Analysis of a Two-Stage Adaptive Negative Binomial Group Testing Model for Estimating Prevalence of a Rare Trait

Jackline Akomboh, Wanyonyi Ronald Waliaula, Cox Tamba, Justine Obwoge Okenye

Department of Mathematics, Egerton University, Egerton, Kenya
Email: akombojacky@gmail.com, ronaldwaliaula@gmail.com, cox.tamba@egerton.ac.ke, justinoke69@gmail.com

Abstract

Group testing is an efficient method for classifying observations and estimating trait prevalence in a population. However, using appropriate group sizes is crucial for maximizing its benefits. Adaptive schemes have been developed to address improper group size selection issues. Existing adaptive schemes are based on a Binomial sampling model, requiring testing of all groups before recording successes. In certain scenarios, like infectious diseases, immediate reporting of estimates upon detection is necessary. A two-stage adaptive Negative Binomial group testing model for such cases was constructed. This adaptive model adjusted group sizes based on estimates from previous stages thus using optimal sizes to minimize the mean squared error and variance of the prevalence rate estimate. The maximum likelihood estimation method was employed to find the model's parameter estimate, and its properties were also investigated. The comparative analysis highlighted the superiority of the adaptive model over the non-adaptive model especially under low prevalence emphasizing the importance of incorporating adaptivity in group testing procedures, particularly in disease screening and surveillance, such as for COVID-19.

Subject Areas

Applied Statistical Mathematics

Keywords

Negative Binomial, Group Testing, Maximum Likelihood Estimation, Prevalence

1. Introduction

Group testing, also known as pool testing or batch testing, is a method that in-

volves combining individuals into various pools and conducting tests on these pools simultaneously. The tests are used to detect the presence of infections or defects, as seen in epidemiological studies. The idea was first introduced by Dorfman in 1943 to improve cost-saving techniques in detecting soldiers with syphilis [1]. Since Dorfman's pioneering work, group testing has been applied in various fields, including epidemiology, quality control, and genetics, with two main objectives including classification [2] [3] and estimation of prevalence [4]. The classification or identification of individuals as either positive or negative of a trait serves as the first objective of Group testing championed by [1]. To enhance cost-effectiveness and reduce the required tests, an adaptation of the Dorfman testing scheme has been investigated and extended to incorporate multi-stage testing [5]. Apart from classification problems, group testing has also been used extensively in the estimation problem of the prevalence rate of a trait. This is the second objective of group testing pioneered by [6] and served as the main focus of this study. In his study, he used the maximum likelihood estimation method, which was found to be reliable when the population was small [7]. Extended the estimation work using the MLE method and incorporating testing errors to account for real practical situations where errors are bound to happen during testing procedure.

Subsequent studies on estimation focused on design matters especially based on selection of group sizes used in group testing procedures [8]. The main aim of putting more emphasis on selection of group sizes was to reduce the chances of obtaining all negative or all positive groups during testing. Another aim was to minimize MSE by incorporating prior information in choosing k [9]. All these studies were carried out using the Binomial sampling model [10]. Later suggested the use of Negative Binomial sampling for estimating efficient p when the prevalence rate is small. Point and interval estimation has been explored under the negative binomial model using equal group sizes and has been found to be efficient in surveillance cases where quick response is desired [11]. Recent works that have considered inverse binomial models include estimation using the Bayesian approach as well as confidence interval estimation [12].

Group testing can be categorized into two forms: non-adaptive and adaptive group testing schemes. Non-adaptive group testing involves testing groups of a fixed size to obtain dichotomous results. On the other hand, adaptive group testing adjusts the group sizes from one stage to the next, allowing for more flexibility and efficiency [13]. Adaptive estimators have been developed to improve efficiency by reducing the Mean Squared Error under the binomial model [9]. Other probabilistic models, such as the beta-binomial, geometric, and hyper-geometric models, have also been considered [14]. For urgent situations where the prevalence rate of a trait needs to be estimated quickly, the Negative Binomial sampling model has been found to be preferable over the Binomial model [11]. The Negative Binomial model has been applied in emergency situations like disease outbreaks and natural disasters to measure risk promptly. Despite the benefits of group testing, its success depends on choosing appropriate

group sizes. Inadequate group size selection can lead to deficiencies in the procedure. To address these deficiencies, this study aimed to investigate a two-stage adaptive Negative Binomial group testing procedure for estimating the prevalence rate of a rare trait.

2. Methodology

2.1. The Model

The technique known as Negative Binomial sampling holds significant importance in the context of biological sample collection. If the proportion of individuals possessing a specific character trait is denoted as p, and sampling continues until a predetermined number, such as x individuals, is observed, the distribution of the sampled individuals follows a negative binomial distribution. Combining Negative Binomial sampling with group testing provides an attractive approach for delivering early estimates during the screening process.

