Alqalam Journal of Medical and Applied Sciences. 2025;8(1):315-318
https://doi.org/10.54361/ajmas.258145

Original article

Binomial Parameter Estimation: Asymptotic analysis of Normality and Confidence Regions Using Moments Estimators

Salma Saad[1], Kamilah Othman[2*], Naeima Abdelati[3]

[1]*Department of Statistics, Faculty of Science, University of Benghazi, Libya*
[2]*Department of Statistics, Faculty of Arts and Science, Al-Marj, University of Benghazi, Libya*
[3]*Department of Mathematics, Faculty of Science, Omar Al-Mukhtar University, Libya*
***Corresponding email**: kamlah.alabd@uob.edu.ly*

Abstract
Confidence regions for binomial distributions are widely used in practical research. Extensive application has interested many statisticians in improving new methods for constructing these regions in asymptotic and exact ways. This research primarily aims to construct confidence ellipses for parameters p and m of the binomial distribution *Bin(m,p)* using a sample of fixed size n in the method of moments estimators. Although the conventional method of moments estimators for p and m are well-established and straightforward to derive, their denominators can reach zero with positive probabilities. This study introduces modified estimators to effectively address this issue. By using the asymptotic normality of the joint distribution of both the conventional and modified estimators, we examine the quadratic forms of these estimators that follow a chi-square distribution, enabling the construction of simultaneous confidence regions (ellipses) for the parameters p and m. The method of moments estimators for binomial parameters p and m are found to be approximately normal, which allows for confidence regions to be constructed based on the chi-square distribution. Presenting modified estimators to handle zero-denominator issues ensures robust and reliable inferences. The theoretical side of the confidence region results show the effectiveness of binomial parameters p and m using modified estimators, solving the limitations of conventional approaches. This method provides a powerful framework for statistical inference, relying on asymptotic normality and a chi-square distribution. This method provides a powerful framework for statistical inference because it depends on asymptotic normality and the chi-square distribution.
Keywords. Binomial Distribution, Method of Moments, Confidence Region, Confidence Ellipse.

Introduction

Binomial distribution is one of the most essential discrete probability distributions in statistics and is defined as *Bin(m,p)*, where m denotes the trial number and p is the probability of success in each trial [1]. It is commonly applied in statistical processes, such as quality control, where it is used as a model of the number of detective objects in a group [2].

Binomial distribution of confidence regions is commonly utilized in practical research. Many statisticians have become interested in improving new methods for constructing these regions in both asymptotic and exact locations owing to their extensive use [3]. Some studies have used many discrete probability distributions to explore and understand the confidence regions and asymptotic properties of estimators with different methods for specific sample sizes. Nkingi and Vrbik studied the moments estimator's method and central limit theorem for building confidence regions for unknown parameters of the negative binomial distribution. Their new method creates two confidence regions for negative binomial and Poisson distributions [4]. Beal developed a new interval method for asymptotic confidence of small samples in the difference between the parameters of the binomial distribution [5]. In addition, Chafaï and Concordet proposed a new method of confidence regions for multinomial parameters with small sample sizes to compute any outcome number [6].

The efficiency of estimates is evaluated against the maximum likelihood for truncated binomial distributions using the asymptotic variance of moments estimates method for truncated binomial and negative binomial distributions by Shah [7]. Furthermore, Brown et al. compared the coverage characteristics and estimated the lengths of different confidence interval methods of binomial proportions. The study found that the Agresti-Coull interval has the longest expected length, whereas the other has a shorter interval method [8].

The main purpose of this research is to build the theoretical considerations of confidence regions for the binomial distribution parameters p and m by applying the moment estimator method. In particular, we address the problem of zero denominators by presenting adjusted estimators that remain definite for all different sample sizes. These estimators, denoted as $\tilde{p}_n$ and $\tilde{m}_n$, are designed to converge the proper parameters when increasing the sample size n, to avoid unreliability related to conventional estimators. Using the normal asymptotic of the joint distribution of the estimators, we obtain the quadratic forms that distribute a chi-square formula, allowing the structure of simultaneous confidence regions (ellipses) for p and m. This research primarily aims to construct confidence ellipses for parameters p and m of the binomial distribution *Bin(m,p)* using a sample of fixed size n in the method of moments estimators.

***Received**: 18-12-2024 - **Accepted**: 12-02-2025 - **Published**: 16-02-2025*

Methods

This section shows the moments estimation methods and confidence religions of binomial distribution as follows:

We consider the binomial distribution as *Bin(m, p)*, where m is the number of trials (a positive integer), whereas p is the probability of success of a single trial, 0 ≤ p ≤ 1. The probability mass function for a random variable Z, following a binomial distribution, is defined as:

$$P(Z = k) = \left\langle {m \atop k} \right\rangle p^k (1-p)^{m-k} \quad , k = 0, 1, 2, \ldots, m$$

The first statistical inference related to the binomial distribution was presented in 1941 by Haldane[9]. According to [9], the estimators of parameters can be evaluated using the method of moment estimation. We refer to [11] and [10] for a detailed discussion of the literature about the statistical inference for the binomial distribution. In this article, we give the confidence regions that provide a whole set of parameters that simultaneously may be true with the confidence$1-\alpha$.

