Part 1. Introduction

1.1. Introduction

In this monograph we present the statistical theory and its application for experiments to estimate survival probabilities (rates) of fish passing through hydroelectric dams and spillways in the Columbia River in the northwestern United States. The application of the methods developed here is more general, however, as it includes experiments to estimate survival of fish as they pass over spillways or through bypass systems and several dams. Additionally, this class of experiments includes studies on many vertebrate populations in which marked animals in control and treatment (dosage) groups are compared. We do not concentrate on these extensions, for they are special cases of the general methodology presented here, but we will illustrate them with some examples.

Fish release programs designed to estimate survival rates have been conducted on the Columbia River for several decades. In the associated literature and reports, little statistical formalism underlies the design or analysis of these research programs. Development of sampling (i.e., statistical) models should be an intrinsic part of knowledge acquisition, opinion, and belief formation (Hacking 1965; Kempthorne and Folks 1971).

1.1.1. Simple Example

A hypothetical example will illustrate a simple survival study and introduce the (often) more complex experiments to be addressed later. Consider a team of fisheries biologists concerned with the survival of young steelhead *Salmo gairdneri* as they pass through turbines in a hydroelectric dam on a large river. As has been the practice for many years, large numbers of hatchery-reared fish are marked and about half are released (releases = R) above the dam into the turbine intakes (let this number be R_{t1}) while the other half are released below the dam in the tailrace area (let this number be R_{c1}). The marks are different for the two groups: t = treatment and c = control. The releases are made at dam 1, and some fish are recaptured at three downstream dams (dams 2, 3, and 4). Let $m_{t\,i}$ and $m_{c\,i}$ be the number of treatment and control fish, respectively, recaptured at downstream dam i (i = 2, 3, 4). These data can be represented symbolically and numerically as:

Treatment group	Released at dam 1	Recaptured at dam 2	Recaptured at dam 3	Recaptured at dam 4
t	R_{t1} = 10,000	m_{t2} = 390	m_{t3} = 480	m_{t4} = 83
c	R_{c1} = 9,000	m_{c2} = 412	m_{c3} = 530	m_{c4} = 83.

For simplicity, we assume that once a fish is captured at one of the three downstream dams, it is removed from the study and not rereleased. The recapture data thus come from a sample of the marked cohorts initially released.

How can we then estimate the survival rate (S) through the turbine from data such as these? The ratio of the two recapture rates at the first downstream dam is an estimate of the survival rate:

$$(m_{t2}/R_{t1})/(m_{c2}/R_{c1}) = (390/10{,}000)/(412/9{,}000) = 0.852.$$

Assuming passage through the turbine has no effect on downstream fish behavior, this estimator is essentially unbiased but it is inefficient because it does not use all the data relevant to estimation of the survival parameters. Likewise, the separate estimators, based only on data from dams 3 or 4,

$$(m_{t3}/R_{t1})/(m_{c3}/R_{c1}) = (480/10{,}000)/(530/9{,}000) = 0.815$$

and

$$(m_{t4}/R_{t1})/(m_{c4}/R_{c1}) = (83/10{,}000)/(83/9{,}000) = 0.900\,,$$

are poor for the same reason. One might try to pool the data, by weighting each of these three estimators, to get an improved estimator of survival. Equal weights might be desirable; however, one could weight the three estimates by the total number of fish captured, or by the number of treatment or control fish captured. Other ways of pooling data are also possible.

Without a formal theory underlying this class of studies, it is impossible to proceed in a rigorous manner. Lacking the necessary basis of a stochastic theoretical model, it is equally difficult to estimate the theoretical precision associated with estimates of survival rate. A proper estimator of the sampling variance is important as a measure of an estimator's precision or repeatability. A number of assumptions must be made in studies of this type, and these assumptions must be stated clearly because they form the basis for a model. Goodness-of-fit tests must be derived in an effort to assess the validity of the assumptions. Finally, one must know the degree to which an estimator of survival rate is sensitive to the partial failure of particular assumptions.

Because intuition was of little help in deriving an estimator and its sampling variance in the simple example, a rigorous approach is required for the more complex cases encountered in real life. These cases include studies in which (1) fish captured at downstream dams are released alive for potential second or third captures, (2) more than one treatment is involved (e.g., three release groups), (3) fish are released with unique tag numbers instead of simple batch marks, (4) fish size is used as a covariate, (5) several replicate releases (lots) are involved, and (6) survival is estimated for several dams and reservoirs in a long reach of the river. We attempt here to establish the analysis and inference theory for this general class of experiments.

This class of experiments is inherently difficult to treat. Unlike more standard experiments (e.g., agricultural field trials), these survival experiments allow less control by the investigators because fish are highly mobile animals whose behavior is poorly understood. Further, the use of sampling methods to reobserve the marked animals imposes substantial complexities. Important assumptions are required, and statistical tests must be made to assess carefully the validity of these assumptions. The fundamental concepts and analysis methods applied here are not trivial.

1.1.2. Historical Note

As our methodological research progressed, we found much existing theory in the scientific literature related directly to the estimation and testing of concern here. Existing theory falls under two broad categories. The first category is band or tag recovery studies, such as the one outlined in Section 1.1.1, where known releases of fish are followed by the removal of recaptured fish from the population upon first capture. The theory for this class of studies dates back to the early 1970s (Seber 1970; Robson and Youngs, unpublished report, 1971) (but see also Seber 1962); much of it was synthesized and extended by the two editions of the handbook of Brownie et al. (1978, 1985). Second, part of the existing theory for the so-called Jolly-Seber model (Jolly 1965; Seber 1965, 1982) was found to be relevant for the problems presented herein in which fish captured at downstream dams are released alive for possible subsequent recapture. In the context of this work, however, we cannot estimate population size or numbers of new recruits (which is possible under the Jolly-Seber framework: e.g., Hightower and Gilbert 1984) because it requires data on, and additional assumptions about, the unmarked members of the population. Brownie et al. (1985) summarized the large literature on Jolly-Seber models. Our primary focus is survival probability within the context of a treatment. We present a series of models, hypothesis tests, and sampling protocols that allow a treatment survival rate to be estimated and evaluated.

These two broad approaches to sampling marked populations are closely related. We have exploited this relationship in the present work, while extending the methodology to enable experiments on marked populations that lead to the assessment of a treatment survival rate. Technical discussions of these relationships were given by Brownie and Robson (1983), Brownie and Pollock (1985), and Brownie et al. (1985:170-175).

Many persons think of the field of statistics in terms of simple *t* and chi-square tests, analysis of variance, regression, and other such methods. In fact, the field is far broader than is suggested by the data analysis methods taught in beginning statistics courses. The field of statistics is not so much a branch of mathematics as it is an area of science concerned with the development of a practical theory of information (White et al. 1982:14). Statistics is primarily concerned with efficient methods of collecting data and establishing rigorous foundations for deriving efficient methods of making inductive inferences from sample data, and it is an integral part of the *scientific method.*

The conclusions drawn from sample data are intended to apply beyond the specific study or experiment. Biologists wish to make generalizations – "inductive inferences" – from a specific study to the population that was sampled. A theorem in logic states that there is uncertainty in inductive inference and that perfect generalizations therefore cannot be made about a population from studying only the sample. However, the degree of uncertainty can be measured if the experiment was performed in accordance with certain scientific principles (Mood et al. 1974). A critical function of the science of statistics is to provide a formalism for making inductive inferences and for measuring their degree of uncertainty (Ostle 1963).

1.1.3. Objectives

The present work is the culmination of several initial objectives. Our overall objective is to present a comprehensive statistical theory to support survival experiments that rely on recapture data collected after release of marked individuals. By an experiment, we explicitly mean at least two treatment levels, i.e., releases of "treatment" and "control" groups (the individuals are marked to reflect what group they are in), with the purpose of comparing results across treatment levels. The methodology developed here addresses both batch and unique marks, more than one treatment, and several other extensions. A series of experimental protocols is defined and maximum-likelihood estimators of parameters are given for each situation. Sampling variances and covariances are given as measures of precision and interdependency, respectively, for parameter estimators.

A second objective deals with hypothesis tests involving various models and assumptions about certain parameters. The interpretation of the test results allows the selection of a proper model for a particular experiment.

In principle, the theory is not complex, but computational details are tedious. Thus, a third objective is to demonstrate the ability of our interactive data analysis computer program RELEASE, which allows biologists to concentrate on the interpretation of experimental results rather than on computational matters. In addition, we wanted to produce a comprehensive monograph giving the relevant background, theory, and application in a way that would be useful to biologists conducting survival experiments. Our monograph illustrates the broad applicability of the underlying theory beyond studies of fish survival and provides an opportunity to illustrate the methods with output from program RELEASE.

Another important objective is to illustrate the performance of the various procedures with sample sizes and recovery rates typically encountered in practice. Therefore, Monte Carlo studies were performed to evaluate bias, confidence interval coverage, robustness, and power of tests. Capabilities for further Monte Carlo studies are incorporated into program RELEASE, allowing the user to study a particular situation.

Our final objective is to outline the design and sample-size requirements for release-recapture survival experiments. In particular, we consider the need for replication in survival studies and provide ways to treat these data in the analysis.

1.1.4. Reader's Guide

We believe everyone should read Part 1 because Part 1 provides background and the notation required for understanding the parts that follow. Persons with good statistical skills could bypass Chapter 1.2 on Statistical Concepts. The concepts in Part 1 are kept at a fairly elementary level, with the exception of Chapter 1.5. We recommend the use of program RELEASE as a learning tool. The data given in Chapter 1.3 are used in many places throughout the monograph. Interpretation of the program output enhances the rate of understanding, especially in becoming familiar with the different models and protocols.

Chapter 2.1 in Part 2 contains difficult material. However, we urge the reader to gain some insight into the concepts given. Chapters 2.2., 2.3, 2.4, and 2.5 are parallel in that they cover the four major sampling protocols. Each of these four chapters contains examples to aid in understanding the analysis theory presented; however, an understanding of Part 1 is assumed. The material in Part 2 deals with experiments involving one treatment and one control. We urge all readers to read at least Chapters 2.1, 2.2, and 2.4. Parts 1 and 2 of this work are written in the context of fish experiments and studies involving large hydroelectric dams. Other parts are less specific, to allow the reader to think more generally about experimental animals and sampling sites or occasions rather than only about fish and dams.

Part 3 is optional reading for biologists, unless their interest lies in more elaborate experiments involving two or more treatments and single or multiple control groups. Understanding Parts 1 and 2 is required to understand Part 3. In contrast, statisticians may want to scan the general theory presented in Part 3 before examining the various special cases given in later chapters.

The subject of replication in Part 4 is essential reading for all user groups. In particular, those considering the design of experiments should understand the need for replication. Biologists may want to postpone reading Part 5, but statisticians probably will want to consider this material in detail.

Part 6 relates to experimental design, and anyone contemplating this problem should find useful suggestions here. However, we caution the reader that only statistical features are stressed in Part 6; biological aspects of these experiments are not provided.

In Part 7, we examine some case studies in an attempt to provide insight into the generality of the methods developed here, and extend the coverage to other vertebrates. We urge all readers to study these examples.

Part 8 is for readers who have a good understanding of the preceding material. Part 8 is not meant to be complete; it is only a brief introduction to some of the extensions that are possible to develop.

Anyone expecting to make intensive use of computer program RELEASE should read Part 9. RELEASE is relatively easy to use if one follows the details given there. Most users will find the interactive features of RELEASE self-explanatory. We urge readers to use this program to work examples as they read Part 2. Finally, we have provided a Glossary of the notation we use.

We wish we could refer biologists to only one short section that would allow a quick understanding of the material presented here. Unfortunately, the subject is large, comprehensive, and complex. At the minimum, we believe most biologists should read most of Parts 1, 2, 4, and 9. We would hope statisticians would focus on Parts 3, 5, and 6, and develop additional theory (e.g., as indicated in Part 8) as well as explore properties of procedures developed using RELEASE.

1.2. Statistical concepts

1.2.1. Maximum Likelihood Theory

Fisher (1922, 1925) presented the method of maximum likelihood as an omnibus procedure for estimating parameters from sample data. Extensions of this method have produced a well-known, powerful means to derive point estimators and estimators of sampling variances and covariances. Likelihood theory allows one to assess the fit of the data to the model, and to test a variety of hypotheses (Mood et al. 1974; Lehmann 1983; Berger and Wolpert 1984). Parameter estimators, and likelihood-based inference in general, have excellent properties such as little or no bias and maximum efficiency. Likelihood methods have been the mainstay of most capture-recapture theory developed over the past 35 years (see Seber 1982), and we use them extensively in this monograph.

Data discussed in the present work are usually modeled as a sample from a multinomial distribution. The multinomial distribution is useful for discrete, mutually exclusive outcomes. The outcomes of each throw of a die ("trial") can be labeled 1, 2, 3, 4, 5, and 6, only one of which is possible. If n throws are made, frequencies of the six possible outcomes can be denoted as $n_1, n_2, ..., n_6$. Given that the n throws of the die are independent, the joint probability distribution (Pr) of the observed data is multinomial:

$$\Pr\{n_1, ..., n_6\} = \frac{n!}{\prod_{i=1}^{6}(n_i)!}\left[\prod_{i=1}^{6}(p_i)^{n_i}\right];$$

p_i = probability of the ith outcome, $i = 1, ..., 6$. The probabilities must satisfy the constraints $0 \leq p_i \leq 1$ for all i, and $p_1 + \cdots + p_6 = 1$. If the die is fair, each p_i is known to be one-sixth. In general, models contain parameters that are unknown; we wish to find estimators of these parameters that have "good" properties. The estimators are functions of the observed sample data and can be derived from the likelihood function L, which is the probability of the observed data viewed as a function of the parameters. The likelihood function for the die trials is

$$L(p_1, ..., p_6 \mid n_1, ..., n_6) = \frac{n!}{\prod_{i=1}^{6}(n_i)!}\left[\prod_{i=1}^{6}(p_i)^{n_i}\right]$$

and is read "the likelihood of the unknown parameters p_i given the observed (sample) data n_i."

The objective is to find the vector of parameter values that maximizes the likelihood function. That is, parameter values are selected to make the sample data seem "most likely." A simple graph of the likelihood can illustrate the concept; however, the above likelihood function cannot be graphed because it has five dimensions (only five because $p_1 + \cdots + p_6 = 1$). Therefore, let us consider a special case where only one unknown parameter exists. We consider a series of simple penny-flipping trials of an unfair penny. The likelihood function L is

$$L(p \mid n_h, n_t) = \frac{n!}{(n_h)!(n_t)!} p^{n_h}(1-p)^{n_t}.$$

Here p is the unknown probability of a "head" and n_h and n_t are the number of heads and tails, respectively, observed from n flips ($n = n_h + n_t$) and $1 - p$ is the probability of a "tail."

If a penny is flipped 16 times and 11 heads and five tails are observed, the likelihood shown in Figure 1.1 is a simple plot of the function $L(p \mid 11, 5)$ versus p where $0 < p < 1$. The likelihood function changes for different sample outcomes n_h and n_t. This change is illustrated in Figure 1.2 for the result of 80 flips ($n_h = 55$ and $n_t = 25$).

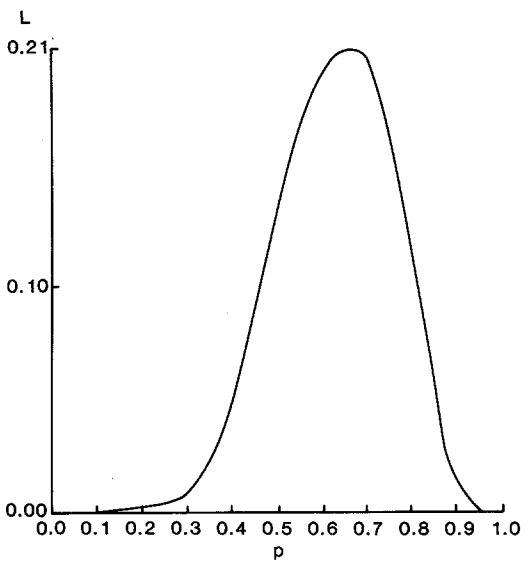

Figure 1.1. – Likelihood function for a penny-flipping study where 11 heads and five tails were observed in 16 flips.

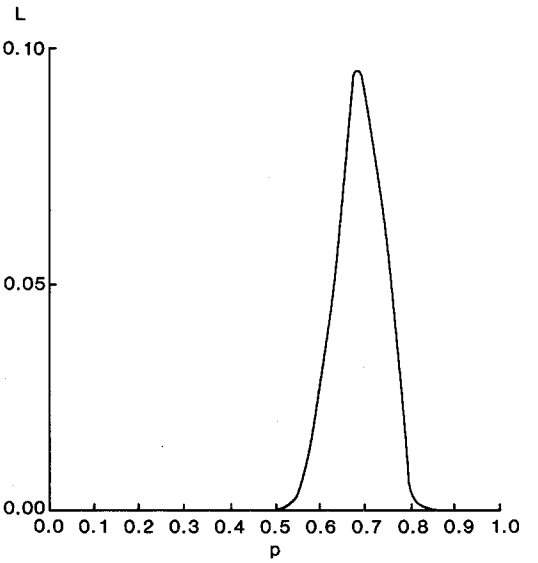

Figure 1.2. – Likelihood function for a penny-flipping study where 55 heads and 25 tails were observed in 80 flips.

The graphs show that some candidate values of the unknown parameter p are "relatively unlikely" (i.e., those with small values of L) given the data observed. This phenomenon is more obvious as sample size increases because more information becomes available. For example, it appears that the penny is indeed unfair (the true p is probably not one-half). In fact, one might speculate that a likely value for the parameter p might be about 0.7 because the maximum value of the likelihood function corresponds to p at about 0.7.

The following extended example considers a likelihood function involving two unknown parameters. Consider the previous example of released fish. For simplicity, it is assumed that the marked fish were recaptured only at downstream dam 2 (a common situation). The data available for consideration are then

Group	Released	Recaptured
t	10,000	390
c	9,000	412.

The likelihood involves two estimable parameters, S, the survival rate of fish through (or over) the dam, and p_2 – the sampling rate at dam 2 (in this simple example it is assumed that $p_{t2} = p_{c2}$), given that the fish were alive just below the first dam (where controls were released).

