	1.017170
	.95	1.00537	1.076419	1.054783
15	.05	1.005474	1.058904	1.075064
	.25	1.006336	1.003908	1.010101
	.5	1.027322	1.045187	1.03867
	.75	1.032886	1.047713	1.016765
	.95	0.979736	1.073113	0.9908763

Table 2.14

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_m$ [t = -1].

m	p	n = 3	n = 9	n = 15
	.05	0.748034	0.638175	0.599662
	.25	0.5910668	0.423564	0.346461
3	.5	0.485880	0.359082	0.279608
	.75	0.475301	0.3692396	0.297316
	.95	0.571936	0.498410	0.446502
	.05	0.827078	0.773413	0.772995
	.25	0.689451	0.460698	0.356935
9	.5	0.531904	0.323947	0.244372
	.75	0.419431	0.296031	0.228841
	.95	0.395580	0.336631	0.325651
	.05	0.935735	0.818368	0.801056
	.25	0.730600	0.501444	0.355690
15	.5	0.25522	0.102064	0.216920
	.75	0.405639	0.260442	0.206015
	.95	0.357410	0.337048	0.276999

Table 2.15

The efficiency of the $\hat{n}^{*}_{m,t}$ with respect to $\hat{n}_L$ [t = 2]

m	p	n = 3	n = 9	n = 15
3	.25	1.70475	0.675629	0.421447
	.5	0.974921	0.658717	0.409237
	.75	1.388537	0.720264	0.564304
9	.25	1.013225	0.466412	0.280216
	.5	1.0095	0.455903	0.334924
	.75	0.304527	0.848754	0.680597
15	.25	0.915272	0.432549	0.231687
	.5	1.118302	0.482368	0.282226
	.75	0.076923	0.748601	0.622032

Table 2.16

The aboluste ratio of the bias of $\hat{n}^{*}_{m,t}$ w.r.t $\hat{n}_{L}$ [t = 2]

		n		
m	p	3	9	15
	.25	0.182088	0.032146	0.021109
3	.5	0.557322	0.078692	0.043493
	.75	0.596626	0.530978	0.014688
	.25	0.073160	0.024005	0.021129
9	.5	0.407792	0.063918	0.005570
	.75	0.290649	0.242811	0.116868
	.25	0.066026	0.024231	0.000621
15	.5	0.068190	0.012228	0.022519
	.75	0.097339	0.199757	0.046082

Table 2.17

The efficiency of $\hat{n}^*_{m,t}$ with respect to $\hat{n}_L$ [t = 0.05]

m	p	n = 3	n = 9	n = 15
3	.25	1.061253	1.019131	1.010989
	.5	0.938007	0.9622576	1.027417
	.75	0.549461	0.969741	1.022429
9	.25	0.920031	1.001176	1.004161
	.5	0.711278	0.964706	0.994063
	.75	0.062942	0.876531	0.915467
15	.25	0.985438	0.990395	1.000318
	.5	0.728820	0.965765	0.987995
	.75	0.007605	0.752461	0.927575

Table 2.18

The absolute ratio of the bias $\hat{n}^*_{m,t}$ w.r.t $\hat{n}_L$ [t = 0.05]

		n				
m	p	3	6	9	12	15
3	.25		0.831000	1.800860		1.370550
	.5		37.424000	5.165000		2.657670
	.75		0.904170	5.381400		4.592300
9	.25		0.649000	1.694830		1.390710
	.5		12.081000	11.329900		2.30314
	.75		0.552550	2.612400		11.094500
15	.25		0.240000	3.310640		1.152700
	.5		2.108000	18.102200		1.626580
	.75		0.035280	7.133700		3.956200

Table 2.19

The efficiency of the $\hat{n}^{*}_{m,t}$ with respect to $\hat{n}_L$ [t = -1]

		n		
m	p	3	9	15
	.25	0.349863	0.373818	0.293825
3	.5	0.482567	0.341241	0.273957
	.75	0.349863	0.373818	0.293825
	.25	0.019324	0.286070	0.243079
9	.5	0.425269	0.323523	0.243527
	.75	0.0193242	0.286070	0.243079
	.25	0.008379	0.231184	0.206069
15	.5	0.366985	0.312207	0.224727
	.75	0.008379	0.231184	0.206069

Table 2.20

The aboluste ratio of the bias $\hat{n}^{*}_{m,t}$ w.r.t $\hat{n}_L$ [t = -1]

		n		
m	p	3	9	15
3	.25	0.403518	0.064980	0.028632
	.5	0.531490	0.106637	0.061991
	.75	0.822629	0.336560	0.098027
9	.25	0.329222	0.086222	0.066088
	.5	0.344678	0.100354	0.043748
	.75	0.085867	0.160238	0.093960
15	.25	0.232705	0.131970	0.026079
	.5	0.312395	0.074568	0.042025
	.75	0.004895	0.165884	0.062894

Table 2.21

The efficiency of the $\hat{n}^*_{m,t}$ with respect to $\hat{n}_m$ [t = 2]

m	p	n = 3	n = 9	n = 15
3	.25	1.728536	0.675398	0.428670
	.5	0.968492	0.653462	0.401246
	.75	1.658352	0.712193	0.548523
9	.25	1.033450	0.468046	0.274840
	.5	1.032969	0.438355	0.328276
	.75	1.572530	0.837549	0.639139
15	.25	0.830554	0.139599	0.229710
	.5	1.39177	0.440586	0.267635
	.75	1.094017	0.799349	0.559198

Table 2.22

The efficiency of the $\hat{n}^*_{m,t}$ with respect to $\hat{n}_m$ [t = 0.05]

		n		
m	p	3	9	15
3	.25	1.082536	1.025549	1.004157
	.5	0.953901	0.985101	1.022615
	.75	0.665770	0.945506	0.991580
9	.25	0.948555	0.977606	0.976089
	.5	0.728633	0.948738	0.970057
	.75	0.985016	0.862916	0.859995
15	.25	0.910055	0.976927	0.953008
	.5	0.851405	0.892987	0.945331
	.75	1.267421	0.78082	0.821252

Table 2.23

The efficiency of the $\hat{n}^*_{m,t}$ with respect to $\hat{n}_m$ [t = -1]

m	p	n = 3	n = 9	n = 15
3	.25	0.570434	0.407678	0.341646
	.5	0.488015	0.346704	0.277185
	.75	0.406346	0.376669	0.288444
9	.25	0.657239	0.467231	0.357964
	.5	0.446343	0.317668	0.238736
	.75	0.315755	0.277771	0.229905
15	.25	0.626678	0.486458	0.353678
	.5	0.435843	0.296588	0.212220
	.75	0.279287	0.240814	0.188921

