The Economic and Social Review, Vol. 27, No. 2, January, 1996, pp. 119-136

Estimation of the Parameter in the Discrete "Taxi" Problem, With and Without Replacement

ANN LARGEY
and
JOHN E. SPENCER
The Queen's University of Belfast

Abstract: In the continuous uniform distribution [O,N], the Maximum Likelihood estimator of N is known to possess high mean square error for large samples. This paper examines this issue in the discrete case, without replacement (extending the work of Tenenbein) and with replacement. Various other estimators including the Minimum Variance Unbiased estimator and Geary's closest estimator are compared in the continuous and two discrete cases. Recommendations are made for choice of estimator in each case, depending on the sample size and on imprecise information on N.

I INTRODUCTION

Given a sample of n observations from a uniform distribution with unknown upper limit, N, can we determine the best estimate of N? The discussion of this problem can be traced back at least as far as the writings of Laplace on probability (1812-1814). There he addressed the question of whether it was possible to estimate N from a random drawing from an urn containing balls numbered 1 up to N. Since then the same issue has been referred to in several writings as the "tramcar" (Jeffreys, 1939), Schrödinger (Geary, 1944), "locomotive" (Mosteller, 1965), "taxi" (Noether, 1971, Kotz and Johnson, 1985) "racing car" (Tenenbein, 1971), or "horse racing" (Rosenberg and Deely, 1976) problem. The scenario invoking these names is the problem faced by a spectator of estimating the total number, N, of trams, taxis, etc., in the town/on the race-course from an observation of n members of the set, knowing that members in the population are numbered consecutively 1,...,N.

The settings for these examples all clearly involve discrete uniform distributions. However, much work on estimation of the upper limit of the uniform distribution has centred around the continuous case, with some references to the discrete case being easily approximated by the continuous case in the instance of large N (e.g., Rao, 1981, Geary, 1944) and often those papers which deal specifically with the problem in the discrete context are constrained to the situation where the population is sampled without replacement (e.g., Noether, 1971, Tenenbein, 1971, Rosenberg and Deely, 1976).

In this paper we will review the results from estimation in the continuous case, but the main emphasis will be on devising and comparing estimators for N in the discrete case with and without replacement. This will extend the results in Spencer and Largey (1993), especially in the with replacement case. Throughout we consider estimation using classical, non-Bayesian techniques, though the "taxi" problem can be treated using a Bayesian approach. For an example see Rosenberg and Deely (ibid), where Bayesian and empirical Bayesian estimators for the zero-one, linear and quadratic loss functions are applied to the discrete case without replacement.

II REVIEW OF THE CONTINUOUS CASE

In the context of the continuous uniform distribution U[0,N], with N unknown, let $x_1, x_2, ..., x_n$ represent the observed sample obtained from n independent drawings from the population, and set $\max[x_1, x_2, ..., x_n] = w$, i.e., w represents the largest value in the sample. Properties of this distribution and the performance of various estimators of N are examined in Spencer and Largey (1993). Estimators considered are:

W = maximum sample value, the maximum likelihood estimator (MLE)

U = (n+1)W/n, the unique minimum variance unbiased estimator (MVUE) (Davis, 1951; Wasan, 1970; Rohatgi, 1976)

J = (n+2)W/(n+1), the minimum mean square error (MSE) estimator (Johnson, 1950)

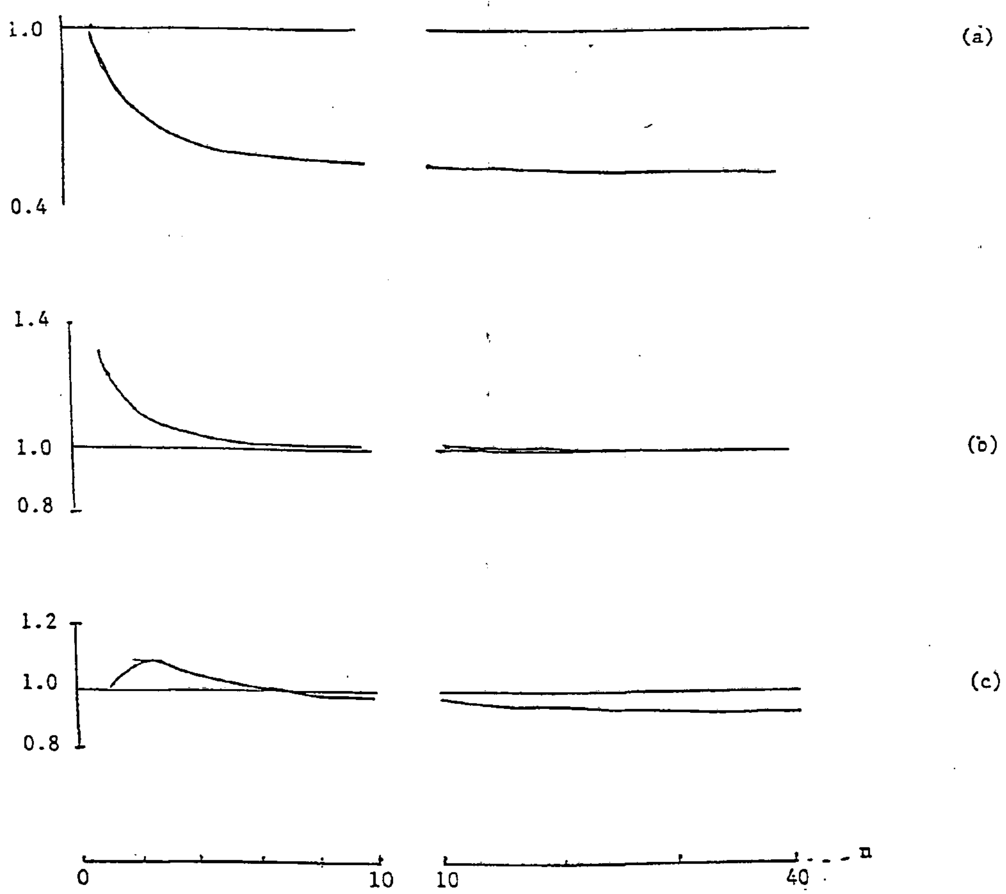
G = $2^{1/n}$W, Geary's estimator, the "closest" estimator following Pitman's definition of closeness (Geary, 1944).

