Estimating Absolute Age Composition of California Salmon Landings

---

THE RESOURCES AGENCY OF CALIFORNIA
DEPARTMENT OF FISH AND GAME

FISH BULLETIN No. 120

Estimating Absolute Age Composition of California Salmon Landings

JOSEPH H. KUTKUHN

MARINE RESOURCES BRANCH1963

— (2) —
— (3) —

INTRODUCTION

Drawing valid inferences about parent populations is directly dependent upon the precision or reliability of estimated parameters. Self-evident as this statement is, however, situations frequently arise wherein unwarranted conclusions may be drawn from sample information whose reliability is imperfectly known. Thus it is clear that wherever possible, estimates of any sort should be supplemented with some measure of their precision. This implies, of course, that assumptions underlying any application of modern sampling theory have been reasonably satisfied.

The present paper represents an attempt to assess estimates that permit delineating the absolute age structure of fish catches, assuming the appropriate attributes are observed in a manner that at least approximates probability sampling. Such procedure further provides a basis for modifying sampling designs so future estimates derived therefrom will possess some *a priori* degree of precision imposed in accordance with available manpower and monetary resources.

Although specifically concerned with determining age composition of California landings of king salmon, Oncorhynchus tshawytscha (Walbaum), the techniques discussed herein could be applied to other species whose catches are handled in similar fashion. Moreover, they would be appropriate where dealing with any of several attributes commonly descriptive of fish populations. Besides the number of scale annuli, constituting the attribute of interest in the present paper, these include: (1) characteristics defining each kind of fish subject to capture; or, perhaps, for a given kind of fish, characteristics that would identify a member of a particular subpopulation, aggregations of which may comprise a commonly fished stock; (2) any artificially applied "mark"; (3) a particular class of lengths or weights; and, (4) sex.

Designed to provide a basis for predicting year-to-year king salmon yield potential, the parent study from which part of my data were selected evolved during the recovery phase of a large-scale salmon marking program, instituted in 1949 by the California Department of Fish and Game in cooperation with the Pacific Marine Fisheries Commission (Hallock, *et al*, 1952). In addition to sampling the California commercial salmon landings for marked fish, it was necessary to sample concurrently for other attributes (Fry and Hughes, 1953). During the program's early stages, it was further decided that still others capable of yielding potentially valuable information, *viz.*, length and age, could be sampled incidentally with little added effort and expense. Fry and Hughes (1952) briefly outlined the commercial catch sampling program set up prior to the age study's outset.

To supplement those secured from the above source, I have used additional data collected during a concomitant study of the ocean sport salmon catch (Dingell-Johnson Project California F-12-R: "Ocean Salmon Study") at various points to help develop the methods discussed.

LIMITATIONS OF THE PROCESSED DATA

Before discussing procedures involved in ascertaining the age structure of an aggregate of fish, it might be well to acknowledge the limitations of such information as determined for king salmon caught by trolling off California. Notwithstanding the fact that age-structure knowledge is important if not necessary for managing any exploited fish population, for salmon, several factors minimized the value of the information derived even though the sampling techniques may have led to very precise descriptions of catch age structure.

of these, primary consideration must be given the fact that none of the several West Coast king salmon subpopulations is definable in the ocean catch. Vital statistics such as year-class abundance and fishing rates, obtained by periodically examining a population's age structure, have little meaning when they relate to a fished stock whose component subpopulations may, with respect to one another, be subject to considerable fluctuation per unit time. It is clear then that prior to considering a salmon subpopulation's dynamics (acknowledged as necessary for implementing a sound management program for that subpopulation), all members thereof must be readily identifiable at any stage of their development.

Assuming that any member of a salmon subpopulation can be recognized in the catch and that a very good picture of that subpopulation's age structure can be obtained (by means of the procedures outlined herein), its utility will be greatly restricted unless it can be complemented with "effective" effort data. Theoretically, computed "catch-effort" or "catch-intensity" ratios for each year-class comprising a subpopulation should constitute efficient measures of the relative abundance of each. Unfortunately, these cannot be materialized because of an inability to obtain, let alone define, fishing effort derivatives representing true indices of the probability a salmon will be encountered and will succumb to a unit of trolling gear. The only way we can measure the effort expended by the California salmon trolling fleet is by troller-days—*but only by troller-days wherein a marketable yield is realized*. Such data obviously have limited usefulness.

Finally, mention should be made of the representativeness of the sample observations intended to help predict year-to-year salmon yields. Obviously, any imposition as to size (length) of fish that can be legally retained renders the total catch unrepresentative of the population or stock. Since only two age groups, 3-year-olds and 4-year-olds (unpublished data; and Appendixes I, II, IX, and XI), consistently make up the bulk of California's king salmon landings, even if all biases were removed from catch sample data, using the relative strength of a year-class to project its future contribution to the fishery over such a short time span appears questionable. Involved is not only the rate at which any king salmon year-class is fished but the differential rates at which it continually excludes itself from the fishery to reproduce. To be of any practicable value, predictions (in relative terms) of future salmon yields based upon the classical process of following year-classes should, in the least, entail full knowledge of the pre-recruitment as well as the fully recruited phases of each subpopulation. At present, such information could only be secured through systematic offshore sampling employing research vessels.

A much more positive and realistic approach to the general problem of prediction appears to lie within the framework of what might be termed the "modern" concept of fishery management. Here all measurable manifestations of a population's dynamic nature (as influenced directly by the environment and fishing) are taken into account and collectively provide a basis for predicting, in absolute terms, that quantity of the end product that can be removed per unit time on a sustained basis, leaving the brood stock more or less static.

Because of the shortcomings just pointed out, I have not attempted to integrate information derived herein and available elsewhere to analyze any aspect of king salmon population dynamics. My main concern was to describe and apply a tool designed to extract and evaluate certain information (required by the theory underlying the above-mentioned concept) which together with its complements would theoretically comprise the framework of a sound salmon management program.

DRAWING THE SAMPLE

Sampling for fishery landing attributes is quite unlike randomly selecting variegated balls from an urn. Accordingly, it was with some reservation that I applied contemporary sampling theory in an attempt to assess the estimates derived as well as the estimation procedures. Implied nonetheless is the blanket assumption that random selection prevailed at all stages of the sampling designs employed; selection and estimation biases are presumed minimal throughout.

Several features of the California oceanic salmon fisheries had to be considered in implementing a catch sampling program.

DISTRIBUTION OF LANDINGS AND SAMPLING EFFORT

King and silver salmon, O. kisutch (Walbaum), are commonly sought by sport and commercial fishermen in California's coastal waters from Avila to the Oregon border. The area of potential exploitation embraces some 12,000 square nautical miles. Generally speaking, however, spatial distribution of fishing effort thereon is nowhere near uniform, both sport and commercial fleets largely confining their activities to areas immediately adjacent to their home ports. Fortuitously, the north-central California coast has relatively few, widely scattered ports out of which operate varying proportions of the salmon fleet. Thus while monthly or seasonal landings collectively may not be representative of all portions of the coastal salmon populations (or subpopulations), gaining access to all landings to sample them is no problem.

All salmon are taken from coastal waters by trolling gear whether it is powered (commercial) or manipulated by rod and reel (sport). Individual fishing trips infrequently exceed 12 hours and, consequently, are made in the vicinity of a vessel's home port. Some commercial trollers fitted with adequate refrigeration may make extended trips. Prorating their catches with respect to origin and time of capture often is a problem.

Principal ports serving both sport and commercial salmon fishing fleets are Avila, Morro Bay, Monterey, Moss Landing, Santa Cruz,

Princeton, San Francisco, Point Reyes, Bodega Bay, Fort Bragg, Eureka, and Crescent City. To insure as complete a coverage of the total annual landings as was practicable, sampling personnel were strategically placed at from three to five of the larger ports (Fry and Hughes, 1952). Permanent (seasonal) samplers were initially (1952) located in the Monterey-Santa Cruz; San Francisco-Point Reyes-Bodega Bay; Fort Bragg; and Eureka-Crescent City areas. In later years, additional personnel frequently assisted in the Monterey vicinity, especially in the outlying Morro Bay-Avila area.

Temporal distribution of salmon fishing effort is, of course, dictated to a large extent by weather. Other factors such as nonavailability of marketable fish and prolonged price disputes often may play major roles in determining when and how much effort is expended. Notwithstanding inconveniences due to uncontrollable factors and the situations they created, time variability in a total annual landing's structure was accounted for by sampling all ports systematically throughout every season. Landings at major commercial ports such as Point Reyes and Eureka usually were sampled five days a week while those sporadically available at smaller ports typically were sampled less often.