2.6. Results of numerical comparison

1. Tables (2.3, 2.5, 2.7, 2.9) present the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_L$ for t = -1, 0.05, 1 and 2.
   a. For t = 1, the tabulated results suggest that, high efficiency (close to 1) is achieved for values of n = 3, 9, m = 3 and p = 0.05, 0.25, 0.5, 0.75.
   b. For t = 0.05, high efficiency (close to 1) is achieved for values of n ≤ 15, m < 15 and p = 0.05, 0.25, 0.5, 0.75.
   c. For t = -1, the efficiency is always less than 1 for all values of n, m and p.
   d. For t = 2, the tabulated results suggest that, thigh efficiency (close to 1) is achieved for n = 3, 9, m = 3, 9, 15 and p = 0.05, also for n = 3, m = 3, 9 , 15 and p = 0.25, 0.5.
2. Tables (2.4, 2.6, 2.8, 2.10) present the absolute ratio of bias of $\hat{m}_{m,t}$ with respect to $\hat{n}_L$ for values of t = -1, 0.05, 1, 2.
   a. For t = -1, the ratio of the absolute bias is less than 1, for all m, n and p.
   b. For t = 0.05, the ratio of the absolute bias is greater than 1, for all m, n and p.
   c. For t = 1, the ratio of the absolute bias is less than 1 for nearly almost m, n and p.
   d. For t = 2, the ratio of the absolute bias is less than 1 for all m, n and p.

3. Tables 2.11-2.14, present the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_m$ for $t = -1, 0.05, 1, 2$.

   a. For $t = 1$, the tabulated results suggest that high efficiency (close to 1) is achieved for values of $n \le 9$, $m \le 15$ and $p = 0.05, 0.25, 0.5, 0.75, 0.95$.

   b. For $t = 0.05$, high efficiency (close to 1) is achieved for values of $n \le 15$, $m \le 15$ and $p = 0.05, 0.25, 0.5, 0.75, 0.95$.

   c. For $t = -1$, the efficiency is less than 1 for all m, n and p.

   d. For $t = 2$, high efficiency (close to 1) is achieved for $n = 3$, $m = 3, 9, 15$ and $p = 0.05, 0.25, 0.5, 0.75, 0.95$ also for $n = 9$, $m = 3, 9$ and $p = 0.05, 0.95$.

4. Tables 2.15, 2.17, 2.19, present the efficiency of $\hat{n}^*_{m,t}$ with respect to $\hat{n}_L$ for $t = -1, 0.05, 2$.

   a. For $t = 0.05$, the tabulated results suggest that high efficiency (close to 1) is achieved for values of $n > 6$, $m \le 15$ and $p = 0.25, 0.5$.

   b. For $t = -1$, the efficiency is always less than 1, for all m, n and $p = 0.25, 0.5, 0.75$.

   c. For $t = 2$, the efficiency (close to 1) for $n = 3$, $m \le 15$ and $p = 0.25, 0.5$.

5. Tables 2.16, 2.18, 2.20, present the absolute ratio of bias of $\hat{n}^*_{m,t}$ with respect to $\hat{n}_L$ for $t = -1, 0.05, 2$.

   a. For $t = 0.05$, the ratio of the absolute bias is greater than 1, for almost all values of $n \le 15$, $m \le$

15 and $p = 0.25, 0.5, 0.75$.

b. For $t = -1$, the ratio of the absolute bias is always less than 1, for all m, n and p.

c. For $t = 2$, the ratio of the absolute bias is always less than 1 for all m, n and p.

6. Tables 2.21-2.23, present the efficiency of $\hat{n}^*_{m,t}$ with respect to $\hat{n}_m$ for $t = -1, 0.05, 2$.

a. For $t = 0.05$, the efficiency is less than 1, but it is still high for m, n and p.

b. For $t = -1$, the efficiency is always less than 1 for all m, n and p.

c. For $t = 2$, high efficiency (close to 1) is achieved for values of $n = 3$, $m \le 15$ and $p = 0.25, 0.5, 0.75$.

2.7. Derivation of the estimator based on moment generating function when p is unknown

Let $\underline{X} = (X_1, \ldots, X_m)$ be a random sample from $b(n,p)$, where both n and p are unknown. The MGF based estimator for n, on the basis of the sample $\underline{X} = (X_1, \ldots, X_m)$ is a solution of the equations

$$E_n\left(e^{t_1 X_1}\right) = \frac{\sum_{i=1}^{m} e^{t_1 x_i}}{m} \qquad (2.7.1)$$

$$E_n\left(e^{t_2 X_1}\right) = \frac{\sum_{i=1}^{m} e^{t_2 x_i}}{m} \qquad (2.7.2)$$

for some t_1 and t_2.

From (1.3.1) and (1.3.2) we have

$$\frac{\ln(q+pe^{t_1})}{\ln(q+pe^{t_2})} = \frac{\ln\left(\sum_{i=1}^{m} e^{t_1 x_i}/m\right)}{\ln\left(\sum_{i=1}^{m} e^{t_2 x_i}/m\right)} = C \text{ , say}$$

This implies,

$$\ln(q+pe^{t_1}) = C \ln(q+pe^{t_2})$$

or

$$\left(q+pe^{t_2}\right)^C - \left(q+pe^{t_1}\right) = 0 \qquad (2.7.3)$$

If a solution of (2.7.3) exists $\hat{p}$ say, then the estimator of n is

$$\hat{n}(\hat{p},t_1) = \frac{\ln\left(\sum_{i=1}^{m} e^{t_1 x_i}/m\right)}{\ln(\hat{q}+\hat{p}e^{t})} ,$$

Also, we suggest the following estimators of p:

i. $\hat{p} = 1/2$.

ii. The maximum likelihood estimator of p is $\bar{X}/n$ and for large m one can estimate n by $\max(X_1, \ldots, X_m)$, thus $\hat{p} = \dfrac{\bar{X}}{\max(X_1,\ldots,X_m)}$.

iii. $\hat{p} = \bar{X}/\hat{n}$, where $\hat{n}$ is the value of $\hat{n}_{L:s}$.

iv. $\hat{p} = \bar{X}/\hat{n}$, where $\hat{n}$ is the value of $\hat{n}_{m:s}$.

2.8. Numerical work

In this section, we analyze the examples listed in Table (2) of Olkin, Petkau and Zidek, (1981) who computed $\hat{n}_m$, $\hat{n}_{m:s}$, $\hat{n}_L$ and $\hat{n}_{L:s}$ for some cases In Tables 2.24-2.28, it is clear that $\hat{n}_m$ and $\hat{n}_L$ are highly unstable. In addition, $\hat{n}_{m:s}$, $\hat{n}_{L:s}$, $\hat{n}(\hat{p},t)$ are clearly stable, with $\hat{n}_{m:s}$, $\hat{n}(\hat{p},t)$ [where $\hat{p} = \overline{X}/\hat{n}$ and $\hat{n}$ is the value of $\hat{n}_{m:s}$ giving rather similar results. Also, $\hat{n}_{L:s}$ and $\hat{n}(\hat{p},t)$ [where $\hat{p} = 1/2$, $\overline{X}/\max(X_1,\ldots,X_m)$ or the root of the equation (2.7.3) for t = 1, 2, giving] rather similar results.