The performance of each is assessed through comparison of MSE = Variance + Bias^2, as advocated by Johnson (1950), and results are summarised in Figure 1.

The most striking result from the analysis is the poor performance of W as an estimator. MSE of W is greater than both MSE of G and MSE of U for $n > 1$. In the limit as $n \to \infty$, while the MSE ratios for G and U relative to J tend to 1.09 and 1 respectively, that for W tends to 2.

For small and moderate n, G performs well on the basis of MSE, while for larger n, U outperforms it.

Figure 1: *Relative Mean Square Errors for Selected Pairs of Estimators of N: Continuous Case*



(a) MSE U/MSE W = $(n+1)/2n \rightarrow 0.5$

(b) MSE U/MSE J = $(n+1)^2/n(n+2) \rightarrow 1.0$

(c) MSE U/MSE G $\rightarrow (1+(\ln\ 2-1)^2)^{-1} = 0.913944$

The poor performance of the MLE is perhaps not surprising when we realise that the likelihood function does not adhere to the standard regularity conditions which are sufficient to ensure the desirable asymptotic properties of MLEs. For example the range of possible values of x_i depends on the upper limit of the distribution, N, which we are attempting to estimate. Moreover, the likelihood function, exhibits a discontinuity at the point w, and is therefore not differentiable at that point (see Spencer and Largey, op. cit., footnote 4).

III DISCRETE CASE

Turning to the case of the discrete uniform distribution, we have two separate situations to consider, i.e., sampling with and without replacement. Patel (1973) provides a summary of these, and the continuous case, focusing on completeness, sufficiency and minimum variance unbiased estimation. (Guenther (1978) discusses some techniques for finding MVU estimators which are usable in the taxi problem.) The ML, MVU, Geary and MMSE estimators will be considered in each situation. (Estimators based on the mean and median are highly inefficient and as a result will not be considered.) To order our findings we proceed by comparing pairs of estimators for the two discrete situations. (References to continuous case results will be made for clarification or where comparisons are of interest.)

Numerical results for all estimators (i.e., values of MSE for various N, n) are presented in the Appendix to Table 1 (sampling without replacement case) and Table 2 (sampling with replacement).

In both discrete cases, f(X;N) = 1/N, X = 1, 2, ..., N where N is the unknown positive integer to be estimated.

(1) MLE vs MVUE

(a) Sampling Without Replacement

The joint pdf of $X_1, X_2, ..., X_n$ is:

$$f(X_1,X_2,\ldots,X_n) = 1/N(N-1)\ldots(N-n+1) \quad X_i = 1,2,\ldots,N \qquad i = 1,2,\ldots,n \quad X_i \neq X_j, \text{ all } i \neq j. \tag{1}$$

Let W = max $(X_1, X_2, ..., X_n)$. This is clearly the maximum likelihood estimator of N since the likelihood function above is a decreasing function of N. It is also straightforward to show (Tenenbein, 1971)

$$E\,W = n(N+1)/(n+1) \tag{2}$$

$$\text{Var}\,W = n(N+1)(N-n)/(n+1)^2(n+2). \tag{3}$$

It is known that W is sufficient for N and that the distribution of W is complete (Tenenbein, op. cit.). W is therefore a complete sufficient statistic so if there is a function of W that is unbiased, this function must be the unique best unbiased estimator. Thus,

$$U = W(n+1)/n - 1 \tag{4}$$

is accordingly the unique minimum variance unbiased estimator.

This estimator can be interpreted as the "average gap" estimator (Noether, 1971; Rohatgi, 1984; Rao, 1981) i.e., W + average gap, where the gaps are the gaps or spacings between successive ordered observations (Spencer and Largey, op. cit.).

While U is unbiased, W is only unbiased for n = N. U has the higher variance, however, with:

$$\begin{aligned} &\text{var } U / \text{var } W = (n+1)^2 / n^2 \\ &\rightarrow (N+1)^2 / N^2 \text{ as } n \rightarrow N. \end{aligned} \tag{5}$$

Relative MSE can be calculated as:

$$\phi = \text{MSE } U / \text{MSE } W = (n+1)(N+1)/n\ (2N-n). \tag{6}$$

For any N, ϕ exceeds unity for n=1 and n≥N−2. If n=2, $\phi>1$ only if N≤6. If n=N−3, $\phi>1$ only if N≤6. Tenenbein, op. cit., calculates ϕ for various n and N=100, 200.

Below we present graphs for N=10, 100 and 500. (See also Table 1 for numerical values of MSE.)

It is clear that U completely outperforms W except for very low or very high n. The rapidity with which the graph declines for low n should be noted. W is of little practical value unless n=1.

