

TESTS FOR MORTALITY AND RECRUITMENT IN A K-SAMPLE
TAG-RECAPTURE EXPERIMENT

BU-422-M

by

July, 1972

K. H. Pollock, D. L. Solomon, and D. S. Robson

Cornell University

Abstract

In many tag-recapture experiments the procedure used for estimation of the population size requires that there be neither mortality nor recruitment between the times of sampling. Often this assumption is unrealistic. Tests for mortality and for recruitment in a K-sample tag-recapture experiment are developed.

The mortality test derived is asymptotically a one-tailed binomial test and is therefore very easy to compute. The asymptotic and exact distributions on which the test is based are compared numerically in the three sample case. The binomial approximation is found to be valid for moderate sample sizes. The test for recruitment is based on a sequence of independent hypergeometric test statistics which may be combined.

TESTS FOR MORTALITY AND RECRUITMENT IN A K-SAMPLE
TAG-RECAPTURE EXPERIMENT

by

K. H. Pollock, D. L. Solomon, and D. S. Robson

Cornell University

SUMMARY

In many tag-recapture experiments the procedure used for estimation of the population size requires that there be neither mortality nor recruitment between the times of sampling. Often this assumption is unrealistic. Tests for mortality and for recruitment in a K-sample tag-recapture experiment are developed.

The mortality test derived is asymptotically a one-tailed binomial test and is therefore very easy to compute. The asymptotic and exact distributions on which the test is based are compared numerically in the three sample case. The binomial approximation is found to be valid for moderate sample sizes. The test for recruitment is based on a sequence of independent hypergeometric test statistics which may be combined.

1. INTRODUCTION

In tag-recapture experiments for estimation of the number of individuals in a population, a typical assumption is that the population is closed (no recruitment or mortality¹) during the sampling period. In some circumstances the validity of

¹ By mortality we mean removals of any type, e.g., death or emigration.

this assumption is not clear from a priori considerations, and a need therefore exists for simple tests for mortality and recruitment.

Section 3 describes a simple asymptotic test for mortality based on the conditional distribution of the recapture data given a statistic which is sufficient under the hypothesis of no mortality. The size of the test is the same for all observed values of this sufficient statistic and the null (conditional) distribution does not involve the unknown population size. For the three sample case, the exact and asymptotic distributions are compared numerically in Section 4 to determine sample sizes required for the approximation to be valid. The section concludes with a discussion of the power of the test. In Section 5 a test for recruitment is developed, again using a distribution conditional on a sufficient statistic.

2. SOME DEFINITIONS

In a K-sample tag-recapture experiment, a population is randomly sampled K successive times. The untagged elements of a sample are serially tagged, and the recapture history of each of the previously tagged elements is recorded. The entire sample is then returned to the population. It is assumed that the force of mortality is the same for tagged and untagged elements. The following notation will be used:

N_i : the population size at the i^{th} sample time, $i = 1, 2, \dots, K-1$ ²

$\underline{N} = (N_1, N_2, \dots, N_{K-1})$

n_i : the number of elements in the i^{th} sample

$n = \sum_{i=1}^K n_i$

R_i : the number of elements in the i^{th} sample that are later recaptured at least once, $i = 1, 2, \dots, K-1$. ($R_K \equiv 0$).

² The parameter N_K is not identifiable.

$$\underline{R} = (R_1, R_2, \dots, R_{K-1})$$

$C_i = \sum_{j=i+1}^K (n_j - R_j)$: the number of distinct elements in all samples after the i^{th} , $i = 0, 1, \dots, K-1$, where $n_j - R_j$ is the number of elements caught in the j^{th} sample and subsequently not recaptured.

The range of R_i is $0, \dots, \min(n_i, C_i)$ and of C_i is $0, \dots, (n_{i+1} + \dots + n_K)$.

3. A TEST FOR MORTALITY

Tentatively assuming that there is no recruitment into the population during the period between the first and K^{th} samples, we wish to test the hypothesis that there is no mortality during this period.