Table 2.24

The last column present the estimator $\hat{n}(\hat{p},t)$ where $\hat{p}$ is the solution of equation (2.7.3)

n	p	m	sample	$\hat{n}_m$	$\hat{n}_{m:s}$	$\hat{n}_L$	$\hat{n}_{L:s}$	$\hat{n}(\hat{p},1)$
75	.32	5	16,18,22,25,27	102	70	99	29	28
			16,18,22,25,28*	195	80	190	30	29
34	.57	4	14,18,20,26	507	77	504	31	27
				< 0	91	∞	31	28
37	.17	20	4,4,4,4,5,5,5	65	25	66	11	12
			5,6,6,6,6,7,9,9	154	27	159	13	13
			10,10,10,11,11					
48	.06	15	0,1,1,2,2,2,3,3	18	10	15	7	7
			3,4,4,4,4,5,6	135	12	125	9	9
40	.17	12	6,7,7,7,8,8,9,9	32	26	42	21	18
			9,10,11,16	61	32	79	32	19
55	.48	20	17,23,24,25,25	71	69	71	43	40
			26,26,26,27,27	79	74	81	45	41
			28,28,28,29,30					
			30,30,31,33,38					
60	.24	15	11,11,12,12,13	67	49	67	24	23
			13,14,16,17,17	88	53	90	28	25
			18,18,20,20,22					

* This is the perturbed sample obtained by adding one to the largest success count. For simplicity, the perturbed samples are not displayed in the remaining cases.

Table 2.25

The last two columns present the estimator $\hat{n}(\hat{p},t)$ where $\hat{p} = \bar{X}/\hat{n}$, $\hat{n}$ is the value of $\hat{n}_{m:s}$ and $t = 1, 2$

n	p	m	sample	$\hat{n}_m$	$\hat{n}_{m:s}$	$\hat{n}_L$	$\hat{n}_{L:s}$	$\hat{n}(\hat{p},t)$	
75	.32	5	16,18,22,25,27	102	70	99	29	59	48
			16,18,22,25,28*	195	80	190	30	68	53
34	.57	4	14,18,20,26	507	77	504	31	68	52
				< 0	91	∞	32	80	60
37	.17	20	4,4,4,4,5,5,5	65	25	66	11	23	19
			5,6,6,6,6,7,9	154	27	159	13	26	21
			9,10,10,10,11,11						
48	.06	15	0,1,1,2,2,2,3,3	18	10	15	7	9	8
			3,4,4,4,4,5,6	135	12	125	9	12	11
40	.17	12	6,7,7,7,8,8,9	32	26	40	21	29	25
			9,9,10,11,16	61	32	79	23	36	30
55	.48	20	17,23,24,25,25	71	69	71	43	67	57
			26,26,26,27,27	79	74	81	45	72	61
			28,28,28,29,30						
			30,30,31,33,38						
60	.24	15	11,11,12,12,13	67	49	67	24	44	37
			13,14,16,17,17	88	53	90	28	49	40
			18,18,20,20,22						

Table 2.26

The last two columns present the estimator $\hat{n}(\hat{p},t)$ where $\hat{p} = \overline{X}/\hat{n}$, $\hat{n}$ is the value of $\hat{n}_{L:s}$ and t = 1, 2

n	p	m	sample	$\hat{n}_m$	$\hat{n}_{m:s}$	$\hat{n}_L$	$\hat{n}_{L:s}$	$\hat{n}(\hat{p},t)$	
75	.32	5	16,18,22,25,27	102	70	99	29	30	29
			16,18,22,25,28*	195	80	190	30	32	31
34	.57	4	14,18,20,26	507	77	504	31	33	31
				< 0	91	∞	32	35	32
37	.17	20	4,4,4,4,5,5,5	65	25	66	11	12	12
			5,6,6,6,6,7,9	154	27	159	13	14	14
			9,10,10,10,11,11						
48	.06	15	0,1,1,2,2,2,3,3	18	10	15	7	7	7
			3,4,4,4,4,5,6	135	12	125	9	10	9
40	.17	12	6,7,7,7,8,8,9	32	26	40	21	24	22
			9,9,10,11,16	61	32	79	23	28	25
55	.48	20	17,23,24,25,25	71	69	71	43	47	44
			26,26,26,27,27	79	74	81	45	50	47
			28,28,28,29,30						
			30,30,31,33,38						
60	.24	15	11,11,12,12,13	67	49	67	24	26	25
			13,14,16,17,17	88	53	90	28	30	28
			18,18,20,20,22						

Table 2.27

The last two columns present the estimator $\hat{n}(\hat{p},t)$

where $\hat{p} = \overline{X}/\max(X_1,\ldots,X_m)$, $t = 1,\ 2$

n	p	m	sample	$\hat{n}_m$	$\hat{n}_{m:s}$	$\hat{n}_L$	$\hat{n}_{L:s}$	$\hat{n}(\hat{p},t)$	
75	.32	5	16,18,22,25,27 16,28,22,25,28*	102 195	70 80	99 190	29 30	29 31	28 30
34	.57	4	14,18,20,26	507 < 0	77 91	504 ∞	31 32	29 31	28 30
37	.17	20	4,4,4,4,5,5,5 5,6,6,6,6,7,9 9,10,10,10,11 11	65 154	25 27	66 159	11 13	12 14	12 13
48	.06	15	0,1,1,2,2,2,3 3,3,4,4,4,5,6	18 135	12 12	15 125	7 9	6 8	6 8
40	.17	12	6,7,7,7,8,8,9 6,6,10,11,16	32 61	26 32	40 79	21 23	20 22	19 21
55	.48	20	17,23,24,25,25 26,26,26,27,27 28,28,28,29,30 30,30,31,33,38	71 79	69 74	71 81	43 45	43 45	42 43
60	.24	15	11,11,12,12,13 13,14,16,17,17 18,18,20,20,22	67 88	49 53	67 90	24 28	24 26	24 25

