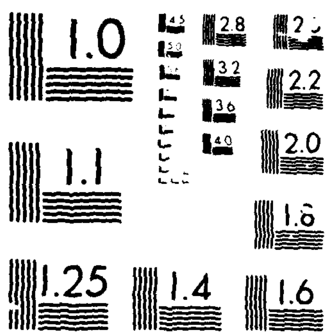
AD-A169 618 STATISTICAL EFFECTS OF IMPERFECT INSPECTION SAMPLING 3 1/1
SCREENING (GROUP T..(U) MARYLAND UNIV COLLEGE PARK DEPT
OF MANAGEMENT SCIENCES AND ST.. N L JOHNSON ET AL.
UNCLASSIFIED MAR 86 N00014-84-K-0301 F/G 12/1 NL


1.0
2.8
2.2
3.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Contract N00014-84-K-0301

3

STATISTICAL EFFECTS OF IMPERFECT INSPECTION SAMPLING
III. SCREENING (GROUP TESTING)

by

Norman L. Johnson
University of North Carolina at Chapel Hill

and

Samuel Kotz
University of Maryland, College Park

and

Robert N. Rodriguez
SAS Institute, Cary, North Carolina

FINAL VERSION MARCH 1986

Abstract

This paper, the third in a series, describes the effects of imperfect inspection on screening (group testing) acceptance sampling procedures. These procedures are based on first grouping items and testing the group as a whole for presence or absence of nonconforming (NC) items proceeding to individual testing only if such presence is indicated. If sampling is perfect, the procedure can reduce the expected amount of testing, especially if the proportion of NC items is low. Modifications of the basic procedure - curtailed and hierarchal group testing - aimed at further reduction in the expected amount of testing, are also discussed. Some new formulas are presented.

KEY WORDS: Acceptance Sampling Plans, Binomial Distribution, Compound Distributions, Hypergeometric Distributions.

INTRODUCTION

This is the third in a series which explores the statistical effects of errors in classifying nonconforming (NC) items in sampling inspection. The first paper covered single sample acceptance procedures. The second paper extended these developments to double, link, and partial link sampling schemes. The reader is advised to consult the prior papers (Johnson, et al. (1985, 1986)) in order to familiarize himself or herself with the previous material. The following terms are used in the subsequent development:

N = lot size

D = number of nonconforming items in a lot of size N.

For 'infinite' lot size (or for sampling with replacement) ω will denote the proportion of NC items. This can be considered as a limiting case, with $N \to \infty$, $D \to \infty$ and $D/N \to \omega$.

The numbering of equations and tables follows from that of the preceding papers in the series (Johnson et al. (1985), (1986)).

In the first two papers, we were concerned with situations wherein the consequences of sample inspection were acceptance or rejection of a lot. We now consider situations in which we are interested in identifying <u>individual</u> items as conforming (C) or NC, and doing this with as much economy in inspection as possible. The procedures to be discussed <u>can</u> also be applied in acceptance sampling though the emphasis in the present paper will be on correct classification of individual items (as well as economy in inspection).

SCREENING

Screening procedures (also called group testing procedures) are based on the ideas of Dorfman (1943). He was concerned with testing large numbers of blood samples for syphilis, and noted that considerable savings in the amount of testing might be attained by mixing together blood from a group of n_0 samples and testing the mixture for presence of syphilis. Only if the test gave a positive result would testing of individual samples be undertaken.

Terming the affected samples as nonconforming (NC) it can be seen that if the proportion of NC items is small, then quite often only the first, single group test will be needed, instead of the n_0 which would be required if each item were tested separately. There is, of course, the possibility that (n_0+1) tests may be needed, instead of n_0. The two effects tend to counterbalance each other. Analysis indicated that (on the assumption of perfect inspection), for maximum saving in expected number of tests, the group size, n_0, should be approximately equal to $\{2(\text{proportion NC})\}^{-\frac{1}{2}}$.

The method is not applicable to situations where the testing is destructive, but it can be used in a variety of problems. For example, Hwang (1984) mentions problems of pollution, leakage, identification of active users, and flow through series of components, and Mundel (1984) describes application in testing sets of cells in electric batteries. An unusual example of application to problems in multiaccess communications is presented in Wolf (1985).