This model operates under the assumption that the number of pools having a trait of interest is pre-established, and the testing process continues until the desired number of positive pools is identified.

The adaptive scheme involved testing groups in stages and adjusting the group sizes from one stage to the next. The group size used at a stage depended entirely on the outcome of the preceding stages. This implies that $k's$ were determined sequentially as the experiment progressed. The value of k_1 is determined by optimizing the variance of the estimator obtained from the non-adaptive Negative Binomial group testing scheme, which serves as prior information. In Stage One, the number of positive pools to be observed X_1 is fixed, and T_1 the number of groups to be tested before observing X_1 positive pools follows a Negative Binomial distribution.

In Stage Two, the group size k_2 is constructed by minimizing the variance of the estimator obtained in Stage One.

$$k_2 = \arg\min_k \left[\operatorname{var}\left(\hat{p}_1 \right) \right]_{p = \hat{p}_1} \tag{1}$$

where k_2 was the value which minimizes the variance of $\hat{p}_1$ which is the estimator obtained in stage 1. The goal is to select the group size that minimizes the variance of the estimator from Stage One. This approach is crucial for enhancing the precision and accuracy of the overall estimation procedure. By minimizing the variance, the estimation process aims to achieve a more reliable and stable outcome, contributing to the effectiveness of the adaptive estimation model.

The number of positive pools to be observed in Stage Two X_2 is fixed, and the number of groups to be tested T_2 to achieve X_2 positive pools follows a Negative Binomial distribution, which depends on the output of Stage One. This adaptive model ensures an efficient allocation of resources and provides a robust strategy for large-scale screening and identification of positive groups in group testing scenarios.

The derivation presented focused on two models; the usual Negative Binomial

model and the proposed two-stage adaptive Negative Binomial group testing model for estimating the prevalence of a rare trait.

Using the usual non-adaptive Negative Binomial to get $\hat{p}_N$ as in Katholi (2006). T follows a Negative Binomial with parameter x and π.

$$f(t/p) = \binom{t-1}{x-1}\left[\pi(p)\right]^x \left[1-\pi(p)\right]^{t-x} \tag{2}$$

The Likelihood function of Equation (2) is expressed as;

$$L(p/t,x) \propto \left[\pi(p)\right]^x \left[1-\pi(p)\right]^{t-x} \tag{3}$$

The log Likelihood function to base 10 is given as;

$$\log L(p/t,x) \propto x\log\left[\pi(p)\right] + (t-x)\log\left[1-\pi(p)\right] \tag{4}$$

The maximum likelihood estimator of the non-adaptive model is obtained as the solution to

$$\frac{\partial \log L(.)}{\partial(p)} = 0 \tag{5}$$

Which is equivalent to;

$$\frac{x}{\pi(p)} - \frac{t-x}{1-\pi(p)} = 0 \tag{6}$$

where $\pi = 1-(1-p)^k$ which is the probability of obtaining a positive group. Equation (6) yields the results obtained by Katholi (2006) as;

$$\hat{p}_N = 1 - \left(1-\frac{x}{t}\right)^{\frac{1}{k}} \tag{7}$$

Finding the variance of $\hat{p}_N$ the Cramer Rao Lower Bound was used where the Fisher's information of the likelihood function was utilized;

$$\text{Variance of } \hat{p}_N = \frac{1}{I(\hat{p}_N)} \tag{8}$$

where the Fisher's information is given as;

$$I(\hat{p}_N) = -E\left[\frac{\partial^2}{\partial p^2}\log L(.)\right]^{-1} \tag{9}$$

The second derivative of the log likelihood function is given as;

$$= \frac{-t\left(\pi'(p)\right)^2}{\pi(p)\left(1-\pi(p)\right)} \tag{10}$$

It is worth noting that $\pi(p) = 1-(1-p)^k$ and $\pi'(p) = k(1-p)^{k-1}$. Substituting in Equation (10) and taking expectation gives;

$$= \frac{tk^2(1-p)^{2k-2}}{1-(1-p)^k(1-p)^k} \tag{11}$$

The $E(T) = \frac{x}{\pi}$. Thus, taking the expectation and the inverse of Equation (11)

will yield

$$\operatorname{Var}(\hat{p}_0) = \frac{1-(1-p)^k}{tk^2(1-p)^{k-2}} \tag{12}$$

The proposed two-stage adaptive model aimed at optimizing resource allocation and estimation efficiency. The study considered two sets of desired positive groups, X_1 and X_2. The first stage estimator was based on the Negative Binomial distribution with a prior derived from the non-adaptive estimator. In the second stage, the study introduces the two-stage adaptive estimator.