Table 2.28

The last two columns present the estimator $\hat{n}(\hat{p},t)$

where $\hat{p} = 1/2$ and $t = 1,\ 2$

n	p	m	sample	$\hat{n}_m$	$\hat{n}_{m:s}$	$\hat{n}_L$	$\hat{n}_{L:s}$	$\hat{n}(\hat{p},t)$	
75	.32	5	16,18,22,25,27	102	70	99	29	36	41
			16,18,22,25,28*	195	80	190	30	37	42
34	.57	4	14,18,20,26	507	77	504	31	35	39
				< 0	91	∞	32	36	41
37	.17	20	4,4,4,4,5,5,5	65	25	66	11	13	14
			5,6,6,6,6,7,9	154	27	159	13	14	15
			9,10,10,,10,11,11						
48	.06	15	0,1,1,2,2,2,3	18	10	15	7	6	6
			3,3,3,4,4,4,4	135	12	125	9	7	7
			5,6						
40	.17	12	6,7,7,7,8,8,9	32	26	40	21	20	21
			9,9,10,11,16	61	32	79	23	21	23
55	.48	20	17,23,24,25,25	71	69	71	34	50	56
			26,26,26,27,27	79	74	81	45	52	58
			28,28,28,29,30						
			30,30,31,33,38						
60	.24	15	11,11,12,12,13	67	49	67	24	28	31
			13,14,16,17,17	88	53	90	28	30	32
			18,18,20,20,22						

CHAPTER THREE

BAYESIAN ESTIMATION OF THE BINOMIAL PARAMETER

3.1. Introduction

In this chapter, we will consider the Bayesian approach for estimating the parameter n. We will take two types of prior, non-informative prior and Poisson prior of n and as we saw in chapter one these two types of priors have been considered by Hamedani and Walter (1988). Assuming p is known and using quadratic loss function, the mean of the posterior distribution of n is the Bayes estimator. The Bayes estimator does not possess a closed form. We used a simulation method for obtaining the mean square error and we compare it with the mean square error of the MGF based estimation $\hat{n}_{m,t}$. For unknown p, we assume that n and p are independent and we take a beta prior distribution for p. Also, the Bayes estimator does not possess a closed form.

3.2. Bayes Estimators

Let $\underline{X} = (X_1, \ldots, X_m)$ be a random sample taken from binomial distribution b(n,p), where n is the parameter of interest, $n \in \{1, 2, \ldots\}$. Then, the likelihood function is

$$L(n,p|\underline{x}) = p^{m\bar{x}}(1-p)^{m(n-\bar{x})} \prod_{i=1}^{m} \binom{n}{x_i} \qquad (3.2.1)$$

where $$\bar{\alpha} = \frac{1}{m} \sum_{i=1}^{m} \alpha_i$$

now, we discuss the cases p known and p unknown separately.

1. p is known. Let g(n) be a prior distribution for n. One sensible form for g(n) is

$$g_1(n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad n = 0, 1, \ldots$$

another prior distribution of n (improper) is $g_2(n) = 1$ for all n.

The posterior distribution for n with respect to $g_1(n)$ is

$$\pi_1(n|\underline{\alpha},p) = \frac{\left[\prod_{i=1}^{m} \binom{n}{\alpha_i}\right] (1-p)^{mn} \frac{\lambda^n}{n!}}{\sum_{n=\alpha_{(m)}}^{\infty} \left[\prod_{i=1}^{m} \binom{n}{\alpha_i}\right] (1-p)^{mn} \frac{\lambda^n}{n!}}, \quad n \geq \alpha(m) \qquad (3.2.2)$$

The posterior distribution for n with respect to $g_2(n)$ is

$$\pi_2(n|\underline{\alpha},p) = \frac{\left(\prod_{i=1}^{m} \binom{n}{\alpha_i}\right) (1-p)^{mn}}{\sum_{n=\alpha_{(m)}}^{\infty} \left(\prod_{i=1}^{m} \binom{n}{\alpha_i}\right) (1-p)^{mn}}, \quad n \geq \alpha(m) \qquad (3.2.3)$$

The Bayes estimator of n w.r.t squared error loss function is given by the mean of the posterior. Thus

$$\hat{n}_{B_1} = \frac{\sum_{n=X_{(m)}}^{\infty} n \left[\prod_{i=1}^{m} \binom{n}{X_i}\right] (1-p)^{mn} \frac{\lambda^n}{n!}}{\sum_{n=X_{(m)}}^{\infty} \left[\prod_{i=1}^{m} \binom{n}{X_i}\right] (1-p)^{mn} \frac{\lambda^n}{n!}} \tag{3.2.4}$$

is the Bayes estimator w.r.t $g_1(n)$

$$\hat{n}_{B_2} = \frac{\sum_{n=X_{(m)}}^{\infty} n \left[\prod_{i=1}^{m} \binom{n}{X_i}\right] (1-p)^{mn}}{\sum_{n=X_{(m)}}^{\infty} \left[\prod_{i=1}^{m} \binom{n}{X_i}\right] (1-p)^{mn}} \tag{3.2.5}$$

is the Bayes estimator w.r.t to $g_2(n)$.

These estimators does not possess a closed form expect in the univariate case. For m = 1 (3.2.4) reduced to $\hat{n}_{B_1} = X+\lambda q$ and (3.2.5) reduced to $\hat{n}_{B_2} = \frac{X}{p} + \frac{q}{p}$.

Hamodani and Walter (1988) considered these estimators just in the univariate case and they apply these formulae to some examples of Draper and Guttman (1971). In this work, we will deal with these estimators in the multivariate case.

2. p is unknown. We assume that n and p are independent. Let $g(n)$, $h(p)$ be the priors. One sensible for $h(p)$ is the beta function $h(p) \propto p^{\alpha}(1-p)^{\beta}$, $0 < p < 1$.

The joint posterior is

$$\Pi(n,p|\underline{x}) = \frac{\left(\prod_{i=1}^{m}\binom{n}{x_i}\right) p^{m\bar{x}+\alpha}\,(1-p)^{m(n-\bar{x})+\beta} g(n)}{\sum_{n=x_{(m)}}^{\infty}\left(\prod_{i=1}^{m}\binom{n}{x_i}\right) g(n)\int_0^1 p^{m\bar{x}+\alpha}(1-p)^{m(n-\bar{x})+\beta}\,dp} \qquad (3.2.6)$$

we can integrate p out from (3.2.6) to get the marginal distribution for n which is

$$\Pi(n|\underline{x}) = \frac{\left(\prod_{i=1}^{m}\binom{n}{x_i}\right) g(n)\,\dfrac{\Gamma(m(n-\bar{x})+\beta+1)}{\Gamma(mn+\alpha+\beta+2)}}{\sum_{n=x_{(m)}}^{\infty}\left(\prod_{i=1}^{m}\binom{n}{x_i}\right) g(n)\,\dfrac{\Gamma(m(n-\bar{x})+\beta+1)}{\Gamma(mn+\alpha+\beta+2)}} \qquad (3.2.7)$$

either with respect to $g_1(n)$ prior or $g_2(n)$ the resulting estimator of n does not appear to have simple closed form. For the case p unknown, our main concern is the stability of the Bayes estimators. We try to analyze the example listed in Table (2) of Olkin, Pethan and Zidek

(1981), but we could not obtain the Bayes estimators for this example. The difficulty of obtain numerical values arises from the gamma function, which is a part of the Bayes estimator.