ACCESS TO INDIVIDUAL CATCHES: CHOICE OF VESSEL

Having accounted for as much space-time variability as was practicable by distributing sampling effort, the problem then became one of sampling specific catches. Two more sources of variability had to be dealt with here, *viz.*, that between vessels and that within. Within-vessel variability, where it pertained to sampling for marked fish, length frequencies, and average weight was eliminated by examining and measuring all salmon taken by individual trollers (commercial) or charter boats and skiffs (sport). When particularly numerous, vessels themselves were selected in a random-systematic manner; *i.e.*, the first vessel was taken objectively with a constant multiple of the remainder rounding out a port-day sample. When vessels were minimal, all catches usually entered the sample.

Salmon were sampled for age, when possible, at the rate of 50 per week. At the risk of not being able to fill a "quota" by systematically sampling throughout a weekly period, these were taken at the first opportunity while sampling for other attributes. Thus, while some weekly age samples represented a series of small catches made throughout a week, many did not, often having been taken during the first one or two days from comparatively large catches of a few vessels.

Although perhaps wanting in a few respects, the foregoing scheme was believed the best that could be implemented under prevailing conditions. It may have been desirable to place more emphasis on minimizing any between-vessel error but because of manpower limitations, this frequently would have necessitated reducing the number of salmon sampled per boat, risking a reduction in weekly (and monthly) sample size. Due to uncontrollable factors which normally govern the frequency and magnitude of salmon catches, age sampling had to be carried out on a day-to-day hit-or-miss basis, thereby obviating any attempt to execute a fully systematic scheme.

CONSOLIDATING RAW DATA

A single criterion usually constitutes the basis for deciding how timeseries data should be consolidated, namely, the rate of change of those parameters being estimated. Frequently, data secured for each time increment in a sampling design are insufficient for precisely estimating an attribute's status during each increment. When those for successive increments are combined, they represent larger increments in which it is presumed the rate of change of the parameter being estimated is either very slight or the parameter itself is effectively constant.

In the present case, estimates of absolute age composition were desired for monthly king salmon landings at each port-area and for over-all annual coastwise landings. Although age samples were obtained at some ports on a daily basis and at practically all on at least a weekly basis, their sizes were often rather small and generally fluctuated widely from period to period and from port to port. This suggested they should be consolidated into larger time units and stratified geographically by general port-areas rather than by individual ports. The idea, of course, was to provide data as amenable to analysis as practicable yet arranged so underlying assumptions as to the spatial and temporal nature of the parameters of interest were not greatly disturbed.

Accordingly, all age samples and resulting data were realigned and grouped on a port-area-month basis. Each such grouping is referred to hereinafter as a port-month cluster. This procedure generally yielded data having increased flexibility computationally but may have resulted in minor losses of information concerning daily or weekly variation patterns in the age structure of landings accumulated at specific coastal points.

The general port-areas recognized are (1) Monterey, which includes all data emanating from sport and commercial landings made between Avila and Santa Cruz; (2) San Francisco, accommodating all sampling data from landings between Half Moon Bay (Princeton) and Bodega Bay; (3) Fort Bragg, with data from the Albion-Fort Bragg-Shelter Cove area; (4) Eureka-Trinidad and (5) Crescent City, to which were assigned all data from catches made in the vicinity of and landed at each. Information from sport catches made in waters adjoining either of the latter two areas was grouped under the heading Crescent City-Eureka. Typically for any given sampling year (1952–57), king salmon age data were made available for 25 commercial and 36 to 40 sport, port-month clusters. During all years covered in this study, the commercial salmon season extended for five months (May-September) and the sport season for nine.

AGE DETERMINATION

Mid-lateral salmon scales selected for aging were cleaned and mounted dry between microscope slides. During the period 1952–57, only one scale per fish was used. It was selected under low magnification from several available to lessen the likelihood of mounting regenerated material. This procedure minimized processing materials and labor and greatly facilitated both scale examination and storage since five fish could be represented on a single mount. Intensive examination

of scales secured during 1952–56 showed that a specimen's age could not always be determined satisfactorily from a single scale so an additional scale was mounted beginning with mid-season 1957.

All scales were read independently by at least two persons and disagreements were resolved by a third. This system proved quite workable, and, considering the large number of scales examined (15 to 18 thousand), disagreements occurred surprisingly seldom (less than two percent). Most of these were eventually resolved.

Assuming correctness in interpreting scale sculpture, ages were recorded on the basis of the number of effective annuli observed. A 3-year-old, for example, would refer to a salmon captured during its third year of life whose scales displayed two effective annuli.

Many scales exhibiting "ocean-type" nuclei (Gilbert, 1914) possessed initial winter marks having obviously been formed a comparatively short time after eclosion. Such scales were construed to represent progeny of early-run subpopulations which had hatched during or just prior to the period of annulus formation. To maintain correspondence with all reproduction during a given brood year, these true winter marks (ineffective annuli) had to be overlooked when assigning ages.

While such characteristics might constitute clues that would assist in separating oceanic salmon stocks into their component subpopulations, the spatial and temporal relationships involved are not clearly understood let alone defined. The same could be said for salmon whose scales indicate measurable first year stream growth. Not being in a position to evaluate or utilize these characteristics properly, we must be satisfied with our ability to determine with reasonable accuracy any specimen's age, and with age structure estimates based upon systematic but composite samples of total landings, each member being standardized according to its brood year.

DEFINITION OF THE PARENT POPULATION

Acquiring the sample and delineating its component's attributes are followed by sample estimation of those population parameters representing the attributes defined. In this case, a parent population comprised the totality of fishery landings (per unit time) from which the required samples were drawn. Whether these landings truly reflected the structure of their universe (*i.e.,* the fish population(s) available for exploiting) depended upon the distribution of fishing effort relative to that of the universe, and any biases introduced during exploitation.

Certain restrictions imposed on the California king salmon fisheries significantly curtailed the amount of information about their populations off California during 1952–1957 that we could obtain from the statistics collected. All commercial landings were composed of king salmon generally having over-all lengths of not less than 26 inches while all taken by the sport fleet had to be at least 22 inches long. Prior to 1955, sport fishermen could have a limited number of smaller king salmon.

Gear selectivity and varying availability-vulnerability of population components were sources of possible bias which, incapable of being measured, may have further delimited the parent populations (landings).

AGE STRUCTURE ESTIMATES: PORT-MONTH LANDINGS

The most widely employed means of describing catch age structure entails computing percentage contribution of age groups represented in simple random catch samples. Patently, such information by itself is of relative value and has no utility in eliciting further details concerning, for example, mortality parameters. On the other hand, when projected to the total landings and effort generating them, it can help estimate absolute numbers or mass contributed by each exploited age group, these data in turn providing estimates of each age group's (relative) abundance in the fishery, and rather good measures of the total mortality suffered by fished portions of the population under scrutiny. Admittedly, total yield and effort data are not always available and relative age structure measures must therefore suffice—though their use is generally quite limited.

Deserving consideration, however, is the fact that simple direct estimates of the proportions of each age group comprising a landing, even when supplemented with total catch and effort data, may not always constitute the most efficient estimates. One of the first to recognize this was Fridriksson (1934), who, working with cod, proposed that information concerning a closely related and easily measured attribute such as length be used to augment oft-meager age data. Since costs of extracting age measurements (counting otolith annuli) ran quite high and consequently curtailed the number of age samples that could be processed (in turn minimizing the precision of resulting estimates), he reasoned and subsequently established that for a given number of age determinations, more information generally accrued if the age composition estimates for a given landing were composites of those from small age subsamples, each selected from lots of fish having common length. Conversely, for the same level of precision that would have been attained had length stratification not been introduced, fewer age determinations would have been required. Not fully realizing the implications involved in attempting to apply a stratified estimator to commercial fishery statistics, Fridriksson merely had as his objective reducing the size of age samples needed to derive relative (percentage) age composition estimates.

With some modification (but still no assessment), Hodgson (1939) extended the procedure to secure improved estimates of absolute numbers of fish in each year-class entering East-Anglian herring landings. The method was assessed in detail by Gulland (1955) who employed it to derive mortality estimates of certain North Sea bottomfishes. Described briefly, it involves acquiring fairly large samples which are ultimately (if not already) sorted into length categories or strata, these in turn yielding smaller samples for age determination. The subsample proportions of each age group are then applied to the number of the estimated total landed that falls into the corresponding length stratum. Numbers in each age group are finally summed for all length strata to provide estimates of the total number of each age landed during a given time increment. Gulland (1955) showed (using a specific set of data) that this "age-length key" approach resulted in better estimates of absolute age composition than could have been attained using the same number of direct-sample age determinations. The frequent exceptions that have occurred cast doubt on his age-length key's general applicability and warrant investigating its limitations further. Moreover, executing certain field sampling operations involved in its construction leave much to be desired from a practical standpoint. In fact, except in situations where price differentials may demand that wholesalers sort landings in accordance with some size criterion (whereupon it is assumed that all size lots would be equally available to sampling personnel), the method has little adaptability.

A similar technique, differing significantly in the procedure of selecting the age subsample, was pursued by Tanaka (1953) when estimating absolute age structure of Japanese sardine landings. It is often referred to as the method of double or two-phase sampling for stratification. It differs from stratified subsampling in that a subsample for age is randomly taken from a larger unsorted length sample; that is, stratification occurs after subsampling for age. Since it seemed best suited to the conditions encountered, this method was used to ascertain the absolute age structure of port-month salmon landings.