Assumptions must always be made in any model. In this model, we assume independent fates of all fish, which allows us to write separate probability models for the treatment and control groups and to assume that the binomial model holds. We also assume that recapture of fish is like the penny-flipping situation with parameter Sp_2 (the probability of surviving the treatment, getting to dam 2, and being recovered) for the treatment group and p_2 for the control group. Finally, we assume that the effect of the treatment is negligible below dam 1. These assumptions allow us to write the two following probability models,

$$\Pr\{m_t \mid R_t\} = K_t(Sp_2)^{m_t}\,(1 - Sp_2)^{R_t - m_t}, \text{ and}$$

$$\Pr\{m_c \mid R_c\} = K_c(p_2)^{m_c}\,(1 - p_2)^{R_c - m_c};$$

$$K_t = \frac{(R_t)!}{(m_t)!(R_t - m_t)!}\;;\; K_c = \frac{(R_c)!}{(m_c)!(R_c - m_c)!}.$$

In this example, one would take

$$\Pr\{m_c = 412 \mid R_c = 9{,}000\} = \frac{9{,}000!}{(412!)(8{,}588!)}\,(p_2)^{412}\,(1 - p_2)^{8{,}588}.$$

The product of these expressions yields the likelihood function for this simple experiment (terms are rearranged below):

$$L\,(S, p_2 \mid R_t, R_c, m_t, m_c) = (K_1K_2)\,(Sp_2)^{m_t}\,(p_2)^{m_c}\,(1 - Sp_2)^{R_t - m_t}(1 - p_2)^{R_c - m_c}.$$

Four terms involving parameters appear in the likelihood because we must account for all fish both recovered and unrecovered from each of the two release groups. The expression on the left is read "the likelihood of the unknown parameters S and p_2, given the number of fish released in each group (R_t and R_c) and the number recovered at dam 2 (m_t and m_c), is equal to." Figure 1.3 provides a two-dimensional graph of the likelihood

$$L\,(S, p_2 \mid \text{data}) = K(Sp_2)^{390}\,(p_2)^{412}\,(1 - Sp_2)^{9{,}610}\,(1 - p_2)^{8{,}588};$$

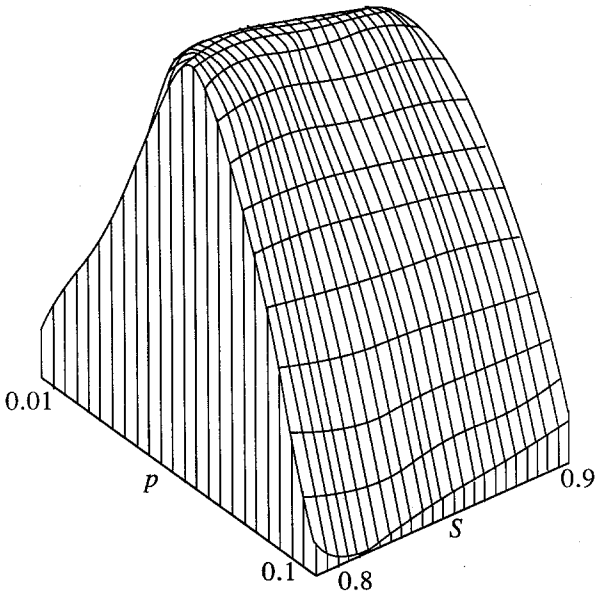

Figure 1.3. – Likelihood function for a simple survival experiment in which treatment and control fish are released at dam 1 and recaptured only at dam 2. Note that many values of the capture probability (p_2) and the treatment survival rate (S) are relatively unlikely.

$$K = K_1 K_2 = \left[\frac{10{,}000!}{390!\,(9{,}610)!} \right] \left[\frac{9{,}000!}{412!\,(8{,}588)!} \right].$$

The expression for K contains no unknown parameters and can therefore be ignored in terms of deriving estimators of parameters (however, K is useful in deriving tests of assumptions). The relative shape of the likelihood function is, of course, identical whether K is included or not.

Figure 1.3 also shows that most possible combinations of the parameters S and p_2 are unlikely for the data observed.

1.2.1.1. Point estimation. – The likelihood principle (Edwards 1972) states that the likelihood function contains all the information in the sample data and is the basis for deriving estimators of parameters and tests of assumptions. Estimators of parameters under the method of maximum likelihood (ML) are the values that maximize the likelihood. The ML procedure is conceptually appealing and has many optimal statistical properties, at least for large samples (Wilks 1962; Rao 1973; Mood et al. 1974) when the assumed model is true: little or no bias and 100% efficiency. In particular, the ML method provides estimators that are asymptotically normal, efficient, and unbiased.

The ML procedure is easy to apply in many situations. Consider a likelihood involving only one unknown parameter θ, $L(\theta|\text{data})$. If L is unimodal, its maximum occurs where the derivative (slope) with respect to θ is zero:

$$\frac{\partial L(\theta \mid \text{data})}{\partial \theta} = 0.$$

In practice, if the derivatives are to be found and the likelihood equations solved analytically, it is easier to work with the natural logarithm of the likelihood function, the log-likelihood, denoted as $\ln L(\theta \mid \text{data})$. Logarithms change the likelihood function from a product of terms to a sum of terms, and thus allows differentiation on a term-by-term basis. For the penny-flipping trials mentioned previously, the log-likelihood function (omitting the constant binomial coefficient) is

$$\ln L(p \mid \text{data}) = n_h \ln(p) + n_t \ln(1-p)$$

(in this case the generic parameter θ is denoted p as before). The partial derivative of $\ln L$ with respect to the unknown parameter p is

$$\frac{\partial[\ln L(p \mid \text{data})]}{\partial p} = \frac{n_h}{p} - \frac{n_t}{1-p} = 0.$$

The maximum likelihood estimator (MLE) of p is found by solving the likelihood equation for p. The resulting estimator is traditionally denoted as $\hat{p}$; in general, hats ("^") are used to distinguish estimators from parameters:

$$\hat{p} = n_h/(n_h + n_t), \text{ or } \hat{p} = n_h/n.$$

For the data observed, $n_h = 55$ and $n_t = 25$,

$$\hat{p} = 55/(55 + 25) = 0.69.$$

The example above provides an estimator said to be in "closed form," meaning that the likelihood equation(s) can be solved analytically for the parameter(s) of interest. If the likelihood equations are "open," i.e., cannot be solved algebraically, ML estimates can be derived by a variety of numerical methods on a digital computer, which performs an "intelligent search" for the parameter values that maximize the likelihood function. When values that simultaneously maximize L are found, they are taken as the MLEs.

In practical applications, one is interested in simultaneously deriving ML estimates of several parameters, say θ_1, θ_2, and θ_3. If these parameters are denoted as $\underline{\theta}$, a column vector containing $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\theta}_3$, the MLEs can be found by solving the system of three likelihood equations,

$$\frac{\partial \ln L(\underline{\theta} \mid \text{data})}{\partial \theta_i} = 0, \quad i = 1, 2, 3.$$

Although the notation and algebra may not appear simple, the underlying concept of ML is both simple and intuitively appealing. If the likelihood equations can be solved analytically, ML estimates can be computed using a calculator. In the present work, we present closed-form estimators of parameters for most of the models under the various protocols. The computer program RELEASE computes ML estimates for an array of specific models to be introduced in Part 2.

Given a set of data and a formal statistical model (i.e., a likelihood function), one can find a reduction of the data. Thus, a smaller set of statistics, which contains all the information in the sample data, can be used instead of the raw data for all statistical estimation purposes. This reduction leads to the concept of sufficient statistics. A sufficient statistic is one containing all the necessary information about the sample. A sufficient statistic that cannot be reduced further is a minimal sufficient statistic (MSS) (see Hogg and Craig 1970; Huzurbazar 1976; Lehmann 1983).

Minimal sufficient statistics are important for a number of statistical reasons (see Mood et al. 1974). If an estimator is not based on the MSS, it is not fully efficient. If an MSS exists, the ML method can be used to find it. MLEs are functions of the MSS. In addition, test derivation often depends on the MSS. The MSS will be identified for the various models in later chapters.

1.2.1.2. Estimation of sampling variances and covariances. – A sampling variance is associated with each estimator and a sampling covariance is associated with each pair of estimators. The sampling variance is a measure of the precision or repeatability of the estimator and is usually a function of sample size and some of the unknown parameters. Sampling covariances measure the degree to which two particular estimators are dependent because they were computed from the same sample data. Often, these quantities are displayed in a *variance-covariance matrix*, usually denoted $\underline{\Sigma}$:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{var}(\hat{\theta}_1) & \text{cov}(\hat{\theta}_1, \hat{\theta}_2) & \text{cov}(\hat{\theta}_1, \hat{\theta}_3) & \cdots & \text{cov}(\hat{\theta}_1, \hat{\theta}_n) \\ \text{cov}(\hat{\theta}_2, \hat{\theta}_1) & \text{var}(\hat{\theta}_2) & \text{cov}(\hat{\theta}_2, \hat{\theta}_3) & \cdots & \text{cov}(\hat{\theta}_2, \hat{\theta}_n) \\ \text{cov}(\hat{\theta}_3, \hat{\theta}_1) & \text{cov}(\hat{\theta}_3, \hat{\theta}_2) & \text{var}(\hat{\theta}_3) & \cdots & \text{cov}(\hat{\theta}_3, \hat{\theta}_n) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \text{cov}(\hat{\theta}_n, \hat{\theta}_1) & \text{cov}(\hat{\theta}_n, \hat{\theta}_2) & \text{cov}(\hat{\theta}_n, \hat{\theta}_3) & \cdots & \text{var}(\hat{\theta}_n) \end{bmatrix}$$

The variances appear on the main diagonal and the covariances are symmetrical about the diagonal: i.e., $\text{cov}(\hat{\theta}_i, \hat{\theta}_j) = \text{cov}(\hat{\theta}_j, \hat{\theta}_i)$ for all i and j. Often, it is convenient to consider the variance of the ith estimator $[\text{var}(\hat{\theta}_i)]$ as the $\text{cov}(\hat{\theta}_i, \hat{\theta}_i)$, referring to a covariance matrix. For any actual study, one will have only estimates of the variance and covariances, and $\hat{\boldsymbol{\Sigma}}$ usually is obtained by substituting estimates for parameter values. Conceptually, the sampling variance is related directly to the curvature of the likelihood function at its maximum. A comparison of Figures 1.1 versus 1.2 suggests that more is known from the large sample shown in Figure 1.2 than from the smaller one. Values at some distance from the ML estimate are relatively "unlikely," and this concept is measured by the sampling variance.

If the likelihood function contains only one parameter, the sampling variance estimator can be derived as the negative inverse of the second partial derivative of the log-likelihood function, evaluated at the ML estimate. In the previously described penny-flipping example,

$$\text{var}(\hat{p}) = -\left[E\left(\frac{\partial^2 \ln L(p \mid \text{data})}{\partial p^2}\right)\right]^{-1},$$

which is estimated by

$$\hat{\text{var}}(\hat{p}) = \left[-\left(\frac{\partial^2 \ln L(p \mid \text{data})}{\partial p^2}\right)\right]^{-1}_{p=\hat{p}}.$$

The procedure yields, for example, the often-used estimator of the variance of a binomial proportion,

$$\hat{\text{var}}(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n}, \; n = n_h + n_t.$$

If the likelihood contains more than one unknown parameter, the sampling variances and covariances can be estimated from the negative of the matrix of mixed second-order partial derivatives of the log-likelihood function. The resulting matrix, expressed as expected partial derivatives, is called the information matrix. Because the information matrix is a function of the parameters, it is denoted here as $I(\underline{\theta})$. Specifically, the (ij)th element of $I(\underline{\theta})$ is given by

$$-E\left[\frac{\partial^2 \ln L(\underline{\theta} \mid \text{data})}{(\partial\theta_i)(\partial\theta_j)}\right],$$

the quantity evaluated at the true parameter value, $\underline{\theta}$.

If the elements of the information matrix are evaluated at the ML estimates, then the estimated variance-covariance matrix is

$$\hat{\underline{\Sigma}} = [I(\underline{\hat{\theta}})]^{-1}.$$

Most numerical methods use a matrix of mixed second partial derivatives (the Hessian matrix) to find the maximum of the likelihood function and, therefore, the ML estimates $\hat{\underline{\theta}}$. In this case, $\hat{\underline{\Sigma}}$ can be estimated with $I(\hat{\underline{\theta}})$ from the final iteration in the numerical procedure (see Kale 1962).

The delta method (Seber 1982:7-8) provides an omnibus procedure for approximating estimates of sampling variances and covariances. The method gives valid large-sample (i.e., asymptotic) estimates of variances and covariances and produces results asymptotically equivalent to the information matrix approach. We use this method at many points in the monograph. In addition, we encourage the use of empirical estimates of sampling variances in an effort to relax the assumptions made in some models and to allow for heterogeneity in the survival and recapture probabilities of individual fish (details on this subject are given in Part 4).

We chose the ML method for our development of statistical theory because it is an excellent omnibus approach (within a frequentist inference approach) and is particularly well suited for models based on the multinomial distribution (a member of the exponential family of distributions). Estimation methods are then based on a general stochastic model for the sampling distribution of the data. MLEs are asymptotically unbiased, fully efficient, and normally distributed. Some estimators have a "small-sample" bias and RELEASE allows bias adjustments to the exact MLEs to be made as an option.

1.2.1.3. Method of expectation. – A critical step in many statistical analysis problems is to specify one or more plausible sampling models for the data. Conceptually, these sampling models take the form of probability distributions; hence, the essence of the model is a mathematical statement. Symbolically the statement is

$$\Pr\{\text{data}|\text{parameters}\},$$

or

$$\Pr\{X_1, ..., X_n|\theta_1, ..., \theta_a\}.$$

As introduced above, once one has actually observed specific values of the variables, then one can convert the probability model to a likelihood in terms of the parameters and derive ML estimates. Rather than do this process on a case-by-case basis, mathematical statisticians have investigated short-cut ways to derive the ML estimates, $\hat{\theta}_1, ..., \hat{\theta}_a$. In particular, when the $\hat{\theta}_i$ exist in closed form, they can often be found by simple methods. We make extensive use of these more sophisticated methods here; we do not, in fact, find it necessary to write out the full likelihood for our recapture models, take partial derivatives, and solve the resultant equations.

The first, and most important, step in an analysis of a statistical model is to use the likelihood to identify the MSS under that model. There is much theory about finding the MSS (e.g., Lehmann 1983). The MSS will take the form of some l functions of the data, $T_i(X_1, ..., X_n)$, $i = 1, ..., l$. For example, in many situations the mean, $\overline{X}$, is one component of the MSS. In the capture-recapture models considered here, the MSS always turns out to be various sums of recapture counts. The dimension of the MSS is l.

Two situations distinguish themselves: (1) the dimension of the MSS is the same as the number of parameters (i.e., $l = a$), or (2) there are more MSS components than parameters ($l > a$). (The case of $l < a$ can occur, in which case not all a parameters can be estimated.) Once the MSS is known, the next step is to find its probability distribution (if possible), or at least find the expected values, variances, and covariances of the MSS. Let

$$E(T_i) = g_i(\theta_1, ..., \theta_a), \quad i = 1, ..., l$$

be these expected values. In the full-rank case of $l = a$, the ML estimators of the θ_i can be found by solving the a equations that result from equating the observed values of T_i to the expected values:

$$T_i = g_i(\theta_1, ..., \theta_a), \quad i = 1, ..., a.$$

Most of the models considered in detail in this monograph are full rank and the "method of expectation" is how we found the ML estimates. Davidson and Solomon (1974) provided insights on the justification of this procedure. Also, Appendix B of Brownie et al. (1985) gives some more technical details of this approach.

When the model is not full rank, but rather has $l > a$ (fewer parameters than MSS elements), iterative numerical methods are usually needed to find the MLE; i.e., closed-form solutions do not usually exist. Many models worth using or trying in capture-recapture are not full rank. (Those models can be analyzed using a combination of programs RELEASE and

SURVIV.) Appendix B in Brownie et al. (1985) gives some details of numerical procedures of parameter estimation with nonfull-rank models.

Whether the ML estimator $\hat{\theta}$ is obtained from closed-form formulae or numerically, it still has a variance-covariance matrix that can be estimated. Theoretically, those variances and covariances are defined in terms of second derivatives of the likelihood function. However, they need not always be found that way; there are short cuts to finding variances just as there are short cuts to finding the ML estimators. We have used those short cuts here. In the full-rank case, all one needs is the expectations and variances and covariances of the MSS. Then the link provided by

$$E(T_i) = g_i(\underline{\theta}), \quad i = 1, ..., a$$

allows one to derive both $\hat{\theta}$ and the variance-covariance matrix of $\hat{\theta}$. Consequently, again we derived our results without having to take first and second partials of the likelihood. Note, however, results we present on the various models considered could have been derived by taking partials of likelihoods; it is just that there are easier, advanced methods available.

Finally, we again note that the key to these derivations of estimators (and tests) is identification of the MSS under any model and model sequence, and then the determination of the sampling distribution of that MSS.

1.2.1.4. Hypothesis testing. – Tests of various hypotheses are important in capture-recapture sampling and experimentation. Cormack (1968) stated, "In all cases every iota of information, both biological and statistical, must be gathered to check and countercheck the unavoidable assumptions." Much of the hypothesis testing herein relates to tests of underlying model assumptions or to selection of an appropriate model. Hypothesis tests, in the context here, fall into two broad classes: goodness of fit tests and between-model tests. The difference is illustrated by considering the null (H_0) and alternative (H_A) hypotheses for each type of test.

The null hypothesis for a goodness of fit test is that the model fits the data; the alternative hypothesis is that the model does not fit the data. The alternative is broad and not specific. We can expect the power of this test to be lower (often much lower) than the between-model test due to the generality of the alternative.

Between-model tests deal with a comparison of two specific models, e.g., model A and model B, wherein model A is a special case because it is a reduced parameter version of model B. The test evaluates model B as an alternative to model A (the null hypothesis). Specifically, the null hypothesis is that model A fits as well as model B; the alternative hypothesis is that model B fits the data better than model A.

