A THEORETICAL ANALYSIS OF AN AUDIO-INDEX

Thesis for the Degree of M. S.
MICHIGAN STATE UNIVERSITY
ROBERT GARDNER HEATH
1961

ABSTRACT

A THEORETICAL ANALYSIS OF

AN AUDIO-INDEX

by Robert Gardner Heath

Population indexes based on systematized counts of an identifying call or sound a species--in this report termed "audio-indexes"--are suggested to be potentially cheaper and more efficient than most other indexes of similar scope. Theyyield estimates of ratios of densities among two or more populations of a species; but unless the total effect of the index-determining variables other than density is equalized between comparisons, estimates will be biased. These variables include the hearing ability of observers, the average frequency of individual sounding, the efficiency of sound transmission, and possible human error in counting and judgment.

This study undertakes, by theoretical analysis, the development of a "bias-free" method using audio-counts in a Latin Square design to estimate relative differences in the densities of several populations. It also proposes a counting procedure to improve the standard audio-index.

After arguing the several causes of variation in a count, the study presents a mathematical model for the count of a single interval. The model depicts a count as the product of (1) the area of an observer's hearing coverage, (2) the density of potential sound producers in this area, and (3) the average frequency with which the capable individuals

in this area sound during the interval. Hearing coverage is assumed to be circular, its area depending on an observer's innate hearing ability and the effect of factors that disturb hearing.

The model is then used to develop the Latin square counting procedure. It associates population densities with areas (treatments), maximum hearing coverages with observers (rows), and the daily complexes of individual sounding activity and transmission efficiency with mornings (columns). The design requires that synchronized counts be taken on as many mornings and with as many observers as there are areas, so that each area is counted once a morning, always by a different observer.

The simplest possible example, that of one counting station per area, is used to demonstrate the analysis of variance for the Latin square. The audio-model is multiplicative, however, while the analysis of variance assumes an additive model, so that the analysis of numerical count values is erroneous. Logarithmic transformation of count values produces an additive model, and the analysis of variance of the transformed values is correct.

Expanding the method to areas sampled from a series of stations (the practical case) involves further complexities. To be accurate a counting procedure must yield morning indexes that for each area are a product of its average sample (station) density, its observer's maximum hearing area, and a factor involving sound activity and sound transmission that on a given morning is constant for all areas. A method termed "out-and-back counting" is described, and suggested as one to approximate

closely the desired index, and to offer an improvement over standard audio-indexes. The adaptation of the "out-and-back count" model to the Latin square analysis is routine.

The Latin square counting design removes from density comparisons the effects of differences in observer hearing abilities and in the daily levels of individual sounding frequency and sound transmission. It should prove useful in evaluating game management practices through measurements of relative population changes before and after management.

A THEORETICAL ANALYSIS

OF AN AUDIO-INDEX

by

Robert G. Heath

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Fisheries and Wildlife

1961

Approved: P. I. Tack

ACKNOWLEDGMENTS

I am indebted, foremost, to the Game Division of the Michigan Department of Conservation and in particular to its Chief, Mr. H. D. Ruhl, for sanctioning my part-time graduate studies at Michigan State University while a Game Biologist with the Department. In this capacity I have been permitted to work a full but revised schedule so as to attend one graduate class a term at the University.

Much of this presentation is an outgrowth of my grouse population studies under Pittman-Robertson Project W-40-R, Research in Farm Game Management. Similarly, I developed much of the paper under this Project.

I owe especial gratitude to Dr. L. L. Eberhardt, Game Division Biometrician, who contributed valuable time to discuss and edit the entire statistical and mathematical content of the report, and offer many helpful suggestions. I am similarly indebted to Dr. Philip J. Clark, Assistant Professor of Zoology, Michigan State University, who reviewed the presentation and offered valuable direction during its development.

To Professor George A. Petrides and Professor Peter I. Tack, Department of Fisheries and Wildlife, Michigan State University, my sincere appreciation for their meticulous reviews of the entire study, and for expert suggestions, corrections, and criticisms.

I thank also Dr. D. W. Douglass and Dr. C. T. Black of the Game Division for editing the entire manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Population studies are indispensable in the scientific management of wildlife. They alone serve to measure the intrinsic characteristics of animal populations and the magnitude of population changes, including those resulting from management. The soundness of those managment decisions that depend on such studies can hardly be expected to exceed the soundness of the population measurements themselves.

Measurements dealing with the sizes of populations are of two basic types: the census and the index. The first is a direct enumeration of a given population; the second is a measurement of some quantity, expressed in units of time or area, that is related to population density. Although absolute counts are more desirable, they are generally more difficult and costly to obtain than indexes and are presently often impossible.

The index, then, though often the expedient measurement of animal abundance, may be highly useful. When, for example, the exact correspondence between index and population density is accurately known, the index can be converted to estimate the true size of a population. More frequently it may only be known that the two variables (i.e., index and density) are directly proportional to each other. In this case index ratios will estimate the ratios of actual population densities about as accurately as will comparisons based on censuses. Unfortunately this assumption of direct proportionality cannot always be justified, a fact not obvious and frequently a source of confusion. Comprehensive experimental design of indexes should alleviate this problem, as this paper will attempt to demonstrate.

This study deals with the general category of indexes that are based on systematized counts of an identifying call or sound of a species. Their application is restricted largely to certain important game birds: Familiar examples include counts of the crowing of pheasants (Phasianus colchicus) and the drumming of ruffed grouse (Bonasa umbellus). The term "audio-index" as used here includes all such indexes. It excludes enumerations of individual animals located by tracing their sounds to their respective points of origin.

Audio-indexes are used most frequently in areas sufficiently large to require sampling from a number of counting stations. The typical counting route is along a stretch of road that intersects a representative portion of the area, with counting stations at designated intervals. An observer then travels the route in some convenient manner, stopping at each station once a morning to count sounds for a specified period of time. He may run his route one or more mornings, and his mean count per stop becomes an index of the area's population density. (e.g., Kimball, 1949.)

The advantages of audio- over visual-type indexes are inferred from the supposition that during periods of active sound production, animals are more readily heard than seen or evidenced by visible sign. If the supposition is correct, it follows that for a given effort by an observer working with a given population, audio-indexes would be expected to yield the larger and less variable measurements. Kimball (1949) and Kozicky (1952) both found this to occur for crowing counts of cock pheasants as compared with visual roadside counts. Consequently audio-indexes should permit either more rapid or more complete sampling of an area, or the sampling of larger areas than is feasible with visual indexes. Kimball (1949) reports establishing a pheasant crowing count index equal in reliability to a roadside count index at about one-third the cost.