3.3. Numerical comparisons

In this section, we make comparisons (as in sec. 2.5) between the Bayes estimators ($\hat{n}_{B_1}$, $\hat{n}_{B_2}$) and the estimator based on the (MGF) $\hat{n}_{m,t}$ for small m, n, 10000, b(n,p), samples of a given m were simulated for m,n = 3, 6, 9, t = 0.05, p = 0.25, 0.5, 0.75 and λ = 3, 6, 9, We describe the method of simulation in the following steps.

1. For a given n and p, generate a random sample of size m from b(n,p),
2. Order the sample obtained in step 1,
3. Sbstitute the sample obtained in step (2) and in (3.2.5) in (3.2.4) to get $\hat{n}_{B_1}$, $\hat{n}_{B_2}$,
4. Repeat stesp (1, 2, 3) 10000 times to get $\hat{n}_{B_1}(i)$, i=1,2, ..., 10000,
5. Approximate the MSE by: $MSE(\hat{n}_{B_1}) = \sum_{i=1}^{10000} \left(\hat{n}_{B_1}(i)-n\right)^2 / 10000$,

and $MSE(\hat{n}_{B_2}) = \sum_{i=1}^{10000} \left(\hat{n}_{B_2}(i)-n\right)^2 / 10000$

6. Approximate the bias $\text{Bias}(\hat{n}_{B_1}) = \sum_{i=1}^{10000} \left(\hat{n}_{B_1}(i) / 10000\right) - n$,

$$\text{Bias}(\hat{n}_{B_2}) = \sum_{i=1}^{10000} \left(\hat{n}_{B_2}(i) / 10000\right) - n$$

We report the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$ in (3.1) - (3.3). Tables (3.4) - (3.6) presents the absolute ratio of bias of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$. Table (3.7) present the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_2}$. Table (3.8) present the absoluet ratio of bias of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_2}$.

Table 3.1

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$, $\lambda = 3$, $t = 0.05$

		n		
m	p	3	6	9
3	.25	0.267958	0.605257	1.36533
	.5	0.549462	0.767981	1.43903
	.75	0.548113	0.760383	1.10750
6	.25	0.439146	0.716706	1.51377
	.5	0.652585	0.830507	1.36353
	.75	0.249351	0.693017	0.96844
9	.25	0.567716	0.777172	1.52195
	.5	0.656713	0.812232	1.27652
	.75	0.083329	0.55744	0.91628

Table 3.2

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$, $\lambda = 6$, $t = 0.05$

		n		
m	p	3	6	9
3	.25	1.06687	0.258719	0.486968
	.5	1.149296	0.565138	0.696598
	.75	0.780394	0.734798	0.782607
6	.25	1.15251	0.444966	0.585179
	.5	1.02189	0.732105	0.803247
	.75	0.346131	0.680863	0.781974
9	.25	1.15667	0.571996	0.66742
	.5	0.91498	0.744915	0.817796
	.75	0.101766	0.565158	0.780233

Table 3.3

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$, $\lambda = 9$, $t = 0.05$

		n		
m	p	3	6	9
3	.25	3.12411	0.639484	0.223289
	.5	2.29087	0.890412	0.543031
	.75	1.16501	0.853714	0.725294
6	.25	2.51021	0.84304	0.399356
	.5	1.57069	0.963691	0.786296
	.75	0.47278	0.749334	0.783296
9	.25	2.23462	0.922563	0.536446
	.5	1.26063	0.933045	0.802864
	.75	0.14436	0.622487	0.765647

Table 3.4

The absolute ratio of bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_{B_1}$, $\lambda = 3$, t=0.05

		n		
m	p	3	6	9
3	.25	4.8933	18.528	40.1027
	.5	3.9304	40.5055	44.381
	.75	4.1651	33.477	66.6125
6	.25	4.9293	30.47	49.5713
	.5	10.282	16.608	38.6005
	.75	5.8522	48.569	21.6123
9	.25	86.8723	85.6492	38.1049
	.5	11.1704	28.279	58.15
	.75	2.9487	15.4958	17.7948

Table 3.5

The absolute ratio of bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_{B_1}$, $\lambda = 6$, t=0.05

m	p	n = 3	n = 6	n = 9
3	.25	59.3265	5.56163	18.9369
	.5	47.8269	17.1085	19.2142
	.75	21.0438	5.2727	27.326
6	.25	61.934	2.0557	21.1293
	.5	64.1745	4.324	65.6269
	.75	15.5992	2.485	7.618
9	.25	14.5117	27.8713	34.9198
	.5	42.0259	9.161	21.2292
	.75	7.308	12.7796	6.5129

Table 3.6

The absolute ratio of bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_{B_1}$, $\lambda = 9$, t=0.05

		n		
m	p	3	6	9
	.25	36.396	17.683	1.23827
3	.5	85.295	22.7252	3.98491
	.75	35.3531	36.5753	2.03318
	.25	32.4101	29.444	0.37052
6	.5	40.2839	36.56	6.19496
	.75	24.0386	11.5171	2.30568
	.25	17.4806	4.20938	1.32262
9	.5	78.5872	75.29	2.97239
	.75	10.86	37.0875	2.07856

Table 3.7

The efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_2}$, $t = 0.05$

m	p	n = 3	n = 6	n = 9
3	.25	1.32724	0.8545	0.478984
	.5	1.06067	1.01654	0.821873
	.75	0.709384	0.87076	0.899851
6	.25	1.13169	0.9903	0.69445
	.5	0.92268	0.99392	0.956701
	.75	0.30215	0.74015	0.862086
9	.25	1.06929	1.03728	0.851327
	.5	0.84010	0.93789	0.98319
	.75	0.08195	0.61911	0.848838

Table 3.8

The absolute ratio of bias of $\hat{n}_{m,t}$ w.r.t $\hat{n}_{B_2}$, t = 0.05

		n		
m	p	3	6	9
3	.25	23.5711	9.039	0.2573
	.5	23.6299	10.4853	5.4142
	.75	9.1713	14.474	11.0105
6	.25	32.823	14.077	4.3366
	.5	31.2699	14.81	25.8636
	.75	8.9613	4.412	4.0762
9	.25	30.3785	13.1098	10.2534
	.5	21.2744	26.17	8.4853
	.75	4.5209	19.918	8.0576

3.4 Results of numerical comparisons

1. Tables 3.1-3.3, presents the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$ for t = 0.05, λ = 3, 6, 9.

   a. For λ = 3, the tabulated results suggest that high efficiency (close to 1) is achieved for values of n = 9 and p = 0.25, 0.5, 0.75.

   b. For λ = 6, high efficiency (close to 1) is achieved for value of n = 3 and p = 0.25, 0.5, 0.75.

   c. For λ = 9, high efficiency (close to 1) is achieved for value of n = 3 and p = 0.25, 0.5, 0.75.