Methods of Moments Estimation

$\bar{z} = \frac{1}{n}\sum_{k=1}^{n} Z_k$ and $S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(Z_k - \bar{z})^2$

Are the sample mean and sample variance. It is well known that the method of moments estimators of parameters p and m by the method of moments are:

$\hat{p}_n = \frac{\bar{Z}-S^2}{\bar{Z}}$ and $\hat{m}_n = \frac{\bar{Z}^2-S^2}{\bar{Z}}$

See for example [11] and (10). Note that the denominator of the statistic can obtain a zero value with positive probability. Because the mean values and the variances do not exist for the method of moment estimators $\hat{p}_n$ and $\hat{m}_n$, modified estimators are introduced as:

$\hat{p}_n = \frac{\bar{Z}-S^2}{\bar{Z}+\epsilon_n}$ and $\hat{m}_n = \frac{\bar{Z}^2-S^2}{\bar{Z}-S^2+\epsilon_n}$

Where $\epsilon_n \to 0$ as n → ∞. For the remainder of this work, the sequence $\epsilon_n = \frac{1}{n}$ will be used as it meets the convergence criteria.

With the help of the famous delta method, it was derived in [11] and [10] that the random vector $\sqrt{n}(\hat{p}_n - p, \hat{m}_n - m)^T$Converges in distribution is to the normal distribution with the mean vector $(0, 0)^T$ and the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_p^2 & p\sigma_p\sigma_m \\ p\sigma_p\sigma_m & \sigma_m^2 \end{pmatrix}$$

Where $\sigma_p^2 = \frac{(1-p)(\alpha+p)}{m}$, $\sigma_m^2 = \frac{m(1-p)\alpha}{p^2}$, $p = \sqrt{\frac{\alpha}{\alpha+p}}$ with $\alpha = 2(1-p)(m-1)$

Also, the $\sqrt{n}(\tilde{p}_n - p, \tilde{m}_n - m)^T$ has the same asymptotic behavior.

Confidence Region

The following result that connects the normal and chi-square distribution is well known, see [9] for example. Suppose $Y = (Y_1, Y_2, \ldots, Y_K)^T$ is a random vector that follows a normal distribution with the mean $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_K)^T$ and covariance matrix Σ. Then,

$\lambda = (Y-\lambda)^T \Sigma^{-1} (Y-\lambda)$ follows a chi-square distribution with n degrees of freedom.

In this article, we are dealing with the two-parameter binomial distribution. Therefore, the value of $k = 2$. This means that the mean vector $\lambda = (p-m)^T$ and the covariance matrix $\Sigma = \begin{pmatrix} \sigma_p^2 & p\sigma_p\sigma_m \\ p\sigma_p\sigma_m & \sigma_m^2 \end{pmatrix}$.

The inverse matrix Σ^{-1} can be evaluated as follows:

$$\Sigma^{-1} = \frac{1}{1-p^2} \begin{pmatrix} \frac{1}{\sigma_p^2} & \frac{-p}{\sigma_p\sigma_m} \\ \frac{-p}{\sigma_p\sigma_m} & \frac{1}{\sigma_m^2} \end{pmatrix} = \frac{1}{p(1-p)} \begin{pmatrix} m & -p \\ -p & \frac{p^2(p+a)}{ma} \end{pmatrix}$$

Now we present the formula required for evaluating the confidence regions for the parameters of the binomial distribution. The chi-square distribution with two degrees of freedom defines the statistics; we can state the confidence regions mathematically as follows:

$$\hat{Z}_2^2(p,m) = (Z-\lambda)^T \Sigma^{-1} (Z-\lambda)$$

Received: 18-12-2024 - **Accepted**: 12-02-2025 - **Published**: 16-02-2025

$$= \frac{1}{p(1-p)}(\hat{p}-p, \hat{m}-m) \begin{pmatrix} m & -p \\ -p & \frac{p^2(p+a)}{ma} \end{pmatrix} \begin{pmatrix} \hat{p}-p \\ \hat{m}-m \end{pmatrix}$$

$$= \frac{1}{p(1-p)} \left(m\,(\hat{p}-p)^2 - 2p(\hat{m}-m)(\hat{p}-p) + \frac{p^2(p+a)}{ma})(\hat{m}-m)^2\right)$$

In the case of the modified estimators, we can simply replace the estimators. $\hat{p}$ and $\hat{m}$ with the modified estimators $\tilde{p}$ and $\tilde{m}$.

Result and Discussion

The principal hypothetical result is the asymptotic normality of the moment estimator approach of $\hat{p}_n$ and $\hat{m}_n$ In particular, with a random vector $\sqrt{n}(\hat{p}_n - p, \hat{m}_n - m)^T$. Converge to a bivariate normal distribution with a covariance matrix Σ [12][11]. Additionally, adjusted estimators addressed zero denominator problems, and modified estimators $\tilde{p}_n$ and $\tilde{m}_n$ were presented, maintaining an approximately normal distribution and confirming robustness [10]. In addition, confidence regions with the chi-square distribution and confidence ellipses for two parameters p and m are derived through precise expressions for the inverse covariance matrix Σ^{-1} assisting the applied computation. Similarly, hypothetical contributions expand the work of [9], proposing a strong background for estimating and understanding binomial parameters.

