Open Access Library Journal
2023, Volume 10, e10141
ISSN Online: 2333-9721
ISSN Print: 2333-9705

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

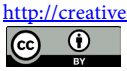
Francis Mwangi Kariuki, Ronald Waliaula Wanyonyi, Ali Salim Islam

Department of Mathematics, Egerton University, Nakuru, Kenya
Email: kariukif93@gmail.com, ronaldwaliaula@gmail.com, asislam54@gmail.com

How to cite this paper: Kariuki, F.M., Wanyonyi, R.W. and Islam, A.S. (2023) Analysis of a Two-Stage Negative Binomial Group Testing Model for Estimating the Prevalence of a Rare Trait. *Open Access Library Journal*, **10**: e10141.
https://doi.org/10.4236/oalib.1110141

Received: April 10, 2023
Accepted: June 24, 2023
Published: June 27, 2023

Copyright © 2023 by author(s) and Open Access Library Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/



Open Access

Abstract

This paper presents the analysis of a two-stage negative binomial group testing estimator of the prevalence of a rare trait when imperfect diagnostic tests with known sensitivity and specificity were used. The study utilized the method of Maximum Likelihood Estimation (MLE) to obtain the estimator and the Cramer-Rao lower bound method to compute the Fischer information of the estimator. The properties of the constructed estimator are discussed and the efficiency of the constructed estimator relative to other estimators in pool testing scheme was determined by computing the Asymptotic Relative Efficiency (ARE) and the Relative Mean Squared Error (RMSE). The procedure was illustrated, and the model was verified by performing Monte Carlo simulations using R programming language.

Subject Areas

Statistics

Keywords

Group Testing, Prevalence, Retesting, Inverse Sampling

1. Introduction

The standard method of screening individuals for the presence or absence of a rare trait of interest (e.g. disease) is uneconomical, especially when the target population is large, and the prevalence is low. A feasible strategy is to pool individuals into groups which are then tested as units. Group testing suggests a considerable amount of savings in efficiency, and the number of tests performed compared to individual testing if rightly applied. The idea dates back to World

War II and was applied to screen for syphilis antigen among army inductees [1].

In statistical literature, group testing literature splits into two distinct areas. The first one is classification whose objective is to identify the positive units having a trait of interest. The first documented work in classification was applied to identify syphilitic men that were called for army induction during World War II [1]. The pooling strategy achieved a significant amount of savings by reducing the number of tests needed by testing groups of blood samples as opposed to testing individual blood samples of the army recruits. Since its inception, different authors have extended and generalized the procedure. Hierarchical testing schemes were shown to achieve greater savings [2] [3]. The procedure has been applied in a variety of fields such as in drug discovery [4], quality control processes [5], and concealing the identity of individuals when screening for HIV antibodies [6].

This paper will focus on the second objective which is on estimation of the prevalence of a rare trait. Research in public health shows that when examining diseases transmitted by insect vectors even if disease identification is the main focus, it is equally important to estimate the prevalence (p) of a pathogen to curb and prevent an outbreak [7]. The first work in estimation was developed by [8] to estimate the proportion of insect vectors that are capable of transmitting the aster-yellow virus. In recent studies, different authors have extended [8] work by considering multi-stage pool testing schemes [9] [10] [11]. Retesting of pools when imperfect diagnostic tests are used has been examined [9] [12] [13] and sample size procedures for estimating the adventitious presence (AP) of transgenic plants in a population have been examined [14] [15]. Their research has largely developed under the binomial model which presumes a fixed number of groups to be tested for a trait of interest. In situations where an unknown number of pools are to be tested until a predetermined number of positive pools are observed, calls for the negative binomial model to be considered.

A procedure called inverse sampling is of great importance when sampling biological samples. It entails continuous sampling until a fixed number of samples having a trait of interest are observed. Inverse sampling was first used to estimate the frequencies of an attribute in a population [16]. A combination of inverse sampling and group testing was suggested to be viable in determining the infection rate of diseases transmitted by insect vectors [17]. It was noted that the sources of errors in group testing designs were attributed to the sampling process and not the pooling process. Thus, if sampling and pooling are done correctly, more desirable results can be achieved. The procedure was established to be suitable in situations that prompt immediate responses like disease outbreaks and natural disasters. Early detection of infectious diseases and quick countermeasures are important to mitigate the effect of an outbreak [18]. Inverse (negative) binomial group testing was applied to assess the transmission of the parasitic worm *Onchocerciasis volvulus* by blackflies which is responsible for causing blindness and skin diseases [19]. Several point estimators that mini-

mized the bias of the estimator were considered. The results showed the potential applicability of the model in the screening of West Nile Virus (WNV) and Foot and Mouth Disease (FMD) [20].

The foundation in estimation under the inverse binomial group testing model was established under the assumption that perfect diagnostic tests and pools of equal sizes are used [20]. The MLE of the proportion p was established to be positively biased, and an almost unbiased estimator within a region of interest was developed by applying a suitable correction method [7]. A comparison of group sizes was considered [21] and applied in the detection of the AP of transgenic plants in a population. In their subsequent work [22] examined the sample size procedures for estimating AP by considering the dilution effect. The MLE was applied and sample sizes that guaranteed a narrow confidence width were obtained. The optimal group sizes that minimized the asymptotic variance of the estimator when imperfect tests are used have also been examined [23]. Sequential estimation under the negative binomial group testing model has been considered and the results established that an unbiased estimator does not exist [24].

The negative binomial group testing model has largely developed under the postulation that perfect tests are used (*i.e.*, pools are not misclassified). However, misclassification errors can occur in an experimental design when the diagnostic tests used are imperfect (*i.e.* the sensitivity and the specificity of the diagnostic test are less than 1). The focus has been to examine efficient estimators and to determine optimal group sizes. Different authors have examined the retesting of pools under the binomial model in group testing, and have established that retesting improves the efficiency of an estimator, and it recovers the sensitivity of the diagnostic tests during estimation [9] [12] [13]. It is on this background that this paper is developed.

In this paper, we construct and analyze a testing scheme developed under the negative binomial model that incorporates retesting of pools when imperfect tests are used. A pool that tests positive in the initial test is sequentially given a retest. The properties of the estimator are discussed and the efficiency of the constructed estimator relative to the one-stage negative binomial group testing model with misclassification was analyzed.

2. Simulation

A generalized Monte-Carlo simulation algorithm used in this study involved the following:

Step 1: Set fixed values of n, k, and p at known sensitivity and specificity values.