The hypothesis of no mortality (given no recruitment) will be referred to as the null hypothesis, H_0 , while the hypothesis that mortality occurs is denoted by H_1 . The problem is thus to test

$$H_0: N_1 = N_2 = \dots = N_{K-1} = N \text{ (unspecified)}$$

vs.

$$H_1: N_1 \geq N_2 \geq \dots \geq N_{K-1}, \text{ at least one inequality strict.}$$

3.1. Some preliminary theory

For $N_1 \geq N_2 \geq \dots \geq N_{K-1}$, Pollock [1972] shows that the joint probability distribution of \underline{R} is

$$P(\underline{R} = \underline{r}; \underline{N}) = \prod_{i=1}^{K-1} \binom{n_i}{r_i} \binom{N_i - n_i}{c_i - r_i} / \binom{N_i}{c_i}. \quad (1)$$

The joint distribution of \underline{R} under H_0 is of the form (1) with $N_i = N$, thus

$$P_{H_0}(\underline{R} = \underline{r}; N) = \prod_{i=1}^{K-1} \binom{n_i}{r_i} \binom{N - n_i}{c_i - r_i} / \binom{N}{c_i}. \quad (2)$$

It is also shown that under the null hypothesis, the total number of recaptures,

$$R. \equiv \sum_{i=1}^{K-1} R_i$$

is a minimal sufficient statistic for the family of distributions (2), while \underline{R} is a minimal sufficient statistic for the family (1). It follows that the distribution of \underline{R} conditional on $R.$ does not involve the unknown population size N under the null hypothesis. In particular, as Pollock [1972] also shows,

$$P_{H_0}(\underline{R} = \underline{r} | R. = r.) = \frac{\prod_{i=1}^{K-1} \binom{n_i}{r_i} \binom{n. - r. - n_i}{c_i - r_i} / \binom{n. - r.}{c_i}}{\sum_{r_1, \dots, r_{K-1}} \prod_{i=1}^{K-1} \binom{n_i}{r_i} \binom{n. - r. - n_i}{c_i - r_i} / \binom{n. - r.}{c_i}}. \quad (3)$$

We could now construct a critical region in the space of \underline{R} (a sufficient statistic under H_1) of size α under (3). However, this distribution is awkward and such a test would involve much computation. We therefore derive an asymptotic approximation to (3) and from it an asymptotic test for mortality.

3.2. Derivation of the asymptotic distributions

We allow the sample sizes n_1, n_2, \dots, n_K to increase at the same rate by setting

$$n_1 = a_1 n, n_2 = a_2 n, \dots, n_K = a_K n, (a_i > 0, i = 1, 2, \dots, K), \quad (4)$$

substituting (4) into (3) and letting $n \rightarrow \infty$. The calculation is performed by replacing each factorial by its Stirling's approximation and simplifying before passing to the limit. Thus for large n we find the distribution (3) is approximately a multinomial distribution (5).

$$P_{H_0}(\underline{R} = \underline{r} | R. = r.) \approx$$

$$\frac{r.!}{r_1! r_2! \cdots r_{K-2}! (r. - r_1 - r_2 - \cdots - r_{K-2})!} p_1^{r_1} p_2^{r_2} \cdots p_{K-2}^{r_{K-2}} (1 - p_1 - p_2 - \cdots - p_{K-2})^{r. - r_1 - r_2 - \cdots - r_{K-2}} \quad (5)$$

$$\text{with } p_i = b_i / \left(1 + \sum_{j=1}^{K-2} b_j\right) \quad i = 1, 2, \dots, K-2,$$

$$\begin{aligned} \text{and } b_i &= a_i (a_{i+1} + \cdots + a_K) / a_{K-1} a_K \quad i = 1, 2, \dots, K-2 \\ &= n_i (n_{i+1} + \cdots + n_K) / n_{K-1} n_K. \end{aligned} \quad (6)$$

