Glossary

Throughout this monograph, we have tried to simplify our notation to the extent feasible by using a few basic symbols in conjunction with several levels of subscripts. An understanding of the general scheme should thus facilitate comprehension of all the specific notation used. The notation is consistent with any statistical modeling needs in capture-recapture and release-resampling of animals. It is often defined in the context of experiments with fish, turbines, dams, etc.; however, the reader will surely recognize more general applications.

Primary Symbols for Statistics

These symbols are subscripted to denote the following factors: treatment-control, release location, recapture location, and subcohort – i.e., capture history.

m, z	Numbers of recaptured fish with reference to a particular sampling site or occasion.
r	Total number of fish ever recaptured from a given release, R.
R	A known number of released fish.
T	A "block total" of the matrix of recaptures, $m + z$.

Primary Symbols for Parameters

The first three symbols are generally subscripted.

ϕ	A survival rate over some reach of river, conditional on fish being alive at the start of the reach.
p	A conditional probability of an animal being captured at a specified site (or on an occasion). This parameter is conditional on the fish arriving alive at that site.
q	$1 - p$, conditional probability of not being captured at a specified site.
S	Treatment effect between the point of release of treatment fish above a turbine at dam 1 and downstream dam 2. A common definition is $S = \phi_{t1}/\phi_{c1}$. If all the turbine mortality is direct and acute, mortality = $1 - S$, and S = the probability of a fish surviving turbine passage (or passage through a bypass system, spillway, etc.). In many applications, S is a probability, but this is not necessary in other applications.

Symbols Used as Subscripts

v — A treatment level. In particular, $v = t$ or c is often used; for example R_{t1} is the number of treatment fish released at dam 1.

t, c — Treatment and control fish, respectively, when only those two cohorts are released at dam 1; e.g., R_{t1} and R_{c1} or r_{t1} and r_{c1}.

i, j — Release and recapture locations, respectively; $i < j$. For example, m_{cij}, where m_{cij} is the number of control fish first recaptured at location j from the numbers R_{ci} released at location i.

. — The dot notation replaces a subscript to denote summation (i.e., pooling) of a statistic over the range of values of that index (e.g., the total number of released fish, pooled over all treatment and control groups, is $R_{.i} = \sum_{v} R_{vi}$). When used with a parameter, the dot notation means that the parameter does not depend on the subscript value. For example, $\phi_{.1}$ means $\phi_{c1} = \phi_{t1}$ ($= \phi_{.1} = \phi_1$); in this example, the v subscript would be dropped and ϕ_1 would be written to show that the parameters are equal.

h — A capture history – for example, R_{cih} and m_{cijh} are components of R_{ci} and m_{cij}. This level of partitioning the data is used for testing of assumptions. Note that when one considers releases at site i having a capture history h, the h depends only on captures that have occurred (or not occurred) at sites 1 to i. We avoid using notation such as $h(i)$ by specifying that a capture history makes sense only with respect to releases at a specific capture site (or time) i.

Miscellaneous Symbols

k — Number of sampling locations. In the current context, release site 1 is the hydroelectric dam at which turbine survival is to be estimated.

H — A hypothesis of interest concerning survival and capture parameters. Specific sampling models for the data arise when one combines a hypothesis about parameters with a particular sampling protocol. Sampling models and hypotheses, such as $H_{1\phi}$, are not equivalent constructs.

Specific Notation

Meanings of the subscripts used here are explained in the previous section. The range of the subscript v includes "." (e.g., $v = t$, c, or "."). In general, $v = 1, ..., V$. Note: recapture always means the first recapture after a (known) release.

R_{vi} — The number of fish of treatment group v released at dam i.

R_{vih} — The number of fish of treatment group v released at dam i that have a particular capture history h at the time of release.

r_{vi} — The number of specific fish ever recaptured from the R_{vi} fish released at dam i, treatment v ($r_{vi} = m_{vi,i+1} + \cdots + m_{vik}$).

r_{vih} — The number of specific fish that had capture history h at the time of release at site i and were ever recaptured; these fish were part of the R_{vih} released at site i.

m_{vj} — The total number of fish of treatment group v recaptured at dam j; $m_{vj} = m_{v1j} + m_{v2j} + \cdots + m_{v,j-1,j}$.

m_{vij} — The number of fish of treatment group v recaptured at dam j from the cohort of R_{vi} fish released at dam i, $i < j$.

m_{vijh} — The number of fish of treatment group v recaptured at dam j that had capture history h at the time they were released at dam i (as part of R_{vih}).

z_{vj} — The total number of fish of treatment group v released before dam j, that were not recaptured at dam j but were recaptured after dam j; hence, fish recaptured at dams $j+1, ..., k$.

T_{vi} — The total number of recaptures for treatment group v at all recapture sites i, $i + 1, ..., k$ from all releases upstream from site i; $T_{vi} = m_{vi} + z_{vi}$.