(b) Sampling With Replacement

The joint pdf of $X_1,\dots,X_n$ is:

$$f(X_1,X_2,\dots,X_n) = (1/N)^n,\ X_i = 1,2,\dots,N \tag{7}$$

Again, $W = \max\ (X_1,X_2,\dots,X_n)$ is the MLE. (8)

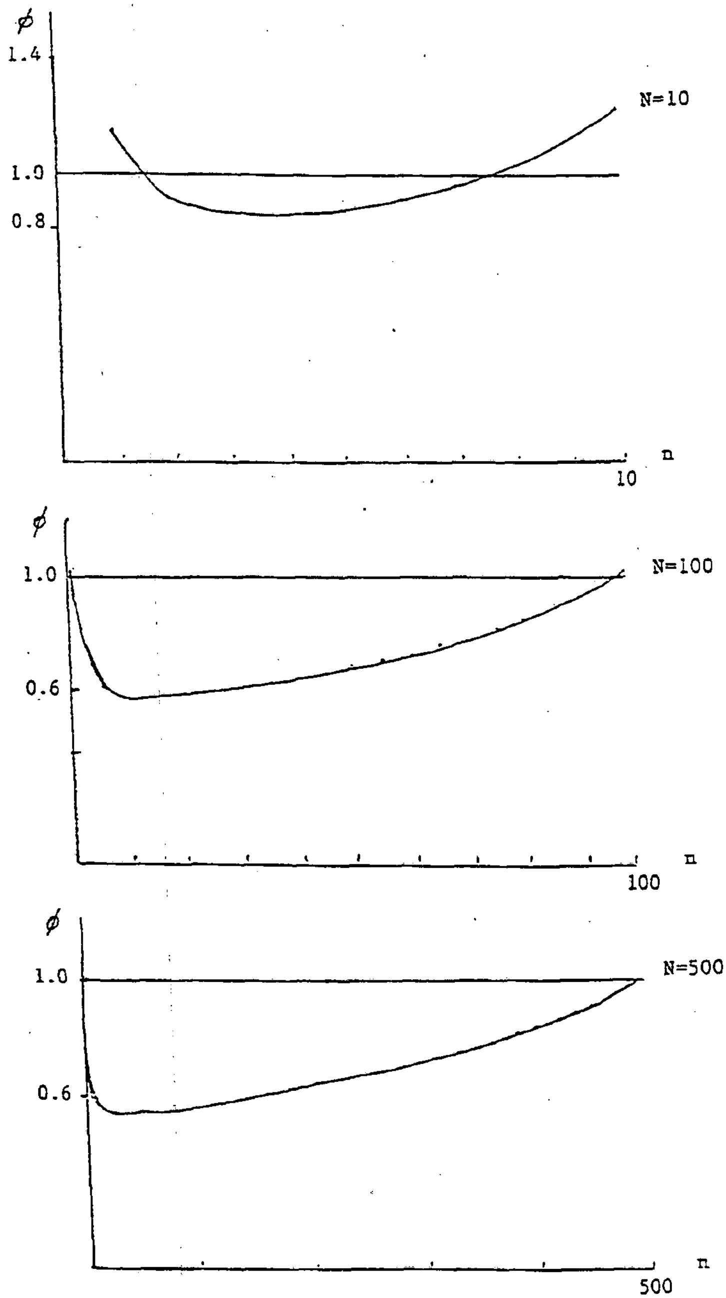
The distribution function of W is $G(w) = (w/N)^n$.

Hence the pdf of W is:

$$g(w) = (w/N)^n - ((w-1)/N)^n,\ w = 1,2,\dots,N \text{ and} \tag{9}$$

$$EW = N - \sum_{j=1}^{N-1} j^n / N^n \tag{10}$$

Figure 2: *MSE U/MSE W for Various N, n: Discrete Case, Sampling Without Replacement*



ϕ = MSE U/MSE W = (n+1)(N+1)/n(2N−n)

$$EW^2 = N^2 - \sum_{j=1}^{N-1} (2j+1)j^n/N^n. \tag{11}$$

(These expressions involve sums of powers of integers which can be calculated by computer or by looking up tables. They can also be rewritten using Bernoulli numbers which in turn can be evaluated using tables e.g., Abramovitz and Stegun, 1965.)

As before, W is sufficient for N (Wasan, 1970; Rohatgi, 1976) and the distribution of W is complete (Wasan, op. cit.; Rohatgi, op. cit.). Accordingly, to find the unique MVUE estimator of N, we look for the function of W that is unbiased. This is:

$$U = \left(W^{n+1} - (W-1)^{n+1}\right) / \left(W^n - (W-1)^n\right). \tag{12}$$

(Wasan, op .cit.; Rohatgi, op. cit.; Guenther, op. cit.)

The notion of an "average gap" in this case is not straightforward and the MVUE above does not seem to have any such interpretation.

Notice that unlike the case without replacement, n may be unboundedly large and Var W will remain positive, as $n \to \infty$. Notice also that approximations for large N are possible (see Rohatgi (1984)).

Defining ϕ = MSE U/MSE W, the results for the with replacement case are summarised in Figure 3. (See also Table 2 for numerical results.)