Similarly for large n if

$$n_1 = a_1 n, n_2 = a_2 n, \dots, n_K = a_K n \text{ and } N_1 = A_1 n, N_2 = A_2 n, \dots, N_{K-1} = A_{K-1} n$$

then under H_1 the joint distribution of \underline{R} conditional on $R.$ is also approximately multinomial with parameters similar to those in (5). Namely,

$$P_{H_1}(\underline{R} = \underline{r} | R. = r.) \approx$$

$$\frac{r.!}{r_1! r_2! \cdots r_{K-2}! (r. - r_1 - r_2 - \cdots - r_{K-2})!} p_1^{*r_1} p_2^{*r_2} \cdots p_{K-2}^{*r_{K-2}} (1 - p_1^* - p_2^* - \cdots - p_{K-2}^*)^{r. - r_1 - r_2 - \cdots - r_{K-2}} \quad (7)$$

$$\text{with } p_i^* = d_i b_i / \left(1 + \sum_{j=1}^{K-2} d_j b_j\right) \quad i = 1, 2, \dots, K-2$$

$$\begin{aligned} \text{and } d_i &= (A_{i+1} - a_{i+1} \cdots - a_K) / (A_i - a_{i+1} \cdots - a_K) \quad i = 1, 2, \dots, K-2 \\ &= (N_{i+1} - n_{i+1} - \cdots - n_K) / (N_i - n_{i+1} - \cdots - n_K). \end{aligned} \quad (8)$$

To insure that (7) be a legitimate probability function we must assume that the sample and population sizes are such that $N_i \leq \sum_{j=i+1}^K n_j \leq N_{i+1}$ does not hold for any $i = 1, 2, \dots, K-2$. This is required in order that we have $d_i \geq 0$.

This should pose no restriction in applications insofar as the sampling fractions are typically small.

3.3. Derivation of the test procedure

We could construct a test of specified size by appealing to the multinomial approximation (5). However, we now show that a simpler test, based on a binomial distribution, is appropriate. Comparing the two distributions (5) and (7) we observe that

$$\sum_{i=1}^{K-2} p_i = \frac{1}{1 + 1/\sum_{j=1}^{K-2} b_j} \quad \text{and} \quad \sum_{i=1}^{K-2} p_i^* = \frac{1}{1 + 1/\sum_{j=1}^{K-2} d_j b_j}.$$

Now since under H_1 , $N_{i+1} \leq N_i$ (for all $i = 1, 2, \dots, K-2$) with at least one strict inequality, we have $d_i \leq 1$ (for all $i = 1, 2, \dots, K-2$) with at least one strict inequality. It follows that under H_1

$$\sum_{i=1}^{K-2} p_i > \sum_{i=1}^{K-2} p_i^* \quad \text{and thus} \quad p_{K-1} < p_{K-1}^*$$

where

$$p_{K-1} = 1 - \sum_{i=1}^{K-2} p_i \quad \text{and} \quad p_{K-1}^* = 1 - \sum_{i=1}^{K-2} p_i^*. \quad (9)$$

From (5) and (7) it is seen that the marginal distribution of R_{K-1} given R is asymptotically binomial with parameter p_{K-1} under H_0 and p_{K-1}^* under H_1 . Thus a test of H_0 against H_1 may be accomplished by the one-tailed binomial test of

$$H_0': p = p_{K-1} \quad \text{vs.} \quad H_1': p > p_{K-1}$$

using the approximate distribution of R_{K-1} given R , i.e.,

$$P(R_{K-1} = r_{K-1} | R = r) \approx \binom{r}{r_{K-1}} p^{r_{K-1}} (1-p)^{r-r_{K-1}}. \quad (10)$$

Note that the constant p_{K-1} can be easily calculated from the sample sizes by

$$p_{K-1} = \frac{n_{K-1}n_K}{K-1 \sum_{i=1}^{K-1} n_i (n_{i+1} + \dots + n_K)} .$$

For sufficiently large r , the test may be further simplified by appealing to the normal approximation to the binomial.

For a given value of the total number of recaptures, r , the test is based on the proportion, r_{K-1}/r , of those recaptures corresponding to elements of sample $K-1$ which are recaptured in the final sample K . If mortality is operating in the population during the experiment, this fraction would tend to be inflated insofar as the elements of sample $K-1$ face mortality for only one additional sampling period.

For a particular alternative, $\underline{N} = (N_1, N_2, \dots, N_{K-1})$, the conditional power of the binomial test is a tail probability of the distribution (10) with $p = p_{K-1}^*$ defined in (9).

In the next section we provide some numerical results for the case $K = 3$.