Both tests involve a statement about one or more parameters in the model. Information regarding the validity of the null hypothesis is based on the value of a test statistic calculated from the experimental data. Most of the test statistics here are distributed approximately as chi-square under the null hypothesis. The approximation to the chi-square distribution improves as sample size increases. Figure 1.4 presents the chi-square distributions for 1, 3, 10, and 25 degrees of freedom (df). The probability of a test statistic being as large as that observed under H_0 (i.e., computed from the data) can be found from the chi-square distribution. Generally, if a test statistic is improbable (e.g., $P = 0.002$), the null hypothesis is rejected. Improbable values ($\alpha \leq 0.05$) of the chi-square distribution are shown in the shaded areas of Figure 1.4. Conversely, if the test statistic is probable under the null hypothesis (e.g., $P = 0.45$), there is no reason to reject the null hypothesis.

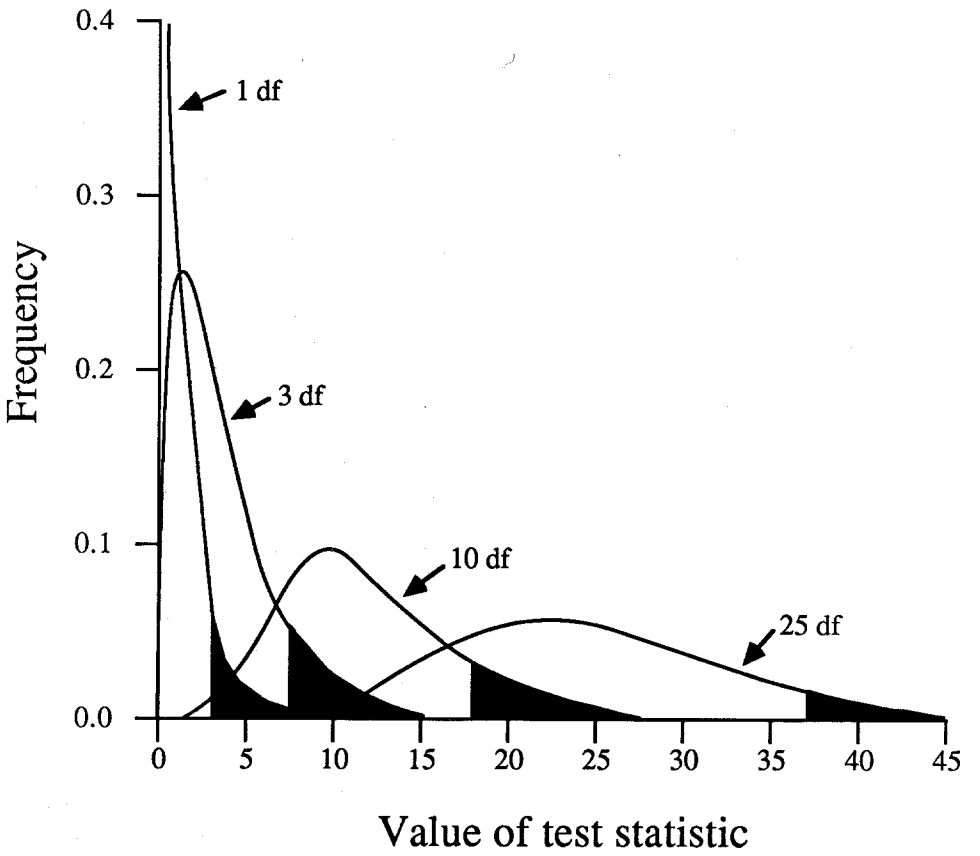

Figure 1.4. – The chi-square distributions for 1, 3, 10, and 25 df. In each case, the 0.05 rejection region is shown as a shaded area. Nearly all test statistics presented in this monograph are distributed asymptotically as chi-square. Program RELEASE computes the exact significance level numerically, making tables or arbitrary rejection levels unnecessary (from White et al. 1982).

As an example, consider the null hypothesis (H_0: $p = 0.5$) that a penny is "fair" (on the average 50% of the tosses will be heads and 50% will be tails). The alternative hypothesis (H_A: $p \neq 0.5$) is that the penny is "unfair." The results of a set of 1,000 trials might be that 506 tosses are heads and 494 tosses are tails. Intuitively, one might accept the null hypothesis that the penny is fair because the observed values are close to 50:50. Statistical hypothesis testing allows this intuition to be quantified and formalized. In this example, $\chi^2 = 0.1$ with 1 df. Figure 1.4 shows that this small value is likely if the null hypothesis is true. If, however, 55 heads are observed in 80 flips, $\chi^2 = 56$ with 1 df. The probability of a value this large, if the null hypothesis is true, is essentially zero and we conclude that the penny is unfair. Quantification allows statistical inferences to be made in complex situations where intuition is of little value.

Most of the tests presented in this monograph are in the form of contingency tables. The subject of contingency tables is covered in most statistical texts on elementary testing methods; we provide only a brief review here. Consider n randomly selected items classified according to two different criteria. The results could be tabulated by rows for one criterion and by columns for the second criterion in a contingency table:

n_{11}	n_{12}	n_{13}	$\cdots$	n_{1c}	$n_{1.}$
n_{21}	n_{22}	n_{23}	$\cdots$	n_{2c}	$n_{2.}$
n_{31}	n_{32}	n_{33}	$\cdots$	n_{3c}	$n_{3.}$
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
n_{r1}	n_{r2}	n_{r3}	$\cdots$	n_{rc}	$n_{r.}$
$n_{.1}$	$n_{.2}$	$n_{.3}$	$\cdots$	$n_{.c}$	n

This $r \times c$ table contains the observed data for each cell. If n is large, a good approximation is to compute the test statistic

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} (n_{ij} - E_{ij})^2 / E_{ij} ;$$

n_{ij} = the observed number in the (ij)th cell;
E_{ij} = the estimated expected number in the (ij)th cell under the null hypothesis.

The null hypothesis is one of homogeneity. The estimated expected value for n_{ij} under this null hypothesis is then

$$E_{ij} = \frac{n_{i.}n_{.j}}{n};$$

$n_{i.}$ = the total of row i, $n_{i.} = n_{i1} + \cdots + n_{ic}$;
$n_{.j}$ = the total in column j, $n_{.j} = n_{1j} + \cdots + n_{rj}$.

Note also that $n \equiv n_{..} = \Sigma\Sigma n_{ij}$. The degrees of freedom for these contingency tables are $(r - 1)(c - 1)$ (Ostle 1963; Snedecor and Cochran 1967).

We present many tests that are computed from a series of (conditionally) independent 2 × 2 tables. Standard statistical texts provide shortcut formulae for the computation of the test statistic from such tables with only two columns and two rows. A convenient computational method for a 2 × 2 table of the form

a	b	(a + b)
c	d	(c + d)
(a + c)	(b + d)	n

is

$$\chi^2 = \frac{n(ad - bc)^2}{(a + c)(b + d)(a + b)(c + d)}.$$

Such 2 × 2 tables have only 1 df (i.e., [2-1][2-1] = 1). By way of interpretation of this test, it is just testing that the expected proportions $a/(a + b)$ and $c/(c + d)$ are the same.

If some expected values (E_{ij}) are small (e.g., <2), the approximation of the test statistic to the chi-square distribution may be poor. This approximation can be improved by pooling cells by row or by column. Each cell pooled results in a loss of 1 df. When recapture data are sparse, rows and columns are often pooled to the extent that only a single 2 × 2 table remains.

Here we present material to provide biologists with some insight into how certain tests are derived using a simple example. No effort is made to provide the theory for the procedures outlined. The example is taken from Section 1.2.1: 10,000 treatment and 9,000 control fish were released at dam 1 and 390 treatment and 412 control fish were recaptured. The probability models for these two cohorts, presented earlier, are

$$\Pr\{m_t \mid R_t\} = K_t (Sp_2)^{m_t} (1 - Sp_2)^{R_t - m_t}, \text{ and}$$

$$\Pr\{m_c \mid R_c\} = K_c (p_2)^{m_c} (1 - p_2)^{R_c - m_c};$$

$$K_t = \binom{R_t}{m_t} = \frac{(R_t)!}{(m_t)!\,(R_t - m_t)!};$$

$$K_c = \binom{R_c}{m_c} = \frac{(R_c)!}{(m_c)!\,(R_c - m_c)!}.$$

Interest is in the hypotheses

H_0: The survival and capture rates for the two cohorts are equal (i.e., $S = 1$; no treatment effect) versus

H_A: The survival and capture rates for the two cohorts are different (i.e., $S \neq 1$; a treatment effect or $p_{t2} \neq p_{c2}$).

Under H_A, the MSS for the treatment cohort is $\text{MSS}_t = m_t$ and the MSS for the control cohort is $\text{MSS}_c = m_c$. Under the null hypothesis the MSS is $\text{MSS}_0 = m_t + m_c$.

A test of H_0 can be derived from the residual distribution of the data, conditional on the MSS, given H_0 is true. Symbolically, the distribution is

$$\Pr_{H_0}\left\{\text{MSS}_t, \text{MSS}_c \mid \text{MSS}_0\right\} = \frac{\binom{R_t}{m_t}\binom{R_c}{m_c}}{\binom{R_t + R_c}{m_t + m_c}}.$$

This type of test was used in similar contexts by Robson and Youngs (unpublished report, 1971), Brownie and Robson (1976), and Pollock et al. (1985). Moreover, this type of test is known to be optimal from the general theory of hypothesis testing (Lehmann 1959). When we define

$$K_0 = \binom{R_t + R_c}{m_t + m_c},$$

the distribution is

$$\Pr_{H_0}\left\{\text{MSS}_t, \text{MSS}_c \mid \text{MSS}_0\right\} = \frac{K_t\, K_c}{K_0}.$$

The distribution is hypergeometric, allowing a 2×2 contingency table and a chi-square test with 1 df to be derived. The table is

m_t	$R_t - m_t$
m_c	$R_c - m_c$

Columns of the contingency table represent "recaptured" versus "not recaptured" for each of the two groups. The values for the example provide the following table

390	10,000 - 390
412	9,000 - 412

Heuristically, this test is comparing the two estimated proportions

$$\hat{Sp}_2 = \frac{390}{10{,}000}$$

and

$$\hat{p}_2 = \frac{412}{9{,}000}.$$

The above contingency table results in $\chi^2 = 5.38$, 1 df, and $P \leq 0.02$; thus, we conclude that the treatment has affected survival of the treatment group or that the capture rates of the treatment and control groups were unequal.

Many between-model tests are made in later chapters. Testing between two models represents a way to test a complex hypothesis. In general, such tests can be derived as likelihood ratio tests, in which one model is a special case (i.e., reduced number of parameters) of the other model. As an example, let

$L(\underline{\theta}_1)$ = a likelihood function with n parameters (e.g., $\phi_1, \phi_2, ..., \phi_6, p_2), ..., p_6$, where $n = 11$);

$L(\underline{\theta}_0)$ = a likelihood function with fewer parameters, $m < n$ (e.g., $\phi_1, ..., \phi_6, p$, where $m = 7$).

The likelihood $L(\underline{\theta}_0)$ corresponds to the null hypothesis that the parameter p is constant (i.e., $p_2 = p_3 = \cdots = p_6 = p$). The alternative hypothesis is that the parameter p varies. Both

hypotheses allow for variation in the parameter ϕ. A test of the null hypothesis is based on

$$\chi^2_{(n-m)} \doteq -2 \ln\left[\frac{L(\theta_0)}{L(\theta_1)}\right]$$

where both likelihoods are evaluated at their MLE values. The test is asymptotically chi-square distributed with $n - m$ df. A significantly large test statistic is taken as evidence that the null hypothesis is false (e.g., that the parameter p is not constant, as in the example above). Further information on likelihood ratio tests was given by Lehmann (1959) and Hogg and Craig (1970).

The contingency table procedure often results in a test that is equivalent to a likelihood ratio test, if no pooling is necessary because expected values are small. The contingency table approach outlined is preferable to the likelihood ratio test because data can be pooled easily if the expected values are small. Consideration of the partitioned test statistics is possible when data from more than two dams are available, thus a finer interpretation of the hypothesis test is allowed (however, the example used here cannot be partitioned). This general approach is used to derive most of the tests here.

In the case of a full-rank MSS for both the null hypothesis and alternative hypothesis models, the contingency table tests can often be found, and are then to be preferred. When the MSS are not full rank, one must usually rely on the likelihood ratio test procedure. The likelihood ratio test is an omnibus procedure justified by large-sample theory, we use it especially in the case of nonfull-rank models. However, for the most part, we were able to find the simpler contingency table tests of hypotheses that go with the full-rank models presented here.

1.2.2. Components of Variance

A conceptually difficult, but important, issue deals with components of variation in sampling and experimentation. A review of these concepts is presented here; more information can be found in White et al. (1982) and in many statistical texts. The two main classes of variation are population and sampling variation. For the moment, we concentrate on the meaning of these terms rather than on how they might be computed or estimated.

1.2.2.1. Spatial and temporal variation. – Variation in biological population parameters occurs commonly. A biological parameter (θ) is likely to vary over space and time (e.g., because of environmental factors), and may also vary among individuals. If the parameter is annual survival, it varies by species and age and perhaps by subpopulation. Survival probabilities may also vary among individuals from the same subpopulation, age, sex, size, and so forth. It is immaterial, for the moment, whether the value of the parameter over time, space, species, or subpopulation is known, but any population parameter may vary. In general, we denote this population variance as σ^2:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(\theta_i - \bar{\theta})^2$$

for N values of θ, θ_1, ..., θ_N corresponding to N different populations; no sampling variation is involved here. If N is large and a sample of n populations has been taken, then (for known θ_i) the estimator of σ^2 is

$$\hat{\sigma}^2 = \sum_{i=1}^{n}(\theta_i - \bar{\theta})^2 / (n - 1).$$

As an example of population variation, consider the number of bluegills *Lepomis macrochirus* in a small pond. A complete census on 10 June each year for 7 years provides the exact annual population size (by definition, this is a parameter). It would be unusual if the true number of fish was the same each of the 7 years. Thus, the population parameters vary. This population variation is conceptually measured as σ^2 and might, in this case, be termed temporal variation. The researcher or manager has no direct control over σ^2 (except, perhaps, by redefining the population itself or by some perturbation); it is a characteristic of the population.

Population variation might also occur spatially. For example, true population sizes are likely to differ among ponds of similar size and type. Spatial variation among population parameters is to be expected.

1.2.2.2. Sampling variation. – Sampling variation occurs because only partial information about the population normally is available. Exactly which members of a population fall into the sample is a result of a stochastic process, if the sampling process is unbiased toward particular members of that population. These processes are fundamentally unpredictable, as is the specific outcome of a flip of a penny.

Sampling variance is a measure of precision or repeatability of a result based on sample data; it is the measure of uncertainty. In general, sampling variance will be relatively small if each sample contains a large fraction of a population and relatively large if each sample contains few members of the population. The precision of results from a properly designed study can be estimated from information collected as part of that study.

Sampling variance of an estimator, $\hat{\theta}$, is denoted $\text{var}(\hat{\theta})$. Technically, one should write $\text{var}(\hat{\theta} \mid \theta)$ because this measure of variation is conditional on the true (unknown) value of θ. Generally, one has only an estimate of the variance $\hat{\text{var}}(\hat{\theta})$. Unlike the population variance, the experimenter has considerable control over the magnitude of the sampling variance by virtue of the study design. The most obvious way to decrease the uncertainty of the sampling process, and decrease the variance, is to increase the size of the sample or the proportion of the total population sampled. Other common ways to decrease the sampling variance include stratification of the population or use of a better estimation method. The standard error (se),

an alternative measure of an estimator's variability, is related to the variance by

$$\text{var}(\hat{\theta}) = [\text{se}(\hat{\theta})]^2 .$$

The terms "high precision" and "unbiasedness" represent indices of high accuracy. One can strive for this accuracy by careful attention in the design of surveys and experiments and in the use of good analytical methods. Consider a lake inhabited by lake trout *Salvelinus namaycush* susceptible to parasitism by sea lampreys *Petromyzon marinus*. An investigator wants to estimate the proportion (p) of these fish bearing at least one lamprey. An initial estimate, $\hat{p}_1 = 0.19$, is computed from the first sample (considered here as one unit of effort). Immediately, three additional samples are drawn and $\hat{p}_2 = 0.12$, $\hat{p}_3 = 0.18$, and $\hat{p}_4 = 0.27$. The variation among these estimates $\hat{p}_i$ is sampling variation, as the parameter p has not changed. In the example, the precision or repeatability for one unit of effort is only fair. However, if unit cost is low, sufficient sampling will lead to reliable results.

Finally, we consider an example where both population and sampling variation occur. Assume that the lake is surveyed once each year to estimate the proportion of lake trout bearing one or more sea lamprey. The data for 5 years are as follows:

Year	Unknown parameter	Estimate	Standard error
1	$p_1 = 0.13$	$\hat{p}_1 = 0.19$	$\hat{\text{se}}(\hat{p}_1) = 0.042$
2	$p_2 = 0.17$	$\hat{p}_2 = 0.12$	$\hat{\text{se}}(\hat{p}_2) = 0.034$
3	$p_3 = 0.16$	$\hat{p}_3 = 0.20$	$\hat{\text{se}}(\hat{p}_3) = 0.043$
4	$p_4 = 0.13$	$\hat{p}_4 = 0.18$	$\hat{\text{se}}(\hat{p}_4) = 0.035$
5	$p_5 = 0.16$	$\hat{p}_5 = 0.11$	$\hat{\text{se}}(\hat{p}_5) = 0.039$
$\overline{x}$	$\overline{p} = 0.15$	$\hat{\overline{p}} = 0.16$	$\hat{\text{se}}(\hat{\overline{p}}) = 0.017$

Here, $\hat{\text{se}}(\hat{\overline{p}}) = 0.017$ includes only the sampling variation. This $\hat{\text{se}}(\hat{\overline{p}})$ is the square root of the estimated theoretical $\hat{\text{var}}(\hat{\overline{p}})$;

$$\hat{\text{var}}(\hat{\overline{p}}) = \frac{1}{5^2} \sum_{i=1}^{k} \left[\hat{\text{se}}(\hat{p}_i)\right]^2 .$$

In contrast, if an average $(\hat{\overline{p}})$ is computed from these five estimates, the empirical variance of $\hat{\overline{p}}$ is

$$\hat{\text{var}}(\hat{\overline{p}}) = \frac{1}{5}\left[\frac{\sum_{i=1}^{5}(\hat{p}_i - \hat{\overline{p}})^2}{4}\right] = \frac{(0.041833)^2}{5} = (0.0187)^2 .$$

This $\hat{\text{var}}(\bar{p})$ includes both population (temporal) variation among the true p_i (σ^2) and conditional sampling variation of the $\hat{p}_i$ [var($\hat{p}_i$)]. Temporal population variation results from changes in the true proportion from year to year. Sampling variation occurs because only a sample of fish was examined, not the entire population. Often, one wishes to estimate σ^2, the variance among the true parameter values. This subject is discussed in Part 4. For now, the reader need only be aware of the distinction between these two types of variation.