X_{vh} — The number of fish of treatment group v having capture history h after occasion k. This symbol is used only with complete capture histories; hence, this notation provides a succinct record of the data.

M_{vj} — The number (unknown) of marked fish from treatment group v that reach recapture dam j alive.

ϕ_{vi} — The survival probability between release site i and recapture site $i + 1$ for fish in treatment group v. When the subscript v is present, it is assumed that this survival may depend on treatment group.

ϕ_i — The survival probability between release and recapture site i and site $i + 1$ when this survival is the same for all treatment groups; hence, $\phi_i \equiv \phi_{.i}$.

p_{vi} — The capture probability at site i for fish in treatment group v that reach the site alive.

p_i — The capture probability at site i when it is the same for all treatment groups, hence, $p_i \equiv p_{.i}$.

λ_i — The probability that a fish released at dam i will be recaptured; thus, $E(r_i \mid R_i) = R_i\lambda_i$.

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

$\lambda_k \equiv 0$

$\lambda_i = E(r_i/R_i)$

τ_i — The expected proportion of fish captured at dam i of those released prior and recaptured after dam i.

$\tau_i = E(m_i/T_i)$

Specific Terms

CH matrix — A presentation of the entire study results as a list of capture histories that occurred and the corresponding numbers of captures, by treatment group, having that capture history. The form for one line is $\{h\}\ X_{1h}, ..., X_{vh}$. Losses on capture are denoted by having the X_{vh} signed negative, a convenient way to enter data for computer analysis.

full m-array — The representation of a complete data set for a given treatment group in terms of the releases R_{vih} and recaptures m_{vijh}; thus, releases and recaptures are shown for all subcohorts.

m-array — By convention, m-array means the "reduced" m-array; the data, for a given treatment, are summarized by cohort in terms of releases R_{vi} and recaptures m_{vij}.

cohort — A known number of fish all released at a given site, which may be defined further by treatment group.

subcohort — A subset of a cohort as defined, e.g., by capture history. Usually the survival and capture parameters are expected to be the same in all subcohorts of a cohort; the value of recognizing subcohorts is in testing this assumption.

lot — A batch of treatment and control fish that are all released at (almost) the same time and place, and that were processed together through the pre-release marking and holding activities. Typically, a lot is homogeneous prior to random assignment to treatment group.

sublot — Any identifiable fraction of a lot based on unique tagging information.

Mathematical Symbols

ln( ) — Natural logarithm (base 2.714).

! — Factorial (e.g., $5! = 5 \times 4 \times 3 \times 2 \times 1$).

Σ — Summation operator; e.g., $\sum_{i=1}^{4} i = 1 + 2 + 3 + 4 = 10$.

Π — Product operator; e.g., $\prod_{i=1}^{6} i = 1 \times 2 \times 3 \times 4 \times 5 \times 6 = 720$.

$\int$	Integral.
$\underset{\sim}{x}$	Vector notation; i.e., $\underset{\sim}{x}' = [x_1, x_2, ..., x_n]$.
$\sim$	Approximately equal to.
exp	Exponential, e.g., $\exp(a) = e^a$.
$\frac{\partial L}{\partial p_2}$	Partial derivative of the likelihood function with respect to the parameter p_2.
bin(n, π)	Denotes the binomial probability distribution based on sample size n and probability parameter π.
$\binom{a}{b}$	Binomial coefficient $= \left(\frac{a!}{b!(a-b)!}\right)$; e.g., $\binom{10}{4} = \frac{10!}{4!6!} = 210$.
$\binom{a}{b\ c\ d\ e}$	Multinomial coefficient $= \left(\frac{a!}{b!c!d!e!}\right)$; requires $b+c+d+e = a$.

Statistical Symbols

Pr{ }	Probability.
^	A hat above a symbol denotes an estimate or estimator.
θ	A generic parameter; used to denote some unspecified parameter (θ) or its estimator ($\hat{\theta}$), or a vector of parameters ($\underset{\sim}{\theta}$).
$\bar{x}$	A bar over a symbol represents a sample mean.
$E(\hat{\theta})$	Expected value of the estimator $\hat{\theta}$.
L( )	Likelihood function.
lnL( )	Log-likelihood function.
$I(\underset{\sim}{\theta})$	Information matrix, i.e., matrix of expectations of the mixed partial derivatives of the log-likelihood with respect to the parameters.
π	Cell probability; here a function of the survival and capture probabilities.

Measures of Variability and Covariability

σ^2	Population variance.
σ	Population standard deviation; $\hat{\sigma}$ is an estimator of the population standard deviation.
var($\hat{\theta}$)	Sampling variance of an estimator $\hat{\theta}$, often written as shorthand for var $(\hat{\theta} \mid \theta)$, the sampling variance of $\hat{\theta}$, given the parameter θ.

