The Korean Communications
in Statistics Vol. 6, No. 1, 1999
pp. 293-303

UMP Unbiased Test for the Infection Rate in Group Testing

Sehyug Kwon[1)]

Abstract

When test outcomes of units are dichotomous and the infection rate is small, group testing is more efficient than one-to-one testing in estimating the true p and classifying units as infected or not. In this paper, two-sided hypothesis testing and confidence intervals are derived based on the UMP(uniformly most powerful) unbiased test. The UMP unbiased approach is compared with Thompson's and Bhattacharyya et al.'s approaches by computing the length of confidence intervals and capture probabilities and shown to have a number of desirable properties. Unequal allocation, one of advantages of the proposed approach, is also mentioned.

1. Introduction

Instead of one-by-one testing, group testing was suggested for classification and estimation when the outcome of test units is dichotomous, meaning units will be classified as infected or not, and the infection rate is small (Dorfman, 1943 and Ungar, 1960). In group testing, groups are formed by units of size k and tested simultaneously. When the test outcome is not infected, it is concluded that there is no infected unit in the group. If infected, the units in the group are regrouped and retested until all units are classified as infected or not. Therefore, experimental design for group testing has focused on the choice of group size (Thompson, 1962 and Swallow, 1985) and the retesting scheme (Dorfman, 1943, Sobel and Groll, 1959, Finucan, 1964, Hwang, 1974, and Chen and Swallow, 1990).

There are two major interests in group testing: (1) the estimation problem to estimate the infection rate and (2) the classification problem to classify each unit in a sample of interest as infected or not. The estimation problem, especially hypothesis testing and confidence intervals, has not been widely studied as classification problem. Thompson (1962) suggested the MLE(Maximum Likelihood Estimator) and an approximate confidence interval for the population infection rate, based on the exact variance and Student's approach. Bhattacharyya, Karandinos and DeFoliart (1979) developed a method for both hypothesis testing and confidence interval using the asymptotic normality of the MLE when the number of groups is large. Kwon (1997) pointed out the problems of both approaches, developed one-sided hypothesis testing and the upper bound for the infection rate based on the UMPT(uniformly

1) Assistant Professor, Department of Applied Statistics, Hannam University, Taejon 306-791, Korea

most powerful test). In this paper, two-sided hypothesis testing and confidence interval are derived based the UMP(uniformly most powerful) unbiased test and compared with two classical approaches by computing the length of confidence intervals and capture probabilities since there in one-to-one correspondence between hypothesis testing and confidence intervals.

2. The MLE

The following assumptions are made: (1) The units are randomly distributed as Bernoulli with the true infection rate. It is fundamental in group testing. (2) The number of groups is fixed and the group size for each of the groups is the same for simplicity. Group size can usually be controlled in practice (3) Units are not to be retested, which is made on the grounds that retesting would only negligibly reduce the mean squared error of the estimator as shown in Chen and Swallow (1990). (4) Classification (testing) of groups is without error. It implies that there are no false negative and false positive effects.

Suppose that the population infection rate is p, the optimal group size is k and the number of group of size k to be tested is n. Let X_i be the test outcome for the i-th group. The probability distribution function of X_i is

$$f(x_i)=[1-(1-p)^k]^{x_i}[(1-p)^k]^{1-x_i} \quad \text{where } x_i = 0(\text{not infected}),\ 1(\text{infected}).$$

The MLE of p is $\hat{p}=1-(1-\sum_{i=1}^{n}X_i/n)^{1/k}$. For computing the $\hat{p}$, the group size k should be obtained in advance. Thompson (1962) gave the integer value of (1.5932/p) as optimal group size and Swallow (1985) constructed a table of optimal group sizes for combinations of the (n, p). Both methods required the knowledge of the true p that is unknown in practice choosing the optimal value. Sometimes, there is at least some prior information about the value of p from preliminary data or other considerations. At least an upper bound on p, not lower bound, may be available and, if so, it has been suggested that it should be used in choosing k (Swallow, 1985).