Step 2: Generate N independent data sets from negative binomial $\left(n,\pi^{*}\left(p\right)\right)$ that is $T_i \sim \text{nbinom}\left(n,\pi^{*}\left(p\right)\right)$ for $i=1,\cdots,N$.

Step 3: Compute the numerical value of the test statistics T for each data set $T_1,T_2,\cdots,T_N$.

Step 4: If N is large enough, summary statistics $T_1,T_2,\cdots,T_N$ should be a good approximation to the true sampling properties of the test condition under the

conditions of interest.

The study simulated different data sets of pool sizes, $k = 5,10,30,50$. The number of positive groups that tested positive on retest was fixed at $n = 1,3,5,10,15$. The sensitivity and specificity values of the tests are assumed to remain constant throughout the entire study and are set at 99%, 98%, 95%, and 90%.

3. Point Estimation

The study assumes that the outcome of individual disease statuses are independently and identically distributed (i.i.d) Bernoulli random variable. We also assume that imperfect diagnostic tests with known sensitivity and specificity defined as $\pi_0 = Pr\left(X_j = 0 \mid D_j = 0\right)$ and $\pi_1 = Pr\left(X_j = 1 \mid D_j = 1\right)$ respectively were used. The testing process is performed sequentially until the nth positive pool that tests positive on retesting is observed.

Suppose that several pools of size k are sequentially tested for the presence of a rare trait in a population until the predetermined nth pool that tests positive on a retest is observed. The total number of pools tested is denoted by $T = \sum_{j=1}^{n} T_j$ where $\sum_{j=1}^{l} T_j$ is the number of groups sequentially tested until the lth positive group that tests positive on retesting is observed on a retest. Let X_i be 1 if the ith group tests positive and 0 otherwise for $i = 1,2,\ldots T$. Let D_i be 1 if the ith group is truly positive and 0 otherwise for $i = 1,2,\ldots T$. If the proportion of the rare trait in a given population is denoted by p, using the theory of probability and indicator function, then the probability that a pool tests positive on retesting is

$$\pi^*(p) = \pi_1^2 + (1-p)^k \left((1-\pi_0)^2 - \pi_1^2 \right). \tag{1}$$

based on the sampling scheme, T follows the negative binomial distribution with parameters n and $\pi^*(p)$. Then the Likelihood function is

$$L(p) = \binom{t-1}{n-1} \pi^{*n} \left(1-\pi^*\right)^{t-n} \tag{2}$$

and the Maximum likelihood Estimator (MLE) of proportion p is

$$\hat{p} = 1 - \left\{ \frac{\pi_1^2 - \frac{n}{t}}{\pi_1^2 - (1-\pi_0)^2} \right\}^{\frac{1}{k}}. \tag{3}$$

The table below shows the MLE of p for different values of p, n, and k when the sensitivity and specificity of the tests are known to be 99%.

From Table 1, it can be observed that for any fixed values of n and k, the MLE of p increases with an increase in the prevalence. Also, for any fixed values of k and p, the MLE of p decreases as the waiting parameter n increases. A close approximation of the prevalence is observed at low values of p when both k and n are large, but high values of p when both k and n are large are sufficient to overestimate the prevalence level. Thus, scrutiny of the table reveals that it is possible to get a combination of n and k values that gives a close approximation of the prevalence.

Table 1. The MLE of p for $n = 1,3,5,10,15$; $k = 5,10,30,50$.

			n		
p	1	3	5	10	15
			$k = 5$		
0.005	0.037083	0.007335	0.006203	0.005540	0.005367
0.01	0.072883	0.014681	0.012411	0.011042	0.010696
0.05	0.277838	0.077258	0.062284	0.055110	0.053197
0.10	0.454758	0.175649	0.128055	0.110799	0.106557
0.20	0.697056	0.419199	0.306039	0.230217	0.215374
0.30	0.837360	0.643550	0.526881	0.392980	0.347869
			$k = 10$		
0.005	0.060468	0.007355	0.006216	0.005534	0.005351
0.01	0.108036	0.015609	0.012384	0.011074	0.010694
0.05	0.421494	0.116445	0.069269	0.055804	0.053448
0.10	0.658052	0.330505	0.199350	0.120100	0.108827
0.20	0.885850	0.709192	0.587668	0.403237	0.309704
0.30	0.957030	0.879560	0.815683	0.696210	0.613936
			$k = 30$		
0.005	0.143000	0.010042	0.006235	0.005524	0.005344
0.01	0.267490	0.027645	0.013566	0.011109	0.010681
0.05	0.774480	0.475560	0.302159	0.120312	0.070893
0.10	0.943910	0.833082	0.740952	0.569289	0.442061
0.20	0.978910	0.940975	0.903185	0.824583	0.754329
0.30	0.980270	0.944985	0.908833	0.834533	0.766407
			$k = 50$		
0.005	0.228030	0.015558	0.006729	0.005549	0.005339
0.01	0.394580	0.067095	0.019821	0.011208	0.010726
0.05	0.911220	0.746833	0.620492	0.394786	0.257628
0.10	0.975650	0.928954	0.883985	0.787707	0.704396
0.20	0.980080	0.944023	0.906544	0.828924	0.756956
0.30	0.980080	0.944120	0.906641	0.829114	0.756958

Next, we plot the relationship between the MLE and the prevalence p for different values of k and n when the assays used are 99% accurate.

We investigate the relationship between $\hat{p}$ and p when the sensitivity and specificity of the test are 99%. A linear relationship exists between $\hat{p}$ and p when the group size $k = 1$. However, when $k > 1$, it can be observed in **Figure 1** that the relationship is monotonic at any predetermined value of the waiting parameter n. A striking feature to note is that the relationship is sensitive to both k and n when the sensitivity and specificity of the diagnostic test are known.