Audio-indexes are subject to seasonal limitations, as are many other census and index methods. Often they measure only a restricted (i.e., male) segment of a population. Despite these limitations, the potentialities of the audio-count are felt to justify its detailed investigation and analysis.

This report attempts to elucidate a general theory for audio-counts. It is not an evaluation of specific counting procedures in current use. Rather it attempts to point out flaws in the assumption that audio-counts are necessarily proportional to population densities, and proposes a method for overcoming this difficulty in certain circumstances.

The report (1) discusses the variables that determine the value of an audio-count; (2) derives a mathematical model for an audio-count as a function of these variables; (3) proposes a technique to measure relative differences in population densities among at least three areas; and (4) suggests a method for improving the standard audio-index.

The analysis is necessarily somewhat theoretical. As in any attempt to develop a theory, certain basic assumptions must be made; but detailed verification of the assumptions is beyond the scope of the present study.

COMPONENT VARIABLES OF AN AUDIO-COUNT

The number of sounds (usually calls) of a species heard during a specified counting interval involves (1) the number of animals capable of sounding within an observer's area of hearing for the sound, and (2) the average number of sounds produced per capable individual during the interval. (This average includes a zero for each capable individual that fails to sound.) The number of individuals present within the observer's hearing area can be computed from the size of this area and the density of the population therein. All of these factors can be subject to considerable variation.

The shape of a person's hearing area has been suggested by several authors including Kimball (1949), Petraborg (1953), and Dorney (1958) as being circular. In this case the area covered would be equal to π times the square of the hearing distance. The true shape of the area, however, could be somewhat elliptical if one tended to hear further to the sides than either to the front or rear. But if the observer turns his head from side to side through at least a 90° angle--a natural action--a circular area of hearing should be approximated. In the following analysis the hearing area is considered to be circular.

Kimball (1949) states: "The accuracy of the cock pheasant crowing count and the accuracy with which it can be used are dependent largely upon the following factors:

1. Variation in the ability of the individuals conducting the survey to hear cock calls.

2. Daily trend and duration of maximum cock crowing.
3. Seasonal trend and duration of maximum cock crowing.
4. Uniformity of results. [Refers to the amount of variation among count values.]
5. Effect of variable factors, such as weather and cover, upon the count."

It seems logical that these same general sources of variation might operate in all audio-counts, regardless of the animal species involved. While this study views the problem somewhat differently, it is most important, as with any measurement approach, to recognize the sources of variation and take them into account as fully as possible.

SOURCES OF VARIATION--Since conversion of audio-indexes to estimates of actual population densities is subject to a number of difficulties, perhaps the best use of such indexes is in estimating ratios of population abundance. The ratios derived from such comparisons are subject to at least the following sources of variation:

1. Chance variations between counts in the average frequency of sounding of individual animals (hereafter "average frequency of individual sounding") under fixed counting conditions.
2. Real differences within a local population in the average frequency of individual sounding either on different days as affected, for instance, by season or weather, or at different hours during the same day, because of animal behavior patterns, weather, or other factors.
3. Real differences between populations in the average frequency of individual sounding due to weather, climate, length of day, possible sub-specific differences, etc. Such differences

might exist between years on the same area as well as between areas regardless of time.

4. Differences between observers' hearing abilities for a specific sound. If hearing coverage is circular, an observer's maximum area of hearing is proportional to the square of his maximum hearing distance under optimum hearing conditions.

5. Differences in sound transmission (a) between days, (b) within days, and (c) between areas even during synchronized counts. Such factors as wind, vegetation, terrain, etc., may all affect sound transmission. At times competition from foreign noises may actually reduce reception; i.e., a sound is not perceivable above an interfering noise. The net effect on the observer's count, in this case, is the same as a reduction in transmission.

6. Changes in the sensitivity of an observer's hearing, especially between years, but possibly even between days. Attentiveness might be temporarily impaired by fatigue or illness, so that the observer fails to detect normally-audible sounds of low intensity. All such causes of changes in hearing ability are denoted beyond as "physiological factors reducing maximum sound reception."

7. The consistency of an observer's counting. Carney and Petrides (1957), for example, demonstrated that the pheasant crowing counts made by experienced observers were less variable than those of inexperienced observers. An observer may bias his count by

tallying misinterpreted "foreign" sounds and possibly imagined "faint sounds," by miscounting, etc. Obviously, observer consistency adds to the reliability of an index.

From the foregoing discussion it would seem plausible to express an audio-count as a function of the following factors: (1) an observer's maximum hearing distance for a sound (i.e., that functional under optimum hearing conditions); (2) the combined effect of external and physiological factors that reduce hearing efficiency during the count; (3) the area of the observer's circle of hearing, itself a function of the first two factors; (4) the density of the potential sound producing individuals within the effective area of hearing; and (5) the average number of sounds produced by each capable individual within hearing during the counting interval.

This concept permits the formulation of a general mathematical model for audio-counts. Error resulting from false interpretations is excluded from the model and is considered separately. Although the "loudness" (intensity) of the sound emitted by a species will undoubtedly vary among individuals, so that some animals will be audible from greater distances than others, the model assumes that such differences will tend to be self-adjusting and cause no appreciable error.

THE AUDIO-COUNT MODEL

Where an animal issues characteristic sounds which are enumerated by a listening observer let:

N = the number of sounds audible during a single counting interval; i.e. the "perfect" count, devoid of human error.

D = the density (animals per unit area) of individuals within the observer's hearing range and capable of sound production.

F = the average number of sounds produced per capable individual within hearing range during the counting interval.

R = the observer's maximum hearing distance (radius) under optimum listening conditions for the type of sound being recorded.

E = the efficiency of the observer in terms of the proportion of R effective during the counting period. The value of E, always between 0 and 1, depends on the combined effect of external physical factors and possible physiological factors which operate to reduce the observer's maximum hearing distance during the count.

Then

RE = the observer's effective hearing distance during the count, and

$\pi(RE)^2$ = the size of his effective hearing circle.

