y====(2 34 4 ----Al Ø

FILE COPY

JE .

EFFECTS OF FALSE AND INCOMPLETE IDENTIFICATION OF DEFECTIVE ITEMS ON THE RELIABILITY OF ACCEPTANCE SAMPLING

by

Samuel Kotz

University of Maryland College Park, Maryland Norman L. Johnson

FLECT

E

149

APR

13

University of North Carolina Chapel Hill, North Carolina

April 1982

ABSTRACT

ocument has been approved

e and sale in

82

04

The effects of false and incomplete identification of nonconforming items on the properties of two-stage acceptance sampling procedures are studied. Numerical tables are presented, and there is some discussion of sensitivity to inspection errors. Methods of taking into account extra costs needed to implement better inspection techniques, when initial grading is inconclusive, are described.

I. INTRODUCTION

Recently, we have considered a number of distributions arising from inspection sampling, when inspection may fail to identify a defective item, or may erroneously classify a nondefective item as 'defective'. (Johnson et al. (1980), Johnson & Kotz (1981), Kotz & Johnson (1982a)). Our interest in these papers was mainly in the distributions (of numbers of items classified as defective) themselves. We now consider some consequences, with special regard to properties of acceptance sampling schemes. Although this is the main purpose of the present paper, we will incidentally encounter some further compound distributions which are of interest on their own account.

We also consider a simple grading situation, allowing for a possible second inspection when first inspection fails to decide whether an item is or is not defective, and introducing some cost functions.

We will suppose sampling is carried out, without replacement, from a lot of size N which contains D defect¹ ve items. The symbol Y (possibly with subscripts) will denote the number of defective items included in a random sample (without replacement) and Z (with subscripts) the number of items classified as 'defective' after inspection.

2. SINGLE-STAGE ACCEPTANCE SAMPLING

Single-stage acceptance sampling schemes have the following simple rule:

"If the number of (alleged) defective items in a sample of size <u>n</u> exceeds <u>a</u>, reject the lot; otherwise accept it." Formally:

"Reject if $Z > \underline{a}$; accept if $Z \leq \underline{a}$ ".

, Y

In order to assess the properties of this procedure, we need only the distribution of Z, which was obtained in Johnson & Kotz (1981) - namely $\Pr[Z=z \ p,p';D] = {\binom{N}{n}}^{-1} \sum_{y} {\binom{D}{y}} {\binom{N-D}{n-y}} \sum_{j=0}^{Z} {\binom{y}{j}} {\binom{n-y}{z-j}} p^{j} (1-p)^{y-j} p^{j} z^{-j} (1-p')^{n-y-2+j}$

$$= {\binom{N}{n}}^{-1} \sum_{y} {\binom{D}{y}} {\binom{N-D}{n-y}} b(z;y,n-y;p,p')$$
(1)

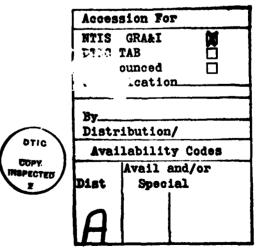
where p = probability that a defective item is detected on inspection and p'= probability that a nondefective item is classified as 'defective', and $max(0,n-N+D) \le y \le min(n,D)$.

In the construction of acceptance sampling schemes (that is, choosing the values of n and a) it is (usually) assumed that inspection is faultless, that is p = 1 and p' = 0. The values of n and a are then chosen to make

 $Pr[Z > a|1,0;D_0] \stackrel{\bullet}{\rightarrow} \alpha$ (the 'Producer's Risk')

while $Pr[Z \le a | 1,0;D^*] \stackrel{*}{\rightarrow} \beta$ (the 'Consumer's Risk')

where α , β , D_0 and D^* are parameters chosen in accordance with the specific circumstances.



-2-

3. TWO-STAGE ACCEPTANCE SAMPLING

These procedures (see e.g. Dodge & Romig (1959)) are of form: "Take a random sample (without replacement) of size n_1 , and observe the number of apparently defective items, Z_1 . If $Z_1 \le a_1$ accept the lot; if $Z_1 > a'_1$, reject the lot; if $a_1 < Z_1 \le a'_1$, take a further random sample, from the remaining items in the lot, of size n_2 and observe the number of apparently defective items in it, Z_2 .