3. Classical approaches

Thompson showed that the MLE of p is distributed asymptotically normally and converges in probability to p as n goes to infinity and suggested the following approximate confidence interval for p.

$$\hat{p} \pm t_{\frac{\alpha}{2},\, n-1}\left(\sum_{i=0}^{n}\left\{(1-\hat{p})-\left(\frac{i}{n}\right)^{1/k}\right\}^2\binom{n}{i}\{(1-\hat{p})^k\}^i\{1-(1-\hat{p})^k\}^{n-i}\right)^{1/2}.$$

He did not consider hypothesis testing for H_o: $p=p_o$. But, since there is a one-to-one between the confidence interval and the hypothesis testing, the two-sided hypothesis can be tested by Thompson's approach. Given a sample, the set of all values of p that would not rejected in testing the null hypothesis H_o: $p=p_o$ at significant level α would be the 100(1-α)% confidence interval for p (Lehmann, 1959).

Bhattacharyya et al. (1979) developed a confidence interval for p using the central limit theorem and a test statistics for testing H_o: $p=p_o$ as follows:

$$\hat{p} \pm z_{\frac{\alpha}{2}}\left[\frac{1}{k}\left\{\frac{r}{n(n-r)}\right\}^{1/2}\left(\frac{n-r}{n}\right)^{1/k}\right] \text{ and}$$

$$\frac{\hat{p}}{s(\hat{p})} \simeq \text{Normal } (0,\ 1) \text{ where } s(\hat{p}) = \left[\frac{1}{k}(1-\hat{p})^2\left[\frac{1}{(1-\hat{p})^2}-1\right]\right]^{1/2}.$$

Both approaches give symmetric confidence intervals for the true p despite the fact that the small-sample distribution of the MLE of p is asymmetric, especially so when p is either very small or close to 1, and their confidence intervals may contain values which lie out of the parameter space, [0, 1].

4. UMP Unbiased Test

Since $\sum X_i$ is distributed as a binomial random variable with (n, $1-(1-p)^k$) and is the sufficient statistic for ($1-(1-p)^k$), it is then considered as a test statistic. The likelihood ratio test statistic for testing H_0: $p \subseteq \theta_0$ vs H_a: $p \subseteq \theta_a$ is

$$\lambda = \frac{[1-(1-p_0)^k]^{\sum X_i}[(1-p_0)^k]^{(n-\sum X_i)}}{[\sum X_i/n]^{\sum X_i}[1-\sum X_i/n]^{(n-\sum X_i)}}$$

Because the likelihood ratio test statistic has the $\sum X_i$ term not only in a base but also in the power (exponent), it is complicated to deal with. Even taking logarithms fails to yield expressions that nicely determine the rejection region.

Consider two separate tests, H_{o1}: {p; $p \le p_o$} vs H_{a1}: {p; $p > p_o$} and H_{o2}: {p; $p \ge p_o$} vs H_{a2}: {p; $p < p_o$}, and combine their results to obtain a two-sided confidence interval. Kwon (1997) showed that for each of these separate tests, the UMPT exists.

Let p^* be $1-(1-p)^k$. $\sum X_i$ is the optimal test statistic for testing H_o: $p^* \le p_o$ vs H_a: $p^* > p_o$ through choosing t and γ in the following test function to satisfy $E_{p_0}[\Phi(\sum X_i)] = \alpha$, where α is the size of the test, which is here called Test 1. The test function is

$$\Phi(x) = \begin{cases} 1 \text{ when } \sum X_i > t \\ \gamma \text{ when } \sum X_i = t \\ 0 \text{ when } \sum X_i < t \end{cases}$$

Similarly, the UMPT for H_o: $p^* \geq p_o$ vs H_a: $p^* < p_o$ can be obtained, which is called Test 2. But, neither Test 1 nor Test 2 yields a UMPT for testing H_o: {p; $p=p_o$} vs H_a: {p; $p \neq p_o$}. For testing H_{o2}, the power of Test 1 is less than that of Test 2, because Test 2 is the UMPT for testing H_{o2}. Likewise, Test 2 has less power than Test 1 for H_{o1}. Although there is no UMPT for two-sided hypothesis testing in group testing, a UMP unbiased test does exists because the distribution of X_i is in the one-parameter exponential family. Lehmann (1959) showed the following:

If the pdf of x is a one-parameter exponential family and T(x) is the sufficient statistic, there exists a UMP unbiased test for testing H_o: {p; $p=p_0$} against the alternative, H_a: {p; $p \neq p_0$} given by

$$\Phi(x) = \begin{cases} 1 \text{ when } T(x) < C_1 \text{ or } T(x) > C_2 \\ \gamma_i \text{ when } T(x) = C_i,\ i=1,2 \\ 0 \text{ when } C_1 < T(x) < C_2 \end{cases}$$

where C_1, C_2, γ_1, and γ_2 are determined by $E_{to}[\psi(x)] = \alpha$ and $E_{to}[T(x)\psi(x)] = E_{to}[T(x)]\alpha$. $\sum X_i$ (=X), the sufficient statistic, can be considered as T(x). $\sum X_i$ is distributed Binomial (n, $1-(1-p_o)^k$) under the null hypothesis. Let p^* be $1-(1-p_o)^k$. Then $E_{to}[\psi(x)] = \alpha$ becomes

$$\sum_{x=C_1+1}^{C_2-1} \binom{n}{x} (p^*)^x (1-p^*)^{n-x} + \sum_{i=1}^{2} (1-\gamma_i) \binom{n}{C_i} (p^*)^{C_i} (1-p^*)^{n-C_i} = 1-\alpha \quad ---(1)$$

and $E_{to}[T(x)\psi(x)] = E_{to}[T(x)]\alpha$ with the help of the identity

$$x\binom{n}{x}(p^*)^x(1-p^*)^{n-x} = np^*\binom{n-1}{x-1}(p^*)^{x-1}(1-p)^{(n-1)-(x-1)}$$

reduces to

$$np^* \left\{ \sum_{x=C_1+1}^{C_2-1} \binom{n-1}{x-1} (p^*)^{x-1} (1-p^*)^{(n-1)-(x-1)} + \sum_{i=1}^{2} (1-\gamma_i) \binom{n-1}{C_i-1} (p^*)^{C_i-1} (1-p^*)^{(n-1)-(C_i-1)} \right\}$$

$$= E_{p^*}[T(x)](1-\alpha) \quad ---(2)$$

The C_i and γ_i are then chosen to satisfy (1) and (2) to obtain the UMP unbiased test for H_o:$\{1-(1-p)^k = 1-(1-p_o)^k\}$ versus H_a:$\{1-(1-p)^k \neq 1-(1-p_o)^k\}$. Because there is one-to-one correspondence between p_o and p^* is a UMP unbiased test for testing H_o: $p=p_o$ vs H_a: $p \neq p_o$.

5. Comparisons

The three approaches are compared by computing the length of confidence intervals and capture probabilities for various combinations of (n, p). Since there is one-to-one correspondence between hypothesis testing and confidence intervals, two-sided confidence intervals based on the UMP unbiased test can be obtained. Two-sided confidence intervals are obtained numerically because the UMP unbiased test requires randomization. For convenience, the upper and lower bounds are obtained by allocating exactly $\alpha/2$ to each tail. For comparison, combinations of (n, p) used are the following:

n=10, 20, 30, 50

p=0.01, 0.02, 0.05, 0.1, 0.2, 0.3

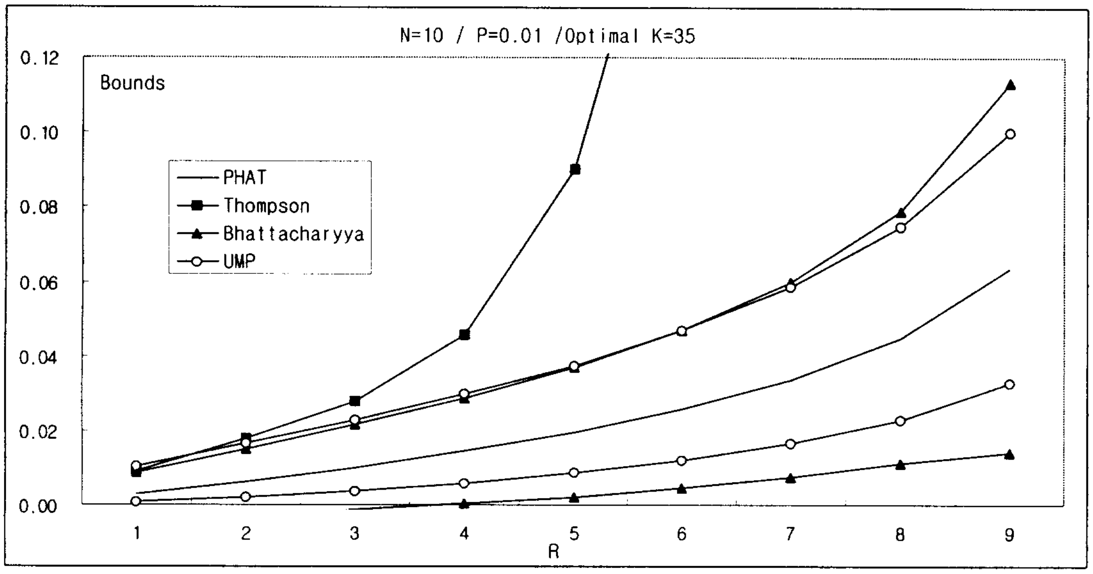
Figure 1 shows the MLE and 95% confidence intervals when n=10 with p=0.01, 0.02, 0.05, and 0.1. The results shown are "best cases" in the sense that the optimal group size that given in Swallow (1995), which depends on n and p, is used in each case. The value shown as r is the number of infected groups found among the n tested. The horizontal lines mark the true p. Figure 2 displays the MLE and 95% confidence intervals when p=0.05 with n=10, 20, 30, and 50. On a relative basis, as n increases with p fixed (Figure 2), the difference between the Bhattacharyya et al. and the UMP unbiased becomes less. Therefore, when circumstances dictate using a rather small n in an application, the advantage of the UMPU approach is greater. As p increases with n fixed (Figure 1), the difference between Bhattacharyya et al.'s and the UMPU approach changes little on a relative basis. Thompson's is never close to them.