TWO-PHASE SAMPLING FOR STRATIFICATION

In delineating the absolute contribution of each age group to a port-month's king salmon landings, the sample was defined as the composite of all specimens measured for length during the period and at the location of interest—subject to certain restrictions. Individual fish constituted the sample elements.

From salmon catches accumulating at a given point, the procedure was to measure fork lengths of as many specimens as practicable, assuming representative selection at all times. Lesser numbers were randomly selected from among those measured for length and scales were taken from them for effecting age determinations. Both first- and second-stage samples were ultimately sorted into 2.5-cm length strata and the second-stage samples were further separated into their component age groups.

Introducing some notation, we have then an initial sample of n fish each measured for length, from which a subsample of m fish is to be aged. Both are subdivided into i length strata, n_i and m_i representing, respectively, the numbers sampled for length and age falling into the ith length stratum. The latter are further separated on the basis of age representation into m_{ij}, the number of m_i falling into the jth age group; then

EQUATION

Treating each subsampled specimen in the ith length stratum as a binomial or Poisson variate, x_{ijk}, taking the values 1 or 0 depending upon whether its scales possess $(j - 1)$ effective annuli

EQUATION

, an unbiased estimate of the proportion of age j salmon (P_j) among the total landed is

EQUATION

where L denotes the number of length strata constructed and k identifies an individual x_{ij}.

Regarding p_i (an unbiased estimate of $P_i = N_i/N$) as a multinomial variate, $p_{j(st)}$ variance is, assuming an infinite population, given by

EQUATION

where $q_i = (1 - p_i)$ and $q_{ij} = (1 - p_{ij})$. In either form, the first term on the right side represents that portion of the total variance due to variation within length strata, the second that due to variation between strata.

EQUATION

where t is a normal deviate corresponding to an appropriate (a) confidence probability.

To apply the estimation methods described, two sets of port-month data (Appendixes I and II) were selected from our 1952–1957 data. Because of space limitations, those for all port-months sampled could not be included.

Besides presenting initial sampling data, Appendixes I and II list those obtained at intermediate stages in the process of formulating estimates of $p_{j(st)}$ for each age group represented. Appendixes III, IV, and V (Parts i), respectively, contain the results of intermediate computations leading to estimates of var ($p_{j(st)}$) for each $p_{j(st)}$ calculated from the data in Appendix II. Table 1 summarizes the results of all computations to which the selected data were subjected, giving also estimates of P_j corresponding to those obtained but derived via the simple random or direct system rather than a stratified estimator. The

simple random estimate, again assuming an infinite population, is denoted by

EQUATION

where $q_{j(ran)} = (1 - p_{j(ran)})$. $N_{j(ran)}$ and its variance are subsequently obtained in the same manner as $N_{j(st)}$ and var $(N_{(st)})$.

The stratified estimator's variance is nearly always much reduced below that of the simple random estimator, indicating a greater degree of efficiency. Comparing the coefficients of variation of the estimated number in each age group,

EQUATION

suggested, however, that the over-all improvement using the stratified estimator was frequently of such small magnitude as to question its general utility.

TABLE 1

Comparison of Age Structure Estimators and Their Variances Derived via Stratified and Simple Random Sampling Techniques. Raw Data Are in Appendixes I and II. Stratification Is by 2.5-cm Intervals

Implementation of the stratified (two-phase) design is advantageous only when gains more than compensate for the loss in precision of the principal attribute's (age) estimate. Such loss is brought about by a decrease in age sample size when part of the total sampling effort is used to measure a secondary attribute (length). In many, if not most, comparable situations, any sampling technique is restricted by the funds or manpower available. Thus to assess truly the two-phase stratified estimator one must bear in mind not only the quantity of available resources, but more important, the costs of measuring the primary attribute compared to those associated with measuring one related to it but of lesser concern. Intrinsically involved here are the size of the secondary sample where a fixed length-sample size is established, and the manner in which the first-, and hence, second-stage samples are stratified. Understanding such factors is prerequisite to determining for a fixed cost what first- to second-stage sample ratio, with varying stratification, provides a "best" estimate of the principal attribute's absolute status.

Over and above any theoretical advantages ascribed to stratified sampling, Gulland (1955) has pointed out one other with practical implications where age structure estimates are concerned. Significantly, a design incorporating stratification yields estimates of numbers in adjacent age groups which tend to be negatively correlated, a desirable condition not attained when employing the simple random sampling approach. Using simple random sampling, a positive correlation could prevail between estimates of numbers of fish falling into adjacent age groups whenever samples destined for aging are biased by length, by weight or by both. Thus if large fish were over-represented in a particular sample, the numbers in each of the older age groups would be over-estimated. Stratifying by length tends to obviate untoward effects of age sampling bias.

AGE SAMPLE SIZE FOR A FIXED LENGTH SAMPLE

Assuming comparatively large numbers of salmon can be measured for length at relatively little expense, it is now desirable to determine approximately how many should be sampled for age to provide reasonably precise estimates of P_j. The problem, for a fixed n and prescribed level of precision, is to find the minimum m for each age group, letting the largest serve as an estimate of the required overall m.

If m is randomly taken from n and the expected value of p_j is P_j, then (for any age group) the size of each m_i will be approximately proportional to p_i, that is,

EQUATION

Substituting this value in (2) and assuming an infinite population, gives, for salmon in the jth age group,

EQUATION

which, solving for m, leads to

EQUATION

the formula for age sample size with a fixed length sample where V is the desired variance of $p_{j(st)}$ and the subscript pro indicates that estimating a proportion is involved. When interest lies mainly in estimating the number in a given landing which fall into the jth age group, *i.e.*, $N_{at}P_j$, this modification can be employed:

EQUATION

where V is the desired variance of $N_{j(st)}$ and the subscript tot indicates that estimating a total is involved. The maximum m_{pro} (or m_{tot}) determined for a particular set of data then represents an approximation to m_{pro} (or m_{tot}), the over-all age sample size required for estimating all P_j's at, or greater than, some predesignated level of precision. For future sampling at corresponding port-months whose landings could be expected to exhibit the same age structure, m_{pro} constitutes a reasonable guide to the required magnitude of m.

To apply these formulas by adducing the extent to which additional age sampling was indicated for a particular situation, and, perhaps, suggest how large the projected samples should be in similar situations, data and subsequent estimates in Appendixes II to V are used. Thus to estimate the proportion of 3-year-old king salmon (±15 percent of the true value) in the Crescent City-Eureka July 1955 sport landings, we have, assuming a length sample size of 2,875.

EQUATION

If the level of precision for $p_{3(st)}$ were increased to ± 10 or 5 percent, m_{pro} would increase to 30 and 123 fish respectively. To estimate with the same precision the number of 3-year-old king salmon in the same landings, we have by means of (5) and using the same n,m_{pro}=9 specimens. The same procedure would be followed for all other age groups and the largest m_{pro} or m_{tot} calculated would determine the desired over-all age sample size.

Table 2 gives approximate values of m_{pro} for estimating P_j at two levels of precision with varying length sample sizes. It will be noted that little would be gained by merely increasing n provided each age group is reasonably represented, $p_{j(st)}$, say, being > 0.15. The size of m depends critically either upon how many age groups occur in a given set of landings or upon the extent to which each overlaps the others.

Obviously, the more age groups represented, the smaller, on the average, will be $p_{j(st)}$, and the larger the over-all age sample size needed for estimating all P_j's at a given level of precision. This points up the difficulty in predicting m's size; hence without *a priori* information, the usual alternative is to take the maximum age sample permitted by the available resource, acknowledging in advance that the status of certain age groups will be determined very precisely while others will be determined with an untenably low degree of reliability (*cf.*[Table 1]).

Though $p_{j(st)}$ can be expected to vary—often quite widely—in space and time, it is nonetheless suggested that if not too stringent precision requirements are imposed in P_j's estimation (*i.e.*, with not more than ± 20 percent allowable error), an m of scales from 200 king salmon per port-month should suffice in practically all cases. For other species exhibiting greater age ranges, m would have to be increased accordingly or the precision requirements for $p_{j(st)}$ relaxed.

TABLE 2

Age Sample Sizes With Fixed Length Samples for Estimating $p_{j(st)}$ at Two Levels of Precision

CHOICE OF STRATIFICATION

Since the age sample, if randomly taken from the length sample, can be expected to distribute itself among length strata in proportion to the latter's size, the number of strata, L, employed can be as large as desired so long as each contains at least one second-phase element. As is usually the case, relatively little is known of the variability of a population's age structure with respect to that of its length. Thus when implementing a scheme such as described, the size of L must be chosen arbitrarily at the outset. The tendency often is to over-stratify, which in turn perhaps leads to unnecessarily high costs of computing the desired estimates and measures of their precision. But while stratification with proportional allocation of sample elements always results in estimates having greater precision than those obtained by simple random sampling, the rate of precision gain with successive stratum division may decrease markedly after a certain point is reached, all gains thereafter being increasingly trifling.