1.3. Release-Recapture Protocols and Data

1.3.1. Introduction

The formal basis for the development of a statistical theory to underly survival experiments is the extensive literature on capture-recapture sampling (see Seber 1982, 1986 and Brownie et al. 1985 for recent reviews). This literature deals almost entirely with the estimation of population parameters (e.g., population size, survival rate, or number of births by time, sex, age-class, or geographic area) or the testing of various hypotheses concerning model assumptions. The theory presented herein extends capture-recapture methodology into survival experiments to assess the effect of a treatment on survival rate. The basis for such assessments is a control group of marked animals which enables the treatment-control comparisons that are standard experimental concepts of long standing.

We present extensive consideration of experiments with a treatment and a single control in Part 2. We consider extensions to multiple treatment and control groups in Part 3, which is more abbreviated, as it concentrates on presenting theory. Assume that a known number of fish in each of two groups is marked and released at dam 1. The first group is marked to denote that they are in the treatment group (t = *t*reatment), while the fish in the second group constitute the control group (c = *c*ontrol). The known numbers marked and released are denoted as R_{t1} and R_{c1} for treatment and control fish, respectively (R = *R*eleased), at dam 1. Throughout this monograph, we use a capital R to denote the number of fish released. We use an initial subscript to denote treatment or control as well as further subscripts to denote the specific release and recapture site (see Glossary).

The treatment may be the passage of the fish over a spillway, through a turbine or bypass system, or around a deflecting screen or barrier. Because survival of fish through various types of hydroelectric turbines is of concern, we use this as a primary example. Assume that a known number of marked fish (R_{t1}) is designated to receive the treatment and is released above dam 1 directly into a turbine intake. Simultaneously, a known number of control fish (R_{c1}) is released immediately below the dam near the end of the draft tube (see Figure 1.5).

As both groups move downstream, they are sampled at one or more downstream dams or other sampling sites. The only difference between the two groups is that some fish in the treatment group may have been killed while passing through the turbine and other parts of the dam structure.

Fish often are given only a batch mark, which is enough to allow their treatment or control status to be recognized when they are recaptured at sampling sites downstream. For example, all the treatment fish (R_{t1}) could be branded with a "T" and all the control fish (R_{c1}) with a "C" (or any other two marks that can be distinguished clearly).

Alternatively, fish may be marked with a unique tag or number. New technology may make this approach more feasible in the future (e.g., passive integrated transponder [PIT] tags). New types of tags are just starting to be evaluated (Prentice and Park 1984, 1985). We will assume here that tags are not lost and that marks remain readable.

A common example of unique marks is the individually numbered bird bands issued by the U.S. Fish and Wildlife Service. The use of unique marks has many advantages in that the

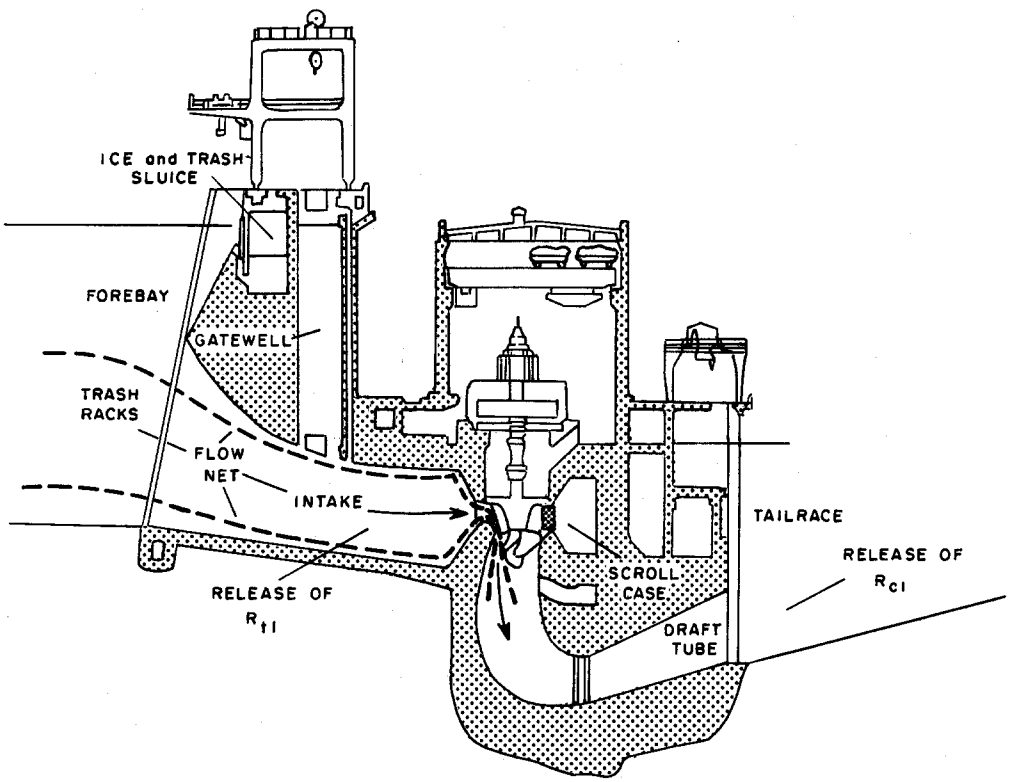

Figure 1.5. – Idealized diagram of a dam, some of its components, and points of release of the treatment and control fish. (Redrawn from Olson and Kaczynski, unpublished report, 1980.)

specific capture history and movements of each fish or other animal can be tabulated and analyzed. In the final assessment, the primary disadvantage is cost. These issues relating to uniqueness of marks will be covered in more detail in the material that follows.

The structure of the release-recapture data can be seen by considering only the control fish released below dam 1 (R_{c1}). If we consider only the first recapture for each fish (regardless of where this might occur), the recapture data have a multinomial sampling distribution if independent fates can be assumed. Thus, a fish released can be first recaptured at only one of the downstream dams (dams 2, ..., k), or it may never be recaptured. These outcomes are mutually exclusive and exhaustive.

Now consider the subset of the R_{c1} fish that are first recaptured at dam 2. We use m_{c12} to denote this number of fish; "m" is used because these recaptures are of marked fish. Three outcomes are possible:

(1) All m_{c12} fish survive recapture and handling and are rereleased at dam 2 as a subset called R_{c2}.

(2) Some fish are accidentally killed or injured or are removed deliberately; the remaining fish are rereleased and called R_{c2}.

(3) The number of m_{c12} fish recaptured that survive is added to a known number of new fish (to be initially released at dam 2) and the entire release is called R_{c2}.

The first two cases (1 and 2) are perhaps the most common in survival experiments; i.e., an initial release at dam 1 followed by the potential recapture and rerelease of the same fish at several downstream sites. In some studies (case 3), new fish are also released at the downstream capture sites along with the recaptured fish.

The term "losses on capture" is used in the literature (Jolly 1965) to describe fish killed accidentally or removed intentionally during recapture and handling. The important point is that the number of fish released or rereleased is known. In the above example, the releases at dam 2 (R_{c2}) could be

(1) less than the m_{c12} captures at dam 2 due to some losses on capture;

(2) equal to m_{c12} if no losses on capture occur;

(3) larger than m_{c12} due to the release of new fish along with "old" fish already in the experiment; or

(4) zero, because all fish were intentionally removed or accidentally killed.

In case 4 (above), only data on first captures are available. This situation is relatively simple and is examined in Chapter 2.2.

The sequence of possible recaptures of each fish leads us to consider the concept of a capture history for each fish used in the study. The capture history of a fish is a succinct way of tabulating the dams at which it was recaptured and possibly rereleased. Capture histories are denoted as a series of ones (= captured or recaptured) and zeros (not captured or recaptured). For example, the string of six values, {100101}, represents the capture history of a fish initially released at dam 1 and recaptured at dams 4 and 6. In general, if there are k release-recapture dams, the capture history consists of k-ordered ones and zeros. The ith digit represents what happened to the fish at the ith dam.

1.3.2. Capture Histories and Data Arrays

Statistical methods for the estimation of unknown parameters or the testing of hypotheses are based on the capture histories of marked fish. Practicalities aside, the most informative experiment is provided by an adequately replicated experiment involving the release of large samples of uniquely marked fish recaptured at a high rate at several downstream dams. Although this experiment may be the ideal, experimentation can be conducted under a host of other conditions. First, we must introduce several levels of data summarization.

1.3.2.1. CH matrix. – The capture history (CH) matrix provides specific capture histories (e.g., {110001}, or {101011}) as rows, along with the number of fish, by treatment group, having that capture history. Consider a particular row of a CH matrix as an example:

{1011101} 37 43,

which represents the results for a particular capture history over seven dams. The interpretation is that, of all treatment and control fish initially released at dam 1, 37 treatment and 43 control fish were recaptured and rereleased at dams 3, 4, 5, and 7 (and these fish were not captured at dams 2 or 6).

Because the CH matrix is a concise summary of the basic data it is important that the reader become familiar with it. A minus sign indicates fish lost on capture or deliberately removed. For example,

{1011101} -4 -3

is similar to the previous example, but signifies that there were four treatment and three control fish with this capture history that were removed at dam 7. Therefore, 37 and 43 fish were recaptured and rereleased alive in addition to four and three fish that were recaptured but removed (e.g., they died accidentally, were seriously injured, or were intentionally removed).

In general, we recommend the use of the CH matrix as a starting point in the analysis. Program RELEASE can compute useful summaries of the data from the CH matrix.

A detailed example of a CH matrix is shown in Table 1.1. We make extensive use of this general numerical example.

Table 1.1. – The capture history (CH) matrix for the example data set. The numbers of fish either recaptured and rereleased or recaptured and removed are shown for each capture history by treatment (t) and control (c) groups.

	Number recaptured	
Capture history	t	c
100000	25925	24605
100001	563	605
100001	-27	-36
100010	508	522
100010	-23	-25
100011	17	23
100011	-1	-1
100100	1500	1678
100100	-81	-57
100101	45	48
100101	-3	-1
100110	37	44
100110	-2	-2
100111	1	2
101000	193	207
101000	-14	-10
101001	5	9
101010	7	4
101100	16	14
101100	-1	-1
101101	1	1
101110	1	1
110000	872	935
110000	-29	-33
110001	26	28
110001	-1	-1
110010	16	18
110010	-1	-1
110100	67	68
110100	-3	-4
110101	1	2
110110	2	1
111000	10	12
111001	0	1
111100	1	0

These data are from a hypothetical study of turbine survival where fish are released at dam 1 and recaptured at five downstream dams. We assume that 30,000 treatment fish are released into a turbine intake of the dam (R_{t1} = 30,000) and 29,000 control fish are released simultaneously immediately below the dam (R_{c1} = 29,000); see Figure 1.5. These data are useful for illustrating analyses because they were generated from known parameter values that we pre-selected. We took the treatment survival rate of fish passing through the turbine and structure of the dam to be 0.9 (S = 0.9), which means there is a 0.9 probability of a treatment fish surviving from the mouth of the turbine intake to the point of release of the control fish just below the dam. This subject is discussed in Chapter 1.5.

We assume that each fish was given a unique tag or mark. Losses on capture were small, averaging about 3.5% of the fish recaptured at each of the five downstream dams. The survival rates (ϕ_i) of fish between dams and recapture rates (p_i) are shown in Figure 1.6.

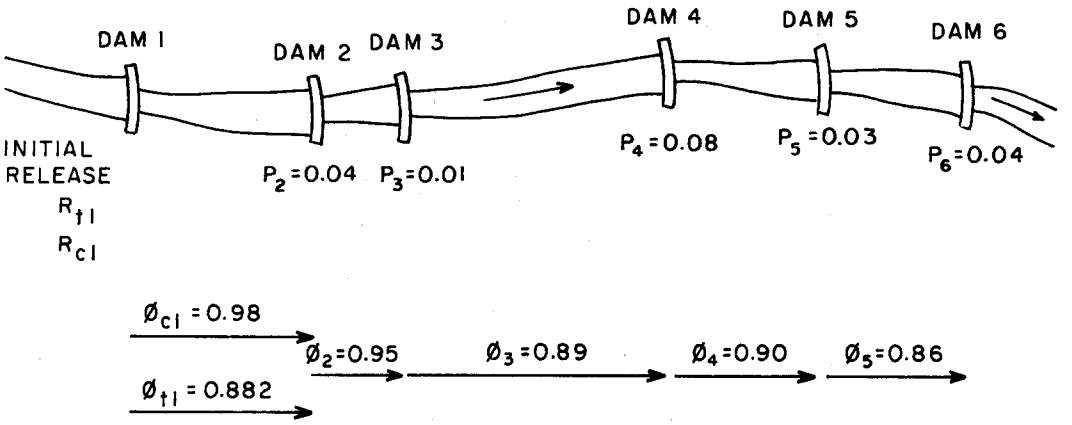

Figure 1.6. – Parameters used in the general numerical example. Because the recapture rates in this example are equal for treatment and control fish, the t or c subscripts are not used (i.e., $p_{c2} = p_{t2} = p_2$). Similarly, the dam-to-dam survival rates for treatment and control fish after the first survival rates are not subscripted for treatment group. The survival from dam 1 to dam 2 differs by treatment and control.

The release and recapture data are shown symbolically in Figure 1.7 and given numerically in Table 1.2. This example is simple but includes some ideal assumptions. These assumptions are relaxed in Part 2, but it is important to understand the formulation of the problem before various extensions and special cases are considered.

The recapture rates ($p_{t\,i}$ and $p_{c\,i}$) represent the probability of a fish being recaptured at dam i, given that the live fish reaches the ith dam. The recapture rates used in the example are small, averaging only about 4%. Although this recapture rate is typical for many studies that have been conducted on the Columbia River, it would be better if these rates were higher because precision and test power would be improved.

The survival rates $\phi_{t\,i}$ and $\phi_{c\,i}$ represent the probability of a fish in one of these two groups surviving from dam i to dam $i + 1$. We chose these parameters to be fairly high (0.86-0.98 for control fish); however, the total survival between dams 1 and 6 for control fish is 0.641 (the product of ϕ_{c1}, ϕ_{c2}, ϕ_{c3}, ϕ_{c4}, and ϕ_{c5}: $\prod_{i=1}^{5}\phi_{ci}$). The parameter representing the treatment effect is the survival rate S. In this simple example, $S = \phi_{t1}/\phi_{c1} = 0.90$. The estimation of the treatment effect S under different models and sampling protocols represents the focus of this monograph.

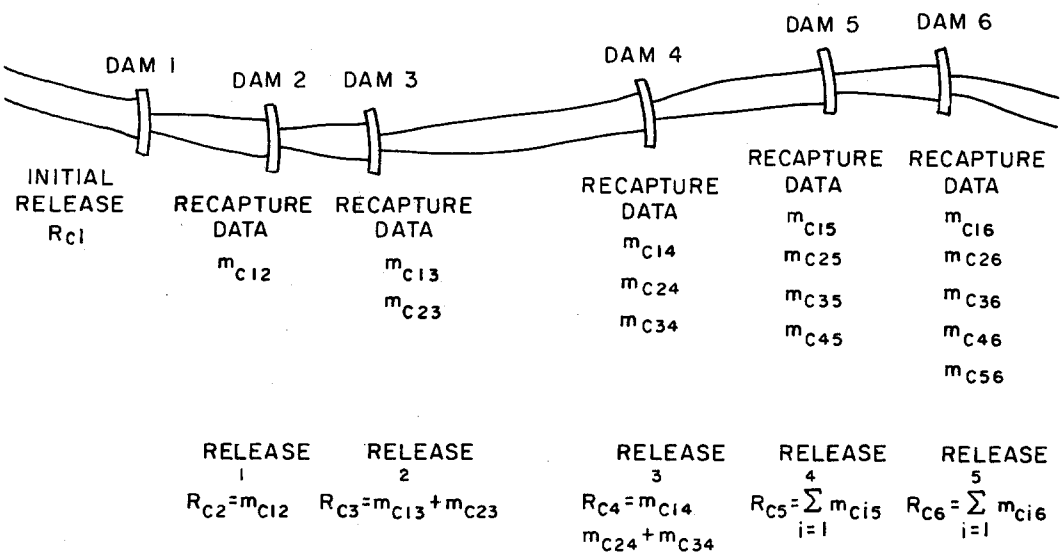

Figure 1.7. – Notation for the recapture and release or rerelease data for the control group in the general numerical example. In the figure, no losses on capture are assumed. If such losses occur, they are not included in the total to be rereleased (R_{c1}).

Table 1.2. – Survival and recapture probabilities for the general numerical example. Treatment and control fish are assumed to have been released at dam 1 and potentially recaptured and rereleased at dams 2-6.

	Recapture probabilities	Survival probabilities[a]		
Dam i	$p_{t1} = p_{c1}$	ϕ_{c1}		ϕ_{t1}
1		0.98	$\neq$	0.88
2	0.04	0.95	=	0.95
3	0.01	0.89	=	0.89
4	0.08	0.90	=	0.90
5	0.03	0.86	=	0.86
6	0.04			

[a] $S = \phi_{t1}/\phi_{c1} = 0.9$.