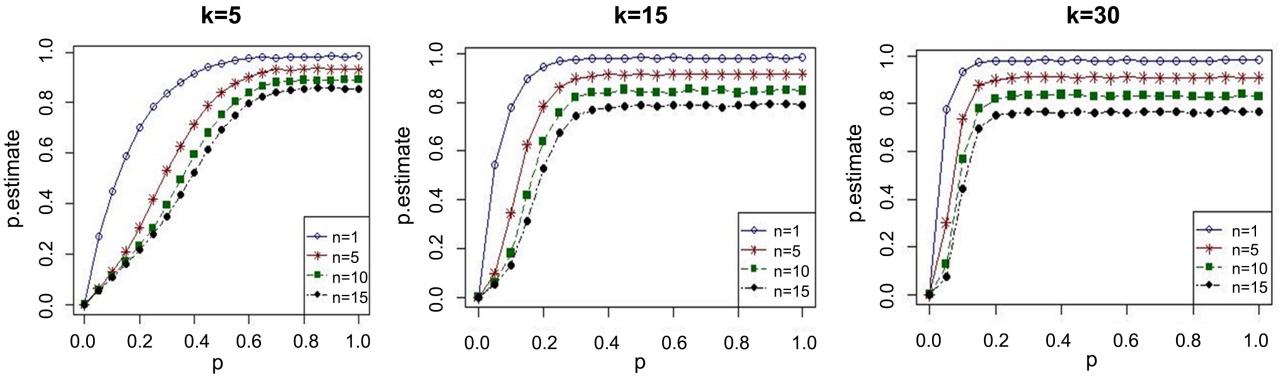


Figure 1. The relationship between $\hat{p}$ and p for $k = 5$, 15, 30 and $n = 1$, 5, 10 and 15.

4. Properties of the Estimator

In this section, the properties of the estimation such as the Biasedness, and the Mean Squared Error (MSE) are discussed.

4.1. Biasedness of the Estimator

The goal in estimation is to find an estimate of the proportion p in a population of interest. The bias of an estimator measures the average error incurred when using the estimate of a parameter. It was pointed out that bias is particularly useful in evaluating point estimates [8]. The exact bias of an estimator is given by:

$$\text{Bias}(\hat{p}) = E(\hat{p}) - p. \tag{4}$$

Because $T \sim \text{Negative Binomial}\left(n, \pi^*(p)\right)$ where $\pi^*(p) = \pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)$. The exact bias and the MSE can be expressed as an infinite sum. It was noted that the sums do not reduce to anything tractable [20]. Thus, the bias of an estimator $\hat{p}$ can be approximated as follows:

$$\text{Bias}(\hat{p}) = \sum_{t=n}^{t^*}(\hat{p} - p)\binom{t-1}{n-1}\pi^{*n}\left(1-\pi^*\right)^{t-n} \tag{5}$$

where $pr\left(T \le t^*\right) \ge 1 - v$ for v small. The value of $v = 0.00001$ was taken throughout, making these approximations very close to the true values of bias and MSE.

Table 2 shows that for any fixed values of the group size and the waiting parameter n, the absolute bias of the estimator increases to a maximum as the optimal value of the prevalence is attained, and afterward decreases as the prevalence increases. Conversely, for a given group size and prevalence level, the absolute bias of the estimator decreases as the waiting parameter n increases. The results show that it is possible to estimate the bias of the estimator at a given level of p for a set of n and k values. For example, when $p = 0.005$, the minimum bias can be observed at $n = 15$ and $k = 10$.

Scrutiny of Table 3 shows that when the diagnostic tests are 95% accurate, the absolute bias of the estimator increases to a maximum as the optimal value of the prevalence is attained and afterward decreases for fixed group sizes and waiting

Table 2. Bias of $\hat{p}$ for various values of p with n = 1, 3, 5, 10, 15 and k = 5, 10, 30, 50 when the sensitivity and specificity of the tests are set at 99%.

Bias × 10^{-4}					
	n				
p	1	3	5	10	15
			k = 5		
0.005	362.80610	37.39135	23.41135	14.38200	11.23476
0.01	653.08290	74.66184	46.54191	28.62074	22.36416
0.05	2371.60600	418.93271	231.33740	140.27249	109.71514
0.10	3840.02890	1037.38995	511.51869	279.00437	217.11517
0.20	5573.34120	2682.07576	1513.74906	650.96485	456.72240
0.30	6035.75530	4163.31983	2864.88066	1455.92420	925.87968
0.40	5685.07450	4860.81024	4057.69919	2519.13300	1829.42647
			k = 10		
0.005	567.82770	38.23104	23.37473	14.36144	11.21754
0.01	1048.95920	80.94800	46.73266	28.64085	22.37064
0.05	3856.78190	821.10962	309.72758	145.44342	112.11377
0.10	5901.92240	2564.13220	1247.39388	391.76074	247.93661
0.20	7167.71460	5611.21106	4330.80178	2387.50558	1421.94666
0.30	6777.53620	6278.09096	5776.61834	4625.86607	3656.79622
0.40	5930.78080	5748.91242	5547.54441	5027.30965	4520.02533
			k = 30		
0.005	1407.68200	61.49968	24.03181	14.39814	11.23803
0.01	2572.09800	228.50082	57.97831	29.02309	22.58504
0.05	7379.39100	4393.14376	2659.68468	818.70792	315.13048
0.10	8494.38800	7535.94242	6669.32861	4886.44017	3558.92440
0.20	7868.38400	7599.10928	7331.94058	6691.10470	6096.45891
0.30	6915.64500	6738.41412	6559.49977	6122.70470	5709.80285
0.40	5955.53000	5855.73383	5751.20042	5487.76257	5231.61935
			k = 50		
0.005	2191.00100	135.81610	28.52969	14.49044	11.28822
0.01	3858.27000	625.75440	131.99412	30.48499	23.01542
0.05	8629.29400	7095.11760	5821.19533	3537.86433	2194.31129
0.10	8796.89600	8395.50360	8007.35776	7103.89562	6294.55529
0.20	7877.70300	7632.13780	7390.91512	6816.34526	6285.50365
0.30	6917.58000	6749.42800	6582.56056	6181.28160	5807.13767
0.40	5957.37300	5866.46570	5773.79640	5545.46300	5327.73594

Table 3. Bias of $\hat{p}$ for various values of p with n = 1, 3, 5, 10, 15 and k = 5, 10, 30, 50 when the sensitivity and specificity of the tests are set at 95%.