Stage two proceeds by testing groups of size k_2 and T_2 is the number of the groups to be tested to obtain X_2 positive groups. Thus, T_2 is conditioned on T_1. This follows that T_2 has a negative binomial distribution. Specifically,

$$T_2/T_1 \sim \text{Negative Binomial}\left(X_2, \pi_{2/1} = 1-(1-p)^{k_2(t_1)}\right)$$

Equation (6) gives the joint distribution of T_1 and T_2 as;

$$\begin{aligned} f(T_2, T_1) &= f(T_2/T_1) \times f(T_1) \\ &= \text{Negative Binomial}\left(X_1, \pi = 1-(1-p)^{k_1}\right) \\ &\quad \times \text{Negative Binomial}\left(X_2, \pi_{2/1} = 1-(1-p)^{k_2(t_1)}\right) \end{aligned} \tag{13}$$

The joint distribution of T_1 and T_2 was used to derive the final two-stage adaptive estimator $\hat{p}_2$ and is given as;

$$\begin{aligned} f(t_1, t_2) &= \binom{t_1-1}{x_1-1}\left[1-(1-p)^{k_1}\right]^{x_1}(1-p)^{k_1(t_1-x_1)} \\ &\quad \times \binom{t_2-1}{x_2-1}\left[1-(1-p)^{k_2}\right]^{x_2}(1-p)^{k_2(t_2-x_2)} \end{aligned} \tag{14}$$

$\binom{t_1-1}{x_1-1}$ and $\binom{t_2-1}{x_2-1}$ are constants of proportionality thus we replace with $\propto$ giving;

$$f(t_1, t_2) \propto \left[1-(1-p)^{k_1}\right]^{x_1}(1-p)^{k_1(t_1-x_1)} \times \left[1-(1-p)^{k_2}\right]^{x_2}(1-p)^{k_2(t_2-x_2)} \tag{15}$$

The log likelihood function to base 10 for the joint distribution of T_1 and T_2 was obtained as.

$$\begin{aligned} \ln L &\propto x_1 \ln\left[1-(1-p)^{k_1}\right] + k_1(t_1-x_1)\ln(1-p) \\ &\quad + x_2\left[1-(1-p)^{k_2}\right] + k_2(t_2-x_2)\ln(1-p) \end{aligned} \tag{16}$$

The adaptive estimator is obtained by solving Equation (17) iteratively since the solution equated to zero was not in a closed form thus was not tractable.

$$\frac{\partial \log L}{\partial p} \propto \frac{x_1}{\pi_1}\pi_1' - \frac{t_1-x_1}{1-\pi_1}\pi_1' + \frac{x_2}{\pi_2}\pi_2' - \frac{t_2-x_2}{1-\pi_2}\pi_2' \tag{17}$$

where $\pi_1 = 1-(1-p)^{k_1}$ and $\pi_2 = 1-(1-p)^{k_2}$. Also $\pi_1' = k_1(1-p)^{k_1-1}$ and $\pi_2' = k_2(1-p)^{k_2-1}$.

The variance of the adaptive estimator was derived using Fisher's information. Where the Fisher's information is given as;

$$I = -E\left[\frac{\partial^2}{\partial p^2}\log L(.)\right] \qquad (18)$$

Thus, the variance of the adaptive estimator was derived and obtained as;

$$\text{Var}(\hat{p}_A) = \frac{1}{\sum_{i=1}^{2}\frac{x_i k_i^2 (1-p)^{2k_i-2}}{\pi_i^2 (1-\pi_i)}} \qquad (19)$$

This variance was used to construct the Wald confidence interval as;