If $Z_1 + Z_2 \le a_2$ accept the lot; if $Z_1 + Z_2 > a_2$ reject it." Formally:

"Accept if $Z_1 \le a_1$, or if $a_1 < Z_1 \le a_1$ and $Z_1 + Z_2 \le a_2$; otherwise reject."

(Popular special cases are $n_2 = n_1$, or $n_2 = 2n_1$ and/or $a_2 = a_1^{\dagger}$)

To assess the properties of this procedure we need the joint distribution of Z_1 and Z_2 . Conditionally on the actual numbers Y_1 , Y_2 of defective items in the two samples, Z_1 and Z_2 are independent, and (for i = 1,2) Z_1 is distributed as the sum of two independent binomial variables with parameters (Y_1, p) and $(n_1 - Y_1, p')$ corresponding to items correctly and incorrectly classified as defective, respectively. Formally

 $Z_i | Y_1, Y_2 \sim \text{Binomial } (Y_i, p) * \text{Binomial } (n_i - Y_i, p')$ (2) (* denotes convolution.)

The joint distribution of Y_1 and Y_2 is a bivariate hypergeometric with parameters $(n_1, n_2; D, N)$ and

$$\Pr[Y_{1} = y_{1}; Y_{2} = y_{2}] = \begin{pmatrix} n_{1} \\ y_{1} \end{pmatrix} \begin{pmatrix} n_{2} \\ y_{2} \end{pmatrix} \begin{pmatrix} N-n_{1}-n_{2} \\ D-y_{1}-y_{2} \end{pmatrix} / \begin{pmatrix} N \\ D \end{pmatrix}$$
(3)

 $(0 \le y_1 \le n_1; D-N+n_1+n_2 \le y_1 + y_2 \le D).$

-3-

The unconditional distribution of (Z_1, Z_2) is a mixture of (2) with mixing distribution (3).

Formally, then

 $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim \begin{pmatrix} \text{Binomial } (Y_1,p) & \text{Binomial } (n_1 - Y_1,p') \\ \text{Binomial } (Y_2,p) & \text{Binomial } (n_2 - Y_2,p') \end{pmatrix} \wedge \begin{array}{c} \text{Biv. Hypg}(n_1,n_2;D,N) & (4) \\ Y_1,Y_2 \end{pmatrix}$

(A denotes the compounding operator (e.g. Johnson & Kotz (1969, p. 184)).)

It would be straightforward to generalize this formula to allow p and p' to vary from sample to sample. (see Johnson & Kotz (1982b)). This will not be done here, as it appears reasonable to suppose p and p' are the same for both the first and second sample.

Explicitly

$$\Pr[Z_{1} = z_{1}, Z_{2} = z_{2} | p, p'; D] = \left\{ \sum_{\substack{y_{1} \\ y_{1} \\ y_{2}}} \sum_{\substack{y_{1} \\ y_{2}}} \frac{\binom{n_{1}}{y_{1}} \binom{n_{2}}{y_{2}}}{\binom{N-n_{1}-n_{2}}{D-y_{1}-y_{2}}} b(z_{1}; y_{1}, n_{1}-y_{1}; p, p') b(z_{2}; y_{2}, n_{2}-y_{2}; p, p') (5) \right\}$$

(Limits for y_1, y_2 as in (3)).

The expected number of items inspected is

 $n_1 + n_2 \Pr[a_1 < Z_1 \le a_1']$.

This can be evaluated using the distribution of Z_1 , which is of the same form as (1), with subscript 'l' attached to n and z. The probability of acceptance at first sample is

$$\sum_{z=0}^{n_{1}} \sum_{y_{1}}^{\binom{n_{1}}{y_{1}}} \frac{\binom{N-n_{1}}{D-y_{1}}}{\binom{N}{D}} b(z_{1}y_{1},n_{1}-y_{1};p,p').$$
 (5)