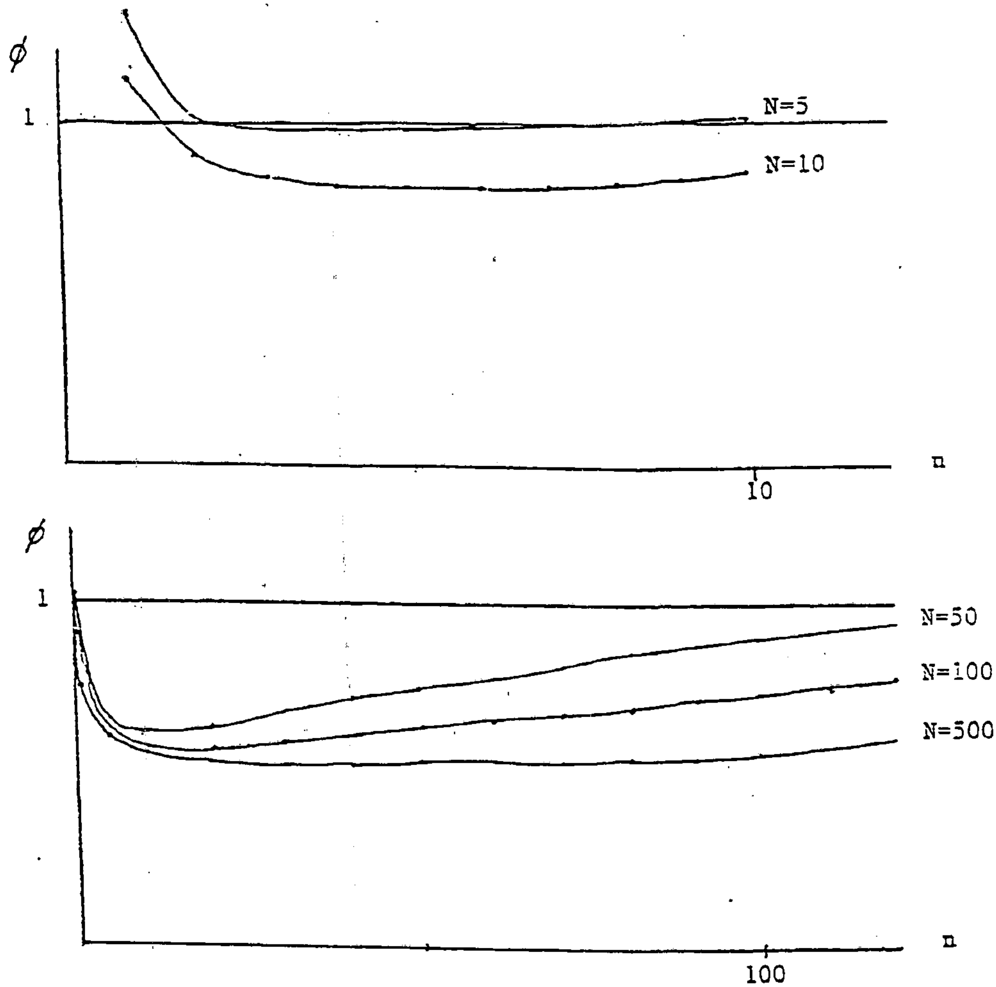
For the case n=1, where with and without replacement coincide, ϕ = (2N+2)/(2N-1) which is greater than one but decreases monotonically towards one as N increases.

As n increases, ϕ decreases very sharply at first, and especially rapidly for large N. The graphs show ϕ, after the initial sharp decline typically continues to decline slowly (to the point n=3 for N=5, to n=4 for N=9, to n=9 for N=50, to n=14 for N=100, to n=30 for N=500, to n=44 for N=1000) and then slowly rises ultimately exceeding 1 (when n=9 for N=5, when n=56 for N=10). As $n \to \infty$, EW $\to$ N, and var W and var U vanish, but our calculations show ϕ well below unity for even quite large relative values of n. In no case does ϕ fall below 0.5 (its limit in the continuous case) but for large N, it does get rather close to it. Some minimum ϕ values are as follows: N=100, n=15, ϕ = 0.578; N=500, n=31, ϕ = 0.533; N=1000, n=45, ϕ = 0.523.)

Note that:

$$1/\phi = \text{var } W/\text{var } U + (\text{bias } W)^2/\text{var } U. \tag{13}$$

Figure 3: *MSE U / MSE W for Various N, n: Discrete Case, Sampling With Replacement*



ϕ = MSE U/MSE W

Calculations show that var W/var U (<1) decreases as N increases, given n and increases with n, given N. (Bias W)^2/var U appears to change in the opposite way.

For very small n the relative variance advantage of W tends to offset the bias disadvantage, so that MLE does reasonably well for very low n, especially if N is not too high.

It is interesting to observe that these graphs are much more similar to the discrete case without replacement than to the continuous case — in the sense that ϕ is above 1 for n=1 but falls below 1 for n=2 (N≥7) reaches a minimum quickly and then rises slowly. (For N=10, ϕ reached 1 at n=56; for N=20, ϕ reached 1 at n=247; and for N=50, ϕ reached 1 at n=816.)

In the continuous case, ϕ fell monotonically as n increased, reaching 0.5 in the limit. In the discrete case without replacement, it is clear that

$$\phi = (n+1)(N+1)/n(2N-n) > \tfrac{1}{2}, \text{ all } n, N \tag{14}$$

and, judging from our calculations, this seems to hold in the with replacement case also.

Examination of the graphs suggests that the with replacement case is closer than the without replacement case to the continuous case, primarily because of the slowness with which ϕ increases after its initial drop. A similar conclusion holds using ψ as the reference — see below.

(2) The Geary Estimator vs MVUE

G is defined to be $2^{1/n}W$, and in the continuous case (and in the discrete case with replacement) is the "closest" estimator.

Its behaviour relative to the MVUE estimator in the three cases is rather similar.

In the continuous case, writing MSE U/MSE G = $\psi(n)$,

$\psi(1) = 1$, $\psi(2) = 1.0928$ and ψ then steadily declines as n rises.

In fact: $\psi(n) < 1$ for $n \geq 7$
$\psi(n) \to .913944$ as $n \to \infty$.

In the discrete cases, write MSE U(n, N)/MSE G(n, N) $\equiv \psi(n, N)$.