An answer to the question of how many strata to employ in a two-phase sampling program defies generalization. Preliminarily, any decision as to L's size would have to be based upon some compromise made in accordance with the expected age structure of the population under scrutiny. And although an optimum or best stratification, could be determined *a posteriori* for a particular situation, it would have little value from the standpoint of general projected utility. Rather, assuming that a large length sample can be secured, a more practical approach would be to weigh the size of over-all age sample commensurate with available resources, against some predesignated minimum m_i which, collectively over i strata, could be expected to yield reasonably precise estimates of the absolute magnitude of most age groups represented.

One possible approach, assuming m_i/m has expectation N_i/N, is

EQUATION

where m is some specified minimum average m_i, and n and N are the sample and population mean stratum sizes respectively. Thus, if n is large, relative to m,

EQUATION

represents a criterion for establishing the stratification to be employed where m, the estimated over-all age sample size determined earlier, is substituted for m.

The size of m can be decided upon once the desired limits of error in estimating each P_{ij} are set. If a coefficient of variation of 0.20 is stipulated, an m_i of 25 to 30 would yield estimates of any P_{ij} being in error by more than ±20 percent of the actual value only once in 20 times. Being conservative and letting $m = 30$ specimens, we obtain where 200 specimens can be aged,

EQUATION

as a rough approximation to the number of length strata needed to obtain estimates having the desired precision. Note that m is an average and most strata will be allocated this number of aged specimens at least, but several will ordinarily be under-represented. This probably would not constitute a serious detriment in the case of California king salmon because individuals at the extremities of observed length distributions rarely fall into more than one age group. Such would not be true with species represented by a greater number of age groups or in situations where a particular age group is markedly dominant. Regardless of the sampling technique, estimates of the total numbers in the infrequently observed age groups occupying the upper extreme of any exploited population's length distribution, will almost never be as precise

as might be desired since the attendant sample size will normally be nowhere near that required to attain such precision.

To determine the effects of reducing the number of length strata (or of reducing strata widths), the length and age distributions given in Appendix II on the basis of 2.5-cm length strata were realigned into 10-cm strata (Appendix VI). Alternative estimates of P_j and its variance, using formulas (1) and (2), are given along with intermediate calculations in Appendix VII. Comparing these with those secured by two-phase sampling for 2.5-cm length strata and simple random sampling (Appendix VIII) suggests that very little would have been lost by 10-cm stratifying. This would be expected, since not only had the initial double-sampling process involving 2.5-cm strata already exhibited little superiority over single-phase sampling (Table 1), but the progressive diminution of L itself leads ultimately to the latter situation. Discounting the deviation reflected by the four-year-olds (which can be attributed to the vagaries of reallocating age data among the larger length strata), note that var N_j calculated with 10-cm strata lies between the values for var N_j computed by double sampling with 2.5-cm strata and simple random or single-phase sampling (Appendix VIII).

Two-phase Versus Simple Random Sampling with Varying Age Sample Costs

As previously indicated, a sampling design can best be appreciated only after taking into account unit costs as opposed to fixed resources and comparing results with those of some predesignated standard technique. We have considered the special case involving two-phase sampling for stratification where the first- to second-stage sampling cost ratio was assumed to be unity, *i.e.*, length and age measurements were equally costly. In the case of salmon, length stratification improved the required estimates only slightly. In fact, if the effort and expense required to make the length measurements noted (Appendixes I and II) had not served other purposes, they would have been virtually wasted.

It is desirable now to compare two-phase and simple random sampling procedures where, in the former, the second-phase sample is always more costly than the first. With either approach, we assume resources can provide an age sample of size m.

As pointed out by Cochran (1953), the values of m_i and n that yield a minimum variance in a particular situation are unavoidably intractable. In approximating a solution, it is suggested that m_i be proportional to the weighted within-stratum standard deviation, *i.e.*, to the root of the numerator of the first term on the right side of (2). This gives

EQUATION

whose substitution in (2) yields, approximately,

EQUATION

the optimum variance of p_j. This can be simplified notationally to

EQUATION

where, assuming optimum sampling, V_m is the variance within length strata and V_n is the variance between. From a practical standpoint allocating a second-phase sample under conditions such as those met with salmon would be highly unrealistic. It can be shown, however, that in two-phase sampling for stratification where the second-phase sample is randomly taken from the first, proportional and optimum allocation are theoretically the same (Deming 1950, p. 240). Hence (3) could have been used instead of (6) though to emphasize that maximum efficiency in stratified sampling is attained when the sample is allocated optimally among strata, formula (7) shall be invoked.

It is appropriate now to introduce a simple cost function of the form

EQUATION

which represents the total cost of securing the age (m) and length (n) samples where the cost of m (c_m) is appreciably greater than that of n (c_n). Under conditions of optimum sampling, the variance of p_j, (7), is now minimized by the appropriate choice of m and n for a given total cost, C. Thus differentiation of

EQUATION

(where ^ is a Lagrange multiplier) with respect to m and n leads to

EQUATION

which, after appropriate manipulation and substitution in (7) (*see* Cochran 1953, p. 272), yields

EQUATION

After several years of collecting and processing king salmon length and age data, it was determined that, on the average, the cost of eliciting a specimen's age was roughly ten times that of measuring its length. Using sampling data emanating from the July 1955 Crescent City-Eureka sport landings (Appendix II), V_m and V_n were calculated for the three age groups and are listed in Parts ii of Appendixes III, IV, and V, and in Appendix VII. Assuming that resources would permit aging no more than 130 specimens regardless of the sampling approach, their substitution in (10) [or in (8) and (9) jointly] gave predicted values for var $(p_j)_{opt}$ ([Table 3], line 3).

For comparative purposes, the same procedure was employed to secure predicted values for var $(p_j)_{opt}$ where the cost of age samples and determinations was presumed to be twice and five times that of the corresponding length sample measurements (Table 3, lines 1 and 2). All predicted var$(p_j)_{opt}$ values were compared with appropriate predicted values for var$(p_j)_{ran}$, the variance resulting in the event a double or two-phase sampling scheme is not employed (*i.e.*, n = 0).

Five salient pieces of information can be gleaned from Table 3: (1) Two-phase sampling for stratification becomes increasingly more efficient for estimating age as the cost of the age sample increases relative to the length sample. (2) For a given cost relationship, diminution of the number of length strata results in a marked increase in variance depending upon the sample age group distribution. (3) Assuming representativeness of these data, there seems to be little point in implementing a two-phase length-age sampling scheme unless the age sample is at least five times costlier than the length samples [*cf.* var$(p_j)_{opt}$ and var$(p_j)_{van}$]. (4) When determining age is much costlier than length (*e.g.*, line 3), age sample sizes required of the two-phase scheme for estimating $_{j(st)}$ (and $N_{at}p_{j(st)}$) with the same precision as attained by the simple random or direct approach, are always reduced (often greatly); conversely, at high second-phase sample costs and for a fixed second-phase sample size, greater precision can be expected to accrue from double sampling. (5) Very little added precision can be expected by increasing first-stage sample sizes (*cf.* n in Appendix II and Table 3).

ESTIMATING OVER-ALL AGE COMPOSITION

Considering the commercial catch only, over-all seasonal estimates of the absolute number of a given age landed could be obtained by merely summing the $N_{j(st)}$'s over all ports and months—provided sampling data are available for all port-month clusters. Since this is usually not the case, some alternative scheme is in order.

A possible estimator appropriate to cluster sampling wherein cluster size is variable is a ratio estimator using the stratified estimates $N_{j(st)}$ obtained for each port-month sampled. For the jth age group it is

EQUATION

is a simple unweighted estimate of $P_{j(r)}$, N_s is the independently estimated over-all total number of fish landed during a season, and k is the

---

TABLE 3

Comparison of Variances of p_j Estimated by Two-phase Sampling at Two Levels of Stratification Presuming Optimum Allocation, and by Simple Random Sampling. Raw Data Are in Appendixes II and VI. Intra-strata Variances, V_n and V_m Are Derived in Appendixes III-V; (parts ii) and Appendix VII

---

number of h port-month clusters sampled. Assuming an infinite population, its estimated variance (Cochran, 1953, p. 125) is approximately

EQUATION

where

EQUATION

is the average cluster size and the remaining notation is as defined earlier. It follows that the variance of $N_S P_{j(r)}$ is simply

EQUATION

Providing sample sizes are large enough so that the normal approximation applies, confidence limits for $N_S P_{j(r)}$ can be calculated

EQUATION

where t is a normal deviate corresponding to an appropriate (a) confidence probability.