The probability of acceptance at second sample is the sum of probabilities (5) over $a_1 < Z_1 \le a_1'$ and $Z_1 + Z_2 \le a_2$. The distribution of $Z_1 + Z_2$ is

$$Z_1 + Z_2 \sim \frac{2}{i=1}^{2} Binomial (Y_i,p)_{i=1}^{*} Binomial(n_i - Y_i,p') \wedge Biv.Hypg(n_1,n_2;D,N)$$
 (6)
 Y_1,Y_2

but it is not directly applicable to calculation of this probability. The acceptance probability is calculated directly as the sum of

$$\sum_{z_1=a_1+1}^{a_1} (1+a_2-z_1) = \frac{1}{2}(a_1'-a_1)(2a_2-a_1-a_1'+1) \text{ terms of type (5)}.$$

Acceptance probabilities for four sampling schemes, with lot sizes N = 100, 200 and defective fractions D/N = 0.05, 0.1, 0.2 are shown in Table 1 for p = 1.00, C.98, 0.95, 0.90, 0.75 and p' = 0.00, 0.01, 0.02, 0.05, 0.10. The sampling schemes have $n_1 = n_2$, the common value corresponding to sample size codes D-G of Military Standard 105D for double sampling (see Duncan (1974)).

As is to be expected, the acceptance probability increases as p decreases, and decreases as p' increases. The latter effect is relatively greater, for the values of p and p' used (which correspond to the situations most likely to be encountered). For a given defective fraction (D/N) probabilities of acceptance for lot sizes N = 100 and N = 200 do not differ much. It is noteworthy that the change with increasing N is sometimes positive and sometimes negative.

When D is small, variation in p has less effect, because it is only the D defectives that are affected. For converse reasons, variation in p' has greater effect when D is small. Effects of changes in p and p' become more marked as the sample size increases.

Roughly speaking, it appears that values of p as low as 95% do not have drastic effect on acceptance probability, but values of p^{\dagger} even as small as 1% do have a noticeable effect.

-5-

4. COST CONSIDERATIONS IN GRADING INDIVIDUAL ITEMS

The topic of grading was discussed by Kotz and Johnson (1982). This differs from acceptance sampling in that we are primarily concerned with the classification assigned to *each* item individually, rather than using the apparent total number of defective items in a sample as a criterion for accepting or rejecting the lot from which it was drawn.

The simplest possible situation to consider is when a single individual is chosen at random and assigned to one of two classes "defective" or "nondefective". (This decision is restricted to the particular item at hand it is not extended to the whole lot.) A natural extension is obtained by allowing for the possibility that on first inspection, no clear decision will be reached - but that this can be resolved, one way or the other, by a second more careful (and probably more efficient and more costly) inspection.

We now introduce π , π' to denote the probability of no decision on first inspection for a defective, nondefective item respectively. Also let p_E , p_E' (E for "expensive") denote the probability that a defective, or nondefective item respectively is classified as 'defective' at the second inspection. Then the probability of a defective item being correctly classified is $(p + \pi p_E)$, and the probability of a nondefective being incorrectly classified as defective is $(p' + \pi' p_E')$. (Note that all the formulae in Section 2 and 3 are still applicable, with p replaced by $(p + \pi p_E)$ and p' by $(p' + \pi' p_E')$.)

Some new points arise if cost is taken into consideration. If c_1 is the cost of the first inspection and c_2 that of the second, the expected cost of inspection for an individual chosen at random from a lot of N items, of which D are defective, is

$$C = c_1 + \left\{ \frac{D}{N} \pi + (1 - \frac{D}{N}) \pi' \right\} c_2 \quad . \tag{7}$$

-6-

If ρ denotes the cost of failing to detect a defective item, and ρ' the cost of classifying a nondefective item as 'defective' then the expected cost of the procedure, per item is

$$R = c_1 + \left\{ \frac{D}{N} \pi + (1 - \frac{D}{N}) \pi' \right\} c_2 + (1 - p - \pi p_E) \frac{D}{N} \rho + (p' + \pi' p_E) (1 - \frac{D}{N}) \rho' \quad . \tag{8}$$

If there is some choice in regard to the amount of effort devoted to second inspections, we say be able to regard p_E and p'_E as functions of c_2 . We would expect p_E to increase and p'_E to decrease with c_2 . We would also expect to have

$$c_2 > c_1$$
, $p_E > p$ and $p'_E < p'$.