(a) Sampling Without Replacement

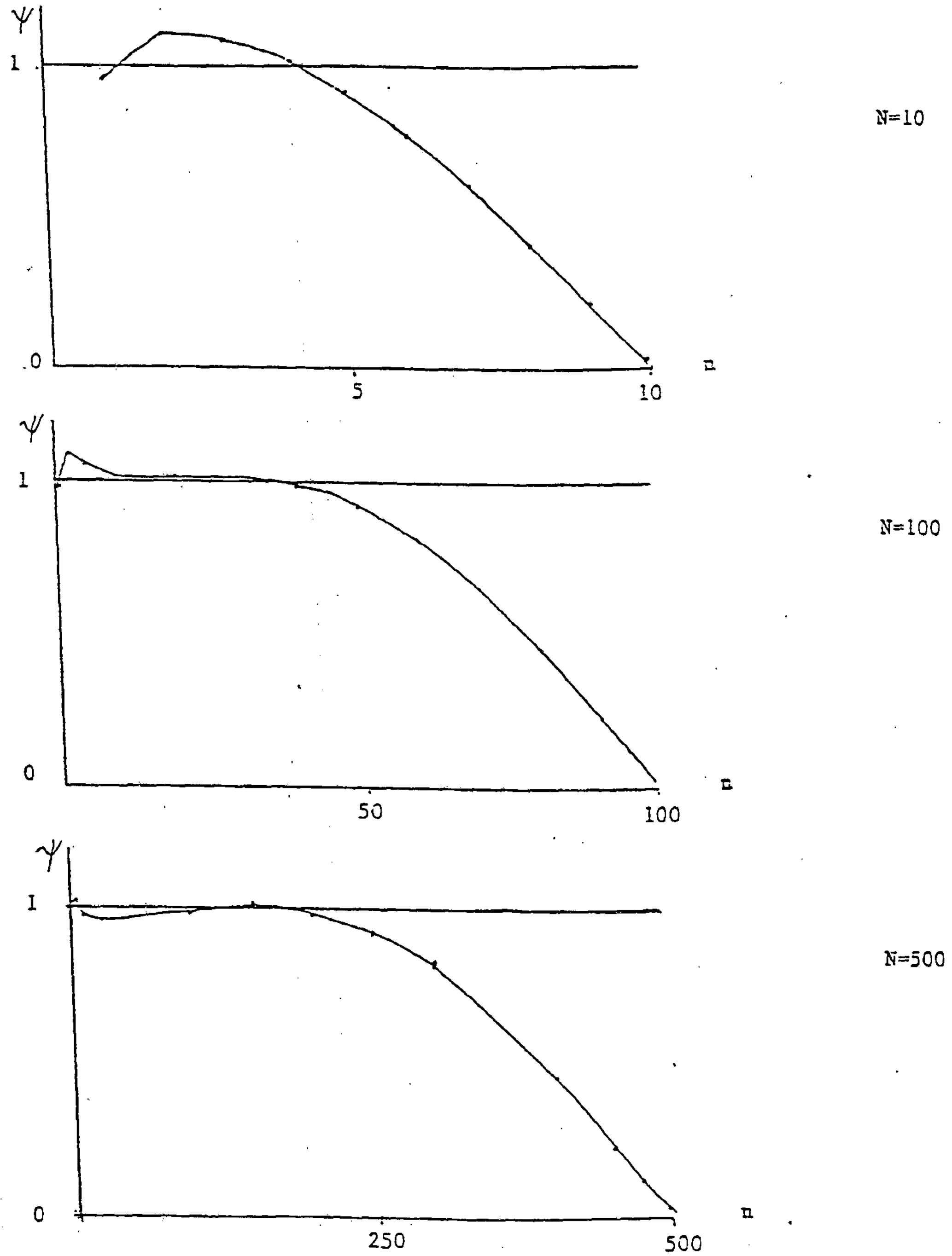
$$\begin{aligned} EG &= N+1 \text{ for } n=1 \\ &= 2^{1/N}N \text{ for } n=N \end{aligned} \tag{15}$$

$$\begin{aligned} \text{Var } G &= 2^{2/n}n(N+1)(N-n)/(n+2)(n+1)^2 \\ &= 0 \text{ for } n=N \end{aligned} \tag{16}$$

$$\begin{aligned} \text{MSE } G(n=N) &= N^2(1-2^{1/N})^2 \\ &\to (\log_e 2)^2 \text{ as } N \to \infty \\ &= .480453. \end{aligned} \tag{17}$$

Since MSE U = var U = $(N+1)(N-n)/(n+2)(n+1)^2$, it is clear that G will not do well for n near N.

Figure 4: *MSE U/MSE G for Various N, n: Discrete Case, Sampling Without Replacement*



ψ = MSE U/MSE G

$$\psi(1,N) = (N^2 - 1)/(N^2 + 2) < 1$$
$$\psi(2,N) > 1 \text{ if and only if } N \geq 6. \tag{18}$$

For low values of $n \geq 2$, it is clear from the graphs (Figure 4) that G does excellently relative to U. Only when n exceeds about 0.4 N, does the graph decline significantly — though after that it does tend to decline rapidly. At that stage, of course, MSE U and MSE G are both rather low.

(b) Sampling With Replacement

For $n = 1$, $\text{var } U = (N^2 - 1)/3$

$$EG = N + 1 \tag{19}$$

$$\text{Var } G = (N^2 - 1)/3 \tag{20}$$

$$\text{MSE } G = (N^2 + 2)/3 \tag{21}$$

$$\psi(1,N) = (N^2 - 1)/(N^2 + 2) < 1. \tag{22}$$

From the graphs shown in Figure 5, for n=2, $\psi > 1$. As n increases, ψ falls slowly initially (remaining above 1 for values of n quite high relative to N), but decreases more rapidly after reaching 1.

G does particularly well for N quite large (say 50 or more) and n between 2 and about 0.9N. The advantage can be quite substantial. For example, if N=50, n=3, MSE U=166.52, MSE G=153.29. The bias in G is more than offset by a low variance.