Using the stratified (two-phase sampling) age estimates calculated for individual clusters (Appendix IX) and treating them as weighted

TABLE 4

Age Composition of California's 1952 Commercial King Salmon Landings. Basic Data Are Total Numbers Estimated by Two-phase Stratified Sampling Within Each Port-month Cluster

sample observations, $P_{j(r)}$'s for the four age groups in the 1952 commercial landings are given in the lower right hand corner of Appendix X. In Appendix X the row values opposite the ages correspond to the ($N_{j(s.)}$'s calculated as described earlier, and over-all P_j's for ports ($P_{j(a.)}$) and months ($P_{j(.t)}$), computed from the $N_{j(st)}$'s summed over ports and months, are in the right and lower margins. These are given merely to indicate on a seasonal basis, relative distribution of the fished ages in the landings by port and month. Their reliability is questionable, however, because of the small number of sampled clusters upon which they are based (*re* Appendix X, a minimum of one and a maximum of three clusters serve to compute the marginal P_j's). Results of applying the estimator of main concern, $P_{j(r)}$, are given in Table 4. Here, too, interval estimates of $NP_{j(r)}$ have questionable utility since the variance (and standard error) of the ratio of two random variables is biased for small samples.

A more useful and realistic means of estimating over-all catch age structure entails fully utilizing the sampling scheme possible with portmonth clusters. Providing reasonably adequate sampling obtained in all clusters, four possible estimators are available for each age landed. They are (for fish of any age j):

EQUATION

TABLE 5

Estimated Proportions and Total Numbers of Age; King Salmon Entering California's 1955 Commercial Landings. For Comparison, Four Estimators Were Employed. Basic Data Are in Appendix XI

Note that each estimator uses basic (untreated) cluster data and, except in the case of (11), the independently estimated catch totals. Estimated port-month totals secured by two-phase sampling were not considered.

To illustrate the role of the two-way cluster scheme in obtaining best point and interval estimates of over-all catch age structure, I used the 1955 commercial age sampling data. Appendix XI contains raw age data and estimated catch totals for each sampled cluster, as well as intermediate calculations needed with estimators (11), (12), and (13). Individual cluster $P_{j(at)}$'s used with estimator (14) are recorded in Appendix XII. Each of the four estimators compared in Table 5 gives approximately the same results. This is to be expected if there is little variability in the proportion of age j fish between ports and months. Estimator (14) should be the least biased since it takes into account the effects (if any) of both ports and months. It is with (14) that we shall be primarily concerned in the following discussion.

Although $N_sP_{j(at)}$ would be expected to yield the best point estimate of the total number of age j fish landed seasonally, a measure of its precision may not be too useful since the available variance estimator for a ratio P is subject to bias in small samples. To secure an appropriate interval estimate (and hence measure of precision) of $N_sP_{j(at)}$ might call for constructing such an interval using the least biased variance estimator, *viz.*, the conventional (but weighted) binomial estimator, and then adjusting it for cluster effects by employing an appropriate correction factor. Such an estimator would be appropriate in the unrealistic case where

EQUATION

is regarded as a simple random sample of m individual fish.

The correction factor used is simply the ratio

EQUATION

where the numerator and denominator are standardized with respect to a common level of P, a ratio representing a measure of the relative change in standard error caused by using the two-way cluster sampling scheme.

The basic variance of $P_{j(a-t)}$>, var ($P_{j(a-t)}$), is, assuming an infinite population and regarding the individual $P_{j(at)}$'s as continuous variates, the mean square for port-month interaction derivable by ordinary analysis of variance techniques; *i.e.*,

EQUATION

For any age group *j*, var $P_{j(a-t)}$> is due to the failure of differences between successive port and month $P_{j(at)}$> to follow a definite pattern from port to port and month to month (Appendix XII).

Since cluster size (m_{at}) varied considerably from one to the next, the interaction sum of squares could not be obtained in the conventional manner. It was necessary therefore to compute a weighted interaction sum of squares by first obtaining the expected P (or regression) value for each port-month cell. Assuming they were independent, the deviations of the expected (or regression) values from their respective observed values (Appendix XII) were used to calculate the approximate sum of squares, and thence an approximation to the interaction mean square, employing the relationship

EQUATION

where k is the number of clusters sampled and

EQUATION

When all clusters are sampled, $k = uv$. Original and intermediate data used in computing this value for each age group encountered in the 1955 landings are in Appendixes XI, XII, and XIII. In the foregoing procedure, data in the two-way table must be complete or nearly so. Values for the data missing from the Monterey-September cluster (*see* Appendixes XI and XII) were obtained using the method described by Anderson and Bancroft (1952, p. 246) where the estimated variates are determined such that the resulting interaction sum of squares will be minimized. This can be seen in Appendix XII (Monterey-September) where, for all practical purposes, the substituted P_{at}'s are their own expected values.

Regarding var ($P_{(a-t)}$) as standardized with respect to the over-all mean P for any corresponding age group *j*, the binomial variance for the same level of P, given by

EQUATION

was then estimated for each age group. Its square root served as the denominator in the ratio 'G' (Appendix VIII).

An approximation to the binomial variance of $P_{(a-t)}$ was made using the weighted estimator

EQUATION

is the weighted mean cluster size. It follows that an approximation to the least biased variance estimator of the total in any age group *j*, *i.e.*,$N_s P_{j(a-t)}$, is $N_S$var($P_{j(a-t)}$) with basic interval estimate for $N_s P_{j(a-t)}$ given by

EQUATION

The desired interval estimate which takes into account the effects of clusters is then simply

EQUATION

Point and corrected interval estimates of the numbers of king salmon in each of four age groups entering the 1955 commercial catch are given in Table 6.

Although somewhat complex computationally, I feel this system provides a convenient means for obtaining best estimates (in absolute terms) of the characters in question, besides permitting some generalizing as to the average size of cluster to employ in future censuses. If it is determined that the required minimum cluster size would be too unwieldy, it might be well to examine critically the between ports variability as regards the proportions of each age group entering the catch (Appendix XI) and, in situations where two-phase (stratified) sampling is used to estimate individual cluster age composition, the age apportionment within corresponding length classes. If over a period of years this variability is generally small (as in the case of certain ages represented in the 1952 and 1955 landings), a reallocation of age samppling effort would be in order. By eliminating some ports and increasing age sample sizes (*i.e.*, cluster sizes) in the remainder, more precise overall estimates of the absolute contributions of each age group to the fishery could be obtained with little or no increase in overall sampling effort.

By making preliminary estimates of mean cluster size using the binomial two-way estimator for $P_{j(a-t)}$ and then inflating it by squaring the 'G' factor to determine the mean size of cluster for a given level of precision (Cochran 1953, p. 204), I found that the cluster sizes for 1955 were generally more than adequate for determining precisely (10 percent coefficient of variation) the total number of 3- and 4-year-old king salmon in the catch. This was not the case for 2- and 5-year-old fish, however. In the former, a mean cluster size of about 3,290 fish would have been needed to provide an estimate of the total number in that age group having a correspondingly high precision. For 5-year-old salmon at about the same level of precision, a mean cluster size of approximately 260 fish would have been required. The actual mean cluster size m was 164.

TABLE 6

Point and Interval Estimates of Total 2-, 3-, 4-, and 5-year-old King Salmon Entering the 1955 California Commercial Landings. (Point Estimates Were Derived Using the Two-way (Port-month) Estimator. Corresponding Interval Estimates Were Corrected for Cluster Effects by an Appropriate 'G' Factor.)

SUMMARY

Because conditions under which king salmon catches must be sampled vary considerably in space and time, it is impossible to generalize as to the width of length strata (i) and sizes of length (n) and age (m) samples which, properly integrated, could be expected to yield satisfactory age composition estimates on a port-month basis. For a fixed stratum width, to improve two-phase sampling estimates requires larger age samples. Only slight improvement can be had by taking larger length samples (Tanaka 1953; Gulland 1955, 1957). Age sample components should adequately represent all length classes involved. Since proper allocation may not be possible, age-within-length samples (m_i) could be fixed at some convenient level for all length strata, admitting wasted effort in those strata exhibiting minimal variability. Under certain conditions (assuming sufficiently large length samples can be obtained), a reasonable compromise might be to sample differentially on the basis of experience; that is, secure an m of any fixed size by judiciously taking for age determination fewer fish from length classes which would be expected to exhibit slight variation (p_{ij} relatively large) than from classes in which greater variability would be expected (p_{ij} relatively small).

In future programs utilizing the two-phase estimator, port-month estimates of the total king salmon catch (by number) in any age group should be based upon age samples consisting of scales from at least 200 specimens. From the standpoint of the labor involved in processing and reading salmon scales, experience has shown this is a reasonable quantity. Also, length strata should have a minimum width of 5 cm and those containing king salmon 80 cm long and longer should be sampled heavier than those with smaller fish. If 5-cm length classes are employed, an average of eight effective strata would likely be involved in commercial salmon landings, four of which would probably include fish 80 cm long and longer. Age samples from large-length strata, twice the size (if possible) of those from small-length strata, should provide, with minimum wasted effort, reasonably precise estimated total numbers of fish landed in each age group.