Bias × 10^{-4}					
			n		
P	1	3	5	10	15
			k = 5		
0.005	380.76220	41.52092	25.95154	15.93273	12.44444
0.01	653.68000	79.14075	49.22275	30.22827	23.61278
0.05	2255.26870	419.97179	238.73971	144.13236	112.44601
0.10	3602.94990	978.32673	521.57424	293.12960	225.82892
0.20	5175.96070	2294.87368	1354.88762	808.08288	519.35439
0.30	5677.46480	3303.48420	2321.27912	1906.67545	1137.52123
0.40	5434.48780	4046.78395	2823.63207	2950.41699	1981.60932
			k = 10		
0.005	560.99250	40.47671	24.74973	15.17049	11.84578
0.01	1008.15300	83.12783	48.40546	29.57685	23.08846
0.05	3593.23780	735.62461	302.91655	153.64969	116.63904
0.10	5458.47350	2124.49065	1032.87077	524.91204	284.51085
0.20	6698.41200	4483.44720	3047.88096	2993.50067	1674.04049
0.30	6418.43550	5219.55184	4137.03897	4614.50812	3137.79043
0.40	5718.11860	5026.75652	4321.83066	4731.85814	3693.31595
			k = 30		
0.005	1325.23900	58.81119	24.68150	14.79048	11.52974
0.01	2396.62500	201.03076	57.00882	29.91896	23.19216
0.05	6817.86500	3479.66278	1848.61660	1315.15555	442.30362
0.10	7881.07700	5984.48217	4502.89514	5074.04357	3235.39221
0.20	7383.34200	6252.38117	5276.53975	6002.40539	4670.04657
0.30	6585.20200	5799.59064	5103.11115	5617.62121	4654.29024
0.40	5780.27900	5329.62519	4905.94283	5201.99972	4597.72182
			k = 50		
0.005	2042.00100	118.19540	28.08787	14.90331	11.57584
0.01	3573.01200	513.19890	108.85200	33.54519	23.92129
0.05	7974.66300	5579.64670	3875.45159	4143.91552	2256.64394
0.10	8167.33000	6701.98660	5484.81003	6367.58059	473092535
0.20	7400.01500	6325.07340	5414.47400	6117.29080	4889.21737
0.30	6595.08500	5855.0844	5217.26849	5701.61852	4832.57489
0.40	5790.07600	5384.89830	5019.79203	5285.59631	4775.47022

parameter n. Secondly, when the group size and the prevalence levels are fixed, the absolute bias decreases to a minimum value as the optimal value of the waiting parameter n is attained, and afterward, it increases. Lastly, the results show that one can estimate the prevalence level at a predetermined optimal value of the waiting parameter n that would register the least bias when the group size is known. For example, when $n=15$ and $k=30$ the least bias is observed at $p=0.005$.

4.2. Mean Squared Error of the Estimator

The Mean squared error of an estimator is the average squared deviation derived from the true value of the parameter, which incorporates measures of both accuracy (bias) and the precision (variance) of the estimator. It is used as a measure for the goodness of an estimator that is given by:

$$\mathrm{MSE}(\hat{p}) = \left(\mathrm{bias}(\hat{p})\right)^2 + \mathrm{var}(\hat{p}). \tag{6}$$

The variance of $\hat{p}$ as annexed in **Appendix** section can easily be shown to be

$$\mathrm{Var}(\hat{p}) = \left[\frac{1-\left(\pi_1^2+(1-p)^k\left((1-\pi_0)^2-\pi_1^2\right)\right)}{nk^2(1-p)^{2k-2}\left((1-\pi_0)^2-\pi_1^2\right)^2}\right]\left(\pi_1^2+(1-p)^k\left((1-\pi_0)^2-\pi_1^2\right)\right)^2. \tag{7}$$

The MSE can also be expressed as an infinite sum, which was shown that they do not reduce to anything tractable [20]. Thus, the MSE of an estimator $\hat{p}$ can be approximated as

$$\mathrm{MSE}(\hat{p}) \approx \sum_{t=n}^{t^*}(\hat{p}-p)^2\binom{t-1}{n-1}\pi^{*n}(1-\pi^*)^{t-n} \tag{8}$$

where $pr\left(T \le t^*\right) \ge 1-v$ for v small. The value of $v=0.00001$ was taken throughout, making these approximations close to the true values of bias and MSE.

Since the sums do not reduce to anything tractable, we use Monte Carlo simulations to generate the MSE tables as outlined in Section 2.

Scrutiny of Table 4 shows that when the group size and the waiting parameter n are fixed, the MSE of $\hat{p}$ increases to a maximum value as the optimal value of the prevalence is attained before subsequently decreasing. It can also be noted that for any fixed values of the prevalence and group size, the MSE of the estimator decreases as the waiting parameter, n increases. Therefore, it is possible to get the minimum MSE of the prevalence from a combination of n and k when the sensitivity and specificity of the tests are 99%. For instance, when p = 0.005 the minimum value of MSE was obtained at n = 15 and k = 10.

When the assay used is 95% accurate, it can be seen from Table 5 that for any fixed values of the group size and the waiting parameter n, the MSE of the estimator increases to a maximum as the optimal value of the prevalence is attained and afterward decreases. Secondly, when both the group size and the prevalence rate are fixed, the MSE of the estimator decreased to a minimum as the optimal

Table 4. MSE of $\hat{p}$ for various values of p with n = 1, 3, 5, 10, 15 and k = 5, 10, 30, 50, and when the sensitivity and specificity of the tests are set at 99%.

MSE × 10^{-4}					
	n				
p	1	3	5	10	15
			k = 5		
0.005	248.0600	0.704640	0.144780	0.041657	0.023382
0.01	482.9300	3.130540	0.571370	0.164670	0.092545
0.05	2016.7300	123.779720	17.538790	3.936600	2.215701
0.10	3261.6100	564.307690	121.615100	16.517000	8.731067
0.20	4247.3500	1857.421800	839.387760	150.740000	49.387109
0.30	4054.6700	2695.858660	1797.260460	685.060000	289.682941
0.40	3325.5500	2732.301220	2228.374240	1346.600000	832.097851
			k = 10		
0.005	478.310000	1.629700	0.147543	0.041564	0.023317
0.01	923.560000	9.917900	0.647720	0.165220	0.092695
0.05	3552.640000	561.137100	95.807460	5.135700	2.374300
0.10	5178.770000	2115.678900	873.769169	107.580000	20.79600
0.20	5622.070000	4306.743200	3295.318680	1694.70000	881.810000
0.30	4692.580000	4272.603200	3879.890790	3040.100000	2381.000000
0.40	3535.480000	3382.286300	3224.232340	2844.800000	2501.200000
			k = 30		
0.005	1356.100000	25.828000	0.625270	0.042023	0.023457
0.01	2501.600000	163.923000	11.184170	0.184380	0.095335
0.05	6949.000000	4117.492000	2440.360710	661.960000	181.220000
0.10	7606.000000	6702.058000	5904.136630	4299.700000	3131.500000
0.20	6271.500000	6017.749000	5772.132500	5197.400000	4677.500000
0.30	4817.600000	4650.590000	4486.036800	4094.100000	3732.900000
0.40	3556.600000	3462.454000	3366.198140	3129.500000	2905.000000
			k = 50		
0.005	2152.033000	101.968400	4.951560	0.045301	0.023790
0.01	3795.008000	569.470800	85.782540	0.930301	0.107290
0.05	8166.129000	6684.168800	5470.766540	3315.612502	2009.934470
0.10	7899.744000	7512.479800	7143.493100	6297.265286	5550.580170
0.20	6279.433000	6042.506400	5813.196300	5275.105616	4785.490050
0.30	4818.755000	4656.493700	4497.833960	4121.942498	3776.736100
0.40	3557.990000	3470.234900	3382.076640	3168.059613	2967.003440

Table 5. MSE of $\hat{p}$ for various values of p with n = 1, 3, 5, 10, 15 and k = 5, 10, 30, 50, and when the sensitivity and specificity of the tests are set at 95%.