IV MINIMUM MEAN SQUARE ERROR

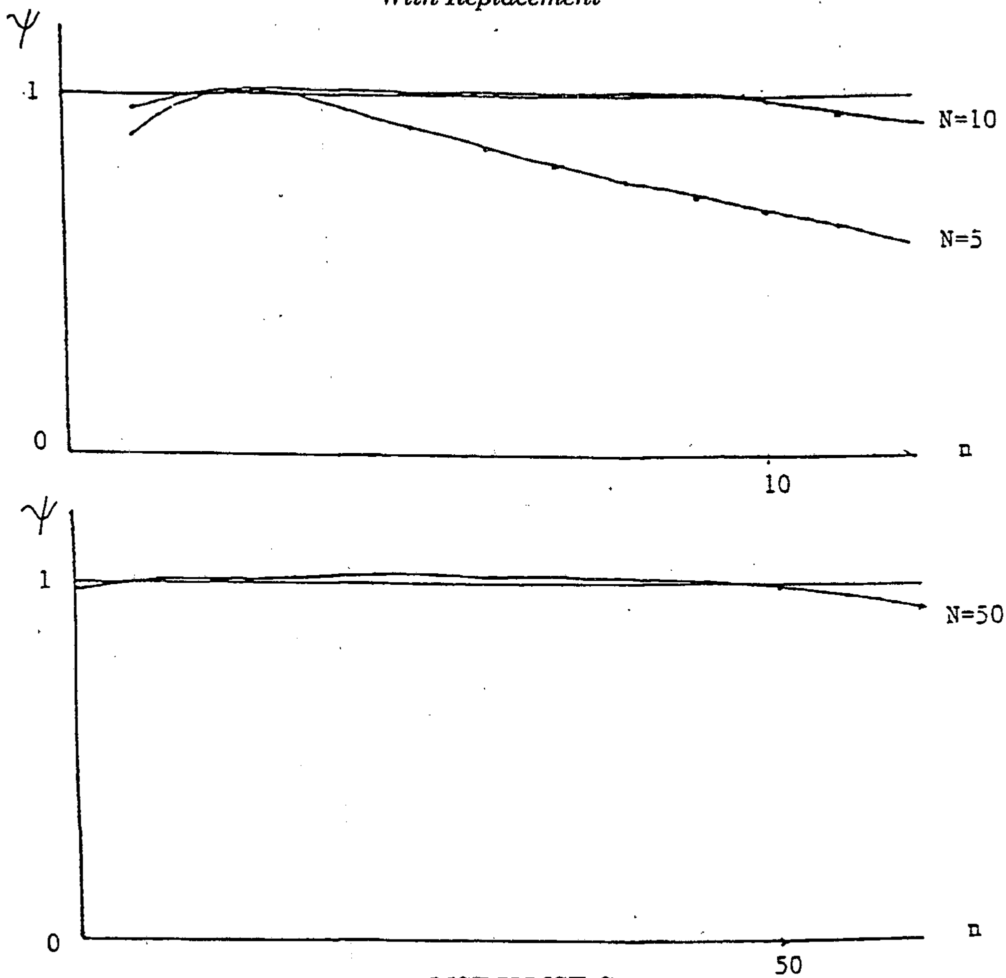
Examining the general problem of finding the minimum MSE estimator for θ of the form $\lambda\hat{\theta} + k$, with k constant, where $\hat{\theta}$ is an estimator of θ with finite mean and variance, the minimum MSE λ is found from differentiation to be:

$$\lambda^* = (\theta - k)E\,\hat{\theta} / E\,\hat{\theta}^2. \tag{23}$$

The corresponding MSE is:

$$\text{MSE } (\lambda^*\hat{\theta} + k) = \lambda^{*2}\text{var }\hat{\theta} + (\lambda^* E\hat{\theta} + k - \theta)^2$$
$$= (\theta - k)^2[(E\hat{\theta} / E\hat{\theta}^2)^2 \text{ var }\hat{\theta} + ((E\hat{\theta})^2 / E\hat{\theta}^2 - 1)^2]. \tag{24}$$

Figure 5: *MSE U/MSE G for Various N, n: Discrete Case, Sampling With Replacement*



ψ =MSE U/MSE G

Clearly this diminishes as k tends towards θ, i.e., as $\lambda^* \rightarrow 0$ and the estimator tends towards θ. As θ is unknown this fact is of little practical significance, though it is worth noting that when θ is a positive integer, as in the discrete taxi problem, a negative k is worse than k = 0. If k = 0, λ^* is not necessarily dependent on θ.

In the continuous taxi problem with k=0, θ=N and $\hat{\theta}$ set to W, λ^* = NEW/EW^2 = (n+2)/(n+1). Thus the estimator J = (n+2)W/(n+1) has minimum mean square error in the class of estimators λW, λ constant, and is not dependent on N (Johnson, 1950). For non-zero k, the estimator would be:

$$(N-k)(n+2)W/N(n+1) + k. \tag{25}$$

(a) The Discrete Case Without Replacement

In the discrete case without replacement, when k=0, the minimum MSE estimator of the form λW is:

$$\begin{aligned} J^* &= \lambda^* W = (n+2)W/(n+1+n/N) \\ &= W \text{ for } n = N. \end{aligned} \tag{26}$$

With N unknown, this is not available, though it shows that J = (n+2)W/(n+1) will have minimum MSE among estimators of the form λW if n is small relative to N. Note also that if n/N is close to one, MLE will have minimum MSE among estimators of the form λW.

Since the MVUE is of the form $\lambda W-1$, we consider other estimators of the same form, where λ depends on n.

Mean Square Error of $(\lambda W-1)$ is λ^2 var W + $[\lambda EW-1-N]^2$. From the equation for λ^* above, the optimal λ is:

$$\lambda^* = (n+2)(N+1)/(n+N(n+1)). \tag{27}$$

The Minimum Mean Square Error estimator in the class λW-1 is thus:

$$\begin{aligned} L^* &= \lambda^* W - 1 \\ &= (n+2)(1+1/N)W/(n+1+n/N) - 1 \end{aligned} \tag{28}$$

which collapses to

$$L = (n+2)W/(n+1) - 1, \text{ for small } n/N. \tag{29}$$

When n=N

$$L^* = (1+1/N)W - 1. \tag{30}$$

Since MSE $(\lambda W + k)$ increases with $(N\text{-}k)^2$, J^* must have lower MSE than L^*. Accordingly, from the formulae for the estimators J and L, J must have lower MSE than L provided n/N is sufficiently small — see Table 1.