Examination of Table 1.1 shows that, of the initial releases, 25,925 treatment and 24,605 control fish were never recaptured at any of the five downstream dams (see row one of the CH matrix); 563 treatment and 605 control fish were recaptured and rereleased only at dam 6 (capture history {100001} in row two), and an additional 27 treatment and 36 control fish were recaptured only at dam 6 and lost on capture. The fish lost on capture are denoted with a minus sign because they were not rereleased. The interpretation of the rest of the CH matrix is similar.

1.3.2.2. Full m-array. – The CH matrix is a compact way to present the basic data from any survival experiment in which marked animals are released and recaptured. All estimation methods and tests can be based on the information contained in the CH matrix. However, the CH matrix can be summarized in what we call the full m-array as an equivalent representation of the data. The full m-array contains every iota of information from the experiment, but is more easily interpreted than the CH matrix because it is directly related in format to our statistical modeling of release-recapture data and to the computational procedures for some hypothesis tests.

For each treatment or control group in the study, there is a separate full m-array. The recapture data are represented as m_{vij}, the number of first recaptures, on occasion j from releases on occasion i ($i < j$), by group (v). The data are represented in a full m-array in terms of releases and first recaptures after release. Moreover, the releases R_{vi} and recaptures m_{vij} are partitioned by capture history h (at release time i) into all possible subcohorts. Thus, an example of one line of this full m-array is of the form

$$h = \{101\} \quad R_{t3h}, m_{t34h}, m_{t35h}, ..., m_{t3kh}.$$

We find that it is useful to append the values r_{vih} and $R_{vih} - r_{vih}$ (total fish ever recaptured from those released and those never seen again) onto the above line.

There will be $k - 1$ major portions (row groupings) of the array ($k - 1$ releases; releases at dam k are irrelevant). Within the ith portion (releases at dam i), the number of subcohorts is variable, as it depends on the number of previous capture occasions. Table 1.3 shows the full m-array for the treatment group data of the numerical example, e.g., for $h = \{101\}$, $R_{t3h} = 224$. The corresponding recaptures at dams 4, 5, and 6 are $m_{t34h} = 19$, $m_{t35h} = 7$, and $m_{t36h} = 5$, respectively. A fish recaptured n times will be represented in n different rows of the full m-array.

A detailed example will make the full m-array more clear. Consider the recapture data on treatment fish at dam 3 in Table 1.3, especially the shaded area. Of the 30,000 fish initially released at dam 1, 238 were first recaptured at dam 3 (their capture history is {101}). Of the 238 recaptured fish, 14 were lost on capture leaving 224 fish available to be rereleased. In addition, 1,029 fish were captured at dam 2 and, of these, 1,000 were released (29 were lost on capture). Of these 1,000 fish released at dam 2, 11 were recaptured at dam 3 and all were rereleased (no fish were lost on capture). These 11 fish have capture history {111}, as they were released at dam 1, captured and rereleased at dam 2, and then captured and rereleased at dam 3. Therefore, 235 fish were released at dam 3 (224 plus 11 = 235). In order to retain all the information in the CH matrix, the full m-array must also present losses on capture (shown in parentheses). The reader is encouraged to work through the meaning of the full m-array, as it is used frequently in material that follows.

Program RELEASE computes the full m-array as an option, if $k \leq 9$, from the CH matrix. Figure 1.6 presents material that aids in the interpretation of Table 1.4 for the control group.

1.3.2.3. Reduced m-array. – A summarization of the data from a survival experiment is the reduced m-array, which combines data over subcohorts within cohorts. The reduced m-array allows the biologist to view critical summary data in a simple format. Capture-recapture data are usually published in the literature in what we call the reduced m-array (see, for example, the data presented by Jolly 1965).

The reduced m-array contains all the information needed for estimation of the ϕ_i and p_i parameters under the usual Jolly-Seber assumptions, but lacks some important information required for components of the full testing procedure. Nonetheless, the reduced m-array is a valuable summary of the data, and extensive use is made of such summaries in this monograph. Program RELEASE allows the user to input the data as either a CH matrix or a reduced m-array.

Table 1.3. – Release-recapture data generated for the treatment group used in the general numerical example. The full *m*-array is given for the complete capture history protocol. Losses on capture are in parentheses and capture histories, up to and including dam *i*, are shown in braces (e.g., {1011} represents a fish initially released at dam 1 and recaptured at dams 3 and 4). Information on the number of fish released or rereleased appears to the left of the line and the number of first recaptures to the right.

	Release-recapture data (R_{tih} and m_{tijh})							
Release	1	2	3	4	5	6	Total recaptured r_{tih}	Never recaptured $R_{tih} - r_{tih}$
Initial release {1}	30,000	1,029(29)	238(14)[a]	1,669(81)	549(23)	590(27)	4,075	25,925
Rereleases at dam 2	{11}	1,000	11(0)	73(3)	17(1)	27(1)	128	872
Rereleases at dam 3		{101}	224	19(1)	7(0)	5(0)	31	193
		{111}	11	1(0)	0(0)	0(0)	1	10
Rereleases at dam 4			{1001}	1,588	40(2)	48(3)	88	1,500
			{1101}	70	2(0)	1(0)	3	67
			{1011}	18	1(0)	1(0)	2	16
			{1111}	1	0(0)	0(0)	0	1
Rereleases at dam 5				{10001}	526	18(1)	18	508
				{11001}	16	0(0)	0	16
				{10101}	7	0(0)	0	7
				{10011}	38	1(0)	1	37
				{11011}	2	0(0)	0	2
				{10111}	1	0(0)	0	1

[a]The shaded area represents the numbers captured, lost on capture, and rereleased. For example, 224 fish were released at dam 3 from the 238 caught; 14 were losses on capture.

Table 1.4. – Release-recapture data generated for the control group used in the general numerical example. The full m-array is given for the complete capture history protocol. Losses on capture are in parentheses and capture histories, up to and including dam i, are shown in braces. Information on the number of fish released or rereleased appears to the left of the line and the number of first recaptures is to the right.

	Release-recapture data ($R_{c\,ih}$ and $m_{c\,ijh}$)						Total recaptured	Never recaptured
Release	1	2	3	4	5	6	$r_{c\,ih}$	$R_{c\,ih} - r_{c\,ih}$
Initial release {1}	29,000	1,104(33)	247(10)	1,832(57)	571(25)	641(36)	4,395	24,605
Rereleases at dam 2	{11}	1,071	13(0)	75(4)	19(1)	29(1)	136	935
Rereleases at dam 3		{101}	237	17(1)	4(0)	9(0)	30	207
		{111}	13	0(0)	0(0)	1(0)	1	12
Rereleases at dam 4			{1001}	1775	48(2)	49(1)	97	1,678
			{1101}	71	1(0)	2(0)	3	68
			{1011}	16	1(0)	1(0)	2	14
Rereleases at dam 5				{10001}	546	24(1)	24	522
				{11001}	18	0(0)	0	18
				{10101}	4	0(0)	0	4
				{10011}	46	2(0)	2	44
				{11011}	1	0(0)	0	1
				{10111}	1	0(0)	0	1

Tables 1.5 and 1.6 show reduced m-arrays for treatment and control fish, respectively, for the general numerical example. Also shown in these tables are row totals r_i, column totals m_j, and a statistic z_j (z_j = the number of fish captured above and below, but not at, dam j). Of course, the specific notation includes t and c in the subscripts to denote group. The m_{vj}, z_{vj}, and r_{vi} for $v = t$ or c are not necessary aspects of the m-array; however, it is useful to show them there (they are redundant, given the R_{vi} and m_{vij} information).

Whereas the full m-array presents R_{vih} and m_{vijh}, the reduced m-array presents only the sums $R_{vi} \equiv R_{vi.} = \sum_h R_{vih}$ and $m_{vij} \equiv m_{vij.} = \sum_h m_{vijh}$ for $v = t$ or c. For example, in Table 1.5,

$$R_{t3} = 235 = 224 + 11 = R_{t3,\{101\}} + R_{t3,\{111\}},$$

and

$$m_{t34} = 19 + 1 = m_{t34,\{101\}} + m_{t34,\{111\}}.$$

Losses on capture may conveniently be shown in parentheses but are not necessary; we omit them from Tables 1.5-1.6.

1.3.3. Four Major Protocols

A technical team planning an experiment to estimate survival due to some treatment (e.g., passage over a spillway) must choose the basic experimental approach to be used in the field. We have identified four broad alternatives and have termed them "protocols." Each protocol is discussed in terms of the type of marking and recapture method and amount of information on the specific capture history of individual fish. We will make extensive use of the general numerical example introduced in Section 1.3.2 to aid in understanding these important concepts.

Table 1.5. – Reduced release-recapture data and summary statistics for the treatment group. Shown are the totals, by release occasion, over subcohorts. This table is the reduced m-array and is a condensation of the information given in Table 1.3.

Dam	Releases	Treatment recapture data at dam j, m_{tij}					
i	R_{ti}	$j = 2$	3	4	5	6	r_{ti}
1	30,000	1,029	238	1,669	549[a]	590	4,075
2	1,000		11	73	17	27	128
3	235			20	7	5	32
4	1,677				43	50	93
5	590					19	19
Totals	m_{tj}	1,029	249	1,762	616	691	4,347
	z_{tj}	3,046	2,925	1,195	672	0	

[a]The sum of elements in the shaded area is $m_{t5} = 616$.

Table 1.6. – Reduced release-recapture data and summary statistics for the control group. Shown are the totals, by release occasion, over subcohorts. This table is the reduced *m*-array and is a condensation of the information given in Table 1.4.

Dam	Releases	Control recapture data, m_{dj}					
i	R_d	$j = 2$	3	4	5	6	r_{c1}
1	29,000	1,104	247	1,832	571	641[a]	4,395
2	1,071		13	75	19	29	136
3	250			17	4	10	31
4	1,862				50	52	102
5	616					26	26
Totals	m_{cj}	1,104	260	1,924	644	758	4,690
	z_{cj}	3,291	3,167	1,274	732	0	

[a]The sum of elements in the shaded area is z_{c3} = 3,167.

1.3.3.1. First capture history protocol. – Under the first capture history protocol, marked fish are released at dam 1 and are removed upon first recapture. Removal can be physical, or another mark or fin clip can be added (and the fish then released) that allows future recaptures to be ignored. Thus, removal data are multinomial, as each fish can be recaptured independently only at a single downstream dam (i.e., at one of dams 2, 3, ..., k) or "never."

Only batch marks are required to distinguish the two releases (e.g., the R_{t1} and R_{c1} fish must have different batch marks). First capture history protocol data can be summarized as a CH matrix. Table 1.7 provides an example of a study involving six dams (k = 6).

Table 1.7. – Capture history matrix for the general numerical example under the first capture history protocol. Negative values indicate that fish were removed upon recapture.

Capture history	Dam of recapture j	Number recaptured	
		t	c
110000	2	-1,029	-1,104
101000	3	-238	-247
100100	4	-1,669	-1,832
100010	5	-549	-571
100001	6	-590	-641
100000	never	25,925	24,605

The CH matrix shows that most fish were released and never recaptured; see capture history $h = \{100000\}$. The minus sign preceding the numbers recaptured in the treatment and control groups denotes the fish that were removed. All fish recaptured were removed; none were rereleased. All the data representations (the CH matrix, full and reduced m-arrays) are essentially equivalent. Only data on first recaptures (that are then removed) are available for analysis. Fish recaptured at the last dam (6 in this example, k in general) need not be removed as no further sampling will be conducted downstream.

The reduced m-array is a more convenient summary of data under this simple protocol (Table 1.8). Note that these data include only the first row of Tables 1.3 and 1.5 for the treatment group and Tables 1.4 and 1.6 for the control group because fish are not rereleased after capture.

The first capture history protocol allows only the treatment survival rate S to be estimated as an individual parameter. (Products of the other parameters [i.e., the ϕ_i and p_i] are estimable.) Limited tests of important assumptions are possible. Overall, the first capture history protocol is simple and useful if the effect of the treatment is acute, and if proper replication is included in the experimental design (see Part 4 for a discussion of replication).

1.3.3.2. Unknown capture history protocol. – In most fish survival experiments conducted in the Columbia River to date, capture histories of individual fish are unknown. Fish are given a batch mark, but the fish are not removed upon recapture; rather they are rereleased. Therefore, a particular fish may be released and recaptured two or more times, but one has no way of knowing its capture history. The resulting recapture data are not multinomial, and information needed to complete the CH matrix is not available with this protocol. The unknown capture history protocol, consequently, has several disadvantages, especially as the recapture probabilities (p_i) increase. The data are not amenable to any exact statistical analysis. However, when capture probabilities are low, the unknown capture history protocol is a potential alternative if there is sufficient empirical replication.

Unknown capture history data include the total number of fish, by treatment and control group, recaptured at dam j (i.e., $m_{tj} \equiv m_{t.j}$ and $m_{cj} \equiv m_{c.j}$, $j = 2, ..., 6$ in the general numerical example). The data are a pooling of recaptures over different capture histories; thus, some important summary statistics cannot be computed (e.g., r_{v1}, the total number of distinct fish

Table 1.8. – Data from the general numerical example under the first capture history protocol for treatment and control groups. The m-array is reduced to a single row for each group because no recaptured fish are rereleased; instead they are removed.

		Number recaptured and removed at dam j, m_{ij}				
Group	Releases R_{i1}	$j = 2$	3	4	5	6
t	30,000	1,029	238	1,669	549	590
c	29,000	1,104	247	1,832	571	641

recaptured from the initial release). Table 1.9 shows the data that would result from the general numerical example if this protocol had been used. The recapture data shown for this protocol are merely column totals from Tables 1.5 and 1.6 for treatment and control fish, respectively. The availability only of totals represents a loss of information and prevents an exact statistical analysis.

1.3.3.3. Complete capture history protocol. – In the complete capture history protocol, each fish bears a unique mark. The use of unique marks allows the capture history of each fish to be known and used in the analysis. Fewer assumptions are required, additional statistical tests about assumptions can be made, and flexibility in the estimation of parameters is increased. The use of unique marks will become increasingly feasible as new technologies develop.

We believe that the complete capture history protocol is often superior to the other three protocols because it provides more tests of key assumptions and provides flexibility in the analysis. This protocol should be given full consideration in the design of future studies. The data derived from all other protocols are special cases of this general approach. The advantages of the complete capture history protocol increase as the recapture rates at downstream dams increase. The CH matrix for the complete capture history protocol is given in Table 1.1, including the fish lost on capture. The full m-arrays are shown in Tables 1.3 and 1.4, and the reduced m-arrays in Tables 1.5 and 1.6.

1.3.3.4. Partial capture history protocol. – Under the partial capture history protocol we consider two useful methods that have many advantages in the field and produce adequate data for statistical analysis. A limitation of this protocol is that the application of a second batch mark is required, and it is crucial that this handling not affect survival. The partial capture history protocol involves the use of site-specific marks at one or more downstream dams (i.e., 2, ..., k-1) in conjunction with "removal" after the maximum number of marks are applied. The use of site-specific marks differs from use of unique marks on fish initially released at dam 1. We develop only two possible partial capture history protocols: scheme A and scheme B.

Table 1.9. – Data from the general numerical example under the unknown capture history protocol for treatment and control groups. Capture histories are pooled and necessitate approximate analysis methods. Losses on capture are shown in parentheses.

		Number recaptured at dam j				
Group	Releases R_{v1}	j = 2	3	4	5	6
t	30,000	1,029(33)	249(14)	1,762(85)	616(26)	691(32)
c	29,000	1,104(33)	260(10)	1,924(62)	644(28)	758(39)

In scheme A, treatment and control fish are batch-marked and then released at dam 1, as in the first and unknown capture history protocols. However, upon first recapture, each fish is given a second mark that is specific to that particular dam. If a fish is recaptured a second time, it is removed (either physically removed or a fin is clipped or a third mark is added to indicate its removal). A possible capture history with scheme A is {1001010}, which indicates a fish marked initially at dam 1, recaptured (and released) at dam 4 after being given an additional mark specific to dam 4, and recaptured and removed at dam 6 (and therefore unavailable for recapture at dam 7). Thus, only partial capture histories are available under this protocol (data cannot be gathered on fish recaptured more than twice).

Data under scheme A contain much of the information available from data under the complete capture history protocol, especially if capture probabilities are low. Marking equipment must be available at each downstream dam, but simple batch marks can be used.

The CH matrix for the general numerical example under scheme A is shown in Table 1.10, including fish lost on capture. Note that, unlike the CH matrix for complete capture his-

Table 1.10. – CH matrix for the general numerical example under the partial capture history protocol, scheme A.

	Number recaptured	
Capture history	t	c
100000	25,925	24,605
100001	563	605
100001	-27	-36
100010	508	522
100010	-23	-25
100011	-18	-24
100100	1,500	1,678
100100	-81	-57
100101	-48	-49
100110	-40	-48
101000	193	207
101000	-14	-10
101001	-5	-9
101010	-7	-4
101100	-19	-17
110000	872	935
110000	-29	-33
110001	-27	-29
110010	-17	-19
110100	-73	-75
111000	-11	-13

histories, no fish under scheme A are recaptured more than twice (i.e., there are no more than three ones in a capture history h: a row of the CH matrix). The m-arrays are shown in Table 1.11. For scheme A, the full and reduced m-arrays are identical, i.e., all releases at a given dam have the same capture history to that point, so there are no subcohorts based on capture histories.

In scheme B, it is assumed that treatment and control fish with distinguishing batch marks are simultaneously released at dam 1. All fish recaptured at the second dam (dam 2) are given a second mark to indicate that they were recaptured and then rereleased. However, all fish recaptured at dams 3, 4, ..., k are removed from the population. This scheme requires additional marking at only one downstream dam (dam 2). In this respect, scheme B is logistically better than scheme A, but the resulting data contain less information than those collected in scheme A. The CH matrix for scheme B is shown in Table 1.12, and the reduced m-arrays are shown in Table 1.13. The data are mostly removals and, therefore, similar to the data from the first capture history protocol (see Table 1.8).

Although some information is lost, we believe scheme B represents an excellent protocol that should be considered further in the design of future experiments. A disadvantage is the potential effect of handling and marking on subsequent survival. Scheme B represents a logistically reasonable protocol and scheme A is a statistically reasonable protocol.