2. Tables 3.4-3.6, presents the absolute ratio of the bias of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_1}$. The tabulated results suggest that, the ratio of the absolute of the bias is always greater than 1 for λ = 3, 6, 9 and p = 0.25, 0.5, 0.75.

3. Tables 3.7, present the efficiency of $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_2}$. High efficiency (close to 1) is achieved for values of n = 3, 6 and p = 0.25, 0.5.

4. Table 3.8, present the absolute ratio of of the bias $\hat{n}_{m,t}$ with respect to $\hat{n}_{B_2}$. The tabulated results suggest that, the ratio of the absolute bias is always greater than 1.

CHAPTER FOUR

CONCLUDING REMARKS AND RECOMMENDATIONS

4.1 Concluding remarks

1. Some members of the family of estimators indexed by t can be used instead of the MLE or MME, in particular, those correspond to small values of t.
2. The estimators underestimate n for positive values of t and overestimate n for negative values of t. So, the negative values of t are recommended when the penalty for underestimation is more than that of overestimation.
3. The estimators (although being stable) are not doing well when p is unknown. This is also applied to the MME and MLE.

4.2 Recommendations

1. Further work should be done on Bayesian estimators of n. For example one may use other types of priors such as negative binomial distribution.
2. Moment generating function approach can be used in some similar problems such as negative binomial.
3. Other methods should be searched for in the case of unknown n and p.

References

[1] Blumenthal, S., and Dahiay, R. C. (1981). "Estimating the n Binomial Parameter n". J. Amer. Statist. Assoc. 76, 903-909.

[2] Carroll, R. J., and Lombard, F. (1985). "A Note on n Estimators for the Binomial Distribution". J. Amer. Statist. Assoc. 80, 423-426.

[3] Casella, G. (1986). "Stabilizing Binomial n Estimators". J. Amer. Statist. Assoc. 81, 172-175.

[4] Draper, N., and Guttman, I. 1971). "Bayesian Estimation of the Binomial Parameter". Technometrics, 13, 667-673.

[5] Gunel, E. and Chilko, D. (1989). "Estimation of Parameter n of the Binomial Distribution". Comm. Statist. Simula., 18(2), 537-551.

[6] Feldman, D. and Fox, M. 1968). "Estimation of the Parameter n in the Binomial Distribution". J. Amer. Statist. Assoc. 63, 150-158.

[7] Ghosh, M. and Meeden, G. (1975). "How Many Tosses of the Coin?" Sankhyā, Ser. A 37, 523-529.

[8] Hamedani, G. G. and Walter, G. G., (1988). "Bayes Estimation of the Binomial Parameter n". Comm. Statist. Theory Math., 17(6), 1829-1843.

[9] Kahn, W. D. (1987). "A Cautionary Note for Bayesian Estimation of the Binomial Parameter n". Amer. Statist. 41, 38-41.

[10] Olkin, I., Petkau, A. J., and Zidek, J. V. (1981). "A Comparison of n Estimators for the Binomial Distribution". J. Amer. Statist. Assoc., 76, 637-642.

[11] Quandt, R. E. and J. B. Ramsey (1978). "Estimating Mixtures of Normal Distribution and Switching Regressions". J. Amer. Statist. Assoc., 73, 730-738.

[12] Rao, C. R. (1973). Linear statistical inference and its applications. Second edition, John Wiley, New York.

[13] Rukhin, A. L. (1975). "Statistical Decision about the Total Number of Obserable Objects". Sankhyā, Ser. A, 37, 514-522.

[14] Sadooghi-Alvandi, S. M. (1986). "Admissible Estimation of the Binomial Parameter n". Ann. Statist., 14, 1634-1641.

[15] Sadooghi-Alvandi, S. M. (1992). "Estimation of the Binomial Parameter n Using A Linex Loss Function". Comm. Statist. Theory Meth., 21(5), 1427-1439.

[16] S. K. Bar-lev, N. Barkan and N. A. Langberg (1993). "Moment Generating Function Based Estimators With Some Optimal Properties". Journal of Statistical Planning and Inference 35, 279-291.

[17] Titterington, D. M., A. F. M. Smith and U. E. Makov (1985). Statistical analysis of finite mixture distributions. John Wiley, New York.

Program 1

```
C-------------------------------------------------------------
C       THIS PROGRAM IS TO CALCULATE THE MEAN SQUAR ERRORE
C         AND THE BIAS FOR THE BAYES ESTIMATOR (n^B2 )
C-------------------------------------------------------------
        INTEGER IR(10000),IOPT,N,NR,RA(10000),WW,MAX,INF,I1,J
       &,I2,II,I6
        REAL P,T, Q,S,PRO,S5,NP,GAMMA,I,EXACT,S2,S10,S11,BIA
        EXTERNAL   RNBIN, RNOPG, SVIGN, BINOM
        PRINT*,' WITHOUT POISSION .... '
        ITER   = 10000
        INF    = 14
        IOPT   = 6
        DO 3000 N= 3,9,3
        PRINT*,'N   = ',N
        DO 4000 NR = 3,9,3
        PRINT*,'M   = ',NR
        DO 5000 P  = .25,.75,.25
        PRINT*,'P   = ',P
        Q       = 1.- P
        S11     = 0.
        S10     = 0.
        DO 1000 II=1,ITER
        CALL RNOPG (IOPT)
        CALL RNBIN (NR, N, P, IR)
        S5    = 0.
        DO 3 J=1,NR
3       S5    =   S5  + IR(J)/FLOAT(NR)
        CALL SVIGN (NR, IR, RA)
        MAX   =   RA(NR)
        S     =   0.
        S2    =   0.
        DO 1 I = MAX,INF
        PRO   = 1.
        DO 2 I1 = 1,NR
        I6    = I
2       PRO   = PRO * BINOM(I6, IR(I1))
        S =S+PRO*(P**(S5*NR))*Q**(NR*(I-S5))
1       S2=S2+PRO*P**(S5*NR)*Q**(NR*(I-S5))*I
        NP = S2/S
        S10 = S10  + (NP - N)**2  / FLOAT(ITER)
        S11 = S11  + NP / FLOAT(ITER)
1000    CONTINUE
        BAI    =   (S11 - N)
        PRINT*,'MSE = ',S10
        PRINT*,'BIAS= ',BAI
5000    CONTINUE
        PRINT*,'==--=== -   END OF P  ====== '
4000    CONTINUE
3000    CONTINUE
        END
```