Since var J and var L are equal, the relative performance of J and L turns on the relative biases.

$$EL = EJ - 1 \doteq N \text{ for n large.} \tag{31}$$

Accordingly L will have lower MSE for n large.

In fact, the necessary and sufficient criterion is:

$$\text{L has lower MSE than J if and only if } 2N+1<n(n+2). \tag{32}$$

Of course, this criterion involves the unknown N, but it has practical value given vague knowledge of N. J is better than L for only quite low values of n: n≤3, N=10; n≤13, N=100; n≤43, N=1000.

The above theory and calculations suggest that if the statistician believes J beats L, then J should be used. It is preferable to U and when not preferable to G is close to G — see Table 1. When L beats J, L has very similar MSE to that of U and the statistician may as well choose between U and G as discussed in the previous section.

(b) The Discrete Case With Replacement

Using the approximation

$$1^n + 2^n + \ldots + (N-1)^n \doteq N^{n+1}/(n+1), \tag{33}$$

to find the MMSE estimator of the form λW, we have

$$\begin{aligned} \lambda^* &= NEW/EW^2 \doteq (n+2)/(n+1-(n+2)/nN) \\ &= (n+2)/(n+1) \text{ for N large.} \end{aligned} \tag{34}$$

$N(EW/EW^2)$ collapses for large N to J also using closer approximations.

Accordingly J should do well if N is large.

The calculations, shown in Table 2, confirm this, and show the three leading estimators as J, G and U.

J in fact does best for small n, any N. For example, J is best for n as high as 3 if N=10, as high as 25 if N=100, and as high as 318 if N=1000. Beyond these limits, either G or U emerges as best with their comparative performance as described in Section III.

V CONCLUSIONS

This paper has reviewed and further considered estimation in the taxi problem in the continuous case and in the discrete cases with and without replacement. The estimators discussed were W, the maximum likelihood and complete sufficient estimator; U, the minimum variance unbiased estimator (the formula for which depended on the situation under consideration); Geary's estimator $G = 2^{1/n}W$; and various forms of Minimum Mean Square Error estimator including that of the form λW which under certain conditions collapsed to Johnson's estimator J=(n+2)W/(n+1). Note that all estimators are based in some way on W.

The maximum likelihood estimator W performs poorly in all cases, especially for moderate sample sizes, using mean square error as the key criterion. Its poor asymptotic behaviour in the continuous case is not replicated in the discrete cases, however, but large samples are typically required for it to match the MVU estimator in those cases. The odd result that the MLE estimator does reasonably well for small samples in the continuous case is replicated in the discrete cases, but its performance relative to MVUE falls away very rapidly at first as n increases, especially if N is large.

Are there any general recommendations which can be made as regards choice of an estimator of N for the various classes of situation we have dealt with? We use MSE as the key criterion, noting that other criteria (bias, closeness, etc.) could point in different directions.

In the continuous case the decision is clear-cut — J should be chosen as it is the Minimum Mean Square Error estimator in a wide class. Its advantage is slight, however, unless n is very low (<5). The Geary estimator performs well for small and moderate n, while for larger n, U=(n+1)W/n is better than G under the criterion of MSE.

In the discrete case both with and without replacement, it is no longer possible to use the exact MMSE estimator since this is now a function of the unknown N. The choice of which estimator to use is therefore no longer clear. However, the following tables give a rough idea of the best estimators, in terms of minimum MSE, to use in various cases. Note that the estimator L does not appear in the without replacement table. L is rejected on the grounds that either J, G or U will better it when N is small, or when n/N is small, whereas with N large, n/N not small, MSE U approximates MSE L so that little is lost by using U in this case.

Use of these tables requires either some prior information on N, or the ability to deduce some information on N from the observed sample values. In the without replacement case n is measured relative to N, whereas in the with replacement case n is measured in absolute terms. S denotes small, M moderate and L large.

Without Replacement

n \ N	S	M	L
S	J,G	J,G	J
M	U,G	U,G	U
L	U	U	U

With Replacement

n \ N	S	M	L
S	J,G	J,G	J
M	U	J,G	J
L	U	U	U

For the special case where n=1 the ordering of J, W, U, G is clear.

For the continuous case, the situation is,

	J	W	U	G
E	3N/4	N/2	N	N
Var	$3N^2/16$	$N^2/12$	$N^2/3$	$N^2/3$
MSE	$N^2/4$	$N^2/3$	$N^2/3$	$N^2/3$

In the discrete case where with and without replacement coincide when n=1,

	J	W	U	G
E	3(N+1)/4	(N+1)/2	N	N+1
Var	$3(N^2-1)/16$	$(N^2-1)/12$	$(N^2-1)/3$	$(N^2-1)/3$
MSE	$(2N^2-3N+3)/8$	$(2N^2-3N+1)/6$	$(N^2-1)/3$	$(N^2+2)/3$

Thus MSE J < MSE W if $N > 2$
MSE W < MSE U if $N > 1$
MSE U < MSE G for all N.

REFERENCES

ABRAMOVITZ, M., and I. STEGUN (eds), 1965. *Handbook of Mathematical Functions*, National Bureau of Standards (1964) and New York: Dover Publications, Inc.

DAVIS, R.C., 1951. "On Minimum Variance in Non-regular Estimation", *Annals of Mathematical Statistics*, Vol. 22, pp. 43-57.

GEARY, R.C., 1944. "Comparison of the Concepts of Efficiency and Closeness for Consistent Estimates of a Parameter", *Biometrika*, Vol. 33, No. 2, pp. 123-128.