MSE × 10^{-4}					
	n				
p	1	3	5	10	15
			k = 5		
0.005	253.980000	0.857360	0.178530	0.051170	0.028702
0.01	470.500000	3.373740	0.643590	0.184030	0.103284
0.05	1879.900000	110.271640	17.963370	4.213000	2.344348
0.10	3018.780000	471.545300	107.371930	20.338320	9.666764
0.20	3919.450000	1480.895760	620.064980	269.110180	73.930991
0.30	3753.850000	2122.574940	1238.619640	1022.769360	407.850012
0.40	3113.380000	2171.630380	1509.867600	1589.588120	885.358940
			k = 10		
0.005	464.030000	1.624700	0.165970	0.046476	0.026033
0.01	873.100000	8.942400	0.686450	0.176990	0.098977
0.05	3285.510000	451.303000	72.453480	7.534300	2.642158
0.10	4776.330000	1667.629100	603.683580	224.190000	35.406000
0.20	5190.500000	3373.496200	2196.958960	2189.500000	1018.600000
0.30	4361.120000	3384.466500	2605.707320	2989.800000	1987.200000
0.40	3339.040000	2774.564400	2265.512850	2591.900000	1910.800000
			k = 30		
0.005	1269.400000	21.336000	0.508233	0.044754	0.024797
0.01	2321.200000	131.450000	7.997270	0.233470	0.101530
0.05	6403.900000	3222.576000	1624.382680	1139.400000	284.890000
0.10	7008.900000	5239.424000	3914.380480	4447.800000	2765.300000
0.20	5799.000000	4745.247000	3875.095560	4549.100000	3396.300000
0.30	4495.700000	3763.511000	3141.892010	3619.400000	2785.400000
0.40	3385.900000	2965.426000	2586.401920	2860.900000	2337.000000
			k = 50		
0.005	2000.300000	82.018000	3.507400	0.056069	0.025197
0.01	3508.900000	450.479000	58.327400	2.725400	0.132810
0.05	7522.200000	5223.102000	3626.523400	3885.200000	2049.800000
0.10	7280.000000	5875.751000	4739.874600	5584.600000	4084.200000
0.20	5809.200000	4779.079000	3929.325800	4601.500000	3476.000000
0.30	4501.300000	3792.030000	3196.577500	3659.800000	2863.100000
0.40	3393.300000	3004.790000	2663.567700	2917.700000	2449.600000

value of the waiting parameter, n is attained and afterward increases. Finally, the results show that one can obtain a combination of n and k values that would yield a minimum approximation of the MSE at a given prevalence. For instance, the minimum value of the MSE is obtained at $p = 0.005$ when $n = 15$ and $k = 30$.

The behavior of the MSE is investigated by plotting $\hat{p}$ against p for different k and n when the tests used are 99% accurate as shown in the figures below.

Next, the relationship between $\mathrm{MSE}(\hat{p})$ and the prevalence is examined for different k and n when the tests used are accurate as presented below.

Figure 2 shows the MSE of the estimator plotted against the prevalence for different values of the waiting parameter n, and group sizes obtained by simulation. A striking feature to note is that the MSE is sensitive to the group size and the waiting parameter n. Also, the maximum value of the MSE increases with an increase in the group size at any predetermined waiting parameter n.

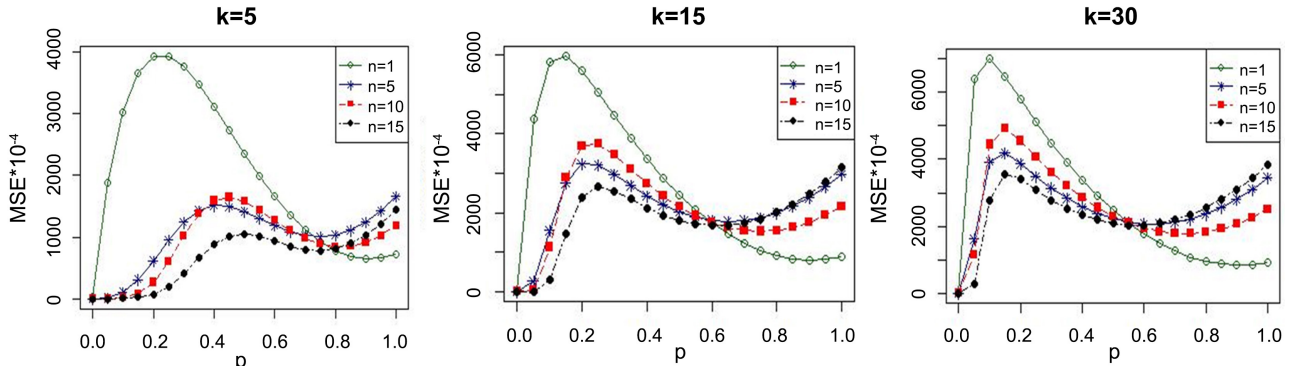


Figure 2. Plot of $\mathrm{MSE}(\hat{p})$ as a function of p for different k = 5, 15, 30 and n = 1, 5, 10, 15 with sensitivity and specificity fixed at 0.99 when the sensitivity and specificity of the tests are set at 99%.

5. Model Comparison

In this section, the efficiency of the constructed estimator relative to other estimators in pool testing scheme was determined by computing the ARE and the RMSE and the results are discussed.