If we also able to give reasonably relevant values to ρ and ρ' we can try to minimize R by appropriate choice of c_2 , by using the value of c_2 satisfying

$$\frac{\partial R}{\partial c_2} = 0, \quad \text{that is}$$

$$\frac{D}{N} \pi + (1 - \frac{D}{N}) \pi' = \rho \pi \frac{D}{N} \cdot \frac{\partial p_E}{\partial c_2} + \rho' \pi' (1 - \frac{D}{N}) \frac{\partial p'_E}{\partial c_2} \quad . \quad (9)$$

If $\partial p_E / \partial c_2 > 0$ and $\partial p'_E / \partial c_2 < 0$, as is to be expected, this equation can have no more than one root in c_2 .

The possibility of using this approach may be rather difficult in practice. In particular, assessment of values of ρ and ρ' requires a very considerable knowledge of the likely financial effects of misclassification. Generally, ρ will reflect the adverse results of accepting a defective item which will commonly have high variability consequent on the actual effects of failure when (and if) it occurs. On the other hand, ρ' corresponds to the loss incurred to the producer by rejecting an item which is really satisfactory, and is likely to be less variable. In this section our aim has been to alert practitioners to the existence of rather straightforward procedures, which, coupled with adequate practical experience can yield helpful results in a variety of applications.

ACKNOWLEDGEMEN T

Samuel Kotz's work was supported by the U.S. Office of Naval Research under Contract N00014-81-K-0301. Norman L. Johnson's work was supported by the National Science Foundation under Grant MCS-8021704.

REFERENCES

- 1. Dodge, H.F. and Romig, H.G. (1959) Sampling Inspection Tables, Wiley, New York. (Second Edition)
- 2. Duncan, A.J. (1974) Quality Control and Industrial Statistics, R.D. Irwin, Homewood, Ill. (Fourth Edition).
- 3. Johnson, N.L. and Kotz, S. (1969) Distributions in Statistics Discrete Distributions, Wiley, New York.
- Johnson, N.L. and Kotz, S. (1981) Faulty Inspection Distributions -Some Generalizations. Institute of Statistics Mimeo Series #1335, University of North Carolina at Chapel Hill. (To be published in Proc. of ONR/ARO Reliability Workshop, April 1981).
- 5. Johnson, N.L. and Kotz, S. and Sorkin, H.L. (1980) Faulty Inspection Distributions, Common. Statist. <u>A9</u>, 917-922.
- 6. Kotz, S. and Johnson, N.L. (1982a) Errors in Inspection and Grading: Distributional Aspects of Screening and Hierarchal Screening, Institute of Statistics Mimeo Series #1385, University of North Carolina at Chapel Hill: (to appear in Commun. Statist. <u>All</u> (1982)).
- 7. Kotz, S. and Johnson, N.L. (1982b) Some Distributions Arising in Multistage Sampling with Inspection Errors, submitted for publication.

A CONTRACT

a statute

ł

0.6089	0.7266	0.7935	0.8146	0.8350	0.6089	0.7280	0.7959	0.8175	0.8383	0.75
0.5201	0.6388	0.7099	0.7332	0.7561	0.5191	0.6393	0.7118	0.7355	0.7589	0.90
0.4911	0.6089	0.6807	0.7043	0.7277	0.4896	0.6089	0.6821	0.7063	0.7303	0.95
0.4739	0.5909	0.6629	0.6867	0.7104	0.4721	0.5906	0.6640	0.6884	0.7126	0.98
0.4625	0.5789	0.6509	0.6749	0.6987	0.4605	0.5784	0.6518	0.6763	0.7007	1.00
		200, D = 40	2				D = 20	N = 100,		
0.7949	0.9009	0.9481	0.9606	0.9713	0.7960	0.9030	0.9507	0.9633	0.9741	0.75
0.7546	0.8698	0.9248	0.9401	0.9538	0.7559	0.8726	0.9285	0.9440	0.9579	0.90
0.7407	0.8586	0.9161	0.9323	0.9469	0.7420	0.8616	0.9201	0.9366	0.9514	0.95
0.7325	0.8517	0.9106	0.9274	0.9425	0.7336	0.8548	0.9148	0.9319	0.9473	0.98
0.7266	0.8470	0.9068	0.9240	0.9395	0.7279	0.8502	0.9112	0.9286	0.9445	1.00
	,-	200, D = 20	Z				D = 10	N = 100,		
0.8735	0.9579	0.9865	0.9923	0.9963	0.8744	0.9592	0.9877	0.9934	0.9974	0.75
0.8570	0.9481	0.9812	0.9884	0.9938	0.8582	0.9500	0.9832	0.9903	0.9955	0.90
0.8513	0.9446	0.9793	0.9870	0.9928	0.8526	0.9467	0.9815	0.9891	0.9947	0.95
0.8478	0.9425	0.9780	0.9860	0.9921	0.8493	0.9447	0.9804	0.9883	0.9942	0.98
0.8455	0.9410	0.9772	0.9854	0.9917	0.8470	0.9433	0.9796	0.9878	0.9939	1.00
		200, D = 10	N H				D = 5	N = 100, 1		
0.10	0.05	0.02	0.01	0	0.10	0.05	0.02	0.01	p =0	ď
			~	U, a1 = a2 =	" " " "	$n_1 = n_2$				
) }	 					