Also, let

K = a constant for converting the squared lineal units measuring hearing area to the units of area used to express population density. (For example, if $\pi(RE)^2$ were expressed in square chains, but density in animals per acre, then K = 0.1).

Then

$D\pi K(RE)^2$ = the population density times the size of the hearing area, which equals the number of potential sound producers within hearing distance.

Multiplying this quantity by F, the average frequency of individual calling, produces the audio-count equation:

$$N = DF\pi K(RE)^2 .$$

The equation may be worded as follows: The number (N) of sounds heard by an observer during a specific interval is the product of the area of the observer's circle of hearing--$\pi K(RE)^2$--for the sound, the density (D) of individuals within hearing capable of sound production, and the average number of sounds (F) produced per capable individual during the counting interval. Clearly, the recording of false sounds and the failure to record valid ones are sources of human error that will cause the recorded count to be inaccurate and bias the index.

According to the above model, two single-interval audio-counts, say N_1 and N_2, will be proportional to their respective population densities, D_1 and D_2, only if the variable factor $F(RE)^2$ is constant for both counts. Otherwise the relative values of this product would have to be determined for each count, and the count values weighted accordingly in order to be proportional to the densities. Since such determinations would be most difficult, a logical approximation is to make counts under conditions such that the individual values of F, R, and E should be relatively constant. Thus their products, though unknown, should be about equal, making count values approximately proportional to densities. This approximation is usually attempted by taking counts when average individual sounding frequencies are believed to be similar and adequate, by using observers with equal hearing ability or whose hearing relationships are known, and by counting when wind, foreign noises and other factors that disturb sound transmission and reception are at a minimum.

The accuracy of the above theory is contingent on the important assumption that the average individual frequency of sounding is reasonably independent of population density.

Various authors have contemplated this assumption. Kimball (1949), for example, speculates that for cock pheasants, two opposing factors could operate to affect individual crowing frequencies. First, he postulates that, since the crowing of one cock may be stimulated by the crowing of others in the vicinity, individuals may tend to crow more frequently as density increases. Conversely, he states that other observations indicate that very dense populations may produce "such a thing as 'whipped-out' cocks which do not crow at all, [so that] an increase in population would then produce something less than a proportional increase in crowing." Dorney, et al, (1958), studying ruffed grouse in Wisconsin, demonstrated that drumming frequencies of individual ruffed grouse were little if at all affected by the various population levels encountered.

In any situation, the true relationship between density and individual sounding frequency would be difficult to establish in the wild. (Gradual, careful removal of individuals from one area compared to a control area of no removal, coupled with a series of audio-counts synchronized between areas, might indicate whether or not individual sounding frequency is independent of population density.) Herein we will make the seemingly reasonable assumption that these two factors (D and F) are, for practical purposes, independent.

According to the audio model, $N = DF\pi K(RE)^2$, N is a function of the four variables D, F, R and E so that its graph can not be represented in 3-dimensional space. Considering N as a function of each of the four variables in turn (holding the other three constant) reveals that N has a lineal relationship with D and F, while varying as the squares of R and E.

The combined effect of R and E is powerful. If, for example, the values of R and E are each only 10 per cent greater in one count than in another, while D and F are constant in both, then the first count will be almost 50 per cent larger than the second. Or, if R and E had been 50 per cent larger in the above example (an extreme case), the first count would have been five times larger. Further, the factor F (as mentioned) can vary considerably between counts. Quite obviously, indiscriminate use of audio-indexes as though they were proportional to population densities can involve considerable error.

PROPOSED AUDIO-INDEX ANALYSIS USING THE LOGARITHMS OF COUNT VALUES IN A LATIN SQUARE DESIGN

Suppose that audio-indexes are to be used to estimate the ratio of the population levels in three or more areas. Let the areas be reasonably similar, and subject to the same weather and seasonal conditions, so that factors influencing animal activity and sound transmission are about equal in each. By using a Latin square design in counting, the effects of variation between the observers and between the daily calling frequency and sound transmission factors can be controlled in the analysis.

The Latin square design requires making counts on as many days (actually mornings in most such counts) and with as many observers as there are areas, so that each area is counted by only one observer each day but always by a different observer. Then "counts by observers" and "counts by mornings" represent "row" and "column" variables, and "area counts" represent "treatments." Thus, from the audio-count model given above, R^2 is associated with observers, FE^2 with mornings, and D with areas.

The analysis assumes that if daily counts are synchronized between areas, the average individual calling frequencies and sound transmission factors on any given day will be about equal between areas. It also assumes a uniform distribution of individuals within sample populations on any given area. Inconsistencies in either case will tend to increase the amount of experimental error and weaken the precision of the analysis. Human error in counting, which for simplicity has not been included in the model, can bias the analysis as well as decrease its sensitivity.

The simplest possible example illustrating the Latin square analysis is that of three areas (A, B, and C), with only <u>one</u> counting station per area.

Use of three areas, requires that three observers (I, II, and III) make counts on three mornings (1, 2, and 3) under the design restrictions of the Latin square. (The same general procedure applies in using more areas.)

When more than one counting station per area is used, as would normally be the case, special problems arise which will be analyzed beyond. Kempthorne (1952) stresses that in practice a 3x3 Latin square is of little value unless replicated, and recommends the use of larger squares. This argument will not weaken the following demonstration, however, as the same viewpoints apply for larger squares.

The first step is to set up a 3x3 table (an n x n table in the case of n areas) letting observers represent rows, for instance, and mornings represent columns. Areas are then randomized within the table cells under the restriction that each area appear once in each row and each column. The following diagram (Fig. 1) represents the basic arrangement before randomization, where on morning 1, observer I would count Area A, observer II would count Area B, etc.

A Latin Square Design
for Audio-Indexes

		Mornings		
		1	2	3
	I	A	B	C
Observers	II	B	C	A
	III	C	A	B

Fig. 1

The measurements, of course, are the numerical counts recorded by a particular observer in a specific area on a given morning. These are represented by the symbols in Figure 2, where, for example, on morning 1, observer I counted $N_{A,I,1}$ sounds at the station in area A; on morning 3, observer III counted $N_{B,III,3}$ sounds in area B, etc.