$\mathrm{se}(\hat{\theta})$	Standard error of an estimator $\hat{\theta}$, $\mathrm{se}(\hat{\theta}) = \sqrt{\mathrm{var}(\hat{\theta})}$. Often only an estimator of this quantity is available, $\widehat{se}(\hat{\theta})$.
$\mathrm{se}(\hat{S})_t$	Theoretical (model-based) standard error of $\hat{S}$ (used only in Part 5).
$\mathrm{se}(\hat{S})_e$	Empirical standard error based on Monte Carlo replicates (used only in Part 5).
$\mathrm{cov}(\hat{\phi}, \hat{p})$	Sampling covariance between the estimators $\hat{\phi}$ and $\hat{p}$.
$\mathrm{corr}(\hat{\phi}, \hat{p})$	Sampling correlation between the estimators $\hat{\phi}$ and $\hat{p}$.
$\boldsymbol{\Sigma}$	Matrix of sampling variances and covariances of estimators.
CI	Confidence interval.
$\mathrm{cv}(\hat{\theta})$	Coefficient of variation, often called "proportional standard error," but expressed as a percentage: $[100\ \mathrm{se}(\hat{\theta})] / \hat{\theta}$.

Hypotheses

H_0	The null hypothesis.
H_A	The alternative hypothesis.

Test Statistics and Related Values

t	A test statistic following Student's t distribution under the null hypothesis.
z	A test statistic distributed normally with $\mu = 0$ and $\sigma = 1$ under the null hypothesis.
χ^2_n	A test statistic distributed as chi-square with n degrees of freedom under the null hypothesis.
df	Degrees of freedom.

Less-Used Notation

d	The number of fish lost on capture; e.g., d_{cj} is the number of control fish lost on capture at dam j.
p	The probability of getting heads when flipping a coin, used in Chapter 1.2. This is not to be confused with capture probability, which appears with subscripts.
w_i	A weight. When normalized, the weights sum to 1.
$h_v(x)$	Instantaneous mortality rate, as a function of location, by treatment group $v = t$ or c (used only in Chapter 1.5).

$\Delta(x)$ — Instantaneous treatment mortality effect, $\Delta(x) = h_t(x) - h_c(x)$ (used only in Chapter 1.5).

$\phi_v(0, d)$ — Survival probability for fish in treatment group v from release point to downstream distance d (used only in Chapter 1.5).

$\phi_v(d_1, d_2)$ — Survival probability for fish in treatment group v between distances d_1 and d_2 ($d_1 < d_2$); distances are measured from the most upstream release point (used only in Chapter 1.5).

d_* — The point downstream from dam 1 after which treatment fish show no more residual treatment effect; hence, after point d_*, treatments and controls have identical responses. Mathematically, $\phi_c(d_*, d) = \phi_t(d_*, d)$ for all $d \geq d_*$, and $\phi_t(d, d_*) < \phi_c(d, d_*)$ for all $d < d_*$ (used only in Chapter 1.5).

Abbreviations

CCH	Complete capture history.
FCH	First capture history.
ML	Maximum likelihood.
MLE	Maximum likelihood estimator.
MSS	Minimal sufficient statistic.
NM	Natural mortality (used only in Chapter 1.5).
PCH	Partial capture history.
PIT	Passive integrated transponder, a new tag technology.
TM	Treatment mortality (used only in Chapter 1.5).
UCH	Unknown capture history.

Technical Terms

Accuracy — Freedom from error or defect: correctness; usually refers to numerical computations.

Bias — (Of an estimator): the difference between the "expected" value of an estimator and the true value of the parameter being estimated. Bias is a measure of how much the average estimate and the true parameter value differ. Bias $= E(\hat{\theta}) - \theta$, where θ is the parameter of interest.

Errors — Type I: rejection of a null hypothesis that is true.
Type II: acceptance of a null hypothesis that is not true.

Estimate	The calculated value of an estimator, given a particular set of sample data, designated by a hat (^) over the symbol for the parameter being estimated.
Estimator	A function of sample data that is used to estimate some parameter. An estimator is a random variable and is also designated by a hat (^) over the symbol for the parameter.
Induction	Generalization from a single experiment to the inference class of all similar experiments. Reasoning from the particular to the general.
Power	The probability of rejecting a null hypothesis.
Precision	A property of an estimator related to the amount of variation among estimates from repeated samples.
Robustness	Insensitivity of an estimation or testing procedure to the breakdown of a specific assumption on which it is based.
Statistic	A function of the sample data.
Test statistic	A value, to be computed from the experimental data, that will determine the decision concerning a null hypothesis. The distribution of the test statistic is known if the null hypothesis is true.
Variance	Theoretical: a variance derived from the particular assumed model by using the ML method. All the models used here are based on the assumption of multinomial sampling variation. Empirical: A variance that is based on some type of replication and is therefore free from specific probability distribution assumptions. If n lots are available, n point estimates of a parameter are available, $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_n$. The simplest empirical variance is then $\hat{\text{var}}\,(\hat{\theta}) = \frac{1}{n-1} \sum_{i=1}^{n} (\hat{\theta}_i - \bar{\hat{\theta}})^2$.