5.1. Asymptotic Relative Efficiency (ARE)

The sample size based on the one-stage negative binomial group testing scheme that has considered misclassification was examined by [23]. The variance of the estimator was computed as follows:

$$\mathrm{var}(\hat{p}) = \frac{\left[\pi_1 - (\pi_0 + \pi_1 - 1)(1-p)^k\right]^2 \left[1 - \pi_1 + (\pi_0 + \pi_1 - 1)(1-p)^k\right]}{n(\pi_0 + \pi_1 - 1)^2 k^2 (1-p)^{2k-2}} \quad (9)$$

If the estimator of the one-stage negative binomial group testing scheme with misclassification is denoted by $\hat{p}_x$, and the computed estimator is denoted $\hat{p}_i$ then

$$\text{ARE} = \frac{\text{var}(\hat{p}_x)}{\text{var}(\hat{p}_i)} \tag{10}$$

Therefore, $\text{ARE} > 1$ implies that the proposed model is more efficient than the one-stage negative binomial group testing model with misclassification.

Table 6. ARE of the proposed model relative to the one-stage negative binomial group testing model with misclassification for $k = 5, 10, 15, 30, 50$ with sensitivity and specificity of the tests set at 99%, 98%, and 95%.

$n = 1$					
			k		
p	5	10	15	30	50
Sensitivity = Specificity = 0.99					
0.005	1.9581	1.4398	1.2841	1.1366	1.0793
0.01	1.4386	1.2089	1.1363	1.0648	1.0348
0.05	1.0774	1.0338	1.0163	0.9818	0.9158
0.10	1.0326	1.0030	0.9798	0.8567	0.6155
0.20	1.0004	0.9418	0.8352	0.5419	0.5131
0.30	0.9695	0.8085	0.6090	0.5133	0.5127
0.40	0.9167	0.6293	0.5242	0.5127	0.5127
Sensitivity = Specificity = 0.98					
0.005	3.2146	1.9585	1.6045	1.2831	1.1626
0.01	1.9557	1.4386	1.2824	1.1324	1.0710
0.05	1.1585	1.0690	1.0339	0.9686	0.8642
0.10	1.0665	1.0079	0.9651	0.7885	0.5832
0.20	1.0029	0.9025	0.7642	0.5409	0.5261
0.30	0.9473	0.7362	0.5792	0.5262	0.5259
0.40	0.8655	0.5919	0.5317	0.5259	0.5259
Sensitivity = Specificity = 0.95					
0.005	8.0954	3.8828	2.7484	1.7763	1.4347
0.01	3.8737	2.2375	1.7741	1.3518	1.1880
0.05	1.4235	1.1829	1.0940	0.9524	0.7990
0.10	1.1763	1.0330	0.9459	0.7232	0.5930
0.20	1.0218	0.8470	0.7031	0.5744	0.5683
0.30	0.9146	0.6818	0.5912	0.5683	0.5682
0.40	0.8005	0.5971	0.5706	0.5682	0.5682

Scrutiny of Table 6 shows that the sequential retesting of a pool that tests positive in the initial test infers that the two-stage negative binomial model is more efficient, especially at low prevalence. This is indicated by $\text{ARE} > 1$ It can also be observed that at a given prevalence of a rare trait, the efficiency of the proposed model decreases with an increase in group size. Similarly, for a fixed group

size, the efficiency of the model decreases with an increase in prevalence. Finally, it can be observed that even when the accuracy of the diagnostic test is relatively low, retesting improves the efficiency of the model in a low-prevalence population. Thus, in cases where group testing is used to screen for a rare trait in a low prevalence population, retesting of pools is more desirable as it yields more accurate results when the accuracy of the diagnostic tests is less than 100%.

5.2. Relative Mean Squared Error (RMSE)

To compare the efficiency of a model, a convenient way is to compare the MSE of estimates of the same p with other existing models obtained using different experimental procedures. The MSE of the proposed model is compared with the MSE of the one-stage negative binomial group testing model with misclassification that is denoted by $\hat{p}_d$. The RMSE is calculated as follows:

$$\text{RMSE} = \frac{\text{MSE}(\hat{p}_d)}{\text{MSE}(p_i)} \tag{11}$$

Table 7. Relative mean squared error of the estimator (RMSE) with sensitivity and specificity value set at 99%.

	n		
p	1	3	5
	$k = 5$		
0.005	38.565142	12733.370308	58400.557061
0.01	19.133470	2643.906573	13086.300135
0.05	3.443933	34.238002	154.682671
0.10	1.463359	3.042883	5.576623
0.20	0.525123	0.194483	0.192145
0.30	0.337283	0.241976	0.365225
	$k = 15$		
0.005	12.899748	1896.419895	39821.816914
0.01	6.205910	226.377645	4210.861195
0.05	0.882280	0.687266	0.538078
0.10	0.279830	0.036577	0.030166
0.20	0.105515	0.070938	0.079776
0.30	0.196267	0.195440	0.205301
	$k = 30$		
0.005	6.228922	238.248501	7177.123081
0.01	2.879419	23.761346	188.761043
0.05	0.287960	0.027176	0.008577
0.10	0.066378	0.014407	0.016307
0.20	0.073766	0.066298	0.069187
0.30	0.193811	0.193251	0.200459

Scrutiny of Table 7 shows that for small group sizes, the proposed model outperforms the one-stage negative binomial group testing model with misclassification when the prevalence is low. This is indicated by the value of $\text{RMSE} > 1$. For instance, when $n = 1$, and $k = 5$, the two-stage negative binomial group testing model is observed to be 38.6 times more efficient than the one-stage negative binomial group testing model at $p = 0.05$. In general, the proposed model is more efficient in a low-prevalence population when the size of the group is fairly small. Lastly, it can be observed that as the group size and the waiting parameter n increases, the efficiency of the model increases in a low-prevalence population.

Table 8. Relative mean squared error of the estimator (RMSE) at 95% sensitivity and specificity value.

	n		
p	1	3	5
	$k = 5$		
0.005	36.204290	9359.937964	39805.251986
0.01	18.897137	2201.840151	9836.147127
0.05	3.589811	35.559549	135.836994
0.10	1.561931	3.537761	6.190444
0.20	0.592202	0.265764	0.272134
0.30	0.395585	0.308493	0.524353
	$k = 15$		
0.005	13.034037	1879.883663	30942.258974
0.01	6.403645	242.737094	3812.958260
0.05	0.961991	0.887832	0.810215
0.10	0.336188	0.055905	0.046155
0.20	0.153077	0.089761	0.118998
0.30	0.242057	0.241927	0.299192
	$k = 30$		
0.005	6.429417	260.242107	7446.285644
0.01	3.022253	27.380735	231.694665
0.05	0.344612	0.044954	0.014607
0.10	0.113660	0.019012	0.024423
0.20	0.119842	0.083489	0.102518
0.30	0.240808	0.237495	0.285215

Table 8 shows that the two-stage negative binomial group testing model outperforms the one-stage negative binomial group testing model with misclassification when the prevalence of a rare trait is low. It can be observed that the effi-

ciency of our model increases with an increase in group size and the waiting parameter n, especially in low prevalence. This indicates that the asymptotic properties of the model are taking hold. Thus, it is possible to get a combination of k and n values where the model would be more efficient for a given prevalence. For instance, when $k=30$ and $n=5$, the model is approximately 7446 more efficient for $p=0.005$.