Schematic Listing of Audio-Counts
in a Latin Square

		Mornings 1	2	3
Observers	I	$N_{A,I,1}$	$N_{B,I,2}$	$N_{C,I,3}$
	II	$N_{B,II,1}$	$N_{C,II,2}$	$N_{A,II,3}$
	III	$N_{C,III,1}$	$N_{A,III,2}$	$N_{B,III,3}$

Fig. 2

ANALYSIS OF VARIANCE AND NEED FOR AN ADDITIVE MODEL - "The analysis of variance," states Cochran (1947), "depends on the assumptions that the treatment [D] and environmental [R, F, and E] effects are additive and that the experimental errors are independent in the statistical sense, have equal variance, and are normally distributed." The model developed in this study,

$$N = DF\ (RE)^2 \pi K,$$

obviously fails to meet the requirement of additivity, i.e., the variables considered in the analysis (D, F, R and E) appear as a product.

Taking the logarithm of a count converts the model to

$$\underset{\text{(Count)}}{\log N} = \underset{\text{(density)}}{\log D} + \underset{\text{(observer)}}{\log R^2} + \underset{\text{(morning)}}{\log (FE^2)} + \underset{\text{(constant)}}{\log \pi K},$$

and this transformation meets the requirement of additivity. (See Kempthorne, 1952.) The other assumptions for the analysis of variance would require extensive field studies for verification. Intuitively, it would seem that the variables are reasonably independent of one another; e.g., the frequency with which individual animals produce sounds should have no effect on a person's hearing ability.

Let us now compare the numerical and logarithmic analyses of variance where "counts" are computed from assigned hypothetical values of the variables D, R, F, and E. The use of actual audio-counts would not appreciably strengthen the theory unless the true values of the component variables could be ascertained for each count. Such determinations obviously would be extremely difficult to make.

To simplify the demonstration of both analyses, the example has been constructed so that no random variation is included. Thus the hypothetical "counts" are made up from the expected values of the D's, R's, F's, and E's. (See Table 1.) Numerical "count" values computed by the audio-model are shown in Table 2; logarithms of these values appear in Table 3; and the standard analysis of variance (given in any elementary statistics text) is set forth in Table 4 for both sets of data. The numerical "count" values have each been divided by πK for simplicity.

Since no random variation was introduced in computing "count" values, one might expect both error mean squares to be zero. The analysis of variance based on the numerical model, however, distorts the relationships among area, observer, and morning levels, and does develop an error term. The analysis using the logarithmic "count" values develops no error term, as would be expected. In a real situation, obviously, the logarithmic analysis would be expected to include some "experimental error."

In the standard analysis of a Latin square the true means of the variables represented by treatments, rows, and columns are estimated from the averages of their respective observations. As is revealed in Table 5, the arithmetic means of the numerical "count" values for areas, observers, and mornings are not proportional to the assigned values of the appropriate variables (D, R^2, and FE^2, respectively). The table further shows that the

TABLE 1

A SET OF HYPOTHETICAL VALUES FOR THE VARIABLES D, R, F, AND E, USED TO COMPUTE AUDIO-COUNT VALUES FOR A 3X3 LATIN SQUARE

Area densities (D)	Maximum hearing distances (R)	R^2	Ratios of R^2	Morning sounding frequencies (F)	Morning trans-mission reception factors (E)	E^2	Morning count inten-sities FE^2	Ratios of morning count intensities
$D_A = 1$	$R_I = 12$	$R_I^2 = 144$	2.25	$F_1 = .9$	$E_1 = .8$	$E_1^2 = .64$	$(FE^2)_1 = .567$	1.00
$D_B = 2$	$R_{II} = 10$	$R_{II}^2 = 100$	1.56	$F_2 = 1.0$	$E_2 = .9$	$E_2^2 = .81$	$(FE^2)_2 = .810$	1.41
$D_C = 3$	$R_{II} = 8$	$R_{III}^2 = 64$	1	$F_3 = 1.2$	$E_3 = 1.0$	$E_3^2 = 1.00$	$(FE^2)_3 = 1.20$	2.08

TABLE 2

NUMERICAL "COUNT VALUES" COMPUTED FROM TABLE 1.

		Mornings 1	2	3
Observers	I	A/ 82.94	B/ 233.28	C/ 518.40
	II	B/ 115.20	C/ 243.00	A/ 120.00
	III	C/ 110.59	A/ 51.84	B/ 153.60

TABLE 3

COMMON LOGARITHMS OF "COUNT VALUES" OF TABLE 2.

		Mornings 1	2	3
Observers	I	A/ 1.91876	B/ 2.36787	C/ 2.71466
	II	B/ 2.06145	C/ 2.38561	A/ 2.07918
	III	C/ 2.04372	A/ 1.71466	B/ 2.18639

TABLE 4

ANALYSES OF VARIANCE OF HYPOTHETICAL NUMERICAL AND LOGARITHMIC COUNT DATA

Source of variation	Analysis of variance of numerical values: Degrees of freedom	Mean square	"F"	Analysis of variance of logarithmic values; Mean square	"F"
Between areas	2	32,163	6.31	.1746	---
Between observers	2	23,459	4.60	.0933	---
Between mornings	2	19,517	3.83	.0763	---
Error	2	5,096		0	
Total	8				

TABLE 5

NUMERICAL VERSUS LOGARITHMIC COMPARISONS OF HYPOTHETICAL AREA, OBSERVER, AND MORNING COUNT AVERAGES

NUMERICAL

	Areas A	Areas B	Areas C	
Numerical "Count" Values	82.94 120.00 51.84	233.28 115.20 153.60	518.40 110.59 243.00	
Avg. "Count"	84.93	167.36	290.66	
Ratios, Area "Count" Averages	1	1.97	3.42	Disagree
Assigned Density Ratios	1	2	3	

LOGARITHMIC (common log.)