If inspection is subject to error, the effectiveness of this procedure can be affected in two ways. Errors at the first (group) test can lead to unnecessary individual testing, or to neglect of testing NC items individually, in addition to the errors incurred at individual testing which are of the same kind as described in the first paper in this series (Johnson et al. (1985)).

Analytical results and some illustrative tables are available in Kotz and Johnson (1982). Table 6 is an extrast from Table 1 of that paper. Here we also summarize the results of the analysis.

We use the following notation:

$$\left.\begin{aligned} &\text{Group Testing:} \quad p_0 = \Pr[\text{declare NC} \mid \text{at least one NC item}] \\ &\qquad\qquad\qquad\quad p_0' = \Pr[\text{declare NC} \mid \text{no NC items}] \end{aligned}\right\} \quad (18)$$

$$\left.\begin{aligned} &\text{Individual Testing:} \quad p = \Pr[\text{declare NC} \mid \text{NC}] \\ &\qquad\qquad\qquad\qquad p' = \Pr[\text{declare NC} \mid \text{not NC}] \end{aligned}\right\} \quad (19)$$

We suppose a random sample of n_0 items is chosen without replacement.

The overall probability of obtaining a positive result (indicating presence of at least one NC item) so that individual testing will be carried out is

$$\{1 - P_0(n_0)\}p_0 + P_0(n_0)p_0' = p_0 - (p_0 - p_0')P_0(n_0) \quad (20)$$

and the expected number of tests is

$$E = 1 + n_0\{p_0 - (p_0 - p_0')P_0(n_0)\} \quad (21)$$

where $P_0(n_0) = (N\cdot D)^{(n_0)}/N^{(n_0)}$ is the probability that the sample contains no NC items $(a^{(b)} = a(a-1)\ldots(a-b+1))$. [For infinite lot size $P_0(n_0) = (1-\omega)^{n_0}$].

The probability that a NC item will be correctly classified is

$$PC(NC) = p_0 p\,; \quad (22)$$

the probability that a conforming item will be correctly classified is

$$PC(C) = P_0^*(n)(1 - p_0'p') + \{1 - P_0^*(n_0)\}(1 - p_0 p')\,, \quad (22)$$

where $P_0^*(n_0) = (N-D-1)^{(n_0-1)}/(N-1)^{(n_0-1)}$ is the probability that the sample contains no NC items, <u>given</u> that it contains at least one conforming item. [For infinite lot size $P_0^*(n_0) = (1-\omega)^{n_0-1}$].

From (20) it is clear that the probability of correct classification of a NC item is decreased by the screening process (since $p_0 p \leq p$). The value of screening must therefore come from increased correct classification of conforming items and/or reduction in the expected number of tests. This is presented in the last column of Table 6 (with $N = \infty$).

HIERARCHAL SCREENING

Instead of proceeding directly from testing the group of n_0 items as a whole, to testing individual items, one or more intermediate stages may be interposed. If testing of the single group of n_0 gives a positive result, the group is split into h_1 subgroups, each containing $n_1 (= n_0/h_1)$ items; each subgroup is then tested. If a subgroup gives a negative result, all items in it are accepted as C; if it gives a positive result then each of h_2 sub^2groups, each containing $n_2 (= n_1/h_2)$ items, and so on. Finally if testing of a sub^kgroup gives a positive result each of the $n_k (= n_0/(\prod_{j=1}^{k} h_j))$ items in it is tested individually. These rules constitute a (k+1)-stage hierarchal screening procedure.

Hierarchal screening procedures are discussed by Kotz & Johnson (1982). They analyze the effects of imperfect inspection on two-stage procedures (i.e. $k = 1$) and provide some illustrative tables. A portion of Table 2 of that paper is reproduced here as Table 7. There are examples of more elaborate hierarchal procedures in Hwang (1984) and Mundel (1984). Interesting relationships between screening procedures and multiaccess communication theory are discussed in Wolf (1985). (See also the references therein.)