Program 2

```
C--------------------------------------------------------------------
C      THIS PROGRAM IS TO CALCULATE THE MEAN SQUAR ERRORE
C
C         AND THE BIAS FOR THE BAYES ESTIMATOR (n̂B1 )
C--------------------------------------------------------------------
       INTEGER    IR(15),IOPT, N,NR, RA(15),WW,MAX,INF,I1,J
      &,I2,II,I6
       REAL P,T,Q,S,PRO,S5,L,NP,GAMMA,G,I,EXACT,S2,S10,S11,BIA
       EXTERNAL   RNBIN, RNOPG, SVIGN, GAMMA, BINOM
       PRINT*,'WITH POSSION ....'
       ITER   = 10000
       INF    = 14
       IOPT   = 6
       DO 2000 N = 3,9,3
       PRINT*,'N   = ',N
       DO 3000 NR = 3,9,3
       PRINT*,'M   = ',NR
       DO 4000 P   = .25,.75,.25
       PRINT*,'P   = ',P
       PRINT*,'+++++++++++++++++++++++++++++++++++++++++++++'
       DO 5000 L = 3.,9.,3.
       PRINT*,'$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$'
       PRINT*,'LAMDA   = ',L
       PRINT*,'$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$'
       Q       = 1.- P
       S11     = 0.
       S10     = 0.
       DO 1000 II=1,ITER
       CALL RNOPG (IOPT)
       CALL RNBIN (NR, N, P, IR)
       S5     = 0.
       DO 3 J=1,NR
 3     S5     =   S5   + IR(J)/FLOAT(NR)
       CALL SVIGN (NR, IR, RA)
       MAX    =   RA(NR)
       S      =   0.
       S2     =   0.
       DO 1 I = MAX,INF
       PRO    = 1.
       DO 2 I1 = 1,NR
       I6     = I
 2     PRO    = PRO * BINOM(I6, IR(I1))
       G      =   GAMMA(I + 1)
       S =S+PRO*(P**(S5*NR))*Q**(NR*(I-S5))*(L**I/G)
 1     S2=S2+PRO*I*P**(S5*NR)*Q**(NR*(I-S5))*(L**I/G)
       NP = S2/S
       S10 = S10  + (NP - N)**2  / FLOAT(ITER)
       S11 = S11  + NP   / FLOAT(ITER)
 1000  CONTINUE
       BAI     =  (S11 - N)
       PRINT*,'MSE = ',S10
```

```
      PRINT*,'BIAS= ',BAI
5000  CONTINUE
      PRINT*,'===-----===   END OF LAMDA ===-------==='
      PRINT*,'                                        '
      PRINT*,'                                        '
4000  CONTINUE
3000  CONTINUE
2000  CONTINUE
      END
```

Program 3

```
C-------------------------------------------------------------
C      THIS PROGRAM IS TO CALCULATE  THE MEAN SQUAR ERROR
C
C    AND THE BIAS FOR THE MAXIUMUM LIKELIHOOD ESTIMATOR (n^l)
C
C   AND THE MOMENT GENARATING FUNCTION BASD ESTIMATORF(n^m,t)
C-------------------------------------------------------------
      INTEGER IR(10000),IOPT,N,NR,RA(1000),L,U,RAN,AA,ZZ,KK,
     &,K1,ITER,VV,MM,THETA
      REAL S2,T,P, Q, S,B, LIK(1000),RAA(1000),OLIK(1000),SS,
     &BIAS,BBB,BBB2,BIAS2,EFF,ESTM,EST2,EESS,
      EXTERNAL   RNBIN, RNOPG, SVIGN
      ITER   = 10000
      P      = 0.5
      Q      = 1.- P
      DO 1000 NR   = 3, 15, 3
 PRINT*,'M     = ', NR
      PRINT*,' MSE(L)  MSE(N^T)  BIAS(T)  BIAS(L)  EFF(L/NT)'
      DO 1000 N = 3, 15, 3
 PRINT*,'                 N = ', N
      DO 2000 T = .1,.1
 PRINT*,'                 T = ',T
      PRINT*,'--------------------------------------------'
      SS     = 0.
      EESS   = 0.
      BBB    = 0.
      BBB2   = 0
      DO 200 WW=1,ITER
      CALL RNOPG (IOPT)
      CALL RNBIN (NR, N, P, IR)
      CALL SVIGN (NR, IR, RA)
      S    = 0.
      DO 1 I=1,NR
1     S  =   S + IR(I)
      S2 = 0.
      DO 4 VV=1,NR
4     S2 = S2 + EXP(IR(VV)*T)/FLOAT(NR)
      EST2 = ALOG(S2)/ALOG(Q+P*EXP(T))
      IF (RA(1) .EQ. 0 .AND. RA(2) .EQ. 0 .AND. RA(3) .EQ. 0
     &.AND. RA(4) .EQ. 0 .AND. RA(5) .EQ. 0 .AND. RA(6).EQ.0
     &.AND. RA(7) .EQ. 0 .AND. RA(8) .EQ. 0 .AND. RA(9).EQ.0
     &.AND. RA(10) .EQ. 0 .AND. RA(11) .EQ. 0 .AND. RA(12)
     &.EQ. 0 .AND. RA(13) .EQ. 0 .AND. RA(14) .EQ. 0
     &.AND. RA(15) .EQ. 0) THEN
           K1 = 1
                 GOTO 333
      ELSE
                 GOTO 35
      ENDIF
35    L   =  INT(RA(NR)/(1. - (Q**NR)))
      U   =  INT(S/(1.-(Q**NR)))
      DO 2 J=L,U
```

```
        LIK(J) = 1.
        DO 24 HH=1,NR
        LIK(J) = LIK(J)*( BINOM(J, IR(HH))*(P**IR(HH))*(Q**(J-IR(H
24      CONTINUE
2       CONTINUE
        DO 3 K=L,U
3       OLIK(K) = LIK(K)
        DO 9 AA=L,U-1
        DO 8 ZZ=AA+1,U
        IF (OLIK(AA) .LT. OLIK(ZZ)) GOTO 8
        QQ = OLIK (ZZ)
        OLIK(ZZ) = OLIK (AA)
        OLIK(AA) = QQ
8       CONTINUE
9       CONTINUE
        DO 22 KK=L,U
        IF (OLIK(U) .NE. LIK(KK)) GOTO 22
        K1 = KK
22      CONTINUE
333     BBB2 = BBB2 + K1/FLOAT(ITER)
        SS   = SS + (K1 - N)**2/FLOAT(ITER)
        EESS = EESS + (EST2 - N)**2/FLOAT(ITER)
200     BBB  = BBB  + EST2/FLOAT(ITER)
        BIAS = (BBB - N)
        BIAS2=(BBB2 - N)
        EFF  = SS/EESS
  PRINT*,'                                              '
        PRINT*,SS,EESS,BIAS,BIAS2,EFF
  PRINT*,'                                              '
2000    CONTINUE
        PRINT*,'++++++++++++++++++++++++++++++++++++++'
1000    CONTINUE
        END
```