	Areas A	Areas B	Areas C	
Logs. Numerical "Counts"	1.91876 2.07918 1.71466	2.36787 2.06145 2.18639	2.71466 2.04372 2.38561	
Avg. Log	1.90420	2.20524	2.38133	
Antilog: (Geom.Mean, Area Counts)	80.2	160.47	240.62	
Ratios, Area Geom.Means	1	2.00	3.00	Agree
Assigned Density Ratios	1	2	3	

	Observers I	Observers II	Observers III	
Avg."Count"	105.34	159.40	278.21	
Ratios, Observer "Count" Averages	1	1.51	2.64	Disagree
Assigned Hearing Ratios	1	1.56	2.25	

	Observers I	Observers II	Observers III	
Avg. Log.	1.98159	2.17541	2.33376	
Geom.Mean, Observers	95.85	149.77	215.66	
Ratios, Observer Geom.Means	1	1.56	2.25	Agree
Assigned Hearing Ratios	1	1.56	2.25	

	Mornings 1	Mornings 2	Mornings 3	
Avg."Count"	102.91	176.04	264.00	
Ratios, Morning "Count" Averages	1	1.71	2.57	Disagree
Assigned Morning Ratios	1	1.41	2.08	

	Mornings 1	Mornings 2	Mornings 3	
Avg. Log.	2.00798	2.15605	2.32674	
Geom.Mean, Mornings	101.86	143.34	212.20	
Ratios, Morning Geom.Means	1	1.41	2.08	Agree
Assigned Morning Ratios	1	1.41	2.08	

geometric means of these "count" values (i.e., the antilogarithms of their average logarithms) display the desired proportionality. This agreement is essentially the basis for the requirement of additivity. (See Kempthorne, 1952.)

It must be remembered that the logarithmic analysis of variance tests only for differences between the logarithmic values of counts. Kempthorne (1952) states: ". . . if a transformation is used, the best estimates of treatment means on the untransformed scale are obtained by transforming back the means of the transformed variate. Similarly, confidence intervals must be obtained on the transformed variate."

AREA COUNTS FROM A SERIES OF STATIONS

As previously mentioned, audio-counts are used most frequently in areas large enough to require sampling from a series of stations. Usually an observer makes one count from each station during a morning by visiting stations in some convenient order. He may count a series one or more mornings, and his average count per station-visit is normally used as an index of the area's population density. This type of procedure introduces several sources of variation that can affect index accuracy, and these should be considered in designing experiments. It is necessary, before considering the use of several stations per area in a Latin Square design, to discuss in some detail the complexities that are likely to be met in counting from a series of stations.

First, population densities (D) are expected to vary somewhat from one counting station to another. Usually the average individual calling frequency (F) for most species will not be constant throughout a morning. Instead, there tends to be a rather sharp rise in activity about an hour before sunrise, followed by maximum activity and then a gradual decline as the morning progresses. (There is often a second but less spectacular increase in activity in the evening, not considered here.) There is further the possibility that the complex sound transmission-reception factor (E) will vary throughout a morning if, for instance, wind velocity changes (usually an increase), cover and terrain vary among counting stations, etc. The fact that the intensity of sound production is seasonally-dependent should not affect the following argument, except that it would seem advantageous to make counts when seasonal sound intensity is near its maximum.

Suppose, for example, that on a given morning an observer makes single counts of equal duration from "n" stations within area "a", all stations being separated sufficiently so that his sampling areas do not overlap. His total morning count, of course, will be the sum of the "n" or individual station counts which, by the audio-count model, will equal the sum of the following products:

$$\sum_{i=1}^{n} N_{a,i} = \pi K D_{a,1} F_1 R^2 E_1^{\,2} + \pi K D_{a,2} F_2 R^2 E_2^{\,2} + \cdots + \pi K D_{a,n} F_n R^2 E_n^{\,2}$$

$$= \pi K R^2 \, (D_{a,1} F_1 E_1^{\,2} + D_{a,2} F_2 E_2^{\,2} + \cdots + D_{a,n} F_n E_n^{\,2}).$$

Equivalently,

$$\frac{1}{\pi K R^2} \sum_{i=1}^{n} N_{a,i} = D_{a,1} F_1 E_1^{\,2} + D_{a,2} F_2 E_2^{\,2} + \cdots + D_{a,n} F_n E_n^{\,2}.$$

Let a second observer with identical hearing ability make a matching set of counts during the same morning from "n" counting stations in a nearby area "b". Suppose that by synchronizing counts between the two areas, average sounding frequencies (F) and sound transmission factors (E) are equal in both areas for any pair of synchronized counts, and that all counts are paired. A parallel equation for area "b" can then be derived as follows:

$$\frac{1}{\pi K R^2} \sum_{i=1}^{n} N_{b,i} = D_{b,1} F_1 E_1^{\,2} + D_{b,2} F_2 E_2^{\,2} + \cdots + D_{b,n} F_n E_n^{\,2}.$$

Suppose now that the total populations of areas "a" and "b" are equal, but distributed somewhat differently throughout each area. Suppose also that in both areas the average density sampled from the counting stations is exactly that of the true area density. Then:

$$\text{density "a"} = \text{density "b"} = \frac{1}{n}\sum_{i=1}^{n} D_{a,i} = \frac{1}{n}\sum_{i=1}^{n} D_{b,i}.$$

It does not necessarily follow that the average counts of the two areas will be equal! The averages, instead, provided that the value of the FE^2 factor changes as counts progress, will be affected by the order in which stations are counted. Further, the greater the variation of the FE^2 factor and the area sample (i.e., station) densities, the greater will be the possible difference between the average counts of the two areas.

The possible inequality of two count averages is illustrated with a simple numerical example. Suppose areas "a" and "b" each have 3 counting stations counted by identical observers in the manner described above, i.e., by using synchronized counts.

Let

$$(FE^2)_{a,1} = (FE^2)_{b,1} = 4;$$
$$(FE^2)_{a,2} = (FE^2)_{b,2} = 3; \text{ and}$$
$$(FE^2)_{a,3} = (FE^2)_{b,3} = 2.$$

Let the average station density in both areas "a" and "b" = 3, but the populations be distributed so that in area "a", $D_{a,1} = 5$, $D_{a,2} = 3$, and $D_{a,3} = 1$; while in area "b", $D_{b,1} = 1$, $D_{b,2} = 3$, and $D_{b,3} = 5$. Then

$$\frac{1}{\pi KR^2}\,\bar{N}_a = \frac{1}{3\pi KR^2}\sum_{i=1}^{3} N_{a,i} = \frac{1}{3}\,(5(4) + 3(3) + 1(2)) = \frac{1}{3}\,(31) = 10\tfrac{1}{3}$$

but

$$\frac{1}{\pi KR^2}\,\bar{N}_b = \frac{1}{3\pi KR^2}\sum_{i=1}^{3} N_{b,i} = \frac{1}{3}\,(1(4) + 3(3) + 5(2)) = \frac{1}{3}\,(23) = 7\tfrac{2}{3}$$

Thus even though the averages of the station densities are equal--$1/3(5+3+1) = 1/3(1+3+5)$--and the "mechanics" of counting are identical--the average counts are not equal. This is not to say that under no circumstances will they be equal, but it would not be probable. The explanation is simply that the average of the products of two variables is not necessarily equal to the product of their individual averages.