$$\hat{p}_A \pm Z_{\frac{\alpha}{2}}\sqrt{\text{Var}(\hat{p}_A)} \qquad (20)$$

2.2. Simulation

Data was simulated using an algorithm which mimics the Negative Binomial process as illustrated in **Figure 1**.
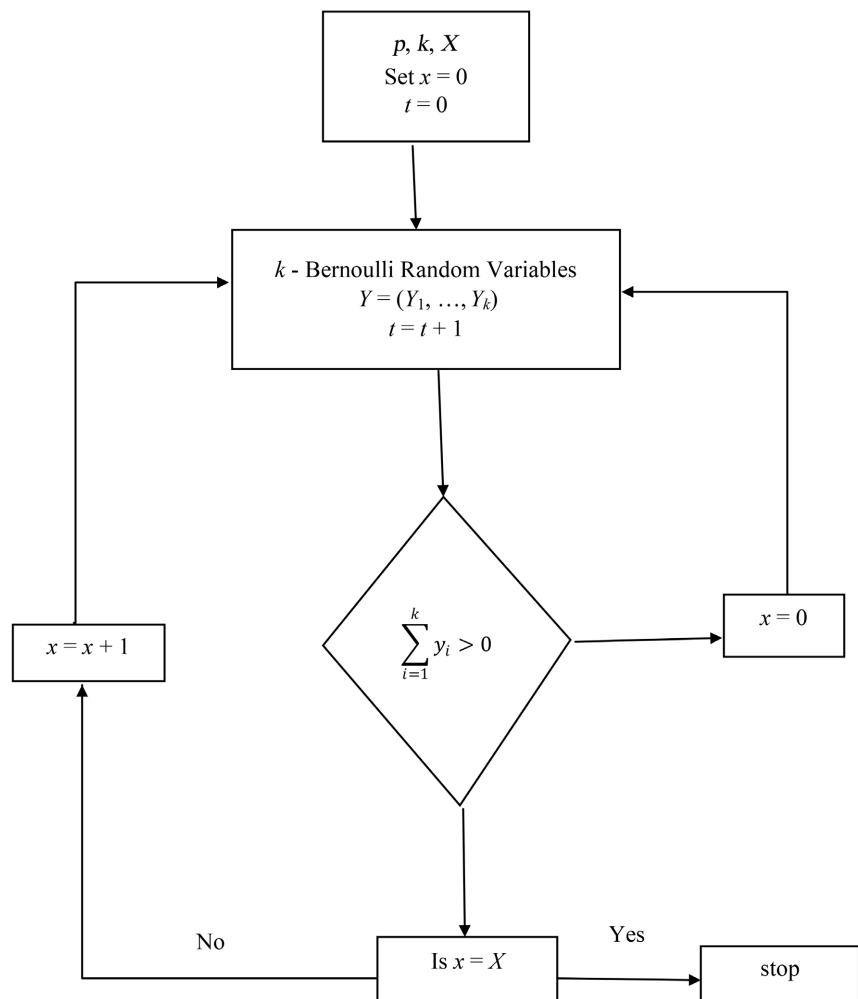

Figure 1. Flow chart for negative binomial group testing simulation.

Steps for Simulation

Step 1: *Specify p, k and X then set x = 0.*

Step 2: *Generate k Bernoulli random variables.*

Set Y = (Y_1, …, Y_k).

Step 3: *If the sum of y_{is} is greater than 0. A success is considered. If the sum of y_{is} is less than the group is considered negative.*

Step 4: *If the success was recorded, repeat the loop if x is not equal to X. If x = X, the procedure stops.*

Step 5: *Report T and calculate p estimate.*

3. Results and Discussion

3.1. Relationship between *t*, *p*, and *k*

The Negative Binomial sampling method used in group testing experiments involves a random number of trials, while the required positive pools and group size are fixed. The number of tests needed to obtain the required positive groups depends on various variables, so it's important to understand how changing the values of *p* and *k* affects the number of testing trials required.

As the probability of success increases in the negative binomial group testing model, the number of trials required to obtain the desired number of positive groups generally decreases as illustrated in Figure 2. Examining the plots, for a fixed value of *k*, it can be observed that as *p* increases, the number of trials decreases. This trend holds true across different values of *k*. This behavior is expected because a higher probability of success implies a greater likelihood of encountering positive groups during testing. Therefore, fewer trials are needed to reach the desired number of positive groups when the probability of success is higher. It is worth noting that increasing the group size from 5 to 100 in the negative binomial group testing model typically leads to a reduction in the number of trials required to achieve the desired number of positive groups as well.
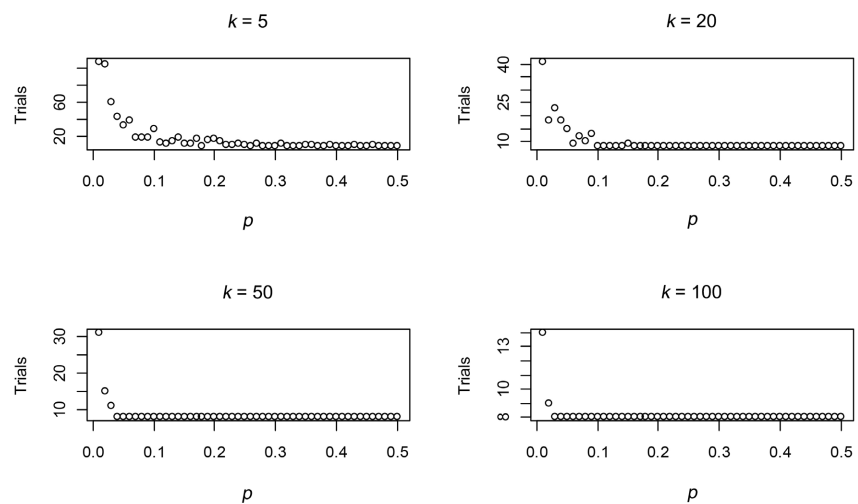

Figure 2. Plots of *T* versus *p* for *k* = 5, 20, 50, 100.