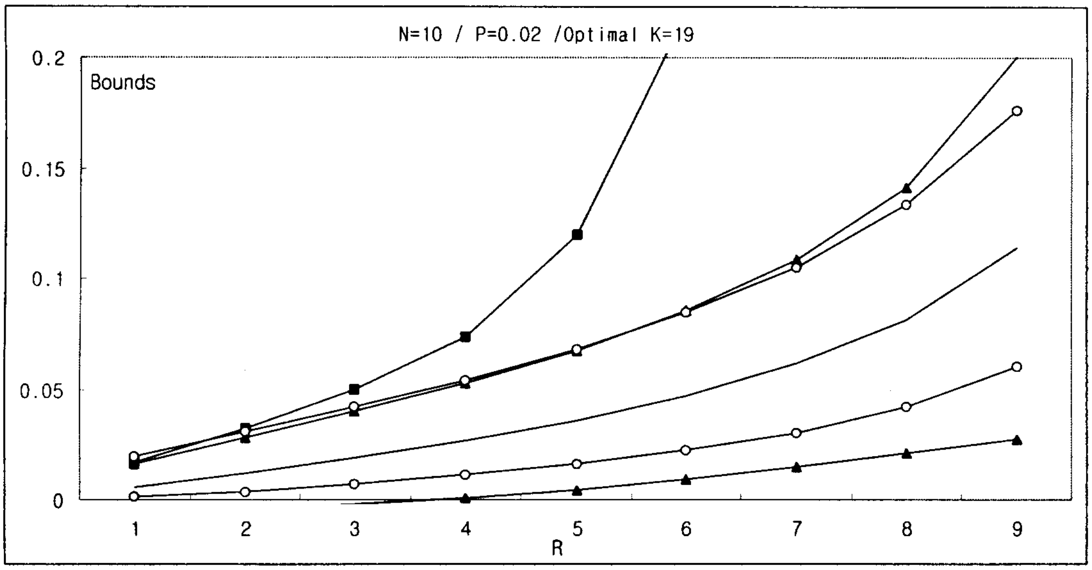
As is suggested in Figures 1 and 2, Bhattacharyya et al.s' and Thompson's approaches can produce a negative value for the lower bound. Moreover, as r is increasing, the length of Thompson's confidence interval becomes larger and larger relative to the other two. When r=9 and p=0.01, and still P(r=9)>0, the Thompson 95% confidence interval spans the entire parameter space. Thompson's approach is the worst of the three on various counts.