Both indirect- and direct-estimation age samples are dependent upon the number and distribution of ages in the catch. For a fixed m, both approaches appear to give comparable results where there are few ages, as with salmon. Indirect estimation does, however, attain increasing precedence over direct as the number of ages increases; such was the case with Japanese sardines (Tanaka 1953) and North Sea plaice (Gulland 1955). In light of the relatively slight gains exhibited for salmon by the two-phase technique over the simple random method, unless age samples are more than five times as costly as length samples, the direct method which is less expensive and less complicated would be preferred over the indirect in future sampling programs. Except

---

where conditions definitely warrant its use, the two-phase method is at least as costly as the simple direct method, is more complex computationally and is more difficult to execute insofar as actual sampling is concerned.

of even more concern than deciding which of several estimators is the most efficient for determining the absolute status of any attribute within individual groups (or clusters) of fish, is the problem of accounting for space-time variability between groups in attempts to secure best estimates for a management unit as a whole. I made such an attempt for salmon employing a two-way cluster sampling scheme that accounted for both port and month effects. My basic data were from direct estimation of individual cluster parameters (P_{at}); data obtained this way are more amenable to the techniques involved than those complicated by using length as an adjunct. Correcting for cluster effects, I determined that an average cluster size of about 250 age determinations could, in future programs, be expected to yield fairly precise (CV = 10 percent) over-all estimates of the number of king salmon in each age group commonly represented.

ACKNOWLEDGMENTS

I am grateful for the generous assistance of Raymond J. Jessen, C-E-I-R, Inc., whose critical review of the manuscript, subsequent comments, and many helpful suggestions added measurably to it. Errors of any sort, however, remain my full responsibility.

Acknowledgment is also given all personnel who helped carry out the California Department of Fish and Game's extensive catch-sampling program described herein.

REFERENCES

Anderson, R. L., and T. A. Bancroft1952. Statistical theory in research. New York, McGraw-Hill Book Co., 399 pp.Cochran, W. G.1953. Sampling techniques. New York, J. Wiley & Sons, 330 pp.Deming, W. E.1950. Some theory of sampling. New York, J. Wiley & Sons, 602 pp.Fridriksson, A.1934. On the calculation of age-distribution within a stock of cod by means of relatively few age-determinations as a key to measurements on a large scale. Rapp. Cons. Explor. Mer, vol. 86, no. 6, pp. 1–14.Fry, D. H., Jr., and E. P. Hughes1952. A sampling program for recovery of marked king and silver salmon. Calif. Fish and Game, vol. 38, no. 4, pp. 535–540.1953. Proportion of king and silver salmon in California's 1952 landings. Calif. Dept. Fish and Game, Fish. Bull. No. 95, pp. 7–16.Gilbert, C. H.1914. Age at maturity of the Pacific coast salmon of the genus *Oncorhynchus*. Bull. U. S. Bur. Fish., vol. 32, pp. 1–22.Gulland, J. A.1955. Estimation of growth and mortality in commercial fish populations. Min. Agric. and Fish., Fish. Invest., ser. 2, vol. 18, no. 9, 46 pp.1957. Sampling problems and methods in fisheries research. FAO Fish Bull., vol. 10, pp. 157–181.

Hallock, R. J., G. H. Warner, and D. H. Fry, Jr.1952. California's part in a three-state salmon fingerling marking program. Calif. Fish and Game, vol. 38, no. 3, pp. 301–332.
Hodgson, W. C.1939. An improved method of sampling herring shoals. Rapp. Cons. Explor. Mer, vol. 110, pp. 31–80.
Tanaka, S.1953. Precision of age-composition of fish estimated by double sampling method using the length for stratification. Bull. Japan. Soc. Sci. Fish., vol. 19, no. 5, pp. 657–670.

APPENDIXES

APPENDIX I

Length-age Data—Two-phase Stratified Sampling From California King Salmon Sport Landings at San Francisco, February 1955. Estimated Total Number Landed (N_{at}) : 8,500. Stratification Is by 2.5-cm Intervals

APPENDIX II

Length-age Data—Two-phase Stratified Sampling From California King Salmon Sport Landings at Crescent City-Eureka, July 1955. Estimated Total Number Landed (N_{at}) : 10,660. Stratification Is by 2.5-cm Intervals

APPENDIX III

Variance Calculations Employing Two-phase Stratified Sampling for Estimating Proportion of Two-year-old King Salmon From Data in Appendix II

APPENDIX IV

Variance Calculations Employing Two-phase Stratified Sampling for Estimating Proportion of Three-year-old King Salmon From Data in Appendix II

Variance Calculations Employing Two-phase Stratified Sampling for Estimating Proportion of Four-year-old King Salmon From Data in Appendix II

APPENDIX VI

Length-age Data—Two-phase Stratified Sampling From California King Salmon Sport Landings Crescent City-Eureka, July 1955. Estimated Number Landed (N_{at}) : 10,660. Stratification Is by Six, 10-cm Intervals

APPENDIX VII

Variance Calculations Employing Two-phase Stratified Sampling for Estimating Proportions of Two-, Three-, and Four-year-old King Salmon From Data in Appendix VI. (Formulas and Explanation in Text.)

APPENDIX VIII

Comparison of Age Structure Estimates and Their Variances Employing Two-phase Sampling for Two Stratification Levels. Raw Data Are in Appendixes II and VI

APPENDIX IX

Estimated Numbers by Age ($N_{j(st)}$) of King Salmon Landed Commercially at California Ports During 1952. Figures in Parentheses Are Variation Coefficients Expressed as Percentages. Empty Cells Indicate No Sampling

APPENDIX X

Estimated Total Numbers ($N_{j(st)}$) and Derived Proportions (P_j) of Age *j* King Salmon. Basic Data From 1952 California Commercial Landing Samples. Empty Cells Indicate No Sampling

APPENDIX XI

Age Sampling Data for 1955 California Commercial King Salmon Landings. Data in Right and Lower Margins Used to Derive Estimates in Table 5

APPENDIX XII

Observed, P_{at}, and Deviation From Regression, $P_{at} — E\ (P)$,* Values for Calculating Port-month Interaction Sum of Squares and Mean Square Error for Two-way Cluster Sampling. P_{at}'s are Cluster Estimates of Proportions of King Salmon in Each of Four Age Groups (*j*) Landed by the 1955 California Commercial Fishery. Sample Data Are in Appendix XI

APPENDIX XIII

Statistics for Calculating Port-month Interaction Mean Square and Correction Factor 'G*j*'. Basic Age Data Are From 1955 California Commercial King Salmon Landing Samples. (*See* Appendixes XI, XII, and Text.)

CALIFORNIA DEPARTMENT OF FISH AND GAME

FISH BULLETINS

[*] No. 1. Report on Fish Conditions. 1913; 48 pp., 3 figs.

[*] No. 2. The Scientific Investigation of Marine Fisheries, as Related to the Work of the Fish and Game Commission in Southern California. By Will F. Thompson. 1919; 27 pp., 4 figs.

[*] No. 3. The Spawning of the Grunion (Leuresthes tenuis). By Will F. Thompson, assisted by Julia Bell Thompson. 1919; 29 pp., 9 figs.

[*] No. 4. The Edible Clams, Mussels and Scallops of California. By Frank W. Weymouth. 1920; 72 pp., 19 pls., 26 figs.

[*] No. 5. A Key to the Families of Marine Fishes of the West Coast. By Edwin C. Starks. 1921; 16 pp., 4 figs.

[*] No. 6. A History of California Shore Whaling. By Edwin C. Starks. 1923; 38 pp., 22 figs.

[*] No. 7. The Life-History and Growth of the Pismo Clam. By Frank W. Weymouth. 1923; 120 pp., 15 figs., 18 graphs.

[*] No. 8. Racial and Seasonal Variation in the Pacific Herring, California Sardine and California Anchovy. By Carl L. Hubbs. 1925; 23 pp., 4 pls.

[*] No. 9. Preliminary Investigation of the Purse Seine Industry of Southern California. By Tage Skogsberg. 1925; 95 pp., 23 figs.

[*] No. 10. The Life History of Leuresthes tenuis, an Atherine Fish with Tide Controlled Spawning Habits. By Frances N. Clark. 1925; 51 pp., 6 graphs, 7 pls.

[*] No. 11. The California Sardine. By the Staff of the California State Fisheries Laboratory. 1926; 221 pp., 74 figs.

[*] No. 12. The Weight-Length Relationship of the California Sardine (Sardina caerulea) at San Pedro. By Frances N. Clark. 1928; 58 pp., 11 figs.

[*] No. 13. The Seasonal Average Length Trends at Monterey of the California Sardine (Sardina caerulea). By Carroll B. Andrews. 1928; 12 pp., 6 figs.

[*] No. 14. Report on the Seals and Sea Lions of California. By Paul Bonnot. 1928; 61 pp., 38 figs.