3.2. Adaptive Estimator and Its Properties

The results of the maximum likelihood estimator and its properties including the variance, bias, and mean squared error for the adaptive group testing model are presented in Table 1. The results are organized based on different group sizes and true probabilities while the number of predetermined desired positive groups set at $X = 30$ as set by [15].

Table 1. Adaptive estimator with its properties for $k = 5$, 10, 20, 50, 100 when $X = 30$.

p	Mle	Var	Bias	MSE
		$k = 5$		
0.001	0.000488	7.94E−09	−0.0004	4.0938E−08
0.005	0.002677	2.38E−07	−0.0025	1.02E−07
0.01	0.004862	7.82E−07	−0.0040	4.05E−06
0.05	0.035433	3.94E−05	−0.0209	7.98E−05
0.1	0.075796	0.000167	−0.0275	0.000322241
		$k = 10$		
0.001	0.000551	1.01E−08	−0.0005	3.88E−08
0.005	0.002264	1.70E−07	−0.0022	8.89E−07
0.01	0.005933	1.16E−06	−0.0042	4.18E−06
0.05	0.027808	2.31E−05	−0.0177	8.29E−05
0.1	0.070342	0.000142	−0.0258	0.00034779
		$k = 20$		
0.001	0.000634	1.34E−08	−0.0006	3.59E−08
0.005	0.003144	3.28E−07	−0.0019	1.05E−06
0.01	0.005263	9.15E−07	−0.0057	3.63E−06
0.05	0.029922	2.82E−05	−0.0198	9.04E−05
0.1	0.074906	0.000158	−0.0268	0.00036145
		$k = 50$		
0.001	0.000588	1.15E−08	−0.0004	3.84E−08
0.005	0.002204	1.61E−07	−0.0026	8.88E−07
0.01	0.004198	5.84E−07	−0.0042	3.28E−06
0.05	0.026812	2.27E−05	−0.0223	9.28E−05
0.1	0.075437	0.000165	−0.0506	0.001282513
		$k = 100$		
0.001	0.000488	7.94E−09	−0.0005	3.93E−08
0.005	0.002503	2.08E−07	−0.0018	8.49E−07
0.01	0.00432	6.18E−07	−0.0048	3.70E−06
0.05	0.033403	3.49E−05	−0.0164	0.00030666
0.1	0.068592	0.000137	−0.0175	0.00438814

A scrutiny of Table 1 shows that the estimated probabilities tend to increase as the probability increases, although the increase is generally small. It is important to note that the MLE values of the adaptive model exhibits monotonic behavior as the model dynamically adjusts the group size based on stage one's outcomes which lead to more consistent and accurate estimations. The MLE generally increases as p increases for all values of k as found by [16]. The variance of the estimated probabilities remains relatively small across different values of p. The bias of the estimation shows negative values indicating a slight underestimation of the true population probability. However, the bias remains relatively small across all values of p. The MSE combines the variance and bias to provide an overall measure of estimation accuracy. The bias remains relatively small and consistent across different values of p indicating the effectiveness of the adaptive group testing model in reducing bias.

3.3. Relationship between $\hat{p}$ and p

The results presented in this section examine the relationship between the adaptive maximum likelihood estimates and the true probability for different values of group size. The results are represented in four graphs for $k = 5, 20, 50, 100$. These findings highlighted that the adaptive nature of the model in adjusting the estimations was based on observed outcomes and the varying performance of the adaptive approach across different group sizes.

The relationship was further investigated by plotting the values $\hat{p}$ against p while varying the waiting parameter X for different values of group size k.

Figure 3 illustrates the relationship between the Maximum Likelihood Estimation values and the true probability for different combinations of X in the adaptive approach. The MLE values generally increase as p increases, as well as with larger values of X and k. However, the rate and pattern of this increase vary depending on the specific combination of X and k. For $X = 30$, the MLE graph
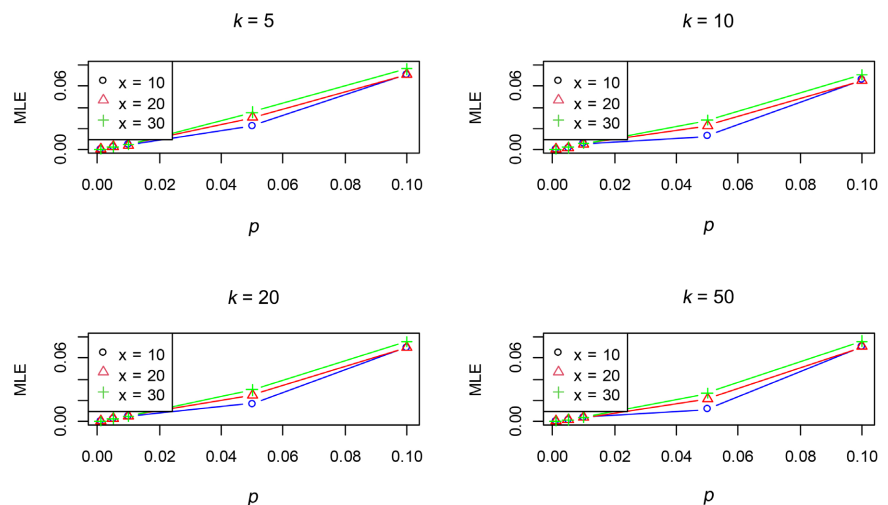