Table 1.11. – The m-array for the general numerical example under the partial capture history protocol, scheme A.

			Number recaptured at dam j, m_{ij}				
Group	i	Releases R_i	$j = 2$	3	4	5	6
t	1	30,000	1,029	238	1,669	549	590
	2	1,000		11	73	17	27
	3	224			19	7	5
	4	1,588				40	48
	5	526					18
c	1	29,000	1,104	247	1,832	571	641
	2	2,071		13	75	19	29
	3	237			17	4	9
	4	1,775				48	49
	5	546					24

Table 1.12. – The CH matrix for the general numerical example under the partial capture history protocol, scheme B.

	Number recaptured	
Capture history	t	c
100000	25,925	24,605
100001	-590	-641
100010	-549	-571
100100	-1,669	-1,832
101000	-238	-247
110000	872	935
110000	-29	-33
110001	-27	-29
110010	-17	-19
110100	-73	-75
111000	-11	-13

Table 1.13. – The m-array for the general numerical example under the partial capture history protocol, scheme B.

			Number recaptured at dam j, m_{ij}				
		Releases					
Group	i	R_i	$j = 2$	3	4	5	6
t	1	30,000	1,029	238	1,669	549	590
	2	1,000		11	73	17	27
c	1	29,000	1,104	247	1,832	571	641
	2	1,071		13	75	19	29

1.4. Release-Recapture Modeling Concepts, Notation, and Assumptions

In Chapter 1.3 we introduced some of the concepts regarding release-recapture, including ways to display the data and different study protocols (which produce different amounts of data). Here we elaborate on the subject by considering concepts essential to the modeling of such data. We also present an overview of the necessary notation required to represent symbolically the data and probability models for the data. Finally, we introduce the philosophy that guides our data modeling and analysis efforts for this class of fish survival experiments. We assume that the complete capture history protocol is the starting point because it is the most general data collection and modeling case; all other protocols should be viewed as special cases of the complete capture history protocol.

1.4.1. Introduction to Release-Recapture Concepts

Release-recapture is used widely to study animal survival processes. In fisheries, it is used more often to estimate population sizes. A cohort of R_1 marked animals is released and then a subsequent sampling process is used to catch (sample) the marked survivors. These survivors may be rereleased at the site (or time) of capture. Consequently, in typical capture-recapture studies (e.g., Jolly 1965), an individual can be captured at several sites. Such multiple recaptures lead to the idea of a capture history, which we introduced in Section 1.3.2. Although capture histories provide a convenient way to record data and enter it for computer analysis, they are not convenient for modeling though they have been so used. Capture histories have been used as the basis of models and subsequent data analysis in the log-linear approach (e.g., Cormack 1979, 1981). Crosbie and Manly (1985) also provided an analysis method based on the capture history representation. We base our models on an alternative conceptualization of the process, as presented by Brownie et al. (1985:170-175).

The first key concept is that one should model the recapture process and then analyze the recapture data, conditional on the known number of releases at each release site (or time). A probability model for the data may then depend on capture history at the time of release. In principle, the released fish at any recapture dam can represent a mix of "new" and recaptured fish. The critical question is whether or not such a mixed cohort meets the assumptions needed for a meaningful data analysis.

The second key concept is that release and recapture are paired; this concept is the essence of release-recapture. Each release of a fish is an experiment in itself. Assume a fish is released at dam 1 and recaptured at dam 3 (but not seen at dam 2). One now knows that the fish survived between dams 1 and 3. The rerelease of the fish at dam 3 starts another survival trial. One conditions (principle 1) on that release and again waits to see if that fish is recaptured.

Consider the capture history $h = \{101011\}$ for a study with six dams. An equivalent representation for this capture history follows.

Release occasion i	First recapture after release time i				
	2	3	4	5	6
1 (released)	0	1			
2					
3 (released)			0	1	
4					
5 (released)					1

This type of representation of data leads to the m-array.

The second concept states that the model is concerned *only* with the *first* recapture after any release. The data are represented as a series of linked release-recaptures. Conditioning on release and then modeling first-only recaptures are the keys to simplified modeling of release-recapture data. Throughout this monograph, recapture (or capture) of marked fish

refers to the first recapture after a release. In this manner, releases and subsequent recaptures are uniquely paired (unless the fish is never observed again). Under the unknown capture history protocol, this pairing still exists in principle, but information about it is not available.

The recognition of different levels of data leads to additional terminology. The "cohort" is the focus of building either single or multiple release-recapture data sets. A cohort is a known number of animals released at a given site (or time). Given such a definition of a cohort, then, each animal in a cohort is either recaptured once or never observed again. Upon (first) recapture, if the animal is (re)released, it automatically becomes a member of a different study cohort.

The subcohort is a partition of a cohort. Suppose 500 fish are recaptured at dam 3 and 480 are released as cohort 3, $R_3 = 480$ (20 were lost on capture). Those fish can have one of two possible capture histories when they are released at dam 3:

h	R_{3h}
101	460
111	20
Total	480 .

The numbers of fish released with each capture history are defined as a subcohort of R_3. Therefore, in this example, cohort 3 has two subcohorts of sizes 460 and 20.

It is possible to define subcohorts on another basis, such as a fish's sex or size. The reason for distinguishing subcohorts within a cohort is that the subcohort data are useful for tests of assumptions. One can test that the capture and survival rates are not affected by the factors (especially capture history) defining subcohorts.

A data set is a collection of cohorts (releases) and the subsequent (first) recapture data from each cohort. Much of the capture-recapture literature deals with the analysis of only one data set. However, many important questions that can be investigated by release-recapture require collecting at least two related data sets (e.g., treatment-control, male-female, age-classes, or different locations; see Manly 1985 for numerous specific examples). Thus, one must be able to cope with the analyses of multiple, related release-recapture data consisting of the following levels:

(1) V release-recapture data sets (groups), e.g., $V = 2$ for t and c;

(2) $k - 1$ cohorts within each data set (for k dams);

(3) subcohorts within each cohort;

(4) first recaptures from each subcohort.

(Again, we note that some of these features vanish under protocols other than the complete capture history protocol.)

One final point regarding concepts: a capture history compiled at time of release is defined only with respect to the previous capture sites. Confusion can be avoided and simpler notation can be used when this is understood. Thus, if h represents a capture history at site i, then h has exactly i components, each component being either a zero or a one.

1.4.2. Release-Recapture Notation

We have already introduced most of our notation. We now present it in greater detail. Its comprehension should be facilitated if the previous ideas are kept in mind.

1.4.2.1. Notation for data. – Notation for data must allow one to distinguish levels of data: data sets (groups), cohorts, subcohorts, and first recaptures given releases. The following symbols are fundamental (see also the Glossary for notation):

k The number of release-recapture sites or times;

R_i Known number released at dam i, $i = 1, ..., k - 1$;

m_{ij} The number of fish recaptured (for the first time) at site j from the cohort released at site $i, j = i + 1, ..., k$.

These basic symbols must be elaborated upon to allow for subcohorts and multiple data groups. Elaboration is in the form of additional subscripts – v for data groups and h for subcohorts:

R_{vih} is the number of released animals in group v and subcohort h at site i.

m_{vijh} is the corresponding number of recaptures at site j for R_{vih}.

When only one data set is involved, the subscript v can be dropped. Thus, R_{ih} and m_{ijh} can arise. When results are pooled over all subcohorts, the h is dropped and R_{vi} and m_{vij} are used. We sometimes replace a subscript with a dot (.) to denote summation (i.e., "pooling") over that subscript. Thus, $R_{vi.}$ is equivalent to R_{vi}, although we prefer the simpler notation of R_{vi} in this situation.

The subscripts always appear in the same order: data group v, release site i, recapture site j, and subcohort h. However, all four subscript levels do not always occur (e.g., R never has subscript j). Various subscripts tend especially to be omitted in summary statistics.

In addition to R_i and m_{ij} (or R_{vih} and m_{vijh}), several summary statistics (i.e., of the m_{ij}) are commonly used. These statistics are sums of the recaptures (i.e., of the m_{ij}).

r_i The total number of the R_i that are recaptured again; a row total, $r_i = \sum_j m_{ij}$.

m_j The total number of marked animals caught at site j; a column total, $m_j = \sum_i m_{ij}$.

T_j The total number of captures at sites $j, j+1, ..., k$ from releases in cohorts $R_1, ..., R_{j-1}$ (hence released prior to site j).

z_j The total number of captures at sites $j+1, ..., k$ from releases in cohorts $R_1, ..., R_{j-1}$.

Many variations of these summary statistics can occur. In particular, these are r_{tih}, r_{cih} (in the case of treatment and control groups and recaptures by subcohort), and r_{ih} (one group only). The quantities m_{vj}, T_{vj}, and z_{vj} also arise, along with $m_{\cdot j}$, $T_{\cdot j}$, and $z_{\cdot j}$ (where summation is over groups). Because there are too many combinations to define explicitly, we imposed a logic on our notation in regard to the order and meaning of subscripts.

Summary statistics can be computed as various totals of the basic data: m_{vijh}; $v = 1, ..., V$; $i = 1, ..., k-1$; $j = i+1, ..., k$; and $h = 1, ..., H_{vi}$. The symbol H_{vi} represents the number of subcohorts in cohort i, data group v. For example,

$$\begin{cases} z_j = \sum_{i=1}^{j-1} \sum_{n=j+1}^{k} m_{in}, & j = 2, ..., k-1; \\ z_k \equiv 0; \end{cases}$$

$$T_j = m_j + z_j;$$

$$T_{j+1} = z_j + r_j.$$

Although there are many other relationships, the reader need not learn them. Program RELEASE computes the necessary summary statistics.

We selected our basic notation to be as consistent as possible with that used in the general capture-recapture literature, e.g., that used by Seber (1982). Our results are applicable to the analysis of Jolly-Seber data with respect to inferences about survival rates. The full analysis of Jolly-Seber data involves estimation of population size, which requires capturing unmarked animals and marking and releasing them. The extra notation needed is primarily u_i, the number of unmarked animals caught at site (or time) i; the notation u_{vi}, and possibly u_{vih}, could also be used. We do not use "u" in our notation, thereby making it available for use in extensions of results to the general open population capture-recapture situation (see, e.g., Chapter 8.3).

Table 1.14. – Symbolic form of the full m-array representation of a single release-recapture data set (see Table 1.3 for an example).

	Release-recapture data at dam j						
	1	2	3	4	5	Total recaptured	Never recaptured
$h = \{1\}$	R_{1h}	m_{12h}	m_{13h}	m_{14h}	m_{15h}	r_{1h}	$R_{1h} - r_{1h}$
	$h = \{11\}$	R_{2h}	m_{23h}	m_{24h}	m_{25h}	r_{2h}	$R_{2h} - r_{2h}$
		$h = \{111\}$	R_{3h}	m_{34h}	m_{35h}	r_{3h}	$R_{3h} - r_{3h}$
		$h = \{101\}$	R_{3h}	m_{34h}	m_{35h}	r_{3h}	$R_{3h} - r_{3h}$
			$h = \{1111\}$	R_{4h}	m_{45h}	r_{4h}	$R_{4h} - r_{4h}$
			$h = \{1011\}$	R_{4h}	m_{45h}	r_{4h}	$R_{4h} - r_{4h}$
			$h = \{1101\}$	R_{4h}	m_{45h}	r_{4h}	$R_{4h} - r_{4h}$
			$h = \{1001\}$	R_{4h}	m_{45h}	r_{4h}	$R_{4h} - r_{4h}$

Tabular forms for presenting data are given in Section 1.3.2. These forms (e.g., Tables 1.3 and 1.5) are directly related to our notation. There is a full m-array for each treatment group. This full m-array presents the data at the level of subcohorts, i.e., the R_{vih} and subsequent m_{vijh}. This form is shown in Table 1.14. Pooling over subcohorts within each cohort gives the data in the reduced m-array, which we usually merely refer to as the m-array (see Table 1.15). That data representation is the most common one used in capture-recapture studies; however, its use results in a loss of all the information (for testing assumptions) that is contained in the subcohort data.

Table 1.15. – Symbolic form of the reduced m-array representation of a single release-recapture data set (see Table 1.5 for an example).

Releases at dam i	Recapture data at dam j				Total recaptured
	2	3	4	5	
R_1	m_{12}	m_{13}	m_{14}	m_{15}	r_1
R_2		m_{23}	m_{24}	m_{25}	r_2
R_3			m_{34}	m_{35}	r_3
R_4				m_{45}	r_4
Summary statistics	m_2	m_3	m_4	m_5	
	z_2	z_3	z_4	z_5	

1.4.2.2. Notation for parameters. – Only two types of parameters are used in release-recapture models: survival rates and capture probabilities. The generic symbols for dam-to-dam survival rate and capture probability at a given dam are ϕ and p, respectively. Also, $1 - \phi$ represents mortality rate and $q = 1 - p$ = probability an animal is not captured. Precise definitions follow:

ϕ_i the conditional probability of a fish surviving from release at site i to site $i + 1$; and

p_i the conditional probability of a fish being captured at site i given that it is alive at (i.e., arrives alive at) site i.

Note the conditional nature of both parameters. Survival, ϕ_i, applies only to fish alive at site i. Also, ϕ_i is unrelated to whether or not the fish was captured at sites i or $i + 1$. Hence, survival and capture processes are evaluated as separate parameters.

The subscripts on these parameters can be expanded in accordance with conventions discussed previously. For example, we also have

ϕ_{vi}, p_{vi} when survival and capture rates vary by data group (e.g., $v = t$ or c).

ϕ_i represents a common survival rate from site i to $i + 1$ for all treatment groups (i.e., $\phi_{vi} = \phi_i$, for all v).

Similarly, we use p_i when the parameters p_{vi} do not vary by treatment v. The use of ϕ_i or p_i is equivalent to $\phi_{\cdot i}$ or $p_{\cdot i}$, except that the latter notation is used only rarely.

Survival and capture probability parameters often enter the models as complicated functions. Therefore, we define other parameters as functions of ϕ_i and p_i; in particular,

$$\lambda_i = \phi_i(p_{i+1} + q_{i+1}\lambda_{i+1}), \quad i = 1, ..., k - 1,$$

$$\lambda_k = 0 \text{ (by definition)},$$

and

$$\tau_i = \frac{p_i}{p_i + q_i\lambda_i}, \quad i = 2, ..., k - 1.$$

We note that λ_i = the probability that a fish released at dam i will be recaptured; thus, $E(r_i \mid R_i) = R_i\lambda_i$. Also, τ_i is the proportion of fish captured at dam i of those released prior to dam i and captured at dams i, $i + 1$, ..., k. Thus, $\tau_i = E\left(\frac{m_i}{T_i}\right)$.

The focus of this monograph is on the estimation of treatment effect S. Often, $S = \phi_{t1}/\phi_{c1}$ and, if the treatment has a detrimental effect on survival, $S < 1$ and can be considered as a probability. In other cases the treatment may enhance survival and, therefore, $S > 1$. If the effect of the treatment extends downstream to dam 3, then $S = (\phi_{t1}\phi_{t2})/(\phi_{c1}\phi_{c2})$.

The treatment effect can be partitioned under some sampling protocols. For example, $S_1 = \phi_{t1}/\phi_{c1}$ and $S_2 = \phi_{t2}/\phi_{c2}$; thus, the overall treatment effect is $S = S_1 S_2$. Other specific definitions of S are possible depending on the application and postulated effect of the treatment on the marked population. The important point is that S is a general measure of a treatment effect.

1.4.3. Release-Recapture Models

1.4.3.1. Modeling approach. – There are two conceptual aspects to the models used here: (1) the structure of the expected number of recaptures given the known releases, and (2) the specification of the nature of the random variation of the recaptures given their expected values. The structural aspects of the expectations are usually readily comprehended; however, a complete model is necessary for purposes of sound scientific inference. The complete model is a specification of the probability distribution of the recaptures m_{ij} given releases R_i.

Numerous modeling approaches have been used as a basis for the analysis of capture data (e.g., Cormack 1979). Conditional on the releases R_i and, given some specific assumptions, the recaptures $m_{i,i+1}$, ..., m_{ik} are multinomial random variables. Multinomial models provide the most convenient approach when the emphasis is on estimating survival rates. Also, multinomial models provide a relatively easy and unified framework for theory development (e.g., Brownie et al. 1985:170-175; Burnham, unpublished report, 1987).

The most critical aspect of modeling release-recapture data is the specification of the expected values of the recaptures m_{ij} given the releases. Symbolically, this expectation is

$$E(m_{ij} \mid R_i) = R_i \pi_{ij}.$$

Thus, $\pi_{ij} = E(m_{ij} \mid R_i)/R_i$, $j=i+1, ..., k$ is the probability that a fish released at site i will be recaptured at site j. By virtue of our definition of a recapture as meaning the first recapture after release, a fish released at site i is either recaptured at exactly one downstream site (site $i+1$, or $i+2$, ..., or k) or it is never observed again. We let $\lambda_i = \pi_{i,i+1} + \pi_{i,i+2} + ... + \pi_{ik}$; then $1 - \lambda_i = \Pr\{$a fish released at site i is never observed again$\}$. Thus, after release at site i, the fish experiences exactly one of $k - i + 1$ mutually exclusive fates. Given the assumptions of statistical independence for the fates of released fish and that the same parameters π_{ij} apply to every fish, then the m_{ij} given R_i are multinomial random variables. The general mathematical form of the multinomial probability distribution for these recapture data is

$$\Pr\{m_{i,i+1}, ..., m_{ik} \mid R_i\} = \binom{R_i}{m_{i,i+1} \cdots m_{ik}\ R_i - r_i} \left[\prod_{j=i+1}^{k} (\pi_{ij})^{m_{ij}}\right] (1 - \lambda_i)^{R_i - r_i}.$$

Multinomial models are completely specified by giving (hypothesizing) the expected values of the m_{ij} given the releases R_i. Thus, by adopting this framework, we reduce the modeling to specification of the π_{ij}.