We now summarize formulae for properties of a (k+1)-stage hierarchal procedure. We use the following notation.

$$\text{Testing of Sub}^s\text{group:} \quad \left.\begin{aligned} p_s &= \Pr[\text{declare NC} \mid \text{at least one NC item}] \\ p'_s &= \Pr[\text{declare NC} \mid \text{no NC items}] \end{aligned}\right\} \qquad (24)$$

[We use sub^0 group to denote the original group of n_0 items.]

For convenience in writing formulae, we will replace p,p' (as defined in (19)) by p_{k+1}, p'_{k+1} respectively. Note also that

$h_{k+1} = n_k$ = number of individual items in each sub^k group. The formulae are (cf. (21)-(23))

$$E = \sum_{s=0}^{k+1} \left(\prod_{j=1}^{s} h_j\right) \pi(s) \tag{25}$$

where

$$\pi(s) = P_0(n_0) \prod_{i=0}^{s-1} p'_i + \sum_{j=1}^{s-1} \{P_0(n_j) - P_0(n_{j-1})\}\left(\prod_{i=0}^{s-1-j} p'_i\right)\left(\prod_{i=s-j}^{s-1} p_i\right)$$

$$+ \{1 - P_0(n_{s-1})\} \prod_{i=0}^{s-1} p_i \tag{26}$$

$$\left(\text{with } \sum_{j=1}^{0} h_j = \pi(0) = 1\right).$$

$$PC(NC) = \prod_{j=0}^{k+1} p_j \tag{27}$$

$$PC(C) = 1 - \sum_{s=0}^{k+1} \{P_0^*(n_s) - P_0^*(n_{s-1})\} \left(\prod_{j=0}^{s-1} p_j\right)\left(\prod_{j=s}^{k+1} p'_j\right) \tag{28}$$

with $P_0^*(n_{-1}) = 0$; $P_0^*(n_{k+1}) = 0$)

(Note that $\pi(s)$ is the probability that a randomly chosen sub^{s-1} group will be inspected (as a whole) and found to give a positive result, so leading to separate inspections of each of the sub^s groups contained in it.)

The ratio of E to the number of groups ($n_0 = \prod_{j=1}^{k+1} h_j$) is

$$\frac{E}{n_0} = \sum_{s=0}^{k+1} \left(\prod_{j=s+1}^{k+1} h_j\right)^{-1} \pi(s) \tag{29}$$

Tables 8 contain values of PC(NC), PC(C) and E for some 3-stage procedures cases. Study of Tables 7 and 8 indicates how use of hierarchal schemes can lead to further savings in expected amount of testing, as compared with a basic Dorfman procedure. As one might expect, savings are obtained when they would be achieved applying a Dorfman procedure to initial groups of the size of the subgroups concerned.

The expected number of tests increases with a, but the probability of correct decision PC(C) does not depend markedly on ω; such dependence as there is increases with n_0. Note that PC(C) is inversed markedly by use of screening procedures.

CURTAILED GROUP TESTING

If inspection is perfect it is possible to obtain a modest further reduction in the number of inspections, without change in probabilities of acceptance by simply noting that if, on individual testing, no item is found to be NC among the first (n_0-1) tested, then the remaining item must be NC (since the group test indicated existence of at least one such item) and so it need not be tested. If inspection were perfect, the resultant reduction in expected number of tests would be

$$n_0^{-1} \binom{D}{1}\binom{N-D}{n_0-1}/\binom{N}{n_0} = D(N-D)^{(n_0-1)}/N^{(n_0)} = D(N-D-n_0+1)^{-1} P_0(n_0). \qquad (30)$$

(For infinite lot size, $(1-\omega)^{n_0-1}\omega$.)

This reduction will usually be quite small (in fact, less than 1). However, Pfeifer and Enis (1978) have considered a screening situation in which similar curtailment may have substantial advantages. They suppose that group testing indicates not only the existence of NC items but also the number (Y) of NC items in the group. Of course, if $Y = 0$ then no individual testing is required. If $Y \geq 1$ then individual testing is continued until (a) Y items are classified as NC, so that no uninspected items are NC, or (b) for some $W(\geq 1)$ only (Y-W) items are classified as NC among the first (n-W) tested, so that the remaining W items 'must' (if inspection is perfect) be NC; whichever comes first.