Figure 3. Plots for Adaptive $\hat{p}$ versus p for $k = 5, 20, 50, 100$ and $x = 10, 20, 30$.

shows a relatively steep curve, indicating that small changes in p lead to noticeable changes in the estimated probabilities of success. For $X = 20$, the MLE values also increase as p increases, but the curve is relatively flat, suggesting a less pronounced change in the MLE values as p varies. The adaptive MLE in this scenario exhibits a slower rate of increase compared to when $X = 30$, implying a less sensitive response to changes in p.

Similar patterns are observed for the other combinations of X and k. The adaptive approach tends to provide more conservative estimates with low MLE values. The MLE values increase with increasing p, but the specific patterns and sensitivities depend on the combination of X and k. The adaptive approach consistently provides more conservative estimates with lower MLE values, indicating a cautious approach in estimating the prevalence of the rare trait.

3.4. Relationship between Variance of $\hat{p}$ and p

We examined the relationship between the variance of the estimated proportion $\hat{p}$ and the true proportion p in the context of the two-stage adaptive negative binomial model. **Figure 4** provides insights into this relationship for different combinations of X and k.

Figure 4 presents the interplay between X, k, and the variance of $\hat{p}$ in the two-stage adaptive negative binomial model. They highlight the impact of the true proportion p, the desired number of positive groups X, and the group size k on the variability of the estimates. As the true proportion p increases, the variance of the estimated proportion $\hat{p}$ also tends to increase, although the magnitude of increase varies across different X and k values. This suggests that higher probabilities of success lead to greater variability in the estimates, highlighting the increased uncertainty associated with higher p values. Comparing different X values, we find that as x increases, the variance of $\hat{p}$ generally tends to increase. This implies that aiming for a higher number of positive groups introduces more variability into the estimates. Analyzing the effect of k, we notice
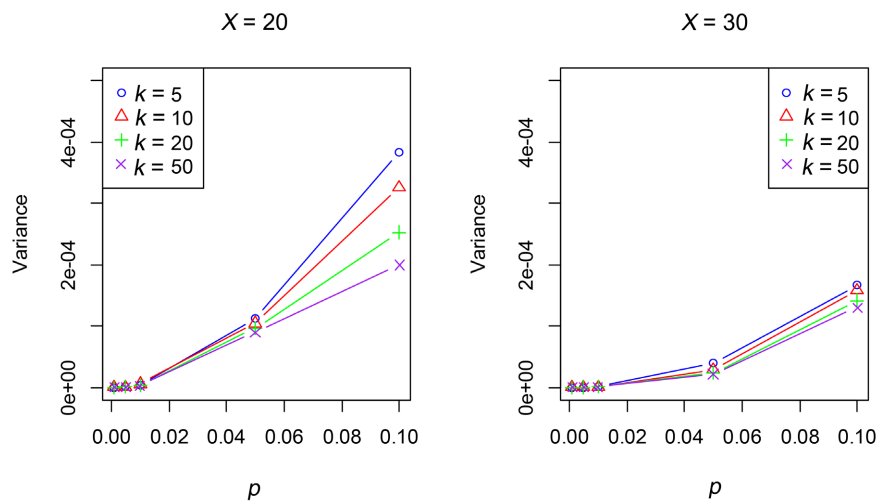

Figure 4. Variance of $\hat{p}$ for $k = 5, 10, 20, 50$ and $x = 20, 30$.

that for a fixed x value, as k increases, the variance of $\hat{p}$ tends to decrease. This indicates that larger group sizes result in more precise estimates and lower variability. Larger group sizes provide more information, reducing the sampling error and enhancing the precision of the estimates. The insights obtained from this analysis can inform the selection of appropriate values for X and k to optimize the accuracy and reliability of the model in estimating the proportion of successes in a population.

4. Comparison of the Model

We conducted a model comparison to evaluate the performance of the two-stage adaptive negative binomial group testing model for estimating the prevalence of a rare trait over the non-adaptive model. In this study, we used two statistical measures, Asymptotic Relative Efficiency and the Relative Mean Squared Error, to compare the efficiency and accuracy of the proposed two-stage adaptive negative binomial group testing model with an existing non-adaptive model by [6].