1.4.3.2. Model structures. – By model structure, we mean the expressions for the π_{ij} in terms of survival (ϕ) and capture (p) probabilities. The estimators of the survival and capture rates are determined by the model structure. The multinomial sampling distribution component of the models really only determines (theoretical) sampling variances and covariances. If sampling variances are obtained from replicate releases, the only critical part of the model is the structure assumed for the $\pi_{ij} = E(m_{ij} \mid R_i)/R_i$.

A convenient initial model structure to consider is that of assuming parameters to be time-specific only. Then, for example,

$$\pi_{i,i+1} = (\phi_i p_{i+1}) ;$$

$$\pi_{i,i+2} = (\phi_i q_{i+1})(\phi_{i+1} p_{i+2}) ;$$

$$\pi_{i,i+3} = (\phi_i q_{i+1})(\phi_{i+1} q_{i+2})(\phi_{i+2} p_{i+3}) .$$

Consider the interpretation of $\pi_{i,i+2}$, which is the probability that a fish released at site i will not be caught at site $i + 1$ but will be caught at site $i + 2$. The probability of survival from site i to $i + 1$ is ϕ_i, while $q_{i+1} = 1 - p_{i+1}$ is the probability that the fish is not caught at site $i + 1$ given that it survives to site $i + 1$. Next, the fish must survive from site $i + 1$ to $i + 2$ (probability = ϕ_{i+1}) and then be caught (probability = p_{i+2}). It is worth noting that presence in the released cohort is represented by the product $(\phi_i q_{i+1})$, whereas removal from the released cohort is represented by the product $(\phi_i p_{i+1})$. All the release-recapture models used here have this basic structure: the π_{ij} are products of $(\phi_n q_{n+1})$ terms, $n = i, ..., j - 2$, and a final $(\phi_{j-1} p_j)$ term. What distinguishes different models is how these survival and capture probability parameters depend on treatment and release or recapture site.

Table 1.15 shows the symbolic form of the reduced m-array for a single release-recapture data set. Any model for such data can be represented by an analogous table giving the $E(m_{ij} \mid R_i) = R_i \pi_{ij}$ (or giving just the π_{ij}). For example, Table 1.16 gives the basic model structure used as our starting point. This model is essentially the Jolly-Seber model (Jolly 1965; Seber 1965). The parentheses enclosing pairs of ϕq or ϕp are shown only to emphasize the way survival from site i to $i + 1$ and being captured or not captured at site $i + 1$ always appear together in the model structure. Seber (1982) used the symbols α_i and β_i to denote the products $\phi_i q_{i+1}$ and $\phi_i p_{i+1}$, respectively.

Table 1.16. – The Jolly-Seber (time-specific parameters) model structure for the symbolic data of Table 1.15.

Releases at dam i	Expected number of recaptures, $E(m_{ij} \mid R_i)$, at dam j			
	$j = 2$	3	4	5
R_1	$R_1(\phi_1 p_2)$	$R_1(\phi_1 q_2)(\phi_2 p_3)$	$R_1(\phi_1 q_2)(\phi_2 q_3)(\phi_3 p_4)$	$R_1(\phi_1 q_2)(\phi_2 q_3)(\phi_3 q_4)(\phi_4 p_5)$
R_2		$R_2(\phi_2 p_3)$	$R_2(\phi_2 q_3)(\phi_3 p_4)$	$R_2(\phi_2 q_3)(\phi_3 q_4)(\phi_4 p_5)$
R_3			$R_3(\phi_3 p_4)$	$R_3(\phi_3 q_4)(\phi_4 p_5)$
R_4				$R_4(\phi_4 p_5)$

1.4.4. Assumptions

The numerous assumptions involved in making inferences from release-recapture data vary in their importance and in terms of what the investigator can do to satisfy them. We next present the necessary assumptions by type of assumption, and order most-to-least important within type of assumption.

Assumptions 1-6 relate to study planning, field procedures, and generality of the desired inferences.

(1) The test fish used are representative of the population of fish about which one seeks mortality information.

(2) Test conditions are representative of the conditions of interest.

(3) Treatment and control fish are biologically identical prior to release at dam 1. A strong version of assumption 3 is that initial handling, marking, and holding do not affect survival rate.

(4) The numbers of fish released are exactly known.

(5) Marking (tagging) is accurate; there are no mark (tag) losses and no misread marks (tags).

(6) All releases and recaptures occur in brief time-intervals, and recaptured fish are released immediately.

Assumptions 7-8 relate to the stochastic component of the models.

(7) The fate of each individual fish, after any known release, is independent of the fate of any other fish.

(8) With multiple lots (or other replication), the data are statistically independent over lots.

Assumptions 9-12 relate to model structure.

(9) Statistical analyses of the data are based on the correct model.

(10) Treatment and control fish move downstream together.

(11) Captured fish that are rereleased have the same subsequent survival and capture rates as fish alive at that site which were not caught, i.e., capture and rerelease do not affect their subsequent survival or recapture.

(12) All fish (in the study) of an identifiable class (e.g., treatment or control, or size, or replicate) have the same survival and capture probabilities; this is an assumption of parameter homogeneity.

It is difficult to specify a set of assumptions that suffice to cover all the protocols and intended inferences presented here. In particular, stronger assumptions are required for estimation of the absolute survival rates ϕ than for the treatment effect $S = \phi_t/\phi_c$. Note particularly that a multiplicative bias that equally affects $\hat{\phi}_t$ and $\hat{\phi}_c$ has no effect on S. We next discuss the role of assumptions from this dual perspective.

No amount or sophistication of data analysis can salvage valid results from an invalid design. Also, statistical inferences cannot validly extend beyond the scope of the design. Assumptions 1 and 2 relate to this point; they are virtually self-evident, but worth bearing in mind. For example, if one wants to know something about a species of fish, then that species should be used as the test fish. Other factors to consider include genetic strain, size, general condition, and so forth. Test conditions (flow, turbine type, power settings, dam design) are also relevant. These types of studies are usually limited to one dam, and often one turbine, at a time. There is little or no random selection of test conditions, nor can there be. That does not affect models or analyses presented here. Assumptions 1 and 2 really just specify the limits of valid statistical inferences regarding fish survival rates and treatment effects. Inferences about salmon or conditions other than the test conditions must be justified on other than statistical grounds.

Assumptions 3, 4, and 5, and to some extent assumption 6, can be influenced by the investigator by the use of careful field procedures. Our reading of the fisheries literature shows that fisheries scientists are well aware of these assumptions; and that these scientists are able to do an excellent job of meeting assumptions 3-6, to the extent they can be met.

For the purposes of making inferences about treatment effect(s), a weak version of assumption 3 suffices: comparability of treatment and control fish when initially released. At this level assumption 3 is met, basically, by random assignment of fish to treatment groups and lots, and identical handling procedures for all treatment groups (often just treatment and control fish). Handling may affect the fish and their survival rates ϕ. If absolute survival rates are a study objective, a strengthened assumption 3 is needed: fish preparation and holding procedures do not affect survival rates. Potentially, assumptions 1 and 3 overlap. The common point being made is that S, or ϕ, is assumed to be the same for the released and wild fish. If the experimental fish are not representative a priori, this assumption fails.

It is critical to know the numbers of live, healthy fish released (assumption 4). Due to handling mortality or natural mortality, the release number R may be less than the number marked and placed in holding facilities. Accurate marking is also critical. Marks must not be lost or become unreadable. Assumption 5 points out that recapture data must be recorded accurately. This assumption is required to obtain unbiased estimates of absolute survival rates ϕ and capture rates p. However, $\hat{S}$ remains unbiased if the rates of tag loss (including, e.g., unreadable freeze-brands) are equal for treatment and control groups.

Assumption 6 relates to fish being recaptured over several days or weeks. In capture-recapture one should release all fish at a given dam (or occasion) simultaneously. Similarly, all recaptures at a given dam (or occasion) should occur simultaneously. When these conditions are met, all fish have been exposed to mortality risks for the same time interval, and the assumption of homogeneous survival and capture rates, by treatment group, is tenable. This assumption may fail when, for example, some control fish move from dam 1 to dam 2 in 1 week and others take 2 weeks. If survival rates depend only on distance moved, this movement time differential is not a problem; if survival rates depend strongly on elapsed time (as well as, or rather than, distance moved), however, then ϕ (absolute survival rate) is affected. At dam 1, fish should be released as quickly as possible. After the initial release the only control the investigator has is the spatial allocation of recapture effort. In terms of meeting assumption 6, it is best to concentrate recapture effort at dam 2, rather than farther downriver.

Assumption 6 can be weakened for purposes of estimating treatment effects; it suffices to have the same time distribution of recaptures and rereleases for all treatment groups (assumption 10). Thus, inferences about S require assumption 10 but not assumption 6.

Although much of the literature on the type of large-scale experiments we are addressing is unpublished, a number of reports nonetheless provide excellent information on, and examples of, careful field procedures. Useful studies include the following: Cramer and Oligher (1964), Semple (unpublished report, 1979), Olson and Kaczynski (unpublished report, 1980), Turbak et al. (unpublished report, 1981), and McKenzie et al. (unpublished report, 1984).

There is little the investigator can do about assumptions 7 and 8. Assumption 7 implies assumption 8, but not vice versa. Assumption 7 is needed to justify the multinomial probability models used herein. Assumption 8 suffices to justify the empirical estimators of variance in Part 4. Failure of assumption 7 has no serious effects on bias of any estimators but can seriously affect variances. Fish fates are expected to be independent. Independence fails if clusters of fish stay together and react together. This positive dependence effectively reduces actual sample sizes and increases actual variances (as compared to theoretical variances). We doubt that assumptions 7 or, especially, 8 are seriously violated.

Failure of assumptions 9, 10, 11, and 12 can seriously affect estimates of parameters. We focus here on these assumptions because they can be investigated by data analysis. Assumption 9 is general but worth stating because it is the essence of all statistical inference: the assumed statistical model is correct. By statistical model, we mean both the structural and stochastic components. Assumption 7 (independence) and assumption 12 (homogeneity) imply the multinomial model of recaptures given releases. Assumption 12 by itself implies the

structural component of the models we use.

Assumption 10 can be examined on the basis of records of recaptures by day at any dam. If the time distribution of captures is the same for treatment and control fish, the results of the experiment are more likely to be valid. Differential movement of treatment and control fish need not invalidate results, although it may require a more complex model for data analysis. If any differential migration is a result of treatment, it is probably unavoidable. It is important to conduct the study so that the design and field procedures do not lead to such differences in movement between treatment versus control fish.

Assumption 11 relates to, among other things, handling mortality (e.g., Arnason and Mills, in press) and behavioral response to capture (Nichols et al. 1984). If capture and handling cause mortality, they will bias estimates of ϕ (and possibly S). In general, capture and release at dam i could affect the next recapture. If only capture probabilities are affected, generalized models can allow adjustment for this effect (e.g., Nichols et al. 1984). However, if assumption 11 is violated, it is most likely the survival rate after release that is affected. There is no analytic way to compensate for a handling effect on survival rate S and still use all the data for the types of experiments considered here. The solution to the problem, at least with respect to estimating S, is to use the first capture history protocol.

Assumption 12 is violated if fish used in the study vary a priori with respect to survival and capture rates. Such heterogeneity in parameter values is likely to happen to some extent. For example, fish size may influence survival ϕ and treatment effect S. If variation in fish size is modest, such heterogeneity causes no problems. That is, the analysis methods have some robustness to heterogeneity (see Nichols et al. 1982; Pollock and Raveling 1982). Studies should be designed to ensure that assumption 12 is true; the investigator should consider stratifying by fish size or eliminating extremes of fish size and using one strain (source) of fish for the entire study, or at least stratifying treatment and control lots by strain. Heterogeneity can result in actual variances exceeding estimated theoretical variances.

1.4.5. Data Analysis Philosophy

Our specification of assumption 9 as simply "the correct model is used" is motivated by our modeling and data analysis philosophy: start with a model sufficiently general so that it is likely to be true for any given experiment of the type of experiments being considered. That umbrella model is expected to be too general. Consequently, one then selects as part of the data analysis process a suitable special case of the general model that best fits one's specific experiment.

Historically, the approach to capture-recapture data analysis was to assume a specific model and use it, often with no testing of assumptions about model fit. Then one's assumptions about the model could be stated as specific assumptions about survival rates and capture probabilities. For the methods presented here, we assume that the Jolly-Seber model applies to each treatment group. This assumption simply means that for all treatment groups, survival and capture are assumed to be time-specific only, with no age, capture history, or handling

effects. More general models are available; however, we do not believe they are useful for the analysis of this class of fisheries experiments where no new (previously unmarked) fish are released at dams 2, ..., k. Although this absence of any new releases simplifies the problem, it also restricts one's ability to use models more general than Jolly-Seber.

The assumption that parameters within a treatment group are time-specific is probably reasonable, especially if there is some stratification or control of fish size. Capture history could, in principle, influence parameters; however, testing could detect this effect. If such a capture history effect is found, it can be adjusted for analytically (i.e., by modeling and data analysis). The one effect that cannot be detected or adjusted for is a release (i.e., handling) effect at dams 2, ..., k - 1. That problem, however, is not circumvented by starting with an umbrella model more general than Jolly-Seber (see Chapter 3.8).

Two important aspects of our data analysis philosophy are as follows:

(1) Assumptions need to be stated explicitly, and tested, insofar as possible. The goodness of fit of the Jolly-Seber model can be tested thoroughly (Pollock et al. 1985). Overall goodness of fit tests can, and should, be partitioned into informative subcomponents. This partitioning is analogous to single-degree-of-freedom contrasts in the analysis of variance. It is important to do a series of such tests focused on specific alternatives because the omnibus (unpartitioned) goodness of fit test has low power.

(2) The analysis of release-recapture data should be thought of as model fitting, in the sense of seeking a "good" model for the data. With knowledge of the ecology of the species of interest as a starting point, one uses a combination of goodness of fit testing and testing between alternative models to search for the most parsimonious model (fewest parameters) that statistically fits the data and makes good ecological sense.

To facilitate this model selection process, we present (in Chapter 2.1) a menu of models of increasing generality and several data collection protocols. Thus, one can search for a model that fits the data. It is not possible to present analytic results for all possible models or model sequences. Thus, it may be necessary to use a model not included here; generally, that will require numerical analysis in which a combination of programs RELEASE and SURVIV is used.

Our focus here is on determining the extent and nature of the treatment effect on survival rates. The extensive testing and model selection, in terms of other parameters (e.g., p, τ, or λ), are essential to determine the extent and magnitude of the treatment effect and the best way to estimate that effect.

Finally, when a model is selected to fit the data from a given experiment, the investigator must remember that it is just a model. That model is not reality; rather it merely provides the best representation of the data at hand. These data may not be refined enough to demonstrate, at statistically significant levels, minor treatment effects. For example, the major effect may be found in $S_1 = \phi_{t1}/\phi_{c1}$; however, $S_2 = \phi_{t2}/\phi_{c2}$ may also be slightly different from one, but this effect may be too small to detect with the available data. One would then select a model wherein $\phi_{t2} = \phi_{c2}$ (thus, $S_2 = 1$) as the best model to describe the given data. As such, it summarizes the statistically significant information in the data and tells what statistical inferences the data justify, not necessarily what reality is.

1.5. Treatment-Control Mortality Concepts

1.5.1. Introduction

The objective of the experiments considered here is to make inferences about the mortality caused by one or more treatments. In particular, we concentrate on mortality at hydroelectric dams caused, for example, by spillways, bypass systems, turbines, or various deflecting screens. This type of mortality is mostly a sudden, "acute," form of treatment effect. However, there are other studies where treatment mortality may occur over a long period of time, for example, when treatment is a mildly toxic substance. In this latter case the treatment effect is referred to as "chronic." Both types of effects sometimes occur: sudden, direct treatment mortality is followed by chronic effects. Also, in principle, it is possible that a treatment is intended to enhance survival rates. This chapter explores these and other mortality concepts, and considers what mortality or survival effects can be estimated and tested for in release-recapture studies.

The complicating factor is the presence of "natural" mortality, which can occur in the treatment releases prior to recapture. The presence of natural mortality necessitates controls. However, using controls does not entirely solve the problems that arise when the treatment effect is partially or totally chronic and natural mortality forces are also present. Conceptually, one would like to estimate "total treatment mortality" (TTM), the total mortality caused by the treatment over a given period of time. Total treatment mortality is difficult to conceptualize and impossible to estimate without bias (independent of the level of natural mortality) unless the treatment mortality is of the acute type.

Further complications arise if the study occurs over any time period long enough to allow possible compensatory population mortality processes. No attempt is made in the methods here to allow for such a problem. We simply note that it is an important consideration in long-term studies of chronic effects.

In general, a treatment effect may be manifested in ways other than a differential survival rate of treatment animals; however, the only effects detectable from classical release-recapture data are altered survival rates (hence mortality effects) or effects on capture probabilities. We regard the latter as nuisance effects because capture probabilities are not fundamental population dynamics parameters. In the actual data analysis the investigator must be concerned with separating any treatment effects on capture probabilities from effects on survival rates. However, in this chapter our considerations are restricted to mortality effects.

Acute or chronic effects are general terms used to describe a treatment effect. When we deal with mortality as the effect, we use the terms direct mortality (an acute effect) or indirect mortality (a result of a chronic effect). Treatment effects are, of necessity, defined only with respect to a control.

1.5.2. Direct versus Indirect Mortality Effects

Direct mortality is effectively instantaneous. Fish either survive the turbine passage relatively unharmed, are killed outright (from a direct hit by a blade), or are fatally injured (e.g., by pressure-caused internal injuries). If a fish is fatally injured after passage through a turbine, its ultimate death is here considered a direct mortality even if it lives for a short time after exiting the turbine housing. All direct mortality occurs upstream from dam 2 (the first recapture site). The important point here is that the actual cause of direct mortality is the turbine, not a cause that operates on both treatment and control fish (e.g., predation or disease).

In contrast, indirect mortality would be revealed by a higher mortality rate in treatment fish than in controls between dams 2 and 3, 3 and 4, and so forth. Moreover, the proximate cause of that mortality would be natural. Basically, an indirect (possibly chronic) mortality effect means that treatment animals are more susceptible than control animals to natural mortality forces (mortality forces other than the treatment). This enhanced susceptibility may last indefinitely. For such indirect mortality, competing risk theory must be used to separate the treatment effect from the control level of natural mortality risk.