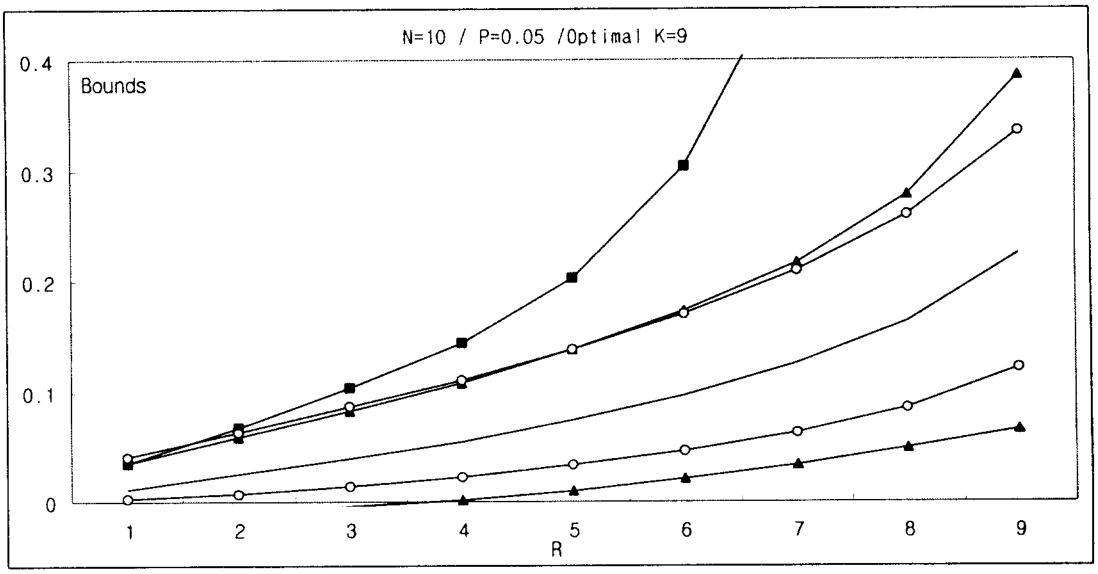
It appears that the length of confidence interval by the UMP unbiased is shorter than that of Bhattacharyya et al. except when r=1. When r=1, Bhattacharyya et al.'s approach is the best. But the lower bound of Bhattacharyya et al.'s confidence interval can be negative (perhaps truncated at 0 in practice) and the unequal allocation can get the shorter one in the UMP unbiased approach. For example, when n is 10 and p varies from 0.01 to 0.05, the value of $\{1-(1-p)^k\}$ is around 0.36. The underlying distribution which is distributed as Binomial(10, 0.36), is skewed to the left. Allocating more than $\alpha/2$ to the longer tail and less than $\alpha/2$ to the shorter tail yields a $(1-\alpha)\%$ confidence interval that has shorter length than that obtained by equal allocation. When p is very small and n is small, unequal allocation may be worthwhile.

For example, suppose n=10 and p=0.01. The recommended group size is 35 from Swallow's (1985). If r=1 is observed, Bhattacharyya et al.'s confidence interval is (-0.0029, 0.0089). The confidence interval given for the UMP unbiased approach is (0, 0.0105). But Bhattacharyya et al.'s approach is cheating a bit in acting as though negative values of the parameter are conceivable, and ignoring the asymmetry of the underlying distribution. If the UMP unbiased approach allocated all 5% to the upper tail, it would give an interval (0, 0.0085).

Figure 1. The MLE and 95% Two-sided Confidence Intervals: n=10






N=10 / P=0.05 /Optimal K=9
Bounds
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
R