The estimator of non-adaptive model was denoted by $\hat{p}_N$ since it is developed under the usual Negative Binomial model while the computed estimator was denoted as $\hat{p}_A$ since is developed under the adaptive Negative Binomial model. Then, ARE was obtained as

$$\text{ARE} = \frac{\text{Var}(\hat{p}_N)}{\text{Var}(\hat{p}_A)} \tag{21}$$

ARE values of greater than one implied that our model is more efficient than the non-adaptive model. ARE measures how much more efficient the adaptive model is compared to the non-adaptive model as the sample size approaches infinity. Higher ARE values indicate that the adaptive model provides better estimates and inferences. The comparison was done for different combinations of group size k and the desired number of positive groups X.

The study found that as X increased, there was an overall increasing trend in the ARE values, indicating improved performance in detecting the desired outcome (Figure 5). On the other hand, increasing k for a fixed X value led to a decreasing trend in the ARE values, implying that larger group sizes result in better performance in detecting the desired outcome.

The RMSE was used to compare the mean squared errors of the estimators obtained from the constructed adaptive model with the one by [6]. This is a convenient way of comparing the MSE of the estimates obtained using different procedures. It is expected that a good model to produce an estimator with a small MSE. For this study, RMSE was computed as;

$$\text{RMSE} = \frac{\text{MSE}(\hat{p}_N)}{\text{MSE}(\hat{p}_A)} \tag{22}$$

The study found that the computed estimator was more efficient as compared to the [10] since the values of RMSE were greater than one (Figure 6). It was worth noting also that as the true probability of success p increases, the RMSE
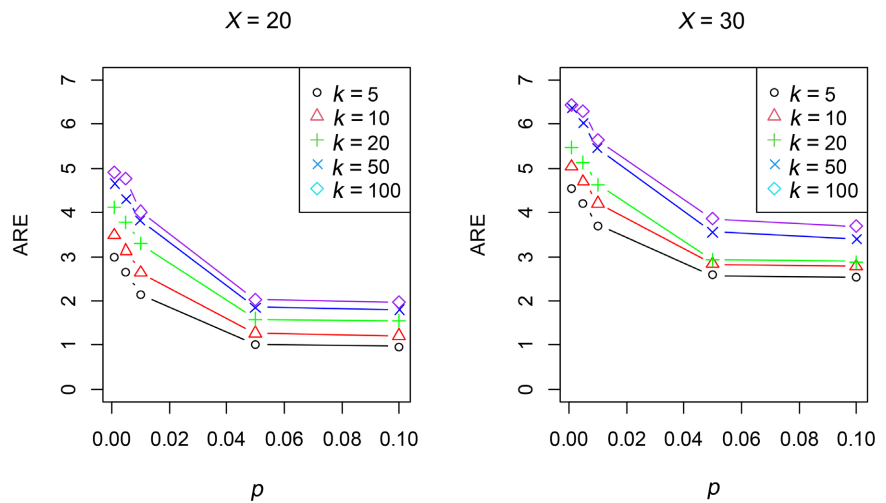
Figure 5. Plots of ARE versus *p* for *k* = 5, 10, 20, 50, 100 and *X* = 20, 30.
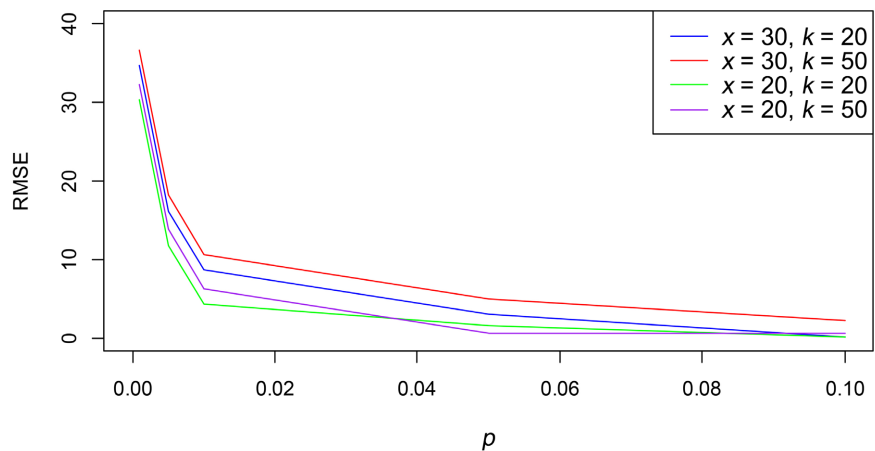

Figure 6. Plots of RMSE versus *p* for *k* = 20, 50 and *X* = 20, 30.

decreases, indicating better fit and improved estimation accuracy at low prevalence. In addition, aiming for a greater number of positive groups and increasing the group size also contributed to higher values of RMSE values, suggesting more accurate estimation and improved model accuracy especially when the prevalence is low.