Kotz et al. (1985) have analyzed the effects of imperfection in inspection on this procedure. We will summarize their analysis and, in addition, provide some illustrative tables. We use a notation similar to that employed in the preceding Sections, except that p_1, p_1' refer to responses from each item contributing to the 'observed' (estimated) total number (y) of NC items as a result of the group test. We also introduce the notation

Z_1: estimated number of NC items, from group test

Z_2: total number of items which would be classified as NC from individual testing, (supposing this to be done)

M: number of individual tests carried out.

The expected total number of tests is $1 + E[M]$ where

$$E[M] = E_Y[E_{Z_1,Z_2}[E[M|Z_1,Z_2]]|Y]$$

(where $E_W[\cdot]$ denotes "expectation with respect to W")

$$= \sum_y P_y \sum_{z_1=0}^{n_0} \sum_{z_2=0}^{n_0} P(z_1|n_0,z,p_1,p_1')P(z_2|n_0,y,p,p')E[M|z_1,z_2] \qquad (31)$$

where $P_y = \Pr[Y=y] = \binom{D}{y}\binom{N-D}{n_0-y}/\binom{N}{n_0}$ $(\max(0,N-D+n_0) \leq y \leq \min(n_0,D))$, (for infinite lot size, $\Pr[Y=y] = \binom{n_0}{y}\omega^y(1-\omega)^{n_0-y}$ $(0 \leq y \leq n_0)$); and (from the basic distribution (4')).

$$P(z|n,y,a,b) = \sum_{w=0}^{z} \binom{y}{w}\binom{n-y}{z-w} a^w b^{z-w}(1-a)^{y-w}(1-b)^{n-y-z+w} \quad ;$$

and

$$E[M|z_1,z_2] = \begin{cases} (n+1)z_1/(z_2+1) & \text{if } z_2 > z_1 \\ z(n-z)\{(z+1)^{-1} + (n-z+1)^{-1}\} & \text{if } z_2 = z_1 = z \\ (n+1)(n-z_1)/(n-z_2+1) & \text{if } z_2 < z_1 \ . \end{cases}$$

The probability of correct classification of a NC item is

$$PC(NC) = (1-P_0)^{-1} \sum_{y>0} P_y \sum_{z_1=1}^{n} \sum_{z_2'=0}^{n-1} P(z_1|n,y,p_1,p_1')P(z_2'|n-1,y-1,p,p')$$
$$\times \left\{ p \min\left(\frac{z_1}{z_2'+1}, 1\right) + (1-p)\max\left(\frac{z_1 - z_2'}{n - z_2'}, 0\right)\right\} \qquad (32)$$

The probability of correct classification of a conforming item is

$$PC(C) = (1-P_n)^{-1} \sum_{y<n} P_y \sum_{z_1=0}^{n} \sum_{z_2'=0}^{n-1} P(z_1|n,y,p_1,p_1')P(z_2'|n-1,y,p,p')$$
$$\times \left\{ (1-p')\min\left(\frac{n-z_1}{n-z_2'}, 1\right) + p' \max\left(\frac{z_2'-z_1}{z_2'+1}, 0\right)\right\} \qquad (33)$$

(See Kotz et al. (1986) for details of derivation.)

Tables 9-1, 2 and 3 give values of expected numbers of tests ($1+E[M]$), and probabilities of correct classification ((32) and (33)) for lot sizes $N = 50$, and 100, with sample sizes $n = 5(5)15$, numbers of NC items $D = 5,10$ and

Group Test	Individual Test
$p_1 = 0.85, 0.95$	$p = 0.75(0.05)0.95$
$p_1' = 0, 0.05$	$p' = 0(0.025)0.1$

Tables 10.1, 2 and 3 gives values of the same quantities for 'infinite' lot size with proportion of NC items $\omega = 0.05 \quad 0.10$ (corresponding to DN^{-1} for a finite size lot) and the same values of the other parameters.

Examination of the tables bring out the following points:

(i) Expected sample size varies only slightly with variation in (1-p) and p' (the error probabilities for individual testing) but markedly with variation in $(1-p_1)$ and p_1' (the error probabilities for group testing). Of course, if the testing were not curtailed, the expected sample size would not depend on p or p' at all.