4. SOME NUMERICAL RESULTS FOR THE CASE $K = 3$

4.1. Comparison of the exact and asymptotic distributions in the test for mortality

In Section 3.2 we showed that the proposed asymptotic test for mortality depends on the distribution $P_{H_0}(R_{K-1} = r_{K-1} | R = r)$. Here we compare the exact (3) and asymptotic (10) forms of this distribution numerically for the three sample case. On specialization to the case $K = 3$, the distributions (3) and (10) become

$$P_{H_0}(R_2 = r_2 | R_1 = r_1) = \frac{\binom{n_2}{r_2} \binom{n_1}{r_1 - r_2} \binom{n_2 + n_3 - r_2}{n_2}}{\sum_{r_2} \binom{n_2}{r_2} \binom{n_1}{r_1 - r_2} \binom{n_2 + n_3 - r_2}{n_2}} \quad (11)$$

$$P_{H_0}(R_2 = r_2 | R_1 = r_1) \approx \binom{r_1}{r_2} p_2^{r_2} (1 - p_2)^{r_1 - r_2} \quad (12)$$

$$\text{where } p_2 = \frac{n_2 n_3}{n_1(n_2 + n_3) + n_2 n_3}.$$

The range of r_2 in (11) is from zero to the minimum of r_1 , n_2 , and n_3 . In (12) the range of r_2 is zero to r_1 . Thus r_1 must exceed neither n_2 nor n_3 . Sample calculations of the distribution of R_2 indicate that this is not a restriction in practice (see Table 3).

In Table 1 the means and variances of the two distributions are tabulated for differing sample sizes and values of r_1 . We notice that the exact mean is slightly underestimated by the binomial mean but the difference is small for all values of R_1 and differing sample sizes. The exact variance is overestimated by the binomial variance. The difference can be large for the larger values of r_1 within a given sampling scheme (fixed n_1, n_2, n_3). If the binomial variance is multiplied by a correction factor f where

$$f = 1 - \frac{r_1}{n_1}, \quad n_1 = n_1 + n_2 + n_3$$

then the corrected variance always gives a better approximation to the exact variance than does the binomial variance. We notice that f has the same form as the finite population correction for variances in finite population sampling. Preliminary calculations indicate that in the range in which the normal approximation is appropriate the correction factor has little effect.

TABLE 1. COMPARISON OF EXACT AND ASYMPTOTIC DISTRIBUTIONS:

$$P_{H_0}(R_2 = r_2 | R_1 = r_1) \text{ FOR THE CASE } K = 3$$

n ₁	n ₂	n ₃	r.	MEAN		VARIANCE		
				EXACT	ASYMPTOTIC	EXACT	ASYMPTOTIC	CORRECTED ASYMPTOTIC
100	100	100	5	1.67	1.67	1.09	1.11	1.09
			25	8.58	8.33	4.88	5.56	5.10
			50	17.70	16.67	8.31	11.11	9.26
			75	27.51	25.00	10.24	16.67	12.50
25	25	25	5	1.70	1.67	1.02	1.11	1.04
			10	3.48	3.33	1.81	2.22	1.92
			15	5.37	5.00	2.36	3.33	2.66
			20	7.38	6.67	2.65	4.44	3.25
			25	9.53	8.33	2.70	5.56	3.71
50	100	100	10	5.06	5.00	2.30	2.50	2.40
			20	10.25	10.00	4.17	5.00	4.60
			30	15.59	15.00	5.61	7.50	6.60
			40	21.10	20.00	6.62	10.00	8.40
			50	26.77	25.00	7.23	12.50	10.00
200	100	50	10	1.44	1.43	1.19	1.22	1.19
			20	2.90	2.86	2.32	2.45	2.31
			30	4.39	4.29	3.37	3.67	3.36
			40	5.91	5.71	4.35	4.90	4.34
			50	7.46	7.15	5.25	6.12	5.25

Let us look at the distributions more closely for $n_1 = n_2 = n_3 = 100$ (that is, the p_2 of (12) is equal to $1/3$). From Table 1 we see that the variances are very close for $r_1 = 5, 25$ and close for $r_1 = 50$. In Table 2 the distributions are compared in detail for $r_1 = 25, 50, 75$. The distributions show very close agreement for $r_1 = 25$. For $r_1 = 50$ the distributions are still close enough for the asymptotic distribution to be useful. If $r_1 = 75$ when $n_1 = n_2 = n_3 = 100$ it is questionable if the asymptotic distribution is close enough for use in practice.