[*] No. 15. The Commercial Fish Catch of California for the Years 1926 and 1927. By the Bureau of Commercial Fisheries. 1929; 93 pp., 52 figs.

[*] No. 16. The Life History of the California Jack Smelt (Atherinopsis californiensis). By Frances N. Clark. 1929; 22 pp., 12 figs.

[*] No. 17. Sacramento-San Joaquin Salmon (Oncorhynchus tschawytscha) Fishery of California. By G. H. Clark. 1929; 73 pp., 32 figs.

[*] No. 18. The Pismo Clam: Further Studies of Its Life History and Depletion. By William C. Herrington. 1930; 67 pp., 16 figs.

[*] No. 19. Sardine Fishing Methods at Monterey, California. By W. L. Scofield. 1929; 61 pp., 27 figs.

[*] No. 20. The Commercial Fish Catch of California for the Year 1928. By the Staff of the Bureau of Commercial Fisheries. 1930; 109 pp., 62 figs.

[*] No. 21. Analysis of Boat Catches of White Sea Bass (Cynoscion nobilis) at San Pedro, California. By S. S. Whitehead. 1930; 26 pp., 20 figs.

[*] No. 22. A Bibliography of the Tunas. By Genevieve Corwin. 1930; 103 pp.

[*] No. 23. Success of the Purse Seine Boat in the Sardine Fishery at Monterey, California (1929–1930 Fishing Season). By J. B. Phillips. 1930; 28 pp., 19 figs.

[*] No. 24. An Analysis of the Catch Statistics of the Striped Bass (Rocous lineatus) Fishery of California. By J. A. Craig. 1930; 41 pp., 22 figs.

[*] No. 25. Fishing Areas Along the California Coast for the Sardine (Sardina caerulea). By

the California State Fisheries Laboratory. 1930; 44 pp., 25 figs.

[*] No. 26. Seasonal Changes in the Daily Average Length of the California Sardine (Sardina caerulea). By Frances N. Clark. 1930; 20 pp., 11 figs.

[*] No. 27. The Ring Net, Half Ring Net, or Purse Lampara in the Fisheries of California. By Donald H. Fry, Jr. 1931; 65 pp., 28 figs.

---

[*] No. 28. Handbook of Common Commercial and Game Fishes of California. By Lionel A. Walford. 1931; 181 pp., 137 figs.

[*] No. 29. The Striped Bass of California (Roccus lineatus). By Eugene C. Scofield. 1931; 82 pp., 47 figs.

[*] No. 30. The Commercial Fish Catch of California for the Year 1929. By the Staff of the Bureau of Commercial Fisheries. 1931; 133 pp., 75 figs.

[*] No. 31. Studies of the Length Frequencies of the California Sardine (Sardina caerulea). By the California State Fisheries Laboratory. 1931; 53 pp., 24 figs.

[*] No. 32. The California Halibut (Paralichthys californicus) and an Analysis of the Boat Catches. By G. H. Clark. 1931; 52 pp., 25 figs.

[*] No. 33. Fishing Methods for the Bluefin Tuna (Thunnus thynnus) and an Analysis of the Catches. By S. S. Whitehead. 1931; 32 pp., 22 figs.

[*] No. 34. Salmon of the Klamath River, California. By John O. Snyder. 1931; 130 pp., 44 figs.

[*] No. 35. A Distributional List of the Species of Freshwater Fishes Known to Occur in California. By Barton W. Evermann and Howard Walton Clark. 1931; 67 pp.

[*] No. 36. A Bibliography of the Sardines. By Genevieve C. Wheeler. 1931; 133 pp.

[*]. No. 37. The California Barracuda (Sphyraena argentea). By Lionel A. Walford. 1932; 120 pp., 32 figs., 6 pls.

[*] No. 38. The California Shrimp Industry. By Paul Bonnot. 1932; 20 pp., 11 figs.

[*] No. 39. Fluctuations in the Abundance of Striped Bass (Roccus lineatus) in California. By G. H. Clark. 1933; 18 pp., 7 figs.

[*] No. 40. The California Mackerel Fishery. By Richard S. Croker. 1933; 149 pp., 73 figs.

[*] No. 41. Early Life History of the California Sardine (Sardina caerulea), with Special Reference to Distribution of Eggs and Larvae. By Eugene C. Scofield. 1934; 48 pp., 24 figs.

No. 42. Maturity of the California Sardine (Sardina caerulea), determined by Ova Diameter Measurements. By Frances N. Clark. 1934; 49 pp., 19 figs.

No. 43. The Sizes of California Sardines Caught by the Different Fishing Gear and in the Different Localities of the Monterey and San Pedro Regions. By the California State Fisheries Laboratory. 1935; 59 pp., 27 figs.

[*] No. 44. The Commercial Fish Catch of California for the Years 1930–1934, Inclusive. By the Staff of the Bureau of Commercial Fisheries. 1935; 124 pp., 19 figs.

[*] No. 45. The Sharks and Rays of California. By Lionel A. Walford. 1935; 66 pp., 58 figs.

[*] No. 46. A Contribution Toward the Life Histories of Two California Shrimps, Crago franciscorum (Stimpson) and Crago nigricauda (Stimpson). By Hugh R. Israel. 1936; 28 pp., 9 figs.

No. 47. Interseasonal and Intraseasonal Changes in Size of the California Sardine (Sardinops caerulea). By Frances N. Clark. 1936; 28 pp., 11 figs.

No. 48. Fishing Localities for the California Sardine, Sardinops caerulea, 1928–1936. By Frances N. Clark. 1937; 11 pp., 5 figs.

[*] No. 49. The Commercial Fish Catch of California for the Year 1935. By the Bureau of Commercial Fisheries. 1937; 170 pp., 114 figs.

No. 50. Sizes of California Sardines Caught in the Different Areas of the Monterey and San Pedro Regions. By J. B. Phillips. 1937; 31 pp., 12 figs.

[*] No. 51. The High Seas Tuna Fishery of California. By H. C. Godsil. 1938; 41 pp., 20 figs.

[*] No. 52. Historical Account of the Los Angeles Mackerel Fishery. By Richard S. Croker. 1938; 62 pp., 37 figs.

No. 53. Measures of Abundance of the Sardine, Sardinops caerulea, in California Waters. By Frances N. Clark. 1939; 45 pp., 19 figs.

[*] No. 54. The Fishes of the Family Sciaenidae (Croakers) of California. By Tage Skogsberg. 1939; 62 pp., 16 figs.

No. 55. Report on Returns of Drift Bottles Released off Southern California, 1937. By Richard B. Tibby. 1939; 36 pp., 22 figs.

[*] No. 56. Development of the Eggs and Early Larvae of Six California Fishes. By Paul L. Budd. 1940; 50 pp., 13 pls.

No. 57. The Commercial Fish Catch of California for the Years 1936–1939, Inclusive. By the Staff of the Bureau of Marine Fisheries. 1940; 100 pp., 9 figs.

No. 58. The Commercial Fish Catch of California for the Year 1940. By the Staff of the Bureau of Marine Fisheries. 1942; 47 pp., 7 figs.

No. 59. The Commercial Fish Catch of California for the Years 1941 and 1942. By the Staff of the Bureau of Marine Fisheries. 1944; 68 pp., 8 figs.

[*] No. 60. A Systematic Study of the Pacific Tunas. By H. C. Godsil and Robert D. Byers. 1944; 131 pp., 76 figs.

[*] No. 61. Results of Tagging Experiments in California Waters on the Sardine (Sardinops caerulea). 1945; 90 pp., 15 figs.

[*] No. 62. Catch per-Unit-of-Effort in California Waters of the Sardine (Sardinops caerulea) 1932–42. By Ralph P. Silliman and Frances N. Clark. 1945; 76 pp., 22 figs.

No. 63. The Commercial Fish Catch of California for the Years 1943 and 1944. By the Staff of the Bureau of Marine Fisheries. 1946; 81 pp., 6 figs.

[*] No. 64. The Biology of the Soupfin, Galeorhinus zyopterus, and Biochemical Studies of the Liver. 1946; 93 pp., 41 figs.

No. 65. Analysis of Populations of the Pacific Sardine on the Basis of Vertebral Counts. By Frances N. Clark. 1947; 26 pp., 3 figs.

No. 66. Drift and Set Line Fishing Gear in California. By W. L. Scofield. 1947; 38 pp., 16 figs.

No. 67. The Commercial Fish Catch of California for the Years 1945 and 1946. By the Staff of the Bureau of Marine Fisheries. 1947; 80 pp., 7 figs.

[*] No. 68. Common Marine Fishes of California. By Phil M. Roedel. 1948; 150 pp., 111 figs.

[*] No. 69. Age and Length Composition of the Sardine Catch off the Pacific Coast of the United States and Canada, 1941–42 through 1946–47. By Frances E. Felin and Julius B. Phillips. 1948; 122 pp.