GUENTHER, W.C., 1978. "Some Easily Found Minimum Variance Unbiased Estimators", *The American Statistician*, Vol. 32, No. 1, pp. 29-34.

JEFFREYS, H., 1939. *Theory of Probability*, Oxford: Clarendon Press.

JOHNSON, N.L., 1950. "On the Comparison of Estimators", *Biometrika*, Vol. 37, Nos. 3/4, pp. 281-287.

KOTZ, S., and JOHNSON, N.L., 1985. *Encyclopedia of Statistical Sciences*, Vol. 6, New York: Wiley.

MOSTELLER, F., 1965. *Fifty Challenging Problems in Probability*, London: Addison-Wesley.

NOETHER, G., 1971. *Introduction to Statistics: A Fresh Approach*, Boston: Houghton Miflin.

PATEL, J.K., 1973. "Complete Sufficient Statistics and MVU Estimators", *Communications in Statistics*, Vol. 2, No. 4, pp. 327-336.

RAO, J.S., 1981. "Estimation Problems for Rectangular Distributions (or the Taxi Problem Revisited)", *Metrika*, Vol. 28, pp. 257-262.

ROHATGI, V., 1976. *An Introduction to Probability Theory and Mathematical Statistics*, New York: Wiley.

ROHATGI, V., 1984. *Statistical Inference*, New York: Wiley.

ROSENBERG, W. J., and DEELY, J.J., 1976. "The Horse-Racing Problem — A Bayesian Approach", *The American Statistician*, Vol. 30, No. 1, pp. 26-29.

SPENCER, J.E., and LARGEY, A., 1993. "Geary on Inference in Multiple Regression and on Closeness and the Taxi Problem", *The Economic and Social Review*, Vol. 24, No. 3, pp. 275-295.
TENENBEIN, A., 1971. "The Racing Car Problem", *The American Statistician*, Vol. 25, No. 1, pp. 38-40.
WASAN, M.T., 1970. *Parametric Estimation*, London: McGraw-Hill.

APPENDIX

Table 1. *MSE of Various Estimators of the Parameters in the "taxi" Problem: Without Replacement*

N=10	*G*	*U*	*J*	*L*	*W*
n					
1	34.0000	33.0000	21.6250	26.1250	28.5000
2	9.9153	11.0000	8.7407	10.1852	12.0000
4	2.7053	2.7500	2.8480	2.7280	3.2000
6	1.1887	0.9167	1.4810	0.9300	1.0000
8	0.6976	0.2750	1.0151	0.2867	0.2667
10	0.5151	0.	0.8264	0.0083	0.
N=100					
n					
1	3334.0000	3333.0000	2462.8750	2512.3750	3283.5000
4	380.9633	404.0000	381.5680	388.6480	627.2000
10	74.4459	75.7500	74.5304	75.1998	129.5454
20	18.0260	18.3636	18.8748	18.3329	31.1688
50	2.0809	1.9423	2.8647	1.9423	2.8280
80	0.6914	0.3079	1.2773	0.3081	0.3613
100	0.4838	0.	0.9802	0.	0.
N=1000					
n					
1	333334.000	333333.000	249625.375	250124.875	332833.500
4	39872.737	41541.500	39808.768	39887.848	66267.200
10	8449.272	8258.250	8175.207	8190.752	14925.000
20	2337.600	2229.500	2221.013	2224.552	4200.000
50	385.716	365.750	365.847	365.617	698.529
100	91.860	88.323	89.120	88.316	165.987
200	20.034	19.822	20.772	19.821	35.466
500	2.142	1.994	2.986	1.994	2.982
800	0.692	0.312	0.309	0.312	0.374
1000	0.481	0.	0.998	0.	0.

Table 2: *MSE of Various Estimators of the Parameters in the "taxi" Problem: With Replacement*

N=10	*G*	*U*	*J*	*L*	*W*
1	34.0000	33.0000	21.6250	26.1250	28.5000
2	11.0675	12.3138	10.0444	11.9778	13.6500
4	3.7053	4.0379	3.7936	4.4736	4.9677
6	1.8700	1.9723	2.0184	2.3976	2.4288
8	1.1267	1.1482	1.2694	1.5523	1.3829
10	0.7520	0.7378	0.8804	1.1344	0.8644
20	0.1985	0.1520	0.2692	0.5974	0.1604
50	0.0023	0.0052	0.0042	0.6602	0.0052
100	0.00486	0.00003	0.00982	0.00003	0.00003
N=100					
1	3334.0000	3333.0000	2462.8750	2512.3750	3283.5000
4	395.3097	416.5365	395.5152	403.3232	646.9967
10	83.1911	83.2326	81.9646	83.5448	142.7493
20	22.8285	22.6356	22.6199	23.0607	38.8451
50	3.7356	3.7605	3.9450	4.0869	5.8745
80	1.4367	1.4417	1.6230	1.7748	2.0491
100	0.9098	0.8995	1.0708	1.2462	1.2105
200	0.2149	0.1779	0.2958	0.6106	0.2004
500	0.0241	0.0067	0.0439	0.6580	0.0067
1000	0.00484	4.32E-05	0.01001	0.81030	0.00004
N=1000					
1	333334.000	333333.000	249625.375	250124.875	332833.500
4	40020.142	41666.536	39952.320	40032.120	66466.999
10	8542.358	8333.233	8255.677	8272.117	15060.939
20	2392.029	2272.635	2265.396	2269.887	4281.717
50	409.360	384.529	384.251	385.009	734.869
80	161.800	152.354	152.430	152.736	289.095
100	103.713	97.954	98.096	98.299	184.561
500	3.920	3.901	4.109	4.198	6.250
1000	0.931	na	1.093	1.258	1.254
1500	0.3997	na	0.5129	0.7544	0.4509