The same argument applies if average sample densities are not equal between areas. If the observers do not hear equally well, or if individual sounding frequency or sound transmission are different between counts, either individually or in combination, the error can become compounded. Obviously, the use of comparisons of such counts to estimate ratios of area densities can be biased.

"OUT-AND-BACK" COUNTS--If a time-dependent trend should be detected during any morning in the product of individual sounding and sound transmission, FE^2, (as might be estimated from evenly-spaced synchronized counts by several observers stationed at individual locations), it could be used to adjust the individual station counts.

Often, beginning shortly after the peak of daily count intensity, there is an approximately linear time-dependent decrease in the value of FE^2. Kozicky (1952) assumes such a decrease in adjusting cock pheasant crowing counts taken along a "circular" route of stations. He first counts all stations once and then recounts the first two stations. He next estimates the decrease in activity from the decrease in the recounts, and assuming a linear rate of decrease throughout the morning, adjusts the remaining counts accordingly. Palmer (1951), summarizing data from 10 grouse drumming routes in the northern half of Michigan's lower peninsula, conjectures such a time-dependent linear decrease in drumming activity. His data represent 153 four-minute counts taken between 45 minutes before sunrise and two hours after sunrise. He has plotted the average number of drums heard per four-minute count on a time scale of half-hour intervals measured from sunrise, choosing the mid-point of each interval as the abscissae (Fig. 7).

Frequency of Ruffed Grouse Drums in Relation
to Time Before and After Sunrise (Palmer, 1951)

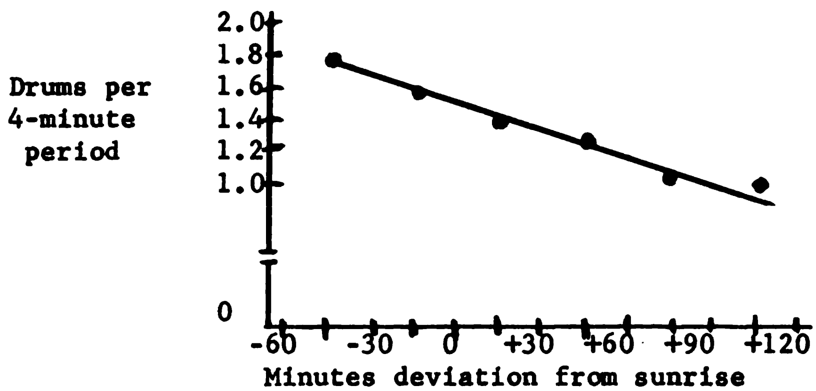

Fig. 7

When, based on previous knowledge, an approximately linear rate of decrease in a morning's count intensity can be expected, the following counting procedure should yield station counts for which, on a given morning, the values of FE^2 will be about equal:

Starting shortly after a morning's peak in count intensity (i.e., in FE^2) the observer first counts each station in some prescribed order, and again in the exact reverse of this order. All counts are of equal length and are spaced evenly in time. These might be described as "out-and-back counts." Thus each station is counted twice a morning, and the sum of each station's two values of FE^2 (i.e., FE^2 of a station's first count plus FE^2 of its second count) should be about equal for all stations.

The result of the procedure is to multiply each sample density by a nearly constant number instead of a varying quantity as when single station counts are taken. The mathematics is easily demonstrated as follows.

Let us refer again to the hypothetical areas "a" and "b" and for simplicity assume that there are only three counting stations per area, each counted twice during a morning using "out-and-back" counts. (The proof for "n" stations is essentially the same.) Again it is stipulated, for the demonstration, that the two observers have identical hearing and the different values of FE^2 are all equal between the two areas for synchronized counts. Let us assume also that between successive counts the values of FE^2 decrease arithmetically by a constant quantity "r." To simplify terms let the value of FE^2 for the first count = V. Then $F_1E_1{}^2 = V$; $F_2E_2{}^2 = V-r$; $F_3E_3{}^2 = V-2r$, etc. (In practice, of course, the value of "r" will be expected to vary slightly between counts.)

The total count for area "a," letting $N_{a,2i}$ denote the sum of the two counts of the "ith" station, can then be expressed as follows:

$$\sum_{i=1}^{3} N_{a,2i} = \pi KR^2 \; (D_{a,1}V + D_{a,2}(V-r)+D_{a,3}(V-2r)+D_{a,3}(V-3r)+D_{a,2}(V-4r)+D_{a,1}(V-5r))$$

$$= \pi KR^2 \; (2D_{a,1}V-5D_{a,1}r + 2D_{a,2}V-5D_{a,2}r + 2D_{a,3}V-5D_{a,3}r)$$

$$= \pi KR^2 \; (2V(D_{a,1} + D_{a,2} + D_{a,3}) - 5r(D_{a,1} + D_{a,2} + D_{a,3}))$$

$$= \pi KR^2 \; (2V-5r) \; (D_{a,1} + D_{a,2} + D_{a,3})$$

$$= \pi KR^2 \; (2V-5r) \; 3\bar{D}_a,$$

where $\bar{D}_a = 1/3 \; (D_{a,1} + D_{a,2} + D_{a,3})$.

With identical algebra, except to replace the $D_{a,i}$ with $D_{b,i}$, we can show that

$$\sum_{i=1}^{3} N_{b,2i} = \pi KR^2 \; (2V-5r) \; (D_{b,1} + D_{b,2} + D_{b,3})$$

$$= \pi KR^2 \; (2V-5r) \; 3\bar{D}_b.$$

Thus in both areas "a" and "b" the expected sum of the station counts is equal to the average sample (station) density $\bar{D}$ times the constant $3\pi KR^2$ (2V-5r). In the general case of "n" counting stations counted twice for a total of 2n counts, the above constant becomes

$$n\pi KR^2(2V-(2n-1)r).$$

Since the total count for each area equals the area's average sample density times a constant, the ratio of the total counts of the two areas should be a close estimate of the ratio of their true population densities. For "n" stations per area, the ratio is expressed as

$$\frac{\sum_{i=1}^{n} N_{a,2i}}{\sum_{i=1}^{n} N_{b,2i}} = \frac{\pi KR^2(2V-(2n-1)r)\ \sum_{i=1}^{n} D_{a,i}}{\pi KR^2(2V-(2n-1)r)\ \sum_{i=1}^{n} D_{b,i}} = \frac{n\bar{D}_a}{n\bar{D}_b} = \frac{\bar{D}_a}{\bar{D}_b} .$$

NUMBER OF STATIONS DIFFER--So far, only comparisons of counts between areas having the same number of counting stations have been considered. If population densities are more variable in some areas, different sampling rates may be advisable. It then becomes necessary to vary the counting procedure slightly to derive a count statistic whose expected value will be equal to an area's sample density times a morning constant common to all areas.