TABLE 2. COMPARISON OF EXACT AND ASYMPTOTIC DISTRIBUTIONS:

$P_{H_0}(R_2 = r_2 | R_1 = r_1)$ WHEN $n_1 = n_2 = n_3 = 100$ ($p_2 = \frac{1}{3}$)

$r_1 = 25$			$r_1 = 50$			$r_1 = 75$		
EXACT	ASYMPTOTIC	r_2	EXACT	ASYMPTOTIC	r_2	EXACT	ASYMPTOTIC	r_2
0.0000	0.0000	0	0.0016	0.0117	<10	0.0055	0.0867	<20
0.0002	0.0005	1	0.0035	0.0157	10	0.0077	0.0477	20
0.0013	0.0030	2	0.0088	0.0286	11	0.0156	0.0623	21
0.0061	0.0114	3	0.0194	0.0465	12	0.0284	0.0765	22
0.0204	0.0313	4	0.0371	0.0679	13	0.0467	0.0881	23
0.0503	0.0658	5	0.0620	0.0898	14	0.0691	0.0954	24
0.0959	0.1096	6	0.0910	0.1077	15	0.0925	0.0973	25
0.1446	0.1487	7	0.1178	0.1178	16	0.1122	0.0936	26
0.1756	0.1673	8	0.1348	0.1178	17	0.1232	0.0849	27
0.1741	0.1580	9	0.1368	0.1080	18	0.1228	0.0728	28
0.1421	0.1264	10	0.1234	0.0910	19	0.1110	0.0590	29
0.0960	0.0862	11	0.0991	0.0705	20	0.0913	0.0452	30
0.0540	0.0501	12	0.0709	0.0503	21	0.0682	0.0328	31
0.0253	0.0251	13	0.0452	0.0332	22	0.0463	0.0226	32
0.0098	0.0108	14	0.0258	0.0202	23	0.0286	0.0147	33
0.0032	0.0040	15	0.0131	0.0113	24	0.0161	0.0091	34
0.0011	0.0016	>15	0.0060	0.0059	25	0.0083	0.0053	35
			0.0024	0.0028	26	0.0038	0.0030	36
			0.0013	0.0021	>26	0.0026	0.0030	>36

In Table 3 we examine the null distribution of R_1 for $n_1 = n_2 = n_3 = 100$.

Note that for moderate population size (N) the values of R_1 which occur with high probability are those for which the asymptotic distribution is a very good approximation to the exact distribution.

TABLE 3. THE EXACT NULL DISTRIBUTION OF $R_.$ FOR
 $n_1 = n_2 = n_3 = 100$ ($p_2 = \frac{1}{3}$)

N	$P_{H_0}(R. \leq 25)$	$P_{H_0}(R. \leq 50)$	$P_{H_0}(R. \leq 75)$
500	0.0000	0.1431	0.9999
1,000	0.1099	1.0000	1.0000
1,500	0.9306	1.0000	1.0000
2,000	0.9978	1.0000	1.0000

For $K = 3$, $E(R_.) = (n_1 n_2 + n_1 n_3 + n_2 n_3) / N - n_1 n_2 n_3 / N^2$. For arbitrary K , the exact distribution is difficult to calculate but Czen Pin and Dzan Dzo-i [1961] have shown that the distribution may be approximated by a Poisson.