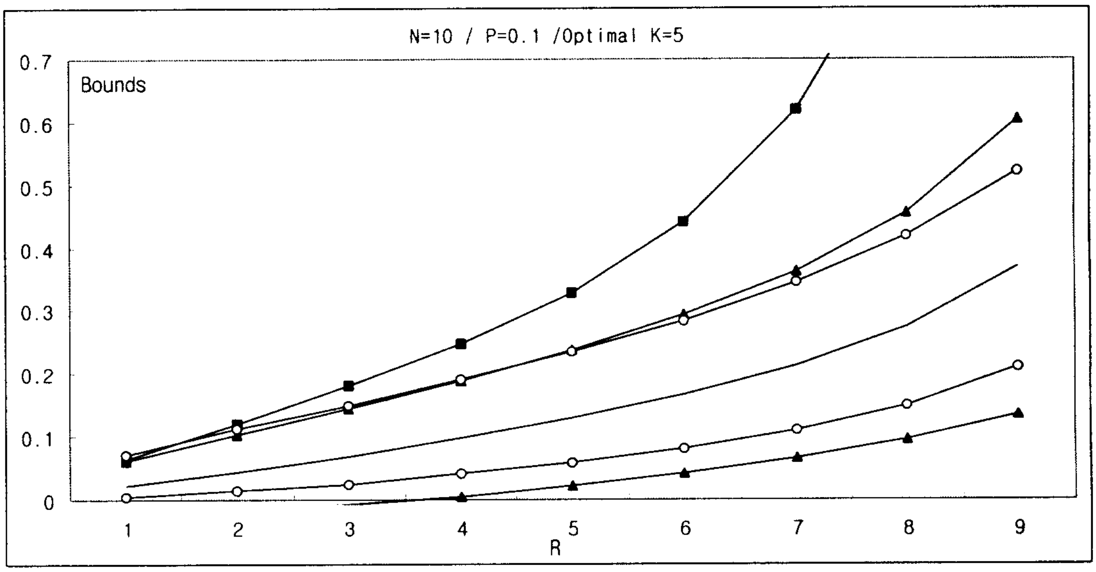
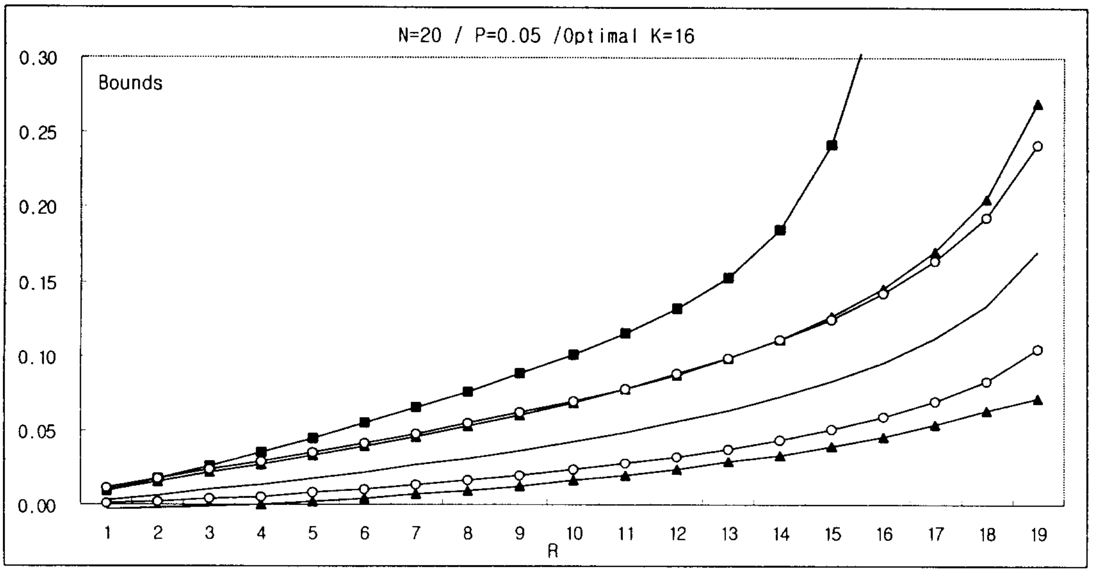
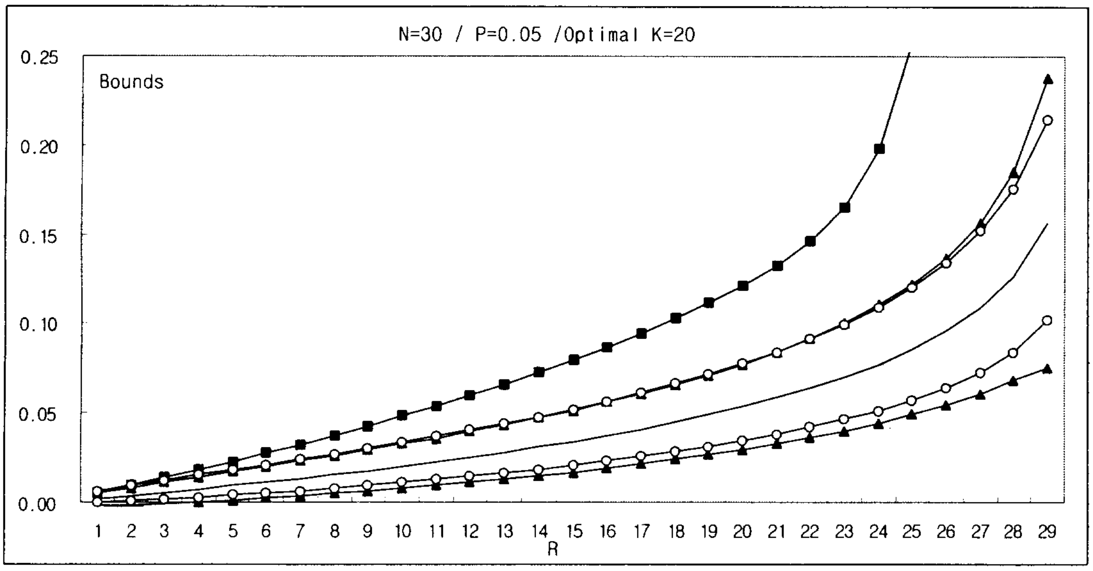