6. Application of the Model to Real Data

The testing scheme was also applied to an investigation that involved the surveillance of West Nile Virus (WNV) conducted in Jefferson County by [25] in Florida following the 2001 outbreak of WNV transmitted by the North American mosquito, *Culex nigripalpus*. The authors documented the first field study on the mosquito transmission rate of WNV which was used by [20] to illustrate their procedure. A total of 11948 mosquitos were captured in the surveillance program and tested in various pool sizes using reverse-transcription polymerase chain reaction assays. A total of 14 mosquito pools tested positive for WNV, and by the end of the outbreak, 12 human cases were reported to have West Nile *Meningoencephalitis* and 483 documented cases among the horses.

The investigation by [25] did not consider inverse sampling when imperfect tests are used. However, this investigation was used as a basis to illustrate our testing procedures when imperfect tests with known sensitivity and specificity values are used. Based on the field study by [25] that focused on pools that tested positive, the estimated prevalence was reported to be approximately 0.005. Hence it was presumed that 0.005 was the true value of the prevalence. We have simulated data sets to illustrate our procedure for the waiting parameters n = 1, 5, 10, and 15 at equal group sizes k = 5, 15, 30, and 50 similar to [25] when the diagnostic tests are 99%, 95%, and 90% accurate.

Table 1 presents the results of the MLE of the population proportion p based on 10,000 Monte Carlo data sets. The performance of the constructed estimator overestimated the population proportion p as the prevalence increased. The MLE of the constructed estimator reaffirms the discovery that group testing is useful in a low-prevalence population. When both the group size k and the waiting parameter n are large, close estimates of the MLE were observed at $p=0.05$ which is consistent with the estimated prevalence level reported by [25]. When the waiting parameter is small say $n=1$, the MLE was observed to be exorbitantly positively biased as the prevalence level increased as observed in Table 2 and Table 3. Alternative estimators have been developed by different scholars to reduce the bias when the waiting parameter n is small [7] [20]. Gart's bias correction to the MLE was recommended by [26] as an effective estimator in reducing the bias when perfect tests are used. A striking feature to note is that the MSE is sensitive to the group size and the waiting parameter n as shown in Figure 2.

To access the efficiency of the testing scheme, the performance of the con-

structed estimator was compared to other estimators by computing the ARE and RMSE. The results showed that in lower values of p, the two-stage negative binomial group testing model was superior to the one-stage negative binomial group testing model with misclassification suggested by [23]. This can be observed in **Table 6** where $ARE > 1$ and in **Table 7** and **Table 8** where $RMSE > 1$. Moreover, the proposed model was established to be efficient even in situations where the accuracy of the diagnostic tests are slightly low. Thus, retesting of positive pools improved the efficiency of the constructed estimator, which agrees with what other authors have already established [9] [12] [27].

7. Conclusion

A two-stage negative binomial group testing procedure for estimating the prevalence of a rare trait has been constructed and analyzed. From the discussions, the two-stage negative binomial group testing model is superior to the one-stage negative binomial group testing model with misclassification. The constructed estimator performed better at a low prevalence level. Also, it performed better in slightly low values of sensitivity and specificity (95%) than at higher values of the diagnostic tests used that are 99% accurate. Thus, we recommend the retesting of pools in negative binomial group testing designs when imperfect diagnostics tests with known sensitivity and sensitivity are used during estimation.

Conflicts of Interest

The authors declare no conflict of interest regarding the publication of this paper.

References

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

[2] Kennedy, N.L. (2004) Multistage Group Testing Procedure (Group Screening). *Communication in Statistics—Simulation and Computation*, **33**, 621-637. https://doi.org/10.1081/SAC-200033231

[3] Brookmeyer, R. (1999) Analysis of Multistage Pooling Studies of Biological Specimens for Estimating Disease Incidence and Prevalence. *Biometric*, **55**, 608-612. https://doi.org/10.1111/j.0006-341X.1999.00608.x

[4] Xie, M., Tatsuoka, K., Sacks, J. and Young, S. (2001) Pool Testing with Blockers and Synergism. *Journal of American Statistical Association*, **96**, 92-102. https://doi.org/10.1198/016214501750333009

[5] Wanyonyi, R.W., Nyongesa, K. and Wasike, A.A.M. (2015) Estimation of Proportion of a Trait by Batch Testing Model in a Quality Control Process. *American Journal of Theoretical and Applied Statistics*, **4**, 619-629.. https://doi.org/10.11648/j.ajtas.20150406.34

[6] Gastwirth, J.L. and Hammick, P.A. (1989) Estimation of the Prevalence of a Rare Disease, Preserving the Anonymity of the Subjects by Group Testing: Application to Estimating the Prevalence of AIDS Antibodies in Blood Donors. *Journal of Statis-*

tical Planning and Inference, **22**, 15-27. https://doi.org/10.1016/0378-3758(89)90061-X

[7] Hepworth, G. (2013) Improved Estimation of Proportions Using Inverse Binomial Group Testing. *Journal of Agricultural, Biological, and Environmental Statistics*, **18**, 102-119. https://doi.org/10.1007/s13253-012-0126-6

[8] Thompson, K.H. (1962) Estimation of the Proportion of Vectors in a Natural Population of Insects. *Biometrics*, **18**, 568-578. https://doi.org/10.2307/2527902

[9] Nyongesa, L.K. (2018) Multiple-Test Pool-Testing Strategy for Estimating HIV/AIDS-Prevalence and Its Extension to Multi-Stage. *Annals of Reviews and Research*, **2**, 58-68. https://doi.org/10.19080/ARR.2018.02.555589

[10] Okoth, A.W., Nyongesa, L.K. and Kwatch, B.O. (2017a) A Comparative Study between a Multi-Stage Adaptive Pool-Testing Model with Test Errors and the Non-Adaptive Model. *International Journal of Applied Mathematical Research*, **6**, 93-97. https://doi.org/10.14419/ijamr.v6i3.7802

[11] Okoth, A.W., Nyongesa, L.K. and Kwatch, B.O. (2017) A Multi-Stage Adaptive Pool-Testing Model with Test Errors; Improved Efficiently. *Journal of Mathematics*, **13**, 43-55.