-9-

TABLE 1: ACCEPTANCE PROBABILITIES

(See formula (5) and (5)')

	0.4199	0.5183	0.5531	0.5887	0.2777	0.4160	0.5155	0.5509	0.5872	0.75
0.1990	0.3103	0.3945	0.4254	0.4575	0.1933	0.3034	0.3878	0.4189	0.4515	0.00
0.1758	0.2781	0.3568	0.3860	0.4166	0.1696	0.2702	0.3488	0.3781	0.4090	0.95
0.1629	0.2598	0.3353	0.3634	0.3929	0.1564	0.2515	0.3264	0.3546	0.3843	0.98
0.1547	0.2481	0.3213	0.3487	0.3775	0.1480	0.2394	0.3120	0.3393	0.3683	1.00
	10	200, D = 40	N II				D = 20	N = 100,		
0.5241	0.7233	0.8366	0.8705	0.9015	0.5230	0.7248	0.8407	0.8756	0.9076	0.75
0.400/	0.0500	0.7765	0.8144	0.8503	0.4583	0.6572	0.7806	0.8198	0.8572	0.90
0.440	0.6340	0.7552	0.7941	0.8313	0.4374	0.6342	0.7590	0.7994	0.8383	0.95
0.4281	0.6204	0.7422	0.7816	0.8195	0.4249	0.6203	0.7458	0.7868	0.8265	0.98
0.4202	0.6114	0.7334	0.7731	0.8115	0.4167	0.6110	0.7369	0.7782	0.8184	1.00
	10	200, D = 20	Z				D = 10	N = 100,		
0.00	0.0040	0.2000	10.5.01	0,9000	0.00//	0.0000	U.9539	6.76.0	0.9887	0.75
0.0341	0.83/3	0.9528	0.9565	0,9754	0.6343	0.8406	0.9379	0.9620	0.9811	0.90
0.022	0.8281	0.9264	0.9514	0.9716	0.6230	0.8315	0.9320	0.9575	0.9780	0.95
0.010.	0.8224	0.9224	0.9481	0.9692	0.6163	0.8259	0.9283	0.9546	0.9760	0.98
0.6118	0.8185	0.9197	0.9459	0.9675	0.6117	0.8221	0.9258	0.9527	0.9747	1.00
	,0	200, D = 10	N				D = 5	N = 100, 1		
0.10	0.05	0.02	0.01	0	0.10	0.05	0.02	0.01	p'=0	ď

-10-

.