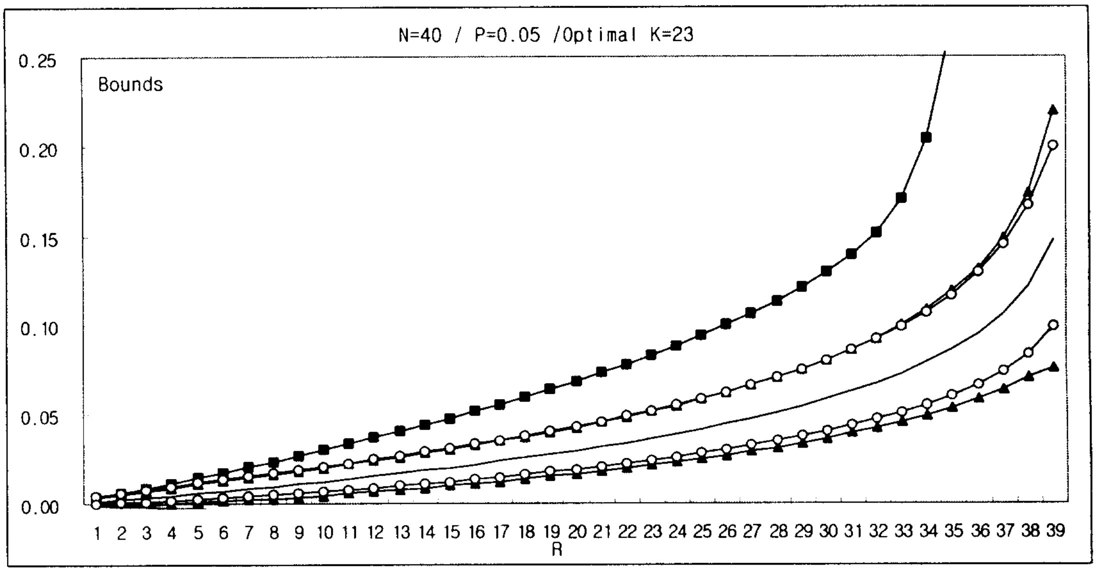
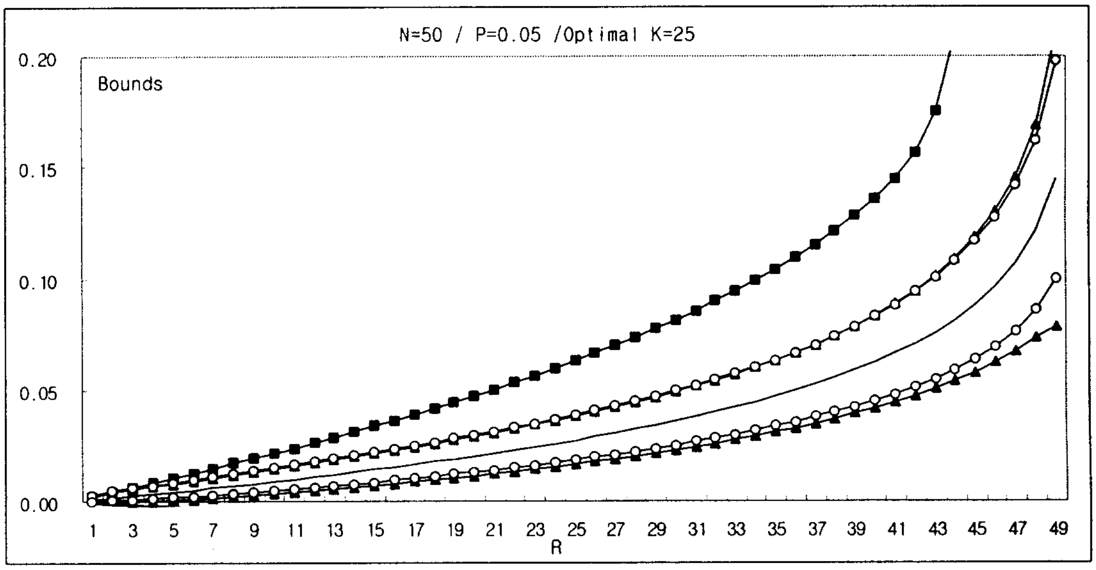