(ii) Probabilities of correct classification vary considerably with variation in p and p'. The amount of variation decreases as the chance of individual testing decreases, and so decreases with D.

(iii) For all three tabulated quantities, values for lot size (N) equal to 100 (see Tables 9-1, 2, 3) differ only slightly from those for infinite lot size (Tables 10-1, 2, 3) with $\omega = D/N$. Values for lot sizes greater than 100 (with the same $D/N = \omega$) can be interpolated from the tabulated values for $N=100$ and $N=\infty$. It appears, from an examination of the values for

$N = 50$, $D = 5$
$N = 100$, $D = 10$
and $N = \infty$, $\omega = 0.10$

that harmonic interpolation (linear with respect to 100/N) will give good results for $N \geq 50$, and probably for N as small as 25 (i.e. $100/N = 4$). We note, for example, that with $n = 10$, $p_1 = 0.80$, $p_1' = 0.05$, $p = 0.85$, $p' = 0.10$ we have the following tabulated values:

	1+E[M]	PC(DEF)	PC(NONDEF)	(100/N)
$N = 50$, $D = 5$	2.4478	0.5950	0.6813	(2)
$N = 100$, $D = 10$	5.4162	0.5978	0.6728	(1)
$N = \infty$, $\omega = 0.10$	5.3864	0.5991	0.6649	(0)

In each case the value for $N = 100$, $D = 10$ differs from the arithmetic mean of the values for $N = 50$, $D = 5$ and $N = \infty$, $\omega = 0.10$ by no more than 0.001.

(See Johnson et al. (1985) for application of harmonic interpolation to the basic distribution (4')).

ACKNOWLEDGEMENT

Dr. Kotz's work was supported by the Office of Naval Research under contract N00014-84-K-0301.

REFERENCES

Dorfman, R. (1943) The detection of defective members of large populations, Ann. Math. Statist. **14**, 436-440.

Hwang, F.K. (1984) Robust group testing, J. Qual. Technol., **16**, 189-195.

Johnson, N.L. and Kotz, S. (1985) Some multivariate distributions arising in faulty sampling inspection. In Multivariate Analysis - VI (P.R. Krishnaiah, ed.) North Holland, Amsterdam, New York and Oxford, pp. 311-325.

Johnson, N.L., Kotz, S. and Rodriguez, R.N. (1985, 1986) Statistical effects of imperfect inspection sampling, J. Qual. Technol., **17**, 1-31, **18**.

Kotz, S. and Johnson, N.L. (1982) Errors in inspection and grading: Distributional aspects of screening and hierarchal screening, Commun. Statist. - Theor. Meth., **11**, 1997-2016.

Kotz, S., Song, M.S. and Johnson, N.L. (1986) Effects of inspection errors on curtailed Dorfman-type procedures, Commun. Statist. Theor. Meth., **15**.

Mundel, A.B. (1984) Group testing, J. Qual. Technol., **16**, 181-188.

Pfeifer, C.G. and Enis, P. (1978) Dorfman-type group testing for a modified binomial model, J. Amer. Statist. Assoc., **73**, 588-592.

Wolf, J.K. (1985) Born again group testing: Multiaccess communications, IEEE Trans. Inf. Theory, **IT-31**, 185-191.

T7

Table 6: Simple Screening

p_1	p_1'	p	p'	ω	n_1	$pc(c) =$ PROBABILITY OF CORRECT CLASSIFICATION OF CONFORMING	$100(1 - E/n_0) =$ EXPECTED PERCENT REDUCTION IN TESTS
0.95	0.05	0.95	0.05	0.05	6	0.9873	54.5
					8	0.9839	52.2
					10	0.9804	48.9
					12	0.9781	45.3
				0.1	6	0.9791	36.2
					8	0.9740	31.2
					10	0.9699	26.4
					12	0.9666	22.1
				0.2	6	0.9672	11.9
					8	0.9619	7.6
					10	0.9589	4.7
					12	0.9564	2.9