Conclusion

An asymptotic analysis provides a challenging foundation for constructing regions for binomial distribution parameters p and m. The procedure of modified estimators addresses a significant limitation of the conventional method of moments estimators, however, is that the asymptotic normality and chi-square distribution theoretical findings confirm that the confidence regions are effective and reliable. These theoretical results expand the existing literature on the estimation of binomial parameters and offer a strong structure for statistical inference in applications in which conventional methods may be unpredictable.

Conflict of interest. Nil

References

1. America C, Walker S, Spohn C, Delone M, Meier H. Copyright 2009 Cengage Learning, Inc. All Rights Reserved. 2018. 333–337 p.
2. Douglas C.Montgomery. Introduction to Statistical Quality Control. Seventh Ed.
3. Tebbs JM, Roths SA. New large-sample confidence intervals for a linear combination of binomial proportions. 2008;138:1884–93.
4. Nkingi E, Vrbik J. Confidence Regions for Parameters of Negative Binomial Distribution. 2016;1–16. Available from: http://arxiv.org/abs/1612.08113
5. Beal SL. Asymptotic Confidence Intervals for the Difference between Two Binomial Parameters for Use with Small Samples. Biometrics. 1987;43(4):941.
6. Chafaï D, Concordet D. Confidence regions for the multinomial parameter with small sample size. J Am Stat Assoc. 2009;104(487):1071–9.
7. Shah SM. The Asymptotic Variances of Method of Moments Estimates of the Parameters of the Truncated Binomial and Negative Binomial Distributions. J Am Stat Assoc. 1961;56(296):990–4.
8. Lawrence D. Brown, T. Cai AD. Confidence intervals for a binomial proportion and asymptotic expansions 1. Ann Stat. 2002;30(1):160–201.
9. J. B. S. Haldane. The fitting of binomial distributions. Ann Eugen. 1941;11(1):179–81.
10. Saad S, Hossain S, Volodin A. Asymptotic analysis of the method of moments estimators of both parameters for the binomial distribution: Theoretical part. Thail Stat. 2021;19(1):43–58.
11. S. Saad, Asymptotic analysis of method of moment estimators of parameters p and m for the binomial distribution. (PhD Thesis, University of Regina, 2019).
12. Holmes SP, Lehmann EL. Elements of Large-Sample Theory. Vol. 95, Journal of the American Statistical Association. 2000. 328 p.

***Received**: 18-12-2024 - **Accepted**: 12-02-2025 - **Published**: 16-02-2025*

المستخلص

تُستخدم مناطق الثقة للتوزيعات ثنائية الحدين على نطاق واسع في البحث العملي. وقد أثار التطبيق الواسع اهتمام العديد من الإحصائيين بتحسين الأساليب الجديدة لبناء هذه المناطق بطرق مقاربة ودقيقة. يهدف هذا البحث في المقام الأول إلى بناء نقاط بيضاوية للثقة للمعلمات p و m للتوزيع ثنائي الحدين Bin(m,p) باستخدام عينة ذات حجم ثابت n في طريقة تقديرات العزوم. على الرغم من أن الطريقة التقليدية لتقديرات العزوم لـ p و m راسخة ومباشرة في الاستنباط، إلا أن مقاماتها يمكن أن تصل إلى الصفر باحتمالات موجبة. تقدم هذه الدراسة تقديرات معدلة لمعالجة هذه المشكلة بشكل فعال. من خلال استخدام التوزيع الطبيعي المقارب للتوزيع المشترك لكل من التقديرات التقليدية والمعدلة، نفحص الأشكال التربيعية لهذه التقديرات التي تتبع توزيع مربع كاي، مما يتيح بناء مناطق ثقة متزامنة (نقاط بيضاوية) للمعلمات p و m. لقد تبين أن طريقة تقدير اللحظات للمعاملات الثنائية p و m طبيعية تقريبًا، مما يسمح بإنشاء مناطق ثقة بناءً على توزيع مربع كاي. إن تقديم مقدرين معدلين للتعامل مع مشكلات المقام الصفري يضمن استدلالات قوية وموثوقة. يُظهر الجانب النظري لنتائج منطقة الثقة فعالية المعاملين الثنائيين p و m باستخدام مقدرين معدلين، وحل قيود الأساليب التقليدية. توفر هذه الطريقة إطارًا قويًا للاستدلال الإحصائي، بالاعتماد على التوزيع الطبيعي المقارب وتوزيع مربع كاي. توفر هذه الطريقة إطارًا قويًا للاستدلال الإحصائي لأنها تعتمد على التوزيع الطبيعي المقارب وتوزيع مربع كاي.

***Received**: 18-12-2024 - **Accepted**: 12-02-2025 - **Published**: 16-02-2025*