N=10 / P=0.1 /Optimal K=5
Bounds
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
4
5
6
7
8
9
R

Figure 2. The MLE and 95% Two-sided Confidence Intervals: p=0.05






N=40 / P=0.05 /Optimal K=23
Bounds
0.25
0.20
0.15
0.10
0.05
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
R


N=50 / P=0.05 /Optimal K=25
Bounds
0.20
0.15
0.10
0.05
0.00
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49
R

When p is unknown, a guessed p, perhaps an upper bound, from prior information must be employed in deciding what group size to use. Therefore, UMP unbiased and Bhattacharyya et al.'s approaches can be compared in the sense of probabilities of the intervals' failure to capture the true p. The difference in capture probabilities are computed and summarized in Table 1 for the true p=0.01 and 0.02 in the UMP unbiased and Bhattacharyya et al.'s approaches. Non-zero values indicate excess capture failures by Bhattacharyya et al.'s approach. The guessed p is the value in choosing the optimal group size, k. In Table 1, the column r_0 gives those values of r for which the upper bound of the UMP unbiased approach includes the true p but Bhattacharyya's does not. When n=10 and p=0.01 or 0.02, there are no differences in capture probabilities. However, differences are observed when p=0.01 for n=20 and 30. Bhattacharyya et al.'s confidence interval has greater probability of not including the true value of the parameter when r is small.

Table 1. Difference of capture probabilities between the Bhattacharyya et al. and the UMP unbiased approaches

(The true p=0.01)

n	Guessed p	k	r_0	Difference
20	0.03	35	3	0.0752
	0.04	25		0
	0.05	19		0
30	0.03	45	4	0.0496
	0.04	30		0
	0.05	25	5	0.0504

(The true p=0.02)

n	Guessed p	k	r_0	Difference
20	0.03	25		0
	0.04	19	3	0.2331
	0.05	16		0
30	0.03	30	9	0.1396
	0.04	25	7	0.1682
	0.05	20	6	0.1739

6. Conclusion and Discussion

The UMP unbiased test for two-sided hypothesis testing was derived and compared with two classical approaches. From the comparison made in the length of confidence intervals and capture probabilities for the various combination of (n, p), the UMP unbiased approach to hypothesis testing and interval estimation was the best. Only equal allocation of α to the two sides of two-sided confidence intervals was computed for comparisons. But, it was shown that unequal allocation can produce short confidence intervals through allocation more that $\alpha/2$ to the longer tail. The smaller p is, where group testing more efficient in estimation and classification than one-to-one testing, the more one can gain.

References

[1] Bhattacharyya, G. K. Karandinos, M. G. and DeFoliart G. R. (1979), Point estimates and confidence intervals for infection rates using pooled organisms in epidemiological studies, *American Journal of Epistemology* 109, 124-131

[2] Chen, C. L. and Swallow, W. H (1990), Using group testing to estimate a proportion, and to test the binomial model, *Biometrics* 46, 1035-1046

[3] Dorfman, R. (1943), The detection of defective members of large population, *Annals of Mathematical Statistics* 14, 436-440

[4] Finucan, H. M. (1964), The blood testing problem, *Applied Statistics* 13, 43-50

[5] Hwang, F. K. (1974), On finding a single defective in binomial group testing, *Journal of the American Statistical Association* 69, 146-150

[6] Kwon, S. (1997), Upper Bounds for the Infection Rate in Group Testing, *The Korean Communications in Statistics*, Vol. 4, No 1, 317-325

[7] Lehmann, E. L. (1959), Testing Statistical Hypothesis, John Wiley and Sons, New York

[8] Sobel, M. and Groll, P. A. (1959), Group testing to eliminate efficiently all defectives in a binomial sample, *The Bell System Technical Journal* 38. 1179-1252

[9] Sobel, M. and Elashoff, R. M. (1975), Group testing with a new goal, estimation, *Biometrika* 62, 181-193

[10] Swallow, W. H. (1985), Group testing for estimation infection rates and probabilities of diseases transmission, *Phytopathology* 75, 1376-1381

[11] Thompson, K. H. (1962), Estimation of the proportion of vectors in a natural population of insects, *Biometrics* 18, 568-578

[12] Ungar, P. (1960), The cutoff point for group testing, *Communications on pure and Applied Mathematics* 13, 49-54

[13] Wilks, S. D. (1962), *Mathematical Statistics.* John Wiley and Sons, New York