T2

Table 7: Two-Stage Hierarchal Screening

p_0	p_D'	p_1	p_1'	p'	ω	n_1	n_2	Pc(c) = PROBABILITY OF CORRECT CLASSIFICATION OF CONFORMING	EXPECTED PERCENT REDUCTION IN TESTS(%)
0.95	0.05	0.95	0.05	0.05	0.05	6	2	0.9972	59.2
						6	3	0.9954	60.2
						12	2	0.9968	57.8
						12	3	0.9947	61.7
					0.1	6	2	0.9947	40.9
						6	3	0.9908	42.0
						12	2	0.9941	37.2
						12	3	0.9902	41.8
					0.2	6	2	0.9898	13.4
						6	3	0.9830	14.3
						12	2	0.9893	11.9
						12	3	0.9825	15.9

Table 7: Two-Stage Hierarchal Screening

P_D	P_D'	P_1	P_1'	p'	ω	n_1	n_2	Pc(c) = PROBABILITY OF CORRECT CLASSIFICATION OF CONFORMING	EXPECTED PERCENT REDUCTION IN TESTS(%)
0.95	0.10	0.95	0.05	0.05	0.25	6	2	0.9971	57.2
						6	3	0.9953	58.8
						12	2	0.9967	56.3
						12	3	0.9946	60.7
					0.1	6	2	0.9946	39.4
						6	3	0.9907	41.0
						12	2	0.9941	36.4
						12	3	0.9902	41.3
					0.2	6	2	0.9898	12.7
						6	3	0.9830	13.8
						12	2	0.9893	11.7
						12	3	0.9825	15.8

Note that in the last set the probabilities of correct classification of nondefectives are very slightly greater (no more than 0.0001) than the corresponding probabilities for the penultimate set, while the expected reduction in number of tests is slightly (c. 1-2%) less. (This comparison reflects the effect of changing p_D' from 0.05 to 0.10).

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a. SECURITY CLASSIFICATION AUTHORITY:

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE:

3. DISTRIBUTION/AVAILABILITY OF REPORT:

4. PERFORMING ORGANIZATION REPORT NUMBER(S):

5. MONITORING ORGANIZATION REPORT NUMBER(S):

6a. NAME OF PERFORMING ORGANIZATION:

6b. OFFICE SYMBOL (If applicable):

7a. NAME OF MONITORING ORGANIZATION: Office of Naval Research

6c. ADDRESS (City, State and ZIP Code):
Dep't. of Management and Statistics
University of Maryland
College Park, Md. 20742

7b. ADDRESS (City, State and ZIP Code):

8a. NAME OF FUNDING/SPONSORING ORGANIZATION: ONR

8b. OFFICE SYMBOL (If applicable):

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-84-K-0301

8c. ADDRESS (City, State and ZIP Code):
Stat. and Probability Program
Office of Naval Research
Arlington, VA 22217

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE (Include Security Classification): "Statistical Effects of Imperfect Inspection Sampling (III. Screening Group Testing)"

12. PERSONAL AUTHOR(S): Johnson, Norman L.; Kotz, Samuel; Rodriguez, Robert N.

13a. TYPE OF REPORT: technical

13b. TIME COVERED: FROM 9/85 TO 8/86

14. DATE OF REPORT (Yr., Mo., Day): January 1986

15. PAGE COUNT: 13 pgs.

16. SUPPLEMENTARY NOTATION:

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Acceptance Sampling Plans, Binomial Distribution, Compound Distributions, Hypergeometric Distributions.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper, the third in a series, describes the effects of imperfect inspection on screening (group testing) acceptance sampling procedures. These procedures are based on first grouping items and testing the group as a whole for presence or absence of nonconforming (NC) items proceeding to individual testing only if such presence is indicated. If sampling is perfect, the procedure can reduce the expected amount of testing, especially if the proportion of NC items is low. Modifications of the basic procedure - curtailed and hierarchal group testing - aimed at further reduction in the expected amount of testing, are also discussed. Some new formulas are presented.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT

UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION:

22a. NAME OF RESPONSIBLE INDIVIDUAL: Samuel Kotz

22b. TELEPHONE NUMBER (Include Area Code): 301 454-6805

22c. OFFICE SYMBOL:

END

DTIC

8-86