Table 1.17 gives a numerical illustration of one aspect of these mortality concepts. For the purposes of this example, some extended notation is needed. For treatment individuals, let $\phi_{t1} = \phi'_{t0}\,\phi'_{t1}$, where $1 - \phi'_{t0}$ is direct treatment mortality. Conditional on fish not experiencing direct mortality, ϕ'_{t1} is the survival rate to recapture site (or time) 2. By definition, $\phi'_{c0} \equiv 1$, hence, $\phi'_{c1} \equiv \phi_{c1}$. Capture methods allow us to estimate, at best, only ϕ_{t1} and ϕ_{c1}; for some protocols, only their ratio $S = \phi_{t1}/\phi_{c1}$ is estimable. In the case of turbine mortality, ϕ'_{t0} is survival through the turbine, and ϕ'_{t1} is the survival of treatment fish between the point where controls are released and dam 2. If all treatment mortality is direct, $\phi'_{t1} = \phi_{c1}$, and $S = \phi_{t1}/\phi_{c1} = \phi'_{t0}$. However, in general, $S = (\phi'_{t0}\phi'_{t1})/\phi_{c1}$, which has the interpretation

$$S = \frac{\Pr_t(\text{survives turbine})\Pr_t(\text{survives to dam 2} \mid \text{survives turbine})}{\Pr_c(\text{survives to dam 2})}.$$

The treatment mortality is well-defined and estimable when it is entirely a direct effect. In the example of this case in Table 1.17, $\phi_{t1} = \phi'_{t0}\phi'_{t1} = (0.9)(0.9)$, and $\phi_{c1} = \phi'_{t1} = 0.9$, thus,

$$1 - S = 1 - \frac{\phi_{t1}}{\phi_{c1}} = 1 - \frac{0.81}{0.90}$$

$$= 1 - 0.9$$

$$= 0.1 = 1 - \phi'_{t0}.$$

Scenarios that include indirect effects are less well-defined. One possibility is to have

$$S = \frac{\phi_{ti}}{\phi_{ci}}$$

be constant (this situation will not occur in turbine studies). The indirect-only case in Table 1.17 has ϕ_{ti}/ϕ_{ci}, i = 1, ..., 5 as 1, 0.99, 0.97, 0.94, and 0.90, respectively. Table 1.17 also shows some mixed cases. Case 1 has the following estimable ratios ϕ_{ti}/ϕ_{ci}, i = 1, 2, 3, 4, 5: 0.81, 0.94, 0.97, 0.99, 1. For mixed-effects case 2, the estimable survival effects are 0.97, 0.95, 0.93, 0.92, 0.90.

1.5.3. Total Treatment Mortality

Generally, in Part 1, we have avoided examining subjects in detail (leaving that to later parts). However, because we do not return elsewhere to mortality concepts, we present the mathematical formulae here. This material is necessary only if one wants a thorough understanding of the limitations of release-recapture experiments in terms of which components of treatment mortality are estimable.

The methodology used here is basically derived from the competing risk theory of survival processes, which is a coherent mathematical theory of survival (or mortality) processes wherein the individuals are subject to several distinct forces of mortality. Much literature exists on competing risk theory (see David and Moeschberger 1978 and Kalbfleisch

Table 1.17. – Numerical illustration of treatment mortality concepts for cases of direct, indirect, and mixed mortality; note that $\phi_1 = \phi'_0 \phi'_1$. Survival rates ϕ_i are between release and recapture sites (and thus are estimable under certain protocols). Treatment occurs at release site 1; ϕ'_0 is the direct effect.

Survival rate	Control survival rate	Possible treatment survival rate			
		Direct only	Mixed case 1	Mixed case 2	Indirect only
ϕ'_0	1.00	0.90	0.900	0.980	1.000
ϕ'_1	0.90	0.90	0.810	0.880	0.900
$\phi_1 = \phi'_0 \phi'_1$	0.90	0.81	0.729	0.873	0.900
ϕ_2	0.85	0.85	0.800	0.807	0.842
ϕ_3	0.95	0.95	0.922	0.888	0.922
ϕ_4	0.78	0.78	0.772	0.718	0.733
ϕ_5	0.93	0.93	0.930	0.837	0.837

and Prentice 1980; also, Anderson and Burnham 1976 provided an ecological application of competing risk theory, and Fletcher 1985 used a competing-risks approach in an analysis of fisheries diversion experiments).

Definitions of some notation used here follow.

$\phi_t(0, d)$	the survival rate of treatment fish from the release point to downstream distance d (results here are also interpretable with d as time).
$\phi_c(\varepsilon, d)$	the survival rate of control fish from their release point (ε) to downstream distance (location) d; for convenience, define $\phi_c(0,\varepsilon) = 1$.
$h_t(x)$	the instantaneous mortality rate (as a function of location, x) for the treatment fish.
$h_c(x)$	the instantaneous mortality rate for the control fish.
$\Delta(x)=h_t(x) - h_c(x)$	the instantaneous treatment mortality effect; this is a fundamental way of conceptualizing the effect of the treatment.

It is assumed that controls are released just downstream from dam 1 at location ε. By defining $\phi_c(0,\varepsilon) = 1$, the notation $\phi_c(0,d)$ becomes equivalent to $\phi_c(\varepsilon,d)$. Direct and indirect instantaneous treatment mortality effects are represented, respectively, as $\Delta(x) = h_t(x)$ for $0 \leq x \leq \varepsilon$ and $\Delta(x) = h_t(x) - h_c(x)$, $\varepsilon < x$.

Finite survival rates can be expressed in terms of the above instantaneous rates:

$$\phi_t(0, d) = \phi_t(0, \varepsilon)\,(e^{-\int_\varepsilon^d \Delta(y)dy})\phi_c(\varepsilon, d)\,;$$

$$\phi_c(\varepsilon, d) = e^{-\int_\varepsilon^d h_c(y)dy}.$$

The objective here is to express mortality in the treatment cohort as the sum of two parts – mortality that is attributed to the treatment, and mortality that is natural ("natural" mortality being defined here as mortality from any risk factor affecting the controls):

$$1 - \phi_t(0, d) = 1 - e^{-\int_0^d h_t(y)dy} = TM(0, d) + NM(0, d),$$

where

$$TM(0, d) = \int_0^d \Delta(x)e^{-\int_0^x [\Delta(y) + h_c(y)]dy}dx,$$

$$NM(0, d) = \int_0^d h_c(x)e^{-\int_0^x [\Delta(y) + h_c(y)]dy}dx.$$

Here, *TM* denotes treatment mortality and *NM* denotes natural mortality. The quantity denoted $TM(0, d)$ is the total mortality over the distance interval 0 to d, that is validly attributed to ("caused" by) the treatment. This $TM(0, d)$ appears to be complex, and it cannot, in general, be simplified. Worse yet, it cannot be estimated without bias except in special cases.

Let there be a short interval $(0, \varepsilon)$ wherein all direct mortality occurs (or is caused). The turbine experiments fit this model. Then $TM(0, d)$ can be partitioned into direct and indirect treatment mortality, assuming that no indirect treatment mortality occurs in 0 to ε. The result is

$$TM(0, d) = 1 - \phi_t(0, \varepsilon) + \phi_t(0, \varepsilon) \int_{\varepsilon}^{d} \Delta(x)\, e^{-\int_{\varepsilon}^{x} [\Delta(y) + h_c(y)]dy} dx.$$

Thus, $TM(0, d)$ is expressed as equal to direct treatment mortality $1 - \phi_t(0, \varepsilon)$ plus indirect treatment mortality over the distance (or time) ε to d.

If treatment effect eventually wears off entirely (as regards mortality), one can conceptualize a distance d_* beyond which $\Delta(d) = 0$. If the notation is extended for finite survival rates, $\phi_t(d_*, d) = \phi_c(d_*, d)$ for all $d > d_*$. Conversely, $\phi_t(0, d) < \phi_c(0, d)$ for all $d < d_*$ (in turbine mortality experiments we expect $\phi_t < \phi_c$). If the treatment mortality effect never entirely vanishes, we take d_* as infinity. The total treatment mortality (TTM) can now be defined as $TM(0,d_*)$. Eventually all the treatment animals die; TTM is the proportion of that 100% mortality that may validly be attributed to the treatment.

Inasmuch as TTM is an unambiguous measure of treatment effect, we would like to be able to estimate it. This estimation can be done without bias only when $d_* = \varepsilon$, that is, when all the treatment mortality is direct. Then $\phi_{t1} = \phi_t(0, \varepsilon)\phi_{c1}$, so that the parameter we denote as $S = \phi_{t1}/\phi_{c1}$ satisfies TTM = $1 - S$ only in this special case. In general, for results over $(0,d)$

$$S = \phi_t(0, \varepsilon)\, e^{-\int_{\varepsilon}^{d} \Delta(y)dy},$$

whereas

$$1 - TM = \phi_t(0, \varepsilon) \left[1 - \int_{\varepsilon}^{d} \Delta(x)\, e^{-\int_{\varepsilon}^{x} [\Delta(y) + h_c(y)]dy} dx \right].$$

Finally, from

$$1 - e^{-\int_{\varepsilon}^{d} \Delta(x)dy} = \int_{\varepsilon}^{d} \Delta(x) e^{-\int_{\varepsilon}^{x} \Delta(y)dy} dx \geq \int_{\varepsilon}^{d} \Delta(x) e^{-\int_{\varepsilon}^{x} [\Delta(y) + h_c(y)]dy} dx,$$

we derive

$$1 - S \geq TM;$$

hence, for $d \geq d_*$, $1 - S \geq TTM$, with equality if and only if $d_* = \varepsilon$.

We add some interpretation of the above. If d_* lies above the first recapture dam, information on the treatment effect, in regard to survival, is entirely contained in ϕ_{t1} and ϕ_{c1}, and their ratio is one valid measure of the treatment effect. If, however, there are indirect effects, $1 - S$ exceeds the total mortality that should be attributed to the treatment effect. For fish passing through turbines, most of the effect will be direct mortality, and any indirect effects are likely to disappear at or before the next downstream dam. Consequently, for turbine (or bypass or spillway) experiments, the interpretation of $1 - (\phi_{t1}/\phi_{c1})$ as TTM should be a reasonable approximation.

The various equations above provide a basis for investigating the matter further. One can specify forms for $h_c(y)$ and $h_t(y)$, or $\Delta(y)$, and the value of d_* and numerically compute S, TTM, and $NM(0, d_*)$. This computation allows comparison of $1 - S$ with TTM over a wide range of conditions. We consider here a simple case which gives analytic formulae. Assume that $h_c(y) = h_c$ is constant over ε to d_* and then replace $\Delta(y)$ over this interval by its average value (say $\overline{\Delta}$) and hence define $\gamma = \overline{\Delta}/h_c$. For $d \geq d_*$, this case leads to the formula

$$\frac{1 - \text{TTM}}{\phi_t(0, \varepsilon)} = 1 - \frac{\gamma}{1 + \gamma}\left[1 - \left(\frac{S}{\phi_t(0, \varepsilon)}\right)^{\frac{1+\gamma}{\gamma}}\right]$$

and

$$S = \phi_t(0, \varepsilon)e^{-(d-\varepsilon)\overline{\Delta}},$$

or equivalently for S,

$$S = \phi_t(0, \varepsilon)[\phi_c(0, d)]^{\gamma},$$

where $1 - \phi_c(0,d)$ is control mortality.

The reader can compute results with these formula for $d \geq d_*$ (modifications are needed to examine results over $d < d_*$). We have calculated representative results. For example, if direct treatment mortality is 0.1 [$\phi_t(0, \varepsilon) = 0.9$], control mortality is 0.15 [i.e., $\phi_c(0\,d) = 0.85$], and indirect treatment mortality is 0.02, then TTM = 0.12, and $1 - S = 0.12167$. Consideration of the theory here and extensive tabulation of results with the above simple formulae lead to some general conclusions:

(1) As d increases (beyond d_*), the error in using $1 - S$ to estimate TTM increases (when natural mortality is occurring in the river reach d_* to d). Consequently, recapturing should occur as close to d_* as possible.

(2) For fixed $d > d_*$, as the control mortality increases [$\phi_c(0, d)$ decreasing], $1 - S$ becomes a progressively poorer approximation of TTM.

(3) If control mortality (over 0 to d) is <0.15, the ratio (1 - S)/TTM is generally <1.1 irrespective of how much of TTM is direct mortality. In fact, if there is direct mortality, (1 - S)/TTM is likely to be <1.05.

(4) If most treatment mortality is direct ($\geq$ 70% TTM), the ratio (1 - S)/TTM appears to be approximately 1.05 at any level of control mortality or TTM.

For turbine mortality studies at hydroelectric dams, we believe that 1 - S is a good approximation of TTM. Viewed alternatively, the fact that 1 - S may not exactly equal TTM is of negligible concern in such large-scale studies compared with the size of se($\hat{S}$) and the many practical problems that constitute possible sources of serious bias in $\hat{S}$.

1.5.4. Problems with Defining a Treatment Effect on Survival

In this section we pursue further the question of what reasonably constitutes the measurable treatment effect in release-resampling experiments. At best, we can only estimate separate survivals ϕ_{ti} and ϕ_{ci} between recapture dams i = 1, ..., k - 2. Thus, we can estimate quantities like $\phi_{c1} - \phi_{t1}$ or ϕ_{t1}/ϕ_{c1} as our measures (indices) of treatment effect. If there is indirect mortality, we cannot separate it from direct mortality. It is, however, possible to test for the existence and extent of the treatment effect under the complete capture history protocol or under the scheme A partial capture history protocol.

In Section 1.5.3 we showed that 1 - S is often a good approximation of TTM; if there is no indirect mortality, S is the best measure of the treatment effect. The primary alternative to S is the difference $\phi_c - \phi_t$. This difference is affected by the choice of d, whereas the ratio of ϕ_t/ϕ_c stabilizes as d increases. To clarify this point, we expand our notation and show the dependence of S on distance: $S(0, d) = \phi_t(0, d)/\phi_c(0, d)$. Of course, we can estimate this ratio at distances d_2, ..., d_{k-2} corresponding to recapture dams (but only for certain protocols). As d increases toward d_*, $S(0, d)$ changes; however, for all $d > d_*$, $\phi_t(0, d)/\phi_c(0, d) = S(0, d_*)$. This stabilization of $S(0, d)$ contrasts with the difference $\phi_c(0, d) - \phi_t(0, d)$, which goes to zero as d increases; it is a valid measure of effect, and many of the tests presented here are actually based on the difference. However, one should be aware of the contrasting properties of the ratio versus the difference of survival rates.

A further reason for our emphasis on S as the measure of treatment effect is that only the ratio S is estimable as a convenient measure of treatment effect for the first capture history and unknown capture history protocols. Only for complete capture history and partial capture history protocols can one separately estimate, at least, ϕ_{t1} and ϕ_{c1}.

Although we have focused here on studies of fish passing through hydroelectric dams, the nature of the treatment effect could be quite different in other treatment-control release-recapture (or capture-recapture) studies. The anticipated pattern of mortality in any study significantly affects the types of models and the strategy of testing one should use. We have concentrated here on effects that are anticipated to be initially large and then vanish. Hence, we assume most of the effect is an acute, initial mortality. If any indirect effects vanish by or before dam 2 (thus, $d_* \leq d_2$), one can characterize the treatment effect with the single

parameter $S(0, d_*)$ and estimate this parameter with the equation $S = \phi_{t1}/\phi_{c1}$. If an indirect effect persists beyond dam 2, then one needs to extend our characterization of the "treatment effect."

The preferred approach would be to compare the entire set of survival curves for treatment versus control cohorts. Thus, one compares $\phi_t(0, d)$ to $\phi_c(0, d)$ as functions of $d > 0$. The comparison(s) might be reduced to looking at functions of these, such as average life time (after release). Such simple comparisons are not possible with release-recapture data because only the discrete survival rates ϕ_{ti}, ϕ_{ci}, $i = 1, ..., k - 2$ can be estimated. Also, these estimators are subject to substantial sampling variances and covariances.

In this monograph, we emphasize intensive testing for differences between these survivals to identify the extent and size of the treatment effect. These "generic" tests can be improved if one has prior knowledge (or beliefs) about the treatment effects. In particular, if effects are chronic with no direct effect, one should consider imposing a parametric model on ϕ. The survival curve might be exponential, Weibull, or logistic; e.g., $\phi(0,d) = \exp[-(d/\alpha)^\beta]$. One then must assess the fit of this model (separately for treatment and control); if it is satisfactory, the treatment effect is reflected by differences in the parameters α and β between treatments and controls.

In studies with long-term chronic effects, one hopes that the effect is monotonic because it is then not difficult to hypothesize a useful parameter to reflect the treatment effect. Monotonic means that $\Delta(d)$ never changes sign. In this case some weighted combination of the $\hat{\phi}_i$s should be considered as the basis of an efficient test for treatment effect. For example,

$$\overline{\phi} = \sum_{i=1}^{k-2} \left(\frac{d_i + d_{i+1}}{2} \right) \phi(d_i, d_{i+1})$$

or

$$\overline{\phi} = \sum_{i=1}^{k-2} i\, \phi(d_i, d_{i+1}) ,$$

where d_i is the time (or location) of the ith release-recapture with respect to $d_1 = 0$, and $\phi(d_i, d_{i+1}) \equiv \phi_i$ for treatment or control. One then gets the estimate of $\overline{\phi}$ for both treatments and controls and tests for a significant difference.

It is clear to us that chronic treatment effects substantially increase difficulties associated with capture studies. Problems worsen as more of the TTM becomes indirect and as the time duration of the study lengthens (i.e., years rather than weeks or months). In fact, simple capture-recapture studies are not suitable as a basis for a serious study of long-term chronic mortality effects. As just one illustration of difficulties, we note the following, using the equations at the end of Section 1.5.3. Let the control mortality be 0.7 and the TTM be 0.3, with no direct mortality. Then, for $d > d_*$ where $\phi_t = \phi_t(0, d)$ and $\phi_c = \phi_c(0, d)$, $1 - S = 0.48$. In this situation, tests for a treatment effect will be powerful, but there is no way to obtain a reliable estimate of TTM without some type of additional information.