4.2. The power of the test

We now examine the power of the asymptotic test under specific alternatives. The alternative distribution is approximated by (12) with p_2 replaced by

$$p_2^* = \frac{1}{1 + d_1 b_1}, \quad b_1 = \frac{n_1(n_2 + n_3)}{n_2 n_3}, \quad d_1 = \frac{N_2 - n_2 - n_3}{N_1 - n_2 - n_3}$$

and the power conditional on $R_.$ ($\beta(N_1, N_2 | R_.)$) can be calculated from (12) which is binomial ($r_., p_2^*$). The power under specific alternatives is given by

$$\beta(N_1, N_2) = \sum_{r_} \beta(N_1, N_2 | R_.) P_{H_1}(R_.)$$

In Table 4 the power is calculated for N_1 equal to 2,000 and different mortality rates for a selected sampling scheme ($n_1 = n_2 = n_3 = 200$). The size of the non-randomized test for each value of $r_.$ was chosen as close as possible to 0.05. The power is low unless the mortality rate is relatively high.

TABLE 4. APPROXIMATE POWER OF TEST FOR MORTALITY
UNDER SPECIFIC ALTERNATIVES

The test has size as close as possible to 0.05.

SAMPLING SCHEME $n_1 = n_2 = n_3 = 200$

N_1	N_2	% MORTALITY	p_2	p_2^*	POWER
2,000	1,000	50	0.33	0.57	0.99
2,000	1,200	40	0.33	0.50	0.88
2,000	1,400	30	0.33	0.44	0.58
2,000	1,600	20	0.33	0.40	0.29
2,000	1,800	10	0.33	0.36	0.12
2,000	2,000	0	0.33	0.33	0.05

5. TESTS FOR RECRUITMENT

If the model H_1 is accepted then population size N_1 is estimated by the familiar maximum likelihood formula $\hat{N}_1 = n_1 C_1 / R_1$, where n_1 may now be viewed as the number of marked individuals in the population of size N_1 , C_1 is the size of the sample drawn from this population, and R_1 is the number of mark-recaptures in the sample. If recruitment is admitted into the model then an individual counted in C_1 is no longer known to have been counted in N_1 unless this individual bears a tag indicating that it was captured and released on or before the i^{th} sample. The effective sample size for estimating N_1 thus becomes substantially less than C_1 , the number of distinct individuals captured after the i^{th} sample, and the estimator of N_1 (developed by G. M. Jolly [1965]) for this case becomes correspondingly less efficient. In situations where this question is in doubt a statistical test for recruitment thus becomes a critical step in determining the estimation procedure.

A test which is statistically independent of the estimators $\hat{N}_i = n_i C_i / R_i$ under the hypothesis $H_0 \cup H_1$ of "no recruitment" can be constructed by defining a critical region in the sample space of a statistic which is minimal sufficient with respect to the alternate hypothesis H_2 that recruitment does occur, and such that this critical region has size α with respect to the conditional probability distributions determined by fixed values of \underline{R} . With both mortality and recruitment admitted into the model (H_2) the minimal sufficient statistic increases in dimension to include both R_1, \dots, R_{k-1} and S_2, \dots, S_{k-1} , where S_i is the number of tag-recaptures in the i^{th} sample (see Robson [1969]). The required conditional distribution is then given by

$$P_{H_0 \cup H_1}(S_i = s_i | S_2, \dots, S_{i-1}, R_1, \dots, R_{k-1}) = \frac{\binom{n_i}{s_i} \binom{C_{i-1} - n_i}{T_{i-1} - s_i}}{\binom{C_{i-1}}{T_{i-1}}} \quad (13)$$

where T_i is the number of individuals tagged on or before the i^{th} sample that are captured again after the i^{th} sample,

$$T_{i-1} = \sum_{v=1}^{i-1} (R_v - S_v) \quad (T_0 = 0, T_1 = R_1).$$

Note that with respect to recruitment the counts S_i and T_{i-1} do not include any individuals recruited into the population after the $(i-1)^{\text{th}}$ sample whereas n_i and C_{i-1} do include such recruits; n_i includes recruits entering the population between the $(i-1)^{\text{th}}$ and i^{th} sample, and C_{i-1} includes all individuals recruited and captured after the $(i-1)^{\text{th}}$ sample. Under the hypothesis H_1 of no recruitment the fractions

$$\frac{S_i}{T_{i-1}} = \frac{\text{No. captured both before and in the } i^{\text{th}} \text{ sample}}{\text{No. captured both before and in and/or after the } i^{\text{th}} \text{ sample}}$$

$$\frac{n_i}{C_{i-1}} = \frac{\text{No. captured in the } i^{\text{th}} \text{ sample}}{\text{No. captured in and/or after the } i^{\text{th}} \text{ sample}}$$

have equal expectations ($E[S_i/T_{i-1} | C_{i-1}] = n_i/C_{i-1}$); but if recruitment occurs after the i^{th} sample then C_{i-1} becomes inflated relative to n_i , and S_i/T_{i-1} then overestimates n_i/C_{i-1} . An optimal critical region in the space of S_i is therefore the upper tail of the hypergeometric distribution (13), evaluated either exactly from the tables of Lieberman and Owen [1961] or approximately from a binomial or normal approximation.