Program 4

```
C-----------------------------------------------------------------
C      THIS PROGRAM IS TO CALCULATE THE MEAN SQUAR AND
C
C     THE BIAS FOR THE MAXIUMUM LIKLEIHOOD ESTIMATOR (n ) AND
C
C    THE MODIFIED MOMENT GENARATING FUNCTION BASD ESTIMATOR
C    (n^*m,t)
C-----------------------------------------------------------------
      INTEGER    IR(10000),IOPT, N,NR, RA(1000),L,U,RAN,AA
     &,K1,ITER,VV,MM,ZZ,KK,THETA
      REAL S2,T,P, Q, S,B, LIK(1000),RAA(1000),OLIK(1000),
     &BIAS,BBB,BBB2,BIAS2,EFF,ESTM,SS,EST2,EESS
      EXTERNAL   RNBIN, RNOPG, SVIGN
      PRINT*,'MODIFY........'
      DO 9000    P      = 0.5,.5
      PRINT*,'P = ',P
      Q      = 1.- P
      DO 3000 NR= 10,10
      PRINT*,'M =',NR
      DO 2000 N=10,10
      PRINT*,'N =',N
      DO 1000 T=-.1,-.05,.05
      PRINT*,'T =',T
      CALL RNOPG (IOPT)
      CALL RNBIN (NR, N, P, IR)
      CALL SVIGN (NR, IR, RA)
      PRINT*,'THE OBSE'
      PRINT*,(IR(AZ),AZ=1,NR)
      S    = 0.
      DO 1 I=1,NR
1     S   =  S + IR(I)
      S2 = 0.
      DO 4 VV=1,NR
4     S2 = S2 + EXP(IR(VV)*T)/FLOAT(NR)
      EST2 = ALOG(S2)/ALOG(Q+P*EXP(T))
      IF (EST2 .LE. 1.) THEN
      ESTM = 1.
      ELSE
         ESTM = INT(EST2+.5)
      ENDIF
      IF (RA(1) .EQ. 0 .AND. RA(2) .EQ. 0 .AND. RA(3) .EQ. 0
     &.AND. RA(4).EQ.0 .AND. RA(5) .EQ. 0 .AND. RA(6) .EQ. 0
     &.AND. RA(7).EQ.0 .AND. RA(8) .EQ. 0 .AND. RA(9) .EQ. 0
     &.AND. RA(10) .EQ. 0 ) THEN
          K1     =    1
       GOTO 333
      ELSE
                     GOTO 222
      ENDIF
222   L     =   INT(RA(NR)/(1. - (Q**NR)))
      U     =   INT(S/(1.-(Q**NR)))
      DO 2 J=L,U
```

```
      LIK(J) = 1.
      DO 24 HH=1,NR
      LIK(J) = LIK(J)*( BINOM(J, IR(HH))*(P**IR(HH))*(Q**
     &(J-IR(HH))))
24    CONTINUE
2     CONTINUE
      DO 3 K=L,U
3     OLIK(K) = LIK(K)
      DO 9 AA=L,U-1
      DO 8 ZZ=AA+1,U
      IF (OLIK(AA) .LT. OLIK(ZZ)) GOTO 8
      QQ = OLIK (ZZ)
      OLIK(ZZ) = OLIK (AA)
      OLIK(AA) = QQ
8     CONTINUE
9     CONTINUE
      DO 22 KK=L,U
      IF (OLIK(U) .NE. LIK(KK)) GOTO 22
      K1 = KK
22    CONTINUE
333   PRINT*,'MGFM =',ESTM
1000  CONTINUE
      PRINT*,'+++++++++++++++++++++++++++++++++++++++++++++++++++++++++'
2000  CONTINUE
3000  CONTINUE
      PRINT*,'--===---    END OF P - = - = '
9000  CONTINUE
      END
```

Program 5

```
C-----------------------------------------------------------------
C     THIS PROGRAM FOR THE ROOT OF EQUATION  (2.7.3) ...
C-----------------------------------------------------------------
      INTEGER    IR(1000),NR,RA(1000)
      REAL S2,T,P, Q, S,B,SS,EST2,VAR,SUM2,ASD6,
     &ESTM,T1,T2,S10,S11,ASD,F,X,ROO,SSS,ASD1,ASD2,ASD4,ESTT
      EXTERNAL   RNBIN, RNOPG, SVIGN
      PRINT*,'INPUT OBS. NN AND SAMPLE SIZE AND T1 T2 '
      READ*,NR,NN,T1,T2
      DO 13 I=1,NR
13    READ*,IR(I)
      S10  = 0.
      SSS  = 0.
      SUM2 = 0.
      CALL SVIGN (NR, IR, RA)
      DO 22 UU=1,NR
22    SSS = SSS + IR(UU)/FLOAT(NR)
      ASD6 = SSS/FLOAT(NN)
      DO 5 AS=1,NR
5     S10 = S10 + EXP(IR(AS)*T1)/FLOAT(NR)
      S11  = 0.
      DO 6 AS1=1,NR
6     S11 = S11 + EXP(IR(AS1)*T2)/FLOAT(NR)
      ASD4 = ALOG(S10)/ALOG(.5+.5*EXP(T1))
      ASD1 = ALOG(S11)/ALOG(.5+.5*EXP(T2))
      ASD3 = SSS/RA(NR)
      ESTT =  ALOG(S10)/ALOG((1.-ASD3)+ASD3*EXP(T1))
      ESTT2 =  ALOG(S11)/ALOG((1.-ASD3)+ASD3*EXP(T2))
      ASD = ALOG(S10)/ALOG(S11)
      ESTT4 =  ALOG(S10)/ALOG((1.-ASD6)+ASD6*EXP(T1))
      ESTT6 =  ALOG(S11)/ALOG((1.-ASD6)+ASD6*EXP(T2))
      DO 9 X=0.0001,1.,.0001
      F =( (1.-X)+X*EXP(T2) )**ASD- ( (1.-X)+X*EXP(T1) )
      IF (F .GT. .00001) GOTO 9
      ROO = X
      GOTO 444
9     CONTINUE
444   PRINT*,'R O O T = ',ROO
      EST2 = ALOG(S10)/ALOG((1.-ROO)+ROO*EXP(T1))
      IF (EST2 .LE. 1.) THEN
       ESTM = 1.
      ELSE
         ESTM = INT(EST2+.5)
      ENDIF
      PRINT*,'n(.5,2)n(.5,1)n(MAX,1) n(MAX,2)n(NN,1)n(NN,2)n
     &(ROOT,1)'
      PRINT*,'=========================================
     &======='
      PRINT*,INT(ASD1),INT(ASD4),INT(ESTT),INT(ESTT2)
     &,INT(ESTT4),INT(ESTT6),INT(ESTM)
      END
```