 $n_1 = n_2 = 8; a.$

H

 $0, a_1' = a_2 = 2$

 $n_1 = n_2 = 13; a_1 = 0, a_1' = a_2 = 2$

-11-

____.

0.0848	0.2114	0.3379	0.3892	0.4445	0.0765	0.1965	0.3222	0.3745	0.4323	0.75
0.0260	0.0//4	0.1389	0.2116	0.2500	0.0203	0.0849	0.1534	0.1849	0.2217	0.90
0.0220	0.0653	0.1188	0.1435	0.1724	0.0162	0.0498	0.0940	0.1153	0.1407	0.98
0.0193	0.0581	0.1067	0.1294	0.1560	0.0139	0.0432	0.0823	0.1013	0.1241	1.00
	10	200, D = 40	Z #				D = 20	N = 100,		
0.3518	0.6720	0.8557	0.9035	0.9419	0.3460	0.6766	0.8703	0.9203	0.9592	0.75
0.2633	0.5587	0.7633	0.8251	0.8798	0.2528	0.5582	0.7793	0.8469	0.9064	0.90
0.2370	0.5204	0.7280	0.7934	0.8530	0.2248	0.5171	0.7430	0.8156	0.8816	0.95
0.2220	0.4976	0.7060	0.7733	0.8355	0.2088	0.4924	0.7200	0.7951	0.8649	0.98
0.2123	0.4824	0.6911	0.7594	0.8232	0.1985	0.4759	0.7041	0.7808	0.8530	1.00
	10	200, D = 20	N H				D = 10	N = 100,		
0.5783	0.8912	0.9824	0.9939	0.9988	0.5786	0.8981	0.9875	0.9970	1.0000	0.75
0.5229	0.8550	0.9705	0.9880	0.9970	0.5222	0.8652	0.9797	0.9944	6666.0	0.90
0.5045	0.8417	0.9655	0.9854	0.9960	0.5032	0.8530	0.9765	0.9932	8666.0	0.95
0.4935	0.8334	0.9622	0.9836	0.9953	0.4918	0.8453	0.9743	0.9925	8666.0	0.98
0.4862	0.8278	0.9599	0.9823	0.9948	0.4842	0.8401	0.9728	6166.0	0.9998	1.00
	10	200, D = 10	N N				D = 5	N = 100, D		
0.10	0.05	0.02	0.01	0	0.10	0.05	0.02	0.01	0='q	טי

 $n_1 = n_2 = 20; a_1 = 1, a_1 = 4, a_2 = 5$

La de Artaria Itan

Same dame

-12-

•

0.0193 0.0220 0.0266 0.0362 0.0848	0.0581 0.0653 0.0774 0.1018 0.2114	0.1067 0.1188 0.1389 0.1778 0.1778 0.3379	0.1294 0.1435 0.1669 0.2116 0.3892	0.1560 0.1724 0.1992 0.2500 0.4445	0.0139 0.0162 0.0203 0.0291 0.0765	0.0432 0.0498 0.0613 0.0849 0.1965	0.0823 0.0940 0.1138 0.1534 0.3222	0.1013 0.1153 0.1387 0.1849 0.3746	0.1241 0.1407 0.1683 0.2217 0.4323	1.00 0.98 0.95 0.90 0.75
	0	200, D = 40	N N				D = 20	N = 100,		
0.2633 0.3518	0.5587 0.6720	0.7633 0.8557	0.9035	0.8798 0.9419	0.2528 0.3460	0.5582 0.6766	0.7793 0.8703	0.8469 0.9203	0.9064 0.9592	0.90 0.75
0.2123	0.4824 0.4976	0.6911 0.7060	0.7594 0.7733	0.8232 0.8355	0.1985 0.2088 0.7748	0.4759 0.4924 0.5171	0.7041 0.7200 0.7430	0.7808 0.7951 0.8156	0.8530 0.8649	1.00 0.98 0.95
		200, D = 20	N H				D = 10	N = 100,		
0.4935 0.5045 0.5229 0.5783	0.8334 0.8417 0.8550 0.8912	0.9622 0.9655 0.9705 0.9824	0.9836 0.9854 0.9880 0.9939	0.9953 0.9960 0.9970 0.9988	0.4918 0.5032 0.5222 0.5786	0.8453 0.8530 0.8652 0.8981	0.9743 0.9765 0.9797 0.9875	0.9925 0.9932 0.9944 0.9970	0.9998 0.9998 0.9999 1.0000	0.98 0.95 0.95 0.75
	,	200, D = 10					D = 5		0000	-
0.10	0.05	0.02	0.01	o	0.10	0.05	0.02	0.01	p'=0	q

 $= n_2 = 20; a_1 = 1, a_1' = 4, a_2 = 5$

ŗ

-12-