Let us consider an example using two areas, each with a different number of counting stations. (A comparable procedure applies for any number of areas.) Suppose, again, that counts are made with identical observers, and that a nearly linear decrease in morning count intensity (FE^2) can be expected that will be approximately equal for both (all) areas. The procedure then is to (1) take "out-and-back" counts, starting the first and the last counts simultaneously among areas, and (2) determine for each area a time interval "t" that will space its remaining counts evenly between its first and last count. (The counts at each station will still be of equal duration.) By this procedure an area's total count divided by its number of counting stations should approximate closely its average sample density times a morning constant common to both (all) areas.

To determine, by areas, the proper time interval (t) to use between counts, a suitable time interval "t_a" is first selected for area "a," which for convenience is designated as the area having the greater number of counting stations. The interval is measured from the beginning of one count to the beginning of the next. Suppose there are "n_a" stations in area "a," and "n_b" stations in area "b." Then the interval "t_b" between counts in area "b" will equal $\frac{(2\ n_a-1)}{(2\ n_b-1)}\ t_a$. Similarly, if the magnitude of the decrease in V (i.e., in FE^2) between counts in area "a" is "r_a," then the amount of decrease "r_b" between counts in area "b" will equal $\frac{(2\ n_a-1)}{(2\ n_b-1)}\ r_a$. In general, in dealing with several instead of only 2 areas, (letting "k" denote any of these areas), then $t_k = \frac{(2\ n_a-1)}{(2\ n_k-1)}\ t_a$, and $r_k = \frac{(2\ n_a-1)}{(2\ n_k-1)}\ r_a$.

The procedure can be illustrated with a simple example: Suppose there are 4 stations in area "a" and 3 stations in area "b," and the first and last counts are synchronized between areas. Thus $2\ n_a = 8$ counts will be made in area "a" and $2\ n_b = 6$ counts in area "b." Let the interval from count to count in area "a" equal "t_a." Then the total time from the start of the first count to the <u>start</u> of the last will equal $t_a(2n_a-1) = 7\ t_a$. This period, by definition, will be synchronized with and equal in length to that from the first to last count in area "b." Since 6 counts are taken in area "b," the period is also equal to $(2n_b-1) = 5$ intervals of length "t_b." Thus $7t_a = 5t_b$ or $t_b = \frac{7}{5} t_a = \frac{(2\ n_a - 1)}{(2\ n_b - 1)} t_a$.

Since t_k and r_k are assumed to be directly proportional to each other, then $\frac{r_a}{r_b} = \frac{t_a}{t_b}$, and therefore $r_b = \frac{7}{5} r_a = \frac{(2\ n_a - 1)}{(2\ n_b - 1)} r_a$.

It was shown previously that the expected total count in area "a" will equal

$$\sum_{i=1}^{4} N_{a,2i} = \pi KR^2\ (2V-7r_a)\ 4\bar{D}_a.$$

The expected total count in area "b" will equal

$$\sum_{i=1}^{3} N_{b,2i} = \pi KR^2\ (D_{b,1}V + D_{b,2}\ (V-\tfrac{7}{5}\ r_a) + D_{b,3}\ (V-\tfrac{14}{5}\ r_a) + D_{b,3}\ (V-\tfrac{21}{5}\ r_a)$$
$$+ D_{b,2}\ (V-\tfrac{28}{5}\ r_a) + D_{b,1}\ (V-\tfrac{35}{5}\ r_a))$$
$$= \pi KR^2 (2V-7r_a)\ 3\bar{D}_b.$$

The total area "a" count divided by $n_a = 4$, the number of stations in "a", is analysed as

$$1/4 \sum_{i=1}^{4} N_{a,2i} = \bar{N}_{a,2i} = \pi KR^2\ (2V - 7r_a)\ \bar{D}_a.$$

Similarly,

$$1/3 \sum_{i=1}^{3} N_{b,2i} = \bar{N}_{b,2i} = \pi KR^2\ (2V- 7r_a)\ \bar{D}_b$$

represents the total area "b" count divided by $n_b = 3$, the number of stations in "b." From the above: $\bar{N}_{a,2i} : \bar{N}_{b,2i} = \bar{D}_a : \bar{D}_b$, and the average counts of the two areas are shown to be proportional to average densities provided, as has been assumed, that the two observers have equal hearing ability.

The actual values of $2V-(2n-1)r$, or identically $2\ FE^2-(2n-1)r$, will normally vary from morning to morning, depending on the morning values of F, E, and r. For any given morning, however, this quantity should be reasonably constant among neighboring areas counted under the conditions specified.

The foregoing analysis of "out-and-back" counts leads next to the theory of their use in combination with a Latin square counting design.

THE LATIN SQUARE ANALYSIS FOR AREAS WITH SEVERAL COUNTING STATIONS

From the above discussion the following model has been developed for an "out-and-back count" index:

$$\bar{N}_{2,i} = \frac{1}{n} \sum_{i=1}^{n} N_{2i} = \bar{D}K\pi R^2 \, (2FE^2 - r(n-1)).$$

This model is identical in form to the representation of a single counting interval,

$$N = DK\pi R^2 FE^2,$$

except that mean values now replace certain of the individual terms.

The formula for "out-and-back counts" can now be utilized in the Latin square design in the same manner as the formula for "single-station" counts.

<u>SYNOPSIS OF PROCEDURE</u>: In using a Latin square counting design where each area is sampled from a series of stations, the following procedure is recommended:

1. Each morning make area counts using the "out-and-back" counting method, making sure that counting is synchronized between areas exactly as described earlier.
2. Total each area's station counts by mornings and divide each total by the number of stations in the respective area. Thus in an "n x n" Latin square there will be "n^2" count statistics, each of which represents the average combined count per station by a given observer for the particular area he counted on a given morning.