[*] No. 70. A Preliminary Population Study of the Yellowfin Tuna and the Albacore. By H. C. Godsil. 1948; 90 pp., 22 figs.

[*] No. 71. Growth of the Sardine, Sardinops caerulea, 1941–42 through 1946–47. By Julius B. Phillips. 1948; 33 pp., 12 figs.

No. 72. Trawling Gear in California. By W. L. Scofield. 1948; 60 pp., 24 figs.

No. 73. Tagging Experiments on the Pacific Mackerel (Pneumatophorus diego). By Donald H. Fry, Jr. and Phil M. Roedel. 1949; 64 pp., 15 figs.

No. 74. The Commercial Fish Catch of California for the Year 1947 With an Historical Review, 1916–1947. By the Staff of the Bureau of Marine Fisheries. 1949; 267 pp., 99 figs.

[*] No. 75. California Sharks and Rays. By Phil M. Roedel and Wm. Ellis Ripley. 1950; 88 pp., 65 figs.

No. 76. Average Lunar Month Catch by California Sardine Fishermen, 1932–33 through 1948–49. By Frances N. Clark and Anita E. Daugherty. 1950; 28 pp., 6 figs.

No. 77. A Comparison of the Bluefin Tunas, Genus Thunnus, from New England, Australia and California. By H. C. Godsil and Edwin K. Holmberg. 1950; 55 pp., 15 figs.

No. 78. The Life History of the Starry Flounder, Platichthys stellatus (Pallas). By Harold George Orcutt. 1950; 64 pp., 52 figs.

No. 79. A Key to Some Southern California Fishes Based on Vertebral Characters. By Charles R. Clothier. 1950; 82 pp., 22 figs., 23 pls.

No. 80. The Commercial Fish Catch of California for the Years 1948–1949 with Yield per Area of the California Sardine Fishing Grounds 1937–1949. By the Staff of the Bureau of Marine Fisheries. 1951; 87 pp., 9 figs.

No. 81. Purse Seines and Other Roundhaul Nets in California. By W. L. Scofield. 1951; 82 pp., 37 figs.

No. 82. A Comparison of the Populations of Yellowfin Tuna, Neothunnus macropterus, from the Eastern and Central Pacific. By H. C. Godsil and E. C. Greenwood. 1951; 33 pp., 6 figs.

No. 83. Age Composition of the Southern California Catch of Pacific Mackerel 1939–40 through 1950–51. By John E. Fitch. 1951; 73 pp., 17 figs.

No. 84. A Racial Study of the Pacific Mackerel, Pneumatophorus diego. By Phil M. Roedel. 1952; 53 pp., 5 figs.

No. 85. The Biology of the Dover Sole, Microstomus pacificus (Lockington). By Frederick B. Hagerman. 1952; 48 pp., 28 figs.

[*] No. 86. The Commercial Fish Catch of California for the Year 1950 with a Description of Methods Used in Collecting and Compiling the Statistics. By the Staff of the Bureau of Marine Fisheries. 1952; 120 pp., 43 figs.

No. 87. Surveys through 1951 of the Distribution and Abundance of Young Sardines (Sardinops caerulea). By Julius B. Phillips and John Radovich. 1952. 63 pp., 13 figs.

No. 88. A revision of the Family Embiotocidae (The Surfperches). By Fred H. Tarp. 1952; 99 pp., 32 figs.

No. 89. The Commercial Fish Catch of California for the Year 1951, with an Evaluation of the Existing Anchovy Case Pack Requirements. By the Staff of the Bureau of Marine Fisheries. 1953; 68 pp., 4 figs.

No. 90. Common Marine Bivalves of California. By John E. Fitch. 1953; 102 pp., 63 figs.

No. 91. Common Ocean Fishes of the California Coast. By Phil M. Roedel. 1953; 184 pp., 175 figs.

No. 92. Studies on Fish Preservation at the Contra Costa Steam Plant of the Pacific Gas and Electric Company. By James E. Kerr. 1953; 66 pp., 36 figs.

No. 93. The Life History of the Cabezon, Scorpaenichthys marmoratus (Ayres). By Charles P. O'Connell. 1953; 76 pp., 40 figs.

No. 94. The Behavior and Reproduction of Salmonid Fishes in a Small Coastal Stream. By John C. Briggs. 1953; 62 pp., 5 figs.

[*] No. 95. The Commercial Fish Catch of California for the Year 1952 with Proportion of King and Silver Salmon in California's 1952 Landings. By the Staff of the Marine Fisheries Branch. 1954; 64 pp., 7 figs.

No. 96. California Fishing Ports. By W. L. Scofield. 1954; 159 pp., 88 figs.

[*] No. 97. A Descriptive Study of Certain Tuna-like Fishes. By H. C. Godsil. 1954; 185 pp., 93 figs.

No. 98. The Life Histories of the Steelhead Rainbow Trout (Salmo gairdneri gairdneri) and Silver Salmon (Oncorhynchus kisutch): with Special Reference to Waddell Creek, California, and Recommendations Regarding Their Management. By Leo Shapovalov and Alan C. Taft. 1954; 375 pp., 32 figs.

[*] No. 99. A Description of Two Species of Bonito Sarda orientalis and S. chiliensis and a Consideration of Relationships within the Genus. By H. C. Godsil. 1955; 43 pp., 14 figs.

No. 100. Catch Localities for Pacific Albacore (Thunnus germo) Landed in California, 1951 through 1953. By Harold B. Clemens. 1955; 28 pp., 24 figs.

No. 101. Age Determination of the Northern Anchovy, Engraulis mordax. By Daniel J. Miller and others. 1955; 66 pp., 19 figs.

No. 102. The Marine Fish Catch of California for the Years 1953 and 1954, with Jack Mackerel and Sardine Yield per Area from California Waters 1946–47 through 1954–55. By the Staff of the Marine Fisheries Branch. 1956; 99 pp., 10 figs.

No. 103. Trolling Gear in California. By W. L. Scofield. 1956; 45 pp., 16 figs.

No. 104. A Review of the Rockfishes of California (Family Scorpaenidae). By Julius B. Phillips. 1957; 158 pp., 66 figs.

No. 105. The Marine Fish Catch of California for the Years 1955 and 1956, with Rockfish Review. By the Staff of the Marine Resources Operations. 1958; 104 pp., 9 figs.

No. 106. Age and Length Composition, Pacific Coast Catches, Sardines and Pacific Mackerel 1955–56 and 1956–57 Seasons, and the Northern Anchovy 1954–55 through 1956–57 Seasons. 1958; 72 pp.

No. 107. Trout and Salmon Culture (Hatchery Methods). By Earl Leitritz. 1959; 169 pp., 58 figs.

No. 108. The Marine Fish Catch of California for the Years 1957 and 1958. By the Biostatistical Section, Marine Resources Operations. 1960; 74 pp., 2 figs.

No. 109. The Barred Surfperch, Amphistichus argenteus (Agassiz) in Southern California. By John G. Carlisle, Jr., Jack W. Schott and Norman J. Abramson. 1960; 79 pp., 40 figs.

No. 110. A Study of the Yellowtail, Seriola dorsalis (Gill). By John L. Baxter and a Staff of Associates. 1960; 96 pp., 46 figs.

No. 111. The Marine Fish Catch of California for the Year 1959. By The Biostatistical Section, Marine Resources Operations. 1960; 44 pp., 2 figs.

No. 112. Relationships of Some Marine Organisms of the Northeast Pacific to Water Temperatures, Particularly During 1957 Through 1959. By John Radovich. 1961; 62 pp., 13 figs.

No. 113. The Ecology of the Salton Sea, California, in Relation to the Sportfishery. Edited by Boyd W. Walker. 1961; 204 pp., 80 figs.

No. 114. An Evaluation of Stocking Hatchery-Reared Steelhead Rainbow Trout (Salmo gairdnerii gairdnerii) in the Sacramento River System. By Richard J. Hallock, William F. Van Woert, and Leo Shapovalov. 1961; 74 pp., 17 figs.

No. 115. The Migration, Age, and Growth of Pacific Albacore (Thunnus germo), 1951–1958. By Harold B. Clemens. 1961; 128 pp., 56 figs.

No. 116. Fitting a Von Bertalanffy Growth Curve by Least Squares, Including Tables of Polynomials. By Patrick K. Tomlinson and Norman J. Abramson. 1961; 69 pp.

No. 117. The Marine Fish Catch of California for the Year 1960. By The Biostatistical Section, Marine Resources Operations. 1961; 45 pp., 2 figs.

No. 118. California Abalones, Family Haliotidae: By Keith W. Cox. 1962; 133 pp., 70 figs.

No. 119. Growth Characteristics of Two Southern California Surffishes, the California Corbina and Spotfin Croaker, Family Sciaenidae. By David C. Joseph. 1962; 54 pp., 18 figs.

No. 120. Estimating Absolute Age Composition of California Salmon Landings. By Joseph H. Kutkuhn. 1962; 47 pp.

FOOTNOTES

1. Out of print.