Under $H_0 \cup H_1$ these $k-2$ one-tailed tests of S_2, \dots, S_{k-1} are statistically independent and hence may be readily combined into a single test; for example, the statistic

$$Z = \frac{1}{\sqrt{k-2}} \sum_{i=2}^{k-1} \left(\frac{S_i}{T_{i-1}} - \frac{n_i}{C_{i-1}} \right) / \sqrt{\frac{n_i(C_{i-1} - T_{i-1})(C_{i-1} - n_i)}{T_{i-1} C_{i-1}^2 (C_{i-1} - 1)}}$$

is approximately distributed as a standard normal deviate under $H_0 \cup H_1$ and hence the critical region $Z > 1.65$ provides a size $\alpha = .05$ test for recruitment. Other alternatives to Z could be constructed by first transforming the observed s_i to $y_i = P_{H_0 \cup H_1}(S_i \leq s_i | C_{i-1}, T_{i-1})$; then, for example, $-2 \sum \ln y_i$ is approximately distributed as chi-square on $2(k-2)$ degrees of freedom, or if Φ is the standard normal c.d.f. then $\sum \Phi^{-1}(y_i) / \sqrt{k-2}$ has approximately the distribution Φ , or $y = \max(y_2, \dots, y_{k-1})$ has the c.d.f. y^{k-2} , etc. As indicated by the earlier heuristic argument the individual one-tailed tests defined by the critical regions $y_i > 1 - \alpha$ provide tests of the hypotheses of no recruitment after the i^{th} sample. Strictly speaking, however, these tests are independent and have size α only if there is no recruitment whatsoever, and hence they should be combined into one test of this null hypothesis.

6. DISCUSSION

We have developed tests of two assumptions typically made in the theory of tag-recapture population estimation. The tests should be applied as follows:

First test $H_0 \cup H_1$ vs. H_2 by the recruitment test of Section 5. If the model H_2 is accepted, then the population sizes are estimated as in Jolly [1965]. If, however, $H_0 \cup H_1$ is accepted, then the test of H_0 vs. H_1 developed in Section 3 is applied. If now H_1 is accepted, the population sizes are estimated by $\hat{N}_1 = n_1 C_1 / R_1$ whereas if H_0 is accepted, a maximum likelihood estimator of the assumed common population size, N , is a solution to $1 - C_0 / \hat{N} = \prod_1 (1 - n_1 / \hat{N})$. (See for example Chapman [1952].) It should be observed that if there is no recruitment, then the two tests proposed here are statistically independent, and that it is only if the hypothesis of no recruitment, $H_0 \cup H_1$, is accepted that the second test is performed.

REFERENCES

- Chapman, D. G. [1952]. Inverse, multiple and sequential sample censuses. Biometrics 8:286-306.
- Czen Pin and Dzan Dzo-i [1961]. On estimating the size of population by capture-mark method. Zastosow. Mat. 6:51-63.
- Jolly, G. M. [1965]. Explicit estimates from capture-recapture data with both death and immigration - stochastic model. Biometrika 52:225-248.
- Lieberman, G. J. and Owen, D. B. [1961]. Tables of the Hypergeometric Probability Distribution. Stanford University Press, 726 pp.
- Pollock, K. H. [1972]. An asymptotic test for mortality in a K sample tag-recapture experiment. M.S. Thesis, Cornell University.
- Robson, D. S. [1969]. Mark-recapture methods of population estimation. New Developments in Survey Sampling. Wiley Interscience, N. L. Johnson and H. Smith, Jr., Editors, pp. 120-140.