3. Take logarithms of the "n^2" averages obtained as above and carry out the analysis of variance. Where significant differences are found, the Multiple Range Test (Duncan, 1955) may aid in further evaluation. Various comparisons and confidence limits are usually transformed back (by taking antilogarithms) to the original scale of measurement (see page 23).

DISCUSSION AND CONCLUSIONS

As often made, estimates of the ratios of population densities based on comparisons of audio-indexes can be badly biased by the effects of differences in observer hearing ability and by conditions affecting sound transmission and the average frequency of individual sounding.

Logical analysis shows that audio-counts are the result of a product of several variables, and that counts by individual observers can be expected to vary as the squares of their respective hearing radii.

When the proportionality of the population densities in several areas is to be estimated by audio-indexes, a Latin square design in counting can serve to cancel out the biasing effects of the index determining variables.

In sampling an area from a series of stations, "out-and-back counts" synchronized between areas are suggested to minimize bias that may result from changes during a morning in the count intensity factor. Valid comparisons of "out-and-back" indexes assume, of course, that differences in the hearing ability of observers are accounted for. If the procedure is combined with a Latin square counting design, the effect of observer differences is negated.

The Latin square counting design should be especially useful in evaluating certain game management practices in that it can measure relative population changes between several areas before and after management. This entails keeping one or more of the areas unmanaged throughout the experiment to act as the experimental "control."

It is hoped that the theoretical analyses advanced here will not only prove usable in many instances as presented, but that they may serve as a guide in designing experiments for specific situations and a basis for expansion and refinement of audio-count techniques in general.

SUMMARY

This study defines the term "audio-index" as any population index based on systematized counts of an identifying sound of a species. It suggests audio-indexes, when applicable, to be cheaper and more efficient than comparable visual-type indexes, and thus worth detailed investigation.

By logical analysis the study (1) discusses the factors that determine the magnitude of an audio-index; (2) derives a mathematical model for an audio-count; (3) proposes a method to measure relative differences in population densities between several areas; and (4) suggests a counting technique to improve the standard audio-index.

An observer's count during a single interval is depicted as the number of potential sound producers within hearing, times the average frequency of individual sounding for this population during the interval. The model suggested for the count (excluding human error) is

$$N = DF\pi K(RE)^2,$$

where N is the numerical count, R is an observer's maximum hearing distance for a sound, E is the efficiency of sound transmission, $\pi K(RE)^2$ is the area of the observer's circle of audio-sensitivity, D is the density of potential sound producers in this area, and F is the population's average frequency of individual sounding during the count.

By this model, the ratio of audio-counts will equal the true ratio of their associated population densities only if the product $F(RE)^2$ is equal for all counts. Since N tends to deviate as the squares of R and E, and F can vary considerably, count values are not necessarily proportional to

densities, so that indiscriminate use of audio-indexes can result in appreciable error.

Audio-counts taken according to a Latin square design, letting areas represent "treatments," and observers and mornings represent "rows" and "columns" respectively, should yield unbiased estimates of the ratios of the population densities of several areas. The simplest example is used to demonstrate the basic analysis, i.e., a "3 x 3" Latin square with only one counting station per area. The use of several stations per area involves complexities considered beyond.

The method assumes that the average frequency of individual sounding is (practically speaking) unaffected by population density, and that this sounding frequency and sound transmission are about equal between areas during synchronized periods.

A standard analysis of variance assumes an additive model, while $N = DF\pi K(RE)^2$ is multiplicative. Additivity is achieved by transforming counts logarithmically, so that

(count)		(Area)		(morning)		(Observer)		(Constant)
$\text{Log } N$	$=$	$\log D$	$+$	$\log FE^2$	$+$	$\log R^2$	$+$	$\log K\pi$

The demonstration reveals that only the analysis of variance of the logarithms of counts is valid, and that the geometric means--not the arithmetic means--of the area, observer, and morning counts are proportional to area densities, observers' hearing, and morning count intensities.

To use the Latin square approach with several stations in each area, it is demonstrated that counts must yield morning indexes that for each area are a product of its average sample (station) density, the observer's area of maximum hearing, and a constant involving sounding activity and sound transmission common to all areas. Based on an assumed linear decrease in counting intensity (i.e., in FE^2), the study devises a system referred to as taking "out-and-back counts" that is expected to yield the necessary index. This index is theoretically superior to the standard index based on one count per station a morning, regardless of Latin square considerations.

The algebraic model for an "out-and-back count" is shown to be identical in form to that for a single count, and its adaptation to the Latin square design is routine. The analysis of variance now uses the logarithms of the "out-and-back" indexes instead of the single station counts; and the appropriate geometric means of these indexes estimate density, observer, and morning relationships.

The Latin square design should be especially useful in evaluating game management practices through measurements of relative population changes between several areas before and after management.

LITERATURE CITED

Carney, Samuel M. and George A. Petrides.
1957. Analysis of variation among participants in pheasant cock-crowing censuses. Jour. Wildl. Mgt., 21(4):392-397.

Cochran, W. G.
1947. Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics, 3(1): 22-38.

Dorney, Robert S., Donald R. Thompson, James B. Hale and Robert F. Wendt.
1958. An evaluation of ruffed grouse drumming counts. Jour. Wildl. Mgt., 22(1):35-40.

Duncan, D. B.
1955. Multiple range and multiple F tests. Biometrics, 11(1): 1-42.

Kempthorne, Oscar.
1952. The design and analysis of experiments. John Wiley and Sons, Inc., New York, pp. xi + 631.

Kimball, James W.
1949. The crowing count pheasant census. Jour. Wildl. Mgt., 13(1):101-120.

Kozicky, Edward L.
1952. Variations in two spring indices of male ring-necked pheasant populations. Jour. Wildl. Mgt., 16(4):429-437.

Palmer, Walter L.
1951. Ruffed grouse management investigations. Quart. Prog. Rept., Federal Aid Project W-46-R, June. Mich. Dept. Cons.

Petraborg, Walter H., Edward G. Wellein and Vernon E. Gunvalson.
1953. Roadside drumming counts, a spring census method for ruffed grouse. Jour. Wildl. Mgt., 17(3):292-295.

ROOM USE ONLY