[12] Wanyonyi, R.W., Mwangi, O.W. and Mwangi, C.W. (2021) Re-Testing in Batch Testing Model Based on Quality Control Process for Proportion Estimation. *Open Journal of Statistics*, **11**, 123-136. https://doi.org/10.4236/ojs.2021.111007

[13] Tamba C.L., Nyongesa K.L. and Mwangi, J.W. (2012) Computational Pool-Testing with Retesting Strategy. *International Journal of Scientific and Statistical Computing* (*IJSSC*), **3**, 47-54.

[14] Montesinos López, O.A., Montesinos López, A., Crossa, J., Eskridge, K. and Hernández Suárez, C.M. (2010) Sample Size for Detecting and Estimating the Proportion of Transgenic Plants with Narrow Confidence Intervals. *Seed Science Research*, **20**, 123-136. https://doi.org/10.1017/S096025851000005X

[15] Montesinos López, O.A., Montesinos López, A., Crossa, J., Eskridge, K. and Sáenz, R.A. (2011) Optimal Sample Size for Estimating the Proportion of Transgenic Plants Using the Dorfman Model with a Random Confidence Interval. *Seed Science Research*, **21**, 235-246. https://doi.org/10.1017/S0960258511000055

[16] Haldane, J.B.S. (1945) On a Method of Estimating Frequencies. *Biometrika*, **33**, 222-225. https://doi.org/10.1093/biomet/33.3.222

[17] 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

[18] Foppa, I.M., Evans, C.L., Wozniak, A. and Wills, W. (2007) Mosquito Fauna and Arbovirus Surveillance in a Coastal Mississippi Community after Hurricane Katrina. *Journal of the American Mosquito Control Association*, **23**, 229-232. https://doi.org/10.2987/8756-971X(2007)23[229:MFAASI]2.0.CO;2

[19] Rodríguez-Pérez, M.A., Katholi, C.R., Hassan, H.K. and Unnasch, T.R. (2006) Large-Scale Entomologic Assessment of *Onchocerca volvulus* Transmission by Poolscreen PCR in Mexico. *The American Journal of Tropical Medicine and Hygiene*, **74**, 1026-1033. https://doi.org/10.4269/ajtmh.2006.74.1026

[20] Pritchard, N.A. and Tebbs, J.M. (2011) 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

[21] Montesinos-López, O.A., Montesinos-López, A., Crossa, J. and Eskridge, K. (2012) Sample Size under Inverse Negative Binomial Group Testing for Accuracy in Parameter Estimation. *PLOS ONE*, **7**, e32250. https://doi.org/10.1371/journal.pone.0032250

[22] Montesinos-López, O.A., Montesinos-Lopez, A., Crossa, J. and Eskridge, K. (2013) Sample Size for Detecting Transgenic Plants Using Inverse Binomial Group Testing with Dilution Effect. *Seed Science Research*, **23**, 279-288. https://doi.org/10.1017/S0960258513000238

[23] Xiong, W. (2016) 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

[24] Haber, G., Malinovsky, Y. and Albert, P.S. (2018) Sequential Estimation in the Group Testing Problem. *Sequential Analysis*, **37**, 1-17. https://doi.org/10.1080/07474946.2017.1394716

[25] Rutledge, C.R., Day, J.F., Lord, C.C., Stark, L.M. and Tabachnick, W.J. (2003) West Nile Virus Infection Rates in *Culex nigripalpus* (Diptera: Culicidae) Do Not Reflect Transmission Rates in Florida. *Journal of Medical Entomology*, **40**, 253-258. https://doi.org/10.1603/0022-2585-40.3.253

[26] Hepworth, G. and Watson, R. (2009) Debiased Estimation of Proportions in Group Testing. *Journal of the Royal Statistical Society Series C: Applied Statistics*, **58**, 105-121. https://doi.org/10.1111/j.1467-9876.2008.00639.x

[27] Kennedy, N.L. (2011) Dual Estimation of Prevalence and Disease Incidence in Pool Testing Strategy. *Communication in Statistics—Theory and Method*, **40**, 3218-3229. https://doi.org/10.1080/03610926.2010.493257

Appendix

The log-likelihood function to base 10 is

$$\ell(p) = n\log\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right) + (t-n)\log\left(1-\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)\right)$$

The first-order derivative of the log maximum Likelihood is

$$\frac{\partial \ell}{\partial p} = \frac{k}{1-p}\left[n\left(\frac{\pi_1^2}{\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)} - 1\right) - (t-n)\left(1 - \frac{1-\pi_1^2}{1-\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)}\right)\right]$$

The second-order derivative is

$$\frac{\partial^2 \ell}{\partial p^2} = \frac{k}{(1-p)^2}\left[n\left(\frac{\pi_1^2}{\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)} - 1\right) - (t-n)\left(1 - \frac{1-\pi_1^2}{1-\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)}\right)\right] + \frac{k}{1-p}\left[n\left(\frac{n\pi_1^2 k(1-p)^{k-1}\left((1-\pi_0)^2 - \pi_1^2\right)}{\left[\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right]^2}\right) - (t-n)\left(\frac{\left(1-\pi_1^2\right)k(1-p)^{k-1}\left((1-\pi_0)^2 - \pi_1^2\right)}{\left[1-\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)\right]^2}\right)\right]$$

Based on the testing process, T follows a negative binomial distribution with parameters n and $\pi^*(p) = \pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)$. Therefore, the expectation of T is $n\Big/\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)$, then

$$E(t-n) = n\left(\frac{1-\pi}{\pi}\right) = n\left(\frac{1-\pi_1^2 - (1-p)^k\left(\pi_1^2 - (1-\pi_0)^2\right)}{\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)}\right)$$

Since $\operatorname{var}(\hat{p}) = \left\{-E\left(\frac{\partial^2 \ell}{\partial p^2}\right)\right\}^{-1}$, we have

$$\operatorname{var}(\hat{p}) = \left[\frac{1-\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)}{nk^2(1-p)^{2k-2}\left((1-\pi_0)^2 - \pi_1^2\right)^2}\right]\left(\pi_1^2 + (1-p)^k\left((1-\pi_0)^2 - \pi_1^2\right)\right)^2.$$