5. Conclusion and Recommendations

In conclusion, this study successfully achieved its objectives by developing and analyzing a two-stage adaptive negative binomial model in group testing for estimating the prevalence of a rare trait using MLE. The adaptive estimator demonstrated superior performance compared to the non-adaptive estimator, providing more accurate and smaller estimates, while maintaining low variance and bias. The comprehensive simulations further confirmed the superiority of the adaptive model, showing better efficiency, lower mean squared error MSE, and improved fit to the data. The study recommends future research to incorporate

imperfect tests in the model to reflect real-world scenarios and evaluate their impact on estimation accuracy. In addition, exploring the extension of the adaptive estimator to multi-stage group testing procedures could enhance the model's applicability to larger populations and improve logistical considerations for estimation.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Dorfman, R. (1943) The Detection of Defective Members of Large Populations. *The Annals of Mathematical Statistics*, **14**, 436-440. https://doi.org/10.1214/aoms/1177731363

[2] Sobel, M. and Elashoff, R.M. (1975) Group Testing with a New Goal, Estimation. *Biometrika*, **62**, 181-193. https://doi.org/10.1093/biomet/62.1.181

[3] Tamba, C.L. (2012) Computational Statistical Model for Group Testing with Retesting. Ph.D. Thesis, Egerton University, Kenya.

[4] Wanyonyi, R.W. (2015) Estimation of Proportion of a Trait by Batch Testing Model in a Quality Control Process. *American Journal of Theoretical and Applied Statistics*, **4**, 610-613. https://doi.org/10.11648/j.ajtas.20150406.34

[5] Bilder, C.R., Tebbs, J.M. and Chen, P. (2010) Informative Retesting. *Journal of the American Statistical Association*, **105**, 942-955. https://doi.org/10.1198/jasa.2010.ap09231

[6] Thompson, K.H. (1962) Estimation of the Proportion of Vectors in a Natural Population of Insects. *Biometrics*, **18**, 568-576. https://doi.org/10.2307/2527902

[7] Nyongesa, L.K. (2004) Testing for the Presence of Disease by Pooling Samples. *New Zealand Journal of Statistics*, **46**, 383-390. https://doi.org/10.1111/j.1467-842X.2004.00337.x

[8] Chiang, C.L. and Reeves, W.C. (1962) Statistical Estimation of Virus Infection Rates in Mosquito Vector Populations. *American Journal of Epidemiology*, **75**, 377-391. https://doi.org/10.1093/oxfordjournals.aje.a120259

[9] Hughes-Oliver, J.M. and Swallow, W.H. (1994) A Two-Stage Adaptive Group-Testing Procedure for Estimating Small Proportions. *Journal of the American Statistical Association*, **89**, 982-993. https://doi.org/10.1080/01621459.1994.10476832

[10] Katholi, C.R. and Unnasch, T.R. (2006) Important Experimental Parameters for Determining Infection Rates in Arthropod Vectors Using Pool Screening Approaches. *The American Journal of Tropical Medicine and Hygiene*, **74**, 779-785. https://doi.org/10.4269/ajtmh.2006.74.779

[11] Pritchard, N.A. and Tebbs, J.M. (2010) Estimating Disease Prevalence Using Inverse Binomial Pooled Testing. *Journal of Agricultural, Biological, and Environmental Statistics*, **16**, 70-87. https://doi.org/10.1007/s13253-010-0036-4

[12] Pritchard, N.A. and Tebbs, J.M. (2011) Bayesian Inference for Disease Prevalence Using Negative Binomial Group Testing. *Biometrical Journal*, **53**, 40-56. https://doi.org/10.1002/bimj.201000148

[13] Okoth, A.W., Nyongesa, L.K. and Kwatch, B.O. (2017) Multi-Stage Adaptive Pool Testing Model with Test Errors; Improved Efficiency. *IOSR Journal of Mathematics*, **13**, 43-55. https://doi.org/10.9790/5728-1301024355

[14] Turechek, W.W. and Madden, L.V. (2003) A Generalized Linear Modeling Approach for Characterizing Disease Incidence in a Spatial Hierarchy. *Phytopathology*, **93**, 458-466. https://doi.org/10.1094/PHYTO.2003.93.4.458

[15] Xiong, W. (2015) The Optimal Group Size Using Inverse Binomial Group Testing Considering Misclassification. *Communications in Statistics—Theory and Methods*, **45**, 4600-4610. https://doi.org/10.1080/03610926.2014.923461

[16] Swallow, W.H. (1985) Group Testing for Estimating Infection Rates and Probabilities of Disease Transmission. *Phytopathology*, **75**, 882.
https://doi.org/10.1094/Phyto-75-882